DERİN ÖĞRENME YAKLAŞIMIYLA GERÇEK ZAMANLI GÖRÜNTÜLER ÜZERİNDE TEHLİKELİ NESNE TESPİTİ

Şuayip Aykut Atmaca

T.C.
BURSA ULUDAĞ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

DERİN ÖĞRENME YAKLAŞIMIYLA GERÇEK ZAMANLI GÖRÜNTÜLER ÜZERİNDE TEHLİKELİ NESNE TESPİTİ

Şuayip Aykut ATMACA
0000-0002-7146-641X

Doç. Dr. Burcu ÇAĞLAR GENÇOSMAN
(Danışman)
0000-0003-0159-8529

YÜKSEK LİSANS
ENDÜSTRİ MÜHENDİSLİĞİ ANABİLİM DALI

BURSA – 2024
Her Hakkı Saklıdır

ÖZET

Yüksek Lisans Tezi

DERİN ÖĞRENME YAKLAŞIMIYLA GERÇEK ZAMANLI GÖRÜNTÜLER ÜZERİNDE TEHLİKELİ NESNE TESPİTİ

Şuayip Aykut ATMACA

Bursa Uludağ Üniversitesi
Fen Bilimleri Enstitüsü
Endüstri Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Burcu ÇAĞLAR GENÇOSMAN

Teknolojinin gelişmesiyle birlikte, yapay zekâ alanında yapılan çalışmalarda ciddi bir artış meydana gelmiştir. Toplumsal yaşamın her alanında yapay zekâ çalışmaları gerçekleştirilmekle beraber güvenlik sistemleri alanında da çalışmalar yapılmaktadır. Bu alan ile ilgili çok karşılaşılan konulardan biri, güvenlik kameraları üzerinde suçun tespit edilmesidir. Bu ihtiyacın en önemli sebebi, günümüzde güvenlik kameralarının yaygın olarak kullanılmasına rağmen suçun ve suçlunun tespiti için insan operatörlerin uzun saatler boyunca ilgili bölgeyi farklı açılardan görüntüleyen video kayıtlarını izlemek zorunda olmasıdır. Bu çalışmada, tehlikeli nesnelerle işlenen suçları derin öğrenme yöntemleri ile güvenlik kameraları üzerinde gerçek zamanlı tespit edebilmek ve ilgili kişileri anlık olarak uyarabilmek amaçlanmıştır. Çalışma kapsamında tehlikeli nesnelerden ‘tabanca’, ‘tüfek’ ve ‘bıçak’ sınıf etiketleri için herkese açık veri setleri kullanılarak popüler derin öğrenme mimarilerinden YOLOv5, YOLO NAS, YOLOv8, YOLOv9 ve MobileNet mimarileri kullanılmıştır. Eğitimler sonucunda en başarılı model seçilerek web tabanlı bir uygulama geliştirilmiş ve gerçek zamanlı olarak uygulama üzerinde ön tanımlı kişilere tehlike durumunda bildirim gönderebilmek hedeflenmiştir. Çalışmadan beklenen fayda, insan kaynaklı hataların önüne geçilmesi ve güvenlik personeli maliyetlerinin azaltılmasının yanında suçun gerçek zamanlı tespitiyle birlikte önlenebilme olasılığını arttırmaktır.

Anahtar Kelimeler: Nesne Tespiti, Derin Öğrenme, YOLO, Mediapipe, Güvenlik, Yapay Zeka
2024, xiv + 40 sayfa.

ABSTRACT

MSc Thesis

DETECTION OF DANGEROUS OBJECTS IN REAL-TIME IMAGES USING A DEEP LEARNING APPROACH

Şuayip Aykut ATMACA

Bursa Uludağ University
Graduate School of Natural and Applied Sciences
Department of Industrial Engineering

Supervisor: Doç Dr Burcu ÇAĞLAR GENÇOSMAN

With the advancement of technology, there has been a significant increase in studies in artificial intelligence. While AI research is being conducted in all areas of social life, studies are also being carried out in the field of security systems. One of the frequently encountered topics in this area is the detection of crime on security cameras. The main reason for this need is that, despite the widespread use of security cameras today, human operators still have to watch video recordings of the relevant area from different angles for long hours to detect crimes and criminals. This study aims to detect crimes committed with dangerous objects on security cameras in real-time using deep learning methods and instantly alert the relevant individuals. Within the scope of the study, popular deep learning architectures such as YOLOv5, YOLO NAS, YOLOv8, YOLOv9, and MobileNet were used with public datasets for the class labels of dangerous objects: ‘gun’, ‘rifle’, and ‘knife’. After the training, the most successful model was selected, and a web-based application was developed to send notifications to predefined individuals in case of danger in real-time. The expected benefit of the study is to increase the possibility of preventing crime through real-time detection, as well as reducing human errors and security personnel costs.

Key words: Object Detection, Deep Learning, YOLO, Mediapipe, Security, Artificial Intelligence
2024, xiv + 40 pages.

TEŞEKKÜR

Yüksek lisans eğitimim ve tez çalışmalarım boyunca her türlü desteği sağlayan, bilgi ve tecrübesi ile çalışmalarıma katkı sağlayan danışman hocam Doç. Dr. Burcu ÇAĞLAR GENÇOSMAN'a teşekkürlerimi sunarım.

Eğitim hayatım boyunca maddi ve manevi desteklerini hiçbir zaman esirgemeyen değerli aileme teşekkürlerimi sunarım.

Şuayip Aykut ATMACA
28/06/2024

İÇİNDEKİLER

Sayfa

ÖZET ……...viii
ABSTRACT ……ix
TEŞEKKÜR ……...x
SİMGELER ve KISALTMALAR DİZİNİ ……xii
ŞEKİLLER DİZİNİ……xiii
ÇİZELGELER DİZİNİ ……xiv
1.GİRİŞ ……...1
2.KURAMSAL TEMELLER VE KAYNAK ARAŞTIRMASI……7
2.1. Kaynak Özetleri ……7
2.2. Yapay Zekâ ……9
2.3. Makine Öğrenmesi ……9
2.4. Yapay Sinir Ağları ……10
2.5. Derin Öğrenme……10
2.6. YOLO Modelleri……13
2.7. MEDIAPIPE Nesne Tespit Modelleri……18
2.8. Performans Metrikleri ……20
3. MATERYAL ve YÖNTEM……22
3.1. Veri Seti ……22
3.2. Veri Setinin Etiketlenmesi ……23
3.3. Kullanılan Yöntemler……25
3.4. Çalışma Ortamı ……26
4. BULGULAR ve TARTIŞMA……27
4.1. Güvenlik Sistemlerinde Kullanım Potansiyeli……32
5. SONUÇ ……36
KAYNAKLAR ……38
ÖZGEÇMİŞ ……41

SİMGELER ve KISALTMALAR DİZİNİ

Simgeler	**Açıklama**
bx, by, bw, bh	Sınırlayıcı kutu koordinat ve boyut bilgileri

Kısaltmalar	**Açıklama**
CCTV	Kapalı Devre Televizyon (Closed Circuit Television)
CNN	Evrişimli Sinir Ağı (Convolutional Neural Network)
GPU	Grafik İşleme Birimi (Graphics Processing Unit)
IoU	Birleşim Üzerinden Kesişim
R-CNN	Bölge Tabanlı CNN (Region Based CNN)
YOLO	Yalnızca Bir Kere Bakarsın (You Only Look Once)
YSA	Yapay Sinir Ağı

ŞEKİLLER DİZİNİ

		Sayfa
Şekil 1.1.	2022 yılı ABD soygun vakalarında kullanılan silah türü adetleri (Number of Robberies in the U.S. by Weapon 2022, t.y.)……..	3
Şekil 1.2.	2003-2015 yılları arası İngiltere ve Galler'de işlenen suçlara göre kullanılan silah tipleri (Offences involving the use of weapons - Office for National Statistics, t.y.)……..	3
Şekil 2.1.	Yapay zekâ disiplinleri hiyerarşisi ……..	9
Şekil 2.2.	Biyolojik ve yapay sinir ağı gösterimleri…….	10
Şekil 2.3.	2000-2023 yılları arası "Nesne Tespiti" alanında yapılan akademik çalışma artışı (Veriler Science Direct arama motoru içinde "Object Detection" anahtar sözcüğü kullanılarak elde edilmiştir.)……	11
Şekil 2.4.	R-CNN Mimarisi Gösterimi(What Is R-CNN?, 2023)……	13
Şekil 2.5.	YOLO Algoritması Tespit Sistemi(Redmon vd., 2016)……..	14
Şekil 2.6.	YOLOv9 Coco veri seti üzerinde elde edilen metrikler	17
Şekil 2.7.	YOLOv9 Ağı Gösterimi	17
Şekil 2.8.	MobileNetV2 konvolüsyon blokları gösterimi (Dong vd., 2020)	18
Şekil 2.9.	Multi-AVG(Solda) ve Multi-MAX(Sağda) model gösterimleri(Chu vd., 2021)…….	19
Şekil 2.10.	Karmaşıklık Matrisi(Confusion Matris)……..	20
Şekil 3.1.	'Tabanca' sınıf etiketine ait örnek görseller…….	22
Şekil 3.2.	'Bıçak' sınıf etiketine ait örnek görseller……..…	23
Şekil 3.3.	'Tüfek' sınıf etiketine ait örnek görseller……..……	23
Şekil 3.4.	Roboflow uygulaması üzerinde "Tabanca" sınıfına ait örnek bir etiketleme işlemi…….	24
Şekil 3.5.	Örnek Pascal VOC yapısı…….	24
Şekil 3.6.	Örnek Yolo etiket yapısı……..	25
Şekil 4.1.	YOLOv5 modeli validasyon veri seti kullanılarak elde edilmiş karışıklık matrisi……..	27
Şekil 4.2.	YOLOv8 modeli validasyon veri seti kullanılarak elde edilmiş karışıklık matrisi……..	28
Şekil 4.3.	YOLOv9 modeli validasyon veri seti kullanılarak elde edilmiş karışıklık matrisi……..	29
Şekil 4.4.	YOLOv8 kullanılarak elde edilmiş 3 ayrı sınıf etiketi için F1 eğrisi grafiği…….	31
Şekil 4.5.	YOLOv9 kullanılarak elde edilmiş 3 ayrı sınıf etiketi için F1 eğrisi grafiği…….	32
Şekil 4.6.	Eğitilen YOLOv9 modelinin kendi oluşturduğumuz bir veri seti örneklemi üzerinde tahminleme performansı……	33
Şekil 4.7.	Normal durumu temsil eden uygulama arayüz görseli……	34
Şekil 4.8.	Alarm durumunu temsil eden uygulama arayüz görseli……..	35

ÇİZELGELER DİZİNİ

Sayfa

Çizelge 4.1. Farklı Dedektörlerin eğitim veri seti üzerindeki eğitim performansları……. 30
Çizelge 4.2. Farklı Dedektörlerin validasyon veri seti üzerindeki tahmin performansları……. 30

1. GİRİŞ

Suç oranlarının yaşam doyumu üzerinde olumsuz etkileri vardır ve suç oranları yüksek topluluklarda toplum memnuniyetinin refah üzerindeki olumlu etkisinin daha az olduğu görülmektedir (Pak & Gannon, 2023).

Dünyada artan nüfusa paralel olarak işlenen suç miktarları da gün geçtikçe artmaktadır. Federal Soruşturma Bürosunun (FBI) paylaştığı istatistiklere göre 2019 yılında dünya çapında 267.988 soygun vakası meydana gelmiştir ve toplam zarar 482.000.000 dolardır (*Robbery*, t.y.).

Silahlı şiddet olayları, toplumu, sağlık psikolojik ve ekonomik olarak olumsuz yönde etkilemektedir. Her yıl binlerce insan silahlı saldırılar sonucu hayatını kaybetmekte ya da maddi zarara uğramaktadır. Ayrıca silah kullanımı toplumlarda çocuk psikolojisi üzerinde de travmalara yol açmakta ve ilerleyen yaşlarda kişilik bozuklukları ve silahlı şiddete meyilli olma olaylarına sebep olmaktadır bu açıdan silahlı saldırıların erken tespiti, önlenebilirlik açısından yüksek önem taşımaktadır (Velastin vd., 2006; Bilgin, 2016).

Geçmişte, görüntüleme ve izleme sistemlerinin olmadığı dönemlerde, suçun ve suçlunun tespiti için görgü tanıklarının ifadeleri doğrultusunda oluşturulan robot resimler kullanılmaktadır. Bu yöntemler zamansal kısıtlar, düşük doğruluk oranları, olay yerinde şahit bulunmaması, şahitlerin tehdit edilmesi gibi kısıtlardan dolayı yüksek verim sağlayamamaktadır. Günümüzde ise, suçun ve suçlunun tespit edilmesi ve önlenmesi faaliyetlerinde, güvenlik kameraları/kapalı devre televizyon (Close Circuit TeleVision - CCTV) yaygın olarak kullanılmaktadır.

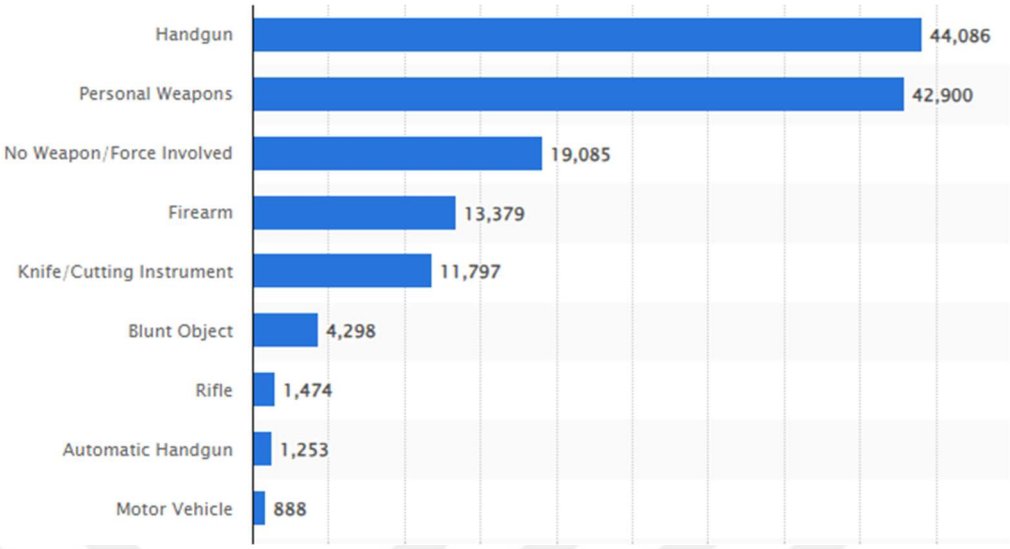
Yapılan tahminlere göre, polis teşkilatlarının ortalama %87'lik bir kısmı güvenlik kameralarını kullanarak suç tespit etmeye çalışmaktadır. Teknolojinin gelişimi ve kamera sistemlerine erişimin kolaylaşması ile kamusal alanlarda güvenlik kameralarının kullanımı yaygınlaşmıştır. Güvenlik kameralarının suç tespitindeki etkinliği bilinen bir gerçektir. Güvenlik, suçun soruşturulması ve tespiti mevut durumda kamusal alanlarda

kullanılan güvenlik kameralarının ana hedefleridir. Hukuksal soruşturmalarda güvenlik kamera görüntüleri en önemli delil kaynakları olarak kabul edilmektedir. Yine kolluk kuvvetleri tarafından da suç mahali araştırmalarında en çok kullanılan kaynak güvenlik kameralarıdır (K U & M, 2023).

