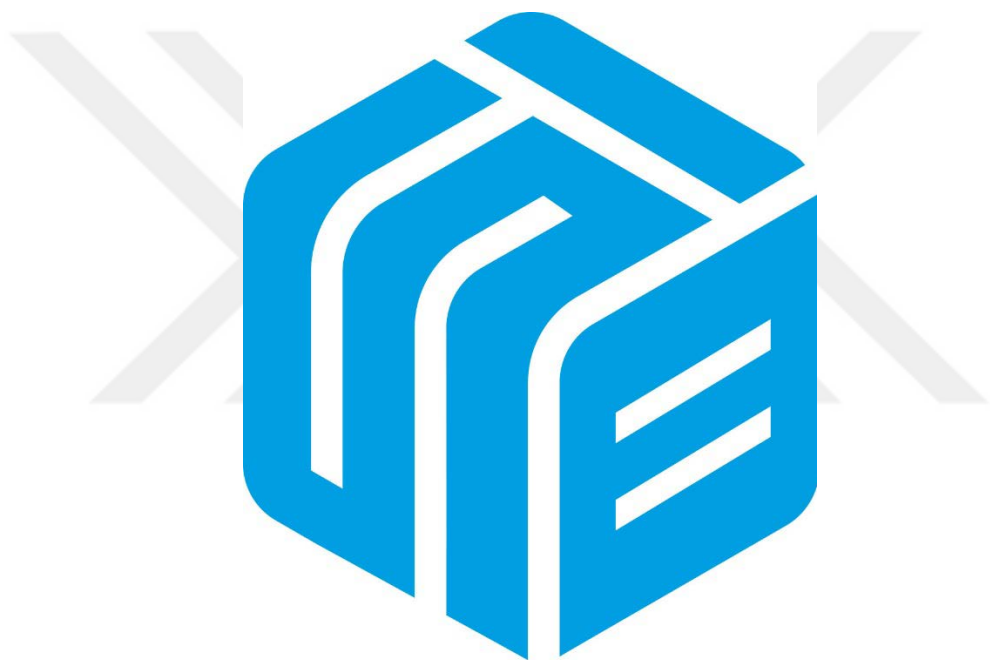




DERİN ÖĞRENME TABANLI KIZIL ÖTESİ HEDEF TESPİTİ

Kevser İrem DANACI

YÜKSEK LİSANS TEZİ

SAVUNMA TEKNOLOJİLERİ ANA BİLİM DALI

SİVAS BİLİM VE TEKNOLOJİ ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

OCAK 2023

Kevser İrem DANACI tarafından hazırlanan “DERİN ÖĞRENME TABANLI KIZIL ÖTESİ HEDEF TESPİTİ” adlı tez çalışması aşağıdaki jüri tarafından OY ÇOKLUĞU ile Sivas Bilim ve Teknoloji Üniversitesi Savunma Teknolojileri Ana Bilim Dalında Yüksek Lisans Tezi olarak kabul edilmiştir.

Danışman: Doç. Dr. Erdem AKAGÜNDÜZ

Çokluortam Bilişimi, Orta Doğu Teknik Üniversitesi

.....

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum/onaylamıyorum.

Başkan: Doç. Dr. Kemal ADEM

Savunma Teknolojileri, Sivas Bilim ve Teknoloji Üniversitesi

.....

Bu tezin, kapsam ve kalite olarak Yüksek Lisans olduğunu onaylıyorum/onaylamıyorum.

Üye: Dr. Öğr. Üyesi Kholoud ELMABRUK

Savunma Teknolojileri, Sivas Bilim ve Teknoloji Üniversitesi

.....

Bu tezin, kapsam ve kalite olarak Yüksek Lisans olduğunu onaylıyorum/onaylamıyorum.

Tez Savunma Tarihi: 19/01/2023

Jüri tarafından kabul edilen bu çalışmanın Yüksek Lisans Tezi olması için gerekli şartları yerine getirdiğini onaylıyorum

.....

Doç. Dr. Emre BİÇER
Lisansüstü Eğitim Enstitüsü Müdürü

ETİK BEYAN

Sivas Bilim ve Teknoloji Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

.....
Kevser İrem DANACI
19/01/2023

DERİN ÖĞRENME TABANLI KIZIL ÖTESİ HEDEF TESPİTİ

(Yüksek Lisans Tezi)

Kevser İrem DANACI

SİVAS BİLİM ve TEKNOLOJİ ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

Ocak 2023

ÖZET

Bilgisayar teknolojisi giderek gelişmekte ve günümüzde birçok alanda kullanılmaktadır. Bu alanlardan en önemlilerinden biri bilgisayarlı görüdür. Gelişmiş bilgisayarlı görü uygulamaları sayesinde görüntülerden nesne tespit çalışmaları yapılmaktadır. Genelde renkli görüntüler üzerinden yapılan nesne tespit çalışmaları, termal algılayıcı teknolojisinin gelişmesiyle birlikte kızılötesi görüntüler üzerinde çalışılmaya başlanmıştır. Kızılötesi görüntüler, başta savunma teknolojileri ve sağlık olmak üzere günlük hayatta dahil birçok alanda yaygın olarak kullanılmaktadır. Yetersiz aydınlatma ve kötü hava koşullarında gerçekleştirilen nesne tespit çalışmalarında kızılötesi görüntülerin kullanılması büyük avantaj sağlamaktadır. Derin öğrenme mimarilerinin giderek gelişmesi ile birlikte nesne tespit uygulamalarında Evrişimli Sinir Ağları yaygın olarak kullanılmaya başlanmıştır. Son zamanlarda literatürde en çok kullanılan, nesne tespit hızı ve doğruluk oranı ile göze çarpan Evrişimli Sinir Ağı modelleri YOLO modelleridir. Bu tez çalışmasında savunma teknolojilerinde kullanılmak üzere, kızılötesi görüntülerden insan ve araç tespiti gerçekleştirilmiştir. Çalışma kapsamında YOLOv3, YOLOv4 ve YOLOv7 modellerinin kızılötesi nesne tespit performansları ve termal alan adaptasyonları kıyaslanmıştır. Bu bağlamda veri çoğaltma yöntemleri ile elde edilen 48000 görüntüden oluşan bir veri seti kullanılmıştır. Veri seti 4 kat çapraz doğrulama yöntemi ile modellere verilmiştir. Her model 100 adım eğitilmiş ve sonuçlar değerlendirilmiştir. Deneysel sonuçlar doğrultusunda YOLOv3 ve YOLOv4 modellerinin en iyi doğruluğa sahip olduğu gözlenmiştir. Modellerin mAP değerleri YOLOv3, YOLOv4 ve YOLOv7 için sırasıyla %84,93; %95,97 ve %57,29 olarak elde edilmiştir.

Bilim Kodu : 92432

Anahtar Kelimeler : Nesne Tespiti, Kızılötesi Görüntüleme, YOLO, Derin Öğrenme, Evrişimli Sinir Ağları

Sayfa Adedi : 88

Danışman : Doç. Dr. Erdem AKAGÜNDÜZ

DEEP LEARNING-BASED INFRARED TARGET DETECTION

(M. Sc. Thesis)

Kevser İrem DANACI

SİVAS UNIVERSITY OF SCIENCE AND TECHNOLOGY

INSTITUTE OF GRADUATE STUDIES

January 2023

ABSTRACT

Computer technology is developing gradually and is used in many fields today. One of the most important of these areas is computer vision. Thanks to advanced computer vision applications, object detection studies are carried out from images. Object detection studies, which are generally done on colour images, have started to work on infrared images with the development of thermal sensor technology. Infrared images are widely used in many areas, especially in defence technologies, medicine, and daily life. The use of infrared images provides a great advantage in object detection studies carried out in bad lighting and weather conditions. With the gradual development of deep learning architectures, Convolutional Neural Networks have been widely used in object detection. The YOLO models are the most widely used Convolutional Neural Network models in the literature, which stand out with their object detection speed and accuracy. In this study, people and vehicles were detected from infrared images to be used in defence technologies. In the scope of the study, infrared object detection performances and thermal domain adaptations of YOLOv3, YOLOv4 and YOLOv7 models were compared. In this context, a dataset consisting of 48000 images obtained by data augmentation methods was used. The dataset was given to the models with the 4-fold cross-validation method. Each model has trained 100 epochs and the results were evaluated. In the experimental results, it was observed that the YOLOv3 and YOLOv4 models had the best accuracy. The mAP of the models were obtained as 84.93%, 95.97% and 57.29% for YOLOv3, YOLOv4 and YOLOv7, respectively.

Science Code : 92432

Key Words : Object Detection, Infrared Imaging, YOLO, Deep Learning, Convolutional Neural Networks

Page Number : 88

Supervisor : Assoc. Prof. Dr. Erdem AKAGÜNDÜZ

TEŞEKKÜR

Kıymetli danışman hocam Doç. Dr. Erdem AKAGÜNDÜZ'e, yüksek lisans sürecim boyunca verdiği tüm destek ve motivasyon için, tecrübeleriyle beni yönlendirip, şu anki bilgi birikimime sahip olmamda bana yol gösterdiği için sonsuz saygı ve teşekkürlerimi sunarım.

Tez çalışmamda dahil her konuda en büyük destekçim, değerli eşim, hayat yoldaşım İlhan DANACI'ya her daim yanımda olduğu, ilgisini ve yardımlarını asla esirgemediği için teşekkür ederim.

Bugüne kadar destekleri ve ilgileriyle benim yanımda olan, beni ben yapan canım annem Belma CÖMERT ve canım babam Ergüenal CÖMERT'e en içten sevgilerimi ve teşekkürlerimi sunarım.

İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
ÇİZELGELERİN LİSTESİ.....	ix
ŞEKİLLERİN LİSTESİ.....	x
RESİMLERİN LİSTESİ	xii
SİMGELER VE KISALTMALAR.....	xiii
1. GİRİŞ	1
2. KAYNAK ARAŞTIRMASI	5
3. MATERYAL VE YÖNTEM.....	11
3.1. Teorik Arka Plan.....	11
3.1.1. Görüntü işleme	11
3.1.2. Kızılötesi ışıma	14
3.1.3. Yapay zeka	22
3.1.4. Evrişim tabanlı nesne tespit modelleri	30
3.2. Yöntem Ve Deneyler.....	38
3.2.1. Verinin elde edilmesi	39
3.2.2. Eğitim işlemleri.....	44
3.2.3. Test işlemleri.....	55
4. BULGULAR VE TARTIŞMA	60
4.1. Doğru Pozitif (DP), Yanlış Negatif (YP) ve Yanlış Pozitif (YP)	60
4.2. Hassasiyet ve Duyarlılık Eğrisi.....	61

4.3. Ortalama Hassasiyet (AP)	62
5. SONUÇ VE ÖNERİLER	67
KAYNAKLAR	70
EKLER.....	80
EK-1. Tezde kullanılan derin öğrenme terminolojileri ve İngilizceleri.....	81
EK-2. Yolov3, Yolov4 Ve Yolov7 modellerinin örnek görüntüler üzerinden nesne tespit performansları	84
ÖZGEÇMİŞ	87

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 3.1. Elektromanyetik yelpaze	15
Çizelge 3.2. IR detektörlerin özellikleri.....	18
Çizelge 3.3. Detektörlerin performanslarını ifade eden parametreler.....	18
Çizelge 3.4. Sıklıkla kullanılan bazı lineer olmayan fonksiyonlar	27
Çizelge 3.5. Yolov3, Yolov4, Yolov7 ve Termal Yolo mimarilerinin genel kısımları	36
Çizelge 3.6. Veri setlerinin özellikleri	40
Çizelge 3.7. Çapraz Doğrulama için resimlerin dağılımı	46
Çizelge 3.8. Verinin kısımlarının eğitim, doğrulama ve test verisi olarak ayrılması	47
Çizelge 3.9. Modellere ait bazı hiper parametreler.....	50
Çizelge 3.10. Modellerin doğrulama veri seti için performansları	54
Çizelge 4.1. Modeller için DP, YN ve YP değerleri.....	60
Çizelge 4.2. Yolov3, Yolov4 ve Yolov7 için nesne tespit sonuçları	63
Çizelge 4.3. Modellerin küçük nesneler üzerinden tespitleri.....	65
Çizelge 4.4. Elde edilen Yolov3 modelinin literatürdeki çalışmalarla karşılaştırılması.....	65
Çizelge 4.5. Elde edilen Yolov4 modelinin literatürdeki çalışmalarla karşılaştırılması.....	66
Çizelge 4.6. Elde edilen Yolov7 modelinin literatürdeki çalışmalarla karşılaştırılması.....	66

ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 3.1. Örnek piksel değerleri.....	11
Şekil 3.2. Bir görüntünün kanalları a) kırmızı kanal, b) yeşil kanal ve c) mavi kanal	12
Şekil 3.3. Kızılötesi dalga boyları.....	16
Şekil 3.4. Kızılötesi kamera türleri, a) birinci nesil kameralar, b) ikinci nesil kameralar, c) üçüncü nesil kameralar.....	17
Şekil 3.5. Görüntü oluşumu	19
Şekil 3.6. Dalga boylarına göre iletim değerleri	20
Şekil 3.7. Yapay zeka ve alt dalları.....	22
Şekil 3.8. Bir algılayıcı yapısı.....	22
Şekil 3.9. Bir ESA yapısı	25
Şekil 3.10. 5x5'lik bir veriye 3x3'lük bir filtrenin uygulanması	26
Şekil 3.11. Havuzlama katmanları, a) maksimum havuzlama, b) ortalama havuzlama ve c) toplam havuzlama	29
Şekil 3.12. Yolo'nun çalışma prensibi.....	31
Şekil 3.13. Yolov3 mimarisi	32
Şekil 3.14. Bir hücredeki çapa kutularının gösterimi	33
Şekil 3.15. Yolov4 mimarisi	34
Şekil 3.16. Yolov7 mimarisi	35
Şekil 3.17. Tez çalışması akış şeması	38
Şekil 3.18. Görüntülerdeki insanların boyutlara göre sınıflandırılması.....	41
Şekil 3.19. Görüntülerdeki araçların boyutlara göre sınıflandırılması	41
Şekil 3.20. Artırılmış resim örnekleri a) AAlartdata veri setinden alınan örnek, b) CAMEL veri setinden alınan örnek, c) LTIR veri setinden alınan örnek.....	42
Şekil 3.21. Alan adaptasyonu sınıflandırılması	48

Şekil	Sayfa
Şekil 3.22. Yolov8 hata grafikleri, a) 1.kat için, b) 2.kat için, c) 3.kat için ve d) 4.kat için	51
Şekil 3.23. Yolov4 hata grafikleri, a) 1.kat için, b) 2.kat için, c) 3.kat için ve d) 4.kat için	51
Şekil 3.24. Yolov7 hata grafikleri, a) 1.kat için, b) 2.kat için, c) 3.kat için ve d) 4.kat için	52
Şekil 3.25. 1.kat için Termal Yolo hata grafiği (giriş boyutu 608x608 için).....	52
Şekil 3.26. 1.kat için Termal Yolo hata grafiği (giriş boyutu 416x416 için).....	53
Şekil 3.27. Gerçek konum (yeşil) ve tahmini konumun (kırmızı) kesişme durumu.....	55
Şekil 4.1. Yolov3 için hassasiyet ve duyarlılık eğrisi.....	61
Şekil 4.2. Yolov4 için hassasiyet ve duyarlılık eğrisi.....	61
Şekil 4.3. Yolov7 için hassasiyet ve duyarlılık eğrisi.....	62
Şekil 4.4. Modeller için hassasiyet ve duyarlılık eğrisi	62

RESİMLERİN LİSTESİ

Resim	Sayfa
Resim 3.1. Nesne tespit örneği	13
Resim 3.2. Örnek görüntü a) görünür yelpaze, b) kızılötesi yelpaze	16
Resim 3.3. Örnek yüz görüntüsü (a) görünür yelpaze, (b) SWIR, (c) MWIR ve (d) LWIR21	
Resim 3.4. Örnek KBO değerleri.....	55
Resim 3.5. DP, YN ve YP için örnekler	58



SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Kısaltmalar	Açıklamalar
AA	Alan Adaptasyonu
AP	Average Precision (Ortalama Hassasiyet)
DN	Doğru Negatif
DP	Doğru Pozitif
ESA	Evrişimsel Sinir Ağları
FIR	Far Infrared (Uzak Kızılötesi)
FLIR	Forward Looking Infrared (İleriye Bakış Kızılötesi)
FPN	Feature Pyramid Network (Özellik Piramit Ağı)
IR	Infrared (Kızılötesi)
KBO	Kesişimin Birleşime Oranı
LWIR	Long-Wave Infrared (Uzun Dalga Kızılötesi)
mAP	Mean Average Precision (Genel Ortalama Hassasiyet)
MWIR	Mid-Wave Infrared (Orta Dalga Kızılötesi)
NEP	Noise Equivalent Power (Gürültü Eşdeğer Güç)
NIR	Near Infrared (Yakın Kızılötesi)
ReLU	Rectified Linear Unit (Doğrultulmuş Doğrusal Birim)
ResNet	Residual Neural Network (Artık Sinir Ağı)
RGB	Red, Green and Blue (Kırmızı, Yeşil ve mavi)
RPN	Region Proposal Network (Bölge Teklif Ağı)
SSD	Single Shot Detector (Tek Atış Detektörü)
SWIR	Short-Wave Infrared (Kısa Dalga Kızılötesi)
VGG	Visual Geometry Group (Görsel Geometri Grubu)
YN	Yanlış Negatif
YOLO	You Only Look Once (Sadece Bir Kez Bakarsın)
YP	Yanlış Pozitif
YSA	Yapay Sinir Ağları

1. GİRİŞ

Bilgisayarlı görü teknolojisinin gelişmesiyle birlikte otonom çalışan uygulamalar ön plana çıkmaktadır. Bunun getirisi olarak bilgisayarlı görünün temel dallarından biri olan nesne tespit uygulamaları giderek yaygınlaşmakta ve bu uygulamalarda kullanılacak yöntemler günbegün artmaktadır. Nesne tespit uygulamalarında kullanılan birçok yöntem literatürde bulunsa da, görüntüleri işleme ve görüntülerden özellik çıkarma işlemlerinde kanıtlanmış başarıları sayesinde en yaygın olarak kullanılan yöntem derin öğrenme yöntemidir (Ji vd., 2021).

Nesne algılama problemi genel olarak bir modelin, bir görüntüdeki (veya videodaki) tüm nesneleri ait olabilecekleri sınıflar için olasılık dağılımlarının hesaplanması işlemidir (Mittal vd., 2020). İşlem hızının ve nesneleri tespit etmedeki doğruluk performansının büyük öneme sahip olduğu bu tarz çalışmalarda derin öğrenme mimarilerinden biri olan Evrişimli Sinir Ağları oldukça fazla araştırılmaktadır (Bhowmick vd., 2022). RCNN (Girshick vd., 2013), FastRCNN (Girshick, 2015), MaskRCNN (He vd., 2017), YOLO (Redmon vd., 2016), SSD (W. Liu vd., 2016) vb. nesne algılama uygulamalarında en çok kullanılan Evrişimli Sinir Ağı modelleridir (Kannadaguli, 2020). Birçok versiyonu bulunan YOLO modelleri, üst düzey özellikleri yakalayarak hızlı ve doğru nesne algılama performansı sayesinde en yaygın kullanılan modellerdir (Fan vd., 2021; Huda vd., 2020b).

Nesne tespit problemleri, günümüzde çoğunlukla görünür yelpazede elde edilen görüntüler üzerinden çalışılmaktadır. Fakat yetersiz aydınlatma ve kötü hava şartlarında görüntü almak oldukça zordur (Guo vd., 2019). Bunun yanı sıra, yetersiz aydınlatma veya farklı hava şartlarında örtüşen renkteki ve dokudaki farklı nesneleri renkli görüntülerde ayırt etmek daha da zor bir problem haline gelmektedir. Bu sebeple görünür yelpaze her nesne tespit uygulaması için uygun olmamaktadır. Bu gibi durumlarda mutlak sıfırın üzerinde sıcaklığa sahip her nesnenin yaydığı kızılötesi ışımaları algılayabilen, kızılötesi yelpazede çalışan, termal kızılötesi kameralardan faydalanılır (C. Li vd., 2019; Wu vd., 2021). Bu kameralar ile elde edilen kızılötesi görüntüler, nesne algılama, nesne bölütleme (segmentasyon), sınıflandırma, hareket algılama vb. gibi birçok uygulamada yaygın olarak kullanılmaya başlanmıştır. Bunun yanı sıra kızılötesi görüntülerden insan ve araç tespit çalışmaları gün geçtikçe önem kazanmaktadır (Kannadaguli, 2020). Fakat kızılötesi

görüntüler, görünür görüntülerden farklı olarak nesnenin sıcaklık haritasını çıkarır. Bu da yanında bazı sıkıntıları beraberinde getirmekte ve nesne tespit çalışmalarını zorlaştırmaktadır. Bu zorluklardan en önemlisi, kızılötesi görüntülerin kızılötesi algılayıcıların kaydettiği verinin işlenmiş hali olmasıdır. Bu da çeşitli hava koşullarında ve ortamlarda nesnelerin farklı şekillerde ifade edilmesine sebep olur (Tumas vd., 2021).

Günümüzde araştırılan nesne tespit uygulamalarında, küçük boyutlu nesneleri tespit edebilme yeteneği, önemli bir gerek olarak ön plana çıkmaktadır. Küçük boyutlu nesnelerden özellikler çıkarmak ve yerlerini tespit etmek derin öğrenme modelleri için oldukça zordur (Tong vd., 2020). Görüntüde piksel boyutu 32x32'den küçük olan nesneler küçük nesneler olarak kabul edilmektedir (Lin vd., 2014). Küçük nesneleri termal kızılötesi görüntülerde arka plandan ayırıp tespit etmesi kolay değildir (Y. Liu, Su, vd., 2021). Özellikler savunma teknolojilerinde önem arz eden küçük hedef tespiti algoritmalarının geliştirilmesi hususunda çok farklı yöntemler sunulmaktadır (Y. Liu, Sun, vd., 2021b). Bu yöntemlerden birisi de veri çoğaltma yöntemidir (Tong vd., 2020). Veri çoğaltma, temelde nesnelerin sayısını artırmak için yapılsada, nesnelerin yerlerini, şekillerini vb. değiştirdiği ve eğitim verisinin çeşitliliğini artırdığı için derin öğrenme modelleri için oldukça önem arz etmektedir.

Bu tez çalışmasında kızılötesi nesne tespit problemi için güncel nesne takip modellerinin performansları karşılaştırılmıştır. Kızılötesi görüntüler ve nesne takip üzerine detaylı araştırmalar yapılmıştır. Fast RCNN'in omurgasını oluşturan Bölge Teklif Ağı gibi yöntemler, renkli görüntüler ve kızılötesi görüntüler arasındaki karakteristik farktan dolayı kızılötesi görüntülerden nesne tespit çalışmalarına pek uygun olmamaktadır (Expansion & Detection, 2020). Bu sebeple ve bu tez kapsamında yapılan literatür taraması sonucunda Yolov3 (Redmon & Farhadi, 2018), Yolov4 (Bochkovskiy vd., 2020), Yolov7 (C.-Y. Wang vd., 2022) ve Thermal Yolo (Huda vd., 2020a) olarak isimlendirilen nesne takip modelleri bu çalışmada karşılaştırma amacıyla seçilmiştir. Yapılan deneylerde bu modellerin kızılötesi alan adaptasyonları incelenmiştir. Çalışma kapsamında, Linux işletim sistemi, Darknet (Bochkovskiy, 2022) iskeleti ve OpenCV (OpenCV, 2000) kütüphanesi kullanılmıştır. OpenCV, bilgisayarlı görü alanında nesneleri tespit etmek için sıkça kullanılan, açık kaynak kodlu bir kütüphanedir. Bu kütüphane yüz tanıma, nesne tanıma ve takip etme vb. uygulamalarda kullanılmasının yanı sıra görüntü/video düzenleme, renk eşitleme, çözünürlük artırma gibi çok çeşitli görüntü/video işleme uygulamalarında da

kullanılmaktadır. Kütüphanenin tüm işletim sistemleri için Python, MATLAB vb. ara yüzleri mevcuttur. Bu çalışmada Linux işletim sisteminde Python programlama dili ile OpenCV 3.2.0 sürümü kullanılmıştır.

Bu tezde yapılan deneylerde kullanılmak amacıyla, insan ve araç sınıfları nesne tespiti için uygun şekilde işaretlenmiş, kamuoyuna açık CAMEL, LTIR ve AALARTDATA veri setlerinden 1000 adet kızılötesi görüntü harmanlanmıştır. Bu harmanlamanın amacı modele farklı boyutlarda ve resim içinde farklı konumlarda nesneleri tanıtmak, modelin bu nesneleri farklı kamera pozisyonlarında ve farklı ortamlardan ayırt etmesini sağlamaktır. Seçilen kızılötesi veri seti, veri artırma yöntemleri ile genişletilmiş, resimlerdeki insanlar ve araçlar bu artırma işlemleri dahilinde otomatik olarak işaretlenmiştir. Elde edilen 48000 resimlik veri seti modeller için eğitim, doğrulama ve test veri seti olarak kullanılmıştır. Veri seti, çapraz doğrulama yöntemi ile eğitim, doğrulama ve test verisi olarak ayrılmıştır. Modeller, eğitim veri seti ile eğitilmiş, doğrulama veri seti ile analiz edilmiş ve test veri seti ile test edilmiştir. Bu testler sonucunda performans değerlendirmesi birçok nesne tespit sonucu değerlendirme metriği ile yapılmıştır. Ayrıca kızılötesi alanda küçük nesne tespiti oldukça önemli bir konudur (Bi vd., 2017). Bu sebeple modellerin küçük nesneleri tespit etmedeki yetenekleri de deneyler dahilinde değerlendirilmiştir.

Bu tez kapsamında;

- Popüler nesne tespit modellerinin kızılötesi görüntü üzerinden nesne tespit performansları incelenmiştir.
- Kızılötesi görüntü özellikleri, veri seti çoğaltma yöntemleri, derin öğrenme modellerinin kullanılması ve modelin çıktısının değerlendirilmesi hakkında bilgiler verilmiştir.
- Kızılötesi görüntü adaptasyon yeteneklerine göre farklı modeller değerlendirilmiştir.
- Kızılötesi alan adaptasyonuna ve küçük nesne tespit çalışmalarına uygun bir veri seti geliştirilmiştir.

Tezin 2. bölümde, örnek nesne tespit çalışmaları ile beraber literatür taraması bulunmaktadır. 3. bölümde, genel olarak tez çalışmasının daha iyi anlaşılması adına görüntü işleme, derin öğrenme, kızılötesi görüntüleme ve evrişimli nesne tespit modelleri konuları alt başlıklarıyla birlikte detaylı incelenmiştir. Ayrıca verinin elde edilmesi, eğitim işlemleri ve test işlemleri olmak üzere kullanılan bütün yöntem ve deneylerden bahsedilmiştir. 4. bölümde tez çalışması kapsamında elde edilen bulgular sunulmuştur. 5.

bölümde tez çalışmasının sonuçları değerlendirilmiştir. Bu tez çalışması Türkçe yazıldığından dolayı terminoloji karmaşıklığını azaltabilmek adına nesne tespit çalışmalarında yaygın olarak kullanılan İngilizce kelimeler Türkçe karşılıkları ile beraber EK-1’de verilmiştir. Modellerin test verilerinde aldıkları sonuçlar EK-2’te verilmiştir.

|



2. KAYNAK ARAŞTIRMASI

Biyolojik nöronlardan esinlenilerek ortaya çıkarılan Yapay Sinir Ağları, sınıflandırma, tanıma, tahmin etme gibi birçok problemde kullanılmaktadırlar. Derin öğrenme ağları, bir giriş, bir çıkış ve gizli katman içeren klasik sinir ağlarından farklı olarak, bin kadar gizli katmana sahip olabilen sinir ağı modelidir (Albawi vd., 2018b). Derin öğrenme modelleri yüksek seviye özellikleri öğrenme ve özelliklerden çıkarım yapma konusunda başarılarından dolayı birçok uygulama alanına sahiptir. Görüntü ve videolar ile işlem yapma konusunda en yaygın olarak kullanılan derin öğrenme modelleri Konvolüsyonel (Evrişimsel) Sinir Ağları olarak geçmektedir (Albawi vd., 2018b). ESA, nesne tanıma, nesne algılama, sınıflandırma gibi birçok alanda araştırılmaktadır. Görüntülerden nesne tespiti, görüntü piksellerinden özellikler elde ederek bu piksellerin ait olabileceği nesneyi tahmin etmeye yarayan bir bilgisayarlı görü uygulamasıdır. Nesne tespiti sayesinde otonom araçlar ve gözetim sistemi teknolojileri giderek gelişmektedir.

Görüntülerin ve videoların piksellerini kullanarak ilgili nesneleri tespit etme çalışmaları görüntünün cinsi, tespit edilecek nesne ve kullanılacak yöntem bakımından oldukça çeşitlidir. Nesne tespit çalışmaları sağlık hizmetleri, savunma teknolojileri, endüstri uygulamaları gibi birçok alanda yaygın olarak kullanılmaktadır. Çoğu nesne tespit çalışmasında gelişmiş derin öğrenme modelleri kullanılsa da literatürde geleneksel görüntü işleme yöntemlerinin kullanıldığı çalışmalarda mevcuttur. (Broggi vd., 2004; Gu vd., 2010; Suard vd., 2006; X. Zhao vd., 2015) numaralı çalışmalar geleneksel görüntü işleme yöntemleri ile yapılan nesne tespit çalışmalarına örnek olarak verilebilir.

Yapay Zeka teknolojisinin gelişmesiyle beraber nesne tespit çalışmalarında kullanılan derin öğrenme modelleri giderek çoğalmakta ve gelişmektedir. Bunun yanı sıra nesne tespit çalışmaları, kullanılacak model ve tespit edilecek nesneye göre de şekillenmektedir. Örneğin, sağlık alanında vücut içyapısı, biyometrik veriler vb. kullanılırken, endüstriyel uygulamalarda çeşitli endüstriyel ürünler, araçlar, cihazlar vb. kullanılabilir. Kullanılacak nesnelerin özellikleri her uygulama için farklı olacağından ona uygun derin öğrenme modelini belirlemek gerekir. Aynı şekilde görüntünün elde edildiği alanda önemlidir. Örneğin, görünür alanda elde edilen bir görüntü ile kızılötesi alanda elde edilen bir

görüntünün özellikleri ve nesneye ait bilgileri farklı olacaktır. Bu sebeple nesne tespit çalışmaları ilk olarak tespit edilecek nesneye, bu nesneyi tespit etmek için uygun bir görüntüye ve tespit işlemini yapacak uygun modele bağlıdır.

