

T.C.
BEYKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ BİLİM DALI

**OPTİK KARAKTER TANIMA TABANLI OTOBÜS
SINIFLANDIRMA UYGULAMASI**

Yüksek Lisans Tezi

Tezi Hazırlayan:

Alican TÜRKER

İSTANBUL, 2014

T.C.
BEYKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ BİLİM DALI

**OPTİK KARAKTER TANIMA TABANLI OTOBÜS
SINIFLANDIRMA UYGULAMASI**

Yüksek Lisans Tezi

Tezi Hazırlayan:

Alican TÜRKER

Öğrenci No:

120820018

Danışman:

Yrd. Doç. Dr. Turhan KARAGÜLER

İstanbul, 2014

YEMİN METNİ

Yüksek Lisans Tezi olarak sunduğum “Optik Karakter Tanıma Tabanlı Otobüs Sınıflandırma Uygulaması” başlıklı bu çalışmanın, bilimsel ahlak ve geleneklere uygun şekilde tarafımdan yazıldığını, yararlandığım eserlerin tamamının kaynaklarda gösterildiğini ve çalışmamın içinde kullandıkları her yerde bunlara atıf yapıldığını belirtir ve bunu onurumla doğrularım. 04.09.2014

Alican TÜRKER

İmza:

T.C.
BEYKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

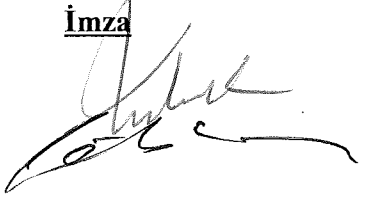

YÜKSEK LİSANS TEZ SAVUNMA SINAVI SONUÇ TUTANAĞI

Beykent Üniversitesi
Fen Bilimleri Enstitüsü Müdürlüğü'ne,

Aşağıda tez adı belirtilen yüksek lisans öğrencisi 120820018 no'lu Alican TÜRKER'in 04/09/2014 tarihinde yapılan tez savunma sınavı¹ sonucunda 45 dakika süreyle sunduğu ve savunduğu tezi hakkında² oybirliğiyle, kabul kararı verilmiştir.

Bilgilerinize saygılarımızla arz ederiz.

Anabilim Dalı : Bilgisayar Mühendisliği
Programı : Bilgisayar Mühendisliği
Proje Konusu³ : Optik Karakter Tanıma Tabanlı Otobüs Sınıflandırma Uygulaması

<u>Tez Sınav Jürisi</u>	<u>Öğretim Üyesi</u>	<u>İmza</u>
Danışman :	Yrd. Doç. Dr. Turhan KARAGÜLER	
Üye :	Doç. Dr. Gökhan SİLAHTAROĞLU	
Üye :	Doç. Dr. Kazım SARI	

¹ Jüri üyeleri söz konusu tezin kendilerine teslim edildiği tarihten itibaren en geç bir ay içinde toplanarak öğrenciyi tez savunma sınavına alır. Belirlenen günde yapılamayan jüri toplantısı, katılanların hazırladığı bir tutanakla enstitü yönetimine bildirilir. Bu durumda jüri en geç onbeş gün içinde toplanarak adayı tez savunma sınavına alır. Tez savunma sınav süresi en az 45 dakikadır. Yüksek lisans tez savunma sınavı, tez çalışmasının sunulması ve bunu izleyen soru-yanıt bölümlerinden oluşur ve dinleyiciye açıktır. (Beykent Lisansüstü eğitim ve Öğretim Yönetmeliği-Madde30-3)

²Tez sınavının tamamlanmasından sonra jüri, tez hakkında “kabul”, “düzeltme” veya “red” kararı verir. Jüri başkanı, jüri üyelerince imzalanmış sınav tutanağını, tez sınavını izleyen üç gün içinde ilgili enstitü yönetimine teslim eder. Tezi başarısız bulunan öğrencinin Enstitü ile ilişkisi kesilir. Tezi hakkında düzeltme kararı verilen öğrenci en geç üç ay içinde gerekli düzeltmeleri yaparak ve yönetmelikte belirtilen usullere uygun olarak tezini aynı jüri önünde yeniden savunur. Bu savunma sınavında da tezi kabul edilmeyen öğrencinin enstitü ile ilişkisi kesilir.(Beykent Lisansüstü eğitim ve Öğretim Yönetmeliği-Madde30-4)

³İleridedoğabilecek aksaklıkların engellenmesi için tezin başlığının yazılması gerekmektedir.

Adı ve Soyadı : Alican Türker
Danışmanı : Turhan Karagüler
Türü ve Tarihi : Yüksek Lisans, 2014
Alanı : Bilgisayar Mühendisliği
Anahtar Kelimeler : Görsel makine öğrenimi, nesne tanıma, geometrik metotlar, erişilebilir yazılım

ÖZET

OPTİK KARAKTER TANIMA TABANLI OTOBÜS SINIFLANDIRMA UYGULAMASI

Bu projede, akıllı telefonlar üzerinden bir otobüsün güzergâh numarasının otomatik olarak okunması için makine öğrenimi ve bilgisayar görüşü algoritmaları bir araya getirilmiştir. Bu çalışma görme engelli insanların şehir senaryolarına adaptasyonunu arttırmak için çok yararlı olabilir. Standart görüntü eşleme metotları yansıtıcıların bolluğu, görüntülerin üst üste gelmesi, aydınlık seviyesinin değişmesi, parlak noktaların oluşması, gölge oluşumu ve bakış açısı değişikliği gibi durumlarda hatalarla karşı karşıya kalabilmektedir. Bu sorun “cascade of classifiers” algoritması ile otobüsün ön yüzünün belirlenmesi ve elde edilen ön yüzde geometrik varlıkların yerlerinin bir model yardımıyla belirlenmesi, ondan sonra bu varlıklar üzerinden görüntü eşleme yapılmasıyla çözülür. Bu metot gerçek zamanlı olarak çalışır ve deney sonuçlarının da gösterdiği gibi tanıma oranı ve güvenilirlik açısından iyi bir performansa sahiptir.

Name and Surname : Alican Türker
Supervisor : Turhan Karagüler
Degree and Date : Master, 2014
Major : Computer Engineer
Key Words : Visual machine learning, object recognition,
geometric methods, accessibility software

ABSTRACT

BUS CLASSIFICATION APPLICATION BASED ON OPTICAL CHARACTER RECOGNITION

In this paper, machine learning and geometric computer vision are combined for reading bus line numbers automatically with a smartphone. This can prove very useful to improve the autonomy of visually impaired people in urban scenarios. The problem is a challenging one, since standard geometric image matching methods fail due to the abundance of distractors, occlusions, illumination changes, highlights and specularities, shadows, and perspective distortions. The problem is solved by locating the main geometric entities of the bus facade through a cascade of classifiers, and then refining the matching with robust geometric matching. The method works in real time and, as experimental results show, has a good performance in terms of recognition rate and reliability.

İÇİNDEKİLER

ÖZET	4
ABSTRACT	5
ŞEKİLLER LİSTESİ	iv
TABLolar LİSTESİ	v
KISALTMALAR VE SEMBOLLER LİSTESİ	vi
1. GİRİŞ	1
1.1. Motivasyon	1
1.1.2. Özürlülere Yönelik Yardımcı Teknoloji Türleri.....	4
1.1.3. Alternatif Haberleşme, Dinleme ve Görme Cihazları	4
1.2. Literatür İncelemesi.....	4
1.2.1. Otomatik Renk Zıtlığı Arttırma.....	4
1.2.2. Videotanim Yazımı (Videodescription Scripting)	7
1.2.3. Gerçek Zaman Nesne Tespiti.....	8
1.2.4. Kuvvetli Gerçek Zaman Yüz Tespiti	11
1.3. Tezin Anahatları.....	13
2. PROJEDE KULLANILAN TEMEL YAPILAR.....	14
2.1. Basamaklı Sınıflandırıcı (Cascading Classifiers)	14
2.1.1. Özellikler.....	14
2.1.1.1. Integral Görüntüsü.....	16
2.1.1.2. Özellik Tartışması	18
2.1.2. Öğrenme Sınıflandırılması İşlevi.....	19
2.1.2.1. Öğrenme Tartışmaları.....	21
2.2. OpenCV (Open Source Computer Vision Library)	23
2.3. OCR (Optical Character Recognition)	23

2.3.1. Matris Eşleme (Matrix Matching)	23
2.3.2. İçerik Analizi (Feature Analysis).....	24
2.3.3. Özgün Tanımlama (Self-Assertion)	24
2.3.4. Verimliliği Arttırmak İçin Kullanılan Yöntemler	24
2.3.4.1. Lineart (Siyah beyaz).....	25
2.3.4.2. Grayscale (Gri tonlama).....	26
3. PROJENİN ÇALIŞMA YAPISI	27
3.1. Otobüsün Tanımlanması ve Lokalizasyonu.....	28
3.2. Hat Numarası Tanımlanması ve Lokalizasyonu	28
3.3. Hat Numarasının Okuması	30
4. DENEY VE ANALİZ SONUÇLARI	32
4.1. Otobüsü Tanımlama ve Lokalize Edilmesi.....	32
4.2. Hat Numarasının Okuması	35
4.2.1. Şablon Uyuşmazlığı.....	36
4.2.2. Hat Numarasının Elde Edilememesi	36
4.2.3. Okuma Sonucu Çıkan Fazla Karakterler.....	37
5. SONUÇ	39
KAYNAKÇA	42
EKLER	44
ÖZGEÇMİŞ.....	45

ŞEKİLLER LİSTESİ

Şekil 1. Renkli fotoğraflar için histogram eşitleme.	6
Şekil 2. Metodumuzun üst düzey görüntüsü.	6
Şekil 3. Tespit penceresi örneği.....	15
Şekil 4. İntegral görüntüsü değer hesaplama.....	17
Şekil 5. Piksel toplamı hesaplama	17
Şekil 6. Projenin çalışma yapısı şeması	27
Şekil 7. Üzerlerine şablon oturtulmuş otobüs resimleri	29
Şekil 8. İşlenmemiş resim üzerinden otobüs hat numarası kesitleri	30
Şekil 9. OCR temel çiftleme metodu sonucu oluşan resimler	30
Şekil 10. Hat numaralarının renk tonuyla filtreleme yapılmış resimler	30
Şekil 11. Sırasıyla farklı eşik değeri denemeleri $\tau = 20$, $\tau = 50$, $\tau = 70$ ve $\tau = 90$	31
Şekil 12. Şehir ortamında çekilmiş otobüs ön yüz resimleri	34
Şekil 13. Kameranın konumu için uygun açı değeri.....	35
Şekil 14. Uygulamanın tanımlaması sonucu oluşan resimler.....	35
Şekil 15. Telefon kamerasından çekilmiş dijital kamera görüntüleri	36
Şekil 16. Otobüs üzerinden alınmış numara kesiti.....	37
Şekil 17. Özel olarak yapılmış çiftleme metodu sonucu oluşan resimler	37
Şekil 18. OCR kütüphanesinin temel çiftleme metodu sonucu oluşan resimler.....	37

TABLÖLAR LİSTESİ

Tablo 1. İstanbul Otobüs A.Ş. 2014 otobüs filo bilgileri ve tanıma sonuçları.....	33
Tablo 2. İETT 2014 otobüs filo bilgileri ve tanıma sonuçları.....	33

KISALTMALAR VE SEMBOLLER LİSTESİ

Apk	: Application package file
CV	: Computer Vision
Doc	: Document (Belge)
DVD	: Digital Versatile Disc (Çok Amaçlı Sayısal Disk)
MLL	: Machine Learning Library
OpenCV	: Open Source Computer Vision Library
Pdf	: Portable Document Format (Taşınabilir Belge Biçimi)
Pps	: Power Point Script
OCR	: Optical Character Recognition
Xml	: eXtensible Markup Language
GÖ	: Görme Özürlü
İGS	: İnsan Görme Sistemi
VT	: Videotanınım
İETT	: İstanbul Elektrik Tramvay ve Tünel İşletmeleri
I(x)	: Mekânsal koordinasyonu x olan görüntü ögesi
L(x)	: Mekânsal koordinasyonu x olan görüntünün aydınlatma ögesi
R(x)	: Mekânsal koordinasyonu x olan görüntünün yansıtma ögesi
R	: Reel Kısım

1. GİRİŞ

1.1. Motivasyon

Görme yeteneđi insanlara verilen en büyük yeteneklerden biridir. Sađlıklı insanlar genellikle günlük yařantılarında bunun farkında olmasalar da, görme özrü olan insanların bunun yokluđunu fazlaca hissettikleri bir gerçektir. Çünkü günlük yařantılarını sürdürebilmek için büyük ölçüde başkalarının yardımına ihtiyaç duymaktadırlar. Dünya sađlık örgütünün tahminine göre, Dünyada yaklaşık 180 milyon görme özürlü insan bulunmaktadır (Nagarajan ve diđerleri, 2004). Bunların 40-45 milyonunun tamamen göremediđi tahmin edilmekte ve 2020 yılına kadar bu sayının ikiye katlanması beklenmektedir.

Çevreyi algılama ve tanıma bütün bireyler için çok önemlidir. Görme özürlülerin en temel ihtiyaçları; çevreyi tanıma, bilgiye erişim ve bađımsız olarak hareket edebilme yeteneđidir. Görme özürlü (GÖ) için bir mağazada veya benzeri yerlerde bir araçla gezinmenin yanında çevredeki işaretleri, önemli mekânları, sokakları tanımak da önemlidir. Bundan dolayı çevreden faydalanma, engellerden sakınmanın yanında, çevre özelliklerinin anlaşılması, tanınması ve yorumlanmasını da kapsamaktadır. Yakın çevre ile her türlü etkileşimde başkalarına bađımlı olmayı azaltan, taşınabilen bir cihazın yaşam kalitesini arttıracakđı açıktır.

Çevremizdeki objelerin görme engelli kişilerin farkına varabilmesi, en azından yerlerini tespit edebilmeleri için şimdiye kadar birçok çalışma vardır. Tüm çalışmalar, görme hücreleri yetersiz çalışan ya da hiç çalışmayan bireylerin, görme hücrelerinin işlevini devralacak ve beyne görme hücresi gibi hizmet edecek yeni algılama yöntemleri çerçevesinde gelişmektedir. En son üzerinde durulan ve iyi sonuçlar vereceđi düşünölen konu ise çevremizdeki görüntüyü sese çevirerek görme engelli kişilere çevrelerindeki algılatmak, bilgi vermek için yapılan çalışmalardır.

