

**BAŐKENT ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ  
BİYOMEDİKAL MÜHENDİSLİĐİ ANABİLİM DALI  
BİYOMEDİKAL MÜHENDİSLİĐİ DOKTORA PROGRAMI**

**DERİN ÖĐRENME YAKLAŐIMLARI İLE MEME KANSERİ  
TESPİTİ VE SINIFLANDIRMASI**

**HAZIRLAYAN**

**BÜŐRA KÜBRA KARACA AYDEMİR**

**DOKTORA TEZİ**

**ANKARA – 2025**



**BAŐKENT ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ  
BİYOMEDİKAL MÜHENDİSLİĐİ ANABİLİMDALI  
BİYOMEDİKAL MÜHENDİSLİĐİ TEZLİ DOKTORA PROGRAMI**

**DERİN ÖĐRENME YAKLAŐIMLARI İLE MEME KANSERİ  
TESPİTİ VE SINIFLANDIRMASI**

**HAZIRLAYAN**

**BÜŐRA KÜBRA KARACA AYDEMİR**

**DOKTORA TEZİ**

**TEZ DANIŐMANI**

**PROF. DR. BERNA DENGİZ**

**TEZ ORTAK DANIŐMANI**

**PROF. DR. ZİYA TELATAR**

**ANKARA – 2025**

**BAŞKENT ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**

Biyomedikal Mühendisliği Anabilim Dalı Biyomedikal Mühendisliği Doktora Programı çerçevesinde Büşra Kübra Karaca Aydemir tarafından hazırlanan bu çalışma, aşağıdaki jüri tarafından Doktora Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 08 / 07 / 2025

**Tez Adı:** Derin Öğrenme Yaklaşımları ile Meme Kanseri Tespiti ve Sınıflandırması

| <b>Tez Jüri Üyeleri</b>      |                      | <b>İmza</b> |
|------------------------------|----------------------|-------------|
| Prof. Dr. Berna Dengiz       | Başkent Üniversitesi | .....       |
| Prof. Dr. Hasan Şakir Bilge  | Gazi Üniversitesi    | .....       |
| Prof. Dr. Hamit Erdem        | Başkent Üniversitesi | .....       |
| Doç. Dr. Selda Güney         | Başkent Üniversitesi | .....       |
| Dr. Öğr. Üyesi Cansel Fıçıcı | Ankara Üniversitesi  | .....       |

**ONAY**

Prof. Dr. Dilek Çökeliler Serdaroğlu

Fen Bilimleri Enstitüsü Müdürü

Tarih : ... / ... / .....

**BAŞKENT ÜNİVERSİTESİ**  
**FEN BİLİMLER ENSTİTÜSÜ**  
**DOKTORA TEZ ÇALIŞMASI ORJİNALLİK RAPORU**

Tarih: 14 / 07 / 2025

Öğrencinin Adı, Soyadı : Büşra Kübra Karaca Aydemir

Öğrencinin Numarası : 21920303

Anabilim Dalı : Biyomedikal Mühendisliği Anabilim Dalı

Programı : Biyomedikal Mühendisliği Doktora Programı

Danışmanın Unvanı/Adı, Soyadı : Prof. Dr. Berna Dengiz

Tez Başlığı : Derin Öğrenme Yaklaşımları ile Meme Kanseri Tespiti ve Sınıflandırması

Yukarıda başlığı belirtilen Doktora tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 177 sayfalık kısmına ilişkin, 14/07/2025 tarihinde tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı %5'dir.

Uygulanan filtrelemeler:

1. Kaynakça hariç
2. Alıntılar hariç
3. Beş (5) kelimedenden daha az örtüşme içeren metin kısımları hariç

“Başkent Üniversitesi Enstitüleri Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Usul ve Esaslarını” inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrenci İmzası:

**ONAY**

Tarih: 14 / 07 / 2025

Öğrenci Danışmanı:

Prof. Dr. Berna Dengiz

## TEŞEKKÜR

Doktora tez çalışmam boyunca bilimsel bakış açımı geliştirmemde, araştırma yetkinliğimi artırmamda ve karşılaştığım her türlü zorluğun üstesinden gelebilmemde büyük rol oynayan; bilgi birikimiyle bana yön veren, sabırlı ve anlayışlı yaklaşımıyla her zaman yanımda olduğunu hissettiren değerli danışmanım Prof. Dr. Berna Dengiz'e en derin saygı ve teşekkürlerimi sunarım.

Doktora sürecime dahil olduğu andan itibaren çalışmama değerli katkılar sunan; teknik bilgi ve deneyimiyle araştırmamın yönünü daha sağlam bir zemine oturtmamı sağlayan, her aşamada sorularımı dikkatle dinleyerek çözüm üretmeme destek olan ortak danışmanım Prof. Dr. Ziya Telatar'a da en derin saygı ve teşekkürlerimi sunarım.

Akademik hayata adım atmam konusunda beni ilk günden itibaren cesaretlendiren, bilimsel gelişimime olan inancını her zaman hissettiren, doktora sürecimde de desteğini ve ilgisini esirgemeyen Prof. Dr. Ömer Faruk Elaldı'ya içten şükranlarımı ve saygılarımı sunarım.

Tez savunma sürecimde yapıcı değerlendirmeleri, kıymetli katkıları ve nazik yaklaşımlarıyla çalışmamın akademik niteliğine değer katan Doç. Dr. Selda Güney, Prof. Dr. Hamit Erdem, Prof. Dr. Hasan Şakir Bilge ve Dr. Öğr. Üyesi Cansel Fıçıncı'ya teşekkürlerimi sunarım.

Akademik yolculuğum boyunca bana bilgi ve tecrübeleriyle rehberlik eden; her zaman moral ve motivasyonlarıyla yanımda olan Dr. Öğr. Üyesi Tansel Uyar ve Öğr. Gör. Dr. Tuğçe Kantar Uğur'a, birlikte çalışmaktan her zaman keyif aldığım tüm çalışma arkadaşlarıma, samimiyeti ve güler yüzüyle her zaman destek veren bölüm sekreterimiz Fatma Nur Turaman'a en içten teşekkürlerimi sunarım.

Akademik yolculuğumun en başından itibaren birlikte yürüdüğüm; lisans bitirme projesinden başlayarak yüksek lisans ve doktora süreçlerinde yan yana ilerlediğimiz, her adımda desteğini hissettirdiği, zorlukları birlikte göğüslediğimiz değerli yol arkadaşım Dr. Burcu Oltu'ya gönülden teşekkür ederim.

Tez çalışmamda kullandığım verilerin temininde gösterdikleri ilgi, destek ve iş birliği için Başkent Üniversitesi Ankara Hastanesi Radyoloji Bölümü'nden Prof. Dr. Ahmet Muhteşem Ağıldere ve Öğr. Gör. Dr. Hazal Selvi Çubuk'a teşekkürlerimi sunarım.

Hayatımın en başındaki öğrenme yolculuğuna ışık tutan; yalnızca okumayı ve yazmayı değil, aynı zamanda merak etmeyi, sorgulamayı ve kendime inanmayı da öğreten sevgili ilkokul öğretmenim Sultan Akkaya'ya, üzerimdeki kıymetli emeği için en derin saygı ve sevgilerimi sunarım.

İlkokul sıralarından bu yana yanımda olan, çocukluk anlarımızdan bugüne uzanan dostluğuyla hayatımda çok özel bir yere sahip olan; bu uzun ve zorlu yolculuğa en yakından şahitlik eden, her zaman desteğini ve varlığını hissettiren can dostum Gülten Büşra Övüt'e teşekkür ederim.

Üniversite yıllarında tanışıp dostlukları bugünlere kadar uzanan; yanımda oluşlarıyla bu süreci benim için daha anlamlı ve dayanılır kılan, her zaman desteğini hissettiren canım arkadaşlarım Aybike Baykan ve İlke Özlem Ünalın Buçan'a teşekkür ederim.

Özellikle doktora sürecimin son dönemlerinde, en az benim kadar bu tezin tamamlanmasını heyecanla bekleyen; desteğini ve inancını her zaman hissettiren değerli dostlarım Burcu Karakaya Gezmez, Kadir Çağrı Gezmez, Merve Durukan Kaya, Mert Kaya, Gizem Türkmenel, Mert Türkmenel ve "8+0.5" grubumuzun neşe kaynağı olan biricik kızları Nil Arya Türkmenel'e, bu süreci birlikte anlamlı ve keyifli hâle getirdikleri için gönülden teşekkür ederim.

Hayatımın her anında yanımda olan; ben mutlu olunca mutlu olan, üzüldüğümde gözleri dolan, strese girdiğimde kalpleri daralan, dualarıyla beni her daim ayakta tutan; varlıklarıyla bana güç veren, sabırlarını, fedakarlıklarını, anlayışlarını, sevgilerini ve desteklerini hiçbir zaman esirgemeyen, başarılarımla gurur duyan annem Sevim Karaca'ya ve babam Kadir Karaca'ya en derin minnet ve sevgilerimi sunarım. Bu zorlu süreçte her zaman yanımda olan; destekleriyle moralimi yüksek tutmamı sağlayan, görüntülü konuşmalarla mesafeleri unutturan, Avusturya'dan gelip savunma sunumumun hemen ardından kapıdan içeri girip beni şaşkınlık ve mutlulukla dolduran sürprizleriyle o anı benim için unutulmaz ve tarifsiz kılan canım abim Tolga Kaan Karaca'ya ve bana her zaman bir abla sıcaklığıyla yaklaşan Funda Karaca'ya en içten teşekkürlerimi sunarım.

Bu zorlu yolculuk boyunca en büyük desteğim, en güçlü motivasyon kaynağım olan; her zaman yanımda durarak bana inanan ve başaracağıma benden önce inanan, yalnızca sevgisiyle değil, anlayışıyla, sabrıyla ve üzerimdeki yükleri hafifletmek için gösterdiği sonsuz çabayla bu süreci benim için çok daha katlanılır kılan ve gözlerindeki gururla bu yolculuğun en kıymetli eşlikçisi olan sevgili eşim Canberk Aydemir'e en derin teşekkürlerimi sunarım.

Bu süreçte bana her zaman sevgiyle yaklaşan, anlayışları ve destekleriyle yanımda olduklarını hissettiren; mutluluğumu en az benim kadar içten yaşayan ve başarılarımla gurur duyan ikinci ailem Alpay Aydemir, Derya Aydemir, Cansu Aydemir ve Alihan Aydemir'e, dualarıyla her zaman yanımda olduklarını hissettiren sevgili dedem Ali Aydemir ve babaannem Sevgi Aydemir'e en içten teşekkürlerimi sunarım.

## ÖZET

**Büşra Kübra KARACA AYDEMİR**

**DERİN ÖĞRENME YAKLAŞIMLARI İLE MEME KANSERİ TESPİTİ VE SINIFLANDIRMASI**

**Başkent Üniversitesi Fen Bilimleri Enstitüsü**

**Biyomedikal Mühendisliği Anabilim Dalı**

**2025**

Meme kanseri, dünya çapında kadınlar arasında kanserle ilişkili ölümlerin başlıca nedenlerinden biridir. Erken teşhis ve doğru tanı, sağkalım oranlarını artırmak ve etkili bir tedavi planlamak açısından çok önemlidir. Mamografi, meme kanseri taramasında anormallikleri erken evrede tespit edebilen temel görüntüleme yöntemidir. Mamografik bulgular arasında kitleler, en yaygın ve tanı açısından en kritik lezyonlardır. Mamografi görüntüleri genellikle radyologlar tarafından manuel olarak değerlendirilmekte, ancak lezyon çeşitliliği, yoğun meme dokusu ve yüksek görüntü sayısı süreci zorlaştırmakta; değerlendirmeyi zaman alıcı ve yorucu hâle getirmektedir. Ayrıca, deneyim farklılıkları ve uzun incelemelere bağlı dikkat dağınıklığı tanı doğruluğunu olumsuz etkileyebilmektedir. Son yıllarda, özellikle derin öğrenmeye dayalı yapay zekâ uygulamaları, tıbbi görüntüleme tanı süreçlerinde önemli ilerlemeler sağlamıştır. Bu alanda öne çıkan You Only Look Once (YOLO) algoritması, gerçek zamanlı nesne tespiti ve sınıflandırma yeteneğiyle öne çıkmaktadır. Bu sayede, mamografi görüntülerinde anormalliklerin hızlı ve doğru şekilde belirlenmesine olanak sağlayarak klinik tanı süreçlerine destek olma potansiyeli taşımaktadır.

Bu tez çalışmasında, mamografi görüntülerinde meme kitlelerini tespit ve sınıflandırmak amacıyla YOLO tabanlı bir derin öğrenme modeli geliştirilmiştir. Bu kapsamda, YOLOv5, YOLOv8 ve YOLOv9 (GELAN) mimarileri temel alınmış; bu modellere çeşitli dikkat modülleri ve evrişimsel bloklar entegre edilerek farklı model varyantları oluşturulmuştur. Modeller, CBIS-DDSM, VinDr-Mammo ve bu iki kümenin birleşimiyle oluşturulan veri setleri kullanılarak sıfırdan eğitilmiş, ardından INBreast veri kümesi üzerinde beş katlı çapraz doğrulama yöntemiyle ince ayar uygulanmıştır. Ayrıca, farklı veri artırma stratejilerinin model performansına etkisi değerlendirilmiştir. INBreast üzerinde en yüksek

başarıyı gösteren model varyantları, klinik uygulanabilirliği test etmek amacıyla Başkent Üniversitesi Ankara Hastanesi'ne ait veri kümesiyle de ince ayara tabi tutulmuştur. CBIS-DDSM + VinDr-Mammo → INBreast senaryosunda GELAN-c mimarisine derinlemesine evrişim (DWConv) ve Evrişimli Blok Dikkat Modülü (CBAM) entegre edilerek geliştirilen model varyantı (GELAN-c + DWConv + CBAM), varsayılan veri artırma konfigürasyonu ile 0.878 mAP@0.5 skoruna ulaşmış ve çalışmadaki en yüksek performansı göstermiştir. CBIS-DDSM + VinDr-Mammo → Başkent senaryosunda ise, YOLOv5s mimarisine Verimli Çok-Ölçekli Dikkat (EMA) modülü entegre edilerek oluşturulan model varyantı (YOLOv5s + EMA), varsayılan + mixup veri artırma kombinasyonu ile 0.848 mAP@0.5 skoruna ulaşmıştır. Bu sonuçlar, geliştirilen YOLO tabanlı model varyantları hem literatür veri kümelerinde hem de gerçek klinik görüntülerde yüksek başarı sağladığını ve klinik karar destek sistemlerine entegre edilebilecek potansiyele sahip olduğunu göstermektedir.

**ANAHTAR KELİMELER:** Meme Kanseri, Mamografi, Tespit ve Sınıflandırma, YOLO, Dikkat Mekanizması.

Başkent Üniversitesi, Proje No: KA 24/376

# ABSTRACT

**Büşra Kübra KARACA AYDEMİR**

## **DETECTION AND CLASSIFICATION OF BREAST CANCER USING DEEP LEARNING APPROACHES**

**Başkent University Institute of Science**

**Department of Biomedical Engineering**

**2025**

Breast cancer is one of the leading causes of cancer-related deaths among women worldwide. Early detection and accurate diagnosis are crucial for improving survival rates and planning effective treatment. Mammography is the primary imaging method for breast cancer screening, detecting abnormalities at an early stage. Among mammographic findings, masses are the most common and diagnostically significant lesions. Mammograms are often evaluated manually by radiologists; however, the diversity of lesions, dense breast tissue, and large image volumes make this process challenging, time-consuming, and labor-intensive. Furthermore, differences in experience levels and attention fatigue caused by prolonged examinations can negatively affect diagnostic accuracy. In recent years, artificial intelligence applications, particularly those based on deep learning, have made significant contributions to medical imaging diagnostics. The You Only Look Once (YOLO) algorithm, a leading approach in this field, stands out with its real-time object detection and classification capabilities. This allows for the rapid and accurate identification of abnormalities in mammograms, offering potential support to clinical diagnostic workflows.

In this thesis, a YOLO-based deep learning model was developed to detect and classify breast masses in mammograms. The YOLOv5, YOLOv8, and YOLOv9 (GELAN) architectures were utilized, and various attention modules and convolutional blocks were integrated into these models to generate different model variants. These models were trained from scratch using CBIS-DDSM, VinDr-Mammo, and their combined dataset. Then, fine-tuning was performed on the INBreast dataset using five-fold cross-validation. Furthermore, the effects of different data augmentation strategies on model performance were analyzed. The best-performing model variants on the INBreast dataset were further fine-tuned using a dataset obtained from Başkent University Ankara Hospital to evaluate clinical applicability. In the CBIS-DDSM + VinDr-Mammo → INBreast scenario, the model variant (GELAN-c

+ DWConv + CBAM), which integrates depth-wise convolution (DWConv) and the Convolutional Block Attention Module (CBAM) into the GELAN-c architecture, achieved a mAP@0.5 score of 0.878 with the default data augmentation configuration, representing the highest performance in the study. In the CBIS-DDSM + VinDr-Mammo → Başkent scenario, the model variant (YOLOv5s + EMA) created by integrating the Efficient Multi-Scale Attention (EMA) module into the YOLOv5s architecture, achieved a mAP@0.5 score of 0.848 using the default + mixup augmentation strategy. These results demonstrate that the proposed YOLO-based model variants achieved high performance on both public datasets and real clinical images and show strong potential for integration into clinical decision support systems.

**KEYWORDS:** Breast Cancer, Mammography, Detection and Classification, YOLO, Attention Mechanism.

Başkent University, Project No: KA 24/376

# İÇİNDEKİLER

|   | Sayfa     |
|---|-----------|
| ÖZET .....  | i         |
| ABSTRACT .....  | iii       |
| İÇİNDEKİLER.....  | v         |
| TABLolar LİSTESİ .....  | ix        |
| ŞEKİLLER LİSTESİ .....  | xiii      |
| SİMGELER VE KISALTMALAR.....  | xvii      |
| <b>1. GİRİŞ .....</b>   | <b>1</b>  |
| <b>1.1. Çalışmanın Konusu, Amacı ve Önemi .....</b>                         | <b>1</b>  |
| <b>1.2. Literatür Özeti .....</b>   | <b>3</b>  |
| <b>1.3. Genel Bakış .....</b>   | <b>12</b> |
| <b>2. KURAMSAL TEMELLER .....</b>   | <b>13</b> |
| <b>2.1. Meme Kanseri .....</b>  | <b>13</b> |
| <b>2.2. Meme Anatomisi ve Fizyolojisi .....</b>                             | <b>13</b> |
| <b>2.3. Meme Kanserinin Gelişimi, Risk Faktörleri ve Belirtileri.....</b>   | <b>14</b> |
| <b>2.4. Meme Kanseri Görüntüleme Yöntemleri .....</b>                       | <b>15</b> |
| <b>2.4.1. Mamografi .....</b>   | <b>16</b> |
| <b>2.4.2. Dijital meme tomosentezi .....</b>                                | <b>18</b> |
| <b>2.4.3. Ultrason .....</b>  | <b>19</b> |
| <b>2.4.4. Manyetik rezonans görüntüleme .....</b>                           | <b>19</b> |
| <b>2.5. Meme Kanserinin Mamografik Bulguları.....</b>                       | <b>20</b> |
| <b>2.5.1. Kitle .....</b>   | <b>20</b> |
| <b>2.5.2. Kalsifikasyon .....</b>   | <b>21</b> |
| <b>2.5.3. Mimari Bozulma.....</b>   | <b>22</b> |
| <b>2.5.4. Asimetri.....</b>   | <b>23</b> |
| <b>3. MATERYAL VE YÖNTEM.....</b>   | <b>24</b> |
| <b>3.1. Çalışmada Kullanılan Mamografi Veri Kümeleri .....</b>              | <b>24</b> |
| <b>3.1.1. INBreast .....</b>  | <b>24</b> |
| <b>3.1.2. Curated Breast Imaging Subset of DDSM (CBIS-DDSM).....</b>        | <b>24</b> |
| <b>3.1.3. VinDr-Mammo .....</b>   | <b>24</b> |
| <b>3.1.4. Başkent Üniversitesi Ankara Hastanesi mamografi verileri.....</b> | <b>25</b> |

|  |           |
|--|-----------|
| <b>3.2. Nesne Tespit Yöntemleri .....</b>  | <b>26</b> |
| 3.2.1. İki aşamalı nesne tespit yöntemleri .....   | 27        |
| 3.2.2. Tek aşamalı nesne tespit yöntemleri.....  | 31        |
| <b>3.3. Nesne Tespit Alanında Kullanılan Veri Kümeleri .....</b>                               | <b>33</b> |
| 3.3.1. PASCAL VOC .....  | 33        |
| 3.3.2. COCO.....   | 34        |
| 3.3.3. ImageNet .....  | 34        |
| 3.3.4. Open Images .....   | 34        |
| <b>3.4. You Only Look Once (YOLO).....</b>   | <b>34</b> |
| 3.4.1. YOLOv1 .....  | 37        |
| 3.4.2. YOLOv2 .....  | 38        |
| 3.4.3. YOLOv3 .....  | 40        |
| 3.4.4. YOLOv4 .....  | 41        |
| 3.4.5. YOLOv5 .....  | 45        |
| 3.4.6. YOLOv6 .....  | 46        |
| 3.4.7. YOLOv7 .....  | 47        |
| 3.4.8. YOLOv8 .....  | 48        |
| 3.4.9. YOLOv8-Ghost .....  | 51        |
| 3.4.10. YOLOv9.....  | 52        |
| <b>3.5. YOLO Algoritmasının Hiperparametreleri ve Veri Artırma Teknikleri....</b>              | <b>55</b> |
| <b>3.6. Dikkat Mekanizmaları (Attention Mechanism).....</b>                                    | <b>57</b> |
| 3.6.1. Sıkıştırma ve uyarma ağları (Squeeze-and-excitation networks,<br>SENet).....            | 58        |
| 3.6.2. Verimli kanal dikkati (Efficient channel attention, ECA) .....                          | 60        |
| 3.6.3. Evrişimli blok dikkat modülü (Convolutional block attention<br>module, CBAM).....       | 61        |
| 3.6.4. Normalizasyon tabanlı dikkat modülü (Normalization-based<br>attention module, NAM)..... | 63        |
| 3.6.5. Koordinat dikkat (Coordinate attention, CA) modülü.....                                 | 66        |
| 3.6.6. Verimli çok-ölçekli dikkat (Efficient multi-scale attention, EMA)<br>modülü.....        | 67        |
| <b>3.7. Farklı Evrişim Modülleri .....</b>   | <b>69</b> |
| 3.7.1. Derinlemesine evrişim (Depth-wise convolution).....                                     | 69        |
| 3.7.2. Ghost evrişim (Ghost convolution).....  | 70        |

|  |     |
|--|-----|
| 3.8. Transfer Öğrenme .....  | 72  |
| 3.9. Performans Değerlendirme Metrikleri .....   | 72  |
| 3.9.1. Nesne tespitinde kullanılan performans metrikleri .....   | 72  |
| 3.9.2. <i>k</i> -katlı çapraz doğrulama .....  | 74  |
| 4. DENEYSSEL ANALİZLER VE SONUÇLAR .....   | 75  |
| 4.1. Tez Kapsamında Önerilen Yaklaşım .....  | 75  |
| 4.2. Veri Kümelerinin Analize Hazırlanması .....   | 76  |
| 4.2.1. Ön işleme adımları .....  | 76  |
| 4.1.1.a. Arka plan artefaktlarının kaldırılması .....  | 76  |
| 4.1.1.b. Meme bölgesinin kırılması .....   | 77  |
| 4.2.2. Veri kümelerinin YOLO formatında etiketlenmesi.....   | 78  |
| 4.1.2.a. INBreast veri kümesinin etiketlenmesi .....   | 79  |
| 4.1.2.b. CBIS-DDSM veri kümesinin etiketlenmesi .....  | 80  |
| 4.1.2.c. VinDr-Mammo veri kümesinin etiketlenmesi .....  | 82  |
| 4.1.2.d. Başkent Üniversitesi Ankara Hastanesi veri kümesinin<br>etiketlenmesi .....   | 83  |
| 4.3. YOLOv5 Analiz Sonuçları .....   | 84  |
| 4.4. YOLOv5 ve Dikkat Modülleri Analiz Sonuçları.....  | 96  |
| 4.5. YOLOv8 ve Dikkat Modülleri Analiz Sonuçları.....  | 108 |
| 4.6. YOLOv5 ve YOLOv8 Modellerine Entegre Edilen Dikkat Modüllerinin<br>Performans Karşılaştırması.....                      | 121 |
| 4.7. Dikkat Modülü Entegre Edilmiş YOLOv5 ve YOLOv8 Modellerinde<br>DWConv Uygulaması .....                                  | 123 |
| 4.7.1. YOLOv5 + dikkat modülü modellerinde DWConv uygulaması ....  | 124 |
| 4.7.2. YOLOv8 + dikkat modülü modellerinde DWConv uygulaması ....  | 130 |
| 4.8. Dikkat Modülleri Entegre Edilmiş YOLOv5 ve YOLOv8 Modellerinde<br>DWConv Uygulamasının Performans Karşılaştırması ..... | 136 |
| 4.9. YOLOv8-Ghost ve Dikkat Modülleri Analiz Sonuçları .....   | 138 |
| 4.10. GELAN ve Dikkat Modülleri Analiz Sonuçları.....  | 145 |
| 4.11. Dikkat Modülü Entegre Edilmiş GELAN Modellerinde DWConv<br>Uygulaması .....  | 155 |
| 4.12. Dikkat Modülleri Entegre Edilmiş GELAN Modellerinde DWConv<br>Uygulamasının Performans Karşılaştırması .....           | 163 |

|   |            |
|---|------------|
| <b>4.13. YOLO Tabanlı En Başarılı Model Varyantlarının Performans</b> |            |
| <b>Karşılaştırması .....</b>  | <b>165</b> |
| <b>5. SONUÇ, TARTIŞMA VE ÖNERİLER .....</b>                           | <b>170</b> |

## TABLolar LİSTESİ

|  | <b>Sayfa</b> |
|--|--------------|
| Tablo 2.1. BI-RADS kategorileri ve tanımları.....  | 18           |
| Tablo 3.1. YOLO eğitimi sırasında kullanılan veri artırma parametreleri ve değer aralıkları.....   | 57           |
| Tablo 4.1. Veri kümelerindeki mamografi görüntülerinin ve kitle lezyonlarının sayısı..   | 84           |
| Tablo 4.2. YOLOv5 eğitiminde kullanılan veri artırma konfigürasyonları.....  | 84           |
| Tablo 4.3. YOLOv5 için sıfırdan eğitim ve ince ayar süreçlerinde kullanılan hiperparametreler.....   | 85           |
| Tablo 4.4. CBIS-DDSM veri kümesinde farklı YOLOv5 modellerinin sıfırdan eğitim performans sonuçları .....                                      | 85           |
| Tablo 4.5. VinDr-Mammo veri kümesinde farklı YOLOv5 modellerinin sıfırdan eğitim performans sonuçları .....                                    | 86           |
| Tablo 4.6. CBIS-DDSM + VinDr-Mammo veri kümesinde farklı YOLOv5 modellerinin sıfırdan eğitim performans sonuçları.....                         | 86           |
| Tablo 4.7. CBIS-DDSM → INBreast: Düşük seviye veri artırma ile eğitilen YOLOv5 modellerinin ince ayar performans sonuçları .....               | 87           |
| Tablo 4.8. VinDr-Mammo → INBreast: Düşük seviye veri artırma ile eğitilen YOLOv5 modellerinin ince ayar performans sonuçları .....             | 87           |
| Tablo 4.9. CBIS-DDSM + VinDr-Mammo → INBreast: Düşük seviye veri artırma ile eğitilen YOLOv5 modellerinin ince ayar performans sonuçları ..... | 88           |
| Tablo 4.10. CBIS-DDSM → INBreast: Orta seviye veri artırma ile eğitilen YOLOv5 modellerinin ince ayar performans sonuçları .....               | 88           |
| Tablo 4.11. VinDr-Mammo → INBreast: Orta seviye veri artırma ile eğitilen YOLOv5 modellerinin ince ayar performans sonuçları .....             | 89           |
| Tablo 4.12. CBIS-DDSM + VinDr-Mammo → INBreast: Orta seviye veri artırma ile eğitilen YOLOv5 modellerinin ince ayar performans sonuçları ..... | 89           |
| Tablo 4.13. CBIS-DDSM → INBreast: YOLOv5s modeli için farklı veri artırma parametrelerinin ince ayar performansına etkisi.....                 | 91           |
| Tablo 4.14. VinDr-Mammo → INBreast: YOLOv5s modeli için farklı veri artırma parametrelerinin ince ayar performansına etkisi.....               | 92           |
| Tablo 4.15. CBIS-DDSM + VinDr-Mammo → INBreast: YOLOv5s modeli için farklı veri artırma parametrelerinin ince ayar performansına etkisi.....   | 93           |

|             |   |     |
|-------------|---|-----|
| Tablo 4.16. | CBIS-DDSM → INBreast: Farklı ölçek değerlerinin Deneme 1 üzerindeki ince ayar performansına etkisi .....  | 94  |
| Tablo 4.17. | VinDr-Mammo → INBreast: Farklı ölçek değerlerinin Deneme 7 üzerindeki ince ayar performansına etkisi .....  | 95  |
| Tablo 4.18. | CBIS-DDSM + VinDr-Mammo → INBreast: Farklı ölçek değerlerinin Deneme 7 üzerindeki ince ayar performansına etkisi .....                                  | 95  |
| Tablo 4.19. | CBIS-DDSM veri kümesinde YOLOv5s ve YOLOv5s + dikkat modüllerinin sıfırdan eğitim performans sonuçları .....  | 97  |
| Tablo 4.20. | VinDr-Mammo veri kümesinde YOLOv5s ve YOLOv5s + dikkat modüllerinin sıfırdan eğitim performans sonuçları .....  | 97  |
| Tablo 4.21. | CBIS-DDSM + VinDr-Mammo veri kümesinde YOLOv5s ve YOLOv5s + dikkat modüllerinin sıfırdan eğitim performans sonuçları .....                              | 97  |
| Tablo 4.22. | CBIS-DDSM → INBreast: YOLOv5s ve YOLOv5s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi .....              | 99  |
| Tablo 4.23. | VinDr-Mammo → INBreast: YOLOv5s ve YOLOv5s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi .....            | 101 |
| Tablo 4.24. | CBIS-DDSM + VinDr-Mammo → INBreast: YOLOv5s ve YOLOv5s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi..... | 103 |
| Tablo 4.25. | CBIS-DDSM + VinDr-Mammo → Başkent: YOLOv5s ve YOLOv5s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi.....  | 106 |
| Tablo 4.26. | YOLOv8 eğitiminde kullanılan varsayılan veri artırma konfigürasyonu.....  | 109 |
| Tablo 4.27. | YOLOv8 için sıfırdan eğitim ve ince ayar süreçlerinde kullanılan hiperparametreler .....  | 109 |
| Tablo 4.28. | CBIS-DDSM veri kümesinde YOLOv8s ve YOLOv8s + dikkat modüllerinin sıfırdan eğitim performans sonuçları .....  | 110 |
| Tablo 4.29. | VinDr-Mammo veri kümesinde YOLOv8s ve YOLOv8s + dikkat modüllerinin sıfırdan eğitim performans sonuçları .....  | 110 |
| Tablo 4.30. | CBIS-DDSM + VinDr-Mammo veri kümesinde YOLOv8s ve YOLOv8s + dikkat modüllerinin sıfırdan eğitim performans sonuçları .....                              | 110 |
| Tablo 4.31. | CBIS-DDSM → INBreast: YOLOv8s ve YOLOv8s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi .....              | 112 |
| Tablo 4.32. | VinDr-Mammo → INBreast: YOLOv8s ve YOLOv8s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi .....            | 114 |

|   |     |
|---|-----|
| Tablo 4.33. CBIS-DDSM + VinDr-Mammo → INBreast: YOLOv8s ve YOLOv8s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi.....             | 116 |
| Tablo 4.34. CBIS-DDSM + VinDr-Mammo → Başkent: YOLOv8s ve YOLOv8s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi.....              | 119 |
| Tablo 4.35. CBIS-DDSM + VinDr-Mammo veri kümesinde YOLOv5s + DWConv + dikkat modülleri sıfırdan eğitim performans sonuçları.....  | 125 |
| Tablo 4.36. CBIS-DDSM + VinDr-Mammo → INBreast: YOLOv5s + DWConv + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi.....               | 126 |
| Tablo 4.37. CBIS-DDSM + VinDr-Mammo → Başkent: YOLOv5s + DWConv + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi.....                | 128 |
| Tablo 4.38. CBIS-DDSM + VinDr-Mammo veri kümesinde YOLOv8s + DWConv + dikkat modülleri sıfırdan eğitim performans sonuçları.....  | 131 |
| Tablo 4.39. CBIS-DDSM + VinDr-Mammo → INBreast: YOLOv8s + DWConv + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi.....               | 132 |
| Tablo 4.40. CBIS-DDSM + VinDr-Mammo → Başkent: YOLOv8s + DWConv + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi.....                | 134 |
| Tablo 4.41. CBIS-DDSM + VinDr-Mammo veri kümesinde YOLOv8s-ghost ve YOLOv8s-ghost + dikkat modüllerinin sıfırdan eğitim performans sonuçları.....                               | 139 |
| Tablo 4.42. CBIS-DDSM + VinDr-Mammo → INBreast: YOLOv8s-ghost ve YOLOv8s-ghost + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi..... | 141 |
| Tablo 4.43. CBIS-DDSM + VinDr-Mammo → Başkent: YOLOv8s-ghost ve YOLOv8s-ghost + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi.....  | 143 |
| Tablo 4.44. GELAN eğitiminde kullanılan varsayılan veri artırma konfigürasyonu .....  | 146 |
| Tablo 4.45. GELAN için model eğitim hiperparametreleri .....  | 147 |
| Tablo 4.46. CBIS-DDSM + VinDr-Mammo veri kümesinde GELAN-s ve GELAN-s + dikkat modüllerinin sıfırdan eğitim performans sonuçları .....  | 148 |
| Tablo 4.47. CBIS-DDSM + VinDr-Mammo veri kümesinde GELAN-m ve GELAN-m + dikkat modüllerinin sıfırdan eğitim performans sonuçları .....  | 148 |

|   |     |
|---|-----|
| Tablo 4.48. CBIS-DDSM + VinDr-Mammo veri kümesinde GELAN-c ve GELAN-c + dikkat modüllerinin sıfırdan eğitim performans sonuçları .....                              | 149 |
| Tablo 4.49. CBIS-DDSM + VinDr-Mammo → INBreast: GELAN-s ve GELAN-s + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi..... | 150 |
| Tablo 4.50. CBIS-DDSM + VinDr-Mammo → INBreast: GELAN-m ve GELAN-m + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi..... | 151 |
| Tablo 4.51. CBIS-DDSM + VinDr-Mammo → INBreast: GELAN-c ve GELAN-c + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi..... | 152 |
| Tablo 4.52. CBIS-DDSM + VinDr-Mammo → Başkent: GELAN-c ve GELAN-c + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi.....  | 154 |
| Tablo 4.53. CBIS-DDSM + VinDr-Mammo veri kümesinde GELAN-s + DWConv + dikkat modüllerinin sıfırdan eğitim performans sonuçları .....                                | 156 |
| Tablo 4.54. CBIS-DDSM + VinDr-Mammo veri kümesinde GELAN-m + DWConv + dikkat modüllerinin sıfırdan eğitim performans sonuçları .....                                | 156 |
| Tablo 4.55. CBIS-DDSM + VinDr-Mammo veri kümesinde GELAN-c + DWConv + dikkat modüllerinin sıfırdan eğitim performans sonuçları .....                                | 156 |
| Tablo 4.56. CBIS-DDSM + VinDr-Mammo → INBreast: GELAN-s + DWConv + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi.....   | 158 |
| Tablo 4.57. CBIS-DDSM + VinDr-Mammo → INBreast: GELAN-m + DWConv + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi.....   | 159 |
| Tablo 4.58. CBIS-DDSM + VinDr-Mammo → INBreast: GELAN-c + DWConv + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi.....   | 160 |
| Tablo 4.59. CBIS-DDSM + VinDr-Mammo → Başkent: GELAN-c + DWConv + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi.....    | 162 |
| Tablo 4.60. CBIS-DDSM + VinDr-Mammo → INBreast için en başarılı model varyantlarının performans karşılaştırması .....   | 166 |
| Tablo 4.61. CBIS-DDSM + VinDr-Mammo → Başkent için en başarılı model varyantlarının performans karşılaştırması .....  | 166 |
| Tablo 5.1. Literatürdeki YOLO tabanlı çalışmalar ile tez kapsamında elde edilen sonuçların karşılaştırılması .....  | 175 |

## ŞEKİLLER LİSTESİ

|  | <b>Sayfa</b> |
|--|--------------|
| Şekil 2.1. Memenin anatomik yapısı [62].   | 14           |
| Şekil 2.2. Meme kanserinin gelişim süreci [66].  | 15           |
| Şekil 2.3. Mamografinin dört temel projeksiyonuna ait görüntüler: (a) R-CC görünüm, (b) L-CC görünüm, (c) R-MLO görünüm ve (d) L-MLO görünüm [70]. | 16           |
| Şekil 2.4. Mamografide kitle görünümü [82].  | 21           |
| Şekil 2.5. Mamografide kalsifikasyon görünümü [83].  | 22           |
| Şekil 2.6. Mamografide mimari bozulma görünümü [85].   | 22           |
| Şekil 2.7. Mamografide asimetri görünümü [88].   | 23           |
| Şekil 3.1. Veri kümelerine ait ortalama normalize piksel yoğunluğu dağılımları.  | 25           |
| Şekil 3.2. R-CNN mimarisinin genel yapısı [100].   | 28           |
| Şekil 3.3. SPP katmanı içeren CNN yapısı [101].  | 29           |
| Şekil 3.4. Fast R-CNN mimarisi [102].  | 29           |
| Şekil 3.5. Faster R-CNN mimarisi [103].  | 30           |
| Şekil 3.6. FPN yapısı [104].   | 31           |
| Şekil 3.7. Mask R-CNN [105].   | 31           |
| Şekil 3.8. SSD ve YOLO mimarilerinin karşılaştırılması [104].  | 32           |
| Şekil 3.9. RetinaNet mimarisi [108].   | 33           |
| Şekil 3.10. YOLO algoritmasının nesne tespit sürecinin genel akışı [111].  | 35           |
| Şekil 3.11. IoU hesaplanması.  | 36           |
| Şekil 3.12. YOLOv1 mimarisi [111].   | 37           |
| Şekil 3.13. YOLOv2 mimarisi [115].   | 38           |
| Şekil 3.14. Anchor kutuları [106].   | 39           |
| Şekil 3.15. YOLOv3 mimarisi [119].   | 40           |
| Şekil 3.16. YOLOv4 mimarisinin genel yapısı [120].   | 41           |
| Şekil 3.17. (a) DenseNet ve (b) CSPDenseNet yapılarının karşılaştırılması [121].   | 42           |

|             |  |    |
|-------------|--|----|
| Şekil 3.18. | (a) Klasik SPP bloğu ve (b) YOLO'ya uyarlanmış SPP bloğu [123].                    | 43 |
| Şekil 3.19. | PANet yapısı, (a) FPN, (b) aşağıdan yukarı yol artırımı [124].                     | 44 |
| Şekil 3.20. | (a) PAN, (b) YOLOv4 değiştirilmiş PAN [124].                                       | 44 |
| Şekil 3.21. | YOLOv5 mimarisi [112].   | 45 |
| Şekil 3.22. | YOLOv8 mimarisi [136].   | 49 |
| Şekil 3.23. | YOLOv8-Ghost mimarisi [43].  | 51 |
| Şekil 3.24. | YOLOv9 PGI mimarisi [146].   | 53 |
| Şekil 3.25. | GELAN mimarisi: (a) CSPNet, (b) ELAN, (c) GELAN [146].                             | 53 |
| Şekil 3.26. | GELAN mimarisi [150].  | 54 |
| Şekil 3.27. | SE blok yapısı [162].  | 59 |
| Şekil 3.28. | ECA mimarisi [162].  | 60 |
| Şekil 3.29. | Kanal dikkat modülü yapısı [170].  | 62 |
| Şekil 3.30. | Uzamsal dikkat modülü yapısı [170].  | 63 |
| Şekil 3.31. | CBAM genel yapısı [170].   | 63 |
| Şekil 3.32. | NAM kanal dikkat modülü yapısı [172].  | 64 |
| Şekil 3.33. | NAM uzamsal dikkat modülü yapısı [172].  | 65 |
| Şekil 3.34. | CA mimarisi [175].   | 66 |
| Şekil 3.35. | EMA mimarisi [178].  | 69 |
| Şekil 3.36. | (a) Standart evrişim, (b) Derinlemesine evrişim [181].                             | 70 |
| Şekil 3.37. | Ghost evrişim modülü [144].  | 71 |
| Şekil 4.1.  | Tez çalışmasına ait önerilen yaklaşım.   | 75 |
| Şekil 4.2.  | CBIS-DDSM veri kümesine ait bir görüntüde arka plan temizleme adımları.            | 77 |
| Şekil 4.3.  | VinDr-Mammo veri kümesine ait bir görüntüde arka plan temizleme adımları.          | 77 |
| Şekil 4.4.  | Kırpma işleminin görüntü üzerindeki temsili ve kırpma sonrası elde edilen görüntü. | 78 |
| Şekil 4.5.  | Bir kitleye ait YOLO sınırlayıcı kutusunun gösterimi.                              | 78 |

|             |   |     |
|-------------|---|-----|
| Şekil 4.6.  | Görüntüdeki iki farklı kitleye ait sınırlayıcı kutu maskeleri.....  | 80  |
| Şekil 4.7.  | INBreast veri kümesinden kitle etiketlenmesine ve etiket dosyasına örnek...   | 80  |
| Şekil 4.8.  | Bir görüntüdeki iki farklı kitleye ait iki adet maske görüntüsü.....  | 81  |
| Şekil 4.9.  | CBIS-DDSM veri kümesinden kitle etiketlenmesine ve etiket dosyasına örnek.....  | 81  |
| Şekil 4.10. | VinDr-Mammo veri kümesinden kitle etiketlenmesine ve etiket dosyasına örnek.....  | 82  |
| Şekil 4.11. | Başkent Üniversitesi Ankara Hastanesi veri kümesinden kitle etiketlenmesine ve etiket dosyasına örnek.....  | 83  |
| Şekil 4.12. | YOLOv5 + dikkat modülü model mimarisinin gösterimi.....   | 96  |
| Şekil 4.13. | YOLOv8 + dikkat modülü model mimarisinin gösterimi.....   | 108 |
| Şekil 4.14. | Farklı kaynak veri kümeleriyle eğitilen YOLO + dikkat modülü model varyantlarının INBreast üzerindeki performans karşılaştırılması.....                 | 121 |
| Şekil 4.15. | En başarılı YOLO + dikkat modülü model varyantlarında kullanılan veri artırma parametre kombinasyonlarının performans üzerindeki etkisi.....            | 122 |
| Şekil 4.16. | Başkent veri kümesi üzerinde en başarılı YOLO + dikkat modülü model varyantlarının ve veri artırma kombinasyonlarının performans karşılaştırılması..... | 123 |
| Şekil 4.17. | YOLOv5 + DWConv + dikkat modülü model mimarisinin gösterimi.....  | 124 |
| Şekil 4.18. | YOLOv8 + DWConv + dikkat modülü model mimarisinin gösterimi.....  | 130 |
| Şekil 4.19. | DWConv entegrasyonu öncesi ve sonrası en başarılı YOLO + dikkat modülü model varyantlarının INBreast üzerindeki performans karşılaştırılması.....       | 137 |
| Şekil 4.20. | DWConv entegrasyonu öncesi ve sonrası en başarılı YOLO + dikkat modülü model varyantlarının Başkent üzerindeki performans karşılaştırılması.....        | 137 |
| Şekil 4.21. | YOLOv8-Ghost + dikkat modülü model mimarisinin gösterimi.....   | 139 |
| Şekil 4.22. | GELAN + dikkat modülü model mimarisinin gösterimi.....  | 146 |
| Şekil 4.23. | GELAN + DWConv + dikkat modülü model mimarisinin gösterimi.....   | 155 |
| Şekil 4.24. | DWConv entegrasyonu öncesi ve sonrası en başarılı GELAN + dikkat modülü model varyantlarının INBreast üzerindeki performans karşılaştırılması.....      | 163 |

- Şekil 4.25. DWConv entegrasyonu öncesi ve sonrası en başarılı GELAN-c + dikkat modülü model varyantlarının Başkent üzerindeki performans karşılaştırması. .... 164
- Şekil 4.26. INBreast veri kümesi için farklı model varyantlarının tespit sonuçlarının karşılaştırılması: (a) gerçek etiket, (b) YOLOv5s + SE, (c) YOLOv8s + DWConv + CBAM, (d) GELAN-c + DWConv + CBAM. .... 168
- Şekil 4.27. Başkent veri kümesi için farklı model varyantlarının tespit sonuçlarının karşılaştırılması: (a) gerçek etiket, (b) YOLOv5s + EMA, (c) YOLOv8s + DWConv + EMA, (d) GELAN-c + DWConv + EMA. .... 169

## SİMGELER VE KISALTMALAR

|           |   |
|-----------|---|
| ABUS      | Otomatik Meme Ultrasonu (Automated Breast Ultrasound)   |
| ACR       | Amerikan Radyoloji Koleji (The American College of Radiology)                                 |
| Adam      | Adaptive Moment Estimation  |
| AP        | Ortalama Kesinlik (Average Precision)   |
| AUC       | Eğri Altında Kalan Alan (Area Under Curve)  |
| B         | Sınırlayıcı Kutu (Bounding Box)   |
| BI-RADS   | Meme Görüntüleme Raporlama ve Veri Sistemi (Breast Imaging Reporting and Data System)         |
| BN        | Toplu Normalizasyon (Batch Normalization)   |
| BoF       | Ücretsiz Çanta (Bag-of-Freebies)  |
| BoS       | Özel Çanta (Bag-of-Specials)  |
| CA        | Koordinat Dikkat (Coordinate Attention)   |
| CBAM      | Evrişimli Blok Dikkat Modülü (Convolutional Block Attention Module)                           |
| CBIS-DDSM | Curated Breast Imaging Subset of DDSM   |
| CC        | Kranio-Kaudal (Cranio-Caudal)   |
| CESM      | Kontrastlı Spektral Mammogramları (Contrast-Enhanced Spectral Mammogram)                      |
| CIOU      | Birleşim Üzerinde Tam Kesişme (Complete Intersection over Union)                              |
| CLA       | Katmanlar Arası Dikkat (Cross-Layer Attention)  |
| CNN       | Evrişimli Sinir Ağı (Convolutional Neural Network)  |
| COCO      | Microsoft Common Objects in COntext   |
| CSP       | Çapraz Aşamalı Kısmi (Cross-Stage Partial)  |
| DBT       | Dijital Meme Tomosentezi (Digital Breast Tomosynthesis)                                       |
| DDSM      | Digital Database for Screening Mammography  |
| DFL       | Dağıtım Odak Kaybı (Distribute Focal Loss)  |
| DM        | Dijital Mamografi (Digital Mammography)   |
| DPM       | Deforme Edilebilir Parça Tabanlı Model (Deformable Part-Based Model)                          |
| DVM       | Destek Vektör Makinesi  |
| DWConv    | Derinlemesine Evrişim (Depth-Wise Convolution)  |
| ECA       | Verimli Kanal Dikkati (Efficient Channel Attention)   |
| E-ELAN    | Genişletilmiş ELAN (Extended ELAN)  |
| ELAN      | Verimli Katman Toplama Ağı (Efficient Layer Aggregation Network)                              |
| EMA       | Verimli Çok-Ölçekli Dikkat (Efficient Multi-scale Attention Module)                           |
| FC        | Tam Bağlı (Fully Connected)   |
| FFDM      | Tam Alan Dijital Mammogramları (Full-Field Digital Mammogram)                                 |
| FPN       | Öznelik Piramit Ağı (Feature Pyramid Network)   |
| GAP       | Global Ortalama Havuzlama (Global Average Pooling)  |
| GELAN     | Genelleştirilmiş Verimli Katman Toplama Ağı (Generalized Efficient Layer Aggregation Network) |
| GhostConv | Ghost Evrişim (Ghost Convolution)   |
| HOG       | Yönlendirilmiş Gradyanların Histogramı (Histogram of Oriented Gradients)                      |
| IARC      | Uluslararası Kanser Araştırmaları Ajansı (International Agency for Research on Cancer)        |

|            |   |
|------------|---|
| ILSVRC     | ImageNet Büyük Ölçekli Görsel Tanıma Mücadelesi (ImageNet Large Scale Visual Recognition Challenge) |
| IoU        | Birleşim Üzerinden Kesişme (Intersection-over-Union)  |
| KNN        | K-En Yakın Komşu  |
| Leaky ReLU | Sızdıran Düzeltilmiş Doğrusal Birim (Leaky Rectified Linear Unit)                                   |
| mAP        | Tüm Sınıflar için Ortalama Kesinlik (mean Average Precision)  |
| MLO        | Mediolateral Oblik (Mediolateral Oblique)   |
| MRG        | Manyetik Rezonans Görüntüleme   |
| NAM        | Normalizasyon Tabanlı Dikkat Modülünü (Normalization-based Attention Module)                        |
| NMS        | Maksimum Olmayan Bastırma (Non-Maximum Supression)  |
| P          | Kesinlik (Precision)  |
| PAN        | Yol Toplama Ağı (Path Aggregation Network)  |
| PASCAL VOC | PASCAL Görsel Nesne Sınıflandırması (PASCAL Visual Object Classification)                           |
| PGI        | Programlanabilir Gradyan Bilgisi (Programmable Gradient Information)                                |
| PR         | Kesinlik-Duyarlılık (Precision-Recall)  |
| R          | Duyarlılık (Recall)   |
| RAD        | Kalan Asimetrik Genişletilmiş (Residual Asymmetric Dilated)   |
| R-CNN      | Bölge-Tabanlı CNN (Region-Based CNN)  |
| RoI        | İlgi Alanı (Region of Interest)   |
| RPN        | Bölge Öneri Ağı (Region Proposal Network)   |
| SENet      | Squeeze-and-Excitation Networks   |
| SFM        | Ekran-Film Mamografisi (Screen-Film Mammography)  |
| SGD        | Stochastic Gradient Descent   |
| SIFT       | Ölçek-Değişmez Öznitelik Dönüşümü (Scale-Invariant Feature Transform)                               |
| SiLu       | Sigmoid Linear Units  |
| SPP        | Uzamsal Piramit Havuzlama (Spatial Pyramid Pooling)   |
| SPP-Net    | Uzamsal Piramit Havuzlama Ağı (Spatial Pyramid Pooling Network)                                     |
| SSD        | Tek Atış Çoklu Kutu Dedektörü (Single Shot Multibox Detector)                                       |
| ViT        | Görü Dönüştürücü (Vision Transformer)   |
| WHO        | Dünya Sağlık Örgütü (World Health Organization)   |
| YOLO       | You Only Look Once  |

# 1. GİRİŞ

## 1.1. Çalışmanın Konusu, Amacı ve Önemi

Kanser, her yıl milyonlarca insanın ölümüne yol açan en ölümcül hastalıklardan biridir [1]. Bu hastalıklar arasında meme kanseri, dünya çapında kadınlar arasında en yaygın görülen türlerden biri olarak öne çıkmaktadır [2]. Ergenlik sonrasında her yaşta kadını etkileyebilmekle birlikte, ileri yaşlarda görülme sıklığı belirgin şekilde artmaktadır. Günümüzde her 12 kadından birisine meme kanseri tanısı konulmakta ve her 71 kadın ise bu hastalık nedeniyle hayatını kaybetmektedir. Kadınlarda görülen tüm kanser vakalarının yaklaşık %11,7'sini oluşturan meme kanseri, günümüzde en sık teşhis edilen kanser türü olmaya devam etmektedir [3]. Meme kanseri, memedeki hücrelerin anormal büyümesi nedeniyle oluşmaktadır. Memenin anatomisi farklı kan damarları, bağ dokuları, süt kanalları, lobüller ve lenf damarlarını içermektedir. Meme dokuları anormal şekilde büyüdüğünde ve hücre bölünmesi kontrolsüz hale geldiğinde süt kanallarında veya lobüllerde bir yumru ortaya çıkmaktadır [1], [4].

Meme kanserinin erken tespiti ve doğru şekilde teşhis edilmesi, hastalığın ilerlemesini önlemek ve olası komplikasyonları azaltmak açısından kritik öneme sahiptir. Bu şekilde zamanında ve etkili bir tedavi planlanarak hastalığın neden olduğu ölüm oranı azaltılabilmektedir. Meme kanseri tespiti için çeşitli görüntüleme yöntemleri kullanılmaktadır. Mamografi, Göğüs Termografisi, Manyetik Rezonans Görüntüleme, Pozitron Emisyon Tomografisi, Bilgisayarlı Tomografi, Ultrason ve Histopatolojik inceleme yaygın olarak kullanılan başlıca tanı yöntemleri arasında yer almaktadır [5], [6], [7]. Bu yöntemler arasında mamografi en sık kullanılan görüntüleme modalitesidir [8]. Mamografi, meme dokusunu görüntüleyen düşük dozlu bir meme grafisidir. Özellikle, herhangi bir semptom ortaya çıkmadan meme kanserini erken teşhis ederek ölüm oranını düşürmeye yardımcı olabileceği için en etkili ve hassas tarama yöntemi olarak kabul edilmektedir [2]. Mamografide, meme dokusunu farklı açılardan incelemek ve detaylı bilgi elde etmek amacıyla Kranio-Kaudal (Cranio-Caudal, CC) ve Mediolateral Oblik (Mediolateral Oblique, MLO) olmak üzere iki standart projeksiyon kullanılmaktadır [9]. Mamografide gözlemlenebilen meme kanseriyle ilişkili başlıca bulgular; kitle, kalsifikasyon, mimari bozulma ve asimetri olmak üzere sınıflandırılabilir. Bu bulgular arasında kitle, diğer anormalliklere kıyasla en belirgin ve klinik açıdan en önemli semptom olarak öne çıkmaktadır. Meme kitleleri, iyi huylu (benign) veya kötü huylu (malign) olarak

sınıflandırılabilir. Benign kitleler genellikle yavaş büyür ve tedavi gerektirmeksizin kaybolabilirken, malign kitleler hızlı büyüyerek vücuda yayılma riski taşıdığından daha tehlikelidir [10], [11].

Geleneksel olarak, meme kanseri tespiti ve teşhisi için radyologlar mamografi görüntülerini manuel olarak incelemekte ve değerlendirme sürecinde diğer uzmanlarla fikir birliği oluşturarak karar vermektedir. Olası meme kanseri tespiti için mamografi görüntülerinin manuel olarak incelenmesi yaygın bir yöntemdir. Ancak bu süreçte karşılaşılan bazı kaçınılmaz sınırlamalar, yanlış algılamalara neden olabilir ve tanı sürecini uzatabilir. Özellikle az gelişmiş bölgelerde radyoloji uzmanlarının yetersizliği, çoklu hastalık bulguları içeren görüntülerin yorumlanmasında deneyim eksikliği, yüksek hacimli görüntülerin günlük olarak değerlendirilmesinin zaman alıcı ve yorucu olması gibi faktörler tanı sürecinin etkinliğini azaltmaktadır. Ayrıca, meme dokusunun karmaşık anatomik yapısı ve tümörlerin heterojen özellikleri de manuel analiz sürecini daha da güçleştirmektedir [1]. Tüm bu sınırlamaların yanında, her bir mamogramın dikkatli bir şekilde değerlendirilmesi gerekmekte olup, bir görüntünün incelenmesi ortalama 30 ila 60 saniye sürebilmektedir. Buna rağmen, mamogramların radyologlar tarafından yorumlanmasında duyarlılık %77–87, özgüllük ise %89–97 olarak bildirilmektedir. Son yıllarda, tanısal doğruluğu artırmak amacıyla çoğu tarama programında çift okuma önerilmekte; ancak bu uygulama radyologların iş yükünü daha da artırmaktadır [12].

Ayrıca, yapılan bazı çalışmalar, daha önce normal olarak değerlendirilen mamogramlarda meme kanseri belirtilerinin sonradan %20–60 oranında tespit edilebildiğini ortaya koymaktadır [13]. Bu durum, manuel yorumlamanın sınırlılıklarını ve tanı sürecinde karşılaşılabilecek yanlış negatif ya da yanlış pozitif sonuçların ciddiyetini göstermektedir. Tanı sürecini uzatan ve hata olasılığını artıran bu tür sorunları azaltmak amacıyla, son yıllarda yapay zekâ tabanlı bilgisayar destekli tanı sistemlerinin geliştirilmesine yönelik çalışmalar önem kazanmıştır. Bu sistemler, radyologlara ikinci bir görüş desteği sağlayarak, bilgi kaybı olmadan tanısal doğruluğu artırmayı hedeflemektedir. Manuel inceleme ile karşılaştırıldığında, yapay zekâ tabanlı otomatik görüntü analizi, zaman alıcı ve tekrarlayan tarama süreçlerini ortadan kaldırmakta; ayrıca büyük hacimli görüntü verilerinden anlamlı tanısal özellikleri daha hızlı ve verimli bir şekilde çıkarabilmektedir [14], [15].

Mamografi görüntülerinden kitle tespiti ve sınıflandırması üzerine çalışan çoğu araştırmacı, makine öğrenimi ve derin öğrenme tekniklerini kullanarak meme kanseri teşhisini kolaylaştırmak için çeşitli algoritmalar geliştirmektedir [16]. Ancak, geleneksel makine öğrenimi temelli bilgisayar destekli tanı sistemlerinin genellikle sınırlı miktarda

veriyle çalışması ve büyük ölçüde elle çıkarılan özneliklere dayanması, bu sistemlerin genel performansını önemli ölçüde kısıtlamaktadır [10], [16]. Buna karşılık, son yıllarda büyük veri kümeleri ile çalışabilen ve otomatik öznelik çıkarımı gerçekleştirebilen derin öğrenme tabanlı nesne tespit algoritmaları, nesne tespiti ve konumlandırma alanında olağanüstü performans göstermiştir. Doğru ve etkili nesne tespit algoritmaları, çeşitli tıbbi durumların teşhisinde ve tedavi süreçlerinin desteklenmesinde hekimlere önemli katkılar sağlamaktadır. Bu algoritmalarından biri olan You Only Look Once (YOLO), anatomik yapılar, lezyonlar ve tümörleri tespit etme ve lokalize etme yeteneği sayesinde tıbbi görüntüleme alanında yoğun ilgi görmektedir. Özellikle tıbbi görüntülerdeki anormallikleri doğru ve etkili bir şekilde tanımlaması sayesinde meme kanseri, akciğer kanseri, kardiyovasküler hastalıklar ve nörolojik bozukluklar gibi pek çok hastalıkların erken tespitine ve teşhisine, tedavi planlamasına ve hastalık sürecinin izlenmesine yardımcı olabilmektedir [17]. Bu nedenlerle, YOLO algoritması son yıllarda bilgisayar destekli tanı sistemleri geliştiren araştırmacıların odak noktalarından biri haline gelmiştir.

Bu tez çalışması kapsamında, mamografi görüntülerinden meme kanseri tespiti ve sınıflandırılması için YOLO tabanlı bir model geliştirilmesi hedeflenmektedir. Bu doğrultuda, farklı YOLO mimarileri üzerinde analizler gerçekleştirilecek ve model performansını artırmak için dikkat mekanizmalarının bu modellere entegrasyonu gerçekleştirilecektir. Ayrıca, farklı evrişim türlerinin model başarımı üzerindeki etkileri de incelenecektir. Bu süreçte hem açık erişimli veri kümeleri hem de gerçek klinik veriler kullanılacaktır. Çalışma sonucunda geliştirilecek modelin, radyologlara ikinci görüş desteği sunarak mamografi değerlendirme sürecine katkı sağlayabileceği düşünülmektedir.

## 1.2. Literatür Özeti

1913 yılında keşfedilmesinden bu yana mamografi, meme kanserine bağlı şüpheli lezyonların erken evrede tespit ve teşhisinde önemli bir araç olarak kabul edilmektedir. Mamografi taramaları, meme kanseri tanısına yardımcı olmuş ve çeşitli çalışmalar, bu yöntem sayesinde ölüm oranlarında anlamlı bir azalma sağlandığını ortaya koymuştur [18]. Bilgisayarlı görü ve yapay zekâ alanındaki kayda değer ilerlemelerle birlikte, son yıllarda yapılan çok sayıda çalışma, bilgisayar destekli tanı sistemlerinin mamografi görüntülerinden şüpheli lezyonları otomatik olarak tespit etmedeki etkinliğini göstermiştir [19], [20], [21], [22], [23]. Sinir ağı tabanlı modellerin kullanıma girmesiyle birlikte, bu sistemlerde

kullanılan yaklaşımlar dönüşüme uğramış; elle çıkarılan öznitelikler yerine, karmaşık öznitelikleri otomatik olarak öğrenebilen derin öğrenme mimarileri ön plana çıkmıştır [24].

Bu gelişmelerin ardından, özellikle son yıllarda derin öğrenme tabanlı nesne tespiti algoritmaları, meme kanserinin tespiti ve sınıflandırılması alanında yaygın şekilde kullanılmaya başlanmıştır. Bu alandaki ilk dikkat çekici uygulamalar, Evrişimli Sinir Ağı (Convolutional Neural Network, CNN) tabanlı iki aşamalı nesne tespit yöntemleri olan Bölge-tabanlı CNN (Region-based CNN, R-CNN), Hızlı R-CNN (Fast R-CNN), Daha Hızlı RCNN (Faster R-CNN) ve Maske RCNN (Mask RCNN) gibi modeller ile gerçekleştirilmiştir [10].

Ribli et al. [25], benign ve malign lezyonları tespit etmek ve sınıflandırmak amacıyla Faster R-CNN tabanlı bir sistem önermişlerdir. Geliştirilen sistem, DDSM veri kümesinde eğitilmiş, INBreast veri kümesi üzerinde değerlendirilerek 0.95 AUC değerine ulaşmıştır. Ayrıca, Dijital Mamografi DREAM Mücadelesinde 0.85'lik bir AUC başarısı ile ikinci sırada yer almıştır. Bununla birlikte, sistem bir dedektör olarak kullanıldığında, INBreast veri kümesinde görüntü başına çok az yanlış pozitif işaretleme ile yüksek hassasiyete ulaşmıştır.

Peng et al. [26], farklı şekil ve boyutlardaki meme lezyonlarının tespitini iyileştirmek için deforme edilebilir evrişimli katmanları içeren Faster R-CNN'e dayalı bir sistem önermişlerdir. Ayrıca, küçük lezyonların algılanabilirliğini artırmak amacıyla çok ölçekli öznitelik piramit ağı (multiscale-feature pyramid network) sisteme entegre edilmiştir. Geliştirilen sistem, CBIS-DDSM ve INBreast veri kümelerinde sırasıyla 0.93 ve 0.95 doğru pozitif oranı elde ederek başarılı bir performans göstermiştir.

Agarwal et al. [27] çalışmalarında meme kitlesi tespiti için Faster R-CNN tabanlı bir çerçeve sunmuşlardır. İlk olarak model OPTIMAM Mamografi Görüntü Veritabanı'nın büyük bir alt kümesinde (OMI-H) eğitilmiş, ardından transfer öğrenmesi yaklaşımı ile daha küçük veri kümelerinde (OMI-G ve INBreast) ince ayar yapılarak performansı değerlendirilmiştir. İnce ayarlı model OMI-G veri kümesinde 0.92 doğru pozitif oranı ve 0.80 F1 skoru elde etmiştir. INBreast veri kümesinde ise malign kitleler için 0.99 doğru pozitif oranı ve 0.86 F1 skoru; benign kitleler için 0.85 doğru pozitif oranı ve 0.74 F1 skoruna ulaşılmıştır.

Raza et al. [28], meme tümörü tespiti, sınıflandırması ve segmentasyonu için Mask R-CNN modeline dayalı bir yaklaşım önermişlerdir. Önerilen yaklaşım, CBIS-DDSM veri kümesinde 0.75 hassasiyet, 0.80 geri çağırma ve 0.825 F1-skoru ile yaklaşık %85 doğruluk elde etmiştir.

Jahangeer et al. [29], çalışmalarında meme kanseri tespiti için hibrit bir derin öğrenme çerçevesi sunmuşlardır. Çalışmada ilk olarak, mamografi görüntüleri geliştirilmiş bulanık tabanlı bir medyan filtre ile gürültüden arındırılmıştır. Ardından, bir Mask R-CNN mimarisi kullanılarak lezyon bölgelerinin segmentasyonu gerçekleştirilmiştir. Son olarak, benign ve malign lezyonları ayırt etmek için topluluk tabanlı bir CNN mimarisiyle sınıflandırma yapılmıştır. Önerilen model, MIAS ve DDSM veri kümelerinde sırasıyla %98,75 ve %97,82 sınıflandırma doğruluğu elde etmiştir.

İki aşamalı nesne tespit yöntemleri başarılı performanslar göstermelerine rağmen, genellikle daha uzun işlem süreleri ve yüksek bellek gereksinimleri gibi sınırlamalara sahiptir. Bu nedenle, çok sayıda mamografinin değerlendirildiği meme kanseri taramalarında, hız ve kaynak verimliliği açısından tek aşamalı nesne tespit yöntemleri (Tek Atış Çoklu Kutu Dedektörü (Single Shot MultiBox Detector, SSD), RetinaNet ve YOLO) daha uygun bir alternatif sunmaktadır.

Singh et al. [30], tarafından yapılan çalışmada, meme tümörlerini lokalize etmek amacıyla ilk adımda SSD modeli kullanılmış, ardından bu bölgeler segmentasyon ve sınıflandırma işlemine tabi tutulmuştur. Önerilen model, INBreast veri kümesinde 0,97'lik doğru pozitif oranına ulaşmıştır.

Jung et al. [31], çalışmalarında RetinaNet tabanlı bir kitle tespit modeli tanıtmışlardır. Önerilen model, INBreast veri kümesi ve özel bir veri kümesi (GURO) üzerinde değerlendirilmiş, çapraz veri kümesi ve transfer öğrenme senaryoları dahil olmak üzere yedi deneysel kurulumda başarılı sonuçlar elde etmiştir. Çalışmanın sonucunda model, görüntü başına düşük yanlış pozitif oranlarıyla birlikte %99'a kadar doğru pozitif oranına ulaşarak, yüksek doğruluk ve genellenebilirlik sergilemiştir.

YOLO, son teknoloji yöntemlerle karşılaştırıldığında tespit ve sınıflandırma görevlerini eşzamanlı ve hızlı bir şekilde gerçekleştirmedeki başarısıyla öne çıkmakta; bu özelliği sayesinde derin öğrenme tabanlı uygulamalarda uygun bir yaklaşım olarak değerlendirilmektedir [17], [32].

Al-masni et al. [33] meme kitlelerinin tespiti ve sınıflandırılması için YOLOv1 tabanlı bir yaklaşım önermişlerdir. DDSM veri kümesinden seçtikleri 600 mamografi görüntüsünü üç farklı açı (90°, 180° ve 270°) ile döndürerek veri kümesini genişletmişlerdir ve toplamda 2400 artırılmış mamografi görüntüsünün %80'ini eğitim, %20'sini test olarak ayırmışlardır. Önerilen sistem hem orijinal hem de genişletilmiş veri kümesinde değerlendirilmiştir. Sonuç

olarak, beş katlı çapraz doğrulama testleri ile sistem genişletilmiş veri kümesinde %99,7 tespit doğruluğu ve %97 sınıflandırma doğruluğu göstermiştir.

Al-antari et al. [34], meme lezyonlarını tespit ve sınıflandırmak amacıyla derin öğrenme tabanlı bir model önermiştir. İlk olarak, mamografi görüntülerinden lezyonları saptamak amacıyla YOLOv2 kullanılmış, ardından CNN, ResNet-50 ve InceptionResNet-v2 olmak üzere üç farklı derin öğrenme sınıflandırıcı modeli ile lezyon sınıflandırması gerçekleştirilmiştir. Önerilen sistem DDSM ve INBreast veri kümelerinde 5 katlı çapraz doğrulama yöntemi ile değerlendirilmiştir. Çalışma sonucunda, YOLOv2 DDSM ve INBreast veri kümeleri için sırasıyla %99,17 ve %97,27 genel tespit doğruluğu, %99,28 ve %98,02 F1 skoru ile yüksek tespit başarısı elde etmiştir. Sınıflandırma aşamasında ise CNN, ResNet-50 ve InceptionResNet-v2 modelleri sırasıyla DDSM veri kümesinde %94,50, %95,83 ve %97,50; INBreast veri kümesinde ise %88,74, %92,55 ve %95,32 genel doğruluk oranlarına ulaşarak umut verici sonuçlar elde etmiştir.

Aly et al. [35] tarafından yapılan çalışmada meme kitlesi tespiti ve sınıflandırması için YOLOv3 modeli YOLOv1 ve YOLOv2'ye karşı değerlendirilmiş ve üç model arasındaki temel farkları tartışılmıştır. Çalışmada, INBreast veri kümesi üzerinde iki farklı veri artırma yaklaşımı uygulanmıştır. İlk yaklaşımda, her bir mamogramı 90°, 180° ve 270°'lik üç açıyla döndürerek tüm veri kümesini genişletmişler, ardından artırılmış verileri %80 eğitim ve %20 test kümesi olacak şekilde bölünmüştür. İkinci yaklaşımda ise orijinal veri kümesi önce %80 eğitim ve %20 test kümesi olarak bölünmüş, ardından sadece eğitim kümesindeki mamogramları üç açıyla döndürerek eğitim kümesini genişletmişlerdir. İlk yaklaşım daha yüksek sonuçlar verse de gerçekçi senaryolarda uygulanacak en doğru yaklaşımın yalnızca eğitim setini genişletmek olduğu sonucuna varılmıştır. Ayrıca bu çalışmada, farklı YOLO modelleri ve veri artırma yaklaşımlarının tespit performansına etkisinin yanında giriş görüntüsünün farklı boyuttaki çözünürlükleri, farklı sayıda anchor kutuları, ResNet ve InceptionV3 gibi farklı öznitelik çıkarıcıların da performansa katkısı araştırılmıştır. Sonuçta, en iyi performans ikinci veri artırma yaklaşımı (yalnızca eğitim kümesinin genişletilmesi) kullanılarak, 832 x 832 giriş görüntüsü boyutunda ve eğitim için kullanılan anchor sayısı "9" iken elde edilmiştir. Kitlelerin %89,4'ü tespit edilmiştir ve bu kitlelerin benign ve malign olarak sınıflandırılması için ortalama hassasiyet değerleri sırasıyla %94,2 ve %84,6'dır.

Yan et al. [36], otomatik meme kitle tespiti amacıyla CC ve MLO mamogramlarını birleştiren yenilikçi bir yaklaşım sunmuşlardır. Bu kapsamda, bölge adaylarını belirlemek için YOLOv3'ten yararlanarak, aynı anda yama düzeyinde kitle sınıflandırması ve çift görünüm eşleştirmesi gerçekleştiren Siamese tabanlı bir mimari önermişlerdir. Çalışmada,

586 CBIS-DDSM mamogram çifti %80 eğitim ve %20 doğrulama için kullanılırken, 35 INBreast mamogram çifti yalnızca test aşamasında değerlendirilmiştir. YOLOv3 tabanlı başlangıç modeli, test verisinde 0.86 doğru pozitif oranı elde ederken, önerilen yöntem bu oranı 0.96'ya yükseltmiştir. Bununla birlikte, önerilen model %94,78 AUC ve 0.8791 genel sınıflandırma doğruluğu ile yüksek performans göstermiştir.

Baccouche et al. [37], meme lezyonlarını eş zamanlı olarak tespit edip kitle veya kalsifikasyon olarak sınıflandırmak için YOLOv3 tabanlı bir sistem sunmuşlardır. Önerilen sistem CBIS-DDSM, INBreast ve kendi özel veri kümelerinde değerlendirilmiştir. Çalışmanın sonucunda, CBIS-DDSM, INBreast ve özel veri kümelerinde kitle için sırasıyla %95,7, %98,1 ve %98 tespit doğruluğu elde edilirken kalsifikasyon için %74,4, %71,8 ve %73,2 tespit doğruluğuna ulaşılmıştır. Çalışmanın devamında, doğru tespit edilen kitle ve kalsifikasyonların bölütlenmesi, benign ve malign olarak sınıflandırılması ve malignite derecesi tahminini yapılması amaçlanmaktadır.

Hamed et al. [32], çalışmalarında mamogramlardaki şüpheli kitleleri lokalize etmek ve sınıflandırmak için YOLOv4 tabanlı bir yaklaşım önermiştir. Bu yaklaşım üç aşamadan oluşmaktadır. İlk olarak, görüntüler ön işleme tabii tutularak artefaktlardan arındırılmış ve üst üste binen küçük dilimler halinde kırılmıştır. İkinci aşamada, lezyonlar tam mamogramlar ve YOLOv4 ile modeli yapılandırıldıktan sonra kırılmış dilimlerin tespiti olmak üzere iki yol ile lokalize edilmiştir. Son olarak, bu lezyonların sınıflandırılması için ResNet, VGG, Inception vb. gibi öznitelik çıkarıcılar kullanılmış ve performansları YOLOv4 ile karşılaştırılmıştır. INBreast veri kümesinde yapılan değerlendirme sonuçlarına göre, önerilen sistem %97,86 oranında tespit doğruluğuna ve Inception-V3 ile %95 sınıflandırma doğruluğuna sahiptir.

Su et al. [38] çalışmalarında meme kitlelerinin eş zamanlı tespiti ve bölütlenmesi için YOLO ve LOGO (Yerel-Global (Local-Global) mimarilerinin kombinasyonundan oluşan bir model önermiştir. İlk aşamada, YOLOv3 ve YOLOv5'in çeşitli varyasyonları CBIS-DDSM ve INBreast veri kümeleri üzerinde değerlendirilmiş; en yüksek başarıyı gösteren YOLOv5L6 modeli ile kitle bölgeleri yüksek çözünürlükte tespit edilip kırılmıştır. İkinci aşamada ise, kırılan görüntüler kullanılarak LOGO tabanlı segmentasyon gerçekleştirilmiş ve YOLO-LOGO olarak adlandırılan model oluşturulmuştur. Önerilen modelin tespit başarımı, CBIS-DDSM ve INBreast veri kümelerinde sırasıyla %65 ve %61,4 mAP değerleriyle ölçülmüştür. Bölütleme aşamasında ise model, CBIS-DDSM veri kümesinde %74,52 F1 skoru ve %64,04 IoU, INBreast veri kümesinde ise %69,37 F1 skoru ve %61.09 IoU değerlerine ulaşarak başarılı sonuçlar elde etmiştir.

Prinzi et al. [39] tarafından yapılan çalışmada meme kanseri tespiti için transfer öğrenme yaklaşımı kullanılarak YOLO tabanlı bir model sunulmuştur. Bu amaçla kaynak veri kümesi olarak CBIS-DDSM sıfırdan eğitilmiş ve hedef veri kümesi olarak INBreast ve özel veri kümesinde ince ayar gerçekleştirilmiştir. Eğitim aşamasından önce benign ve malign sınıfları arasındaki veri sayısını dengelemek için INBreast ve özel veri kümelerine 180° döndürme, yatay ve dikey çevirme gibi bazı veri artırma teknikleri uygulanmıştır. Çalışmada, YOLOv3, farklı YOLOv5 versiyonları (YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l) ve YOLOv5-Transformer dahil olmak üzere çeşitli YOLO mimarileri karşılaştırılmıştır. Analizler sonucu en başarılı model YOLOv5s olup INBreast ve özel veri kümesinde sırasıyla 0.835 ve 0.621 mAP@0.5 değerleri elde edilmiştir. Ek olarak, Eigen-CAM, mamogramdaki tüm şüpheli ilgi alanlarını vurgulayarak modelin geliştirilmiş çıktısını görselleştirilmesi için uygulanmıştır.

