

T.C.
BURDUR MEHMET AKİF ERSOY ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
YÜKSEK LİSANS TEZİ

MİKROSERVİS İLE GÖRÜNTÜ İŞLEME TABANLI
KİMLİK TANIMA VE CANLILIK ANALİZ
SİSTEMİ

Ömer Can ESKİCİOĞLU

Danışman: Prof. Dr. Ali Hakan IŞIK

BURDUR, 2024

ETİK KURALLARA UYGUNLUK BEYANI

Burdur Mehmet Akif Ersoy Üniversitesi Fen Bilimleri Enstitüsü Lisansüstü Eğitim Öğretim ve Sınav Yönetmeliğinin ilgili hükümleri uyarınca Yüksek Lisans Tezi olarak sunduğum “**Mikroservis ile Görüntü İşleme Tabanlı Kimlik Tanıma ve Canlılık Analiz Sistemi**” başlıklı bu tezin;

- Kendi çalışmam olduğunu,
- Sunduğum tüm sonuç, doküman, bilgi ve belgeleri bizzat ve bu tez çalışması kapsamında elde ettiğimi,
- Bu tez çalışmasıyla elde edilmeyen bütün bilgi ve yorumlara atıf yaptığımı ve bunları kaynaklar listesinde usulüne uygun olarak verdiğimi,
- Kullandığım verilerde değişiklik yapmadığımı,
- Tez çalışması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığını,
- Bu tezin herhangi bir bölümünü bu üniversite veya diğer bir üniversitede başka bir tez çalışması içinde sunmadığımı,
- Bu tezin planlanmasından yazımına kadar bütün safhalarda bilimsel etik kurallarına uygun olarak davrandığımı,

bildirir, aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul edeceğimi beyan ederim.

19/01/2024

(İmza)

Ömer Can ESKİCİOĞLU

ÖNSÖZ

Bu araştırma için beni yönlendiren, karşılaştığım zorlukları bilgi ve tecrübesi ile aşmamda yardımcı olan değerli Danışman Hocam Prof. Dr. Ali Hakan IŞIK'a teşekkürlerimi sunarım.

Araştırmalarım sırasında yardımlarını gördüğüm değerli can dostum ve çalışma arkadaşım Ali Tezcan SARIZEYBEK'e teşekkür ederim.

0832-YL-22 No'lu Proje ile tezimi maddi olarak destekleyen Burdur Mehmet Akif Ersoy Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü'ne teşekkür ederim.

Eğitim hayatımın her aşamasında beni her anlamda destekleyen aileme sonsuz sevgi ve saygılarımı sunarım.

Ocak, 2024

Ömer Can ESKİCİOĞLU

İÇİNDEKİLER

Sayfa

ÖNSÖZ	i
İÇİNDEKİLER	ii
ŞEKİLLER DİZİNİ	iv
ÇİZELGELER DİZİNİ	vi
SİMGELER VE KISALTMALAR DİZİNİ	vii
ÖZET	ix
ABSTRACT	x
1. GİRİŞ	1
2. GENEL BİLGİLER	3
2.1. Kaynak Özetleri	3
2.2. Görüntü İşleme	8
2.2.1. Görüntü Edinme ve Dönüştürme	8
2.2.2. Görüntü İyileştirme	9
2.2.3. Görüntü Filtreleme	10
2.2.4. Kenar ve Kontur Tespiti	11
2.2.5. Öznitelik Çıkarma ve Tanıma	12
2.2.6. Morfolojik İşlemler	13
2.2.7. Görüntü Sıkılaştırma	13
2.3. Yapay Zeka	14
2.3.1. Genetik Algoritma	14
2.3.2. Yapay Sinir Ağları	15
2.3.3. Evrimsel Sinir Ağı	16
2.3.4. Optik Karakter Tanıma	18
2.3.5. Metin Tespiti	19
2.3.6. Poz Tahmini	19
2.4. Veritabanı	20
3. MATERYAL VE YÖNTEM	21
3.1. Materyal	21
3.1.1. Veri Kümesi	21
3.1.2. Performans Metrikleri	21
3.2. Yöntem	25
3.2.1. Global Görüntü İyileştirme için Sinirsel Eğri Katmanları (CURL)	28
3.2.2. YOLOv8	29
3.2.3. Dışbükey Gövde Maksimum Alan Arama Algoritması ve GA Optimizasyonu	30
3.2.4. CloDSA	31
3.2.5. RetinaNet	32
3.2.6. Öklid Mesafesi	33
3.2.7. Perspektif Düzenlemesi	33
3.2.8. Metin Tespiti İçin Karakter-Bölge Farkındalığı (CRAFT)	34
3.2.9. OTSU Eşik Belirleme Yöntemi	35
3.2.10. Görüntü Süper Çözünürlüğü (ISR)	35
3.2.11. RDN	35
3.2.12. RRDN	36
3.2.13. EDSR	36
3.2.14. Optik Karakter Tanıma (OKT)	37

3.2.15. TesseractOCR	37
3.2.16. DonutOCR	38
3.2.17. TrOCR.....	38
3.2.18. Kosinüs Benzerliği	39
3.2.19. Yüz Doğrulama Sistemi (RetinaNet + Arcface)	39
3.2.20. Canlılık Analiz Sistemi (6DRepNet)	40
3.2.21. SQLite Veritabanı	40
3.2.22. Mikroservisler	41
4. ARAŞTIRMA BULGULARI VE TARTIŞMA.....	42
4.1. Kimlik Kartı Tanıma Performans Sonuçları	42
4.2. Canlılık ve Benzerlik Analizi Performans Sonuçları.....	50
4.3. Mikroservis Mimarisi ile Kullanıcı Arayüzü	51
4.4. Tartışma	52
5. SONUÇ VE ÖNERİLER	54
KAYNAKLAR	56



ŞEKİLLER DİZİNİ

	Sayfa
Şekil 2.1. Görüntü işlemenin temel başlıkları.....	8
Şekil 2.2. Görüntü boyut dönüşümü	9
Şekil 2.3. Görüntünün kontrast, parlaklık ve gürültü bakımından iyileştirilmesi.....	10
Şekil 2.4. Görüntü filtreleme uygulaması	10
Şekil 2.5. Derin öğrenme ile görüntü filtrelenmesi.....	11
Şekil 2.6. SIFT, BRISK ve ORB algoritmalarının görüntü üzerinde bulduğu özellikler.....	12
Şekil 2.7. Görüntü üzerinde morfolojik işlemler	13
Şekil 2.8. Genetik algoritma akış diyagramı.....	14
Şekil 2.9. Yapay sinir ağının yapısı	15
Şekil 2.10. Evrimsel sinir ağının yapısı	18
Şekil 2.11. Optik karakter tanıma uygulaması.....	19
Şekil 2.12. Metin tespiti uygulaması	19
Şekil 2.13. Örnek bir poz tahmini uygulaması (Serpa vd., 2020).....	20
Şekil 3.1. Karmaşıklık matrisi	22
Şekil 3.2. Önerilen sistemin akış diyagramı	28
Şekil 3.3. CURL mimarisi	29
Şekil 3.4. YOLOv8 mimarisi (RangeKing, 2023)	30
Şekil 3.5. Maksimum alan arama algoritması GA ile optimizasyonu	31
Şekil 3.6. RetinaNet mimarisi (Del Prete vd., 2021)	33
Şekil 3.7. CRAFT mimarisi (Baek vd., 2019)	34
Şekil 3.8. OTSU eşikleme yöntemi (Elen, 2020).....	35
Şekil 3.9. RDN mimarisi (Noura vd., 2023).....	36
Şekil 3.10. RRDN mimarisi (Idealo, 2021)	36
Şekil 3.11. EDSR mimarisi (Fallah vd., 2022)	37
Şekil 3.12. TesseractOCR mimarisi (Akhil, 2016).....	37
Şekil 3.13. DonutOCR mimarisi (Kim vd., 2021).....	38
Şekil 3.14. TrOCR mimarisi (Liv vd., 2023).....	39
Şekil 3.15. 6DRepNet mimarisi (Hempel vd., 2023).....	40
Şekil 3.16. Örnek bir mikroservis mimarisi.....	41
Şekil 4.1. CURL modelinin kimlik kartlarına uygulanması	42
Şekil 4.2. YOLOv8 örnek segmentasyon modelinin performans grafikleri	43

Şekil 4.3. Kimlik kartlarına örnek segmentasyonu uygulanması	44
Şekil 4.4. Kimlik kartında maksimum alan arama algoritmasının çıktısı.....	46
Şekil 4.5. Kimlik kartında Türk Bayrağı ve yüz görüntüsünün orta noktası	47
Şekil 4.6. Köşeler ile bulunan noktalar arasındaki öklid uzaklığı	47
Şekil 4.7. Kimlik kartı üzerinde perspektif düzenlenmesi	48
Şekil 4.8. Kimlik kartı üzerinde CRAFT modelinin uygulanması	48
Şekil 4.9. OTSU yönteminin ROI'lere uygulanması.....	49
Şekil 4.10. Orjinal görüntüye RDN, RRDN ve EDSR modellerinin uygulanması	49
Şekil 4.11. Arcface + RetinaNet ile benzerlik analizi.....	50
Şekil 4.12. 6DRepNet ile canlılık analizi sistemi (Hempel vd., 2023).....	51
Şekil 4.13. Mikroservis mimarisi ile kimlik tanıma ve canlılık analiz sistemi.....	52

ÇİZELGELER DİZİNİ

Çizelge 3.1. CloDSA'da kullanılan yöntemler ve parametreleri.....	32
Çizelge 4.1. YOLOv8 örnek segmentasyonu performans sonuçları.....	42
Çizelge 4.2. YOLOv8 nesne tespiti performans sonuçları	46
Çizelge 4.3. OKT modellerinin karşılaştırılması	50



SİMGELER VE KISALTMALAR DİZİNİ

AHE	: Adaptif Histogram Eşitleme
AUC	: Eğri Altında Kalan Alan
BRISK	: İkili Güçlü Değişmez Ölçeklenebilir Anahtar Noktalar
CLAHE	: Kontrast Sınırlı Adaptif Histogram Eşitleme
CNN	: Evrişimli Sinir Ağı
CRAFT	: Metin Tespiti için Karakter-Bölge Farkındalığı
CRNN	: Evrişimli Tekrarlayan Sinir Ağı
CTC	: Bağlantısal Zamansal Sınıflandırma
CTPN	: Bağlantılı Metin Öneri Ağı
CURL	: Global Görüntü İyileştirme için Sinirsel Eğri Katmanları
DFE	: Yoğun Özellik Birleştirme
DNN	: Derin Sinir Ağı
EAST	: Verimli ve Doğru Sahne Metni Tespiti
EDSR	: Gelişmiş Derin Artık Ağlar
Faster RCNN	: Bölge Temelli Evrişimli Sinir Ağı Hızlı R-CNN
FPR	: Yanlış Pozitif Oranı
GA	: Genetik Algoritma
GAN	: Çekişmeli Üretici Ağlar
GN	: Gerçek Negatif
GP	: Gerçek Pozitif
HE	: Histogram Eşitleme
HED	: Histogram Gradyan Kenar Dedektörü
HOG	: Yönlü Gradyan Histogramı
HPE	: İnsan Poz Tahmini
IoU	: Kesişim ve Birleşim Oranı
ISR	: Görüntü Süper Çözünürlüğü
LDCNet	: Uzun Yönü İşaretlerine Duyarlı Ağ
LSTM	: Uzun Kısa Süreli Bellek
MAE	: Ortalama Mutlak Hata
mAP	: Ortalama Hassasiyet Oranı
MSE	: Ortalama Kare Hata
OKT	: Optik Karakter Tanıma

ORB	: Yönlü Hızlı ve Döndürülmüş Kısa ve Bağımsız Özellikler
RDB	: Artık Yoğun Bloklar
RDBMS	: İlişkisel Veritabanı Yönetim Sistemi
RDN	: Artık Yoğun Ağ
ReLu	: Düzleştirilmiş Doğrusal Birim Katmanı
RNN	: Tekrarlayan Sinir Ağı
ROC	: Alıcı İşletim Karakteristiği
ROI	: İlgi Bölgesi
SFENet	: Yüzeysel Özellik Çıkarma Ağı
SIFT	: Ölçekten Bağımsız Özellik Dönüşümü
SSD	: Tek Atışlı Çoklu Kutu Dedektörü
SURF	: Hızlandırılmış ve Güvenilir Özellikler
SVM	: Destek Vektör Makinesi
TPR	: Doğru Pozitif Oranı
UPNet	: Yükseltme Ağı
YN	: Yanlış Negatif
YP	: Yanlış Pozitif
YSA	: Yapay Sinir Ağı

ÖZET

Yüksek Lisans Tezi

Mikroservis ile Görüntü İşleme Tabanlı Kimlik Tanıma ve Canlılık Analiz Sistemi

Ömer Can ESKİCIOĞLU

Burdur Mehmet Akif Ersoy Üniversitesi

Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Prof. Dr. Ali Hakan IŞIK

Ocak, 2024

Dijital kimlik doğrulama, bireylerin veya sistemlerin dijital ortamlarda kimliklerini tespit etmek ve doğrulamak için kullanılan bir süreçtir. Bu süreç, genellikle kullanıcı adları, parolalar, biyometrik veriler (parmak izi, yüz tanıma, retina tarama gibi), tek kullanımlık şifreler ve diğer güvenlik protokollerini içerir. Çevrimiçi hesaplar, finansal işlemler, sağlık kayıtları ve diğer dijital platformlarda güvenliği artırmak için yaygın olarak kullanılan dijital kimlik doğrulama yöntemleri, kullanıcı bilgilerini koruyarak yetkisiz erişimlere karşı savunma sağlamayı amaçlamaktadır. Bu çalışmada, Türkiye Cumhuriyet Kimlik Kartı görüntülerinden kart bilgilerinin optik karakter tanıma yöntemiyle çıkarılması hedeflenmiştir. Ayrıca, kimlik kartında bulunan yüz görüntüsü ile kart sahibinin kimlik kartını tanıyan kişi olup olmadığını belirlemek için yüz benzerlik analizi ve canlılık analizi yapılmıştır. Görüntü iyileştirmesi için CURL derin öğrenme mimarisi kullanılmış, ardından YOLOv8 Örnek Segmentasyonu ile kimlik kartının konumu %99,2 mAP skoru ile belirlenmiştir. Dışbükey noktalar üzerinden maksimum alanı tespit eden bir arama algoritması geliştirilmiş ve bu algoritma genetik algoritma ile optimize edilmiştir. Kimlik kartındaki Türk Bayrağı ve yüz görüntüsü tespit edilmiş, perspektif düzenlemesi yapılarak kartın köşe noktaları belirlenmiştir. Türk Bayrağının tespiti için YOLOv8 nesne tespiti modeli %97,6 mAP skoru ile eğitilmiştir. Metin tespiti için CRAFT modeli kullanılmış, ardından Otsu yöntemi ile optimal eşik değeri tespiti yapılmıştır. Görüntü kalitesini artırmak için Görüntü Süper Çözünürlüğü (ISR) modelleri kullanılmış ve TesseractOCR, DonutOCR ve TrOCR modelleri ile metin okuma işlemi gerçekleştirilmiştir. Yüz tanıma ve benzerlik analizi için RetinaNet + Arcface modeli kullanılmış, canlılık analizi için poz tahminleme mimarisi olan 6DRepNet modeli kullanılmıştır. Sistemin mikroservis mimarisine göre tasarımı tamamlanmıştır. Yapılan sistemin performansları çalışmada sunulmuştur.

Anahtar Kelimeler: dijital kimlik doğrulama, optik karakter tanıma, nesne tespiti, canlılık analizi, görüntü doğrulama

Hazırlanan bu Yüksek Lisans tezi Burdur Mehmet Akif Ersoy Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü tarafından 0832-YL-22 proje numarası ile desteklenmiştir.

ABSTRACT

M.Sc Thesis

Image Processing Based Identity Recognition and Viability Analysis System with Microservices

Ömer Can ESKİCİOĞLU

**Burdur Mehmet Akif Ersoy University
Graduate School of Natural and Applied Sciences
Department of Computer Engineering**

Supervisor: Prof. Dr. Ali Hakan IŞIK

January, 2024

Dijital identity verification is a process used to identify and authenticate individuals or systems in digital environments. This process typically involves usernames, passwords, biometric data (such as fingerprints, facial recognition, retina scanning), one-time passwords, and other security protocols. Digital identity verification methods commonly used to enhance security in online accounts, financial transactions, health records, and other digital platforms aim to provide defense against unauthorized access while safeguarding user information. In this study, the goal was to extract card information from images of the Turkish Republic Identity Card using optical character recognition (OCR) methods. Additionally, facial similarity analysis and liveness analysis were conducted to determine if the person presenting the ID card is its legitimate owner. The CURL deep learning architecture was utilized for image enhancement, followed by YOLOv8 Instance Detection for card positioning with a 99.2% mAP score. A search algorithm was developed to detect the maximum area based on convex points, and this algorithm was optimized using genetic algorithms. The Turkish flag and facial image on the ID card were identified, and corner points of the card were determined through perspective correction. YOLOv8 object detection model for the Turkish flag achieved a 97.6% mAP score. The CRAFT model was used for text detection, followed by determining the optimal threshold value using the Otsu method. Image Super-Resolution (ISR) models were employed to enhance image quality, and text reading was performed using TesseractOCR, DonutOCR, and TrOCR models. For facial recognition and similarity analysis, the RetinaNet + Arcface model was used, while the 6DRepNet model with pose prediction architecture was utilized for liveness analysis. The system design was completed according to a microservices architecture, and the performance of the implemented system is presented in the study.

Keywords: digital identity verification, optical character recognition, object detection, liveness analysis, image authentication

The present M.Sc Thesis was supported by Burdur Mehmet Akif Ersoy University Scientific Research Projects Coordinatorship Under the Project number of 0832-YL-22.

1. GİRİŞ

Günümüzde teknolojinin gittikçe artan bir şekilde hayatımıza girmesi ile beraber yeni ihtiyaçlar oluşmuştur. Sağlık, finans, eğitim, enerji, turizm, ulaşım, telekomünikasyon, eğlence, medya, güvenlik ve savunma gibi sektörlerin dijitalleşen dünya karşısında teknolojiye hızlı ve etkili bir şekilde adapte olması gerekmektedir. Sektörlerin ve kullanıcıların dijital olarak birer kimliği olup, bu kimlikler arasında haberleşmenin sağlanması iletişim ve gelişim açısından çok önemlidir (Kavut, 2020). Sektörler ve kullanıcılar arasındaki iletişim ve etkileşimler internet ortamında dijital varlık olarak bulunmalarına sebep olmaktadır. Şirketlerin ve kullanıcıların varlığını sürdürebilmesi için dijital dünya ile iç içe olması gereklidir. Bu etkileşimler ile birlikte dijital varlığa sahip olmak isteyen kullanıcıların bir dizi tanıma ve doğrulama aşamalarından geçmesi gerekmektedir. Kullanıcının dijital bir kimliğe sahip olabilmesi için yetkili kurumlardan alınan fiziksel kimlik kartı ile dijitalleşme sağlanmaktadır.

Dijital kimlik tanıma ve doğrulama sistemleri fiziksel olarak yetkili kurumlarca verilmiş olan kimlik, kart veya dökümanın dijitalleşmesi için tanımlama ve doğrulama yapılması gerekmektedir. Gerçek kimliklerini dijital ortamlarda kanıtlamaya çalışılan sürece dijital kimlik doğrulama süreci denilmektedir. Dijital kimliklerde bir dizi güvenlik önlemi bulunmaktadır. Biyometrik doğrulama, çift faktörlü kimlik doğrulama, güvenli şifre politikaları, güvenli bağlantılar, güvenli oturum yönetimi, token ve şifrelemeler bu önemlere örnek verilebilir (Varlık ve Çorumluoğlu, 2011). Bu güvenlik önlemleri dijital kimlik tanıma ve doğrulama süreçlerini daha güvenli olarak sağlamasını amaçlamaktadır.

Dijital kimlik tanımlama ve doğrulama sistemlerinde sahtekarlığı önlemek için yapay zeka teknikleri sıklıkla kullanılmaktadır. Ses tanıma, yüz tanıma, parmak izi tanıma, anomali tespiti, davranış biyometrisi, biyometrik doğrulama ve konum tabanlı doğrulama örnek verilebilmektedir. Yapay zeka temel güvenlik süreçlerinin yanında bu yöntemler ile birlikte çalışarak ekstra bir güvenlik sağlamayı da amaçlamaktadır.

Günümüzde dijital kimlikler; Çevrimiçi hizmetlerde, sosyal medya platformlarında, iş dünyasında, devlet hizmetlerinde, alışveriş sitelerinde ve sağlık hizmetlerinde sıklıkla kullanılmaktadır (Kaya, 2021). Dijital kimliklerin hayatımızda kullandığımız çoğu uygulama için önem arz etmektedir.

Dijital kimlikler günümüzde çoğu uygulama ve servislerde kullanılmaktadır. Kullanıcı ile servisler arasındaki etkileşimler ve yeni dijital bir kimlik oluşturmak isteyen bireyler için çeşitli doğrulama metotları bulunmaktadır. Kullanım açısından daha kolay

ve esnek olduđu için yapay zeka teknolojileri sıklıkla tercih edilmektedir. Yapay zeka, dijital kimlik tanıma ve doğrulama süreçlerini iyileştirmek ve güvenliğı arttırmak amacıyla kullanılmaktadır.

Dijital Kimlik tanıma süreçlerinde, fiziksel kimliğin görüntü veya video yoluyla analiz edilmesi gerekmektedir (Mir vd., 2022). Bu analiz kapsamında kartın konumu, yönü, parlaklığı ve netliği gibi çeşitli özelliklere de bakılmalıdır. Kart görüntüsü üzerinde çeşitli görüntü işleme teknikleri uygulanarak en uygun renk skalası bulunması hedeflenmektedir. Köşe noktalarının tespiti ve perspektif düzenlemesi işlemlerinden sonra kartın üzerinde metin çıkarma işlemi yapılması gereklidir. Kart üzerindeki bilgiler çeşitli algoritmalar ile alınıp dijital kimlik tanıma işlemi gerçekleştirilmektedir. Dijital kimlik tanıma sürecinde biyometrik verilerde kullanılabilir. Retina taraması, parmak izi ve yüz görüntüsü örnek verilebilir. Her insanda bulunan benzersiz doğal kimliklerde dijital kimliğiniz için yeterli bir argüman olarak gösterilebilir.

Dijital Kimlik doğrulama kapsamında kimlik ile kullanıcının eşleştirilmesi yapılması gereklidir. Kimlikteki görüntü ile kullanıcının görüntüleri karşılaştırılır ve bir benzerlik oranı elde edilir. Bu karşılaştırma sonucunda kimlik sahibi doğrulaması yapılır. Benzerlik ile eşleştirme yeterli olmamaktadır. Kimlikte bulunan bireyin fotoğrafı da bir eşleşme sağlayabilmektedir. Bu nedenden dolayı kullanıcıya canlılık analizi yapılması gerekmektedir.

Canlılık analizi kapsamında kullanıcıdan bir akış veya video boyunca çeşitli komutlar ile yönlendirme yapılır. Yapılan yönlendirmeler ile ışığında kullanıcıdan belirli bir doğruluk oranı elde edilir. Edilen doğruluk oranı ile kullanıcının fizik kimlik ile aynı kişi olduğuna karar veren sistemdir. Kullanıcı verilen komutları uygulama aşamasında yüz görüntüsü alınıp kimlikteki eşleşmeye bakılmaktadır (Okumura vd., 2019). Eğer uygun olma durumunda kullanıcı ile dijital kimliği birbiri ile eşleşir ve dijital kimlik tanıma ve doğrulama sistemi başarıyla çalıştırılmış olmaktadır.

Çalışmamızda dijital kimlik tanıma ve doğrulama kapsamında kimliğin görüntüsünden kimlik doğrulama aşamasında bulunan canlılık analizine kadar olan tüm süreçler yapılmış olup, yapay zeka ve görüntü işleme modellerinin performanslarının artırılması amaçlanmıştır. Çalışmada dijital kimlik tanıma ve doğrulama sürecini hızlı, güvenilir, esnek ve modüler bir yapıda tasarlanması hedeflenmiştir. Yüksek doğrulukta kimlik tanıma ve doğrulama işlemleri gerçekleştirilmiş olup, literatürde bulunan mevcut modeller ile karşılaştırılması yapılmıştır.