Wu ve arkadaşları (Wu vd., 2021) numaralı çalışmada, kızılötesi görüntülerden insan, araba ve bisiklet gibi küçük nesnelerin tespiti için YOLOV3 modeli Toplu Normalleştirme (BN) katmanını eklenerek kullanılmıştır. Eğitim için kamuya açık FLIR Termal veri setini kullanılmıştır. Küçük nesnelerin tespiti için özellik ölçekleri 13x13, 26x26, 52x52 ve 104x104 olmak üzere genişletilmiştir. Kayıp fonksiyonunu azaltmak için GIoU olarak isimlendirdikleri ve IoU yönteminden farklı olarak nesnenin bulunduğu gerçek kutu ve tahmin kutusu arasındaki mesafeyi ölçmeye yarayan bir yöntemi önermişlerdir. Bu sayede gerçek kutu ve tahmin kutusunun ne kadar yakın olduğu bilgisi ile modelin iyileştirmesini yapmışlardır. Toplamda 37000 iterasyon eğitilen modelde momentum = 0,9 olarak seçilmiştir. Öğrenme oranı ise 0,001'den başlayarak 29600 ve 33300. iterasyonlarda 0,1 oranında düşürülmüştür. NVIDIA Xavier gömülü sistemi üzerinden yapılan test sonucunda mAP %67,9 olarak ölçülmüştür.

Liu ve arkadaşları (Y. Liu, Su, vd., 2021) kızılötesi görüntülerden nesneleri tespit etmek için Yolov3 modelinin özellik piramit katmanları olan 79, 91 ve 103'üncü katmanlarına bağlanan, çapa (anchor) içermeyen, özellik seçici FSAF modülü eklemişlerdir. FSAF modülünün ağırlık parametresi $\beta = 0,6$ optimum değer olarak elde edilmiştir. Bu modülün küçük nesne tespit performansını geliştirdiği ifade edilmiştir. Genel tespit sonucunu verecek katman NMS (maksimum olmayan bastırma) yerine Soft-NMS olarak seçilmiştir. Araba, insan, bisiklet gibi etiket verisi içeren veri seti, birçok farklı arka plana sahip 640x480 boyutunda 5322 adet kızılötesi resim içermektedir. Elde ettikleri model, orijinal Yolov3 modeline göre daha iyi performans sergilemektedir. mAP, %72,2 olarak elde edilmiştir.

Du ve arkadaşları (Du, Zhang, Zhang, & Xu, 2021) araç takibi için kızılötesi hava aracı resimleri ile Yolov4 modelini kullanmışlardır. Çalışmada çift aşamalı bir eğitim modeli kullanılmış ve model ilk olarak UCAS-AOD ve VIVID renkli görüntü veri setleri ile eğitilmiştir. Daha sonra VIVID kızılötesi görüntü veri seti ile eğitim gerçekleştirilmiştir.

Çalışmada kullanılan UCAS-AOD veri setinden 510 görüntü seçilmiş ve döndürme, çevirme, gürültü ekleme yöntemleri kullanılarak 3060 görüntüye çoğaltılmıştır. VIVID renkli görüntü veri seti 320 görüntüyü döndürme ve çevirme yöntemleri ile 960 görüntüye çoğaltılarak kullanılmıştır. Kızılötesi görüntü seti ise 310 resimden oluşmaktadır. Modele arka plandan kaynaklı yanlış tespitleri azaltmak adına sert olumsuz örnek madencilik bloğu ekleyerek modelin performansını attırdıklarını ifade etmişlerdir. Elde edilen yeni model için AP ve F1 değerleri sırasıyla %91,92 ve %87,98'dir.

Tumas ve arkadaşları (Tumas vd., 2021) gerçekleştirdikleri çalışmada insan tespiti için TinyV3, TinyL3, YOLOv3, YOLOv4, ResNet50 ve ResNext50 modellerini kullanmışlardır. Modeller üzerinde bazı değişiklikler yapılmıştır. Giriş görüntü boyutu 640x480 olarak değiştirilmiş, çapa kutusu boyutları k-ortalama algoritması ile yeniden hesaplanmıştır. Çalışmada kötü hava şartlarında elde edilen kızılötesi görüntü eksikliği için iyi hava koşullarında elde edilen resimleri kötü hava şartlarına benzeyecek şekilde üretilmiş ve ayrı ayrı eğitim gerçekleştirmişlerdir. Artırılmış verilerin insan tespit performansını artırdığını belirtmişlerdir. Test aşamasında en iyi sonuç Yolov4 tarafından elde edilmiştir. mAP, Yolov3 ve Yolov4 için sırasıyla %84,91 ve %87,02 olarak ölçülmüştür.

Jiang ve arkadaşları (C. Jiang vd., 2022) insansız hava araçlarından elde edilen kızılötesi video ve görüntülerden insan ve araç tespiti gerçekleştirmişlerdir. Kullanılan veri seti 10288 kızılötesi görüntü ve 4224 adet kızılötesi video içermektedir. Yolov3, Yolov4 ve Yolov5 modelleri toplamda 300 adım eğitilmiştir. Başlangıç öğrenme oranı 0,1; momentum ve ağırlık düşüşü sırasıyla 0,9 ve 0,005 olarak seçilmiştir. Eğitim aşamasında Yolov3'ün görüntülerdeki çoğu nesneyi tespit ederek kesinlik ve hassasiyet değerlerinin en yüksek değerlere sahip olduğu görülmüştür. Doğrulama veri seti için Yolov3 ve Yolov5'in en yüksek performansa sahip olduğu gözlenmiştir. Test aşamasında Yolov3, Yolov4 ve Yolov5 modelleri için sırasıyla %82,2, %62,2 ve %80,3 mAP değerleri elde edilmiştir.

Hou ve arkadaşları (Hou vd., 2022) tarafından gerçekleştirilen çalışmada özellik çıkarıcı katmanı yani omurgası MobileNetV2den oluşan bir M-YOLO modeli önerilmektedir. Güçlü yönleri ön planda tutularak modelin temeli Yolov4 üzerine kurulmuştur. Modelde Piramit Havuzlama Modülü (PPM) gibi yüksek çözünürlüklü özellik haritalama modülleri yerine, Küresel Bağlam Bilgisi Toplama Modeli (GCIAM) isimli bir modül

geliştirmişlerdir. Bu bağlam modülü kızılötesi görüntülerdeki nesneleri arka plandan ayırmak için kullanılmıştır. Özellikleri birleştirmek ve anlamlandırmak için Özellik Füzyon Modeli (FFM) kullanılmıştır. Aktivasyon fonksiyonu olarak ReLU6 tercih edilmiştir. FLIR ve GIR veri setleri ile yapılan deneylerde M-YOLO'nun Yolov4'e göre insan, araç, bisiklet gibi nesneleri tespit etme performansı veri setleri için sırasıyla %5,3 ve %6,4 geliştirdiği gözlenmiştir. Bunun yanı sıra tespit hızının arttığı ve yanlış algılama oranının azaldığı belirtilmiştir. Elde edilen yeni model için AP %83,4 olarak ölçülmüştür.

Du ve arkadaşları (Du, Zhang, Zhang, Xiang, vd., 2021) yaptıkları çalışmada kızılötesi görüntülerde nesneyi arka plandan ayırma zorluğuna karşılık odak ve dikkat (Focus-Attention) ismi verdikleri bir FA-YOLO modelini üretmişlerdir. Küçük nesne tespiti için modele genişletilmiş CBAM modülü eklenmiştir. Aynı zamanda Yolov4 modelinden türetilen bu modele YP olarak değerlendirilen nesneleri hesaba katmaya yarayan bir negatif örnek odaklama mekanizması eklenmiştir. Transfer Öğrenme sürecinde model ilk olarak UCAS AOD veri seti ile eğitilir. Daha sonra CycleGAN kullanarak VIVID görünür görüntü veri setinden araçları içeren 500 adet kızılötesi görüntü üretmişlerdir. Üretilen yeni modelin %92,95 AP ve %88,13 F1 değerlerine sahip olduğu gözlenmiştir.

Kowalski ve arkadaşları (Kowalski vd., 2021) tarafından gerçekleştirilen çalışmada insan ve tekne tespiti için Fast RCNN, Yolov2 ve Yolov3 modellerini karşılaştırılmıştır. Modelleri GoogLeNet, ResNet101 gibi farklı özellik çıkarma ağları ile denemişlerdir. Eğitim için nehirde çekilen kızılötesi videolar kullanılmıştır. Veri seti 75/25 oranında eğitim ve test veri seti olarak ayrılmıştır. Eğitim işlemi 0,95 doğruluk derecesine erişinceye dek her model için gerçekleştirilmiştir. GoogLeNet + Yolov2, %47 ortalama KBO'ya sahip olurken, Darknet53 + Yolov3 için bu değer %69 olarak elde edilmiştir.

Kristo ve arkadaşları (Kristo vd., 2020) termal-kızılötesi hedef tespitinde Faster RCNN, SSD ve Yolov3 gibi modelleri karşılaştırmıştır. Yolov3'ün, aralarında en hızlı performansa sahip detektör olduğu görülmüştür. Kullandıkları UNIRI TID veri seti farklı hava şartlarında elde edilen insan ve hayvanlara ait LWIR ve MWIR videolardan oluşmaktadır. Modeller orijinal yapıları değiştirilmeden 40000 iterasyon eğitilmiştir. En iyi AP sonuçları Faster RCNN, Cascade RCNN ve Yolov3 için sırasıyla %98,96, %98,80 ve %97,93 olarak ölçülmüştür.

Shao ve arkadaşları (Shao vd., 2022) yaptıkları çalışmada havadan kızılötesi nesne algılama problemi için ağ budama ve YOLOv3'ü beraber kullandıkları AIR-YOLO modelini önermişlerdir. Modelin hızını artırmak ve modeli küçültmek için hem kanal budama hem de katman budama yöntemi kullanılmıştır. Çalışmada ilk olarak drone görüntüleri toplanmış ve etiketlenmiştir. Bu görüntüler ile eğitilen model %91,5 mAP değerine ulaşmıştır.

Nesne tespit çalışmalarında görünür yelpaze ve kızılötesi yelpaze ayrı ayrı kullanılabileceği gibi, bu iki yelpazenin beraber kullanıldığı çalışmalarda literatürde yer almaktadır. (Bhowmick vd., 2022) numaralı çalışmada görünür yelpaze ve yakın kızılötesi yelpazede elde edilen görüntülerin barındırdıkları farklı bilgiler birleştirilerek multispektral bir nesne tespit çalışması yapılmıştır. YOLOv3 ile geliştirilen nesne tespit modeli ilk olarak renkli görüntü ve kızılötesi görüntüler ile ayrı ayrı eğitilmiş ve ardından iki yelpaze bir araya getirilerek 4 kanallı bir yapı (renkli görüntü için 3 kanal ve kızılötesi görüntü için 1 kanal) halinde modele verilmiştir. Model parametreleri, iyileştirici olarak ADAM tercih edilmiştir ve momentum = 0,9; öğrenme oranı = 0,005'den 0,001'e azalacak şekilde seçilmiştir. Eğitim toplam 200 adım gerçekleştirilmiştir. Multispektral görüntülerin kullanılması sadece IR görüntülere göre performansı %5 artırdığı gözlenmiştir.

Roszyk ve arkadaşları (Roszyk vd., 2022) tarafından yapılan çalışma son yıllarda yapılan multispektral nesne tespiti çalışmalarına bir diğer örnek olarak verilebilir. Otonom araçların gerçek zamanlı çalışabilmesi için Yolov4-Tiny modelini ve KAIST veri setini kullanmışlardır. Nesne tespit performansını değerlendirmek adına erken, orta ve geç füzyon olmak üzere 3 farklı füzyon yöntemi kullanmışlardır. Erken füzyon, modele RGB ve termal görüntüyü 4 kanalda birleştirilip verildiği yöntemdir. Önceden eğitilmiş ağırlıklar kullanılarak 25 adımda eğitilmiştir. Sonuçta %61,8 mAP değeri elde edilmiştir. Aynı şekilde geç füzyon yöntemi uygulanmıştır. Bu yöntem RGB ve termal görüntülerin ayrı ayrı işlenip son aşamada tahmin eşleştirme sistemi (PMS) ismi verilen bir yapı ile birleştirilmesidir. Geç füzyon ile elde edilen model, %64,5 mAP değerine sahip bulunmuştur. Orta füzyon ise, 3 kanal RGB ve 3 kanal termal görüntünün ayrı ayrı omurgalarda işlenip birleştirilmesi yöntemidir. Bu yöntem sonucunda %68,6 mAP değeri elde edilmiştir.

Yolov7, 2022 yılının son aylarında sunulmuş yeni bir model olmasından dolayı literatürde pek fazla örnek çalışma yer almamaktadır. Liu ve arkadaşları (S. Liu vd., 2022) tarafından gerçekleştirilen ve trafik kazalarını önlemek adına sürücülerin dikkatini dağıtan davranışların tespit edildiği çalışma örnek olarak verilebilir. Çalışmada tespit yeteneğini artırmak için kullanılan küresel dikkat mekanizması (global attention mechanism- GAM) modülü Yolov7 modelinin omurga ve kafa kısımlarına eklenmiştir. Önerdikleri modele CEAM-YOLOV7 ismini vermişlerdir. Kızılötesi görüntülerden özellik çıkarma yeteneğini artırmak için kanal genişletme algoritması kullanmışlardır. HNUST ve HNU isimli iki veri seti ile 300 adımda eğitilen yeni modelin, orijinal Yolov7'ye göre %20,26 daha yüksek mAP değerine sahip olduğu söylenmektedir.

Literatürde multispektral nesne tespit çalışmaları oldukça yaygın kullanılsa da (Yun vd., 2022), multispektral (renkli + termal) görüntüleri günlük uygulamalarda kullanmak hem maliyet hem de pratiklik açısından pek uygun değildir (Guo vd., 2019). Bu tarz çalışmalar için özel üretim çok kanallı sensörler gerekmektedir ki bu sensörler oldukça pahalıdır. Ayrıca düşük ışık yoğunluğunda nesnelerin termal görüntülerde renkli görüntülere göre daha iyi ayırt edildiği literatürde belirtilmiştir (González vd., 2016; Guo vd., 2019). Bu sebeple bu çalışmada multispektral görüntüler yerine sadece kızılötesi alanda elde edilen görüntüler tercih edilmiştir.

3. MATERYAL VE YÖNTEM

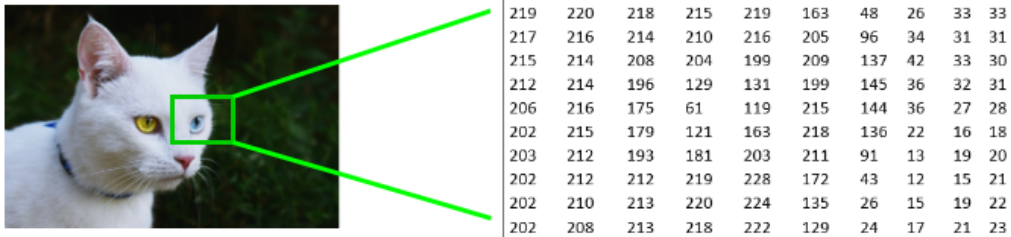
Bu bölümde tez çalışması kapsamı içinde kalan alanlara dair teorik arka plan üzerine genel bilgiler verilmiştir. Daha sonra çalışma kapsamında gerçekleştirilen yöntemler ve deneylerden bahsedilmiştir.

3.1. Teorik Arka Plan

Bu bölümde sırasıyla görüntü işleme, kızılötesi görüntüleme, derin öğrenme ve evrişimli nesne algılama modelleri konularından teker teker bahsedilmiştir.

3.1.1. Görüntü işleme

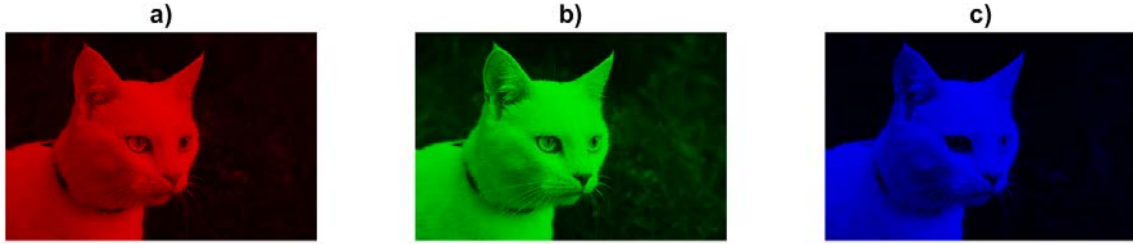
Dijital görüntü, temelde belli bir koordinattaki parlaklık verisini ifade eder ve $f(x,y)$ şeklinde gösterilir. Burada x , y uzamsal koordinatları temsil ederken, f fonksiyonu o koordinattaki parlaklık değerini verir (Hlavac, 2011; Petrou & Petrou, 2010). Bu sebeple bir görüntü, her bir elemanı piksel olarak isimlendirilen ve parlaklık verilerini içeren 2 boyutlu bir matristen oluşur. Renkli bir görüntü için parlaklık verisi 8 bit olup, Şekil 3.1’de gösterildiği gibi $[0, 255]$ arasında değerler alır.



Şekil 3.1. Örnek piksel değerleri

Renkli görüntüler kırmızı, mavi ve yeşil olmak üzere 3 ana renkten yani 3 kanaldan oluşur ve bu parlaklık değerleri 3 kanalın birleşmesi ile elde edilir (Petrou & Petrou, 2010). Şekil 3.2’de bir görüntünün kanalları ayrı ayrı verilmiştir. Bu 3 kanalın birleşmesi ile orijinal resim elde edilir. Burada her kanalın pikselleri kendi içerisinde $[0, 255]$ arasında değerler alır ve bu değer o pikselin parlaklık değerini ifade eder. 3 kanaldaki piksellerin üst üste

gelmesi ile diğer renkler oluşur. Bu resimler RGB (red-green-blue) olarak isimlendirilen 3 boyutlu bir koordinat sistemiyle ifade edilir.



Şekil 3.2. Bir görüntünün kanalları a) kırmızı kanal, b) yeşil kanal ve c) mavi kanal

Bilgisayarlar görüntüleri matris formatında algırlar. Bu matrisler aslında görüntü hakkında bilgileri barındıran elemanlardır. Görüntü işleme yöntemleri bu elemanları analiz etmeye, değiştirmeye, düzenlemeye vb. yarar. Günümüzde sayısız alanda görüntü işleme yöntemleri kullanılmaktadır. Görüntü işleme yöntemlerine;

- Görüntü dönüşümleri,
- Görüntü iyileştirme,
- Görüntü onarma,
- Görüntü bölümleme,
- Görüntü sıkıştırma,
- Görüntü algılama

gibi örnekler verilebilir (Ince vd., 2020).

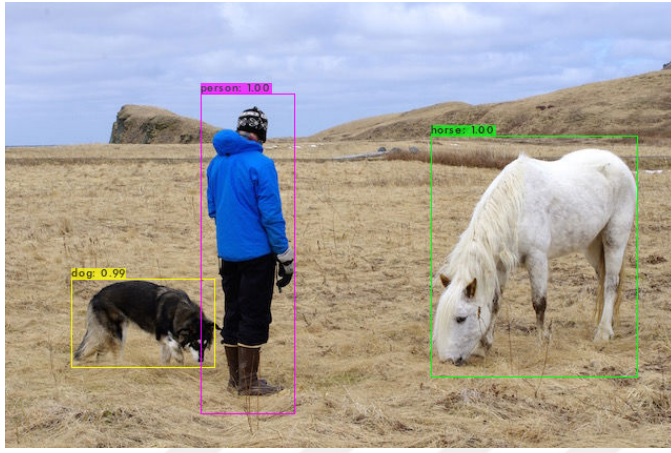
Bilgisayarlı görü ve nesne tespiti

Bilgisayarlı görü, bir görüntü işleme dalıdır. Görüntünün piksellerinden özellikler çıkarma işlemidir. Bilgisayarlı görünün en çok kullanılan 2 alt başlığı sınıflandırma ve nesne tespittir. Görüntü sınıflandırma, görüntüde nesneye ait özelliklerin aranması ve sonuç olarak görüntüde bu nesnenin olup olmadığının çıkarılması işlemidir.

Nesne tespit ise, görüntülerden nesnelerin piksel yer bilgilerinin elde edilmesi yöntemidir (Bhowmick vd., 2022). Bir başka deyişle, nesnelerin olduğu piksellerin işaretlenmesi ve bu piksellerin hangi nesneye ait olduğunun bulunması işlemidir (Ju vd., 2019). Derin öğrenme tabanlı bir nesne tespit çalışması genel olarak aşağıdaki işlem sıralamalarını içerir.

- İlk olarak nesne tespit uygulamasında kullanılacak bir derin öğrenme modeli seçilir.
- Bu model, bilgisayarlı görü sayesinde tespit edilmek istenilen nesnelerin özelliklerini öğrenir. Bu özellikler öznitelik olarak isimlendirilir.
- Burada özellik öğrenme, örnek görüntü verisi üzerinde daha önceden nesnenin bilgilerini barındıran piksellerin işaretlenmesi ve bu piksellerdeki özelliklerin model tarafından öğrenilmesi ile yapılır.
- Modelin eğitilmesinin ardından model, yeni görüntüler üzerinde ilgili nesnelere ait özellikler tarar ve nesneleri tespit etmeye çalışır.

Resim 3.1’de bir nesne tespit örneği verilmiştir.



Resim 3.1. Nesne tespit örneği¹

Nesne tespit algoritmaları günümüzde;

- Trafik işareti tanıma (Gavrila, 1999),
- Otonom araçlar (Tumas vd., 2021),
- Medikal çalışmalar (Lyra vd., 2021),
- Akıllı gözetim (Nguyen-Meidine vd., 2018),
- Terör tehditleri ve yasa dışı göç tespiti (Kristo vd., 2020),
- Trafo merkezlerinde hata algılama (Gao vd., 2022),
- Tarım uygulamaları (K. Jiang vd., 2022),
- Alev algılama (S. Li vd., 2022),

vb. gibi birçok alanda kullanılmaktadır. Bu uygulamaların bazıları kızılötesi görüntüler kullanılarak gerçekleştirilen örnekler arasından seçilmiştir. Çünkü bu tez çalışmasında odak nokta kızılötesi görüntülemedir.

¹ Yolov3 modeli kullanılarak Colab platformu üzerinden elde edilmiştir.

Son yıllarda artan bilgisayarlı görü teknolojisi ile birlikte küçük hedef tespiti çalışmaları yaygınlaşmıştır (Y. Liu, Sun, vd., 2021a). Bunların bir kısmı, gözetleme, uzaktan kontrol gibi uygulamalar kızılötesi görüntüler üzerinden gerçekleştirilmektedir. Küçük nesneler, görüntülerde az sayıda piksellerden oluşur. Kızılötesi görüntülerden küçük hedef tespitinin en zor kısmı, nesneleri arka plandan ayırmaktır. Kızılötesi görüntülerde bulutlar, ağaçlar, deniz vb. gibi arka plan öğeleri bazen nesnelerle karıştırılabilir (Bi vd., 2017; Gu vd., 2010). Aynı zamanda soğuk nesneler görüntüde kaybolabilir (D'Acremont vd., 2019). Bu gibi sebepler tespit performansını düşürmektedir. Ayrıca küçük nesnelerin sahip olduğu özellikler tespit modellerinin evrişim katmanları boyunca giderek kaybedilir (Xiao, 2021). Özellikle maksimum havuzlama gibi katmanlarda bu özellikler tamamen silinebilmektedir. Bu sebeple günümüz nesne tespit detektörleri büyük ve orta boyutlu nesneler için üretilmiştir (Melik Ahmet Daye, 2021).

Küçük nesnelerin literatürde kesin bir tanımı ve açıklaması bulunmamaktadır. Torralba ve diğerleri (Torralba vd., 2008), 32x32 ve daha küçük piksel boyutlarına sahip nesneleri küçük nesne olarak tanımlarken, Zhu ve diğerleri (Zhu vd., 2016), görüntü boyutunun %20'sinden küçük olan nesneleri küçük nesne olarak ifade etmişlerdir. COCO veri setine göre ise boyutu 32x32'den küçük olanlar küçük, 32x32 ve 96x96 arasında olan orta ve 96x96'dan büyük olanlar büyük nesne olarak kabul edilir (Lin vd., 2014).

3.1.2. Kızılötesi ışıma

Bilinen fizik yasalarına göre, her nesne Elektromanyetik Yalpaze olarak adlandırılan geniş bir dalga boyu bölgesinde benzersiz bir ışıma yayar (Karim & Andersson, 2013). Çizelge 3.1'de listelendiği gibi Elektromanyetik Yalpaze, farklı dalga boylarında farklı ışımalara sahiptir. Bu geniş dalga boyu bölgesinin 0,7 μm ila 1 mm arasındaki dalga boyuna kızılötesi ışıma denir.

Çizelge 3.1. Elektromanyetik yelpaze

İşıma Türü	Dalga boyu aralığı
Gama Işınları	$<10^{-12}$ m
Röntgen	1 nm- 1 μ m
Ultraviyole	400 nm- 1 nm
Görünür	750 nm- 400 nm
Kızılötesi	1 mm- 0,7 μ m
Mikrodalgalar	1 mm- 25 μ m
Radyo Dalgaları	>1 mm

Radyometre, ışımayı ölçmek için kullanılan bir cihazdır. Işıma ölçümü, analizi ve hesaplanması üzerine eğilen bilim dalına radyometri denilmektedir (Datla & Parr, 2005). Optik radyometri için bazı önemli terminolojiler vardır. Bunlar kısaca maddeler halinde verilmiştir.

- Parlaklık (L), tüm frekanslarda katı bir açı içindeki bir alandan gelen optik güçtür. Parlaklık, $W/(m^2 \text{ sr})$ birimlerine sahiptir; burada W watt, m metre ve sr steradyandır.
- Işınım (E), uzayın bir bölgesindeki bir alan tarafından tüm frekanslarda alınan ışıma akısıdır. Işınım W/m^2 birimlerine sahiptir.
- Radyan Çıkış (M), tüm frekanslarda yüzeyden birim alan başına yayılan radyan güçtür. W/m^2 birimlerine sahiptir.

Elektromanyetik (EM) ışıma, Eş. 3.1’de verildiği gibi dalga boyun (λ) ve optik frekansa (ν) sahip salınımlı bir sinüzoidal dalga biçimi olarak tanımlanır.

$$U = c/\lambda \quad (3.1)$$

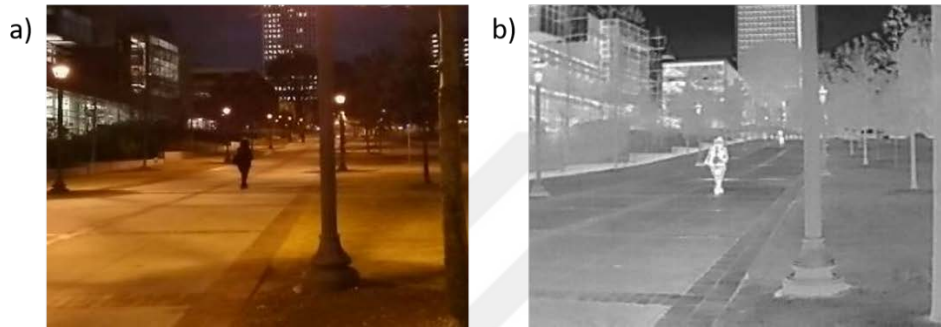
Termal ışıma teorisine göre, gelen tüm ışımaları emen ve sürekli bir spektrum yayan cisimler kara cisim olarak isimlendirilir. Kara cisim tarafından yayılan güç Planck yasasına göre Eş. 3.2’deki gibi ifade edilir.

$$q = \epsilon T^4 \sigma A \quad (3.2)$$

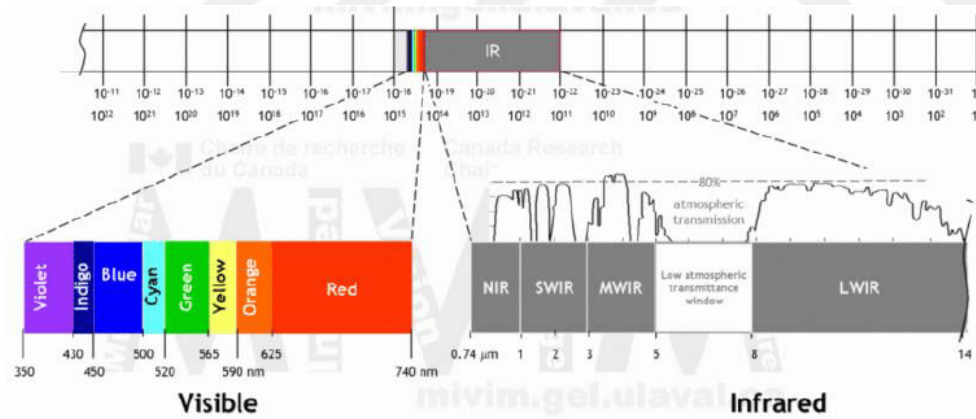
Burada A, yüzey alanıdır (m^2), T, sıcaklıktır (K), σ , Stefan–Boltzmann sabitidir ($\sigma = 5,676 \times 10^{-8} W m^{-2} K^{-4}$) ve ϵ , salım gücüdür (Bagavathiappan vd., 2013).

Kızılötesi (IR) ışımanın kullanıma başlanması, modern fizik, kuantum teorisi, Plank yasası ve kara cisim ışımasının anlaşılmasıyla 20. yüzyılın sonlarına denk gelmektedir. Ancak

bazı kaynaklar IR ışımasının ilk keşfini 200 yıl önce William Herschel tarafından yapılan bir deneye dayandırmaktadır (Delimata, 2019; Karim & Andersson, 2013). Kızılötesi ışıma, insan gözüyle görülemeyen, uzun bir dalga boyuna sahip olan ve her türlü nesne tarafından yayılabilen küçük bir enerji ışımasıdır (Hamamatsu Photonics K.K., 2011). Bu ışıma, bir nesnenin yüzey sıcaklık dağılımının elde edilmesini sağlar. Bu sayede IR bandı sivil, havacılık, tıp ve askeri alanlarda yaygın olarak kullanılmaktadır (Lozano & Hassanipour, 2019; Piñeiro-Ave vd., 2014; Wen vd., 2018; Xu vd., 2019). Resim 3.2’de aynı alanın görünür yelpaze ve kızılötesi yelpazedeki görüntüleri verilmiştir.



