

Kısıtsız Renkli İmgelerde Güncel Yüz Sezimi Yöntemlerinin Karşılaştırılması

Bahattin Mutlu

YÜKSEK LİSANS TEZİ

Elektrik Elektronik Mühendisliği Anabilim Dalı

Mayıs 2018

Comparison of Current Face Detection Methods in Unconstrained RGB Images

Bahattin Mutlu

MASTER OF SCIENCE THESIS

Department of Electrical Electronical Engineering

May 2018

Kısıtsız Renkli İmgelerde Güncel Yüz Sezimi Yöntemlerinin Karşılaştırılması

Bahattin Mutlu

Eskişehir Osmangazi Üniversitesi
Fen Bilimleri Enstitüsü
Lisansüstü Yönetmeliği Uyarınca
Elektrik Elektronik Mühendisliği Anabilim Dalı
Telekomünikasyon-Sinyal İşleme Bilim Dalında
YÜKSEK LİSANS TEZİ
Olarak Hazırlanmıştır

Danışman: Dr. Öğr. Üyesi Helin DUTAĞACI

Mayıs 2018

ONAY

Elektrik Elektronik Mühendisliđi Anabilim Dalı Yüksek Lisans öđrencisi Bahattin Mutlu'nun YÜKSEK LİSANS tezi olarak hazırladıđı “Kısıtsız Renkli İmgelerde Güncel Yüz Sezimi Yöntemlerinin Karşılaştırılması” başlıklı bu çalıřma, jürimizce lisansüstü yönetmeliđin ilgili maddeleri uyarınca deđerlendirilerek oybirliđi ile kabul edilmiřtir.

Danıřman : Dr.Öđr.Üyesi Helin Dutađacı

İkinci Danıřman : -

Yüksek Lisans Tez Savunma Jürisi:

Üye : Dr.Öđr.Üyesi Helin Dutađacı

Üye : Dr.Öđr.Üyesi Hasan Serhan Yavuz

Üye : Prof.Dr.Ömer Nezh Gerek

Fen Bilimleri Enstitüsü Yönetim Kurulu'nun tarih ve
..... sayılı kararıyla onaylanmıřtır.

Prof. Dr. Hürriyet ERŐAHAN
Enstitü Müdürü

ETİK BEYAN

Eskişehir Osmangazi Üniversitesi Fen Bilimleri Enstitüsü tez yazım kılavuzuna göre, Dr.Öğr.Üyesi Helin Dutağacı danışmanlığında hazırlamış olduğum “Kısıtsız Renkli İmgelerde Güncel Yüz Sezimi Yöntemlerinin Karşılaştırılması” başlıklı tezimin özgün bir çalışma olduğunu; tez çalışmamın tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı; tezimde verdiğim bilgileri, verileri akademik ve bilimsel etik ilke ve kurallara uygun olarak elde ettiğimi; tez çalışmamda yararlandığım eserlerin tümüne atıf yaptığımı ve kaynak gösterdiğimi ve bilgi, belge ve sonuçları bilimsel etik ilke ve kurallara göre sunduğumu beyan ederim. 14/05/2018

Bahattin Mutlu

İmza

ÖZET

Yüz bulma konusu zorlu bir alan olması ve henüz yeterli başarıyı yakalayamamış olmasının yanında birçok uygulamanın temelini oluşturmasından dolayı ilgi odağı haline gelmiştir. Son zamanlarda bu konuda yapılan sayısız çalışma bulunmaktadır. Fakat bu çalışmaların sınırlandırılmamış imgeler üzerinde başarı oranları düşüktür. Bu tezde, yüz bulma üzerine yapılan güncel çalışmaların özgün ve zorlu sınırlandırılmamış imgeler üzerinde gösterdikleri performanslar kıyaslanmıştır. Çalışma sonucunda, sınırlandırılmamış görüntülerde henüz başarı oranlarının yeterli olmadığı ve basit yaklaşımların hızlı sonuçlar verdiği görülmüştür.

Anahtar Kelimeler : Yüz bulma, Yüz sezimi, Viola-Jones Algoritması, Sınırlandırılmamış imge, Kinect

SUMMARY

Face detection is a widely studied problem since it is an important step for many computer vision algorithms. However, the performance of current face detection methods is low on images containing unconstrained faces. In this thesis, the performances of state-of-the-art face detection methods are compared using original, challenging images containing unrestricted faces. As a result of the study, it has been concluded that simpler methods are faster, and the performance of current techniques are still not sufficient on unconstrained images.

Anahtar Kelimeler : Face detection, Viola-Jones Algorithm, Unconstrained image, Kinect

TEŐEKKÜR

Yüksek lisans öğrencisi olduğum günden beri bu tez çalışmasının her aşamasında desteğini esirgemeyen, engin bilgi ve tecrübelerinden yararlandığım, yönlendirme ve bilgilendirmeleriyle çalışmamı bilimsel temeller ışığında şekillendiren değerli hocam Sayın Dr.Öğr.Üyesi Helin Dutağacı'na teşekkürlerimi sunuyorum.



İÇİNDEKİLER

Sayfa

ÖZET	vi
SUMMARY	vii
TEŞEKKÜR	viii
İÇİNDEKİLER	ix
ŞEKİLLER DİZİNİ	xi
ÇİZELGELER DİZİNİ	xii
SİMGELER VE KISALTMALAR DİZİNİ	xiii
1. GİRİŞ VE AMAÇ	1
2. LİTERATÜR ARAŞTIRMASI	3
3. MATERYAL VE YÖNTEM	5
3.1. Viola-Jones Algoritması	5
3.1.1. İntegral görüntü	6
3.1.2. Adboost sınıflandırıcı	7
3.1.3. Kaskat sınıflandırıcı	9
3.2. Yüz Bulma, Poz Tahmin Etme Ve Yüz Nirengi Noktalarını Bulma	9
3.2.1. Karma ağaç modeli	10
3.2.2. Nirengi noktalarını bulma	12
3.3. Hızlı Ve Doğru Kısıtlanmamış Yüz Sezimi	13
3.3.1. Normalize edilmiş piksel farkı (NPD) öznitelikleri	14
3.3.2. Derin kuadratik ağaç yapısı	15
3.4. SVM Kullanarak HOG Tabanlı Yüz Algılama	16
3.4.1. Yönlü gradyan histogramı (HOG)	17
3.4.2 Destek vektör makineleri (SVM)	18

İÇİNDEKİLER (devam)

	<u>Sayfa</u>
4. BULGULAR VE TARTIŞMA	20
4.1. Veritabanı	20
4.2. Başarı Ölçütleri	21
4.3. Yöntemlerin Uygulanması	22
4.3.1. Viola&Jones algoritması	23
4.3.2. Poz tahmin etme, yüz nirengi noktalarını bulma	25
4.3.3. Hızlı ve doğru kısıtlanmamış yüz sezimi	26
4.3.4. SVM kullanarak HOG tabanlı yüz algılama	28
5. SONUÇ VE ÖNERİLER	29
KAYNAKLAR DİZİNİ	32

ŞEKİLLER DİZİNİ

<u>Sekil</u>	<u>Sayfa</u>
3.1. Dikdörtgen filtreler	5
3.2. Filtreler ve uygulaması	6
3.3. İntegral görüntünün 1 konumundaki değeri, A dikdörtgeni içindeki piksellerin yeğinlik değerlerinin toplamıdır. 2 konumundaki değeri A+B, 3 konumundaki değeri A+C, 4 konumundaki değeri A+B+C+D'dir. D içindeki piksellerin yeğinlik değerleri toplamı $4+1-(2+3)$ olarak hesaplanır	7
3.4. Kaskad sınıflandırıcı	9
3.5. Poz durumuna göre oluşturulan modeller	10
3.6. Deformasyona uğramış modeller	11
3.7. Bosphorus veritabanından alım açılarına örnek	11
3.8. Nirengi bölgeleri	12
3.9. Poz açısı, nirengi noktaları ve yüz bölgesinin gösterimi	13
3.10. $f(x,y)$ NPD fonksiyonunun gösterimi	14
3.11. NPD özelliklerinin derin kuadratik ağaç yapısı ile birleştirilmesi ve öğrenilmesi	16
3.12. Gradyan yönelimleri ve bir hücreye ait histogramın gösterimi	18
3.13. a) İki sınıflı verileri ayıran hiperdüzlemler b) Optimum hiperdüzlem	19
3.14. Doğrusal veriler için optimum hiperdüzlem oluşumu	19
4.1. Veritabanından örnek imgeler	21
4.2. Yüz bulma işlemi sonucu: a) Doğru bir şekilde tespit edilen yüz görüntüsü, b) Hatalı tespit edilen ve kaçırılan yüz görüntüsü	23
4.3. V&J Algoritmasının sonuçlarına örnekler	24
4.4. Nirengi noktalarını bulma yöntemi ile tespit edilen yüzler	26
4.5. Kısıtlanmamış yüz sezimi ile tespit edilen yüzler	27
4.6. HOG algoritması ile tespit edilen yüzler	28
5.1. Sonuç değerlerin grafik gösterimi: a) Hassasiyet, b) Geri çağırma	31

ÇİZELGELER DİZİNİ

<u>Cizelge</u>	<u>Sayfa</u>
3.1. Adaboost Algoritması	8
4.1. Viola&Jones algoritması ile elde edilen sonuçlar	23
4.2. Orijinal imge boyutlarında elde edilen sonuçlar	25
4.3. İmge boyutlarının iki katına çıkarıldığı durumda elde edilen sonuçlar	25
4.4. İmge boyutlarının üç katına çıkarıldığı durumda elde edilen sonuçlar	25
4.5. Hızlı ve doğru kısıtlanmamış yüz sezimi yöntemi ile elde edilen sonuçlar	27
4.6. SVM kullanılarak HOG tabanlı yüz bulma ile elde edilen sonuçlar	28

SİMGELER VE KISALTMALAR DİZİNİ**Simgeler** θ ε_j Ω_{npd} **Açıklama**

Eşik Değer

Hata oranı

NPD öznitelik uzayı

Kısaltmalar

PASCAL, LFW, FDDB

Açıklama

Yüz algılama ve kıyaslama veri tabanları



1. GİRİŞ VE AMAÇ

Yüz sezimi problemi, verilen bir görüntü üzerinde yüz bulunup bulunmadığını, bulunan yüzlerin imgedeki konumunu tespit etme ve sunma problemidir (Yang, 2002). Yüz sezimi mevcut görüntüler üzerinde yapılabildiği gibi, gerçek zamanlı video çerçevelerine de uygulanabilir.

Yüz bulma, birçok uygulamanın (yüz tanıma ve takip etme, yüz ifade analizi, cinsiyet ve yaş tespiti vb.) ilk basamağı olması nedeniyle üzerinde çok çalışılmış bir alandır. Yüzün doğru bulunması sonraki aşamalar için büyük önem taşımaktadır. Örneğin, çoğu robotik uygulamalarda, yüz bulma ve tanıma işlemlerini otomatik olarak yapan modüllerin hazır bulunması gerekmektedir. Ayrıca, günümüzde hızla yaygınlaşan telefonlarda yüksek çözünürlük kameralarda otomatik yüz algılama ve sosyal medyada yüz algılama uygulamaları yer almaktadır.

