

T.C.
ERZİNCAN BİNALİ YILDIRIM ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
YAPAY ZEKA ve ROBOTİK ANABİLİM DALI

YAPAY ZEKA TABANLI GÖRÜNTÜ İŞLEME TEKNİKLERİ İLE İŞ
GÜVENLİĞİNDE KULLANILAN KİŞİSEL EKİPMANLARIN TESPİT EDİLMESİ

Mehmet Ertuğrul EVRENSEL

Danışman: Dr. Öğr. Üyesi İsmail AKGÜL

TEZ JÜRİ ÜYELERİ

Doç. Dr. Fulya ASLAY

Dr. Öğr. Üyesi İsmail AKGÜL

Dr. Öğr. Üyesi Hatice ARSLANTAŞ DALBOY

YÜKSEK LİSANS TEZİ

ERZİNCAN, 2025

© 2025 [Mehmet Ertuğrul Evrensel]. Tüm hakları saklıdır.

Kabul ve Onay Sayfası

Dr. Öğr. Üyesi İsmail AKGÜL danışmanlığında, Mehmet Ertuğrul EVRENSEL tarafından hazırlanan bu çalışma 26/06/2025 tarihinde aşağıdaki jüri tarafından Yapay Zeka ve Robotik Anabilim Dalı'nda Yüksek Lisans Tezi olarak oybirliği ile kabul edilmiştir.

Başkan : Doç. Dr. Fulya ASLAY İmza:

Üye : Dr. Öğr. Üyesi İsmail AKGÜL İmza:

Üye : Dr. Öğr. Üyesi Hatice ARSLANTAŞ DALBOY İmza:

Yukarıdaki Yüksek Lisans Tezi Enstitü Yönetim Kurulunun / / 20.... tarih ve/..... sayılı kararı ile onaylanmıştır.

Doç. Dr. Kemal Volkan ÖZDOKUR

Enstitü Müdür V.

Not: Bu tezde kullanılan özgün ve başka kaynaklardan yapılan bildirişlerin, şekil ve tabloların kaynak olarak kullanımı, 5846 sayılı Fikir ve Sanat Eserleri Kanunundaki hükümlere tabidir.

Bilimsel Etięe Uygunluk Sayfası

“Yapay zeka tabanlı görüntü işleme teknikleri ile iş güvenliğinde kullanılan kişisel ekipmanların tespit edilmesi” isimli “Yüksek Lisans” tezim tarafımca intihal tespit programı ile incelenmiştir. Buna göre tezimde bilimsel etik ihlali ve intihal olarak nitelendirilebilecek herhangi bir durum olmadığını taahhüt ederim.

Bu çalışmadaki tüm bilgilerin, akademik ve etik kurallara uygun bir biçimde elde edildiğini; aynı zamanda bu kural ve davranışların gerektirdiğı gibi, bu çalışmanın özünde olmayan tüm materyal ve sonuçları tam olarak aktardığımı ve referans gösterdiğimi beyan ederim. 26/06/2025

(İmza)

Mehmet Ertuğrul EVRENSEL

ÖZET

YAPAY ZEKA TABANLI GÖRÜNTÜ İŞLEME TEKNİKLERİ İLE İŞ GÜVENLİĞİNDE KULLANILAN KİŞİSEL EKİPMANLARIN TESPİT EDİLMESİ

Mehmet Ertuğrul EVRENSEL

Yüksek Lisans Tezi, Erzincan Binali Yıldırım Üniversitesi, Fen Bilimleri Enstitüsü,

Yapay Zeka ve Robotik Anabilim Dalı

Danışman: Dr. Öğr. Üyesi İsmail AKGÜL

2025, 91 sayfa

Yapay zeka (YZ), günümüzde pek çok sektörde dönüşüm yaratan en önemli teknolojilerden biri haline gelmiştir. Özellikle bilgisayarla görme alanında yaşanan gelişmeler, gerçek zamanlı veri analizi ile insan hayatını doğrudan etkileyen çözümler sunmaktadır. Bu çalışmada, YZ destekli nesne tespiti teknikleri kullanılarak işçi güvenliğini artırmaya yönelik bir uygulama geliştirilmiştir. Araştırmada, Kaggle platformunda sunulan “Railroad Worker Detection Dataset” adlı veri seti kullanılarak, demiryolu işçilerinin kişisel koruyucu donanım (KKD) kullanıp kullanmadıkları, özellikle baret ve yelek takıp takmadıkları, gerçek zamanlı görüntüler üzerinden tespit edilmeye çalışılmıştır. Bu amaçla derin öğrenmeye dayalı YOLOv5, YOLOv7 ve YOLOv8 algoritmaları kullanılarak bu modellerin performansları karşılaştırılmış ve en uygun çözüm yöntemi belirlenmiştir. Elde edilen bulgulara göre, çalışmada ele alınan nesne tespiti modelleri farklı ortamlarda çalışan işçilerin konumlarını, üzerlerindeki güvenlik ekipmanlarının varlığını yüksek doğrulukla belirleyebilmekte ve böylece iş kazalarının önüne geçilmesine katkı sağlamaktadır. Bu çalışma, yalnızca akademik bir araştırma olmanın ötesinde, kullanılan yöntemlerin endüstride doğrudan uygulanabilirliği sayesinde güvenli çalışma ortamlarının oluşturulmasına destek olacak pratik çözümler sunmaktadır. Ayrıca, derin öğrenme temelli nesne tespiti algoritmalarının iş sağlığı ve güvenliği alanında ne denli etkin bir şekilde kullanılabileceğini de ortaya koymaktadır.

Anahtar Kelimeler: Yapay zeka, nesne tespiti, algoritma, görüntü işleme, iş güvenliği

ABSTRACT

DETECTION OF PERSONAL EQUIPMENT USED IN OCCUPATIONAL SAFETY WITH ARTIFICIAL INTELLIGENCE BASED IMAGE PROCESSING TECHNIQUES

Mehmet Ertuğrul EVRENSEL

**Master's Thesis, Erzincan Binali Yıldırım University, Institute of Science and
Technology,
Department of Artificial Intelligence and Robotics**

Advisor: Asst. Prof. İsmail AKGÜL

2025, 91 pages

Artificial intelligence (AI), has emerged as one of the most transformative technologies across various industries today. In particular, advances in computer vision have enabled real-time data analysis, offering solutions that directly impact human life. In this study, an application was developed to enhance worker safety using AI-powered object detection techniques. The research utilizes the "Railroad Worker Detection Dataset" available on the Kaggle platform to determine whether railway workers are wearing personal protective equipment (PPE)—specifically helmets and safety vests—based on real-time imagery. To this end, deep learning-based object detection algorithms YOLOv5, YOLOv7, and YOLOv8 were employed, and their performance was compared to identify the most effective solution. The findings indicate that the selected object detection models can accurately identify workers in various environments and detect the presence of safety equipment with high precision, thereby contributing to the prevention of occupational accidents. This study not only constitutes an academic research effort but also offers practical solutions that can be directly implemented in industrial settings to support the creation of safer working environments. Furthermore, it demonstrates the potential of deep learning-based object detection algorithms to be effectively utilized in the field of occupational health and safety.

Keywords: Artificial intelligence, object detection, algorithm, image processing, work safety

TEŞEKKÜR

Yüksek lisans eğitimim süresince bilgi birikimi, akademik rehberliği ve kıymetli katkılarıyla yolumu aydınlatan, her aşamada sabrını ve desteğini esirgemeyen değerli danışmanım Dr. Öğr. Üyesi İsmail AKGÜL'e en içten teşekkürlerimi sunarım. Kendisinden aldığım akademik ve insani değerler, bu çalışmanın ötesinde bana hayat boyu ışık tutacaktır.

Bu süreçte yanımda olduklarını her daim hissettiren, sevgileri ve dualarıyla bana güç veren aileme minnettarım. Zor zamanlarda moral kaynağım olan, desteğini esirgemeyen kıymetli arkadaşlarıma da ayrıca teşekkür ederim.

Hayatımın en özel yol arkadaşlarından biri olan nişanlıma, bu zorlu süreci birlikte aşarken gösterdiği anlayış, sabır ve sevgi için gönülden teşekkür ederim. Varlığı, bu sürecin daha anlamlı ve güçlü geçmesini sağlamıştır.

Bu tez çalışmasında emeği geçen, katkı sunan tüm kişi ve kurumlara şükranlarımı sunarım.

Mehmet Ertuğrul EVRENSEL

Haziran, 2025

İÇİNDEKİLER

ÖZET	i
ABSTRACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER	iv
TABLolar DİZİNİ	vii
ŞEKİLLER DİZİNİ	viii
SİMGELER VE KISALTMALAR DİZİNİ	x
1. GİRİŞ	1
1.1. Araştırmanın Amacı	2
1.2. Araştırmanın Önemi	2
1.3. Varsayımlar	3
1.4. Sınırlılıklar	3
2. LİTERATÜR ÖZETİ	4
3. KAVRAMSAL ÇERÇEVE	16
3.1. Yapay Zeka	16
3.2. Makine Öğrenmesi ve Derin Öğrenme	16
3.3. Bilgisayarla Görme	17
3.4. Nesne Tespiti	18
3.4.1. Nesne tespiti türleri	19
3.4.1.1. Nesne tespiti süreci	20
3.4.2. Nesne tespiti algoritma çeşitleri	20
3.4.2.1. İki aşamalı nesne tespiti	20
3.4.2.2. Tek aşamalı nesne tespiti	23
3.4.3. Isı haritası tabanlı nesne tespiti	24
3.5. İş Güvenliği ve Görsel Denetim Sistemleri	24
4. NESNE TESPİTİNDE KULLANILAN YÖNTEMLER	26
4.1. CNN Modelinin Gelişimi	26
4.2. CNN Temelli İki Aşamalı Tespit Modelleri	28
4.2.1. Regions with CNN (R-CNN)	28
4.2.2. Spatial pyramid pooling network (SPPNet)	29
4.2.3. Fast R-CNN	30

4.2.4. Faster R-CNN	31
4.2.5. Feature pyramid networks (FPN)	32
4.3. CNN Temelli Tek Aşamalı Tespit Modelleri	33
4.3.1. You only look once (YOLO).....	33
4.3.1.1. YOLOv1 - YOLOv3.....	34
4.3.1.2. YOLOv4	36
4.3.1.3. YOLOv5	38
4.3.1.3.1. YOLOv5'in mimari yapısı.....	38
4.3.1.3.2. YOLOv5 sürümleri ve uygulama esnekliği.....	38
4.3.1.3.3. YOLOv5'in eğitim süreci ve teknik özellikleri	39
4.3.1.3.4. YOLOv5'in gerçek zamanlı uygulamaları	39
4.3.1.3.5. YOLOv5'in avantajları ve sınırlılıkları	40
4.3.1.4. YOLOv6	40
4.3.1.5. YOLOv7	42
4.3.1.5.1. YOLOv7'nin yenilikçi özellikleri	43
4.3.1.5.2. YOLOv7 mimarisinin teknik ayrıntıları	43
4.3.1.5.3. YOLOv7'nin performans karşılaştırmaları	44
4.3.1.5.4. YOLOv7'nin gerçek dünya uygulamaları	44
4.3.1.5.5. YOLOv7'nin avantajları ve sınırlılıkları	44
4.3.1.6. YOLOv8	45
4.3.1.6.1. YOLOv8'in mimari yapısı.....	45
4.3.1.6.2. YOLOv8'in özellikleri ve avantajları.....	46
4.3.1.6.3. YOLOv8'in uygulama alanları	46
4.3.1.6.4. YOLOv8'in dezavantajları ve sınırlılıkları.....	46
4.3.1.6.5. YOLOv8'in avantajları.....	47
4.3.2. Single shot multiBox detector (SSD)	47
4.3.3. RetinaNet.....	47
4.3.4. Güncel tek aşamalı modeller ve gelişmeler	48
4.3.4.1. EfficientDet	48
4.3.4.2 PP-YOLOE.....	48
4.4. Derin Öğrenme Algoritmalarında Tespit İşleminin Hızlandırılması.....	48
4.4.1. Hafif mimari kullanımı	49
4.4.2. Model sıkıştırma ve pruning.....	49
4.4.3. Kuantizasyon ve düşük hassasiyetli hesaplama.....	49

4.4.4. Donanım tabanlı hızlandırma	49
4.4.5. Mimari iyileştirmeler	50
5. MATERYAL VE YÖNTEM.....	51
5.1. Veri Seti ve Önışleme	51
5.1.1 Veri setinin tanıtımı ve genel yapısı	51
5.1.2 Etiketleme yapısı ve kalitesi	53
5.1.3 Önışleme süreci ve teknikleri	53
5.1.3.1 Boyutlandırma (Resizing).....	53
5.1.3.2 Veri artırımı (Data augmentation)	53
5.1.3.3 Etiket uyumlaştırması ve normalize etme	54
5.1.4 Veri setinin eğitim-doğrulama-test ayrımı	54
5.2. Yöntem	54
5.2.1 Ölçüm metrikleri.....	56
5.2.1.1. Intersection over union (IoU)	57
5.2.1.2. Kesinlik (Precision)	57
5.2.1.3. Duyarlılık (Recall).....	57
5.2.1.4. F1-score	58
5.2.1.5. Mean average precision (mAP)	58
5.2.2. Çalışmada kullanılan YOLO sürümlerinin karşılaştırılması	58
6. BULGULAR	61
6.1. YOLOv5 Modeline Ait Bulgular.....	61
6.2. YOLOv7 Modeline Ait Bulgular.....	67
6.3. YOLOv8 Modeline Ait Bulgular.....	72
6.4. Karşılaştırmalı Performans Analizi	78
7. TARTIŞMA.....	81
8. SONUÇ VE ÖNERİLER.....	83
8.1. Sonuçlar	83
8.2. Öneriler.....	83
KAYNAKÇA	85

TABLÖLAR DİZİNİ

Tablo 1. Etiketli nesne sayısı	52
Tablo 2. YOLO sürümleri ağırlık dosyaları	55
Tablo 3. Eğitim sırasında kullanılan parametreler.....	56
Tablo 4. YOLOv5, YOLOv7 ve YOLOv8 sürümlerinin karşılaştırması	59
Tablo 5. YOLOv5 üzerinde yapılan deneyler	61
Tablo 6. YOLOv7 üzerinde yapılan deneyler	67
Tablo 7. YOLOv8 üzerinde yapılan deneyler	73
Tablo 8. YOLO sürümlerinin karşılaştırmalı performans analizi.....	79



ŞEKİLLER DİZİNİ

Şekil 1. YOLO-EA ağ yapısı diyagramı.....	4
Şekil 2. Derin öğrenme kullanarak işçi güvenliği baret algılama için genel mimari	5
Şekil 3. Sınıflandırmaları gösteren video çıktı örnek kareler.....	6
Şekil 4. Nesne tespit metodolojisi	7
Şekil 5. YOLO modellerine dayalı KKD tespiti için önerilen çerçeve	9
Şekil 6. CNN mimarisi	10
Şekil 7. Değiştirilmiş YOLOv5 yapısı	11
Şekil 8. CA-Yolov5-6.2 yapı şeması	12
Şekil 9. YOLOv5 hedef tespit ağının yapısı.....	14
Şekil 10. Geliştirilmiş YOLOv8 tespit modelinin genel mimarisi	15
Şekil 11. Bilgisayarla görme ve algoritmalar	17
Şekil 12. İki aşamalı dedektörlerin temel mimarisi	19
Şekil 13. R-CNN.....	21
Şekil 14. Fast R-CNN	21
Şekil 15. Faster R-CNN.....	22
Şekil 16. SPPNet.....	23
Şekil 17. Isı haritası tabanlı nesne tespiti	24
Şekil 18. CNN temel yapı.....	26
Şekil 19. ReLU katmanları	27
Şekil 20. CNN temelli iki aşamalı tespit modeli	28
Şekil 21. R-CNN.....	29
Şekil 22. SPPNet'te piramit ile üç seviyeli mekansal piramit havuzlaması	30
Şekil 23. Fast R-CNN mimarisi.....	31
Şekil 24. Faster R-CNN mimarisi.....	32
Şekil 25. Feature Pyramid Networks	33
Şekil 26. YOLOv1	34
Şekil 27. YOLOv2 mimarisi.....	35
Şekil 28. YOLOv4 mimarisi.....	37
Şekil 29. YOLOv6 mimarisi.....	41
Şekil 30. Veri seti örnek görseller	52
Şekil 31. YOLOv5 eğitim sonuçları	62

Şekil 32. YOLOv5 karmaşıklık matrisi	63
Şekil 33. YOLOv5 F1-Güven eğrisi	64
Şekil 34. YOLOv5 precision-recall eğrisi	65
Şekil 35. YOLOv5 modelinin doğrulama verisi üzerindeki tahminlerine ait örnek çıktılar	66
Şekil 36. YOLOv7 eğitim sonuçları	67
Şekil 37. YOLOv7 karmaşıklık matrisi	69
Şekil 38. YOLOv7 F1-Güven eğrisi	70
Şekil 39. YOLOv7 precision-recall eğrisi	71
Şekil 40. YOLOv7 modelinin doğrulama verisi üzerindeki tahminlerine ait örnek çıktılar	72
Şekil 41. YOLOv8 eğitim sonuçları	73
Şekil 42. YOLOv8 karmaşıklık matrisi	75
Şekil 43. YOLOv8 F1-Güven eğrisi	76
Şekil 44. YOLOv8 precision-recall eğrisi	77
Şekil 45. YOLOv8 modelinin doğrulama verisi üzerindeki tahminlerine ait örnek çıktılar	78

SİMGELER VE KISALTMALAR DİZİNİ

AI	Artificial Intelligence
CNN	Convolutional Neural Network
DÖ	Derin Öğrenmesi
FN	False Negative
FP	False Positive
FPN	Feature Pyramid Networks
FPS	Frame Per Second
GPU	Graphics Processing Unit
IoU	Intersection over Union
İSG	İş Sağlığı ve Güvenliği
KKD	Kişisel Koruyucu Donanım
mAP	Mean Average Precision
MÖ	Makine Öğrenme
P	Precision
PPE	Personal Protective Equipment
R	Recall
ReLU	Rectified Linear Unit
SSD	Single Shot Multibox Detector
TN	True Negative
TP	True Positive
YOLO	You Only Look Once
YZ	Yapay Zeka

1. GİRİŞ

21. yüzyılın en dikkat çeken teknolojik gelişmelerinden biri olan yapay zeka, farklı disiplinlerde dönüşümsel etkiler yaratmaya devam etmektedir. Özellikle sağlık, finans, eğitim, savunma ve sanayi gibi alanlarda kullanılan yapay zeka uygulamaları, süreçleri hızlandırmakta, verimliliği artırmakta ve insan kaynaklı hataları azaltmaktadır. Bu bağlamda, yapay zekanın alt disiplinlerinden biri olan bilgisayarla görme, dijital görüntülerin analiz edilmesi, anlamlandırılması ve yorumlanması süreçlerinde büyük bir potansiyel sunmaktadır.

Bilgisayarla görme, makinelerin görsel veriyi tıpkı insanlar gibi algılamasını ve bu veriden anlam çıkarmasını sağlayan bir teknolojidir. Bu teknoloji sayesinde makineler görüntülerdeki nesneleri, yüzleri, hareketleri, renkleri ve desenleri tanımlayabilir. Bilgisayarla görme teknolojilerinin en önemli uygulama alanlarından biri ise nesne tespiti olarak öne çıkmaktadır. Nesne tespiti, bir görüntü ya da video içerisinde yer alan nesnelerin türünü belirleme ve konumlarını saptama sürecidir. Bu süreç hem sınıflandırma hem de konumlandırma işlemlerini içerdiğinden, bilgisayarla görme uygulamalarının en karmaşık ve hesaplama açısından en yoğun görevlerinden biridir.

Günümüzde nesne tespiti; yüz tanıma sistemlerinden otonom araçlara, medikal görüntü analizinden endüstriyel üretim hatlarına kadar geniş bir yelpazede kullanılmaktadır. Aynı zamanda güvenlik kameralarıyla yapılan izleme sistemlerinde, alışveriş davranışlarını analiz eden perakende uygulamalarında ve tarımsal verimliliği ölçen akıllı tarım sistemlerinde de etkin bir rol oynamaktadır. Son yıllarda ise iş sağlığı ve güvenliği (İSG) kapsamında, özellikle riskli çalışma ortamlarında işçilerin uygun ekipman kullanıp kullanmadığını tespit etmeye yönelik sistemler geliştirilmeye başlanmıştır. Bu sistemlerde amaç, görsel veriler üzerinden işçilerin kişisel koruyucu donanım (KKD) kullanım durumlarını otomatik olarak algılayarak olası kazaların önüne geçmek ve güvenli çalışma alanları oluşturmaktır.

Bu çalışmada görüntü işleme ve derin öğrenme teknikleri kullanılarak gerçek zamanlı bir işçi güvenliği izleme sistemi geliştirilmesi amaçlanmaktadır. Bu kapsamda, Kaggle platformunda sunulan “Railroad Worker Detection Dataset” veri seti kullanılarak, demiryolu işçileri üzerinde baret ve yelek tespiti gerçekleştirilmiştir. Demiryolları gibi zorlu ve tehlikeli çalışma alanlarında, çalışanların koruyucu ekipman kullanıp kullanmadığının tespiti hem iş güvenliği hem de yasal zorunluluklar açısından hayati bir rol oynamaktadır. Bu doğrultuda, çalışmada

derin öğrenme tabanlı üç farklı tek aşamalı nesne tespiti algoritması olan YOLOv5, YOLOv7 ve YOLOv8 modelleri kullanılarak hem doğruluk hem de performans açısından karşılaştırmalı analizler yapılmıştır.

Çalışmada kullanılan You Only Look Once (YOLO) algoritmaları, yüksek hız ve doğruluk kombinasyonu sunmaları nedeniyle gerçek zamanlı nesne tespiti uygulamalarında öne çıkmaktadır. Bu modeller, görüntü üzerinde birden fazla nesneyi eşzamanlı olarak tespit edebilmekte ve pratik uygulamalar için oldukça elverişli çözümler sunmaktadır. Yapılan deneyler sonucunda elde edilen başarı oranları ve performans ölçütleri, bu teknolojinin iş güvenliği sistemlerine nasıl entegre edilebileceğine dair önemli bulgular ortaya koymuştur.

Bu çalışmanın bilimsel katkısı yalnızca yüksek doğrulukla çalışan bir sistem geliştirmekle sınırlı kalmamaktadır. Aynı zamanda, farklı derin öğrenme modellerinin güvenlik alanındaki uygulanabilirliğini karşılaştırmalı olarak analiz ederek literatüre katkı sunmakta ve ilerleyen çalışmalara temel oluşturmaktadır.

1.1. Araştırmanın Amacı

Bu araştırmanın temel amacı, görüntü işleme ve derin öğrenme tabanlı nesne tespiti yöntemlerini kullanarak işçi güvenliğini artırmaya yönelik bir sistem geliştirmektir. Özellikle, riskli çalışma alanlarından biri olan demiryolu sahalarında görev yapan işçilerin KKD kullanıp kullanmadıklarının tespit edilmesi hedeflenmiştir. Bu kapsamda, açık kaynaklı bir veri seti kullanılarak işçilerin baret ve yelek takma durumlarının analiz edilmesi ve bu amaçla YOLOv5, YOLOv7 ve YOLOv8 nesne tespiti algoritmalarının performanslarının karşılaştırılması amaçlanmaktadır. Elde edilecek bulgular ile gerçek zamanlı izleme sistemlerinin iş sağlığı ve güvenliği alanına katkı sağlayacak biçimde uygulanabilirliği değerlendirilecektir.

1.2. Araştırmanın Önemi

İSG, hem yasal zorunluluklar hem de etik sorumluluklar açısından tüm işverenler ve toplumlar için kritik bir konudur. Günümüzde hala birçok sektörde, KKD kullanımı ihmali nedeniyle ciddi iş kazaları yaşanmakta ve bu kazalar can kayıplarına veya kalıcı sakatlanmalara neden olabilmektedir. Bu çalışmanın önemi, yapay zeka destekli nesne tespiti yöntemlerinin, iş güvenliği uygulamalarına entegrasyonu ile kazaların önlenmesinde proaktif bir çözüm sunma

potansiyelidir. Ayrıca, çalışmada farklı nesne tespiti modellerinin etkinliği karşılaştırılarak, sektörel uygulamalarda tercih edilebilecek en uygun modelin belirlenmesine katkı sağlanacaktır. Bu yönüyle çalışma, yalnızca akademik bilgi üretimine değil, aynı zamanda uygulamalı endüstriyel çözümlerin geliştirilmesine de hizmet etmektedir.

1.3. Varsayımlar

Bu araştırma kapsamında aşağıdaki varsayımlar yapılmıştır:

- Kullanılan “Railroad Worker Detection Dataset” veri seti, işçilerin baret ve yelek kullanımı açısından yeterli çeşitliliğe ve temsil gücüne sahiptir.
- Kullanılan etiketleme verileri doğru ve tutarlıdır.
- Eğitim ve test aşamalarında kullanılan YOLO algoritmaları, doğru şekilde yapılandırılmış ve aynı koşullar altında eğitilmiştir.
- Model performanslarını değerlendirmede kullanılan metrikler (precision, recall, mAP, FPS vb.) sonuçları anlamlı şekilde yansıtır.
- Gerçek zamanlı sistemlerde, elde edilen sonuçlar üretim ortamlarına uygun şekilde entegre edilebilir.

1.4. Sınırlılıklar

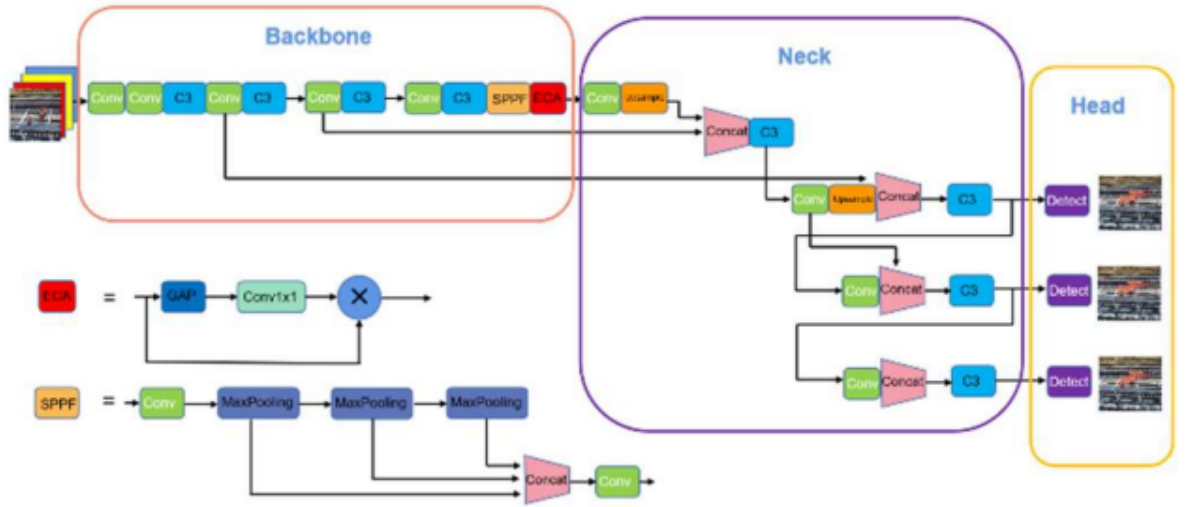
Araştırmanın belirli sınırlamaları bulunmaktadır. Bu sınırlamalar şunlardır:

- Kullanılan veri seti yalnızca belirli bir sektöre (demiryolu işçileri) odaklıdır ve diğer sektörlerdeki işçi görüntülerini temsil etmeyebilir.
- Veri seti, farklı hava koşulları, ışık düzeyleri ve kamera açılarında çekilmemiş olabilir; bu durum modelin genelleme kabiliyetini sınırlayabilir.
- Bu çalışmada sadece tek aşamalı nesne tespiti algoritmaları olan YOLOv5, YOLOv7 ve YOLOv8 incelenmiş, iki aşamalı yöntemler kapsam dışı bırakılmıştır.
- Eğitim sürecinde kullanılan donanım sınırlamaları nedeniyle modellerin eğitim süresi ve ölçeklenebilirliği kısıtlanmıştır.
- Geliştirilen sistemin gerçek dünyada uygulanabilirliği test ortamı dışında, saha koşullarında uygulanmamış ve doğrulanmamıştır.

2. LİTERATÜR ÖZETİ

Bu bölümde literatürde işçi güvenliğinde kullanılan ekipmanların tespiti üzerine yapay zeka tekniklerinin kullanıldığı güncel çalışmalara yer verilmiştir.

Liu ve Qin (2024) tarafından yapılan “Target Detection of Safety Protective Gear Using the Improved YOLOv5” adlı çalışmada, demiryolu inşaat sahalarında KKD’lerin tespiti için geliştirilen YOLO-EA adlı yeni bir derin öğrenme modeli tanıtılmaktadır. Bu model, özellikle küçük ve sıklıkla gizlenmiş nesnelerin tespiti gibi zorlukların üstesinden gelmeyi amaçlamaktadır. Demiryolu inşaat sahaları gibi yüksek riskli ortamlarda, işçilerin güvenliğini sağlamak için KKD'lerin doğru kullanımı hayati öneme sahiptir. Ancak, bu ekipmanların otomatik olarak tespiti, küçük boyutları ve diğer nesneler tarafından sıklıkla engellenmeleri nedeniyle zordur. Geleneksel yöntemler, bu tür zorluklarla başa çıkmakta yetersiz kalmaktadır. YOLO-EA modeli, YOLOv5 mimarisi üzerine inşa edilmiştir ve iki ana teknik yenilik olan Efficient Channel Attention (ECA) Mekanizması ve Enhanced Intersection over Union (EIoU) Kayıp Fonksiyonu özelliklerini içermektedir.



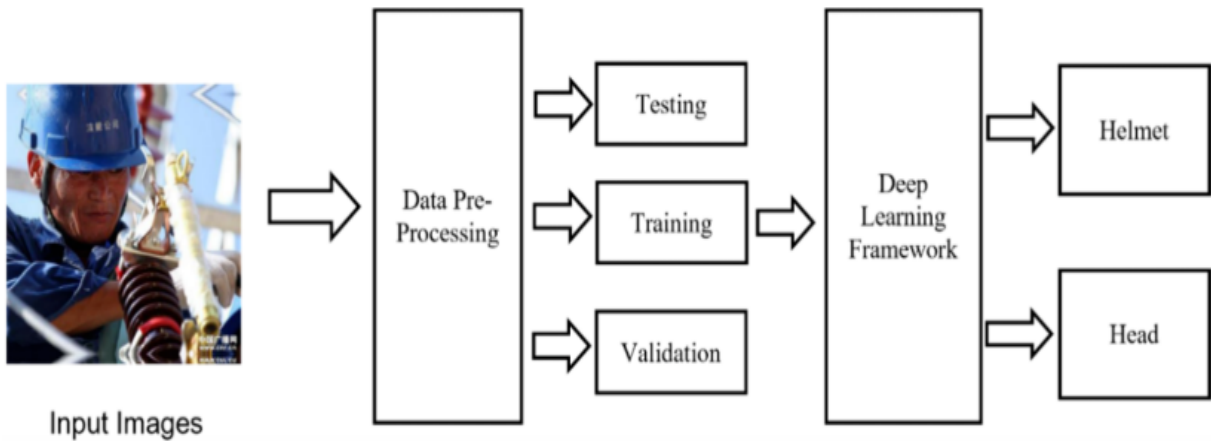
Şekil 1. YOLO-EA ağ yapısı diyagramı (Liu ve Qin, 2024)

Şekil 1’de gösterilen modelin yapısı, giriş katmanı, omurga (Conv, C3, SPPF modülleri), boyun ağı ve çıkış katmanından oluşmaktadır. SPPF modülü, görüntülerden etkili özellikler çıkarmak için küçük konvolüsyon çekirdekleri kullanırken, ECA modülü, önemli kanal etkileşimlerine odaklanarak küçük nesnelerin tespitini iyileştirir. YOLO-EA modeli, gerçek demiryolu inşaat sahalarından elde edilen gözetim görüntülerinden oluşan bir veri kümesi üzerinde test

edilmiştir. Bu veri kümesi, farklı aydınlatma koşulları, gizlenme durumları ve ölçek varyasyonları içeren çeşitli senaryoları kapsamaktadır. Deneysel sonuçlar, YOLO-EA'nın YOLOv5'e kıyasla üstün performans sergilediğini göstermektedir:

- Doğruluk: %98.9
- Duyarlılık: %94.7
- Gerçek Zamanlı Performans: 70.774 FPS

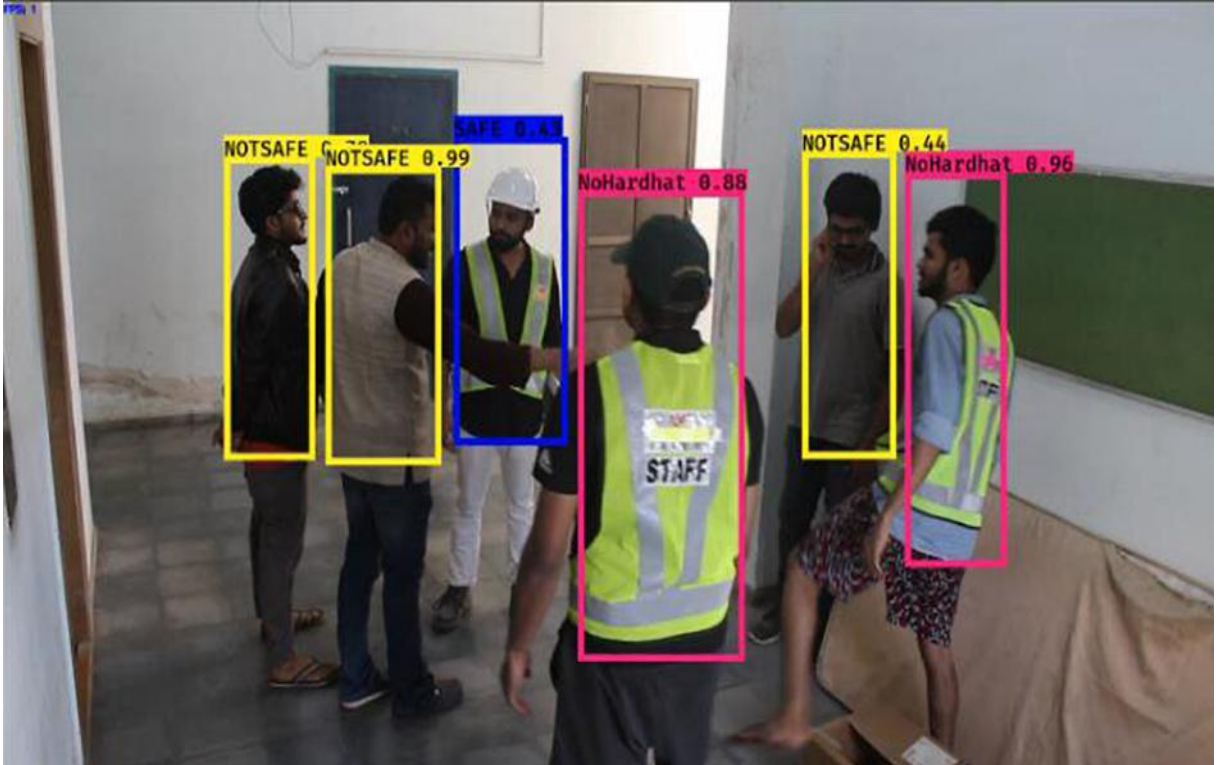
Hayat ve Morgado-Dias (2022) tarafından yapılan “Deep Learning-Based Automatic Safety Helmet Detection System for Construction Safety” adlı çalışmada, inşaat alanlarında işçi güvenliğini artırmak amacıyla geliştirilen derin öğrenme tabanlı bir baret tespit sistemi sunulmuştur (Şekil 2). Özellikle düşük ışık koşulları ve küçük nesne boyutları gibi gerçek zamanlı uygulamalarda sorun teşkil eden durumlara odaklanılmıştır. Bu çalışmada, YOLO algoritmasının farklı versiyonları (YOLOv3, YOLOv4 ve YOLOv5x) karşılaştırılmış; en yüksek başarı, YOLOv5x ile elde edilmiştir. Çalışmada, 5000 etiketli görüntü içeren Hard Hat Workers veri kümesi kullanılmıştır. Bu görüntüler, baret takılı (helmet) ve baret takılmamış (head) olmak üzere iki sınıfa ayrılmıştır. Görsellerin %60'ı eğitim, %20'si doğrulama, %20'si ise test amacıyla kullanılmıştır. YOLOv5x mimarisi, CSPDarknet53 omurgası ve Path Aggregation Network (PANet) boyun yapısı sayesinde yüksek doğrulukla küçük nesneleri tespit edebilmektedir. Modelin eğitiminde görüntü iyileştirme ve yeniden boyutlandırma gibi ön işleme adımları uygulanmıştır.