Resim 3.2. Örnek görüntü a) görünür yelpaze, b) kızılötesi yelpaze ²



Şekil 3.3. Kızılötesi dalga boyları (Akhroufi & Bendada, 2013)

Kızılötesi ışıma aralığı geleneksel olarak beş spektral bantta sınıflandırılır. 0,7 μm ile 1,4 μm arasındaki dalga boyu bölgesi yakın kızılötesi (NIR), 1,4 μm ile 3 μm arası kısa dalga kızılötesi (SWIR), 3 μm ile 8 μm arası orta dalga kızılötesi (MWIR), 8 μm ile 15 μm uzun dalga kızılötesi (LWIR) ve son olarak 15 μm ile 1000 μm uzak kızılötesi (FIR) olarak isimlendirilir. Şekil 3.3’te kızılötesi dalga boyları verilmiştir. Bu bölgelerde mutlak sıcaklığı 0 K’nin üzerinde olan her nesne, nesnenin sıcaklığı ve yüzey durumu ile

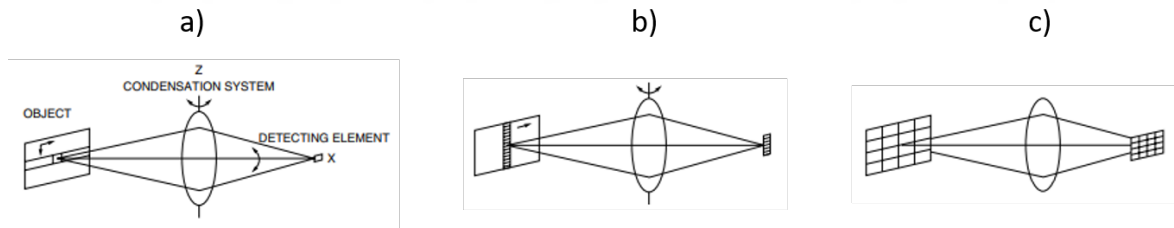
² Görüntüler CAMEL veri setinden alınmıştır.

belirlenen bir kızılötesi ışıma yayar (Boreman, 1998; Dobrzański vd., 1997; Hamamatsu Photonics K.K., 2011).

Kızılötesi detektörler

Kızılötesi detektörler, nesnelerden yayılan enerjiyi veya ısıyı algılamaya yararlar ve termal kameraların en önemli parçasıdır. Çalışma prensibi farklı sıcaklıklardaki nesnelerin farklı kızılötesi enerji yaymasına dayanmaktadır. Algılanan bu enerji, elektrik sinyalleri olarak depolanır ve sonuçta kızılötesi görüntülerin oluşturulmasında kullanılır.

Yüksek performanslı IR detektörlerin gelişmesiyle birlikte kızılötesi kamera teknolojisi, uzaktan algılama uygulamalarında yaygın olarak kullanılmaya başlanmıştır. Genel olarak optik ve elektronik sistem, detektör elemanları ve soğutma sisteminden oluşan birimin tümü kızılötesi kameralar olarak adlandırılabilirler (Bagavathiappan vd., 2013). Kamera teknolojisi üç başlıkta incelenebilir. Şekil 3.4'te kızılötesi kamera türleri verilmiştir.



Şekil 3.4. Kızılötesi kamera türleri, a) birinci nesil kameralar, b) ikinci nesil kameralar, c) üçüncü nesil kameralar (Hamamatsu Photonics K.K., 2011)

Birinci nesil kameralar, Şekil 3.4a'da gösterildiği gibi, tek elemanlı bir detektörden oluşur ve bir görüntü oluşturmak için hem yatay hem de dikey tarama aynalarına sahiptir. İkinci nesil kameralar Şekil 3.4b'de verilmiştir, yalnızca z eksenini dönen bir optik sisteme sahip detektörlerden oluşur. Modern üçüncü nesil kameralar, FPA (odak düzlemi dizileri) olarak bilinen Şekil 3.4c'de gösterildiği üzere iki boyutlu detektörlere sahiptir (Bagavathiappan vd., 2013; Hamamatsu Photonics K.K., 2011). IR detektörler temelde iki gruba ayrılır: foton (kuantum) detektörler ve termal detektörler. Çizelge 3.2'de, çeşitli IR detektörlerinin özellikleri verilmiştir.

Çizelge 3.2. IR detektörlerin özellikleri (Choi vd., 2005; Hamamatsu Photonics K.K., 2011; Rogalski, 2002)

Tip		Özellikleri
Termal Detektörler	Termokupl, Termopil, Bolometre, Pnömatik hücre, Piroelektrik dedektör	Oda sıcaklığında çalışırlar. Düşük maliyetlidirler. Ancak dezavantajları olarak, düşük algılama, yüksek frekans ve yavaş tepkiye sahiptirler.
Foton Detektörler	Foto iletkenler	Tek bir yarı iletken malzemeden oluşur. Uygulanan voltaj veya akım ile hassasiyet artar. Genellikle fotoiletkenler termal görüntüleme için çok uygun bir seçim değildir (Vollmer & Möllmann, 2010).
	Foto diyotlar	Bir yarı iletkendeki p-n bağlantısından kaynaklanır. Fotona duyarlı bir diyottur. Bir akım veya voltaj çıkışı üretir.
	Schottky Bariyer	Schottky bariyer detektörleri düşük kuantum verimliliğine sahiptir. IR foton absorpsiyonu zayıftır.
	Kuantum Kuyusu IR Foto detektörler (QWIP)	Geniş bir algılama dalga boyu aralığına sahiptir, düşük maliyetlidir. Bir dezavantaj olarak, QWIP'ler düşük sıcaklıklarda çalışır.

Kızılötesi detektörlerin performansını gösteren üç ana özellik vardır; foto duyarlılığı, NEP ve D*'dir. Bunlar Çizelge 3.3'te verilmiştir.

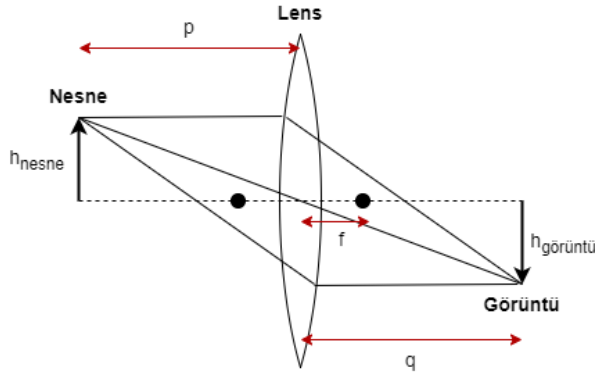
Çizelge 3.3. Detektörlerin performanslarını ifade eden parametreler (Boreman, 1998; Vollmer & Möllmann, 2010)

Açıklama	Simge	Formül	Birimi
Foto Duyarlılığı (Photo Sensitivity- Responsivity)	R	$R = \frac{S}{PA}$	V/W
Gürültü Eşdeğer Güç (Noise Equivalent Power)	NEP	$NEP = \frac{PA}{\frac{S}{N} \cdot \sqrt{\Delta}}$	W/Hz ^{1/2}
Algılama (Normalleştirilmiş Algılama- Detectivity)	D*	$D^* = \frac{\sqrt{A}}{NEP}$	cm.Hz ^{1/2} /W

Burada S sinyal çıkışı, P gelen enerji, A detektörlerin aktif alanı, N gürültü çıkışı ve Δ gürültü bant genişliğidir.

Kızılötesi görüntüleme

Temelde bir cismin sıcaklık haritasının elde edilmesi prensibine dayanan IR görüntüleme teknolojisi, 1920'lerin sonlarında foto emisyon ile başlamış olup günümüzde gelişmeye devam etmektedir (Delimata, 2019). Termal-Kızılötesi görüntülemeyi anlamak için temel optik bilgilerinden yararlanmak gerekir. Bir ışın modelini ele alırsak, bu modelde kırınım olmadığı ve ışınların bir kaynak noktasından yani nesneden düz çizgiler halinde yayıldığı varsayılır. Dolayısıyla her nesnenin yaydığı ışının enerjisi birbirinden farklı olacaktır (Boreman, 1998). Şekil 3.5'te görüntü oluşumu için örnek bir optik sistem verilmiştir.



Şekil 3.5. Görüntü oluşumu (Boreman, 1998)

Lensler, nesnenin görüntü düzleminde benzer enerji dağılımı ile yeniden üretilmesini sağlar. Şekil 3.5'te gösterildiği gibi lensin iki yanında, lense f odak uzaklığında odak noktaları mevcuttur. Bu odak noktaları merceğe gelen ışınların mercekten çıkarken birleştiği noktayı ifade eder (Boreman, 1998). Böyle bir görüntüleme sisteminde görüntü ile nesne arasındaki ilişki Eş. 3.3 ve Eş. 3.4'teki gibidir.

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{f} \quad (3.3)$$

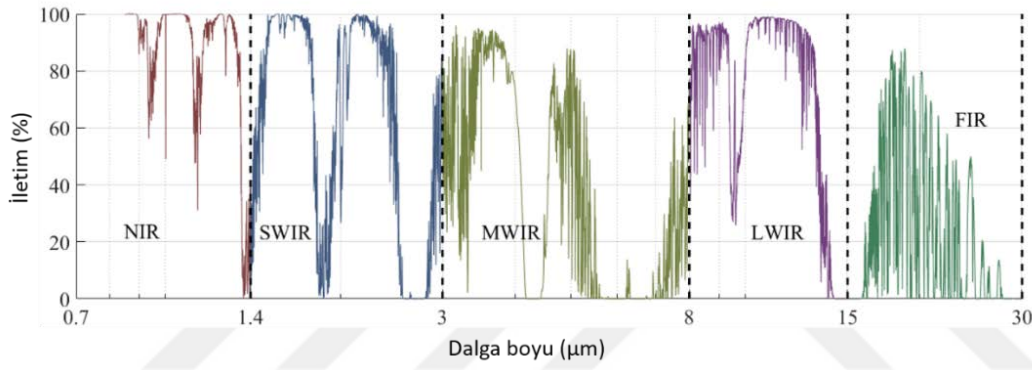
$$\frac{h_{\text{görüntü}}}{h_{\text{nesne}}} = -\frac{q}{p} \quad (3.4)$$

Burada p ve q sırasıyla merceğin cisme ve görüntüye olan uzaklıklarıdır. f, merceğin odak uzaklığıdır. Görüntü yüksekliği ($h_{\text{görüntü}}$), nesne yüksekliği (h_{nesne}) ile doğru orantılıdır ve

oranlarının negatif olmasının nedeni, merceğin oluşturduğu ters görüntüdür (Boreman, 1998).

Kızılötesi görüntüleme için dalga boyları

IR yelpazesi 0,75 μm ile 1000 μm arasındaki dalga boylarından oluşsada, atmosfer şartlarında 25 μm 'nin üzerindeki dalga boylarında tüm IR ışımları IR kamera sistemine ulaşmadan atmosfer tarafından emilmektedir. Dolayısıyla 25 μm ve 1000 μm arasındaki dalga boylarında iletim gerçekleştirilmemektedir. Şekil 3.6'da dalga boylarına göre IR ışımların iletim yüzdeleri verilmiştir.



Şekil 3.6. Dalga boylarına göre iletim değerleri (Danaci & Akagunduz, 2022)³

Şekil 3.6'da görülebileceği gibi emilim hızının daha düşük olduğu yerler “atmosferik pencereler” olarak adlandırılır. Bu pencereler 2-5 μm ve 8-12 μm bant genişlikleri arasında yer alır ve uzaktan algılama uygulamalarında yaygın olarak kullanılmaktadır.

IR dalga boyları farklı özelliklere sahiptir, bu nedenle belirli uygulamalar için doğru dalga boyu bandı tercih edilmelidir. Optimum dalga boyu bandını seçmek için çalışma mesafesi, ortam sıcaklığı (iç veya dış), nesnelerin emisyonu ve IR ışımlarının geçtiği ortam dikkate alınmalıdır. Örneğin, uzun dalga boyu bandı (LWB) çoğunlukla ortam sıcaklığının altındaki nesneleri görüntülemek için kullanılır. Fakat kısa dalga boyu bandında emisyon sıcaklıkla doğru orantılıdır. Dolayısıyla sıcaklık artarsa, emisyon (yayılım) da artacaktır. Dalga boylarının özellikleri göz önünde bulundurularak tek tek açıklamak gerekirse; NIR bandında, iletim yüksek olup gece görüş uygulamalarında sıklıkla kullanılmaktadır. LWIR

³ Veriler ATRAN modülünden alınmıştır.

diğer adıyla termal bant, nesnelerin sıcaklıklarını ölçerek, çeşitli durum ve hava şartlarından dış mekanda iyi görüntüleme kalitesine sahiptirler. Otonom araç uygulamalarında far ve ışıklardan az etkilenmeleri sayesinde sıklıkla kullanılırlar (Guo vd., 2019). Otonom araç uygulamalarında NIR ve FIR bantları da kullanılmaktadır (Tumas vd., 2021; Venkataraman & Raj, 2003). LWIR bandı, insan tespit için çok önemli bir banttır. Bunun nedeni, insan vücudunun 9.3 μm dalga boyunda ışın yaymasıdır. SWIR görüntüleme, yüksek çözünürlüğe sahip olma özelliği sayesinde MWIR ve LWIR görüntüleyicilerden daha iyidir. SWIR görüntüleyiciler, düşük hassasiyet performansları nedeniyle termal görüntüleme için iyi bir seçim olmasa da, askeri hedefleme, uzun menzilli görüntüleme ve tıbbi görüntüleme gibi yüksek çözünürlüklü termal görüntüleme uygulamalarında kullanılırlar (Wen vd., 2018). Askeri uygulamalarda çoğunlukla MWIR kullanılmaktadır (Tumas vd., 2021). LWIR, sisli ve soğuk iklimlerde daha iyi çalışırken, MWIR sıcak iklimlere daha uygundur (Kanti Bhowmik vd., 2011; Kristo vd., 2020). Resim 3.3'te görüldüğü üzere, SWIR, ışık yansımalarını kullanırken, MWIR ve LWIR nesne sıcaklıklarını kullanır (Jangblad, 2018).

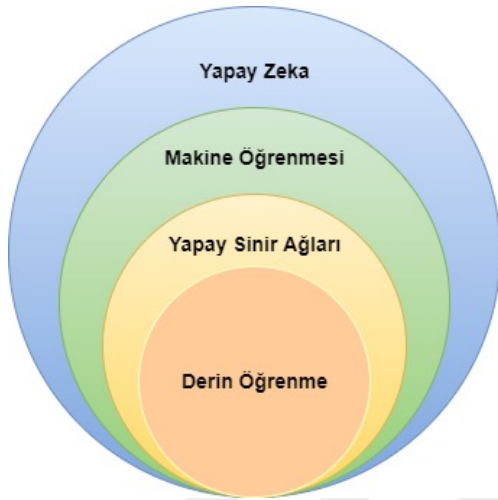


Resim 3.3. Örnek yüz görüntüsü (a) görünür yelpaze, (b) SWIR, (c) MWIR ve (d) LWIR (Wolff vd., 2005)

Kısaca, bir IR kameranın dalga boyu kısaldıkça ısıma algılaması daha hassas hale gelecektir (Vollmer & Möllmann, 2010). Kızılötesi dalga boylarının farklı özelliklere ve çalışma koşullarına sahip olması, uygulamalar için dalga boyu seçiminde zorluklar çıkarmaktadır. Bunun çözümü olarak çoğunlukla hem LWIR hem de MWIR bandında çalışabilen FLIR (Forward Looking Infrared- İleriye Bakış Kızılötesi) sistemler tercih edilir (Kristo vd., 2020).

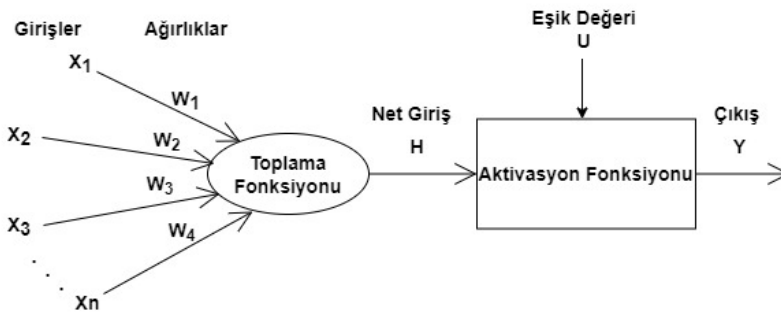
3.1.3. Yapay zeka

İnsanların çalışma sistemine benzer çalışan bilgisayarlı sistemler oluşturmayla başlayan yapay zeka günümüzde birçok alanda bulunmakta ve çalışma alanları belli kategorilere ayrılmaktadır. Şekil 3.7'deki diyagramda görüleceği gibi makine öğrenmesi, yapay zekanın bir alt dalıdır ve verilerden faydalanarak bir bilgisayarın belli bir görevi yapması olarak ifade edilir.



Şekil 3.7. Yapay zeka ve alt dalları

Makine öğrenmesinin gelişmesi ile beraber YSA (Yapay Sinir Ağı) kavramı ortaya çıkmıştır. Algılayıcılar, YSA'ların temel parçalarıdır. İnsan sinir ağından etkilenilerek modellenen algılayıcıların yapısı Şekil 3.8'de verilmiştir.



Şekil 3.8. Bir algılayıcı yapısı

Yapay sinir ağları, giriş katmanını, insan nöronundaki sinapslardan esinlenilen ağırlık verisi ve girdi verisini bulunduğu kısım; gizli katman, girdi verisini çıkışa ileten kısım; çıkış katmanını, girdi verisine göre çıkış üreten kısım olmak üzere 3 parçadan oluşur. Gizli katman, çıkış verisinin durumuna göre giriş katmanının ağırlıklarını ayarlar ve buna

öğrenme adı verilir (O'Shea & Nash, 2015). Öğrenme, yapay sinir ağının belli bir işlevi yerine getirebilmesi için eğitilmesi sonucu bağlantı ağırlıklarının ayarlanması sürecidir. Eğitim ise, eğitim için ayrılmış bir miktar verinin yapay sinir ağının girdisi olarak verilmesi ile başlar. Daha sonra YSA, bu eğitim verisi ile bağlantı ağırlıklarını güncellemiş ve öğrenme yoluyla eğitilmiş olur (Jain vd., 1996).

Farklı eğitim türleri vardır; denetimli, denetimsiz, yarı denetimli, pekiştirmeli ve kendi kendine denetimli öğrenme.

- Denetimli öğrenmede, modele eğitim verisi ile birlikte, eğitim verisine ait çıktı verisi de verilir. Bu çıktı verisi genelde etiket olarak ifade edilir. Model, girdi-çıktı arasındaki ilişkiyi bularak, etiket verisine en yakın sonucu bulacak şekilde ağırlıkları güncellemektedir (Jain vd., 1996). Denetimli öğrenme, sınıflandırma ve regresyon problemlerini kapsamaktadır.
- Denetimsiz öğrenmede bir etiket verisi mevcut değildir. Model çıktısının yorumlanması ile öğrenme gerçekleştirilir. Burada sonuca ulaşmak yerine, modelin etiketlenmemiş veriden çıkarımlar yapması amaçlanır. Bu öğrenme çeşidi kümeleme ve ilişkilendirme çalışmalarında kullanılır.
- Yarı denetimli öğrenme, hem etiketli hem de etiketsiz verilerin bir arada kullanıldığı yöntemdir.
- Pekiştirmeli öğrenmede, diğer türlerin aksine bir veriye ihtiyaç duyulmaz. Bu yöntemde ödül ve eylem sistemi mevcuttur. Yöntem deneme yanılma üzerine kurulmuştur. Burada eğitim işlemi çevreden gelen geri bildirim ile sağlanır.
- Kendi kendine denetimli öğrenme, modelin verinin bir kısmını kullanarak diğer kısım için etiket verisi elde etme yöntemidir. Dolayısıyla tahmine dayalı bir öğrenme türüdür.

Eğitime sıfırdan başlarken, bağlantı ağırlıkları 0 ve 1 arasında rastgele değerlerden seçilir. Ne kadar girdi verisi varsa o kadar sayıda ağırlık verisi ayarlanır. Bağlantı ağırlıkları güncellendikçe, YSA'nın doğruluk performansı da artar. Bir eğitimi gerçekleştirmenin iki yolu vardır. Birincisi YSA'yı rastgele ağırlıklar ile sıfırdan eğitmek, ikincisi ise Transfer Öğrenimi olarak isimlendirilen ve daha önce eğitilmiş bir modelin ağırlıklarını kullanarak üzerine tekrar eğitim yapmaktır.

Yapay sinir ağlarının kurucularından olan McCulloch ve Pitts, yapay sinir ağlarını, giriş verisinin ağırlıklı toplamı olacak şekilde, Eş. 3.5 ile ifade etmişlerdir (McCulloch & Pitts,

1943). Girdi verisinin (x), ağırlık verisi (w) ile çarpımından eşik değeri (u) çıkarılır. Bu işlem girdi verisinin tüm elemanları için yapılır ve toplamı alınır. Toplam sonucu bir aktivasyon fonksiyonundan (θ) geçirilir.

$$y = \theta \sum_{j=1}^n (w_j x_j - u) \quad (3.5)$$

Bu çalışmada denetimli öğrenme gerçekleştirildiği için bu konu üzerinde durulmuştur. Denetimli öğrenmede 3 temel unsur vardır; girdi verisi, girdi verisine ait etiket verisi ve modelin girdi verisi için oluşturduğu çıktı verisi. Model girdi verisini ağırlıklarla çarpar ve aktivasyon fonksiyonu sonucunda bir çıktı üretir. Bu çıktı her zaman girdi verisinin gerçek sonucu yani etiket verisi ile aynı olmayabilir. Bu duruma hata denir. Hata düzeltme yöntemi kullanılarak bağlantı ağırlıkları güncellenir ve çıktı verisi etiket verisine eşitlenmeye çalışılır (Jain vd., 1996). Hata düzeltme algoritması, w ağırlıklar, u eşik değeri, y modelin çıktısı, d gerçek çıktı, x girdi olmak üzere ilk olarak başlangıç ağırlıklarının ve eşik değerinin rastgele sayılarla belirlenmesi ile başlar. Daha sonra x girdi verisine göre model tarafından bir y çıktısı oluşturulur. d ve y kıyaslanarak ağırlıklar Eş. 3.6'daki gibi güncellenir.

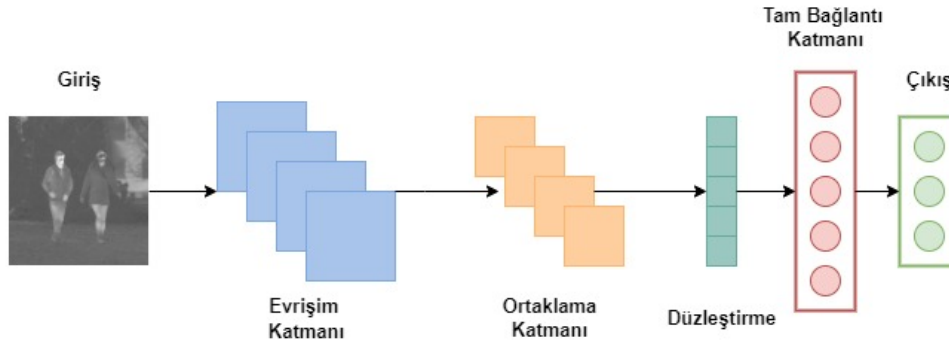
$$w_j(t + 1) = w_j(t) + n(d - y)w_j \quad (3.6)$$

Burada t yineleme sayısı ve n adım boyutudur. Bu işleme hata değeri minimum seviyeye gelene kadar devam edilir.

Evrişimsel Sinir Ağları (CNN- ESA)

Yapay zekanın bir alt başlığı olan derin öğrenme temelde birden fazla gizli katmana sahip bir YSA modelidir. Evrişimsel Sinir Ağları (ESA), derin öğrenme modellerinin en gelişmişlerinden ve en çok kullanılanlarından biridir. Genel çalışma mantığı olarak YSA'lara çok benzeyen ESA'ların başlıca kullanıldığı alan görüntü sinyalleridir ve girdi olarak matris formatında resimler kullanılmaktadır. Çıktı olarak ise resimlerin ait oldukları sınıflarla ilgili olan hata fonksiyonlarını vermektedir (O'Shea & Nash, 2015). Nesne tespiti çalışmalarında bu hata fonksiyonu hem nesnenin sınıfını hem de konum bilgisini barındırır. Dolayısıyla çıktı olarak sınıflandırma hatası ve regresyon hatası bir arada bulunan bir hata fonksiyonu elde edilir.

ESA'lar bir görüntü üzerindeki özellikleri çıkarmada oldukça başarılılardır. Derin katmanlara sahip bu modeller, ilk katmanlarında görüntü üzerinden basit özellikleri çıkarırken, katmanlar derinleştikçe yüksek seviye özellikler olarak isimlendirilen özellikleri sınıflandırmaya başlarlar (Albawi vd., 2018a). Bu sayede nesne tespit uygulamalarında yaygın olarak kullanılırlar.

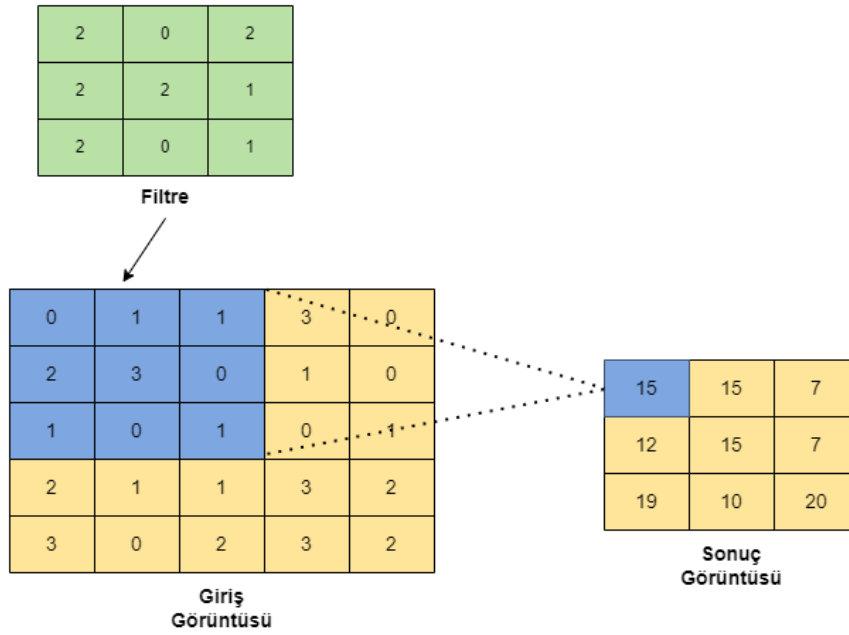


Şekil 3.9. Bir ESA yapısı

Şekil 3.9'da bir ESA'nın yapısı verilmiştir. ESA'lar 6 kısımdan oluşmaktadır. Giriş verisi, bir resmin pikselleri olup; çıkış verisi, o girdiye ait sınıf olasılıklarını verir. Çıkış sayısı modelin sınıf sayısı kadardır. Giriş verisi eğer bir resim ise, genişlik (w), yükseklik (h) ve kanal sayısı (c) cinsinden $w * h * c$ şeklinde ifade edilir. Aynı zamanda bu değer giriş verisinin uzunluğunu vermektedir. Renkli görüntülerde kanal sayısı bir diğer adıyla derinlik 3 olurken, gri görüntülerde 1 olmaktadır. Giriş ve nöronlar arasındaki bağlantıyı yapabilmek için, $w * h * c * n$ kadar bağlantı ağırlığı gerektirmektedir, burada n nöron sayısını ifade eder. Örneğin, 128×128 boyutunda renkli bir görüntü için 10 adet nöron kullanılmak istenirse, bu $128 \times 128 \times 3 \times 10 = 491520$ adet ağırlık verisine denk gelmektedir. Bunun yerine ESA'lara evrişim katmanı ve ortaklama katmanı eklenmiştir. Daha sonra ise çıkış verisini elde edebilmek amacıyla düzleştirme ve tam bağlantı katmanları bulunmaktadır (Albawi vd., 2018a). Tam bağlantı katmanına kadar olan kısım öznitelik çıkarmaya yararken, sonraki kısım bu öznitelikleri sınıflandırmaya yarar.

İki katmanı arasındaki her nöronun bağlantı kurması mantığına dayanan YSA'ların aksine, ESA'larda kendinden önceki katmanın sadece belli bir kısmına bağlanan evrişim katmanı bulunmaktadır. Bu katman kendinden önceki katmanın öznitelik haritasını çıkarmaya yarar (O'Shea & Nash, 2015). Evrişim katmanı, temelde 3×3 veya 5×5 gibi bir filtrenin evrişim

katmanının girişine uygulanmasıdır, yani filtre matrisinin giriş matrisinde gezdirilmesi ve filtrenin denk geldiği alanlardaki elemanların çarpımlarının toplamının alınması işlemidir. Evrişim işlemi Şekil 3.10'da verilmiştir. Bu sayede girişteki değerin boyutu küçültülmüş ve özellikleri korunmuş olur olur. Bu da hesaplama kapasitesini olumlu yönde etkiler. Evrişim katmanı, bir modelde bir tane olabileceği gibi yüzlerce de olabilmektedir. Evrişim katmanının çok sayıda olması modelin yüksek seviye öznitelikler olarak isimlendirilen öznitelikleri ayırt etmesini sağlar. Bu yüksek seviye öznitelikler, nesnelerin kısımları, renkleri, şekilleri vb. gibi olabilir.



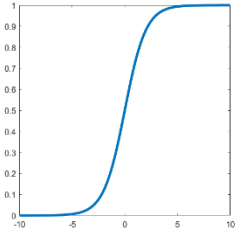
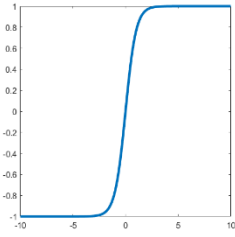
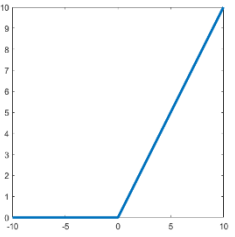
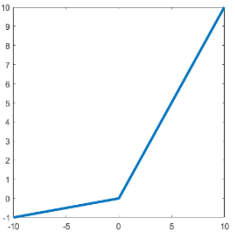
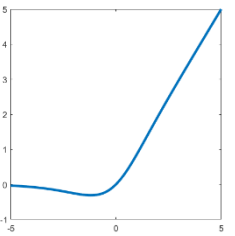
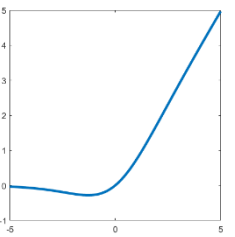
Şekil 3.10. 5x5'lik bir veriye 3x3'lük bir filtrenin uygulanması

Şekil 3.10'da görüldüğü gibi çıkış, giriş verisinden küçük boyutlu ama giriş verisinin özelliklerini saklayan bir matristir. Burada uygulanacak filtrenin boyutu, doğrudan çıkış matrisinin boyutunu belirleyeceğinden dikkat edilmesi gereken hiper parametrelerden biridir. Bir diğer hiper parametre ise filtrenin adım sayısı yani, kaç eleman sıklıkla giriş verisi üzerinde gezilecek olmasıdır. Bu hiper parametre de çıkış boyutunu etkiler. Giriş görüntü boyutu $N * N$, filtre boyutu $F * F$ ve filtre adım sayısı S olarak ifade edilirse çıkış matrisi Eş. 3.7'deki gibi hesaplanabilir.