Teknolojinin gelişmesi, depolama kapasitelerinin artması, işlemci hızları ve internetteki gelişmelerle birlikte bu alandada önemli gelişmeler kaydedilmiştir. Bunların başında araştırmalarda kullanılmak üzere PASCAL (Everingham, 2010, 2015), LFW (Huang, 2007) ve Fddb (Jain, 2010) gibi sınırlanmamış şartlarda veritabanları oluşturulmuş ve internet üzerinden paylaşılmıştır. Bu veritabanları özellikle karmaşık, deforme olabilen yüz modellerinin eğitim çalışmaları başta olmak üzere birçok uygulamada eğitim aşamasında çokça kullanılmaktadır (Felzenszwalb, 2010).

Yüz sezimi üzerinde uzun yıllardır yoğun olarak çalışılan bir problem olmasına rağmen, yüz bulma yöntemleri henüz istenen başarı düzeyine ulaşmamıştır. Yüz sezimi problemini zorlaştıran nedenler arasında, duruş (poz) farklılıkları, aydınlatma farklılıkları, çözünürlükten kaynaklanan problemler, kapanma problemi, yüz ifadeleri ve yüz aksesuarlarından kaynaklanan farklılıklar sayılabilir.

Bu tezde, Kinect algılayıcısının renk kamerasıyla, dağınık ve rasgele ev ve ofis ortamlarında elde edilmiş, poz farklılıkları, kapatma, ışıklandırma farklılıkları sınırlandırılmamış olan görüntülerden oluşan bir veritabanı üzerinde, literatürdeki en güncel ve gelişmiş yüz bulma yöntemleri test edilmiş ve karşılaştırılmıştır. Sözkonusu görüntülerin çözünürlükleri de düşük düzeydedir. Veritabanında 500 imge bulunmaktadır. Bu imgelerin ilk 300 tanesi, karşılaştırılan yöntemlerin parametrelerinin optimize edilmesi için kullanılmış, diğer 200'ü ise test için ayrılmıştır. Sonuçlar bölümünde, yüz sezimi alanındaki gelişmelere rağmen, varolan algoritmaların henüz istenen başarı seviyesine ulaşılmadığı gösterilmektedir.



2. LİTERATÜR ARAŞTIRMASI

Eğitmenli öğrenmeye dayalı yüz seziminde, eğitim aşamasında çok sayıda görüntüden yüz örnekleri ve konumları kullanılarak bir model oluşturulur; daha sonra bu model kullanılarak test görüntülerindeki yüzler tespit edilir. Eğitim verisinin sayısının artması, yüz sezimi başarımını genel olarak artırmaktadır.

Yüz sezimi yöntemleri Yang'a (2002) göre bilgi, öznitelik, şablon ve görünüm tabanlı yöntemler olarak sınıflandırılmıştır. Bilgi tabanlı yöntemler, yüzün temel öğelerinin (gözler, kaşlar, ağız, burun gibi) özelliklerini, konumlarını ve birbirleriyle uzamsal ilişkilerini temel alan yöntemlerdir (Yang, 1994). Öznitelik tabanlı yöntemler ise, yüzün aydınlatma, bakış açısı gibi farklılıklara duysız yapısal özelliklerinin çıkarılmasına dayanmaktadır (Leung, 1995). Şablon eşleme tabanlı yöntemler, önceden oluşturulmuş yüz şablonlarının test görüntülerine oturduğu bölgeleri bulmayı amaçlar (Sakai, 1969; Yuille, 1992). Görünüm tabanlı yöntemler ise, eğitim imgelerinden toplanan küçük boyutlardaki pencerelerden bir model öğrenmeyi amaçlar (Sung, 1998; Viola, 2001). Bu modeller kullanılarak test görüntülerindeki yüzlerin bulunması hedeflenir.

Yüz ve yüz özelliklerinin tespitine yönelik ilk çalışmalar 1945 yılında başlamıştır (Bledsoe, 1965). 2000'li yılların başlangıcında yapılan çalışmalarda büyük ilerlemeler kaydedilmiştir (Hjelmås, 2001; Yang, 2002). Bu algoritmalar, belirli şartlarda başarı sağlarken, sınırlandırılmamış şartlarda (in-the-wild) istenen başarı düzeyini yakalayamamaktadır. Gerçek dünya şartlarında uygulanabilir ilk yüz sezimi metodu, Viola ve Jones tarafından geliştirilmiştir (Viola, 2001). Viola-Jones algoritması hızlı ve başarılı olması nedeniyle, günümüzde dijital kameraların yazılımlarında da kullanılmaktadır.

2005'li yıllarda yüz bulma üzerine yapılan çalışmalarda Yönlü Gradyan Histogramları (HOG) öznitelikleri kullanılmıştır (Dalal, 2005). Yüz bulma çalışmalarında HOG algoritması genellikle öznitelik çıkarımı işlemlerinde kullanılmış ve çıkarılan bu öznitelikler Destek Vektör Makineleri (SVM) gibi sınıflandırıcılar kullanılarak sınıflandırma işlemine tabi tutulmuştur (Mathias, 2014; Yadav, 2015; Dalal, 2005)

Yüz üzerinde deđişmeyen bazı özel bölgelerin bulunması yüz bulma işlemini kolaylaştırmıştır. Yüz nirengi noktaları olarak adlandırılan bu bölgelerden çıkarılan özellikler karşılaştırılarak yüz bulma işlemi gerçekleştirilmiştir (Çavdarođlu, 2013; Zhu, 2012). Örnek resimler alınarak bu resimler eğitilmiş ve test verisi üzerinde örnek resimler yardımı ile yüz bulma işlemi gerçekleştirilmiştir (Kumar, 2015; Li, 2014). Örnek tabanlı yüz algılama işlemi olarak adlandırılan bu çalışmalarda yaygın olarak kullanılmaktadır.

Yüz bulma işlemlerindeki karmaşıklık ve hız sorununu çözmek için özellikle öznitelik çıkarımı ve eğitim aşamasına yönelik 2016 yılında yeni bir yöntem sunulmuştur. Normalleştirilmiş Piksel Farkı ile öznitelik çıkarımı çok hızlı ve basit bir şekilde gerçekleştirilmiştir(Liao, 2016). Deforme edilebilir parça modelleme ile yüz bulma çalışmaları etkili ve yeni bir metod olarak karşımıza çıkmaktadır (Girshick, 2012).

Bütün bu çalışmalara rağmen yüzün kapanma, poz ve aydınlatma deđişiklikleri gibi sorunlardan dolayı sınırlandırılmamış imgelerde istenilen başarı düzeyine ulaşılamamıştır. Son yıllarda bu alanda yapılan çalışmalar arasında en etkili yöntem yapay sinir ađlarıdır. Yapay sinir ađlarında derin öğrenme üzerine gerçekleştirilen çalışmalar başarılı sonuçlar vermiştir. Bunlardan en dikkat çekenleri konvolüsyonel sinir ađı ve derin konvolüsyonel sinir ađlarıdır(Garcia, 2004; Zhang, 2014). Bu çalışmalarda yüz dokuları hakkında herhangi bir varsayım yapmadan eğitim örneklerinden problem odaklı özel öznitelikler çıkartarak sonuca ulaşılır. Derin öğrenme konusu çok eski bir geçmişe sahip olmasına karşın donanımsal ihtiyaçların yeteri kadar karşılanamamış olmasından henüz çalışmalar yenidir. Fakat gelişen teknoloji ile bu ihtiyaçlar giderilmekte ve derin öğrenme sınıflandırma işlemleri, tahmin ve modelleme başta olmak üzere birçok alanda uygulanmaktadır.

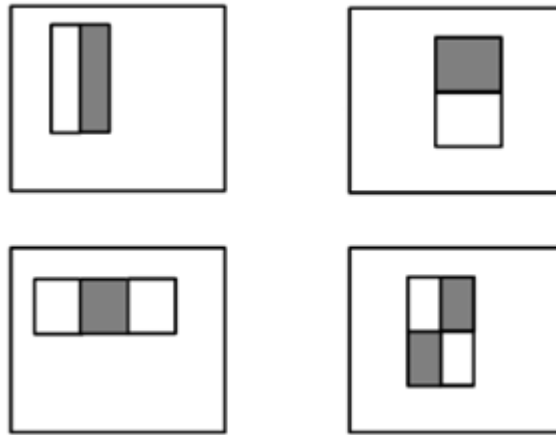
3. MATERYAL VE YÖNTEM

Bu bölümde, bu tez kapsamında Kinect'in renk kamerası ile elde edilmiş renkli imgeler üzerinde karşılaştırılan yüz sezimi yöntemleri ayrıntılı olarak anlatılmıştır.

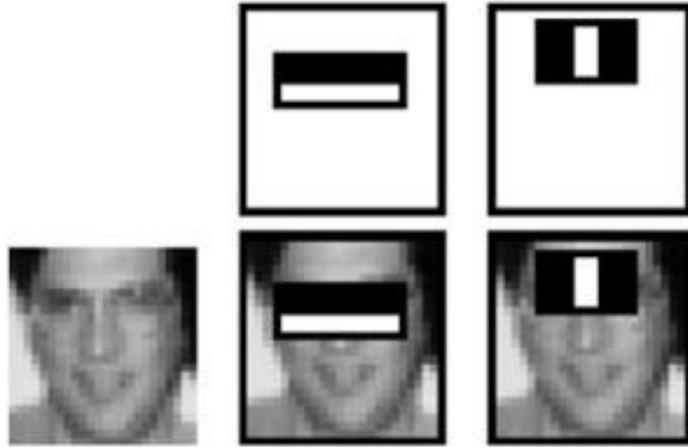
3.1. Viola-Jones Algoritması

Viola-Jones (V&J) algoritması, nesnelere hızlı ve etkin bir şekilde bulunmasını sağlaması nedeniyle en sık kullanılan yüz bulma yöntemlerinden bir tanesidir. Bu algoritma, ön cepheden alınan yüz görüntülerinden Haar dalgacık şablonlarını anımsatan dikdörtgen özneliklerin çıkarılmasına ve sınıflandırılmasına dayanmaktadır (Viola, 2004).

V&J algoritmasının eğitim aşamasında, nesne (yüz) bulunan pozitif resimler ile nesne bulunmayan negatif resimler kullanılmaktadır. Yüzün betimlenmesi için kullanılacak öznelikler, Haar filtrelerini anımsatan dikdörtgen filtreler yardımıyla çıkarılır (Şekil 3.1). Bu kutucukların en boy oranı çeşitlilik gösterebilir. Filtreler resim üzerinde kaydırılarak bütün resme uygulanır. Filtrelerin yüz üzerinde uygulama örneği Şekil 3.2'de gösterilmiştir. Filtrelerden elde edilen özellikler, siyah dikdörtgen kutucuklarda yer alan piksellerin yeğlilik değerlerinin toplamından beyaz kutucuklarda bulunan piksellerin yeğlilik değerlerinin toplamı çıkartılarak hesaplanır.



Şekil 3.1. Dikdörtgen filtreler (Viola, 2001).



Şekil 3.2. Filtreler ve uygulaması (Viola, 2001).