Kentsel fonksiyonlara erişimi sađlayan en önemli unsur olan ulaşım ile ilgili detaylı olarak çalışılmalıdır. Ulaşılamayan faaliyetlerin olması engelli ya da engelsiz bireylere fayda sađlamayacaktır. Engellilerin toplumsal katılımını geliřtirmek için önce ulaşımı düzenlememiz gerekmektedir.

Ulaşımında yol, trafik, otopark ve yaya sirkülasyon sisteminin bütünlük içinde, etkin biçimde hizmet etmesi sağlanmalıdır. Erişilebilirlik, otobüsten durağa, yaya yoluna, otoparka, konuta kadar uzanmalıdır. Bu konuda toplu taşıma servislerinin engelli insanlara da ulaşım olanağı sağlamak üzere genişletilmesi öncelikler arasında yer almalıdır.

Ulaşılabilirlik yapısal çevrenin temel bir özelliği olup, özürhüelerin günlük yaşamlarının sürdürebilmelerinde, öncelikle evlerinden çıkabilmeleri ve başkalarına muhtaç bütün ticari ve idari kamu binalarındaki işlerine gidip gelebilmeleri, sosyal-kültürel, spor, sağlık, dini faaliyetler gibi aktivitelere ulaşılabilirliğin sağlanması için gerekli düzenlemeler yapılmalıdır.

1.1.1. Özürhülere Yönelik Teknolojik Uygulamaların Önemi

Teknoloji alanındaki gelişmelerin bütün topluma sağladığı bağımsızlık ve yaşam kalitesindeki artışın özürhü insanlar içinde sağlanması mümkündür. Hatta uygulanabilmesi durumunda özürhü bireylere daha çok katkı sağlayacağı açık bir durumdur. Bu çerçevede modern bilgi ve iletişim teknolojileri, özürhü insanların toplumla bütünleşmesini sağlayacak bir potansiyele sahiptir.

Mevcut durumda ise teknolojik uygulamalar bu katkıdan oldukça uzaktır. Bu durum özellikle özürhü insanların teknolojik uygulamalardan soyutlanması anlamına gelmektedir. Yeni teknolojilerin gittikçe küçülmesi ve karmaşık hale gelmesi ise bu kullanımı daha da zorlaştırmaktadır. Sağlıklı insanların bile kullanmakta zorlandıkları bu yeni teknolojiler özürhü insanlar için adeta bir çileye dönüşmektedir.

Son dönemlerde yönetimlerin yasal düzenlemeler yoluyla ve sivil toplum örgütlerinin de etkin toplumsal baskı ve güçleriyle, teknoloji üreticileri özürhü grupların pazar içerisindeki önemlerini anlayarak bu doğrultuda çalışmalar gerçekleştirilmeye başlamıştır.

Artık günümüzde hem özürhü hem de yaşlıların toplumun diğer kesimleri gibi benzer standartlara sahip olma hakkı olduğu bilinci gittikçe gelişmektedir. “Özürhü ve yaşlı insanlar için yapılan iyi bir tasarım herkes için yapılmış iyi bir tasarımdır” biçiminde formüle edilen görüş teknoloji üreticileri ve servis sağlayıcıları açısından

da hızla kabul görmeye başlayan bir görüştür. Bu zihniyet deęişiklięinin tek nedeni elbette hükümetlerin yasal düzenlemeleri veya servis sağlayıcıları ile teknoloji üreticilerinin insafından kaynaklanmamaktadır. Sayısal olarak büyük bir yekûn tutan yaşlı ve özürlü kesimlerin pazar payı da bu zihniyet deęişiminde önemli bir nedendir.

Ergonomik bir tasarımda amaç; araç ve gereci, kullanıcının yetenek ve kapasitesini en üst düzeye çıkaracak, vücudun ilgili organ ve özelliklerine uygun olacak ve oluşabilecek kaza ve hata sayısını minimuma indirebilecek şekilde tasarlamaktır. Böylece daha güvenli, emniyetli, rahat ve etkili olarak işlerin yürütülmesi sağlanacaktır.

Herkes için tasarım ile aynı sistem içinde farklı kullanıcılar için farklı tasarım ve herkese uygun olarak ayarlanabilirlik ilkesi temel alınarak tasarımda farklılık ve esneklik yaratılmaktadır. Bu tasarımda hedef olarak sınırlı yeteneklere sahip olan (yaşlı, özürlü gibi) bireylerin ihtiyaç ve istekleri esas alınmaktadır. Böylece dięer nüfus grubundaki insanların bu araçları çok daha rahat olarak kullanacakları vurgulanmaktadır.

Örneęin, telefon günlük hayatımızın temel bir aracıdır. Hemen her evde ve cebimizde bulunan telefon pek çok insan için dış dünya ile kolayca iletişimimizi sağlamaktadır. Bu özellięi ile telefon özürlü ve yaşlı kişiler için daha da önemlidir. Fakat iyi tasarlanmamış ve kullanımı ergonomik olmayan bir telefon iletişimin önünde bir engel olabilmektedir. Çünkü pek çok özürlü ve yaşlı insan telefonları kullanmakta zorluklarla karşılaşmaktadır. Çünkü pek çok telefonun tuşları bu grup insanlar için ergonomik tasarlanmamıştır. Özel tasarım gerektiren telefonlar ise ya çok pahalıdır ya da kolay erişilebilir deęildir.

Bilgi ve iletişim teknolojilerinin günlük yaşamı kolaylaştırıcı yönü yanında, bu teknolojilerin özörlülerin mesleki ve genel eęitiminde de kullanılması son derece önemlidir.

1.1.2. Özürlülere Yönelik Yardımcı Teknoloji Türleri

Son birkaç yıldır daha da gelişen yardımcı teknolojiler ile ilgili olarak çeşitli özörlere sahip çocuklar ve yetişkinlere yardım eden yeni cihazlar ve yazılımlar geliştirilmesine devam edilmektedir. Yardımcı teknoloji cihazlar, özörlü insanların normalde yapamadığı işlerin gerçekleştirilmesine olanak sağlayan mekanik, elektrik ya da bilgisayarlı öğelerdir. Özörlü insanlara yönelik yardımcı teknolojilerin çeşitli türleri, çeşitli görevleri yerine getiremeyen insanlar için uyarlanabilir ve iyileştirici teknikler sağlamaktadır. Yardımcı teknoloji, birçok farklı özörlü için geniş bir uygulamaya cevap vermektedir.

1.1.3. Alternatif Haberleşme, Dinleme ve Görme Cihazları

Artırımı ve alternatif iletişim, sözlerle kendilerini ifade etme zorluğu içinde olan özörlüler için tahsis edilen bir yardımcı teknoloji türüdür. Microsoft'a göre, metin okuma cihazları genellikle özörlü bireyler tarafından kullanılmaktadır. Bu cihazlar, düşüncelerin sesli söylenmesi amacıyla bilgisayar destekli sesi kullanacak olan bir bilgisayara düşüncelerin yazılmasına olanak tanımaktadır. Bu cihazlar aynı zamanda bilgisayar ekranında görülebilir türü, sesli konuşurmaya dönüştürerek görme engellilerin duymasına yardımcı olabilmektedir.

1.2. Literatür İncelemesi

Bu proje otobüsleri tanımlayabilmek için görüntüyü analiz eden bir sınıflandırıcı ve hat numarasının görüntüsünü ön plana çıkarmak için görüntü işleme metotları kullanmaktadır. Bu gibi metotlar önceden denenmiş uygulanmış ve geliştirilmiş metotlardır ve örnekleri incelenmiştir.

1.2.1. Otomatik Renk Zıtlığı Arttırma

İnsan görme sistemi (İGS), detayın ve rengin tam olarak yansıtıldığı bir görüntüyü yakalamaya muktedir, karmaşık bir mekanizmadır. İGS'de, bireysel algılayıcılar belli bir düzeye kadar ayırt edebiliyorken, sistemin bütünü çeşitli biyolojik edimler yoluyla çok daha fazla dinamik erimle baş edebilmektedir. Güncel çekim veya görüntüleme aygıtları görüntünün bütün dinamik erimini güvenilir bir şekilde yansıtamamaktadır, böylelikle bir kameradan çekilen veya ekran yahut

görüntüleme cihazında görüntülenen fotoğrafların bazı kısıtlamalardan mustarip olmaktadır. Sonuç olarak, fotoğrafın parlak kısımları aşırı ışıklı; karanlık kısımları ise düşük ışıklı görünebilmektedir. Kontrast artışının amacı fotoğrafların görsel niteliğini arttırmaktır.

Fotoğrafın kontrastını arttırmak için kullanılan yaygın yöntemlerden bir tanesi histogram eşitlemeyi uygulamaktır. Bu tekniğin avantajı, gri tonlu fotoğraflara kolayca uygulanabilir olmasıdır ama bununla birlikte histogram eşitleme renkli fotoğraflarda kullanıldığında, **Şekil 1**'de gösterildiği üzere fotoğrafın renginde yapaylığa ve dengesizliğe yol açarak renk ölçeğinde bir kaymaya sebebiyet verebilmektedir. Bu istenmeyen yapaylıklar hoş değildir zira kalitelerini arttırırken fotoğrafların renk özelliklerini muhafaza etmek gereklidir.

Yaklaşımın üst düzey tanıtımı **Şekil 2**'de görülebilmektedir. Herhangi bir görüntünün aydınlatmanın (görüntüye düşen ışık) piksel piksel bölünmüş ürünü olduğunu ve görüntünün yansıtma ögesi olduğunu farz edilir. Bu, şu şekilde ifade edilebilir:

$$\mathbf{I}(\mathbf{x}) = \mathbf{L}(\mathbf{x}) \mathbf{R}(\mathbf{x}),$$

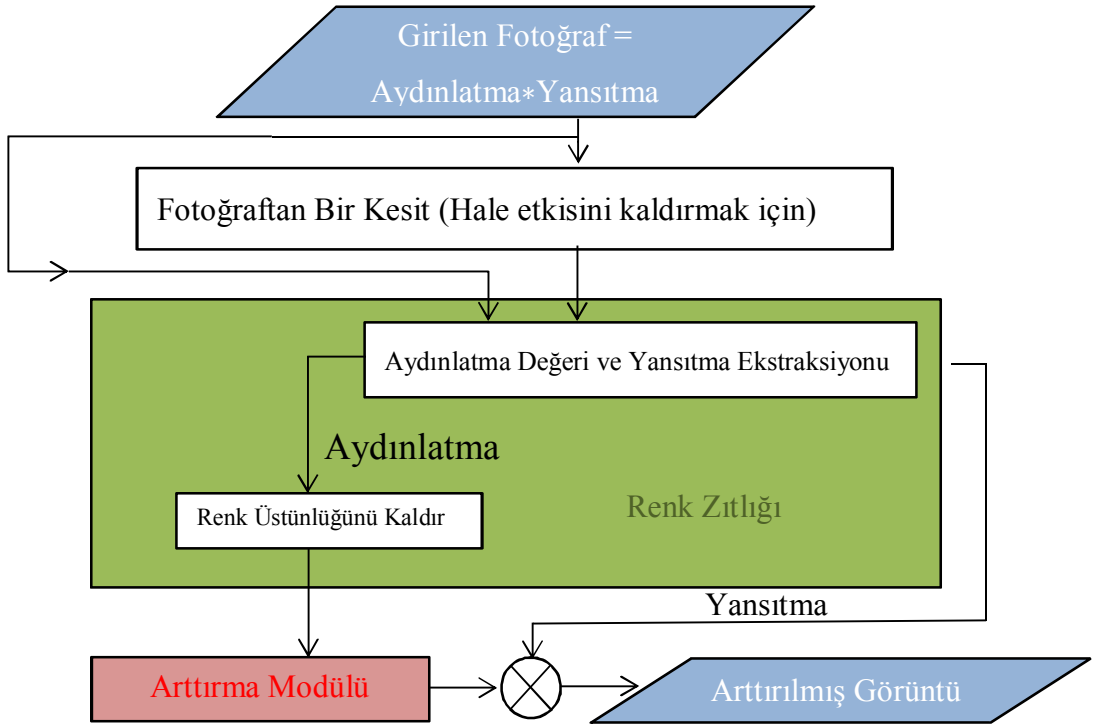
Burada, $\mathbf{L}(\mathbf{x})$ aydınlatma ögesidir ve $\mathbf{R}(\mathbf{x})$, mekânsal koordinasyonu \mathbf{x} olan $\mathbf{I}(\mathbf{x})$ görüntüsünün yansıtma ögesidir. Bu projede renkli görüntüleri ele almaktadır. Dolayısıyla, $\mathbf{I}(\mathbf{x})$, $\mathbf{L}(\mathbf{x})$ ve $\mathbf{R}(\mathbf{x})$ 'in her bir renk kanalı için bir tane olmak üzere 3 ögesi vardır. Mesela, aydınlatma görüntüsü $\mathbf{L}(\mathbf{x})$ için $\bar{L}_{\text{kırmızı}}(\mathbf{x})$ diyerek kırmızı rengi, $\bar{L}_{\text{yeşil}}(\mathbf{x})$ için yeşil rengi, $\bar{L}_{\text{mavi}}(\mathbf{x})$ içinse mavi rengi ifade edilmiştir. Benzer şekilde, renk kanallarını diğer görüntüler için de kullanabilir. Büyük koyu font çok bantlı anlamına gelmektedir. Büyük koyu fontun üzerindeki çizginin olması ise tek olduğunun göstergesidir. Benzer şekilde, mekânsal koordinasyon, $\mathbf{x} = (x, y) \in \mathbb{R}^2$ şeklindedir.



(a) Orijinal

(b) Histogram Eşitleme

Şekil 1. Renkli fotoğraflar için histogram eşitleme.



Şekil 2. Metodun üst düzey görüntüsü.