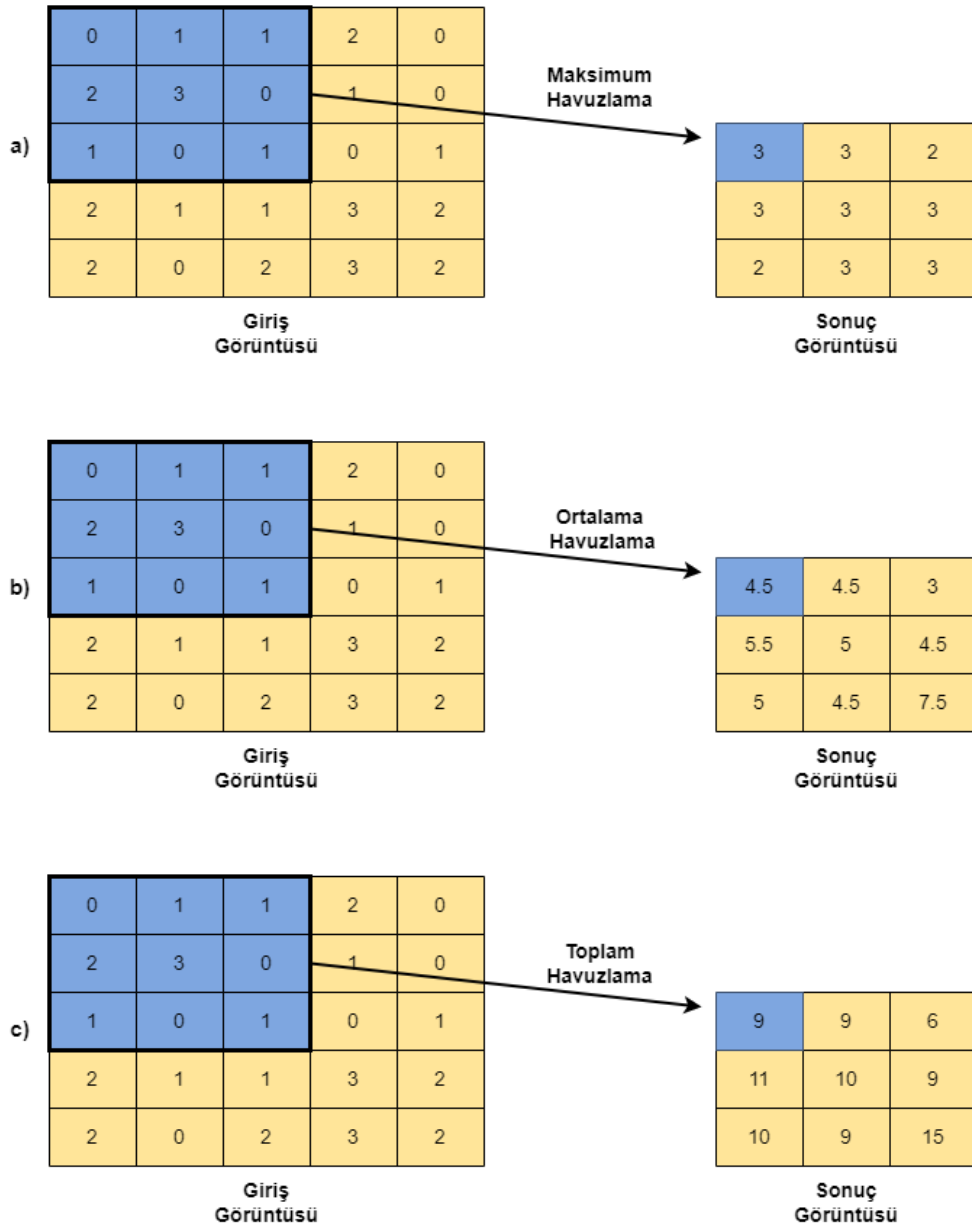
$$\frac{N - F}{S} + 1 \quad (3.7)$$

Evrişim katmanının ardında çıkış verisini düzenlemek için lineer olmayan bir fonksiyon kullanılır. Sıklıkla kullanılan bazı lineer olmayan fonksiyonlar Çizelge 3.4'te verilmiştir.

Çizelge 3.4. Sıklıkla kullanılan bazı lineer olmayan fonksiyonlar

Fonksiyonun Adı	Grafiği	Formülü
Sigmoid		$\frac{1}{1 + e^{-x}}$
Tanh		$\frac{e^x - e^{-x}}{e^x + e^{-x}}$
RELU		$\max(0, x)$
Leaky-RELU		$\max(0.1x, x)$
Mish		$x * \tanh(\ln(1 + e^x))$
SiLU		$x * \frac{1}{1 + e^{-x}}$

ESA'nın bir diğerkatmanı olan havuzlama (ortaklama) katmanında, katmanlar arasındaki karmaşıklığı azaltmak amacıyla çıktının boyutunu küçültme işlemi yapılmaktadır. Girdinin bir resim olduğu bir modelde, havuzlama katmanı görselin matris boyutunu yani çözünürlüğünü düşürmeye yarar (Albawi vd., 2018a). Günümüzde Evrişimsel Sinir Ağları'nda kullanılan havuzlama katmanları; maksimum havuzlama, ortalama havuzlama ve toplam havuzlama olarak isimlendirilir. Burada girdi belli alanlara ayrılır, bu işlem evrişim katmanındaki filtrenin görüntü üzerinde gezmesi ile aynı mantıktır ve her alana ilgili havuzlama işlemi uygulanarak çıkışa aktarılır. Şekil 3.11'de üç farklı havuzlama örneği verilmiştir. Maksimum havuzlama yöntemi en çok kullanılan yöntemdir (Albawi vd., 2018a). Şekil 3.11a'da gösterilmiştir. Bu havuzlama katmanı, girdinin ilgili bölgesindeki elemanların en büyüğünü çıktı olarak verir. Bu havuzlama yönteminin yaygın olarak kullanılmasının sebebi, çıkışının girişindeki elemanlardan bazılarını saklıyor olmasıdır. Toplama havuzlama, girdi görüntüsündeki havuzlama katmanının boyutuna denk gelen elemanlarının toplamının alınmasıdır. Şekil 3.11b'de toplama havuzlama çalışma mantığı verilmiştir. Ortalama havuzlama ise, toplam havuzuna benzer olarak bu sefer toplamın eleman sayısına bölünüp aritmetik ortalamasının alınması işlemidir. Şekil 3.11c'de örneği verilmiştir.



Şekil 3.11. Havuzlama katmanları, a) maksimum havuzlama, b) ortalama havuzlama ve c) toplam havuzlama

Burada havuzlama katmanı 3x3, girdi 5x5 boyutlarda ve filtre adım sayısı 1 olan örnekler verilmiştir. Fakat en sıklıkla kullanılan havuzlama katmanı boyutu 2x2 ve adım sayısı 2 olarak alınmaktadır (Albawi vd., 2018a). Burada dikkat edilmesi gereken havuzlama katmanının evrişim katmanının aksine konuma dayalı bilgiyi saklamıyor olmasıdır. Havuzlama katmanının çıktı boyutu (o), girdi boyutu (i), filtre boyutu (k) ve adım sayısı (s) olmak üzere Eş. 3.8'deki gibi hesaplanabilir.

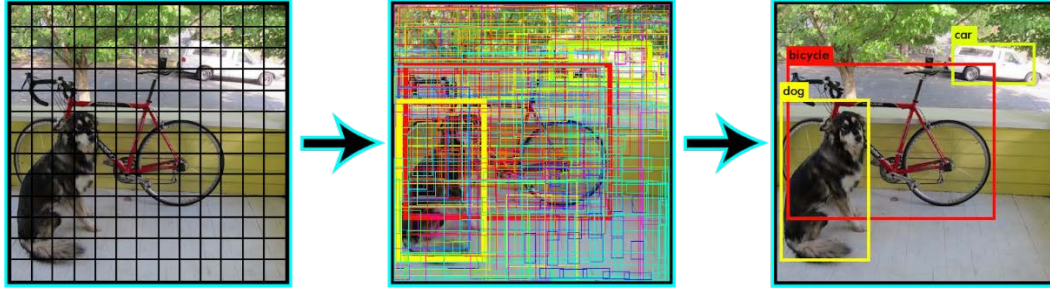
$$o = \frac{i - k}{s} + 1 \quad (3.8)$$

ESA'ların son katmanı tam bağlantı katmanı olup, bu katman yapay sinir ağlarının çalışma mantığı ile aynıdır. Evrişim katmanları ve havuzlama katmanları sonucunda üretilen çıktı tam bağlantı katmanına gelir ve burada Evrişimli Sinir Ağı'nın girişine verilen görüntünün sınıfı tahmin edilir. Tam bağlantı katmanının çıkışı sınıf sayısı adedince. Bu sayede girdi görüntüsünün her sınıf için olasılık dağılımı elde edilir ve belli bir eşik değerinin üzerinde olan olasılık değerine ait sınıf doğru sınıf olarak kabul edilir.

3.1.4. Evrişim tabanlı nesne tespit modelleri

Nesne tespit için kullanılan birçok derin öğrenme mimarisi mevcuttur. Bunlar kendi içlerinde tek aşamalı ve çift aşamalı detektörler olmak üzere ikiye ayrılmaktadır. Tek aşamalı detektörler yüksek performansa sahip olma ve hızlı sonuç verme özellikleri sayesinde sıkça kullanılmaktadırlar (Redmon & Farhadi, 2018). Bu detektörlere YOLO (Redmon vd., 2016), RetinaNet (Lin vd., 2017) ve SSD (W. Liu vd., 2016) modelleri örnek olarak verilebilir.

Yolo, eğitim ve test aşamasında görüntüyü tek seferde inceleme özelliği sayesinde nesne tespit uygulamalarında sıkça kullanılmaktadır. Bu özellik sayesinde Yolo modeli, nesneleri arka plandan daha iyi ayırmakta ve daha hızlı sonuç vermektedir (Redmon vd., 2016). Yolo'nun çalışma prensibi, bir resmi parçalara ayırması ve her parçada nesne aramasıdır. Sonuç olarak her parçada tahmini kutucuklar ve kutucuklara ait güven skorları elde edilir. Güven skorları kutucukların bir nesneye denk gelip gelmediğini ve nesnenin sınıfını ne kadar doğru tahmin ettiğini ifade eden bir birimdir (Redmon vd., 2016). Resmin her parçası için nesne tahminleri yapıldıktan sonra bu tahminler arasından en yüksek güven skoruna sahip olanlar alınır. Bu yöntem maksimum olmayan bastırma (MOB) olarak isimlendirilir. Yolo'nun çalışma prensibi Şekil 3.12'de verilmiştir.



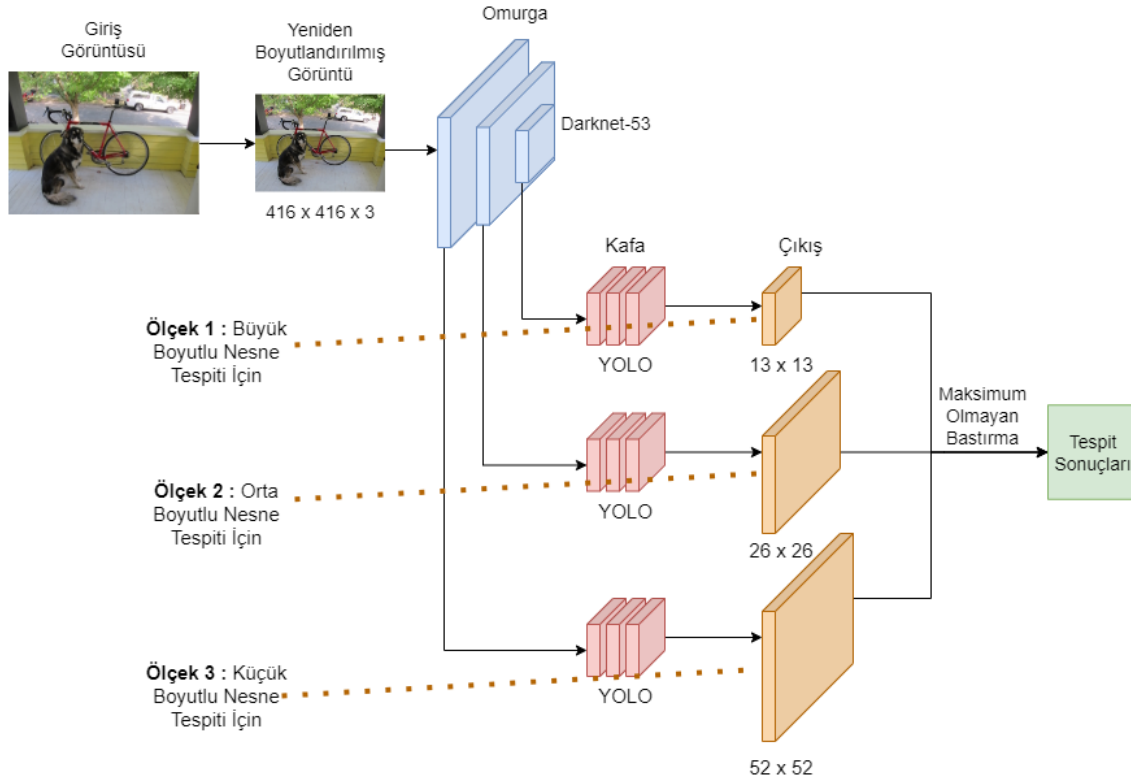
Şekil 3.12. Yolo'nun çalışma prensibi (Redmon vd., 2016)

Yolo, sınıflandırma hatası, konum hatası ve güven hatasından oluşan, kayıp fonksiyonu için toplam-kare hatasını kullanır. Sınıflandırma hatası, her sınıf için olasılık tahminin karesidir. Konum hatası, tahmini sınırlayıcı kutunun konumuna ve boyuna ait hatadır. Güven hatası ise tahmini sınırlayıcı kutuda nesne olup olmadığını belirtir.

Yolo, çok ölçekli eğitim modeline sahiptir yani eğitim sürecinde görüntünün boyutunda değişimler olmaktadır. Bu değişim 32 ile orantılıdır, bu sebeple modelin giriş boyutu 32'nin katları olacak şekilde seçilir (Redmon vd., 2016). Yolo'nun temeli Evrişimli Sinir Ağı'na dayanmakta ve v3, v4, v7 gibi birçok gelişmiş modeli bulunmaktadır.

YOLOv3

Yolo'nun tasarımının geliştirilerek elde edilen yeni modellerinden biri Redmon ve arkadaşları tarafından üretilen Yolov3'tür (Redmon & Farhadi, 2018). Model genel olarak Şekil 3.13'te gösterildiği gibi özellik çıkarıcı (omurga), özellik haritası çıkarma (boyun) ve detektör (kafa) olmak üzere üç kısımdan oluşur. Özellik çıkarıcı ve özellik haritası, derin öğrenme modelleri için en önemli kısımlardandır (Ren vd., 2017).



Şekil 3.13. Yolov3 mimarisi

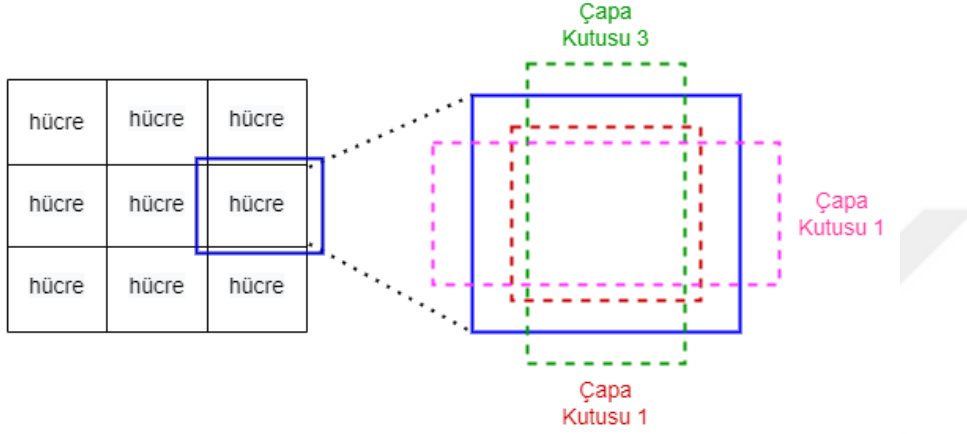
Omurga ismi verilen özellik çıkarma bloğunda görsellerden öznetelik çıkarımı yapılır. Yolov3 omurga olarak 53 evrişim katmanına sahip Darknet53'ü kullanır. Evrişim katmanlarının ardından Leaky-ReLU aktivasyon fonksiyonu eklenmiştir. Bu blokta ayrıca aşırı uyum gösterme olarak isimlendirilen modelin girdi verisini ezberleme durumuna karşılık toplu normalleştirme yöntemi kullanılmıştır. Özellik çıkarma gerçekleştikten sonra detektör bloğu sayesinde tahmini sınırlayıcı kutular ve sınıf bilgileri elde edilir. Bu iki blok arasında bulunan boyun adı verilen blokta, sınıflandırma için özellik haritaları çıkarılır (Bochkovskiy vd., 2020). Bu özellik haritaları bir özellik piramit ağına (FPN) ile farklı ölçeklerde gerçekleştirilir. Bunun amacı daha sıkı bilgilere sahip ilk katmanların daha derin bilgilere sahip son katmanların bilgilerine erişebilmesidir (Redmon & Farhadi, 2018).

Literatürde omurga, boyun ve kafa kısımları için birçok model kullanılmaktadır. Bazı örnekler aşağıda verilmiştir.

- Omurga; VGG16, ResNet, Darknet53, vb.
- Boyun; Özellik Piramit Ağı (FPN), Yol Toplama Ağı (PAN), BiFPN, NAS-FPN, vb.

- Kafa; Yolo, Ssd, RetinaNet, vb.

Yolo'nun aksine Yolov3'te sınıflandırma işlemi, softmax fonksiyonu yerine lojistik kayıp fonksiyonu ile gerçekleştirilir. Aynı şekilde sınıflandırma hatasını hesaplamak için Yolo'nun aksine Yolov3, ikili çapraz entropi hata fonksiyonunu kullanır (Redmon & Farhadi, 2018). Bu iki değişiklik modelin hesaplama karmaşıklığını azaltmak için yapılan değişikliklerdir.



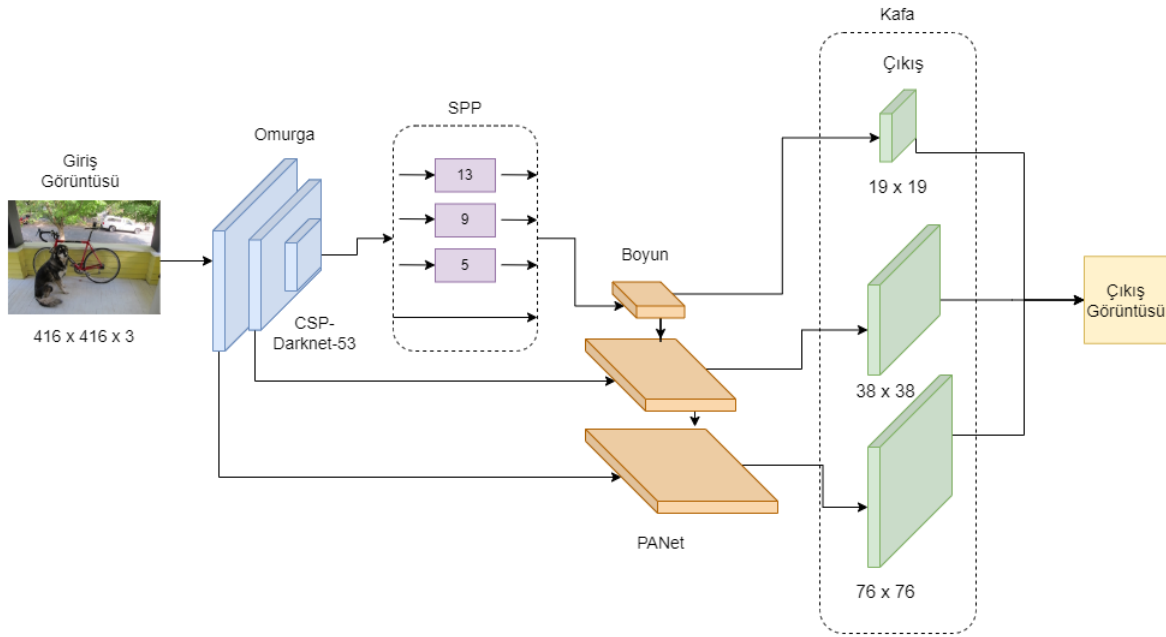
Şekil 3.14. Bir hücredeki çapa kutularının gösterimi

Yolov3, aynı Yolo'da olduğu gibi görüntüyü hücrelere ayırmaktadır. Şekil 3.14'te gösterildiği gibi her parça için 3 adet çapa kutusu bulunmaktadır. Bunlar 52x52, 23x23 ve 13x13 boyutlarındadır. Çapa kutuları, çeşitli ölçek ve oranlarda önceden belirlenmiş kutulardır (Y. Liu, Sun, vd., 2021a). Bu kutular sayesinde farklı boyutlarda özellik haritaları üretilir, bu da farklı boyutlardaki ve şekillerdeki nesnelerin tespit edilmesini sağlar (Bhowmick vd., 2022).

YOLOv4

Yolov3'ün bir üst modeli Bochkovski ve arkadaşları tarafından geliştirilmiştir (Bochkovski vd., 2020). Yolov4, omurga olarak Çapraz Aşamalı Kısmi Ağın (CSPNet) bir çeşidi olan CSP-Darknet53, boyun olarak Yol Toplama Ağı (PANet) ve kafa olarak Yolov3'ü kullanır. Öznitelik çıkarma Mekânsal Piramit Havuzlaması (SPP) ile gerçekleştirilir. Evrişim katmanlarının ardından kullanılan aktivasyon fonksiyonu Mish olarak seçilmiştir (Bochkovski vd., 2020).

Yolov4, eğitim aşamasında girdi görüntülerine bazı veri çoğaltma yöntemleri uygular. Bunlar, görüntünün parlaklığını, kontrastını değiştirmek, görüntüye gürültü eklemek, görüntüyü kırpmak, çevirme ve döndürme gibi yöntemlerdir ve piksel bazlı olarak gerçekleştirilir (Bochkovskiy vd., 2020). Düzenleştirme metodu olarak, Yolov3'ün aksine DropBlock yöntemi kullanılmıştır. Normalleştirme yönteminin için ise syncBN yöntemi seçilmiştir. Şekil 3.15'te Yolov4'e ait mimari verilmiştir.

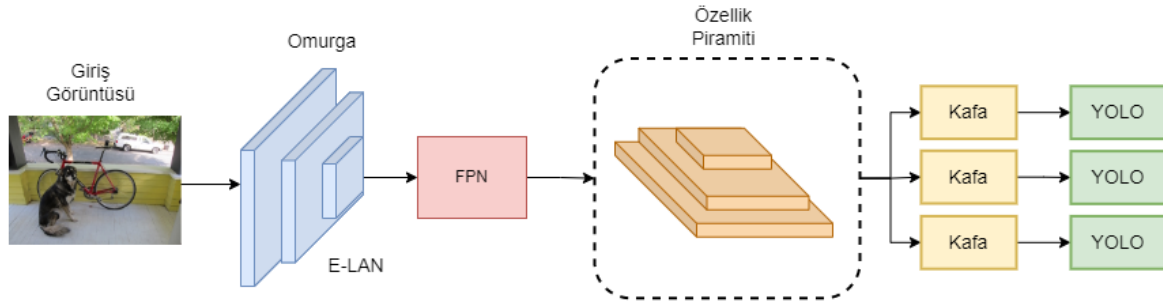


Şekil 3.15. Yolov4 mimarisi

YOLOv7

Yolov7 (C.-Y. Wang vd., 2022), Yolov4'ten türetilmiş yeni bir modeldir. MS COCO veri seti ile sıfırdan eğitilerek elde edilmiştir. Nesne tespitteki hızı ve doğruluğu artırmak adına Yolov7 omurgasında E-LAN (Genişletilmiş Verimli Katman Toplama Ağı) yapısı kullanılmaktadır. Yolov7'de birden fazla kafa yapısı mevcuttur, bunlardan en son çıktıyı üreten kısma öncelikli kafa (the lead head) denilmektedir. Diğerleri ise yardımcı kafa (the auxiliary head) olarak isimlendirilmektedir. Aktivasyon fonksiyonu olarak SiLU fonksiyonu kullanılmıştır (C.-Y. Wang vd., 2022). Yolov7 modeli, doğruluğu ve hızı ayarlamak için Model Ölçeklendirme yöntemini kullanır. Bu ölçekleme, giriş görüntüsünün boyutu, kanal sayısı, katman sayısı ve özellik piramitlerinin sayısı gibi

parametreler üzerinden yapılır. Yolov7, en iyi ölçeklendirme faktörlerini bulmak için NAS (Ağ Mimarisi Araması) yapısını kullanır. Yolov7’de kanal sayısı yani derinlik 1,5 kat ve katman sayısı yani genişlik 1,25 kat artırılarak ölçekleme yapılmıştır. Yolov7, model düzeyinde yeniden parametrelendirme kullanılarak farklı eğitim setleriyle eğitilen farklı modellerin belli adımlardaki ağırlıklarının ortalaması kullanılarak üretilmiştir (C.-Y. Wang vd., 2022). Şekil 3.16’da Yolov7 mimarisi verilmiştir.



Şekil 3.16. Yolov7 mimarisi

Termal YOLO

Huda ve arkadaşlar tarafından insan tespiti için geliştirilen Termal Yolo (Huda vd., 2020a), Yolov3’ün önceden eğitilmiş ağırlıkları kullanılarak, yine Huda ve arkadaşları tarafından geliştirilen bir kızılötesi veri seti ile eğitilmesi sonucu üretilmiştir. Eğitim aşaması öğrenme oranı 0,001, momentum 0,9 ve sönüm oranı 0,0005 olacak şekilde seçilmiştir. Toplamda 40000 iterasyon eğitilmiştir. İnsan tespiti için kamuya açık veri setlerinde Termal Yolo modelinin iyi performanslarının sergilediğini ve yeni alana adapte olma özelliğinin bulunduğunu belirten Huda ve arkadaşlar, görüntülerle beraber, bakış açısı, ısı efektleri, görüntü artefaktları ve hava efektleri özellikleri göz önüne alındığında modelin performansının daha da arttığı ifade etmişlerdir. Yolov3, Yolov4, Yolov7 ve Termal Yolo modellerinin mimarileri özetle Çizelge 3.5’teki gibidir.

Çizelge 3.5. Yolov3, Yolov4, Yolov7 ve Termal Yolo mimarilerinin genel kısımları

	Yolov3	Yolov4	Yolov7	Termal Yolo
Omurga	Darknet53	CSP-Darknet53	E-LAN	Darknet53
Aktivasyon fonksiyonu	Leaky-ReLU	Mish	SiLU	Leaky-ReLU
Boyun	FPN	SPP + PAnet	SPPCSP	FPN
Kafa	Yolo	Yolo	Yolo	Yolo
Hata Fonksiyonu	İkili çapraz entropi kaybı	CIoU Kaybı	CIoU Kaybı	İkili çapraz entropi kaybı

Model hiper parametreleri

Tüm yapay sinir ağlarında eğitimden önce belirlenmesi gereken bazı özellikler vardır. Bu özellikler hiper parametreler olarak isimlendirilir. Hiper parametreler sınıf sayısı, filtre sayısı ve boyutu, aktivasyon fonksiyonu, öğrenme oranı gibi modelin performansını etkileyen parametrelerdir. Yolo modelleri için kullanılan hiper parametreler aşağıdaki gibi ele alınabilir.

- Toptan sayısı, tek seferde modele verilecek resim sayısını ifade eder. Biz bu çalışmada 64 olarak kullandık.
- Alt bölümler sayısı, toptan sayısının kaç seferde modele verileceğini ifade eder. GPU'nun gücü ile alakalıdır. 8'in katları şeklinde artmaktadır. Biz bu çalışmada GPU'nun gücüne göre alt bölüm sayısını 64 olarak seçtik. Sonuç olarak görüntüler $64/64=1$ şeklinde modele verilir.
- Genişlik ve yükseklik, modelin giriş boyutunu ifade ederler. 32'nin katları şeklindedir. Bu çalışmada küçük nesne tespit performansını artırabilmek için 608×608 şeklinde seçilmiştir. Fakat modelin giriş boyutu büyüdükçe eğitim süresi artmaktadır.
- Öğrenme Oranı, modelin ağırlıkları güncelleme hızı olarak ifade edilebilir. Eğer öğrenme hızı büyük bir değer seçilirse optimum ağırlık değerleri elde edilemezken, çok düşük seçildiğinde de optimum ağırlık değerlerine ulaşmak oldukça fazla zaman alacaktır. Bazı çalışmalarda eğitim aşaması boyunca öğrenme oranı azaltılmaktadır (Bhowmick vd., 2022; Wu vd., 2021). Bu çalışmada öğrenme oranı başlangıç değeri 0,001 olacak şekilde toplam iterasyon sayısının %80 ve %90'ında öğrenme oranını 10 kat düşürecek şekilde seçilmiştir. Öğrenme oranı GPU sayısına

göre de değişmektedir. Bizim kullandığımız iş istasyonu tek GPU'ya sahip olduğu için öğrenme oranımız 0,0001/1 şeklinde hesaplanmaktadır.

- Momentum, modeli iyileştirmek için kullanılan bir parametredir. Hata fonksiyonunun düşürülmesinde kullanılır. Bu çalışmada 0,9 olarak seçilmiştir.
- Sönüm, modelin karmaşıklığını azaltabilmek için hata fonksiyonunun küçük bir sayı ile çarpılmasıdır. Bu çalışmada 0,0005 olarak kullanılmıştır.
- Doygunluk, parlaklık ve ton değerleri eğitim aşamasında farklı resimler üretmek için kullanılır. Bu çalışmada orijinal Yolo modelleriyle aynı olacak şekilde sırasıyla 1,5; 1,5 ve 0,1 olarak seçilmiştir.

Yukarıda model için gerekli genel hiper parametreler verilmiştir. Bu hiper parametrelerin ardından evrişim katmanı parametreleri gelmektedir. Daha sonra ise Yolo katmanı parametreleri verilmiştir.

