



T.C.
KONYA TEKNİK ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



DERİN ÖĞRENME TABANLI YÜZ
ALGILAMA SİSTEMİYLE TEMASSIZ ATEŞ
ÖLÇÜMÜ

Kubilay TUNA

YÜKSEK LİSANS TEZİ

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Aralık-2022
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Kubilay Tuna tarafından hazırlanan “Derin Öğrenme Tabanlı Yüz Algılama Sistemiyle Temassız Ateş Ölçümü” adlı tez çalışması 22/12/2022 tarihinde aşağıdaki jüri tarafından oy birliği ile Konya Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü Elektrik-Elektronik Mühendisliği Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan

Dr. Öğr. Üyesi Sabri ALTUNKAYA

.....

Danışman

Doç. Dr. Bayram AKDEMİR

.....

Üye

Dr. Öğr. Üyesi Hasan KOYUNCU

.....

Yukarıdaki sonucu onaylarım.

Prof. Dr. Saadettin Erhan KESEN
Enstitü Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Kubilay TUNA

22/12/2022

ÖZET

YÜKSEK LİSANS TEZİ

DERİN ÖĞRENME TABANLI YÜZ ALGILAMA SİSTEMİYLE TEMASSIZ ATEŞ ÖLÇÜMÜ

Kubilay TUNA

**Konya Teknik Üniversitesi
Lisansüstü Eğitim Enstitüsü
Elektrik-Elektronik Mühendisliği Anabilim Dalı**

Danışman: Doç. Dr. Bayram AKDEMİR

2022, 69 Sayfa

Jüri

**Doç. Dr. Bayram AKDEMİR
Dr. Öğr. Üyesi Sabri ALTUNKAYA
Dr. Öğr. Üyesi Hasan KOYUNCU**

Pandemik hastalıkların hemen hemen hepsinin ortak semptomu yüksek ateştir. Böylece en basit şekilde, anormal vücut sıcaklığı olarak nitelendirilen 38 °C ve üstünde ateşe sahip kişilerin tespiti, bulaş riskini ortadan kaldırarak salgının kontrol altına alınmasını sağlayacaktır. Fakat yüzlerce kişinin giriş çıkış yaptığı kalabalık insan topluluğunun olduğu ortamlarda bu kişilerin kontrolünü sağlamak güçtür. Bu durum iş gücü, maliyet gerektirerek zaman kaybına neden olmaktadır. Haliyle içinde bulunduğumuz gelişmiş teknoloji dünyasında, bu süreci insan etkisini ortadan kaldırarak otomatikleştirmek gerekliliktir.

Bu tezde, Covid-19 nedeniyle enfekte olmuş kişilerin yüzlerini tespit etmek için özel Single Shot Detection (SSD) modeli kullanılmıştır. Tespit edilen bu yüzler üzerinde Ensemble of Regresyon Trees (ERT) modeliyle yüz işaret noktaları belirlenerek kişinin vücut sıcaklığının en doğru olduğu göz çevresinin tespit edilmesi önerilmiştir. Son olarak, termal değer, sensör füzyonu kullanılarak temassız bir şekilde göz çevresinden ölçülmüştür. Yapılan analizler sonucunda önerilen sistemin farklı ölçüm yöntemlerine yakın sonuçlar verdiği gözlemlenmiştir (Tuna ve Akdemir, 2022).

Anahtar Kelimeler: Covid-19, ERT, Pandemi, Sensör Füzyonu, SSD, Temassız, Yüz İşaret Noktaları

ABSTRACT

MS THESIS

NON-CONTACT FEVER MEASUREMENT WITH DEEP LEARNING BASED FACE DETECTION SYSTEM

Kubilay TUNA

**Konya Technical University
Institute of Graduate Studies
Department of Electrical-Electronics Engineering**

Advisor: Doç. Dr. Bayram AKDEMİR

2022, 69 Pages

Jury

**Assoc. Prof. Bayram Akdemir
Asst. Prof. Sabri ALTUNKAYA
Asst. Prof. Hasan KOYUNCU**

The common symptom of almost all pandemic diseases is high fever. Thus, in the simplest way, the detection of people with a fever of 38 °C and above, which is described as abnormal body temperature, will eliminate the risk of transmission, and ensure that the epidemic is brought under control. However, it is difficult to control these people in environments where hundreds of people enter and exit. This situation causes loss of time by requiring labor and cost. Therefore, in the advanced technology world we live in, it is necessary to automate this process by eliminating human influence.

In this thesis, custom Single Shot Detection (SSD) was used to detect infected people faces because of Covid-19. It has been suggested to determine around the eyes area where the body temperature of the person most accurate by determining the facial landmarks with Ensemble of Regression Trees (ERT) model on these detected faces. Finally, the thermal value was measured from around the eyes area in a non-contact way using sensor fusion. As a result of the analyzes made, it was observed that the proposed system gave results close to the different measurement methods (Tuna & Akdemir, 2022).

Keywords: Covid-19, ERT, Facial Landmark, Non-contact, Pandemic, Sensor Fusion, SSD

ÖNSÖZ

Bu tezde son dönemlerde tüm dünyayı derinden etkileyen ve birçok sektörü yıkıma uğratan Covid-19'un etkisini kontrol altına alabilecek ve ileride yaşanması muhtemel pandemik rahatsızlıkların tespiti için de rahatlıkla uyarlanabilecek bir derin öğrenme sistemi geliştirilmiştir. Bu sistem ön süreçleri otomatikleştirerek insan görüşüne destek olmaktadır. Önerilen derin öğrenme tabanlı yüz algılama sistemiyle riskli kişilerin hızlı, verimli ve kolay bir şekilde tespiti sağlanacaktır. Böylece, her türlü kurum ya da kuruluşlarda kolaylıkla kullanılabilir olan bu proje, günlük hayatın akışını olumsuz etkileyen pandemik rahatsızlıkların negatif etkilerini azaltarak, karşılaşılabilecek aksaklıkları giderebilir.

Tez danışmanım Doç. Dr. Bayram Akdemir'e tez konusunu belirlerken isteklerim doğrultusunda bana yol gösterici olduğu için ve çalışmalarım boyunca değerli desteğini esirgemediğinden dolayı sonsuz teşekkürlerimi iletirim.

Çalışmaya ayırdığım zamanı anlayışla karşılayıp bana her zaman destek olan ve bu doğrultuda güvenini hiçbir zaman esirgemeyen sevgili nişanlım Seher Çelik'e çok teşekkür ederim. Son olarak, hayatım boyunca maddi, manevi desteklerini esirgemeyip, bu yolda sürekli yanımda olan, beni bugünlere ulaştıran ve başarımın temeli olan aileme de sonsuz teşekkürlerimi sunarım.

Kubilay Tuna
KONYA-2022

İÇİNDEKİLER

ÖZET	iv
ABSTRACT	v
ÖNSÖZ	vi
SİMGELER VE KISALTMALAR	ix
1. GİRİŞ	1
2. KAYNAK ARAŞTIRMASI	2
3. MATERYAL VE YÖNTEM	7
3.1. İnsan Beyni ve Görü Sistemi	7
3.2. Yapay Sinir Ağları	10
3.2.1. Tekli nöron modeli.....	10
3.2.2. Aktivasyon fonksiyonları.....	11
3.2.3. İleri beslemeli sinir ağlar	11
3.2.4. Geri beslemeli sinir ağları.....	13
3.3. Derin Öğrenme	14
3.3.1. Konvolüzyonel sinir ağları	15
3.3.1.1. Özellik çıkarım aşaması.....	16
3.3.1.1.1. Konvolüsyon katmanı	17
3.3.1.1.2. Havuzlama katmanı	20
3.3.1.1.3. Relu (rectifier linear unit) katmanı	21
3.3.1.2. Sınıflandırma aşaması.....	21
3.3.1.2.1. Softmax fonksiyonu	22
3.3.2. Regularizasyon teknikleri.....	22
3.3.2.1. Dropout	22
3.3.3. Öğrenme süreci	23
3.3.3.1. Denetimli öğrenme	24
3.3.4. Veri setinin hazırlanması	25
3.4. Temel KSA yapıları ve Transfer Öğrenmesi	26
3.4.1. Temel ksa yapıları.....	26
3.4.1.1. AlexNet.....	26
3.4.1.2. VGGNet	27
3.4.1.3. GoogleNet.....	28
3.4.1.4. ResNet.....	29
3.4.1.5. DenseNet.....	29
3.4.2. Transfer öğrenme	30
3.4.3. Derin öğrenme kütüphaneleri	30
3.4.3.1. Caffe.....	31
3.4.3.1. TensorFlow	32
3.4.3.1. PyTorch.....	32

3.5. Nesne Tanıma Tabanlı Konvolüsyonel Sinir Ağları	33
3.5.1. IoU (Intersection over union)	33
3.5.2. Maksimum olmayan bastırma (NMS-Non-maximum suppression).....	34
3.5.3. Temel nesne tanıma tabanlı ksa yapıları.....	35
3.5.3.1. Bölgesel tabanlı konvolüsyonel sinir ağları.....	35
3.5.3.2. Hızlı bölgesel tabanlı konvolüsyonel sinir ağları	36
3.5.3.2. Daha hızlı bölgesel tabanlı konvolüsyonel sinir ağları.....	36
3.5.3.3. You only look once (YOLO)	38
3.5.3.4. Tek atış dedektörü (single shot detector-SSD)	39
3.6. Yüz işaretleri.....	40
3.7. Kızılötesi sıcaklık ölçüm teorisi	42
3.8. Nvidia Jetson Nano Modülü	45
3.9. Raspberry Pi Kamera Modülü	46
3.10. Optris CS LT IR Sıcaklık Sensörü.....	47
3.10. Önerilen Yöntem.....	47
4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA	51
5. SONUÇLAR VE ÖNERİLER.....	55
KAYNAKLAR.....	56

SİMGELER VE KISALTMALAR

Simgeler

°C	Santigrat Derece
°K	Kelvin Derece
mm	Milimetre
µm	Mikrometre
s	Saniye

Kısaltmalar

API	Application Programming Interface (Uygulama Programlama Arayüzü)
BVLC	Berkley Vision and Learning Center (Berkeley Görü ve Öğrenme Merkezi)
COCO	Common Objects in Context
CPU	Central Processing Unit (Merkezi İşlem Birimi)
CUDA	Compute Unified Device Architecture (Hesaplama Birleşik Cihaz Mimarisi)
DNN	Deep Neural Network (Derin Sinir Ağları)
DenseNet	Dense Convolutional Network (Yoğun Konvolüsyonel Sinir Ağları)
ERT	Ensemble of Regression Trees
Fddb	Face Detection Data Set and Benchmark
GPU	Graphic Processing Unit (Grafik İşlemci Ünitesi)
IoT	Internet of Things (Nesnelerin İnterneti)
IoU	Intersection Over Union (Kesiştirilmiş Bölgeler Ölçütleri)
IR	Infrared (Kızılötesi)
KSA (CNN)	Konvolüsyonel Sinir Ağları (Convolution Neural Networks)
LED	Light Emitting Diode (Işık Yayan Diyot)
MAFA	Masked Faces (Maskeli Yüzler)

NMS	Non Maximum Suppression (Maximum Olmayan Bastırma)
mAP	Mean Average Precision (Ortalama Hassasiyet)
RAM	Random Access Memory (Rastgele Erişimli Bellek)
ReLU	Rectified Linear Unit (İyileştirilmiş Doğrusal Birim)
ResNet	Residual Network (Artık Ağlar)
RGB	Red-Green-Blue (Kırmızı-Yeşil-Mavi)
R-CNN	Region Based CNN (Bölge Bazlı KSA)
ROI	Region of Interest (İlgili Alan)
ROM	Read Only Memory (Yalnızca Okunabilen) Bellek
RPN	Region Proposal Network (Bölge Öneri Ağları)
USB	Universal Serial Bus (Evrensel Seri Veriyolu)
VGG	Visual Geometry Group (Görsel Geometri Grubu)
YSA	Yapay Sinir Ağları

1. GİRİŞ

Hızla büyüyen toplumların arttığı bu dönemlerde her geçen gün insanoğlu ile vahşi yaşam giderek iç içe girmektedir. Bu da ekolojik bazı sorunların yanı sıra, insanlığı derinden etkileyerek birçok kayba neden olan salgınları ortaya çıkarmakta ve bazen bu salgınlar, dünyada birden fazla ülkede, çok geniş bir alanda yayılıp etkisini göstererek pandemileri meydana getirmektedir.

Tarihte de birçok örneğinin yer aldığı pandemiler insanlığın en büyük sorunlarından biridir. Bu tür bulaşıcı hastalıklarla mücadelede en önemli adım, mümkün olan en kısa sürede hastalıklı kişileri tespit ederek daha fazla kişiye hastalığın bulaşmasına engel olmaktır. Dolayısıyla kalabalık insan topluluğunun olduğu ortamlarda kişilerin kontrolünü sağlamak zorunluluktur. Fakat binlerce kişinin bulunduğu bu ortamlarda her bir bireyin kontrollü ciddi bir iş gücü gerektirmektedir. Bu noktada insan etkisini ortadan kaldırarak süreci otomatikleştirmek etkili bir çözümdür (Gupta ve ark., 2021).

Minimum hata, maksimum verimlilik ilkesiyle her geçen gün insanların yerini alan teknoloji, özellikle yapay zekâ algoritmalarının ve elektronik görüntüleme/ölçüm sistemlerinin geliştirilmesiyle birlikte medikal alanda da birçok soruna çözüm getirmektedir. Dolayısıyla küresel dünyada büyük kitleleri etkileyen salgınlarla mücadelede de teknolojinin rolü büyüktür.

Bu tezde derin öğrenme tabanlı bir yüz algılama sistemi geliştirilerek pandemik hastalıkların ortak semptomu olan yüksek ateşli kişilerin (38 °C ve üstü) tespiti amaçlanmıştır (Lun ve ark., 2021). Böylece son teknolojik gelişmeleri kullanarak riskli kişilerin hızlı, verimli ve kolay bir şekilde tespiti sağlanabilir. Her türlü kurum ya da kuruluşlarda kolaylıkla kullanılabilmesi planlanan bu proje, günlük hayatın akışını olumsuz etkileyen pandemik rahatsızlıkların negatif etkilerini azaltarak, karşılaşılabilecek aksaklıkları giderebilir.

Gerçekleştirilen tez çalışması beş bölümden oluşmaktadır. Tezin birinci bölümü giriş kısmı olup ikinci bölümde kaynak araştırması başlığı altında yüz algılama ve temasız ateş ölçümü üzerine yapılan çalışmalar incelenmiştir. Üçüncü bölümde ise tez kapsamında kullanılan materyal ve yöntemler den bahsedilmiştir. Dördüncü bölümde gerçekleştirilen çalışmalar ve elde edilen bulgular anlatılmıştır ve son olarak beşinci bölümde sonuçlar ve öneriler kısmı aktarılmıştır.

2. KAYNAK ARAŞTIRMASI

Oldukça popüler bir araştırma alanı olan yapay zekâ günümüzde birçok probleme etkili çözümler üretmektedir. Görüntü, metin, ses gibi farklı türde verileri anlamlandırıp işleyerek oldukça başarılı sonuçlar vermektedir. Bugün gündelik yaşantımızın önemli bir kısmında yapay zekâ çalışmalarından faydalanılmaktadır. Yapay zekâ, gelişmeye yeni başladığı dönemlerde bir problemi karmaşık matematiksel işlemlerle ve matematiksel formüllerle çözülebilmekteydi. Günümüzdeyse böyle problemlerin çözümü birbiri üzerine kurulu çok sayıda katmandan oluşan, eğitimle elde edilen deneyimler sayesinde sağlanmaktadır (Goodfellow ve ark., 2016).

Zhang ve arkadaşları (2014) yüz işaretlerinin (facial landmark) algılanmasını tek ve bağımsız bir sorun olarak ele almak yerine çok görevli öğrenme (multi-task learning) ile algılama sağlamlığının geliştirilmesinin önermiştir. Özellikle, heterojen ancak ince bir şekilde ilişkilendirilmiş görevlerle birlikte yüz işareti algılaması optimize edilmiştir. Yapılan çalışmalar sonucunda önerilen görev kısıtlamalı öğrenmenin (proposed task-constrained learning), özellikle şiddetli tıkamaya sahip yüzlerle başa çıkmada ve poz değişiminde mevcut yöntemlerden daha iyi performans gösterdiğini ve kademeli derin modele dayalı son teknoloji yöntemle karşılaştırıldığında model karmaşıklığını büyük ölçüde azalttığı gözlemlenmiştir.

Li H. ve arkadaşları (2015) gerçek dünyadaki yüzleri ışıklandırma gibi büyük görsel farklılıklardan soyutlayarak arka plandan doğru bir şekilde ayırt edebilen bir CNN kaskad modeli önermiştir. Dedektör, yüz olmayan bölgeleri hızlı bir şekilde reddetmek ve doğru algılama için zorlu bölgeleri daha yüksek çözünürlükte dikkatlice işlemek için giriş görüntüsünü düşük çözünürlükte değerlendirmektedir. Algılamayı hızlandırmak ve sınırlayıcı kutu kalitesini iyileştirmek için kademeli olarak kalibrasyon ağı eklenmiştir. Önerilen dedektör çok hızlıdır, CPU'daki tipik VGA görüntüleri için 14 FPS'ye ulaşır ve GPU'da 100 FPS'ye kadar hızlandırılabilir.

Liu ve arkadaşları (2015) vahşi doğada yüz nitelik tahmini için yeni bir derin öğrenme çerçevesi (framework) önermiştir. Ağ iki farklı kaskad CNN modelinden oluşmaktadır. LNET ve ANET basamakları önceden farklı şekilde eğitilmiştir. LNET yüz lokasyon tahmini için eğitilmişken ANET özellik tahmini için büyük yüz kimlikleriyle eğitilmiştir. Model tasarlanmış ön eğitim stratejileri ile arka plandaki karışıklıklara karşı dayanıklıdır. Gereksiz hesaplardan tasarruf etmek için hızlı ileri besleme algoritması kullanılmıştır.

Yang ve arkadaşları (2015) FDDB, PASCAL Face ve AFW'de olağanüstü performans sağlayan yeni bir derin konvolüsyonel ağ öneriyor. Spesifik olarak, yöntemleri zorlu FDDB kıyaslamasında %90,99 gibi yüksek bir geri çağırma oranına ulaşmaktadır. Daha da önemlisi, yüz bölümlerinin tepkilerini uzaysal yapıları ve düzenlemelerine göre puanlayarak yeni bir perspektiften yüz algılamayı sağlamaktadır. Bu ciddi tıkanmaya sahip ve kısıtlanmamış poz varyasyonu altında yüzlerin algılamasına olanak tanır.

Yang ve arkadaşları (2016) mevcut veri kümelerinden 10 kat daha büyük olan WIDER FACE veri kümesini sunmuşlardır. Veri kümesi, tıkamalar, pozlar, olay kategorileri ve yüz sınırlayıcı kutular dahil olmak üzere zengin açıklamalar içerir. Ölçek, poz ve tıkanmadaki büyük farklılıklar nedeniyle önerilen veri kümesindeki yüzler son derece zordur. Bu faktörler birçok gerçek dünya uygulamasında her yerde bulunur. Dolayısıyla bu veri seti ile eğitilen modeller gerçek dünyada oldukça iyi performans sergilemektedir.

Ranjan ve arkadaşları (2017) Derin evrişimli sinir ağlarını (CNN) kullanarak eşzamanlı yüz algılama, önemli noktaların yerleştirilmesi, poz tahmini ve cinsiyet tanıma için bir algoritma önermiştir. HyperFace adı verilen önerilen yöntem, derin bir CNN'nin ara katmanlarını ayrı bir CNN kullanarak birleştirir ve ardından birleştirilmiş özellikler üzerinde çalışan çok görevli bir öğrenme algoritmasını (multi-task learning) izler. Kapsamlı deneyler, önerilen modellerin hem küresel hem de yerel bilgileri yüzlerde yakalayabildiğini ve bu dört görevin her biri için birçok rakip algoritmadan önemli ölçüde daha iyi performans sergilediğini göstermektedir.

Jiang ve Learnend-Miller (2017), yakın zamanda çeşitli nesne algılama kriterlerinde etkileyici sonuçlar sergileyen Faster RCNN'i yüz algılamaya uygulamayı önermiştir. Daha hızlı bir R-CNN modelini geniş ölçekli WIDER yüz veri kümesinde eğiterek, WIDER test setindeki en son sonuçlar ile kıyaslamışlardır. Yapılan çalışma neticesinde Faster R-CNN genel nesne algılama için tasarlanmış olmasına rağmen uygun bir yüz algılama eğitim setinde yeniden eğitildiğinde etkileyici bir yüz algılama performansı gösterdiği sonucuna varılmıştır.

Sun ve arkadaşları (2017) derin öğrenmeyi kullanan yeni bir yüz algılama şeması önermiştir. Özellikle, zorlu negatif madencilik, özellik birleştirme, çok ölçekli eğitim, model ön eğitimi (pre-training) ve anahtar parametrelerin doğru kalibrasyonu dahil olmak üzere bir dizi stratejiyi birleştirerek son teknoloji Hızlı RCNN çerçevesi geliştirilmiştir. Sonuç olarak önerilen yöntemin son teknoloji yüz algılama performansını elde ettiği

gözlemlenmiş ve FDDDB kıyaslamasında yayınlanan yöntemlerin ROC eğrileri açısından en iyi modellerden biri olarak sıralanmıştır.