Öncelikle, görüntünün aydınlatma ögesi olan $L(x)$ 'ine değer vermeye çalışır ve böylelikle aydınlatmayı yansıtmadan ayırmış olur. $L(x)$ ve $R(x)$ 'i ayırma sebebi, Land [1], deneysel olarak bize gösterir ki İGS bir görüntüdeki rengi aydınlatmaya bakmaksızın ayırt edebilir. İGS'nin bu özelliği renk zıtlığı olarak adlandırılır ve teori

de Retinex Teorisi olarak bilinmektedir. Metot, aydınlatmayı değerlendirdikten sonra görüntüden aydınlatmanın etkilerini kaldırarak insan yetisinin taklidini yapmaya da çalışmaktadır. $L(x)$ 'i bir görüntüden kurtarmak ve $L(x)$ 'i $R(x)$ 'ten ayırmak matematiksel olarak kötü konumlanmış bir problemdir, dolayısıyla $L(x)$ 'i kurtarmak için başka varsayımlara ihtiyaç vardır. $L(x)$ 'in yumuşak olduğu varsayımı yapılmaktadır. Bu varsayıma dayanarak, $L(x)$ 'i değerlendirebilmek için görüntüyü algoritma kullanarak yumuşatılmaktadır. Bu bize görüntünün tam aydınlatmasını vermese de aydınlatma rengine dair iyi bir tahmin vermektedir. Ama bununla birlikte, yumuşatmanın sonucu olarak, görüntünün belli sınırları boyunca güçlü hale etkileri görülebilir. Böylelikle, $L(x)$ 'i tahmin etmek için yumuşatma yapmadan evvel, görüntüyü ön işleme sokulur ve kesitler alınır. Yumuşatma uygun bir şekilde icra edilebilir ve sınırlara kadar indirgenebilir. Renk üstünlüğü görüntüden kaldırıldığında sadece $L(x)$ ögesini işleme koyulur. Görüntünün sadece aydınlatma kısmında değişiklikler yapmanın ardındaki sebep görüntünün orijinal özelliklerini muhafaza etmektir ki bu görüntünün yansıtılmasıdır. Aynı zamanda, aydınlatmanın dinamik erimi hayli geniş olabilir ve böylece aydınlatmanın dinamik erimini sıkıştırmış olur. Piksellerin yoğunluğunun dağılımına dayanarak ve arttırılmış aydınlatmanın değerini görmek için logaritmik fonksiyonları kullanarak görüntünün aydınlatması üzerinde değişiklik yapılır ve sonrasında arttırılmış görüntüyü elde etmek için yansıtma ögesi ile çarpılır. Bu aynı zamanda insan görme sisteminin başka bir özelliği olan bölgesel görüntü zıtlığını arttırmada yardımcıdır. Sonuçlar göstermektedir ki karmaşık birleşmemiş ışıklandırma koşullarında dahi, arttırılmış sonuçlar görsel olarak daha iyi durmaktadır. [2]

1.2.2. Videotanın Yazımı (Videodescription Scripting)

Videotanın (VT) veya bilinen diğer ismiyle ses tanım veya tanımlanmış video, görme engellilere yardımcı olmak veya görsel olarak problem yaşayan insanlara ekranda ne olup bittiğini zihinlerinde canlandırmaları için müzik parçalarının sessiz kısımlarına eklenen önemli görsel elementlerin tanımıdır.

VT üretim süreci çoğunlukla manuel olarak yapılan gayet sıkıcı bir süreçtir. VT tanımının eklenebileceği ses parçalarını belirlemek adına filmi 1 veya 2 kez gözden geçirmek, tanımlanabilecek uygun görsel içeriği belirlemek ve sonrasında bir

VT metni yazmak, metnin yetkin bir ses tarafından (genellikle profesyonel bir aktör) okunmasını, senkronize edilmesini ve yeniden düzenlemesini sağlamak anlamına gelmektedir. Bu, filmler ve belgeseller için 25:1 civarında bir iş yükü, daha çok “artistik” veya “izlenimci” üretimler için (sadece müzik ve görsel elementleri içerenler [3]) 40:1 dolaylarında iş yükünü temsil eder. Buna rağmen, VT endüstrisinin büyümesi görme engelli toplumun istekleri ve ulusal yayıncılık organizasyonlarınca dayatılan düzenlemeler dolayısıyla artış göstermektedir.

VT üretim süreci film endeksleme ve özetleme görevi ile alakalıdır zira sahne geçişleri gibi hikâye akışlarını anlamak, aktörlerin kim olduğunu ve ne zaman göründüklerini, olayın nerede geçtiğini, metin bilgisinin ne dediğini, sessiz kısımların ses parçasında nerede olduğunu anlamakla alakalı görsel içerikleri belirlemeyi gerektirir.

Endeksleme ve özetleme değerlerini VT üretimine uygulamak gereklidir zira kullanım durumları farklıdır. İçerik temelli video kurtarımı için kullanıcı, özel bir görsel işitsel içeriğe sahip filmin özel bir kısmının arayışında olan tipik bir video arşivcisidir. Temelde, görsel içeriğin şifrlenmesinde bir kısıtlama yoktur (sanatın güncel durumu dâhilinde) çünkü video sorgulamasında hepsi potansiyel olarak ilgi çekicidir. Diğer taraftan, VT için tanımlanacak görsel içerik ses bandının içeriği tarafından kısıtlanır (bir aktör konuşurken veya aksiyon sahnelerinde olduğu gibi ses düzeyi yüksekken ekstra VT eklenemez). Aynı zamanda, sahnede ne olup bittiğini hayal etmeye yardımcı olmuyorsa görsel bilgiyi tanımlamaya gerek yoktur. Böylelikle otomatik bir VT yazım üreticisi, üretim sürecine uydurulmuş bilgisayar insan etkileşimiyle bütünleşmiş birbirini etkileyen video ve ses parçalarının birleşimini gerektirir.

Önceden yapılmış bu projede, otomatik ses denetleme, bilgisayar görüntüsü ve sentetik ses teknolojilerinin birleşimini kullanarak ilk bütünüyle işlevsel bilgisayar yardımcı VT yazılım aracı yapılmıştır. [4]

1.2.3. Gerçek Zaman Nesne Tespiti

Gerçek zaman nesne tespiti çoğu gerçek dünya görsel uygulamaları için önemlidir. Son çalışmalar göstermektedir ki, yüz [5,6,7,8] ve metin [9,10,11,12] gibi

belli başlı nesnelere güvenilir bir şekilde ve oldukça hızlı bir şekilde tespit edilebilir. Lakin tespit hızları her zaman uygulamalar için yeterince hızlı olmamaktadır. Bu projenin amacı nesne tespitini hızlandırabilecek bir teknik olan zaman açısından etkin yapıları sağlamaktır. Görme engellilere ve görsel açıdan yetersizlere yardım etmek amacıyla bilgisayar görsel sistemleri tasarlamakla ilgilenilmektedir. Özellikle, metin görsel açıdan yetersizler için geliştirilsin diye ve görme engellilere sesli bir şekilde okunsun diye şehir görüntülerinde metin tespit edecek algoritmalar geliştirilmelidir. Bu uygulamalar yüksek kalite performansı (yanlış pozitif ve yanlış negatif oranlarınca ölçülmüş) ve büyük görüntüler (1600x1200px mesela) yoluyla gerçek zaman araştırmalarını gerektirir. Önceki çalışmalar kabul edilebilir bir performans sergilemiştir ama hızı görsel olarak sorunlu kişiler tarafından değerlendirildiğinde yeterince hızlı değildir. Bu projede tanımlanan teknikler gerçek zaman tespitini elde etmeyi mümkün kılmaktadır. Tartışılır ama gerçek zaman nesne tespit sistemi geliştirmek için en etkili yol cascade olarak bilinen [13] hiyerarşik yapıyı kullanmaktır. Cascade bir yozlaşmış karar ağacıdır. Ağacın her seviyesi veya tabakası, görüntünün özellikleri üzerinde değerlendirilen bir sınıflandırıcı tarafından gerçekleştirilir.

Yüz tespiti [8] veya metin tespiti gibi nesne tespit işlerinde bu yapıların başarısı bir görüntünün çoğu parçasının merak konusu içermemesi gerçeğine bağlıdır. Bununla ilgili önemli nokta, görüntünün nesnelere içeren parçalarını muhafaza ederken nesnelere içermeyen kısımları da reddeden basit, hızla sayılabilen bir sınıflandırıcı kullanmaktır. Böylelikle daha karmaşık ve zaman alan sınıflandırıcılar görüntünün sadece belli kısımlarına uygulanmak durumunda olurlar.

Sezgisel olarak, yapı, bir heykeli oymak gibidir. Başlangıçta kaba bir şekil vermek amacıyla büyük kesimler yapılır. Sonrasında, detayları rötuş etmek daha belli başlı kesimler yapılır.

Etkili yapı tasarımı, sınıflandırıcıların zaman karmaşıklığı ve yanlış pozitif oranları gibi performans faktörlerinin değiş tokuşunu gerektirir. Ama kademeli diziler üzerinde yapılan önceki çalışmalar ve daha genel karar ağaçları sınıflandırıcıların zaman karmaşıklığını göz ardı etmişlerdir. Yapıların farklı düzeylerinde kullanılmış sınıflandırıcılar aynı zaman karmaşıklığındalarsa, bu

mantıklıdır. Ama her zaman olduğu üzere eğer sınıflandırıcılar farklı zaman karmaşıklıklarındaysa bu standardın altında olabilir.

Mesela, Viola ve Jones [8, 14] her yapı tabakası için maksimum kabul edilebilir yanlış pozitif oranları ve minimum kabul edilebilir tespit oranlarına dayanarak yüz tespiti için bir yapı tasarladılar. Zayıf sınıflandırıcılar benzer zaman karmaşıklıklarındalar ama güçlü sınıflandırıcılar değişken sayıda zayıf sınıflandırıcılardan meydana gelebiliyorlar ve farklı zaman karmaşıklığına sahipler.

Sınıflandırıcıların önceki çalışmalarında, benzer zaman karmaşıklığına sahip olmaları pek muhtemel değildir. Bu projede sunulan sınıflandırıcılar zayıf sınıflandırıcılardan AdaBoost tarafından da öğrenilmiştir. Fakat bu zayıf sınıflandırıcılar pek çok farklı görüntü özelliklerine ve farklı zaman karmaşıklıklarına dayanmaktadır. Bazı zayıf sınıflandırıcılar basit yapı özellikleri temellidir ve hızlıca sayılabilir. Diğer zayıf sınıflandırıcılar uçların arasındaki mekânsal ilişkiler gibi mekânsal özelliklere dayalıdır ve daha çok sayım gerektirirler.

Sınıflandırıcılar, Destek Vektör Makineleri [16] gibi diğer teknikleri içermesi için sınıflandırıcıların kelime haznelerini genişletirsek daha değişken zaman karmaşıklıklarına sahip olabilirler. Ötesi, Viola ve Jones [8] ve önceki çalışmalar gibi algoritmalar, zaman karmaşıklığı hesaplarına da dâhil edilmesi gereken ön işlem aşamalarını da içerir. Şu sonuca varabiliriz ki iyi tasarlanmış bir yapı zaman karmaşıklığını dikkate almalıdır.

Görüntünün algılanması için karar ağaçları tasarlamada Geman ve Blanchard tarafından benzer kanıtlar öne sürülmüştür [15]. Bu sorunun matematiksel olarak analizi yapılmıştır ve strateji geliştirmenin yararlılığı üzerinde durulmuştur. Ama yaklaşımları bu projedekinden farklıdır ve sonuçları çözmeye çalışılan sorunlara nasıl uygulanacağı da net değildir.

Bu projede, yanlış pozitif oranını (yanlış negatif oranının eğitme verilerinde sıfır olduğunu söylemelidir) aldığı kadar her yapı tabakasının ortalama işlem sürecini de dikkate alan zaman açısından etkin yapılar tasarlamada en uygun şekle sokma ölçütü verilmiştir. Yapı tasarımları bilinen çok terimli zaman algoritmalarıyla

çözülememektedir. Daha ziyade, ölçütlere sonuçlar bulan algoritma tasarlanmıştır. Bu yaklaşımı, zaman etkili sınıflandırıcıyı eski sınıflandırıcılarla karşılaştırabileceğimiz metin tespitinde denenmiştir. (Viola ve Jones'un sınıflandırıcı tasarımını kullanarak elde edilmiştir [8]).

Bu problem için, zaman etkili sınıflandırıcının oluşturulan önceki sınıflandırıcıya göre 2,5 kat daha hızlı işlediğini gösterilmektedir. Bu hızlanma metin tespiti algoritmasının gerçek hayatta çalışmasını ve böylece görsel olarak yetersiz olanlara yardım eden uygulamalar için yararlı olmaktadır. [17]

1.2.4. Kuvvetli Gerçek Zaman Yüz Tespiti

Bu proje kuvvetli ve oldukça hızlı görsel tespit için bir yapı inşa etmede yeni algoritmaları ve anlayışları bir araya getirmektedir. Sonuna doğru tespite ve bilinen en iyi sonuçlara eşdeğer yanlış pozitif orana ulaşan ön yüz tespit sistemi inşa edilmiştir. Bu yüz tespit sistemi önceki yaklaşımlardan yüzleri hızlı bir şekilde tespit etmesi özelliğiyle ayrılmaktadır. 384 – 288 piksel resimlerde işlenen yüzler sıradan 700 MHz Intel Pentium III'te saniye başına 15 çerçevede tespit edilmiştir. Diğer yüz tespit sistemlerinde, video dizinindeki görüntü farklılığı veya renkli görüntülerdeki piksel rengi gibi ikincil bilgiler yüksek çerçeve oranlarını elde etmek amacıyla kullanılmıştır. Bu sistem bunu tek bir gri tonlu görüntüde var olan bilgiyle çalışarak elde etmektedir. Bu alternatif bilgi kaynakları daha da yüksek oranları elde etmede sistemle bütünleştirilebilir.