Evrişim katmanı parametreleri;

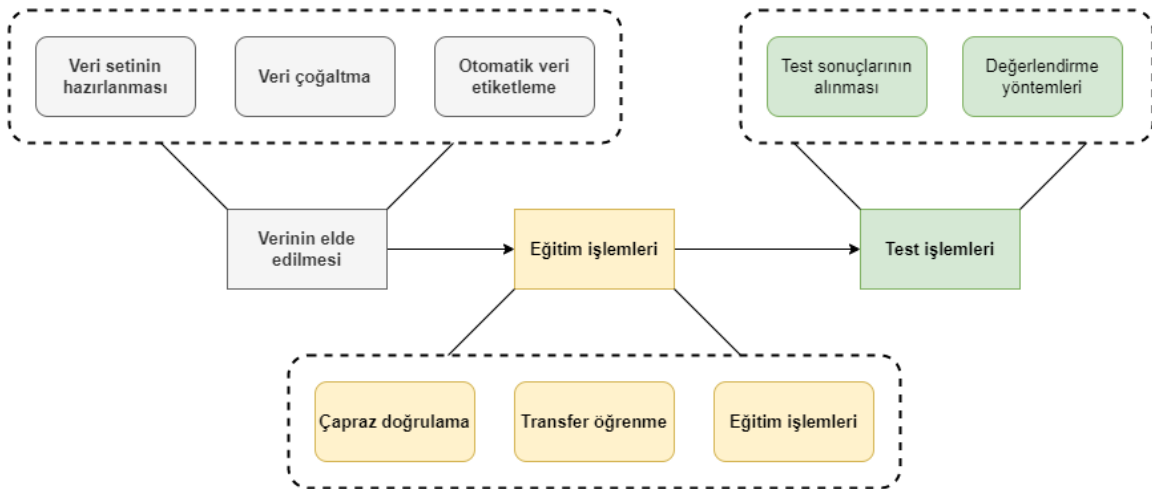
- Toplu normalleştirme, Evrişimli Sinir Ağları'na aşırı uyum göstermeyi engellemek için evrişim katmanı ve aktivasyon fonksiyonu arasına eklenen bir katmandır (Wu vd., 2021).
- Filtreler, resimden anlamsal bilgileri çıkarmak için kullanılırlar.
- Aktivasyon fonksiyonu, evrişim katmanının ardından gelen lineer olmayan katmandır. Katman çıktısını yorumlamayı sağlar. En yaygın kullanılan aktivasyon fonksiyonları, Leaky Relu, Relu, Sigmoid, Mish vb.dir.

Yolo katmanı parametreleri;

- Çapa kutusu, görüntünün ayrılmış parçalarındaki nesneleri bulmak için belirlenir. 12, 16, 19, 36, 40, 28, 36, 75, 76, 55, 72, 146, 142... gibi değerler alır.
- Sınıf sayısı, tespit edilecek etiketleri belirtir. Bu çalışma için sınıf sayısı insan ve araç olmak üzere 2 adettir.
- Filtre boyutu, Yolo katmanındaki filtre boyutudur. Sınıf sayısına göre belirlenir. Bu çalışma için $3 \times (2+5) = 21$ olarak belirlenmiştir. Burada 3 Yolo için tahmin edilen çerçeve sayısını, 2 sınıf sayısını ve 5 ise sınırlayıcı kutu bilgilerini ifade eder.
- Titreşim (jitter), Yolo katmanına gelindiğinde resmin boyutunu değiştirerek ve resmi kırarak veri sayısının artırılması yöntemidir.

3.2. Yöntem Ve Deneyler

Bu tez çalışması kapsamında, literatürde nesne tespit uygulamalarındaki gerçek zamanlı çalışmaları ile dikkat çeken Yolov3, Yolov4, Yolov7 ve Termal Yolo modellerinin kızılotesi hedef tespiti için performans analizleri yapılmıştır. Bu çalışma için AALARTDATA, LTIR ve CAMEL isimli kamuya açık veri setlerinden 1000 adet resim seçilmiş ve bu resimler döndürme, kırpmaya, kenar doldurma ayna görüntüsü alma gibi görüntü işleme yöntemleri kullanılarak çoğaltılmıştır. Sonuçta elde edilen 48000 adet görüntü 4 kat çapraz doğrulama yöntemi ile eğitim, doğrulama ve test veri seti olarak ayrılmıştır. Modeller, bu eğitim verileri ile her adımda 375 iterasyon olacak şekilde toplamda 100 adım eğitilmiştir (Eğitim aşamasında kullanılan tüm parametreler sırasıyla verilecektir.). Nesne tespit için savunma sanayi uygulamaları ön planda tutularak insanlar ve araçlar tercih edilmiştir. Çalışma Linux Ubuntu işletim sistemi üzerinden NVIDIA Quadro P4000 GPU ile gerçekleştirilmiştir. Tez çalışmasının akış şeması Şekil 3.17’de verilmiştir.



Şekil 3.17. Tez çalışması akış şeması

Şekil 3.17’de görüldüğü gibi ilk aşama, verinin elde edilmesi olup bu aşamada ilk olarak veri seti hazırlanmıştır. Daha sonra veri seti çoğaltılmış ve son olarak ta otomatik veri etiketlenmesi yapılır. İkinci aşama olan eğitim kısmına geçilmiştir. Eğitim sürecinde çapraz doğrulama ve transfer öğrenmesi metotları kullanılmıştır. Eğitim sürecinin ardından test işlemleri ve sonuçların değerlendirilme kısmı bulunmaktadır. Test işlemleri bir sonraki bölüm olan “Bulgular ve Tartışma” başlığı altında incelenmiştir.

3.2.1. Verinin elde edilmesi

Bu bölümde kızılötesi görüntü veri setinin oluşturulma süreçlerinden bahsedilmiştir. Veri setinin elde edilme süreci sırasıyla veri setinin hazırlanması ve veri çoğaltma başlıkları altında incelenmiştir. Daha sonra derin öğrenme modellerinin eğitimi için gerekli olan etiketleme işlemlerinden bahsedilmiştir.

Veri setinin hazırlanması

Veri toplaması, saklaması, bu verileri etiketlemesi günümüzde hala zaman alıcı ve zor bir süreçtir (Huda vd., 2020a). Bu durum veri sayısı arttıkça daha zor bir hal almaktadır. Bu zorluk araştırmacıları internette ulaşılabilir olan, başkaları tarafından üretilmiş ve etiketlenmiş verileri kullanmaya itmektedir. Fakat söz konusu kızılötesi görüntüler olduğunda, kızılötesi görüntülerin askeri ve tıbbi uygulamalarda kullanımlarından dolayı ulaşılabilir olmaları pek mümkün olmamaktadır. Ayrıca kızılötesi görüntü veri setleri bir renkli görüntü veri seti gibi bol miktarda görüntü içermemektedir.

Çalışma kapsamında internette ulaşılabilir olan 3 adet kızılötesi veri setinden insan ve araçların ön planda olduğu 1000 adet resim seçilip harmanlanmıştır. Veri setlerinin harmanlamanın amacı, tek tip bir veri seti yerine, çeşitli boyutlarda, renk yoğunluklarında ve pozisyonlarda nesneler elde etmektir. Bu sayede veri seti modele verildiğinde model nesnenin olması muhtemel birçok farklı durumunu görmüş olacaktır. Hali hazırda kullanılan kızılötesi veri setleri genellikle belli bir alanda belli bir süre çekilen görüntüleri içermektedir. Ve bu görüntülerdeki nesneler çoğu zaman görüntünün büyük bir çoğunluğunu kaplamaktadır. Bu da veri setinin gerçek hayat senaryolarından uzak kalmasına sebep olur (H. Zhang vd., 2018). Bu sebeple bu çalışmada veri setini olabildiğince çeşitlendirerek bu eksikliğin önüne geçilmeye çalışılmıştır. Veri setinde yer alan bazı resimler özellikle nesne tespit modelini zorlayacak şekilde seçilmiştir. Seçilen görüntülerde insan ve araçların farklı pozisyonlarda yer alıyor olmasına özen gösterilmiştir. Ayrıca modellerin nesneleri arka plandan ayırma performanslarını inceleyebilmek için farklı arka plana sahip resimler seçilmiştir. Veri setinden bazı örnek görüntüler Çizelge 3.6'da verilmiştir. Veri setlerin özellikleri, 1000 resim içerisindeki

dağılım yüzdeleri ve örnek görselleri çizelgede verilmiştir. (Toplamda 18 senaryodan sadece 3'ü örnek olarak gösterilmiştir.)

Çizelge 3.6. Veri setlerinin özellikleri

Veri Seti İsmi	Görüntülerin Boyutu	Dağılım Yüzdesi	Bulunduğu Yelpaze	Örnek Görüntü
AAlart Data (H. Zhang vd., 2018)	640x513	%33	8–15 μm (LWIR)	
CAMEL (Gebhardt & Wolf, 2019)	336x256	%4	7,5-13,5 μm (MWIR ve LWIR)	
LTIR (Berg vd., 2015)	320x256 1920x480 324x256 640x480	%63	3-8 μm (MWIR)	

Veri setindeki görüntülerde bulunan nesnelerin, boyutlarına göre dağılımları Şekil 3.18 ve Şekil 3.19'da verilmiştir. COCO değerlendirme metrikleri göz önüne alınarak, nesnenin boyutu 32x32'den küçük nesneler küçük, 32x32 ve 96x96 arasında olan nesneler orta ve 96x96'dan fazla olan nesneler büyük nesne olarak kategorize edilmiştir. Fakat nesnelerin tamamı (Zhu vd., 2016) numaralı kaynakta verilen küçük nesne tarifine göre küçük nesne sayılmaktadır. Yani aslında nesneler görüntü boyutlarının %20'sinden küçük boyuttadır.

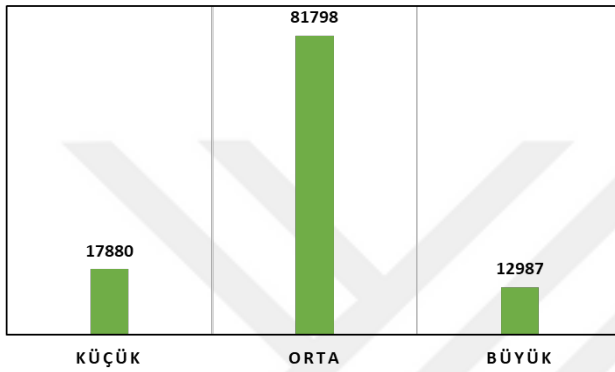
Veri setinin Yolo modelleri ile eğitilebilmesi için uygun formatta nesnelerin işaretlenmesi gerekmektedir. Etiketleme işlemi yapılırken makesense.ai programı kullanılmıştır. Toplamda 112665 adet insan 19986 adet araç kutu içerisine alınıp işaretlenmiştir. Etiketleme verisi Yolo formatı olarak kaydedilmiştir. Yolo formatında her nesnenin işaretlenmiş gerçek sınırlayıcı kutusu 5 elemandan oluşur (Redmon vd., 2016). Bunlar nesnenin ait olduğu sınıf, sınırlayıcı kutunun merkezinin x koordinatı, y koordinatı, kutunun uzunluğu ve genişliği olmak üzere sırasıyla c, x, y, w, h şeklinde ifade edilir. Bunun yanı sıra tahmini sınırlayıcı kutular için birde güven skoru (G) bulunmaktadır. Bu

değerler görüntünün enine ve boyuna bölünerek normalize edilmiş halde bulunur (Redmon vd., 2016). Örnek bir etiket verisi Eş. 3.9 ve Eş. 3.10'da verilmiştir. Bu çalışma için sınıf verisi, insanlar için 0 değerini alırken araçlar için 1 değerini alır.

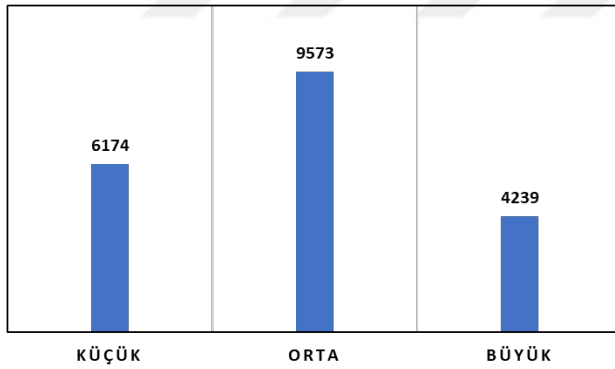
$$\text{gerçek sınırlayıcı kutu} : \{c, x/W, y/H, w/W, h/H\} \quad (3.9)$$

$$\text{tahmini sınırlayıcı kutu} : \{c, x/W, y/H, w/W, h/H, G\} \quad (3.10)$$

Burada W, H görüntünün boyutlarıdır.



Şekil 3.18. Görüntülerdeki insanların boyutlara göre sınıflandırılması



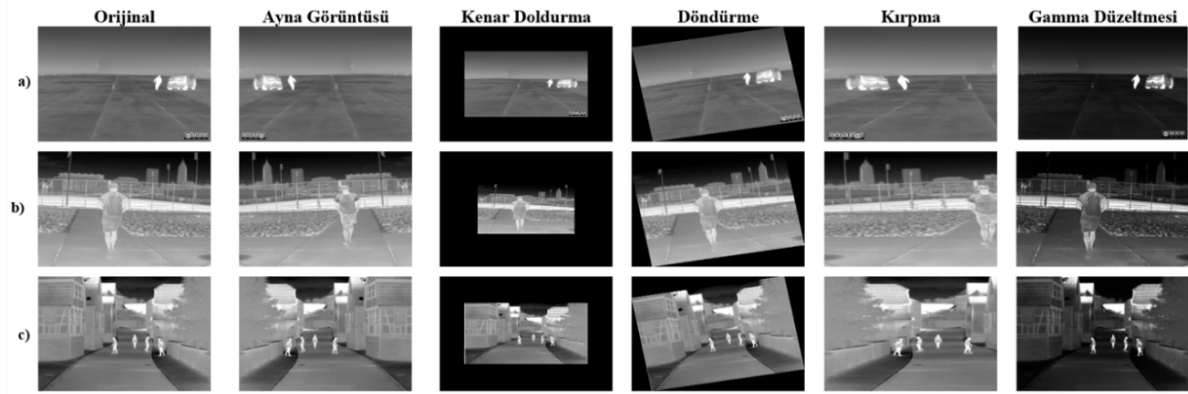
Şekil 3.19. Görüntülerdeki araçların boyutlara göre sınıflandırılması

Veri çoğaltma

Günümüzde çoğu görüntü veri kümesi renkli görüntülerden oluşmaktadır. Bunun yanı sıra kızılötesi görüntü veri kümesi sayısı renkli görüntü veri kümesi sayısına oranla oldukça azdır. Bu durum kızılötesi nesne tespit çalışmalarını sınırlandırmaktadır (Du, Zhang, Zhang, & Xu, 2021). Bu çalışma için elde edilen toplamda 1000 adet kızılötesi görüntü verisi verimli bir nesne tanıma modeli oluşturmak için yeterli olmamaktadır. Ayrıca renkli görüntüler ile eğitilmiş ve transfer öğrenmesi yöntemi ile kullanılacak modelleri yeni

termal alana adapte etmek için de yeterli olmayacaktır. Bu sebeple veri çoğaltma adı verilen yöntemler ile verilerin çoğaltılması ve yeterli sayıda veri elde edilmesi sağlanır. Aynı zamanda veri çoğaltma yöntemi ile modelin veriyi ezberlemesinin önüne geçilir (Jangblad, 2018).

Veri çoğaltma, görüntüye döndürme, kırpma, gürültü ekleme gibi görüntü işleme yöntemlerini uygulayarak yeni görüntüler elde etme özelliğidir (Shorten & Khoshgoftaar, 2019). Bu yöntem derin öğrenme uygulamalarında yaygın olarak kullanılmaktadır. Bu çalışmada topladığımız 1000 adet resimden 48000 adet resim elde etmek için, Şekil 3.20’de görüldüğü gibi ayna görüntüsü alma, kenar doldurma, döndürme, kırpma ve gama düzeltmesi yöntemlerini uygulanmıştır.



Şekil 3.20. Artırılmış resim örnekleri a) AAlartdata veri setinden alınan örnek, b) CAMEL veri setinden alınan örnek, c) LTIR veri setinden alınan örnek

Uygulanan veri çoğaltma yöntemleri ve ayrıntıları maddeler halinde verilmiştir.

- Ayna Görüntüsü Alma: Görüntünün x eksenine veya y eksenine göre simetrisinin alınmasıdır. Bu çalışmada y eksenine göre simetri kullanıldı.
- Kenar Doldurma: Görüntünün kenarlarını doldurarak boyut değiştirme yöntemidir. Bu çalışmada küçük boyutlu nesneler elde edebilmek için kenar doldurma yöntemi, orijinal ve ayna görüntüsü alınmış resimlere uygulandı.
- Döndürme: Resimlerin belli açılarda döndürülmesi ile elde edilir. Bu yöntem nesnelerin pozisyonlarının ve şekillerinin değişmesi için yapılır. Gerçeğe yakın resimler elde edebilmek için 25 dereceden büyük döndürmeler yapılmamaktadır. Bu çalışma için orijinal görüntüler ve ayna görüntüsü alınmış görüntüler 5, -5, 10 ve -10 derecelik açılarla döndürüldü.
- Kırpma: Görüntünün sağdan veya soldan belli bir kısmının kesilmesidir. Bu çalışmada büyük ve orta boyutlu nesneler elde etmek için sağdan ve soldan kırpma yöntemleri kullanıldı. Orijinal ve ayna görüntüsü alınmış resimlere uygulandı.

- Gama Düzeltmesi: Resmin parlaklık değerinin değiştirilmesidir. Bu çalışmada termal adaptasyonu güçlendirmek amacıyla 2/5 ve 5/2 gama değerleri için koyu renkli ve açık renkli resimler elde edildi.

Veri çoğaltma yöntemlerine ait kod blokları EK-1’de yer almaktadır.

Otomatik veri etiketleme

Derin öğrenme problemleri için oldukça fazla veriye ihtiyaç duyulur. Fakat bu verileri teker teker manuel olarak işaretlemesi ve etiketlemesi zaman alan ve yorucu bir iştir (Tumas vd., 2021). Bu sebeple biz bu çalışmada, manuel olarak etiketlediğimiz 1000 görüntü verisini çoğaltarak otomatik olarak etiketleyecek bir kod geliştirdik. Her veri artırma yöntemi için ilgili otomatik veri etiketleme koduna ait sözde kodlar Algoritma 1, 2 ve 3’te verilmiştir.

Algoritma 1: Ayna Görüntüsü Alma ve Otomatik Gerçek Kutu Etiketlerini Üretme

Giriş: gerçek_kutu = [Xg1,Yg1,Xg2,Yg2, wg, hg], resim_boyutları = [w, h]
Sonuç: yeni_gerçek_kutu = [Xg1y,Yg1y,Xg2y,Yg2y, wgy, hgy]
 if (Xg1 > w/2) ise
 sonuç Xg1y $\leftarrow$ w/2 - (Xg1 - w/2) ve Xg2y $\leftarrow$ w/2 - (Xg2 - w/2);
 elif (Xg1 < w/2) ise
 sonuç Xg1y $\leftarrow$ w/2 + (Xg1 - w/2) ve Xg2y $\leftarrow$ w/2 - (Xg2 - w/2) ;
 else ise
 sonuç Xg1y $\leftarrow$ Xg1 ve Xg2y $\leftarrow$ Xg2;

 sonuç
 Yg1y $\leftarrow$ Yg1/ (h/2)
 Yg2y $\leftarrow$ Yg2/ (h/2)
 wgy $\leftarrow$ wg/ (w/2)
 hgy $\leftarrow$ hg/ (h/2)

Algoritma 2: Kenar Doldurma ve Otomatik Gerçek Kutu Etiketlerini Üretme

Giriş: gerçek_kutu = [Xg1,Yg1,Xg2,Yg2, wg, hg], resim_boyutları = [w, h]
Sonuç: yeni_gerçek_kutu = [Xg1y,Yg1y,Xg2y,Yg2y, wgy, hgy],
 yeni_resim_boyutları = [wy, hy]
 boyut $\leftarrow$ max (w,h) / 100
 wy, hy $\leftarrow$ (boyut + 3)*100
 Xg1y $\leftarrow$ Xg1 + wy/2 - w/2
 Yg1y $\leftarrow$ Yg1 + hy/2 - h/2
 Xg2y $\leftarrow$ Xg2 + wy/2 - w/2
 Yg2y $\leftarrow$ Yg2 + hy/2 - h/2

 sonuç
 Xg1y $\leftarrow$ Xg1y/ (wy/2)
 Yg1y $\leftarrow$ Yg1y/ (hy/2)
 Xg2y $\leftarrow$ Xg2y/ (wy/2)
 Yg2y $\leftarrow$ Yg2y/ (hy/2)

```

wgy ← wg/ (wy/2)
hgy ← hg/ (hy/2)
Algoritma 3: Döndürme ve Otomatik Gerçek Kutu Etiketlerini Üretme
Giriş: gerçek_kutu = [Xg1,Yg1,Xg2,Yg2, wg, hg], resim_boyutları = [w,
h], döndürme_açısı = angle
Sonuç: yeni_gerçek_kutu = [Xg1y,Yg1y,Xg2y,Yg2y, wgy, hgy],
yeni_resim_boyutları = [wy, hy]
image_center ← (w/2, h/2)
rotation_mat ← cv2.getRotationMatrix2D(image_center, angle, 1.)
abs_cos ← abs(rotation_mat[0, 0])
abs_sin ← abs(rotation_mat[0, 1])
bound_w ← height * abs_sin + width * abs_cos
bound_h ← height * abs_cos + width * abs_sin
rotation_mat[0, 2] ← rotation_mat[0, 2] + bound_w / 2 - image_center[0]
rotation_mat[1, 2] ← rotation_mat[1, 2] + bound_h / 2 - image_center[1]
image_center2 ← (bound_w / 2, bound_h / 2)
rotation_angle ← -1*angle * np.pi / 180
x_sapma ← abs(image_center[0]-image_center2[0])
y_sapma ← abs(image_center[1]-image_center2[1])
Xg1y ← Xg1y/ (wy/2)
Yg1y ← Yg1y/ (hy/2)
Xg2y ← Xg2y/ (wy/2)
Yg2y ← Yg2y/ (hy/2)
wgy ← wg/ (wy/2)
hgy ← hg/ (hy/2)

sonuç
Xg1y ← Xg1y*math.cos(rotation_angle)-Yg1y*math.sin(rotation_angle)
+image_center[0]*(1-math.cos(rotation_angle))
+image_center[1]*math.sin(rotation_angle)+x_sapma
Yg1y ← Xg1y*math.sin(rotation_angle)+Yg1y*math.cos(rotation_angle)
+image_center[1]*(1-math.cos(rotation_angle))-
image_center[0]*math.sin(rotation_angle)+y_sapma
Xg2y ← Xg2y*math.cos(rotation_angle)-Yg2y*math.sin(rotation_angle)
+image_center[0]*(1-math.cos(rotation_angle))
+image_center[1]*math.sin(rotation_angle)+x_sapma
Yg2y ← Xg2y*math.sin(rotation_angle)+Yg2y*math.cos(rotation_angle)
+image_center[1]*(1-math.cos(rotation_angle))-
image_center[0]*math.sin(rotation_angle)+y_sapma

```

Kırpma işlemi, genişliğin 3'te biri olacak şekilde sağdan ve soldan olmak üzere Python komutu sayesinde yapılmıştır. Daha sonra kesilen kısım hariç kalan nesnelerin gerçek kutu değerleri alınmıştır. Gama Düzeltmesi ise yine aynı şekilde Python komutu ile sağlanmıştır ve nesnelerin gerçek kutuları değişmediği için direk aynı değerler alınmıştır.

3.2.2. Eğitim işlemleri

Eğitim aşamasında farklı Yolo modellerinin önceden eğitilmiş ağırlıkları kullanılmıştır. Veri seti 4 katlı çapraz doğrulama ile ayrılmış ve her kat için eğitim gerçekleştirilmiştir. Eğitim Quadro P4000 GPU ile Ubuntu işletim sistemi üzerinden Darknet (Bochkovskiy, 2022) arayüzü kullanılarak gerçekleştirilmiştir.

Çapraz doğrulama

Çapraz Doğrulama, model seçimi için kullanılan yöntemlerden biridir (Arlot & Celisse, 2010; Berrar, 2018; Y. Zhang & Yang, 2015). Genel olarak veriyi bölmeye dayalıdır. Normal şartlarda veri seti, eğitim, doğrulama ve test verisi olarak 3 kısma ayrılır. Bu ayrım genellikle verini %70-90'ı eğitim kısmında ve %10-30'u test kısmında kullanılmak üzere yapılır (Berrar, 2018). Bu gibi bir ayırmda model test verisini görmez sadece eğitim verisi ile eğitilir. Fakat bu durumda verinin hangi kısmının test verisi, hangi kısmının eğitim verisi seçileceği bir muammadır. Modelin, verinin hangi kısmında iyi performans elde edeceği bilinemez. Bu sebeple Çapraz Doğrulama yöntemi kullanılır ve modelin veri setinin tüm kısımları için performans sonucu elde edilir.

Çapraz doğrulama bir orana göre veriyi eşit parçalara bölerek uygulanır (Kerbaa vd., 2019). Bu çalışmada 4 Katlı Çapraz Doğrulama kullanılmıştır, yani veri 4 eşit parçaya bölünmüştür. Elde edilen veride veri çoğaltma yöntemleri kullanıldığı için çoğaltılan görüntülerde 4 eşit parçaya ayrılacak şekilde bir dağılım elde edilmiştir. Dağılım Çizelge 3.7'de gösterilmektedir.

Çizelge 3.7. Çapraz Doğrulama için resimlerin dağılımı

	Birinci Parça		İkinci Parça		Üçüncü Parça		Dördüncü Parça	
	Başlangıç Sırası	Bitiş Sırası	Başlangıç Sırası	Bitiş Sırası	Başlangıç Sırası	Bitiş Sırası	Başlangıç Sırası	Bitiş Sırası
Görüntü Sırası	0	499	500	999	1000	1499	1500	1999
	2000	2499	2500	2999	3000	3499	3500	3999
	4000	4499	4500	4999	5000	5499	5500	5999
	6000	6499	6500	6999	7000	7499	7500	7999
	8000	8499	8500	8999	9000	9499	9500	9999
	10000	10499	10500	10999	11000	11499	11500	11999
	12000	12499	12500	12999	13000	13499	13500	13999
	14000	14499	14500	14999	15000	15499	15500	15999
	16000	16499	16500	16999	17000	17499	17500	17999
	18000	18499	18500	18999	19000	19499	19500	19999
	20000	20499	20500	20999	21000	21499	21500	21999
	22000	22499	22500	22999	23000	23499	23500	23999
	24000	24499	24500	24999	25000	25499	25500	25999
	26000	26499	26500	26999	27000	27499	27500	27999
	28000	28499	28500	28999	29000	29499	29500	29999
	30000	30499	30500	30999	31000	31499	31500	31999
	32000	32499	32500	32999	33000	33499	33500	33999
	34000	34499	34500	34999	35000	35499	35500	35999
	36000	36499	36500	36999	37000	37499	37500	37999
	38000	38499	38500	38999	39000	39499	39500	39999
	40000	40499	40500	40999	41000	41499	41500	41999
	42000	42499	42500	42999	43000	43499	43500	43999
	44000	44499	44500	44999	45000	45499	45500	45999
	46000	46499	46500	46999	47000	47499	47500	47999

Çizelge 3.7’de görüldüğü gibi veri seti, veri çoğaltma yöntemleri göz önünde bulundurularak 4 eşit parçaya ayrılmıştır. Daha sonra bu kısımlar 4 kat çapraz doğrulama ile Çizelge 3.8’deki şekilde eğitim, doğrulama ve test veri setleri olarak sıralanmıştır.

Çizelge 3.8. Verinin kısımlarının eğitim, doğrulama ve test verisi olarak ayrılması

Kat Numarası	Eğitim Verisi	Doğrulama Verisi	Test Verisi
1. Kat	Birinci ve İkinci Parça	Üçüncü Parça	Dördüncü Parça
2. Kat	İkinci ve Üçüncü Parça	Dördüncü Parça	Birinci Parça
3. Kat	Üçüncü ve Dördüncü Parça	Birinci Parça	İkinci Parça
4. Kat	Birinci ve Dördüncü Parça	İkinci Parça	Üçüncü Parça

Çizelge 3.8’de görüldüğü gibi katlar arasında bağlantı olmayıp, hepsinin eğitim, doğrulama ve test verileri birbirinden farklıdır. Bu sayede bu 4 Kat veri ile baştan teker teker eğitim gerçekleştirilerek modellerin tüm veri seti üzerindeki performansları elde edilmiştir. Modellerin performans değerlendirmeleri 5 numaralı kısımda her kat için ayrı ayrı ve son olarak 4 katın ortalaması şeklinde verilecektir.

Transfer öğrenme

Nesne tespit çalışmalarında, performans kadar zamanda önemli bir etkidir. Sıfırdan bir model üretmek zaman alıcı ve oldukça zor bir süreçtir. Bunun yerine başarıları kanıtlanmış modelleri kullanmak ve bu modeller üzerinde değişiklikler yapmak daha efektiftir. Transfer Öğrenmesi, önceden eğitilmiş modelleri kullanması anlamına gelmektedir. Bu sayede farklı alanda veya farklı nesnelerle eğitilen modelleri kendi çalışmalarımıza entegre etmek mümkün olmaktadır. Bunun yanı sıra kızılotesi görüntü veri setlerinin, renkli görüntü veri setleri kadar bol olmaması ve bu durumun modelin eğitimini etkileyecek olması araştırmacıları Transfer Öğrenme yöntemini kullanmaya itmektedir (Du, Zhang, Zhang, Xiang, vd., 2021).

Bu çalışmada YOLOv3, YOLOv4, YOLOv7 ve Termal YOLO modellerinin önceden eğitilmiş verileri kullanılmıştır. Bu modellerden sadece Termal YOLO daha önce kızılotesi görüntülerle eğitilmiş olup, diğer modeller için Transfer Öğrenmenin bir alt dalı olan Alan Adaptasyonu sağlanmaya çalışılmıştır.

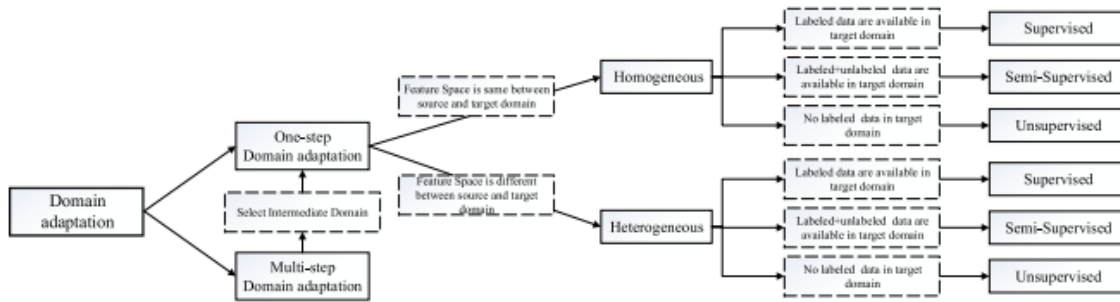
Çalışma kapsamında YOLOv3, YOLOv4 ve YOLOv7 modellerinin önceden eğitilmiş ağırlık dosyaları Github’taki Darknet sayfasından (Bochkovski, 2022), Termal Yolo’nun önceden eğitilmiş ağırlık dosyası ise Kaggle’daki AAU-PD-T isimli veri setinden (AAU-

PD-T | Kaggle, 2020) alınmıştır. Modellerin orijinal dosyalarında bulunan temel hiper parametreler aynı kalmak suretiyle, diğerleri çalışmaya uygun olarak değiştirilmiştir. Temel hiper parametrelerin değiştirilmeme sebebi bu modellerin Kızılötesi Alan Adaptasyon performanslarını değerlendirmek içindir.