Güvenlik kameraları sayesinde muhtemel suç bölgeleri 7 gün 24 saat boyunca izlenip kayıt altına alınabilmektedir ve suçun ve suçlunun tespitinde başarılı sonuçlar elde edilmektedir. Ancak bu işlem oldukça maliyetli bir süreçtir. Çünkü bir suç tespiti için izlenmesi gereken binlerce saatlik güvenlik kamerası görüntüsü kaydı bulunabilir ve bunları izlemek için personel maliyetlerine katlanmak zorunda kalınabilir. Bu süreçteki diğer bir risk ise kayıtları izleyen personelin dikkat eksikliği nedeniyle gözden kaçırdığı durumlardır. Bu alanda yapılan bazı çalışmalar, video kayıtları izleyen insan operatörlerin 20-40 dakikalık bir yoğun izleme süreci sonrasında video körlüğü yaşadığı ve ekrandaki aksiyonun 95%'lik bir kısmını kaçırdığını ortaya koymaktadır (Velastin vd., 2006).

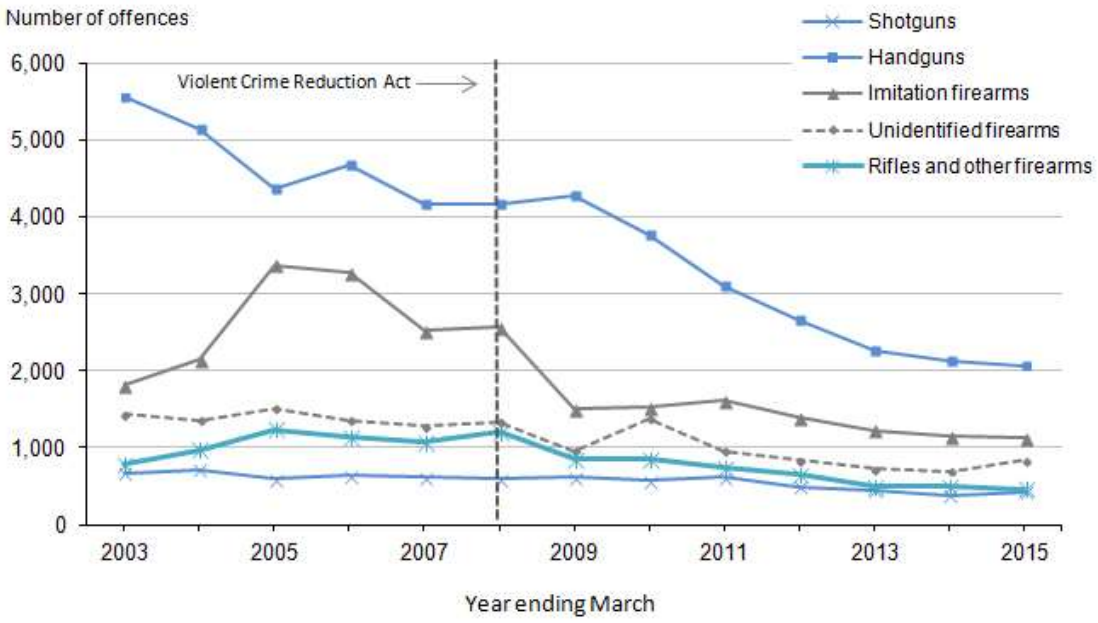
Son yıllarda, gelişen yapay zekâ ve görüntü işleme teknolojisi ile yapılan çalışmalar, insan kaynaklı hataların ve personel maliyetlerinin önüne geçmekte ve yüksek başarı oranı göstermektedir. Derin öğrenme ile nesne tespiti alanında meydana gelen son gelişmeler, söz konusu denetim sürecinin otomatikleştirilmesine önemli bir katkı sağlamaktadır. Derin öğrenme yöntemleri ile elde edilen görüntü veri setleri üzerindeki zararlı nesnelerin özellikleri çıkarılıp öğrenilebilir ve görüntü üzerinde istenen nesneler tespit edilebilir.

İstatistikler, dünyanın belirli bölgelerinde, özellikle de ateşli silah bulundurmanın yasak olduğu ülkelerde, silahla ilgili önemli sayıda suç yaşandığını göstermektedir (*Estimating Global Civilian-Held Firearms Numbers*, 2018). Amerikan Adalet bürosu tarafından yapılan bir çalışmaya göre silahlı suçların 86%'si tabanca ile işlenmiştir (Zawitz, t.y.). Şekil 1.1'de görüldüğü üzere tespit edilecek silah türleri içerisinde tahmin edilecek sınıf olarak "tabanca" diğer tiplere göre daha büyük bir öneme sahiptir.



Şekil 1.1. 2022 yılı ABD soygun vakalarında kullanılan silah türü adetleri (*Number of Robberies in the U.S. by Weapon 2022*, t.y.)

Şekil 1.2'de görüleceği üzere, "Birleşik Krallık Ulusal İstatistik Ofisi" verilerine göre silahlı suç olaylarında en çok kullanılan saldırı aracı tabancadır. Tabancaya oransal olarak en yakın değere sahip saldırı aracı ise sahte silahlar (*imitation firearms*) olmuştur.



Şekil 1.2. 2003-2015 yılları arası İngiltere ve Galler'de işlenen suçlara göre kullanılan silah tipleri (*Offences involving the use of weapons - Office for National Statistics*, t.y.)

Tehlikeli nesne tespitinde kullanılan derin öğrenme yöntemleri son yıllarda üstel olarak artmıştır. Bu yöntemlerin başında evrişimli sinir ağları (convolutional neural network - CNN), daha hızlı bölge tabanlı evrişimli sinir ağları (faster-RCNN) ve YOLO gelmektedir. YOLO- you only look once, Ultralytics (https://www.ultralytics.com) tarafından geliştirilen bir derin öğrenme ağıdır. YOLO (You Only Look Once), bilgisayarla görme alanında devrim yaratan, son teknoloji ürünü, gerçek zamanlı olarak nesne tespiti yapabilen bir derin öğrenme mimarisidir. Bölge önerisi ve sınıflandırma için aşamalı bir tespit yapısı kullanan geleneksel yöntemlerin aksine YOLO, nesne algılama görevini tek bir regresyon problemi olarak ele alır ve yalnızca tek aşamada görüntü üzerindeki sınırlayıcı kutuları ve sınıf olasılıklarını tahmin eder. Bu yapı sayesinde YOLO, geleneksel nesne tespit modellerine göre nesne tespit işlemini daha hızlı bir şekilde gerçekleştirir.

Zararlı nesne tespiti, görüntüler veya video akışları içindeki potansiyel olarak tehlikeli veya tehdit edici nesneleri tanımaya odaklanan özel bir bilgisayarlı görü alanıdır. Bu teknoloji, silahlar, tehlikeli maddeler veya güvenlik açısından risk oluşturabilecek diğer nesneleri tespit etmek için genellikle YOLO gibi derin öğrenme modelleri olmak üzere gelişmiş makine öğrenimi algoritmaları kullanır. Tespit süreci, sistemin bu öğeleri çeşitli bağlamlarda ve ortamlarda doğru bir şekilde algılamasını sağlamak için zararlı nesnelerin etiketli görüntülerini içeren kapsamlı veri kümeleri üzerinde eğitim modellerini içerir. Bu yetenek, tehditlerin hızlı ve güvenilir bir şekilde tespit edilmesinin önemli olduğu kamusal alanlarda, ulaşım merkezlerinde ve kısıtlı alanlarda güvenlik önlemlerinin artırılması için önem arz etmektedir.

Bu çalışmanın amacı, anlık olarak güvenlik kameraları üzerinden elde edilen görüntüler üzerinde derin öğrenme yöntemleri ile suçun ve suç unsurlarının tespiti, uyarıcı sistemlerin aktive edilmesi ve önceden belirlenmiş yetkili kişilerin bildirim ile uyarılmasıdır. Günümüzde güvenlik tehditlerinin hızla artması, güvenlik sistemlerinin de eş zamanlı olarak gelişmesini zorunlu kılmıştır. Geleneksel yöntemler, genellikle suç meydana geldikten sonra devreye girerken, bu çalışma, suçun anlık olarak tespit edilip olaya daha hızlı müdahale edilebilmesine odaklanmaktadır.

Görüntü işleme ve derin öğrenme yöntemleri, kamera görüntülerindeki tehlikeli nesneleri tespit edebilme kapasitesine sahiptir. Bu çalışmada geliştirilecek sistem, tehlikeli nesne tespitini gerçekleştirdikten sonra uyarı sinyalini ilgili kişi veya kurumlara iletecektir. Yapılan araştırmalar kapsamında, suç aleti olarak tabancanın diğer nesnelere oranla daha fazla kullanıldığı gözlemlenmiştir. Bu nedenle çalışmanın öncelikli amacı, kullanılacak sınıf etiketleri içerisinde ('tabanca', 'tüfek', 'bıçak') tabanca tespitinde yüksek doğruluk elde etmektir.

Çalışmada çeşitli derin öğrenme yöntemleri kullanılmış olup, literatürde bilindiği kadarıyla tehlikeli nesne tespitinde ilk kez YOLOv9 kullanılmıştır. YOLOv9, daha önceki versiyonlarına göre daha yüksek doğruluk ve hız sunmakta olup, gerçek zamanlı uygulamalarda kullanılabilmektedir.

Araştırma sürecinde, üç farklı veri seti birleştirilerek modeller eğitilmiştir. Bu veri setleri, farklı açılardan ve koşullardan elde edilen görüntülerden oluşmakta olup, farklı çalışmalarda da kullanılmıştır. Eğitim süreci, geniş ve çeşitli veri setleri ile yürütülmüş, böylece modelin gerçek dünya senaryolarına uyum sağlaması hedeflenmiştir. YOLOv9'un kullanımına ek olarak, farklı derin öğrenme algoritmaları da karşılaştırmalı olarak incelenmiş ve en uygun sonuçları veren yöntem seçilmiştir.

Çalışma kapsamında geliştirilen model, suçun erken tespiti ve önlenmesi hususunda gerekli donanımsal altyapı ile birlikte kullanılarak bir ürüne dönüştürülecektir. Bu ürün sayesinde, eğitilen yüksek doğruluk oranına sahip model, güvenlik kameralarından aldığı görüntüler üzerinde anlık olarak tehlikeli nesne tespiti yapıp, ilgili kişi ve kurumlara mesaj veya alarm yoluyla bildirim gönderecektir. Tehlikeli nesne tespitinde YOLOv9 destekli bir erken uyarı sistemi yazılımı bilindiği kadarıyla henüz mevcut değildir. Bu anlamda da bu çalışma bir ilk olacaktır. Bu sistemin özellikle yoğun nüfuslu ve yüksek suç oranlarına sahip bölgelerde, suç oranlarının azaltılmasında ve genel güvenlik seviyesinin artırılmasına katkı sağlayacağı düşünülmektedir.

Bunun yanı sıra, bu sistemin entegrasyonu ile güvenlik güçlerinin olaylara müdahale süreleri kısalacak, potansiyel tehlikeler daha erken tespit edilerek gerekli önlemler

alınabilecektir. Bu çalışmanın sonucunda elde edilen veriler ve geliştirilen yöntemler, güvenlik sektöründe önemli bir yenilik sağlayacak, ileriye dönük araştırmalar ve uygulamalar için de bir temel oluşturacaktır.

Tezin ilerleyişi şu şekildedir: Bölüm 2'de konu ile ilgili literatür çalışması verilmiş ve yapay zekadan derin öğrenmeye giden süreçte kullanılan yaklaşımlardan bahsedilmiştir. Bölüm 3'de çalışmada ele alınan veri setleri ve uygulanan görüntü işleme yöntemleri anlatılmıştır. Bölüm 4'de anlık suç tespitinde kullanılmak üzere eğitilen derin öğrenme ağları detaylandırılmıştır. Bölüm 5'de eğitim sonuçları incelenmiş ve gerçek bir veri seti kullanılarak tahminler gerçekleştirilmiştir. Bununla beraber, uyarı mekanizması anlatılmıştır. Bölüm 6'da ise çalışmanın bir özeti verilmiş ve yaygınlaştırılması halinde literatüre ve gerçek hayata katkısı anlatılmıştır.

2.KURAMSAL TEMELLER VE KAYNAK ARAŞTIRMASI

2.1. Kaynak Özetleri

Derin öğrenme, geniş uygulama alanı sebebiyle farklı nesne tanıma görevlerinde kullanılmaktadır. Örnek olarak yüz tespiti (Rowley vd., 1998), yaya tespiti (Dollar vd., 2012), araç tanıma (Fan vd., 2016), imalat sektöründe hata tespiti (Parlak & Emel, 2023), stok kontrolü (Babila vd., 2022) gibi uygulamalar verilebilir.

Evrişimli Sinir Ağları (CNN- Convolutional Neural Network) kullanılarak silah, bıçak vb. tehlikeli nesne tespiti, etkili gözetleme ve güvenlik önlemlerine yönelik artan talep nedeniyle günümüzde büyük ilgi görmektedir. Çeşitli ortamlarda, özellikle video görüntülerinde, silah, bıçak vb. nesnelerin otomatik olarak tespitinde CNN model uygulamalarını araştıran birçok akademik çalışma bulunmaktadır.