Yu ve arkadaşları (2021), yaptıkları çalışmada işlem maliyetini azalttığı ve hızlı olduğu için Yolov4 e dayalı yeni bir yüz maskesi algılama algoritması kullanma önermişlerdir. Pandemiyle birlikte maske takılması zorunlu olan bölgelerde maske takıp takmama durumu manuel olarak kontrol edilmekte olup hem zaman kaybına hem de kaynak israfına sebep olmaktadır. Bu sebeple yazarlar geliştirilmiş oldukları özel yolov4 nesne algılama algoritmasını önererek kontrol işleminin daha pratik ve verimli bir şekilde gerçekleştirilmesi sağlanmaktadır. Önerilen yeni yolov4 yapısında CSPDarkNet53, CSP1_X ve CSP2_X olarak geliştirilmiş, özellik çıkarma parametrelerini azaltmak için ağ modülleri azaltılmış ve görüntü ölçekleme yönteminin yerini almak için uyarlanabilir görüntü ölçekleme yöntemi kullanılmıştır. Elde edilen deneysel sonuçlar önerilen yöntemin zorlu maske algılama görevleri için diğer algılayıcılara kıyasla en yüksek algılama performansına sahip olduğunu göstermektedir.

Sun ve arkadaşları (2017) bulaşıcı hastalıkları olan hastaları taramak için IRT ile donatılmış bir CMOS kamera kullanan birleşik bir görünür ve termal görüntü işleme yaklaşımının uygulamasını önerdi. Görünür ve termal görüntüler üreten bir IRT sistemi, görüntü elde etmek için kullanıldı. Deneklerin solunum hızları, termal görüntülerde nazal alanların etrafındaki sıcaklık değişiklikleri izlenerek ölçüldü; aynı anda yüz cilt sıcaklıkları ölçüldü. Sonuç olarak önerilen sistem şüpheli bulaşıcı hastalıkları olan hastaları verimli bir şekilde tespit etti.

Somboonkaew ve arkadaşları (2017), kızıl ötesi temassız ateş ölçerlere alternatif olarak çift sensörlü termal görüntüleme cihazları üzerinden otomatik alın ateşi taraması için bir mobil platform önermiştir. Kamera üzerine yerleşik yüz algılama algoritması kullanarak sırasıyla görüntü hizalama, hedef yerleştirme ve sıcaklık ölçümü işlemleri gerçekleştirilmektedir. Gerçek tabanlı uygulamalar için düşük kare hızına sahip olan sistem, yapılan testler sonucunda, yüksek ateşli hastaları taramak için %100 duyarlılık ve %70 özgüllük göstermiştir.

Li ve arkadaşları (2018) kızılötesi termografi kullanarak gerçek zamanlı termal konfor yorumu için yeni bir çerçeve sunmuştur. Bu çalışmanın ana katkısı, yüz cilt sıcaklığı verilerini kesintisiz ve otomatik olarak elde etmek, almak ve analiz etmek ve gerçek operasyonel ortamlarda her bina kullanıcısı için termal konfor koşullarını yorumlamak için önerilen veri toplama ve analiz çerçevesidir. Önerilen çerçeve, termoregülasyon teorisi, bilgisayar vizyonu ve makine öğrenimi dahil olmak üzere

disiplinler arası teknikleri kullanır. Sonuç olarak, müdahaleci olmayan düşük maliyetli kızılötesi termal kameralardan toplanan yüz cildi sıcaklığının gerçek zamanlı olarak sağlam bir termal konfor tahmini elde edilmesine yardımcı olabileceğini ve bina işlerinde minimum kesinti ile iç mekân ortamlarının senkron kontrolü olasılığını sunduğunu göstermektedir.

Mohammed ve arkadaşları (2020), ateşin yaygın bir semptom olması nedeniyle IoT'nin ateş taraması için kullanılmasını önermiştir, çünkü IoT kullanılarak teşhis daha az insan etkileşimli hale getirilebilir. Termometre tabancalarının kullanılması, ateşi kontrol eden kişiyi de yüksek riske maruz bıraktığından savunmasız hale getirir, ayrıca toplu tarama için uygun değildir. Bu nedenle, insanları ateş için taramak için drone kullanımı tasvir edilmiştir. Bu çalışmada, şüpheli bölgeye termal kamera takılı olarak drone kamera gönderilir ve şüpheli bölge termal kamera yardımı ile taranarak vücut ısısı yüksek olan bir kişi olup olmadığı kontrol edilir. Yapay zekâ ile desteklenmiş bu sistemin kullanımı pozitif vakaların tespitini kolaylaştırarak sağlık birimlerinin talebini karşılamaktadır.

Maghdid ve arkadaşları (2020) sadece sensörleri kullanarak bir akıllı telefona dayalı Covid-19'u tespit etmek için yeni bir çerçeve önerdiler. Bu çerçeve, yerleşik sensöre dayalı olarak normal ateşi Covid-19 ateş seviyesi ile ayırt etmeye çalışır. Buna ek olarak koronavirüs hastalığında var olan semptomları tespit etmek için ayarlanmış bir dizi sensör kullanılır. Önerilen yöntemde sensör verilerini okumak için giriş ve sensör konfigürasyon, hastalığı belirlemek içinse semptomları analiz etme ve tahmin etme katmanları olmak üzere toplamda dört ayrı katman oluşturulmuştur. CNN tabanlı bir sınıflandırıcı eğiterek bu dört katmandan gelen verilere göre kişinin pozitif veya negatif olduğu tespit edilmeye çalışılmıştır.

Wang ve arkadaşları (2020) termal görüntüleme ve derin öğrenme mimarisine dayalı temassız bir sağlık sistemi önerdi. GoogLeNet derin öğrenme modelini kullanılan deneyde, eğitim ve doğrulama için 800 görüntü ve test için 200 görüntü kullanılmıştır. Termal görüntüler FLIR-One kullanılarak yakalanmıştır. Yapılan testler sonucunda, bu termal görüntülemeye dayalı yeni özellik çıkarma yöntemiyle bireylerin uzun vadeli sağlık takibinin gerçekleştirilebileceği sonucuna varılmıştır.

Lun ve arkadaşları (2021) insan trafiğinin yoğun olduğu alanlar için uzaktan, QR kod tanımlama, kızılötesi termal vücut sıcaklığı izleme, uyarı sinyali ve veri yönetimi ile otomatikleştirilmiş, etkin bir uçtan uca sıcaklık izleme mekanizması önermişlerdir. Raspberry Pi üzerinde oturttukları bu sistem de 36.5 ve üstünü yüksek ateş olarak

değerlendirmişler. Minimum entegrasyon çabasıyla geleceğin donanımlarını desteklemek için yapılandırılabilen yenilikçi ve yüksek düzeyde ölçeklenebilir bir çözüm sunduklarını savunarak geliştirdikleri yapıya yüz tanıma, makine öğrenimi ve yapay zekâ geliştirmelerini entegre ederek daha etkili bir çözüm sunabileceklerini belirtmişlerdir.

Gupta ve arkadaşları (2021), Covid-19'un yayılımını engellemek için yapay zekâ tabanlı ateş ölçüm sistemiyle personel girişlerinin kontrolünün sağlanmasını önermişlerdir. İş yeri giriş çıkışlarına konumlandırılan bu sistem ilk olarak kişinin yüzünü algılayarak ateşini ölçer ve ardından elde edilen sıcaklık değeri eşik değerin üstündeysen personelini geçişine izin verilmez. Böylece çalışma ortamındaki diğer kişilerin Covid-19'a yakalanma riski düşürülmüş olur.

Ulleri ve arkadaşları (2021), maske algılama ve temassız ateş ölçümlü çalışan yönetim sistemi önermişlerdir. Bu sistemde, yüz tanıma işlemleri 5 MP bir CMOS kamera ve sıcaklık ölçümü bir kızıl ötesi termal kamera üzerinden gerçekleştirilerek iki farklı tür kamera kullanılmıştır. Personel devamlılığının yönetimi için bir API oluşturmuş ve yüz tanıma sisteminin beslenmesi için bir veri tabanı yapış kurulmuştur. Ayrıca yüz algılama ve maske tespiti için Tensorflow-OpenCV tabanlı bir derin öğrenme algoritması kullanılmıştır. Termal sıcaklık ölçümü içinse RGB görüntü üzerinden algılanan yüz termal görüntü üzerinde eşleştirilmiş ve bir görüntü işleme algoritması kullanılarak sıcaklık değerleri belirlenmiştir. Sistem ölçüm bölgesi olarak alını referans almaktadır. Yapılan testler sonucunda önerilen algoritmanın salgınla mücadelede etkin olduğu gözlemlenmiştir.

3. MATERYAL VE YÖNTEM

Bu tez çalışmasında Python programlama dili kullanılmıştır. Python yapay zekâ uygulamaları üzerinde popüler olarak kullanılmaktadır. Bu dil Cafe, Pytorch, Tensorflow gibi birçok derin öğrenme çerçevesini destekleme birlikte önerilen sistem üzerinde OpenCV-Cafe modülü tercih edilmiştir.

3.1. İnsan Beyni ve Görü Sistemi

Yapay sinir ağlarının çalışma mekanizmasının insan beyninin yapısını ve işleyişini baz aldığı bilinmektedir. Bu durumda yapay sinir ağlarını tam anlamıyla kavrayabilmek için beyin yapısına ve işleyişine de hâkim olmak önemlidir. Beyin, sinir sisteminin en önemli organıdır. Vücuttaki diğer doku ve organların kontrolünü sağlarken, birçok fonksiyonun yönetiminde de rol oynar.

Sinir sistemi vücudun en karmaşık sistemlerinden biridir. Beyin, omurilik ve ilişkili tüm sinir ve duyu organlarından oluşmaktadır. Uyarılar duyu organları aracılığıyla elde edilir, beyinde saklanır ve işlenir.

Beyin glia adı verilen beyin hücrelerini ve nöron adı verilen sinir hücrelerini kapsamaktadır. Nöronlar görevleri bakımından üç başlık altında incelenebilir. Bunlar duysal nöronlar, motor nöronlar ve internöronlardır.

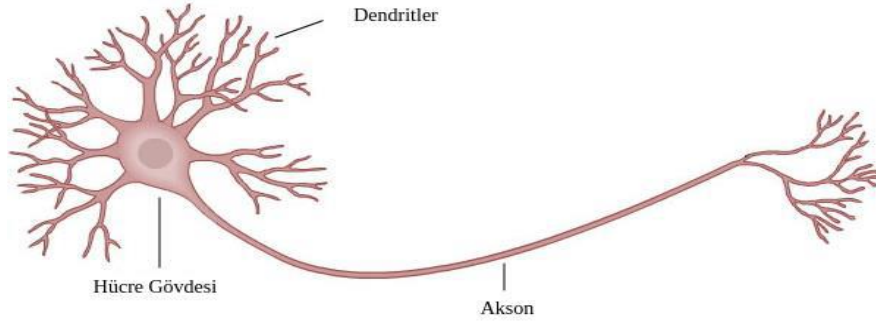
Duysal nöronlar doku ve organlardan gelen sinir uyarılarını beyne taşır. Duysal nöronlar daha çok içsel ve dışsal faktörlere tepki oluşturabilmek için özelleşmiştir. Temelde tat alma, koklama, işitme, dokunma, görme duyularının alınmasını sağlar.

Motor nöronlar beyinden gelen uyarıları ilgili doku ve organlara taşır. Motor nöronlar kas aktivitelerini kontrol eder, istemli hareketlerden sorumludur.

İternöronlar, duysal nöronlar ve motor nöronlar arasındaki bağlantıyı sağlar. İnsan beyinde sayısı en fazla olan nöron çeşididir. Öğrenme, muhakeme gibi daha çok bilişsel ve karmaşık süreçleri yönetir.

Sinir sistemi vücudun elektrokimyasal iletişim ağıdır. İnsan beyinde yaklaşık 100 milyar nöron ve 3,2 milyon km sinaptik bağlantı bulunmaktadır.

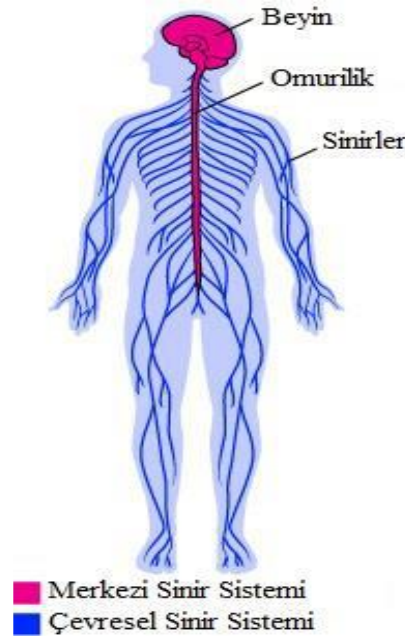
Nöronlar sinir sisteminin yapıtaşlarıdır. Nöronlar üç kısımdan oluşur. Bunlar hücre gövdesi, elektriksel uyarı sinyallerini (impuls) hücre gövdesinden bir sonraki nörona ileten tek bir uzun akson ve diğer hücrelerden bu uyarıları alan ve hücre gövdesine ileten birkaç dallı dentritlerdir. Şekil 3.1.'de nöronun modellenmiş biçimi gösterilmiştir.



Şekil 3.1. Biyolojik sinir ağı modeli

Nöronlar birbirleriyle fiziksel temas halinde değildir, ancak birbirlerine o kadar yakınlardır ki elektrokimyasal uyarılar sinaptik boşluklara kolaylıkla atlayabilirler. Sinaptik boşluklar, bir nöronun aksonunun bir sonrakinin dentritiyle bağlantı kurduğu, boşluk gibi görünmesine rağmen oldukça aktif kullanılan küçük bölgelerdir.

Beyin ve omurilik merkezi sinir sistemini oluşturur. Merkezi sinir sistemi sinirler yoluyla vücudun geri kalanıyla iletişim kurar. Gelen tüm uyarılar ve giden tüm sinyaller merkezi sinir sistemi aracılığıyla işlenir. Merkezi sinir sistemine uyarıları getiren ise çevresel sinir sistemidir. Çevresel sinir sistemi organlardan gelen sinyalleri toplayarak merkezi sinir sistemine iletmekle görevlidir. Kısacası organlar ve merkezi sinir sistemi arasındaki köprüdür. Çevresel sinir sistemi kalp atış hızı, nefes alıp verme gibi hayati fonksiyonlar üzerinde rol oynamaktadır.



Şekil 3.2. Biyolojik sinir sistemi

Beyin dört loptan oluşmaktadır. Bunlar frontal, parietal, oksipital ve temporal loptur. Oksipital lop, beynin görme fonksiyonu ile özelleşmiş bölümüdür. Kabaca bir hesaplama beynin yaklaşık dörtte birinin görme için ayrıldığı sonucuna ulaşılabılır. Bu da diğer fonksiyonlar ve duyuyla birlikte düşünüldüğünde görme duyusunu önemini gözler önüne sermektedir.

Görme sanıldığından daha karmaşık bir süreçtir. Görme duyusu organı gözdür. Gözler içinde bulunan çevreyi algılar ve insanın dış dünya ile bütünleşmesine olanak sağlar. Çevredeki cisimlerden yansıyan ışıklar göze gelen ışıklar görme sürecini başlatır. Işık, göze korneaya aracılığıyla girer. Kornea saydam, parlak bir görünüme sahiptir ve havayla temas halindedir. Kornea sinirler yönünden oldukça zengindir. Işık buradan odaklanmayı ayarlayan lense geçer. Lens, göz üzerine düşen ışınları birbirine yaklaştırdıktan sonra kırar. Cisme ait olan tam ve gerçek görüntüyü retina üzerine düşürür. Retina tabakasının sarı noktası üzerine ters bir biçimde düşen görüntü, retinada bulunan fotoreseptörler aracılığıyla absorbe edilir ve görme sinirleri üzerinden beyinde görmeyle özelleşmiş alana iletilir. Görüntü beyinde düzeltilir ve cisimler düz biçimde algılanır. Sonuçta ise görme gerçekleşmiş olur.

Göze giren ışık miktarını göz bebeği ayarlar, göz bebeği ise iris tarafından kontrol edilir. Yakındaki ve uzaktaki nesnelere odaklanılabilmesi için lenste birtakım şekil değişiklikleri yapılır, bu şekil değişiklikleri irisin arkasında bulunan kaslar tarafından kontrol edilir.

Fotoreseptörler, üzerine düşen ışık enerjisini elektrik enerjisine çevirir. İki tip fotoreseptör vardır. Bunlar çubuk ve koni hücreleridir. Çubuk hücreleri yalnızca ışığı algılar, düşük yoğunluklu ışıklara bile duyarlı olacak kadar özelleşmişlerdir. Koni hücreleri ise nesnelere rengini ayırt etme görevini üstlenirler, renkli görmeyi sağlarlar. İnsan gözünde yaklaşık olarak 125 milyon fotoreseptör bulunmaktadır.

Görmede son derece önemli bir işleve sahip olan retinayı yapay sinir ağlarıyla eşleştirecek olarak şemalaştırmak gerekirse üç aşamaya yer vermek doğru olacaktır. Birinci katman sinyallerin alındığı ve bir üst katmana iletiildiği çubuk ve koni hücrelerini içerir. Sinyaller ikinci katmandan üçüncü katmana iletilir. İletilen bu son katmanda sinyaller görsel kortekse ulaşır. Son katmandaki optik sinirin ganglionları elektriksel uyarı sinyallerini hücre gövdesinden bir sonraki nörona ileten aksonları oluşturur. Kabaca üç katmandan oluşan bu şema karmaşık süreçleri anlamak ve çözmek için olanak sağlar.

3.2. Yapay Sinir Ağları

Daha önce de bahsedildiği gibi yapay sinir ağları insan beyninin yapısını ve işleyişini kendisine temel almaktadır. İnsan beyniyle doğrudan bağlantılı olan ve kökeni 1940'lara kadar uzanan YSA, yapay nöronları ve nöron gruplarından oluşan katmanları bünyesinde barındırmaktadır. YSA bünyesindeki bu katmanlar tek bir katmandan ibaret olabileceği gibi birden fazla sayıda katman da içerebilmektedir.

İnsanların temel fonksiyonlarına örnek olabilecek işitme, duyma, konuşma gibi fonksiyonlar insanlar için oldukça basit gibi görünse de bilgisayar algoritmasının bu süreçleri algılaması, çözümlenmesi ve işlememesi o kadar da kolay değildir. YSA bahsi geçen bu karmaşık süreçleri anlamlı hale getirmek için olanak sağlamaktadır. YSA, kendisine sunulan girdileri çözümler, bünyesinde biriktirir ve sonuçta önemli deneyimler edinir. Elde ettiği bu deneyimleri ise daha sonrasında istenilen herhangi bir problemi çözme sırasında aktif olarak kullanabilir.

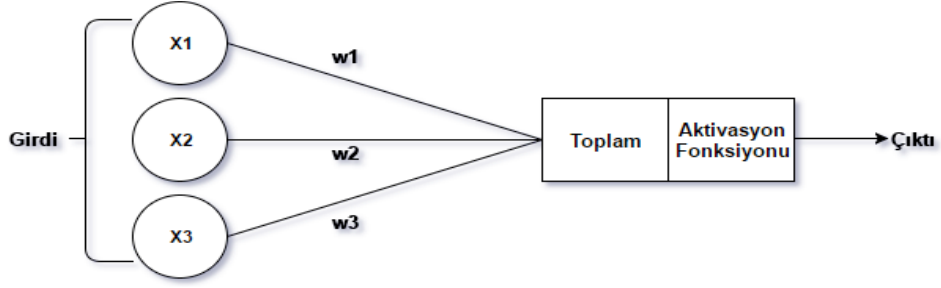
3.2.1. Tekli nöron modeli

Sinir hücreleri daha önce nöron olarak isimlendirilmişti. Nöron, en basit hesaplama elemanıdır. Giriş değerlerini doğrusal biçimde alıp onlara uygun çıkış değeri oluşturur. Şekil 3.3.'te gösterildiği biçimde giriş değerleri ile önceden belirlenmiş olan ağırlık değerleri çarpma işlemine girer. Bu sonuca eğer varsa nörona ait bias değeri de eklenir, sonuçta toplam değere ulaşılmış olur. Herhangi bir katmanda bulunan bütün nöronların sayısı n olarak belirlenmek üzere çıktı fonksiyonu;

$$T(x) = \sum_{i=1}^3 w_i x_i + b \quad (3.1)$$

$$T(x) = w_1 x_1 + w_2 x_2 + w_3 x_3 + b \quad (3.2)$$

biçiminde formüleleştirilebilir. Ulaşılmış olan toplam değer, aktivasyon fonksiyonuna giriş değeri olarak sunulur ve çıkış değerine göre aktif olur.



Şekil 3.3. Tek nöronun çalışma yapısı

YSA' da ağırlık ve bias değerleri eğitim aşamasına göre güncellenen değişkenlerdir. Nöronlara ait olan bias değerlerinin kullanılması ile aktivasyon fonksiyonunun eksenindeki hareketi amaçlanır. Ağın farklı durumları temsil etmesi bias değeri kullanılarak sağlanabilir.

