



**T.C.
İSTANBUL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**



Doktora Tezi

**Videolarda Derin Öğrenme Yaklaşımları ile Anormal Durum Tespiti
ve Boyut İndirgeme Sistemi**

Mehmet Tevfik AĞDAŞ

Enformatik Anabilim Dalı

Enformatik Programı

**DANIŞMAN
Prof. Dr. Sevinç GÜLSEÇEN**

Aralık, 2022

İSTANBUL

Bu alıřma, 30.12.2022 tarihinde ařağıdaki jüri tarafından Enformatik Anabilim Dalı, Enformatik Programında Doktora tezi olarak kabul edilmiştir.

Tez Jürisi

Prof. Dr. Sevin GÜLSEÇEN (Danışman)
İstanbul Üniversitesi
Enformatik Bölümü

Prof. Dr. Ümit GÜZ
Iřık ÜNİVERSİTESİ
Mühendislik ve Doęa Bilimleri Fakültesi

Do. Dr. Çiğdem SELÇUKCAN EROL
İstanbul ÜNİVERSİTESİ
Enformatik Bölümü

Do. Dr. Muammer TÜRKOĞLU
Samsun ÜNİVERSİTESİ
Mühendislik Fakültesi

Do. Dr. Murat GEZER
İstanbul ÜNİVERSİTESİ
Enformatik Bölümü

İntihal Programı Beyanı

20.04.2016 tarihli Resmi Gazete’de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince; Bu Lisansüstü teze, İstanbul Üniversitesi’nin aboneli olduğu intihal yazılım programı kullanılarak Fen Bilimleri Enstitüsü’nün belirlemiş olduğu ölçütlere uygun rapor alınmıştır.

Tezden Üretilmiş Yayınların Künye Bilgileri

AĞDAŞ, M. T., & GÜLSEÇEN, S. (2022). Güvenlik Kameralarında Otomatik Silah ve Bıçak Tespit Sistemi: Karşılaştırmalı YOLO Modelleri. Avrupa Bilim ve Teknoloji Dergisi, (41), 16-22.



ÖNSÖZ

Doktora eğitimimin başladığı gün itibariyle, eğitimlerim ve tez çalışmalarımın her aşamasında bilgi ve birikimlerini sabırla paylaşan, anlayışlı, yol ve yön gösteren ve çok değerli akademik ve insani deneyimler edindiğim, ulusal ve uluslararası birçok akademik çalışmaya çok önemli katkıları olan ve değerli danışmanım Prof. Dr. Sevinç GÜLSEÇEN'e,

Doktora çalışmalarımın her aşamasında akademik deneyimlerini benimle paylaşan, yol ve yön gösteren Samsun ÜNİVERSİTESİ öğretim üyelerinden Doç. Dr. Muammer TÜRKOĞLU'na

Doktora eğitimim ve tez çalışmalarında yol gösteren ve deneyimlerinden faydalandığım İstanbul ÜNİVERSİTESİ öğretim üyelerinden Doç. Dr. Çiğdem SELÇUKCAN EROL'a,

Doktora tez çalışmalarında yol ve yön gösteren ve deneyimlerinden faydalandığım Işık ÜNİVERSİTESİ öğretim üyelerinden Prof. Dr. Ümit GÜZ'e,

Doktora eğitimlerime başladığım günden beri yenilikçi düşünmemi sağlayan, hayata ve iş hayatıma bakış açışımı değiştiren İstanbul ÜNİVERSİTESİ öğretim üyelerinden Prof. Dr. Zümrüt ECEVİT SATI'ya,

Hayatımın her aşamasında karşılaştığım stres ve sıkıntıları aşılamamda yardımlarını hiçbir zaman esirgemeyen, hayat yoldaşım ve kıymetli eşim Asuman AĞDAŞ'a ve hayatımı daha anlamlı kılan ve benim için çok kıymetli olan evlatlarım Ahmet Mete ve Melis Azra'ya,

Bu günlere gelmemde emeğini, sevgisini ve desteklerini eksik etmeyen haklarını hiçbir zaman ödeyemeyeceğim çok kıymetli annem ve babama,

Şükranlarımı sunar, teşekkür ederim.

Aralık 2022

Mehmet Tevfik AĞDAŞ

İÇİNDEKİLER

Sayfa No

ÖNSÖZ	iv
İÇİNDEKİLER.....	v
ŞEKİL LİSTESİ	viii
TABLO LİSTESİ.....	xii
SİMGE VE KISALTMA LİSTESİ	xiii
ÖZET	xiv
SUMMARY	xvi
1. GİRİŞ	1
1.1. TEZİN KAPSAMI.....	2
1.2 TEZİN ÖZGÜN DEĞERİ.....	4
1.3 TEZİN TASLAĞI	11
2. GENEL KISIMLAR.....	14
2.1 MAKİNE ÖĞRENMESİ.....	14
2.2 DERİN ÖĞRENME	14
2.2.1. Yapay Sinir Ağları.....	16
2.2.1.1 Tekrarlayan Sinir Ağları (Recurrent Neural Network (TSA-RNN)).....	17
2.2.1.1.1 Uzun-Kısa Süreli Bellek (UKSB- LSTM).....	18
2.2.1.1.2 Kapılı Tekrarlayan Birimler (KTP-GRU)	19
2.2.2. Evrişimsel Sinir Ağları (ESA-CNN)	21
2.2.2.1.1 Giriş Katmanı.....	21
2.2.2.1.2 Evrişim Katmanı	21
2.2.2.1.3 Havuzlama Katmanı	23
2.2.2.1.4 Aktivasyon Fonksiyonu	23
2.2.2.1.5 Normalizasyon Katmanı	24
2.2.2.1.6 Tam Bağlı Katman.....	24
2.2.2.1.7 DropOut Katmanı	24
2.2.2.1.8 Sınıflandırma Katmanı.....	25
2.2.2.2 ESA Mimarileri	25
2.2.2.2.1 AlexNet.....	26
2.2.2.2.2 GoogleNet.....	27

2.2.2.2.3 VGG16.....	29
2.2.2.2.4. VGG19.....	29
2.2.2.2.5. SqueezeNet	30
2.2.2.2.6. Inception	31
2.2.2.2.7. ResNet.....	32
2.2.2.2.8 DenseNet201.....	33
2.2.3 Görüntülerde Nesne Algılama ve Tespiti Modelleri	34
2.2.3.1 EfficientDet.....	34
2.2.3.2 Faster RCNN	35
2.2.3.3 YOLOv4	35
2.2.3.4 YOLOv5	37
2.2.3.5 YOLOX	37
2.2.3.6 YOLOR.....	38
2.3. BOYUT İNDİRGEME.....	39
2.3.1. Video Boyut İndirilmesi	40
2.3.2 Görüntü İşleme ve Filtreleme Yaklaşımları	41
2.3.2.1. Görüntü İşleme İkili Görüntü Dönüşümü.....	41
2.3.2.2. Görüntünün Gri Ölçek Dönüşümü.....	42
2.3.2.3. Histogram Eşitleme Yöntemi	43
2.3.2.4 Thresholding Yöntemi	44
2.3.3. Görüntü Filtreleme	47
2.3.3.1. Konvolüsyon İşlemi.....	48
2.3.3.2. Ortalama Filtre	48
2.3.3.3. Orta Değer Medyan Filtre.....	49
3. MALZEME VE YÖNTEM.....	50
3.1 TEZİN BİRİNCİ AŞAMASI	50
3.1.1 Yöntemler	50
3.1.2 Veri Kümeleri.....	50
3.1.2.1 Veri Kümelerinin Özellikleri.....	50
3.1.2.2. Veri Kümelerinin Toplanması.....	51
3.2 TEZİN İKİNCİ AŞAMASI.....	58
3.2.1 Boyut İndirgeme Modeli	58
4. BULGULAR.....	67
4.1 BİRİNCİ BÖLÜM BULGULARI	67

4.1.1 Anlık Görüntülerde Silah ve Bıçak Algılanması Çalışması Bulguları.....	67
4.1.2 Anlık Görüntülerde Şiddet Algılanması Çalışması Bulguları	70
4.2 İKİNCİ BÖLÜM BULGULARI.....	73
4.2.1. Güvenlik Kamerası Kayıtlarındaki Silah ve Bıçak Görüntülerinin Sınıflandırılması.	74
4.2.2. Güvenlik Kamerası Kayıtlarındaki Şiddet Görüntülerinin Sınıflandırılması.	78
4.2.3 Boyut İndirgeme Çalışması ve Kullanıcı Arayüzü.....	92
4.2.3.1 Canlı Kayıt ve Boyut Analizi Sekmesi	94
4.2.3.2 Görüntü Analizi Sekmesi.....	97
4.2.3.3 Sistem Sekmelerinde Ortak Kullanılan Menü ve Komutlar	100
5. TARTIŞMA VE SONUÇ	107
5.1 Boyut İndirgeme ve Görüntü Analizi Çalışmalarının Amaç ve Sonuçları.....	107
5.2 Öneriler.....	116
KAYNAKLAR.....	117
ÖZGEÇMİŞ	126

ŞEKİL LİSTESİ

Sayfa No

Şekil 1.1: Boyut indirgeme sistemi şeması.	7
Şekil 1.2: Boyut indirgeme algoritması kısa taslağı.....	12
Şekil 2.1: Tez çalışmasının derin öğrenme tabanlı modelleri.	15
Şekil 2.2: Sinir hücrelerinin biyolojik yapısı (Yakıt & Özkan, 2017).	16
Şekil 2.3: YSA hücresinin matematiksel yapısı.	17
Şekil 2.4: TSA açılımı.....	18
Şekil 2.5: UKSB yapısı blok şeması (Greff vd., 2017).	19
Şekil 2.6: GRU modeli hücre yapısı (Pavithra vd., 2019).	20
Şekil 2.7: Evrişim katmanında 5x5x3 görüntüye 3x3 filtrenin uygulanması örneği.	22
Şekil 2.8: Havuzlama katmanı örneği.	23
Şekil 2.9: DropOut düğüm kaldırma örneği.....	25
Şekil 2.10: AlexNet mimarisinin katmanları ve parametreleri (Krizhevsky vd.,2017).	27
Şekil 2.11: GoogleNet mimari yapısı (Szegedy vd.,2015).	28
Şekil 2.12: VGG16 mimarisinin katmanlı ağ yapısı (Simonyan & Zisserman, 2015).	29
Şekil 2.13: VGG19 mimarisinin ağ yapısı (Bansal vd., 2016).	30
Şekil 2.14: SqueezeNet mimarisi genişletilmiş AğYapısı (Lee vd., 2019).	31
Şekil 2.15: Inception mimarisi ağ yapısı (Journal t.y., 2019).	32
Şekil 2.16: Artık blok modül çalışma yapısı.	32
Şekil 2.17: EffcientDet mimarisi (Tan vd., 2020).....	34
Şekil 2.18: Faster R-CNN mimarisi (Yılmaz vd., 2020).....	35
Şekil 2.19: Görüntü Kesit Örtüşmesi (IoU).	36
Şekil 2.20: YOLOX Modeli ile YOLOv3-v5 Modelinin Yapısı Arasındaki Fark (Ge vd., 2021).....	38

Şekil 2.21: YOLOR modelinin çalışma yapısı.....	39
Şekil 2.22: Renkli görüntü ikili form çıktısı.	42
Şekil 2.23: Renkli görüntünün gri renk dönüşümü.	43
Şekil 2.24: Histogram eşitleme örneği.	44
Şekil 2.25: Thres holding yöntemleri ile iyileştirilmiş görüntü örnekleri.....	45
Şekil 2.26: Histogram eşitleme yöntemlerinin OpenCV kodlaması.	45
Şekil 2.27: Histogram eşitleme yöntemlerinin OpenCV örneği çıktısı.....	46
Şekil 2.28: Bölgesel uyarlanabilir eşileme kodlaması ve ekran çıktısı.....	47
Şekil 2.29: Konvolüsyon matris çarpımı örneği.	48
Şekil 2.30: Orta değer medyan filtresi örneği.....	49
Şekil 3.1: Bıçak veri kümesi örnek görüntüleri.	53
Şekil 3.2: Silah veri kümesi örnek görüntüleri.....	53
Şekil 3.3: Şiddet veri kümesi örnek görüntüler,(a) Şiddet içeren görüntüler,(b) Şiddet içermeyen görüntüler.	54
Şekil 3.4: The crowd violence veri kümesi örnek görüntüler.....	55
Şekil 3.5: The fight detection surv veri kümesi örnek görüntüler.....	56
Şekil 3.6: Movie violence veri kümesi örnek görüntüler.	56
Şekil 3.7: Silah bıçak algılanması veri kümesi örnek, (a) silah verileri, (b) bıçak verileri, (c) normal görüntüler.	57
Şekil 3.8: Boyut indirgeme modeli örnek şeması.	60
Şekil 3.9: Boyut indirgeme modeli algoritması.	62
Şekil 4.1: Deneysel çalışma analizlerinin grafiksel gösterimi.	68
Şekil 4.2: Silah ve bıçak algılanması çalışması nesnel sonuç.	69
Şekil 4.3: Silah ve bıçak algılanması çalışması yüzdesel sonuç.	69
Şekil 4.4: EfiicientDet D0-D7 modeli grafiksel çıktısı.	71
Şekil 4.5: YOLOv5 modeli grafiksel çıktısı.	72
Şekil 4.6: YOLO anlık görüntülerde şiddet algılama sistemi örnek çıktıları.v5 modeli grafiksel çıktısı.	73

Şekil 4.7: Anlık görüntülerde şiddet algılama sistemi örnek çıktıları(Yüzdesel).	73
Şekil 4.8: 2-Sınıflı veri kümelerinin deneysel çalışmalarının Karışıklık Matrisi çıktısı.	76
Şekil 4.9: 3-Sınıflı veri kümesi Karışıklık Matrisi çıktıları.	77
Şekil 4.10: Şiddet görüntülerinin sınıflandırılması için önerilen yaklaşımın genel yapısı.....	79
Şekil 4.11: Önerilen modelin özellik çıkarma mimarisi.	80
Şekil 4.12: SURV veri kümesi için önerilen modellerin Karışıklık Matrisi değerleri, a) Özellik birleştirme yaklaşımı, b) Ortalama Havuzlama yaklaşımı, c) Standart Sapma Havuzlama yaklaşımı.	82
Şekil 4.13: Crown veri kümesi için önerilen modellerin Karışıklık Matrisi değerleri, a) Özellik Birleştirme yaklaşımı, b) Ortalama Havuzlama yaklaşımı, c) Standart Sapma Havuzlama yaklaşımı.	83
Şekil 4.14: Movie Violence veri kümesi için önerilen modellerin Karışıklık Matrisi değerleri, a) Füzyon yaklaşımı, b) Ortalama Havuzlama yaklaşımı, c) Standart Sapma Havuzlama yaklaşımı.	85
Şekil 4.15: Önerilen Kombine modelin mimarisi.	88
Şekil 4.16: Crowd veri kümesi (a,b,c) ve Surv veri kümesi (d,e,f), a-d) GRU yöntemi, b-e) LSTM yöntemi c-f) Önerilen Yaklaşım	91
Şekil 4.17: Görüntü analizi ve boyut indirgeme sistemi arayüzü.	92
Şekil 4.18: Görüntü analizi ve boyut indirgeme sistemi arayüzü, görüntü sekmesi.	94
Şekil 4.19: Güvenlik kamerası canlı kayıt görüntüsü.	95
Şekil 4.20: Güvenlik kamerası arşiv görüntüsü.	96
Şekil 4.21: Görüntü analizi sekmesi komutları, a) açılış, b) sistem aktif, c) aktif ve şiddet tespiti.....	97
Şekil 4.22: Görüntü analizi sekmesi, şiddet tespiti.	99
Şekil 4.23: Boyut indirgeme işlemi örnek görüntüsü.....	101
Şekil 4.24: Boyut indirgeme örnek kazanım bilgileri.	102
Şekil 4.25: Anasayfa menü sayfası görüntüsü.	102
Şekil 4.26: Ayarlar menü sayfası görseli.	103
Şekil 4.27: Kazanım menü sayfası görseli.	104
Şekil 4.28: Olay günlüğü menü sayfası görseli.....	105

Şekil 4.29: Arşiv olay günlüğü menü sayfası görseli.....	105
Şekil 4.30: Destek menü sayfası görseli.	105
Şekil 5.1: Görüntü analizi ve boyut indirgeme kullanıcı arayüzü uygulaması.	113
Şekil 5.2: Görüntü analizi ve boyut indirgeme kolay kurulum ve tanıtım görseli.	115



TABLO LİSTESİ

	Sayfa No
Tablo 1.1: Kayıt cihazlarının depolama birimi kapasitesine göre gün hesabı.....	3
Tablo 1.2: Çalışmada parametrelerinin seçiminde incelenen çalışmalar.....	9
Tablo 2.1: ESA mimarilerinin karakteristik özellikleri.	26
Tablo 2.2: ResNet Bağlantı ve Katmanları.....	33
Tablo 3.1: Analiz Türleri ve Veri Toplama Kaynakları.	51
Tablo 3.2: Boyut indirgeme sınıfları.	60
Tablo 4.1: Çalışmada elde edilmiş karşılaştırmalı yüzdesel skorlar.	67
Tablo 4.2: Anlık Görüntülerde Şiddet Algılama Deneysel Sonuçlar.	70
Tablo 4.3: Anlık Görüntülerde Şiddet Algılama Deneysel Sonuçlar-2.....	71
Tablo 4.4: Deneysel Çalışmalarda Kullanılan Ağ Parametreleri.....	74
Tablo 4.5: 2-Sınıflı Veri Setinin Başarım Skorları(%).	75
Tablo 4.6: 3-Sınıflı Veri Setinin Başarım Skorları(%).	76
Tablo 4.7: Deneysel Sonuçların Literatür Karşılaştırması.	77
Tablo 4.8: Surveillance Camera Fight veri kümesi deneysel sonuçları.....	82
Tablo 4.9: Crowd Violence veri kümesi için önerilen modelin performans sonuçları (%)..	83
Tablo 4.10: Movie Violence veri kümesi için önerilen modelin performans sonuçları (%).	84
Tablo 4.11: Önerilen model ile önceki çalışmaların doğruluk puanlarının (%) karşılaştırılması.....	85
Tablo 4.12: Deneysel çalışma analizlerinin grafiksel gösterimi.....	89
Tablo 4.13: Deneysel çalışma analizlerinin grafiksel gösterimi.....	91
Tablo 5.1: Silah ve bıçak görüntülerinin algılanması başarım skorları.	108
Tablo 5.2: Görüntülerde anlık şiddet olaylarının algılanması.	108
Tablo 5.3: Silah ve bıçak tespiti deneysel başarım skorları.....	109
Tablo 5.4: Şiddet görüntülerinin sınıflandırılması modeller ve başarım değerleri.....	111

SİMGE VE KISALTMA LİSTESİ

Simgeler	Açıklama
----------	----------

tb	: Terabyte
gb	: Gigabayte
b	: Byte

Kısaltmalar	Açıklama
-------------	----------

CNN	: Konvolüsyonel Sinir Ağları
DWT	: Ayrık Dalgacık Dönüşümü Yöntemi
FPS	: Birim Resim Karesi
CCTV	: Kapalı Devre Gözetleme Sistemi
GRU	: Kapalı Tekrarlayan Birimler Modeli
LSTM	: Kısa - Uzun Süreli Bellek Modeli
MCML	: Daralan Metrik Öğrenme Yöntemi
PCA	: Temel Bileşen Analizi Yöntemi
RNN	: Tekrarlayan Sinir Ağları Modeli
SVM	: Destek Vektör Makineli Modeli
YSA	: Yapay Sinir Ağları
TSA	: Tekrarlayan Sinir Ağları
ESA	: Evrişimli Sinir Ağları

ÖZET

DOKTORA TEZİ

VIDEOLARDA DERİN ÖĞRENME YAKLAŞIMLARI İLE ANORMAL DURUM TESPİTİ VE BOYUT İNDİRGEME SİSTEMİ

Mehmet Tevfik AĞDAŞ

İstanbul Üniversitesi

Fen Bilimleri Enstitüsü

Enformatik Anabilim Dalı

Danışman : Prof. Dr. Sevinç GÜLSEÇEN

Kurum ve kuruluşlarda, videolarda anormal durumların tespit edilmesi için genellikle güvenlik kamera sistemleri kullanılmaktadır. Güvenlik kamera sistemlerinin yönetilmesinde ise güvenlik personeli kaynağı kullanılmaktadır. Güvenlik personellerinin dikkati ve takibatları sayesinde, güvenlik kamerası sistemlerindeki video görüntülerinde anormal bir durum tespit edilip gerekli müdahaleler yapılmaktadır. Güvenlik personeli kaynağının insan olması sebebiyle sahip olmaları gereken dikkat ve takibatlarını, fizyolojik, biyolojik ve ruhsal gibi birçok insani yapı etkilemektedir.

Güvenlik kamerası videolarının anlık takip edilmesinin yanı sıra depolanması da oldukça önemlidir. Videolarda detaylı analiz işlemleri ve hukuki takibatının yapılabilmesi için videoların kayıt edilmesi ve depolanması gerekmektedir. Güvenlik kamera sistemlerindeki video kayıtlarının depolanmasında geçmişe dönük zamansal artış doğrudan sistemin kullanışlılığını, güvenilirliğini ve beraberinde dezavantajlı bir durum olan sistemin maliyetini artırmaktadır.

Bu tez çalışması iki aşamalı olarak planlanmıştır. Birinci aşamada, videolardaki anormal durumların, derin öğrenme tabanlı ve insani yapılardan etkilenmeyen bilgisayarlı görüntüleme teknikleri ile tespit edilmesi amaçlanmıştır. Bu aşamadaki çalışma neticesinde, anlık video görüntülerinde anormal durumlardan otomatik silah, bıçak ve şiddet tespiti sistemi yapılmıştır. Bu çalışma Covid-19 pandemisi dönemlerinde yapıldığından, videolarda anormal durum tespit işlemlerinin yapılabilmesi için gerçekleştirilen model eğitimlerinde veri kümeleri

anonim ve telif hakkı içermeyen internet ortamlarından, YouTube görüntülerinden ve açık kaynak erişimli olan akademik çalışma veri kümelerinden derlenmiştir. Deneysel çalışmalarda derin öğrenme tabanlı, görüntü işleme tekniklerinden YOLOv4, YOLOv5, YOLOX ve YOLOR modelleri kullanılmıştır. En yüksek başarımlı değeri, YOLOR modeli ile % 97,6 olarak elde edilmiştir. Şiddet tespiti için ise, EfficientDet D0-D7, YOLOv5 ve Faster R-CNN modeli uygulanmış ve en iyi başarımlı değeri EfficientDET D0-D7 modeli ile %94,1 elde edilmiştir.

Bu tez çalışmasının ikinci aşamasında, güvenlik kamerası video kayıt görüntülerinde anormal durumların sınıflandırılması ve görüntülerin depolamasında boyut indirgenmesi çalışması yapılmıştır. Görüntü sınıflandırması için veri kümeleri, tezin birinci aşamasındaki anormal durum algılanması ve tespitinde olduğu gibi anonim telif hakkı içermeyen veri kümelerinin derlemesinden oluşturulmuştur. Bu sınıflandırma işlemleri için derin öğrenme temelli önceden eğitilmiş derin ESA mimarilerinden AlexNet, GoogleNet, VGG16, VGG19, SqueezeNet, Inception, ResNet18 ve ResNet50 ile TSA mimarilerinden LSTM ve GRU kullanılmıştır. Anormal durumlardan silah ve bıçak sınıflandırılması işleminde en iyi başarımlı değeri ince ayar yaklaşımına dayalı eğitilmiş VGG16 modeli ile elde edilmiştir. Şiddet sınıflandırılması çalışmasında en iyi başarımlı değeri ise AlexNet modeli ile %100 doğruluk skoru elde edilmiştir.

Boyut indirgeme çalışması için, tezin ikinci aşamasında yapılmış olan sınıflandırma verileri ve temel görüntü işleme teknikleri ile video resim çerçeveleri farkları hesaplanarak özgün boyut indirgeme modeli üretilmiştir. Bu model, kullanıcıların rahatlıkla kullanabilmesi, kurum ve kuruluşların gereksinimlerine göre hassasiyetinin kontrol edilmesi ve sistem sorumlusunun teknik destek verebilmesi gibi ayrıntılar dikkate alınarak geliştirilmiştir. Bu sistem, insan bilgisayar etkileşimine uygun olarak Matlab AppDesigner platformu aracılığıyla tasarlanmıştır.

Bu tez çalışmasında, şiddet, silah ve bıçak tespitine ve sınıflandırılmasına dayalı deneysel çalışmalar yapılmıştır. Ayrıca, bu deneysel çalışmaların uygulanabilir bir hale getirilmesi ve görüntülerin depolamasında kazanç elde edilebilecek özgün bir boyut indirgeme sistemi geliştirilmiştir. Bu çalışmalar neticesinde video sistemlerinde anlık ve görüntü arşiv kayıtlarında analiz işlemlerinde kullanılmak üzere şiddet, silah ve bıçak tespiti ve sınıflandırma çalışmaları başarı ile gerçekleştirilmiştir. Görüntülerde anlık tespit çalışmaları, güvenlik kamera sistemlerine direk olarak bağlanıp kullanılabilir. Gerçekleştirilen arşiv görüntü analizi çalışmaları ise, bu çalışmaya özgü insan bilgisayar etkileşimleri kuralları gözetilerek hazırlanan görüntü analizi ve boyut indirgeme sistemine bağlanmıştır. Bu çalışmada, gerçekleştirilen özgün boyut indirgeme sistemi sayesinde hareketsiz ortamlarda % 10 ile %40 arasında kazanç sağladığı gözlemlenmiştir. Sonuç olarak, bu tez çalışmasında bütün amaç ve hedefler başarılı bir şekilde yerine getirilmiştir.

Aralık 2022, 128 sayfa.

Anahtar kelimeler: Derin Öğrenme, Görüntü İşleme, Bıçak Tespiti, Silah Tespiti, Şiddet Tespiti, Boyut İndirgeme, Arayüz Uygulama.

SUMMARY

Ph.D. THESIS

ANOMALY DETECTION AND SIZE REDUCTION SYSTEM WITH DEEP LEARNING APPROACHES IN VIDEOS

Mehmet Tevfik AĞDAŞ

İstanbul University

Institute of Graduate Studies in Sciences

Department of Informatics

Supervisor : Prof. Dr. Sevinç GÜLSEÇEN

In institutions and organizations, security camera systems are usually used to detect abnormal situations in videos. Security personnel resources are used in the management of security camera systems. Thanks to the attention and follow-up of the security personnel, an abnormal situation is detected in the video images on the security camera systems and the necessary interventions are made. Due to the fact that the source of security personnel is human, many human structures such as physiological, biological and spiritual affect the attention and monitoring they should have.

In addition to instant monitoring of security camera videos, it is also very important to store them. The videos must be recorded and stored in order for detailed analysis operations and legal proceedings to be carried out on the videos. The retrospective temporal increase in the storage of video recordings in security camera systems directly increases the usefulness, reliability of the system and the cost of the system, which is a disadvantageous situation.

This thesis study is planned in two stages. In the first stage, it is aimed to detect abnormal situations in videos with computerized imaging techniques based on deep learning and not affected by human structures. As a result of the work at this stage, an automatic weapon, knife and violence detection system was made from abnormal situations in instant video images. Since this study was conducted during the Covid-19 pandemic, the data sets were compiled from anonymous and copyright-free internet environments, YouTube images and academic study data sets with open source access in the model trainings conducted to enable abnormal

situation detection operations in videos. YOLOv4, YOLOv5, YOLOX and YOLOR models based on deep learning based image processing techniques were used in experimental studies. The highest performance value was achieved as 97.6% with the YOLOR model. For the detection of violence, EfficientDet D0-D7, YOLOv5 and Faster R-CNN models were applied and the best success rate was achieved with the EfficientDET D0-D7 model at 94.1%.

In the second phase of this thesis study, abnormal situation classification in security camera video recording images and size reduction in the storage of images were studied. The data sets for image classification were created from the compilation of anonymous copyright-free data sets, as in the detection and detection of abnormal situations in the first stage of the thesis. For these classification operations, AlexNet, GoogLeNet, VGG16, VGG19, SqueezeNet, Inception, ResNet18 and ResNet50, which are pre-trained deep CNN architectures based on deep learning, and LSTM and GRU from RNN architectures were used. The best performance value in the classification of weapons and knives from abnormal situations was obtained with the VGG16 model, which was trained based on a fine-tuning approach. The best performance value in the violence classification study was obtained with a 100% accuracy score using the AlexNet model.

For the size reduction study, an original size reduction model was produced by calculating the classification data and basic image processing techniques and video picture frame differences made in the second stage of the thesis. This model has been developed taking into account details such as the ability of users to use it easily, checking its sensitivity according to the requirements of institutions and organizations, and the ability of the system supervisor to provide technical support. This system is designed through the Matlab AppDesigner platform in accordance with human-computer interaction.

In this thesis study, experimental studies based on the detection and classification of violence, weapons and knives were carried out. In addition, an original size reduction system has been developed to make these experimental studies feasible and to make gains in the storage of images. As a result of these studies, violence detection, weapon and knife detection and classification studies have been successfully carried out for use in analysis operations in instant and image archive recordings in video systems. Instant detection studies in images can be connected directly to security camera systems and used. The archival image analysis studies carried out were connected to the image analysis and size reduction system prepared by observing the rules of human computer interactions specific to this study. In this study, it has been observed that it provides a gain between 10% and 40% in stationary environments thanks to the original Decimation system. As a result, all the goals and objectives have been successfully fulfilled in this thesis study.

December 2022, 128 pages.

Keywords: Deep Learning, Image Processing, Knife Detection, Weapon Detection, Violence Detection, Size Reduction, Interface Application.

1. GİRİŞ

Güvenlik insanın temel ihtiyaçlarından biridir. Ünlü psikolog Maslow, insanların temel ihtiyaçları piramidini tasarlarken insanların fiziksel ihtiyaçlarından daha üst basamakta güvenlik ve korunma ihtiyacı olduğunu belirtmiştir (Maslow, 1943). İnsanlar doğal olarak güvenilir ortamlarda hayatını devam ettirmek istemektedirler. Dünya nüfusunun sürekli artması ile birlikte yeni yaşam alanları oluşmuş ve özellikle şehirler her geçen gün daha kalabalık bir hal almıştır. Birçok yaşam alanında birçok canlı ortak yaşam sürmektedir. İnsanlarda kendi güvenliğini sağlayabilmek için öncelikli olarak güvenlik personeli kaynağına başvurmaktadır. Ortak yaşam alanlarında kalabalık insan gruplarının ve bazen diğer canlıların birlikte yaşaması personel kaynağının öneminin ve sayısının artmasına sebep olmaktadır. Personel kaynağının sayısının artması güvenlik ihtiyacının karşılanması konusunda rahatlama yaratsa da ekonomik olarak kayıplara yol açmaktadır. Ortak yaşam alanlarında güvenlik personeli kaynağına harcanan maliyetleri azaltmak için güvenlik kameraları kullanılmaktadır. Güvenlik kameraları sayesinde daha az personelle daha çok ortak yaşam alanı izlenebilmekte, depolama aygıtları ile de kayıt altına alınabilmektedir.

Standart güvenlik kamerası sistemlerinin denetimini insanlar yapmaktadır. Bu sistemlerin güvenirliliğini ve etkinliğini etkileyen en önemli unsur, sistemin başındaki insanın dikkati olmaktadır. Bu tür sistemlerde ister istemez güvenlik zafiyetleri olmaktadır. Ancak günümüz teknolojisinin gelişmesi ile bu tür sistemlerin zafiyetlerini ve insan yükünü azaltmak için sık sık yapay zekâya başvurulmakta ve daha modern güvenlik kamerası sistemleri oluşturulmaktadır. Modern yapay zekâ destekli sistemlerde denetimi hem bilgisayarlar hem de insanlar yapabilmektedir. Bu sayede insan gözünden kaçabilecek durumların tespiti bilgisayarlı bir görü ile yapılmaktadır.

Güvenlik kameraları sayesinde her olay kamera sistemlerinin depolama birimlerinde kayıt altına alınmaktadır. Bu kayıtlar sayesinde kameraların gözlemlemiş olduğu mekânlar tekrar izlenebilmekte ve herhangi bir olay gözden kaçmış olsa bile tekrar incelenebilmektedir. Bu özellik sayesinde olay delilleri oluşturulmaktadır. Güvenlik kameralarında geçmişe dönük depolama işlemi depolama biriminin kapasitesi kadar olmaktadır. Güvenlik kameraları sistemlerinde 1 Terrabyte'lık depolama birimi ve 4 adet kamera kullanılıyorsa 28 gün kayıt tutabilmektedir (Teknoloji, t.y.). Bu süre zarfı sonunda bu kayıtların alınıp farklı bir depolama

aygıtına aktarılması gerekmekte aksi durumda kayıtlar ya silinmekte ya da yeni kayıt işlemi yapılamamaktadır. Kayıtların silinmesi veya kayıt işleminin durması sonucunda da olası olay delillerinin erişimi düşeceğinden sistemin güvenilirliği ve kullanışlılığı da düşmektedir. Güvenlik kamerası sistemlerinde sürekli depolama birimi artırımı ihtiyacının olması maliyeti sürekli artıracığından, bu tür sistemlerin bu dezavantajlı bir durum olarak kabul edilmektedir.

Bu tez çalışmasında, modern güvenlik sistemlerine katkı sağlaması amacıyla makine öğrenmesi tabanlı derin öğrenme yaklaşımları ile bilgisayarlı görü kontrolünü arttırmak ve bu tür sistemlerin dezavantajı kabul edilen depolama kapasitelerinde maliyeti azaltacak bir yöntem önermek ve uygulamasını geliştirmek amaçlanmıştır.

Bu bölümde tez çalışmasının amacı ve önemi aşağıdaki başlıklar altında anlatılmıştır: “Tezin kapsamı”, “Tezin özgün değeri” anlaşılır ifade edilmesi için bu konu alt başlıklar halinde açıklanmıştır. “Tezin kapsamı”, başlığı altında bu tez çalışmasının neden yapıldığı ve hangi deneysel çalışmalarla oluşturulması gerektiği detaylı olarak açıklanmıştır. Tezin özgün değeri başlığı altında bu çalışmaya hangi çalışmaların ilham olduğu ve bu çalışmanın diğer çalışmalardan özgün taraflarından bahsedilmiştir. Tezin taslağı başlığı altında ise bu çalışmanın akış diyagramı oluşturulmuş ve detaylı olarak adımlar açıklanmıştır.

1.1. TEZİN KAPSAMI

Bu tez çalışmasında, güvenlik kameraları sistemleri veya bir başka deyişle kapalı devre kamera sistemlerinde görüntüleri sınıflandırılması ve bu sınıflandırma sonucuna göre görüntülerin kamera sistemlerinin depolama birimlerinde boyut indirgenmesi uygulaması yapılması hedeflenmiştir. Tez çalışmasının konusuna karar verilirken güvenlik kameraları örneği incelenmiştir. Yapılan araştırmalar sonucunda güvenlik kameraları sisteminde geri dönük kayıt tutulduğu ve geçmişe dönük vakaların incelendiği ve bu vakaların delil niteliği taşıdığı gözlemlenmiştir. Ancak güvenlik kameralarında depolama işleminin sınırsız olmadığı, bu işlemin kamera sisteminin depolama biriminin kapasitesi doğrultusunda olduğu gözlemlenmiştir. Ülkemizde yaygın kullanılan güvenlik kamerası sistemleri incelendiğinde yaygın olarak üreticilerin standart olarak sundukları depolama birimi kapasite boyutunun 500 gb ile 1 tb boyutlarında olduğu anlaşılmıştır. Ülkemizde ayrıca kamera sistemlerinin büyük

çoğunluğunun 4 kameralı sistemden oluştuğu bilgisine varılmıştır. Güvenlik kamerası sistemi veya kapalı devre kamera sistemlerinde standart özellikler göz önüne alınırsa 4 kameralı 500 gb olursa veya 4 kameralı 1 tb sistem olursa bu kameralarda kaç günlük geri dönük kayıt tutulabiliyor araştırması yapılmıştır. Güvenlik kamerası görüntü kayıtlarında önemli bir diğer faktör ise kayıt kalitesi yani saniye başına düşen görüntü sayısının (Fps) kaç olması gerektiğidir. Fps değeri olarak yaygın kullanılan değerinin 24 olduğu, yani bir saniyelik görüntüde 24 resim karesinin olduğu durumudur. Bu özelliklere göre depolama birimlerinin ne kadar süre depolama yapacağı tablo 1.1’de gösterilmiştir.

Tablo 1.1: Kayıt cihazlarının depolama birimi kapasitesine göre gün hesabı (Teknoloji, t.y.).

KAPASİTE	FPS	ÇÖZÜNÜRLÜK	SIKIŞTIRMA TEKNOLOJİSİ	KAMERA SAYISI	KAYIT GÜNÜ SAYISI
500 GB	24	720P	H265	4	12
1 TB	24	720P	H265	4	24
500 GB	11	720P	S+265	4	30
1 TB	18	720P	H265	4	30

Tablo 1’de yaygın olarak kullanılan kamera sistemlerinde kullanılan standart depolama birimleri ve bu birimlerin kamera sisteminin standart özellikleri ile çalıştığı zaman kaç günlük geriye dönük depolama yapabileceği gösterilmiştir. Tablo 1’de verilen değerler ülkemizde ve dünya genelinde yaygın olarak kamera sistemi üretimi yapan firmanın web sitesinden bugünün koşullarında alınmıştır (Teknoloji, t.y). Veriler incelendiğinde 4 kameralı, 24 Fps değeri ile kayıt yapan, görüntü çözünürlüğü kamera sisteminin minimum değeri olan 720 piksel ve sıkıştırma teknolojisi olarak en iyi sıkıştırmayı yapabilen h265 yazılımı kullanılırsa 500 gb depolama biriminde 12 günlük kayıt tutulmakta, 1tb depolama biriminde ise 24 günlük geriye dönük kayıt tutulmaktadır. Yaygın olarak kullanılan güvenlik kamera sistemlerinin standart depolama birimleri ile 30 günlük kayıt yapılmak istenirse 1 tb kapasiteli sistem için Fps değerinin 18’e düşürülmesi, 500 gb kapasiteli sistem için 11’e düşürülmesi gerekmektedir. 500 gb sistemin 30 günlük kayıt yapabilme performansına ulaşabilmesi için ayrıca sıkıştırma yazılımının da s+265 teknolojisine yükseltilmesi gerekmektedir. Güvenlik kamerası sistemlerinde boyut indirgenip günlük depolama sayısının artırılması işlemi çok kolay olamamaktadır. Çünkü güvenlik kamerası sistemlerinin kayıtlarının güvenilirliğinin zarar görmesi söz konusu olabilmektedir. Tablo 1.1 de gösterildiği gibi 30 günlük kayıt alınabilmesi ve görüntü Fps değerinin en düşük değerlere çekilmesi önemli sayılabilecek görüntülerde

atlamalarına neden olabilecektir. Güvenlik kamera sistemlerinde veya kapalı devre kamera sistemlerinde günlük depolama sayısının artırılmasının en güvenli yolu depolama birimlerinin sürekli olarak artırılmasıdır. Ancak bu tarz kamera sistemleri sürekli güvenlik sağlanması amacıyla kullanıldığı için depolama birimlerinin sürekli artırılması sistemin kullanıcıları açısından sürekli maliyet oluşturacaktır. Bu şekilde sistemin güvenilirliği sürekli sağlanmış olacak fakat sistemin artan maliyetleri nedeniyle sürekli erişilebilirliğini düşürecektir.

Bu tez çalışmasında gelişen teknolojik imkânlar kullanılarak güvenlik kamerası sistemlerinin veya kapalı devre kamera sistemlerinin hem güvenilirliğinin sağlanarak hem de erişilebilirlik konusunda sürekli zarar görülmesinin önüne geçilmesi için deneysel çalışmalar ışığında uygulama gerçekleştirilmesi hedeflenmiştir.

Günümüz bilgisayar teknolojisinde geline noktada bu tarz uygulamalar için en etkili yöntemlerin bir sonraki konu başlığı olan tezin özgün değeri başlığında anlatıldığı gibi literatür kaynakları ışında makine öğrenmesi yaklaşımları altında derin öğrenme ve yapay zeka tekniklerinin olduğu gözlemlenmiştir. Bu tez çalışmasında da derin öğrenme ve yapay zeka tekniklerinin kullanılarak sınıflandırma yöntemleri ile videonun özetlenmesi çalışması yapılmıştır. Özetlenen kamera görüntülerinin analizleri yapılarak yeni özgün bir yöntemle gereksiz Fps azaltarak değil verimli, güvenilir ve erişebilir güvenlik kameralarının günlük depolama sayılarının artırılması çalışması yapılmıştır. Kamera sistemlerinin depolama birimlerinin artırılmasında derin öğrenme, yapay zekâ teknikleri kullanılarak elde edilen sınıflandırmalar ışığında yeni özgün bir algoritma oluşturulmuş ve başarısı yüzlerce deneysel çalışmalarla ispatlanması yolu tercih edilmiştir.

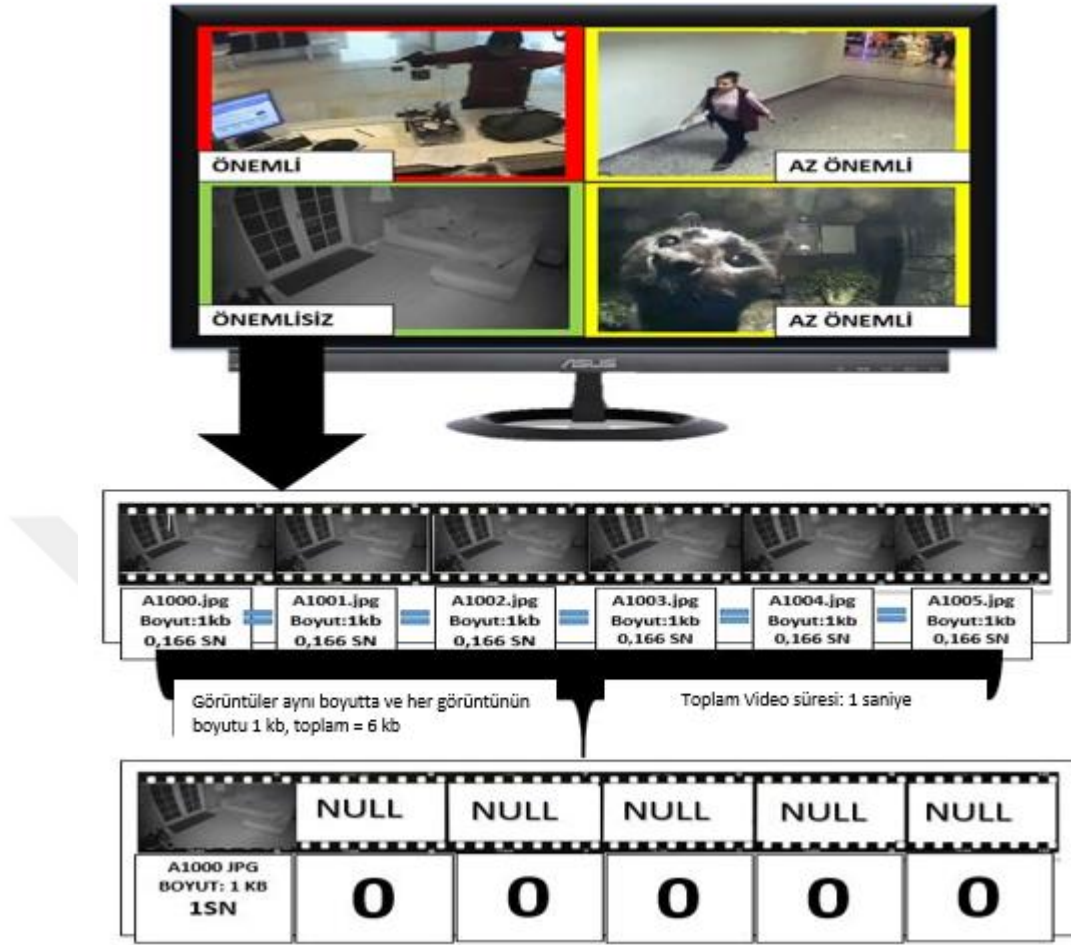
1.2 TEZİN ÖZGÜN DEĞERİ

Bu tez çalışmasının özgün değerinin belirlenmesi için literatürde bu ve benzeri konularda yapılmış olan bilimsel çalışmalar detaylı olarak incelenmiştir. Günümüzde gelişen teknolojik gelişmeler ışında bilimsel çalışmanın daha verimli olması açısından bu tarz çalışmaların literatür araştırmaları yapılırken günümüze yakın tarihli yayınlanmış çalışmalar olmasına dikkat edilmiştir. Görmüş ve Akar (2019) çalışmasında, Bergama bölgesindeki kentsel ve kırsal alanlarda VNIR kameralar ile kayıtlar alınmıştır. Bu kayıtların indirgenmesi için ilk olarak atmosferik ve radyometrik düzenlemeler yapılmıştır. Ardından indirgeme işlemi için bozuk ve kullanılmayan kayıt bantlarının silinmesi için 5 farklı yöntem kullanılmıştır.

Bu yöntemler Temel Bileşen Analizi (PCA), Ayrık Dalgacık Dönüşümü (DWT), Faktör Analizi, Çok Boyutlu Ölçekleme ve En fazla Daralan Metrik Öğrenme (MCML) yöntemini kullanılmıştır. Çalışmanın sonucunda en verimli yöntemin DWT yöntemi olduğunu savunulmuştur. Masala vd. (2014) çalışmasında, görüntülerin boyutlarının indirgenmesi için veri madenciliği algoritmaları ile boşlukların tespit edilmesini ve bu boşlukların videolar üzerinde az hasar oluşturabilecek şekilde tespit edilip temizlenmesini önerilmiştir. Çalışmada veri madenciliği algoritmalarından SVM algoritması ile en başarılı görüntülerdeki boşluk tespiti işlemi yapılmıştır. Başkurt (2019) tez çalışmasında, çok kameralı nesne takibi ve video sinopsis üzerine algoritma geliştirilmiştir. Çok kameralı sistemlerde, etkili bir şekilde görsel tarama ve ilgilenilen olayların veri setleri içerisinde çıkarılması büyük zaman almaktadır. Yazar çalışmasında bu konu üzerine algoritma geliştirmiştir. Algoritma video görüntüleri üzerinde ayrıntılı nesne takibi ve istenilen konu üzerinde başarılı bir sınıflandırma yapmıştır. Çalışmanın gerçekleştirilmesinde derin öğrenme yöntemleriyle sistem eğitilmiş ve SVM algoritması ile sınıflandırma yapılmıştır. Chakraborty vd. (2015) çalışmasında, video özetleme için uyarlanabilir ana kare seçimi üzerine yöntem geliştirilmiştir. Video özetleme işlemi için video içerisindeki önemsiz verilerin soyutlaştırılması işlemi yapılmıştır. Soyutlaştırma işlemi için anlık olmayan video kayıtlarından parametreler belirlenmiş ve sistem eğitilmiştir. Çalışmalar sonucunda video özetleme için bir model ortaya çıkarılmış ve bu modelin başarılı sonuçlar verdiği yazarlar tarafından gözlemlenmiştir. Çalışmanın video özetleme uygulamaları için kullanılmasının umut verici olduğu kanısına varılmıştır. Hussain vd. (2019) çalışmasında, CNN ve çift yönlü LSTM kullanılarak bulut destekli çoklu görüntü video özetleme işlemi yapılmıştır. Kurum ve kuruluşların gözetleme birimleri tarafından elde edilen video verilerinin video özetleme, indeksleme ve analiz için kullanılmasında çeşitli zorluklarla karşılaşmaktadır. Özellikle çoklu görüntü video özetleme (MVS) sistemlerinin devasa veri boyutları, fazlalık, görünümde üst üste binme, ışık değişimleri ve röportaj korelasyonları nedeniyle çalışmada zorluklar yaşanmaktadır. Yazarlar yaşanan zorluklardan üstesinden gelebilmek için MVS'den tam olarak yararlanılmayan çeşitli düşük seviyeli özellikler ve kümeleme tabanlı çalışan bilgi işlem tekniklerinin kullanılmasını önermişlerdir. Derin sinir ağı tabanlı yumuşak bilgi işlem teknikleri iki katmanlı bir çerçeveye bütünleştirilerek yeni bir MVS elde edilmiştir. İlk katman çevrimiçi katmandır. Çevrimiçi katmanda hedef görüntülere dayalı çekim segmentasyonu gerçekleştirilmiş ve bir buluta aktarılması işlemi yapılmıştır.

