
GAZİ ÜNİVERSİTESİ
Gazi
1926

VİDEO DOSYALARINDA YÜZ TESPİTİ, TANIMA VE VİDEO İÇERİSİNDEKİ ZAMANA GÖRE İŞARETLEME

Hüseyin GÖZE

YÜKSEK LİSANS TEZİ
BİLGİSAYAR BİLİMLERİ ANA BİLİM DALI

GAZİ ÜNİVERSİTESİ
BİLİŞİM ENSTİTÜSÜ

HAZİRAN 2022

ETİK BEYAN

Gazi Üniversitesi Bilişim Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmasında;

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

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

İmza
Hüseyin GÖZE
...../……/……

VİDEO DOSYALARINDA YÜZ TESPİTİ, TANIMA VE VİDEO İÇERİSİNDEKİ ZAMANA GÖRE İŞARETLEME

(Yüksek Lisans Tezi)

Hüseyin GÖZE

GAZİ ÜNİVERSİTESİ

BİLİŞİM ENSTİTÜSÜ

Nisan 2022

ÖZET

Akan video görüntülerinde gerçek zamanlı yüz tespiti, tanımlanması, görüntü içerisinde geçtiği zamanın ve konumun bulunması, video içindeki diğer yüzlerle ilişkilendirilmesi günümüzde yüksek öneme sahip bir konudur. Son yıllarda, hızlı bir şekilde gelişen bilgisayar donanımlarının ve yapay zeka, sinir ağları, derin öğrenme, makine öğrenmesi gibi metotların da yardımıyla yüz tanıma problemlerinde daha hızlı ve doğru sonuçlar almak mümkün hale gelmiştir. Fakat video görüntülerinde aynı kare içerisinde yer alan yüzlerin sayısı ve boyutları, yüzlerin doğru bir şekilde tespit edilmesini ve uygulanan modelin yanıt verme süresini olumsuz yönde etkilemektedir. Gerçekleştirilen bu çalışmada, video görüntüleri içerisinde yer alan yüzlerin daha hızlı ve doğru bir şekilde tespit edilmesi ve video görüntüsü içerisinde bulunma ve birliktelik bilgileri ile birlikte etiketlenmesi için derin öğrenme tabanlı yeni bir model önerilmiştir. Gerçekleştirilen deneysel çalışmada önerilen Evrişimli Sinir Ağı tabanlı modelin daha önce önerilen MTCNN (Multitask Cascaded Convolutional Networks), OPENCV-CNN(Open Source Computer Vision Library - Convolutional Neural Network), HOG+SVM(Histogram of Gradient + Support Vector Machine), SSD-CAFFEMODEL(Single Shot MultiBox Detector with Caffemodel) modellerine göre daha yüksek bir performans ve doğruluk oranına sahip olduğu gösterilmiştir. Gerçekleştirilen çalışmada kullanılan farklı çözünürlükteki video dosyalarının ilgili video görüntüsünün gerçek zamanından 1,54 ile 2,61 kat arasında daha hızlı şekilde işlenmiştir. İşlenme sonucunda tanımlanan kişiler veri tabanına başarılı bir şekilde kayıt edilmiştir. Elde edilen sonuçlara göre ortalama 78,36 FPS değerlerine ulaşan hızlarda yüz tespiti ve tanımlaması yapılmıştır.

Bilim Kodu	:	92432
Anahtar Kelimeler	:	Yüz tespiti, yüz tanıma, zaman işaretleme, derin öğrenme
Sayfa Adedi	:	47
Danışman	:	Doç. Dr. Oktay YILDIZ

FACE DETECTION, RECOGNITION AND TIME MARKING IN VIDEO FILES
(M. SC. THESIS)

Hüseyin GÖZE

GAZİ UNIVERSITY
INSTITUTE OF INFORMATICS
April 2022

ABSTRACT

Real-time face detection and identification in videos, finding them according to the time and location in the video, and associating them with other faces in the video is a matter of high importance nowadays. In recent years, the rapidly growing computer hardware and artificial intelligence, neural networks, deep learning, machine learning methods such as with the help of face recognition problems, it is possible to get faster and more accurate results. But the number and sizes of faces located in the same frame in video images affects negatively the correct detection of faces and the response time of the applied model. In this study, it was carried out to identify the faces in the video images more quickly and accurately and a new model based on deep learning has been suggested for being in a video image and tagging it together with association information. In an experimental study, it was shown that the proposed Convolutional Neural Network-based model has a higher performance and accuracy rate than the MTCNN (Multitask Cascaded Convolutional Networks), OPENCV-CNN(Open Source Computer Vision Library - Convolutional Neural Network), HOG+SVM(Histogram of Gradient + Support Vector Machine), SSD-CAFFEMODEL(Single Shot MultiBox Detector with Caffemodel) models as previously proposed. The video files of different resolutions used in the study were processed between 1.54 and 2.61 decker faster than the real-time of the corresponding video. As a result of the processing, the identified persons were successfully registered in the database. According to the results obtained, face recognition and identification were performed at speeds reaching an average of 78.36 FPS in this extent.

Science Code	:	92432
Key Words	:	Face detection, face recognition, time marking, deep learning
Page Number	:	47
Supervisor	:	Assoc. Prof. Oktay YILDIZ

TEŞEKKÜR

Tez çalışmalarım boyunca kıymetli bilgi, birikim ve tecrübeleri ile bana yol gösterici ve destek olan danışman hocam Doç. Dr. Oktay Yıldız'a tüm yardım ve katkıları için sonsuz teşekkür ve saygılarımı sunarım.

Hayatımın her aşamasında beni destekleyen, her zaman yanımda olan aileme ve yardımlarını hiçbir zaman esirgemeyen değerli çalışma arkadaşlarıma en içten dileklerimle teşekkür ederim.

İÇİNDEKİLER

Sayfa

ÖZET iv
ABSTRACT....v
TEŞEKKÜR....vi
İÇİNDEKİLER vii
ÇİZELGELERİN LİSTESİ....ix
ŞEKİLLERİN LİSTESİ....x
RESİMLERİN LİSTESİ xi
SİMGELER VE KISALTMALAR.... xii
1. GİRİŞ 1
2. LİTERATÜR 5
3. GENEL KAVRAMLAR.... 9
3.1. Görüntü İşleme ve Yüz Tanıma....9
3.2. Makine Öğrenmesi....9
3.2.1 Makine öğrenmesinde sıklıkla kullanılan algoritmalar 12
3.3. Derin Öğrenme 13
3.3.1. McCulloch ve Pitts modeli 20
3.3.2. Hebb kuralı 22
3.3.3. Perseptron 23
3.3.4. İleri yayılım....23
3.3.5. Geri yayılım 24
3.3.6. Aktivasyon fonksiyonları....25
3.3.7. Optimizasyon algoritmaları 25
3.4. Model Performans Değerlendirme....26
4. VİDEO GÖRÜNTÜLERİNDE YÜZ TESPİTİ VE ZAMAN İŞARETLEME....27
4.1. Veri Kümesi....27
4.2. Deneysel Çalışma 28

Sayfa

5. DENEYSEL ÇALIŞMA SONUÇLARI.....35
6. SONUÇ VE DEĞERLENDİRME.....41
KAYNAKLAR.....43
ÖZGEÇMİŞ.....47

ÇİZELGELERİN LİSTESİ

Çizelge **Sayfa**

Çizelge 3.1. Aktivasyon fonksiyonları..... 25

Çizelge 5.1. Uygulanan mimarilerin karşılaştırma tablosu..... 37

Çizelge 5.2. Uygulanan mimarinin parametre tablosu..... 38

Çizelge 5.3. Önerilen modelin karmaşıklık matrisi 40

ŞEKİLLERİN LİSTESİ

Şekil **Sayfa**

Şekil 3.1. Temel bir ESA mimarisi 13

Şekil 3.2. Biyolojik sinir hücresi 14

Şekil 3.3. Yapay sinir ağları 14

Şekil 3.4. ESA mimarisinde bulunan gizli nöronların yapısı 17

Şekil 3.5. ESA mimarisinde kullanılan havuzlama katmanı 18

Şekil 3.6. 28 katmanlı Evrişimli Sinir Ağları modeli 19

Şekil 3.7. 34 katmanlı ResNet modeli 20

Şekil 3.8. ResNet mimarisi……. 18

Şekil 3.9. Matematiksel sinir hücresi modeli 21

Şekil 3.10. Perseptron yapısı ……..... 23

Şekil 3.11. Yapay sinir ağı ve katmanları 24

Şekil 4.1. Önerilen sistem mimarisi 28

Şekil 4.2. Uygulanan algoritmanın sözde kodu 30

Şekil 4.3. Tespit edilen yüz örüntüsü 31

Şekil 4.4. Uygulanan model ve tespit edilen yüz sayısı 32

Şekil 4.5. Uygulanan mimarilerin çalışma süreleri 33

Şekil 5.1. Video işlemesi sonucunda oluşan veri tabanı kayıtları 35

Şekil 5.2. Önerilen modelin farklı video çözünürlüklerine ait işleme hızları 36

Şekil 5.3. Önerilen modelin eğitim, test ve doğruluk grafiği 39

RESİMLERİN LİSTESİ

Resim **Sayfa**

Resim 4.1. Labeled Faces in the Wild Home veri kümesi yapısı 27

Resim 5.1. Örnek görüntü içerisinde tespit edilen yüzler 39

Resim 5.2. Örnek görüntü içerisinde tespit edilemeyen yüzler 40

SİMGELER VE KISALTMALAR

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

Simgeler	**Açıklamalar**
GB	Gigabyte
GHz	Gigahertz
σ	Sigma

Kısaltmalar	**Açıklamalar**
CACHE	Uygulama ön belleği
CAFFEMODEL	Derin öğrenme modeli
CNN	Evrişimsel sinir ağları (Convolutional neural network)
CPU	Merkezi işlem birimi (Central process unit)
ESA	Evrişimli sinir ağı
FPS	Bir saniyedeki kare sayısı (Frame per second)
HOG	Yönlendirilmiş gradyanların histogramı (Histogram of oriented gradients)
JPG	Birleşmiş fotoğraf uzmanları grubu (Joint photographic experts group)
LTSM	Uzun kısa-süreli bellek (Long short-term memory)
MTCNN	Çok görevli basamaklı evrişimsel ağlar (Multi-task cascaded convolutional networks)
OPENCV	Açık kaynak bilgisayarla görme (Open source computer vision)
RAM	Rasgele erişim belleği (Random access memory)
SSD	Tek çekim dedektörü (Single shot detector)
SVM	Güvenli sanal makine (Secure virtual machine)
JSON	JavaScript nesnesi gösterimi (JavaScript object notation)
NoSql	İlişkisel olmayan sorgulama dili (Non-relational SQL)

1. GİRİŞ

Görüntü işleme herhangi bir aygıt aracılığı ile kayıt altına alınmış görüntü veya videolardaki nesne veya örüntüleri tarama, tanımlama ve anlamlandırma amacıyla uzun yıllardır bilgisayar biliminin ve araştırmacılarının üzerinde çalıştığı güncel bir konudur. Görüntü işleme uygulamalarının temel araştırma konuları arasında görüntü içerisinde tespit edilen ve tanımlanan nesne veya örüntülerin sınıflandırılması ve gruplandırılması konuları da yer almaktadır. Günümüzde görüntü işleme çalışmaları ve uygulamalarının birçok sektörde ihtiyaç duyulması ve kullanılması nedeni ile popülerliği artarak devam etmektedir.

Son yıllarda gelişen donanım teknolojileri ve bu teknolojilere erişimin her kesime yayılması sayesinde özellikle internet ortamında bulunan sosyal medya platformları, kişisel görüntü arşivleri ve özel veya kamusal güvenlik amaçlı uygulamalar ile üretilen ve depolanan görüntülerin sayı ve boyutları her geçen gün artmaktadır. Nitelik ve nicelik bakımından yüksek hızla büyüyen bu görüntüler içerisinde yer alan kişi veya kişilerin görüntü içerisinde tespit edilmesi, tanımlanması, ihtiyaca göre sınıflandırılması ve doğru bir şekilde etiketlenmesi günümüzde yüksek öneme sahip bir konu haline gelmiştir. Herhangi bir dijital görüntü içerisinde yer alan kişilerin yüz tespiti, yaş tespiti, cinsiyet tespiti veya duygu tespiti gibi farklı durumlar üzerinde çeşitli çalışmalar yapılmıştır. Özellikle görüntü içerisinde bulunan kişilerin yüz tespiti ve tanımlanması, görüntü karesi içerisinde bulunan tüm kişilerin tespit edilmesi ve tanımlanması gereken durumlarda yüksek öneme sahiptir. Farklı format ve çözünürlüklere sahip özellikle video görüntüleri içerisinde geleneksel yöntemler kullanılarak insan gözünün ayırt edemeyeceği kadar yüksek sayıdaki yüzün tespiti ve tanımlanması mümkün olamamaktadır. Ayrıca geleneksel video görüntüleri taramalarında çok uzun zamanlar kaybedilmekte, ilgili durumun tespiti zorlaşmakta ve yüksek maliyetler ortaya çıkmaktadır. Bu türden durumlarda özellikle video görüntüleri için otomatik tarama, tespit ve tanımlama işlemlerinin daha az süre ve maliyet ile daha doğru bir şekilde yapılabilmesi ihtiyacı her geçen gün artmaktadır. Bu problemin çözümü için zaman içerisinde birçok farklı yol, yordam ve algoritma geliştirilmiş ve uygulanmıştır. Farklı sonuç ve performansa sahip bu yöntemlerin en doğru sonucu en hızlı ve en düşük maliyet ile verebilmesi beklenmektedir.

Tezin ilk bölümü, çalışmanın geneli hakkında bilgi veren giriş bölümüdür. İkinci bölümde konu ile ilgili yapılan literatür taraması hakkında bilgiler verilmiştir. Üçüncü bölümde Görüntü işleme, makine öğrenmesi ve derin öğrenme hakkında bilgilere yer verilmiştir. Dördüncü bölümde veri seti, gerçekleştirilen deneysel çalışma ve oluşturulan mimari ile ilgili bilgiler bulunmaktadır. Beşinci bölümde gerçekleştirilen deneysel çalışma ile elde edilen sonuçlar gösterilmiştir. Altıncı bölümde de elde edilen sonuçlar çerçevesinde bir değerlendirme yapılmıştır.

Tezin Amacı

Bu tez kapsamında akan video görüntülerinde kişi ve nesnelere ait yüz tanımlama, tespit etme, görüntüyü anlamlandırma, video içerisinde görüntünün geçtiği zamanı bulma, tespit edilen kişileri etiketleme, birden fazla kişinin aynı anda aynı karede yer alması durumunda karedeki yüzleri ayrı ayrı tespit etme işlemlerini gerçek zamanlı yapabilen derin öğrenme tabanlı yeni bir model önerilmiştir. İnsan gözünün zorlanacağı durumlarda manuel tespitten daha hızlı ve doğru tanımlama ile çalışabilecek bu model için deneysel çalışmalar gerçekleştirilmiştir. Bu deneysel çalışmalarda daha önce önerilen MTCNN, OPENCV-CNN, HOG+SVM, SSD-CAFFEMODEL modelleri ile bu çalışmada önerilen Evrişimli Sinir Ağı tabanlı modeli karşılaştırılmıştır. Karşılaştırmaya dahil edilen modellerin hem çalışma hızı hem de doğru tespit oranları açısından avantaj ve dezavantajları irdelenmiştir. Elde edilen sonuçlara göre önerilen Evrişimli Sinir Ağı tabanlı modelin diğer modellere göre daha yüksek bir performans ve doğruluk oranı ile çalıştığı gösterilmiştir.

Hipotez

Gerçekleştirilen çalışmanın ele aldığı ana problem, farklı boyut ve çözünürlüklere ait akan video görüntülerinde bulunan kişilerin görüntü karesi bazında çoklu tespiti, tanımlanması ve zaman-konum damgası ile etiketlenmesidir. Aynı görüntü karesi içinde birden fazla kişiye ait yüzün bulunması, ilgili yüzün görüntü karesi içerisindeki konumu ve büyüklüğü, yüzün pozisyonu ve aynı kare içerisinde bulunan yüz sayısına bağlı olarak tespit ve tanımlama hızı dramatik şekilde düşmektedir. Yüz tespit ve tanımlamasının doğru şekilde yapılmasına rağmen çalışma hızının yukarıda bahsedilen durumlara bağlı olarak düşmesi istenmeyen bir durumdur. Diğer yandan da bazı durumlarda çalışma hızının yüksek olmasına rağmen doğru

tespit ve tanımlama oranının düşmesi de yapılan taramaların doğru sonuç üretmemesi nedeniyle anlamsız kılmaktadır. Bu çalışma her iki duruma da optimum düzeyde cevap verebilecek bir mimari geliştirilmiştir.