3.2.2. Aktivasyon fonksiyonları

Diğer ismi Doğrusal Olmayan Fonksiyon olarak da bilinen aktivasyon fonksiyonları, nöronlara doğrusal olmayan bir form vererek modellerin karmaşıklık seviyesini artırma işlevini üstlenir. YSA bu karmaşık yapıları çözümlenerek daha başarılı sonuçlar elde edebilme fırsatı bulur. Dolayısıyla gerçek dünyadaki karmaşık problemlere daha etkili çözümler getirebilme becerisi kazanmış olur. Şayet aktivasyon fonksiyonunu seçerken doğrusal fonksiyon tercih edilecekse ağın doğrusal olarak çözülebilecek bir nitelik taşıması gerekmektedir.

Aktivasyon fonksiyonuna göre nöronu aktive edebilmek için gereken eşik değerinin altında kaldığında, nöronlar aktive olmazlar. Dolayısıyla aktive olamayan nöronlar bu süreçte işlevsel bir yere sahip olamazlar.

Literatürde birbirinden farklı aktivasyon fonksiyonlarına rastlamak mümkündür. Bunlardan en yaygın kullanılanlara örnek vermek gerekirse ReLU (Rectified Linear Unit), Sigmoid, Tanh gibi fonksiyonlardan bahsetmek doğru olacaktır.

3.2.3. İleri beslemeli sinir ağlar

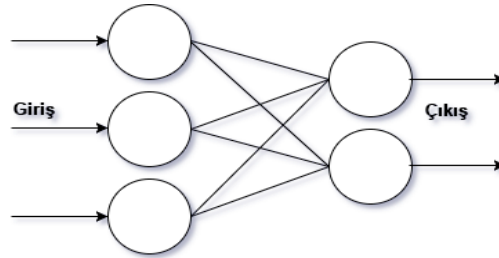
İleri Beslemeli Ağlar, temel olarak üç katmandan oluşmaktadır. Bu katmanlar giriş katmanı, gizli katman ve çıkış katmanıdır. Her katman farklı sayıda nöron gruplarını bünyesinde barındırabilir. Nöronların görevi bilgiyi ileri katmanlara doğru iletmektir.

Birbirleriyle ağırlık değerini içeren bir bağlantı kurmuş olan bu nöronlar düğüm olarak adlandırılmaktadır. Bilgi iletiminin ileri katmanlara doğru tek yönlü yapıldığı hatırlanacak olursa aynı katmanda bulunan düğümlerin birbiriyle bağlantı halinde olmadığından ve bu ağda bir döngü bulunmadığından rahatlıkla söz edilebilir.

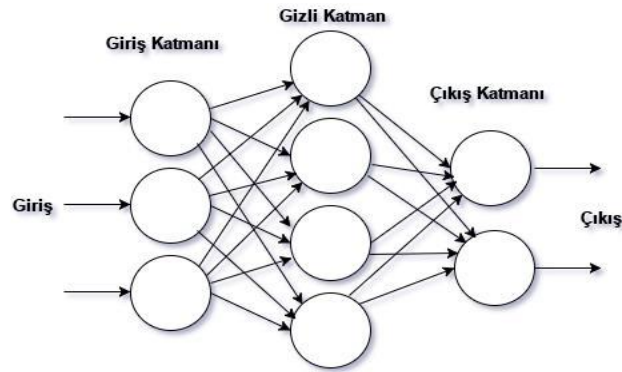
Giriş katmanında öncelikli olarak düğüm sayısı belirlenir. Belirlenen bu sayıya göre veri girdisi sağlanır. Gizli katman sıralama olarak giriş katmanı ve çıkış katmanı arasında yer almaktadır, isteğe bağlı olarak eklenme seçeneği mevcuttur. Ağ derinleştirilmek istenirse, gizli katman sayısında artışa gidilebilir. Girdi değerleri ile birtakım hesaplamalar yapıldıktan sonra elde edilen değerler ileri yönlü bir akış sağlanarak çıkış katmanına ulaşmış olur.

Yönlü düz çizgeleri (directed acyclic graph) takip etmek, ileri beslemeli sinir ağları bünyesinde bulunan fonksiyonların birbiriyle nasıl uyum içinde çalıştıklarını kavrayabilmek adına izlenmesi gereken bir adımdır.

İleri beslemeli ağlar modellenmek istenirse, iki alt modelden söz etmek mümkündür. Bunlar tek katmanlı algılayıcı modeli ve çok katmanlı algılayıcı modelidir. Şekil 3.4. ile tek katmanlı algılayıcı modeline, Şekil 3.5. ile çok katmanlı algılayıcı modeline yer verilmiştir.



Şekil 3.4. Tek katmanlı algılayıcı modeli

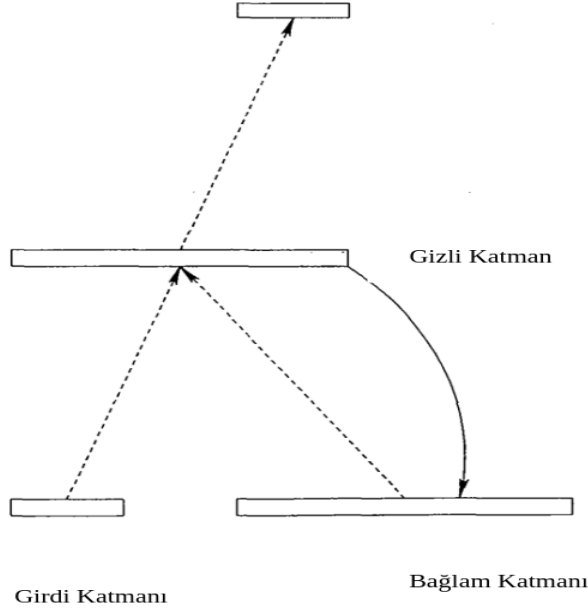


Şekil 3.5. Çok katmanlı algılayıcı modeli

3.2.4. Geri beslemeli sinir ağıları

Geri beslemeli YSA'da, ileri beslemeli YSA'dan farklı olarak nöronlar bulunduğu katmandan sonra gelen katmandaki nöronlara tek yön içeren belirli bir sıra halinde bağlanmazlar. Yani katman kendinden önceki katmanın çıktısını girdi olarak alabileceği gibi, aynı zamanda kendinden sonraki katmanın çıktısını da girdi olarak alabilmektedir. Diğer adı tekrarlı ağılar olarak da bilinen geri beslemeli ağılar, genellikle zaman serisi gibi sıralı problemlerin çözümünde kullanılırlar. Örneğin, yazı tahmin sistemlerinde girilen metne karşılık olarak sonraki karakter tahmin edilmeye çalışılır.

Oldukça güçlü olan bu geri beslemeli ağılar, sinyalleri her iki yönde de taşıyabilir ve son derece karmaşık bir yapıya sahip olabilirler. Giriş değişinceye kadar denge noktasında kalan geri beslemeli sinir ağıları, girişteki değişikliklere bağlı olarak durumları dengelemek için dinamik bir davranış gösterirler. Şekil 3.6.'da geri beslemeli yapay sinir ağı gösterilmiştir.



Şekil 3.6. Geri beslemeli sinir ağı modeli

İleri beslemeli yapay sinir ağlarına eklenen bir bağlam katmanı ile kurulmuş yapıya sahip olan Elman ağı, 1990 yılında Jeffrey L. Elman tarafından ortaya atılmıştır. Bahsi geçen ağ, basit tekrarlı bir ağ modeline sahiptir. Ağırlık değerleri sabit olacak şekilde, gizli katman yönünden bağlam katmanı yönüne doğru aktivasyonlar tamamen

aynı olacak şekilde çoğaltılır. Şekil 3.6.' da gösterilen kesik çizgili kısımlar öğrenilmesi mümkün bağlantılara karşılık gelmektedir. Ağda belirlenen girdiye uygun olacak doğru çıktıyı elde edebilmek için kullanılan kalıpları, gizli birimler kodlamaktadır. Yaptığı kodlamayla girdinin kalıbına uygun olacak içsel durumları oluşturmakla görevlidir. Önceki içsel durumu hatırlamak bağlam birimleri aracılığıyla sağlar. Gizli katmanda bulunan bu bağlam birimleri dışsal girdiyi haritalama görevinin yanı sıra önceden belirlenmiş birtakım çıktılarının ilk durumlarını haritalama görevini de üstlenmektedir. İçsel temsillerin zamansal bağlama duyarlı olabilmesi, gizli birimlerde kurulan örüntülerin bağlam olarak saklanmasıyla mümkün olur. Bahsedilen bu model bir sıra halinde sunulmuş olan girdilerin gerekli işlemlere tâbi tutulmasını sağlar. Böyle modeller hava durumunu tahmin edebilen uygulamalarda kullanılabilir (McClelland, 2015).

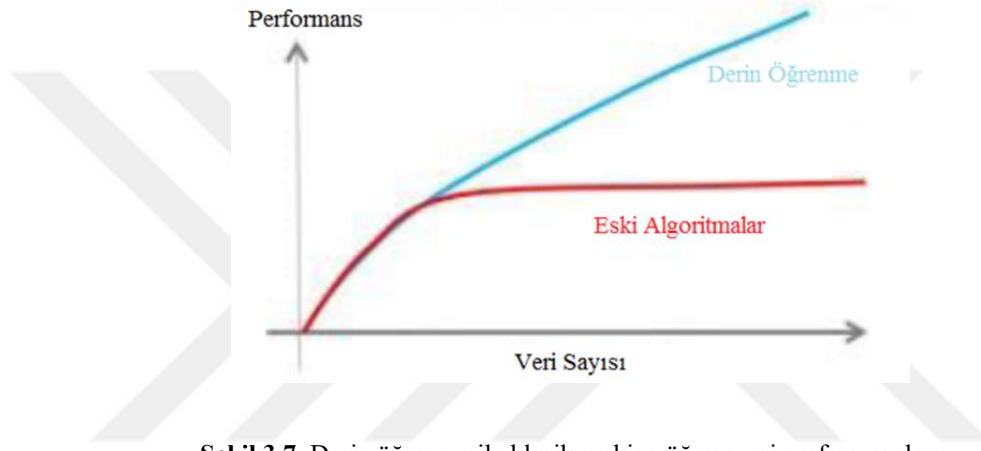
3.3. Derin Öğrenme

YSA hesaplamalarına dayanan derin öğrenme popüler bir makine öğrenimi yaklaşımıdır. YSA'nın derinleşmesiyle oluşan bu yöntem, YSA kıyasla daha iyi başarı elde eden bir yöntemdir. YSA eğitilmesi zor bir model olduğu için ve geçmiş yıllarda kısıtlı olan donanım olanakları sebebiyle modellerdeki katman sayıları az sayıda seçilmekteydi. Daha derin modellerin eğitilme olanağı, işlem gücü yüksek makinelerin hayatımıza girmesiyle birlikte gerçekleşmiştir. Özellikle GPU'nun matris işlemlerini hızlı yapabilmesi sebebiyle modeller GPU'lar üzerinde eğitime başlanmıştır. Bu durum derinleştirilmiş modellerin CPU'ya görece çok daha kısa bir zaman diliminde eğitilebilmesiyle sonuçlanmıştır.

Derin öğrenmenin aktif biçimde kullanılmasını sağlayan önemli faktörlerin başında şunlar gelebilir: İşlemci becerisinin büyük oranda artmış olması (GPU gibi), eğitim sürecine dahil edilen veri miktarının artırılması. Verilerden deneyim kazanılan derin öğrenmede, bilgisayarların sorun çözmesini sağlama konusunda gerçek dünyayı yansıtacak verilerin önemi yadsınamaz.

YSA'nın doğrusal olmayan problemleri çözebilme becerisi, derin öğrenmenin problem çözme becerisine zemin hazırlamaktadır. Karmaşık problemlerin çözümünü kolaylaştıran, modelin derinliğine paralel olarak artan katman sayısıdır. Bunların yanı sıra derin öğrenme modelleri, minimum değişiklik ile farklı uygulama alanlarındaki problemlerin çözümüne geniş bir perspektif sunar.

Derin öğrenmede, klasik makine öğrenmesi tekniklerinde olduğu gibi özellik mühendisliği çalışmalarına ihtiyaç duyulmamakta, dolayısıyla verilerden otomatik bir şekilde özellik çıkarımı yapılmaktadır. Bu da derin öğrenmenin en önemli özelliklerindedir. Konvolüzyonel sinir ağları, görüntü sınıflandırma problemlerinde yoğun bir biçimde kullanılmaktadır. Görüntüde bulunan özellikleri eşleyebilmek için ağ görüntüyü analiz eder. Aslında karmaşık bir sahneyi analiz eden bu eğitilmiş ağ, görüntüdeki nesnenin sınırlı bir alanını görmüş olsa bile özellik eşleme yardımıyla bahsi geçen parçanın hangi nesneye ait olduğunu ortaya koyabilmektedir.



3.3.1. Konvolüzyonel sinir ağları

KSA, iki alt bölüme ayrılmaktadır. Birinci bölümde girdi sunulduktan sonra sayısı bir ya da birden daha fazla olan konvolüsyon ve havuzlama katmanlarını içeren özellik çıkarımı yapılmaktadır. Sonraki bölümdeyse tam bağlı katmanları içeren sınıflandırma işlemini yapılarak çıktıya ulaşılmaktadır. Şekil 3.8.' de KSA adımları şema halinde gösterilmiştir.



Şekil 3.8. KSA aşamaları

Çoğunlukla görüntüyü sınıflandırmak için kullanılan bu ağlar, işleyiş olarak insanın görme mekanizması gibi belirli bir özelliğin çıkarımlarıyla öngöründe

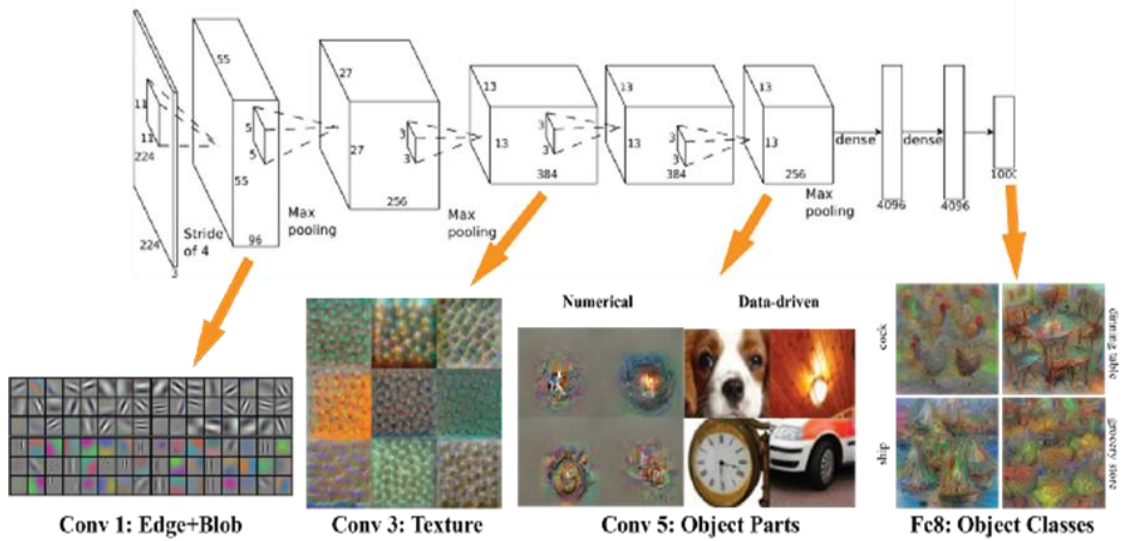
bulunmaktadır. Görüntü sınıflandırma problemlerinde süreci asıl zorlaştıran görüntüye ait çok sayıda piksel içinden istenilen özelliğin ayrıştırılmasıdır. KSA zor olan bu süreci ilk bölümde otomatik bir şekilde işletir ve özellik haritalarına ulaşır. Sonrasında ise özelliklere dair ulaşılan bu bilgiler tam bağlı katmanlara iletilir. Belirli sayı değerlerine çok katmanlı algılayıcılar aracılığıyla ulaşılır. Bu değerler yardımıyla görüntünün ait olduğu sınıf ile ilgili öngöründe bulunulur.

Bilgisayarlı görü uygulamalarında KSA görüntüye ait piksel değerleri girdi olarak kullanılır. Görüntünün içeriği renk kanal sayısına bağlı olarak değişen giriş katmanındaki nöron sayısı, gri tonlamalı tek kanallı bir görüntüde bir iken, renkli bir görüntüde üçtür.

Girdi olarak alınan görüntünün pikselleri daha sonra tensör olarak adlandırılan çok boyutlu matrislere dönüştürülür. Örneğin yükseklik ve genişlik değeri 640 olan bir renkli görüntünün tensör boyutu $1 \times 640 \times 640$ olarak belirlenerek matris işlemlerini gerçekleştirmesi sağlanır.

3.3.1.1. Özellik çıkarım aşaması

KSA'nın özellik çıkarımı sırasında birbirini takip eden bir veya birden fazla konvolüsyon, havuzlama ve Relu işlemi uygulanır. Bu aşamanın sonunda elde edilen özellik haritaları, bir görüntüden elde edilen özelliklerin farklı ölçeklerde temsilidir. 3.9.'da ağır özellik haritalarında belirli bir görüntünün nasıl gördüğü gösterilmektedir.



Şekil 3.9. VGG görselleştirilmiş özellik haritaları

3.3.1.1.1. Konvolüsyon katmanı

Bir görüntünün piksel matrislerini kullanarak özellik haritalarını çıkarmak, konvolüsyon katmanlarının bir amacıdır. Bu katmanda girdi olarak alınan görüntünün piksel matrisleri üzerine başlangıç noktası görüntünün sol üst köşesi olacak şekilde soldan sağa adım sayısı kadar filtreler uygulanır. Görüntü matrisinin en sağ noktasına ulaşan filtreler, yine adım sayısı kadar aşağı kaydırılır ve bu süreç görüntünün sağ alt köşesine kadar devam eder. Genellikle 3x3, 5x5 gibi birbirinden farklı boyutlarda olan ağırlıklar her kaydırma işleminde filtre ile aynı boyuta seçilen piksel matrisi ile çarpılır. Sonrasında elde edilen matris elemanlarının değerleri toplanır ve sonuca ulaşılır. Bu işlemler son piksel değerine kadar sürekli tekrarlanır. Böylece özellik haritalarına ulaşılmış olur.

Giriş katmanından hemen sonraki katmanda yapılan filtreleme miktarı görüntünün derinliğine denktir ve düşük seviyede özellik çıkarımı yapılmasını sağlamaktadır. Sonraki konvolüsyon katmanlarıysa görüntünün belirli kısımlarını içeren yüksek seviyeli özellik çıkarımı yapılmasını sağlamaktadır. İlk katmanda uygulanacak olan filtrenin boyutunu belirlemek için görüntüde odaklanılması gereken asıl özelliklerin ortalama boyutlarını saptamak, önemli bir faktör olabilmektedir. Filtreler aracılığıyla bir görüntünün farklı özelliklerine odaklanmış birden çok sayıda sürümüne ulaşılmaktadır. Ağın eğitimi ile öğrenilen ağırlıklar, uygulanan filtreleri meydana getirmektedir. KSA'yı diğer yapay sinir ağlarından ayıran asıl kısım, özellik çıkarımı yapılan kısımdır. Ağda bulunan konvolüsyon katmanı sayısı, yine o sayının çokluğuyla bağlantılı olarak karmaşık özelliklerin belirlenebilmesini sağlamaktadır.

Örneğin, üç kanallı 25x25 boyutundaki bir görüntüye ilk konvolüsyon katmanında 3x3 boyutunda filtreler uygulandığını düşünelim. Konvolüsyon işlemi tüm katmanlara başarı ile uygulandıktan sonra 23x23 boyutunda özellik matrisleri elde edilmiş olacaktır. Kaydırma işlemi sırasında adım(stride) ve boşluk(padding) gibi başka parametreler de dahil edilmektedir. Adım sayısı, filtre uygulama sırasında filtrenin hangi tarifelerle uygulanacağını belirtirken, boşluk ise kenar değerlerinin bahsedilen bu işleme dahil edilme sürecini belirtmektedir. Konvolüsyon katmanındaki çerçevelerin görüntü matrislerine uygulanma adımları aşağıdaki gibidir:

- 1) Görüntü matrislerine uygulanılması istenen filtrenin boyutu, boşluk ve adım sayısı belirlenir.

- 2) Görüntü matrisinin en üst sol köşesine seçilen değerlere uygun olarak konvolüsyon çerçevesi uygulanır.
- 3) Uygulanan çerçeve yerleştirildiği noktadan özellik matrisinin en sağına adım sayısı kadar kaydırarak uygulanır.
- 4) En sağa gelindikten sonra başa dönülür ve çerçeve, adım parametresi kadar aşağı kaydırılarak uygulanır.
- 5) Eşleşme sonlanana kadar üçüncü aşamaya geri dönülür.

0	0	0
1	1	0
0	1	0

Şekil 3.10. 3x3 boyutunda filtre

5x5 boyutunda bir piksel matrisine Şekil 3.10.' da görülen 3x3 boyutundaki filtre, bir adımlı konvolüsyon işlemi şeklinde görüldüğü gibi uygulanmaktadır. Bu işlem sırasında kullanılan numaralandırılmış matrisler filtrelerin görüntü pikselleri ile eşleştiği noktalarlardır. Çarpım işlemi sırasında eşleşen noktaların sayısal değerleri dikkate alınır ve hesaba katılır.