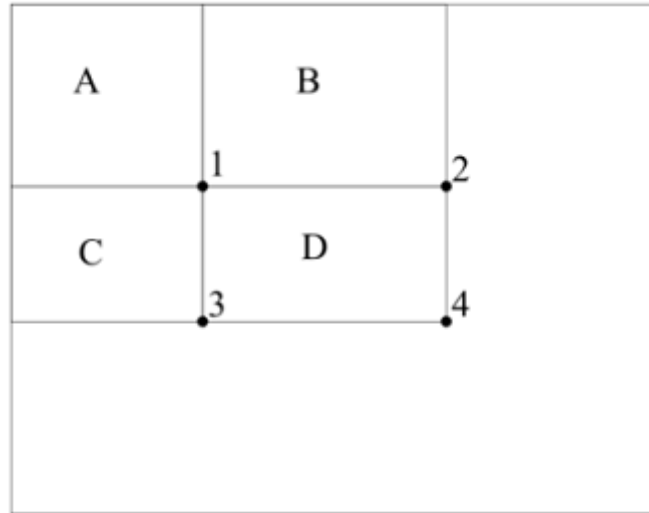
V&J algoritması diğer uygulamalara oranla hem çok daha hızlıdır, hem de doğruluk oranı yüksektir. Bunu sağlamak için üç temel fikir kullanılmıştır: Özniteliklerin çıkarılması için integral görüntülerin oluşturulması, zayıf sınıflandırıcılara dayanan Adaboost öğrenme yönteminin kullanılması, ve kaskad sınıflandırıcıların uygulanması. Bu aşamaların ayrıntıları aşağıda verilmektedir.

3.1.1. İntegral görüntü

Öznitelikleri hesaplarken her seferinde filtrelerin kutucukları içerisine düşen piksellerin yeğlilik değerlerinin toplamalarını bulmak zaman alan bir işlemdir. Bu toplamaların hızlı hesaplanmasını sağlamak için V&J algoritmasının bir adımı olarak integral görüntü tekniği geliştirilmiştir. İntegral görüntünün (x,y) konumundaki değeri, (x,y) 'nin üstünde ve solunda kalan alanın piksellerin yeğlilik değerlerinin toplamıdır. Şu formülle ifade edilir:

$$ii(x,y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (3.1)$$

Burada $ii(x,y)$ integral görüntüyü $i(x,y)$ ise orijinal görüntüyü göstermektedir. İntegral imge kullanarak herhangi bir dikdörtgensel toplam dört referansla hesaplanabilir (Şekil 3.3).



Şekil 3.3. İntegral görüntünün 1 konumundaki değeri, A dikdörtgeni içindeki piksellerin yeğlilik değerlerinin toplamıdır. 2 konumundaki değer A+B, 3 konumundaki değer A+C, 4 konumundaki değer A+B+C+D'dir. D içindeki piksellerin yeğlilik değerleri toplamı $4+1-(2+3)$ olarak hesaplanır (Viola, 2001).

3.1.2. Adaboost sınıflandırıcı

Adaboost öğrenme algoritması, birçok uygulama için etkili bir sınıflandırma sağlar. Adaboost algoritması özellikle iki sınıflı problemler için başarılı sonuçlar vermektedir.

Adaboost algoritması, yapısı gereği, hem öznitelik seçimini hem de sınıflandırıcı eğitimini sağlar. Bunu yaparken özniteliklerin eşiklenmesine dayalı basit öğrenme adımları kullanır. Bu basit öğrenme adımlarına, tek başlarına iyi sınıflandırma sağlayamadıkları için, zayıf sınıflandırıcılar denir. Adaboost, zayıf sınıflandırıcı fonksiyonları bir araya getirerek kuvvetli bir sınıflandırıcı oluşturur.

Eğitim aşamasında, yer gerçekliği bilinen pozitif ve negatif örneklerin bulunduğu eğitim kümesi ve dikdörtgen filtreler kullanılarak elde edilen bütün özniteliklerin kümesi kullanılır. AdaBoost algoritması eğitim sürecinde, bu öznitelikler arasından küçük bir altküme seçer ve her biri tek bir özniteliğe dayanan zayıf sınıflandırıcıları, $h_j(x)$, eğitir. Zayıf sınıflandırıcı, bir imge bölgesini, aşağıdaki eşikleme işlemi ile pozitif (yüz) ya da negatif (yüz değil) olarak sınıflandırır.

$$h_j(x) = \begin{cases} 1, & p_j f_j(x) < p_j \theta_j \\ 0, & \text{diğer durumda} \end{cases} \quad (3.2)$$

Burada, $f_j(x)$, x imge penceresinden çıkarılan j 'ninci öznelik, $p_j \in \{-1, +1\}$ eşitsizliğin yönünü belirleyen sayıdır. Her öznelik için, eğitim kümesi üzerinde en küçük sınıflandırma hatasını veren eşik değeri θ_j bulunur.

AdaBoost algoritması, yinelemeli bir algoritmadır ve her iterasyonda tüm öznelik kümesinin içinden bir adet zayıf sınıflandırıcı seçer. Algoritmanın detayları Çizelge 3.1'de verilmiştir(Bulut, 2016; Freund, 1999).

Çizelge 3.1. Adaboost algoritması.

1. Veri setinde yer alan n adet örnek imge, $(x_1, y_1), \dots, (x_n, y_n)$ şeklinde verilmiştir ve $y_i = 0, 1$ sırası ile negatif ve pozitif örnekleri temsil etmektedir.

2. Başlangıçtaki ağırlık değeri $(w_{1,i})$ $y_i = 0, 1$ için; (Burada m negatif, l pozitif örnek sayılarını içermektedir.)

$$w_{1,i} = \frac{1}{2m} + \frac{1}{2l}$$

3. T toplam iterasyon sayısı olmak üzere $t = 1, \dots, T$ için;

a. Ağırlıklar normalize edilir.

$$w_{t,i} = \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$$

b. Herbir j özneliği için sadece bu özelliği kullanan herbir h_j sınıflandırıcısı eğitilir.

Bu eğitim esnasında hata oranı aşağıdaki şekilde hesaplanır.

$$\epsilon_j = \sum_i w_i |h_j(x_i) - y_i|$$

c. Hesaplama sonucunda en az hata oranına sahip h_t sınıflandırıcısı seçilir.

d. Ağırlıklar güncellenir. Burada düşük hata oranı ile doğru sınıflandırma yapıldıysa $e_t = 0$ aksi durumda $e_t = 1$ olur.

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i} \quad (\beta_t = \frac{e_t}{1-e_t})$$

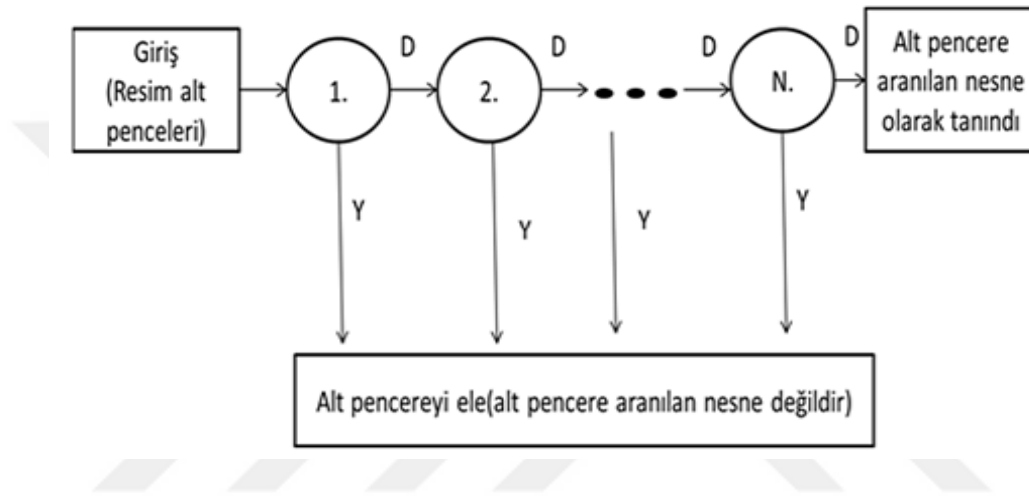
4. $\alpha_t = \log \frac{1}{\beta_t}$ olarak kabul edilirse h_x sınıflandırıcısının son durumu;

$$h(x) = \begin{cases} 1, & \sum_{t=1}^T a_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T a_t \\ 0, & \text{farklı durumlarda} \end{cases}$$

şeklinde ifade edilir.

3.1.3. Kaskat sınıflandırıcı

Bu aşamada, kuvvetli sınıflandırıcıların veriyi birbiri peşi sıra elemesine dayanan kaskat sınıflandırma yöntemi kullanılmaktadır. Kaskat sınıflandırma, daha hafif yapılı olması nedeniyle daha verimli olan sınıflandırıcıların kullanılmasını sağlayarak, algılama performansını artırırken işlem süresini azaltır. Bu sistemde sınıflandırıcılar Şekil 3.4’te görüldüğü gibi art arda sıralanarak bir kaskad yapı oluşturmaktadır.



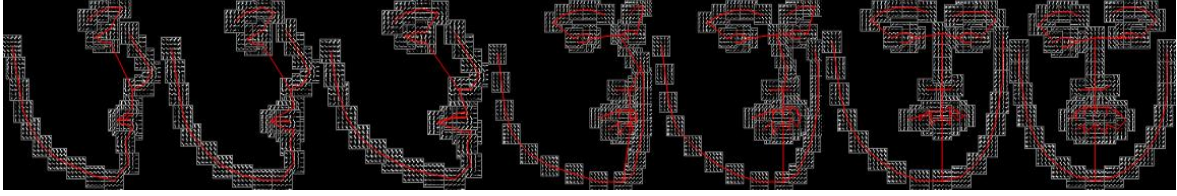
Şekil 3.4. Kaskad sınıflandırıcı (Viola, 2001).

Kaskat yapı içinde organize edilen her bir sınıflandırıcı, pencereleri “negatif” ve “pozitif” olmak üzere sınıflandırır. “Pozitif” olarak sınıflandırılan pencereler bir sonraki sınıflandırıcıya gönderilir. Herhangi bir sınıflandırıcı tarafından “negatif” olarak sınıflandırılan örnekler ise bir sonraki sınıflandırıcıya gönderilmeden “negatif” olarak etiketlenir. Bütün sınıflandırıcılardan “pozitif” olarak geçen pencereler ise “pozitif” olarak etiketlenir.

3.2. Yüz Bulma, Poz Tahmin Etme Ve Yüz Nirengi Noktalarını Bulma

Zhu ve Ramanan (2012), genellikle birbirinden ayrı olarak elen alınan yüz bulma, poz tahmin etme ve nirengi noktalarını belirleme problemlerini bir arada çözmeyi amaçlayan bir yüz bulma yöntemi geliştirmişlerdir. Bu yöntemde, ortak bir parça havuzu kullanan karma ağaç yapıları kullanılmıştır. Her bir yüz nirengi noktası yüzün bir parçası olarak modellenmiş ve bakış açısından kaynaklanan topolojik farklılıklar global karışım modelleri

ile ifade edilmiştir. Farklı pozlar için nirengi noktalarının yerel özelliklerini ve uzamsal ilişkilerini kodlayan farklı modeller oluşturulmuştur (Şekil 3.5).



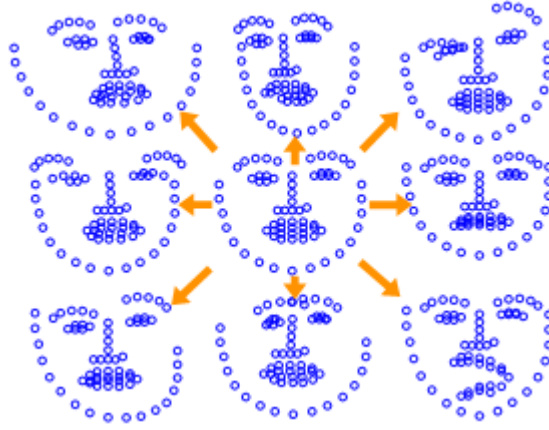
Şekil 3.5. Poz durumuna göre oluşturulan modeller (Zhu, 2012).