Şekil 2. Derin öğrenme kullanarak işçi güvenliği baret algılama için genel mimari (Hayat ve Morgado-Dias, 2022)

Modelin deęerlendirilmesi iin doęruluk, keskinlik, duyarlılık, F1 skoru ve mAP metrikleri kullanılmıřtır. YOLOv5x modeli %92.44 mAP, %92 doęruluk ve %89 duyarlılık ile dięer versiyonları geride bırakmıřtır. Ayrıca modelin dūřuk ıřıkta, kūuk nesnelerde ve arka pozisyonlu iřilerde dahi bařarılı tahminlerde bulunduęu gōsterilmiřtir.

Delhi vd. (2020) tarafından yapılan “Detection of Personal Protective Equipment (PPE) Compliance on Construction Site Using Computer Vision Based Deep Learning Techniques” adlı alıřmada, inřaat sahalarında KKD’lerin doęru kullanımının derin ōęrenme destekli bilgisayarlı gōrme teknikleriyle otomatik olarak tespit edilmesi hedeflenmiřtir. alıřma, KKD uyumunun manuel kontrolūnūn sınırlılıklarına dikkat ekerek, gerek zamanlı ve otomatik bir denetim sisteminin geliřtirilmesinin ōnemini vurgulamaktadır. alıřmada, YOLOv3 tabanlı bir derin ōęrenme modeli kullanılarak, iřilerin baret ve yelek gibi KKD’leri takıp takmadıklarının algılanması amalanmıřtır. Model, dōrt sınıfta tahmin yapmaktadır: "SAFE", "NOTSAFE", "NoHardhat" ve "NoJacket" (řekil 3). Transfer ōęrenme yaklařımı kullanılarak ōnceden eęitilmiř YOLOv3 aęı, inřaat gūvenlięi senaryosuna ōzel olarak yeniden eęitilmiřtir.

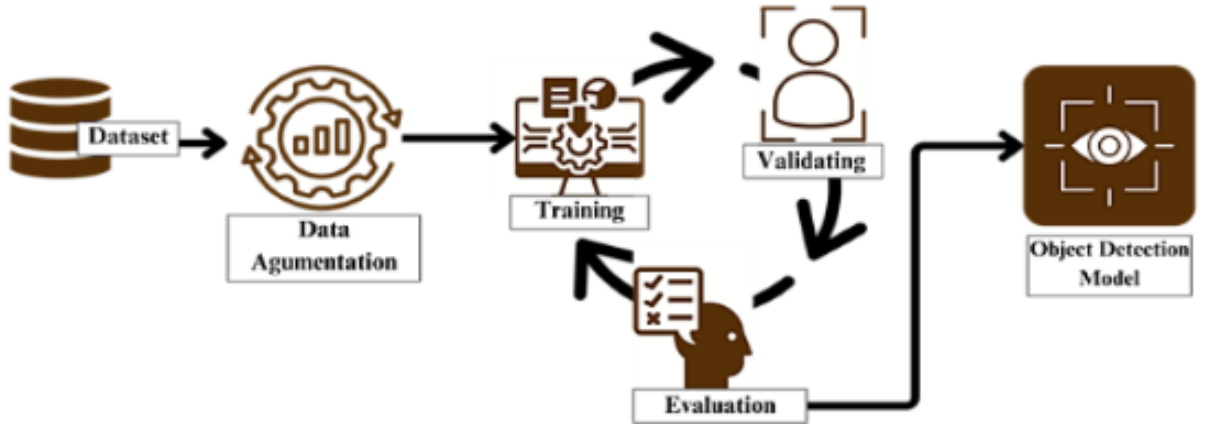


řekil 3. Sınıflandırmaları gōsteren video ıktı ōrnek kareler (Delhi vd., 2020)

Model, üç aşamalı bir eğitim sürecinden geçirilmiş ve her aşamada katmanlar kademeli olarak serbest bırakılarak optimize edilmiştir. Eğitim sonucunda modelin kayıp (loss) değeri 12.06'ya düşmüştür. Kullanılan YOLOv3 modeli, sınıflandırma başarımını değerlendirmek için doğruluk, hassasiyet, duyarlılık ve F1 skor gibi metriklerle test edilmiştir.

- Doğruluk: %96.27 (doğrulama setinde), %96.51 (test setinde)
- Ortalama F1 Skoru: 0.96 (doğrulama), 0.97 (test)
- Yeni (görülmemiş) veri seti ile doğruluk: %96.92

Ahmed vd. (2023) tarafından yapılan “Personal Protective Equipment Detection: A Deep-Learning-Based Sustainable Approach” adlı çalışmada, inşaat ve tehlikeli endüstriyel sahalarda işçilerin KKD kullanımını gerçek zamanlı ve otomatik olarak algılayan derin öğrenme tabanlı bir sistem geliştirilmiştir (Şekil 4). Çalışma, manuel denetimin yetersizliklerine çözüm olarak bilgisayarla görme ve yapay zeka yöntemlerini bir araya getirerek sürdürülebilir bir güvenlik yaklaşımı önermektedir. Araştırma, KKD eksikliğinin neden olduğu ölümlü ve yaralanmalı kazaların önüne geçmeyi hedeflemektedir. Özellikle Suudi Arabistan gibi endüstriyel yoğunluğa sahip ülkelerde, KKD uyumunun denetimi büyük önem arz etmektedir. Bu bağlamda, KKD takmama nedenleri (yetersiz eğitim, düşük risk algısı, denetim eksikliği gibi) literatür taramasıyla incelenmiş ve bu duruma yönelik teknolojik çözümler sunulmuştur.

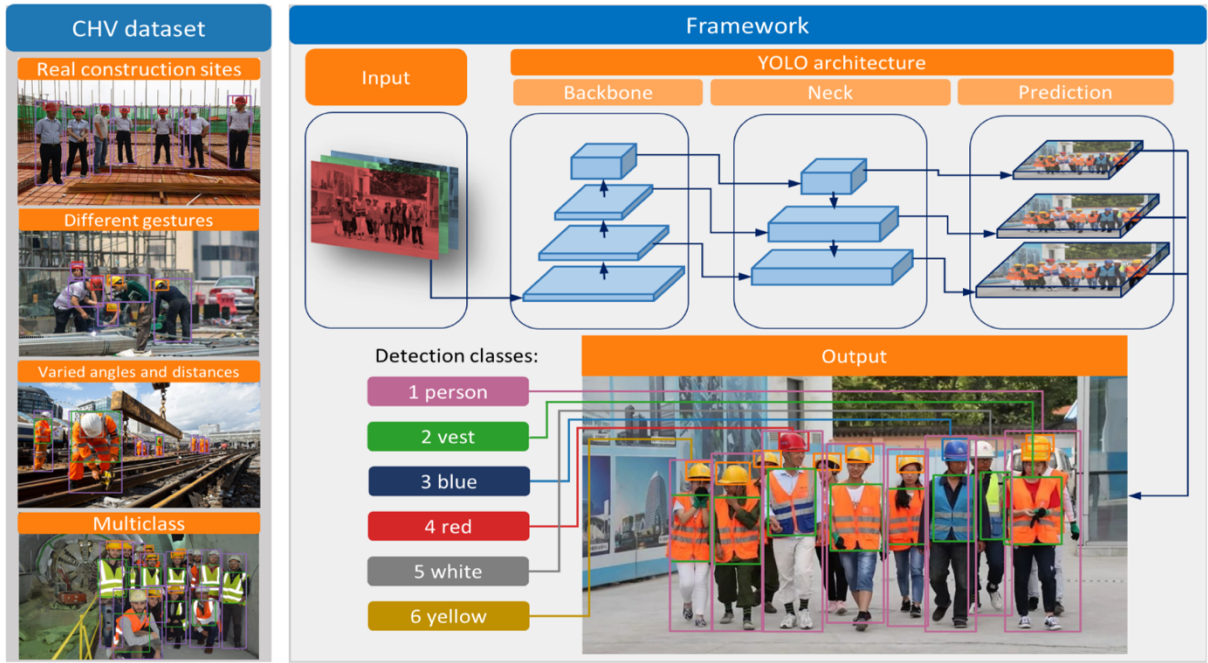


Şekil 4. Nesne tespit metodolojisi (Ahmet vd., 2023)

Araştırmada CHVG adlı bir veri kümesi kullanılmıştır. Veri kümesinde 1189 etiketli görüntü yer almakta ve sekiz sınıf (kırmızı, mavi, beyaz, sarı baretler, kişi, baş, yelek, gözlük) bulunmaktadır. Görüntüler eğitim (%62), doğrulama (%23) ve test (%15) setlerine ayrılmıştır. Görüntü artırımı (flipping, scaling, HSV değişimi vb.) ile modelin genelleme kabiliyeti artırılmıştır.

Modelleme sürecinde karşılaştırılan iki farklı derin öğrenme mimarisinden YOLOv5, gerçek zamanlı algılama avantajı sunmasına rağmen bu çalışmada daha düşük doğruluk (mAP50: %63.9) elde ederken, ResNet50 tabanlı Faster R-CNN daha yüksek doğruluk sağlamış (mAP50: %96) ve özellikle kişi, baş, yelek ile gözlük gibi sınıflarda ortalama %70'in üzerinde doğruluk ve %80'e yakın geri çağırma başarısı göstermiştir. Faster R-CNN modeli, test verisi üzerinde %96 mAP50 başarımı elde etmiştir. Modelin ortalama işlem süresi sadece 0.17 saniyedir ve bu hız, benzer çalışmalarla karşılaştırıldığında kayda değer bir iyileşme olarak sunulmaktadır. Bu modelin performansı YOLOX-m gibi literatürdeki güçlü rakip modellerin bile önüne geçmiştir. Bunun temel nedeni, ağ bileşenlerinin dikkatli seçimi ve katmanların optimize edilmesidir.

Wang vd. (2021) tarafından yapılan "Fast Personal Protective Equipment Detection for Real Construction Sites Using Deep Learning Approaches" adlı çalışmada, gerçek inşaat sahalarında KKD'lerin tespiti için hızlı ve güvenilir bir sistem geliştirilmesi amaçlanmıştır. Yazarlar, özellikle baretin farklı renkleri (mavi, kırmızı, beyaz, sarı), yelek ve kişinin kendisinin otomatik olarak tanımlanmasını hedeflemiş ve bu amaçla YOLO algoritmasının çeşitli versiyonlarını kapsamlı bir biçimde incelemişlerdir (Şekil 5). Çalışmada öncelikle, mevcut KKD tespit yöntemlerinin genellikle sınırlı sayıda ekipmanı algılayabildiği ve gerçek inşaat ortamlarında yeterince başarılı sonuçlar veremediği ifade edilmiştir. Bu sorunu çözmek adına araştırmacılar, Color Helmet and Vest (CHV) adını verdikleri yeni ve yüksek kaliteli bir veri seti oluşturmuşlardır. Bu veri seti, 1330 görüntüden oluşmakta ve gerçek inşaat sahalarındaki çeşitli açılar, pozlar ve mesafelerden çekilmiş altı farklı sınıfı kapsamaktadır: kişi, yelek, mavi baret, kırmızı baret, beyaz baret ve sarı baret. Çalışmada üç farklı YOLO sürümü (v3, v4 ve v5) üzerinde deneyler yürütülmüştür. YOLOv3'te üç ve beş katmanlı yapıların küçük nesne tespiti üzerindeki etkisi test edilmiş, YOLOv4 için farklı eğitim görüntüsü boyutlarının doğruluk üzerindeki etkisi incelenmiştir. En kapsamlı değerlendirme ise YOLOv5 ailesi üzerinde yapılmış, bu kapsamda dört farklı model büyüklüğü (v5s, v5m, v5l, v5x) test edilmiştir. Sonuçlara göre, en yüksek mAP %86.55 ile YOLOv5x modelinde elde edilmiştir.

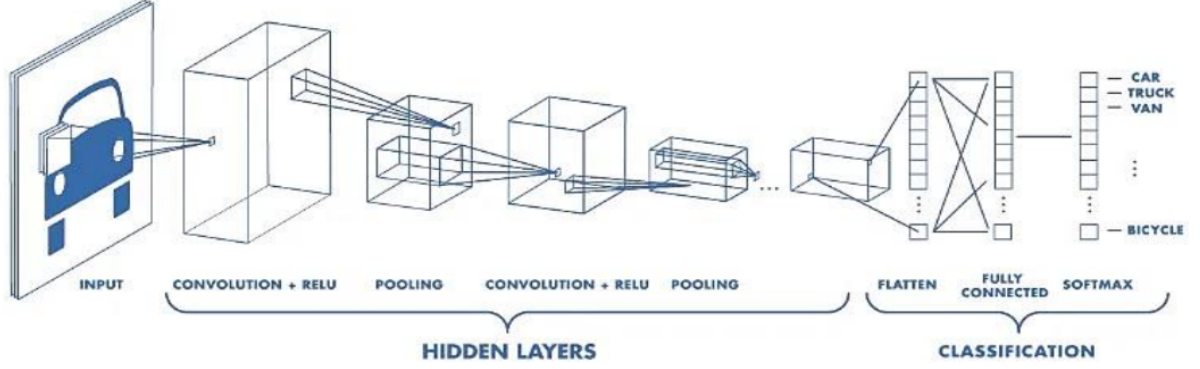


Şekil 5. YOLO modellerine dayalı KKD tespiti için önerilen çerçeve (Wang vd., 2021)

Bununla birlikte, en hızlı çalışan model ise 52 FPS işlem hızı ile YOLOv5s olmuştur. Bu da sistemin gerçek zamanlı uygulamalar için uygun olduğunu göstermektedir.

Cheng vd. (2023) tarafından yapılan “Real-Time Detection of Personal Protective Equipment in Construction Sites Using YOLOv5 and Deep Learning” isimli çalışma, inşaat alanlarında iş güvenliğini sağlamak amacıyla KKD’lerin otomatik olarak tespit edilmesine yönelik gerçek zamanlı bir sistem geliştirmeyi amaçlamaktadır. Özellikle manuel denetim süreçlerinin yetersiz kaldığı veya yüksek riskli alanlarda insan faktörüne bağımlı olmayan bir denetim mekanizması sunulması hedeflenmiştir. Yazarlar bu kapsamda, derin öğrenme mimarisi olan YOLOv5 algoritmasını kullanarak; baret, güvenlik yeleği ve işçi sınıflarını içeren bir KKD tespit modeli oluşturmuşlardır (Şekil 6). Çalışmada, inşaat sahası koşullarını yansıtan 5000’den fazla görüntüden oluşan özel bir veri seti etiketlenmiş ve modelin eğitim sürecinde kullanılmıştır. Modelin doğruluğu, precision, recall, F1-score ve mAP gibi performans ölçütleri ile detaylı biçimde değerlendirilmiştir. Sonuçlara göre, YOLOv5 modeli yüksek doğruluk (%92’nin üzerinde mAP@0.5) ve düşük hata oranı ile baret ve yelek tespitinde başarılı bir performans göstermiştir. Bunun yanı sıra, farklı kamera açıları, ışık değişimleri ve arka plan gürültüsü gibi çevresel değişkenlere karşı sistemin dayanıklılığı da test edilmiştir. Çalışmanın bir diğer katkısı, modelin edge cihazlar veya düşük işlem gücüne sahip donanımlar üzerinde çalıştırılabilirliğini test etmesidir. Bu sayede geliştirilen sistemin pratikte, gerçek zamanlı gözetim ve denetim

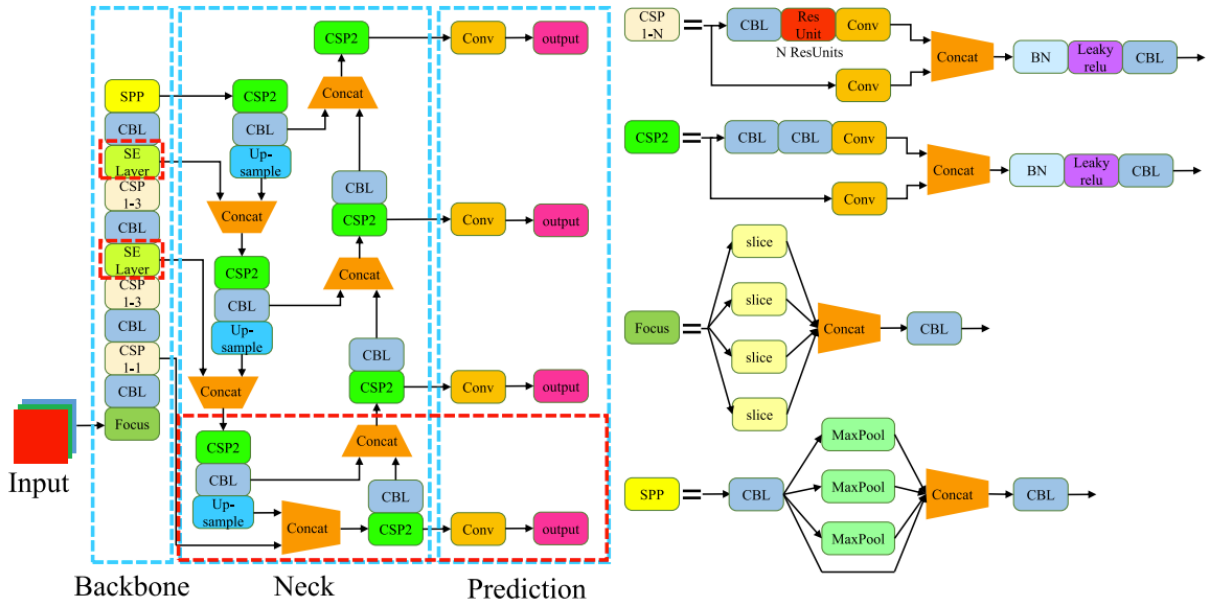
sistemlerine entegre edilebileceği ve iş sahalarında hızlı şekilde uygulanabileceği ortaya konmuştur. Sonuç olarak, bu araştırma hem veri seti oluşturma hem de uygulamalı nesne tespiti açısından literatüre önemli katkı sunmakta, yapay zekâ destekli otomatik KKD denetim sistemlerinin iş sağlığı ve güvenliği alanındaki potansiyel kullanım alanlarını başarılı şekilde ortaya koymaktadır.



Şekil 6. CNN mimarisi (Cheng vd., 2023)

Han ve Zeng (2021) tarafından yapılan “Deep learning-based workers safety helmet wearing detection on construction sites using multi-scale features” adlı çalışmada, inşaat alanlarında işçi güvenliğini artırmak amacıyla güvenlik kaskı takma durumunun tespiti için derin öğrenme tabanlı bir yöntem önermiştir (Şekil 7). Bu çalışmada YOLOv5 algoritması temel alınmış; küçük nesneleri daha iyi tanıyabilmek amacıyla dördüncü bir tespit ölçeği eklenmiş ve dikkat mekanizmaları ile ağın temsil gücü artırılmıştır. Ayrıca, eğitim verisi eksikliğini aşmak amacıyla hedefli veri artırma ve transfer öğrenme teknikleri uygulanmıştır. Mevcut literatürde yer alan geleneksel yöntemler güvenlik kaskı tespitinde sınırlı başarı göstermekte ve genellikle düşük genelleme kapasitesine sahiptir. Bu yöntemler ya sınırlı özellik çıkarımıyla çalışmakta ya da küçük ve çok ölçekli nesneler karşısında yetersiz kalmaktadır. Derin öğrenmenin yükselişiyle birlikte, iki aşamalı (Faster R-CNN) ve tek aşamalı (YOLO, SSD) tespit mimarileri bu alanda kullanılmaya başlanmıştır. Ancak çoğu yöntem ya yeterli doğruluğa ulaşamamakta ya da gerçek zamanlı performans kriterlerini karşılamamaktadır. Bu bağlamda Han ve Zeng’in önerdiği model, YOLOv5-S mimarisi üzerine inşa edilerek dördüncü ölçekli katmanla desteklenmiş, Squeeze-and-Excitation (SE) blokları ile dikkat düzeyi artırılmış ve modelin küçük nesneleri tespit kabiliyeti yükseltilmiştir. Yapılan deneyler sonucunda model, ortalama %92.2 mAP değerine ulaşarak orijinal YOLOv5-S mimarisine göre %6.3 oranında bir iyileşme sağlamıştır. Ayrıca modelin 640x640 çözünürlükte bir görüntüyü yalnızca 3.0 ms’de

işleyebilmesi, gerçek zamanlı uygulamalarda kullanılabilirliğini göstermektedir. Ek olarak, geliştirilen model kullanıcı dostu bir arayüz (GUI) ile desteklenmiş ve kullanıcıların görseller veya video akışları üzerinden doğrudan kask tespiti yapabilmesi sağlanmıştır. Bu sistem sayesinde, geleneksel manuel izleme yöntemlerine kıyasla iş yükü önemli ölçüde azaltılmakta ve anlık müdahale imkânı sunulmaktadır. Sonuç olarak, çalışma hem doğruluk hem hız açısından dengeli bir performans sunarak inşaat sahalarında işçi güvenliği denetiminde etkili ve uygulanabilir bir çözüm olarak öne çıkmaktadır.

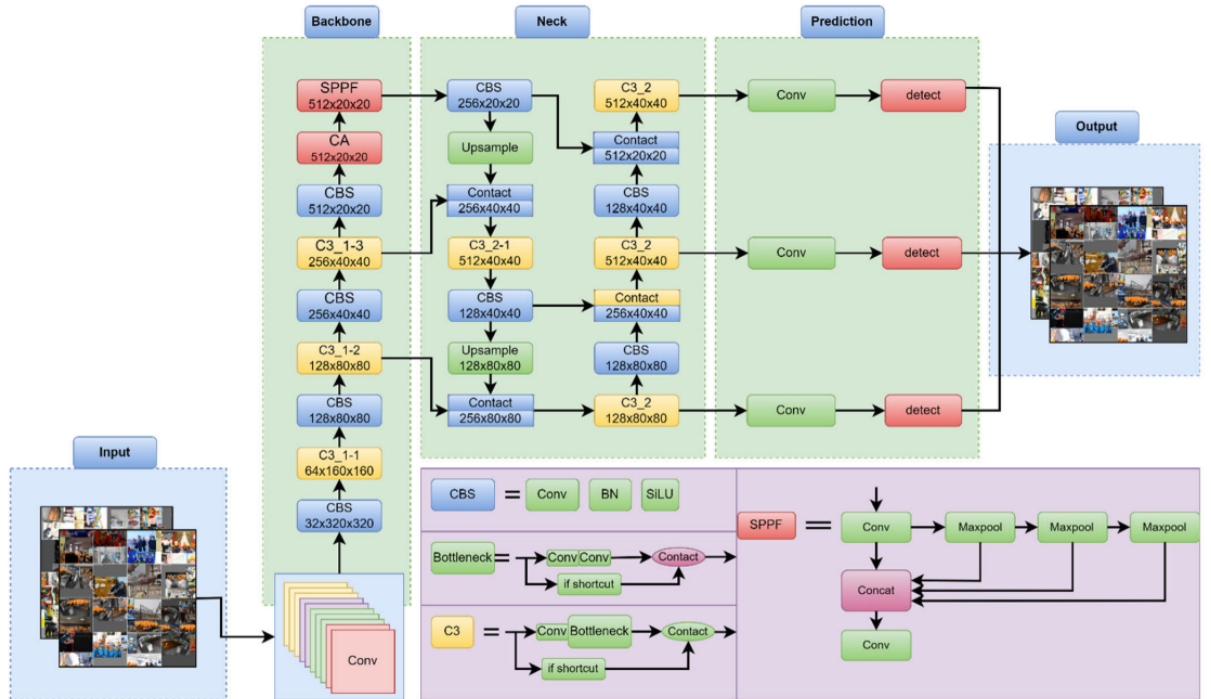


Şekil 7. Değiştirilmiş YOLOv5 yapısı (Han ve Zeng, 2021)

Yipeng ve Junwu (2024) tarafından yapılan “Personal protective equipment detection for construction workers: A novel dataset and enhanced YOLOv5 approach” adlı çalışmada, KKD tespitine yönelik mevcut çalışmalarda çoğunlukla baret tespitine odaklanıldığını, ancak reflektif yelek gibi diğer hayati donanımların göz ardı edildiğini belirterek bu eksikliği gidermeyi amaçlamıştır. Çalışmada, yarı otomatik etiketleme yöntemleriyle oluşturulan yeni bir veri seti sunulmuş hem baret hem reflektif giysi içeren dört kategorilik bu veri seti, gerçek inşaat sahası koşullarına uygun şekilde çeşitlendirilmiştir. Bu veri seti üzerine geliştirilen AL-YOLOv5 adlı model, YOLOv5 mimarisi üzerine dikkat mekanizmaları ve geliştirilmiş bir SEIoU kayıp fonksiyonu eklenerek optimize edilmiştir (Şekil 8). Çalışma, inşaat alanlarında işçilerin yeterli güvenlik bilincine sahip olmamaları, KKD kullanımını ihmal etmeleri ve manuel denetim yöntemlerinin yetersizliği nedeniyle ciddi güvenlik açıkları olduğu vurgulanmıştır. Özellikle reflektif giysilerin işçilerin görünürlüğünü artırarak ağır makine çarpışmalarını önlemede kritik

rol oynadığına dikkat çekilmiş, buna rağmen literatürde bu unsurun çoğu zaman ihmal edildiği belirtilmiştir. Reflektif giysilerin tespitinde yaşanan zorluklar; giysilerin deformasyona açık yapısı, işçilerin duruşu, çevresel etkiler ve görüntü örtüşmeleri gibi faktörlere bağlanmıştır. Mevcut tespit sistemlerinin sınırlı veriyle çalıştığı, dengesiz örnek dağılımı içerdiği ve gerçek saha verilerinden uzak olduğu ifade edilmiştir. Bu sebeple çalışmada hem dengeli ve kapsamlı bir veri seti oluşturulmuş hem de tespit doğruluğunu artırmak için SEIoU tabanlı yeni bir kayıp fonksiyonu entegre edilmiştir. Ayrıca, dikkat mekanizmaları sayesinde modelin görüntüdeki kritik KKD bileşenlerine daha hassas odaklanması sağlanmıştır.

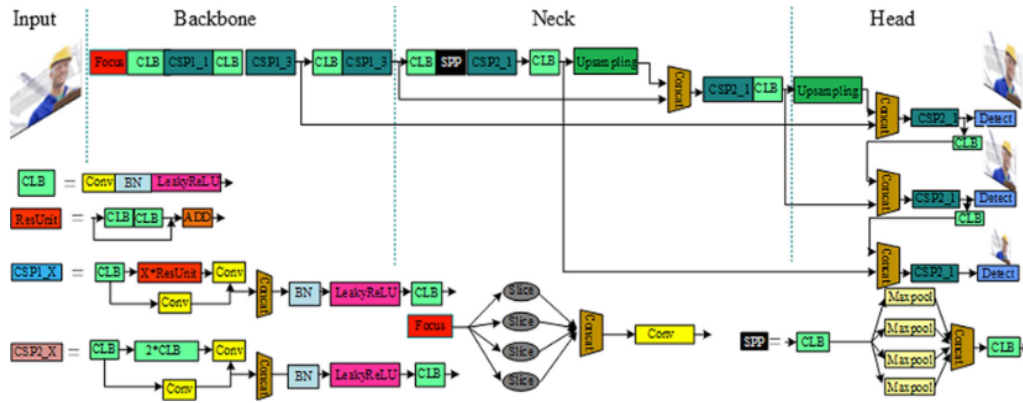
Deneyisel sonuçlar, geliştirilen AL-YOLOv5 modelinin özellikle reflektif giysi tespitinde ciddi performans artışı sağladığını; sınırlı veriye sahip kategorilerde %0.9 AP ve genel doğrulukta %0.4 mAP artışı elde ettiğini göstermiştir. Modelin örtüşen nesnelerdeki kutu tahmin doğruluğu da belirgin şekilde geliştirilmiştir. Bu sayede çalışma, otomatik, hassas ve kapsamlı bir KKD izleme altyapısına katkı sunarak inşaat sahalarında iş güvenliğine yönelik akıllı karar destek sistemleri için önemli bir temel oluşturmuştur.



Şekil 8. CA-Yolov5-6.2 yapı şeması (Yipeng ve Junwu, 2024)

Liang vd. (2024) tarafından yapılan “Detection and tracking of safety helmet wearing based on deep learning” isimli çalışmada, inşaat ve endüstriyel üretim sahalarında işçilerin güvenlik

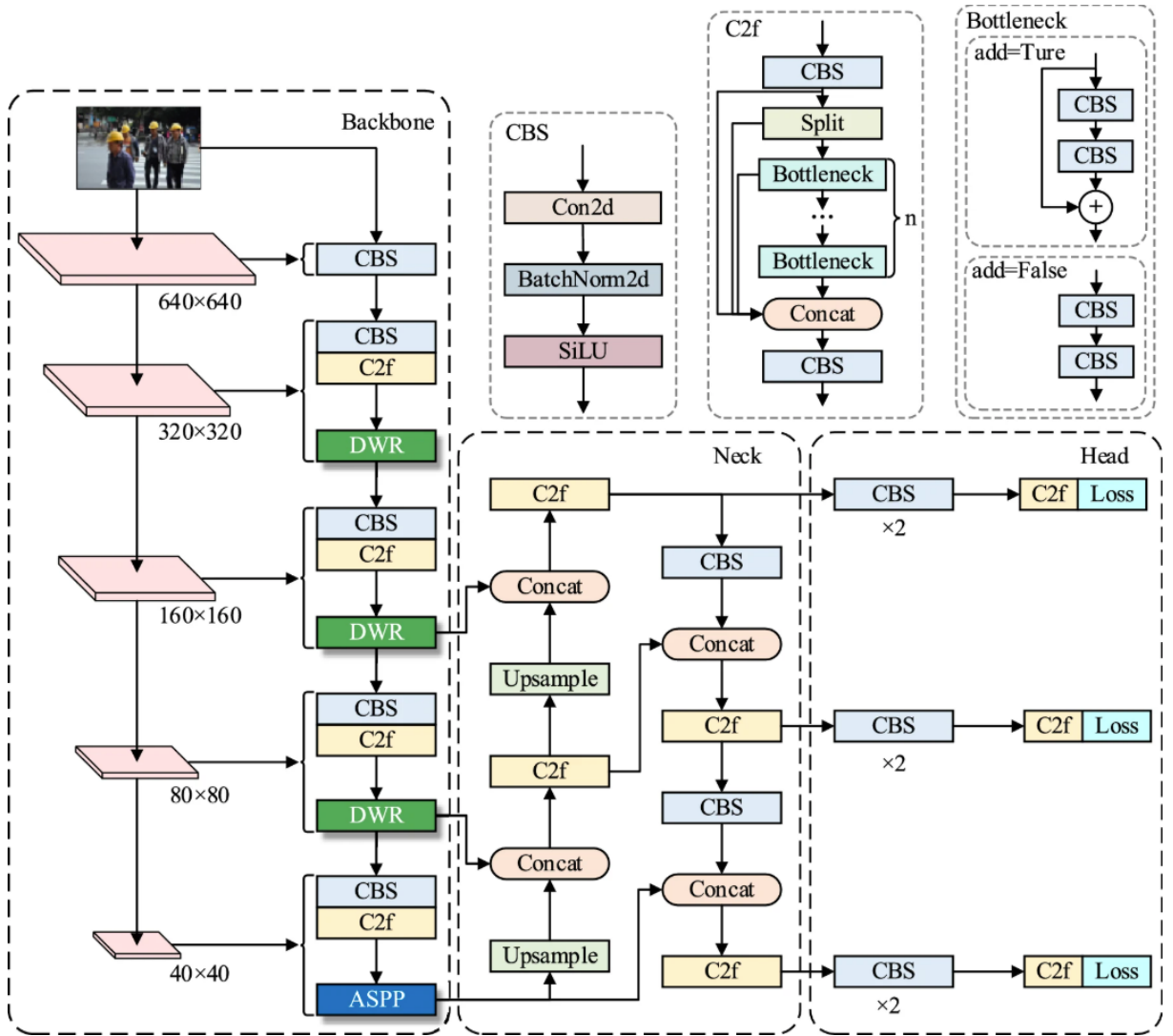
kaskı takma durumlarının gerçek zamanlı ve yüksek doğrulukla tespit edilmesini amaçlayan, derin öğrenme tabanlı bir yöntem önerilmiştir (Şekil 9). Geleneksel denetim yöntemlerinin yüksek maliyetli ve düşük verimli olmasından hareketle, araştırmacılar YOLOv5 mimarisi temelinde geliştirilen, iyileştirilmiş bir nesne tespit sistemi önermiştir. Sistem, personelin kask kullanımını anlık olarak analiz edebilmekte ve uygunsuzluk durumlarında uyarı üretebilmektedir. Modelin performansını artırmak amacıyla Generalized IoU (GIoU), Distance IoU (DIOU) ve Complete IoU (CIOU) gibi gelişmiş kayıp fonksiyonları optimize edilmiş; ayrıca Mosaic-9 adlı veri artırma algoritması kullanılarak, eğitim sürecine çeşitlilik ve çevresel zorluklara karşı dayanıklılık kazandırılmıştır. Bu geliştirmeler sayesinde, model karmaşık arka planlarda, düşük ışık koşullarında ve örtüşen nesnelerde dahi yüksek doğrulukla çalışabilmektedir. Deneysel sonuçlara göre geliştirilen sistem %93.16 doğruluk ve %88.96 duyarlılık oranlarına ulaşmıştır. Çalışmada, güvenlik kasklarının baş travmalarını azaltmadaki kritik rolü vurgulanmış, buna rağmen geleneksel manuel denetimlerin modern şantiye ortamlarında yetersiz kaldığı ifade edilmiştir. Erken dönem çalışmalar genellikle makine görüşü temelli yöntemlerle sınırlı kalmış; HOG, SVM, Hough dönüşümü gibi elle çıkarılmış özelliklere dayalı tekniklerin karmaşık sahalarda genelleme kapasitesinin düşük olduğu gösterilmiştir. Derin öğrenmeye geçişle birlikte, R-CNN, Fast/Faster R-CNN gibi iki aşamalı modeller geliştirilmiş, ancak bunlar gerçek zamanlı uygulamalar için yeterince hızlı çalışmamıştır. Bu bağlamda, tek aşamalı YOLO mimarileri dikkat çekmiş; önerilen modelde YOLOv5s ağı, BiFPN (çift yönlü özellik piramidi ağı), Soft-NMS (geliştirilmiş son işlem) ve dikkat mekanizmaları ile desteklenmiştir. Ayrıca, Focal-EIoU gibi kayıp fonksiyonlarıyla eğitim süreci daha verimli hale getirilmiştir. Sonuç olarak, önerilen BiFEL-YOLOv5s modeli, önceki modellere kıyasla hem ortalama doğrulukta (%0.9 artış) hem de çağrı oranında (%2.8 artış) belirgin iyileşmeler göstermiştir. Bu çalışma, kask takmama kaynaklı iş kazalarının azaltılmasında, otomatik ve hızlı tespit sağlayan yapay zekâ temelli çözümlerin uygulanabilirliğini ortaya koymakta ve iş güvenliğinde yeni bir denetim standardı oluşturma potansiyeli taşımaktadır.



