

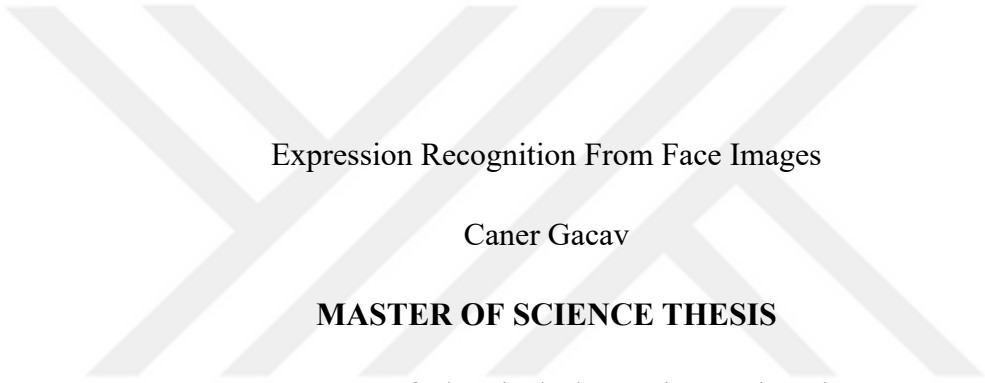
Yüz İmgelerinden Duygu Tanıma

Caner Gacav

YÜKSEK LİSANS TEZİ

Elektrik Elektronik Mühendisliği Anabilim Dalı

Mayıs 2018



Expression Recognition From Face Images

Caner Gacav

MASTER OF SCIENCE THESIS

Department of Electrical Electronics Engineering

May 2018

Yüz İmgelerinden Duygu Tanıma

Caner Gacav

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: Doç. Dr. Kemal Özkan

Mayıs 2018

ONAY

Elektrik Elektronik Mühendisliği Anabilim Dalı YÜKSEK LİSANS öğrencisi Caner Gacav'ın YÜKSEK LİSANS tezi olarak hazırladığı “**Yüz İmgelerinden Duygu Tanıma**” başlıklı bu çalışma, jürimizce lisansüstü yönetmeliğin ilgili maddeleri uyarınca değerlendirilerek oybirliğiyle kabul edilmiştir.

Danışman : Doç. Dr. Kemal Özkan

Yüksek Lisans Tez Savunma Jürisi:

Üye : Doç. Dr. Kemal Özkan

Üye : Dr. Öğr. Üyesi Hasan Serhan Yavuz

Üye : Dr. Öğr. Üyesi Cihan Topal

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

Prof.Dr. Hürriyet ERŞAHAN
Enstitü Müdürü

ETİK BEYAN

Eskişehir Osmangazi Üniversitesi Fen Bilimleri Enstitüsü tez yazım kılavuzuna göre, Doç. Dr. Kemal Özkan danışmanlığında hazırlamış olduğum “**Yüz İmgelerinden Duygu Tanıma**” 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. 31/05/2018

Caner Gacav

ÖZET

Yüz ifadesi tanıma, insan-bilgisayar etkileşimi ve davranış analizi gibi birçok konuda uygulaması olan popüler bir yapay görme problemidir. Birçok bilgisayarlı görü uygulamasında olduğu gibi yüz ifadesi tanıma da aydınlatma ve kontrast farklılıkları problemi zorlaştıran etkenlerdir. Özellikle yüzün düzlemsel olmayan yapısı, burun ve çene gibi çıkıntılı bölgeler ile göz çukurları gibi girintili alanlar yüz üzerindeki ışık dağılımının farklı olmasına yol açmaktadır. Yüz ifadesi tanımadaki problemlerden bir diğeri de yüz tespit yöntemlerinin çok-ölçekli çalışmalarından dolayı buldukları yüzlerin hizalanmamış olmasıdır. Bu durum bazı yüz bölgelerinin farklı yüz bölgelerinden gelen vektörlerle karşılaştırılmasına neden olacağından tanıma başarısını olumsuz yönde etkilemektedir. Bu çalışma yüz ifadesini tanıma problemini çözmeye odaklanan iki çalışmadan oluşmaktadır. İlk çalışmada, yüz imgesinden elde edilen nirengi noktaları kullanılarak elde edilen özniteliklerden hangilerinin yüz ifadesi tanıma da daha etkili olduğu incelenmektedir. Bir yüzde tespit edilen nirengi noktalarının ikili kombinasyonları arasındaki uzaklıklardan bir vektör oluşturulur. Kullanılan öznitelikler, yüzün nötr ve belirgin bir ifadeye sahip olduğu hallerinden elde edilen uzaklık vektörlerindeki değişimdir. Elde edilen öznitelik vektörü, yüz ifadelerini tanıma da fayda sağlamayacak elemanları da içermektedir. Ardışık ileri öznitelik seçimi uygulanarak bu öznitelik vektöründeki en faydalı elemanlardan oluşan bir altküme belirlenmiştir. İkinci çalışmada ise, birçok nesne tespit probleminde başarılı sonuçlar veren yönelimli gradyan histogramlarının HOG kontrast farklılıklarına daha duyarlı bir versiyonu olan Felzenszwalb HOG FHOG'un yüz ifadesi tanıma problemindeki başarımı irdelenmiştir. Bununla beraber hizalanma probleminin neden olduğu performans düşümlerine engel olmak için de nirengi noktaları kullanılarak hizalama yapılması önerilmiştir.

Anahtar Kelimeler: yüz ifadesi tanıma, öznitelik seçimi, ardışık ileri öznitelik seçimi, Felzenszwalb'ın yönelimli gradyan histogramı, Genişletilmiş Cohn-Kanade veri seti, destek vektör makineleri

SUMMARY

Facial expression recognition is a popular computer vision subject that has many applications such as human-computer interaction and behaviour analysis. As for many computer vision problems, lighting and contrast differences increase the difficulty of the problem. Especially the non-planar structure of the face, protruding regions such as the nose and chin and recessed regions such as eye sockets cause variations in lighting. Another problem with facial expression recognition problems is that the multi-scale detection methods do not align the faces accurately. This leads to comparing features that are extracted from different facial regions, which degrades performance. In this work, to solve facial expression recognition problem, it was focused on two different solutions. In the first study, it is investigated the effectiveness of features derived from facial landmarks in facial expression recognition. Distances between two combinations of facial landmarks constitute a distance vector. Features we use are the changes in the distance vectors extracted from expressive and neutral states of the face. The obtained feature vector contains elements that are relatively useless in expression recognition. By applying forward sequential feature selection, a subset of the most effective elements is formed. The chosen features are classified using a multi-class support vector machine. In the second study, the performance of Felzenszwalb HOG (FHOG) features at facial expression recognition is investigated. Additionally, aligning with respect to the facial landmarks is proposed to prevent performance degradation due to misalignment.

Keywords: facial expression recognition, feature selection, sequential forward feature selection, Felzenszwalb Histogram of oriented gradients, Extended Cohn-Kanade dataset, Support Vector Machines

TEŞEKKÜR

Öncelikle benden desteklerini hiç bir zaman esirgemeyen aileme, daha sonra yüksek lisansa başlamamda katkısı olan ve bu süreçte bana desteklerinden dolayı Dr. Öğr. Üyesi Cihan Topal'a ve danışmanım Doç. Dr. Kemal Özkan'a teşekkür ederim.



İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	vi
SUMMARY	vii
TEŞEKKÜR	viii
İÇİNDEKİLER	ix
ŞEKİLLER DİZİNİ	xi
ÇİZELGELER DİZİNİ	xiii
1. GİRİŞ VE AMAÇ	1
2. LİTERATÜR ARAŞTIRMASI	3
3. YÖNTEM	10
3.1. İmgelerden Yüz Tespiti	10
3.1.1. Yüz Tespiti için Kullanılan Öznitelikler	10
3.1.2. Sınıflandırma Fonksiyonlarını Öğrenme	12
3.1.3. Kaskad Yöntemi	13
3.2. Nirengi Noktaları Bulabilmek için Hizalama Yöntemi	16
3.2.1. Regresörlerin Kaskadları	16
3.2.2. Kaskad İçindeki Herbir Regresörü Öğrenme	16
3.2.3. Ağaç Tabanlı Regresör	17
3.2.3.1. Şekil İnvaryant Bölünmüş Testleri	17
3.2.3.2. Düğüm Ayrımlarını Seçme	18
3.2.3.3. Öznitelik Seçimi	19
3.2.3.4. Kayıp Etiketlerin Üstesinden Gelme	19
3.3. Çok Sınıflı Destek Vektör Makinesi	19
3.4. Duygu Tanıma İçin Geometrik Öznitelikler Üzerinde Ardışık İleri Seçim Algoritması	21
3.4.1. Geometrik Öznitelik Vektörünün Oluşturulması	22
3.4.2. Ardışık İleri Seçim Algoritması	23
3.5. FHOĞ Öznitelikleri ile Duygu Tanıma	24
3.5.1. Yüz İmgeleri, Nirengi Noktaları Tespiti ve Ön İşlemler	24

İÇİNDEKİLER (devam)

3.5.2. FHOG Özniteliklerin Elde Edilmesi	25
4. BULGULAR VE TARTIŞMA	29
4.1. Extended Cohn Kanade Veri Seti	29
4.2. Ardışık İleri Seçim Algoritması Bulguları ve Tartışma	30
4.3. FHOG Bulguları ve Tartışma	32
5. SONUÇ VE ÖNERİLER	36
KAYNAKLAR DİZİNİ	37

ŞEKİLLER DİZİNİ

Şekil	Sayfa
3.1 Örnek dikdörtgen öznitelikler tespit penceresini kapsama şekilllerine bağlı olarak gösterilmiştir. Siyah dikdörtgenler altında kalan piskellerin toplamından beyaz dikdörtgenler altında kalan piksellerin toplamı çıkartılır. (A) ve (B) iki-dikdörtgen öznitelikleri, (C) üç-dikdörtgen öznitelikleri ve (D) dört-dikdörtgen özniteliklerini temsil eder.	11
3.2 (x,y) noktasındaki integral imgenin değeri, yukarıda ve solda kalan bütün piksel değerlerinin toplamına eşittir.	12
3.3 İntegral imgede bir bölgedeki değeri hesaplamak için gereken referans noktaları gösterilmiştir.	13
3.4 Adaboost tarafından birinci ve ikinci seçilmiş öznitelikler. Üst satırdaki bu iki özneliğin gösterilmiştir. Alt satırda bu öznitelikler bir eğitim yüzüne bindirilmiştir. Birinci öznitelik göz bölgesi ile üst yanak arasındaki yoğunluk farkını ölçer. Bu öznitelikten göz bölgesinin yanak bölgesine göre daha karanlık olduğu gözlemlenebilir. İkinci öznitelik ise iki göz ve arada kalan burun kaş bölgesi arasındaki geçişteki yoğunluk farklarını karşılaştırır.	14
3.5 Şekil’de kaskad tespitinin şematigi gösterilmiştir. Her bir alt pencerede sınıflandırıcı serisi uygulanmıştır. İlk sınıflandırıcı oldukça fazla yanlış negatifleri eler ve çok az bir hesaplama süresi alır. Daha sonraki tabakalarda ise bunlara ek olarak daha fazla negatif eler, ama bu da demektir ki ek olarak bir hesaplama süresi gerektirir. Sonraki ilerlemede bu şekilde evreler oluşur. Kaskad yönteminde tespit evreleri bu sistemde ve diğer sistemlerde şekildeki gibi gerçekleştirilir.	15
3.6 Dikdörtgen ile yüz tespiti algoritması ile bulunan yüzler, noktalar ile tespit edilen yüzlerdeki nirengi noktaları ifade edilmiştir. (a) ve (b) imgelerinde nötr yüz ifadeleri, (c) imgesinde kızgınlık ve (d) imgesinde üzüntü ifadesi temsil edilmiştir.	20
3.7 Şekil’de nötr yüzden duygu ifade eden yüz arasındaki geçiş gösterilmiştir. . . .	22
3.8 Şekilde sadece bir noktanın diğer noktalarla nötr ve duygu yüzünde oluşturduğu nokta çiftleri renkli çizgilerle gösterilmiştir. Bütün noktalar arasındaki uzaklıklar aynı şekilde bulunur.	23
3.9 Önerilen algoritmanın akış diyagramı.	24
3.10 Hizalanmadan kesilen yüz ile hizalanarak kesilen yüz imgeleri gösterilmiştir. .	25

ŞEKİLLER DİZİNİ (devam)

3.11	Üst sıradaki imgelerde tespit edilen yüzler kırmızı dikdörtgenler ile gösterilmiştir. Yeşil noktalar nirengi noktalarını temsil etmektedir. Tespit edilen yüzlerden kırpma, yeniden boyutlandırma ve histogram eşitleme sonucu elde edilen imgeler alt sırada gösterilmiştir.	26
3.12	Yüzden çıkarılan FHOG'ların görselleştirilmiş şekilleri gösterilmiştir.	27
4.1	Şekil'de AİS algoritması için veri setinin 10 kat çapraz doğrulamada bölünmesi gösterilmiştir.	29
4.2	Mavi barlar iki nirengi noktası arasındaki yatay uzaklıkları, kırmızı barlar ise dikey uzaklıkları temsil eder. Her bir bar ilk nirengi noktası ile başlar ve o noktanın sonu ikinci nokta ile kesik çizgiler ile bağlanır. Bu barlardaki uzunluklardaki değişimler sınıflandırma için kullanılan özniteliklerdir. Görüldüğü üzere, en az bir bar her bir duygu için ayırt edici bir şekilde kısalır ya da uzar.	30
4.3	Şekilde Ardışık İleri Seçim Algoritması ile seçilen öznitelikler kullanılarak yapılan deneylerde ilk iki sütun doğru bulunanları temsil ederken, son sütun yanlış bulunan imgeleri göstermektedir.	31
4.4	Şekilde FHOG öznitelikleri ile yapılan deneylerde ilk iki sütun doğru bulunanları temsil ederken, son sütun yanlış bulunan imgeleri göstermektedir.	33

ÇİZELGELER DİZİNİ

<u>Çizelge</u>	<u>Sayfa</u>
4.1 CK+ veri kümesindeki Sınıflar ve sınıflara ait örnek sayıları	29
4.2 Ardışık İleri Seçim Algoritması ile Seçilen Öznitelikler kullanılarak CK+ verisetinde duygu tanıma karışıklık matrisi.	32
4.3 FHOG CK+ verisetinde duygu tanıma karışıklık matrisi.	34
4.4 Geometrik, HOG ve FHOG özniteliklerinin CK+ veri setinde karşılaştırılması. .	34



1. GİRİŞ VE AMAÇ

Bireylerarası iletişimde sözlü ifadeler kadar jest ve mimikler de önemli rol oynar. Bu tür iletişimde en önemli işaretlerden birisi şüphesiz duygusal mod hakkında bilgi içeren yüz ifadeleridir. Özellikle insan davranış analizi, insan-makine etkileşimi ve pazar araştırmaları gibi konularda bireylerin yüz ifadeleri üzerinden duygusal ve psikolojik durum analizi yapılması giderek yaygınlaşmaktadır. Bu tür çalışmalarda veri toplamak ve ölçüm yapmak için görüntü işleme teknikleri kullanılması araştırmacıların işini kolaylaştırmaktadır.

Duygu tespiti sözlü iletişimi destekleyen oldukça önemli bir parçadır. Yüzselsel ifadeler tarafından kişiye ait psikolojik ve duygusal durumların analiz edilmesi dünyada insan davranış analizi ve insan-bilgisayar iletişimi çalışmalarında oldukça geniş bir yer kapsamaya başlamıştır (Ekman ve Rosenberg, 2012; Bartlett vd., 2003). Yüzselsel ifade verileri ve otomatikleştirilmiş bilgisayarla görü yöntemleri, bu konu hakkında daha az uğraşı ile çalışmaların ilerlemesinde katkıda bulunulmuştur (Zeng vd., 2009; Rehman vd., 2013). Teknolojinin ilerlemesiyle görü sistemleri insanların bile zorlukla algılayabileceği duyguları ve ifadeleri görü sistemleri ile olanak sağlamıştır.

Geometrik ve görünüm tabanlı öznitelikler yüzselsel ifadeler için kullanılan ortak özelliklerdendir. Bu çalışmada geometrik tabanlı uzaysal öznitelikler ile görünüm tabanlı olan Felzenszwalb HOG (FHOG) öznitelikleri üzerinde odaklanılmıştır. Bu iki farklı öznitelik tipleri iki ayrı çalışma olarak incelenmiştir. Geometrik öznitelikler üzerinde boyut indirgeme için seçim algoritması yapılmış ve sonrasında seçilen öznitelikler ile yüz ifadesi bulunmaya çalışılmıştır. FHOG öznitelikleri ise yüzün tamamından öncelikle hizalama yapılacak daha sonra yüzden bu görünüm tabanlı öznitelikler çıkarılarak yüz ifadesi bulunmaya çalışılmıştır.