2. LİTERATÜR

Son yıllarda görüntü işleme ve analizi alanında Derin Öğrenme modellerinin kullanımında artış olduğu ve daha sık uygulandığı görülmektedir [1-4]. 1943 yılının ilk çeyreğinde McCulloch ve Pitts [5] adlı bilim insanları tarafından ortaya atılmış olan Derin Öğrenme, insan beynini referans alan, işleyişini inceleyip modelleyen ve ilgili verileri ise hesaplayan bir yöntemidir. Zaman içerisinde bu yöntemin daha fazla destek görmesi ve daha fazla kullanılmış olması yöntemin gelişim göstermesini ve birçok alanda kullanılmasını sağlamıştır. Pitts ve McCulloch tarafından ortaya atılmış olan bu yöntem, gösterdiği gelişim ve uygulanabilirliği ile günden güne daha fazla kullanılmaya başlanmıştır. 1997 yılında Hochreiter ve Schmidhuber tarafından önerilmiş olan LTSM, bilginin saklanması ve korunması probleminin çözülebilmesi ve bunun öğrenilmesi amacıyla geri yayılımla bilgiyi periyodik tekrarlayan bir yöntemdir [6]. 2012 yılına gelindiğinde Google ekiplerinin üzerinde uzun yıllar çalıştığı yüksek işlem gücüne sahip yapay desen tanıma algoritmalarının performansı derin öğrenme için yapılmış en büyük atılım olmuştur. Böylelikle Derin Öğrenme performansı ciddi seviyelerde artış göstererek insanların algı seviyelerine yaklaşmıştır [7].

Dijital görüntülerdeki nesnelerin analiz edilmesi ve bunun hem etkin hem de hızlı bir şekilde tespitinin sağlanması için geliştirilen ilk algoritma Viola Jones olmuştur [8]. Donanımların hızlı ve daha efektif çalışma konusundaki gelişimi, birçok konuda iyileşme ve büyümeyi tetiklemiştir. Bu donanımların gelişmesi ile Evrişimli Sinir Ağı (Convolutional Neural Network – CNN) tabanlı çözümler daha etkin çalışmaya ve tercih edilmeye başlanmıştır. Evrişimli Sinir Ağı tabanlı yaklaşımlar sayesinde yaşanan problemlerin giderilmesi amacıyla yüksek doğruluk oranlarına sahip yeni ve çeşitli çözüm yolları bulunmuştur. Sırasıyla görüntünün; ön işlemlerden geçirilip hazırlanması, nesnenin uyumluluk tespiti, elde edilen nesnelerin sınıflandırılıp kategorize edilmesi ve nesnelerin takibi Evrişimli Sinir Ağı tabanlı yöntemlerin problemleri genellikle ele alıp incelediği 4 ana adımdır [9]. Görüntülerde yer alan nesnelerin tespiti problemlerin çözümleri için büyük öneme sahiptir. Nesnelerin tespitindeki başarı oranı, daha sonraki aşamaların başarı oranlarını birebir etkilemektedir.

Derin öğrenme tabanlı görüntü işleme uygulamaları günümüzde oldukça yaygın kullanılmaktadır. Luo, X. ve arkadaşlarının çalışmalarında geliştirdiği araç tespiti ve akıllı trafik izleme sistemi de derin öğrenme tabanlı görüntü işleme uygulamasıdır [10]. Luo, X ve arkadaşlarının çalışmasında en az 9 katmandan oluşan bir evrişimli sinir ağı önerilip tercih edilmiştir. Bu uygulama için içerisinde araçlar bulunan birçok farklı açılardan alınmış görüntü dosyaları kullanılmıştır. Önerilen 9 katmanlı yöntem şu katmanları içermektedir; evrişim katmanları, havuzlama katmanları ve tam bağlantılı katmanlar. Nesnelerin özelliklerini çıkartmadan ve sesli görüntüler için ayrıca gürültü gidermeden Evrişim katmanları sorumludur. Parametrelerin azaltılması ve işlem hızının arttırılmasını sağlamak Havuzlama katmanın görevidir. Sınıflandırma ve regresyondan sorumlu olan katman ise tamamen bağlantılı katmanlardır. Uygulama için yaklaşık olarak 90.000 görüntü içeren veri kümesi kullanılmıştır. Geliştirme aşamasında yapılan çalışmalarda yaklaşık olarak belirtilen verilerin eğitim için %80'i, test için %20'si kullanılmıştır. Önerilen dokuz katmanlı model ile çalışma tamamlandığında %97,51 oranında başarı sağlanmıştır.

Yüz tanıma uygulamaları için Derin Öğrenme tercih edilebilir durumdadır. Prav, K. ve arkadaşları yüksek çözünürlüğe sahip görüntü üreten cihazlar için yüz yakalama ve tanımlama yapılabilmesi için ESA tabanlı bir mimari önerisinde bulunmuşlardır [11]. Veri kümesi olarak bu yöntem de AT&T yüz veri kümesi kullanılmıştır. Basitleştiriliş olarak, önerilen ESA tabanlı mimarinin katmanları şu şekildedir; INPUT, CONV, RELU, POOL, FC'dir. Görüntülerin ham piksel değerleri INPUT katmanında tutulmaktadır. CONV katmanı, pencereli görüntü üzerinde evrişim işlemini yapabilmek için bir pencere tarzında kayan sabit boyutlu bir çekirdek veya filtreden oluşmaktadır ve bu şekilde özellikleri çıkarır. Görüntü eşleşmesi filtre boyutuna uygun olmaması durumunda bu problemi çözmek için giriş görüntüsü boyutuna dolgu uygulanır. RELU ise etkinleştirme işlevi olan düzeltilmiş doğrusal birimler anlamına gelmektedir. Bu işlevi görüntünün gizli birimlerine sıfır değeri atayarak sağlar. Verileri işlemek için gerekli olan hesaplama gücünü azaltan aşağı örneklemelerden ve boyutsallığının azaltılmasından sorumlu olan havuzlama katmanı, POOL'dur. Ayrıca havuzlama katmanı baskın özelliklerinden dönme ve konumsal olarak sabit olanları çıkarmak amacıyla girdinin üzerinde bir pencere gibi kayan bir çekirdeğe veya işleve sahiptir. Havuzlama katmanında kullanılan iki yaygın işlevler maksimum havuzlama ve ortalama havuzlamadır. Girdideki oluşturulmuş olan her sinir ağının çıktıdaki her bir sinir ağına bağlandığı katman FC katmanıdır. Ayrıca bu katman, belirli bir sınıfın puanını

hesaplamaktan sorumludur tam bağlantılı katmandır ve bu hesaplama n'nin sınıflandırılacak (sınıf / kategori) sayısını gösterdiği kabul edildiğinde n tane çıktı ile sonuçlanır. ESA mimarisinin öngörülen sınıfı en yüksek puana sahip sınıf olarak belirlenir. FC katmanı DENSE katmanı olarak da adlandırılmıştır. Sistem tasarım gereksinimlerine ve sistemin performansına göre ESA mimarisinin değiştirilebileceği ve düzenlenebileceği ortaya konulmuştur. ESA mimarisinde birçok katman bulunmaktadır. Bu katmanların bazıları DROPOUT ve FLATTEN içermektedir. ESA'nın aşırı uyumunu ve kontrolünü önlemek için bir düzenleme tekniği olan DROPOUT katmanı, burada bulunan girişlerin bir kısmı (bırakma oranı olarak adlandırılmaktadır), eğitim sırasındaki yapılan her güncellemede değerleri 0 olarak eğitimden çıkarılır. Böylece elde edilmiş olan girdilerin değerleri ölçeklendirilmiş olur ve toplamları eğitim sırasında değişmez. İki boyutlu olan verilerin tek boyuta düşürülmesi FLATTEN katmanlar tarafından yapılmaktadır. Bu nedenle FLATTEN katmanlar FC katmanlarından önce eklenmektedir. Viola Jones algoritması, bu çalışmada kameradan alınan canlı yayın akışı sırasında yakalanan yüzleri algılamak için kullanılmıştır. Akış sırasında elde edilen yüz görüntüsü kırpıldıktan sonra 120x120 piksel ölçülerinde yeniden boyutlandırılır. Kırpılıp yeniden ölçeklendirildikten sonra bu görüntü ESA (ilk 32 katman) katmanlarına iletilir. 3x3 piksel boyutundaki görüntüler bu katmandan elde edildikten sonra diğer katmana iletilerek tasarlanmış olan pooling ve dropout katmanlarından geçirilir. Tanımlama, belirlenen başarım parametrelerine göre gerçekleştirilmiştir. Bu model AT&T veri kümesi üzerinde çalıştırılmıştır ve en yüksek doğru tanımlama oranı olan %98,75'e varan değere ulaşılmıştır.

Yapmış oldukları çalışmalarda elde ettikleri sonuçlar ile Mahmood, Z. ve arkadaşları, Yüz Tanıma (FR) algoritmalarına genel bir bakış sunmuşlardır. Bu çalışmada her yaklaşımın kendine ait olan tanıma oranı ile ilgili olarak görüntü veri tabanlarının koşullarına da değinilmektedir. Yüz tanıma ile alakalı olan güncel metodolojilere ve genel kurallara hızlı bir bakış için bilgi kaynağı olarak yaklaşılmasının yanı sıra algoritma tercihleri açısından geniş çaplı bir ön bilgi kaynağı olarak düşünülen yine bu çalışma olmuştur [12]. Çalışmada yer alan deneyler sırasında, 300 adet insan yüzü görüntüleri, FERET veri tabanından eğitim amacıyla rastgele seçilmiştir. Eğitim aşamasında 300 adet resimden 80 tanesi kullanılmıştır. Etiketleme için eğitim kümesinde yer alan 80 tane görüntünün tamamının doğru şekilde etiketlenmiş olduğu ve geri kalan 220 görüntü içerisinde 207 görüntünün doğru bir şekilde belirlendiği görülmüştür. Elde edilen ortalama sonucunda %95,67 oranında bir tanıma oranı

görülmüştür. Deneyler, FERET veri tabanından sonra 3 veri tabanı ile devam etmiştir. CAS-PEALR1, LFW ve HFB veri tabanları üzerinde gerçekleştirilen deneylerin elde edilen ortalama doğruluk oranları sırasıyla %86,8, %84,02 ve %92,2 olmuştur. FERET bu veri tabanlarından en büyük başarım yakalanan veri tabanı olmuştur.

Gerçekleştirilmiş olan bu çalışmada, video görüntüleri içerisinde geçen tanınan ya da tanınmayan kişileri, düşük sistem ve donanım gereksinimi ile birlikte gerçek zamanlı olarak tanıma ve yakalama işlemi yapabilen bir mimari geliştirilmiştir. Anlık olarak tanınan kişiler için yakalama, etiketleme ve belirtme işlemleri yapılabilmesinin yanı sıra, tanınmamış kişileri ayrı bir havuzda tutması sağlanmıştır. Havuzda tutulan bu kişilerin, uygulamanın devam eden aşamalarında tekrar tanıması ve sisteme dahil etmesi sağlanmıştır. Bu sayede düşük donanım koşulları, düşük işlem gücü, minimum zaman gereksinimi ve iş gücü maliyeti ile daima öğrenen ve geliştirilmeye açık, sürekli tanıma ve etiketleme mantıkları çerçevesine bir mimari oluşturulmuştur. Bu mimari zaman içerisinde gerçek zamanlı tespit gerektiren koşullar ve durumlar için kullanılabilecektir.

3. GENEL KAVRAMLAR

3.1. Görüntü İşleme ve Yüz Tanıma

Günümüzde yaygın olarak aygıtlar aracılığı ile kayıt altına alınan görüntülerin netliğinin arttırılması, görüntüden bir nesne veya kişi elde edilmesi gibi işlemleri sağlayan tekniklere görüntü işleme denir. Görüntünün işlenebilmesi için önce sayısallaştırılması gerekmektedir. Sayısallaştırma işlemi ise, görüntüdeki renklerin sayısal değerler ile ifade edilmesiyle yapılır. Görüntü İşleme tıp, trafik, astronomi, güvenlik ve askeri endüstriler gibi birçok farklı alanda yoğun olarak kullanılmaktadır. Görüntü işleme dijital ortamda bulunan bir fotoğraf üzerinde yapılabileceği gibi dijital fotoğraf dizisinden oluşan video dosyaları üzerinde de yapılabilmektedir.

Yüz tanıma ise, yine bir aygıt aracılığı ile alınmış olan görüntünün içerisinde yer alan kişilerin yüz bölgesinin bulunduktan sonra veri tabanında önceden kaydedilmiş diğer resimler ile karşılaştırılması ve eşleştirme yapmasıdır. Mevcut resimler ile yeni bulunan resim arasında benzerlik aranırken, bulunan benzerlik en yüksek oranda olan ile eşleşme yapılır.

3.2. Makine Öğrenmesi

Gün geçtikçe önemi artan Makine Öğrenmesi, insanların öğrenme yöntemlerini taklit ederek kullanılan verilere göre makinenin eğitilmesi ve performansının iyileştirilmesi üzerine çalışmalar yapılan bir yapay zeka alt kümesi ve bilgisayar bilimi dalıdır. 1950 yılında yazılan "Makinelerin İşleyişi ve Zeka" makalesinde bahsi geçen "Turing Testi" ile makine öğrenimi Alan Turing tarafından öngörülmüş ve bir çok çalışmaya öncülük etmiştir [20].

İlk kez 1959 yılında Arthur Samuel, dama oyunu ile ilgili araştırmaları sırasında "Makine Öğrenmesi" terimini ortaya atmıştır. Samuel, bu fikri ortaya attığında "Program yazan kişinin oynayabileceğinden daha iyi dama oyunu oynamayı öğrenecek." demiştir. 1962 yılında, uzman dama oyuncusu olan Robert Nealey, IBM 7094 bilgisayarına karşı dama oyununu kaybetmiştir.

Henry J. Kelly'nin 1960 yılında yayınladığı "Optimal Uçuş Noktaları için İrtifa Kuramı" makalesi, Kontrol Kuramı ilgili birçok fikri ortaya koymuştur. Girdili sistemlerin davranışları ve geri bildirimler ile bu davranışların nasıl değiştirilebileceği hakkında yazılanlar ile doğrudan Yapay Zeka ve Yapay Sinir Ağları alanlarına uygulanabilmiştir. Ayrıca, sinir ağları eğitiminde kullanılan sürekli geri yayılım modellerinin temellerinin atılımını da etkilemiştir.

George Devol tarafından icat edilen ve dünyanın ilk yapay zeka kullanan endüstriyel robotu olan Unimate, 1961 yılında New Jersey'deki General Motors fabrikasında kullanılmıştır. İnsan gücü gerektiren işlerde robotların kullanılmaya başlamış olması, günümüzde de devam eden tartışmaları başlatmıştır.

1966 yılında Joseph Weizenbaum tarafından MIT Yapay Zeka Laboratuvarı'nda geliştirilen ilkel doğal dil işleme çalışması olan ELIZA, psikolog gibi tasarlanmış ve iletişim esnasında karşısındaki ile insani diyaloglar kurması planlanmıştır [20]. Ancak, henüz doğal dil işleme ve makine öğrenmesi teknolojilerinin gelişmemiş olmasından dolayı konuşmalar sırasında sohbet ediyor gibi görünmesine karşın, sadece karşısındaki kişinin cümlelerinde geçen kelime eşleştirmeleri ile cevaplama yapmıştır. İnsan - makine iletişimini sağlamak amacıyla metin veya konuşma yoluyla diyalog kurma yazılımı olan Chatbot için kabul edilen ilk geliştirme ELIZA olmuştur.

Shakey, 1966 yılında geliştirilen ve kendi hareketleri hakkında akıl yürütebilen ilk mobil robot olmuştur [20]. Diğer robotlar gibi görevleri gerçekleştirirken her adımda komut verilmesine gerek duymadan bir görevi analiz ederek parçalara ayırabilmiştir. Gelişimi, birçok robotik, yapay zeka ve bilgisayar bilimi çalışmalarına katkı sağlamıştır.

Deep Blue, IBM'in satranç oynayan yapay zekası 1997 yılında dönemin en iyi satranç ustası olan Garry Kasparov'u altı maçlık seride yenmiştir [22]. Bu olay ile yapay zeka tekrar gündeme gelerek Makine Öğrenme için yapılan çalışmaları duraklama döneminden çıkarmıştır.

Cynthia Breazeal tarafından 1998'de geliştirilen, insan duygularını tanıyabilen ve simüle edebilen robot kafasına KISMET adı verilmiştir. İnsanlarla iletişim kurabilmesi için işitsel

ve görsel yetenekleri sağlayan aygıtlar sayesinde yüz ifadeleri ve hareketleri yaparak duyguları simüle edebilmiştir. Dünyanın ‘’ilk sosyal robotu’’ unvanı ile ödüllendirilmiştir [20].

AIBO, 1999’da Sony tarafından piyasaya sürülmüştür. Bir robot köpek olan AIBO, ilerleyen yıllarda geliştirilerek yeni yetkinlikler kazandırılmıştır.