2. GENEL BİLGİLER

2.1. Kaynak Özetleri

Dijital kimlik, günümüz dünyasındaki bireylerin internet üzerinde biyometrik ve kişisel bilgilerini içeren bilgiler bütünüdür. Bu bilgiler kullanıcılar için son derece önemli olup kullanımı yaygınlaşmaktadır (Kavut, 2023). Dijital dünyada bireylerin kendini tanıtabilmesi, iletişim kurabilmesi ve internet tabanlı uygulamaların kullanabilmesini sağlamaktadır. Ayrıca hızlı bir biçimde dijitalleşmeye geçen ülkelerin ve özel kuruluşların doğrulama hizmetleri için gerekli görülmektedir. Dijital kimlik, dijital toplum olma yolunda büyük bir adımdır. Bu durum kimlik doğrulama sistemlerinin daha akıllı hale gelmesine ve güvenliğin önemsenmesi gibi sonuçlar yaratmaktadır. Dijital kimlikler kullanıcıların biyometrik veya fiziksel kimliklerinin bir dijital kopyası olarak düşünülmelidir.

Dijital kimlikler birçok alanda kullanılmaktadır. Devlet hizmetlerinde, eğitim hizmetlerinde, sağlık hizmetlerinde, ulaşım hizmetlerinde, finansal hizmetlerde, e-ticaret hizmetlerinde ve sosyal medya platformlarında sıklıkla faydalanılmaktadır. Bu hizmetlerde dijital kimlik tanıma ve doğrulama durumlarında yeni teknolojiler de kullanılmaktadır. Görüntü işleme, derin öğrenme, 5G iletişimi ve blok zinciri gibi teknolojiler mevcut sektörlerle ve hizmetlere adaptasyon sağlayabilmek ve kaynakların verimli kullanımı açısından sürekli olarak geliştirilmektedir. Bu teknolojilerin toplumda dijitalleşme adına çok sayıda yararı bulunmaktadır. Maliyetlerin azaltılması, güvenliği artırılması, şeffaflığın sağlanması, daha iyi bir müşteri deneyiminin sunulması, daha verimli ve hızlı sistemlerin oluşturulması gibi olanaklar sağlamaktadır (Parate vd., 2023).

Kimlik doğrulama için maliyet, güvenlik ve şeffaflık doğrultusunda blok zinciri tabanlı sistemlerin yeterli olduğu düşünülmektedir. Bu kapsamda Jamal ve arkadaşları tarafından gerçekleştirilen çalışmada, blok zinciri tabanlı bir kimlik doğrulama sistemi oluşturulmuştur. Oluşturulan sistem, veri ihlallerinin, yetkisiz erişimin, belge kaybının ve merkezi yapıdan kaynaklanan sorunların çözümü olarak gösterilmektedir (Jamal vd., 2019). Böylelikle güvenli bir kimlik doğrulamanın, blok zinciri ile dijitalleşmesi gerekliliğini çalışmalarında vurgulamıştır.

Kimlik doğrulama süreçlerinin gerçekleştirilmesi için fiziksel kimliğin gerçek zamanlı tanımlanması gerekmektedir. Görüntü işleme tabanlı derin öğrenme modelleri ile kimlik tanımlanma aşamasın gerçekleştirilmektedir. Çetiner ve arkadaşları tarafından yapılan çalışmada Türkiye Cumhuriyeti Kimlik kartlarında bulunan kimlik numarasının

görüntü işleme teknikleri ve Yapay Sinir Ağları (YSA) kullanılarak tespiti ve tanımlaması yapılmıştır. Öznitelik çıkarma işleminde İki Boyutlu Ayırık Dalgacık yöntemi kullanılmıştır. Bu yöntemde bir 3. Seviyeden ‘db2’ türü kullanılarak hesaplama zamanı kısaltılmıştır. Test aşamasında 88 adet kimlik numarasının %100 başarı ile tespiti gerçekleştirilmiştir. Kimlik numarasının tanımlama işlemi yaklaşık 0.25 saniye olduğu gözlemlenmiştir. Deney aşamasında görüntü tanıma problemlerinde Levenberg-Marquardt algoritmasından yeterli performans elde edilemediği için kullanılmamıştır (Çetiner vd., 2012). Satyawan ve arkadaşları tarafından gerçekleştirilen bir çalışmada Endonezya Elektronik Kimlik Kartlarının görüntü işleme ve Optik Karakter Tanıma (OKT) ile kimlik kartı tespiti ve bilgi çıkarımı yapılmıştır. Bu çalışmada öncelikle ikili görüntü sonuçlarını elde etmek için görüntüler gri tonlamaya çevrilmiştir. Daha sonra bir eşik değeri verilerek, görüntüler üzerinde kenar tespiti için Sobel filtresi uygulanmıştır. Elde edilen görüntüde gürültüyü azaltmak için morfolojik dönüşüm teknikleri kullanılmıştır. Görüntü üzerindeki optimal eşik değerini belirleyebilmek Otsu metodu kullanılmıştır. Kimlik kartı üzerinde bulunan metin alanlarını çıkarım işlemlerinden sonra tesseract 4.0 ile OKT işlemi gerçekleştirilmiştir. Yapılan sistemde NIK numarasını %100 doğrulukta ve kimlik kartının ismini %98,6 doğrulukta tespit, sağlanmıştır (Satyawan vd., 2019). Tan ve arkadaşları tarafından gerçekleştirilen çalışmada Vietnam kimlik tanıma sistemindeki bozuk belge sorununu çözmek için bir yöntem önerilmiştir. Önerilen yöntemde öncelikli olarak ilgi bölgesinin (ROI) tespiti, görüntü segmentasyonu ve köşe noktalarının tespiti yapılmıştır. Elde edilen sonuçlar doğrultusunda bozuk olan görüntü alanını düzeltme işlemi gerçekleştirilmiştir. Çalışmada nesne tespiti için optimize edilmiş modüller içeren, hızlı ve kolay bir mimari olan PP-YOLO kullanılmıştır. Kimlik kartı tespit edilme işleminden daha sonra kartın köşeleri belirlenip, perspektif düzenlemesi yapılmıştır. PP-YOLO tabanlı kimlik kartı karakter tanıma modeli %99,85 ile %99,95 aralığında hassasiyet sonucu elde edilmiştir. Ayrıca ortalama düzeltme süresi verilen her görüntü için 0,03 ile 0,06 saniye arasındadır. Çalışmada %97,65 doğruluk oranı elde edilmiştir (Tan vd., 2023). Tamirat tarafından yapılan tez çalışmasında kimlik kartlarının dijitalleşmesini sağlamak için tam otomatik bir yöntem önermiştir. Görüntü ön işleme yöntemlerinden perspektif düzenlemesi uygulanmıştır. Piksellerin çevresindeki yoğunluğa göre adaptif bir eşik belirleyen Sauvola eşikleme yöntemini kullanmıştır. Segmentasyon teknikleri, bağlantı bileşenleri ve morfolojik genişletme yöntemleri ile metin alanlarının ayrımı yapılmıştır. Herhangi bir perspektif ve eğri düzeltme yöntemleri kullanılmadan %90,87 hassasiyet oranı elde edilmiştir. Perspektif ve eğri düzeltmeleri

uygulandıktan sonra ise %93,6 hassasiyet oranına ulaşılmıştır. Çalışmada Google Tesseract OKT modelini kullanarak kimlik kartı tanımlama işlemi gerçekleştirmiştir (Tamirat, 2023).

Kimlik tanımlama ve doğrulama işlemlerinde kart üzerindeki bilgilerin yüksek doğrulukla alınması, bilgi güvenliği ve bütünlüğü bakımından önemlidir. Literatürde yapılan çalışmalarda OKT modelinin kimlik üzerindeki verileri okuma doğruluğu, uygulanan ön işleme metotlarının performansına bağlı olduğu anlaşılmaktadır. Bu kapsamda Das ve arkadaşları tarafından geliştirilen DewarpNet mimarisinde 2B ve 3B regresyon ağlarını kullanarak bütünlüğü fiziksel olarak bozulan dökümanların düzeltilmesi ve iyileştirilmesi amaçlanmıştır. DewarpNet ile OKT performansı önemli bir derecede artmış olup hata miktarının %42 oranında azaldığı çalışmada tespit edilmiştir (Das vd., 2019).

OKT, kimliği tanımlama işlemindeki temel unsurlardan biridir. Metin içeren görüntülerde güçlü öznitelik çıkarma yöntemleriyle birlikte derin öğrenme algoritmalarının bir arada kullanılması model doğruluğu açısından istenilen bir durumdur. Van Hoai ve arkadaşları tarafından yapılan çalışmada Vietnam Kimlik Kartlarını Derin Özellikli Ağlar kullanarak metin tanıma işlemi gerçekleştirilmiştir. Düşük ışık, flaş ışığı ve güneş ışığı görüntü niteliklerine sahip 2500 adet veriden oluşan bir veri kümesi kullanılmıştır. Evrişimli Tekrarlayan Sinir Ağı (CRNN) mimarisi Adam optimizasyon yöntemi ile 6 turluk bir eğitim yapılmıştır. Önceden eğitilmiş bir Bağlantılı Metin Öneri Ağı (CTPN) kullanarak tüm kırılmış görüntüler CRNN modeline, Tencent OKT, Google OKT API'lerine gönderilmektedir. CRNN modelinin diğer OKT API'lerine oranla daha başarılı olduğu anlaşılmaktadır. Bu model ile kelime düzeyinde %89,7 ve karakter düzeyinde ise %96,7 doğruluk elde edilmiştir (Van Hoai vd., 2021). Rusli ve arkadaşlarının yapılan çalışmada Endonezya Kimlik Kartlarında bulunan metinlerin tespiti için PyOCR, Tesseract ve Pytesseract olmak üzere 3 ayrı OKT modeli kullanılmış ve karşılaştırılması yapılmıştır. Yapılan karşılaştırma sonucunda 50 test verisi üzerinde gerçekleştirilen çalışmada %78 F1-Skoru ve 4,51 saniye hesaplama süresi ile Pytesseract en performansı göstermiştir (Rusli vd., 2021). Fang ve arkadaşları tarafından yapılan çalışmada Çin Kimlik Kartında bulunan karakterlerin tanımlaması yapılmıştır. Tekli karakter işleme, şablon eşleme yöntemi ve Destek Vektör Makinesi (SVM) makine öğrenmesi modeli uygulanmıştır. Şablon eşleme yöntemi olarak XOR işlemi ve Öklid uzaklığı kullanılmıştır. Oluşturulan sistem dijital kart tanımlamada %90 doğruluk oranına sahipken Çin karakterleri tanımlamada %70 doğruluk oranına sahiptir (Fang vd., 2017).

Baek tarafından yapılan çalışmada metin alanlarının tespiti için karakterler arasında bağlantı durumlarını hesaplayan yeni bir sahne metin tespit yöntemi önerilmiştir. Önerilen yöntemin, karmaşık metin görüntülerini yüksek algılama hassasiyetine sahip olduğu gösterilmiştir. VGG-16 tabanlı bir ağ mimarisi kullanılmıştır. Çalışmada %97,4 oranında kesinlik, %93,1 oranında geri bildirim skoru elde etmiştir. Çalışmada ayrıca diğer literatürdeki yöntemler ile karşılaştırılmış olup bu yöntemin yüksek başarı oranına dikkat çekilmiştir (Baek vd., 2019).

Kimlik tanıma işlemleri kapsamında kimlik tespiti, perspektif düzenlemesi, kimlik kartından öznitelik çıkarımı, kelime ve metin tespiti, OKT ve derin öğrenme modelleri sıklıkla kullanılmaktadır. Kimlik kartındaki bilgiler bahsedilen yöntemler ile elde edildiği takdirde doğrulama aşamasına geçilmesi gerekmektedir. Kimlik doğrulama aşaması ise kimliği dijital hale getirmek isteyen birey ile kimlik sahibinin eşleşmesini sağlamak ve karşıdaki kişinin gerçek bir birey olduğunu anlamaktır. Bu kapsamda karşıdaki kişiye canlılık analizi uygulanması gerekmektedir. Gerçek zamanlı olarak çeşitli davranışsal komutlar vererek karşıdaki kişinin canlı olduğunu anlamak kimlik güvenliği anlamında önemlidir. Kimliğin çalıntı olup olmadığını ve kimlik görüntüsündeki kişinin gerçek zamanlı olarak davranışsal komutlarla işlem yapan kişinin aynı olduğunu anlamamız gereklidir. Bu gerekçelerden dolayı kimlik tanıma ve canlılık analizi sistemi birlikte kullanılmalıdır. Liu ve arkadaşları tarafından yapılan çalışmada 2B İnsan Poz Tahmini (HPE) kimlik doğrulama davranış anlama gibi birçok alanda kullanımına vurgu yapılmış olup sınırlı senaryolara ve sıkı kurallar gibi kısıtlama ile karşılaştığı belirtilmiştir. Bu sıkıntıları çözmek için Uzun Yönü İşaretlerine Duyarlı Ağ (LDCNet) modeli önerilmiştir. MS-COCO ve MPII veri kümeleri üzerinde gerçekleştirilen çalışmada LDCNet modelinin state-of-the-art yaklaşımlara üstünlüğü anlaşılmaktadır (Liu vd., 2023). Dang ve Tran'ın yaptığı çalışma kapsamında sahte giriş görüntülerini engellemeyi amaçlayan bir yüz tanıma yaklaşımı önerilmiştir. MTCNN ve FaceNet ağlarının birlikte kullanılmasıyla oluşturulmuş yöntem, tanıma verimliliğinin artırılmasını amaçlamaktadır. Poz tahminleme sırasında başın 5 durumu üzerine yoğunlaşmıştır. Model, başın sola, sağa, yukarı, aşağı ve düz pozisyonlarında tespitini yapmaktadır. Jetson Nano kartının üzerinde çalışan bu sistemin doğruluk oranı %90 ile %95 arasında ve görüntü işleme hızı ise 21-25 fps arasındadır. Önerilen sistem son kullanıcı tanıma işlemlerinde kullanılabilir olduğu vurgulandı (Dang ve Tran, 2023). Wu ve Trivedi tarafından yapılan çalışmada baş poz tahmini yapabilmek için 2 aşamalı bir çerçeve sunulmuştur. Çalışmada Gabor dalga dönüşümü ile öznitelik çıkarımı yapılmıştır. Veriyi

daha iyi anlayabilmek için PCA, FDA, KPCA ve KDA kullanılmıştır. İkinci aşama olan entegre yaklaşımında ise poz tahmini geometrik yapılar oluşturularak ve yapı ayrıntıları analiz edilerek iyileştirilmiştir. Çalışma kapsamında -90 ile +90 derece ve eğimi -60 ve 45 derece 86 poz incelenmiştir. Elde edilen sonuçlara göre entegre yaklaşımın alt uzay analizine göre daha iyi bir performans sergilediği ortaya konulmuştur (Wu ve Trivedi, 2008). Najafzadeh ve arkadaşları tarafından yapılan çalışmada yüz görüntü kalitesi için bir vektör üreten çok-görevli bir Evrişimli Sinir Ağı (CNN) önerilmiştir. Önerilen yöntem ile görüntü ve yüz kalitesiyle ilgili bilgiler içeren bir vektör yardımıyla önceden eğitilmiş ResNet-50 mimarisini temel olarak almaktadır. Çalışmada uygulanan algoritma CMU Multi-PIE yüz veri kümesinden 50 farklı kimlik seçilmiştir. Test aşamasında model, poz tahminleme için %92,71 ve yüz ifadesi tanıma için %87,35 doğruluk oranı elde edilmiştir. State-of-the-art yöntemlerine göre daha yüksek başarımlar sağladığı görülmüştür (Najafzadeh vd, 2023).

Literatürdeki çalışmalardan anlaşıldığı gibi poz tahminleme ile kimlik doğrulama işlemi yapılabilmektedir. Kimlik tanıma işlemi ile kimlik doğrulama işleminin ard arda senkron olarak çalışması gerekmektedir. Çeşitli yazılım çözümleriyle bu durum aşılabılır. Mikroservis mimarisi ile istemci-sunucu mantığında sistemler günümüzde sıklıkla kullanılmaktadır. Gadea ve arkadaşları tarafından geliştirilen çalışmada mikroservisler ile zengin bir metin düzenleme mimarisi önerilmiştir. Önerilen sistem ile birden fazla kullanıcının belgeyi düzenleyebilmesine ve içeriğin bir parçası olarak yüz tanıma ile eklenen yüzleri belgenin bir parçası olarak göstermiştir. Çalışmada konteyner, selenium, nodejs, websocket ve REST API gibi çerçeveler ve teknolojiler kullanılmıştır (Gadea vd.,2016). Silva ve arkadaşları tarafından yapılan çalışmada suç delillerinin analizinde görüntü sınıflandırıcıları kullanarak mikroservisler ile bir sistem oluşturulmuştur. Oluşturulan sistem ile Brezilya Rio Grande do Norte Bölge Savcılığı'nın adli tıp analistlerinin ihtiyaçları göz önünde bulundurularak kimlik kartı, metin belgeleri, ekran görüntüleri, çıplaklık içeren görseller, ateşli silahlar ve mühimmat gibi nesne türlerinin sınıflandırılması yapılmaktadır. Çalışmada önerilen mimari mikroservis geliştirme yaklaşımı ve Apache Spark gibi büyük veri işleme ve analiz çerçevesi kullanılmıştır (Silva vd.,2021).

Literatürdeki araştırmalar ve çalışmalar incelendiğinde kimlik tanıma ve doğrulama sistemleri ile benzer çalışmalar, projeler ve araştırmalar olduğu görülmüştür. Ancak incelenen çalışmalarda kimlik tanıma ve doğrulama aşamasındaki canlılık analizi ayrı olarak ele alınmıştır. Çalışmamızda kimlik tanıma ve doğrulama aşamalarını görüntü

işleme ve yapay zeka yöntemleri ile destekleyen bir mikroservis tabanlı genel bir yazılım çözümü önerilmiştir. Önerilen yöntemde kimlik kartı tespiti, tespit edilen kimliğin perspektif düzenlemesi, OKT modeli, poz tahmini ve benzerlik analizleri yapılmıştır.

2.2. Görüntü İşleme

Dijital görüntüler ile çeşitli operasyonlar gerçekleştirilmesi sonucu istenilen amaca yönelik görüntü oluşturma sürecini ifade etmektedir (Baskar vd., 2023). Bu süreç, çeşitli aşamalardan oluşmaktadır. Şekil 2.1’de gösterildiği üzere görüntü edinme ve dönüştürme, görüntü iyileştirme, renk dönüşümleri ve renk uzayları, görüntü filtreleme, kenar ve kontur tespiti, öznitelik çıkarma ve tanıma, nesne tespiti ve takibi, morfolojik işlemler, doku ve desen analizi ve görüntü sıkılaştırma uygulamaları bu temel başlıklar arasındadır. Tıp, fizik, sanayi, güvenlik, finans, tarım teknolojileri, bilgisayar teknolojilerinde ve askeri sanayi gibi birçok alanda sıklıkla kullanılmaktadır.

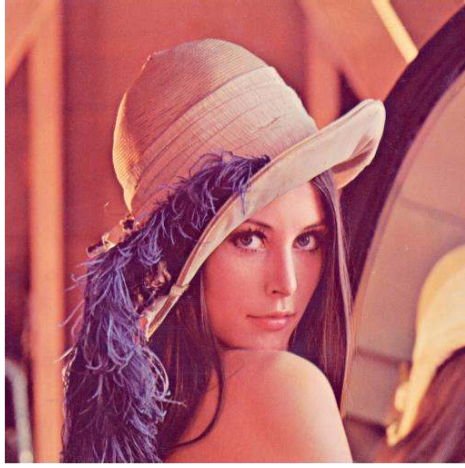


Şekil 2.1. Görüntü işlemenin temel başlıkları

2.2.1. Görüntü Edinme ve Dönüştürme

Optik tabanlı çalışan elektronik bir cihaz, sensör veya kamera kullanarak görüntü elde edilme sürecine denilmektedir (Yıldız ve Duru, 2023). Elde edilen görüntü üzerinde işlem yapılabilir formata dönüştürülmesi gerekmektedir. Dönüştürme işleminde farklı özellik ve biçimlere çevrilmesi yapılmaktadır. Görüntü dönüştürme işlemleri boyut ve renk uzay dönüşümleri, görüntü filtreleme ve ışık düzeltmeleri konu başlıklarını içermektedir.

512 x 512 Boyutlarındaki Görüntü



224 x 224
Boyutlarındaki Görüntü



100 x 100
Boyutlarındaki
Görüntü



Şekil 2.2. Görüntü boyut dönüşümü

Derin öğrenme algoritmalarına önerilen görüntü giriş boyutuna göre boyut dönüşümü yapılması gerekmektedir. Şekil 2.2’de örnek bir görüntü boyut dönüşümü uygulaması yapılmaktadır. Algoritmalara verilecek dijital görüntülerde renk uzay dönüşümleri, filtreleme işlemleri ve ışık düzeltmelerini yapılması performans açısından önem arz etmektedir.

2.2.2. Görüntü İyileştirme

Dijital bir görüntünün kalitesini arttırmak, daha iyi bir öznitelikler ve anlamlı bilgiler çıkarmak amacıyla yapılan işlemlere denir. Görüntü işlemenin önemli bir aşaması olarak kabul edilmektedir. Birçok alanda farklı uygulamalarda kullanılmaktadır. Dijital görüntü iyileştirme için; ışıklandırma düzeltmeleri, gürültü giderme, kontrastın iyileştirilmesi, histogram denkleştirme işlemleri, renk düzenlemeleri, morfolojik işlemler, görüntü filtrelemeleri ve çeşitli derin öğrenme modelleri ile görüntü iyileştirme işlemi yapılmaktadır. Şekil 2.3’de görüntü iyileştirme kapsamında bir örnek gösterilmektedir.

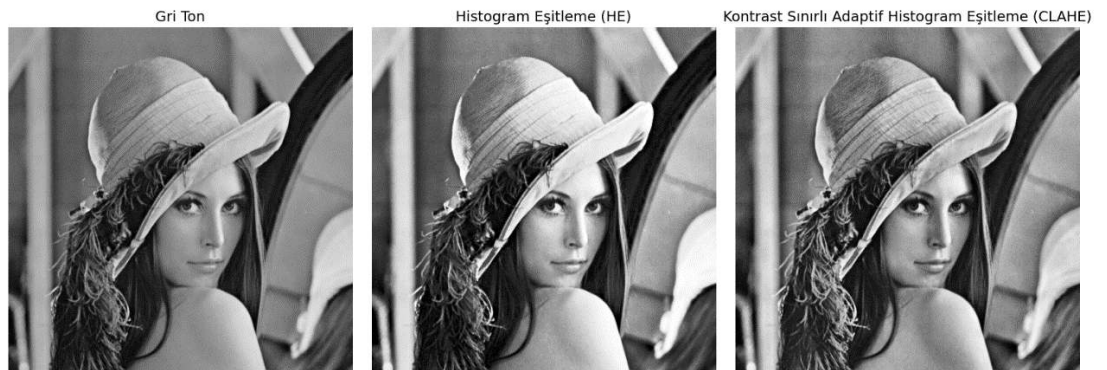


Şekil 2.3. Görüntünün kontrast, parlaklık ve gürültü bakımından iyileştirilmesi

Görüntü iyileştirme aşaması filtreler veya aritmetik işlemlerin yanında derin öğrenme yöntemleriyle de yapılabilmektedir Genellikle Çekişmeli Üretici Ağlar (GAN) kullanılarak yapılan işlemlerde daha iyi bir şekilde görüntü iyileştirme uygulamaları bulunmaktadır (Şenalp, 2023).

2.2.3. Görüntü Filtreleme

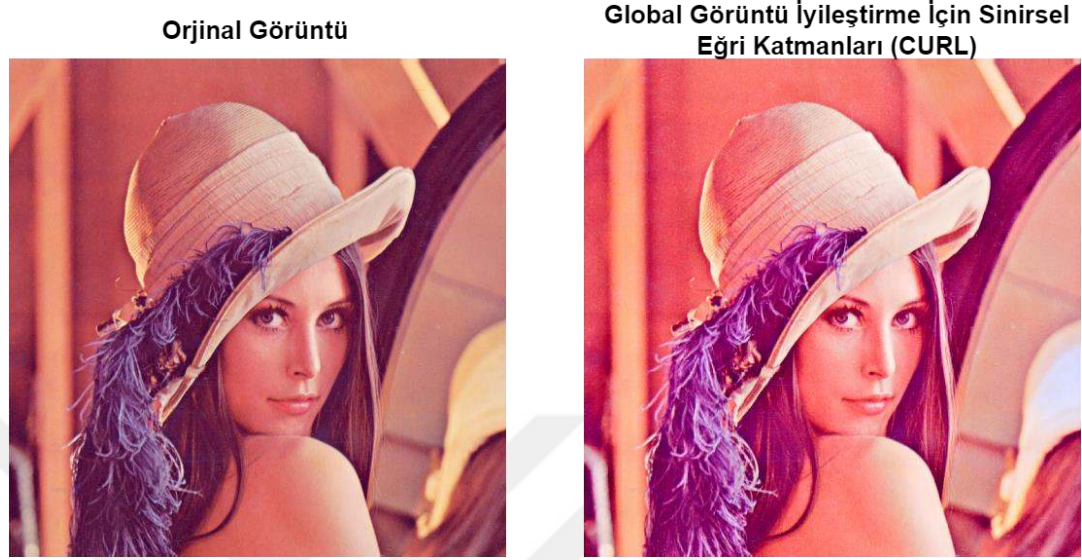
Dijital bir görüntü üzerinde özel filtreler, matematiksel ve aritmetiksel işlemler sonucunda görüntüdeki bilgilerin kısmi olarak değiştirilmesi veya güncellenmesi işlemidir. Görüntü filtreleme genellikle görüntüdeki öznitelikleri ve belirli bilgileri ortaya çıkarmak için uygulanmaktadır. Bilgisayarda görü, medikal görüntülerde, kameralarda ve birçok alanda kullanılmaktadır (Gezer, 2021).