İkinci katmanda ise, bulut verisindeki her bir verinin ayıklanması ve derin özelliklerin çıkarılması işlemi yapılır ve derin çift yönlü uzun kısa süreli belleğe aktarılması (LSTM) ve video özetleme işlemi gerçekleştirilmiştir. Sonuç olarak sistemin YouTube karşılaştırılmalı veri kümeleri üzerindeki deneylerde en iyi MVS yöntemlerinden daha iyi performans gösterdiği gözlemlenmiştir Hassanien (2011) çalışmasında, makine öğrenmesi tabanlı bir olay tespit ve özetleme sistemi sunulmuştur. Önerilen sistem 6 aşamadan oluşturulmuştur. Bu aşamalar, ön işleme, çekim işleme, tekrar algılama, skor tablosu algılama, heyecan olayı tespit ve özetleme aşamasıdır. Çalışmada, makine öğrenmesi algoritmalarından SVM ve K-ortalamlar algoritmaları kullanılmıştır. Sonuç olarak futbol videolarında başarılı bir şekilde özet çıkarma işlemi gerçekleştirilmiştir.

Bu çalışmada boyut indirgeme işlemi için boşluk (NULL) atama yöntemi kullanılacaktır. Güvenlik kameralarındaki kayıtlarda önemsiz parametre olarak kabul edilecek tekrarlı hareketsiz veriler üzerinde işlem yapılacaktır. Bu işlemi gerçekleştirebilmek için görüntülerde analiz ve indirgeme işlemlerinin birlikte yürütülmesi gerekmektedir. Video kayıtlarında güvenlik açısından önemli ve önemsiz kabul edilecek parametreler gözlem yoluyla belirlenecek ve bu parametreler makine öğrenmesinin analiz algoritmaları ile algılanacak ya da tespit edilecektir. Bu tespit edilen veriler önemli haricindekiler önemsiz veri olarak kabul edilecektir. Bu önemsiz kabul edilen veriler üzerinde yine makine öğrenmesi algoritmaları kullanılarak tekrarlı veriler yani görüntüler üzerinde hiç değişiklik olmaması durumları tespit edilecek. Bu tespit edilen tekrarlı verilerin ilk görüntüsü alınacak tekrar eden diğer görüntülere null atanacaktır. Bu şekilde hiçbir görüntü kaybı yaşanmayacaktır. Video kayıt izleme aşamasında ise null atanan verilerin yerine null verilerin süresi kadar ilk görüntü gösterilecektir. Bu çalışma sonucunda hem önemli parametreler tespit edilip sistem uyarılacak hem de güvenli bir şekilde boyut indirgemesi işlemi yapılacaktır. Şekil 1.1'de bu tez çalışmasında hedeflenen boyut indirgemesi sisteminin şeması gösterilmiştir.



Şekil 1.1: Boyut indirgeme sistemi şeması.

Parametre seçimleri için güvenlik için önemli sayılacak sık görülen (silah, bıçak ve şiddet içeren olaylar) güvenlik olayları önemli parametre, parametreler dışında kalan ve sadece hareketli görüntüler az önemli ve hiçbir önem arz etmeyen hareketsiz görüntüler önemsiz parametre kabul edilecektir.

Önemli parametreleri içeren algılama/tespit çalışmaları; Grega vd. (2016), çalışmasında, güvenlik kameralarından tabanca ve bıçak algılamak için MPEG7 videolardan özellikli çıkarımı yapmış SVM algoritmaları ile sınıflandırma yapmıştır. Gonzalez vd. (2020), kapalı devre güvenlik kamerası (CCTV) sisteminde gerçek zamanlı silah tespiti çalışması yapmıştır. R-CNN tabanlı önceden eğitilmiş ResNet-50 mimarisi kullanmışlardır. Verma ve Dhillon (2017), derin öğrenme tekniklerinden Faster R-CNN kullanarak elde taşınabilir silah algılama çalışması yapmıştır. Dağınık sahnelerden otomatik silah algılaması için derin evrişimli ağı kullanmışlardır. Verma & Tiwari (2015), Haris ilgi noktası dedektörü kullanarak bilgisayarlı

görü tabanlı görsel silah algılama çalışması yapmıştır. K-Means algoritması kullanarak ilgisiz nesneyi görüntülerden çıkarma işlemi yapmışlardır. Görüntülerdeki renk tabanlı bölümlerden yararlanmışlardır. Kmiec vd. (2012), Görüntülerdeki bıçakları algılamak için Bilgisayarlı görüş tekniklerinden Aktif Görünüm Modellerinin yeni bir uygulamasını sunmuşlardır. Güvenlik kamerası görüntülerinde bıçak olup olmadığı sorusuna cevap aramışlardır. Castillo vd. (2019), derin öğrenme özellikli güvenlik kamerası gözetim videolarında otomatik soğuk çelik silah tespiti için parlaklık rehberliğinde ön işleme işlemi yapmışlardır. Algılama işlemi için ESA tabanlı ResNet mimarilerini kullanmışlardır. Ayrıca parlaklık kılavuzlu ön işleme için DaCoLT (Öğrenme ve Test aşamalarında Koyulaştırma ve Kontrast) tekniği kullanmışlardır. Fernandez-Carrobles vd. (2019), Video gözetimi için Faster R-CNN tabanlı silah ve bıçak algılama çalışması yapmışlardır. Çalışmalarında R-CNN tabanlı SqueezeNet mimarisi ve GoogleNet mimarisini kullanmışlardır. Algılama işlemi test ölçümlerinde GoogleNet mimarisi ile %46,68, SqueezeNet mimarisi ile %85.44 başarımlı skor elde etmişlerdir. Jain vd. Güvenlik uygulamaları için yapay zeka ve derin öğrenme teknikleri kullanarak silah algılanması çalışması yapmıştır. Çalışmalarında R-CNN sinir ağı SSD VGG-16 mimarisini kullanmış ve ortalama %73,8 başarımlı skor elde etmişlerdir. Olmos vd. (2018), derin öğrenme yaklaşımları kullanılarak videolarda otomatik tabanca algılanması yapmıştır. Açık erişim veri kümesi kullanmışlardır. Çalışmalarında ESA tabanlı R-CNN modeli ile eğitim işlemini yapmışlar. Kakadiva vd. (2019) çalışmasında, bankalarda akıllı güvenlik kameraları ile yapay zeka tabanlı otomatik soygun/hırsızlık algılama sistemi geliştirmişlerdir. Çalışmada derin öğrenme yöntemleri ve görüntü işleme algoritmaları kullanılmıştır. Sonuç olarak sensör kullanılmadan hırsızlık tespiti yapabilen sistem geliştirmişlerdir. Gao vd. (2016) çalışmasında, güvenlik kameralarından alınan veriler üzerinden algoritma geliştirerek şiddet algılama sistemi yapılmıştır. Çalışmada, videolarda pratik şiddet tespiti için istatistiksel hareket yönelimlerdeki hareket büyüklüğü değişiklik bilgisinden tam olarak yararlanan yönlendirilmiş violent flows (OVİF) adlı yeni bir özellik çıkarma yöntemi önerilmiştir. Deneyler, VİF+OVİF'nin ADABOOST'u SVM ile kullanırken en gelişmiş şiddet algılama performansını elde ettiğini göstermektedir. OVİF ve taban çizgisi yaklaşımlarının iki kamu veri tabanında karşılaştırılması, önerilen yöntemin etkinliğini göstermektedir. İkinci olarak, özellik kombinasyonu ve çok sınıflı birleşim stratejileri benimsenmiş ve başarılı sonuçlar elde edilmiştir. Zhang vd. (2016) çalışmasında, kalabalık ortamlardaki güvenlik kameralarından otomatik olarak şiddet algılanması için bir algoritma geliştirmiştir. Bunun için örnek kalabalık

alanlardan çok sayıda örneklem çıkarılmış ve sistem eğitilmiştir. SVM algoritmaları ile beslenen sistemleri çalışmanın yapıldığı dönemdeki en teknolojik cihazlardan bile başarılı sonuç verdiği çalışma yazarları tarafından belirtilmiştir. Ayrıca tablo 1.2’de ek literatür araştırmaları gösterilmiştir.

Tablo 1.2: Çalışmada parametrelerinin seçiminde incelenen çalışmalar.

KONU	REFERANS	DETAY
Şiddet Algılama	Pexioto vd. (2020),	Videolarda çok modlu şiddet algılama (Peixoto vd., 2020).
	Parez vd. (2019),	Gözetim videolarında gerçek dünya kavgalarının tespiti (Perez vd., 2019).
	Jiang vd. (2020),	Şiddet tespiti için 3D konvolüsyon ağları (Jiang vd., 2019).
Şüpheli Durumlar	Chuang vd. (2013),	Şüpheli nesne algılama ve soygun olay analizi (Chuang vd., 2007).
	Kakadiva vd. (2019),	Bankalarda akıllı gözetim kullanılarak AI tabanlı otomatik soygun/hırsızlık algılama (Kakadiya vd., 2019).
	Joshi vd. (2016),	Şüpheli nesne algılama (Joshi vd., 2016).
	Rakumpthong vd. (2014),	Mevcut nesnenin yerinin değiştirilmesine dayalı katılımsız ve çalıntı nesne

		algılama (Rakumthong vd., 2014).
	Kamthe ve Patil (2019),	Video gözetim sistemlerinde şüpheli etkinlik tanıma (Kamthe & Patil, 2018).
Şüpheli Durumlar	Lai vd. (2010),	Nesne hareketinin algılanması ve sınıflandırılması- video gözetim sistemi için bir uygulama (Kong Chien Lai vd., 2010).
	Çetin vd. (2013),	Videolarda yangın algılama (Çetin vd., 2013).
	Liu ve Ahuja (2004),	Görüş tabanlı yangın algılama (Che-Bin Liu & Ahuja, 2004).
Yangın Algılama	Marbach vd. (2006),	Video görüntülerinde yangın tespiti için bir görüntü işleme tekniği (Marbach vd., 2006).
	Li ve Zhao (2020),	Evrişimli sinir ağlarına dayalı görüntü yangın algılama algoritmaları (Li & Zhao, 2020).

Tablo 1.2 'de bu çalışmada belirlenen sınıflandırma için kullanılan parametrelerin literatürde başarılı sonuçlar almış örnekleri gösterilmiştir. Bu parametrelerin sınıflandırılması için kullanılan yöntemler detaylı olarak incelenmiştir.

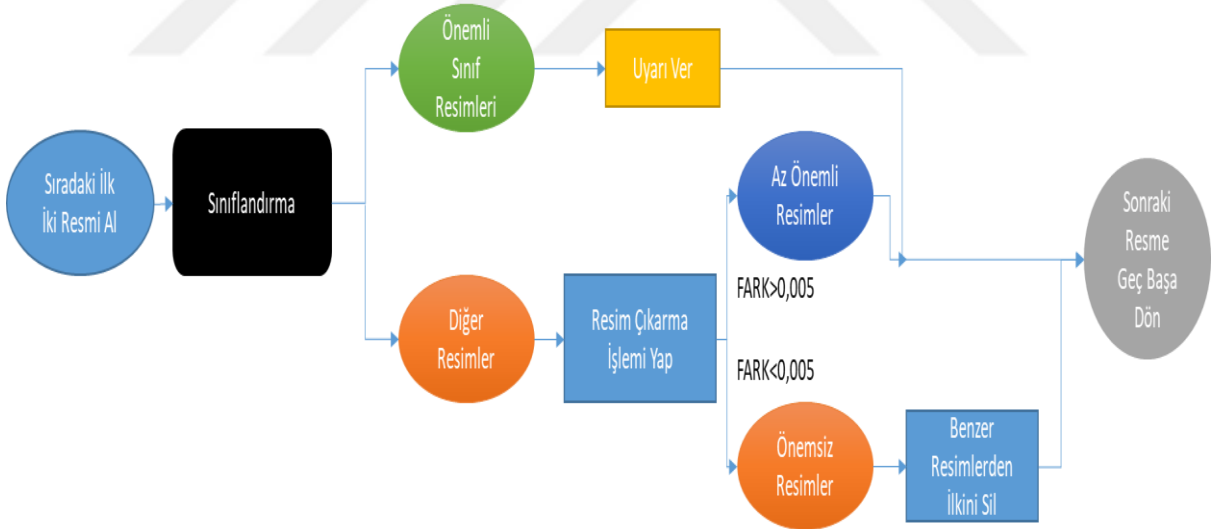
Bu çalışmanın özgün değeri, daha önceki çalışmalardan farklı olarak boyut indirgeme süreçlerinde derin öğrenme tabanlı görüntü sınıflandırması tekniğinin kullanılmasıdır. Sınıflandırma sonucuna göre güvenlik kamera sisteminin veya kapalı devre kamera sistemlerinin güvenliğini tehlikeye düşürmeyecek şekilde boyut indirgenmesi işlemi yapılmıştır. Çalışmada görüntüler önemli, az önemli ve önemsiz sınıflara ayrılmıştır. Önemli sınıf içerisinde görüntülerin analizleri ile elde edilmiş silah, bıçak, şiddet ve yangın görüntülerinin özetlenmesi işlemi yapılmıştır. Az önemli sınıfı hareketli önemli sınıfın haricindeki her durum oluşturmaktadır. Önemsiz sınıfı ise hiçbir hareketin olmadığı durumlar oluşturmaktadır. Bu çalışmanın tezi güvenilir ve erişilebilirliğin daha az zarar gördüğü güvenlik kamerası kullanılabilirliğini artırılması için uygulama geliştirmektir. Bu sebeple güvenlik kamerasında günlük depolama artırılırken bütün Fps'lerin azaltılmasının yerine daha kazançlı ve güvenli yol olan önemsiz videolardaki tekrarlı görüntülerin içerisindeki tekrarların silinmesi yöntemi kullanılmıştır. Bu tez çalışmasının benzerine literatür kaynaklarında rastlanılmamıştır.

1.3 TEZİN TASLAĞI

Bu tez çalışması ilk olarak taslak olarak hazırlanmış ve taslağın her adımı süreç içerisinde deneysel test edilmiş güncellenmiş ve son halini almıştır. Bu çalışmanın taslağının son haline göre ilk adım olarak literatür kaynakları detaylı olarak incelenmiş ve karşılaştırılmıştır. Yöntem olarak makine öğrenmesi tabanlı derin öğrenme ve yapay zekâ teknikleri kullanılarak sınıflandırılma işlemine tabi tutulmuştur. Sınıflandırma işleminde parametreler seçilmeden önce ana sınıflar oluşturulmuştur. Güvenlik kameralarının boyut indirgenmesi işlemi için kullanılacak ana sınıflar önemli sınıf, az önemli sınıf ve önemsiz sınıf şeklinde belirlenmiştir. Ana sınıfların iç sınıflandırılması işlemi için parametre belirlenmesi ihtiyacı doğmuştur. Parametrelerin belirlenmesinde birçok güvenlik şirketi ile temasa geçilmiş sözlü olarak hangi parametrelerin tespitinin önemli olacağı konusunda fikir alışverişi yapılmıştır. Fikir alışverişleri sonucunda en verimli anasınıfın alt parametreleri şu şekilde belirlenmiştir.

1. Önemli sınıf: Silah algılama, bıçak algılama ve şiddet algılama şeklinde belirlenmiştir.
2. Az önemli sınıf: Önemli sınıfın haricindeki her türlü hareketli durumları kapsamakta,
3. Önemsiz sınıf: Hiçbir hareketin olmadığı ve gözle görüntünün değişip değişmediğinin tespit edilemediği durumları kapsamaktadır.

Tez çalışmasının taslağının bir sonraki adımında derin öğrenme ve yapay zekâ teknikleri kullanılarak güvenlik kamerası sistemleri veya kapalı devre kamera sistemlerinin kayıtlarının alınması ve bir veri kümesi oluşturulması işlemi yapılmıştır. Bu tez çalışmasına Covid-19 pandemisi döneminde başlanıldığı ve süreç içerisinde çalışmaları zorunlu olarak gerçekleştirilmesi gerektiğinden güvenlik kamerası görüntüleri internet ortamından paylaşan kişilerin izinleri doğrultusunda temin edilip kullanılmıştır. Elde edilen görüntülerden ham veri kümesi oluşturulmuştur. Bu görüntüler ön işleme teknikleri ile güvenlik kameralarından elde edilmiş ve halen depolama birimlerinde kullanılıyormuş gibi hazırlanmıştır. Veri kümesindeki video görüntüleri, Matlab programı sayesinde resim çerçevesi haline getirilmiş ve bu görüntüler sınıflandırma işlemine tabi tutulmuştur. Sınıflandırma işleminde derin öğrenme tekniklerinin birçoğu denenmiş ve daha önceleri bu tarz çalışmalardan daha başarılı sonuçlar alınması hedeflenmiş ve gerçekleştirilmiştir. Elde edilen sınıflandırma parametreleri ana sınıflara dâhil edilmiştir. Sınıflandırma işlemlerinin tamamlanması ile boyut indirgenme işleminin gerçekleştirilmesi için bir boyut indirgenmesi algoritması oluşturulmuştur. Boyut indirgenme algoritmasının kısa taslağı şekil 1.2’de gösterilmiştir.



Şekil 1.2: Boyut indirgeme algoritması kısa taslağı.

Güvenlik kamerası sistemlerinin video görüntüleri ilk olarak resimlere ayrılması işleminde resimler 1 rakamından başlanılarak ardışık olarak adlandırılacaktır. Şekil 1.2’de gösterilen kısa taslak görüntüsünde ilk “*sıradaki ilk resmi al*” adımı sıralı olan görüntülerden 1 ve 2 adında olan resimler alınacaktır. Bir sonraki “*sınıflandırma*” adımı ise alınan bu iki görüntü

üzerinde sınıflandırma algoritmaları tekniği uygulanacaktır. Sınıflandırma işlemi sonunda görüntülerde herhangi önemli sınıf parametresi tespit edilirse görüntüler “**önemli sınıf resimleri**” ana sınıfına dâhil edilecek ve kullanıcıya parametre hakkında bilgi vermesi için anlık “**uyarı ver**” adımına geçecektir. Uyarı işleminden sonra sistem resimler üzerinde hiçbir değişiklik yapmayıp resim adlarının rakamsal değerleri bir artırılıp tekrar işleme başlanacaktır. Sınıflandırma işleminde eğer görüntülerde herhangi parametre tespit edilmez ise resimler “**diğer resimler**” geçici sınıfına dâhil edilecektir. Bu adımdan sonra ardışık olan resimler üzerinde filtreleme işlemleri yapılarak resimlerin arasında fark olup olmadığı kontrol edilecektir. Bu elde edilen fark 0,005 değerinden küçük ise adları rakam olarak verilen resimlerden adının rakamsal değeri küçük olan resim benzer “**resimlerden ilkini sil**” adımında silinecektir. Bu adımdan sonra yine resimlerin adlarının rakam değeri 1 artırılıp yeni adda olan resimlerle sistem baştan başlayacaktır. Ancak fark 0,005 değerinden küçük değilse o zaman görüntüler üzerinde hiçbir değişiklik yapılmayıp görüntüler “**az önemli sınıf**” adımında az önemli sınıfa dâhil edilecektir. Buradaki görüntülerde yine adlarının rakam değerleri 1 artırılarak yeni resimler alınıp sistem tekrar başlatılacaktır. Bu adımlar arka planda çalışacak ve hiçbir şekilde kullanıcının sistemi kullanmasına engel olmayacaktır. Sistem kullanıcının gün sonunda sistemin tasarlanmış ara yüzünde boyut azalt sekmesini seçmesiyle çalışacaktır. Bu şekilde sistem gece yoğunluğun az olduğu saatte çalışacak ve bilgisayarın performansı düşmesi gibi sorunlarda yaşanmayacaktır.

2. GENEL KISIMLAR

Bu tez çalışmasının gerçekleşmesi için tezin taslağı başlığında anlatıldığı gibi birçok işlemin gerçekleştirilmesi gerekmektedir. Bu işlemlerin gerçekleştirilebilmesi için birçok bilgi birikimine sahip olunması gerekmektedir. Çalışmada kullanılan bilgi ve birikimlerinden ilki, bu tarz çalışmaların temelini oluşturan makine öğrenmesi konu başlığının çok iyi şekilde öğrenilmesidir. Makine öğrenmesi konusu, makine öğrenmesi başlığı altında açıklanmıştır. Çalışmada sınıflandırma, algılama ve tespit işlemlerinde derin öğrenme teknikleri kullanılmıştır. Derin öğrenme konuları alt konuları başlıklar halinde açıklanmıştır. Tez çalışmasının ikinci adımı olan boyut indirgeme işlemleri için ise algoritma geliştirime, temel görüntü işleme ve filtreleme teknikleri kullanılmıştır. Görüntü işleme ve görüntü filtreleme konusu resim boyut indirgeme başlığı altında detaylı açıklanmıştır.

2.1 MAKİNE ÖĞRENMESİ

Son yıllarda birçok bilimsel araştırma için çalışma konusu olan makine öğrenmesi, verilerden öğrenerek modelleme sağlayabilen bir yapay zeka alanıdır. Kavramsal olarak makine öğrenimi, ilk olarak 1959 yılında Arthur Samuel tarafından literatüre kazandırılmıştır (Samuels, 1959). Makine öğrenimi yöntemleri kullanılarak, sınıflandırma, anlama, veri kümeleri arasındaki ilişki hesaplanması, kümeleme ve görüntü işleme gibi birçok bilimsel araştırma yapılabilmektedir.

Makine öğrenmesi, günümüzde sağlık, ekonomi, meteoroloji, müzik, sinema teknolojisi ve güvenlik gibi alanlarda sıklıkla kullanılmaktadır.

Bu çalışmada da makine öğrenmesiyle tekniklerinden, verilerden derin öğrenme, yöntemlerden, görüntü işleme ve çalışma alanlarından, güvenlik alanı üzerine araştırma ve uygulama gerçekleştirilmesi hedeflenmiştir.

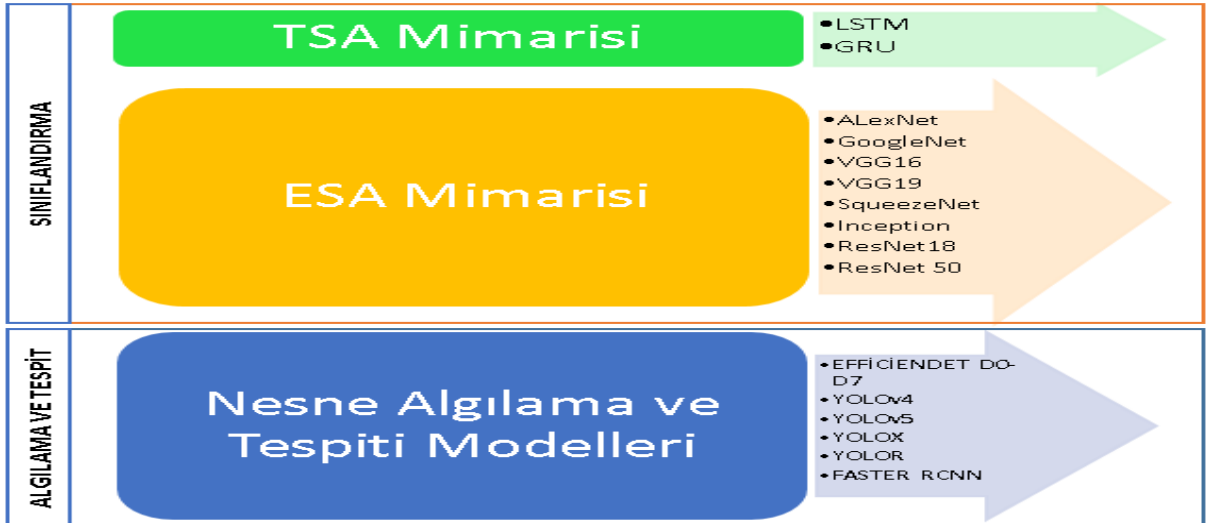
2.2 DERİN ÖĞRENME

Öğrenme becerisine sahip olan beynin işlevsel yapılarından esinlenerek tasarlanmış ve çok katmanlı yapılara sahip olan bir makine öğrenmesi algoritmasıdır. Derin öğrenme, makine öğrenmesinin kapsamı içerisinde ve yapay sinir ağlarından (YSA) daha karmaşık bir yapıya sahiptir. İnsandaki gibi öğrenme işlevine benzer şekilde geliştirilen YSA, insanın biyolojik

yapısında bulunan nöronları taklit ederek bu nöronlar arasındaki ilişkilere ve bu nöronların bu nöronların birbirleri ile olan bağlantılarına dayanmaktadır (Türkoğlu, 2019).

Derin öğrenme, verilerin birden fazla öğrenme seviyesine ve temsillerinin öğrenmesine dayalı bir sistemdir. Verilerin öğrenilecek üst düzey özellikleri ve alt düzey özellikleri üretilerek hiyerarşik bir temsili oluşturulmaktadır. Bu oluşturulan temsil, somutlaştırmanın farklı seviyelerine karşılık bir temsil seviyesinin öğrenilmesini sağlamaktadır (Bengio, 2009). Derin öğrenme teknikleri genel bir ifadeyle verilerin özellik temsillerinden öğrenmeye dayalıdır. Bir görüntü örneğinin temsili ifadesi; piksel başına yoğunluk değerlerinin vektörleri veya kenar kümeleri, özel şekiller gibi özelliklerin düşünülmesi anlamına gelmektedir. Verilerden çıkarılan özellik temsillerinin bazıları veriyi daha iyi temsil etmektedir. Bu aşamada yine bir avantaj olarak, derin öğrenme yöntemleri, elle çıkarılan özellikler (handcrafted features) yerine veriyi en etkili temsil eden hiyerarşik özellik çıkarım algoritmaları kullanılmaktadır (H. A. Song & Lee, 2013).

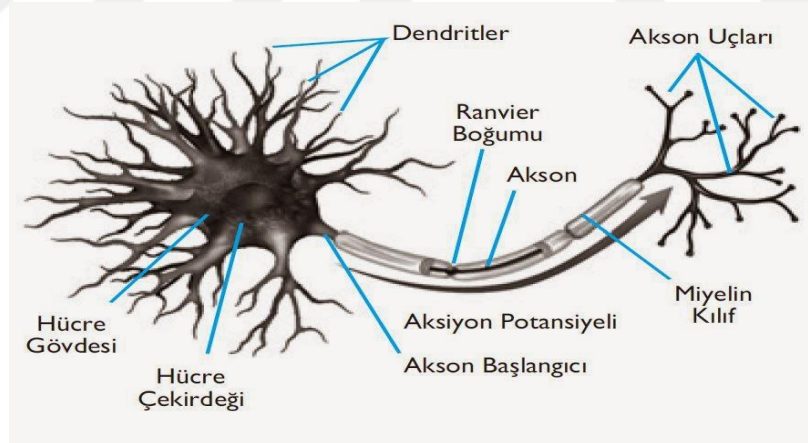
Bu çalışmada, görüntü sınıflandırılması, nesne algılanması ve tespiti işlemleri yapılmıştır. Görüntü sınıflandırılması işlemlerinde yapay sinir ağlarının sınıflandırma modelleri ve evrişimli sinir ağları modelleri kullanılmıştır. Görüntülerde nesne algılanması ve tespiti işlemlerinde ise, derin öğrenme tabanlı nesne algılama ve tespiti modelleri kullanılmıştır. Şekil 2.1’de bu çalışmada kullanılan derin öğrenme tabanlı modellerin şeması gösterilmiştir.



Şekil 2.1: Tez çalışmasının derin öğrenme tabanlı modelleri.

2.2.1. Yapay Sinir Ağları

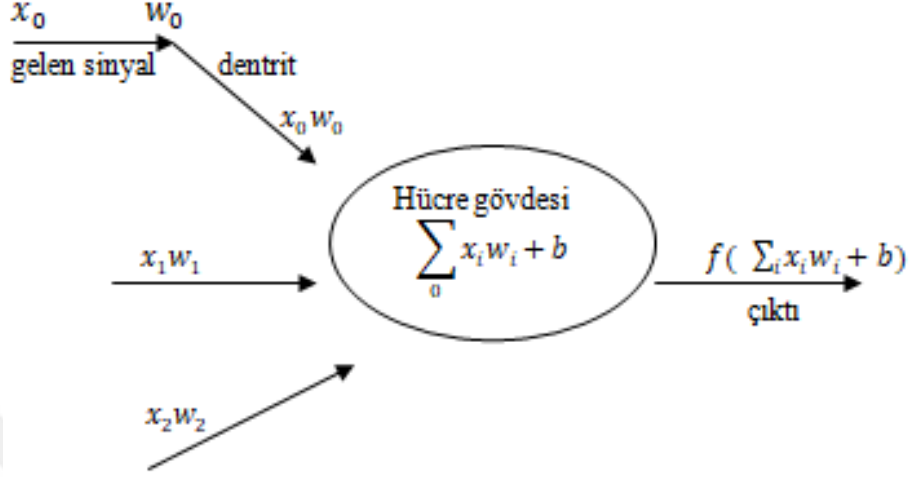
İnsan beyninin doğayı öğrenmesinin ve doğayı taklit eden komutlar üretmesinin en önemli çalışmalarından olan Yapay Sinir Ağları (YSA) hem bir teknolojik yaklaşım hem de; biyolojik sinir sistemlerini simüle edebilen programlar olarak tanımlanmaktadır. Bu sayede insan gibi düşünme becerisi olan, hafıza depolaması yapabilen ve deneyimlerinden çıkarımlar yapabilen bilgisayar programları üretilmesi amaçlanmıştır. Örneğin insanlar geçmiş deneyimleri sayesinde çevresel şartlara göre hava durumu tahmini yapabilir. Veya döviz kurlarını ülkenin ve dünyanın şartlarını ve birçok parametreyi göz önüne alarak yorumlar yapabilir. Hava durumunun ve döviz kurunun tahmin edilmesi örneğinde olduğu gibi geçmiş deneyimler ve birçok parametrenin insan beyninde işlenmesi yerine daha hızlı ve güvenilir sonuçlar alabilmek için YSA alt yapısını kullanan programlar tasarlanmıştır. YSA alt yapısıyla çalışan programlar sayesinde sadece bir ortamın hava durumu yerine ve bir döviz kuru yerine birçok döviz kurunun anlık tahminleri hızlı bir şekilde yapılabilmektedir. Bu programların çalışma mantığını anlayabilmek için bu programların esin kaynağı olan insanlara ait biyolojik sinir yapısının incelenmesi gerekmektedir (Salvi & Kadam, 2021; Yakıt & Özkan, 2017).



Şekil 2.2: Sinir hücrelerinin biyolojik yapısı (Yakıt & Özkan, 2017).

İnsanlara ait biyolojik sinir hücrelerinin yapısı şekil 2.2’de gösterilmiştir. Şekil 2.2’de anlaşılabileceği üzere sinir yapısı üç temel yapıdan oluşmaktadır. Bunlar dendritler, hücre gövdesi ve aksonlardır. Bilgiler, dendritlerden hücre gövdesine ve hücre gövdesinden aksonlara iletilmektedir. Bir sinir hücresinin aksonu ile diğer bir sinir hücresinin dendriti arasında boşluk bulunmaktadır. Öğrenme aşamasında bu boşluk kimyasal bir sıvı ile dolmakta ve bilgilerin

iletilmesi sağlanmaktadır. YSA da ise bu işlem basamaklarındaki görevlerin yapılmasını sağlayan katmanlar bulunmaktadır.

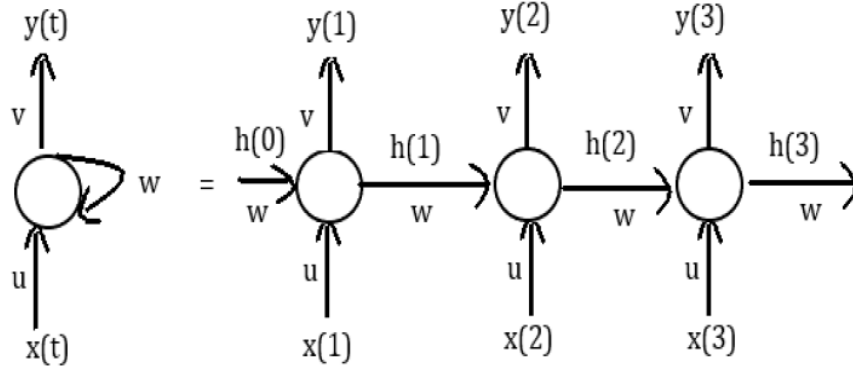


Şekil 2.3: YSA hücresinin matematiksel yapısı.

YSA hücresinin matematiksel yapısı şekil 2.3'te gösterilmiştir. Şekilden anlaşılacağı üzere insan sinir hücresinin yapısına benzer bir yapı vardır. Dendritler üzerinden gelen bilgiler hücre gövdesine ve hücre gövdesinde işlendikten sonra çıktı olarak başka nörona aktarılmaktadır. Burada x : gelen bilgiyi, w : ağırlık anlamına gelen weight ve b : bias değerini göstermektedir. Dendritlerden gelen bilgi hücre gövdesinde $\sum_a x_i w_i + b$ işlemine tabi tutmakta ve yeni veri oluşturmaktadır. Yapay nöronlar aldığı bu veriyi $f(\sum_i x_i w_i + b)$ aktivasyon fonksiyonundan geçirerek diğer nöronlara iletmektedir. YSA x değeri en başta verilmekte, model elde ettiği hatalara göre yeni w ve b değeri üretmektedir (Yıldız vd., 2021).

2.2.1.1 Tekrarlayan Sinir Ağları (Recurrent Neural Network (TSA-RNN))

Tekrarlayan sinir ağları, yapay sinir ağlarının bir türüdür. TSA çıktı verileri sinir ağının ardışık, zaman serisi girdilerinin sonucuna bağlıdır. TSA diğer ileri beslemeli ağlardan farklı olarak, girdi verilerini işleme alırken ağın kendi hafızasından faydalanmaktadır. TSA, eylem, ses ve metin sınıflandırmalarında zaman serileriyle çalışmaktadır (Mikolov vd., 2010). Şekil 2.4'te TSA açılımı gösterilmektedir.



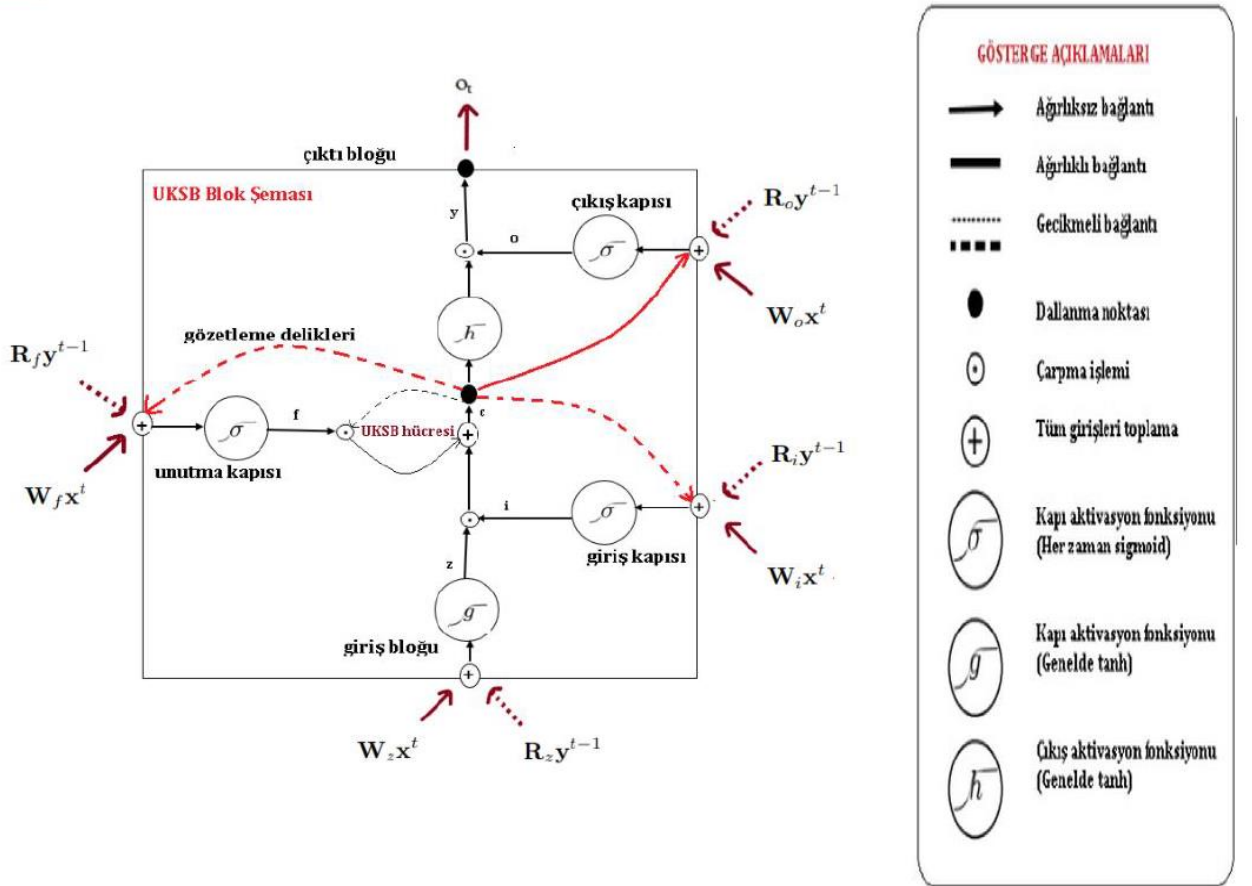
Şekil 2.4: TSA açılımı.

Şekil 2.4’de gösterildiği gibi ağın eğitilmesinde diğer sinir ağlarında olduğu gibi geri yayılım (Backpropagation) işleminin yapılması gerekmektedir. Ancak TSA’da önemli bir sorun meydana gelmektedir. Gradyan, tüm ağırlıkların ayarlamasının yapıldığı değeri temsil etmektedir. Tekrarlayan ağlarda gradyan değeri, sadece bulunduğu katmana değil bir önceki katmana da bağlıdır. Gradyan değeri 1’den küçük olması durumunda, zaman aralıklarında geri yayılım sürekli tekrarlanması işleminde sonuç sürekli olarak küçülmekte, bu nedenle azalan ve kaybolan gradyan değeri sorunu çıkmaktadır. Gradyan değerinin 1’den büyük olması durumunda ise, benzer işlemlerde sonuç sürekli büyümekte ve gradyan patlaması meydana gelmektedir (Metin & Karasulu, 2019).

2.2.1.1.1 Uzun-Kısa Süreli Bellek (UKSB- LSTM)

UKSB, 1997 yılında Juergen Schmidhuber tarafından gradyan sorunlarına çözüm olarak önerilen bir TSA türüdür (Hochreiter & Schmidhuber, 1997). Zaman serilerinde özel iç kapılara (unutma (forget gate)) sahip olması sebebiyle, normal TSA türlerinden daha tutarlı ve daha iyi performans göstermektedir. UKSB, ESA mimarileri, Kısıtlı Boltzman Makineleri (KBM-RBM) ve otomatik kodlayıcı (Auto Encoder) gibi ağlarla karşılaştırıldığında, UKSB’nin yapısı; ardışık eylemler, hava durumu, ses ve metin işleme gibi zaman serisi içeren işlemlerde daha başarılı sonuç verdiği gözlemlenmiştir. Bu ağ hücre mimarisinde, gradyanlar zamansal yayılımlarında zarar görmemektedir. UKSB girdileri uzun süreli bellekte tutmaktadır. UKSB ve TSA’nın mimarileri genel olarak benzerdir; ancak gizli durumları hesaplamak için farklı fonksiyonlar kullanılmaktadır. UKSB hücrelerinin hafızasında bulunan önceki durum verileri ile o anki durum verileri bir sonraki durum için girdi olarak aktarılmaktadır. UKSB hücreleri, hafızada hangi verilerin tutulup veya silineceğine karar verirler. Bu işlemlerin sonunda, daha

önceki durum verileri ile mevcut durum verilerini birleştirirler. UKSB sisteminde hücrelerin uzun süreli bağımlılıkları öğrenmede oldukça önemlidir. Şekil 2.5'te UKSB hücrelerinin blok yapısı gösterilmiştir. Yapıda W terimi ağırlıkları, R terimi ise tekrarlayıcı ağırlıkları göstermektedir. x^t terimi, girdi verilerinin t zamanına göre çekimini göstermektedir. z terimi girdi bloğunu, i girdi kapısını, f unutmaya kapısını, c terimi hücreyi, o terimi çıktı kapısını ve y terimi ise çıktı bloğunu temsil etmektedir (Metin & Karasulu, 2019).

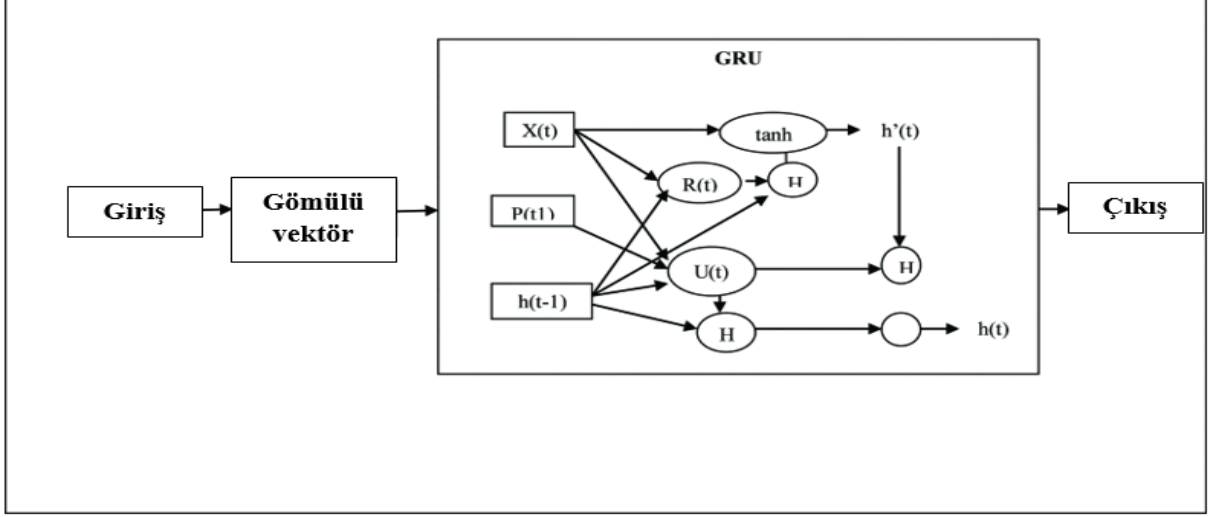


Şekil 2.5: UKSB yapısı blok şeması (Greff vd., 2017).

2.2.1.1.2 Kapılı Tekrarlayan Birimler (KTP-GRU)

LSTM ağları, 1997 yılında Hochriter ve Schmidhuber tarafından dil işleme çalışmaları için uzun ve kısa süreli ezberleme yeteneğine sahip model olarak önerilmiştir (Hochreiter & Schmidhuber, 1997). Ancak LSTM mimarilerinin karmaşık yapısı nedeniyle dil işleme eğitim süreçleri uzun almaktadır. Bu problemin aşılması için Cho ve diğerleri, 2014 yılında LSTM

modeline göre eğitim süresi daha az zaman alan ve daha basit yapıya sahip olan LSTM ağlarının bir modifikasyonu olan GRU modelini önermişlerdir (Cho vd., 2014).



Şekil 2.6: GRU modeli hücre yapısı (Pavithra vd., 2019).

GRU modelinin hücre yapısı şekil 2.6'da gösterilmiştir. Bu modelde gizli durum parametresi h_t ve hücre durumu parametresi c_t birleştirilerek bir giriş olarak ve bir diğer gizli durum parametresi x_t ayrı giriş olarak kullanılmaktadır. GRU modelinde LSTM modelinden farklı olarak iki kontrol kapısı kullanılmaktadır. Bu kapılar güncelleme kapısı (h_t) ve sıfırlama kapısıdır (x_t). Önceki kapılara müdahale geçmişlerini (p_t) almaktadır. Güncelleme kapısı (h_t), bir önceki adımda ne kadar durum bilgisi elde edildiğini kontrol etmek için kullanılmaktadır. Güncelleme kapısının değeri ne kadar fazla olursa bir önceki adımda o kadar çok durum bilgisi getirilmektedir. Sıfırlama kapısı (h_t), mevcut aday kümesine bir önceki adımdan ne kadar bilgi yazıldığı bilgisini kontrol etmektedir. Sıfırlama kapısının değeri ne kadar küçük ise bir önceki adımdan az bilgi yazıldığı bilgisi elde edilmektedir (Gao vd., 2020).

GRU modelinin hücre yapısının denklemleri aşağıda gösterildiği gibi hesaplanmaktadır.

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (2.1)$$

$$i_t = \sigma(W_{xi}x_i + W_{hi}h_{t-1} + b_i) \quad (2.2)$$

$$\hat{c}_t = \tanh(W_{xc}x_i + W_{hc}h_{t-1} + b_c) \quad (2.3)$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \hat{c}_t \quad (2.4)$$

$$o_t = \sigma(W_{xo}x_i + W_{ho}h_{t-1} + b_o) \quad (2.5)$$

$$h_t = o_t \tanh \otimes c_t \quad (2.6)$$

Denklemlerde kullanılan σ lojistik sigmoid fonksiyonudur. $\otimes$ iki vektörün eleman bazlı çarpımıdır. W ağırlık matrislerini, b önyargı vektörlerini, f unutma kapısı, i giriş kapısı ve o çıkış kapısının aktivasyon değerlerini göstermektedir.

2.2.2. Evrişimsel Sinir Ağları (ESA-CNN)

Evrişimsel Sinir Ağları (ESA) görüntüler üzerinde nesnelerin sınıflandırılmasını, tespitini ve takibini yapabilen algoritmalar olarak tanımlanmaktadır. YSA dan farklı görüntüleri giriş vektörü olarak kullanmasıdır. Bunun yanı sıra birden fazla katman kullanmakta ve sınıflandırma yeteneğine sahiptir (Türkoğlu & Hanbay, 2019). ESA mimarisi kullanılarak oluşturulan modeller birçok önemli avantajlar sunmaktadır. Bunlar;

- Görüntülerde, bölütleme işlemleri ve hedef görüntü ile arka plan görüntünün ayrılması gibi işlemlerde ön-işleme adımlarına ihtiyaç duyulmaksızın ileri düzey görevleri yerine getirebilmektedir.
- Birçok görüntü işleme ve algılama modellerine göre kendi yapısında bulunan katmanlı yapılar sayesinde daha iyi performans göstermektedir.
- ESA mimarilerinde görüntülerde tek bir nesne yerine birçok nesnenin algılanması işlemleri yapılabilmektedir.

ESA alt yapısını oluşturan katmanlar ve ESA temelli mimariler alt başlıklar halinde açıklanmıştır.