2002 yılına gelindiğinde Roomba adında zemin alanında gezinebilen sensörlere sahip otonom robotik elektrikli süpürge geliştirilmiştir. İlerleyen yıllarda robotun özellikleri geliştirilmiştir [20].

Siri adı verilen yapay zekaya sahip telefon asistanı Apple firması tarafından 2011’de duyurulmuştur. Ses ile kontrol edilebilmesi sayesinde tarihte ilklerden olmuştur.

IBM tarafından 2011 yılında Watson adı verilen bir bot geliştirilmiştir. Soruları cevaplayabilmesi için tasarlanan bu yapay zeka programı, “Jeopardy” isimli TV programında $ 1 Milyon dolarlık soru yarışmasında birinci olmuştur [20].

2014 yılında, Eugene Goostman, Turing Testi ilk defa geçebilen yapay zeka olmuştur. Testin geçme koşulu, deneye katılan kişilerin karşısındaki makine ile diyalog kurduğunda, makinenin gerçek insan olduğunu düşünmeleri gerekmekteydi. Deney sırasında, insanlar Eugene’nin gerçek insan olduğu görüşüne varmışlardır. Yine aynı yılda, Amazon, Amazon-Alexa adlı bir çevrimiçi sanal asistanı piyasaya sürmüştür [20].

Microsoft tarafından 2016 yılında geliştirilen yapay zekaya sahip sohbet botuna Tay adı verilmiştir. “Twitter” adlı yaygın kullanılan bir sosyal medya platformunda bir kullanıcı hesabı ile kullanıma açılmıştır. Diğer platform kullanıcılarının paylaştığı veriler ile çıkarımlarda bulunarak kendini geliştirmiştir. Tay, paylaşımlarında ırkçı söylemleri arttırdığı için büyük tepki görmüş ve Microsoft tarafından kaldırılmıştır [20].

2017 yılında, Google DeepMind tarafından geliştirilen yapay zeka Dünya Go şampiyonu olan Ke Jie’yi yenmiştir. AlphaGo adı verilen bu yapay zeka, 19x19'luk tahtada avantaj verilmeden profesyonel bir Go oyuncusunu yenen ilk bilgisayar programı olmuştur.

Go oyunu, satranç ve dama gibi oyunlara göre bilgisayar tarafından kazanılması zor bir oyundur çünkü oyunda yaklaşık olarak 2^{170} olası hamle bulunmaktadır. DeepMind şirketi CEO'su Demis Hassabis, AlphaGo'ya karşı oynadığı oyunda AlphaGo'yu tüm limitlerine kadar zorlayarak güzel bir oyun çıkardığı için Ke Jei'ye büyük saygı duyulması gerektiğine değinmiştir [20].

3.2.1 Makine öğrenmesinde sıklıkla kullanılan algoritmalar

Makine öğrenmesi için kullanılan farklı algoritmalar vardır. Algoritma seçimleri girdilere, performanslara, kullanım amacına göre değişiklik göstermektedir. Sıklıkla kullanılan algoritmalar arasında; Naive Bayes, en yakın k komşu, karar ağaçları, random forest, gradient boosting trees algoritmaları yer almaktadır.

Naive Bayes, temel seviyede sınıflandırma için önerilen ve diğer algoritmalarla karşılaştırmak için kullanılan bir istatistiksel öğrenme algoritmasıdır. Çalışma mantığına göre, belirlenmiş bir sınıf için girdileri birbirinden bağımsız kabul eder ve sınıf koşulu olasılıkları tarafsız tahmin eder. Birbiri ile ilişkili girdiler olmadığı sürece diğer algoritmalara oranla iyi performans gösterir [32].

En yakın k komşu algoritması, sınıflandırılmak istenen bir bireyin diğer k tane bireye olan yakınlığına bakıldığı bir algoritmadır. Komşuluk mesafesini hesaplamak için Öklid uzunluğu kullanılmaktadır. Komşu örneklerin oy çoğunluğuna dayalı tahminler yapılır [32].

Karar ağaçları, bir ağaç yapısına benzetilerek çalışılan bir algoritmadır. Ağacın başlangıçtan çıkan kolları ile yaprak seviyesine kadar giden sınıf etiketleri arasındaki işlemleri ifade edilir. Temel algoritmalardan biridir. Anlaşılması ve yorumlanabilir olması açısından basit bir algoritmadır. Kullanımı hem sayısal hem de sınıfsal verilerin işlenebilmesi açısından uygundur [32].

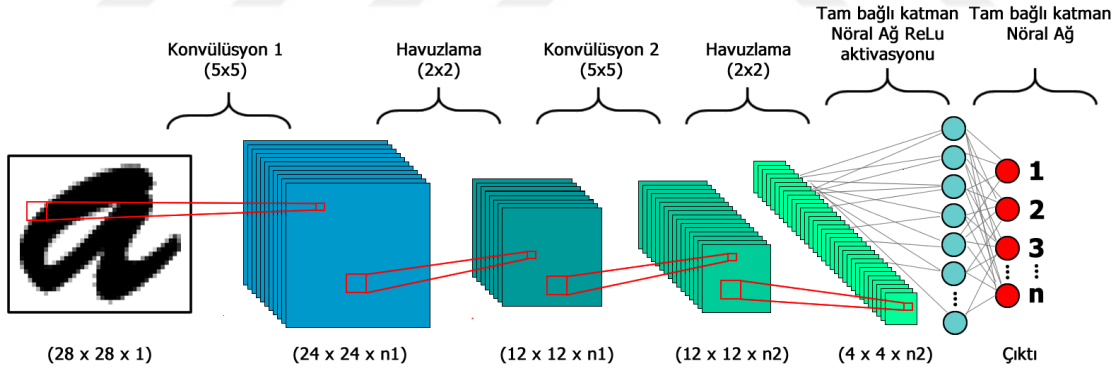
Random forest, sınıflandırma işlemi yaparken birden fazla sayıda karar ağacı üreten ve bununla sınıflandırma oranını yükselterek çalışan bir algoritma türüdür. Birden fazla kara ağaçlar arasında rastlantısal olarak seçilen karar ağaçlarının bir araya gelmesi ile karar ormanını oluşur. Veri setinin fazla değişkeni olması, sınıf etiketine sahip kategorik değişken

içermesi, kayıp verilerinin olması veya dengesiz bir dağılım sergilemesi durumlarında bile iyi sonuç verir [32].

Gradient boosting trees algoritması, random forest algoritmasına benzer fakat temel fark, son tahminin tüm ağaçların doğrusal bir toplamı olmasıdır. Diğer bir farkı ise her ağacın amacının, önceki ağaçların artık hatasını en aza indirgemesidir. Tüm senaryo tiplerinde ölçeklendirilebilen bu algoritmanın verimli bir şekilde uygulanmasını sağlar [32].

3.3. Derin Öğrenme

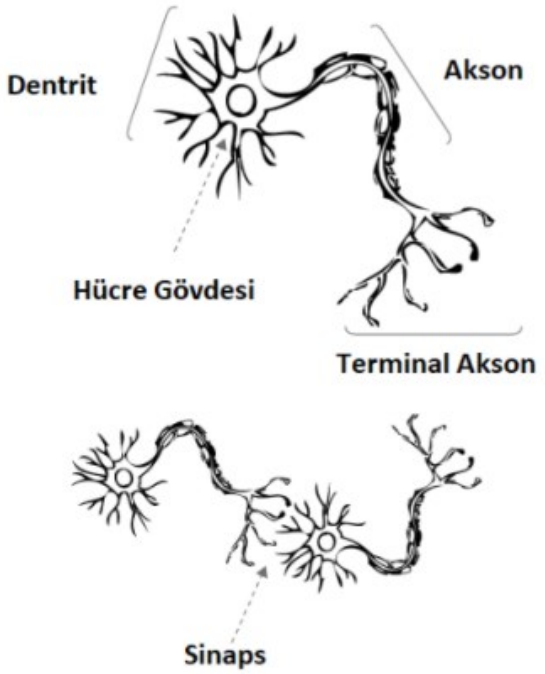
Günümüzde Evrişimli Sinir Ağı modelleri görüntü işleme ve analizinde yaygın olarak kullanılmaktadır [2]. Şekil 3.1'de temel çalışma prensibi gösterilmiştir. Deneysel çalışmada Evrişimli Sinir Ağı tasarlanırken derin öğrenme yöntemi benimsenmiştir. Bu Evrişimli Sinir Ağı, Evrişimsel katman, lineer olmayan katman, havuzlama katmanı ve tam bağlı katman olmak üzere dört farklı katmandan oluşmaktadır.



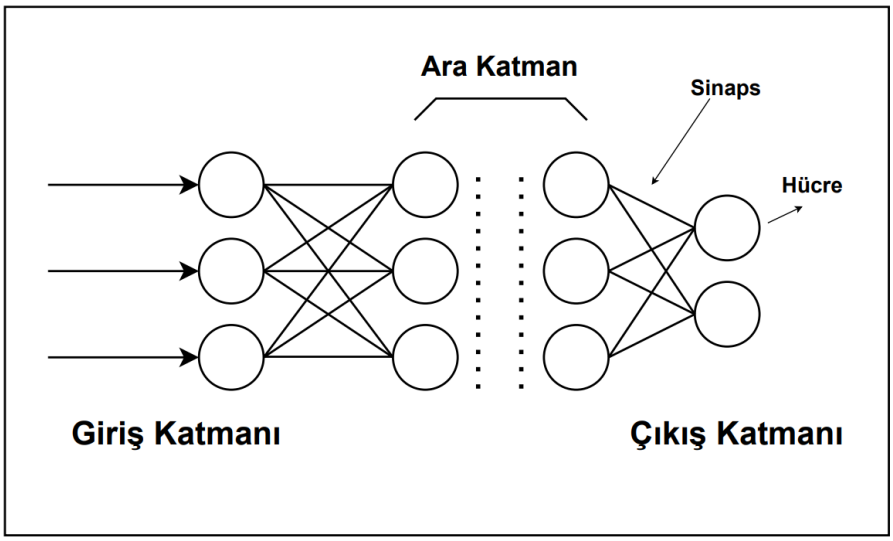
Şekil 3.1. Temel bir ESA mimarisi

İnsan beyni örnek alınarak tasarlanmış, bir veya birden fazla gizli katman içeren yapay sinir ağları ve benzeri makine öğrenme algoritmalarını kapsayan bir çalışma alanıdır. İnsanlar gibi algılama, analiz etme, çıkarım yapma ve bu çıkarıma göre aksiyon alarak karşılık verme, öğrenme, kategorize etme gibi özellikleri çok fazla veri kullanarak yapabilme özellikleri barındıran makine öğrenmesi tekniğidir. En az bir tane yapay sinir ağı kullanılarak bir çok algoritma desteği ile makinenin mevcut verilerden yeni veri elde ederek bu verileri anlamlandırmasıdır.

Biyolojik olan sinir hücresinde yer alan, Hücre, Dentrit, Akson ve Sinapslar taklit edilerek yapay sinir ağları geliştirilmiştir. Şekil 3.2'de biyolojik sinir hücresi modeli, Şekil 3.3'te ise yapay sinir ağı modeli verilmiştir.



Şekil 3.2. Biyolojik sinir hücresi [33]



Şekil 3.3. Yapay sinir ağları

Karmaşık yapıdaki insan beyninin daha anlaşılır hale getirilebilmesi için biyolojik sinir hücreleri modellenmektedir. Güçlü insan beynine oranla bilgisayarların karşılayabilmesi için 10^{11} tane bilgisayar işlemcisi çalışması gerekmektedir. Biyolojik olarak sinir hücreleri kendi görevleri ile aktivasyon üretmekte ve diğer sinir hücrelerine sinaps ile sağladığı bağlantılarla aktarmaktadır. Sinir hücreleri arasında yer alan bağlantılar zaman içinde güçlenmekte veya zayıflayabilmektedir. Bu durumlarda yeni bağlantılar eklenebilir ya da unutulabilir [21].

Derin Öğrenme son yıllarda gelişimi hızlanmış ve kullanımı artmış bir yöntemdir. Literatür Taraması bölümünde tarihçesinde anlatıldığı gibi ilk olarak McCulloch ve Pitts [5] adlı bilim insanları tarafından ortaya atılmıştır. 1943 yılında ortaya atılan bu fikir zaman içerisinde kabul görmüş ve özellikle 2000'li yıllardan sonra oldukça hızlı bir şekilde gelişimini sürdürmüştür. Derin öğrenme temelde insan beyninin çalışma prensiplerini taklit ederek, öğrenme, anlama, analiz yapma, özellik çıkarma ve karar verme mekanizmalarını uygulayabilen bir yapıdır. Biyolojik insan sinir hücresi; ana gövde, çekirdek, akson ve sinir uçlarından oluşur. Yapay sinir hücresi de biyolojik sinir hücresinin matematiksel karşılığına denk gelecek bir model olarak tanımlanabilir. Yapay sinir hücreleri tıpkı biyolojik sinir hücreleri gibi kendi aralarında çok sayıda bağlantı kurar. Bu bağlantılar bazen zayıf bazen de güçlü bir yapıya sahip olurken bazen de gereksiz görülen bağlantılar koparılır. Bu şekilde biyolojik beyin matematiksel olarak modellenmiş olur. Yapay sinir hücreleri arasındaki bağlantıların duruma göre değişiklik göstermesi fikri ilk olarak Donald Hebb tarafından ortaya atılmıştır [20]. Hebb, yapay sinir hücreleri arasındaki bağlantı gücünün, kullanım oranı ile doğru orantılı olacak şekilde artıp azaldığını kural olarak belirlemiştir. Dolayısıyla iki sinir hücresi arasındaki bağlantının kullanımı arttıkça hücreler arasındaki bağlantı gücü de artar, kullanım azaldıkça hücreler arasındaki bağlantı gücü de azalır ve duruma göre koparılır [21].

Bir Derin Öğrenme algoritması olan ESA, herhangi bir girdi görüntüsünü aldıktan sonra, görüntüdeki yönlere ve nesnelere öğrenilebilir ağırlıklarla ve önyargılarla önem atamaktadır ve bu şekilde görüntüleri birbirlerinden ayırabilmektedir. Diğer sınıflandırma algoritmaları ile kıyaslandığında, ESA modeli için gerekli olan ön işleme algoritması çok daha düşük kalmaktadır. İlkel yöntemlerde filtreler yeterli eğitim ile elle üretilirken, ESA için bu farklıdır. ESA bu filtre ve özellikleri öğrenme yeteneğine sahiptir. Biyolojik insan beyninden

esinlenilmiş olan ESA'nın mimarisi, insan beynindeki nöronların bağlantı şekilleriyle benzer modele sahiptir. Her bireysel nöron sadece Alıcı Alan olarak bilinen görme alanının sınırlı bir bölgesinde uyarıya yanıt vermektedir. Buna benzer alanların tamamından oluşan bir koleksiyon, görselin alanının tamamını kapsayacak biçimde örtüşmektedir.

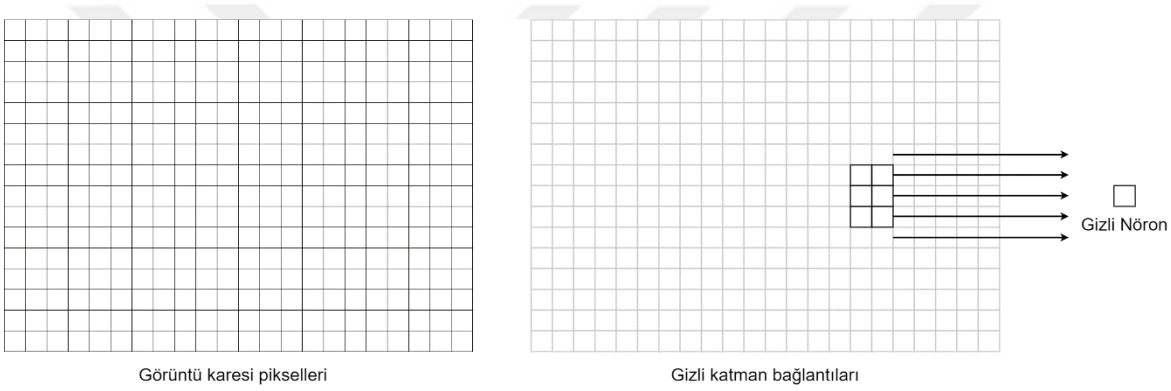
MTCNN, yüz algılama ve hizalama için geliştirilmiş ve yaygın olarak tercih edilen bir modeldir. MTCNN insan yüzlerini farklı ağ aşamaları ile tanımlamaktadır. Bunlar; göz, burun ve ağız gibi dönüm noktası konumunu tanımlayan ağ aşamalarıdır. MTCNN modeli P-Net, R-Net ve O-Net olmak üzere üç evrişimli ağa sahiptir. MTCNN, gerçek zamanlı performans sağlıyor olmasıyla beraber performans açısından diğer yüz algılama modellerinden daha iyi sonuçlar verdiği görülmüştür. Bu model alt resimleri farklı boyutlarda üretir ve piramit yapısı oluşturarak görüntü içindeki yüzleri bu yöntemle tespit eder. Oluşturulan alt resimlerde yüz taraması yapılır ve yüz tespit edilirse bir sonraki aşamaya geçilir. İlk aşamada, MTCNN yukarıda bahsedilen model ağının girdisi olan görüntüyü farklı ölçeklerde yeniden boyutlandırarak bir görüntü piramidi oluşturur [3].