Şekil 9. YOLOv5 hedef tespit ağının yapısı (Liang vd., 2024)

Song vd. (2024) tarafından yapılan “An improved YOLOv8 safety helmet wearing detection network” adlı çalışmada, karmaşık endüstriyel ortamlarda küçük hedefli güvenlik kaskı tespitindeki doğruluğu artırmak için DWR dikkat modülü, ASPP havuzlama ve NWD kayıp fonksiyonları ile geliştirilmiş bir YOLOv8 mimarisi önermek amaçlanmıştır. Geleneksel yöntemlerin gerçek zamanlı ve doğru tespit konusunda yetersiz kaldığı inşaat ve endüstriyel üretim alanlarında, güvenlik kaskı takılmaması kaynaklı kazalar önemli bir sorun teşkil etmektedir. Bu çalışmada, küçük nesne tespiti ve karmaşık arka planlardan kaynaklı yanlış ya da kaçırılan tespit sorunlarını azaltmaya yönelik, geliştirilmiş bir YOLOv8 temelli kask tespit ağı önerilmiştir (Şekil 10). Modelde, çok ölçekli özellik çıkarımı ve detay yakalama yeteneğini artırmak amacıyla Dilation-wise Residual Attention (DWR) modülü, Atrous Spatial Pyramid Pooling (ASPP) katmanı ve normalize edilmiş Wasserstein mesafesi (NWD) kayıp fonksiyonu entegre edilmiştir. Çalışmanın temel motivasyonu, geleneksel denetim sistemlerinin (manuel gözetim, kamera izleme) geniş alanlı ve değişken yapıları şantiye ortamlarında yeterli kapsayıcılığı ve tepki hızını sağlayamamasıdır. Bu manuel sistemler yüksek iş gücü gerektirmekte, denetim kalitesinde süreklilik sorunlarına yol açmakta ve çoğu zaman subjektif değerlendirmelere dayalı olduğundan hataya açıktır. Bu nedenle, daha sistematik ve otomatik bir denetim altyapısına ihtiyaç duyulmaktadır. Bu ihtiyaca cevap olarak geliştirilen akıllı denetim sistemleri, derin öğrenme destekli hedef tespiti sayesinde insan faktörünü ortadan kaldırarak daha hızlı, doğru ve kapsamlı çözümler sunmaktadır. Ancak, özellikle düşük ışık, değişken açılar, ekipmanla kısıtlı görüş alanları ve küçük hedeflerin (örneğin uzaktaki kask) varlığı gibi çevresel zorluklar, bu sistemlerin performansını etkilemektedir. Bu bağlamda önerilen model, özellikle bu sınırlamaları aşmayı hedeflemektedir. Geliştirilmiş YOLOv8 mimarisi, çok ölçekli konvolüsyonlarla donatılmış DWR dikkat modülü sayesinde farklı uzaklıklardaki kaskları daha iyi tanımlamakta; ASPP ile çeşitli ölçeklerden gelen bilgi işlenerek

bağlam farkındalığı artırılmakta; NWD kaybı ile de küçük hedeflerin lokalizasyon başarımı iyileştirilmektedir. SHWD veri seti üzerinde yapılan deneyler, modelin %92.0 mAP değerine ulaştığını ve geleneksel yöntemlere kıyasla hem doğruluk hem çağrı oranlarında anlamlı gelişmeler sağladığını göstermiştir. Sonuç olarak, bu çalışma, derin öğrenme temelli gelişmiş nesne tespiti tekniklerini entegre ederek küçük ölçekli hedeflerde bile yüksek doğruluk sağlayan, gerçek zamanlı ve güvenilir bir güvenlik kaskı denetim sistemi sunmaktadır. Bu sistem, şantiyelerdeki üretim güvenliğini artırma ve kazaları önleme yönünde önemli bir katkı niteliğindedir.



Şekil 10. Geliştirilmiş YOLOv8 tespit modelinin genel mimarisi (Song vd., 2024)

3. KAVRAMSAL ÇERÇEVE

Bu bölümde, çalışmanın kuramsal temelini oluşturan yapay zeka, makine öğrenmesi, derin öğrenme, bilgisayarla görme ve nesne tespiti gibi kavramlar ayrıntılı olarak ele alınmakta, bu teknolojilerin iş güvenliği bağlamındaki kullanımı değerlendirilmekte ve literatürdeki teorik yaklaşımlar incelenmektedir.

3.1. Yapay Zeka

Yapay zeka (YZ), makinelerin insan zekâsına benzer şekilde öğrenme, akıl yürütme, problem çözme ve karar verme gibi görevleri gerçekleştirmesini sağlayan bilimsel ve teknolojik bir disiplindir. YZ'nin tarihi 1950'li yıllarda Alan Turing'in "Makineler düşünebilir mi?" sorusuyla başlamış, 1956'da John McCarthy tarafından Dartmouth Konferansı'nda resmi olarak bilimsel bir alan olarak tanımlanmıştır (Russell ve Norvig, 2010).

YZ sistemleri; bilgi temsili, otomatik planlama, makine öğrenmesi, doğal dil işleme ve bilgisayarla görme gibi alt alanlara ayrılmaktadır. Günümüzde yapay zeka, büyük veri kümeleri ile entegre biçimde çalışarak insan müdahalesine gerek kalmadan öğrenme ve uyarlanabilirlik sağlayabilmektedir. Bu da sağlık, ulaşım, güvenlik, perakende ve üretim gibi sektörlerde verimlilik ve otomasyon açısından devrim yaratmaktadır (Kaplan ve Haenlein, 2019).

Özellikle endüstriyel iş güvenliği bağlamında YZ, tehlikeli ortamları denetlemek, güvenlik kurallarına uyumu sağlamak ve kazaları önlemek amacıyla giderek daha fazla kullanılmaktadır.

3.2. Makine Öğrenmesi ve Derin Öğrenme

Makine öğrenmesi (MÖ), makinelerin açık bir şekilde programlanmaksızın, veriden öğrenmesini sağlayan bir YZ alt alanıdır. Algoritmalar, verilerdeki örüntüleri öğrenerek gelecekteki durumları tahmin eder (Alpaydin, 2020). MÖ teknikleri denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenme gibi başlıklarda sınıflandırılmaktadır (Goodfellow vd., 2016).

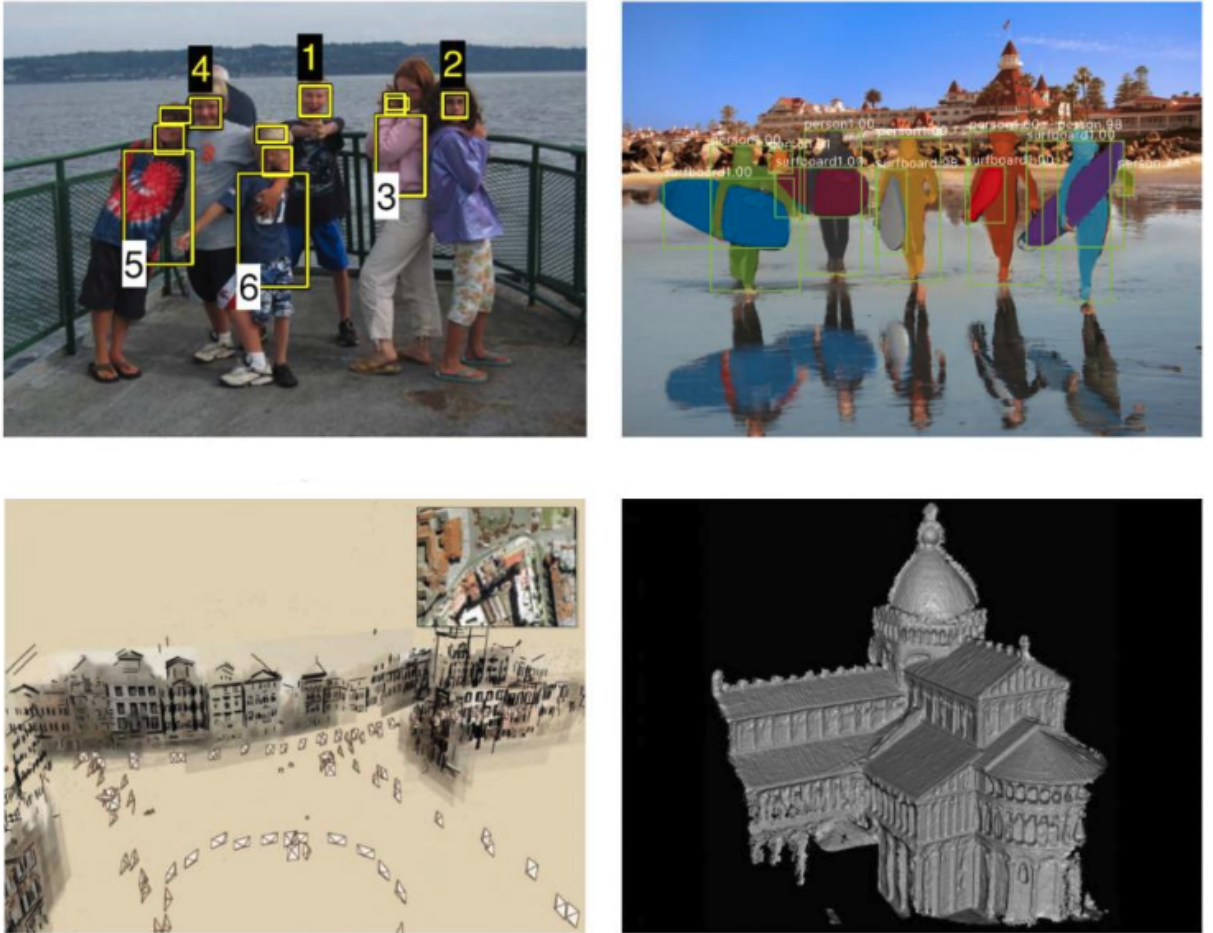
Derin öğrenme (DÖ) ise, MÖ'nün bir uzantısı olarak çok katmanlı yapay sinir ağlarını kullanarak veriden yüksek seviyeli özellik çıkarımı gerçekleştiren yöntemler bütünüdür.

Convolutional neural network (CNN) mimarisi, derin öğrenmenin özellikle görsel veriler üzerindeki başarısının temel yapı taşıdır (LeCun vd., 2015).

Görüntü tanıma, yüz tespiti, nesne sınıflandırma gibi görevlerde insan performansını aşan başarı oranlarına ulaşılması, derin öğrenmenin bilgisayarla görme alanında devrim yaratmasına neden olmuştur. Bu çalışmada kullanılan YOLO algoritmaları da CNN tabanlı derin öğrenme yöntemleridir.

3.3. Bilgisayarla Görme

Bilgisayarla görme, dijital görüntülerin analiz edilmesi yoluyla bilgisayarların görsel dünyayı algılamasını ve anlamlandırmasını sağlayan YZ alt disiplindir (Szeliski, 2010). Bu teknoloji; kenar tespiti, segmentasyon, hareket takibi, nesne tespiti, yüz tanıma gibi görevleri yerine getirmektedir (Şekil 11).



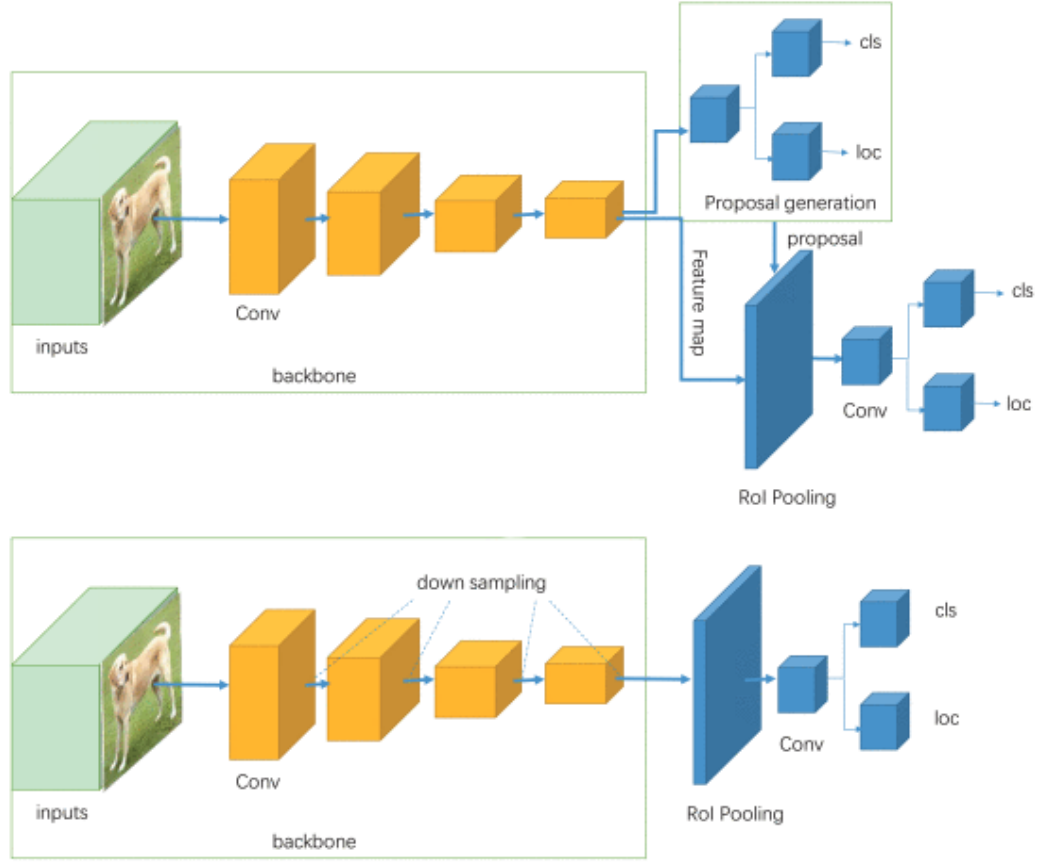
Şekil 11. Bilgisayarla görme ve algoritmalar (Szeliski, 2010)

İnsan gözü ve beyninin iş birliğiyle gerçekleştirdiği görsel algının dijital ortamda simüle edilmesi, bilgisayarla görmenin temel felsefesini oluşturur. Görüntü işleme teknikleriyle desteklenen bu yapı, güvenlik sistemleri, endüstriyel otomasyon, sağlık tanı sistemleri ve otonom araçlar gibi birçok alanda aktif olarak kullanılmaktadır.

Günümüzde iş güvenliği alanında da bilgisayarla görme, işçilerin KKD kullanımını denetlemek amacıyla kullanılmakta; özellikle baret, yelek, maske gibi ekipmanların varlığını analiz edebilen sistemlerin geliştirilmesine olanak tanımaktadır (Gholami vd., 2020).

3.4. Nesne Tespiti

Nesne tespiti, bilgisayarla görme alanının temel konularından biri olarak, bir görüntü veya video akışındaki nesnelerin tanımlanması ve konumlandırılmasını amaçlayan bir süreçtir. Nesne tespiti yalnızca "bu nesne nedir?" sorusunu değil, aynı zamanda "bu nesne görüntüde nerededir?" sorusunu da yanıtlar (Liu vd., 2020). Günümüzde nesne tespiti, güvenlik gözetim sistemlerinden otonom araçlara, tarımdan sağlık sektörüne kadar birçok farklı alanda yaygın bir şekilde kullanılmaktadır. Derin öğrenme teknikleri ile desteklenen modern nesne tespiti algoritmaları, yüksek doğruluk ve gerçek zamanlı performans sunabilmektedir.



Şekil 12. İki aşamalı dedektörlerin temel mimarisi (Liu vd., 2020)

3.4.1. Nesne tespiti türleri

Nesne tespiti genel olarak iki ana soruna yanıt verir: nesnelerin sınıflandırılması ve lokalizasyonu. Bu süreç, çeşitli veri türlerine ve uygulama alanlarına göre farklı türlerde nesne tespiti biçimlerine ayrılabilir. Başlıca türler şunlardır:

- **Tek Sınıf Nesne Tespiti:** Belirli bir nesne türünü tespit etmeyi amaçlar. Örneğin sadece baret tespiti.
- **Çoklu Sınıf Nesne Tespiti:** Aynı anda birden fazla nesne sınıfını tespit eder. Örneğin hem baret hem de yelek tespiti.
- **Çoklu Nesne Tespiti:** Bir görüntüde birden fazla nesne örneğini tanımlar ve konumlandırır.
- **Zaman Serili Nesne Tespiti:** Video gibi zamanla değişen veri akışlarında nesnelerin takibini içerir.

3.4.1.1. Nesne tespiti süreci

Nesne tespiti süreci genellikle şu adımları izler:

1. **Giriş Verisinin Hazırlanması:** Görüntü verisi uygun formatta ve boyutta modele sunulur.
2. **Özellik Çıkarımı:** CNN gibi yöntemlerle görüntüden anlamlı özellikler çıkarılır.
3. **Bölge Önerileri:** Potansiyel nesnelerin yer aldığı bölgeler belirlenir (özellikle iki aşamalı yöntemlerde).
4. **Sınıflandırma:** Her bölgeye ait nesne sınıfı belirlenir.
5. **Bounding Box Regresyonu:** Nesnenin kesin konumu (x, y, genişlik, yükseklik) tahmin edilir.

Güncel yaklaşımlar bu adımları tek bir yapı içinde entegre ederken, geleneksel yöntemler her adımı ayrı gerçekleştirir (Girshick vd., 2014).

3.4.2. Nesne tespiti algoritma çeşitleri

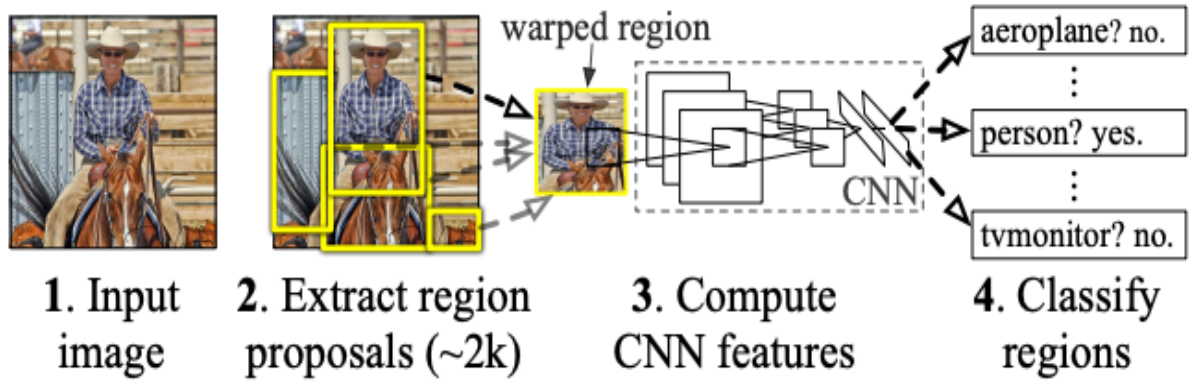
Nesne tespiti algoritmaları, genel olarak iki aşamalı (two-stage) ve tek aşamalı (one-stage) mimariler olarak ikiye ayrılır.

3.4.2.1. İki aşamalı nesne tespiti

İki aşamalı yaklaşımlar ilk olarak olası nesne bölgelerini belirleyip, daha sonra bu bölgeleri sınıflandırır. Bu yapılar, yüksek doğruluk sağlar ancak gerçek zamanlı uygulamalar için genellikle yavaştır.

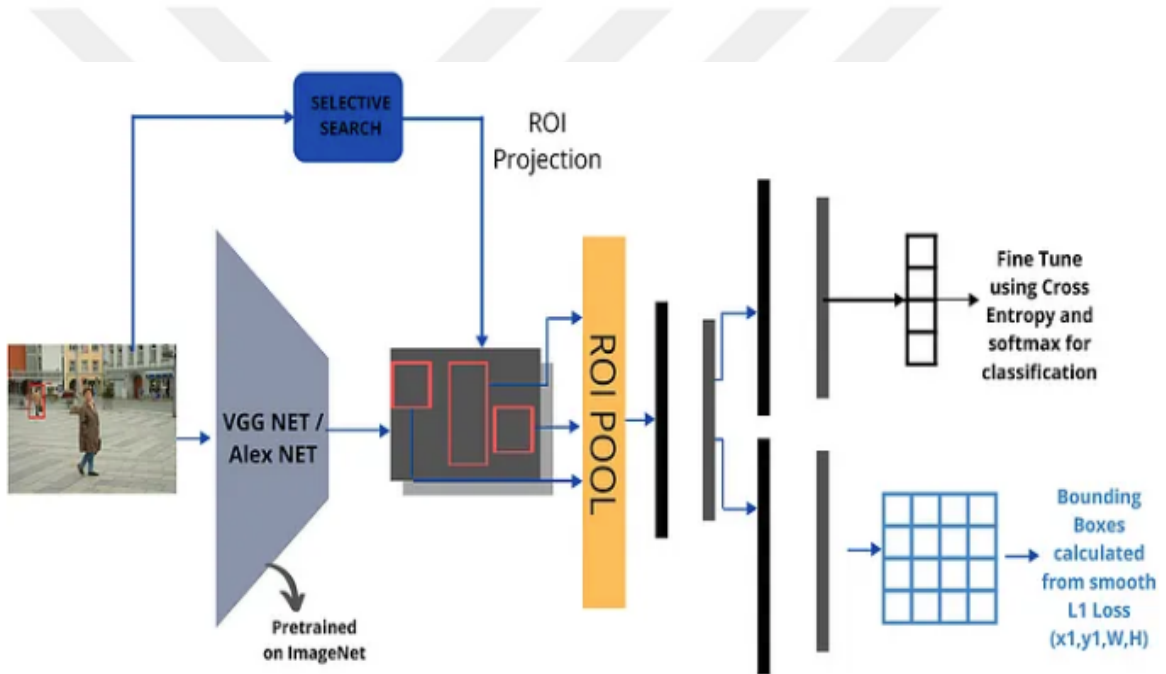
Başlıca algoritmalar:

- **Region-based CNN (R-CNN):** İlk olarak Girshick vd. (2014) tarafından önerilmiştir. Bölge önerileri Selective Search ile çıkarılır ve her bölge ayrı CNN'e gönderilir (Şekil 13).



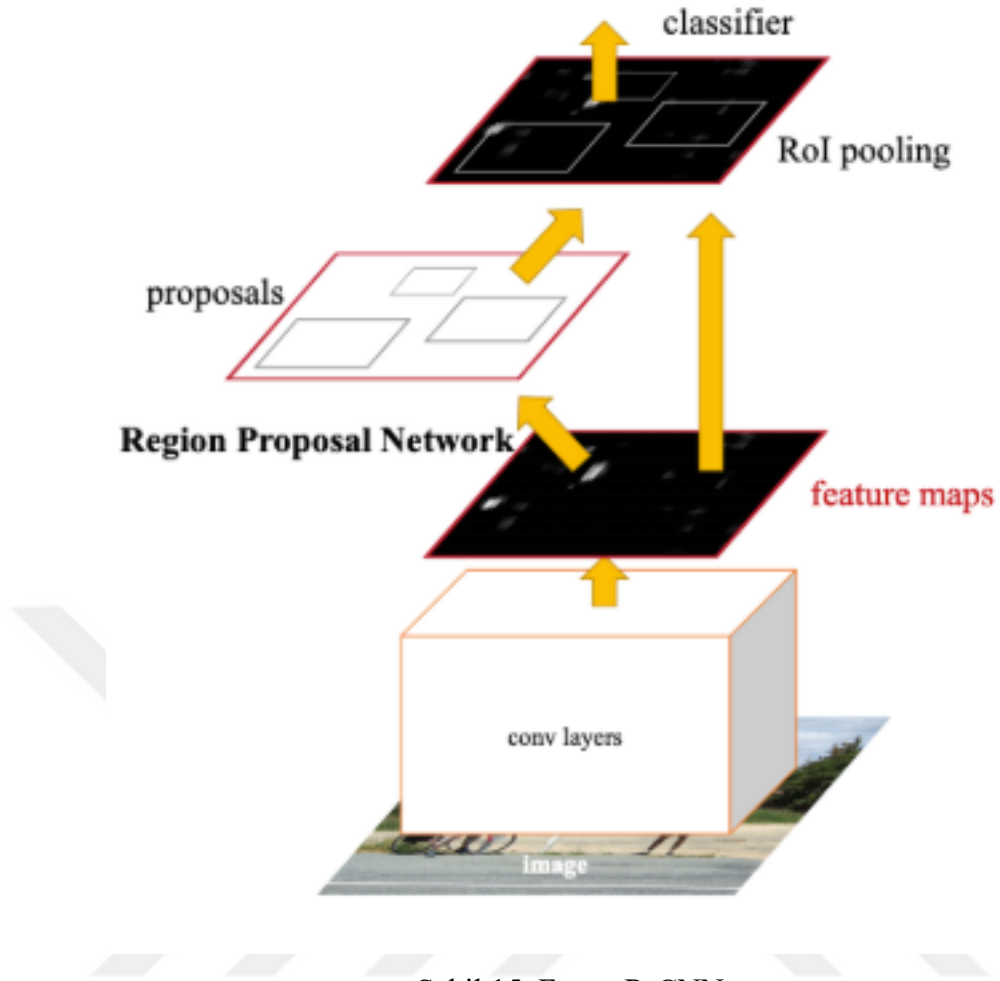
Şekil 13. R-CNN (Girshick vd., 2014)

- **Fast R-CNN:** R-CNN'in hesaplama maliyetini azaltan geliştirilmiş versiyondur. Özellik haritası tek seferde çıkarılır, ROI Pooling uygulanır (Şekil 14).



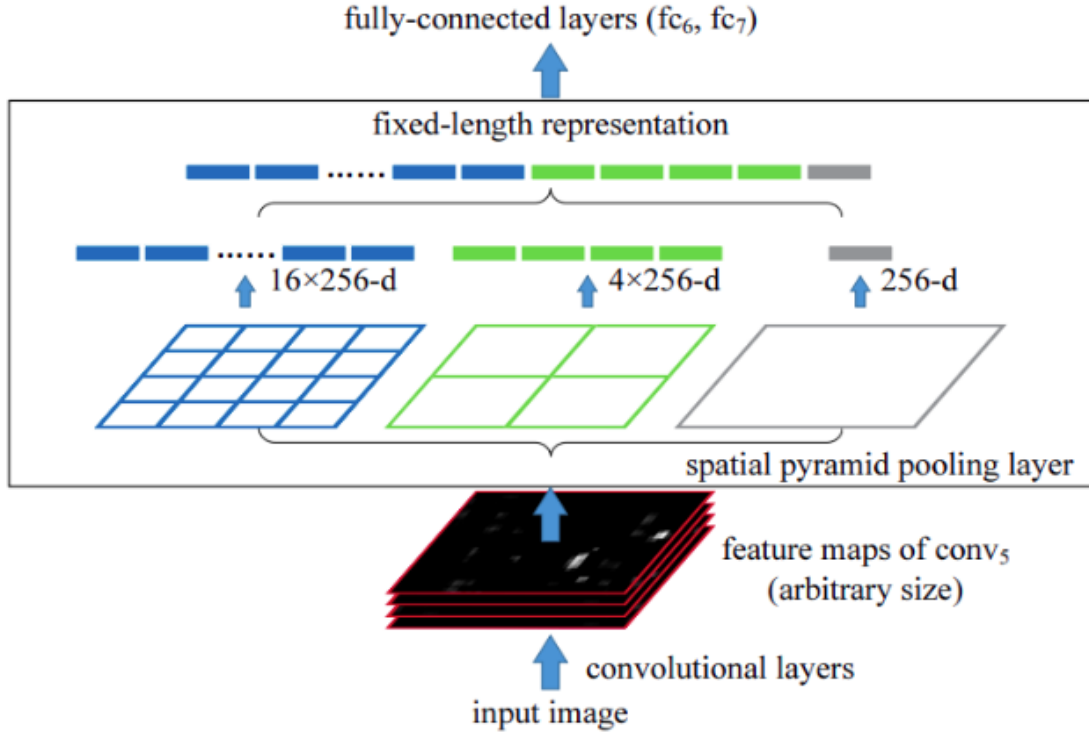
Şekil 14. Fast R-CNN

- **Faster R-CNN:** Region Proposal Network (RPN) ile bölge önerme sürecini tamamen öğrenilebilir hale getirir (Ren vd., 2015).



Şekil 15. Faster R-CNN

- **Spatial Pyramid Pooling Network (SPPNet):** Sabit boyut zorunluluğunu ortadan kaldırır, farklı boyutlardaki görüntülerle çalışmayı mümkün kılar (He vd., 2015).



Şekil 16. SPPNet

İki aşamalı yaklaşımlar, genellikle daha karmaşık ve detaylı tespitlerin gerektiği medikal görüntüleme veya uydu fotoğrafları gibi alanlarda tercih edilir.

3.4.2.2. Tek aşamalı nesne tespiti

Tek aşamalı modeller, nesne tespiti işlemini doğrudan ve hızlı bir şekilde gerçekleştirir. Bölge önerileri üretmeden doğrudan sınıflandırma ve lokalizasyon yaparlar.

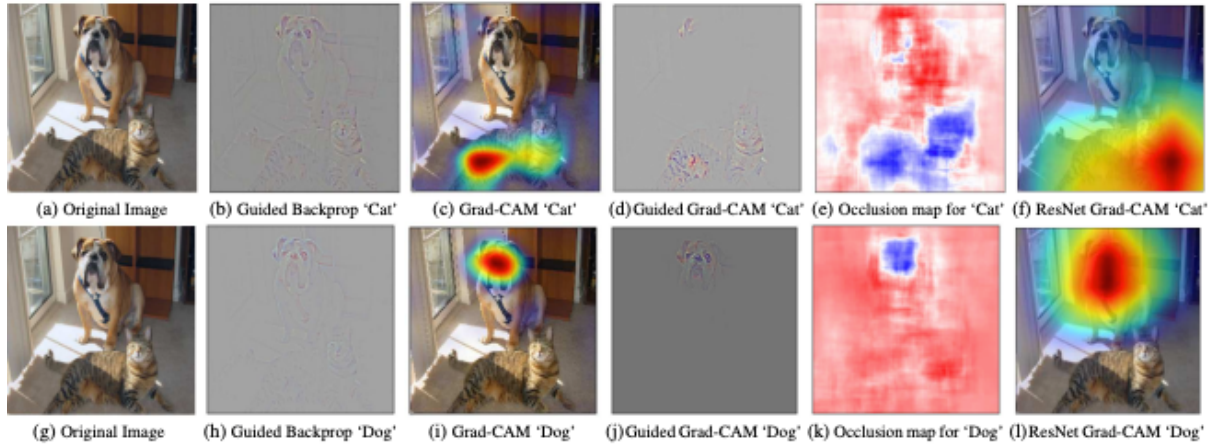
Öne çıkan yöntemler:

- **YOLO:** Görüntüyü sabit sayıda bölgeye ayırır, her birinden sınıf ve konum tahmini yapar (Redmon vd., 2016).
- **YOLOv3/v5/v7/v8:** YOLO mimarisinin evrimsel sürümleri, daha yüksek doğruluk ve hız sağlar. YOLOv5, PyTorch tabanlı kolay kullanım; YOLOv7 ise hız/performans dengesi; YOLOv8 ise küçük nesneler için yüksek başarı sağlamaktadır.
- **Single Shot MultiBox Detector (SSD):** Farklı ölçeklerdeki öznetelik haritaları üzerinden tahmin yapılmasını sağlar (Liu vd., 2016).
- **RetinaNet:** Focal Loss fonksiyonuyla nadir görülen sınıflara odaklanır (Lin vd., 2017).

Tek aşamalı mimariler, özellikle gözetim sistemleri, dronelar ve mobil cihazlar gibi gerçek zamanlı uygulamalarda yaygın olarak kullanılır.

3.4.3. Isı haritası tabanlı nesne tespiti

Isı haritası tabanlı yöntemler, görüntü üzerinde belirli bölgelerin etkinliğini ya da önemini vurgulamak amacıyla renk kodlaması kullanır. Özellikle modelin karar verdiği alanların görselleştirilmesi için kullanılır. Bunlar genellikle Class Activation Map (CAM) veya Grad-CAM yöntemleriyle gerçekleştirilir (Selvaraju vd., 2017). Isı haritaları, modelin hangi bölgelerden ne kadar bilgi edindiğini anlamak ve açıklanabilir yapay zeka (Explainable AI - XAI) bağlamında model kararlarını yorumlamak için kritik rol oynar.



Şekil 17. Isı haritası tabanlı nesne tespiti

Bu yöntemler ayrıca eğitim sırasında attention mekanizmalarının anlaşılması, modelin overfitting yaptığı bölgelerin tespiti ve güvenlik alanlarında modelin hangi nesnelere odaklandığını belirlemek açısından önemli analiz araçlarıdır.

3.5. İş Güvenliği ve Görsel Denetim Sistemleri

İş sağlığı ve güvenliği (İSG), çalışanların iş ortamındaki tehlikelerden korunmasını ve iş kazalarının önlenmesini amaçlayan multidisipliner bir çalışma alanıdır. Dünya Sağlık Örgütü'ne göre dünya genelinde her yıl 2.78 milyon kişi iş kazaları veya işle ilgili hastalıklar nedeniyle hayatını kaybetmektedir (ILO, 2022). Bu durum, iş güvenliği sistemlerinin teknolojik araçlarla desteklenmesini zorunlu hale getirmiştir.

KKD kullanımı, iş güvenliği yönetmeliklerinin temel bileşenlerinden biridir. Ancak manuel denetim mekanizmalarının yetersiz kaldığı durumlarda, YZ ve bilgisayarla görme destekli sistemler bu denetimi otomatikleştirerek ciddi katkılar sunmaktadır (Fang vd., 2018).

Bu bağlamda geliştirilen görsel denetim sistemleri, kamera görüntülerinden işçilerin baret ve yelek gibi zorunlu KKD'leri kullanıp kullanmadığını anında tespit edebilmekte, olası ihlalleri raporlayarak önleyici tedbirlerin hızla alınmasına olanak tanımaktadır. Bu çalışmada, söz konusu teknoloji gerçek zamanlı bir şekilde uygulanarak iş güvenliği süreçlerine yenilikçi bir katkı sunmaktadır.

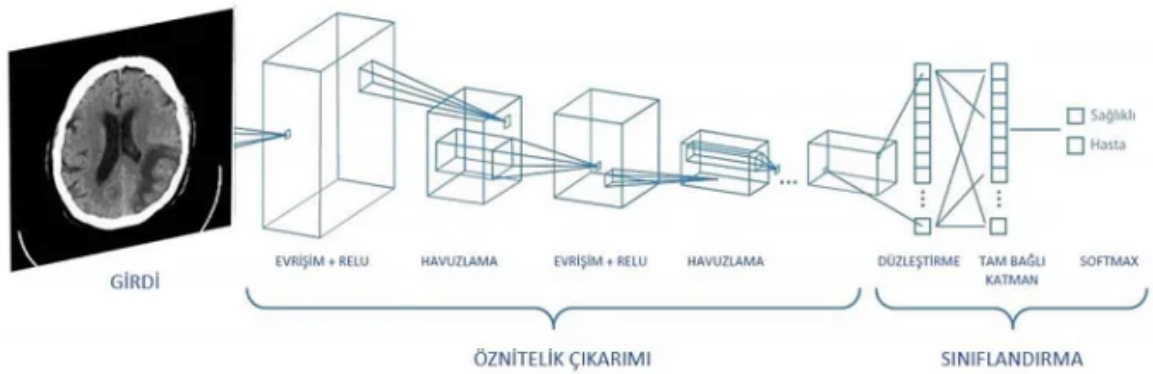


4. NESNE TESPİTİNDE KULLANILAN YÖNTEMLER

Derin öğrenme, özellikle CNN mimarileri, nesne tespiti alanında devrim yaratmıştır. Bu bölümde, CNN tabanlı nesne tespiti yöntemlerinin evrimi, iki aşamalı ve tek aşamalı tespit modelleri, tespit işleminin hızlandırılması ve model seçimi konuları detaylı bir şekilde incelenecektir.