Quiñones-Espín et al. [40] tarafından yapılan çalışmada, meme kitlesi tespiti için YOLOv5'e dayalı otomatik bir sistem tasarlanması amaçlanmıştır. Bu doğrultuda YOLOv5x ve YOLOv5s modelleri VinDr-Mammo veri kümesinde değerlendirilmiştir. VinDr-Mammo veri kümesinin %80'i eğitim, %20'si doğrulama için ayrılmıştır. MIAS veri kümesi ise test için kullanılmıştır. Çalışmanın sonucunda, en iyi model YOLOv5x ile validasyon kümesinde %80 hassasiyet ve 0.602 mAP@0.5, test kümesinde ise %72 hassasiyet ve 0.488 mAP@0.5 değerlerine ulaşılmıştır.

Anas et al. [41], meme kitlesi tespiti ve sınıflandırması için YOLOv5 ve Mask R-CNN modellerini birleştiren hibrit bir derin öğrenme modeli geliştirmiştir. YOLOv5 modeli meme kitlelerini etkili bir şekilde tespit edip sınıflandırmasını yapabilmektedir; ancak özellikle karmaşık meme dokularında daha iyi performans elde edebilmek için mimari bazı iyileştirmeler gerekmiştir. Bu doğrultuda, tümör sınırlarını ve boyutlarını hassas şekilde belirlemek ve evre tahmini yapılmasına olanak tanımak amacıyla Mask R-CNN modeli sisteme entegre edilmiştir. Önerilen hibrit model INBreast ve CBIS-DDSM veri kümeleri üzerinde eğitilmiş ve değerlendirilmiştir. Elde edilen sonuçlar, temel YOLOv5 modeline kıyasla anlamlı performans artışları göstermiştir; mAP değeri %88,90'dan %95,20'ye, hassasiyet %91,08'den %96,93'e, Matthews Korelasyon Katsayısı ise %83'ten %92,20'ye yükselmiştir.

Tan et al. [42], çalışmalarında küçük meme kitlelerinin tespiti ve sınıflandırmasını iyileştirmek amacıyla, YOLOv5 mimarisi üzerine inşa edilmiş gelişmiş bir model olan RCM-YOLO'yu tanıtmıştır. Önerilen model, üç özgün bileşen içermektedir: alıcı alanı genişletmek için tasarlanan Kalan Asimetrik Genişletilmiş (Residual Asymmetric Dilated,

RAD) modülü, katmanlar arası öznitelik aktarımını güçlendiren Katmanlar Arası Dikkat (Cross-Layer Attention, CLA) modülü ve pozitif örnek seçimini optimize eden Ortalama-Standart Sapma Uyarlamalı Seçim (Mean-Standard Deviation Adaptive Selection, MSAS) algoritması. Bu modüller, YOLOv5'in küçük lezyonları tespit etmedeki sınırlamalarını aşmayı hedeflemektedir. INBreast ve CBIS-DDSM veri kümeleri üzerinde değerlendirilen RCM-YOLO, %90,34 mAP elde ederek temel YOLOv5 modelini (%85,82 mAP) geride bırakmış ve kaçırılan tespit oranını %34'ten %11'e düşürmüştür.

Lan et al. [43], meme kanseri tespitinin doğruluğunu ve verimliliğini artırmak için YOLOv8 tabanlı iyileştirilmiş bir model geliştirmiştir. Bu kapsamda, daha zengin öznitelik yakalarken parametre sayısını neredeyse yarıya indiren GhostConv ve C3Ghost modüllerini içeren YOLOv8-GHOST modeli ile öznitelik haritasındaki kanal sayısını düzene sokarak gerekli parametre sayısını azaltan YOLOv8-P2 algoritmasının avantajlarını birleştirerek "YOLOv8-P2" modelini önermiştir. Önerilen model ile MIAS ve DDSM veri kümelerinde veri artırımı yapılarak analizler gerçekleştirilmiştir. MIAS veri kümesinde, YOLOv8-P2 modeli %99 hassasiyet, %98,14 duyarlılık ve %98,57 F1 skoru elde etmiştir. DDSM veri kümesinde ise %87,10 hassasiyet, %96,46 duyarlılık ve %91,53 F1 skoru elde edilmiş olup, bu sonuçların orijinal YOLOv8 performansından daha düşük olduğu gözlemlenmiştir. YOLOv8-P2 modeli artırılmış veri kümesi üzerinde yüksek doğruluk göstermiştir; ancak, sınırlı sayıdaki orijinal veriden 10.000'den fazla veri üretilmesi performans ölçümlerinin gerçekçi olmadığını düşündürmektedir. Bu nedenle, modelin gerçek veri çeşitliliğine karşı genelleme yeteneği sorgulanabilir.

Son yıllarda dikkat mekanizmaları, tıbbi görüntüleme alanında kullanılan derin öğrenme modellerine giderek daha fazla entegre edilmektedir. Bu eğilimin temelinde, klinik olarak anlamlı öznitelikleri vurgulama ve küçük ya da belirsiz lezyonlara odaklanma yeteneği yatmaktadır. Bu tür avantajlar, dikkat tabanlı mimarileri modern tanı sistemleri için etkili bir bileşen haline getirmektedir [44].

Coskun et al. [45], kitle tespiti ve sınıflandırma performansını artırmak amacıyla dönüştürücü (transformer) tabanlı bir model önermiştir. İlk olarak, INBreast veri kümesi kullanılarak YOLOv3, YOLOv4 ve farklı YOLOv5 versiyonlarının performansları karşılaştırılmıştır. Ardından, çok ölçekli öznitelik çıkarımını iyileştirmek amacıyla YOLOv5x omurgasına Swin Dönüştürücü entegre edilmiştir. Dönüştürücü tabanlı YOLOv5x modeli, %94,7 mAP, %94,4 hassasiyet, %88 duyarlılık ve %91 F1 skoru ile 832×832 çözünürlükte en iyi performansı elde ederek diğer YOLO versiyonlarına kıyasla daha yüksek başarı göstermiştir. Çalışmanın bulguları, dikkat mekanizmalarının bir türü olan

dönüştürücü yapılarının entegrasyonunun, kitle tespitinde performans artışı açısından umut verici bir yön olduğunu ortaya koymaktadır.

Hassan et al. [46], tam alan dijital mammogramları (full-field digital mammogram, FFDM) ve kontrastlı spektral mammogramları (contrast-enhanced spectral mammogram, CESM) kullanarak, kitle tespiti için YOLOv4 ve kitle sınıflandırması için ViT tabanlı bir model önermiştir. Geliştirilen model, INBreast ve CDD-CESM (CE-CESM ve DM-CESM) veri kümeleri üzerinde değerlendirilmiştir. Model, kitle tespitinde INBreast, CE-CESM ve DM-CESM veri kümeleri için sırasıyla %98,69, %81,52 ve %71,65 mAP değerlerine ulaşmıştır. Sınıflandırma doğruluğu ise sırasıyla %95,65, %97,61 ve %80 olarak elde edilmiştir. Buna ek olarak, model yalnızca genel kitle tespiti yapmakla kalmamış; aynı zamanda INBreast veri kümesinde benign ve malign kitlelerin tespiti ve sınıflandırılması görevinde %84,43 mAP performansı göstermiştir.

Zhang et al. [47], mamografi görüntülerinde kitlelerin küçük boyutlu olması, düşük kontrast ve belirsiz sınırlar gibi tespit zorluklarını aşmak için OMHIB-YOLO adlı geliştirilmiş bir model önermiştir. Model, YOLOv5 mimarisi temel alınarak tasarlanmış olup, tespit performansını artırmak için üç temel bileşen birleştirilmiştir: hafif ve hızlı bir yapı için MobileNetV2 omurgası, farklı boyutlardaki özellikleri yakalayabilen çoklu çıkış başlıkları (Multi-scale Dynamic Head), ve daha güçlü özellik çıkarımı için dinamik evrişim katmanları. Önerilen model, DDSM, MIAS ve INBreast veri kümeleri üzerinde değerlendirilmiştir. DDSM veri kümesinde %98,5 hassasiyet, %94,8 duyarlılık ve %94,8 mAP elde edilmiştir. Beş katlı çapraz doğrulama yöntemiyle, MIAS veri kümesinde %98,1 hassasiyet, %97,9 duyarlılık ve %98,8 mAP, INBreast veri kümesinde ise %97,9 hassasiyet, %86,9 duyarlılık ve %90,7 mAP elde edilmiştir.

Du et al. [48] çalışmalarında, meme tümörü tespiti ve sınıflandırması için geliştirilmiş bir YOLOv7 mimarisine dayanan bir model sunmuştur. Model, özellikle arka plan parazitini azaltmak ve anlamlı özniteliklere odaklanmak amacıyla CBAM dikkat modülü ile güçlendirilmiştir. Ayrıca tespit doğruluğunu ve hesaplama verimliliğini artırmaya yönelik çeşitli yapısal iyileştirmeler yapılmıştır. Çalışma kapsamında oluşturulan özel veri kümesinde eğitilen model, %97,01 hassasiyet, %97,98 duyarlılık ve %90,78 AP elde ederek başarılı bir performans göstermiştir. Temel YOLOv7 modeli ile karşılaştırıldığında, yapılan iyileştirmelerle modelin doğruluğu ve başarısı önemli ölçüde artmıştır.

Al-Tam et al. [49], meme kanserini tespit etmek ve sınıflandırmak amacıyla yenilikçi bir tanı sistemi önermiştir. Önerilen sistem, kitle tespiti için YOLOv8'i ve ikili sınıflandırma (benign-malign) için Görü Dönüştürücü (Vision Transformer, ViT) tabanlı ResNet50

modellerinden oluşan bir topluluk (ensemble) yapısını birleştirmiştir. Model, INBreast ve MIAS veri kümeleri üzerinde test edilmiştir. Tespit aşamasında, INBreast ve MIAS veri kümelerinde sırasıyla %96,10 ve %95,65 mAP değerlerine ulaşılmıştır. Sınıflandırma aşamasında, INBreast veri kümesinde %95,65 doğruluk, %96,67 AUC, %94 F1 skoru, MIAS veri kümesinde ise %91,66 doğruluk, %92,86 AUC, %91,5 F1 skoru elde edilmiştir. Çalışmada önerilen hibrit yapı, hem YOLOv8'in başarılı şekilde tespit ettiği lezyonlarda yüksek doğrulukta sınıflandırma yapılmasına hem de YOLOv8'in tespit edemediği durumlarda tüm görüntü üzerinden analiz gerçekleştirilmesine olanak tanımıştır.

Güncel YOLO sürümleri arasında, YOLOv8 ile yapılan sınırlı sayıda çalışma bulunmakla birlikte, literatürdeki gelişmeleri yansıtabilmek amacıyla YOLOv8 ve YOLOv9 modellerini temel alan konferans çalışmaları da özetlenmiştir.

Mahoro et al. [50], meme kitlesi tespiti ve sınıflandırması için YOLOv7 ve YOLOv8 modellerinin performanslarını karşılaştırmıştır. Bu amaçla, VinDr-Mammo veri kümesindeki görüntülere kontrast sınırlı adaptif histogram eşitleme, medyan filtreleme ve bilateral filtreleme gibi iyileştirme teknikleri uygulanmış ve her bir tekniğin performansı ayrı ayrı analiz edilmiştir. Deneysel sonuçlar, YOLOv8'in VinDr-Mammo veri kümesinde YOLOv7'ye kıyasla daha iyi bir performans gösterdiğini ortaya koymuş; özellikle medyan filtreli görüntülerde, YOLOv7 modelinin 0.53 mAP@0.5 değerine karşılık, YOLOv8 modelinin 0.65 mAP@0.5 değeri elde ettiği gözlemlenmiştir.

Alhsnony et al. [51] çalışmalarında, MIAS ve DDSM veri kümelerini kullanarak mamografi görüntülerinden meme kanseri tespiti için YOLOv8x modelini kullanmıştır. Kapsamlı veri artırma teknikleri uygulanmış ve analizler, MIAS, DDSM ve birleşik veri kümelerinin hem orijinal hem de artırılmış halleri üzerinde olmak üzere altı farklı senaryoda gerçekleştirilmiştir. Deneysel sonuçlar, YOLOv8x modelinin artırılmış birleşik veri kümesi üzerinde eğitildiğinde %96,21 mAP elde ettiğini ve bireysel ya da artırılmamış veri kümeleri üzerinde yapılan eğitimlere kıyasla daha yüksek performans sergilediğini göstermiştir.

Marchi et al. [52], CBIS-DDSM veri kümesini kullanarak mamografi görüntülerinden meme kanseri tespiti için YOLOv8 (YOLOv8s ve YOLOv8l) ve YOLOv9c modellerinin performansını karşılaştırmıştır. Elde edilen sonuçlar, YOLOv9c modelinin 0.677 mAP@0.5 ve 0.686 duyarlılık değerleri ile YOLOv8s ve YOLOv8l modellerini geride bırakarak en yüksek performansı sergilediğini göstermiştir. Bununla birlikte, en yüksek hassasiyet (0.788) YOLOv8l modeline aittir. Sonuçlar, YOLOv9c modelinin gelecekteki uygulamalar için umut verici bir potansiyele sahip olduğunu göstermektedir.

### 1.3. Genel Bakış

Sunulan tez çalışması temel olarak beş bölümden oluşmaktadır. Birinci bölümde çalışmanın konusu, amacı ve önemi açıklanmış, mamografi görüntülerinden kitle tespiti ve sınıflandırması üzerine yapılan literatür çalışmaları özetlenmiştir. İkinci bölümde, meme kanseri, meme anatomisi, görüntüleme yöntemleri ve mamografik bulgular gibi tez çalışmasının kuramsal altyapısını oluşturan kavramlara yer verilmiştir. Üçüncü bölümde, çalışmada kullanılan veri kümeleri, nesne tespit yöntemleri, YOLO modelleri, dikkat mekanizmaları ve evrişim modülleri detaylı bir şekilde ele alınmıştır. Dördüncü bölümde, önerilen modellerin uygulama süreçleri ve deneysel analizleri aktarılmış; farklı yapıların performansa etkileri karşılaştırmalı olarak değerlendirilmiştir. Beşinci ve son bölümde ise çalışmanın genel değerlendirmesi yapılmış, literatürdeki çalışmalarla karşılaştırmalar gerçekleştirilmiş ve ileriye dönük önerilere yer verilmiştir.

Sonuç olarak, bu tez çalışmasında mamografi görüntülerinden meme kanseri tespiti ve sınıflandırılması için YOLO tabanlı bir model önerilmiştir.

## 2. KURAMSAL TEMELLER

### 2.1. Meme Kanseri

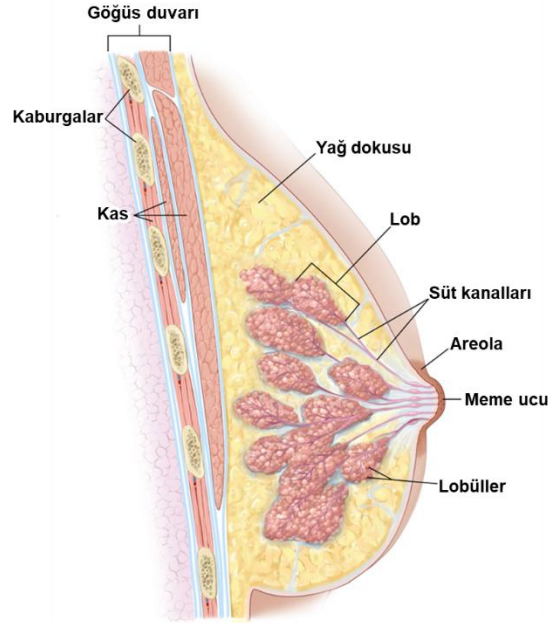
Kanser, vücut dokularına zarar veren anormal hücrelerin kontrolsüz çoğalmasıyla karakterize edilen ve dünya çapında önde gelen ölüm nedenleri arasında yer alan ciddi bir hastalıktır [53], [54]. Dünya Sağlık Örgütü (World Health Organization, WHO), 2022 yılındaki yaklaşık 20 milyon kanser vakasının 2050 yılında 35 milyonu aşabileceğini tahmin etmektedir. Bu da yaklaşık %77'lik bir artış anlamına gelmektedir [55].

Meme kanseri, dünya çapında kadınlar arasında en sık görülen kanser türü olarak öne çıkmaktadır. 2022 yılında dünya genelinde 2,3 milyon kadına meme kanseri teşhisi konmuş ve 670.000 ölüm gerçekleşmiştir [56]. Vakaların ve ölümlerin büyük bir kısmı 50 yaş ve üzeri bireylerde gözlemlenmekte olup, bu yaş grubundaki bireyler yeni vakaların %71'ini ve meme kanserine bağlı ölümlerin %79'unu oluşturmaktadır [57]. Uluslararası Kanser Araştırmaları Ajansı (International Agency for Research on Cancer, IARC) ve işbirlikçileri tarafından yürütülen bir çalışmaya göre, dünya çapında her 20 kadından birinin yaşamı boyunca meme kanseri tanısı alması beklenmektedir. Mevcut eğilimlerin devam etmesi halinde, 2050 yılına kadar yıllık yeni meme kanseri vaka sayısının 3,2 milyona, bu hastalıkla ilişkili ölüm sayısının ise 1,1 milyona ulaşacağı öngörülmektedir [57].

### 2.2. Meme Anatomisi ve Fizyolojisi

Meme ön torasik duvarda, ikinci ve altıncı interkostal kıkırdakları arasında yer almaktadır [58]. Bir yetişkin meme, farklı tipte glandüler, yağlı ve lifli (fibröz) dokudan oluşmaktadır. Şekil 2.1'de memenin anatomik yapısı gösterilmiştir. Glandüler doku, 15-20 lobdan oluşur ve bu lobların her biri süt üreten bez yapısına sahip lobül adı verilen birçok küçük yapı içermektedir. Loblar ve lobüller ile bağlantılı olan süt kanalları, sütü areolanın (meme ucunu çevreleyen koyu renkli alan) ortasında bulunan meme uçlarına taşımaktadır. Her meme ucunda yaklaşık dokuz süt kanalı ve yüzlerce sinir bulunmaktadır. Lifli dokular, yağ dokusunu ve lobülleri yerinde tutmaktan sorumludur. Yağlı doku (yoğun olmayan meme dokusu) ise glandüler ve lifli doku arasındaki boşlukları doldurmakta ve büyük ölçüde meme boyutunu belirlemektedir [59], [60]. Ayrıca, meme içindeki lenf düğümleri ve damarlar, lenf sıvısının ve beyaz kan hücrelerinin iletilmesini sağlamaktadır. Lenf damarları ve kan damarları ise memenin her yerinde bulunabilir [60], [61]. Ek olarak, meme dokusunu yerinde tutmak için deriden göğüs duvarına uzanan bağlar adı verilen destekleyici, esnek bağ dokusu

bantları da vardır. Pektoral kas, her iki memenin altında göğüs duvarına yaslanarak onlara destek sağlamaktadır [60].



Şekil 2.1. Memenin anatomik yapısı [62].

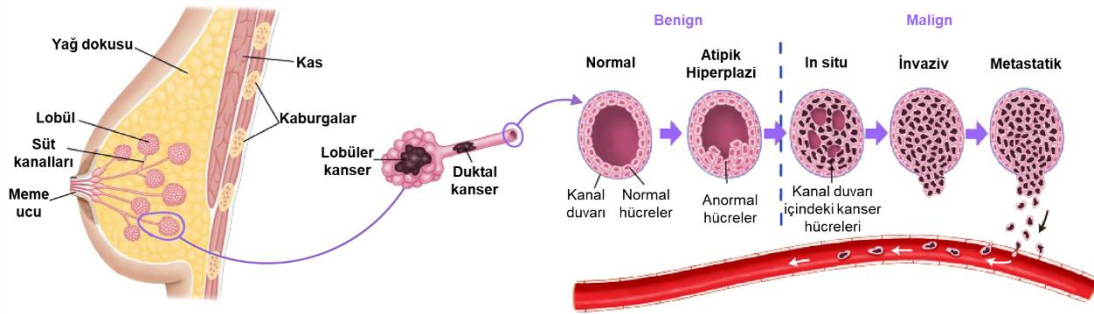
İnsan yaşamı boyunca, meme; doğum, ergenlik, hamilelik, emzirme ve menopoz dönemlerinde, vücudun diğer herhangi bir bölümüne kıyasla çok daha fazla yapısal ve işlevsel değişime uğrayan bir organdır [59]. Meme gelişimi, ergenliğin başlangıç aşamalarında çeşitli hormonlar tarafından uyarılır. Östrojen ve progesteron, bu dönemde meme büyümesinden ve gelişiminden sorumlu olan ana hormonlardır. Östrojen, duktal gelişimi uyarır; progesteron, lobüler gelişimi ve epitel farklılaşmasını uyarır [59], [60].

### 2.3. Meme Kanserinin Gelişimi, Risk Faktörleri ve Belirtileri

Meme kanseri, meme dokusundaki hücrelerin anormal ve kontrolsüz çoğalması sonucu gelişen bir kanser türüdür. Şekil 2.2’de meme kanserinin gelişim süreci gösterilmektedir. Meme kanseri hücreleri, süt kanallarının içinde veya memenin süt üreten lobüllerinin içinde köken alır. Duktal kanserler kanallardan kaynaklanan meme kanserini tanımlarken; lobüler kanserler, lobüllerden kaynaklanan kanseri ifade etmektedir. Daha nadir görülen vakalarda ise meme kanseri, memenin yağlı ve lifli bağ dokularını içeren stromal dokularda da başlayabilir [63], [64].

Meme dokusunda gelişen bir tümör, benign veya malign özellik gösterebilir. Atipik hiperplazi gibi bazı erken evre lezyonlar, başlangıçta benign özellikler sergilemeler de zaman

içerisinde malign bir sürece evrilebilme potansiyeline sahiptir [64]. Benign tümörler kanserli yapılar olarak kabul edilmez, genellikle yavaş büyür, çevre dokulara zarar vermez ve vücudun diğer bölgelerine yayılmaz. Malign tümörler kanserlidir, vücudun diğer bölgelerine ve yakındaki sağlıklı meme dokusuna yayılabilir [65]. Kanserli bir tümör meme bezlerine yerleştiğinde ve kanserli hücreler vücudun diğer bölgelerine yayılmadığında, meme kanseri “invaziv olmayan” veya “in situ” olarak kabul edilir. İnvaziv olan meme kanserleri, bitişik lenf düğümlerine veya diğer organlara yayılabilir. Bu sürece metastaz denir ve geç bir evrede yakalandığında yaşamı tehdit edebilir [64].



Şekil 2.2. Meme kanserinin gelişim süreci [66].

Meme kanseri gelişiminde genetik ve çevresel faktörler rol oynar. Yaşlanma, aile öyküsü ve genetik mutasyonlar riski artıran başlıca faktörlerdir. Bunun yanı sıra, beslenme alışkanlıkları, çevresel maruziyetler, sigara ve alkol kullanımı da hastalığın gelişiminde etkili olabilmektedir [67].

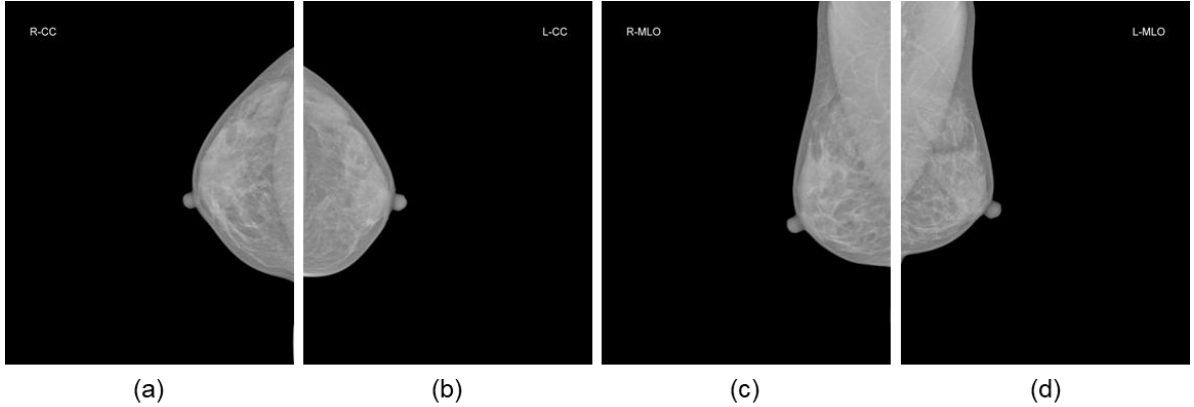
Meme kanserinin belirtileri arasında memede bir yumru, memenin boyutunda veya şeklinde bir değişiklik, meme ucundan gelen sıvı, meme ucunun veya çevresinin görünümünde değişiklik, çukurlaşma veya kızarıklık gibi cilt değişiklikleri yer alabilir [68].

## 2.4. Meme Kanseri Görüntüleme Yöntemleri

Meme kanserinin erken teşhisi ve tedavisi için görüntüleme çok önemlidir. Klinikte meme kanseri taramasına yönelik yaygın olarak kullanılan görüntüleme yöntemleri arasında mamografi, Dijital Meme Tomosentezi (Digital Breast Tomosynthesis, DBT), ultrason ve Manyetik Rezonans Görüntüleme (MRG) yer almaktadır. Bu görüntüleme yöntemleri farklı avantajlara ve dezavantajlara sahip olduğundan, erken tanı için doğru ve etkili yöntemin seçilmesi kritik öneme sahiptir.

### 2.4.1. Mamografi

Mamografi, meme kanserini erken tespit etmeye yardımcı olan düşük doz X-ışını kullanan bir görüntüleme yöntemidir [65]. Bu yöntemde, meme dokusu bir plaka ile bastırılır ve ardından düşük enerjili X-ışınları dokulara nüfuz ederek radyografik bir görüntü oluşturur. Mamogram, meme kanserinde morfolojik olarak şüpheli bulguları (kitle, kalsifikasyon, mimari bozulma ve asimetri) belirlemeye yardımcı olan iki boyutlu bir görüntüdür [69]. Mamografi görüntülemesi, CC ve MLO olmak üzere iki temel projeksiyon ile gerçekleştirilir. CC projeksiyonunda, glandüler dokuyu ve çevresindeki yağ dokusunu ortaya çıkarmak için meme yukarıdan aşağıya iki plaka arasında sıkıştırılır. MLO projeksiyonunda ise meme, 45 derecelik bir açı ile plakalar arasında çapraz olarak sıkıştırılır ve bu sayede CC projeksiyonuna kıyasla meme dokusunun daha büyük bir bölümü görüntülenir. Ayrıca, MLO projeksiyonu pektoral kasların mamogramda görünmesini sağlar [2]. Mamografide standart olarak her iki meme, CC ve MLO projeksiyonlarında görüntülenir. Sağ meme için CC (Right CC, R-CC) ve MLO (Right MLO, R-MLO), sol meme için ise CC (Left CC, L-CC) ve MLO (Left MLO, L-MLO) projeksiyonları kullanılarak dört temel görüntü elde edilir. Şekil 2.3'te bu projeksiyonlara ait örnek mamografi görüntüleri sunulmaktadır.



**Şekil 2.3.** Mamografinin dört temel projeksiyonuna ait görüntüleri: (a) R-CC görünüm, (b) L-CC görünüm, (c) R-MLO görünüm ve (d) L-MLO görünüm [70].

Ekran-film mamografisi (screen-film mammography, SFM) ve dijital mamografi (digital mammography, DM) olmak üzere iki tür mamografi yöntemi bulunmaktadır. SFM, şüpheli lezyonların erken aşamada tespiti için standart bir görüntüleme yöntemidir. Bu yöntem, öncelikli olarak yağ dokusundan oluşan memelerde şüpheli lezyonları tespit etmede yüksek bir duyarlılığa sahipken yoğun glandüler dokuya sahip memelerde bu oran önemli ölçüde azalmaktadır. Sonuç olarak, meme kanserlerinin %10-20'si görüntülenememektedir

[6], [10]. Bununla birlikte, SFM bazı dezavantajlara sahiptir. SFM ile üretilen bazı görüntüler düşük kontrastlı olabilir ve SFM görüntülemesinin daha yüksek radyasyon dozuyla tekrarlanması gerekebilir. Ayrıca, SFM’de film üretildikten sonra görüntü iyileştirme yapılamamaktadır. SFM’nin bir diğer önemli sınırlılığı da dijital olmamasıdır [6], [71]. DM ise erken evre meme kanseri taraması için önemli ve etkili olan bir diğer önemli görüntüleme yöntemidir [10]. DM’de film kullanımı yerine, doğrudan dijital görüntü elde edilmesini sağlayan bir dijital X-ışını dedektörü kullanılmaktadır [72]. Ancak, DM’nin düşük özgüllüğe sahip olması bazı sınırlamalara yol açmaktadır. Bu durum, çok sayıda gereksiz biyopsiye neden olabilir ve bu da hem hasta üzerindeki maliyet yükünü hem de stres düzeyini artırmaktadır. Ayrıca, düşük özgüllük ve yüksek maliyet gibi etkenler, hastaların gereksiz yere iyonlaştırıcı radyasyona maruz kalmasına neden olabilir. Bununla birlikte, DM’nin SFM’ye göre avantajı dijitalleştirilebilmesi ve böylece radyologların hassasiyetini artırmak için bilgisayar destekli tanı sistemlerinden yararlanılabilesidir [6], [10].

Mamografi ile yapılan iki tür inceleme bulunmaktadır: tarama mamografisi ve tanısal mamografi. Tarama mamografisi, asemptomatik bireylerde meme kanserini erken tespit etmek amacıyla uygulanmaktadır. Tanısal mamografi ise, memede kitle veya diğer anormal klinik bulgular saptanan hastalarda daha ayrıntılı değerlendirme yapmak için kullanılmaktadır. Tanısal mamografi, genellikle tarama mamografisinde şüpheli bulgular saptandığında ek görüntüleme gereksinimi doğrultusunda gerçekleştirilir ve şüpheli alanın daha ayrıntılı incelenmesini sağlar. Ayrıca, gerek duyulması halinde ek meme görüntüleme yöntemleri ya da biyopsi ile kesin tanı sürecine katkı sağlar [73].

Amerikan Radyoloji Koleji (The American College of Radiology, ACR), mamografi raporlarının terminolojisini, bulguların değerlendirilmesini ve izlenilecek eylemlerin yönünü standart hale getirmek için Meme Görüntüleme Raporlama ve Veri Sistemi (Breast Imaging Reporting and Data System, BI-RADS) ölçeğini geliştirmiştir [74]. BI-RADS, radyologlar ve klinisyenler tarafından mamogram incelemelerini raporlamak için kullanılan temel araçlardan biridir [75].

BI-RADS skorları eksik değerlendirme (kategori 0) ve tam değerlendirme (kategori 1, 2, 3, 4, 5, 6) olarak sınıflandırılmaktadır [76]. Her bir kategorinin açıklaması Tablo 2.1’de sunulmaktadır.

**Tablo 2.1.** BI-RADS kategorileri ve tanımları

| BI-RADS Kategorisi | Tanım   | Kanser olma olasılığı   | Klinik Öneri   |
|--------------------|---|---|--|
| 0                  | Eksik değerlendirme   | -   | Ek tarama  |
| 1                  | Normal  | 0   | Rutin mamografi taraması   |
| 2                  | Benign  | 0   | Rutin mamografi taraması   |
| 3                  | Muhtemelen benign   | <%2   | Kısa süreli (6 ay) takip taraması                                  |
| 4                  | Malignite şüphesi<br>4a: düşük şüpheli malignite<br>4b: orta derecede şüpheli malignite<br>4c: yüksek şüpheli malignite | 4a: $\geq\%2$ ile $<\%10$<br>4b: $\geq\%10$ ile $<\%50$<br>4c: $\geq\%50$ ile $<\%95$ | Biyopsi gerekli  |
| 5                  | Malignite şüphesi kuvvetli  | $\geq\%95$  | Biyopsi ve tedavi  |
| 6                  | Biyopsi ile kanıtlanmış malignite   | %100  | Bekleyen/devam eden tedavi tamamlanmalı, uygunsa cerrahi operasyon |

#### 2.4.2. Dijital meme tomosentezi

DBT, farklı açılardan alınan düşük dozlu X-ışınlarını kullanarak memeye ait üç boyutlu görüntülerin elde edilmesini sağlayan bir görüntüleme yöntemidir [6]. DBT’de görüntüleme sırasında meme, mamografide olduğu gibi yerleştirilip sıkıştırılır; ancak mamografiden farklı olarak X-ışını tüpü meme etrafında dairesel bir yay çizerek hareket eder [10]. Tomosentez sırasında uygulanan X-ışını dozu, standart bir mamogramda uygulanan doza benzerdir. DBT’de görüntüleme işlemi kısa sürede tamamlanır ve üç boyutlu meme görüntüleri, X-ışınlarından alınan bilgilere dayanarak bilgisayar aracılığıyla üretilir [6]. DBT’nin temel avantajı, yoğun meme dokusu nedeniyle geleneksel mamografide gizlenebilen kitlelerin ve lezyonların daha etkin şekilde tespit edilebilmesidir. Tomosentezin hassasiyeti yüksektir ve yanlış pozitif tespitlerin oranı azalır [69]. Ancak, DBT’nin hassasiyeti yüksek olmasına rağmen bazı önemli sınırlamaları bulunmaktadır. Bu sınırlamalardan biri, DBT’nin dilim düzleminde yer almayan kötü huylu mikrokalsifikasyonları tespit etmede yetersiz kalabilmesidir. Bunun yanında, mimari bozulma tipindeki meme kanseri vakalarında duyarlılık oranlarında artış sağlayabilmektedir [6]. Tarama sırasında DBT ile kanser tespitine yönelik yapılan çalışmalarda umut verici sonuçlar elde edilmiştir. Ancak DM ile karşılaştırıldığında, üç boyutlu veri setlerinin

yorumlanması daha uzun zaman almakta ve bu durum önemli bir dezavantaj olarak değerlendirilmektedir [72].

### **2.4.3. Ultrason**

Meme ultrasonu, yüksek frekanslı ses dalgaları kullanarak memenin iç yapısının gerçek zamanlı görüntülerini oluşturan bir görüntüleme yöntemidir [2]. Gerçek zamanlı ve dinamik bir değerlendirme imkânı sunarak, malign tümörlerin ince bulgularının tespit edilmesini ve hastaların semptomlarının lokalizasyonunun doğrudan değerlendirilmesini sağlar. Ultrason, kistik lezyonları katı lezyonlardan güvenilir bir şekilde ayırt edebilir; ayrıca katı lezyonların benign ya da malign karakterde olup olmadığını belirlemeye yardımcı olabilir [77]. Mamografi yoğun meme dokusundaki anormallikleri gizleme eğilimindeyken, ultrason meme yoğunluğundan etkilenmemektedir. Bu nedenle, yoğun meme dokusuna sahip hastalarda ultrason tamamlayıcı bir tarama aracı olarak kullanılabilir [54]. Ayrıca ultrason; genç, emziren ve hamile hastalarda klinik semptomların değerlendirilmesinde tercih edilen ilk görüntüleme yöntemidir [77]. Mamografi ve DBT gibi diğer görüntüleme yöntemleriyle karşılaştırıldığında ultrason; invaziv olmayan, radyasyon içermeyen, düşük maliyetli ve güvenilir bir alternatif olarak öne çıkmaktadır [6]. Ayrıca, belirsiz veya şüpheli meme lezyonları için ultrason eşliğinde biyopsi işlemleri de kolay bir şekilde gerçekleştirilebilmektedir [78]. Bununla birlikte, ultrasonun bazı sınırlamaları mevcuttur. Ultrason görüntülerinin yorumlanması oldukça zordur ve bu görüntüleri kapsamlı bir şekilde anlamak için uzman bir radyoloğa ihtiyaç duymaktadır. Bunun nedeni, ultrason görüntülerinin karmaşık yapısı ve görüntülerde sıkça rastlanan benek gürültüsünün varlığıdır. Bu sorunla başa çıkmak için meme görüntülemesinde otomatik meme ultrasonu (automated breast ultrasound, ABUS) gibi yeni teknolojiler geliştirilmiştir. ABUS, daha geniş problemler kullanarak üç boyutlu görüntüler üretmektedir [71]. Bu yöntemin başlıca avantajları arasında yüksek tanısal doğruluk ve lezyon boyutunun daha iyi belirlenebilmesi yer almaktadır [69].

### **2.4.4. Manyetik rezonans görüntüleme**

MRG, güçlü bir manyetik alan ve radyo frekanslı dalgaları kullanarak dokuların farklı açılardan ayrıntılı görüntülerinin elde edilmesine dayanır. Elde edilen çoklu görüntülerin bir araya getirilmesiyle, dokuların net ve ayrıntılı görünümü oluşturulur. Bu nedenle MRG, diğer görüntüleme yöntemlerine kıyasla meme kanseri analizi için daha yüksek görüntü

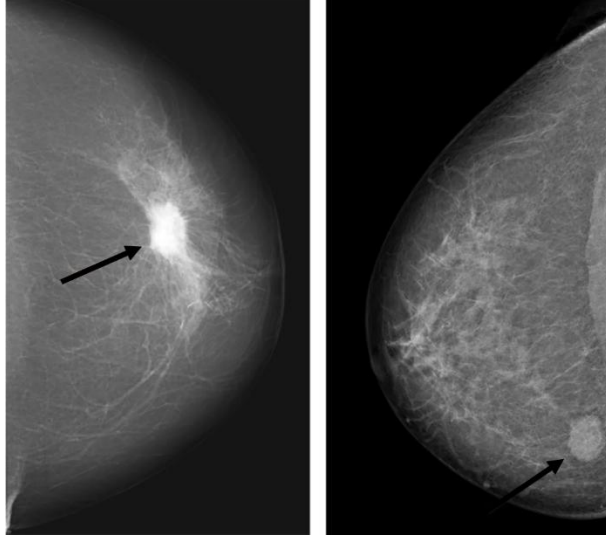
kalitesi sunmaktadır [71]. Ayrıca, iyonlaştırıcı radyasyon içermemesi nedeniyle hastalar için daha güvenli bir görüntüleme yöntemidir. MRG; yoğun meme dokusuna sahip kadınlar, aile öyküsünde meme kanseri bulunanlar ve mamografide patolojiyi maskeleyebilecek silikon implantları olan hastalar için tarama amacıyla önerilmektedir [79]. Meme kanseri tespitinde MRG'nin hassasiyeti, ultrason veya mamografiye kıyasla daha yüksektir. Bununla birlikte, MRG'nin bazı sınırlamaları da bulunmaktadır. Yüksek maliyeti, intravenöz gadolinyum kontrast maddesi enjeksiyonu gerektirmesi, sınırlı erişilebilirlik ve yüksek yanlış pozitiflik oranları gibi nedenlerle, nadir özel durumlar dışında meme kanseri taramasında rutin olarak tercih edilmemektedir [77]. Buna ek olarak, MRG görüntülerinin değerlendirme süresinin uzun olması, işlemi zaman alıcı hale getirmektedir [71].

## **2.5. Meme Kanserinin Mamografik Bulguları**

Meme kanseri, mamografik görüntülerde kitle, kalsifikasyon, mimari bozulma, asimetri veya bu bulguların kombinasyonu şeklinde ortaya çıkabilir. Mimari bozulma ve asimetri, kitlelerle benzer görünüm sergileyebileceğinden, radyologlar tarafından sıklıkla kitlelerle karıştırılabilmektedir [80]. Mamografik incelemelerde en sık rastlanan bulgular ise kitleler ve kalsifikasyonlardır [81].

### **2.5.1. Kitle**

Kitle, memede anormal bir şişlik, çıkıntı veya yumru olarak tanımlanan genel bir terimdir [81]. Şekil 2.4'te mamografik bir kitle görünümüne örnek sunulmaktadır. Kitleler, şekil ve kenar (sınır) gibi morfolojik özelliklerine göre sınıflandırılabilir. Şekil açısından kitleler; yuvarlak, lobüler, oval veya düzensiz; kenar yapısı bakımından belirsiz, sivri uçlu veya belirgin sınırlı olarak gözlemlenebilir. Benign kitleler genellikle yuvarlak veya oval şekle, pürüzsüz ya da belirgin sınırlara sahiptir. Kitlenin şekli ne kadar düzensizse, malign olma riski o kadar artar. Özellikle düzensiz veya sivri uçlu kenarlar, bir kitlenin malign olabileceğini gösteren önemli morfolojik göstergeler arasında yer almaktadır [5], [80].

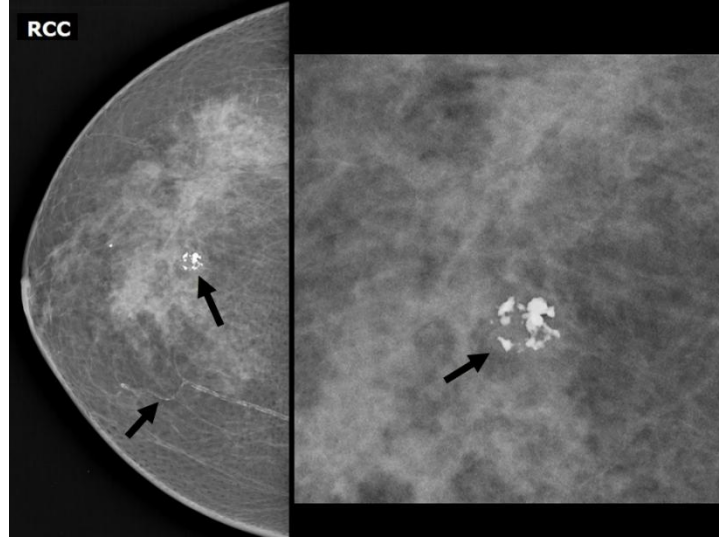


Şekil 2.4. Mamografide kitle görünümü [82].

Benign kitleler vücudun diğer bölgelerine yayılım (metastaz) göstermez; ancak diğer dokulara zarar verebileceği için çıkarılmaları gerekebilir. Buna karşılık, malign kitleler vücudun diğer bölgelerine yayılabilir ve komşu dokulara zarar verebilir [11].

### 2.5.2. Kalsifikasyon

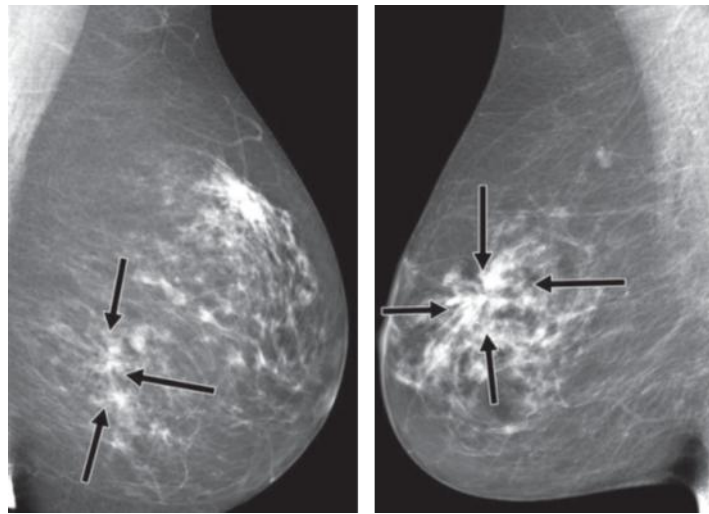
Kalsifikasyonlar, meme dokularında gelişen kalsiyum birikintileridir ve Şekil 2.5'te mamografik kalsifikasyon görünümü sunulmaktadır. Kalsifikasyonlar, makrokalsifikasyon ve mikrokalsifikasyon olmak üzere iki gruba ayrılmaktadır. Makrokalsifikasyonlar, genellikle meme bölgesine rastgele dağılmış, beyaz renkli, büyük ve kaba yapılı noktalarlardır. Pürüzsüz konturlara sahiptirler ve çoğunlukla benign karakterdedirler. Bu nedenle rutin takip dışında ek incelemeye ihtiyaç duyulmamaktadır. Mikrokalsifikasyonlar ise genellikle benign yapıda olan küçük kalsiyum birikintileridir. Ancak belirli desenler ve kümeler halinde ortaya çıktıklarında, malign hastalığın erken bir belirtisi olabilirler. Malign mikrokalsifikasyonlar çok sayıda, kümelenmiş, küçük boyutlu, morfolojik olarak değişken ve düzensiz konturlu yapılar şeklinde izlenir. Mamografik görüntülerde, bu kalsifikasyonlar küçük ve parlak noktalar şeklinde görülebilir [5], [80]. Yoğun meme dokusunda, mikrokalsifikasyonlar ile çevre dokular arasındaki düşük kontrast nedeniyle bu yapıların tespiti güçleşebilmektedir [4].



Şekil 2.5. Mamografide kalsifikasyon görünümü [83].

### 2.5.3. Mimari Bozulma

Mimari bozulma, belirgin bir kitle olmaksızın meme parankiminin normal yapısının bozulması olarak tanımlanır. Bu durum; düz ince çizgiler, speküle (ışınsal) çizgiler veya fokal retraksiyon (lokalize çekilme) gibi bulgularla karakterize edilebilir. Şekil 2.6'da mamografide gözlenen mimari bozulma görünümünü sunulmaktadır. Mimari bozulma, fiziksel muayene ile saptanamayan meme kanserinin mamografide en sık gözlenen üçüncü bulgusu olarak kabul edilmektedir. Mimari bozulmaya eşlik eden bir kitle varlığında, malignite riski önemli ölçüde artmaktadır. Ayrıca, mimari bozulma sıklıkla asimetri veya kalsifikasyonlarla birlikte de izlenebilir [80], [84].

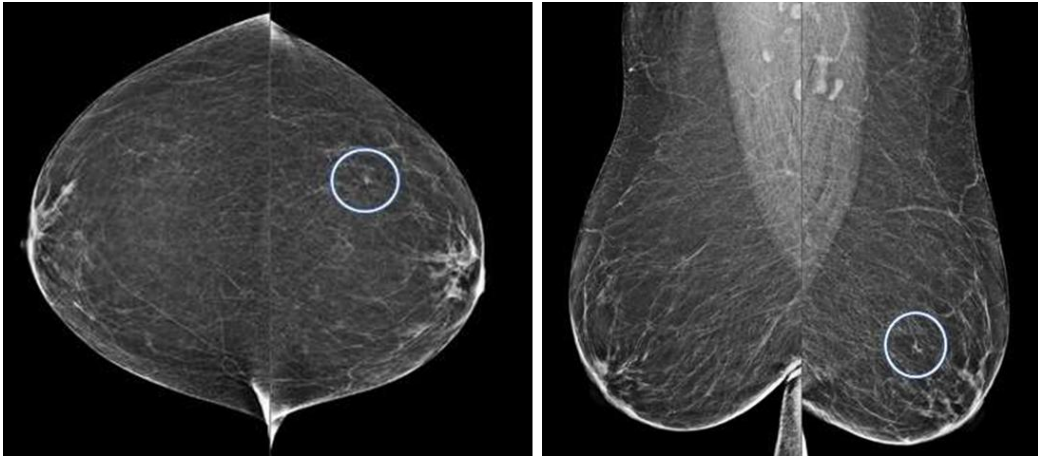


Şekil 2.6. Mamografide mimari bozulma görünümü [85].

Mimari bozulma, mamografik görüntülerde üst üste binen normal meme dokusu görünümünü taklit edebileceğinden, tespiti zor olabilir. Özellikle tarama mamografilerinde mimari bozulma, gözden kaçan veya yanlış yorumlanan meme kanseri vakalarının %12 ila %45'ini oluşturabilmektedir [84].

#### 2.5.4. Asimetri

Genellikle her iki memenin boyut ve fibroglandüler doku dağılımı açısından simetrik bir görünüme sahip olduğu kabul edilir. Asimetri, Şekil 2.7'de gösterildiği gibi karşı memedeki aynı anatomik bölgeyle kıyaslandığında, bir memede yoğunluğun arttığı alan olarak tanımlanır. Genellikle, standart mamografi projeksiyonlarından yalnızca birinde görülür ve çoğunlukla normal meme dokularının üst üste binmesi (summasyon artefaktı) sonucu oluşur [80], [86]. Çoğu durumda benign karakterde olmasına rağmen, tanısal açıdan zorluk oluşturabilen bir mamografik bulgudur. Tarama mamografilerinin yaklaşık %3'ünde izlenmektedir. Kitle, mikrokalsifikasyon veya malignite şüphesini artıran mimari bozulma gibi bulgularla birlikte görülmediği sürece, genellikle normal bir varyant olarak değerlendirilir. Ancak bu bulgularla birlikte izlenmesi durumunda, malignite ihtimaline işaret edebileceğinden ek tanısal değerlendirme gerektirir [87].



Şekil 2.7. Mamografide asimetri görünümü [88].

## 3. MATERYAL VE YÖNTEM

### 3.1. Çalışmada Kullanılan Mamografi Veri Kümeleri

Meme kanserinin tespiti ve sınıflandırılmasına yönelik derin öğrenme tabanlı modellerin başarılı bir şekilde geliştirilebilmesi için, kaliteli ve büyük ölçekli veri kümelerine ihtiyaç duymaktadır. Bu ihtiyacı karşılamak amacıyla, farklı araştırmacılar ve kuruluşlar tarafından çeşitli mamografi görüntü veri kümeleri oluşturulmuştur. Literatürde, meme kanseri tanısına yönelik birçok açık erişimli mamografi veri kümesi bulunmaktadır. Bu doğrultuda, tez kapsamında hem açık erişimli veri kümeleri hem de Başkent Üniversitesi Ankara Hastanesi'nden retrospektif olarak elde edilen mamografi verileri kullanılmıştır.

#### 3.1.1. INBreast

INBreast veri kümesi, 115 vakadan elde edilen toplam 410 görüntüden oluşmaktadır. Bu görüntüler DICOM formatında olup, farklı lezyon türlerini (kitle, kalsifikasyon, asimetrisi ve mimari bozulma) içermektedir [89]. Bu lezyon türleri arasında yalnızca kitle etiketine sahip 107 mamografi görüntüsü bulunmaktadır. Bazı görüntüler birden fazla kitle içermesi nedeniyle toplam kitle sayısı 116'ya ulaşmıştır. BI-RADS kategori değerlendirmesine göre BI-RADS  $\in \{2,3\}$  olan 40 adet kitle benign, BI-RADS  $\in \{4,5,6\}$  olan 76 adet kitle ise malign olarak etiketlenmiştir.

#### 3.1.2. Curated Breast Imaging Subset of DDSM (CBIS-DDSM)

CBIS-DDSM veri kümesi, Digital Database for Screening Mammography (DDSM) veri kümesinin düzenlenmiş ve geliştirilmiş bir versiyonudur. Güncellenmiş kitle bölütlemeleri ve bunlara karşılık gelen maske görüntüleri içermektedir. DICOM formatında sunulan CBIS-DDSM, 1566 vakaya ait kitle ve kalsifikasyonları içeren toplam 6671 mamografi görüntüsünden oluşmaktadır. Bu görüntülerden 1600'ü kitle görüntüsü olup, bu görüntüler toplam 1696 kitle lezyonu barındırmaktadır [90].

#### 3.1.3. VinDr-Mammo

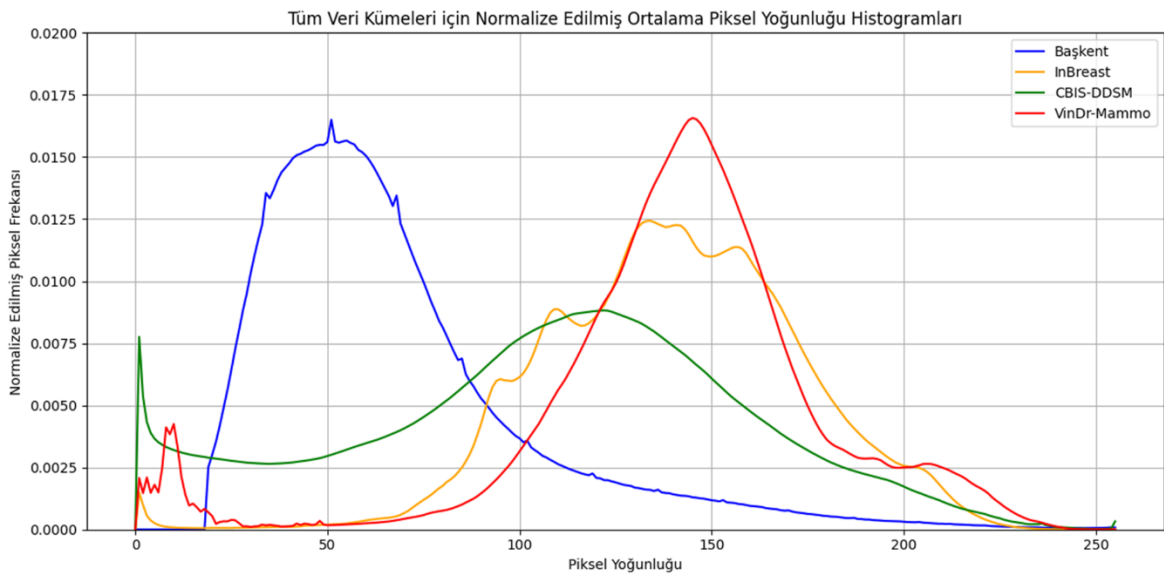
VinDr-Mammo veri kümesi, bilgisayar destekli tespit ve teşhis algoritmalarının geliştirilmesi ve değerlendirilmesi amacıyla oluşturulmuştur. Bu veri kümesi, 5000 vakadan elde edilen dört farklı görünümünden (sağ meme-CC/MLO, sol meme-CC/MLO) oluşan

toplam 20.000 DICOM formatında mamografi görüntüsünü içermektedir. Görüntüler, meme düzeyinde BI-RADS kategori (1, 2, 3, 4, 5) değerlendirmelerine ve çeşitli lezyon türlerine (kitle, kalsifikasyon, asimetri vb.) sahiptir. Kitle kategorisine ait 1029 görüntüde toplam 1123 kitle lezyonu bulunmaktadır [70]. INBreast veri kümesinde olduğu gibi BI-RADS  $\in \{2,3\}$  olan kitleler benign, BI-RADS  $\in \{4,5\}$  olan kitleler ise malign olarak kabul edilmiştir.

### 3.1.4. Başkent Üniversitesi Ankara Hastanesi mamografi verileri

Başkent Üniversitesi Ankara Hastanesi Radyoloji Bölümü'nde yürütülen klinik uygulamalardan retrospektif olarak 102 mamografi görüntüsü toplanmıştır. Bu görüntülerde, 65'i benign ve 53'ü malign olmak üzere toplam 118 kitle bulunmaktadır. Kitlelerin konumu ve türü, uzman bir radyolog tarafından işaretlenmiş ve etiketlenmiştir. Verilerin toplanması sürecinde, DICOM formatındaki orijinal mamogramlar .png formatına dönüştürülerek anonimleştirilmiştir. Bu çalışma, Başkent Üniversitesi Tıp ve Sağlık Bilimleri Araştırma Kurulu ve Etik Kurulu tarafından onaylanmış (Proje no: KA 24/376) ve Başkent Üniversitesi Araştırma Fonunca desteklenmiştir.

Farklı veri kümelerindeki mamografi görüntülerinin kontrast düzeyleri, parlaklık yapıları ve piksel yoğunluğu dağılımları arasında belirgin farklar gözlemlenmiştir. Bu farklılıkları daha somut bir şekilde ortaya koymak amacıyla, her bir veri kümesindeki tüm görüntülerden elde edilen ortalama normalize edilmiş piksel yoğunluğu histogramları Şekil 3.1'de karşılaştırmalı olarak sunulmuştur. Görselde sunulan histogramlar, veri kümelerinin yoğunluk dağılımı açısından taşıdığı yapısal farklılıkları ortaya koymaktadır.



Şekil 3.1. Veri kümelerine ait ortalama normalize piksel yoğunluğu dağılımları.

### 3.2. Nesne Tespit Yöntemleri

Bilgisayarlı görü alanında nesne tespiti oldukça önemli ve zorlu bir konudur. Temel amacı, bir görüntüdeki nesnelerin varlığını ve konumunu doğru bir şekilde tespit etmek ve bunları doğru etiket ile sınıflandırmaktır [91]. Nesne tespiti; otomatik araç sürüşü, video ve görüntü algılama, uzaktan algılama, yüz algılama, metin algılama, yaya algılama, logo algılama, tıbbi görüntü analizi, endüstriyel inceleme ve diğer birçok alanda önemli gelişmelere olanak sağlamıştır [92], [93].

Nesne tespiti, geleneksel nesne tespit algoritmaları (1998-2014) ve derin öğrenmeye dayalı nesne tespit algoritmaları (2014-günümüz) olmak üzere iki döneme ayrılabilir [94]. Geleneksel nesne tespit yöntemleri genellikle özneliklerin elle çıkarıldığı ve karmaşık temsillere dayanan algoritmalarıdır. Bu alandaki önemli ilerlemelerden biri, kayar pencere yaklaşımını kullanarak yüz algılama problemini çözen Viola-Jones dedektörü olmuştur. Viola-Jones algoritması, görüntülerde nesnelere tespit etmek için Haar benzeri öznelikleri ve aşamalı bir sınıflandırıcı dizisini kullanır. Haar benzeri öznelikler, nesne şekillerini yaklaşık olarak tanımlamak amacıyla kullanılan basit matematiksel işlevlerdir [95], [96]. Bir diğer önemli yaklaşım ise, Yönlendirilmiş Gradyanların Histogramı (Histogram of Oriented Gradients, HOG) algoritmasıdır. HOG, ölçek, çeviri ve aydınlatma değişimlerine karşı dayanıklı olacak şekilde tasarlanmış olup, özellikle yaya tespiti gibi uygulamalarda başarılı olmuştur [95], [96], [97]. Geleneksel yöntemler arasında öne çıkan bir diğer model, Deforme Edilebilir Parça Tabanlı Model (Deformable Part-Based Model, DPM), başlangıçta HOG tabanlı dedektörlerin bir uzantısı olarak geliştirilmiş ve zamanla iyileştirilmiştir. DPM, nesnelere parçalarına ayırarak modelleme yapar ve nesnenin farklı parçalarının tespiti ile genel nesne tanıma gerçekleştirir. Zayıf gözetimli öğrenme ile eğitilen bu model, geleneksel nesne tespiti yaklaşımlarının en etkili örneklerinden biri olmuştur [95], [96], [97].

Geleneksel nesne tespit yöntemleri; aday bölgelerin oluşturulması, öznelik çıkarma ve sınıflandırma olmak üzere üç aşamadan oluşmaktadır:

- **Aday bölgelerin oluşturulması:** İlk adımda, görüntüde nesne olabilecek bölgeleri belirlemek amacıyla kayar pencere yöntemi kullanılır. Görüntü, farklı konum ve ölçeklerde taranarak binlerce aday bölge oluşturulur. Ancak, tüm görüntüyü bu şekilde incelemek, çok sayıda aday pencerenin oluşmasına ve yüksek hesaplama maliyetine yol açar. Bu durum, özellikle eğitim sürecinde tekrar eden işlemler nedeniyle verimliliği azaltabilir [97], [98].
- **Öznelik çıkarma:** Belirlenen aday bölgelerdeki nesnelere tanımlayabilmek için, bu bölgelerden anlamlı özneliklerin çıkarılması gerekir. HOG, Haar ve SIFT gibi yöntemler

nesne tanımaya yönelik öznitelikleri çıkarmak için kullanılmaktadır. Bu aşama, algoritmanın başarısını büyük ölçüde etkilediği için doğru özniteliklerin seçilmesi kritik öneme sahiptir [97], [98].

- **Sınıflandırma:** Son adımda, belirlenen bölgelerin nesne içerip içermediği sınıflandırma algoritmaları ile belirlenir. Bu amaçla Destek Vektör Makinesi (DVM), Adaboost, DPM ve K-En Yakın Komşu (KNN) gibi yöntemler kullanılmaktadır [97], [98].

Geleneksel nesne tespit algoritmaları; kayan pencereye yönteminin yeterince amaca yönelik olmaması, manuel olarak seçilen özniteliklerin temsil yeteneğinin zayıf olması, zaman karmaşıklığı ve yüksek hesaplama maliyetleri gibi çeşitli sınırlamalara sahiptir. Bu sınırlamaların üstesinden gelmek amacıyla son yıllarda derin öğrenme tabanlı yöntemlere olan ilgi artmıştır [98].

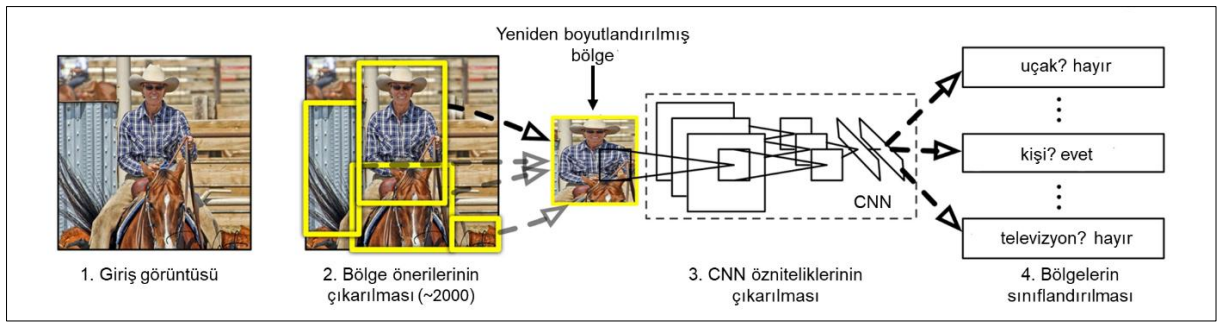
Derin öğrenmedeki son gelişmelerle birlikte, CNN nesne tespit algoritmalarında geleneksel yöntemlerin manuel öznitelik çıkarma sürecinin yerini almıştır. Yüksek öznitelik temsil gücüne sahip olmaları sayesinde, CNN'ler nesne tespit mimarilerinde yaygın olarak kullanılmaktadır [98]. Bu yetenek, hem tespit doğruluğunu artırmakta hem de modellerin farklı veri kümeleri ve değişen koşullar arasında daha güçlü genelleme yapabilmesini sağlamaktadır [95].

Derin öğrenme tabanlı nesne tespit yaklaşımları, iki aşamalı nesne dedektörü ve tek aşamalı nesne dedektörü olmak üzere iki kategoriye ayrılmaktadır.

### 3.2.1. İki aşamalı nesne tespit yöntemleri

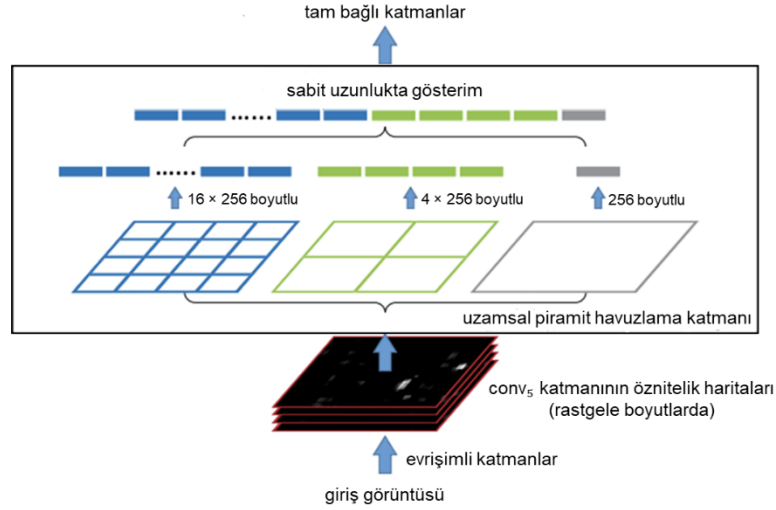
İki aşamalı nesne detektörleri, nesne lokalizasyonu ve sınıflandırma görevlerini birbirini izleyen iki aşamada gerçekleştiren yöntemlerdir. İlk aşamada, görüntüde nesne içerebilecek bölgeler belirlenir. Bu işlem, bazı mimarilerde Seçici Arama (Selective Search) ile, daha gelişmiş modellerde ise Bölge Öneri Ağı (Region Proposal Network, RPN) aracılığıyla gerçekleştirilir. İkinci aşamada ise, bu önerilen bölgeler üzerinden sınıflandırma ve sınırlayıcı kutu tahmini yapılır [99]. Bu gruptaki başlıca algoritmalar; Bölge-tabanlı CNN (Region-based CNN, R-CNN), Uzamsal Piramit Havuzlama Ağı (Spatial Pyramid Pooling Network, SPP-Net), Hızlı R-CNN (Fast R-CNN), Daha Hızlı RCNN (Faster R-CNN) ve Maske RCNN (Mask RCNN)'dir [93], [99]. Her ne kadar iki aşamalı dedektörler yüksek tespit doğruluğu ile öne çıksa da, düşük tespit hızları nedeniyle gerçek zamanlı uygulamalarda sınırlı bir kullanım alanına sahiptir [98].

CNN'lerin nesne tespiti alanındaki ilk büyük atılımı, Şekil 3.2'de verilen R-CNN modeli ile gerçekleştirmiş [100]. Bu modelde, öncelikle Seçici Arama algoritması ile yaklaşık 2000 bölge önerisi oluşturulur. Ardından, bu bölgeler CNN öznetelik çıkarıcısına uygun boyutlara getirilerek modelden geçirilir ve öznetelik vektörleri elde edilir. Elde edilen bu öznetelikler, ilgili bölgede nesne bulunup bulunmadığını belirlemek amacıyla DVM ile sınıflandırılır. Son aşamada ise, tespit edilen her nesne için doğrusal regresyon modeli kullanılarak sınırlayıcı kutular tahmin edilir. R-CNN, nesne tespitinde önemli bir ilerleme sağlamış olsa da, her sınıf için ayrı bir DVM modeli eğitilmesi ve öznetelik çıkarım sürecinin her bölge önerisi için tekrarlanması, hesaplama maliyetini oldukça artırmaktadır [92], [98].



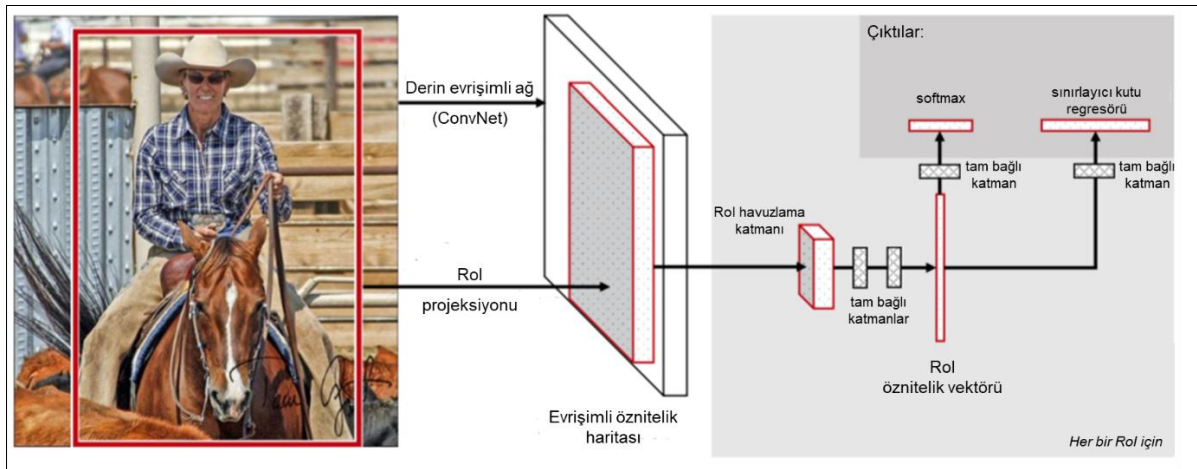
Şekil 3.2. R-CNN mimarisinin genel yapısı [100].

R-CNN, her görüntü için yaklaşık 2000 bölge önerisi ürettiğinden, bu bölgelerin her biri için CNN ile öznetelik çıkarımı hesaplama açısından oldukça maliyetlidir. Ayrıca sabit giriş boyutu gereksinimi, ağdaki tam bağlı (fully connected, FC) yapısından kaynaklanmaktadır. Bu zorluğun üstesinden gelmek için SPP-Net önerilmiştir. SPP katmanı, ilgi alanının (Region of Interest, RoI) boyutundan bağımsız olarak, yeniden ölçeklendirme gerektirmeden sabit uzunlukta öznetelik vektörleri üretmek için ağın son evrişimli katmanına entegre edilmiştir [97], [98]. SPP-Net, R-CNN'e kıyasla önemli ölçüde daha hızlı çalışsa da, benzer mimariye sahip olması nedeniyle çok aşamalı eğitim gereksinimi, yüksek hesaplama yükü ve uzun eğitim süresi gibi dezavantajları da korumaktadır [96]. Şekil 3.3'te SPP katmanı içeren CNN yapısı gösterilmektedir. Burada 256, conv<sub>5</sub> katmanındaki filtre sayısını ifade etmektedir ve conv<sub>5</sub>, ağdaki son evrişimli katmandır [101].



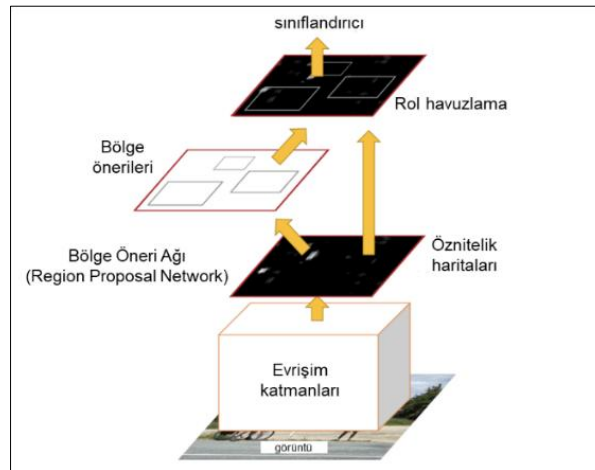
Şekil 3.3. SPP katmanı içeren CNN yapısı [101].

SPP-Net ve R-CNN modellerinin sınırlamalarını gidermek amacıyla, Şekil 3.4'te gösterilen Fast R-CNN modeli geliştirilmiştir [102]. Fast R-CNN, önceki yöntemlerden farklı olarak, softmax sınıflandırıcı ve sınıfa özgü sınırlayıcı kutu regresyonunu eş zamanlı olarak öğrenerek dedektörün uçtan uca eğitilmesine olanak tanır. Bu modelde, CNN yalnızca bir kez çalıştırılarak tüm görüntü için öznelik haritası oluşturulur. Ardından, Seçici Arama algoritması ile bölge önerileri üretilir ve her öneri için sabit uzunlukta öznelik vektörleri elde etmek amacıyla RoI havuzlama (RoI Pooling) katmanı uygulanır. Çıkarılan bu öznelikler, sınıflandırma ve yerleştirme işlemleri için softmax ve doğrusal regresyon katmanlarına iletilir. Fast R-CNN, hem hesaplama süresini önemli ölçüde azaltmış hem de tespit doğruluğunu artırmıştır. Ancak, bölge önerisi için hala zaman alıcı olan seçici arama yöntemine bağımlıdır [92], [98].



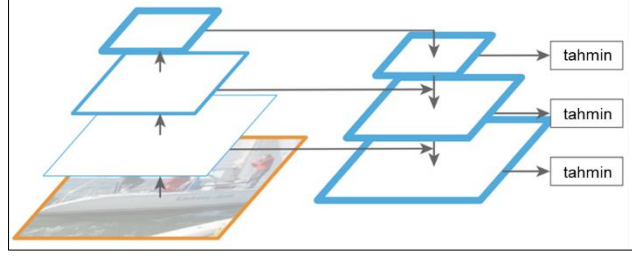
Şekil 3.4. Fast R-CNN mimarisi [102].

R-CNN, SPP-Net ve Fast R-CNN modelleri, bölge önerilerini belirlemek için seçici arama algoritmasını kullanmaktadır. Ancak bu yöntem, yavaş ve zaman alıcı yapısıyla tüm sürecin darboğazı hâline gelmiş ve ağın genel performansını olumsuz yönde etkilemiştir. Bu sorunu çözmek için, bölge önerilerini otomatik olarak üretebilen ve seçici arama ihtiyacını ortadan kaldıran Şekil 3.5’te sunulan Faster R-CNN modeli geliştirilmiştir [103]. Faster R-CNN’de geleneksel bölge önerisi algoritması, RPN adı verilen bir CNN ile değiştirilmiştir. Bu modelde, giriş görüntüsü CNN’e verilerek öznitelik haritaları elde edilir. Ardından RPN, bu öznitelik haritaları üzerinde çalışarak nesne önerilerini (RoI) ve nesnellik puanlarını üretir. RoI’ler çıkarıldıktan sonra, tüm önerileri sabit bir boyuta getirmek için RoI havuzlama katmanı uygulanır. Elde edilen bu öznitelikler, nesnelerin sınıflandırılması ve sınırlayıcı kutuların belirlenmesi amacıyla, softmax ve regresyon katmanlarıyla sonlanan tam bağlı katmanlara iletilir [96], [98].



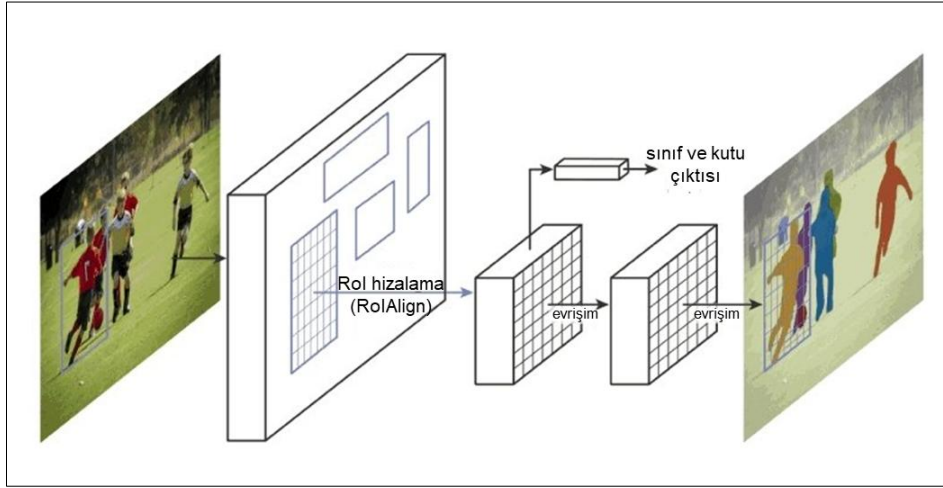
Şekil 3.5. Faster R-CNN mimarisi [103].

Öznitelik Piramit Ağı (Feature Pyramid Network, FPN), derin sinir ağlarının çok ölçekli yapısından yararlanarak düşük maliyetli ve etkili öznitelik piramitleri oluşturmak amacıyla geliştirilmiş, nesne dedektörleriyle birlikte kullanılan bir öznitelik çıkarıcı yapıdır. Şekil 3.6’da gösterildiği üzere, bu yapı, yanal bağlantılarla entegre edilmiş aşağıdan yukarıya (bottom-up) ve yukarıdan aşağıya (top-down) olmak üzere iki yönlü bir yol içerir. Aşağıdan yukarıya yol, giriş görüntüsünden farklı çözünürlüklerde öznitelik haritaları üretirken; yukarıdan aşağıya yol, üst düzey öznitelikleri geri örnekleyerek alt seviyelere taşır. Bu süreçte yanal bağlantılarla yukarıdan gelen öznitelikler, karşılık gelen aşağıdan gelenlerle birleştirilir. Böylece, her çözünürlük seviyesinde hem yerel hem de küresel bağlam içeren güçlü öznitelikler elde edilir. Bu mimari, farklı ölçeklerdeki nesnelerin daha etkili biçimde algılanmasını sağlar [98], [104].