40	47	53	63	74	40	47	53	63	74	40	47	53	63	74
42	50	57	65	76	42	50	57	65	76	42	50	57	65	76
46	52	62	68	77	46	52	62	68	77	46	52	62	68	77
48	56	64	69	79	48	56	64	69	79	48	56	64	69	79
54	58	66	73	80	54	58	66	73	80	54	58	66	73	80
I	II	III												
40	47	53	63	74	40	47	53	63	74	40	47	53	63	74
42	50	57	65	76	42	50	57	65	76	42	50	57	65	76
46	52	62	68	77	46	52	62	68	77	46	52	62	68	77
48	56	64	69	79	48	56	64	69	79	48	56	64	69	79
54	58	66	73	80	54	58	66	73	80	54	58	66	73	80
IV	VI	VI												
40	47	53	63	74	40	47	53	63	74	40	47	53	63	74
42	50	57	65	76	42	50	57	65	76	42	50	57	65	76
46	52	62	68	77	46	52	62	68	77	46	52	62	68	77
48	56	64	69	79	48	56	64	69	79	48	56	64	69	79
54	58	66	73	80	54	58	66	73	80	54	58	66	73	80
VII	VIII	IX												
I	II	III	144	169	190									
IV	V	VI	154	178	199									
VII	VIII	IX	162	186	206									

Şekil 3.11. Bir adımlı konvolüsyon işlemi

Çarpım işlemi sırasında piksel matrisine ait hacim küçülme olmaktadır. Eğer hacmin aynı kalması istenirse, uygun sayıda sıfır boşluk (zero padding) yöntemi, adım sayısı ve filtre boyutu uygulanmalıdır. $K \times K$ boyutunda görüntü matrisi üzerinde L sayıda sıfır boşluk değeri seçildiğinde görüntüye ait piksellerin kenarlarına ve köşe noktalarına sıfır değerleri eklenir ve sonuç olarak $(K+(2xL)) \times (K+(2xL))$ boyutunda bir matris elde edilir. Yukarıda bahsi geçen işlem bir sıfır boşluk değeri ile işleme koyulmak istendiğinde 5×5 boyutundaki matris 7×7 boyutundaki bir başka matrise dönüşmüş olur. Eğer bu işlemin basamakları iki adımla sınırlandırılırsa 3×3 boyutunda bir matrise ulaşılmış olunur.

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
0	40	47	53	63	74	0	0	0	40	47	53	63	74	0	0	0	40	47	53	63	74	0	0			
0	42	50	57	65	76	0	0	0	42	50	57	65	76	0	0	0	42	50	57	65	76	0	0			
0	46	52	62	68	77	0	0	0	46	52	62	68	77	0	0	0	46	52	62	68	77	0	0			
0	48	56	64	69	79	0	0	0	48	56	64	69	79	0	0	0	48	56	64	69	79	0	0			
0	54	58	66	73	80	0	0	0	54	58	66	73	80	0	0	0	54	58	66	73	80	0	0			
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
I								II								III										
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
0	40	47	53	63	74	0	0	0	40	47	53	63	74	0	0	0	40	47	53	63	74	0	0			
0	42	50	57	65	76	0	0	0	42	50	57	65	76	0	0	0	42	50	57	65	76	0	0			
0	46	52	62	68	77	0	0	0	46	52	62	68	77	0	0	0	46	52	62	68	77	0	0			
0	48	56	64	69	79	0	0	0	48	56	64	69	79	0	0	0	48	56	64	69	79	0	0			
0	54	58	66	73	80	0	0	0	54	58	66	73	80	0	0	0	54	58	66	73	80	0	0			
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
IV								VI								VI										
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
0	40	47	53	63	74	0	0	0	40	47	53	63	74	0	0	0	40	47	53	63	74	0	0			
0	42	50	57	65	76	0	0	0	42	50	57	65	76	0	0	0	42	50	57	65	76	0	0			
0	46	52	62	68	77	0	0	0	46	52	62	68	77	0	0	0	46	52	62	68	77	0	0			
0	48	56	64	69	79	0	0	0	48	56	64	69	79	0	0	0	48	56	64	69	79	0	0			
0	54	58	66	73	80	0	0	0	54	58	66	73	80	0	0	0	54	58	66	73	80	0	0			
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
VII								VIII								IX										
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
0	40	47	53	63	74	0	0	0	40	47	53	63	74	0	0	0	40	47	53	63	74	0	0			
0	42	50	57	65	76	0	0	0	42	50	57	65	76	0	0	0	42	50	57	65	76	0	0			
0	46	52	62	68	77	0	0	0	46	52	62	68	77	0	0	0	46	52	62	68	77	0	0			
0	48	56	64	69	79	0	0	0	48	56	64	69	79	0	0	0	48	56	64	69	79	0	0			
0	54	58	66	73	80	0	0	0	54	58	66	73	80	0	0	0	54	58	66	73	80	0	0			
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
X								X								X										
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
0	40	47	53	63	74	0	0	0	40	47	53	63	74	0	0	0	40	47	53	63	74	0	0			
0	42	50	57	65	76	0	0	0	42	50	57	65	76	0	0	0	42	50	57	65	76	0	0			
0	46	52	62	68	77	0	0	0	46	52	62	68	77	0	0	0	46	52	62	68	77	0	0			
0	48	56	64	69	79	0	0	0	48	56	64	69	79	0	0	0	48	56	64	69	79	0	0			
0	54	58	66	73	80	0	0	0	54	58	66	73	80	0	0	0	54	58	66	73	80	0	0			
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
X								X								X										
I	II	III				82	157	213	IV	V	VI				94	178	224	VII	VIII	IX				54	124	153

Şekil 3.12. İki adımlı, sıfır boşluk konvolüsyon işlemi

3.3.1.1.2. Havuzlama katmanı

Basitçe havuzlama katmanı, konvolüsyon katmanında elde edilen özellik haritalarının boyutunu indirgemeyi amaçlar. KSA'da özellik çıkarımı aşamasında yer alan bu işlem ayrıca aşırı uyum (overfitting) probleminin çözümü için de uygulanan bir yöntemdir.

Farklı boyutlarda, özellikte ve değerlerde filtreler kullanılarak özellik matrisleri üzerinde hesaplamalar gerçekleştirir. Bahsedilen bu farklı özellikler, maksimum ve ortalama havuzlama işlemleridir. Maksimum havuzlama, filtre çerçevesi ile özellik matrisinde eşleşme sağlamış olan alan içerisindeki değerlerden en büyüğünü çıktı olarak verir. Ortalama havuzlamaysa, filtre çerçevesi ile özellik matrisinde eşleşme sağlamış olan alan içerisindeki değerlerin ortalamasını çıktı olarak verir. Yapılan bu çalışmada maksimum havuzlama katmanlarına yer verilmiştir. Havuzlama katmanının özellik matrisleri üzerine uygulanma sırası aşağıdaki gibidir:

- 1) Havuzlama çerçevesinin özelliği, değerleri ve adım sayısı seçilir.
- 2) Özellik matrisinin en üst sol köşesine seçilen değerlere uygun olarak havuzlama çerçevesi uygulanır.
- 3) Uygulanan çerçeve yerleştirildiği noktadan özellik matrisinin en sağına adım sayısı kadar kaydırarak uygulanır.
- 4) En sağa gelindikten sonra başa dönülür ve çerçeve, adım parametresi kadar aşağı kaydırılarak uygulanır.
- 5) Eşleşme sonlanana kadar üçüncü aşamaya geri dönülür.

Şekil 3.13.'de 4x4 boyundaki matriste, 2x2 boyutunda ortalama ve maksimum havuzlama katmanı uygulandığında elde edilen sonuçlar verilmiştir.



Şekil 3.13. Ortalama havuzlama ve maksimum havuzlama işlemleri

3.3.1.1.3. Relu (rectifier linear unit) katmanı

Bu katmanın aktif olabilmesi için, giriş değerinin belirli bir miktarın üzerinde olması gerekmektedir. Giriş değeri sıfırın altında bir değere sahipse, çıkış değeri sıfır olacaktır. Ancak giriş değeri belirli bir miktarın üzerindeyse, bağımlı değişkenle aralarında doğrusal bir ilişki var olacaktır. Rectified Linear Fonksiyonu birbirinden farklı birçok durumda performansını kanıtlanmış bulunmaktadır.

3.3.1.2. Sınıflandırma aşaması

KSA girdi olarak aldığı görüntülerden özellikleri çıkardıktan sonra bu özellikleri sınıflandırır. Tam bağı katmanları kullanarak sınıflandırma yapılmasını sağlayan KSA'nın çıkışında elde edilen değerler ondalık sayılardır. Büyük bir matris dağılımında sadece sınıfların olasılık dağılımını gösteren bir sayı değerinin elde edilmesi özellik matrislerinde boyut indirgeme yapılmasını gerektirir. Bu sayede tek boyutlu veri üzerinden sınıflandırma işlemi gerçekleştirilebilir.

Sınıflandırma iki şekilde gerçekleştirilebilir. Bunlar ikili sınıflandırma ve çok sınıflı sınıflandırmadır. İkili sınıflandırmada sınıf sayısı ikidir. Görüntüde nesne bulunuyorsa pozitif işaretleme, nesne bulunmuyorsa negatif işaretleme yapılır. Böylece görüntü içerisinde nesnenin varlığı ya da yokluğu kolaylıkla anlaşılabilir. Çoklu sınıflandırmada ise sınıf sayısı N adettir. Verinin N adet sınıf arasından hangisine ait olduğunu saptamak amaçlanır. Topun şekli, dokusu, rengi, boyutları gibi değerlere bakarak bir topun futbol ya da voleybol için uygun olduğunu belirleyen YSA modeli ikili sınıflandırmaya örnek olarak verilebilir. Çoklu sınıflandırmaya örnek olarak da renk, yolcu kapasitesi, logo şekli, kapı sayısı, şase yapısı gibi özelliklere bakarak N adet araç sınıfından tahminler yürüterek uygun sınıfı belirleyen YSA modeli gösterilebilir.

Sınıflandırma aşamasında ağın çıktısı olasılık değerleridir. Her sınıf için bir olasılık değeri üretilir. Dolayısıyla her sınıf için üretilen bu değerlerin toplamı 1,0 (%100) olmalıdır. En yüksek olasılık değerine sahip olan sınıf doğru sınıf olarak kabul edilir ve sonuç olarak yansıtılır. Softmax fonksiyonu bu süreçleri gerçekleştirerek 0,0 ile 1,0 arasında çıktılar üreterek uyumlu olan sınıfı belirlememizi sağlar.

3.3.1.2.1. Softmax fonksiyonu

KSA'nın son aşmasında kullanılan softmax, bir görüntüdeki objenin hangi sınıfa ait olduğunu belirler.

$$f(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (3.3)$$

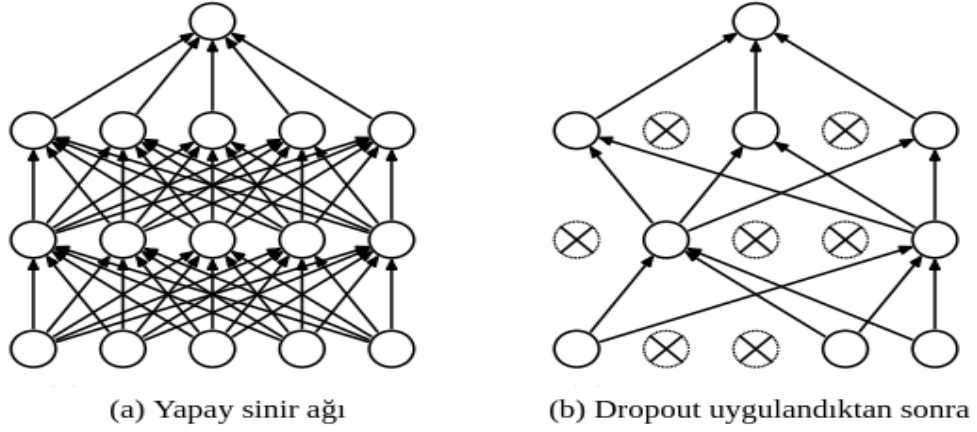
Sonuç olarak 0 ve 1 değerleri arasında değerler üreten bu fonksiyon bir sayısal dizgiyi girdi olarak alır. Sonuç değerlerinin toplamı %100 yani 1 olmalıdır. Buna göre çok sınıflı sınıflandırma probleminde en büyük değere sahip sınıf doğru tahmindir.

3.3.2. Regularizasyon teknikleri

Derin öğrenme modelleri çok sayıda sayıda parametreye sahiptir. Bu yüzden derin öğrenme modelleri; yeterli veri ile eğitilmeme, kullanılan veri çeşitliliğinin yetersiz oluşu gibi durumlarda aşırı uyum problemiyle karşı karşıya kalmaktadır. Yani sistem eğitim verilerini ezberlemektedir. Bu problemin çözümü içinse regularizasyon teknikleri uygulanmaktadır. Yaygın olarak kullanılan regularizasyon teknikleriyse L1, L2 ve Dropout'tur. Bu çalışmada ise Dropout tekniği kullanılmıştır.

3.3.2.1. Dropout

En yaygın kullanılan regularizasyon tekniklerinden biri olan dropout, L1 ve L2 methotlarından farklılık göstermektedir. Eğitim sırasında ağ üzerindeki sinir hücrelerinden bazılarını rastgele silerek sinir hücresinin aynı özellikleri temsil etmesi engellenir. Birçok ağın birleştirilmesinde de aktif görev alan dropout tekniği bir sinir ağı modelinin performansı artırarak iyi sonuçlar elde edilmesini sağlar.

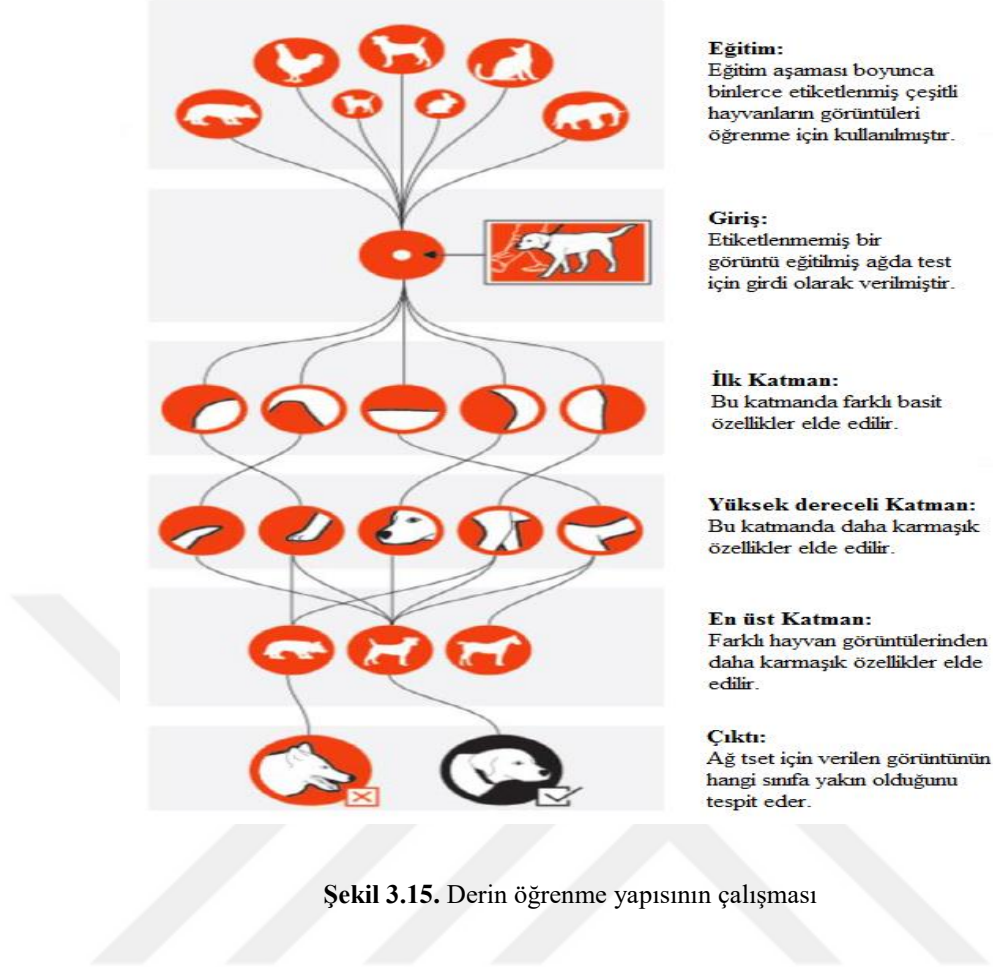


Şekil 3.14. Dropout'un YSA üzerinde kullanımı

3.3.3. Öğrenme süreci

Öğrenme süreci klasik makine öğrenmesinde olduğu gibi derin öğrenmede de veriler üzerinden gerçekleştirilir. Öğrenme süreçleri üç alt başlığa ayrılmaktadır. Bunlar; denetimli, denetimsiz ve takviyeli öğrenmedir. Bütün bu süreçlerdeyse sistemin yapıtaşlarından biri veridir. Dolayısıyla eğitim aşamasında kullanılacak olan verilerin gerçek dünyayı yansıtacak şekilde seçilmesi önem arz etmektedir.

Eğitim aşamasında bir ağ üzerindeki parametreler belirli aralıklarla güncellenir ve böylece ağların en optimum değeri yakalaması beklenir. Buna göre farklı veriler üzerinde önceden eğitilmiş olan modeller eğitim aşamasında kullanılacak olan yeni veriler üzerinde eğitildiğinde başka bir problemin çözümüne hizmet edeceklerdir.



Bu çalışmada denetimli öğrenme sürecine yer verilmiştir.

3.3.3.1. Denetimli öğrenme

Etiketli veriler üzerinden bir modelin anlamlar çıkartarak makinelerin öğrenmesine denetimli öğrenme denir. Sınıflandırma ve lokalizasyon problemlerinde kullanılan denetimli öğrenme, eğitim aşamasında değişik sahneleri barındırılan çok sayıda veri setiyle gerçekleştirilir. Öğrenme sırasında amaç ise sistemin hata oranını minimize ederek modeli optimize etmektir. Böyle elde edilen ağırlıklar da istenen durumları sağlayacak şekilde güncellenmektedir. Tahmin edilmiş olan değer ile gerçek değer arasındaki fark hata değeri şeklinde tanımlanmaktadır.

$$\text{Hata} = \text{Tahmin} - \text{Gerçek} \quad (3.4)$$

Öğrenme sürecinde çok miktarda veri üzerinde çalışıldığında, veri setleri parçalara ayrılarak yığınlar hâlinde eğitim aşamasına dahil edilmektedir. Örneğin X tane

veri seti Y tane yığınlar biçiminde sisteme dahil edildiğinde sistem, $Z=X/Y$ kadar döngüde eğitilir ve gerçekleştirilen bu işlem dönem (epoch) olarak adlandırılır.

3.3.4. Veri setinin hazırlanması

Bir derin öğrenme modelinin eğitim aşamasında, modeller eğitim veri setleri üzerinde eğitilirken, test verileriyle de doğrulanır. Böylece daha önce görmediği bu test verileriyle, modelin gerçek dünyayı temsil edip etmediği belirli aralıklarla kontrol edilmeye çalışılır. Denetimli öğrenmede amaç kaybı en aza indirerek doğrulu yükseltmektir. Bu nokta veri setinin seçimi oldukça önem arz etmektedir. Örneğin veri setlerinin olması gerektiği şekilde ayarlanmaması durumunda aşırı uyum (overfitting) ve alçak uyum (underfitting) adı verilen problemlerle karşılaşılır. Aşırı uyum, modelin gerçek dünyaya verilerine karşı başarısız olduğu anlamına gelirken alçak uyum, modelin veriler ile tam olarak örtüşmediğinin ya da modelin yeteri kadar eğitilmediğinin bir göstergesidir. Asıl ulaşılmak istenen ise modelin eğitim verileriyle test verilerinin her ikisinde de benzer sonuçlar ortaya koymasıdır.

Aşırı uyum sorununu ekarte edebilmek için nitelik ve nicelik bakımından yeterli olarak adlandırılabilir veri kullanılmalıdır. Bu çözümün yanı sıra kullanılan model basitleştirebilir veya dönem sayısı azaltılarak eğitim verilerinin ezberlenmesi önlenir. Bunun için erken durdurma (early stopping) kullanılarak ağ yeterli tekrar sayısına ulaşıldığında eğitim durdurulur. Bunların yanında regularizasyon teknikleri de kullanılabilir. Bu çalışmada regularizasyon yöntemi olarak dropout yöntemine yer verilmiştir.

Eğitim, test ve doğrulama olarak üçe ayrılmış veri setleri üzerinde modeller eğitilerek aşırı uyum probleminin saptanması sağlanabilir. Böylece eğitim sırasında, o ana kadar eğitilmiş model üzerinde doğrulama veri seti denenerek eğitim verilerinin haricindeki verilere dair sonuçların nasıl olduğu gözlemlenebilir ve buna göre model parametreleri ayarlanabilir.

Bu tez kapsamında WIDER Face (Yang ve ark., 2015) ve MAFA (Ge ve ark., 2017) açık kaynaklı veri setleri üzerinden alınan çeşitli görsellerinin bir kombinasyonu olan veri kümesi kullanılmıştır. Veri seti hazırlanırken tüm görüntüler bir dizi geliştirilmiş algoritma ile VOC formatına uygun olarak etiketlenmiş ve kontrol edilmiştir.