Verma vd. (2017) VGG16 ve Daha Hızlı RCNN (faster-RCNN) modelleri kullanarak herkese açık internet ateşli silahlı filmler veri tabanı (IMFDB) üzerinden elde ettikleri görselleri kullanarak tabanca tespiti yapmışlardır (Verma & Dhillon, 2017). Warsi vd. (2019) yaptıkları çalışmada, YOLOv3 algoritması kullanarak silah tespiti yapmışlardır (Warsi vd., 2019). Jesus vd. (2021) yaptıkları çalışmada, güvenlik kamerası görüntülerinden yararlanarak otomatik tabanca tespiti için üç farklı CNN modelini kullanan bir çalışma gerçekleştirmiş ve tabanca tespitinde yüksek başarı oranı elde etmişlerdir (Jesus vd., 2021). Narejo vd. (2021) yaptıkları çalışmada, akıllı güvenlik sistemleri için YOLOv3 algoritmasından yararlanarak silah tespit projesi geliştirmişlerdir (Narejo vd., 2021). Ashraf vd. (2021) çalışmalarında, CNN ve YOLOv5 algoritmaları ile güvenlik ve video izleme sistemlerinde silah tespiti uygulaması geliştirmişlerdir (Ashraf vd., 2021). Akbulut ve Halaf (2021), YOLO algoritmaları ve OpenCV kütüphanelerini kullanarak geliştirdikleri uygulama ile silah tespiti yapmışlardır (Akbulut & Khalaf, 2021). Duran-Vega vd. (2021) yaptıkları çalışmada YOLOv5 algoritmalarından faydalanarak, gerçek zamanlı tabanca tespitinde kullanılmak üzere, yarı tekrarlı sinir ağlarına dayanan bir zamansal YOLOv5 dedektörü geliştirmişlerdir (Duran-Vega vd., 2021). Lim vd. (2021) gerçek zamanlı görüntüler

üzerindeki farklı boyuttaki silahları algılayabilmek için M2Det ağı eğitmiş ve test verileri üzerinde 87.42% oranında başarı elde etmişlerdir (Lim vd., 2021). Raju vd. (2022), silah tespiti için silah, bıçak vb. tehlikeli nesnelerin bulunduğu herkese açık bir veri seti kullanarak, YOLO ve Faster R-CNN tabanlı bir model geliştirmişlerdir (Raju vd., 2022). Rehman vd. (2022) suçun erken tespiti için yaptıkları çalışmada YOLOv5 kullanarak kendi oluşturdukları üç farklı silah sınıfı (kısa silah, uzun sihah, bıçak) için tespit yapmışlardır (Rehman & Fahad, 2022). KU vd. (2023) güvenlik kamerası görüntülerini kullanarak silah, bıçak vb. nesneleri tespit edebilmek için 6000 adet etiketlenmiş görsel ile bir YOLOv8 modeli eğitmişlerdir ve yaklaşık 93%'lik bir başarı elde etmişlerdir (K U & M, 2023). Khalid vd. (2023) silah tespiti için herkese açık bir veri seti üzerinde YOLOv5 temelli bir model geliştirmişler ve %95,43'lük F1 skoru elde etmişlerdir (Khalid vd., 2023). Tram vd. (2023) güvenlik kameralarından elde ettikleri görüntüler üzerinde silah tespiti yapan bir erken uyarı sistemi için YOLO mimarisi kullanarak bir çalışma yapmışlardır (Tram vd., 2023). Yadav vd. (2024) ve DarkVision modellerini kullanarak, aydınlatmanın az olduğu ortamlarda silah tespit edebilmek için bir uygulama geliştirmişlerdir ve %95,50 değerinde ortalama hassasiyet skoru elde etmişlerdir (Yadav vd., 2024). Berardini vd. (2024) kapalı alanlarda güvenlik kameralarından elde ettikleri görüntüler üstünde kullanmak üzere, tabanca ve bıçak tespiti yapabilmek için iki aşamalı bir derin öğrenme yaklaşımı uygulamışlardır. İlk aşamada insan tespiti yapan bir model kullanılmış, bu modelin rehberliğinde ortamdaki silah ve bıçak türü nesnelerin tespiti yapılmıştır (Berardini vd., 2024).

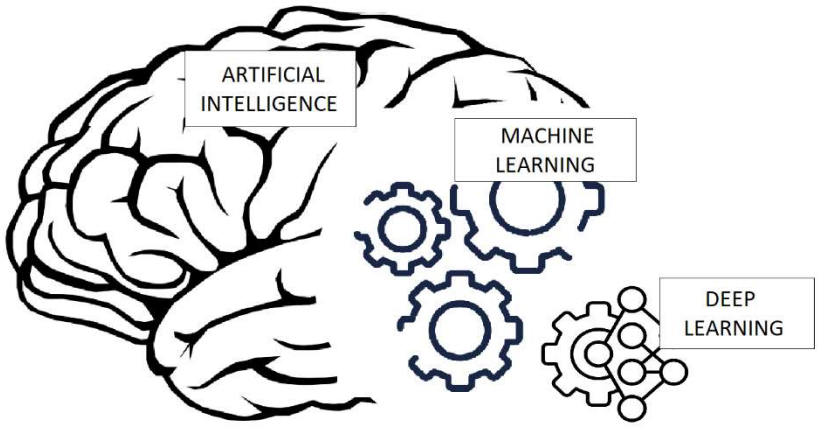
Sonuç olarak, CNN mimarileriyle tehlikeli nesne (tabanca, tüfek, bıçak vb) tespiti yapan çalışmalar incelendiğinde farklı senaryolar için uygulama ve araştırmaların yapıldığı görülmektedir. Yapılan çalışmalar ile birlikte CNN ağlarından, transfer öğreniminden, uzun kısa süreli bellek (LSTM) ağlarından ve diğer ileri teknolojilerden yararlanılarak, gelişmiş doğruluk, verimlilik ve gerçek zamanlı tespit yapabilen otomatik silah tespit sistemlerinin geliştirilmesinde önemli ilerlemeler kaydedilmiştir.

Bu çalışma kapsamında, tehlikeli nesne tespitinde literatürde başarısı kanıtlanmış YOLO modellerinden yararlanılacaktır. Literatürdeki çalışmalardan farklı olarak ilk kez

YOLOv9 modeli analizlere dahil edilecek ve yine literatürde ilk kez üç veri seti birleştirilerek önerilen modellerin performansları değerlendirilecektir.

2.2. Yapay Zekâ

Yapay zekâ genel olarak, bir bilgisayarın veya bilgisayar kontrollü robotun, genellikle akıllı varlıklarla ilişkili görevleri yerine getirme yeteneği olarak tanımlanabilir. Bu terim sıklıkla akıl yürütme, anlam keşfetme, genelleme veya geçmiş deneyimlerden öğrenme gibi insanlara özgü entelektüel süreçlerle donatılmış sistemler geliştirmek amacıyla kullanılmaktadır. Kavramsal olarak yapay zekâ ve alt disiplinlerinin temsili Şekil 2.1'de gösterilmiştir.



Şekil 2.1. Yapay zekâ disiplinleri hiyerarşisi

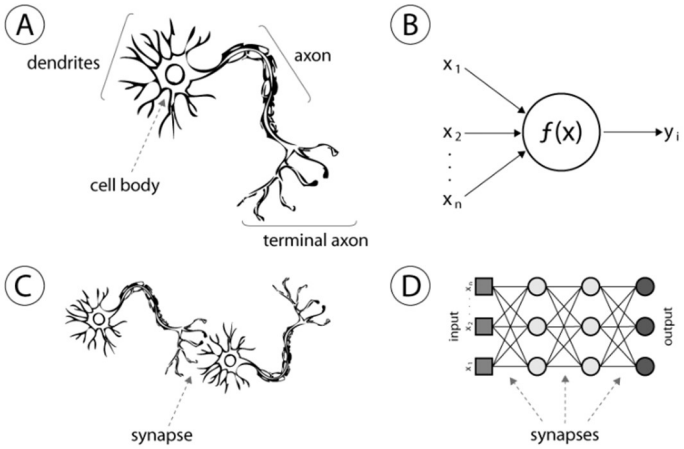
2.3. Makine Öğrenmesi

Makine öğrenmesi (ML-*Machine Learning*), bir bilgisayarın doğrudan, yönlendirme olmadan öğrenmesine yardımcı olmak için matematiksel modelleri kullanma işlemi olarak tanımlanır. Makine öğrenmesi, kavram olarak yapay zekanın (AI - *Artificial Intelligence*) bir alt kümesi olarak kabul edilmektedir. Makine öğrenmesi, veri kümelerindeki örüntüleri belirlemek için matematiksel algoritmaları kullanır. Tahmin

yapabilen bir veri modeli oluşturmak için de bu kalıplar kullanılır. İnsanların daha fazla bir konu üzerinde alıştırma yaptıkça gelişmesi gibi, veri ve deneyim arttıkça makine öğrenmesinin sonuçları da daha doğru hale gelir.

2.4. Yapay Sinir Ağları

Yapay sinir ağları insan beyninin en temel özelliği olan öğrenme fonksiyonunu gerçekleştiren bilgisayar sistemleridir. Şekil 2.2'de yapay sinir ağları ve biyolojik sinir ağlarının yapılarının temsili gösterimi yer almaktadır.



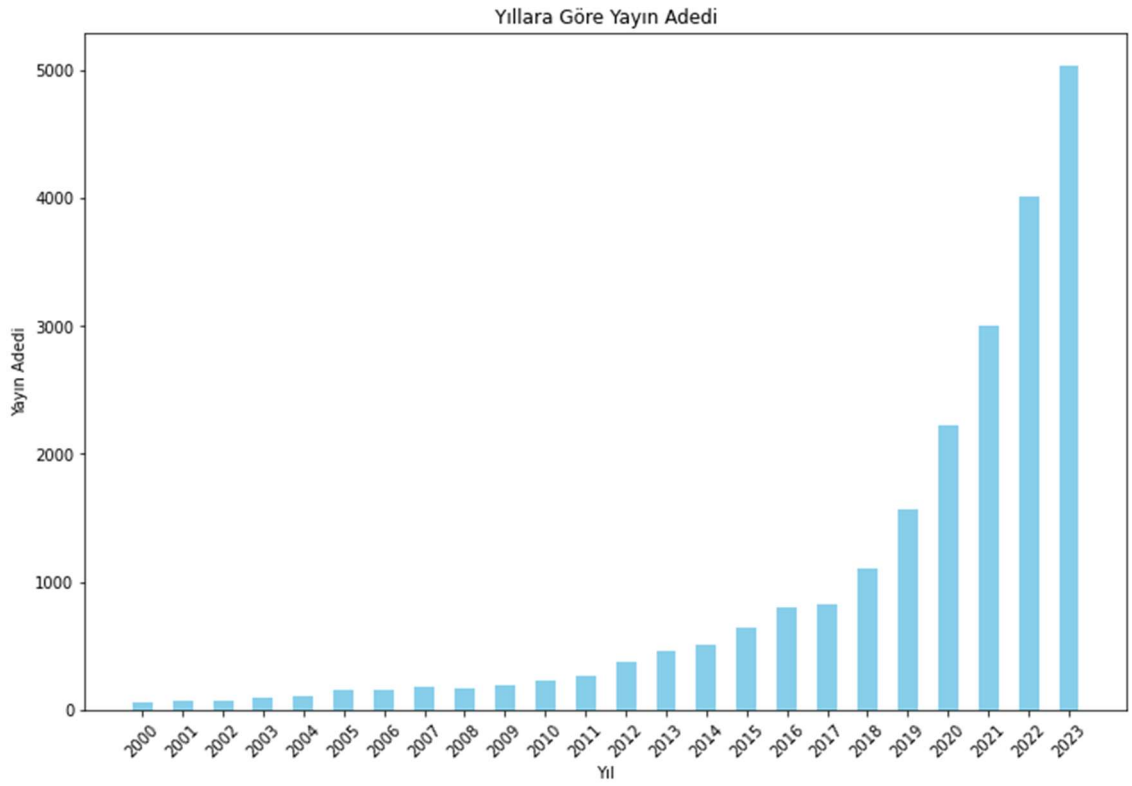
Şekil 2.2. Biyolojik ve yapay sinir ağı gösterimleri

2.5. Derin Öğrenme

Evrişimli Sinir Ağları (CNN-*Convolutional Neural Network*) derin öğrenme kavramını en iyi temsil eden modeldir (LeCun vd., 2015). Derin Sinir Ağları (DNNs-*Deep Neural Networks*) bilgisayarlı görü uygulamalarında yaygın olarak kullanılmaktadır. Bu teknolojilerin temel taşlarından biri olan Evrişimli Sinir Ağları (CNN), özellikle görüntü ve video işleme konularında devrim yaratmıştır. CNN'ler, görüntülerden otomatik olarak karmaşık özellikler çıkarabilme yeteneğine sahiptirler.

2.5.1. Nesne Tespiti

Nesne tespiti, dijital görüntülerdeki belirli bir sınıftaki (insan, hayvan, araba… vb.) görsel nesnelerin örneklerinin tespit edilmesiyle ilgilenen önemli bir bilgisayarlı görme görevidir (Zou vd., 2023). Nesne tespitinin amacı, verilen bir görüntüde nesnelerin nerede bulunduğunu belirlemek (nesne konumlandırma) ve her nesnenin hangi sınıfa ait olduğunu belirlemektir (nesne sınıflandırma). Bu nedenle, geleneksel nesne tespiti modellerinin iş akışı genellikle bilgilendirici bölge seçimi, özellik çıkarma ve sınıflandırma olmak üzere üçe ayrılmaktadır (Zhao vd., 2019). Şekil 2.3'de görüldüğü üzere nesne tanıma alanında yapılan çalışma adedi her geçen yıl artmaktadır.



Şekil 2.3. 2000-2023 yılları arası "Nesne Tespiti" alanında yapılan akademik çalışma artışı (Veriler Science Direct arama motoru içinde "Object Detection" anahtar sözcüğü kullanılarak elde edilmiştir.)

2.5.2. Evrişimli Sinir Ağlarının Temelleri (CNN)

Evrişimli Sinir Ağları, Yapay Sinir Ağları'nın bir alt kümesidir ve özellikle görsel verinin analizi için tasarlanmıştır. CNN'ler, veriden herhangi bir insan müdahalesi gerektirmeksizin özellikler çıkarma yetenekleri sayesinde, görsel tanıma görevlerinde yüksek başarı oranları elde etmektedirler. Evrişimli sinir ağları, birden fazla katmandan oluşur; her bir katman, girdi verisinden daha soyut öz ellikler çıkarmak için tasarlanmıştır.

Bir CNN, genellikle üç tür katmandan oluşur: evrişim katmanları, havuzlama (*pooling*) katmanları ve tam bağlantılı (*fully connected*) katmanlar.

Evrişim Katmanı: Evrişim katmanı, görüntü üzerinde bir filtre uygulayarak yerel özellikleri (kenarlar, doku vb.) tespit eder. Bu işlem, farklı konumlardaki özelliklerin tanınmasını sağlar.

Havuzlama Katmanı: Havuzlama katmanı, bir sonraki katmana aktarılacak veri miktarını azaltarak, hesaplama yükünü hafifletir ve ağın aşırı uyumunu (*overfitting*) önler. En yaygın kullanılan havuzlama yöntemi, maksimum havuzlamadır (*max pooling*), burada belirli bir pencere boyutu içindeki maksimum değer seçilir.

Tam Bağlantılı Katman: CNN'nin son katmanları genellikle tam bağlantılı katmanlardır ve öğrenilen özelliklerin sınıflandırma veya diğer görevler için kullanıldığı yerdir. Bu katman, ağın çıkarım yapmasını sağlayan son tahminleri üretir.

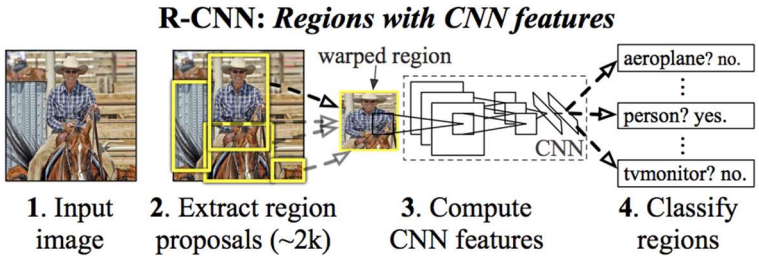
2.5.3. Bölge Tabanlı Evrişimli Sinir Ağı (R-CNN)

Bölge Tabanlı Evrişimli Sinir Ağı (R-CNN/*Region-based CNN*) görüntü tanıma ile nesne tespiti görevlerinde kullanılan bir derin öğrenme mimarisidir.

Çalışma Prensibi şu şekildedir:

1. **Bölge Önerisi Üretme**: Giriş görüntüsünden potansiyel nesne bölgeleri belirlenir. Bu işlem için seçici arama (Selective Search) gibi arama algoritmaları kullanılır.