Şekil 2.4. Görüntü filtreleme uygulaması

Görüntü filtreleme kapsamında Ortalama filtresi, Gauss filtresi, Medyan filtresi, Sobel filtresi, Prewitt filtresi, Laplacian filtresi, Histogram Eşitleme (HE), Adaptif

Histogram Eşitleme (AHE), Kontrast Sınırlı Adaptif Histogram Eşitleme (CLAHE) ve Bilateral filtreler gibi birçok filtre bulunmaktadır (Yakut ve Ulukaya, 2022). Şekil 2.4’de örnek bir görselin gri tonlama, HE ve CLAHE filtrelerinin uygulanması verilmiştir.



Şekil 2.5. Derin öğrenme ile görüntü filtrelenmesi

Derin öğrenme yöntemleriyle de filtreleme işlemleri yapılmaktadır. Derin öğrenme algoritmalarıyla görüntülerden alınan bilgilerin daha kaliteli olmasını sağlamak ve bilgi çıkarımını daha iyi yapabilmek için çeşitli modeller eğitilmiştir. Şekil 2.5’de Global Görüntü İyileştirme için Sinirsel Eğri Katmanları (CURL) mimarisi kullanılarak görüntü filtreleme işlemi yapılmıştır (Moran vd., 2021).

2.2.4. Kenar ve Kontur Tespiti

Kenar ve kontur tespiti görüntü üzerindeki bölgelerin veya nesnelerin dış hatlarını belirleme sürecine denir. Görüntüdeki yoğunlukları veya renk değişimlerini analiz ederek görüntünün içerisinde bulunan nesnelerin sınırlarını belirlemeyi amaçlamaktadır (Hanbay, 2022). Nesne tanıma, nesne takibi ve görüntü analizi kapsamında sıklıkla kullanılmaktadır. Güvenlik sistemlerinde, otomasyon sistemlerinde, otomotiv endüstrisinde, kalite süreçlerinde, meteoroloji alanında, robotik sistemlerde, yüz ve nesne tanıma gibi alanlarında kullanılmaktadır. Görüntüdeki yoğunluk veya renk değişimini hesaplayan birçok algoritma bulunmaktadır. Sobel operatörü, Prewitt operatörü, Canny kenar dedektörü, Scharr filtresi, Histogram Gradyan Kenar Dedektörü (HED) algoritmaları bu algoritmalarından bazılarıdır.

2.2.5.Öznitelik Çıkarma ve Tanıma

Dijital görüntü işleme kapsamında belirli bir deseni veya özneliği tespit etme bilgi çıkarma sürecine öznitelik çıkarma ve tanıma denilmektedir. Görüntülerde bulunan öznitelikler algılanan desenler ile ilişkilendirme yapıp tanıma işlemi gerçekleştirilmektedir. Bu kapsamda Yönlü Gradyan Histogramı (HOG), Hızlandırılmış ve Güvenilir Özellikler (SURF), Ölçekten Bağımsız Özellik Dönüşümü (SIFT), Yönlü Hızlı ve Döndürülmüş Kısa ve Bağımsız Özellikler (ORB), İkili Güçlü Değişmez Ölçeklenebilir Anahtar Noktalar (BRISK) ve Evrimsel Sinir Ağları (CNN) yöntemleri kullanılmaktadır (Tareen ve Raza, 2023). Şekil 2.6’da örnek bir kullanım gösterilmektedir.



Şekil 2.6. SIFT, BRISK ve ORB algoritmalarının görüntü üzerinde bulduğu özellikler

2.2.6. Morfolojik İşlemler

Morfolojik işlemler görüntüdeki yapı ve nesneleri ortaya çıkarmak ve vurgulamak amacıyla yapılmaktadır. Genellikle siyah-beyaz görüntüler üzerine uygulanır (Lotufo vd., 2023). Nesne ayırma, nesne ve kenar tespiti gibi uygulamalarda da kullanılmaktadır. Erozyon, genişleme, açılma, kapanma ve yarıma gibi morfolojik işlemlere sahiptir. Şekil 2.7’de örnek bir morfolojik işlem uygulaması verilmiştir.



Şekil 2.7. Görüntü üzerinde morfolojik işlemler

2.2.7. Görüntü Sıkılaştırma

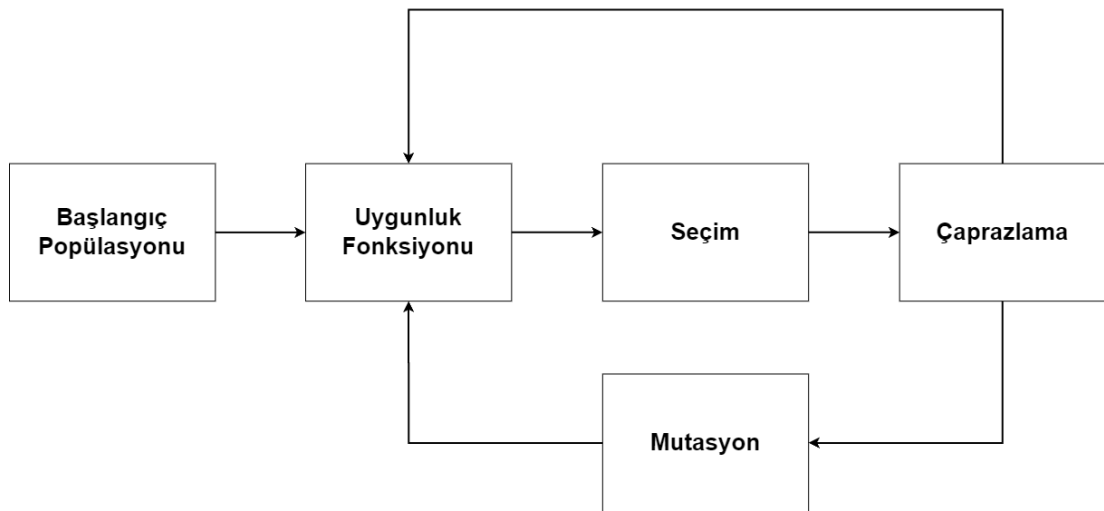
Görüntü verilerinin daha verimli depolanmasını veya iletilmesini amacıyla kullanılan yöntemlere görüntü sıkılaştırma teknikleri denir. Bu teknikler ile depolama alanını optimize ederek, görüntüde bulunan gereksiz bilgileri ortadan kaldırmaktadır (Fraihat, S. ve Al-Betar, 2023). Görüntüyü iletmek amacıyla görüntü boyutunun azaltılması veri iletim hızının yükselmesine sebep olmaktadır. Kayıpsız ve kayıplı görüntü sıkılaştırma olarak 2 temel kategoriye ayrılmaktadır. Görüntü sıkılaştırma yöntemleri frekans dönüşümleri veri tekrarı ve renk skalasındaki istatistikleri kullanarak çalışmaktadır. Görüntü sıkılaştırma algoritmalarına örnek olarak JPEG, MPEG, LZW ve RLE algoritmaları gösterilebilir.

2.3. Yapay Zeka

Alan Mathison Turing tarafından “Makineler düşünebilir mi?” sorusunu ortaya atarak makinenin düşünebilme ve zeka yeteneğini tartışmaya açmıştır (Shieber, 2004). Daha sonra bu süreçte 1956 yılında Dartmouth Üniversitesi'nde bir konferansta “yapay zeka” terimi önerilmiştir (Zhang ve Lu, 2021). Yapay zeka, makinelerin insan benzeri zeka, öğrenebilme ve düşünebilme kapasitelerini taklit edebilmesi amacıyla geliştirilmiş bir bilim dalıdır (Whitby, 2019). Geliştirme çalışmaları günümüzde de hala devam etmektedir. Yapay zeka kendi içerisinde temel konu alanlarına ayrılmıştır. Bu konu alanları olarak derin öğrenme, makine öğrenmesi, doğal dil işleme, robotik, görüntü işleme, bilgi madenciliği, evrimli algoritmalar, büyük veri analizi, bilişsel robotik, üretici sistemler gibi kendi içerisinde birçok alana sahiptir. Günümüzde yapay zeka; finans, eğitim, sağlık, perakende, otomotiv, telekomünikasyon, enerji, havacılık, uzay, hukuk, oyun, görüntü ve ses işlemlerinde sıklıkla kullanılmaktadır.

2.3.1. Genetik Algoritma

Biyolojik evrim sürecinden ilham alarak tasarlanmış olan bir arama ve optimizasyon algoritmasıdır. Genetik algoritma (GA), doğal seçim, çaprazlama, mutasyon ve popülasyon aşamalarından oluşmaktadır (Mirjalili, 2019). Bu aşamalarda popülasyondaki bireylerin evrimsel süreçlerini simüle etmektedir.



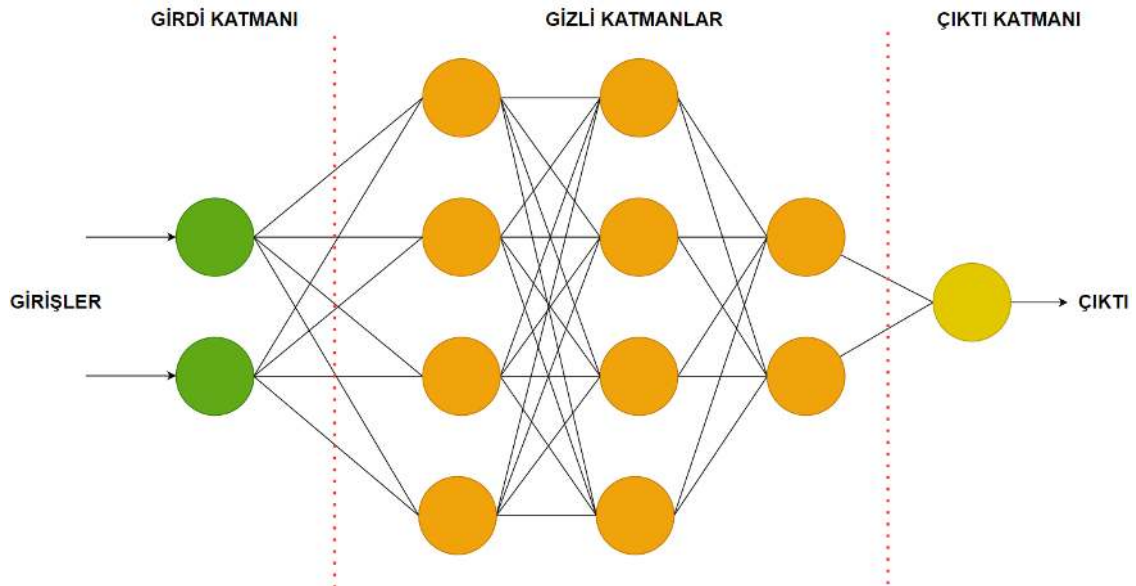
Şekil 2.8. Genetik algoritma akış diyagramı

GA'da öncelikle başlangıç popülasyonunun oluşturulması gereklidir. Bu popülasyon potansiyel çözümlerden meydana gelmektedir. Popülasyondaki bireylerin

başarısını belirleyebilmek için uygunluk fonksiyonunun değerlendirme yapması gereklidir. Bu fonksiyon, problemin çözülme oranını ölçmektedir. Uygunluk fonksiyonundan çıkarılan sonuçlara göre ebeveyn seçimi için seçim yapılması gerekmektedir. Seçilen ebeveynlerden yeni bireyler üretilmektedir (Mathew, 2012). Çaprazlama adı verilen bu aşamada mevcut ebeveynlerin genetik materyali birleştirme ve değiştirilme işlemleri uygulanmaktadır. Mutasyon aşamasında, yeni üretilen bireylerde rastgele değişiklikler yapılarak popülasyona çeşitlilik eklenmektedir. Böylelikle yeni çözümler keşfetmesi sağlanmaktadır. Popülasyonun yeni oluşturulan bireyler ve eski bireyler için güncellenmesi gerekmektedir. Uygunluk fonksiyonunda değerlendirme yapılır ve daha iyi çözümler ile popülasyon güncelleştirilmektedir. Algoritma belirli bir iterasyon veya uygunluk durumuna ulaştığında sonlanmaktadır. Şekil 2.8’de GA’nın akış diyagramı gösterilmekte olup algoritmanın adımları belirtilmiştir.

2.3.2. Yapay Sinir Ağları

İnsanların beyinde bulunan sinir ağlarını taklit ederek oluşturulan yapay zeka modeline Yapay Sinir Ağı (YSA) denir (Abraham, 2005). YSA ile bilgisayar sistemlerinin ve makinelerin tanıma ve öğrenme işlemleri ile problem çözmesi amaçlanmaktadır. Şekil 2.9’da YSA’nın yapısı gösterilmektedir.



Şekil 2.9. Yapay sinir ağının yapısı

YSA temel olarak girdi katmanı, gizli katman ve çıktı katmanı olarak 3 katmanda incelenmektedir. Girdi katmanına verilen giriş değerleri ve ağırlık değerleri birbiri ile

çarpılıp bias değeri eklenerek bir sonuç elde edilmektedir. Çıkan sonuç, aktivasyon fonksiyonuna verilerek bir sonraki nörona aktarılmaktadır. Böylelikle nöronlar mevcut bilgisini bir sonraki nöronlara geçirmektedir. Son katmana gelene kadar devam eden bu süreç, son katmanda çıkan değer üzerinde maliyet fonksiyonu çalıştırılarak geri besleme işlemi gerçekleştirilir. Bu işlemde türev işlemi yapılarak mevcut ağırlıkların güncellenmesi yapılmaktadır. Bu işlemler dışarıdan verilen tur sayısı kadar sürekli devam etmektedir. Gizli katmandaki nöron sayısı problemin karmaşıklığına göre değiştirilebilir (Hassoun, 1995). Fazla sayıda gizli katmanı olan YSA'lara Derin Sinir Ağları (DNN) denilmektedir. Tüm katmanlardaki nöronlar birbirine bağlı bir şekildedir (Zou vd., 2009). Ancak derin öğrenme uygulamalarında ezberden kaçınmak için belirli parametrenin altındaki nöronlar ağdan çıkarılarak algoritmanın farklı yolları deneyip keşfederek daha uygun çözüme gitmesi amaçlanmaktadır.

2.3.3. Evrimsel Sinir Ağı

Görsel verilerin analizi için özel olarak tasarlanmış bir yapay sinir ağı türüdür. Alınan bir giriş görüntüsünü, görüntüdeki içerisindeki çeşitli nesnelerin ve görünüşlerin birbiri arasındaki farkı ayırabilen derin öğrenme algoritmasına denir. Nesne tespiti, görüntü tanıma, görüntü sınıflandırma, doğal dil işleme ve ses tanıma gibi uygulamalarda sıklıkla kullanılmaktadır. Evrimsel Sinir Ağları (CNN) birden fazla ardışık eğitilebilir bölümden oluşmaktadır (İnik ve Ülker, 2017). Pikseller, kenar kombinasyonlarından meydana gelen motifleri, motifler bir araya gelerek nesne parçalarını oluşturmaktadır. Nesne parçalarının birleşerek bir araya gelmesiyle görüntüdeki nesneler oluşmaktadır (LeCun vd., 2015).

CNN mimarisinde giriş katmanında mimariye girecek veriler ham olarak verilmektedir. Modelin başarımı için bu katmanda ağa verilen görüntü boyutu önemlidir.

CNN mimarisi evrim katmanlarına sahiptir. Bu katmanlar CNN'nin temelini oluşturmaktadır. Giriş katmanından gelen girdi verisi üzerindeki belirli özelliklerin öğrenilmesi ve önemli bilgilerin vurgulanması amacıyla mimaride kullanılmaktadır. Bu katmanda filtreler 2x2, 3x3 veya 5x5 gibi farklı boyutlarda olabilmektedir. Filtreler, önceki katmanlardan gelen görüntü verilerine konvolüsyon işlemini uygulayarak çıktı verisi üretmektedir. Uygulanan işlemler sonucunda aktivasyon haritaları oluşmaktadır (Bouvrie, 2006). Bu haritalar bulunan her bir filtreye özgün niteliklerin keşfedildiği bölgelere denir. Giriş görüntüleri üzerinde belirli bir adım (stride) değeri ile kaydırılarak dolaşma sağlanmaktadır. Dolaşma işleminde görüntü matrisi sınır değerine gelindiğinde

bir basamak aşağıdan devam edilmektedir. Filtrelerde bulunan katsayılar renk kanallarındaki değerler ile çarpılıp toplama işlemi yapılmaktadır. Her bir kanal üzerinde bahsedilen işlem yapıldıktan sonra her bir kanal değerinin toplamı ile Aktivasyon haritası oluşturulmaktadır (İnik ve Ülker, 2017). Aktivasyon haritasında bulunan tüm değerler giriş boyutu ve çıkış boyutu arasındaki yoğunluk oranının sağlanması için normalize edilmektedir.

Düzleştirilmiş Doğrusal Birim Katmanında (ReLU), giriş değerlerindeki negatif değerleri sıfıra çekme işlemi yapılmaktadır. Böylelikle sinir ağına doğrusallık kazandırmaktadır. Formül 2.1'deki tanımlanmaktadır.

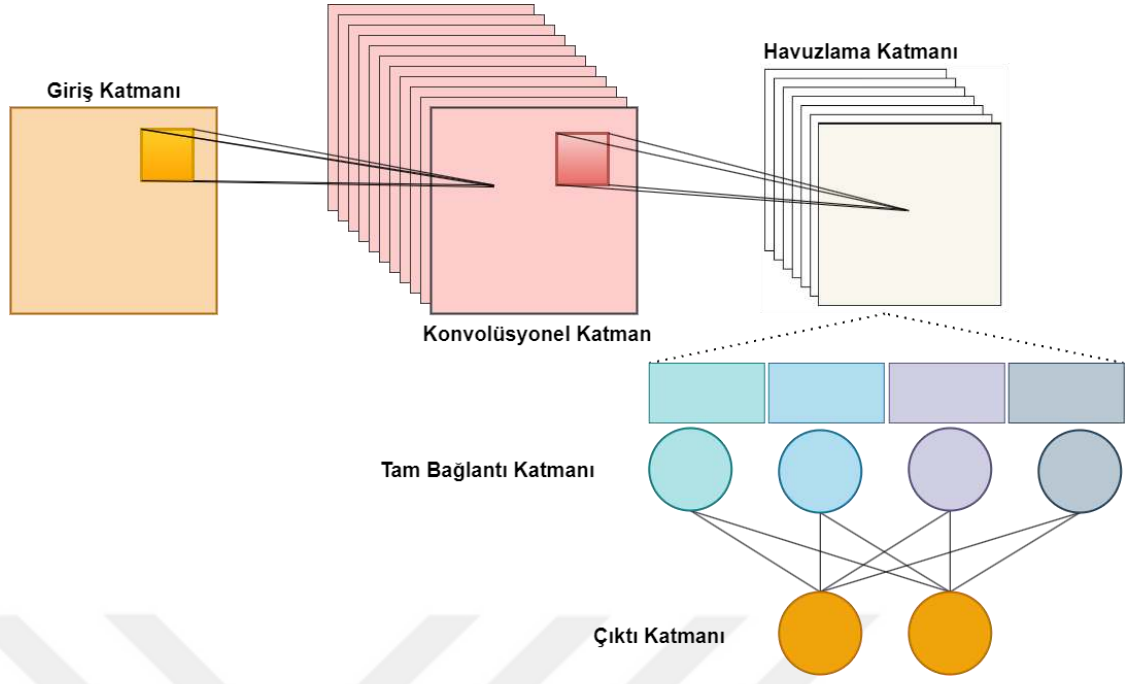
$$f(x) = \{0 \text{ eğer } x < 0, \text{ eğer } x \geq 0\} \quad (2.1)$$

Havuzlama katmanlarında evrişim katmanlarından elde edilen özellik haritalarının küçültülmesi işlemi yapılmaktadır. Bu işlem, önemli bilgileri vurgulamak içinde kullanılmaktadır. Genellikle ortalama ve maksimum havuzlama yöntemleri kullanılmaktadır.

CNN mimarisinde havuzlama ve ReLu katmanlarından sonra tam bağlantı katmanı gelmektedir. Tam bağlantı katmanı kendinden önceki katmandaki tüm değerlere bağlıdır.

CNN'de eğitim işlemi sırasında bazen nöral ağ ezber yapabilmektedir. Bu sorunu çözebilmek için Dropout katmanı derin öğrenme modellerinde ezberi azaltmak amacıyla sıklıkla kullanılan bir regularizasyon tekniğidir (Albawi vd., 2017).

CNN'de bulunan tam bağlantılı katmandan sonra Sınıflandırma Katmanı gelmektedir. Bu katmanda ağa verilen sınıf sayısı kadar çıkış değeri bulunmaktadır. Bu katmanda farklı sınıflandırıcılarda kullanılabilir. Genellikle bu katmanda diğer aktivasyon fonksiyonlarından daha başarılı olduğu için Softmax aktivasyon fonksiyonu tercih edilmektedir. CNN mimarisinin yapısı Şekil 2.10'da verilmektedir.



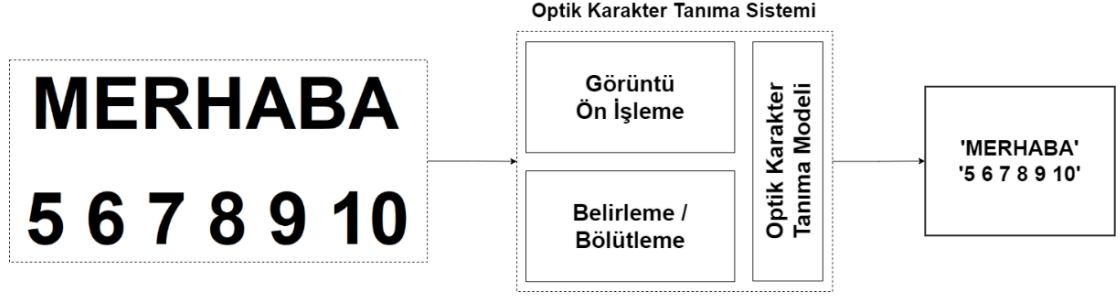
Şekil 2.10. Evrişimsel sinir ağının yapısı

2.3.4. Optik Karakter Tanıma

Görüntüde yazılı bir metni bilgisayar ortamına metin olarak tanıtılmasını sağlayan yeteneğe denilmektedir. OKT'nin temel amacı basılı metinleri ve dökümanları elektronik ortama taşımaktır (Koyun ve Afşin, 2017). Veri kümeleri oluşturulmasında, kimlik ve doküman tanımlamasında, yapay zeka araçlarında ve finans sektöründe sıklıkla kullanılmaktadır.

OKT'nin çalışma süreci genellikle 4 adımda gösterilmektedir. Görüntü edinme, görüntü önileme, belirleme veya bölütleme süreci, OKT algoritmasının uygulanmasıdır (Saray vd., 2017). Elde edilen sonuçlar ile gerçek sonuçlar karşılaştırılıp, yöntemin doğruluk oranı bulunmaktadır. Görüntü işleme yöntemlerinin ve derin öğrenme algoritmalarının sürekli olarak gelişmesiyle çok daha yüksek başarıma sahip modeller elde edilmektedir (Shaffi ve Hajamohideen, 2021).

