



T.C.
KIRŞEHİR AHİ EVRAN ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İLERİ TEKNOLOJİLER ANABİLİM DALI

YÜZ İFADELERİNİN DERİN ÖĞRENME METODU İLE SINIFLANDIRILMASI

SALAH HARAJ MESHAL MESHAL

YÜKSEK LİSANS TEZİ

KIRŞEHİR/2023



T.C.
KIRŞEHİR AHİ EVRAN ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İLERİ TEKNOLOJİLER ANABİLİM DALI

YÜZ İFADELERİNİN DERİN ÖĞRENME METODU İLE SINIFLANDIRILMASI

SALAH HARAJ MESHAL MESHAL

YÜKSEK LİSANS TEZİ

DANIŞMAN
Doç. Dr. Şekip Esat HAYBER

KIRŞEHİR/2023

TEZ BİLDİRİMİ

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

Salah Haraj Meshal MESHAL



20.04.2016 tarihli Resmî Gazete’de yayımlanan Lisansüstü Eğitim ve Öğretim Yönetmeliğinin 9/2 ve 22/2 maddeleri gereğince; Bu Lisansüstü teze, Kırşehir Ahi Evran Üniversitesi’nin aboneli olduğu intihal yazılım programı kullanılarak Fen Bilimleri Enstitüsü’nün belirlemiş olduğu ölçütlere uygun rapor alınmıştır.



ÖNSÖZ

Rahmân ve Rahîm olan Allah'ın adıyla, bu tezi tamamlayabilmem için bana verdiği kuvvet ve nimetler için Allah'a hamdolsun. Başta tez danışmanım Doç. Dr. Şekip Esat HAYBER'e göstermiş olduğu ilgi ve desteklerinden dolayı teşekkür ederim. Bilimsel tartışma ve bu çalışmanın gidişatındaki yardımları için Doç. Dr. Mohammed Basil ABDULKAREEM'e özel olarak teşekkür ederim. Son olarak, destekleri ve yardımları için anneme, babama ve eşime en içten teşekkürlerimi sunarım.

Ocak, 2023

Salah Haraj Meshal MESHAL



İÇİNDEKİLER

TEZ BİLDİRİMİ	IV
ÖNSÖZ	iv
İÇİNDEKİLER.....	v
ŞEKİL LİSTESİ	vii
TABLO LİSTESİ.....	viii
SİMGE VE KISALTMA LİSTESİ.....	ix
ÖZET	x
ABSTRACT	xi
1. GİRİŞ.....	1
1.1. Arka Plan	1
1.2. Sorun Bildirimi	3
1.3. Hedefler	4
1.4. Katkı.....	5
1.5. Tezin Yapısı	6
2. LİTERATÜR İNCELEMESİ	7
2.1 Yüz İfadesi Tanıma Sistemleri	9
2.1.1 Yüz görüntüsü temelleri	10
3. MATERYAL VE YÖNTEM.....	11
3.1 Yüz Lokalizasyonu	11
3.2 Yüz Lokalizasyonuna Giriş	14
3.2.1 Şablon Eşleme	16
3.2.2 Görünüme Dayalı Yöntemler	18
3.3 FAFE'nin Yüz Lokalizasyon Aşaması	20
3.3.1 Cilt Segmentasyonu.....	20
3.3.2 Elips Uydurma.....	21
3.4 Yüz Özellik Çekimi	22
3.4.1 Yüz Özelliği Çıkarma Teknikleri	22
3.4.2 Hedeflenen Yüz Görevleri.....	24
3.4.3 Bütünsel Açıklama	27
3.4.4 Bölgesel Açıklama.....	28
3.5 Sınıflandırma	28
3.5.1 En yakın komşu	28
3.5.2 Destek Vektör Makineleri	29
3.5.3 Nöral Ağlar	30

4. ÖNERİLEN YÖNTEM	32
4.1 Sistemin Ana Hatları.....	32
4.1.1 Önerilen CNN Algoritması.....	32
4.1.2 Evrişimli (Konvolüsyonlu) Sinir Ağlarının Mimarisi	33
4.1.3 Aktivasyon Fonksiyonları	36
4.1.4 CNN’de Önerilen Değişiklikler.....	36
4.1.5 Özellik Füzyonu	37
4.2 Veri Tabanı	39
4.3 Video Çekim Modülü	41
4.4 İfadelerin Sınıflandırılması	42
4.5 Sonuçlar	43
5. SONUÇ.....	46
5.1 Öneriler	48
KAYNAKLAR.....	49
ÖZGEÇMİŞ.....	54

ŞEKİL LİSTESİ

Şekil 1.1. Yüz görüntüsü veri kümeleri.....	1
Şekil 1.2. Yüz ifadesi tanıma sisteminin yapısı [3].....	3
Şekil 3.1. Yüzün MRI taraması [20].	11
Şekil 3.2. Yüz özelliği noktaları [25].	13
Şekil 3.3. Yüz özelliklerine göre yüz lokalizasyonu [32].	15
Şekil 3.4. Yüzün tespit edilmesinde şablon eşleştirme [39].....	17
Şekil 3.5. Cilt bölgesinin segmentasyonu [51].....	20
Şekil 3.6. PCA ekstraksiyonundan sonra yüzün özelliği [64].	23
Şekil 3.7. Yüz simgelerine göre yüz özelliklerinin sınıflandırılması [69].	26
Şekil 3.8. Destek vektör makinesi kullanılarak sınıflandırmanın yapılması.	30
Şekil 3.9. Yapay sinir ağlarının yapısı.	31
Şekil 4.1. Önerilen sistemin ana adımları.	32
Şekil 4.2. CNN algoritması için önerilen optimizasyon yöntemleri [72].....	37
Şekil 4.3. Paylaşılan CNN katmanlarında özelliklerin birleştirilmesi.....	38
Şekil 4.4. Convent mimarisinde iki özelliği birleştirme.....	39
Şekil 4.5. JAFFE veri tabanından örnek ifadeler; soldan sağa tiksinti, şaşkınlık, mutluluk, üzüntü ve korku ifadeleri [75-77].	40
Şekil 4.6. CNN'ye dayalı Haar özellik çıkarma yöntemi.....	40
Şekil 4.7. Video yakalama aşamasında algılanan yüzler.....	41
Şekil 4.8. Video çekiminde algılanan duygular.	42
Şekil 4.9. Model doğruluğu.....	44
Şekil 4.10. Yanlış ifade sınıflandırması ve doğru ifade sınıflandırması örnekleri.....	44

TABLO LİSTESİ

Tablo 4.1 Önerilen algoritmanın parametreleri	35
Tablo 4.2 Füzyondan önceki ve sonraki özellik boyutları.....	39
Tablo 4.3 Önerilen sistemin karışıklık matrisi.....	43
Tablo 4.4 JAFFE 1M veri setindeki döneme göre doğruluk	45
Tablo 4.5 CNN'e dayalı farklı algılama algoritmalarının doğruluğunun karşılaştırılması.	45



SİMGE VE KISALTMA LİSTESİ

Simgeler **Açıklama**

%	: Yüzde
g	: Gram
>	: Büyüktür

Kısaltmalar **Açıklama**

CNN	: Konvolüsyonlu Sinir Ağı
LBP	: Yerel ikili model
FACS	: Floresan Aktif Hücre Sıralama
MRI	: Manyetik Rezonans Görüntüleme
SVM	: Destek vektör makineleri
HOG	: Gradient histograms
ABD	: Amerika Birleşik Devletleri
MLP	: Çok katmanlı algılayıcılar
CONV	: İlk konvolüsyon katmanı
SGD	: Stokastik gradyan iniş

ÖZET

YÜKSEK LİSANS TEZİ

YÜZ İFADELERİNİN DERİN ÖĞRENME METODU İLE SINIFLANDIRILMASI

Salah Haraj Meshal MESHAL

Kırşehir Ahi Evran Üniversitesi

Fen Bilimleri Enstitüsü

İleri Teknolojiler Anabilim Dalı

Danışman: Doç. Dr. Şekip Esat HAYBER

Duygu, insan zihninde var olmanın psikolojik bir halidir. Birçok alanda yapılan çeşitli çalışmalar, duygu oluşum sürecine ilişkin farklı bakış açılarını desteklemektedir. Bazı filozoflara göre duygu, kişisel koşullardaki veya çevredeki değişikliklerin sonucudur. Bununla birlikte, bazı biyologlar, nörolojik ve hormonal sistemlerin, duyguların oluşumunda başlıca sorumlu olduğuna inanırlar. Tezimizde, konvolüsyonlu sinir ağlarına (ConvNets) dayalı bir yüz ifadesi tanıma yaklaşımı oluşturmak için motive olduk. Konvolüsyonlu sinir ağı modelleri, sinir, mutluluk, öfke, sürpriz, üzüntü veya tiksinti gibi kategorilerden birine ait bir yüz ifadesi etiketini tahmin etmek için kullanılır. Yüz ifadesinin etiketi sinir, mutluluk, öfke, şaşkınlık, üzüntü veya tiksinti gibi kategorilerden birine karşılık gelmelidir. Görüntü, sistemde bir girdi görevindedir. Sonuç olarak, bu modeller sıkı testlerden geçmektedir. Yakalama ve sınıflandırma modülü, bir bilgisayar sisteminin kullanıcı üzerinde meydana getirebileceği duygusal etkinin bilinmesini ve bu bilgilerle kullanıcı etkileşimini iyileştirmek için önlemler alınmasına olanak sağlar. Böylece, kullanılabilirlik testlerine uygulanan ifadeleri sınıflandırmak için CNN kullanmanın mümkün olduğu, hatta testi yapanlara gerçek zamanlı yanıtlar verebileceği sonucuna varmaktayız. Bu tezin ana katkısı, CNN algoritmasını lineer aktivasyon fonksiyonlarıyla modifiye ederek algılamının doğruluğunu artırmak için CNN'nin yerleşik softmax katmanını diğer sınıflandırıcılarla değiştirmemizdir. Değiştirilen CNN algoritması hem Conv-2 katmanını kullanarak özellikleri ayıklar, hem de iş akışının hızını artırmak için özellik şablonlarını ve Haar benzeri özellikleri kullanır.

Ocak 2023, 66 Sayfa

Anahtar Kelimeler: CNN, DL, ML, ANN, AI

ABSTRACT

M.Sc. THESIS

CLASSIFICATION OF FACE EXPRESSIONS BY DEEP LEARNING METHOD

Salah Haraj Meshal MESHAL

Kirsehir Ahi Evran University

Graduate School of Sciences and Engineering

Advanced Technologies Department

Supervisor: Assoc. Prof. Dr. Şekip Esat HAYBER

Emotion is a psychological state of being in the human mind. Various studies in various fields support various viewpoints on the process of emotion formation. Emotion, according to some philosophers, is the outcome of changes in personal conditions or the environment. Some biologists, however, believe that the neurological and hormonal systems are principally responsible for the genesis of emotion. We were motivated to create a facial expression recognition approach based on convolutional neural networks (ConvNets). Convolutional neural network models are used to predict a facial expression label belonging to one of the following categories: neutral, happiness, anger, surprise, sadness, or disgust. The label for the facial expression must correspond to one of the following categories: neutral, happiness, anger, surprise, sadness, or disgust. The image serves as an input to the system. Consequently, these models undergo rigorous testing. The capture and classification module makes it possible to know the emotional impact that a computer system can cause on the user, and with this information to take measures to improve user interaction. Thus, we conclude that it is possible to use CNN to classify expressions applied to usability tests and can even provide real-time responses to those who perform the test. The main contribution of this thesis is that we modify the CNN algorithm with linear activation functions and replace the built-in SoftMax layer of the CNN with other classifiers to increase the accuracy of the detection, the modified CNN algorithm not only extracts the features using the Conv-2 layer but also uses feature templates and Haar like features to increase the speed of the workflow.