Şekil 3.6. FPN yapısı [104].

Mask R-CNN, nesne sınıfı ve sınırlayıcı kutuya ek olarak nesne maskesini tahmin eden bir dal ekleyerek Faster R-CNN modelini genişletmiştir. Şekil 3.7’de Mask R-CNN çerçevesi sunulmaktadır. Modelin temel amacı, nesnelere piksel düzeyinde segmentasyon ile ayırt etmektir. Her pikseli analiz ederek, bir nesnenin parçası olup olmadığını tahmin eder. Mask R-CNN, Faster R-CNN mimarisini temel alır ve her iki model de aynı RPN’yi kullanır. Ancak Mask R-CNN, her nesne önerisi için sınıf etiketi, sınırlayıcı kutu ofseti ve nesne segmentasyon maskesi olmak üzere üç ayrı çıktı üretir. Ayrıca, daha hassas hizalama sağlamak amacıyla RoI havuzlama yerine RoI hizalama (RoIAlign) katmanı kullanılır. RoIAlign, örnekleme sırasında kesirli koordinatları yuvarlamak yerine doğrusal enterpolasyon uygulayarak maskelerin konumsal doğruluğunu artırır [96], [98].



Şekil 3.7. Mask R-CNN [105].

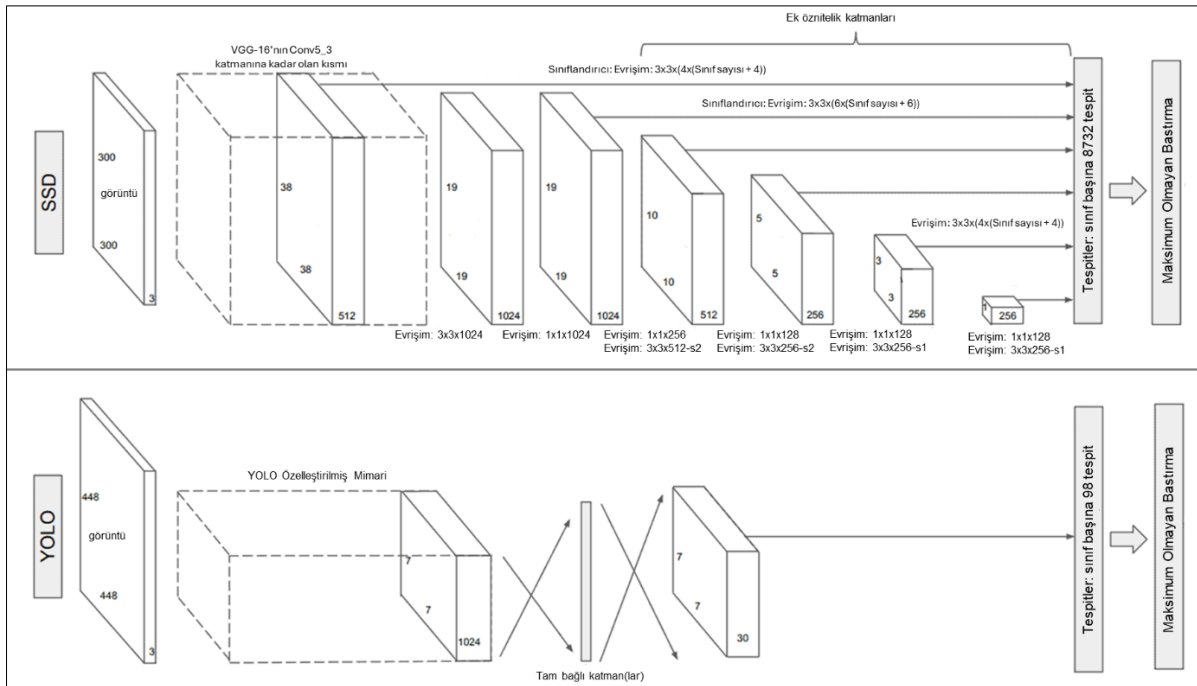
### 3.2.2. Tek aşamalı nesne tespit yöntemleri

Daha basit mimariye olan sahip tek aşamalı nesne dedektörleri, bölge önerisi adımını ortadan kaldırarak nesnelere için sınırlayıcı kutuları ve sınıf olasılıklarını tek aşamada üretir. Tek aşamalı nesne dedektörleri, Tek Atış Çoklu Kutu Dedektörü (Single Shot MultiBox Detector, SSD), RetinaNet ve YOLO algoritmalarını içermektedir [91], [99].

YOLO, kayan pencerelerle tüm görüntüyü tarayan ya da görevi iki aşamada gerçekleştiren mimarilerin aksine, tespit görevini tek bir ağ geçişiyle gerçekleştirebilen bir yaklaşımdır. YOLO, her bir nesne için sınırlayıcı kutu koordinatlarını ve ilgili sınıf olasılıklarını doğrudan tahmin eder [98], [106].

SSD, gerçek zamanlı hızını koruyarak Faster R-CNN gibi iki aşamalı dedektörlerin doğruluğuna yaklaşan ilk tek aşamalı modeldir. VGG16 omurgası üzerine inşa edilen SSD, ağdan yalnızca bir kez geçirilen görüntü üzerinde hem sınıf etiketlerini hem de sınırlayıcı kutuları öznetelik haritalarından doğrudan tahmin eder. Model, öznetelik piramidinin farklı katmanlarına küçük evrişim filtreleri uygulayarak farklı ölçeklerdeki nesnelere algılayabilir. SSD'lerin temel avantajı, tüm görüntü üzerinde aynı anda çok sayıda nesne tespiti yapabilmesidir ve bu sayede önceki yöntemlere kıyasla hesaplama maliyetini önemli ölçüde azaltır. Farklı ölçek ve en-boy oranlarına sahip sabit varsayılan kutular kullanarak, tek bir ağ içinde farklı boyutlardaki nesnelere etkili şekilde algılar ve daha hızlı bir tespit süreci sağlar [95], [107].

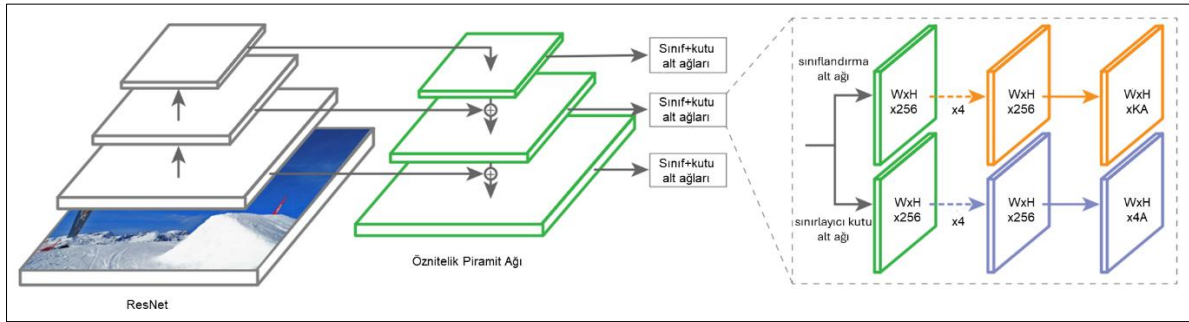
SSD ve YOLO mimarilerinin karşılaştırılması Şekil 3.8'de gösterilmektedir.



Şekil 3.8. SSD ve YOLO mimarilerinin karşılaştırılması [104].

RetinaNet, eğitim sırasında sınıf dengesizliği sorununu ele almak amacıyla odaksal kayıp (focal loss) işlevi kullanan tek aşamalı bir nesne tespit modelidir. Tek ve birleşik bir yapıya sahip olan RetinaNet, bir omurga ağı ve iki göreve özgü alt ağdan oluşur. Omurga olarak FPN ile desteklenmiş ResNet mimarisi kullanılır ve bu geniş görüntüsü üzerinde

evrişim işlemleri uygulayarak çok ölçekli öznitelik haritaları üretir. İlk alt ağ, bu öznitelik haritaları üzerinde nesne sınıflandırması görevini yerine getirir. İkinci alt ağ ise sınırlayıcı kutu koordinatlarının regresyonunu gerçekleştirir. RetinaNet, kolay eğitilebilir yapısı ve doğruluk-hız dengesi ile birçok iki aşamalı dedektörlerden daha iyi performans sergilemiştir [107], [108]. Şekil 3.9’da RetinaNet mimarisi gösterilmektedir, burada “K” sınıf sayısını, “A” her konum için kullanılan çapa (anchor) kutu sayısını, “W” genişliği ve “H” ise yüksekliği ifade etmektedir [108].



Şekil 3.9. RetinaNet mimarisi [108].

### 3.3. Nesne Tespit Alanında Kullanılan Veri Kümeleri

Nesne tespit araştırmalarında eğitim verileri, model başarımını doğrudan etkileyen temel bileşenlerden biridir. Nesne tespit modellerinin eğitimi ve test değerlendirilmesi bazı büyük veri kümelerinde gerçekleştirilmelidir. Kullanılan farklı senaryolar ve tespit edilen nesne çeşitliliği nedeniyle birçok veri kümesi türü vardır. Nesne tespit görevlerinde en yaygın olarak kullanılan popüler veri setleri aşağıda özetlenmiştir [98]:

#### 3.3.1. PASCAL VOC

PASCAL Görsel Nesne Sınıflandırması (PASCAL Visual Object Classification, PASCAL VOC) nesne tespiti için yaygın olarak kullanılan bir veri kümesidir [109]. İlk olarak 2005 yılında yalnızca dört sınıfla başlatılan bu veri kümesi, zamanla yapılan katkılarla 2007 itibarıyla 20 sınıfa ulaşmıştır. Bu sınıflar; araçlar, insanlar, ev eşyaları ve hayvanlar olmak üzere dört ana kategori altında gruplanmaktadır. PASCAL VOC 2007 ve 2012, nesnelere ve sınırlayıcı kutular içeren yaklaşık 10.000 eğitim ve doğrulama görüntüsü içermektedir [97], [98].

### **3.3.2. COCO**

Microsoft Common Objects in COntext (COCO) veri kümesi, Microsoft tarafından geliştirilmiş en zorlu nesne tespiti veri kümelerinden biridir [110]. COCO, 91 farklı nesne sınıfı içermekte olup bunların 20'si PASCAL VOC ile ortaktır. Veri kümesi, 328.000 görüntü ve 2.500.000'den fazla etiketli nesne içermektedir [92]. COCO, özellikle küçük nesne yoğunluğu ve ölçek değişkenliği açısından daha karmaşık yapısı sayesinde PASCAL VOC'e kıyasla daha zorlayıcı bir veri kümesi olarak değerlendirilmektedir [98].

### **3.3.3. ImageNet**

ImageNet, algoritmaların performansını değerlendirmek amacıyla oluşturulmuş, büyük ve çeşitli bir görüntü veri kümesidir. Bu kümeden türetilen ImageNet Büyük Ölçekli Görsel Tanıma Mücadelesi (ImageNet Large Scale Visual Recognition Challenge, ILSVRC), özellikle sınıflandırma ve nesne tespiti görevlerinde sıkça kullanılan bir veri kümesidir. En yaygın kullanılan alt kümeleri ILSVRC 2014 ve ILSVRC 2017 olup, her biri 200 kategori içeren yaklaşık 500.000 etiketli görüntü barındırmaktadır. Tüm görüntü anotasyonları, PASCAL VOC formatındaki XML dosyalarıyla sağlanmaktadır [94], [98].

### **3.3.4. Open Images**

Open Images, Google tarafından geliştirilen ve nesne sınırlayıcı kutuları, görüntü düzeyinde etiketler ve segmentasyon maskeleri ile açıklanmış, genel kullanıma açık en büyük veri kümelerinden biridir. Open Images v5, 600 farklı nesne kategorisi için 16 milyon sınırlayıcı kutu anotasyonuna ve toplamda 1,9 milyon görüntüye sahiptir. Veri kümesindeki görüntüler, doğaları gereği heterojen olup karmaşık sahnelerde çok sayıda nesne içerebilmektedir. Ortalama olarak her bir görüntüde 8,3 farklı nesne kategorisi bulunmaktadır [96], [98].

## **3.4. You Only Look Once (YOLO)**

YOLO, 2016 yılında Redmon et al. tarafından geliştirilen ve nesne tespitini bir regresyon problemi olarak ele alan uçtan uca çalışan, tek aşamalı bir nesne dedektörüdür [111]. Algoritma, bir görüntüdeki tüm nesnelere ve bu nesnelere konumlarını sınırlayıcı kutular aracılığıyla yalnızca tek bir ileri geçişte belirlediği için Sen Sadece Bir Kez Bak (You Only Look Once, YOLO) adını almıştır [99]. Bu ifade, görüntüdeki tüm nesnelere

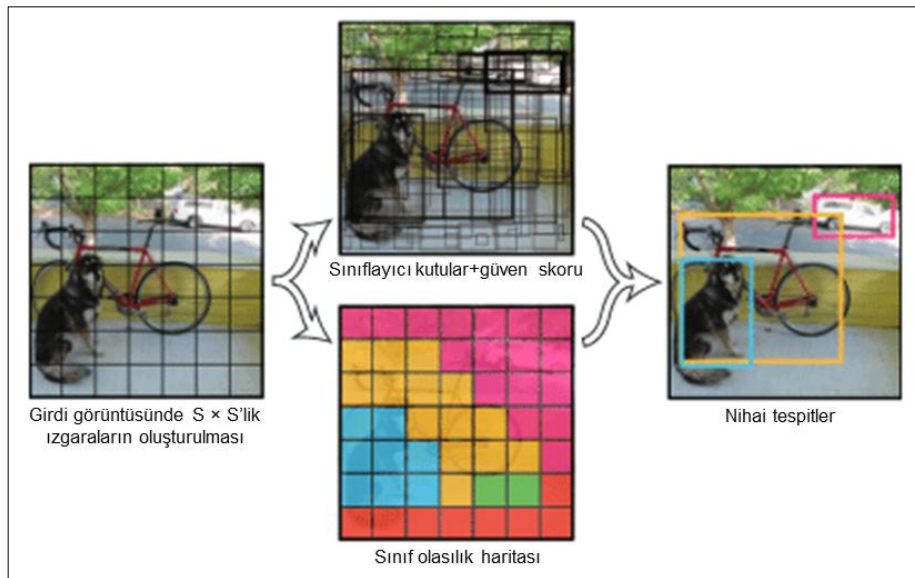
sınırlayıcı kutularını ve sınıf olasılıklarını tahmin etmek için tek bir CNN mimarisinin kullanıldığı anlamına gelmektedir [17].

YOLO algoritmasında, ilk olarak girdi görüntüsü  $S \times S$  ızgara (grid) hücrelerine bölünür ve her bir ızgara hücresi, kendi içinde bulunan nesnenin merkezine karşılık gelen sınırlayıcı kutunun tespitinden sorumludur [99]. Bu süreç Şekil 3.10'da şematik olarak gösterilmektedir.

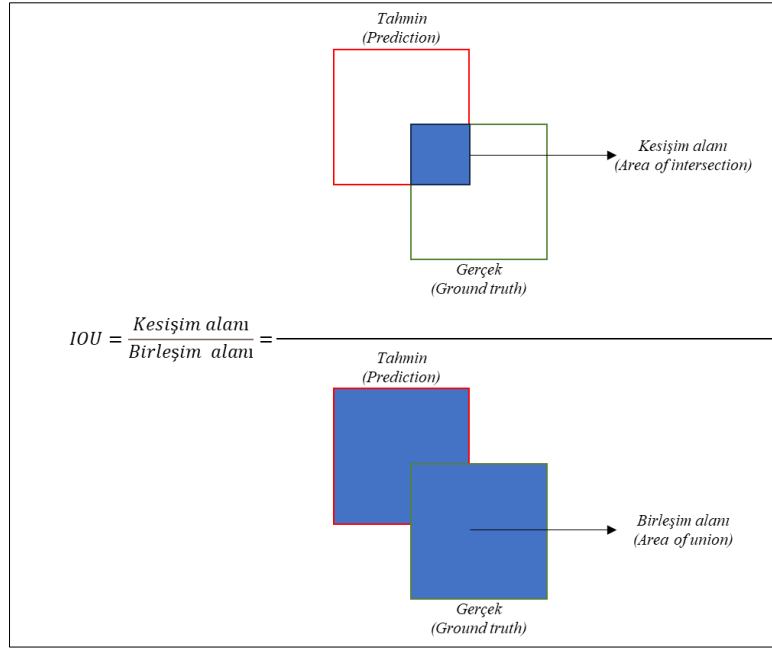
Her ızgara hücresi birden fazla sınırlayıcı kutu (bounding box, B) tahmin eder. Tahmin edilen bu kutular 5 parametre ile tanımlanır:  $p_c$ ,  $b_x$ ,  $b_y$ ,  $b_w$ ,  $b_h$ . Burada  $p_c$  güven skorunu,  $b_x$  ve  $b_y$  tahmin edilen sınırlayıcı kutunun merkez koordinatlarını,  $b_w$  ve  $b_h$  ise sınırlayıcı kutunun genişliğini ve yüksekliğini ifade etmektedir [17]. Güven skoru, tahmin edilen kutuda bir nesnenin bulunma olasılığını yansıtmakta ve Eşitlik 3.1 ile hesaplanmaktadır [17]:

$$\text{Güven skoru} = p(\text{nesne}) * IOU_{\text{tahmin}}^{\text{gerçek}} \quad (3.1)$$

Eşitlik 3.1'de  $p(\text{nesne})$ , ızgara hücre içerisinde bir nesne olma olasılığıdır. Birleşim Üzerinden Kesişme (Intersection-over-Union, IoU), tahmin edilen sınırlayıcı kutular ile nesnenin gerçek (ground truth) sınırlayıcı kutuları arasındaki örtüşmeyi ifade eden bir ölçümdür. Şekil 3.11'de gösterildiği gibi iki kutunun kesişme alanının, iki kutunun birleşim alanına oranı olarak hesaplanmaktadır. IoU, nesne algılamanın doğruluğunu ölçmek için kullanılmaktadır. IoU değerleri 0 ile 1 arasında değişmektedir; tahmin edilen sınırlayıcı kutu alanı ile gerçek sınırlayıcı kutu alanı birbiriyle mükemmel şekilde örtüşüyorsa IoU değeri 1'dir, aksi halde birbirleriyle kesişmiyorsa IoU 0'dır [98].



Şekil 3.10. YOLO algoritmasının nesne tespit sürecinin genel akışı [111].



Şekil 3.11. IoU hesaplanması.

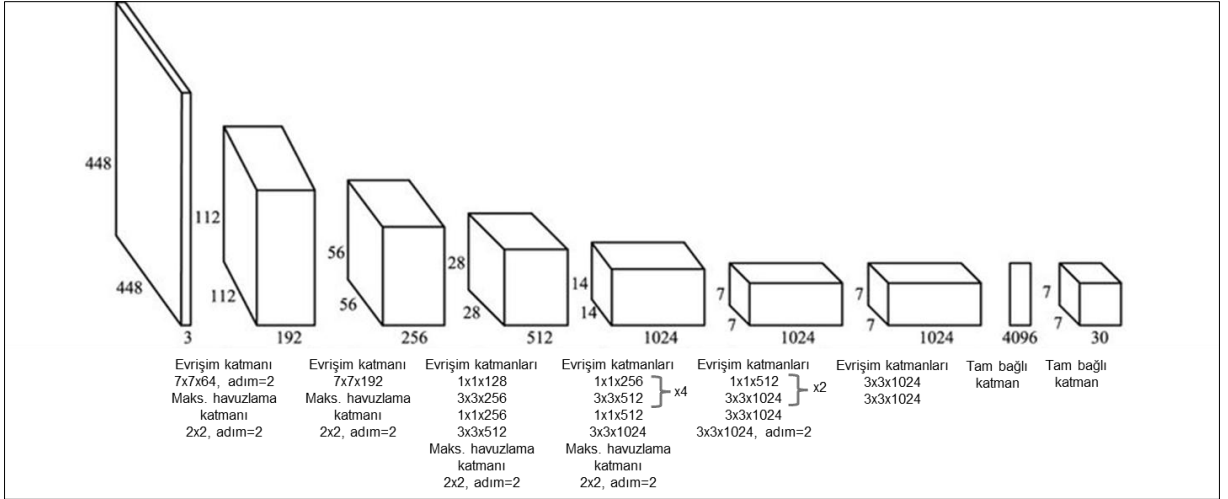
YOLO algoritması, her bir ızgara hücresi için tahmin edilen çıktı boyutunu ( $B \times 5 + c$ ) şeklinde tanımlar. Burada B, her ızgara hücresinin tahmin ettiği sınırlayıcı kutu sayısını, 5, her sınırlayıcı kutunun konum ve güven skoru parametrelerini, c ise tahmin edilecek sınıf sayısını ifade etmektedir. Dolayısıyla,  $S \times S$  boyutlarına bölünmüş bir görüntü için modelin çıktısı  $S \times S \times (B \times 5 + c)$  boyutunda bir tensör şeklinde elde edilir [99].

YOLO'nun nihai çıktısı, gereksiz sınırlayıcı kutuların azaltılması amacıyla maksimum olmayan bastırma (Non-Maximum Supression, NMS) yöntemine tabi tutulmaktadır. NMS, üst üste binen sınırlayıcı kutuların sayısını azaltmak ve genel tespit kalitesini iyileştirmek için kullanılan bir işlemdir [106]. Nesne tespit algoritmaları genellikle aynı nesne etrafında farklı güven skorlarına sahip birden fazla sınırlayıcı kutu oluşturur. İlk olarak tüm sınırlayıcı kutular güven skoru değerlerine göre büyükten küçüğe doğru sıralanır ve en yüksek güven skoruna sahip sınırlayıcı kutu belirlenir. Ardından bu yüksek güven skoruna sahip kutu ile önerilen diğer sınırlayıcı kutular arasındaki örtüşme hesaplanır. NMS'deki önemli parametre, bir sınırlayıcı kutunun gereksiz olarak kabul edilmesinden önce iki kutu arasında ne kadar örtüşmeye izin verileceğini belirleyen IoU eşiğidir. Bu eşik, nesne tespit algoritmalarının çoğunda 0,5 olarak ayarlanır. Eğer iki kutu arasındaki hesaplanan IoU değeri 0,5'ten büyükse ( $IoU > 0,5$ ) en yüksek güven skoruna sahip kutu ile karşılaştırılan diğer kutu kaldırılır. Seçilen sınırlayıcı kutudan daha düşük güven skoruna sahip sınırlayıcı kutu kalmayana kadar bu adımlar tekrarlanır. Bu sayede, gereksiz ve fazla sınırlayıcı kutular temizlenerek nihai sınırlayıcı kutu elde edilmektedir [99], [106], [112].

### 3.4.1. YOLOv1

YOLOv1, Darknet çerçevesi üzerine inşa edilmiş ilk YOLO mimarisidir. Şekil 3.12’de gösterildiği gibi 24 evrişim katmanı ile sınırlayıcı kutu koordinatlarını ve olasılıklarını tahmin eden 2 tam bağlı katmandan oluşmaktadır. Bu mimaride, ilk katmanda  $7 \times 7$  evrişim filtresi kullanılırken, sonraki katmanlarda kanal sayısını azaltmak amacıyla  $3 \times 3$  evrişim katmanlarından önce  $1 \times 1$  evrişim filtreleri tercih edilmiştir. Bu yaklaşım GoogLeNet (Inception) modelinden esinlenmiştir. Modelde son katmanda doğrusal aktivasyon fonksiyonu kullanılırken, diğer tüm katmanlarda sızdıran düzeltilmiş doğrusal birim (Leaky Rectified Linear Unit, Leaky ReLU) aktivasyonu uygulanmıştır [99], [113].

YOLOv1, temel YOLO yaklaşımıyla PASCAL VOC veri kümesi üzerinde  $S=7$ ,  $B=2$  ve  $C=20$  parametreleriyle değerlendirilmiştir. Bu parametreler, son öznitelik haritalarının  $7 \times 7$  boyutunda olmasını ve modelin çıktı boyutunun  $(7 \times 7 \times (2 \times 5 + 20))$  şekilde hesaplanmasını sağlamaktadır [106], [113].



Şekil 3.12. YOLOv1 mimarisi [111].

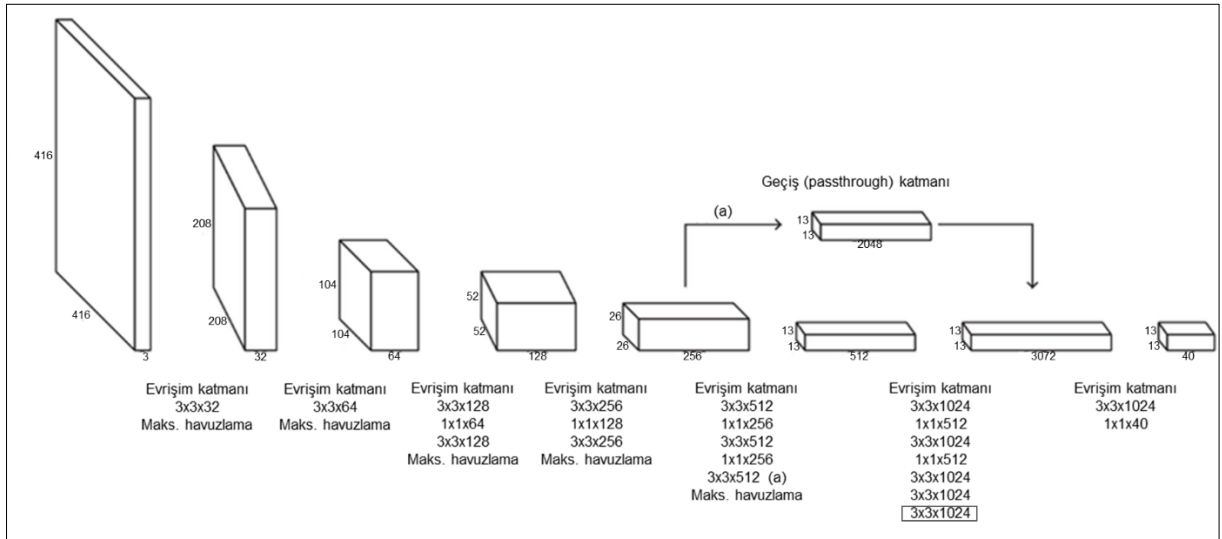
YOLOv1’in ilk 20 katmanı, bir ortalama havuzlama katmanı ve ardından bir tam bağlı katman eşliğinde ImageNet veri kümesi üzerinde  $224 \times 224$  çözünürlükte eğitilmiştir. Nesne tespitine uyarlamak amacıyla, bu yapıya dört ek evrişim katmanı ile rastgele başlatılmış ağırlıklara sahip iki tam bağlı katman eklenmiştir. Son olarak, modelin nesne algılama yeteneğini geliştirmek için,  $448 \times 448$  çözünürlükteki PASCAL VOC 2007 ve 2012 veri kümeleri ile ince ayar yapılmıştır [91], [106].

YOLOv1, büyük lokalizasyon hataları ve düşük duyarlılık gibi bazı sınırlamalara sahiptir [99]. Özellikle küçük nesnelere tespit etmede zorlanmaktadır; çünkü her ızgara hücresi yalnızca iki sınırlayıcı kutu tahmin edebilmekte ve aynı anda yalnızca tek bir sınıfa

ait nesneyi tanıyabilmektedir. Bu durum, birden fazla küçük nesnenin aynı ızgara hücresi içine düştüğü senaryolarda performans kaybına yol açmaktadır. Ayrıca, model eğitim sırasında karşılaşmadığı en-boy oranlarına sahip nesnelere tahmin etme konusunda da yetersiz kalabilmektedir [92], [106].

### 3.4.2. YOLOv2

YOLOv2, 2017 yılında YOLOv1 modelinin hız ve doğruluk açısından geliştirilmiş bir versiyonu olarak sunulmuştur. Model daha sonra, 9000'den fazla nesne kategorisini tespit edebilme yeteneğini vurgulamak amacıyla YOLO9000 adıyla yeniden adlandırılmıştır. YOLOv2, VGGNet'in daha sade ve hızlı bir versiyonu olan Darknet-19 adlı bir CNN omurgasını temel almaktadır. Darknet-19 yapısı, öznetelik gösterimini sıkıştırmak amacıyla, çoğunlukla  $3 \times 3$  evrişim ve havuzlama katmanları ile birlikte aralara serpiştirilmiş birkaç  $1 \times 1$  evrişim katmanı içermektedir. Ayrıca model eğitimi sürecinde, rastgele kırpma ve rotasyon gibi çeşitli veri artırma teknikleri uygulanmaktadır [99], [113], [114]. Şekil 3.13'te YOLOv2 mimarisi gösterilmektedir.



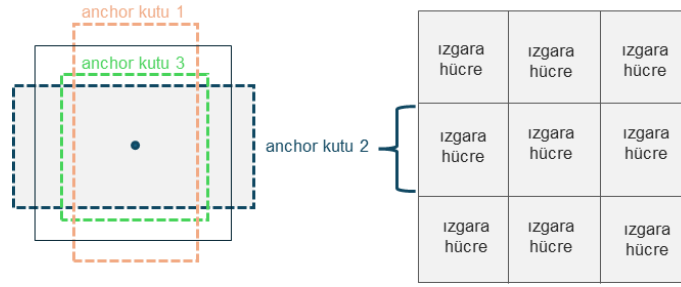
Şekil 3.13. YOLOv2 mimarisi [115].

YOLOv2'deki önemli iyileştirmeler kısaca şu şekilde tanımlanabilir:

Yüksek çözünürlüklü sınıflandırıcı: YOLOv1, eğitim aşamasında  $224 \times 224$  boyutunda giriş görüntüleri kullanmış, ancak tespit aşamasında bu çözünürlük  $448 \times 448$ 'e yükseltilmiştir. YOLOv2'de ise sınıflandırma ağı, ImageNet üzerinde doğrudan  $448 \times 448$  çözünürlükte 10 epok boyunca eğitilmiş ve böylece yüksek çözünürlüklü girdilerde ağıın performansı iyileştirilmiştir [99], [106].

Toplu normalizasyon (Batch normalization, BN): Bir düzenleyici görevi gören toplu normalizasyon katmanı, aşırı uyumu önlemek ve modelin yakınsama hızını artırmak amacıyla her evrişim katmanından sonra uygulanmaktadır [116].

Çapa (anchor) kutuları: Görüntüdeki bir nesne öyle bir sınırlayıcı kutu tarafından algılanmalıdır ki nesnenin merkezi bu sınırlayıcı kutunun içinde bulunmalıdır. Ancak, birden fazla nesnenin merkezinin aynı sınırlayıcı kutuda bulunma olasılığı olabilir. Bu nedenle YOLOv2’de, tek bir ızgara hücresine karşılık gelen sınırlayıcı kutuları temsil etmek amacıyla “anchor kutuları” terimi kullanılmıştır. Bu kutular, veri kümesindeki tüm nesnelere temsil edebilecek farklı genişlik-yükseklik kombinasyonlarını içeren, önceden tanımlanmış kutulardır. Bu kutulara manuel olarak karar vermek yerine, bu kombinasyonları bulmak için k-ortalama kümeleme algoritması kullanılmaktadır [99], [117]. Şekil 3.14’te gösterildiği gibi YOLOv2 her ızgara hücresi için birden fazla anchor kutusu tanımlar.



Şekil 3.14. Anchor kutuları [106].

Maliyet fonksiyonu: YOLOv2’nin maliyeti üç bileşenden oluşur. Yerelleştirme kaybı, sınırlayıcı kutunun koordinatlarını belirlemek için kullanılır. Güven kaybı, bir görüntüdeki bir nesnenin var olup olmadığını temsil eder ve sınıflandırma kaybı ise her kategoriye doğru tahmin etmek için kullanılır [113].

Çok-ölçekli tespit: YOLOv2, esas olarak giriş görüntüsünü farklı çözünürlüklerde örnekleme ve modeli bu çözünürlüklere göre eğitmeyi içerir. Daha sonra, elde edilen bireysel sonuçlar nihai tahminleri üretmek için birleştirilir [113]. Model 320×320 ile 608×608 arasında değişen boyutlardaki giriş görüntülerinde çalışabilir. Her 10 epokta bir, 32’nin katı olan rastgele bir yeni boyut seçilir. Bu yaklaşım, modelin farklı boyutlardaki görüntüler üzerinde tahmin yapma esnekliğini artırır [99].

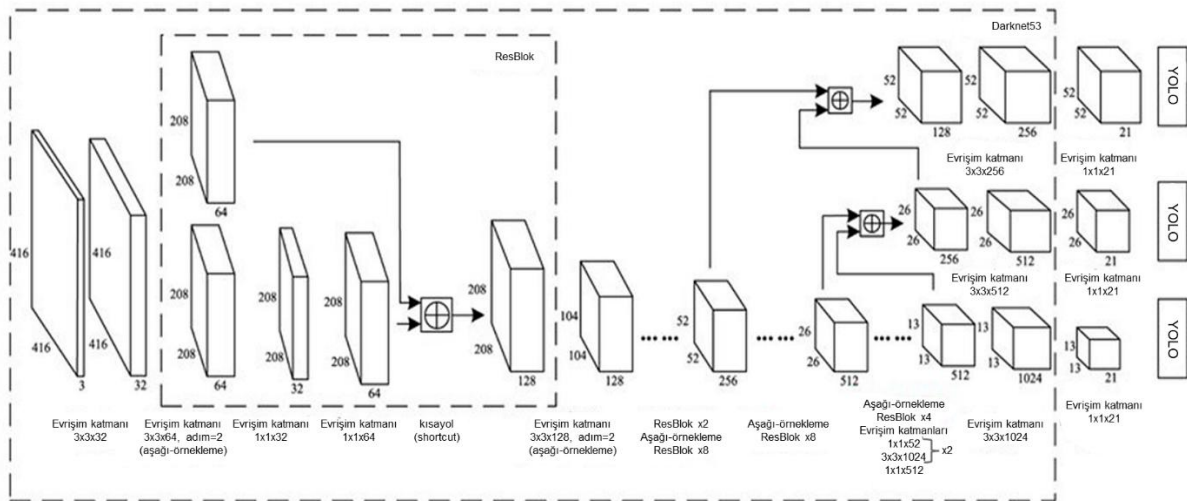
YOLOv2 ile birlikte modelin tespit doğruluğu az da olsa iyileştirilmiş ve hesaplama maliyeti azaltılmıştır. Bu versiyonda, ince taneli öznetelikleri (fine-grained features) elde edebilmek için geçiş (passthrough) katmanı eklenmiş olsa da öznetelik haritasının çözünürlüğünün aşağıya örneklenmesi (down sampling), bu detaylı özneteliklerin

kaybolmasına neden olmuştur. Bu durum, küçük boyutlu nesnelerin tespitinde zorluklara yol açmaktadır [116].

### 3.4.3. YOLOv3

YOLOv3, Joseph Redmon ve Ali Farhadi tarafından 2018 yılında tanıtılmıştır [118]. YOLOv1'deki büyük lokalizasyon hataları ve YOLOv2'nin küçük boyutlu nesneleri tespit etmedeki yetersizliği nedeniyle, bu versiyon ilk iki sürümün dezavantajlarını gidermeyi amaçlamaktadır [91], [99]. YOLOv3, COCO veri seti ile bu dezavantajların üstesinden gelerek nesneleri tespit etmek için kolay ve etkili bir çözüm sunmaktadır. Bununla birlikte, model daha küçük nesneler için daha iyi performans gösterirken, orta ve büyük boyutlu nesneler için doğru sonuçlar üretmekte zorlanmaktadır [99].

Şekil 3.15'te gösterildiği gibi YOLOv3 mimarisi  $3 \times 3$  ve  $1 \times 1$  evrişim filtrelerinin yanı sıra bazı kısayol bağlantılarını (residual connections) içeren ve toplamda 53 evrişim katmanından oluşan Darknet53 çerçevesine dayanmaktadır. Sınıflandırıcı katmanında, çok etiketli sınıflandırma amacıyla Softmax yerine her sınıf için bağımsız çalışan lojistik sınıflandırıcı kullanılmaktadır. YOLOv3, FPN'den ilham alarak farklı ölçeklerdeki nesneleri algılayabilmek için çok ölçekli tahmin yaklaşımını benimsemiştir. Farklı evrişim çekirdek boyutlarına sahip öznelik piramitleri, farklı boyutlardaki nesneleri tespit etmek için kullanılmaktadır. Daha spesifik olarak YOLOv3, büyük, orta ve küçük nesneleri tespit edebilmek amacıyla sırasıyla  $13 \times 13$ ,  $26 \times 26$  ve  $52 \times 52$  boyutlarında üç farklı öznelik haritası kullanılmaktadır [99], [106], [116].

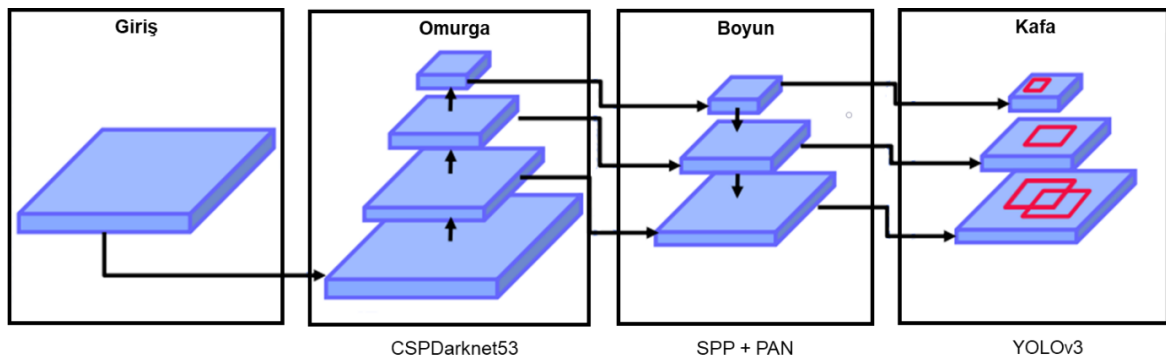


Şekil 3.15. YOLOv3 mimarisi [119].

Bununla birlikte, YOLOv3'ün çıkarımı daha yavaştır, çünkü tespit sürecinde birden fazla evrişimsel katmanın çalıştırılması gerekmektedir. Ayrıca, bazı durumlarda hatalı nesne konumlandırma sorunları bulunmaktadır [116].

### 3.4.4. YOLOv4

2020 yılında Alexey Bochkovskiy et al. tarafından sunulan YOLOv4 hem hız hem de doğruluk açısından önceki tüm versiyonları geride bırakmıştır. YOLOv4 itibariyle nesne dedektörlerinin mimarisi Şekil 3.16'da görüldüğü gibi omurga (backbone), boyun (neck) ve kafa (head) olmak üzere üç bileşende tanımlanmaya başlanmıştır [106], [120]:

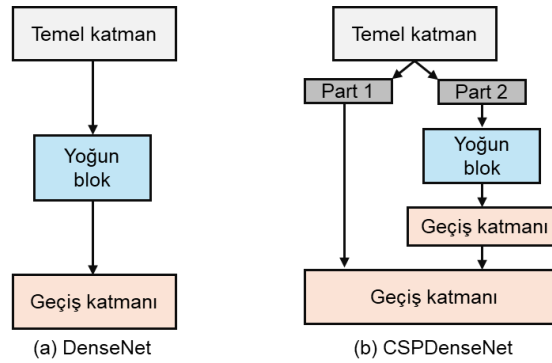


Şekil 3.16. YOLOv4 mimarisinin genel yapısı [120].

- Omurga (Backbone): Bu bileşen, farklı derinliklerdeki katmanlar aracılığıyla hem alt düzey öznitelikleri (kenarlar ve dokular) hem de yüksek düzey öznitelikleri (nesne parçaları ve anlamsal özellikler) yakalamayı amaçlamaktadır. Genellikle, ImageNet gibi büyük ölçekli görüntü sınıflandırma görevlerinde önceden eğitilmiş evrişimli sinir ağları bu amaçla kullanılmaktadır [17], [106].
- Boyun (Neck): Omurga ile kafa arasında konumlanan bu katman, çıkarılan öznitelikleri zenginleştirerek daha güçlü temsiller oluşturmayı hedefler. Uzamsal ve anlamsal bilgileri birleştirerek, tespit performansını artıran bir köprü işlevi görür. YOLO modellerinde yaygın olarak kullanılan boyun yapıları arasında FPN ve Yol Toplama Ağı (Path Aggregation Network, PAN) bulunmaktadır [17], [106].
- Kafa (Head): Mimarinin son bileşeni olan kafa katmanı, omurga ve boyun tarafından sağlanan öznitelikleri kullanarak sınırlayıcı kutuların koordinatlarını ve sınıf etiketlerini tahmin eder. Bu katman, ağın nihai çıktısını üretir [17], [106].

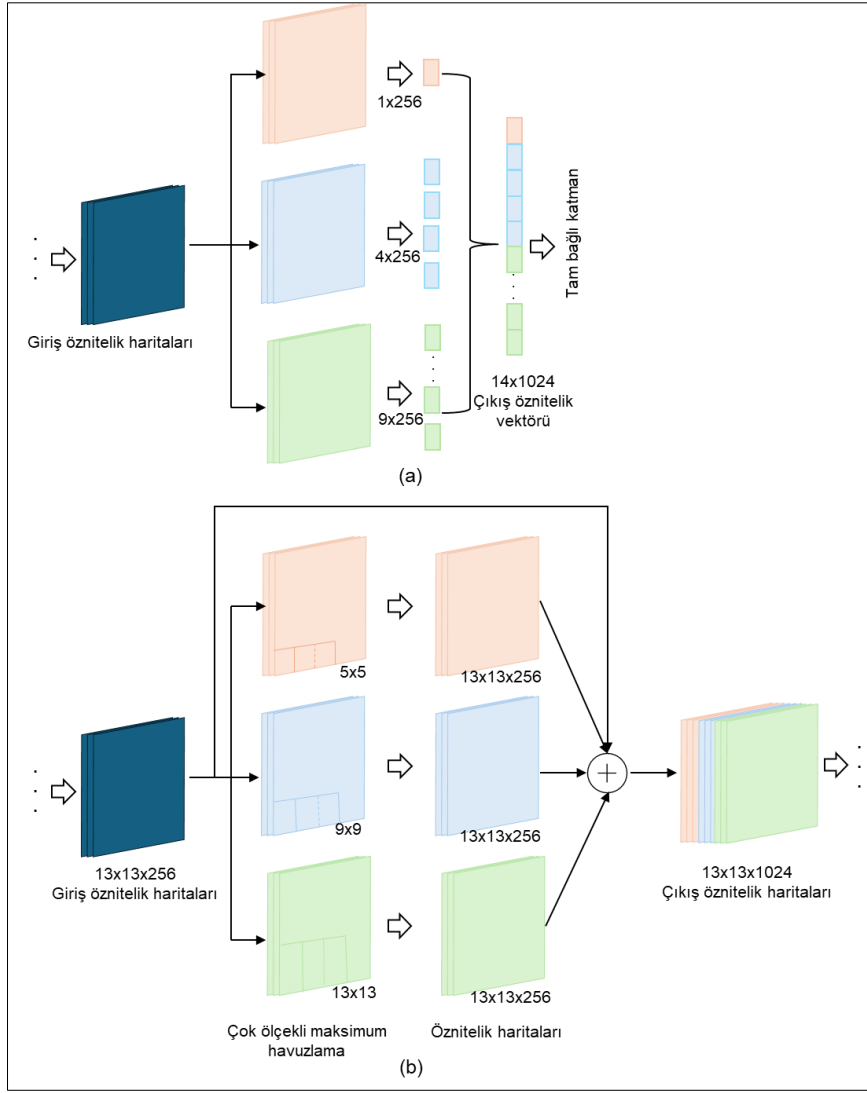
YOLOv4 omurga katmanı için CSP-ResNeXt50, CSP-Darknet53 ve EfficientNetB3 mimarileri karşılaştırılmış ve bu mimariler arasında en başarılı olan CSP-Darknet53,

öznitelik çıkarımı için omurga ağı olarak belirlenmiştir [99]. CSP-Darknet53, kaybolan gradyan sorunlarını azaltmak amacıyla katmanlar arası bağlantılar kuran DenseNet mimarisinden türetilmiştir. DenseNet, her yoğun katmanın çıktısının bir sonraki katmana geçmeden önce girdisiyle birleştirildiği yoğun bloklardan ve geçiş katmanlarından oluşmaktadır. Şekil 3.17’de görüldüğü gibi Çapraz Aşamalı Kısmi (Cross-Stage Partial, CSP), girişi iki parçaya bölerek DenseNet yapısını değiştirmektedir; bu parçalardan biri yoğun bloğu takip ederken, diğeri doğrudan bir sonraki aşamaya iletilmektedir. Bu yaklaşım, gereksiz gradyan bilgi akışını azaltır ve öznitelik yayılımını iyileştirir [121].



Şekil 3.17. (a) DenseNet ve (b) CSPDenseNet yapılarının karşılaştırılması [121].

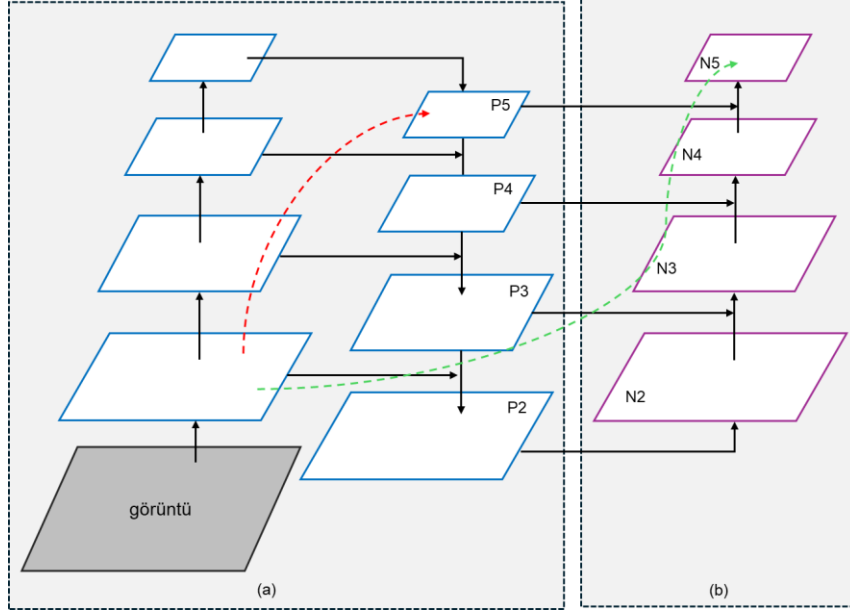
CSPDarknet53 omurgasından gelen öznitelik haritası, boyun kısmına iletilmeden önce Uzamsal Piramit Havuzlama (Spatial Pyramid Pooling, SPP) bloğundan geçmektedir. Bu bloğun amacı öznitelik haritasının alıcı alanını artırmaktır [99]. Klasik SPP bloğu, farklı çekirdek boyutlarında çok ölçekli havuzlama yaparak çeşitli öznitelikleri çıkarmaktadır. Ardından bu bilgileri tek boyutlu vektöre dönüştürerek tam bağlı katmanlara göndermektedir. Ancak bu süreçte, öznitelik haritasının uzamsal boyut bilgisi kaybolmaktadır. Buna karşılık, YOLO’daki değiştirilmiş SPP bloğu ise dolgu (padding) ve birleştirme tekniklerini kullanarak uzamsal boyutları korumaktadır ve böylece önemli detayların kaybolmasını önlemektedir. YOLO’da, SPP bloğu omurgaya bitişik olarak konumlandırılmış olup, giriş öznitelik haritasını azaltmak için omurga ağı ile SPP bloğu arasında 1x1 evrişim katmanı kullanılmaktadır [122]. Klasik SPP bloğu ve YOLO’daki SPP bloğu Şekil 3.18’de gösterilmektedir.



Şekil 3.18. (a) Klasik SPP bloğu ve (b) YOLO'ya uyarlanmış SPP bloğu [123].

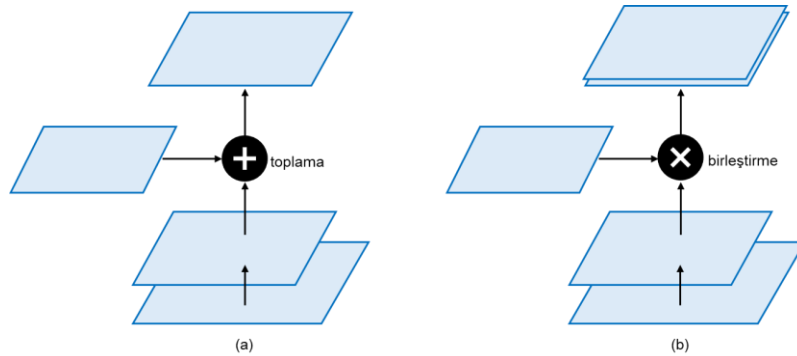
Öznitelik birleştirme aşamasında, FPN ve PAN gibi çeşitli teknikler denenmiş ve mimarinin boyun kısmında öznitelik toplayıcı olarak PANet seçilmiştir. PAN, Şekil 3.19'da gösterildiği gibi FPN'nin gelişmiş bir versiyonudur. FPN, yukarıdan aşağıya bilgi akışı sayesinde büyük ölçekli dedektörlerin hem yüksek seviyeli anlamsal (semantik) bilgilere hem de düşük seviyeli detaylı özniteliklere erişmesini sağlar. Bu mimaride, üst katmanlardan gelen güçlü anlamsal bilgiler, alt seviyelere yanal bağlantılar yoluyla aktarılır (Şekil 3.19, kırmızı yol). Ancak bu yaklaşımda, küçük ölçekli dedektörler yalnızca üst katmanlardan türeyen anlamsal bilgilere dayanmakta ve bu durum, konum hassasiyeti gibi detaylarda performans kaybına yol açabilmektedir. Bu eksikliği gidermek amacıyla PANet, FPN yapısına aşağıdan yukarıya yol artırımı (bottom-up path augmentation) ekleyerek çift yönlü bilgi akışını mümkün kılmıştır. Bu mekanizma sayesinde, alt katmanlarda bulunan detaylı konumsal öznitelikler, doğrudan üst katmanlara taşınabilir hale gelir (Şekil 3.19, yeşil yol).

PANet'in sunduğu bu kısayol bağlantıları, FPN'deki uzun bilgi aktarım yollarına kıyasla daha az katmandan geçerek daha hızlı ve etkili bir öznitelik yayılımı sağlar.



Şekil 3.19. PANet yapısı, (a) FPN, (b) aşağıdan yukarı yol artırımı [124].

Aşağıdan yukarıya yol artırımı, her aşamada aynı uzamsal boyutlara sahip öznitelik haritaları üreten katmanlar içerdiğinden, FPN'nin yukarıdan aşağıya yolunun bir kopyası olarak görülebilmektedir. Bu öznitelik haritaları, eleman bazında toplama işlemiyle yanıl mimariye entegre edilmektedir (Şekil 3.20 (a)). Ancak YOLOv4'te kullanılan değiştirilmiş PAN mimarisinde (Şekil 3.20 (b)), bu toplama işlemi yerine kanal boyunca birleştirme (concatenation) uygulanmaktadır. Bu değişiklik sayesinde, hem FPN'den hem de aşağıdan yukarıya gelen öznitelikler bilgi kaybı olmadan birleştirilebilmektedir [124].



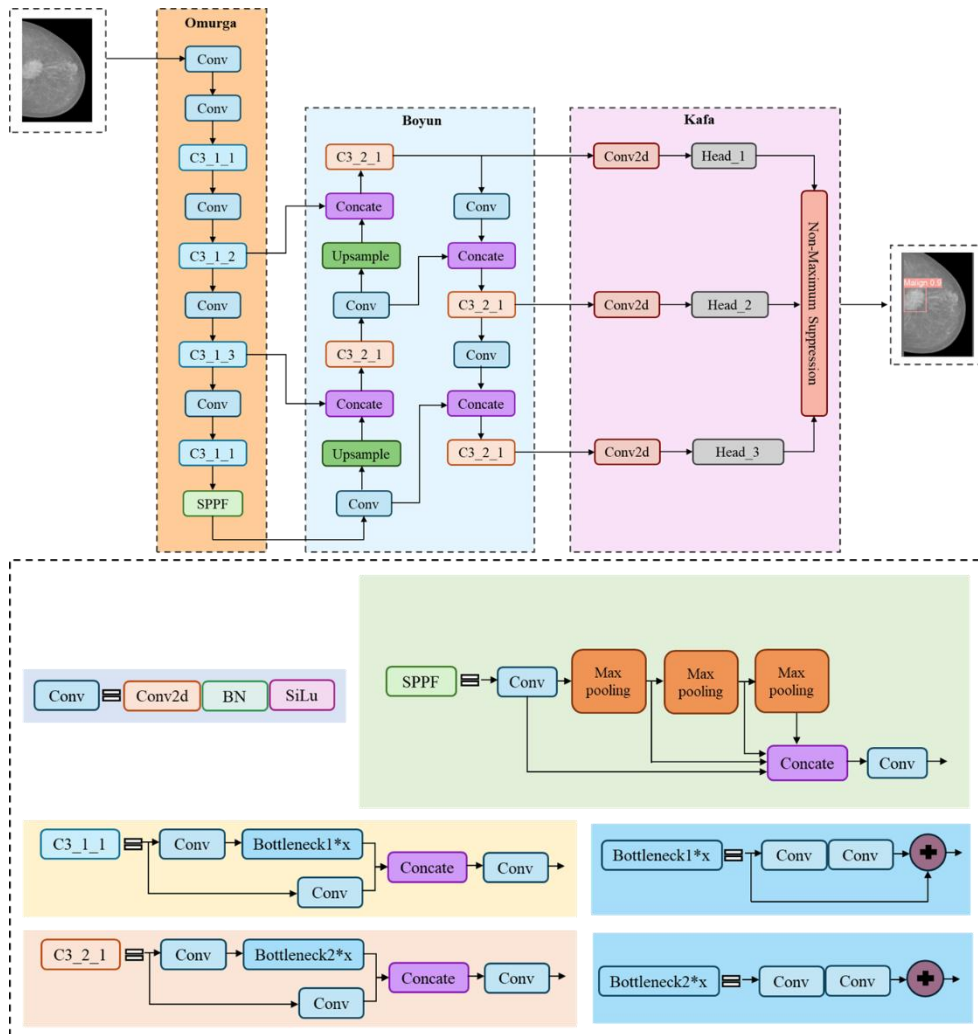
Şekil 3.20. (a) PAN, (b) YOLOv4 değiştirilmiş PAN [124].

Son olarak, YOLOv4 modelinin tespit kafası kısmı, YOLOv3 ile aynı şekilde tasarlanmıştır [107].

YOLOv4, performansı optimize etmek amacıyla ücretsiz çanta (Bag-of-Freebies, BoF) ve özel çanta (Bag-of-Specials, BoS) yaklaşımlarını bir arada kullanmaktadır. BoF, yalnızca eğitim stratejisini değiştirerek eğitim maliyetini artıran ancak çıkarım süresini etkilemeyen yöntemleri içermektedir. En yaygın örneği veri artırmadır. BoS ise çıkarım süresinde küçük bir maliyet artışı yaratırken doğruluğu önemli ölçüde iyileştiren yöntemlerden oluşmaktadır. Bu yöntemlere alıcı alanın genişletilmesi, özneliklerin birleştirilmesi ve son-işleme (post-processing) gibi teknikler örnek verilebilir [106], [107].

### 3.4.5. YOLOv5

YOLOv5, 2020 yılında Ultralytics CEO'su Glen Jocher tarafından, YOLOv4'ten kısa bir süre sonra GitHub üzerinden yayımlanmıştır [125]. YOLOv5 mimarisi Şekil 3.21'de gösterildiği üzere omurga katmanı, boyun katmanı ve tahmin kafası katmanı olmak üzere üç ana bileşenden oluşmaktadır [112].



Şekil 3.21. YOLOv5 mimarisi [112].

Omurga katmanı, öznitelik çıkarımı amacıyla CSP-DarkNet53 ağını benimsemektedir ve sırasıyla Conv (Conv2d-BN-SiLu), C3 ve Uzamsal Piramit Havuzlama Hızlı (Spatial Pyramid Pooling Fast, SPPF) modüllerinden oluşmaktadır. Conv modülü, ilk olarak öznitelik çıkarımı için iki boyutlu evrişim (Conv2d) işlemini, ardından eğitim sürecini hızlandırmak ve modelin genelleme yeteneğini iyileştirmek amacıyla BN işlemi ve son olarak modelin doğrusal olmayan uyum yeteneğini geliştirmek için SiLu (Sigmoid Linear Units) aktivasyon fonksiyonunu içermektedir [112], [126]. C3 modülü, öznitelik gösterimini daha verimli hale getirmek için Conv modülü tarafından çıkarılan öznitelikleri istifleyen bir modüldür [112]. C3'ün temel fikri CSPNet'ten gelmektedir [127]. Omurgada C3\_1 ve boyunda C3\_2 olmak üzere iki farklı C3 yapısı kullanılmaktadır [128]. SPPF modülü, farklı ölçeklerdeki öznitelikleri havuzlama aracılığıyla sabit boyutlu bir öznitelik haritasında birleştirmektedir. Bu yapı sayesinde, modelin farklı ölçeklere yönelik bilgileri yakalama ve algılama yeteneği geliştirilmektedir [106], [129].

Boyun katmanı, öznitelik birleştirme (füzyon) amacıyla PANet yapısını temel almaktadır [130]. PANet, yukarıdan aşağıya (top-down) ve aşağıdan yukarıya (bottom-up) olmak üzere çift yönlü bilgi akışı sağlayarak, farklı seviyelerdeki öznitelikleri birleştirir [131]. Bu yaklaşım, modelin farklı ölçekli hedeflerin özniteliklerini daha iyi yakalamasına yardımcı olmaktadır ve tespit performansını arttırmaktadır [129].

Tahmin kafası, nesnelerin konumunu ve sınıf etiketini tespit etmektir sorumludur. Bu aşamada, NMS işlemi ile gereksiz kutular ortadan kaldırılarak nihai sonuçlar elde edilmektedir [128], [129].

YOLOv5, farklı model boyutları ve donanım gereksinimlerine göre YOLOv5n (nano), YOLOv5s (small), YOLOv5m (medium), YOLOv5l (large) ve YOLOv5x (extra-large) olmak üzere beş farklı versiyon sunmaktadır [51]. YOLOv5n'den YOLOv5x'e doğru gidildikçe modelin genişliği ve derinliği artmakta, buna bağlı olarak parametre sayısı yükselmekte ve işlem hızı azalmaktadır [130].

### **3.4.6. YOLOv6**

Eylül 2022'de Meituan Görsel Zeka Departmanı tarafından yayınlanan YOLOv6, endüstriyel uygulamalarda karşılaşılan pratik sorunları ele almayı amaçlamaktadır [132]. Endüstriyel uygulama gereksinimlerini karşılamak için mimarinin, yüksek hız ve doğruluğu koruyarak farklı donanım seçeneklerinde yüksek performans göstermesi gerekmektedir. Bu doğrultuda YOLOv6, en az parametreye sahip ve en hızlı sürüm olan YOLOv6-nano'dan

başlayarak, daha yüksek doğruluk sunan ancak hızdan bir miktar ödün veren YOLOv6-large sürümüne kadar çeşitli model versiyonlarıyla sunulmaktadır [117].

YOLOv6'nın en önemli yenilikleri arasında EfficientRep Omurgası, Rep-PAN Boynu, Ayırıştırılmış Kafa (Decoupled Head) ve etkili eğitim stratejileri yer almaktadır. YOLOv6, önceki versiyon YOLOv5'in CSP omurgasından farklı olarak, güçlü temsili kapasiteye sahip olacak ve donanım işleme kaynaklarını optimize edecek şekilde tasarlanmıştır. Modelin Rep-PAN boyun yapısı, donanım sinir ağı tasarımına dayalı olarak geliştirilmiş olup, öznetelik birleştirme sürecini daha etkili hale getirmektedir. Bu yapı, donanım tüketimini azaltırken doğruluk ve hız arasındaki dengeyi optimize etmektedir. Ayrıca, YOLOv6'nın ayırıştırılmış kafa yapısı sayesinde, sınıflandırma ve regresyon görevleri ayrı dallarda gerçekleştirilmekte, bu da hem temsili kapasiteyi artırmakta hem de hesaplama yükünü daha dengeli dağıtarak modelin genel verimliliğini artırmaktadır. Modelin tespit doğruluğunu artırmak amacıyla, anchor içermeyen (anchor-free) paradigma, SimOTA etiket atama algoritması ve SIOU (Scaled IoU) sınırlayıcı kutu regresyon kaybı gibi ileri düzey eğitim teknikleri uygulanmaktadır. Bu yeniliklerin birleşimi sayesinde YOLOv6, endüstriyel uygulamalara uygun, yüksek hız ve doğruluk sunan güçlü bir nesne tespit modeli olarak öne çıkmaktadır [91], [117].

### 3.4.7. YOLOv7

YOLOv6'nın yayınlanmasından yalnızca bir ay sonra, YOLOv4'ün yazarları tarafından YOLOv7 geliştirilmiş ve yayınlanmıştır. YOLOv7, önceki sürümlere kıyasla daha yüksek doğruluk ve tespit hızı için çeşitli mimari yenilikler önermektedir [133]. Önerilen bu yenilikler, mimari tasarım değişiklikleri ve eğitilebilir ücretsiz çantalar (trainable BoF) yaklaşımı olmak üzere iki kategoriye ayrılmaktadır.

Mimari tasarımdaki yeniliklerden ilki, Verimli Katman Toplama Ağı'nın (Efficient Layer Aggregation Network, ELAN) genişletilmiş sürümü olan Genişletilmiş ELAN (Extended ELAN, E-ELAN) yapısıdır. ELAN, ağ içindeki gradyan akışını kontrol ederek daha verimli derin öğrenme modelleri oluşturmayı hedeflemektedir. Bu kontrollü bilgi akışı sayesinde, geri yayılım süreci iyileşmekte ve modelin öğrenme kapasitesi artmaktadır. E-ELAN, tıpkı ELAN gibi karıştırma (mixing), genişletme (expansion) ve birleştirme (aggregation) işlemlerini gerçekleştiren paralel yol yapısını temel almaktadır. Ancak E-ELAN'da yalnızca yapısal düzeyde değişiklik yapılmakta, kanal sayısı (kanal çarpanı) değiştirilmemektedir. Böylece gradyanların geçiş yolları korunurken öğrenme performansı

daha da geliştirilmektedir [113], [133]. İkinci yenilik ise bağlantı tabanlı ölçekleme yöntemidir. Bu yöntem, modelin yapısal bloklarının derinlik ve genişliğini eş zamanlı olarak ölçeklendirerek farklı boyutlarda modeller oluşturmayı hedeflemektedir. Geleneksel derinlik ölçekleme yöntemleri, giriş ve çıkış kanalları arasındaki oranı değiştirerek donanım gereksinimlerini azaltmaktadır. YOLOv7 ise model mimarisini bozmadan tüm blokları orantılı şekilde ölçekleyerek yapısal bütünlüğü koruyan daha verimli bir ölçekleme stratejisi sunmaktadır [91], [113].

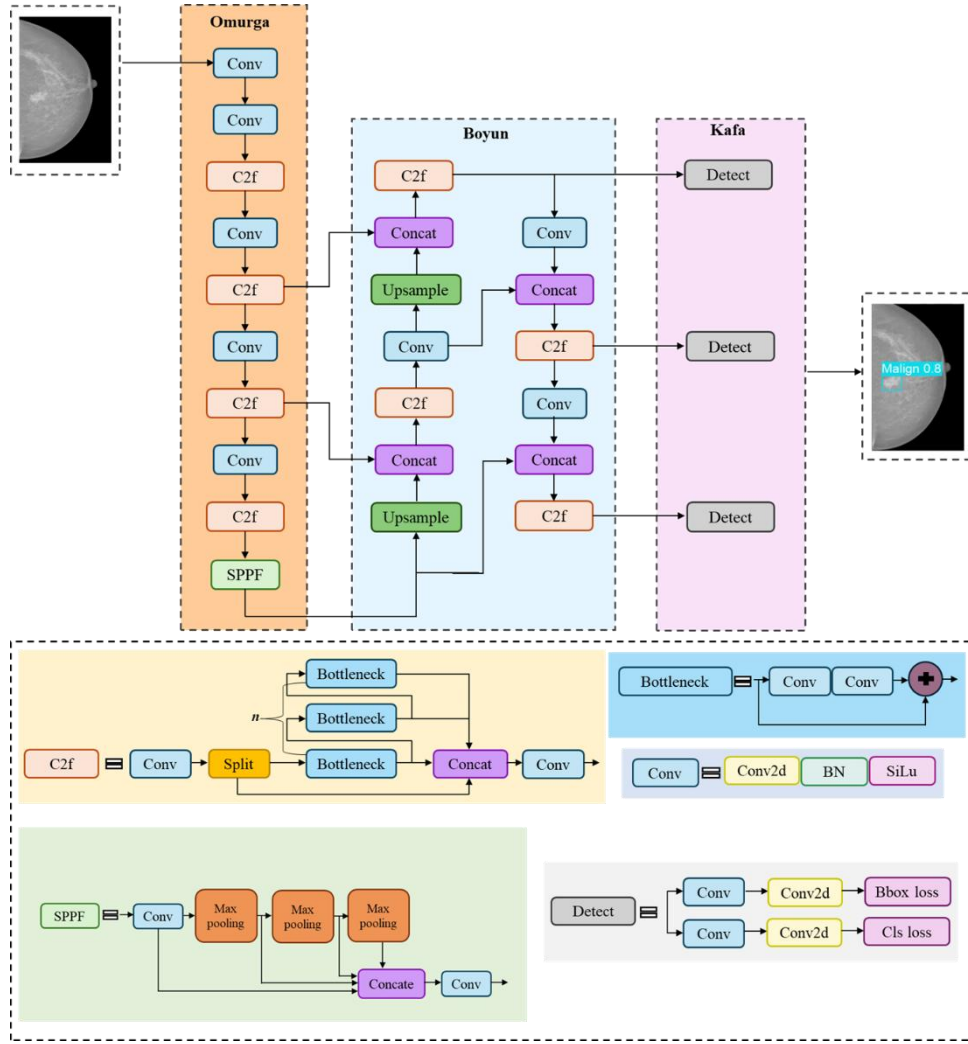
Eğitilebilir BoF yaklaşımındaki yeniliklerden ilki, planlanmış yeniden parametrelendirilmiş evrişimdir. YOLOv7, yeniden parametrelendirilmiş evrişim (Re-parameterized Convolution, RepConv) yapısından esinlenerek geliştirilmiştir. Bu yapıda kullanılan kimlik (identity) bağlantısı, bazı durumlarda kalıntı (residual) ve birleştirme (concatenation) bağlantılarına dayanan DenseNet ve ResNet gibi mimarilerin performansını olumsuz etkileyebilir. Bu sorunu önlemek amacıyla, YOLOv7 modelinde kimlik bağlantılarını kaldırarak RepConv kullanılmıştır [113].

Diğer bir yenilik ise kaba-ince (coarse-fine) etiket atama yaklaşımıdır. Bu yöntemde, yardımcı kafa (auxiliary head) ve öncü kafa (lead head) olmak üzere çift başlı bir yapı kullanılmaktadır. Öncü kafa, son tahminleri üretmekle sorumludur ve daha detaylı (ince) etiketlerle eğitilmektedir. Yardımcı kafa ise eğitim sürecine rehberlik etmekte ve daha genel (kaba) etiketlerle destekleyici öğrenmeyi sağlamaktadır [91], [113].

YOLOv7'nin farklı varyasyonları arasında YOLOv7-tiny, YOLOv7-X, YOLOv7-E6 ve YOLOv7-D6 modelleri yer almaktadır [133].

### **3.4.8. YOLOv8**

YOLOv8 [134] mimarisi, YOLOv5'in geliştirilmesinden sorumlu olan Ultralytics tarafından Ocak 2023'te yayımlanmıştır. Artırılmış doğruluk ve daha yüksek tespit hızı sayesinde önceki YOLO serilerin çoğundan daha iyi performans göstermektedir [135]. YOLOv8 mimarisi Şekil 3.22'de gösterildiği gibi omurga katmanı, boyun katmanı ve tahmin kafası katmanı olmak üzere üç bileşenden oluşmaktadır.



Şekil 3.22. YOLOv8 mimarisi [136].

YOLOv8'in omurga katmanı temel olarak YOLOv5 ile benzer yapıdadır. Ancak CSP mimarisine dayalı C3 modülü yerine C2f modülü kullanılmıştır [137]. C2f modülü, YOLOv7'deki ELAN yapısının faydasından ilham alarak C3 modülü ile ELAN'ı birleştirmiştir. Bu sayede model, hem daha hafif (lightweight) hale gelmiş hem de daha güçlü gradyan akışı sağlayarak öğrenme kapasitesini artırmıştır [137], [138]. Omurga katmanında, giriş görüntüsü farklı ölçeklerde öznetelik haritaları çıkarmak için birden fazla Conv ve C2f modülü aracılığıyla işlenmektedir. Bu katmanın sonunda da SPPF modülü, çıktı öznetelik haritalarını farklı çekirdek boyutlarında havuzlama işlemi ile birleştirmektedir [139].

Boyun katmanı, öznetelik birleştirme (füzyon) için PANet yapısını korumaktadır. Bu yapı, yukarı ve aşağı örnekleme işlemleri yoluyla yüksek ve düşük düzeyli öznetelikleri birleştirerek hem anlamsal hem de konumsal bilgilerin etkili bir şekilde aktarılmasını sağlar [139].

YOLOv5 ile karşılaştırıldığında, YOLOv8 tahmin kafası, sınıflandırma ve sınırlayıcı kutu tahminini ayıran ayrıştırılmış kafa (decoupled head) yapısı benimsenmiştir [136], [140]. Sınıflandırma görevi için Eşitlik 3.2’de verilen İkili Çapraz Entropi Kaybı (Binary Cross-Entropy (BCE)) kullanılmaktadır.

$$Loss_{BCE} = -w[y_n \log x_n + (1 - y_n) \log(1 - x_n)] \quad (3.2)$$

Eşitlik 3.2’de  $w$  ağırlığı belirtir;  $y_n$  etiketli değeri temsil eder ve  $x_n$  ise modelin tahminini ifade etmektedir.

Sınırlayıcı kutu tahmini için Dağıtım Odak Kaybı (Distribute Focal Loss, DFL) ve Birleşim Üzerinde Tam Kesişme (Complete Intersection over Union, CIoU) olmak üzere iki temel kayıp fonksiyonu kullanılmaktadır. Bu iki yapı, tespit doğruluğunu arttırmayı ve modelin yakınsama hızını hızlandırmayı amaçlamaktadır [136]. DFL, hedef kutunun konumunu genel bir dağılım olarak modellemeyi amaçlar, ağırlık  $y$  etiketine yakın değerlere daha hızlı odaklanmasını sağlayarak olasılıklarını artırır [141]. Başka bir deyişle DFL işlevi,  $y_n$  ve  $y_{n+1}$  olasılık değerlerinin ( $y_n \leq y \leq y_{n+1}$ )  $y$  nesnesi etrafındaki genişlemesini vurgulamak için tasarlanmıştır [140]. DFL Eşitlik 3.3 ile hesaplanır.

$$Loss_{DFL} = -[(y_{n+1} - y) \log \frac{y_{n+1} - y}{y_{n+1} - y_n} + (y - y_n) \log \frac{y - y_n}{y_{n+1} - y}] \quad (3.3)$$

CIoU kaybı, tahmin edilen sınırlayıcı kutu ile gerçek kutu arasındaki en-boy oranını dikkate alarak benzerliğin daha doğru ve kapsamlı bir ölçümünü amaçlar ve Eşitlik 3.4 ile ifade edilir [140], [141].

$$Loss_{CIoU} = 1 - IoU + \frac{d^2}{c^2} + \frac{v^2}{(1 - IoU) + v'} \quad (3.4)$$

Eşitlik 3.4’te  $d$ , tahmin edilen ve gerçek sınırlayıcı kutuların merkez noktaları arasındaki öklid uzaklığıdır.  $c$  her iki kutuyu çevreleyen en küçük kutunun köşegen uzunluğunu;  $v$  ise Eşitlik 3.5 ile tanımlanan en boy oranının tutarlılığını ölçen parametreyi temsil eder [140].

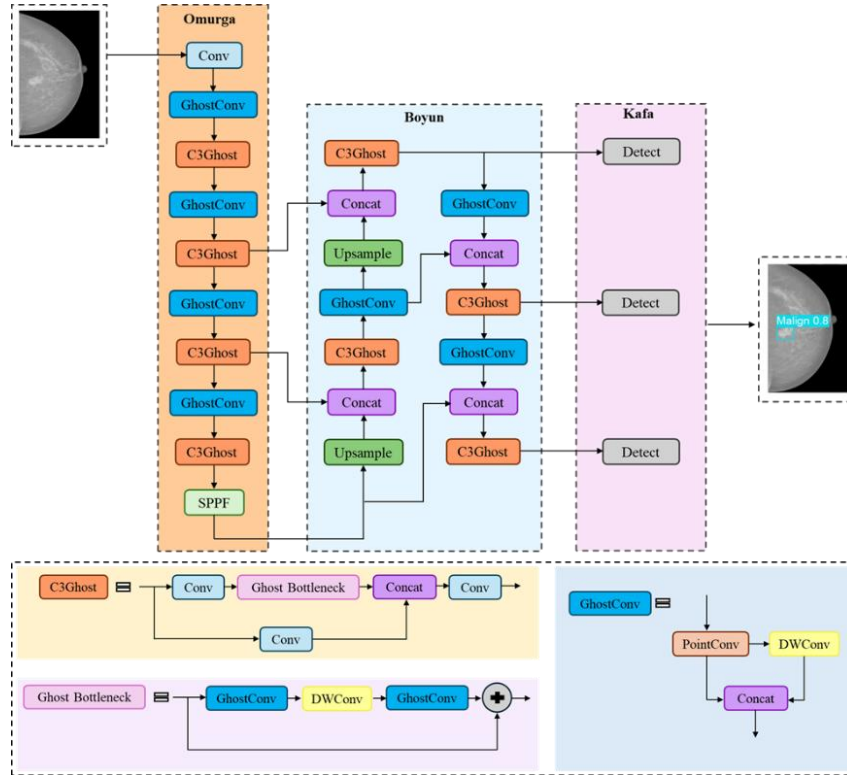
$$v = \frac{4}{\pi^2} \left( \arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w_p}{h_p} \right)^2 \quad (3.5)$$

Eşitlik 3.5’te  $w$  ve  $h$  sırasıyla sınırlayıcı kutunun genişliğini ve yüksekliğini belirtir;  $gt$  temel gerçek (ground truth) ve  $p$  ise tahmin (prediction) anlamına gelir.

YOLOv8, “anchor-free” nesne tespiti yaklaşımını benimseyerek, sınırlayıcı kutuların öğrenimini daha verimli hale getirmiştir. Bu yaklaşım, yalnızca en olası konumları değerlendirerek NMS işlemini daha hızlı ve hesaplama açısından daha verimli biçimde gerçekleştirmektedir [142].

### 3.4.9. YOLOv8-Ghost

YOLOv8-Ghost mimarisi Şekil 3.23’te gösterilmektedir. Standart YOLOv8’den farklı olarak, bu mimaride omurga katmanındaki ilk Conv modülü dışındaki tüm Conv ve C2f modülleri, sırasıyla GhostConv ve C3Ghost modülleri ile değiştirilmiştir. C3Ghost, Ghost modüllerinin ve Ghost Bottleneck yapısının avantajlarını kullanan, orijinal C3 modülünün daha hafif ve verimli bir versiyonudur [43], [143]. Bu modül, birden fazla evrimsel işlemde ağırlık paylaşımı yaparak hem hesaplama karmaşıklığını azaltmakta hem de modelin toplam parametre sayısını düşürmektedir. Bu sayede, doğruluktan ödün vermeden daha az parametreye sahip bir model elde edilmektedir [144]. Ghost Bottleneck yapısında, klasik bottleneck yapısındaki standart evrişim katmanları, GhostConv modülleriyle değiştirilmiştir. İlk Ghost evrişimi, kanal sayısını artırmak ve öznelik boyutunu genişletmek amacıyla; ikinci Ghost evrişimi ise bu boyutu azaltmak üzere kullanılmaktadır [145].