Yüzün nirengi noktalarına dayanan önceki çalışmalar, optimize edilmesi masraflı olan yoğun bağlantılı elastik çizgeler kullanmışlardır. Bu bölümde bahsedilen çalışmada ise karma ağaç yapıları kullanılmıştır. Bu sayede, dinamik programlama ile kolaylıkla optimize edilebilen bir model elde edilmiştir (Şekil 3.6).

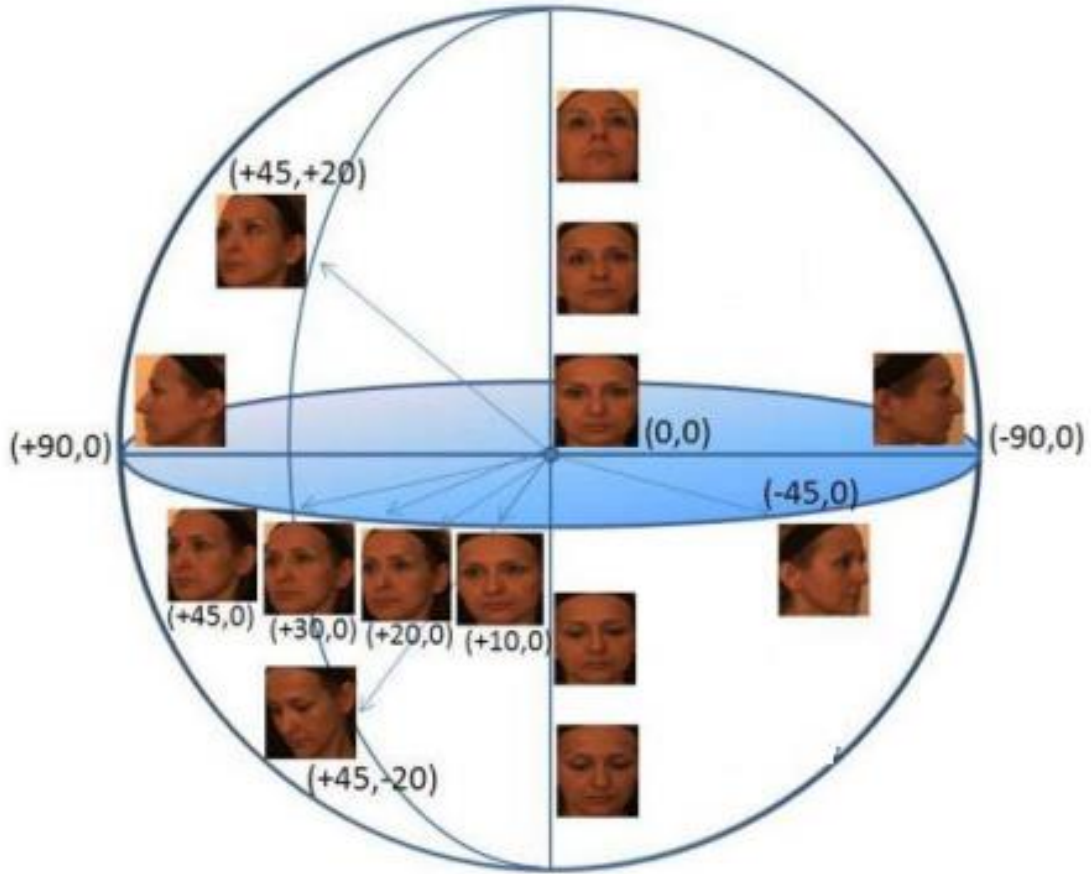
3.2.1. Karma ağaç modeli

Model, bir parça havuzunu paylaşan karma ağaçlara dayanmaktadır. Yüzün her bir nirengi noktası bir parça olarak modellenmektedir. Her ağaç çizgesinin düğümleri nirengi noktalarıyla, çizgenin kenarları ise nirengi noktaları arasındaki komşuluk ilişkileriyle tanımlanır. Belli bir ağaç modeli ile gözlem imgesinin ne kadar uyduğu, görünüm skoru ve şekil skorunun toplamıyla hesaplanır. Görünüm skoru, her bir modelle ilişkilendirilen parçaların görüntüden çıkarılan nirengi noktalarındaki parçalarla ne kadar benzeştiğinin (örneğin HOG öznitelikleri kullanarak) ağırlıklı toplamıdır. Şekil skoru ise, ağaç modelinde birbirine bağlı olan nirengi noktalarının, imgede birbirlerine göre olan konum değişikliklerinin ağırlıklı toplamıdır. Bu toplamdaki her bir terim birbirine bağlı olan nirengi noktaları arasında olduğu varsayılan bir yayın ne kadar gerildiğinin bir ifadesi olarak düşünülebilir.

Her bir yüz pozunu için farklı bir ağaç eğitilmiş, eğitim sırasında yukarıda bahsedilen ağırlıklar öğrenilmiştir. Ağaçlar, parçalar için ortak bir öznitelik havuzu kullanmaktadır. Önyüz görüntü poz olarak 0 derece kabul edilmektedir. Buna göre yüzün görünüm açısındaki değişiklikler +90 ile -90 derece arasında sınıflandırılmıştır. Görüntünün farklı açılarda değişimleri Şekil 3.7’de gösterilmiştir (Savran, 2008).



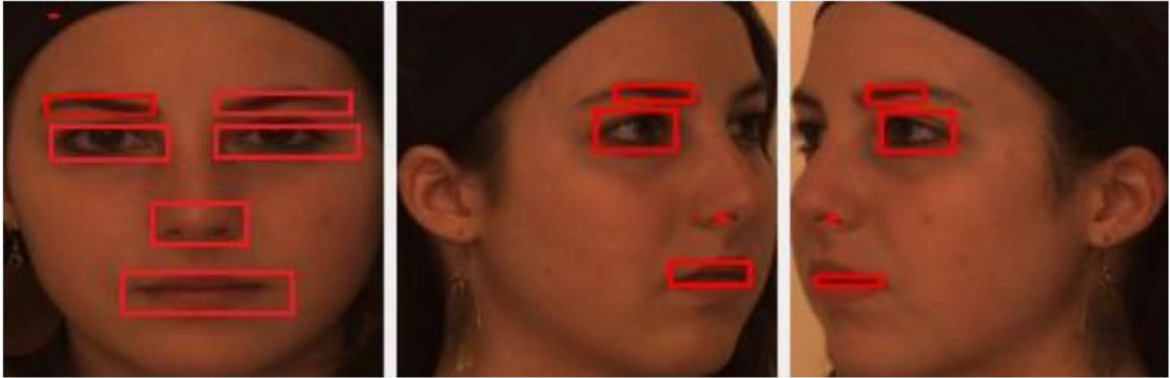
Şekil 3.6. Deformasyona uğramış modeller (Zhu, 2012).



Şekil 3.7. Bosphorus veritabanından alım açlarına örnek (Savran, 2008).

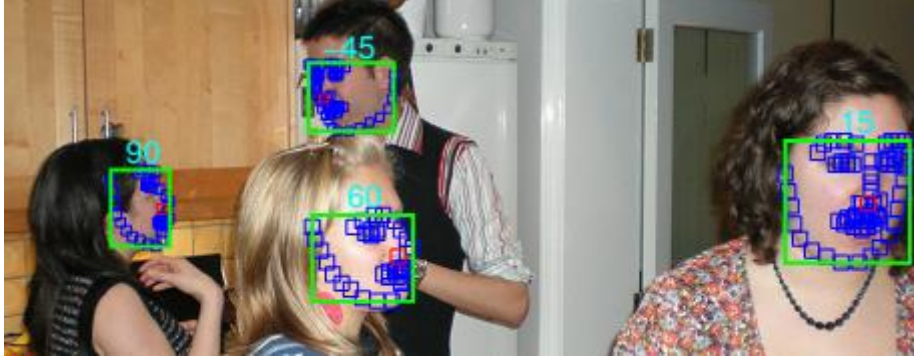
3.2.2. Nirengi noktalarını bulma

Nirengi noktaları, yüz bölgesinde değişmeyen özel bölgeler olarak tanımlanmaktadır. Nirengi noktalarına örnekler, göz merkezleri, burun ucu, ağız kenarları ve merkezidir (Şekil 3.8). Bu odak noktaları çerçevesinde nirengi noktaları poz durumuna göre belirlenir. Bir önyüzde; ikişer tane kaş ve göz bölgesi, birer tane burun ve dudak bölgesi yer alır. Zhu ve Ramanan'ın (2012) çalışmasında toplamda 68 adet nirengi noktası hesaplanır. Bunlar +45 ile -45 derece arasındaki önyüz görüntülerinde bu 68 nokta bulunabilmektedir. Profil resimlerde ise her nirengi bölgesinden birer tane yer alır ve toplamda 36 adet nirengi noktası hesaplanabilmektedir. Nirengi bölgeleri yüz tepe noktasından aşağıya doğru kaş-göz-burun-dudak bölgeleri olarak sıralanmıştır. Nirengi bölgeleri arasında geometrik ilişkiler bulunmaktadır. Nirengi bölgeleri bulunurken bu sıralama ve geometrik ilişkiler göz önüne alınarak hesaplamalar gerçekleştirilir.



Şekil 3.8. Nirengi bölgeleri (Güney, 2015).

Poz durumuna göre nirengi noktaları da dikkate alınarak 14 adet yüz modeli oluşturulmuştur (Bkz. Şekil 3.5). Bu modeller ile test verileri kıyaslanarak yüz nirengi noktaları ve poz açıları belirlenmektedir. Bulunan noktalar yardımı ile de yüz bölgesi tespit edilebilmektedir (Şekil 3.9).



Şekil 3.9. Poz açısı, nirengi noktaları ve yüz bölgesinin gösterimi (Zhu, 2012).

Nirengi noktalarında işaretlemeyen kaynaklanan hatalar genellikle göz içi mesafeye göre normalize edilerek düzeltilmektedir (Belhumeur, 2013). Burada iki gözün de görünür olduğu kabul edilir. Fakat her zaman bu doğru olmayabilir. Bu yüzden piksel hatalarını düzenlemek için yüz ölçülerine (yüzün yükseklik ve genişlik ölçülerinin ortalaması alınarak) göre ayarlama yapılır (Zhu, 2012).

3.3. Hızlı Ve Doğru Kısıtlanmamış Yüz Sezimi

Bu bölümde, rasgele poz farklılıkları, kapatmalar gibi durumların bulunduğu kısıtsız imgelerde yüz sezimi için Liao ve diğerleri tarafından geliştirilen bir yöntemden bahsedilecektir (Liao, 2016). Öznitelik olarak normalleştirilmiş piksel farkı (NPD) kullanılmıştır.

NPD özniteliklerinin optimal alt kümesini ve kombinasyonlarının belirlemek için derin karesel ağaç yöntemi önerilmiştir. Bu sayede tek bir kaskad sınıflandırıcı ile hızlı yüz tespiti yapılabilmektedir.