Alan adaptasyonu

Alan (domain) adaptasyonu (AA), Transfer Öğrenmenin alt alanlarından biridir (Patel vd., 2015). Bilindiği üzere derin öğrenme çalışmaları için veri üretme ve bu verileri etiketleme oldukça zahmetli ve zaman alıcı işlemlerdir. Bu iş yükünden kurtulmak için Transfer Öğrenmesi yöntemi sunulmuştur. Aynı zamanda Transfer Öğrenmesinin bir alt konusu olan Alan adaptasyonu sayesinde az sayıda veri ile derin öğrenme çalışmaları gerçekleştirilebilmektedir (M. Wang & Deng, 2018).

Alan adaptasyonu, bir derin öğrenme modelinin Transfer Öğrenme yöntemi kullanılarak farklı bir alan üzerinde kullanılması demektir. Bu alan farkı kaynak ve hedef alan arasındaki resmin boyutu, kamera açısı, nesne farklı olabileceği gibi görüntünün çekildiği sensörün farklı olmasından da kaynaklanmaktadır. Bu durum 2 şekilde ifade edilir. Kaynak alanı ve hedef alanı birbirine benzerse bu homojen alan adaptasyonu olarak isimlendirilirken, kaynak alanı ve hedef alanı birbirinden farklıysa bu heterojen alan adaptasyonu olarak isimlendirilir (M. Wang & Deng, 2018). Bunların her biri kendi içinde gözetimli, gözetimsiz ve yarı gözetimli olarak ayrılır. Şekil 3.21’de alan adaptasyonu sınıfları verilmiştir.



Şekil 3.21. Alan adaptasyonu sınıflandırılması (M. Wang & Deng, 2018)

Özellik uzayı X ve olasılık dağılımı $P(X)$ olmak üzere, bir alan $D = \{X, P(X)\}$ şeklinde ifade edilir. Buradan yola çıkarak, kaynak alanı $D_k = \{X_k, P(X)_k\}$ ve hedef alanı $D_h = \{X_h, P(X)_h\}$ olmaktadır. Bir derin öğrenme uygulamasında, D_k ve D_h eğitim verilerini ifade etmektedir. Test verileri ise, T olarak gösterilir ve özellik uzayı Y ile koşullu olasılık dağılımı $P(Y|X)$ 'den oluşur. Kaynak alanı için yapılacak görev $T_k = \{Y_k, P(Y_k|X_k)\}$ ve hedef alanı için $T_h = \{Y_h, P(Y_h|X_h)\}$ şeklinde elde edilir. Burada $P(Y_k|X_k)$ kaynak etiketli verisinden ve $P(Y_h|X_h)$ hedef etiketli verisinden öğrenilen çıktıdır. Homojen AA'da özellik uzayları ve alanları aynıdır ($X_k = X_h$ ve $D_k = D_h$). Heterojen AA'da ise, kaynak ve hedef özellik uzayı ve alanları aynı değildir ($X_k \neq X_h$) ve $D_k \neq D_h$) (M. Wang & Deng, 2018). Heterojen alan farklılığı, kaynak alan ve hedef alan arasındaki farktan (görünür alan veya kızılotesi alan gibi) veya yapılacak görevin farklı olmasından (kaynakta alanda olmayan bir sınıfın hedef alanda bulunması gibi) meydana gelir.

Bu çalışmada görünür alanda eğitilmiş Yolo tabanlı modellerin kızılotesi alana adaptasyonları sağlanmıştır. Bu şekilde yapılan adaptasyona tek adımlı alan adaptasyonu denilmektedir (M. Wang & Deng, 2018). Yolo modellerinde, toplu normalleştirme ve nöron bırakma gibi alan adaptasyonunda faydalı katmanlar bulunmaktadır. Ayrıca öncede eğitilmiş Yolo modelleri insan ve araç kategorilerine aşinadır (Termal Yolo hariç). Bu iki özellik, bu çalışmada Yolo modellerinin seçilmesinde etkili olmuştur.

Eğitim işlemleri

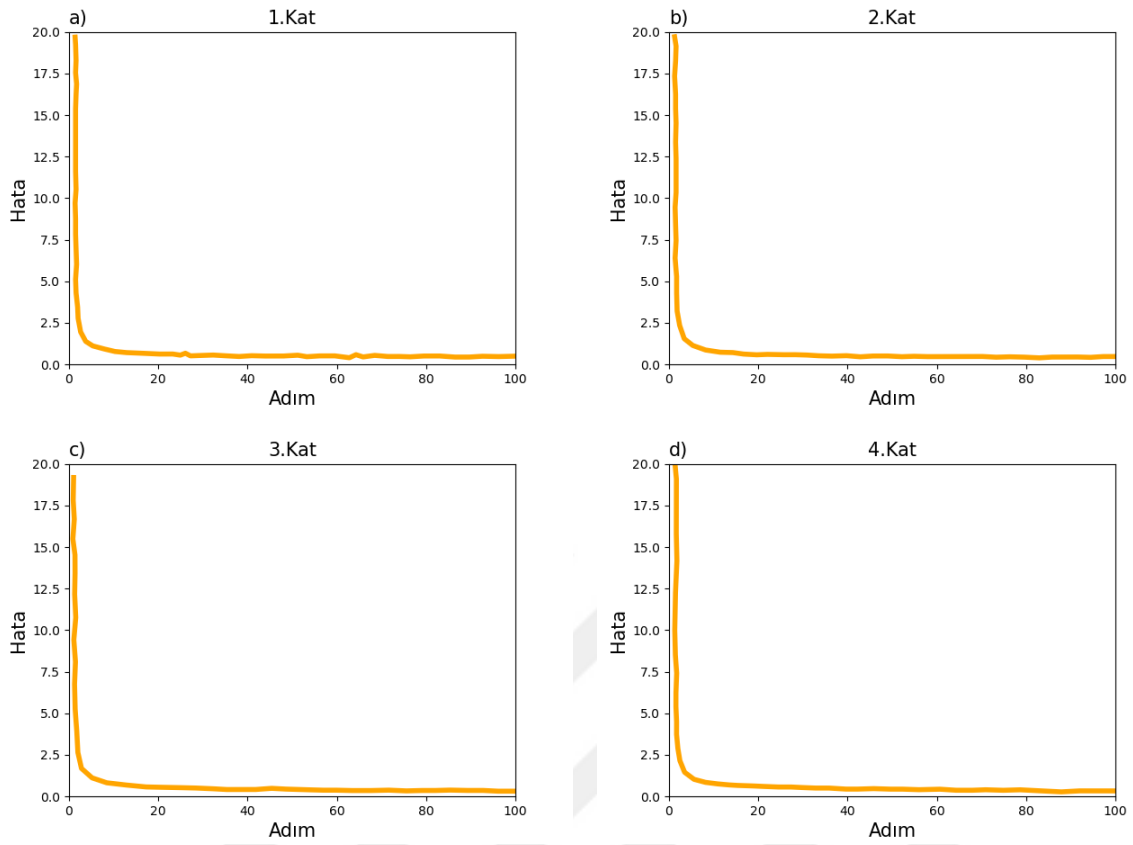
Yolov3, Yolov4, Yolov7 ve Termal Yolo modelleri 4 kat çapraz doğrulama ile ayrılmış veri seti ile eğitilmiştir. Modellerin önceden eğitilmiş ağırlık dosyaları Transfer Öğrenmesi yöntemi ile kullanılmıştır. Bu sayede sıfırdan eğitim gerçekleştirmek yerine, modellerin var olan nesne tanıma yetenekleri kızılotesi alana adapte edilmiştir. Modellere ait yapılandırma dosyalarında bulunan temel hiper parametreler aynı tutulmuş bazı değişiklikler yapılmıştır. Bu hiper parametreler Çizelge 3.9'da verilmiştir.

Çizelge 3.9. Modellere ait bazı hiper parametreler

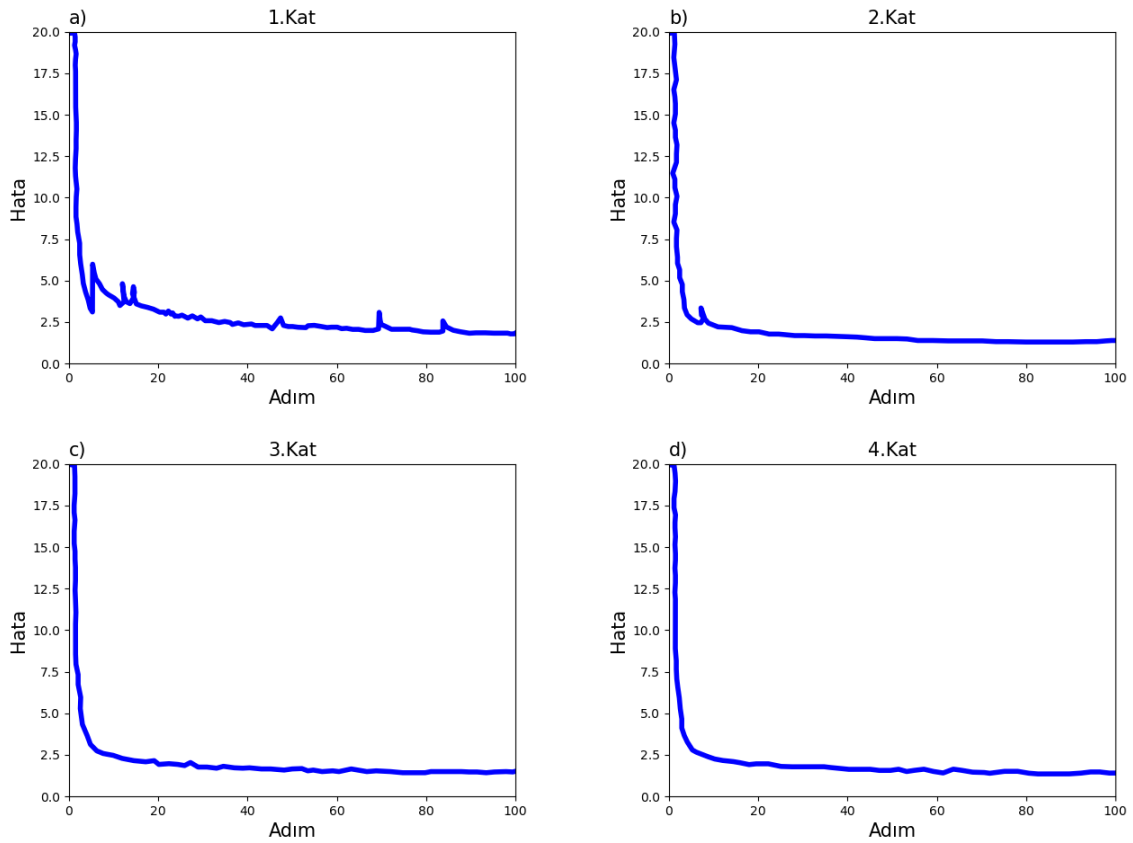
	Yolov3	Yolov4	Yolov7	Termal Yolo
Giriş Görüntü Boyutu	608x608	608x608	608x608	608x608
Momentum	0,9	0,949	0,9	0,9
Sönüm	0,0005	0,0005	0,0005	0,0005
Öğrenme Oranı	0,001	0,001	0,001	0,001
Aktivasyon Fonksiyonu	leaky	mish ve leaky	swish	leaky
Çapa Boyutu	10, 13, 16, 30, 33, 23, 30, 61, 62, 45, 59, 119, 116, 90, 156, 198, 373, 326	12, 16, 19, 36, 40, 28, 36, 75, 76, 55, 72, 146, 142, 110, 192, 243, 459, 401	12, 16, 19, 36, 40, 28, 36, 75, 76, 55, 72, 146, 142, 110, 192, 243, 459, 401	10, 13, 16, 30, 33, 23, 30, 61, 62, 45, 59, 119, 116, 90, 156, 198, 373, 326

Çizelge 3.9’da görüldüğü üzere, modellerin girişi 608x608 boyutunda seçilmiştir. Bunun sebebi giriş boyutu büyüdükçe küçük nesne tespit performansının artmasıdır. Eğitim veri setinin her katı için ayrı ayrı ve bir adım 375 iterasyon olacak şekilde toplamda 100 adımda yani 37500 adet iterasyonda gerçekleştirilmiştir.

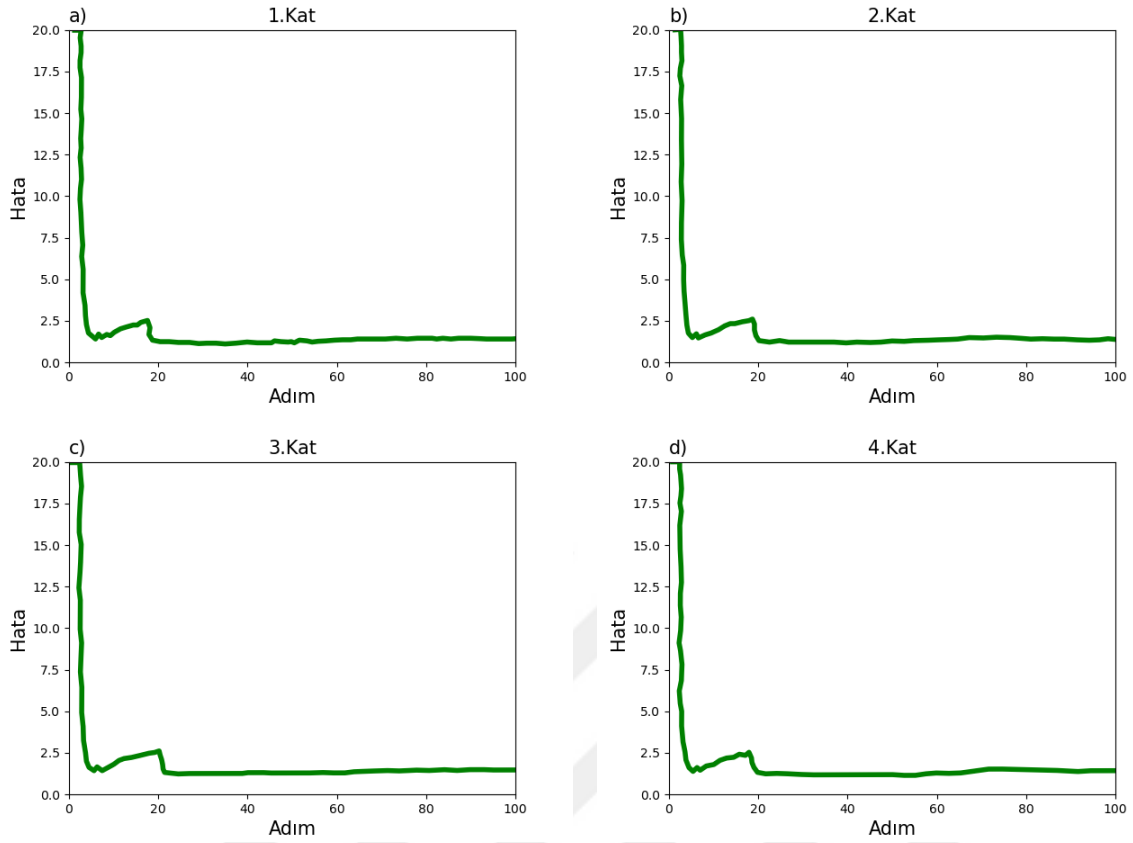
Veri setini her katında eğitim için 24000, doğrulama için 12000 ve test için 12000 kızılotesi görüntü bulunmaktadır. Bu sayede modeller eğitim veri seti ile eğitilmiş ve doğrulama veri seti ile eğitim performansları değerlendirilmiştir. Eğitim aşamasında Yolov3, Yolov4 ve Yolov7 modellerinin hata grafikleri ve doğrulama mAP değerleri sırasıyla Şekil 3.22, 3.23 ve 3.24’te verilmiştir.



Şekil 3.22. Yolov8 hata grafikleri, a) 1.kat için, b) 2.kat için, c) 3.kat için ve d) 4.kat için

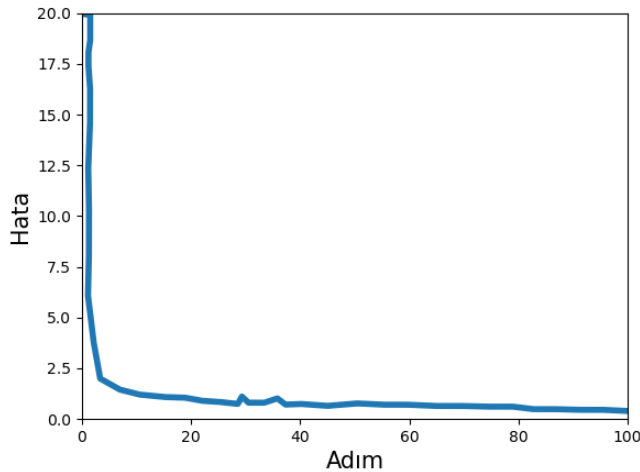


Şekil 3.23. Yolov4 hata grafikleri, a) 1.kat için, b) 2.kat için, c) 3.kat için ve d) 4.kat için



Şekil 3.24. Yolov7 hata grafikleri, a) 1.kat için, b) 2.kat için, c) 3.kat için ve d) 4.kat için

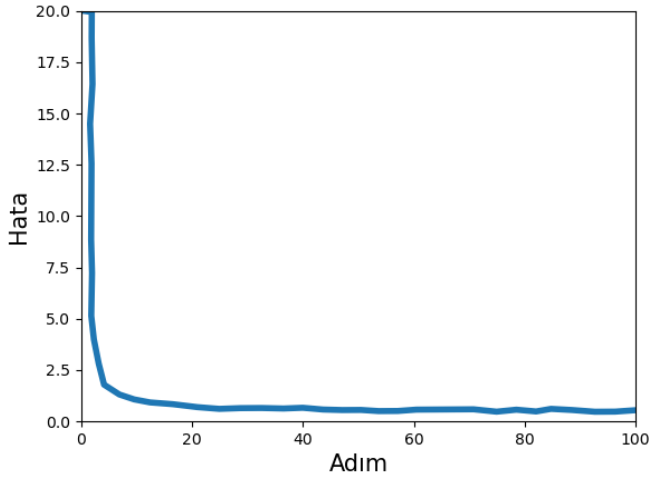
Termal Yolo'nun ise, veri setinin 1.katı ile eğitimi sonucu iyi bir performans sergileyemediği gözlenmiştir. Şekil 3.25'te Termal Yolo için eğitim hatası grafiği bulunmaktadır.



Şekil 3.25. 1.kat için Termal Yolo hata grafiği (giriş boyutu 608x608 için)

Grafikte hata değeri her ne kadar düşük gibi görünmesine rağmen, doğrulama performansına ve test performansına bakıldığında modelin iyi bir nesne tespit uygulaması için yeterli olmadığı görülüp eğitim aşaması burada kesilmiş yani diğer katlar için eğitim

gerçekleştirilmemiştir. Doğrulama ve test mAP değerleri sırasıyla %4 ve %2,84 şeklindedir. Bu da modelin aşırı öğrenmiş (overfit) olabileceğini göstermektedir. Bu sonuca dayanarak giriş görüntü boyutu orijinal Termal Yolo makalesindeki gibi 416x416 olarak seçilip eğitim tekrarlanmıştır. Fakat bu eğitim süreci sonucunda da yeteri kadar iyi bir nesne tespit modeli elde edilememiştir. 416x416 giriş görüntü boyutuna sahip modelin eğitim hata grafiği Şekil 3.26’da verilmiştir.



Şekil 3.26. 1.kat için Termal Yolo hata grafiği (giriş boyutu 416x416 için)

Aynı şekilde doğrulama ve test mAP değerlerine bakıldığında sırasıyla %54,96 ve %32,91 olarak elde edilmiştir. Sonuç olarak giriş görüntü boyutu küçüldüğünde performans artmış olsa da Yolov3 modelinden elde edilen Termal Yolo modelinin, seçilen veri seti için iyi bir performans sergileyemediği görülmüştür. Bunun sebebi Termal Yolo modelinin sadece insan tespiti için kızılötesi görüntülerle eğitilmesi ve sonucunda orijinal Yolov3'ün sahip olduğu bazı öznelilikleri unutması olabilmektedir. Aynı zamanda Termal Yolo'nun, Yolov3'ten eğitilmiş olması ve bunun üzerine bizimde 100 adımlık bir eğitim gerçekleştirmemiz, aşırı eğitim sebebiyle performansı düşürmüş olabilmektedir.

Nesne tespit uygulamalarında performans değerlendirmesinde kullanılan bazı metrikler vardır. Bu metrikler 5 numaralı bölümde detaylı incelenmiştir. Çizelge 3.10'da modellerin eğitim performanslarını değerlendirebilmek için bazı değerlendirme metrikleri için aldıkları değerler ve tüm katlar için alınan ortalama değerleri verilmektedir.

Çizelge 3.10. Modellerin doğrulama veri seti için performansları

	Kat	AP _{insan}	AP _{araç}	mAP
Yolov3	1.Kat	%86,53	%84,63	%85,58
	2.Kat	%94,58	%84,17	%89,38
	3.Kat	%79,14	%77,25	%78,20
	4.Kat	%95,01	%87,13	%91,07
ortalama		%88,81	%83,29	%86,05
Yolov4	1.Kat	%87,42	%96,01	%91,72
	2.Kat	%97,39	%99,77	%98,58
	3.Kat	%91,15	%95,76	%93,46
	4.Kat	%97,40	%99,47	%98,43
ortalama		%93,34	%97,75	%95,54
Yolov7	1.Kat	%88,31	%25,48	%56,89
	2.Kat	%95,72	%25,11	%60,42
	3.Kat	%87,20	%25,94	%56,57
	4.Kat	%95,37	%29,28	%62,33
ortalama		%91,65	%26,45	%59,05

Burada mAP, PASCAL VOC (*The PASCAL Visual Object Classes Homepage*, 2005) ile aynı olarak şekilde KBO = 0,5 iken elde edilen değerdir ve doğrulama performansını temsil eder.

Elde edilen sonuçlar;

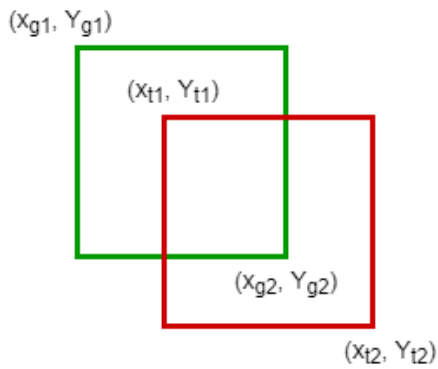
- Yolov3 modelinden elde edilen Termal Yolo, daha önce kızılötesi alanda eğitilmiş bir model olmasına rağmen, görünür alanda eğitilen diğer modellere karşı performansı oldukça düşüktür. Bunun sonucunda aynı model yapısına ve parametrelerine sahip Termal Yolo ve Yolov3 arasından, hem kızılötesi alan adaptasyonu için hem de nesne tespit çalışması için Yolov3'ün tercih edilmesi önerilmektedir.
- Doğrulama veri setinde her üç modelde insanlar için iyi sonuç verse de araçlar için Yolov7 başarılı olamamıştır.
- Yolov4 doğrulama veri setinde en iyi sonuç veren model olarak gözlenmiştir.

3.2.3. Test işlemleri

Nesne tespit performanslarını değerlendirmek için yaygın olarak kullanılan bazı metrikler vardır. Bunlar detaylı olarak sıralanmıştır.

Kesişimin birleşime oranı (KBO- IoU)

Nesne tespit uygulamalarında, nesnenin sınıfının doğru olarak tahmin edilmesi dışında nesnenin yerinin de doğru olarak tahmin edilmesi gerekir. Nesnenin tahmin edilen konumu ve gerçek konumu arasındaki yakınlığı tespit etmeye yarayan değerlendirme metriği KBO olarak isimlendirilir. Şekil 3.27’de örnek bir görüntü üzerinde tahmin edilen konum ve gerçek konumlar verilmiştir.



Şekil 3.27. Gerçek konum (yeşil) ve tahmini konumun (kırmızı) kesişme durumu

$$KBO = \frac{\text{Kesişimin Alanı}}{\text{Birleşimin Alanı}} \quad (3.11)$$



Resim 3.4. Örnek KBO değerleri

KBO, Eş. 3.11'deki gibi ifade edilir ve $[0,1]$ arasında bir değer alır. Değer 1'e ne kadar yakınsa nesnenin tahmini konumu gerçek konumuna o kadar yakın demektir. Resim 3.4'te farklı KBO değerlerine ait örnekler verilmiştir.

KBO değerinin her uygulama için illa 1'e yakın olması gerekmemektedir. Çok hassas uygulamalar için 0,9 üzeri bir sonuç yeterli olurken, hassasiyetin yüksek olması gerekmeyen uygulamalar için 0,5 üzeri bir sonuç yeterli olacaktır. Bu sebeple sonucu değerlendirmek için bir eşik değeri belirlenir ve bu eşik değerinin üzerinde KBO'ya sahip nesneler doğru tespit edilmiş kabul edilir. Örnek bir KBO hesaplaması Eş. 3.12, Eş. 3.13, Eş. 3.14, Eş. 3.15, Eş. 3.16 ve Eş. 3.17 ile açıklanmıştır.

Görüntünün boyutu, W, H olmak üzere, sınırlayıcı kutunun bilgilerini barındıran elemanlar normalize haldedir. Bu elemanlar görüntünün eni ve boyu ile çarpılarak normalize olmayan halleri elde edilir. Sınırlayıcı kutu: $\{c, x, y, w, h\}$ şeklinde ifade edilir. Bu sayede sınırlayıcı kutunun sol üst ve sağ alt noktaları (yani başlangıç ve bitiş noktaları) elde edilebilir.

$$\text{Başlangıç Noktası: } \left\{x - \frac{w}{2}, y - \frac{h}{2}\right\} \quad (3.12)$$

$$\text{Bitiş Noktası: } \left\{x + \frac{w}{2}, y + \frac{h}{2}\right\} \quad (3.13)$$

Bu noktalar gerçek sınırlayıcı kutu için (x_g, y_g) ve tahmini sınırlayıcı kutular için (x_t, y_t) şeklinde ele alınabilir. Buna göre gerçek sınırlayıcı kutunun alanı ve tahmini sınırlayıcı kutunun alanları hesaplanabilir.

$$\text{Alan}_{\text{gerçek}} = w_g * h_g \quad (3.14)$$

$$\text{Alan}_{\text{tahmini}} = w_t * h_t \quad (3.15)$$

Buradan hareketle iki alanın birleşimi elde edilir. Birleşim kutusu yine başlangıç ve bitiş noktaları ile ifade edilir. Birleşim kutusunun başlangıç noktalarının koordinatları (X_{b1}, Y_{b1}) ve bitiş noktalarının koordinatları (X_{b2}, Y_{b2}) olarak isimlendirilebilir.

$$(X_{b1}, Y_{b1}): \{\max(x_{g1}, x_{t1}), \max(y_{g1}, y_{t1})\} \quad (3.16)$$

$$(X_{b2}, Y_{b2}): \{\max(x_{g2}, x_{t2}), \max(y_{g2}, y_{t2})\} \quad (3.17)$$

KBO hesaplamasına ait pseudocode (sözde kod) Algoritma 4'te verilmiştir.

Algoritma 4: KBO hesaplaması

Giriş: gerçek_kutu = [Xg1,Yg1,Xg2,Yg2, wg, hg], tahmini_kutu = [Xt1,Yt1,Xt2,Yt2, wt, ht] - iki sınırlayıcı kutu için başlangıç, bitiş noktaları ve kutuların boyutları

Sonuç: KBO - gerçek_kutu ve tahmini_kutu için kesişimin birleşime oranı

$Xb1 \leftarrow \max(xg1, xt1),$
 $Yb1 \leftarrow \max(yg1, yt1)$
 $Xb2 \leftarrow \min(xg2, xt2)$
 $Yb2 \leftarrow \min(yg2, yt2)$
 if (Xb1 > Xb2) veya (Yb1 > Yb2) ise
 sonuç 0;
 son

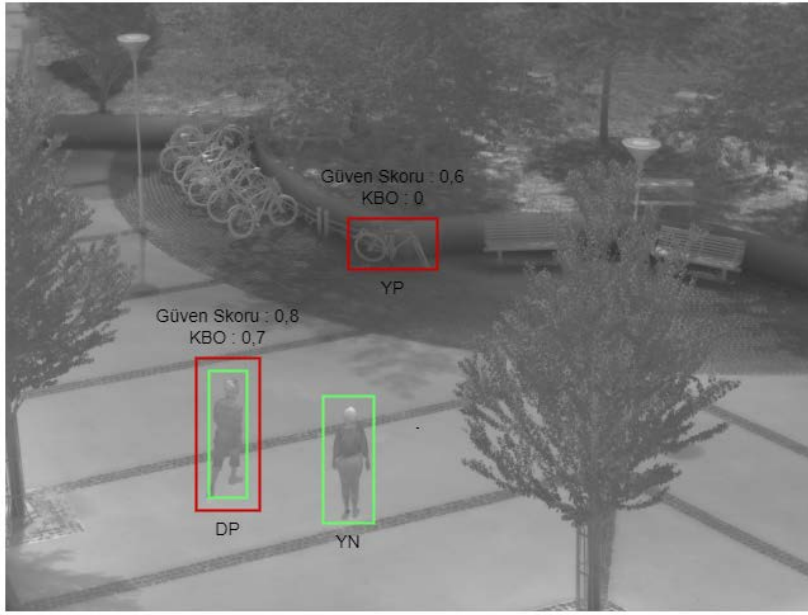
Alan_birleşim $\leftarrow |Xb1 - Xb2| * |Yb1 - Yb2|$
 Alan_gerçek $\leftarrow wg * hg$
 Alan_tahmini $\leftarrow wt * ht$
 Alan_toplam $\leftarrow \text{Alan_gerçek} + \text{Alan_tahmini} - \text{Alan_birleşim}$
 KBO $\leftarrow \text{Alan_birleşim} / \text{Alan_toplam}$
 sonuç KBO;

Doğru negatif, doğru pozitif, yanlış negatif ve yanlış pozitif

Nesne sınıflandırma veya tespit çalışmalarında doğruluğu ifade eden bazı terimler vardır. Bunlar DP (doğru pozitif), DN (doğru negatif), YN (yanlış negatif) ve YP (yanlış pozitif) olarak isimlendirilir.

- DP, tespit edilen nesneler arasından sınıfı doğru tahmin edilen nesnelerin sayısını ifade eder.
- DN, ilgili sınıfa ait nesnelerin görüntüde bulunmaması demektir. Dolayısıyla görüntüde nesne olmayan tüm kısımları temsil eder ve nesne tespit için kullanılan bir terim değildir.
- YP, görüntüde olmadığı halde tespit edilen yanlış tespitlerin sayısını verir.
- YN ise, nesnenin bulunup model tarafından tespit edilememesi durumudur.