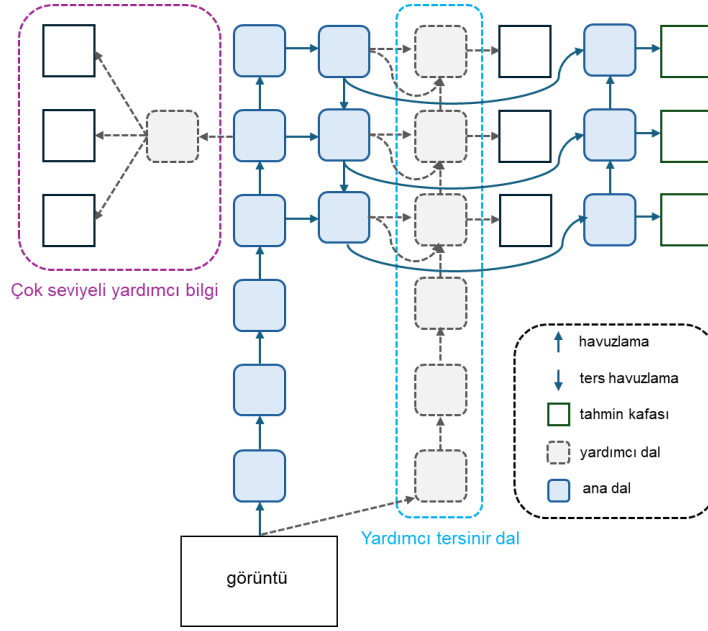


Şekil 3.23. YOLOv8-Ghost mimarisi [43].

### 3.4.10. YOLOv9

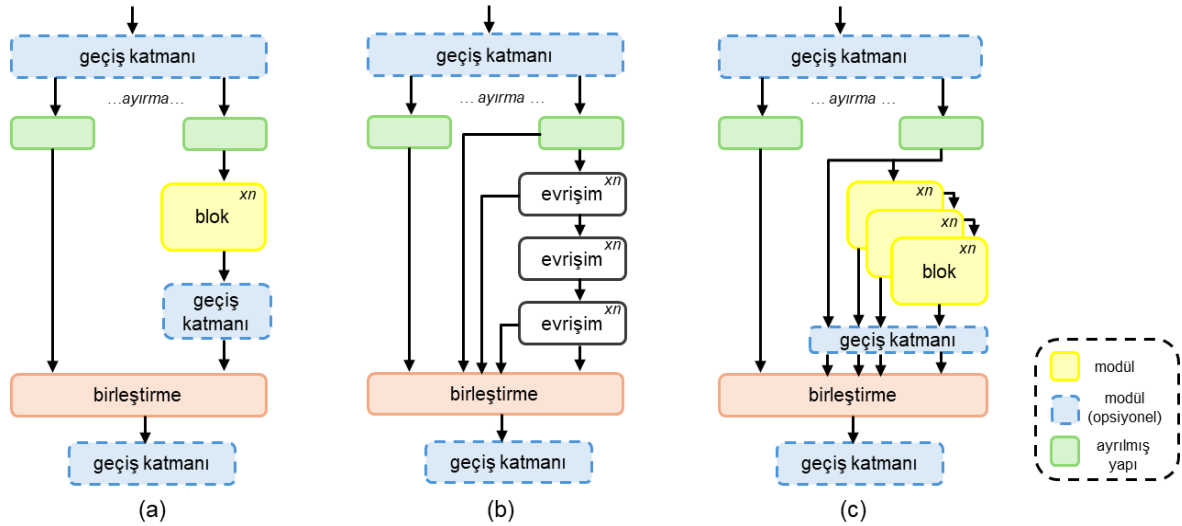
Şubat 2024'te yayınlanan YOLOv9, önceki YOLO sürümlerine kıyasla iki temel yenilik sunmaktadır. Bunlardan ilki, Programlanabilir Gradyan Bilgisi (Programmable Gradient Information, PGI) yaklaşımıdır. Derin sinir ağlarında girdi verilerinin ardışık katmanlardan geçişi bilgi kaybı problemini meydana getirir; bu durum da modelin eğitimi sırasında eksik bilginin kullanılmasına neden olur. PGI yaklaşımı, eğitim sürecinde daha güvenilir gradyan bilgileri üreterek modelin öğrenme kapasitesini artırmayı hedeflemektedir. Diğer bir yenilik ise Genelleştirilmiş Verimli Katman Toplama Ağı (Generalized Efficient Layer Aggregation Network, GELAN) olarak adlandırılan, gradyan yollarını daha verimli planlayan yeni ve hafif bir ağ mimarisidir [146].

PGI, modelin tespit performansını artırmak amacıyla, farklı anlamsal düzeylerde gradyan bilgisinin yayılımını yöneten destekleyici bir çerçevedir. Şekil 3.24'te gösterildiği gibi PGI temel olarak üç bileşenden oluşur: (1) ana dal (main branch), (2) yardımcı tersinir dal (auxiliary reversible branch) ve (3) çok seviyeli yardımcı bilgi (multi-level auxiliary information). Çıkarım sürecinde ana dal hem ileri hem de geri yayılım işlemlerinden sorumludur. Ancak ağ derinleştikçe bilgi kaybı oluşabilir ve bu durum, kayıp fonksiyonlarının güvenilir gradyanlar üretememesine yol açabilir. Bu sorunu gidermek amacıyla, yardımcı geri dönüşümlü dal, ana daldaki bilgi kaybını azaltmak ve daha sağlam gradyan üretimini desteklemek için tersinir fonksiyonlardan yararlanır. Buna ek olarak, çok seviyeli yardımcı bilgi bileşeni, derin denetimde ortaya çıkabilecek hata birikimini azaltmayı hedefler. Bu bileşen, farklı katmanlara ek bilgi sunarak modelin genel öğrenme kapasitesini ve performansını artırır [147], [148].



Şekil 3.24. YOLOv9 PGI mimarisi [146].

GELAN model eğitimi sırasında katmanlar arası bilgi etkileşimini artırmak ve bilgi yayılımı sürecinde meydana gelebilecek kayıpları en aza indirmek amacıyla, Şekil 3.25'te gösterildiği gibi CSPNet ve ELAN yapılarını birleştirir. Bu mimari, parametre sayısının azaltılması ve hesaplama verimliliğinin artırılması sayesinde yüksek doğruluğu korurken, çıkarım sürecini de hızlandırmaktadır [147], [149].

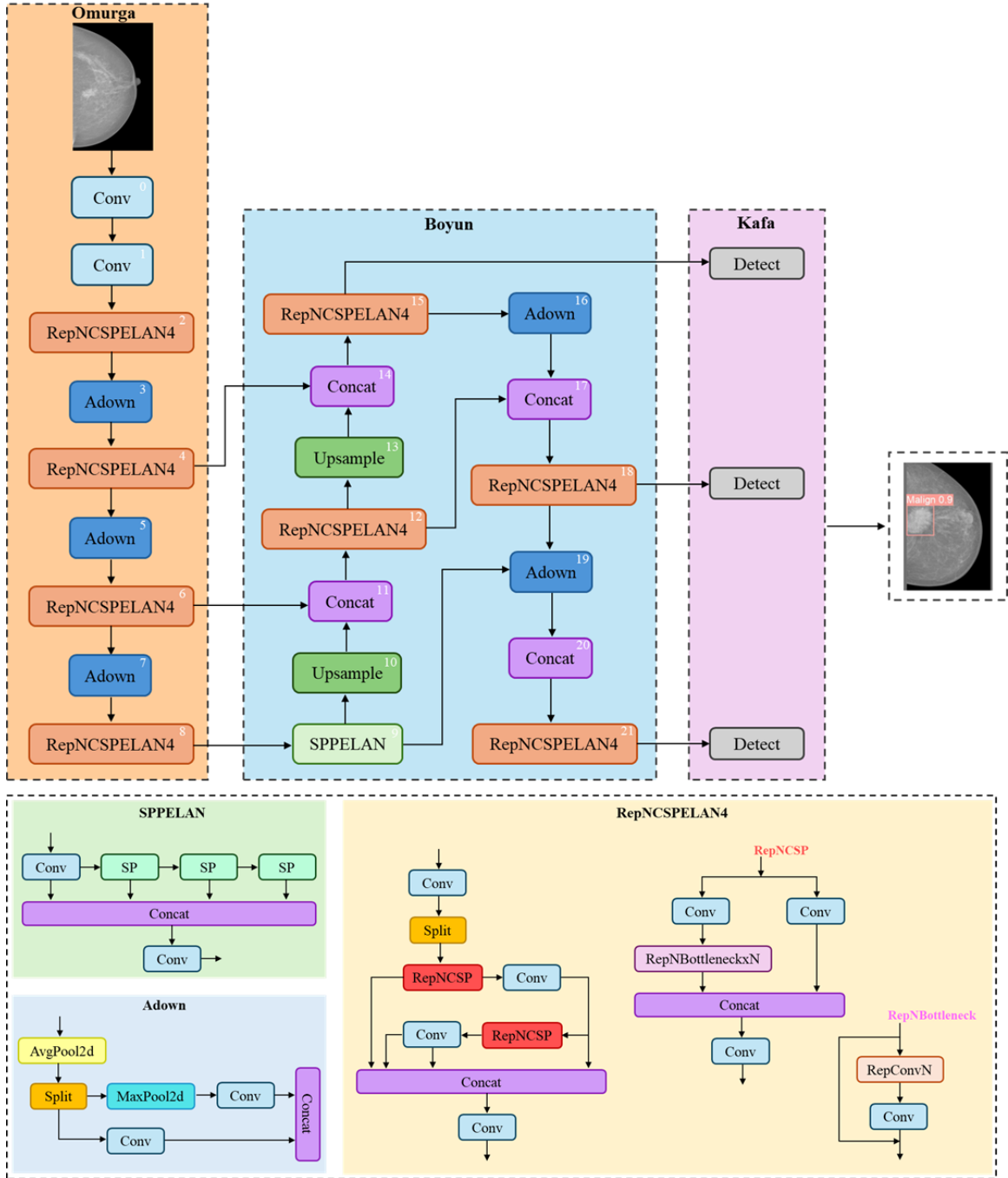


Şekil 3.25. GELAN mimarisi: (a) CSPNet, (b) ELAN, (c) GELAN [146].

YOLOv9, her biri farklı performans ve kaynak gereksinimlerini karşılayan farklı modeller (YOLOv9-tiny, YOLOv9-small, YOLOv9-medium, YOLOv9-compact, YOLOv9-extended, GELAN-tiny, GELAN-small, GELAN-medium, GELAN-compact,

GELAN-extended) sunmaktadır. Bu modeller, azaltılmış hesaplama gereksinimleriyle yüksek doğruluk sağlayacak şekilde tasarlanmıştır.

Çalışmada hafif ağ mimarisine sahip olması nedeniyle YOLOv9'un GELAN modelleri tercih edilmiştir. Şekil 3.26 GELAN mimarisini göstermektedir.



Şekil 3.26. GELAN mimarisi [150].

### 3.5. YOLO Algoritmasının Hiperparametreleri ve Veri Artırma Teknikleri

YOLO algoritmasının eğitim sürecinde kullanılan hiperparametreler, model eğitimini ve performansını önemli ölçüde etkilemektedir. YOLO eğitiminde yaygın olarak kullanılan temel hiperparametreler ve veri artırma teknikleri sırasıyla özetlenmiştir [151], [152]:

- Öğrenme oranı, modelin eğitim sırasında ağırlıklarını güncelleme hızını belirler. Yüksek öğrenme oranı, daha hızlı yakınsamayı sağlayabilir; ancak bu durum, optimum ağırlıkların atlanmasına ve modelin kararsız öğrenmesine yol açabilir. Düşük öğrenme oranı ise daha kararlı bir eğitim sağlar ancak modelin en iyi çözüme ulaşması daha uzun sürebilir [151].
- Toplu işlem (batch) boyutu, modelin her adımda ağırlıklarını güncellemek için kullandığı eğitim örneği sayısını ifade eder. Daha büyük bir toplu işlem boyutu, daha doğru bir gradyan tahmini sağlayabilir, ancak hesaplama maliyeti açısından zorlayıcı olabilir. Buna karşılık, daha küçük bir toplu işlem boyutu ise daha hızlı ancak daha gürültülü bir öğrenme süreci sunabilir [152].
- Epok sayısı, eğitim verisinin tamamı üzerinde gerçekleştirilen tam tarama sayısını ifade eder. Düşük epok sayısı yetersiz öğrenmeye (underfitting), yüksek epok sayısı ise aşırı öğrenmeye (overfitting) neden olabilir [152].
- Giriş görüntü boyutu, modelin hem doğruluk hem de hız performansını etkiler. Daha büyük boyutlu görüntüler detay açısından avantaj sağlasa da işlem yükünü artırır [151].
- Ağırlık azaltma, modelin aşırı öğrenmesini engellemek amacıyla ağırlıklara uygulanan bir düzenleme tekniğidir. Bu teknik, kayıp fonksiyonuna bir ceza terimi ekleyerek modelin daha sade yapılar öğrenmesini teşvik eder [152].
- Optimizasyon algoritması, eğitim sırasında modelin parametrelerini güncellemek için kullanılır ve kayıp fonksiyonunu en aza indirmeyi hedefler. Genellikle YOLO için Stochastic Gradient Descent (SGD) veya Adaptive Moment Estimation (Adam) algoritmaları tercih edilmektedir [151].
- Momentum, optimizasyon sürecinde model ağırlıklarının güncellenmesini iyileştirmek için kullanılan bir hiperparametredir. Temel olarak, önceki gradyan bilgisinin bir kısmını mevcut gradyana ekleyerek, gradyanların doğru yönde ilerlemesini hızlandırır ve dalgalanmaları azaltır. Bu sayede, model daha kararlı ve hızlı bir şekilde yakınsamaya ulaşabilir [153].
- Erken durdurma, eğitim sırasında modelin doğrulama kümesindeki performansı

izlenerek, doğrulama kaybında iyileşme gözlenmediğinde eğitim sürecinin sonlandırılmasıdır. Bu yaklaşım, aşırı öğrenmeyi engelleyerek eğitim verilerinin ezberleme riskini azaltır [154].

Veri artırma teknikleri, nesne tespit performansını iyileştirmek için önemli araçlardır. Bu teknikler sayesinde, nesne sayısı artırılarak modelin çeşitli örnekler üzerinde eğitilmesi sağlanır. Böylece sınıf etiketleri korunurken, genelleme yeteneği yüksek ve daha sağlam bir nesne tespit modeli elde edilir [54].

Nesne tespit alanında yaygın olarak kullanılan veri artırma teknikleri; döndürme, kaydırma, ölçekleme, kesme, yatay ve dikey çevirme gibi yöntemlerden oluşmaktadır. Bu temel tekniklere ek olarak, son yıllarda YOLO tabanlı nesne tespiti algoritmalarının geliştirilmesiyle birlikte farklı artırma yaklaşımları da tanıtılmıştır [3]. YOLO mimarilerinde veri artırma işlemleri klasik anlamda yeni görüntüler üretmek yerine, eğitim sırasında her görüntünün veri artırma parametrelerine göre rastgele dönüştürülerek modele aktarılması şeklinde gerçekleştirilir. Bu sayede, veri kümesi fiziksel olarak büyütülmeden, modelin her epokta farklı varyasyonlar görmesi sağlanır.

- HSV dönüşüm, görüntülerin renk tonu (hue), doygunluk (saturation) ve değer (value) bileşenlerinde yapılan değişiklikler ile farklı aydınlatma ve renk koşullarına uyum sağlamasına yardımcı olur [155], [156].
- Döndürme (degrees), görüntünün belirli bir açıyla bir eksen etrafında sağa veya sola çevrilmesidir.
- Kaydırma (translate), görüntünün belirli bir oranda yatay veya dikey yönde ötelenmesi ile konumunun değiştirilmesidir.
- Ölçekleme (scale), bir görüntünün belirli bir oranla yakınlaştırılması veya uzaklaştırılmasıdır [155], [157].
- Kesme (shear), görüntünün yatay veya dikey eksen boyunca belirli bir açı ile bir kenarının kaydırılmasıyla görüntünün paralelkenar biçimini almasıdır [156].
- Perspektif (perspective) dönüşümü, aynı üç boyutlu nesnenin alternatif izdüşümleri olan iki farklı görüntüyü iki farklı izdüşüm düzlemine ilişkilendirir [158].
- Yatay veya dikey çevirme (flip left-right/up-down), bir görüntünün yatay veya dikey ekseninde yansıtılarak ayna görüntüsünün oluşturulmasıdır [156].
- Mozaik (mosaic), dört farklı görüntüyü tek bir görüntüde birleştirir. Bu işlem sonucunda, birleştirilmiş görüntüdeki nesnelere orijinal görüntüden daha küçük bir ölçekte görünür. Bu tür bir artırma, görüntülerdeki küçük nesnelere algılanmasını

iyileştirmede faydalıdır ve ilk kez YOLOv4 tarafından kullanılmıştır [120].

- Karıştırma (mixup), rastgele seçilen iki görüntü ile bunlara ait etiketlerin doğrusal olarak birleştirilmesiyle yeni bir görüntü oluşturulmasıdır [159].
- Kopyala-yapıştır (copy-paste), bir görüntüdeki nesnelerin kopyalanıp rastgele seçilen başka bir görüntüye yapıştırılmasıyla yeni bir görüntünün elde edilmesini sağlayan bir veri artırma tekniğidir [160].

YOLO eğitimi sırasında kullanılan veri artırma parametreleri ve değer aralıkları Tablo 3.1’de sunulmuştur.

**Tablo 3.1.** YOLO eğitimi sırasında kullanılan veri artırma parametreleri ve değer aralıkları

| Parametre ve tanımı  | Değer aralığı |
|--|---------------|
| <b>hsv_h:</b> görüntü HSV-ton artırma (kesir)                    | 0.0 - 1.0     |
| <b>hsv_s:</b> görüntü HSV-doygunluk artırma (kesir)              | 0.0 - 1.0     |
| <b>hsv_v:</b> görüntü HSV-değer artırma (kesir)                  | 0.0 - 1.0     |
| <b>degrees:</b> görüntü döndürme (+/- derece)                    | 0.0 – 45.0    |
| <b>translate:</b> görüntü kaydırma (+/-kesir)                    | 0.0 – 0.3     |
| <b>scale:</b> görüntü ölçekleme (+/- kazanç)                     | 0.0 – 0.9     |
| <b>shear:</b> görüntü kesme (+/- derece)                         | 0.0 - 10.0    |
| <b>perspective:</b> görüntü perspektifi (+/-kesir)               | 0.0 – 0.001   |
| <b>flipud:</b> görüntü yukarı-aşağı çevirme (olasılık)           | 0.0 - 1.0     |
| <b>fliplr:</b> görüntü sola-sağa çevirme (olasılık)              | 0.0 - 1.0     |
| <b>mosaic:</b> görüntü mozaik (olasılık)                         | 0.0 - 1.0     |
| <b>mixup:</b> görüntü karıştırma (olasılık)                      | 0.0 - 1.0     |
| <b>copy_paste:</b> segmentasyon için kopyala-yapıştır (olasılık) | 0.0 - 1.0     |

### 3.6. Dikkat Mekanizmaları (Attention Mechanism)

Dikkat, insanlarda tüm algısal ve bilişsel işlemlerin temel bir özelliğidir [161]. Bir görüntüde önemli bölgelere odaklanma ve ilgisiz kısımları göz ardı etme yeteneği, dikkat mekanizmaları olarak adlandırılır. Görsel algı sistemlerinde dikkat, girdinin önemine göre özniteliklerin uyarlamalı olarak ağırlıklandırılmasıyla gerçekleştirilen dinamik bir seçim süreci olarak ele alınabilir. Bu yaklaşım, araştırmacılara bilgisayarlı görü sistemlerinin performansını artırmak için dikkat mekanizmalarının entegrasyonu konusunda ilham vermiştir [162].

Dikkat mekanizmaları; görüntü sınıflandırma, nesne algılama, semantik bölütleme, video anlama, görüntü üretimi, 3 boyutlu görselleştirme, tıbbi görüntü işleme, çok modlu görevler ve kendi kendini denetleyen öğrenme dahil olmak üzere pek çok görsel görevde başarıyla uygulanmaktadır [162].

Derin CNN mimarileri, çok katmanlı veya geniş yapılar nedeniyle yüksek hesaplama maliyetlerine yol açabilir. Bu sorunun üstesinden gelmek amacıyla geliştirilen dikkat mekanizmaları, modelin yalnızca göreve özgü ve anlamlı bilgilere odaklanmasını sağlayarak ilgisiz verilerin etkisini azaltır. Böylece hem hesaplama yükü düşürülür hem de ağırlık genelleme yeteneği iyileştirilir [163].

### 3.6.1. Sıkıştırma ve uyarma ağları (Squeeze-and-excitation networks, SENet)

Hu ve et al. [164], kanal dikkatine dayalı ilk dikkat modülü olan Sıkıştırma ve Uyarma Ağları (Squeeze-and-Excitation Networks, SENet) yapısını önermiştir. Kanal tabanlı bu modül, global bilgileri toplamak, kanal bazında ilişkileri (bağımlılıkları) yakalamak ve temsil yeteneğini geliştirmek için kullanılan bir SE bloğuna dayanmaktadır. SE blokları, Şekil 3.27’de gösterildiği gibi bir sıkıştırma modülü (squeeze module) ve bir uyarma modülü (excitation module) olmak üzere iki bileşenden oluşmaktadır [162], [164]:

- **Sıkıştırma modülü:** Bu modül, her bir kanal için global uzamsal bilgileri toplamak amacıyla global ortalama havuzlama (global average pooling, GAP) işlemini uygular. Bu işlem sonucunda her kanal için bir istatistiksel özet ( $Z \in R^C$ ) çıkarılır.  $z_k$  değeri, k. kanala ait öznitelik haritasının ortalaması olarak Eşitlik 3.6 ile hesaplanır [164]:

$$z_k = F_{sq}(u_k) = \frac{1}{HxW} \sum_{i=1}^H \sum_{j=1}^W u_k(i, j) \quad (3.6)$$

Eşitlik 3.6’da  $F_{sq}$  sıkıştırma işlemi,  $u_k$  ise  $HxW$  uzamsal boyuta sahip k’ncı ( $k=1,2,\dots,C$ ) öznitelik haritasıdır [164].

- **Uyarma modülü:** Kanal bazında ilişkileri yakalamak amacıyla tasarlanmıştır, parametre ve hesaplama sayısını büyük ölçüde azaltabilir. Bu modül, iki FC ve doğrusal olmayan iki aktivasyon katmanı (ReLU ve sigmoid) içerir ve Eşitlik 3.7 ile hesaplanır [164]:

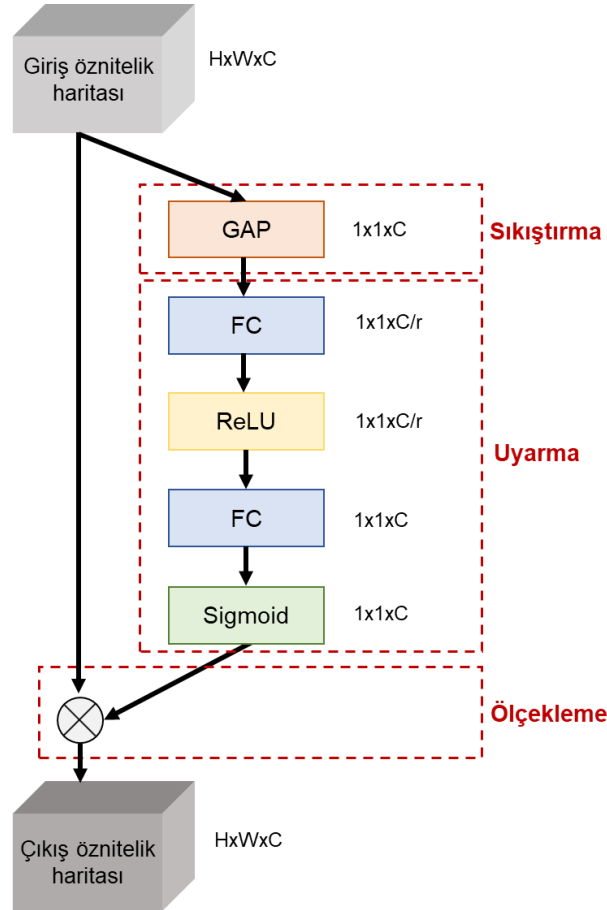
$$S = F_{ex}(Z, W) = \sigma(W_2 \delta(W_1 Z)) \quad (3.7)$$

Eşitlik 3.7’de  $F_{ex}$  uyarma işlemini temsil eder.  $W_1 \in R^{C/r \times C}$  ve  $W_2 \in R^{C \times C/r}$  ağırlıkları, sırasıyla boyut azaltma ve artırma işlemleri için kullanılan ağırlık matrisleri;  $r$  ise kanal azaltma oranı yani kapasiteyi ve hesaplama maliyetini değiştirebilen hiperparametredir.  $\delta$  ve  $\sigma$ , ReLU ve sigmoid aktivasyon fonksiyonlarıdır. Bu işlem sonucunda,  $S = \{s_1, s_2, \dots, s_C\}$  olmak üzere her kanal için dikkat ağırlıkları üretilir [164], [165].

Nihai çıkış ( $\tilde{x}_k (k = 1, 2, \dots, C)$ ), Eşitlik 3.8’de verildiği gibi giriş kanallarının ilgili dikkat ağırlıkları ile çarpılmasıyla elde edilir [162].

$$\tilde{x}_k = F_{scale}(u_k, s_k) = u_k \cdot s_k \quad (3.8)$$

Eşitlik 3.8’de  $\tilde{x}_k \in R^{H \times W}$ ,  $F_{scale}(u_k, s_k)$  ise skaler  $s_k$  ile  $u_k$  öznelik haritası arasındaki kanal bazında çarpmayı ifade eder [162].



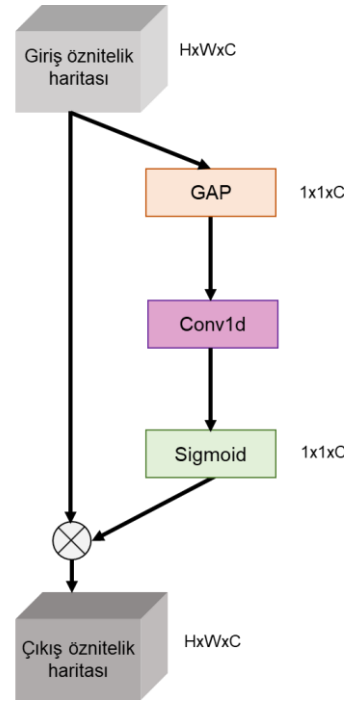
Şekil 3.27. SE blok yapısı [162].

SE blokları, gürültüyü bastırırken önemli kanalları vurgulamada rol oynar. Ancak bazı sınırlamaları da bulunmaktadır. Sıkıştırma modülünde, global ortalama havuzlama,

karmaşık global bilgileri yakalamak için yetersiz kalabilir. Bununla birlikte, uyarma modülündeki tam bağlı katmanlar ise modelin parametre sayısını artırarak hesaplama maliyetini yükseltebilir [162].

### 3.6.2. Verimli kanal dikkati (Efficient channel attention, ECA)

Verimli Kanal Dikkati (Efficient Channel Attention, ECA) modülü, kanal dikkatini modellemek için SE yapısına dayansa da bazı temel farklılıklar içermektedir. SE blokları, kanal bağımlılıklarını modellemek için boyut azaltma işlemleri kullanırken, bu işlem ilave parametre yüküne ve bilgi kaybına neden olabilmektedir. ECA modülü ise boyut indirgeme yerine, kanallar arası etkileşimi doğrudan ve daha verimli bir şekilde modellemeyi hedefler [162], [166]. Şekil 3.28’de ECA mimarisi gösterilmektedir.



Şekil 3.28. ECA mimarisi [162].

İlk olarak, boyut azaltma yapılmadan global öznitelikleri elde etmek için, giriş öznelik haritası üzerinde GAP işlemi gerçekleştirilir [166], [167]. Bu işlem sonucunda, giriş öznelik haritası, Eşitlik 3.9 ile tek boyutlu vektöre ( $1 \times 1 \times C$ ) dönüştürülür [168].

$$z_c = \frac{1}{HxW} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (3.9)$$

Eşitlik 3.9’da  $z_c$  GAP işlemi sonucunda her kanal için elde edilen çıktı öznitelik değerini ifade eder ( $z \in R^C$ ).

GAP işleminin ardından, kanallar-arası etkileşim stratejisini uygulamak üzere çekirdek boyutu  $k$  olan 1-boyutlu evrişim (Conv1d) işlemi gerçekleştirilir. Bu çekirdek boyutu ( $k$ ) parametresi, yerel kanallar arası etkileşimin verimli temsili için uyarlanabilir bir fonksiyon aracılığıyla üretilir [166], [167], [169]. Bu fonksiyon Eşitlik 3.10’da verilmiştir:

$$k = \psi(C) = \left\lfloor \frac{\log_2(C) + b}{\gamma} \right\rfloor_{odd} \quad (3.10)$$

Eşitlik 3.10’da  $C$  kanal sayısını,  $|x|_{odd}$  ise değeri en yakın tek sayıya yuvarlayan işlemi göstermektedir. Bu çalışmada  $\gamma$  ve  $b$  hiperparametreleri sırasıyla 2 ve 1 olarak ayarlanmıştır. Son olarak, kanal ağırlıkları aktivasyon fonksiyonu ile elde edilir ve bu ağırlıklar, giriş öznitelik haritası ile kanal bazında çarpılarak dikkat ağırlıklı çıktı üretilir [166], [167].

### 3.6.3. Evrişimli blok dikkat modülü (Convolutional block attention module, CBAM)

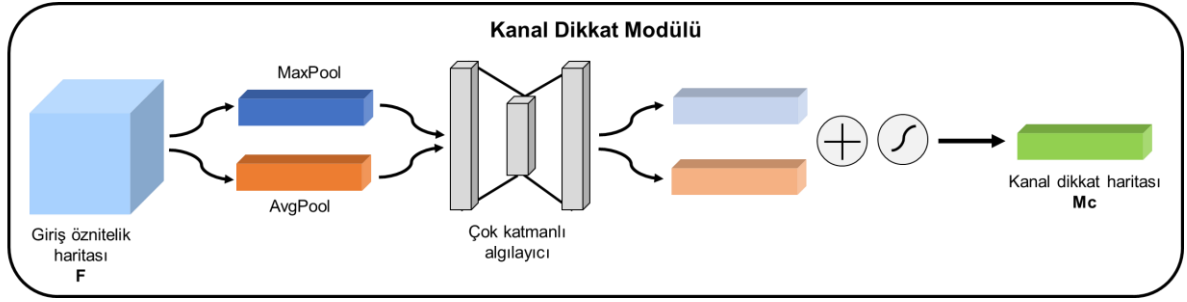
Kanal dikkati ile uzamsal dikkat arasındaki tamamlayıcı özellikler nedeniyle, Woo et al. [170] kanal dikkati ve uzamsal dikkati seri olarak birleştiren Evrişimli Blok Dikkat Modülü (Convolutional Block Attention Module, CBAM) yapısını önermişlerdir. Bu yapı, ağın öznitelik haritalarındaki önemli bilgileri daha iyi yakalayabilmesi için hem kanallar arası hem de uzamsal ilişkileri göz önünde bulundurarak “neye” ve “nereye” odaklanması gerektiğini öğrenmesini sağlar [162].

- **Kanal dikkat modülü:** Woo et al. [170] SENet’te kullanılan yalnızca GAP’a dayalı kanal dikkatinin, ince detayların öğrenilmesinde yetersiz kaldığını belirtmişlerdir. Bu nedenle, GAP’a ek olarak maksimum havuzlama özniteliklerinin de kullanılması gerektiğini önermişlerdir. CBAM’de önerilen kanal tabanlı dikkat modülüne göre, giriş öznitelik haritasındaki ( $F \in \mathbb{R}^{C \times H \times W}$ ) her kanal, görüntünün farklı özniteliklerini içerir ve bu öznitelikler farklı öneme sahiptir. Bu doğrultuda ilk olarak, giriş öznitelik haritası hem ortalama hem de maksimum havuzlama işlemlerine tabi tutularak, uzamsal boyutta sıkıştırılır. Bu işlemler sonucunda aynı boyuta sahip iki kanal öznitelik vektörü elde edilir: ortalama havuzlanmış öznitelik vektörü ( $F_{avg}^C \in \mathbb{R}^{C \times 1 \times 1}$ ) ve maksimum havuzlanmış öznitelik vektörü ( $F_{max}^C \in \mathbb{R}^{C \times 1 \times 1}$ ). Bu iki vektör, yalnızca bir gizli katmana sahip çok katmanlı algılayıcıya

aktarılarak iki yeni kanal öznitelik vektörü elde edilir. Parametreleri azaltmak için, gizli katmanın boyutunu  $\mathbb{R}^{C/r \times 1 \times 1}$  olarak ayarlanır. Burada  $r$ , gizli katmanın küçültme oranını temsil eden bir hiperparametredir ve  $r=16$  olduğunda ağın en iyi performansı elde edebileceği ifade edilmiştir. Son olarak, Şekil 3.29’da gösterildiği gibi kanal dikkat haritasını  $M_c \in \mathbb{R}^{C \times 1 \times 1}$  elde etmek için kanal öznitelik vektörleri eleman bazında toplanır [170], [171]. Dolayısıyla, kanal dikkat haritası Eşitlik 3.11 ile hesaplanabilir:

$$\begin{aligned} M_c(F) &= \sigma \left( MLP(AvgPool(F)) + MLP(MaxPool(F)) \right) \\ &= \sigma \left( W_1 \left( W_0(F_{avg}^c) \right) + W_1 \left( W_0(F_{max}^c) \right) \right) \end{aligned} \quad (3.11)$$

Eşitlik 3.11’de  $\sigma$  sigmoid aktivasyon fonksiyonudur.  $W_0 \in \mathbb{R}^{C/r \times C}$  ve  $W_1 \in \mathbb{R}^{C \times C/r}$  ise MLP ağırlıklarıdır ve iki kanal vektörü için paylaşılır.

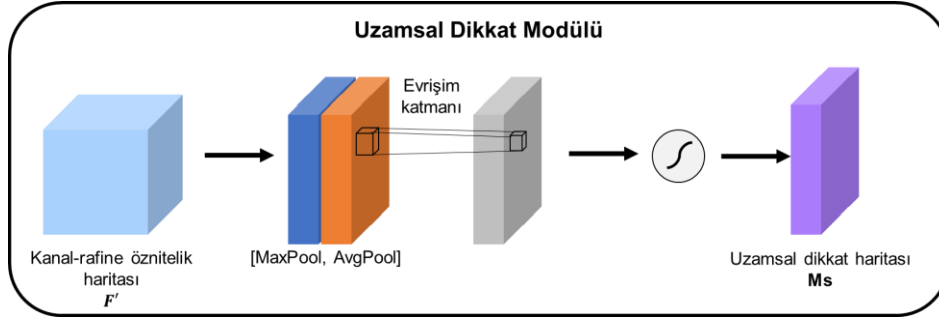


Şekil 3.29. Kanal dikkat modülü yapısı [170].

- **Uzamsal dikkat modülü:** Woo et al. tarafından önerilen uzamsal tabanlı dikkat modülü, öznitelik haritasının konumsal bilgisini dikkate alarak, her konumun görelî önemine göre seçici biçimde artırılmasını veya bastırılmasını sağlar. İlk olarak, giriş öznitelik haritası ( $F \in \mathbb{R}^{C \times H \times W}$ ) kanal boyutunda maksimum havuzlama ve ortalama havuzlama aracılığıyla sıkıştırılır. Bu sıkıştırma işlemi sonucunda, iki adet iki boyutlu (2B) sıkıştırılmış öznitelik haritası ( $F_{avg}^s \in \mathbb{R}^{1 \times H \times W}$  ve  $F_{max}^s \in \mathbb{R}^{1 \times H \times W}$ ) elde edilir. Ardından, iki sıkıştırılmış öznitelik haritası bir evrişim işlemiyle birbirine bağlanır ve sonuçta Şekil 3.30’da gösterildiği gibi 2B uzamsal dikkat haritası ( $M_s(F) \in \mathbb{R}^{1 \times H \times W}$ ) oluşturulur. Uzamsal dikkat haritasının hesaplanması Eşitlik 3.12 ile ifade edilebilir [170], [171]:

$$\begin{aligned} M_s(F) &= \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) \\ &= \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \end{aligned} \quad (3.12)$$

Eşitlik 3.12’de  $\sigma$  sigmoid fonksiyonunu,  $f^{7 \times 7}$  ise filtre boyutu  $7 \times 7$  olan bir evrişim işlemini temsil eder.

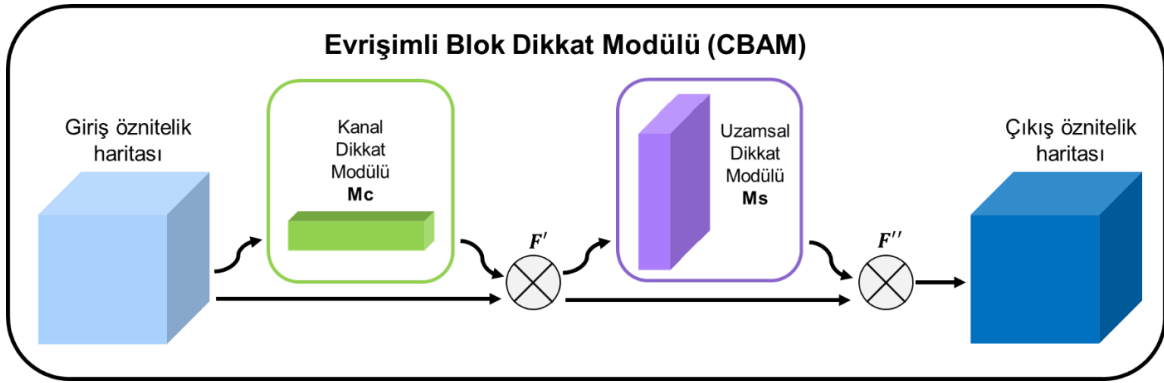


Şekil 3.30. Uzamsal dikkat modülü yapısı [170].

CBAM’ın genel yapısı Şekil 3.31’de gösterilmiştir. İlk olarak, kanal dikkat haritası ( $M_c(F)$ ) ile giriş öznitelik haritası ( $F$ ) eleman bazında çarpılarak ( $\otimes$ ) rafine edilmiş öznitelik haritası ( $F'$ ) elde edilir (Eşitlik 3.12). Ardından bu rafine edilmiş öznitelik haritası ( $F'$ ) ile uzamsal dikkat haritası ( $M_s(F)$ ) çarpılarak, CBAM’ın nihai çıkış öznitelik haritası ( $F''$ ) hesaplanır (Eşitlik 3.13) [170], [171].

$$F' = M_c(F) \otimes F \quad (3.12)$$

$$F'' = M_s(F') \otimes F' \quad (3.13)$$



Şekil 3.31. CBAM genel yapısı [170].

### 3.6.4. Normalizasyon tabanlı dikkat modülü (Normalization-based attention module, NAM)

Liu et al. [172], CBAM yapısını temel alan, kanal ve uzamsal dikkat alt modüllerini yeniden tasarlayarak etkili ve hafif bir dikkat mekanizması olan Normalizasyon Tabanlı Dikkat Modülünü (Normalization-based Attention Module, NAM) önermişlerdir. NAM’da ağırlıkların önem derecelerini belirlemek amacıyla standart sapmaya dayanan bir BN

ölçekleme faktörü kullanılmaktadır [172], [173]. Kanal dikkat alt modülü için kullanılan BN tabanlı ölçekleme faktörü Eşitlik 3.14'te verilmiştir:

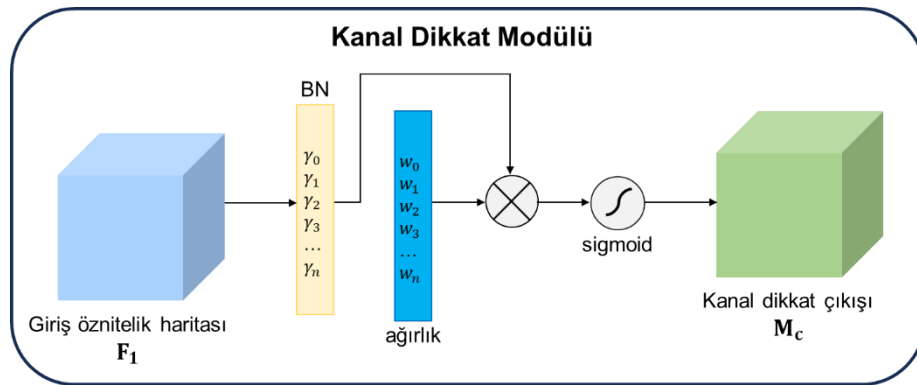
$$B_{\text{çıkış}} = BN(B_{\text{giriş}}) = \gamma \frac{B_{\text{giriş}} - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} + \beta \quad (3.14)$$

Eşitlik 3.14'te  $B_{\text{giriş}}$  ve  $B_{\text{çıkış}}$ , sırasıyla BN'nin giriş ve çıkışlarını ifade eder. Her B mini parti (batch) için  $\mu_B$  ortalamayı ve  $\sigma_B$  standart sapmayı temsil eder.  $\epsilon$  sayısal kararlılık için kullanılan küçük bir sabitken,  $\gamma$  ve  $\beta$  ise sırasıyla eğitilebilir ölçek ve kaydırma parametreleridir. Kanal dikkat alt modülünün yapısı Şekil 3.32'de gösterilmiş olup, Eşitlik 3.15 ve Eşitlik 3.16 ile hesaplanmaktadır:

$$M_c = \text{sigmoid}(W_\gamma(BN(F_1))) \quad (3.15)$$

$$W_\gamma = \frac{\gamma_i}{\sum_{j=0} \gamma_j} \quad (3.16)$$

Eşitlik 3.15'te  $F_1 \in \mathbb{R}^{CxHxW}$  giriş öznetelik haritasını,  $M_c$  kanal dikkat çıkışını,  $\gamma$  BN'den elde edilen ölçekleme faktörünü ve  $W_\gamma$  ise bu faktörlerden türetilen normalize edilmiş ağırlıkları temsil etmektedir. Ölçekleme faktörü ( $\gamma$ ), BN işleminde gözlemlenen varyansa karşılık gelir. Daha yüksek bir varyans, artan değişkenliği ve daha zengin bilgiyi ifade eder.  $\gamma$ 'den türetilen normalleştirilmiş korelasyon ağırlıkları ( $W_\gamma$ ) kullanılarak, kanal dikkat modülü bilgi açısından zayıf kanalları bastırırken anlamlı kanalları vurgular [174].



Şekil 3.32. NAM kanal dikkat modülü yapısı [172].

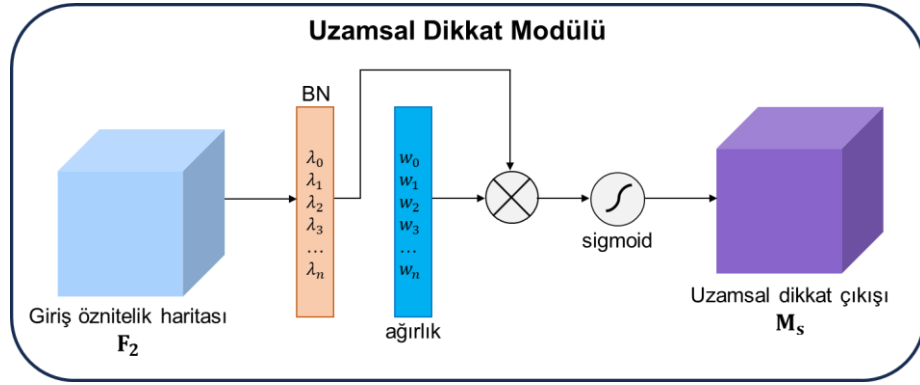
Uzamsal dikkat modülünde, piksellerin önem düzeyini belirlemek amacıyla uzamsal boyuttaki piksellere BN ölçekleme faktörü uygulanır. Bu yaklaşım piksel normalizasyonu

olarak adlandırılır [172]. Uzamsal dikkat alt modülünün yapısı Şekil 3.33'te gösterilmiş olup, Eşitlik 3.17 ve Eşitlik 3.18 ile hesaplanmaktadır:

$$M_s = \text{sigmoid}(W_\lambda(BN_s(F_2))) \quad (3.17)$$

$$W_\lambda = \frac{\lambda_i}{\sum_{j=0} \lambda_j} \quad (3.18)$$

Eşitlik 3.17'de  $M_s$  uzamsal dikkat çıkışı,  $\lambda$  piksel düzeyinde BN'den elde edilen ölçekleme faktörünü ve  $W_\lambda$  ise buna dayalı normalize edilmiş ağırlıkları ifade etmektedir. Uzamsal dikkat modülü, BN'den elde edilen ölçekleme faktörüne ( $\lambda$ ) göre daha bilgilendirici piksellere öncelik verir ve karşılık gelen korelasyon ağırlıklarını ( $W_\lambda$ ) ayarlar. Benzer şekilde, her piksele atanan ağırlık değeri giriş öznitelik haritasına ( $F_2 \in \mathbb{R}^{C \times H \times W}$ ) uygulanır ve sigmoid aktivasyon fonksiyonunun kullanımıyla uzamsal dikkat çıkışı elde edilir [174].



Şekil 3.33. NAM uzamsal dikkat modülü yapısı [172].

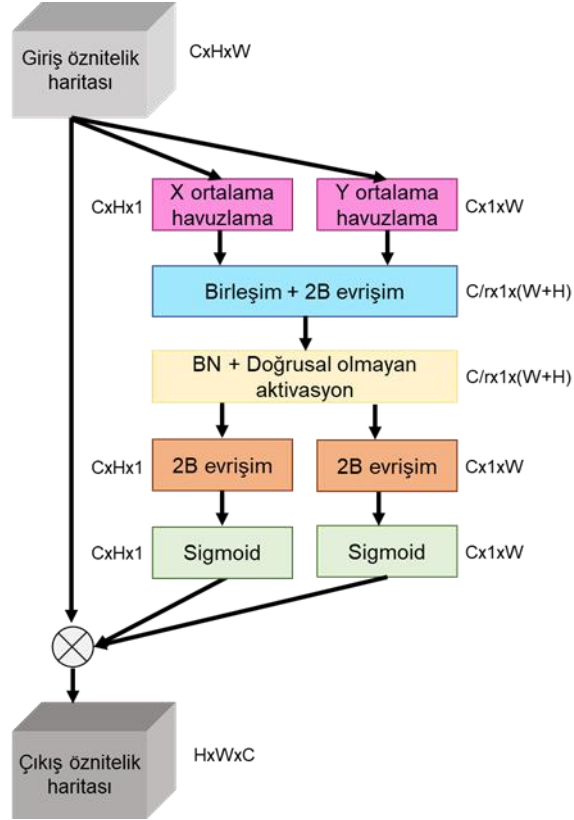
NAM'da, önemsiz ağırlıkları bastırmak için Eşitlik 3.19'da gösterildiği gibi kayıp fonksiyonuna bir düzenleme (regularization) terimi eklenmiştir [172].

$$Kayıp = \sum_{(x,y)} l(f(x, W), y) + p \sum g(\gamma) + p \sum g(\lambda) \quad (3.19)$$

Eşitlik 3.19'da  $x$  girdi,  $y$  çıktı ve  $W$  ise model ağırlıklarıdır.  $l(\cdot)$  kayıp fonksiyonunu,  $g(\cdot)$  L1 normuna dayalı ceza fonksiyonunu,  $p$  ise  $g(\gamma)$  ve  $g(\lambda)$  terimlerini dengeleyen ceza katsayısını belirtir. Bu yaklaşım, modelin anlamlı öznitelikleri daha etkili şekilde öğrenmesini sağlarken, genelleme yeteneğini ve genel tespit performansını artırabilir [173].

### 3.6.5. Koordinat dikkat (Coordinate attention, CA) modülü

SENet, kanallar arası ilişkileri (bağımlılıkları) başarılı bir şekilde yakalarken uzamsal (konumsal) bilgiyi göz ardı etmektedir. CBAM ise uzamsal bilgiyi yakalamak için evrişim çekirdeklerini kullanır, ancak uzun menzilli bağımlılıkları modellemede yetersiz kalmaktadır. Bu dezavantajların üstesinden gelmek için uzamsal bilginin kanal dikkatine entegre edildiği Koordinat Dikkat (Coordinate Attention, CA) modülü geliştirilmiştir [175]. CA mimarisi Şekil 3.34'te gösterilmektedir.



Şekil 3.34. CA mimarisi [175].

CA modülü,  $CxHxW$  boyutundaki giriş öznelik haritası (X) üzerinde, her kanal için hem yatay hem de dikey yönde ortalama havuzlama uygular. Bu işlemler sonucunda, her kanal için biri yatay ekseninde ( $Hx1$ ), diğeri ise dikey ekseninde ( $1xW$ ) olmak üzere iki ayrı öznelik gösterimi elde edilir [167], [176]. Bu işlemler, Eşitlik 3.20 ve Eşitlik 3.21 ile hesaplanır:

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(h, i) \quad (3.20)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq i \leq H} x_c(j, w) \quad (3.21)$$

Eşitlik 3.20 ve Eşitlik 3.21’de  $z^h$  ve  $z^w$ , sırasıyla yatay ve dikey yönde ortalama havuzlama çıktılarıdır.  $z_c^h(h)$  ve  $z_c^w(w)$  ise  $c$  numaralı kanalın  $h$ . satır ve  $w$ . sütundaki öznitelik değerlerini temsil eder [167].

Havuzlama işlemi sonrası elde edilen öznitelik haritaları birleştirilir ve ardından 1x1 boyutlu standart bir evrişim ( $F_1$ ) uygulanır. Çıktı, doğrusal olmayan h-swish ( $\delta$ ) aktivasyon fonksiyonu aracılığıyla işlenerek her iki yöndeki uzamsal bilgiyi temsil eden ara öznitelik haritası ( $f$ ) Eşitlik 3.22 ile elde edilir [175], [177].

$$f = \delta(F_1([z^h, z^w])) \quad (3.22)$$

Bu ara öznitelik haritası ( $f$ ), uzamsal boyutta ikiye bölünerek  $f^h$  ve  $f^w$  olmak üzere iki ayrı öznitelik haritası oluşturulur. Daha sonra bu haritalar, 1x1 boyutlu iki adet evrişim ( $F_h$  ve  $F_w$ ) işlemi aracılığıyla işlenerek giriş öznitelik haritasının kanal sayısına göre genişletilir. Son olarak, sigmoid aktivasyon fonksiyonu ile dikkat ağırlıkları üretilir [175], [176], [177]. Hesaplamalar Eşitlik 3.23 ve Eşitlik 3.24’te gösterilmiştir.

$$g^h = \sigma(F_h(f^h)) \quad (3.23)$$

$$g^w = \sigma(F_w(f^w)) \quad (3.24)$$

Eşitlik 3.23 ve Eşitlik 3.24’te  $g^h$  ve  $g^w$ , sırasıyla yatay ve dikey yönlerdeki koordinat dikkat ağırlıklarıdır. Son olarak, Eşitlik 3.25’te gösterildiği gibi bu dikkat ağırlıkları giriş öznitelik haritası ile kanal bazında çarpılması sonucu nihai çıktı ( $y_c(i, j)$ ) elde edilir [176].

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (3.25)$$

### 3.6.6. Verimli çok-ölçekli dikkat (Efficient multi-scale attention, EMA) modülü

Verimli Çok-Ölçekli Dikkat (Efficient Multi-scale Attention Module, EMA) modülü, çapraz-uzamsal öğrenmeye dayanan bir dikkat modülüdür. Boyut azaltmaya ihtiyaç duymadan, grublama temelli bir yapı aracılığıyla çok ölçekli dikkat işlemini verimli biçimde gerçekleştirebilir [178].

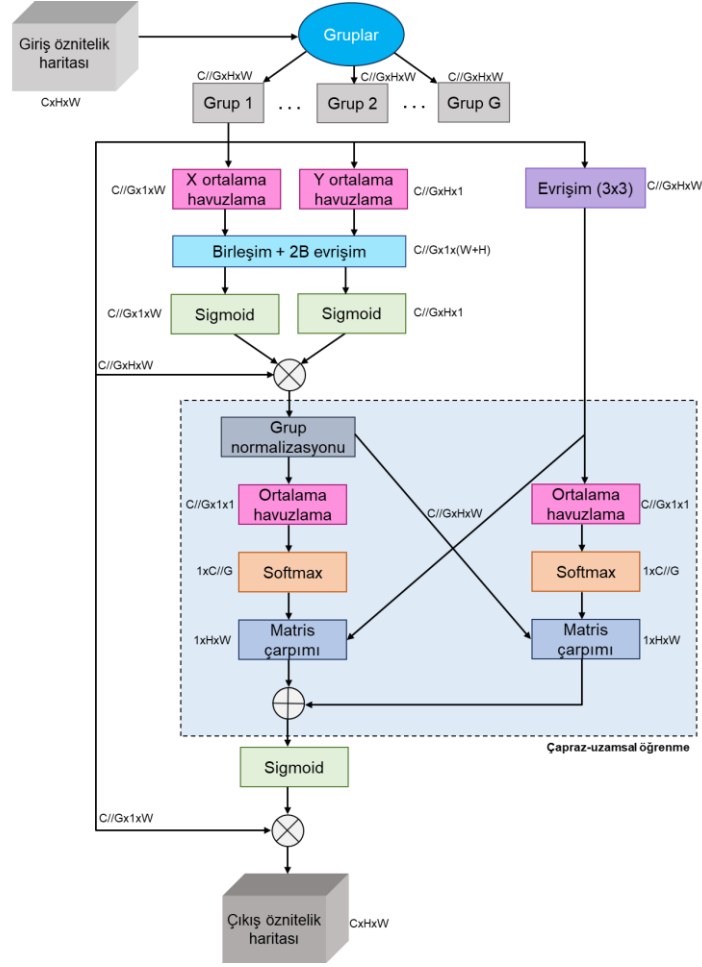
EMA modülü yapısı Şekil 3.35’te gösterilmektedir. Giriş öznitelik haritası ( $X \in \mathbb{R}^{C \times H \times W}$ ) kanal boyutu boyunca  $G$  gruba ayrılır ve böylece farklı anlamsal özelliklerin

yakalanması için birden fazla alt öznitelik haritası oluşturulur. Her bir grup  $X = [X_1, X_2, \dots, X_{G-1}]$ ,  $X_i \in \mathbb{R}^{C//G \times H \times W}$  şeklinde ifade edilir. EMA, bu gruplanmış öznitelik haritaları için dikkat ağırlıklarını çıkarmak üzere üç paralel dal kullanır: bunlardan ikisi  $1 \times 1$ , biri ise  $3 \times 3$  evrişim dalıdır.

$1 \times 1$  dallarında, öznitelikler hem yatay hem de dikey yönde kodlanarak, her ekseninde 1-boyutlu (1B) ortalama havuzlama işlemine tabi tutulur. Elde edilen öznitelikler birleştirilir ve  $1 \times 1$  evrişimle işlenerek sigmoid aktivasyon fonksiyonundan geçirilen iki ayrı tek boyutlu dikkat vektörü oluşturulur. Bu vektörler, her kanal grubu için dikkat haritalarının hesaplanmasında kullanılır ve kanallar arası öznitelik etkileşimini kolaylaştırır.

$3 \times 3$  dalında ise, çok ölçekli uzamsal bilgileri genişletmek için  $3 \times 3$  evrişim çekirdeği kullanılır. Sonuç olarak, Böylece yalnızca kanal önem dereceleri belirlenmekle kalmaz, aynı zamanda kanallar içindeki hassas yapısal ayrıntılar da korunmuş olur [178], [179], [180].

Ek olarak EMA, öznitelik etkileşimlerini işlemek için çapraz-uzamsal öğrenme (cross-spatial learning) stratejisini benimser. Bu kapsamda,  $1 \times 1$  dalında global uzamsal bilgiyi kodlamak ve uygun boyuta dönüştürmek amacıyla iki boyutlu 2B global ortalama havuzlama işlemi uygulanır. Aynı zamanda,  $3 \times 3$  dalından elde edilen öznitelikler ortak bir aktivasyon mekanizması içerisinde değerlendirilir. Paralel işleme dallarından elde edilen çıktılar, matris nokta çarpımı (dot product) yöntemiyle birleştirilerek tek bir uzamsal dikkat haritası oluşturulur. Böylece, farklı ölçeklerdeki uzamsal bilgilerin aynı aşamada etkili biçimde bütünleştirilmesi sağlanır. Buna ek olarak,  $3 \times 3$  dalı üzerinden ikinci bir uzamsal dikkat haritası daha üretilerek uzamsal bilginin korunması desteklenir. Son adımda ise, her grup için çıktı öznitelik haritasını elde etmek üzere bu iki uzamsal dikkat haritası, sigmoid aktivasyon fonksiyonu aracılığıyla birleştirilerek son öznitelik eşlemesi gerçekleştirilir [179], [180].

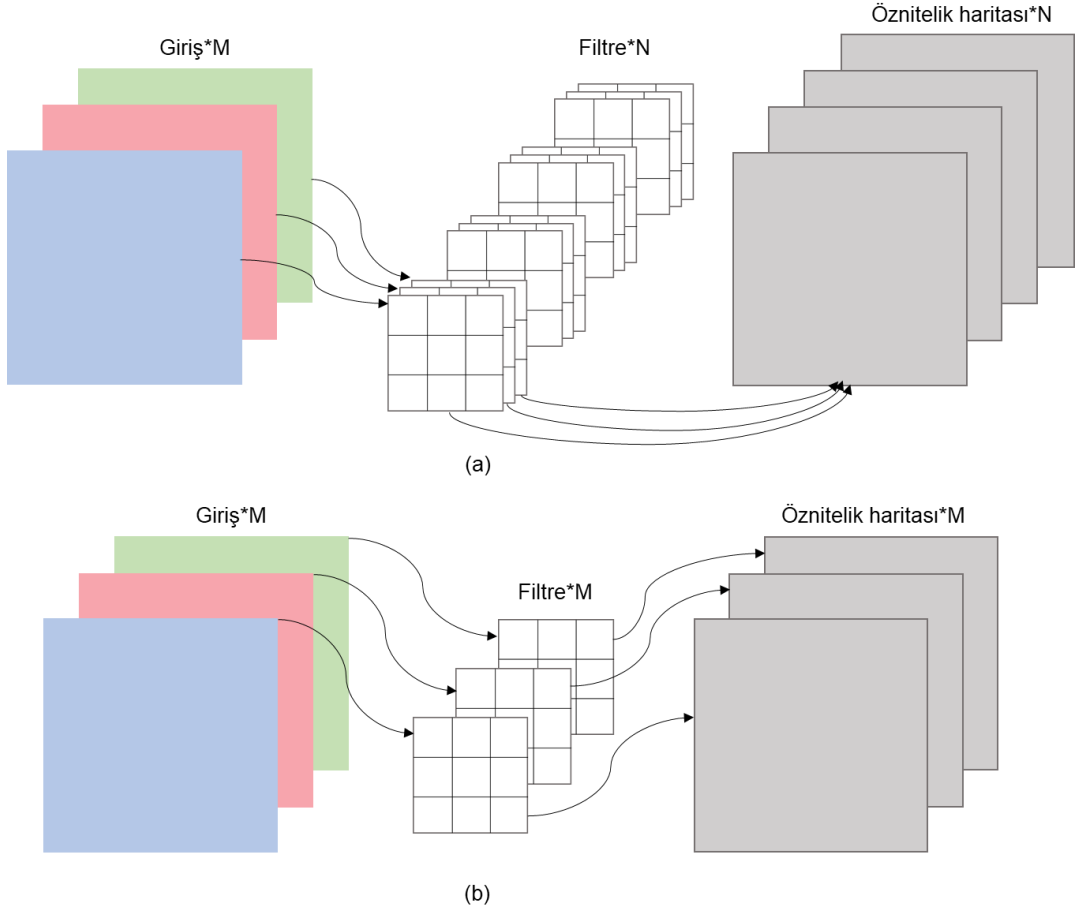


Şekil 3.35. EMA mimarisi [178].

### 3.7. Farklı Evrişim Modülleri

#### 3.7.1. Derinlemesine evrişim (Depth-wise convolution)

Derinlemesine evrişim (Depth-wise convolution, DWConv), standart evrişim işlemine kıyasla daha az hesaplama maliyeti ve parametre sayısına sahip bir yöntemdir. Şekil 3.36'da gösterildiği gibi, standart evrişimde her bir filtre tüm giriş kanallarına uygulanarak bir çıktı öznelik haritası oluşturulur. Buna karşılık, DWConv'da giriş öznelik haritasının her kanalına yalnızca bir adet filtre uygulanır. Bu yapıda, filtre sayısı giriş kanalı sayısıdır, bu nedenle kanal sayısı korunur [181], [182]. Örneğin, giriş kanal sayısı  $M$ , filtre boyutu  $k \times k$  ve çıkış kanal sayısı  $N$  olmak üzere; toplam parametre sayısı standart evrişimde  $M \cdot N \cdot k \cdot k$  iken, DWConv'da ise  $M \cdot k \cdot k$  olmaktadır. Bu durum, hesaplama maliyetinde ve parametre sayısında yaklaşık olarak  $1/N$  oranında bir azalma anlamına gelir.

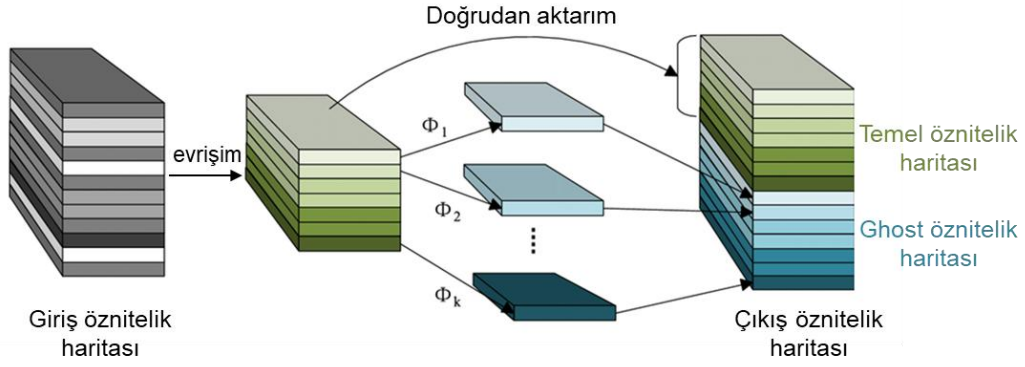


Şekil 3.36. (a) Standart evrişim, (b) Derinlemesine evrişim [181].

### 3.7.2. Ghost evrişim (Ghost convolution)

Ghost evrişim (Ghost convolution, GhostConv), Noah's Ark Lab tarafından tasarlanan hafif bir ağ olan GhostNet'in evrişimsel bir modülüdür [183]. Standart evrişim ile karşılaştırıldığında, GhostConv çıktı öznitelik haritasının boyutunu korurken, parametre sayısını ve hesaplama karmaşıklığını azaltmaktadır [184].

GhostConv modül yapısı Şekil 3.37'de gösterilmektedir. Bu yapıda, giriş öznitelik haritası iki aşamada işlenmektedir: İlk olarak, standart bir evrişim işlemi uygulanarak temel (işsel) öznitelik haritaları oluşturulur ve bu haritalar doğrudan çıktı olarak kullanılır. Ardından, bu temel öznitelik haritaları üzerinde bir dizi hafif doğrusal dönüşümler ( $\phi$ ) gerçekleştirilerek ek (ghost) öznitelik haritaları üretilir. Son öznitelik haritası, temel öznitelik haritaları ve ghost öznitelik haritalarının birleştirilmesi ile elde edilir [43], [183].



Şekil 3.37. Ghost evrişim modülü [144].

Standart bir evrişim işlemi,  $X \in \mathbb{R}^{c \times h \times w}$  giriş öznelik haritasını temsil etmek üzere Eşitlik 3.26 ile tanımlanır:

$$Y = X * f + b \quad (3.26)$$

Eşitlik 3.26'da  $*$  evrişim operatörünü,  $b$  sapma (bias) terimini,  $X \in \mathbb{R}^{c \times h' \times w'}$   $n$  adet kanal içeren çıktı öznelik haritasını ve  $f \in \mathbb{R}^{c \times k \times k \times n}$  ise bu katmanda kullanılan evrişim çekirdeğini ifade eder. Burada,  $h'$  ve  $w'$  çıktı haritasının yüksekliğini ve genişliğini,  $k \times k$  çekirdek boyutunu tanımlar. Bu bilgiler doğrultusunda, bias terimi hariç tutulduğunda standart bir evrişim işleminin hesaplama maliyeti " $n \times h' \times w' \times c \times k \times k$ " şeklinde ifade edilir. Özellikle ağırlık katmanlarında  $h'$  ve  $w'$  daha büyükken, derin katmanlarda ise  $n$  ve  $c$  daha büyük değerlere sahiptir. Bu nedenle, Ghost evrişim kavramı bu maliyeti azaltmak için geliştirilmiştir. Ghost evrişim yapısı, temel (içsel) öznelik haritaları üreten Eşitlik 3.27 ve ek (ghost) öznelik haritalarını üreten Eşitlik 3.28 ile tanımlanır [144], [185]:

$$Y' = X * f' \quad (3.27)$$

$$Y_{ghost} = \phi_{i,j}(y_i') \quad (3.28)$$

Eşitlik 3.27'de  $Y' \in \mathbb{R}^{h' \times w' \times m}$  temel (içsel) öznelik haritasını,  $f' \in \mathbb{R}^{c \times k \times k \times m}$  ise evrişim çekirdeğinin boyutunu temsil eder. GhostConv katmanında elde edilen çıktı kanal sayısı, standart evrişim katmanındakinden daha azdır; yani  $m < n$  koşulu sağlanır. Eşitlik 3.28 ise ek (ghost) öznelik haritalarını üreten doğrusal dönüşüm katmanını temsil eder. Burada,  $y_i'$   $Y'$  içindeki  $m$  adet öznelik haritasından birini belirtir.  $Y'$  içindeki her öznelik haritası,  $s$  öznelik haritası üreten hafif bir doğrusal dönüşüm  $\phi_{i,j}$  ( $j = 1, 2, \dots, s$ ) işlemine tabi tutulur. Eğer doğrusal dönüşüm işlemi olarak evrişim kullanılırsa, bu dönüşüm sabit bir

yapı olarak tanımlanır. Bu durumda, her biri temel öznitelik haritasından türetilmiş olan toplam  $m \times (s - 1)$  adet ek (ghost) öznitelik haritası üretilir ve son aşamada tekrar  $m$  öznitelik haritasına indirgenir. Bu işlem sonucunda Ghost evrişiminin hesaplama maliyeti “ $(s - 1) \times m \times h' \times w' \times k \times k$ ” şeklinde ifade edilir [144], [185].

### 3.8. Transfer Öğrenme

Transfer öğrenme, bir görev için edinilen bilginin benzer başka bir görevi çözmek için kullanılma sürecini ifade eder [186]. YOLO gibi derin öğrenme modelleri, optimum performansa ulaşmak için genellikle büyük miktarda etiketli eğitim verisi gerektirir. Ancak, birçok tıbbi görüntü veri kümesinin sınırlı boyutu, etkili eğitim için önemli bir zorluk oluşturur. Bu sınırlamanın üstesinden gelmek amacıyla, transfer öğrenme yaklaşımları bir çözüm olarak önerilmiştir [130], [187]. İlgili bir kaynak veri kümesinde önceden eğitilmiş bir model, hedef veri kümesine ince ayar (fine-tuning) yöntemi ile uyarlanır. İnce ayar, özellikle etiketli verinin sınırlı olduğu durumlarda, modeli sıfırdan eğitmeye kıyasla göre daha etkili ve avantajlıdır [188].

Transfer öğreniminin başarısı büyük ölçüde kaynak ve hedef veri kümesi arasındaki benzerliğe bağlıdır. Bu iki veri kümesi ortak görsel veya anlamsal özellikler taşıyorsa, aktarılan bilginin hedef görevde model performansını artırma olasılığı yüksektir [189].

### 3.9. Performans Değerlendirme Metrikleri

#### 3.9.1. Nesne tespitinde kullanılan performans metrikleri

Nesne tespit modellerinin performansını değerlendirmek amacıyla yaygın olarak kullanılan metrikler; kesinlik (Precision, P), duyarlılık (Recall, R), ortalama kesinlik (Average Precision, AP) ve tüm sınıflar için ortalama kesinlik (mean Average Precision, mAP) şeklinde sıralanabilir [97].

Kesinlik ve duyarlılık performans metriklerinin hesaplanabilmesi için öncelikle aşağıdaki temel kavramların tanımlanması gerekmektedir [190]:

- **Doğru Pozitif (DP):** Gerçek sınırlayıcı kutunun model tarafından doğru şekilde tespit edilmesidir.
- **Yanlış Pozitif (YP):** Gerçek sınırlayıcı kutuya karşılık gelmeyen bir nesnenin yanlışlıkla tespit edilmesi veya mevcut bir nesnenin yanlış konumda tespit edilmesidir.
- **Yanlış Negatif (YN):** Modelin, mevcut bir gerçek sınırlayıcı kutuyu tespit

edememesi durumudur.

- **Doğru Negatif (DN):** Bu kavram genellikle karışıklık matrisi kapsamında kullanılan bir metriktir. Ancak nesne tespiti görevlerinde, bir görüntüde tespit edilmemesi gereken sonsuz sayıda sınırlayıcı kutu olabileceği için, DN metriği performans değerlendirme hesaplamalarına dahil edilmemektedir.

*Bir tespitin doğru ya da yanlış olduğuna karar vermek için IoU değeri kullanılır*

**Şekil 3.11.** IOU değeri, önceden tanımlanmış bir eşik değeri ( $t$ ) ile karşılaştırılır.  $IoU \geq t$  olduğunda tespit doğru pozitif (DP),  $IoU < t$  olduğunda ise tespit yanlış pozitif (YP) olarak kabul edilir [107].

Kesinlik, modelin pozitif olarak tespit ettiği nesnelere arasında gerçekten doğru olanların oranını ifade eder. Başka bir deyişle, modelin doğru tespit etme yüzdesini gösterir ve Eşitlik 3.29 ile hesaplanır [98], [187].

$$Kesinlik = \frac{DP}{DP + YP} = \frac{DP}{Tüm\ tespitler} \quad (3.29)$$

Duyarlılık, modelin tespit etmesi gereken tüm gerçek nesnelere arasından doğru şekilde bulabildiği nesnelere oranını ifade eder. Modelin gerçek pozitifleri tespit etme yüzdesini gösterir ve Eşitlik 3.30 ile hesaplanır [98], [187].

$$Duyarlılık = \frac{DP}{DP + YN} = \frac{DP}{Tüm\ gerçekler} \quad (3.30)$$

Ortalama kesinlik (AP), farklı eşik değerleri için oluşturulan kesinlik-duyarlılık (Precision-Recall, PR) eğrisinin altında kalan alanı (Area Under Curve, AUC) temsil eder. mAP ise her nesne kategorisi için hesaplanan AP değerlerinin ortalaması olarak tanımlanır ve burada “ $n$ ” toplam kategori sayısını ifade eder. Bununla birlikte, mAP bir veri kümesindeki tüm sınıflar üzerinde nesne tespit modellerinin genel doğruluğunu ölçmek için kullanılır. mAP değeri ne kadar yüksek olursa, modelin tespit ve sınıflandırma performansı o kadar iyi kabul edilir [112], [130], [190]. AP ve mAP metrikleri, sırasıyla Eşitlik 3.31 ve Eşitlik 3.32 ile hesaplanır.

$$AP = \int_0^1 P(R) dR \quad (3.31)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \quad (3.32)$$

mAP hesaplaması için farklı IoU eşik değerleri kullanılmaktadır. En yaygın olarak uygulanan IoU eşiği 0,5'tir ve mAP@0.5 şeklinde gösterilir [40].

### **3.9.2. *k*-katlı çapraz doğrulama**

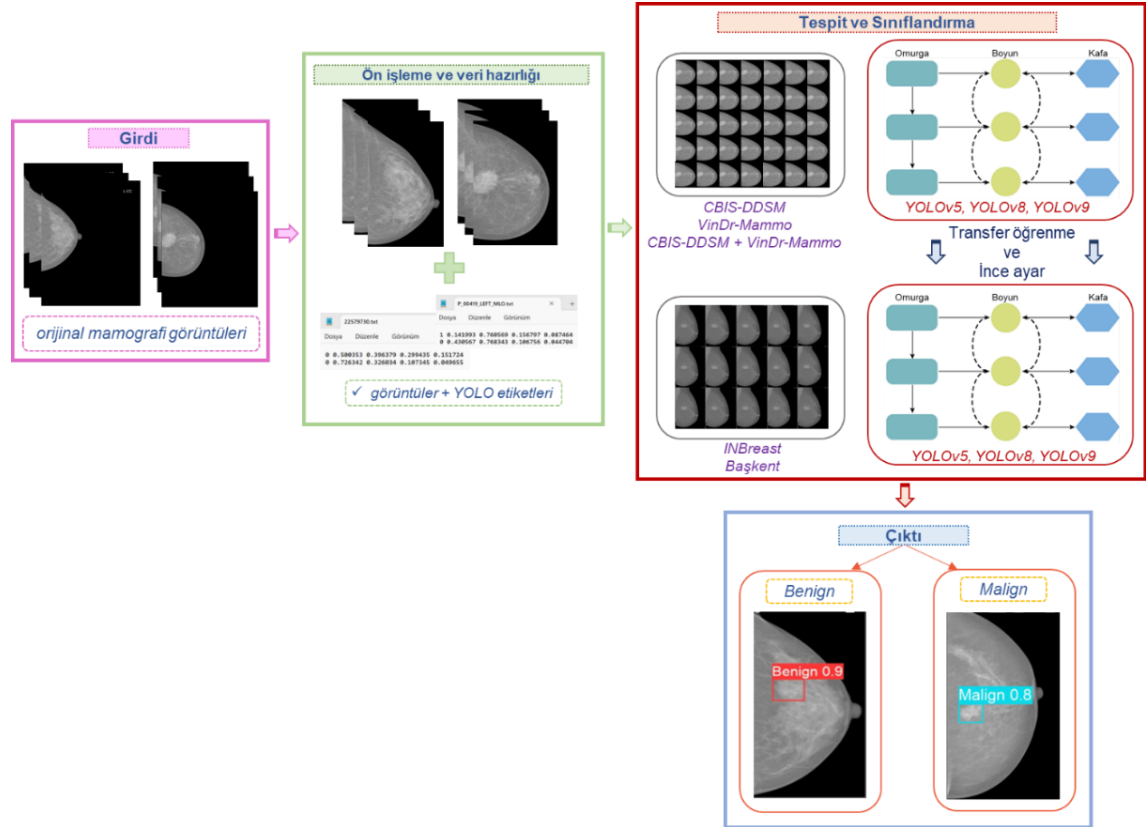
*k*-katlı çapraz doğrulama, derin öğrenme modellerinin performansını etkili bir şekilde değerlendirmek amacıyla yaygın olarak kullanılan bir yöntemdir. Bu yöntem, verilerin daha verimli kullanılmasını sağlar ve eşit olmayan veri bölümlendirmelerinden kaynaklanan önyargıyı azaltır. Ayrıca, modelin genelleme yeteneğini, yani daha önce görülmemiş verilere uyum sağlama başarısını değerlendirir [191].