Konvolüzyonel sinir ağı eğitim, doğrulama ve test aşamalarından oluşan süreçlere sahiptir. Bu sebeple oluşturulan karma veri setindeki yüz görüntülerin %70'i eğitim, %10'u doğrulama ve %20'si test aşamaları için kullanılmıştır.



Şekil 3.16. Oluşturulan karma veri seti

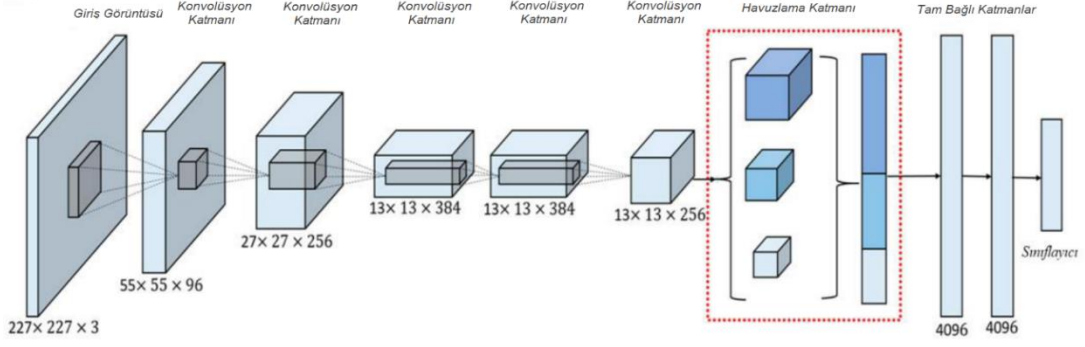
3.4. Temel KSA yapıları ve Transfer Öğrenmesi

Bu bölümde literatürün temelini oluşturan bazı sinir ağı yapıları açıklanmıştır. Literatürde konvolüzyonel sinir ağlarını içeren çok fazla çalışma olmasına rağmen bu tez kapsamında AlexNet, VGGNet, GoogleNet, ResNet ve DenseNet modelleri gibi günümüz modellerinin omurgasını oluşturan yapılara yer verilmiştir. Bu ağ yapıları aslında birbirlerinden son derece farklı olmakla birlikte kendi özelliklerinde hepsi bazı problemlerin çözümüne yönelik çalışmak için tasarlanmıştır.

3.4.1. Temel ksa yapıları

3.4.1.1. AlexNet

Krizhevsky ve arkadaşları (2012) nesne tanıma için AlexNet adını verdikleri bir derin öğrenme modeli geliştirmişlerdir. Şekil 3.17.'de görüldüğü gibi bu KSA yapısı beş konvolüzyonel katmanı ve iki tam bağlı katmandan oluşmaktadır. AlexNet, ImageNet adı verilen veri seti üzerinde eğitilmiş ve %16,7 bir hata oranı elde etmiştir. Hem lokalizasyon hem de görüntü sınıflandırma görevlerinde oldukça iyi performans gösteren AlexNet üzerinde yüksek doğruluk oranı elde etmek için gerçek dünyayı yansıtan daha büyük veri setlerin kullanılması gerekmektedir.

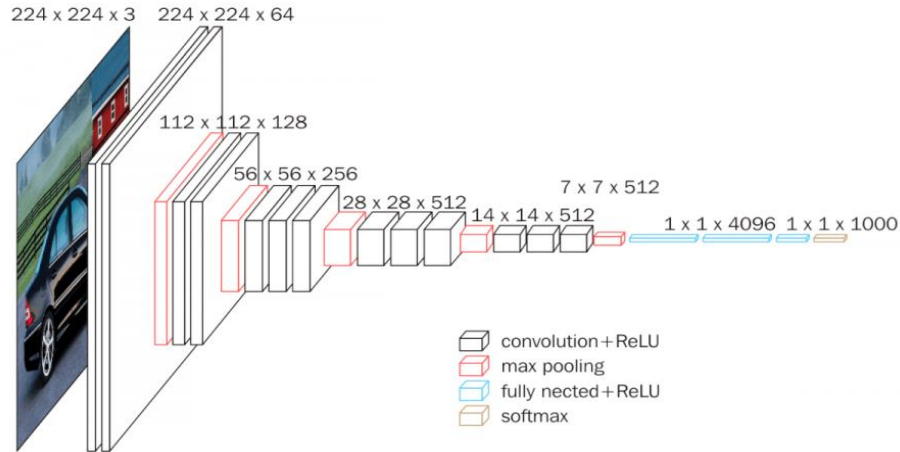


Şekil 3.17. AlexNet mimarisi

AlexNet tahmini olarak 60 milyon parametre içermektedir ve 650.000 yapay siniri bünyesinde barındırır. Bu yüksek parametre sayısına bakarak düşünüldüğünde, aşırı uyum problemi AlexNet için son derece önemli bir sorun haline gelir. Aşırı uyum problemini çözmek için veri artırımı ve kaçınma tekniklerini uygulanmaktadır.

3.4.1.2. VGGNet

Simonyan ve arkadaşları (2014) VGGNet geliştirilmiştir. Böylece derinlik kavramı literatüre kazandırılmıştır. 3×3 boyutlu filtreler kullanılarak, girdi görüntüsü adım kaydırma işlemi 1 ile sabit tutulup konvolüsyon katmanından geçirilmektedir. Konvolüsyon katmanından sonra tamamen bağlı (fully connected) katmanlar yer almaktadır. Yapılan testler sonucunda %10,4 hata oranı elde edilmiştir. Fakat çoklu kırpma ağı yöntemiyle bu oran %7,1'e kadar düşürülmüştür. VGG-16 modelinin mimarisi şekil 3.18'de gösterilmiştir.

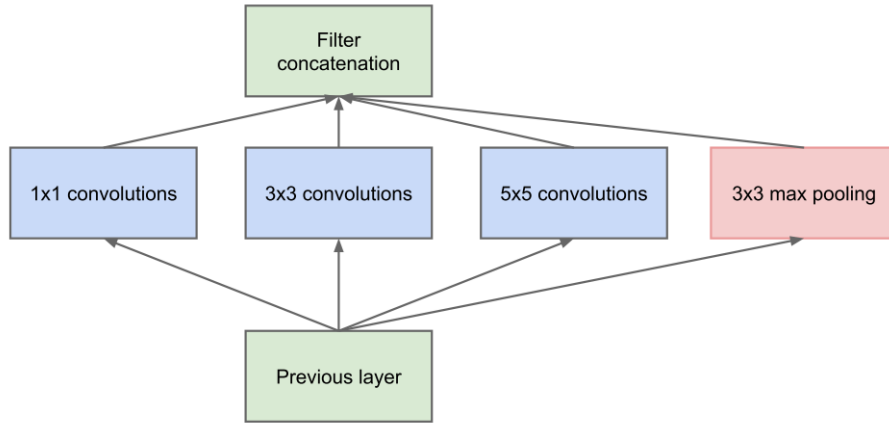


Şekil 3.18. VGG-16 mimarisi

224×224×3 giriş boyutuna sahip olan VGG-16 mimarisi, 3×3 filtre boyutuna sahip 13 adet konvolüsyon katmanı bulundurur. Bu konvolüsyon katmanlarından sonra 2×2 boyutunda maksimum havuzlama işlemi uygulanmaktadır.

3.4.1.3. GoogleNet

Szegedy ve arkadaşları (2015) tarafından 22 katmana sahip GoogLeNet modeli geliştirilmiştir. 2014 senesinde düzenlenen ImageNet görüntü sınıflandırma yarışmasında birinci olan bu modeli diğer modellerden farklı yapan noktaysa görü modülü (Inception module) geliştirilmesidir. Böylece modelin ihtiyaç duyduğu hesaplama gücü sabit kalırken derinleşmesi sağlanmıştır. Diğer modellerden daha az parametreye sahip olan bu model başarısını ispatlamıştır. Şekil 3.19.'da inception katmanının yapısına yer verilmiştir.



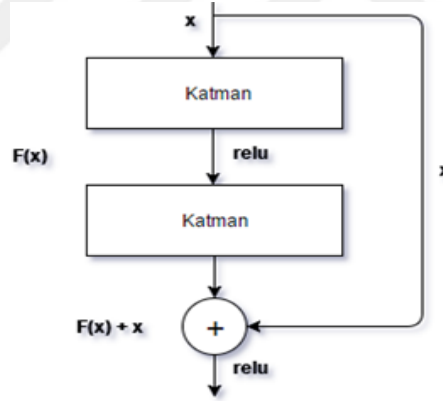
Şekil 3.19. Inception katmanı

Görüldüğü gibi bu katmanda her biri birbirinden bağımsız olarak çalışan ve kendisinden önce gelen katmandan aktarılan girdiyle işlem yapan 1x1, 3x3, 5x5 konvolüsyon katmanı ve 3x3 maksimum havuzlama katmanı kullanılmıştır (Yiğit, 2017). Böylece daha büyük ağların daha iyi sonuçlar vereceği belirlenmiştir. Daha büyük ağları elde etmek için iki yöntem bulunmaktadır. Bunlardan ilki ağın derinliğini artırmakken, ikincisiyse genişliğini (her katmandaki işlem sayısı) artırmaktır. Ağın derinliğini artırmak parametre sayısında artışına sebep olur ve bu da ezberlemeye yol açar. Genişliğini artırırsa hesaplama maliyetinde artışa sebep olur. Sonuç olarak bu mimariyle birlikte peş peşe konvolüsyon ve havuzlama katmanları kullanımından uzaklaşmıştır. Model üzerinde yapılan testler sonucunda yaklaşık olarak %6,7'lik bir hata elde edilmiş olup bu

oran neredeyse insan görüsü hata oranına (%5) denktir. Sonuç olarak KSA modellerinin insan görüsü hata oranına son derece yakın olduğu ve insan görüsü ile yarışa girebilecek durumda olduğu anlaşılmıştır.

3.4.1.4. ResNet

He ve arkadaşları (2016) ResNet adını verdikleri bir model geliştirmişlerdir. Microsoft tarafından desteklenen bu model 2015 yılında ImageNet yarışmasında birincilik elde ederek GoogLeNet'i geride bırakmıştır. ResNet, içerisinde artık blok (residual block) yapısını bulundurarak daha derin ağların oluşturulmasına olanak sağlamıştır. 152 katmandan oluşan ResNet, yapılan testlerde %3.57 hata oranı sergilemiştir. Böylece %1.43 farkla insan görüsünü geride bırakmıştır. ResNet50, ResNet101 ve ResNet150 gibi çeşitli versiyonları vardır ve birçok kütüphane tarafından kullanılmaktadır. Sonuç olarak artan derinlikle birlikte doğruluk oranının da arttığı gözlemlenmiştir. Şekil 3.20.'de artık ağ bloğu gösterilmektedir.



Şekil 3.20. Artık ağ bloğu

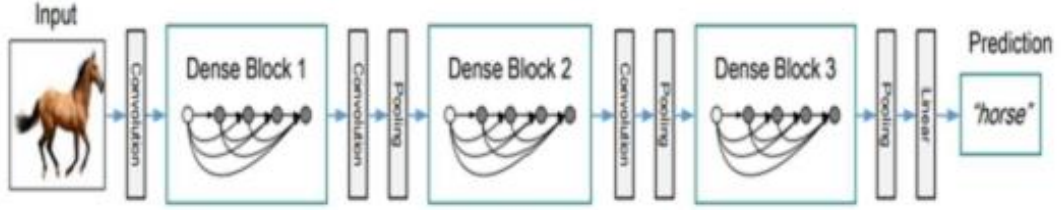
Artık ağlar literatüre girmeden önce 18 katmanlı bir sinir ağı, 34 katmanlı bir sinir ağına kıyasla çok daha iyi performans göstermekteydi. Fakat artık ağların kullanımıyla bu oran tersine dönerek derinlik artışı modellerin performansında artışa sebep olmuştur.

3.4.1.5. DenseNet

Huang ve arkadaşları (2017) tarafından DenseNet geliştirilmiştir. ResNet'e benzeyen bu ağ temelde iki önemli farka sahiptir. İlk fark, DenseNet öznetelikleri birbirine

eklemek yerine birleştirmesidir. İkinci farksa, DenseNet'in her katmandan elde edilen öznetelikleri bir önceki özneteliğe eklemek yerine tüm katmanlardakileri birleştirmeyi tercih etmesidir.

DenseNet ağı, katman başına 12 filtre düşecek şekilde modüler olarak inşa edilmiştir ve inşa edilen her modül birden fazla konvolüsyon katmanına sahiptir. Şekil 3.21.'de 3 adet yoğun bloğa sahip olan DenseNet ağı gösterilmiştir.



Şekil 3.21. DenseNet ağı

Derinliğine göre DenseNet-121, DenseNet-169, DenseNet-201, DenseNet-264 gibi çeşitli versiyonları olan DenseNet, ResNet göre 3 kat daha az parametreye sahiptir.

3.4.2. Transfer öğrenme

Transfer öğrenme, daha önceden özel bir amaç için oluşturulmuş bir veri seti üzerinde eğitilen modelin farklı bir görev için başlangıç parametreleri olarak kullanılmasıdır (Torrey ve ark., 2008). Transfer öğrenmedeki temel amaç farklı modeller üzerinden gelen ağırlıkları kullanarak daha az veri ve dönem sayısı ile yüksek başarıyı yakalamaktır. Bir yapay zekâ modelinin sıfırdan oluşturulması maliyetli olmakla birlikte ciddi zaman harcamaktadır. Transfer öğrenme doğru şekilde yapılırsa hiç uygulanmamış senaryoların bile maliyetli veri etiketleme çalışmalarına gerek kalmadan yüksek performanslı bir şekilde dağıtımını sağlanabilir.

3.4.3. Derin öğrenme kütüphaneleri

Var olan derin öğrenme uygulamalarının gelişen teknolojiyle birlikte bir üst seviyeye çıkarılabilmesi için uzman görüşü ve bilgisine ihtiyaç duyulurken bu süreç donanımsal maliyetleri de beraberinde getirmektedir. Ağ eğitimi sırasında kaynakların doğru ve verimli kullanılıyor olması performansı azımsanamayacak ölçüde

etkilemektedir. Örneklendirmek gerekirse CPU çekirdeklerinin paralel kullanılıyor olması, CPU kümelerinin oluşturulması, GPU'nun aktif kullanılıyor olması bahsi geçen süreçlere zaman bakımından olumlu katkılar sağlamaktadır. Bu gibi yapıları inşa edebilmek için bilgi birikimi ve zamana ihtiyaç duyulmaktadır. Derin öğrenme kütüphaneleri birbirinden farklı özelliklere sahip yapıları baz alarak çalışmasını sürdürürken, bunların dolaylı sonucu olarak da eğitim sürecini kısaltıp zamandan tasarruf etmeyi sağlamaktadır.

Bu çalışmada OpenCV'nin DNN modülü kullanılarak gerçekleştirilmiştir. OpenCV ile birlikte; Cuda ve cuDNN araçları kullanılmıştır. OpenCV DNN GoogLeNet, VGG-16, VGG-19, AlexNet, ResNet-50, ResNet-101 ve MobileNet sinir ağı modellerini bünyesinde bulundurmaktadır. Ayrıca GPU'lar üzerinden işlem yapılmasını sağlayarak algoritmaları hızlandırmaktadır. OpenCV Caffe ve TensorFlow-Keras derin öğrenme kütüphanelerini desteklemektedir. Bu tezde Caffe framework alt yapısı kullanılmıştır.



Şekil 3.22. OpenCV tarafından desteklenen derin öğrenme kütüphaneleri

3.4.3.1. Caffe

BSD lisansına sahip açık kaynaklı bir C++ kütüphanesi olan Caffe, bu özellikleri sayesinde araştırmacılara üzerinde değişiklik yapılabilir ve temiz bir çalışma alanı sunmaktadır. Berkeley Görü ve Öğrenme Merkezi (BVLC) tarafından geliştirilen kütüphane, ara yüzleri aracılığıyla Python ve Matlab ile geliştirme yapmaya imkan sunar. CPU ve GPU üzerinde paralel geliştirme yapılabilir. Akademik çalışmalarda ve endüstriyel uygulamalarda bu kütüphane yer verilmektedir.

Bu kütüphane CUDA destekli tek K40 veya Titan GUP üzerinde yaklaşık 2.5 milisaniyede bir görüntü işleme yeteneğine sahiptir. Bu da bir günde 40 milyon üzerinde görüntü işleyebildiğini gösterir. Bu yeteneği sebebiyle endüstride ve araştırmalarda sıkça kullanılarak araştırmacıların ihtiyaç duyduğu sistem eksikliğini gidermektedir. Kütüphaneye dair hazırlanmış olan detaylı dokümanda görüntü sınıflandırma, özellik çıkarımı gibi ihtiyaç duyulabilecek birçok uygulamanın örneklerine ulaşmak mümkündür. Doküman; eğitim, test etme, modelin üzerinde hassas ayarların yapılması ve modelin dağıtımı için eksiksiz bir araç seti sunmaktadır. Ayrıca algoritmaların hızlı bir şekilde uygulanmasını sağlayarak endüstri dağıtımlarında zamandan tasarruf edilmesine olanak sağlamaktadır. Bir öğrenme sistemi geliştirilmek istendiğinde kütüphane tarafından sunulan iyileştirme tanımlamaları ve tasarlanan modeller satırlarca kod yazmayı zorunlu kılmadan, yapılandırma dosyaları aracılığıyla pratik bir biçimde gerçekleştirilmektedir.

3.4.3.1. TensorFlow

TensorFlow, Google Brain ekibi tarafından oluşturulmuş ve ilk kez 2015 yılında halka sunulmuştur. Sayısal hesaplama yapmak ve büyük ölçekli makine öğrenimini gerçekleştirmek için son derece uygun açık kaynaklı bir kütüphanedir. TensorFlow, bir dizi makine öğrenimi ve derin öğrenme modelini ve algoritmasını bir araya getirir. Uygulamaları oluşturmak için uygun bir ön uygulama programlama arayüzü sağlamak için Python veya JavaScript'i kullanır ve bu uygulamaları yüksek performanslı C++'da yürütür (Li ve ark., 2017).

Tensorflow hem MİB hem de GİB desteği sağlayarak ekstra işlemlere gerek kalmadan modellerinizi ilgili metotları kullanarak oluşturmanızı sağlar. Özellikle bilgisayarlı görü uygulamaları başta olmak üzere birçok alanda kullanılmaktadır. Buna ek olarak dağıtılmış hesaplamalarını destekleyerek çok daha büyük ağ yapılarını kısa süre içerisinde eğitilmesini sağlar (Mo ve ark., 2017).

3.4.3.1. PyTorch

PyTorch, Torch'a dayalı optimize edilmiş olan açık kaynaklı bir derin öğrenme kütüphanesidir. Tensör hesaplamalarına dayanan PyTorch, çoğunlukla GPU'ları ve CPU'ları kullanan uygulamalar için kullanılır. PyTorch, dinamik hesaplama grafikleri

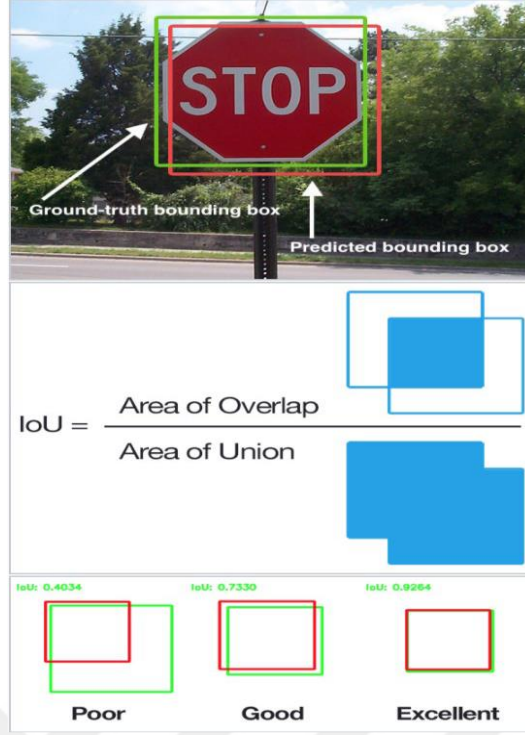
kullandığı ve son derece esnek betik bir dil olan Luna'yla geliştirildiği için TensorFlow ve Keras gibi diğer Derin Öğrenme çerçevelerine göre tercih edilir (R. Collobert ve ark., 2016). Pytorch'u cazip kılan bir başka noktaysa, geliştiricilerin sinir ağı kod bölümlerini gerçek zamanlı olarak çalıştırmasına, test etmesine ve hata ayıklamasına olanak tanınmasıdır. Böylece kullanıcılar, kodun bir bölümünün çalışıp çalışmadığını kontrol etmek için tüm kodun uygulanmasını beklemek zorunda kalmazlar.

3.5. Nesne Tanıma Tabanlı Konvolüsyonel Sinir Ağları

Nesne tanıma tabanlı KSA'lar, klasik KSA'ların daha gelişmiş bir halidir. Böylece bir sahnedeki nesnelere konumunu ve bu nesnelere karşılık gelen sınıf değerini tek bir çıktı olarak verirler. Nihai sonuç üzerinde elde edilen bu konumların temiz ve gerçeğe en yakın şekilde gösterilmesi için IoU ve NMS gibi bazı yaklaşımların kullanılması gerekmektedir.