Uzaysal öznitelikler yüzdeki nirengi noktalarının kombinasyonları oluşturularak yer değiştirmeleri sonucu oluşan uzaklık farkları öznitelik olarak kullanılmıştır. Bu ikili kombinasyonlardan gelen oldukça fazla uzaysal özniteliklerden, bazıları potansiyel olarak daha tanımlayıcı ve bazıları az tanımlayıcıdır. Bu özniteliklerin tanımlayıcı olan artıklardan arınmış alt kümesi ile bir sınıflandırıcı sağlamak için önceki çalışmalara dayanan bir seçim yapılabilir. Örneğin; Facial Action Coding System (FACS) olarak adlandırılan yüzdeki hareketleri kodlama sistemi yüzselsel ifadeleri tespit etmek için öznitelikler arasından bir alt küme oluşturulmasını sağlar (Ekman ve Rosenberg, 2012). Diğer bir yaklaşım ise açık bir şekilde boyut indirgeme uygulamadır (Huang vd., 2014; Ekenel ve Sankur, 2004) ya da

sınıflandırıcının öznitelik seçiminin üstesinden gelmesini sağlamaktır (Chen ve Chen, 2015). Bu yöntemlere alternatif olarak bu yöntemde tanımlayıcı öznitelik alt kümesi seçebilmek için bir seçim algoritması uygulanmıştır.

FHOG öznitelikleri ise, Dalal ve Triggs'in önerdiği HOG'tan (Dalal ve Triggs, 2004) türetilen daha düşük boyutlu Felzenszwalb HOG (FHOG)'un (Felzenszwalb vd., 2009) yüzdeki imgelerde ifade tanıma amaçlı başarımı seçim algoritması uygulanmadan incelenmiştir.



2. LİTERATÜR ARAŞTIRMASI

Ryu vd., 2017, çalışmalarında yüz ifadelerini tanımak için yeni bir yüz tanımlayıcısı olan Yerel Yönlü Üçlü Deseni (YYÖD) tanıtmışlardır. YYÖD, yön bilgisini ve üçlü deseni kullanarak duygu ile ilgili öznitelikleri verimli bir şekilde kodlar. Bu özelliklerin kullanılmasının sebebi, kenar-tabanlı metodların pürüzsüz alanlardaki zayıflığının üstesinden gelmek ve aynı zamanda kenar olan alanlardaki kenar desenlerinin gürbürlüğünün avantajlarından yararlanmaktır. Diğer histogram tabanlı yüzü farklı bölgelere ayırarak örneklemeyi eşit homojen bir şekilde yaparken; bu yöntemde, yüz tanımlayıcıları oluşturulurken 2-kademeli ızgara kullanılır. Bu aynı zamanda duygu ile ilgili olan öznitelikleri farklı ölçeklerde örneklemeye olanak sağlar. Deneyler yapılırken 6 farklı veri-setinde çapraz-doğrulama ile Destek Vektör Makineleri (DVM) kullanılarak gerçekleştirilmiştir.

Çiğdem Turan ve Kim Man Lam, 2014, göz ve ağız bölgelerinden Yerel Faz Nicelemesi (YFN) ve Piramit HOG (PHOG) tanımlayıcılarını kullanarak öznitelik çıkarımı yapmışlardır. Kanonik Korelasyon Analizine dayalı bir öznitelik birleştirme yapmışlardır. Deneyler DVM ile Genişletilmiş Cohn-Kanade (CK+) veriseti üzerinde gerçekleştirilmiştir.

Chen ve Chen, 2015, görünüm tabanlı ve geometrik öznitelikler kullanmışlardır. Görünüm tabanlı öznitelikler Üç Dikey Düzlem'den gelen Yönlü Gradyan Histogramları (HOG - TOP) ile temsil edilirken, geometrik öznitelikler ise kendi içinde iki kategoriye bölünerek katı ve katı-olmayan değişimli öznitelikler ile temsil edilmiştir. Bu iki çeşit özniteliklerin optimal kombinasyonlarını bulabilmek için Çoklu-Kernel Öğrenme yöntemi uygulanmıştır. Son olarak bu çoklu-kerneller ile DVM'de eğitilip test edilmiştir. Deneyler Genişletilmiş Cohn-Kanade (CK+) verisetinde gerçekleştirilmiştir.

Arı vd., 2011, duygu problemini irdeleyip, gerçek-zamanlı olarak çalışabilen bir yöntem önermişlerdir. Bu yöntemde şaşkınlık, kızgınlık, mutluluk, üzülmeye, korku, iğrenme ve yalın ifadeyi tanıyabilen bir yöntemdir. Önerdikleri yöntemde, çok-çözünürlüklü aktif şekil modeline dayalı takipçi ile nirengi noktaları takip edilmekte ve ardından nokta konumlarından ve yüzün belli bölgelerindeki değişimlerden çıkarılan üst seviye öznitelik vektörlerinin duygusal ifadelerle ait öznitelik vektörlerine olan uzaklıkları kullanılarak sınıflandırma yapılmıştır. Deneyleri 5 özne üzerinde gerçekleştirmişlerdir.

Huang vd., 2012, hızlandırılmış gürbüz öznitelikler (SURF) kullanarak elde ettikleri öznitelik vektörlerini normalize etmişler ve olasılık dağılım fonksiyonlarını hesaplamışlardır. Elde edilen vektörler arasındaki uzaklık Kullback-Leiber ıraksaması kullanılarak hesaplanmış ve sınıflandırma ağırlıklı çoğunluk oylaması ile gerçekleştirilmiştir. Deneyler Japanese Female Facial Expression (JAFPE) verisetinde gerçekleştirilmiştir.

Sarawagi ve Arya, 2013 çalışmalarında renk normalizasyonunun vurgu yapmışlar ve yüz özniteliklerini çıkarmak için öznitelik çıkarmada etkin olan Yerel İkili Desen (LBP) kullanmışlardır. Sınıflandırma için DVM'den faydalanmışlardır. Deneyler, kendileri tarafından oluşturulan toplamda 10 (4 kız 6 erkek) Hintli öznenin oluşturduğu verisetinde gerçekleştirilmiştir.

Akakin ve Sankur, 2012, işaret yörüngelerini etkin bir tespitçi ile bularak, bu yörüngelerden bağımsız bileşke analizi (ICA) kullanılarak öznitelik çıkarımı yapılmıştır. Ayrıca, zaman-uzamsal yüz küpüne hizalanmış ve normalize edilmiş yüzler istifleterek, yüzdeki duygunun evrimi de yakalanmıştır. Son olarak duygu tanımlayıcıları bu prizmadan ve ICA özniteliklerinden oluşturulmuştur. Altı temel duyguda sınıflandırma yapılmıştır. Sınıflandırma birden fazla sınıflandırıcı kullanılarak tespit edilmiştir. Deneyler CK verisetinde gerçekleştirilmiştir.

Hong ve Cong, 2014, çalışmalarında aktif şekil modeli (ASM) ile öznitelik çıkarımı yapmışlardır. Sınıflandırma için DVM üzerinde iyileştirme içeren yeni bir yöntem kaba küme DVM adı altında bir yöntem önerilmiştir. Bunlara ek olarak, kaba küme teorisinde, elde edilen özniteliklerde seçim ve sınıflandırma yapabilmek için boyut indirgeme algoritması tanıtmışlardır. Alakasız ve yanlış olan öznitelikler filtrelenmiştir. Sınıflandırma DVM algoritması üzerinde yapılmıştır. Deneyler JAFPE verisetinde gerçekleştirilmiştir.

Huang vd., 2014, kötü hizalama ve büyük varyasyonlarınların duygu tanımda iki önemli problem olduğuna değinmişlerdir. Bu yüzden bu iki problemi çözebilmek için kanonik korelasyon metodunu tekrar gözden geçirmişlerdir. Sebebi ise bu iki problemten kaynaklanan duygu tanıma problemini, yanlış bulmaya ve yanlış hizalamalara karşı gürbüz bir hale getirmektir. İlk başta, öznitelik olarak görünüm tabanlı olan LBP'yi sunmuşlardır. Daha sonra ise video klibin kanonik alt-uzayını oluşturmak için geçici ortogonal bölgesi korunmuş izdüşümü önermişlerdir. Böylece, duygu tanımda hareketten kaynaklanan değişimleri yalayabilmek amaçlanmıştır. Deneyler Ck+ ve MAHNOB-HCI verisetlerinde gerçekleştirilmiştir.

Lucey vd., 2010, aktif şekil modelini kullanarak yüzdeki nirengi noktalarını tespit etmişlerdir. Daha sonra bu noktaları benzerlik – normalize edilmiş şekilde (SPTS) ve kanonik görünüm (CAPP) özniteliklerini hesaplamada kullanmışlardır. Sınıflandırma için DVM kullanılmıştır. Deneyler CK+ verisetinde gerçekleştirilmiştir.

Suk ve Prabhakaran, 2015, gerçek-zamanlı olarak mobilde çalışabilen duygu tanıma yöntemi önermişlerdir. Önerdikleri yöntem sınıflandırıcı için DVM kullanmıştır. 6 temel duyguyu ve normal yüzü tanımak için ağız bölgesinin durumunu kontrol etmişlerdir. Duygu tanıma için öznitelik çıkarımında aktif şekil modeli kullanılarak elde edilen noktalar kullanılmıştır. Öznitelikler, duygu ifade eden ve normal yüz arasındaki bu noktaların yer değişim farklarından oluşturulmuştur. Deneyler CK+ verisetinde gerçekleştirilmiştir.

Shan vd., 2017, otomatik duygu tanıma için Derin Konvolüsyonel Sinir Ağları (CNN) üzerinde durmuşlardır. Sebebi CNN'in daha derinlerdeki özniteliklere sahip bir yapıya sahip olmasıdır. Sistem 4 modülden oluşturulmuştur. Bunlar; Giriş modülü, ön-işlem modülü, tanıma modülü ve çıkış modülüdür. Deneylerinde JAFFE ve CK+ verisetlerini kullanmışlardır. Ayrıca karşılaştırmak için k-en yakın komşuluk (KNN) algoritmasını kullanmışlardır.

Zhou ve Shi, 2017, hareket ünitelerine seçici öznitelik haritalarının faydasını ve oluşumunu kurmaya çalışmışlardır. Bunu tranfer yöntemiyle öğrenilmiş CNN üzerinde kurmuşlardır. Image-Net üzerinde önceden eğitilmiş bir ağı, Karolinska Directed Emotional Faces (KDEF), Radboud Faces Database (RaFD) CK+ verisetleri üzerinde duygu tanıma görevine tranfer etmişlerdir. Daha CNN'i genel imgelerde eğiterek daha yüksek tabakalar oluşturup yüzsel hareket ünitelerine seçici hale getirilmiştir. Sebebi öznitelik seçiminin sistemin gürbüz olmasında önemli rol oynamasıdır.

Ding vd., 2017, ilk önce en tepe duygu çerçevesini yakalamak için çift LBP (DLBP)'yi önermişler. Önerilen DLBP daha küçük boyutta ve başarılı bir şekilde tanıma zamanını düşürmüştür. Ayrıca, LBP'deki ışık varyasyonlarının üstesinden gelmek için logarithm-laplace (LL) önerilmiştir. Böylece sistemin ışık varyasyonlarında daha gürbüz olması amaçlanmıştır. En son olarak ise, LBP tabanlı Taylor öznitelik deseni (TFP) ve genişletilmiş Taylor'ı önermişlerdir. Sebebi Taylor öznitelik haritasından etkili olan yüzsel öznitelikleri elde etmektir. Deneyler JAFFE ve CK+ veri setinde gerçekleştirilmiştir.

Sulistiyo, 2017, çalışmalarında, radyal tabanlı fonksiyonunu (RBF) yüz hareket sistemindeki bir harekete uydurarak işaretçi transferini referans alan bir metod olarak ele almışlardır. Farklı radyal fonksiyonları kullanmışlardır. Bunlar; RBF Gaussian, Inverse Quadratic, Inverse Multiquadratic ve Multiquadratic olarak adlandırılan fonksiyonlardır. En

iyi sonucu RBF Gaussian da almışlardır. Deneyler kendilerinin tek bir kameradan alınan görüntülerden oluşan verisetinde gerçekleştirilmiştir.

Swinkels vd., 2017, çalışmalarında gerçek zamanlı duygu tanıma üzerine yoğunlaşmışlardır. Bu yüzden yüzden elde edilen 19 öznelik noktanın hareketlerine dayalı bir duygu tanıma sistemi yapmışlardır. Noktaların bulunduğu bölgeler ağız, gözler, kaşlar ve burun bölgelerindedir. Bu noktaları elde etmek için Ensemble of Regression Trees (ERT) adından bir algoritma sunmuşlardır. Bu noktalar elde edildikten sonra 12 farklı uzaklık hesaplanmıştır. Son olarak öznelikler elde edildikten sonra çok-sınıflı DVM'ye beslenmiştir. Deneyler CK+ verisetinde gerçekleştirilmiştir.

Ruiz vd., 2017, çalışmalarında yüz ifadesi ile ona karşılık gelen görsel öznelikleri arasındaki ilişkiyi modelleyen Yapay Sinir Ağlarını (ANN) önermişlerdir. Sınıflandırma için de yine ANN kullanılmıştır. 7 temel duygu içermektedir. İmge verisetinden elde edilen öznelik vektörleri üzerine Principle Component Analysis (PCA) uygulanarak boyut küçültme yapılmıştır. Deneyler JAFFE verisetinde gerçekleştirilmiştir.

Çam vd., 2016, çalışmalarında, yedi yüz ifadesini ayırt etmekte kullanmak üzere şekil tabanlı yeni bir öznelik çıkarma yöntemi önermişlerdir. Yüz nirengi noktaları, eğitimli birinin metodu kullanılarak tespit edildikten sonra bu noktaların video dizisi boyunca değişimlerinden üst seviye öznelikler çıkarılmıştır. Sınıflandırma olarak DVM kullanılmıştır. Deneyler CK+ verisetinde gerçekleştirilmiştir.

d. A. Fernandes vd., 2016, çalışmalarında iki geometrik tabanlı bir yaklaşım ile yüzde duygu tanıma sistemi önermişlerdir. İlk metodta insan bakışı ile bularak deneysel bir geometrik öznelik çıkarımı yapmışlardır. Diğer metodta ise Korelasyon Öznelik Seçimi uygulayarak duygu sistemi geliştirmişlerdir. Sınıflandırıcı olarak DVM kullanılmıştır. Deneyler CK+ verisetinde gerçekleştirilmiştir.

Mousavi vd., 2016, yüz imgelerinden duyguları CNN'i ve Lateral Inhibition Pyramidal Neural Network (LIPNet)' i eğitmişlerdir. Daha sonra dekonvolüsyon ile CNN'de öğrenilen öznelikleri işlemek için görselleştirmişlerdir. Ayrıca LIPNet'in iç temsilini görselleştirmek için yeni bir mekanizma tanıtmışlardır. Bu sinir ağlarını eğitmek için CK+ veriseti kullanılmıştır. Daha sonra eğitilen yüzsel ifade tanıma mekanizmasını JAFFE verisetinde genellemek için test etmişlerdir. Son olarak invaryansın prensiplerini, fazlalıkları ve filtrelemenin sinir ağlarında genel olarak duygu tanıma nasıl etkili olduğunu açıklamışlardır.

Zhao vd., 2017, çalışmalarında yeni bir metod olarak işaretçi manifoldu olarak adlandırılan bir yöntem sunmuşlardır. Böyle yüzdeki işaretçi bilgilerinin kendi başlarına kompetitif olan bir performans elde edildiğini göstermişlerdir. Deneyler CK+ verisetinde gerçekleştirilmiştir.

Talele vd., 2016, çalışmalarında öznitelik çıkarımı için verimli LBP ve sınıflandırıcı için Yapay Sinir Ağlarını (ANN) kullanmışlardır. Verimli LBP öznitelikleri 4 farklı bloklar ele alınarak elde edilmişlerdir. Bunlar; 256x256, 128x128, 64x64 ve 32x32'lik bloklardır. 6 temel duygu ANN kullanılarak sınıflandırılmıştır. Ayrıca Genel Regresyon Sinir Ağları (GRNN) sınıflandırıcı olarak da uygulamışlardır. GRNN sistemi daha iteratif bir eğitim süreci gerektirmediği için daha hızlı bir şekilde gerçekleştirdiğinden bahsetmişlerdir. Deneyler JAFFE, CK+, Taiwanese Facial Expression Database (TAFED) ve Indian Face Database (IFD) veri setlerinden gerçekleştirilmiştir. En iyi sonucu 64x64 bloklarda elde etmişlerdir.

Afshar ve Salah, 2016, çalışmalarında, yüzü ve nirengi noktalarını video imgelerinden bulabilmek için son yıllarda çıkan 2 yönteminin kombinasyonunu kullanmışlardır. Yüz kayıtları için Procrustes yöntemini kullanmışlardır. Öznitelikleri Dense Trajectories (DT), geometrik ve Local Gabor Binary Patterns (LGBP-TOP) kullanarak elde etmişlerdir. Sınıflandırıcı için Extreme Learning Machines (ELM) kullanmışlardır. Deneyler CK+ verisetinde gerçekleştirilmiştir.