4.1. CNN Modelinin Gelişimi

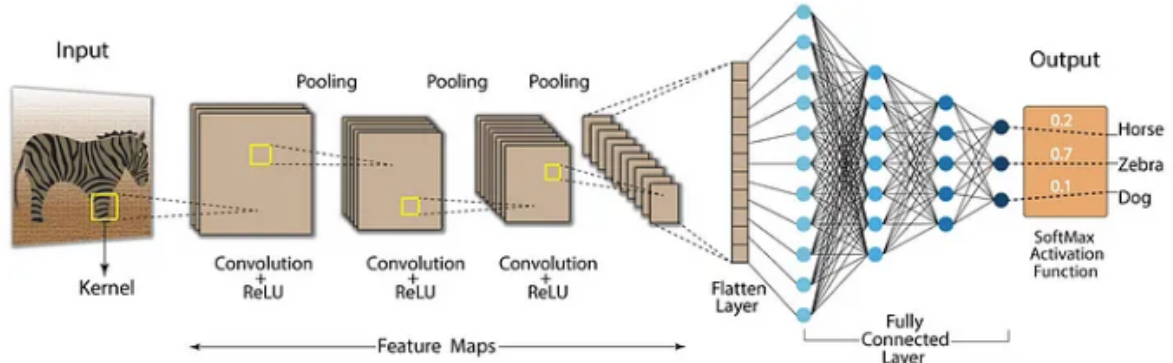
Convolutional Neural Networks (CNN), görüntü işleme ve nesne tespiti gibi görsel algı görevlerinde devrim yaratan bir derin öğrenme mimarisi olarak öne çıkmaktadır. CNN'ler, verilerdeki mekânsal ve zamansal bağıntıları öğrenme yetenekleri sayesinde, özellikle bilgisayarla görme alanında büyük başarı sağlamışlardır. İlk olarak LeCun ve arkadaşları tarafından geliştirilen LeNet-5 mimarisi, el yazısı karakter tanıma gibi temel görevlerde etkili olmuş; ancak modern CNN mimarilerinin temelini oluşturan sıçrama noktası AlexNet'in 2012 yılında ImageNet yarışmasında büyük başarı kazanmasıyla gerçekleşmiştir (Krizhevsky vd., 2012).



Şekil 18. CNN temel yapı (Özcan, 2024)

AlexNet'in başarısı, GPU'ların hesaplama gücünün kullanılmaya başlanması ve ReLU aktivasyon fonksiyonu gibi yenilikçi bileşenleri sayesinde mümkün olmuştur. Bu başarıyı takiben geliştirilen VGGNet (Simonyan & Zisserman, 2014), GoogLeNet (Szegedy vd., 2015), ResNet (He vd., 2016) gibi mimariler, daha derin ve karmaşık ağ yapılarının nasıl daha doğru

tahminler yapılmasına katkı sağladığını göstermiştir. CNN'lerin gelişimi, aynı zamanda nesne tespitinde daha etkili ve hızlı modellerin ortaya çıkmasına zemin hazırlamıştır.



Şekil 19. ReLU katmanları (Özcan, 2024)

Son yıllarda yapılan çalışmalar, geleneksel CNN mimarilerinin ötesine geçerek dikkat mekanizmaları, transformer mimarileri ve hafif yapılar (mobil CNN'ler) gibi daha verimli ve özelleştirilebilir modellerin geliştirilmesini sağlamıştır. Örneğin, EfficientNet (Tan ve Le, 2019) model ölçekleme ilkelerini kullanarak daha az parametreyle yüksek doğruluk sağlayan bir mimari sunarken, MobileNetV3 (Howard vd., 2019) mobil cihazlar için optimize edilmiştir.

Özellikle nesne tespiti bağlamında, CNN temelli mimariler; öz nitelik çıkarımı, bölge önerisi ve sınıflandırma işlemlerini birbirine entegre ederek yüksek doğruluk oranlarına ulaşmaktadır. Bu süreçler, iki aşamalı ve tek aşamalı nesne tespiti modellerinin gelişmesine temel oluşturmuştur.

Ayrıca, son dönemde yapılan çalışmalarda CNN'lerin transformer tabanlı yapılarla hibrit olarak kullanımı dikkat çekmektedir. Vision Transformer (Dosovitskiy vd., 2020) ve Swin Transformer (Liu vd., 2021) gibi modeller, CNN'lerin lokal özellik çıkarım yeteneklerini, transformerların küresel dikkat mekanizması ile birleştirerek daha üstün performanslar elde edilmesini sağlamıştır.

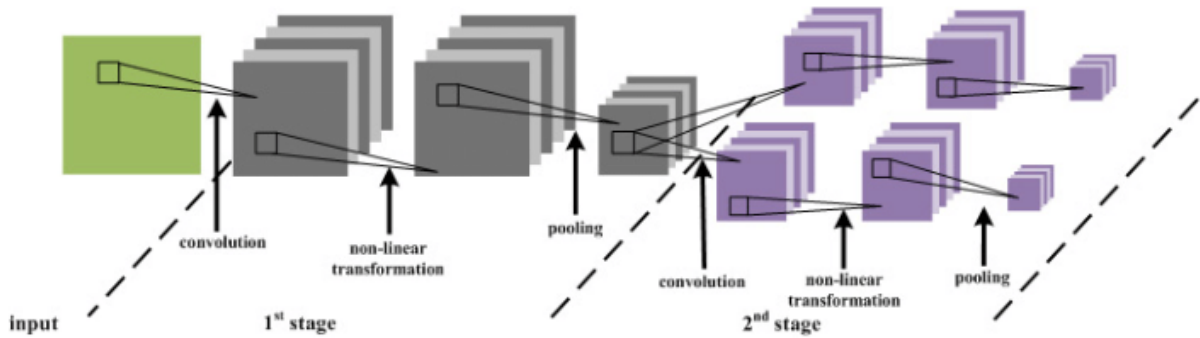
CNN'lerin bu evrimi, yalnızca doğruluğu artırmakla kalmamış, aynı zamanda daha az donanım kaynağı kullanımı ve gerçek zamanlı sistemlerde uygulanabilirliği de beraberinde getirmiştir. Bu bağlamda, özellikle endüstriyel alanlarda (örneğin işçi güvenliği, otomotiv, üretim tesisleri

vb.) gerçek zamanlı görüntü işleme uygulamalarında tercih edilen temel yapı taşlarından biri haline gelmiştir.

4.2. CNN Temelli İki Aşamalı Tespit Modelleri

İki aşamalı (two-stage) nesne tespit yaklaşımları, nesne tespitinde yüksek doğruluk sağlamasıyla öne çıkan yöntemlerdir. Bu yaklaşımlar, genel olarak ilk aşamada görüntü üzerindeki olası nesne konumlarını belirleyen bir bölge öneri ağı (Region Proposal Network - RPN) veya benzeri bir mekanizma kullanır. İkinci aşamada ise, bu aday bölgeler sınıflandırılır ve sınırlayıcı kutular hassas bir şekilde ayarlanır. İki aşamalı mimariler, işlem süresi açısından daha yavaş olabilese de özellikle doğruluk gereksinimi yüksek olan uygulamalarda tercih edilmektedir (Ren vd., 2015).

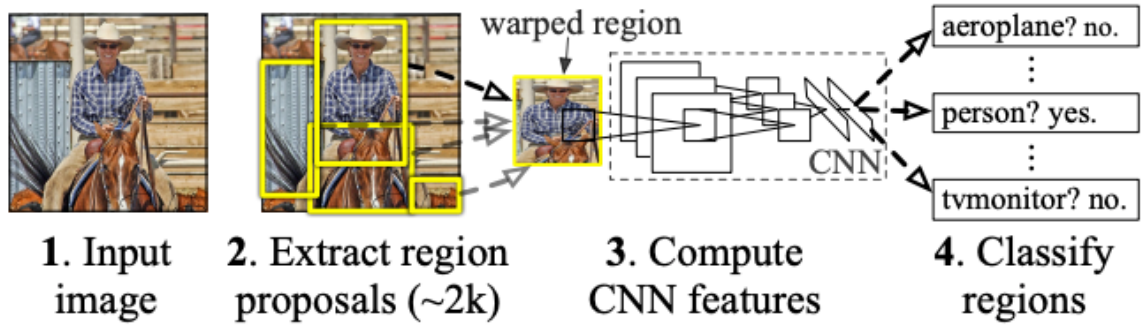
Bu yaklaşım, R-CNN ile başlamış ve zamanla daha hızlı ve verimli hale getirilmiş türev modeller (Fast R-CNN, Faster R-CNN, SPPNet, FPN vb.) geliştirilmiştir. Aşağıdaki alt başlıklarda bu modeller detaylı bir şekilde ele alınacaktır.



Şekil 20. CNN temelli iki aşamalı tespit modeli (Zhai vd., 2016)

4.2.1. Regions with CNN (R-CNN)

R-CNN (Girshick vd., 2014), nesne tespiti alanında derin öğrenme tabanlı ilk büyük adımlardan biridir. Bu model, klasik bölge önerme yöntemleri (örneğin Selective Search) kullanarak bir görüntüdeki potansiyel nesne bölgelerini tespit eder. Her bir bölge için CNN ile öznetelik çıkarımı yapılır ve bu öznetelikler, SVM sınıflandırıcıları ve sınır kutusu regresörleri ile değerlendirilir.



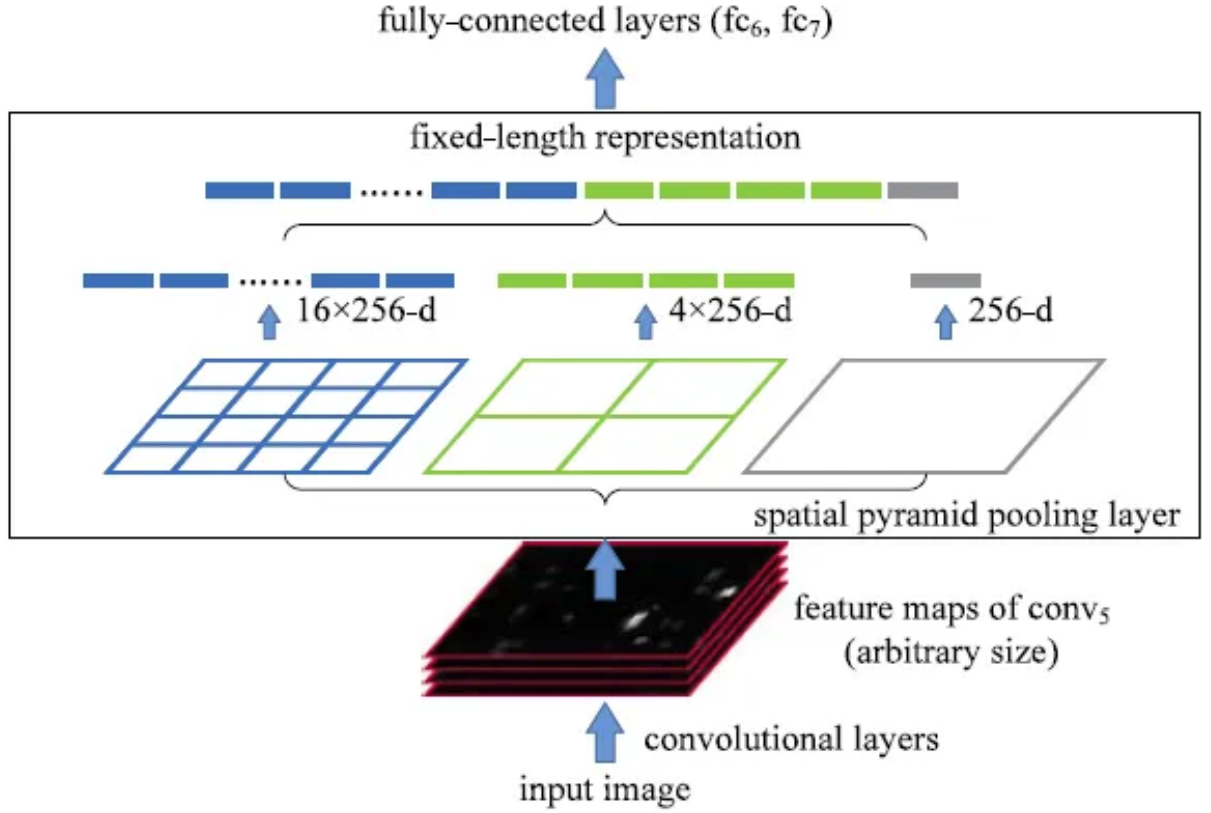
Şekil 21. R-CNN (Girshick vd., 2014)

Ancak R-CNN’in en büyük dezavantajı, her bir bölge önerisi için ayrı ayrı CNN çalıştırmasıdır. Bu durum, işlem süresini oldukça uzatmakta ve gerçek zamanlı uygulamaları olanaksız hale getirmektedir.

Her ne kadar günümüzde R-CNN doğrudan kullanılsa da önerdiği bölge tabanlı yapı birçok modern mimarinin temelini oluşturmuştur. Bu mimarideki bölge temelli yaklaşım daha sonra Fast R-CNN ve Faster R-CNN gibi daha verimli modellerin geliştirilmesine zemin hazırlamıştır.

4.2.2. Spatial pyramid pooling network (SPPNet)

SPPNet, R-CNN’in zaman maliyetini azaltmak amacıyla geliştirilmiştir (He vd., 2015). Bu model, tüm görüntü için yalnızca bir defa CNN çalıştırarak öznitelik haritasını çıkarır. Daha sonra, her bir önerilen bölge için öznitelik haritası üzerinden “spatial pyramid pooling” uygulanarak sabit boyutta bir öznitelik vektörü elde edilir. Bu yöntem, çok sayıda bölge için CNN’i yeniden çalıştırma ihtiyacını ortadan kaldırır.



Şekil 22. SPPNet'te piramit ile üç seviyeli mekansal piramit havuzlaması (Pham, 2018)

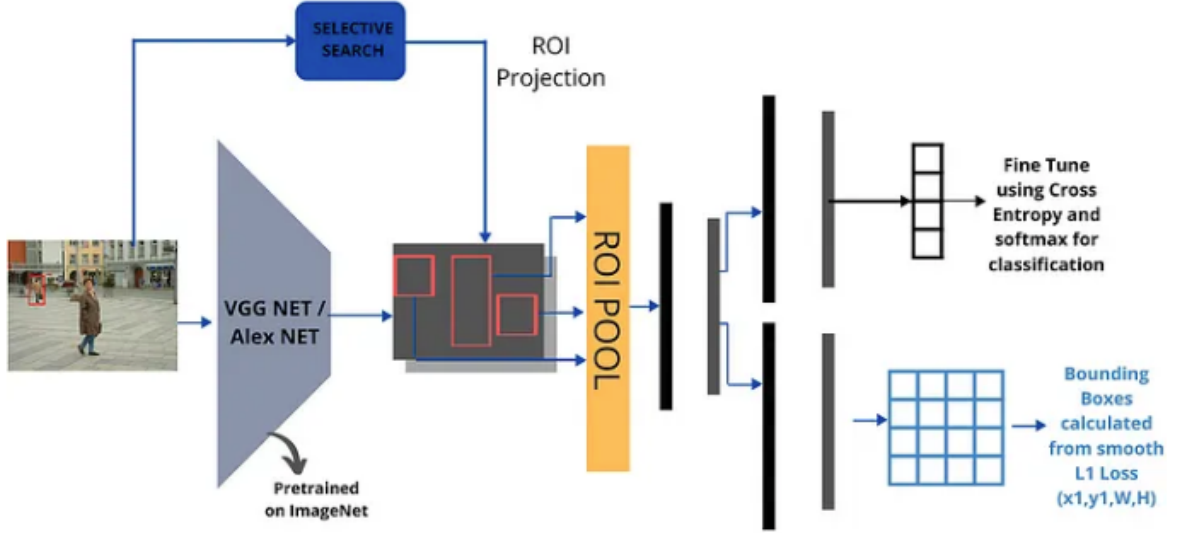
SPPNet'in getirdiği en önemli katkılardan biri, değişken boyutlardaki girişlerin sabit boyutta öznitelik vektörüne dönüştürülmesidir. Bu hem işlem süresini azaltmış hem de doğruluğu koruyarak R-CNN'den daha hızlı bir yapı sunmuştur.

SPPNet'in temel fikirleri günümüzde hala kullanılmakta, özellikle FPN ve bazı YOLO mimarilerinde bu tür sabitleme yaklaşımları entegre edilmektedir.

4.2.3. Fast R-CNN

Fast R-CNN, hem R-CNN'in doğruluğunu koruyarak hem de SPPNet'in hız avantajlarını benimseyen bir modeldir (Girshick, 2015). Bu modelde tüm görüntü için bir kez konvolüsyon uygulanır ve bölge önerileri bu öznitelik haritası üzerinde işlenir. Bu bölgeler, ROI pooling katmanı aracılığıyla sabit boyuta getirilir ve tam bağlantılı katmanlarla sınıflandırma ve sınır kutusu düzeltmeleri yapılır.

Fast R-CNN, önceki modellere göre çok daha hızlı ve daha az disk belleği kullanır. En önemli farklardan biri, SVM yerine softmax katmanı ile sınıflandırma yapması ve tüm bileşenlerin tek bir ağda entegre çalışmasıdır.

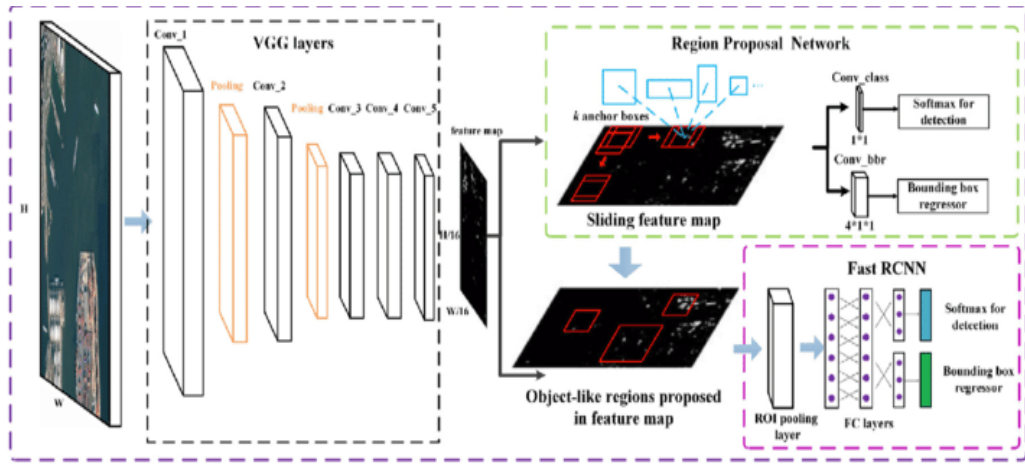


Şekil 23. Fast R-CNN mimarisi (Vinayak, 2020)

4.2.4. Faster R-CNN

Faster R-CNN, bölge önerilerini dış sistemlere (selective search gibi) bırakmak yerine, doğrudan sinir ağı içinde bir Region Proposal Network (RPN) bileşeniyle üretir (Ren vd., 2015). Bu yapı, tam anlamıyla uçtan uca eğitim yapılabilen ilk bölge tabanlı nesne tespit modelidir. RPN, öznitelik haritası üzerinde kayan pencerelerle aday bölgeler üretir ve bunlar ROI pooling ile sınıflandırma ve regresyon işlemlerine gönderilir.

Faster R-CNN, doğruluk ve işlem süresi açısından önceki modellere kıyasla önemli avantajlar sunmuş; günümüzde ise hâlâ akademik çalışmalarda ve endüstriyel uygulamalarda yaygın biçimde kullanılmaktadır. Ayrıca, ResNet-101, ResNeXt ve EfficientNet gibi daha derin backbone'lar ile çeşitli eklenti yapılar entegre edilerek güncel versiyonları geliştirilmiştir (Zhao vd., 2019).

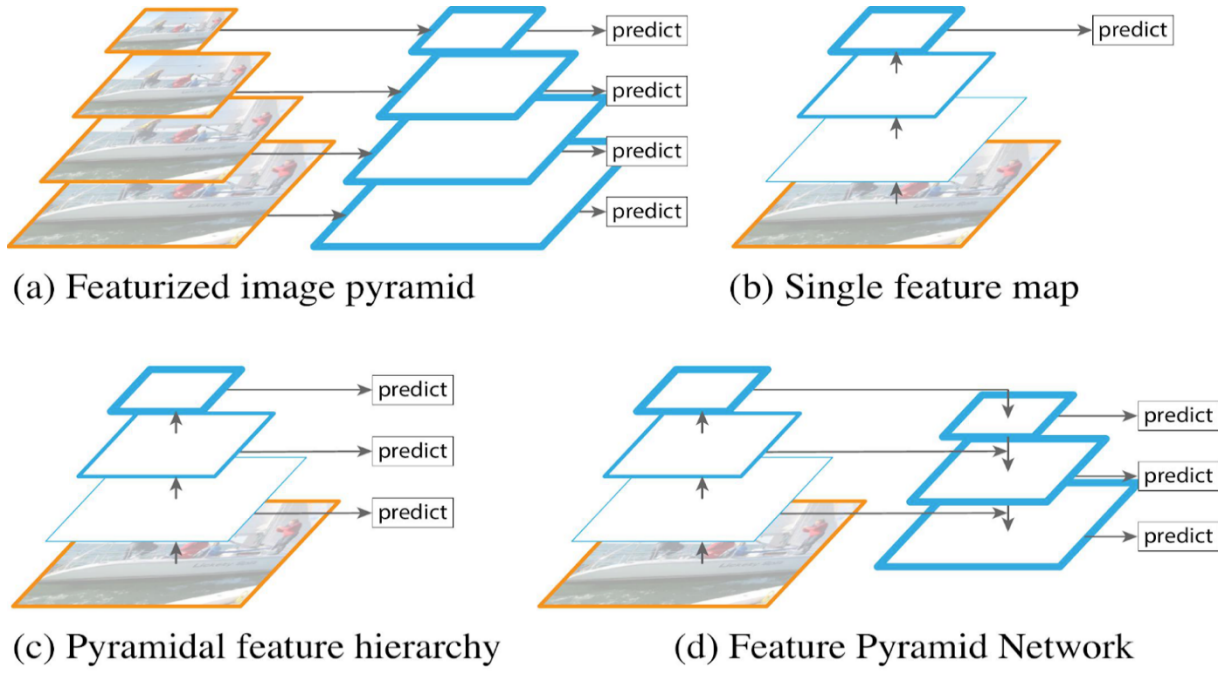


Şekil 24. Faster R-CNN mimarisi (Li vd., 2018)

4.2.5. Feature pyramid networks (FPN)

FPN, farklı seviyelerdeki öznitelik haritalarını birleştirerek küçük nesnelerin tespitinde performansı artırmayı amaçlayan bir yapıdır (Lin vd., 2017). CNN mimarilerinde, ağ derinleştikçe semantik bilgi artarken çözünürlük düşer. FPN, bu sorunu çözmek için düşük seviyeli (yüksek çözünürlüklü) ve yüksek seviyeli (düşük çözünürlüklü) öznitelikleri birleştirerek daha dengeli bir temsil oluşturur.

FPN, Faster R-CNN gibi modellere entegre edilerek hem küçük hem büyük nesnelerin tespitinde güçlü sonuçlar vermektedir. Günümüzde RetinaNet, Mask R-CNN gibi birçok gelişmiş modelde FPN yapısı kullanılmaktadır.



Şekil 25. Feature Pyramid Networks (Lin vd., 2017)

4.3. CNN Temelli Tek Aşamalı Tespit Modelleri

Tek aşamalı nesne tespit modelleri, sınıflandırma ve konumlandırma işlemlerini tek bir adımda gerçekleştirerek gerçek zamanlı uygulamalarda büyük avantaj sağlamaktadır. Bu tür modeller, bölge öneri adımını atlayarak daha hızlı çalışır; bu da özellikle gömülü sistemler, mobil cihazlar ve düşük gecikmeli güvenlik uygulamaları için kritiktir (Liu vd., 2020).

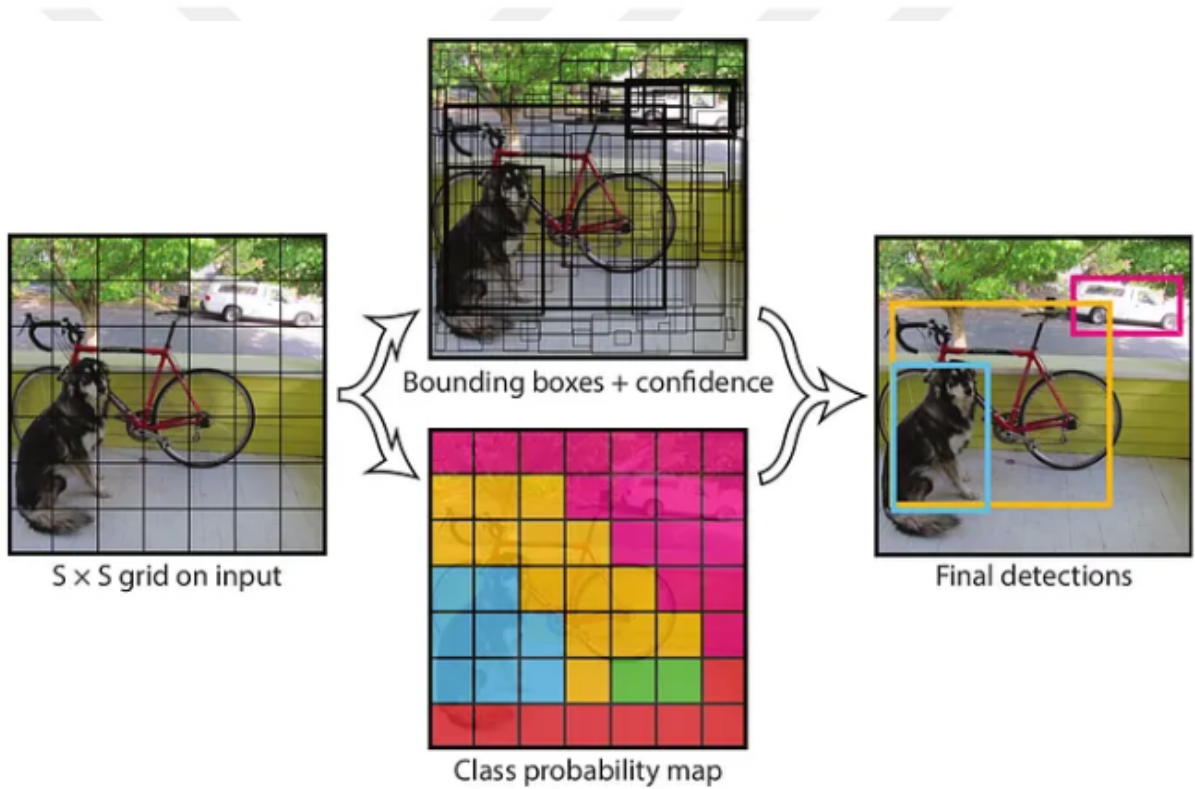
YOLO ailesi, SSD ve RetinaNet, bu sınıfta yer alan en yaygın modellerdendir. Bu bölümde, bu modellerin her biri güncel versiyonlarıyla birlikte detaylı olarak ele alınacaktır.

4.3.1. You only look once (YOLO)

YOLO mimarisi, nesne tespitini bir sınıflandırma problemine benzeterek tek bir CNN ağı ile giriş görüntüsünü birden fazla bölgeye böler ve her bir bölge için sınırlayıcı kutu tahmini yapar. YOLO'nun temel avantajı yüksek hızıdır; bu nedenle gerçek zamanlı uygulamalarda sıkça tercih edilir (Redmon vd., 2016).

4.3.1.1. YOLOv1 - YOLOv3

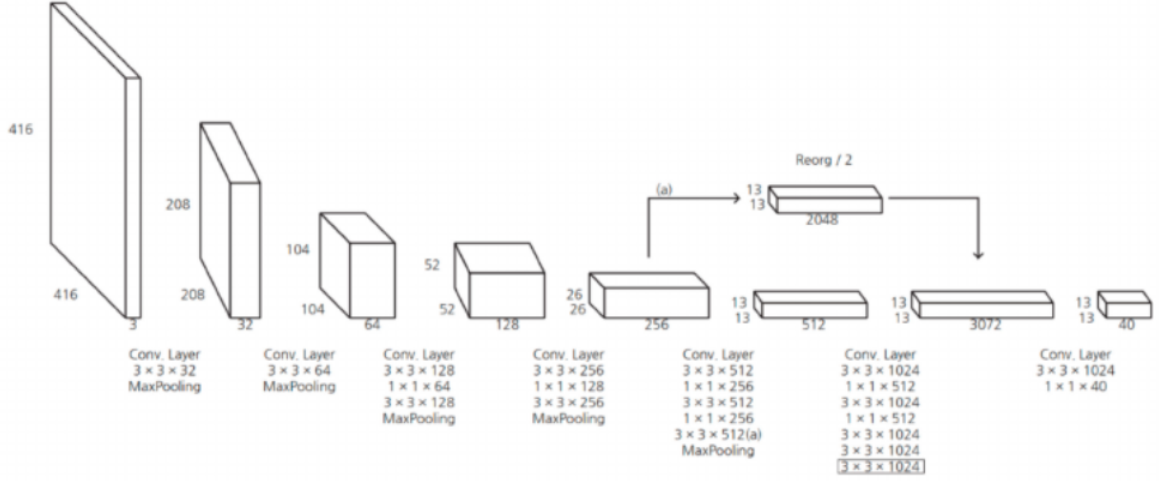
YOLO algoritması, gerçek zamanlı nesne tespiti alanında devrim niteliğinde bir yaklaşım sunarak, geleneksel böl-bul yöntemlerinin ötesine geçmiştir. İlk kez Joseph Redmon ve arkadaşları tarafından geliştirilen YOLOv1, 2016 yılında tanıtılmış ve hızlı bir şekilde popülerlik kazanmıştır (Redmon vd., 2016). Bu modelde görüntü, $S \times S$ boyutunda eşit grid hücrelerine bölünür ve her hücre belirli sayıda sınırlayıcı kutu ve sınıf olasılığı tahmini üretir. Bu yaklaşım sayesinde tespit işlemi tek bir evrede (tek aşamalı) gerçekleştirilir; böylece hem sınıf hem de koordinat tahmini aynı anda yapılır. YOLOv1, klasik R-CNN ve türevlerine göre çok daha hızlı çalışmasına karşın, özellikle küçük nesneleri tespit etme ve nesnelerin birbirine yakın olması durumunda başarımlarının düşüklüğü gibi sınırlamalara sahipti.



Şekil 26. YOLOv1 (Zhang, 2020)

YOLOv2 (YOLO9000 olarak da bilinir) ile birlikte model önemli yapısal iyileştirmeler geçirmiştir. Bu sürümde “batch normalization” gibi derin öğrenme mimarilerini stabilize eden teknikler uygulanmış, anchor box kullanımı sayesinde kutu tahminleri daha gerçekçi hale gelmiştir (Redmon & Farhadi, 2017). Ayrıca, giriş görüntü boyutu 416×416 'ya çıkarılarak doğruluk oranları artırılmıştır. YOLOv2 aynı zamanda WordTree yapısı kullanarak sınırlı

etiketli verilerle geniş sınıf kümesine karşı eğitim alabilmiş ve 9000 sınıfa kadar tespit yapabilme kabiliyeti kazanmıştır. Bu yönüyle, yalnızca hızlı değil aynı zamanda büyük veri kümesiyle etkili çalışabilen bir sistem haline gelmiştir.



Şekil 27. YOLOv2 mimarisi (Seong vd., 2019)

YOLOv3 sürümü ise önceki versiyonların hem hız hem de doğruluk açısından eksik kalan yönlerini geliştirmek amacıyla tasarlanmıştır. Bu modelde Darknet-53 adını taşıyan yeni bir backbone mimarisi benimsenmiştir. Bu yapı, 53 katmandan oluşan derin bir konvolüsyonel ağ olup residual bağlantılar içermektedir (Redmon & Farhadi, 2018). Bu sayede daha derin ve öğrenilebilir bir yapı ortaya çıkmıştır. YOLOv3 ayrıca çoklu ölçekli tespit özelliği sunar. Yani model, farklı boyutlardaki nesneleri daha doğru bir şekilde tanıyabilmek için üç farklı ölçekte (13×13, 26×26 ve 52×52) çıkış üretir. Bu mekanizma küçük nesnelerin de daha isabetli bir şekilde tespit edilmesini sağlamıştır.

YOLOv3, COCO veri setinde yapılan karşılaştırmalarda oldukça başarılı sonuçlar elde etmiş ve gerçek zamanlı uygulamalarda yüksek doğruluğu koruyarak kullanılabilirliğini artırmıştır. Bununla birlikte, YOLOv3 halen tek aşamalı yapısının doğası gereği bazı karmaşık sahnelerde, özellikle çok küçük ve çok sayıda nesnenin bulunduğu durumlarda çift aşamalı modellere göre daha düşük doğruluklar gösterebilmektedir.

Sonuç olarak, YOLOv1'den YOLOv3'e kadar olan gelişim süreci, modelin hızdan ödün vermeden daha doğru tespitler yapabilmesini mümkün kılmış, günümüzdeki YOLO sürümleri için sağlam bir temel oluşturmuştur. Bu sürümler, gerçek zamanlı nesne tespiti gerektiren akıllı

güvenlik sistemleri, otonom araçlar, endüstriyel denetim sistemleri gibi birçok uygulama alanında etkili biçimde kullanılmıştır (Bochkovskiy vd., 2020; Hanyao vd., 2021).

4.3.1.2. YOLOv4

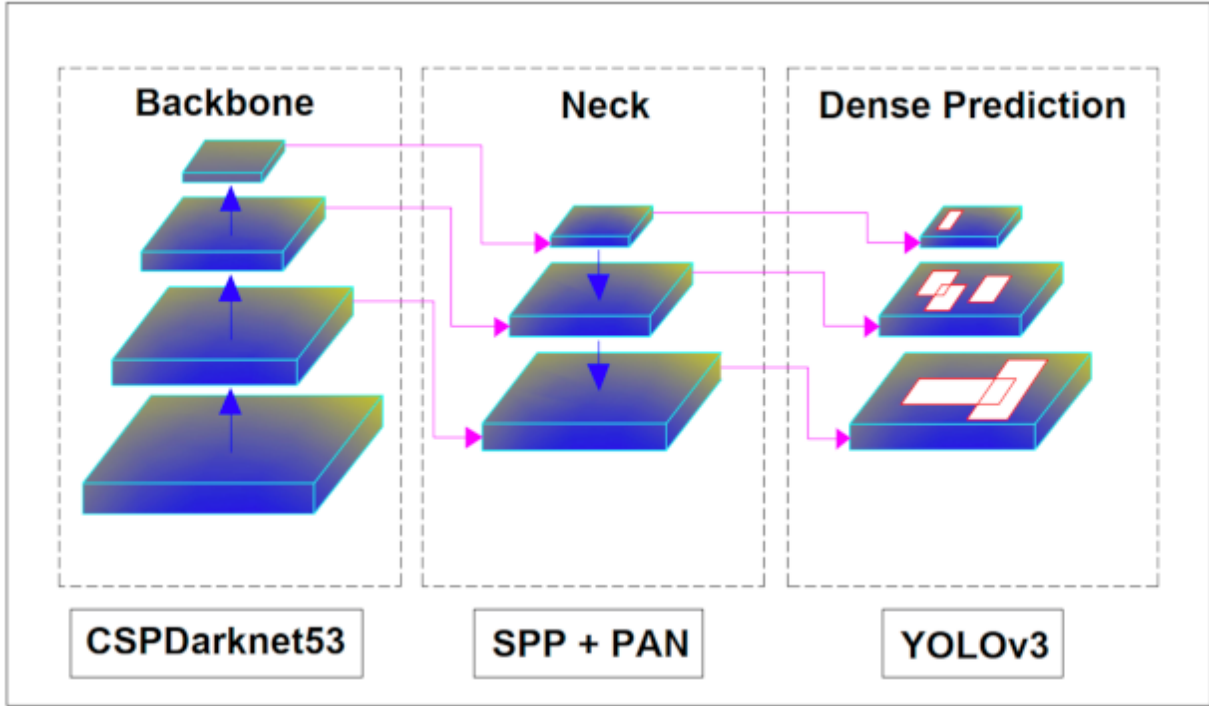
YOLOv4, gerçek zamanlı nesne tespiti alanında hız ve doğruluk arasında optimal bir denge sağlayan önemli bir geliştirme olarak 2020 yılında Alexey Bochkovskiy, Chien-Yao Wang ve Hong-Yuan Mark Liao tarafından tanıtılmıştır (Bochkovskiy vd., 2020). YOLOv4, selevi YOLOv3'ün temel prensiplerini korurken, birçok yeni bileşen ve optimizasyon tekniği ile performansını belirgin şekilde artırmıştır. Bu sürüm hem araştırmacılar hem de geliştiriciler için erişilebilir olacak şekilde tasarlanmış ve özellikle tek GPU ile eğitilebilecek kapasitede optimize edilmiştir.