2. **Özellik Çıkarma:** Potansiyel nesne bölgeleri üzerinde önceden eğitilmiş bir CNN modeli kullanılarak özellik haritaları elde edilir.
3. **Bölge Sınıflandırma ve Sınırlayıcı Kutu Regresyonu:** Özellik haritaları, nesne sınıflandırması ve sınırlayıcı kutu regresyonu yapabilmek için ayrı ayrı birer makine öğrenmesi modeline (genellikle bir destek vektör makinesi SVM) verilir. Sınıflandırma modeli, bölgenin hangi sınıfa ait olduğunu tahmin ederken, regresyon modeli, nesnenin sınırlayıcı kutusunu (*bounding box*) daha doğru bir şekilde hizalamak için kullanılır.
4. **Sonuçların Birleştirilmesi:** Son adımda, önerilen bölge adaylarından gelen sınıflandırma sonuçları birleştirilir ve çakışan bölge önerileri birleştirilerek tek bir nesne tahmini oluşturulur. Aynı nesneye ait olan ve birbirine yakın bölge önerileri birleştirilir.



Şekil 2.4. R-CNN Mimarisi Gösterimi (*What Is R-CNN?*, 2023)

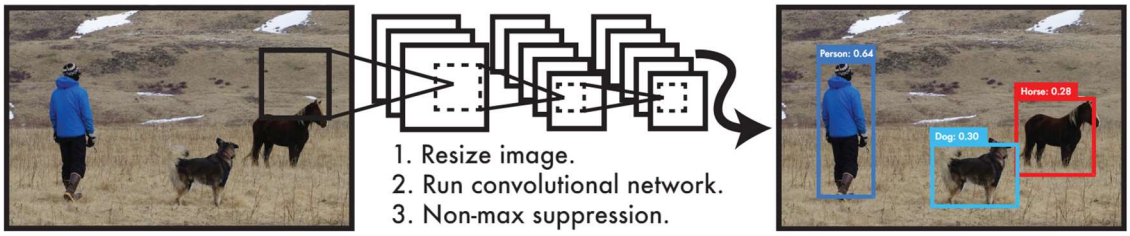
R-CNN'in en büyük dezavantajı, ilgili öznitelikleri çıkarma görevinde, önerilen bölgelerin miktarının çok olması sebebiyle yavaş olmasıdır. Tüm bu bölgeler, binlerce ileri beslemeli hesaplama ve büyük bir bilgi işlem yükü gerektiren CNN ağından geçirilir ve bu da onu gerçek zamanlı nesne algılama uygulamaları için olanaksız hale getirir (Jiao vd., 2019).

2.6. YOLO Modelleri

İnsan, bir görüntüye baktığında hangi nesnenin nerede olduğunu ve görsel üzerinde nasıl bir etkileşim gösterdiğini anlık olarak algılar. İnsan algılarının hızlı ve doğru

olması sebebiyle, araç kullanmak gibi karmaşık bir görevi daha az bir zihinsel aktiviteyle gerçekleştirebilir (Redmon vd., 2016). Nesne tespitinde kullanılan hızlı ve doğru algoritmalar, bilgisayarların üst düzey sensörlere ihtiyaç duymadan görsel verilerle bir otomobili otonom olarak kontrol etmesine olanak sağlar.

Nesne tespitinde kullanılan R-CNN gibi güncel algoritmaların bir kısmı, tahminlerinde kayan pencereler ve bölge önerilerinden yararlanmaktadır. YOLO ise farklı olarak eğitim ve test süresi boyunca görüntünün tamamını görmektedir. Bu durum, ilgili sınıfların yanı sıra nesnelerin görünümü hakkındaki bağlamsal bilgileri de sağlamaktadır. Şekil 2.5'de YOLO algoritmasının çalışma prensibi gösterilmiştir.



Şekil 2.5. YOLO Algoritması Tespit Sistemi(Redmon vd., 2016)

Bir YOLO (*You Only Look Once*) mimarisinde temel olarak şu katmanlar bulunmaktadır: Giriş katmanı, konvolüsyon katmanları, tam bağlantılı katmanlar ve çıkış katmanı. Model, giriş katmanında bir görüntü alır ve bu görüntüyü konvolüsyon katmanlar aracılığıyla işler. Konvolüsyon katmanları, ilgili görüntüdeki özellikleri çıkarır ve bu özellikleri daha yüksek seviyede temsil eder. Tam bağlantılı katmanlar, bu özellikleri daha ileri düzeyde analiz eder ve son olarak çıkış katmanında nesne tespiti ve sınıflandırma sonuçlarını üretirler. YOLO mimarisinin genel çalışma mantığı, görüntüyü ızgaralara (*grid*) bölerek her ızgara hücresinde nesne olup olmadığını ve varsa nesnenin sınıfını ve konumunu tahmin etmek şeklindedir. Bu tahminler sınırlayıcı kutular (*bounding boxes*) kullanılarak yapılır. Sınırlayıcı kutular, nesnelerin konumlarını belirten dikdörtgen alanlardır. YOLO modeli, her grid hücresinde belirli sayıda bağlantı (*anchor*) kutusu kullanarak nesnelerin olası konumlarını ve boyutlarını tahmin etmektedir. Bu bağlantı kutuları, çeşitli boyut ve şekillerde olabilmekte ve modelin daha geniş bir nesne kümesini doğru bir şekilde tespit etmesine yardımcı

olmaktadırlar. Temel gerçek kutusu (*ground truth box*), modelin eğitimi sırasında kullanılan ve nesnelerin gerçek konumlarını ve boyutlarını belirten kutulardır. Modelin performansı, tahmin edilen sınırlayıcı kutuların temel gerçek kutularla ne kadar örtüştüğüne göre değerlendirilir. Bu örtüşme, IoU (*Intersection over Union*) yani "Birleşim üzerinden kesişim" metriği kullanılarak hesaplanır. IoU, tahmin edilen kutu ile temel gerçek kutu arasındaki kesişim alanının birleşim alanına oranıdır ve modelin doğruluğunu değerlendirmede önemli bir başarı metriğidir. YOLO modellerinin zaman içinde farklı versiyonları geliştirilmiştir ve sürekli yeni modeller geliştirilmeye devam edilmektedir. İlk versiyon olan YOLO, ardından gelen YOLOv2 ve YOLOv3 ile daha da geliştirilmiş, performans ve doğruluk açısından iyileştirilmiştir. Sonrasında YOLOv4 ve YOLOv5 gibi versiyonlar, daha yüksek hız ve doğruluk sunarken, modelin karmaşıklığını da artırmıştır. Günümüzde ise YOLOv6, YOLOv7, YOLOv9 gibi daha yeni ve gelişmiş versiyonlar bulunmaktadır. Her yeni versiyon, önceki versiyonların üzerine eklemeler yaparak ve çeşitli optimizasyonlar sağlayarak nesne tespitinde daha yüksek başarı oranları ve daha hızlı işlem süreleri sunmaktadır. Bu çalışmada YOLOv5 ve sonrası açıklanmıştır.

YOLOv5: YOLOv5, YOLOv4 sürümünden yaklaşık bir ay sonra yayınlanmış bir ardıl sürümdür. Mimari yapı olarak YOLOv4'e benzemektedir. YOLOv4'ten farklı olarak bağlantı kutusu (*anchor box*) entegrasyonudur.

YOLO NAS: YOLO NAS (*You Only Look Once Neural Architecture Search*), nesne tespit modelleri için özelleştirilmiş, derin öğrenme modelleri oluşturmak için kullanılan bir yöntemdir. Bu yöntemin amacı, evrişimli sinir ağlarının (CNN) mimari yapılarını otomatik olarak araştırmak ve optimize etmektir. YOLO NAS, önceki YOLO modellerinin mimari kısıtlarına takılmamak ve daha iyi performans elde etmek için kullanılmaktadır. YOLO NAS, belirli bir nesne tespit görevine en uygun mimari yapıyı bulmak ve modelin performansını arttırmak için özelleştirilmiş bir CNN mimarisi oluşturmak için genetik algoritma ve diğer optimizasyon tekniklerini kullanmaktadır.

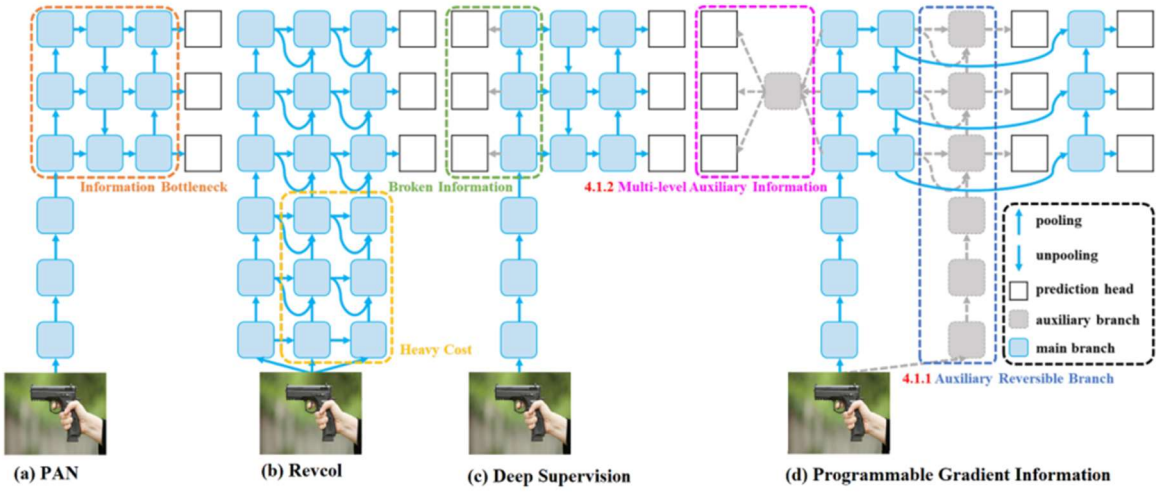
YOLOv8: YOLOv8, YOLOv5 ile benzer bir mimari yapı kullanmaktadır. YOLOv5'ten farklı olarak YOLOv8 için, sabitleyici Olmayan Algılama (*Anchor-Free Detection*) geliştirmesi yapılmıştır.

YOLOv9: Derin öğrenme yöntemlerinin mevcut uygulamaları, modelin tahmin sonuçlarının gerçek değerlere yaklaşmasını sağlayacak hedef fonksiyonların tasarımına ve yeterli bilginin elde edilmesini kolaylaştıracak uygun mimarilerin geliştirilmesine büyük önem vermektedir. Ancak, bu yöntemler genellikle girdi verisinin katmanlar arası geçişi sırasında yaşanan önemli bir sorunu - büyük miktarda bilgi kaybı - göz ardı eder. Bu çalışma, derin ağlar aracılığıyla veri iletiminde meydana gelen bilgi kaybının, bilgi darboğazı ve tersinir fonksiyonlar gibi kritik konularını ele almakta ve bu zorlukların üstesinden gelmek için 'programlanabilir gradyan bilgisi' (PGI) kavramını ve yeni bir hafif ağ mimarisi olan 'Genelleştirilmiş Verimli Katman Birleştirme Ağı' (GELAN) tasarımını önermektedir. Önerilen çözümler, MS COCO veri seti temelli nesne algılama görevinde test edilmiş ve GELAN'ın (*Generalized Efficient Layer Aggregation Network*), geleneksel evrişim operatörlerini kullanarak derinlik tabanlı evrişimlere dayalı yöntemlere kıyasla daha iyi parametre kullanımı sağladığı kanıtlanmıştır (Wang vd., 2024). Şekil 2.6'da herkesin kullanımına açık 80 farklı sınıf, 200.000 adedi etiketli, 330.000 adet görselden oluşan COCO veri setinde gerçekleştirilen YOLOv9 modelleri performans karşılaştırması yapılmıştır. Şekilde birinci sütun model adını, ikinci sütun resimlerin boyutlarını, 3.sütun 0.50 (50%) ila 0.95 (95%) arasındaki IoU eşiklerinde hesaplanan ortalama AP değerini, 4. sütun mAP 50, IoU eşiği 0.50 (50%) olduğunda hesaplanan AP değerlerinin ortalamasını, 5. sütun params(M) ilgili satırdaki modelin kaç milyon parametre olduğunu, 6. Sütun FLOP(B) milyar cinsinden ilgili modelin kayan nokta sayısını (karmaşıklığını) temsil etmektedir. Bu çalışmada ise kullanıma açık son sürüm olan YOLOv9m modeli kullanılmıştır.

Model	boyut (piksel)	mAPval 50-95	mAPval 50	params (M)	FLOP'lar (B)
YOLOv9t	640	38.3	53.1	2.0	7.7
YOLOv9s	640	46.8	63.4	7.2	26.7
YOLOv9m	640	51.4	68.1	20.1	76.8
YOLOv9c	640	53.0	70.2	25.5	102.8
YOLOv9e	640	55.6	72.8	58.1	192.5

Şekil 2.6. YOLOv9 Coco veri seti üzerinde elde edilen metrikler

YOLOv9 ağı mimarisi Şekil 2.7'de gösterilmiştir. Birinci katman (a) PAN (*Path Aggregation Network*) katmanı, bilgi kaybını minimize ederek farklı seviyelerdeki özellikleri birleştirmektedir. (b) RevCol (*Reversible Columns*) katmanı, bilgi aktarımındaki kayıpları ve kırılmaları iyileştirmektedir. (c) Derin Gözetim (*Deep Supervision*) katmanı derin katmanların daha iyi öğrenmesini sağlamaktadır. (d) Programlanabilir Gradyan Bilgisi (*Programmable Gradient Information*) (PGI) katmanı modelin geri yayılım sırasında daha iyi öğrenmesini sağlamaktadır.

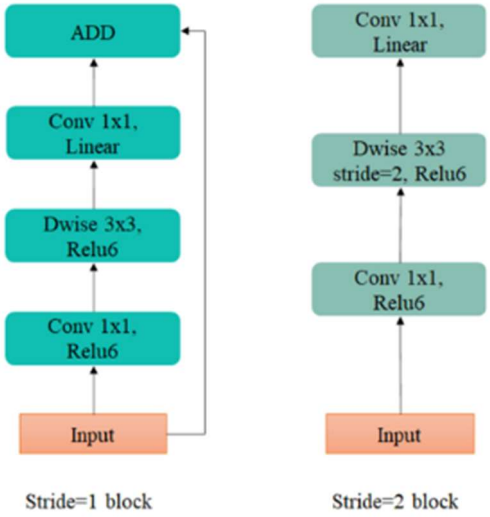


Şekil 2.7. YOLOv9 Ağı Gösterimi

2.7. MEDIAPIPE Nesne Tespit Modelleri

MediaPipe, ses ve görüntü gibi duyusal veriler üzerinde çıkarım yapmak için Google tarafından geliştirilmiş açık kaynak bir çerçevedir (*Framework*) (Lugaresi vd., 2019). MediaPipe ile makine öğrenmesi tabanlı üretken yapay zekâ modelleri kullanılarak görüntü üretmek mümkündür. Genel olarak Mediapipe kullanılarak yapılabilecek görevler; nesne tanıma, sınıflandırma, görüntü bölütleme, etkileşimli bölütleme, mimik tanıma, poz kestirimi şeklindedir. Mediapipe ile nesne tespit görevleri için MobileNet-V2 ve MobileNet-MultiHW-AVG modelleri eğitim için kullanılabilmektedir.

MobileNet-V2: MobileNet-V2, gömülü sistemler ve mobil cihazlarda kullanım için oluşturulmuş hafif bir derin öğrenme modelidir. MobileNetV2, MobileNetV1'in üzerine geliştirilmiş bir modeldir. Kullanım alanları genel olarak, görüntü sınıflandırma, nesne tespiti ve yüz tanıma gibi görsel algılama görevleridir. Mobil cihazların donanım ve işlem gücü kısıtları sebebiyle hafif modeller önem kazanmaktadır. MobileNetV2 mimarisi şekil 2.8'de gösterilmiştir.