January 2023, 66 Pages

Keywords: CNN, DL, ML, ANN, AI

1. GİRİŞ

1.1. Arka Plan

Yüz ifadesi analiz algoritmalarının kullanımı, öğrencilerin akademik başarısını artırmanın bir yöntemi olarak rağbet görmektedir. Bir öğrencinin zihinsel durumunu anlamak için, araştırmacılar genellikle yüz analizini içeren teknolojileri kullanırlar (Şekil 1.1).



Şekil 1.1. Yüz görüntüsü veri kümeleri.

Öğrenciler, bu tür taktikleri kullanmanın, akademik çabaları sırasında ortaya çıkabilecek engellerin üstesinden gelinmesinde ve ilerlemek için cevaplar bulunmasında onlara yardımcı olduğunu görebilir. Yüz izleme, çıkarma ve sınıflandırmayı etkin bir şekilde yürütebilmek için araştırmacıların gerçek zamanlı olarak güncellenen verilere erişimi olması gereklidir. Bu geri bildirim sağlama yöntemi, kabul edilebilir sonuçlar sağlama yeteneğine ve öğrenme sürecini yavaşlatma potansiyeline sahiptir. Bu veriler, bir öğrencinin duygusal durumundan kaynaklanabilecek hataları belirlemek için kullanılabilir ve daha sonra bu hatalara yanıt olarak düzeltici önlemler alınabilir. Son yıllardaki teknolojik atılımlar, bir kişinin sadece fiziksel konumunu değil, aynı zamanda zihinsel durumunu ve yüz özelliklerini de

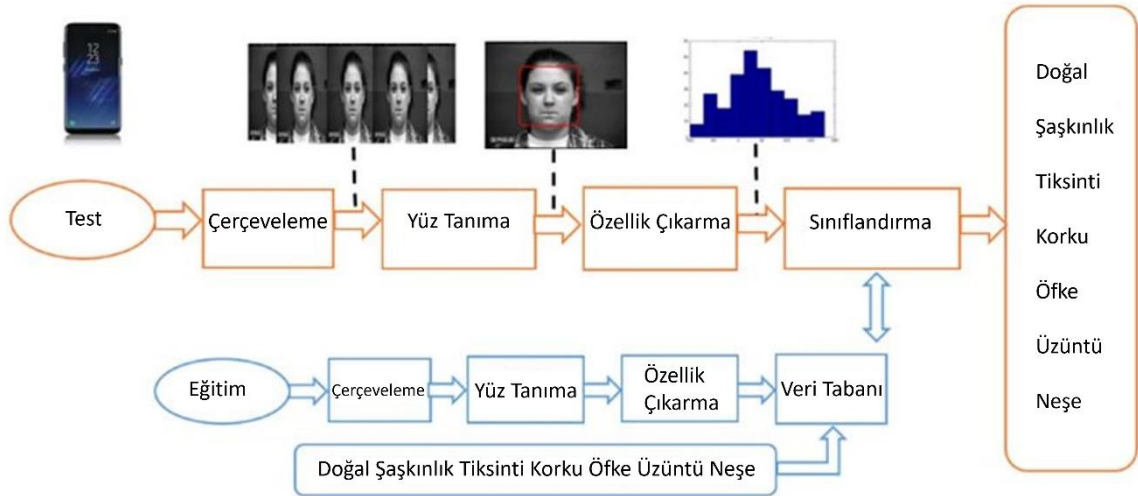
sınıflandırmayı mümkün kılmıştır. Bu yöntem hem dijital kameralar hem de Facebook'un yüz ifadesi özelliği tarafından kullanılmaktadır. Ancak, önceden var olan sınıflandırıcıları kullanmakla ilgili bir sorun söz konusudur.

Çok çeşitli bağlamlar olduğundan, sınıflandırıcılar gerçek ve uydurma duygusal tepkiler arasında ayırım yapmak için eğitilmelidir [1]. Bu ifade, benzerleriyle birlikte, çeşitli veri tabanlarına dahil edilmiştir. Son beş yılda, politikanın etkisizliği daha belirgin hale gelmiştir. Şiddetli duygu sergileyen bir yüz ifadesi üretmesi için talimat verilen bir katılımcının duygusal durumunu kaydetmek ve değerlendirmek mümkündür. Bu sınıflandırıcıların başarı oranı yüzde doksandan fazladır. Bunun doğrudan bir sonucu olarak, yüz veri tabanı toplama metodolojileri oluşturulmuştur. Eğer öyle yapılmış olsaydı, duyguları eğer varsa, bu şekilde iletilirdi. Halihazırda mevcut araçların otantik deneyimleri kaydetme ve farklı duygusal durumları ayırt etme yeteneğine sahip olup olmadığı henüz belli değildir. Nasıl hissettiğine bağlı olarak, kişinin zihinsel durumu "duygusal" olarak tanımlanabilir. Konuyla ilgili çok sayıda çalışma yapılmış ve bu çalışmaların sonuçları tutarsız sonuçlar vermiştir. Bazı filozoflar, duyguların bir kişinin koşulları veya çevresi gibi (iyi veya kötü) dış veya iç unsurlardan kaynaklandığını düşünmektedir. Bazı araştırmalar, beynin nörolojik ve hormonal sistemlerinin bir kişinin duygusal gelişiminden başlıca sorumlu olduğunu düşündürmektedir. Hepimiz duygularımıza bağlı olarak görünüşümüzün değişebileceğinin farkında olsak da bu değişikliklerin nedenleri hakkında bir fikir birliği yoktur. Yüz ifadeleri, insan ilişkilerinin ve sözsüz iletişimin temel bir bileşeni olmasının yanı sıra, bir kişinin duygusal durumu hakkında önemli bir bilgi kaynağıdır. Sözlü iletişimle birlikte kullanılmasalar bile, yine de çok fazla bilgi aktarabilirler. Bir mesajın etkisinin yüzde 7'si söylenen gerçek kelimelere atfedilebilirken, yüzde 38'i mesajın sesli kısmına (tonlama) ve yüzde 55'i konuşmacının yüz ifadesine bağlanabilir. Bir kişinin duygusal durumunu yüz ifadelerine dayalı olarak yorumlamanın belirsizliğine iki etken sebep olur. Birinci etken, kişinin ruhsal durumunun yüz ifadesine tam olarak yansımama olasılığının varlığıdır [2]. Diğer etken ise yüz ifadelerini doğru bir şekilde değerlendirmek için çok hassas teknolojilerin gerekliliğidir. İlk zorluk, araştırmaya katılanların kendi duygularını bastırma veya üretme yeteneklerine bağlı olduğu için üstesinden gelinemez. Bu durum sorunu çözülemez hale getirmektedir. Katılımcıların yüz ifadesi tanıma araştırmaları için uygun oldukları varsayılarak, otomatik tanıma sistemleri bunu çözebildiğinden yalnızca ikinci sorun ele alınmıştır.

Bir kişinin yüz ifadelerini ayırt etme yeteneğinin çoğu, yüzün göz ve ağız alanlarının benzersiz özelliklerine bağlıdır. Bu nedenle bilgisayar bu alanları tanımalı ve kullanıcının yüzü tarafından gönderilen duyguyu doğru bir şekilde tanımlamak için gerekli özellikleri çıkarmalıdır. Özellikler, makine öğrenimi, örüntü tanıma ve görüntü işlenmesindeki ilk veri koleksiyonundan çıkarılır. Elde edilen özellikler değer sağlamalı ve gereksiz olmamalıdır. Resim özelliklerinin örnekleri; görüntünün noktaları, kenarları ve nesnelere gibi yapılarını içermektedir. Ek örnekler, bir dizideki bir görüntüden diğerine geçişi, bir görüntünün çeşitli bölgeleri arasındaki eğriler ve sınırlar tarafından üretilen formları ve sınırları ve ayrıca belirli alanlarda bulunabilecek renkleri veya dokuları içermektedir. Bir özellik fikrinin açık uçlu doğası nedeniyle, bir bilgisayarlı görme sistemindeki özelliklerin seçimi, o sırada yapılan etkinliğe bağlı olabilir.

1.2. Sorun Bildirimi

Bu projenin uygulanmasıyla, insanların yüzlerinden üst düzey bilgiler otomatik olarak çıkarılacaktır. Yüz tanıma teknolojisini en çok kullanan iki uygulama yüz tanıma ve yüz ifadesi tanıma uygulamalarıdır. Bu uygulamalar bu alanda yaygın olarak kullanılmaktadır.



Şekil 1.2. Yüz ifadesi tanıma sisteminin yapısı [3].

Bir görüntüde kimin veya neyin gösterildiğinin belirlenmesi ve aynı kişinin iki fotoğrafının benzer olup olmadığının belirlenmesi gibi görevler için yüz tanıma gereklidir. Örneğin, bir sırtıma veya kaş çatma başlı başına bir yüz duygusu (gülme, ağlama veya konuşma gibi) veya daha büyük bir yüz ifadesinin (gülümseme gibi) bir bileşeni (kapalı gözler, açık ağız) olarak kabul edilebilir. Bu alanda araştırmalar yapılırken, çeşitli faktörlerin ele alınması gereklidir. Çünkü yüzler, üretilen birçok ürünün aksine, çeşitli boyut, şekil ve fiziksel görünümlere sahiptirler. Bilgisayar görüşüne dayanan uygulamalar bundan faydalanabilirler

ancak bir takım olası dezavantajlar da vardır. Yüzler son derece tanımlanabilir özelliklerdir, bu nedenle bir kişiyi sadece onlara bakarak tanımlamak mümkündür. Orijinal ve sahte ürünler arasındaki bariz benzerlik nedeniyle, ikisini birbirinden ayırt etmek için ürün numaraları veya plakalar gibi tanımlayıcı işaretlerin kullanılması hayati önem taşımaktadır. Görünür benzerliklere rağmen, hiçbir iki insanın görünüşü tamamen aynı değildir. Sonuç olarak, kendi yaşam deneyimlerimizin doğrudan bir sonucu olarak, diğer insanları sadece görünüşlerini görerek tanıyabiliriz. Bilgisayarlı görme perspektifinden bakıldığında, yüz tanıma sistemleri yüksek düzeyde güvenilirliğe ek olarak bir dereceye kadar uyarlanabilirliğe sahip olmalıdır. Yüzün; iki gözün, bir burnun ve bir ağzın varlığı gibi temel nitelikleri yaygın olarak bilinse dahi bir yüzün yapısı ve dokusu ancak internet araştırması ile keşfedilebilir. Yüz modelleri, belirli bir uygulama için gerekli miktarda ölçeklenebilirlik ve güvenilirlik sağlamak için kullanılabilir. Yüz modelleri, yalnızca birkaç tanımlayıcı model parametresi kullanarak bir fotoğrafta yer alan bilgileri ortadan kaldırabilir. “Yüz hizalama” olarak da bilinen yüz modeli yerleştirme sürecinde, izleyicilere uygun bilgiler sunmak için modelin resmin konusuna karşılık gelen özellikleri belirlenir. Fotoğrafın aydınlatması ve arka planı, genel soruna sebep olan unsurlar olabilir. Yüz görüntüsü analizi genellikle iyi yapılandırılmamış veya kontrol edilmemiş durumlarda kullanılır. Sonuç olarak, kamuya açık verilerin büyük bir kısmı, aydınlatmanın kontrol edildiği ve öznenin sürekli olarak doğrudan kameraya baktığı daha doğal bir ortamda değil, laboratuvarlarda ve ofislerde çekilen yüz görüntülerini göstermektedir. Gerçek dünyadan sınırlı veri bulunması nedeniyle, sağlam algoritmalar oluşturmak ve bir algoritmanın sağlamlığını göstermek zorlu bir iştir.