YOLOv4'ün en dikkat çeken yeniliklerinden biri, Cross Stage Partial Network (CSPNet) mimarisinin backbone (omurga) olarak benimsenmesidir. CSPNet, daha önce kullanılan Darknet-53 yerine, CSPDarknet53 adı verilen yeni bir yapıyı entegre ederek, ağır gradyan akışını iyileştirmiş, parametre tekrarını azaltmış ve genel öğrenme yeteneğini artırmıştır (Wang vd., 2020). Bu sayede model daha derinlemesine bilgi öğrenebilirken aynı zamanda hesaplama yükünü azaltmayı başarmıştır.

YOLOv4 aynı zamanda iki temel teknik setiyle öne çıkmıştır: Bag of Freebies (BoF) ve Bag of Specials (BoS).

- **Bag of Freebies (BoF):** Modelin test zamanındaki hızını etkilemeden yalnızca eğitim sürecine yönelik yapılan optimizasyonları içerir. Bu bağlamda Mosaic data augmentation, DropBlock regularization, Class label smoothing ve CIOU loss gibi teknikler, modelin genelleme kabiliyetini artırmak amacıyla kullanılmıştır (Bochkovskiy vd., 2020). Özellikle Mosaic veri artırma yöntemi, dört farklı görüntünün tek bir karede birleştirilmesini sağlayarak çeşitliliği önemli ölçüde artırmış ve küçük nesne tespitinde kayda değer iyileşmeler sağlamıştır.
- **Bag of Specials (BoS):** Test zamanı performansını artıran fakat hesapsal maliyeti biraz yükselten modülleri kapsar. Bu kısımda Spatial Pyramid Pooling (SPP), Path Aggregation Network (PANet) ve Mish aktivasyon fonksiyonu gibi yapıların

entegrasyonu ile hem kontekst bilgisi zenginleştirilmiş hem de bilgi akışı daha verimli hale getirilmiştir (Lin vd., 2017; Misra, 2019).



Şekil 28. YOLOv4 mimarisi (Park vd., 2021)

YOLOv4, COCO ve Open Images gibi büyük veri kümelerinde yapılan karşılaştırmalarda önceki sürümlere kıyasla daha yüksek mAP değerleri elde etmiş ve özellikle gerçek zamanlı uygulamalarda dikkat çekici sonuçlar sunmuştur. Ayrıca, modelin açık kaynak kodlu olması, birçok farklı donanım ortamında kolayca uygulanabilir hale gelmesini sağlamıştır.

YOLOv4'ün başka bir önemli avantajı ise tek GPU üzerinde yüksek doğrulukla eğitilebilir olmasıdır. Bu özellik, özellikle akademik çalışmalar ve kaynakları sınırlı uygulama senaryoları için büyük bir kolaylık sağlamıştır. YOLOv4, derin öğrenmede “herkes için erişilebilir nesne tespiti” vizyonunu sürdürerek hem verimli hem de güçlü bir tespit mimarisi olarak kabul görmüştür.

Sonuç olarak YOLOv4, tek aşamalı nesne tespit algoritmalarının gelişiminde önemli bir dönüm noktası olmuş, yüksek doğruluk, eğitim verimliliği ve uygulama kolaylığı ile akademik literatürde ve endüstriyel projelerde yoğun şekilde kullanılmaya devam etmektedir (Zhao vd., 2019; Zhang vd., 2022).

4.3.1.3. YOLOv5

YOLOv5, resmi olarak YOLO ailesinin bir parçası olmasa da, Ultralytics tarafından PyTorch tabanlı olarak geliştirilmiştir. YOLOv5'in modüler yapısı, kullanıcı dostu arayüzü ve hızlı eğitim performansı sayesinde büyük popülerlik kazanmıştır (Jocher vd., 2020).

Nesne tespiti alanında çığır açan yöntemlerden biri olan "You Only Look Once" (YOLO) mimarisi, gerçek zamanlı tespit sistemlerinde hem hız hem de doğruluk açısından çığır açıcı bir performans sunmuştur. YOLO ailesinin 5. nesli olan YOLOv5, Haziran 2020'de Ultralytics tarafından duyurulmuş ve resmi bir akademik makaleye dayanmamasına rağmen, endüstri ve araştırma camiasında yaygın kabul görmüştür (Jocher vd., 2020). YOLOv5'in başarısı, kod tabanının açık kaynaklı olması, kolay eğitilebilirliği, yüksek doğruluğu ve gerçek zamanlı performansı ile doğrudan ilişkilidir.

4.3.1.3.1. YOLOv5'in mimari yapısı

YOLOv5, üç ana bileşenden oluşur: Backbone, Neck ve Head.

- Backbone kısmında CSPDarknet53 yapısı kullanılmaktadır. Bu yapı, Cross Stage Partial Network (CSPNet) mimarisine dayanmaktadır ve bilgi kaybını minimize ederek daha verimli bir özellik çıkarımı sağlamaktadır (Wang vd., 2021).
- Neck bölümünde ise Path Aggregation Network (PANet) kullanılarak çok katmanlı özellik haritaları daha etkin bir şekilde birleştirilmekte ve tespit doğruluğu artırılmaktadır.
- Head kısmı ise her bir nesnenin sınıfı, konumu (bounding box) ve nesneye ait güven skoru gibi çıktıları üretmektedir. Bu yapı sayesinde model hem küçük hem büyük nesneleri daha tutarlı bir şekilde tespit edebilmektedir (Zhao vd., 2022).

4.3.1.3.2. YOLOv5 sürümleri ve uygulama esnekliği

YOLOv5, kullanıcıya işlem gücüne ve tespit doğruluğu gereksinimlerine göre çeşitli versiyonlar sunar:

- **YOLOv5s (small):** Düşük işlem gücü gerektiren gerçek zamanlı uygulamalarda kullanılır.

- **YOLOv5m (medium):** Dengeli hız/doğruluk gerektiren uygulamalar içindir.
- **YOLOv5l (large):** Daha fazla parametre içerir ve doğruluk oranı yüksektir.
- **YOLOv5x (extra-large):** En yüksek doğruluğu sağlar, ancak hesaplama yükü fazladır.

Bu çeşitlilik sayesinde model, gömülü sistemlerden yüksek performanslı sunuculara kadar geniş bir yelpazede uygulanabilmektedir (Jocher vd., 2020).

4.3.1.3.3. YOLOv5'in eğitim süreci ve teknik özellikleri

YOLOv5'in öne çıkan avantajlarından biri kolay ve optimize edilmiş eğitim sürecidir.

- Otomatik hiperparametre optimizasyonu,
- Öğrenme oranı sıcaklık planlaması,
- Mosaic veri artırımı,
- MixUp, Label Smoothing gibi modern eğitim teknikleri desteklenmektedir (Zhang vd., 2023).

YOLOv5, PyTorch tabanlıdır ve bu durum, TensorFlow tabanlı önceki YOLO sürümlerine kıyasla daha esnek ve özelleştirilebilir bir yapı sunar. Model, aynı zamanda TensorRT, ONNX, CoreML gibi farklı formatlara dönüştürülerek mobil ve uç cihazlara kolayca entegre edilebilmektedir.

4.3.1.3.4. YOLOv5'in gerçek zamanlı uygulamaları

YOLOv5, yalnızca akademik çalışmalarla sınırlı kalmamış, aynı zamanda birçok endüstriyel uygulamada da başarıyla kullanılmaktadır.

- Güvenlik kameralarında şüpheli nesne takibi,
- Otonom araçlarda trafik işareti ve yaya tespiti,
- Tarımda meyve ve zararlı tespiti,
- İnşaat alanlarında KKD tespiti gibi birçok farklı senaryoda yüksek performans göstermektedir (Li vd., 2022).

Özellikle KKD tespitine yönelik yapılan çalışmalarda, YOLOv5 ile %92-96 arasında mAP değerleri elde edildiği, gerçek zamanlı uygulamalarda başarılı sonuçlar alındığı belirtilmiştir (He vd., 2021; Zhang vd., 2023).

4.3.1.3.5. YOLOv5'in avantajları ve sınırlılıkları

YOLOv5'in başlıca avantajları şunlardır:

- Gerçek zamanlı yüksek doğruluk,
- PyTorch tabanlı olması nedeniyle geniş geliştirici desteği,
- Transfer öğrenme yeteneği,
- Hızlı eğitim ve düşük donanım gereksinimi.

Ancak bazı sınırlamaları da bulunmaktadır:

- Çok küçük nesnelerde (örneğin uzaktaki baret gibi) tespit performansı düşebilir.
- Çok karmaşık arka planlı görüntülerde yanlış pozitif oranı artabilir.
- Düşük çözünürlüklü görüntülerde doğruluk kaybı yaşanabilir.

Bu nedenlerle YOLOv5'in uygulamaya özel detaylı optimizasyonlarla desteklenmesi gerekmektedir.

4.3.1.4. YOLOv6

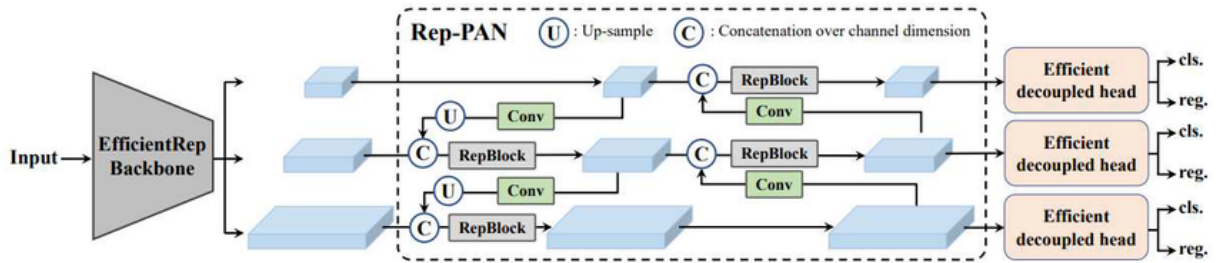
YOLOv6, 2022 yılında Çin merkezli teknoloji firması Meituan tarafından geliştirilmiş ve açık kaynaklı olarak yayımlanmıştır. YOLO mimarisinin önceki sürümlerine kıyasla hem endüstriyel uygulamalarda daha verimli çalışması hem de yüksek doğrulukla nesne tespiti gerçekleştirmesi hedeflenmiştir (Meituan, 2022). YOLOv6, özellikle gerçek zamanlı uygulamalarda kullanılan nesne algılama sistemleri için optimize edilmiştir ve yüksek verimlilik, düşük gecikme süresi ve daha fazla doğruluk oranı sunarak dikkat çekmiştir.

YOLOv6'nın en temel farkı, çapa bağımsız (anchor-free) bir tasarımı benimsemesidir. Bu tasarım sayesinde model, geleneksel anchor tabanlı yöntemlerde karşılaşılan kutu ön tanımı ve eşleştirme karmaşıklığından kurtulmuştur. Anchor-free yapılar, kutu koordinatlarını doğrudan

tahmin ederek, modelin daha esnek ve daha az hiperparametreye bağılı bir şekilde öğrenmesini sağlar (Zhou vd., 2019). Bu da hem eğitim süresini kısaltmakta hem de genel hata oranını azaltmaktadır.

YOLOv6 ayrıca Adaptive Training Sample Selection (ATSS) algoritmasını kullanarak pozitif ve negatif örneklerin seçimini otomatikleştirmekte ve daha dengeli bir eğitim süreci sunmaktadır (Zhang vd., 2020). ATSS, tespit edilen nesneye uygun olan örneklerin daha etkili seçilmesini sağlayarak sınıflandırma ve yerleştirme hatalarını azaltmakta büyük katkı sunmaktadır. Bu yöntem, sabit IoU eşiklerine dayalı klasik seçim algoritmalarına göre daha esnek ve öğrenmeye dayalı bir örnek seçimi sunar.

Backbone mimarisi olarak YOLOv6, Efficient Visual Architecture (EVA) adını verdikleri güçlü ve optimize edilmiş bir yapı kullanır. EVA hem hafif hem de yüksek doğruluklu tespit için tasarlanmış bir ağ yapısı olup, Meituan mühendislerinin özelleştirdiği ve PyTorch ortamında yüksek performans gösteren bir mimaridir. EVA ile birlikte gelen RepVGG block ve Depthwise Separable Convolutions gibi yapılar sayesinde model daha az parametre kullanarak daha derin öğrenme gerçekleştirir (Ding vd., 2021).



Şekil 29. YOLOv6 mimarisi (Tao vd., 2023)

YOLOv6 ayrıca aşağıdaki teknik yeniliklerle donatılmıştır:

- **SimOTA Label Assignment:** Gelişmiş bir örnek eşleme algoritması olup, başta Decoupled Head mimarisi olmak üzere modelin öğrenme sürecini daha sağlam hale getirir.
- **Efficient Decoupled Head:** Kayıp fonksiyonlarının sınıflandırma ve regresyon başlıkları arasında ayrıştırılması ile daha doğru sonuçlar alınmasını sağlar.

- **Quantization-aware Training (QAT)** desteđi ile düşük seviyeli donanımlarda bile yüksek dođruluk korunur. Bu özellik sayesinde YOLOv6, edge cihazlar ve mobil uygulamalar için de uygun hale gelmiştir.

YOLOv6, özellikle endüstriyel uygulamalarda nesne tespiti, kalite kontrol, trafik analizi ve robotik görüş sistemleri gibi alanlarda kullanılmak üzere optimize edilmiştir. Yapılan değerlendirmelerde, YOLOv5 ve YOLOv4 gibi seleflerine kıyasla hem FPS hem de mAP değerlerinde anlamlı iyileşmeler sunduđu gösterilmiştir (Meituan, 2022).

Sonuç olarak YOLOv6, nesne tespitinde anchor-free yaklaşımı, modern mimarisi ve endüstriyel uygulamalara uygunluđu ile literatürde önemli bir yer edinmiştir. Akademik araştırmalarda da giderek artan şekilde kullanılmakta ve özellikle düşük gecikme süresi ile yüksek dođruluđu bir arada sunması bakımından dikkat çekici bir mimaridir.

4.3.1.5. YOLOv7

YOLOv7 (Wang vd., 2022), E-ELAN blokları ve dynamic label assignment gibi tekniklerle daha hızlı ve dođru sonuçlar üretmektedir. YOLOv7 hem nesne tespiti hem de örnek segmentasyonu gibi görevlerde optimize edilmiştir.

YOLOv7, 2022 yılında Wong Kin-Yu, Alexey Bochkovskiy ve Chien-Yao Wang tarafından geliştirilen ve yayımlanan, nesne tespiti alanındaki en güncel ve ileri düzey YOLO sürümlerinden biridir. "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors" adlı çalışmayla duyurulan bu model, özellikle gerçek zamanlı nesne tespiti görevlerinde üstün performans sunmak üzere optimize edilmiştir (Wang vd., 2022).

YOLOv7 hem tespit dođruluđu hem de işlem süresi açısından önceki YOLO sürümlerini ve birçok çağdaş rakibini geride bırakmış; COCO veri kümesinde yüksek mAP ve FPS skorları ile dikkat çekmiştir.

4.3.1.5.1. YOLOv7'nin yenilikçi özellikleri

YOLOv7, sadece mimari güncellemelerle değil aynı zamanda eğitim sürecine entegre ettiği yenilikçi tekniklerle de dikkat çeker:

- **Trainable Bag-of-Freebies (TBoF):** Eğitim sürecinde doğrudan doğruluğu artırmak için parametresiz, ek maliyet getirmeyen ancak öğrenmeyi iyileştiren yöntemlerin (örneğin veri artırımı, DropBlock, EMA vb.) birlikte kullanılmasıyla elde edilen bir yapı.
- **Extended Efficient Layer Aggregation Network (E-ELAN):** Bu yapı, farklı derinliklerdeki özellik haritalarını daha etkili bir biçimde birleştirerek modelin genelleme kabiliyetini artırır. Aynı zamanda ağı derinliği artırıldığında eğitim kararlılığını korur.
- **Model Scaling for Concatenation-Based Models:** YOLOv7, model ölçeklemesini hem derinlik hem de genişlik boyutunda optimize eder. Bu, farklı boyutlardaki uygulamalar için çok yönlü kullanım imkânı sağlar.
- **Auxiliary Head:** Eğitim sırasında kullanılan bu yapı, geri yayılımı daha verimli hale getirerek daha hızlı yakınsama sağlar. Bu yardımcı başlık, sadece eğitim sırasında aktiftir ve çıkarım sırasında kullanılmaz.
- **Re-parameterized Convolution (RepConv):** Eğitim ve çıkarım aşamasında farklı yapıların kullanıldığı, eğitim sırasında çoklu dallanma sağlayan, çıkarımda ise tek konvolüsyonla sadeleştirilen bu yapı hem performansı hem de hızı artırır (Ding vd., 2021).

4.3.1.5.2. YOLOv7 mimarisinin teknik ayrıntıları

YOLOv7'nin mimarisi, backbone, neck ve head katmanlarını içerir. Aşağıda bu katmanlar özetlenmiştir:

- **Backbone:** E-ELAN modülünü barındıran yapı, temel özelliklerin çıkarımını yapar. CSPDarknet yapısına benzer, ancak gelişmiş bağlantı ve yoğunluk sağlar.
- **Neck:** PANet benzeri çok seviyeli özellik birleştirme mekanizmaları kullanılarak farklı boyutlardaki nesnelerin tespiti desteklenir.

- **Head:** YOLO kafa yapısı ile sınıf, konum ve nesne varlığı skorları çıkarılır. Bu çıkışlar, farklı ölçeklerde (small, medium, large objects) hesaplanır.

Ayrıca YOLOv7, Neural Architecture Search (NAS) teknikleriyle de uyumlu olacak şekilde tasarlanmıştır. Bu, modelin farklı platformlara otomatik optimize edilebilmesini kolaylaştırır.

4.3.1.5.3. YOLOv7'nin performans karşılaştırmaları

YOLOv7, birçok benchmark testinde state-of-the-art başarılar elde etmiştir. Örneğin: COCO test-dev üzerinde 51.4% AP değeri ve 160 FPS hız ile önceki sürümler ve EfficientDet, Faster R-CNN, RetinaNet gibi modellere karşı üstünlük sağlamıştır (Wang vd., 2022).

Bu özellikleri ile YOLOv7 hem kaynak kısıtlı cihazlar hem de yüksek doğruluk isteyen sunucu uygulamaları için ideal bir çözüm haline gelmiştir.

4.3.1.5.4. YOLOv7'nin gerçek dünya uygulamaları

YOLOv7, birçok pratik uygulamada kullanılmaktadır:

- Otonom sürüş sistemlerinde yol işaretleri ve yayaların tespiti,
- Endüstriyel kalite kontrolü,
- Kişisel koruyucu ekipman (baret, yelek) tespiti,
- Güvenlik ve izleme sistemlerinde anomali ve kişi takibi,
- Tarımda mahsul, zararlı veya hastalık tespiti (Li vd., 2023).

Modelin hem hızlı hem de yüksek doğrulukta sonuç vermesi, uç cihazlarda bile kullanılmasını mümkün kılmaktadır.

4.3.1.5.5. YOLOv7'nin avantajları ve sınırlılıkları

Avantajları:

- YOLOv5'e göre %10-20 daha yüksek doğruluk,
- Gerçek zamanlı tespit performansı (160+ FPS),
- Daha iyi küçük nesne tespiti başarımı,
- Esnek model ölçekleme ve eğitim yöntemleri.

Sınırlılıkları:

- Eğitim süreci daha karmaşıktır ve ileri düzey ayar gerektirir.

- YOLOv8 gibi daha yeni sürümler bazı durumlarda daha optimize çalışabilir.
- Veri seti çeşitliliği yüksek değilse overfitting riski vardır.

4.3.1.6. YOLOv8

YOLOv8 (Ultralytics, 2023), tamamen yeniden tasarlanmış bir mimari ile gelmiş, model yapısını daha hafif ve esnek hale getirmiştir. Özellikle ONNX ve TensorRT gibi platformlara doğrudan destek vermesi, edge computing uygulamaları için önemlidir. YOLOv8 aynı zamanda segmentasyon, sınıflandırma ve vücut duruşu tahmini görevlerini entegre destekleyen ilk YOLO versiyonudur.

YOLOv8, Ultralytics tarafından 2023 yılında geliştirilen ve tanıtılan, YOLO serisinin en güncel ve gelişmiş sürümüdür. Daha önceki versiyonlardan bağımsız olarak sıfırdan inşa edilen bu sürüm, segmentasyon, sınıflandırma ve insan pozu tahmini gibi görevleri desteklemekte ve kullanım kolaylığı sağlamaktadır (Ultralytics, 2023).

4.3.1.6.1. YOLOv8'in mimari yapısı

YOLOv8'in mimarisi, önceki YOLO sürümlerine göre daha hafif ve daha verimli olacak şekilde tasarlanmıştır. Bu sürümde önceki bazı mimari bileşenler çıkarılmış veya modern yaklaşımlarla değiştirilmiştir. Örneğin:

- Cross Stage Partial (CSP) yapısı yerine daha sade bir mimari olan Convolution-BatchNorm-LeakyReLU (CBL) ve Conv-BN-SiLU kombinasyonları tercih edilmiştir.
- Anchor-free bir nesne tespiti yaklaşımı benimsenmiştir. Bu sayede özellikle küçük nesnelerin daha doğru bir şekilde tespit edilmesi sağlanmış ve modelin genel doğruluğu artırılmıştır.
- Decoupled Head kullanılarak sınıflandırma ve regresyon görevleri birbirinden ayrılmıştır. Bu mimari değişiklik, özellikle daha kararlı bir öğrenme süreci ve daha iyi genel performans sağlamıştır (Ultralytics, 2023).

4.3.1.6.2. YOLOv8'in özellikleri ve avantajları

- **Modüler Yapı:** YOLOv8, kullanıcı dostu modüler bir yapıya sahiptir. Bu durum, araştırmacıların model üzerinde değişiklik yapmasını, transfer öğrenme uygulamasını veya yeni veri setleriyle kolayca eğitmesini mümkün kılar.
- **Performans:** COCO gibi benchmark veri setleri üzerinde yüksek AP değerlerine ulaşarak önceki sürümleri geride bırakmıştır. Örneğin, YOLOv8x modeli COCO veri setinde yaklaşık %53 AP'ye ulaşmıştır (Jocher vd., 2023).
- **Segmentasyon ve Pose Estimation:** YOLOv8 yalnızca nesne tespitiyle sınırlı değildir; aynı zamanda anlamsal ve örnek segmentasyonu, vücut duruşu tahmini gibi görevlerde de kullanılabilir.
- **Optimize Edilmiş Çıktılar:** ONNX, CoreML, TensorRT gibi birçok farklı platforma kolaylıkla dönüştürülebilir, bu da edge cihazlarda çalışmayı kolaylaştırır.
- **Yüksek Uyumluluk:** Ultralytics'in sağladığı Command Line Interface (CLI) ve Python API desteği ile modelin eğitilmesi ve kullanımı oldukça basitleştirilmiştir.

4.3.1.6.3. YOLOv8'in uygulama alanları

YOLOv8, endüstriyel otomasyon, güvenlik sistemleri, akıllı şehirler, sağlık teknolojileri ve otonom araçlar gibi birçok alanda uygulanabilmektedir. Özellikle iş sağlığı ve güvenliği kapsamında KKD tespiti için ideal bir tercih sunmaktadır. YOLOv8'in anchor-free yapısı sayesinde işçilerin farklı uzaklık ve açılardaki görüntülerinde daha güvenilir tespit yapılabilmektedir (Guo vd., 2023).

4.3.1.6.4. YOLOv8'in dezavantajları ve sınırlılıkları

- YOLOv8'in anchor-free yapısı, belirli durumlarda yüksek oranda örtüşen nesneleri ayırt etmekte zorlanabilir.
- Daha fazla özellik sunması, bazı modellerin boyutunun artmasına neden olabilir. Bu da gömülü sistemlerde modelin küçültülmesini gerekli kılar.
- YOLOv8 ile birlikte gelen bazı özel formatlar (örneğin Ultralytics'in kendi .yaml ve .pt formatları), diğer açık kaynaklı platformlarla doğrudan uyumsuzluk yaratabilir.

4.3.1.6.5. YOLOv8'in avantajları

YOLOv8, YOLO ailesinin en gelişmiş üyesi olarak hem doğruluk hem de kullanım kolaylığı açısından kullanıcı dostu ve güçlü bir modeldir. Modern yapısı, anchor-free tasarımı, segmentasyon ve pose tahmini gibi çok yönlü destekleriyle, nesne tespitinde çok sayıda endüstriyel ve akademik uygulama için öncelikli tercih haline gelmiştir. Bu çalışmada kullanılan YOLOv8, özellikle KKD tespiti gibi kritik güvenlik uygulamalarında yüksek doğruluğu ve gerçek zamanlı performansı ile tercih edilmiştir.

YOLOv8, COCO ve custom veri setlerinde çok yüksek mAP değerlerine ulaşmıştır. Kullanımı kolay arayüzü ve transfer learning desteği, onu birçok akademik ve ticari projede tercih edilen model haline getirmiştir.

4.3.2. Single shot multiBox detector (SSD)

SSD, her bir feature map konumundan doğrudan sınırlayıcı kutu tahmini yapar. Bu yönüyle anchor box tabanlı ilk tek aşamalı modellerdendir (Liu vd., 2016). SSD, farklı boyuttaki feature map'lerde farklı ölçeklerde nesne tespiti yaparak hem küçük hem büyük nesneler için esnek çözüm sunar.

SSD'nin avantajları arasında:

- Gerçek zamanlı çalışabilme,
- Daha az karmaşık yapı,
- YOLOv1'e göre daha yüksek doğruluk yer alır.

Modern Geliştirmeler: 2020 sonrası, SSD'ye dayalı olarak mobil cihazlar için Lightweight SSD ve Tiny SSD gibi sürümler geliştirilmiştir. Bunlar MobileNet, EfficientNet gibi hafif backbone'lar ile entegre çalışmaktadır (Zhao vd., 2022).

4.3.3. RetinaNet

RetinaNet (Lin vd., 2017), focal loss fonksiyonunu tanıtarak tek aşamalı modellerde sık karşılaşılan sınıf dengesizliği sorununa çözüm getirmiştir. Bu model, FPN yapısı üzerine kuruludur ve düşük örnek frekansına sahip nesnelerin daha iyi tanınmasına olanak sağlar.

Avantajları:

- Yüksek doğruluk
- Tek aşamalı tespitlerde mAP başarımı
- Karmaşık sahnelerde iyi performans

Modern Kullanım: RetinaNet, tıbbi görüntüleme ve güvenlik sistemlerinde sık tercih edilen bir yöntemdir. EfficientDet gibi yeni mimarilerde de focal loss ve FPN yapısı kullanılmaya devam etmektedir.

4.3.4. Güncel tek aşamalı modeller ve gelişmeler**4.3.4.1. EfficientDet**

EfficientDet, EfficientNet tabanlı backbone ve Bidirectional Feature Pyramid Network (BiFPN) yapısıyla hem hız hem de doğruluk açısından dengeli bir yapı sunar (Tan vd., 2020). Model büyüklükleri D0'dan D7'ye kadar ölçeklenebilir.

4.3.4.2 PP-YOLOE

PP-YOLOE, Baidu tarafından geliştirilen PaddlePaddle framework'ü üzerinden yürütülen bir projedir. ResNet-DCN ve ESE attention mekanizması kullanarak tespit doğruluğunu artırmıştır (Xu vd., 2022).

4.4. Derin Öğrenme Algoritmalarında Tespit İşleminin Hızlandırılması

Derin öğrenme tabanlı nesne tespit sistemleri, yüksek doğruluk ve esneklik sağlamalarına rağmen, özellikle gerçek zamanlı uygulamalarda yavaş çalışabilmektedir. Bu nedenle, tespit algoritmalarının hızlandırılması kritik öneme sahiptir. Hızlandırma; mimari düzeyde iyileştirme, model sıkıştırma, kuantizasyon, edge cihazlara uygunlaştırma ve donanım odaklı optimizasyonlarla mümkündür.

4.4.1. Hafif mimari kullanımı

Geleneksel ağlar (VGG, ResNet) yerine daha az parametreye sahip, düşük bellek tüketen mimariler kullanmak tespit süresini ciddi ölçüde azaltır:

- **MobileNetV2/V3:** Özellikle mobil ve gömülü cihazlarda kullanılan hafif CNN mimarileridir (Howard vd., 2019). SSD, YOLOv5 ve YOLOv8 gibi tespit modelleri bu tür backbonelarla entegre edilmiştir.
- **EfficientNet:** Parametre sayısını optimize ederek daha az hesaplama gücüyle yüksek doğruluk sağlar (Tan & Le, 2019).

4.4.2. Model sıkıştırma ve pruning

Ağ parametrelerinin optimize edilmesi, gereksiz düğümlerin budanması (pruning), seyreltilmiş ağ yapıları sayesinde modelin hem boyutu hem de hesaplama yükü azaltılabilir (Molchanov vd., 2017). YOLOv5'in "nano" ve "small" versiyonları, budama ve katman azaltımı yoluyla düşük kaynaklı ortamlara uygun hale getirilmiştir.

4.4.3. Kuantizasyon ve düşük hassasiyetli hesaplama

Model parametrelerinin 32-bit kayan nokta (FP32) yerine 16-bit (FP16) ya da 8-bit (INT8) ile temsil edilmesi, performansı önemli ölçüde artırabilir. TensorRT ve ONNX Runtime gibi kütüphaneler bu tür optimizasyonları destekler.

4.4.4. Donanım tabanlı hızlandırma

GPU, TPU, NPU gibi özel işlemciler kullanılarak paralel hesaplama desteklenir. Özellikle YOLOv8 gibi modellerin TensorRT desteği ile edge cihazlarda dahi yüksek FPS değerleri elde edilebilir.

- **Edge cihazlar:** NVIDIA Jetson, Google Coral, Raspberry Pi gibi platformlar düşük güç tüketimiyle gerçek zamanlı nesne tespiti sağlar (Hanyao vd., 2021).

4.4.5. Mimari iyileřtirmeler

- Depthwise Separable Convolution (DSConv) gibi teknikler, parametre sayısını azaltarak hızlandırma sağlar.
- Attention mekanizmaları (SE blokları, ECA-Net) seçici bilgi işleme ile hem doğruluęu artırır hem de gereksiz hesaplamayı azaltır (Wang vd., 2020).



5. MATERYAL VE YÖNTEM

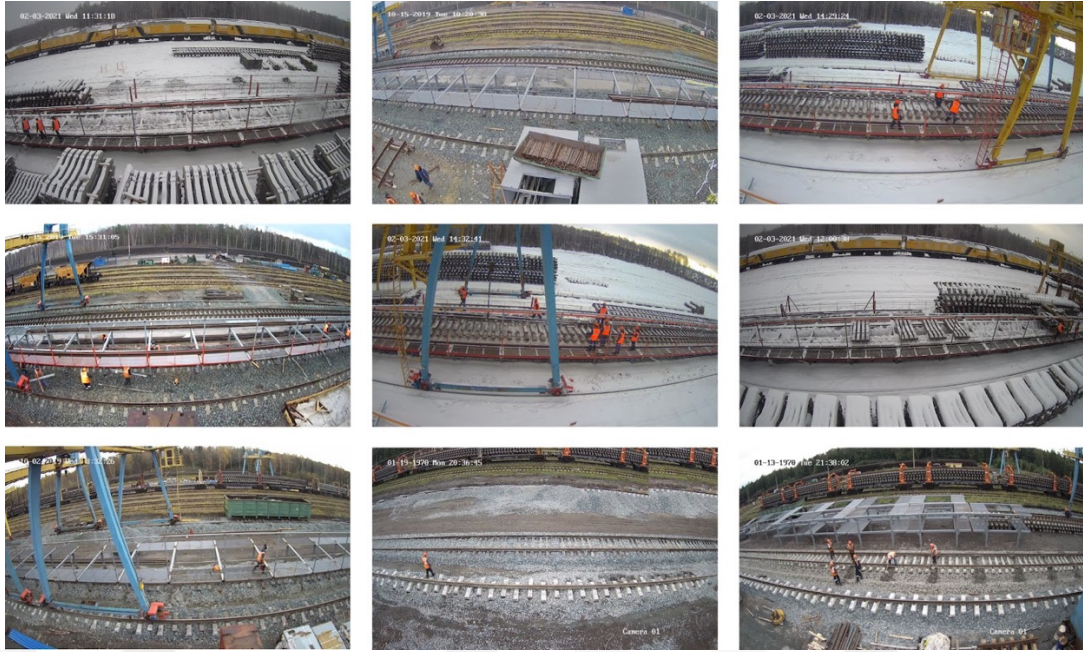
5.1. Veri Seti ve Önışleme

Bu alıřmanın temelini, iř saėlıėı ve ėüvenliėi kapsamında KKD kullanım durumunun otomatik olarak tespit edilmesini amalayan bir bilgisayarla ėörme uygulaması oluřturmaktadır. Bu doėrultuda, makine ėrenmesi ve zellikle derin ėrenme yöntemlerinin uygulanabilirliėi aısından zengin ieriėe ve etiketleme yapısına sahip aık kaynaklı bir veri seti olan Railroad Worker Detection Dataset (Mazurov, 2021) kullanılmıřtır.

5.1.1 Veri setinin tanıtımı ve genel yapısı

Veri seti, Kaggle platformu üzerinde Mike Mazurov (2021) tarafından yayımlanmıřtır. Veri seti temel olarak demiryolu řantiyelerinde ėörev yapan iřilerin farklı pozisyonlarda ve aılarda ekilmiş ėörüntülerini iermekte olup, iřilerin baret (helmet), ėüvenlik yeleėi (vest) ve iři (worker) sınıflarına ait etiketlerini barındırmaktadır. řekil 30’da veri setine ait rnek ėörseller verilmiřtir.

Bu veri seti, ėerek dnya sahnelerinden elde edilmiř olması bakımından sentetik veya laboratuvar ortamlarında oluřturulmuř veri setlerinden ayrılmaktadır. Bu da modelin ėerek zamanlı uygulamalarda genelleme kabiliyetini artırmaktadır. Veri setindeki ėörsellerin eřitliliėi — farklı kamera aılarından, ıřık kořullarından ve sahne dzenlerinden elde edilmesi — derin ėrenme modellerinin performansını test etmek aısından oldukça elveriřlidir (Zhou vd., 2020).



Şekil 30. Veri seti örnek görseller (Mazurov, 2021)

Veri setinin temel özellikleri:

- **Görüntü sayısı:** 3.222 adet
- **Etiket sınıfları:** Baret (helmet), yelek (vest), işçi (worker)
- **Dosya biçimi:** JPEG formatında görseller, TXT formatında YOLO uyumlu etiket dosyaları
- **Görüntü boyutları:** Değişken; en yaygın boyutlar 1280×720, 640×480
- **Etiket formatı:** YOLO etiketi biçiminde: class_id x_center y_center width height

Baret yelek ve işçi görüntülerinin yer aldığı 3222 görüntüde her bir sınıfa ait etiket sayıları tablo 1’de gösterilmiştir.