HOG, genellikle üç boyutlu görüntülerde kullanılmasının yanı sıra görüntülerde nesnelerin tespit edilmesi amacıyla kullanılan özellik tabanlı bir tanımlayıcıdır. İşleyişin yavaşlığı konusunda bu nedenlere dayalı genel bulunmaktadır. HOG, sadece yüz tanıma için değil, aynı zamanda arabalar, evcil hayvanlar ve meyveler gibi nesnelerin tespit edilmesi için de sıklıkla kullanılan basit ve güçlü bir özellik tanımlayıcısıdır. HOG, nesne şekli, yerel yoğunluk gradyan dağılımı ve kenar yönü kullanılarak karakterize edilmesinden dolayı, nesne algılama konusunda güvenilirdir. Görüntü üzerinde gradyan değerleri hesaplanırken hem yatay hem de dikey yönde olmak üzere iki yönlü bir filtre uygulanmaktadır. Görüntüye ait renk ve doygunluk verileri bu yöntemle birlikte filtrelenebilmektedir. Her piksel için bir ağırlıklandırma işlemi, ikinci aşama olan yönlendirme gruplaması ile hesaplanmış olan gradyan değerlerine göre yapılmaktadır. Kontrast ve aydınlık bölgelerdeki gradyan değerinin güçlü yönleri ise son aşamada normalleştirilir [1].

SSD, ESA tabanlı sınıflandırıcılar olan bir temel derin öğrenme ağını kullanarak özellik haritasını çıkardıktan sonra nesneleri algılamak için evrişim filtreleri uygulamaktadır. Berkeley AI Research ve topluluğa katkıda bulunan kişiler tarafından geliştirilip kullanıma sunulan Caffe framewok ise bir derin öğrenme çerçevesidir. Caffe modelin birçok modele

göre daha verimli çalışmasının yanında diğer modellere kıyasla daha hızlı çalışmaktadır. [4].

ESA, Yerel Alıcı Alanlar, paylaşılan ağırlıklar ve havuzlama olmak üzere üç temel yöntemi görüntülerden özellik çıkarmak için kullanır [14]. Yerel Alıcı Alanlar: Evrişimsel sinir ağlarında, gizli nöronların katmanı küçük bağlantılardan oluşmaktadır fakat diğer sinir ağlarından farklıdır. Bu farkın sebebi giriş piksellerinin lokal bölgesi, her giriş pikselinin birebir bağlantısına sahip olmasıdır. Böylelikle giriş nöronlarının küçük bir bölgesi ile ilk gizli katmandaki her nöron bağlantılı olarak çalışmaktadır. Bağlantı, belirli gizli nöronlar için Şekil 3.4'teki gibi olacaktır [15].



Şekil 3.4. ESA mimarisinde bulunan gizli nöronların yapısı

Giriş görüntüsünde bulunan bu bölge, giriş pikselleri üzerindeki küçük bir pencere olan gizli nöron için yerel alıcı alan olarak adlandırılır. Alıcı alan, kayarak çalıştığı için tüm gizli katman nöronlarına ulaşıp bağlantı kurabilmektedir.

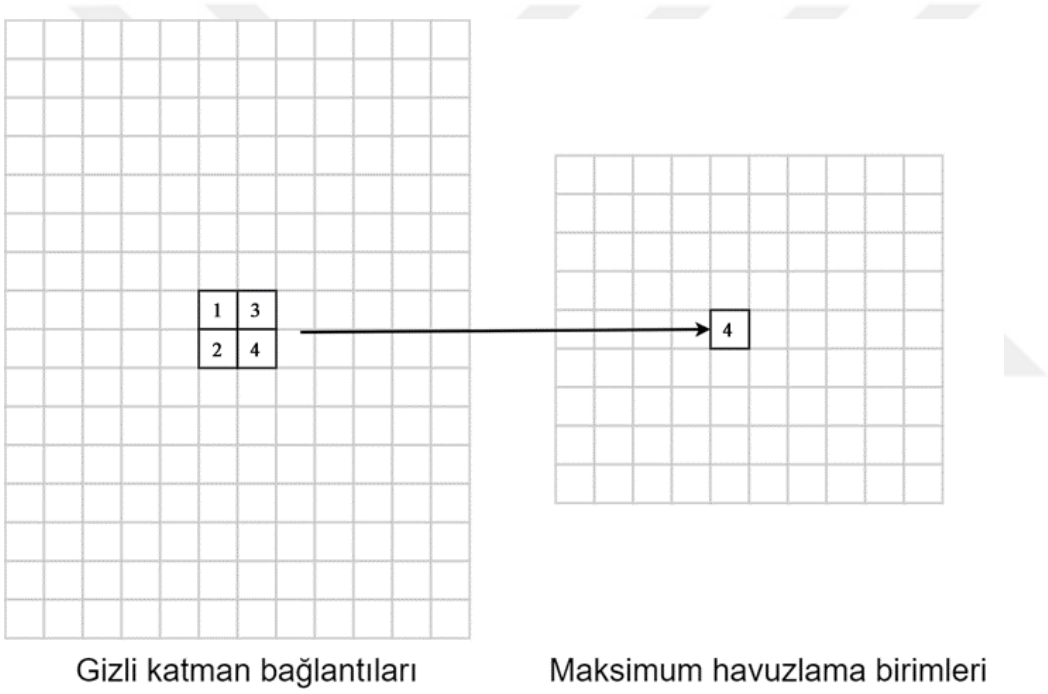
Paylaşılan Ağırlıklar: Gizli katmana bağlı yerel alıcı alanlardan aynı uzaklığa sahip olanlar, gizli katman süresince aynı sapmayı ve ağırlığı paylaşmaktadır. j_{th} için, k_{th} gizli nöron çıktısı Eş. 3.1'de gösterildiği gibi hesaplanmaktadır.

$$o(j, k) = \sigma\left(b + \sum_{l=0}^{4} \sum_{m=0}^{4} W(l,m) \alpha_j^{+l,k+m}\right) \quad (3.1)$$

σ nöral aktivasyon fonksiyonu veya sigmoid fonksiyon olduğunda, b paylaşılan sapma, W(l,m) paylaşılan ağırlıklardır ve $\alpha_{x,y}$ x, y konumunda girdi aktivasyonunu beslemektedir. İlk gizli katmandaki nöronlar, girişte bulunan özellikleri farklı konumlarda algılamaktadır. Bu nedenle, girişin gizli katmanından gelen harita, özellik haritası için kastedilen sapma

değeri ortak sapma olarak adlandırılmaktadır. Paylaşılan ağırlıklar ve sapmalar, çekirdek veya filtre olarak adlandırılmaktadır. Yüz tanıma birçok özellik eşler, ancak çekirdek parametreleri etkili bir şekilde azaltır [15].

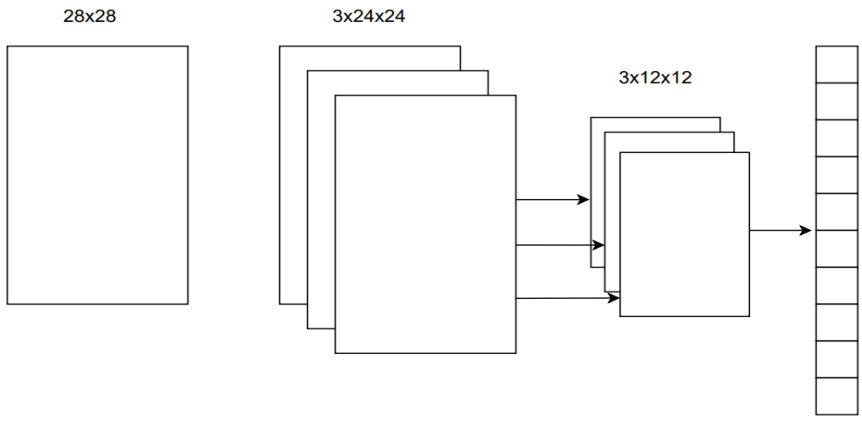
Havuzlama Katmanı: Havuzlama katmanı, çıktıdaki bilgileri yoğunlaştırılmış özellik haritasına basitleştirmek amacıyla kullanılan konvolüsyon katmanının özellik haritasından sonra konvolüsyon katmanı tarafından takip edilmektedir. Maksimum havuzlama, 2x2 giriş bölgesindeki bağlantıların maksimum etkinleştirme çıktısını döndürdüğü Şekil 3.5'te gösterilmiştir [15].



Şekil 3.5. ESA mimarisinde kullanılan havuzlama katmanı

Çalışmalarda, 28 dönüşüm katmanına sahip bir ResNet ağı modeli yüz tanıma aşamasında kullanılmıştır. Bu ResNet ağı, He ve diğerleri tarafından yayınlanmış olan "Görüntü Tanıma için Derin Kalıntı Öğrenme" [16] çalışmasında önerilen ResNet-34 ağından birkaç katmanın kaldırılmasıyla ve katman başına filtre sayısının yarı yarıya azaltılmış, Şekil 3.6'da temel yapısı gösterilen bir versiyonudur. Zhang ve diğerleri çalışmada önerilen ResNet-34 özellikle görüntü sınıflandırma modeli olarak kullanılabilen 34 katmanlı bir evrişimsel sinir ağıdır. Bu ağ olan ImageNet veri kümesinde önceden eğitilmiş bir modeldir. Bu veri kümesinde 200 farklı sınıfta 100.000 görüntü bulunmaktadır. ResNet 34 modeli çalışmada

referans model olarak seçilmiştir. Bu model katmanları ve filtreleri üzerinde aynı senaryo için çoklu-çapraz testler gerçekleştiren bir yazılım geliştirilmiştir. Aynı koşullar altında modelin farklı seviyelerde katmanları ve filtreleri aktif/pasif hale getirilmiş ve her döngü sonunda elde edilen sonuçlar kaydedilmiştir. Bu işlemler sonucunda en yüksek çalışma performansını en düşük kayıpla gerçekleştirebilen ResNet 34 modelinin bir versiyonu olan 28 katmanlı model elde edilmiş ve çalışmamızda kullanılmıştır.



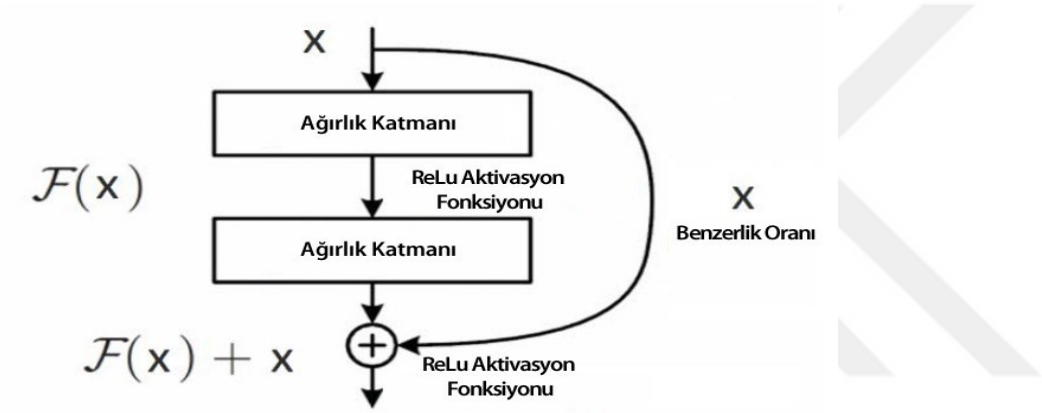
Şekil 3.6. 28 katmanlı Evrişimli Sinir Ağları modeli

ResNet modeli, He ve diğerleri tarafından önerilmiş bir evrişimli sinir ağı CNN modelidir. VGGNet mimarisine göre sekiz kat daha fazla derinliğe sahiptir [2]. ResNet 2015 yılında gerçekleştirilen ImageNet yarışmasında %3,57 hata oranı ile birincilik derecesini elde etmiştir. ResNet modeli barındırdığı katman sayısına göre farklı şekillerde isimlendirilmektedir. Örnek verilecek olursa, ResNet34 modeli 34, ResNet101 modeli ise 101 adet katmandan oluşmaktadır. Şekil 3.7'de 34 katmanlı bir ResNet modeli gösterilmiştir. ResNet, sahip olduğu mikro mimari modül yapısı ile geleneksel ardışık ağ mimarileri arasında yer alan VGGNet ve AlexNet mimarilerinden ayrılır. Ortaya çıkış amacı, çok katmanlı derin ağlarda yaşanan gradyan değerini sıfıra yakınsama probleminin çözümüdür. ResNet bir köprü görevi görerek alt katmanlarda elde edilen öznitelik bilgilerini üst katmanlara sıralı şekilde taşır [16]. Katmanlar arasında gerçekleşen atlama işlemi olan ResBlock ile bir önceki bloktan bilgi alınamamış olsa bile eski katmanlardan elde edilmiş olan bilgiler yeni katmanlara uygulanır ve böylece model daha güçlü hale getirilir. Bu sayede, gradyan silinmesinde yaşanan problemin de ResBlock ile önüne geçilmiş olunur [16].



Şekil 3.7. 34 katmanlı ResNet modeli

Kullanılan ağın derinleşmeye başlaması ile derin öğrenme mimarilerinde bozulmalar başlar. Bozulmalar sonucunda ise eğitimler için alınan sonuçlarda fazla oranda hata meydana gelir. Ayrıca sistemin optimizasyonu da zorlaşır [16].



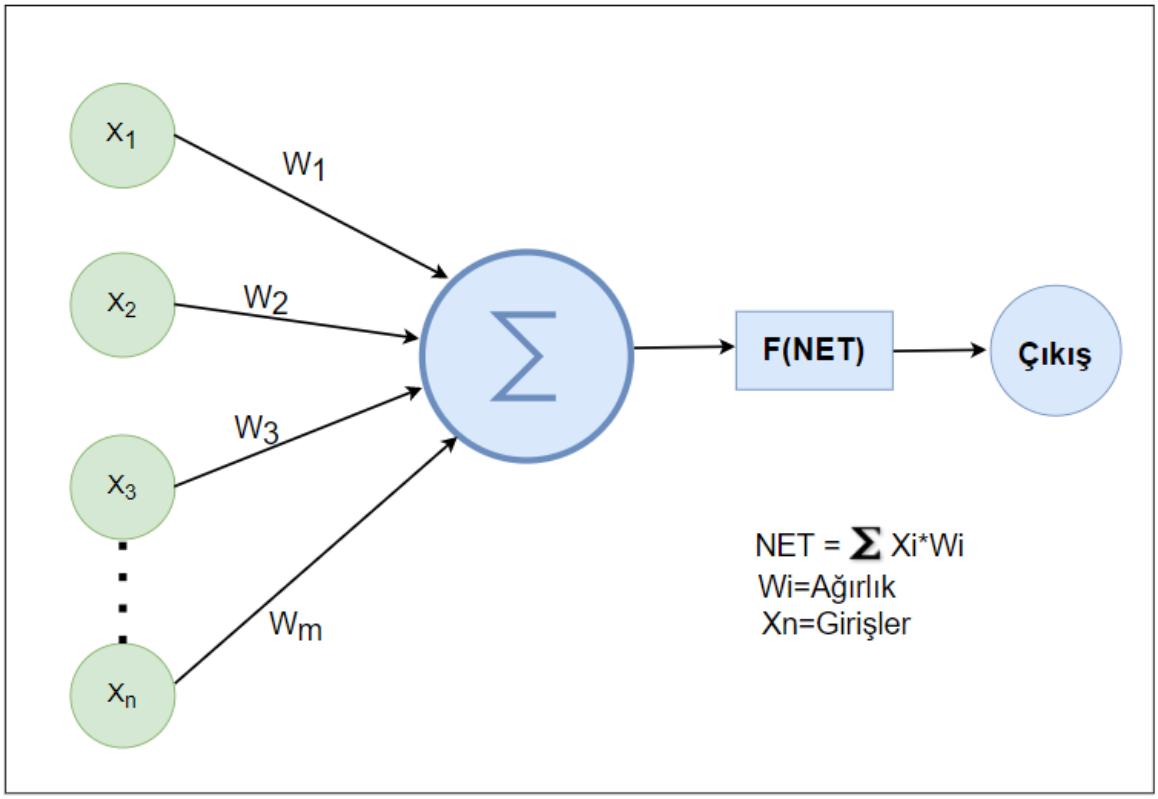
Şekil 3.8. ResNet mimarisi [16]

Microsoft tarafından geliştirilen derin artık öğrenme modeli bu problemin çözümü için kullanılır. Kullanılma mantığı, birkaç yığın katmanı seçilen bir haritaya doğrudan uydurmak yerine bu katmanların artık bir haritaya uymasına izin verilmesi şeklindedir. Kısayol bağlantılarına sahip ileri beslemeli sinir ağlar ile F(x)+x 'in formülasyonu gerçekleştirilebilmektedir. Şekil 3.8'de kısayol bağlantıları gösterilmiştir. Bu bağlantılar bir veya birden fazla katmanı atlayan bağlantılardır. Kısayol bağlantıları, kimlik eşlemesi yaparak çıktıları yığın katmanların çıktılarına ekler. Bu yöntem ile ağdaki derinlik arttıkça meydana gelen eğitim hataları önlenir ve diğer mimarilerden daha iyi sonuçlar alınır [16].