3.5.1. IoU (Intersection over union)

MultiBox (çoklu kutu) ile önceden hesaplanmış, anchor box (çapa kutus) olarak adlandırılan orijinal kesin referans kutularıyla neredeyse birebir eşleşen sabit boyutlu sınırlayıcı kutular oluşturulur. Bu çapa kutuları, IoU oranı 0,5'ten daha fazla olacak biçimde seçilir. Fakat aşağıdaki şekilde de görüleceği gibi, 0,5'lik bir IoU hala yeterince iyi değildir. Sınırlayıcı kutu regresyon algoritması için güçlü bir başlangıç noktası sağlayan bu durum, tahminlere rastgele koordinatlarla başlamaktan çok daha iyi bir stratejidir. Dolayısıyla MultiBox, çapa kutuları tahminleriyle başlar ve orijinal kesin referans kutularıyla daha yakın bir yere gerilemeye çalışır. Sonuç olarak; MultiBox, hem konum hem de güven kayıplarını en aza indiren en iyi K tahminlerini tutar.



Şekil 3.23. IoU'yu açıklayan şema

3.5.2. Maksimum olmayan bastırma (NMS-Non-maximum suppression)

Çıkarım sırasında üretilen sınırlayıcı kutuların (bounding box) çoğu maksimum olmayan bastırma olarak bilinen bir teknik uygulayarak ortadan kaldırılabilir. Güven kaybı ve IoU'su belirlenen eşik değerlerinden az olan kutular atılır. Bu, gürültülü tahminlerin kaldırılmasını ve sadece en olası tahminlerin sonuç olarak gösterilmesini sağlar. Şekil 3.24.'de NMS örneği gösterilmiştir.

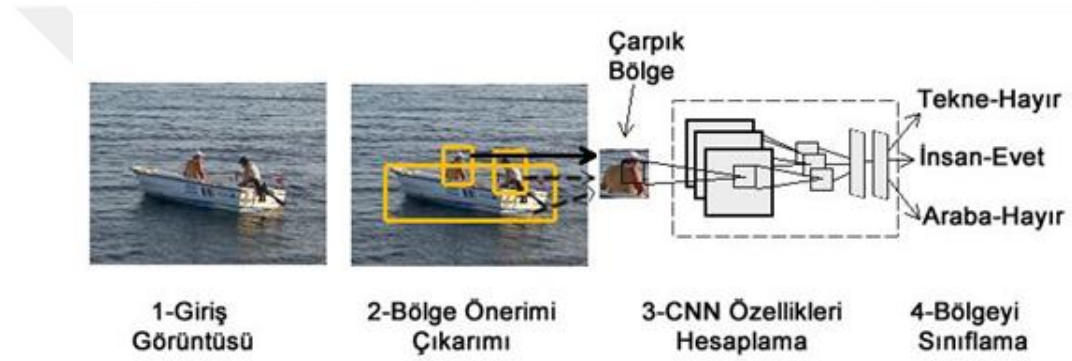


Şekil 3.24. NMS örneği

3.5.3. Temel nesne tanıma tabanlı ksa yapıları

3.5.3.1. Bölgesel tabanlı konvolüsyonel sinir ağları

Girshick ve arkadaşları (2014) bölge bazlı konvolüsyonel sinir ağlarını (B-KSA) öne sürülmüştür. Öncelikle girdi olarak alınan görüntüler üzerinde nesnenin bulunma ihtimalinin olduğu yaklaşık iki bin farklı bölge tanımlanır. Seçici arama algoritması kullanarak elde edilen bu nesnelere her biri KSA'ya sokulur ve imge üzerindeki özellik haritalarının elde edilmesi sağlanır. Elde edilen bu özellik haritaları sınıflandırma işlemine tabi tutularak görüntü üzerinde bulunan nesnelere algılanır. Şekil 3.25.'de R-CNN çalışma şekli gösterilmiştir.



Şekil 3.25. R-CNN

R-CNN, oldukça güçlü bir bölgesel KSN modelidir fakat mevcut sahne üzerindeki tüm nesnelere yerlerini bulmaya çalıştığından maliyeti yüksektir. Bu sebeple mevcut arama bölgesini kısıtlamak gereklidir.

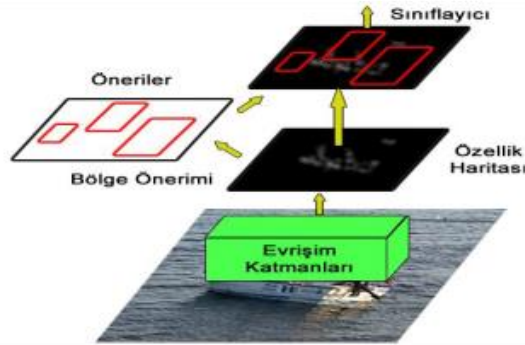
Aynı sahne içerisindeki nesnelere tespiti zor bir süreçtir. Örneğin görüntü içerisindeki nesnelere birbirinin parçası olabilir veya aynı nesnelere farklı ölçeklerde bulunabilir. Herhangi bir görüntü içerisinde sadece renk bazlı bir ayırım yapılmaya çalışılması birçok nesnenin parçasının ayrı bir bölüm olarak algılanmasına sebep olacaktır. Örneğin otobüsün tekerlekleri araçtan farklı bir nesne olarak algılanabilir. Bu sebeple nesnelere belirli bir yapıya göre tanımlamak gereklidir. Segmentasyon işlemlerinde Seçici Arama Algoritması küçük bölgelerden büyük bölgelere doğru hiyerarşi oluşturur. Bu süreçte izlenen 3 temel adım:

- 1) Başlangıç bölgeleri oluşturulur.
- 2) Oluşturulan bu bölgeler arasındaki komşuluklar belirlenir

- 3) Aç Gözlü (Greedy) Algoritması kullanılarak öncelikle birbirine en fazla oranda benzemekte olan iki tane başlangıç alanı belirlenir ve birleştirilir. Böylece tek bir bölge oluşturulması sağlanarak işleme devam edilir. Bu adım tüm başlangıç noktaları tamamlanıp sadece bir bölge kalıyorsa kadar devam eder. (Uijlings ve ark., 2013).

3.5.3.2. Hızlı bölgesel tabanlı konvolüsyonel sinir ağları

Girshick ve arkadaşları (2015) bölge bazlı konvolüsyonel sinir ağlarının yavaşlık problemine çözüm getirmek amacıyla Fast R-CNN olarak bilinen hızlı bölgesel ağları öne sürdüler. Fast R-CNN, çok aşamalı işlemlerden oluşan R-CNN'nin getirdiği hesaplama maliyetine çözüm bulmayı ve dolayısıyla zaman kaybını azaltmayı amaçlamaktadır. Böylece Faster R-CNN ile elde edilen 47 saniyelik gecikme 0.3 saniyeye kadar indirilmiştir. Birkaç konolüsyon ve havuzlama katmanından oluşurlar. Konvolüsyon aşaması sonunda elde edilen özellik haritaları havuzlama katmanına sokulur ve sabit boyutlu özellik matrisleri elde edilir. Elde edilen bu matrisler en sonunda tam bağlı katmanlara iletilerek sınıf skorlarının optimize edilmesi sağlanır. Son olarak ağdan iki çıktı elde edilir. Bunlardan biri softmax olasılık değeri, diğeryse nesnenin sınırlarını belirten koordinat değerlerdir. Böylece tek bir ağ ile hem nesnelerin algılanması hem de sınıflandırılması sağlanır.

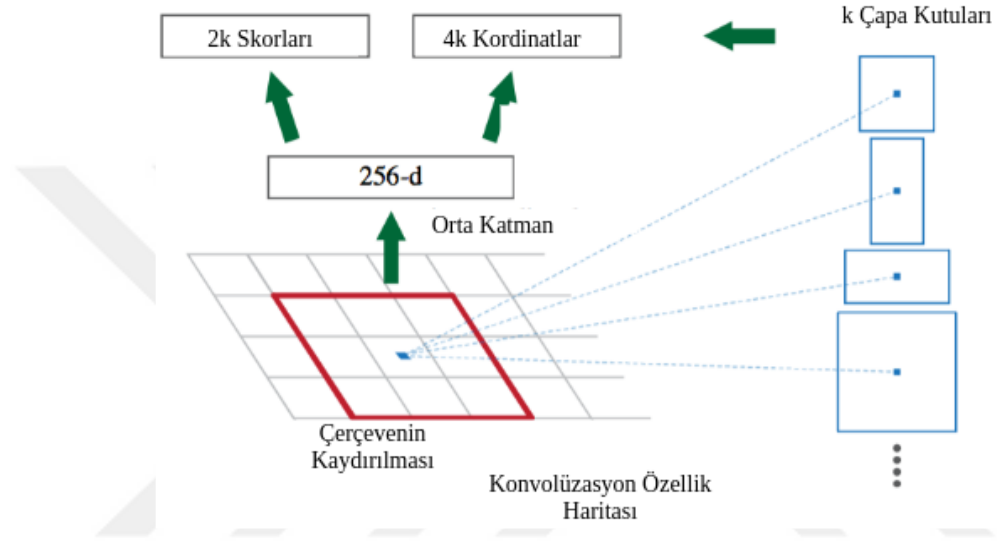


Şekil 3.26. Fast R-CNN

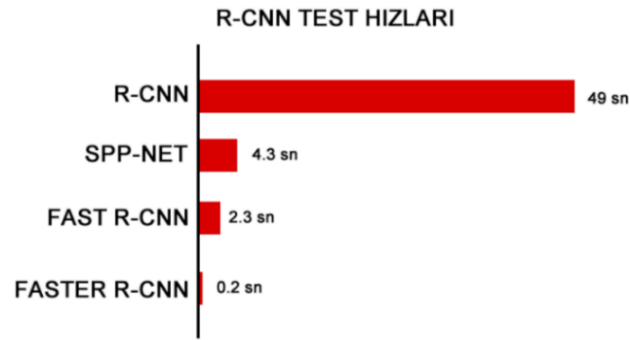
3.5.3.2. Daha hızlı bölgesel tabanlı konvolüsyonel sinir ağları

Günün sonunda mevcut donanım alt yapısıyla birlikte Fast-RCNN'nin de gerçek zamanlı uygulamalar için yeteri kadar performanslı olmadığı görülmüş ve Fast-

RCNN'nin daha hızlı bir modeli olan Faster R-CNN önerilmiştir. Ren ve arkadaşları (2015) Fast R-CNN'ni hızlandırmak için konvolüsyonel katmanların çıkışına Bölge Öneri Ağlarının (RPN) uygulanmasını önermiştir. Özellik haritalarını girdi olarak alan bu ağlar elde edilen nesnelerin skorlarını ve sınır kutularını çıktı olarak verir. İlk olarak özellik haritasından belirli bir boyuta kutu belirlenir ve her kutunun ortasına çapa (anchor) yerleştirilir. Sonuç olarak en boy oranına göre asıl bölgeler saptanır. (Ren ve ark., 2015). Şekil 3.27.'de Faster R-CNN çalışma yapısı gösterilmiştir.



Şekil 3.27. Daha hızlı R-CNN



Şekil 3.28. Saniye bazında nesne algılama algoritmalarının kıyaslanması

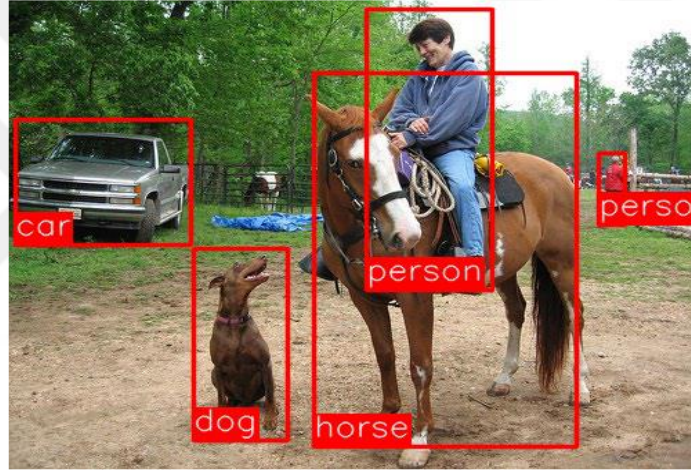
Günün sonunda Faster R-CNN'le belirli bir hız yakalansa da gerçek zamanlı uygulamalar için yeterli değildir. Bunun başlıca nedenleriyse;

- Ağların büyük veri setleri üzerinden eğitiminin çok uzun olması,

- Sınıflandırıcı ve bölge önerici ayrı ayrı eğitilir. Dolayısıyla normal eğitim süresinin yaklaşık iki katı bir zaman tüketimine neden olması,
- Çıkarım zamanında (eğitim dışı veriler üzerinde) ağıın çok yavaş çalışmasıdır.

Şekil 3.28.'de görüldüğü gibi Faster R-CNN en az test süresine sahip olan bir ağ modelidir ve bu da gerçek zamanlı uygulamalarda kullanılabilirliğini kanıtlar.

Dolayısıyla son birkaç yılda bu problemlerin çözümü için yeni yaklaşımlar öne sürüldü ve böylece gerçek zamanlı uygulamalar üzerinde çalışabilen nesne dedektörleri oluşturuldu. Bunlardan en ünlüleri YOLO (You Only Look Once) ve SSD (Single Shot Detector)'dir. Şekil 3.29.'da tipik bir R-CNN'nin çıktıları gösterilmektedir.

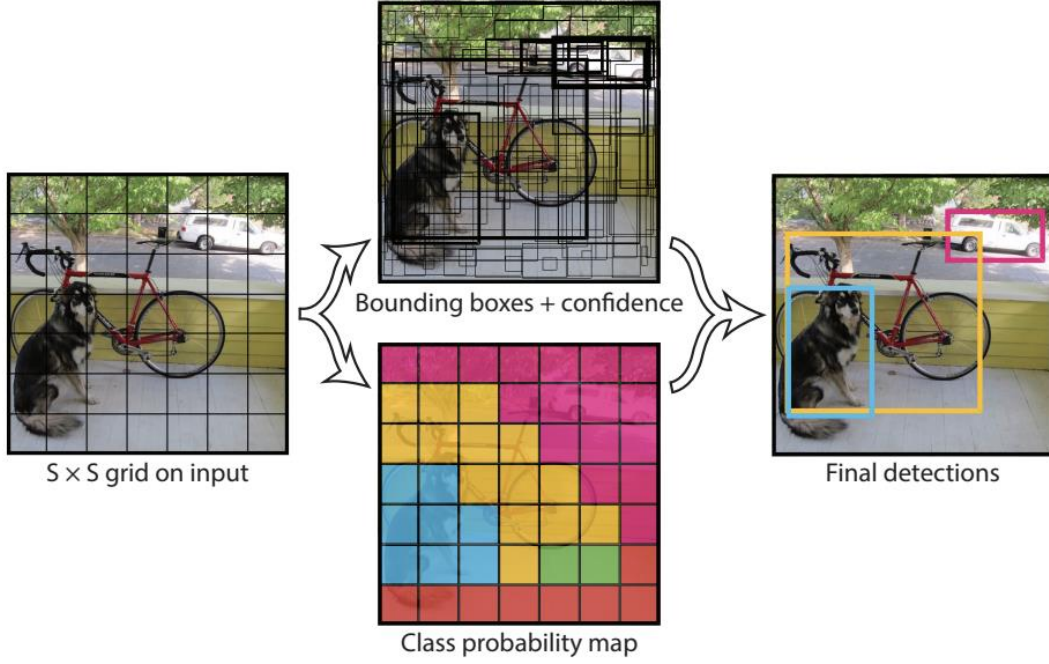


Şekil 3.29. R-CNN çıktıları

3.3.2.3. You only look once (YOLO)

J. Redmon ve arkadaşları (2016) bir görüntüdeki nesnelere ve onların konumlarını belirlemek için tasarladıkları YOLO mimarisini, tespit edilen nesnelere etrafına sınırlayıcı kutular çizer, eş zamanlı olarak birden fazla nesneyi algılamayı sağlar. Diğer algoritmalar, nesnelere algılamak ve sınıflandırmak için görüntü üzerinde çok katmanlı görevler gerçekleştirirken YOLO tek bir ağ geçişinde görevi başarıyla tamamlar. Bu da YOLO'yu çok hızlı ve performanslı yapan unsurdur. Omurgasında açık kaynak kodlu, C/Cuda ile yazılmış olan Darknet çerçevesini kullanır. Hem CPU hem de GPU üzerinde hesaplamaları destekleyen YOLO literatürde bir dönüm noktası olarak nitelendirilmektedir.

YOLO, aynı anda hem nesne konumu ve bu nesneye ait sınıf tahminini çıktı olarak vermesi, basitliği ve hızlı olması nedeniyle gerçek zamanlı uygulamalara uygundur. 24 konvolüsyon ve 2 tam bağlı katmandan oluşan YOLO, giriş görüntüsünü girdi olarak $S \times S$ ızgaralarına böler ve ardından her bir ızgara üzerinde N sınırlayıcı kutuyu bir güven eşiğiyle tahmin eder.



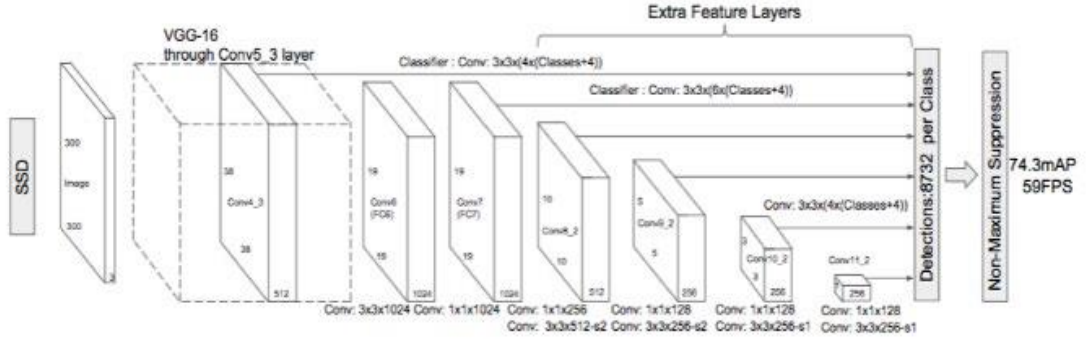
Şekil 3.30. YOLO mimarisi

3.5.3.4. Tek atış dedektörü (single shot detector-SSD)

Kasım 2016 sonunda ortaya çıkan bu model, nesne algılama görevlerini performans ve hassasiyet açısından yeni rekorlara ulaştırdı. PascalVOC ve COCO gibi standart veri kümeleri üzerinde yapılan denemeler sonucunda, saniyede 59 kare hızında (FPS) %74 mAP (ortalama hassasiyet) sahip performans sergilerdi (Liu ve ark., 2016). SSD adından da anlaşılacağı üzere kendine özgü üç temel yöntemi ele alarak çalışır. Bunlar;

- Tek çekim: Bu, nesnenin konumunun belirlenmesi ve sınıflandırılması işlemlerinin ağır tek bir ileri geçişinde yapıldığı anlamına gelir.
- Çoklu kutu: Szegedy ve diğerleri tarafından geliştirilen sınırlayıcı kutu regresyon tekniğine verilmiş olan addır.

- Dedektör: Ağ, algılanan nesnelere aynı zamanda sınıflandırmaya yarayan bir nesne dedektörüne verilen addır.

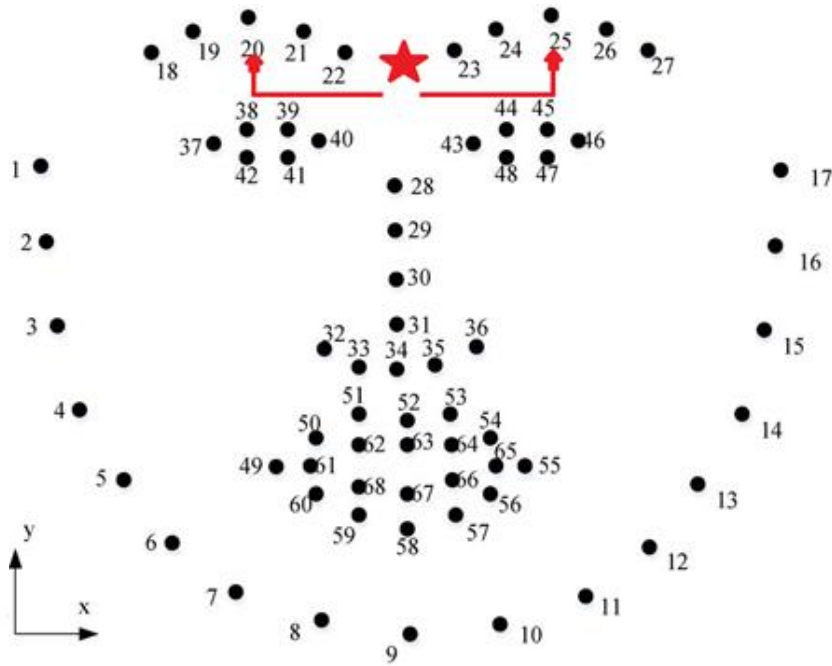


Şekil 3.31. SSD mimarisi

Şekil 3.31.'de görüldüğü gibi, SSD mimarisi VGG-16 mimarisine baz alarak oluşturulmuştur. Fakat VGG'de bulunan tam bağımlı katmanlar (full connected layers) atılmıştır. VGG-16'nın temel ağ olarak kullanılmasının nedeni, yüksek kaliteli görüntü sınıflandırma görevlerindeki güçlü performansı ve aktarımlı öğrenme (transfer learning) ile sonuçları iyileştirmeye yardımcı olmasıdır. Orijinal VGG'deki tam bağımlı katmanlar (full connected layers) yerine, bir dizi yardımcı konvolüzyon katmanı (conv6'dan itibaren) eklenerek özelliklerin birden çok ölçekte çıkarılması ve sonraki her katmanın giriş boyutunun kademeli olarak azaltılması sağlandı.