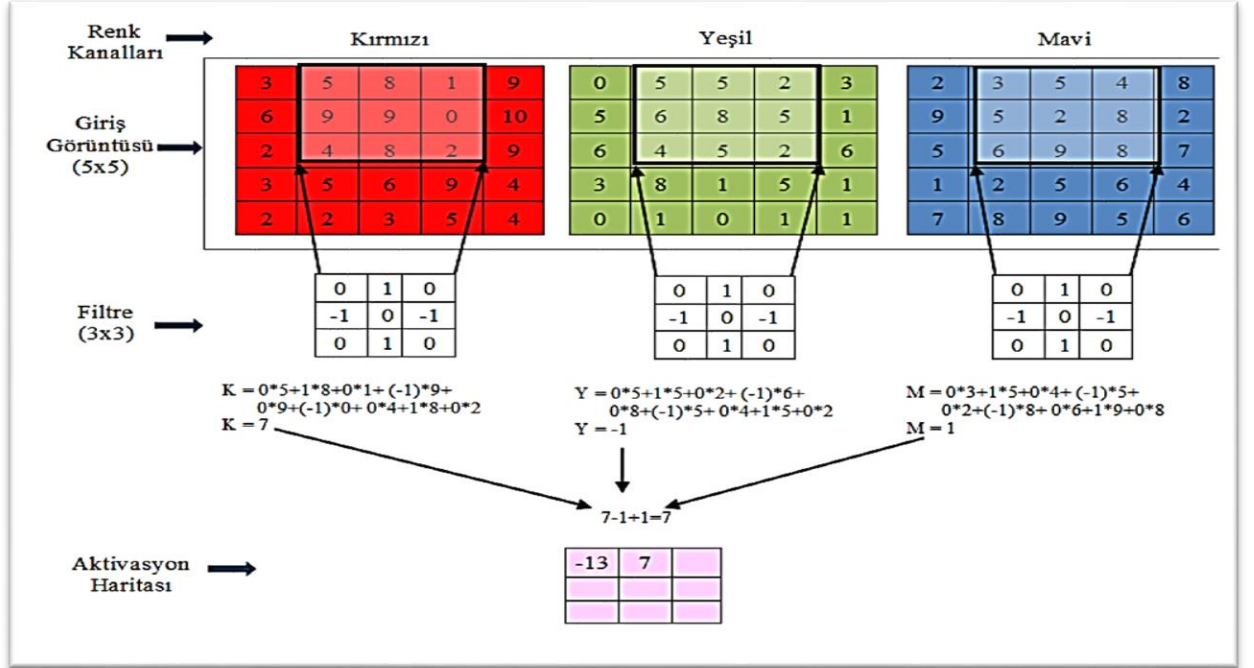
2.2.2.1.1 Giriş Katmanı

ESA modellerinin ilk katmanını giriş katmanı oluşturmaktadır. Giriş katmanında veri herhangi bir ön-işleme adımına maruz kalmaksızın ham olarak ağı iletilmektedir. Bu katmana iletilecek görüntünün boyutunun yüksek seçilmemesi eğitim süresinin hızlı tamamlanabilmesi için çok önemlidir. Oluşturulan modelle yüksek performans ve donanımsal yeterlilik sağlanması için uygun görüntü boyutu seçilmelidir (Arı & Hanbay, 2019).

2.2.2.1.2 Evrişim Katmanı

Bu katman dönüşüm katmanı olarak ta bilinmektedir. Bu dönüşüm katmanında filtreler görüntünün tamamında dolaştırılmaktadır. Bu filtreler katmanlı yapının olmazsa olmaz bileşenidir. Bir önceki katmandan gelen görüntülere filtreler aracılığı ile konvolüsyon işlemi

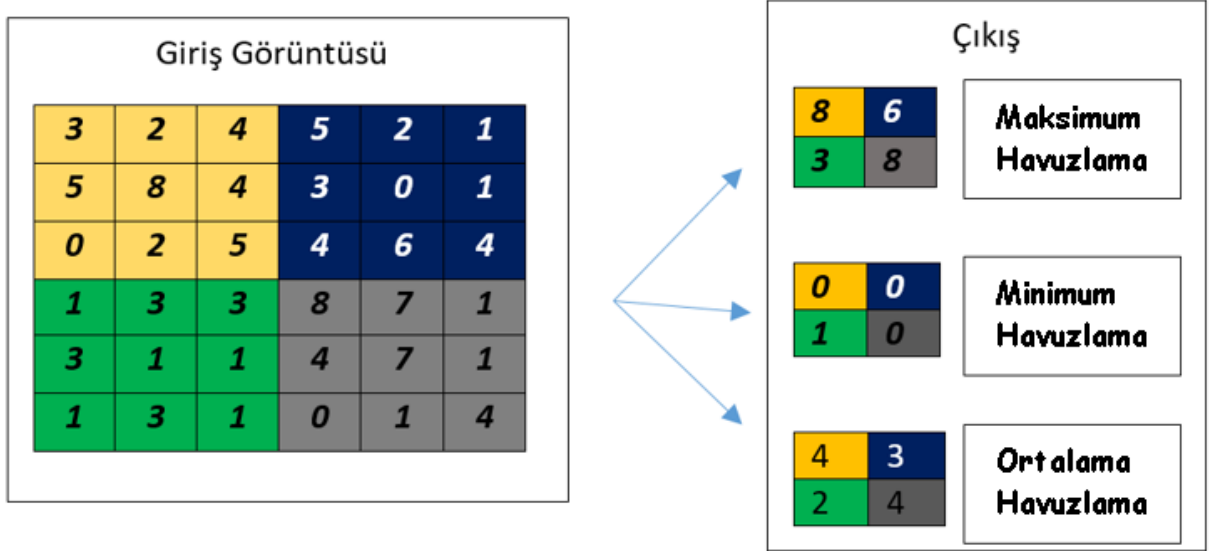
uygulanmakta ve çıkışa verilmektedir. Uygulanan konvolüsyon işlemi sonucunda bir özellik haritası (aktivasyon haritası) oluşturulmaktadır. Özellik haritaları, her bir filtreleme işlemi sonucu çıkarılan özelliklerin haritasıdır. Kullanılan filtrelerin katsayıları, modelin her eğitim yinelemesinde değişmektedir. Bu sebeple ESA'da elde edilen özellik haritasının her bölgesindeki bilginin ayrı bir önemi vardır. ESA mimarisinin evrişim katmanında bir görüntü üzerinde filtreleme işleminin örneği şekil 2.7'de gösterilmiştir. Şekil 9'da gösterildiği gibi renkli görüntü (RGB), boyutuda 5x5 olduğu düşünülürse resim boyutu 5x5x3 olacaktır. Evrişim katmanının filtre işleminin boyutunun 3x3 olduğu kabul edilirse görüntü üzerinde 3x3 boyutlarında tarama işlemi yapılacaktır. Bu tarama veya bir diğer adıyla dolaşma işleminde görüntü üzerinde soldan sağa 3x3'lük tarama yapıp sınıra gelindiğinde ise bir alt basamağa geçilip tekrar sağa doğru işlem devam edecektir. Görüntü haritasındaki 3x3 alandaki değerler filtre alanında 3x3 lük alandaki değerler ile çarpılıp toplanarak bir özellik(aktivasyon) değeri elde edilmektedir. Bu değerlerin bir araya gelmesiyle özellik haritaları oluşmaktadır. Özellik haritalarında giriş çıkış değerlerinin yoğunluk uygunluğunun sağlanması için normalizasyon işlemi yapılmaktadır. Normalizasyon işlemi için her bir renk kanalına ait değerlerin toplamının filtre değerlerinin toplamına bölümü ile elde edilmektedir (Kızrak & Bolat, 2018; Özkan & Ülker, 2017).



Şekil 2.7: Evrişim katmanında 5x5x3 görüntüye 3x3 filtrenin uygulanması örneği.

2.2.2.1.3 Havuzlama Katmanı

Havuzlama katmanları, derin öğrenme modellerinde boyut azaltmak için kullanılmaktadır. Boyut azaltma ve indirgeme işlemleri sınıflandırma performansını ve güvenilirliği düşürdüğü düşünülse de literatür çalışmaları incelendiğinde havuzlama işlemlerinin modelin ezberlenmesinin önlemesi ve çalışma sürelerinin kısaltılması gibi avantajlar sağladığı gözlemlenmiştir (Kizrak & Bolat, 2018). Bu katmanda da belirli fitreleme işlemleri ile görüntü üzerinde gezilir ve elde edilen maksimum ve minimum değerler ve bu değerlerin ortalaması alınarak bir çıktı elde edilir. Şekil 2.8’de örnek bir görüntü olduğu varsayılp havuzlama işlemi yapılmış ve maksimum, minimum ve ortalama değerlerle oluşturulan çıktı değerleri gösterilmiştir (Doğan vd., 2015).



Şekil 2.8: Havuzlama katmanı örneği.

2.2.2.1.4 Aktivasyon Fonksiyonu

Derin sinir ağlarında görüntü, ses ve sinyal gibi verilerin özelliklerini çıkarmak için aktivasyon fonksiyonlarına ihtiyaç duyulmaktadır. Aktivasyon fonksiyonları olarak genellikle literatürdeki kaynaklar incelediğinde görülmektedir ki Relu fonksiyonu kullanılmaktadır (Türkoğlu, 2019). Relu fonksiyonunun matematiksel denklemi aşağıda gösterilmiştir.

$$f(x) = \begin{cases} 0 & \text{eğer } x < 0 \\ x & \text{eğer } x \geq 0 \end{cases} \quad (2.7)$$

Denklem 2.7' de gösterildiği gibi Relu aktivasyon fonksiyonu giriş değeri eğer negatif ise çıktıyı sıfır yapmaktadır. Eğer giriş değeri pozitif ise giriş değeri olduğu gibi çıkışa aktarılmaktadır (Doğan vd., 2015).

2.2.2.1.5 Normalizasyon Katmanı

Derin evrişimli sinir ağlarının eğitilmesi işleminin süresi bazen oldukça uzun olabilmektedir. Bu sürelerin kısaltılabilmesi için nöron aktivasyonlarının normalize edilmesi gerekmektedir. Normalizasyon katmanı gizli nöronlardaki durumları stabil hale getirmede etkili olmaktadır. Genellikle evrişimli sinir ağlarının Relu aktivasyon katmanından sonra normalleştirme işlemi yapılmaktadır (Li vd., 2015).

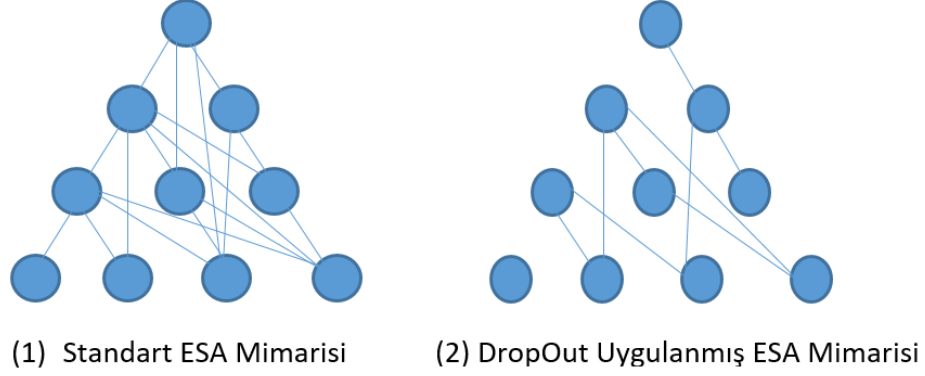
Normalizasyon işleminin yapılması ağın performansını olumlu yönde etkilemektedir. Bu nedenle ağ içerisindeki nöronların bir düzen içerisinde olması sağlanmalıdır. Bazen girdi verileri çok büyük veya çok küçük olabilmektedir. Normalizasyon işlemi ile verilerin dağılımları bir ortalama değere çekilmektedir. Normalizasyon işlemi sayesinde belirli bir bölgeyi temsil edebilme kabiliyeti de artmaktadır (Ioffe ve Szegedy, 2015).

2.2.2.1.6 Tam Bağlı Katman

Bu katman evrişim, havuzlama ve Relu aktivasyon katmanından sonra gelmektedir. Kendisinden önce gelen katmanlardan alınan verileri bağlamaktadır. Farklı mimarilerde katman sayıları değişebilmektedir. ESA mimarisinde $25 \times 25 \times 256$ matrisin son katmana geldiğini düşünürsek 160000×1 büyüklüğünde bir matris gelir. Tam bağlı katmanında matris değerinin 4096×1 olduğunu düşünürsek 160000 kendisinden önce gelen matris, 4096 tam bağlı nöronla bağlanacağı anlamına gelmektedir (İnik & Ülker, 2017).

2.2.2.1.7 DropOut Katmanı

Bu katman ESA mimarilerinde ezberlemenin önüne geçmek için kullanılmaktadır. Bu katmana gelen veriler birbirine bağlı düğümler halinde gelmektedir. Bu katmanda düğümlü gelen bu verilerin içerisinde bazı düğümlerin kaldırılması işlemi yapılmaktadır. Aşağıdaki şekil 2.9'da bir DropOut katmanı işlemi örneği gösterilmiştir. 1 numaralı şekilde düğümlerin daha çok olduğu 2 numaralı şekilde ise bazı düğümlerin kaldırıldığı gösterilmektedir. Bu tarz bir işlemle mimarinin ezberlenmesi ihtimali azaltılmaya çalışılmıştır (Kizrak & Bolat, 2018).



Şekil 2.9: DropOut düğüm kaldırma örneği.

2.2.2.1.8 Sınıflandırma Katmanı

Sınıflandırma katmanı, ESA mimarilerinin katman yapısında tam bağlantılı katmandan sonra gelmektedir. Bu katmanda sınıflandırma işlemi yapılmaktadır. Sınıflandırma yapılacak nesne sayısı bu katmanın çıkış değerine eşit olmaktadır. Örnek olarak 20 farklı tür nesne sınıflandırması işlemi yapılacak ise sınıflandırma çıkış değeri 20 olması gerekmektedir. Tam bağlantı katmanda 4096 özellik çıkarımı yapılacaksa sınıflandırma ağırlık matrisinin 4096x20 olması gerekmektedir. Sınıflandırma katmanında farklı sınıflandırıcı fonksiyonlar kullanılmaktadır. Genellikle softmax sınıflandırma kullanılmaktadır. Sınıflandırma işleminde 0-1 arasında bir çıkış elde edilmektedir. Çıkışın 1'e yakın değerleri, ağırlık tahmin ettiği nesne olduğu anlaşılmaktadır (İnik & Ülker, 2017).

2.2.2.2 ESA Mimarileri

Derin öğrenme tabanlı ve büyük veri kümeleri ile geliştirilmiş birçok ESA mimarisi bulunmaktadır. Bu mimariler, ILSVRC yarışmasındaki 1000 sınıflı çok çeşitli nesnelerden oluşan milyonlarca görüntü içeren ImageNet veri kümeleri ile eğitilmiştir. Bu çalışmada, görüntü sınıflandırma işlemlerinde AlexNet, VGG16, VGG19, GoogleNet, SqueezeNet, InceptionResNet, ResNet(18,50,101), InceptionResNet ve DenseNet201 mimarileri kullanılmıştır. Çalışmada ESA tabanlı nesne algılanması ve tespiti işlemlerinde ise, Efficient D0-D7, Faster R-CNN ve YOLO algoritmaları kullanılmıştır. Bu çalışmada, derin evrişimli sinir ağları tabanlı sınıflandırma işlemlerinde kullanılan mimarilerin derinlik, boyut, parametre ve görüntü boyutu gibi karakteristik özellikleri bulunmaktadır. ESA sınıflandırma

mimarilerinin karakteristik özellikleri tablo 2.1’de gösterilmiş ve alt başlıklarda detaylı olarak açıklanmıştır.

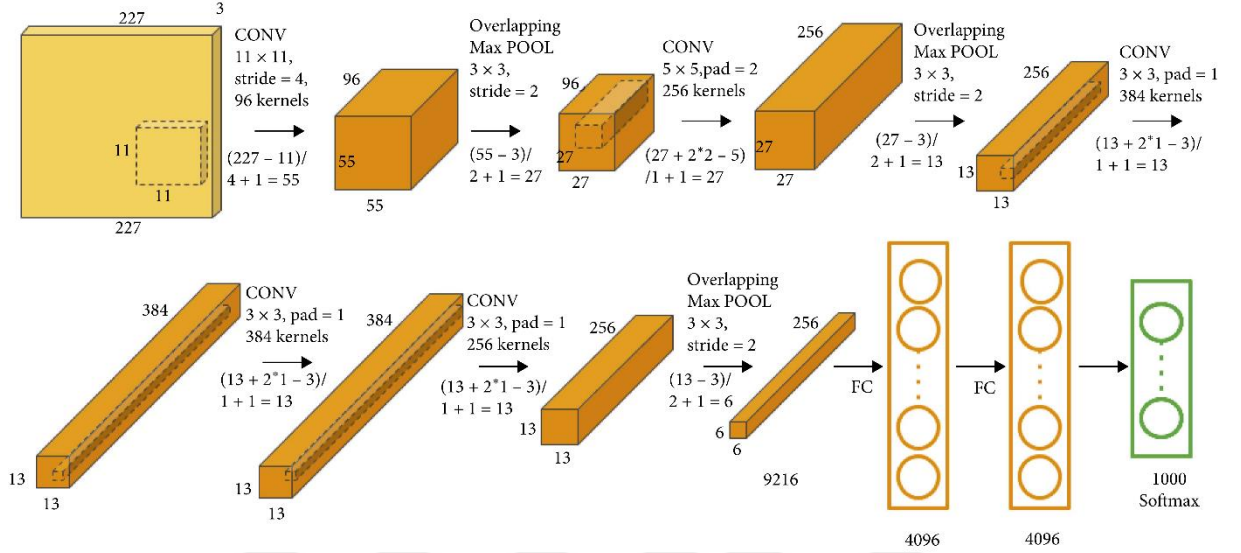
Tablo 2.1: ESA mimarilerinin karakteristik özellikleri.

Algoritmalar	Parametre	Derinlik	Boyut	Görüntü Boyutu
AlexNet (Krizhevsky vd., 2012)	61 Milyon	8	227	227x227
Google Net (Szegedy vd., 2015)	7 Milyon	27	27	224x224
VGG16 (Simonyan vd., 2015)	138 Milyon	16	515	224x224
VGG19 (Simonyan vd., 2015)	144 Milyon	19	535	224x224
SqueezeNet (Landola vd., 2016)	1,24 Milyon	18	4,6	227x227
Inception (Szegedy vd., 2016)	55,9 Milyon	164	209	299x229
ResNet18 (He vd., 2016)	11,7 Milyon	18	44	224x224
ResNet50 (He vd., 2016)	25,6 Milyon	50	96	224x224
ResNet101 (He vd., 2016)	44,6 Milyon	101	167	224x224
DenseNet (Huang, 2017)	20 Milyon	201	77	224x224

2.2.2.2.1 AlexNet

AlexNet mimarisi, 5 evrimsel, 3 maksimum havuzlama ve 3 tam bağlı katmandan oluşmaktadır. Alexnet mimarisinde giriş görüntü boyutu olarak 227x227 çözünürlüklü görüntü kullanılmaktadır(Krizhevsky vd., 2017). Bu mimaride doğrusallığı azaltmak için Relu aktivasyon fonksiyonuna ihtiyaç duyulmaktadır. Havuz katmanları sayesinde alt örneklerinden özellik haritaları çıkarılmaktadır. İlk tam bağlı katman, özelliklerin yüzde 50’sini veren özellik vektörünü düzleştirmek için kullanılmaktadır. Tam bağlı katmanlar 4096 nöron içermektedir. Son tam bağlı katman ise çıktı katmanı olarak görev yapmakta ve 1000 çıktı üretmektedir. Bu

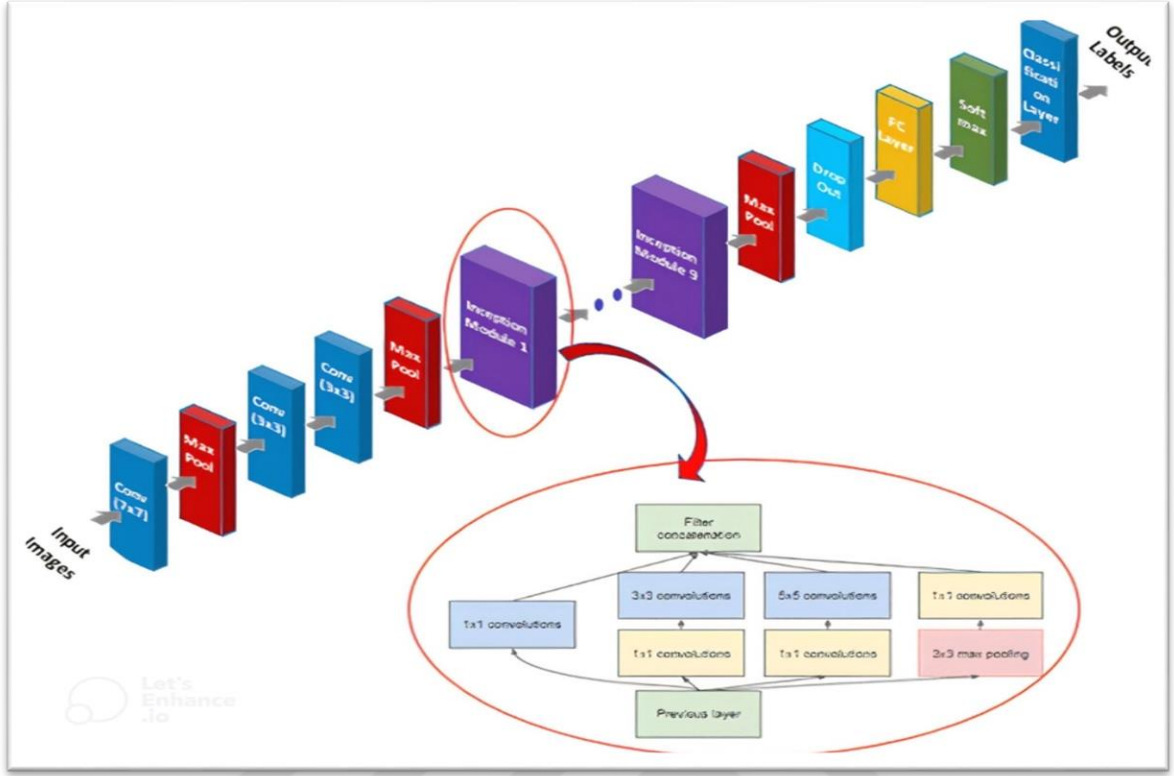
üretilen çıkışlar SoftMax aktivasyonundan geçirilmektedir. SoftMax aktivasyon fonksiyonu, olasılık dağılımlarını normalleştirmek için kullanılmaktadır. Son olarak normal ve anormal görüntüleri sınıflandırabilen bir ESA modeli eğitilmiş olacaktır. Şekil 2.10'da AlexNet mimarisinin katmanları ve parametreleri gösterilmiştir (Shaheen vd., 2021).



Şekil 2.10: AlexNet mimarisinin katmanları ve parametreleri (Krizhevsky vd.,2017).

2.2.2.2.2 GoogleNet

GoogleNet, diğer popüler ESA mimarilerinden farklı olarak 22 katmana sahip derin evrişimli sinir ağ mimarisidir. Daha önceki mimarilerden daha iyi performans göstermesi için tasarlanmıştır (Szegedy vd., 2015). 2014 yılındaki ILSVRSC yarışmasında ödül kazanmış bir mimaridir. Görsel katmanları tam olarak kaynaktan algılamada üstün performans sergilemiştir. GoogleNet mimari yapısı şekil 2.11'de gösterilmiştir.



Şekil 2.11: GoogleNet mimari yapısı (Szegedy vd.,2015).

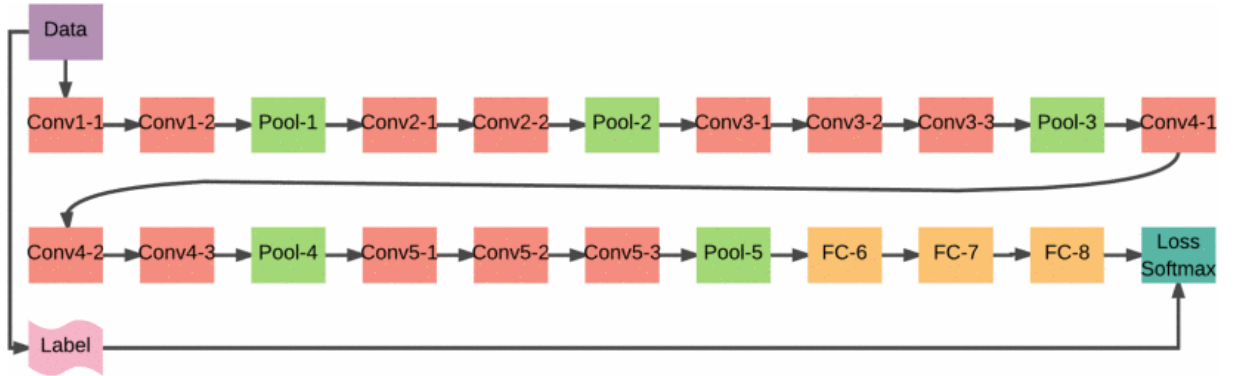
Tek katman içerisinde birden fazla çekirdeğin maksimum havuzlama ile bir veya daha fazla evrişime imkan tanıyan IM ağı en iyi ağırlıklarla eğitilmesini ve özellik haritasının çıkarılmasını sağlamaktadır. Bu işlem için, her giriş katmanı 1x1, 3x3 ve 5x5 matris boyutlarında değişken boyutlu çekirdek ve ayrıca 3x3 maksimum havuzlamaya sahiptir. IM mimarisi şekil 2.11’de gösterilmektedir. IM mimarisinde havuzlama katmanı ve beraberinde 1x1 filtreler, ikili boyutsallık azaltma ve kendinden önceki katmanlardaki içeriklerin özetlenmesi görevinde yerine getirmektedir. Şekil 2.11’de gösterildiği gibi GoogleNet, 9 IM katmanından meydana gelmektedir. Bu katmanlar görüntülerin daha iyi ayırt edicilik özelliklerini ortaya çıkarmaktadır. Ancak bu IM katmanlarının, bazı dezavantaj ve açıklanamayan karmaşıklıkları vardır. Büyük boyutlu özellik haritalarının oluşturulması hızlı eğitim süreci ve yeni ağda işlenmesini kolaylaştıracağı için tercih edilmektedir. GoogleNet, bu konuda küçük optimum özellikler seçen IM’leri ile çalışmaktadır. Bu çelişkinin üstesinden gelmek için GoogleNet büyük özellik haritalarını çarpanlarına ayırarak kullanmaktadır(Szegedy vd., 2016). Yani örnek verilmesi gerekirse, nxn evrişim, 1xn evrişim ve ardından nx1 evrişim ile uygulanabilir. Bu çarpanlara ayırma işlemi n boyutunun ilk ve son

için değil orta basamakları için iyi sonuç vermektedir. GoogleNet'te bazı diğer mimariler gibi kaybolan gradyan problemi görülmektedir. Bu problemin dezavantajının azaltılması için aktivasyon fonksiyonlarına başvurulmaktadır (Balagourouchetty vd., 2020; Huang vd., 2017; Luo & Zhang, 2014).

2.2.2.2.3 VGG16

VGG16 mimarisi, ImageNet'in milyonlarca verisinden oluşan veri tabanından eğitilmiştir. Bu kapsamlı eğitim sayesinde küçük boyutlu görüntülerde bile başarılı doğruluk skorları verebilmektedir.

VGG16 mimarisi 16 evrişim katmanı ve 3x3'lük matris boyutunda giriş verileri ile çalışmaktadır. Bu mimari 2x2 matris boyutunda Maksimum havuzlama katmanına sahip ve bunun gibi 5 katmana sahiptir. En sonuncu havuzlama katmanından sonra tam bağlı 3 katman bulunmaktadır. Son katman olarak Softmax sınıflandırıcı katmanı kullanılmaktadır. Relu aktivasyon katmanı tüm gizli katmanlara uygulanmaktadır. VGG16 mimarisinin katmanlı ağ yapısı Şekil 2.12'de gösterilmiştir (Simonyan & Zisserman, 2015; Theckedath & Sedamkar, 2020).

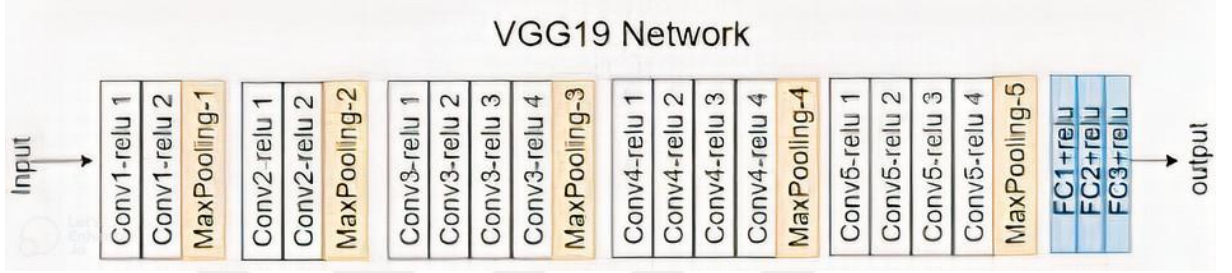


Şekil 2.12: VGG16 mimarisinin katmanlı ağ yapısı (Simonyan & Zisserman, 2015).

2.2.2.2.4. VGG19

VGG19 mimarisi 2014 yılında Simonyan ve Zisserman tarafından önerilmiştir (Simonyan & Zisserman, 2014). Bu mimari görüntülerin 1000 nesne kategorisinde sınıflandırılması için 16 evrişim katmanı ve 3 tam bağlı katmandan yani toplamda 19 katmandan oluşan evrişimli sinir

ağı mimarisidir. VGG19 mimariside VGG16 gibi ImageNet görüntüleri ile eğitilmiştir. Bu mimari, VGG16'dan farklı olarak 19 katmandan oluşmaktadır. Her evrişim katmanında 3x3 filtre kullanmaktadır. Bu özellikte VGG19'un popüler olarak kullanılmasını sağlamaktadır. VGG19 mimarisinde katmanlar 5 bölüme ayrılmıştır ve son olarak 3 tane Relu fonksiyonu ile birleşmiş tam bağlı katmandan oluşmaktadır. Bu mimaride giriş görüntülerinin boyutunun 224x224 olması gerekmektedir. VGG19 mimarisinin ağ yapısı şekil 2.13'te gösterilmiştir (Bansal vd., 2021).

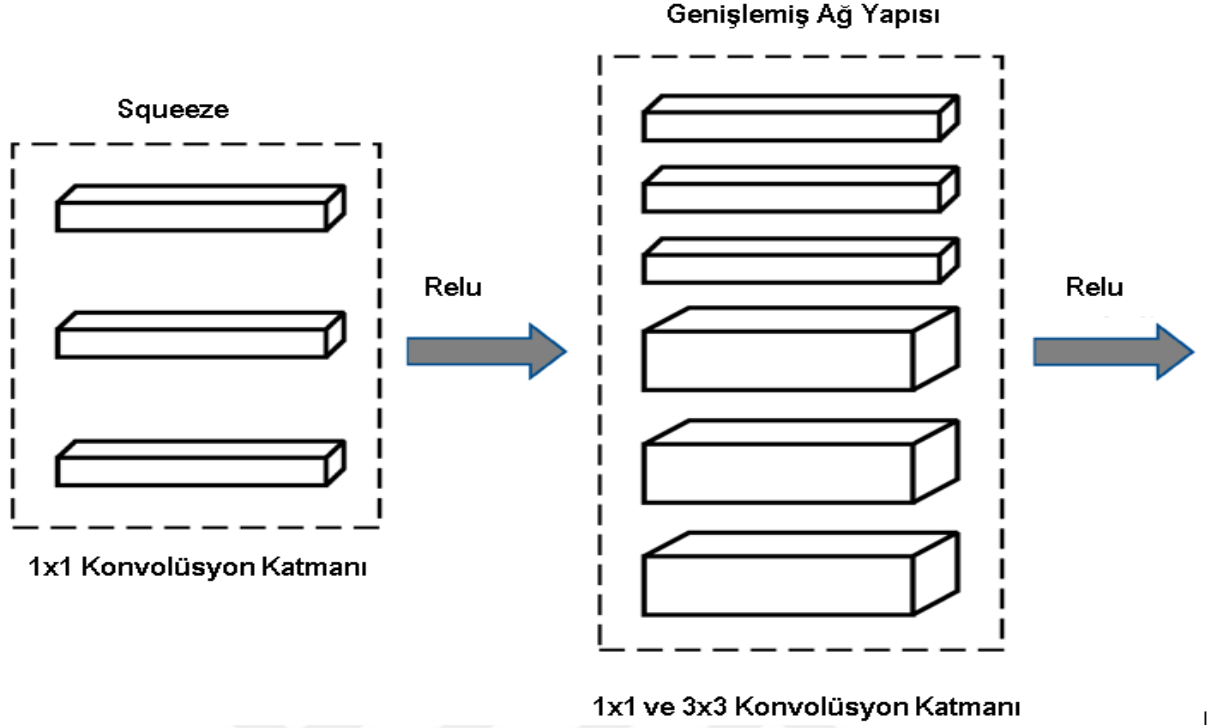


Şekil 2.13: VGG19 mimarisinin ağ yapısı (Bansal vd., 2016).

2.2.2.2.5. SqueezeNet

SqueezeNet mimarisi, AlexNet ve VGG mimarilerinin parametre sayıları aşırı büyüdüğü için AlexNet mimarisinden 50 kat daha az parametre ile çalışan mimari olarak önerildi. Bu mimari 2 ayrı yapı ile oluşmuştur. Bunlar Squeeze ve Expand katman yapılarıdır. Squeeze katmanı, S 1x1 evrişim çekirdekleri içermektedir. Expand katmanı ise, 1x1 ve 3x3 evrişim çekirdekleri içermektedir.

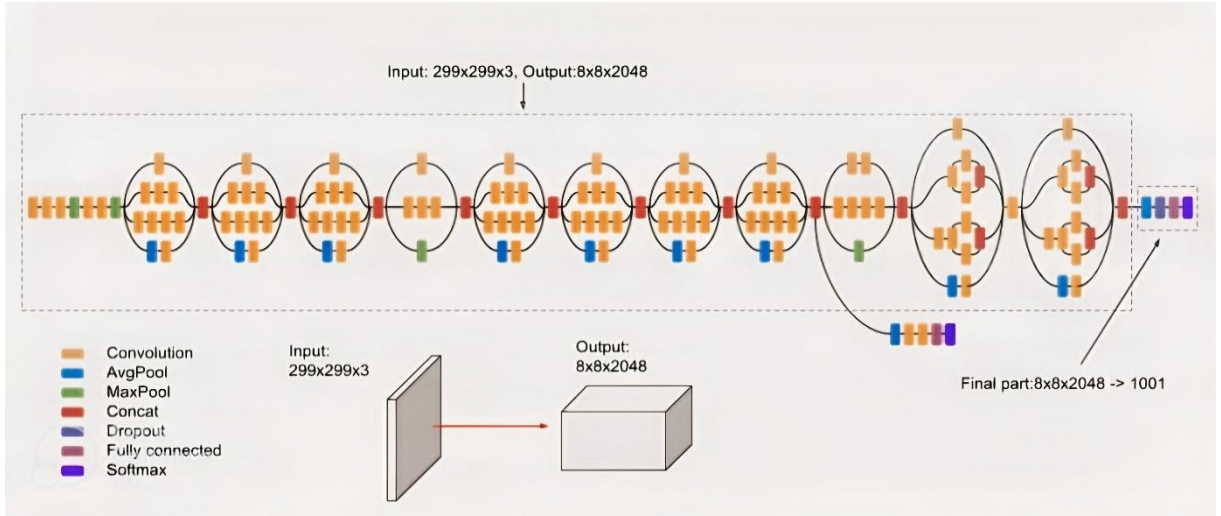
SqueezeNet, parametre sayısını 9 kat azaltmak için 1x1 evrişimleri 3x3 olarak genişlemektedir. Bu mimaride, örnekleme sayısı azaltma işlemi ertelenip, daha büyük aktivasyon grafiği ve daha yüksek bir sınıflandırma doğruluğu sağlanabilmekte ve bu sayede bilgiler korunabilmektedir. SqueezeNet, 1BatchNorm katmanı, 3 blok katmanı, 4 evrişim katmanı ve bir global ortalama havuzlama katmanından oluşmaktadır (Xu vd., 2020). SqueezeNet mimarisi ağ genişleme yapısı şekil 2.14'te gösterilmiştir.



Şekil 2.14: SqueezeNet Mimarisi Genişletilmiş Ağ Yapısı (Lee vd., 2019).

2.2.2.2.6. Inception

Inception, Szegedy vd. tarafından önerilen bir ESA mimarisidir. Bu sinir ağı modeli birçok maksimum havuzlama ve dolayısıyla evrişim katmanından meydana gelmiştir. Son katmanında ise tam bağlı katman bulunan bir sinir ağı yapısına sahiptir. Bu modelin ağırlıkları hem VGGNet ve ResNet mimarilerine göre daha küçüktür. Bu modelin birçok farklı özelliklere sahip sürümleri vardır. Bunların içerisinde en yaygın kullanılan sürümü InceptionV3 mimarisidir. InceptionV3 mimarisinde, 350 bağlantı ve 316 katman bulunmaktadır. Ayrıca bu modelin giriş görüntüsü boyutu 299x299x3 ebatlarındadır. Inception mimarisinin ağ yapısı şekil 2.15’de gösterilmiştir (Szegedy vd., 2016).

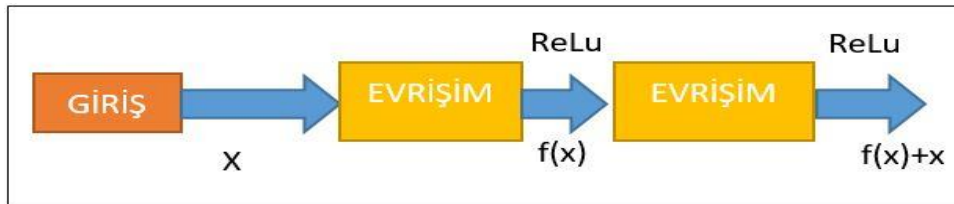


Şekil 2.15: Inception mimarisi ağ yapısı (Journal t.y., 2019).

2.2.2.2.7. ResNet

He vd. tarafından geliştirilen bu mimari daha yüksek derinlikli sinir ağlarını eğitmek için geliştirilmiştir. ResNet mimarisi, mikro mimarilere dayanmaktadır (He vd., 2016). VGG ve Alexnet mimarilerinden farklı olarak ardışık geleneksel yapıdan ziyade mikro bir yapıya sahiptir. ResNet mimarisi, VGG ağından daha derin çalışır ve ağıın parametre ve boyutları düşüktür. Bu mimari, Residual blok yapısından oluşmaktadır. Bu nedenle artık değerlerin sonraki katmalara eklenmesi sistemine dayanmaktadır. Bu mimarinin öne çıkan özelliği derinlik artsa bile hızlı çalışmaktadır.

Mimarinin katmanlarında oluşan artık blokta x bir girişin sırasıyla evrişim-ReLu-evrişim katmanlarını izlemesinden sonra bir $f(x)$ fonksiyon değeri elde edilmektedir. Bu sonuç giriş değerine eklenir ve son katmana $f(x)+x$ olarak aktarılır. Şekil 2.16'da ResNet artık modül çalışma yapısı gösterilmiştir (Doğan & Türkoğlu, 2019).



Şekil 2.16: Artık blok modül çalışma yapısı.

ResNet mimarisinde birçok versiyon model geliştirilmiştir. Bunlardan yaygın olarak kullanılanları ResNet18, ResNet50 ve ResNet101 modelleridir. Bu modellerin yanlarında verilen rakamsal ifadeler modelin derinliğini temsil etmektedir. Ayrıca bu modellerin bağlantı ve katman sayıları tablo 2.2’de gösterildiği gibidir (Doğan & Türkoğlu, 2019).

Tablo 2.2: ResNet Bağlantı ve Katmanları.

MODEL	BAĞLANTI	KATMAN
ResNet18	79	72
ResNet50	192	177
ResNet101	379	347

2.2.2.2.8 DenseNet

Huang vd. tarafından önerilen bu modelde, diğer ağ modellerinden farklı olarak her katman kendisinden sonra gelen diğer katmanlara bağlıdır. DenseNet mimarisinde bütün katmanlar için, özellik haritaları ayrı bir giriş olarak değerlendirilmekte ve her katmanın özellik haritası bir sonraki katmana girdi olarak aktarılmaktadır. Ayrıca bu modelin, karmaşıklığını azaltmak ve özellik parametrelerinin verimli kullanılabilmesi için dar boğaz yapısı ve havuzlama katmanları kullanılmaktadır (Bahad vd., 2019; Chen vd., 2017; Liang vd., 2018; Nguyen vd., 2018).

DenseNet mimarisi ResNet mimarisi ile benzer özelliklere sahiptir. Ancak ResNet Mimarisinde, her blok kendisinden önceki bloktan haber alınabilirken, DenseNet mimarisinde bütün bloklardan bilgi alınmaktadır. DenseNet mimarisinin, ResNet mimarisine göre ayrıca avantajlı farkları şunlardır;

- DenseNet, ResNet mimarisinden daha güçlü gradyan akışı sağlamaktadır.
- Hesaplama ve parametreler daha verimlidir.
- Özellikleri daha çeşitlidir.
- Sınıflandırmada daha etkilidir.

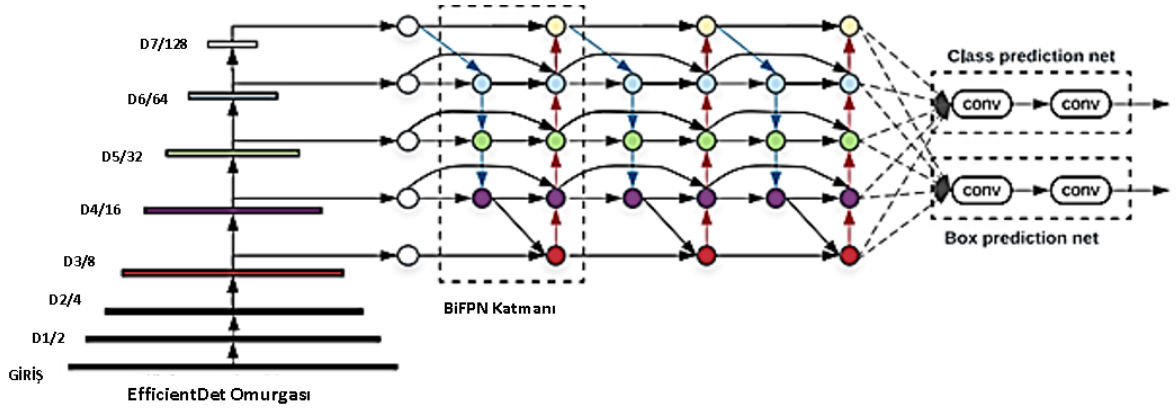
DenseNet mimarisinde giriş katmanında görüntü boyutu olarak 224x224x3 kullanılmaktadır (Vapnik, 2010).

2.2.3 Görüntülerde Nesne Algılama ve Tespiti Modelleri

Bu çalışmada nesne algılama ve tespiti çalışmaları için kullanılan derin öğrenme tabanlı modellerin karakteristik özellikleri alt başlıklar halinde açıklanmıştır.

2.2.3.1 EfficientDet

Bu mimari, Tan ve Pang (2020), tarafından ölçeklenebilir nesne algılamaya yönelik olarak geliştirilmiştir (Tan vd., 2020). EfficientDet mimarisi, EfficientDet D0'dan D7'ye kadar sekiz ayrı algoritma modeli içermektedir (S. Song vd., 2021). EfficientDet, hafif, ölçeklenebilir bir algılama Ağı'dır. D0'dan D7'ye, modelin doğruluğu ve zaman karmaşıklığı, model boyutuyla birlikte artmaktadır. Sekiz model, geniş bir kaynak kısıtlamaları yelpazesini karşılamaktadır. EfficientDet'in omurgasında, modeli daha hafif hale getirmek için bir derin ayrılabilir evrişim31 kütlesi kullanan EfficientNet30 kullanılmaktadır. Ayrıca, ağıın boyun kısmı çift yönlü ağırlıklı özellik piramit ağı (BiFPN) kullanır. Yol toplama ağından (PANet)32 başlayarak BiFPN, yalnızca bir giriş kenarı olan düğümleri kaldırır, ardından aynı seviyede giriş ve çıkış düğümü arasına bir kısayol ekler. Amaç, hesaplama maliyetini artırmadan daha fazla özelliği birleştirmektir. Farklı giriş özelliklerinin çözünürlüğü nedeniyle, özelliklerin birleştirilmesi sırasında her giriş özelliğine ek ağırlık eklenmiştir. Bu ağırlıklı füzyon, ağda farklı özelliklerin önemini öğrenmesini sağlamaktadır. EfficientDet mimarisi şekil 2.17'de gösterilmiştir.

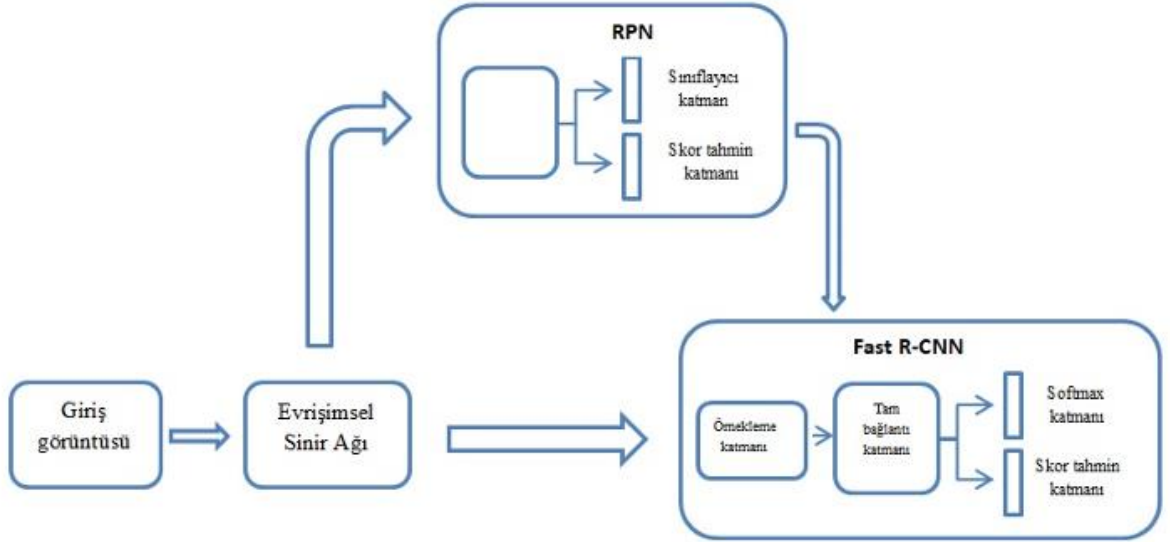


Şekil 2.17: EfficientDet mimarisi (Tan vd., 2020).

2.2.3.2 Faster R-CNN

Faster R-CNN, Ross Girshick, Shaoqing Ren, Kaiming He ve Jian Sun tarafından 2015 yılında geliştirilen bir nesne algılama modelidir. Bu mimari, konvolüsyon sinir ağı ile nesne tespiti için kullanılabilir. Daha hızlı R-CNN, evrişim katmanları, bölge öneri ağı ve sınıflar ve sınırlayıcı kutu tahmini olarak üç bölümden oluşmaktadır (Ren vd., 2015).

Faster R-CNN, R-CNN mimarisinin neden olduğu zaman kaybını önlemek için geliştirilen bir mimari. Bu mimaride, R-CNN mimarisinden farklı olarak önce bölgeyi belirlemek yerine, önce görüntüyü evrişimli sinir ağına gönderir ve daha sonra bölge önerisi yapılmaktadır. Faster R-CNN'nin R-CNN'den diğer bir farkı ise CNN (ESA)'nın sınıflandırma için kullandığı destek vektör makineleri yöntemini değil, derin öğrenme yöntemi olan softmax sınıflandırmasını kullanmasıdır. Tüm işlemlerde CNN (ESA) bir kez kullanıldığı için işlem süresi oldukça kısalmaktadır. Fast R-CNN, önce görüntüyü evrişimli sinir ağı ile işlemekte ve bir özellik haritası elde etmektedir. Fast R-CNN, R-CNN'den farklı olarak, haritada bölgesel seçim yapmaktadır (Yılmaz vd., 2020). Faster R-CNN mimarisi şekil 2.18'de gösterilmiştir.



Şekil 2.18: Faster R-CNN mimarisi (Yılmaz vd., 2020).