3.3.1. McCulloch ve Pitts modeli

İlk sinir hücresi modeli 1943 yılında Warren McCulloch ve Walter Pitts tarafından yapılmıştır. İlk matematiksel yapay sinir modeli olarak kabul gören bu model "Sinir Sisteminin İçinde Olan Fikirlerin Mantıksal Hesabı" makalesinde duyurulmuştur. Adım

fonksiyonları kullanılmasından dolayı model, doğrusal eşik kapısı olarak da bilinmektedir. Hesaplamalarda ara değerler için bir eşik değeri belirlenir ve bu eşik değeri üzerinde kalan ara değerler tepe değere yuvarlanır. Bu modelde yer alan aktivasyon fonksiyonu (0,1) veya (-1, 1) değerlerini alabilen çıktı üretmesi, girdi verilerinin iki farklı sınıfa ayrılması sağlayarak ikili sayı sisteminde çıktı verir. Böylece, belirlenmiş girdiye karşılık bir çıktı değeri üretilmiş olur. Yapay sinir hücresinde elde edilen bu çıktı, her girdi değerinin dentritler ile çekirdeğe iletilmesi, çekirdekte gelen sinyallerin toplanarak aksona iletilmesi ve aksonda işlenen sinyallerin sinapslar ile sonraki hücreye iletilmesi şeklinde olur [5].



Şekil 3.9. Matematiksel sinir hücresi modeli

Şekil 3.7'de verilen sinir hücresi modelinde X_1, X_2, …, X_n ifadeleri giriş sinyallerini, w_1, w_2, …, w_n ifadeleri ise ağırlıkları belirtmektedir. X değerleri, ağırlıklar ile çarpılıp toplamları alınır. Aktivasyon fonksiyonuna F(NET) tabi tutularak çıkış bulunur. Bir biyolojik sinir hücresi ile akson karşılaştırıldığında Akson - Çıktı, Dentrit – Toplam Fonksiyon, Çekirdek – Aktivasyon Fonksiyonu, Sinaps – Ağırlık olacak şekilde denk gelmektedir. Gerçek bir sinir ağının karmaşık yapısına bakıldığında, McCulloch ve Pits modelinin çok benzer olmadığı görülür. McCulloch ve Pits modeli doğrusal aktivasyon sinyalleri gönderirken

gerçek sinir hücrelerinde bu doğrudan olmak zorunda olmadığı gibi tek bir çıkış yerine birden fazla çıkış verebilmektedir. Ayrıca bağlantı değerlerinin bilgisayarda sıralı olarak güncellenirken gerçek sinir hücrelerinde ağırlıkları rastgele ve asenkron yapılmaktadır [5]. McCulloch ve Pits modelinde yapay sinir hücresi; Giriş, Çıktı, Ağırlıklandırma, Hedef ve Hata bölümlerinden oluşmaktadır. Giriş kısmı yapay sinir hücresine iletilen dizi şeklindeki verilerden oluşmaktadır. Genellikle x_i şeklinde gösterilmektedir. Çıktı kısmı da giriş kısmından alınıp ara katmanlarda işlenen verilerin sonuç olarak yine bir dizi şeklinde veren kısımdır. Ağırlıklandırma kısmı 2 farklı sinir hücresi arasındaki bağlantıyı ifade eder. Hedef kısmının genellikle t ile gösterimi yapılır, çıktı ile aynı değere sahip olması durumunda sonuç doğru kabul edilir, farklı olması durumundan ise hata olduğu kabul edilir ve geriye doğru bir dönüş olur. Hata kısmı ise E ile ifade edilir ve hedef ve çıktı kısmı arasındaki uyumsuzluktan kaynaklanır. Birden fazla hesaplanma metodu vardır. Genellikle hedef ve çıktı arasındaki Eş. 3.2'de gösterilen Manhattan ve Eş. 3.3'te Öklid mesafeleri kullanılarak hesaplanır [5].

Manhattan mesafesi;

$$m = |x1 - x2| + |y1 - y2| \tag{3.2}$$

Öklid mesafesi;

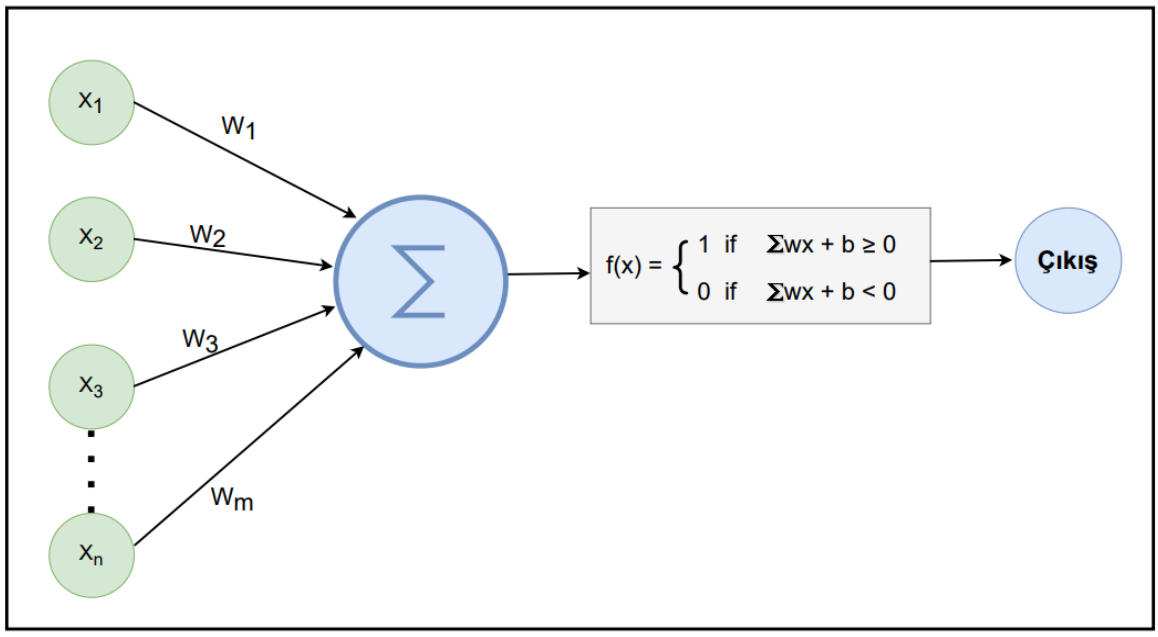
$$o = \sqrt{(x1 - x2)^2 + (y1 - y2)^2} \tag{3.3}$$

3.3.2. Hebb kuralı

1949 yılında Donald Hebb tarafından yapılan çalışma ile sinir hücreleri ağırlıklarının değişebileceği ortaya çıkmıştır. Bilgisayar bilimlerinde yapay sinir ağları için gerçekleştirilen eğitimlerde kullanılan bir kuraldır. Sinir hücreleri arasındaki bağlantıların zamanla kullanıma bağlı olarak güçlerinde meydana gelen artma ve azalması Hebb Kuralı ile açıklanmaktadır. Sık kullanılan bağlantıların ağırlık kat sayıları artar ve güçlenirken, az kullanılan bağlantıların ağırlık kat sayıları yani güçleri azalır [21]. Biyolojik çalışmalar sırasında elde edilen bu kural yapay sinir ağları için de uygulanmıştır ve başarılı sonuçlar elde edilmiştir.

3.3.3. Perseptron

McCulloch ve Pits sinir hücrelerinin bir araya getirilmesi ile oluşan bir yapı olan Perseptron, 1958 yılında F.Rosenblatt tarafından bulunmuştur. Perseptron, başka bir sinir hücresine bağlı olmayan bir sinir hücresidir ve bu nedenle diğer sinir hücrelerinin ne tür değer ürettiği ile ilgilenmez. Ancak, girdi değerlerini paylaşırlar ve bu girdi değerleri ağırlıklı bir bağlantı ile tüm sinir hücrelerine ulaşır [23]. Şekil 3.8'de Perseptron yapısı verilmiştir.

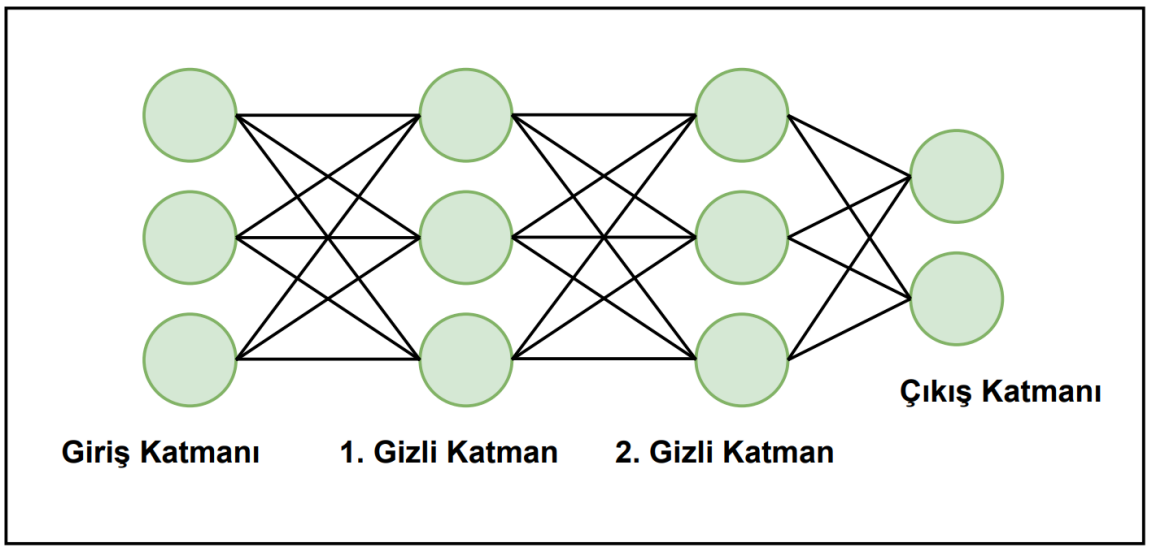


Şekil 3.10. Perseptron yapısı

3.3.4. İleri yayılım

Karmaşık bir problemin biyolojik olarak çözülebilmesi için sinir hücreleri arasında bağlantılılar kullanılır. Sinir ağları girdi katmanı, gizli katman ve çıktı katmanı olmak üzere üç ana katmandan oluşur. Girdi katmanının görevi, girdi verilerini gizli katmana iletmektir [24].

En az bir ara katman bulunduran sinir ağları, doğruluk skoru kabul edilebilir ve fonksiyonu tahmin edilebilir olarak görülür. Problem çözümlerinde düşük veya yetersiz kalan doğruluk skorları ile karşılaşıldığında, gizli katman sayısı ve girdi-çıktı bağlantıları ile ilişkili olduğu düşünülmektedir [24]. Şekil 3.9'da Yapay Sinir Ağı ve katmanları verilmiştir.



Şekil 3.11. Yapay sinir ağı ve katmanları

Girdi katmanına, gelen veriler arasında resim veya bir sayı dizisi verileri olabilmektedir. Girdi katmanı ise bu verileri alarak vektör veya matris olarak ara katmanlara iletir. Önceden atanmış olan ağırlık değeri ile girdi katmanından çıkan veriler ilişkilendirilir ve gizli katmana iletilir. Gizli katmanlar ise kendinden sonraki gizli katmanlara veya gizli katman olmaması durumunda çıktı katmanına iletmektedir. İleri yayılımlı ağlar, bu şekilde sıralı olarak bağlı ağlara verilen isimdir. İletim verilerin sürekli bir sonraki katmana aktarılması şeklinde gerçekleşmektedir [24].

3.3.5. Geri yayılım

Yapay sinir ağları için Geri Yayılım yaklaşımı ilk olarak Rumelhart tarafında ortaya atılmıştır. Gerçek verilerle tahmin edilen veriler arasındaki farkı en asgari düzeye indirgemek için ortaya atılmış yeni bir yaklaşımdır [25]. İleri yayılım yaklaşımı gerçek değerler ile tahmin edilen değerler arasında oluşan farka göre işlem yaparken, geriye yayılım yaklaşımı bu iki değer arasındaki fark en küçük değerine ulaşıncaya kadar devam eder. Geri yayılım yaklaşımında ağırlıklar geriye doğru güncellenebilmekte ve yeni ağırlık değerleriyle ileriye yayılım devam etmektedir. Dolayısıyla gerçek veriler ile tahmin edilen veriler arasındaki doğruluk oranı artmaktadır.

Geri yayılım yaklaşımında aktivasyon fonksiyonlarının türevi alındıktan sonra elde edilen eğim değerlerine göre arada bulunan ağırlıklar güncellenmekte ve en düşük maliyet elde edilinceye kadar devam etmektedir [25].

3.3.6. Aktivasyon fonksiyonları

Bir yapay sinir hücresinin giriş verilerine karşılık gelen çıktı değerini elde etmek için kullanılan fonksiyonlardır. Sigmoid, ReLU, LeaklyReLU, ELU, SELU, PReLU, Tanh ve Softmax gibi birden fazla türde aktivasyon fonksiyonları kullanılmaktadır [26]. Çizelge 3.1'de bu aktivasyon fonksiyonları verilmiştir.

Çizelge 3.1. Aktivasyon fonksiyonları

Aktivasyon Adı	Aktivasyon Fonksiyonu
Sigmoid [26]	$\delta(\mathrm{x}) = \frac{1}{1+e^{-x}}$
ReLU [26]	$f(x) = \max(0, x)$
LeakyReLU [26]	$f(x) = \begin{cases} ax, & \mathrm{x} < 0 \\ \mathrm{x}, & \mathrm{x} \geq 0 \end{cases}$
ELU [28]	$f(x) = \begin{cases} \mathrm{x}, & \text{eğer } \mathrm{x} > 0 \\ a(e^x - 1), & \text{eğer } \mathrm{x} \leq 0 \end{cases}$
SELU [29]	$f(x) = \lambda \begin{cases} \mathrm{x}, & \mathrm{x} > 0 \\ ae^x - a, & \mathrm{x} \leq 0 \end{cases}$

Aktivasyon fonksiyonları formüllerinin parametreleri [27]

x : değişken

e : üstel fonksiyon

a : sabit katsayısı

δ : delta

k : vektör boyutu

3.3.7. Optimizasyon algoritmaları

Geri yayılım yaklaşımında ağırlıkların güncellendiğinden ve güncel ağırlık değerleri ile ileriye yayılımın tekrar yapıldığından önceki bölümlerde bahsedilmişti. Ağırlıkların güncellenmesi aşamasında oluşan hata değerlerine göre en optimal ağırlık değerinin bulunması için geliştirilmiş çeşitli optimizasyon algoritmaları kullanılmaktadır. Momentum,

Nesterov Hızlandırılmış Eğimi, Adagrad ve Adamax gibi birçok farklı optimizasyon algoritması mevcuttur [27].

3.4. Model Performans Değerlendirme

Bu çalışmada Eş. 3.4'te gösterilmiş olan bir eşleşme eşik değeri, yüz tespitinin ve tanımlanmasının doğru kabul edilebilmesi için kullanılmıştır. Literatürdeki [17, 18] gerçekleştirilen deneysel çalışmalarda en uygun eşik değeri 0,6 olarak belirlenmiştir fakat yüz tespiti için farklı eşik değerleri de kullanılmıştır.

Aynı şekilde literatürdeki diğer çalışmalarda [1, 3, 4, 11, 19], Accuracy, Precision, Recall ve F1 Skoru ilgili modellerin başarım ve performans değerlerini ölçmede sıklıkla kullanılmıştır.

Çalışmada kullanılmış olan diğer değerlendirme ölçütleri sınıflandırma Accuracy, Precision, Recall ve F1 Skoru aşağıda sırasıyla Eş. 3.5, Eş. 3.6, Eş. 3.7 ve Eş. 3.8'de gösterilmiştir. Bir saniyede işlenen toplam görüntü karesi (FPS) ile beraber F1 Skor değerleri dikkate alınmıştır. İlgili modellerin çalışma hızları, aynı video dosyasının ilgili modeller ile baştan sona kadar taranması sırasında geçen toplam süreler ile hesaplanmıştır.