*k*-katlı çapraz doğrulama yönteminde, veri kümesi rastgele seçilerek eşit büyüklükte *k* parçaya bölünür. Her adımda bir parça test verisi olarak seçilirken, kalan *k* – 1 parça eğitim verisi olarak kullanılır. Bu işlem, her parça bir kez test kümesi olarak kullanılana kadar tekrarlanır. Böylece tüm veri kümesi hem eğitim hem de test aşamalarında değerlendirilmiş olur. Nihai model performansı ise tüm katlardan elde edilen sonuçların ortalaması alınarak hesaplanır [192], [193].

## 4. DENEYSEL ANALİZLER VE SONUÇLAR

### 4.1. Tez Kapsamında Önerilen Yaklaşım

Tez çalışması kapsamında, meme kanseri tespiti ve sınıflandırmasına yönelik olarak önerilen yaklaşımın adımları Şekil 4.1’de sunulmuştur.



Şekil 4.1. Tez çalışmasına ait önerilen yaklaşım.

Çalışmada ilk olarak açık erişimli veri kümelerine erişilmiş, ardından bu görüntüler üzerinde gerekli veri dönüşümleri, ön işleme adımları ve YOLO algoritması için etiketleme işlemleri gerçekleştirilmiştir. Tespit ve sınıflandırma sürecinde transfer öğrenme yaklaşımı benimsenmiştir. Bu doğrultuda, herhangi bir hazır ön eğitilmiş ağırlık kullanılmadan, CBIS-DDSM, VinDr-Mammo ve bu iki veri kümesinin birleşimi olan CBIS-DDSM + VinDr-Mammo kaynak veri kümeleri olarak değerlendirilmiş ve farklı YOLO modelleri bu veri kümelerinde rastgele başlatılan ağırlıklarla sıfırdan eğitilmiştir. Ardından, bu aşamada elde edilen en iyi model ağırlıkları hem INBreast veri kümesine hem de klinik mamografi veri kümesine aktarılmış ve her iki veri kümesinde ayrı ayrı ince ayar işlemi gerçekleştirilmiştir. Ayrıca çalışmada, farklı dikkat modülleri ve evrişim bloklarının entegrasyonu ile mimari iyileştirmeler gerçekleştirilmiş; bunun yanında çeşitli veri artırma parametrelerinin model

performansına etkisi de kapsamlı olarak incelenmiştir. Gerçekleştirilen tüm deneysel analizler sonucunda, mamografi görüntülerinden kitle tespiti ve sınıflandırması amacıyla en yüksek performansı gösteren model nihai model olarak belirlenmiştir.

## **4.2. Veri Kümelerinin Analize Hazırlanması**

Veri kümelerinin analize hazırlanması aşamasındaki ilk adım, tüm görüntülerin aynı formatta standartlaştırılmasıdır. Bu nedenle, analiz öncesinde tüm mamografi görüntüleri .png formatına dönüştürülmüştür. Bu bölümde açıklanan tüm işlemler, MATLAB ortamında gerçekleştirilmiştir.

### **4.2.1. Ön işleme adımları**

Ön işlemenin temel amacı, önemli tanısal özellikleri koruyarak görüntü kalitesini iyileştirmek, modelin performansını olumsuz etkileyebilecek artefaktları ve gürültüyü ortadan kaldırmaktır. Bu doğrultuda aşağıdaki adımlar uygulanmıştır:

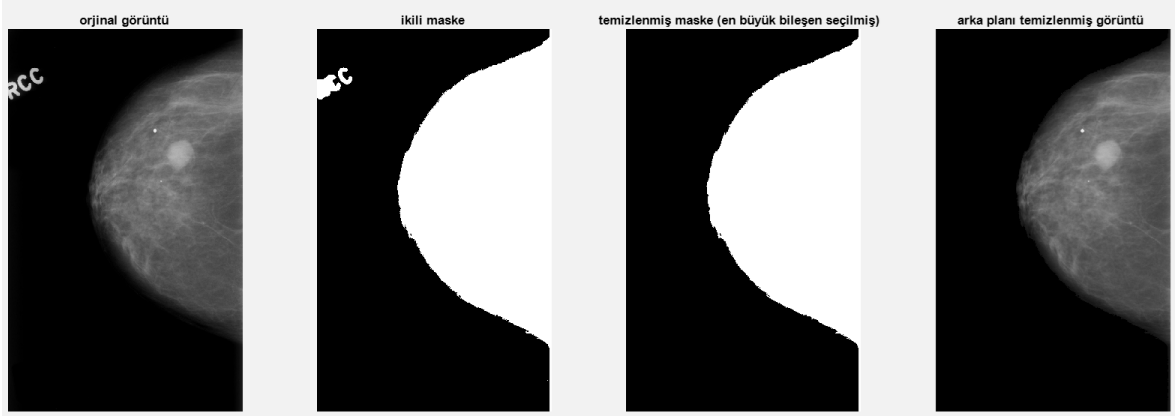
#### **4.1.1.a. Arka plan artefaktlarının kaldırılması**

Mamografi görüntülerinde yer alan tıbbi etiketler, projeksiyon bilgileri (CC, MLO) ve çekime özgü diğer artefaktlar, modellerin performansını olumsuz etkileyebilmektedir. Bu nedenle, giriş görüntüsündeki bu tür istenmeyen yapıları temizlemek amacıyla aşağıdaki işlem adımları uygulanmıştır:

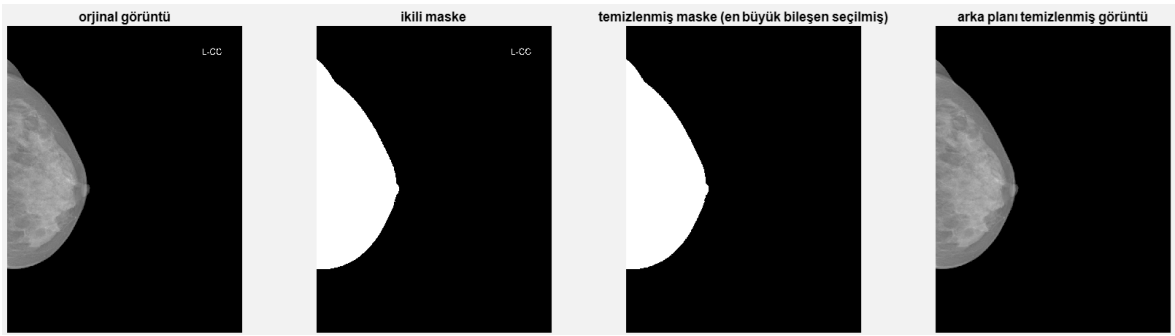
- İlk olarak, gri seviyeli giriş görüntüsü ikili (siyah-beyaz) formata dönüştürülmüştür.
- Ardından, ikili görüntüde beyaz bölgelerin alanları tespit edilmiş ve yalnızca en büyük bileşen (genellikle meme) korunarak, geri kalan artefakt bölgeleri kaldırılmıştır.
- Son olarak, arka plan artefaktları temizlenen ikili görüntü tekrar gri seviyeye dönüştürülerek nihai görüntü elde edilmiştir.

Bu işlem, yalnızca CBIS-DDSM ve VinDr-Mammo veri kümelerine uygulanmıştır. INBreast ve klinik veri kümesinde görüntülerde belirgin artefakt bulunmadığından bu adıma ihtiyaç duyulmamıştır.

CBIS-DDSM ve VinDr-Mammo veri kümeleri için uygulanan arka plan artefaktlarını temizleme adımlarına ait örnekler Şekil 4.2 ve Şekil 4.3'te gösterilmiştir.



Şekil 4.2. CBIS-DDSM veri kümesine ait bir görüntüde arka plan temizleme adımları.

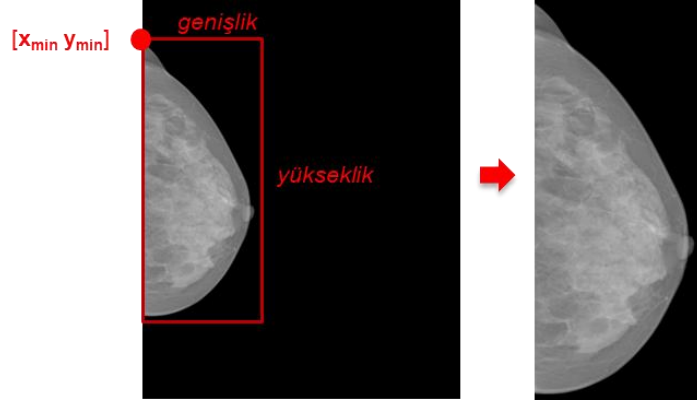


Şekil 4.3. VinDr-Mammo veri kümesine ait bir görüntüde arka plan temizleme adımları.

#### 4.1.1.b. Meme bölgesinin kırılması

Arka plan artefaktları temizlenmiş görüntülerde, geniş siyah alanların analizi olumsuz etkilemesini önlemek ve yalnızca meme dokusunu içeren bölgeyi çalışmaya dahil etmek amacıyla görüntülerin kırılması gerekmektedir.

YOLO formatında etiketleme sürecinde kullanılan kitle anotasyonları, görüntülerin orijinal boyutlarına göre tanımlanmıştır. Ancak kırpma işlemi sırasında görüntü boyutları değiştiği için, bu anotasyonların doğruluğunu koruyabilmek adına kırpma kutusuna ait koordinat bilgilerinin kaydedilmesi gerekmektedir. Görüntü kırpma işlemleri MATLAB ortamında imcrop fonksiyonu kullanılarak gerçekleştirilmiştir. Bu fonksiyonda kullanılan “rect” parametresi, kırılacak bölgenin orijinal görüntü üzerindeki konumunu ve boyutunu ( $x_{min}$ ,  $y_{min}$ , genişlik, yükseklik) tanımlar. Kırpma sırasında elde edilen bu bilgiler, kitle koordinatlarının kırılmış görüntü boyutlarına göre yeniden hesaplanmasında kullanılacaktır. Şekil 4.4’te, kırpma işlemi için kullanılan rect parametresinin görüntü üzerindeki temsili gösterimi ile bu işlem sonucunda elde edilen görüntü sunulmuştur.



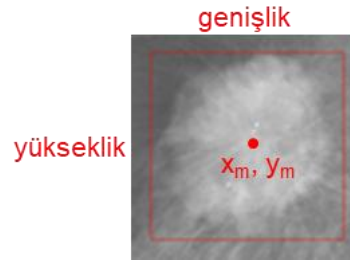
Şekil 4.4. Kırma işleminin görüntü üzerindeki temsili ve kırma sonrası elde edilen görüntü.

#### 4.2.2. Veri kümelerinin YOLO formatında etiketlenmesi

YOLO modellerinin eğitilmesi için ilk adım, veri kümesinin YOLO formatına uygun şekilde etiketlenmesidir. Bu formatta, her görüntüde yer alan nesnelere için sınırlayıcı kutular tanımlanır ve bu bilgilere karşılık gelen açıklamalar, görüntü ile aynı ismi taşıyan bir metin (.txt) dosyasına kaydedilir. Her .txt dosyası, görüntüde bulunan her bir nesne için aşağıdaki ek açıklamaları içermektedir:

< sınıf etiketi > <  $x_{\text{merkez}}$  > <  $y_{\text{merkez}}$  > < genişlik > < yükseklik >

Burada sınıf etiketi, nesnenin sınıfını tanımlayan ve 0'dan başlayan bir tamsayıdır. Bu çalışmada, meme kitlesi tespiti için benign sınıfı "0", malign sınıfı ise "1" olarak etiketlenmiştir. Diğer ek açıklamalar ise sınırlayıcı kutunun merkez koordinatları ( $x_{\text{merkez}}$ ,  $y_{\text{merkez}}$ ), genişlik ve yükseklik değerlerini temsil eder. Mamografi görüntüsündeki bir kitleye ait bu koordinatların görsel karşılığı Şekil 4.5'te sunulmuştur.



Şekil 4.5. Bir kitleye ait YOLO sınırlayıcı kutusunun gösterimi.

Bu değerler, .txt dosyalarında sırasıyla Eşitlik 4.1, Eşitlik 4.2, Eşitlik 4.3 ve Eşitlik 4.4'te gösterildiği gibi normalize edilerek saklanır.

$$x_{\text{merkez}} = \frac{x_{\text{merkez}}}{\text{görüntü genişliği}} \quad (4.1)$$

$$y_{\text{merkez}} = \frac{y_{\text{merkez}}}{\text{görüntü yüksekliği}} \quad (4.2)$$

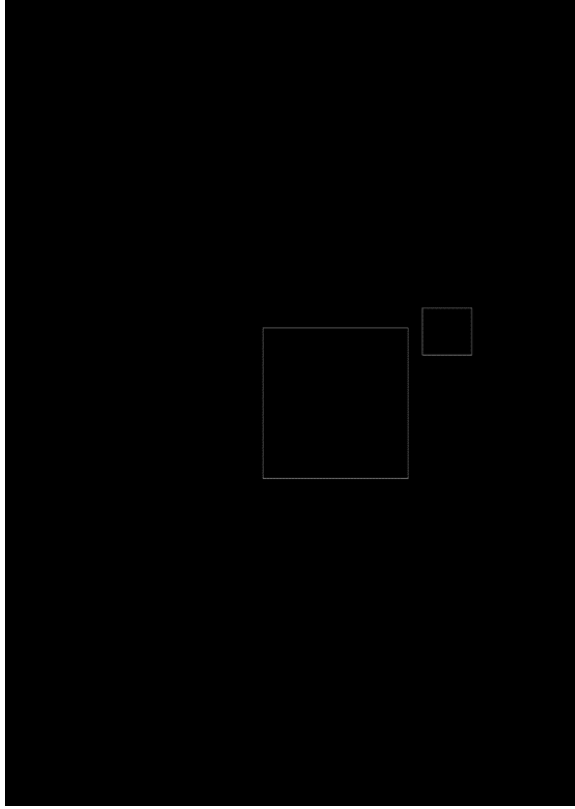
$$\text{genişlik} = \frac{\text{genişlik}}{\text{görüntü genişliği}} \quad (4.3)$$

$$\text{yükseklik} = \frac{\text{yükseklik}}{\text{görüntü yüksekliği}} \quad (4.4)$$

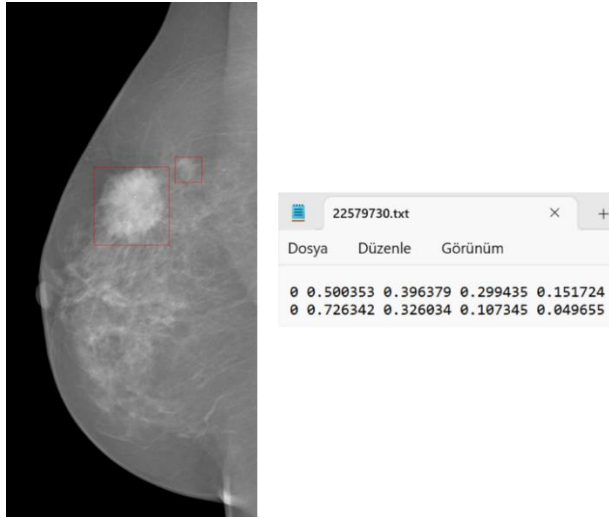
Mamografi görüntülerinde birden fazla kitle bulunması durumunda, her nesne bilgisi .txt dosyasına yeni satır olarak eklenir.

#### 4.1.2.a. INBreast veri kümesinin etiketlenmesi

INBreast veri kümesinde her vakaya ait lezyon türüne göre piksel düzeyinde ek açıklamaları içeren bir XML dosyası bulunmaktadır. Bu dosyalarda kitle sınırları, bir dizi koordinat noktası ile tanımlanmıştır. Kitle tespiti için model girdi olarak sınırlayıcı kutuları kullandığından, bu nokta dizisinin sınırlayıcı kutulara dönüştürülmesi gerekmektedir. Bu dönüşüm için, kitle sınırını ifade eden nokta dizisinden x ve y koordinatlarının minimum ve maksimum noktaları belirlenmiştir. Ardından, kitle dikdörtgen sınırlayıcı kutu içerisinde kalacak şekilde bu koordinat değerleri kullanılarak tüm görüntüler için Şekil 4.6'da gösterildiği gibi sınırlayıcı kutuları temsil eden maske görüntüleri oluşturulmuştur. Bu maske görüntüleri, orijinal görüntülere ait kırpma konum bilgilerine göre kırpılarak sınırlayıcı kutuların koordinat bilgileri güncellenmiştir. Böylece, her kitlenin konumu kırpılmış görüntüye göre yeniden hesaplanmış ve etiketleme işlemi bu yeni koordinatlar üzerinden gerçekleştirilmiştir. Ardından, her sınırlayıcı kutunun merkez koordinatları ( $x_{\text{merkez}}$ ,  $y_{\text{merkez}}$ ), yüksekliği ve genişliği hesaplanmış; ardından bu değerler görüntü boyutuna göre normalize edilmiştir. Son olarak, her görüntü için oluşturulan .txt dosyasına kitlenin sınıf etiketi ve normalize edilmiş koordinat değerleri yazdırılmıştır. Şekil 4.7'de iki adet benign kitlesine sahip örnek görüntü ve bu kitlelerin sınıfları ve koordinat değerlerini içeren etiket dosyası gösterilmiştir.



Şekil 4.6. Görüntüdeki iki farklı kitleye ait sınırlayıcı kutu maskeleri.

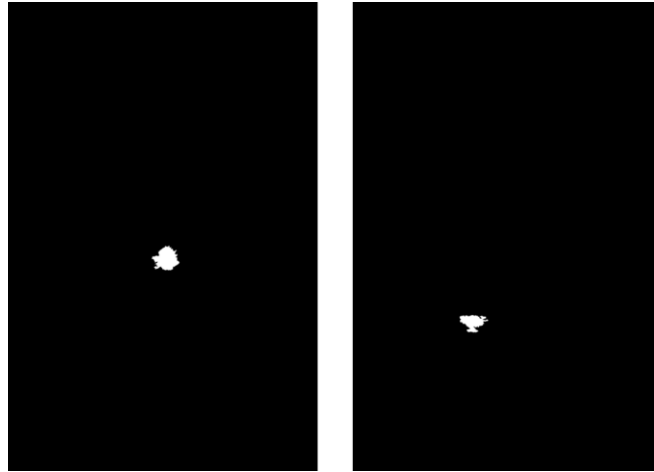


Şekil 4.7. INBreast veri kümesinden kitle etiketlenmesine ve etiket dosyasına örnek.

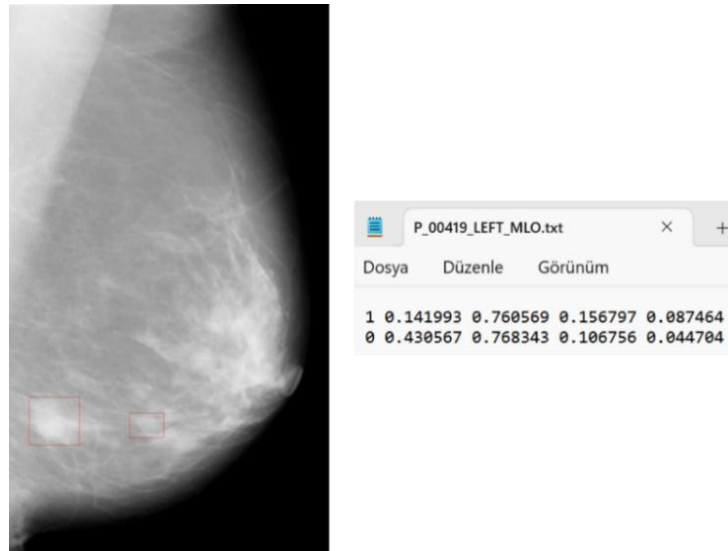
#### 4.1.2.b. CBIS-DDSM veri kümesinin etiketlenmesi

CBIS-DDSM veri kümesinde kitleler için maske görüntüleri bulunmaktadır. Şekil 4.8’de gösterildiği gibi bu maske görüntüleri siyah-beyaz formatta olup, kitle ROI’si beyaz alanla ifade edilmiştir. İlk olarak bu maske görüntüleri, karşılık geldiği orijinal görüntülerin kırpma konum bilgilerine göre kırpılmıştır. Kırpılmış maske görüntüleri üzerinde, beyaz alan tespit edilerek kitleye karşılık gelen bölge belirlenmiştir. Ardından, bu bölgenin

sınırlayıcı kutusu MATLAB'daki "regionprops" fonksiyonu ile elde edilmiştir. Bu fonksiyon, bir görüntüdeki her nesne için sınırlayıcı kutu bilgilerini hesaplamak için kullanılmıştır. Elde edilen sınırlayıcı kutulara ait merkez koordinatları ( $x_{\text{merkez}}$ ,  $y_{\text{merkez}}$ ), yükseklik ve genişlik bilgileri hesaplanmış; ardından bu değerler görüntü boyutuna göre normalize edilmiştir. Sonuçta her görüntü için oluşturulan .txt dosyasına kitlenin sınıf etiketi ve normalize edilmiş koordinat değerleri yazdırılmıştır. Şekil 4.9'da hem benign hem de malign kitlesine sahip örnek bir görüntü ve bu görüntüye ait etiket dosyası gösterilmiştir.



Şekil 4.8. Bir görüntüdeki iki farklı kitleye ait iki adet maske görüntüsü.



Şekil 4.9. CBIS-DDSM veri kümesinden kitle etiketlenmesine ve etiket dosyasına örnek.

Tüm bu işlemler sonucunda hem düşük görüntü kalitesi hem de görüntü boyutu ile maskesi arasındaki uyumsuzluk sebebiyle 94 görüntü veri kümesinden ayıklanmıştır. Sonuç olarak, analizlerde kullanılmak üzere 1506 mamografi görüntüsünden elde edilen, 853'ü

benign ve 756'sı malign olmak üzere toplam 1609 kitle lezyonu YOLO formatında etiketlenmiştir.

#### 4.1.2.c. VinDr-Mammo veri kümesinin etiketlenmesi

VinDr-Mammo veri kümesinde, her görüntü ye ait kitlelerin sınırlayıcı kutu koordinatları ( $x_{\min}$ ,  $y_{\min}$ ,  $x_{\max}$  ve  $y_{\max}$ ) ve ilgili tüm ek açıklamalar CSV dosyalarında saklanmaktadır. Verilen bu koordinat değerleri orijinal görüntülere ait olduğundan kırılan görüntüler için bu koordinat değerleri kırpma konum bilgilerine göre güncellenmiştir. Ardından, sırasıyla Eşitlik 4.5, Eşitlik 4.6, Eşitlik 4.7 ve Eşitlik 4.8 kullanılarak YOLO formatında gerekli olan merkez koordinatları ve kutu boyutları hesaplanmıştır:

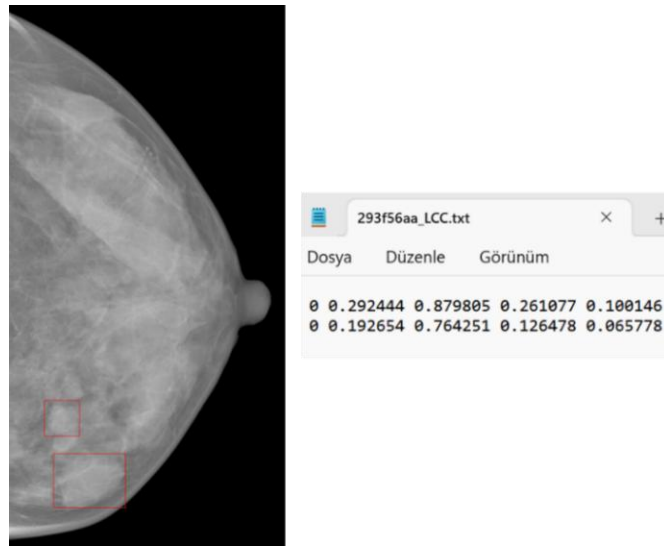
$$x_{\text{merkez}} = \frac{x_{\min} + x_{\max}}{2} \quad (4.5)$$

$$y_{\text{merkez}} = \frac{y_{\min} + y_{\max}}{2} \quad (4.6)$$

$$\text{genişlik} = x_{\max} - x_{\min} \quad (4.7)$$

$$\text{yükseklik} = y_{\max} - y_{\min} \quad (4.8)$$

Elde edilen bu değerler, daha önce tanımlanan normalizasyon işlemleri (Eşitlik 4.5-Eşitlik 4.8) doğrultusunda görüntü boyutuna göre normalize edilmiştir. Normalize edilen bilgiler, her görüntüyle aynı ismi taşıyan .txt dosyalarına, sınıf etiketleriyle birlikte kaydedilmiştir. Şekil 4.10'da iki adet benign kitlesine sahip örnek bir görüntü ve görüntüye ait etiket dosyası gösterilmiştir.

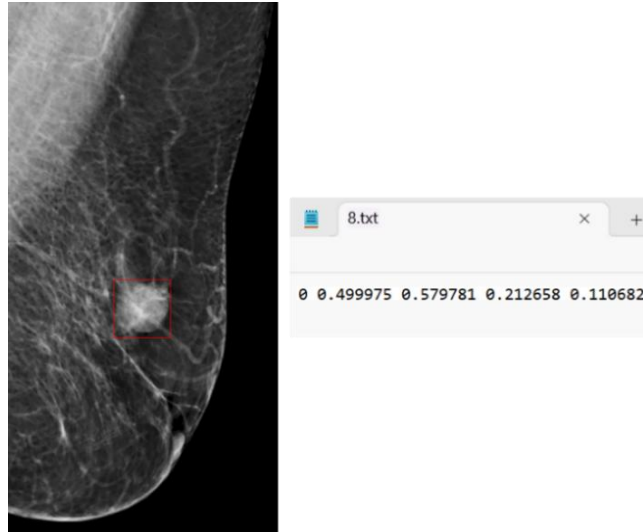


Şekil 4.10. VinDr-Mammo veri kümesinden kitle etiketlenmesine ve etiket dosyasına örnek.

Veri kümesinde 98 görüntüde bulunan kitleler, verilen sınırlayıcı kutular ile uyuşmadığından bu görüntüler analiz dışı bırakılmıştır. Bu işlem sonucunda, geriye 931 görüntü kalmış ve görüntülerde tespit edilen toplam 1019 kitle YOLO formatında etiketlenmiştir. Bu lezyonların 477'si benign, 542'si ise malign olarak sınıflandırılmıştır.

#### 4.1.2.d. Başkent Üniversitesi Ankara Hastanesi veri kümesinin etiketlenmesi

Başkent Üniversitesi Ankara Hastanesi'nden elde edilen anonimleştirilmiş mamografi görüntüleri, MATLAB ortamına aktarılmış ve bu çalışma kapsamında geliştirilen bir arayüz aracılığıyla uzman bir radyolog tarafından manuel olarak işaretlenmiş ve etiketlenmiştir. Her görüntü üzerinde, radyolog tarafından kitleler sınırlayıcı kutular çizilerek belirlenmiş ve her kutuya ait sınıf etiketi (benign için 0, malign için 1) kullanıcı girdisiyle atanmıştır. Bir görüntüde birden fazla kitle işaretlenebilmekte olup, bu süreç radyoloğun onayıyla tekrarlanabilir şekilde tasarlanmıştır. İşaretleme ve etiketleme işlemi tamamlandığında, her kitlenin sınırlayıcı kutu koordinatları ve sınıf etiketleri kaydedilmiş; ayrıca anotasyon içeren görseller de ayrı bir klasörde saklanmıştır. Elde edilen bu koordinat bilgileri, daha önce uygulanan kırpma işlemlerine göre güncellenmiş, ardından normalize edilerek her görüntüye karşılık gelen .txt dosyalarına YOLO formatında yazdırılmıştır. Şekil 4.11'de bir adet benign kitlesine sahip örnek bir görüntü ve görüntüye ait etiket dosyası gösterilmiştir.



Şekil 4.11. Başkent Üniversitesi Ankara Hastanesi veri kümesinden kitle etiketlenmesine ve etiket dosyasına örnek.

Sonuç olarak, çalışma kapsamında kullanılan mamografi veri kümelerindeki görüntü ve kitle lezyon sayıları Tablo 4.1'de verilmiştir.

**Tablo 4.1.** Veri kümelerindeki mamografi görüntülerinin ve kitle lezyonlarının sayısı

| Veri kümesi                              | Mamografi görüntü sayısı | Kitle lezyon sayısı |        |        |
|--|--------------------------|---------------------|--------|--------|
|  |                          | Benign              | Malign | Toplam |
| CBIS-DDSM                                | 1506                     | 853                 | 756    | 1609   |
| VinDr-Mammo                              | 931                      | 477                 | 542    | 1019   |
| INBreast                                 | 107                      | 40                  | 76     | 116    |
| Başkent Üniversitesi<br>Ankara Hastanesi | 102                      | 65                  | 53     | 118    |

### 4.3. YOLOv5 Analiz Sonuçları

Bu tez kapsamında deneysel analizlere, literatürde yaygın olarak kullanılan ve dengeli bir performans sunmasıyla öne çıkan YOLOv5 modeliyle başlanmıştır. Analizlerin planlandığı dönemde YOLOv5, nesne tespiti alanında sıkça tercih edilen bir mimari olduğundan, çalışmanın başlangıç modeli olarak seçilmiştir. YOLOv5 eğitimi sırasında, genelleme yeteneğini artırmak ve aşırı uyumu (overfitting) önlemek amacıyla, YOLOv5 eğitim sürecinde düşük (low), orta (medium) ve yüksek (high) seviye olmak üzere üç farklı veri artırma konfigürasyonu sunulmaktadır. Bu konfigürasyonlara ait veri artırma parametreleri ve değerleri Tablo 4.2’de sunulmuştur.

**Tablo 4.2.** YOLOv5 eğitiminde kullanılan veri artırma konfigürasyonları

| Parametre ve tanımı  | düşük (low) | orta (medium) | yüksek (high) |
|--|-------------|---------------|---------------|
| <b>hsv_h:</b> görüntü HSV-ton artırma (kesir)                    | 0.015       | 0.015         | 0.015         |
| <b>hsv_s:</b> görüntü HSV-doygunluk artırma (kesir)              | 0.7         | 0.7           | 0.7           |
| <b>hsv_v:</b> görüntü HSV-değer artırma (kesir)                  | 0.4         | 0.4           | 0.4           |
| <b>degrees:</b> görüntü döndürme (+/- derece)                    | 0.0         | 0.0           | 0.0           |
| <b>translate:</b> görüntü kaydırma (+/-kesir)                    | 0.1         | 0.1           | 0.1           |
| <b>scale:</b> görüntü ölçekleme (+/- kazanç)                     | 0.5         | 0.9           | 0.9           |
| <b>shear:</b> görüntü kesme (+/- derece)                         | 0.0         | 0.0           | 0.0           |
| <b>perspective:</b> görüntü perspektifi (+/-kesir)               | 0.0         | 0.0           | 0.0           |
| <b>flipud:</b> görüntü yukarı-aşağı çevirme (olasılık)           | 0.0         | 0.0           | 0.0           |
| <b>fliplr:</b> görüntü sola-sağa çevirme (olasılık)              | 0.5         | 0.5           | 0.5           |
| <b>mosaic:</b> görüntü mozaikçi (olasılık)                       | 1.0         | 1.0           | 1.0           |
| <b>mixup:</b> görüntü karıştırma (olasılık)                      | 0.0         | 0.1           | 0.1           |
| <b>copy_paste:</b> segmentasyon için kopyala-yapıştır (olasılık) | 0.0         | 0.0           | 0.1           |

YOLOv5 için sıfırdan eğitim ve ince ayar süreçlerinde kullanılan hiperparametreler ise Tablo 4.3'te verilmiştir.

**Tablo 4.3.** YOLOv5 için sıfırdan eğitim ve ince ayar süreçlerinde kullanılan hiperparametreler

| hiperparametre                     | değeri    |
|------------------------------------|-----------|
| görüntü boyutu                     | 544 x 544 |
| epok (sıfırdan eğitim / ince ayar) | 300 / 200 |
| toplu işlem boyutu                 | 8         |
| başlangıç öğrenme oranı            | 0.01      |
| optimizasyon algoritması           | SGD       |
| momentum katsayısı                 | 0.937     |
| ağırlık azaltma katsayısı          | 0.0005    |
| erken durdurma değeri (epok)       | 100       |

Tablo 4.2'de görüldüğü üzere yüksek seviye konfigürasyonda yer alan kopyala-yapıştır parametresi yalnızca segmentasyon problemleri için kullanılmakta olup, bu çalışma nesne tespit problemi üzerine olduğu için ilgili parametre uygulanmamıştır. Orta ve yüksek seviye veri artırma konfigürasyonları arasındaki tek fark bu parametre olduğu için, analiz sürecinde yalnızca düşük ve orta seviye veri artırma konfigürasyonları dikkate alınmıştır.

Önerilen transfer öğrenme yaklaşımı doğrultusunda, CBIS-DDSM, VinDr-Mammo ve bu iki veri kümesinin birleşiminden oluşan CBIS-DDSM + VinDr-Mammo veri kümesi, rastgele olacak şekilde %80 eğitim, %10 doğrulama ve %10 test alt kümelerine ayrılmıştır. Bu üç veri kümesi üzerinde, farklı YOLOv5 modelleri (YOLOv5n, YOLOv5s, YOLOv5m ve YOLOv5s-Transformer) düşük ve orta seviye veri artırma konfigürasyonları kullanılarak sıfırdan ayrı ayrı eğitilmiş ve test edilmiştir. Her bir veri kümesine ait test performans sonuçları sırasıyla Tablo 4.4, Tablo 4.5 ve Tablo 4.6'de verilmiştir.

**Tablo 4.4.** CBIS-DDSM veri kümesinde farklı YOLOv5 modellerinin sıfırdan eğitim performans sonuçları

| Model   | Veri artırma konfigürasyonu | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5     |
|---------|-----------------------------|----------|------------|-------|-------|-------------|
| YOLOv5n | düşük                       | 0.517    | 0.444      | 0.375 | 0.509 | 0.442       |
|         | orta                        | 0.489    | 0.453      | 0.322 | 0.519 | 0.42        |
| YOLOv5s | düşük                       | 0.461    | 0.463      | 0.403 | 0.558 | <b>0.48</b> |
|         | orta                        | 0.439    | 0.504      | 0.372 | 0.531 | 0.452       |

**Tablo 4.4.** devam ediyor.

|                         |       |       |       |       |       |       |
|-------------------------|-------|-------|-------|-------|-------|-------|
| YOLOv5s-<br>transformer | düşük | 0.615 | 0.423 | 0.401 | 0.499 | 0.45  |
|                         | orta  | 0.461 | 0.472 | 0.378 | 0.541 | 0.46  |
| YOLOv5m                 | düşük | 0.454 | 0.466 | 0.35  | 0.51  | 0.43  |
|                         | orta  | 0.48  | 0.518 | 0.364 | 0.583 | 0.474 |

**Tablo 4.5.** VinDr-Mammo veri kümesinde farklı YOLOv5 modellerinin sıfırdan eğitim performans sonuçları

| Model                   | Veri artırma<br>konfigürasyonu | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|-------------------------|--------------------------------|----------|------------|-------|-------|--------------|
| YOLOv5n                 | düşük                          | 0.541    | 0.494      | 0.484 | 0.404 | 0.444        |
|                         | orta                           | 0.495    | 0.496      | 0.524 | 0.451 | 0.487        |
| YOLOv5s                 | düşük                          | 0.535    | 0.476      | 0.467 | 0.477 | 0.472        |
|                         | orta                           | 0.571    | 0.466      | 0.546 | 0.464 | <b>0.505</b> |
| YOLOv5s-<br>transformer | düşük                          | 0.523    | 0.506      | 0.525 | 0.435 | 0.48         |
|                         | orta                           | 0.495    | 0.496      | 0.524 | 0.451 | 0.487        |
| YOLOv5m                 | düşük                          | 0.497    | 0.437      | 0.435 | 0.453 | 0.444        |
|                         | orta                           | 0.529    | 0.502      | 0.489 | 0.456 | 0.473        |

**Tablo 4.6.** CBIS-DDSM + VinDr-Mammo veri kümesinde farklı YOLOv5 modellerinin sıfırdan eğitim performans sonuçları

| Model                   | Veri artırma<br>konfigürasyonu | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|-------------------------|--------------------------------|----------|------------|-------|-------|--------------|
| YOLOv5n                 | düşük                          | 0.471    | 0.502      | 0.417 | 0.517 | 0.467        |
|                         | orta                           | 0.455    | 0.525      | 0.402 | 0.554 | 0.478        |
| YOLOv5s                 | düşük                          | 0.498    | 0.51       | 0.438 | 0.528 | 0.483        |
|                         | orta                           | 0.499    | 0.564      | 0.447 | 0.581 | <b>0.514</b> |
| YOLOv5s-<br>transformer | düşük                          | 0.547    | 0.467      | 0.416 | 0.53  | 0.473        |
|                         | orta                           | 0.451    | 0.536      | 0.414 | 0.50  | 0.457        |
| YOLOv5m                 | düşük                          | 0.395    | 0.543      | 0.389 | 0.509 | 0.449        |
|                         | orta                           | 0.459    | 0.503      | 0.411 | 0.524 | 0.468        |

Tablo 4.4, Tablo 4.5 ve Tablo 4.6’da sıfırdan eğitilmiş YOLOv5 modellerine ait performans sonuçları karşılaştırmalı olarak değerlendirilmiştir. CBIS-DDSM veri kümesi üzerinde en iyi performans (mAP@0.5=0.48), düşük seviye veri artırma konfigürasyonu ile

eğitilen YOLOv5s modeli ile elde edilmiştir. VinDr-Mammo veri kümesi için en iyi başarıya ( $mAP@0.5=0.505$ ), YOLOv5s modeli ile orta seviye veri artırma konfigürasyonu kullanıldığında ulaşılmıştır. CBIS-DDSM + VinDr-Mammo birleşim veri kümesi ise diğer iki veri kümesine kıyasla daha yüksek performansa ( $mAP@0.5=0.514$ ) sahip olup, bu başarı YOLOv5s modeli ile orta seviye veri artırma konfigürasyon kombinasyonunda elde edilmiştir.

İnce ayar sürecinde, CBIS-DDSM, VinDr-Mammo ve CBIS-DDSM + VinDr-Mammo birleşim veri kümesi ile sıfırdan eğitilmiş YOLOv5 modellerinin düşük ve orta seviye veri artırma konfigürasyonlarıyla elde edilen en iyi ağırlıkları kullanılarak INBreast veri kümesi üzerinde ince ayar gerçekleştirilmiştir. INBreast veri kümesinin küçük boyutlu olmasından dolayı tüm veri kümesini değerlendirmek amacıyla 5 katlı çapraz doğrulama stratejisi kullanılmıştır. Bu üç veri kümesine ait ince ayar sonuçları sırasıyla Tablo 4.7, Tablo 4.8, Tablo 4.9, Tablo 4.10, Tablo 4.11 ve Tablo 4.12’de sunulmuştur.

**Tablo 4.7.** CBIS-DDSM → INBreast: Düşük seviye veri artırma ile eğitilen YOLOv5 modellerinin ince ayar performans sonuçları

| Model               | Veri artırma konfigürasyonu | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|---------------------|-----------------------------|----------|------------|-------|-------|--------------|
| YOLOv5n             | düşük                       | 0.732    | 0.585      | 0.515 | 0.815 | 0.665        |
|                     | orta                        | 0.707    | 0.676      | 0.544 | 0.832 | 0.688        |
| YOLOv5s             | düşük                       | 0.92     | 0.695      | 0.709 | 0.858 | 0.783        |
|                     | orta                        | 0.783    | 0.779      | 0.755 | 0.868 | <b>0.812</b> |
| YOLOv5s-transformer | düşük                       | 0.774    | 0.72       | 0.711 | 0.818 | 0.765        |
|                     | orta                        | 0.82     | 0.717      | 0.738 | 0.848 | 0.793        |
| YOLOv5m             | düşük                       | 0.814    | 0.656      | 0.602 | 0.854 | 0.728        |
|                     | orta                        | 0.777    | 0.548      | 0.778 | 0.893 | 0.623        |

**Tablo 4.8.** VinDr-Mammo → INBreast: Düşük seviye veri artırma ile eğitilen YOLOv5 modellerinin ince ayar performans sonuçları

| Model   | Veri artırma konfigürasyonu | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|---------|-----------------------------|----------|------------|-------|-------|--------------|
| YOLOv5n | düşük                       | 0.733    | 0.704      | 0.613 | 0.806 | 0.709        |
|         | orta                        | 0.812    | 0.619      | 0.587 | 0.833 | 0.71         |
| YOLOv5s | düşük                       | 0.808    | 0.726      | 0.709 | 0.865 | 0.787        |
|         | orta                        | 0.828    | 0.739      | 0.761 | 0.844 | <b>0.803</b> |

**Tablo 4.8.** devam ediyor.

|                         |       |       |       |       |       |       |
|-------------------------|-------|-------|-------|-------|-------|-------|
| YOLOv5s-<br>transformer | düşük | 0.774 | 0.72  | 0.711 | 0.818 | 0.765 |
|                         | orta  | 0.774 | 0.772 | 0.702 | 0.876 | 0.789 |
| YOLOv5m                 | düşük | 0.857 | 0.621 | 0.609 | 0.853 | 0.731 |
|                         | orta  | 0.733 | 0.704 | 0.613 | 0.806 | 0.709 |

**Tablo 4.9.** CBIS-DDSM + VinDr-Mammo → INBreast: Düşük seviye veri artırma ile eğitilen YOLOv5 modellerinin ince ayar performans sonuçları

| Model                   | Veri artırma<br>konfigürasyonu | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|-------------------------|--------------------------------|----------|------------|-------|-------|--------------|
| YOLOv5n                 | düşük                          | 0.735    | 0.694      | 0.607 | 0.868 | 0.738        |
|                         | orta                           | 0.809    | 0.644      | 0.633 | 0.865 | 0.749        |
| YOLOv5s                 | düşük                          | 0.833    | 0.673      | 0.729 | 0.899 | 0.814        |
|                         | orta                           | 0.842    | 0.677      | 0.721 | 0.916 | <b>0.818</b> |
| YOLOv5s-<br>transformer | düşük                          | 0.732    | 0.748      | 0.729 | 0.867 | 0.798        |
|                         | orta                           | 0.868    | 0.688      | 0.748 | 0.846 | 0.797        |
| YOLOv5m                 | düşük                          | 0.792    | 0.725      | 0.695 | 0.883 | 0.789        |
|                         | orta                           | 0.849    | 0.641      | 0.652 | 0.86  | 0.756        |

**Tablo 4.10.** CBIS-DDSM → INBreast: Orta seviye veri artırma ile eğitilen YOLOv5 modellerinin ince ayar performans sonuçları

| Model                   | Veri artırma<br>konfigürasyonu | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5 |
|-------------------------|--------------------------------|----------|------------|-------|-------|---------|
| YOLOv5n                 | düşük                          | 0.777    | 0.633      | 0.603 | 0.857 | 0.73    |
|                         | orta                           | 0.758    | 0.649      | 0.569 | 0.851 | 0.71    |
| YOLOv5s                 | düşük                          | 0.81     | 0.654      | 0.618 | 0.84  | 0.729   |
|                         | orta                           | 0.721    | 0.712      | 0.65  | 0.85  | 0.75    |
| YOLOv5s-<br>transformer | düşük                          | 0.731    | 0.702      | 0.635 | 0.873 | 0.754   |
|                         | orta                           | 0.883    | 0.642      | 0.588 | 0.881 | 0.735   |
| YOLOv5m                 | düşük                          | 0.832    | 0.709      | 0.685 | 0.857 | 0.771   |
|                         | orta                           | 0.902    | 0.65       | 0.676 | 0.872 | 0.774   |

**Tablo 4.11.** VinDr-Mammo → INBreast: Orta seviye veri artırma ile eğitilen YOLOv5 modellerinin ince ayar performans sonuçları

| Model               | Veri artırma konfigürasyonu | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5 |
|---------------------|-----------------------------|----------|------------|-------|-------|---------|
| YOLOv5n             | düşük                       | 0.737    | 0.638      | 0.556 | 0.838 | 0.697   |
|                     | orta                        | 0.736    | 0.63       | 0.538 | 0.839 | 0.689   |
| YOLOv5s             | düşük                       | 0.803    | 0.673      | 0.637 | 0.832 | 0.734   |
|                     | orta                        | 0.714    | 0.716      | 0.651 | 0.844 | 0.748   |
| YOLOv5s-transformer | düşük                       | 0.767    | 0.672      | 0.649 | 0.816 | 0.732   |
|                     | orta                        | 0.744    | 0.643      | 0.60  | 0.829 | 0.714   |
| YOLOv5m             | düşük                       | 0.865    | 0.636      | 0.667 | 0.796 | 0.732   |
|                     | orta                        | 0.806    | 0.658      | 0.705 | 0.808 | 0.756   |

**Tablo 4.12.** CBIS-DDSM + VinDr-Mammo → INBreast: Orta seviye veri artırma ile eğitilen YOLOv5 modellerinin ince ayar performans sonuçları

| Model               | Veri artırma konfigürasyonu | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5 |
|---------------------|-----------------------------|----------|------------|-------|-------|---------|
| YOLOv5n             | düşük                       | 0.80     | 0.703      | 0.696 | 0.889 | 0.792   |
|                     | orta                        | 0.752    | 0.673      | 0.674 | 0.801 | 0.738   |
| YOLOv5s             | düşük                       | 0.868    | 0.693      | 0.649 | 0.873 | 0.761   |
|                     | orta                        | 0.823    | 0.682      | 0.678 | 0.847 | 0.763   |
| YOLOv5s-transformer | düşük                       | 0.861    | 0.622      | 0.626 | 0.863 | 0.745   |
|                     | orta                        | 0.839    | 0.658      | 0.648 | 0.87  | 0.759   |
| YOLOv5m             | düşük                       | 0.722    | 0.646      | 0.531 | 0.865 | 0.70    |
|                     | orta                        | 0.891    | 0.627      | 0.644 | 0.845 | 0.746   |

Düşük ve orta seviye veri artırma konfigürasyonlarıyla sıfırdan eğitilen YOLOv5 modellerine ait ince ayar sonuçları (Tablo 4.7-Tablo 4.12) karşılaştırmalı olarak değerlendirildiğinde;

- Her üç veri kümesi için de en yüksek ince ayar performansı, düşük seviye veri artırma konfigürasyonu ile eğitilmiş modellerden elde edilmiştir. Bu doğrultuda, model eğitimi aşamasında uygulanan düşük seviye veri artırma stratejisinin, ince ayar sürecindeki performans üzerinde daha olumlu bir etki yarattığı sonucuna varılmıştır.
- İnce ayar aşamasında ise, tüm veri kümeleri için en yüksek performans YOLOv5s modeli ile orta seviye veri artırma konfigürasyonunun kombinasyonu kullanıldığında

elde edilmiştir. Bu kombinasyon ile ulaşılan mAP@0.5 değerleri CBIS-DDSM, VinDr-Mammo ve birleşim veri kümeleri için sırasıyla 0.812, 0.803 ve 0.818 olarak hesaplanmıştır.

- En yüksek ince ayar performansının YOLOv5s modeli ve orta seviye veri artırma konfigürasyonu ile elde edilmesi nedeniyle, sonraki deneysel analizlerde bu kombinasyon temel alınmıştır.

Çalışmanın devamında, ince ayar sürecinde kullanılan varsayılan orta seviye veri artırma konfigürasyonuna dahil olmayan bazı ek veri artırma parametrelerinin model performansına etkisi incelenmiştir. Bu kapsamda; görüntü yukarı-aşağı çevirme (flipud), görüntü döndürme (degrees) ve görüntü kesme (shear) parametreleri analizlere dahil edilmiştir. Yapılan ön analizlerde perspektif parametresi anlamlı bir katkı sağlamadığı için bu parametre analiz dışı bırakılmıştır. İlgili parametreler sırasıyla flipud (0.5), degrees (10°) ve shear (10°) olarak belirlenmiştir. Bu veri artırma parametrelerinin INBreast veri kümesinin ince ayar performansı üzerindeki etkisi Tablo 4.13, Tablo 4.14 ve Tablo 4.15'te sunulmuştur.

**Tablo 4.13.** CBIS-DDSM → INBreast: YOLOv5s modeli için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Deneme | Veri artırma parametresi |                  |                | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|--------|--------------------------|------------------|----------------|----------|------------|-------|-------|--------------|
|        | flipud<br>(0.5)          | degrees<br>(10°) | shear<br>(10°) |          |            |       |       |              |
| 0      | -                        | -                | -              | 0.783    | 0.779      | 0.755 | 0.868 | 0.812        |
| 1      | ✓                        | -                | -              | 0.778    | 0.828      | 0.778 | 0.902 | <b>0.839</b> |
| 2      | -                        | ✓                | -              | 0.895    | 0.711      | 0.70  | 0.907 | 0.803        |
| 3      | -                        | -                | ✓              | 0.805    | 0.697      | 0.646 | 0.905 | 0.776        |
| 4      | ✓                        | ✓                | -              | 0.841    | 0.695      | 0.679 | 0.897 | 0.789        |
| 5      | ✓                        | -                | ✓              | 0.858    | 0.719      | 0.701 | 0.897 | 0.799        |
| 6      | -                        | ✓                | ✓              | 0.817    | 0.718      | 0.723 | 0.853 | 0.788        |
| 7      | ✓                        | ✓                | ✓              | 0.868    | 0.752      | 0.769 | 0.893 | 0.831        |

**Tablo 4.14.** VinDr-Mammo → INBreast: YOLOv5s modeli için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Deneme | Veri artırma parametresi |                  |                | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5     |
|--------|--------------------------|------------------|----------------|----------|------------|-------|-------|-------------|
|        | flipud<br>(0.5)          | degrees<br>(10°) | shear<br>(10°) |          |            |       |       |             |
| 0      | -                        | -                | -              | 0.828    | 0.739      | 0.761 | 0.844 | 0.803       |
| 1      | ✓                        | -                | -              | 0.765    | 0.746      | 0.726 | 0.872 | 0.799       |
| 2      | -                        | ✓                | -              | 0.754    | 0.785      | 0.703 | 0.869 | 0.786       |
| 3      | -                        | -                | ✓              | 0.824    | 0.727      | 0.697 | 0.881 | 0.789       |
| 4      | ✓                        | ✓                | -              | 0.783    | 0.797      | 0.791 | 0.879 | 0.835       |
| 5      | ✓                        | -                | ✓              | 0.804    | 0.778      | 0.742 | 0.894 | 0.818       |
| 6      | -                        | ✓                | ✓              | 0.83     | 0.689      | 0.719 | 0.87  | 0.795       |
| 7      | ✓                        | ✓                | ✓              | 0.829    | 0.787      | 0.784 | 0.896 | <b>0.84</b> |

**Tablo 4.15.** CBIS-DDSM + VinDr-Mammo → INBreast: YOLOv5s modeli için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Deneme | Veri artırma parametresi |                  |                | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|--------|--------------------------|------------------|----------------|----------|------------|-------|-------|--------------|
|        | flipud<br>(0.5)          | degrees<br>(10°) | shear<br>(10°) |          |            |       |       |              |
| 0      | -                        | -                | -              | 0.842    | 0.677      | 0.721 | 0.916 | 0.818        |
| 1      | ✓                        | -                | -              | 0.832    | 0.774      | 0.781 | 0.894 | 0.838        |
| 2      | -                        | ✓                | -              | 0.756    | 0.771      | 0.682 | 0.868 | 0.775        |
| 3      | -                        | -                | ✓              | 0.759    | 0.74       | 0.714 | 0.883 | 0.798        |
| 4      | ✓                        | ✓                | -              | 0.853    | 0.664      | 0.908 | 0.864 | 0.749        |
| 5      | ✓                        | -                | ✓              | 0.85     | 0.75       | 0.774 | 0.897 | 0.835        |
| 6      | -                        | ✓                | ✓              | 0.768    | 0.723      | 0.704 | 0.876 | 0.79         |
| 7      | ✓                        | ✓                | ✓              | 0.911    | 0.687      | 0.792 | 0.905 | <b>0.848</b> |

Tablo 4.13, Tablo 4.14 ve Tablo 4.15 incelendiğinde;

- CBIS-DDSM → INBreast için 0.839'luk en yüksek mAP@0.5 değeri, varsayılan parametrelere yalnızca flipud (0.5) parametresinin eklendiği kombinasyonda (deneme 1) elde edilmiştir.
- VinDr-Mammo → INBreast için en yüksek mAP@0.5 değeri 0.84 olup, bu başarı tüm parametrelerin (flipud, degrees, shear) birlikte uygulandığı kombinasyonda (deneme 7) gözlenmiştir.
- CBIS-DDSM + VinDr-Mammo → INBreast için ise 0.848'lik en yüksek mAP@0.5 değeri yine flipud + degrees + shear kombinasyonu (deneme 7) ile elde edilmiştir.

Sonuç olarak, her üç veri kümesi için de uygun veri artırma parametrelerinin doğru kombinasyonlarla uygulanması, modelin ince ayar performansını belirgin bir şekilde artırmıştır. Özellikle flipud parametresinin her üç durumda da yüksek başarıya katkı sağladığı gözlemlenmiştir.

Çalışmanın devamında, ölçek (scale) parametresinin model performansı üzerindeki etkisi, önceki analizlerde en iyi performansı sağlayan parametre kombinasyonları temel alınarak incelenmiştir. Bu analiz, meme kitlelerinin boyut ve çözünürlük açısından farklılık göstermesi nedeniyle, uygun ölçek değerinin tespit performansını etkileyebileceği varsayımıyla gerçekleştirilmiştir. Farklı ölçek değerlerinin, en iyi parametre setleriyle birlikte uygulandığı durumlara ait ince ayar performans sonuçları sırasıyla Tablo 4.16, Tablo 4.17 ve Tablo 4.18'de gösterilmiştir.

**Tablo 4.16.** CBIS-DDSM → INBreast: Farklı ölçek değerlerinin Deneme 1 üzerindeki ince ayar performansına etkisi

| Scale                     | Kesinlik | Duyarlılık | BAP   | M AP  | mAP@0.5      |
|---------------------------|----------|------------|-------|-------|--------------|
| scale=0.3                 | 0.788    | 0.789      | 0.749 | 0.868 | 0.808        |
| scale=0.5                 | 0.855    | 0.774      | 0.794 | 0.892 | <b>0.843</b> |
| scale=0.7                 | 0.835    | 0.778      | 0.772 | 0.879 | 0.825        |
| scale=0.9<br>(varsayılan) | 0.778    | 0.828      | 0.778 | 0.902 | 0.839        |

**Tablo 4.17.** VinDr-Mammo → INBreast: Farklı ölçek değerlerinin Deneme 7 üzerindeki ince ayar performansına etkisi

| Scale                     | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5     |
|---------------------------|----------|------------|-------|-------|-------------|
| scale=0.3                 | 0.864    | 0.733      | 0.736 | 0.861 | 0.799       |
| scale=0.5                 | 0.788    | 0.726      | 0.668 | 0.884 | 0.776       |
| scale=0.7                 | 0.80     | 0.767      | 0.782 | 0.862 | 0.822       |
| scale=0.9<br>(varsayılan) | 0.829    | 0.787      | 0.784 | 0.896 | <b>0.84</b> |

**Tablo 4.18.** CBIS-DDSM + VinDr-Mammo → INBreast: Farklı ölçek değerlerinin Deneme 7 üzerindeki ince ayar performansına etkisi

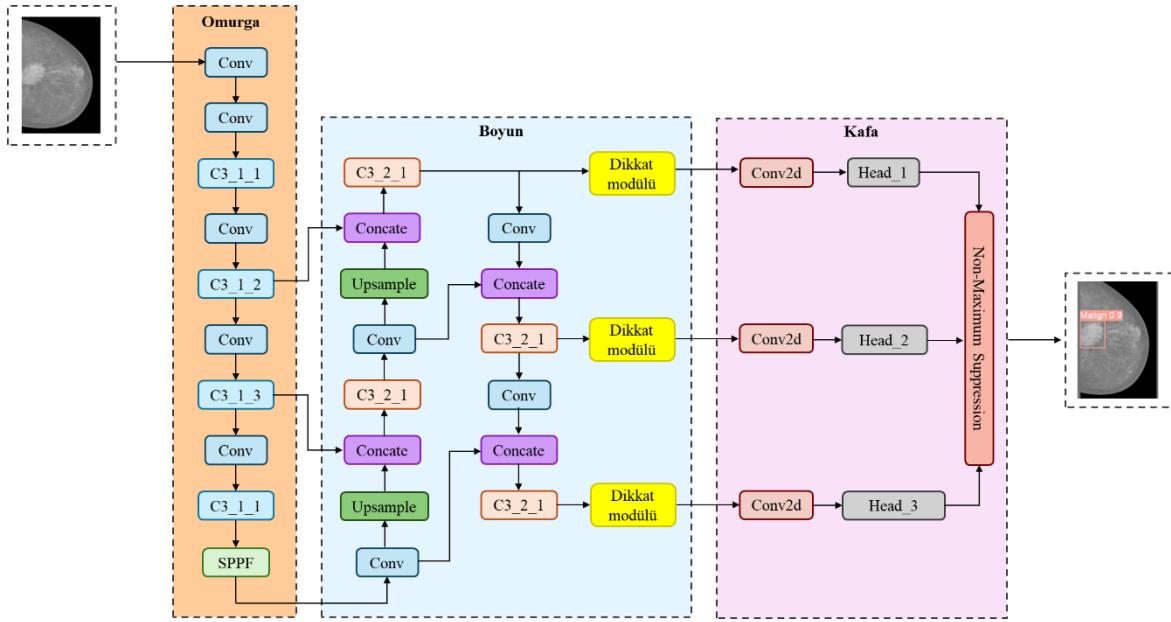
| Scale                     | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|---------------------------|----------|------------|-------|-------|--------------|
| scale=0.3                 | 0.779    | 0.754      | 0.757 | 0.897 | 0.827        |
| scale=0.5                 | 0.848    | 0.742      | 0.749 | 0.885 | 0.817        |
| scale=0.7                 | 0.778    | 0.78       | 0.747 | 0.895 | 0.821        |
| scale=0.9<br>(varsayılan) | 0.911    | 0.687      | 0.792 | 0.905 | <b>0.848</b> |

Tablo 4.16, Tablo 4.17 ve Tablo 4.18'e göre CBIS-DDSM → INBreast ince ayar modeli 0.5 ölçek değeri ile daha iyi performans gösterirken, VinDr-Mammo → INBreast ve CBIS-DDSM + VinDr-Mammo → INBreast ince ayar modelleri ise en yüksek performansa 0.9 ölçek değeri ile ulaşmıştır. Sonuç olarak;

- CBIS-DDSM → INBreast için en başarılı ince ayar modeli, orta seviye veri artırma konfigürasyonunda flipud = 0.5 ve scale = 0.5 parametrelerinin uygulandığı kombinasyon ile elde edilmiştir. Bu kombinasyonda model, 0.843 mAP@0.5, 0.855 kesinlik ve 0.774 duyarlılık değerleri göstermiştir.
- VinDr-Mammo → INBreast için en yüksek performans, flipud = 0.5, degrees = 10 ve shear = 10 parametrelerinin uygulandığı orta seviye veri artırma konfigürasyonunda elde edilmiş olup; model 0.840 mAP@0.5, 0.829 kesinlik ve 0.787 duyarlılık değerlerine ulaşmıştır.
- CBIS-DDSM + VinDr-Mammo → INBreast modelinde ise en iyi sonuç, yine aynı üç parametrenin (flipud, degrees, shear) kullanıldığı kombinasyon ile elde edilmiş; bu durumda model 0.848 mAP@0.5, 0.911 kesinlik ve 0.687 duyarlılık değerleri göstermiştir.

#### 4.4. YOLOv5 ve Dikkat Modülleri Analiz Sonuçları

Bu bölümde, bir önceki bölümde en başarılı sonuçları veren YOLOv5s modeli temel model olarak seçilmiş ve bu modele farklı dikkat modülleri (SE, ECA, CBAM, NAM, CA ve EMA) entegre edilerek yeni model varyantları elde edilmiştir. Farklı şekil ve boyuta sahip olan meme kitlelerinin tespit doğruluğunu etkili bir şekilde iyileştirmek amacıyla, dikkat modülleri Şekil 4.12’de gösterildiği gibi YOLOv5 mimarisinin boyun katmanı çıkışına eklenmiştir. Dikkat modülleri, modelin boyun katmanında üretilen üç farklı çözünürlük seviyesine ait öznetelik haritalarını işleyerek, her bir seviyede en kritik ve anlamlı özneteliklerin ön plana çıkarılmasını sağlamaktadır.



Şekil 4.12. YOLOv5 + dikkat modülü model mimarisinin gösterimi.

İlk olarak, VinDr-Mammo, CBIS-DDSM ve bu iki veri kümesinin birleşiminden oluşan VinDr-Mammo + CBIS-DDSM veri kümesi, rastgele olacak şekilde %80 eğitim, %10 doğrulama ve %10 test alt kümelerine ayrılmıştır. YOLOv5s modeline altı farklı dikkat modülü (SE, ECA, CBAM, NAM, CA ve EMA) entegre edilerek oluşturulan yeni model varyantları, bu veri kümeleri üzerinde varsayılan (düşük seviye) veri artırma konfigürasyonu ile sıfırdan eğitilmiştir. Elde edilen sıfırdan eğitim performans sonuçları Tablo 4.19, Tablo 4.20 ve Tablo 4.21’de sunulmuştur.

**Tablo 4.19.** CBIS-DDSM veri kümesinde YOLOv5s ve YOLOv5s + dikkat modüllerinin sıfırdan eğitim performans sonuçları

| <b>Model</b>   | <b>Kesinlik</b> | <b>Duyarlılık</b> | <b>B AP</b> | <b>M AP</b> | <b>mAP@0.5</b> |
|----------------|-----------------|-------------------|-------------|-------------|----------------|
| YOLOv5s        | 0.461           | 0.463             | 0.403       | 0.558       | 0.48           |
| YOLOv5s + SE   | 0.505           | 0.462             | 0.431       | 0.54        | 0.486          |
| YOLOv5s + ECA  | 0.481           | 0.52              | 0.53        | 0.502       | 0.466          |
| YOLOv5s + CBAM | 0.506           | 0.488             | 0.426       | 0.551       | <b>0.488</b>   |
| YOLOv5s + NAM  | 0.469           | 0.541             | 0.401       | 0.564       | 0.482          |
| YOLOv5s + CA   | 0.517           | 0.412             | 0.312       | 0.498       | 0.405          |
| YOLOv5s + EMA  | 0.566           | 0.493             | 0.34        | 0.596       | 0.468          |

**Tablo 4.20.** VinDr-Mammo veri kümesinde YOLOv5s ve YOLOv5s + dikkat modüllerinin sıfırdan eğitim performans sonuçları

| <b>Model</b>   | <b>Kesinlik</b> | <b>Duyarlılık</b> | <b>B AP</b> | <b>M AP</b> | <b>mAP@0.5</b> |
|----------------|-----------------|-------------------|-------------|-------------|----------------|
| YOLOv5s        | 0.535           | 0.476             | 0.467       | 0.477       | 0.472          |
| YOLOv5s + SE   | 0.581           | 0.523             | 0.461       | 0.552       | 0.506          |
| YOLOv5s + ECA  | 0.556           | 0.561             | 0.451       | 0.562       | 0.507          |
| YOLOv5s + CBAM | 0.475           | 0.58              | 0.426       | 0.544       | 0.485          |
| YOLOv5s + NAM  | 0.56            | 0.429             | 0.525       | 0.45        | 0.487          |
| YOLOv5s + CA   | 0.482           | 0.644             | 0.506       | 0.531       | <b>0.519</b>   |
| YOLOv5s + EMA  | 0.566           | 0.493             | 0.34        | 0.596       | 0.468          |

**Tablo 4.21.** CBIS-DDSM + VinDr-Mammo veri kümesinde YOLOv5s ve YOLOv5s + dikkat modüllerinin sıfırdan eğitim performans sonuçları

| <b>Model</b>   | <b>Kesinlik</b> | <b>Duyarlılık</b> | <b>B AP</b> | <b>M AP</b> | <b>mAP@0.5</b> |
|----------------|-----------------|-------------------|-------------|-------------|----------------|
| YOLOv5s        | 0.498           | 0.51              | 0.438       | 0.528       | 0.483          |
| YOLOv5s + SE   | 0.507           | 0.529             | 0.437       | 0.554       | 0.496          |
| YOLOv5s + ECA  | 0.455           | 0.556             | 0.404       | 0.557       | 0.481          |
| YOLOv5s + CBAM | 0.55            | 0.476             | 0.415       | 0.566       | 0.49           |
| YOLOv5s + NAM  | 0.566           | 0.509             | 0.428       | 0.584       | <b>0.506</b>   |
| YOLOv5s + CA   | 0.559           | 0.498             | 0.423       | 0.516       | 0.47           |
| YOLOv5s + EMA  | 0.528           | 0.538             | 0.43        | 0.555       | 0.493          |

Tablo 4.19, Tablo 4.20 ve Tablo 4.21’de sunulan sıfırdan eğitim sonuçlarına göre;

- YOLOv5s modeline entegre edilen dikkat modülleri, tespit ve sınıflandırma başarısını çoğu durumda artırmıştır.
- CBIS-DDSM veri kümesi için en yüksek başarı ( $mAP@0.5=0.488$ ), YOLOv5s + CBAM modeli ile elde edilmiştir.
- VinDr-Mammo veri kümesinde, en başarılı performans ( $mAP@0.5=0.519$ ) YOLOv5s + CA modeline aittir.
- CBIS-DDSM + VinDr-Mammo birleşim veri kümesinde ise en yüksek başarı ( $mAP@0.5=0.506$ ) YOLOv5s + NAM modeli ile sağlanmıştır.

Ardından, eğitilen modellerin en iyi ağırlıkları kullanılarak INBreast veri kümesi üzerinde ince ayar gerçekleştirilmiştir. İnce ayar sırasında, INBreast veri kümesinin küçük boyutlu olmasından dolayı tüm veri kümesini değerlendirmek için 5 katlı çapraz doğrulama stratejisi kullanılmıştır. Bu aşamada, YOLOv5 eğitiminde kullanılan varsayılan (düşük seviye) veri artırma parametrelerine ek olarak, kullanılmayan bazı parametreler analizlere dahil edilerek modellerin performansının artırılması hedeflenmiştir. Bu kapsamda, görüntü yukarı-aşağı çevirme (flipud) ve görüntü karıştırma (mixup) parametreleri uygulanmıştır. İlgili parametreler sırasıyla flipud (0.5) ve mixup (0.1) olarak belirlenmiştir. Bununla birlikte, yapılan ön analizlerde etkili olmadığı gözlemlenen diğer parametreler (degrees, shear ve perspective) bu çalışmaya dahil edilmemiştir. Yapılan analizler ve elde edilen sonuçlar sırasıyla Tablo 4.22, Tablo 4.23 ve Tablo 4.24’te sunulmuştur.

**Tablo 4.22.** CBIS-DDSM → INBreast: YOLOv5s ve YOLOv5s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model       | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5 |
|-------------|--------------------------|-------------|----------|------------|-------|-------|---------|
|             | flipud (0.5)             | mixup (0.1) |          |            |       |       |         |
| YOLOv5s     | -                        | -           | 0.92     | 0.695      | 0.709 | 0.858 | 0.783   |
|             | ✓                        | -           | 0.848    | 0.709      | 0.747 | 0.84  | 0.793   |
|             | -                        | ✓           | 0.847    | 0.758      | 0.741 | 0.883 | 0.811   |
|             | ✓                        | ✓           | 0.82     | 0.783      | 0.786 | 0.886 | 0.836   |
| YOLOv5s+SE  | -                        | -           | 0.763    | 0.725      | 0.639 | 0.874 | 0.756   |
|             | ✓                        | -           | 0.783    | 0.761      | 0.726 | 0.89  | 0.808   |
|             | -                        | ✓           | 0.808    | 0.653      | 0.653 | 0.867 | 0.76    |
|             | ✓                        | ✓           | 0.884    | 0.744      | 0.713 | 0.899 | 0.806   |
| YOLOv5s+ECA | -                        | -           | 0.806    | 0.688      | 0.684 | 0.863 | 0.774   |
|             | ✓                        | -           | 0.836    | 0.699      | 0.727 | 0.891 | 0.809   |
|             | -                        | ✓           | 0.785    | 0.699      | 0.675 | 0.874 | 0.775   |
|             | ✓                        | ✓           | 0.828    | 0.755      | 0.75  | 0.917 | 0.834   |

Tablo 4.22. devam ediyor.

|              |   |   |       |       |       |       |              |
|--------------|---|---|-------|-------|-------|-------|--------------|
| YOLOv5s+CBAM | - | - | 0.932 | 0.755 | 0.767 | 0.894 | 0.831        |
|              | ✓ | - | 0.879 | 0.734 | 0.749 | 0.875 | 0.812        |
|              | - | ✓ | 0.92  | 0.733 | 0.77  | 0.891 | 0.831        |
|              | ✓ | ✓ | 0.837 | 0.761 | 0.804 | 0.887 | <b>0.845</b> |
| YOLOv5s+NAM  | - | - | 0.839 | 0.705 | 0.698 | 0.847 | 0.833        |
|              | ✓ | - | 0.899 | 0.71  | 0.739 | 0.875 | 0.807        |
|              | - | ✓ | 0.86  | 0.758 | 0.747 | 0.893 | 0.82         |
|              | ✓ | ✓ | 0.89  | 0.757 | 0.794 | 0.858 | 0.826        |
| YOLOv5s+CA   | - | - | 0.861 | 0.653 | 0.592 | 0.86  | 0.726        |
|              | ✓ | - | 0.805 | 0.736 | 0.653 | 0.908 | 0.781        |
|              | - | ✓ | 0.806 | 0.676 | 0.634 | 0.843 | 0.738        |
|              | ✓ | ✓ | 0.894 | 0.704 | 0.715 | 0.893 | 0.804        |
| YOLOv5s+EMA  | - | - | 0.796 | 0.767 | 0.727 | 0.876 | 0.802        |
|              | ✓ | - | 0.838 | 0.736 | 0.726 | 0.87  | 0.798        |
|              | - | ✓ | 0.839 | 0.734 | 0.742 | 0.893 | 0.817        |
|              | ✓ | ✓ | 0.82  | 0.792 | 0.764 | 0.921 | 0.843        |

**Tablo 4.23.** VinDr-Mammo → INBreast: YOLOv5s ve YOLOv5s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model       | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|-------------|--------------------------|-------------|----------|------------|-------|-------|--------------|
|             | flipud (0.5)             | mixup (0.1) |          |            |       |       |              |
| YOLOv5s     | -                        | -           | 0.808    | 0.726      | 0.709 | 0.865 | 0.787        |
|             | ✓                        | -           | 0.858    | 0.72       | 0.747 | 0.846 | 0.796        |
|             | -                        | ✓           | 0.856    | 0.738      | 0.797 | 0.907 | 0.852        |
|             | ✓                        | ✓           | 0.848    | 0.76       | 0.764 | 0.902 | 0.833        |
| YOLOv5s+SE  | -                        | -           | 0.871    | 0.722      | 0.73  | 0.883 | 0.806        |
|             | ✓                        | -           | 0.889    | 0.763      | 0.835 | 0.872 | <b>0.853</b> |
|             | -                        | ✓           | 0.776    | 0.796      | 0.739 | 0.881 | 0.81         |
|             | ✓                        | ✓           | 0.842    | 0.728      | 0.742 | 0.876 | 0.809        |
| YOLOv5s+ECA | -                        | -           | 0.808    | 0.658      | 0.701 | 0.83  | 0.766        |
|             | ✓                        | -           | 0.796    | 0.758      | 0.773 | 0.86  | 0.817        |
|             | -                        | ✓           | 0.809    | 0.698      | 0.714 | 0.886 | 0.80         |
|             | ✓                        | ✓           | 0.846    | 0.69       | 0.711 | 0.88  | 0.795        |

**Tablo 4.23.** devam ediyor.

|              |   |   |       |       |       |       |       |
|--------------|---|---|-------|-------|-------|-------|-------|
| YOLOv5s+CBAM | - | - | 0.86  | 0.69  | 0.743 | 0.869 | 0.806 |
|              | ✓ | - | 0.842 | 0.746 | 0.747 | 0.889 | 0.818 |
|              | - | ✓ | 0.789 | 0.724 | 0.753 | 0.869 | 0.811 |
|              | ✓ | ✓ | 0.784 | 0.761 | 0.827 | 0.886 | 0.804 |
| YOLOv5s+NAM  | - | - | 0.883 | 0.689 | 0.725 | 0.843 | 0.784 |
|              | ✓ | - | 0.853 | 0.754 | 0.724 | 0.875 | 0.799 |
|              | - | ✓ | 0.756 | 0.689 | 0.644 | 0.849 | 0.746 |
|              | ✓ | ✓ | 0.804 | 0.733 | 0.74  | 0.88  | 0.81  |
| YOLOv5s+CA   | - | - | 0.815 | 0.745 | 0.742 | 0.862 | 0.802 |
|              | ✓ | - | 0.83  | 0.769 | 0.771 | 0.876 | 0.824 |
|              | - | ✓ | 0.877 | 0.726 | 0.73  | 0.901 | 0.816 |
|              | ✓ | ✓ | 0.872 | 0.789 | 0.778 | 0.918 | 0.848 |
| YOLOv5s+EMA  | - | - | 0.796 | 0.767 | 0.727 | 0.876 | 0.802 |
|              | ✓ | - | 0.838 | 0.736 | 0.726 | 0.87  | 0.798 |
|              | - | ✓ | 0.839 | 0.734 | 0.742 | 0.893 | 0.817 |
|              | ✓ | ✓ | 0.82  | 0.792 | 0.764 | 0.921 | 0.843 |

**Tablo 4.24.** CBIS-DDSM + VinDr-Mammo → INBreast: YOLOv5s ve YOLOv5s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model         | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|---------------|--------------------------|-------------|----------|------------|-------|-------|--------------|
|               | flipud (0.5)             | mixup (0.1) |          |            |       |       |              |
| YOLOv5s       | -                        | -           | 0.845    | 0.773      | 0.784 | 0.896 | 0.84         |
|               | ✓                        | -           | 0.924    | 0.757      | 0.815 | 0.884 | 0.849        |
|               | -                        | ✓           | 0.929    | 0.727      | 0.807 | 0.898 | 0.852        |
|               | ✓                        | ✓           | 0.833    | 0.797      | 0.768 | 0.926 | 0.847        |
| YOLOv5s + SE  | -                        | -           | 0.844    | 0.785      | 0.774 | 0.884 | 0.829        |
|               | ✓                        | -           | 0.806    | 0.773      | 0.78  | 0.896 | 0.838        |
|               | -                        | ✓           | 0.876    | 0.748      | 0.782 | 0.898 | 0.84         |
|               | ✓                        | ✓           | 0.801    | 0.814      | 0.79  | 0.917 | <b>0.854</b> |
| YOLOv5s + ECA | -                        | -           | 0.903    | 0.731      | 0.805 | 0.889 | 0.847        |
|               | ✓                        | -           | 0.849    | 0.761      | 0.81  | 0.871 | 0.841        |
|               | -                        | ✓           | 0.854    | 0.727      | 0.775 | 0.904 | 0.839        |
|               | ✓                        | ✓           | 0.79     | 0.844      | 0.789 | 0.895 | 0.842        |

**Tablo 4.24.** devam ediyor.

|                |   |   |       |       |       |       |       |
|----------------|---|---|-------|-------|-------|-------|-------|
| YOLOv5s + CBAM | - | - | 0.873 | 0.747 | 0.783 | 0.857 | 0.82  |
|                | ✓ | - | 0.848 | 0.766 | 0.784 | 0.902 | 0.843 |
|                | - | ✓ | 0.763 | 0.802 | 0.767 | 0.907 | 0.837 |
|                | ✓ | ✓ | 0.834 | 0.778 | 0.792 | 0.915 | 0.853 |
| YOLOv5s + NAM  | - | - | 0.768 | 0.783 | 0.727 | 0.891 | 0.809 |
|                | ✓ | - | 0.834 | 0.745 | 0.753 | 0.915 | 0.834 |
|                | - | ✓ | 0.807 | 0.742 | 0.772 | 0.862 | 0.817 |
|                | ✓ | ✓ | 0.872 | 0.761 | 0.767 | 0.897 | 0.832 |
| YOLOv5s + CA   | - | - | 0.85  | 0.777 | 0.777 | 0.874 | 0.826 |
|                | ✓ | - | 0.855 | 0.71  | 0.735 | 0.877 | 0.806 |
|                | - | ✓ | 0.793 | 0.786 | 0.739 | 0.888 | 0.814 |
|                | ✓ | ✓ | 0.748 | 0.793 | 0.753 | 0.867 | 0.811 |
| YOLOv5s + EMA  | - | - | 0.798 | 0.764 | 0.772 | 0.89  | 0.831 |
|                | ✓ | - | 0.895 | 0.696 | 0.731 | 0.882 | 0.806 |
|                | - | ✓ | 0.844 | 0.745 | 0.75  | 0.882 | 0.816 |
|                | ✓ | ✓ | 0.837 | 0.739 | 0.733 | 0.90  | 0.816 |

Tablo 4.22, Tablo 4.23 ve Tablo 4.24 incelendiğinde;

- CBIS-DDSM → INBreast için 0.845'lik en yüksek mAP@0.5 değeri, YOLOv5s + CBAM modeli ve varsayılan + flipud + mixup veri artırma kombinasyonu ile elde edilmiştir.
- VinDr-Mammo → INBreast için en yüksek mAP@0.5 değeri 0.853 olup, bu başarı YOLOv5s + SE modeli ve varsayılan + flipud veri artırma kombinasyonunda gözlenmiştir.
- CBIS-DDSM + VinDr-Mammo → INBreast için en yüksek mAP@0.5 değeri olan 0.854, YOLOv5s + SE modeli ve varsayılan + flipud + mixup veri artırma kombinasyonu ile elde edilmiştir.
- Genel olarak değerlendirildiğinde, CBAM modülü, CBIS-DDSM ve VinDr-Mammo veri kümeleri için en yüksek performansı sağlamış; birleşim veri kümesinde ise SE modülü öne çıkmıştır. Ayrıca, flipud ve mixup parametrelerinin ince ayar sürecine dahil edilmesi, tüm modellerde performansı belirgin ölçüde artırmıştır.