1.3. Hedefler

Bir sonraki bölümde de gösterileceği üzere, bu sorunları etkili bir şekilde ele almak için duygu tanıma teknolojilerinin gerçek zamanlı kullanımı gereklidir. Yaklaşım, daha sonra çizgi grafikler olarak gösterilen iki boyutlu duygusal ifadeleri haritalamak için yüz fotoğraflarını kullanır. Modelin uyarılma ve değerlik bileşenleri, bir bireyin duygusal durumunu analiz etmek için kullanılan iki farklı türde değerlendirmeyi sınıflandırmak amacıyla kullanılabilir. Her ikisi de bireyin ruh halinin ve enerjisinin yoğunluğunu gösterdiğinden, değerlik ve uyarılma yönleri arasında ayırım yapmak önemlidir. Bu ölçeğin kullanılmasıyla sevinç veya korku belirtilebilir, bir kişinin mevcut uyarılma düzeyi değerlendirilerek düşükten yüksek uyarılma derecelerine mi yoksa değerlik derecelerine mi ilerlediği anlaşılabilir. Bir robot verilere dayalı olarak etkinlikleri duygusal tepkilerle

eşleştirebilecektir. Diğer sistemlerde olduğu gibi bu sınıflandırma metodolojisi, sadece ayrı duyguların tanımlanmasına dayanmaz, aynı zamanda bir kullanıcının ruh halinin üzgünden mutluya değiştiği zamanları da algılayabilir. Bu araştırma, sistemi gerçek dünyada insan-robot etkileşimine uygun hale getirmeyi amaçlamıştır. Bu amaçla insanların gerçek yüz ifadelerinin çok sayıda görüntüsü çekilmiştir. Yüz özelliklerini tanımlamak ve duygusal durumları gerçek zamanlı olarak sınıflandırmak için kısıtlı yerel modeller ve destek vektör makineleri kullanılacaktır. Bu sayede, sistemin doğru çalışmasını sağlanacaktır.

Yüz özellikleri oldukça açık hedeflerdir. Bir özellik çıkarma programı kullanılarak hem resimlerden hem de filmlerden özellikler çıkarılabilir. Yüz ifadelerini ayırt edecek bir sistemin eğitilmesi amacıyla blok eşleştirme hareket tahmin yöntemleri kullanılmalıdır. Çeşitli hareket biçimlerini ayırt edebilen bir sınıflandırma sisteminin geliştirilmesi gerekmektedir.

1.4. Katkı

Bu bölümde, tezin aşağıdaki bileşenlerini daha ayrıntılı olarak inceleyeceğiz: Bu bölüm, çalışmanın metodolojisinden veri toplama ve hazırlamaya kadar her şeyi ve ayrıca veri artırma üzerine bir deneyi, önerilen sistemin uygulanmasını, çalışmanın sonuçlarını, bulguların tartışılmasını ve gelecekteki araştırmalar için önerileri içermektedir. Kullanılacak yöntem; veri toplama, veri hazırlama ve görüntü iyileştirme olmak üzere üç bölümden oluşmaktadır. Başlamak için bir model oluşturularak, bu model üzerinde uygulama yapılır. Bu çalışmanın öncelikli amacı, bir grafik kullanıcı arayüzü kullanarak yüz ifadelerini tanımlama ve tasvir etme yöntemi oluşturmaktır. Görüntü işleme işlemi sırasında konuyla ilgili çeşitli çalışmalar dikkate alınacaktır. Bu araştırma, yüzlerin tanınmasını, ilgi alanlarının çıkarılmasını ve yüz duygularının sınıflandırılmasını içermektedir. Bu çalışmanın amacı, çeşitli araştırma projelerinin iyi özelliklerini bir araya getirerek, insani açıdan verimli ve olabildiğince insani algoritmalar kullanan bir çözüm sunmaktır. Bu çalışmanın sonuçları, ifadelerle etkileşim için yenilikçi yöntemlerin tavsiye edilmesine yol açacaktır. Yöntemin başarısı, kullanıcının duygusal durumu hakkında ifadeye dayalı bilgileri iletmek için kullanıcının çevresinin görüntülerini toplayıp analiz etmeye dayanır. Bunun doğrudan bir sonucu olarak, gerçek dünya senaryolarında algoritmaların performansını değerlendirmek için bir dizi prototip oluşturmak gereklidir. Bu amaca ulaşmak için aşağıdaki üç yöntem oluşturulmuş ve her biri ayrıntılı olarak 4. bölümde ele alınmıştır:

- Yüz ifadelerinin kapsamlı bir şekilde toplanması ve sınıflandırılması ve ayrıca kategorizasyonla ilgili diğer bilgiler,
- Herhangi bir aktif pencereye ifadelerin enjeksiyonu,
- Bir müzik çalar çerçevesinde şarkıların kullanıcı ifadelerine göre sınıflandırılması.

Bu tezin ana katkısı, CNN algoritmasını lineer aktivasyon fonksiyonlarıyla modifiye etmemiz ve algılamının doğruluğunu artırmak için CNN'nin yerleşik softmax katmanını diğer sınıflandırıcılarla değiştirmemizdir. Değiştirilmiş CNN algoritması, yalnızca Conv-2 katmanını kullanarak özellikleri çıkarmakla kalmaz, aynı zamanda iş akışının hızını artırmak için özellik şablonları ve Haar benzeri özellikler kullanır. Bundan sonra, hesaplamalar için gereken süreyi optimize etmek için farklı katmanlardan aynı ağırlıklara sahip bazı özellikleri birleştiriyoruz.

Ek olarak, sınıflandırıcının etkinliği bir dizi bilgisayar tabanlı test kullanılarak değerlendirilmiştir. Bir deneyin etkili olup olmadığını değerlendirmek için, kullanıcının duygularını doğru bir şekilde yansıtan bir ifade ile diğer tüm olası ifadeler arasında ayrım yapmanın mümkün olup olmadığı incelenmektedir.

1.5. Tezin Yapısı

Tezimizin temel konuları aşağıdaki sıra ile işlenecektir: 2. bölümde, en eski akıllı ağ araştırma ve uygulama etkinliklerinden bazıları incelenecektir. 3. bölümde, çeşitli bileşenlerin bir özeti verilecektir. 4. Bölümde, modelimizin kapsamlı bir tanımı sunulacaktır. 5. Bölümde modelimizin simülasyon yoluyla nasıl kullanılabileceği açıklanacak, simülasyonumuzun sonuçları sunulacak ve gelecekte yapılacak çalışmalar için önerilerde bulunulacaktır.

2. LİTERATÜR İNCELEMESİ

Araştırmacılar, sadece yüzünü gözlemleyerek bir kişinin duygusal durumunu belirlemek için kullanılabilir yüz ifadelerinde bir model olduğuna inanmaktadırlar. Yunan filozofları, MÖ dördüncü yüzyılda bir bireyin kişiliğini ve karakterini değerlendirmenin bir yolu olarak yüz ifadelerinin incelenmesine öncülük etmişlerdir. Yüz ifadelerinin bu bakış açısıyla incelenmesini yapan ilk kişiler onlar olmuştur. 1970'lerde Paul Ekman, yüz ifadelerinin incelenmesinde bilimsel yöntem uygulayan ilk kişi olarak birkaç kültürde yüz duygularının tasviri üzerine araştırmalar yapmıştır. O zamanlar insanlar tarafından, dünyanın farklı bölgelerinin farklı dilleri yorumlamalarına benzer şekilde, yüz ifadelerinin sosyal normlar ve edinilmiş alışkanlıklar tarafından belirlendiği düşünülüyordu. Ekman, çeşitli yüz ifadelerini betimleyen bir resim koleksiyonuyla Japonya, Brezilya ve Arjantin'e uçtu. Deneyi dünyanın her yerinden gönüllüleri içermekte olup hepsinin yüz ifadelerinin önemi konusunda aynı anlayışı paylaştığını öğrenince çok şaşırıldı. Çalışmasını gerçekleştirdiği Papua Yeni Gine'nin en ıssız bölgelerinde bile, yerlilerin insanların yüzlerinde gösterilen duyguları tanımlayabildiğini keşfetti. Papua Yeni Gine ormanlarında yapılmış olan bu deneylerden sonra yüz ifadelerinin incelenmesi büyük ilgi görmüştür [4]. Ekman'ın grubu tarafından oluşturulan Yüz Hareketi Kodlama Sistemi (FACS), yüz ifadelerini kodlamak için oluşturulmuş ilk sistemdir. Hem animatörler hem de psikologlar, vücuttaki duyguların fiziksel tezahürlerini tam olarak tanımlamak için bir kriter olarak FACS'yi faydalı bulmuşlardır [5]. FACS kullanılarak, tıbbi olarak akla gelebilecek hemen hemen her yüz ifadesi manuel olarak sınıflandırılabilir. Bu durum, insan etiketleyicilerinin ifadeyi bileşen parçalarına ayırmasına ve oluşturulmasından sorumlu geçici bölümleri kesin Eylem Birimlerine (AU) dönüştürmesine olanak tanımaktadır [6]. Zeka bağlamında AU'lar, temel duyguların tanınması ve önceden programlanmış komutların verilmesi gibi kapsamlı karar verme süreçleri için kullanılma potansiyeline sahiptir. Bu konunun önemi, hem Akademik Birimler (AU'lar) hem de Dr. Ekman için bin sayfadan fazla FAC rehber kitabında özetlenmiştir. Belirli koşullar altında, otomatik yüz ifadesi tanıma sistemleri AU'ları kullanmaktadır. Robotik ve otomatik işler için yazılımdaki son gelişmeler nedeniyle, güvenilir bir yüz ifadesi tanıma sistemine duyulan ihtiyaç artmaktadır. Bilgisayarların ve diğer otomatik sistemlerin bir gün insanların duygusal durumlarını tıpkı insanlar gibi okuyabileceğine inanılmaktadır [7]. Bilim insanları, insan-makine etkileşimlerinin geliştirilmesinde büyük ilerleme kaydetmiştir. Aynı zamanda daha fazla kişi, insanların ve robotların birbirleriyle nasıl iletişim kurduklarını araştırmaktadır. Kinect ve Wii gibi video

oyunlarında hareket sensörlerini kullanarak gerçek ve sanal dünyaları tek bir deneyimde birleştirmek mümkündür. Sürücü uykusunu algılayabilen sensörlerle donatılmış araçlar, trafik kazası olasılığını en aza indirebilir. Kısa bir süre önce, yapay zekaya (AI) sahip refakatçi robotlar ticari olarak erişilebilir hale gelmiştir. Yüz ifadesi analizi bu uygulamalar için oldukça değerlidir [8].