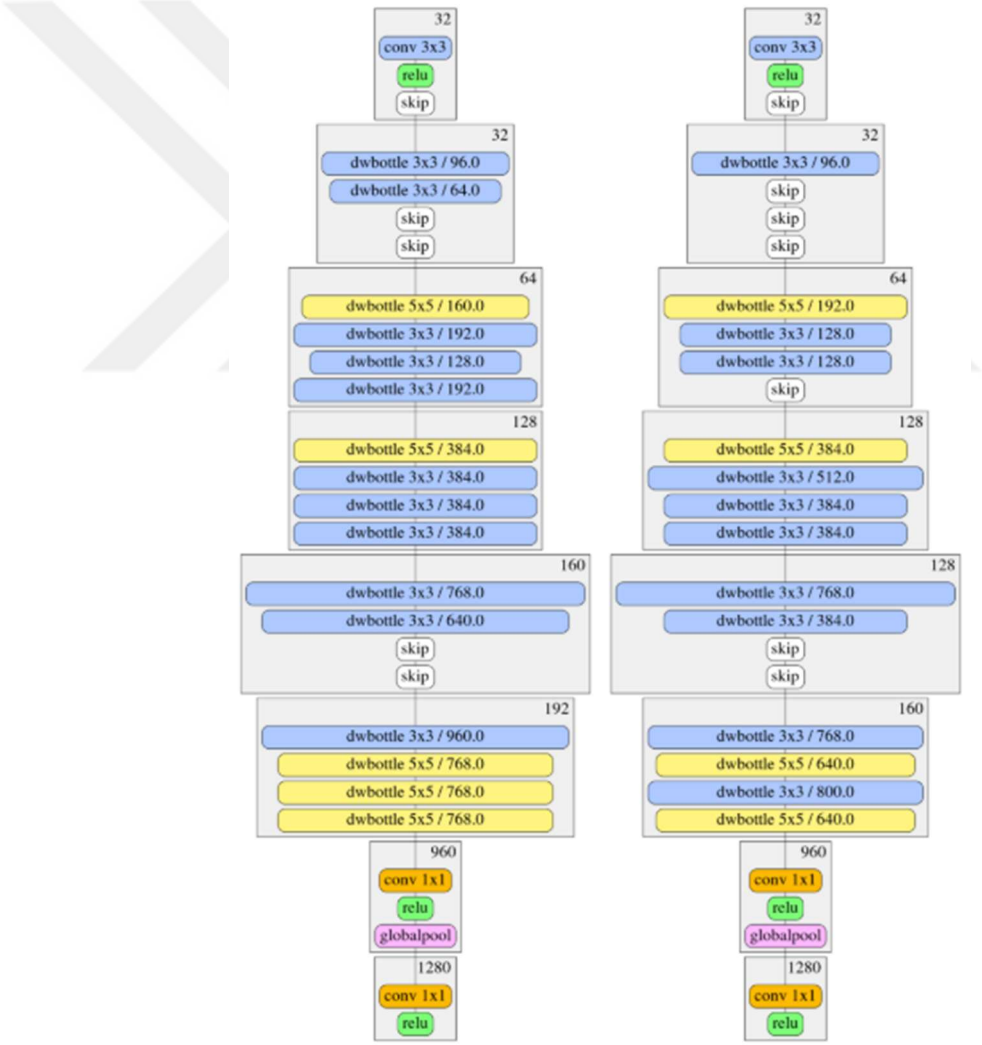


Şekil 2.8. MobileNetV2 konvolüsyon blokları gösterimi (Dong vd., 2020)

MobileNet-MultiHW-AVG: MobileNet-MultiHW-AVG, MediaPipe için özel olarak tasarlanmış bir görüntü işleme modelidir. Bu model, farklı donanımsal ortamlarda

(MultiHW) etkili bir şekilde çalışacak şekilde optimize edilmiştir. "AVG" ise "average" kavramının kısaltması olup, modelin çıkarımlarını hesaplarken ortalama alma işlemini vurgulamaktadır.

MobileNet-MultiHW-AVG modeli, görüntü işleme görevlerinde kullanılmak üzere oluşturulmuş hafif bir modeldir. Mobil aygıtlar, gömülü sistemler vb. donanım platformları için düşük hesaplama gücü ve bellek kullanımı sağlamaktadır. MobileNet-MultiHW-AVG modeli, donanımsal kısıtların bulunduğu bu tür uygulamalarda yüksek performans ve düşük kaynak tüketimi sağlamak amacıyla tasarlanmıştır. MobileNet-MultiHW-AVG mimarisi Şekil 2.9'da gösterilmiştir.



Şekil 2.9. Multi-AVG(Solda) ve Multi-MAX(Sağda) model gösterimleri(Chu vd., 2021)

2.8. Performans Metrikleri

Nesne tespit yöntemlerinin performanslarını objektif bir şekilde değerlendirmek için karmaşıklık matrisi, ortalama kesinlik (*average precision-AP*), AP değerlerinin ortalaması (*mean average precision-*mAP), kesinlik (*Precision-*Pr), duyarlılık (*Recall-*Re) ve $F1$ skoru gibi metrikler kullanılır. Bu şekilde, kullanılan yöntemin performansı değerlendirilebilir. Bu metrikler kullanılarak nesne tespit yöntemleri arasında karşılaştırmalar yapılır. Bu metrikleri hesaplayabilmek için bazı değerlerin kullanılması gerekir. Doğruluk, kesinlik ve duyarlılık değerleri hesaplanırken Confusion Matrix yararlanılır. Şekil 2.10 üzerinde karmaşıklık matrisi yapısı gösterilmiştir.

		Predicted Class	
		P	N
Actual Class	P	TP	FN
	N	FP	TN

Şekil 2.10. Karmaşıklık Matrisi(Confusion Matris)

True Positive (TP), doğru sınıflandırılmış pozitif örnekleri temsil eder. True Negative (TN), doğru sınıflandırılmış negatif örnekleri temsil eder. False Positive (FP), yanlış sınıflandırılmış pozitif örnekleri temsil eder. False Negative (FN), yanlış sınıflandırılmış negatif örnekleri temsil eder.

Nesne tespiti sonucunda elde edilen sınırlayıcı kutuların başarılı birer tahmin olup olmadığının belirlenebilmesi için tahmin edilen sınırlayıcı kutu değeri ile temel-gerçek kutusu değeri arasındaki birleşim üzerinden kesişim değerini gösteren $IoUp$ değeri ile eşik değeri olarak kullanılan $IoUt$ değerinin birbiriyle karşılaştırılması gerekmektedir. $IoUt$, birleşim üzerinden kesişim hesaplandıktan sonra eşik değeri olarak kullanılır. Hesaplanan bu değer, tahmin edilen sınırlayıcı kutunun, doğru bir tahmin olarak değerlendirilebilmesi için gereken en alt limittir. $IoUp \geq IoUt$ ise, nesnenin sınırlayıcı

kutunun içinde olduğuna karar verilir. IoU hesabı denklem (2.1)'de verilmiştir (SK: Sınırlayıcı Kutu, TG: Temel Gerçek Kutusu).

$$IoU_p = \frac{alan(SK \cap TG)}{alan(SK \cup TG)} \tag{2.1}$$

Kesinlik (*Precision*): Doğru olarak tespit edilen nesnelerin toplam tespit edilen nesneler içindeki oranını temsil eder. Formül, denklem (2.2)'de verilmiştir.

$$Pr=\frac{TP}{TP+FP} \tag{2.2}$$

Duyarlılık (*Recall*): Gerçekten var olan tüm nesnelerin doğru şekilde tespit edilme oranını ifade eder. Formül, denklem (2.3)'de verilmiştir.

$$Re=\frac{TP}{TP+FN} \tag{2.3}$$

F1 Skoru: Hassasiyet ve duyarlılığın harmonik ortalamasından elde edilir. Formül, denklem (2.4)'de verilmiştir.

$$F_1=2\ x\ \frac{Pr\ x\ Re}{Pr+Re} \tag{2.4}$$

AP metriği, model eğitiminde kullanılan her bir sınıf için ayrı ayrı hesaplanır. AP metriği Pr-Re eğrisi altında kalan alana eşittir. Formül, denklem (2.5)'de verilmiştir.

$$AP=\int_0^1 \mathrm{Pr(Re)}\,.\,\mathrm{dPr} \tag{2.5}$$

Ortalama Hassasiyet (*mAP - Mean Average Precision*): Her bir sınıf için hesaplanan hassasiyetlerin ortalamasıdır ve nesne tespit yöntemlerinde genel başarı değerlendirmesi yapılırken kullanılır. Formül, denklem 2.6'de verilmiştir.

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{2.6}$$

3. MATERYAL ve YÖNTEM

3.1. Veri Seti

Veri seti; tabanca, tüfek ve bıçak olmak üzere 3 nesneyi tespit etmek üzere oluşturulmuştur. Bunun için herkese açık 3 ayrı veri seti kullanılmıştır. Tabanca tespiti için kullanılan veri seti toplam 2986 adet, bıçak tespiti için kullanılan veri seti toplam 2078 ve tüfek tespiti için kullanılan veri seti toplam 719 adet RGB formatında 640x640'lık görselden oluşmaktadır.

Tabanca görselleri internet görselleri ve güvenlik kameralarından alınmıştır. Bu görüntüler daha önce Olmos vd. (2018) tarafından çalışmalarında kaynak olarak kullanılmıştır (Olmos vd., 2018).



Şekil 3.1. 'Tabanca' sınıf etiketine ait örnek görseller

Bıçak görselleri en az bir adet bıçak görüntüsü içermektedir. Bu görseller internet ortamı, Youtube videoları ve güvenlik kameraları görüntülerinden derlenmiştir. Bıçak veri seti daha önceleri Castillo vd. (2019) tarafından çalışmalarında kaynak olarak kullanılmıştır (Castillo vd., 2019).



Şekil 3.2. ‘Bıçak’ sınıf etiketine ait örnek görseller

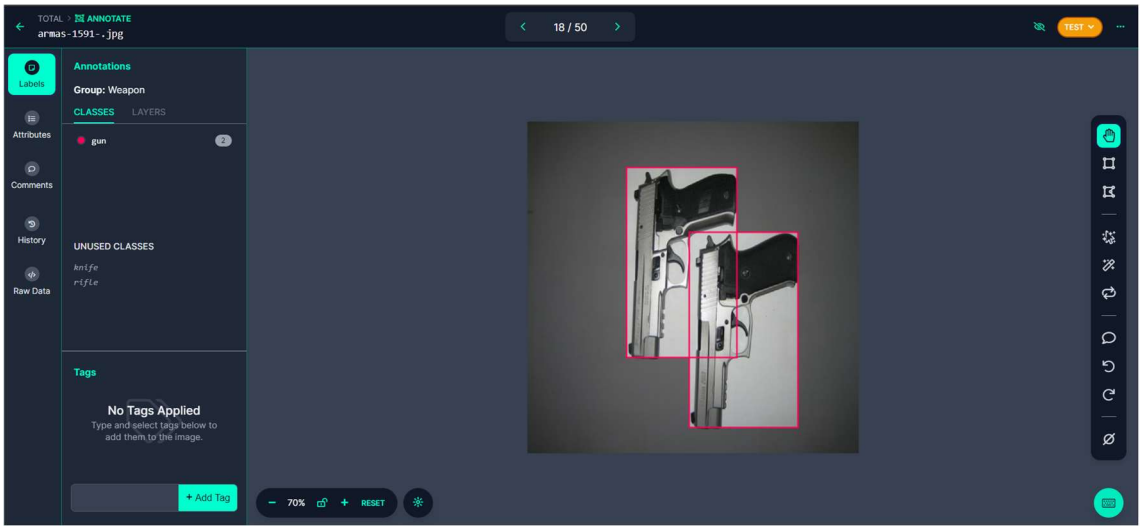
Tüfek görselleri internette herkese açık rastgele bir veri setinden alınmıştır.



Şekil 3.3. ‘Tüfek’ sınıf etiketine ait örnek görseller

3.2. Veri Setinin Etiketlenmesi

Veri seti etiketleme yeniden boyutlandırma işlemi yapılırken web tabanlı Roboflow uygulaması (https://roboflow.com) kullanılmıştır. Bu uygulama üzerinde veri seti etiketleme, veri arttırımı (*Data Augmentation*), yeniden boyutlandırma ve farklı etiket formatlarına dönüştürme işlemleri yapılabilmektedir. Şekil 3.4 üzerinde örnek bir görsel etiketleme işlemine yer verilmiştir.



Şekil 3.4. Roboflow uygulaması üzerinde "Tabanca" sınıfına ait örnek bir etiketleme işlemi

Pascal VOC formatı: Mediapipe model eğitimi için Pascal VOC tipi görüntü etiket yapısı kullanılmıştır. PASCAL VOC XML yapısı ile, veri setindeki her bir görsele ait etiket bilgileri bir XML dosyasında saklanır. Bu XML dosyası, kök düzeyde sınıf etiketi içerir ve bu etiket, görüntü ve nesne bilgilerini tanımlayan çeşitli alt etiketlerden oluşur. Örnek bir PASCAL VOC XML görseli Şekil 3.5. üzerinde gösterilmektedir

```
<annotation>
  <folder>VOC2007</folder>
  <filename>sharpenKnife303_jpg.rf.d5a0803472b25d87b8e9af919b6a8f1a.jpg</filename>
  <source>
    <database>Unknown</database>
  </source>
  <size>
    <width>640</width>
    <height>640</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  <object>
    <name>0</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>199</xmin>
      <ymin>117</ymin>
      <xmax>590</xmax>
      <ymax>335</ymax>
    </bndbox>
  </object>
</annotation>
```

Şekil 3.5. Örnek Pascal VOC yapısı

Etiket dosyası içerisinde klasör adı, dosya adı, görselin boyutları tespit edilecek nesnenin sınıf bilgisi, sınırlayıcı kutunun koordinatları gibi bilgiler yer almaktadır.

PASCAL VOC etiket formatı, veri setindeki her bir görüntü için XML dosyaları kullanarak sınırlayıcı kutuları ve diğer ilgili bilgileri ayrıntılı bir şekilde tanımlar. Bu format, nesne tespiti ve sınıflandırma görevlerinde yaygın olarak kullanılmaktadır ve birçok derin öğrenme uygulaması tarafından desteklenmektedir.