Yüz tespiti yapısının üç genel katkısı bulunmaktadır. Bu projenin ilk katkısı gayet hızlı özellik değerlendirmesi sağlayan yeni görüntü temsidir. Kısmen Papageorgiou'ın (1998) çalışmalarına dayanan tespit sistemi direkt olarak görüntü yoğunlukları ile çalışmamaktadır. Çoğu gibi bu projede de Haar Basis işlevlerini andıran bir dizi özellik kullanılmıştır. Birden fazla ölçekte bu özellikleri hızlı bir şekilde sayabilmek amacıyla görüntünün integral görüntü temsilini tanıtılır. İntegral görüntü, yapı haritalama için bilgisayar grafiklerinde kullanılan toplam alan tablosuna benzerdir. İntegral görüntü, piksel başına birkaç işlem kullanarak bir görüntüden hesaplanabilmektedir. Hesaplandığında, Haar benzeri bu özellikler sabit sürede her ölçekte veya konumda hesaplanabilir.

Bu projenin ikinci katkısı, AdaBoost (Freund ve Schapire, 1995) kullanan çok sayıda potansiyel özelliklerden az sayıda önemli özellik seçilerek geliştirilen basit ve etkili sınıflandırıcıdır. Görüntü alt penceresinde toplam Haar benzeri özelliklerin sayısı oldukça fazladır ve piksel sayısından da oldukça çoktur. Hızlı sınıflandırmayı mümkün kılmak için, öğrenme süreci uygun özelliklerin pek çoğunu çıkarmalıdır ve önemli birtakım özelliklere yoğunlaşmalıdır. Tieu ve Viola'nın (2000) çalışmalarından yola çıkan özellik seçmeye tek bir özelliğe dayanması için her zayıf sınıflandırıcıyı sınırlandırmak suretiyle AdaBoost'u kullanarak ulaşılabilir. Sonuç olarak, yeni bir zayıf sınıflandırıcı seçen her bir hızlandırma süreci özellik seçme süreci olarak görülebilmektedir. AdaBoost etkili öğrenme algoritması ve genelleştirme performansı üzerinde güçlü bağlar sağlamaktadır. (Schapire, 1998).

Bu projenin üçüncü en önemli katkısı, görüntünün önemli bölgelerine yoğunlaşmak suretiyle algılayıcının hızını önemli ölçüde arttıran yapıdaki karmaşık sınıflandırıcıları art arda birleştirme yöntemidir. Yoğunlaşma yaklaşımının ardındaki fikir görüntünün neresinde bir yüz meydana gelebileceğini hızlı bir şekilde belirleyebilmenin mümkün olduğudur. Daha karmaşık işlemler sadece bu önemli bölgelerde tutulur. Böylesi bir yaklaşım için ana ölçüt, dikkat sürecinin “yanlış negatif” oranıdır. Yüz örneklerinin tamamı ya da tamamına yakın bir kısmının dikkat filtresi tarafından seçilmesi söz konusu olmalıdır.

Denetlenen ve dikkat yoğunlaştırma operatörü olarak kullanılan oldukça basit ve etkili sınıflandırıcıları çalıştırmak için bir süreç tanımlanmıştır. Yüzlerin %99'unu muhafaza ederken görüntünün %50'den fazlasını filtreleyen bir yüz tespit dikkat operatörü öğrenilebilmektedir. Bu filtre fazlasıyla etkilidir; lokasyon/ölçek başına 20 basit işlemde değerlendirilebilmektedir.

Başlangıçtaki sınıflandırıcılar tarafından reddedilmeyen bu alt pencereler, her biri bir öncekinden daha da karmaşık olan bir dizi sınıflandırıcı tarafından işlenir. Herhangi bir sınıflandırıcı alt pencereyi reddederse, başka işlem gerçekleşmemektedir. Tespit sürecinin yapısı yozlaşmış bir karar ağacı gibidir ve Fleuret ve Geman'ın (2001) ve Amit ve Geman'ın (1999) çalışmalarıyla bağlantılıdır.

Tam bir yüz tespit yapısının, 80,000 üzerinde işlemi olan 38 tane sınıflandırıcısı vardır. 507 yüz ve 75 milyon alt penceresi olan zor bir veri setinde, yüzler alt pencere başına ortalama 270 mikro işlem talimatı kullanarak tespit edilir. Kıyasla, bu sistem Rowley ve diğerleri (1998) tarafından geliştirilen tespit sisteminin uygulanmasından 15 kat daha fazla hızlıdır.

Oldukça hızlı bir yüz tespit algılayıcısının geniş uygulamaları olacaktır. Bunlar kullanıcı ara yüz, görüntü veri tabanı ve telekonferansı içerecektir. Hızdaki bu artış, öncesinde mümkün olmadıkları sistemler üzerindeki gerçek zaman yüz tespiti uygulamalarını mümkün kılacaktır. Resim karesi hızının çok olmasına gerek olmayan uygulamalarda, sistemin önemli ekstra işlem sonrasına ve analize izin verecektir. Buna ek olarak sistem, elle taşınır ve saklı işlemcileri içeren çok sayıda düşük gerilimli aletlere uygulanabilmektedir.

1.3. Tezin Ana Hatları

Bu bölümde görme engelli insanların yaşadıkları zorluklara ve bu projenin onlar için önemine değinilmiştir. Ayrıca projede yer alan aşamaların daha önceden nerelerde ve nasıl kullanıldığını öğrenmek amacıyla bir literatür taraması yer almaktadır. İkinci bölümde, projede kullanılmış yapılar gösterilmiş ve nasıl çalıştıkları hakkında bilgi verilmeye çalışılmıştır. Üçüncü bölümde, projenin işleyiş yapısı anlatılmış olup her adımda gerçekleşen işlemler anlatılmaya çalışılmıştır. Projenin temel yapısı hakkında bilgiler verilmiştir. Dördüncü bölüm, proje için yapılmış deney ve analizler hakkında bilgi vermektedir. Ayrıca deney sonuçları ve analizleri de yer almaktadır. Analiz sonucu bulunan eksikliklere ve çözümlerine değinilmektedir. Beşinci bölümde, genel bir sonuç çıkarılmış olup çıkarılan sonuç tartışılmıştır.

2. PROJEDE KULLANILAN TEMEL YAPILAR

2.1. Basamaklı Sınıflandırıcı (Cascading Classifiers)

Basamaklı sınıflandırıcı projenin görüntü içerisinde yer alan otobüsün ön yüzünün tanımlanmasında kullanılmaktadır. Proje için büyük öneme sahip bu algoritma, ilk olarak Paul VIOLA ve Michael J. JONES tarafından 2001 yılında bir yüz tanıma projesi üzerinde uygulanmıştır. Bu algoritmada üç anahtar kalıtım vardır. İlk kalıtım “integral görüntü (Integral Image)” olarak adlandırılan yeni bir resmin tanıtımındır. Bu resim detektörler için kullanılacak özellikleri sağlar. İkinci adım ise kolay ve verimli bir sınıflandırıcıdır. Bu sınıflandırıcı AdaBoost öğrenim algoritmasını kullanarak üretilmiştir. Sınıflandırıcı çok büyük potansiyel özellik setinden az sayıda bulunan eleştirel görsel özelliklerin seçilmesinde kullanılmıştır. Üçüncü kalıtım ise bir kademeli dizi (cascade) içinde sınıflandırıcıları kombinlemek için kullanılan bir metottur. Bu kademeli dizi resimde istenilen alanda hesaplamalar yapılırken istenilmeyen arka plan görüntüsünden kurtulmak için kullanılmaktadır. İlk olarak resim üzerinde özellikler olarak sayılabilecek piksellerden daha çok şey bulunmaktadır. Özellik tabanlı sistemler piksel tabanlı sistemlerden daha hızlıdır. Basamaklı sınıflandırıcıda kullanılan üç ana özellik vardır. İlki “iki-dikdörtgen özelliği (two-rectangle feature)”. Bu özellik eşit ölçülere sahip iki alanın içinde kalan piksellerin toplamının farkına dayanır. İkincisi “üç-dikdörtgen özelliği (three-rectangle feature)”. Bu özellik yan yana üç eşit alanın, dışarıda kalan alanların piksellerinin toplamının ortada kalan alanın piksel toplamından farkına dayanır. Üçüncü “dört-dikdörtgen özelliği (four-rectangle feature)”. Bu özellikte köşeleri kesişmiş ikili alanların toplamının farkına dayanır. Pozitif ve negatif resimler üzerinde az rastlanır özellikler aranır ve bir set oluşturulur ve bu set makine öğrenimi algoritmasında kullanılır [8].

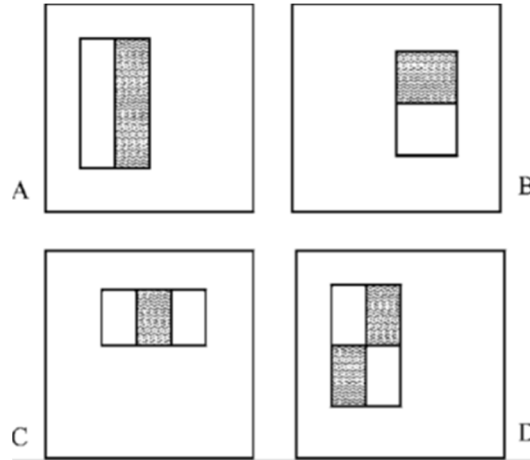
2.1.1. Özellikler

Yüz tespit prosedür sınıflandırıcısı basit özelliklerin değerine dayalıdır. Direkt olarak piksel kullanmaktansa özellikleri kullanmanın ardında pek çok sebep vardır. En yaygın olan sebeplerden bir tanesi, bu özelliklerin sınırlı sayıda eğitim verileri kullanarak öğrenmenin zor olduğu geçici alan bilgisini şifrelemede

kullanılabilir. Bu sistem söz konusu olduğunda özelliklerin kullanımı için ikincil önemli sebep vardır: özellik temelli sistem piksel temelli sistemden daha hızlı işlem yapmaktadır.

Papageorgiou ve diğerleri (1998) tarafından kullanılan Haar temelli fonksiyonları andıran basit özellikler kullanılır. Daha spesifik olarak, üç çeşit özellik kullanılmaktadır. İki dikdörtgen özelliğinin değeri, iki dikdörtgenel bölgedeki piksellerin toplamı arasındaki farktır. Bölgelerin boyları ve şekilleri aynıdır ve yatay veya dikey olarak bitişiktirler. (Şekil 3.) Üç dikdörtgen özelliği, merkez dikdörtgende toplamdan düşülen iki dış dikdörtgenin toplamını hesaplar. Sonuç olarak, dört dikdörtgen özelliği, diyagonal dikdörtgen çiftlerinin arasındaki farkları hesaplar.

Algılayıcının alt çözünürlüğü 24 x 24 olarak veriliyor, ayrıntılı dikdörtgen özelliği oldukça geniştir. Haar temellinin aksine, dikdörtgen özellik dizilerinin tamamlanmış olduğunu göz ardı edilmemelidir.



Şekil 3. Tespit penceresi örneği

Tespit penceresini çevrelemeye ilişkin gösterilen örnek dikdörtgen özelliği. Beyaz dikdörtgenlerde bulunan piksellerin toplamı gri dikdörtgenlerde piksellerin toplamından çıkarılmıştır. İki dikdörtgen özelliği (A) ve (B)'de gösterilmiştir. Şekil (C) üç dikdörtgen özelliğini ve (D) de dört dikdörtgen özelliğini göstermektedir.

2.1.1.1. Integral Görüntüsü

Dikdörtgen özellikleri, integral görüntü adını verdiğimiz görüntünün orta derecedeki temsilini kullanarak hızlı bir şekilde ölçülebilir. X, y konumundaki integral görüntü yukarıdaki piksellerin toplamını içerir ve x, y 'nin soluna doğru kapsar:

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'),$$

$ii(x, y)$ integral görüntüsüdür ve $i(x, y)$ orijinal görüntüdür (**Şekil 4**). Aşağıdaki tekrar çiftini kullanarak:

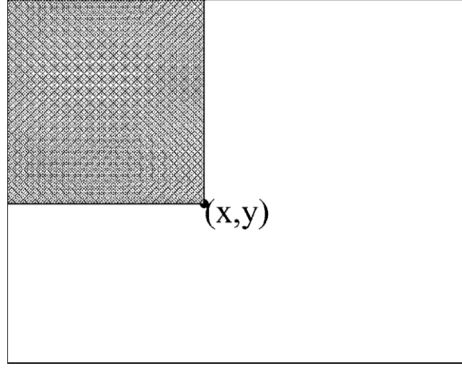
$$s(x, y) = s(x, y - 1) + i(x, y) \quad (1)$$

$$ii(x, y) = ii(x - 1, y) + s(x, y) \quad (2)$$

$(s(x, y)$ birikmiş satır toplamı, $s(x, -1) = 0$, ve $ii(-1, y) = 0$) integral görüntüsü, orijinal görüntünün üzerinden tek geçişte hesaplanabilir.

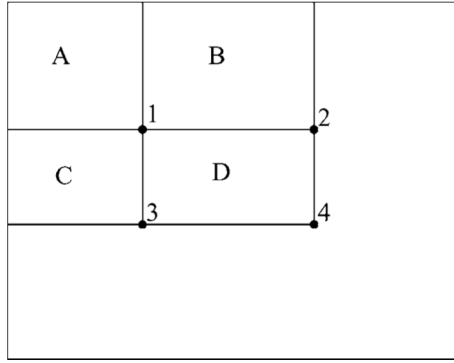
İntegral görüntüsünü kullanarak dikdörtgen toplamı, dört dizi referansta hesaplanabilir (**Şekil 5**). Açıkça, iki dikdörtgen toplamı arasındaki fark sekiz referansta hesaplanabilir. Yukarıda açıklanan iki dikdörtgen özelliği bitişik dikdörtgen toplamları içerdiğinden, altı dizi referansta; üç dikdörtgen özelliği durumunda sekiz ve dört dikdörtgen durumunda ise dokuz referansta hesaplanabilir.

İntegral görüntüsü için alternatif sebeplerden birisi Simard'ın "Boxlets" çalışmasından ileri gelmektedir.



Şekil 4. İntegral görüntüsü değer hesaplama

(x, y) noktalarındaki integral görüntüsünün değeri yukarıdaki ve soldaki piksellerin toplamıdır.



Şekil 5. Piksel toplamı hesaplama

D dikdörtgenindeki piksellerin toplamı dört dizi referansla hesaplanabilir. Konum 1'deki integral görüntüsünün değeri A dikdörtgenindeki piksellerin toplamıdır. Konum 2'deki değer A+B'de, konum 3'teki A+C'de ve konum 4'teki A+B+C+D'dedir. D'deki toplam $4+1 - (2+3)$ şeklinde hesaplanabilir.