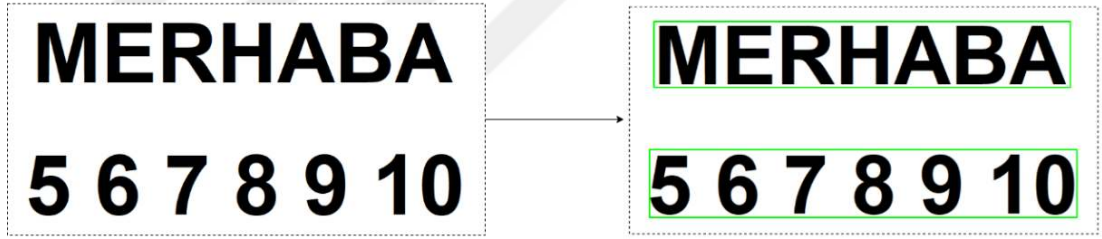
OKT sistemlerinde farklı yaklaşımlar ve mimariler mevcuttur. Karakter sınıflandırılması, nesne tespiti ve temel görüntü işleme teknikleriyle yapılabilmektedir. Özellikle Transformer, Tekrarlayan Sinir Ağı (RNN), Bağlantısal Zamansal Sınıflandırma (CTC) ve Uzun Kısa Süreli Bellek (LSTM) tabanlı modeller yüksek doğruluk oranı vermektedir. Şekil 2.11'de örnek bir OKT uygulaması gösterilmektedir.



Şekil 2.11. Optik karakter tanıma uygulaması

2.3.5. Metin Tespiti

Video veya görüntü içerisinde bulunan bir metni tespit ve algılamasını sağlayan bir bilgisayar görü uygulamasıdır. Genellikle OKT'dan önce uygulanan bir yöntemdir (Long vd., 2021). Birçok metin tespiti algoritması bulunmaktadır. Bunlardan bazıları; Verimli ve Doğru Sahne Metni Tespiti (EAST), Metin Tespiti için Karakter-Bölge Farkındalığı (CRAFT), Bölge Temelli Evrişimli Sinir Ağı Hızlı R-CNN (Faster R-CNN), Tek Atışlı Çoklu Kutu Dedektörü (SSD), Evrişimli Tekrarlayan Sinir Ağı (CRNN). Şekil 2.12'de örnek bir metin tespiti gösterilmektedir.



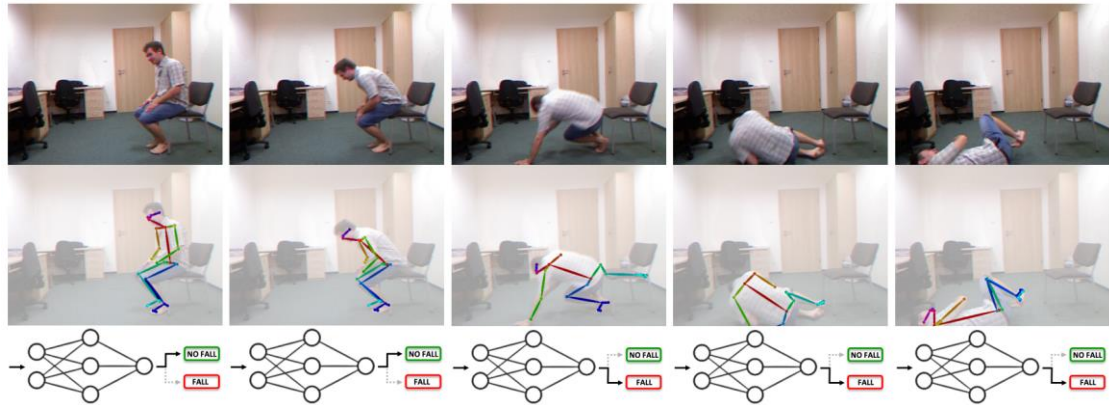
Şekil 2.12. Metin tespiti uygulaması

2.3.6. Poz Tahmini

Poz tahmini, nesnelerin veya insanların yaptıkları belirli değişiklikleri izleyerek anlama ve analiz etmek amacıyla kullanılan bir yöntemdir. Bu yöntem hareket analizi ve bilgisayarda görü uygulamalarında sıklıkla kullanılmaktadır (Murphy-Chutorian, ve Trivedi, 2008). Birçok uygulama alanı bulunmaktadır. Robotik, otonom araçlar, spor bilimlerinde, bilgisayarda görü kapsamında ve kamera tabanlı sistemlerde görüş sistemlerinde sıklıkla kullanılmaktadır.

İnsanlarda poz tahmini, vücutta bulunan eklem bölgelerini tahmin etmek amacıyla kullanılmaktadır. Bu tahmin modellini geliştirmek için derin öğrenme modellerinden yararlanılmaktadır. OpenPose, PoseNet, AlphaPose, HRNet, EfficientPose ve DeepPose

en yaygın kullanılan poz tahminleme algoritmalarıdır (Serpa vd., 2020). Şekil 2.13’de örnek bir insanda poz tahmini gösterilmektedir.



Şekil 2.13. Örnek bir poz tahmini uygulaması (Serpa vd., 2020)

2.4. Veritabanı

Veritabanı, bilgilerin belirli bir düzenle depolanabildiği, yönetilebildiği ve hızlı bir şekilde verilerin işlenebildiği bir depolama sistemidir. Büyük miktardaki veriler ile çalışabilen etkili bir bilgi yönetim aracı olarak tanımlanabilir. Veritabanları çok sayıda tablo ile çalışabilir ve tablolar arasında ilişkiler kurabilmektedir. Her tablo, satır ve sütunlardan oluşur ve verileri depolar. Anahtarlar kullanılarak tablolar arasındaki ilişkiler kurulmaktadır.

Veritabanı birkaç temel bileşenden oluşmaktadır. Tablolar, satırlar, sütunlar, anahtarlar, sorgular ve indeksler temel bileşenleridir (Köseoğlu, 2008). Veritabanları kuruluşlar, şirketler ve işletmeler tarafından sıklıkla kullanılmaktadır.

İlişkisel veritabanları, anahtar-değer veritabanları, sıralı veritabanları, NoSQL veritabanları, graf veritabanları ve zaman serisi veritabanları şeklinde farklı yapıda veritabanı türleri bulunmaktadır. Sektördeki farklı ihtiyaçlar ve senaryolara göre veritabanı seçimi yapılmaktadır.

3. MATERYAL VE YÖNTEM

3.1. Materyal

3.1.1. Veri Kümesi

Çalışmamızda kart tespiti ve perspektif düzenleme aşamaları için 2 ayrı veri kümesi etiketlenerek hazırlanmıştır. Öncelikle kart tespiti için bir Örnek Segmentasyonu modeli eğitilmiştir. Bu model eğitiminde 1224 görüntü verisi etiketlenmiş olup 1021 görüntü eğitim, 102 görüntü validasyon 101 görüntü ise test olarak ayrılmıştır. Eğitim kümesi üzerinde veri arttırımı uygulanıp eğitim kümesindeki veri sayısı 5104'e çıkarılmıştır.

Türkiye Cumhuriyeti Kimlik kartının üzerindeki Türk Bayrağı'nın tespiti için 389 görüntü etiketlenmiştir. Bu görüntülerden 311 görüntü eğitim, 39 görüntü validasyon ve 39 görüntü test olarak ayrılmıştır. 311 görüntüye sahip eğitim kümesi üzerinde veri arttırımı uygulanıp görüntü sayısı 933'e çıkarılmıştır.

Yüz tespiti için WIDER FACE hazır veri kümesi kullanılmıştır. Bu veri kümesi üzerinde 32,203 görüntü bulunmakta olup, 393,703 yüz görüntüsü etiketlenmiştir (Deng vd., 2020).

TrOCR, OKT mimarisinde birinci aşama ön eğitimde yaklaşık 684 milyon metin satırı, ikinci aşama ön eğitimde ise 16 milyon metin resmi ile birlikte model eğitilmiştir (Liv d., 2023). Çalışmamızda bu model önceden eğitilmiş ağırlıklar ile kullanılmıştır.

Poz tahminleme kısmında CMU Panoptic ve 300W-LP veri kümeleri kullanarak eğitilen bir model kullanılmıştır (Hempel vd., 2023). Çalışmamızda poz tahminleme modelimizde önceden eğitilmiş ağırlıklar içermektedir.

Kimlik kartında bulunan yüz ile poz tahminleme aşamasındaki tanımlanan yüzü karşılaştırma işlemi için RetinaFace + ArcFace mimarisi kullanarak LFW, CFP-PF ve AgeDB-30 veri kümelerinde testi gerçekleştirilmiş önceden eğitilmiş ağırlıklar ile kullanılmıştır (Deng vd., 2020).

3.1.2. Performans Metrikleri

Yapay zeka modellerini değerlendirme aşamasında algoritmaların genel performansını belirlemek için çeşitli metrikler kullanılmaktadır. Problemin türüne göre farklılık gösteren bu metrikler yapay zeka modelinin veriyi öğrenme durumu hakkında çeşitli sonuçlar vermektedir. Eğitilen bir yapay zeka modelinin başarılı veya başarısız şeklinde yorumlanabilmesi için bu metrikler faydalanılmaktadır.

Bir sınıflandırma modelinin performansını analiz edilmesini ve değerlendirmesini sağlayan ölçülere sınıflandırma metrikleri denilmektedir. Oluşturulan bir sınıflandırma modelinin test aşamasında verilerin gerçek sınıf değeri ve tahmin edilen sınıf değeri bulunmaktadır. Bu değerler üzerinde karmaşıklık matrisi ile model başarımının değerlendirilmesi yapılmaktadır.

		Tahmini Değerler	
		Pozitif	Negatif
Gerçek Değerler	Pozitif	Gerçek Pozitif	Yanlış Negatif
	Negatif	Yanlış Pozitif	Gerçek Negatif

Şekil 3.1. Karmaşıklık matrisi

Şekil 3.1’de karmaşıklık matrisinde gerçek pozitif (GP), sınıflandırma modelinin gerçekte pozitif olarak tahmin ettiği doğru örnekleri temsil etmektedir. Yanlış negatif (YN), sınıflandırma modelinin gerçekte pozitif olan ancak negatif olarak tahmin ettiği yanlış örnekleri temsil etmektedir. Yanlış pozitif (YP), sınıflandırma modelinin gerçekte negatif ancak pozitif olarak tahmin ettiği yanlış örnekleri temsil etmektedir. Gerçek negatif (GN) ise sınıflandırma modelinin gerçekte negatif olarak tahmin ettiği doğru örnekleri temsil etmektedir.

Karmaşıklık matrisinden elde edilen GP, YN, YP ve GN değerleri ile sınıflandırma modelimiz hakkında doğruluk, hassasiyet, duyarlılık ve f1 puanı hakkında çıkarımlarda bulunmaktadır (Krstinić vd., 2020).

Doğruluk oranı, modelimiz tahmin ettiği doğru sonuçları temsil etmektedir. GP ve GN sayılarının, karmaşıklık matrisindeki tüm veri sayısına oranı modelin doğruluğunu vermektedir. Formül 3.1’de modelimizin doğruluk oranını hesaplayan fonksiyon verilmiştir.

$$Doğruluk = \frac{GP + GN}{GP + YN + YP + GN} \quad (3.1)$$

Hassasiyet değeri, gerçekte pozitif olan değerlerin, gerçek pozitif ve yanlış pozitif olan değerlere bölünmesiyle elde edilmektedir. Formül 3.2’de modelimizin hassasiyet oranını hesaplayan fonksiyon verilmiştir.

$$Hassasiyet = \frac{GP}{GP + YP} \quad (3.2)$$

Duyarlılık değeri, gerçekte pozitif olan değerlerin, gerçek pozitif ve yanlış negatif olan değerlere bölünmesiyle elde edilmektedir. Formül 3.3’de modelimizin duyarlılık oranını hesaplayan fonksiyon verilmiştir.

$$Duyarlılık = \frac{GP}{GP + YN} \quad (3.3)$$

Hassasiyet ve duyarlılık değerinin harmonik ortalaması ile F1 puanı bulunmaktadır (Krstinić vd., 2020). Formül 3.4’de modelimizin F1 puanını hesaplayan fonksiyon verilmiştir.

$$F1 Puanı = \frac{Hassasiyet \times Duyarlılık}{Hassasiyet + Duyarlılık} \quad (3.4)$$

Doğruluk, hassasiyet, duyarlılık ve f1 puanı değerleri 0 ile 1 arasında değer almakta olup 1’e yaklaşma durumunda iyi bir performans gösterdiği anlamına gelmektedir.

Eğri Altında Kalan Alan (AUC) ve Alıcı İşletim Karakteristiği (ROC) sınıflandırma modellerinde kullanılan bir performans değerlendirme metriğidir. ROC eğrisi özgüllük ve duyarlılığı görselleştirmektedir. Bu eğri ile farklı eşik değerlerindeki özgüllük ve duyarlılık ikililerini göstermektedir. X ekseninde Yanlış Pozitif Oranı (FPR), y ekseninde ise Doğru Pozitif Oranı (TPR) bulunmaktadır (Huang ve Ling, 2005). Formül 3.5’de TPR ve Formül 3.6’da FPR gösterilmektedir.

$$TPR = \frac{GP}{GP + YN} \quad (3.5)$$

$$FPR = \frac{YP}{YP + GN} \quad (3.6)$$

ROC eğrisi farklı eşik değerlerinin altında kalan kısmın model performansını değerlendirmektedir. Çizilen ROC eğrisinin sol üst köşeye yaklaşması modelin performansının iyi olduğunun bir göstergesidir (Marzban, 2004).

AUC 0 ile 1 arasında değer alan, ROC eğrisinin altında bulunan alanı hesaplayarak model performansını gösteren bir metriktir.

- AUC = 1 durumunda sınıflandırma modeli yüksek başarımlı,
- AUC = 0,5 durumunda sınıflandırma modeli rastgele tahminle aynı,
- AUC < 0,5 durumunda rastgele tahmininde altında

bir sonuç göstermektedir.

Regresyon modellerinin tahmin performansını hata oranı üzerinden değerlendiren ölçütlere regresyon metrikleri denir. Ortalama Mutlak Hata (MAE), Ortalama Kare Hata (MSE) ve R-Kare en çok kullanılan regresyon metrikleridir (Liu vd., 2014).

MAE, regresyon modelinin gerçek değerler ile tahmini değerler arasındaki mutlak farklarının ortalamasına denilmektedir. Formül 3.7’de MAE değerini hesaplayan fonksiyon gösterilmiştir. Formülde gösterilen N veri sayısı, x_i gerçek çıkış değeri ve $\hat{x}$ ise tahmin edilen çıkış değeridir.

$$MAE = \frac{1}{N} \sum_{i=1}^N (|x_i - \hat{x}|) \quad (3.7)$$

MSE, regresyon modelinin gerçek değerler ile tahmini değerler arasındaki kare farklarının ortalamasına denilmektedir (Chicco vd., 2021). Formül 3.8’de MSE değerini hesaplayan fonksiyon gösterilmiştir.

$$MSE = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x})^2 \quad (3.8)$$

R-Kare, regresyon modelinin bağımsız değişkenler tarafından bağımlı değişkenler varyans değerinin açıklama ölçüsüne denir (Chicco vd., 2021). Formül 3.9’da R-Kare

değerini hesaplayan fonksiyon gösterilmiştir. Fonksiyonda bulunan $\hat{x}$ tahmin edilen x değerlerini, $\bar{x}$ ise x değerlerinin ortalamasını temsil etmektedir.

$$R - Kare = 1 - \frac{\sum (x_i - \hat{x})^2}{\sum (x_i - \bar{x})^2} \quad (3.9)$$

Nesne algılama ve segmentasyon modellerinin performansını analiz edilip değerlendirmek için nesne algılama ve segmentasyon metrikleri kullanılmaktadır. Ortalama Hassasiyet Oranı (mAP) ve Kesişim ve Birleşim Oranı (IoU) sıklıkla kullanılan nesne algılama ve segmentasyon metriklerindendir (Padilla vd., 2020).

mAP, hesaplanan her nesne sınıfları için hassasiyet değerlerinin ortalamasıdır. Formül 3.10'da mAP değerini hesaplayan fonksiyon gösterilmiştir. Fonksiyonda bulunan k değeri nesne sınıflarının sayısını temsil etmektedir.

$$mAP = \frac{1}{k} \sum_i^k AP_i \quad (3.10)$$

IoU, gerçek bölge ile modelin tahmin ettiği bölgenin birbiriyle örtüşme değeridir (Padilla vd., 2020). Formül 3.11'de IoU değerini hesaplayan fonksiyon gösterilmiştir.

$$IoU = \frac{GP}{GP + YP + YN} \quad (3.11)$$

3.2. Yöntem

Bu çalışmada Türkiye Cumhuriyet Kimlik Kartı görüntüleri üzerinden kart bilgilerin optik karakter tanıma yapılması amaçlanmıştır. Ayrıca kimlik kartında bulunan yüz görüntüsü ile kartı tanıtan kişinin aynı olduğunu belirleyebilmek için yüz benzerlik tanımlaması yapılmıştır. Yüzü tanıtan kişinin sahte bir görüntü olmadığını kanıtlamak amacıyla canlılık analizi kapsamında poz tahminlemesi yapılmıştır.

Öncelikle kimlik kartı görüntülerinin iyileştirilmesi için CURL derin öğrenme mimarisi kullanılarak görüntü iyileştirilmesi yapılmıştır. İyileştirilen görüntüler üzerinde kimlik kartının tespiti için YOLOv8 Örnek Segmentasyonu kullanılarak kimlik kartının bulunduğu noktalar dizini elde edilmiştir. Bu noktalar dizinin içerisinde kimlik kartının köşelerini bulabilmek için dışbükey noktaları üzerinden maksimum alanı tespit edebilmek için bir arama algoritması geliştirilmiştir. Bu algoritma örnek segmentasyonundan gelen değerler üzerinde maksimum alanı veren 4 köşe noktasını

bulmayı amaçlamaktadır. Algoritmadan gelen noktaları daha verimli bulabilmek için GA ile optimizasyon yapılmıştır. Bulunan noktaların kartın hangi köşeleriyle eşleştiğini bulabilmek için 2 ayrı nesne tanıma uygulaması yapılmıştır.

Türk Bayrağını tespit eden YOLOv8 nesne tanıma modelini kullanarak kimlik kartı üzerinde bulunan Türk Bayrağının orta noktası belirlenmiştir. 4 köşe noktası ile belirlenen nokta arasında öklid uzaklığına bakılarak en yakın nokta kartın sağ üst, 2. en yakın nokta ise kartın sağ alt köşesi olarak tespit edilmiştir. Kimlik kartındaki yüz görüntüsünün tespiti için önceden eğitilmiş bir RetinaNet modeli kullanılmıştır. Bu model ile yüz tespiti yapıp, yüzün ortasında bulunan burnu merkezi nokta olarak belirlenmiştir. Bu nokta ile 4 köşe arasındaki öklid uzaklığına bakılıp en yakın mesafedeki nokta sol alt, 2. en yakın mesafedeki nokta ise sol üst olarak kimlik kartında belirlenmiştir.

Kimlik kartındaki köşe noktalarının belirleyip kimlik kartının hangi açı ile yerleştiğini ve köşelerin hangi noktalarda bulunduğu anlaşılmıştır. Bu noktalar ile perspektif düzenlemesi yapılmış olup kimlik kartı 1700x1100 boyutlarına getirilerek ortalanmıştır.

Ortalanmış görüntü üzerinde ROI noktalarının tespiti için 2 ayrı yaklaşım uygulanmıştır. Select ROI sınıfını kullanarak statik olarak belirlenen noktalara bakılmış olup, kimlik kartının her köşesi bulunmayan görüntülerde ufak derecede kayma tespit edilmiştir. Bu sebepten dolayı CRAFT derin öğrenme modelini kullanarak perspektif düzenlemesinden sonra metinlerin tespiti, dinamik bir şekilde yapılmış olup herhangi bir kayma söz konusu olmamaktadır.

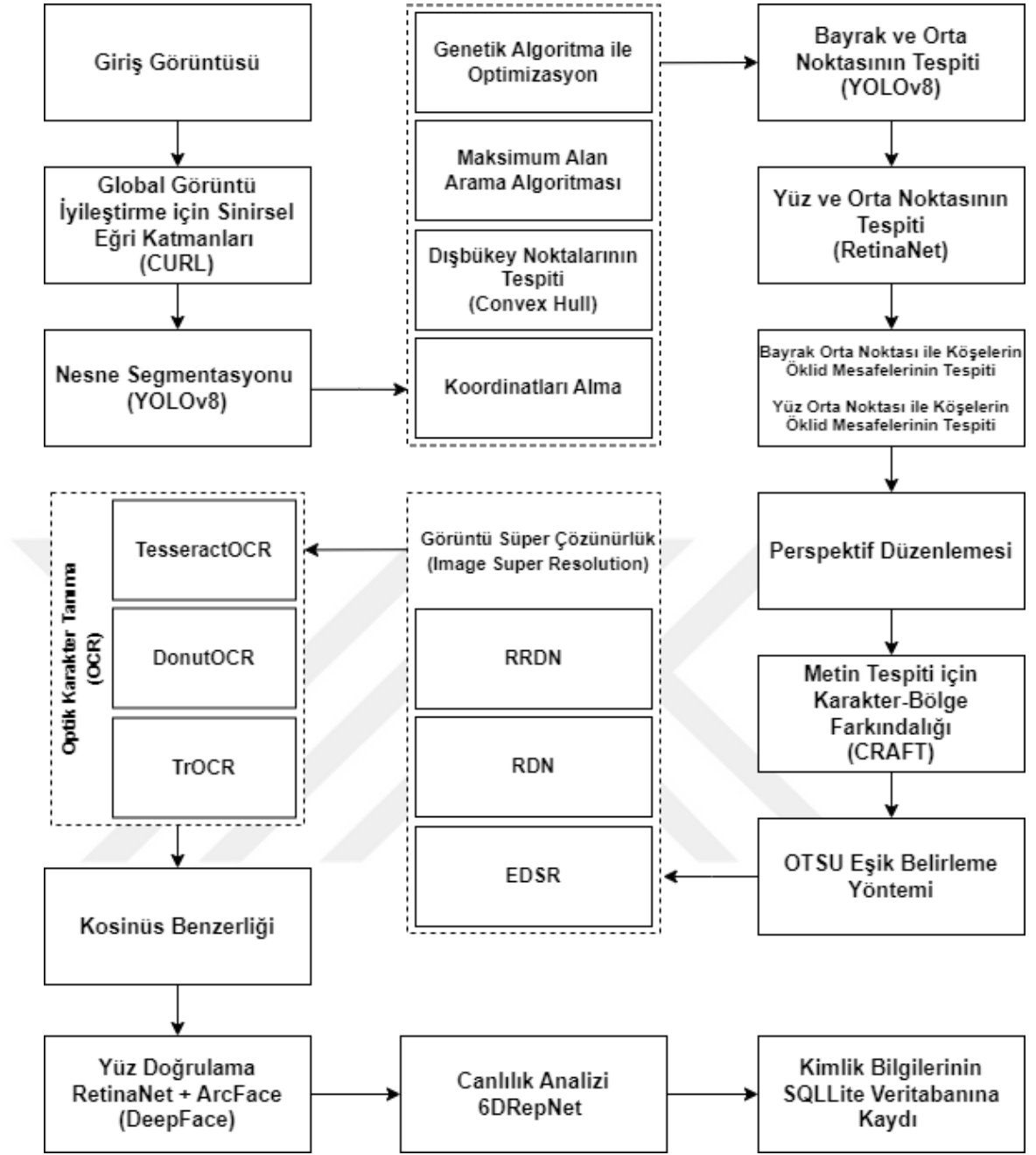
CRAFT metin dedektöründen gelen metin koordinatları Otsu ve Sauvola yöntemleri ile optimal eşik değeri tespiti yapılmıştır. Alınan sonuçlara göre Otsu yönteminin Sauvola yöntemine göre kimlik verilerinde daha başarılı olduğu anlaşılmıştır. Otsu yöntemi uygulanan ROI noktaları üzerinde Görüntü Süper Çözünürlük (ISR) derin öğrenme mimarilerinden olan RRDN, RDN, EDSR ve SRGAN modelleri, ROI noktalarındaki görüntü kalitelerini yükseltmeye çalışılmıştır.

Görüntü filtreleme ve çözünürlük yükseltme işlemlerinden sonra TesseractOCR, DonutOCR ve TrOCR modelleri elde edilen görüntüler üzerinde çalıştırılarak kimlik kartı üzerinden OKT uygulanmıştır. Diğer modellere göre Transformer temelli TrOCR'ın yüksek başarısıyla kimlik kartından metin okuma işlemi tamamlanmıştır. OKT algoritmasından gelen çıkarılacak bilgilerin dışındaki diğer bilgiler metin parçacıklarına dönüştürülüp Kosinüs Benzerliği ile elimine edilecektir.