Sonuç olarak, giderek daha fazla kişi duygularını yorumlamak için yüz ifadelerine bağımlı hale gelmektedir. Yaş, etnik köken, cinsiyet ve sakal gibi özellikler, birini sadece fiziksel görünümüne göre tanımlamaya çalışırken göz önünde bulundurulması gereken unsurlardan sadece birkaçıdır. 1990'lar boyunca ve 2000'lerin başında, araştırma ve geliştirme çalışmalarının çoğu, otomatik yüz duygu tespiti için teoriler ve teknikler üzerinde yoğunlaşmıştır. Bir kişinin duygularının ifadesi, öznenin yüzünün statik veya dinamik görüntülerinin kullanılması da dahil olmak üzere çeşitli şekillerde iletilebilir. Sony kameralar, tıbbi görüntü analizi için kullanılan gülümseme tanıma ek olarak ağrı algılama özelliğine sahiptir [9]. Onlar için diğer olası kullanım, mevcut özelliklerini geliştirmek olacaktır. Son yıllarda, yüz tanıma ve yüz duygularının tespiti için kullanılan yerel ikili model (LBP) olarak bilinen basit ve verimli bir operatörün kullanımında artış olmuştur. Bu iyimser sonuçlara rağmen, LBP yüz ifadesi tanıma sistemi henüz tam olarak geliştirilmemiştir. Bazı araştırmalar, statik bir duruşta çekilen fotoğrafların yüz ifadelerindeki ince değişiklikleri yakalayamayacağını göstermektedir [10]. Bunlar, yüz duygu algılamasını iyileştirmek için önerilen LBP'lerin dinamik uzantılarından biri olan dinamik resim LBP'lerinde genel bir çerçevenin ne kadar kolay genişletilebileceğini göstermiştir [11]. Ayrıca bu tasarımın işlevselliği maliyet tasarrufu açısından oldukça verimlidir. Aslında yüz ifadelerini yorumlayabilen yepyeni, dinamik bir LBP geliştirilmektedir. Yüzlerin statik ve dinamik fotoğraflarına dayalı çeşitli teknolojilerin geliştirilmesi, otomatik yüz ifadesi tanımlamanın uygulanmasını mümkün kılmıştır. Yüz özelliklerini göstermek için çok sayıda yöntem vardır. Üç farklı yöntem incelenecektir ve her birinin kendine özgü geometrisi ve görünümü vardır (hibrit özellik tabanlı strateji) [12]. Yüz işaretleri güvenilir bir şekilde tanınıp takip edilebildiğinde, yüz duygularını kavramak ve ayırt etmek için geometrik özellikler tercih edilir [13]. Yüz özelliği noktalarının geometrik yer değiştirmesini tanımlamak için görsel bilgileri geometrik yer değiştirme ölçümleriyle birleştirmek ve yüzün genel bir geometrik grafik temsilini üretmek önemlidir. Bir dizi yüz ifadesini analiz ederek yüz anatomisini daha iyi anlamak için kapsamlı araştırmalar yapılmıştır. Yüze geometrik noktaların yerleştirilmesi sayesinde yüzün

yapısının tespiti hızlı ve kolay bir şekilde yapılabilir. Zhang vd. çok faydalı bir örnek vaka sunmaktadır (1998). Bu deneyde, 34 referans noktası yüzün geometrisini temsil etmiştir. Resim verilerini kullanarak, bu noktalar daha sonra doğrusal bir özellik vektörü oluşturmak için kullanılmıştır. Durağan fotoğraflarda bir kişinin yüzünde görünen duygu, Rudovic vd. tarafından sağlanan 39 yüz referans noktası kullanılarak kategorize edilebilir (2013). Son zamanlarda özellikle dinamik yüz fotoğrafçılığı alanında bu yöntemlere büyük ilgi söz konusudur [14].

2.1 Yüz İfadesi Tanıma Sistemleri

Herhangi bir yüz tanıma sisteminin birincil amacı, deneğin yüz ifadesindeki dalgalanmaların neden olduğu yanlış pozitif tanımlamaların sıklığını azaltmak olmalıdır [15]. Bu bilgi, bir kişinin duygusal durumunu tanımlayabilen ve ardından uygun yüz ifadesiyle tepki verebilen bir aracı meydana getirmek amacıyla kullanılabilir [16]. Bir bilgisayar, bir bireyin duygusal durumunu ayırt edebildiğinde, insanların gerçek dünyada nasıl davrandığına benzer şekilde insan duygularını, ruh hallerini ve niyetlerini kavrayabilecek ve bunlara tepki verebilecektir. Yüz tanıma yazılımının temel amacı, yüz özellikleri ve yüz hareketlerindeki değişiklikleri inceleyip tespit edebilen görsel girdi tabanlı sistemler oluşturmaktır. Bir ifade tanıma sisteminin, yüzünde belirli bir duyguya sahip olduğu belirlenmeden önce, bir dizi prosedürü tamamlaması gereklidir. Öncelikle sistemin eğitime ve test edilmesine temel oluşturacak fotoğraf veya filmler belirlenmelidir. Bazı alınan her nitelik, onu diğerlerinden ayıran benzersiz özelliklere sahiptir. Bu özellikler, yüzün görüntülerde görüldüğü açıyı, erişilebilir doğal ışığın miktarını ve diğer değişkenleri içermektedir. Bir ifade tanıma sisteminin doğruluğunun, yüz ifadelerini tespit etmeye çalışırken kullandığı temelin türüne bağlı olarak değişmesi düşünülebilir. Taban çizgisinin belirlenmesinden sonraki aşama, sisteme girdi görevi görecektir bir resim veya videoda bireyin yüzünün yerinin belirlenmesidir. Bu çok önemli bir adımdır. Bundan sonra sistem, duygunun kaynağını belirlemek için yüz görünümünde veya yüz kaslarında herhangi bir değişiklik olup olmadığına dikkat etmelidir. Öznenin yüzünü kısmen veya tamamen gizleyebilecek engeller nedeniyle, öznenin yüzündeki duygunun belirli konumunu bulmak ve izlemek zor olacaktır. Bir yüz tanındığında, sistem, gözlenen niteliklere dayalı olarak gösterilen duyguyu belirlemek için dudakların, kaşların ve yanakların hareketi gibi belirli özellikleri incelemektedir.

2.1.1 Yüz görüntüsü temelleri

Herhangi bir yüz duygu tanıma sisteminin geliştirilmesi için esas olan, fotoğraf temelinin seçimidir. Tüm araştırmacıların aynı fotoğraf temeline erişimi olsaydı, geliştirilen sayısız sistem arasında testler yapmak ve karşılaştırmalar yapmak çok daha kolay olurdu. Araştırmacılar, bir özelliğin kalıcı mı yoksa geçici mi olduğunu belirlemek için gözlemlenebilir özelliklere dayalı bir strateji kullanır. Dudaklar ve gözler gibi kalıcı yüz özellikleri için izleyiciler vardır. Kenar algılama algoritmaları, kırıksıklıklar gibi yüz özelliklerini tanımlamak için kullanılabilir. Buna karşılık, iyi bir estetik temelin mevcudiyeti çok önemlidir. Yüz Tanıma Teknolojisi (FERET) veri tabanında, insanların yüzlerinin görüntüleri gelişim ve izolasyon olmak üzere iki benzersiz kategoriye ayrılır [17]. Geliştirme grubu, çeşitli disiplinlerden araştırmacıları içerirken, izolasyon grubu sadece yüz tanıma algoritmalarını test etmek için kullanılır. Bir yüz tanıma sistemini değerlendirmek için FERET değerlendirme işlemini bağımsız bir taraf yönetir. Testin amaçları; çeşitli algoritmaların doğrudan karşılaştırılması, uygun metodolojilerin belirlenmesi, yüz tanımadaki teknolojinin gözden geçirilmesi ve gelecekteki olası araştırma yollarının belirlenmesidir. Ayrıca araştırmada katılımcıların yüz tanıma becerilerinin artırılması hedeflenmiştir. Ağustos 1993 ile Temmuz 1996 arasında, FERET veri tabanından alınan veriler, on beş farklı oturum sırasında toplanmıştır. Bu koleksiyonda 1.199 kişinin fotoğrafları ve 365 adet birbirinin aynısı görsel olmak üzere toplam 14.126 fotoğraf bulunmaktadır [18]. Bir kişi, ilk günkü çekim ile aynı zamana denk olmayan başka bir günde ikinci fotoğraf setini çekerek bir dizi görüntü oluşturabilir. Bazı kişiler için, birinci ve ikinci resim grupları arasında iki yıllık bir ara vardır. Aradaki zaman farkı nedeniyle, araştırmacılar bir kişinin fiziksel görünümünde bir yıl veya daha uzun süre boyunca meydana gelen değişiklikleri ilk kez görebilmişlerdir [19].

3. MATERYAL VE YÖNTEM

3.1 Yüz Lokalizasyonu

Yüz tanıma, son yıllarda pek çok çalışmanın odak noktası olmuştur. Bilgisayarla görme, robotlara görme kabiliyeti kazandırmayı amaçlayan bir yapay zeka alt alanı olarak ortaya çıkmış ve çok daha büyük bir çalışma alanına doğru genişlemiştir. Bilgisayarla görme uygulamaları söz konusu olduğunda, resimler elektromanyetik spektrumun birçok bölümü kullanılarak işlenebilir. X-ışınları, ameliyat gerektirmeden insan vücudunun resimlerini oluşturmak için kullanılabilir. Gama radyasyonu ve radyo dalgalarının kullanımıyla bu resimleri oluşturan manyetik rezonans görüntüleme (MRI), kanser gibi hastalıkları tanımlamak ve tedavi etmek için kullanılabilir.



Şekil 3.1. Yüzün MRI taraması [20].

Jant dökümleri gibi görsel olarak tespit edilmesi zor olan malzemeleri kırıklar, kabarcık boşlukları ve döküm işleminde füzyonun olmadığı arızalar açısından kontrol ederken, değerlendirme sürecine yardımcı olması için X-ışınları kullanılabilir. Gıda şirketleri, ürünlerinin güvenliğini ve kalitesini sağlamak için denetim süreci boyunca X-ray ve gama

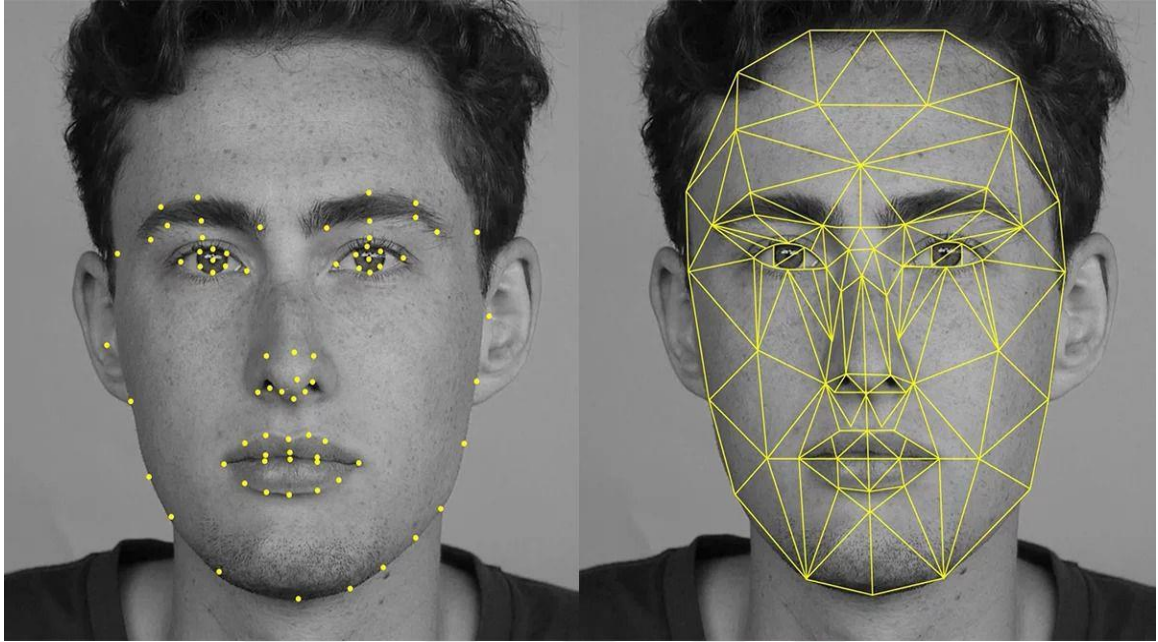
radasyonu kullanır [21]. Paketlenmiş balıklarda kılçık bulunup bulunmadığının belirlenmesi, turuncgillerin böcekler tarafından istila edilip edilmediğinin tespit edilmesi, yemeklerde su seviyesinin bölünmüş çukurlar veya düzensiz su dağılımı açısından kontrol edilmesi gibi işlemler gıda güvenliği uygulamalarına örnek olarak verilebilir.