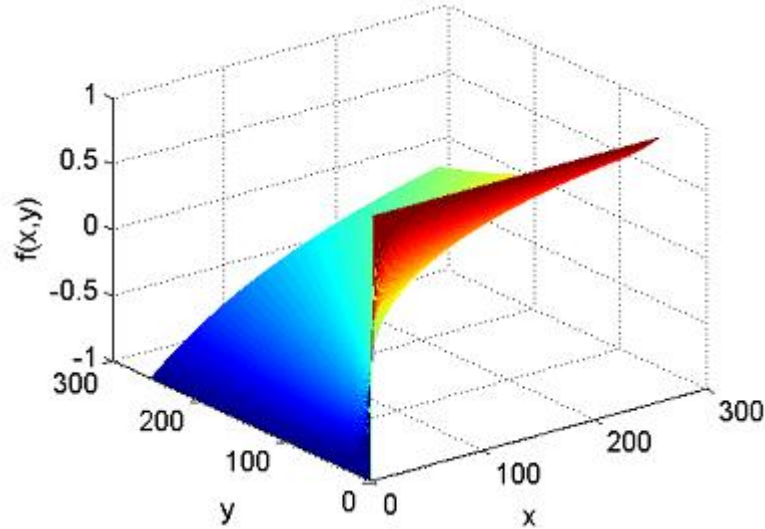
Viola Jones algortimasından hareketle (Viola, 2001), sınırlandırılmamış görüntülerde yüz tespiti için Haar benzeri özelliklere alternatif farklı karmaşık özellikler sunulmuş (Jones, 2003; Lienhart, 2002; Mita, 2005; Yan, 2008) ve kaskad sınıflandırıcı sayısı arttırılmıştır. Fakat bu çalışmalarda özellik sayısı ve kaskad sınıflandırıcı katman sayısı çok fazla artmış bu da işlem yükünün artması ve işlem süresinin uzamasına neden olmuştur. Bu bölümde verilen yöntemde (Liao, 2016) sınırlandırılmamış yüz ifadeleri için etkili bir özellik ve hızlı bir sınıflandırıcı kullanmak hedeflenmiştir.

3.3.1. Normalize edilmiş piksel farkı (NPD) öznitelikleri

NPD öznitelikleri, herhangi iki pikselin yeğlilik değerlerinin farkının toplamına oranı olarak hesaplanmaktadır. Bir imge içerisindeki x ve y pikselleri arasındaki NPD özniteliği aşağıdaki şekilde formülize edilir.

$$f(x,y) = \frac{x-y}{x+y} \quad (3.3)$$

x, y iki pikselin yeğlilik değerlerini gösterir ve $x, y \geq 0$ olduğu kabul edilir. Ayrıca $x = 0, y=0$ olması durumunda $f(x,y)=0$ sayılır.



Şekil 3.10. $f(x,y)$ NPD fonksiyonunun gösterimi (Liao, 2016).

Şekil 3.10'da görüldüğü gibi bir NPD fonksiyonu sabit ölçeklidir, sınırlıdır ve doğrusal olmayan bir yüzeyi tanımlayabilmektedir. NPD özellikleri iki piksel değeri üzerinden hesap yaptığından diğer Haar benzeri özelliklere nazaran daha hassas bir yapıya sahiptir ve yüz bulmada tıkanmalara, aydınlatma farklılıklarına karşı daha güçlüdür. Aynı zamanda düşük çözünürlüklü resimlerde de etkilidir.

NPD öznitelikleri antisimetriktir, dolayısıyla $f(x,y)$ ya da $f(y,x)$ öznitelik gösterimi için kullanılabilir. $p \times 1$ boyutunda bir vektör olarak tanımlanmış $s \times s$ büyüklüğündeki bir imge parçasında ($p = s.s$), $d = p(p-1)/2$ sayıda NPD özniteliği çıkarılır. Her bir öznitelik $1 \leq i \leq j \leq p$ piksel çifti arasında hesaplanan $f(x_i, x_j)$ ile tanımlanır. Ortaya çıkan NPD öznitelik uzayı $\Omega_{\text{npd}}(\in \mathbb{R}^d)$ ile gösterilir.

3.3.2. Derin kuadratik ağaç yapısı

NPD özniteliklerinin seçilmesi ve sınıflandırılması için, derin karesel ağaç öğrenme yöntemi kullanılmıştır. Önceki araştırmalarda (Viola, 2001) kullanılan temel ağaç yapısı ve tek bir eşik değeri ile basit sınıflandırma, farklı özellikler arasındaki etkileşimi kavrayamadığı gibi bütün bilgiyi de değerlendirememektedir. Öğrenme aşamasında derin bir ağaç yapısı ve ikinci dereceden bir eşikleme metodu kullanılarak bu sorun giderilmeye çalışılmıştır. Burada ağaç düğümünün ayrımı işlemi bir x özniteliğinin karesel bir fonksiyonunun eşiklenmesi şeklinde formüle edilmiştir:

$$(ax^2+bx+c) < t \quad (3.4)$$

a, b, c sabit değerleri ifade ederken, t eşik değerini vermektedir. Uygun katsayılarla bu eşikleme, t değerinin $[\theta_1, \theta_2]$ aralığında olup olmadığını kontrol etmektedir. Eğitim aşamasında öğrenilen θ_1 ve θ_2 eşik değerlerine göre, NPD özniteliği, üç farklı yapıyı öğrenmektedir:

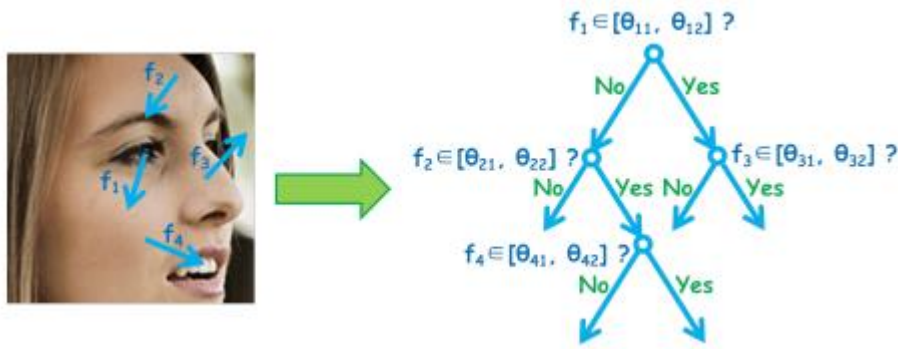
$$-1 \leq \frac{x-y}{x+y} \leq \theta < 0$$

$$0 < \theta \leq \frac{x-y}{x+y} \leq 1 \quad (3.5)$$

$$\theta_1 \leq \frac{x-y}{x+y} \leq \theta_2$$

Eğer x pikseli y pikselinden daha karanlık ise (şekil 3.11 f1) 1. Formül X pikseli y pikselinden daha parlak olduğunda ise (şekil 3.11 f2) 2. Formül uygulanır.

Şekil 3.11 f_3, f_4 'te görüldüğü gibi x pikseli ile y pikseli arasında kenar bölgeleri veya karşıtlı olmasına rağmen kutuplaşma belli olmayabilir. Bu yapılar 3. Formül ile kuadratik bölünme ile bulunabilir.



Şekil 3.11. NPD özelliklerinin derin kuadratik ağaç yapısı ile birleştirilmesi (Liao, 2016).

Karesel eşikleme kullanılarak derinliği 8 olan bir ağaç yapısı kullanılmıştır. Bu sayede, çeşitli NPD öznitelikleri yüzün içsel yapısını temsil etmek üzere optimal olarak bir araya getirilmiş olur. Ağaç yapısına örnek Şekil 3.11'de verilmiştir. Benzer bakış açısından çekilmiş yüz imgelerinin ağacın aynı yaprağında kümeleşmesini sağlamasından dolayı, bu derin karesel ağaç yapısının poz farklılıkları altında yüz sezimi için uygun olduğu öngörülmüştür.

Yazarlar, Liao'nun (2016) önerdiği bu yöntemi AFLW yüz veritabanında eğitmişler, testleri ise FDDB, GENKI, ve CMU-MIT veritabanları üzerinde gerçekleştirmişlerdir.

3.4. SVM Kullanarak HOG Tabanlı Yüz Algılama

Bu bölümde, nesne bulmada sık kullanılan öznitelikler olan Yönlü Gradyan Histogramları (HOG) ve popüler bir sınıflandırıcı olan Destek Vektör Makineleri (SVM)'ne dayanan bir yüz sezimi yöntemi betimlenmektedir. Yönlü Gradyan Histogramları yüz özniteliklerinin çıkartılmasında kullanılmış ve bu öznitelikler Destek Vektör Makineleri yardımı ile sınıflandırılmıştır.

3.4.1. Yönlü gradyan histogramı (HOG)

HOG özniteliklerinin çıkarılması için imge, hücre olarak adlandırılan dikdörtgen bölgelere ayrılır (Dalal, 2005; Kadota, 2009). Her bir bölgenin içerdiği yönlü imge gradyanlarının histogramı hesaplanır. İmge gradyanları piksellerin yeğinlik değerleri arasındaki yerel farklar hesaplanarak bulunur. HOG özniteliklerin çıkarılması için aşağıdaki adımlar izlenir:

- 1) İmge yeğinlik değerleri kullanılarak gradyan büyüklüklerinin ve yönlerinin belirlenmesi
- 2) Her bir hücre için gradyan yönelimlerinin histogramının hesaplanması
- 3) Histogramların bloklandırılarak normalizasyon uygulanması
- 4) Hücelere ait histogramların ardarda eklenmesiyle öznitelik vektörünün oluşturulması

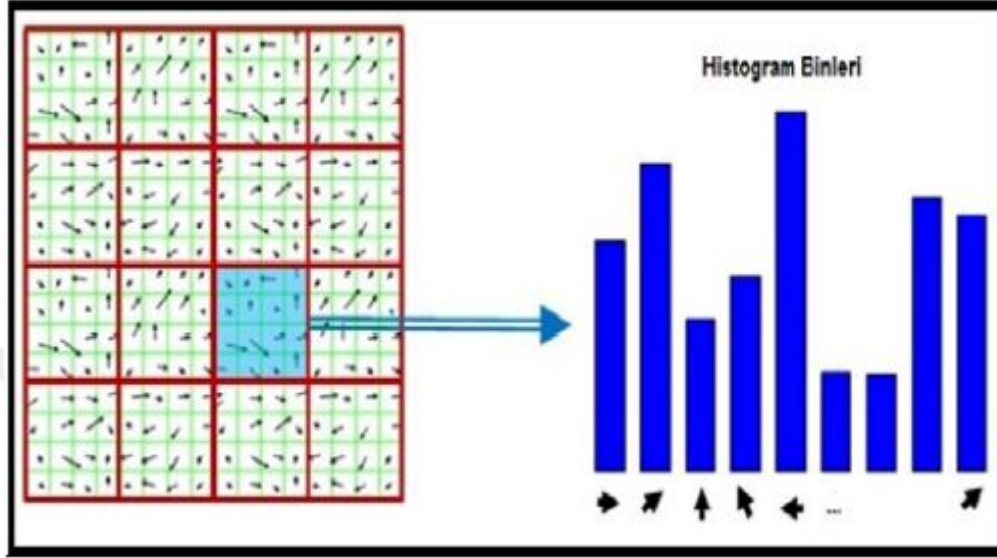
Gradyan hesaplamasında imge üzerindeki her bir pikselin yeğinlik değerlerinin yatay ve dikey yönlerdeki değişimi kullanılır baz alınır. Bu değişimler gradyanın iki boyutlu veri üzerindeki yatay $f_x(x,y)$ ve dikey $f_y(x,y)$ yönlerdeki değerlerini verir. $I(x,y)$ ifadesini (x,y) noktasındaki yeğinlik değeri olarak kabul edersek yatay ve dikey gradyan değerleri aşağıdaki şekilde ifade edilir.

$$\begin{aligned} f_x(x,y) &= I(x+1, y) - I(x-1, y) \\ f_y(x,y) &= I(x, y+1) - I(x, y-1) \end{aligned} \quad (3.6)$$

Buradan gradyan değeri (büyüklüğü) $m(x,y)$ ve yönelimini (açısını) $\theta(x,y)$ hesaplayabiliriz.

$$\begin{aligned} m(x,y) &= \sqrt{f_x(x,y)^2 + f_y(x,y)^2} \\ \theta(x,y) &= \arctan \frac{f_x(x,y)}{f_y(x,y)} \end{aligned} \quad (3.7)$$

Bulunan gradyan yönelimlerinin histogramı gradyan büyüklükleri ile ağırlıklandırılarak çıkarılır. Şekil 3.12’de 360 derecenin 9 bölgeye nicemlenmesiyle elde edilmiş histogram çıkarımına örnek verilmiştir.