Tablo 1. Etiketli nesne sayısı

Etiket Sınıfı	Etiket Sayısı
Baret	6515
Yelek	7883
İşçi	7973
Toplam	22371

5.1.2 Etiketleme yapısı ve kalitesi

Veri seti, YOLO formatına uygun olarak her görüntüye karşılık gelen metin dosyalarında nesnelerin sınıfı ve konum bilgilerini içerecek şekilde düzenlenmiştir. Koordinatlar, görüntü boyutlarına göre normalize edilmiş olup, her satırda bir nesne tanımı yer almaktadır. Ancak veri seti üzerinde yapılan manuel kontrollerde bazı eksik etiketlemeler, çakışmalar ve sınıf dengesizlikleri tespit edilmiştir. Bu durum, derin öğrenme modelinin eğitimi öncesinde veri temizleme ve düzeltme işlemleri ile giderilmeye çalışılmıştır.

Benzer şekilde veri setindeki sınıflar arasında gözle görülür bir dengesizlik olduğu belirlenmiştir. Örneğin, işçi sınıfına ait nesnelerin sayısı baret veya yelek sınıflarına kıyasla daha fazladır. Bu durum sınıf dengesizliği sorununa neden olabileceğinden, eğitim aşamasında özel kayıp fonksiyonları ve ağırlıklı örnekleme teknikleri dikkate alınmıştır (Johnson ve Khoshgoftaar, 2019).

5.1.3 Önişleme süreci ve teknikleri

Veri setinin model eğitime uygun hale getirilebilmesi için aşağıda açıklanan önişleme adımları gerçekleştirilmiştir:

5.1.3.1 Boyutlandırma (Resizing)

Tüm görüntüler, YOLOv5, YOLOv7 ve YOLOv8 modelleri için önerilen 640×640 piksel boyutuna yeniden boyutlandırılmıştır. Bu boyut hem eğitim süresini optimize etmekte hem de bellek tüketimini dengeli tutmaktadır (Jocher vd., 2020).

5.1.3.2 Veri artırımı (Data augmentation)

Modelin farklı durumları genelleme kabiliyetini artırmak amacıyla eğitim verileri üzerinde YOLO algoritmalarının varsayılan değerleri esas alınarak şu dönüşümler uygulanmıştır:

- Yatay çevirme (horizontal flip)
- Mosaic (görüntü birleştirme)
- HSV renk dönüşümü
- Rastgele parlaklık/arttırma (brightness adjustment)

- Dönme (rotation)

Bu işlemler, modelin farklı kamera açıları, ışık koşulları ve çevresel faktörlere karşı daha dayanıklı hale gelmesini sağlamıştır (Shorten ve Khoshgoftaar, 2019).

5.1.3.3 Etiket uyumlaştırması ve normalize etme

Tüm etiketler, görüntü boyutlarına göre yeniden normalize edilmiştir. Etiketleme formatı şu biçimdedir: class_id x_center y_center width height. Bu yapı, YOLO algoritmalarının doğrudan uyumlu biçimde eğitilmesine olanak tanır.

5.1.4 Veri setinin eğitim-doğrulama-test ayrımı

Modelin başarısının tarafsız şekilde ölçülebilmesi için veri seti üç parçaya ayrılmıştır:

- **Eğitim verisi (train):** %80 (2577 görüntü)
- **Test verisi (test):** %20 (645 görüntü)

5.2. Yöntem

YOLO (You Only Look Once) algoritması, nesne tespitinde tek aşamalı bir yaklaşım sunarak diğer yöntemlere kıyasla önemli avantajlar sağlamaktadır. Özellikle R-CNN, Fast R-CNN ve Faster R-CNN gibi iki aşamalı modeller, yüksek doğruluk oranlarına ulaşabilseler de, bölge öneri adımları ve katmanlı işleme süreçleri nedeniyle gerçek zamanlı uygulamalarda yetersiz kalabilmektedir (Ren et al., 2015). Buna karşın YOLO, görüntüyü tek bir ileri besleme adımıyla (single forward pass) işleyerek tüm nesne tespiti sürecini son derece hızlı bir şekilde tamamlar. Bu yapı, hem sınıflandırma hem de konumlandırma işlemlerinin aynı ağ üzerinden eşzamanlı olarak gerçekleştirilmesine olanak tanır (Redmon et al., 2016). Ayrıca YOLO'nun daha yeni sürümleri (YOLOv5, YOLOv7, YOLOv8 gibi) hem tespit doğruluğunu hem de işlem hızını artıracak şekilde optimize edilmiştir. Bu sürümler, daha verimli backbone mimarileri, veri artırma teknikleri ve gelişmiş kayıp fonksiyonları sayesinde hem düşük gecikme süresi hem de yüksek hassasiyet sağlar (Jocher et al., 2020; Wang et al., 2022). Özellikle endüstriyel uygulamalarda ve gerçek zamanlı güvenlik izleme sistemlerinde hızın kritik bir öneme sahip olması, bu çalışmada YOLO algoritmasının tercih edilmesini gerekli kılmıştır.

İş güvenliğinde KKD ekipmanlarının tespit edilmesi için YOLO versiyonlarından en güncel sürümler olan YOLOv5, YOLOv7 ve YOLOv8 tercih edilmiştir. YOLOv5, YOLOv7 ve YOLOv8 sürümleri kendi içerisinde farklı ağırlık dosyaları barındırmaktadır (Tablo 2).

Tablo 2. YOLO sürümleri ağırlık dosyaları

Model	Ağırlık Dosyası	Açıklama
YOLOv5	yolov5s.pt	Küçük model, hızlı ve hafif, düşük doğruluk
	yolov5m.pt	Orta model, hız ve doğruluk dengesi
	yolov5l.pt	Büyük model, daha yüksek doğruluk
	yolov5x.pt	Çok büyük, en yüksek doğruluk
YOLOv7	yolov7.pt	Standart YOLOv7 modeli, genel amaçlı
	yolov7-tiny.pt	Daha küçük ve hızlı, düşük doğruluk
	yolov7-w6.pt	Daha büyük, daha yüksek doğruluk için
YOLOv8	yolov8n.pt	Nano model, çok hızlı ve hafif
	yolov8s.pt	Küçük model, iyi hız ve doğruluk dengesi
	yolov8m.pt	Orta model, daha yüksek doğruluk
	yolov8l.pt	Büyük model, yüksek doğruluk
	yolov8x.pt	Çok büyük model, en yüksek doğruluk

Tablo 2’de verilen YOLO versiyonlarının ağırlık dosyalarından, kullanılan veri setinin küçük olması nedeniyle S sürümü seçilmiştir.

Çalışmada ele alınan YOLO versiyonları kullanılarak Railroad Worker Detection Dataset veri seti üzerinde, aynı ortam ve şartlarda eğitimler gerçekleştirilmiştir. Eğitimler Nvidia geforce 1070TI ekran kartı, Intel Core İ7 7700HQ işlemci ve 16GB RAM’e sahip bir bilgisayarda, Python programlama dili kullanılarak yürütülmüştür. YOLO versiyonlarının eğitimleri sırasında kullanılan parametreler Tablo 3’te verilmiştir.

Tablo 3. Eğitim sırasında kullanılan parametreler

Parametre	Değer	Açıklama
Kullanılan Modeller	YOLOv5s, YOLOv7s, YOLOv8s	Her üç modelin küçük (small - s) versiyonları kullanılmıştır
Epoch Sayısı	50	Modeller 50 dönem (epoch) boyunca eğitilmiştir
Batch Size	16	Her iterasyonda eğitime giren görüntü sayısı
Görüntü Boyutu	640 × 640	Giriş görüntülerinin yeniden boyutlandırıldığı sabit çözünürlük
Optimizasyon Algoritması	SGD	YOLO modellerinde genellikle varsayılan olarak kullanılan optimizasyon algoritması
Veri Artırımı	Aktif (varsayılan)	Flip, scale, mosaic, HSV gibi görüntü artırma teknikleri kullanılmıştır
Loss Fonksiyonları	Box, Cls, DFL	Kayıp fonksiyonları: kutu regresyonu (box), sınıf tahmini (cls), dağılım odaklı kayıp (df)
Önceden Eğitilmiş Ağırlıklar	yolov5s.pt, yolov7.pt, yolov8s.pt	Her model için kendi küçük ağırlık dosyası kullanılmıştır
Eğitim-Test Oranı	%80 / %20	Eğitim ve doğrulama için veri ayrımı

Tablo 3'te görüldüğü üzere epoch 50, batch size 16, image size 640 olarak belirlenmiştir. Tabloda verilen diğer parametreler modellerin varsayılan değerlerinde alınmıştır.

5.2.1 Ölçüm metrikleri

Nesne tespiti modellerinin başarımını değerlendirmek amacıyla çeşitli ölçüm metrikleri kullanılmaktadır. Bu metrikler; modelin doğruluğunu, duyarlılığını ve genel performansını nicel olarak ifade etmeye yardımcı olur. Bu çalışmada kullanılan YOLOv5, YOLOv7 ve YOLOv8 modellerinin karşılaştırılmasında yaygın olarak kullanılan Precision, Recall, F1-Score, Intersection over Union (IoU) ve mean Average Precision (mAP) metrikleri esas alınmıştır.

5.2.1.1. Intersection over union (IoU)

IoU, modelin tahmin ettiği sınırlayıcı kutu (bounding box) ile gerçek kutu (ground truth) arasındaki kesişim oranını ölçer. Formülü denklem 1’de verilmiştir.

$$IoU = \frac{A_{intersection}}{A_{union}} \quad (1)$$

Burada:

- $A_{intersection}$: Tahmin kutusu ile gerçek kutunun kesişim alanı
- A_{union} : Tahmin kutusu ile gerçek kutunun birleşim alanı

IoU değeri 0 ile 1 arasında değişir. 0, tamamen yanlış tespiti; 1 ise mükemmel örtüşmeyi ifade eder. Genellikle IoU eşik değeri olarak %50 (0.5) kullanılır.

5.2.1.2. Kesinlik (Precision)

Precision, modelin pozitif olarak sınıflandırdığı nesnelerin ne kadarının gerçekten doğru olduğunu ölçer. Formülü denklem 2’de verilmiştir.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Burada:

- **True Positive (TP)**: Doğru olarak tespit edilen nesneler
- **False Positive (FP)**: Yanlış olarak tespit edilen nesneler

5.2.1.3. Duyarlılık (Recall)

Recall, modelin tespit etmesi gereken tüm gerçek nesnelerden kaç tanesini doğru şekilde tespit ettiğini gösterir. Formülü denklem 3’te verilmiştir.

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

Burada:

- **False Negative (FN):** Gerçek nesne olup da model tarafından tespit edilemeyen durumlar

5.2.1.4. F1-score

Precision ve Recall'un harmonik ortalamasıdır. Dengeleyici bir metriktir, özellikle Precision ve Recall arasında bir denge kurmak gerektiğinde önemlidir. Formülü denklem 4'te verilmiştir.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

5.2.1.5. Mean average precision (mAP)

mAP, birden fazla sınıf için ortalama olarak modelin ne kadar başarılı olduğunu gösterir. Bir sınıf için Average Precision (AP) hesaplanır ve tüm sınıfların ortalaması alınarak mAP elde edilir. Formülü denklem 5'te verilmiştir. IoU eşik değerine göre farklı mAP hesaplamaları yapılabilir. Örneğin:

- $mAP_{@0.5}$: IoU ≥ 0.5 için
- $mAP_{@0.5:0.95}$: IoU 0.5'ten 0.95'e kadar her 0.05'lik artışla hesaplanan ortalama

$$AP = \int_0^1 p(r) dr \quad (5)$$

Burada:

- $p(r)$: Recall'a göre değişen precision değeridir.

mAP değeri, genellikle grafiksel olarak Precision-Recall eğrisi altında kalan alanın (AUC) hesaplanmasıyla elde edilir.

5.2.2. Çalışmada kullanılan YOLO sürümlerinin karşılaştırılması

Son yıllarda bilgisayarla görme alanında, özellikle gerçek zamanlı nesne tespiti uygulamalarında YOLO algoritmaları oldukça yaygın olarak kullanılmaktadır. YOLO'nun her yeni sürümü, bir önceki versiyona göre gerek doğruluk oranı gerekse de işlem hızı açısından önemli gelişmeler sunmaktadır. Bu bağlamda, YOLOv5, YOLOv7 ve YOLOv8 sürümleri;

mimari yapı, eğitim kolaylığı, görev çeşitliliği ve performans açısından değerlendirilerek Tablo 4'te kapsamlı bir biçimde karşılaştırılmıştır.

Tablo 4. YOLOv5, YOLOv7 ve YOLOv8 sürümlerinin karşılaştırması

Özellik	YOLOv5	YOLOv7	YOLOv8
Yayın Tarihi	Haziran 2020	Temmuz 2022	Ocak 2023
Model Türleri	v5s, v5m, v5l, v5x	E6, E6E, D6, X	v8n, v8s, v8m, v8l, v8x
Mimari Yapı	CSPDarknet + PANet	E-ELAN + RepConv	Anchor-free, decoupled head
Anchor Yapısı	Anchor-based	Anchor-based	Anchor-free
Çıkış Katmanı	Coupled head	Auxiliary head	Decoupled head
Desteklenen Görevler	Detection	Detection	Detection, segmentation, classification, pose
Segmentasyon/Pose Estimation	Harici modül ile	Desteklenmiyor	Entegre olarak mevcut
Kullanım Kolaylığı	CLI & Python API	Derin öğrenme bilgisi gerektirir	Geliştirilmiş CLI ve Python API
Model Eğitimi	Kolay	Orta	Çok kolay (modüler yapı)
Donanım Uyumluluğu	Yüksek (Edge GPU'lar)	Orta-üst	Yüksek, ONNX/RT/CoreML destekli
Performans (COCO mAP)	~50 (v5x)	~52 (E6E)	~53 (v8x)
Hız (FPS)	Çok hızlı	Hızlı	Hızlı
Model Boyutu	Küçükten büyüğe seçenekli	Daha büyük	Optimize edilmiş, daha esnek
Kodlama Desteği	PyTorch, TorchScript, ONNX	PyTorch	PyTorch, ONNX, CoreML, TensorRT
Öne Çıkan Özellikler	Hafif ve esnek	Performans odaklı yapı	Çok görevli yapı, anchor-free mimari
Zorluklar	Mimari olarak eski	Daha karışık yapı	Yeni formatlar ve uyumluluk sorunları

YOLOv5, genellikle kullanıcı dostu arayüzü ve PyTorch tabanlı yapısıyla eğitim ve uygulama kolaylığı sunması nedeniyle en yaygın kullanılan sürümlerden biri olmuştur (Jocher vd., 2020). Özellikle daha küçük modelleri (v5s, v5m) edge cihazlar üzerinde başarılı şekilde çalıştırılabilirken, daha büyük versiyonları (v5x) yüksek doğruluk oranları elde etmede etkili olmuştur.

YOLOv7, daha karmaşık bir mimariye sahiptir ve yüksek performans sunmak amacıyla birçok yenilikçi yaklaşımı entegre etmiştir. Özellikle "RepConv" ve "Coarse-to-Fine Lead Head" gibi bileşenler, tespit doğruluğunu artırmada önemli rol oynamıştır (Wang vd., 2022). Bununla birlikte, eğitim süreci YOLOv5'e kıyasla daha fazla uzmanlık gerektirir.

YOLOv8, Ultralytics tarafından yayınlanan en yeni ve gelişmiş sürümdür. Bu sürümde anchor-free yapı benimsenmiş, decoupled head kullanılarak sınıflandırma ve regresyon işlemleri ayrılmıştır. Ayrıca segmentasyon, sınıflandırma ve insan pozu tespiti gibi görevler için de doğal destek sunmaktadır. Bu sayede, çok görevli projelerde modüler ve güçlü bir çözüm haline gelmiştir (Ultralytics, 2023). Eğitim süreci oldukça kullanıcı dostudur ve CLI üzerinden rahatlıkla uygulanabilmektedir.

Sonuç olarak, model seçimi yapılırken projenin ihtiyaçları, donanım kapasitesi ve istenen doğruluk-hız dengesi göz önünde bulundurulmalıdır. YOLOv8, günümüzde en kapsamlı ve gelişmiş seçenek olarak öne çıksa da YOLOv5'in basit yapısı ve YOLOv7'nin doğruluk potansiyeli, belirli senaryolarda halen etkili çözümler sunmaktadır.

6. BULGULAR

Bu bölümde, çalışmada kullanılan YOLO mimarisinin üç farklı versiyonu olan YOLOv5, YOLOv7 ve YOLOv8 algoritmalarının eğitim ve doğrulama süreçlerinden elde edilen performans bulguları ayrıntılı olarak sunulmaktadır. Her bir model, aynı veri kümesi üzerinde eğitilmiştir ve benzer hiperparametre ayarları kullanılmıştır. Böylece modellerin göreceli başarımları adil ve doğrudan karşılaştırılabilir hale getirilmiştir. Bulgular hem grafiksel çıktılar hem de sayısal metrikler üzerinden yorumlanmış; her bir mimarinin güçlü ve zayıf yönleri teknik olarak analiz edilmiştir.

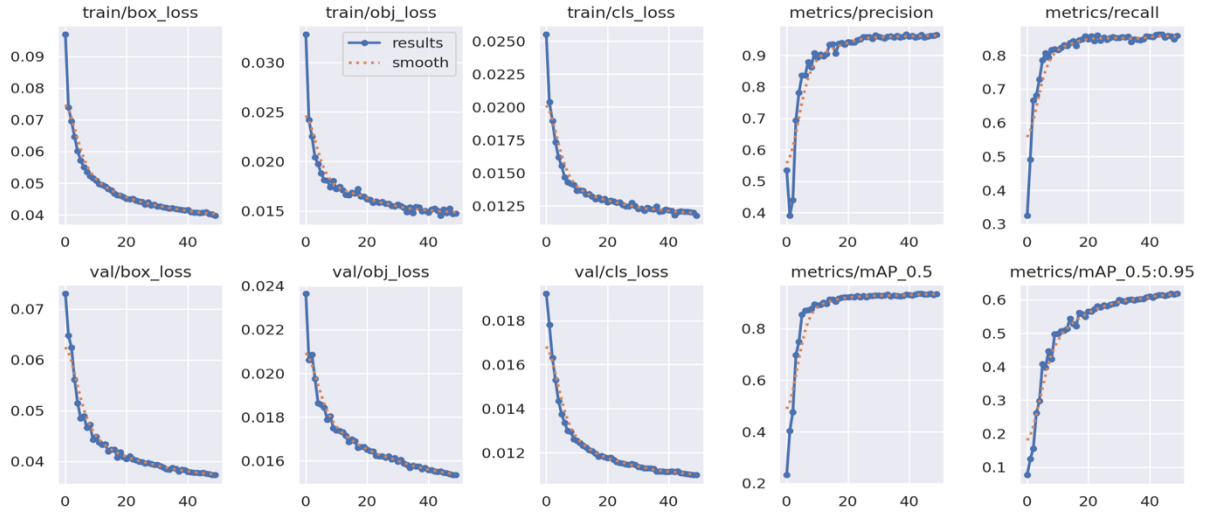
6.1. YOLOv5 Modeline Ait Bulgular

Bu çalışma kapsamında ilgili YOLOv5 modeli üç kez eğitime tabi tutulmuş ve her bir eğitim sonucunda Tablo 5’de sunulan değerlendirme metrikleri elde edilmiştir. Modeller, precision, recall, mAP@0.5 ve mAP@0.5:0.95 performans kriterlerine göre karşılaştırılmıştır. Elde edilen sonuçlar incelendiğinde, en yüksek başarıya sahip olan modelin "Deney 1" olduğu görülmektedir. Bu nedenle, sonraki aşamalarda analiz ve değerlendirmeler "Deney 1" eğitiminden elde edilen model üzerinden yürütülmüştür.

Tablo 5. YOLOv5 üzerinde yapılan deneyler

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95
Deney 1	%95	%88	%96	%62
Deney 2	%92	%86	%94	%58
Deney 3	%90	%85	%93	%55

YOLOv5 modelinin eğitim sürecine ait Şekil 31’de verilen eğitim sonuçlarına göre, modelin başlangıçtaki yüksek box_loss, obj_loss ve cls_loss değerlerinin yaklaşık 10. epoch’tan itibaren hızla düşüşe geçtiği ve 40. epoch sonrasında stabil bir hale geldiği görülmüştür. Bu durum, modelin kısa sürede konverge ettiğini ve hızlı öğrenme kapasitesine sahip olduğunu göstermektedir.



Şekil 31. YOLOv5 eğitim sonuçları

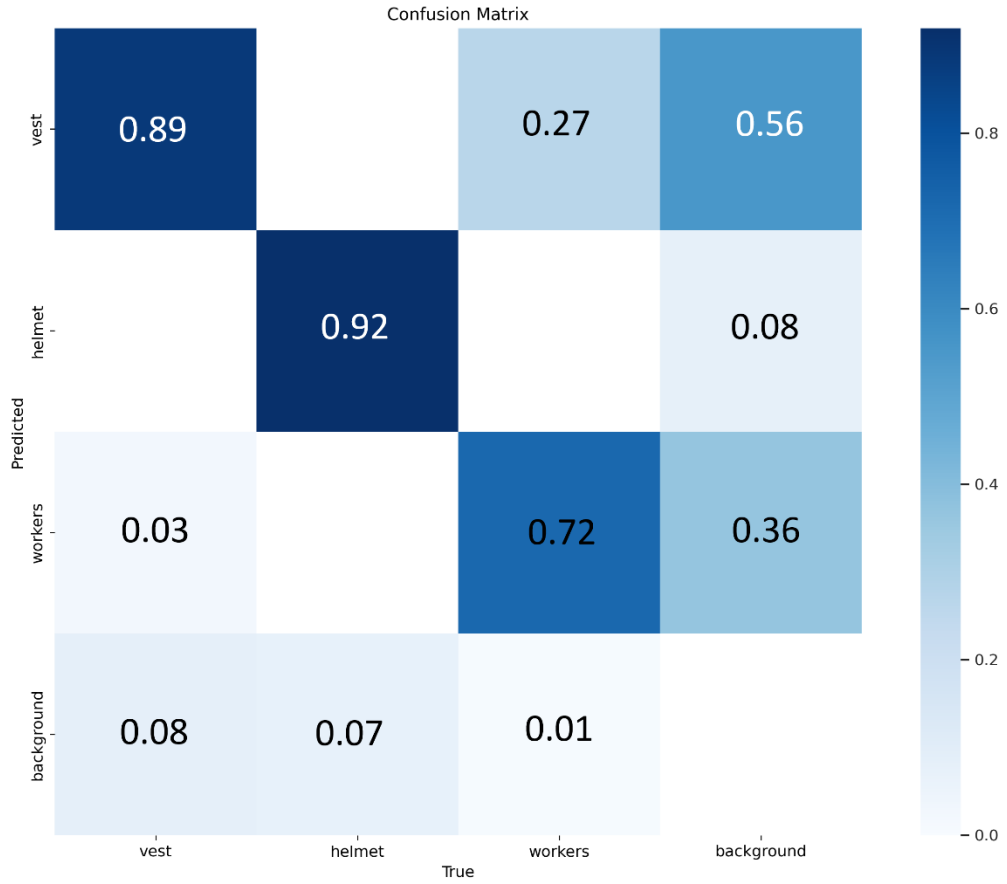
Eğitim süreci boyunca train/box_loss değeri ~0.09 seviyesinden başlayarak yaklaşık 0.04 seviyesine kadar düşmüştür. obj_loss da benzer şekilde %50'den fazla azalmış ve 0.015 civarında dengelenmiştir. Bu metrikler, modelin kutu tahmini (bounding box regression) ve nesne varlığına dair güven skorlarını başarılı bir şekilde öğrenebildiğini göstermektedir. sınıflandırma kaybı (cls_loss) da benzer şekilde belirgin bir düşüş göstermiş ve bu durum, modelin sınıf etiketlerini doğru şekilde ayırt etme yetisinin güçlendiğini kanıtlamaktadır.

Kesinlik (precision) metriği başlangıçta %50 civarındayken hızlı bir artış göstermiş ve eğitim sonunda %95 seviyesine ulaşmıştır. Duyarlılık (recall) ise başlangıçta %40 civarındayken kademeli bir şekilde artmış ve %88 düzeyine ulaşmıştır. Bu metrikler modelin hem doğru pozitif oranını yüksek tuttuğunu hem de çok sayıda nesneyi kaçırmadan tespit edebildiğini göstermektedir.

Özellikle metrics/mAP@0.5 değeri %96 ile oldukça yüksek bir başarı göstermektedir. Bu metrik yalnızca doğru sınıf tahmini değil, aynı zamanda tespit edilen nesnelerin doğru konumlandırıldığını da gösterir. Daha hassas bir metrik olan metrics/mAP@0.5:0.95 ise %62 düzeyindedir. Bu değer, modelin farklı IoU eşiklerinde dahi yüksek genel doğruluk gösterdiğini ortaya koymaktadır.

YOLOv5 modeline ait Şekil 32'de verilen karışıklık matrisi (confusion matrix), modelin sınıf bazında doğru ve yanlış sınıflandırma performansını değerlendirmek amacıyla kullanılmaktadır. Matrisin diyagonalindeki yüksek değerler —sırasıyla 0.89 yelek (vest), 0.92

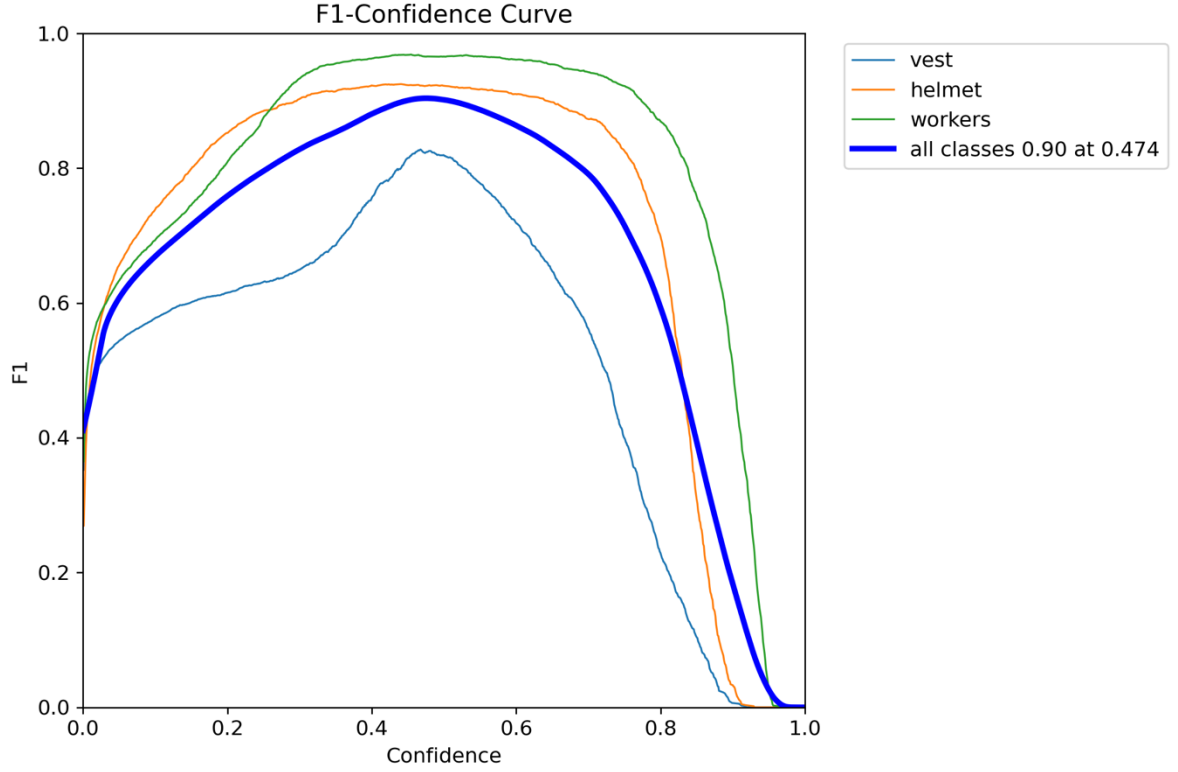
baret (helmet) ve 0.72 işçi (workers)— modelin çoğunlukla doğru sınıflandırmalar gerçekleştirdiğini ortaya koymaktadır. Bununla birlikte, 0.27 ve 0.56 gibi nispeten yüksek diyagonal dışı değerler, bazı sınıflar arasında karışıklıkların yaşandığını göstermektedir. Özellikle yelek sınıfının, diğer sınıflarla karıştırılma eğiliminde olduğu gözlemlenmektedir. Ayrıca arka plan (background) için oluşan yanlış pozitif ve yanlış negatif tahminler de çevresel faktörlerin sınıflandırma performansını etkileyebileceğine işaret etmektedir. Genel olarak, bu matris YOLOv5'in sınıflar arasında dengeli bir performans sergilediğini, ancak bazı sınıflarda iyileştirmeye açık alanlar bulunduğunu göstermektedir.



Şekil 32. YOLOv5 karmaşıklık matrisi

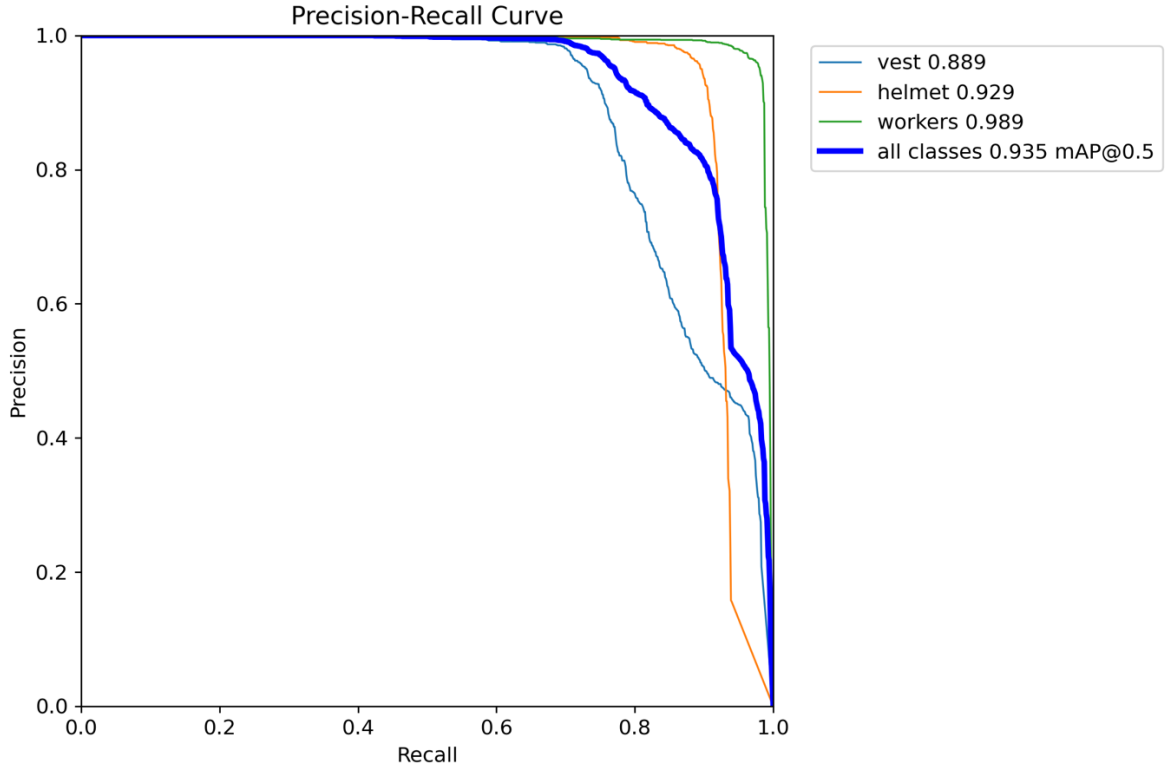
YOLOv5 modeline ait Şekil 33'te verilen F1-Güven eğrisi (F1-Confidence curve), modelin farklı güven skoru eşiklerinde sergilediği sınıflandırma başarımını detaylı şekilde ortaya koymaktadır. Grafik üzerinde tüm sınıflar için birleşik F1 skoru 0.90 seviyelerinde maksimuma ulaşmakta ve bu başarı 0.474 güven eşiği seviyesinde gerçekleşmektedir. Eğriler incelendiğinde, düşük eşik değerlerinde daha yüksek duyarlılık gözlemlenirken, artan eşik değerleriyle birlikte modelin yalnızca yüksek güvenli tahminlere yönelmesi nedeniyle kesinlik

artmakta, ancak duyarlılık azalmaktadır. Bu eğilimler, F1 skorunun güven eşiğiyle optimize edilmesinin performans açısından kritik olduğunu göstermektedir. YOLOv5 modelinin bu bağlamda dengeli bir yapı sergilediği ve uygun bir eşik değeriyle yüksek performans elde edilebildiği görülmektedir.



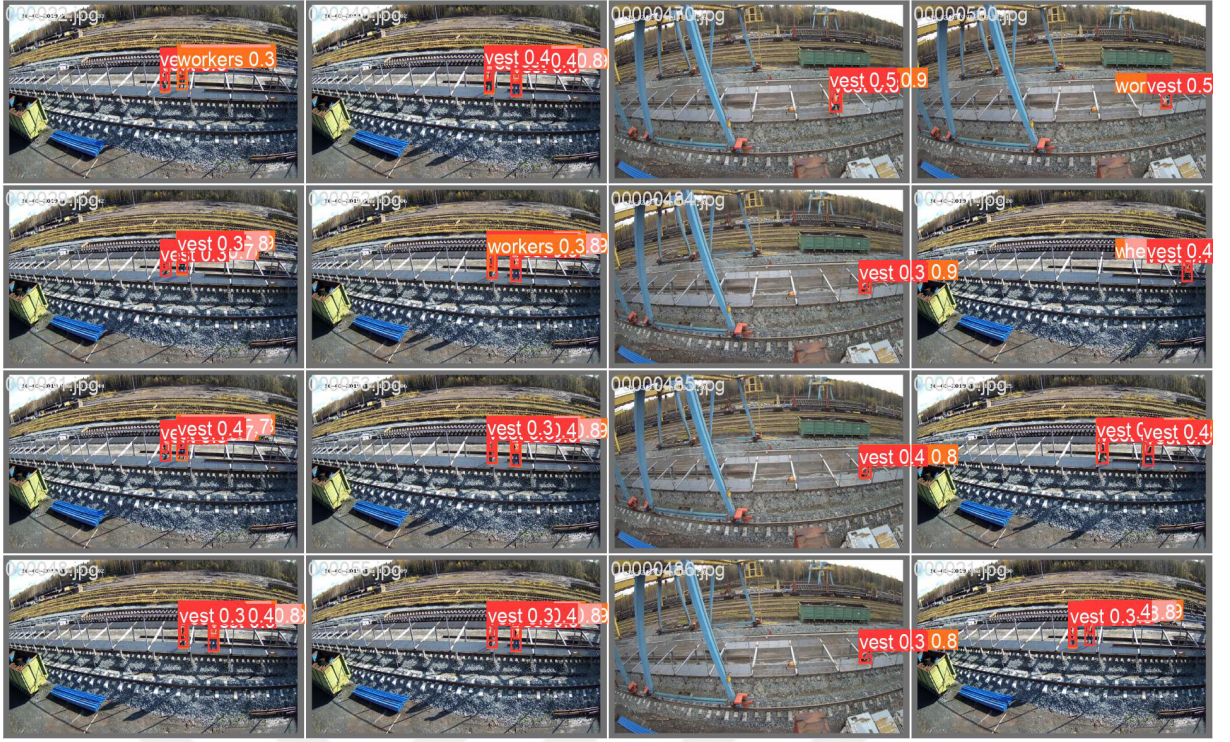
Şekil 33. YOLOv5 F1-Güven eğrisi

YOLOv5 modeline ait Şekil 34'te verilen kesinlik-duyarlılık (precision-recall) eğrisi, her bir sınıfın pozitif örneklerini tespit etme başarısını ve bu tahminlerin doğruluğunu göstermektedir. Eğriler, özellikle yüksek kesinlik ve duyarlılık değerleriyle sınıflandırma performansının güçlü olduğunu vurgulamaktadır. Sınıf bazlı değerlendirildiğinde, yelek için 0.889, baret için 0.929 ve işçi için 0.989 düzeyinde doğruluk sağlanmıştır. Tüm sınıflar için birleşik olarak hesaplanan ortalama $mAP@0.5$ değeri 0.935'tir. Bu sonuç, modelin genel başarımının yüksek olduğunu ve her sınıf için başarılı şekilde tahminlerde bulunabildiğini göstermektedir. Eğrilerin eğimi ve alan büyüklüğü, YOLOv5'in hem yanlış pozitifleri azaltmak hem de gerçek pozitifleri yüksek oranda yakalamak konusunda yetkin bir performans sergilediğini kanıtlamaktadır.