Bilgisayarla görme ile karşılaştırıldığında yüz tanıma, gözetim ve biyometrik tanımlama uygulamaları için çok az miktarda görünür ışık kullanır. Biyometri, her bireye özgü olan iris, parmak izi veya el şekli gibi metrikleri ifade eden bir kelimedir. Gözetim altındaki bireyler tespit edilebilir ve bu önlemlerin kullanılması yoluyla kaynaklara erişimleri kısıtlanabilir [22]. Yüz tanımanın girişimsel olmayan doğası ve artan kullanım basitliği, bir kişinin kimliğini doğrulamak için tercih edilen araçlar olarak geleneksel biyometriyi hızla gölgede bırakmaktadır. Örneğin, bir uzman, kişinin parmağını bir parmak izi tarayıcısına yerleştirmesini sağlayarak bir parmak izi sistemini doğrular. Uzman olmayan kişiler, ölçüm cihazı olarak kullanıcı yüzünü inceleyerek bulgularını doğrulayabilirler.

1. Bilgisayarla görme hakkında bu kadar zor olan nedir?

Tüm görüntüler bir fotoğraf makinesi ile çekilmeli ve daha sonra herhangi bir işlemde önce incelenmek üzere bir bilgisayarlı görme programına iletilmelidir. Fotoğraf makinelerine programın gözü, işleme yazılımına ise beyin denir. Burada işlem, bir nesneden yansıyan ışığın toplanması ve ışığın yoğunluğunun daha sonra resmi kaydeden fotoğraf makinesinin dahili sensörlerine aktarılmasıyla gerçekleştirilir. Daha sonra, sensörlerin hücre yoğunluklarının her biri sayı ızgarası kullanılarak 0-255 arasında bir tam sayıya çevrildikten sonra elde ettiği resmi temsil etmek için kullanılır. Işık, bir frekans aralığında meydana gelen bir tür elektromanyetik enerji veya görsel spektrum olarak bilinir [23]. Daha önceki analog fotoğraf makinelerinde, artık dijital fotoğraf makinelerinde bulunmayan film sensörleri kullanılmıştır. Beynimiz saniyeden kısa bir sürede yarı saydamlıktaki çok küçük farklılıklar sayesinde öğeyi bir kedi olarak doğru bir şekilde tanıyabilir. İnsan gözü ve beyni, zaman içinde gelişen bir tanıma modelini kullanarak görüntüleri kavrayabilir. Ayrıca insan görsel sistemi, bağlamsal özellikleri olmayan üç boyutlu nesnelere yakalananın yanı sıra, derinlik, renk, şekil ve görünüm gibi bağlamsal özelliklere sahip üç boyutlu öğeleri üç boyutlu nesnelere biçiminde toplar. Ancak bir görüntü bir kamera tarafından yakalandığında ve bir bilgisayar görme sistemi tarafından değerlendirildiğinde, tüm bu özellikler kaybolur. Bir bilgisayar görüş sistemi, iki boyutlu bir tamsayı ızgarasından kayıp bağlamsal bilgileri başarılı bir şekilde kurtarmak için fotoğraf makinesi toplama işlemini beş kez çevirmelidir

[24]. Bilgisayarla görme karmaşıktır, çünkü seyrek verilerden görüntülerin yeniden oluşturulmasını, yeniden oluşturma işlemi sırasında kaybolan bağlamsal öğelerin yeniden oluşturulmasını ve ortaya çıkan görüntülerin yorumlanmasını gerektirir.



Şekil 3.2. Yüz özelliği noktaları [25].

2. Bir kişinin sadece görünüşüne göre tanımlanması

“Yüz tanıma” ifadesi, bir yüzün tanındığı veya tanınmadığı işlemi ifade etmektedir. Verileri toplamak, tespit etmek, işlemek ve tanımlamak için bir işlem hattı kullanılır. Resim koleksiyonu, hayvan eğitimi ve tanımlamasına yardımcı olmak için kullanılabilir. Yüz görüntüleri, ilgilenilen bölgeleri belirlemek için analiz edilir ve bu ilgi alanlarının yerleri kaydedilir. Bu adımda gölgeler ve aşırı ışık gibi istenmeyen unsurlar resimden temizlenir [26]. İşlem hattının son adımı olan tanıma, bir yüzün tanımlanıp tanımlanmadığını değerlendirmekten sorumludur.

3. Yüzün fiziksel özelliklerinin derlemesi

Bir dizi eğitim görüntüsünü gözden geçirerek, yüz tanıma sistemine yüzleri tanıması öğretilir. Koleksiyonun içeriği, varyansın geçerli bir tanımlamayı garanti edecek kadar küçük olup olmadığını değerlendirmek için yeni oluşturulan bir resimle karşılaştırılabilir [27]. İyi bir tanıma elde etmek için otomobiller, ağaçlar ve diğer yüz dışı görüntüler gibi hem olumlu hem de olumsuz örnekler de dahil olmak üzere çeşitli resimlerin koleksiyona dahil edilmesi hayati önem taşımaktadır. Ayrıca koleksiyon, kişinin bakışı yukarıya veya

aşağıya dönük olanlar ve çeşitli aydınlatma koşullarında çekilenler de dahil olmak üzere kişinin yüzünün çeşitli çekimlerini içermelidir.

4. Yüz Tanıma, bireyleri tanıma yöntemidir.

Yüz algılama, bir görüntüdeki bir kişinin yüzünü, o kişinin kimliğinin tespit edilmesini gerektirmeden algılamak için kullanılan bir yöntemdir. 2001'den önce, "Basit Özelliklerden Arttırılmış Bir Basamak Kullanarak Hızlı Nesne Tanımlama" olarak bilinen teknik, yüz tanıma alanında kayda değer bir başarı olmuştur.

Viola-Jones, önceki yüz tanıma sistemlerinde olduğu gibi piksel analizi yerine Haar benzeri özelliklere dayanan "Haar-sınıflandırıcı" tekniğini geliştirmiştir. Nature Communications algoritmayı yayınlamıştır [28]. Fotoğrafları sınıflandırmak için Haar sınıflandırıcısı olarak bilinen teknik hem olumlu hem de olumsuz örneklerden kapsamlı bir şekilde yararlanmıştır. Sınıflandırıcının başarılı olabilmesi için eğitim setindeki görüntü boyutları, nesne tanımlama için kullanılan görüntü boyutlarıyla aynı veya daha küçük olmalıdır.

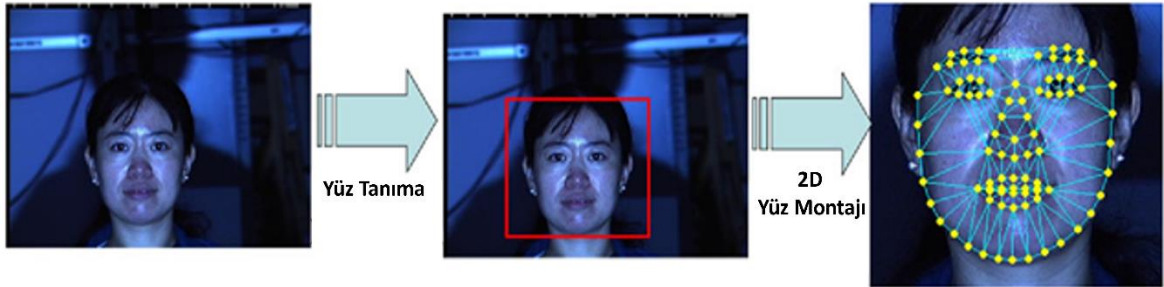
3.2 Yüz Lokalizasyonuna Giriş

Başlamak için, otomatik bir yüz özelliği çıkarıcı, yüzün alanını ve dış çevresini belirlemelidir. Kapsamlı bir yüz yerleştirme sistemi geliştirmek için aşağıdaki sorunları çözmek çok önemlidir. Yüz ve fotoğraf makinesi arasındaki etkileşim nedeniyle bazı görüntüler gözler, burun ve ağız gibi yüz özelliklerini bozabilir veya gizleyebilir. Sakal, bıyık ve gözlükler, engelleyici yüz özelliklerine örnek olarak verilebilirler. Ayrıca makyaj yüzde yapay alanlar oluşturabilir veya yüzün çirkin görünen doğal sınırlarını gizleyebilir.

Yüz özelliklerinin biçimi, iletilen duyguya yanıt olarak önemli ölçüde değişir. Bazı yönler bastırılabilirken, diğerleri yalnızca kişi uygun zihin durumunda olduğunda belirginleşir. Fotoğraflar çeşitli aydınlatma ve kamera koşullarında çekildiğinde, yüz özelliklerinin rengi önemli ölçüde dalgalanır. Yüzün bazı özellikleri bulanıksa ya da gölgeler veya parıltılarla birleşiyorsa, renk bilgisi kaybolabilir [29].

Fotoğraflardaki kişileri tanımlamak için çeşitli teknikler vardır. Bu stratejiler arasındaki büyük ölçüde örtüşme göz önüne alındığında, kesin olarak sınıflandırılmadıkları gerçeğine rağmen, birlikte yerleştirilmeleri önerilir.

Yukarıdan aşağıya yöntemi: Birkaç yüz özelliği arasında kurala dayalı bağlantılar oluşturmak mümkündür. Bir yüzde iki göz, bir burun ve bir ağız vardır ve bunların hepsinin farklı yerleşimleri vardır. Yukarıdan aşağıya yöntemler, hiyerarşik kuralların kullanımı yoluyla hedeflerine ulaşır. Araştırmacılar, en temel ilkeleri değerlendirmek için düşük çözünürlüklü görüntülerin, en spesifik standartları analiz etmek için ise yüksek çözünürlüklü fotoğrafların kullanıldığını gözlemlemişlerdir. Yang ve Huang araştırmalarını yayınladıklarında, öncelikle yüz bölgesinin bir örneğini elde ettiler [30]. İşlemin başlangıcında her komşu hücrenin yoğunluğunun, etrafındaki dört hücrenin yoğunluğuna eşit olduğu basit bir dört komşu hücre modeli kullandılar. Kenar algılamayla tanımlanan yüz özelliklerinin konfigürasyonları, kabul edilebilir olup olmadığına karar vermek için en yüksek kaliteli resimlerle karşılaştırılır. Ayrıca Kotropoulos ve Pitas, yukarıdan aşağıya yöntemine ek bir yaklaşım önermiştir. Buna göre resimlerdeki yüzler, resimlerin yatay ve dikey yoğunluk histogramları analiz edilerek tespit edilebilir. Yüz özelliklerini keşfetmek için, çıkarılan yüz bölgesinin dikey ve yatay histogramları değerlendirilmelidir [31].



Şekil 3.3. Yüz özelliklerine göre yüz lokalizasyonu [32].

Yüzle ilgili insan bilgisini kurallara dönüştürmek, çözülmesi zor bir iştir. Sınırların hatalı tespit edilmesinden kaçınmak için, kuralların çok özel veya çok genel olmamasını sağlamak gereklidir. Kural tabanlı doğrulamanın bir diğer dezavantajı, çok çeşitli yüz ifadelerini içerecek şekilde genişletilmesinin gerekliliğidir. Yaş, yön, aydınlatma ve cilt tonu, bir kişinin yüzünün nasıl sınıflandırıldığı üzerinde önemli bir etkiye sahiptir. Görüntü hazırlama yöntemleri kullanılarak, nihai görüntüdeki varyasyon miktarı azaltılabilir. Gauss türev filtreleri ve Gabor dalgacıkları ise daha stabil temsiller üretir [33]. Işık farklılıklarını hesaba katmak için görüntü, gölge düzeltme filtreleri ve histogram eşitleme kullanılarak işlenebilir. Yukarıdan aşağıya sistemlerde işlenmesi gereken veri miktarı genellikle etkin bir şekilde yönetilemeyecek kadar büyüktür. Bu stratejiyi hesaplama açısından uygulanabilir kılmak için girdi görüntüsünün boyutu minimumda tutulmalıdır. Ancak araştırma yüzün iç kısmıyla sınırlı olsa bile, analiz edilmesi gereken veri miktarı çok büyük olacaktır.