YOLOv5s + dikkat modülleriyle yapılan analizlerde, en yüksek ince ayar performansı birleşim veri kümesiyle eğitilen modeller tarafından elde edilmiştir. Bu nedenle, ince ayar sürecinde yalnızca birleşim veri kümesi ile eğitilmiş modeller kullanılarak klinik veri kümesi olan Başkent veri kümesi üzerinde ince ayar gerçekleştirilmiştir. Elde edilen ince ayar sonuçları Tablo 4.25'te sunulmuştur.

**Tablo 4.25.** CBIS-DDSM + VinDr-Mammo → Başkent: YOLOv5s ve YOLOv5s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model         | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5 |
|---------------|--------------------------|-------------|----------|------------|-------|-------|---------|
|               | flipud (0.5)             | mixup (0.1) |          |            |       |       |         |
| YOLOv5s       | -                        | -           | 0.833    | 0.753      | 0.837 | 0.837 | 0.837   |
|               | ✓                        | -           | 0.778    | 0.793      | 0.806 | 0.838 | 0.822   |
|               | -                        | ✓           | 0.818    | 0.76       | 0.82  | 0.842 | 0.831   |
|               | ✓                        | ✓           | 0.818    | 0.763      | 0.797 | 0.853 | 0.825   |
| YOLOv5s + SE  | -                        | -           | 0.795    | 0.727      | 0.815 | 0.772 | 0.793   |
|               | ✓                        | -           | 0.771    | 0.728      | 0.786 | 0.773 | 0.779   |
|               | -                        | ✓           | 0.834    | 0.719      | 0.801 | 0.803 | 0.802   |
|               | ✓                        | ✓           | 0.736    | 0.735      | 0.781 | 0.821 | 0.801   |
| YOLOv5s + ECA | -                        | -           | 0.791    | 0.722      | 0.804 | 0.759 | 0.782   |
|               | ✓                        | -           | 0.765    | 0.755      | 0.795 | 0.821 | 0.808   |
|               | -                        | ✓           | 0.786    | 0.741      | 0.782 | 0.821 | 0.802   |
|               | ✓                        | ✓           | 0.758    | 0.776      | 0.771 | 0.855 | 0.813   |

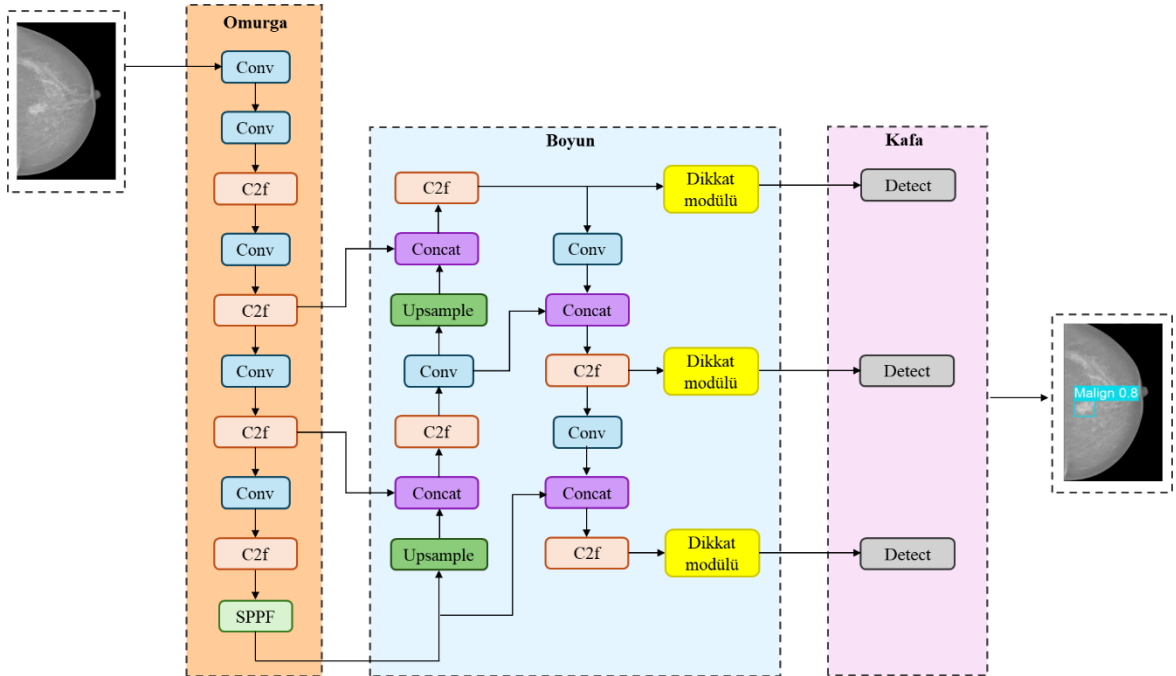
Tablo 4.25. devam ediyor.

|                |   |   |       |       |       |       |              |
|----------------|---|---|-------|-------|-------|-------|--------------|
| YOLOv5s + CBAM | - | - | 0.763 | 0.721 | 0.779 | 0.805 | 0.792        |
|                | ✓ | - | 0.795 | 0.764 | 0.777 | 0.82  | 0.798        |
|                | - | ✓ | 0.835 | 0.726 | 0.781 | 0.844 | 0.812        |
|                | ✓ | ✓ | 0.71  | 0.812 | 0.787 | 0.854 | 0.82         |
| YOLOv5s + NAM  | - | - | 0.825 | 0.705 | 0.826 | 0.754 | 0.79         |
|                | ✓ | - | 0.88  | 0.675 | 0.814 | 0.786 | 0.80         |
|                | - | ✓ | 0.841 | 0.745 | 0.82  | 0.811 | 0.815        |
|                | ✓ | ✓ | 0.821 | 0.754 | 0.825 | 0.834 | 0.829        |
| YOLOv5s + CA   | - | - | 0.792 | 0.734 | 0.775 | 0.764 | 0.77         |
|                | ✓ | - | 0.84  | 0.72  | 0.795 | 0.79  | 0.793        |
|                | - | ✓ | 0.828 | 0.783 | 0.782 | 0.81  | 0.796        |
|                | ✓ | ✓ | 0.784 | 0.744 | 0.785 | 0.817 | 0.801        |
| YOLOv5s + EMA  | - | - | 0.773 | 0.749 | 0.815 | 0.798 | 0.806        |
|                | ✓ | - | 0.846 | 0.688 | 0.81  | 0.797 | 0.804        |
|                | - | ✓ | 0.856 | 0.757 | 0.849 | 0.848 | <b>0.848</b> |
|                | ✓ | ✓ | 0.771 | 0.813 | 0.807 | 0.825 | 0.816        |

Tablo 4.25’te sunulan Başkent veri kümesi üzerinde yapılan ince ayar sonuçları değerlendirildiğinde, en yüksek mAP@0.5 değeri 0.848 ile YOLOv5s + EMA modeli ve varsayılan + mixup veri artırma kombinasyonu kullanıldığında elde edilmiştir.

#### 4.5. YOLOv8 ve Dikkat Modülleri Analiz Sonuçları

YOLOv5 modelleri ile elde edilen analizlerin ardından, literatürdeki son gelişmeler doğrultusunda daha güncel ve gelişmiş mimariye sahip olan YOLOv8 modelleri değerlendirmeye alınmıştır. Bu bölümde öncelikle, yapılan ön analizlerde YOLOv8n, YOLOv8s ve YOLOv8m modelleri karşılaştırılmış; en yüksek performansı gösteren YOLOv8s temel model olarak seçilmiştir. Ardından, bu modele altı farklı dikkat modülü (SE, ECA, CBAM, NAM, CA ve EMA) entegre edilerek yeni model varyantları oluşturulmuştur. Elde edilen bu modellerin, meme kitlelerinin tespiti ve sınıflandırmasındaki başarısı incelenmiştir. Dikkat modülleri, modelin daha anlamlı ve odaklanmış öznelikler öğrenmesini sağlamak amacıyla, Şekil 4.13’te gösterildiği gibi YOLOv8 mimarisinin boyun katmanına entegre edilmiştir. Bu yapı sayesinde, boyun katmanından sınıflandırma kafasına aktarılan özneliklerin daha odaklanmış ve bilgi açısından anlamlı olması hedeflenmiştir.



Şekil 4.13. YOLOv8 + dikkat modülü model mimarisinin gösterimi.

YOLOv5’ten farklı olarak, YOLOv8 modeli için eğitimde kullanılan veri artırma parametreleri tek bir varsayılan konfigürasyon dosyasında tanımlanmıştır. Bu dosyada yer alan parametreler ve bunlara ait varsayılan değerler Tablo 4.26’da sunulmuştur. YOLOv8

için sıfırdan eğitim ve ince ayar süreçlerinde kullanılan hiperparametreler ise Tablo 4.27’de verilmiştir.

**Tablo 4.26.** YOLOv8 eğitiminde kullanılan varsayılan veri artırma konfigürasyonu

| <b>Parametre ve tanımı</b>                                       | <b>değeri</b> |
|--|---------------|
| <b>hsv_h:</b> görüntü HSV-ton artırma (kesir)                    | 0.015         |
| <b>hsv_s:</b> görüntü HSV-doygunluk artırma (kesir)              | 0.7           |
| <b>hsv_v:</b> görüntü HSV-değer artırma (kesir)                  | 0.4           |
| <b>degrees:</b> görüntü döndürme (+/- derece)                    | 0.0           |
| <b>translate:</b> görüntü kaydırma (+/-kesir)                    | 0.1           |
| <b>scale:</b> görüntü ölçekleme (+/- kazanç)                     | 0.5           |
| <b>shear:</b> görüntü kesme (+/- derece)                         | 0.0           |
| <b>perspective:</b> görüntü perspektifi (+/-kesir)               | 0.0           |
| <b>flipud:</b> görüntü yukarı-aşağı çevirme (olasılık)           | 0.0           |
| <b>fliplr:</b> görüntü sola-sağa çevirme (olasılık)              | 0.5           |
| <b>mosaic:</b> görüntü mozaiği (olasılık)                        | 1.0           |
| <b>mixup:</b> görüntü karıştırma (olasılık)                      | 0.0           |
| <b>copy_paste:</b> segmentasyon için kopyala-yapıştır (olasılık) | 0.0           |

**Tablo 4.27.** YOLOv8 için sıfırdan eğitim ve ince ayar süreçlerinde kullanılan hiperparametreler

| <b>hiperparametre</b>              | <b>değeri</b> |
|------------------------------------|---------------|
| görüntü boyutu                     | 544 x 544     |
| epok (sıfırdan eğitim / ince ayar) | 300 / 200     |
| toplu işlem boyutu                 | 8             |
| başlangıç öğrenme oranı            | 0.01          |
| optimizasyon algoritması           | AdamW         |
| momentum katsayısı                 | 0.937         |
| ağırlık azaltma katsayısı          | 0.0005        |
| erken durdurma değeri (epok)       | 100           |

İlk olarak, VinDr-Mammo, CBIS-DDSM ve bu iki veri kümesinin birleşiminden oluşan CBIS-DDSM + VinDr-Mammo veri kümesi, rastgele olacak şekilde %80 eğitim, %10 doğrulama ve %10 test alt kümelerine ayrılmıştır. Temel model olarak belirlenen YOLOv8s’e, altı farklı dikkat modülü (SE, ECA, CBAM, NAM, CA ve EMA) entegre

edilerek yeni model varyantları oluşturulmuştur. Bu modeller, üç veri kümesinde ayrı ayrı varsayılan veri artırma konfigürasyonu ile sıfırdan eğitilmiştir. Elde edilen sıfırdan eğitim performans sonuçları Tablo 4.28, Tablo 4.29 ve Tablo 4.30’da sunulmuştur.

**Tablo 4.28.** CBIS-DDSM veri kümesinde YOLOv8s ve YOLOv8s + dikkat modüllerinin sıfırdan eğitim performans sonuçları

| <b>Model</b>   | <b>Kesinlik</b> | <b>Duyarlılık</b> | <b>B AP</b> | <b>MAP</b> | <b>mAP@0.5</b> |
|----------------|-----------------|-------------------|-------------|------------|----------------|
| YOLOv8s        | 0.428           | 0.483             | 0.343       | 0.466      | 0.404          |
| YOLOv8s + SE   | 0.57            | 0.386             | 0.348       | 0.479      | 0.413          |
| YOLOv8s + ECA  | 0.481           | 0.498             | 0.36        | 0.551      | <b>0.455</b>   |
| YOLOv8s + CBAM | 0.511           | 0.451             | 0.385       | 0.497      | 0.441          |
| YOLOv8s + NAM  | 0.502           | 0.402             | 0.349       | 0.494      | 0.422          |
| YOLOv8s + CA   | 0.475           | 0.414             | 0.311       | 0.47       | 0.391          |
| YOLOv8s + EMA  | 0.569           | 0.422             | 0.364       | 0.52       | 0.442          |

**Tablo 4.29.** VinDr-Mammo veri kümesinde YOLOv8s ve YOLOv8s + dikkat modüllerinin sıfırdan eğitim performans sonuçları

| <b>Model</b>   | <b>Kesinlik</b> | <b>Duyarlılık</b> | <b>B AP</b> | <b>MAP</b> | <b>mAP@0.5</b> |
|----------------|-----------------|-------------------|-------------|------------|----------------|
| YOLOv8s        | 0.502           | 0.458             | 0.389       | 0.495      | 0.442          |
| YOLOv8s + SE   | 0.532           | 0.493             | 0.459       | 0.433      | 0.446          |
| YOLOv8s + ECA  | 0.481           | 0.591             | 0.493       | 0.487      | 0.49           |
| YOLOv8s + CBAM | 0.494           | 0.504             | 0.56        | 0.444      | <b>0.502</b>   |
| YOLOv8s + NAM  | 0.486           | 0.423             | 0.423       | 0.422      | 0.423          |
| YOLOv8s + CA   | 0.558           | 0.523             | 0.513       | 0.477      | 0.495          |
| YOLOv8s + EMA  | 0.492           | 0.477             | 0.461       | 0.428      | 0.44           |

**Tablo 4.30.** CBIS-DDSM + VinDr-Mammo veri kümesinde YOLOv8s ve YOLOv8s + dikkat modüllerinin sıfırdan eğitim performans sonuçları

| <b>Model</b>  | <b>Kesinlik</b> | <b>Duyarlılık</b> | <b>B AP</b> | <b>MAP</b> | <b>mAP@0.5</b> |
|---------------|-----------------|-------------------|-------------|------------|----------------|
| YOLOv8s       | 0.524           | 0.466             | 0.43        | 0.536      | 0.483          |
| YOLOv8s + SE  | 0.572           | 0.486             | 0.459       | 0.549      | 0.504          |
| YOLOv8s + ECA | 0.517           | 0.49              | 0.444       | 0.568      | 0.506          |

**Tablo 4.30.** devam ediyor.

|                |       |       |       |       |              |
|----------------|-------|-------|-------|-------|--------------|
| YOLOv8s + CBAM | 0.549 | 0.475 | 0.449 | 0.566 | <b>0.507</b> |
| YOLOv8s + NAM  | 0.497 | 0.517 | 0.464 | 0.483 | 0.473        |
| YOLOv8s + CA   | 0.496 | 0.49  | 0.404 | 0.561 | 0.483        |
| YOLOv8s + EMA  | 0.575 | 0.483 | 0.429 | 0.521 | 0.475        |

Tablo 4.28, Tablo 4.29 ve Tablo 4.30’da sunulan sıfırdan eğitim sonuçlarına göre;

- YOLOv8s modeline entegre edilen dikkat modülleri, tespit ve sınıflandırma başarısını çoğu durumda artırmıştır.
- CBIS-DDSM veri kümesi için en yüksek başarı ( $mAP@0.5=0.455$ ), YOLOv8s + ECA modeli ile elde edilmiştir.
- VinDr-Mammo veri kümesinde, en başarılı performans ( $mAP@0.5=0.502$ ) YOLOv8s + CBAM modeline aittir.
- CBIS-DDSM + VinDr-Mammo birleşim veri kümesinde ise en yüksek başarı ( $mAP@0.5=0.507$ ) YOLOv8s + CBAM modeli ile sağlanmıştır.

Ardından, eğitilen modellerin en iyi ağırlıkları kullanılarak INBreast veri kümesi üzerinde ince ayar gerçekleştirilmiştir. İnce ayar sırasında, INBreast veri kümesinin küçük boyutlu olmasından dolayı tüm veri kümesini değerlendirmek için 5 katlı çapraz doğrulama stratejisi kullanılmıştır. Bu aşamada, YOLOv8 eğitiminde kullanılan varsayılan veri artırma parametrelerine ek olarak, daha önce kullanılmayan bazı parametrelerin etkisi de analizlere dahil edilerek modellerin performansının artırılması hedeflenmiştir. Bu kapsamda, görüntü yukarı-aşağı çevirme ( $flipud = 0.5$ ) ve görüntü karıştırma ( $mixup = 0.1$ ) parametreleri uygulanmıştır. Daha önce YOLOv5 modelleriyle yapılan analizlerde olduğu gibi, degrees, shear ve perspective gibi parametrelerin etkili olmadığı gözlemlendiğinden, bu parametreler YOLOv8 ince ayar sürecinde de değerlendirmeye alınmamıştır. Yapılan ince ayar analizleri ve elde edilen sonuçlar sırasıyla Tablo 4.31, Tablo 4.32 ve Tablo 4.33’te sunulmuştur.

**Tablo 4.31.** CBIS-DDSM → INBreast: YOLOv8s ve YOLOv8s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model         | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5 |
|---------------|--------------------------|-------------|----------|------------|-------|-------|---------|
|               | flipud (0.5)             | mixup (0.1) |          |            |       |       |         |
| YOLOv8s       | -                        | -           | 0.79     | 0.698      | 0.752 | 0.832 | 0.792   |
|               | ✓                        | -           | 0.826    | 0.711      | 0.729 | 0.883 | 0.806   |
|               | -                        | ✓           | 0.834    | 0.675      | 0.713 | 0.859 | 0.786   |
|               | ✓                        | ✓           | 0.819    | 0.719      | 0.743 | 0.876 | 0.81    |
| YOLOv8s + SE  | -                        | -           | 0.806    | 0.727      | 0.746 | 0.858 | 0.802   |
|               | ✓                        | -           | 0.874    | 0.724      | 0.722 | 0.885 | 0.804   |
|               | -                        | ✓           | 0.828    | 0.759      | 0.742 | 0.882 | 0.812   |
|               | ✓                        | ✓           | 0.82     | 0.757      | 0.749 | 0.866 | 0.808   |
| YOLOv8s + ECA | -                        | -           | 0.769    | 0.747      | 0.686 | 0.90  | 0.793   |
|               | ✓                        | -           | 0.79     | 0.72       | 0.656 | 0.888 | 0.772   |
|               | -                        | ✓           | 0.815    | 0.757      | 0.764 | 0.877 | 0.82    |
|               | ✓                        | ✓           | 0.75     | 0.753      | 0.684 | 0.883 | 0.784   |

Tablo 4.31. devam ediyor.

|                |   |   |       |       |       |       |              |
|----------------|---|---|-------|-------|-------|-------|--------------|
| YOLOv8s + CBAM | - | - | 0.874 | 0.728 | 0.728 | 0.875 | 0.802        |
|                | ✓ | - | 0.795 | 0.778 | 0.722 | 0.829 | 0.776        |
|                | - | ✓ | 0.848 | 0.738 | 0.762 | 0.903 | <b>0.833</b> |
|                | ✓ | ✓ | 0.859 | 0.734 | 0.727 | 0.875 | 0.801        |
| YOLOv8s + NAM  | - | - | 0.737 | 0.765 | 0.64  | 0.867 | 0.754        |
|                | ✓ | - | 0.804 | 0.688 | 0.659 | 0.877 | 0.768        |
|                | - | ✓ | 0.784 | 0.695 | 0.642 | 0.865 | 0.754        |
|                | ✓ | ✓ | 0.802 | 0.71  | 0.704 | 0.892 | 0.80         |
| YOLOv8s + CA   | - | - | 0.864 | 0.718 | 0.713 | 0.876 | 0.794        |
|                | ✓ | - | 0.844 | 0.734 | 0.764 | 0.855 | 0.81         |
|                | - | ✓ | 0.827 | 0.735 | 0.74  | 0.872 | 0.806        |
|                | ✓ | ✓ | 0.813 | 0.768 | 0.773 | 0.876 | 0.825        |
| YOLOv8s + EMA  | - | - | 0.767 | 0.689 | 0.65  | 0.845 | 0.747        |
|                | ✓ | - | 0.757 | 0.694 | 0.644 | 0.88  | 0.762        |
|                | - | ✓ | 0.772 | 0.747 | 0.725 | 0.861 | 0.793        |
|                | ✓ | ✓ | 0.795 | 0.70  | 0.656 | 0.892 | 0.774        |

**Tablo 4.32.** VinDr-Mammo → INBreast: YOLOv8s ve YOLOv8s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model         | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5 |
|---------------|--------------------------|-------------|----------|------------|-------|-------|---------|
|               | flipud (0.5)             | mixup (0.1) |          |            |       |       |         |
| YOLOv8s       | -                        | -           | 0.75     | 0.691      | 0.702 | 0.853 | 0.778   |
|               | ✓                        | -           | 0.811    | 0.688      | 0.712 | 0.862 | 0.787   |
|               | -                        | ✓           | 0.798    | 0.682      | 0.677 | 0.864 | 0.77    |
|               | ✓                        | ✓           | 0.8      | 0.693      | 0.728 | 0.844 | 0.786   |
| YOLOv8s + SE  | -                        | -           | 0.816    | 0.75       | 0.707 | 0.843 | 0.775   |
|               | ✓                        | -           | 0.849    | 0.704      | 0.711 | 0.837 | 0.774   |
|               | -                        | ✓           | 0.783    | 0.727      | 0.709 | 0.85  | 0.779   |
|               | ✓                        | ✓           | 0.811    | 0.756      | 0.726 | 0.875 | 0.8     |
| YOLOv8s + ECA | -                        | -           | 0.816    | 0.75       | 0.707 | 0.843 | 0.775   |
|               | ✓                        | -           | 0.846    | 0.67       | 0.697 | 0.853 | 0.775   |
|               | -                        | ✓           | 0.778    | 0.714      | 0.692 | 0.876 | 0.784   |
|               | ✓                        | ✓           | 0.816    | 0.70       | 0.743 | 0.858 | 0.80    |

**Tablo 4.32.** devam ediyor.

|                |   |   |       |       |       |       |              |
|----------------|---|---|-------|-------|-------|-------|--------------|
| YOLOv8s + CBAM | - | - | 0.86  | 0.701 | 0.714 | 0.872 | 0.793        |
|                | ✓ | - | 0.789 | 0.746 | 0.729 | 0.889 | <b>0.809</b> |
|                | - | ✓ | 0.865 | 0.723 | 0.717 | 0.873 | 0.795        |
|                | ✓ | ✓ | 0.863 | 0.672 | 0.687 | 0.886 | 0.787        |
| YOLOv8s + NAM  | - | - | 0.772 | 0.691 | 0.661 | 0.832 | 0.747        |
|                | ✓ | - | 0.822 | 0.674 | 0.637 | 0.874 | 0.755        |
|                | - | ✓ | 0.74  | 0.712 | 0.627 | 0.851 | 0.739        |
|                | ✓ | ✓ | 0.744 | 0.712 | 0.652 | 0.753 | 0.753        |
| YOLOv8s + CA   | - | - | 0.745 | 0.754 | 0.71  | 0.856 | 0.783        |
|                | ✓ | - | 0.826 | 0.712 | 0.691 | 0.873 | 0.782        |
|                | - | ✓ | 0.831 | 0.714 | 0.716 | 0.871 | 0.793        |
|                | ✓ | ✓ | 0.794 | 0.736 | 0.707 | 0.878 | 0.792        |
| YOLOv8s + EMA  | - | - | 0.805 | 0.68  | 0.70  | 0.859 | 0.778        |
|                | ✓ | - | 0.80  | 0.734 | 0.715 | 0.848 | 0.781        |
|                | - | ✓ | 0.824 | 0.738 | 0.742 | 0.867 | 0.805        |
|                | ✓ | ✓ | 0.81  | 0.734 | 0.736 | 0.868 | 0.802        |

**Tablo 4.33.** CBIS-DDSM + VinDr-Mammo → INBreast: YOLOv8s ve YOLOv8s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model         | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5 |
|---------------|--------------------------|-------------|----------|------------|-------|-------|---------|
|               | flipud (0.5)             | mixup (0.1) |          |            |       |       |         |
| YOLOv8s       | -                        | -           | 0.774    | 0.739      | 0.693 | 0.865 | 0.779   |
|               | ✓                        | -           | 0.752    | 0.724      | 0.711 | 0.846 | 0.779   |
|               | -                        | ✓           | 0.861    | 0.755      | 0.755 | 0.877 | 0.816   |
|               | ✓                        | ✓           | 0.798    | 0.738      | 0.706 | 0.866 | 0.786   |
| YOLOv8s + SE  | -                        | -           | 0.821    | 0.705      | 0.717 | 0.864 | 0.789   |
|               | ✓                        | -           | 0.827    | 0.757      | 0.749 | 0.882 | 0.816   |
|               | -                        | ✓           | 0.936    | 0.679      | 0.746 | 0.894 | 0.82    |
|               | ✓                        | ✓           | 0.836    | 0.755      | 0.751 | 0.882 | 0.816   |
| YOLOv8s + ECA | -                        | -           | 0.81     | 0.693      | 0.742 | 0.839 | 0.79    |
|               | ✓                        | -           | 0.877    | 0.702      | 0.734 | 0.866 | 0.8     |
|               | -                        | ✓           | 0.832    | 0.683      | 0.754 | 0.851 | 0.802   |
|               | ✓                        | ✓           | 0.775    | 0.812      | 0.776 | 0.846 | 0.811   |

Tablo 4.33. devam ediyor.

|                |   |   |       |       |       |       |              |
|----------------|---|---|-------|-------|-------|-------|--------------|
| YOLOv8s + CBAM | - | - | 0.816 | 0.738 | 0.708 | 0.877 | 0.792        |
|                | ✓ | - | 0.772 | 0.767 | 0.786 | 0.858 | 0.822        |
|                | - | ✓ | 0.788 | 0.81  | 0.73  | 0.882 | 0.806        |
|                | ✓ | ✓ | 0.805 | 0.789 | 0.77  | 0.904 | <b>0.837</b> |
| YOLOv8s + NAM  | - | - | 0.796 | 0.792 | 0.755 | 0.882 | 0.818        |
|                | ✓ | - | 0.911 | 0.682 | 0.752 | 0.881 | 0.817        |
|                | - | ✓ | 0.825 | 0.786 | 0.775 | 0.896 | 0.836        |
|                | ✓ | ✓ | 0.867 | 0.744 | 0.803 | 0.869 | 0.836        |
| YOLOv8s + CA   | - | - | 0.78  | 0.701 | 0.678 | 0.882 | 0.78         |
|                | ✓ | - | 0.846 | 0.705 | 0.723 | 0.881 | 0.802        |
|                | - | ✓ | 0.823 | 0.675 | 0.71  | 0.857 | 0.783        |
|                | ✓ | ✓ | 0.847 | 0.686 | 0.727 | 0.884 | 0.805        |
| YOLOv8s + EMA  | - | - | 0.778 | 0.684 | 0.675 | 0.845 | 0.76         |
|                | ✓ | - | 0.754 | 0.701 | 0.698 | 0.809 | 0.754        |
|                | - | ✓ | 0.781 | 0.721 | 0.712 | 0.88  | 0.798        |
|                | ✓ | ✓ | 0.871 | 0.73  | 0.718 | 0.887 | 0.803        |

Tablo 4.31, Tablo 4.32 ve Tablo 4.33 incelendiğinde;

- CBIS-DDSM → INBreast için 0.833'lük en yüksek mAP@0.5 değeri, YOLOv8s + CBAM modeli ve varsayılan + mixup veri artırma kombinasyonu ile elde edilmiştir.
- VinDr-Mammo → INBreast için en yüksek mAP@0.5 değeri 0.809 olup, bu başarı YOLOv8s + CBAM modeli ve varsayılan + flipud veri artırma kombinasyonunda gözlenmiştir.
- CBIS-DDSM + VinDr-Mammo → INBreast için en yüksek mAP@0.5 değeri olan 0.837, YOLOv8s + CBAM modeli ve varsayılan + flipud + mixup veri artırma kombinasyonu ile elde edilmiştir.
- Genel olarak değerlendirildiğinde, YOLOv8s modeline entegre edilen CBAM dikkat modülü, CBIS-DDSM, VinDr-Mammo ve birleşim veri kümeleri için en yüksek mAP@0.5 skorlarını sağlamıştır. Ayrıca, flipud ve mixup gibi ek veri artırma parametrelerinin dahil edilmesi, özellikle CBAM modeliyle birlikte kullanıldığında model performansını belirgin şekilde artırmıştır.

YOLOv8s + dikkat modülleriyle yapılan analizlerde, en yüksek ince ayar performansı birleşim veri kümesiyle eğitilen modeller tarafından elde edilmiştir. Bu nedenle, ince ayar sürecinde yalnızca birleşim veri kümesi ile eğitilmiş modeller kullanılarak klinik veri kümesi olan Başkent veri kümesi üzerinde ince ayar gerçekleştirilmiştir. Elde edilen ince ayar sonuçları Tablo 4.34'te sunulmuştur.

**Tablo 4.34.** CBIS-DDSM + VinDr-Mammo → Başkent: YOLOv8s ve YOLOv8s + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model         | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5 |
|---------------|--------------------------|-------------|----------|------------|-------|-------|---------|
|               | flipud (0.5)             | mixup (0.1) |          |            |       |       |         |
| YOLOv8s       | -                        | -           | 0.826    | 0.662      | 0.77  | 0.748 | 0.759   |
|               | ✓                        | -           | 0.756    | 0.68       | 0.765 | 0.75  | 0.759   |
|               | -                        | ✓           | 0.856    | 0.66       | 0.73  | 0.83  | 0.779   |
|               | ✓                        | ✓           | 0.684    | 0.732      | 0.755 | 0.787 | 0.771   |
| YOLOv8s + SE  | -                        | -           | 0.821    | 0.709      | 0.789 | 0.825 | 0.807   |
|               | ✓                        | -           | 0.78     | 0.75       | 0.79  | 0.803 | 0.797   |
|               | -                        | ✓           | 0.794    | 0.722      | 0.767 | 0.83  | 0.799   |
|               | ✓                        | ✓           | 0.791    | 0.747      | 0.773 | 0.816 | 0.795   |
| YOLOv8s + ECA | -                        | -           | 0.82     | 0.721      | 0.828 | 0.733 | 0.781   |
|               | ✓                        | -           | 0.76     | 0.761      | 0.788 | 0.817 | 0.803   |
|               | -                        | ✓           | 0.80     | 0.669      | 0.782 | 0.77  | 0.776   |
|               | ✓                        | ✓           | 0.728    | 0.786      | 0.777 | 0.833 | 0.805   |

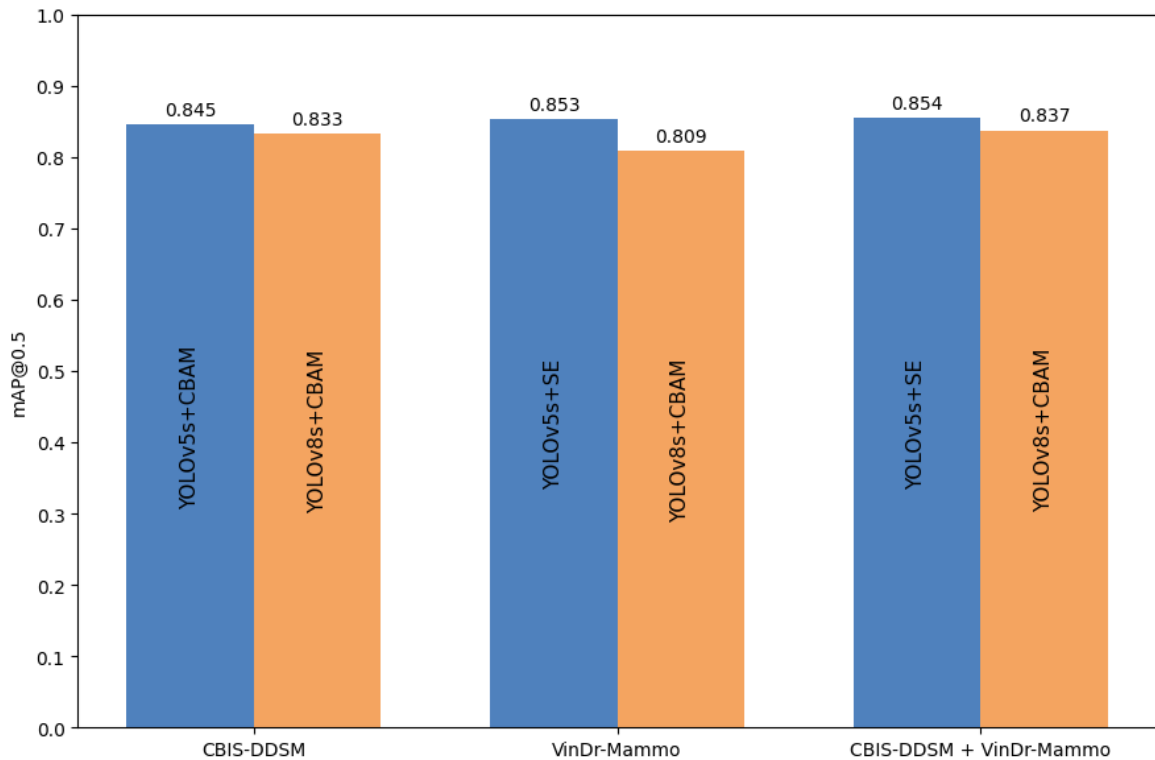
**Tablo 4.34.** devam ediyor.

|                |   |   |       |       |       |       |              |
|----------------|---|---|-------|-------|-------|-------|--------------|
| YOLOv8s + CBAM | - | - | 0.831 | 0.71  | 0.781 | 0.81  | 0.80         |
|                | ✓ | - | 0.78  | 0.762 | 0.772 | 0.83  | 0.80         |
|                | - | ✓ | 0.808 | 0.71  | 0.76  | 0.82  | 0.79         |
|                | ✓ | ✓ | 0.787 | 0.775 | 0.781 | 0.814 | 0.80         |
| YOLOv8s + NAM  | - | - | 0.75  | 0.685 | 0.792 | 0.735 | 0.763        |
|                | ✓ | - | 0.743 | 0.769 | 0.831 | 0.813 | <b>0.822</b> |
|                | - | ✓ | 0.722 | 0.692 | 0.737 | 0.777 | 0.757        |
|                | ✓ | ✓ | 0.726 | 0.745 | 0.769 | 0.793 | 0.781        |
| YOLOv8s + CA   | - | - | 0.797 | 0.68  | 0.762 | 0.78  | 0.771        |
|                | ✓ | - | 0.758 | 0.717 | 0.796 | 0.79  | 0.794        |
|                | - | ✓ | 0.769 | 0.716 | 0.773 | 0.769 | 0.77         |
|                | ✓ | ✓ | 0.746 | 0.731 | 0.762 | 0.797 | 0.779        |
| YOLOv8s + EMA  | - | - | 0.748 | 0.718 | 0.809 | 0.771 | 0.79         |
|                | ✓ | - | 0.75  | 0.727 | 0.764 | 0.828 | 0.796        |
|                | - | ✓ | 0.811 | 0.69  | 0.778 | 0.79  | 0.784        |
|                | ✓ | ✓ | 0.781 | 0.76  | 0.77  | 0.82  | 0.795        |

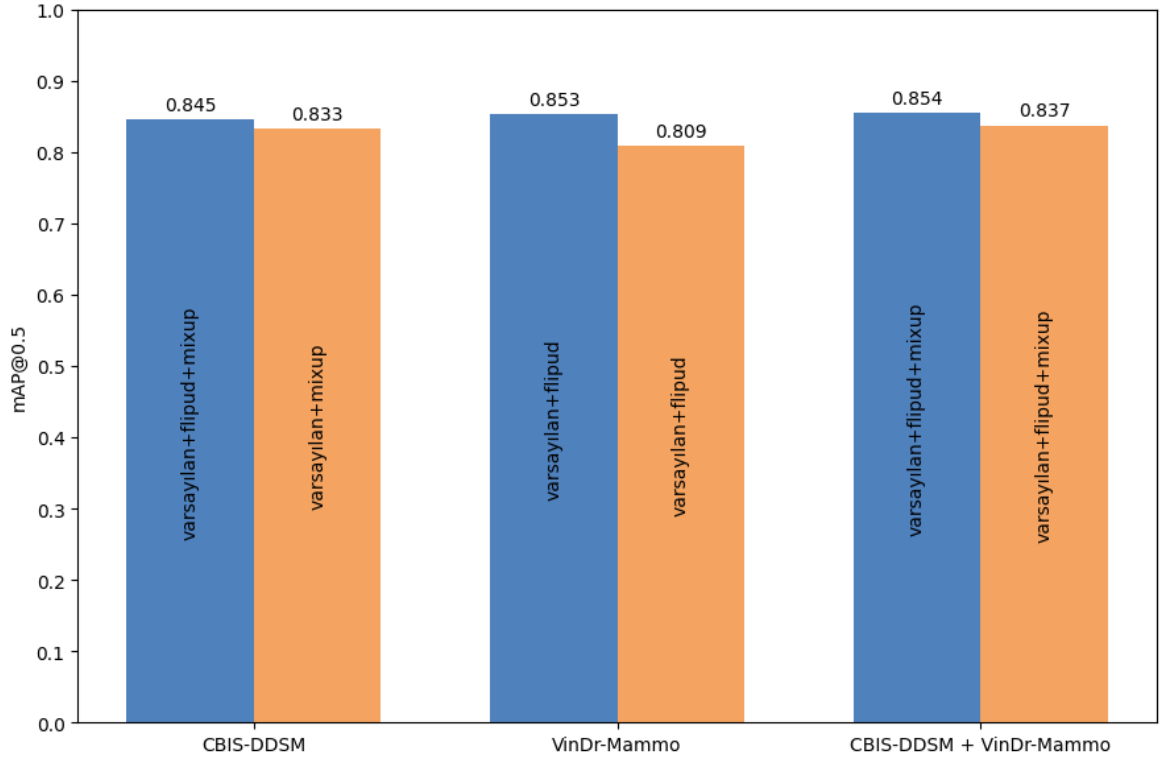
Tablo 4.34’te sunulan Başkent veri kümesi üzerinde yapılan ince ayar sonuçları değerlendirildiğinde, en yüksek  $mAP@0.5$  değeri 0.822 ile YOLOv8s + NAM modeli ve varsayılan + flipud veri artırma kombinasyonu kullanıldığında elde edilmiştir.

#### 4.6. YOLOv5 ve YOLOv8 Modellerine Entegre Edilen Dikkat Modüllerinin Performans Karşılaştırması

Bu bölümde, YOLOv5s ve YOLOv8s modellerine entegre edilen farklı dikkat modülleri ile oluşturulan model varyantlarının, kitle tespiti ve sınıflandırma performansı üzerindeki etkileri değerlendirilmiştir. Her iki model yapısı, farklı kaynak veri kümeleri (CBIS-DDSM, VinDr-Mammo ve CBIS-DDSM + VinDr-Mammo) ile sıfırdan eğitilmiş ve ardından INBreast veri kümesi üzerinde ince ayar uygulanarak elde edilen en yüksek başarılar analiz edilmiştir. Karşılaştırmalar, her bir veri kümesinde iyi sonucu veren model varyantını ve buna karşılık gelen veri artırma parametre kombinasyonunu belirlemek amacıyla,  $mAP@0.5$  performans ölçütü üzerinden gerçekleştirilmiştir. Elde edilen karşılaştırmalı sonuçlar sırasıyla Şekil 4.14 ve Şekil 4.15’te görselleştirilmiştir.



Şekil 4.14. Farklı kaynak veri kümeleriyle eğitilen YOLO + dikkat modülü model varyantlarının INBreast üzerindeki performans karşılaştırılması.



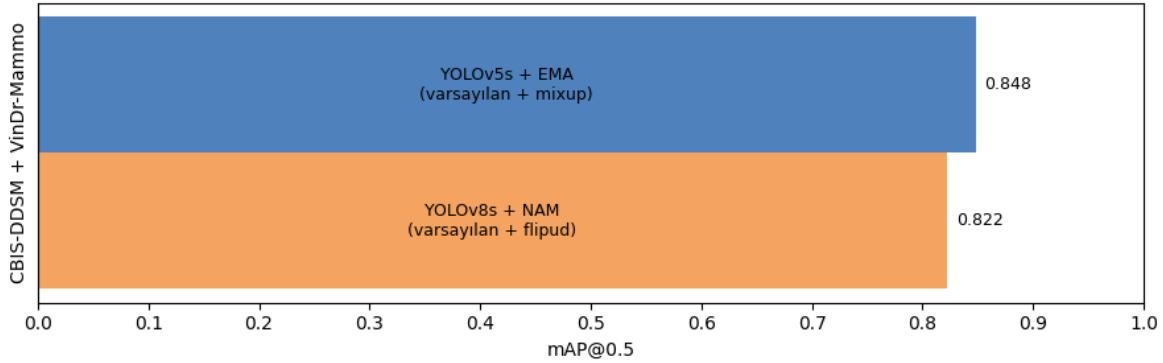
Şekil 4.15. En başarılı YOLO + dikkat modülü model varyantlarında kullanılan veri artırma parametre kombinasyonlarının performans üzerindeki etkisi.

Şekil 4.14 ve Şekil 4.15'te görüldüğü gibi;

- Üç farklı eğitim senaryosu kapsamında, YOLOv5s + dikkat modülü model varyantları, YOLOv8s + dikkat modülü model varyantlarına kıyasla göre daha yüksek performans göstermiştir.
- YOLOv5s modeli için en etkili dikkat modülü SE olurken, YOLOv8s modeli için tüm senaryolarda en başarılı ve tutarlı performans CBAM modülü ile elde edilmiştir.
- Her iki model yapısı için de en iyi performans, birleşim veri kümesi (CBIS-DDSM + VinDr-Mammo) üzerinden yapılan ince ayar ile elde edilmiştir. Bu durum, kaynak veri kümesinin sahip olduğu veri sayısı ve çeşitliliğinin model performansı üzerinde önemli bir etkisi olduğunu göstermektedir.
- Veri artırma parametreleri açısından değerlendirildiğinde, her iki model için en başarılı performansa varsayılan + flipud + mixup kombinasyonu ile ulaşılmıştır. Özellikle sınırlı veriyle gerçekleştirilen ince ayar analizlerinde, veri artırma parametreleri model performansını iyileştirmede önemli bir rol oynadığı gözlemlenmiştir.

En başarılı eğitim senaryosu, CBIS-DDSM ve VinDr-Mammo birleşim veri kümesi ile elde edildiği için, Başkent veri kümesi üzerinde yalnızca bu birleşim veri kümesi

kullanılarak ince ayar gerçekleştirilmiştir. Şekil 4.16’da en yüksek performans gösteren model varyantları ve bunlara ait veri artırma parametre kombinasyonları karşılaştırmalı olarak sunulmuştur.



**Şekil 4.16.** Başkent veri kümesi üzerinde en başarılı YOLO + dikkat modülü model varyantlarının ve veri artırma kombinasyonlarının performans karşılaştırılması.

Şekil 4.16’da görüldüğü gibi;

- YOLOv5s + dikkat modülü model varyantı, YOLOv8s + dikkat modülü model varyantına kıyasla göre daha yüksek performans göstermiştir.
- YOLOv5s modeli için en başarılı dikkat modülü EMA olurken, YOLOv8s modeli için NAM modülü öne çıkmıştır.
- Veri artırma parametreleri açısından değerlendirildiğinde, her iki model yapısı için en yüksek performansa farklı veri artırma kombinasyonları ile ulaşıldığı gözlemlenmiştir. Bu durum, veri artırma parametrelerinin model mimarisine bağlı olarak performans üzerindeki etkisinin değişebileceğini ortaya koymaktadır.
- Bununla birlikte, INBreast ve Başkent veri kümeleri üzerinde yapılan ince ayar analizleri, her iki model yapısı için de en başarılı dikkat modülü ve veri artırma kombinasyonlarının değiştiğini göstermektedir.

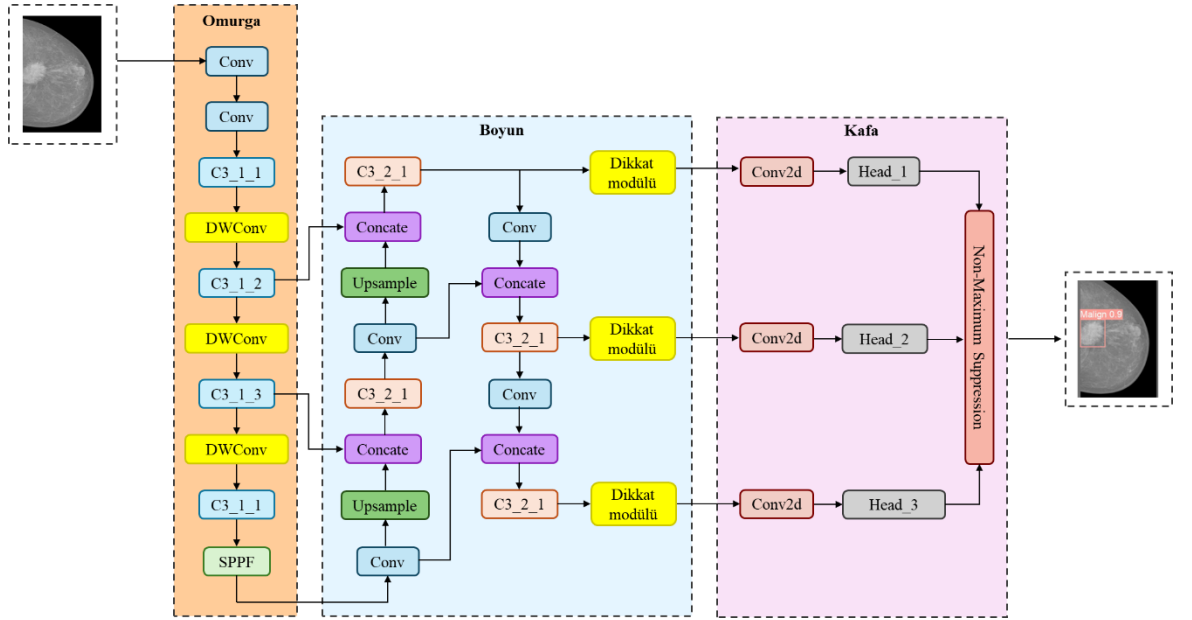
#### 4.7. Dikkat Modülü Entegre Edilmiş YOLOv5 ve YOLOv8 Modellerinde DWConv Uygulaması

YOLOv5s ve YOLOv8s modellerine entegre edilen farklı dikkat modülleri ile yapılan analizlerde, en yüksek performans CBIS-DDSM + VinDr-Mammo birleşim veri kümesi kullanılarak INBreast veri kümesi üzerinde gerçekleştirilen ince ayar sonucunda elde edilmiştir. Bu nedenle, bu bölümde söz konusu ince ayar yaklaşımının performansını artırmak amacıyla YOLOv5s ve YOLOv8s modellerinin omurga katmanlarına DWConv

uygulanmış ve bu yapının model performansı üzerindeki etkisi analiz edilmiştir. Bu doğrultuda elde edilen yeni modeller, sırasıyla YOLOv5s + DWConv + dikkat modülü ve YOLOv8s + DWConv + dikkat modülü olarak adlandırılmıştır.

#### 4.7.1. YOLOv5 + dikkat modülü modellerinde DWConv uygulaması

Şekil 4.17’de gösterildiği üzere, bu bölümde YOLOv5s + dikkat modülü model varyantlarının performansını iyileştirmek amacıyla, modelin omurga katmanında yer alan son üç standart evrişim (Conv) bloğu, DWConv blokları ile değiştirilmiştir. Bu yapısal değişiklik ile, modelin parametre verimliliği artırılırken, dikkat modülleriyle birlikte daha etkili öznitelik çıkarımı sağlanması hedeflenmiştir.



Şekil 4.17. YOLOv5 + DWConv + dikkat modülü model mimarisinin gösterimi.

Bu analiz kapsamında yalnızca CBIS-DDSM + VinDr-Mammo birleşim veri kümesi kullanılarak önce sıfırdan eğitim gerçekleştirilmiş, ardından tüm model varyantları ve veri artırma parametre kombinasyonları için INBreast veri kümesi üzerinde ince ayar analizleri uygulanmıştır. Sıfırdan eğitim sonuçları Tablo 4.35’te, INBreast ve Başkent veri kümeleri için gerçekleştirilen ince ayar sonuçları ise sırasıyla Tablo 4.36 ve Tablo 4.37’de sunulmuştur.

**Tablo 4.35.** CBIS-DDSM + VinDr-Mammo veri kümesinde YOLOv5s + DWConv + dikkat modülleri sıfırdan eğitim performans sonuçları

| <b>Model</b>            | <b>Kesinlik</b> | <b>Duyarlılık</b> | <b>B AP</b> | <b>M AP</b> | <b>mAP@0.5</b> |
|-------------------------|-----------------|-------------------|-------------|-------------|----------------|
| YOLOv5s + DWConv + SE   | 0.598           | 0.459             | 0.432       | 0.548       | 0.49           |
| YOLOv5s + DWConv + ECA  | 0.477           | 0.537             | 0.441       | 0.54        | <b>0.491</b>   |
| YOLOv5s + DWConv + CBAM | 0.446           | 0.54              | 0.413       | 0.527       | 0.47           |
| YOLOv5s + DWConv + NAM  | 0.568           | 0.482             | 0.439       | 0.536       | 0.488          |
| YOLOv5s + DWConv + CA   | 0.451           | 0.512             | 0.401       | 0.538       | 0.469          |
| YOLOv5s + DWConv + EMA  | 0.518           | 0.521             | 0.426       | 0.525       | 0.476          |

Tablo 4.35'e göre, YOLOv5s + dikkat modülü model varyantlarına DWConv uygulanması, genel olarak mAP@0.5 performansında bir düşüğe neden olmuştur. Bununla birlikte, yalnızca YOLOv5s + ECA modeli, DWConv uygulanmadan önce 0.481 mAP@0.5 değeri elde etmişken (Bkz. Tablo 4.21), DWConv sonrası bu değer 0.491'e yükselmiştir.

**Tablo 4.36.** CBIS-DDSM + VinDr-Mammo → INBreast: YOLOv5s + DWConv + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model                   | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5 |
|-------------------------|--------------------------|-------------|----------|------------|-------|-------|---------|
|                         | flipud (0.5)             | mixup (0.1) |          |            |       |       |         |
| YOLOv5s + DWConv + SE   | -                        | -           | 0.863    | 0.695      | 0.719 | 0.868 | 0.794   |
|                         | ✓                        | -           | 0.78     | 0.767      | 0.716 | 0.894 | 0.805   |
|                         | -                        | ✓           | 0.82     | 0.766      | 0.719 | 0.889 | 0.804   |
|                         | ✓                        | ✓           | 0.845    | 0.754      | 0.749 | 0.895 | 0.822   |
| YOLOv5s + DWConv + ECA  | -                        | -           | 0.877    | 0.655      | 0.747 | 0.865 | 0.806   |
|                         | ✓                        | -           | 0.813    | 0.694      | 0.692 | 0.864 | 0.779   |
|                         | -                        | ✓           | 0.837    | 0.725      | 0.724 | 0.881 | 0.802   |
|                         | ✓                        | ✓           | 0.808    | 0.715      | 0.704 | 0.884 | 0.794   |
| YOLOv5s + DWConv + CBAM | -                        | -           | 0.859    | 0.712      | 0.68  | 0.885 | 0.783   |
|                         | ✓                        | -           | 0.888    | 0.65       | 0.663 | 0.894 | 0.779   |
|                         | -                        | ✓           | 0.767    | 0.739      | 0.673 | 0.908 | 0.79    |
|                         | ✓                        | ✓           | 0.812    | 0.721      | 0.676 | 0.906 | 0.791   |

Tablo 4.36. devam ediyor.

|                        |   |   |       |       |       |       |              |
|------------------------|---|---|-------|-------|-------|-------|--------------|
| YOLOv5s + DWConv + NAM | - | - | 0.786 | 0.738 | 0.679 | 0.866 | 0.772        |
|                        | ✓ | - | 0.851 | 0.723 | 0.72  | 0.905 | 0.812        |
|                        | - | ✓ | 0.777 | 0.706 | 0.702 | 0.879 | 0.79         |
|                        | ✓ | ✓ | 0.864 | 0.737 | 0.735 | 0.913 | <b>0.824</b> |
| YOLOv5s + DWConv + CA  | - | - | 0.786 | 0.703 | 0.647 | 0.881 | 0.765        |
|                        | ✓ | - | 0.822 | 0.686 | 0.706 | 0.848 | 0.777        |
|                        | - | ✓ | 0.87  | 0.665 | 0.658 | 0.877 | 0.768        |
|                        | ✓ | ✓ | 0.807 | 0.722 | 0.647 | 0.90  | 0.774        |
| YOLOv5s + DWConv + EMA | - | - | 0.858 | 0.704 | 0.663 | 0.895 | 0.779        |
|                        | ✓ | - | 0.825 | 0.737 | 0.749 | 0.859 | 0.804        |
|                        | - | ✓ | 0.775 | 0.746 | 0.683 | 0.877 | 0.78         |
|                        | ✓ | ✓ | 0.888 | 0.723 | 0.719 | 0.89  | 0.805        |

**Tablo 4.37.** CBIS-DDSM + VinDr-Mammo → Başkent: YOLOv5s + DWConv + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model                   | Veri artırma parametresi |             | Kesinlik | Duyarlılık | BAP   | MAP   | mAP@0.5 |
|-------------------------|--------------------------|-------------|----------|------------|-------|-------|---------|
|                         | flipud (0.5)             | mixup (0.1) |          |            |       |       |         |
| YOLOv5s + DWConv + SE   | -                        | -           | 0.805    | 0.696      | 0.782 | 0.797 | 0.789   |
|                         | ✓                        | -           | 0.806    | 0.727      | 0.755 | 0.81  | 0.783   |
|                         | -                        | ✓           | 0.775    | 0.759      | 0.815 | 0.805 | 0.81    |
|                         | ✓                        | ✓           | 0.768    | 0.73       | 0.763 | 0.821 | 0.792   |
| YOLOv5s + DWConv + ECA  | -                        | -           | 0.792    | 0.664      | 0.749 | 0.76  | 0.755   |
|                         | ✓                        | -           | 0.732    | 0.698      | 0.77  | 0.779 | 0.774   |
|                         | -                        | ✓           | 0.769    | 0.705      | 0.743 | 0.791 | 0.767   |
|                         | ✓                        | ✓           | 0.775    | 0.698      | 0.741 | 0.788 | 0.765   |
| YOLOv5s + DWConv + CBAM | -                        | -           | 0.776    | 0.768      | 0.824 | 0.807 | 0.816   |
|                         | ✓                        | -           | 0.754    | 0.777      | 0.791 | 0.82  | 0.805   |
|                         | -                        | ✓           | 0.815    | 0.789      | 0.803 | 0.815 | 0.809   |
|                         | ✓                        | ✓           | 0.715    | 0.721      | 0.728 | 0.795 | 0.762   |

**Tablo 4.37.** devam ediyor.

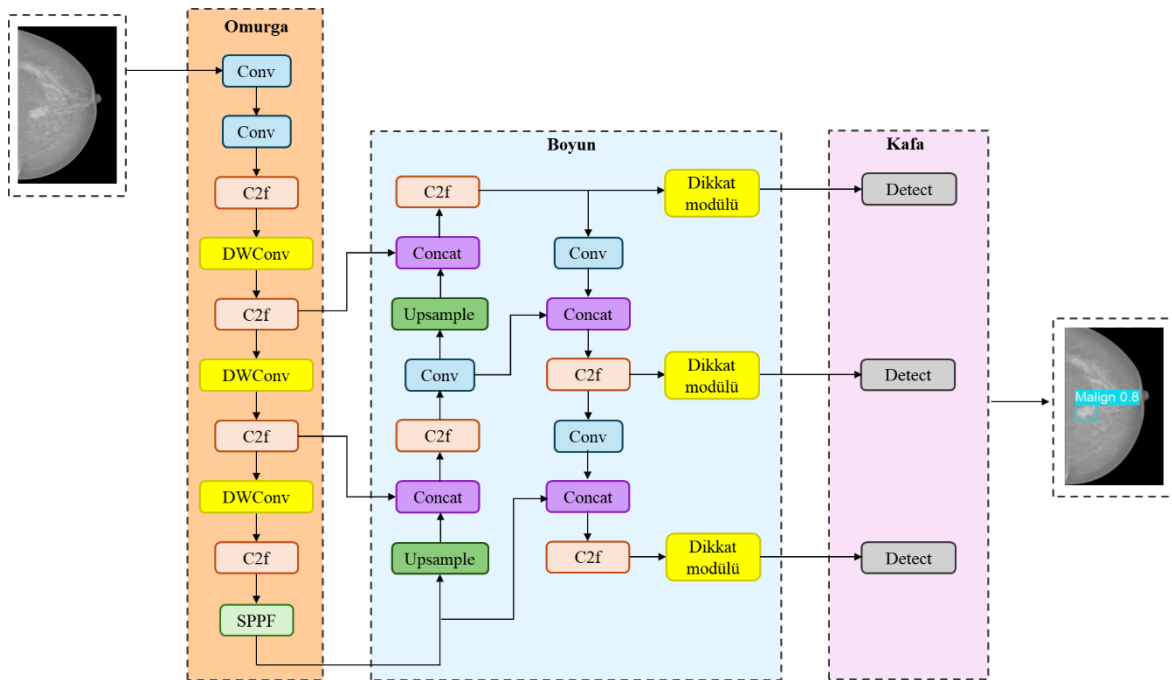
|                        |   |   |       |       |       |       |              |
|------------------------|---|---|-------|-------|-------|-------|--------------|
| YOLOv5s + DWConv + NAM | - | - | 0.778 | 0.701 | 0.793 | 0.776 | 0.785        |
|                        | ✓ | - | 0.797 | 0.719 | 0.762 | 0.79  | 0.775        |
|                        | - | ✓ | 0.769 | 0.721 | 0.756 | 0.778 | 0.767        |
|                        | ✓ | ✓ | 0.753 | 0.765 | 0.787 | 0.809 | 0.798        |
| YOLOv5s + DWConv + CA  | - | - | 0.771 | 0.732 | 0.771 | 0.797 | 0.784        |
|                        | ✓ | - | 0.805 | 0.664 | 0.786 | 0.758 | 0.772        |
|                        | - | ✓ | 0.788 | 0.72  | 0.738 | 0.815 | 0.777        |
|                        | ✓ | ✓ | 0.781 | 0.78  | 0.782 | 0.823 | 0.803        |
| YOLOv5s + DWConv + EMA | - | - | 0.824 | 0.738 | 0.845 | 0.818 | 0.831        |
|                        | ✓ | - | 0.786 | 0.741 | 0.804 | 0.798 | 0.801        |
|                        | - | ✓ | 0.863 | 0.743 | 0.843 | 0.827 | <b>0.835</b> |
|                        | ✓ | ✓ | 0.795 | 0.795 | 0.786 | 0.806 | 0.797        |

Tablo 4.36’da sunulan INBreast veri kümesi üzerinde gerçekleştirilen ince ayar analizlerinde, omurga katmanındaki DWConv uygulamasının dikkat modüllerinin etkisini genel olarak zayıflattığı görülmektedir. DWConv entegrasyonu sonrasında, YOLOv5s + dikkat modülü model varyantlarının performansında düşüş gözlemlenmiştir (Bkz. Tablo 4.24). Her ne kadar en yüksek performans YOLOv5s + DWConv + NAM modeli ile 0.824 mAP@0.5 olarak elde edilse de bu değer DWConv öncesi ulaşılan en yüksek performansın altında kalmıştır.

Benzer şekilde Tablo 4.37’de sunulan Başkent veri kümesi üzerinde gerçekleştirilen ince ayar analizlerinde ise, DWConv uygulaması bazı dikkat modülleri ile birlikte model performansını kısmen korusa da genel olarak önceki sonuçlara kıyasla daha düşük performans elde edilmiştir (Bkz. Tablo 4.25). Bu kapsamda, en yüksek performans YOLOv5s + DWConv + EMA modeli ile 0.835 mAP@0.5 olarak edilmiştir, ancak bu değer DWConv öncesinde aynı dikkat modülü ile elde edilen 0.848’lik değerden daha düşüktür.

#### 4.7.2. YOLOv8 + dikkat modülü modellerinde DWConv uygulaması

Şekil 4.18’de gösterildiği üzere, YOLOv8s + dikkat modülü model varyantlarının da performansını iyileştirmek amacıyla, modelin omurga katmanında yer alan son üç standart evrişim (Conv) bloğu, DWConv blokları ile değiştirilmiştir.



Şekil 4.18. YOLOv8 + DWConv + dikkat modülü model mimarisinin gösterimi

Bu analiz kapsamında yalnızca VinDr-Mammo + CBIS-DDSM birleşim veri kümesi kullanılarak önce sıfırdan eğitim gerçekleştirilmiş, ardından tüm model varyantları ve veri artırma parametre kombinasyonları için INBreast veri kümesi üzerinde ince ayar analizleri uygulanmıştır. Sıfırdan eğitim sonuçları Tablo 4.38’de, INBreast ve Başkent veri kümeleri için gerçekleştirilen ince ayar sonuçları ise sırasıyla Tablo 4.39 ve Tablo 4.40’ta sunulmuştur.

**Tablo 4.38.** CBIS-DDSM + VinDr-Mammo veri kümesinde YOLOv8s + DWConv + dikkat modülleri sıfırdan eğitim performans sonuçları

| Model                   | Kesinlik | Duyarlılık | BAP   | MAP   | mAP@0.5      |
|-------------------------|----------|------------|-------|-------|--------------|
| YOLOv8s + DWConv + SE   | 0.522    | 0.504      | 0.419 | 0.555 | 0.487        |
| YOLOv8s + DWConv + ECA  | 0.509    | 0.555      | 0.451 | 0.543 | 0.497        |
| YOLOv8s + DWConv + CBAM | 0.515    | 0.534      | 0.466 | 0.562 | <b>0.514</b> |
| YOLOv8s + DWConv + NAM  | 0.525    | 0.498      | 0.428 | 0.532 | 0.48         |
| YOLOv8s + DWConv + CA   | 0.508    | 0.486      | 0.443 | 0.515 | 0.479        |
| YOLOv8s + DWConv + EMA  | 0.534    | 0.502      | 0.46  | 0.522 | 0.491        |

Tablo 4.38’e göre DWConv entegrasyonu sonrasında yalnızca YOLOv8s + CBAM modelinde anlamlı bir performans artışı gözlemlenmiştir. Bu model, DWConv eklenmeden önce 0.507 mAP@0.5 değeri elde etmişken (Bkz. Tablo 4.30), DWConv sonrası bu değer 0.514’e yükselmiştir.

**Tablo 4.39.** CBIS-DDSM + VinDr-Mammo → INBreast: YOLOv8s + DWConv + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model                  | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|------------------------|--------------------------|-------------|----------|------------|-------|-------|--------------|
|                        | flipud (0.5)             | mixup (0.1) |          |            |       |       |              |
| YOLOv8s + DWConv + SE  | -                        | -           | 0.871    | 0.66       | 0.696 | 0.899 | 0.797        |
|                        | ✓                        | -           | 0.812    | 0.754      | 0.708 | 0.886 | 0.797        |
|                        | -                        | ✓           | 0.806    | 0.722      | 0.71  | 0.901 | 0.806        |
|                        | ✓                        | ✓           | 0.822    | 0.734      | 0.721 | 0.886 | 0.804        |
| YOLOv8s + DWConv + ECA | -                        | -           | 0.841    | 0.699      | 0.722 | 0.874 | 0.798        |
|                        | ✓                        | -           | 0.791    | 0.742      | 0.712 | 0.883 | 0.798        |
|                        | -                        | ✓           | 0.808    | 0.732      | 0.787 | 0.877 | 0.832        |
|                        | ✓                        | ✓           | 0.84     | 0.736      | 0.738 | 0.884 | 0.811        |
| YOLOv8s+ DWConv + CBAM | -                        | -           | 0.812    | 0.796      | 0.762 | 0.908 | 0.835        |
|                        | ✓                        | -           | 0.82     | 0.79       | 0.797 | 0.894 | 0.846        |
|                        | -                        | ✓           | 0.799    | 0.793      | 0.744 | 0.905 | 0.824        |
|                        | ✓                        | ✓           | 0.857    | 0.804      | 0.799 | 0.916 | <b>0.858</b> |

**Tablo 4.39.** devam ediyor.

|                        |   |   |       |       |       |       |       |
|------------------------|---|---|-------|-------|-------|-------|-------|
| YOLOv8s+ DWConv + NAM  | - | - | 0.824 | 0.693 | 0.702 | 0.877 | 0.702 |
|                        | ✓ | - | 0.799 | 0.785 | 0.694 | 0.902 | 0.798 |
|                        | - | ✓ | 0.81  | 0.734 | 0.733 | 0.89  | 0.812 |
|                        | ✓ | ✓ | 0.825 | 0.774 | 0.746 | 0.896 | 0.821 |
| YOLOv8s + DWConv + CA  | - | - | 0.78  | 0.759 | 0.693 | 0.888 | 0.79  |
|                        | ✓ | - | 0.862 | 0.777 | 0.759 | 0.912 | 0.835 |
|                        | - | ✓ | 0.84  | 0.729 | 0.741 | 0.864 | 0.803 |
|                        | ✓ | ✓ | 0.826 | 0.75  | 0.757 | 0.893 | 0.825 |
| YOLOv8s + DWConv + EMA | - | - | 0.764 | 0.747 | 0.696 | 0.868 | 0.781 |
|                        | ✓ | - | 0.802 | 0.766 | 0.731 | 0.878 | 0.805 |
|                        | - | ✓ | 0.859 | 0.706 | 0.704 | 0.881 | 0.793 |
|                        | ✓ | ✓ | 0.87  | 0.751 | 0.745 | 0.881 | 0.813 |

**Tablo 4.40.** CBIS-DDSM + VinDr-Mammo → Başkent: YOLOv8s + DWConv + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model                  | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5 |
|------------------------|--------------------------|-------------|----------|------------|-------|-------|---------|
|                        | flipud (0.5)             | mixup (0.1) |          |            |       |       |         |
| YOLOv8s + DWConv + SE  | -                        | -           | 0.77     | 0.744      | 0.78  | 0.80  | 0.79    |
|                        | ✓                        | -           | 0.723    | 0.702      | 0.739 | 0.808 | 0.773   |
|                        | -                        | ✓           | 0.785    | 0.671      | 0.76  | 0.777 | 0.769   |
|                        | ✓                        | ✓           | 0.782    | 0.646      | 0.767 | 0.774 | 0.77    |
| YOLOv8s + DWConv + ECA | -                        | -           | 0.813    | 0.718      | 0.757 | 0.819 | 0.788   |
|                        | ✓                        | -           | 0.885    | 0.672      | 0.802 | 0.808 | 0.805   |
|                        | -                        | ✓           | 0.734    | 0.775      | 0.765 | 0.851 | 0.808   |
|                        | ✓                        | ✓           | 0.802    | 0.703      | 0.823 | 0.775 | 0.799   |
| YOLOv8s+ DWConv + CBAM | -                        | -           | 0.725    | 0.717      | 0.789 | 0.752 | 0.771   |
|                        | ✓                        | -           | 0.818    | 0.681      | 0.796 | 0.776 | 0.786   |
|                        | -                        | ✓           | 0.799    | 0.692      | 0.793 | 0.80  | 0.796   |
|                        | ✓                        | ✓           | 0.766    | 0.774      | 0.766 | 0.83  | 0.798   |

**Tablo 4.40.** devam ediyor.

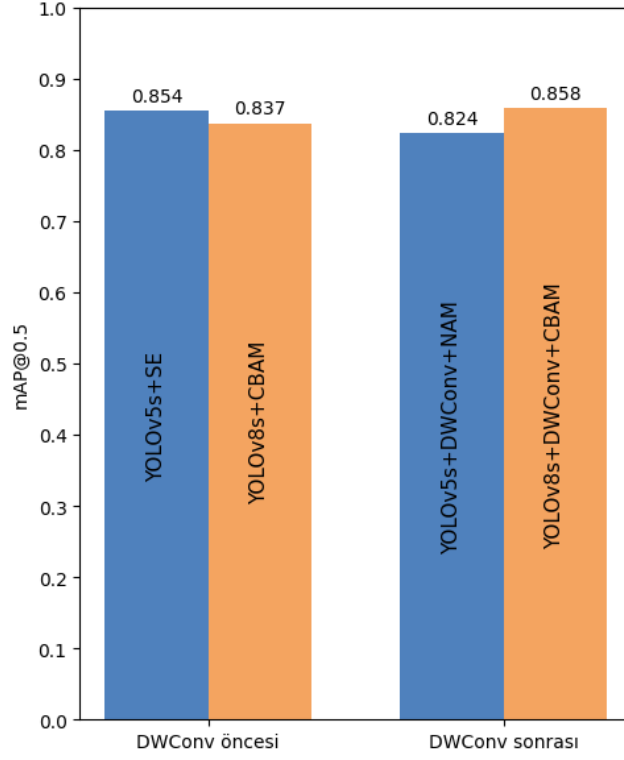
|                        |   |   |       |       |       |       |              |
|------------------------|---|---|-------|-------|-------|-------|--------------|
| YOLOv8s+ DWConv + NAM  | - | - | 0.753 | 0.723 | 0.76  | 0.821 | 0.79         |
|                        | ✓ | - | 0.766 | 0.742 | 0.801 | 0.802 | 0.801        |
|                        | - | ✓ | 0.831 | 0.695 | 0.779 | 0.841 | 0.81         |
|                        | ✓ | ✓ | 0.783 | 0.758 | 0.784 | 0.787 | 0.785        |
| YOLOv8s + DWConv + CA  | - | - | 0.762 | 0.772 | 0.826 | 0.801 | 0.814        |
|                        | ✓ | - | 0.792 | 0.747 | 0.796 | 0.82  | 0.808        |
|                        | - | ✓ | 0.85  | 0.701 | 0.773 | 0.811 | 0.792        |
|                        | ✓ | ✓ | 0.829 | 0.686 | 0.771 | 0.812 | 0.791        |
| YOLOv8s + DWConv + EMA | - | - | 0.846 | 0.709 | 0.794 | 0.783 | 0.788        |
|                        | ✓ | - | 0.826 | 0.755 | 0.811 | 0.842 | <b>0.827</b> |
|                        | - | ✓ | 0.798 | 0.73  | 0.772 | 0.808 | 0.79         |
|                        | ✓ | ✓ | 0.87  | 0.709 | 0.814 | 0.808 | 0.811        |

Tablo 4.39’da sunulan INBreast veri kümesi üzerinde gerçekleştirilen ince ayar analizlerinde, omurga katmanına DWConv uygulanması, YOLOv8s + dikkat modülü model varyantlarının çoğunun performansında küçük artışlara yol açmıştır (Bkz. Tablo 4.33). DWConv entegrasyonu sonrası en yüksek performans, 0.858 mAP@0.5 değeri ile YOLOv8s + DWConv + CBAM modelinde elde edilmiştir. Bu performans, DWConv uygulanmadan önce aynı dikkat modülü ile elde edilen 0.837’lik değerden daha yüksektir.

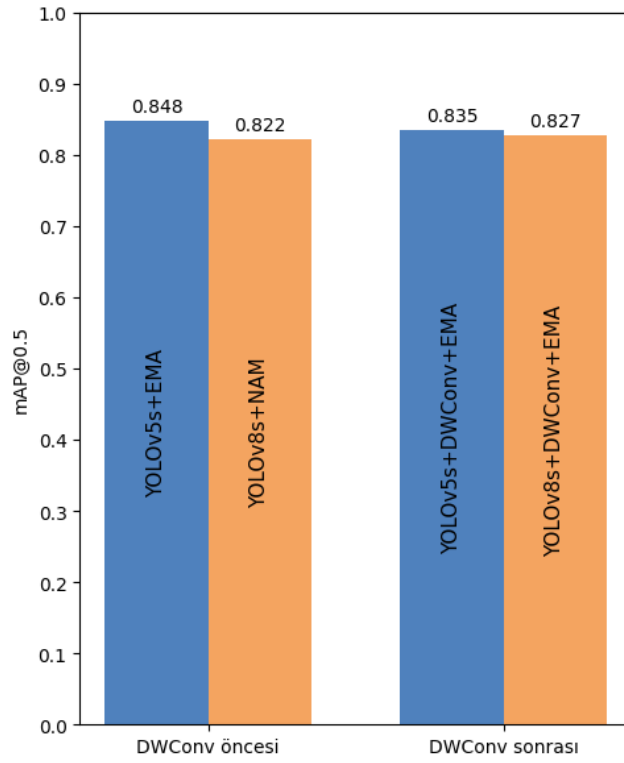
Benzer şekilde Tablo 4.40’ta sunulan Başkent veri kümesi üzerindeki ince ayar analizlerinde ise, DWConv entegrasyonunun dikkat modülleri ile birlikte model performansı üzerinde genel olarak sınırlı ama olumlu etkiler yarattığı gözlemlenmiştir (Bkz. Tablo 4.34). DWConv öncesinde en yüksek mAP@0.5 değeri YOLOv8s + NAM modeli ile 0.822 olarak elde edilirken, DWConv sonrası en yüksek performans YOLOv8s + DWConv + EMA modeli ile 0.827’lik mAP@0.5 değerine ulaşmıştır.