Şekil 34. YOLOv5 precision-recall eğrisi

Ayrıca YOLOv5 modelinin doğrulama verisi üzerindeki tahminlerine ait örnek çıktıları Şekil 35'te yer almaktadır. Her bir küçük karede, farklı doğrulama görüntüleri üzerine modelin tespit ettiği sınıflar (örneğin yelek, işçi) ve bu tahminlerin güven skorları etiketlenmiştir. Görüntülerde genel olarak yelek sınıfının tespiti baskındır ve çoğu durumda tahmin güven skorları %80 üzerinde seyretmektedir. Örneğin, yelek 0.5 0.9 gibi bir etiket, modelin yelek nesnesini %90 güven ile tespit ettiğini ifade eder.



Şekil 35. YOLOv5 modelinin doğrulama verisi üzerindeki tahminlerine ait örnek çıktıları

Ancak görsel incelendiğinde, bazı sınıflarda kararsızlık ve düşük güven skorları da gözlemlenmektedir. Özellikle işçiler (workers) sınıfında yapılan tahminlerde bazı kutuların skorları %30 civarında olup (workers 0.3) modelin bu sınıfı ayırt etmekte daha fazla belirsizlik yaşadığını göstermektedir. Benzer şekilde, arka planla benzer özellikler taşıyan nesnelerin model tarafından yanlış sınıflandırılması ya da düşük skorla işaretlenmesi, modelin bazı durumlarda sınıf ayrımını zor yaptığını işaret etmektedir.

Buna ek olarak, bazı etiketlerin yazım hatalı olarak ("worvest", "whevest", "cvest") gösterilmesi ya da kaymaların yaşanması, görselleştirme sırasında modelden bağımsız sistematik hata ihtimaline de işaret edebilir.

Genel olarak bu görsel, YOLO tabanlı modelin özellikle yelek sınıfında yüksek başarı gösterdiğini, ancak diğer sınıflarda — özellikle işçi — daha düşük güven skorlarıyla tahmin yaptığını ve modelin sınıf ayrımında bazı örneklerde zorlandığını açıkça ortaya koymaktadır. Bu durum modelin optimizasyonunda sınıf dengesine dikkat edilmesi gerektiğini ve bazı sınıflarda veri artırımı ya da örnek çeşitliliği artırılarak başarımın daha da iyileştirilebileceğini düşündürmektedir.

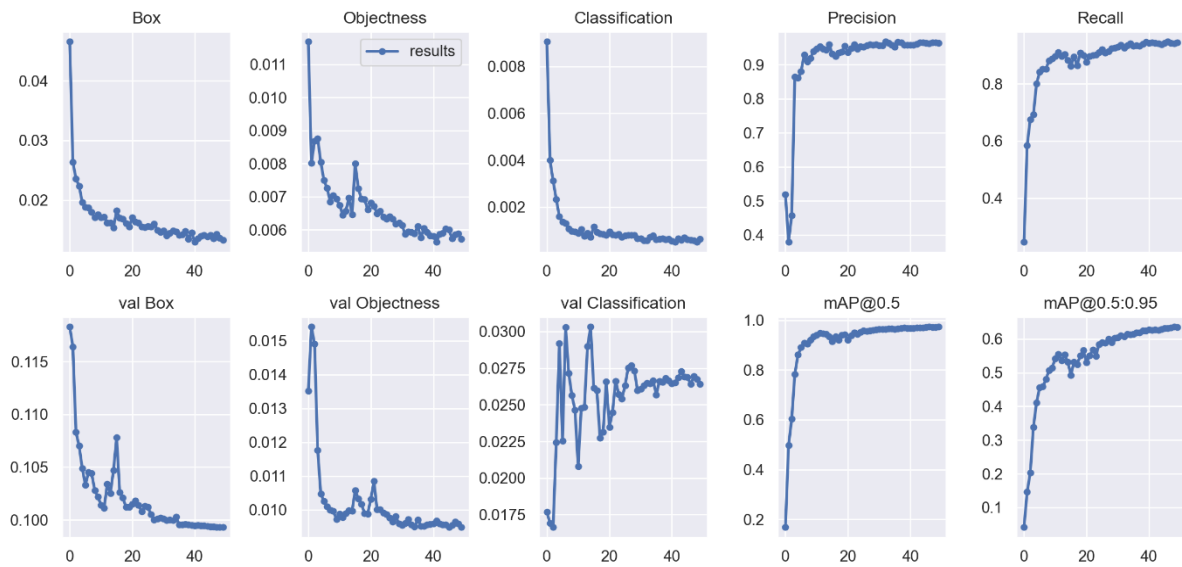
6.2. YOLOv7 Modeline Ait Bulgular

Bu çalışma kapsamında ilgili YOLOv7 modeli üç kez eğitime tabi tutulmuş ve her bir eğitim sonucunda Tablo 6’da sunulan değerlendirme metrikleri elde edilmiştir. Modeller, precision, recall, mAP@0.5 ve mAP@0.5:0.95 performans kriterlerine göre karşılaştırılmıştır. Elde edilen sonuçlar incelendiğinde, en yüksek başarıya sahip olan modelin "Deney 2" olduğu görülmektedir. Bu nedenle, sonraki aşamalarda analiz ve değerlendirmeler "Deney 2" eğitiminden elde edilen model üzerinden yürütülmüştür.

Tablo 6. YOLOv7 üzerinde yapılan deneyler

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95
Deney 1	%92	%88	%94	%60
Deney 2	%94	%90	%96	%63
Deney 3	%91	%87	%93	%58

YOLOv7 modelinin eğitim sürecine ait Şekil 36’da verilen eğitim sonuçlarına göre, box, obj ve cls kayıplarında benzer bir iyileşme görülmekle birlikte, bu modelin daha agresif bir öğrenme eğrisi çizdiği fark edilmiştir. Özellikle obj_loss metriği eğitim süresince oldukça hızlı bir şekilde düşmüş ve doğrulama tarafında (val obj_loss) daha düşük minimumlara ulaşmıştır. Bu durum, YOLOv7’nin nesne tespitiye yönelik skorlamada daha tutarlı sonuçlar verebildiğini göstermektedir.



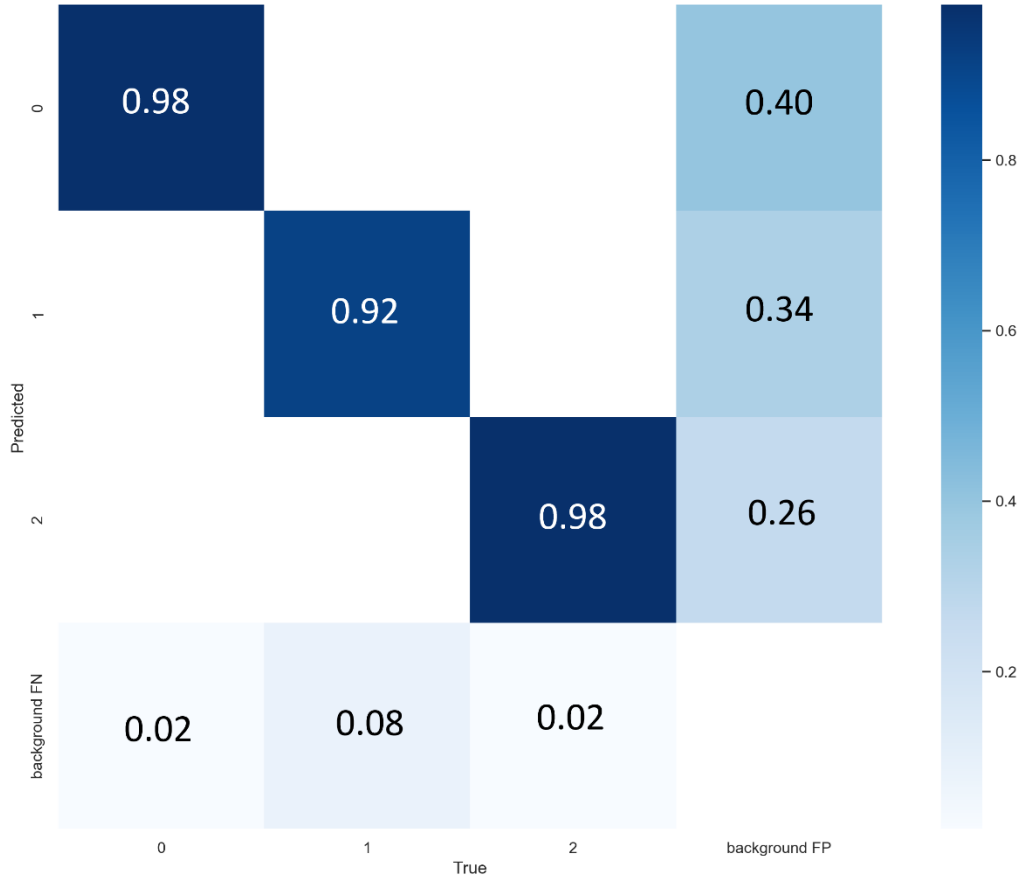
Şekil 36. YOLOv7 eğitim sonuçları

Bununla birlikte, val classification loss grafiğinde zaman zaman belirgin dalgalanmalar gözlemlenmiştir. Bu dalgalanmalar modelin sınıf ayrımı konusunda zaman zaman kararsızlık yaşadığını veya sınıflar arası veri dengesizliklerinden etkilendiğini gösterebilir. Ancak bu dalgalanmalara rağmen genel eğilim azalan bir yönde ilerlemiştir.

Kesinlik (precision) ve duyarlılık (recall) metriklerinde YOLOv7'nin performansı sırasıyla %94 ve %90 seviyelerine ulaşmıştır. mAP@0.5 metriği, YOLOv5'le benzer şekilde %96 düzeyindedir. mAP@0.5:0.95 metriği ise %63 seviyesinde sabitlenmiştir. Bu değerler, YOLOv7'nin YOLOv5'e benzer bir doğruluk sunarken, belirli sınıflarda daha iyi ayrıştırma yaptığına dair ipuçları vermektedir.

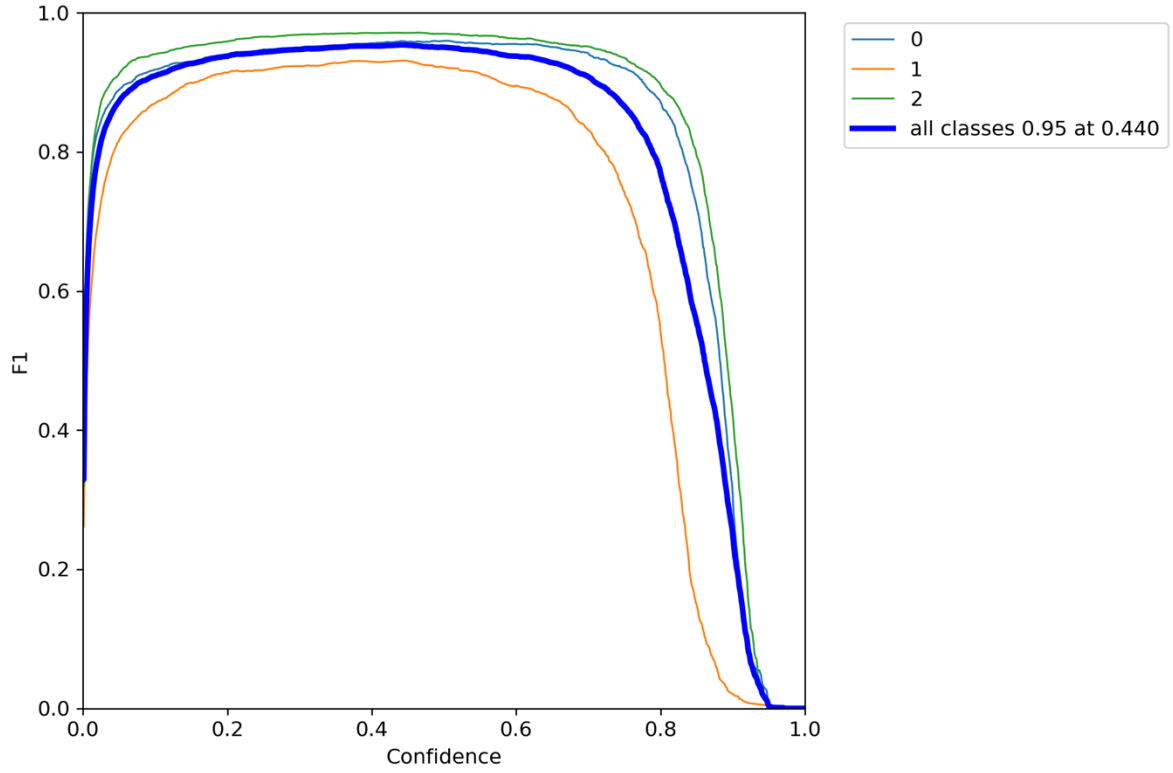
YOLOv7, literatürde genellikle hız ve doğruluk açısından dengeli bir model olarak kabul edilir. Bu çalışmada da eğitim süresi açısından daha kısa sürede konverge olmuş ve düşük kayıplarla yüksek doğruluk değerleri sunmuştur. Bu nedenle gerçek zamanlı sistemlerde veya gömülü sistemlerde tercih edilebilir niteliktedir.

YOLOv7 modeline ait Şekil 37'da verilen karışıklık matrisi (confusion matrix), modelin test veri seti üzerindeki sınıflandırma doğruluğunu detaylı bir biçimde göstermektedir. Matrisin diyagonalinde yer alan değerler, her bir sınıfın doğru sınıflandırılma oranlarını ifade etmektedir. Buna göre sırasıyla 0 (yelek), 1 (baret) ve 2 (işçi) etiketli sınıflar için doğruluk oranları 0.98, 0.92 ve 0.98 olarak gerçekleşmiştir. Bu oranlar, modelin sınıf ayrımını yüksek doğrulukla gerçekleştirebildiğini göstermektedir. Diyagonal dışı kalan değerler, modelin diğer sınıfları yanlış tahmin ettiği durumlara işaret etmektedir. Özellikle arka plan (background) için elde edilen yanlış pozitif (false positive) ve yanlış negatif (false negative) oranları, sınıflar arasında bazı karışıklıkların yaşandığını göstermekle birlikte, bu sapmalar genel doğruluk seviyesini anlamlı düzeyde düşürmemektedir. Karışıklık matrisinin genel yapısı, YOLOv7 modelinin sınıf bazında başarılı ve tutarlı bir sınıflandırma performansı sergilediğini kanıtlamaktadır.



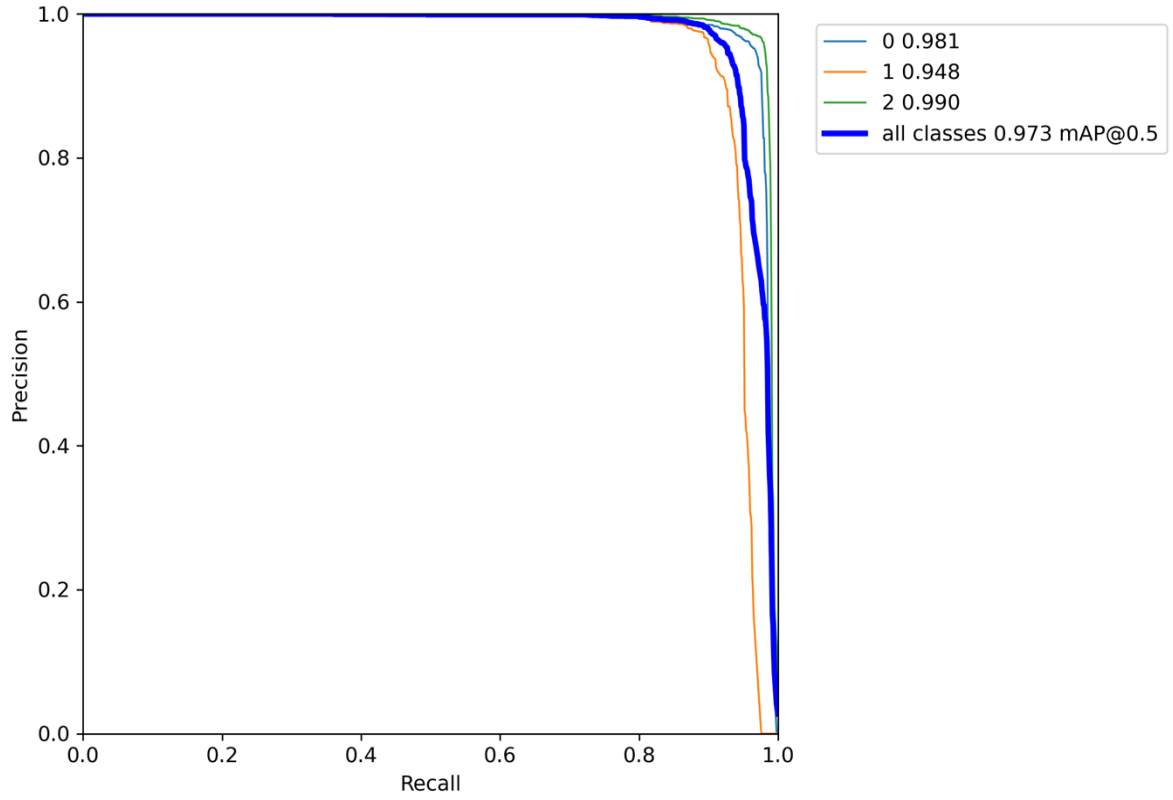
Şekil 37. YOLOv7 karmaşıklık matrisi

YOLOv7 modeline ait Şekil 38’de verilen F1-Güven eğrisi (F1-Confidence curve), farklı güven skoru (confidence score) eşik değerleri altında modelin sınıf bazlı F1 skorlarını gözlemlemeye olanak sağlamaktadır. F1 skoru, hem kesinlik (precision) hem de duyarlılık (recall) metriklerinin harmonik ortalaması olduğundan, modelin genel sınıflandırma başarısını dengeleyen kritik bir ölçüttür. Grafik incelendiğinde, üç sınıfın her biri için yüksek F1 skorları elde edildiği ve özellikle güven skoru 0.44 civarında tüm sınıflar için optimum bir denge noktası yakalandığı gözlemlenmektedir. Bu noktada genel F1 skoru yaklaşık 0.95 seviyesinde gerçekleşmiştir. Özellikle düşük güven eşiklerinde daha fazla pozitif tespit yapılmasına bağlı olarak duyarlılığın artarken, yanlış pozitiflerin artması nedeniyle kesinlikte hafif bir azalma görülmektedir. Buna karşın, yüksek eşiklerde tam tersi bir eğilim mevcuttur. Grafik, modelin güven eşiği optimizasyonuna duyarlı olduğunu ve sınıflandırma başarımının güven skoruna göre istikrarlı şekilde değiştiğini göstermektedir.



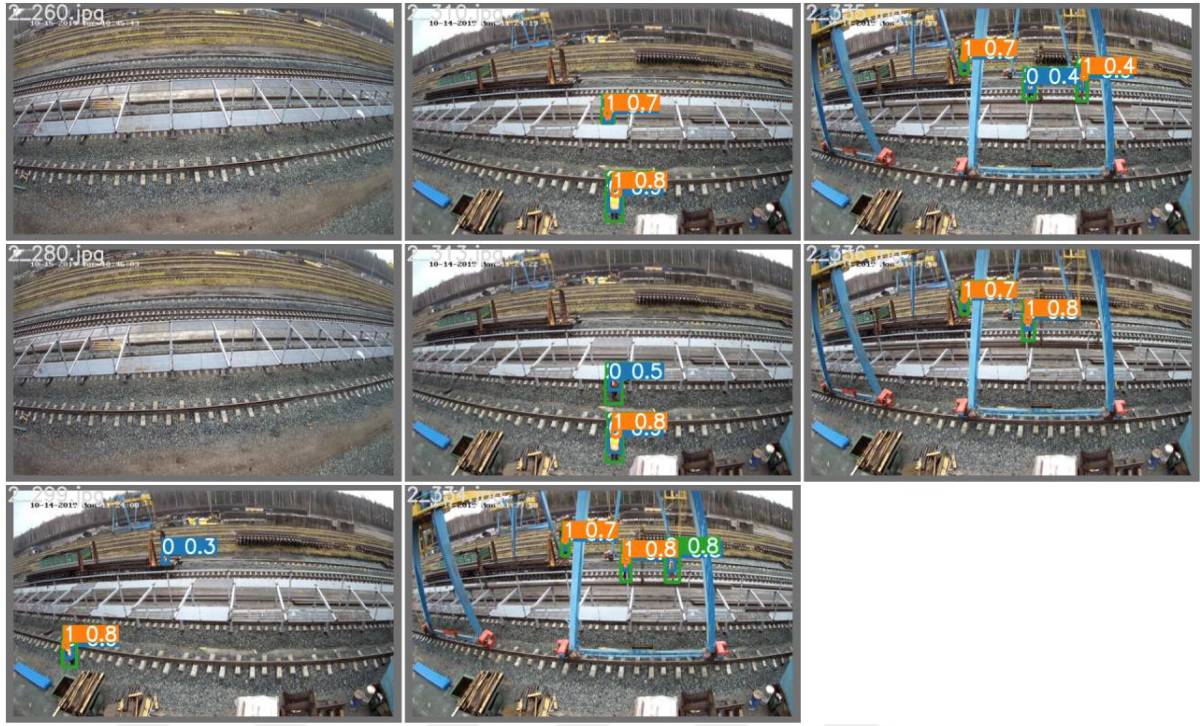
Şekil 38. YOLOv7 F1-Güven eğrisi

YOLOv7 modeline ait Şekil 39’da verilen kesinlik-duyarlılık (precision-recall) eğrisi, her bir sınıf için elde edilen tahminlerin doğruluğu ve kapsamı arasındaki ilişkiyi ayrıntılı şekilde sunmaktadır. Eğrilerdeki alanın büyüklüğü Area Under Curve (AUC), modelin başarılı tahminlerde bulunma kabiliyetinin önemli bir göstergesidir. Bu bağlamda, modelin sınıf 0, 1 ve 2 için sırasıyla 0.981, 0.948 ve 0.990 düzeyinde ortalama kesinlik değerleriyle yüksek başarı sağladığı görülmektedir. Tüm sınıflar için ortalama alınarak elde edilen $mAP@0.5$ değeri 0.973’tür. Bu sonuç, YOLOv7 modelinin hem hassasiyet hem de duyarlılık bakımından dengeli ve güçlü bir performans sergilediğini ortaya koymaktadır. Özellikle eğrilerin hızlı bir şekilde üst sınırdan yoğunlaşması, modelin pozitif örnekleri ayırt etme ve yanlış sınıflandırmaları minimize etme becerisinin yüksek olduğunu göstermektedir. Bu başarı düzeyi, YOLOv7 mimarisinin nesne tespiti görevlerinde güçlü bir alternatif olduğunu desteklemektedir.



Şekil 39. YOLOv7 precision-recall eğrisi

Ayrıca YOLOv7 modelinin doğrulama verisi üzerindeki tahminlerine ait örnek çıktıları Şekil 40'da yer almaktadır. Görseller üzerinde tespit edilen nesneler (0: baret, 1: yelek, 2: işçi) sınıf etiketleri ve modelin verdiği güven skoru ile birlikte gösterilmiştir. Modelin çoğu tahmini %70 ila %90 arasında yüksek güven skorlarıyla yapılmıştır. Örneğin, bazı karelerde işçi sınıfına ait tahminlerin 0.8 skoru ile doğrulandığı görülmektedir. Aynı şekilde, farklı sınıflara ait tespitlerin birden fazla nesne içeren görüntülerde bile başarılı şekilde ayrıştırıldığı dikkat çekmektedir.



Şekil 40. YOLOv7 modelinin doğrulama verisi üzerindeki tahminlerine ait örnek çıktılar

Bununla birlikte, düşük skorla yapılan (örneğin 0.3–0.5 arası) bazı tahminler de mevcuttur. Bu durum modelin bazı sınırlı senaryolarda kararsızlık yaşadığını veya zorlu çevresel koşullarda (açık alan, uzaklık, benzer zemin rengi gibi) nesneleri sınıflandırmakta zorlandığını göstermektedir. Ancak genel performans açısından bakıldığında, YOLOv7 modelinin sahne içi kalabalık, farklı perspektifler ve zayıf kontrast gibi etkenlere rağmen oldukça tutarlı ve doğru tespitler gerçekleştirdiği görülmektedir.

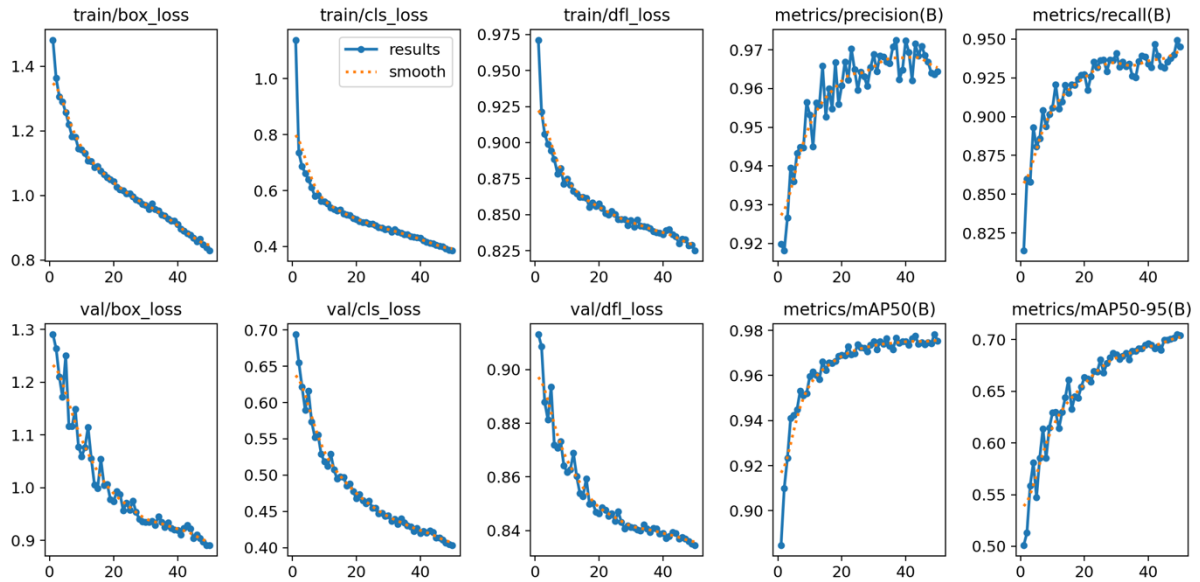
6.3. YOLOv8 Modeline Ait Bulgular

Bu çalışma kapsamında ilgili YOLOv8 modeli üç kez eğitime tabi tutulmuş ve her bir eğitim sonucunda Tablo 7’de sunulan değerlendirme metrikleri elde edilmiştir. Modeller, precision, recall, mAP@0.5 ve mAP@0.5:0.95 performans kriterlerine göre karşılaştırılmıştır. Elde edilen sonuçlar incelendiğinde, en yüksek başarıya sahip olan modelin "Deney 1" olduğu görülmektedir. Bu nedenle, sonraki aşamalarda analiz ve değerlendirmeler "Deney 1" eğitiminden elde edilen model üzerinden yürütülmüştür.

Tablo 7. YOLOv8 üzerinde yapılan deneyler

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95
Deney 1	%94	%93	%97	%70
Deney 2	%92	%91	%95	%66
Deney 3	%90	%90	%94	%63

YOLOv8 modeli, önceki versiyonlara göre mimari olarak önemli yenilikler içermektedir. Özellikle distribution focal loss (dfl_loss) gibi ileri düzey kayıp fonksiyonları ile model, kutu regresyonunda daha hassas bir öğrenme sağlayabilmektedir. YOLOv8 modelinin eğitim sürecine ait Şekil 41’de verilen eğitim sonuçlarına göre, train/box_loss, cls_loss ve dfl_loss metriklerinde oldukça yüksek başlangıç değerleri gözlemlenmiş olsa da, bu kayıpların hızla düşüşe geçerek oldukça stabil ve düşük seviyelere indiği görülmüştür.



Şekil 41. YOLOv8 eğitim sonuçları

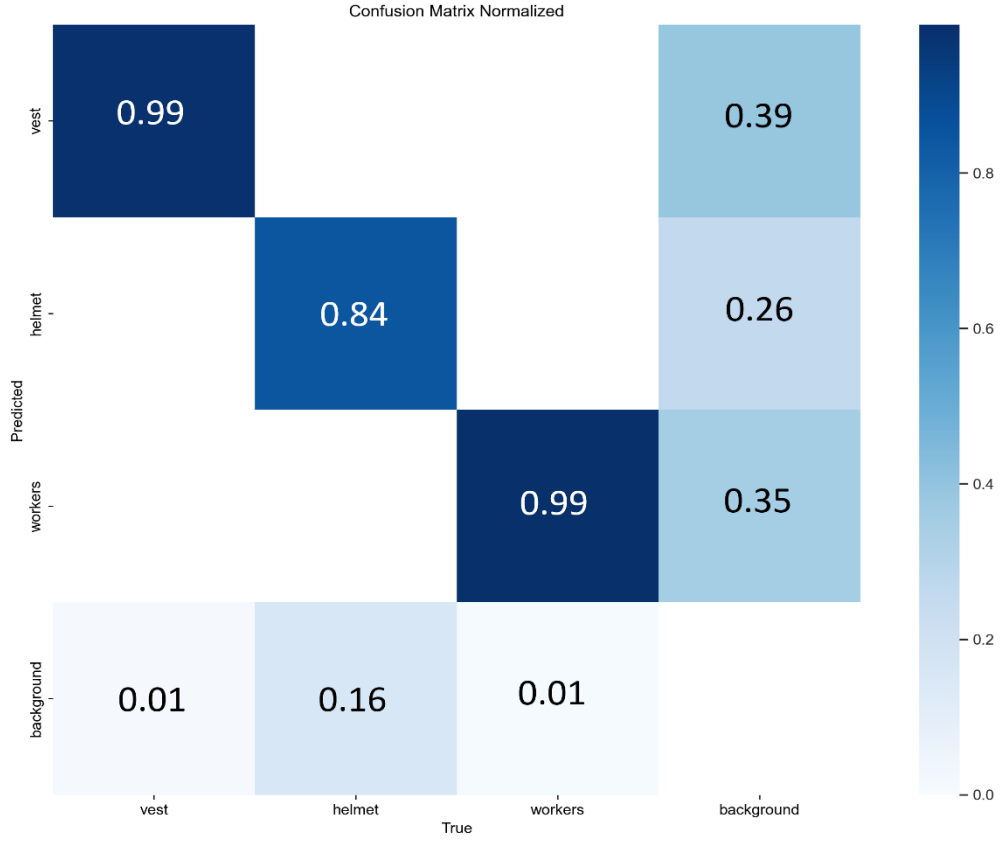
Doğrulama tarafında da benzer bir eğilim gözlemlenmiştir. val/dfl_loss yaklaşık 0.91 seviyesinden başlayarak 0.83 seviyelerine kadar düşmüştür. val/box_loss ve val/cls_loss metrikleri de benzer şekilde yaklaşık %30 oranında iyileşme göstermiştir.

En dikkat çekici bulgu, metrics/mAP50-95 değerinin %70 seviyesine kadar çıkmış olmasıdır. Bu değer, YOLOv8’in daha sıkı IoU eşiklerinde dahi üstün performans sergileyebildiğini ve hem sınıflandırma hem de konumlandırmada yüksek isabet oranına sahip olduğunu

göstermektedir. YOLOv8, kesinlik (precision) ve duyarlılık (recall) metriklerinde de sırasıyla %94 ve %93 gibi oldukça yüksek değerlere ulaşmıştır. Bu durum, modelin hem yanlış pozitif oranını düşürdüğünü hem de eksik tespiti minimize ettiğini göstermektedir.

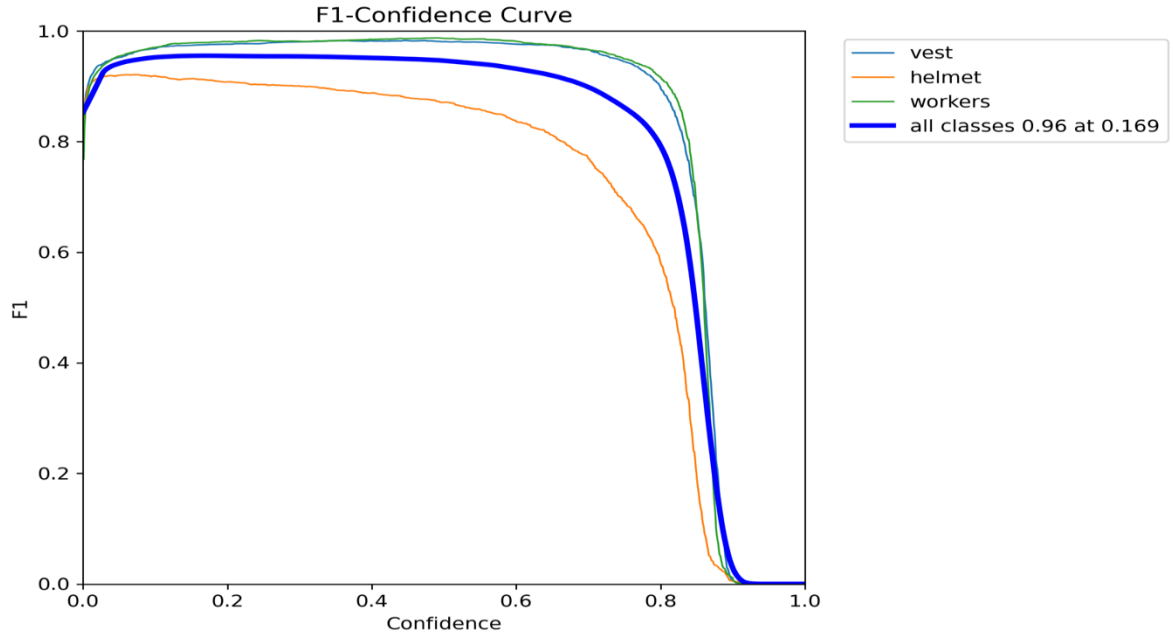
Bu sonuçlar, YOLOv8'in daha karmaşık yapılar, küçük nesneler veya sınıf benzerliği yüksek veriler gibi zorlayıcı senaryolarda daha iyi genelleme yapabildiğini ortaya koymaktadır.

Şekil 42'de YOLOv8 modelinin dört sınıfa ait sınıflandırma performansı, sütun bazlı normalize edilmiş karışıklık matrisi üzerinden görselleştirilmiştir. Model, yelek (vest) sınıfında %99 oranında doğru tahmin yapmış ve yalnızca %1 oranında arka plan (background) ile karıştırmıştır. Bu sınıfı baret (helmet) ve işçi (workers) sınıflarıyla neredeyse hiç karıştırmamıştır. Baret (helmet) sınıfında ise doğruluk oranı %84 seviyesinde olup, %16 oranında arka plan sınıfına yanlış tahmin yapılmıştır. Modelin baret ile diğer sınıflar arasında anlamlı bir karışıklık yapmadığı görülmektedir. İşçi (workers) sınıfı için doğruluk %99 olup, yalnızca %1 oranında arka plan olarak yanlış tahmin edilmiştir. Bu da modelin işçi sınıfını oldukça yüksek doğrulukla tespit ettiğini göstermektedir. Özellikle dikkat çekici olan ise arka plan (background) sınıfıdır. Arka plan olan bölgelerin %39'u yelek, %26'sı baret ve %35'i işçi olarak yanlış tahmin edilmiştir. Bu durum, modelin arka planı tanımakta zorlandığını ve arka planın nesnelerle olan görsel benzerliğinin model üzerinde hata oluşturduğunu göstermektedir. Bu durum, arka planın nesnelerle olan görsel benzerliğinden kaynaklı olarak modelin nesne olmayan bölgeleri sınıflandırmakta daha fazla hata yaptığını göstermektedir. Genel olarak matriste diyagonal değerlerin yüksekliği, YOLOv8'in sınıflar arasında güçlü bir ayırım kabiliyetine sahip olduğunu ortaya koymaktadır.