Kumar vd., 2017, duyguyu HOG özniteliklerini kullanarak yüz imgelerinden tanıyabilen gerçek zamanlı bir yöntem önermişlerdir. Ayırıcı HOG öznitelikleri çıkarabilmek için yüzde bulunan nirengi noktalarından faydalanılmıştır. Sistemi hızlandırmak için takip algoritması uygulanmıştır. Bütün yüzden HOG özniteliklerini elde etmek yerine sadece aktif olan bölgelerden öznitelikler toplanmıştır. Çıkarılan öznitelikler DVM'ye beslenip sınıflandırılma yapılmıştır. Deneyler CK+ verisetinde gerçekleştirilmiştir.

Suk ve Prabhakaran, 2015, çalışmalarında, akıllı telefonlarda gerçek zamanlı çalışabilecek yüzden duygu tanıma sistemi geliştirmişlerdir. Bunun için, gerçek zamanlı geçici video segmentasyon yaklaşımını sunmuşlardır. Sistem Finite System Machine (FSM)'i normal ifadeden tepe ifadeye giden gerçek zamanlı videoyu geçici faza segmente etmek için kullanmışlardır. FSM'nin, Lucas-Kanade'nin farklı hızlardaki durum geçişlerine adapte olabilen optik akış vektör tabanlı bir algoritmayı kullandığından bahsetmişlerdir. Sistemin herhangi bir örneklemeden kaynaklanan bir gecikme içermediğini söylemişlerdir. Sınıflandırıcı olarak DVM kullanmışlardır. Deneyler CK+ verisetinde gerçekleştirilmiştir.

Kamarol vd., 2015, çalışmalarında Spatio-Temporal Texture Map (STTM) diye adlandırılan yüzden spontane duygu tanımak için uzaysal-zamansal bir öznitelik çıkarımı önermişlerdir. Hem görünüm hem de geometrik öznitelik çıkarma yaklaşımında bulunulmuştur. Görünüm tabanlı öznitelikler olarak Volume LBP (VLBP) ve LBP Three Orthogonal Planes (LBP-TOP) ve geometrik tabanlı olarak da çok açılı ağaç tabanlı öznitelikler düşünülmüştür. Sınıflandırıcı olarak DVM kullanılmıştır. Deneyler CASME II verisetinden gerçekleştirilmiştir.

Lin Zhong vd., 2012, çalışmalarında farklı duygularda yüzdeki ortak ve özel bilgileri keşfetmeyi amaçlamışlardır. Sadece yüzdeki belli parçaların aktif olarak hareket etmekte olduğu görülmüştür. Bu parçaları bulabilmek için 2 seviyeli çok görevli seyrek öğrenme (MTSL) metodunu önermişlerdir. Bu iki seviyeli görevde her duyguda ayırt edici yüzdeki parçaları bulmaya çalışmışlardır. Bu öğrenilen öznitelikler ve eğitilmiş MTSL, CK+, GEMEP-FERA ve MMI verisetlerinde denemiştir.

Li ve Lam, 2015, çalışmalarında, yüzden duyguları tanımak için derin sinir ağlarında yeni bir teknik geliştirmişlerdir. İnsan yüz imgelerini işlerken fotometrik normalizasyon ve histogram maniplasyonu yapmışlardır. Bunu yapmalarının sebebi ışık farklılıklarında doğan problemleri gidermektir. Daha sonra öznitelikler Gabor filtreleri kullanılarak elde edilmiştir. Derin sinir ağlarına bu öznitelikleri beslemeden önce boyut indirgeme için PCA uygulamışlardır. Sinir ağlarında 1 giriş tabakası, 2 gizli tabaka ve 1 de softmax diye adlandırdıkları sınıflandırıcı içermektedir. Deneyler CK+ verisetinde gerçekleştirilmiştir.

Mistry vd., 2015, çalışmalarında, FACS'tan esinlenerek insansı robot için akıllı yüz hareketi ve duygu tanıma sistemi sunmuşlardır. 7 temel duyguya ve 18 yüzdeki hareket ünitesine odaklanmışlardır. Bağımsız Aktif Görünüm Modeli (AAM) ile rotasyon invariant öznitelik nokta detektörü olan Binary Robust Invariant Scalable Keypoints (BRISK)'i entegre ederek yeni bir şekil ve görünüm tabanlı öznitelik çıkarım metodu önermişlerdir. Sınıflandırıcı olarak sinir ağlarını kullanmışlardır. Deneyler CK+ verisetinde gerçekleştirilmiştir.

Wang vd., 2015, çalışmalarında, yüzden duygu tanımada, sınıf etiketlerinin olmaması durumunda yeni bir kontrol edilmeyen öznitelik seçim metodu önermişlerdir. Yüksek boyutlu öznitelik uzayında regresyon analiz yardımı ile özniteliklere verilen skorlar ile boyut indirgeme yapılmıştır. Böylece daha az skoru olan öznitelikler öznitelik uzayından elenerek, daha küçük boyutlu bir öznitelik vektörü ile temsil edilmiştir. 4 farklı sınıflandırıcı ile performansı ölçülmüştür. Bu sınıflandırıcılar; DVM, FDA, LPP ve NPE sınıflandırıcılarıdır. Deneyler 3 farklı verisetinde gerçekleştirilmiştir. Bu verisetleri; MMI, JAFFE ve CK+ verisetleridir.

Lee vd., 2015, çalışmalarında, iki seviye MLP metodunu önermişlerdir. Ayrıca etkin olan bu öğrenme metodunu etkin bir şekilde destekleyen alt küme ön eğitimi metodunu (SPT) önermişlerdir. SPT’de yapılan eğitim verisinin bir kısmı MLP’nin ilk seviyesinde kullanılmaktadır. İkinci seviyede ise bütün eğitim verisi kullanılarak MLP eğitilmektedir. Öznitelik olarak LBP kullanmışlardır. Deneyler CK+ verisetinde gerçekleştirilmiştir.

Zhu vd., 2015, çalışmalarında, çapraz veriseti kullanarak duygu tanıma için transfer öğrenme önermişlerdir. İki seviyeli bir eğitim süreci içerir. İlk kaynak veriden bilgiyi öğrenmek, ikincisi ise öğrenilen bilgiyi hedef veriye adapte etmektir. Bu yaklaşım da öznitelikler gabor filtrelerinden elde edilen öznitelikler, regresyon ağaç sınıflandırıcıları, Adaboost algoritması ve DVM tabanlı methodları baz alarak oluşturulmuştur. Deneyler JAFFE, FEED ve CK+ veri setleri üzerinde gerçekleştirilmiştir.

Zhang vd., 2015, çalışmalarında gerçek zamanlı 3B Hareket Ünitesi (AU) yoğunluk kararı ve yüzden duygu bulma sistemi önermişlerdir. 16 hareket tabanlı yüzsel öznitelikleri minimal-redundancy-maximal-relevance criterion tabanlı optimizasyonu kullanarak bulmuşlardır. Bu öznitelikleri Support Vector Regressor ve sinir ağlarına beslemişlerdir. Ayrıca 6 yeni adapte olabilen sınıflandırıcıyı 6 farklı duygu için birleştirmişlerdir. Deneyler Bosphorus 3D veri setinde gerçekleştirilmiştir.

Lo Presti ve La Cascia, 2015, çalışmalarında, yüzden duygu tanıma için yeni bir yaklaşım olan yüzsel ifadelerinin seri dinamik geçiş modelini önermişlerdir. Yüz İmge Tanımlayıcıları (FID)’na Linear Time Invariant (LTI) sisteminin çıkışları gibi bakmışlardır. Bu serilerin geçiş dinamiklerine Henkel Matrisleri ile temsil etmişlerdir. Bu temsiller duygu tanıma ve acı tanımada doğrulanmıştır. Deneyler CK+ veri setinde gerçekleştirilmiştir.

3. YÖNTEM

Çalışma iki farklı yöntemden oluşmaktadır. Bu yöntemler iki ayrı başlık altında incelenmiş olup ortak adımları tek bir bölüm altında anlatılmıştır. Çalışmalardan ilki geometrik öznitelik havuzundan duygu tespiti için en uygun özniteliklerin bulunduğu alt küme Ardışık İleri Seçim algoritması ile seçmeye çalışılmıştır (Gacav vd., 2017). İkinci çalışmada ise Felzenszwalb'ın önerdiği HOG'dan türetilen daha düşük boyutlu ve HOG'a kıyasla daha fazla bilgi içeren FHOG öznitelikleri ile duygu tespiti gerçekleştirilmiştir (Gacav vd., 2018). İmgelerden yüz tespiti, bulunan yüzler üzerinde nirengi noktalarının hizalanması, sınıflandırıcı için kullanılan Çok-Sınıflı Destek Vektör Makineleri iki yöntemde de ortak olduğu için aynı başlık altında anlatılmıştır. Daha sonra geometrik öznitelikler için Ardışık İleri Seçim Algoritması anlatılmıştır. En son olarak FHOG öznitelikleri ile yapılan çalışma anlatılıp iki çalışmaya ait tespit sonuçları verilmiştir.

3.1 İmgelerden Yüz Tespiti

İki yöntemin ilk adımı olan ve imgelerden yüz tespiti olan Viola ve Jones, 2001'nin nesne bulma için geliştirdiği ve yüz bulmak için kullandığı yöntemden bahsedilmiştir.

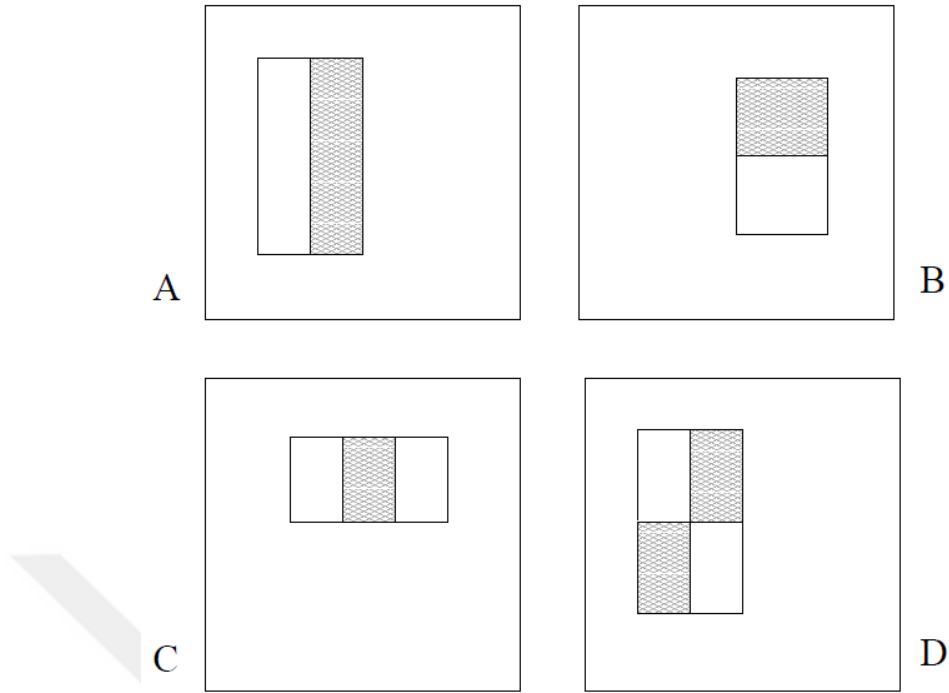
Yüz tespit algoritmasında, ilk önce kullanılan öznitelikler, sonrasında bu özniteliklerin kombinasyonu ile oluşturulan sınıflandırıcıdan, Adaboost'un bir çeşidi olan, öznitelik seçmek için de kullanılan makine öğrenme yönteminden ve son olarak güvenilir ve hızlı nesne tespiti oluşturmak için kaskad sınıflandırıcıları oluşturan yöntemden bahsedilmiştir.

3.1.1 Yüz Tespiti için Kullanılan Öznitelikler

Çalışmada Haar'ın temel fonksiyonlarını anımsatan öznitelikler kullanılmıştır. 3 çeşit öznitelik kullanılmıştır. Bunlar 2-dikdörtgen, 3-dikdörtgen ve 4-dikdörtgen hesaplamalarından gelen değer öznitelikleridir. Şekil 3.1'de dikdörtgenler gösterilmiştir.

İntegral İmge: İmgeler "integral imge" adı ile de bilinen bir temsil ile dikdörtgen öznitelikler çok hızlı bir şekilde hesaplanabilir. İntegral imgede herhangi bir x,y noktasındaki değer, yukarı ve solda kalan bütün piksellerin toplamına eşittir. Şekil 3.2

(3.1)'de integral imgedeki bir (x,y) noktasındaki toplamın hesaplanması genel olarak verilmiştir. (3.2) ve (3.3)'te ise hesaplanması açılmıştır. Bu formüllerde $ii(x,y)$



Şekil 3.1: Örnek dikdörtgen öznelikler tespit penceresini kapsama şekillerine bağlı olarak gösterilmiştir. Siyah dikdörtgenler altında kalan piksellerin toplamından beyaz dikdörtgenler altında kalan piksellerin toplamı çıkarılır. (A) ve (B) iki-dikdörtgen öznelikleri, (C) üç-dikdörtgen öznelikleri ve (D) dört-dikdörtgen özneliklerini temsil eder.

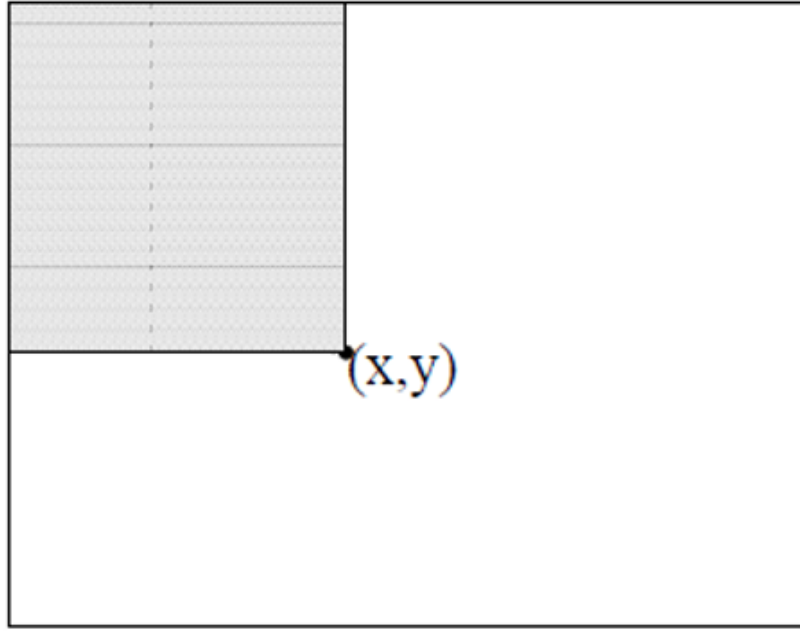
integral imgeyi, $i(x, y)$ orjinal imgeyi temsil etmektedir. ($s(x, y)$ kümülatif satır toplamını temsil etmektedir. ($s(x, -1) = 0$ ve $ii(-1, y) = 0$)

$$ii(x, y) = \sum_{x'=0}^x \sum_{y'=0}^y i(x', y') \quad (3.1)$$

$$s(x, y) = s(x, y - 1) + i(x, y) \quad (3.2)$$

$$ii(x, y) = ii(x - 1, y) + s(x, y) \quad (3.3)$$

Şekil 3.3'te D bölgesini hesaplayabilmek için 4 adet referans noktanın bilinmesi gerekir. 1. noktadaki değer A, 2. noktadaki değere A + B, 3. noktadaki değer A + C ve 4. noktadaki değer A + B + C + D olarak hesaplanır. D bölgesindeki toplam ise $4 + 1 - (2 + 3)$ olarak hesaplanmaktadır. Böylece integral imge sayesinde, her seferinde verilen noktaya göre bütün pikselleri toplamak yerine yukarıda verilen yöntem ile integral imge oluşturulduktan sonra hesaplama oldukça verimli bir şekilde gerçekleştirilir.

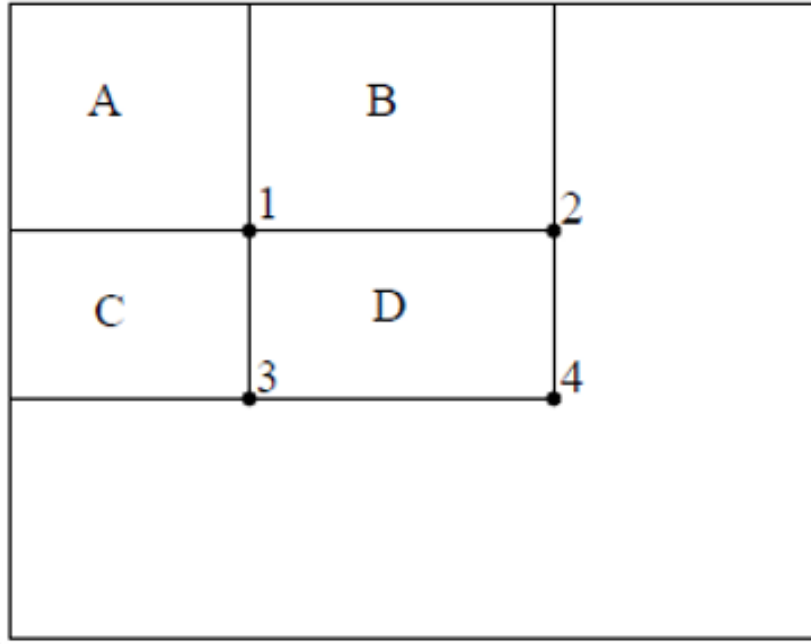


Şekil 3.2: (x,y) noktasındaki integral imgenin değeri, yukarıda ve solda kalan bütün piksel değerlerinin toplamına eşittir.