Yazarlar belirtiyor ki, doğrusal işlem durumunda (örneğin $f \cdot g$), çevrilebilir doğrusal işlemler, aksi sonuca uygulandığı müddetçe f veya g 'ye uygulanabilir.

Sarım durumunda, türev operatörü hem görüntüye hem de çekirdeğe uygulanırsa sonuç iki katlı olmak durumundadır.

$$f * g = \int \int (f' * g').$$

Yazarlar, sarımın f ve g türevlerinin aralıklı olduğunda (ya da öyle yapıldıklarında) önemli ölçüde arttırılabileceğini göstermeye devam etmektedirler.

Benzer bir anlayış da şudur, çevrilebilir doğrusal işlem zıttı g' 'ye uygulandığında f' 'ye uygulanıyorsa:

$$(f'') * \left(\int \int g \right) = f * g.$$

Bu sistemde gösterilen dikdörtgen toplam hesabı üründe bir nokta olarak gösterilebilir

$i \cdot r$, i görüntüdür ve r de kapalı görüntüdür. Bu işlem tekrar yazılabilir:

$$i \cdot r = \left(\int \int i \right) \cdot r''.$$

İntegral görüntüsü aslında görüntünün iki katlı türevidir (ilki satırdaki ve ikincisi de sütundaki). Dikdörtgenin ikincil türevi (ilk satırlar boyunca sonra sütunlar boyunca) dikdörtgenin köşelerinde dört delta işlevi getirir. İkinci nokta ürününün değerlendirilmesine dört erişimde ulaşılır.

2.1.1.2. Özellik Tartışması

Dikdörtgen özellikleri, yönlendirilebilir filtrelerle karşılaştırıldığında bir bakıma ilkel kalmaktadır. Yönlendirilebilir filtreler ve benzerleri sınırların detaylı analizi, görüntü sıkıştırma ve metinsel inceleme açısından mükemmeldir. Dikdörtgen özellikler köşelere, çizgilere ve diğer basit görüntü yapılarına karşı hassas olmalarına rağmen kabadırlar. Yönlendirilebilir filtrelerden farklı olarak, uygun olanlar dikey, yatay ve diyagonal uygulamalardır. Ortogonalite bu özelliğin merkezinde olmadığından daha geniş ve farklı dikdörtgen özellikleri üretmeye çalışılmıştır. Tipik olarak, temsil 400 kere tamamlanmıştır. Bu tamamlama setleri, keyfi en boy oranı ve güzel bir şekilde örneklendirilmiş konum sağlamaktadır. Deneysel olarak, dikdörtgen özellik dizisi etkili öğrenmeyi destekleyecek zengin görüntü temsili sağlamaktadır.

Ayrıca da Dikdörtgen özelliğın sayısal etkinliğı kısıtlılıklarını telafi etme olanağı sağlamaktadır.

İntegral görüntü tekniğinin sayısal avantajını anlayabilmek için, görüntü piramitlerinin hesaplandığı daha geleneksel bir yaklaşım göz önünde bulundurulur. Çoğu yüz tespit sistemi gibi, bu sistem de girdiyi pek çok ölçüde tarar, yüzlerin 24 x 24 piksel boyutunda tespit edildiğı alt düzeyde başlar, 384'e 288 piksellik bir görüntü her biri bir öncekinden 1.25 daha büyük olan 12 ölçüde taranır. Geleneksel yaklaşım, her biri bir öncekinden 1,25 kat daha küçük olan 12 görüntüden oluşan piramidi hesaplamaktır.

Sabit ölçekli algılayıcı bu görüntülerin her biri boyunca taranır. Piramidin hesaplanması önemli zaman gerektirir. Geleneksel donanıma etkili bir şekilde uygulanabilir (piramidin her seviyesini ölçmek için çift doğrusal değerlendirme yapılır) ve bu büyüklükteki 12 seviyeli piramidi hesaplamak için 0,5 saniyeye ihtiyaç duymaktadır. (Intel PIII 700 MHz işlemcide)

Buna zıt olarak, anlamlı bir dizi dikdörtgen özelliğı tanımlamıştır bunun da birkaç işlemde herhangi ölçüde ve konumda tek bir özelliğın değerlendirilebilme olanağı vardır. Özelliklerin sayısal etkinlikleri, verilen yüz tespit süreci tüm bir görüntü için, 12 seviyeli görüntü piramidinin tek başına değerlendirilmesi için gerekli olan zamanla aynı zamanda saniye başı 15 çerçevede herhangi bir ölçekte tamamlanabilir. Bu tür bir piramit gerektiren herhangi bir prosedür oluşturulan yeni algılayıcıdan daha yavaş çalışacaktır.

2.1.2. Öğrenme Sınıflandırılması İşlevi

Özellik seti ve pozitif ile negatif görüntü eğitim seti verilirse, çok sayıda makine öğrenme yaklaşımları sınıflandırma yaklaşımını öğrenmede kullanılabilir. Sung ve Poggio Gaus modelinin bir karışımını kullanmışlardır. (Sung ve Poggio, 1998). Rowley ve diğerleri (1998) bir dizi basit görüntü özelliğı ve sınır ağı kullanmıştır. Osuna ve diğerleri (1997) destek vektör makinesi kullanmıştır. Daha yakın bir zamanda, Roth ve diğerleri (2000) yeni ve alışılmıřın dışında bir görüntü temsili ileri sürdü ve Winnow öğrenim prosedürünü kullanılmıştır.

Her görüntü alt penceresiyle alakalı 160,000 dikdörtgen özelliği olduğunu hatırlayın ki bu sayı da piksellerin sayısından fazladır. Her özellik etkili bir şekilde hesaplanırsa da bütün bir seti hesaplamak da olmaması gereken bir şekilde pahalıdır. Deneyler sonunda ortaya çıkan hipotez şudur ki bu özelliklerin çok az bir kısmı etkili bir sınıflandırıcı oluşturmak için birleştirilebilir. Asıl mesele bu özellikleri bulabilmektedir.

Sistem, özellikleri seçmek ve sınıflandırıcıları çalıştırmak için (Freund ve Schapire, 1995) AdaBoost değişkeni kullanılır. Orijinal yapısında, AdaBoost öğrenme algoritması basit bir öğrenme algoritmasının sınıflandırma performansını arttırmak amacıyla kullanılır. Bunu, daha güçlü bir sınıflandırıcı oluşturmak için bir takım zayıf sınıflandırıcıları birleştirerek yapar. Arttırma dilinde basit öğrenme algoritmasına zayıf öğrenici denir. Dolayısıyla, algısal öğrenme algoritması olası algıları araştırır ve en az sınıflandırma hatasıyla geri getirir. Öğrenene zayıf denir zira en iyi sınıflandırma işlevinden bile veriyi doğru düzgün sınıflandırmasını beklenmemektedir. (örneğin, verilen bir problemde en iyi algı veriyi %51 doğru olarak sınıflandırabilir). Zayıf öğreneni geliştirmek için, öğrenme problemlerini çözmeye çalışılmıştır. İlk öğrenme turundan sonra, önceki zayıf sınıflandırıcı tarafından yanlış şekilde sınıflandırılanları vurgulamak amacıyla örneklere yeniden göz atılmıştır. Nihai güçlü sınıflandırıcı bir eşik tarafından takip edilen zayıf sınıflandırıcıların kombinasyonu anlamına gelen perseptron şeklini alır.

AdaBoost öğrenme prosedürü tarafından sağlanan resmi güvence hayli güçlüdür. Freund ve Schapire güçlü sınıflandırıcının çalışma hatasının sıfıra yaklaştığını kanıtlamışlardır. Daha da önemlisi birkaç sonuç genelleştirme işlemlerini kanıtlamıştır (Schapire ve diğerleri, 1997). Önemli kısım şudur ki genelleştirme işlemi örneklerin marjinleri ile ilgilidir ve AdaBoost geniş marjnlere hızlı bir şekilde ulaşabilmektedir.

Geleneksel Adaboost prosedürü özellik seçim süreci olarak yorumlanabilir. Geniş çaplı sınıflandırma fonksiyonlarının çoğunluğun oyunun kullanılarak bütünleştirildiği arttırma problemini düşünelim. Yapılması gereken büyük ağırlığı iyi sınıflandırma fonksiyonu ile ve küçük ağırlığı da zayıf fonksiyonlarla ilişkilendirmektir. AdaBoost, önemli değişkenlikler gösterse de az sayıda iyi

sınıflandırma fonksiyonu seçmede saldırgan bir mekanizmadır. Zayıf sınıflandırıcılar ve özellikler arasında bir analogi kuran AdaBoost az sayıdaki iyi “özellik” araştırmada etkili bir prosedürdür.

Bu analogiyi tamamlamada pratik yöntem, zayıf öğreniciyi her biri tek bir fonksiyona dayanan bir dizi sınıflandırma fonksiyonuyla kısıtlamaktır. Bu amacı desteklemek için, zayıf öğrenme algoritması pozitif ve negatif örnekleri en iyi ayıran tek dikdörtgen özelliği seçmek için tasarlanmıştır. (Tieu ve Viola'nın (2000) yaklaşımına benzemektedir). Her özellik için, zayıf öğrenici uygun eşik sınıflandırma işlevine karar verir. Zayıf sınıflandırıcı ($h(x, f, p, \theta)$), böylelikle özellik (f), eşik (θ), ve eşitliğin yönünü belli eden kutuplaşmadan (p) meydana gelmektedir:

$$h(x, f, p, \theta) = \begin{cases} 1 & pf(x) < p\theta \\ 0 & \end{cases}$$

Burada x , bir görüntünün 24 x 24 piksellik alt penceresidir.

Pratikte, hiçbir tek özellik sınıflandırma işini az hata oranıyla gerçekleştirememektedir. Sürecin başlangıcında seçilen özellikler 0,1 ve 0,3 arasında bir hata oranı verirler. Sonraki aşamalarda seçilen özellikler, görev daha da zorlaştıkça, 0,4 ve 0,5 arasında bir hata oranı verirler.

2.1.2.1. Öğrenme Tartışmaları

Projede betimlenen algoritma olası zayıf sınıflandırıcılardan önemli zayıf sınıflandırıcıları seçmek için kullanılır. AdaBoost süreci çok etkili olmasına karşın zayıf sınıflandırıcılar dizini oldukça geniştir. Her belirgin özellik birleşimi için bir tane zayıf sınıflandırıcı olduğu için, K 'nin özellik sayısı ve N 'nin de örnek sayısı olduğu etkili KN zayıf sınıflandırıcıları vardır. N 'ye bağlılığı anlayabilmek için örneklerin verilen özelliğin değerine göre sınıflandırıldığı farz edilir. Çalıştırma süreci uyarınca, aynı sıralandırılmış örnekler arasında duran herhangi iki eşik eşittir. Böylelikle toplam belirgin eşik sayısı N 'dir. N 'nin 20000 ve K 'nin 160000 olarak verildiği bir işte 3,2 milyar tane belirgin zıt zayıf sınıflandırıcı vardır.

Wrapper metodu M zayıf sınıflandırıcısını kullanan bir algıyı öğrenmek için kullanılabilir. (Hohn ve diğerleri, 1994) Bu metot, her turda perseptrona zayıf bir sınıflandırıcı eklemek suretiyle adım adım işler. Zayıf sınıflandırıcı ekleme güncel sete eklendiğinde en düşük hata oranını verir. Her tur en azından O(NKN) (veya 60 trilyon işlem) civarındadır; tüm zıt özellikleri numaralandırmak ve özelliği kullanarak her örneği değerlendirme zamanıdır. Perseptron ağırlıklarını öğrenme zamanını ihmal eder. Böyle olsa bile, 200 özellikli bir sınıflandırıcıyı öğrenmek için yapılan nihai çalışma 10^{16} işleme tekabül eden O(MNKN) civarında bir şey olacaktır. AdaBoost'un özellik seçme mekanizması olarak wrapper metodu gibi rakiplere karşı önemli avantajı öğrenme hızıdır. AdaBoost kullanarak 200 özellikli sınıflandırıcı 10^{11} işlem civarında O(MNK)'da öğrenilebilir. Önemli bir avantaj her turda daha önce seçilmiş olan özelliklere olan bağlılığın örnek ağırlıklar kullanılarak etkili sıkıştırılmış bir halde şifrelenebilirliğidir. Bu ağırlıklar sabit zamanda verilen bir zayıf sınıflandırıcıyı değerlendirmede kullanılabilir.

Zayıf sınıflandırma seçimi algoritması şöyle işlemektedir. Her özellik için, örnekler özellik değerine göre sıralandırılır. O özellik için AdaBoost uygun eşiği bu sıralandırılmış listede tek seferde hesaplanabilir. Sıralandırılmış listede her bir öge için dört toplam korunur ve değerlendirilir: Toplam pozitif örnek ağırlığı T^+ , negatif örnek ağırlığı T^- , aşağıdaki örnekte pozitif ağırlıklar toplamı S^+ ve negatif ağırlıklar toplamı S^- 'dir. Şimdiki ve önceki örnek arasındaki mesafeyi bölen eşiğin hatası şöyle gösterilebilir:

$$e = \min (S^+ + (T^- - S^-), S^- + (T^+ - S^+)),$$

Bu toplamlar araştırma devam ettikçe kolayca güncellenebilir.

Çoğu genel özellik seçim prosedürleri öne sürülmüştür (Webb, 1999). Nihai uygulama, özelliklerin çoğunu dışarıda bırakacak olan saldırgan bir süreci gerektirmektedir. Benzer tanıma problemi için, Papageorgiou (1998) özellik değişkeni temelli özellik seçimi projesi öne sürmüştür. Toplam 1734 özellikten 37 özelliği seçerek güzel sonuçlar gösterilmiştir. Bu önemli bir azaltma olsa da her görüntü alt penceresi için değerlendirilen özelliklerin sayısı hala oldukça fazladır.