Eşitliklerde belirtilmiş olan TP, FP, TN ve FN değerler sırasıyla Doğru Pozitif, Yanlış Pozitif, Doğru Negatif ve Yanlış Negatif şeklindedir.

$$e = \begin{cases} \geq 0.6 \text{ Doğru} \\ < 0.6 \text{ Yanlış} \end{cases} \tag{3.4}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \tag{3.5}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{3.6}$$

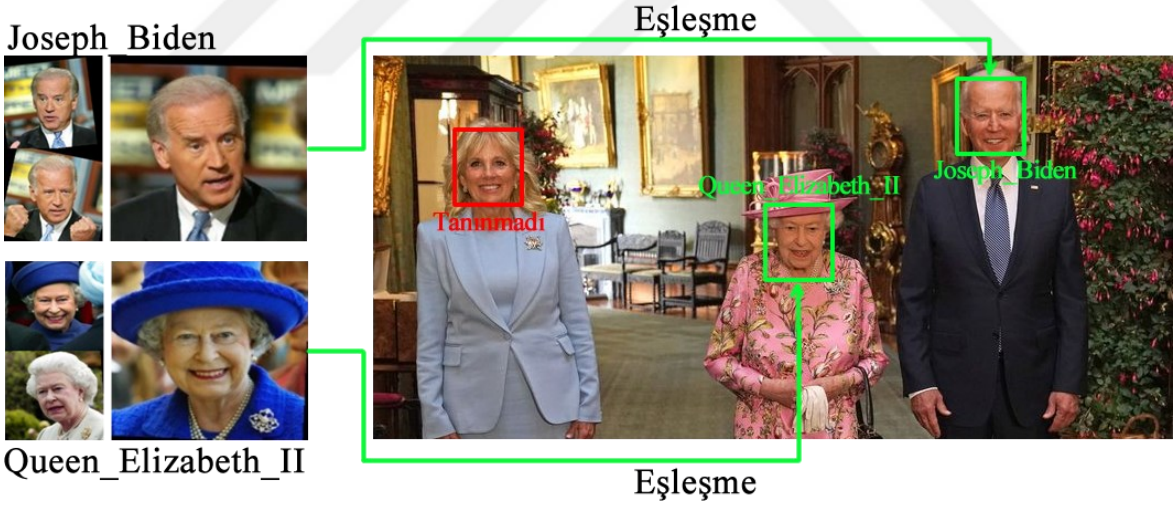
$$\text{Recall} = \frac{TP}{TP+FN} \tag{3.7}$$

$$F1 = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3.8}$$

4. VİDEO GÖRÜNTÜLERİNDE YÜZ TESPİTİ VE ZAMAN İŞARETLEME

4.1. Veri Kümesi

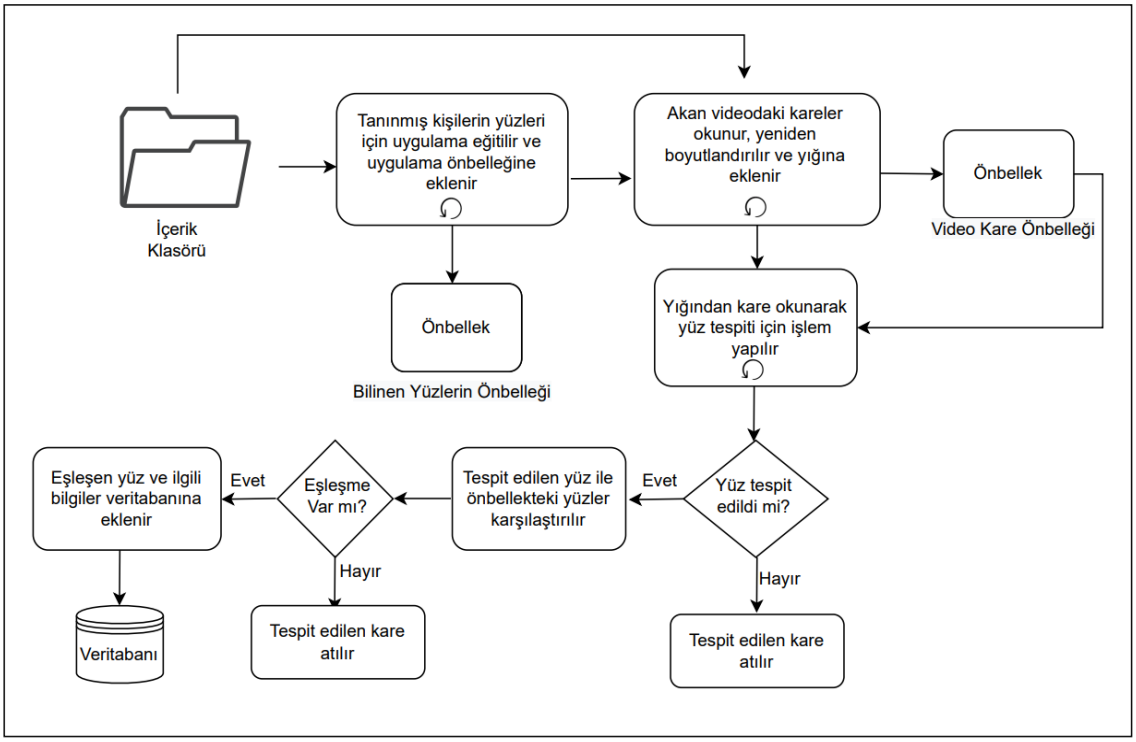
Gerçekleştirilen deneysel çalışmalarda Labeled Faces in the Wild Home [13] veri kümesi kullanılmıştır. Bu veri kümesinde 5749 farklı kişi ve bu kişilere ait 13 000'den fazla yüz görüntüsü bulunmaktadır. Kullanılmış olan bu veri setinde her yüz görüntüsü için kişilerin adı etiketlenmiş durumdadır. Labeled Faces in the Wild Home veri kümesi içerişindeki tanımlı olan 1680 kişinin en az 2 görüntüsü bulunmaktadır. Viola-Jones yüz detektörü kullanılarak yüzlerin tespit edilmiş olması yüzlerdeki tek kısıtlamadır. Veri kümesinde yer alan resimlerin dosya uzantıları .jpg olup boyutları 250x250 pikseldir. Resim-4.1'de veri kümesinin dosya yapısı gösterilmiştir.



Resim 4.1. Labeled Faces in the Wild Home veri kümesi yapısı

Yapılmış olan deneysel çalışmada kullanılmış olan video görüntüleri, Labeled Faces in the Wild Home veri kümesinde yer almış olan kişilerin de içerisinde bulunduğu, herkese açık bir şekilde paylaşılan ve ulaşılabilen internet ortamında bulunan videolardan oluşturulmuştur. Kullanılmış olan bu video görüntüleri farklı çözünürlüklerde ayrı dosyalara dönüştürülmüştür.

4.2. Deneysel Çalışma



Şekil 4.1. Önerilen sistem mimarisi

Video üzerinde akan karelerde görüntülenen kişilere ait gerçek zamanlı yüz tespitlerinin yapılabilmesi için geliştirilen uygulamada yapılan deneysel çalışmalar sırasında Şekil 4.1'de gösterilen işlem adımları izlenmiştir. Çalışmalar sırasında kullanılan bilgisayar 3.80 GHz CPU, 32 GB RAM ve 8 GB Nvidia Cuda ekran kartına sahiptir.

Yukarıda verilen adımlarda da belirtildiği gibi veri kümesi içerisinde yer alan kişilere ait yüz görüntülerinin işlenmesi ve bu kişilere ait yüzlere denk gelen eğitim modelleri, çalışmaların ilk aşamasında oluşturulmuştur. Geçmişten günümüze farklı zaman ve mekanlarda kayıt altına alınmış videolar için akan görüntülere ait karelerde aynı anda birden fazla kişinin yer aldığı ve aynı kişilerin farklı zamanlarda geçen karelerde bulunduğu görülmüştür. Bu nedenle deneysel çalışmalar için kullanılan veri setlerinin benzer içeriklere sahip olması gerektiği göz önünde bulundurulmuştur. Kullanılan veri setlerinde modele dahil edilmek için seçilen görüntü dosyaları, görüntüde yer alan kişiler için kişi başına en az bir görüntü en fazla beş görüntü olacak şekilde her bir kişi için belirlenmiştir. Bir sonraki aşamaya geçildiğinde, kurulan Cache yapısına eğitim modelleri dahil edilmiştir. Bunun nedeni

kullanılan uygulamanın yaptığı işlemlerin hızını arttırmak ve kişileri tanımlama süresini azaltmasını sağlamaktır.

Geliştirilen mimari ile görüntü işleme ve nesne tanıma için literatürde sıklıkla kullanılan ESA modellerine göre daha az sayıda katmana sahiptir. Geliştirilen ESA modeli toplamda 28 katmana sahip olması nedeniyle benzer modellere göre ardışık görüntü karesi işlenme hızının arttığı deneysel çalışmalarda gözlemlenmiştir. Geliştirilen mimaride içerisinde ön işleme, tarama, tespit ve kayıt aşamaları için birden fazla ön bellek yapısı oluşturulmuş bu sayede disk okuma-yazma işlemlerinden kaynaklı gecikmeler ortadan kaldırılmıştır. Ayrıca oluşturulan mimaride kullanılan paralel programlama teknikleri ile görüntü işleme ve karşılaştırma aşamalarındaki gecikme seviyesi en aza düşürülmüştür. Görüntü karelerinin okunması, boyutlandırılması, yüz tespiti, yüz tanımlanması ve veri tabanı kayıt aşamalarının her biri paralel şekilde çalışacak ayrı ayrı prosesler şeklinde tasarlanmıştır. Tüm bunlar sayesinde diğer benzer yüz tanıma mimarilerine göre %33'e varan bir hız artışı sağlanmıştır.

Yapılan çalışmalar sonucunda uygulama performansının arttığı, disk okuma ve yazma işlemlerinin neden olduğu süre uzatımları ve gecikmelerin en aza indirgenmiş olduğu gözlemlenmiştir. Video dosyalarının ardışık görüntü kareleri FIFO mantığıyla oluşturulan maksimum 100 eleman kapasiteli kuyruk yapısına sırası ile yeniden boyutlandırılarak eklenmiştir. Daha sonrasında veri kümesinde bulunan kişiler ile karşılaştırma için sırasıyla kuyruktan okunmuş ve işleme tabi tutulmuştur. Karşılaştırma sonucunda herhangi bir eşleşme bulunmaması durumunda ilgili görüntü karesi kuyruktan atılmış ve bu işlem tüm görüntü kareleri için devam ettirilmiştir. Tek tek karşılaştırma yapılırken yüz görüntüleri arasında birbiri ile eşleşme olması durumunda eşleşen yeni görsel veri tabanına kaydedilmektedir.

Bu mimariye ait Sözde Kod (Pseudocode) Şekil 4.2'de gösterilmiştir.

Önerilen Mimarinin Sözde Kodu

Video dosyalarında gerçek zamanlı yüz tanımlama.

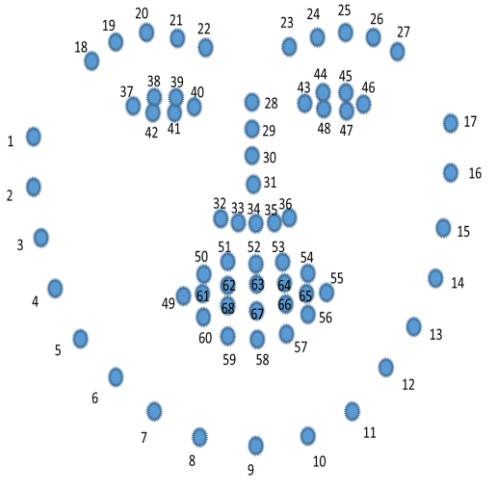
```
Giriş: Resim ve video görüntü kareleri f ∈ {1,...,N}, T toplam video kare sayısını ifade eder.
Çıkış: Tespit edilen ve tanınan yüzler, zaman damgaları, kare içindeki yüz koordinatları,
kare numarası
Prosedür:
1: Tanınmış yüzleri eğit.
2: Her yüzün sayısal karşılığını ön belleğe al
3:Video karelerini oku, yeniden boyutlandır ve kare kuyruğuna ekle
4: for t = 1 de T'ye kadar do
5:      Geçerli kareden yüzleri bul
6:      if Geçerli çerçevede herhangi bir yüz bulundu ise
7:          Bulunan yüzleri eğitilmiş yüzlerle karşılaştır
8:              if Herhangi bir yüz eşleşmesi var ise
9:                      Veritabanına eşleşen yüzü ekle
10:             end if
11:     end if
12: end for
```

Şekil 4.2. Uygulanan algoritmanın sözde kodu

Günümüzde teknolojik gelişmeler ile kullanıcıların uygulamalara ait beklentileri de çalışma performansının geliştirilmesi yönünde olmuştur. Bu çerçeveden bakıldığında kullanıcılar uygulamaları kıyaslarken çalışma hızını ve sonuç doğruluk oranını göz önünde bulundurmaktadır. Bunların yanında arşivlerde biriken çok sayıda ve büyük boyutlu videolar, inceleme ve yüz tespiti açısından yüksek performans gösteren uygulamalara ihtiyaç duymaktadır. Akan videolar için geliştirilen yüz tanımlama uygulamalarında, görüntü dosyaları içerisinde yer alan nesne ve/veya örüntülerin tespit edilmesi ve tanımlanması için yürütülen aşamalarda öne çıkan iki temel problem vardır. Bu problemler; uygulamanın çalışma sonucu verdiği çıktının başarı oranı ve uygulamanın çalışma performansıdır. Bu problemlerin giderilmesi için, görselde yer alan nesnelerin en doğru şekilde tespitinin gerçekleşmesi, tanımlanması ve etiketlenmesinin yanı sıra tüm bu işlemler gerçekleştirilirken uygulamanın daha hızlı çalışması ve daha az maliyet çıkaracak şekilde geliştirilmesi gerekmektedir. Çalışmalar esnasında kullanılabilecek farklı dosyalar için ön çalışma yapılırken her zaman yüksek çözünürlüklü dosya bulmanın mümkün olmayacağı görülmüş ve uygulama performansının arttırılması için bu kriter göz önünde

bulundurulmuştur. Bu nedenle kullanılan veri kümesi için görüntü dosyaları seçilirken videodaki kişiler etiketlenmiş, etiketlenen kişiler için birden fazla ve birbirinden farklı olan video dosyaları belirlenerek bu dosyaların farklı çözünürlükte olmasına önem verilmiştir. Deneysel çalışmada, veri kümesinden kullanılan video dosyalarının içerisinde yer alan kişiler için yapılan tanımlamalarda minimum gerçek zamanlı tespitler yapılması ve kişilerin video içerisinde geçtiği zamanların da tespit edilerek bu zamanlara göre etiketlenmesi amaçlanmıştır.