3.1.2 Sınıflandırma Fonksiyonlarını Öğrenme

Sistemde öznitelikleri seçmek ve sınıflandırıcıyı eğtmek için Adaboost kullanılmıştır. Adaboost öğrenme algoritması, basit öğrenme algoritmalarının performansını arttırmakta kullanılır. İşlem sürecinde zayıf sınıflandırıcıları birleştirip daha gürbüz bir sınıflandırıcı oluşturarak yapmaktadır. Adaboost öğrenme algoritması kısaca aşağıda gösterilmiştir.

- Verilen örnek imgeler $(x_1, y_1), \dots, (x_n, y_n)$ olsun. Burada $y_i = 0, 1$ negatif ve pozitif örnekler için değerlerdir.
- Ağırlıklar sırasıyla $y_i = 0, 1$ için $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ olarak başlatılsın. Burada m ve l pozitif ve negatiflerin sayısıdır.
- $t = 1, \dots, T$ için:
 1. Her bir w_t olasılık dağılımı için ağırlıkları normalize etme $w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$
 2. Her bir öznitelik j için, sadece bir öznitelik kullanma sınırı olan sınıflandırıcı h_j 'yi eğit. Hata değerlendirmesi $w_t, err_j = \sum_i w_i |h_j(x_i) - y_i|$
 3. En düşük hata değeri err_t 'ye sahip olan sınıflandırıcıyı seç.



Şekil 3.3: İntegral imgede bir bölgedeki değeri hesaplamak için gereken referans noktaları gösterilmiştir.

4. Ağırlıkları güncelle: $w_{t+1,i} = w_{t,i}\beta_t^{1-e_i}$. Burada eğer örnek doğru sınıflandırıldıysa $e_i = 0$, değilse $e_i = 1$ ve $\beta_t = \frac{err_t}{1-err_t}$

- Son güçlü sınıflandırıcı:

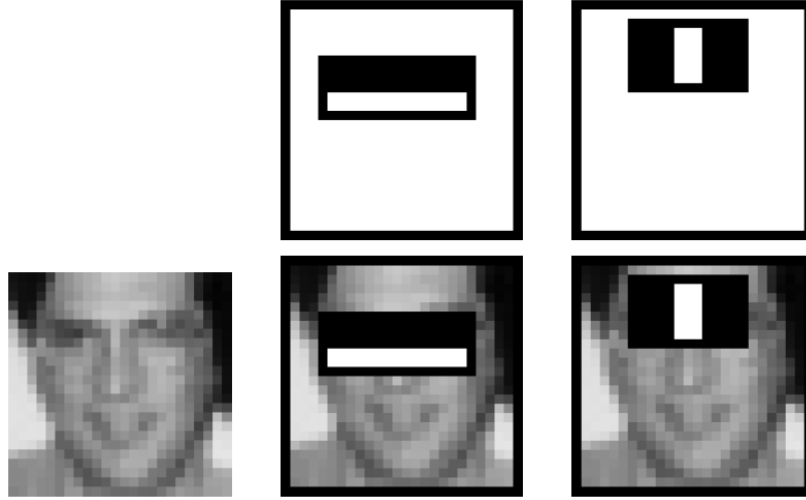
$a_t = \log \frac{1}{\beta_t}$ olduğunda

$$\begin{cases} 1 & \sum_{t=1}^T a_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T a_t \\ 0 & \text{tersi} \end{cases} \text{şeklinde gösterilir.}$$

Algoritma sonucunda seçilen birinci ve ikinci öznitelikler Şekil 3.4'te gösterilmiştir.

3.1.3 Kaskad Yöntemi

Kaskad yönteminde evreler Adaboost ile eğitilerek oluşturulmuştur. Öncelikle iki-öznitelik güçlü sınıflandırıcısı ile başlayıp, yanlış negatif sayısını minimize edecek eşik değeri değiştirilerek iyi bir yüz filtresi elde edilebilir. İlk eşik değeri (3.4)'de gösterilmiştir. Bu formül eğitim setinde düşük bir hata oranı sağlaması için tasarlanmıştır. Düşük eşik değeri yüksek tespit oranı ve yüksek yanlış negatif oranı beraberinde getirmiştir. Sonuç olarak iki-öznitelik sınıflandırıcısı kabul edilebilecek performanstan oldukça uzak olduğundan alt pencerelerde zayıf sınıflandırıcılardan oluşan daha güçlü bir sınıflandırıcı



Şekil 3.4: Adaboost tarafından birinci ve ikinci seçilmiş öznitelikler. Üst satırdaki bu iki özneliğin gösterilmiştir. Alt satırda bu öznelikler bir eğitim yüzüne bindirilmiştir. Birinci öznelik göz bölgesi ile üst yanak arasındaki yoğunluk farkını ölçer. Bu öznelikten göz bölgesinin yanak bölgesine göre daha karanlık olduğu gözlemlenebilir. İkinci öznelik ise iki göz ve arada kalan burun kaş bölgesi arasındaki geçişteki yoğunluk farklarını karşılaştırır.

oluşturabilmek için "Karar Ağacı" da denilen kaskad yöntemi önerilmiştir (Viola ve Jones, 2001). Bu yöntemde, ilk sınıflandırıcıdan gelen pozitif sonuç ikinci sınıflandırıcıyı tetikler. İkinci sınıflandırıcıdan gelen pozitif sonuç bir sonrakini tetikler. Daha sonraki adımlar aynı şekilde devam eder. Herhangi bir yerde gelen negatif bir sonuç, alt pencerenin tamamen yanlış olarak geri bildirim yapması ile sonuçlanır.

$$\frac{1}{2} \sum_{t=1}^T \alpha_t \quad (3.4)$$

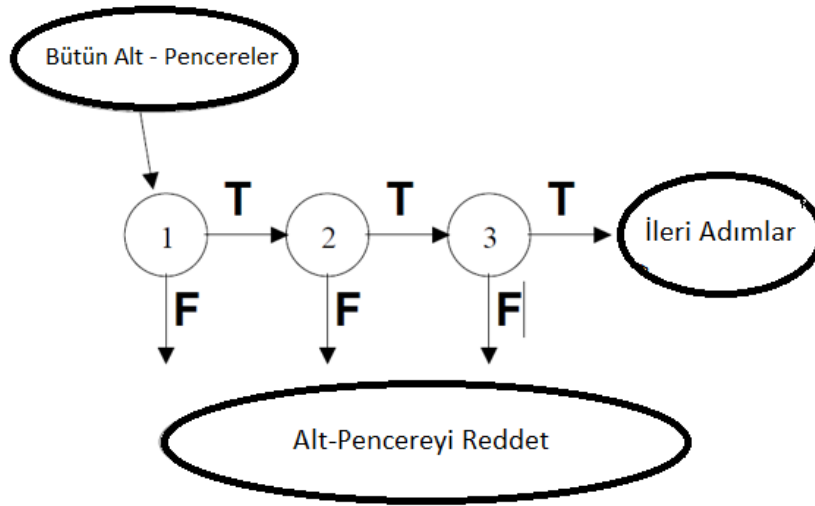
Kaskad Sınıflandırıcılarının Eğitilmesi: Eğitilmiş bir kaskad sınıflandırıcılarında, bir kaskadın yanlış pozitif oranı (3.5)'deki gibi gösterilmektedir. Burada F kaskad yapılmış sınıflandırıcının yanlış pozitif sayısını, K sınıflandırıcı sayısını ve f_i i'ninci sınıflandırıcının üzerinden geçtiği örneklerdeki yanlış pozitif oranını göstermektedir.

$$F = \prod_{i=1}^K f_i \quad (3.5)$$

Tespit oranı ise (3.6)'daki gibi gösterilmektedir. D kaskad yapılmış sınıflandırıcının tespit oranını, K sınıflandırıcı sayısını ve d_i i'ninci sınıflandırıcının üzerinden geçtiği örneklerdeki tespit oranını temsil etmektedir.

$$D = \prod_{i=1}^K d_i \quad (3.6)$$

Bu formüllerden çıkarılarak örneğin; eğer her evre 0.99 başarısı olan bir tespitten oluşuyor



Şekil 3.5: Şekil’de kaskad tespitinin şematığı gösterilmiştir. Her bir alt pencerede sınıflandırıcı serisi uygulanmıştır. İlk sınıflandırıcı oldukça fazla yanlış negatifler alır ve çok az bir hesaplama süresi alır. Daha sonraki tabakalarda ise bunlara ek olarak daha fazla yanlış negatifler, ama bu da demektir ki ek olarak bir hesaplama süresi gerektirir. Sonraki ilerlemede bu şekilde evreler oluşur. Kaskad yönteminde tespit evreleri bu sistemde ve diğer sistemlerde şekildeki gibi gerçekleştirilir.

ise, 0.9 tespit başarısına 10 evrede ulaşılır ($0.9 \approx 0.99^{10}$). Aynı şekilde her bir evre yanlış pozitif oranı 0.3 ise yaklaşık 6×10^{-6} gibi düşük bir değer elde edilir.

Değerlendirilen öznitelik sayısı, gerçek imgeler üzerinde tarama yapılarak olasılıksal bir süreç gerektirir. Herhangi bir alt-pencere tek bir seferde kaskad süreci boyunca sadece bir sınıflandırıcı ile devam edebilmektedir. Bu süreç pencere negatif ya da nadir şartlarda sağlanan her bir testin pozitif olarak etiketlenen süreç ile sonlanacaktır. Bu sürecin beklenen davranışı tipik bir testte imge pencerelerindeki dağılım ile karar verilmektedir. Her bir sınıflandırıcının anahtar ölçütü onun ”pozitif oran”ıdır. Beklenen öznitelik sayıları (3.7)’deki gibi gösterilmektedir. N değerlendirilen özniteliklerin sayısı, K sınıflandırıcı sayısı ve p_j j ’nci sınıflandırıcının pozitif oranı ve n_i i ’nci sınıflandırıcıdaki özniteliklerinin sayısıdır.

$$N = n_0 + \sum_{i=1}^K (n_i \prod_{j<i} p_j) \quad (3.7)$$

Daha önceki bölümde bahsedilen Adaboost öğrenme tekniği hata oranını minimize etmek için sunulmuştur. Adaboost yüksek yanlış pozitif oranlarda daha yüksek tespit oranı için özelleştirilmemiştir. Adaboost’ta eşik değerinin değiştirilmesi ile sınıflandırıcılarda daha az yanlış pozitif alması sağlanırken aynı zamanda daha düşük tespit oranı da sağlamaktadır. Bu durumun tam tersi de geçerlidir.

3.2 Nirengi Noktaları Bulabilmek için Hizalama Yöntemi

İki yöntemin diğer bir adımı olan ve imgelerde tespit edilen yüzlerde nirengi noktalarını hizalayabilen Kazemi ve Sullivan, 2014 çalışmasından faydalanılmıştır.

Yöntemde yüzdeki nirengi noktalarının yerleri verimli bir şekilde kesin olarak tahmin edilen algoritma kullanılmıştır. Kaskad regresörlerden faydalanılmıştır. Yüz hizalama yöntemi kaskad regresörleri, tekli bileşkelerin oluşturulma aşaması ve kaskadların eğitim aşamalarından sonraki bölümlerde ayrıntılarıyla bahsedilecektir.

3.2.1 Regresörlerin Kaskadları

$x_i \in R^2$, I imgesinde i 'ninci nirengi noktasının x, y koordinatları olsun. Vektör $S = (x_1^T, x_2^T, \dots, x_p^T) \in R^{2p}$, I 'daki bütün p nirengi noktalarını tanımlar. Bu yöntemde vektör S şekil olarak atıflanacaktır. $\hat{S}^{(t)}$, S 'in güncel tahmini değerini göstermekte kullanılacaktır. Kaskaddaki her bir regresör, $r_t(\cdot, \cdot)$, imgeden gelen güncelleme vektörünü ve güncel şekile eklenen $\hat{S}^{(t)}$ 'yi tahmin eder. $\hat{S}^{(t)}$ tahmini gelişimi (3.8)'te gösterilmiştir.

$$\hat{S}^{(t+1)} = \hat{S}^{(t)} + r_t(I, \hat{S}^{(t)}) \quad (3.8)$$

Kaskadın kritik noktası, I ve $\hat{S}^{(t)}$ (indisilenmiş güncel şekil tahmini)'den hesaplanmış piksel yoğunluğu değerleri gibi özniteliklerini temel alarak regresör r_t tahminlerini yapmasıdır. Bu da sürece geometrik invaryans sağlamaktadır.

Genişletilen çıktılarının uzaklığı, eğer ilk tahmin olan $\hat{S}^{(0)}$ eğitim verisinin lineer altuzayına ait ise, bu altuzayda uzandığı kesin olarak bilinir. Bu yüzden tahminlerde fazladan bir kısıtlamaya ihtiyaç duyulmamaktadır. Bu da oldukça yöntemi basitleştirir. İlk şekil basitçe eğitim verisinin ortalaması olarak seçilebilir.

r_t 'yi eğitmek için hata kayıplarının karesi ile gradyan ağaç arttırma kullanılmıştır (Hastie vd., 2001).

3.2.2 Kaskad İçindeki Herbir Regresörü Öğrenme

I_i yüz imgesi ve S_i onun şekil vektörü olarak tanımlanan, $(I_1, S_1, \dots, I_n, S_n)$ şeklinde bir eğitim seti olduğu varsayalım. Kaskad'taki ilk regresyon fonksiyonu olan r_0 'ı öğrenmek için, veri setinden yüz imgesinin üçlülere oluşturulur. Bir başlangıç şekli ve hedef güncelleme adımı $(I_{\pi_i}, \hat{S}_i^{(0)}, \Delta \hat{S}_i^{(0)})$ (3.9, 3.10 ve 3.11)'de tanımlanmıştır.

$$\pi_i \in (1, \dots, n) \quad (3.9)$$

$$\hat{S}_i^{(0)} \in (S_1, \dots, S_n)/S_{\pi_i} \quad (3.10)$$

$$\Delta \hat{S}_i^{(0)} = S_{\pi_i} - \hat{S}_i^{(0)} \quad (3.11)$$

$i = 1, \dots, N$ Bu üçlülerin toplam sayısı $N = nR$ olarak belirtilir. R imge I_i başına kullanılan başlatma sayısını temsil eder. Regresyon fonksiyonu r_0 bu veriden hata kayıp karesi ile gradyan ağaç arttırması algoritması ile öğrenilir. Öğrenme algoritması aşağıda gösterilmiştir. Eğitim üçlülerinin seti, kaskadtaki bir sonraki regresör olan r_1 için eğitim verisi $(I_{\pi_i}, \hat{S}_i^{(1)}, \Delta \hat{S}_i^{(1)})$ sağlamak için güncellenir. Bu işlem $r_0, r_1, \dots, r_{T-1}$ olan T regresörlerinin kaskadları yeterli bir doğruluk seviyesine ulaşana kadar tekrarlanır.

3.2.3 Ağaç Tabanlı Regresör

Herbir regresyon fonksiyonu r_t 'nin çekirdeği ağaç tabanlı regresörlerdir. Sonraki bölümlerde herbir regresyon ağacı için en önemli uygulama ayrıntıları ile incelenecektir.