2.2.3.3 YOLOv4

YOLOv4, DarkNet Kütüphanesi tabanlı sinir ağı çerçevesinde oluşturulmuş bir nesne algılama ve tespiti algoritmasıdır. YOLOv4, performans ve süre bakımından başarılı sonuçlar vermesi

nedeniyle birçok uygulamada yaygın kullanılmaktadır. Derin ESA mimarisi kullanarak nesneleri tespit edebilmekte ve nesne koordinatlarını göstermektedir. Veriler eğitilip YOLOv4 ağına görüntülerin aktarılmasından sonra görüntüler üzerindeki nesnelerin hangi sınıfa ait olduğunu göstermektedir. Bu işlem sonucunda bir güven skoru oluşturmaktadır. Güven skoru, tahmin doğruluk performansını göstermektedir. Güven skoru 0 ile 1 arasında olmaktadır (Güçlü vd., 2021). Güven skorunun doğru ölçülebilmesi için görüntünün örtüşme ölçütü olan IoU (Intersection over Union) değerinin önceden bilinmesi gerekmektedir. Görüntüler üzerinde analizler yapılırken, görüntülerin en dip pikselinden itibaren kesitler alınıp ve kesitlerin bütünü temsil ne kadar temsil ettiğinin belirlenmesi için IoU değerine bakılması gerekmektedir. IoU değerinin hesaplanması için kullanılan matematiksel formül denklem (2) de, Güven skorunun hesaplanması için kullanılan matematiksel formül denklem (1) de verilmiştir. Görüntü kesit örtüşmesi örneği şekil 2.19’da verilmiştir.

$$\text{Güven Skoru} = Pr(Nesne) * IoU \quad (2.8)$$

$$IoU = \frac{\text{Kesişim(Gerçek-Referans} \cap \text{Tahmin)}}{\text{Birleşim(Gerçek-Referans} \cup \text{Tahmin)}} \quad (2.9)$$



Şekil 2.19: Görüntü Kesit Örtüşmesi (IoU).

YOLOv4 ağı, üç bölümden meydana gelmiştir. Bu bölümler omurga, boyun ve kafa yapılarıdır. Omurga, farklı görüntüleri analiz edip özellik çıkarıp işleyen evrimsel sinir ağı bölümü olarak tanımlanmaktadır. YOLOv4’de alt yapı olarak CSPDarknet53 mimarisi kullanılmaktadır (Wang vd., 2020). Ayrıca omurga bölümü içinde Bag of Specials (BoS) ve Bag of Freebies (BoF) bölümleri bulunmaktadır. Bu bölümler, veri kümelerinin zenginleştirilmesi için kullanılmaktadır (Güçlü vd., 2021). Boyun, nesnelerin farklı boyutlarda da tespit edilebilmesi için kullanılmaktadır. Kafa bölümü, YOLOv3 ile benzer bir yapıya sahip Dense Prediction

katmanı olarak bilinmektedir. Kafa bölümünde nesnelerin görüntü içi yerleşim koordinatlarını, etiketlerini ve güven skorlarını içeren bir vektör oluşturulmaktadır. Ayrıca YOLOv4'te SAM, SPP, PANet ağları da etkin olarak kullanılmaktadır (He vd., 2015; Liu vd., 2018; Woo vd., 2018).

2.2.3.4 YOLOv5

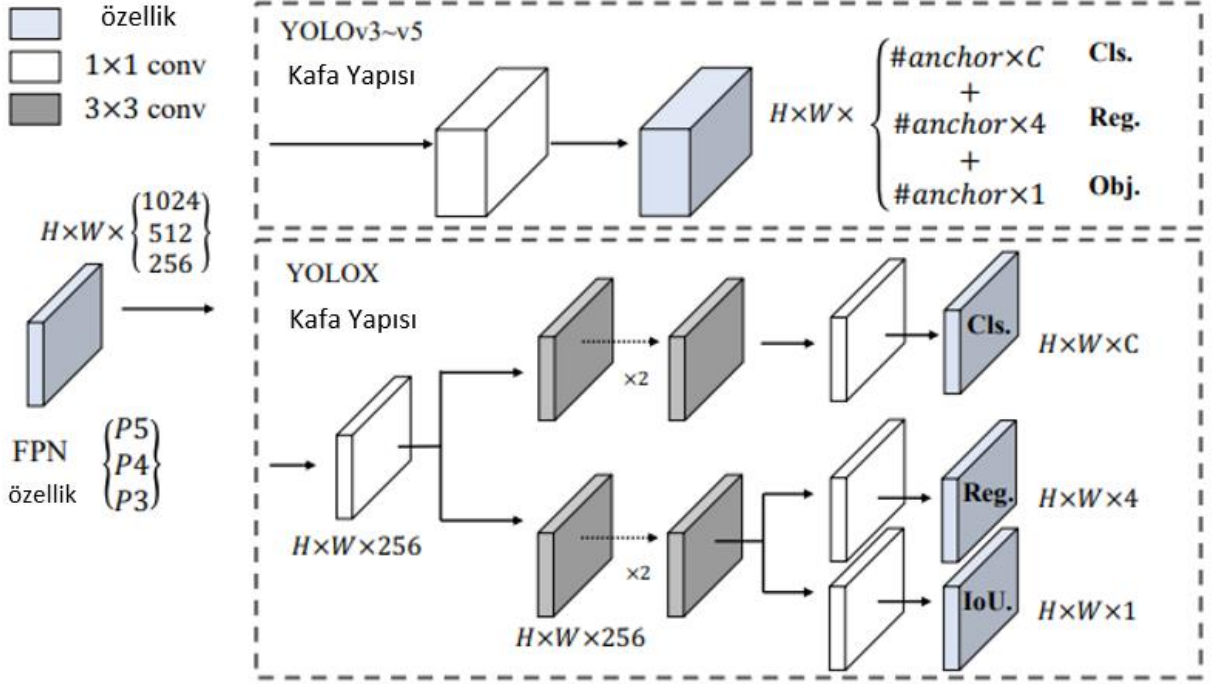
YOLOv5 modelinin mimarisinde, YOLOv4 modelinde olduğu gibi omurga, boyun ve kafa yapıları kullanılmıştır. Giriş görüntüleri omurga (backbone) yapısı kullanılarak özellik çıkarma işlemi yapılır. Omurga ile elde edilen özellikler kullanılarak, özellik haritası şeması oluşturulmaktadır. YOLOv5, özellik haritalarını kafa (head) yapılarına aktaran boyun (neck) aşağıdan yukarı, yukarıdan aşağı gibi çift yönlü oluşturmaktadır. Tek boyutlu nesne tespiti işlemlerinde, kafa yapılarında performanslı güven skoru ve başarımlı skor sonuçları elde edilmektedir.

YOLOv5, YOLOv4'den farklı olarak PyTorch kütüphanesini kullanmaktadır. (Murat, t.y., 2018).

2.2.3.5 YOLOX

YOLOX, Akış Algı Yarışmasında (CVPR2021) Otomatik Sürüş Workshop'u olarak ilk olarak tanıtılmıştır. DarkNet kütüphanesinin alt yapısını kullanan bir modeldir. YOLOv4 ve YOLOv5'den farklı olarak, bu yapı YOLOv3-SPP ağ yapısı ile geliştirilmiştir. YOLOv3-SPP, gelişmiş bir CSPNet omurgasını ve ek bir PAN kullanılmasını benimseyerek YOLOv5 mimarisi kullanılarak geliştirilmiş ağ yapısıdır (Ge vd., 2021).

Bu modelde, nesne algılama ve tespiti sistemlerinde işlem çakışıklığına neden olan ve doğruluğun azalmasına neden olabilecek sınıflandırma ve regresyon görevleri aynı anda gerçekleştirilir. Bu problemin çözümünde ayrıştırılmış kafa yapısı kullanılmaktadır. Geleneksel YOLO serilerinde, omurga ve özellik piramitleri hala klasik kafa yapıları ile elde edilmektedir. Ancak YOLOX modelinde, ayrıştırılmış bir kafa yapısı kullanılmış ve daha performanslı doğruluk skorları elde edilmesi hedeflenmiştir.



Şekil 2.20: YOLOX Modeli ile YOLOv3-v5 Modelinin Yapısı Arasındaki Fark (Ge vd., 2021).

Yukarıdaki şekil 2.20’de YOLOv3-v5 modellerinin yapısı ile YOLOX modelinin yapısı arasındaki fark gösterilmiştir. Her bir FPN seviyesi için özellik kanalının 256’ya dönüştürülmesi için önce 1x1 konvolüsyon katmanı benimsenmiş ve iki paralel dal eklenmiştir. Sırasıyla sınıflandırma ve regrasyon görevleri için 3X3 dönüşüm katmanı eklenmiştir. Ayrıca regrasyon dalına IoU dalı eklenmiştir (Ge vd., 2021).

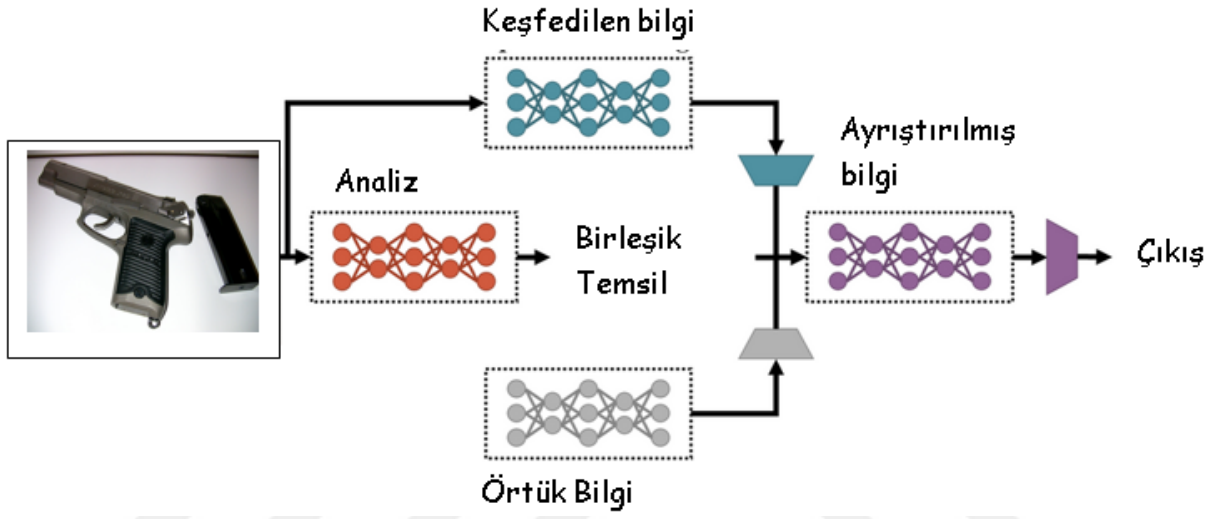
YOLOX, Mozaik ve karıştırma stratejileri kullanılarak güçlü bir şekilde güçlendirilmiş bir veri kümesi üzerinde eğitilmiştir.

2.2.3.6 YOLOR

YOLOR, YOLOv5 modeli ile benzer bir şekilde Pytorch Kütüphanesi kullanılarak geliştirilmiş bir modeldir (YOLOR object detection in Gradient, 2022). CSPNet, doğruluk performansları üzerinde azalmaya neden olsa da YOLO modellerinde algılama hızını artırabilmek için ultra verimli evrişimli katmanlar kullanmaktadır.

Günümüzde ESA tabanlı kullanılan modeller, ağı önce genel bir görev için büyük veri kümesi ile eğitmekte ve sonra özel görevler için veri kümeleri üzerinde ince ayarlarla düzenlemeler yapmaktadır. Bu eğitim aşamalarında ince ayar düzeltme işlemleri ile modelin ağırlıkları da

ayarlanabilmektedir. YOLOR, COCO tanım sınıflarında mevcut görevlerle, yani nesne algılama, nesne segmentasyonu, optik segmentasyon, anahtar nokta algılama ve tespiti, mikro segmentasyon, resim yazısı, çok sınıflı görüntülerin ayrıştırılıp sınıflandırılması ve uzun kuyruklu nesne tanıma ile örtük bir bilgi ağını önceden eğitmesini sağlamaktadır (Wang vd., 2021).



Şekil 2.21: YOLOR modelinin çalışma yapısı.

Şekil 2.21’de YOLOR modelinin çalışma yapısı gösterilmektedir. YOLOR, COCO kümesini tespit ederken açık bilgiyi temsil eden başka bir parametre kümesini de eğitebilmektedir. Tespit ve algılama tahmini için ise hem gizli hem de açık bilgiyi kullanmaktadır (Wang vd., 2021).

2.3. BOYUT İNDİRGEME

Bu çalışmada, videoların depolanması işlemlerinde kullanılan depolama birimlerinden kar edilebilmesi için videoların boyutlarında indirgeme çalışması yapılmaktadır. Çalışmanın kapsamı nedeniyle bu başlık altında videolarda boyut indirgenmesi konusu açıklanmıştır.

Videoların uygun koşullarda saklanabilmesi için depolama birimine ihtiyaç duyulmaktadır. Videonun sayısı, süresi ve kalitesi arttıkça kullanılan depolama birimlerinin de kapasitesi aynı doğrultuda artmaktadır. Depolama birimlerinin arttırılması da dolaylı olarak maliyetlerin artması anlamına gelmektedir. Depolama birimlerinin tasarruflu kullanılması gerekmektedir. Tasarruflu depolama birimi kullanımı için kullanılan yöntemlerden biri depolama içeriklerinin

boyutlarının indirgenmesi işlemidir. Bu çalışmada da videolarda boyut indirgemesi işlemi yapılacaktır.

2.3.1. Video Boyut İndirgemesi

Videoların boyut indirgemesi işlemi daha önceleri birçok yöntem kullanılmıştır. Video indirgeme ve özetleme konusunda farklı yöntemler kullanan literatür çalışmalarından bazıları;

Görmüş ve Akar (2019) çalışmasında, Bergama Bölgesindeki kentsel ve kırsal alanlarda VNIR kameralar ile kayıtlar almışlardır. Bu kayıtların indirgenmesi için ilk olarak atmosferik ve radyometrik düzenlemeler yapmışlardır. Kayıtlar üzerinde boyut indirgeme işlemi için bozuk kullanılmayan kayıt bantlarının silinmesi için 5 farklı yöntem kullanmışlardır. Bu yöntemler temel bileşen analizi (PCA), ayırık dalgacık dönüşümü (DWT), faktör analizi, çok boyutlu ölçekleme ve en fazla daralan metrik öğrenme (MCML) yöntemini kullanılmıştır. Çalışmanın sonucunda en verimli yöntemin DWT yöntemi olduğunu savunmuşlardır (Görmüş & Akar, 2019).

Masala vd. (2014) çalışmasında, görüntülerin boyutlarının indirgenmesi için veri madenciliği algoritmaları ile boşlukların tespit edilmesini ve bu boşlukların videolar üzerinde az hasar oluşturabilecek şekilde tespit edilip temizlenmesini önermişlerdir. Çalışmada veri madenciliği algoritmalarından SWM algoritması ile görüntüler üzerindeki boşlukların tespit edilmesi işlemini gerçekleştirmişlerdir (Masala vd., 2014).

Chakraborty vd. (2015) çalışmasında, video özetleme için uyarlanabilir ana kare seçimi üzerine yöntem geliştirmiştir. Bu çalışmada görüntülerde ana karenin önemli olduğunu savunmuş ve bazı parametreler geliştirmiştir. Bu parametrelere göre ana karelerin dışındaki verilerde silme işlemi yapılarak özetme gerçekleştirilmiştir (Chakraborty vd., 2015).

Hussain vd. (2020) çalışmasında, ESA ve çift yönlü LSTM kullanılarak bulut destekli çoklu görüntü video özetleme işlemi yapılmıştır. Kurum ve kuruluşların gözetleme birimleri tarafından elde edilen video verilerinin video özetleme, indeksleme ve analiz için kullanılmasında çeşitli zorluklarla karşılaşmaktadır. Özellikle çoklu görüntü video özetleme (MVS) sistemlerinin devasa veri boyutları, fazlalık, görünümde üst üste binme, ışık değişimleri ve röportaj korelasyonları nedeniyle çalışmada zorluklar yaşanmaktadır. Yazarlar yaşanan zorluklardan üstesinden gelebilmek için MVS' den tam olarak yararlanılmayan çeşitli düşük seviyeli özellikler ve kümeleme tabanlı çalışan bilgi işlem

tekniklerinin kullanılmasını önermişlerdir. Derin sinir ağı tabanlı yumuşak bilgi işlem teknikleri iki katmanlı bir çerçeveye entegre edilerek yeni bir MVS elde edilmiştir. İlk katman çevrimiçi katmandır. Çevrimiçi katmanda hedef görüntülere dayalı çekim segmentasyonu gerçekleştirilmiş ve bir buluta aktarılması işlemi yapılmıştır. İkinci katmanda ise, bulut verisindeki her bir verinin ayıklanması ve derin özelliklerin çıkarılması işlemi yapılır ve derin çift yönlü uzun kısa süreli belleğe aktarılması (LSTM) ve video özetleme işlemi gerçekleştirilmiştir. Sonuç olarak sistemin YouTube karşılaştırılmalı veri kümeleri üzerindeki deneylerde en iyi MVS yöntemlerinden daha iyi performans gösterdiği gözlemlenmiştir (Hussain vd., 2020).

Literatür çalışmaları incelendiğinde kullanılan yöntemlerle genellikle ya veri silme ya da görüntü kalitesini bozma işlemi yapıldığı görülmektedir. Bu tez çalışmasında, örnek kullanım alanı olarak güvenlik kamera sistemleri tercih edilmiş ve verilerin kaybolmaması yani sadece ana kare seçimi yapılması güvenlik zafiyeti doğuracaktır. Bu sebeple ana kare seçimi yöntemi bu çalışma için uygun olmamaktadır. Bu çalışmanın kullanım alanından dolayı her türlü verinin önemi vardır. Her türlü farklı verinin kayıt izleme aşmasında gösterilmesi gerekmektedir. Bu sebeple güvenlik kamerası sistemlerinde güvenlik için her türlü verinin bilinmesi gerekliliği ile hareket edilirse güvenlik kamera sistemlerine uygun bir yöntem tercih edilmesi gerekmektedir. Bu sebeple bu çalışmada güvenlik zafiyeti oluşturmayacak yeni yöntem sunulmuştur.

Bu çalışmada, derin öğrenme yaklaşımları ile görüntülerin analizleri ve çeşitli görüntü filtreleme yöntemleri kullanılarak boyut indirgemesi işleminin yapılabileceği savunulmuştur. Derin öğrenme yaklaşımları, derin öğrenme başlığı altında açıklanmıştır. Görüntü filtreleme yaklaşımları konusu ise, görüntü filtreleme yaklaşımları başlığı altında açıklanmıştır.

2.3.2 Görüntü İşleme ve Filtreleme Yaklaşımları

Bu bölümde boyut indirgeme işlemi için, derin öğrenme ile elde edilen görüntü işlemenin haricinde görüntünün piksel özellikleri için uygulanan algoritmik görüntü işleme ve filtreleme yaklaşımları açıklanmıştır.

2.3.2.1. Görüntü İşleme İkili Görüntü Dönüşümü

İkili görüntü veya binary görüntü dönüşümü, görüntülerin ikili (0-1) görüntüye dönüştürülmesi işlemine denilmektedir. Görüntüyü iki renk meydana getireceğinden bu renkler siyah ve beyaz olacaktır. Siyah renkleri 0 rakamı, Beyaz renkleri 1 rakamı temsil etmektedir (Russ, 2006).

Görüntünün ikili biçime dönüştürülmesi için görüntünün orijinal haline bakılmaktadır. Görüntünün piksel değerlerinin eşik değeri hesaplanır. Eşik değeri, görüntünün en büyük piksel değeri ile en küçük piksel değerinin ortalaması alınarak hesaplanmaktadır. Eğer görüntünün piksel değerleri, eşik değerinin altında ise 0, üstünde ise 1 kabul edilir ve görüntü ikili görüntü biçimine dönüştürülmüş olmaktadır.

Renkli görüntünün, ikili görüntü biçimine dönüştürülmesinin kodu ve ekran çıktısı şekil 2.20’de gösterilmiştir.



Şekil 2.22: Renkli Görüntü İkili Form Çıktısı.

2.3.2.2. Görüntünün Gri Ölçek Dönüşümü

Gri seviyeli veya ölçekli görüntüler, farklı tonlardaki gri görüntülerden meydana gelmektedir. Gri renk değeri 0-255 arasında değişkenlik gösterebilmektedir. Görüntü işlemede yaygın olarak kullanılan gri renk dönüşümü, görüntü hesaplama maliyetlerini önemli düşürmektedir. Bu sayede görüntü işleme süresini kısaltılacağı için anlık görüntülerde görüntü işlemeyi kullanışlı hale getirmektedir. Görüntünün gri haline yapılacak her türlü işlem görüntünün renkli haline de yapılabilir. Şekil 2.23’te renkli görüntünün gri tona dönüştürülmesinin kodu ve ekran çıktısı verilmiştir (Aydın, 2021).



Şekil 2.23: Renkli görüntünün gri renk dönüşümü.

2.3.2.3. Histogram Eşitleme Yöntemi

Histogram eşitleme, görüntü üzerindeki renklerin dağılımının düzgün yapılabilmesi için kullanılan renk iyileştirme tekniğidir. Eğer görüntünün tamamında yapılırsa global histogram, sadece bir bölgesinde bu işlem yapılırsa lokal histogram olarak adlandırılmaktadır (Aydın, 2021).

Histogram eşitleme işlemi, ilk olarak görüntünün histogramının bulunması ile başlamaktadır. Histogramı bulunan görüntünün birleştirilmiş kümülatif histogramı çıkarılır. Elde edilmiş olan birleştirilmiş kümülatif histogram değerleri ile maksimum renk değerleri çarpılarak toplam piksele bölünür ve normalleşme işlemi yapılmış olmaktadır. Yeni renk değeri ile tüm renklerin güncellenmesi işe histogram eşitleme işlemi tamamlanmış olmaktadır. Histogram eşitleme işlemi, genellikle yetersiz ortam ışıklarında çekilmiş görüntülerin iyileştirilmesinde kullanılmaktadır.

Şekil 2.24'te histogram eşitleme yöntemine örnek kod ve ekran çıktısı paylaşılmıştır.

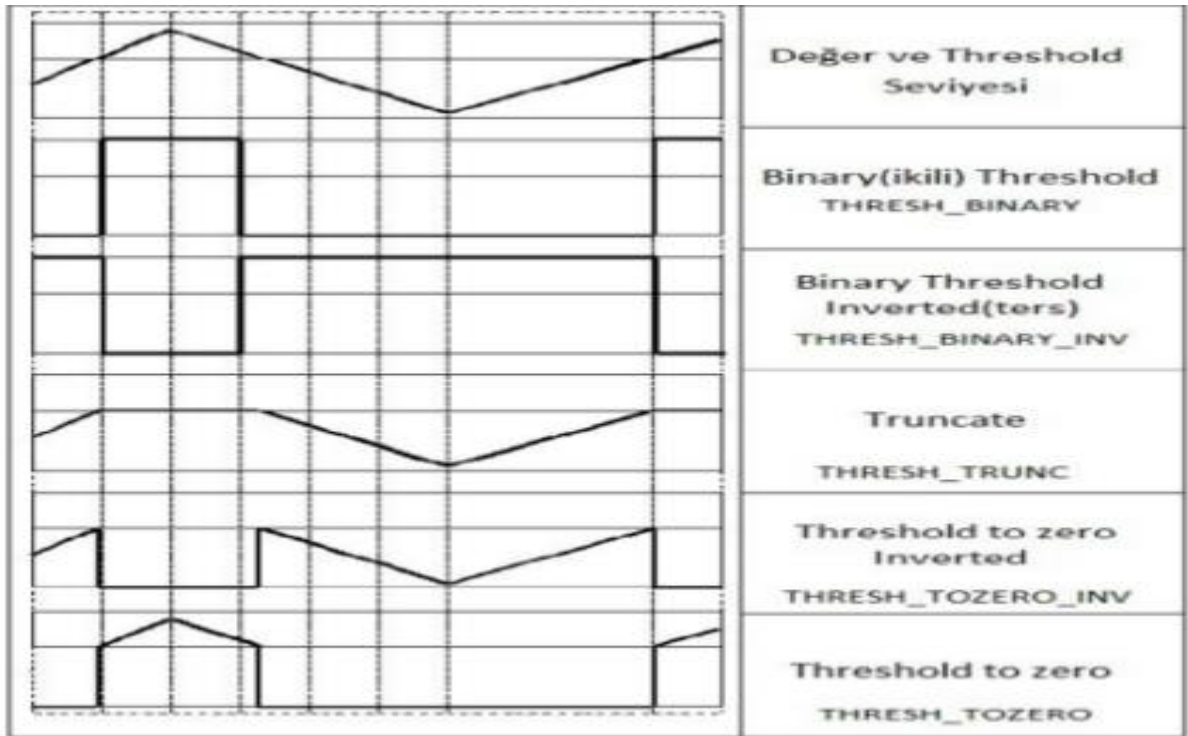


Şekil 2.24: Histogram eşitleme örneği.

2.3.2.4 Thresholding Yöntemi

Thresholding eşikleme olarakta adlandırılmaktadır. Thresholding, belirli bir değere sahip, belirli bir değerin altında veya üstünde kalan pikseller için eşikleme işlemi yapma işlemidir. Eşikleme yöntemi, basitliği ve kolaylığı nedeniyle görüntü bölütlemeye en yaygın kullanılan teknik olarak bilinmektedir (Sahoo, Soltani & Wong, 1988). Eşiklemede amaç, hedef görüntüden arka planı çıkarmaktır. OpenCv kütüphanesi içerisinde 5 temel eşikleme yöntemi vardır. Bunlar;

- THRESH_BINARY
- THRESH_BINARY_INV
- THRESH_TRUNC
- THRESH_TOZERO
- THRESH_TOZERO_INV



Şekil 2.25: Thres holding yöntemleri ile iyileştirilmiş görüntü örnekleri.

Şekil 2.25'te Thres holding yöntemleri ile iyileştirilmiş görüntü örnekleri gösterilmiştir. Thres holding yöntemlerinin OpenCV üzerinde kodlaması da şekil 2.26'da gösterilmiştir.

```
import cv2
image1 = cv2.imread("lenna.png")
ret, thresh1 = cv2.threshold(image1, 127, 255, cv2.THRESH_BINARY)
ret, thresh2 = cv2.threshold(image1, 127, 255, cv2.THRESH_BINARY_INV)
ret, thresh3 = cv2.threshold(image1, 127, 255, cv2.THRESH_TRUNC)
ret, thresh4 = cv2.threshold(image1, 127, 255, cv2.THRESH_TOZERO)
ret, thresh5 = cv2.threshold(image1, 127, 255, cv2.THRESH_TOZERO_INV)
cv2.imwrite('BinaryThreshold.png', thresh1)
cv2.imwrite('BinaryThresholdInverted.png', thresh2)
cv2.imwrite('TruncatedThreshold.png', thresh3)
cv2.imwrite('Thresholdto0.png', thresh4)
cv2.imwrite('Thresholdto0Inverted.png', thresh5)
```

Şekil 2.26: Histogram eşitleme yöntemlerinin OpenCV kodlaması.

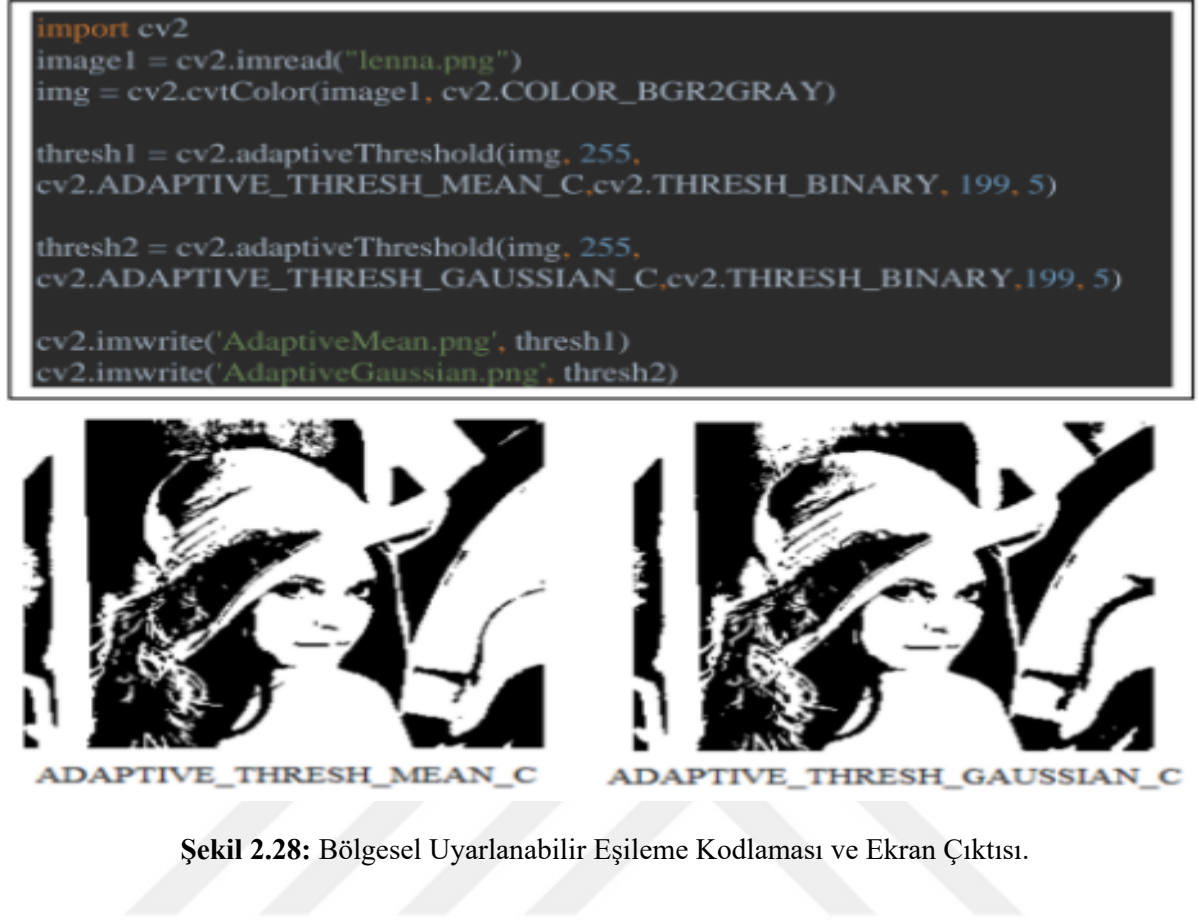


Şekil 2.27: Histogram eşitleme yöntemlerinin OpenCV örneği çıktısı.

Şekil 2.27’de, Şekil 2.26’da verilen histogram eşitleme türlerinin OpenCv kodlarının ekran çıktısı gösterilmiştir. Yukarıda belirtilen eşikleme yöntemlerinden farklı olarak bölgesel uyarlanabilir eşikleme yöntemleri de bulunmaktadır. OpenCv kütüphanesinde bölgesel uyarlanabilir (adaptif) eşikleme yöntemleri şunlardır;

- ADAPTIVE_THRESH_MEAN_C
- ADAPTIVE_THRESH_GUESSIAN_C

Bu yöntemlerden ilki ADAPTIVE_THRESH_MEAN_C, uygulandığı bölgelerin ortalamasını alarak eşik değeri oluşturmaktadır. Yöntemlerden ikincisi ise ADAPTIVE_THRESH_GUESSIAN_C, bloklardaki değerlerin merkez bloğa uzaklığının guess fonksiyonu ile ağırlıklarının hesaplanması ile eşik değeri hesaplamaktadır. Şekil 2.28’de bu iki modelin OpenCv üzerinde uygulanmasının kodları ve ekran çıktısı gösterilmiştir.



Şekil 2.28: Bölgesel Uyarlanabilir Eşileme Kodlaması ve Ekran Çıktısı.

2.3.3. Görüntü Filtreleme

Görüntüler, çoğu zaman görüntü gürültüsü ile bozulabilmektedir. Dijital görüntüler elde edilme, işleme, boyut sıkıştırması, depolanması ve çoğaltılması işlemlerinde gürültülere maruz kalmaktadır. Bu nedenle gürültülü görüntülerin gürültülerden kurtarılması ihtiyacı doğmuştur. Bilgisayar ortamında görüntülerden gürültülerin kaldırılması orijinal görüntünün çıkarılması için filtreleme yöntemlerine başvurulmaktadır.

Dijital görüntülerde filtreleme işlemleri yapılarak, görüntülerdeki gürültüyü gidermek, ayrıntıları ortaya çıkarmak ve bulanıklaştırma gibi birçok işlem yapılabilir (Kim, 1997).

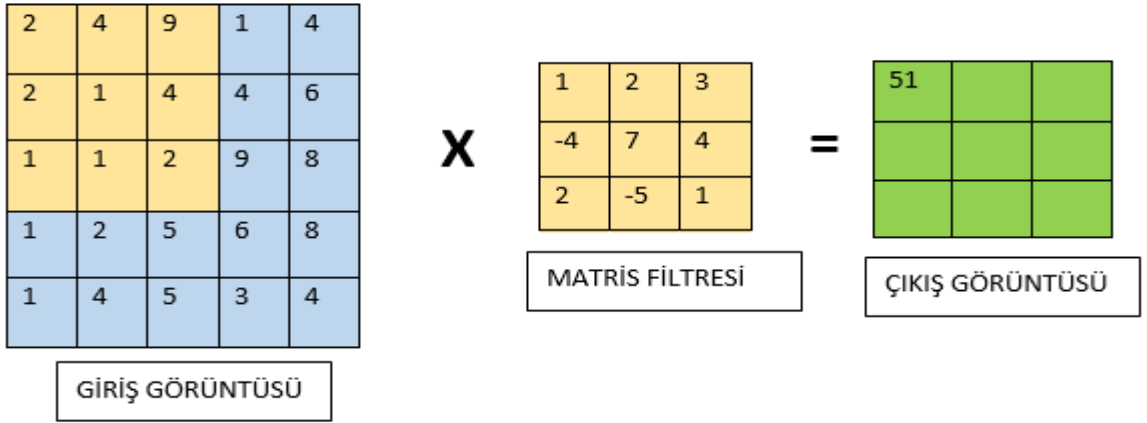
Filtreler 3x3 veya 2x2 boyutunda olabilmektedir. Filtreler görüntünün tamamı üzerinde gezdirilerek çeşitli konvolüsyon işlemleri uygulanmaktadır. Sayısal görüntünün pikseli ile filtre matrisi değeri çarpılarak ortalama alınır ve ortanca eleman belirlenmektedir. Bu şekilde tüm piksellere benzer işlemler uygulanarak filtreleme tamamlanmış olacaktır. Yaygın olarak kullanılan filtreleme işlemleri aşağıdaki gibidir. Bunlar;

- Ortalama (mean) filtresi,
- Orta değer (median) filtresi,
- Sobel kenar çıkarma filtresi,
- Gauss bulanıklaştırma filtresi.

Yukarıda belirtilen filtreleme türleri ve konvolüsyon işlemi alt başlıklar halinde açıklanmış ve örneklendirilmiştir.

2.3.3.1. Konvolüsyon İşlemi

Çekirdek olarak belirlenen bir matrisin hedef görüntü üzerinde gezdirip çarpılması işlemine konvolüsyon denilmektedir. Şekil 2.29'da örnek bir konvolüsyon işlemi gösterilmiştir.



Şekil 2.29: Konvolüsyon matris çarpımı örneği.

Şekil 2.29'da görüldüğü gibi giriş matrisinin üzerinde matris filtresinin konvolüsyon işlemi yapılmıştır. Giriş görüntüsündeki hedef matrisin, matris filtresi ile eşleşen kısımları çarpılıp toplanmış ve yeni çıkış değeri elde edilmiştir. Bu işlem bir görüntüde bütün piksellere uygulanmaktadır.

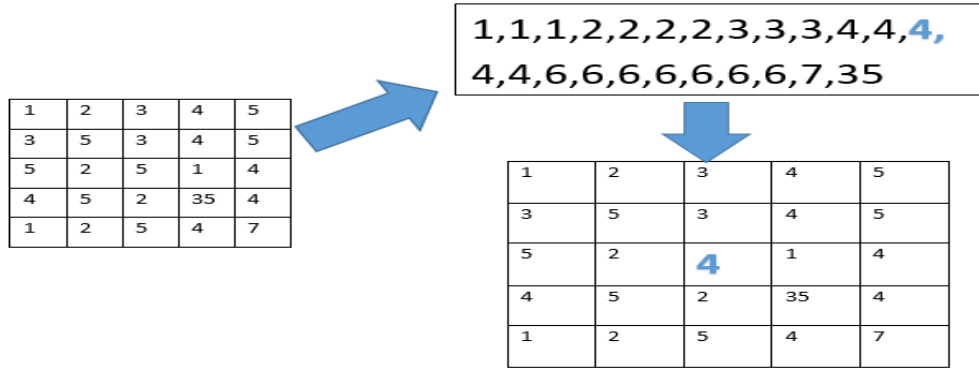
2.3.3.2. Ortalama Filtre

Ortalama filtresi, averaging, yumuşatma veya kutu filtre olarak bilinmekte ve yaygın olarak kullanılmaktadır. Ortalama filtre, görüntüleri yumuşatmak için kullanılan basit bir yöntemdir. Yani görüntülerdeki, piksellerin bir önceden bir sonraya geçişinde sert ve ani geçişleri azaltmaktır. Bu filtre yönteminde de konvolüsyon işlemi gibi yapılmakta ve tüm görüntü

taranmaktadır. Ancak konvolüsyon işleminden farkı konvolüsyonda filtreler çarpılmakta bu işlemde bölünmektedir. Örnek verilecek olursa, 3x3 lük bir görüntü matrisinde 9 kutu bulunmaktadır. Yani burada yumuşatma işlemi yapılacak ise her kutunun değeri 9'a bölünecek anlamına gelmektedir (Aydın, 2021).

2.3.3.3. Orta Değer Medyan Filtre

Orta değer medyan filtresi, sıralama filtresi olarak da bilinmektedir. Görüntülerde gürültü azaltma işlemlerinde başarılı sonuçlar vermektedir. Görseller üzerindeki detayların kaybolmaması için orta değer medyan filtre, ortalama filtreye göre daha çok tercih edilmektedir. Bu filtrede, konvolüsyon işlemi uygulanmamaktadır. Bu filtreleme işleminde, matris üzerindeki değerler büyükten küçüğe veya küçükten büyüğe sıralanır ortanca değer ile yeni matris oluşturulur. Şekil 2.30'da örnek orta değer medyan örneği gösterilmiştir (Aydın, 2021).



Şekil 2.30: Orta Değer Medyan Filtresi Örneği.

3. MALZEME VE YÖNTEM

Bu çalışma, iki aşamalı deneysel çalışmalar neticesinde gerçekleşmektedir. Birinci aşaması görüntülerin derin öğrenme ile analiz edilmesi, ikinci aşaması ise deneysel analizleri yapılan görüntüler kullanılarak görüntünün depolanmasında boyut indirgemesi çalışması yapılmasıdır. Tezin iki aşaması için ayrı malzeme ve yöntemler kullanılmıştır.

Bu bölümde, tezin birinci aşamasında nasıl bir veri kümesi kullanılması gerektiği, veri kümelerinin nasıl elde edildiği, veri kümeleri için nasıl ön işleme çalışmaları yapıldığı ve hangi deneysel yöntemlerle analizlerin yapılması gerektiği açıklanmıştır. Ayrıca bu bölümde tezin ikinci aşamasında, nasıl bir veri kümesi kullanılması gerektiği, veri kümesinin kaynağı ve boyut indirme işlemi için hangi filtreleme yöntemlerinin kullanıldığı açıklanmıştır.

3.1 TEZİN BİRİNCİ AŞAMASI

Bu bölümde tez çalışmasında yapılan derin öğrenme tabanlı görüntü işleme teknikleri için malzeme ve yöntemler açıklanmıştır.

3.1.1 Yöntemler

Bu tez çalışmasının birinci aşamasında, görüntülerde silah, bıçak ve şiddet görüntülerinin sınıflandırılması ve görüntülerde anormal nesne algılanması ve tespitlerinin yapılabilmesi için deneysel çalışmalar yapılmıştır. Bu deneysel çalışmalarda, sınıflandırmalar için ayrı nesne algılanması işlemleri için ayrı yöntemler kullanılmıştır.

3.1.2 Veri Kümeleri

3.1.2.1 Veri Kümelerinin Özellikleri

Tez çalışmasında, güvenlik kamerası görüntüleri üzerinde deneysel çalışmalar yapılmıştır. Güvenlik kamerası görüntülerinden şiddet algılama, silah algılama ve bıçak algılama gibi deneysel çalışmalar yapılmıştır. Görüntü çeşitliliğinin çok olması veri kümelerinin de çeşitli olması anlamına gelmektedir. Çeşitli veri kümelerini seçerken belirli özelliklerde olması gerekliliği gözlemlenmiştir. Bu özellikler;

- Veri kümelerinin, analiz yapılacak görüntülerin belirleyici özelliklerini taşıması,

- Veri kümelerinin çeşitliliğin olması sebebiyle, elde edilen verilerin anonim telif olması ve kişisel bilgileri ve hakları ihlal etmemesi,
- Veri kümelerinin hedef görüntüyü doğru temsil etmesi için yeterli sayıda veri içermesi,
- Veri kümesindeki görüntülerin, sınıflandırma işlemi yapılacak ise sınıflandırmaya uygun olması, görüntü algılama işlemi var ise algılama işlemine uygun olması,
- Çalışmanın güvenlik kamerası görüntüleri üzerinde bir analiz yapabilmesi için görüntü kalitesinin güvenlik kamerası standartlarında ve zorluğunda olması.
- Veri kümelerinin, ön işleme ve düzenleme gibi işlemlerin yapılabilmesine elverişli olması gerekmektedir.

3.1.2.2. Veri Kümelerinin Toplanması

Çalışmada yapılacak deneysel çalışmalara uygun ve yeterince çalışmanın amacını temsil edebilen veri kümeleri toplanmıştır. Veri kümeleri, deneysel çalışmaların amacına uygun olacak şekilde seçilmiştir. Veri kümelerinin kullanılacağı analizler ve toplandığı kaynaklar tablo 3.1’de açıklanmıştır.

Tablo 3.1: Analiz Türleri ve Veri Toplama Kaynakları.

Kaynaklar	SİLAH Algılama	BIÇAK Algılama	ŞİDDET Algılama	SİLAH/BIÇAK Sınıflandırma	ŞİDDET Sınıflandırma
Bilimsel	X	X	X	X	X
Çalışmalar					
YouTUBE		X			X
Kaagle/	X	X	X	X	X
Github					
Google	X		X	X	X
Açık Erişim	X	X	X	X	X
Veri					
tabanları					

Bu çalışmanın deneysel analizleri yapılırken, ülkemizi ve dünyayı saran Covid19 pandemisinin olması sebebiyle kaynak seçiminde kurum ve kuruluşlardan güncel veri görüntüleri alınamamıştır. Çalışmanın işleyişinin aksamaması için kaynak seçiminde, telif ve etik ihlal etmeyecek açık erişimli internet üzerinden erişilebilen kaynaklar olmasına dikkat edilmiştir. Tablo 3.1’de görüldüğü üzere tüm analizlerin verilerinin toplanması için benzer kaynaklar

seçilmiştir. Ancak farklı olarak şiddet algılama anlık ve gerçek zamanlı olması gerektiğinden verilerinin sadece resim görüntüsü değil video görüntüsü olması da gerekmektedir. Bu sebeple şiddet veri kümesinde, YouTube üzerinde yaratıcı birliktelik yani “CC” kodlu eserlerde seçilmiş ve toplanmıştır.

Bu tez çalışmasında kullanılan veri kümeleri, tez çalışması tamamlandıktan sonra internet sayfaları ve çeşitli akademik kaynak paylaşım sitelerinde açık erişimli olarak paylaşılacaktır.

A. Silah/Bıçak Algılanması Veri Kümesi

Silah ve bıçak algılanması tez çalışmasının birinci aşaması yani anlık görüntülerde anormal durumların tespiti için yapılan bir çalışmadır. Silah ve bıçak algılanması çalışmasında hem silah algılanması hem de bıçak algılanması tek bir deneysel çalışma şeklinde yapılmıştır. Yani tek deneysel çalışma ile görüntülerde silah varsa veya bıçak varsa uyarı veren sistem gerçekleştirilmiştir. Bu çalışmada, veri kümesi iki ayrı nesne türü içermektedir. Bunlardan biri bıçak algılanması bir diğeri ise silah algılanması için oluşturulan veri kümesidir. Silah verisinde, farklı ortam şartlarında çekilmiş 3000 adet görüntü bulunmaktadır. Silah görüntüleri, internet ortamı ve güvenlik kamera görüntülerinden oluşmaktadır. Bu çalışmada, kullanılan veri kümesi daha önceleri Olmos vd.(2018), çalışmalarına kaynak veri kümesi olarak kullanılmıştır. Bıçak veri kümesi ise, 2078 adet farklı ortam şartlarında çekilmiş görüntülerden oluşturulmuştur. Bu görüntülerde en az bir bıçak görüntüsü bulunmaktadır. Bu veri kümesinin görüntüleri ise, YouTube videoları ve güvenlik kameraları görüntülerinden toplanmıştır. Bu görüntülerin seçiminde bazı özelliklerin olması tercih edilmiştir. Bu özellikler şunlardır;

- Nesnelerinden çeşitli tür, şekil, boyut ve farklı malzemeler ile yapılmış olması,
- Görüntülerin yakın ve uzak çekim şartlarında elde edilmiş olması,
- Bıçak görüntülerinin insan eliyle tutulmuş veya işaret edilmiş olması,
- Bıçak ve silah görüntülerinin çeldiricilerinin olması,
- Görüntülerin iç ve dış mekan şartlarında çekilmiş olmasına dikkat edilmiştir.

Bıçak veri kümesindeki görüntüler Castillo vd. (2019) çalışmasında kaynak veri kümesi olarak kullanılmıştır. Bu veri kümelerine ait örnek görüntüler şekil 3.1 ve şekil 3.2’de gösterilmiştir.



Şekil 3.1: Bıçak veri kümesi örnek görüntüleri.



Şekil 3.2: Silah veri kümesi örnek görüntüleri.

Bıçak ve silah veri kümeleri birleştirilmiş ve çeşitli ön işleme adımları uygulanmıştır. Görüntü sayıları görüntü açıları değiştirilerek artırılmıştır. Son olarak, veri kümesi olarak kullanılan görüntü sayısı 5078 adet görüntüden 9253 adet görüntüye çıkarılmıştır.

B. Şiddet Algılama Veri Kümesi

Bu çalışma tez çalışmasının birinci aşama analizleri için yapılmıştır. Şiddet algılama çalışması anlık görüntülerde anlık olarak şiddet görüntülerinden şiddet bölümlerini tespit etmektedir. Bu çalışma için akademik çalışmalardan, internet sayfalarından ve açık erişim veri tabanlarından veri toplanması işlemi yapılmıştır. Şiddet görüntüsü içeren 3333 adet görüntü elde edilmiştir. Elde edilen görüntülerin seçiminde bazı özellikler aranmıştır. Bu özellikler şunlardır;

- Görüntülerin bir kısmının seçiciliğinin zor olması,
- Görüntülerin güvenlik kamerası standartlarında elde edilmiş olması,
- Görüntülerin içerisinde sadece şiddet ortamındaki şiddet eylemine karışmış insanların değil, ayrıca şiddet olaylarına ayırma rolüyle veya hiç karışmamış insanların olması,
- Veri kümesi için sadece şiddet içeren görüntülerden seçilmeyip ayrıca şiddet içermeyen görüntülerde seçilmesi,
- Şiddet görüntülerinin bir akış halinde olması, yani sadece şiddeti gösteren bir görüntü değil ardışık şiddet resim karelerinin olması gibi özellikler tercih edilmiştir.

Şiddet veri kümesi için farklı paylaşım ortamlarından elde edilen görüntüler titizlikle birleştirilmiş bir bütün haline getirilmiş ve bir bütün olarak sistem eğitilmiştir. Şiddet algılama veri setine ait örnek görüntüler şekil 3.3'te verilmiştir.



(a)



(b)

Şekil 3.3: Şiddet veri kümesi örnek görüntüler, (a) Şiddet içeren görüntüler, (b) Şiddet içermeyen görüntüler.