Kimlik kartındaki bilgileri okuma işleminden sonra kimlik kartını sisteme yükleyen kişinin gerçek kişi olduğunu tespit edebilmek ve çalıntı kart olabilme ihtimaline karşın yüz tanıma işlemi uygulanacaktır. Yüz tanıma işleminde DeepFace çerçevesinde RetinaNet + Arcface modelini kullanarak kimlik kartındaki yüz görüntüsü ile mevcut görüntü arasında benzerlik oranına bakılarak doğrulama işlemi yapılmaktadır. Doğrulama işleminden sonra karşıdaki kişiye çeşitli komutlar vererek canlılık analizi yapılmıştır. Karşıdaki kişinin kafasını yukarı, aşağı, sağ, sol yönünde ve yakınlaştırma gibi komutlar kullanarak hareket ettirilmesi istenilmiştir. Bireyin kamerada rastgele komutlar ile hareketler yapılmasını sağlanıp karşıdaki kişinin canlı olmasının teyit edilmesi gereklidir. Bu işlemlerde 6DRepNet modeli kullanılmıştır.

Yapılan işlemler sonucunda bir kullanıcının karşıdan herhangi bir açıyla verilen bir kimliğin tanımlanarak bilgi çıkarımı yapılmıştır. Kimlik kartı sahibi tarafından sisteme yüklendiğini teyit edebilmek için yüz tanıma ve benzerlik algoritmaları kullanılmış olup herhangi bir sahtekarlığa karşın canlılık analizi yapılarak 2 aşamalı doğrulama yapılmıştır.

Tüm görüntü işleme yöntemleri ve derin öğrenme modellerin eğitimi ve testleri tamamlanıp mikroservis mimarisine entegre bir tasarım sunulmuştur. Önerilen sistemin akış diyagramı Şekil 3.2’de gösterilmektedir.

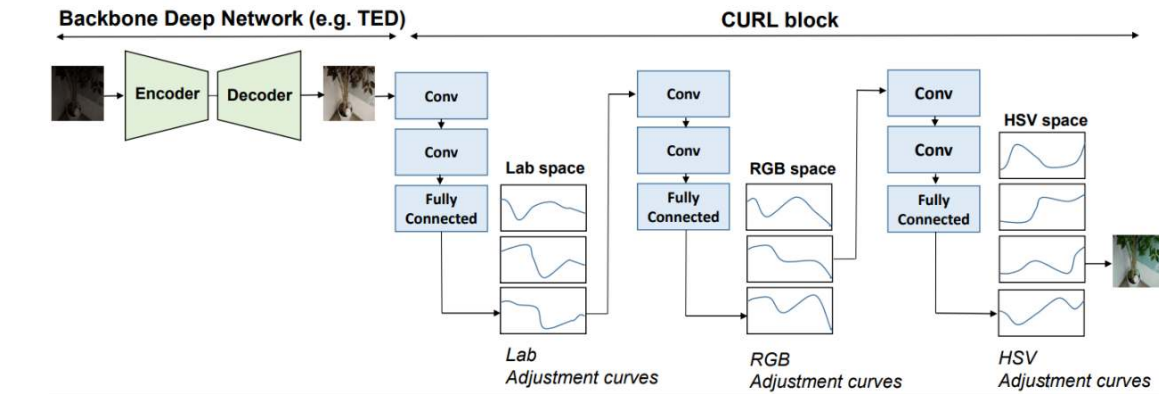


Şekil 3.2. Önerilen sistemin akış diyagramı

3.2.1. Global Görüntü İyileştirme için Sinirsel Eğri Katmanları (CURL)

CURL mimarisi ile doygunluk, parlaklık ve renk gibi görüntü özelliklerini yorumlanabilir görüntü geliştirme eğrileriyle ayarlama yapılabilmektedir. HSV, CIELab ve RGB renk uzaylarında eğitilmiştir. Eğriler tamamıyla türevlenebilir olmaktadır. Görüntü oluşturma ve görüntü iyileştirme işlemleri için uçtan uca eğitilmiştir. CURL'ün etkinliğini gösterebilmek için CURL bloğunu bir kodlayıcı-çözücü yapısıyla birleştirilmiştir (Moran vd., 2021). CURL bilgisayar görüşü problemleri için RGB-RGB

veya RAW-RGB dönüşümleri için uçtan uca eğitilmiştir. CURL bloğunun yapısında her renk uzayı için konvolüsyonel işlemler ve tam bağlantı katmanı bulunmaktadır. Şekil 3.3’de CURL mimarisi gösterilmektedir.



Şekil 3.3. CURL mimarisi

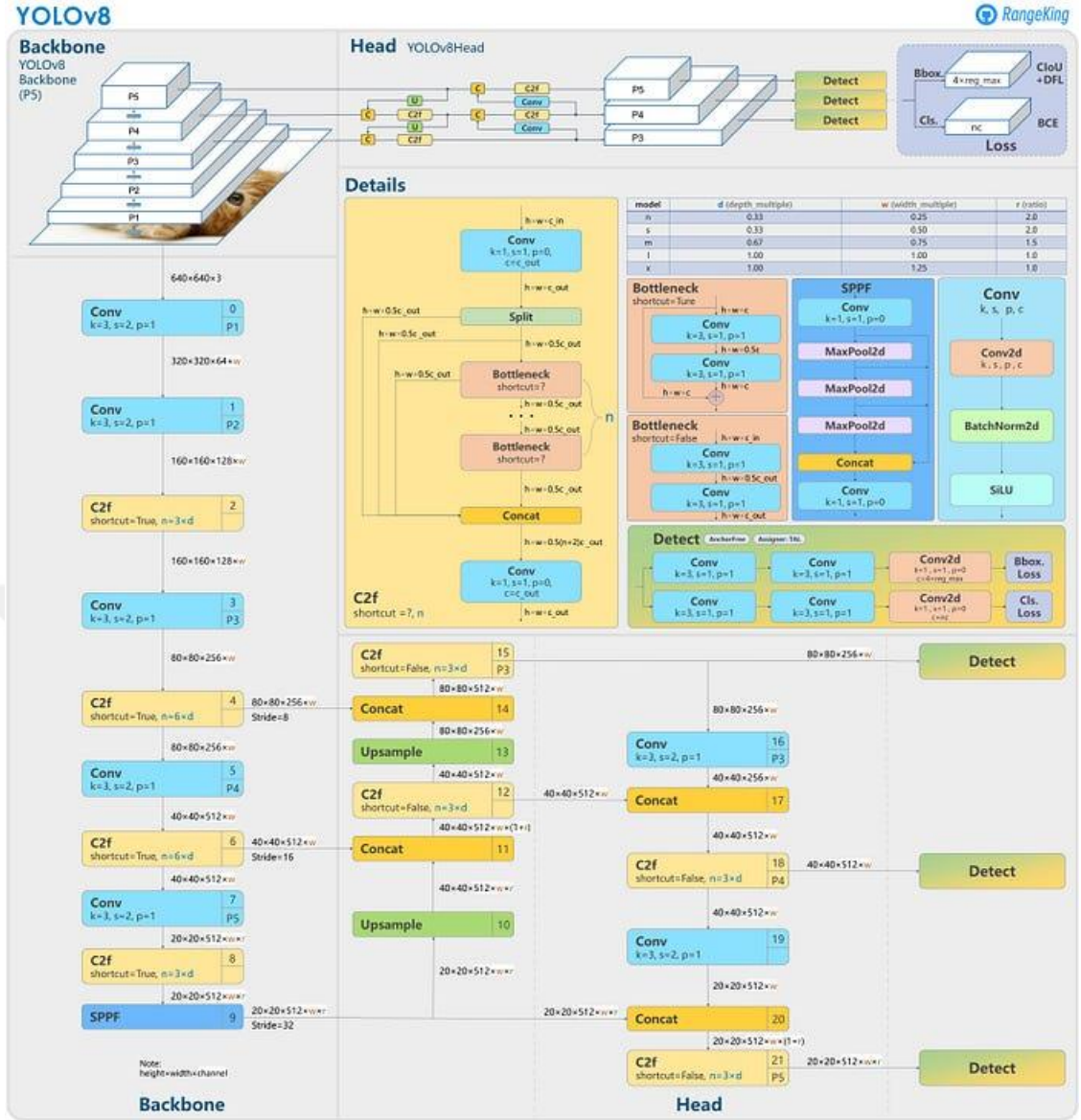
3.2.2. YOLOv8

YOLOv8 Ocak 2023 yılında Ultralytics tarafından piyasaya sürülen açık kaynaklı güncel bir nesne tespit modelidir. Bu model temel olarak Head ve Backbone denilen 2 parçadan oluşmaktadır. Backbone kısmında özellik çıkarımı yapıp, özellik piramitleri oluşturmaktadır (Tamang vd., 2023). Head kısmında ise nesne tanıma ve nesne skorlarıyla birlikte sınırlama kutularının tanımını yapmaktadır (Terven, J. ve Cordova-Esparza, 2023).

Modelin doğruluğunu arttırmak için model eğitiminde veri arttırımı yapılmaktadır. Mozaik arttırma 4 adet eğitim görüntü verisinin bir araya gelerek yeni bir görüntü oluşturma tekniğine denir. Bu teknik ile modelin daha etkili bir performans göstermesine neden olmaktadır (Hao ve Zhili, 2020). Mozaik arttırma tekniği eğitim aşamasında sürekli olarak kullanımından kaynaklı modelin hassasiyeti azalabilmektedir. Bu durumdan dolayı son 10 tur boyunca devre dışı bırakılmaktadır.

YOLOv8 modelin esnekliği ve performansını önemli bir ölçüde arttırmak için tasarlanmış hızlı ve esnek bir modeldir. Daha önceki versiyonlarına göre daha yüksek bir doğruluk oranı vermektedir. Bu model ile görüntü segmentasyonu, nesne tespiti ve görüntü sınıflandırma işlemleri yapılabilmektedir. Şekil 3.4’de YOLOv8 mimarisi gösterilmektedir.

Çalışmamızda kimlik kartının görüntü segmentasyonu ile tespitinde ve Türk Bayrağının tespitinde YOLOv8 modeli kullanılmaktadır.



Şekil 3.4. YOLOv8 mimarisi (RangeKing, 2023)

3.2.3. Dışbükey Gövde Maksimum Alan Arama Algoritması ve GA Optimizasyonu

Dışbükey gövde (Convex Hull), bir nokta kümesinden en küçük çokgeni bulmasını sağlayan bir geometrik kavramdır. Herhangi bir bölgenin rastgele 2 noktasını birleştiren çizgi bu bölgenin içerisindeyse şeklin dışbükey olduğunu göstermektedir. Nokta kümesinden bir çokgenin dışbükey gövdesini bulabilmek için farklı algoritmalar bulunmaktadır. Literatürde genellikle Graham Scan, QuickHull, Jarvis March ve Chan algoritması kullanılmaktadır (Gomes, 2016).

Bu çalışmada scipy kütüphanesindeki ConvexHull metodu ile QuickHull algoritması kullanılmıştır. Bu algoritmanın ortalama zaman karmaşıklığı $O(n \log n)$, en kötü durumdaki karmaşıklığı ise $O(n^2)$ 'dir. Bu algoritma bölme ve ele geçirme stratejisini

kullanmaktadır. Algoritma, noktalar arasındaki en uzak mesafeyi bularak dışbükey gövdeyi 2 alt kümede böler ve 2 küme üzerinde de bu işlem devam etmektedir (Barber vd., 1996).

Optimizasyon bir sistemi en ideal duruma getirme sürecine denilmektedir. Bu süreç bir amaç fonksiyonu istenilen hedefe minimize veya maksimize edilmesini içermektedir (Emel ve Taşkın 2002). Çalışmamızda noktalar üzerine QuickHull algoritması uygulandıktan sonra elde edilen dışbükey geometrik şeklin noktaları bir diziye atanmıştır. Atanan noktalar üzerinde 4 nokta ile maksimum alanı verecek bir arama algoritması yazılmıştır. Yazılan algoritma GA ile optimize edilerek çalışma zamanı hızlandırılmıştır. Şekil 3.5’de algoritmanın temel adımları gösterilmektedir.

1. Verilen noktaları içeren bir NumPy array oluşturun.
2. Bu noktaları kullanarak bir Convex Hull oluşturun.
3. Convex Hull'ın köşe noktalarını bulun.
4. Genetik algoritma için DEAP kütüphanesi yapılandırın.
5. Amaç fonksiyonunu tanımlayın:
 - a. Genetik algoritma Convex Hull'ın içerisindeki maksimum alan hesabını optimize etmeye çalışsın.
 - b. Amaç fonksiyonu, Convex Hull'un içindeki alanı hesaplasın.
6. Genetik algoritma için popülasyon, çaprazlama, mutasyon ve seçim operatörlerini tanımlayın.
7. Belirli sayıda nesne ve parametre ile genetik algoritmayı çalıştırın.
8. En iyi çözümü seçin.
9. En iyi çözümün köşe noktalarını bulun.
10. En iyi çözümün köşe noktalarını ekrana bastırın.
11. Noktaları, Convex Hull'ı ve en iyi çözümü grafik olarak çizin.

Şekil 3.5. Maksimum alan arama algoritması GA ile optimizasyonu

3.2.4. CloDSA

CloDSA, açık kaynaklı bir görüntü çoğaltma aracıdır. Bu araç ile nesne algılama, nesne sınıflandırması, nesne tespiti ve nesne bölümlendirme işlemlerinden görüntü artırımı yapabilmektedir. Video dosyaları, görüntüler veya z-yığınlarında da uygulanabilmektedir. (Casado-García vd., 2019).

Çalışmamızda CloDSA açık kaynaklı görüntü çoğaltma aracını kullanarak veri çoğaltma işlemi yapılmıştır. Bu veriler, görüntü segmentasyonu ve nesne tanıma derin öğrenme modellerinde kullanılmış olup modellerdeki başarımların oranının önemli bir ölçüde

artmasına neden olmuştur. Çizelge 3.1’de CLoDSA kullanarak görüntü arttırımı yapılmış olup kullanılan yöntemler ve parametreler verilmiştir.

Çizelge 3.1. CloDSA'da kullanılan yöntemler ve parametreleri

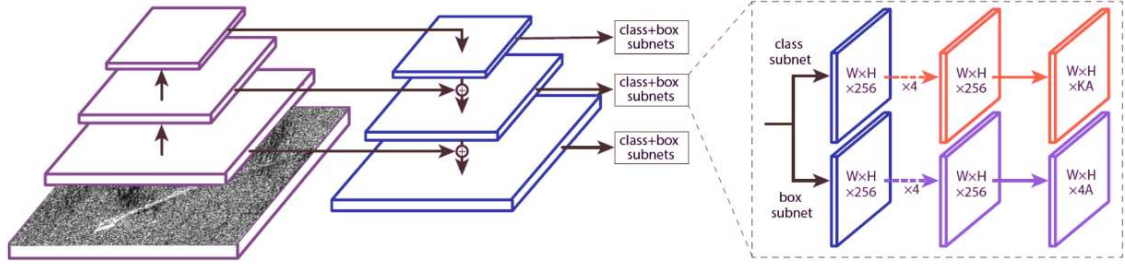
Yöntemler	Parametreler
Kesme	Yatay %25, Dikey %25
Keskinleştirme	-
Döndürme	-1 (Dikey ve Yatay)
Yeniden Boyutlandırma	INTER_LINEAR
Çevirme	45°
Medyan Bulanıklığı	5
Elastik Deformasyon	alpha: 4 sigma:1
Histogram Eşitleme	-
Gamma Düzeltmesi	1

3.2.5. RetinaNet

RetinaNet, genellikle nesne tespiti işlemleri için tasarlanmış bir derin öğrenme modelidir. Bu model, ölçeklenebilir ve yüksek doğrulukta sonuçlar vermektedir (Coelho vd., 2021). RetinaNet tek bir derin ağ içerisinde nesne sınıflandırma ve nesne tespitini aynı anda yapabilmektedir. RetinaNet tek aşamalı bir nesne algılama modelidir. Bu modelin temelinde nesne tespiti için özellik piramit ağları ve odak kaybı olmak üzere 2 temel bileşen içermektedir (Archana vd., 2022).

RetinaNet 2 aşamalı nesne algılaması ile tek aşamalı algılama yaklaşımlarının avantajlarını birleştirmektedir. 2 aşamalı nesne algılamasında ilk olarak öneri kutularını oluşturmaktadır. Daha sonra sınıflandırma ve konumun regresyonu yapılmaktadır. Böylece nesne tespiti işlemlerinde daha yüksek doğruluk oranı elde edilmesi sağlanmaktadır.

Genellikle RetinaNet çoklu sınıflandırma modellerinde tercih edilmektedir. Tensorflow ve PyTorch gibi açık kaynaklı derin öğrenme çerçevelerinde bu model kolay bir şekilde uygulanabilmektedir. Ölçeklenebilir yapısı, dengesiz veriler ile çalışabilme ve yüksek doğruluk oranı gibi sebeplerden dolayı günümüzde sıklıkla kullanılan popüler bir nesne tespit modelidir. Şekil 3.6’da RetinaNet mimarisi gösterilmektedir.



Şekil 3.6. RetinaNet mimarisi (Del Prete vd., 2021)

3.2.6. Öklid Mesafesi

Öklid mesafesi, bir uzayda 2 nokta arasındaki uzaklığı ölçen en kısa doğrusal çizgiyi ifade etmektedir. 2 noktanın x ve y eksenlerindeki koordinatlarının farkların karelerinin toplamının karekökü olarak hesaplanmaktadır (Suwanda vd., 2020). Formül 3.12’de Öklid mesafesi gösterilmektedir.

$$\text{öklid mesafesi} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (3.12)$$

Öklid mesafesi, fizik, matematik, bilgisayar bilimi ve mühendislik uygulamalarında sıklıkla kullanılmaktadır. Çalışmamızda nesne tespit modellerinin tespit ettiği nesne köşeleri ile kimlik kartının köşeleri arasındaki mesafeyi bulmak için kullanılmıştır. Böylece alınan mesafe değerlerine göre görüntünün hangi perspektifte olduğu anlaşılmaktadır.

3.2.7. Perspektif Düzenlemesi

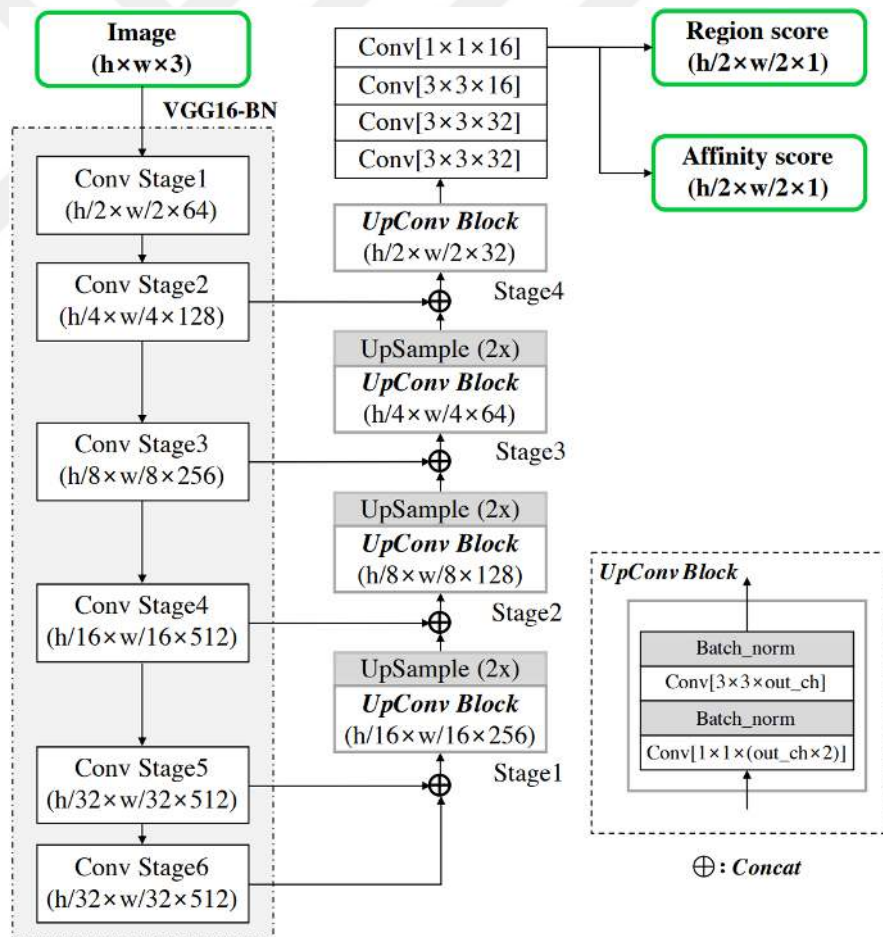
Perspektif düzenlemesi, görüntü işlemede sıklıkla kullanılan popüler bir yöntemdir. Bu yöntem bir görüntünün perspektiften kaynaklanan etkileri veya bozuklukları düzeltmeyi amaçlamaktadır (Gangal vd., 2021) Bakış açısından veya görüntüdeki nesnenin konumundan kaynaklı olan bozuklukları değiştirilmesi ve giderilmesi için kullanılmaktadır.

Perspektif düzenlemesinde öncelikle dörtgenin köşelerini belirlemek ile başlamaktadır. Köşeleri belli olan dörtgenin perspektif dönüşümü matrisinin hesaplanması gerekmektedir. Hesaplanan matris ile nesnelerin açıdan kaynaklanan perspektif etkilerinin düzeltilmesi amaçlanmaktadır. Görüntü üzerinde elde edilen matris ile perspektif düzenlemesi yapılarak ilgili nesnenin konumu ve boyutları değiştirilmektedir.

Çalışmamızda OpenCV kütüphanesinde bulunan `getPerspectiveTransform` ve `warpPerspective` fonksiyonları kullanılarak perspektif düzenlemesi yapılmıştır.

3.2.8. Metin Tespiti İçin Karakter-Bölge Farkındalığı (CRAFT)

CRAFT, CNN kullanarak tekli karakter bölgelerini belirleyip, bunları metin örnekleriyle birleştiren bir çerçeve olarak tasarlanmıştır. Karakter bölge puanı, görüntüdeki karakterleri lokalize etmek için, uygunluk puanı ise algılanan karakterleri tek bir örnek içinde gruplayabilmek için kullanılmaktadır. Yöntem, karakterlerde etiket eksikliğini gidermek için denetimli öğrenme kullanmaktadır (Baek vd., 2019). Sentetik ve gerçek veri kümelerinde elde edilen karakter düzeyi etiketleri birleştirilerek bu eksiklik giderilmektedir. CRAFT'ın özellikle eğri ve deforme metinler içeren TotalText ve CTW-1500 veri kümelerinde diğer metin tespit yöntemlerini önemli derecede geride bırakmıştır (Baek vd., 2019). Şekil 3.7'de CRAFT mimarisi gösterilmektedir.



Şekil 3.7. CRAFT mimarisi (Baek vd., 2019)

3.2.9. OTSU Eşik Belirleme Yöntemi

OTSU eşikleme yöntemi gri skaladaki bir görüntüyü 2 ayrı sınıfta siyah ve beyaz olarak ayırarak kullanılan bir otomatik eşikleme yöntemidir. (Elen, 2020). OTSU eşikleme yöntemi kontrastı yüksek olan görüntülerde etkili bir şekilde çalışmaktadır. Bilgisayarda görü, görüntü işleme, nesne tespiti ve görüntü analizi kapsamında sıklıkla kullanılan popüler bir eşik belirleme yöntemidir. Şekil 3.8’de algoritmanın kaba-kodu verilmiştir.

1. Her yoğunluk seviyesinin histogramı ve olasılıkları hesaplanmaktadır.
2. İlk $m_n(0)$ ve ayarlanmaktadır.
3. $k = 1, \dots$ olmak üzere tüm olası eşikler maksimum yoğunluk için hesaplanmaktadır.
4.
 1. m_n ve h_n güncellenmektedir.
 2. $t_a^2(k)$ hesaplanmaktadır.
5. Hedeflenen eşik maksimum $t_a^2(k)$ değerine karşılık gelmektedir.

Şekil 3.8. OTSU eşikleme yöntemi (Elen, 2020)

3.2.10. Görüntü Süper Çözünürlüğü (ISR)

ISR, görüntüleri iyileştirmek, çözünürlüğünü arttırmak ve kalitelerini yükseltmek amacıyla kullanılan bir görüntü işleme tekniğidir (Yue vd., 2016). ISR temel amacı verilen düşük çözünürlüklü ve gürültülü görüntüyü yüksek çözünürlüklü kaliteli bir çıktı oluşturmayı amaçlamaktadır. ISR genellikle video analizi, medikal görüntüleme, uzay ve savunma sanayi çalışmalarında, uydu görüntülerinde ve fotoğrafçılık uygulamalarında sıklıkla kullanılmaktadır. Çalışmamızda RDN, RRDN ve EDSR görüntü süper çözünürlüğü mimarileri kullanılmıştır.