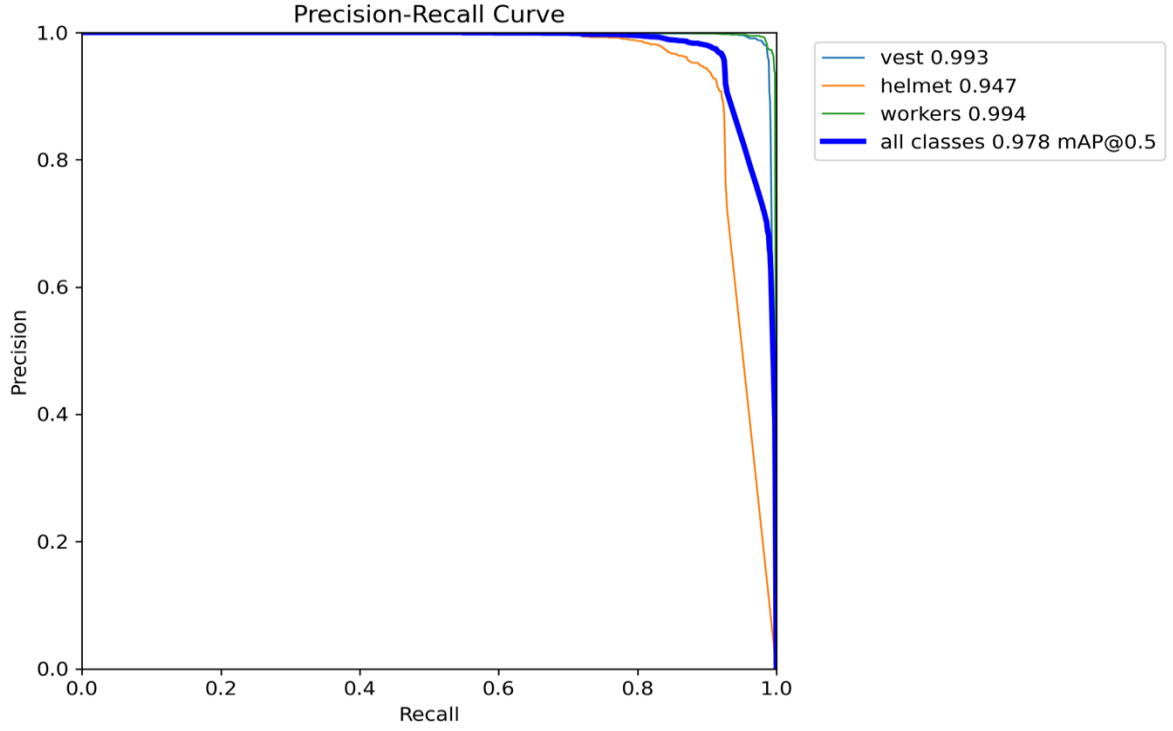
Şekil 42. YOLOv8 karmaşıklık matrisi

YOLOv8 modeline ait Şekil 43’te verilen güven skoru eşiklerine göre F1 skorlarındaki değişim görselleştirilmiştir. Modelin genel F1 skoru, %96 gibi yüksek bir seviyeye ulaşmakta olup, bu maksimum değer 0.169 güven skorunda elde edilmiştir. Bu eşik değeri, modelin daha düşük güven seviyelerinde dahi yüksek bir genel başarı sağladığını göstermektedir. Sınıf bazında incelendiğinde, işçi sınıfı eğrisi %99’a yakın bir F1 skoru ile üst düzey bir doğruluk sunarken; yelek sınıfı da benzer şekilde güçlü bir eğri profili sergilemiştir. Ancak baret sınıfı için F1 skoru diğerlerine göre daha düşük seyretmiş, yüksek güven eşiklerinde belirgin bir düşüş gözlenmiştir. Bu, baret sınıfındaki tespitlerin daha az güvenli olduğunu ve sınıfın genel başarıyı sınırlayan bir faktör olabileceğini göstermektedir. Eğrilerin genel yapısı incelendiğinde, YOLOv8’in düşük güven skoru eşiklerinde bile hatırı sayılır F1 değerleri sunduğu, dolayısıyla esnek ve geniş kapsamlı bir tespit kabiliyetine sahip olduğu sonucuna varılabilir.



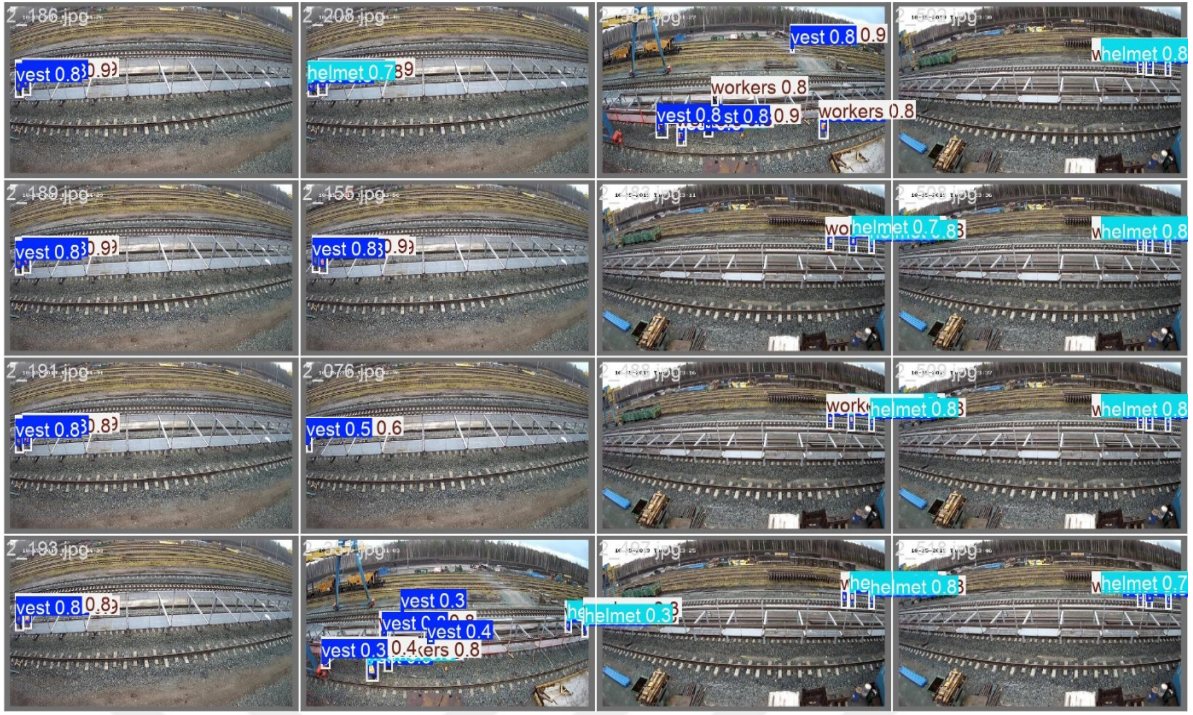
Şekil 43. YOLOv8 F1-Güven eğrisi

YOLOv8 modeline ait Şekil 44'te verilen kesinlik-duyarlılık (precision-recall) eğrisi, modelinin sınıf bazında kesinlik ve duyarlılık dengelerini görselleştirmektedir. Model, işçi sınıfında %99.4 kesinlik ve oldukça yüksek bir duyarlılık değeri elde ederek bu sınıfta neredeyse kusursuz bir performans göstermiştir. Benzer şekilde yelek sınıfında da %99.3 kesinlik ile yüksek doğruluk sağlanmıştır. Ancak baret sınıfında kesinlik %94.7 seviyesinde kalmış, bu da diğer sınıflara göre daha fazla yanlış pozitif tespit gerçekleştiğini göstermektedir. Tüm sınıflar için hesaplanan ortalama doğruluk ($mAP@0.5$) ise %97.8 gibi son derece yüksek bir seviyededir. Bu yüksek ortalama değer, YOLOv8'in tüm sınıflar için dengeli, güvenilir ve güçlü bir genel performans sergilediğini ortaya koymaktadır. Eğrilerin biçimi itibarıyla, özellikle düşük duyarlılık aralıklarında bile yüksek kesinlik değerlerinin korunmuş olması, modelin yalnızca doğru tahminler yapmadığını, aynı zamanda bu tahminlerde yüksek güven düzeyine de sahip olduğunu göstermektedir.



Şekil 44. YOLOv8 precision-recall eğrisi

Ayrıca YOLOv8 modelinin doğrulama verisi üzerindeki tahminlerine ait örnek çıktıları Şekil 45'te yer almaktadır. Görüntülerin üzerinde tespit edilen nesneler; ilgili sınıf etiketleri yelek (vest), baret (helmet), işçi (workers) ve modelin tahminine ait güven skorları ile birlikte gösterilmiştir. Görselde, modelin çoğu tespiti yüksek doğrulukla gerçekleştirdiği gözlemlenmektedir. Özellikle yelek sınıfında 0.8 ve üzeri güven skoru ile yapılan tahminler dikkat çekmektedir. Aynı şekilde, baret sınıfında da 0.73–0.83 aralığında tutarlı tahminler yer almakta, bu da modelin KKD öğelerini etkin bir biçimde ayırt edebildiğini göstermektedir.



Şekil 45. YOLOv8 modelinin doğrulama verisi üzerindeki tahminlerine ait örnek çıktılar

YOLOv8 modeli, karmaşık sahnelerdeki çoklu nesne varlığına rağmen sınıflar arası ayrımı başarılı bir şekilde gerçekleştirmiştir. Bazı görüntülerde düşük güven skorlarıyla (örneğin 0.3–0.4 arası) yapılan tahminler görülsede genel olarak tespit edilen nesnelerin konumlandırması doğru ve sınıflandırmaları tutarlıdır. Özellikle zorlu açılardan veya uzak mesafeden görülen nesnelerin bile belirli güven aralıklarında doğru şekilde sınıflandırılması, modelin yüksek genel performansını yansıtmaktadır.

Bu sonuçlar, YOLOv8'in hem sınıf ayrımı hem de konum doğruluğu açısından selef modellerine kıyasla daha gelişmiş bir tespit kabiliyeti sunduğunu desteklemektedir.

6.4. Karşılaştırmalı Performans Analizi

YOLO mimarisinin üç farklı sürümü olan YOLOv5, YOLOv7 ve YOLOv8 modelleri, bu çalışmada aynı veri seti ve benzer eğitim parametreleriyle değerlendirilmiş; elde edilen sonuçlar kesinlik (precision), duyarlılık (recall) ve ortalama doğruluk (mAP) bakımından Tablo 8'de karşılaştırmalı olarak sunulmuştur.

Tablo 8. YOLO sürümlerinin karşılaştırmalı performans analizi

Model	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLOv5	%95	%88	%96	%62
YOLOv7	%94	%90	%96	%63
YOLOv8	%94	%93	%97	%70

Tablo 8 incelendiğinde, YOLOv5 modeli, hızlı konverjans kabiliyeti ve düşük eğitim kayıpları ile dikkat çekmiştir. Eğitim boyunca hem kutu kaybı (box loss) hem de sınıflandırma kaybı (cls loss) istikrarlı bir biçimde düşmüş, doğrulama kayıplarında da paralel bir azalma gözlemlenmiştir. Elde edilen kesinlik oranı yaklaşık %95, duyarlılık oranı ise %88 seviyesine ulaşmıştır. Bu durum, modelin doğru tespit oranının oldukça yüksek, ancak bazı nesneleri gözden kaçırma ihtimalinin görece daha fazla olduğunu göstermektedir. mAP@0.5 değeri %96 ile oldukça yüksek iken, mAP@0.5:0.95 değeri %62 civarında seyretmiştir. Bu da modelin genel anlamda başarılı olduğunu, ancak daha sıkı IoU eşiklerinde bazı sınırlamalar yaşadığını göstermektedir.

YOLOv7 modeli, genel performans açısından YOLOv5'e oldukça benzer sonuçlar vermiştir; ancak dikkat çekici farklar da mevcuttur. YOLOv7'nin duyarlılık oranı %90 ile YOLOv5'ten daha yüksektir, bu da modelin daha fazla nesneyi doğru şekilde tespit edebildiğini göstermektedir. Kesinlik değeri ise %94 olarak ölçülmüş ve YOLOv5'e çok yakın bir performans sergilemiştir. mAP@0.5 yine %96 düzeyindedir; ancak mAP@0.5:0.95 değeri %63'e çıkarak, sıkı eşiklerde biraz daha iyi performans sergilemiştir. Bununla birlikte, doğrulama sırasındaki sınıflandırma kaybı grafiklerinde zaman zaman dalgalanmalar görülmüştür. Bu durum, modelin sınıf ayrımı yaparken bazı belirsizlikler yaşadığını ya da eğitim verisinde sınıflar arası dengesizliklerden etkilendiğini göstermektedir. Ancak genel stabilitesi ve overfitting riski açısından YOLOv7'nin oldukça dengeli bir öğrenme süreci sergilediği söylenebilir.

Öte yandan, YOLOv8 modeli bu çalışmada en yüksek başarıyı gösteren mimari olarak öne çıkmıştır. Bu model, yalnızca klasik loss fonksiyonlarını değil, aynı zamanda gelişmiş distribution focal loss (dfl_loss) gibi modern yöntemleri de içermektedir. Bu sayede hem kutu tahmininde hem de sınıf ayrımında çok daha hassas sonuçlar elde edilebilmiştir. Eğitim sürecinde kayıpların oldukça yüksek seviyelerden başlamasına rağmen hızlı ve dengeli bir

azalma göstermesi, modelin hem karmaşık yapıları hem de sınıflar arası farklılıkları öğrenmede yüksek bir kapasiteye sahip olduğunu göstermektedir.

YOLOv8 modelinin kesinlik değeri %94 ve duyarlılık değeri %93'tür. Bu, modelin hem doğru pozitif oranının yüksek hem de çok az sayıda nesneyi atladığını göstermektedir. Bu iki metriğin birlikte yüksek çıkması, modelin genel başarımı açısından oldukça olumlu bir göstergedir. Ayrıca mAP@0.5 metriği %97 ile diğer modellerin önünde yer almakta; mAP@0.5:0.95 değeri ise %70 gibi oldukça etkileyici bir seviyeye ulaşmıştır. Bu değer, modelin yalnızca genel tespitlerde değil, aynı zamanda daha zorlayıcı, hassas ölçümlerde de başarılı olduğunu ortaya koymaktadır. Eğitim süreci boyunca herhangi bir overfitting belirtisine rastlanmamış ve eğitim-doğrulama kayıpları arasında önemli bir sapma oluşmamıştır. Bu da modelin yüksek genelleme kabiliyetine sahip olduğunu göstermektedir.

Tüm bu bulgular bir araya getirildiğinde, YOLOv8 modeli, özellikle doğruluk, hassasiyet ve eğitim stabilitesi açısından en üstün performansı göstermiştir. YOLOv5, daha düşük sistem kaynakları gerektiren uygulamalar için tercih edilebilecek, hızlı ve dengeli bir seçenek olarak değerlendirilebilirken; YOLOv7, daha güçlü duyarlılık değeriyle sahadaki nesneleri daha eksiksiz tespit etme konusunda avantaj sağlamaktadır. Ancak gerek mAP performansı gerek eğitim kararlılığı bakımından YOLOv8 açık ara öne çıkmış ve genel başarıda en yüksek verimi sağlamıştır.

7. TARTIŞMA

Bu çalışma kapsamında, KKD kullanan çalışanların tespitine yönelik bir yapay zeka tabanlı görüntü işleme sistemi geliştirilmiştir. Bu sistem, YOLOv5, YOLOv7 ve YOLOv8 gibi güncel nesne tespiti mimarileri kullanılarak modellenmiştir. Her bir model, farklı performans metrikleri açısından değerlendirilerek, model doğruluğu, hassasiyeti ve genel etkinliği üzerinde karşılaştırmalı analizler yapılmıştır.

Deneysel çalışmalarda, YOLOv5 modelinin %95 kesinlik ve %88 duyarlılık oranı ile başarılı bir performans sergilediği gözlemlenmiştir. Bu model, He vd. (2021) ve Zhang vd. (2023) çalışmalarına benzer şekilde, sahadaki KKD bileşenlerini tespit etmede etkili bulunmuştur. Bununla birlikte, Li vd. (2022) tarafından demiryolu üzerindeki tespit çalışmalarında YOLOv5'in zaman duyarlı uygulamalar için de yeterli olduğu rapor edilmiştir.

YOLOv7 modeli ise, %94 kesinlik ve %90 duyarlılık oranı ile özellikle eksik tespitlerin azaltılmasında öne çıkmıştır. Wang vd. (2022) tarafından sunulan YOLOv7 mimarisi, eğitilebilir parametre çeşitliliği ve hafif mimari tasarımı sayesinde hem isabetli hem de verimli bir çözüm sunmaktadır. Bu modelin bu çalışmadaki uygulamasında da benzer şekilde, KKD bileşenlerinin yoğun bulunduğu sahnelerde daha fazla tespiti başarıyla gerçekleştirebildiği gözlemlenmiştir.

YOLOv8 modeli ise %94 kesinlik ve %93 duyarlılık oranının yanı sıra, %97 mAP@0.5 ve %70 mAP@0.5:0.95 gibi önemli genel başarı değerleri elde ederek en üst düzey performansı göstermiştir. Chen vd. (2023) tarafından sunulan YOLOv8 mimarisi, endüstriyel görüsel denetim senaryoları için optimize edilmiştir ve bu çalışmadaki sonuçlar da bu görüşü desteklemektedir. Ek olarak, Guo vd. (2023) çalışmasında da YOLOv8 tabanlı bir modelin akıllı şantiye uygulamalarında başarıyla kullanılabileceği vurgulanmıştır.

Sistemin genel yapısı, Ahmed vd. (2023) tarafından geliştirilen sürdürülebilir iş güvenliği yaklaşımlarıyla paralellik göstermektedir. Her üç sistem de yapay zeka ve otomasyon teknolojilerini entegre ederek, insan faktörüne bağlı hataların azaltılmasını ve sahada anlık izleme yeteneğinin artmasını amaçlamaktadır.

Bununla birlikte, bu çalışmanın bazı sınırlılıkları da mevcuttur. Kullanılan veri setleri sabit kamera açılarından elde edilmiş olup, farklı kamera perspektifleri, değişken aydınlatma koşulları veya yoğun kalabalık ortamlar gibi daha karmaşık senaryolar test edilmemiştir. Delhi vd. (2020) ve Gholami vd. (2020) tarafından yapılan çalışmalarda bu tarz çeşitli ortamlarda sistemlerin performansları daha genelleyici şekilde değerlendirilmiştir.

Ayrıca, bu çalışmadaki sistem yalnızca durağan görüntüler üzerinden test edilmiş olup, gerçek zamanlı video akışı üzerinde performans değerlendirmesi yapılmamıştır. Zhao vd. (2022), video bazlı sistemlerin daha farklı zorluklara sahip olduğunu ortaya koymaktadır.

Sonuç olarak, bu çalışmada sunulan sistem, hem literatürdeki çalışmalara paralel bir başarı düzeyi sunmuş hem de farklı nesne tespiti mimarilerinin özelliklerinin ayrıntılı biçimde analiz edilmesine olanak tanımıştır. Genişletilmiş veri setleri, video tabanlı analizler ve daha çeşitli saha koşullarında yapılacak testler ile bu sistemin pratik uygulanabilirliği ve sürekliliği daha sağlam biçimde ortaya konabilecektir.

8. SONUÇ VE ÖNERİLER

8.1. Sonuçlar

Bu çalışmada İSG kapsamında çalışan bireylerin KKD kullanımının tespit edilmesine yönelik, yapay zeka tabanlı bir görüntü işleme sistemi geliştirilmiştir. Sistem; YOLOv5, YOLOv7 ve YOLOv8 gibi derin öğrenme temelli nesne tespiti algoritmaları kullanılarak tasarlanmış ve farklı sahne koşullarında elde edilen işçi görselleri üzerinde eğitilip test edilmiştir. Bu bağlamda, sistemin performansı kesinlik, duyarlılık ve mAP gibi standart metriklerle değerlendirilmiş ve modellerin üstünlükleri karşılaştırmalı olarak analiz edilmiştir.

Elde edilen bulgular, her üç modelin de yüksek doğruluk oranları sunduğunu ve iş güvenliğinde kullanılabilecek potansiyele sahip olduklarını göstermektedir. YOLOv5 modeli, kısa sürede kararlı bir öğrenme eğrisi sergileyerek istikrarlı doğrulama sonuçları üretmiştir. YOLOv7 ise özellikle duyarlılık oranında daha başarılı sonuçlar vererek daha fazla sayıda KKD ögesini tespit edebilmiştir. En iyi genel performans başarısı ise YOLOv8 modeliyle elde edilmiştir. Bu model, %70'e varan mAP@0.5:0.95 değeriyle konum doğruluğu ve sınıf ayırımında en yüksek başarıyı göstermiştir.

Bu sonuçlar, yapay zeka temelli nesne tespit sistemlerinin sahada kullanılabilirliğini ve insan denetimine alternatif oluşturabilecek düzeyde başarılı çalışabildiğini ortaya koymaktadır. Böylece iş kazalarının önlenmesine, denetim süreçlerinin hızlanmasına ve işyeri güvenliğinin sürdürülebilirliğine katkı sağlanabileceği anlaşılmıştır.

8.2. Öneriler

Çalışmada geliştirilen sistem başarılı sonuçlar üretmiş olsa da daha kapsamlı ve esnek bir yapı oluşturulması için aşağıdaki öneriler dikkate alınabilir:

- Veri çeşitliliği artırılmalıdır. Modelin genelleme gücünü artırmak için farklı sektörlerden, farklı ışık ve çevre koşullarında çekilmiş daha zengin ve dengeli bir veri seti oluşturulmalıdır.
- Gerçek zamanlı video akışında test yapılmalıdır. Bu çalışmada yalnızca durağan görseller kullanılmıştır. Sistem, video tabanlı sahnelerde de test edilerek sahaya uygulanabilirliği pratik olarak değerlendirilmelidir.

- Farklı mimarilerle karşılaştırmalı çalışmalar yapılmalıdır. EfficientDet, RetinaNet veya Transformer tabanlı modeller gibi alternatif yapılarla performans karşılaştırmaları yapılarak daha optimum çözümler araştırılabilir.
- Gömülü sistem entegrasyonu sağlanmalıdır. Raspberry Pi, Jetson Nano gibi düşük maliyetli donanımlara entegrasyon testleri yapılarak sistemin taşınabilirliği ve mobil saha uygulamaları için uygunluğu artırılabilir.
- Semantik segmentasyon gibi ileri yöntemlerle desteklenmelidir. Nesne tespitine ek olarak segmentasyon yöntemleriyle her bir nesnenin tam alanı belirlenebilir; bu sayede daha hassas analizler yapılabilir.
- İşyeri özelinde özelleştirilmiş modeller geliştirilebilir. Her sektör için yaygın kullanılan KKD türlerine göre etiketleme ve eğitim süreçleri özelleştirilerek, modele göre optimize edilmiş sistemler kurulabilir.

Bu öneriler doğrultusunda yapılacak ileri çalışmalar, iş güvenliği denetimlerinde daha akıllı, daha hızlı ve daha etkin sistemlerin kurulmasına olanak sağlayacak; yapay zeka destekli denetim sistemlerinin sektörel yaygınlığını artıracaktır.

KAYNAKÇA

- Ahmed, M. I. B., Saraireh, L., Rahman, A., Al-Qarawi, S., Mhran, A., Al-Jalaoud, J., Al-Mudaifer, D., Al-Haidar, F., AlKhulaifi, D., Youldash, M., & Gollapalli, M. (2023). Personal protective equipment detection: A deep-learning-based sustainable approach. *Sustainability*, 15(18), 13990. <https://doi.org/10.3390/su151813990>
- Alpaydin, E. (2020). *Introduction to Machine Learning* (4th ed.). MIT Press.
- Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv preprint arXiv:2004.10934. <https://arxiv.org/abs/2004.10934>
- Chen, X., Wang, L., & Zhao, Y. (2023). Lightweight YOLOv8 for industrial visual inspection tasks. *Sensors*, 23(4), 1988. <https://doi.org/10.3390/s23041988>
- Cheng, Y., Lin, W., Bai, J., Zhang, Y., & Bai, C. (2023). Real-Time Detection of Personal Protective Equipment in Construction Sites Using YOLOv5 and Deep Learning. *Sustainability*, 15(17), 13990. <https://doi.org/10.3390/su151713990>
- Delhi, V. S. K., Sankarlal, R., & Thomas, A. (2020). Detection of personal protective equipment (PPE) compliance on construction site using computer vision based deep learning techniques. *Frontiers in Built Environment*, 6, 136. <https://doi.org/10.3389/fbuil.2020.00136>
- Ding, X., Zhang, X., Han, J., & Ding, G. (2021). RepVGG: Making VGG-style ConvNets Great Again. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 13733–13742. <https://doi.org/10.1109/CVPR46437.2021.01353>
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Hounsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.
- Fang, Q., Ding, L., Luo, H., & Love, P. E. (2018). Falls from heights: A computer vision-based approach for safety harness detection. *Automation in Construction*, 91, 53–61. <https://doi.org/10.1016/j.autcon.2018.02.022>
- Gholami, M. F., Faeghi, S., & Gharehbaghi, A. (2020). Computer vision-based PPE detection using deep learning for construction safety monitoring. *Journal of Safety Research*, 75, 1–11.
- Girshick, R. (2015). Fast R-CNN. *Proceedings of the IEEE International Conference on Computer Vision*, 1440–1448.

- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 580–587.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Cambridge, MA: MIT Press.
- Guo, J., Li, Y., & Zhang, W. (2023). Real-time PPE detection in smart construction using YOLOv8-based deep learning. *Automation in Construction*, 153, 105099. <https://doi.org/10.1016/j.autcon.2023.105099>
- Han, K., & Zeng, X. (2021). Deep learning-based workers safety helmet wearing detection on construction sites using multi-scale features. *Ieee Access*, 10, 718-729.
- Hanyao, M., Jin, Y., Qian, Z., Zhang, S., & Lu, S. (2021, May). Edge-assisted online on-device object detection for real-time video analytics. In *IEEE INFOCOM 2021-IEEE Conference on Computer Communications* (pp. 1-10). IEEE.
- Hayat, A., & Morgado-Dias, F. (2022). Deep learning-based automatic safety helmet detection system for construction safety. *Applied Sciences*, 12(16), 8268. <https://doi.org/10.3390/app12168268>
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9), 1904–1916.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.
- He, Q., Lin, J., & Fu, Y. (2021). Smart PPE compliance detection for construction safety using YOLOv5. *Automation in Construction*, 129, 103843. <https://doi.org/10.1016/j.autcon.2021.103843>
- Howard, A., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., ... & Le, Q. V. (2019). Searching for MobileNetV3. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 1314–1324.
- ILO. (2022). Safety and health at the heart of the future of work: Building on 100 years of experience. <https://www.ilo.org>
- Jocher, G., Chaurasia, A., Qiu, J., & Stoken, A. (2020). YOLOv5. GitHub Repository. <https://github.com/ultralytics/yolov5>

- Jocher, G., Chaurasia, A., Qiu, J., & Ultralytics. (2023). YOLOv8: Cutting-edge object detection and image segmentation model. Retrieved from <https://github.com/ultralytics/ultralytics>
- Johnson, J. M., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, 6(1), 1–54. <https://doi.org/10.1186/s40537-019-0192-5>
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15–25.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 1097–1105.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Li, C., Liu, C., Kang, W., & Zhang, Y. (2018). Multi-scale object detection in remote sensing imagery with convolutional neural networks. https://www.researchgate.net/publication/324903264_Multi-scale_object_detection_in_remote_sensing_imagery_with_convolutional_neural_networks
- Li, W., Liu, Y., & Wang, Z. (2022). Application of YOLOv5 in real-time detection of railway defects. *Computational Intelligence and Neuroscience*, 2022, 1–11. <https://doi.org/10.1155/2022/7236502>
- Li, Y., Zhang, M., & Zhao, Q. (2023). Real-time detection of protective gear using YOLOv7 in construction sites. *Journal of Safety Research and Technology*, 5(2), 85–93.
- Liang, H., Yang, L., Chen, J., Liu, X., & Hang, G. (2024). Detection and tracking of safety helmet wearing based on deep learning. *Open Computer Science*, 14(1), 20240017.
- Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2117–2125).
- Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. *Proceedings of the IEEE International Conference on Computer Vision*, 2980–2988.
- Liu, H., & Qin, X. (2024). *Target Detection of Safety Protective Gear Using the Improved YOLOv5*. arXiv preprint arXiv:2408.05964. <https://doi.org/10.48550/arXiv.2408.05964>

- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. *European Conference on Computer Vision*, 21–37.
- Liu, Y., Wang, L., Zhang, Z., Hou, D., & Xue, W. (2020). Deep learning-based object detection algorithms: A survey. *Pattern Recognition*, 110, 107598.
- Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., ... & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 10012–10022.
- Mazurov, M. (2021). Railroad Worker Detection Dataset. Kaggle. <https://www.kaggle.com/datasets/mikhailma/railroad-worker-detection-dataset>
- Meituan. (2022). YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications. GitHub repository. <https://github.com/meituan/YOLOv6>
- Misra, D. (2019). Mish: A Self Regularized Non-Monotonic Neural Activation Function. *arXiv preprint arXiv:1908.08681*. <https://arxiv.org/abs/1908.08681>
- Molchanov, P., Tyree, S., Karras, T., Aila, T., & Kautz, J. (2017). Pruning convolutional neural networks for resource efficient inference. *arXiv preprint arXiv:1611.06440*. <https://doi.org/10.48550/arXiv.1611.06440>
- Özcan, K. (2024). Convolutional Neural Networks (CNN) nedir? Medium. https://medium.com/@kubra.ozcan_98680/convolutional-neural-networks-cnn-nedir-a3a9d5318835
- Park, S. S., Tran, V. T., & Lee, D. E. (2021). Application of various YOLO models for computer vision-based real-time pothole detection. *Applied Sciences*, 11(23), 11229.
- Pham, T. (2018). [Review] SPPnet — 1st runner-up (Object Detection) & 2nd runner-up (Image Classification) in ILSVRC. Medium. <https://medium.com/coinmonks/review-sppnet-1st-runner-up-object-detection-2nd-runner-up-image-classification-in-ilsvrc-906da3753679>
- Redmon, J., & Farhadi, A. (2017). YOLO9000: Better, Faster, Stronger. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 7263–7271. <https://doi.org/10.1109/CVPR.2017.690>
- Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *arXiv preprint arXiv:1804.02767*. <https://arxiv.org/abs/1804.02767>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788. <https://doi.org/10.1109/CVPR.2016.91>

- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, 28, 91–99.
- Russell, S., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach* (3rd ed.). Pearson Education.
- Seong, S., Song, J., Yoon, D., Kim, J., & Choi, J. (2019). Determination of vehicle trajectory through optimization of vehicle bounding boxes using a convolutional neural network. *Sensors*, 19(19), 4263. <https://doi.org/10.3390/s19194263>
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Proceedings of the IEEE International Conference on Computer Vision*, 618–626.
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 60. <https://doi.org/10.1186/s40537-019-0197-0>
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Song, X., Zhang, T., & Yi, W. (2024). An improved YOLOv8 safety helmet wearing detection network. *Scientific reports*, 14(1), 17550.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1–9).
- Szeliski, R. (2010). *Computer Vision: Algorithms and Applications*. Springer.
- Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *International Conference on Machine Learning*, 6105–6114.
- Tan, M., Pang, R., & Le, Q. V. (2020). EfficientDet: Scalable and efficient object detection. *CVPR*, 10781–10790.
- Tao, N., Shengteng, D., & Xiangkun, J. (2023). Highway forecasting for weather factor and traffic flow interaction scenarios. *ResearchGate*. https://www.researchgate.net/publication/374632617_Highway_forecasting_for_weather_factor_and_traffic_flow_interaction_scenarios
- Ultralytics. (2023). YOLOv8 Documentation. Retrieved from <https://docs.ultralytics.com/>
- Vinayak, N. (2020). Fast R-CNN. *Medium*. <https://naveenvinayak.medium.com/fast-rcnn-5c86e648e054>
- Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2020). CSPNet: A New Backbone That Can Enhance Learning Capability of CNN. *Proceedings of the IEEE/CVF Conference on*

- Computer Vision and Pattern Recognition Workshops, 390–391.
<https://arxiv.org/abs/1911.11929>
- Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv preprint arXiv:2207.02696*.
<https://doi.org/10.48550/arXiv.2207.02696>
- Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2021). Scaled-YOLOv4: Scaling Cross Stage Partial Network. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 13029–13038.
- Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., & Hu, Q. (2020). ECA-Net: Efficient channel attention for deep convolutional neural networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 11531–11539.
https://openaccess.thecvf.com/content_CVPR_2020/papers/Wang_ECA-Net_Efficient_Channel_Attention_for_Deep_Convolutional_Neural_Networks_CVPR_2020_paper.pdf
- Wang, Z., Wu, Y., Yang, L., Thirunavukarasu, A., Evison, C., & Zhao, Y. (2021). Fast personal protective equipment detection for real construction sites using deep learning approaches. *Sensors*, 21(10), 3478. <https://doi.org/10.3390/s21103478>
- Xu, S., Wang, X., Lv, W., Chang, Q., Cui, C., Deng, K., ... & Lai, B. (2022). PP-YOLOE: An evolved version of YOLO. *arXiv preprint arXiv:2203.16250*.
- Yipeng, L., & Junwu, W. (2024). Personal protective equipment detection for construction workers: A novel dataset and enhanced YOLOv5 approach. *IEEE Access*.
- Zhang, J. (2020). YOLOv1 explained by A.J. Medium.
<https://medium.com/@jesse419419/yolov1-explained-by-a-j-7fe408764c55>
- Zhang, M., He, J., & Liu, Y. (2023). An intelligent detection system for safety gear wearing using YOLOv5. *Applied Sciences*, 13(5), 2780. <https://doi.org/10.3390/app13052780>
- Zhang, S., Chi, C., Yao, Y., Lei, Z., & Li, S. Z. (2020). Bridging the Gap Between Anchor-Based and Anchor-Free Detection via Adaptive Training Sample Selection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 9759–9768. <https://doi.org/10.1109/CVPR42600.2020.00978>
- Zhang, X., Wang, Y., & Zhang, D. (2022). Enhancing YOLOv4 for Real-Time Object Detection. **Journal of Real-Time Image Processing*, 19*, 805–817.
<https://doi.org/10.1007/s11554-021-01147-4>
- Zhai, S., Cheng, Y., Lu, W., & Zhang, Z. (2016). Mixed pooling for convolutional neural networks.

https://www.researchgate.net/publication/300020038_Mixed_Pooling_for_Convolutional_Neural_Networks

- Zhao, C., Zhang, L., & Wang, Y. (2022). Lightweight single shot object detection based on MobileNetV3 for embedded systems. *Sensors*, 22(11), 4031.
- Zhao, H., Zhang, Y., & Li, Z. (2022). Real-time helmet detection for construction workers based on YOLOv5. *Journal of Safety Science and Technology*, 18(3), 45–52.
- Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems*, 30(11), 3212–3232.
- Zhou, X., Wang, D., & Krähenbühl, P. (2019). Objects as Points. *arXiv preprint arXiv:1904.07850*. <https://arxiv.org/abs/1904.07850>
- Zhou, Z., Sodhi, H., Majumdar, A., & Dey, D. (2020). Domain adaptation in computer vision applications. *Neurocomputing*, 412, 418–431. <https://doi.org/10.1016/j.neucom.2020.06.009>