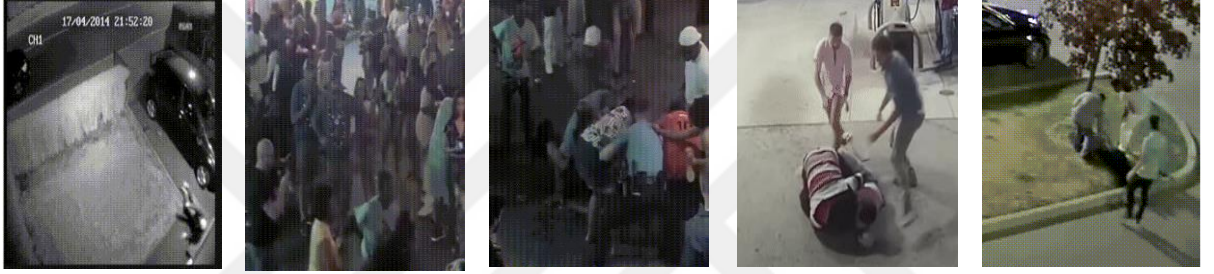
Şiddet algılama veri setinde, görüntüler üzerinde ön işleme adımları izlenmiş ve augmentasyon işlemi yapılmıştır. Görüntü sayısı, görüntü açılarının değiştirilmesi ile 8650 görüntüye çıkarılmıştır.

C. Şiddet Sınıflandırılması Veri Kümesi

Bu veri kümesi, tez çalışmasının ikinci aşamasında boyut indirgeme uygulaması için kullanılmak üzere oluşturulmuştur. Bu veri kümesinin hazırlanmasında üç genel veri kümesi kullanılmış ve üç veri kümesi için ayrı ayrı deneysel çalışma yapılmıştır. Bu genel veri kümeleri daha önceleri akademik çalışmaların veri aşamalarında kaynak görevi görmüştür. Veri kümelerinin isimleri sırasıyla, The Crown Violence(Hassner vd., 2012), The Fight Detection Surv Master veri kümesi (Aktı vd., 2019)ve Movie Violence (Serrano Gracia vd., 2015) veri kümesi şeklindedir.

Bu veri kümeleri, iki sınıflı olarak oluşturulmuştur. Bu sınıflar, şiddet içeren sınıf ve şiddet içermeyen sınıf verileri ile oluşturulmuştur.

The Crown Violence: Bu veri kümesi, kalabalık ortamlardaki şiddet içeren ve içermeyen görüntüleri kapsamaktadır. Görüntüler karşılaştırmalar içermektedir. Hem şiddet içeren ve içermeyen görüntülerin sınıflandırılması hem de şiddet içeren ve içermeyen ortamların sınıflandırılması için oluşturulmuştur. Dünyanın farklı yerlerinden farklı şartlarda görüntüler içermektedir. Bu veri setinin görüntüleri, YouTube videolarından elde edilmiştir. Veri setindeki, en kısa görüntü 1,04 saniye, en uzun görüntü ise 6,52 saniyelik görüntülerden oluşmakta ve görüntülerin ortalama süresi ise 3,6 saniye olmak üzere 246 görüntüden oluşmaktadır. Veri kümesine ait örnek görüntüler şekil 3.4'te gösterilmiştir (Hassner vd., 2012).



Şekil 3.4: The crowd violence veri kümesi örnek görüntüler.

The Fight Detection Surv: Bu veri kümesi, kavga görüntüleri içeren görüntü videolarından oluşmaktadır. Veri kümesi, YouTube videolarından toparlanmıştır. Ayrıca bazı internet sayfalarında açık kaynak olarak temin edilebilen güvenlik kameraları kayıtlarından alınan kavga içermeyen görüntülerde mevcuttur. Bu veri kümesi, 150 kavga içeren ve 150 kavga içermeyen 300 adet görüntüden oluşmaktadır. Görüntüler ortalama 2 saniye uzunluğundadır.

Bu veri kümesi hazırlanırken, hareketsiz arka plana sahip görüntüler tercih edilmiştir. Ayrıca nesnelerle vurma, tekmeleme, yumruk atma ve güreş gibi çeşitli kavga senaryoları içermektedir (Aktı vd., 2019).

Şekil 3.5'te veri kümesine ait örnek görüntüler gösterilmiştir.



Şekil 3.5: The fight detection surv veri kümesi örnek görüntüler.

Movie Violence: Filmlerden derlenen veri kümesidir. Filmlerde karşılaşılan şiddet içeren görüntülerden toparlanmıştır. Farklı sahnelerde çekilmiş 200 kavga görüntüsü içermektedir. Ayrıca şiddet içermeyen görüntü sahneleri de veri kümesine dâhil edilmiştir (Serrano Gracia vd., 2015). Şekil 3.6’da veri kümesine ait örnek görüntüler verilmiştir.



Şekil 3.6: Movie violence veri kümesi örnek görüntüler.

D. Silah/Bıçak Sınıflandırılması Veri Kümesi

Silah ve bıçak sınıflandırılması çalışmanın ikinci aşamasında yani boyut indirgeme aşamasında kullanılmak amacıyla gerçekleştirilen bir çalışmadır. Bu sınıflandırma işlemi, boyut indirgeme aşamasında görüntülerin boyutunu düşürme işleminde sayaç artırmak ve not defterine bu bilginin işlenebilmesi için çok önemlidir.

Bu çalışmada, literatürde daha önceleri birçok çalışmaya tekrar tekrar kaynaklık etmiş veri kümelerinden (Castillo vd., 2019; Olmos vd., 2018, 2019; Perez vd., 2019) ve orijinal internet tarayıcıları ve video sayfalarından görüntüler toplanarak amaca uygun bir veri kümesi oluşturulmuştur (Grega vd., 2016; Mاتیolański vd., 2016). Bu veri kümesinde 9500 adet bıçak görüntüsü, 3500 silah görüntüsü ve 3000 adet normal görüntü içeren 16000 görüntü bulunmaktadır. Bu veri setindeki sınıflara ait örnek görüntüler şekil 3.7’de verilmiştir.



Şekil 3.7: Silah bıçak algılanması veri kümesi örnek, (a) silah verileri, (b) bıçak verileri, (c) normal görüntüler.

Şekil 3.7’de görüldüğü üzere veri kümesi üç sınıftan meydana gelmektedir. A sınıfı silah görüntüleri içeren sınıf, b sınıfı bıçak görüntüleri içeren sınıf ve c sınıfı ise normal herhangi bir silah ve bıçak görüntüsü içermeyen sınıftır. Görüntülerin seçiminde, görüntülerin her şartta

temsil etmesi gereken sınıfı doğru bir şekilde ortaya çıkarması için titizlikle araştırma yapmıştır.

Veri setindeki görüntülerin, uygulanacak deneysel modele uygun olabilmesi ve sağlıklı sonuç verebilmesi için boyut ve augmentasyon gibi bir takım ön işleme adımları uygulanmıştır.

3.2 TEZİN İKİNCİ AŞAMASI

Tezin bu aşamasında, güvenlik kamerası videolarının depolanması işlemlerinde boyut indirgeme çalışması yapılmaktadır. Boyut indirgeme, işlemlerinin gerçekleşmesi için yeni bir boyut indirgeme modeli, derin öğrenme tabanlı görüntü içerik analizleri ve indirgeme işlemlerinin yapılabilmesi için güvenlik kamerası görüntülerine ihtiyaç vardır.

3.2.1 Boyut İndirgeme Modeli

Bu tez çalışmasında amaç, güvenlik kameralarında derin öğrenme yaklaşımları ile anormal durum tespiti yapılması ve görüntülerin depolanmasında boyut indirgemesi sisteminin gerçekleştirilmesidir. Bu sistemin herhangi örneğinin bulunmaması sebebiyle tamamen özgün bir model oluşturulmuştur.

Modeli en basit haliyle ifade etmek gerekirse, güvenlik kamerası görüntülerinden alınan görüntülerde silah, bıçak ve şiddet içeren herhangi bir durumun görülmesi durumunda görüntülerin koruma altına alınması ve herhangi bir olay ya da hareketin görülmediği durumlarda boyut indirgeme işleminin yapılmasıdır. Boyut indirgeme aşamasında, görüntü kayıtları yani güvenlik kamerası video kayıtları resim çerçevelerine (frame) ayrılmakta görüntüler üzerinde karşılaştırma yapılmaktadır. Karşılaştırma işlemlerinde, çeşitli görüntü filtreleme ve görüntü işleme yöntemleri kullanılmakta ve eğer herhangi anormal durum yoksa birbirine eşit olan görüntülerden ikinci silinmekte ve yerine ilk görüntünün gölgesi yansıtılmaktadır.

Boyut indirgeme işlemlerinde sınıf analizlerine ihtiyaç duyulmaktadır. Bu sınıflar;

1. **Silah Görüntü Sınıfı:** Güvenlik kamerası görüntülerinde eğer herhangi bir silah içeren görüntü varsa bu görüntü silah görüntü sınıfına alınmaktadır. Bu sınıf içerisine herhangi bir veri düştüğü zaman sistem yöneticisi uyarılmakta ve mail olarak bilgi gönderilmektedir. Ayrıca bu sınıf içerisinde yer alan herhangi bir verinin bulunduğu

- görüntü kümesinde güvenlik zafiyeti olmaması için boyut indirgemesi çalışması yapılmamaktadır. Bu verilerin bulunduğu görüntü kümesinden sonraki görüntü kümesinde tekrar boyut indirme çalışması devam etmektedir.
2. **Bıçak Görüntü Sınıfı:** Bu görüntü sınıfı güvenlik kamera sistemlerinden gözlemlenen bıçak verilerinden oluşmaktadır. Bu sınıfa veriler, anlık olarak görüntülerin analizlerinin yapılması ile aktarılmaktadır. Bu sınıfta eğer herhangi bir bıçak verisi varsa bu görüntü kümesinde de herhangi bir boyut indirme çalışması yapılmamakta ve bir sonraki görüntü kümesine geçilmektedir. Bu sınıfın verileri de tespit edildiği zaman sistem yöneticisine anlık bilgi verilmekte ve mail olarak rapor gönderilmektedir.
 3. **Şiddet Görüntü Sınıfı:** Bu görüntü sınıfı da güvenlik kamera sistemlerinde gözlemlenen bir şiddet olayının görüntülerinden oluşmaktadır. Bu sınıfın verileri güvenlik kamera sistemlerinin kayıtlarında herhangi bir şiddet sahnesi görüldüğü zaman, görüntünün bulunduğu kümede boyut indirme işlemi yapılmamaktadır. Bir sonraki görüntü kümesi ile boyut indirme işlemi devam etmektedir. Bu görüntü sınıfına da yeni veri geldiği zaman diğer sınıflar gibi sistem yöneticisine anlık bilgi verilmekte ve mail olarak rapor gönderilmektedir.
 4. **Hareketli Görüntü Sınıfı:** Bu görüntü sınıfı, diğer görüntü sınıflarından farklı olarak silah, bıçak ve şiddet görüntüsü içermeyen bütün hareketli görüntüleri içermektedir. Bu sınıfın içerisinde hareketli insan, araç, havyan, nesne ve aydınlatma değişikliği verileri bulunmaktadır. Görüntü verileri anlık olarak güvenlik kamerası kayıtlarından alınmaktadır. Bu görüntülerin bulunduğu veri kümelerinde de diğer sınıflar gibi boyut indirme çalışması yapılmamaktadır. Ancak bu görüntü kümesi verileri ilgili sistem yöneticisine herhangi bir bilgi veya mail raporu atılmamaktadır.
 5. **Hareketsiz Normal Görüntü Sınıfı:** Bu görüntü sınıfı, boyut indirme işlemleri için ana görüntü sınıfı olarak kullanılmaktadır. Bu görüntü sınıfında boyut indirme işlemleri yapılmaktadır. Bu görüntü sınıfının verilerini, hareketsiz görüntüler oluşturmaktadır. Yani hiçbir hareketin olmadığı ve özellikle gece çekimlerinde saatlerce hiç hareketin olmadığı görüntülerdir. Bu görüntüler üzerinde bir güvenlik zafiyet olmaması için bilgisayarlı görü ile birebir aynı görüntüler seçilir ve aynı görüntülerden ikincisi sistem tarafından silinir ve aynı olan ilk görüntünün gölgesi aktarılır. Bu sayede görüntünün bütünlüğünde hiçbir bozulma olmaz ve ikinci aynı olan görüntünün boyutu kadar görüntü kazancı olmaktadır. Görüntüler aynı olduğu içinde herhangi bir güvenlik

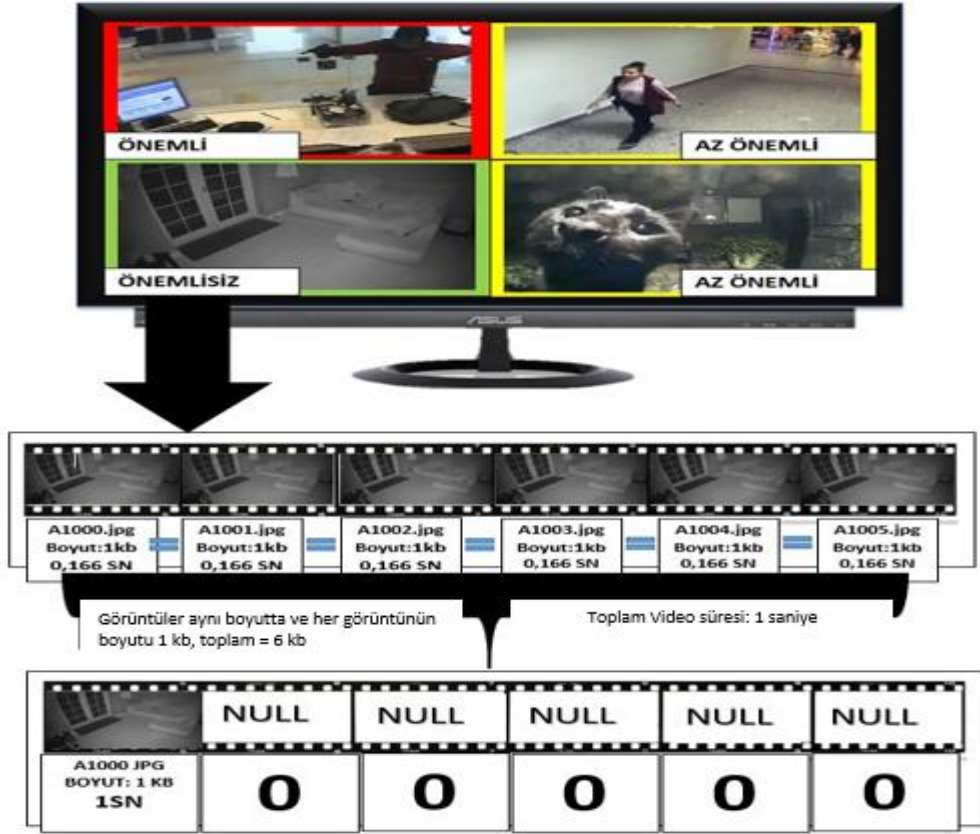
zafiyeti ortaya çıkmamaktadır. Tablo 3.2’de boyut indirgeme modeli sınıfları gösterilmiştir.

Boyut indirgeme modelinin sınıflarının sistem için önemi ifade edilirse;

Tablo 3.2: Boyut indirgeme sınıfları.

Önem Derecesi	Görüntü Sınıfı
Önemli	Silah, bıçak ve şiddet görüntüleri
Az Önemli	Silah, bıçak ve şiddet haricindeki her hareketli olay görüntüleri
Önemsiz	Hiçbir hareketli bir olayın olmadığı görüntüler

Boyut indirgeme modelinin örnek şematik gösterimi şekil 3.8’da gösterilmiştir.



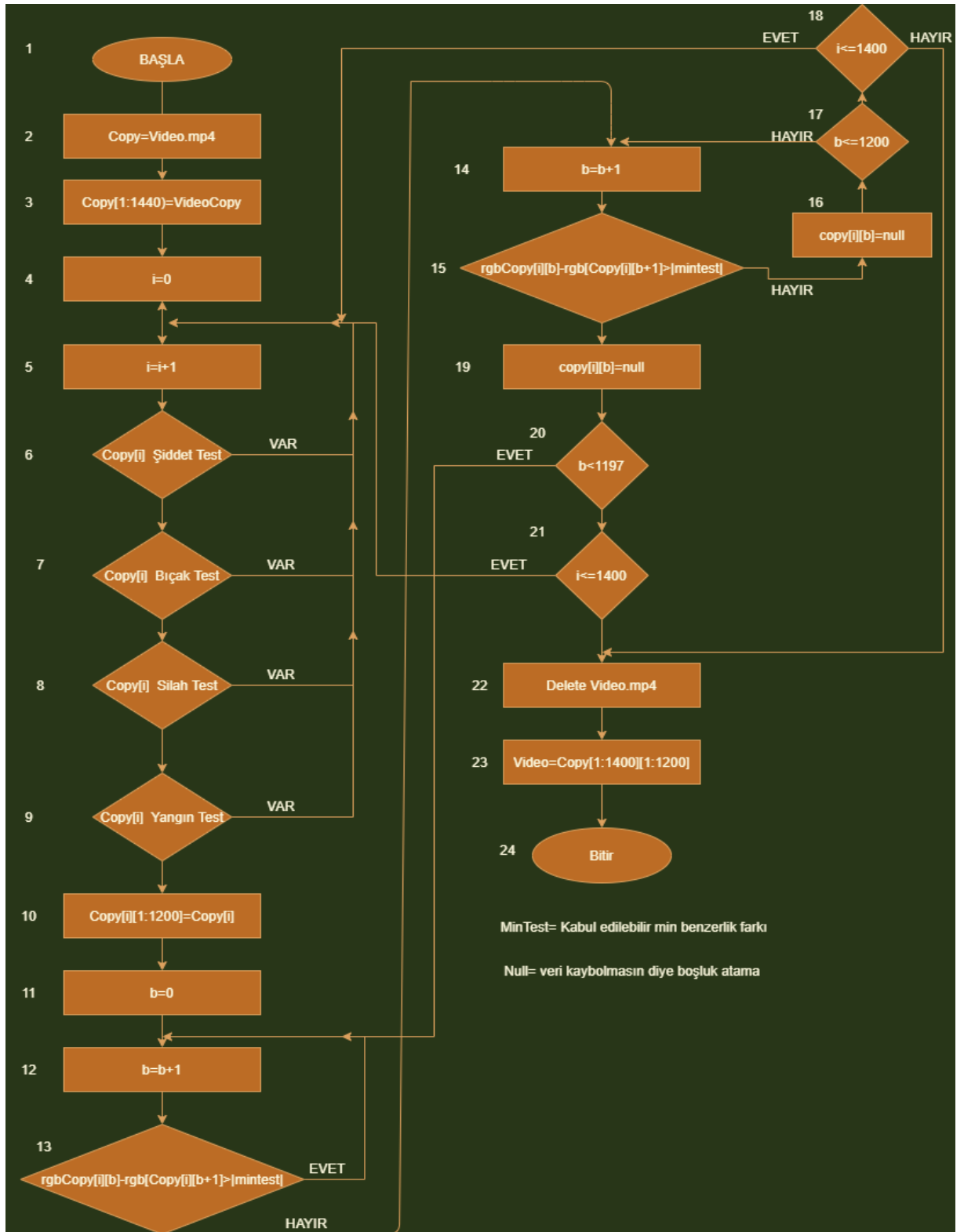
Şekil 3.8: Boyut indirgeme modeli örnek şeması.

Şekil 3.8’de gösterildiği gibi boyut indirgeme işlemi sadece önemsiz görüntüde yapılmıştır. Örnek şemada önemsiz görüntüde bir odanın gece çekimi gösterilmiştir. Herhangi bir anda

alınan bir saniyelik görüntü resim çerçevelerine (frame) ayrılmıştır. Bir biri ile aynı olan görüntülerden birinci görüntüye hiçbir işlem yapılmamış ve kendisinden sonraki beş görüntüde boyut indirgeme işlemi yapılmıştır. Bu işlem sayesinde örnek olarak bir saniyelik görüntüyü 6 kilobyte olarak varsayarsak 5 kilobyte'lık bir boyut alanı kazancı olduğunu kabul edebiliriz.

Boyut indirgeme modeli özgün olarak oluşturulmuştur. Bu modelin işleyişi için öncelik olarak algoritma taslağı oluşturulmuş ve sürekli tekrar deneysel çalışmaları yapılarak modelin kesinleşmiş algoritması çıkarılmıştır.





Şekil 3.9: Boyut indirgeme modeli algoritması.

Şekil 3.9’da boyut indirgeme için oluşturulan modeli gösterilmiştir. Bu model ve algoritma, bu tez çalışması için özgün olarak tasarlanmıştır. Şekil 3.9’da gösterilen algoritma üzerinde algoritma adımları detaylı olarak gösterilmiştir. Adımların açıklanmasında kolaylık olması için adımı temsil eden rakamlar verilmiş ve aşağıda bu adımlar açıklanmıştır.

Algoritma adımları;

1. Bu adım sistemin ilk başlama adımıdır. Bu adımın başlaması için bu çalışmaya özgü bir arayüz tasarlanmıştır. Bu arayüz üzerinden *Boyut İndirgeme İşlemi Başlat* komutunun seçilmesi ile boyut indirgeme işlemi başlamaktadır.
2. Boyut indirgeme işleminin yapılabilmesi için daha önceleri güvenlik kamerası vb cihazlardan bir kayıt işleminin gerçekleşmiş olması gerekmektedir. Bu adımda daha önceleri sistem tarafından edinilmiş olan kayıtların bir kopyası alınmaktadır. Bu adımın amacı, sistem çalışırken herhangi bir nedenle sistemin kapatılması, donması veya elektrik kesintisinde işlem yapılan kayıtların kaybedilmesinin önüne geçmektir.
3. Boyut indirgeme işleminde daha önceleri alınmış olan güvenlik kamerası kayıtları 24 saatlik üzerinden olacaktır. Bu görüntüleri bu şekilde bir bütün olarak kullanmak zor olacağından görüntüyü, bütün olarak değilde dakikalık küçük parçalara ayrılarak kullanılmıştır. Bu adımda görüntüler 24 saatlik boyuttan 1 dakikalık boyutlara dönüştürülmüştür. Ancak bu işlem için görüntünün direk olarak 1440 parçaya bölünmesi görüntülerin sırasında ve bütünlüğünde bozukluk oluşturacağı için 1440 boyutunda tek boyutlu bir matris dizi yardımıyla görüntüler bir seri şeklinde tekrar birleştirilmiştir.
4. Bu adım boyut indirgeme işlemlerinde, görüntü analizlerinin başlatılması için oluşturulan sayaç kurma adımıdır. Bu adımda bir önceki adımda oluşturulan 1440 boyutundaki matris dizisi biçiminde olan video görüntülerinin sırasıyla çağırılması için sayaç kurulduğu adımdır.
5. Bu adımda bir önceki adımda kurulan sayacın arttırılması işlemi yapılmaktadır. Boyut indirgeme işlemlerinde analiz videonun birinci dakikasında yapılmamaktadır. Bu işleme, deneysel çalışmalar sonunda sistem karmaşıklığı olmaması için karar verilmiştir.
6. Bu adımda, sayacın gösterdiği değerlere denk gelen *copy[sayaç değeri]* sırasındaki görüntü analiz edilip herhangi bir şiddet olayının olup olmadığına bakılır. Eğer şiddet

- var ise, sistem günlüğüne şiddet olayının detayı kaydedilir ve boyut indirgeme işlemi durdurulup sayaç adımına geri dönülür. Ancak şiddet yoksa görüntüde, bir sonraki algoritma görevine geçilir.
7. Bu adımda bir önceki adıma benzer görevler yapılmaktadır. Sayacın gösterdiği adıma denk gelen *copy[sayaç değeri]* görüntüsünde analizler neticesinde bir silah olayı tespit edilirse, olay detayı bilgisi video sistem günlüğüne kaydedilir ve boyut indirgeme işlemi durdurulur tekrar sayaç artırma adımı olan 5. adıma geri dönülür. Ancak görüntüde bir silah olayı yoksa bir sonraki adıma geçilir.
 8. Bu adımda, 6. ve 7. Adımla benzer görevler yapılmaktadır. Bu adımda, *copy[sayaç değeri]* görüntüsünde bir bıçak olayı olup olmadığı kontrol edilir. Eğer bir bıçak olayı varsa olay detayı video sistem günlüğüne kaydedilir, boyut indirgeme işlemi durdurulur ve sistem sayaç artırım adımına geri döndürülür. Ancak bir bıçak olayı algılanmazsa bir sonraki adıma geçilir.
 9. Bu adım tercihe bağlı olarak opsiyoneldir. Bu sistemiN kullanılacağı alanlarda yangın alarmı sistemi yoksa sisteme eklenebilir. Bu adım eklenirse işleyişi 6.,7. ve 8. Adımla benzer şekilde olacaktır.
 10. Boyut indirgeme işlemi, bir biriyle aynı olduğu kabul edilen ardışık video görüntülerindeki resim karelerinden ikinci görüntünün silinip birinci görüntünün gölgesinin aktarılması işlemine dayanmaktadır. Bu adımda da 1 dakikalık olan video görüntüleri, güvenlik kamerası kayıtlarında saniyede 20 resim karesi boyutu birimine dönüştürülür ve her 1 dakikalık görüntü 1200 resim karesine çevrilir. Bu resim kareleri de iki boyutlu *copy[sayacdeğeri,1200]* birimlik matrise kaydedilir.
 11. Bu adımda matris üzerinde kayıtlı resim karelerini çağırabilmek için ikinci sayaç kurulumu işlemi yapılmaktadır.
 12. Bu adımda, bir önceki adımda kurulmuş olan matrisin değerinin bir arttırılması işlemi yapılmaktadır. Resim karelerinde de ilk kare üzerinde işlem yapılmamaktadır.
 13. Bu adımda iki resim karesi arasında bir benzerlik olup olmadığı kontrol edilir. Sayaç değerine göre alınmış olan resim karesi ile bir sonraki resim karesi arasındaki benzerlik kabul edilebilir değerde ise boyut indirgeme işlemine geçilmektedir. Bu adımda görüntü işleme ve görüntü filtreleme başlığı altında belirtilen yöntemler uygulanmaktadır. Bu yöntemler sonucunda iki resimden ilki, ikinci resimden çıkarılır ve görüntülerin birbirine eşit kabul edeceğimiz eşik değerinden (*mintest*) küçük ise boyut indirgeme

işlemine devam edilir. Eğer resim farkları *mintest* değerinden küçük çıkmaz ise 11. adımda sayaç arttırma işlemine geri dönülür ve bir sonraki resim karesine geçilir. Mintest değeri, 0.001 ila 0.005 değerleri arasında olmaktadır. Bu değer 0.001 olması çok hassas ölçüm yapılacağı, 0.005 olması en az hassas ölçüm yapılacağı anlamına gelmektedir. Bu değer sistem yöneticisi tarafından, bu uygulama için oluşturulmuş boyut indirgeme arayüzünden seçilmektedir. Örneğin kullanıcı çok hassas ölçüm seçeneğini seçerse sistem, iki resim arasındaki farkın 0.001 değerinden küçük olması durumunda boyut indirgemesi işlemi yapacaktır.

14. Bu adımda bir önceki adımdan gelen benzer kabul edilen iki görüntüden ikincisiyle, kendisinden sonra gelen resmin karşılaştırılmasının yapılabilmesi için burada da bir sayaç kurma işlemi yapılır.
15. Bu adımda benzer kabul edilen iki görüntüden ikinci ile kendisinden sonra gelen görüntülerin farklarına bakılır, yine fark mintest değerinden küçük ise diğer görüntüler üzerinde de boyut indirgeme işlemi 16. adıma geçilerek devam edilir. Ancak fark mintest değerinden büyük çıkarsa sadece ikinci görüntü üzerinde boyut indirgeme işlemi doğrudan 19. adıma geçilerek devam edilir.
16. Bu adımda, 15.adımda ikinci görüntü ile kendisinden sonra gelen görüntüler arasında benzerlik olması durumunda devam edilir. İkinci görüntü silinir ve bir sonraki adıma geçilir.
17. Bu adımda sayaç kontrolü yapılır. İşlem akışının devam edebilmesi için 1200. resim karesinin dışına çıkılmaması gerekir. Eğer 1200. resim karesi sayacında işlem yapılıyorsa boyut indirgeme işlemi burada durdurulur ve 18. adıma geçilir. Ancak 1200. Resim karesinde işlem yapılmıyorsa 14. Adıma dönülerek boyut indirgeme işlemine devam edilir. 14. Adımdaki sayacın görevi artık üçüncü resim karesini ikinci resim karesiymiş gibi işlem yapılmasının sağlanmasıdır. Üçüncü resim karesi 16. adımda dördüncü resim ile karşılaştırılır benzerlik varsa bir döngü şeklinde 17. adıma yoksa 19.adıma geçilerek işleme devam edilir.
18. Bu adımda görüntü kayıt sayacı kontrolü yapılır. Eğer sistem 1400. görüntü kayıt sırasında ise boyut indirgeme işlemi çalıştırılmadan ve 22. adıma geçilip boyut indirgeme işlemi tamamlanır. Ancak görüntü sayacı 1400'den küçük ise sistem 5. adıma gidilip tekrar buradan sırasıyla çalıştırılır.

19. Bu adımda, ikinci görüntü ile üçüncü görüntü arasında benzerlik olmaması durumunda ikinci görüntü silinir ve bir sonraki adıma devam edilir.
20. Bu adımda, sayaç kontrolü yapılır. Resim karesi sayacı 1200 değerine ulaşmamış ise sistem 12.adıma gidip tekrar başlatılır. Ancak resim karesi sayacı 1200 değerine ulaşmış ise görüntü kayıt sayacına bakılmak üzere bir sonraki adıma geçilir.
21. Bu adımda, görüntü kayıt sayacı kontrol edilir. Eğer sayaç değeri 1400 değerinden küçük ise 5. adıma dönülerek tekrar sistem çalıştırılır. Ancak sayaç değeri 1400 değerine eşit ise sistem tamamlanmak üzere bir sonraki adımla devam eder.
22. Bu adım, sistemin boyut indirgeme işlemini tamamladığını işaret etmektedir. Bu adımda 2. adımda kopyası alınan video görüntüsünün orijinal halinin silinmesi işlemi yapılır.
23. Bu adımda, kopya video görüntüsünün parçaları birleştirilip tekrar video haline getirilmesi işlemi yapılmaktadır. Bu birleştirilen kopya görüntüler sistemin orijinal görüntüsü olarak kabul edilip depolama birimine kaydedilir.
24. Sistem görevini tamamladığı için sistem kapatılır.

4. BULGULAR

Bu tez çalışması, iki bölümden oluşmaktadır. Birinci bölümde anlık güvenlik kamerası görüntülerinde anormal durumların tespiti çalışması, ikinci bölümünde ise güvenlik kamerası aracılığı ile alınmış olan video kayıtlarında normal durum sınıflandırılması ve boyut indirgemesi çalışması yapılmıştır.

Bu bölümde, tez çalışmasının birinci bölüm çalışmaları ile ikinci bölüm çalışmalarının bulguları paylaşılmıştır.

4.1 BİRİNCİ BÖLÜM BULGULARI

4.1.1 Anlık Görüntülerde Silah ve Bıçak Algılanması Çalışması Bulguları

Bu çalışmada, güvenlik kamerası görüntülerinden otomatik silah ve bıçak algılanması için literatürdeki diğer çalışmalardan farklı olarak popüler ve başarılı sonuçlar veren daha önceden eğitilmiş model olan YOLO modelleri kullanılmış ve sonuçları karşılaştırılmıştır. YOLO modellerinden YOLOv4, YOLOv5, YOLOX ve YOLOR modelleri kullanılmıştır.

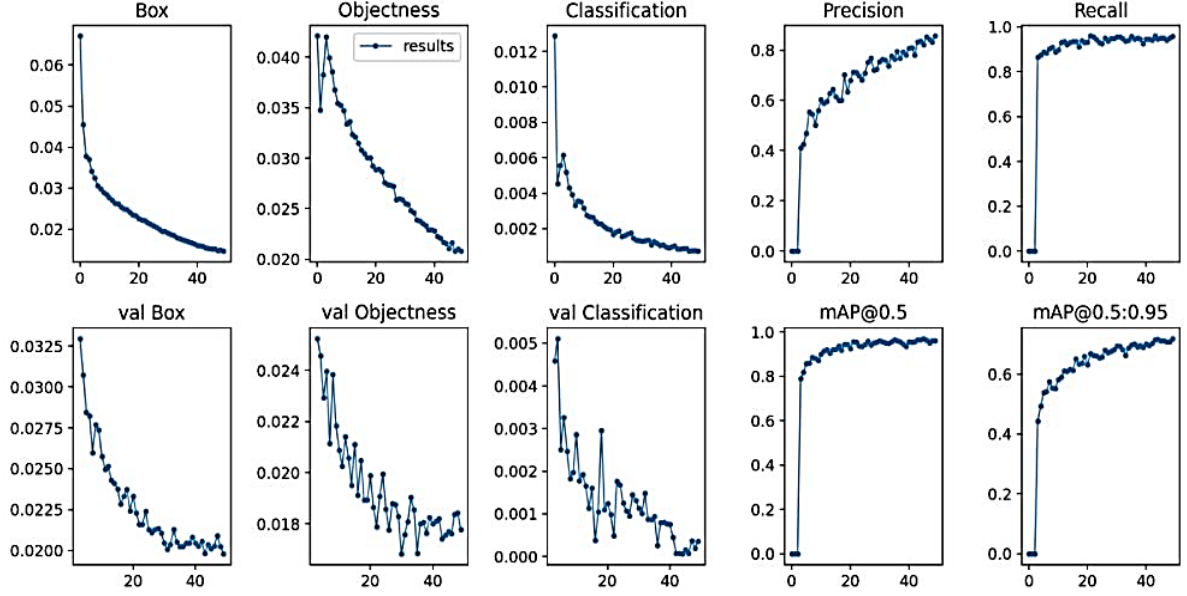
Karşılaştırma işleminin sağlıklı ve güvenilir olması için benzer Epoch, Batch ve Image Size değerleri seçilmiştir. Tablo 4.1’de çalışmada elde edilmiş karşılaştırmalı yüzdesel skorlar görülmektedir.

Tablo 4.1: Çalışmada elde edilmiş karşılaştırmalı yüzdesel skorlar.

MODEL	BATCH	EPOCH	Image Size	Recall	Precisio	Map@0,5
YOLOv4	8	50	416x416	0,91	0,90	0,9412
YOLOv5	8	50	416x416	0,816	0,86	0,864
YOLOX	8	50	416x416	0,89	0,894	0,9069
YOLOR	8	50	416x416	0,958	0,967	0,976

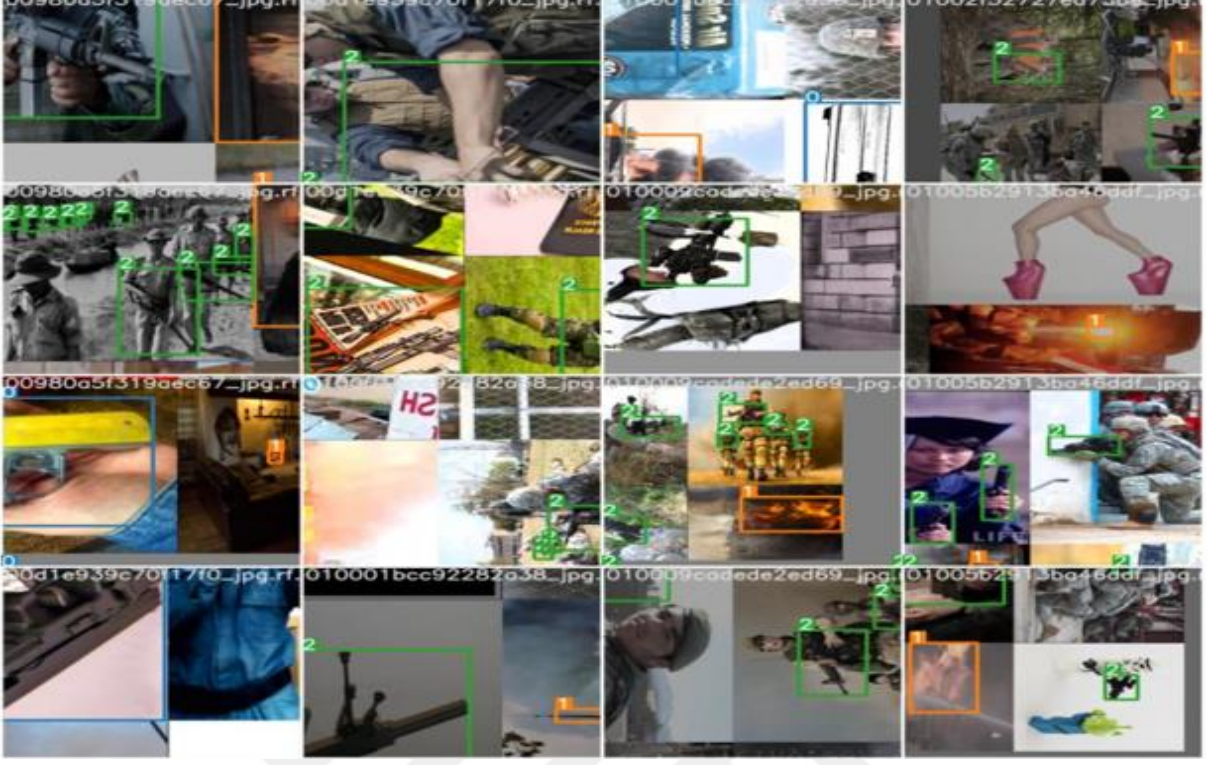
Tablo 4.1’de incelendiğinde, çalışmada uygulanan modeller arasında en başarılı sonuçların YOLOR algoritmasının vermiş olduğu gözlemlenmiştir. YOLOR modeli çalışma süresi olarak ta diğer modellere göre çok daha kısa sürede tamamlanmıştır. YOLOR modeline göre silah ve

bıçak tespiti işlemi sonuçları recall değeri 0,958, precision değeri 0,967 ve nesnelerin kesişiminin ortalama hassasiyet değeri olan $\text{Map}@0,5$ 0,976 olarak hesaplanmıştır. YOLOR modelinde hesaplanan diğer değerlerde grafik halinde sonuçları da aşağıda şekil 4.1’de verilmiştir.



Şekil 4.1: Deneysel çalışma analizlerinin grafiksel gösterimi.

Deneysel çalışma neticesinde, elde edilen bulgulara ait görsel sonuçlar şekil 4.2 ve şekil 4.3’te gösterilmiştir. Şekil 4.2’de nesnel olarak bulgular, şekil 4.3’te ise bulguların yüzdesel sonuçları gösterilmiştir.



Şekil 4.2: Silah ve bıçak algılanması çalışması nesnel sonuçlar.



Şekil 4.3: Silah ve bıçak algılanması çalışması yüzdesel sonuçlar.

Bu çalışmada, görüntülerdeki silah ve bıçakların otomatik algılanması için güncel YOLO algoritmaları ile karşılaştırmalı deneysel çalışmalar yapılmıştır. Literatürde yaygın olarak kullanılan YOLOv3 modeli yerine daha güncel daha hızlı performans sonuçları veren YOLOv4, YOLOv5, YOLOX ve YOLOR modelleri tercih edilmiştir. Modellerin performanslarının doğruluğunun kıyaslanabilmesi için modellerde 8 batch, 50 epoch ve görüntülerin boyutlarının 416x416 ebatlarında olması sağlanmıştır. Çalışmada yapılan deneysel

çalışmalar neticesinde modeller arasındaki en başarılı performansı % 95,8 recall, 0,967 Precision ve %97,6 Map@0,5 değerleri YOLOR algoritması ile elde edilmiştir. YOLOR algoritmasından sonra ise en başarılı sonuçlar sırasıyla daha önceki versiyon olmasına rağmen %94,12 map@0,5 başarımla YOLOv4 algoritmasından, %90,69 map@0,5 değeriyle YOLOX ve %86,5 map@0,5 değeriyle YOLOv5 algoritmasından elde edilmiştir.

4.1.2 Anlık Görüntülerde Şiddet Algılanması Çalışması Bulguları

Bu çalışmada, güvenlik kamera sistemlerinde anlık olarak şiddet olaylarının otomatik tespit edilmesi amaçlanmıştır. Bu çalışmada hedeflenen şiddet görüntülerinin algılanması işlemine, literatür kaynaklarında örneklerine oldukça az rastlanılmaktadır. Bunun sebebi, şiddet görüntülerinin algılanmasında diğer nesnelerin algılanmasından farklı olarak hareketli ve şekil değiştiren nesnelerin olmasıdır. Ancak şiddet görüntülerinin sınıflandırılması örneklerine literatürde rastlanılmaktadır. Bu durumun sebebi ise, şiddet görüntülerinin sınıflandırılmasında doğrudan görüntü kayıtları kullanılabilmektedir. Sınıflandırma işleminde bir görüntünün kaydının bütünü kullanılabilmektedir. Ancak algılama(detection) işlemlerinde bir bütünün parçalara ayrılması gerekmektedir. Bu sebeple güvenlik kameralarında anlık yani canlı olarak şiddet olayının tespit edilmesi için algılama (detection) işlemi yapılması gerekmektedir.

Bu çalışmada, güvenlik kamera sistemlerinde şiddet olaylarının otomatik algılanması için görüntü işleme modellerinden, Tensorflow 2.0 Object Detection EfficientDet D0-D7, YOLO modellerinden YOLOv5 ve Tensorflow 2.0 Object Detection Faster R-CNN modelleri kullanılarak deneysel çalışmalar yapılmıştır. Deneysel çalışmalar neticesinde elde edilen başarımlar skorları tablo 4.2’de verilmiştir.

Tablo 4.2: Anlık Görüntülerde Şiddet Algılama Deneysel Sonuçlar.

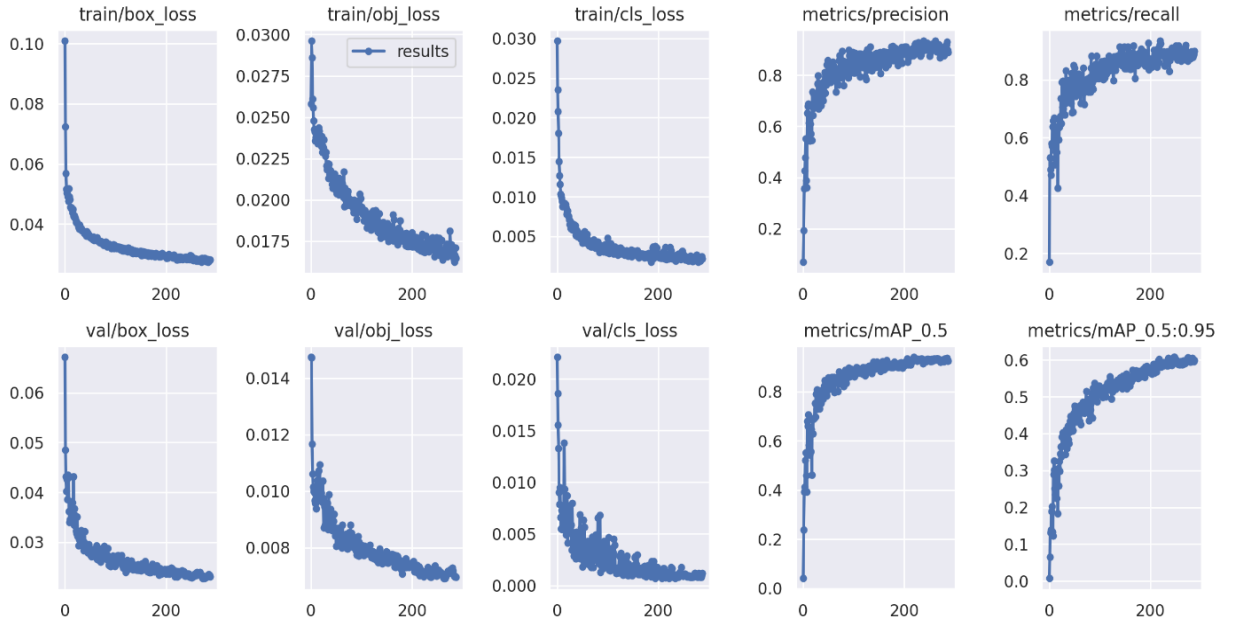
Model	Görüntü Sayısı	Görüntü Boyut	Batch	Epoch	Map
EfficientDet	3333	416X416	16	200	%94.1
YOLOv5	3333	416X416	16	200	%93.5
R-CNN	3333	416X416	16	200	%85.1

En başarılı sonuçlar EfficientDet D0-D7 modeli ve YOLO modellerinden YOLOv5 modelinden alınmıştır. R-CNN modelinin uygulanması aşaması, diğer modellerin yaklaşık 2 katı kadar süre almasına rağmen diğer modeller gibi başarılı sonuçlar vermemiştir.

EfficientDet D0-D7 ve YOLOv5 modellerinin uygulanması neticesinde elde edilen diğer başarımlar skorları tablo 4.3'te ve grafikler şekil 4.4'te verilmiştir.

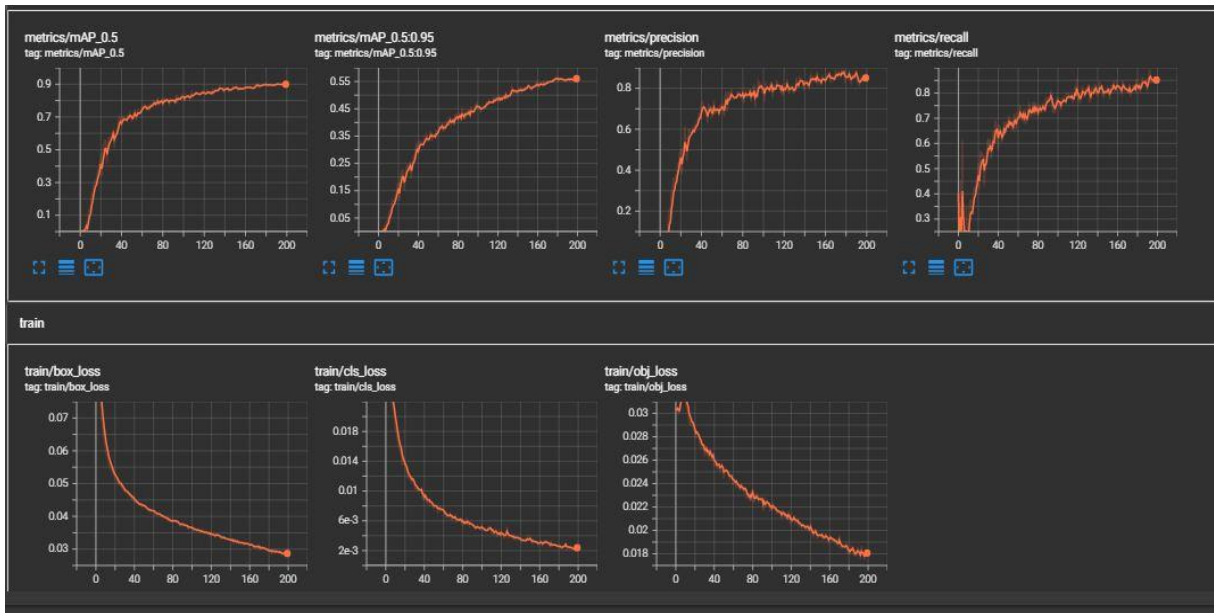
Tablo 4.3: Anlık Görüntülerde Şiddet Algılama Deneysel Sonuçlar-2.

model	recall	precision
EfficientDet D0-D7	%93.5	%88
YOLOv5	%85.5	%90,3



Şekil 4.4: EfficientDet D0-D7 modeli grafiksel çıktısı.

Şekil 4.4 incelendiğinde, veri kayıp (loss) değerleri düştükçe, başarımlar skoru (mAP) değerlerinin yükselmekte olduğu görülmektedir. Bu sonuçlar, modelin sağlıklı çalıştığını göstermektedir.



Şekil 4.5: YOLOv5 modeli grafiksel çıktısı.

Şekil 4.5 incelendiğinde, bir önceki model gibi, loss değerleri düşerken mAP değerleri yükselmiş olduğu görülmektedir. Diğer model gibi bu modelinde sağlıklı çalışmış olduğu gözlemlenmiştir.

Şekil 4.6’da deneysel çalışmalar neticesinde, test verilerinin algılanması işlemlerinin örnek görüntüleri verilmiştir.