Aşağıdan yukarıya yöntemi: Bütünsel tekniklerin aksine, aşağıdan yukarıya yöntemler, bireysel yüz özelliklerini tanımlamaya ve potansiyel yüz kombinasyonlarını bir bütün yerine bir yüz olarak kategorize etmeye çalışır. Bunun için kişilerin yüzleri, duruşlarından, yönlerinden ve aydınlatmalarından bağımsız olarak belirli özellikler sergilemelidir. “Özellik tabanlı teknikler” terimi ile yüzün ayırt edici özelliklerini ortaya çıkarmak ve kullanmak için çaba sarfetmek ifade edilir. Bu ayrıntılar, bir yüzün var olup olmadığını ve nerede bulunduğunu belirlemek için gereklidir. Göz dedektörü, burun dedektörü ve ağız dedektörü kullanılmadan bu işlem başarılı olamaz. Tek bir yüz yerine form değiştirebilen bileşenlerin bir spektrumunu tanımlamak için daha yüksek çözünürlüklü fotoğraflar ve daha fazla işleme kapasitesi gerekecektir [34]. Bu tekniği kullanırken kişinin yüzüne düşen saç, gözlük ve gölgelerin neden olduğu tıkanıklıklar nedeniyle yüz hatlarını tanımlamak zor olabilir. Sınırlılıklarına rağmen özellik tabanlı yaklaşımlar, basit olmaları nedeniyle yüz tanıma ve izleme uygulamalarında sıklıkla kullanılmaktadır.

3.2.1 Şablon Eşleme

Şablon eşleştirme, aşağıda tanımlanan şablonlar kullanarak bir yüzü temsil etmektir. Bir yüzün varlığı veya yokluğu, görüntü düzenleyicide şablon ve görüntü arasındaki korelasyon değerleri hesaplanarak belirlenir. Yüzün yeri belirlenirken temel sorunlardan biri, tekniğin duruş değişikliğine dayanmasıdır [35]. Tek bir yüz görüntüsünde çok çeşitli yüz ifadelerini düzgün bir şekilde tasvir etmek için tek bir şablon yerine birden çok şablon kullanmak daha etkili bir çözümdür. Bir yüz şablonunun ayarlarını değiştirerek yüzün görünümünü değiştirmek mümkündür.

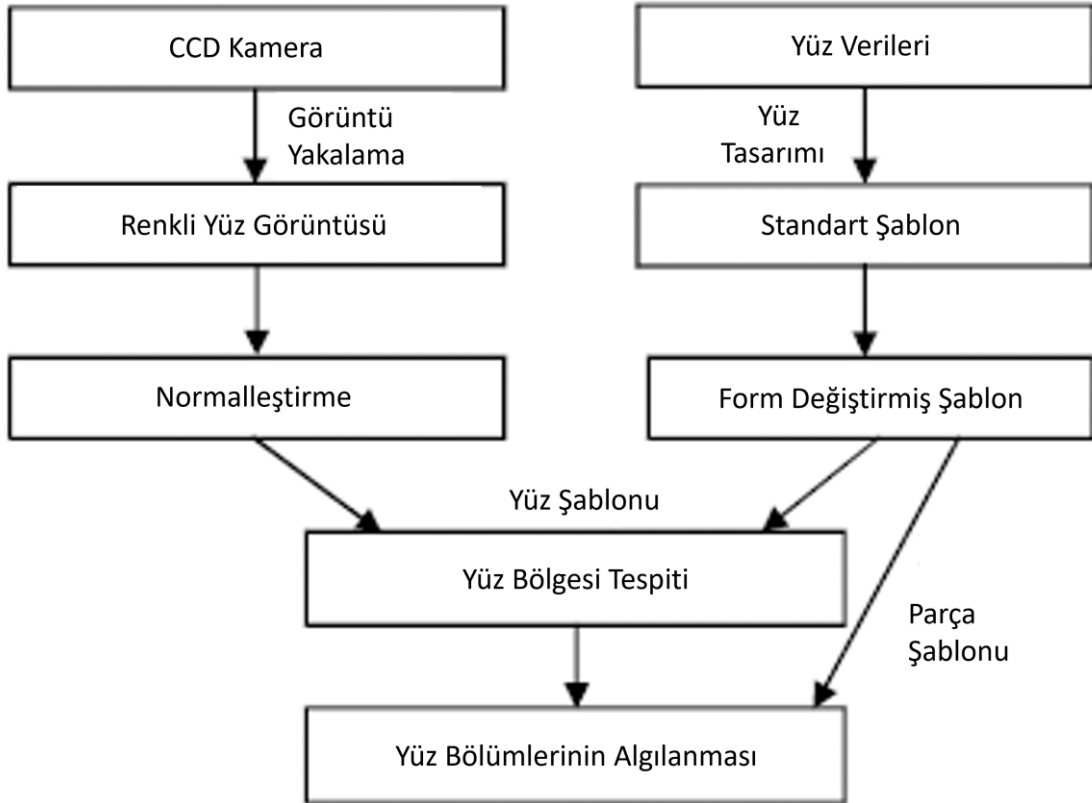
1. Önceden Geliştirilmiş Şablonların Kullanılması

Girişe yanıt olarak şekillerini değiştirmek yerine, önceden belirlenmiş şablonlar girişten bağımsız olarak ilk şekillerini korur. Söz konusu resmin üst-örneklenmesi veya alt-örneklenmesiyle, değişen boyutlarda yüzlerin varlığı tanımlanabilir. Önden görünüm yüz şablonuna dayalı bir yüz yerelleştirme yöntemi sunulmuştur [36]. Belirli sınırlamalar kullanılarak bir resmin kenarlarını tanımlamak için Sobel filtreleri kullanılabilir ve daha sonra bu kenarlar bir küme oluşturmak için bir araya getirilebilir. Gözler ve ağız gibi yüz özelliklerini tanımlamak için daha küçük şablonlar gerekirken, yüzün dış hatlarını resimden çıkarmak için tam yüz şablonu kullanılır. Yüz algılama şablonu, eşleştirme yaklaşımında çeşitli çözünürlük ve yönlerden yararlanan bir yüz algılama şablonu eşleştirme tekniği sağlamıştır. Elde edilen görüntü, orijinal yönüne göre -20 ile 20 derece arasında bir değerle

döndürülür. Daha önce elde edilen görüntüleri döndürerek çeşitli yüksek çözünürlüklü görüntüler oluşturmak mümkündür [37]. Laplace operatörleri kullanılarak görüntülerden kenarlar çıkarılır ve bir yüzün mevcut olup olmadığını belirlemek için fotoğraflar ile bir yüz şablonu arasındaki ilişki incelenir.

2. Özelleştirilebilir Tasarım Şablonlarının Kullanılması

Form değiştirebilir şablon, kendisinin uyumlu hale getirdiği hedef nesnenin bir temsilidir. Mevcut bir şablonun parametreleri değiştirilerek, söz konusu öge arama yapmak için kullanılabilir. Bu şablonu oluşturmak basittir. Değiştirilebilir şablon, görüntü ile dinamik etkileşim sağlar. Bir enerji fonksiyonunun kullanılması, giriş görüntüsünün kenarlarının, girinti ve çıkıntılarının şablon parametrelerine bağlanmasını sağlar [38]. Bu enerji fonksiyonu minimum değerine düşürülerek görüntü uyumluluğu açısından en iyi sonuçlar elde edilir. Genellikle şablon parametreleri, şablonun en uyumlu şeklini yansıtacak şekilde ayarlanır ve ardından şablonun enerji fonksiyonunu azaltmak için şablon parametreleri en dik düşme yöntemi kullanılarak güncellenir.



Şekil 3.4. Yüzün tespit edilmesinde şablon eşleştirme [39].

Bulanıklaştırma filtreleri ve morfolojik operatörler kullanılarak girdi fotoğraflarından elde edilen kenar bilgileri geliştirilebilir. Hough dönüşümü, bir görüntüde en sık hangi eliptik şekillerin görüldüğünü belirlemek için kullanılır. En olası dört elips belirlenirken ve özellikleri çıkarılırken, başlangıç noktası olarak form değiştirebilen şablonlar kullanılır. Bu özellikler arasındaki ilişki, çıkarılan elipslerin içinde bir yüzün olup olmadığını belirlemek için kullanılır.

3.2.2 Görünüme Dayalı Yöntemler

Şablon eşleştirme tekniği, insan tarafından belirlenen şablonlara dayanmaktadır. Eğitim görüntülerini daha önce kullanılan şablonlardan ve benzeri tekniklerden oluşan bir kütüphane ile karşılaştırmak yerine, şablonlar tamamen eğitim görüntülerinden öğrenilmelidir. Yüz modelleri, görünüme dayalı yaklaşımlarda diğer tekniklerin yanı sıra makine öğrenimi ve istatistiksel yöntemler kullanılarak oluşturulur. Farklılığı oluştururken, yüzler ve yüz dışı nesnelere arasında ayırım yapmak için dağılımlar ve ayırt edici işlevler biçiminde eğitilmiş modeller kullanılır. Görünüme dayalı yüz lokalizasyonu düşünüldüğünde, istatistiksel bir teknik düşünülebilir. Bir öznelik vektöründen rastgele bir değişken oluşturulur ve bu rastgele değişken daha sonra öznelik vektörüne dahil edilen özelliklere dayalı olarak 'yüz' veya 'yüz olmayan' olarak sınıflandırılır [40]. Olma ihtimalinin maksimum değerini veya Bayes sınıflandırma yaklaşımlarını etkili kılmak için, özellik vektörü mümkün olduğunca kompakt olmalı ve küçük boyuta sahip olmalıdır. Daha düşük boyutlu bir özellik vektörünü daha yüksek boyutlu bir özellik vektörüne yansıtarak, 'yüz' ve 'yüz olmayan' sınıfları ayırt etmek mümkündür. Daha sonra lineer diskriminant fonksiyonları kullanılarak 'yüz' ve 'yüz olmayan' sınıflar ayırt edilir. Bununla birlikte, doğrusal olmayan sınıf sınırları oluşturmak için de sinir ağları kullanılabilir [41].

- ***Bir dizi göstergede doğruluk (Eigen)***

Daha çok temel bileşen analizi olarak adlandırılan Karhunen-Loeve açılımı, sinyal ve fonksiyon analizinde geniş uygulama yelpazesine sahip bir sinyal ve fonksiyon analizi yaklaşımıdır. Araştırmacılar, yüz görüntü analizinde bu teknolojinin kullanımı üzerinde yoğun çalışmalar yapmıştır. 40 görüntülü özgün bir yüz setinin, %3 temsil hatasıyla bir resmi düzgün bir şekilde temsil edebileceğini göstermişlerdir. Turk ve Pentland, bir yüz resmindeki özyüz alt uzay koordinatlarının çığır açan araştırmalarda yüz tanıma için kritik olduğunu ilk kez göstermişlerdir. Ayrıca özyüzlerin oluşturulacağı en iyi resimlerin seçilmesi için bir yöntem açıklamışlardır [42, 43]. Açıklanan bu yüz sınıfı projeksiyon

yaklaşımı, eğitim resimleri için ortalama kare yeniden oluşturma hatalarını en aza indirir. Bu alt uzay için temel vektörlere özyüzler denir ve yüz analizinde ana bileşenler olarak da adlandırılırlar.