Roth ve diğeri (2000), Winnow üslu öğrenme kuralı temelli özellik seçme süreci öne sürmüştür. Bu yazarlar, her pikselin d boyutlu zıt vektörlerle işaretlendiği hayli geniş ve alışılmıřın dıřında özellik setleri kullanırlar. ($[0, d - 1]$ aralığında piksel x değeri aldığında, x . boyut 1'e ayarlanır ve diğeri 0'a ayarlanır). Her piksel için ikili vektörler nd boyutlu (n , piksel sayısıdır)tek bir ikili vektör oluşturmak için bağlanırlar. Sınıflandırma kuralı perseptrondur ki bu da bir ağırlığın, girilen vektörün her boyutuna atanmasıdır. Winnow öğrenme süreci, bu ağırlıklarının çoğunun sıfır olduđu bir çözüme kavuşur. Ama bununla birlikte çok sayıda özellik de korunur.

2.2. OpenCV (Open Source Computer Vision Library)

OpenCV açık kaynak kodlu bilgisayarla görü ve makine öğrenimi ile ilgili bir yazılım kütüphanesidir. OpenCV bilgisayarla görü uygulamaları için altyapı sağlar. Bu kütüphane 2500'den fazla optimize edilmiş algoritmayı içerisinde bulundurur. Bu algoritmalar yüz bulma ve tanıma, obje tanımlama, video içindeki insan hareketlerini sınıflandırma, kamera hareketlerini takip etme, obje hareketlerini takip etme gibi olaylarda kullanılabilir [18].

2.3. OCR (Optical Character Recognition)

OCR, yazılı dokümanlardan tarayıcı yardımıyla alınan görüntünün içeriğinde yer alan metnin bilgisayarınızın anlayacağı dile çevrilebilmesini sağlayan bir teknolojidir. Bu sayede metni elle yazmaya gerek kalmadan bilgisayar üzerinde yeniden düzenlenebilir, birçok program altında kullanılabilir ve yeniden düzenlenebilir hale getirebilir. OCR teknolojisi, karakterlerin tanımlanması için temel olarak üç farklı yöntem kullanır.

2.3.1. Matris Eşleme (Matrix Matching)

Bu yöntemde OCR yazılımı, karşılaşması muhtemel tüm karakterlerin görüntüsünü bir veri tabanında tutar. Daha sonra taranmış dokümandaki karakterleri bir bir bu kütüphanedeki şekillerle eşleyerek hangisine uygun olduğunu bulmaya çalışır. Ucuz ve hızlı çalışan bir yöntemdir, ancak okunacak fontun şekli biraz olsun değiştiğinde hassasiyeti kaybolur.

2.3.2. İçerik Analizi (Feature Analysis)

Bu yöntemde yazılım, karakterleri görüntüsünden değil, kendilerine özgü özelliklerinden hareketle tanımlamaya çalışır. Kullanılan harfte kaç tane düz çizgi var, bunların kaç dikey kaç yatay, yuvarlak köşelerin konumları neler, karakterde delikler var mı, vb. Bu sayede tanımlama kriterlerine ve genel kuralına uygun olarak yazılmış hemen her karakter içerik analizi yöntemiyle tanımlanabilir. İçerik analizinde karakterin şekli genel karakter kurallarına uyduğu sürece hangi karakter setinin kullanıldığının pek bir önemi yoktur.

2.3.3. Özgün Tanımlama (Self-Assertion)

Bu yöntemde yukarıdaki iki yöntem bir arada kullanılır. Önce doküman içerik analizi yöntemiyle taranarak genel bir tablo oluşturulur ve bu sayede kesinliği yüksek karakterler bir kenara ayrılarak bir veri tabanına kaydedilir. Daha sonra ilk tarama sonucunda ne olduğundan emin olunamayan karakterler, veri tabanındaki ne olduğu kesinleşmiş karakterlerle karşılaştırılarak “benzetilmeye” çalışılır. Yani sistem önce içerik analiziyle kesinlikle A olduğuna emin olduğu bir A yakalayarak bunun fotoğrafını çeker. Daha sonra da bu A harfinin görüntüsünü A olduğundan emin olamadığı, fakat A'ya benzediğini düşündüğü karakterlerle karşılaştırıp ortak noktalarına bakarak gerçekten A olup olmadığına karar verir.

Bu karar verme aşamaları tamamen yazılım tarafından işletilen bir süreçtir ve çoğu zaman en uygun yöntemin hangisi olduğuna yazılım karar vererek uygular. Dolayısıyla bu aşamada sizin yapabileceğiniz pek bir şey yoktur; karakterleri başarılı bir şekilde tanıma yüzdesi sadece OCR yazılımının yeteneklerine bağlıdır.

2.3.4. Verimliliği Arttırmak İçin Kullanılan Yöntemler

OCR işleminin doğruluğu açısından kullanılan yöntemlerin verimliliği OCR programının yeteneklerine bağlı olsa da, OCR işlemi öncesinde yapacağımız bazı ön hazırlıklarla doğruluğun artmasına katkıda bulunabilir.

Öncelikle OCR teknolojisinin, metnin sayısal ortama aktarılan görüntüsü üzerindeki harflerin ve sembollerin önceden belirlenmiş birtakım karakter bilgileriyle eşleştirilmesi temeline dayanmaktadır. Yani elde edilen görüntüdeki örneğin bir A

harfinin, yazılım tarafından mevcut karakter haritalarıyla ya da tanımlarıyla eşlenerek A olup olmadığına karar verilmesi gerekmektedir. Temelde bir metni OCR işlemine tabi tutacak biçime getirmek için üç farklı seçenek kullanılabilir: Renkli, gri tonlama (grayscale) ve siyah beyaz (lineart).

2.3.4.1. Lineart (Siyah beyaz)

Lineart, pikselleri sadece siyah ve beyaz olarak tanımlayabilen bir format olduğu için görüntü üzerinde oynayabileceğiniz sadece bir tek parametreye sahiptir. Bu yegâne parametre “threshold”, yani algılama eşiği adını taşımaktadır. Algılama eşiğini, elde ettiğiniz görüntü üzerinde bir pikselle karşılaştığında bunun siyah mı, yoksa beyaz mı olacağına dair verdiği kararı etkileyen bir değer olarak düşünülebilir. 0-255 Arasında değişen ve orijinali 128 olan algılama eşiği değerinin 0’a yaklaşması okunan piksellerin renginin siyah ya da beyaz oluşuna dair verilen kararın beyaza doğru, 255’e yaklaşması ise siyaha doğru kaymasına neden olur. Bu nedenle algılama eşiği değerini 255’e yaklaştırdıkça tamamen siyah, 0’a yaklaştırdıkça tamamen beyaz bir sayfanın önünüzde belirdiğini görülmektedir.

OCR yazılımı ideal sonuçlara ulaşabilmek için karakter sınırlarının iyi ve temiz belirlenmiş olmasını ister, gördüğü harfi referanslarıyla karşılaştığında doğru sonuca ulaşabilir. Eğer ortada karakterin tanınmasını zorlaştıracak kadar yabancı piksel varsa, karakterler birbiri içine geçmişse ya da karakterlerin sınırlarını oluşturan pikseller eşleştirmeyi zorlaştıracak ölçüde kaybolmuşsa, OCR yazılımının doğruluğu bunlardan etkilenecektir. Dolayısıyla algılama eşiğini öyle bir ayarlanmalıdır ki ne çok koyu olup karakterlerin iyice birbirine geçmesine sebep olmalı, ne de çok açık olup karakterlerin sınırlarının kaybolmasına yol açmalıdır. Lineart’ın en büyük avantajı, lineart taramada elde edilen her bir piksel sadece 1 bitlik (siyah veya beyaz) renk bilgisi içerdiğinden dolayı dosya boyutunun düşük ve tarama hızının yüksek olmasıdır. Dezavantajları ise görüntülere keskinleştirme ya da otomatik aydınlık seviyesi belirleme gibi fonksiyonların uygulanamaması ve güzel bir algılama eşiği dengesi tutturmanın nispeten zahmetli oluşu.

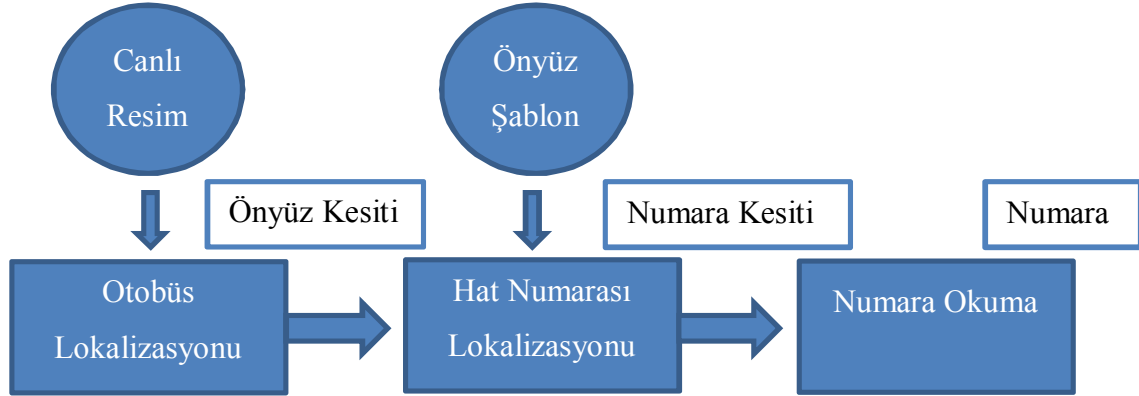
2.3.4.2. Grayscale (Gri tonlama)

Grayscale modunda her piksel için 256 gri tonu (8 bit) tanımlaması yapılabiliyor ve haliyle sahip olduğu parametrelerle özelleştirme imkânları daha fazla. Bu formattaki bir görüntüye sadece parlaklık ayarı değil, keskinleştirme gibi OCR işlemini kolaylaştıran efektler de eklenebilmektedir.

Grayscale birçok doküman için OCR amaçlı tavsiye edilen en uygun kullanım şekli, ancak birkaç ufak kusuru da yok değil. Birincisi, grayscale modunda resim tarandığında arka plan hafif gri bir renge bürünüyor ve bu da lineart'ın keskin siyah-beyaz ayrımlarıyla karşılaştırıldığında karakter tanımlama yeteneğine az da olsa etki edebilecek bir faktör. İkincisi, taranan görüntünün dosya boyu lineart eşiyile karşılaştırıldığında 8 kat daha fazla ve tarama süresi biraz daha uzundur. Bu faktörler çok sayıda sayfanın taranıp işlenmesini gerektiren durumlarda göz önüne alınabilir.

3. PROJENİN ÇALIŞMA YAPISI

Bu proje, görme engelli insanların ellerinde cep telefonları ile otobüs durağında durarak gelen otobüsün hat numarasını öğrenmesine olanak sağlar. Kullanıcı otobüsün sesini duyduğu zaman telefonun kamerasını o yöne doğru çevirmesi uygulamanın başlaması için yeterlidir. Bu uygulama otobüsün bir resminden numarasını okuması için tasarlanmıştır. Bununla birlikte daha yüksek tanıma güvenilirliği için birkaç ardışık görüntü art arda işlenir ve en sık rastlanan sonuç kullanıcıya sesli olarak bildirilir. Bu yöntem birkaç hesaplama modülüne ayrılmıştır. Her şeyden önce, yaklaşmakta olan otobüs görüntü içinde lokalize edilmektedir. Elde edilen kesit farklı bir modül de işleme tabi tutulur. Bu seçim sayesinde gereksiz görsel bilgidan kurtulur. Böylece hesaplamalar hızlanır. Otobüs kesiti bir kez bulunduktan sonra hat numara bölgesini izole etmek için bir başka görüntü kırpmaya gerçekleştirilir. Otobüs cephesini içiren bir şablonla otobüsün ön yüz kesiti eşleştirilir. Şablonda otobüs numarasının bulunduğu kısma denk gelen kesit alınır. Uygulamanın son aşamasında numaranın okunması yer almaktadır. Numaranın bulunduğu kesit görsel işlemden geçer ve sınıflandırıcı kullanılır. Sonuç olarak elde edilen numara kullanıcıya bildirilir.



Şekil 6. Projenin çalışma yapısı şeması

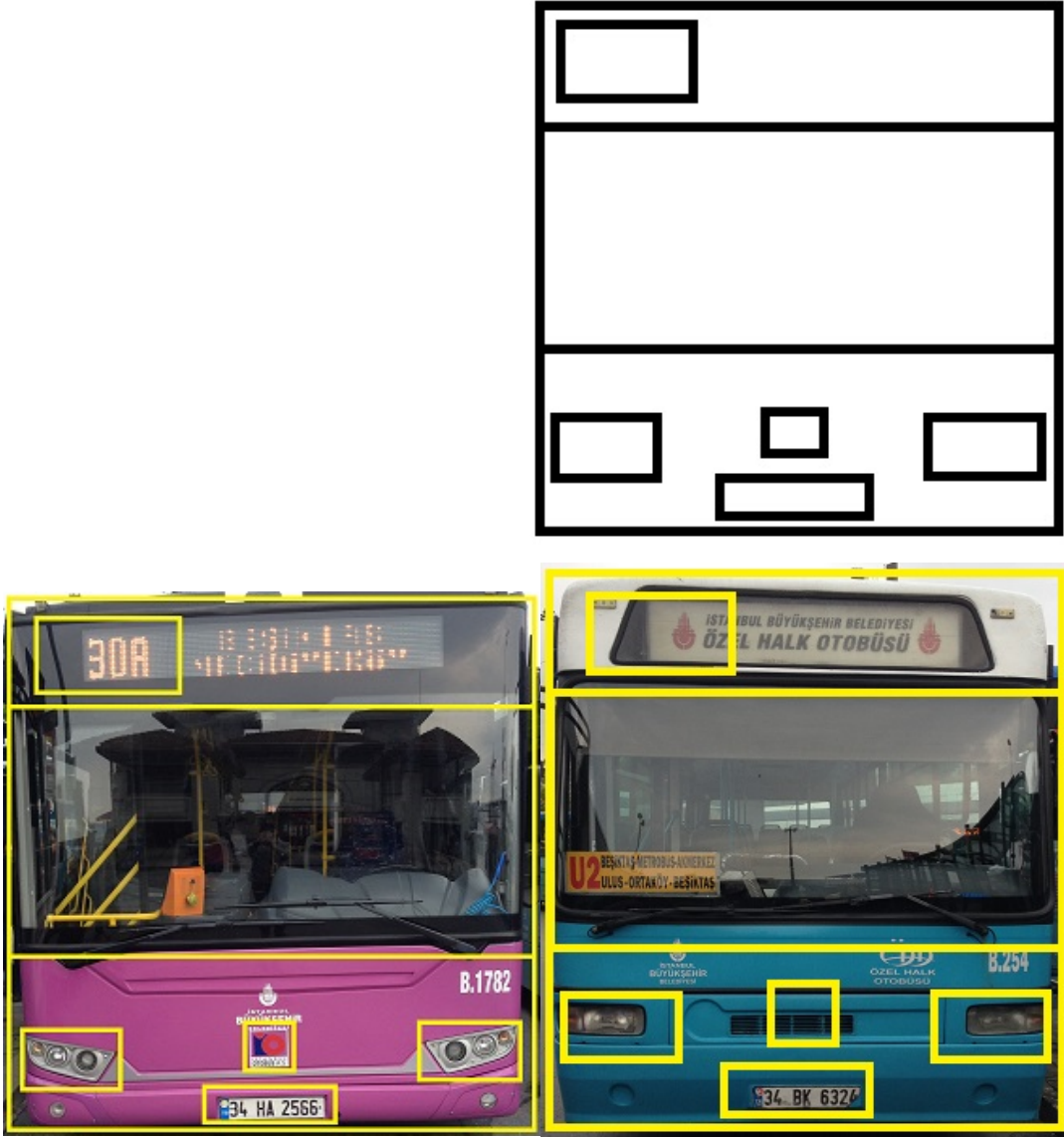
3.1. Otobüsün Tanımlanması ve Lokalizasyonu