Şekil 4.6: YOLO anlık görüntülerde şiddet algılama sistemi örnek çıktıları.v5 modeli grafiksel çıktısı.

Şekil 4.6’da gösterildiği gibi 1 rakamı şiddet olma durumunu, 0 rakamı şiddet olmama durumunu temsil etmektedir. Görüntülerde görüldüğü gibi şiddet olan ortamlarda şiddet olayına karışmış bireyler 1 rakamıyla kare içerisine alınmıştır. Şiddet olaylarının yaşandığı ortamlarda şiddet olaylarına karışmamış bireyler 0 rakamı ile kare içerisine alınmıştır. Sistemin doğru bir algılama yaptığı görüntülerden de anlaşılmaktadır.



Şekil 4.7: Anlık Görüntülerde Şiddet Algılama Sistemi Örnek Çıktıları(Yüzdesel).

Şekil 4.7’de ise sistemin çıktılarının yüzdesel iki sonucu gösterilmiştir. Bu görüntülerde incelendiğinde, ilk görüntüde şiddet olayının yakınında olan bir bireyin olaya karışmadığı bilgisi doğru tespit edildiği görülmektedir.

Anlık görüntülerde silah ve bıçak algılanması çalışması ve anlık görüntülerde şiddet algılanması çalışması, bu tez çalışmasının birinci bölümü olan güvenlik kamerası canlı kayıtlarında anlık görüntü tespiti sistemi için yapılmıştır. Bu başlıktan sonraki sınıflandırma çalışmaları ise tez çalışmasının ikinci bölümü yani boyut indirgeme sistemi için kullanılacaktır.

4.2 İKİNCİ BÖLÜM BULGULARI

Bu bölümde, tez çalışmasının boyut indirgeme sistemi için yapılmış olan deneysel çalışmalar, boyut indirgeme modeli ve oluşturulan insan bilgisayar etkileşimine uygun hazırlanan boyut indirgeme arayüzü açıklanmıştır.

4.2.1. Güvenlik Kamerası Kayıtlarındaki Silah ve Bıçak Görüntülerinin Sınıflandırılması.

Bu tez çalışmasında, güvenlik kamerası kayıtlarında boyut indirgemesi modeli önerilmektedir. Sistemin çalışma alanının güvenlik kamerası sistemi olması sebebiyle, güvenlik zafiyeti olup olmadığı konusu önemlidir. Rastgele olarak boyut indirgemesi çalışması yapılması, sistem için önemli sayılacak görüntülerin zarar görmesine sebebiyet verebilmektedir. Bu nedenle, güvenlik kamerası kayıtlarında boyut indirgeme işlemine başlamadan bazı önemli sayılacak görüntülerin ana kayıta koruma altına alınması gerekmektedir. Bu görüntülerin koruma altına alınması için ilk olarak bu görüntülerin tespit edilmesi gerekmektedir. Önemli sayılacak görüntüler, silah, bıçak ve şiddet görüntüleridir. Bu görüntülerin hem tespit edilmesi, hem de sistem yöneticisinin uyarılması sistemin sağlık çalışması açısından önemlidir.

Güvenlik kamerası sistemi kayıtlarında silah ve bıçak görüntüsünün tespit edilip sınıflandırılması için derin öğrenme tabanlı görüntü işleme tekniklerinden yararlanılmıştır. Bu tekniklerden evrişimli sinir ağlarına dayalı AlexNet, VGG16 ve VGG19 mimarileri ile deneysel çalışmalar yapılmıştır. Çalışmada kullanılan bu mimarilerin kullanımında ilk olarak, bu mimarilerin daha önceleri milyonlarca görüntülerden eğitilmiş ağırlıkları sonrasında ise, sıfırdan eğitilmemiş ağırlıkları kullanılmış ve sıfırdan tarafımızdan eğitilmiş yeni ağırlıkları ile deneysel ölçümler yapılmıştır.

Deneysel çalışmalar için, RTX 2080 ekran kartı, 32 GB Ram bellek, İ7 işlemci ve MATLAB (R2020) yazılımı kullanılmıştır. Veri kümeleri 2 ve 3 sınıflı olarak düzenlenmiştir. Eğitim ve test işlemlerinde veriler, 20-80%, 40-60%, 60-40% ve 80-20% eğitim ve test verisi olarak ayrılmış ve her biri için ayrı ayrı deneysel çalışmalar yapılmıştır. Ayrıca deneysel çalışmalarda çeşitli ağ parametreler kullanılmış ve bunlar tablo 4.4'te gösterilmiştir.

Tablo 4.4: Deneysel Çalışmalarda Kullanılan Ağ Parametreleri.

Minimum/Maksimum Batch Boyutu	8-16
Minimum/Maksimum Epoch Boyutu	5-20
Ağırlık Geçikme Faktörü	0.01
Tekrarlanan Öğrenme Oranı	0.0001
Optimal Metot	Momentumlu Stokastik Gradyan İnişi (SGDM)

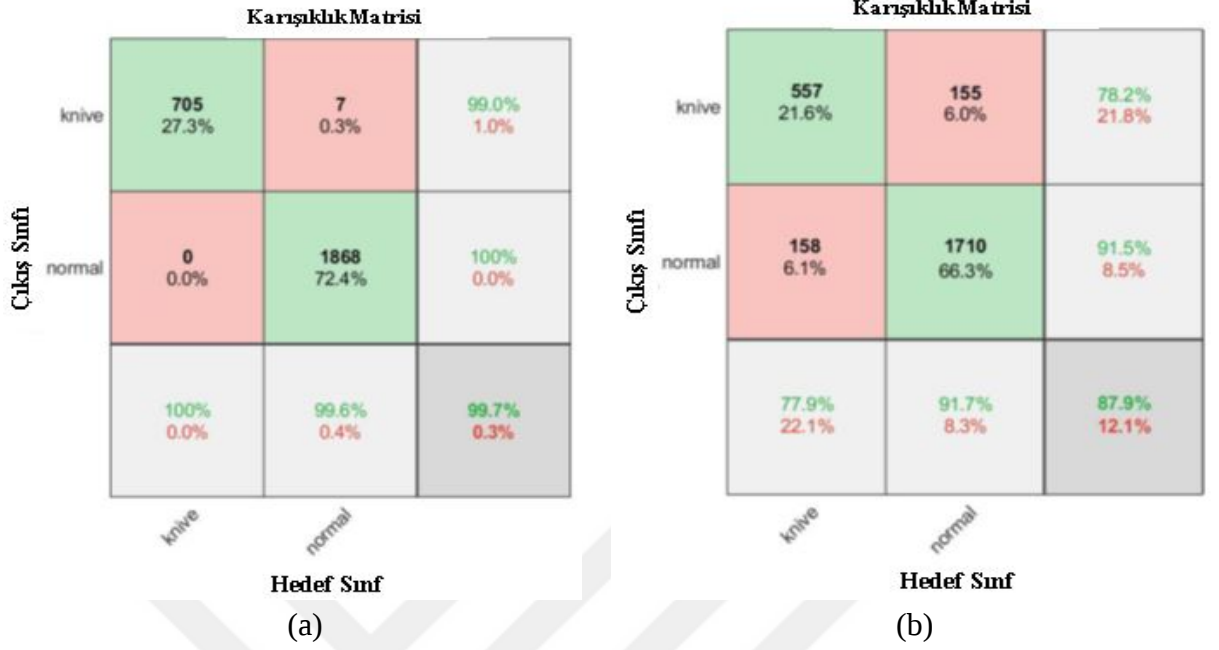
Veri kümelerinin 2 sınıflı(bıçak ve normal görüntüler) olarak düzenlenmesi sonucunda elde edilen deneysel çalışma başarımları skorları tablo 4.5'te gösterilmiştir.

Tablo 4.5: 2-Sınıflı veri kümelerinin başarımları skorları(%).

Yaklaşımlar	Modeller	80-20	60-40	40-60	20-80	Ortalama
Eğitilmiş model (fine-tuning)	AlexNet	98.80	98.31	97.39	96.48	97.74
	VGG16	99.73	99.65	99.42	98.77	99.38
	VGG19	99.65	99.53	99.25	98.66	99.27
Eğitilmemiş Model (Training from scratch)	AlexNet	87.87	87.09	84.05	79.40	84.58
	VGG16	84.73	83.57	80.23	74.38	80.72
	VGG19	83.76	79.46	77.46	75.40	79.02

İki sınıflı veri kümelerinde en başarılı sonuç, eğitilmiş VGG16 modellinden alınmıştır. Eğitilmemiş modellerde başarımları skoru daha düşük çıkmıştır. Bunun sebebi, eğitilmiş modeller daha önceleri milyonlarca ImageNet görüntüsü ile bir eğitim süreci geçirmiş olmasıdır.

Tablo 4.5'te dört farklı eğitim/test kümesi için ayırma değerleri kullanılmış ve her biri için ayrı sonuçlar verilmiştir. Bu sonuçlara göre ince ayar yaklaşımına dayalı derin modeller arasında ortalama olarak en yüksek doğruluğu VGG16 mimarisi elde ederken, AlexNet mimarisi en düşük performansı göstermiştir. Öte yandan, sıfırdan eğitim yaklaşımına dayalı derin modeller arasında, Alexnet mimarisi ortalama olarak en yüksek doğruluğu elde ederken, VGG19 mimarisi en düşük performansı elde etmiştir. Ek olarak, ince ayar yaklaşımına dayalı VGG16 mimarisi, eğitim/test kümelerini 80-20 olarak ayırmada %99,73 başarı ile derin modeller arasında en yüksek doğruluk puanını alırken, AlexNet mimarisi ile %87,87 puan elde edilmiştir. Bu ince ayar ve sıfırdan eğitim yaklaşımlarına dayalı en yüksek performansın karışıklık matrisleri şekil 4.8'de verilmiştir.



Şekil 4.7: 2-Sınıflı veri kümelerinin deneysel çalışmalarının Karışıklık Matrisi çıktısı.

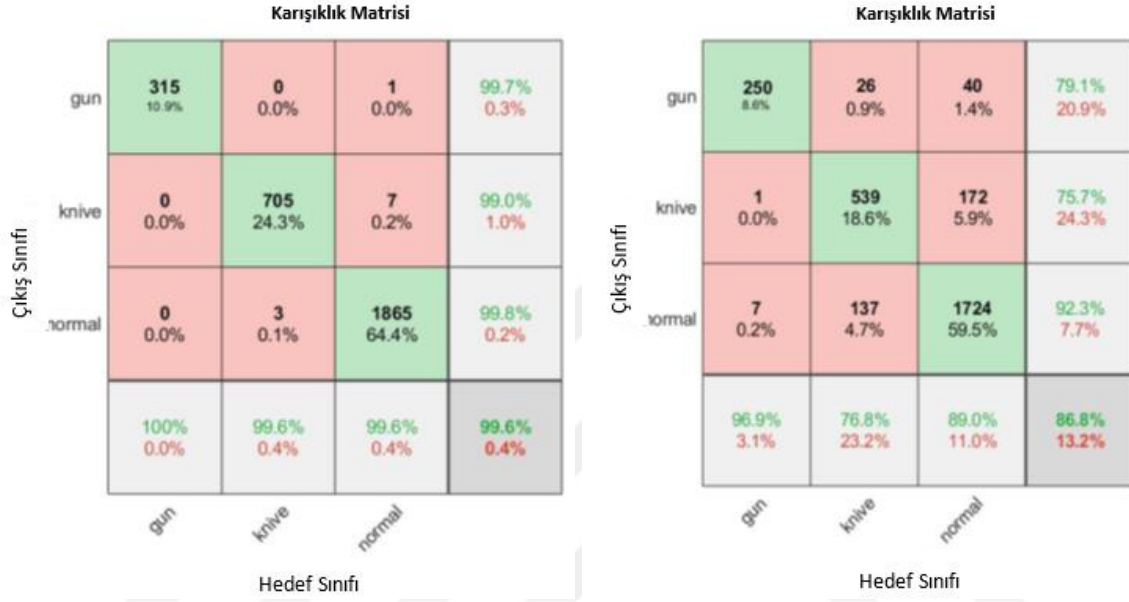
Veri kümelerinin 3 sınıfa (Silah, Bıçak ve Normal) ayrılması işlemi yapılmış ve deneysel çalışmalar tekrarlanmıştır. Bu işlemler neticesinde elde edilen başarımlar skorları tablo 4.6’da gösterilmiştir.

Tablo 4.6: 3-Sınıflı veri kümelerinin başarımlar skorları(%).

Yaklaşımlar	Modeller	80-20	60-40	40-60	20-80	Ortalama
Eğitilmiş model (fine-tuning)	AlexNet	98.86	98.68	97.82	96.72	98.02
	VGG16	99.62	99.41	99.29	98.64	99.24
	VGG19	99.59	99.27	99.15	98.59	99.15
Eğitilmemiş Model (Training from scratch)	AlexNet	86.77	85.25	80.48	72.86	81.34
	VGG16	85.74	83.31	78.18	70.03	79.31
	VGG19	83.67	79.69	74.04	64.64	75.51

Tablo 4.6’da görüldüğü gibi, İnce ayar yaklaşımına dayalı derin modeller arasında en yüksek ortalama doğruluk oranı VGG16 modeli ile %99,24 olarak elde edilirken, sıfırdan eğitim

yaklaşımının kullanıldığı AlexNet modeli %81,34 başarı oranına ulaşmıştır. Eğitim/test kümelerinin 80-20 ayırma değeri açısından, ince ayar yaklaşımına dayalı VGG16 mimarisi ile derin modeller arasında en yüksek doğruluk puanı %99,62 ile elde edilmiştir. İnce ayar ve sıfırdan yaklaşımlara dayalı derin modeller arasında en yüksek performansa sahip karışıklık matrisleri şekil 4.9’da verilmiştir.



Şekil 4.9: 3-Sınıflı veri kümesi Karışıklık Matrisi çıktıları.

Önceden eğitilmiş derin modeller için ince ayar yaklaşımının sıfırdan eğitim yaklaşımına göre daha yüksek performans sağladığı gözlemlenmiştir. Ayrıca tüm deneysel çalışmalardan elde edilen sonuçlara göre ince ayar yöntemine dayalı derin modeller arasında en yüksek doğruluk VGG16 modeli ile elde edilirken, AlexNet modeli sıfırdan eğitim yaklaşımı kullanılarak en yüksek puanı almıştır.

Literatürdeki birçok silah bıçak sınıflandırması yapan çalışma incelenmiş ve bu tez çalışmasında bulunan deneysel sonuçlar ile karşılaştırılmıştır. Karşılaştırma sonuçları tablo 4.7’de gösterilmiştir.

Tablo 4.7: Deneysel sonuçların literatür karşılaştırması.

Kaynaklar	Metotlar	Sınıf Sayısı	Görüntü Sayısı	Skorlar (%)
(Salur vd., 2019)	Fine-Tuning LeNet	3	12,000	99

(Mery vd., 2017)	Visual Vocabularies and Deep Features	4	1,950	95
(Kibria & Hasan, 2017)	AlexNet+SVM	2	2,000	95
(Matiolański vd., 2016)	Fuzzy Sınıflandırma	2	12,899	86
(Dwivedi vd., 2019)	VGGNET	3	5,504	98.41
Bu çalışma	Fine-Tuning Based VGG16	2	12,899	99.73
		3	14,481	99.62

Tablo 4.7’de görüldüğü üzere en yüksek başarımlar skorları bu çalışma neticesinde alınmıştır.

Bu çalışma, suç araçlarının sınıflandırılması için farklı yaklaşımlara dayalı olarak derin evrişimli sinir ağlarının performanslarını değerlendirmiştir. Bu amaçla önceden eğitilmiş AlexNet, VGG16 ve VGG19 modelleri kullanılmıştır. Bu modellerin performansları İnce Ayar ve Sıfırdan Eğitim yaklaşımları kullanılarak hesaplanmıştır. Kapsamlı deneysel çalışmalar sonucunda, suç araçlarının tespitinde derin mimarilerin öğrenilen ağırlıkları kullanılarak %100’e yakın doğruluk elde edildiği açıkça görülmüştür. Sonuç olarak çalışmada, derin mimarilerin öğrenilmiş ağırlıkları kullanılarak güvenlik kameraları ile silahların tespit edilebileceği gösterilmiştir.

4.2.2. Güvenlik Kamerası Kayıtlarındaki Şiddet Görüntülerinin Sınıflandırılması.

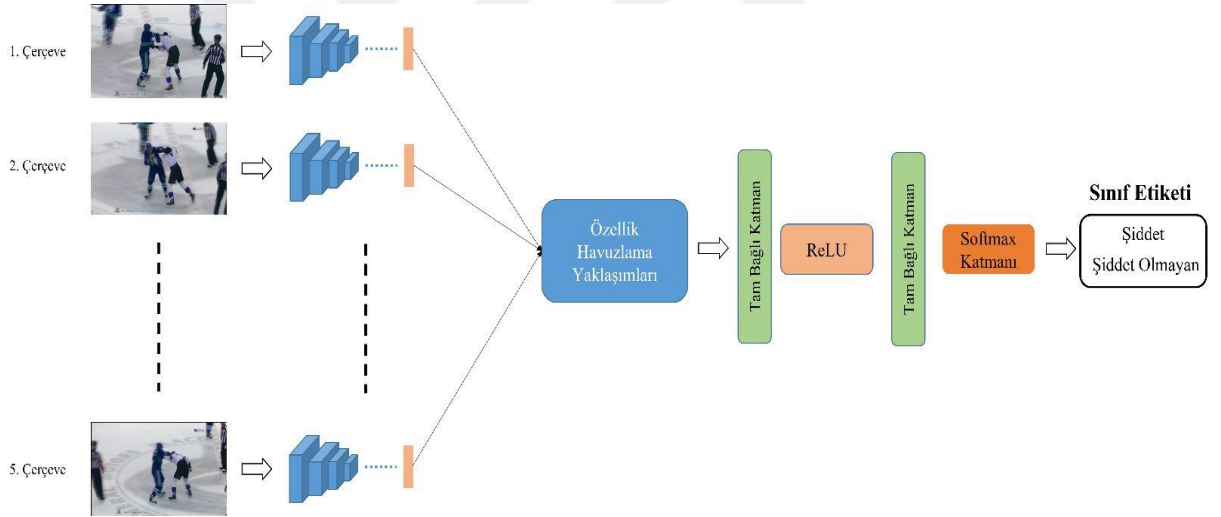
Tez çalışmasının güvenlik kamerası kayıtlarındaki şiddet görüntülerinin sınıflandırılması için iki farklı deneysel çalışma yöntemi kullanılmıştır. Bunlardan ilki AlexNet mimarisinde alt katmanlarda yeni yaklaşımlarla yeni model önerilmesi ve deneysel çalışmalarının yapılmasıdır. İkinci çalışma ise, ESA ve TSA tabanlı modellerin birleştirilmesi ile kombine bir model önerilmesi ve deneysel çalışmalarının yapılmasıdır.

4.2.2.1 Şiddet Sınıflandırılması AlexNet Alt Katmanlar Modeli

Sosyal alanlardaki şiddet davranışlarının kontrolü ve tespiti, güvenliğin ve sosyal istikrarın devamlılığı için çok önemlidir. Güvenlik kameralarına dayalı şiddet ve suç tespit problemlerinde görüntü işleme yöntemlerine dayalı otomatik sistemlere ihtiyaç duyulmaktadır. Bu amaç doğrultusunda bu çalışmada, gözetleme videolarında şiddet tespiti için sunulan öznetelik havuzlama yaklaşımlarına dayalı, derin evrişimli sinir ağları incelenmiştir. İlk olarak,

video görüntülerinden ardışık çerçeveler elde edilmiştir. Daha sonra, transfer öğrenme yaklaşımına dayalı önceden eğitilmiş bir derin evrişimli sinir ağı kullanılarak bu görüntülerden derin öznitelikler çıkarılmıştır. Bu derin özellikler, çalışmada önerilen Özellik Birleştirme, Özellik Ortalama Havuzlama ve Özellik Standart Sapma yaklaşımları kullanılarak birleştirildi. Son olarak, havuzlama işlemlerinden sonra sırasıyla tam bağlı, Relu aktivasyon, tam bağlı ve Softmax katmanları eklendi. Önerilen modeli test etmek için şiddet içeren videoları içeren üç genel veri kümesi kullanıldı. Deneysel sonuçlarda MOVIE, SURV ve Crowd veri kümeleri kullanılmıştır.

Bu çalışma için, video gözetiminde şiddet tespiti için derin evrişimli sinir ağlarına dayalı özellik birleştirme yaklaşımları kullanılmıştır. Bu doğrultuda üç yaklaşım geliştirilmiştir: 1) Özellik Birleştirme, 2) Özellik Ortalama Havuzlama ve 3) Özellik Standart Sapma Havuzlama.

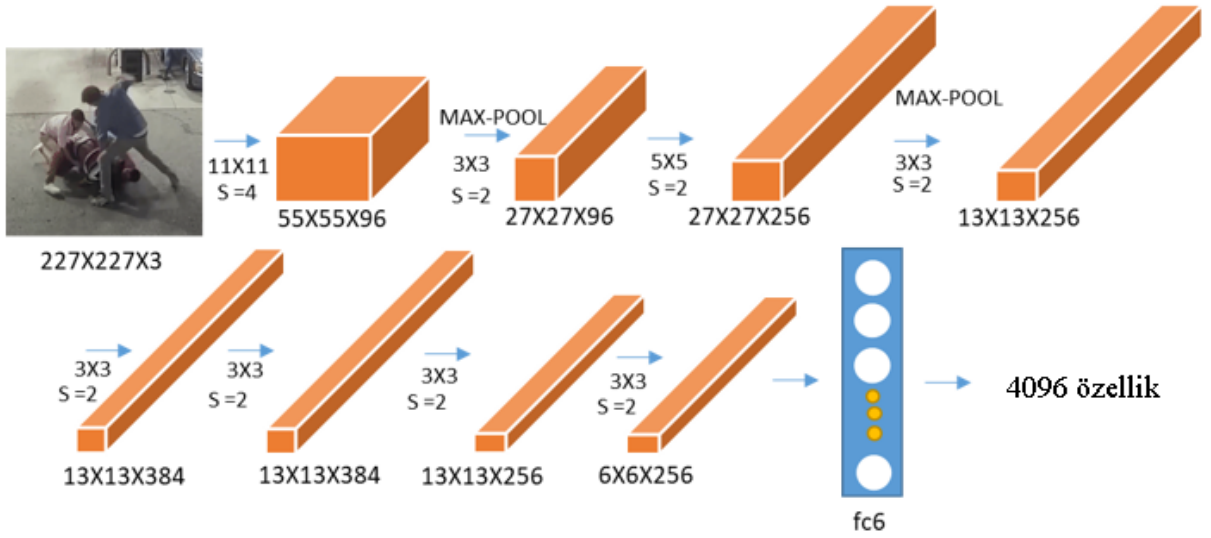


Şekil 4.10: Şiddet görüntülerinin sınıflandırılması için önerilen yaklaşımın genel yapısı.

Şekil 4.10’da görüldüğü gibi, önerilen model dört aşamadan oluşmakta: ön işleme, özellik çıkarma, özellik birleştirme ve sınıflandırma aşamalarıdır. Bu aşamalar alt başlıklarda detaylandırılmıştır.

Aşama 1: Ön işleme aşamasında videolar sıralı ve belirtilen sayıda kareye bölünmüştür. Bir video m sayıda kareden oluştuğunda ve n , kaldırılacak kare sayısı olduğunda, sıralı aralık yuvarlak (m/n) olarak belirlenir. Mevcut çalışmada çerçeve sayıları (n) 5 ve 10 olarak belirlenmiştir.

Aşama 2: Özellik çıkarma aşamasında, 1. Aşamada şiddet içeren videolar kullanılarak elde edilen her bir çerçeve görüntüsü için derin ESA katmanları uygulanmıştır. Bu amaçla önceden eğitilmiş AlexNet mimarisi kullanılmıştır. Önceden eğitilmiş derin mimarinin sıfırdan geliştirilen modellere göre daha az veriye sahip olması ve genellikle daha yüksek performansa sahip olması bu mimarinin tercih edilme sebebi olmuştur. Mevcut çalışmada, AlexNet mimarisinin fc6 adı verilen tam bağlantılı katmanından 4096 öznitelik çıkarılmıştır. Önerilen modelde kullanılan derin öznitelik çıkarma mimarisi şekil 4.11’de verilmiştir.



Şekil 4.11: Önerilen modelin özellik çıkarma mimarisi.

Aşama 3: Özellik birleştirme aşamasında, Denklem kullanılarak her bir çerçeve görüntüsünden derin öznitelikler çıkarılmıştır. Fc6 katmanında özellik çıkarmak için kullanılan denklem aşağıda verilmiştir.

$$feat_k = activation(pretrainedparameters, frame_k, 'fc6') \quad (4.1)$$

$$k = 1, 2, 3, \dots, N$$

Denkleme göre n, çerçeve sayısını temsil ederken, N, özellik sayısını temsil eder. Buna göre önerilen füzyon yaklaşımlarının detayları aşağıdaki gibidir:

- 1) Özellik Birleştirme: Bu yaklaşım, özellik çıkarımı sırasında her bir kare görüntüsünden 1xN sayıda özelliği birleştirmeye dayanır.

$$feat_c = [feat_1, feat_2, \dots, feat_k] \quad (4.2)$$

2) Özellik Ortalama Havuzlama: Bu yaklaşım, özellik çıkarma aşamasında ortalama havuzlama yaklaşımına dayalı olarak her bir çerçeve görüntüsünden elde edilen derin öznitelikleri birleştirmektedir. Bu yaklaşımda sütun tabanlı bir $(1 \times N) \times (\text{çerçeve sayısı})$ özelliğin ortalaması alınmaktadır.

$$feat_{AP} = \frac{1}{K} \sum_{i=1}^K feat_i \quad (4.3)$$

3) Özellik Standart Sapma Havuzlaması: Her bir çerçeve görüntüsünden elde edilen derin öznitelikleri, öznitelik çıkarma aşamasında standart sapma havuzlama yaklaşımına göre birleştirmektedir.

$$feat_{SD} = \sqrt{\frac{1}{K} \sum_{i=1}^K |feat_i - feat_{AP}|^2} \quad (4.4)$$

Aşama 4: Sınıflandırma aşamasında, sırasıyla tam bağlı, Relu, tam bağlı ve softmax katmanları ile öznitelik birleştirme süreci takip etmekte ve gözetim videolarında şiddet tespiti yapılmaktadır.

Önerilen model, bir uygulama platformu olarak MATLAB 2021a'yı kullanmıştır. Çalışmada bu uygulamaları gerçekleştirmek için 32 GB RAM ve NVIDIA Quadro P4000 karta sahip bir iş istasyonu bilgisayarı kullanılmıştır. Ayrıca önerilen yöntem için 10 kat çapraz doğrulama yöntemi kullanılmıştır.

Önerilen modeli test etmek için farklı şiddet içeren üç kamuya açık veri kümesi kullanılmıştır. Bu veri kümeleri kullanılarak yapılan deneysel çalışmalardan elde edilen sayısal ve görsel sonuçlar alt başlıklarda yer almaktadır.

Surveillance Camera Fight Veri Kümesi Deneysel Sonuçları

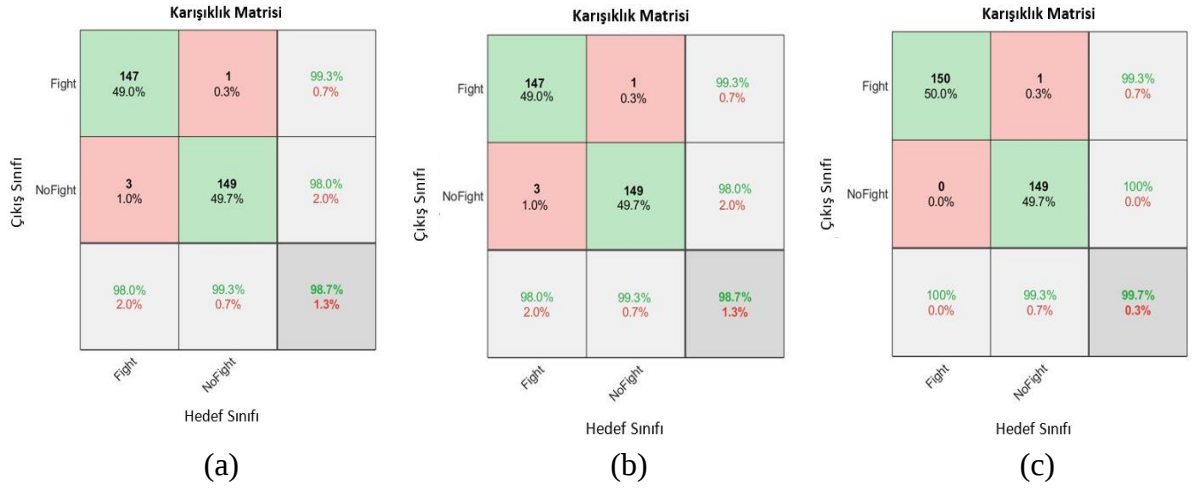
Bu veri setindeki şiddet türleri, farklı ışık ve renk koşullarına bağlı olarak güvenlik kamerasının ortamından bağımsızdır. Tablo 4.8, bu veri kümesindeki çerçeve sayısına göre önerilen yaklaşımların performans sonuçlarını özetlemektedir. Bu sonuçlara göre önerilen yaklaşımlar arasında en iyi doğruluk puanı Önerilen Özellik Standart Sapma tabanlı derin model ile %99,67 iken, Önerilen Özellik Ortalama Havuzlama ve Özellik Birleştirme

yaklaşımlarının doğruluk puanları %98,67 olmuştur.

Tablo 4.8: Surveillance Camera Fight veri kümesi deneysel sonuçları.

Model	Resim Sayısı	Hassaslık	Özgüllük	F1-Skoru	Doğruluk
Özellik Birleştirme tabanlı derin model	5	99.32	98.03	98.66	98.67
	10	98.66	98.01	98.33	98.33
Özellik Ortalama Havuzlama tabanlı derin model	5	98.63	96.10	97.30	97.33
	10	99.32	98.03	98.66	98.67
Özellik Standart Sapma tabanlı derin model	5	99.34	100	99.67	99.67
	10	99.33	99.33	99.33	99.33

Videolardan elde edilen kare sayılarına göre, Önerilen Özellik Standart Sapma ve Birleştirme yaklaşımlarında beş kare ile en yüksek performans elde edilirken, Önerilen Özellik Ortalama Havuzlama yaklaşımında on kare elde edilmiştir. Ek olarak, şekil 4.12’de, tablo 4.8’de önerilen her bir yöntem için elde edilen en yüksek performansların Karışıklık Matrisi değerlerini göstermektedir.



Şekil 4.12: SURV veri kümesi için önerilen modellerin Karışıklık Matrisi değerleri, a) Özellik birleştirme yaklaşımı, b) Ortalama Havuzlama yaklaşımı, c) Standart Sapma Havuzlama yaklaşımı.

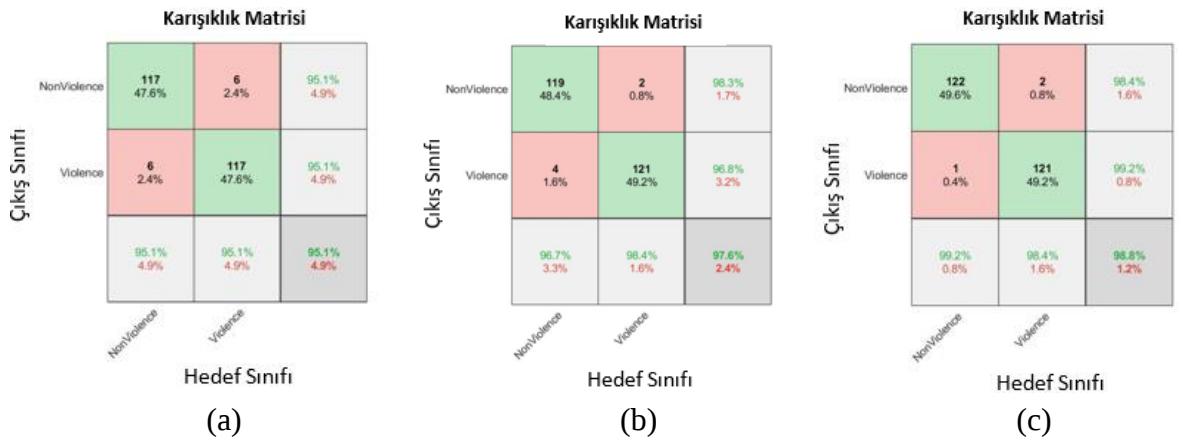
Crowd Violence Veri Kümesi Deneysel Sonuçları

Bu veri kümesi, farklı yerlerden kalabalık ortamlardaki şiddet içeren ve içermeyen vakaları içermektedir. Tablo 4.9, bu veri kümesindeki çerçeve sayısına göre önerilen yaklaşımların performans sonuçlarını özetlemektedir. Bu sonuçlara göre önerilen yaklaşımlar arasında en iyi doğruluk puanı Önerilen Özellik Standart Sapma tabanlı derin model ile %98,78 iken, Önerilen Özellik Ortalama Havuzlama ve Özellik Birleştirme yaklaşımlarının doğruluk puanları sırasıyla %97,56 ve %95,12'dir.

Tablo 4.9: Crowd Violence veri kümesi için önerilen modelin performans sonuçları (%).

Model	Resim Sayısı	Hassaslık	Özgüllük	F1-Skoru	Doğruluk
Özellik Birleştirme tabanlı derin model	5	93.50	93.50	93.50	93.50
	10	95.12	95.12	95.12	95.12
Özellik Ortalama Havuzlama tabanlı derin model	5	97.50	95.24	96.30	96.34
	10	98.35	96.80	97.54	97.56
Özellik Standart Sapma tabanlı derin model	5	95.20	96.69	95.97	95.93
	10	98.39	99.18	98.79	98.78

Videolardan elde edilen kare sayılarına göre önerilen tüm yöntemlerde en yüksek performans on kare kullanılarak elde edilmiştir. Şekil 4.13'te, tablo 4.9'da önerilen her bir yaklaşım için elde edilen en yüksek performansların Karışıklık Matrisi değerleri gösterilmektedir.



Şekil 4.13: Crown veri kümesi için önerilen modellerin Karışıklık Matrisi değerleri, a) Özellik Birleştirme yaklaşımı, b) Ortalama Havuzlama yaklaşımı, c) Standart Sapma Havuzlama yaklaşımı.

Movie Violence Veri Kümesi Deneysel Sonuçları

Bu veri kümesi, film sahnelerinde karşılaşılan dövüş görüntülerini içerir. Tablo 4.10, bu veri kümesindeki çerçeve sayısına göre önerilen yaklaşımların performans sonuçlarını özetlemektedir. Bu sonuçlara göre, önerilen yaklaşımlar arasında en iyi doğruluk puanı Özellik Standart Sapma tabanlı derin model ile %100 iken, Özellik Ortalama Havuzlama ve Özellik Birleştirme yaklaşımlarının doğruluk puanları sırasıyla %99,5 ve %98 olmuştur.

Tablo 4.10: Movie Violence veri kümesi için önerilen modelin performans sonuçları (%).

Model	Resim Sayısı	Hassaslık	Özgüllük	F1-Skoru	Doğruluk
Özellik Birleştirme tabanlı derin model	5	97.98	97.03	97.49	97.50
	10	98	98	98	98
Özellik Ortalama Havuzlama tabanlı derin model	5	100	99.01	99.50	99.50
	10	100	99.01	99.50	99.50
Özellik Standart Sapma tabanlı derin model	5	100	100	100	100
	10	100	100	100	100

Videolardan alınan kare sayılarına göre Önerilen Özellik Ortalama Havuzlaması ve Özellik Standart Sapması için 5 ve 10 kare sayıları kullanılması aynı sonuçları vermiştir. Şekil 4.14'te, tablo 4.10'da önerilen her bir yaklaşım için elde edilen en yüksek performansların Karışıklık Matrisi değerleri gösterilmektedir.

Karışıklık Matrisi

Çıkış Sınıfı	Fight	98 49.0%	2 1.0%	98.0% 2.0%
	NoFight	2 1.0%	98 49.0%	98.0% 2.0%
		98.0% 2.0%	98.0% 2.0%	98.0% 2.0%
		Fight	NoFight	
		Hedef Sınıfı		

(a)

Karışıklık Matrisi

Çıkış Sınıfı	Fight	99 49.5%	0 0.0%	100% 0.0%
	NoFight	1 0.5%	100 50.0%	99.0% 1.0%
		99.0% 1.0%	100% 0.0%	99.5% 0.5%
		Fight	NoFight	
		Hedef Sınıfı		

(b)

Karışıklık Matrisi

Çıkış Sınıfı	Fight	100 50.0%	0 0.0%	100% 0.0%
	NoFight	0 0.0%	100 50.0%	100% 0.0%
		100% 0.0%	100% 0.0%	100% 0.0%
		Fight	NoFight	
		Hedef Sınıfı		

(c)

Şekil 4.14: Movie Violence veri kümesi için önerilen modellerin Karışıklık Matrisi değerleri, a) Füzyon yaklaşımı, b) Ortalama Havuzlama yaklaşımı, c) Standart Sapma Havuzlama yaklaşımı.

Şekil 4.14'te verilen Karışık Matrisi değerlerine göre, Standart Sapma Havuzlama yaklaşımına dayalı olarak önerilen model kullanılarak tüm test videoları doğru bir şekilde sınıflandırılmıştır (Şekil 4.14 (c)). Ayrıca, Ortalama Havuzlama yaklaşımına dayalı önerilen model tüm test videolarından sadece 1 tanesini yanlış sınıflandırırken, Standart havuzlama yaklaşımı 4 videoyu yanlış sınıflandırmaktadır.

Önerilen modelin önceki çalışmalarla karşılaştırılması: Sosyal alanlardaki şiddet davranışlarının kontrolü ve tespiti, güvenliğin ve sosyal istikrarın devamlılığı için önemlidir. Güvenlik sistemlerinde bu tür tespitleri yapmak zor ve yorucu olmaktadır. Bu nedenle son zamanlarda araştırmacılar, görüntü işleme ve makine öğrenmesine dayalı olarak geliştirilen yöntemlerle videolardaki şiddet algılama sorununu çözmeye çalışmışlardır. Daha önceki çalışmalarda video gözetiminde Şiddet tespiti için birçok veri kümesi kullanılmıştır. Bu çalışmada da kullanılan MOVIE, SURV ve Crowd veri kümeleri önceki çalışmalarda en çok tercih edilen veri kümeleridir. Bu veri kümeleri ve önerilen modele dayalı son çalışmaların performans sonuçları tablo 4.11'de karşılaştırılmıştır.

Tablo 4.11: Önerilen model ile önceki çalışmaların doğruluk puanlarının (%) karşılaştırılması.

Kaynaklar	Model	Surv	Crowd	Movie
(Soliman vd., 2019)	VGG16 ve LSTM	-	90	99
(Aktı vd., 2019)	Xception ve Bi-LSTM	72	-	-
(Cheng vd., 2020)	Flow Gated Network based on Optical Flow ve CNN	-	88.8	100
(Su vd., 2020)	Skeleton Points Interaction Learning (SPIL)	-	94.5	98.5
(Traoré & Akhloufi, 2020)	VGG16 ve Bi-GRU	-	95.5	
(Kang vd., 2021)	Motion Saliency Map (MSM), EfficientNet-B0, ve Temporal Squeeze-and-Excitation (T-SE)	92	98	100
(W. Song vd., 2019)	3D ConvNet	-	94.3	99.97

(Sudhakaran & Lanz, 2017)	Conv LSTM	-	94.57	100
(Tran vd., 2015)	Deep 3-dimensional convolutional networks (C3D)	-	84.44	100
Ullah, (2021)	Optical Flow based CNN ve LSTM	74	98.21	-
	Özellik Birleştirme tabanlı derin model	98.67	95.1	98
Bu Çalışma	Özellik Ortalama Havuzlama tabanlı derin model	98.67	97.56	99.5
	Özellik Standart Sapma tabanlı derin model	99.67	98.78	100

Tablo 4.11'de görülebileceği, Movie veri kümesi için önerilen model %100 doğruluk puanına sahip bir puan elde edilmiştir. Ullah (2021) tarafından Crowd veri kümesine dayalı çalışmalar arasında en yüksek doğruluk %98.21 iken, önerilen model %98.78 olarak elde edilmiştir. Literatürde SURV veri kümesini kullanan az sayıda çalışma yapılmıştır. Bu çalışmalar arasında en yüksek doğruluk Kang (2021) ile %92 iken, önerilen model %99.67'ye ulaşmıştır. Tüm bu sonuçlara göre, her üç veri kümesi için önerilen model diğer önceki çalışmalara göre en yüksek performansı sağlamıştır. Ayrıca tablo 4.11'de görüldüğü gibi son zamanlarda video gözetiminden kaynaklanan şiddetin tespitinde derin öğrenme yöntemlerinin kullanıldığı ve yüksek performans elde edildiği görülmektedir.

Önerilen havuzlama yaklaşımlarının analizi;

Önerilen yaklaşımların avantajları ve sınırlamaları aşağıdaki gibidir:

Avantajlar:

- Bu çalışma, önceden eğitilmiş mimariden elde edilen yüksek performanslı özellikler sağlar.
- Özellik havuzlama yaklaşımları, çerçeve görüntülerinden elde edilen çok sayıda derin özneteliği azaltır.
- Önerilen özellik havuzu yaklaşımı, geleneksel derin füzyon yaklaşımından daha yüksek performansa sahiptir.

Sınırlamalar:

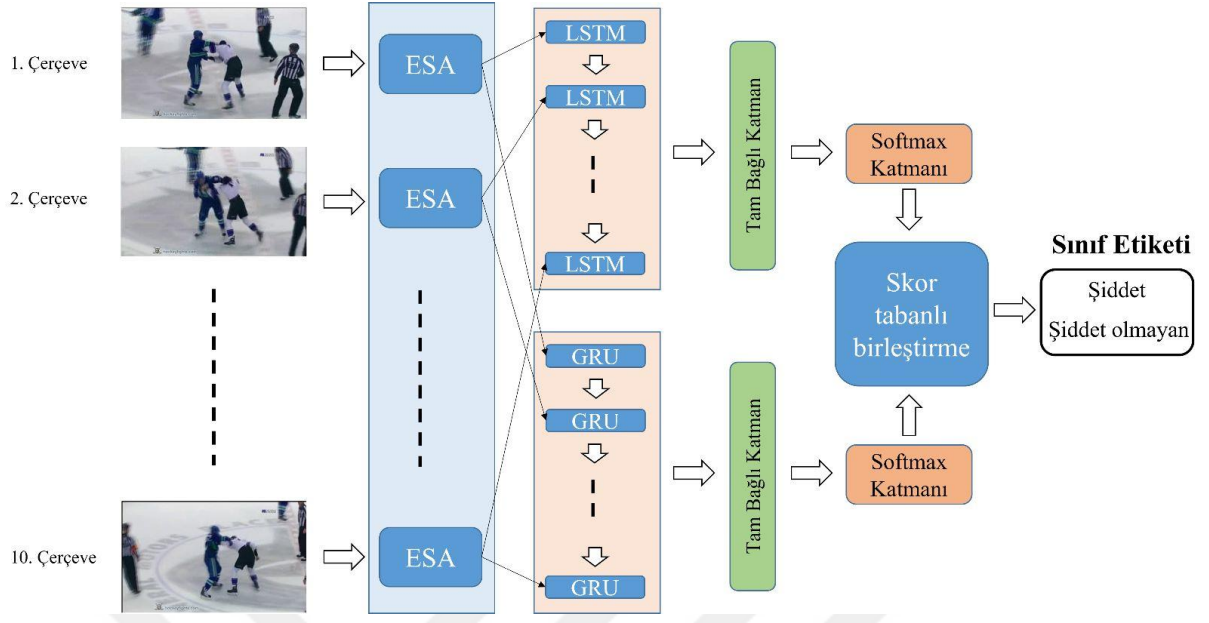
- ESA tabanlı Softmax sınıflandırıcı için optimal parametrelerin varlığı, çalışma performansında bir sınırlama olarak görülebilmektedir.

Bu çalışmada, güvenlik kameralarında şiddet tespiti için derin evrişim sinir ağına dayalı özellik havuzu yaklaşımları geliştirilmiştir. Önerilen bu model, havuzlama yaklaşımlarını kullanarak önceden eğitilmiş derin sinir ağından elde edilen derin öznitelikleri birleştirdi. Deneysel çalışmalarda, önerilen modelin performansını değerlendirmek için halka açık üç veri kümesi kullanılmıştır. Önerilen model kullanılarak MOVIE, SURV ve Crowdveri kümeleri için elde edilen doğruluk değerleri sırasıyla %100, %99,67 ve %98,78'dir. Veri kümeleri için sırasıyla hassaslık değerleri %100, %99,34 ve %98,38 olarak elde edilmiştir. Veri kümeleri için sırasıyla özgüllük değerleri %100, %100 ve %99,18 elde edilmiştir. Veri kümeleri için sırasıyla F1-skorları %100, %99,67 ve %98,78 elde edilmiştir. Bu sonuçlar önerilen modelin önceki çalışmalardan daha başarılı olduğunu göstermiştir. Ek olarak, deneysel çalışmalar, video gözetiminde şiddet tespiti için önerilen modelin güvenlik hizmetlerinde umut verici sonuçlar elde ettiğini kanıtlamıştır.

4.2.2.2 Şiddet Sınıflandırılması ESA-TSA Kombine Model

Önerilen Model

Bu çalışmada, video gözetiminde şiddet tespiti için çok aşamalı sinir ağı topluluklarına dayalı bir Otomatik Şiddet Tespit sistemi geliştirilmiştir. Önerilen model dört aşamadan oluşmaktadır: ön işleme, derin öznitelik çıkarımı, TSA, puana dayalı sınıflandırmadır. Bu kullanımı içeren çizilmiş modelin genel gösterimi şekil 56'da verilmiştir.



Şekil 4.15: Önerilen Kombine modelin mimarisi.

Şekil 4.15'te anlaşılabacağı üzere, model dört aşamadan oluşmaktadır: ön işleme, derin öznitelik çıkarımı, TSA, puana dayalı sınıflandırma. Bu aşamalar alt başlıklar halinde detaylandırılmıştır.

Ön İşleme Aşaması: Ön işleme aşamasında videolar sıralı ve belirli sayıda kareye bölünmüştür. Bir video m sayıda kareden oluştuğunda ve n , kaldırılacak kare sayısı olduğunda, sıra aralığı yuvarlak (m/n) olarak belirlenir. Mevcut çalışmada çerçeve numaraları (n) 5 ve 10 olarak belirlenmiştir.

Derin Özellik Çıkarma: Özellik çıkarma aşamasında, 1. Adımda şiddet içerikli videolar kullanılarak elde edilen her kare görüntü için derin ESA modeli uygulanmıştır. Bu amaçla önceden eğitilmiş AlexNet mimarisi kullanılmıştır. Önceden eğitilmiş derin mimarinin sıfırdan geliştirilen modellere göre daha az veriye sahip olması ve genel olarak daha yüksek performansa sahip olması bu mimarinin tercih sebebi olmuştur. Mevcut çalışmada, AlexNet mimarisinin fc6 olarak adlandırılan tam bağlantılı katmanından 4096 öznitelik çıkarılmıştır. Önerilen modelde kullanılan derin özellik çıkarma mimarisi şekil 4.15'te verilmiştir.

TSA (RNN): Her çerçeve görüntüsünden elde edilen derin özellikler, paralel olarak kullanılan LSTM ve GRU olmak üzere iki RNN tabanlı yöntemle beslenmektedir. Daha sonra bu TSA ağları sırayla birbirine bağlanmaktadır. Son LSTM ve GRU katmanlarından sonra, giriş

videolarının sınıf etiketlerini tespit etmek için tam bağlantılı bir katman, bir softmax katmanı ve bir sınıflandırma katmanı kullanılmaktadır.

Performans Verileri

Önerilen bu kombine model için, iki veri kümesi için k-katlı çapraz doğrulama algoritması ile eğitilip test edilmiştir. Deneysel kullanım adım (epoch) sayısı 100 ve yoğunluk boyutu (batch size) 16 olarak ayarlanmıştır. Uygulamalarda faktör = 0,1 ve patience=8 olarak ayarlanmıştır. Başlangıç öğrenme oranı 0,001 ve bulunması gereken minimum öğrenme oranı ise 0,00001 olarak ayarlanmıştır. Eğitim için Adam optimizasyonu yöntemi kullanılmıştır.