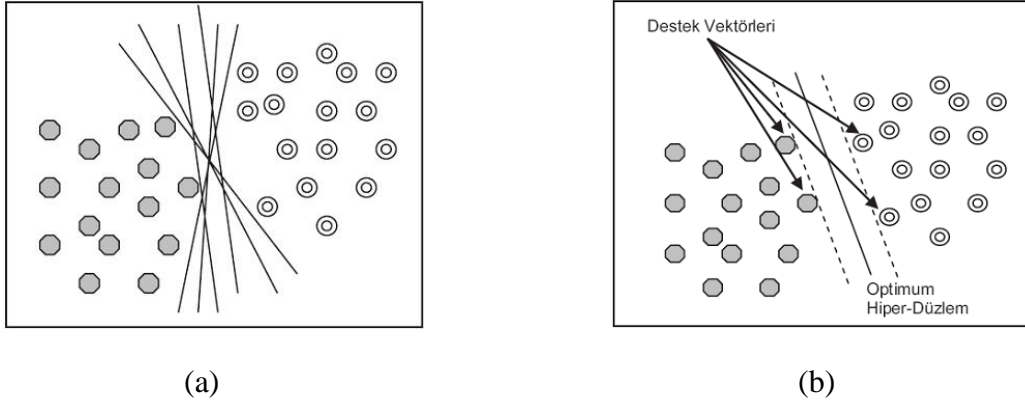


Şekil 3.12. Gradyan yönelimleri ve bir hücreye ait histogramın gösterimi.

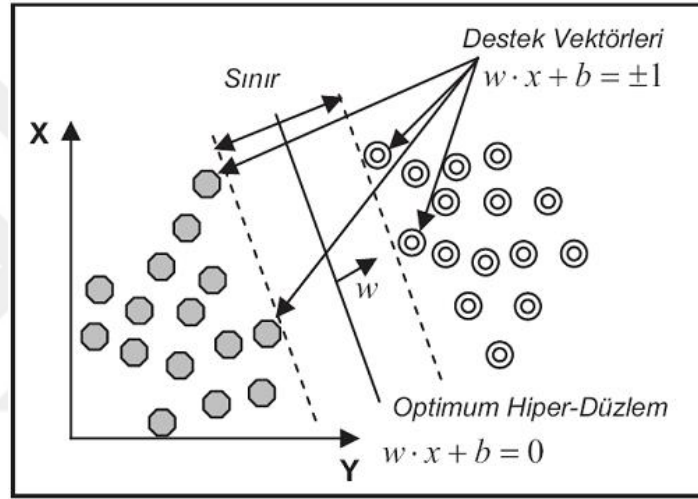
3.4.2 Destek vektör makineleri (SVM)

Destek vektör makineleri, iki sınıfa ait öznitelik vektörlerinin doğrusal sınıflandırılması için geliştirilmiştir (Vapnik, 1995, 2013). Amaç, iki sınıftan gelen verilerin doğrusal bir karar fonksiyonu ile birbirinden ayrılmasıdır. Destek vektör makineleri, çok sınıflı problemlere ve doğrusal olarak ayrılmayan özniteliklere de uygulanmak üzere geliştirilmiştir.

İki sınıflı öznitelik uzayında birbirinden ayırabilecek çok sayıda hiperdüzlem bulunabilir (şekil 3.13). Ancak SVM sınıflandırıcı, iki sınıflı, destek vektörleri arasındaki marjı maksimize edecek şekilde ayıran bir hiperdüzlem bulmak üzere eğitilir (Şekil 3.14).



Şekil 3.13. a) İki sınıflı verileri ayıran hiperdüzlemler b) Optimum hiperdüzlem.



Şekil 3.14. Doğrusal veriler için optimum hiperdüzlem oluşumu.

Optimum hiperdüzlem aşağıdaki ifadeler ile formülize edilir.

$$\begin{aligned} y_i = +1 \text{ için, } w \cdot x_i + b &\geq +1 \\ y_i = -1 \text{ için, } w \cdot x_i + b &\leq -1 \end{aligned} \quad (3.8)$$

Burada x öznitelik vektörünü, $y \in \{-1, +1\}$ sınıf etiketini, w ağırlık vektörünü ve b eğilim değerini ifade etmektedir (Osuna, 1998).

4. BULGULAR VE TARTIŞMA

Yüz bulma yöntemlerin kıyaslanması ve performanslarının değerlendirilmesi oldukça önemli bir problemdir [1]. Bu tezdeki amacımız, iç ortamlarda Kinect algılayıcısının renkli kamerasıyla çekilmiş düşük çözünürlüklü RGB imgelerinden, farklı pozlarda, kapatmalarda ve ışıklandırma koşulları altındaki insan yüzlerinin bulunmasında, güncel yöntemlerin başarımının değerlendirilmesidir.

Bu tezde açık kaynak koduna sahip 4 farklı güncel yöntem değerlendirmeye alınmıştır. Burada doğru bir değerlendirme yapabilmek için veritabanının her yöntem için sabit olması gerekmektedir. Tezin bu bölümünde, öncelikle veritabanından ve kullanılan başarımların ölçütlerinden bahsedilecektir. Ardından sırası ile yöntemlerin başarımları verilecektir.

4.1. Veritabanı

Uygulamaların test aşamasında kullanılmak üzere 300 farklı renkli imge Kinect cihazı vasıtasıyla elde edilmiştir. Görüntü seçiminde özellikle farklı pozisyonlardaki insan görüntülerinin alınmasına özen gösterilmiştir. Bütün imgelerde insanlar bulunmaktadır; ancak imgedeki insanların yüzleri görünür olmayabilir (arkası dönük insanlar, kapatma). Yüz bölgesinin görünür olduğu yerler elle dikdörtgen içine alınarak yer gerçekliği oluşturulmuştur. 300 imgede toplam 489 görünür yüz bölgesi işaretlenmiştir. Görüntülerde farklı pozlar, farklı aydınlatma düzeyleri ve farklı insanlar bulunmaktadır. Veritabanındaki imgelere örnekler Şekil 4.1’de verilmiştir.



Şekil 4.1. Veritabanından örnek imgeler.

4.2. Başarı Ölçütleri

Birçok farklı başarımlı değerlendirme metodu olsa da her metot için ortak olan yüz bulma yöntemlerinin iki tür hatası kullanılır: Bunlar yanlış negatif ve yanlış pozitif hatalarıdır. Bu hatalara ek olarak doğru pozitif oranı da hesaplanır.

Yanlış negatif (false negative); İmgedeki bir yüzün yöntem tarafından bulunamadığı, kaçırıldığı hata şeklindedir.

Yanlış pozitif (false positive); Yüz olmayan bir bölgede yöntemin yüz bulduğu hata şeklindedir.

Doğru pozitif (true positive); İmgedeki bir yüzün yöntem tarafından doğru bir şekilde bulunmasıdır.

Yanlış negatif hata sayısı (YN), yöntemin kaçırdığı yüz sayısı, Yanlış Pozitif hata sayısı (YP), yöntemin yüz olmayan yerde bulunduğu yüzlerin sayısı ve Doğru Pozitif sayısı (DP) yöntemin doğru yerde bulunduğu yüzlerin sayısı olarak tanımlanabilir.

Performans değerlendirmesinde başarı ölçütleri olarak, Hassasiyet (Precision) ve Geri Çağırma (Recall) değerleri kullanılmıştır. Hassasiyet ve geri çağırma değerleri şu şekilde hesaplanır:

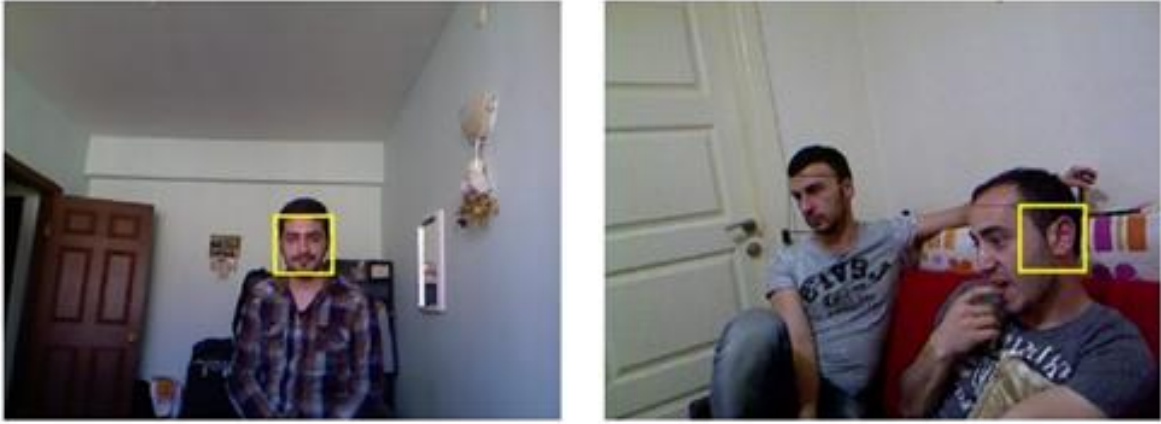
$$\begin{aligned} \text{Hassasiyet} &= \text{Doğru Tahmin Sayısı} / (\text{Doğru Tahmin Sayısı} + \text{Yanlış Tahmin Sayısı}) \\ &= \text{DP} / (\text{DP} + \text{YP}) \end{aligned}$$

$$\begin{aligned} \text{Geri Çağırma} &= \text{Doğru Tahmin Sayısı} / \text{Toplam Yüz Sayısı} \\ &= \text{DP} / (\text{DP} + \text{YN}) \end{aligned}$$

4.3. Yöntemlerin Uygulanması

Yukarıda bahsedilen **300 adet imge** ayrı ayrı 4 farklı yöntemle giriş verisi olarak kullanılmıştır. Test işlemi sonucu elde edilen kuşatan kutular ve yöntemlerin parametreleri kayıt altına alınmış ve sonuçlar değerlendirilmiştir. Yanlış negatif, yanlış pozitif ve doğru pozitif sayıları bulunarak hassasiyet ile geri çağırma hesaplanmıştır.

Doğruluk oranlarının tespiti amacıyla test görüntülerinde yer alan **toplam 489 yüz bölgesinin** herbirinin yer gerçekliği (ground truth), elle kuşatan kutuları belirlenerek elde edilmiştir. Bu gerçek kuşatan kutular, yüz sezimi yöntemleri ile elde edilen kuşatan kutular ile kıyaslanmıştır. Gerçek kuşatan kutu ile bulunan kuşatan kutunun %50 üzerinde örtüşmesi durumu doğru yüz tespiti olarak kabul edilmiştir ve doğru pozitif olarak adlandırılmıştır. Örtüşme oranı iki dikdörtgenin kesişimlerinin birleşimlerine oranı olarak hesaplanmaktadır. %50 oranın altına kalan veya hiç örtüşmeyen sonuçlar yanlış olarak değerlendirilmiş ve yanlış pozitif değerine atanmıştır. Örnek imgelerde (Bkz.Şekil 4.1) yöntemin bulunduğu yüz sarı kutucuk ve yer gerçekliği siyah kutucuk ile gösterilmektedir. Şekil 4.2 (a)'da doğru pozitif, Şekil 4.2 (b)'de ise yanlış negatif ve yanlış pozitif örnek verilmiştir.