Resimdeki otobüsün varlığını ve yerini tespit etmek için Viola ve Jones sunmuş olduğu makine öğrenimi algoritması kullanıldı. Orijinalinde bu algoritma yüz tanımak için kullanılmıştı. Bu proje otobüs tanımak için kullanılmıştır. Aralarındaki resimlerdeki az bulunan nadir özelliklerin belirlenmesi için fark yüz resimlerinin negatif ve pozitiflerinde yapılan çalışmaların otobüs resimleri üzerinde yapılmasıdır. Makine öğrenimine sunulan resimleri, farklı marka otobüslerin ön yüz fotoğraflarının ve negatiflerinin birleşimi oluşturmaktadır. Birçok resim üzerinde nadir özellik bulma işlemi yapılmıştır. Ayrıca her otobüs markası için farklı açıdan resimlerde işleme dâhil edilmiştir.

3.2. Hat Numarası Tanımlanması ve Lokalizasyonu

Bir sefer otobüs bulundu mu, akabindeki tüm çalışmalar otobüs kesiti ile sınırlanır. Farklı açılarda metottu sağlamlaştırmak ve projektif bozulmaları düzeltmek için otobüsün ön yüzünü tanımlayan bir şablon kullanıldı. Şablon kullanılmadan yapılan lokalizasyonlarda ışık varyasyonları ve yansıma olaylarında dolayı tam sonuca ulaşılamamaktadır. Bu tür görüntü sorunları sebebiyle otobüsün ön yüz lokalizasyonu için kullanılan metot numara lokalizasyonu için kullanılamamaktadır. Görüntü üzerinde önemli özellikler yakalanamamaktadır.

Şablon otobüslerin belirgin ve ayır edici özellikleri belirlenerek oluşturulmuştur. Bu özelliklerin bazılarını alt kısımda yer alan farlar, amblem ve plaka oluşturmuştur. Şablon ayrıca numaranın bulunduğu kısmı da içermektedir. Bu kısım otobüslerin sol üst köşesine yakın bir noktada bulunur. Fakat bazı eski otobüsler bu şablona uymamaktadır. Bu sebepten dolayı onlarda numara okuma yapılamamaktadır.

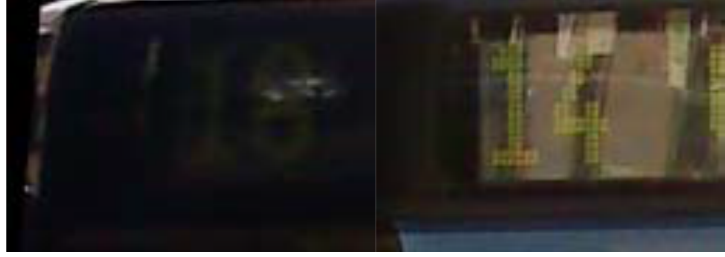


Şekil 7. Üzerlerine şablon oturtulmuş otobüs resimleri

Otobüs kesitinden alınan numara kesiti üzerinde “Canny kenar belirleme” ve “Hough transform” gibi farklı görüntü işleme metotları kullanılmıştır. Üst kısımda oluşan gökyüzü, ağaç ya da bina gibi yansımalar numara almayı epeyce zorlaştırmıştır.

3.3. Hat Numarasının Okuması

Öncelikle otobüs tanımlanır, ardından otobüs kesiti üzerinden şablon yardımıyla numaranın bulunduğu kesit alınır ve bu kesit üstünde işlemler yapılır. Aşağıdaki **Şekil 8**'de normal bir numara kesiti görülmektedir.



Şekil 8. İşlenmemiş resim üzerinden otobüs hat numarası kesitleri



Şekil 9. OCR temel çiftleme metodu sonucu oluşan resimler



Şekil 10. Hat numaralarının renk tonuyla filtreleme yapılmış resimler

Okumayı gerçekleştirmek için, kesit üzerinde işlemlerin uygulanması büyük önem taşır. Okumada kullanılacak OCR teknolojisinde de bu işlemler ortak ve kritik işlemlerdir. Bu işlemlerden en temel olanı “çiftleme (Binarization)” operasyonudur. O işlem sonunda resimde sadece siyah ve beyaz renk içerir. **Şekil 9**'da görebildiğiniz

gibi, fazlaca gölge ve yansıma olayları OCR'ın temel çiftleme (Binarization) metodunu numara okuma için kullanışsız hale getirmektedir. Bu sorunun çözümün için, tamamlayıcı renklere dayalı alternatif ve özgün bir çiftleme yöntemi kullanılmıştır. Bu yöntem otobüsün numarasının renk bilgisinden yararlanmaya çalışır. HSV renk uzayını kullanarak numara renginin tamamlayıcısını kolaylıkla bulunabilir. Aynı renk tonuna sahip olamayan tüm renkleri filtreleme yöntemleri sayesinde siyah renge çekilir. Bu işlemin sonucunu **Şekil 10**'da görülmektedir. Fark edildiği gibi çiftleme yöntemi sonunda resim kalitesinde okunabilirlik açısından iyileşme olmuştur. Işık varyasyonlarından işlemin etkilenmemesi için, resim üzerinde eşikleme (thresholding) işlemi yapılmıştır. Eşikleme yönteminde optimum eşik değerinin bulunması için farklı eşik değerleri (τ) ta ki istenilen netliğe ulaşılan kadar denenmiştir. **Şekil 11**'de bir eşikleme sürecinin sonuçları görülebilmektedir. Denemeler sonucunda anlaşılışmıştır ki en iyi eşik değerinin (τ) 50 olduğu saptanmıştır.



Şekil 11. Sırasıyla farklı eşik değeri denemeleri $\tau = 20$, $\tau = 50$, $\tau = 70$ ve $\tau = 90$

Tüm çiftleme işlemleri bitirildiğinde otobüsün hat numarası elde edilebilir hale gelmiştir. Yine makine öğrenimi bu görev için fazlaca kullanılmıştır. Böylelikle her hat numarası için yani her rakam veya harf için tek tek değil tümünü tanıyabilecek bir yapı oluşturulmaya çalışılmıştır.

4. DENEY VE ANALİZ SONUÇLARI

Bu bölümde, uygulama üzerinde yapılan deney ve analizlerin sonuçları ayrıca gözlemler sonucu yetersiz olduğu düşünülen kısımlar anlatılmaktadır. Proje kendi ana işlevlerinden olan otobüs bulma modülü ve hat numarasını okuma modülü olmak üzere iki önemli işlevler üzerinde deneyler yapılarak değerlendirilmeye alınmıştır.

4.1. Otobüsü Tanımlama ve Lokalize Edilmesi

Sınıflandırıcının otobüs tanımlama görevini test etmek için önceden farklı marka ve tipteki otobüslerin farklı açılardan birçok resmi çekilmiştir ve sınıflandırıcı ile bir veri oluşturulmuştur. Bu veri oluşumunda İstanbul ulaşımında kullanılan otobüsler temel alınmıştır. İstanbul Elektrik Tramvay ve Tünel İşletmeleri'nden alınan bilgiye göre, İstanbul'da on farklı model otobüs bulunmaktadır. Bu otobüslere özel halk otobüsleri dâhil değildir. Oluşturulan sistem İETT' on markadan dokuzunda olumlu sonuç göstermektedir. Özel halk otobüsleri de %90 oranında aynı markalardan oluşmaktadır. Fakat belirtilen markalarda olma zorunlulukları yoktur. Özel halk otobüslerinde İETT araçlarında farklı olarak tercih edilen başlıca markaları aşağıdaki listede yer almaktadır. Bu listede yer alan tüm otobüslerin verileri oluşturulan sistemde mevcuttur.

Marka/Tip	Adet	Sonuç
Temsa Avenue LF	342	Olumlu
Gülyüz Cobra	180	Olumlu
BMC Procity Euro5	114	Olumlu
Tezeller Autobus	67	Olumlu
TCV Karat	66	Olumlu
Man Lions City	20	Olumlu
Mercedes Conecto	139	Olumsuz
Otokar Kent 290 LF (Solo)	162	Olumlu
TOPLAM	1090	10/10

Tablo 1. İstanbul Otobüs A.Ş. 2014 otobüs filo bilgileri ve tanıma sonuçları

Marka/Tip	Adet	Sonuç
Mercedes 0345 (Solo)	477	Olumlu
Mercedes Citaro (Solo)	393	Olumlu
Mercedes 0345K (Körüklü)	84	Olumlu
Mercedes Citaro 0530 (Körüklü)	100	Olumlu
Mercedes Capacity (Körüklü)	250	Olumlu
Mercedes Conecto G (Körüklü)	265	Olumlu
APTS-PHILEAS (Körüklü)	50	Olumsuz
OTOKAR KENT 290 LF (Solo)	900	Olumlu
KARSAN B. Avancity+CNG (Solo)	240	Olumlu
KARSAN B. Avancity+	300	Olumlu
TOPLAM	3060	9/10

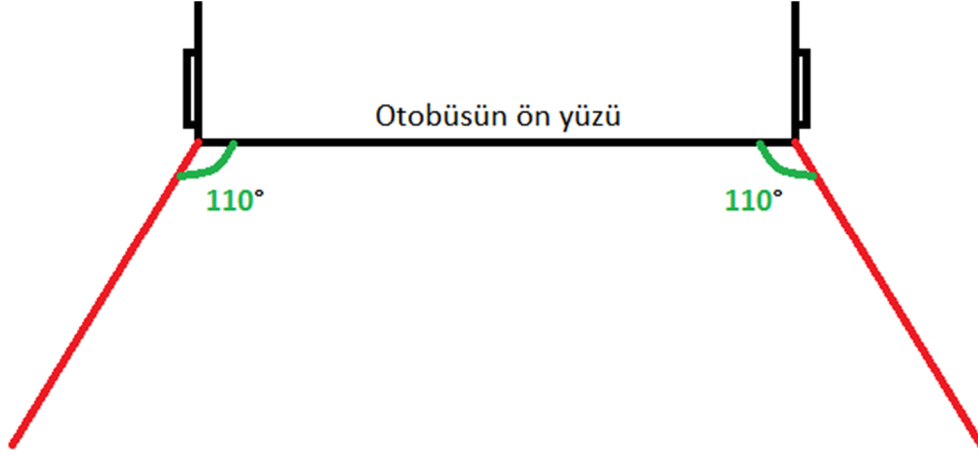
Tablo 2. İETT 2014 otobüs filo bilgileri ve tanıma sonuçları

Yukarıda belirtilen bazı otobüs markalarına ait şehir ortamında çekilmiş ön yüz resimleri aşağıda yer almaktadır.



Şekil 12. Şehir ortamında çekilmiş otobüs ön yüz resimleri

Otobüs tanımlama bir başka kriterde otobüsün kamera açısıdır. Tanımlamanın başarılı olabilmesi için otobüsün ön yüzünün kameranın kadrajına doğru açıyla girmesi gerekmektedir. Tanımlamalarda kamera ile otobüs arasındaki yakınlık mesafesi kamera özelliklerine göre değişim gösterebilir. Fakat otobüsün ön yüz görüntüsünün belli açılardan fazla yana kaymaması gerekmektedir. Farklı açılardan denemeler yapılmıştır ve sonuç olarak kameranın bulunduğu noktayla otobüsün ön yüzü arasındaki açının yetmiş dereceden az olmaması gerektiği saptanmıştır. En uygun açı değerleri aşağıdaki şekilde gösterilmeye çalışılmıştır.



Şekil 13. Kameranın konumu için uygun açı değeri

Sistem her çekilen görüntü karesi üzerinde sınıflandırma ve tanımlama işlemleri yapmaktadır. Dolayısıyla açı ya da mesafe yüzünden tanımlanamayan otobüsler size yaklaştıkça ya da doğru açığa geldikçe sistem tarafından algılanacaktır.



Şekil 14. Uygulamanın tanımlaması sonucu oluşan resimler

4.2. Hat Numarasının Okuması

Bu bölümde projenin otobüs hattı numarası tanıma görevi ve performansı gösterilmektedir. Otobüsün tanımlanmasının ardından tanımlanan lokasyondan otobüsün ön yüz görüntüsü arka plan görüntüden ayrılır ve elde edilen yeni görüntü hat numarasının tespiti için bir şablon üzerine oturtulur. Tüm ön yüz görüntüsü üzerinden hat numarası okunmaya çalışıldığında fazla zaman almaktadır ve yanlış sonuç elde edilmektedir. Daha küçük bir parça üzerinde numara okuma yapılması

sistemi hızlandırmaktadır. Yeni yapılan otobüslerin hepsinde hat numaraları sol üst köşede bulunduğundan şablon yardımıyla hat numarasının bulunduğu kısım alınıp okuma işlemine sokulmaktadır. Bu sürecin denemelerinde bazı sorunlarla karşılaşmıştır.