#### **4.8. Dikkat Modülleri Entegre Edilmiş YOLOv5 ve YOLOv8 Modellerinde DWConv Uygulamasının Performans Karşılaştırması**

Bu bölümde, YOLOv5s + dikkat modülü ve YOLOv8s + dikkat modülü modellerine DWConv entegrasyonu ile oluşturulan model varyantlarının, kitle tespiti ve sınıflandırma performansı üzerindeki etkileri değerlendirilmiştir. Bu kapsamda, her iki model yapısı için birleşim veri kümesi (CBIS-DDSM + VinDr-Mammo) kullanılarak gerçekleştirilen eğitim sürecinin ardından, DWConv entegrasyonu öncesi ve sonrası elde edilen ince ayar performansları karşılaştırmalı olarak analiz edilmiştir. Karşılaştırmalar mAP@0.5 performans ölçütü üzerinden gerçekleştirilmiştir. INBreast ve Başkent veri kümeleri üzerinde elde edilen karşılaştırmalı en iyi ince ayar sonuçları sırasıyla Şekil 4.19 ve Şekil 4.20’de görselleştirilmiştir.



Şekil 4.19. DWConv entegrasyonu öncesi ve sonrası en başarılı YOLO + dikkat modülü model varyantlarının INBreast üzerindeki performans karşılaştırması.



Şekil 4.20. DWConv entegrasyonu öncesi ve sonrası en başarılı YOLO + dikkat modülü model varyantlarının Başkent üzerindeki performans karşılaştırması.

Şekil 4.19’da INBreast veri kümesi üzerinde sunulan karşılaştırmalı sonuçlar incelendiğinde;

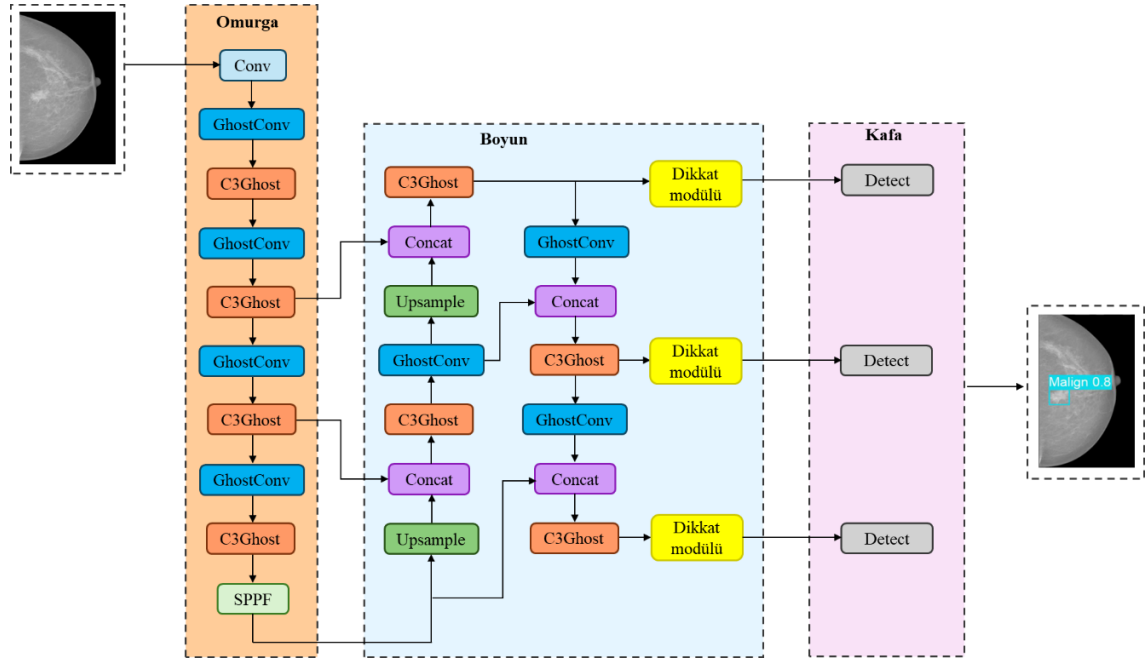
- Karşılaştırılan en başarılı model varyantları dikkate alındığında, DWConv uygulaması YOLOv5s + dikkat modülü modelinde performans düşüşüne, YOLOv8s + dikkat modülü modelinde ise performans iyileşmesine neden olmuştur.
- YOLOv8s için hem DWConv öncesi hem de DWConv sonrası en iyi performansın CBAM ile elde edilmesi, bu dikkat modülünün tutarlı ve güvenilir bir başarı sağladığını göstermektedir.
- Ayrıca, YOLOv8s + DWConv + CBAM modeli, şimdiye kadar INBreast veri kümesi üzerinde elde edilen en yüksek mAP@0.5 değerine (0.858) ulaşmıştır.

Şekil 4.20’de Başkent veri kümesi üzerinde sunulan karşılaştırmalı sonuçlar incelendiğinde;

- Karşılaştırılan en başarılı model varyantları dikkate alındığında, DWConv uygulaması YOLOv5s + dikkat modülü modelinde performans düşüşüne, YOLOv8s + dikkat modülü modelinde ise hafif bir performans artışına neden olmuştur.
- DWConv sonrası YOLOv8s + dikkat modülü modelinde elde edilen sınırlı başarı artışına rağmen, bu modelin performansı hem DWConv öncesi hem de DWConv sonrası YOLOv5s + dikkat modülü modellerine kıyasla daha düşük kalmıştır.
- YOLOv5s + DWConv + EMA modeli, DWConv entegrasyonu sonrası 0.835 mAP@0.5 değeri ile en başarılı sonuçlardan birini sunmuştur. Ancak bu skor, DWConv uygulanmadan önce aynı dikkat modülü ile elde edilen 0.848’lik değerinin altında kalmıştır.

#### 4.9. YOLOv8-Ghost ve Dikkat Modülleri Analiz Sonuçları

Önceki analizlerde DWConv entegrasyonunun sağladığı olumlu sonuçlar doğrultusunda, alternatif bir hafif evrişim yapısı olan GhostConv’un da model performansını iyileştirme potansiyeli taşıdığı düşünülmüştür. Bu nedenle, bu bölümde temel model olarak standart YOLOv8s yerine YOLOv8s-Ghost modeli tercih edilmiştir. Önceki bölümlerdeki yaklaşım temel alınarak dikkat modülleri, Şekil 4.21’de gösterildiği üzere boyun katmanından sınıflandırma kafasına geçişte daha anlamlı ve odaklanmış özneteliklerin taşınmasını sağlamak amacıyla entegre edilmiştir.



Şekil 4.21. YOLOv8-Ghost + dikkat modülü model mimarisinin gösterimi.

Bir önceki bölümde sunulan analiz süreci, YOLOv8s-Ghost ve YOLOv8s-Ghost + dikkat modülleri model varyantları için aynı deneysel koşullar altında yeniden uygulanmıştır. Bu doğrultuda, VinDr-Mammo + CBIS-DDSM birleşim veri kümesi kullanılarak önce sıfırdan eğitim gerçekleştirilmiş, ardından tüm model varyantları ve veri artırma parametre kombinasyonları için INBreast ve Başkent veri kümelerinde ince ayar analizleri gerçekleştirilmiştir. Sıfırdan eğitim sonuçları Tablo 4.41’de, ince ayar sonuçları ise sırasıyla Tablo 4.42 ve Tablo 4.43’te sunulmuştur.

**Tablo 4.41.** CBIS-DDSM + VinDr-Mammo veri kümesinde YOLOv8s-ghost ve YOLOv8s-ghost + dikkat modüllerinin sıfırdan eğitim performans sonuçları

| Model                | Kesinlik | Duyarlılık | BAP   | MAP   | mAP@0.5      |
|----------------------|----------|------------|-------|-------|--------------|
| YOLOv8s-ghost        | 0.513    | 0.525      | 0.415 | 0.537 | 0.476        |
| YOLOv8s-ghost + SE   | 0.529    | 0.527      | 0.439 | 0.52  | 0.48         |
| YOLOv8s-ghost + ECA  | 0.501    | 0.502      | 0.414 | 0.54  | 0.477        |
| YOLOv8s-ghost + CBAM | 0.475    | 0.548      | 0.412 | 0.558 | <b>0.485</b> |
| YOLOv8s-ghost + NAM  | 0.473    | 0.552      | 0.446 | 0.512 | 0.479        |
| YOLOv8s-ghost + CA   | 0.444    | 0.475      | 0.363 | 0.469 | 0.416        |
| YOLOv8s-ghost + EMA  | 0.454    | 0.502      | 0.358 | 0.511 | 0.435        |

Tablo 4.41’de sunulan sıfırdan eğitim analiz sonuçlarına göre, YOLOv8s-ghost modeline entegre edilen dikkat modülleri (CA ve EMA hariç) genel olarak tespit ve sınıflandırma performansını artırmıştır. En yüksek başarı, mAP@0.5 değeri 0.485 ile YOLOv8s-ghost + CBAM modeline aittir. Ancak dikkat modüllerinin YOLOv8s modeline eklendiği önceki analizlerle (Bkz. Tablo 4.30) karşılaştırıldığında, NAM dışındaki tüm modüller için mAP@0.5 değerlerinde bir miktar performans düşüşü gözlemlenmiştir.

**Tablo 4.42.** CBIS-DDSM + VinDr-Mammo → INBreast: YOLOv8s-ghost ve YOLOv8s-ghost + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model               | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|---------------------|--------------------------|-------------|----------|------------|-------|-------|--------------|
|                     | flipud (0.5)             | mixup (0.1) |          |            |       |       |              |
| YOLOv8s-ghost       | -                        | -           | 0.796    | 0.726      | 0.694 | 0.883 | 0.789        |
|                     | ✓                        | -           | 0.835    | 0.70       | 0.746 | 0.91  | 0.828        |
|                     | -                        | ✓           | 0.833    | 0.736      | 0.696 | 0.92  | 0.806        |
|                     | ✓                        | ✓           | 0.893    | 0.773      | 0.795 | 0.927 | <b>0.861</b> |
| YOLOv8s-ghost + SE  | -                        | -           | 0.781    | 0.753      | 0.696 | 0.901 | 0.799        |
|                     | ✓                        | -           | 0.792    | 0.719      | 0.704 | 0.883 | 0.793        |
|                     | -                        | ✓           | 0.728    | 0.667      | 0.615 | 0.858 | 0.736        |
|                     | ✓                        | ✓           | 0.808    | 0.734      | 0.706 | 0.893 | 0.799        |
| YOLOv8s-ghost + ECA | -                        | -           | 0.87     | 0.675      | 0.698 | 0.891 | 0.794        |
|                     | ✓                        | -           | 0.882    | 0.759      | 0.737 | 0.927 | 0.832        |
|                     | -                        | ✓           | 0.795    | 0.735      | 0.712 | 0.886 | 0.779        |
|                     | ✓                        | ✓           | 0.781    | 0.764      | 0.726 | 0.888 | 0.807        |

**Tablo 4.42.** devam ediyor.

|                      |   |   |       |       |       |       |       |
|----------------------|---|---|-------|-------|-------|-------|-------|
| YOLOv8s-ghost + CBAM | - | - | 0.832 | 0.697 | 0.706 | 0.883 | 0.795 |
|                      | ✓ | - | 0.80  | 0.758 | 0.703 | 0.91  | 0.806 |
|                      | - | ✓ | 0.892 | 0.715 | 0.724 | 0.899 | 0.812 |
|                      | ✓ | ✓ | 0.882 | 0.718 | 0.714 | 0.907 | 0.81  |
| YOLOv8s-ghost + NAM  | - | - | 0.825 | 0.705 | 0.715 | 0.881 | 0.798 |
|                      | ✓ | - | 0.837 | 0.734 | 0.734 | 0.882 | 0.808 |
|                      | - | ✓ | 0.806 | 0.792 | 0.715 | 0.906 | 0.811 |
|                      | ✓ | ✓ | 0.831 | 0.772 | 0.771 | 0.876 | 0.824 |
| YOLOv8s-ghost + CA   | - | - | 0.803 | 0.749 | 0.691 | 0.902 | 0.797 |
|                      | ✓ | - | 0.824 | 0.782 | 0.79  | 0.901 | 0.846 |
|                      | - | ✓ | 0.884 | 0.745 | 0.722 | 0.902 | 0.812 |
|                      | ✓ | ✓ | 0.826 | 0.793 | 0.784 | 0.914 | 0.849 |
| YOLOv8s-ghost + EMA  | - | - | 0.796 | 0.693 | 0.677 | 0.867 | 0.772 |
|                      | ✓ | - | 0.821 | 0.696 | 0.691 | 0.891 | 0.791 |
|                      | - | ✓ | 0.718 | 0.741 | 0.672 | 0.88  | 0.776 |
|                      | ✓ | ✓ | 0.739 | 0.74  | 0.705 | 0.881 | 0.793 |

**Tablo 4.43.** CBIS-DDSM + VinDr-Mammo → Başkent: YOLOv8s-ghost ve YOLOv8s-ghost + dikkat modülleri için farklı veri artırma parametrelerinin ince ayar performansına etkisi

| Model               | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|---------------------|--------------------------|-------------|----------|------------|-------|-------|--------------|
|                     | flipud (0.5)             | mixup (0.1) |          |            |       |       |              |
| YOLOv8s-ghost       | -                        | -           | 0.808    | 0.689      | 0.752 | 0.795 | 0.773        |
|                     | ✓                        | -           | 0.814    | 0.73       | 0.753 | 0.84  | 0.795        |
|                     | -                        | ✓           | 0.816    | 0.706      | 0.75  | 0.79  | 0.772        |
|                     | ✓                        | ✓           | 0.764    | 0.722      | 0.747 | 0.791 | 0.769        |
| YOLOv8s-ghost + SE  | -                        | -           | 0.781    | 0.746      | 0.793 | 0.793 | 0.793        |
|                     | ✓                        | -           | 0.83     | 0.72       | 0.79  | 0.803 | 0.796        |
|                     | -                        | ✓           | 0.745    | 0.806      | 0.816 | 0.807 | 0.811        |
|                     | ✓                        | ✓           | 0.782    | 0.762      | 0.781 | 0.837 | 0.809        |
| YOLOv8s-ghost + ECA | -                        | -           | 0.83     | 0.731      | 0.808 | 0.801 | 0.805        |
|                     | ✓                        | -           | 0.80     | 0.745      | 0.805 | 0.829 | <b>0.817</b> |
|                     | -                        | ✓           | 0.77     | 0.733      | 0.789 | 0.838 | 0.813        |
|                     | ✓                        | ✓           | 0.786    | 0.754      | 0.794 | 0.808 | 0.801        |

**Tablo 4.43.** devam ediyor.

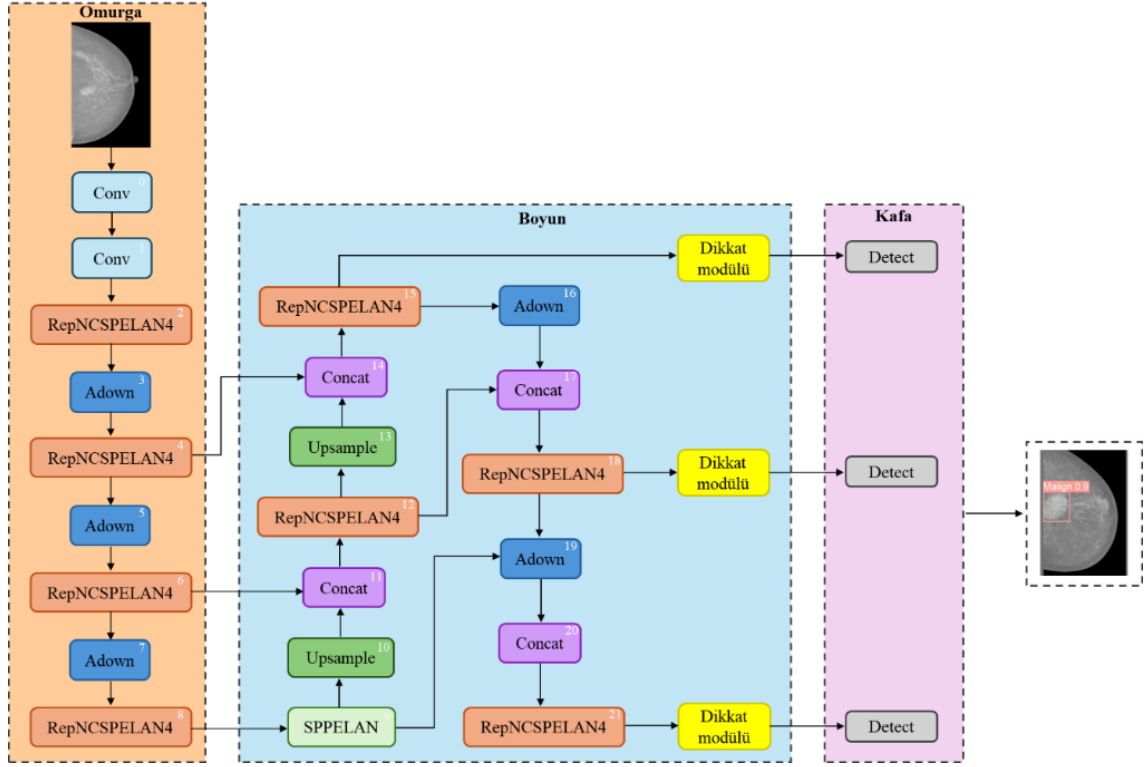
|                      |   |   |       |       |       |       |       |
|----------------------|---|---|-------|-------|-------|-------|-------|
| YOLOv8s-ghost + CBAM | - | - | 0.856 | 0.68  | 0.766 | 0.834 | 0.80  |
|                      | ✓ | - | 0.79  | 0.704 | 0.747 | 0.80  | 0.776 |
|                      | - | ✓ | 0.779 | 0.724 | 0.735 | 0.839 | 0.787 |
|                      | ✓ | ✓ | 0.681 | 0.772 | 0.733 | 0.795 | 0.76  |
| YOLOv8s-ghost + NAM  | - | - | 0.77  | 0.692 | 0.738 | 0.799 | 0.769 |
|                      | ✓ | - | 0.743 | 0.703 | 0.778 | 0.767 | 0.773 |
|                      | - | ✓ | 0.777 | 0.723 | 0.749 | 0.817 | 0.783 |
|                      | ✓ | ✓ | 0.804 | 0.736 | 0.755 | 0.821 | 0.788 |
| YOLOv8s-ghost + CA   | - | - | 0.797 | 0.68  | 0.778 | 0.82  | 0.799 |
|                      | ✓ | - | 0.727 | 0.747 | 0.78  | 0.837 | 0.811 |
|                      | - | ✓ | 0.742 | 0.716 | 0.751 | 0.868 | 0.81  |
|                      | ✓ | ✓ | 0.789 | 0.752 | 0.748 | 0.835 | 0.791 |
| YOLOv8s-ghost + EMA  | - | - | 0.841 | 0.666 | 0.787 | 0.826 | 0.806 |
|                      | ✓ | - | 0.859 | 0.682 | 0.738 | 0.836 | 0.787 |
|                      | - | ✓ | 0.767 | 0.75  | 0.78  | 0.84  | 0.812 |
|                      | ✓ | ✓ | 0.753 | 0.75  | 0.77  | 0.824 | 0.797 |

Tablo 4.42’de sunulan INBreast veri kümesi üzerinde gerçekleştirilen ince ayar analizleri incelendiğinde, YOLOv8s-Ghost modelinin dikkat modülleri entegre edilmeden de yüksek performans sergilediği görülmektedir. Özellikle flipud ve mixup parametrelerinin birlikte uygulandığı durumda, temel YOLOv8s-Ghost modeli 0.861 mAP@0.5 değeri ile çalışmada şimdiye kadar elde edilen en yüksek başarıya ulaşmıştır.

Benzer şekilde Tablo 4.43’te sunulan Başkent veri kümesi üzerinde gerçekleştirilen ince ayar sonuçları değerlendirildiğinde ise, YOLOv8s-Ghost modeli ile dikkat modüllerinin entegrasyonu sonucunda bazı varyantların başarılı performanslar sergilediği görülmektedir. Bu kapsamda, en yüksek performans 0.817 mAP@0.5 değeri ile YOLOv8s-ghost + ECA modeli ve varsayılan + flipud veri artırma kombinasyonunda elde edilmiştir. Ancak bu performans, daha önce aynı veri kümesi üzerinde elde edilen en yüksek sonuçların altında kalmıştır.

#### **4.10. GELAN ve Dikkat Modülleri Analiz Sonuçları**

Bu bölümde, YOLOv9 ailesine ait GELAN varyantları (GELAN-s, GELAN-m ve GELAN-c) ile farklı dikkat modüllerinin (SE, ECA, CA, CBAM, NAM ve EMA) entegrasyonu sonucu oluşturulan model varyantlarının kitle tespiti ve sınıflandırması üzerindeki performansı değerlendirilmiştir. Önceki analizlerde dikkat modüllerinin sağladığı performans artışları dikkate alınarak, Şekil 4.22’de gösterildiği gibi GELAN mimarisinin boyun katmanına entegre edilmiştir. Bu sayede, boyun katmanından sınıflandırma kafasına aktarılacak özniteliklerin daha anlamlı ve odaklanmış hale getirilmesi amaçlanmıştır.



Şekil 4.22. GELAN + dikkat modülü model mimarisinin gösterimi.

GELAN modellerinin eğitimi sırasında kullanılan veri artırma parametreleri, yüksek seviye bir varsayılan konfigürasyon dosyasında tanımlanmıştır. Bu dosyada yer alan parametreler ve bunlara ait varsayılan değerleri Tablo 4.44’te sunulmuştur. Bu yüksek seviye konfigürasyon dosyasında, önceki YOLO versiyonlarından farklı olarak mixup ve copy-paste gibi gelişmiş veri artırma tekniklerine de yer verilmiştir. Ayrıca, GELAN için sıfırdan eğitim ve ince ayar süreçlerinde kullanılan hiperparametreler ise Tablo 4.45’te verilmiştir.

Tablo 4.44. GELAN eğitiminde kullanılan varsayılan veri artırma konfigürasyonu

| Parametre ve tanımı                                 | değeri |
|---|--------|
| <b>hsv_h:</b> görüntü HSV-ton artırma (kesir)       | 0.015  |
| <b>hsv_s:</b> görüntü HSV-doygunluk artırma (kesir) | 0.7    |
| <b>hsv_v:</b> görüntü HSV-değer artırma (kesir)     | 0.4    |
| <b>degrees:</b> görüntü döndürme (+/- derece)       | 0.0    |
| <b>translate:</b> görüntü kaydırma (+/-kesir)       | 0.1    |
| <b>scale:</b> görüntü ölçekleme (+/- kazanç)        | 0.9    |
| <b>shear:</b> görüntü kesme (+/- derece)            | 0.0    |
| <b>perspective:</b> görüntü perspektifi (+/-kesir)  | 0.0    |

**Tablo 4.44.** devam ediyor.

|  |      |
|--|------|
| <b>flipud:</b> görüntü yukarı-aşağı çevirme (olasılık)           | 0.0  |
| <b>fliplr:</b> görüntü sola-sağa çevirme (olasılık)              | 0.5  |
| <b>mosaic:</b> görüntü mozaığı (olasılık)                        | 1.0  |
| <b>mixup:</b> görüntü karıştırma (olasılık)                      | 0.15 |
| <b>copy_paste:</b> segmentasyon için kopyala-yapıştır (olasılık) | 0.3  |

**Tablo 4.45.** GELAN için model eğitim hiperparametreleri

| <b>hiperparametre</b>              | <b>değeri</b> |
|------------------------------------|---------------|
| görüntü boyutu                     | 544 x 544     |
| epok (sıfırdan eğitim / ince ayar) | 300 / 200     |
| toplu işlem boyutu                 | 8             |
| başlangıç öğrenme oranı            | 0.01          |
| optimizasyon algoritması           | SGD           |
| momentum katsayısı                 | 0.937         |
| ağırlık azaltma katsayısı          | 0.0005        |
| erken durdurma değeri (epok)       | 100           |

Daha önceki analizlerde, CBIS-DDSM ve VinDr-Mammo birleşim veri kümesi ile eğitilen modeller en yüksek performansı gösterdiği için bu bölümde analizler yalnızca birleşim veri kümesi üzerinden yürütülmüştür. Bu kapsamda, CBIS-DDSM + VinDr-Mammo veri kümesi, rastgele olacak şekilde %80 eğitim, %10 doğrulama ve %10 test alt kümelerine ayrılmıştır. GELAN-s, GELAN-m ve GELAN-c modellerine altı farklı dikkat modülü (SE, ECA, CBAM, NAM, CA ve EMA) entegre edilerek yeni model varyantları oluşturulmuştur. Tüm model varyantları, varsayılan veri artırma konfigürasyonu kullanılarak ayrı ayrı sıfırdan eğitilmiştir. Elde edilen sıfırdan eğitim performans sonuçları sırasıyla Tablo 4.46, Tablo 4.47 ve Tablo 4.48’de sunulmuştur.

**Tablo 4.46.** CBIS-DDSM + VinDr-Mammo veri kümesinde GELAN-s ve GELAN-s + dikkat modüllerinin sıfırdan eğitim performans sonuçları

| <b>Model</b>   | <b>Kesinlik</b> | <b>Duyarlılık</b> | <b>BAP</b> | <b>MAP</b> | <b>mAP@0.5</b> |
|----------------|-----------------|-------------------|------------|------------|----------------|
| GELAN-s        | 0.536           | 0.524             | 0.445      | 0.528      | 0.486          |
| GELAN-s + SE   | 0.575           | 0.536             | 0.46       | 0.567      | 0.513          |
| GELAN-s + ECA  | 0.599           | 0.506             | 0.457      | 0.60       | 0.529          |
| GELAN-s + CBAM | 0.558           | 0.521             | 0.51       | 0.56       | 0.535          |
| GELAN-s + NAM  | 0.547           | 0.529             | 0.509      | 0.521      | 0.515          |
| GELAN-s + CA   | 0.59            | 0.509             | 0.53       | 0.607      | <b>0.569</b>   |
| GELAN-s + EMA  | 0.591           | 0.452             | 0.478      | 0.538      | 0.508          |

**Tablo 4.47.** CBIS-DDSM + VinDr-Mammo veri kümesinde GELAN-m ve GELAN-m + dikkat modüllerinin sıfırdan eğitim performans sonuçları

| <b>Model</b>   | <b>Kesinlik</b> | <b>Duyarlılık</b> | <b>BAP</b> | <b>MAP</b> | <b>mAP@0.5</b> |
|----------------|-----------------|-------------------|------------|------------|----------------|
| GELAN-m        | 0.548           | 0.533             | 0.492      | 0.531      | 0.512          |
| GELAN-m + SE   | 0.557           | 0.525             | 0.483      | 0.541      | 0.512          |
| GELAN-m + ECA  | 0.609           | 0.521             | 0.50       | 0.571      | <b>0.535</b>   |
| GELAN-m + CBAM | 0.48            | 0.575             | 0.466      | 0.537      | 0.502          |
| GELAN-m + NAM  | 0.626           | 0.459             | 0.498      | 0.565      | 0.532          |
| GELAN-m + CA   | 0.559           | 0.51              | 0.506      | 0.506      | 0.506          |
| GELAN-m + EMA  | 0.57            | 0.521             | 0.485      | 0.551      | 0.518          |

**Tablo 4.48.** CBIS-DDSM + VinDr-Mammo veri kümesinde GELAN-c ve GELAN-c + dikkat modüllerinin sıfırdan eğitim performans sonuçları

| <b>Model</b>   | <b>Kesinlik</b> | <b>Duyarlılık</b> | <b>BAP</b> | <b>MAP</b> | <b>mAP@0.5</b> |
|----------------|-----------------|-------------------|------------|------------|----------------|
| GELAN-c        | 0.589           | 0.525             | 0.507      | 0.566      | 0.536          |
| GELAN-c + SE   | 0.551           | 0.537             | 0.483      | 0.558      | 0.52           |
| GELAN-c + ECA  | 0.53            | 0.525             | 0.466      | 0.561      | 0.514          |
| GELAN-c + CBAM | 0.55            | 0.587             | 0.522      | 0.568      | 0.545          |
| GELAN-c + NAM  | 0.598           | 0.521             | 0.522      | 0.577      | <b>0.55</b>    |
| GELAN-c + CA   | 0.542           | 0.482             | 0.483      | 0.56       | 0.521          |
| GELAN-c + EMA  | 0.538           | 0.568             | 0.47       | 0.619      | 0.545          |

Tablo 4.46, Tablo 4.47 ve Tablo 4.48’de sunulan sıfırdan eğitim sonuçlarına göre;

- GELAN modeline entegre edilen dikkat modüllerinin, tespit ve sınıflandırma başarısını çoğu durumda artırdığı gözlemlenmiştir.
- GELAN-s modelleri arasında en iyi performans ( $mAP@0.5=0.569$ ), GELAN-s + NAM modeline aittir.
- GELAN-m modelleri içinde en başarılı performans ( $mAP@0.5=0.535$ ), GELAN-m + ECA modeli ile elde edilmiştir.
- GELAN-c modelleri arasında ise en yüksek performansa ( $mAP@0.5=0.550$ ), GELAN-c + NAM modeli ile ulaşılmıştır.

Ardından, eğitilen modellerin en iyi ağırlıkları kullanılarak INBreast veri kümesi üzerinde amacıyla 5 katlı çapraz doğrulama stratejisi ile ince ayar gerçekleştirilmiştir. Önceki analizlerde, özellikle mixup ve flipud gibi veri artırma parametrelerinin model performansı üzerindeki etkileri detaylı şekilde incelenmiştir. GELAN eğitiminde, mixup parametresi varsayılan konfigürasyon içerisinde yer aldığı için yeniden değerlendirilmemiştir. Bu nedenle, yalnızca görüntü yukarı-aşağı çevirme (flipud = 0.5) parametresi analizlere dahil edilmiştir. Bununla birlikte, degrees, shear ve perspective gibi parametrelerin de model performansı üzerinde anlamlı bir etkisi olmadığı gözlemlendiğinden, bu parametreler analiz dışında tutulmuştur. GELAN-s/m/c modelleri için gerçekleştirilen ince ayar analizleri ve elde edilen sonuçlar sırasıyla Tablo 4.49, Tablo 4.50 ve Tablo 4.51’de sunulmaktadır.

**Tablo 4.49.** CBIS-DDSM + VinDr-Mammo → INBreast: GELAN-s ve GELAN-s + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi

| Model          | Veri artırma parametresi flipud (0.5) | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|----------------|---------------------------------------|----------|------------|-------|-------|--------------|
| GELAN-s        | -                                     | 0.80     | 0.756      | 0.73  | 0.865 | 0.798        |
|                | ✓                                     | 0.719    | 0.707      | 0.752 | 0.822 | 0.787        |
| GELAN-s + SE   | -                                     | 0.842    | 0.738      | 0.764 | 0.865 | 0.814        |
|                | ✓                                     | 0.862    | 0.758      | 0.76  | 0.858 | 0.809        |
| GELAN-s + ECA  | -                                     | 0.767    | 0.759      | 0.68  | 0.879 | 0.78         |
|                | ✓                                     | 0.825    | 0.749      | 0.739 | 0.883 | 0.811        |
| GELAN-s + CBAM | -                                     | 0.74     | 0.756      | 0.696 | 0.864 | 0.78         |
|                | ✓                                     | 0.778    | 0.74       | 0.694 | 0.876 | 0.785        |
| GELAN-s + NAM  | -                                     | 0.842    | 0.721      | 0.74  | 0.88  | 0.81         |
|                | ✓                                     | 0.831    | 0.725      | 0.739 | 0.869 | 0.804        |
| GELAN-s + CA   | -                                     | 0.823    | 0.747      | 0.729 | 0.887 | 0.808        |
|                | ✓                                     | 0.739    | 0.779      | 0.745 | 0.867 | 0.806        |
| GELAN-s + EMA  | -                                     | 0.854    | 0.765      | 0.776 | 0.891 | <b>0.833</b> |
|                | ✓                                     | 0.839    | 0.684      | 0.714 | 0.89  | 0.80         |

**Tablo 4.50.** CBIS-DDSM + VinDr-Mammo → INBreast: GELAN-m ve GELAN-m + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi

| Model          | Veri artırma parametresi flipud (0.5) | Kesinlik | Duyarluluk | B AP  | M AP  | mAP@0.5      |
|----------------|---------------------------------------|----------|------------|-------|-------|--------------|
| GELAN-m        | -                                     | 0.832    | 0.748      | 0.733 | 0.897 | 0.815        |
|                | ✓                                     | 0.782    | 0.714      | 0.674 | 0.873 | 0.773        |
| GELAN-m + SE   | -                                     | 0.788    | 0.787      | 0.739 | 0.895 | <b>0.817</b> |
|                | ✓                                     | 0.783    | 0.759      | 0.764 | 0.862 | 0.813        |
| GELAN-m + ECA  | -                                     | 0.75     | 0.776      | 0.736 | 0.882 | 0.809        |
|                | ✓                                     | 0.769    | 0.777      | 0.707 | 0.878 | 0.792        |
| GELAN-m + CBAM | -                                     | 0.764    | 0.741      | 0.721 | 0.871 | 0.796        |
|                | ✓                                     | 0.80     | 0.72       | 0.708 | 0.864 | 0.786        |
| GELAN-m + NAM  | -                                     | 0.889    | 0.677      | 0.724 | 0.892 | 0.809        |
|                | ✓                                     | 0.722    | 0.725      | 0.696 | 0.875 | 0.786        |
| GELAN-m + CA   | -                                     | 0.769    | 0.79       | 0.703 | 0.889 | 0.796        |
|                | ✓                                     | 0.72     | 0.75       | 0.736 | 0.855 | 0.796        |
| GELAN-m + EMA  | -                                     | 0.761    | 0.752      | 0.725 | 0.857 | 0.791        |
|                | ✓                                     | 0.816    | 0.704      | 0.726 | 0.855 | 0.79         |

**Tablo 4.51.** CBIS-DDSM + VinDr-Mammo → INBreast: GELAN-c ve GELAN-c + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi

| Model          | Veri artırma parametresi flipud (0.5) | Kesinlik | Duyarluluk | B AP  | M AP  | mAP@0.5      |
|----------------|---------------------------------------|----------|------------|-------|-------|--------------|
| GELAN-c        | -                                     | 0.899    | 0.738      | 0.762 | 0.894 | 0.827        |
|                | ✓                                     | 0.831    | 0.77       | 0.769 | 0.896 | 0.832        |
| GELAN-c + SE   | -                                     | 0.855    | 0.746      | 0.785 | 0.891 | <b>0.838</b> |
|                | ✓                                     | 0.817    | 0.782      | 0.758 | 0.892 | 0.825        |
| GELAN-c + ECA  | -                                     | 0.782    | 0.755      | 0.693 | 0.894 | 0.792        |
|                | ✓                                     | 0.786    | 0.783      | 0.779 | 0.893 | 0.836        |
| GELAN-c + CBAM | -                                     | 0.863    | 0.705      | 0.707 | 0.89  | 0.798        |
|                | ✓                                     | 0.846    | 0.702      | 0.728 | 0.862 | 0.795        |
| GELAN-c + NAM  | -                                     | 0.813    | 0.776      | 0.777 | 0.882 | 0.83         |
|                | ✓                                     | 0.862    | 0.764      | 0.795 | 0.88  | <b>0.838</b> |
| GELAN-c + CA   | -                                     | 0.754    | 0.717      | 0.697 | 0.891 | 0.794        |
|                | ✓                                     | 0.804    | 0.764      | 0.767 | 0.894 | 0.83         |
| GELAN-c + EMA  | -                                     | 0.809    | 0.742      | 0.73  | 0.887 | 0.809        |
|                | ✓                                     | 0.836    | 0.707      | 0.74  | 0.854 | 0.797        |

Tablo 4.49, Tablo 4.50 ve Tablo 4.51 incelendiğinde;

- GELAN-s ve dikkat modülü içeren modellerde en yüksek  $mAP@0.5$  değeri 0.833 ile GELAN-s + EMA modeli tarafından elde edilmiş olup, bu başarı yalnızca varsayılan veri artırma parametreleri kullanıldığında gözlemlenmiştir.
- GELAN-m ve dikkat modülü modelleri için en iyi performans, 0.817'lik  $mAP@0.5$  değeri ile GELAN-m + SE modeline ait olup, bu sonuç yine varsayılan veri artırma konfigürasyonunda elde edilmiştir.
- GELAN-c ve dikkat modülü modellerinde ise en yüksek  $mAP@0.5$  değeri olan 0.838, hem GELAN-c + SE (varsayılan veri artırma) hem de GELAN-c + NAM (varsayılan + flipud) modelleri tarafından sağlanmıştır.
- Bu sonuçlar, flipud veri artırma parametresinin bazı modellerde küçük performans artışları sağlarken, bazı modellerde ise performans düşüşüne neden olduğunu göstermiştir. Genel olarak, en iyi sonuçlara çoğunlukla varsayılan veri artırma konfigürasyonu ile ulaşılmıştır.

GELAN + dikkat modülü model varyantları ile yapılan analizler sonucunda, en yüksek ince ayar performansı GELAN-c model varyantlarında elde edilmiştir. Bu doğrultuda, Başkent veri kümesi üzerinde yalnızca GELAN-c ve GELAN-c + dikkat modülü modelleri ile ince ayar gerçekleştirilmiştir. Elde edilen ince ayar sonuçları Tablo 4.52'de sunulmuştur.

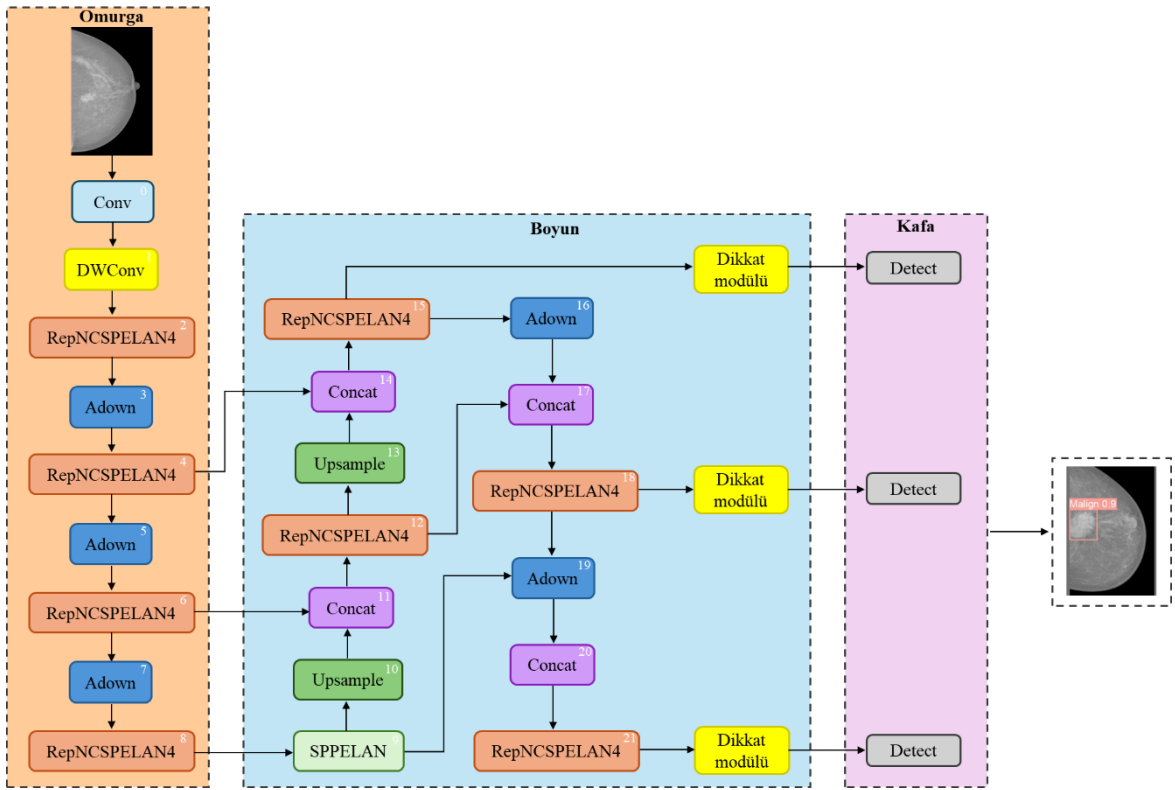
**Tablo 4.52.** CBIS-DDSM + VinDr-Mammo → Başkent: GELAN-c ve GELAN-c + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi

| Model          | Veri artırma parametresi flipud (0.5) | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|----------------|---------------------------------------|----------|------------|-------|-------|--------------|
| GELAN-c        | -                                     | 0.763    | 0.756      | 0.818 | 0.819 | 0.819        |
|                | ✓                                     | 0.825    | 0.72       | 0.823 | 0.817 | 0.82         |
| GELAN-c + SE   | -                                     | 0.753    | 0.753      | 0.793 | 0.823 | 0.808        |
|                | ✓                                     | 0.808    | 0.748      | 0.776 | 0.81  | 0.793        |
| GELAN-c + ECA  | -                                     | 0.764    | 0.714      | 0.765 | 0.809 | 0.787        |
|                | ✓                                     | 0.755    | 0.754      | 0.779 | 0.834 | 0.807        |
| GELAN-c + CBAM | -                                     | 0.793    | 0.722      | 0.778 | 0.79  | 0.785        |
|                | ✓                                     | 0.768    | 0.693      | 0.753 | 0.803 | 0.797        |
| GELAN-c + NAM  | -                                     | 0.853    | 0.739      | 0.831 | 0.819 | 0.82         |
|                | ✓                                     | 0.771    | 0.72       | 0.812 | 0.78  | 0.797        |
| GELAN-c + CA   | -                                     | 0.801    | 0.766      | 0.842 | 0.814 | <b>0.828</b> |
|                | ✓                                     | 0.756    | 0.73       | 0.767 | 0.796 | 0.78         |
| GELAN-c + EMA  | -                                     | 0.805    | 0.745      | 0.84  | 0.811 | 0.826        |
|                | ✓                                     | 0.822    | 0.694      | 0.78  | 0.831 | 0.805        |

Tablo 4.52’de sunulan Başkent veri kümesi üzerindeki ince ayar sonuçları değerlendirildiğinde, en yüksek mAP@0.5 değeri 0.828 ile GELAN-c + CA modeli ve varsayılan veri artırma parametreleri kullanıldığında elde edilmiştir.

#### 4.11. Dikkat Modülü Entegre Edilmiş GELAN Modellerinde DWConv Uygulaması

Bu bölümde, GELAN + dikkat modülü model varyantlarının performansını iyileştirmek amacıyla, Şekil 4.23’te gösterildiği üzere, omurga katmanındaki 1 numaralı standart evrişim (Conv) bloğu, DWConv bloğu ile değiştirilmiştir. Bu yapısal değişikliğin, modelin genel performansı üzerindeki etkisi detaylı olarak analiz edilmiştir.



Şekil 4.23. GELAN + DWConv + dikkat modülü model mimarisinin gösterimi.

Bu analiz kapsamında, VinDr-Mammo + CBIS-DDSM birleşim veri kümesi kullanılarak önce sıfırdan eğitim gerçekleştirilmiş, ardından tüm model varyantları için INBreast veri kümesi üzerinde ince ayar analizleri uygulanmıştır. Sıfırdan eğitim sonuçları sırasıyla Tablo 4.53, Tablo 4.54 ve Tablo 4.55’te, elde edilen ince ayar sonuçları ise sırasıyla Tablo 4.56, Tablo 4.57 ve Tablo 4.58’de sunulmuştur.

**Tablo 4.53.** CBIS-DDSM + VinDr-Mammo veri kümesinde GELAN-s + DWConv + dikkat modüllerinin sıfırdan eğitim performans sonuçları

| Model                   | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|-------------------------|----------|------------|-------|-------|--------------|
| GELAN-s + DWConv + SE   | 0.593    | 0.493      | 0.479 | 0.568 | 0.523        |
| GELAN-s + DWConv + ECA  | 0.579    | 0.556      | 0.481 | 0.59  | 0.533        |
| GELAN-s + DWConv + CBAM | 0.494    | 0.49       | 0.439 | 0.487 | 0.463        |
| GELAN-s + DWConv + NAM  | 0.58     | 0.479      | 0.498 | 0.532 | 0.515        |
| GELAN-s + DWConv + CA   | 0.637    | 0.486      | 0.461 | 0.559 | 0.51         |
| GELAN-s + DWConv + EMA  | 0.538    | 0.568      | 0.513 | 0.593 | <b>0.553</b> |

**Tablo 4.54.** CBIS-DDSM + VinDr-Mammo veri kümesinde GELAN-m + DWConv + dikkat modüllerinin sıfırdan eğitim performans sonuçları

| Model                   | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5     |
|-------------------------|----------|------------|-------|-------|-------------|
| GELAN-m + DWConv + SE   | 0.518    | 0.496      | 0.491 | 0.529 | 0.51        |
| GELAN-m + DWConv + ECA  | 0.606    | 0.476      | 0.48  | 0.57  | 0.525       |
| GELAN-m + DWConv + CBAM | 0.572    | 0.57       | 0.505 | 0.555 | <b>0.53</b> |
| GELAN-m + DWConv + NAM  | 0.53     | 0.529      | 0.47  | 0.55  | 0.51        |
| GELAN-m + DWConv + CA   | 0.582    | 0.498      | 0.475 | 0.576 | 0.526       |
| GELAN-m + DWConv + EMA  | 0.548    | 0.531      | 0.498 | 0.555 | 0.527       |

**Tablo 4.55.** CBIS-DDSM + VinDr-Mammo veri kümesinde GELAN-c + DWConv + dikkat modüllerinin sıfırdan eğitim performans sonuçları

| Model                   | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|-------------------------|----------|------------|-------|-------|--------------|
| GELAN-c + DWConv + SE   | 0.53     | 0.482      | 0.434 | 0.542 | 0.531        |
| GELAN-c + DWConv + ECA  | 0.601    | 0.573      | 0.542 | 0.605 | 0.574        |
| GELAN-c + DWConv + CBAM | 0.552    | 0.519      | 0.466 | 0.572 | 0.519        |
| GELAN-c + DWConv + NAM  | 0.599    | 0.575      | 0.548 | 0.606 | <b>0.577</b> |

**Tablo 4.55.** devam ediyor.

|                           |       |       |       |       |       |
|---------------------------|-------|-------|-------|-------|-------|
| GELAN-c + DWConv<br>+ CA  | 0.587 | 0.533 | 0.519 | 0.578 | 0.549 |
| GELAN-c + DWConv<br>+ EMA | 0.604 | 0.575 | 0.53  | 0.609 | 0.569 |

Tablo 4.53, Tablo 4.54 ve Tablo 4.55’te sunulan DWConv uygulaması ile elde edilen sıfırdan eğitim sonuçlarına göre;

- GELAN-s modelleri arasında, DWConv eklenmeden önce GELAN-s + CA modeli 0.569 mAP@0.5 değeri elde etmişken (Bkz. Tablo 4.46), DWConv sonrası en yüksek mAP@0.5 değeri 0.553 ile GELAN-s + DWConv + EMA modelinde gözlemlenmiştir.
- GELAN-m modellerinde, DWConv eklenmeden önce GELAN-m + ECA modeli 0.535 mAP@0.5 değeri elde etmişken (Bkz. Tablo 4.47), DWConv sonrası en yüksek mAP@0.5 değeri 0.53 ile GELAN-m + DWConv + CBAM modeline aittir.
- DWConv entegrasyonu bazı GELAN-s ve GELAN-m model varyantlarında performans artışı sağlamış olsa da genel olarak DWConv öncesinde elde edilen en iyi performans değerini aşamamıştır.
- GELAN-c modelleri arasında ise DWConv eklenmeden önce GELAN-c + NAM modeli 0.55 mAP@0.5 değeri elde etmişken (Bkz. Tablo 4.48), DWConv sonrası bu modelin mAP@0.5 değeri 0.577’ye yükselmiştir.

**Tablo 4.56.** CBIS-DDSM + VinDr-Mammo → INBreast: GELAN-s + DWConv + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi

| Model                   | Veri artırma parametresi<br>flipud (0.5) | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|-------------------------|--|----------|------------|-------|-------|--------------|
| GELAN-s + DWConv + SE   | -  | 0.81     | 0.714      | 0.74  | 0.883 | 0.812        |
|                         | ✓  | 0.839    | 0.73       | 0.719 | 0.87  | 0.794        |
| GELAN-s + DWConv + ECA  | -  | 0.833    | 0.749      | 0.759 | 0.874 | 0.816        |
|                         | ✓  | 0.796    | 0.686      | 0.686 | 0.884 | 0.785        |
| GELAN-s + DWConv + CBAM | -  | 0.79     | 0.73       | 0.77  | 0.866 | 0.818        |
|                         | ✓  | 0.772    | 0.741      | 0.753 | 0.837 | 0.795        |
| GELAN-s + DWConv + NAM  | -  | 0.827    | 0.722      | 0.73  | 0.90  | 0.82         |
|                         | ✓  | 0.826    | 0.722      | 0.733 | 0.865 | 0.799        |
| GELAN-s + DWConv + CA   | -  | 0.785    | 0.753      | 0.696 | 0.864 | 0.78         |
|                         | ✓  | 0.837    | 0.77       | 0.775 | 0.876 | <b>0.826</b> |
| GELAN-s + DWConv + EMA  | -  | 0.825    | 0.762      | 0.767 | 0.861 | 0.814        |
|                         | ✓  | 0.822    | 0.732      | 0.723 | 0.876 | 0.80         |

**Tablo 4.57.** CBIS-DDSM + VinDr-Mammo → INBreast: GELAN-m + DWConv + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi

| Model                   | Veri artırma parametresi flipud (0.5) | Kesinlik | Duyarlılık | BAP   | MAP   | mAP@0.5      |
|-------------------------|---------------------------------------|----------|------------|-------|-------|--------------|
| GELAN-m + DWConv + SE   | -                                     | 0.848    | 0.74       | 0.765 | 0.881 | 0.823        |
|                         | ✓                                     | 0.864    | 0.749      | 0.772 | 0.901 | 0.837        |
| GELAN-m + DWConv + ECA  | -                                     | 0.77     | 0.75       | 0.746 | 0.893 | 0.82         |
|                         | ✓                                     | 0.836    | 0.775      | 0.762 | 0.882 | 0.822        |
| GELAN-m + DWConv + CBAM | -                                     | 0.849    | 0.738      | 0.781 | 0.852 | 0.817        |
|                         | ✓                                     | 0.82     | 0.67       | 0.692 | 0.856 | 0.774        |
| GELAN-m + DWConv + NAM  | -                                     | 0.812    | 0.767      | 0.721 | 0.897 | 0.809        |
|                         | ✓                                     | 0.775    | 0.745      | 0.729 | 0.889 | 0.809        |
| GELAN-m + DWConv + CA   | -                                     | 0.865    | 0.81       | 0.824 | 0.903 | <b>0.864</b> |
|                         | ✓                                     | 0.91     | 0.70       | 0.733 | 0.908 | 0.821        |
| GELAN-m + DWConv + EMA  | -                                     | 0.86     | 0.741      | 0.776 | 0.845 | 0.811        |
|                         | ✓                                     | 0.805    | 0.733      | 0.761 | 0.871 | 0.82         |

**Tablo 4.58.** CBIS-DDSM + VinDr-Mammo → INBreast: GELAN-c + DWConv + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi

| Model                   | Veri artırma parametresi flipud (0.5) | Kesinlik | Duyarlılık | BAP   | MAP   | mAP@0.5      |
|-------------------------|---------------------------------------|----------|------------|-------|-------|--------------|
| GELAN-c + DWConv + SE   | -                                     | 0.857    | 0.75       | 0.783 | 0.891 | 0.837        |
|                         | ✓                                     | 0.83     | 0.751      | 0.784 | 0.893 | 0.838        |
| GELAN-c + DWConv + ECA  | -                                     | 0.844    | 0.789      | 0.771 | 0.899 | 0.835        |
|                         | ✓                                     | 0.79     | 0.812      | 0.755 | 0.894 | 0.824        |
| GELAN-c + DWConv + CBAM | -                                     | 0.86     | 0.793      | 0.853 | 0.90  | <b>0.878</b> |
|                         | ✓                                     | 0.744    | 0.797      | 0.706 | 0.875 | 0.79         |
| GELAN-c + DWConv + NAM  | -                                     | 0.889    | 0.775      | 0.821 | 0.884 | 0.85         |
|                         | ✓                                     | 0.815    | 0.809      | 0.793 | 0.90  | 0.845        |
| GELAN-c + DWConv + CA   | -                                     | 0.852    | 0.779      | 0.809 | 0.906 | 0.858        |
|                         | ✓                                     | 0.832    | 0.779      | 0.746 | 0.902 | 0.82         |
| GELAN-c + DWConv + EMA  | -                                     | 0.845    | 0.80       | 0.80  | 0.91  | 0.855        |
|                         | ✓                                     | 0.764    | 0.809      | 0.78  | 0.869 | 0.822        |

Tablo 4.56, Tablo 4.57 ve Tablo 4.58’de sunulan DWConv uygulaması ile elde edilen ince ayar sonuçlarına göre;

- GELAN-s modelleri arasında, DWConv eklenmeden önce GELAN-s + EMA modeli 0.833 mAP@0.5 değeri elde etmişken (Bkz. Tablo 4.49), DWConv sonrası en yüksek mAP@0.5 değeri 0.826 ile GELAN-s + DWConv + CA modelinde gözlemlenmiştir.
- GELAN-m modellerinde, DWConv eklenmeden önce GELAN-m + SE modeli 0.817 mAP@0.5 değeri elde etmişken (Bkz. Tablo 4.50), DWConv sonrası en yüksek mAP@0.5 değeri 0.864 ile GELAN-m + DWConv + CA modeline aittir.
- GELAN-c modelleri arasında ise DWConv eklenmeden önce GELAN-c + SE ve GELAN-c + NAM modelleri 0.838 mAP@0.5 değeri elde etmişken (Bkz. Tablo 4.51), DWConv sonrası GELAN-c + DWConv + CBAM modeli 0.878 ile en yüksek mAP@0.5 değerine ulaşmıştır.
- Genel olarak değerlendirildiğinde, DWConv uygulaması GELAN-s modellerinin ince ayar performansına anlamlı bir katkı sağlamasa da GELAN-m ve GELAN-c modellerinde performansı belirgin şekilde arttırmıştır. Bu sonuçlar, DWConv entegrasyonunun daha küçük GELAN-s modeline kıyasla, daha büyük ve derin mimarilere sahip GELAN-m ve GELAN-c modelleri üzerinde daha etkili olduğunu göstermektedir. Ayrıca, en yüksek performansa varsayılan veri artırma konfigürasyonu ile ulaşılmıştır.

GELAN + DWConv + dikkat modülleri ile yapılan analizler sonucunda, en yüksek ince ayar performansı GELAN-c modellerinde elde edilmiştir. Bu doğrultuda, Başkent veri kümesi üzerinde yalnızca GELAN-c + DWConv + dikkat modülü içeren modeller ile ince ayar gerçekleştirilmiştir. Elde edilen sonuçlar Tablo 4.59’da sunulmuştur.

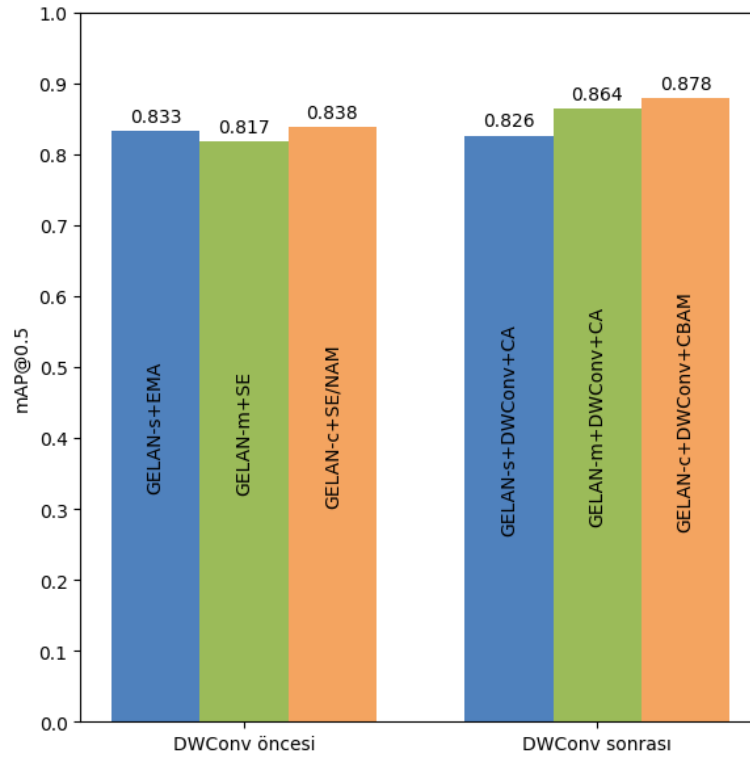
**Tablo 4.59.** CBIS-DDSM + VinDr-Mammo → Başkent: GELAN-c + DWConv + dikkat modülleri için veri artırma parametresinin (flipud) ince ayar performansına etkisi

| Model                   | Veri artırma parametresi flipud (0.5) | Kesinlik | Duyarlılık | BAP   | MAP   | mAP@0.5      |
|-------------------------|---------------------------------------|----------|------------|-------|-------|--------------|
| GELAN-c + DWConv + SE   | -                                     | 0.821    | 0.75       | 0.823 | 0.841 | 0.832        |
|                         | ✓                                     | 0.79     | 0.766      | 0.795 | 0.871 | 0.833        |
| GELAN-c + DWConv + ECA  | -                                     | 0.789    | 0.734      | 0.794 | 0.818 | 0.806        |
|                         | ✓                                     | 0.79     | 0.717      | 0.794 | 0.806 | 0.80         |
| GELAN-c + DWConv + CBAM | -                                     | 0.776    | 0.772      | 0.797 | 0.823 | 0.81         |
|                         | ✓                                     | 0.799    | 0.71       | 0.782 | 0.806 | 0.794        |
| GELAN-c + DWConv + NAM  | -                                     | 0.814    | 0.755      | 0.82  | 0.812 | 0.816        |
|                         | ✓                                     | 0.76     | 0.711      | 0.763 | 0.81  | 0.785        |
| GELAN-c + DWConv + CA   | -                                     | 0.852    | 0.706      | 0.808 | 0.795 | 0.802        |
|                         | ✓                                     | 0.801    | 0.69       | 0.783 | 0.776 | 0.779        |
| GELAN-c + DWConv + EMA  | -                                     | 0.873    | 0.715      | 0.826 | 0.848 | 0.837        |
|                         | ✓                                     | 0.789    | 0.772      | 0.833 | 0.844 | <b>0.838</b> |

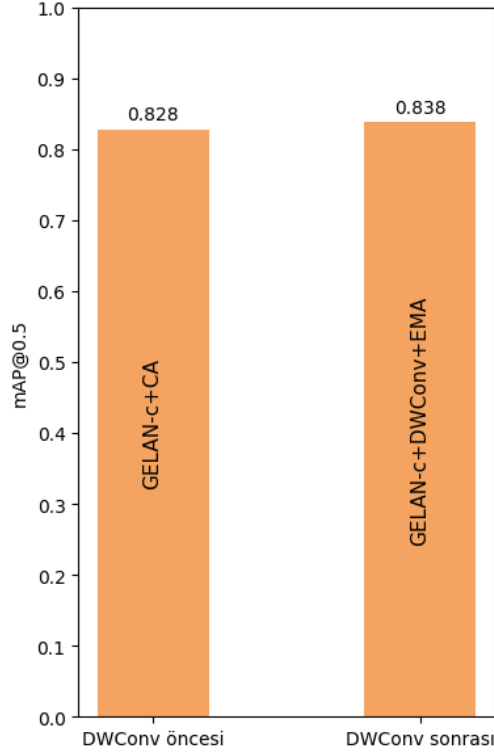
Tablo 4.59’da sunulan Başkent veri kümesi üzerindeki ince ayar sonuçları değerlendirildiğinde, DWConv öncesinde en yüksek mAP@0.5 değeri GELAN-c + CA modeliyle 0.828 olarak elde edilirken (Bkz. Tablo 4.52), DWConv sonrası en yüksek performans GEAN-c + DWConv + EMA modeli ve varsayılan + flipud kombinasyonu ile 0.838’lik mAP@0.5 değerine ulaşmıştır.

#### 4.12. Dikkat Modülleri Entegre Edilmiş GELAN Modellerinde DWConv Uygulamasının Performans Karşılaştırması

Bu bölümde, GELAN-s/m/c + dikkat modülü modellerine DWConv uygulanarak oluşturulan model varyantlarının, kitle tespiti ve sınıflandırma performansı üzerindeki etkileri değerlendirilmiştir. Bu kapsamda, birleşim veri kümesi (CBIS-DDSM + VinDr-Mammo) üzerinden DWConv entegrasyonu öncesi ve sonrası elde edilen sonuçlar karşılaştırmalı olarak analiz edilmiştir. Karşılaştırmalar mAP@0.5 performans ölçütü üzerinden gerçekleştirilmiştir. INBreast ve Başkent veri kümeleri üzerinde elde edilen karşılaştırmalı en iyi ince ayar sonuçları sırasıyla Şekil 4.24 ve Şekil 4.25’te görselleştirilmiştir.



Şekil 4.24. DWConv entegrasyonu öncesi ve sonrası en başarılı GELAN + dikkat modülü model varyantlarının INBreast üzerindeki performans karşılaştırması.



**Şekil 4.25.** DWConv entegrasyonu öncesi ve sonrası en başarılı GELAN-c + dikkat modülü model varyantlarının Başkent üzerindeki performans karşılaştırması.

Şekil 4.24’te INBreast veri kümesi üzerinde sunulan karşılaştırmalı sonuçlar incelendiğinde;

- DWConv entegrasyonu öncesi ve sonrası en başarılı GELAN + dikkat modülü model varyantları değerlendirildiğinde, DWConv uygulaması GELAN-s modelinde performans düşüşüne neden olurken, GELAN-m ve GELAN-c modellerinde belirgin bir performans artışı sağlamıştır.
- Ayrıca, DWConv entegrasyonu ile birlikte en yüksek başarıyı sağlayan dikkat modüllerinde farklılık gözlemlenmiştir.
- Özellikle, GELAN-c + DWConv + CBAM modeli varsayılan veri artırma konfigürasyonu ile 0.878’lik mAP@0.5 değerine ulaşmış ve şimdiye kadar INBreast veri kümesi üzerinde elde edilen en yüksek performansı göstermiştir.

Şekil 4.25’te Başkent veri kümesi üzerinde sunulan sonuçlar incelendiğinde ise;

- Karşılaştırılan en başarılı model varyantları dikkate alındığında, DWConv entegrasyonu GELAN-c + dikkat modülü modelinde performans artışı sağlamıştır.
- DWConv öncesinde GELAN-c modeli CA ile en iyi performansı verirken, DWConv sonrası EMA ile daha yüksek bir başarı elde edilmiştir. Bu durum, DWConv sonrası en başarılı dikkat modülünün değiştiğini göstermektedir.

#### 4.13. YOLO Tabanlı En Başarılı Model Varyantlarının Performans Karşılaştırması

Bu bölümde, yürütülen deneysel çalışmalar sonucunda her bir YOLO versiyonu (YOLOv5, YOLOv8 ve GELAN) için uygulanan dikkat modülleri ve DWConv entegrasyonu sonrası elde edilen en yüksek performansa sahip model varyantları sunulmaktadır. En başarılı sonuçlar, CBIS-DDSM ve VinDr-Mammo veri kümelerinin birleşimi ile gerçekleştirilen eğitimler sonucunda elde edilmiştir. İlgili model varyantlarının INBreast ve Başkent veri kümeleri üzerindeki performansları sırasıyla Tablo 4.60 ve Tablo 4.61’de özetlenmiştir.

Bu karşılaştırmada, GELAN-c modelleri için yalnızca flipud veri artırma parametresi ayrı olarak değerlendirilmiştir. Bunun nedeni, mixup parametresinin GELAN mimarilerinde varsayılan veri artırma konfigürasyonuna dahil edilmiş olmasıdır. Dolayısıyla, GELAN-c model varyantlarında mixup sabit tutulmuş, yalnızca ek olarak uygulanan flipud parametresinin etkisi analiz edilmiştir.

Ayrıca, bu modellerin farklı veri kümeleri üzerindeki başarılarında gözlemlenen değişikliklerin yalnızca mimari yapıdan değil, aynı zamanda görüntü özelliklerinden de kaynaklanmış olabileceği düşünülmektedir. Şekil 3.1’de sunulan histogram analizine göre, INBreast veri kümesinde piksel yoğunluğu dağılımının daha dengeli ve geniş aralıklı olduğu; buna karşılık, Başkent veri kümesinin daha sınırlı bir dinamik aralığa ve düşük kontrasta sahip olduğu görülmektedir. Bu farklılık, modelin öznelik öğrenme kapasitesini ve dolayısıyla genel performansını doğrudan etkileyebilecek önemli bir unsur olarak değerlendirilebilir.

**Tablo 4.60.** CBIS-DDSM + VinDr-Mammo → INBreast için en başarılı model varyantlarının performans karşılaştırması

| Model                   | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|-------------------------|--------------------------|-------------|----------|------------|-------|-------|--------------|
|                         | flipud (0.5)             | mixup (0.1) |          |            |       |       |              |
| YOLOv5s + SE            | ✓                        | ✓           | 0.801    | 0.814      | 0.79  | 0.917 | 0.854        |
| YOLOv8s + DWConv + CBAM | ✓                        | ✓           | 0.857    | 0.804      | 0.799 | 0.916 | 0.858        |
| GELAN-c + DWConv + CBAM | -                        | varsayılan  | 0.86     | 0.793      | 0.853 | 0.90  | <b>0.878</b> |

**Tablo 4.61.** CBIS-DDSM + VinDr-Mammo → Başkent için en başarılı model varyantlarının performans karşılaştırması

| Model                  | Veri artırma parametresi |             | Kesinlik | Duyarlılık | B AP  | M AP  | mAP@0.5      |
|------------------------|--------------------------|-------------|----------|------------|-------|-------|--------------|
|                        | flipud (0.5)             | mixup (0.1) |          |            |       |       |              |
| YOLOv5s + EMA          | -                        | ✓           | 0.856    | 0.757      | 0.849 | 0.848 | <b>0.848</b> |
| YOLOv8s + DWConv + EMA | ✓                        | -           | 0.826    | 0.755      | 0.811 | 0.842 | 0.827        |
| GELAN-c + DWConv + EMA | ✓                        | varsayılan  | 0.789    | 0.772      | 0.833 | 0.844 | 0.838        |

Tablo 4.60’da INBreast veri kümesi üzerinde elde edilen en başarılı model varyantlarına ait sonuçlar incelendiğinde;

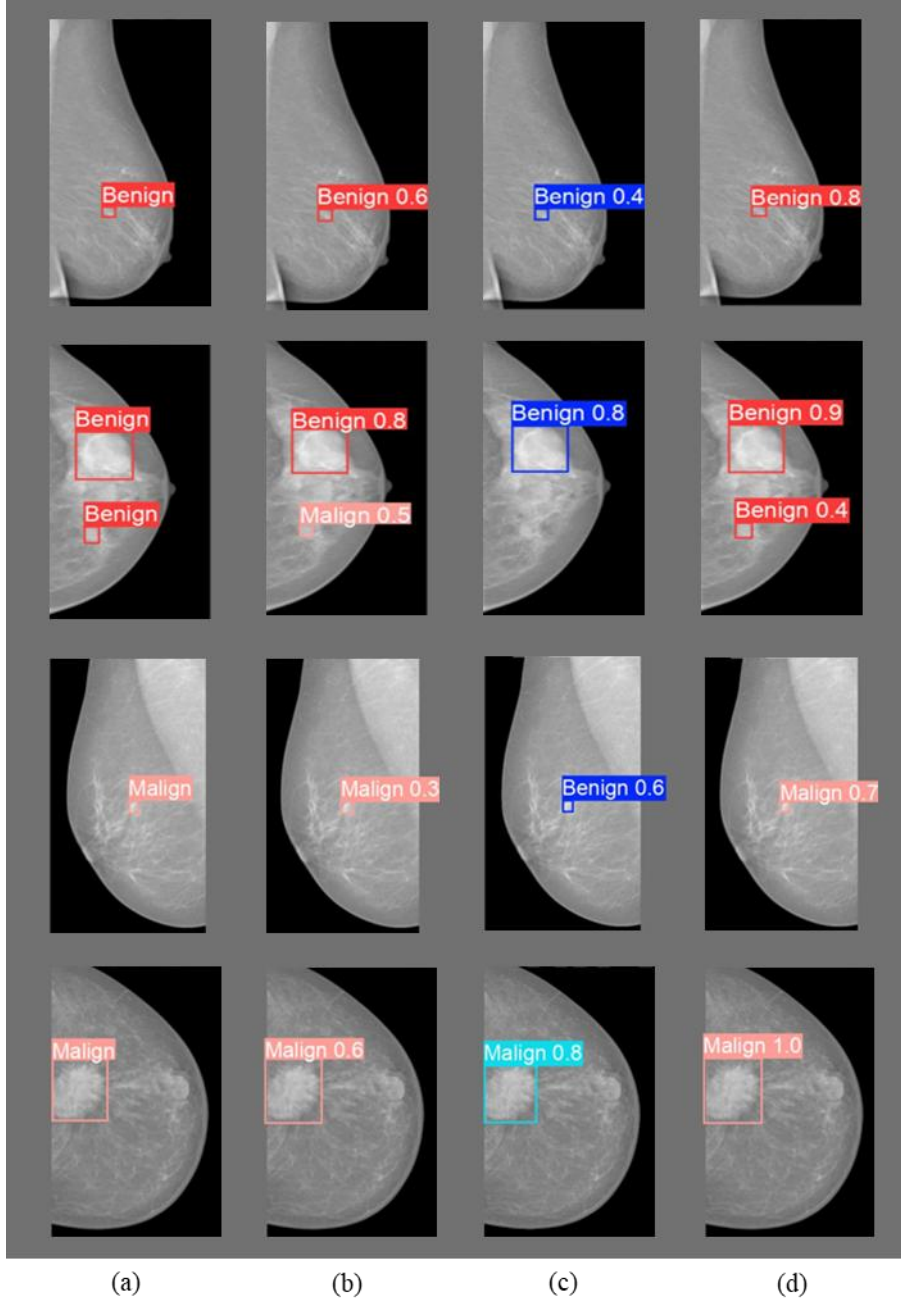
- GELAN-c + DWConv + CBAM model varyantı, 0.878 mAP@0.5 değeriyle en yüksek performansı göstermiştir. Bu model, ek veri artırma yöntemi uygulanmadan dahi diğer modelleri geride bırakmıştır. Bu durum, GELAN mimarisinin güçlü yapısı ile CBAM ve DWConv kombinasyonunun etkili bir şekilde çalıştığını göstermektedir.
- YOLOv5 ve YOLOv8 tabanlı modellerde ise performans artışı, flipud ve mixup veri artırma parametrelerinin varsayılan konfigürasyonlara dahil edilmesiyle sağlanmıştır. Bu parametrelerin özellikle daha hafif mimarilerde başarıyı artırmada önemli katkı sunduğu gözlemlenmiştir.

Tablo 4.61’de, Başkent veri kümesi üzerinde elde edilen en başarılı model varyantlarına ait sonuçlar incelendiğinde;

- YOLOv5s + EMA model varyantı, 0.848 mAP@0.5 değeriyle en yüksek performansı göstermiştir. Bu sonuç, YOLOv5 mimarisinin daha küçük ve hafif yapısının, klinik veri kümesi üzerinde daha iyi genelleme yapabildiğini göstermektedir.
- Ayrıca tüm model varyantlarında en yüksek başarı, EMA dikkat modülü ile elde edilmiştir. Bu durum, EMA’nın bu veri kümesinde tutarlı bir şekilde performansı artırmada etkisi olduğunu göstermektedir.
- Bu veri kümesinde uygulanan modellerde, yalnızca tek bir veri artırma parametresi (flipud veya mixup) kullanılması, performans açısından yeterli olmuştur. Bu durum, veri artırma stratejilerinin etkisinin veri kümesinin yapısına göre değişebileceğini göstermektedir.
- Elde edilen sonuçlar, daha güncel veya karmaşık mimarilerin her zaman daha yüksek performans sağlamadığını göstermektedir. INBreast veri kümesinde en iyi sonuç GELAN tabanlı model ile elde edilirken, Başkent veri kümesinde en yüksek başarı YOLOv5s tabanlı model ile sağlanmıştır. Bu durum, model performansının veri kümesinin yapısal özelliklerine bağlı olarak değişkenlik gösterebileceğini ortaya koymaktadır.

Şekil 4.26’da INBreast veri kümesine ait örnek mamografi görüntüleri üzerinde gerçek etiketler ile sırasıyla en başarılı model varyantlarına ait tespit çıktıları karşılaştırmalı olarak sunulmuştur. Görsel incelemeye göre, GELAN-c + DWConv + CBAM modeli, diğer model

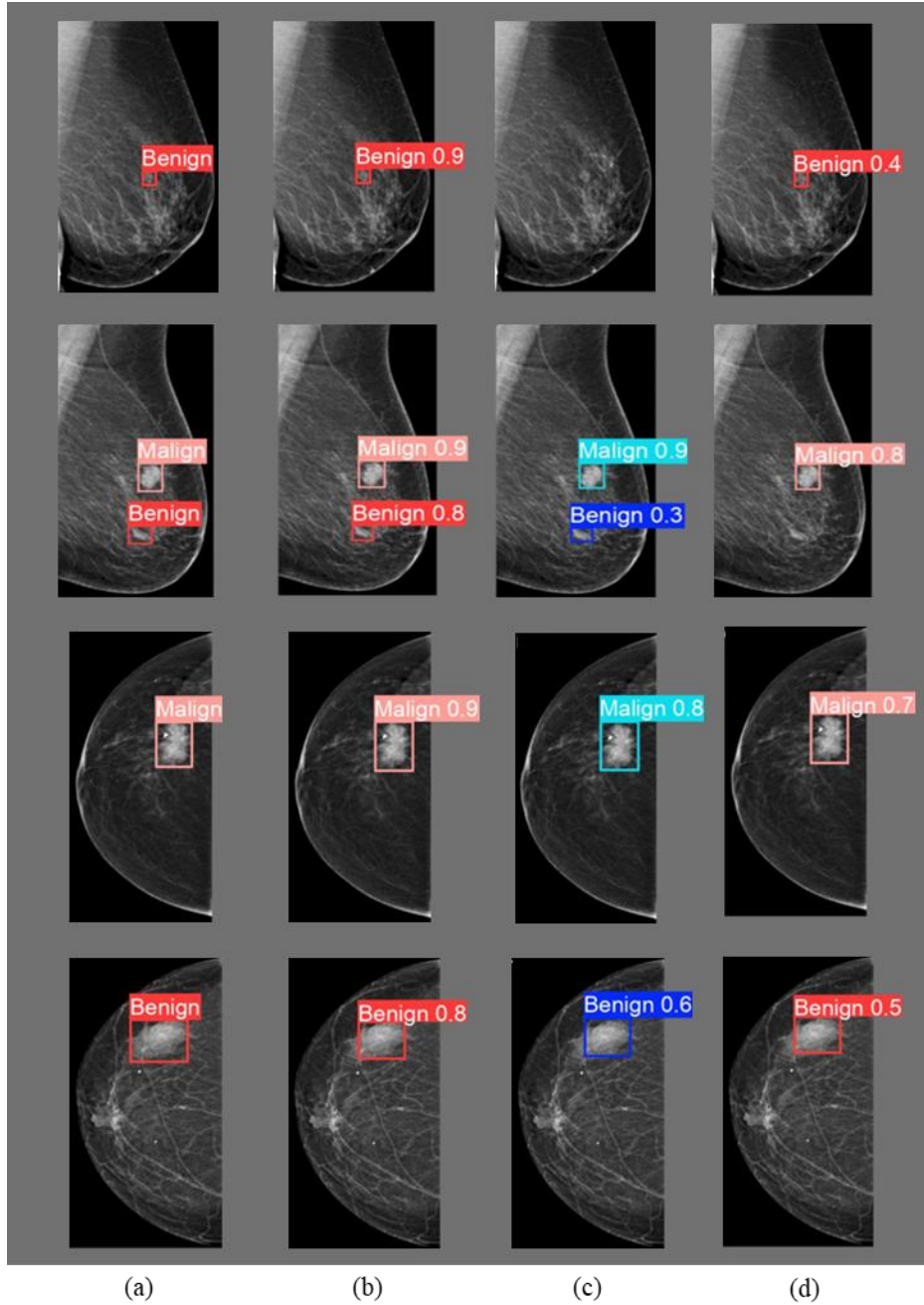
varyantlarına kıyasla hem benign hem de malign kitleleri doğru bir şekilde tespit etme ve sınıflandırmada üstün performans sergilemektedir. Ayrıca bu modelin küçük boyutlu kitlelerde dahi yüksek olasılıklarla doğru tahmin yapabilmesi, GELAN-c mimarisinin hassas hedef tespiti konusunda avantaj sağladığını göstermektedir.