Güven skoru ve KBO, tespit edilen nesnelerin DN, DP, YN veya YP olup olmadığını belirlemede en önemli etkenlerdir. Güven skoru ve KBO için ayrı ayrı belirlenen eşik değerlerine göre değerlendirme yapılır. İki eşik değerini de geçen tespitler DP olarak değerlendirilir. Sadece güven skoru eşik değerini geçen tespitler YP, sadece KBO eşik değerini geçenler ise YN olarak değerlendirilir. Resim 3.5'te DP, YN ve YP için örnekler güven skorları ve KBO değerleri ile birlikte verilmiştir. Burada güven skoru ve KBO için eşik değerleri sırasıyla 0,25 ve 0,5 olarak ele alınmıştır.



Resim 3.5. DP, YN ve YP için örnekler

Hassasiyet ve duyarlılık

Hassasiyet, tahmin edilen pozitifler arasında kaç tanesinin doğru olduğunu ifade eder. Duyarlılık ise, bulunan gerçek nesneleri ifade eder. Yani, iyi bir model için duyarlılığın yüksek olması beklenir. Fakat hassasiyetin yüksek olması beklenmez. Çünkü hassasiyet değerinde yanlış pozitiflerde yer almaktadır. Hassasiyet ve duyarlılık $[0,1]$ arasında değerler alırlar. Formülleri sırasıyla Eş. 3.18 ve Eş. 3.19’da verilmiştir.

$$\text{Duyarlılık} = \frac{DP}{DP + YN} \quad (3.18)$$

$$\text{Hassasiyet} = \frac{DP}{DP + YP} \quad (3.19)$$

Burada DP+YP toplam tahmini nesne sayısını ifade ederken, DP+YN toplam gerçek nesne sayısını ifade eder.

Hassasiyet ve duyarlılık eğrisi

Hassasiyet ve duyarlılık eğrisi, duyarlılığa göre hassasiyeti belirten bir eğridir ve nesne tespit çalışmalarında en çok kullanılan değerlendirme metriklerinden biridir. Burada

hassasiyet ve duyarlılık değerleri güven skoru eşik değerine göre çizdirilir. Eşik değeri $[0,1]$ arasında değiştirilerek grafik elde edilir. Eşik değeri 0 iken, hassasiyet en düşük, duyarlılık ise en yüksek değerindedir. Eşik değeri artırıldıkça hassasiyet değeri artarken, duyarlılık düşer.

Ortalama hassasiyet (AP)

Ortalama hassasiyet, duyarlılık 0 ve 1 arasındayken bütün hassasiyet değerlerinin ortalaması demektir, Eş. 3.20'deki gibi ifade edilir.

$$AP = \frac{1}{N} \sum_{r \in \{0, \dots, 1\}} p(r) \quad (3.20)$$

Burada N duyarlılık değeridir, p ise o duyarlılık değerine ait hassasiyet değerini ifade eder.

Bir modelde AP her sınıf için hesaplanabilir. Bu sınıflara ait AP değerlerinin ortalaması modelin performansını ifade eder ve mAP olarak isimlendirilir (Y. Liu, Su, vd., 2021).

4. BULGULAR VE TARTIŞMA

Tez çalışması kapsamında elde edilen sonuçlar, nesne tespit performanslarını değerlendirmek için yaygın olarak kullanılan metrikler çerçevesinde sunulmuştur. Bu metrikler 3.bölümde detaylı olarak ele alınmıştır. Bu bölümde ise, sonuçlar doğru pozitif, yanlış negatif ve yanlış pozitif sayıları, hassasiyet ve duyarlılık eğrileri ve ortalama hassasiyet değerleri aracılığıyla değerlendirilmiştir.

4.1. Doğru Pozitif (DP), Yanlış Negatif (YN) ve Yanlış Pozitif (YP)

Modellere ait doğru pozitif (DP), yanlış negatif (YN) ve yanlış pozitif (YP) değerleri, modellerin doğrudan performans analizinin yapılmasında kullanılsa da modeller hakkında bilgi elde etmemizi sağlayan değerlerdir. Bu değerlere ait detaylı açıklama 3. Bölümde yapılmıştır. Çizelge 4.1’de KBO eşik değeri 0,5 ve güven skoru eşik değeri 0,25 iken elde edilen DP, YN ve YP değerleri verilmiştir.

Çizelge 4.1. Modeller için DP, YN ve YP değerleri

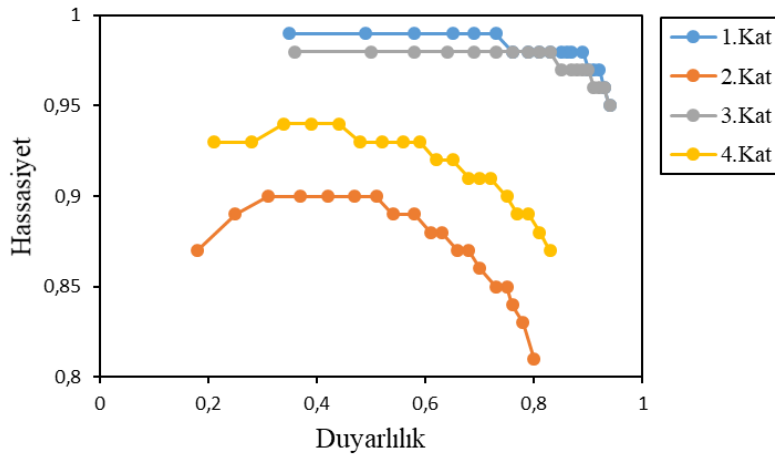
		DP	YP	YN
Yolov3	1.Kat	26335	720	3002
	2.Kat	26837	4623	10111
	3.Kat	26502	891	2883
	4.Kat	27610	3040	9371
ortalama		26821	2318	6341
Yolov4	1.Kat	28644	15795	693
	2.Kat	35945	34283	1003
	3.Kat	28954	17249	431
	4.Kat	36202	34294	779
ortalama		32436	25405	726
Yolov7	1.Kat	26922	12891	2415
	2.Kat	27536	21014	9412
	3.Kat	29232	15727	7716
	4.Kat	28959	15842	8022
ortalama		28162	16368	6891

Buna göre, tespit edilen nesneler arasında en doğru tahminler Yolov4 tarafından yapılmıştır. Fakat görüntüde olmadığı halde tespit edilen yanlış tahminlerin sayısı en fazla Yolov4’te bulunmaktadır. En az yanlış tahmin sayısı ise Yolov3’te bulunmaktadır.

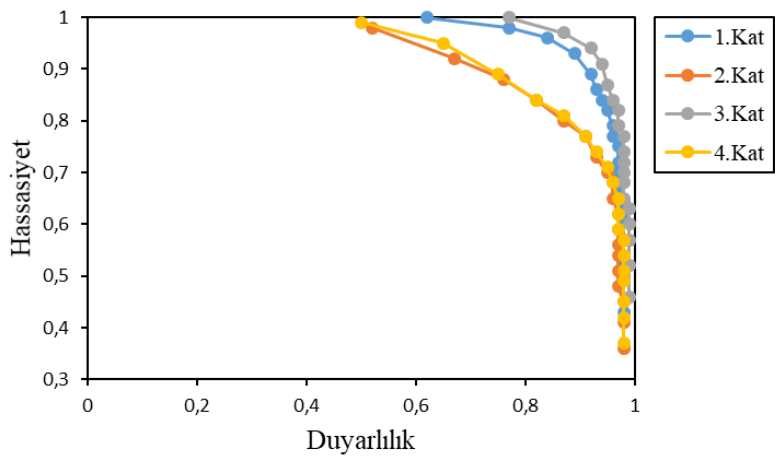
Görüntüde olup tespit edilemeyen nesne sayısı en yüksek Yolov7’de, en düşük Yolov4’te bulunmaktadır.

4.2. Hassasiyet ve Duyarlılık Eğrisi

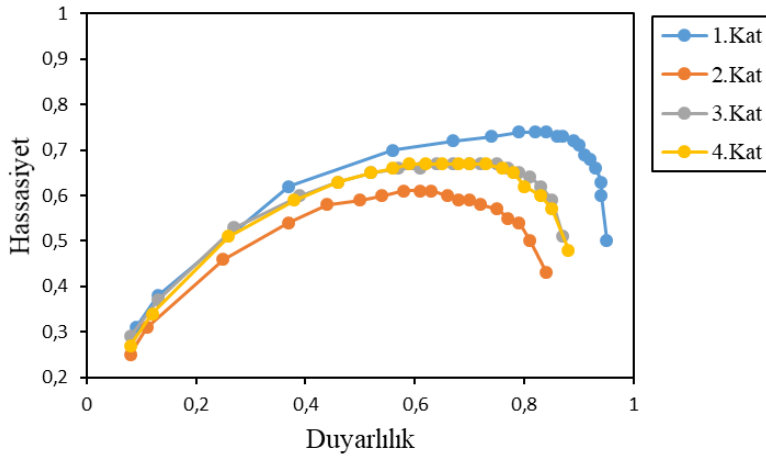
Hassasiyet ve duyarlılık eğrisi, en yaygın kullanılan değerlendirme metrikleri arasında yer alır. Modeller arasında performansı en iyi olan modeli analiz etmemize yarar. Detaylı açıklama 3. Bölümde bulunmaktadır. Şekil 4.1, 4.2 ve 4.3’te modellerin 4 kat için ayrı ayrı hassasiyet ve duyarlılık eğrileri verilmiştir. Burada güven skoru 0,5 ve 0,95 arasında 0,5 adım boyutuyla artırılarak hassasiyet ve duyarlılık değerleri elde edilmiştir.



Şekil 4.1. Yolov3 için hassasiyet ve duyarlılık eğrisi

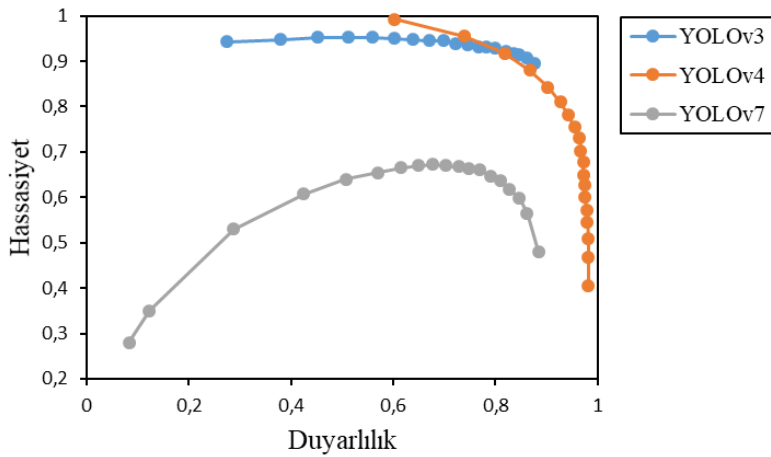


Şekil 4.2. Yolov4 için hassasiyet ve duyarlılık eğrisi



Şekil 4.3. Yolov7 için hassasiyet ve duyarlılık eğrisi

Buna göre model veri setinin 1. ve 3. Katlarında en iyi performansı sergilemişlerdir. 4 katın ortalaması ile modellerin beraber performansları Şekil 4.4'te verilmiştir.



Şekil 4.4. Modeller için hassasiyet ve duyarlılık eğrisi

Hassasiyet ve duyarlılık eğrisinde en ideal model en tepe noktasında bulunmaktadır. Şekil 5.5'te görüldüğü üzere burada ideal modele en yakın olan model Yolov4 olarak gözlenmiştir.

4.3. Ortalama Hassasiyet (AP)

Ortalama hassasiyet, nesne tespit çalışmasının performansını ifade etmede kullanılan en yaygın yöntemdir. Her sınıf için ayrı ayrı elde edilen AP değerlerinin ortalaması alınarak

hesaplanan genel ortalama hassasiyet (mAP) değeri modelin nihai performansını ifade eder. Burada AP değerleri COCO veri kümesi nesne algılama değerlendirme metriklerine göre $KBO=0,5$ ve $KBO=0,75$ iken elde edilmiştir. $KBO=0,5$ için elde edilen AP, AP_{50} ve $KBO=0,75$ için elde edilen AP, AP_{75} olarak isimlendirilmektedir. AP ve mAP değerleri detaylı olarak 3. Bölümde sunulmuştur. Bütün modeller için AP ve mAP değerleri Çizelge 4.2’de verilmiştir.

Çizelge 4.2. Yolov3, Yolov4 ve Yolov7 için nesne tespit sonuçları (En iyi ortalama değer koyu renkle ve en iyi ikinci ortalama değer altı çizili olarak işaretlenmiştir)

Model		Ortalama Hassasiyet – AP (%), KBO:					
		İnsan		Araç		mAP (%)	
		0,5	0,75	0,5	0,75	0,5	0,75
Yolov3	1.Kat	95,53	66,38	89,82	54,97	92,67	60,68
	2.Kat	80,78	41,93	64,58	28,48	72,68	35,21
	3.Kat	95,33	61,91	89,45	54,78	92,39	58,34
	4.Kat	82,68	43,98	81,27	39,29	81,97	41,63
	ortalama	88,58	53,55	<u>81,28</u>	<u>44,38</u>	<u>84,93</u>	<u>48,96</u>
Yolov4	1.Kat	95,83	80,89	99,48	92,80	97,66	86,85
	2.Kat	90,97	68,45	96,47	81,56	93,72	75,00
	3.Kat	97,30	86,48	99,55	94,95	98,43	90,72
	4.Kat	91,72	71,18	96,45	83,21	94,08	77,20
	ortalama	93,95	76,75	97,98	88,13	95,97	82,44
Yolov7	1.Kat	95,81	76,44	29,15	8,79	62,48	42,61
	2.Kat	86,31	50,87	20,66	5,25	53,49	28,06
	3.Kat	87,20	52,05	25,94	5,67	56,57	28,86
	4.Kat	87,83	51,08	25,46	5,93	56,64	28,50
	ortalama	<u>89,28</u>	<u>57,61</u>	25,30	6,41	57,29	32,00

Çizelge 4.2’ye göre Yolov4’ün açık ara en iyi performansı sergilediği görülmüştür. Yolov7, insan tespitinde en iyi ikinci model olmuştur. Fakat araç tespitinde yeterli sonuç veremeyip mAP metriğinde Yolov3’ün gerisinde kalmıştır. Yolov3 araç tespitinde ve mAP metriğine göre ikinci en iyi model olmaktadır. Modellerin örnek görüntüler üzerinden tespit ettikleri nesneler EK-2’te verilmiştir.

EK-2’te verilen görsellere göre;

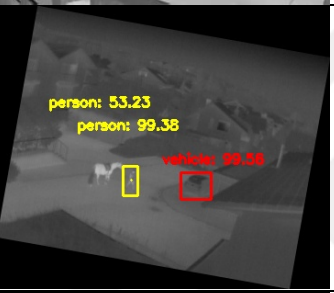
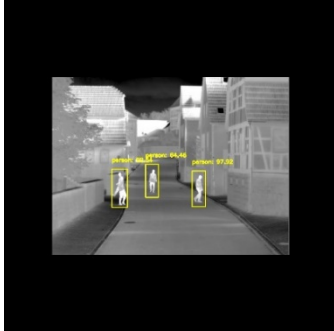
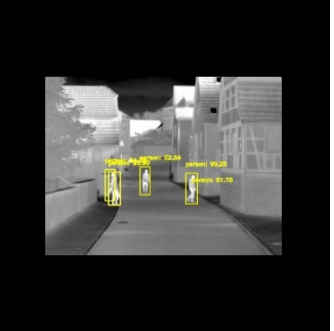
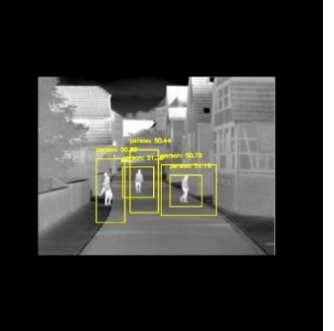
- Nesne tespit yeteneği en yüksek modeller Yolov3 ve Yolov4 olarak gözlenmiştir. Bu modeller farklı pozisyonlarda ve şekillerde nesneleri tespit etmede Yolov7 modeline göre daha iyi performans sergilemiştir.

- Arka plan ile benzer renk yoğunluğa sahip nesneleri ayırt etmede Yolov4 modeli diğer modellere göre daha iyi sonuç vermiştir.
- Üç model içinde nesneleri yanlış sınıflandırma olasılığının düşük olduğu gözlenmiştir. Arka planda bulunan diğer nesneler insan ya da araç olarak sınıflandırılmamıştır. Bu bağlamda üç modelde doğru nesneleri tespit etme konusunda başarılı bulunmuşlardır.
- Görüntülerden ve AP değerlerinden yola çıkarak en iyi nesne tespit yeteneğine sahip model Yolov4 olarak seçilmiştir. Fakat Yolov3 modeli de Yolov4 kadar iyi performans sergilemektedir. Hatta bazı durumlarda nesneleri daha iyi pencere içine alıp gerçek tahmini kutucukları tam yakalamaktadır. Dolayısıyla KBO değeri daha yüksektir. Bunun yanı sıra bazı durumlarda Yolov3 tarafından elde edilen güven skorunun, Yolov4 tarafından elde edilen güven skorundan daha yüksek olduğu ve Yolov4'ün tespit edemediği bazı nesneleri tespit edebildiği görülmektedir.
- Yolov7 ise, nesnelerin konumlarını bulmada başarılı olsa da nesnelerin sayısını tam tespit edememekte olduğu görülmüştür. KBO ve güven skoru değerlerinin düşük olması nedeniyle hassasiyet, duyarlılık, AP değerleri de düşük çıkmaktadır.

Küçük nesne tespiti için modellerin performanslarının değerlendirilmesi adına 32x32 boyutundan daha küçük nesnelere sahip görseller ve modellerin tespitleri Çizelge 4.3'te verilmiştir.

Çizelge 4.3'te görüldüğü üzere, Yolov3 ve Yolov4 küçük nesneleri tespit etme konusunda Yolov7'ye göre daha iyi performans sergilemişlerdir. Yolov4 daha doğru tespitler gerçekleştirmiş olsada bazı durumlarda nesne sayısını ayırt etmede zorlanmıştır. Genel olarak bakıldığında modeller, hiçbir arka plan nesnesini araç veya insanlarla karıştırmayarak sınıflara ait öznitelikleri gayet iyi öğrenebildiklerini göstermişlerdir. Bu bağlamda iyi bir eğitim aşaması geçirmiş oldukları görülmektedir.

Çizelge 4.3. Modellerin küçük nesneler üzerinden tespitleri

Yolov3	Yolov4	Yolov7
		
		
		
		

Modeller Pascal Voc Yarışması'na kullanıldığı gibi KBO=0.5 iken hesaplanan AP-mAP değerine göre literatürdeki diğer çalışmalarda kıyaslandığında sonuçlar Çizelge 4.4, 4.5 ve 4.6'da verilmiştir.

Çizelge 4.4. Elde edilen Yolov3 modelinin literatürdeki çalışmalarla karşılaştırılması (En yüksek değer koyu renkle ve ikinci yüksek değer altı çizili olarak işaretlenmiştir)

Model – Çalışma	mAP (%)
Yolov3 (Tumas vd., 2021)	84,91
Yolov3 (Wu vd., 2021)	67,9
Yolov3 (Y. Liu, Su, vd., 2021)	72,2
Yolov3-spp (C. Jiang vd., 2022)	82,2
Yolov3 (Kristo vd., 2020)	97,93
Yolov3 (bizimki) ⁴	<u>84,93</u>

⁴ Oluşturduğumuz veri seti ile eğitilen Yolov3 modelini ifade etmektedir.

Çizelge 4.5. Elde edilen Yolov4 modelinin literatürdeki çalışmalarla karşılaştırılması (En yüksek değer koyu renkle ve ikinci yüksek değer altı çizili olarak işaretlenmiştir)

Model – Çalışma	AP (%)	mAP (%)
Yolov4 (Tumas vd., 2021)	-	<u>87,20</u>
Yolov4 (Du, Zhang, Zhang, & Xu, 2021)	91,92	-
Yolov4-pacsp-mish (C. Jiang vd., 2022)	-	62,2
M-YOLO (Hou vd., 2022)	83,4	-
FA_YOLO (Du, Zhang, Zhang, Xiang, vd., 2021)	<u>92,95</u>	-
Yolov4 (bizimki) ⁵	95,97	95,97

Çizelge 4.6. Elde edilen Yolov7 modelinin literatürdeki çalışmalarla karşılaştırılması (En yüksek değer koyu renkle ve ikinci yüksek değer altı çizili olarak işaretlenmiştir)

Model – Çalışma	AP (%)	mAP (%)
CEAM-YOLOV7 (S. Liu vd., 2022)	-	73,60
Yolov7 (bizimki) ⁶	57,29	57,29

Çizelge 4.4’te görüldüğü üzere Yolov3 modeli ile literatürde kızılötesi görüntülerden insan ve araç tespiti üzerine yapılan çalışmalar arasında bizim modelimiz en iyi ikinci mAP değerine sahiptir. Yolov4 modelimiz en yüksek mAP değerine sahip model olmaktadır. Yolov7 modelimiz ise literatürde pek fazla örnek çalışma olmaması sebebiyle değerlendirme açısından eksik kalmış olsa da iyi bir performans sergileyemediği gözlenmektedir. Literatürde Yolov7 modelinin, kızılötesi görüntüler ile renkli görüntüler bir arada kullanıldığında daha iyi sonuç verdiği çalışmalarda sunulmuştur (Gallagher & Oughton, 2022).

⁵ Oluşturduğumuz veri seti ile eğitilen Yolov4 modelini ifade etmektedir.

⁶ Oluşturduğumuz veri seti ile eğitilen Yolov7 modelini ifade etmektedir.

5. SONUÇ VE ÖNERİLER

Gerçekleştirilmiş olan tez çalışmasında kızılötesi nesne tespiti için etkili bir derin öğrenme modeli elde etmek adına renkli görüntüler üzerinde popüler olarak kullanılan Yolov3, Yolov4 ve Yolov7 derin öğrenme modelleri üzerinde eğitimler gerçekleştirilmiştir. Bu eğitimler için transfer öğrenme yöntemi ile modellerin önceden eğitilmiş ağırlıkları kullanılmıştır. Kızılötesi veri seti olarak internette ulaşılabilir olan üç farklı veri setinin harmanlanmış, bu sayede modellerin insan ve araçlar için çeşitli öznelilikleri görmesi amaçlanmıştır. Veri setlerinden elde edilen 1000 adet kızılötesi görüntü, ayna görüntüsü alma, kenar doldurma, döndürme, kırpma ve gama düzeltmesi gibi veri çoğaltma yöntemleri kullanılarak 48000 adet kızılötesi görüntüye çoğaltılmıştır. Veri setinin tamamı için modellerin performanslarını değerlendirmek adına, popüler model seçme yöntemlerinden biri olan çapraz doğrulama yöntemi 4 katlı olacak şekilde kullanılmıştır. Bu bağlamda veri seti 4 eşit parçaya ayrılmış ve bu parçalar her kat için değişecek şekilde eğitim, doğrulama ve test veri seti olarak ayarlanmıştır. Veri setinin %50'si eğitim, %25'i doğrulama ve %25'i test verisi olacak şekilde her kat için ayrılmıştır.

Eğitim sürecinde en uygun eğitim adımı 100 olarak bulunmuş ve modeller 100 adımda eğitilmişlerdir. Öğrenme oranı 0,001 olarak seçilmiş ve adım sayısının %80 ve %90'ında sonra azaltılmıştır. Modeller veri setinin çapraz doğrulama ile ayrılmış her katında 24000 görüntü ile eğitilmiş ve 12000 görüntü ile doğrulama performanslarına bakılmıştır. Yolov3 modelinin doğrulama performansı her kat için sırasıyla %85,58; %89,38; %78,20; ve %91,07 mAP değerlerine sahiptir. Yolov4 için bu değerler sırasıyla %91,72; %98,58; %93,46 ve %98,43 olarak elde edilmiştir. Yolov7 için ise %56,89; %60,42; %56,57 ve %62,33'tür. Eğitim gerçekleştirildikten sonra elde edilen modeller her kat için 12000 görüntüye sahip test veri setleri ile test edilmiştir. Bu testlerin sonucunda 4 kat için elde edilen mAP değerlerinin ortalaması alınarak modellerin veri setinin genelindeki performansları elde edilmiştir. 4 kat için ortalama mAP değerleri Yolov3, Yolov4 ve Yolov7 için sırasıyla %84,93; %95,97 ve %57,29 olarak hesaplanmıştır. Yolov3 ve Yolov4 için elde edilen sonuçlar literatürdeki çalışmalarda kıyaslandığında iyi sonuçlar veriyor olsa da Yolov7 modelinin iyileştirilmesi gerekmektedir.

Özet olarak bu tez çalışması kapsamında;

- Nesne tespit performansını artırmak için veri çoğaltma yöntemleri sunulmuştur.
- Kızılötesi görüntülerden ve bu görüntüleri elde etme yöntemlerinden bahsedilmiştir.
- Literatürde güncel olarak çalışılan bir konu olan kızılötesi nesne tespit uygulamaları için üç adet YOLO modeli sunulmuştur.
- Çizelge 4.5'te görüldüğü üzere literatürdeki çalışmalarla kıyaslandığında yüksek sahip bir Yolov4 modeli elde edilmiştir.

Elde edilen sonuçlar değerlendirildiğinde;

- En iyi sonucu veren modellerin Yolov3 ve Yolov4 olmasında, modellerin omurgalarında yer alan Darknet53'ün kızılötesi nesne özelliklerini başarılı bir şekilde çıkarması olduğu düşünülmektedir. Bunun yanı sıra Yolov4'ün Yolov3'e göre daha yüksek performansa sahip olmasında, Darknet53 omurgasına eklenen CSPNet olduğu düşünülmektedir. Bunun sebebinin CSPNet'in özellik karmaşıklığı ile başa çıkması olduğu düşünülmektedir.
- Yolov3 ve Yolov4 modellerinin boyun kısmında yer alan FPN ve PANet modüllerinin benzer olması bu modellerin performanslarının benzer olmasında etkili olduğu düşünülmektedir. Bunun yanı sıra Yolov3'ün nesnelerin gerçek sınırlayıcı kutularını daha iyi tahmin etmesinin sebebi FPN modülünün yüksek seviye özellikleri tespit etmedeki başarısından kaynaklanabilmektedir (G. Zhao vd., 2021).
- Yolov4 modelinin performansının daha yüksek olmasındaki bir diğer sebep PANet modülünün uzamsal bilgileri ve piksel konumlarının çıkarmadaki başarısından kaynaklanmaktadır. Dolayısıyla küçük nesneleri diğer modellere göre daha iyi tespit etmiştir.
- Yolov7 omurgasında bulunun ELAN modülü, teoride derin ağların eğitim süresini azaltırken, tespit yeteneklerini artırmaya yöneliktir. Ayrıca Yolov7 makalesine göre, nesnelerin gerçek sınırlayıcı kutularının tahmin etmede Yolov7'nin daha başarılı olması beklenmektedir. Fakat bu tez çalışmasında Yolov7 başarılı bir sonuç verememiştir. Bunun sebepleri; ELAN modülünün özellik haritalarını birleştirmesi ve bizim düşüncemize göre kızılötesi nesne özelliklerinin kaybolması; ELAN modülünün kızılötesi nesne özelliklerinin öğrenmede yetersiz kalması; Yolov7'nin kafa yapısında yer alan öncelikli kafa ve yardımcı kafa ayırımından dolayı insan ve araç etiketlerinin yetersiz kalması olarak düşünülmektedir. Bu durumların, Yolov7 modelinin kızılötesi alana adaptasyonunu engellediği varsayılmaktadır.

Bundan sonraki çalışmalarda;

- Nesne sınıfları artırılabilceği gibi kızılötesi görüntülerdeki bantlar da artırılabilir.

- Farklı veri çoğaltma yöntemleri ile veri karmaşıklığı artırılabilir.
- Henüz yeni çıkan Yolov7 modelinin performansını artırmak için hiper parametreler değiştirilebilir. Ayrıca insan ve araç gibi genel etiketler yerine erkek-kadın, otomobil-kamyon gibi daha özel etiketler kullanılabilir.
- Yolov7 modeli için eğitim veri setine kızılötesi görüntülerin yanı sıra renkli görüntü verileri de eklenerek performans analizi yapılabilir. Bunun yanı sıra
- Yolov4'ün halı hazırda elde ettiği başarıyı artırmak için geniş çaplı bir hiper parametre iyileştirmesi yapılabilir.
- Bu tez çalışması sırasında en zorlayıcı aşama yüksek GPU performansı ihtiyacından kaynaklı modellerin eğitim aşamasının çok uzun sürmesiydi. Bundan dolayı ilerleyen çalışmalar, şu anda güncel çalışılan bir konu olarak, daha hızlı veya daha küçük boyutlu modellerin geliştirilmeye çalışması olabilir.