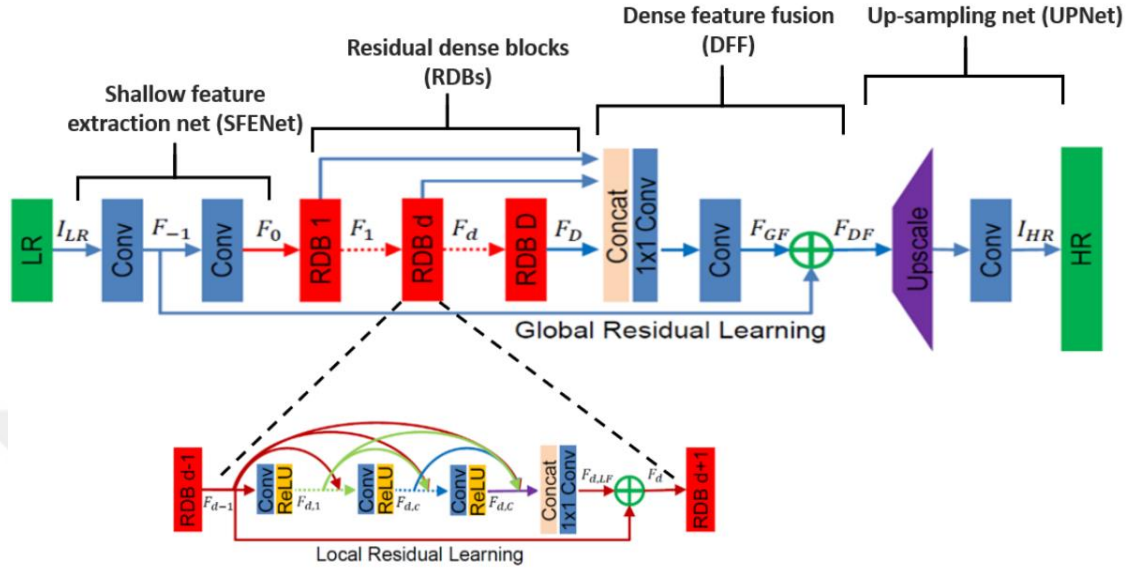
3.2.11. RDN

Artık Yoğun Ağ (RDN), mimarisi temel olarak 4 ana parçaya ayrılmaktadır (Zhang vd., 2018).

1. Yüzeysel Özellik Çıkarma Ağı (SFENet)
2. Artık Yoğun Bloklar (RDB)
3. Yoğun Özellik Birleştirme (DFF)
4. Yükseltme Ağı (UPNet)

SFENet, 2 konvolüsyonel katmandan oluşmaktadır. Bu katmanlar yüzeysel özellikleri çıkarılmasını sağlamaktadır. RDB ile bu özellikler bloklara girdi olarak

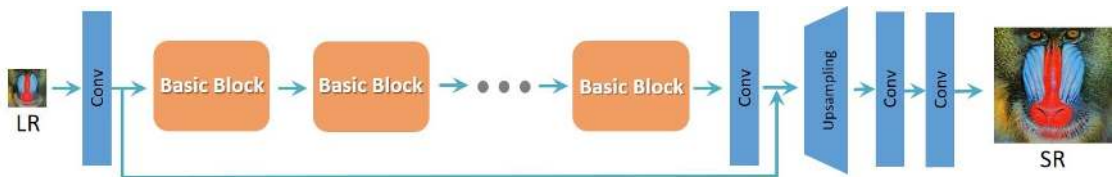
verilmektedir. DFF bileşenleri, daha önceki katmanlardaki tüm özellikleri birleştirmektedir. UPNet ile amaçlanan çıktı verilmektedir (Zhang vd., 2018). Şekil 3.9’da RDN mimarisi gösterilmektedir.



Şekil 3.9. RDN mimarisi (Noura vd., 2023)

3.2.12. RRDN

RRDN, RDN temelli VGG özellik kaybı değerleriyle eğitilmiş RRDN modeli “gans” ağırlık değerini kullanmaktadır. RRDN ağ mimarisi başlıca RRDB sayısı (T), RRDB içerisindeki RDB’lerin sayısı (D), RDB’lerin içerisindeki yığılmış konvolüsyon katman sayısı (C), RDB’lerin içerisinde bulunan her konvolüsyon katmanının özellik haritası sayısı (G) ve RDB’lerin dışındaki konvolüsyonlar ile RDB çıkışlarının özellik haritaları sayısı (G0) parametrelerini ile önceden eğitilmiş ağdan bir çıktı alınmaktadır (idealo, 2021). Şekil 3.10’da RRDN mimarisi gösterilmektedir.

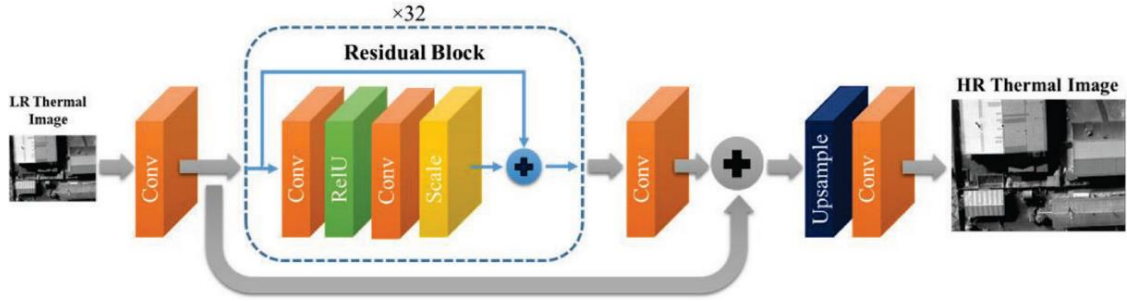


Şekil 3.10. RRDN mimarisi (idealo, 2021)

3.2.13. EDSR

Gelişmiş Derin Artık Ağlar (EDSR), NTIRE2017 Super-Resolution Challenge (Timofte vd., 2017) yarışmasını kazanmış güçlü bir görüntü süper çözünürlük

mimarisidir. Bu mimaride geleneksel ağırlarda gereksiz görülen modüllerin kaldırılmasının ve model üzerinde uygulanan optimizasyon ile önemli bir performans artışı sağlamıştır (Lim vd., 2017). Şekil 3.11’de EDSR mimarisi gösterilmektedir.



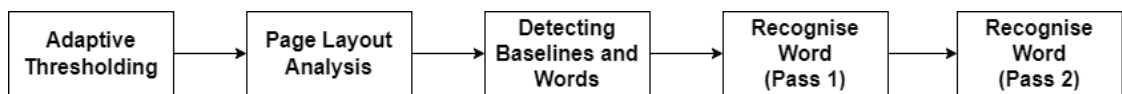
Şekil 3.11. EDSR mimarisi (Fallah vd., 2022)

3.2.14. Optik Karakter Tanıma (OKT)

OKT, yazılı bir belge veya görüntüler üzerinde bulunan yazıları bilgisayar tarafından anlamlandırılması için görüntü-metin dönüşümü yapan bir teknolojidir (Ullmann, 2019). Çalışmamızda kimlik kartı görüntülerini TesseractOCR, DonutOCR ve TrOCR OKT modelleriyle metin verilerine dönüştürme işlemi gerçekleştirilmiştir.

3.2.15. TesseractOCR

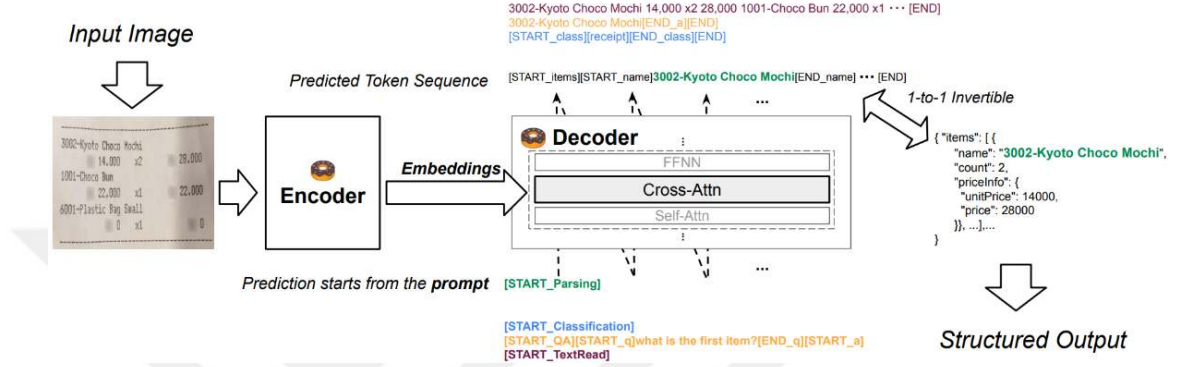
TesseractOCR, öncelikle OTSU eşik belirleme yöntemini kullanarak görüntüyü ikilik düzene çevirmektedir. İkilik düzene çevrilmiş görüntü ile belgedeki metinlerin bloklarının çıkartılması için sayfa düzeni analizi yapılacaktır. Analiz kapsamında her satırın taban çizgileri, belirsiz boşluklar ve metinlerdeki boşluklar kullanılarak kelimelere ayrılma işlemi yapılmaktadır (Akhil, 2016). Elde edilen kelimelerden karakter çizgileri çıkarılarak metin tanıma işlemi 2 süreçte yönetilmektedir. İlk süreç kapsamında statik bir sınıflandırma algoritması kullanarak gerçekleştirilmektedir. İlk süreçte düzgün anlaşılabilen metin, ikinci süreçte adaptif bir sınıflandırıcıya iletilerek model çalışmaktadır (Smith, 2007). Bu süreçte yeterince iyi tanınamayan kelime veya metinler tekrar tanınmaktadır. Şekil 3.12’de TesseractOCR mimarisinin akış diyagramı gösterilmektedir.



Şekil 3.12. TesseractOCR mimarisi (Akhil, 2016)

3.2.16. DonutOCR

DonutOCR, herhangi bir modüle bağılı olmayan transformer tabanlı kodlayıcı-kod çözücü OKT mimarisidir (Kim vd., 2021). Transformer temelli tasarlanmış basit bir mimari oluşturulması amaçlanmıştır (Vaswani vd., 2017). Bu mimari görsel kodlayıcıdan ve metinsel kod çözücü modüllerden meydana gelmektedir. Şekil 3.13’de DonutOCR mimarisi gösterilmektedir.



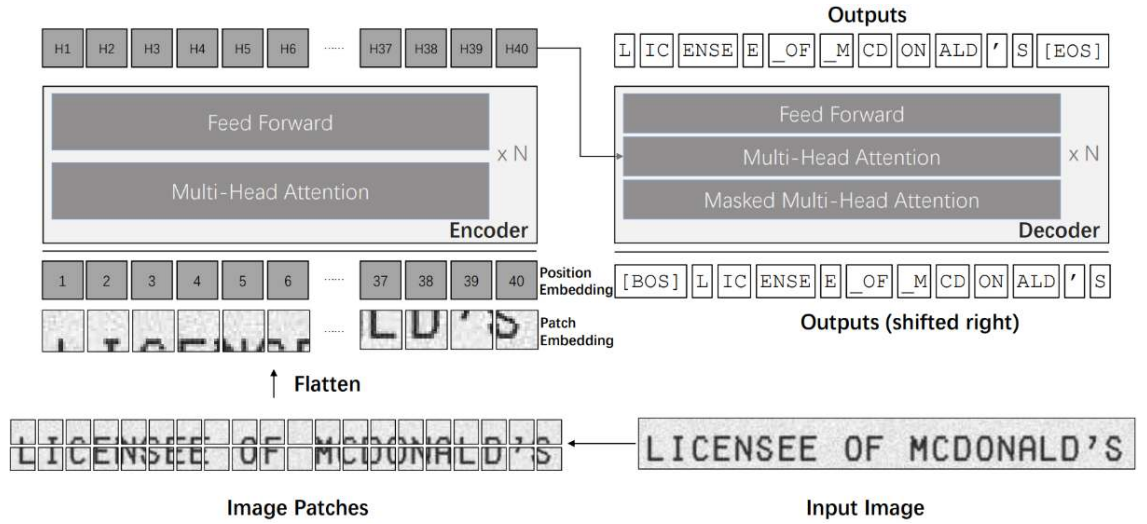
Şekil 3.13. DonutOCR mimarisi (Kim vd., 2021)

3.2.17. TrOCR

TrOCR, görüntüdeki kelime parçalarının anlaşılması ve metne çevrilmesi için Transformer tabanlı bir yaklaşım sunmaktadır. TrOCR modeli basit ve anlaşılabilir bir yapıya sahiptir. Önceden eğitilmiş verilerle eğitilebilir veya herhangi bir veri seti ile ince ayar yapılabilir (Liv d., 2023).

Deneysel aşamada TrOCR modeli doküman tarama, sahne metni tanımlamada ve el yazısı tanımlamada en gelişmiş modelleri geride bırakmaktadır. TrOCR standart bir kodlayıcı-kod çözücü modeli ile iyi sonuçlar elde etmektedir. Bu model herhangi bir konvolüsyonel işlemler içermemektedir (Liv d., 2023). Ayrıca karmaşık ön işleme adımlarına da sahip değildir.

Çalışmamızda TrOCR ile diğer OKT modellerinden daha yüksek bir doğruluk oranı elde edilmiştir. Önerilen modelimizde OKT sisteminde TrOCR mimarisinin kullanılması önerilmektedir. Şekil 3.14’de TrOCR’ın mimarisi gösterilmektedir.



Şekil 3.14. TrOCR mimarisi (Liv vd., 2023)

3.2.18. Kosinüs Benzerliği

Kosinüs benzerliği, genellikle metin madenciliğinde kullanılan 2 kelime vektörü arasındaki açıya dayanarak benzerliği ölçebilen bir hesaplama yöntemidir. Bu yöntem ile vektörlerin birbirleri olan ilişkisini ve benzerlik derecesini değerlendirmektedir (Gunawan vd., 2018).

2 vektör arasındaki açı kosinüs benzerliğine göre 1'e yaklaştıkça vektörler birbirlerine benzemektedir. Ancak aralarındaki açı 0'a yaklaştıkça benzerlik azalmaktadır. Negatif değerlikli benzerlik oranları ile bu durumda pek karşılaşmamaktadır (Vijaymeena ve Kavitha 2016). Vektörler arasındaki benzerlikler pozitif değerler ile ifade edilmektedir.

Çalışmamızda kimlik kartının üzerinde bulunan metinlerin ilgi durumlarına göre önemli bilgilerin çıkarımında kullanılacaktır. Kimlik kartının üzerinde bulunan açıklayıcı bilgileri gidermek ve önemli görülen bilgileri OKT yöntemlerinden gelen bilgiler doğrultusunda düzeltecektir.

3.2.19. Yüz Doğrulama Sistemi (RetinaNet + Arcface)

Yüz doğrulama modelleri kapsamında Lightface hafif hibrit yüz tanıma çerçevesi, FaceNet, VGG-Face, OpenFace, DeepID, Dlib ve DeepFace gibi önde gelen modelleri desteklemektedir. Temsil aşamasında, yüz görüntülerinin vektörlerle temsil edilmesi üzerinde durulmuş ve çeşitli metriklerle benzerlik hesaplamaları yapılmıştır (Serengil ve Ozpınar, 2020). Görüntülerin benzerliklerinin belirlenmesi için optimum eşik değeri, pozitif ve negatif örnekler kullanılarak bir karar ağacı algoritmasıyla belirlenmiştir. Bu

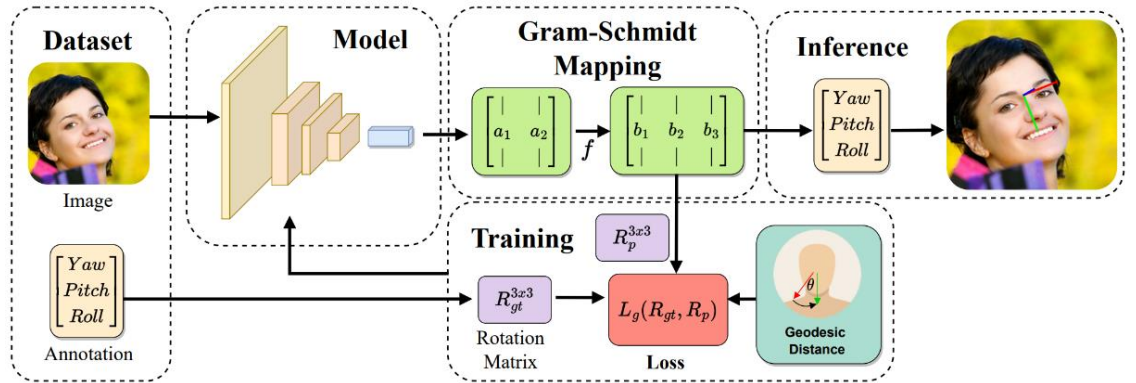
çerçeve ile farklı yüz tanıma modelleri arasında geçiş yapabilme yeteneği ile esnek bir yapı sunmaktadır (Serengil ve Ozpınar, 2020).

Çalışmamızda Lightface çerçevesini kullanarak RetinaNet + ArcFace mimarisi ile bir yüz doğrulama sistemi yapılmıştır.

3.2.20. Canlılık Analiz Sistemi (6DRepNet)

Canlılık analizi kapsamında çalışmamızda Hempel ve arkadaşlarının geliştirdiği 6DRepNet modeli kullanılmıştır (Hempel vd., 2023). Bu model ile baş yöneliminin tahmini için uçtan uca bir yaklaşım sunmaktadır. 6D döndürme matrisi kullanılarak oluşturulan modelde başın tam dönüş görünümünü etkili bir şekilde öğrenilmesi hedeflenmiştir.

Yeni birikmiş bir eğitim verisiyle beraber daha istikrarlı öğrenme için çeşitli kayıp yaklaşımları ile geniş bir baş yönelim aralığını tahmin eden bir model oluşturulmuştur. Şekil 3.15’de 6DRepNet mimarisi gösterilmiştir.



Şekil 3.15. 6DRepNet mimarisi (Hempel vd., 2023)

3.2.21. SQLite Veritabanı

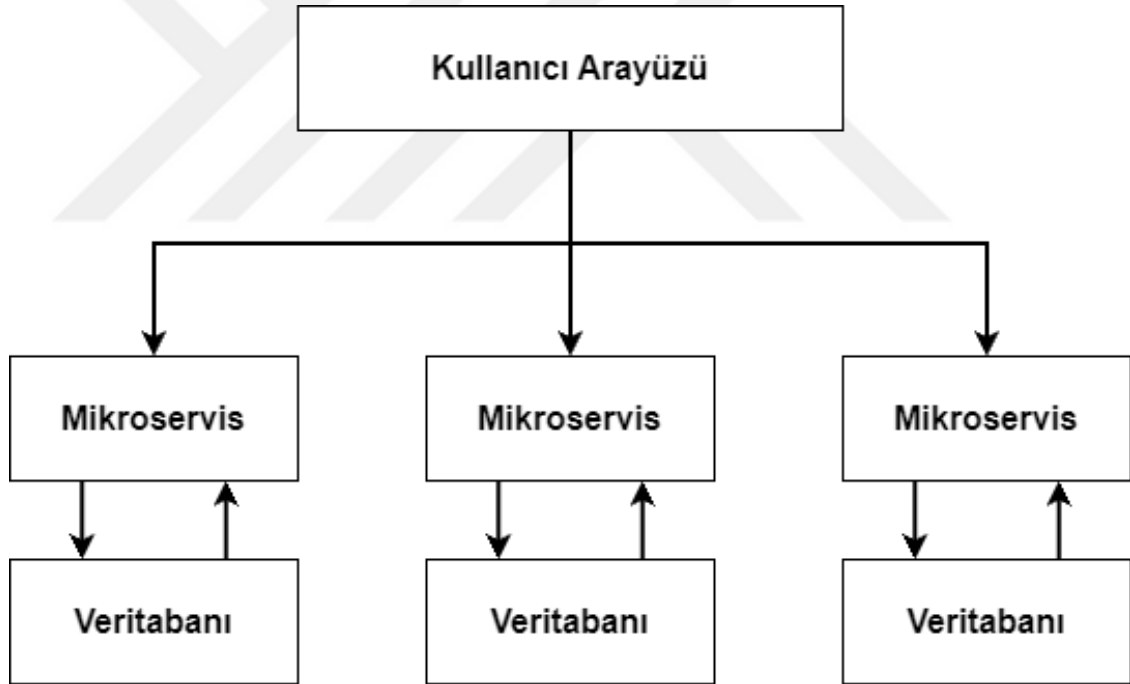
SQLite, ilişkisel veritabanı yönetim sistemidir (RDBMS). SQLite, sunucu gerektirmeyen, herhangi bir konfigrasyon ayarı istenilmeyen, platformlar arası, kendi kendine yetebilen, hafif, tam özellikli, güvenilir ve işlemsel bir veritabanıdır (Kreibich, 2010). Mobil uygulamalarda, web sitelerinde ve yerel uygulamalarda sıklıkla kullanılmaktadır.

Çalışmamızda SQLite veritabanını, kimlik kartından alınan verileri ve görüntüleri yerel olarak saklamak için kullanılacaktır. Görüntüler base64 formatında kaydedilecektir.

3.2.22. Mikroservisler

Mikroservisler, bir yazılım veya uygulamanın daha modüler, esnek ve bağımsız bir hale getirebilmek için dağıtım ve geliştirme sürecinde kullanılan bir yaklaşımdır. Bu kapsamda büyük ve karmaşık kabul edilen bir çalışmayı daha küçük bileşenlere ayırarak daha etkin, verimli ve modüler bir şekilde kullanılması sağlanmaktadır (Velepucha ve Flores, 2023).

Monolitik mimarilere bir alternatif olarak gösterilmektedir ve kullanıcılara çok sayıda avantaj sağlamaktadır. Bu kapsamda ölçeklenebilirlik, bağımsız geliştirme, teknoloji çeşitliliği, dayanıklılık, hata izolasyonu ve yeniden servis kullanımı avantajlarına örnek gösterilmektedir. Veri bütünlüğü, izleme ve güvenlik konularında mikroservisler, karmaşık ve yönetsel olarak zorluklara sahiptir. Mikroservis mimarisi uygulamaların ihtiyaç ve özelliklerine bağlı olarak seçilmeli ve tasarlanmalıdır. Şekil 3.16’da örnek bir mikroservis mimarisi gösterilmektedir.

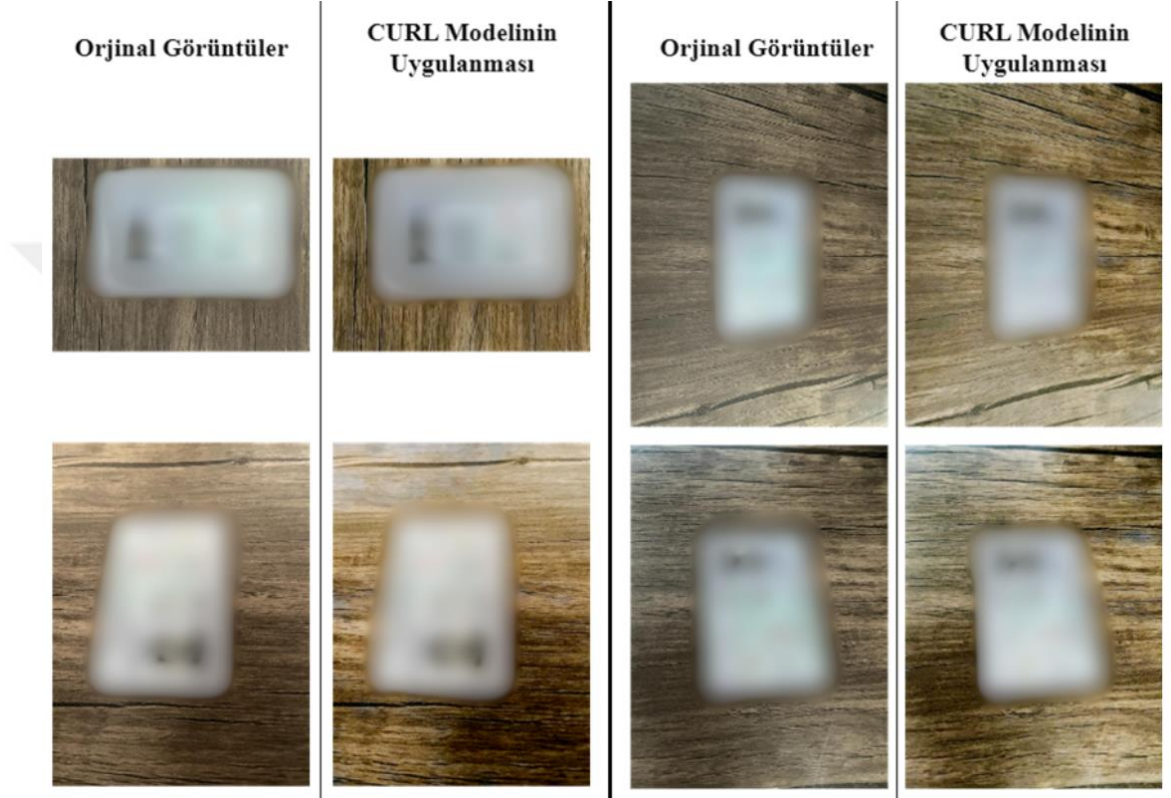


Şekil 3.16. Örnek bir mikroservis mimarisi

4. ARAŞTIRMA BULGULARI VE TARTIŞMA

4.1. Kimlik Kartı Tanıma Performans Sonuçları

Çalışmamızda kimlik kartı görüntüsünün öncelikle iyileştirilmesi için Şekil 4.1’de gösterildiği gibi CURL modeli uygulanmıştır. Böylelikle görüntüler üzerinde bilgi çıkarımı daha sağlıklı bir şekilde yapılması amaçlanmıştır.