**Şekil 4.26.** INBreast veri kümesi için farklı model varyantlarının tespit sonuçlarının karşılaştırılması: (a) gerçek etiket, (b) YOLOv5s + SE, (c) YOLOv8s + DWConv + CBAM, (d) GELAN-c + DWConv + CBAM.

Şekil 4.27’de Başkent veri kümesine ait örnek mamografi görüntüleri üzerinde gerçek etiketler ile sırasıyla en başarılı model varyantlarına ait tespit çıktıları karşılaştırmalı olarak

sunulmuştur. Görsele göre, YOLOv5s + EMA modeli hem benign hem de malign kitlelerde genel olarak doğru tespit ve sınıflandırma yaparak tutarlı bir başarı sergilemiştir. Diğer model varyantlarında ise bazı kitlelerin hiç tespit edilemediği örnekler dikkat çekmektedir. Bu durum, özellikle küçük boyutlu kitlelerde model hassasiyetinin yetersiz kalabildiğini göstermektedir.



**Şekil 4.27.** Başkent veri kümesi için farklı model varyantlarının tespit sonuçlarının karşılaştırılması: (a) gerçek etiket, (b) YOLOv5s + EMA, (c) YOLOv5s + DWConv + EMA, (d) GELAN-c + DWConv + EMA.

## 5. SONUÇ, TARTIŞMA VE ÖNERİLER

Meme kanseri, dünya çapında milyonlarca kadının sağlığını tehdit eden en yaygın kanserlerden biri haline gelmiştir. Son yıllarda bu hastalığın görülme sıklığı belirgin şekilde artmaktadır. Ancak, erken tanı ve tarama yöntemlerinin yaygınlaşması sayesinde meme kanserine bağlı ölüm oranları küresel düzeyde nispeten düşük kalmaktadır. Meme kanserinin semptomları; mimari bozulma, asimetri, kalsifikasyon ve kitle olmak üzere sınıflandırılabilir. Bu semptomlar arasında, kitle en sık rastlanan ve tanısız açıdan en önemli bulgulardan biridir. Meme kanserinin erken taranması için kullanılan görüntüleme yöntemleri arasında, kullanım kolaylığı ve maliyet etkinliği açısından en çok tercih edileni mamografidir. Ancak, meme dokularının karmaşıklığı, radyologların deneyim düzeyi ve yorgunluk gibi faktörler nedeniyle mamogramların manuel yorumlanması zaman alıcı ve hataya açık olabilir. Bu nedenle, mamografi değerlendirme sürecini desteklemek amacıyla yapay zekâ tabanlı otomatik analiz sistemlerine duyulan ihtiyaç giderek artmaktadır.

Bu tez çalışmasında, mamografi görüntülerinden kitle tespiti ve sınıflandırması için YOLO tabanlı bir model önerilmiştir. Bu amaç doğrultusunda, farklı YOLO versiyonları temel alınarak model varyantları oluşturulmuş; çeşitli dikkat modüllerinin ve evrimsel iyileştirmelerin bu modeller üzerindeki etkileri analiz edilmiştir. Ayrıca, YOLO eğitimi sırasında kullanılacak farklı veri artırma parametrelerinin performans üzerindeki etkileri de değerlendirilmiştir. Önerilen yöntemde, üç farklı eğitim senaryosu oluşturularak CBIS-DDSM, VinDr-Mammo ve CBIS-DDSM + VinDr-Mammo birleşim veri kümeleri ile modeller sıfırdan eğitilmiş; ardından transfer öğrenme yaklaşımıyla INBreast veri kümesi üzerinde 5 katlı çapraz doğrulama yöntemiyle ince ayar gerçekleştirilmiştir. Son olarak, modellerin klinik uygulanabilirliğini değerlendirmek amacıyla Başkent klinik veri kümesi üzerinde de ince ayar analizi yapılmıştır.

Çalışmanın ilk aşamasında, YOLOv5 modeli kullanılmıştır. Bu kapsamda, üç eğitim senaryosu altında hem farklı YOLOv5 modelleri (YOLOv5n, YOLOv5s, YOLOv5m ve YOLOv5s-Transformer), hem de farklı veri artırma konfigürasyonları (düşük ve orta seviye) sıfırdan eğitim ve ince ayar süreçlerinde kapsamlı bir şekilde değerlendirilmiştir. En yüksek performans, YOLOv5s modeli ile orta seviye veri artırma konfigürasyonunun birlikte kullanıldığı kombinasyonda elde edilmiştir. Devamında, veri artırmanın etkisini daha ayrıntılı incelemek amacıyla, varsayılan konfigürasyonda yer almayan üç parametre orta (görüntü yukarı-aşağı çevirme (flipud), görüntü döndürme (degrees) ve görüntü kesme (shear)) analizlere dahil edilmiştir. Sonuç olarak, en yüksek mAP@0.5 skoru olan 0.848,

CBIS-DDSM + VinDr-Mammo → INBreast senaryosuna ait olup, bu başarı orta seviye veri artırma konfigürasyonuna ek olarak üç parametrenin (flipud, degrees, shear) birlikte uygulandığı kombinasyonda elde edilmiştir. Ardından, kitle tespit ve sınıflandırma performansını artırmak amacıyla, YOLOv5s modeline çeşitli dikkat modülleri (SE, ECA, CBAM, NAM, CA ve EMA) entegre edilerek YOLOv5s +dikkat modülü model varyantları oluşturulmuştur. Bu modeller, yine üç eğitim senaryosu altında varsayılan (düşük seviye) veri artırma konfigürasyonu ile analizler gerçekleştirilmiştir. Ayrıca bu aşamada, veri artırmanın etkisini incelemek amacıyla yalnızca görüntü yukarı-aşağı çevirme (flipud) ve görüntü karıştırma (mixup) parametreleri kullanılmıştır. Bu seçimin nedeni, bu aşamada yapılan ön analizlerde diğer veri artırma parametrelerinin (degrees, shear, perspective) model performansı üzerinde anlamlı bir katkı sağlamamış olmasıdır. İlk analizlerde daha fazla veri artırma parametresiyle başarı artışı elde edilirken, dikkat modüllerinin entegrasyonu sonrasında daha az veri artırma parametresiyle daha yüksek performansa ulaşılabildiği gözlemlenmiştir. Yapılan analizler sonucunda, en yüksek mAP@0.5 skoru olan 0.854 yine CBIS-DDSM + VinDr-Mammo → INBreast senaryosuna aittir. Bu başarıya YOLOv5s + SE modeli ve varsayılan + flipud + mixup veri artırma kombinasyonunda ulaşılmıştır. Bu kapsamlı deneyler sonucunda, YOLOv5s + dikkat modülü model varyantları birleşim veri kümesine ait senaryoda daha başarılı olduğu için Başkent klinik veri kümesi üzerinde sadece bu başarılı kombinasyon için ince ayar süreci gerçekleştirilmiştir. CBIS-DDSM + VinDr-Mammo → Başkent senaryosunda, YOLOv5s + EMA modeli ve varsayılan + mixup veri artırma kombinasyonu ile 0.848 mAP@0.5 skoru elde edilmiştir.

YOLOv5s modeline dikkat modüllerinin entegre edilmesiyle elde edilen başarılı sonuçların ardından, çalışmada ikinci model olarak YOLOv8s seçilmiştir. Bu kapsamda, YOLOv8s modeline de farklı dikkat modülleri entegre edilerek oluşturulan YOLOv8s + dikkat modülü model varyantlarının, kitle tespiti ve sınıflandırma performansı üzerindeki etkileri değerlendirilmiştir. Modeller, önceki aşamalarda olduğu gibi üç eğitim senaryosu altında, varsayılan veri artırma konfigürasyonu ile analiz edilmiştir. Bu aşamada yalnızca görüntü yukarı-aşağı çevirme (flipud) ve görüntü karıştırma (mixup) parametreleri kullanılmış; tıpkı YOLOv5 analizlerinde olduğu gibi, bu iki parametrenin YOLOv8 modellerinde de performansa anlamlı katkı sağladığı, diğer veri artırma parametrelerinin ise belirgin bir etkisinin olmadığı gözlemlenmiştir. Bu sonuç, her iki model mimarisi için de veri artırma açısından benzer eğilimler gösterildiğini ortaya koymaktadır. Yapılan değerlendirmeler sonucunda, en yüksek mAP@0.5 skoru olan 0.837, CBIS-DDSM + VinDr-Mammo → INBreast senaryosunda, YOLOv8s + CBAM modeli ve varsayılan + flipud +

mixup veri artırma kombinasyonu ile elde edilmiştir. Bununla birlikte, CBIS-DDSM + VinDr-Mammo → Başkent senaryosunda ise YOLOv8s + NAM modeli ve varsayılan + flipud kombinasyonu ile 0.822 mAP@0.5 skoru gözlemlenmiştir.

Her iki model için INBreast ve Başkent veri kümeleri üzerinde elde edilen en yüksek ince ayar performansları değerlendirildiğinde, en başarılı sonuçların CBIS-DDSM ve VinDr-Mammo veri kümelerinin birleşimiyle eğitilen modeller tarafından sağlandığı görülmüştür. Bu durum, kaynak veri kümesinin sahip olduğu örnek sayısı ve çeşitliliğinin model performansı üzerinde doğrudan etkili olduğunu ortaya koymaktadır. Bu nedenle, çalışmanın devamındaki analizlerde yalnızca birleşim veri kümesine dayalı eğitim senaryosu kullanılmıştır. YOLOv5s + dikkat modülü model varyantları, genel olarak YOLOv8s + dikkat modülü yapılarına kıyasla daha yüksek performans sergilemiş; bu da YOLOv5 modelinin, dikkat modülleri ile birlikte daha etkili sonuçlar üretebildiğini göstermiştir. Dikkat modülleri özelinde değerlendirildiğinde ise, en başarılı modülün hem kullanılan veri kümesine hem de model mimarisine bağlı olarak değiştiği gözlemlenmiştir. Bu bulgu, dikkat modüllerinin etkisinin yalnızca mimari yapıya değil, aynı zamanda ince ayarın uygulandığı hedef veri kümesine bağlı olarak değiştiğini ortaya koymaktadır. Veri artırma parametreleri açısından değerlendirildiğinde ise, her iki modelde de INBreast veri kümesi üzerinde en yüksek başarı, varsayılan + flipud + mixup kombinasyonu ile elde edilmiştir. Öte yandan, Başkent veri kümesinde başarıyı artıran veri artırma parametresi, kullanılan modele göre farklılık göstermiştir. Bu durum, veri artırma parametrelerinin başarısının model mimarisi ve hedef veri kümesinin özelliklerine bağlı olarak değiştiğini göstermektedir.

Çalışmanın devamında, YOLOv5s + dikkat modülü ve YOLOv8s + dikkat modülü model varyantlarının performanslarını artırmak amacıyla, bu modellere DWConv entegrasyonu uygulanmış ve omurga katmanlarının yapısal olarak iyileştirilmesi hedeflenmiştir. Önceki analizlerde, en yüksek performansların CBIS-DDSM + VinDr-Mammo birleşim veri kümesi kullanılarak elde edilmiş olması nedeniyle, DWConv entegrasyonu da yalnızca bu senaryo üzerinden değerlendirilmiştir. CBIS-DDSM + VinDr-Mammo → INBreast senaryosunda DWConv entegrasyonu, YOLOv5s + dikkat modülü modellerinde performans düşüşüne neden olurken, YOLOv8s + dikkat modülü modellerinde belirgin bir performans artışı sağlamıştır. Söz konusu senaryoda, YOLOv8s + DWConv + CBAM modeli ve varsayılan + flipud + mixup veri artırma kombinasyonu ile 0.858 mAP@0.5 skoru ile en yüksek performans elde edilmiştir. CBIS-DDSM + VinDr-Mammo → Başkent senaryosunda ise, DWConv entegrasyonu sonrasında YOLOv5s + dikkat modülü modellerinde yine performans düşüşü gözlenmiş; YOLOv8s + dikkat modülü modellerinde

ise yalnızca sınırlı bir iyileşme sağlanmıştır. Ancak bu iyileşme, önceki analizlerde elde edilen başarı seviyesinin altında kalmıştır. Başkent veri kümesi üzerinde hâlen en yüksek mAP@0.5 skoru olan 0.848, YOLOv5s + EMA modeli ve varsayılan + mixup veri artırma kombinasyonuna aittir.

DWConv entegrasyonunun bazı model varyantlarında sağladığı başarı üzerine, alternatif hafif bir evrişim yapısı olan GhostConv tabanlı YOLOv8s-Ghost modelinin de kitle tespiti ve sınıflandırma performansı araştırılmıştır. Bu kapsamda, YOLOv8s-Ghost ve YOLOv8s-Ghost + dikkat modülleri model varyantları, önceki deneysel koşullarla tutarlı şekilde yeniden değerlendirilmiştir. CBIS-DDSM + VinDr-Mammo → INBreast senaryosunda, YOLOv8s-Ghost modeli, varsayılan + flipud + mixup veri artırma kombinasyonunda 0.861 mAP@0.5 skoru ile şimdiye dek elde edilen en yüksek başarıya ulaşmış; dikkat modülü entegrasyonu olmadan da yüksek performans sergileyebildiği gözlemlenmiştir. CBIS-DDSM + VinDr-Mammo → Başkent senaryosunda ise, en yüksek mAP@0.5 skoru olan 0.817, YOLOv8s-Ghost + ECA modeli ve varsayılan + flipud veri artırma kombinasyonu ile elde edilmiştir. Ancak bu değer, daha önce aynı veri kümesinde YOLOv5s + EMA modeli ve varsayılan + mixup kombinasyonu ile elde edilen 0.848 mAP@0.5 skorunun altında kalmıştır.

Çalışmada son olarak, YOLOv9 ailesine ait GELAN model varyantları (GELAN-s, GELAN-m ve GELAN-c) kullanılmıştır. Bu modeller ile farklı dikkat modüllerinin (SE, ECA, CA, CBAM, NAM ve EMA) entegrasyonu sonucu oluşturulan model varyantlarının, kitle tespiti ve sınıflandırma performansları değerlendirilmiştir. Tüm analizler, yalnızca birleşim veri kümesine dayalı eğitim senaryosu altında ve yüksek seviye varsayılan veri artırma konfigürasyonu ile yürütülmüştür. Bu aşamada, önceki analizlerden elde edilen sonuçlara dayanarak özellikle mixup ve flipud veri artırma parametrelerinin etkisinin incelenmesi planlanmıştır. Ancak mixup zaten varsayılan konfigürasyona dâhil olduğu için yalnızca flipud parametresinin katkısı analiz edilmiştir. CBIS-DDSM + VinDr-Mammo → INBreast senaryosunda, en yüksek mAP@0.5 skoru olan 0.838, hem GELAN-c + SE (varsayılan veri artırma) hem de GELAN-c + NAM (varsayılan + flipud) model varyantlarıyla elde edilmiştir. Ardından, omurga katmanlarını yapısal olarak iyileştirmek amacıyla tüm model varyantlarına DWConv entegrasyonu uygulanmıştır. Bu entegrasyon sonucunda, GELAN-c + DWConv + CBAM modeli, varsayılan veri artırma konfigürasyonunda 0.878'lik mAP@0.5 skoru ile çalışmadaki en yüksek başarıya ulaşmıştır. Elde edilen en başarılı GELAN-c model varyantları, Başkent veri kümesi üzerinde de değerlendirilmiştir. CBIS-DDSM + VinDr-Mammo → Başkent senaryosunda, DWConv

öncesinde en yüksek başarı GELAN-c + CA modeli ile 0.828 mAP@0.5, DWConv sonrasında ise en yüksek skor GELAN-c + DWConv + EMA modeli ve varsayılan + flipud kombinasyonu ile 0.838 mAP@0.5 olarak elde edilmiştir. Ancak bu sonu, hâlen Başkent veri kümesinde ulaşılan en yüksek skor olan 0.848 mAP@0.5'e sahip YOLOv5s + EMA modelinin altında kalmaktadır.

Bu tez çalışmasında elde edilen sonuçlar ile literatürdeki YOLO tabanlı meme kitlesi tespiti çalışmaları Tablo 5.1'de karşılaştırılmıştır. Literatürde birçok çalışmada MIAS ve DDSM gibi mamografi veri kümeleri kullanılmıştır, ancak bu veri kümelerinin bazı temel sınırlılıkları bulunduğu bilinmektedir. MIAS veri kümesi düşük çözünürlüklü görüntüler içermekte olup bu durum, özellikle derin öğrenme tabanlı yöntemlerde modelin öğrenme kapasitesini kısıtlayarak hatalı sonuçlara yol açabilmektedir. Benzer şekilde, DDSM veri kümesinde ise anotasyonların net olmaması ve mamogram kalitesinin düşük olması; hedef bölgelerin yanlış tanımlanmasına, dolayısıyla tespit performansının düşmesine neden olabilmektedir [35]. Bu nedenle, söz konusu veri kümeleriyle yapılan çalışmalar tabloya dahil edilmemiştir.

**Tablo 5.1.** Literatürdeki YOLO tabanlı çalışmalar ile tez kapsamında elde edilen sonuçların karşılaştırılması

| <b>Çalışma</b>                | <b>Veri kümesi</b>                 | <b>Model</b>               | <b>mAP@0.5</b> |
|-------------------------------|------------------------------------|----------------------------|----------------|
| Prinzi et al.<br>[39]         | CBIS-DDSM → INBreast               | YOLOv5s                    | 0.835          |
| Quiñones-Espín et al.<br>[40] | VinDr-Mammo                        | YOLOv5x                    | 0.602          |
| Coskun et al.<br>[45]         | INBreast                           | YOLOv5x + Swin dönüştürücü | 0.947          |
| Hassan et al.<br>[46]         | INBreast                           | YOLOv4 + ViT               | 0.844          |
| Zhang et al.<br>[47]          | INBreast                           | OMHIB-YOLO                 | 0.907          |
| Al-Tam et al.<br>[49]         | INBreast                           | YOLOv8x + ViT              | 0.961          |
| Mahoro et al.<br>[50]         | VinDr-Mammo                        | YOLOv8                     | 0.65           |
| Marchi et al.<br>[52]         | CBIS-DDSM                          | YOLOv9c                    | 0.677          |
| Tez çalışması                 | CBIS-DDSM + VinDr-Mammo → INBreast | GELAN-c + DWConv + CBAM    | 0.878          |
|                               | CBIS-DDSM + VinDr-Mammo → Başkent  | YOLOv5s + EMA              | 0.848          |

Tablo 5.1'deki INBreast veri kümesi üzerinde sonuç sunan literatür çalışmaları incelendiğinde, bazı çalışmaların bu tez çalışmasından daha yüksek başarıya sahip olduğu görülse de bu çalışmalar ortak bazı sınırlılıklar barındırmaktadır.

Coskun et al. [45], INBreast veri kümesi üzerinde gerçekleştirdikleri deneylerde, YOLOv5x modeli ile Swin dönüştürücüyü birleştirerek 0.947 mAP@0.5 skoruna ulaşmışlardır. Model, 5 katlı çapraz doğrulama yöntemiyle değerlendirilmiş ve eğitim setinde yer alan her mamogram, yalnızca döndürme (90°, 180° ve 270°) işlemleriyle manuel üç yeni görüntüye dönüştürülerek artırılmıştır.

Zhang et al. [47] tarafından önerilen OMHIB-YOLO modeli, INBreast veri kümesi üzerinde 0.907 mAP@0.5 doğruluk oranına ulaşmıştır. Çalışmada, INBreast veri kümesinden 112 görüntünün alınarak veri artırma yoluyla toplam 4480 görüntüye çıkarıldığı belirtilmiştir. Ancak bu veri kümesinde kitle anotasyonuna sahip yalnızca 107 mamogram bulunduğu bilinmektedir. Hem bu sayısal tutarsızlık hem de veri kümesinin önemli ölçüde artırılmasına rağmen kullanılan tekniklerin açıkça belirtilmemiş olması, çalışmanın metodolojik şeffaflığını azaltmaktadır.

Al-Tam et al. [49] çalışmasında, INBreast veri kümesi %80 eğitim ve %20 test olacak şekilde rastgele ikiye ayrılmış, eğitim kümesinin %10'u da doğrulama seti olarak kullanılmıştır. Bu sabit bölme yöntemi, özellikle düşük örnek sayısına sahip INBreast veri kümesi gibi küçük boyutlu veri kümelerinde, modelin gerçek performansını yansıtmakta yetersiz kalabilir. Bu tür sınırlı verilerde genellikle 5-katlı çapraz doğrulama yöntemleri kullanılarak modelin genellenebilirliği daha objektif bir şekilde test edilmelidir. Aksi halde test kümesinin rastgele seçimi, özellikle uygun örneklerin denk gelmesi durumunda model başarımını olduğundan yüksek gösterebilir. Ayrıca hem eğitim hem de doğrulama kümelerine veri artırma uygulanması, modelin daha önce görmediği özgün örnekler yerine, benzer türetilmiş görüntüler üzerinden değerlendirilmesine yol açabilir.

Bu çalışmalarda dikkat çeken ortak bir unsur ise manuel yöntemlerle artırılmış verilerin, eğitim sürecinde YOLO mimarisinin varsayılan olarak uyguladığı artırma teknikleriyle de işlenmiş olmasıdır. Bu çifte artırma süreci, aynı mamogramların çok sayıda varyasyonu şeklinde modele tekrar tekrar sunulmasına yol açmış olabilir. Modelin gerçek görüntü çeşitliliği yerine, büyük ölçüde türetilmiş örnekler üzerinden öğrenmesi ise aşırı uyum (overfitting) riskini artırmakta ve elde edilen yüksek başarıların genellenebilirliğini tartışmalı hale getirmektedir.

Metodoloji açısından sunulan tez çalışmasına en yakın çalışma, Prinzi et al. [39] tarafından sunulmuştur. Bu çalışmada, transfer öğrenme yaklaşımını temel alan CBIS-

DDSM → INBreast senaryosu ve YOLOv5s modeli ile elde edilen mAP@0.5 değeri, bu tez kapsamında hem CBIS-DDSM hem de CBIS-DDSM + VinDr-Mammo veri kümeleri ile eğitilen modellerin INBreast üzerindeki ince ayar performanslarının gerisinde kalmıştır. Ayrıca, yalnızca açık erişimli veri kümelerinde değil, gerçek klinik veriler içeren Başkent veri kümesi üzerinde elde edilen sonuçlar da geliştirilen yaklaşımın yüksek başarısını ortaya koymaktadır. Bu durum, yalnızca sınırlı veri ve tek bir senaryo ile ulaşılan başarıların sınırlı kalabileceğini; buna karşılık, veri çeşitliliği ve farklı eğitim stratejileriyle modelin hem genellenebilirliğinin hem de klinik uygulanabilirliğinin artırılabilirliğini ortaya koymaktadır. Özellikle dikkat modülleri ve evrimsel yapıların entegre edilmesiyle model performansında sağlanan artışlar, bu tür mimari iyileştirmelerin katkısını açıkça göstermektedir.

Bu çalışmanın en önemli kısıtlarından biri, daha güncel ve kapsamlı veri kümelerine erişilememiş olmasıdır. Özellikle daha fazla sayıda ve çeşitli örnekler içeren veri kümeleriyle yapılacak çalışmalar, geliştirilen modelin genellenebilirliğini ve güvenilirliğini artıracaktır. Bu nedenle, gelecekte yapılacak çalışmalarda daha geniş ve çok merkezli veri kümeleri kullanılması önerilmektedir.

Ayrıca, geliştirilen modelin klinik ortamda uygulanabilirliğini artırmak için farklı hastanelerden elde edilen bağımsız veri kümeleri ile ek doğrulama çalışmaları yapılmalıdır. Bunun yanı sıra, modelin radyologlarla birlikte kullanılabilirliğini değerlendirmek amacıyla kullanıcı dostu bir arayüz geliştirilmesi ve sistemin klinik iş akışlarına entegre edilebilirliği araştırılmalıdır.

Bununla birlikte, YOLO tabanlı modellerin performansını daha da iyileştirmek amacıyla, ileriye dönük çeşitli mimari geliştirme stratejileri değerlendirilebilir. Farklı dikkat mekanizmalarının kullanımı, dönüştürücü (transformer) tabanlı mimarilere geçiş ve standart evrim katmanları yerine daha hafif veya dinamik yapıların entegrasyonu, modelin hem verimliliğini artırabilir hem de farklı veri koşullarına karşı genellenebilirliğini güçlendirebilir. Ayrıca, mamografi dışındaki görüntüleme modalitelerinin (ultrason veya MR) entegre edilmesi, sistemin genel tanılabilirliğini ve klinik uygulanabilirliğini artıracak önemli bir adım olabilir.

## KAYNAKLAR

- [1] S. M. Shah, R. A. Khan, S. Arif, and U. Sajid, “Artificial intelligence for breast cancer analysis: Trends & directions,” *Comput. Biol. Med.*, vol. 142, no. October 2021, p. 105221, Mar. 2022, doi: 10.1016/j.combiomed.2022.105221.
- [2] N. M. Hassan, S. Hamad, and K. Mahar, “Mammogram breast cancer CAD systems for mass detection and classification: a review,” *Multimed. Tools Appl.*, vol. 81, no. 14, pp. 20043–20075, Jun. 2022, doi: 10.1007/s11042-022-12332-1.
- [3] H. K. Elaibi, F. F. Mutlag, E. Halvaci, A. Aygun, and F. Sen, “Review: Comparison of traditional and modern diagnostic methods in breast cancer,” *Meas. J. Int. Meas. Confed.*, vol. 242, no. PE, p. 116258, 2025, doi: 10.1016/j.measurement.2024.116258.
- [4] S. Z. Ramadan, “Methods Used in Computer-Aided Diagnosis for Breast Cancer Detection Using Mammograms: A Review,” *J. Healthc. Eng.*, vol. 2020, pp. 1–21, Mar. 2020, doi: 10.1155/2020/9162464.
- [5] Y. Jiménez-Gaona, M. J. Rodríguez-Álvarez, and V. Lakshminarayanan, “Deep-Learning-Based Computer-Aided Systems for Breast Cancer Imaging: A Critical Review,” *Appl. Sci.*, vol. 10, no. 22, p. 8298, Nov. 2020, doi: 10.3390/app10228298.
- [6] T. G. Debelee, F. Schwenker, A. Ibenthal, and D. Yohannes, “Survey of deep learning in breast cancer image analysis,” *Evol. Syst.*, vol. 11, no. 1, pp. 143–163, Mar. 2020, doi: 10.1007/s12530-019-09297-2.
- [7] V. Nemade, S. Pathak, and A. K. Dubey, “A Systematic Literature Review of Breast Cancer Diagnosis Using Machine Intelligence Techniques,” *Arch. Comput. Methods Eng.*, vol. 29, no. 6, pp. 4401–4430, Oct. 2022, doi: 10.1007/s11831-022-09738-3.
- [8] J. Mendes and N. Matela, “Breast Cancer Risk Assessment: A Review on Mammography-Based Approaches,” *J. Imaging*, vol. 7, no. 6, p. 98, Jun. 2021, doi: 10.3390/jimaging7060098.

- [9] P. Hogg, J. Kelly, and C. Mercer, “Digital mammography: A holistic approach,” *Digit. Mammogr. A Holist. Approach*, no. January 2015, pp. 1–309, 2015, doi: 10.1007/978-3-319-04831-4.
- [10] X. Yu, Q. Zhou, S. Wang, and Y. Zhang, “A systematic survey of deep learning in breast cancer,” *Int. J. Intell. Syst.*, vol. 37, no. 1, pp. 152–216, Jan. 2022, doi: 10.1002/int.22622.
- [11] S. Zahoor, I. U. Lali, M. A. Khan, K. Javed, and W. Mehmood, “Breast Cancer Detection and Classification using Traditional Computer Vision Techniques: A Comprehensive Review,” *Curr. Med. Imaging Former. Curr. Med. Imaging Rev.*, vol. 16, no. 10, pp. 1187–1200, Jan. 2021, doi: 10.2174/1573405616666200406110547.
- [12] Y. J. Suh, J. Jung, and B. J. Cho, “Automated breast cancer detection in digital mammograms of various densities via deep learning,” *J. Pers. Med.*, vol. 10, no. 4, pp. 1–11, 2020, doi: 10.3390/jpm10040211.
- [13] W. Lotter *et al.*, “Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach,” *Nat. Med.*, vol. 27, no. 2, pp. 244–249, 2021, doi: 10.1038/s41591-020-01174-9.
- [14] L. Kerschke, S. Weigel, A. Rodriguez-Ruiz, N. Karssemeijer, and W. Heindel, “Using deep learning to assist readers during the arbitration process: a lesion-based retrospective evaluation of breast cancer screening performance,” *Eur. Radiol.*, vol. 32, no. 2, pp. 842–852, Feb. 2022, doi: 10.1007/s00330-021-08217-w.
- [15] W. T. Tran *et al.*, “Computational Radiology in Breast Cancer Screening and Diagnosis Using Artificial Intelligence,” *Can. Assoc. Radiol. J.*, vol. 72, no. 1, pp. 98–108, Feb. 2021, doi: 10.1177/0846537120949974.
- [16] S. J. Malebary and A. Hashmi, “Automated Breast Mass Classification System Using Deep Learning and Ensemble Learning in Digital Mammogram,” *IEEE Access*, vol. 9, pp. 55312–55328, 2021, doi: 10.1109/ACCESS.2021.3071297.
- [17] R. Qureshi *et al.*, “A Comprehensive Systematic Review of YOLO for Medical Object

- Detection (2018 to 2023),” *TechRxiv*. Jul. 17, 2023. doi: 10.36227/techrxiv.23681679.v1.
- [18] S. W. Duffy *et al.*, “Mammographic density and breast cancer risk in breast screening assessment cases and women with a family history of breast cancer,” *Eur. J. Cancer*, vol. 88, pp. 48–56, 2018, doi: 10.1016/j.ejca.2017.10.022.
- [19] B. Mughal, M. Sharif, and N. Muhammad, “Bi-model processing for early detection of breast tumor in CAD system,” *Eur. Phys. J. Plus*, vol. 132, no. 6, 2017, doi: 10.1140/epjp/i2017-11523-8.
- [20] R. Rouhi, M. Jafari, S. Kasaei, and P. Keshavarzian, “Benign and malignant breast tumors classification based on region growing and CNN segmentation,” *Expert Syst. Appl.*, vol. 42, no. 3, pp. 990–1002, 2015, doi: 10.1016/j.eswa.2014.09.020.
- [21] M. Dong, X. Lu, Y. Ma, Y. Guo, Y. Ma, and K. Wang, “An Efficient Approach for Automated Mass Segmentation and Classification in Mammograms,” *J. Digit. Imaging*, vol. 28, no. 5, pp. 613–625, 2015, doi: 10.1007/s10278-015-9778-4.
- [22] S. Punitha, A. Amuthan, and K. S. Joseph, “Benign and malignant breast cancer segmentation using optimized region growing technique,” *Futur. Comput. Informatics J.*, vol. 3, no. 2, pp. 348–358, 2018, doi: 10.1016/j.fcij.2018.10.005.
- [23] D. A. Ragab, M. Sharkas, S. Marshall, and J. Ren, “Breast cancer detection using deep convolutional neural networks and support vector machines,” *PeerJ*, vol. 2019, no. 1, pp. 1–23, 2019, doi: 10.7717/peerj.6201.
- [24] N. Dhungel, G. Carneiro, and A. P. Bradley, “A deep learning approach for the analysis of masses in mammograms with minimal user intervention,” *Med. Image Anal.*, vol. 37, pp. 114–128, 2017, doi: 10.1016/j.media.2017.01.009.
- [25] D. Ribli, A. Horváth, Z. Unger, P. Pollner, and I. Csabai, “Detecting and classifying lesions in mammograms with Deep Learning,” *Sci. Rep.*, vol. 8, no. 1, pp. 1–7, 2018, doi: 10.1038/s41598-018-22437-z.

- [26] J. Peng, C. Bao, C. Hu, X. Wang, W. Jian, and W. Liu, “Automated mammographic mass detection using deformable convolution and multiscale features,” *Med. Biol. Eng. Comput.*, vol. 58, no. 7, pp. 1405–1417, 2020, doi: 10.1007/s11517-020-02170-4.
- [27] R. Agarwal, O. Díaz, M. H. Yap, X. Lladó, and R. Martí, “Deep learning for mass detection in Full Field Digital Mammograms,” *Comput. Biol. Med.*, vol. 121, p. 103774, 2020, doi: 10.1016/j.combiomed.2020.103774.
- [28] S. K. Raza and S. M. Syed, “Classification and Segmentation of Breast Tumor Using Mask R- CNN on Mammograms Classification and Segmentation of Breast Tumor using Mask,” 2021, [Online]. Available: <https://www.researchsquare.com/article/rs-523546/latest.pdf>
- [29] G. S. B. Jahangeer and D. R. Thambidurai, “Detecting breast cancer using novel mask R-CNN techniques,” *Expert Syst.*, vol. 39, no. 9, pp. 1–20, 2022, doi: 10.1111/exsy.12954.
- [30] V. K. Singh *et al.*, “Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network,” *Expert Syst. Appl.*, vol. 139, p. 112855, 2020, doi: 10.1016/j.eswa.2019.112855.
- [31] H. Jung *et al.*, “Detection of masses in mammograms using a one-stage object detector based on a deep convolutional neural network,” *PLoS One*, vol. 13, no. 9, pp. 1–16, 2018, doi: 10.1371/journal.pone.0203355.
- [32] G. Hamed, M. Marey, S. E. Amin, and M. F. Tolba, “Automated Breast Cancer Detection and Classification in Full Field Digital Mammograms Using Two Full and Cropped Detection Paths Approach,” *IEEE Access*, vol. 9, pp. 116898–116913, 2021, doi: 10.1109/ACCESS.2021.3105924.
- [33] M. A. Al-masni *et al.*, “Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system,” *Comput. Methods Programs Biomed.*, vol. 157, pp. 85–94, Apr. 2018, doi: 10.1016/j.cmpb.2018.01.017.

- [34] M. A. Al-antari, S. M. Han, and T. S. Kim, "Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms," *Comput. Methods Programs Biomed.*, vol. 196, p. 105584, 2020, doi: 10.1016/j.cmpb.2020.105584.
- [35] G. H. Aly, M. Marey, S. A. El-Sayed, and M. F. Tolba, "YOLO Based Breast Masses Detection and Classification in Full-Field Digital Mammograms," *Comput. Methods Programs Biomed.*, vol. 200, p. 105823, Mar. 2021, doi: 10.1016/j.cmpb.2020.105823.
- [36] Y. Yan, P. H. Conze, M. Lamard, G. Quellec, B. Cochener, and G. Coatrieux, "Towards improved breast mass detection using dual-view mammogram matching," *Med. Image Anal.*, vol. 71, p. 102083, 2021, doi: 10.1016/j.media.2021.102083.
- [37] A. Baccouche, B. Garcia-Zapirain, C. C. Olea, and A. S. Elmaghraby, "Breast lesions detection and classification via YOLO-based fusion models," *Comput. Mater. Contin.*, vol. 69, no. 1, pp. 1407–1425, 2021, doi: 10.32604/cmc.2021.018461.
- [38] Y. Su, Q. Liu, W. Xie, and P. Hu, "YOLO-LOGO: A transformer-based YOLO segmentation model for breast mass detection and segmentation in digital mammograms," *Comput. Methods Programs Biomed.*, vol. 221, p. 106903, 2022, doi: 10.1016/j.cmpb.2022.106903.
- [39] F. Prinzi, M. Insalaco, A. Orlando, S. Gaglio, and S. Vitabile, "A Yolo-Based Model for Breast Cancer Detection in Mammograms," *Cognit. Comput.*, vol. 16, no. 1, pp. 107–120, Jan. 2024, doi: 10.1007/s12559-023-10189-6.
- [40] A. E. Quiñones-Espín, M. Perez-Diaz, R. M. Espín-Coto, D. Rodriguez-Linares, and J. D. Lopez-Cabrera, "Automatic detection of breast masses using deep learning with YOLO approach," *Health Technol. (Berl.)*, vol. 13, no. 6, pp. 915–923, Nov. 2023, doi: 10.1007/s12553-023-00783-x.
- [41] M. Anas, I. U. Haq, G. Husnain, and S. A. F. Jaffery, "Advancing Breast Cancer Detection: Enhancing YOLOv5 Network for Accurate Classification in Mammogram Images," *IEEE Access*, vol. 12, no. February 2024, pp. 16474–16488, 2024, doi:

10.1109/ACCESS.2024.3358686.

- [42] L. Tan, Y. Liang, J. Xia, H. Wu, and J. Zhu, “Detection and Diagnosis of Small Target Breast Masses Based on Convolutional Neural Networks,” *Tsinghua Sci. Technol.*, vol. 29, no. 5, pp. 1524–1539, 2024, doi: 10.26599/TST.2023.9010126.
- [43] Y. Lan, Y. Lv, J. Xu, Y. Zhang, and Y. Zhang, “Breast mass lesion area detection method based on an improved YOLOv8 model,” *Electron. Res. Arch.*, vol. 32, no. 10, pp. 5846–5867, 2024, doi: 10.3934/era.2024270.
- [44] T. Goncalves, I. Rio-Torto, L. F. Teixeira, and J. S. Cardoso, “A Survey on Attention Mechanisms for Medical Applications: Are we Moving Toward Better Algorithms?,” *IEEE Access*, vol. 10, pp. 98909–98935, 2022, doi: 10.1109/ACCESS.2022.3206449.
- [45] D. Coşkun *et al.*, “A comparative study of YOLO models and a transformer-based YOLOv5 model for mass detection in mammograms,” *Turkish J. Electr. Eng. Comput. Sci.*, vol. 31, no. 7, pp. 1294–1313, 2023, doi: 10.55730/1300-0632.4048.
- [46] N. M. Hassan, S. Hamad, and K. Mahar, “YOLO-based CAD framework with ViT transformer for breast mass detection and classification in CESM and FFDM images,” *Neural Comput. Appl.*, vol. 36, no. 12, pp. 6467–6496, 2024, doi: 10.1007/s00521-023-09364-5.
- [47] Y. Zhang, P. Li, Y. Lan, X. Jia, and Y. Lv, “Research on Breast Cancer Detection Methods Based on ODMV-MulDyHead-YOLO,” *IEEE Access*, vol. 12, no. October, pp. 186819–186835, 2024, doi: 10.1109/ACCESS.2024.3508780.
- [48] Y. Du, W. Liu, Y. Wang, R. Li, and L. Xie, “YOLO-CPC: a breast tumor detection and identification algorithm based on improved YOLOv7,” *Signal, Image Video Process.*, vol. 19, no. 3, pp. 1–11, 2025, doi: 10.1007/s11760-024-03811-z.
- [49] R. M. Al-Tam, A. M. Al-Hejri, S. S. Alshamrani, M. A. Al-antari, and S. M. Narangale, “Multimodal breast cancer hybrid explainable computer-aided diagnosis using medical mammograms and ultrasound Images,” *Biocybern. Biomed. Eng.*, vol. 44, no. 3, pp. 731–758, 2024, doi: 10.1016/j.bbe.2024.08.007.

- [50] E. Mahoro and M. A. Akhloufi, “Breast masses detection on mammograms using recent one-shot deep object detectors,” *BioSMART 2023 - Proc. 5th Int. Conf. Bio-Engineering Smart Technol.*, pp. 1–4, 2023, doi: 10.1109/BioSMART58455.2023.10162036.
- [51] F. H. Alhsnony and L. Sellami, “Enhancing Breast Cancer Detection through Advanced Deep Learning: An Application of YOLOv8x on Mammographic Images,” *7th IEEE Int. Conf. Adv. Technol. Signal Image Process. ATSIP 2024*, vol. 1, pp. 128–133, 2024, doi: 10.1109/ATSIP62566.2024.10638868.
- [52] R. Marchi, S. Hau, K. M. Suryaningrum, and R. Yunanda, “Comparing YOLOv8 and YOLOv9 Algorithm on Breast Cancer Detection Case,” *Procedia Comput. Sci.*, vol. 245, no. C, pp. 239–246, 2024, doi: 10.1016/j.procs.2024.10.248.
- [53] I. Ahmad and F. Alqurashi, “Early cancer detection using deep learning and medical imaging: A survey,” *Crit. Rev. Oncol. Hematol.*, vol. 204, no. August, p. 104528, 2024, doi: 10.1016/j.critrevonc.2024.104528.
- [54] X. Wen, X. Guo, S. Wang, Z. Lu, and Y. Zhang, “Breast cancer diagnosis: A systematic review,” *Biocybern. Biomed. Eng.*, vol. 44, no. 1, pp. 119–148, 2024, doi: 10.1016/j.bbe.2024.01.002.
- [55] World Health Organization (WHO), “Global cancer burden growing, amidst mounting need for services,” WHO News. Accessed: Mar. 02, 2025. [Online]. Available: <https://www.who.int/news/item/01-02-2024-global-cancer-burden-growing--amidst-mounting-need-for-services>
- [56] W. H. O. (WHO), “Breast cancer,” WHO News. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>
- [57] J. Kim *et al.*, “Global patterns and trends in breast cancer incidence and mortality across 185 countries.,” *Nat. Med.*, 2025, doi: 10.1038/s41591-025-03502-3.
- [58] A. Alex, E. Bhandary, and K. P. McGuire, “Anatomy and physiology of the breast during pregnancy and lactation,” *Adv. Exp. Med. Biol.*, vol. 1252, pp. 3–7, 2020, doi:

10.1007/978-3-030-41596-9\_1.

- [59] G. Bistoni and J. Farhadi, “Anatomy and Physiology of the Breast,” *Plast. Reconstr. Surg. Approaches Tech.*, no. March 2015, pp. 477–485, 2015, doi: 10.1002/9781118655412.ch37.
- [60] K. P. McGuire, “Breast Anatomy and Physiology,” in *Breast Disease*, vol. 1, Cham: Springer International Publishing, 2016, pp. 1–14. doi: 10.1007/978-3-319-22843-3\_1.
- [61] M. Paruch, “Mathematical modeling of breast tumor destruction using fast heating during radiofrequency ablation,” *Materials (Basel)*, vol. 13, no. 1, 2020, doi: 10.3390/ma13010136.
- [62] Terese Winslow LLC, “Female Breast Anatomy,” National Cancer Institute Visuals Online. Accessed: Apr. 26, 2025. [Online]. Available: <https://visualsonline.cancer.gov/details.cfm?imageid=7127>
- [63] R. M. Tamimi *et al.*, “Population Attributable Risk of Modifiable and Nonmodifiable Breast Cancer Risk Factors in Postmenopausal Breast Cancer,” *Am. J. Epidemiol.*, vol. 184, no. 12, pp. 884–893, Dec. 2016, doi: 10.1093/aje/kww145.
- [64] M. Akram, M. Iqbal, M. Daniyal, and A. U. Khan, “Awareness and current knowledge of breast cancer,” *Biol. Res.*, vol. 50, no. 1, pp. 1–23, 2017, doi: 10.1186/s40659-017-0140-9.
- [65] E. H. Houssein, M. M. Emam, A. A. Ali, and P. N. Suganthan, “Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review,” *Expert Syst. Appl.*, vol. 167, no. April 2020, p. 114161, 2021, doi: 10.1016/j.eswa.2020.114161.
- [66] “Types of Breast Cancer,” Saint John’s Cancer Institute. Accessed: Apr. 26, 2025. [Online]. Available: <https://www.saintjohnscancer.org/breast/breast-cancer/types-of-breast-cancer/>

- [67] S. Winters, C. Martin, D. Murphy, and N. K. Shokar, *Breast Cancer Epidemiology, Prevention, and Screening*, vol. 151. Elsevier Inc., 2017. doi: 10.1016/bs.pmbts.2017.07.002.
- [68] L. Muçaraku, “Breast Cancer Detection using Deep Learning,” Epoka University, 2024.
- [69] S. Iranmakani *et al.*, “A review of various modalities in breast imaging: technical aspects and clinical outcomes,” *Egypt. J. Radiol. Nucl. Med.*, vol. 51, no. 1, 2020, doi: 10.1186/s43055-020-00175-5.
- [70] H. T. Nguyen *et al.*, “VinDr-Mammo: A large-scale benchmark dataset for computer-aided diagnosis in full-field digital mammography,” *Sci. Data*, vol. 10, no. 1, p. 277, May 2023, doi: 10.1038/s41597-023-02100-7.
- [71] M. Madani, M. M. Behzadi, and S. Nabavi, “The Role of Deep Learning in Advancing Breast Cancer Detection Using Different Imaging Modalities: A Systematic Review,” *Cancers (Basel)*, vol. 14, no. 21, pp. 1–36, 2022, doi: 10.3390/cancers14215334.
- [72] I. Sechopoulos, J. Teuwen, and R. Mann, “Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art,” *Semin. Cancer Biol.*, vol. 72, no. May 2020, pp. 214–225, 2021, doi: 10.1016/j.semcancer.2020.06.002.
- [73] J. Tang, R. M. Rangayyan, J. Xu, I. E. El Naqa, and Y. Yang, “Computer-aided detection and diagnosis of breast cancer with mammography: Recent advances,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 2, pp. 236–251, 2009, doi: 10.1109/TITB.2008.2009441.
- [74] S. Siddeeq, J. Li, H. M. A. Bhatti, A. Manzoor, and U. Subhan Malhi, “Deep Learning RN-BCNN Model for Breast Cancer BI-RADS Classification,” *ACM Int. Conf. Proceeding Ser.*, pp. 219–225, 2021, doi: 10.1145/3447587.3447620.
- [75] L. Falconi, M. Perez, W. Aguilar, and A. Conci, “Transfer learning and fine tuning in mammogram bi-rads classification,” *Proc. - IEEE Symp. Comput. Med. Syst.*, vol.

2020-July, pp. 475–480, 2020, doi: 10.1109/CBMS49503.2020.00096.

- [76] B. Ibromkhimov and J. Y. Kang, “Two-Stage Deep Learning Method for Breast Cancer Detection Using High-Resolution Mammogram Images,” *Appl. Sci.*, vol. 12, no. 9, 2022, doi: 10.3390/app12094616.
- [77] M. Zhang, B. Mesurolle, M. Theriault, S. Meterissian, and E. A. Morris, “Imaging of breast cancer—beyond the basics,” *Curr. Probl. Cancer*, vol. 47, no. 2, p. 100967, 2023, doi: 10.1016/j.currproblcancer.2023.100967.
- [78] H. R. F. D. Pria *et al.*, “Current Status of Imaging for Breast Cancer Staging,” *Curr. Breast Cancer Rep.*, vol. 16, no. 2, pp. 126–133, 2024, doi: 10.1007/s12609-024-00530-9.
- [79] P. Pathak, A. S. Jalal, and R. Rai, “Breast Cancer Image Classification: A Review,” *Curr. Med. Imaging Former. Curr. Med. Imaging Rev.*, vol. 17, no. 6, pp. 720–740, 2020, doi: 10.2174/0929867328666201228125208.
- [80] P. Oza, P. Sharma, S. Patel, and A. Bruno, “A bottom-up review of image analysis methods for suspicious region detection in mammograms,” *J. Imaging*, vol. 7, no. 9, 2021, doi: 10.3390/jimaging7090190.
- [81] A. Oliver *et al.*, “A review of automatic mass detection and segmentation in mammographic images,” *Med. Image Anal.*, vol. 14, no. 2, pp. 87–110, 2010, doi: 10.1016/j.media.2009.12.005.
- [82] L. Sturesdotter, M. Sandsveden, K. Johnson, A. M. Larsson, S. Zackrisson, and H. Sartor, “Mammographic tumour appearance is related to clinicopathological factors and surrogate molecular breast cancer subtype,” *Sci. Rep.*, vol. 10, no. 1, pp. 1–10, 2020, doi: 10.1038/s41598-020-77053-7.
- [83] International Agency for Research on Cancer (IARC), “Atlas of breast cancer early detection, Calcifications.” Accessed: Apr. 27, 2025. [Online]. Available: <https://screening.iarc.fr/atlasbreastdetail.php?Index=055&e=>

- [84] I. Rodriguez-Suarez and V. Cruz-Cruz, “Architectural Distortion of the Breast: Imaging Approach and Management,” *Curr. Radiol. Rep.*, vol. 12, no. 9, pp. 87–95, Jul. 2024, doi: 10.1007/s40134-024-00429-4.
- [85] S. Gaur, V. Dialani, P. J. Slanetz, and R. L. Eisenberg, “Architectural distortion of the breast,” *Am. J. Roentgenol.*, vol. 201, no. 5, pp. 662–670, 2013, doi: 10.2214/AJR.12.10153.
- [86] B. Johnson, “Asymmetries in Mammography.,” *Radiol. Technol.*, vol. 92, no. 3, pp. 281M-298M, Jan. 2021, [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/33472885>
- [87] A. L. Chesebro, N. S. Winkler, R. L. Birdwell, and C. S. Giess, “Developing asymmetries at mammography: A multimodality approach to assessment and management,” *Radiographics*, vol. 36, no. 2, pp. 322–334, 2016, doi: 10.1148/rg.2016150123.
- [88] UCLA Health, “Breast Asymmetry.” Accessed: Apr. 27, 2025. [Online]. Available: <https://www.uclahealth.org/departments/radiology/education/breast-imaging-teaching-resources/birads/breast-asymmetry>
- [89] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso, “INbreast: Toward a Full-field Digital Mammographic Database.,” *Acad. Radiol.*, vol. 19, no. 2, pp. 236–248, 2012, doi: 10.1016/j.acra.2011.09.014.
- [90] R. S. Lee, F. Gimenez, A. Hoogi, K. K. Miyake, M. Gorovoy, and D. L. Rubin, “A curated mammography data set for use in computer-aided detection and diagnosis research,” *Sci. Data*, vol. 4, no. 1, p. 170177, Dec. 2017, doi: 10.1038/sdata.2017.177.
- [91] U. Sirisha, S. P. Praveen, P. N. Srinivasu, P. Barsocchi, and A. K. Bhoi, “Statistical Analysis of Design Aspects of Various YOLO-Based Deep Learning Models for Object Detection,” *Int. J. Comput. Intell. Syst.*, vol. 16, no. 1, p. 126, Aug. 2023, doi: 10.1007/s44196-023-00302-w.
- [92] Y. Xiao *et al.*, “A review of object detection based on deep learning,” *Multimed. Tools*

- Appl.*, vol. 79, no. 33–34, pp. 23729–23791, Sep. 2020, doi: 10.1007/s11042-020-08976-6.
- [93] Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu, “Object Detection with Deep Learning: A Review,” *IEEE Trans. Neural Networks Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, 2019, doi: 10.1109/TNNLS.2018.2876865.
- [94] W. Chen, J. Luo, F. Zhang, and Z. Tian, *A review of object detection: Datasets, performance evaluation, architecture, applications and current trends*, vol. 83, no. 24. Springer US, 2024. doi: 10.1007/s11042-023-17949-4.
- [95] M. Trigka and E. Dritsas, “A Comprehensive Survey of Machine Learning Techniques and Models for Object Detection,” *Sensors*, vol. 25, no. 1, p. 214, Jan. 2025, doi: 10.3390/s25010214.
- [96] S. S. A. Zaidi, M. S. Ansari, A. Aslam, N. Kanwal, M. Asghar, and B. Lee, “A survey of modern deep learning based object detection models,” *Digit. Signal Process. A Rev. J.*, vol. 126, p. 103514, 2022, doi: 10.1016/j.dsp.2022.103514.
- [97] A. B. Amjoud and M. Amrouch, “Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review,” *IEEE Access*, vol. 11, no. March, pp. 35479–35516, 2023, doi: 10.1109/ACCESS.2023.3266093.
- [98] R. Kaur and S. Singh, “A comprehensive review of object detection with deep learning,” *Digit. Signal Process.*, vol. 132, p. 103812, Jan. 2023, doi: 10.1016/j.dsp.2022.103812.
- [99] T. Diwan, G. Anirudh, and J. V. Tembhurne, “Object detection using YOLO: challenges, architectural successors, datasets and applications,” *Multimed. Tools Appl.*, vol. 82, no. 6, pp. 9243–9275, Mar. 2023, doi: 10.1007/s11042-022-13644-y.
- [100] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 580–587, 2014, doi: 10.1109/CVPR.2014.81.

- [101] K. He, X. Zhang, S. Ren, and J. Sun, “Spatial pyramid pooling in deep convolutional networks for visual recognition,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8691 LNCS, no. PART 3, pp. 346–361, 2014, doi: 10.1007/978-3-319-10578-9\_23.
- [102] R. Girshick, “Fast R-CNN,” *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2015 Inter, pp. 1440–1448, 2015, doi: 10.1109/ICCV.2015.169.
- [103] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2017, doi: 10.1109/TPAMI.2016.2577031.
- [104] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, “Feature Pyramid Networks for Object Detection,” *Proc. - 2019 IEEE Intl Conf Parallel Distrib. Process. with Appl. Big Data Cloud Comput. Sustain. Comput. Commun. Soc. Comput. Networking, ISPA/BDCLOUD/SustainCom/SocialCom 2019*, pp. 1500–1504, Dec. 2016, doi: <https://doi.org/10.48550/arXiv.1612.03144>.
- [105] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 2, pp. 386–397, 2020, doi: 10.1109/TPAMI.2018.2844175.
- [106] J. Terven and D. Cordova-Esparza, “A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS,” pp. 1–34, Apr. 2023, doi: 10.3390/make5040083.
- [107] S. S. A. Zaidi, M. S. Ansari, A. Aslam, N. Kanwal, M. Asghar, and B. Lee, “A survey of modern deep learning based object detection models,” *Digit. Signal Process.*, vol. 126, p. 103514, Jun. 2022, doi: 10.1016/j.dsp.2022.103514.
- [108] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal Loss for Dense Object Detection,” Aug. 2017.
- [109] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “The pascal visual object classes (VOC) challenge,” *Int. J. Comput. Vis.*, vol. 88, no. 2, pp.

303–338, 2010, doi: 10.1007/s11263-009-0275-4.

- [110] T. Y. Lin *et al.*, “Microsoft COCO: Common objects in context,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8693 LNCS, no. PART 5, pp. 740–755, 2014, doi: 10.1007/978-3-319-10602-1\_48.
- [111] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You only look once: Unified, real-time object detection,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 779–788, 2016, doi: 10.1109/CVPR.2016.91.
- [112] H. Chen, Z. Chen, and H. Yu, “Enhanced YOLOv5: An Efficient Road Object Detection Method,” *Sensors*, vol. 23, no. 20, p. 8355, Oct. 2023, doi: 10.3390/s23208355.
- [113] A. Vijayakumar and S. Vairavasundaram, *YOLO-based Object Detection Models: A Review and its Applications*, no. 0123456789. Springer US, 2024. doi: 10.1007/s11042-024-18872-y.
- [114] J. Redmon and A. Farhadi, “Yolo V2.0,” *Cvpr2017*, no. April, pp. 187–213, 2017, [Online]. Available: [http://www.worldscientific.com/doi/abs/10.1142/9789812771728\\_0012](http://www.worldscientific.com/doi/abs/10.1142/9789812771728_0012)
- [115] S. Seong, J. Song, D. Yoon, J. Kim, and J. Choi, “Determination of vehicle trajectory through optimization of vehicle bounding boxes using a convolutional neural network,” *Sensors (Switzerland)*, vol. 19, no. 19, 2019, doi: 10.3390/s19194263.
- [116] X. Cong, S. Li, F. Chen, C. Liu, and Y. Meng, “A Review of YOLO Object Detection Algorithms based on Deep Learning,” *Front. Comput. Intell. Syst.*, vol. 4, no. 2, pp. 17–20, 2023, doi: 10.54097/fcis.v4i2.9730.
- [117] M. Hussain, “YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection,” *Machines*, vol. 11, no. 7, 2023, doi: 10.3390/machines11070677.
- [118] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” 2018,

[Online]. Available: <http://arxiv.org/abs/1804.02767>

- [119] X. Zhang, W. Wang, Y. Zhao, and H. Xie, “An improved YOLOv3 model based on skipping connections and spatial pyramid pooling,” *Syst. Sci. Control Eng.*, vol. 9, no. S1, pp. 142–149, 2021, doi: 10.1080/21642583.2020.1824132.
- [120] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” Apr. 2020, [Online]. Available: <http://arxiv.org/abs/2004.10934>
- [121] C. Y. Wang, H. Y. Mark Liao, Y. H. Wu, P. Y. Chen, J. W. Hsieh, and I. H. Yeh, “CSPNet: A new backbone that can enhance learning capability of CNN,” *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2020-June, pp. 1571–1580, 2020, doi: 10.1109/CVPRW50498.2020.00203.
- [122] Z. Huang, J. Wang, X. Fu, T. Yu, Y. Guo, and R. Wang, “DC-SPP-YOLO: Dense connection and spatial pyramid pooling based YOLO for object detection,” *Inf. Sci. (Ny)*, vol. 522, pp. 241–258, 2020, doi: 10.1016/j.ins.2020.02.067.
- [123] M. H. Junos, A. S. Mohd Khairuddin, and M. Dahari, “Automated object detection on aerial images for limited capacity embedded device using a lightweight CNN model,” *Alexandria Eng. J.*, vol. 61, no. 8, pp. 6023–6041, 2022, doi: 10.1016/j.aej.2021.11.027.
- [124] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, “PANet: Path Aggregation Network for Instance Segmentation. (arXiv:1803.01534v3 [cs.CV] UPDATED),” *Cvpr*, pp. 8759–8768, 2019, [Online]. Available: <http://arxiv.org/abs/1803.01534>
- [125] G. Jocher, “YOLOv5 by Ultralytics.” Accessed: Jan. 03, 2024. [Online]. Available: <https://github.com/ultralytics/yolov5>
- [126] F. Wu, Z. Cai, S. Fan, R. Song, L. Wang, and W. Cai, “Fish Target Detection in Underwater Blurred Scenes Based on Improved YOLOv5,” *IEEE Access*, vol. 11, no. November, pp. 122911–122925, 2023, doi: 10.1109/ACCESS.2023.3328940.

- [127] J. Liu, M. Hu, J. Dong, and X. Lu, “The application of a lightweight model FA-YOLOv5 with fused attention mechanism in insulator defect detection,” *Front. Energy Res.*, vol. 11, no. November, pp. 1–15, Nov. 2023, doi: 10.3389/fenrg.2023.1283394.
- [128] W. Meng and Y. Yuan, “SGN-YOLO: Detecting Wood Defects with Improved YOLOv5 Based on Semi-Global Network,” *Sensors*, vol. 23, no. 21, p. 8705, Oct. 2023, doi: 10.3390/s23218705.
- [129] Q. Sun, X. Zhang, Y. Li, and J. Wang, “YOLOv5-OCDS: An Improved Garbage Detection Model Based on YOLOv5,” *Electronics*, vol. 12, no. 16, p. 3403, Aug. 2023, doi: 10.3390/electronics12163403.
- [130] S. Wu *et al.*, “Enhanced YOLOv5 Object Detection Algorithm for Accurate Detection of Adult Rhynchophorus ferrugineus,” *Insects*, vol. 14, no. 8, p. 698, Aug. 2023, doi: 10.3390/insects14080698.
- [131] Y. Jiang, “Surface defect detection of steel based on improved YOLOv5 algorithm,” *Math. Biosci. Eng.*, vol. 20, no. 11, pp. 19858–19870, 2023, doi: 10.3934/mbe.2023879.
- [132] C. Li *et al.*, “YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications,” 2022, [Online]. Available: <http://arxiv.org/abs/2209.02976>
- [133] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, “YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors,” pp. 7464–7475, 2023, doi: 10.1109/cvpr52729.2023.00721.
- [134] J. Glenn, “Ultralytics YOLOv8.” [Online]. Available: <https://github.com/ultralytics/ultralytics>
- [135] N. U. A. Tahir, Z. Long, Z. Zhang, M. Asim, and M. ELAffendi, “PVswin-YOLOv8s: UAV-Based Pedestrian and Vehicle Detection for Traffic Management in Smart Cities Using Improved YOLOv8,” *Drones*, vol. 8, no. 3, pp. 1–20, 2024, doi: 10.3390/drones8030084.

- [136] T. Jiang and S. Chen, "A Lightweight Forest Pest Image Recognition Model Based on Improved YOLOv8," *Appl. Sci.*, vol. 14, no. 5, p. 1941, 2024, doi: 10.3390/app14051941.
- [137] H. Lou *et al.*, "DC-YOLOv8: Small-Size Object Detection Algorithm Based on Camera Sensor," *Electron.*, vol. 12, no. 10, pp. 1–14, 2023, doi: 10.3390/electronics12102323.
- [138] W. Yang *et al.*, "Deformable convolution and coordinate attention for fast cattle detection," *Comput. Electron. Agric.*, vol. 211, no. June, p. 108006, 2023, doi: 10.1016/j.compag.2023.108006.
- [139] X. Wang, H. Gao, Z. Jia, and Z. Li, "BL-YOLOv8: An Improved Road Defect Detection Model Based on YOLOv8," *Sensors (Basel)*, vol. 23, no. 20, 2023, doi: 10.3390/s23208361.
- [140] C.-T. Chien, R.-Y. Ju, K.-Y. Chou, and J.-S. Chiang, "YOLOv8-AM: YOLOv8 with Attention Mechanisms for Pediatric Wrist Fracture Detection," pp. 1–25, 2024, [Online]. Available: <http://arxiv.org/abs/2402.09329>
- [141] Q. Zeng, G. Zhou, L. Wan, L. Wang, G. Xuan, and Y. Shao, "Detection of Coal and Gangue Based on Improved YOLOv8," *Sensors*, vol. 24, no. 4, pp. 1–12, 2024, doi: 10.3390/s24041246.
- [142] H. Hwang, D. Kim, and H. Kim, "A Deep-Learning-Based Quality Control Evaluation Method for CT Phantom Images," *Appl. Sci.*, vol. 14, no. 5, p. 1971, 2024, doi: 10.3390/app14051971.
- [143] M. Hussein and W. X. Zhu, "A real-time ghost machine learning model built on YOLOv8 for traffic road signs detection and classification in Germany," *Multimedia Systems*, vol. 30, no. 6. 2024. doi: 10.1007/s00530-024-01527-1.
- [144] X. Xiao and X. Feng, "Multi-Object Pedestrian Tracking Using Improved YOLOv8 and OC-SORT," *Sensors*, vol. 23, no. 20, p. 8439, 2023, [Online]. Available: <https://www.mdpi.com/1424-8220/23/20/8439>

- [145] Y. Bin Liu, Y. H. Zeng, and J. H. Qin, “GSC-YOLO: a lightweight network for cup and piston head detection,” *Signal, Image Video Process.*, vol. 18, no. 1, pp. 351–360, 2024, doi: 10.1007/s11760-023-02746-1.
- [146] C.-Y. Wang, I.-H. Yeh, and H.-Y. M. Liao, “YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information,” 2024, [Online]. Available: <http://arxiv.org/abs/2402.13616>
- [147] N. Chandra, H. Vaidya, S. Sawant, and S. R. Meena, “A Novel Attention-Based Generalized Efficient Layer Aggregation Network for Landslide Detection from Satellite Data in the Higher Himalayas, Nepal,” *Remote Sens.*, vol. 16, no. 14, 2024, doi: 10.3390/rs16142598.
- [148] R. An, X. Zhang, M. Sun, and G. Wang, “GC-YOLOv9: Innovative smart city traffic monitoring solution,” *Alexandria Eng. J.*, vol. 106, no. July, pp. 277–287, 2024, doi: 10.1016/j.aej.2024.07.004.
- [149] Y. Wang, Q. Rong, and C. Hu, “Ripe Tomato Detection Algorithm Based on Improved YOLOv9,” *Plants*, vol. 13, no. 22, 2024, doi: 10.3390/plants13223253.
- [150] W. Xu, D. Zhu, R. Deng, K. L. Yung, and A. W. H. Ip, “Violence-YOLO: Enhanced GELAN Algorithm for Violence Detection,” *Appl. Sci.*, vol. 14, no. 15, 2024, doi: 10.3390/app14156712.
- [151] N. Dahiya *et al.*, “Hyper-parameter tuned deep learning approach for effective human monkeypox disease detection,” *Scientific Reports*, vol. 13, no. 1. 2023. doi: 10.1038/s41598-023-43236-1.
- [152] L. Ramos, E. Casas, E. Bendek, C. Romero, and F. Rivas-Echeverría, “Hyperparameter optimization of YOLOv8 for smoke and wildfire detection: Implications for agricultural and environmental safety,” *Artif. Intell. Agric.*, vol. 12, pp. 109–126, 2024, doi: 10.1016/j.aiaa.2024.05.003.
- [153] M. M. Bhosale, “Performance Analysis of Momentum of Adam Optimizer on YOLO-V8 Using Traffic Object Dataset Performance Analysis of Momentum of Adam

Optimizer on YOLO-V8 Using Traffic Object,” 2024.

- [154] B. M. Hussein and S. M. Shareef, “An Empirical Study on the Correlation between Early Stopping Patience and Epochs in Deep Learning,” *ITM Web Conf.*, vol. 64, p. 01003, 2024, doi: 10.1051/itmconf/20246401003.
- [155] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *J. Big Data*, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0197-0.
- [156] E. Goceri, *Medical image data augmentation: techniques, comparisons and interpretations*, vol. 56, no. 11. Springer Netherlands, 2023. doi: 10.1007/s10462-023-10453-z.
- [157] P. Kaur, B. S. Khehra, and E. B. S. Mavi, “Data Augmentation for Object Detection: A Review,” in *2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, IEEE, Aug. 2021, pp. 537–543. doi: 10.1109/MWSCAS47672.2021.9531849.
- [158] K. Wang, B. Fang, J. Qian, S. Yang, X. Zhou, and J. Zhou, “Perspective Transformation Data Augmentation for Object Detection,” *IEEE Access*, vol. 8, pp. 4935–4943, 2020, doi: 10.1109/ACCESS.2019.2962572.
- [159] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, “mixup: Beyond Empirical Risk Minimization,” pp. 1–13, Oct. 2017, [Online]. Available: <http://arxiv.org/abs/1710.09412>
- [160] G. Ghiasi *et al.*, “Simple Copy-Paste is a Strong Data Augmentation Method for Instance Segmentation,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 2917–2927, 2021, doi: 10.1109/CVPR46437.2021.00294.
- [161] A. de Santana Correia and E. L. Colomhini, *Attention, please! A survey of neural attention models in deep learning*, vol. 55, no. 8. Springer Netherlands, 2022. doi: 10.1007/s10462-022-10148-x.
- [162] M. H. Guo *et al.*, “Attention mechanisms in computer vision: A survey,” *Comput. Vis.*

- Media*, vol. 8, no. 3, pp. 331–368, 2022, doi: 10.1007/s41095-022-0271-y.
- [163] M. Zheng *et al.*, “Attention-based CNNs for Image Classification: A Survey,” *J. Phys. Conf. Ser.*, vol. 2171, no. 1, 2022, doi: 10.1088/1742-6596/2171/1/012068.
- [164] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, “Squeeze-and-Excitation Networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 8, pp. 2011–2023, 2020, doi: 10.1109/TPAMI.2019.2913372.
- [165] Z. Huangfu and S. Li, “Lightweight You Only Look Once v8: An Upgraded You Only Look Once v8 Algorithm for Small Object Identification in Unmanned Aerial Vehicle Images,” *Appl. Sci.*, vol. 13, no. 22, 2023, doi: 10.3390/app132212369.
- [166] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, “ECA-Net: Efficient channel attention for deep convolutional neural networks,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 11531–11539, 2020, doi: 10.1109/CVPR42600.2020.01155.
- [167] Z. Feng, J. Yang, L. Chen, Z. Chen, and L. Li, “An Intelligent Waste-Sorting and Recycling Device Based on Improved EfficientNet,” *Int. J. Environ. Res. Public Health*, vol. 19, no. 23, 2022, doi: 10.3390/ijerph192315987.
- [168] C. Li, R. Xu, Y. Lv, Y. Zhao, and W. Jing, “Edge Real-Time Object Detection and DPU-Based Hardware Implementation for Optical Remote Sensing Images,” *Remote Sens.*, vol. 15, no. 16, 2023, doi: 10.3390/rs15163975.
- [169] D. Luo, Q. He, M. Ma, K. Yan, D. Liu, and P. Wang, “ECANodule: Accurate Pulmonary Nodule Detection and Segmentation with Efficient Channel Attention,” *Proc. Int. Jt. Conf. Neural Networks*, vol. 2023-June, pp. 1–8, 2023, doi: 10.1109/IJCNN54540.2023.10191210.
- [170] S. Woo, J. Park, J. Y. Lee, and I. S. Kweon, “CBAM: Convolutional block attention module,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11211 LNCS, pp. 3–19, 2018, doi: 10.1007/978-3-030-01234-2\_1.

- [171] C. Zhang, J. Zhao, J. Niu, and D. Li, “New convolutional neural network model for screening and diagnosis of mammograms,” *PLoS One*, vol. 15, no. 8 August, pp. 1–20, 2020, doi: 10.1371/journal.pone.0237674.
- [172] Y. Liu, Z. Shao, Y. Teng, and N. Hoffmann, “NAM: Normalization-based Attention Module,” no. NeurIPS 2021, 2021, [Online]. Available: <http://arxiv.org/abs/2111.12419>
- [173] Y. Zhang, J. Su, T. Wang, C. Xu, and A. Yu, “Intelligent detection method of microparticle virus in silkworm based on YOLOv8 improved algorithm,” *J. Supercomput.*, vol. 80, no. 12, pp. 18118–18141, 2024, doi: 10.1007/s11227-024-06159-w.
- [174] P. Selvam, P. Shanmuga Sundari, T. Suresh, M. Tamilselvi, M. Murugappan, and M. E. H. Chowdhury, “YOLO-SAIL: Attention-Enhanced YOLOv5 with Optimized Bi-FPN for Ship Target Detection in SAR Images,” *IEEE Access*, vol. 13, no. February, pp. 29523–29540, 2025, doi: 10.1109/ACCESS.2025.3536621.
- [175] Q. Hou, D. Zhou, and J. Feng, “Coordinate attention for efficient mobile network design,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 13708–13717, 2021, doi: 10.1109/CVPR46437.2021.01350.
- [176] W. Xu and R. Wang, “ALAD-YOLO:an lightweight and accurate detector for apple leaf diseases,” *Front. Plant Sci.*, vol. 14, no. August, pp. 1–15, 2023, doi: 10.3389/fpls.2023.1204569.
- [177] H. Zou and Z. Wang, “An enhanced object detection network for ship target detection in SAR images,” *J. Supercomput.*, no. 0123456789, 2024, doi: 10.1007/s11227-024-06136-3.
- [178] D. Ouyang *et al.*, “Efficient Multi-Scale Attention Module with Cross-Spatial Learning,” in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, Jun. 2023, pp. 1–5. doi: 10.1109/ICASSP49357.2023.10096516.

- [179] W. Hao, C. Ren, M. Han, L. Zhang, F. Li, and Z. Liu, “Cattle Body Detection Based on YOLOv5-EMA for Precision Livestock Farming,” *Animals*, vol. 13, no. 22, 2023, doi: 10.3390/ani13223535.
- [180] M. Huang, W. Mi, and Y. Wang, “EDGS-YOLOv8: An Improved YOLOv8 Lightweight UAV Detection Model,” *Drones*, vol. 8, no. 7, 2024, doi: 10.3390/drones8070337.
- [181] Z. Jin, X. Li, H. Yang, B. Wu, and X. Zhu, “Depthwise separable convolution Unet for 3D seismic data interpolation,” *Front. Earth Sci.*, vol. 10, no. January, pp. 1–18, 2023, doi: 10.3389/feart.2022.1005505.
- [182] M. S. Kim, A. A. Del Barrio, H. J. Kim, and N. Bagherzadeh, “The Effects of Approximate Multiplication on Convolutional Neural Networks,” *IEEE Trans. Emerg. Top. Comput.*, vol. 10, no. 2, pp. 904–916, 2022, doi: 10.1109/TETC.2021.3050989.
- [183] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, “GhostNet: More features from cheap operations,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1577–1586, 2020, doi: 10.1109/CVPR42600.2020.00165.
- [184] Y. Li, C. Yin, Y. Lei, J. Zhang, and Y. Yan, “RDD-YOLO: Road Damage Detection Algorithm Based on Improved You Only Look Once Version 8,” *Appl. Sci.*, vol. 14, no. 8, p. 3360, Apr. 2024, doi: 10.3390/app14083360.
- [185] L. Gao *et al.*, “A Lightweight YOLOv8 Model for Apple Leaf Disease Detection,” *Appl. Sci.*, vol. 14, no. 15, 2024, doi: 10.3390/app14156710.
- [186] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, “A survey on deep transfer learning,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 11141 LNCS, pp. 270–279, 2018, doi: 10.1007/978-3-030-01424-7\_27.
- [187] H.-K. Jung and G.-S. Choi, “Improved YOLOv5: Efficient Object Detection Using Drone Images under Various Conditions,” *Appl. Sci.*, vol. 12, no. 14, p. 7255, Jul.

2022, doi: 10.3390/app12147255.

- [188] N. Tajbakhsh *et al.*, “Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1299–1312, 2016, doi: 10.1109/TMI.2016.2535302.
- [189] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio, “Transfusion: Understanding transfer learning for medical imaging,” *Adv. Neural Inf. Process. Syst.*, vol. 32, no. NeurIPS, 2019.
- [190] R. Padilla, S. L. Netto, and E. A. B. Da Silva, “A Survey on Performance Metrics for Object-Detection Algorithms,” *Int. Conf. Syst. Signals, Image Process.*, vol. 2020-July, no. July, pp. 237–242, 2020, doi: 10.1109/IWSSIP48289.2020.9145130.
- [191] Z. Ji *et al.*, “Lung Nodule Detection in Medical Images Based on Improved YOLOv5s,” *IEEE Access*, vol. 11, no. July, pp. 76371–76387, 2023, doi: 10.1109/ACCESS.2023.3296530.
- [192] B. K. Karaca, M. F. Akşahin, and R. Öcal, “Detection of multiple sclerosis from photic stimulation EEG signals,” *Biomed. Signal Process. Control*, vol. 67, no. November 2020, pp. 1–6, 2021, doi: 10.1016/j.bspc.2021.102571.
- [193] W. M. Salama, A. M. Elbagoury, and M. H. Aly, “Novel breast cancer classification framework based on deep learning,” *IET Image Process.*, vol. 14, no. 13, pp. 3028–3038, 2020, doi: 10.1049/iet-ipr.2020.0122.