3.2.3.1 Şekil İnvaryant Bölünmüş Testleri

Regresyon ağacındaki birebir ayırık düğümde, iki piksel yoğunluğu arasındaki eşik değerlemeye göre karar verilmiştir. Ortalama şeklin koordinat sisteminde testlerde kullanılan pikseller u ve v olarak tanımlanan şekiller ile ilintili olan aynı pozisyondaki noktalar indekslenmek istenmiştir. Bu işlem, öznitelikler çıkarılmadan önce o anda olan imgenin ortalama şekli tahmin edecek şekilde çarpıtma ile gerçekleştirilebilir. Çünkü imgede seyrek temsil kullanıldığı için, noktaların pozisyonlarını çarpıtmak, bütün imgeyi çarpıtmaya göre kolay ve daha verimlidir.

u 'ya en yakın olan ortalama şekildeki yüzsel nirengi noktanın indeksi k_u olsun. k_u 'nın u 'dan ofseti (3.12)'deki gibi tanımlanır.

$$\delta x_u = u - \bar{x}_{k_u} \quad (3.12)$$

Daha sonra I_i imgesinde tanımlanan şekil S_i için, I_i imgesindeki pozisyonu, yani kalitatif olarak ortalama imgede verilen u 'ya benzer (3.13)'de verilmiştir.

$$u' = x_{i,k_u} + \frac{1}{s_i} R_i^T \delta x_u \quad (3.13)$$

Burada s_i ve R_i , S_i 'yi ortalama şekil $\hat{S}$ 'e dönüştüren benzerlik dönüşümünün ölçeği ve rotasyon matrisleridir.

$$\sum_{j=1}^p ||\bar{x}_j - (s_i R_i x_{i,j} + t_i)||^2 \quad (3.14)$$

(3.14)'de ölçek ve rotasyon çarpıtılmış şekiller ve ortalama şeklin nirengi noktaları arasında karelerin toplamını minimize etmek için bulunur. v' 'de aynı şekilde tanımlanır. Her bir ayrığın şeklen 3 parametre içerir. Bunlar $\theta = (\rho, u, v)$ olarak tanımlanır. Bu parametrelerin herbir eğitim ve test örneğine uygulanması (3.15)'de gösterilmiştir.

$$h(I_{\pi_i}, \hat{S}_i^{(t)}, \theta) = \begin{cases} 1 & I_{\pi_i}(u') - I_{\pi_i}(v') > \rho \\ 0 & \text{tersi} \end{cases} \quad (3.15)$$

Burada u' ve v' , (3.13)'e göre $\hat{S}_i^{(t)}$ 'yi $\bar{S}$ 'e en iyi şekilde çarpıtan ölçek ve rotasyon matrisleri kullanılarak tanımlanmıştır.

Pratikte, atamalar ve yerel dönüşümler eğitim sırasında tanımlanır. Benzerlik dönüşümünü hesaplamak, kaskadın her evresinde bir kez yapıldığı için test sırasında işlemin en çok zaman alan kısmıdır.

3.2.3.2 Düğüm Ayrımını Seçme

Her bir regresyon ağacı için, parçalı olan sabit fonksiyon ile esas fonksiyona yakınsanmıştır. Burada parçalı sabit fonksiyon her bir yaprak düğüme uygun olan sabit vektördür. Regresyon ağacını eğitmek için her bir düğümde rastgele aday ayrımları seti (θ) oluşturulmuştur. Daha sonra θ^* aç gözlü bir şekilde bu adaylardan kare hatalar toplamını minimize edecek şekilde seçilmiştir.

Eğer bir düğümdeki eğitim örneklerinin indislerinin seti Θ ise, buna karşılık gelen minimize etme (3.16)'da gösterilmiştir.

$$E(\Theta, \theta) = \sum_{s \in (l, r)} \sum_{i \in \Theta_{\theta, s}} \|r_i - \mu_{\theta, s}\|^2 \quad (3.16)$$

Burada Θ_{θ} , sol düğüme gönderilen örneklerin indisleridir. r_i , gradyan artırma algoritmasındaki imge i için hesaplanmış kalanların vektörüdür. $s \in (l, r)$ için;

$$\mu_{\theta, s} = \frac{1}{|\Theta_{\theta, s}|} \sum_{i \in \Theta_{\theta, s}} r_i \quad (3.17)$$

(3.16) tekrar düzenlenirse, en uygun ayrık çok verimli bir şekilde bulunabilir (3.17). θ 'ya bağlı olmaya faktörler ihmal edilerek (3.18) elde edilir.

$$\operatorname{argmin}_{\theta} E(\Theta, \theta) = \operatorname{argmax}_{\theta} \sum_{s \in (l, r)} |\Theta_{\theta, s}| \mu_{\theta, s}^T \mu_{\theta, s} \quad (3.18)$$

Burada farklı θ 'ları değerlendirirken hesaplanması gereken tek üye $\mu_{\theta, l}$ 'dir. $\mu_{\theta, r}$ ana düğüm μ ve $\mu_{\theta, l}$ 'deki hedeflerin ortalamalarından hesaplanabilir. (3.19)'da gösterilmiştir.

$$\mu_{\theta, r} = \frac{|\Theta| \mu - |\Theta_{\theta, l}| \mu_{\theta, l}}{|\Theta_{\theta, r}|} \quad (3.19)$$

3.2.3.3 Öznitelik Seçimi

Her bir düğümdeki karar, piksel çiftlerindeki yoğunluk değerlerinin farkları arasındaki eşiklemeye bağlıdır. Bu oldukça basit testtir. Fakat bu tek bir değer eşiklemesine göre ışıktandırmadaki global değişimlere bağlı yoğunluktan dolayı daha güçlüdür. Ne yazık ki, piksel farklarını kullanmanın zorluğu, adaylar arasındaki potansiyel ayırım (öznitelik) sayılarıdır. Bu durum iyi bir θ bulmayı, yüksek sayılarda olan bu adaylardan arama yapmadan zorlaştırır. Ama bu sınırlama faktörü imge verisinin yapısını hesaba katarak kolaylaştırılabilir. Bunun için pikseller arasındaki uzaklığında eksponansiyel önceliği olan bir ayırmada daha yakın piksel çiftininin seçilmesi sağlanmıştır(3.20).

$$P(u, v) \propto e^{-\lambda \|u-v\|} \quad (3.20)$$

3.2.3.4 Kayıp Etiketlerin Üstesinden Gelme

(3.16)'da eğitim imgelerinde etiketlenmeyen bazı nirengi noktalarının kolaylıkla üstesinden gelinebilir. Her bir eğitim imgesi için $W_{i,j} \in [0, 1]$ değişkenleri tanımlanır, i her bir eğitim imgesini, j ise her bir nirengi noktasını temsil eder. Etiketlenmeyen i 'ninci imgedeki j 'ninci nirengi noktası $W_{i,j}$ 0 olarak tanımlanır. 1 ise etiketlendiğini belirtir. Böylece (3.16), (3.22)'deki gibi güncellenir. W_i vektör $(w_{i1}, w_{i1}, w_{i2}, w_{i2}, \dots, w_{ip}, w_{ip})^T$ kendisi üzerindeki diyagonalın, diyagonal matrisidir.

$$E(\Theta, \theta) = \sum_{s \in (l, r)} \sum_{i \in \Theta_{\theta, s}} (r_i - \mu_{\theta, s})^T W_i (r_i - \mu_{\theta, s}) \quad (3.21)$$

$s \in (r, l)$ için

$$\mu_{\theta, s} = \left(\sum_{i \in \Theta_{\theta, s}} W_i \right)^{-1} \sum_{i \in \Theta_{\theta, s}} W_i r_i \quad (3.22)$$

Gradyan artırma algoritması da bu ağırlık faktörleri hesaba katılarak modifiye edilmelidir. Bu durumda hedeflerin ağırlıklı ortalaması ile bütün modeli başlatma aracılığı ile kolay bir şekilde yapılabilir. Uyarlama regresyon ağaçlarındaki ağırlıklı kalanlar (3.23) gibi yapılır.

$$r_{ik} = W_i (\Delta S_i^{(t)} - f_{k-1}(I_{\pi_i}, \hat{S}_i^{(t)})) \quad (3.23)$$

Şekil 3.6'de yüz tespiti ve nirengi noktalarını yüz imgesinde hizalama yöntemleri uygulandıktan sonra gösterilmiştir. Tespit edilen yüz diktörtgen içine alınmıştır. Hizalanan noktalar yüz üzerinde noktalar ile gösterilmiştir.

3.3 Çok Sınıflı Destek Vektör Makinesi

DVM takip eden marjın kriterini (3.24) maksimize ederek iki-sınıflı birbirinden ayıran en uygun hiperdüzlemleri bulmaya çalışır. Verilen iki sınıfa ait eğitim vektörleri $x_i \in R^d$, $I =$



Şekil 3.6: Dikdörtgen ile yüz tespiti algoritması ile bulunan yüzler, noktalar ile tespit edilen yüzlerdeki nirengi noktaları ifade edilmiştir. (a) ve (b) imgelerinde nötr yüz ifadeleri, (c) imgesinde kızgınlık ve (d) imgesinde üzüntü ifadesi temsil edilmiştir.

$1, \dots, l$ ve etiket vektörü $y \in [1, -1]^l$, destek vektör tekniği (3.24)'teki problemi optimize etmeyi gerektirir.

$$\min_{w \in H, b \in R, \varepsilon_i \in R}$$

$$\frac{1}{2}w^T w + C \sum_{i=1}^l \varepsilon_i \quad (3.24)$$

$$y_i(w^T \rho(x_i) + b) \geq 1 - \varepsilon_i$$

$$\varepsilon_i \geq 0, i = 1, \dots, l$$

Burada $w \in R^d$ ağırlık vektörü, $C \in R_+$ düzenleme sabiti ve haritalama fonksiyonu ρ , eğitim verisinin uygun öznitelik uzayı H üzerine izdüşümüdür. Bu yaygın yöntemi çözmek için dual'i aracılığıyla (3.25)'teki yöntem kullanılır.

$$\min_{a \in R^l}$$

$$\frac{1}{2}a^T (K \circ (yy^T))a - e^T a \quad (3.25)$$

$$0 \leq a \leq Ce, y^T a = 0$$

Burada $e \in R^I$ bütün 1'lerin vektörü, $K \in R^{I \times I}$ kernel matrisi ve $\circ$ Hadamard-Schur çarpımıdır. K ve y sembolleri matrisleri ya da vektörleri temsil eder. 0 sembolü 0'a ayarlanmış bütün bileşke vektörlerini temsil eder.

(3.24) yerine yeni bir DVM formülasyonu önerilmiştir. (3.26)'de formülasyon gösterilmiştir. Bunun çifti sınırlı-zorlamalı bir probleme dönüşmüş hali (3.3)'de gösterilmiştir.

$$\min_{w \in H, b \in R, \varepsilon_i \in R}$$

$$\frac{1}{2}(w^T w + b^2) + C \sum_{i=1}^l \varepsilon_i \quad (3.26)$$

$$y_i(w^T \rho(x_i) + b) \geq 1 - \varepsilon_i$$

$$\varepsilon_i \geq 0, i = 1, \dots, l$$

$$\min_{a \in R}$$

$$\frac{1}{2}a^T((K + ee^T) \circ (yy^T))a - e^T a \quad (3.27)$$

$$0 \leq a \leq Ce$$

Buradaki motivasyon, dual problemini lineer zorlama olmadan daha kolay üstesinden gelmektir.

1vs1 Yaklaşımı: Yöntemde kullanılan çok sınıflı DVM yöntemi, 1vs1 yada ikili dağılma yaklaşımıdır. Bu yöntem bütün ihtimal olan ikili sınıflandırıcıları değerlendirir. Toplamda $k(k-1)/2$ bireysel ikili sınıflandırıcı oluşturur. Her bir test örneğini her bir sınıflandırıcıya uygulayınca, tespit edilen sınıfa +1 oy verilir. En çok oy olan sınıf test örneğinin o sınıfa etiketlenmesi ile sonuçlanır. 1vs1 yönteminde oluşturulan sınıflandırıcı sayısı 1vsHepsi yöntemine göre daha fazla sayıda sınıflandırıcı oluşturur. Fakat, her bir sınıflandırıcıda karesel programlamanın (QP) boyutu daha küçüktür. Bu durum eğitimin daha hızlı yapılmasına olanak sağlar. Dahası 1vs1 yaklaşımı 1vsHepsi yaklaşımına göre daha fazla simetriktr.

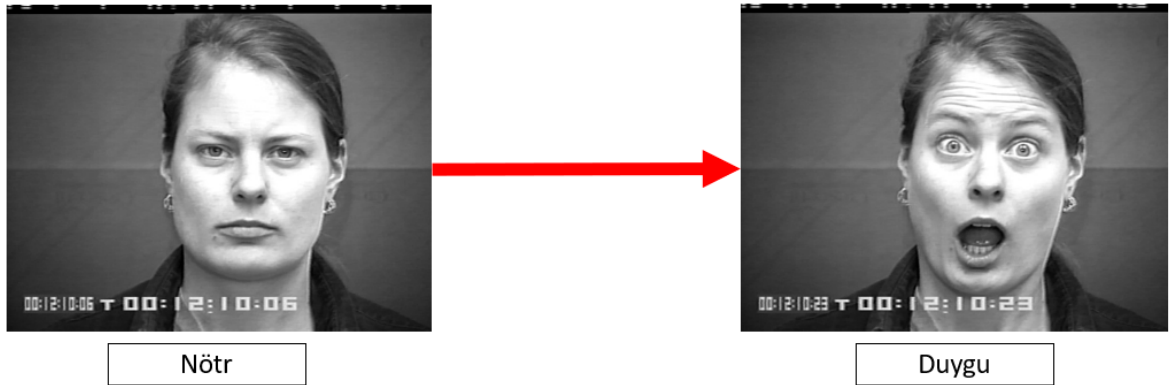
3.4 Duygu Tanıma İçin Geometrik Öznitelikler Üzerinde Ardışık İleri Seçim Algoritması

Gerçekleştirilen çalışmanın ilk adımında imgelerdeki yüzlerin bulunması için Viola ve Jones'un yüz bulma algoritması kullanılmıştır. Daha sonra Kazemi ve Sullivan'ın

yöntemiyle yüzlerdeki 68 adet nirengi noktası tespit edilmiştir. 68 farklı nirengi noktasının ikili kombinasyonlarından 2278 farklı nokta çifti elde edilerek bu noktaların birbirlerine olan uzaklıklarının yatay ve dikey eksenlerdeki bileşenleri hesaplanmıştır. Nötr ve belirgin bir ifadeye sahip yüzler için hesaplanan iki uzaklık vektörü arasındaki fark öznitelik vektörünü oluşturur. Hesaplanan öznitelik vektöründe ardışık ileri seçim algoritması (AIS) uygulanarak sadece faydalı özniteliklerden oluşan bir alt küme seçilmiştir. Çok sınıflı destek vektör makinesi sadece seçilen öznitelikler ile eğitilerek sınıflandırma yapılmıştır.

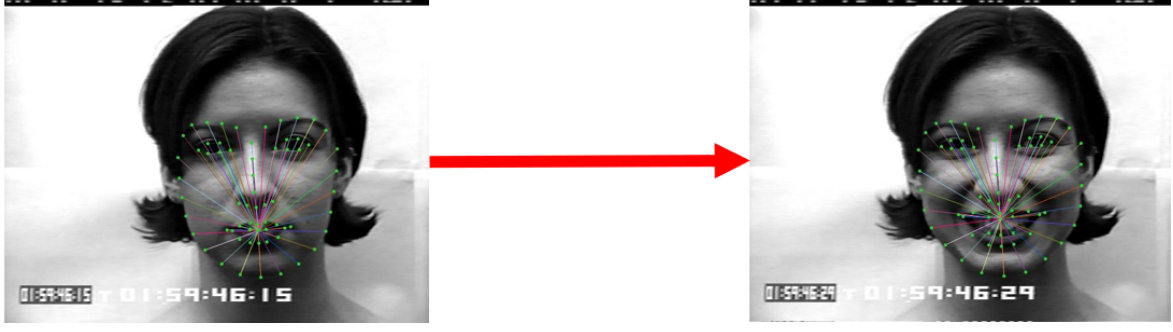
3.4.1 Geometrik Öznitelik Vektörünün Oluşturulması

Öznitelik vektörü oluşturulurken aynı kişiye ait nötr ve duygu içeren iki yüz ifadesinden nirengi noktaları tespit edilmekte (bkz. Şekil 3.7), sonrasında tespit edilen nirengi noktalarının tüm ikili kombinasyonlarının yatay ve dikey eksenlerdeki değişimleri elde edilmektedir (bkz. Şekil 3.8). Nirengi noktalarının tekil hareketleri yerine nirengi noktalarının kombinasyonları arasındaki hareketin öznitelik olarak kullanılması, kişinin yüz yapısına göre bir normalizasyon sağlamaktadır. Bu yaklaşım, kişinin nötr yüz ifadesini kullanarak yapılan bir ölçümleme olarak da görülebilir. (3.28)'de nirengi noktalarının ikili



Şekil 3.7: Şekil’de nötr yüzden duygu ifade eden yüz arasındaki geçiş gösterilmiştir.

kombinasyonları arasındaki uzaklıkları içeren vektörün hesaplanması gösterilmiştir. Burada l nirengi noktalarının konumlarını içeren vektör, d ise uzaklık vektörüdür. Elimizde 68 adet nirengi noktası olduğu için i ve j 1’den 68’e kadar değerler alacaktır. Bir nirengi noktasının kendisine uzaklığını öznitelik olarak kullanmak anlamsız olacağı için i ’nin j ’ye eşit olduğu durumlar uzaklık vektörüne eklenmemiştir. l_i ve l_j arasında hesaplanan yatay ve dikey uzaklıklar d vektörüne ayrı elemanlar olarak yerleştirilir. Yatay ve dikey uzaklıkların ayrıca ele alındığı göz önünde bulundurulursa 68 nirengi noktasından elde edilen 2278



Şekil 3.8: Şekilde sadece bir noktanın diğer noktalarla nötr ve duygu yüzünde oluşturduğu nokta çiftleri renkli çizgilerle gösterilmiştir. Bütün noktalar arasındaki uzaklıklar aynı şekilde bulunur.

kombinasyondan 4556 ayrı uzaklık elde edilir. Bir öznitelik kullanılacağı zaman o uzaklığa ait yatay ve dikey uzaklıklar ayrı kullanılır.