3.6. Yüz işaretleri

Yüzdeki işaretlerin belirlenmesi son zamanlarda hemen hemen her alanda sıklıkla kullanılan bir metottur. Yüzdeki işaretlerin belirlenmesiyle, yüz hizalama, baş poz tahmini, yüz değiştirme, göz kırpma algılama vb. birçok uygulama başarıyla gerçekleştirildi. Yüzdeki işaretlerin algılanması, şekil tahmin probleminin bir alt kümesidir. Bir girdi görüntüsü (ve normalde ilgilenilen nesneyi belirten bir ROI) verildiğinde, bir şekil öngörücü, şekil boyunca önemli ilgi noktalarını lokalize etmeye çalışır. Şekil 3.32.'de yüz işaretleri gösterilmektedir.



Şekil 3.32. Yüz işaretleri

Yüzdeki işaretlerin belirlenmesindeki amaç şekil tahmin yöntemlerini kullanarak yüzdeki önemli yüz yapılarını tespit etmektir. Yüzdeki işaretlerin tespit edilmesi bu nedenle iki aşamalı bir süreçtir. İlk olarak herhangi bir görüntüde bulunan yüzler tespit edilmeli ve daha sonra buluna yüzlerin ROI'sindeki temel yüz işaretleri tespit edilmelidir. İlk adım önceki başlıkta belirtildiği gibi bir derin öğrenme modeli olan SSD kullanılarak gerçekleştirilecektir.

Bu çalışmada Kazemi ve Sullivan (2014) tarafından yazılan “One Millisecond Face Alignment with an Ensemble of Regression Trees” makalesindeki yüz işaret detektörden yararlanılacaktır. Bu yöntem aşağıdakileri kullanarak başlar:

- Etiketli yüz işaretlerinden oluşan bir eğitim seti oluşturulur. Veri setindeki görüntüler, her yüz yapısını çevreleyen bölgelerin belirli koordinatlarını belirterek manuel olarak etiketlenir.
- Önceki değerler, giriş piksel çiftleri arasındaki mesafe oranına göre özel olarak elenir.

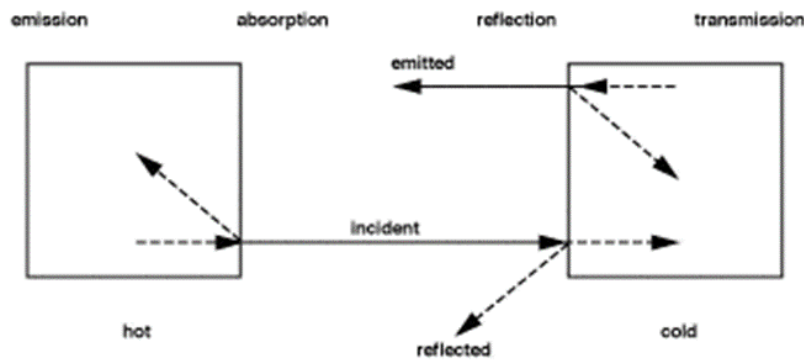
Bu eğitim verileri göz önüne alındığında, bir regresyon ağaçları topluluğu, yüzün dönüm noktası konumlarını doğrudan piksel yoğunluklarından tahmin etmek için eğitilir. Yani, bu aşamada özellik çıkarımı yapılmaz.

Sonuç olarak yüksek doğruluk oranına sahip, gerçek zamanlı olarak yüzdeki yer işaretlerini tespit etmek için kullanılabilen bir yüz işareti detektörü oluşturulmuştur.

3.7. Kızılötesi sıcaklık ölçüm teorisi

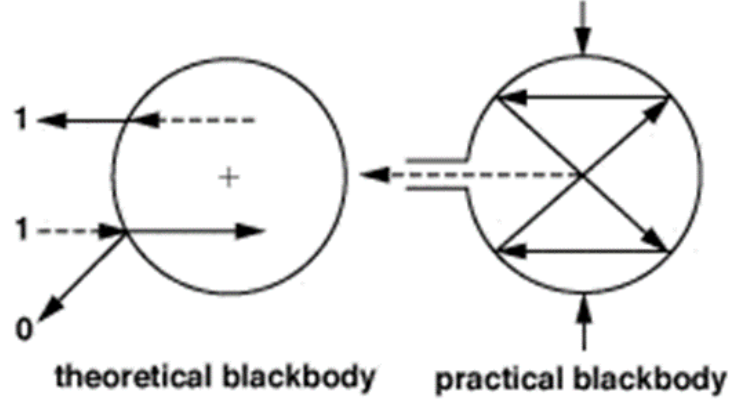
Kızılötesi (IR) enerji, 0 °K üzerindeki tüm malzemeler tarafından yayılır. Kızılötesi radyasyon, Elektromanyetik Spektrumun bir parçasıdır ve görünür ışık ile radyo dalgaları arasındaki frekansları işgal eder. Spektrumun IR kısmı 0.7 μm 'den 1000 μm (mikron) kadar olan dalga boylarını kapsar. Bu dalga bandında, pratik, günlük sıcaklık ölçümü için yalnızca 0,7 mikron ila 20 mikron frekanslar kullanılır.

IR radyasyonu insan gözüyle görülemez de ölçüm ilkeleri ile uğraşırken ve uygulamaları değerlendirirken onu görünür olarak hayal etmek yararlıdır, çünkü birçok bakımdan görünür ışıkla aynı şekilde davranır. IR enerjisi kaynaktan düz çizgiler halinde ilerler ve yolundaki malzeme yüzeyleri tarafından yansıtılabilir ve absorbe edilebilir. Şekil 3.33.'te radyoaktif ısı değişimi gösterilmiştir.



Şekil 3.33. Radyoaktif ısı değiştirme

Herhangi bir IR enerjisini yansıtmayan veya iletmeyen malzemeler kara cisimler (black body) olarak bilinir ve doğal olarak var oldukları bilinmemektedir. Bununla birlikte, teorik hesaplama amacıyla, gerçek bir kara cisme 1.0 değeri verilir. Gerçek hayatta elde edilebilen 1.0'lık kara cisim emisyonuna en yakın yaklaşım, aşağıdaki şekilde gösterildiği gibi küçük bir boru şeklinde girişi olan bir IR opak, küresel boşluktur.



Şekil 3.34. Emisyon

Farklı malzeme ve gaz türlerinin farklı emisyonları vardır ve bu nedenle belirli bir sıcaklık için farklı yoğunluklarda IR yayarlar.

Görünür ışıpta olduğu gibi, bazı yüzeyler ne kadar çok parlatılırsa, yüzey o kadar fazla IR enerjisi yansıtır. Bir malzemenin yüzey özellikleri bu nedenle aynı zamanda yayıcılığını da etkileyecektir. Sıcaklık ölçümünde bu, doğası gereği düşük bir salım gücüne sahip kızılötesi opak malzemeler durumunda en önemlisidir. Bu nedenle, yüksek derecede parlatılmış bir paslanmaz çelik parçası, pürüzlü, işlenmiş bir yüzeye sahip aynı parçadan çok daha düşük bir salım gücüne sahip olacaktır. Bunun nedeni, işleme tarafından oluşturulan olukların IR enerjisinin çoğunun yansıtılmasını önlemesidir. Moleküler yapı ve yüzey durumuna ek olarak, bir malzemenin veya gazın görünen yayılımını etkileyen üçüncü bir faktör, sensörün spektral tepkisi olarak bilinen sensörün dalga boyu duyarlılığıdır. Daha önce belirtildiği gibi, Pratik sıcaklık ölçümü için yalnızca 0,7 mikron ile 20 mikron arasındaki IR dalga boyları kullanılır. Bu genel bant içinde, sensörler 0.78 ila 1.06 veya 4.8 ila 5.2 mikron gibi, bandın yalnızca dar bir bölümünde çalışabilir.

3.8. Performans Metrikleri

Gerçekleştirilen test aşamasında eğitim modellerinin doğruluk değerleriyle birlikte karmaşıklık matrisinin de hesaplaması kullanılmış ve F1 skor, kesinlik(precision) ve anma(recall) değerleri hesaplanmıştır. Aşağıda verilen şekil 3.35.'de karmaşıklık matrisine yer verilmiştir.

		Tahmin Değerleri		Toplam
		A	B	
Gerçek Değerler	A	DP (Doğru Pozitif)	YN (Yanlış Negatif)	N _A
	B	YP (Yanlış Pozitif)	DN (Doğru Negatif)	N _B
Toplam		N' _A	N' _B	N

Şekil 3.35. Karmaşıklık matrisi

Matriste tüm örneklerin toplam sayısı N ile, tahmin toplamları N'_A ve N'_B ile, sınıf temelli toplam örnek sayısı ise N_A ve N_B değerleriyle belirtilmiştir. Doğru tahminde bulunulan A sınıf sayısı DP , Doğru tahmin edilen B sınıf sayısı DN , gerçekte A sınıfına ait olup da B sınıfı olarak tahmin edilen örnekleri sayısı YN , gerçekte B sınıfına ait olup da A sınıfı olarak tahmin edilen örneklerin sayısı ise YP ile belirtilmiştir. Matriste doğruluk hesaplaması yapılmak istenirse doğru tahmin edilen değerlerin toplam örnek sayısına oranına bakılır. Sonuç değeri 0,1 veya 0 ve 1 arasındadır.

$$\text{Doğruluk} = \frac{DP + DN}{N} \quad (3.5)$$

Bir sınıfa ait doğru tahmin edilen örnek sayısının, toplam tahmin sayısına bölünmesiyle kesinlik değeri elde edilir. A sınıfının kesinlik değerini elde ederken doğru tahminlerin sayısı, doğru tahminlerin sayısı ile gerçek değeri B olan fakat A olarak yanlış tahmin edilenlerin örnek sayılarının toplamına oranlanarak bulunur. Hem A hem B sınıflarında A sınıfı olarak tahmin edilen durumun başarısına aşağıdaki denklem 3.6'da değinilmiştir.

$$\text{Keskinlik} = \frac{DP}{N'_A} = \frac{DP}{DP+YP} \quad (3.6)$$

Bir sınıfa ait doğru sınıflandırılan örnek sayısının, doğru ve yanlış sınıflandırılan toplam örnek sayısına oranlanmasıyla anma değeri elde edilir. A sınıfının anma değeri elde edilirken doğru tahminlerin sayısının, doğru tahminlerin sayısı ve gerçek değeri A olan

fakat B olarak adlandırılmış yanlış tahminlerin örnek sayılarının toplamına oranlanması ile elde edilir. Aşağıdaki denklem 3.7' de A sınıfının tahmin düzeyinin başarılı olarak nitelendirilebilecek düzeyde olduğu gösterilmiştir.

$$\text{Anma Değeri} = \frac{DP}{N_A} = \frac{DP}{DP + YN} \quad (3.7)$$

Kesinlik değeri ve anma değerinin harmonik ortalaması hesaplanarak F1 skoruna ulaşılır. Kesinlik değeri ve anma değeri arasında büyük bir fark olursa F1 değerinin küçük olan değere yakın bir değer alması harmonik ortalama sayesinde olmaktadır.

$$\text{F1 Skoru} = \frac{2 * \text{Kesinlik} * \text{AD}}{\text{Kesinlik} + \text{AD}} \quad (3.8)$$

3.8. Nvidia Jetson Nano Modülü

Bu tez çalışmasında önerilen sistem NVIDIA Jetson Nano modülü kullanılarak gerçekleştirilmiştir. Geliştirilen algoritma yapısının ve modellerin çalıştırılmasını sağlayan Nvidia Jetson Nano görüntü sınıflandırma, nesne algılama, segmentasyon ve konuşma işleme gibi uygulamalar için birden çok sinir ağını paralel olarak çalıştırmanıza izin veren küçük, güçlü bir bilgisayardır. Sadece 5 watt kadar enerji tüketimine sahip olan bu modül üzerinde NVIDIA Maxwell 128 CUDA çekirdekli GIB'ni barındırır. 4 GB dahili hafızası bulunan bu modül, sistem görüntüsünü içeren bir micro sd kart üzerinden işlemlerini gerçekleştirir. Literatürdeki birçok yapay zekâ çerçevesini destekler. Ayrıca iki kamera girişi bulunan kart üzerinde USB bağlantı girişleri de mevcuttur. Şekil 3.36'da Jetson Nano modülü gösterilmiştir.



Şekil 3.36. Jetson Nano modülü

3.9. Raspberry Pi Kamera Modülü

Bu tez kapsamında RGB görüntülerin elde edilmesinde 8 megapiksel sabit odaklı Raspberry Pi kamera modülü kullanılmıştır. Pi Kamera modülü, fotoğraf ve yüksek çözünürlüklü video çekmek için kullanılabilen bir kameradır. Bu modül, 15 pimli şerit kablo kullanarak Jetson Nanonun CSI bağlantı noktasına takılmaktadır. Şekil 3.37’de Raspberry Pi kamera modülü gösterilmiştir.



Şekil 3.37. Raspberry Pi kamera modülü

3.10. Optris CS LT IR Sıcaklık Sensörü

Bu tez kapsamında sıcaklık değerlerinin elde edilmesinde 80°C'ye kadar ek soğutma gerektirmeyen ve kolay yerleştirilebilen 12x1x87 mm ölçülerinde kompakt paslanmaz çelik kutuya sahip, programlanabilir Optris CS LT IR sıcaklık sensörü kullanılmıştır. Analog ve dijital çıktının yanı sıra alarm kontağı olarak da kullanılabilen bu sensörün ölçüm alanı üretici firma tarafından sağlanan lensleri ve değerleri kullanarak ayarlanabilmektedir. Ayrıca üzerinde yerleşik olarak bulunan yeşil LED alarm lambası programlanarak teşhis veya sıcaklık kodu göstergesi olarak kullanılabilir.



Şekil 3.38. Optris CS LT IR sıcaklık sensörü

3.10. Önerilen Yöntem

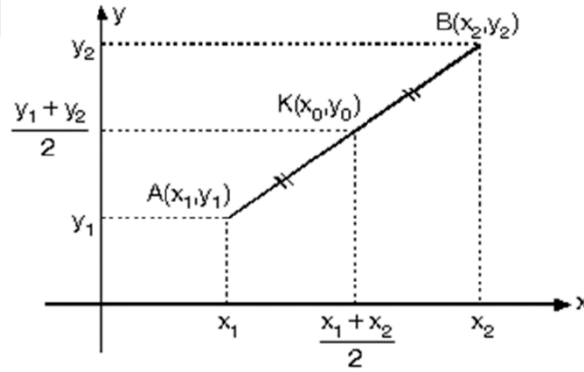
Bu çalışmada günlük hayattaki kullanımının yanı sıra Covid-19 sonrası gittikçe yaygınlaşan, alından temassız ölçüm yapan ateş ölçerlere alternatif olarak kullanılacak, insan görüşünü derin öğrenme algoritmasıyla destekleyen ve kızıl ötesi sıcaklık ölçüm teorisine dayanan bir sistem gerçekleştirilmiştir. Pandeminin olumsuz etkisini kontrol altına alarak hastalık şüphesine sahip kişilerin tespiti için yapılan temassız ateş ölçümü çoğu zaman kapı girişlerinde görevlendirilmiş bir personel tarafından yapılmaktadır. Bu durum iş gücünü gerekli kılmakta ve insan etkisine bağlı hataları beraberinde getirerek maliyetin artmasına neden olmaktadır. Özellikle tabanca tipi termometreler, ölçüm süresi açısından verimli olmamakla birlikte personel ile ölçüm yapılan kişi arasında doğrudan teması zorunlu kılmaktadır (Gupta ve ark., 2021). Geliştirilen sistem bu açığı kapatmayı hedeflemiştir.

Bu tezde, uç birimi üzerinde en iyi performansı gösterecek modeli bulmak için literatürde kendini ispatlamış birbirinden farklı özelliklere sahip nesne tanıma tabanlı derin öğrenme modelleri, oluşturulan karma veri seti üzerinde eğitilerek doğruluk ve FPS açısından karşılaştırılmıştır.

En iyi performansı gösteren yüz algılayıcısı belirlendikten sonra bir video akışı üzerinden yüz görüntüleri elde edilmiştir. Elde edilen bu yüz görüntülerine ise ERT modelinin uygulamasını içeren DLib kütüphanesi kullanılarak yüz işaretleri dedektörü uygulanmış ve yüz noktaları belirlenmiştir.

Diğer ölçüm yöntemlerine kıyasla maliyeti düşük, yüksek doğruluk oranına sahip ve hızlı olduğu için Taylor ve arkadaşları bu çalışma kızıl ötesi ölçüm yöntemi kullanılarak gerçekleştirilmiştir.

Sıcaklık ölçümünün yapılabilceği en uygun alan göz içi olarak bilinmekle birlikte gözün doğrudan ışına maruz kalması kalıcı hasarlara yol açabilmektedir. Dolayısıyla alın bölgesi ölçüm yapmak için göze göre daha uygun bir alandır ve geniş yüzey alanı sunar (Foster ve ark., 2021). Ortalama bir insan yüzü, göz önüne alındığında göz bebeği ve kaş arasındaki uzaklık kaş orta noktasından alına doğru eklendiğinde elde edilen noktayla alın orta noktasının benzer yatay seviyede olduğu gözlemlenmiştir (Yalçınkaya ve ark., 2016). Bu doğrultuda şekil 3.32’de yüz işaretlerinin 20 ve 25 numaralı koordinatları orta nokta formülüne tabi tutularak kaş orta noktası bulunur. Herhangi iki noktanın orta noktasının hesaplanması şekil 3.40’ta gösterilmiştir.



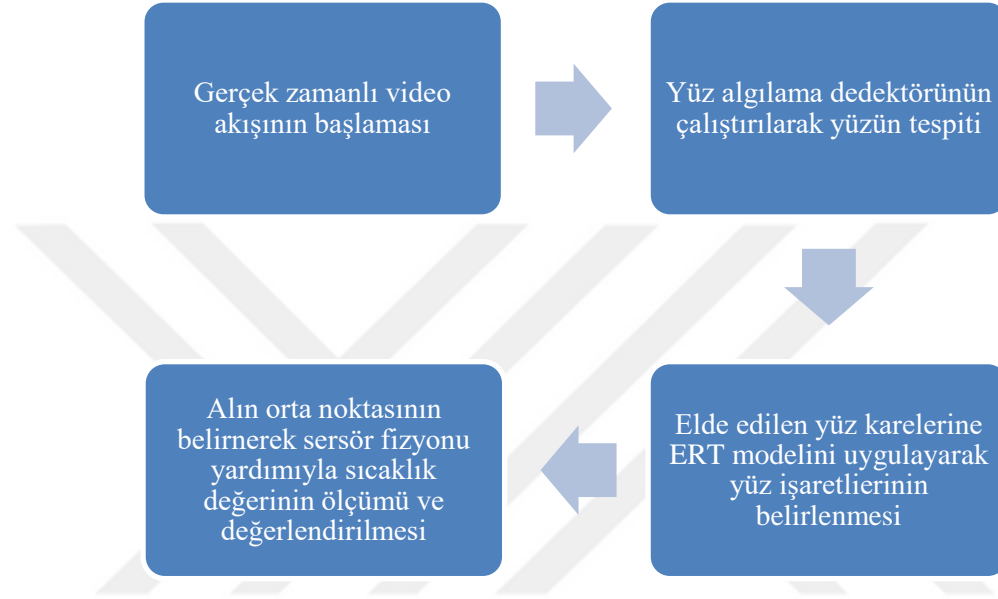
Şekil 3.39. Herhangi iki noktanın orta noktasının bulunması

Daha sonra şekil 3.32’de yüz işaretlerinin 20 ve 42 veya 25 ve 47 numaralı koordinatları iki nokta arası uzaklık formülüne tabi tutulur ve elde edilen sonuç kaş orta noktasına eklenerek alın orta noktası belirlenir. İki nokta arasındaki uzaklık,

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (3.9)$$

formülüyle hesaplanır.

Elde edilen bilgiler ışığında beyin bölgesine yakınlığı, kolay ulaşabilirliği ve giysilerle etkileşiminin azlığı da göz önünde bulundurularak seçilen alın orta noktasından Optris CS-LT kızılötesi termometre (sensör füzyonu) kullanılarak sıcaklık ölçümü gerçekleştirilmiştir. Son olarak, ölçülen sıcaklık değeri 38 °C ve üzerinde olan kişiler riskli olarak değerlendirilir. Şekil 3.40'ta uygulama adımları ve Şekil 3.41'de sistemin nihai çıktıları gösterilmiştir.