(a)

(b)

Şekil 4.2. Yüz bulma işlemi sonucu: a) Doğru bir şekilde tespit edilen yüz görüntüsü, b) Hatalı tespit edilen ve kaçırılan yüz görüntüsü.

4.3.1. Viola&Jones algoritması

Viola&Jones Algoritması ile elde edilen sonuçlar Çizelge 4.1’de verilmiştir. Tabloda, birleştirme eşiği, birbirine yakın olarak bulunan kuşatan kutuları tek bir kutuda birleştirmeye karar vermek için kullanılan eşik değerini vermektedir.

Çizelge 4.1. Viola&Jones algoritması ile elde edilen sonuçlar.

Birleştirme eşiği	DP sayısı	YP sayısı	Kaçırma	Hassasiyet	Geri çağırma
0	288	7755	201	0.0358	0.5890
1	275	1571	214	0.1490	0.5624
2	270	214	219	0.5579	0.5521
3	256	71	233	0.7829	0.5235
4	244	28	245	0.8971	0.4990
5	235	18	254	0.9289	0.4806
6	225	10	264	0.9574	0.4601
7	220	5	269	0.9778	0.4499
8	210	4	279	0.9813	0.4294
9	203	3	286	0.9854	0.4151
10	194	3	295	0.9848	0.3967

Şekil 4.3'te ise Viola&Jones algoritmasının örnek resimler üzerinde verdiği sonuçlar gösterilmektedir. Sarı kutucuklar algoritmanın yüz olarak işaretlediği bölgeleri göstermektedir.



Şekil 4.3. V&J Algoritmasının sonuçlarına örnekler.

Viola&Jones Algoritması yüz bulmada özellikle önyüz algılama çalışmalarında çok etkili bir yöntemdir. İmgedeki yüzlerin boyutlarından az etkilenmektedir. Ancak profilden çekilen yüzlerde ya da yüzün düzlemsel olarak döndüğü durumlarda başarısız olmaktadır. Veritabanı üzerindeki geri çağırma oranı yüzde 60'ın üzerine çıkamamaktadır.

4.3.2. Poz tahmin etme, yüz nirengi noktalarını bulma

Poz Tahmini, Nirengi Noktalarının Tespiti İle Yüz Bulma yönteminin sonuçları Çizelge 4.2’de verilmiştir. Eşik değeri, test edilen yüz bölgesinin modelle uyumluluğunun ölçüsünü vermektedir. Eşik değeri küçüldüğünde doğru kabuller artmakta, bununla birlikte yanlış kabul oranı yükselmektedir. Bu yöntem daha çok 80x80 boyutlarından daha yüksek ebatlara sahip yüz bölgesinin tespiti için kullanıldığından, veritabanımızdaki imge çözünürlüğünde yüksek sonuçlar vermemiştir. Çizelge 4.3 ve Çizelge 4.4’te, sırasıyla, imge boyutlarının 2 ve 3 katına çıkarılması sonucu elde edilmiş sonuçlar verilmiştir. İmge boyutunun artırılması, sonuçların iyileşmesini sağlamıştır. Ancak yanlış kabul oranı da yükselmiştir.

Çizelge 4.2. Orijinal imge boyutlarında elde edilen sonuçlar.

Birleştirme eşiği	DP sayısı	YP sayısı	Kaçırma	Hassasiyet	Geri çağırma
-0.65	21	6	468	0.7778	0.0429
-0.7	23	8	466	0.7419	0.047
-0.8	38	29	451	0.5672	0.0777

Çizelge 4.3. İmge boyutlarının iki katına çıkarıldığı durumda elde edilen sonuçlar.

Birleştirme eşiği	DP sayısı	YP sayısı	Kaçırma	Hassasiyet	Geri çağırma
-0.65	221	64	268	0.7754	0.4519
-0.7	237	112	252	0.6791	0.4847
-0.8	281	470	208	0.3741	0.5746

Çizelge 4.4. İmge boyutlarının üç katına çıkarıldığı durumda elde edilen sonuçlar.

Birleştirme eşiği	DP sayısı	YP sayısı	Kaçırma	Hassasiyet	Geri çağırma
-0.65	237	402	252	0.3708	0.4846
-0.7	245	924	244	0.2096	0.501
-0.8	271	3754	218	0.0673	0.5542

Şekil 4.4’te poz tahmini, nirengi noktalarının tespiti ile yüz bulma metodunun örnek resimler üzerinde verdiği sonuçlar gösterilmektedir.



Şekil 4.4. Nirengi noktalarını bulma yöntemi ile tespit edilen yüzler.

4.3.3. Hızlı ve doğru kısıtlanmamış yüz sezimi

Hızlı ve doğru kısıtlanmamış yüz sezimi yöntemi ile elde edilen sonuçlar Çizelge 4.5'te verilmiştir. Tabloda, birleştirme eşiği, birbirine yakın olarak bulunan kuşatan kutuları tek bir kutuda birleştirmeye karar vermek için kullanılan eşik değerini vermektedir. Bu yöntem, sınırlandırılmamış yüzleri çok hızlı ve etkin bir şekilde, hedef yüz boyutlarında herhangi bir sınırlama da olmaksızın bulabilmektedir. Hız ve doğruluk olarak diğer uygulamalara nazaran daha iyi olmasına karşın yanlış bulma oranı biraz yüksektir.

Çizelge 4.5. Hızlı ve doğru kısıtlanmamış yüz sezimi yöntemi ile elde edilen sonuçlar.

Birleştirme Eşiği	DP sayısı	YP sayısı	Kaçırma	Hassasiyet	Geri çağırma
0	63	209	426	0.2316	0.1288
0.2	248	382	241	0.3937	0.5072
0.4	245	399	244	0.3804	0.501
0.5	245	405	244	0.3769	0.501
0.6	243	421	246	0.366	0.4969
0.8	238	471	251	0.3357	0.4867
1	275	1718	214	0.1635	0.5624

Şekil 4.5'te kısıtlanmamış yüz sezimi yönteminin örnek resimler üzerinde verdiği sonuçlar gösterilmektedir.



Şekil 4.5. Kısıtlanmamış yüz sezimi ile tespit edilen yüzler.

4.3.4. SVM kullanarak HOG tabanlı yüz algılama

SVM Kullanarak HOG Tabanlı Yüz Algılama ile elde edilen sonuçlar Çizelge 4.6'da verilmiştir. Bu yöntem küçük ve sabit ebatlardaki (24x32) yüzler için tasarlanmış, oldukça yavaştır. Çok yüzlü resimlerde işevselliğini yitirmektedir. Doğruluk oranı düşük hata oranı fazladır. Zor test verilerinde başarısızdır.

Çizelge 4.6. SVM kullanılarak HOG tabanlı yüz bulma ile elde edilen sonuçlar.

DP sayısı	YP sayısı	Kaçırma	Hassasiyet	Geri çağırma
38	262	451	0,1266	0,0777

Şekil 4.6'da HOG algoritmasının örnek resimler üzerinde verdiği sonuçlar gösterilmektedir.



Şekil 4.6. HOG algoritması ile tespit edilen yüzler.

5. SONUÇ VE ÖNERİLER

Bu çalışmada dört güncel yüz sezimi yönteminin zor bir veritabanı üzerindeki başarımları karşılaştırılmıştır. Veritabanındaki imgelerdeki yüzler, ışıklandırma farklılıkları ve kapatma gibi pekçok zorluk barındırmaktadır. İnsanların ve yüzlerinin pozları üzerinde hiçbir kısıt bulunmamaktadır. Sahneler karmaşık bir yapıya sahiptir. Bu sebeple, bu tezde karşılaştırılan algoritmalar sınırlı başarıma ulaşabilmişlerdir.

Viola&Jones Algoritması, yüz sezimi problemi için çok yaygın olarak kullanılan bir yöntemdir. Öznitelik çıkarımında kullanılan Haar benzeri filtrelerin boyutlarının değişkenlik arz etmesinden dolayı imgedeki yüzlerin boyutlarından ve çözünürlüklerinden çok az etkilenmektedir. İçerisinde barındırdığı integral imge özelliği ve kaskad sınıflandırıcı sayesinde çok hızlı sonuç veren bir yöntemdir.

Ancak Viola&Jones algoritmasının başarı oranı yüzün düzlemsel olarak döndüğü durumlarda düşüktür. Viola&Jones Algoritması tek bir eşik değeri ile basit sınıflandırma işlemi yapmasından dolayı farklı özellikler arasındaki etkileşimi kavrayamamakta ve bütün bilgiyi birarada değerlendirememektedir. Bu yüzden zor profil resimlerinde yüz bulmada istenilen seviyede değildir. Öğrenme aşamasında ikinci dereceden bir eşikleme metodu kullanılarak, basit eşikleme metodunun eksik kaldığı doğrusal olmayan durumlarda da yüz bulmadaki etkinlik artırılabilir gibi, özellikler arasındaki etkileşimi de kullanan derin bir ağaç yapısı benzeri bir sınıflandırıcı ile de desteklenmesi halinde yeterli başarı oranını sağlayabileceği öngörülerek hızlı ve doğru kısıtlanmamış yüz sezimi yöntemi geliştirilmiştir. Bu yöntemle ilgili değerlendirme aşağıda verilmektedir.

Hızlı ve doğru kısıtlanmamış yüz sezimi yöntemi sınırlanmamış yüzleri çok hızlı ve etkin bir şekilde, hedef yüz boyutlarında herhangi bir sınırlama da olmaksızın bulabilmeye yönelik olarak geliştirilmiştir. NPD özellikleri iki piksel değeri üzerinden hesap yaptığı için diğer Haar benzeri özelliklere kıyasla daha gürbüz bir yapıya sahiptir ve yüz bulmada tıkanmalara, aydınlatma farklılıklarına karşı daha güçlüdür. Aynı zamanda düşük çözünürlüklü resimlerde de etkilidir.

Öğrenme aşamasında derin bir ağaç yapısı ve ikinci dereceden bir eşikleme metodu kullanılarak, temel ağaç yapısı ve tek bir eşik değeri ile basit sınıflandırmanın sağlayamadığı, farklı özellikler arasındaki etkileşim kavranmakta ve bilgi bütünsel olarak değerlendirilebilmektedir.