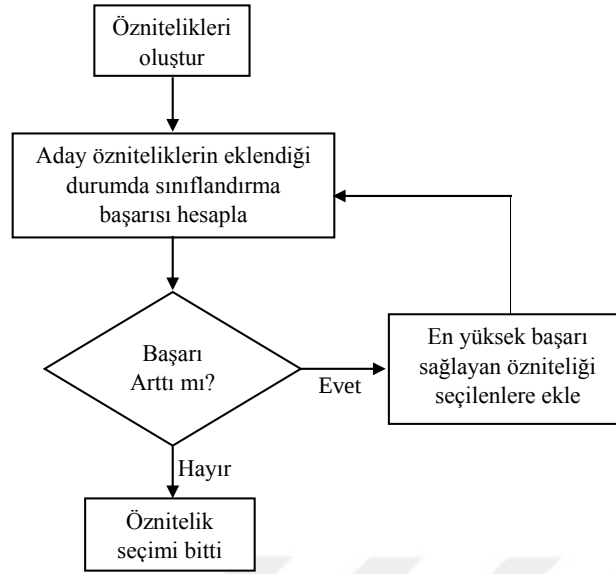
$$d_k = l_i - l_j \quad \begin{cases} 1 \leq i, j \leq 68 \\ i \neq j \end{cases} \quad i, j \in \mathbb{Z} \quad (3.28)$$

(3.28)'deki işlem nötr ve duygu ifadesine sahip yüz imgeleri için ayrı ayrı uygulanır. Öznitelik vektörü f , nötr yüzden elde edilen uzaklık vektörü d_n ve duygu ifadesine sahip yüzden elde edilen uzaklık d_d 'nin farkı ile elde edilir (bkz. 3.29).

$$f = d_n - d_d \quad (3.29)$$

3.4.2 Ardışık İleri Seçim Algoritması

(3.29)'deki öznitelik vektörü f , bütün nirengi noktalarının kombinasyonlarıyla oluşturulmuştur. Bu kombinasyonların en faydalı olanların seçilmesi için ardışık ileri seçim algoritması kullanılmıştır. Şekil 3.9'de Ardışık İleri Seçim (AİS) algoritmasının akış diyagramı verilmiştir. Bu işlemi imgelerin birini dışarıda bırakarak gerçekleştirmek çok uzun sürdüğü için CK+ veri seti rastgele şekilde eğitim ve test kısımlarına bölünmüştür. Seçilen öznitelikler ve eğitim imgeleri kullanılarak eğitilen sınıflandırıcının test imgelerindeki başarısı, seçilen özniteliklerin başarısını temsil etmektedir. Algoritmada her iterasyonun başında seçilen öznitelikler vektöründe k eleman vardır. Mevcut 4456 öznitelik arasından öznitelik vektöründe bulunmayan özniteliklerin uygunlukları sırayla test edilir. Bu amaçla her aday öznitelik için ayrı ayrı k elemanlı vektöre eklendiklerinde elde edilen sınıflandırma başarısı ölçülmelidir. Bu aşamada 4456 öznitelikten vektördeki $k + 1$ tanesi her imge için kendi içinde L2 normalizasyona tabi tutulur. Eğitim imgelerinin normalize edilen öznitelikleri hata düzelten çıktı kodları tabanlı bir çok-sınıflı DVM'ye beslenir. Eğitilen DVM'ye test imgelerinin normalize edilmiş öznitelikleri sınıflandırılarak eklenen



Şekil 3.9: Önerilen algoritmanın akış diyagramı.

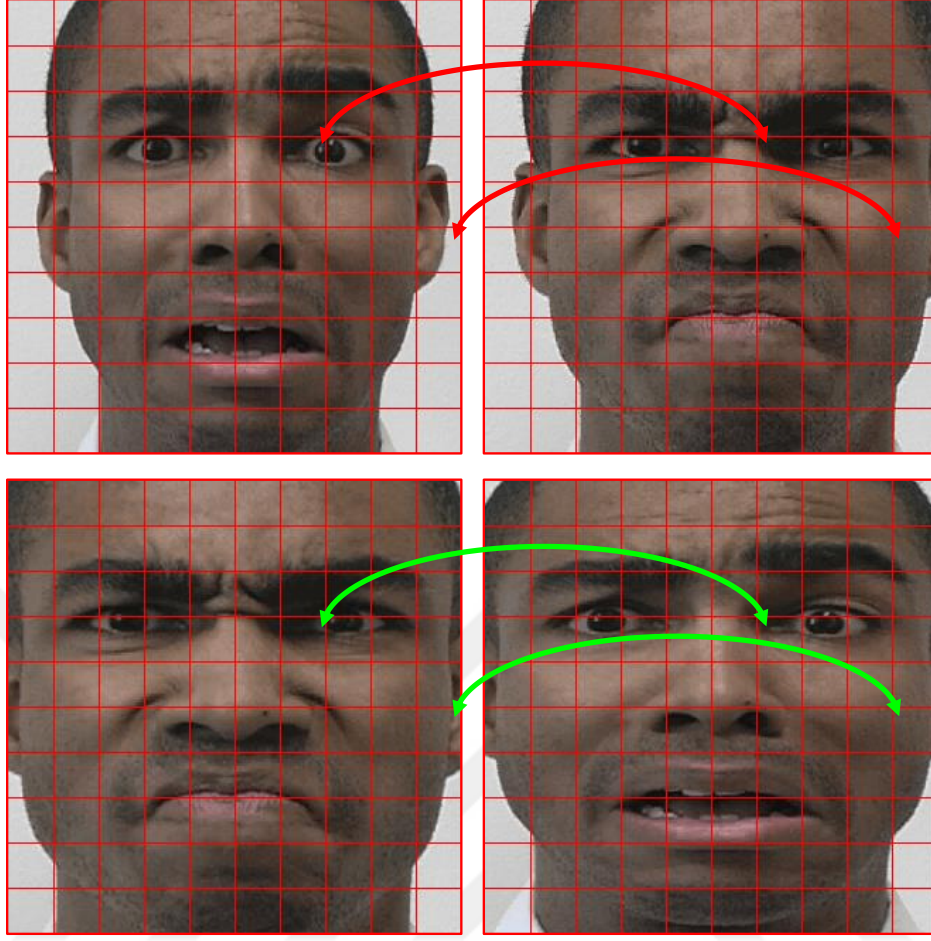
öznitelik için sınıflandırma başarısı hesaplanır. Öznitelikler arasından sınıflandırma başarısını en çok artıran seçilen öznitelikler vektörüne eklenir ve bir sonraki iterasyona geçilir. Bu işlem, yeni bir öznitelik eklemek sınıflandırma başarısını düşürmeye başlayana kadar tekrarlanır.

3.5 FHOG Öznitelikleri ile Duygu Tanıma

Önerilen yöntemin ilk adımında imgelerdeki yüzler Viola ve Jones'un algoritması ile tespit edilir. Sonra yüz nirengi noktalarının bulunması için Kazemi ve Sullivan'ın yüz hizalama algoritması kullanılır. İmgeler, tespit edilen yüzler üzerinde nirengi noktaları da kullanılarak ön işleminden geçirilir. Daha sonra FHOG (Felzenszwalb vd., 2009) öznitelikleri elde edilir. Elde edilen öznitelikler DVM'ye beslenerek imgeler sınıflandırılır. İzlenilen yöntemin adımları aşağıdaki bölümlerde detaylandırılmaktadır.

3.5.1 Yüz İmgeleri, Nirengi Noktaları Tespiti ve Ön İşlemler

Öncelikle imgeler gri seviyeye dönüştürülmüştür. İmgelerden yüzlerin tespiti için Viola ve Jones'un 2004'te önerdiği Haar kaskad yöntemi kullanılmıştır. Yüz bulma algoritmasında aynı yüze ait farklı imgelerden bulunan yüz sınır kutuları, pozisyon farklılarından dolayı yüzü aynı şekilde çevrelememektedir. Bu da direk yüzlerin bulunduğu kutulardan çıkarılacak özniteliklerde örtüşme sağlamadığı için performans düşüklüğüne sebep olur. Bu yüzden bir sonraki adımda, bulunan yüzde hizalanmış nirengi noktalarını bulabilmek için Kazemi ve Sullivan'ın 2014 yılında önerdiği yöntem kullanılmıştır.

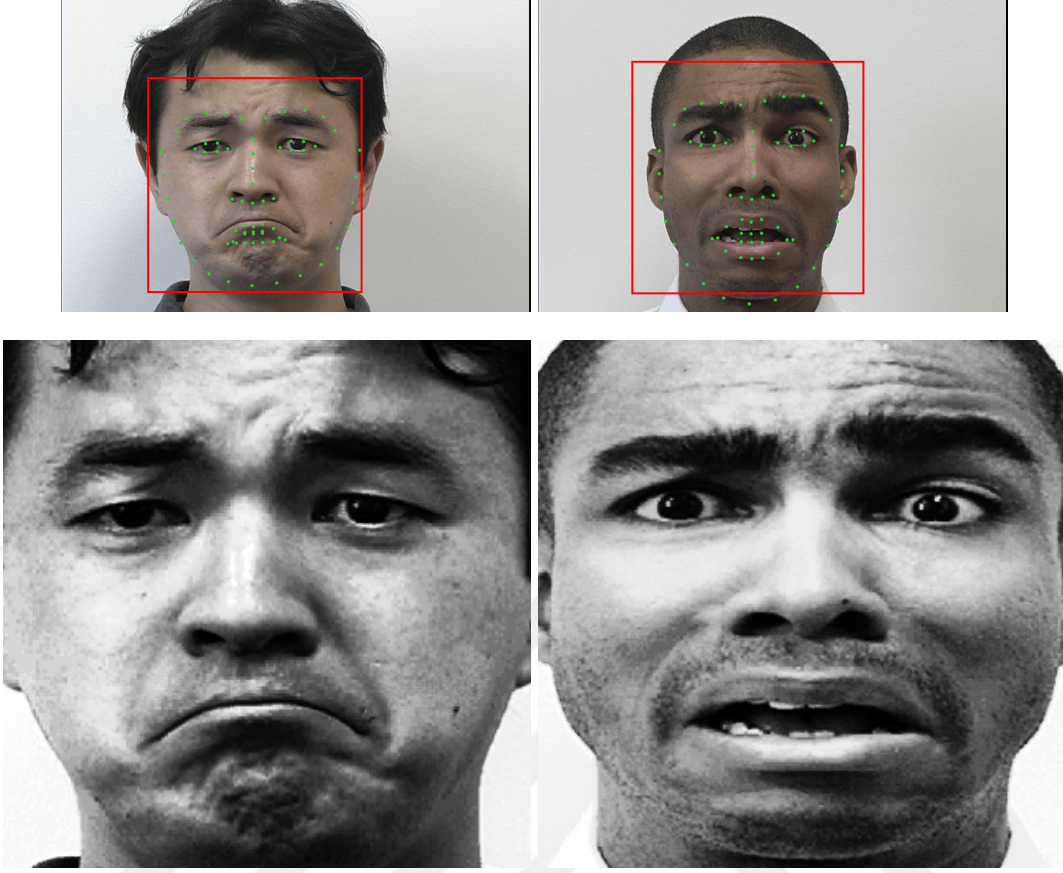


Şekil 3.10: Hizalanmadan kesilen yüz ile hizalanarak kesilen yüz imgeleri gösterilmiştir.

Bulunan yüzler, hizalanmış nirengi noktalarından yatay olarak en sağ ve en sol yanak noktaları, dikey olarak çenenin en alt noktası ve kaşın en üst noktası referans alınıp yüz hizalaması yapılmıştır. Alın bölgesinin de alınabilmesi için kaşın en üst noktasından alın bölgesine doğru %20 kaydırılma yapıp imgeler kırılmıştır ve 320×320 ebatına yeniden boyutlandırılmıştır. Şekil 3.10’de hizalanmamış yüzler ile hizalanmış yüzler arasındaki fark gösterilmiştir. Yeniden boyutlandırılan imgelerde ışık etkisini azaltmak ve imgelerdeki renk karışıklığını arttırmak için histogram eşitlemesi uygulanmıştır. Şekil 3.11’de CK+ veri kümesindeki (Lucey vd., 2010), duygu ifade eden imgelerden orjinal imge ve ön işlemlerden sonraki halleri gösterilmiştir.

3.5.2 FHOG Özniteliklerin Elde Edilmesi

HOG öznitelikleri, ilk kez Dalal ve Triggs, 2004 yaya tespiti yönteminde uygulanmıştır . Bunlar görünüm tabanlı öznitelikler olup, imgelerden alınan hücrelerde her piksel için gradyan hesabı yapılarak o hücreye ait yönelimli gradyan histogramı hesaplanarak elde edilir. Bu histogramda, imgenin o hücredeki gradyanların yönelimlerinin

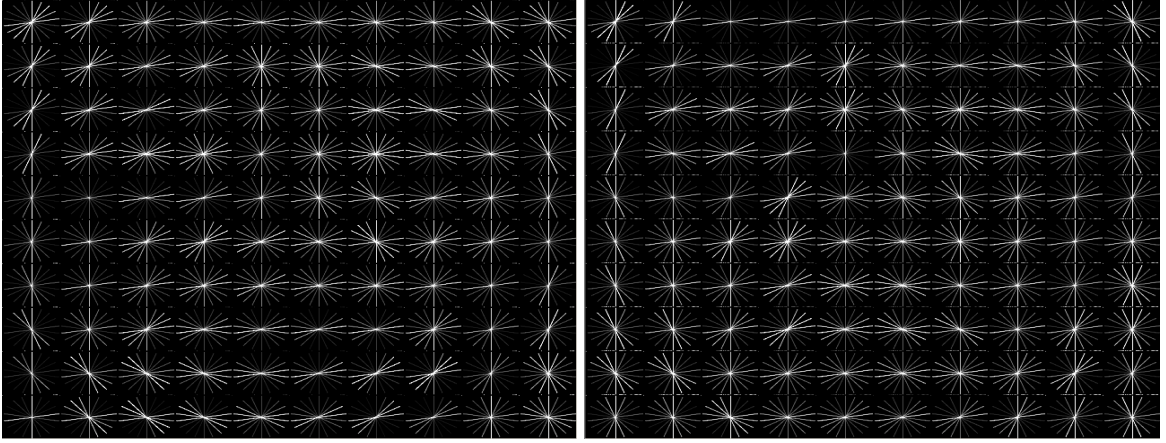


Şekil 3.11: Üst sıradaki imgelerde tespit edilen yüzler kırmızı dikdörtgenler ile gösterilmiştir. Yeşil noktalar nirengi noktalarını temsil etmektedir. Tespit edilen yüzlerden kırpma, yeniden boyutlandırma ve histogram eşitleme sonucu elde edilen imgeler alt sırada gösterilmiştir.

istatistikleri tutulur. HOG özneliliklerinin elde edilmesinde izlenen adımlar kısaca şu şekildedir:

- İmgeler hücrelere bölünür,
- Bu hücrelerde gradyanlar hesaplanır,
- Hesaplanan gradyanlar ile hücrelere özel histogramlar oluşturulur.
- Oluşturulan histogramlar bloklar halinde gruplanıp aralarında düzgelenir.

Gradyan hesaplamak için imgeye yatay ve dikey filtreler uygulanır. Bu uygulamada, Dalal ve Triggs'in çalışmasında en iyi başarıyı elde ettiği belirlenen $S_x = [-1, 0, 1]$ ve transpozu olan S_y filtreler kullanılmıştır. Bu filtrenin yatay ve dikey olarak imgeye uygulanmasından sonra I_x ve I_y olarak gösterilen yatay ve dikey değişimlerin olduğu imgeler bulunur. Daha



Şekil 3.12: Yüzden çıkarılan FHO'ların görselleştirilmiş şekilleri gösterilmiştir.

sonra bu imgeler kullanılarak bu gradyanların yönelim şiddetleri (G) ve açıları (θ) bulunur. (3.30, 3.31, 3.32 ve 3.33)'te bu değerlerin nasıl hesaplandığı gösterilmiştir.

$$I_x = I * S_x \quad (3.30)$$

$$I_y = I * S_y \quad (3.31)$$

$$G = \sqrt{I_x^2 + I_y^2} \quad (3.32)$$

$$\theta = \arctan \frac{I_x}{I_y} \quad (3.33)$$

Gradyanlar hesaplandıktan sonra yönelim gruplaması yapılır. Hücredeki her piksel ağırlıklı oylama ile yönelim-tabanlı histograma katkıda bulunur. Histogram kanalları 0-180 derece arasında eşit 9 parçaya bölünmüştür. Ağırlık oylaması için gradyanların yönelim şiddetleri ve açıları kullanılır. Her hücre, 9×1 'lik boyutu olan bir öznitelik vektörü oluşturur. Histogramlar hesaplandıktan sonra hücreleri aralarında düzgelemek için bloklar oluşturulur. 16×16 boyutlu hücreler 32×32 boyutlu blokları oluşturur. Her bir hücreye ait 9×1 boyutlu vektörler uç uca eklenerek 36×1 'lik bir vektör elde edilir. Bu blok içinde L2 düzgelemesi uygulanır. Böylece özniteliklerin ışıık koşulları ve diğer olumsuz şartlara karşı daha gürbüz olması sağlanır.