4.2.1. Şablon Uyuşmazlığı

Şablon uyumsuzluğu otobüsün hat numarasının şablon üzerinde belirtilen noktaya denk gelmemesinden oluşmaktadır. Sistemin daha hızlı çalışması için çoğunlukla hat numaralarının yazıldığı sol üst köşe üzerinden kesit alınıp okuma işlemine sokulmaktadır. Farklı bir lokasyondan yazılan hat numaralarının okuna bilmesi için şablona ek yapılması ve o ek üzerinden elde edilen kesitinde okuma işlemine dâhil edilmesi gerekmektedir. Bu tür eklemeler süreci uzatmakta ve mobil cihazlarda uygulamanın yavaş çalışmasına yol açmaktadır.

4.2.2. Hat Numarasının Elde Edilememesi

Bu sorun mobil cihaz kameralarının LED tabelalar üzerindeki hat numarasını net biçimde çekememesinden kaynaklanmaktadır. Çekilen görüntü üzerinde hat numarasının bütünüyle görünmemesinden dolayı okuma işlemi sonucu yanlış çıkmaktadır.



Şekil 15. Telefon kamerasından çekilmiş dijital kamera görüntüleri

4.2.3. Okuma Sonucu Çıkan Fazla Karakterler

Bu sorun genellikle hat numaralarının yanında fazladan belirtici amaçlı kullanılmış çizgilerden kaynaklanmaktadır. Hat numarasının siyah üzerine beyaz gibi daha kolay ayırt edilebilir bir şekilde yazılmış olması sistemin okuma sürecindeki hataların azalmasını sağladığı gözlenmiştir. OCR kütüphanesinin numarayı net bir şekilde okuyabilmesi için kesit üzerinde numarayı ön plana çıkaracak işlemler uygulanmaktadır. İşlemler sonrası kalan fazlalıklar OCR işleminden çıkan sonucu etkilememektedir.



Şekil 16. Otopüs üzerinden alınmış numara kesiti



Şekil 17. Özel olarak yapılmış çiftleme metodu sonucu oluşan resimler



Şekil 18. OCR kütüphanesinin temel çiftleme metodu sonucu oluşan resimler

Şekil 16'da görülen resimler işlenmemiş numara kesitleridir. Şekil 17'de projede uygulanan çiftleme işlemi sonucu çıkan resimler yer almaktadır. Şekil 18'de ise temel çiftleme işleminin sonuçları yer almaktadır.

Yazının sade olmayışı ve kesit üzerinde yazının belirginleştirilmesi için yapılan işlemleri bazen yetersiz olmasından sonuçlar görülebilmektedir. Aranılan numaranın dışında fazladan karakterler tanımlanabilmektedir. Net sonucu bulmak için birden fazla OCR Kütüphanesi denenmiştir. İşleme tabi tutulmuş görselleri OCR işlemine tabi tutmamıza rağmen kesin sonuç alınamamaktadır. En iyi sonuçları veren ve açık kaynak kodlu Tesseract OCR Kütüphanesi projede kullanılmıştır.

5. SONUÇ

Son birkaç yıl içerisinde, gelişmiş teknolojiler engelli insanların hareketliliklerini ve özerkliklerini geliştirmek için fazlaca katkıda bulunmuştur. Örnek olarak, yeni etkileşim paradigmaları sesli komutlara ya da özellikle hareket engelli insanlar için geliştirilmiş göz takip mekanizmalarına dayanmaktadır. Bu paradigmalar tekerlekli sandalyelerin kontrolünde ve bilgisayarlar aracılığı ile iletişim kurmaları için kullanılmıştır.

Bilgisayarla görü, görme engelli insanları içerideki ve dışarıdaki senaryolar kullanabilecekleri en yararlı ve önemli teknolojilerden biridir. En basit düşünceyle kameralar birer göz gibi kullanılabilir. Bu mekanik gözlerin amacı yazılımlar yardımıyla görüntüyü otomatik analiz edip kullanıcılarının günlük görevlerine yardımcı olmaktır. Mevcut bulunan tüm işler ve görme engelli insanlar için yapılmış cihazlar, görüntü işleme ya da analiz etme üzerine metotlar geliştirmeye odaklanmıştır.

Bu yazıda, görme engelli insanlar için gelen bir otobüsün bulunması ve hat numarasının okunması amaçlı bir proje sunulmuştur. Bu projeye benzer parçalar içeren, yapılar kullanan projeler bulunmaktadır. Önceden yapılmış bir bilgisayarla görü projesinde sunulduğu üzere, bir bilgisayarla görü aracı video üzerindeki birçok nesneyi tanımlayabilmiştir. Bu tanımlama, görme engelli kullanıcıların sevdikleri dizilerinin az diyalog içeren kısımlarını daha iyi anlayabilmelerinde yardımcı olmuştur. Bir başkasında, özel bir görüntü kontrast geliştirme yöntemi gösterdi ki resim üzerindeki yazı ya da diğer görsel materyallerde yapılan kontrast ayarlamaları sayesinde yarı görme engelliler resmi anlayabilmişlerdir. Son olarak bir diğer projede omuzda taşınan bir kamera ve sırt çantasında taşınan bir bilgisayar ile şehir ortamındaki yazıların belirlenip okunduğu gözlemlenmiştir.

Şehir senaryolarının çoğunda, görme engelli insanların otobüsü bulmaları dışında doğru otobüsü bulmalarında büyük zorluklar çektikleri gözlemlenmiştir. Bu zorluk otobüs duraklarında otobüs numaralarının sadece görsel olarak insanlara sunulmuş olmasından dolayı oluşmuştur. Bu proje, tamamen otomatik olarak, otobüs numarasının bulunduğu yerin saptanmasını ve numaranın okunmasını sağlayacak bir

yöntem bulmayı amaçlanmaktadır. Bu proje akıllı telefonlar üzerinde kullanıldığı için ekstra bir cihaz gerektirmemektedir. Bu görüntü işleme projesi, ışık seviyesindeki değişiklikler, yansımalar, değişik bakış açıları gibi tüm açık alanda çıkabilecek etkenler göz önüne alınarak hazırlanmaya çalışılmıştır.

Projede önceden geliştirilmiş farklı yapılar kullanılmıştır. Otobüslerin tanımlanmasında basamaklı sınıflandırıcı (cascading classifiers) algoritma ve elde edilen hat numarası görüntülerini okumak için OCR kütüphanelerinden yararlanılmıştır.

Kullanılan sınıflandırıcı görüntü üzerindeki piksel değerlerinin toplamalarını kıyaslayarak ayırt edici özellikler bulmak üzerine dayanmaktadır. Bu özelliklerin bulunması için gereken resimler tanımlanmak isteyen öğenin net bir biçimde yer aldığı resimlerdir. Birçok resim bir araya getirilerek sınıflandırıcı için bir veri seti meydana getirilmiştir.

Tanımlanan ve resim üzerinde kesiti alınan her otobüs görüntüsü hat numarasının daha kolay elde edilebilmesi için bir şablon üzerine oturtulmaktadır. Şablon üzerinde hat numarasının lokasyonu önceden belirlenmiştir. Numaranın bulunduğu kesim arka plandan şablon yardımıyla ayrılır ve OCR taramasına girmeden önce görüntü işleme işlemleri yapılır.

Görüntü işlemlerin başlıca hedefi var olan doğal etkilerin azaltılmasıdır. Bunun için geometrik bilgisayarla görü ve makine öğrenimi algoritmaları bir araya getirilmiştir. Deney sonuçları, geometrik eşleştirme metotlarının OCR tekniklerinde daha verimli sonuçlar alınmasına yardımcı olduğunu göstermiştir. Elde edilen görüntüler önce işleme sokulup ardından numaranın okunması için OCR teknikleri uygulanmaktadır. Bu işlemlerden biri Işık varyasyonlarından işlemin etkilenmemesi için, resim üzerinde yapılan eşikleme (thresholding) işlemidir. İşlem sonucunda elde edilen resim siyah beyaz hale getirilmiş ve numara belirgin bir şekilde ön plana alınmıştır. İşlenmiş görüntü üzerinde son olarak OCR teknikleri uygulanır ve metin elde edilmeye çalışılır.

Proje şehir ortamında deney ve analizlere tabi tutulmuştur. Oluşabilecek her durum değerlendirilmeye alınmıştır ve eksikler belirlenmiştir. İstanbul ulaşımında yer alan on yedi farklı marka otobüsten on altısında olumlu sonuçlar vermiştir. Fakat projenin tanımlamada kullanılan veri seti, geliştirilebilir ve daha fazla markayı tanınması sağlanabilir durumdadır. Test aşamasında oluşturulmuş şablon üzerinde eksikler bulunmuştur. Bazı otobüslerde hat numaraları şablonun belirttiği yerde bulunmamaktadır. Bu gibi durumlarda hat numarası okunamamaktadır. Eksiklerin giderilmesi için şablon üzerinde ilaveler yapılmalıdır. Fakat her ilave fazla işlem anlamına geldiği için uygulama yavaşlayacaktır. Ayrıca kameranın net bir görüntü elde edememesi ve hat numarasının sade bir arka plan üzerine yazılmamasından dolayı da sorunlar yaşanmaktadır. Bu gibi durumlarda OCR işleminin sonucunda fazla ve anlamsız karakterler elde edilmiştir.

İleriki çalışmalarda OCR için daha iyi sonuç verebilecek görüntü işleme metodunun geliştirilmesi gerekmektedir. Ayrıca otobüslerin zamanla değişeceği için farklı model otobüslerle çekilmiş resimler ile makine öğreniminin geliştirilmelidir.

KAYNAKÇA

- [1] E. H. Land, “The retinex theory of color vision,” Scientific American dergisi, sayı: 237, no: 6, sayfa aralığı: 108–128, 1977.
- [2] Choudhury, A. - 18 Haziran 2010 – “Color contrast enhancement for visually impaired people”
- [3] J. Lakritz, A. Salway, “The Semi-Automatic Generation of Audio Description from Screenplays”, Bilgisayar Bölümü Teknik Raporu, Surrey Üniversitesi, 2006
- [4] Gagnon, L. - 18 Haziran 2010 – “Computer-vision-assisted system for video description scripting”
- [5] H. Rowley, S. Baluja, and T. Kanade, “Neural network-based face detection”, Desen Analizi ve Makine İstihbarat hakkındaki EEME (IEEE - Institute of Electrical and Electronics Engineers) işlemleri (IEEE transactions on PAMI), Cilt: 20, 1998.
- [6] F. Fleuret, and D. Geman, “Coarse-to-Fine face detection”, International Journal of Computer Vision, Haziran, 2000.
- [7] H. Schniederman ve T. Kanade, “A Statistical method for 3D object detection applied to faces and cars”, 2000.
- [8] P. Viola and M. Jones, “Robust Real-Time Face Detection”, International Journal of Computer Vision, sayı:57(2), sayfa aralığı:137–154, 2004
- [9] A. K. Jain and B. Yu, “Automatic text localication in images and video frames”, Pattern Recognition, sayı:31(12), 1998.
- [10] T. Sato, T. Kanade, E. Hughes, ve M. Smith, “Video OCR for Digital News Archives,” EEME (IEEE - Institute of Electrical and Electronics Engineers) Uluslararası, Görüntü ve Video Veritabanları İçerik Tabanlı Erişim Çalıştayı (IEEE Intl. Workshop on Content-Based Access of Image and Video Databases), Ocak, 1998.

- [11] H. Li, D. Doermann and O. Kia. "Automatic Text Detection and Tracking in Digital Video". EEME (IEEE - Institute of Electrical and Electronics Engineers) görüntü işleme işlemleri (IEEE Transactions on Image Processing), 9(1):147–156, 2000.
- [12] J. Xi, X. Hua, X. Chen, W. Liu, H.-J. Zhang. "A Video Text Detection and Recognition System". EEME (IEEE - Institute of Electrical and Electronics Engineers) uluslararası (ICME 2001), sayfa 1080-1083. 2001.
- [13] J. Quinlan, "Induction of decision trees. Machine Learning", 1986.
- [14] P. Viola and M. Jones, "Fast and Robust Classification using Asymmetric AdaBoost and a Detector Cascade", 2001.
- [15] G. Blanchard, D. Geman. "Hierarchical testing designs for pattern recognition."
- [16] V.N. Vapnik. "Statistical Learning Theory". Wiley, New York, 1998.
- [17] Xiangrong Chen - 25 Haziran 2005 – "A time-efficient cascade for real-time object detection"
- [18] OpenCV resmi hakkında sayfası <http://opencv.org/about.html>

EKLER

a) Projede kullanılan yazılım programları:

- Java, Android
- Tesseract OCR Kütüphanesi
- OpenCV 2.4.8 (Open source computer vision library)

b) Proje kitapçığıyla birlikte teslim edilecek DVD'nin içeriği:

- Bitirme projesi kitapçığı (doc ve pdf formatlarında)
- Proje sunum dosyası (pps ve pdf formatlarında)
- Proje yazılım dosyaları (Kodlar ve Cascade Xml Dosyası)
- Proje android kurulum dosyası (apk formatında)
- Tesseract OCR kütüphanesinin kullandığı veri dosyası

ÖZGEÇMİŞ

Alican TÜRKER

16 Eylül 1989 tarihi, İstanbul ili Şişli ilçesi doğumluyum. İlköğretimi Etiler Hasan Ali Yücel İlköğretim Okulu'nda, Liseyi ise Nişantaşı Nuri Akın Anadolu Lisesi'nde tamamladım. Beykent Üniversitesi, Mühendislik Mimarlık Fakültesi, Bilgisayar Mühendisliği Bölümü İngilizce mezunuyum. Stajımı İnfina Yazılım A.Ş'de yaptım. Üniversiteden 2012 yılı Haziran ayında mezun oldum. Lisanstan sonra Beko Ticaret A.Ş de webmaster olarak işe başladım. Aynı zamanda Beykent Üniversitesi'nde yüksek lisansa başladım. Sekiz ay kadar Beko'da çalıştım sonra ayrıldım ve şuan çalışmakta olduğum BtcTrader A.Ş de işe başladım.