KAYNAKLAR

AAU-PD-T | Kaggle. (2020). <https://www.kaggle.com/datasets/noorulhuda90/aaupdt>

Akhloufi, M. A., & Bendada, A. (2013). Fusion of active and passive infrared images for face recognition. *Thermosense: Thermal Infrared Applications XXXV*, 8705(May 2013), 87050B. <https://doi.org/10.1117/12.2017942>

Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2018a). Understanding of a convolutional neural network. *Proceedings of 2017 International Conference on Engineering and Technology, ICET 2017, 2018-Janua*, 1–6. <https://doi.org/10.1109/ICEngTechnol.2017.8308186>

Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2018b). Understanding of a convolutional neural network. *Proceedings of 2017 International Conference on Engineering and Technology, ICET 2017, 2018-January*, 1–6. <https://doi.org/10.1109/ICENGTECHNOL.2017.8308186>

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4, 40–79. <https://doi.org/10.1214/09-SS054>

Bagavathiappan, S., Lahiri, B. B., Saravanan, T., Philip, J., & Jayakumar, T. (2013). Infrared thermography for condition monitoring - A review. *Infrared Physics and Technology*, 60, 35–55. <https://doi.org/10.1016/j.infrared.2013.03.006>

Berg, A., Ahlberg, J., & Felsberg, M. (2015). A thermal Object Tracking benchmark. *AVSS 2015 - 12th IEEE International Conference on Advanced Video and Signal Based Surveillance*. <https://doi.org/10.1109/AVSS.2015.7301772>

Berrar, D. (2018). Cross-validation. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, 1–3(April), 542–545. <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>

Bhowmick, S., Kuiry, S., Das, A., Das, N., & Nasipuri, M. (2022). Deep Learning-Based Outdoor Object Detection Using Visible and Near-Infrared Spectrum. *Multimedia Tools and Applications*, 81(7), 9385–9402. <https://doi.org/10.1007/s11042-021-11848-2>

Bi, Y., Bai, X., Jin, T., & Guo, S. (2017). Multiple feature analysis for infrared small target detection. *IEEE Geoscience and Remote Sensing Letters*, 14(8), 1333–1337. <https://doi.org/10.1109/LGRS.2017.2711047>

Bochkovskiy, A. (2022). *Darknet*. github. <https://github.com/AlexeyAB/darknet>

- Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). *YOLOv4: Optimal Speed and Accuracy of Object Detection*. <https://doi.org/10.48550/arxiv.2004.10934>
- Boreman, G. . D. (1998). *Basic Electro-Optics for Electrical Engineers*. <https://doi.org/10.1117/3.294180>
- Broggi, A., Fascioli, A., Carletti, M., Graf, T., & Meinecke, M. (2004). A multi-resolution approach for infrared vision-based pedestrian detection. *IEEE Intelligent Vehicles Symposium, Proceedings*, 7–12. <https://doi.org/10.1109/IVS.2004.1336347>
- Choi, K. K., Monroy, C., Goldberg, A., Dang, G., Jhabvala, M., La, A., Tamir, T., Leung, K. M., Majumdar, A., Li, J., & Tsui, D. C. (2005). Designs and applications of corrugated QWIPs. *Infrared Physics and Technology*, 47(1–2), 76–90. <https://doi.org/10.1016/j.infrared.2005.02.013>
- D’Acremont, A., Fablet, R., Baussard, A., & Quin, G. (2019). CNN-based target recognition and identification for infrared imaging in defense systems. *Sensors (Switzerland)*, 19(9), 1–16. <https://doi.org/10.3390/s19092040>
- Danaci, K. I., & Akagunduz, E. (2022). *A Survey on Infrared Image and Video Sets*. <https://doi.org/10.48550/arxiv.2203.08581>
- Datla, R. U., & Parr, A. C. (2005). 1. Introduction to Optical Radiometry. *Experimental Methods in the Physical Sciences*, 41(C), 1–34. [https://doi.org/10.1016/S1079-4042\(05\)41001-2](https://doi.org/10.1016/S1079-4042(05)41001-2)
- Delimata, N. (2019). Historical Overview. *Philosophy and Medicine*, 131, 7–17. https://doi.org/10.1007/978-3-030-21898-0_2
- Dobrzański, L., Nossarzewska-Orłowska, E., Nowak, Z., & Piotrowski, J. (1997). Micromachined silicon bolometers as detectors of soft X-ray, ultraviolet, visible and infrared radiation. *Sensors and Actuators, A: Physical*, 60(1–3), 154–159. [https://doi.org/10.1016/S0924-4247\(97\)01512-4](https://doi.org/10.1016/S0924-4247(97)01512-4)
- Du, S., Zhang, B., Zhang, P., Xiang, P., & Xue, H. (2021). FA-YOLO: An Improved YOLO Model for Infrared Occlusion Object Detection under Confusing Background. *Wireless Communications and Mobile Computing*, 2021. <https://doi.org/10.1155/2021/1896029>
- Du, S., Zhang, P., Zhang, B., & Xu, H. (2021). Weak and Occluded Vehicle Detection in Complex Infrared Environment Based on Improved YOLOv4. *IEEE Access*, 9, 25671–25680. <https://doi.org/10.1109/ACCESS.2021.3057723>

- Expansion, F. C., & Detection, P. (2020). *Feature Channel Expansion and Background Pedestrian Detection*.
- Fan, J., Lee, J. H., Jung, I. S., & Lee, Y. K. (2021). Improvement of Object Detection Based on Faster R-CNN and YOLO. *2021 36th International Technical Conference on Circuits/Systems, Computers and Communications, ITC-CSCC 2021*. <https://doi.org/10.1109/ITC-CSCC52171.2021.9501480>
- Gallagher, J. E., & Oughton, E. J. (2022). *Assessing thermal imagery integration into object detection methods on ground-based and air-based collection platforms*. 1–29.
- Gao, S., Ruan, Y., Hong, Q., & Yin, D. (2022). Infrared Thermal Image Fault Detection based on YOLOV3-L. *2022 IEEE International Conference on Artificial Intelligence and Computer Applications, ICAICA 2022*, 175–178. <https://doi.org/10.1109/ICAICA54878.2022.9844534>
- Gavrila, D. M. (1999). *Traffic Sign Recognition Revisited*. 86–93. https://doi.org/10.1007/978-3-642-60243-6_10
- Gebhardt, E., & Wolf, M. (2019). CAMEL Dataset for Visual and Thermal Infrared Multiple Object Detection and Tracking. *Proceedings of AVSS 2018 - 2018 15th IEEE International Conference on Advanced Video and Signal-Based Surveillance*. <https://doi.org/10.1109/AVSS.2018.8639094>
- Girshick, R. (2015). Fast R-CNN. *Proceedings of the IEEE International Conference on Computer Vision, 2015 Inter*, 1440–1448. <https://doi.org/10.1109/ICCV.2015.169>
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2013). Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 580–587. <https://doi.org/10.48550/arxiv.1311.2524>
- González, A., Fang, Z., Socarras, Y., Serrat, J., Vázquez, D., Xu, J., & López, A. M. (2016). Pedestrian Detection at Day/Night Time with Visible and FIR Cameras: A Comparison. *Sensors 2016*, Vol. 16, Page 820, 16(6), 820. <https://doi.org/10.3390/S16060820>
- Gu, Y., Wang, C., Liu, B., & Zhang, Y. (2010). A kernel-based nonparametric regression method for clutter removal in infrared small-target detection applications. *IEEE Geoscience and Remote Sensing Letters*, 7(3), 469–473. <https://doi.org/10.1109/LGRS.2009.2039192>

- Guo, T., Huynh, C. P., & Solh, M. (2019). Domain-Adaptive Pedestrian Detection in Thermal Images. *Proceedings - International Conference on Image Processing, ICIP, 2019-Septe*, 1660–1664. <https://doi.org/10.1109/ICIP.2019.8803104>
- Hamamatsu Photonics K.K., S. S. D. (2011). *Characteristics and Use of Infrared Sensors*.
- He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2017). *Mask R-CNN* (ss. 2961–2969).
- Hlavac, V. (2011). Fundamentals of Image Processing. *Optical and Digital Image Processing: Fundamentals and Applications*, July, 71–96. <https://doi.org/10.1002/9783527635245.ch4>
- Hou, Z., Sun, Y., Guo, H., Li, J., Ma, S., & Fan, J. (2022). M-YOLO: an object detector based on global context information for infrared images. *Journal of Real-Time Image Processing*, 19(6), 1009–1022. <https://doi.org/10.1007/S11554-022-01242-Y/FIGURES/9>
- Huda, N. U., Hansen, B. D., Gade, R., & Moeslund, T. B. (2020a). The effect of a diverse dataset for transfer learning in thermal person detection. *Sensors (Switzerland)*, 20(7). <https://doi.org/10.3390/s20071982>
- Huda, N. U., Hansen, B. D., Gade, R., & Moeslund, T. B. (2020b). The Effect of a Diverse Dataset for Transfer Learning in Thermal Person Detection. *Sensors 2020, Vol. 20, Page 1982*, 20(7), 1982. <https://doi.org/10.3390/S20071982>
- Ince, O., Senel, I. K., & Yilmaz, F. (2020). Image Processing and Analysis in Health: Advantages, Challenges, Threats and Examples. *Archives of Health Science and Research*, 7(1), 66–74. <https://doi.org/10.5152/ArcHealthSciRes.2020.535808>
- Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, 29(3), 31–44. <https://doi.org/10.1109/2.485891>
- Jangblad, M. (2018). Object Detection in Infrared Images using Deep Convolutional Neural Networks. *undefined*.
- Ji, Y., Zhang, H., Zhang, Z., & Liu, M. (2021). CNN-based encoder-decoder networks for salient object detection: A comprehensive review and recent advances. *undefined*, 546, 835–857. <https://doi.org/10.1016/J.INS.2020.09.003>
- Jiang, C., Ren, H., Ye, X., Zhu, J., Zeng, H., Nan, Y., Sun, M., Ren, X., & Huo, H. (2022). Object detection from UAV thermal infrared images and videos using YOLO models.

International Journal of Applied Earth Observation and Geoinformation, 112(June), 102912. <https://doi.org/10.1016/j.jag.2022.102912>

Jiang, K., Xie, T., Yan, R., Wen, X., Li, D., Jiang, H., Jiang, N., & Feng, L. (2022). *An Attention Mechanism-Improved YOLOv7 Object Detection Algorithm for Hemp Duck Count Estimation*.

Ju, M., Luo, J., Zhang, P., He, M., & Luo, H. (2019). A Simple and Efficient Network for Small Target Detection. *IEEE Access*, 7, 85771–85781. <https://doi.org/10.1109/ACCESS.2019.2924960>

Kannadaguli, P. (2020). YOLO v4 Based Human Detection System Using Aerial Thermal Imaging for UAV Based Surveillance Applications. *2020 International Conference on Decision Aid Sciences and Application, DASA 2020*, 1213–1219. <https://doi.org/10.1109/DASA51403.2020.9317198>

Kanti Bhowmik, M., Kankan, Majumder, S., Majumder, G., Saha, A., Nath, A., Bhattacharjee, D., Basu, D. K., & Nasipuri, M. (2011). Thermal Infrared Face Recognition – A Biometric Identification Technique for Robust Security system. *Reviews, Refinements and New Ideas in Face Recognition*. <https://doi.org/10.5772/18986>

Karim, A., & Andersson, J. Y. (2013). Infrared detectors: Advances, challenges and new technologies. *{IOP} Conference Series: Materials Science and Engineering*, 51, 12001. <https://doi.org/10.1088/1757-899x/51/1/012001>

Kerbaa, T. H., Mezache, A., & Oudira, H. (2019). Model Selection of Sea Clutter Using Cross Validation Method. *Procedia Computer Science*, 158, 394–400. <https://doi.org/10.1016/j.procs.2019.09.067>

Kowalski, M. Ł., Pałka, N., Młyńczak, J., Karol, M., Czerwińska, E., Życzkowski, M., Ciurapiński, W., Zawadzki, Z., & Brawata, S. (2021). Detection of Inflatable Boats and People in Thermal Infrared with Deep Learning Methods. *Sensors 2021*, Vol. 21, Page 5330, 21(16), 5330. <https://doi.org/10.3390/S21165330>

Kristo, M., Ivasic-Kos, M., & Pobar, M. (2020). Thermal Object Detection in Difficult Weather Conditions Using YOLO. *IEEE Access*, 8, 125459–125476. <https://doi.org/10.1109/ACCESS.2020.3007481>

Li, C., Liang, X., Lu, Y., Zhao, N., & Tang, J. (2019). RGB-T object tracking: Benchmark and baseline. *Pattern Recognition*, 96, 106977. <https://doi.org/10.1016/J.PATCOG.2019.106977>

- Li, S., Wang, Y., Feng, C., Zhang, D., Li, H., Huang, W., & Shi, L. (2022). A Thermal Imaging Flame-Detection Model for Firefighting Robot Based on YOLOv4-F Model. *Fire*, 5(5), 1–18. <https://doi.org/10.3390/fire5050172>
- Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2017). Focal Loss for Dense Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2), 318–327. <https://doi.org/10.48550/arxiv.1708.02002>
- Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8693 LNCS(PART 5), 740–755. https://doi.org/10.1007/978-3-319-10602-1_48/COVER
- Liu, S., Wang, Y., Yu, Q., Liu, H., & Peng, Z. (2022). CEAM-YOLOv7: Improved YOLOv7 Based on Channel Expansion and Attention Mechanism for Driver distraction behavior detection. *IEEE Access*, 10(December), 129116–129124. <https://doi.org/10.1109/ACCESS.2022.3228331>
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9905 LNCS, 21–37. https://doi.org/10.1007/978-3-319-46448-0_2/FIGURES/5
- Liu, Y., Su, H., Zeng, C., & Li, X. (2021). A robust thermal infrared vehicle and pedestrian detection method in complex scenes. *Sensors (Switzerland)*, 21(4), 1–15. <https://doi.org/10.3390/s21041240>
- Liu, Y., Sun, P., Wergeles, N., & Shang, Y. (2021a). A Survey and Performance Evaluation of Deep Learning Methods for Small Object Detection. *Expert Systems with Applications*, 114602. <https://doi.org/10.1016/j.eswa.2021.114602>
- Liu, Y., Sun, P., Wergeles, N., & Shang, Y. (2021b). A survey and performance evaluation of deep learning methods for small object detection. *Expert Systems with Applications*, 172, 114602. <https://doi.org/10.1016/J.ESWA.2021.114602>
- Lozano, A., & Hassanipour, F. (2019). Infrared imaging for breast cancer detection: An objective review of foundational studies and its proper role in breast cancer screening. *Infrared Physics and Technology*, 97(December 2018), 244–257. <https://doi.org/10.1016/j.infrared.2018.12.017>

- Lyra, S., Mayer, L., Ou, L., Chen, D., Timms, P., Tay, A., Chan, P. Y., Ganse, B., Leonhardt, S., & Antink, C. H. (2021). A deep learning-based camera approach for vital sign monitoring using thermography images for ICU patients. *Sensors*, 21(4), 1–18. <https://doi.org/10.3390/s21041495>
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics* 1943 5:4, 5(4), 115–133. <https://doi.org/10.1007/BF02478259>
- Melik Ahmet Daye, B. (2021). *a Comprehensive Survey on Small Object Detection*.
- Mittal, P., Singh, R., & Sharma, A. (2020). Deep learning-based object detection in low-altitude UAV datasets: A survey. *Image and Vision Computing*, 104, 104046. <https://doi.org/10.1016/J.IMAVIS.2020.104046>
- Nguyen-Meidine, L. T., Granger, E., Kiran, M., & Blais-Morin, L. A. (2018). A comparison of CNN-based face and head detectors for real-time video surveillance applications. *Proceedings of the 7th International Conference on Image Processing Theory, Tools and Applications, IPTA 2017, 2018-January*, 1–7. <https://doi.org/10.1109/IPTA.2017.8310113>
- O’Shea, K., & Nash, R. (2015). *An Introduction to Convolutional Neural Networks*. 1–11. <http://arxiv.org/abs/1511.08458>
- OpenCV. (2000). <https://opencv.org/>
- Patel, V. M., Gopalan, R., Li, R., & Chellappa, R. (2015). Visual Domain Adaptation: A survey of recent advances. *IEEE Signal Processing Magazine*, 32(3), 53–69. <https://doi.org/10.1109/MSP.2014.2347059>
- Petrou, M., & Petrou, C. (2010). Image Processing: The Fundamentals. İçinde *Image Processing: The Fundamentals*. <https://doi.org/10.1002/9781119994398>
- Piñeiro-Ave, J., Blanco-Velasco, M., Cruz-Roldán, F., & Artés-Rodríguez, A. (2014). Target detection for low cost uncooled MWIR cameras based on empirical mode decomposition. *Infrared Physics and Technology*, 63, 222–231. <https://doi.org/10.1016/j.infrared.2014.01.007>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Computer Society Conference on*

- Computer Vision and Pattern Recognition*, 2016-Decem, 779–788. <https://doi.org/10.1109/CVPR.2016.91>
- Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. *arXiv*.
- Ren, S., He, K., Girshick, R., Zhang, X., & Sun, J. (2017). Object detection networks on convolutional feature maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(7), 1476–1481. <https://doi.org/10.1109/TPAMI.2016.2601099>
- Rogalski, A. (2002). Infrared detectors: An overview. *Infrared Physics and Technology*, 43(3–5), 187–210. [https://doi.org/10.1016/S1350-4495\(02\)00140-8](https://doi.org/10.1016/S1350-4495(02)00140-8)
- Roszyk, K., Nowicki, M. R., & Skrzypczyński, P. (2022). Adopting the YOLOv4 Architecture for Low-Latency Multispectral Pedestrian Detection in Autonomous Driving. *Sensors* 2022, Vol. 22, Page 1082, 22(3), 1082. <https://doi.org/10.3390/S22031082>
- Shao, Y., Zhang, X., Chu, H., Zhang, X., Zhang, D., & Rao, Y. (2022). AIR-YOLOv3: Aerial Infrared Pedestrian Detection via an Improved YOLOv3 with Network Pruning. *Applied Sciences (Switzerland)*, 12(7). <https://doi.org/10.3390/app12073627>
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(1), 1–48. <https://doi.org/10.1186/S40537-019-0197-0/FIGURES/33>
- Suard, F., Rakotomamonjy, A., Bensch, A., & Broggi, A. (2006). Pedestrian detection using infrared images and histograms of oriented gradients. *IEEE Intelligent Vehicles Symposium, Proceedings*, 206–212. <https://doi.org/10.1109/IVS.2006.1689629>
- The PASCAL Visual Object Classes Homepage. (2005). <http://host.robots.ox.ac.uk/pascal/VOC/>
- Tong, K., Wu, Y., & Zhou, F. (2020). Recent advances in small object detection based on deep learning: A review. *Image and Vision Computing*, 97, 103910. <https://doi.org/10.1016/j.imavis.2020.103910>
- Torrallba, A., Fergus, R., & Freeman, W. T. (2008). 80 million tiny images: A large data set for nonparametric object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(11), 1958–1970. <https://doi.org/10.1109/TPAMI.2008.128>
- Tumas, P., Serackis, A., & Nowosielski, A. (2021). Augmentation of severe weather

- impact to far-infrared sensor images to improve pedestrian detection system. *Electronics (Switzerland)*, 10(8), 1–16. <https://doi.org/10.3390/electronics10080934>
- Venkataraman, S. B., & Raj, B. (2003). Performance parameters for thermal imaging systems. *Insight: Non-Destructive Testing and Condition Monitoring*, 45(8), 531–535. <https://doi.org/10.1784/insi.45.8.531.52914>
- Vollmer, M., & Möllmann, K.-P. (2010). Basic Properties of IR Imaging Systems. *Infrared Thermal Imaging*, 73–155. <https://doi.org/10.1002/9783527630868.CH2>
- Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. 1–15. <http://arxiv.org/abs/2207.02696>
- Wang, M., & Deng, W. (2018). Deep Visual Domain Adaptation: A Survey. *Neurocomputing*, 312, 135–153. <https://doi.org/10.1016/j.neucom.2018.05.083>
- Wen, M., Wei, L., Zhuang, X., He, D., Wang, S., & Wang, Y. (2018). High-sensitivity short-wave infrared technology for thermal imaging. *Infrared Physics and Technology*, 95(October), 93–99. <https://doi.org/10.1016/j.infrared.2018.10.020>
- Wolff, L. B., Socolinsky, D. A., & Eveland, C. K. (2005). Face Recognition in the Thermal Infrared. *Computer Vision Beyond the Visible Spectrum*, 167–191. https://doi.org/10.1007/1-84628-065-6_6
- Wu, Y., Pan, F., An, Q., Wang, J., Feng, X., & Cao, J. (2021). Infrared Target Detection Based on Deep Learning. *Chinese Control Conference, CCC, 2021-July*, 8175–8180. <https://doi.org/10.23919/CCC52363.2021.9549852>
- Xiao, J. (2021). ExYOLO: A small object detector based on YOLOv3 Object Detector. *Procedia CIRP*, 188(2019), 18–25. <https://doi.org/10.1016/j.procs.2021.05.048>
- Xu, Z., Zhuang, J., Liu, Q., Zhou, J., & Peng, S. (2019). Benchmarking a large-scale FIR dataset for on-road pedestrian detection. *Infrared Physics and Technology*, 96(October 2018), 199–208. <https://doi.org/10.1016/j.infrared.2018.11.007>
- Yun, J.-S., Park, S.-H., & Yoo, S. B. (2022). Infusion-Net: Inter- and Intra-Weighted Cross-Fusion Network for Multispectral Object Detection. *Mathematics*, 10(21), 3966. <https://doi.org/10.3390/math10213966>
- Zhang, H., Luo, C., Wang, Q., Kitchin, M., Parmley, A., Monge-Alvarez, J., & Casaseca-de-la-Higuera, P. (2018). A novel infrared video surveillance system using deep

- learning based techniques. *Multimedia Tools and Applications*, 77(20), 26657–26676. <https://doi.org/10.1007/s11042-018-5883-y>
- Zhang, Y., & Yang, Y. (2015). Cross-validation for selecting a model selection procedure. *Journal of Econometrics*, 187(1), 95–112. <https://doi.org/10.1016/j.jeconom.2015.02.006>
- Zhao, G., Ge, W., & Yu, Y. (2021). GraphFPN: Graph Feature Pyramid Network for Object Detection. *Proceedings of the IEEE International Conference on Computer Vision*, 2743–2752. <https://doi.org/10.1109/ICCV48922.2021.00276>
- Zhao, X., He, Z., Zhang, S., & Liang, D. (2015). Robust pedestrian detection in thermal infrared imagery using a shape distribution histogram feature and modified sparse representation classification. *Pattern Recognition*, 48(6), 1947–1960. <https://doi.org/10.1016/j.patcog.2014.12.013>
- Zhu, Z., Liang, D., Zhang, S., Huang, X., Li, B., & Hu, S. (2016). Traffic-Sign Detection and Classification in the Wild. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-December, 2110–2118. <https://doi.org/10.1109/CVPR.2016.232>

EKLER

EK-1. Tezde kullanılan derin öğrenme terminolojileri ve İngilizceleri

Türkçe Terminoloji	İngilizce Terminoloji
Ağ Mimarisi Araması	Network Architecture Search (Nas)
Alan Adaptasyonu	Domain Adaptation
Algılama	Detectivity
Alıcı İşletim Karakteristiği	Receiver Operating Characteristic
Alt Bölümler	Subdivision
Aşırı Uyum Gösterme	Overfitting
Bilgisayarlı Görü	Computer Vision
Bölge Teklif Ağı	Region Proposal Network
Bölgesel Evrişimsel Sinir Ağları	Region Based Convolutional Neural Network (Rcnn)
Çapa Kutusu	Anchor Box
Çapraz Aşamalı Kısmi Ağ	Cross Stage Partial Network (Cspnet)
Çapraz Doğrulama	Cross Validation
Denetimli Öğrenme	Supervised Learning
Denetimsiz Öğrenme	Unsupervised Learning
Derin Öğrenme	Deep Learning
Duyarlılık	Recall
Düzenlileştirme	Regularization
Evrişim Katmanı	Convolutional Layer
Evrişimsel Sinir Ağları	Convolutional Neural Network
Foto Duyarlılığı	Photo Sensitivity
Genişletilmiş Verimli Katman Toplama Ağı	Extended Efficient Layer Aggregation Network (E-Lan)
Gürültü Eşdeğer Güç	Noise Equivalent Power

EK-1. Tezde Kullanılan derin öğrenme Terminolojileri ve İngilizceleri (devam)

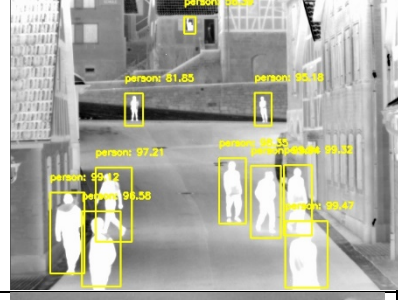
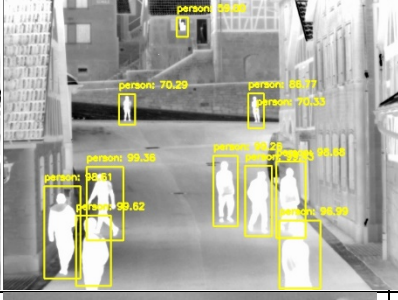
Güven Skoru	Confidence Score
Hassasiyet	Precision
Havuzlama Katmanı	Pooling Layer
Hızlı Bölgesel Evrişimsel Sinir Ağları	Fast Region Based Convolutional Neural Network (Fast Rcn)
Hibrit Öğrenme	Reinforcement Learning
Kendi Kendine Denetimli Öğrenme	Self-Supervised Learning
Kesişimin Birleşime Oranı	Intersection Over Union
Makine Öğrenmesi	Machine Learning
Maksimum Olmayan Bastırma	Nun Maximum Supression
Maskelenmiş Hızlı Bölgesel Evrişimsel Sinir Ağları	Mask Region Based Convolutional Neural Network (Mask Rcn)
Mekânsal Piramit Havuzlaması	Spatial Pyramid Pooling (Spp)
Ortalama Hassasiyet	Average Precision
Genel Ortalama Hassasiyet	Mean Average Precision (mAP)
Öğrenme Oranı	Learning Rate
Özellik Piramit Ağı	Feature Pyramid Network
Pekiştirmeli Öğrenme	Reinforcement Learning
Sadece Bir Kez Bakarsın	You Only Look Once (Yolo)
Sert Olumsuz Örnek Madencilik Bloğu	Hard Negative Example Mining Block
Sönüm	Decay
Tam Bağlantı Katmanı	Fully Connected Layer
Tek Atış Dedektörü	Single Shot Detector (Ssd)
Toplam-Kare Hatası	Sum-Squared Error
Toplu Normalleştirme	Batch Normalization

EK-1. Tezde Kullanılan derin öğrenme Terminolojileri ve İngilizceleri (devam)

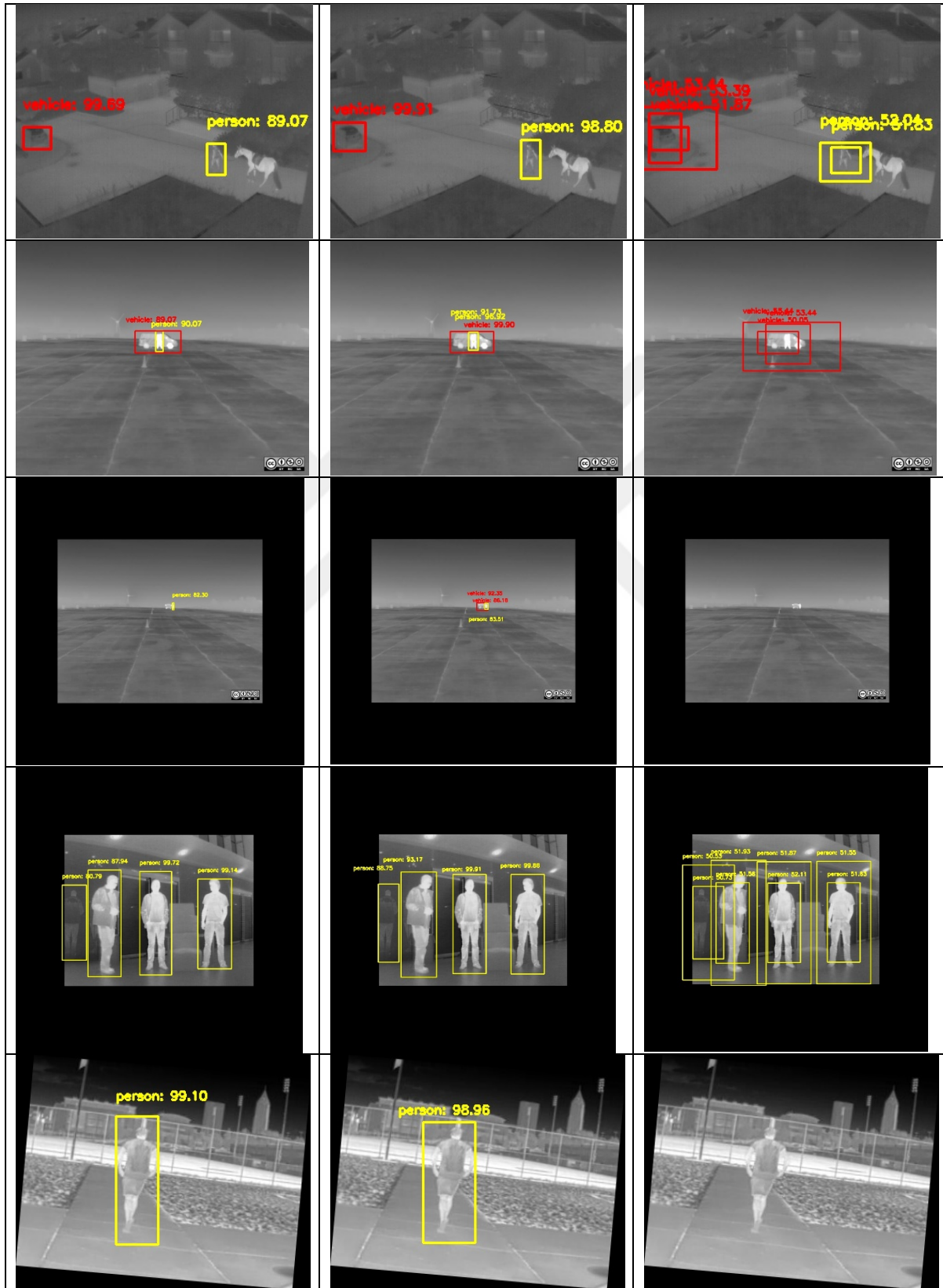
Toptan	Batch
Yapay Sinir Ağları	Artificial Neural Networks
Yarı Denetimli Öğrenme	Semi-Supervised Learning
Yol Toplama Ağı	Path Aggregation Network (Panet)



EK-2. Yolov3, Yolov4 Ve Yolov7 modellerinin örnek görüntüler üzerinden nesne tespit performansları

YOLOv3	YOLOv4	YOLOv7
		
		
		
		
		
		

EK-2. Yolov3, Yolov4 Ve Yolov7 modellerinin örnek görüntüler üzerinden nesne tespit performansları (devam)



EK-2. Yolov3, Yolov4 Ve Yolov7 modellerinin örnek görüntüler üzerinden nesne tespit performansları (devam)



ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : D*****, K**** İ***
 Uyruğu :
 Doğum tarihi ve yeri :
 Medeni hali :
 e-mail :

Eğitim

Derece	Eğitim Birimi	Mezuniyet Tarihi
Yüksek lisans	SBTÜ/ Savunma Teknolojileri	2023
Lisans	Selçuk Üniversitesi/ Elektrik Elektronik Müh.	2018

İş Deneyimi

Yıl	Yer	Görev
2019-Halen	Sivas Bilim ve Teknoloji Üniversitesi	Araştırma Görevlisi

Yabancı Dil

İngilizce

Yayınlar

- Danacı, K. İ., & Akagunduz, E. (2022). A Survey on Infrared Image and Video Sets. <https://doi.org/10.48550/arxiv.2203.08581>
- DANACI, İ., ÖNEL, M. N., ÖZBAY, S., DANACI, K. İ., ERDEN, F., & KATIRCI, R. (2022). Determination And Optimization Of Parameters Affecting Cu-Zn-Sn Layers In Czts Solar Cell. 2nd INTERNATIONAL ISTANBUL CONGRESS OF MULTIDISCIPLINARY SCIENTIFIC RESEARCH, 305–311.
- ÖNEL, M. N., DANACI, İ., ÖZBAY, S., DANACI, K. İ., ERDEN, F., & KATIRCI, R. (2022). Effect Of Surface Hydrophilicity On The Sulphurization Performance In Czts Films. 14th INTERNATIONAL CONFERENCE ON ENGINEERING & NATURAL SCIENCES, 219–226.
- K.İ. Danacı, E. Akagündüz, KIZILÖTESİ NESNE ALGILAMA UYGULAMALARINDA VERİ ÇOĞALTMA YÖNTEMİNİN PERFORMANSA ETKİSİ, Int. AEGEAN Conf., 2022: ss. 60–61.



**SİVAS
BİLİM VE TEKNOLOJİ
ÜNİVERSİTESİ**

KÖKLERDEN GÖKLERE...