YOLO Formatı: YOLO (You Only Look Once) etiket formatı, nesne tespiti görevlerinde yaygın olarak kullanılan formatlardan biridir. .txt uzantılı bu format, veri setindeki her bir görüntü için bir metin dosyası kullanarak sınırlayıcı kutuları ve sınıf etiketlerini depolamaktadır. YOLO formatı, basit bir yapı sunar ve bu nedenle büyük ölçekli veri kümelerinde hızlı ve etkili bir şekilde kullanılabilir. YOLO Etiket Formatı, her bir görüntü için bir metin dosyasından oluşmaktadır ve bu dosya, her nesne için ayrı bir satır içermektedir. Her satır, aşağıdaki bilgileri içerir:

- Sınıf etiketi,
- Sınırlayıcı kutu merkezi x koordinatı (0 ve 1 arasında normalize edilmiş halde),
- Sınırlayıcı kutu merkezi y koordinatı (0 ve 1 arasında normalize edilmiş halde)
- Sınırlayıcı kutunun genişliği (0 ve 1 arasında normalize edilmiş halde)
- Sınırlayıcı kutunun yüksekliği (0 ve 1 arasında normalize edilmiş halde)

Üzerinde iki adet nesne bulunan bir görsele ait örnek etiket dosyası Şekil 3.6'da görülmektedir.

```
1 0.253125 0.52578125 0.50625 0.6375
1 0.775 0.5890625 0.4375 0.81796875
```

Şekil 3.6. Örnek Yolo etiket yapısı

3.3. Kullanılan Yöntemler

Tehlikeli nesne tespitinde derin öğrenme yöntemlerinden güncel YOLO versiyonları ve Mediapipe Framework kullanılmıştır. YOLO özelinde YOLO NAS, YOLOv5, YOLOv8 ve YOLOv9 mimarilerinin m (medium) ölçekli modelleri kullanılmıştır. YOLOv9 için m(medium) ölçekli modeli henüz kullanıma açılmadığından, benzer ölçekli GELAN-C modeli kullanılmıştır. Diğer YOLO versiyonları için *m* (*medium*)

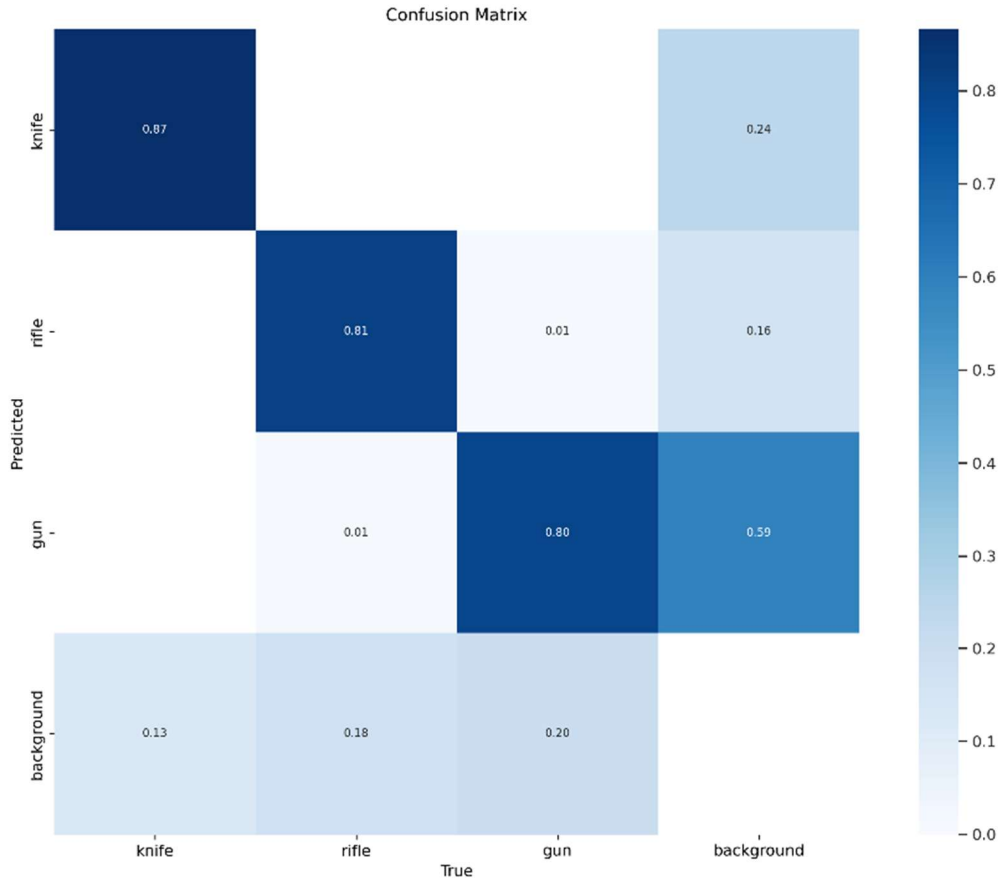
ölçekli model kullanılmıştır. Modeller 100 epoch ve 16 yığın boyutu (*batch size*) ile eğitilmiştir ve eğitilen model kullanılarak modelin daha önce bilmediği veri örneklerinden oluşan test verileri ile model performansları değerlendirilmiştir. Bunun dışında en iyi sonucu veren model kullanılarak gerçek zamanlı görüntüler üzerinde eğitilen modeli kullanarak tespit yapan bir uygulama geliştirilmiş ve model eğitim ve değerlendirilme aşamaları dışında anlık verilerle başarılı bir şekilde nesne tespiti yapılmıştır.

3.4. Çalışma Ortamı

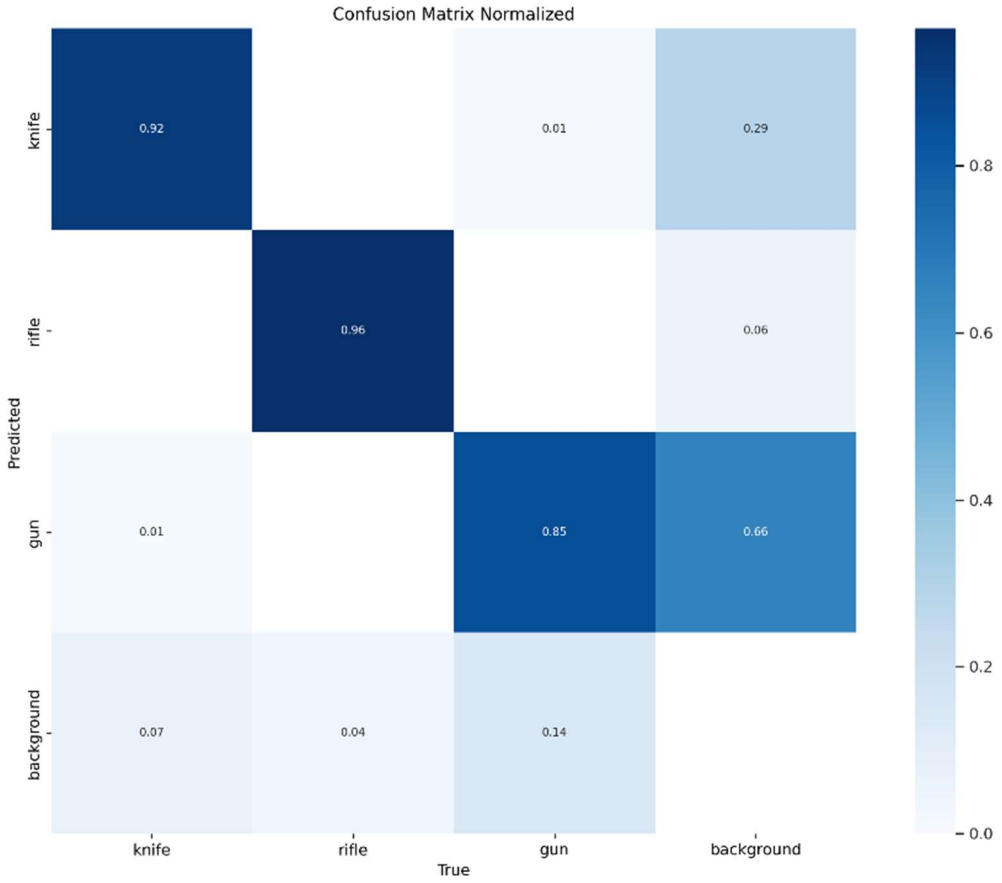
Çalışma boyunca modeller eğitilirken, Google Colab (https://colab.research.google.com) ortamı kullanılmıştır. Servis üzerinde NVIDIA V100 ekran kartı (GPU), 12.7 GB sistem belleği (System RAM) ve 15 GB GPU belleği (GPU RAM) ile kullanılmıştır.

4. BULGULAR ve TARTIŞMA

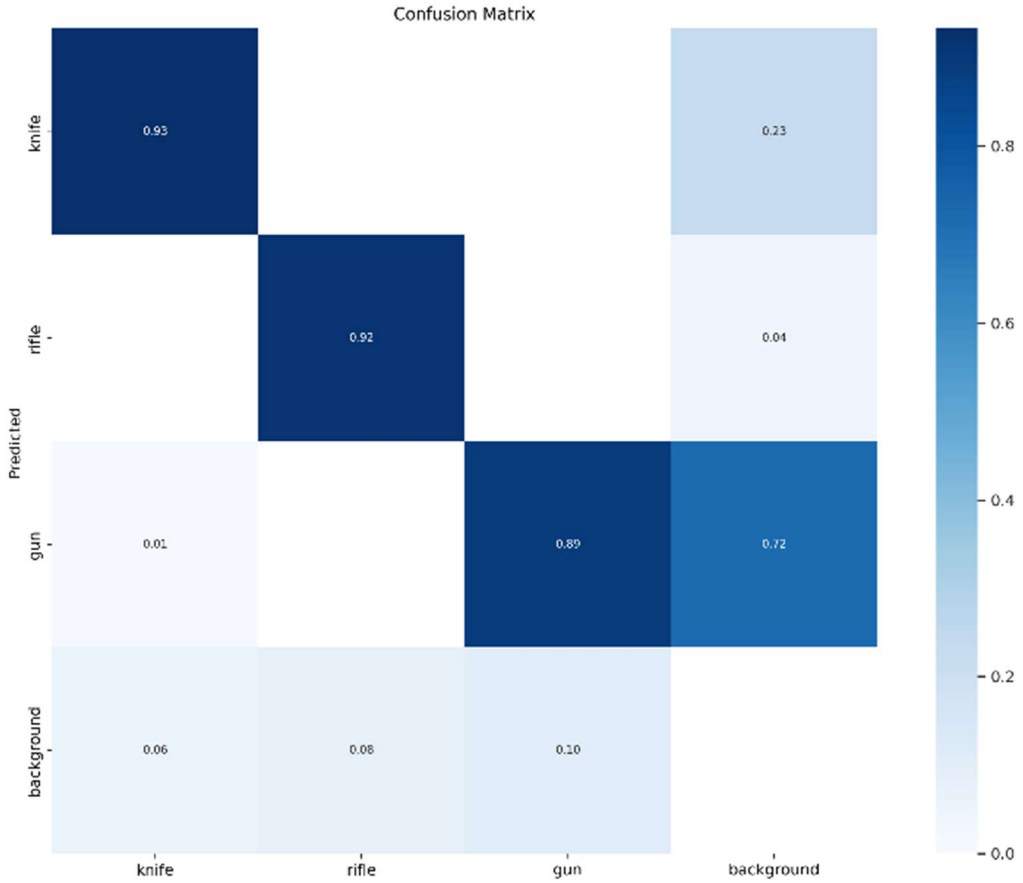
Seçilen modeller etiketlenmiş veri setleri ile eğitilmiştir. Çalışmada YOLOv9 GELAN-C modeli ile doğruluk ve güvenilirlik değeri en yüksek sonuç elde edilmiştir. Eğitilen modeller arasında en iyi sonucu veren 3 model için karmaşıklık matrisleri Şekil 4.1 – Şekil 4.4'te verilmiştir. YOLOv8 ve YOLOv9'un diğer modellerden açık ara daha başarılı sonuç verdiği görülmektedir. YOLOv9 ise YOLOv8 ile karşılaştırıldığında birbirine yakın değerler elde edilmiştir ancak karşılaştırılan sınıf örneklerinde 'Tabanca' ve 'Bıçak' sınıfları için YOLOv9 daha başarılı iken, 'Tüfek' sınıfı için YOLOv8 daha yüksek doğruluk oranı ile tahmin yapmıştır.



Şekil 4.1. YOLOv5 modeli validasyon veri seti kullanılarak elde edilmiş karışıklık matrisi



Şekil 4.2. YOLOv8 modeli validasyon veri seti kullanılarak elde edilmiş karışıklık matrisi



Şekil 4.3. YOLOv9 modeli validasyon veri seti kullanılarak elde edilmiş karışıklık matrisi

Farklı modeller kullanılarak elde edilen sonuçlar Çizelge 4.1'deki gibidir.

Çizelge 4.1. Farklı Dedektörlerin eğitim veri seti üzerindeki eğitim performansları

Dedektör	Epoch	Yığın Boyutu	mAP50	Pr	Re	F1	Eğitim Süresi (dk)
Mediapipe MobileNet-MultiHW AVG	100	16	0.7240	0.7240	0.6080	0.6603	37
YOLO Nas m	100	16	83.40	90.5	67.5	0.7739	142
YOLOv5 m	100	16	0.847	0.85	0.77	0.8083	292
YOLOv8 m	100	16	0.94	0.938	0.876	0.9045	303
YOLOv9 GELAN-C	**100**	**16**	**0.944**	**0.924**	**0.906**	**0.9152**	**328**

Çizelge 4.1'de en başarılı iki modelin YOLOv8 ve YOLOv9 olduğu görülmektedir. YOLOv8 ve YOLOv9 arasında büyük bir fark olmamasına rağmen YOLOv9 daha güvenilir bir sonuç vermiştir.

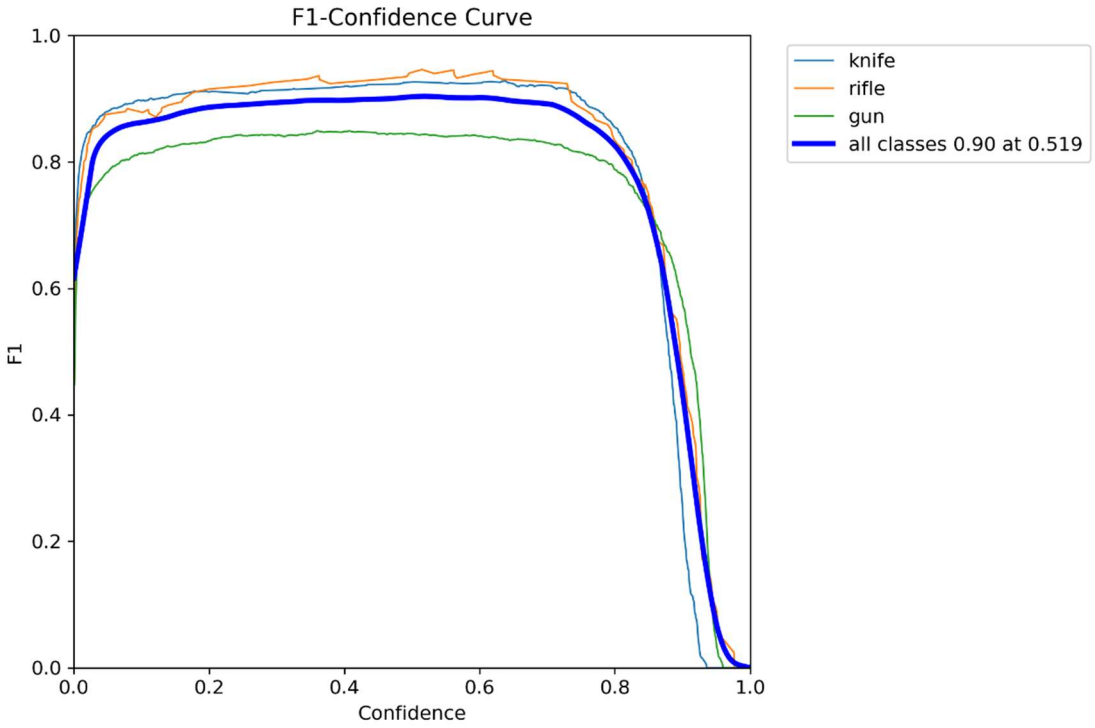
Giriş bölümünde yapılan araştırmalara ve Şekil 1.1'e göre silahlı suçlarda açık ara en çok kullanılan saldırı aracı 'Tabanca' olarak tespit edilmiştir. Bu duruma istinaden sonuçların daha objektif bir şekilde karşılaştırılıp çalışma amaçları doğrultusunda en iyi sonucu veren modelin tespiti için, validasyon veri seti üzerinde, sınıf etiketi düzeyinde YOLOv8 ve YOLOv9 modelleri karşılaştırılmıştır.