FHO (Felzenszwalb vd., 2009), HOG'tan (Dalal ve Triggs, 2004) daha küçük boyutta öznitelik vektörü üreten ve daha iyi başarımlar gösteren bir öznitelik çeşididir. HOG'daki 36×1 boyutundaki vektörün muadili FHO'da 31×1 boyutundadır ve daha çok bilgi içerir.

FHOG öznitelikleri elde edilirken, yönelim gradyan hesaplamaları ve histogram oluşturmaları aynı adımlarda gerçekleştirilir. Daha sonra 36×1 boyutundaki vektöre 4 hücreden gelen 9×1 vektörlerden oluştuğu için 4×9 boyutunda bir matris olarak bakılır. Daha sonra 13 boyutlu öznitelik vektörü hesaplanmaktadır. Bu 13 boyutlu vektör 36 boyutlu HOG vektörü ile aynı performansı sağlamaktadır. Sonra düşük boyutlu kontrasta duyarlı öznitelik bazı kategorilerde yarar sağladığı, aynı zamanda kontrasta duyarlı olmayan özniteliklerin de başka kategorilerde yarar sağladığı gözlemlenmiştir. Bu yüzden FHOG’da bu iki öznitelik vektöründen gelen bilgilerin hepsini içermiştir. En son olarak 31 boyutlu öznitelik vektörü elde edilirken, 27 boyuta karşılık gelen (9 kontrasta duyarlı ve 18 kontrasta duyarlı olmayan) farklı yönelim kanalları ve 4 hücreli kare bloklarda toplam gradyan enerjisini yakalayan 4 boyutlu öznitelik vektörü ile toplamda 31 boyutlu olan FHOG öznitelik vektörü oluşturulmuştur.

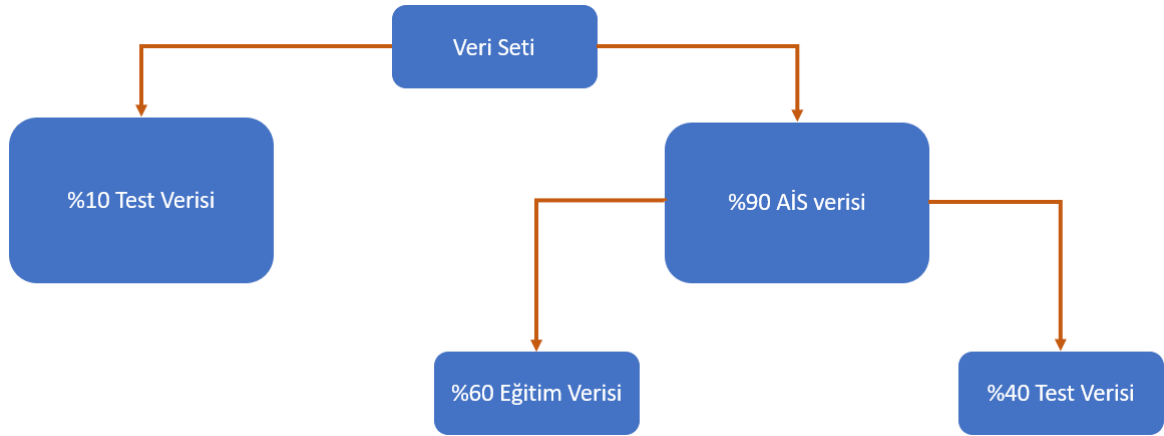
4. BULGULAR VE TARTIŞMA

4.1 Extended Cohn Kanade Veri Seti

Çalışmalarda deneyler Extended Cohn-Kanade (CK+) veri kümesinde gerçekleştirilmiştir (Lucey vd., 2010). CK+ veri seti; kızgınlık, küçümseme, iğrenme, korku, mutluluk, üzülmeye ve şaşkınlıktan oluşan toplam yedi farklı yüz ifadesi sınıfına ait örneklerden oluşur. 123 farklı öznenen normal yüzden başlayarak duygu ifade eden yüze doğru video kareleri olarak alınan toplamda 327 yüz ifadesi örneği bulunur. Çalışmalarda test ve eğitim için duygu ifade eden video karelerinin sonuncu kareleri alınmıştır. Çizelge 4.1’de sınıf dağılımları gösterilmiştir.

Kızgınlık	Küçümseme	İğrenme	Korku	Mutluluk	Üzüntü	Şaşırma	Toplam
45	19	59	25	69	28	82	327

Çizelge 4.1: CK+ veri kümesindeki Sınıflar ve sınıflara ait örnek sayıları



Şekil 4.1: Şekil’de AİS algoritması için veri setinin 10 kat çapraz doğrulamada bölünmesi gösterilmiştir.

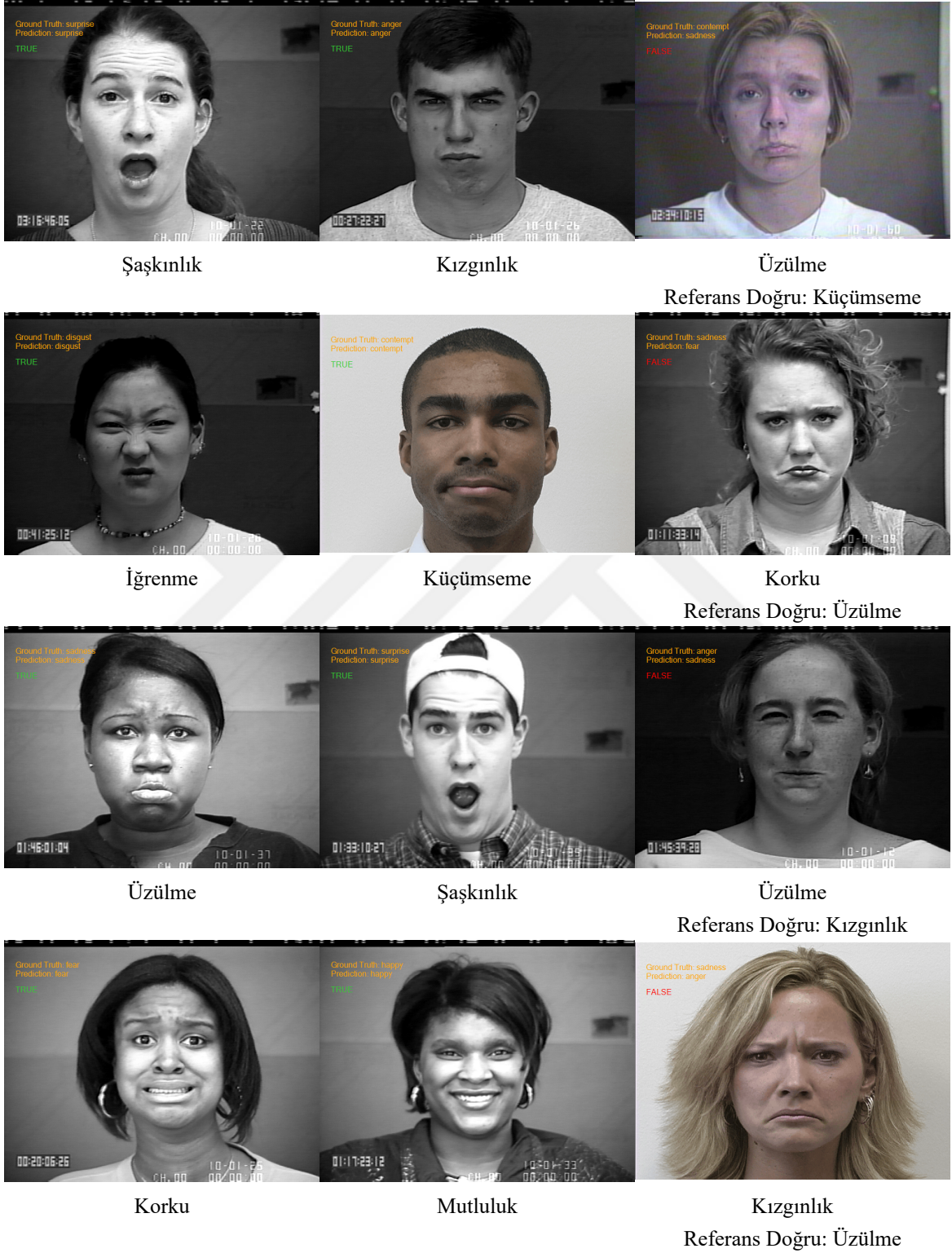


Şekil 4.2: Mavi barlar iki nirengi noktası arasındaki yatay uzaklıkları, kırmızı barlar ise dikey uzaklıkları temsil eder. Her bir bar ilk nirengi noktası ile başlar ve o noktanın sonu ikinci nokta ile kesik çizgiler ile bağlanır. Bu barlardaki uzunluklardaki değişimler sınıflandırma için kullanılan özniteliklerdir. Görüldüğü üzere, en az bir bar her bir duygu için ayırt edici bir şekilde kısalır ya da uzar.

4.2 Ardışık İleri Seçim Algoritması Bulguları ve Tartışma

Deneysel Sonuçlar: Deneyler gerçekleştirilirken bütün veri setinde bir performans ölçümü yapabilmek için 10 kat çapraz doğrulama yöntemi kullanıldı. Bu yöntemde veri seti 10'a bölünerek 1 parça test için ayrılmıştır ve diğer 9 parça AİS Algoritmasını uygulamak için kullanılmıştır. Bu adım bütün parçalar test edilene kadar tekrarlanarak yapılmıştır. AİS algoritmasını uygulamak için verisetinde test haricinde kalan 9 parça %60 ve %40 oranlarında bölünmüştür. %60 olan kısım eğitim %40 kısım AİS testi için kullanılmıştır(bknz. Şekil 4.1). Bu bölünme yapılırken sınıf frekansları korunmuştur. Seçilen öznitelikler Şekil 4.2'de gösterilmiştir. Öznitelik olarak kullanılan öznitelikler dikey ya da yatay barlar olarak çizilmiştir. Seçilen öznitelikler kalan 1 parça ile test edilmiştir.

Öznitelik vektörü L2 normalize edilmiştir. Çoklu sınıf DVM sınıflandırıcısı kullanılmıştır. Çizelge 4.2'de karışıklık matrisi gösterilmiştir. Başarım %88.7 olarak bulunmuştur.



Şekil 4.3: Şekilde Ardışık İleri Seçim Algoritması ile seçilen öznetelikler kullanılarak yapılan deneylerde ilk iki sütun doğru bulunanları temsil ederken, son sütun yanlış bulunan imgeleri göstermektedir.

Çizelge 4.2: Ardışık İleri Seçim Algoritması ile Seçilen Öznitelikler kullanılarak CK+ verisetinde duygu tanıma karışıklık matrisi.

	Kızgınlık	Küçümseme	İğrenme	Korku	Mutluluk	Üzülme	Şaşırma
Kızgınlık	0.78	0.04	0.09	0	0	0.09	0
Küçümseme	0.21	0.64	0	0.05	0.05	0.05	0
İğrenme	0.05	0	0.93	0.02	0	0	0
Korku	0	0	0	0.80	0.04	0	0.16
Mutluluk	0	0	0	0.01	0.99	0	0
Üzülme	0.11	0.11	0.03	0.03	0	0.64	0.08
Şaşırma	0	0	0	0	0	0	1

Bu barların boyları duyguya bağlı olarak ayırıcı bir şekilde değişmiştir. Bu öznitelikler otomatik olarak seçilmesine rağmen bireysel olarak incelendiğinde mantıklı sonuçlar çıkarılabilir. Buradan ağız ve kaş hareketleri duygu tanımda oldukça etkili olduğu anlaşılabılır. 3 öznitelik ağız hareketlerinde, 4 öznitelik ise kaş hareketlerinin sonucu oluşmuştur. En alttaki mavi bar ağızın genişlemesini tanımlar. En soldaki kırmızı bar ise kaşın dikey hareketinin tanımlayıcı bir öznitelik olduğunu açıklar.

Bir beklenmedik öznitelik olarak kulak ile burun arasındaki dikey uzaklık öznitelik olarak seçilmiştir. AİS öznitelikleri, doğruluk oranında oldukça düşük yükselme olmasına rağmen seçer. Bu seçilen öznitelik çıkarıldığında %3 lük bir düşüş yaşanmıştır. Bu öznitelik, kafa pozu değişimlerine duyarlılığında bir rol oynadığı varsayılmıştır. Çünkü bir kişinin kulak ile burun arasındaki mesafe değişmeyeceğinden, sadece imgelerde bulunan yüzlerdeki kafa pozu değişikliğinden etkileneceğinden böyle bir varsayım yapılmıştır.

4.3 FHOG Bulguları ve Tartışma

DeneySEL Sonuçlar: Deneyler gerçekleştirilirken bütün veri setinde bir performans ölçümü yapabilmek için 10 kat çapraz doğrulama yöntemi kullanıldı. Bu yöntemde veri seti 10'a bölünerek 1 parça test için ayrılmıştır ve diğer 9 parça DVM'de eğitim için kullanılmıştır. Bu adım bütün parçalar test edilene kadar tekrarlanarak yapılmıştır. Deneyler sonucu toplamda %93.27 başarımla elde edilmiştir. Çizelge 4.3'de her bir sınıfa ait başarımlar ayrı ayrı olarak gösterilmiştir. Şekil 4.4'te deneyler sonucu test edilen imge, imgeye ait referans doğrusu ve yöntemden tahmin edilen sonuçlar gösterilmiştir.



Referans Doğru: Kızgınlık
Tahmin Edilen: Kızgınlık

Referans Doğru: Şaşkınlık
Tahmin Edilen: Şaşkınlık

Referans Doğru: Üzülme
Tahmin edilen: Üzülme



Referans Doğru: İğrenme
Tahmin Edilen: İğrenme

Referans Doğru: Mutluluk
Tahmin edilen: Mutluluk

Referans Doğru: Küçümseme
Tahmin edilen: Küçümseme



Referans Doğru: Kızgınlık
Tahmin Edilen: Üzülme

Referans Doğru: Korku
Tahmin Edilen: Mutluluk

Referans Doğru: Korku
Tahmin Edilen: Şaşkınlık

Şekil 4.4: Şekilde FHOG öznetelikleri ile yapılan deneylerde ilk iki sütun doğru bulunanları temsil ederken, son sütun yanlış bulunan imgeleri göstermektedir.

Çizelge 4.3: FHOG CK+ verisetinde duygu tanıma karışıklık matrisi.

	Kızgınlık	Küçümseme	İğrenme	Korku	Mutluluk	Üzülme	Şaşırma
Kızgınlık	0.89	0.04	0.02	0	0	0.04	0
Küçümseme	0	0.84	0	0.06	0	0.1	0
İğrenme	0.05	0	0.93	0	0.02	0	0
Korku	0.04	0	0.04	0.76	0.12	0	0.04
Mutluluk	0	0	0	0	1	0	0
Üzülme	0.07	0.04	0	0	0	0.89	0
Şaşırma	0	0	0	0.01	0	0	0.99

Çizelge 4.4: Geometrik, HOG ve FHOG özniteliklerinin CK+ veri setinde karşılaştırılması.

	Geometrik	HOG Dalal ve Triggs, 2004	Lee vd., 2015	FHOG Felzenszwalb vd., 2009
Kızgınlık	0.78	0.86	0.84	0.89
Küçümseme	0.64	0.73	0.78	0.84
İğrenme	0.93	0.91	0.91	0.93
Korku	0.80	0.68	0.80	0.76
Mutluluk	0.99	0.97	0.99	1
Üzülme	0.64	0.61	0.68	0.89
Şaşkınlık	1	0.99	1	0.99
Toplam	0.88	0.88	0.90	0.93

Daha sonra FHOG'un duygu tanımda HOG'a ve geometrik özniteliklerine karşı başarımları aynı deneyler uygulanarak karşılaştırılmıştır. FHOG'un %5'e varan bir performans üstünlüğü olduğu görülmüştür. Çizelge 4.4'de FHOG, HOG ve geometrik öznitelikler arasındaki bütün duygu sınıflarındaki başarımları gösterilmiştir. Ayrıca güncel yöntemlerden olan Lee vd., 2015'nin yöntemi ile karşılaştırma yapıldığında %3'e varan bir performans üstünlüğü de gözlemlenmiştir.