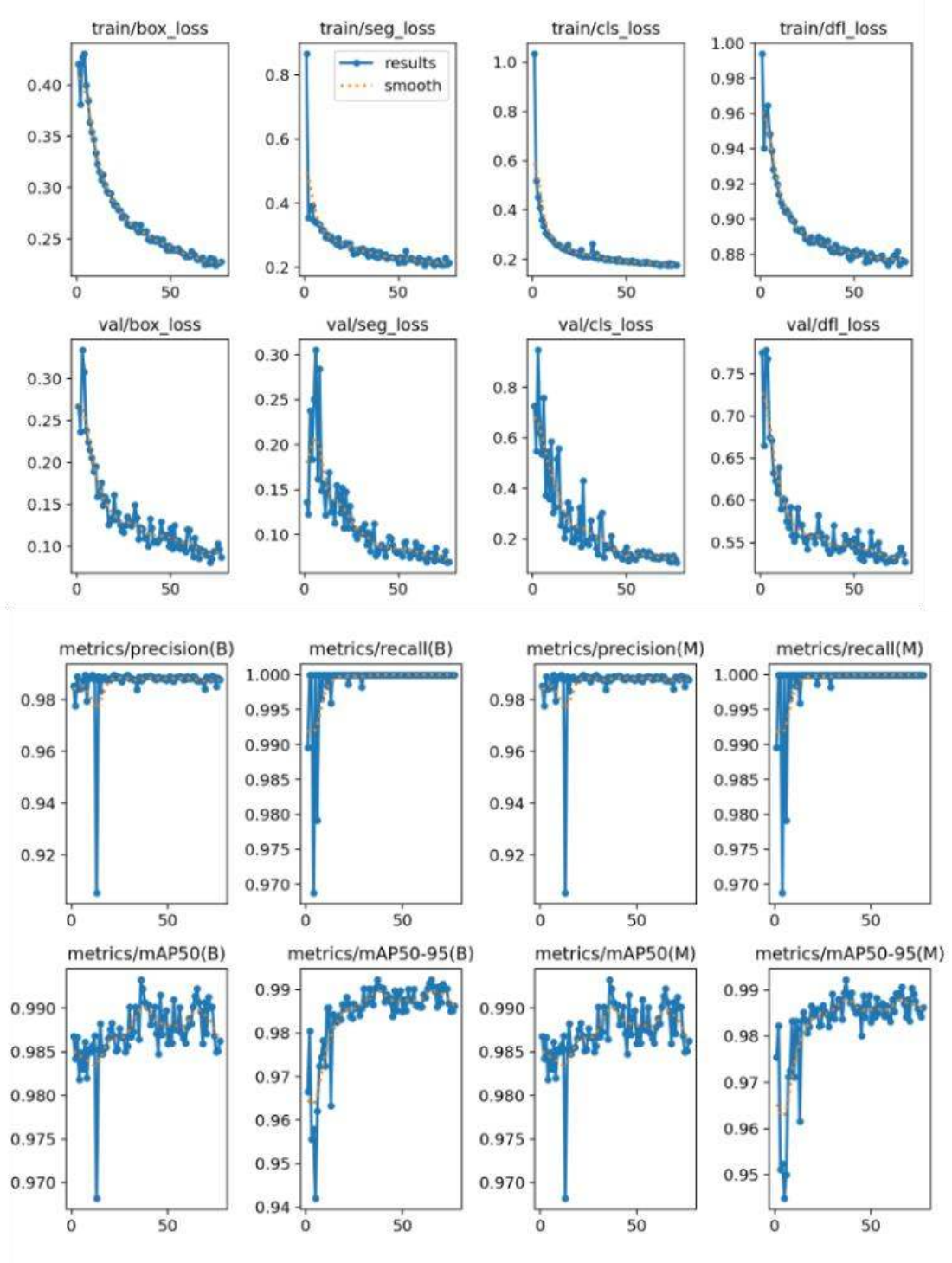
Şekil 4.1. CURL modelinin kimlik kartlarına uygulanması

YOLOv8 örnek segmentasyon modelinin performans parametreleri Çizelge 4.1’de verilmektedir.

Çizelge 4.1. YOLOv8 örnek segmentasyonu performans sonuçları

Model	Duyarlılık	Hassasiyet	mAP
YOLOv8 (Instance Segmentation)	99,1	99,2	99,2

YOLOv8 örnek segmentasyon modelinin performansı grafikleri Şekil 4.2’de verilmiştir.



Şekil 4.2. YOLOv8 örnek segmentasyon modelinin performans grafikleri

CURL filtresinden gelen görüntüler üzerinde Şekil 4.3’de görüldüğü gibi YOLOv8 örnek segmentasyonu yapılmıştır ve orijinal görüntüler ile karşılaştırılmıştır.

Orjinal Görüntüler



Nesne Segmentasyonu (YOLOv8)



Şekil 4.3. Kimlik kartlarına örnek segmentasyonu uygulanması

YOLOv8 ile örnek segmentasyonu aşamasından sonra segmentasyon yapılan kimlik kartının koordinatları elde edilmiştir. Koordinatlar üzerinde kimlik kartının maksimum alanını veren 4 nokta tespit edilecektir. Bu arama aşaması GA ile optimize edilmiştir.

GA kullanarak bir konveks yamuk içinde en büyük alanı kaplayan 4 noktayı bulmayı amaçlayan bir optimizasyon probleminin çözülmesi amaçlanmıştır. Bu problemi aşağıda bulunan adımlar ile çözülmüştür.

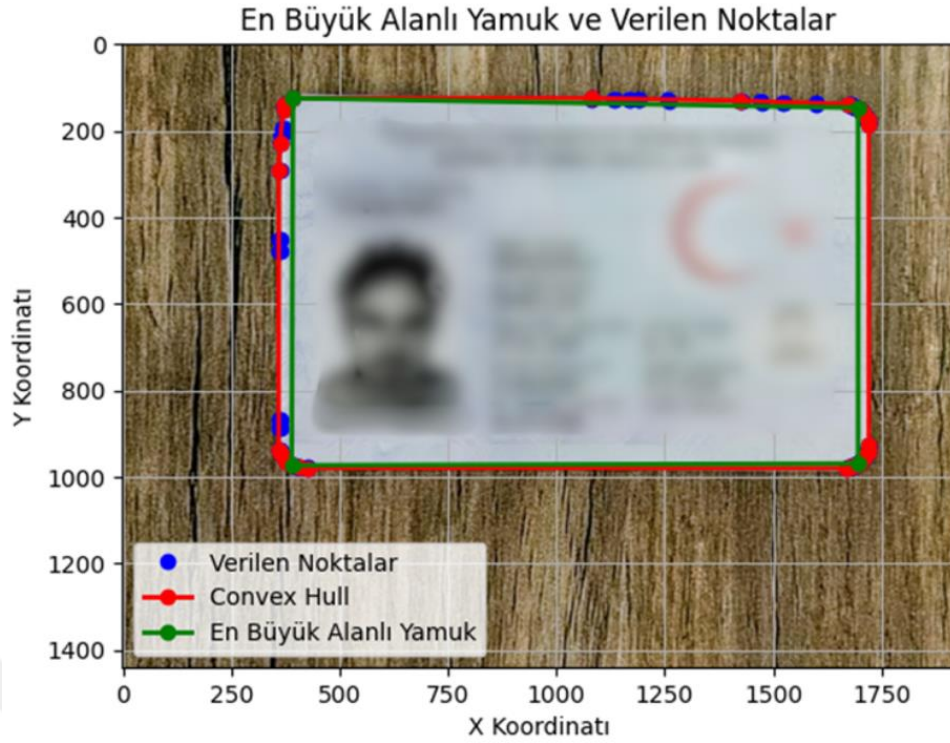
1. Genetik algoritma konfigürasyonu DEAP kütüphanesi kullanılarak yapılmaktadır.
2. Popülasyon, bireylerin koordinatlarını temsil eden gen özelliklerini içeren rastgele noktalardan oluşturulmaktadır.
3. Aptitude fonksiyonu, her bir bireyin convex hull içindeki alanını hesaplamaktadır.
4. Genetik operatörler tanımlanır: çaprazlama (blend crossover) ve mutasyon (Gaussian mutation).
5. Genetik algoritma, belirli sayıda nesil üzerinde çalıştırılmaktadır.
6. En iyi birey, konveks poligon içinde en büyük alanı kapsayan 4 noktayı temsil etmektedir.
7. Sonuçlar görselleştirilir: Verilen noktalar, konveks poligon ve en büyük alanı kaplayan 4 nokta grafik olarak gösterilmektedir.

Çalışmada;

- Popülasyon büyüklüğü ($n=100$): Başlangıçta oluşturulan popülasyonun birey sayısını temsil etmektedir.
- $\mu = 50$ (mu): GA'nın kullanacağı elit birey sayısıdır. Her nesilde en iyi bireylerin seçileceği kısmı belirtmektedir.
- $\lambda = 50$ (lambda): her nesilde oluşturulacak yeni birey sayısıdır. Çaprazlama ve mutasyon işlemleri sonucu üretilen bireylerin sayısını belirtmektedir.
- $c_{xpb}=0.7$: Her iki ebeveynin çaprazlanma olasılığını belirtmektedir. Bu durumda, %70 olasılıkla çaprazlama gerçekleştirilmektedir.
- $m_{utpb}=0.2$: Bir bireyin genlerinden herhangi birinin mutasyona uğrama olasılığını belirtmektedir. Bu durumda, %20 olasılıkla mutasyon gerçekleştirilmektedir.
- $n_{gen}=100$: Genetik algoritmanın çalıştırılacağı toplam nesil sayısını belirtmektedir.

Bu hiperparametreler, GA'nın performansını ve çözüm kalitesini etkileyen önemli faktörlerdir. Optimal hiperparametre değerleri, spesifik bir probleme ve veri kümesine bağlı olarak değişebilmektedir.

Şekil 4.4'de kimlik kartı üzerinde yapılan arama işlemi gösterilmektedir. Görselde kırmızı ile gösterilen koordinatlar QuickHull algoritmasından alınan sonuçlardır. Bu sonuçlar üzerinde maksimum alanı elde eden 4 noktayı tespit edilmek istenilmektedir. Yeşil nokta ile gösterilen koordinatlar ise maksimum alanı veren köşelerdir.



Şekil 4.4. Kimlik kartında maksimum alan arama algoritmasının çıktısı

Kimlik kartının 4 köşesi elde edilmesiyle birlikte kimliğin sağ alt, sağ üst, sol alt ve sol üst kısımları perspektif düzenlemesi için bilinmesi gereklidir. 4 noktanın hangi köşelere denk gelebildiğini anlayabilmek için 2 farklı nesne tanıma algoritması kullanılmıştır. Yüz tespiti ile kimlik kartının sol alt ve sol üst köşeleri elde edilmiştir. Türk Bayrağını tespit ederek kimlik kartının sağ alt ve sağ üst köşeleri elde edilmiştir.

Çalışmamızda yüz tespiti için PyPI’de bulunan RetinaFace kütüphanesi kullanıldı. Bu kütüphane ile yüksek doğrulukta yüz tespiti sağlandı. Kimlik kartında bulunan Türk Bayrağının tespiti için YOLOv8 nesne tespiti modeli kullanılarak Çizelge 4.2.’deki sonuçlar elde edilmiştir.

Çizelge 4.2. YOLOv8 nesne tespiti performans sonuçları

Model	Duyarlılık	Hassasiyet	mAP
YOLOv8 (Object Detection)	97,4	97,6	97,6

Köşelerin hangi bölgede olduğunu bulabilmemiz için yüz tespiti ve Türk Bayrağında algoritmalar ile bulunan işaretli dörtgenlerin orta noktası bulunmuştur. Şekil 4.5’de kimlik kartı üzerinde bulunan yüzün ve Türk Bayrağının orta noktası tespit edilmiştir.



Şekil 4.5. Kimlik kartında Türk Bayrağı ve yüz görüntüsünün orta noktası

Şekil 4.6'da bulunan orta nokta ile köşeler arasındaki Öklid mesafesi hesaplanmaktadır.

Yüz ile en yakın mesafeler:

Köşe 1 ile Yüz arasındaki mesafe: 627.465536902227

Köşe 2 ile Yüz arasındaki mesafe: 1209.4560760937125

Köşe 3 ile Yüz arasındaki mesafe: 1103.5524455140317

Köşe 4 ile Yüz arasındaki mesafe: 354.9380227589036

Yüz ile en yakın 1. mesafe Köşe = [389, 973] Mesafe: 354.9380227589036

Yüz ile en yakın 2. mesafe Köşe = [392, 124] Mesafe: 627.465536902227

Bayrak ile en yakın mesafeler:

Köşe 1 ile Bayrak arasındaki mesafe: 1095.3214140150826

Köşe 2 ile Bayrak arasındaki mesafe: 378.1904282236662

Köşe 3 ile Bayrak arasındaki mesafe: 595.9060328608865

Köşe 4 ile Bayrak arasındaki mesafe: 1186.996630155284

Bayrak ile en yakın 1. mesafe Köşe = [1696, 147] Mesafe: 378.1904282236662

Bayrak ile en yakın 2. mesafe Köşe = [1696, 969] Mesafe: 595.9060328608865

Şekil 4.6. Köşeler ile bulunan noktalar arasındaki öklid uzaklığı

Öklid uzaklıklarını baz alarak kart köşelerinin hangi konuma geldiği hesap edilerek Şekil 4.7'de gösterildiği gibi perspektif düzenlemesi yapılmıştır.



Şekil 4.7. Kimlik kartı üzerinde perspektif düzenlenmesi

Elde edilen kimlik görüntüsü 1700x1100 piksel çözünürlüğüne getirilmiştir. Daha sonra kart üzerindeki yazıların tespiti için CRAFT modeli kullanılmıştır. Şekil 4.8’de bu model ile alınan sonuçlar gösterilmektedir.



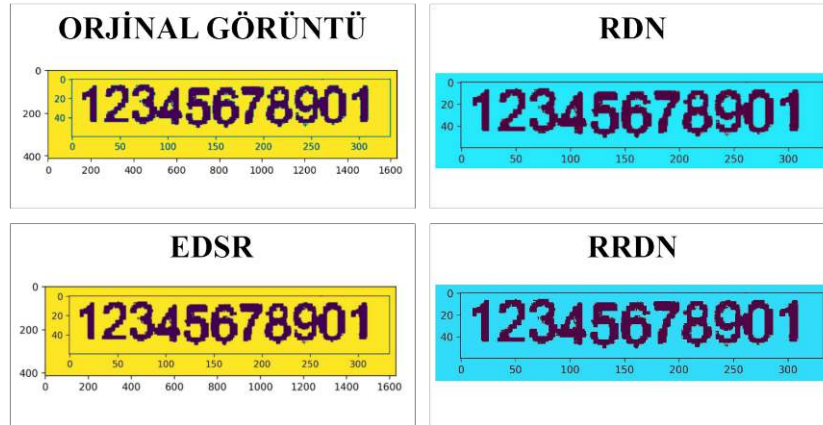
Şekil 4.8. Kimlik kartı üzerinde CRAFT modelinin uygulanması

Şekil 4.9’da CRAFT ile alınan örnek görüntü perspektif düzenleme aşamasından sonra OTSU yöntemi ile filtreleme işlemi yapılmıştır.



Şekil 4.9. OTSU yönteminin ROI'lere uygulanması

Otsu yönteminden elde edilen görüntü üzerine RRDN, RDN ve EDSR olmak üzere ISR modelleri uygulanmıştır. Şekil 4.10’da gösterilmektedir.



Şekil 4.10. Orjinal görüntüye RDN, RRDN ve EDSR modellerinin uygulanması

Elde edilen yüksek çözünürlüklü görüntüler ile OKT uygulaması yapılmıştır. Yapılan OKT uygulaması kapsamında 100 adet farklı arka planlar görüntülerine sahip kimlik kullanılarak TesseractOCR, DonutOCR ve TrOCR modellerinin performans sonuçları Çizelge 4.3.’de gösterilmektedir.

Çizelge 4.3. OKT modellerinin karşılaştırılması

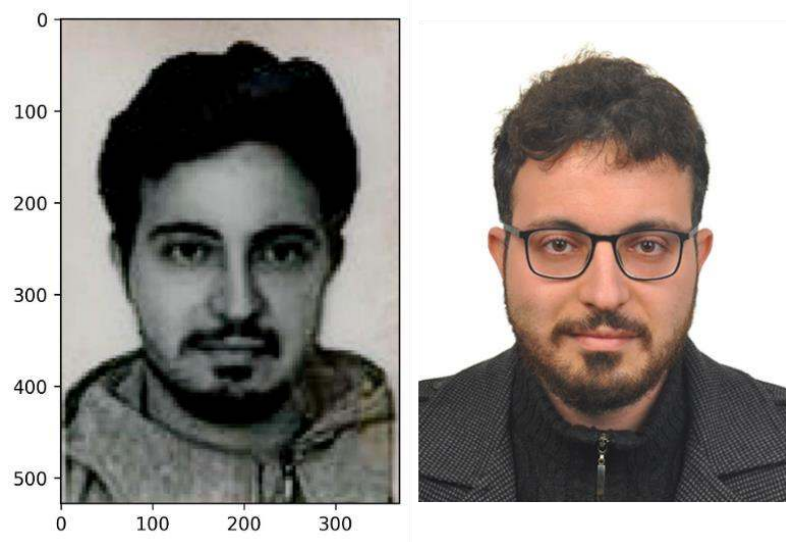
OKT Mimarisi	Doğru Bilinen Kelime Sayısı
TesseractOCR	%81
DonutOCR	%89
TrOCR	%94

Kimlik kartında üzerinde bulunan yazıların ROI'leri alınarak OKT modelleriyle metin haline dönüştürülmüştür. Ancak kimlik numarası, soyadı, adı, doğum tarihi, geçerlilik tarihi, seri numarası, cinsiyeti ve uyruğunun dışında açıklama yapan metinleri de algoritma okumaktadır. Elimizde bulunan bu metin verisi üzerinden kosinüs benzerliğini kullanarak süzgeçleme işlemi yapılmıştır.

Yapılan işlemler kapsamında açıklama metinlerini ayrı kelimeler ile bir metin dosyasında tutulmuştur. Metin dosyasındaki verilerimiz ile tespit edilen veriler arasında bir kosinüs benzerlik matrisi oluşturulmuştur. Benzer veriler yani açıklamaların kelime vektör değerleri 1'e yaklaşırken diğer veriler önemli kart bilgileri 0'a doğru yaklaşmaktadır. 0,6 eşik değerince algoritmamız kusursuz olarak çalışmakta her bir veriyi elde etmektedir.

4.2. Canlılık ve Benzerlik Analizi Performans Sonuçları

Çalışmamızın benzerlik analizi kapsamında Arcface + RetinaNet kullanarak benzerlik analizi yapılmıştır. Şekil 4.11'de bulunan görüntülere 0,68 puanla başarıyla doğrulama yapılmıştır.



Şekil 4.11. Arcface + RetinaNet ile benzerlik analizi

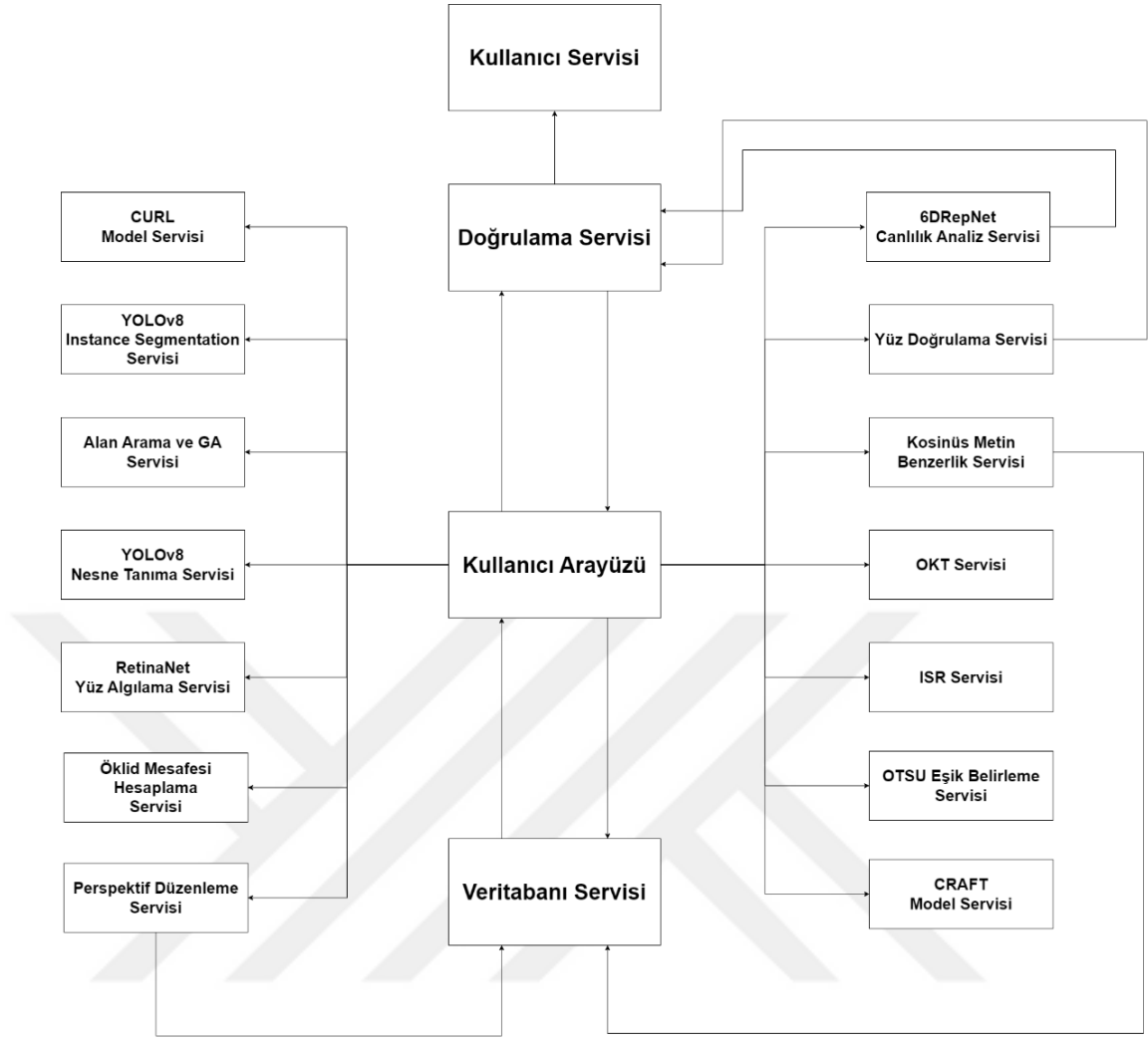
Kimlik doğrulama aşamasında karşıdakinin kişinin canlı bir kullanıcı olduğunu kanıtlamak için canlılık analizi yapılmıştır. Yapılan çalışmada baş sağa, sola, yukarı, aşağıya ve kameraya yaklaşması istenilmiştir. Şekil 4.12’de bulunan örnek bir görüntüde 6 eksenli yüksek hassasiyetli bir poz tahmini yöntemi kullanılmıştır. Kullanılan yöntem ile pitch, yaw ve roll değerlerini hesap ederek karşıdaki bireyin başı izlemeye alınmıştır. Verilen komutlar ışığında canlılık tespiti yapılmıştır.



Şekil 4.12. 6DRepNet ile canlılık analizi sistemi (Hempel vd., 2023)

4.3. Mikroservis Mimarisi ile Kullanıcı Arayüzü

Çalışmamızda 15 adet hesaplama birimi bulunmaktadır. Bu birimlerin içinde derin öğrenme, görüntü işleme ve çeşitli matematiksel ifadelerde vardır. Çok kullanıcıli sunucularda veya büyük sistemlerde monolitik mimarilerde ciddi sorunlar oluşabilmektedir. Bu birimlerin esnek ve modüler bir yapıda olması gerekmektedir. Bu modüler servislerin birbirleriyle anlaşması ve konuşması gereklidir. Mikroservis yapısı ile bu yazılım problemi çözülebilmektedir. Şekil 4.13’de verilen görselde çalışmamızın mikroservis tasarımı yapılmıştır.