Çizelge 4.2. Farklı dedektörlerin validasyon veri seti üzerindeki tahmin performansları

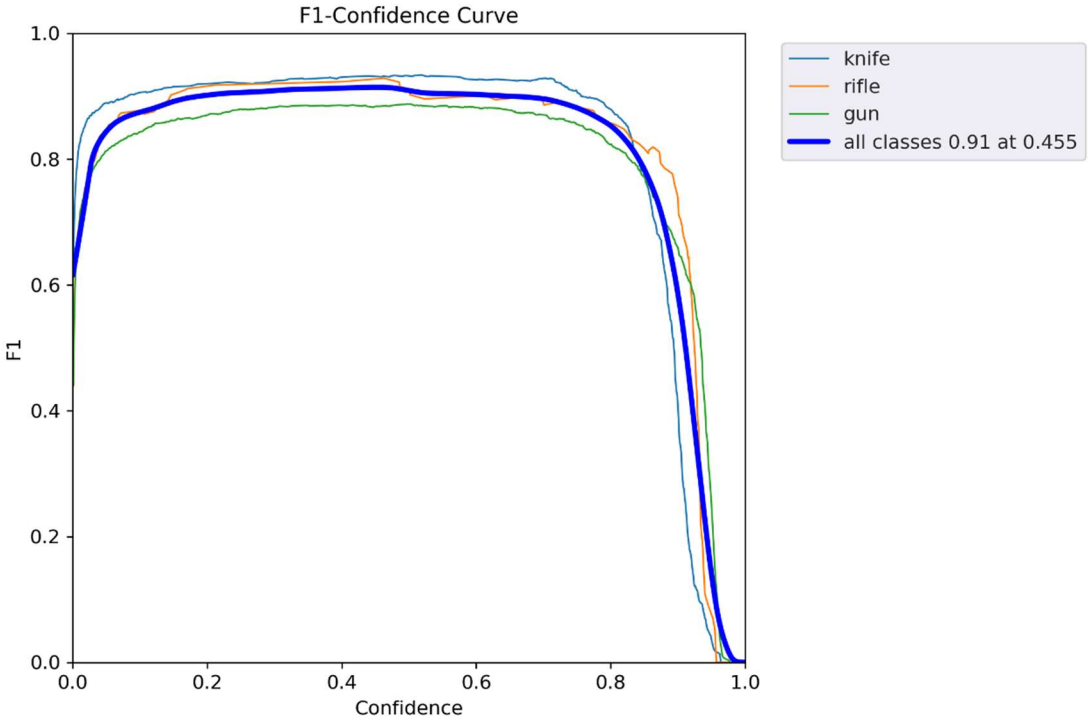
Dedektör	mAP50	Pr	Re	F1
YOLOv8 m Genel	0.94	0.938	0.876	0.9
• Tabanca	0.886	0.908	0.788	0.844
• Bıçak	0.953	0.954	0.9	0.926
• Tüfek	0.982	0.951	0.94	0.946
YOLOv9 GELAN-C	**0.94**	**0.931**	**0.86**	**0.9**
• Tabanca	**0.924**	**0.928**	**0.865**	**0.897**
• Bıçak	**0.964**	**0.952**	**0.912**	**0.931**
• Tüfek	**0.924**	**0.928**	**0.865**	**0.895**

Çizelge 4.2'de görüleceği gibi YOLOv9, validasyon veri seti üzerinde 'Tabanca' sınıf etiketi için daha başarılı tahminleme yapmıştır. Şekil 4.4 – Şekil 4.5'te verilen F1 skor

grafiklerinde de bu durum görülebilir. Elde edilen sonuçlara istinaden YOLOv9 ile hem eğitim hem validasyon veri setleri üzerinde daha başarılı sonuçlar elde edilmiştir.



Şekil 4.4. YOLOv8 kullanılarak elde edilmiş 3 ayrı sınıf etiketi için F1 eğrisi grafiği (Confidence: Modelin tahminlerinin güven seviyesini belirtir.)



Şekil 4.5. YOLOv9 kullanılarak elde edilmiş 3 ayrı sınıf etiketi için F1 eğrisi grafiği (Confidence: Modelin tahminlerinin güven seviyesini belirtir.)

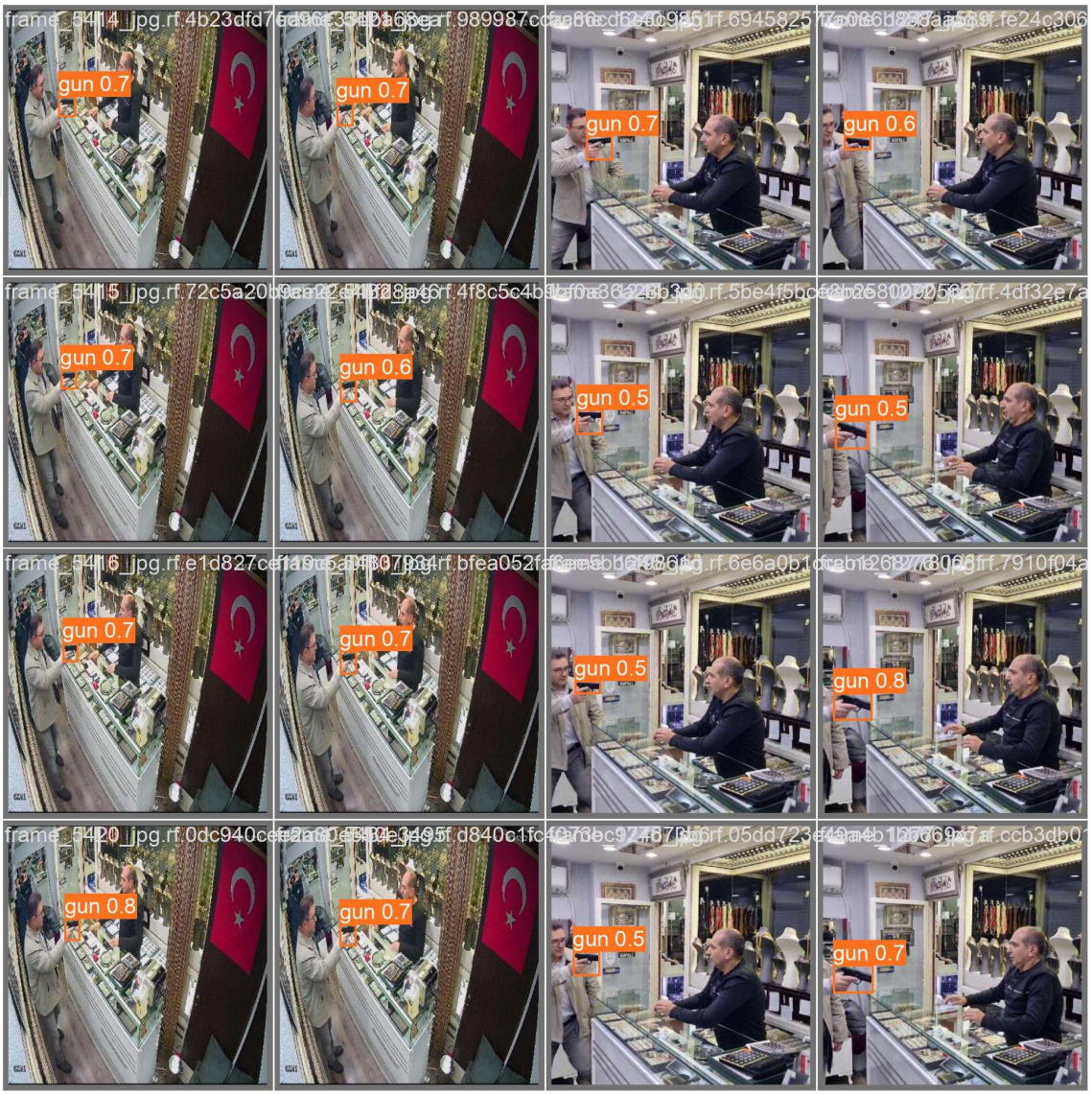
Yapılan deneyler, YOLOv9 GELAN-C modelinin diğer modellere kıyasla üstün bir performans sergilediğini göstermektedir. Bu modelin, diğer modellere göre yüksek doğruluk oranı ve düşük yanlış pozitif oranı, ön plana çıkartmaktadır.

4.1. Güvenlik Sistemlerinde Kullanım Potansiyeli

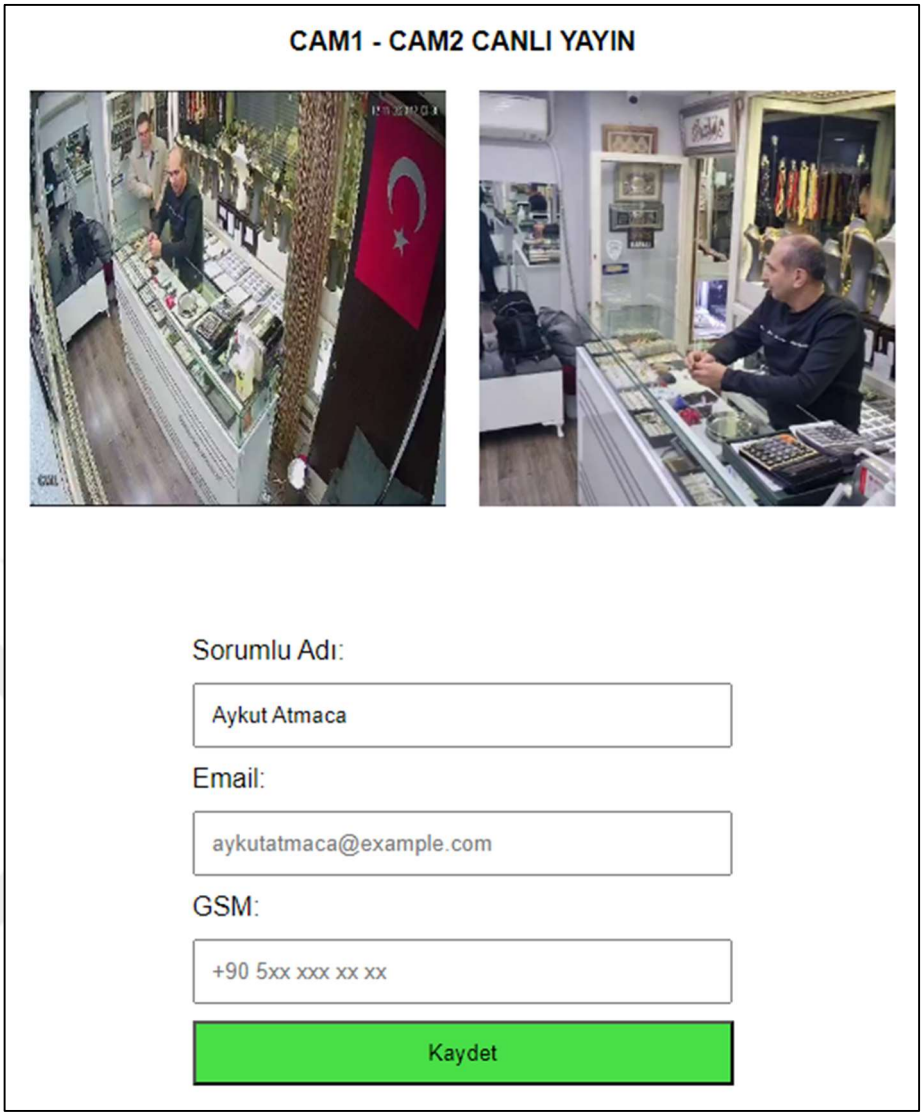
YOLOv9 GELAN-C modelinin üstün performansı, bu modelin gerçek zamanlı güvenlik sistemlerinde etkili bir şekilde kullanılabileceğini göstermektedir. Model, yüksek doğruluk ve hız gerektiren uygulamalarda, örneğin bankalar, kuyumcu ve döviz büroları, havaalanları, alışveriş merkezleri ve kamu binalarındaki güvenlik taramalarında uygulanabilir. Ayrıca, bu modelin düşük yanlış pozitif oranı, gereksiz alarmların sayısını azaltarak güvenlik personelinin verimliliğini artırabilir.

Bu durumu temsilen bir kuyumcuda tabancalı ve bıçaklı soygun simülasyonları gerçekleştirilmiş ve güvenlik kamerası görüntüleri kullanılmıştır. Toplanan

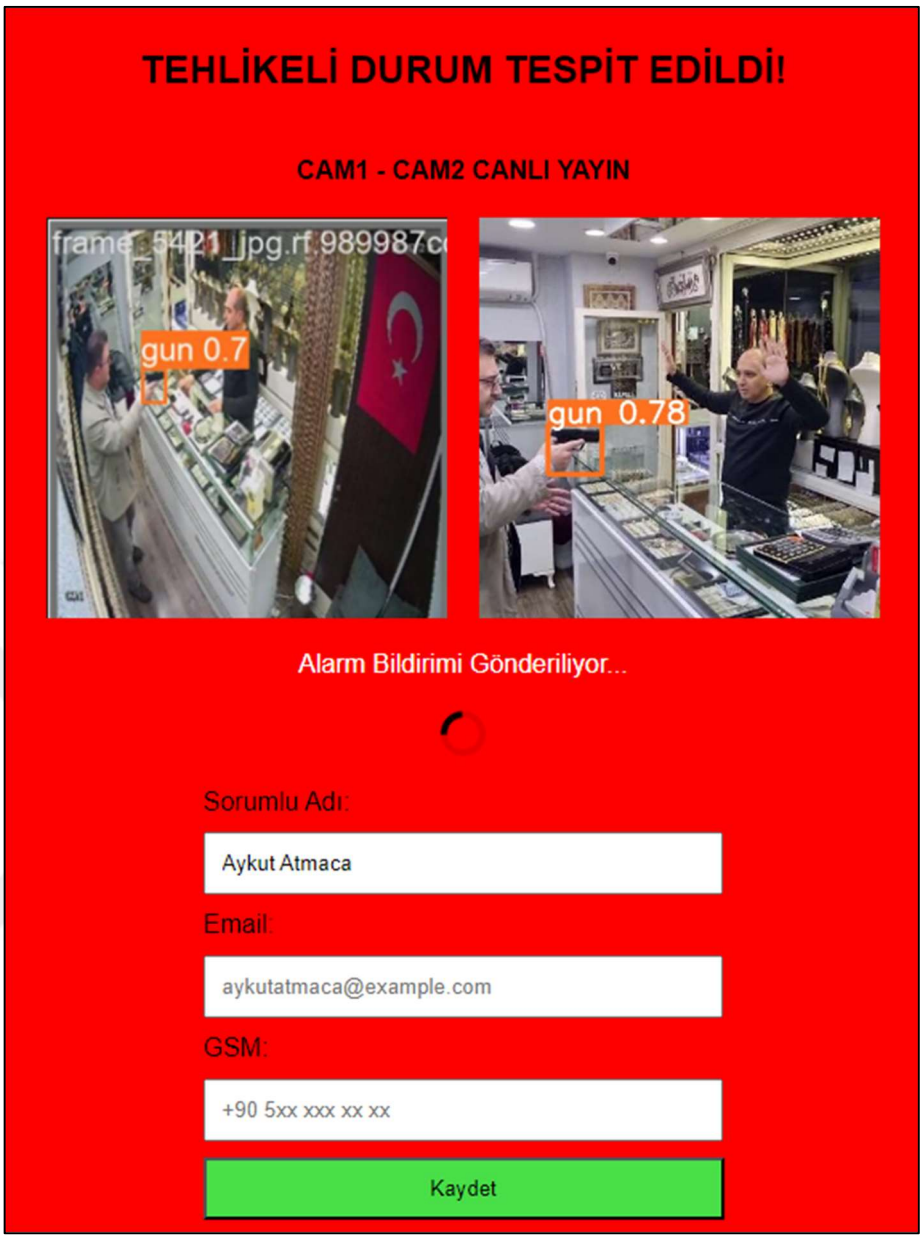
görüntülerden bir kısmı Şekil 4.6'deki gibidir. Bu görüntüler ve YOLOv9 GELAN-C modelini içeren yazılım ile tehlikeli nesne tespiti tahmini gerçekleştirilmiş ve Şekil 4.6'daki tahminler elde edilmiştir. Geliştirilen uyarı sistemi yazılımında Şekil 4.7 herhangi bir alarm durumu olmayan normal durumu temsil ederken, Şekil 4.8 modelin tehlike algıladığı durumu temsil etmektedir. Tehlike algılanan durumda arayüzden girişi sağlanan kullanıcı bilgilerine gerçek zamanlı bildirim gönderilmektedir.