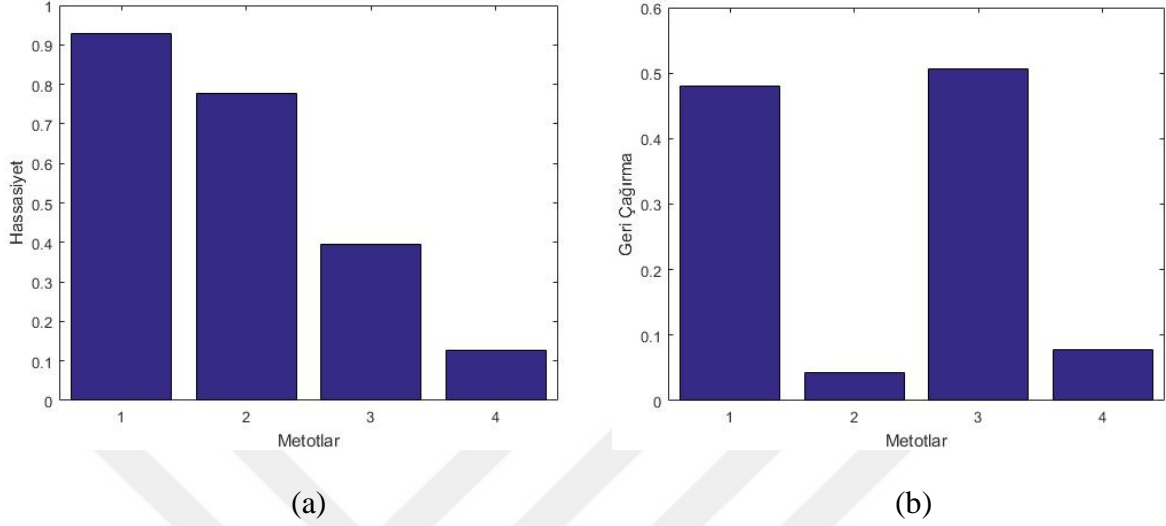
Hız ve doğruluk olarak diğer uygulamalara kıyasla daha iyi olmasına karşın yanlış bulma oranı yüksektir. NPD özniteliklerinin yapısı, karmaşık arka plana sahip imgelerdeki yüz dışındaki detaylı yapıları yüz olarak tanımasına neden olmaktadır.

Poz Tahmini, Nirengi Noktalarının Tespiti İle Yüz Bulma Yöntemi, önyüzlerde olduğu gibi profil yüzlerde de başarı sağlamaktadır. Ancak, 80x80 piksel boyutlarından daha yüksek ebatlara sahip yüz bölgesinin tespiti için tasarlandığından, imgedeki yüzlerin boyutlarından etkilenmektedir. Test verilerindeki imge boyutunun artırılması, sonuçların iyileşmesini sağlamıştır. Diğer uygulamalara kıyasla profil resimlerinde daha iyi bir başarı oranı yakalamasına karşın oldukça yavaş çalışan bir yöntemdir.

Bu yöntemde, farklı pozlar için nirengi noktalarının yerel özelliklerini ve uzamsal ilişkilerini kodlayan farklı modeller oluşturulmuştur. Ayrıca karma ağaç modeli ile her bir yüz nirengi noktası için parça modeller oluşturularak, genel bir yüz modeli elde edilmiştir. Karma ağaç yapısından dolayı farklı pozları ve elastik deformasyonları modelleyebilmektedir. Zorlu profil resimlerinin ağırlıkta olduğu, elastik deformasyonların çoğunlukla yer aldığı imgelerde nirengi noktalarını bulmak, poz tahmin etmek ve yüz bölgesini belirlemek için elverişli bir yöntem olarak kullanılabilir. Ancak, yüz nirengi noktalarının betimlenebilmesi için yüksek çözünürlüklü girdi imgelerine ihtiyaç duymakta ve çok sayıda ağaç yapısını değerlendirdiği için yavaş çalışmaktadır.

Bu tez kapsamında kullanılan **SVM Kullarak HOG Tabanlı Yüz Algılama kodları**, küçük ve sabit ebatlardaki (24x32) yüzler için tasarlanmıştır ve oldukça yavaş çalışmaktadır. Çok yüzlü resimlerde işevselliğini yitirmektedir. Doğruluk oranı düşük hata oranı fazladır. Zor test verilerinde başarı sağlayamamıştır.

Yapılan testler sonucunda, elde edilen Hassasiyet ve Geri Çağırma değerlerinin grafiği Şekil 5.1’de verilmiştir.



Şekil 5.1. Sonuç değerlerin grafik gösterimi: a) Hassasiyet, b) Geri çağırma.

Metotlar:

1. Viola&Jones Algoritması
2. Poz Tahmin Etme, Yüz Nirengi Noktalarını Bulma
3. Hızlı Ve Doğru Kısıtlanmamış Yüz Sezimi
4. SVM Kullarak HOG Tabanlı Yüz Algılama

Bu tezde, karşılaştırılan dört güncel yöntem içerisinde, en etkili olan yöntemin Viola&Jones Algoritması olduğu sonucuna varılmıştır. Fakat önyüz resimlerdeki başarısını, düzlem içi dönmeye uğramış yüz pozları için sağlayamamıştır. Yüz bulma yöntemleri üzerinde yoğun olarak çalışılmış olsa da, zorlu imgelerde ve zorlu pozlar için de hızlı bir şekilde yüksek başarı oranı sağlamaya yönelik yöntemler geliştirilmesi hala açık bir çalışma alanıdır.

KAYNAKLAR DİZİNİ

- Belhumeur, P. N., Jacobs, D. W., Kriegman, D. J., & Kumar, N. (2013). Localizing parts of faces using a consensus of exemplars. *IEEE Transactions on pattern analysis and machine intelligence*, 35(12), 2930-2940.
- Bledsoe, W., & Chan, H. (1965). A Man-Machine Facial Recognition System-Some Preliminary Results, Technical Report PRI 19A. *Panoramic Research, Inc., Palo Alto, California*.
- Bulut, F. (2016). Determining Heart Attack Risk Ration Through AdaBoost/AdaBoost ile Kalp Krizi Risk Tespiti. *Celal Bayar Üniversitesi Fen Bilimleri Dergisi*, 12(3), 459-472.
- Çavdaroğlu, G. Ç. (2013). *Sanal gerçeklikte ilinti operatörleri geliştirilerek yüz tanıma analizi*. YTÜ Fen Bilimleri Enstitüsü,
- Dalal, N., & Triggs, B. (2005). *Histograms of oriented gradients for human detection*. Paper presented at the Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on.
- Everingham, M., Eslami, S. A., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A. (2015). The pascal visual object classes challenge: A retrospective. *International journal of computer vision*, 111(1), 98-136.
- Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A. (2010). The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88(2), 303-338.
- Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. (2010). Object detection with discriminatively trained part-based models. *IEEE Transactions on pattern analysis and machine intelligence*, 32(9), 1627-1645.
- Freund, Y., Schapire, R., & Abe, N. (1999). A short introduction to boosting. *Journal-Japanese Society For Artificial Intelligence*, 14(771-780), 1612.
- Garcia, C., & Delakis, M. (2004). Convolutional face finder: A neural architecture for fast and robust face detection. *IEEE Transactions on pattern analysis and machine intelligence*, 26(11), 1408-1423.
- Girshick, R. B., Felzenszwalb, P. F., & McAllester, D. (2012). Discriminatively trained deformable part models, release 5.
- Güney, M., & Arıca, N. (2015). Desen Tabanlı İlgi Bölgesi Tespiti. *Deniz Bilimleri ve Mühendisliği Dergisi*, 5(1), 94-106.

KAYNAKLAR DİZİNİ (devam)

- Hjelmås, E., & Low, B. K. (2001). Face detection: A survey. *Computer vision and image understanding*, 83(3), 236-274.
- Huang, G. B., Ramesh, M., Berg, T., & Learned-Miller, E. (2007). *Labeled faces in the wild: A database for studying face recognition in unconstrained environments*. Retrieved from
- Jain, V., & Learned-Miller, E. (2010). Fddb: A benchmark for face detection in unconstrained settings. *University of Massachusetts, Amherst, Tech. Rep. UM-CS-2010-009*, 2(7), 8.
- Jones, M., & Viola, P. (2003). Fast multi-view face detection. *Mitsubishi Electric Research Lab TR-20003-96*, 3(14), 2.
- Kadota, R., Sugano, H., Hiromoto, M., Ochi, H., Miyamoto, R., & Nakamura, Y. (2009). *Hardware architecture for HOG feature extraction*. Paper presented at the Intelligent Information Hiding and Multimedia Signal Processing, 2009. IHH-MSP'09. Fifth International Conference on.
- Kumar, V., Namboodiri, A., & Jawahar, C. (2015). *Visual phrases for exemplar face detection*. Paper presented at the Proceedings of the IEEE International Conference on Computer Vision.
- Leung, T. K., Burl, M. C., & Perona, P. (1995). *Finding faces in cluttered scenes using random labeled graph matching*. Paper presented at the Computer Vision, 1995. Proceedings., Fifth International Conference on.
- Li, H., Lin, Z., Brandt, J., Shen, X., & Hua, G. (2014). *Efficient boosted exemplar-based face detection*. Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
- Liao, S., Jain, A. K., & Li, S. Z. (2016). A fast and accurate unconstrained face detector. *IEEE Transactions on pattern analysis and machine intelligence*, 38(2), 211-223.
- Lienhart, R., & Maydt, J. (2002). *An extended set of haar-like features for rapid object detection*. Paper presented at the Image Processing. 2002. Proceedings. 2002 International Conference on.
- Mathias, M., Benenson, R., Pedersoli, M., & Van Gool, L. (2014). *Face detection without bells and whistles*. Paper presented at the European Conference on Computer Vision.
- Ming-Husan Yang, D. J. K., and Narendra Ahuja. (January 2002). Detecting Faces in Images: A Survey IEEE transaction on pattern analysis and machine intelligence, vol.24 no.21.

KAYNAKLAR DİZİNİ (devam)

- Mita, T., Kaneko, T., & Hori, O. (2005). *Joint haar-like features for face detection*. Paper presented at the Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on.
- Osuna, E. E. (1998). *Support vector machines: Training and applications*. Massachusetts Institute of Technology,
- Sakai, T., Nagao, M., & Fujibayashi, S. (1969). Line extraction and pattern detection in a photograph. *Pattern recognition, 1*(3), 233-248.
- Savran, A., Alyüz, N., Dibekliolu, H., Çeliktutan, O., Gökberk, B., Sankur, B., & Akarun, L. (2008). *Bosphorus database for 3D face analysis*. Paper presented at the European Workshop on Biometrics and Identity Management.
- Sung, K.-K., & Poggio, T. (1998). Example-based learning for view-based human face detection. *IEEE Transactions on pattern analysis and machine intelligence, 20*(1), 39-51.
- Vapnik, V. (1995). *The nature of statistical learning theory* Springer New York Google Scholar.
- Vapnik, V. (2013). *The nature of statistical learning theory*: Springer science & business media.
- Viola, P., & Jones, M. (2001). *Rapid object detection using a boosted cascade of simple features*. Paper presented at the Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on.
- Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International journal of computer vision, 57*(2), 137-154.
- Yadav, R. P., Senthilarasu, V., Kutty, K., & Ugale, S. P. (2015). Implementation of robust HOG-SVM based pedestrian classification. *International Journal of Computer Applications, 114*(19).
- Yan, S., Shan, S., Chen, X., & Gao, W. (2008). *Locally assembled binary (LAB) feature with feature-centric cascade for fast and accurate face detection*. Paper presented at the Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on.
- Yang, G., & Huang, T. S. (1994). Human face detection in a complex background. *Pattern recognition, 27*(1), 53-63.
- Yang, M.-H., Kriegman, D. J., & Ahuja, N. (2002). Detecting faces in images: A survey. *IEEE Transactions on pattern analysis and machine intelligence, 24*(1), 34-58.

KAYNAKLAR DİZİNİ (devam)

Yuille, A. L., Hallinan, P. W., & Cohen, D. S. (1992). Feature extraction from faces using deformable templates. *International journal of computer vision*, 8(2), 99-111.

Zhang, C., & Zhang, Z. (2014). *Improving multiview face detection with multi-task deep convolutional neural networks*. Paper presented at the Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on.

Zhu, X., & Ramanan, D. (2012). *Face detection, pose estimation, and landmark localization in the wild*. Paper presented at the Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on.