Şekil 4.13. Mikroservis mimarisi ile kimlik tanıma ve canlılık analiz sistemi

4.4. Tartışma

Çalışmamızda dijital kimlik kartı tespiti yapılmıştır. Kimlik kartının gerçek kişi tarafından verildiğini tespit edebilmek için yüz tanımlama ve canlılık analizi yapılarak 2 faktörlü doğrulama ile gerçek kullanıcı ile kimlik eşleştirilmesi sağlanmıştır. Yapılan sistemin iyileştirilmesi gereken çeşitli noktaları bulunmaktadır. YOLOv8 nesne tespit modeli daha fazla veri kümesi ile daha yüksek başarımlar elde edilebilir. Görüntü Süper Çözünürlüğü kısmında ise metin süper çözünürlüğü kısmına bakılıp metin için özelleşen algoritmalar kullanılmalıdır. Böylelikle sistemimizin için özelleşen mimariler ile birlikte daha güçlü doğruluk oranına sahip bir model kullanılabilir.

Bu çalışmanın görüntü işleme ve derin öğrenme kısımlarında eksiklik bulunmayıp mevcut modelin iyileştirilmesi düşünülmelidir. Çalışma kapsamında yeni derin öğrenme yaklaşımları ve görüntü işleme yöntemleri ile çalışılmış olup yüksek performanslı bir

model oluşturulmuştur. Oluşturulan bu sistem muadili olan sistemlerden kimlik kartı tespitinde daha başarılı olduğu görülmektedir.

Çalışma kapsamında köşe arama algoritması GA ile optimizasyonu yapılmış olup literatüre kazandırılmıştır. Ayrıca kosinüs benzerlik matrisini kullanarak ilgili verilerin kusursuz ayrımı da yapılmıştır. Bu tekniklerin bir arada kullanılmasıyla mevcut sistemin güçlü bir şekilde kimlik tanıma ve doğrulama yaptığı anlaşılmaktadır. Çalışmada alınan sonuçlar sektör için yeterli olduğunu ancak sistem açısından da geliştirilebilir olduğu vurgulanmaktadır.



5. SONUÇ VE ÖNERİLER

Günümüz dünyası, dijitalleşmeye doğru evrilmektedir. Klasikleşmiş olan belgeler, dökümanlar veya yazılı metinler artık daha akıllı algoritmalarla çıkarım yapabilen modeller ile okunması kolaylaşmaktadır. Bireylerin bu dijitalleşmeye girebilmesi ve internet ortamında dijital bir kimliğe sahip olması son derece önemlidir. Dijital kimlikler ile birlikte dijital kimlik güvenliği problemi de ortaya çıkmaktadır. Bu sebepten dolayı dijital kimlik doğrulama ve onaylama algoritmalarının da aynı hızda gelişmesi gereklidir. Bu çalışmayla beraber sektör ihtiyaçlarına gözeterek kullanıcıların dijital kimliklere sahip olmasını ve bu süreci güvenilir ve verimli şekilde sağlanması amaçlanmıştır.

Bu çalışma, dijital kimlik doğrulamanın önemli bir yönü olan kimlik kartlarının güvenliğini artırmak amacıyla çeşitli teknolojik yöntemlerin entegre edildiği kapsamlı bir yaklaşım sunmaktadır. Türkiye Cumhuriyet Kimlik Kartları üzerinde gerçekleştirilen detaylı görüntü analizi ve işleme adımlarıyla, optik karakter tanıma, nesne tespiti, yüz tanıma ve canlılık analizi gibi çeşitli teknikler kullanılarak güvenli kimlik doğrulama sağlanmaktadır.

Özellikle, derin öğrenme modelleri aracılığıyla yapılan görüntü iyileştirmesi, örnek segmentasyonu ve metin tespiti adımları, kimlik kartlarının üzerindeki bilgilerin doğru ve güvenilir bir şekilde çıkarılmasını sağlamak için etkili bir şekilde bir araya getirilmiştir. Yüz tanıma ve benzerlik analizi adımları, kart sahibinin kimlik kartını tanıyan kişi olup olmadığını belirlemede önemli bir rol oynamaktadır.

Sisteme verilen kimlik kartı üzerinde CURL modeli, YOLOv8 ile örnek segmentasyonu, arama algoritması, YOLOv8 ile nesne tespiti, RetinaNet ile yüz tespiti, Öklid mesafe hesaplaması, perspektif düzenlemesi, CRAFT modeli, OTSU eşik belirleme yöntemi, görüntü süper çözünürlüğü mimarileri, optik karakter tanıma mimarileri, kosinüs benzerliği, RetinaNet + ArcFace ile yüz doğrulama sistemi ve canlılık analizi kapsamında 6DRepNet poz tahmini mimarisi çalıştırılmıştır. Alınan tüm sonuçlar bu çalışma kapsamında paylaşılmış olup doğruluğu yüksek bir model sunulmuştur.

Bu sistem aynı zamanda mikroservis mimarisi ile tasarlanmıştır, bu da hem ölçeklenebilirlik hem de esneklik açısından avantajlar sunmaktadır. Güvenlik ve etkinlik odaklı bu entegre çözüm, dijital kimlik doğrulama süreçlerini daha güvenilir ve etkili hale getirerek çeşitli uygulama alanlarına katkı sağlamaktadır. Elde edilen performans sonuçları, sistemin başarılı bir şekilde çalıştığını ve güvenlik standartlarına uygun bir dijital kimlik doğrulama çözümü sunduğunu göstermektedir.

Elde edilen bulgular kapsamında sektör ihtiyalarına uygun bir dijital kimlik tanıma ve doėrulama yazılımı geliştirilmiř ve mevcut kořullar altında kullanımı önerilmiřtir.



KAYNAKLAR

- Abraham, A. (2005). Artificial Neural Networks. *Handbook of Measuring System Design*, John Wiley & Sons, (Baskıda).
- Akhil, S. (2016). *An Overview of Tesseract OCR Engine*. (Seminar Report). Department of Computer Science and Engineering National Institute of Technology, Calicut Monsoon.
- Albawi, S., Mohammed, T. A., and Al-Zawi, S. (2017). Understanding of a Convolutional Neural Network. *In 2017 international conference on engineering and technology (ICET)*, 1-6.
- Archana, V., Kalaiselvi, S., Thamaraiselvi, D., Gomathi, V., and Sowmiya, R. (2022). A Novel Object Detection Framework using Convolutional Neural Networks (CNN) and RetinaNet. *In 2022 International Conference on Automation, Computing and Renewable Systems*, 1070-1074.
- Baek, Y., Lee, B., Han, D., Yun, S., and Lee, H. (2019). Character Region Awareness for Text Detection. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 9365-9374.
- Barber, C. B., Dobkin, D. P., and Huhdanpaa, H. (1996). The Quickhull Algorithm for Convex Hulls. *ACM Transactions on Mathematical Software (TOMS)*, 22(4), 469-483.
- Baskar, A., Rajappa, M., Vasudevan, S. K., and Muruges, T. S. (2023). *Digital Image Processing*. CRC Press.
- Bouvier, J. (2006). Notes on Convolutional Neural Networks, Massachusetts Institute of Technology.
- Casado-García, Á., Domínguez, C., García-Domínguez, M., Heras, J., Inés, A., Mata, E., and Pascual, V. (2019). CLoDSA: a Tool for Augmentation in Classification, Localization, Detection, Semantic Segmentation and Instance Segmentation Tasks. *BMC bioinformatics*, 20, 1-14.
- Chicco, D., Warrens, M. J., and Jurman, G. (2021). The Coefficient of Determination R-Squared is More Informative than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation. *PeerJ Computer Science*, 7, e623.
- Coelho, C., Costa, M. F. P., Ferrás, L. L., and Soares, A. J. (2021). Object Detection with Retinanet on Aerial Imagery: The Algarve Landscape. *In International Conference on Computational Science and Its Applications*. 501-516. Cham: Springer International Publishing.
- Çetiner, H., Cetişli, B., ve Çetiner, İ. (2012). Gerçek Zamanlı Kimlik Numarası Tanıma. *Sakarya University Journal of Science*, 16(2), 123-129.

- Dang, T. V., and Tran, H. L. (2023). A Secured, Multilevel Face Recognition based on Head Pose Estimation, MTCNN and FaceNet. *Journal of Robotics and Control (JRC)*, 4(4), 431-437.
- Das, S., Ma, K., Shu, Z., Samaras, D., and Shilkrot, R. (2019). Dewarpnet: Single-Image Document Unwarping with Stacked 3D and 2D Regression Networks. *In Proceedings of the IEEE/CVF International Conference on Computer Vision*, 131-140.
- Del Prete, R., Graziano, M. D., and Renga, A. (2021). RetinaNet: A Deep Learning Architecture to Achieve A Robust Wake Detector in SAR Images. *In 2021 IEEE 6th International Forum on Research and Technology for Society and Industry (RTSI)*, 171-176.
- Deng, J., Guo, J., Ververas, E., Kotsia, I., and Zafeiriou, S. (2020). Retinaface: Single-Shot Multi-Level Face Localisation in the Wild. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5203-5212.
- Elen, A. (2020). Görüntü İkileştirme İçin Global Eşikleme Yöntemleri Üzerine Bir İnceleme. *Mühendislik Bilimleri ve Araştırmaları Dergisi*, 2(2), 38-49.
- Emel, G. G., ve Taşkın, Ç. (2002). Genetik Algoritmalar ve Uygulama Alanları. *Uludağ Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 21(1), 129-152.
- Fallah, A., Samadzadegan, F., and Dadrass Javan, F. (2022). Intensifying the Spatial Resolution of 3D Thermal Models from Aerial Imagery Using Deep Learning-Based Image Super-Resolution. *Geocarto International*, 37(26), 13518-13538.
- Fang, X., Fu, X., and Xu, X. (2017). ID Card Identification System Based on Image Recognition. *In 2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 1488-1492.
- Fraihat, S., and Al-Betar, M. A. (2023). A Novel Lossy Image Compression Algorithm Using Multi-Models Stacked AutoEncoders. *Array*, 19, 100314.
- Gadea, C., Trifan, M., Ionescu, D., Cordea, M., and Ionescu, B. (2016). A Microservices Architecture for Collaborative Document Editing Enhanced With Face Recognition. *In 2016 IEEE 11th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, 441-446.
- Gangal, A., Kumar, P., and Kumari, S. (2021). Complete Scanning Application Using Opencv. *arXiv preprint arXiv:2107.03700*.
- Gezer, M. (2021). Tıpta Görüntü İşleme Yöntemleri. İstanbul: İstanbul University. *Tıp Bilişimi*, 203-218.
- Gomes, A. J. (2016). A Total Order Heuristic-Based Convex Hull Algorithm For Points in The Plane. *Computer-Aided Design*, 70, 153-160.

- Gunawan, D., Sembiring, C. A., and Budiman, M. A. (2018). The Implementation of Cosine Similarity to Calculate Text Relevance Between Two Documents. *In Journal of Physics: Conference Series*, 978, 012120. IOP Publishing.
- Hanbay, K. (2022). Gauss Filtreleri Kullanılarak Geliştirilen Bölge Temelli Yeni Bir Aktif Kontur Yöntemi. *Computer Science*, 7(1), 29-35.
- Hao, W., and Zhili, S. (2020). Improved Mosaic: Algorithms for More Complex Images. *In Journal of Physics: Conference Series*, 1684(1), 012094, IOP Publishing.
- Hassoun, M. H. (1995). Fundamentals of Artificial Neural Networks. *MIT press*, (Baskıda).
- Hempel, T., Abdelrahman, A. A., and Al-Hamadi, A. (2023). Towards Robust and Unconstrained Full Range of Rotation Head Pose Estimation. *arXiv preprint arXiv:2309.07654*.
- Huang, J., and Ling, C. X. (2005). Using AUC and Accuracy in Evaluating Learning Algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 17(3), 299-310.
- İdealo. (2021). Image Super Resolution. Erişim adresi: <https://github.com/idealo/image-super-resolution>
- İnik, Ö., ve Ülker, E. (2017). Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri. *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, 6(3), 85-104.
- Jamal, A., Helmi, R. A. A., Syahirah, A. S. N., and Fatima, M. A. (2019). Blockchain-Based Identity Verification System. *In 2019 IEEE 9th International Conference on System Engineering and Technology (ICSET)*, 253-257.
- Jeong, H. J., Park, K. S., and Ha, Y. G. (2018). Image Preprocessing for Efficient Training of YOLO Deep Learning Networks. *In 2018 IEEE International Conference on Big Data and Smart Computing (BigComp)*, 635-637.
- Kavut, S. (2020). Kimliğin dönüşümü: Dijital kimlikler. *Selçuk İletişim*, 13(2), 987-1008.
- Kavut, S. (2023). Türkiye’de Dijital Kimlik ve Dijital Gözetim Üzerine Bir Değerlendirme. *Dijital Çağda Medya*, 67. Erişim adresi: https://books.google.com/books?hl=tr&lr=&id=uyTFEAAAQBAJ&oi=fnd&pg=PA67&dq=kimlik+tan%C4%B1ma+sistemleri&ots=9vzei6OBOJ&sig=8eaLHD8KpKcIQ6_BPYGTL_9g7lM
- Kaya, S. (2021). Sosyal Medya Kullanıcılarının Dijital Kimlik İnşası Sürecinde Kültürel Faaliyet ve Sanat ile İlgili Paylaşımlarının Rolü. *Turkish Online Journal of Design Art and Communication*, 11(4), 1403-1419.

- Kim, G., Hong, T., Yim, M., Park, J., Yim, J., Hwang, W., Yun, S., Han, D., and Park, S. (2021). Donut: Document Understanding Transformer Without Ocr. *arXiv preprint arXiv:2111.15664*, 7, 15.
- Koyun, A., ve Afşin, E. (2017). Derin Öğrenme İle İki Boyutlu Optik Karakter Tanıma. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 10(1), 11-14.
- Köseoğlu, K. (2008). Veritabanı Mantığı. *Pusula*, (Baskıda).
- Kreibich, J. (2010). Using SQLite. *O'Reilly Media, Inc*, (Baskıda).
- Krstinić, D., Braović, M., Šerić, L., and Božić-Štulić, D. (2020). Multi-Label Classifier Performance Evaluation with Confusion Matrix. *Computer Science & Information Technology*, 1.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep Learning. *Nature*, 521(7553), 436-444.
- Li, M., Lv, T., Chen, J., Cui, L., Lu, Y., Florencio, D., Zhang C., Li, Z., and Wei, F. (2023). Trocr: Transformer-Based Optical Character Recognition with Pre-Trained Models. *In Proceedings of the AAAI Conference on Artificial Intelligence*, 37(11), 13094-13102.
- Lim, B., Son, S., Kim, H., Nah, S., and Mu Lee, K. (2017). Enhanced Deep Residual Networks for Single Image Super-Resolution. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 136-144.
- Liu, T., Liu, H., Yang, B., and Zhang, Z. (2023). LDCNet: Limb Direction Cues-aware Network for Flexible Human Pose Estimation in Industrial Behavioral Biometrics Systems. *IEEE Transactions on Industrial Informatics*.
- Liu, Y., Zhou, Y., Wen, S., and Tang, C. (2014). A Strategy On Selecting Performance Metrics for Classifier Evaluation. *International Journal of Mobile Computing and Multimedia Communications (IJMCMC)*, 6(4), 20-35.
- Long, S., He, X., and Yao, C. (2021). Scene Text Detection and Recognition: The Deep Learning Era. *International Journal of Computer Vision*, 129, 161-184.
- Lotufo, R. A., Audigier, R., Saúde, A. V., and Machado, R. C. (2023). Morphological Image Processing. *In Microscope Image Processing*, Academic Press.
- Marzban, C. (2004). The ROC Curve And The Area Under it As Performance Measures. *Weather and Forecasting*, 19(6), 1106-1114.
- Mathew, T. V. (2012). Genetic Algorithm. *Report submitted at IIT Bombay*, 53.
- Mir, U., Kar, A. K., and Gupta, M. P. (2022). AI-Enabled Digital Identity—Inputs for Stakeholders and Policymakers. *Journal of Science and Technology Policy Management*, 13(3), 514-541.

- Mirjalili, S. (2019). Genetic Algorithm. *Evolutionary Algorithms and Neural Networks: Theory and Applications*, 43-55.
- Moran, S., McDonagh, S., and Slabaugh, G. (2021, January). Curl: Neural Curve Layers for Global Image Enhancement. *In 2020 25th International Conference on Pattern Recognition (ICPR)*, 9796-9803.
- Murphy-Chutorian, E., and Trivedi, M. M. (2008). Head Pose Estimation in Computer Vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(4), 607-626.
- Najafzadeh, N., Kashiani, H., Saadabadi, M. S. E., Talemi, N. A., Malakshan, S. R., and Nasrabadi, N. M. (2023). Face Image Quality Vector Assessment for Biometrics Applications. *In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 511-520.
- Noura, H. N., Azar, J., Salman, O., Couturier, R., and Mazouzi, K. (2023). A Deep Learning Scheme for Efficient Multimedia Iot Data Compression. *Ad Hoc Networks*, 138, 102998.
- Okumura, A., Komeiji, S., Sakaguchi, M., Tabuchi, M., and Hattori, H. (2019). Identity Verification Using Face Recognition for Artificial-Intelligence Electronic Forms with Speech Interaction. In *HCI for Cybersecurity, Privacy and Trust: First International Conference, HCI-CPT 2019, Held as Part of the 21st HCI International Conference*, Springer International Publishing, (Baskıda).
- Padilla, R., Netto, S. L., and Da Silva, E. A. (2020). A Survey on Performance Metrics for Object-Detection Algorithms. *In 2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 237-242.
- Parate, S., Josyula, H. P., and Reddi, L. T. (2023). Digital Identity Verification: Transforming KYC Processes in Banking Through Advanced Technology and Enhanced Security Measures. *International Research Journal of Modernization in Engineering Technology and Science*, 5(9), 128-137.
- RangeKing. (2023). YOLOv8. Erişim adresi: <https://github.com/open-mmlab/mmyolo/blob/dev/configs/yolov8/README.md/>
- Rusli, F. M., Adhiguna, K. A., and Irawan, H. (2021). Indonesian ID Card Extractor Using Optical Character Recognition and Natural Language Post-Processing. *In 2021 9th International Conference on Information and Communication Technology (ICoICT)*. 621-626.
- Saray, T., Çetinkaya, A., ve Okatan, A. (2017). Optik Karakter Tanıma Yöntemi ile Otomatik Tabela Okuyucu. *In 2017 International Conference on Computer Science and Engineering (UBMK)*, 1061-1063. IEEE.
- Satyawan, W., Pratama, M. O., Jannati, R., Muhammad, G., Fajar, B., Hamzah, H., Rusnandi F., and Kristian, K. (2019). Citizen ID Card Detection Using Image

- Processing and Optical Character Recognition. *In Journal of Physics: Conference Series*, 1235(1), 012049. IOP Publishing.
- Serengil, S. I., and Ozpinar, A. (2020). Lightface: A Hybrid Deep Face Recognition Framework. *In 2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*, 1-5.
- Serpa, Y. R., Nogueira, M. B., Neto, P. P. M., and Rodrigues, M. A. F. (2020). Evaluating Pose Estimation as a Solution to the Fall Detection Problem. *In 2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH)*, 1-7.
- Shaffi, N., and Hajamohideen, F. (2021). uTHCD: A New Benchmarking for Tamil Handwritten OCR. *IEEE Access*, 9, 101469-101493.
- Shieber, S. M. (2004). The Turing test: Verbal Behavior as the Hallmark of Intelligence. *Mit Press*, (Baskıda).
- Silva, I., do Valle, J. M., Souza, G., Budke, J., Araújo, D., Carvalho, B., Cacho, N., Sales, H., Lopes, F., and Júnior, R. S. (2021). Using Micro-Services and Artificial Intelligence to Analyze Images in Criminal Evidences. *Forensic Science International: Digital Investigation*, 37, 301197.
- Smith, R. (2007). An Overview of the Tesseract OCR Engine. *In Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)* 2, 629-633.
- Suwanda, R., Syahputra, Z., and Zamzami, E. M. (2020). Analysis of Euclidean Distance and Manhattan Distance in the K-Means Algorithm for Variations Number of Centroid K. *In Journal of Physics: Conference Series* 1566(1), 012058. IOP Publishing.
- Şenalp, F. M. (2023). Tek Görüntü Süper Çözünürlük Uygulamaları için Görsel Kaliteyi İyileştirmeye Yönelik Yeni Bir Yaklaşım. *Journal of Intelligent Systems: Theory and Applications*, 6(1), 52-58.
- Tamang, S., Sen, B., Pradhan, A., Sharma, K., and Singh, V. K. (2023). Enhancing Covid-19 Safety: Exploring YOLOv8 Object Detection for Accurate Face Mask Classification. *International Journal of Intelligent Systems and Applications in Engineering*, 11(2), 892-897.
- Tamirat, C. (2023). *Customers Identity Card Data Detection and Recognition Using Image Processing*, (Doctoral Dissertation). St. Mary's University, Addis Ababa, Ethiopia.
- Tan, N. T. T., Van Huy, H., Kim, D. H., and Ngoc, L. A. (2023). An Intelligent Threats Solution for Object Detection and Resource Perspective Rectification of Distorted Anomaly Identification Card Images in Cloud Environments. *Applied Intelligence*, 53(1), 385-404.

- Tareen, S. A. K., and Raza, R. H. (2023). Potential of SIFT, SURF, KAZE, AKAZE, ORB, BRISK, AGAST, and 7 More Algorithms for Matching Extremely Variant Image Pairs. *4th International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, 1-6.
- Terven, J., and Cordova-Esparza, D. (2023). A Comprehensive Review of YOLO: From YOLOv1 to YOLOv8 and Beyond. *arXiv preprint arXiv:2304.00501*.
- Timofte, R., Agustsson, E., Van Gool, L., Yang, M. H., and Zhang, L. (2017). Ntire 2017 Challenge on Single Image Super-Resolution: Methods and Results. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 114-125.
- Ullmann, J. R. (2019). Advances in Character Recognition. In *Applications of Pattern Recognition*, CRC Press, (Baskıda).
- Van Hoai, D. P., Duong, H. T., and Hoang, V. T. (2021). Text Recognition for Vietnamese Identity Card Based on Deep Features Network. *International Journal on Document Analysis and Recognition (IJDAR)*, 24, 123-131.
- Varlık A., ve Çorumluoğlu, Ö. (2011). Dijital Fotogrametri Teknikleri ile Kişi Tanıma. *Harita Teknolojileri Elektronik Dergisi*, 3(2), 1-24.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is All You Need. *Advances in Neural Information Processing Systems*, 30.
- Velepucha, V., and Flores, P. (2023). A Survey on Microservices Architecture: Principles, Patterns and Migration Challenges. *IEEE Access* 2023, 11, 88339–88358.
- Vijaymeena, M. K. and Kavitha, K. (2016). A Survey on Similarity Measures in Text Mining. *Machine Learning and Applications: An International Journal*, 3(2), 19-28.
- Whitby, B. (2009). *Artificial Intelligence*. The Rosen Publishing Group Inc, (Baskıda).
- Wu, J., and Trivedi, M. M. (2008). A Two-Stage Head Pose Estimation Framework and Evaluation. *Pattern Recognition*, 41(3), 1138-1158.
- Yakut, C., ve Ulukaya, S. (2022). Görüntü Filtreleme ile Denetimsiz Retina Damar Bölütlenmesi için Parametre En iyileştirilmesi. *Mühendislik Bilimleri ve Tasarım Dergisi*, 10(3), 844-855.
- Yıldız, G., ve Duru, N. (2023). İki Boyutlu (2D) Arkeolojik Parçalanmış Nesnelerin Görüntü İşleme ve Geometrik Metotları Kullanarak Birleştirilmesi. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, 16(1), 53-59.

Yue, L., Shen, H., Li, J., Yuan, Q., Zhang, H., and Zhang, L. (2016). Image Super-Resolution: The Techniques, Applications, and Future. *Signal Processing*, 128, 389-408.

Zhang, C., and Lu, Y. (2021). Study on Artificial Intelligence: The state of the Art and Future Prospects. *Journal of Industrial Information Integration*, 23, 100224.

Zhang, Y., Tian, Y., Kong, Y., Zhong, B., and Fu, Y. (2018). Residual Dense Network for Image Super-Resolution. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2472-2481.

Zou, J., Han, Y., and So, S. S. (2009). Overview of Artificial Neural Networks. *Artificial Neural Networks: Methods and Applications*, 14-2.