FHOG özniteliklerinin HOG'tan %5 daha yüksek performans göstermesinin sebebi daha fazla bilgiyi daha sıkıştırılmış bir şekilde içerebilmesidir. HOG sadece 36×1 boyutlu vektörde kontrast duyarlı öznitelikleri içerirken, FHOG 31×1 boyutundaki vektörde hem kontrast duyarlı hem kontrast duyarlı olmayan ve hem de gradyan enerjisini yakalayan bilgiler içermektedir. Böylece hem hız olarak hem de performans olarak HOG'a göre daha

yüksek başarımlı elde edilmektedir. Kısacası FHOG'un daha yüksek performans sağlamasının sebebi eklenen 18 boyutlu kontrast duyarlı olmayan özniteliklerinin de FHOG vektörü içinde bulunmasıdır.



5. SONUÇ VE ÖNERİLER

Duygu Tanıma İçin Geometrik Öznitelikler Üzerinde Ardışık İleri Seçim Algoritması Sonuç ve Öneriler: Yüz ifadesi tanıma yöntemleri genelde yüz ifadelerini ayırt edici bir şekilde temsil etmek amacıyla özel tasarlanmış öznitelikler kullanırlar. Bu çalışmada az sayıda yüksek performanslı öznitelik tasarlamak ya da elle seçmek yerine çok sayıda olası öznitelik üretilmiştir. Bir özniteliğin temsil kabiliyeti en başarılı şekilde kullanılacağı sınıflandırıcıya sağlayacağı fayda ile ölçülebilir. Bu fikirden yola çıkarak, aday öznitelikler arasında beraber kullanıldığında temsil kabiliyeti yüksek olacak bir seti bulmak amacıyla ardışık ileri seçim algoritması kullanılmıştır. Bu yöntemle seçilen öznitelikler kullanılarak eğitilen DVM ile yapılan sınıflandırma, %88.7 doğruluk oranı elde edilmiştir.

Önerilen yöntemde sadece yüzde tespit edilen nirengi noktalarının geometrik hareketleri kullanılmıştır. Daha iyi sonuçlar elde edebilmek için görünüm ya da şekil tabanlı öznitelikler gibi ek öznitelikler kullanılabilir. Eklenen öznitelikleri de dahil ederek yapılacak otomatik seçimin sonuçları iyileştirmesi beklenebilir. Buna ek olarak, öznitelik seti uzayını daha kapsamlı bir şekilde arayacak bir seçim algoritması da başarıyı artıracaktır.

FHOG Öznitelikleri ile Duygu Tanıma Sonuç ve Öneriler: Bu çalışmada aydınlanma ve kontrast farklılıklarına karşı duyarsız olan FHOG öznitelikleri kullanılarak anlık yüz ifadesi tanıma gerçekleştirilmiştir. Bu sayede yüz gibi düzlemsel olmayan ve aydınlatma açısıyla bağlantılı olarak kontrast farklılıkları oluşan bir bölgeden daha gülbüz özniteliklerin elde edilmesi sağlanmıştır. Elde edilen özniteliklerin daha etkili olabilmesi için tespit edilen yüzler nirengi noktaları kullanılarak hizalanmıştır. Genişletilmiş Cohn-Kanade veri kümesi kullanılarak gerçekleştirilen deneylerde geometrik ve HOG öznitelikleriyle %88 gibi bir doğruluk seviyesine ulaşılırken; FHOG öznitelikleri ile %93 doğruluk değerlerine ulaşılmıştır.

Gelecek çalışmalarda FHOG hücrelerinden elde edilen yerel öznitelikler arasından seçim yapılarak daha yüksek doğruluk seviyelerinin elde edilmesi hedeflenmektedir. Bununla beraber görünüm tabanlı özniteliklerle beraber, yüz tanımadaki diğer bir öznitelik tipi olan geometrik özniteliklerin birleştirilmesiyle de denemeler yapılması planlanmaktadır.

KAYNAKLAR DİZİNİ

- Afshar, S. ve Salah, A. A. (2016), Facial Expression Recognition in the Wild Using Improved Dense Trajectories and Fisher Vector Encoding, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 1517–1525. ISSN: 21607516. DOI: 10.1109/CVPRW.2016.189.
- Akakin, H. n. ve Sankur, B. (2012), Spatiotemporal Features for Effective Facial Expression Recognition, 207–218. URL: http://www.busim.ee.boun.edu.tr/~sankur/SankurFolder/Conf_ECCV_Girit_2010.pdf.
- Arı, s., Alsaran, F. O. ve Akarun, L. (2011), Görü-tabanlı Gerçek-zamanlı Duygu Tanıma, *Signal Processing and Communications Applications Conference (SIU), 2011 IEEE 19th*.
- Bartlett, M. S., Littlewort, G., Fasel, I. ve Movellan, J. R. (2003), Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction., *2003 Conference on Computer Vision and Pattern Recognition Workshop*, 53–53. ISSN: 1063-6919. DOI: 10 . 1109 / CVPRW . 2003 . 10057. URL: <http://ieeexplore.ieee.org/document/4624313/>.
- Chen, J. ve Chen, Z. (2015), DYNAMIC TEXTURE AND GEOMETRY FEATURES FOR FACIAL EXPRESSION RECOGNITION IN VIDEO Junkai Chen Department of Electronic and Information Engineering , The Hong Kong Polytechnic University Department of Computer Science , Chu Hai College of Higher Education , 4967–4971.
- Dalal, N. ve Triggs, W. (2004), Histograms of Oriented Gradients for Human Detection, *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR05* 1.3, 886–893. ISSN: 1063-6919. DOI: 10 . 1109 / CVPR . 2005 . 177. arXiv: 9411012 [chao-dyn]. URL: <http://eprints.pascal-network.org/archive/00000802/>.

KAYNAKLAR DİZİNİ (devam)

Ding, Y., Zhao, Q. I. N., Li, B. ve Yuan, X. (2017), Facial Expression Recognition From Image Sequence Based on LBP and Taylor Expansion,

Ekenel, H. K. ve Sankur, B. (2004), Feature selection in the independent component subspace for face recognition, *Pattern Recognition Letters* 25.12, 1377–1388. ISSN: 01678655. DOI: 10.1016/j.patrec.2004.05.013.

Ekman, P. ve Rosenberg, E. L. (2012), *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*, 1–672. ISBN: 9780199847044. DOI: 10.1093/acprof:oso/9780195179644.001.0001. arXiv: arXiv:1011.1669v3.

Felzenszwalb, P. F., Girshick, R. B., Mcallester, D. ve Ramanan, D. (2009), Object Detection with Discriminatively Trained Part Based Models, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32.9, 1–20. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2009.167. URL: <http://cs.brown.edu/~pff/papers/lsvm-pami.pdf>.

Gacav, C., Benligiray, B. ve Cihan, T. (2017), Greedy search for descriptive spatial face features, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1497–1501. DOI: 10.1109/ICASSP.2017.7952406.

Gacav, C., Benligiray, B., Cihan, T. ve Özkan, K. (2018), FHOG Öznitelikleri ile Yüz İfadesi Tanıma, Facial Expression Recognition with FHOG Features, *Signal Processing and Communications Applications Conference (SIU)*.

Hastie, T., Tibshirani, R. ve Friedman, J. (2001), The Elements of Statistical Learning, *The Mathematical Intelligencer* 27.2, 83–85. ISSN: 03436993. DOI:

KAYNAKLAR DİZİNİ (devam)

10 . 1198 / jasa . 2004 . s339. arXiv: 1010 . 3003. URL:
<http://www.springerlink.com/index/D7X7KX6772HQ2135.pdf{\%}255Cnhttp://www-stat.stanford.edu/{~}tibs/book/preface.ps>.

Hong, Y. Z. ve Cong, L. (2014), Research of facial expression recognition based on ASM model and RS-SVM, *2014 Fifth International Conference on Intelligent Systems Design and Engineering Applications*.

Huang, X., Zhao, G., Zheng, W. ve Pietikäinen, M. (2012), Spatiotemporal local monogenic binary patterns for facial expression recognition, *IEEE Signal Processing Letters* 19.5, 243–246. ISSN: 10709908. DOI: 10.1109/LSP.2012.2188890.

Huang, X., Zhao, G., Pietikainen, M. ve Zheng, W. (2014), Robust Facial Expression Recognition Using Revised Canonical Correlation, *2014 22nd International Conference on Pattern Recognition*, 1734–1739. ISSN: 1051-4651. DOI: 10.1109/ICPR.2014.305. URL: <https://www.computer.org/csdl/proceedings/icpr/2014/5209/00/5209b734-abs.html>.

Kamarol, A., Khairuni, S., Meli, N. S., Jaward, M. H. ve Kamrani, N. (2015), Spatio-Temporal Texture-Based Feature Extraction for, 3–6.

Kazemi, V. ve Sullivan, J. (2014), One millisecond face alignment with an ensemble of regression trees, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1867–1874. ISSN: 10636919. DOI: 10.1109/CVPR.2014.241.

Kumar, P., Happy, S. L. ve Routray, A. (2017), A real-time robust facial expression recognition system using HOG features, *International Conference on Computing*,

KAYNAKLAR DİZİNİ (devam)

Analytics and Security Trends, CAST 2016, 289–293. DOI: 10.1109/CAST.2016.7914982.

Lee, K.-w., Kim, T.-h., Kim, S.-h., Lee, S.-h. ve Lee, H.-s. (2015), Facial Expression Recognition using Dual Stage MLP with Subset Pre-Training, 8.October, 1–7.

Li, J. ve Lam, E. Y. (2015), Facial Expression n Recognition Using Deep Neural Networks, 263–268.

Lin Zhong, Qingshan Liu, Peng Yang, Bo Liu, Junzhou Huang ve Metaxas, D. N. (2012), Learning active facial patches for expression analysis, *2012 IEEE Conference on Computer Vision and Pattern Recognition* 45.8, 2562–2569. DOI: 10.1109/CVPR.2012.6247974. URL: <http://ieeexplore.ieee.org/document/6247974/>.

Lo Presti, L. ve La Cascia, M. (2015), Using Hankel matrices for dynamics-based facial emotion recognition and pain detection, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops* 2015-October, 26–33. ISSN: 21607516. DOI: 10.1109/CVPRW.2015.7301351. arXiv: 1506.05001.

Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z. ve Matthews, I. (2010), The extended cohn-kande dataset (CK+): A complete facial expression dataset for action unit and emotionspecified expression, *Cvprw* July, 94–101. ISSN: 01628828. DOI: 10.1109/ISIEA.2010.5679500. arXiv: NIHMS150003.

Mistry, K., Zhang, L. ve Barnden, J. (2015), Intelligent facial expression recognition with adaptive feature extraction for a humanoid robot, *Proceedings of the International Joint Conference on Neural Networks* 2015-September. ISSN: 2161-4393. DOI: 10.1109/IJCNN.2015.7280323.

KAYNAKLAR DİZİNİ (devam)

- Mousavi, N., Siqueira, H., Barros, P., Fernandes, B. ve Wermter, S. (2016), Understanding how deep neural networks learn face expressions, *Proceedings of the International Joint Conference on Neural Networks 2016-October*, 227–234. DOI: 10.1109/IJCNN.2016.7727203.
- Rehg, J. M., Abowd, G. D., Rozga, A., Romero, M., Clements, M. A., Sclaroff, S., Essa, I., Ousley, O. Y., Li, Y., Kim, C., Rao, H., Kim, J. C., Presti, L. L., Zhang, J., Lantsman, D., Bidwell, J. ve Ye, Z. (2013), Decoding Children's Social Behavior, *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 3414–3421. ISSN: 10636919. DOI: 10.1109/CVPR.2013.438. URL: <http://ieeexplore.ieee.org/document/6619282/>.
- Ruiz, L. Z., Alomia, R. P. V., Dantis, A. D. Q., Diego, M. J.S. S., Tindugan, C. F. ve Serrano, K. K. D. (2017), Human emotion detection through facial expressions for commercial analysis, *2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)*, 1–6. DOI: 10.1109/HNICEM.2017.8269512. URL: <http://ieeexplore.ieee.org/document/8269512/>.
- Ryu, B., Rivera, A. R., Kim, J. ve Chae, O. (2017), Local Directional Ternary Pattern for Facial Expression Recognition, *IEEE Transactions on Image Processing* 26.12, 6006–6018. ISSN: 10577149. DOI: 10.1109/TIP.2017.2726010.
- Sarawagi, V. ve Arya, K. V. (2013), Automatic facial expression recognition for image sequences, *2013 Sixth International Conference on Contemporary Computing (IC3)* i, 278–282. DOI: 10.1109/IC3.2013.6612205. URL: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84886525816&partnerID=tZ0tx3y1>.

KAYNAKLAR DİZİNİ (devam)

- Shan, K., Guo, J., You, W., Lu, D. ve Bie, R. (2017), Automatic facial expression recognition based on a deep convolutional-neural-network structure, *2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA)*, 123–128. DOI: 10.1109/SERA.2017.7965717. URL: <http://ieeexplore.ieee.org/document/7965717/>.
- Suk, M. ve Prabhakaran, B. (2015), Real-time facial expression recognition on smartphones, *Proceedings - 2015 IEEE Winter Conference on Applications of Computer Vision, WACV 2015*, 1054–1059. ISSN: 1550-5790. DOI: 10.1109/WACV.2015.145.
- Sulistiyono, A. (2017), Toward to An Exaggeration Engine for Facia Animation: Evaluating the Difference of RBF Implementation in Expression-Marker Transfer, *2017 International Conference on Smart Cities, Automation and Intelligent Computing Systems (ICSC)*, 8–12.
- Swinkels, W., Claesen, L., Xiao, F ve Shen, H (2017), SVM point-based real-time emotion detection, *2017 IEEE Conference on Dependable and Secure Computing*, 86–92. DOI: 10.1109/DESEC.2017.8073838.
- Talele, K., Shirsat, A., Uplenchwar, T. ve Tuckley, K. (2016), Facial expression recognition using general regression neural network, *IEEE Bombay Section Symposium*, 1–6. DOI: 10.1109/IBSS.2016.7940203. URL: <http://ieeexplore.ieee.org/document/7940203/>.
- Viola, P. ve Jones, M. (2001), Rapid object detection using a boosted cascade of simple features, *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001* 1, 1–511–I–518. ISSN: 1063-6919. DOI:

KAYNAKLAR DİZİNİ (devam)

10 . 1109 / CVPR . 2001 . 990517. arXiv: arXiv : 1011 . 1669v3. URL: <http://ieeexplore.ieee.org/document/990517/>.

Wang, L., Wang, K. ve Li, R. (2015), Unsupervised feature selection based on spectral regression from manifold learning for facial expression recognition, *IET Computer Vision* 9.5, 655–662. ISSN: 1751-9632. DOI: 10 . 1049 / iet - cvi . 2014 . 0278. URL: <http://digital-library.theiet.org/content/journals/10.1049/iet-cvi.2014.0278>.

Zeng, Z., Pantic, M., Roisman, G. I. ve Huang, T. S. (2009), A survey of affect recognition methods: Audio, visual, and spontaneous expressions, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31.1, 39–58. ISSN: 01628828. DOI: 10.1109/TPAMI.2008.52.

Zhang, Y., Zhang, L. ve Hossain, M. A. (2015), Adaptive 3D facial action intensity estimation and emotion recognition, *Expert Systems with Applications* 42.3, 1446–1464. ISSN: 09574174. DOI: 10 . 1016 / j . eswa . 2014 . 08 . 042. URL: <http://dx.doi.org/10.1016/j.eswa.2014.08.042>.

Zhao, K., Yang, S., Wiliem, A. ve Lovell, B. C. (2017), Landmark manifold: Revisiting the Riemannian manifold approach for facial emotion recognition, *Proceedings - International Conference on Pattern Recognition*, 1095–1100. ISSN: 10514651. DOI: 10.1109/ICPR.2016.7899782.

Zhou, Y. ve Shi, B. E. (2017), Action unit selective feature maps in deep networks for facial expression recognition, *2017 International Joint Conference on Neural Networks (IJCNN)*, 2031–2038. DOI: 10 . 1109 / IJCNN . 2017 . 7966100. URL: <http://ieeexplore.ieee.org/document/7966100/>.

KAYNAKLAR DİZİNİ (devam)

Zhu, R., Zhang, T., Zhao, Q. ve Wu, Z. (2015), A Transfer Learning Approach to Cross-Database Facial Expression Recognition, 293–298.

Çam, B. C., Yalçın, M. ve Yavuz, H. S. (2016), Konumsal Öznitelikler Tabanlı Yüz İfadesi Tanıma Facial Expression Recognition Based on Locational Features, *Signal Processing and Communications Applications Conference (SIU)*.

d. A. Fernandes, J. Matos, L. N. ve d. S. Aragão, M. G. (2016), Geometrical Approaches for Facial Expression Recognition Using Support Vector Machines, *2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, 347–354. doi: 10.1109/SIBGRAPI.2016.055.

Çiğdem Turan ve Kim Man Lam (2014), REGION-BASED FEATURE FUSION FOR FACIAL-EXPRESSION RECOGNITION Cigdem Turan , Kin-Man Lam The Hong Kong Polytechnic University, *International Conference on Image Processing(ICIP)*, 5966–5970.