- ***Sinirsel Ağlar***

Örüntü tanıma için sinir ağlarının kullanımı popülerlik kazanmaktadır. Yüz algılama, sinir ağlarının yaygın bir uygulamasıdır ve iki çözüm sınıfına sahip bir örüntü tanıma sorunu olarak düşünülebilir. Geniş bir yüz görüntüsü koleksiyonu üzerinde çalışılmış olan nöron ağları, yüz sınıfının ayrıntılı sınıf yoğunluğunu düzgün bir şekilde yakalama yeteneğine sahiptir. Nöronal ağların bazı dezavantajları vardır. Bunlardan biri optimal topolojiyi seçememeleridir. Bu nedenle gizli katmanların ideal sayısını ve boyutunu seçmenin zordur. Eğitim performansını artırmak için, yüz olmayan fotoğrafların (söz konusu yüz sınıfı ile aynı yüz sınıfında olması gereken) seçimini incelemek önemlidir. Propp ve Sanal, yüz tanıma için sinir ağlarının kullanımına öncülük etmiştir. Verilerin sınıflandırılması, bir sinir ağı kullanılarak geliştirilen 1024 giriş birimine sahip dört katmanlı bir ağ kullanılarak gerçekleştirilmiştir. Fotoğraflardaki yüzleri algılamak için iki aşamalı bir sinir ağı yaklaşımı gereklidir. Sinir ağı, mevcut teknikleriyle sonuçlanan filtreleme teknolojilerini geliştirmek için iki aşamadan geçmiştir [44]. Öncelikle çeşitli çözünürlüklerde yüzler aranmaktadır. Birinci sinir ağı, ikinci sinir ağının bilgisini kullanarak resimdeki her noktayı filtrelemek için kullanılır. Giriş resminin alt örnekleme yapılar ve arama penceresinden daha büyük yüzleri bulmak amacıyla bu çözünürlükte başka bir arama yapmak için aynı sınıflandırıcı kullanılır. Tekrarlanan ve yanlış yüz algılamalarını azaltmak için bir filtre kullanılabilir. Her konumdaki veri tabanından elde edilen yüzlerin sayısını (olumlu sonuçlar) saymak için kayan bir pencere kullanılır. Önceden tanımlanmış bir eşik üzerindeki pozitif bulgular, bir bölgeyi “yüze ait” olarak sınıflandırmak için kullanılırken, bu kriteri karşılamayan bölgeler daha fazla araştırmanın dışında bırakılır [45].

- ***Yapısal Sistem Modelleme ve Analizi***

Kalabalık ortamlarda destek vektör makineleri (SVM'ler), yüz tanımda yüksek doğruluğa ulaşma yeteneğine sahiptir. Durum için en uygun sınıf sınırını seçmek amacıyla yapısal risk azaltma kullanarak eğitim verilerini incelemektedir. Destek vektörleri (bir seçimin yapılabileceği veriler), karar sınırlarını oluşturmak için karar sınırları ile kullanılmıştır. Bu durum, tasarımın koyu renkli destek vektörlerinin kullanımıyla gösterilir. SVM'ler, duygusal bir karar sınırı seçmek için büyük miktarda eğitim verisine ihtiyaç duyar ve

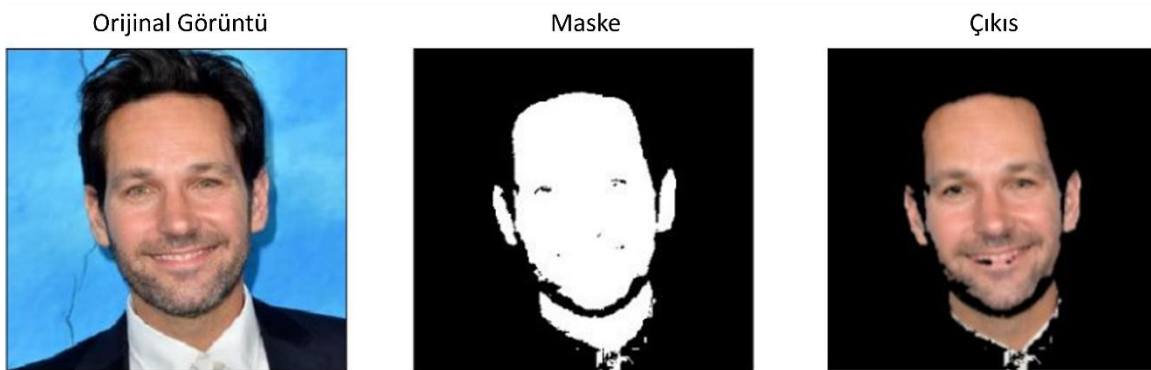
kendimizi tek bir poz (ön) algılamayla sınırlasak bile, bilgi işlem maliyeti makul olmayan bir şekilde yüksek kalır [46].

3.3 FAFE'nin Yüz Lokalizasyon Aşaması

FAFE yöntemi yüzün konumunu, boyutunu ve yönünü tanımlamakla başlar. Hedef fotoğraflarda (veya çerçevelerde) sadece bir kişinin omzunun görünmesi planlanmıştır. Yüz, görüntüdeki en büyük elipsi bularak bu kısıtlamayı kullanır ve cilt segmentli ikili resimden tanımlanır [47]. Sonraki bölümlerde, cilt segmentasyonu ve elips uydurma daha ayrıntılı olarak incelenecektir.

3.3.1 Cilt Segmentasyonu

Yüz belirlemede ayırt edici bir özellik olarak ten renginin kullanılmasının etkinliği araştırılmıştır. Her kişinin cilt tonu benzersiz olsa da birden fazla çalışma, birincil değişkenin renkten ziyade yoğunluk olduğunu göstermiştir. Renk segmentasyonu, RGB, normalize RGB, YCrCb ve YIQ dahil olmak üzere bir dizi farklı renk uzayında gerçekleştirilebilir. Bu, yalnızca renk bilgilerini içeren ve yoğunlukla ilgili bilgileri atlayan en temel renk uzayıdır [48]. "Normalize RGB" terimi bunu ifade etmektedir. Bu renk uzayı, bu deney için seçilmiştir çünkü daha önceki deneylerde yüksek doğrulukta segmentasyon sağladığı kanıtlanmıştır. Eğitim setinden alınan deri alanı örneklerinin histogramları toplanır ve (r,g) uzayına yansıtıldıktan sonra daha detaylı olarak görüntülenir. Veri tabanındaki cilt alanının dağılımı dikkate alınır. (r,g) düzleminde cilt tonlarının dar dağılımı nedeniyle, görüntü cilt tonlarının çok sınırlı bir dağılıma sahip olduğunu göstermektedir. Bu grafikteki (r,g) değerlerinden görüldüğü gibi, basit bir seviye dilimleme işlemi kesin veri segmentasyonu ile sonuçlanabilir.



Şekil 3.5. Cilt bölgesinin segmentasyonu [51].

Bir görüntünün karanlık alanlarında yetersiz renk bilgisi olduğunda yanlış segmentasyon meydana gelir ve bu durum da hatalı segmentasyona neden olur [49, 50]. Bu sorun, daha önce bahsedildiği gibi giriş görüntülerinde renk eşiğine ek olarak gri seviye eşiği kullanılarak önlenir. Bilim insanları, cilt bölgesini çıkarmak için gri seviye kriterleri 0,4r0,5 0,2g0,4 ve I>50'yi kullanmışlardır.

3.3.2 Elips Uydurma

Bu çalışma için girdi fotoğrafları çekildiğinde, kişinin yüzünün 45 derecelik bir yüz dönüşü ile kameraya dönük olduğu varsayılmaktadır. Yüzün açısı, pozisyonu ve konumu, ciltten çıkarılan ikili resimde mümkün olan en büyük elips aranarak belirlenebilir [52-54]. Dairesel alanın bu durumda makul bir yüz olduğuna inanılmaktadır. İkili görüntü, yüz organları, sakal ve gözlükle ilgili küçük hatalı tespitler ve boşluklar içerebilir. Yüzün yanlış tespitlerini ve eksik kısımlarını düzeltmek için, ciltten çıkarılan ikili resimde ekstraksiyon sonrası morfolojik değişiklikler yapmak gereklidir. Açılış yaklaşımı, $n \times n$ yapılandırılmış bileşenlerini kullanır. Bu senaryoda daha küçük görüntünün kenar boyutunun $n=1/32$ 'si dikkate alınır. Yüz organları ve saçla ilgili sorunların yanı sıra, ikili gösterimin yüzün diğer çeşitli bölgeleriyle ilgili sorunları vardır. Özgün renk dağılımları nedeniyle, çıkarılan ikili resimde delikler olarak görünürler. Elips uydurmaya uygun bir alan sağlamak için ikili resmin kapatılması gereklidir. Bu yöntem, $3n \times 3n$ yapısal elemanı kullanır. Yüz belirleme yöntemindeki bir sonraki aşama, bir elips ile en yakından eşleşen alanı belirlemektir. Zamandan tasarruf etmek için çıkarılan cilt bölgesinin içine uyan en büyük dikdörtgen aranır. Ardından, bu alanın genişletilmiş bir kopyası, orijinaline en yakın olan elips için aranır [55]. Aşındırma tekniği, renk, şekil ve simetriye dayalı olarak fotoğraflardan yüz özelliklerinin çıkarılmasına yönelik bir yöntemi tanımlayan değeri yinelemeli olarak artırarak en büyük 31×21 dikdörtgene sahip ikili görüntüyü bulmak için kullanılır. Yüzü bir elips olarak temsil etme yaklaşımları, dikey ve yatay yönlerdeki farklılıklara dayanmakta olup, hangi elipsin yüzü en iyi şekilde temsil ettiği belirlenmiştir [56]. Bu tezde kullanılan yüz elips uydurma tekniği, açıklanan bu yöntemle temelde aynıdır. Yüz elips parametrelerinin aranacağı alan seçildikten sonra, yüz elips parametrelerini çıkarmak için bu yaklaşım kullanılır. Bir dizi ayırt edici özellik için en iyi eşleşmenin nasıl bulunacağını gösteren bir örnek olarak elips şeklinde bir kenarlık kullanılacaktır.