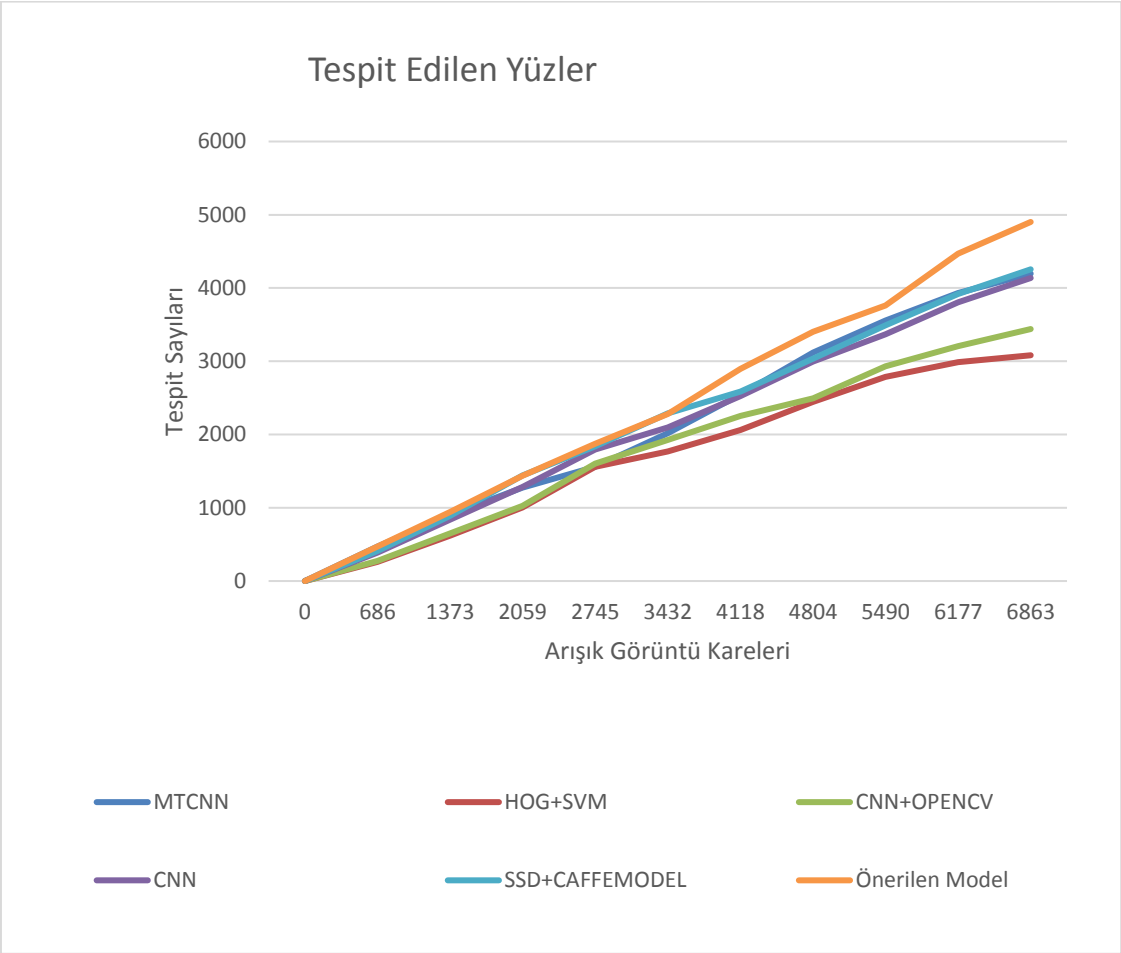
Bu çalışmada, önerilen ESA tabanlı mimari kullanılarak akan video görselinde tespit edilen görüntü içerisinde bulanan ve yüz olarak kabul edilen örüntüye sahip desenler Şekil 4.3'te gösterilmektedir. Şekilde de görüldüğü üzere çalışmamızda yüzün göz, kaş, burun, ağız ve çene bölgelerini ifade eden ve 68 noktadan oluşmakta olan yüz tanımlayıcı bir model kullanılmıştır. Videoda geçen görüntü karelerinde yer alan kişilerin yüz tespitlerinin yapılabilmesi için ESA tabanlı yüz tanımlayıcı modeller kullanılmıştır. Uygulamanın aşamalarında, ilk olarak yüz tespiti yapıldıktan sonra görüntü karesinde bulunan yüz için dikey hizalama işlemi uygulanır. Çalışmanın son aşamasına geçildiğinde ise hizalama işlemi uygulanmış olan yüz görüntüsü önceden belirlenmiş olan çözünürlükte kırpılır ve sonra kişinin tanınması işlemi için hazırlanmış olur.



Şekil 4.3. Tespit edilen yüz örüntüsü

Gerçek zamanlı yüz tespit uygulamalarının, çalışma hızı ve doğruluk oranları açısından değerlendirilebilmesi amacıyla aynı veri setleri ve ortamlar kullanılarak testler edilmeleri

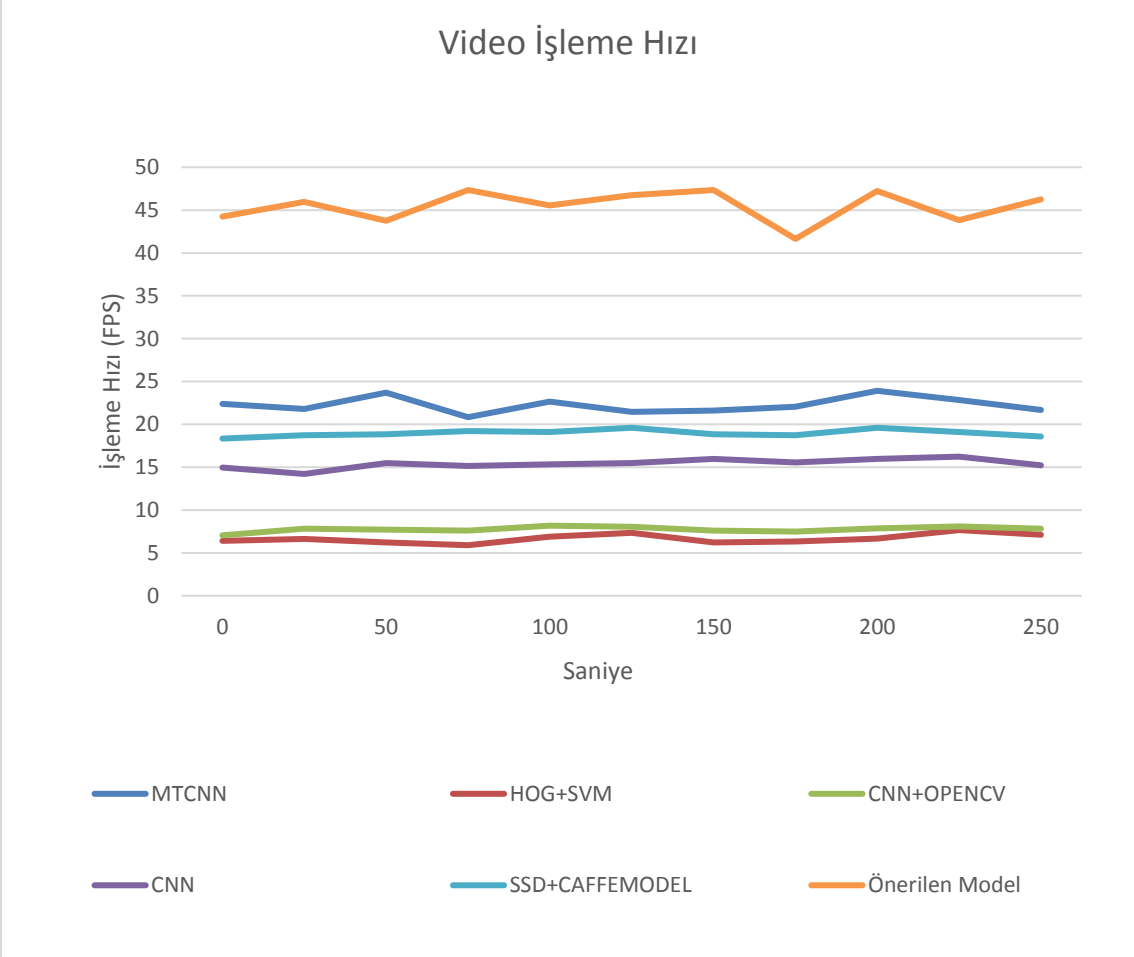
sonucunda alınan değerler kıyaslandığında performansı yüksek olan modele ulaşılır. Bu çalışmada ESA modeli ile beraber diğer tüm modellerin performans ve başarım oranları hesaplanmıştır. Bu hesaplamalar yapılırken aynı şartlar kullanılmış ve aynı veri kümesi ile testler yapılmıştır. Önerilen model ve diğer modeller için yapılan testlere ait alınan sonuçlar ve tespit edilen yüzlere ait sayılar ve görsel karşılaştırmaları Şekil 4.4'te görülmektedir.



Şekil 4.4. Uygulanan model ve tespit edilen yüz sayısı

Çalışma esnasında test amacıyla kullanılmış olan video görüntüsünde geçen görüntü karesi sayısı 6863'tür ve bu görüntü karelerine ait 4941 tane görüntüde yüz örüntüsü bulunmaktadır. Yapılan çalışmada alınan sonuçlara göre önerdiğimiz ESA tabanlı model 4902 adet yüz örüntüsünü başarılı bir şekilde tamamlamıştır. Kullanılmış olan diğer mimariler ise 4941 yüz örüntüsünden sırayla; HOG+SVM mimarisi 3082 yüz örüntüsünü, CNN+OPENCV mimarisi 3440 yüz örüntüsünü, CNN mimarisi 4137 yüz örüntüsünü, MTCNN mimarisi 4196 yüz örüntüsünü, SSD+CAFFEMODEL mimarisi 4255 yüz

örüntüsünü başarılı tamamlamıştır. Alınan sonuçlara bakıldığında sayısal olarak büyük bir fark ile ESA tabanlı modelin en yüksek performansı verdiğini görülmektedir. Aşağıda verilen Şekil 4.5'te kullanılan video görüntüsü için toplam işlenme süreleri gösterilmiştir.



Şekil 4.5. Uygulanan mimarilerin çalışma süreleri

5. DENEYSEL ÇALIŞMA SONUÇLARI

Yapılan deneysel çalışmalar sonucunda, ESA tabanlı tasarlanmış olan mimari, akan video görüntülerinde yer alan kişilerin yüzlerini başarılı bir şekilde tespit etmiş, bu kişileri tanımlamış ve edinilen bu verileri veri tabanına kaydettiği görülmüştür. Uygulama, yüz tespiti yaptıktan sonra tespit edilen yüz videoda tanımlanmış olan kişilere ait ise kişinin ismi, video içinde bulunduğu zamanı, girdideki görüntünde bulunduğu koordinatları, kare numarası ve tanımlama zamanına ait veriler kayıt altına alınmıştır. Kayıt işlemleri için, videoda tespit edilmiş kişilerin ismi, video içerisinde geçtiği zaman milisaniye cinsinden, görüntü karesi içindeki x1, y1, x2 ve y2 koordinatları, kare numarası ve tespit edilme tarihi ile veri tabanına kaydedilmesi için bir örnek Şekil 5.1'de gösterilmiştir.

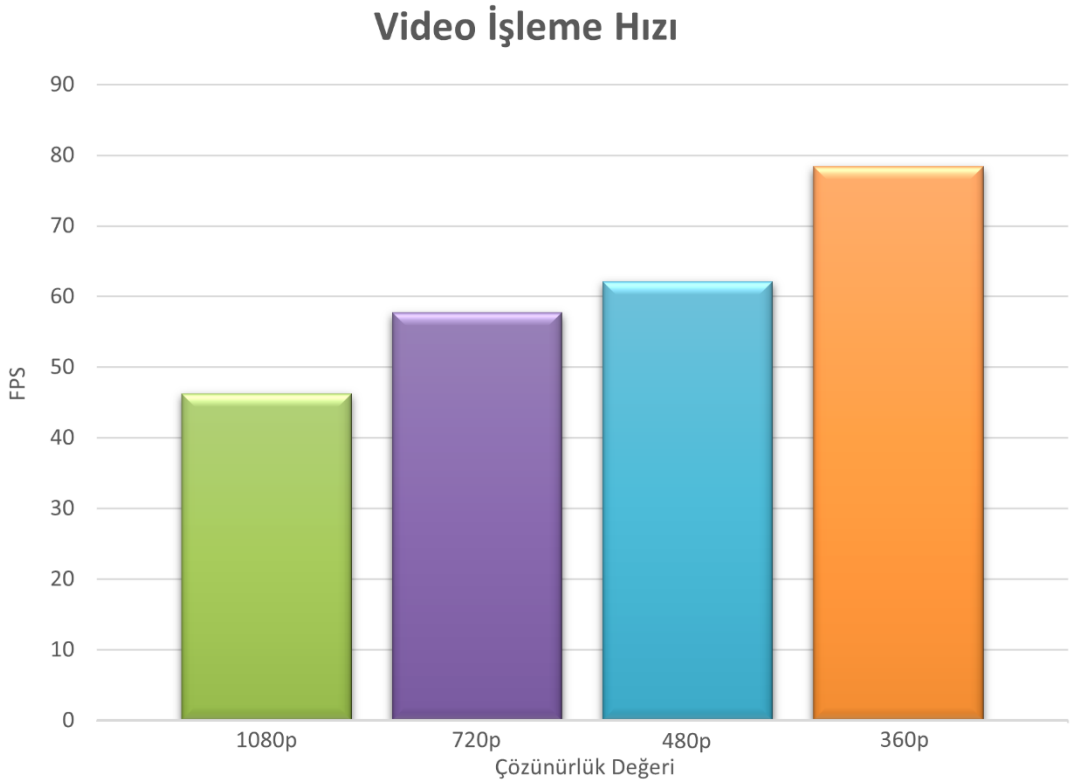

```
[  {
        "isim": "Kişi 1",
        "gectigiZaman": 8108,
        "konumu": [57, 336, 97, 297],
        "kareNumarasi": 243,
        "kayitTarihi":"30/11/2021 17:49:09.988"
    }, {
        "isim": "Kişi 2",
        "gectigiZaman": 8208,
        "konumu": [43, 366, 96, 274],
        "kareNumarasi": 246,
        "kayitTarihi": "30/11/2021 17:49:10.030"
    }
 ]
```


Şekil 5.1. Video işlemesi sonucunda oluşan veri tabanı kayıtları

Elde edilen sonuçlar JSON formatında NoSql veritabanı ortamında kayıt altına alınmıştır. Büyük veriler üzerinde veritabanı işlemlerini klasik ilişkisel veritabanlarına göre daha yüksek performans ile gerçekleştirebilen yapısından dolayı NoSql kullanılmıştır. Çalışmalarda kullanılan mimari ile deneyler yapılmış, örnek videoların 1920x1080, 1280x720, 854x480 ve 640x360 piksel genişliklerine sahip çözünürlüklerdeki versiyonları da, aynı algoritma üzerinde ayrı ayrı çalıştırılmış olup elde edilen sonuçlar karşılaştırılmıştır. Elde edilmiş olan bu sonuçlar aşağıda verilen Şekil 5.2'de gösterilmiştir. FPS değeri 29.97 olan ham video dosyasının geliştirilmiş olan algoritma üzerinde ortalama 46.24 FPS ile 1920x1080 piksel çözünürlükte gerçek zamanlı video görüntüsünden 1,54 kat daha hızlı bir

şekilde işlenmiş ve içerisindeki yüzler tanımlanmıştır. 360 piksel boyutlarındaki video görüntüsünde 78.36 FPS ile gerçek zamanlı video görüntüsünden 2,61 kat daha hızlı işlendiği ve bulunmuş olan yüzlerin tanımlandığı görülmüştür. Kısaca, yapılan deneylerle video görüntüsünün çözünürlüğü düşürüldüğünde uygulamanın video işleme hızında artış kaydedildiği görülmüştür.

Deneysel çalışma sonucunda alınan ve ilgili diğer modeller tarafından elde edilen başarı oranı Çizelge 5.1'de verilmiştir.



1080p	46,24
720p	57,7
480p	62,14
360p	78,36

Şekil 5.2. Önerilen modelin farklı video çözünürlüklerine ait işleme hızları

Çizelge 5.1. Uygulanan mimarilerin karşılaştırma tablosu

Karşılaştırma Tablosu	Elde edilen Sonuçlar		
Model	Kullanılan Veri Seti	Doğruluk (%)	F1 Skor (%)
HOG+SVM [1]	Yale	92,00	-
CNN+OPENCV [19]	Kaggle's Medical Mask Dataset	99,00	99,00
CNN [11]	AT&T	98,75	-
MTCNN [3]	Wider Face	95,40	-
SSD+CAFFEMODEL [4]	Pascal Voc	90,70	-
Önerilen Model	Labeled Faces in the Wild Home	99,43	99,60

ESA modelleri günümüzde görüntü işleme konusundan performans açısından tercih edilmekte ve yaygın olarak kullanılmaktadır. Geliştirilmiş olan mimari ise geleneksel ESA modellerinden daha az sayıda katman bulundurmaktadır. Katman sayısı 28 olarak belirlenmiş ve geliştirilmiş olan bu ESA modeli ile görüntü karelerinin işlenme hızı arttırılmıştır. Geliştirilen mimaride verilerin ön işlemesi aşamasında işlemleri hızlı ilerletmek adına ön bellekler oluşturulur. Böylece, ön bellekler kullanılarak verileri diske okuma-yazma süreçlerinden kaynaklı yaşanan gecikmeler de ön belleklere hızlı erişim sayesinde ortadan kaldırılmıştır. Oluşturulmuş olan algoritmanın paralel programlama tekniklerine uyumlu yazılmış olması görüntü işleme ve karşılaştırma aşamalarındaki gecikme seviyesini en aza indirgemiştir. Tüm bu eklenen özellik ve geliştirmelerden kaynaklı olarak birçok geleneksel yüz tanıma uygulamalarına oranla geliştirilmiş olan mimarinin performans açısından %33'e varan artış sağladığı görülmüştür.