Deneysel çalışmalarda, önerilen modelin şiddet tespitindeki performansını değerlendirmek için doğruluk, duyarlılık, özgüllük ve F1-Score metrikleri kullanılmaktadır.

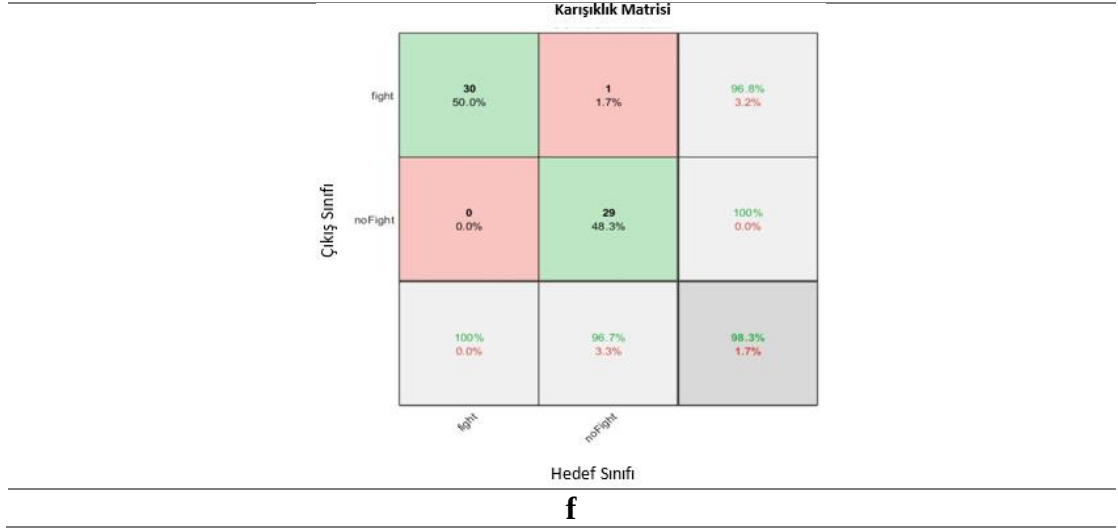
Önerilen modelin performansını hesaplamak için karışıklık matrisi kullanılmıştır. Bu matrisine dayalı olarak duyarlılık, özgüllük, f1 skoru ve doğruluk gibi performans ölçütleri kullanılmıştır. Bu başarımlar ölçütleri ile ilgili matematiksel sonuçlar aşağıda tablo 4.12’ de verilmiştir:

Tablo 4.12: Deneysel çalışma analizlerinin grafiksel gösterimi.

Önerilen Modeller	Performans Ölçütleri	Veri Kümeleri	
		Crowd	Surv
ESA-GRU	Duyarlılık	100	100
	Özgüllük	91.67	90
	F1-Skoru	96.15	95.24
	Doğruluk	95.92	95
ESA-LSTM	Duyarlılık	100	100
	Özgüllük	95.83	93.33
	F1-Skoru	98.04	96.77
	Doğruluk	97.96	96.67
Kombine Model	Duyarlılık	100	100
	Özgüllük	100	96.67
	F1-Skoru	100	98.36
	Doğruluk	100	98.33

Şekil 4.16' da çalışma için kullanılan veri kümeleri ve uygulanan modellerin karışıklık matrisi tabloları verilmiştir.





Şekil 4.16: Crowd veri kümesi (a,b,c) ve Surv veri kümesi (d,e,f), a-d) GRU yöntemi, b-e) LSTM yöntemi c-f) Önerilen Yaklaşım.

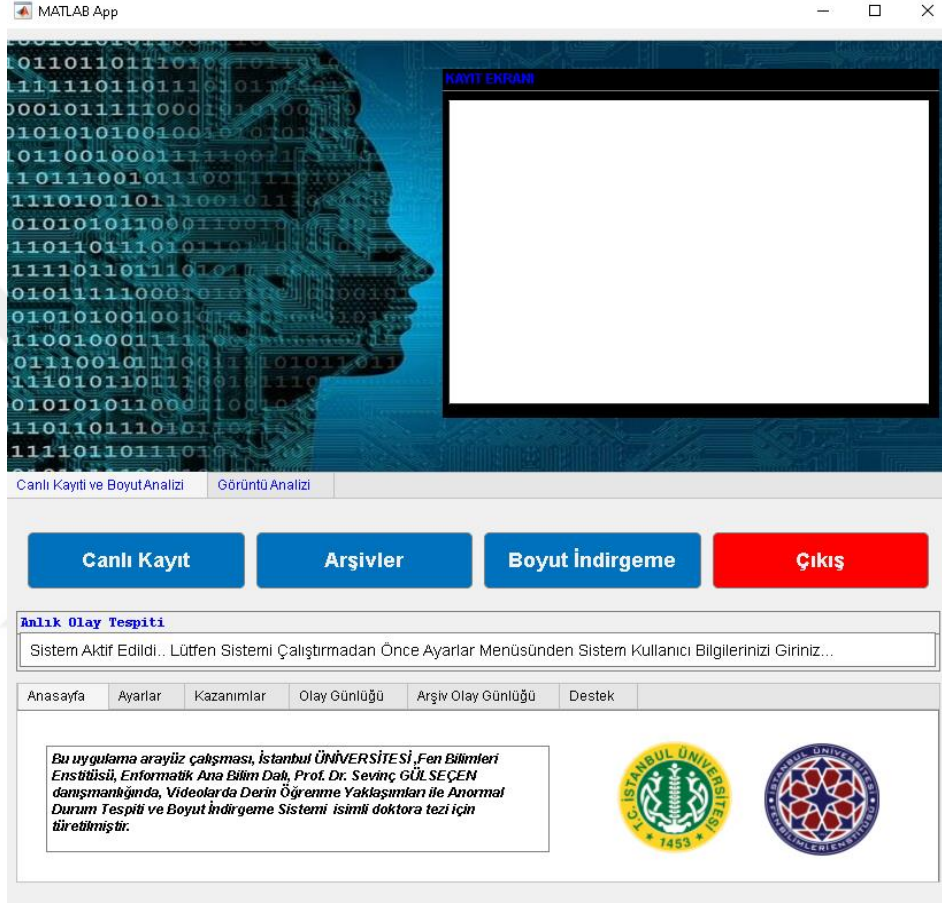
Şekil 4.16’da veri kümesi ve kullanılan deneysel yöntemlerin karşılık matrisi performans çıktıları gösterilmiştir. Önerilen modelde kullanılan ESA modelleri ve performans sonuçları tablo 4.13’ te gösterilmiştir.

Tablo 4.13: Deneysel çalışma analizlerinin grafiksel gösterimi.

Model	Crowd	Surv
AlexNet	97.96	93.33
VGG16	98	95
GoogleNet	98	93.33
SqueezeNet	95.92	90
ResNet	100	96.67
Önerilen DenseNet201 tabanlı yaklaşım	100	98.33

4.2.3 Boyut İndirgeme Çalışması ve Kullanıcı Arayüzü

Boyut indirgeme sisteminin verimli çalışabilmesi için kullanışlı bir kullanıcı arayüzü tasarlanmış ve tüm sistem çalışmalarının bu arayüz üzerinden kontrol edilebilmesi sağlanmıştır. Şekil 4.17’de sistemin arayüzünün açılış ekranı gösterilmiştir.



Şekil 4.17: Görüntü analizi ve boyut indirgeme sistemi arayüzü.

Bu sistemin kullanıcısı insan olacağından arayüz tasarımı, daha kullanışlı ve insanca sistemlerin gelişimini destekleyen ve etkileşimli sistemlerin tasarım ve kullanımı konusunda çalışan insan bilgisayar etkileşimi bilimi alanından destek alınmıştır.

İnsan bilgisayar etkileşimine göre hazırlanan bu sistemde, kullanıcı, görev, amaç ve bağlam temel bileşen olarak kullanılmıştır. Temel bileşenlere göre;

Kullanıcı: Sistemin kullanışlı, etkili ve verimli olmasına dikkat edilmiştir. Sistem hazırlanırken kullanıcının rahat kullanabilmesi için sistem karmaşıklıktan uzak ancak çok fonksiyonlu olarak

tasarlanmıştır. Arayüz tasarımı, her yetişkin bireyin yardım almaksızın rahatlıkla kullanabilmesi için deneysel birçok tasarım değerlendirmeleri yapılarak tasarlanmıştır. Ayrıca sistemin tasarımını, sistem hakkında hiçbir bilgisi olmayan ve tasarımı daha önceleri görmemiş kişilere deneysel olarak kullandırılmış ve rahatlıkla yardım almadan kullanabildikleri gözlemlenmiştir.

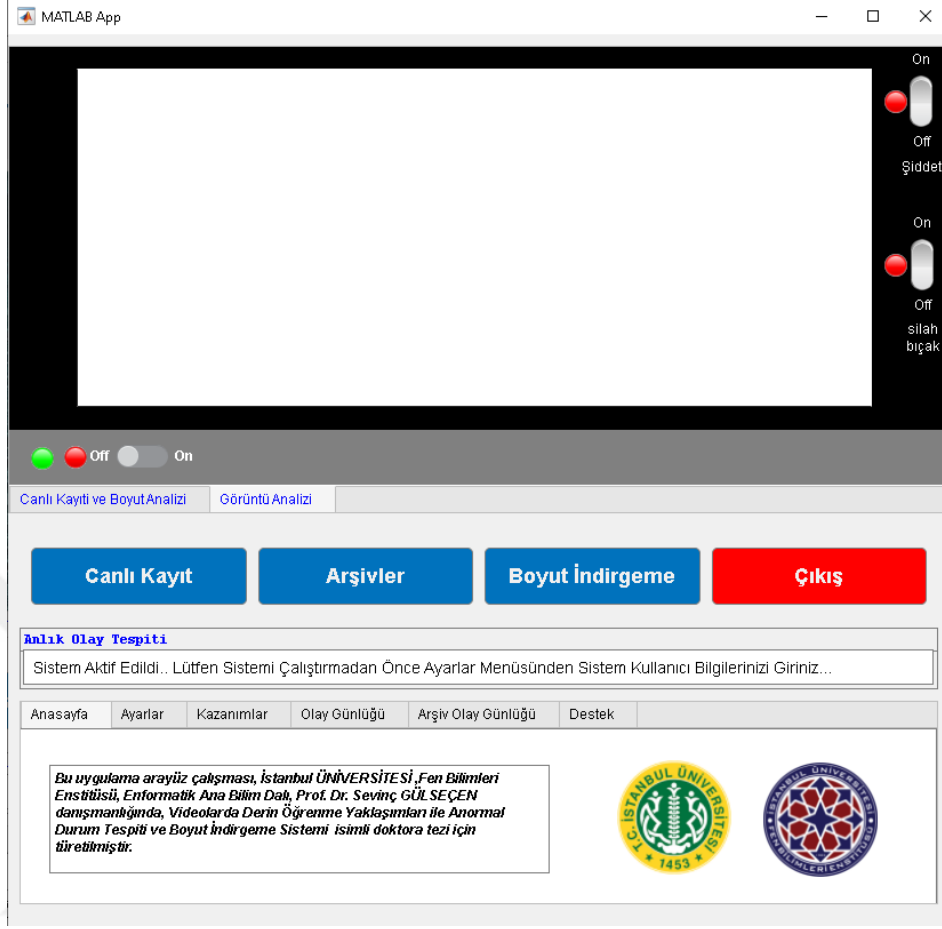
Görev: Sistemin görevleri eksiksiz ve hızlı bir şekilde yerine getirebilmesi için görevlerin sağlanması yapılmış ve sistemin gereksiz hızını düşürebilecek bazı gereksiz özellikler kaldırılmıştır.

Amaç: Sistem tasarımı, bu tez çalışmasının tez adı savunmasından itibaren yapılmış olan toplantılarda ifade edilen amaçlara uygun olarak hazırlanmıştır.

Bağlam: Tezin amaçları, daima tezin çalışma alanları ve bağlamlarına uygun olarak çalışıp çalışmadığına uygunluk göstermiştir. Tasarım hazırlanmasında da tez çalışmasının bağlamları dikkate alınmıştır.

Sistemin kodlanmasında ve tasarımda birçok aşamada Matlab programı 2022b versiyonundan yararlanılmıştır. Tasarım aşamasında özellikle Matlab 2022b AppDesigner, yani matlab arayüz hazırlama paketi kullanılmıştır.

Şekil 4.17’de gösterildiği gibi tasarım iki sekmeden oluşmaktadır. Bu sekmeler canlı kayıt ve boyut analizi sekmesi ve görüntü analizi sekmeleridir. Şekil 4.17’de canlı kayıt ve boyut analizi sekmesinin aktif olduğu sekme görülmektedir. Şekil 4.18’de görüntü analizi sekmesi ve sayfası gösterilmiştir.



Şekil 4.18: Görüntü analizi ve boyut indirgeme sistemi arayüzü, görüntü sekmesi.

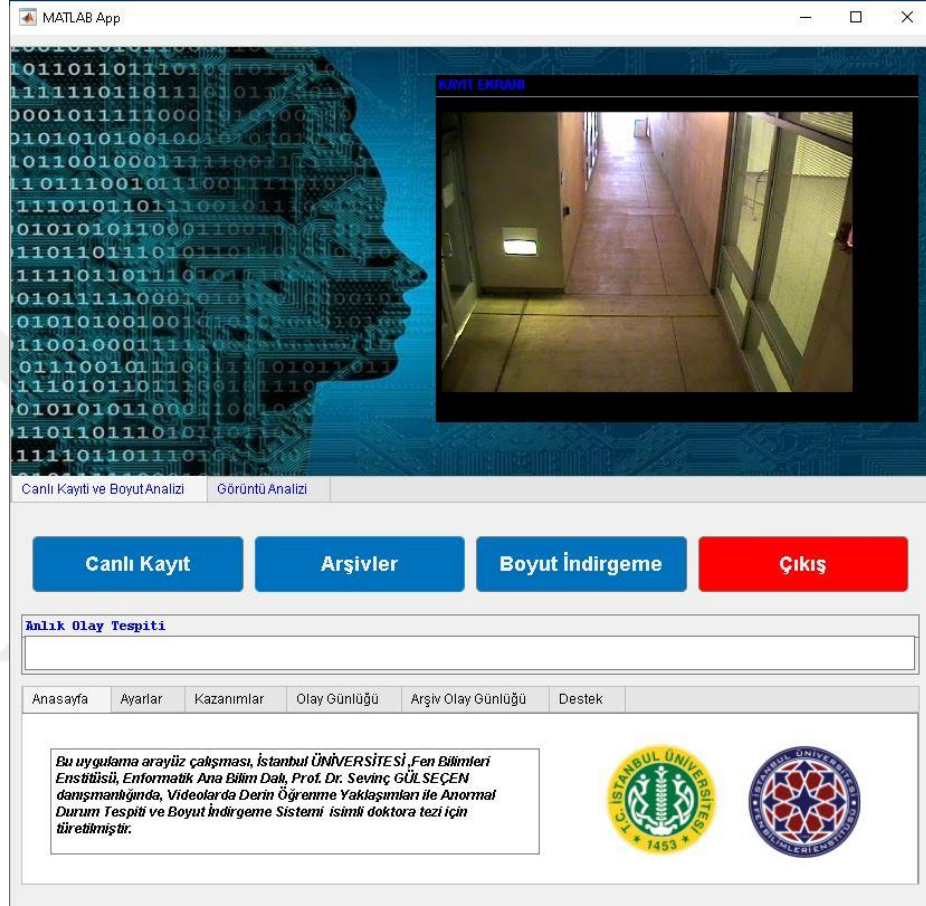
Şekil 4.17 ve şekil 4.18’de gösterilen her iki sekme görüntüsünde görüldüğü gibi bu sekmelerde ortak metin kutusu ve menü bulunmaktadır. Buradan da anlaşılacağı gibi her iki sekmede birbirine bağlantılı olarak çalışmaktadır.

Canlı kayıt ve boyut analizi sekmesinde ve görüntü analizi sekmesinde, sistemin amacına uygun olarak ayrı ayrı görevler yerine getirilebilmektedir. Bu görevler ve ortak menüler alta başlıklar halinde açıklanmıştır.

4.2.3.1 Canlı Kayıt ve Boyut Analizi Sekmesi

Bu arayüz sekmesinde, şekil 4.17’de gösterildiği gibi sistemin anlık canlı güvenlik kamera kayıtlarının ve arşiv kayıtlarının görüntülenebilmesi için ekran bulunmaktadır. Canlı kayıt ve Arşivler menüsü bu sekmenin ekranına bağlı olarak çalışmaktadır.

Canlı Kayıt Menüsü: Bu menü sayesinde canlı kayıt ve boyut analizi sekmesinde anlık olarak güvenlik kamerası sisteminin görüntüleri izlenebilmektedir. Canlı kayıt menüsü doğrudan kayıt izleme ekranına (axes) bağlıdır. Bu menünün çalışmasının örnek görüntüsü şekil 4.19’da gösterilmiştir.



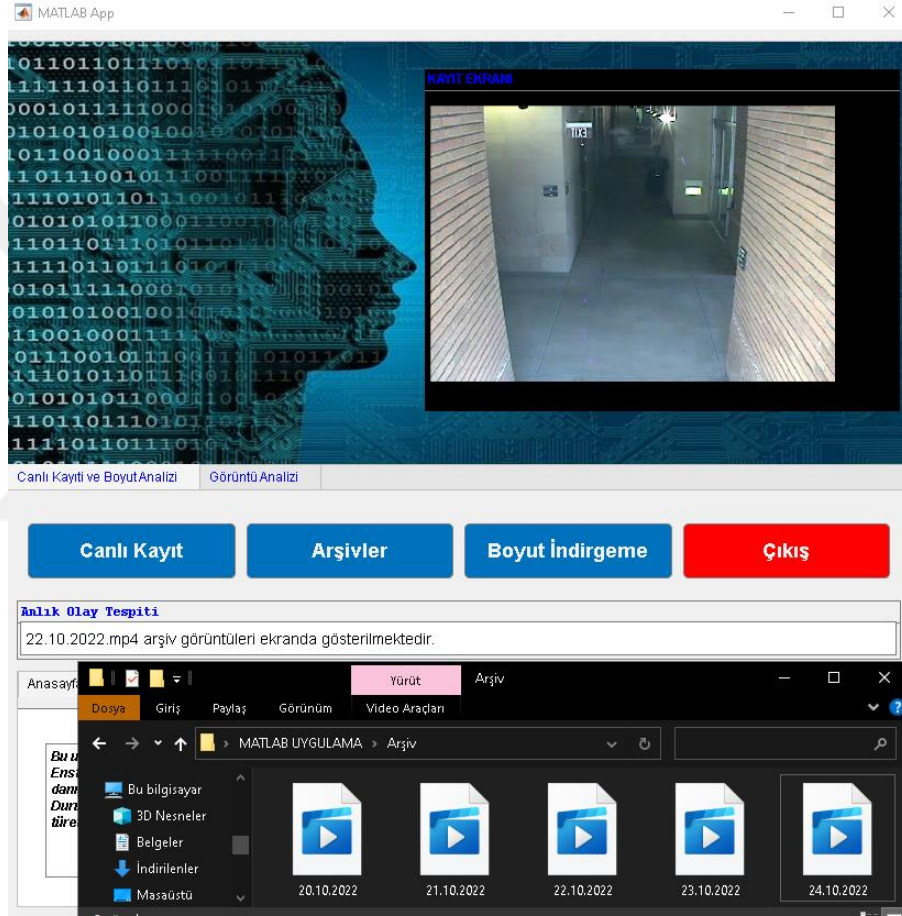
Şekil 4.19: Güvenlik kamerası canlı kayıt görüntüsü.

Bu menünün görüntü ekranına bağlanabilmesi için Matlab Appdesigner kodlarından yararlanılmıştır. Kullanılan örnek kod satırında paylaşılmıştır.

```
app.cam = webcam(1);
frame = snapshot(app.cam);
im = image(app.axes1, zeros(size(frame),'uint8'));
axis(app.axes1,'image');
preview(app.cam,im)
```

Kod satırında, sistem üzerinde tanımlı kamera sisteminin doğrudan axis ekranına bağlandığı görülmektedir.

Arşivler: Bu menü sayesinde daha önceleri kaydedilmiş güvenlik kamerası kayıtları ekran üzerinde izlenebilmektedir. Bu menü çalıştığında direk olarak kayıtlar klasörü çağrılmakta ve arşiv kayıtlarının seçilebilmesi imkanı sunulmaktadır. Arşivler menüsünün çalışmasına örnek görüntü şekil 4.20’de gösterilmiştir.



Şekil 4.20: Güvenlik kamerası arşiv görüntüsü.

Şekil 4.20’de gösterildiği gibi arşiv klasöründen, 22.10.2022 tarihli ekran görüntüsü sistem üzerinden görüntülenmiştir. Ayrıca anlık olay tespiti rapor alanından izlenen görüntüye ait anlık bilgi paylaşılmıştır. Sistemin çalışmasında, ekran görüntüsünün alınması ve anlık rapor paylaşımı için kullanılan Matlab Appdesigner kodu, kod satırında paylaşılmıştır.

```
[file, path] = uigetfile({'*.avi'; '*.mp4'; '*.*'}, 'Select Video File');
yazi= ' arşiv görüntüleri ekranda gösterilmektedir.'
```

```

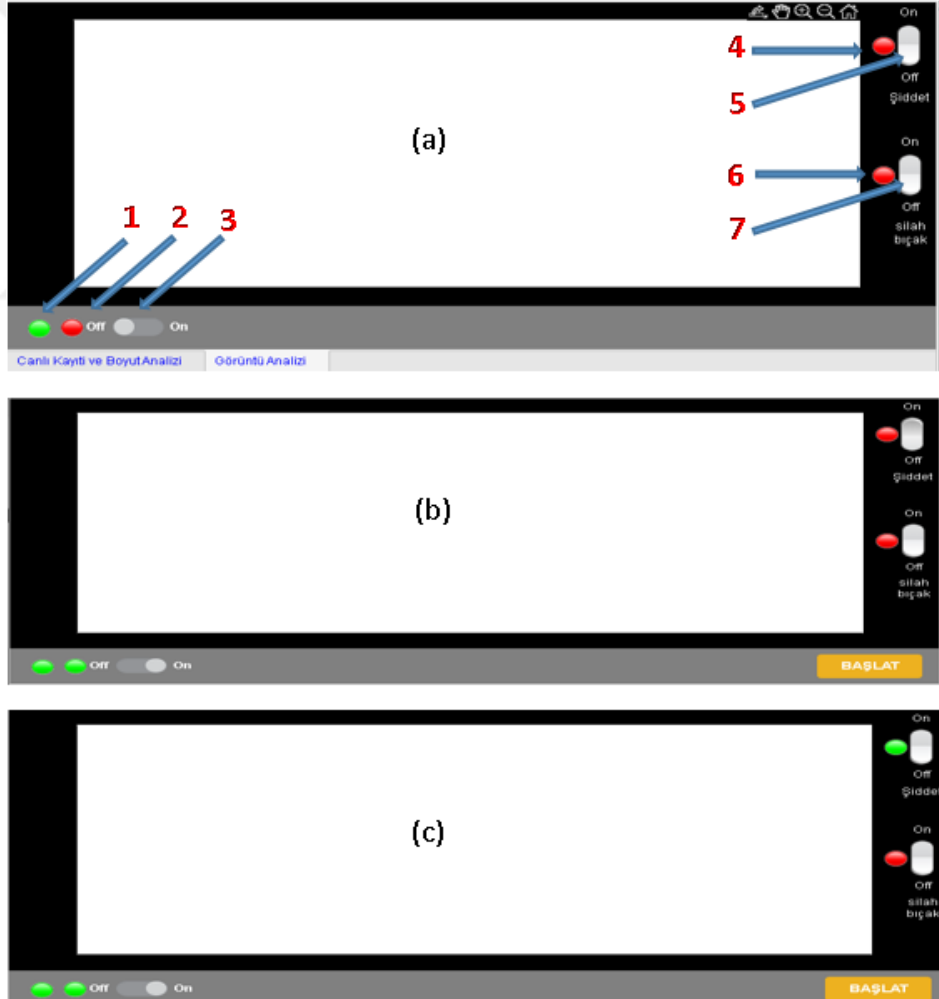
app.EditField.Value= strcat(file, yazi);
.
.
.
imshow(frame, 'Parent', app.axes1);

```

Bu menüde doğrudan canlı kayıt ve boyut analizi sekmesi ekranına bağlı olarak çalışmaktadır.

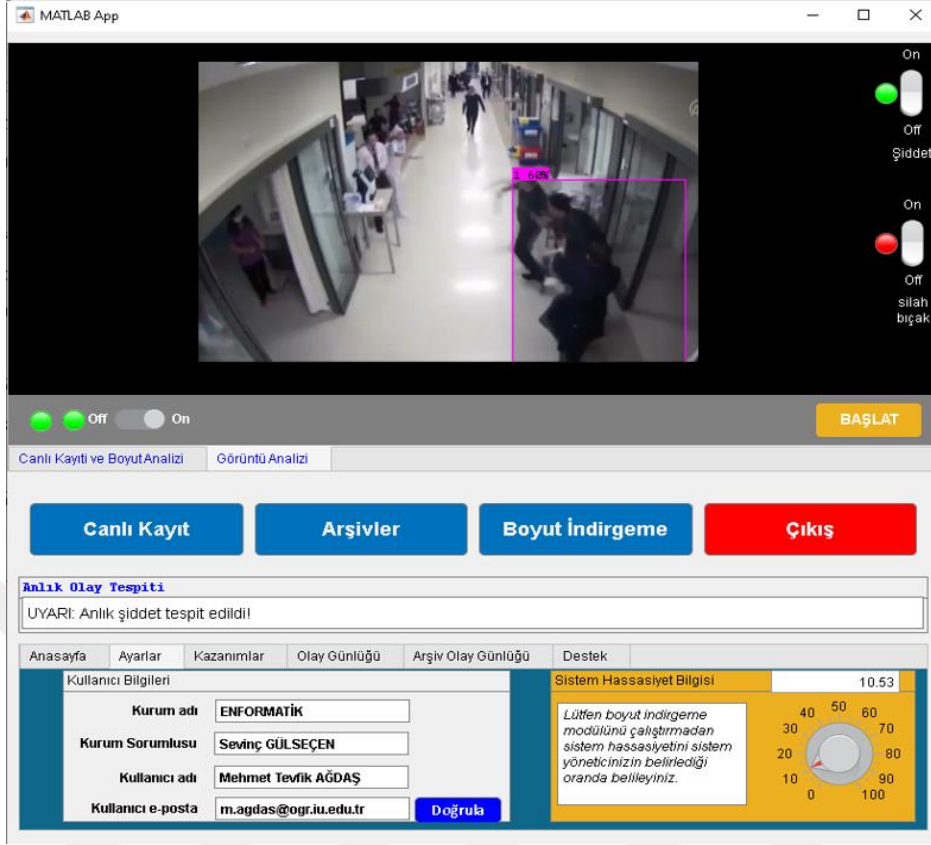
4.2.3.2 Görüntü Analizi Sekmesi

Bu sekme, görüntülerde şiddet ve şiddet unsuru içeren durumların tespitinin yapılabilmesi için tasarlanmıştır. Bu sekme, görüntüye ait örnek görüntü şekil 4.18’de gösterilmiştir. Bu sekmede, dört adet lamba, üç adet anahtar, bir adet görüntü aktarıcı ekran ve bir adet gizli algoritmali komut butonu kullanılmıştır. Bu komutlar şekil 4.21’de gösterilmiştir.



Şekil 4.21: Görüntü analizi sekmesi komutları, a) açılış, b) sistem aktif, c) aktif ve şiddet tespiti.

Şekil 4.21’de açılış görüntüsünde gösterildiği gibi sistem ilk çalıştığında 1 numaralı lamba yeşil renkte çalışmaktadır. Ancak boyut indirgeme işlemi başlatıldığında, işlem bitinceye kadar lamba kırmızı yanmakta ve bittiğinde tekrar yeşil renge dönmektedir. Lambalarda, yeşil renk sistemin uygun olduğunu, kırmızı renk ise sistemin çalıştırılmaya uygun olunmadığını göstermektedir. Şekil 4.21’de gösterilen 2 numaralı lamba, sistem sekmesi açıldığında ilk olarak kırmızı renkte çalışmaktadır. Şekildeki 2 numaralı lamba, 3 numarası ile gösterilen sistem açma anahtarı ile aktif edilmektedir. Sistem açma anahtarı açıldığında 2 numaralı lamba yeşil renge yani sistemin açıldığını göstermektedir. 3 numaralı sistem açma anahtarı aktif edildiğinde ayrıca sistemin başlatılması için kullanılan şekil 4.21’de sistem aktif görüntüsünde gösterilen gizli buton aktif olmaktadır. Başlat butonu sayesinde hem hedef görüntü seçilebilmekte hem de sistem analizi işlemi başlatılmaktadır. Şekil 4.21’de açılış ekranında gösterilen 5 numaralı anahtar, şiddet analizi komutunun aktif edilmesini sağlamaktadır. Sekmedeki 5 numaralı anahtar aktif edildiğinde, 4 numaralı lamba yeşil olmakta ve kapatıldığında lamba kırmızı renge dönmektedir. Şekilde açılış ekranında gösterilen 7 numaralı anahtar, silah ve bıçak analizi komutunun aktif edilmesini sağlamaktadır. Bu anahtar aktif edildiğinde, 6 numaralı lamba yeşil renkte olmakta ve anahtar pasif olduğunda kırmızı renge dönmektedir. Ayrıca şekilde açılış ekranında gösterilen 3 numaralı anahtar kapatıldığında, 2,4 ve 6 numaralı lamba kırmızı renge dönmekte ve başlat butonu gizlenmektedir.



Şekil 4.22: Görüntü analizi sekmesi, şiddet tespiti.

Şekil 4.22’de görüntü analizi sekmesinde şiddet tespitine ait örnek görüntü verilmiştir. Görüntü incelendiğinde, ekranda şiddet tespiti yapıldığı ve anlık olay tespiti rapor alanında şiddet tespit edildiğine dair uyarı verildiği görülmektedir. Sistemin çalışmasında, iki farklı işlem ile şiddet ve şiddet unsuru tespiti işlemi yapılmaktadır. Sistem kullanıcısı talep ederse tespit işlemi için, bu çalışma için eğitilmiş olan şiddet görüntülerinin sınıflandırılması veya algılanması uygulamaları kullanılabilir. Bu uygulamaların deneysel başarımları daha önceki bulgular konu başlıklarında açıklanmıştır.

Şekil 4.22’de gösterildiği gibi ayarlar menüsünden, kullanıcı bilgileri alanında sistem kullanıcısının bilgilerinin girilmesi halinde, sistem anlık raporları ve görüntü bilgileri ayrıca sistem kullanıcısına e-posta bildirisi olarak gönderilmektedir.

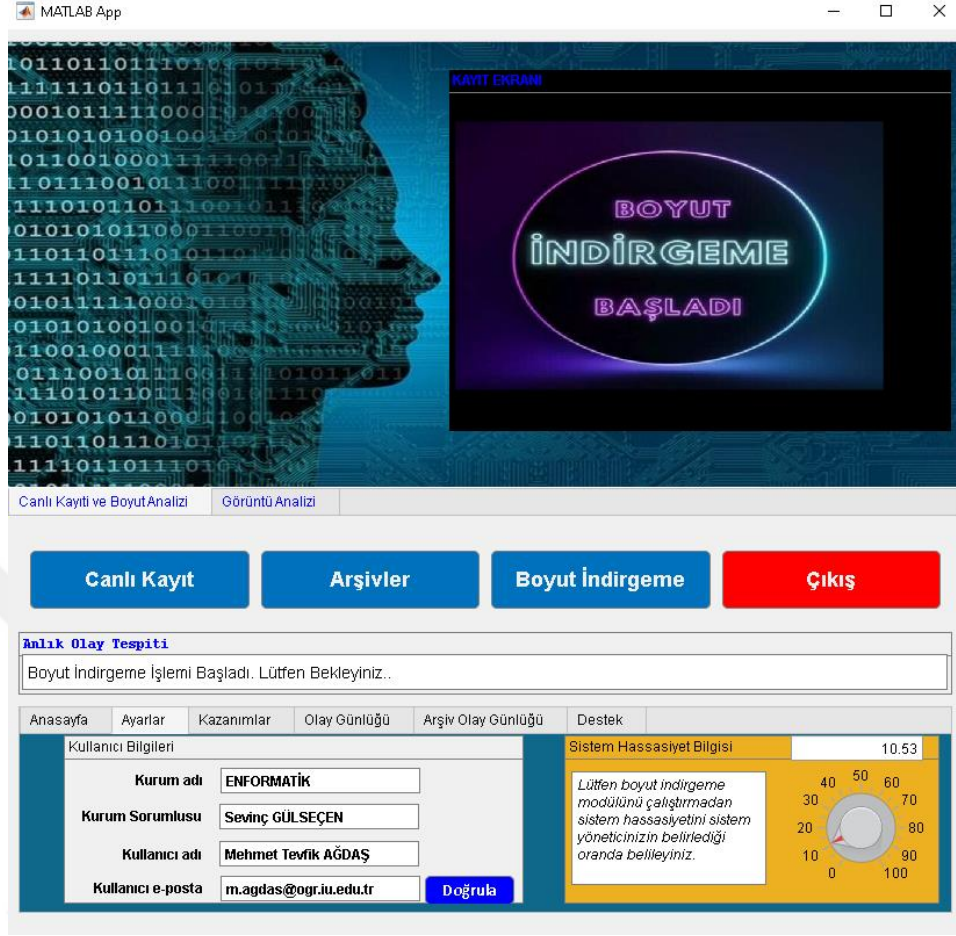
4.2.3.3 Sistem Sekmelerinde Ortak Kullanılan Menü ve Komutlar

Bu tez çalışmasının görüntü analizi ve boyut indirgeme arayüzünde kullanılan sekmeler için ortak menü ve komutlar hazırlanmıştır. Bu menü ve komutlar, boyut indirgeme ana butonu, çıkış ana butonu, anasayfa menüsü, ayarlar menüsü, kazanımlar menüsü, olay günlüğü menüsü, arşiv olay günlüğü ve destek menüsüdür. Butonlar, menüler ve menülerin içeriğindeki komutlar ayrı ayrı alt başlıklarda açıklanmıştır.

Boyut İndirgeme Butonu

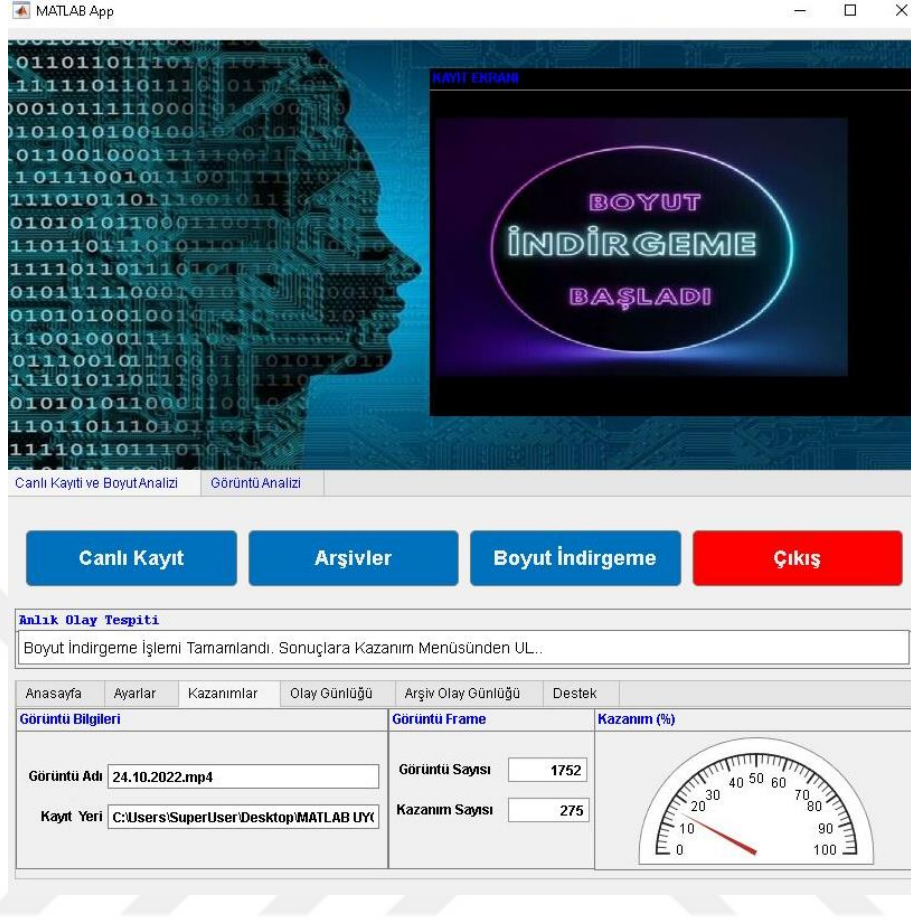
Boyut indirgeme butonu, daha önceleri kayıt edilmiş görüntüler üzerinde boyut indirgeme işlemini başlatmaktadır. Bu butona tıklanmadan önce ayarlar menüsünde kullanıcı bilgilerinin girilmesi ve hassasiyet ayarlamasının yapılması gerekmektedir. Ayarlar menüsündeki kullanıcı bilgilerinin doğru girilmesi boyut indirgeme işleminin bitiminde bazı bilgilerin kullanıcıya e-posta olarak gönderilmesini ve hassasiyet ayarlamasının kurallarını belirlemektedir. Hassasiyet ayarlamasının kuralları, kullanıcı bilgilerindeki kurum sorumlusu tarafından belirlenmekte ve kullanıcı tarafından tanımlaması yapılmaktadır. Hassasiyet ayarlamasında, hassasiyet anahtarı kullanılmaktadır. Hassasiyet anahtarı 1000'de oransal değer ile sistemin hassasiyetini belirlemektedir. Boyut indirgemedede ardışık iki görüntü arasındaki görüntü farkının değerlendirilmesi, hassasiyet değerlerine göre yapılabilmektedir.

Boyut indirgeme butonuna tıklandığında, boyut indirgeme sistemi başlamakta ve arşiv görüntülerinden görüntü seçilmesi ile boyut indirgeme işlemi çalışmaktadır. Boyut indirgeme çalışmasında ilk olarak görüntüler resim karelerine ayrılmaktadır. Resim kareleri üzerinde görüntü işleme ve görüntü filtreleme çalışmaları yapılarak resimler arasındaki farklar ve belirgin özellikler ortaya çıkarılmaktadır. Farkları ve belirgin özellikleri çıkarılan resimler birbirinden çıkarılarak 1000'de oransal farklılık değeri oluşturulmaktadır. Sistem kullanıcısı tarafından girilen hassasiyet değerinin altında kalan farklılık değerine sahip görüntülerden ikincisi silinmekte ve çalışma tamamlanmaktadır. Görüntüler tekrar sıralanıp arşiv klasörüne kaydedilmektedir. Sistem boyut indirgeme çalışması şekil 3.9'da boyut indirgeme modeli algoritmasındaki adımlara göre çalışmaktadır. Şekil 4.23'te boyut indirgeme çalışmasının örnek görüntüsü gösterilmiştir.



Şekil 4.23: Boyut indirgeme işlemi örnek görüntüsü.

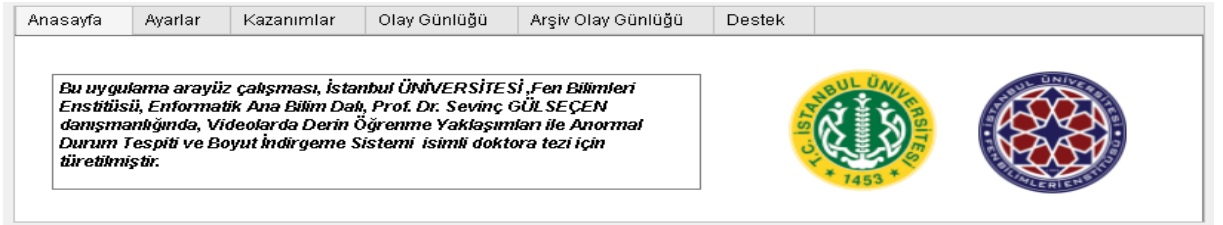
Şekil 4.23'te gösterildiği gibi sistem çalışmaya başladığında, canlı kayıt ve boyut indirgeme sekmesindeki ekranda boyut indirgeme görseli gösterilmektedir. Anlık olay tespiti raporunda ise, boyut indirgeme işlemi başladığı bilgisi paylaşılmaktadır. Ayrıca kazanım menüsüne, kazanım değerleri ve görüntü bilgileri otomatik olarak yüklenmektedir. Çalışmaya ait örnek kazanım bilgileri şekil 4.24'de gösterilmiştir.



Şekil 4.24: Boyut indirgeme örnek kazanım bilgileri.

Anasayfa Menüsü

Bu menü, sistemin anasayfasında sistem hakkında genel bilgiler paylaşılabilmesi için tasarlanmıştır. Bu menü sayfasında, sistem tasarımcısı hakkında bilgi ve logo bilgileri içermektedir. Daha sonraları sistem kullanıcıları tarafından bu bilgiler değiştirilebilir. Anasayfa menüsünün görüntüsü şekil 4.25'te gösterilmiştir.



Şekil 4.25: Anasayfa menü sayfası görüntüsü.

Ayarlar Menüsü

Bu menü sayfasından, kurum bilgisi, kurum sorumlusu, kullanıcı adı, e-posta bilgisi ve boyut indirgeme hassasiyet değeri bilgileri ayarlanmaktadır. Bu sayfada ayarlanan özellikler sistemin birçok aşamasında kullanılmaktadır. Şekil 4.26'da ayarlar menü sayfasının görseli gösterilmiştir.

Şekil 4.26: Ayarlar menü sayfası görseli.

Kullanıcı bilgileri alanındaki e-posta bilgisi doğru girildiğinde anlık olay bilgisi raporları ve olay günlükleri sistem kullanıcısına e-posta olarak anlık gönderilmektedir. Sistem kullanıcısının doğru e-posta girdiğinin test edilmesi için doğrula butonuyla doğrulama e-postası gönderilmesi sağlanmıştır. E-posta sistemine ait kod satırı aşağıda paylaşılmıştır.

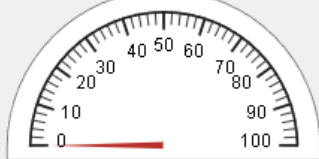
```
mail = 'm.agdas@ogr.iu.edu.tr';
psswd = '*****';
host = 'smtp.mail.gmail.com';
port = '465';
emailto = app.eposta.Value;
m_subject = 'subject';
m_text = 'test';
setpref( 'Internet', 'E_mail', mail );
setpref( 'Internet', 'SMTP_Server', host );
setpref( 'Internet', 'SMTP_Username', mail );
setpref( 'Internet', 'SMTP_Password', psswd );
props = java.lang.System.getProperties();
props.setProperty( 'mail.smtp.user', mail );
props.setProperty( 'mail.smtp.host', host );
props.setProperty( 'mail.smtp.port', port );
props.setProperty( 'mail.smtp.starttls.enable', 'true' );
props.setProperty( 'mail.smtp.debug', 'true' );
props.setProperty( 'mail.smtp.auth', 'true' );
props.setProperty( 'mail.smtp.socketFactory.port', port );
props.setProperty( 'mail.smtp.socketFactory.class',
'javax.net.ssl.SSLSocketFactory' );
props.setProperty( 'mail.smtp.socketFactory.fallback', 'false' );
sendmail( emailto , app.EditField.Value );
```

Sistem hassasiyet bilgisi, boyut indirgeme sisteminde hassaslık oranının belirlenmesi için tasarlanmıştır. Hassaslık değerinin boyut indirgeme sistemine bağlayan kod satırı aşağıda paylaşılmıştır.

```
hassas=app.Knob.Value;
.
.
if degisimyuzde<hassas
```

Kazanımlar Menüsü

Bu menü sayfasında, boyut indirgeme sisteminin tamamlanması ile kazanımlar değerleri sunulmaktadır. Sayfada ayrıca görüntünün adresi, görüntünün adı, görüntüdeki resim karesi sayısı ve kazanım (silinen) resim sayısı gösterilmektedir. Şekil 4.27’de kazanım menü sayfasının görseli paylaşılmıştır.

Anasayfa	Ayarlar	Kazanımlar	Olay Günlüğü	Arşiv Olay Günlüğü	Destek
Görüntü Bilgileri		Görüntü Frame		Kazanım (%)	
Görüntü Adı		Görüntü Sayısı 0			
Kayıt Yeri		Kazanım Sayısı 0			

Şekil 4.27: Kazanım menü sayfası görseli.

Kazanım değeri, yüzdesel gösterge üzerinde gösterilmektedir. Kazanımın yüzdesel değeri aşağıdaki matematiksel denklem ile hesaplanmaktadır.

$$kazanım(\%) = \frac{kazanım\ sayısı}{görüntü\ sayısı} \times 100 \quad (4.5)$$

Olay Günlüğü Menüsü

Anlık olay bilgisi alanında sürekli olarak güncel bilgiler paylaşılmaktadır. Bu bilgilerin kaybolmaması için olay günlüğü menüsü sayfası hazırlanmıştır. Olay günlüğü kayıtları arşiv olay günlüğü klasörüne kaydedilmektedir. Olay günlüğü menü sayfası görüntüsü şekil 4.28’de gösterilmiştir.

Anasayfa	Ayarlar	Kazanımlar	Olay Günlüğü	Arşiv Olay Günlüğü	Destek	
No- Olay Tanımı	Kayıt Bilgisi	Kayıt Yeri	Tarih			
01- Şiddet Tespit Edildi	30ekim2022.mp4	C:\Users\SuperUser\Desktop\MATLAB UYGULAMA\uygulama\	11:27 01/11/2022			

Şekil 4.28: Olay günlüğü menü sayfası görseli.

Arşiv Olay Günlüğü Menüsü

Arşiv olay günlüğü menü sayfasında, olay günü ve geçmiş tarihli olaylar toplu olarak gösterilmektedir. Ayrıca olay günlükleri excel raporu halinde indirilebilir. Şekil 4.29’da arşiv olay günlüğü menü sayfası görseli gösterilmiştir.

Anasayfa	Ayarlar	Kazanımlar	Olay Günlüğü	Arşiv Olay Günlüğü	Destek	
No- Olay Tanımı	Kayıt Bilgisi	Kayıt Yeri	Tarih			
01- Şiddet Tespit Edildi	30ekim2022.mp4	C:\Users\SuperUser\Desktop\MATLAB UYGULAMA\uygulama\	11:27:26 01/11/2022			
02- Bıçak Tespit Edildi	30ekim2022.mp4	C:\Users\SuperUser\Desktop\MATLAB UYGULAMA\uygulama\	11:35:15 01/11/2022			
03- Bıçak Tespit Edildi	31ekim2022.mp4	C:\Users\SuperUser\Desktop\MATLAB UYGULAMA\uygulama\	14:12:41 01/11/2022			

Şekil 4.29: Arşiv olay günlüğü menü sayfası görseli.

Destek Menüsü

Bu menü sayfası, sistem kullanıcılarının sistemsel sorunlarla karşılaştıklarında sistem sorumlusu veya tasarımcısı ile iletişim ve destek alabilmesi için hazırlanmıştır. Destek menü sayfasında, sistemsel problemlerin yazılabileceği mesaj bölümü ve mesajın gönderilebilmesi için gönder butonu bulunmaktadır. Şekil 4.30’da destek menü sayfasının görseli gösterilmiştir.

Anasayfa	Ayarlar	Kazanımlar	Olay Günlüğü	Arşiv Olay Günlüğü	Destek	
						Sayın Kullanıcımız,
						Destek taleplerinizi anlaşılır bir şekilde yandaki mesaj kutusuna yazınız.
						Destek talabinizi oluşturmadan önce mutlaka Ayarlar menüsünden Kullanıcı Adınızın ve E-posta bilginizin doğru olduğundan emin olunuz. Teşekkürler
Gönder						

Şekil 4.30: Destek menü sayfası görseli.