Şekil 4.6. Eğitilen YOLOv9 modelinin kendi oluşturduğumuz bir veri seti örneklemi üzerinde tahminleme performansı



Şekil 4.7 Normal durumu temsil eden uygulama arayüz görseli



Şekil 4.8 Alarm durumunu temsil eden uygulama arayüz görseli

Gerçek zamanlı görüntüler ile yapılan çalışmalar sonucunda YOLOv9 GELAN-C modelinin anlık tehlikeli nesne tespitinde başarılı olduğu kanıtlanmış ve geliştirilen uyarı sisteminin etkin bir şekilde çalıştığı gözlenmiştir. Bu yazılım gerekli donanım ile bir ürüne dönüştüğünde çeşitli mağazalarda, kuyumcularda, banka ve hastaneler gibi kamu kurumlarında kullanımı ile tehlikeli durumların daha erken tespit edilmesi sağlayacağı ve güvenlik sistemlerine destek olacağı tahmin edilmektedir.

5. SONUÇ

Bu çalışmada derin öğrenme yöntemleri ile güvenlik kameraları görüntülerinde tehlikeli nesne tespiti yapılarak uyarıcı sistemlerin aktive edilmesi ve önceden belirlenmiş yetkili kişilerin bildirim ile uyarılmasıyla tehlikeli durumların önlenebilmesi amaçlanmıştır. Literatür çalışmaları sonucunda suç aleti olarak tabancanın diğer nesnelere oranla daha fazla kullanıldığı tespit edilmiştir. Çalışmada tabanca ile birlikte tüfek ve bıçak etiketleri ile toplamda üç sınıf etiketi çıktı olarak kullanılmış ve görüntülerde tespit edilmeye çalışılmıştır. Tabanca tespiti öncelikli olduğundan bu etiketi en yüksek doğrulukta tahmin eden yöntem sonraki aşamalarda kullanılmak üzere seçilmiştir.

Çalışmada, tehlikeli nesnelerin tespiti için farklı derin öğrenme modelleri kullanılmış ve performansları farklı metriklerle karşılaştırılmıştır. Nesne tespit çalışmalarında sıklıkla kullanılan Google MediaPipe ve YOLO modellerinden Mediapipe için MobileNet-MultiHW-AVG, Yolo için YOLO NAS-m, YOLOv5-m, YOLOv8-m ve YOLOv9 GELAN-C modelleri aynı setleri ile eğitilmiş ve performansları karşılaştırılmıştır.

Çalışmada çeşitli derin öğrenme yöntemleri kullanılmış olup, literatürde bilindiği kadarıyla tehlikeli nesne tespitinde ilk kez YOLOv9 kullanılmıştır. YOLOv9, daha önceki versiyonlarına göre daha yüksek doğruluk ve hız sunmakta olup, gerçek zamanlı uygulamalarda kullanılabilmektedir.

Çalışma kapsamında eğitimi sağlanan YOLOv9 modeli, suçun erken tespiti ve önlenmesi hususunda bir yazılıma dönüştürülerek literatürde ilk kez YOLOv9 modeli içeren bir erken uyarı sistemi geliştirilmiştir. Ayrıca bir mağazada tabanca ve bıçaklı soygun simülasyonları gerçekleştirilmiş ve güvenlik kamerası görüntüleri toplanmıştır. Bu görüntüler ile geliştirilen yazılım kullanıldığında gerçek hayat güvenlik kamerası görüntüleri ile de başarılı tespitler yapıldığı görülmüştür. Gerekli donanımsal altyapı ile desteklenen yazılım bir ürüne dönüştürüldüğünde, güvenlik kameralarından gelen görüntüler üzerinde anlık olarak tehlikeli nesne tespiti yapılıp, ilgili kişi ve kurumlara mesaj veya alarm yoluyla bildirim gönderilebilecektir. Bu sistemin özellikle yoğun

nüfuslu ve yüksek suç oranlarına sahip bölgelerde, suç oranlarının azaltılmasında ve genel güvenlik seviyesinin artırılmasına katkı sağlayacağı düşünülmektedir.

Bu çalışmada elde edilen sonuçlar, gelecekte yapılacak çalışmalar için önemli bir temel oluşturmaktadır. YOLOv9 GELAN-C modelinin performansını daha da artırmak için, farklı veri seti arttırma yöntemleri ve ön işleme metotları kullanılabilir. Ayrıca, farklı senaryolarda ve koşullarda test edilmesi, modelin güvenilirliğini ve yaygınlaşmasını arttıracaktır. Gelecekteki çalışmalar, daha hafif ve hızlı modellerin geliştirilmesine de odaklanabilir, bu da donanımsal olarak daha maliyet etkin çözümler sunabilir.

Sonuç olarak, yapılan çalışma ile, YOLOv9 GELAN-C modeli, tehlikeli nesne tespitinde üst düzey performans sergileyen bir yöntem olarak öne çıkmaktadır ve güvenlik sistemlerinde geniş bir kullanım potansiyeli bulunmaktadır. Gelecekte yapılacak iyileştirmeler, genişletilmiş testler, modelin daha da etkili hale gelmesini sağlayacaktır.

KAYNAKLAR

A, R. R., Maddileti, T., J, S., Srinivas, R., & Saikumar, K. (2022). Pseudo Trained YOLO R_CNN Model for Weapon Detection with a Real-Time Kaggle Dataset. *International Journal of Integrated Engineering*, *14*(7), Article 7.

Akbulut, Y., & Khalaf, R. (2021). Smart Arms Detection System Using YOLO Algorithm and OpenCV Libraries. *Turkish Journal of Science and Technology*, *16*(1), Article 1.

Ashraf, A., Imran, M., Qahtani, A., Alsufyani, A., Almutiry, O., Mahmood, A., Attique, M., & Habib, M. (2021). Weapons Detection for Security and Video Surveillance Using CNN and YOLO-V5s. *Computers, Materials & Continua*, *70*(2), 2761-2775. https://doi.org/10.32604/cmc.2022.018785

Babila, I. F. E., Villasor, S. A. E., & Dela Cruz, J. C. (2022). Object Detection for Inventory Stock Counting Using YOLOv5. *2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA)*, 304-309. https://doi.org/10.1109/CSPA55076.2022.9782028

Berardini, D., Migliorelli, L., Galdelli, A., Frontoni, E., Mancini, A., & Moccia, S. (2024). A deep-learning framework running on edge devices for handgun and knife detection from indoor video-surveillance cameras. *Multimedia Tools and Applications*, *83*(7), 19109-19127. https://doi.org/10.1007/s11042-023-16231-x

Bilgin, R. (2016). ÇATIŞMA VE ŞİDDET ORTAMINDA BÜYÜYEN ÇOCUKLAR SORUNU. *Fırat Üniversitesi Sosyal Bilimler Dergisi*, *24*(1), Article 1. https://doi.org/10.18069/fusbed.51605

Castillo, A., Tabik, S., Pérez, F., Olmos, R., & Herrera, F. (2019). Brightness guided preprocessing for automatic cold steel weapon detection in surveillance videos with deep learning. *Neurocomputing*, *330*, 151-161. https://doi.org/10.1016/j.neucom.2018.10.076

Chu, G., Arikan, O., Bender, G., Wang, W., Brighton, A., Kindermans, P.-J., Liu, H., Akin, B., Gupta, S., & Howard, A. (2021). *Discovering Multi-Hardware Mobile Models via Architecture Search* (arXiv:2008.08178). arXiv. http://arxiv.org/abs/2008.08178

Dollar, P., Wojek, C., Schiele, B., & Perona, P. (2012). Pedestrian Detection: An Evaluation of the State of the Art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *34*(4), 743-761. https://doi.org/10.1109/TPAMI.2011.155

Dong, K., Zhou, C., Ruan, Y., & Li, Y. (2020). MobileNetV2 Model for Image Classification. *2020 2nd International Conference on Information Technology and Computer Application (ITCA)*, 476-480. https://doi.org/10.1109/ITCA52113.2020.00106

Duran-Vega, M. A., Gonzalez-Mendoza, M., Chang, L., & Suarez-Ramirez, C. D. (2021). *TYolov5: A Temporal Yolov5 Detector Based on Quasi-Recurrent Neural Networks for Real-Time Handgun Detection in Video* (arXiv:2111.08867). arXiv. https://doi.org/10.48550/arXiv.2111.08867

Estimating Global Civilian-held Firearms Numbers. (2018, Haziran 5). https://www.smallarmssurvey.org/resource/estimating-global-civilian-held-firearms-numbers

Fan, Q., Brown, L., & Smith, J. (2016). A closer look at Faster R-CNN for vehicle detection. *2016 IEEE Intelligent Vehicles Symposium (IV)*, 124-129. https://doi.org/10.1109/IVS.2016.7535375

Jesus, S. M., Lomas, V., Ruíz-Santaquiteria, J., & Deniz, O. (2021). Automatic Handgun Detection with Deep Learning in Video Surveillance Images. *Applied Sciences*, *11*(13). https://doi.org/10.3390/app11136085

Jiao, L., Zhang, F., Liu, F., Yang, S., Li, L., Feng, Z., & Qu, R. (2019). A Survey of Deep Learning-Based Object Detection. *IEEE Access*, *7*, 128837-128868. https://doi.org/10.1109/ACCESS.2019.2939201

K U, F. S., & M, S. (2023). Subduing Crime and Threat in Real-Time by Detecting Weapons Using Yolov8. *2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)*, 864-868. https://doi.org/10.1109/ICCPCT58313.2023.10245146

Khalid, S., Waqar, A., Ain Tahir, H. U., Edo, O. C., & Tenebe, I. T. (2023). Weapon detection system for surveillance and security. *2023 International Conference on IT Innovation and Knowledge Discovery (ITIKD)*, 1-7. https://doi.org/10.1109/ITIKD56332.2023.10099733

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436-444. https://doi.org/10.1038/nature14539

Lim, J., Al Jobayer, M. I., Baskaran, V. M., Lim, J. M., See, J., & Wong, K. (2021). Deep multi-level feature pyramids: Application for non-canonical firearm detection in video surveillance. *Engineering Applications of Artificial Intelligence*, *97*, 104094. https://doi.org/10.1016/j.engappai.2020.104094

Narejo, S., Pandey, B., Esenarro Vargas, D., Rodriguez, C., & Anjum, M. (2021). Weapon Detection Using YOLO V3 for Smart Surveillance System. *Mathematical Problems in Engineering*, *2021*, 1-9. https://doi.org/10.1155/2021/9975700

Number of robberies in the U.S. by weapon 2022. (t.y.). Statista. Geliş tarihi 19 Mayıs 2024, gönderen https://www.statista.com/statistics/251914/number-of-robberies-in-the-us-by-weapon/

Offences involving the use of weapons—Office for National Statistics. (t.y.). Geliş tarihi 19 Mayıs 2024, gönderen https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/compendium/focusonviolentcrimeandsexualoffences/yearendingmarch2015/chapter3offencesinvolvingtheuseofweapons#types-of-offences-where-firearms-are-used

Olmos, R., Tabik, S., & Herrera, F. (2018). Automatic handgun detection alarm in videos using deep learning. *Neurocomputing*, *275*, 66-72. https://doi.org/10.1016/j.neucom.2017.05.012

Pak, A., & Gannon, B. (2023). The effect of neighbourhood and spatial crime rates on mental wellbeing. *Empirical Economics*, *64*(1), 99-134. https://doi.org/10.1007/s00181-022-02256-y

Parlak, İ. E., & Emel, E. (2023). Deep learning-based detection of aluminum casting defects and their types. *Engineering Applications of Artificial Intelligence*, *118*, 105636. https://doi.org/10.1016/j.engappai.2022.105636

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer*

Vision and Pattern Recognition (CVPR), 779-788. https://doi.org/10.1109/CVPR.2016.91

Rehman, A., & Fahad, L. G. (2022). Real-Time Detection of Knives and Firearms using Deep Learning. *2022 24th International Multitopic Conference (INMIC)*, 1-6. https://doi.org/10.1109/INMIC56986.2022.9972915

Robbery. (t.y.). FBI. Geliş tarihi 07 Nisan 2024, gönderen https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/topic-pages/robbery

Rowley, H. A., Baluja, S., & Kanade, T. (1998). Neural network-based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(1), 23-38. https://doi.org/10.1109/34.655647

Tram, N., Son, D., & Thái, A. (2023). *Weapon Detection Using Deep Learning* (s. 109). https://doi.org/10.1145/3628797.3628967

Velastin, S., Boghossian, B., & Vicencio-Silva, M. (2006). A Motion-Based Image Processing System for Detecting Potentially Dangerous Situations in Underground Railway Stations. *World Transit Research*. https://www.worldtransitresearch.info/research/2001

Verma, G., & Dhillon, A. (2017). *A Handheld Gun Detection using Faster R-CNN Deep Learning* (s. 88). https://doi.org/10.1145/3154979.3154988

Warsi, A., Abdullah, M., Husen, M. N., Yahya, M., Khan, S., & Jawaid, N. (2019). *Gun Detection System Using Yolov3* (s. 4). https://doi.org/10.1109/ICSIMA47653.2019.9057329

What is R-CNN? (2023, Eylül 25). Roboflow Blog. https://blog.roboflow.com/what-is-r-cnn/

Yadav, P., Gupta, N., & Sharma, P. K. (2024). Robust weapon detection in dark environments using Yolov7-DarkVision. *Digital Signal Processing*, *145*, 104342. https://doi.org/10.1016/j.dsp.2023.104342

Zawitz, M. W. (t.y.). *Guns Used in Crime*. https://bjs.ojp.gov/content/pub/pdf/GUIC.PDF