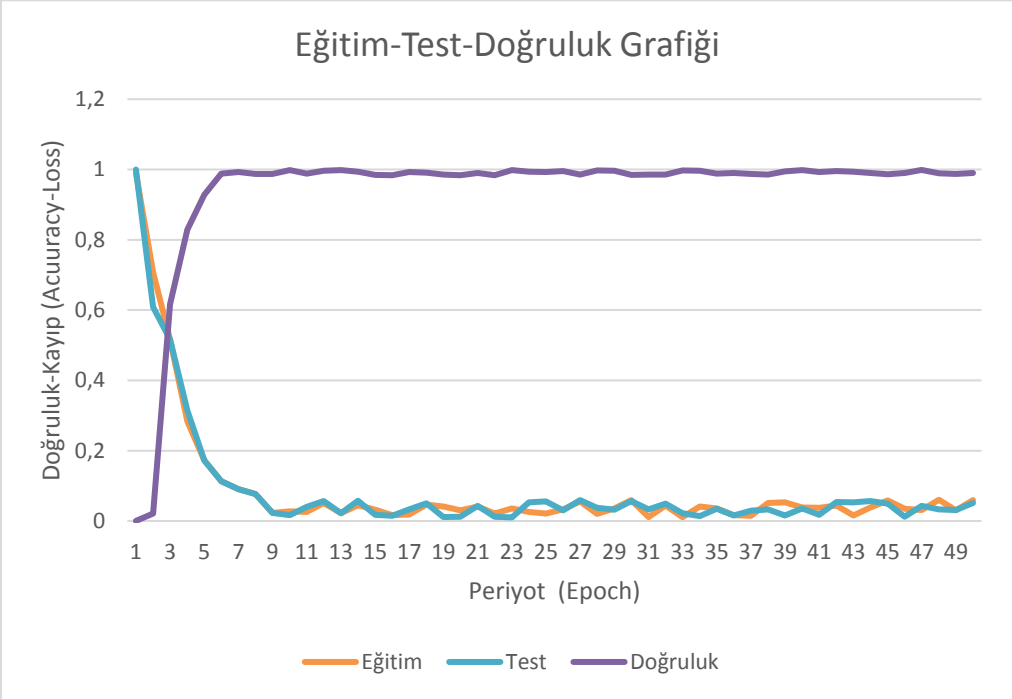
Deneysel çalışmalarda yüz tespiti için seçilmiş olan veri setindeki görüntülerden, uygulamayı eğitmek için %60'lık kısmı test etmek için %40'lık kısmı kullanılmıştır. Geliştirilen ESA modelinde aynı anda eğitilecek veri sayısı (Batch size) 16 olarak belirlenmiştir ve her bir verinin modelden geçiş sayısı (Epoch) 50 olarak seçilmiştir. Ayrıca veriler eğitim ve test süreçlerinde kullanıldıktan sonra her süreç kendi grubu içerisinde karıştırılmıştır (Shuffle). Böylece daha doğru sonuçlar alınması amaçlanmıştır.

Tüm bu değerler ESA modeli için hem literatürdeki kullanıma hem de model geliştirilmesi aşamasında elde edilen sonuçlara göre en optimal değerler olarak belirlenmiş değerlerdir. Aşağıda yer alan Çizelge 5.2'de bu parametreler ve değerleri verilmiştir. Eğitim-Test-Doğruluk grafiği Şekil 5.3'te gösterilmiştir. Grafik oluşturulurken ESA modelinin belirlenen parametreleri kullanılmıştır.

Çizelge 5.2. Uygulanan mimarinin parametre tablosu

ESA Parametre Tablosu	Kullanılan Değerler	
	Parametre	Değer
	Periyot (Epoch)	50
	Periyot Veri Sayısı (Batch Size)	16
	Karıştırma (Shuffle)	Evet

Akan video görüntülerinde tespit edilmiş olan kişilerin doğru bir şekilde tanımlanması ve kayıt altına alınması, geliştirilen mimari aracılığıyla sağlanmaktadır. Düşük çözünürlüklerdeki videolara ait görüntüler ile yüksek çözünürlüklerdeki görüntülerde yüz tespitleri arasında, uygulama performansı, doğruluk oranı ve çalışma hızı açısından karşılaştırılmıştır. Bu nedenle bir videonun farklı çözünürlüklerdeki görüntüleri de geliştirilen mimari ile test edilmiştir. Yapılan testler sonucunda tespit edilme ve işleme sürecinin gerçek video zamanından daha hızlı olduğu görülmüştür. Geliştirilen mimari kullanılarak 6863 kare bulunduran video dosyası örnek olarak taranmış ve toplamda 4941 yüz örüntüsü bulunmuş olup bunun 4902'sinin tespitinin başarılı olduğu görülmüştür. Video karesinde yer alan diğer 39 yüz örüntüsü tespit edilememiştir.



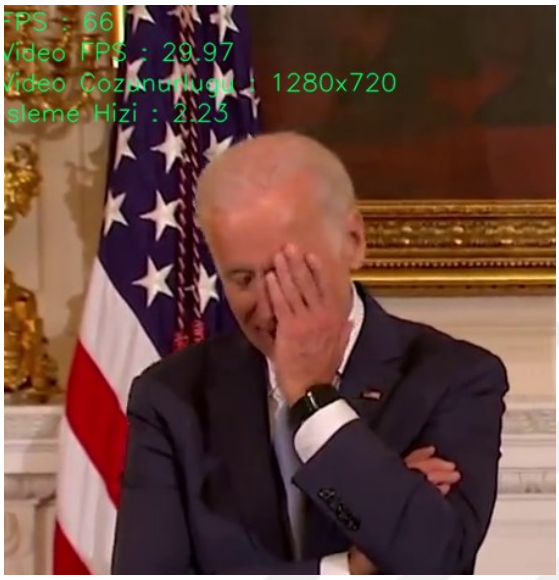
Şekil 5.3. Önerilen modelin eğitim, test ve doğruluk grafiği

Başarılı yüz ve kişi tespitine ait elde edilmiş örnek görüntü Resim 5.1'te gösterilmiştir.



Resim 5.1. Örnek görüntü içerisinde tespit edilen yüzler

Başarısız yüz ve kişi tespitine ait elde edilmiş örnek görüntü Resim 5.2'te gösterilmiştir. Resim 5.2'de de görüldüğü gibi ilgili kişi yüzünün bir kısmını veya tamamını kapatacak veya görünmesine engel teşkil edecek nesneler yüz tespit ve tanınmasını engellemektedir.



Resim 5.2. Örnek görüntü içerisinde tespit edilemeyen yüzler

Oluşturulmuş olan mimari kullanılarak yüz tespitine ait karmaşıklık matrisi elde edilmiş ve Çizelge 5.3'te gösterilmiştir. Karmaşıklık matrisinde de görüldüğü üzere 4902 görüntü karesinde yer alan yüzler başarılı bir şekilde tespit edilmiş ve tanımlanmıştır. 39 görüntü karesinde bulunan yüzler tespit edilememiş veya tanımlanamamıştır. İçerisinde yüz örüntüsü bulunmayan karelerin tamamında tespit ve tanımlama yapılmamıştır. Akan videoda kişilere ait yüzler, yüz tanımlama ile tespit edildikten sonra başarılı bir şekilde tanımlanmış olup elde edilen bilgilerin veri tabanına kayıt edilmesi sağlanmıştır.

Çizelge 5.3. Önerilen modelin karmaşıklık matrisi

Karmaşıklık Matrisi		Tahmin	
		Pozitif	Negatif
Gerçek	Pozitif	4902	39
	Negatif	0	1961

6. SONUÇ VE DEĞERLENDİRME

Her geçen gün teknolojik gelişmelerin getirdiği değişikliklerin insan yaşantısına yansıması sonucunda video kayıtları ve dijitale aktarılma işlemleri artmıştır. Saklanan videoların arşivlerde ve diğer dijital depolama alanlarında birikmesi ile bu video görsellerinin insan gözü ile taranması zorlaşmıştır. Hemen her sektörde hataya açık olmayan yüz tespiti ve kişi tanımlaması güvenlik ve tutarlılık için büyük önem taşımaktadır. Video depolarında biriken ve sayıları oldukça yüksek olan kayıtlar farklı boyut ve çözünürlükte bulunmaktadır. Bu nedenler ile akan video görüntülerinde sınıflandırma, kişileri etiketleme, kişilerin videoda geçme zamanlarını belirtme ihtiyaçlarının karşılanması için çalışmalar yapılmaktadır. Geliştirmeler ile elde edilen çalışmalarda aynı zamanda yüksek performans ve en kısa sürede doğru sonuç alma taleplerine de çözüm sağlanması amaçlanmıştır. Akan video üzerinde gerçek zamanlı insan yüzü tespitleri, tanımlama, anlamlandırma, işaretleme, zaman tespiti ve verimli bir şekilde etiketleme işlemleri geliştirilen mimaride ESA tabanlı model ile sağlanmıştır.

Deneysel çalışmalar aşamasında kişilere ait etiketlenmiş yüz görüntülerini bulunduran veri kümeleri kullanılmıştır. Veri kümeleri farklı çözünürlüklerde ve farklı boyutlarda video dosyaları içermektedir. Bununla beraber içerisinde birden fazla kişiyi aynı anda aynı karede bulunduran videolar da veri kümesinde yer almaktadır. Yapılan deneysel çalışmalar esnasında kullanılan farklı çözünürlükteki video dosyalarının ilgili video görüntüsünün gerçek zamanından 1,54 ile 2,61 kat arasında daha hızlı şekilde işlenmiş olduğu görülmüştür. İşlenme sonucunda tanımlanan kişiler veri tabanına başarılı bir şekilde kayıt edilmiştir. Yaklaşık olarak 30 FPS hızda ham video dosyaları için yüksek performanstaki grafik işlemci desteği ile test yapılmıştır. Alınan sonuçta ortalama 78,36 FPS değerlerine ulaşan hızlarda yüz tespiti ve tanımlaması yapıldığı görülmüştür. Elde edilen bilgilere bakıldığında uygulanan Evrişimli Sinir Ağı modeli ile veri kümelerindeki önceden etiketlenmiş kişilerin %99,43 oranında doğru tanımlandığı ve başarılı sonuç alındığı görülmüştür. Arşivlerde yer alan video sürelerinin uzun olması nedeniyle ihtiyaç duyulan diğer bir konu da videolarda tespit, tanıma ve etiketleme işlemlerinin hızlı yapılmasıdır. Modellerin hız açısından karşılaştırılması amacıyla yine aynı modellerin çalışma süreleri ile önerilen mimarinin çalışma süresi kıyaslanmıştır. Bunun sonucunda aynı video görüntüsü üzerinde çalışan modellerin çalışma süreleri arasında 1,84 ile 5,77 kat arasında hız artışı olduğu görülmüştür.

Sonuç olarak literatürde yer alan MTCNN, OPENCV-CNN, HOG+SVM, SSD-CAFFEMODEL modellerine kıyasla önerilen ESA tabanlı modelin performansının daha yüksek olduğu ve daha yüksek doğruluk oranına sahip olduğu yapılan deneysel çalışmalar ile gösterilmiştir. Gelecek çalışmalarda önerilen ESA modeli ile akan video görüntüleri içinde kişi sorgulama sistemi oluşturulması hedeflenmektedir.

KAYNAKLAR

1. Dadi, H.S. and Pillutla, G. M. (2018). Improved face recognition rate using HOG features and SVM classifier. *IOSR Journal of Electronics and Communication Engineering*, 11(4), 34-44.

2. Yıldız, O. (2019). Derin öğrenme yöntemleriyle dermoskopi görüntülerinden melanom tespiti: Kapsamlı bir çalışma. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 34(4), 2241-2260.

3. Zhang, K., Zhang, Z., Li Z. and Y. Qiao, (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10), 1499-1503.

4. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y. and Berg, A. C. (2016). Ssd: Single shot multibox detector. *European conference on computer vision,* 21-37.

5. McCulloch, WS. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics,* 5(4), 115–133, 1943.

6. Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 735–1780.

7. Lohr, S. (2012). The age of big data. *New York Times*, 11(2012).

8. Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition. Proceedings of the 2001 IEEE Computer Society Conference (CVPR)* on, vol. 1, 511–518.

9. Özbaysar, E. and Borandağ, E. (2018). Vehicle plate tracking system, in 2018 26th *Signal Processing and Communications Applications Conference* (SIU). IEEE, 1–4.

10. Luo, X., Shen, R., Hu, J., Deng, J., Hu, L. and Guan, Q. (2017). A deep convolution neural network model for vehicle recognition and face recognition. *Procedia Computer Science*, 107, 715-720.

11. Pranav, K. B. and Manikandan, J. (2020). Design and Evaluation of a Real-Time Face Recognition System using Convolutional Neural Networks. *Procedia Computer Science*, 171, 1651-1659.

12. Mahmood, Z., Muhammad, N., Bibi, N. and Ali, T. (2017). A review on state-of-the-art face recognition approaches. *Fractals*, 25(02).
13. Huang, G. B., Ramesh, M., Berg, T. and Learned-Miller, E. (2007). Labeled faces in the wild: A database for studying face recognition in unconstrained environments, *Technical Report*, 07-49.

14. Sharma, S., Shanmugasundaram, K. and Ramasamy, S. K. (2016). FAREC—CNN based efficient face recognition technique using Dlib. In *International Conference on Advanced Communication Control and Computing Technologies* (ICACCCT), 192-195, IEEE.

15. Schroff, F., Kalenichenko, D. and Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 815-823.

16. He, K., Zhang, X., Ren, S. and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770-778, 2016.

17. Jiang, H. and Learned-Miller, E. (2017). Face detection with the faster R-CNN. *12th IEEE international conference on automatic face & gesture recognition*, 650-657, IEEE.

18. Kalinovskii, I. and Spitsyn, V. (2016). Compact convolutional neural network cascade for face detection. In *Proceedings of the 10th Annual International Scientific Conference on Parallel Computing Technologies* (PCT), 1576,375–387.

19. Sagayam, K. M. (2021, May). CNN-based Mask Detection System Using OpenCV and MobileNetV2. In *2021 3rd International Conference on Signal Prosessin and Communication* (ICPSC), 115-119.

20. İnternet: Yapay Zekanın Tarihçesi ve Gelişim Süreci, *Medium blog*, URL: https://medium.com/türkiye/yapay-zekanın-tarihçesi-ve-gelişim-süreci-cb4c73deb01d, Son Erişim Tarihi: 01.05.2022.

21. Tezbaşaran, E. ve Gelbal, S. (2018). Temel Bileşenler Analizi ve Yapay Sinir Ağı Modellerinin Ölçek Geliştirme Sürecinde Kullanılabilirliğinin İncelenmesi. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 14 (1), 225-252.

22. İnternet: Gartner Hype Curve, URL: https://view.ceros.com/gartner/digital-marketing-and-advertising-hype-cycle/p/2 , Son Erişim Tarihi: 01.05.2022.

23. Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol Reviev*,65(6), 386-408.

24. Hornik, K., Stinchcombe, M. and White, H. (1989), Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5), 359-366.

25. Rumelhart, E., Hinton, E. and Williams, R. (1986). Learning representations by backpropagating errors. Nature,323.

26. İnternet: CS231n Convolutional Neural Networks for Visual Recognition, URL: http://cs231n.github.io/neural-networks-1/#actfun, Son Erişim Tarihi: 03.05.2022.

27. Kurt, F., (2018). Evrişimli Sinir ağlarında Hiper Parametrelerin Etkisinin İncelenmesi, Yüksek Lisans Tezi, *Hacettepe Üniversitesi Bilgisayar Mühendisliği Ana Bilim Dalı*, Ankara.

28. Clevert, D., Unterthiner, T. and Hochreiter, S. (2016). Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). *Johannes Kepler University Institute of Bioinformatics*, Austria.

29. Klambauer, G., Unterthiner, T., Mayr, A. and Hochreiter, S. (2017). Self-Normalizing Neural Networks. *LIT AI Lab. & Institute of Bioinformatics*.

30. He, K., Zhang, X., Ren, S. and Sun, J. (2015). Delving Deep into Rectifiers: Suprass,ng Human-Level Performance on Imagenet Classification. *Microsoft Research*.

31. Gao, B. and Pavel, L., (2018). On the Properties of the Softmax Function with Application in Game Theory and Reinforcement Learning,1-10.

32. Kaynar, O., Görmez, Y., Yıldız, M., ve Albayrak, A. (2016). Makine öğrenmesi yöntemleri ile Duygu Analizi. *In International Artificial Intelligence and Data Processing Symposium* (Vol. 234, p. 241).

33. Öztürk, K. ve Şahin M., (2018). Yapay Sinir Ağları ve Yapay Zeka'ya Genel Bir Bakış. *Recep Tayyip Erdoğan Üniversitesi, Mühendislik Fakültesi, Rize.*

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : GÖZE, Hüseyin
Uyruğu : T.C.

Eğitim

Yüksek lisans	Gazi Üniversitesi / Bilişim Enstitüsü	Devam Ediyor.
Lisans	Fırat Üniversitesi / Bilgisayar Mühendisliği	2012
Lise	Midyat Aziz Önen Lisesi	2007

İş Deneyimi

Yıl	Yer	Görev
2015-Halen	Estep Bilişim	Bilgisayar Mühendisi
2014-2015	EES Bilişim	Bilgisayar Mühendisi
2012-2014	Figen Yazılım Evi	Bilgisayar Mühendisi

Yabancı Dil

İngilizce, Arapça

Yayınlar

1. Göze, H. & Yıldız, O. (2022). Video Görüntülerinde Gerçek Zamanlı Yüz Tanıma ve Zaman İşaretleme için Yeni Bir Derin Öğrenme Modeli. *Bilişim Teknolojileri Dergisi*. 15 (2), 167-175 (Tezden Üretilmiştir).

Hobiler

-



GAZİLİ OLMAK AYRICALIKTIR.