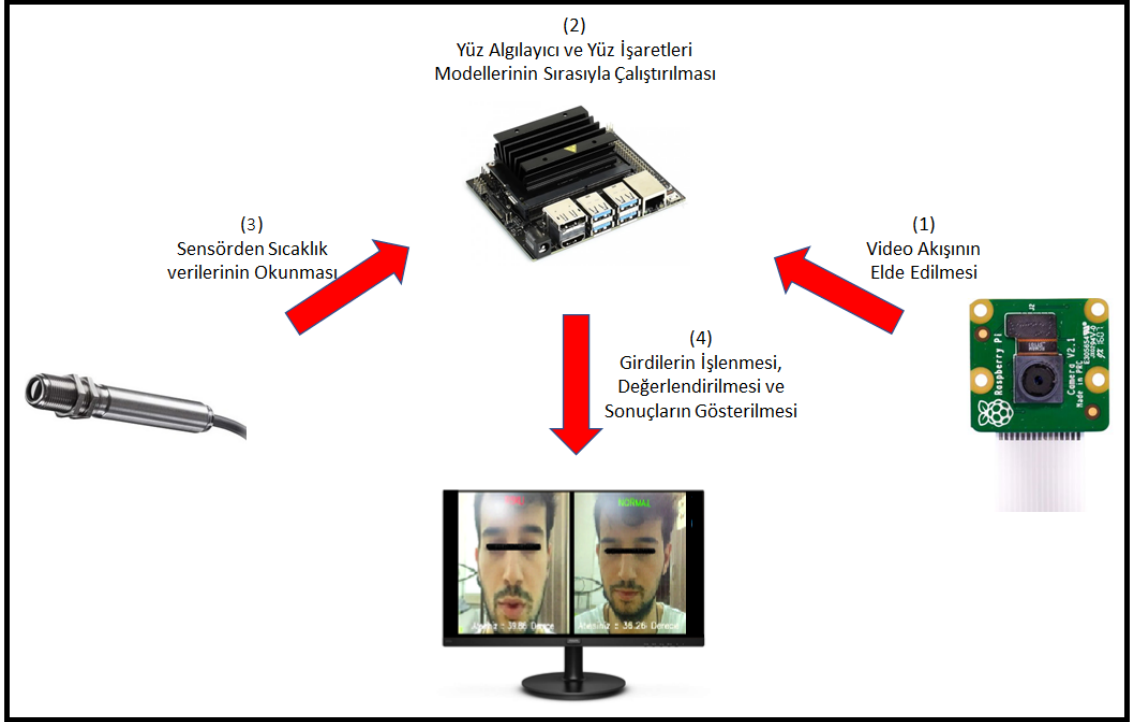
Şekil 3.40.Uygulama adımları



Şekil 3.41. Sistemin nihai çıktıları

Modellerin eğitimleri Google Colab üzerinde oluşturulan Tesla P100 donanımına sahip sanal bir ortam üzerinde gerçekleştirilmiştir. Eğitilen bu modellerin çıkarım işlemi daha önce tanıtılmış olan Nvidia Jetson Nano donanımı üzerinde yapılmıştır. Geliştirme için Python programlama dili seçilmiş ve derin öğrenme kütüphanesi olarak TensorFlow, Caffe ve Pytorch kütüphanelerinden yararlanılmıştır. Buna ek olarak yüz

görüntüleri Raspberry Pi CSI modülü kullanılarak video akışı şeklinde elde edilmiştir. Şekil 3.42.'de sistemin genel çalışma modeli gösterilmektedir



Şekil 3.42. Sistemin genel çalışma modeli

Nvidia Jetson Nano ve dual-sensörlü kamera (RGB+IR) Şekil 3.43.'te gösterilmektedir.



Şekil 3.43. Nvidia Jetson Nano ve dual-sensör (RGB+IR) yapısı

4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA

Bu çalışmada Fast-RCNN, Faster-RCNN, YOLO ve SSD modelleri hazırlanan karma veri seti üzerinde eğitilmiştir. Birbirinden farklı özelliklere sahip nesne tanıma tabanlı bu dört farklı model; farklı aydınlatma koşullarına, tıklama derecesine ve açığa sahip 1000 farklı yüz görüntüsünden oluşan doğrulama test veri seti üzerinde keskinlik, F1 skor gibi performans metriklerinin yanı sıra FPS’ide hesaplayarak değerlendirilmiştir. Şekil 4.1.’de oluşturulan doğrulama veri setinden örnekler gösterilmektedir.



Şekil 4.1. Oluşturulan doğrulama veri seti

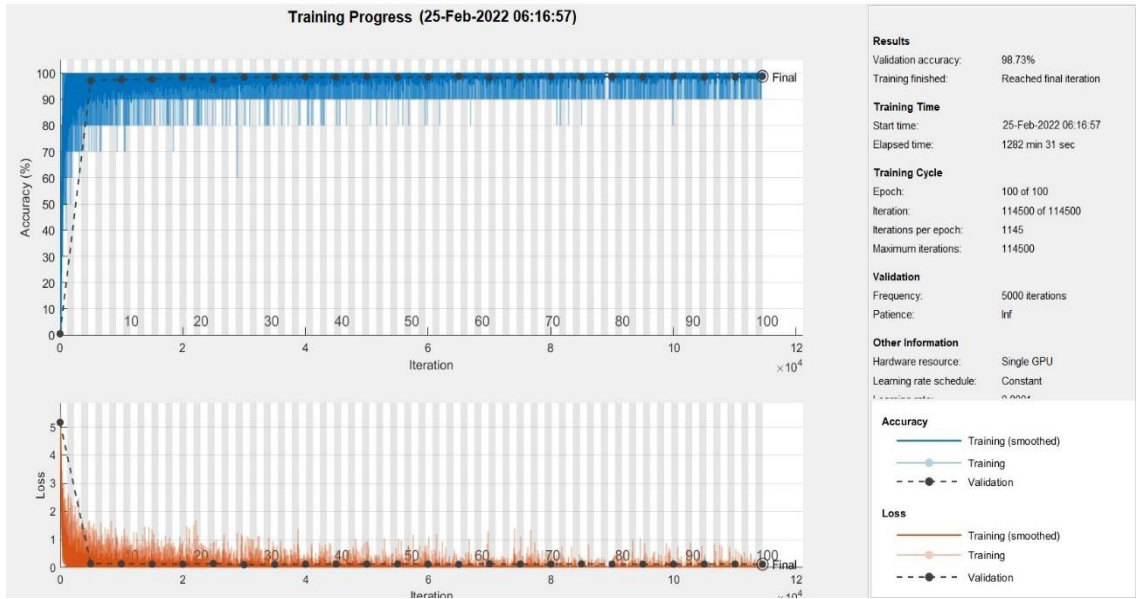
Gerçekleştirilen testlerin tamamı Nvidia Jetson Nano üzerinde yapılmış olup elde edilen sonuçlar bu kartın performansı baz alınarak hesaplanmıştır. Değerlendirilen dört farklı modelin karşılaştırılması tablo 4.1.’de gösterilmiştir.

Tablo 4.1. Modellerin performans metriklerinin ve FPS’lerinin karşılaştırması

Modeller	Doğruluk (%)	Keskinlik (%)	Duyarlılık (%)	F1-skor (%)	FPS
Fast-RCNN	98.26	98.05	98.45	98.25	0.52
Faster-RCNN	98.31	98.08	98.75	98.75	1.02
YOLO (DarkNet-53)	96.36	96.61	96.21	94.41	7.64
SSD (ResNet-10)	96.84	96.97	96.71	96.84	16.49

Tablo 4.1.'de de görüldüğü gibi elde edilen bulgular sonucunda, bölgesel konvolüsyonel sinir ağların YOLO ve SSD benzeri tek geçiş algılayıcılarına göre yaklaşık %1.5 oranında daha yüksek doğruluğa sahip olduğu fakat gerçek zamanlı uygulamalarda uygulanamayacak kadar düşük FPS'te çalıştığı gözlemlenmiştir. YOLO ve SSD modelleriyse kendi arasında kıyaslandığında SSD'nin FPS açısından YOLO dan daha iyi olduğu ve ayrıca yakın mesafeli yüz görüntülerinde yüzün açısına ve tıkaama derecesine daha dayanıklı olduğu gözlemlenmiştir. Bu doğrultuda ResNet-10 alt yapısına sahip özel SSD modelini, yüz algılayıcısı olarak kullanmanın her açıdan daha avantajlı olduğu sonucuna varılmıştır.

Şekil 4.2.'de özel SSD modelinin eğitim ve test süreci gösterilmektedir. Şekilde görüldüğü gibi 100 dönem için eğitim yapılmıştır. Burada bir dönem, tüm eğitim veri setinin ağ üzerinde tam bir döngüsüne karşılık gelmektedir. Öğrenme oranı transfer katmanlarda oranı 0,0001 olarak ayarlamıştır ve böylece bu katmanlarda eğitimin yavaşlatılması amaçlanmıştır. Eğitim sonucunda %98,73 doğruluk oranına sahip bir SSD yüz dedektörü elde edilmiştir.



Şekil 4.2. Doğruluk ve maliyet fonksiyonunun değişim grafiği

SSD modelinin tıkanmaya karşı performansı Şekil 4.3.'te gösterilmektedir.



Şekil 4.3. Eğitilen SSD modelin tıkamalara karşı performansı

Eğitilen SSD modeli sonraki aşmaların gerçekleştirilmesinde sistem yüz dedektörü olarak kullanılmıştır. Dolayısıyla ilerleyen aşamalarda yapılan testlerden yüz dedektörü bağımsız tutulmuştur.

Bu sistem literatürde var olan temassız ölçüm sistemlerine alternatif olarak sunulduğundan sistemin doğruluğu yaygın ölçüm sistemleriyle kıyaslanarak doğrulanmıştır. Önerilen sistemin haricinde tabanca tipi termometre ve koltuk altı dijital termometre olmak üzere iki farklı ölçüm yönetimi baz alınarak 20'si kadın, 20'si erkek 40 farklı birey üzerinde ölçümler yapılmıştır. Eldeki 40 ayrı veriden yola çıkarak varılan sonuçta 24 °C oda sıcaklığında en az 10 dk bekleme koşulu (standart temassız ölçüm yöntemleri) göz önüne alındığında;

- Tabanca tipi kızıl ötesi termometre için kadınların sıcaklık ortalaması 36.8 °C, erkeklerin sıcaklık ortalaması ise 36.4 °C olarak tespit edilmiş olup, grubun toplam sıcaklık ortalaması 36.6 °C olarak,

- Klasik dijital termometre için kadınların sıcaklık ortalaması 35.9 °C, erkeklerin sıcaklık ortalaması ise 37.0 °C olarak tespit edilmiş olup, grubun toplam sıcaklık ortalaması 36.45 °C olarak,
- Önerilen sensör fizyon yöntemi için kadınların sıcaklık ortalaması 36.2 °C, erkeklerin sıcaklık ortalaması ise 36.8 °C olarak tespit edilmiş olup, grubun toplam sıcaklık ortalaması 36.5 °C olarak,

belirlenmiştir. Ayrıca bu ölçüm işlemleri gerçekleştirilirken önerilen ölçüm sisteminin performansını kanıtlamak açısından ölçüm süreleri de not edilmiştir. Önerilen yöntemin farklı ölçüm yöntemleriyle karşılaştırılması Tablo 4.2.'de gösterilmiştir.

Tablo 4.2. Farklı ölçüm yöntemlerine göre sıcaklık karşılaştırması

Termometre Tipi	Vücut Bölümü	Ortalama Sıcaklık (°C)	Ortalama Ölçüm Süresi (s)
Tabanca Tipi Kızıl Ötesi Termometre	Alın	36.6	4
Klasik Dijital Termometre	Koltuk altı	36.45	80
Sensör füzyonu (önerilen yöntem)	Alın	36.0	3

En nihayetinde yapılan bir dizi testler sonucunda önerilen yöntemin sağlam, güvenilir, kullanımı kolay, gerçek zamanlı uygulamalara (~15 FPS) uygun, zaman ve kaynak tasarrufu sağladığı görülmüştür.

5. SONUÇLAR VE ÖNERİLER

Bu çalışmada, Covid-19'dan korunmak için termal değeri ölçerek insan görüsünü destekleyebilecek, derin öğrenme tabanlı bir yüz algılama sistemi önerilmiştir. Önerilen sistemin doğruluğunu ispatlamak ve literatürde var olan yöntemlerden farklılığını ortaya koymak için bir dizi test gerçekleştirilmiştir. İlk olarak kullanılan yüz algılama modelinin doğruluğu ve uç birimi üzerinde çalışma performansı literatürde var olan popüler bir dizi nesne tanıma tabanlı modeli aynı veri setiyle eğitilerek kıyaslanmıştır. Daha sonra önerilen sistem metodolojiye uygun olarak kurgulanmıştır. Son olarak temassız ölçüm yapan ateş ölçerlere bir alternatif olup olamayacağını doğrulamak üzere 20'si kadın 20'si erkek 40 farklı kişi üzerinde ölçüm doğruluğu diğer ölçüm sistemleriyle kıyaslanmıştır. Nitekim elde edilen sonuçlar neticesinde önerilen sistemin mevcut temassız ölçüm sistemlerinin yerini alarak halk sağlığı ölçüm ve analizleri için rahatlıkla kullanılacağı sonucuna varılmıştır.

Bu sistem; havaalanları, hastaneler, kamu binaları, alışveriş merkezleri, eğitim kurumları, adalet sarayları ve ceza infaz kurumları, askeri kurumlar, üniversiteler ve kısacası büyük bir insan nüfusunun olduğu herhangi bir ortamda iş gücünü ortadan kaldırarak minimum hata, maksimum verim sunmaktadır. Ayrıca ölçüm süresini azaltarak sürdürülebilirliği sağlamakta ve personel ile ölçüm yapılan kişi arasındaki doğrudan teması ortadan kaldırmaktadır.

Önerilen sistem mevcut yapısıyla pandemi süreçlerinin yönetiminde etkin olarak kullanılabilir ve hali hazırda kullanılan sistemlerin yerini maliyeti etkin bir şekilde azaltıp verimi artırarak alabilir. Fakat yapılan literatür taraması sonucunda personel katılım (yönetim) sisteminin önerilen çalışmaya uyarlanarak personel kimlikleri üzerinden pandemi kontrolünün daha etkin bir şekilde gerçekleştirilebileceği düşünülmektedir. Bu şekilde yüz algılama modeline eklenecek olan yüz tanıma yapısı giriş çıkış işlemlerinde yapılan kontrol işlemini doğrudan personel kimliği ile eşleştirilerek kişilerin takibi sağlanabilecektir. Böylece önerilen sistem pandemik hastalıkların kontrolünde gelecekte çok daha güçlü bir yapıda olması sağlanarak uygulanabilirliği artırılabilir.

KAYNAKLAR

- Foster, J., Lloyd, A. B., & Havenith, G. (2021). Non-contact infrared assessment of human body temperature: The journal Temperature toolbox. *Temperature*, 8(4), 306-319.
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580-587).
- Girshick, Ross. "Fast r-cnn." In *Proceedings of the IEEE international conference on computer vision*, pp. 1440-1448. 2015.
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning*. vol. 1.
- Gupta, A., Maurya, S., Mehra, N., & Kapil, D. (2021, January). Covid-19: Employee fever detection with thermal camera integrated with attendance management system. In *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 355-361). IEEE.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
- Jiang, H., & Learned-Miller, E. (2017, May). Face detection with the faster R-CNN. In *2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017)* (pp. 650-657). IEEE.
- Karpathy, A., & Fei-Fei, L. (2015). Deep visual-semantic alignments for generating image descriptions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3128-3137).
- Kazemi, V., & Sullivan, J. (2014). One millisecond face alignment with an ensemble of regression trees. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1867-1874).
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).
- Li, A., Li, Y. X. ve Li, X. H., 2017, TensorFlow and Keras-based Convolutional Neural Network in CAT Image Recognition, 2017 2nd International Conference on Computational Modeling, Simulation and Applied Mathematics (CMSAM), 529-533.

- Li, H., Lin, Z., Shen, X., Brandt, J., & Hua, G. (2015). A convolutional neural network cascade for face detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 5325-5334).
- Li, D., Menassa, C. C., & Kamat, V. R. (2018). Non-intrusive interpretation of human thermal comfort through analysis of facial infrared thermography. *Energy and Buildings*, 176, 246-261.
- Liu, Z., Luo, P., Wang, X., & Tang, X. (2015). Deep learning face attributes in the wild. In Proceedings of the IEEE international conference on computer vision (pp. 3730-3738).
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single shot multibox detector, in European conference on computer vision (ECCV).
- Lun, L. T., Chin, T. S., Ishak, M. K., & Asaari, M. S. M. (2022). CLOUD-BASED TEMPERATURE MONITORING MECHANISM FOR COVID-19 PANDEMIC TRACKING. *Jurnal Teknologi*, 84(1), 241-247.
- Maghded, H. S., Ghafoor, K. Z., Sadiq, A. S., Curran, K., Rawat, D. B., & Rabie, K. (2020, August). A novel AI-enabled framework to diagnose coronavirus COVID-19 using smartphone embedded sensors: design study. In 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI) (pp. 180-187). IEEE.
- McClelland, J. L. (2015). Explorations in parallel distributed processing: A handbook of models, programs, and exercises. Stanford Üniversitesi.
- Mesnil, G., Dauphin, Y., Yao, K., Bengio, Y., Deng, L., Hakkani-Tur, D., ... & Zweig, G. (2014). Using recurrent neural networks for slot filling in spoken language understanding. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(3), 530-539.
- Mo, Y. J., Kim, J., Kim, J. K., Mohaisen, A. ve Lee, W., 2017, Performance of Deep Learning Computation with TensorFlow Software Library in GPU-Capable Multi-Core Computing Platforms, 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN 2017), 240-242.
- Mohammed, M., Hazairin, N. A., Al-Zubaidi, S., AK, S., Mustapha, S., & Yusuf, E. (2020). Toward a novel design for coronavirus detection and diagnosis system using IoT based drone technology. *International Journal of Psychosocial Rehabilitation*, 24(7), 2287-2295.
- R. Collobert, L.J.P. van der Maaten, and A. Joulin. Torchnet: An Open-Source Platform for (Deep) Learning Research. ICML Machine Learning Systems Workshop, (2016).

- Ranjan, R., Patel, V. M., & Chellappa, R. (2017). Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. *IEEE transactions on pattern analysis and machine intelligence*, 41(1), 121-135.
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems* (pp. 91-99).
- Redmon J., D. S., Girshick R., Farhadi A., 2016, You Only Look Once: Unified, Real-Time Object Detection.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Somboonkaew, A., Prempee, P., Vuttivong, S., Wetcharungsri, J., Porntheeraphat, S., Chanhorm, S., ... & Sumriddetchkajorn, S. (2017, July). Mobile-platform for automatic fever screening system based on infrared forehead temperature. In *2017 Opto-Electronics and Communications Conference (OECC) and Photonics Global Conference (PGC)* (pp. 1-4). IEEE.
- Sun, G., Nakayama, Y., Dagdanpurev, S., Abe, S., Nishimura, H., Kirimoto, T., & Matsui, T. (2017). Remote sensing of multiple vital signs using a CMOS camera-equipped infrared thermography system and its clinical application in rapidly screening patients with suspected infectious diseases. *International Journal of Infectious Diseases*, 55, 113-117.
- Sun, X., Wu, P., & Hoi, S. C. (2018). Face detection using deep learning: An improved faster RCNN approach. *Neurocomputing*, 299, 42-50.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1), 1929-1958.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).
- Taylor, W., Abbasi, Q. H., Dashtipour, K., Ansari, S., Shah, S. A., Khalid, A., & Imran, M. A. (2020). A Review of the State of the Art in Non-Contact Sensing for COVID-19. *Sensors*, 20(19), 5665.
- Torrey, L., Shavlik, J., Walker, T. ve Maclin, R., 2008, Relational macros for transfer in reinforcement learning, *Inductive Logic Programming*, 4894, 254-257.
- Tuna, K. & Akdemir, B. (2022). Face Detection by Measuring Thermal Value to Avoid Covid-19 . *Avrupa Bilim ve Teknoloji Dergisi* , ICAENS Mayıs 2022 , 191-196 . DOI: 10.31590/ejosat.1113302

- Ulleri, P., Mukund, S., Sudarshan, K., Zenith, K., & NB, S. S. (2021, March). Development of contactless employee management system with mask detection and body temperature measurement using TensorFlow. In 2021 Sixth International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET) (pp. 235-240). IEEE.
- Uijlings, J. R., Van De Sande, K. E., Gevers, T., & Smeulders, A. W. (2013). Selective search for object recognition. *International journal of computer vision*, 104(2), 154-171.
- Yalçınkaya, E., Cingi, C., Söken, H., Ulusoy, S., & Muluk, N. B. (2016). Aesthetic analysis of the ideal eyebrow shape and position. *European Archives of Oto-Rhino-Laryngology*, 273(2), 305-310.
- Yang, S., Luo, P., Loy, C. C., & Tang, X. (2015). From facial parts responses to face detection: A deep learning approach. In *Proceedings of the IEEE international conference on computer vision* (pp. 3676-3684).
- Yang, S., Luo, P., Loy, C. C., & Tang, X. (2016). Wider face: A face detection benchmark. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 5525-5533).
- Yiğit, A. (2017). İş süreçlerinde insan görüşünü derin öğrenme ile destekleme (Master's thesis, Trakya Üniversitesi Fen Bilimleri Enstitüsü).
- Zeiler, M. D., & Fergus, R. (2014, September). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818-833). Springer, Cham.
- Zhang, Z., Luo, P., Loy, C. C., & Tang, X. (2014, September). Facial landmark detection by deep multi-task learning. In *European conference on computer vision* (pp. 94-108). Springer, Cham.
- Wang, Z. H., Horng, G. J., Hsu, T. H., Chen, C. C., & Jong, G. J. (2020). A novel facial thermal feature extraction method for non-contact healthcare system. *IEEE Access*, 8, 86545-86553.