Çıkış Butonu

Çıkış butonu sistemden çıkış yapmak için tasarlanmıştır. Bu butona tıklandığında görüntü analizi ve boyut indirgeme programı tamamen kapatılmaktadır. Çıkış butonu için kullanılan kod satırı aşağıda paylaşılmıştır.

```
delete(app.UIFigure);
```



5. TARTIŞMA VE SONUÇ

Çalışmanın bu bölümünde, çalışmanın araştırma ve geliştirme uygulamalarının sonuçları yorumlara dayalı olarak tartışılmış ve açıklanmıştır. Literatürdeki benzer çalışmalara kıyasla sonuçların değerlendirilmesi yapılmıştır. Ayrıca gelecekte yapılacak çalışmalara yol gösterici olması sebebiyle öneriler ve deneyimler paylaşılmıştır.

5.1 Boyut İndirgeme ve Görüntü Analizi Çalışmalarının Amaç ve Sonuçları

Bu tez çalışmasında, güvenlik kamera sistemlerinde derin öğrenme yaklaşımları kullanılarak canlı görüntülerdeki anormal durumların tespit edilip sistem yöneticisinin bilgilendirilmesi amaçlanmıştır. Ayrıca, güvenlik kamerası kayıtları üzerinde derin öğrenme sınıflandırma ve algılama işlemleri yapılarak görüntülerdeki gereksiz görüntülerin silinerek boyut indirgemesi çalışması amaçlanmıştır. Görüntü analizleri ve boyut indirgeme sisteminin gerçekleştirilmesi için amaçlanan detay çalışmalar ve sonuçları altta maddeler halinde sunulmuştur.

Görüntü analizi çalışmasının detaylı hedefleri;

- **Amaç 1:** Güvenlik kamerası anlık görüntülerinde derin öğrenme temelli silah ve bıçak görüntülerinin anlık algılanması işlemi yapılabilmesi için veri kümesi oluşturulması.
- **Sonuç 1:** Güvenlik kamerası sistemlerinde silah ve bıçak görüntülerinin anlık algılanması için internet ortamında ve akademik çalışmalardan elde edilmiş olan güvenlik kamerası kayıtlarından oluşan silah ve bıçak içeren görüntüler toparlanmıştır. Elde edilen veri kümesi 3000 adet silah ve 2078 adet bıçak görüntüsü içermektedir.
- **Amaç 2:** Güvenlik kamerası anlık görüntülerinde derin öğrenme temelli silah ve bıçak görüntülerinin anlık algılanması için deneysel çalışmanın yapılması.
- **Sonuç 2:** Görüntülerde anlık silah ve bıçak algılanması işleminin yapılabilmesi için deneysel çalışmalar yapılmıştır. Deneysel çalışmalarda derin öğrenme temelli ve görüntü tespitlerinde başarılı sonuçlar vermiş olan ve literatür araştırmalarında popüler kullanılan dört model denenmiştir. Bunlar, YOLOv4, YOLOv5, YOLOX ve YOLOR modelleridir. Deneysel çalışmalar neticesinde elde edilen modellerin karşılaştırmalı başarımları skorları tablo 5.1’de gösterilmiştir.

Tablo 5.1: Silah ve bıçak görüntülerinin algılanması başarımları.

MODEL	Recall	Precision	Map@0,5
YOLOv4	0,91	0,90	0,9412
YOLOv5	0,816	0,86	0,864
YOLOX	0,89	0,894	0,9069
YOLOR	0,958	0,967	0,976

- **Amaç 3:** Güvenlik kamerası anlık görüntülerinde derin öğrenme temelli şiddet olayı görüntülerinin anlık algılanması işlemi yapılabilmesi için veri kümesi oluşturulması.
- **Sonuç 3:** Şiddet algılama çalışması, görüntülerde anlık olarak şiddet içeren görüntülerin şiddet bölümlerini tespit etmektedir. Bu çalışma için akademik çalışmalardan, internet sayfalarından ve açık erişim veri tabanlarından veri toplanması işlemi yapılmıştır. Şiddet görüntüsü içeren 3333 adet görüntü elde edilmiştir.
- **Amaç 4:** Güvenlik kamerası anlık görüntülerinde derin öğrenme temelli şiddet olayları görüntülerinin anlık algılanması için deneysel çalışmanın yapılması.
- **Sonuç 4:** Bu çalışmada, güvenlik kamera sistemlerinde şiddet olaylarının otomatik algılanması için görüntü işleme modellerinden Tensorflow 2.0 Object Detection, EfficientDet D0-D7 ve YOLO modellerinden YOLOv5 modeli kullanılarak deneysel çalışmalar yapılmıştır. Deneysel çalışmalar neticesinde elde edilen başarımları tablo 5.2’de verilmiştir.

Tablo 5.2: Görüntülerde anlık şiddet olaylarının algılanması.

Model	Map
EfficientDet	%94.1
YOLOv5	%93.5
R-CNN	%85.1

- **Amaç 5:** Güvenlik kamerası anlık görüntülerinde derin öğrenme temelli silah ve bıçak görüntülerinin sınıflandırılması işleminin yapılabilmesi için veri kümesi oluşturulması.

- **Sonuç 5:** Şiddet görüntüsü içeren literatürde daha önceleri birçok çalışmaya tekrar tekrar kaynaklık etmiş veri kümelerinden, internet tarayıcılarından ve video sayfalarından görüntüler toplanarak amaca uygun bir veri kümesi oluşturulmuştur. Sonuç olarak veri kümesi, 9500 adet bıçak görüntüsü, 3500 silah görüntüsü ve 3000 adet normal görüntü içeren toplamda 16000 görüntü ile hazırlanmıştır.
- **Amaç 6:** Güvenlik kamerası anlık görüntülerinde derin öğrenme temelli silah ve bıçak görüntülerinin sınıflandırılması için deneysel çalışmanın yapılması.
- **Sonuç 6:** Güvenlik kamerası sistemi kayıtlarında silah ve bıçak görüntüsünün tespit edilip sınıflandırılması için derin öğrenme tabanlı görüntü işleme tekniklerinden yararlanılmıştır. Bu tekniklerden evrişimli sinir ağlarına dayalı AlexNet, VGG16 ve VGG19 mimarileri ile deneysel çalışmalar yapılmıştır. Çalışmada kullanılan bu mimarilerin kullanımında ilk olarak, bu mimarilerin daha önceleri milyonlarca görüntülerden eğitilmiş ağırlıkları sonrasında ise, sıfırdan eğitilmemiş ağırlıkları kullanılmış ve sıfırdan tarafımızdan eğitilmiş yeni ağırlıkları ile deneysel ölçümler yapılmıştır.

Deneysel çalışmalarda, veri kümeleri 2 ve 3 sınıflı kullanılarak ayrıca sistem eğitilmiş deneysel sonuçlar elde edilmiştir. 2 sınıflı veri kümesi içeriğinde, bıçak ve bıçak olmayan görüntüler, 3 sınıflı veri kümesi içeriğinde ise bıçak, silah ve şiddet aleti içermeyen görüntüler bulunmaktadır. Her iki deneysel çalışma neticesinde de literatürdeki çalışmalardan daha başarılı sonuçlar elde edilmiştir. Tablo 5.3'te bu çalışma neticesinde elde edilen başarımlar ve başarımlar elde edilen model gösterilmiştir.

Tablo 5.3: Silah ve bıçak tespiti deneysel başarımlar ve başarımlar.

Sınıf Sayısı	Model	Başarımlar
2 Sınıflı	Eğitilmiş model	99.38
	(fine-tuning)	
	AlexNet	84.58
	Eğitilmemiş Model	
	(Training from scratch)	
	AlexNet	

3 Sınıflı	Eğitilmiş model (fine-tuning)	99.24
	VGG16	
	Eğitilmemiş Model (Training from scratch)	81.34
	VGG16	

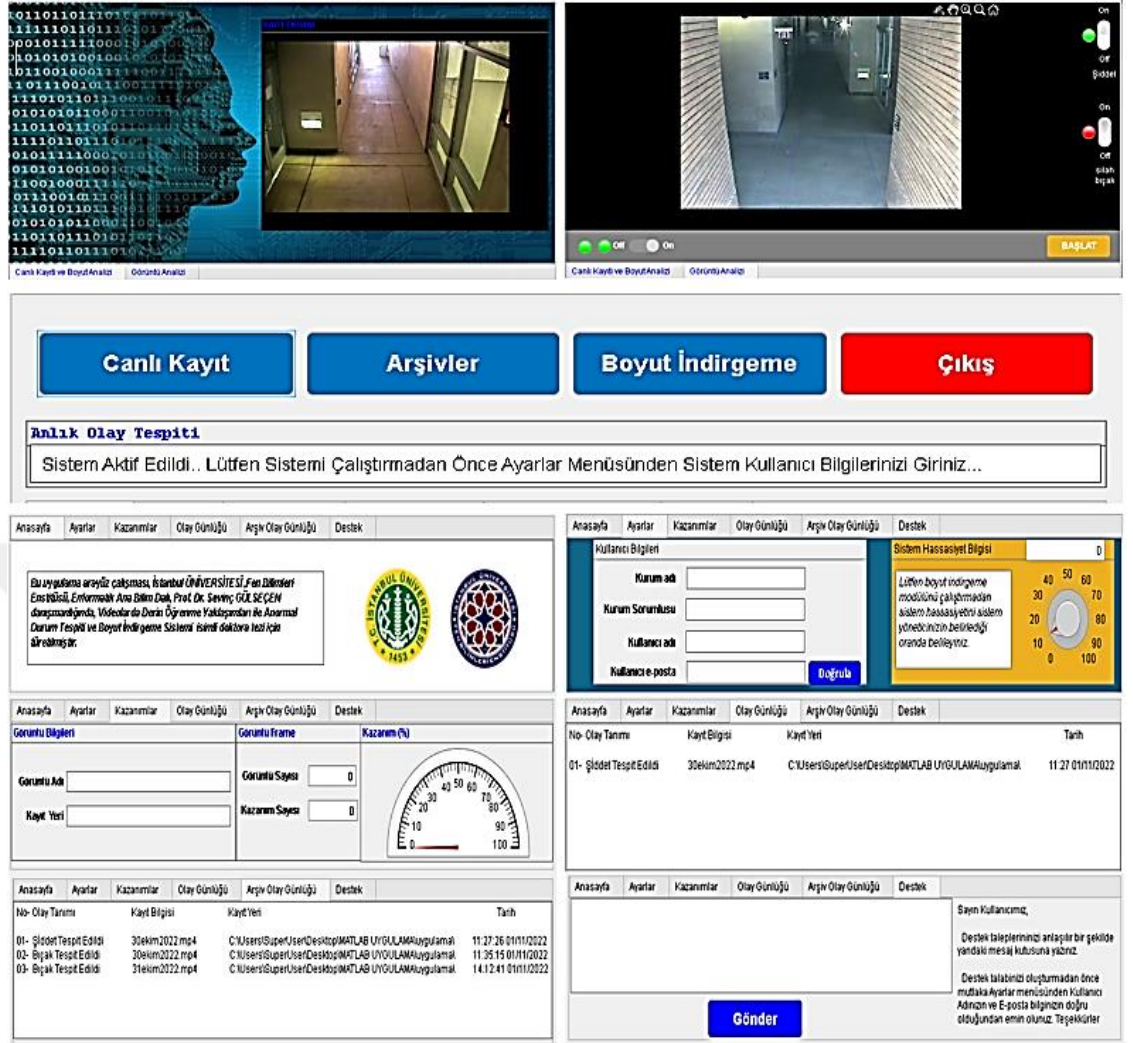
- **Amaç 7:** Güvenlik kamerası anlık görüntülerinde derin öğrenme temelli şiddet sınıflandırılması işlemi yapılabilmesi için veri kümesi oluşturulması.
- **Sonuç 7:** Şiddet olaylarının sınıflandırılabilmesi çalışmasının yapılması için 3 ayrı deneysel veri kümesinden yararlanılmıştır. Veri setlerinin isimleri sırasıyla, The Crown Violence, ve Movie Violence veri kümesi şeklindedir. The Crown Violence veri kümesi, ortalama süresi 3,6 saniye olmak üzere 246 görüntüden oluşmaktadır. The Fight Detection Surv Master veri kümesi, 150 kavga içeren ve 150 kavga içermeyen 300 adet görüntüden oluşmaktadır. Görüntüler ortalama 2 saniye uzunluğundadır. Movie Violence, filmlerden derlenen veri kümesidir. Filmlerde karşılaşılan şiddet içeren görüntülerden toparlanmıştır. Farklı sahnelerde çekilmiş 200 kavga görüntüsü içermektedir.
- **Amaç 8:** Güvenlik kamerası anlık görüntülerinde derin öğrenme temelli şiddet olayı görüntülerinin sınıflandırılması için deneysel çalışmanın yapılması.
- **Sonuç 8:** Şiddet olayı görüntülerinin sınıflandırılması için veri kümeleri ile ayrı ayrı deneysel çalışmalar yapılmıştır. Bütün veri kümeleri için ESA tabanlı özellik standart sapma tabanlı derin model, özellik ortalama havuzlama tabanlı derin model ve özellik standart sapma tabanlı derin model kullanılmıştır. Bu üç model neticesinde üç veri kümesinde de en iyi başarımlar ESA tabanlı özellik standart sapma tabanlı derin model ile elde edilmiştir. Başarımların değerleri tablo 5.4'te gösterilmiştir.

Tablo 5.4: Şiddet görüntülerinin sınıflandırılması modeller ve başarımlar değerleri.

Veri Kümesi	Model	Görüntü	Hassaslık	Özgüllük	F1-Skoru	Doğruluk
Surveillance Camera Fight	Özellik	5	99.32	98.03	98.66	98.67
	Birleştirme	10	98.66	98.01	98.33	98.33
	tabanlı					
	derin					
	model					
	Özellik	5	98.63	96.10	97.30	97.33
	Ortalama	10	99.32	98.03	98.66	98.67
	Havuzlama					
	tabanlı					
	derin					
Crowd Violence	model					
	Özellik	5	99.34	100	99.67	99.67
	Standart	10	99.33	99.33	99.33	99.33
	Sapma					
	tabanlı					
	derin					
	model					
	Özellik	5	93.50	93.50	93.50	93.50
	Birleştirme	10	95.12	95.12	95.12	95.12
	tabanlı					
Movie Violence	derin model					
	Özellik	5	97.50	95.24	96.30	96.34
	Ortalama	10	98.35	96.80	97.54	97.56
	Havuzlama					
	tabanlı					
	derin					
	model					
	Özellik	5	95.20	96.69	95.97	95.93
	Standart	10	98.39	99.18	98.79	98.78
	Sapma					
Movie Violence	tabanlı					
	derin					
	model					
	Özellik	5	97.98	97.03	97.49	97.50
	Birleştirme	10	98	98	98	98
	tabanlı					
Movie Violence	derin model					
	Özellik	5	100	99.01	99.50	99.50
	Ortalama	10	10	99.01	99.50	99.50
	Havuzlama					
	tabanlı					
	derin					
	model					

Movie Violence	Özellik	5	100	100	100	100
	Standart	10	100	100	100	100
	Sapma					
	tabanlı					
	derin					
	model					

- **Amaç 9:** Güvenlik kamerası görüntülerinde görüntü analizi ve boyut indirgeme uygulaması ve kullanıcı arayüzünün geliştirilmesi.
- **Sonuç 9:** Güvenlik kamera kayıtlarının analiz edilebilmesi ve boyutlarının indirgenebilmesi için kullanıcı etkileşimli, insan bilgisayar etkileşimi kurallarına uygun arayüz uygulaması çalışması yapılmıştır. Şekil 5.1’de görüntü analizi ve boyut indirgeme kullanıcı arayüzünün genel özellikli yapısı gösterilmiştir.



Şekil 5.1: Görüntü analizi ve boyut indirgeme kullanıcı arayüzü uygulaması.

Görüntü analizi ve boyut indirgeme arayüzü uygulamasında birçok özellik tanımlanmıştır. Bu özellikler;

- İki ayrı sekme bulunmaktadır. Bu sekmeler Canlı Kayıt ve Boyut İndirgeme sekmesi ve Görüntü Analizi sekmesidir. Canlı kayıt ve Boyut İndirgeme sekmesinde, canlı olarak güvenlik kamerası görüntüsünün takip edilebilmesi, arşiv görüntülerin izlenebilmesi ve boyut indirgeme işleminin takip edilebilmesi özellikleri bulunmaktadır. Görüntü Analizi sekmesinde, görüntü kayıtlarının analiz edilmesi ve anormal durumların tespit edilmesi özellikleri bulunmaktadır.
- Canlı kayıtların takibinin aktif ekranda aktif edilebilmesi için Canlı Kayıt butonu bulunmaktadır.

- Eski kayıt görüntülerinin ekranda izlenmesi için Arşivler butonu bulunmaktadır.
- Boyut indirgeme sisteminin başlatılması için Boyut İndirgeme butonu bulunmaktadır.
- Görüntü analizi ve boyut indirgeme uygulamasının kapatılması için Kapat butonu bulunmaktadır.
- Sistem üreticisi bilgilendirilmesinin yapılması, yönlendirme ve klavuz işlemlerinin yapılabilmesi için Anasayfa menüsü bulunmaktadır.
- Kullanıcı ve boyut indirgeme sistemi hassasiyet ayarlamalarının yapılabilmesi için Ayarlar menüsü bulunmaktadır.
- Kazanımlar ve sistem çıktıları ile ilgili bilgilendirmeler yapılabilmesi için Kazanımlar menüsü bulunmaktadır.
- Günlük anlık olay bilgilendirilmesinin yapılabilmesi için Olay Günlüğü menüsü bulunmaktadır.
- Arşiv olay kayıtlarının takip edilebilmesi için Arşiv Olay Günlüğü menüsü bulunmaktadır.
- Sistem yöneticisi ile kullanıcının yardımlaşmasına imkan tanıyan Destek menüsü bulunmaktadır.
- Anlık sistem olay bilgisinin kullanıcıya bildirilmesi için Anlık Olay Tespiti mesaj kutusu bulunmaktadır.
- Sistemin görüntü analizlerinin başlatılabilmesi için Görüntü Analizi sekmesi ekranında yönlendirici anahtarlar ve gizli butonlar bulunmaktadır.
- Sistem arayüzü uygulaması, iş çakışmalarının önlenmesi için akıllı algoritmalar ile işlenmiştir. Bütün buton, menü ve anahtarlar kendi aralarında etkileşimli olarak kodlanmıştır.
- Kullanıcının eğitim almadan sistemi kullanışlı ve ergonomik kullanabilmesi için sayfa düzenlemesinde basitlik ve renklerde göz rahatlığına uygun renkler seçilmiştir.
- Sistem çalıştırılmasında harici bir programa ihtiyaç olmaksızın sistem kurulum (Kurulum.exe) dosyası oluşturulmuştur. Kurulum yapılan bilgisayarda kamera bağlantısının otomatik yapılması için gerek program paketleri uygulama kurulum dosyasına gömülü program olarak yerleştirilmiştir. Kolay kurulum ve

tanıtım görseli hazırlanmıştır. Kolay kurulum ve tanıtım görseli şekil 5.2' de gösterilmiştir.



Şekil 73: Görüntü analizi ve boyut indirgeme kolay kurulum ve tanıtım görseli.

Sonuç olarak, bu tez çalışmasında, şiddet, silah ve bıçak tespiti için kapsamlı deneysel çalışmalar gerçekleştirilmiştir. Bu çalışmalardan elde edilen sonuçlar kısaca şu şekilde özetlenebilir:

- Silah ve bıçak sınıfı tespiti iki sınıflı olarak yapılmıştır. Bu çalışmada, YOLOR modeli kullanılarak %97,6 başarımla elde edilmiştir.
- Videolarda şiddet tespiti için EfficientDet modeli, %94,1 skor üretmiştir.
- Silah ve bıçak sınıflandırılması çalışmasında VGG16 modelinin daha önceleri eğitilmiş ağırlıklarının kullanılması yöntemiyle tekrar eğitim yapılmış ve %99,24 doğruluk değeri elde edilmiştir.
- Videolarda şiddet sınıflandırılması için iki ayrı model geliştirilip kullanılmıştır. Birinci modelde ESA mimarisinin havuzlama katmanında özellik standart sapma

yöntemi kullanılarak Surveillance veri kümesinde %99,67, Crowd veri kümesinde %98,78 ve Movie veri kümesinde %100 doğruluk değeri elde edilmiştir. İkinci şiddet sınıflandırılması çalışmasında ise ESA, LSTM ve GRU mimarilerinin kombinasyonuna dayalı bir yaklaşım geliştirilmiştir. Sonuç olarak, önerilen model Crowd ve Surveillance veri kümelerinde sırasıyla %100 ve % 98,33 doğruluk değerleri elde edilmiştir.

Bu çalışmalardan elde edilen tespit işlemleri, güvenlik kamera sistemlerinin otomasyon sistemlerine bağlanması planlanmıştır. Görüntü sınıflandırma çalışmalarının ise bu tez çalışmasında geliştirilen Görüntü Analizi ve Boyut İndirgeme uygulamasına bağlanması sağlanmıştır. Uygulama üzerinden doğrudan kayıt edilmiş görüntülerde otomatik şiddet, silah ve bıçak sınıflandırılması işlemi yapılabilmektedir. Şekil 3.29'da gösterilen boyut indirgeme modeli algoritmasına dayalı boyut indirgemesi modeli geliştirilmiştir. Bu model gece ve hareketsiz ortam kayıtlarında %10-%40 arasında boyuttan kazanım sağlamaktadır. Bu modelde Görüntü Analizi ve Boyut İndirgeme uygulamasına bağlanmıştır. Uygulama üzerinden sistem kullanıcısı boyut indirgeme işlemi yapabilecektir. Ayrıca uygulama üzerinden canlı ve arşiv kamera kayıtlarına erişebilmekte ve izleyebilmektedir. Uygulama üzerinde modül olarak, anasayfa, ayarlar ve kullanıcı doğrulama, günlük olay kaydı, arşiv olay kaydı ve mail destekli sistem- kullanıcı etkileşimli destek modülü sistemleri yapılmıştır. Görüntü olay algılama sınıflandırmalarında otomatik mail desteği sisteme eklenmiştir.

Bu tez çalışmasında geliştirilen boyut indirgeme modeli, güvenlik kamera sistemlerinde kullanılması konusunda hukuki geçerlilik sınırlılığı bulunmaktadır. Bu sınırlılığın aşılması ve ticari olarak kullanıma geçebilmesi için ise patent ve ülkesel geçerlilik belgesine ilerleyen dönemlerde başvuru yapılacaktır.

5.2 Öneriler

Tez çalışması süresince yapılmış olan araştırmalar ve deneysel çalışmalar neticelerine dayalı olarak ileri zamanlarda yapılması düşünülen çalışmalar şunlardır.

- Bu tez çalışmasında geliştirilen boyut indirgeme işleminin, hukuki geçerlilik konusunda netlik sınırlılığı bulunmaktadır. Bu sınırlılığın aşılması için patent ve ispatlanmış deneyler ile geçerlilik sertifikalarının alınması gerekmektedir.

- Türkiye şartlarında sektör tarzı özgün veri kümeleri ile deneysel çalışma ve analiz türlerinin kapsamının genişletilmesi düşünülebilir.
- Görüntü analizi ve boyut indirgeme uygulamasının mobil cihazlarda kullanılabilir hale getirilmesi yapılabilir.
- Sistemin sektörel olarak günlük hayatta kullanılması şartlarının oluşması durumunda web sayfası tasarlanarak sistem yönetici paneli yapılabilir.



KAYNAKLAR

- Aktı, Ş., Tataroğlu, G. A., & Ekenel, H. K. (2019). Vision-based Fight Detection from Surveillance Cameras. *2019 Ninth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, 1-6. <https://doi.org/10.1109/IPTA.2019.8936070>.
- Arı, A., & Hanbay, D. (2019). Tumor detection in MR images of regional convolutional neural networks. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 34(3), 1395-1408.
- Bahad, P., Saxena, P., & Kamal, R. (2019). Fake News Detection using Bi-directional LSTM-Recurrent Neural Network. *Procedia Computer Science*, 165, 74-82. <https://doi.org/10.1016/j.procs.2020.01.072>.
- Balagourouchetty, L., Pragatheeswaran, J. K., Pottakkat, B., & Ramkumar, G. (2020). GoogLeNet-Based Ensemble FCNet Classifier for Focal Liver Lesion Diagnosis. *IEEE Journal of Biomedical and Health Informatics*, 24(6), 1686-1694. <https://doi.org/10.1109/JBHI.2019.2942774>.
- Bansal, M., Kumar, M., Sachdeva, M., & Mittal, A. (2021). Transfer learning for image classification using VGG19: Caltech-101 image data set. *Journal of Ambient Intelligence and Humanized Computing*. <https://doi.org/10.1007/s12652-021-03488-z>.
- Başkurt, B. (2019). *10235829.Pdf* [Doktora Tezi]. Ankara Üniversitesi.
- Bengio, Y. (2009). Learning Deep Architectures for AI. *Foundations and Trends® in Machine Learning*, 2(1), 1-127. <https://doi.org/10.1561/22000000006>
- Bengio, Y., Lecun, Y., & Hinton, G. (2021). Deep learning for AI. *Communications of the ACM*, 64(7), 58-65.
- 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>.
- Chakraborty, S., Tickoo, O., & Iyer, R. (2015). Adaptive Keyframe Selection for Video Summarization. *2015 IEEE Winter Conference on Applications of Computer Vision*, 702-709. <https://doi.org/10.1109/WACV.2015.99>.
- Che-Bin Liu, & Ahuja, N. (2004). Vision based fire detection. *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, 4, 134-137 Vol.4. <https://doi.org/10.1109/ICPR.2004.1333722>.
- Cheng, M., Cai, K., & Li, M. (2020). RWF-2000: An Open Large Scale Video Database for Violence Detection. *ArXiv:1911.05913 [Cs]*. <http://arxiv.org/abs/1911.05913>.

- Cheng, M., Cai, K., & Li, M. (2021). RWF-2000: An Open Large Scale Video Database for Violence Detection. *2020 25th International Conference on Pattern Recognition (ICPR)*, 4183-4190. <https://doi.org/10.1109/ICPR48806.2021.9412502>.
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
- Chuang, C.-H., Hsieh, J.-W., & Fan, K.-C. (2007). Suspicious Object Detection and Robbery Event Analysis. *2007 16th International Conference on Computer Communications and Networks*, 1189-1192. <https://doi.org/10.1109/ICCCN.2007.4317981>.
- Crews, K. D. (1995). Copyright law and information policy planning: Public rights of use in the 1990s and beyond. *Journal of Government Information*, 22(2), 87-99. [https://doi.org/10.1016/1352-0237\(94\)00041-7](https://doi.org/10.1016/1352-0237(94)00041-7).
- Çetin, A. E., Dimitropoulos, K., Gouverneur, B., Grammalidis, N., Günay, O., Habiboğlu, Y. H., Töreyin, B. U., & Verstockt, S. (2013). Video fire detection – Review. *Digital Signal Processing*, 23(6), 1827-1843. <https://doi.org/10.1016/j.dsp.2013.07.003>.
- Doğan, A., Calp, M. H., Arı, E. S., & Özköse, H. (2015). İnsan Bilgisayar Etkileşimi Kapsamında Beyin Bilgisayar Arayüzleri Üzerine bir İnceleme: Özellikleri ve Çalışma Prensipleri. *Yönetim Bilişim Sistemleri Dergisi*, 1(2), 1-10.
- DOĞAN, F., & TÜRKOĞLU, İ. (2019). Derin öğrenme modelleri ve uygulama alanlarına ilişkin bir derleme. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, 10(2), 409-445.
- Dwivedi, N., Singh, D. K., & Kushwaha, D. S. (2019). Weapon Classification using Deep Convolutional Neural Network. *2019 IEEE Conference on Information and Communication Technology*, 1-5. <https://doi.org/10.1109/CICT48419.2019.9066227>.
- Fernandez-Carrobles, M. M., Deniz, O., & Maroto, F. (2019). Gun and Knife Detection Based on Faster R-CNN for Video Surveillance. İçinde A. Morales, J. Fierrez, J. S. Sánchez, & B. Ribeiro (Ed.), *Pattern Recognition and Image Analysis* (ss. 441-452). Springer International Publishing. https://doi.org/10.1007/978-3-030-31321-0_38.
- Gao, S., Huang, Y., Zhang, S., Han, J., Wang, G., Zhang, M., & Lin, Q. (2020). Short-term runoff prediction with GRU and LSTM networks without requiring time step optimization during sample generation. *Journal of Hydrology*, 589, 125188. <https://doi.org/10.1016/j.jhydrol.2020.125188>.
- Gao, Y., Liu, H., Sun, X., Wang, C., & Liu, Y. (2016). Violence detection using Oriented Violent Flows. *Image and Vision Computing*, 48-49, 37-41. <https://doi.org/10.1016/j.imavis.2016.01.006>.
- Ge, Z., Liu, S., Wang, F., Li, Z., & Sun, J. (2021). YOLOX: Exceeding YOLO Series in 2021. <https://doi.org/10.48550/ARXIV.2107.08430>.

- Görmüş, E. T., & Akar, Ö. (2019). Headwall Hyperspec VNIR Kamerası İle Elde Edilen Hiperspektral Hava Fotoğrafı İçin Boyut İndirgeme Yöntemlerinin Performanslarının Analizi. *Geomatik*, 4(3), 201-214. <https://doi.org/10.29128/geomatik.538838>.
- Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2017). LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10), 2222-2232. <https://doi.org/10.1109/TNNLS.2016.2582924>.
- Güçlü, E., Aydın, İ., Şahbaz, K., Akin, E., & Karaköse, M. (2021). Demiryolu Bağlantı Elemanlarında Bulunan Kusurların YOLOv4 ve Bulanık Mantık Kullanarak Tespiti. *Demiryolu Mühendisliği*, 14, Art. 14. <https://doi.org/10.47072/demiryolu.939830>.
- Hassner, T., Itcher, Y., & Kliper-Gross, O. (2012). Violent flows: Real-time detection of violent crowd behavior. *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 1-6. <https://doi.org/10.1109/CVPRW.2012.6239348>.
- 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. <https://doi.org/10.1109/TPAMI.2015.2389824>.
- 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.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4700-4708.
- Hussain, T., Muhammad, K., Ullah, A., Cao, Z., Baik, S. W., & de Albuquerque, V. H. C. (2020). Cloud-Assisted Multiview Video Summarization Using CNN and Bidirectional LSTM. *IEEE Transactions on Industrial Informatics*, 16(1), 77-86. <https://doi.org/10.1109/TII.2019.2929228>.
- İNiK, Ö., & Ülker, E. (2017). *Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri*. 20.
- Ioffe, S., & Szegedy, C. (2015, June). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning* (pp. 448-456). PMLR.
- Jain, H., Vikram, A., Mohana, Kashyap, A., & Jain, A. (2020). Weapon Detection using Artificial Intelligence and Deep Learning for Security Applications. *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 193-198. <https://doi.org/10.1109/ICESC48915.2020.9155832>.
- Jiang, B., Xu, F., Tu, W., & Yang, C. (2019). Channel-wise Attention in 3D Convolutional Networks for Violence Detection. *2019 International Conference on Intelligent*

- Computing and its Emerging Applications (ICEA)*, 59-64. <https://doi.org/10.1109/ICEA.2019.8858306>.
- Joshi, T., Aarya, H., & Kumar, P. (2016). Suspicious object detection. *2016 2nd International Conference on Advances in Computing, Communication, & Automation (ICACCA) (Fall)*, 1-6. <https://doi.org/10.1109/ICACCAF.2016.7748964>.
- Kakadiya, R., Lemos, R., Mangalan, S., Pillai, M., & Nikam, S. (2019). AI Based Automatic Robbery/Theft Detection using Smart Surveillance in Banks. *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)*, 201-204. <https://doi.org/10.1109/ICECA.2019.8822186>.
- Kamthe, U. M., & Patil, C. G. (2018). Suspicious Activity Recognition in Video Surveillance System. *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBE)*, 1-6. <https://doi.org/10.1109/ICCUBE.2018.8697408>.
- Kang, M.-S., Park, R.-H., & Park, H.-M. (2021). Efficient Spatio-Temporal Modeling Methods for Real-Time Violence Recognition. *IEEE Access*, 9, 76270-76285. <https://doi.org/10.1109/ACCESS.2021.3083273>.
- Kibria, S. B., & Hasan, M. S. (2017). An analysis of Feature extraction and Classification Algorithms for Dangerous Object Detection. *2017 2nd International Conference on Electrical Electronic Engineering (ICEEE)*, 1-4. <https://doi.org/10.1109/CEEE.2017.8412846>.
- Kizrak, M. A., & Bolat, B. (2018). Derin öğrenme ile kalabalık analizi üzerine detaylı bir araştırma. *Bilişim Teknolojileri Dergisi*, 11(3), 263-286.
- Kmieć, M., Głowacz, A., & Dziech, A. (2012). Towards Robust Visual Knife Detection in Images: Active Appearance Models Initialised with Shape-Specific Interest Points. İçinde A. Dziech & A. Czyżewski (Ed.), *Multimedia Communications, Services and Security* (ss. 148-158). Springer. https://doi.org/10.1007/978-3-642-30721-8_15.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- Li, P., & Zhao, W. (2020). Image fire detection algorithms based on convolutional neural networks. *Case Studies in Thermal Engineering*, 19, 100625. <https://doi.org/10.1016/j.csite.2020.100625>.
- Li, H., Lin, Z., Shen, X., Brandt, J., & Hua, G. (2015). A convolutional neural network cascade for face detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 5325-5334).
- Liang, S., Zhang, R., Liang, D., Song, T., Ai, T., Xia, C., Xia, L., & Wang, Y. (2018). Multimodal 3D DenseNet for IDH Genotype Prediction in Gliomas. *Genes*, 9(8), Art. 8. <https://doi.org/10.3390/genes9080382>.

- Luo, M., & Zhang, K. (2014). A hybrid approach combining extreme learning machine and sparse representation for image classification. *Engineering Applications of Artificial Intelligence*, 27, 228-235.
- Maksimova, A., Matiolański, A., & Wassermann, J. (2014). Fuzzy Classification Method for Knife Detection Problem. İçinde A. Dziech & A. Czyżewski (Ed.), *Multimedia Communications, Services and Security* (C. 429, ss. 159-169). Springer International Publishing. https://doi.org/10.1007/978-3-319-07569-3_13.
- Marbach, G., Loepfe, M., & Brupbacher, T. (2006). An image processing technique for fire detection in video images. *Fire Safety Journal*, 41(4), 285-289. <https://doi.org/10.1016/j.firesaf.2006.02.001>.
- Masala, E., Yu, Y., & He, X. (2014). Content-based group-of-picture size control in distributed video coding. *Signal Processing: Image Communication*, 29(3), 332-344. <https://doi.org/10.1016/j.image.2014.01.013>.
- Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50(4), 370-396. <https://doi.org/10.1037/h0054346>.
- Matiolański, A., Maksimova, A., & Dziech, A. (2016). CCTV object detection with fuzzy classification and image enhancement. *Multimedia Tools and Applications*, 75(17), 10513-10528. <https://doi.org/10.1007/s11042-015-2697-z>.
- Mery, D., Svec, E., Arias, M., Rizzo, V., Saavedra, J. M., & Banerjee, S. (2017). Modern Computer Vision Techniques for X-Ray Testing in Baggage Inspection. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 47(4), 682-692. <https://doi.org/10.1109/TSMC.2016.2628381>.
- Metin, İ. A., & Karasulu, B. (2019). *İnsan Aktivitelerinin Sınıflandırılmasında Tekrarlayan Sinir Ağları Kullanan Derin Öğrenme Tabanlı Yaklaşım*. 10.
- Mikolov, T., Karafiát, M., Burget, L., Cernocký, J., & Khudanpur, S. (2010). Recurrent neural network based language model. *Interspeech*, 2(3), 1045-1048.
- N, V. V. (1995). The Nature of Statistical Learning. *Theory*. <https://cir.nii.ac.jp/crid/1570854175550242816>.
- 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>.
- Olmos, R., Tabik, S., Lamas, A., Pérez-Hernández, F., & Herrera, F. (2019). A binocular image fusion approach for minimizing false positives in handgun detection with deep learning. *Information Fusion*, 49, 271-280. <https://doi.org/10.1016/j.inffus.2018.11.015>.
- Pavithra, M., Saruladha, K., & Sathyabama, K. (2019). GRU Based Deep Learning Model for Prognosis Prediction of Disease Progression. 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), 840-844. <https://doi.org/10.1109/ICCMC.2019.8819830>.

- Peixoto, B., Lavi, B., Bestagini, P., Dias, Z., & Rocha, A. (2020). Multimodal Violence Detection in Videos. *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2957-2961. <https://doi.org/10.1109/ICASSP40776.2020.9054018>.
- Pérez-Hernández, F., Tabik, S., Lamas, A., Olmos, R., Fujita, H., & Herrera, F. (2020). Object Detection Binary Classifiers methodology based on deep learning to identify small objects handled similarly: Application in video surveillance. *Knowledge-Based Systems*, 194, 105590. <https://doi.org/10.1016/j.knosys.2020.105590>.
- Rakumthong, W., Phetcharaladakun, N., Wealveerakup, W., & Kamnoonwatana, N. (2014). Unattended and stolen object detection based on relocating of existing object. *2014 Third ICT International Student Project Conference (ICT-ISPC)*, 115-118. <https://doi.org/10.1109/ICT-ISPC.2014.6923231>.
- Review: DenseNet—Dense Convolutional Network (Image Classification) | by Sik-Ho Tsang | Towards Data Science. (t.y.). Geliş tarihi 28 Kasım 2022, gönderen <https://towardsdatascience.com/review-densenet-image-classification-b6631a8ef803>.
- Russ, J. C. (2006). *The Image Processing Handbook* (5. bs). CRC Press. <https://doi.org/10.1201/9780203881095>.
- Salazar González, J. L., Zaccaro, C., Álvarez-García, J. A., Soria Morillo, L. M., & Sancho Caparrini, F. (2020). Real-time gun detection in CCTV: An open problem. *Neural Networks*, 132, 297-308. <https://doi.org/10.1016/j.neunet.2020.09.013>.
- Salur, M. U., Aydin, İ., & Karaköse, M. (2019). Gömülü Derin Öğrenme ile Tehdit İçeren Nesnelerin Gerçek Zamanda Tespiti. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, 10(2), Art. 2. <https://doi.org/10.24012/dumf.492433>.
- Salvi, S., & Kadam, A. (2021). Breast Cancer Detection Using Deep learning and IoT Technologies. *Journal of Physics: Conference Series*, 1831(1), 012030.
- Samuels, I. (1959). Reticular mechanisms and behavior. *Psychological Bulletin*, 56(1), 1-25. <https://doi.org/10.1037/h0044488>.
- Serrano Gracia, I., Deniz Suarez, O., Bueno Garcia, G., & Kim, T.-K. (2015). Fast Fight Detection. *PLOS ONE*, 10(4), e0120448. <https://doi.org/10.1371/journal.pone.0120448>.
- Shaheen, M., Khan, R., Biswal, R. R., Ullah, M., Khan, A., Uddin, M. I., Zareei, M., & Waheed, A. (2021). Acute Myeloid Leukemia (AML) Detection Using AlexNet Model. *Complexity*, 2021, 1-8. <https://doi.org/10.1155/2021/6658192>.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Soliman, M. M., Kamal, M. H., El-Massih Nashed, M. A., Mostafa, Y. M., Chawky, B. S., & Khattab, D. (2019). Violence Recognition from Videos using Deep Learning Techniques. *2019 Ninth International Conference on Intelligent Computing and Information Systems (ICICIS)*, 80-85. <https://doi.org/10.1109/ICICIS46948.2019.9014714>.

- Song, S., Jing, J., Huang, Y., & Shi, M. (2021). EfficientDet for fabric defect detection based on edge computing. *Journal of Engineered Fibers and Fabrics*, 16, 155892502110083. <https://doi.org/10.1177/15589250211008346>.
- Song, W., Zhang, D., Zhao, X., Yu, J., Zheng, R., & Wang, A. (2019). A Novel Violent Video Detection Scheme Based on Modified 3D Convolutional Neural Networks. *IEEE Access*, 7, 39172-39179. <https://doi.org/10.1109/ACCESS.2019.2906275>.
- Su, Y., Lin, G., Zhu, J., & Wu, Q. (2020). Human Interaction Learning on 3D Skeleton Point Clouds for Video Violence Recognition. In A. Vedaldi, H. Bischof, T. Brox, & J.-M. Frahm (Eds.), *Computer Vision – ECCV 2020* (ss. 74-90). Springer International Publishing. https://doi.org/10.1007/978-3-030-58548-8_5.
- Sudhakaran, S., & Lanz, O. (2017). Learning to detect violent videos using convolutional long short-term memory. *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 1-6. <https://doi.org/10.1109/AVSS.2017.8078468>.
- Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. *Thirty-first AAAI conference on artificial intelligence*.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2818-2826. <https://doi.org/10.1109/CVPR.2016.308>.
- Szegedy, C., Wei Liu, Yangqing Jia, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>.
- Talan, T., & Aktürk, C. (t.y.). *Bilgisayar Bilimlerinde Teorik Ve Uygulamalı Araştırmalar*. Efe Akademi Yayınları.
- Tan, M., Pang, R., & Le, Q. V. (2020). EfficientDet: Scalable and Efficient Object Detection. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10778-10787. <https://doi.org/10.1109/CVPR42600.2020.01079>.
- Theckedath, D., & Sedamkar, R. R. (2020). Detecting Affect States Using VGG16, ResNet50 and SE-ResNet50 Networks. *SN Computer Science*, 1(2), 79. <https://doi.org/10.1007/s42979-020-0114-9>.
- Tiwari, R. K., & Verma, G. K. (2015). A Computer Vision based Framework for Visual Gun Detection Using Harris Interest Point Detector. *Procedia Computer Science*, 54, 703-712. <https://doi.org/10.1016/j.procs.2015.06.083>.
- Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). *Learning Spatiotemporal Features With 3D Convolutional Networks*. 4489-4497. https://openaccess.thecvf.com/content_iccv_2015/html/Tran_Learning_Spatiotemporal_Features_ICCV_2015_paper.html.

- Traoré, A., & Akhloufi, M. A. (2020). 2D Bidirectional Gated Recurrent Unit Convolutional Neural Networks for End-to-End Violence Detection in Videos. İçinde A. Campilho, F. Karay, & Z. Wang (Ed.), *Image Analysis and Recognition* (ss. 152-160). Springer International Publishing. https://doi.org/10.1007/978-3-030-50347-5_14.
- Teknoloji, A. W. T. Y. ve. (t.y.). HDD Hard Disk Hesaplama—CENOVA. HDD Hard Disk Hesaplama - CENOVA. Geliş tarihi 1 Kasım 2022, gönderen <https://www.cenova.com.tr/hdd-harddisk-hesaplama>.
- Türkoğlu, M. *İNÖNÜ ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ*. Doktora Tezi, İnönü Üniversitesi, 2019.
- Türkoğlu, M., & Hanbay, D. (2019). Plant disease and pest detection using deep learning-based features. *Turkish Journal of Electrical Engineering and Computer Sciences*, 27(3), 1636-1651.
- Verma, G. K., & Dhillon, A. (2017). A Handheld Gun Detection using Faster R-CNN Deep Learning. *Proceedings of the 7th International Conference on Computer and Communication Technology*, 84-88. <https://doi.org/10.1145/3154979.3154988>.
- Verma, S. K., & Tiwari, A. K. (2015). Progress of nanofluid application in solar collectors: A review. *Energy Conversion and Management*, 100, 324-346. <https://doi.org/10.1016/j.enconman.2015.04.071>.
- Wang, C.-Y., Liao, H.-Y. M., Wu, Y.-H., Chen, P.-Y., Hsieh, J.-W., & Yeh, I.-H. (2020). CSPNet: A New Backbone That Can Enhance Learning Capability of CNN. 390-391. https://openaccess.thecvf.com/content_CVPRW_2020/html/w28/Wang_CSPNet_A_New_Backbone_That_Can_Enhance_Learning_Capability_of_CVPRW_2020_paper.html.
- Wang, C.-Y., Yeh, I.-H., & Liao, H.-Y. M. (2021). You Only Learn One Representation: Unified Network for Multiple Tasks. <https://doi.org/10.48550/ARXIV.2105.04206>
- Xu, Y., Yang, G., Luo, J., & He, J. (2020). An Electronic Component Recognition Algorithm Based on Deep Learning with a Faster SqueezeNet. *Mathematical Problems in Engineering*, 2020, 1-11. <https://doi.org/10.1155/2020/2940286>.
- Yakıt, O., & Özkan, Y. (2017). *Kurumsal Kaynak Planlama Sistemlerinde Yapay Sinir Ağlarının Değerlendirilmesi Yaklaşımı*. Siyaset, Ekonomi ve Yönetim Araştırmaları Dergisi, 5(2), 287-296.
- Yildiz, Z., Danişmani, T., & Telatar, D. Z. (2021). *Çoklu Spektural Bantlardaki Görüntüleri Kullanarak Derin Öğrenme ile Nesne Algılama ve Tanıma*. 94.
- Yılmaz, O., Aydın, H., & Çetinkaya, A. (2020). Faster R-CNN Üzerinde Geliştirilen Model ile Object Detection Api Üzerinde Doğruluk Tahmini ve Analizi. *European Journal of Science and Technology*. <https://doi.org/10.31590/ejosat.753896>.
- YOLOR object detection in Gradient. (2022, Nisan 27). Paperspace Blog. <https://blog.paperspace.com/yolor/>

- Zhang, S., Song, L., Liu, S., Ge, Z., Li, Z., He, X., & Sun, J. (2021). *Workshop on Autonomous Driving at CVPR 2021: Technical Report for Streaming Perception Challenge*. <https://doi.org/10.48550/ARXIV.2108.04230>.
- Zhang, T., Yang, Z., Jia, W., Yang, B., Yang, J., & He, X. (2016). A new method for violence detection in surveillance scenes. *Multimedia Tools and Applications*, 75(12), 7327-7349. <https://doi.org/10.1007/s11042-015-2648-8>.
- Zhang, X., & Shen, Z. (2018). Copyright protection method for 3D model of geological body based on digital watermarking technology. *Journal of Visual Communication and Image Representation*. <https://doi.org/10.1016/j.jvcir.2018.12.013>.



ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı Soyadı	Mehmet Tevfik AĞDAŞ
Doğum Yeri	-
Doğum Tarihi	-
Uyruğu	☒ T.C. ☐ Diğer:
E-Posta Adresi	-
Web Adresi	-

Eğitim Bilgileri	
Lisans	
Üniversite	Fırat ÜNİVERSİTESİ
Fakülte	Teknik Eğitim Fakültesi
Bölümü	Bilgisayar Öğretmenliği
Mezuniyet Yılı	2011

Yüksek Lisans	
Üniversite	Fırat ÜNİVERSİTESİ
Enstitü Adı	Fen Bilimleri ENSTİTUSU
Anabilim Dalı	Elektronik ve Bilgisayar Eğitimi
Programı	Telekomünikasyon

Doktora	
Üniversite	İstanbul Üniversitesi
Enstitü Adı	Fen Bilimleri Enstitüsü
Anabilim Dalı	Enformatik Anabilim Dalı
Programı	Enformatik Programı

Makale ve Bildiriler	
Ağdaş, M. T., Türkoğlu, M., & Gülseçen, S. (2021). <i>Deep neural networks based on transfer learning approaches to classification of gun and knife images</i> . Sakarya University journal of computer and information sciences, 4(1), 131-141.	
Ağdaş, M. T., & Gülseçen, S. (2022). <i>Güvenlik Kameralarında Otomatik Silah ve Bıçak Tespit Sistemi: Karşılaştırmalı YOLO Modelleri</i> . Avrupa Bilim ve Teknoloji Dergisi, (41), 16-22.	
Ağdaş, M. T. (2021). Copyright Sensitivity in Choosing Ways to Acquire Visual Data on the Internet Environment.	
Ağdaş, M. T., & Arık, K. (2022). Classification of Alzheimer's Disease Using Visual Based ViT (Vision Transformers) Model.	

Ağdaş, M. T., & Karabatak, M. (2013). *Google Tabanlı Siber Suç Haritaları*. Book, P. 1 ST INTERNATIONAL SYMPOSIUM ON DIGITAL FORENSICS AND SECURITY.