3.4 Yüz Özellik Çekimi

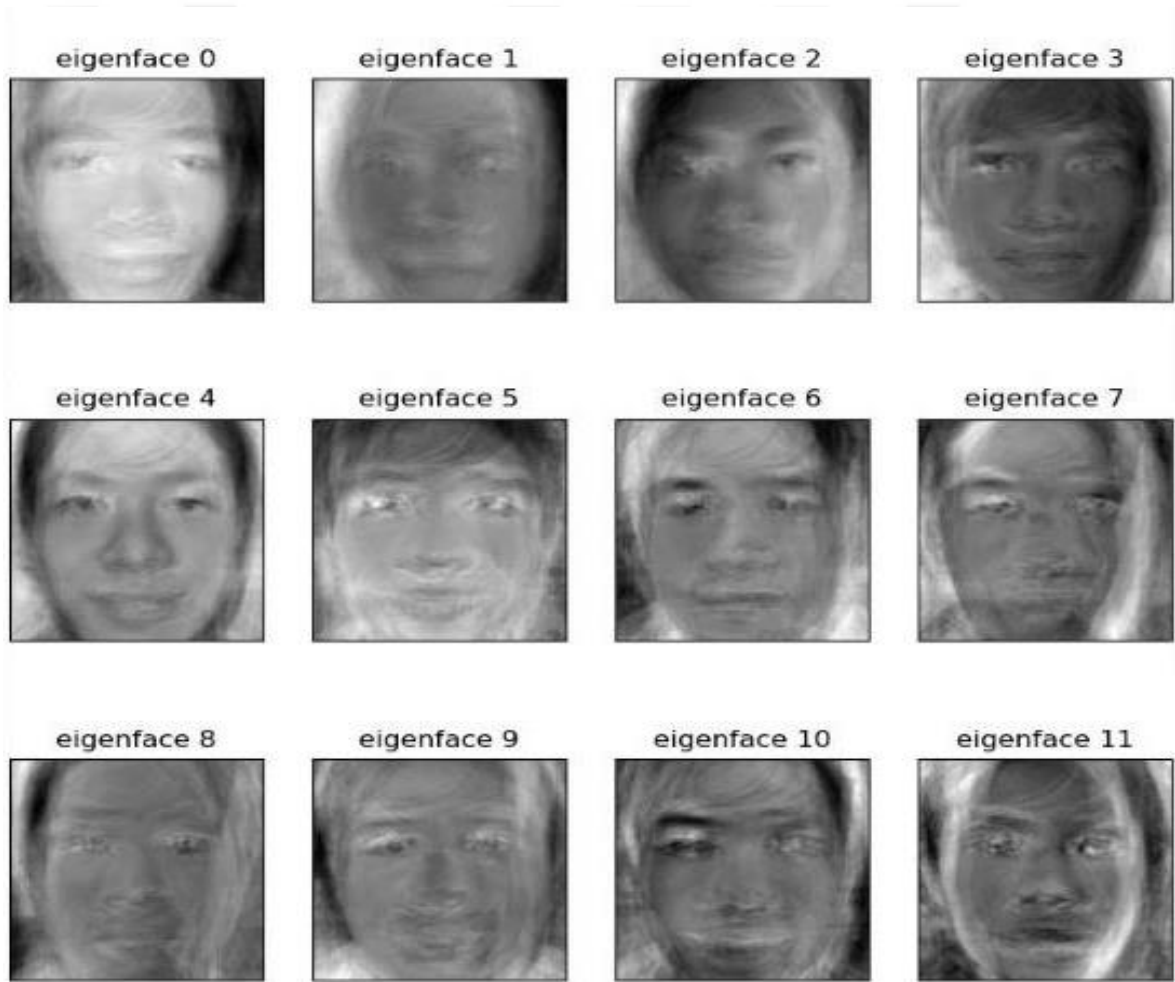
Biyometri, yirminci yüzyılın başından beri kişileri tanımlamak için kullanılmıştır. Yüz tanıma, mevcut biyometrik tanımlama tekniklerinden sadece biridir. Yüz tanıma, iki kişinin yüzlerini karşılaştırmanın otomatik bir yoludur. Yüz tanımanın düzgün çalışması büyük ölçüde ayıklamaya bağlıdır. Aynı yüze ait görüntüler aynı kişiden elde edilse bile mükemmel bir eşleştirme yapılamaz. Bireyler arasındaki etkili iletişim ve katılım, altta yatan insan davranışının anlaşılmasını gerektirir. Yüz özelliklerini sınıflandırmak için bir sistem tasarlamak önerilmiştir [57]. Önerilen bu sistemde çeşitli eğrileri karşılaştırmak için bir bağımsız ölçüm vektörünün geliştirilmesi tavsiye edilmiştir. Veri tabanının vektör kullanımı, gerçek dünya senaryolarında yüz tanıma algoritmalarının gösterilebileceği noktaya kadar ilerlemiştir. Devam eden algoritma araştırmalarının yanı sıra, geniş bir yüz görüntüleri veri tabanının mevcudiyeti ve yüz tanıma algoritmalarının etkinliğini değerlendiren bir sistemin varlığı çok önemlidir. Ayrıca bu alanın hızlı büyümesine katkıda bulunan başka etkenler de mevcuttur. Görüntünün hissedilmesini ve biçimini tanımlayan yerel ikili desenleri (LBP) kullanarak dijital görüntülerden yüz tanıma özelliklerini çıkarmak için dört farklı strateji kullandık. Bu teknikler arasında Harris ve gradyan histogramları (HOG) aracılığıyla yüz algılama da yer almıştır. Gradyan yönünün kopyaları, görüntünün gradyan yönünden türetilen hızlandırılmış sağlam özellikler (SURF) kullanılarak görüntünün izole bölgelerinde tahmin edilir. SURF tanımlayıcıları öncelikle nesnelere, insanları veya yüzleri bulmaktan ve çevrelerinden ilgi alanlarını çıkarmaktan sorumludur [58].

3.4.1 Yüz Özelliği Çıkarma Teknikleri

Yüz analizinde sıklıkla göz ve kaş çevresindeki çizgiler, dudakların şekli gibi yüz özellikleri kullanılır. Yüz tanıma, dudak takibi ve göz takibi, yüz sınırlarındaki değişiklikleri doğrudan izlemeyi amaçlayan konulardan sadece birkaçıdır. Tıp literatüründe bir yüz özelliği, yüz sınırları ile birbirinin yerine kullanılmaktadır [59, 60]. Aşağıda verilenler, yüz özelliklerini tanımlamak için en sık kullanılan yöntemlerdir:

- Şablonların karşılaştırılması ve kıyaslamaların yapılması.
- Dalgacıkların tekniklere uygulanması.
- Mekânsal segmentasyona dayalı teknikler.

Şablon eşleştirme, bu üç yöntem arasında en sık kullanılan yaklaşım olsa da aynı zamanda en zordur. Daha önce açıklandığı gibi, şablon eşleştirme, mevcut şablon şekli ile hedef alan arasındaki farklılığı nicelleştiren bir maliyet fonksiyonunu en aza indirmek için bir örnek şablonun formunun değiştirilmesiyle gerçekleştirilir. Herhangi bir şey ararken akılda belirli bir şekle sahip olmak gerekir ancak yüz sınırları insanlar arasında geniş bir çeşitlilik göstermektedir. Bir kişinin ruh hali, cinsiyeti, sakalı ve gözlüğü, yüz görüntüsünün sınırlarının nasıl görüldüğü üzerinde bir etkiye sahiptir. Bu engeller göz önüne alındığında, şablonlar için ortak bir format benimsemek zorlu bir girişimdir [61]. Üst ve alt dudak kenarlarını tasvir etmek için, dört satıra indirgenmiş bir MPEG-4 yüz modelinin kullanılması seçilmiştir. Diyagramda gözler dört nokta ile temsil edilmektedir. Yüzün öznelikleri, çok durumlu bir şablon eşleştirme yaklaşımı kullanılarak çıkarılmıştır.



Şekil 3.6. PCA uygulanmasından sonra yüzün özelliği [64].

Bu araştırmada üç farklı ağız türünü temsil etmek için açık ağız, kapalı ağız ve sıkıca kapatılmış ağız olmak üzere üç farklı ağız şablonu kullanılmıştır. Şablon, ilk karede görüntülenmiş ve daha sonra şablonun özellikleri gelecekteki karelerde ayarlanmıştır.

Mevcut çerçevedeki aktif ağız şablonunu belirlemek için renk bilgisi kullanılmıştır. Ayrıca gözleri tasarlarken açık ve kapalı gözleri tasvir etmek için iki durumlu bir şablon kullanılmıştır. Mevcut şablon modelini seçmek için bir iris dedektörü kullanılmıştır. Picard, kızılötesi tabanlı göz tanıma ve şablon eşleştirme kombinasyonunu kullanarak üst yüz özelliklerini çıkarmış ve izlemiştir [62]. Şablon parametrelerini kurtarmak için PCA kullanılmış ve saniyede 30 kare özellik izleme hızı bildirilmiştir. Bu sistemin özellik çıkarma yaklaşımı, önceki tekniğin bir uzantısıdır. Şablon eşleştirme işlemlerinin çoğunu başlatmak ve şablon parametre veri tabanını başlatmak için genellikle bir görüntü dizisinin ilk karesi kullanılır. Bununla birlikte, önerilen yaklaşımın göz kürelerini algılamak için bir kızılötesi kamera yapısına ihtiyaç duyduğunu belirtmekte fayda vardır. Gözleri tespit ettikten sonra, en belirgin yüz özellikleri için şablonlar bulunur ve bu şablonlar kullanılarak parametreler türetilir. Üst yüz şablonunu hesaplamak için yenilikçi bir yöntemi sunulmuştur [63]. Önerilen bu yöntemle üst yüz şablonunu bulmak için gözlere güvenmek yerine, COE kullanılır. Bu bilgilerle, dudaklar ve kaşlar gibi alt yüz özelliklerinin yanı sıra gözlerin ve kaşların konumu ve boyutu/yönü gibi üst yüz özelliklerinin türetilmesi mümkündür. Bu bölgelere özellik konumlarını çıkarmak için UFR ve LFR olmak üzere iki farklı şablon kullanılır. Şablon parametreleri COE ve COM dosyalarına eklendikten sonra, bunları hesaplamak için PCA tekniği kullanılır.

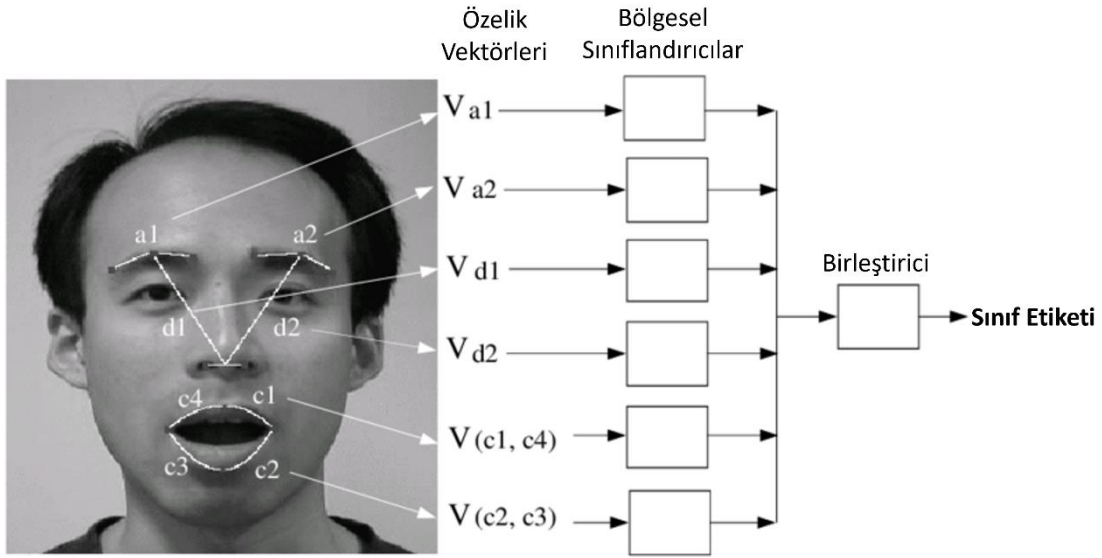
3.4.2 Hedeflenen Yüz Görevleri

Bu tezin amacı, resim segmentasyonu ve alan kategorizasyonu için yeni fikirler oluşturmaktır. Renk, doku ve biçim gibi tanımlayıcılar kullanılarak alanlar oluşturulabilir ve adli tıp da dahil olmak üzere bir dizi yüz analizi uygulaması için yeni derin öğrenme mimarileri geliştirilebilir. Yüz tanıma ve kişiden bağımsız yüz ifadesi kategorizasyonu gibi problemler, özellikle kısıtlamasız koşullarda daha zor olduğu için bu konuya odaklanılmıştır. Artan sayıda kullanıcı, yüz tanıma teknolojilerini kullanarak fotoğraf ve videolardaki kişileri tanımlamak için yapay zeka kullanmaktadır. Görsel özellikleri toplayıp bunları verilere dönüştürdükten sonra, düşük boyutlu bir özellik uzayı(piksel) hesaplanabilir [65, 66]. Yüzü tanımlamak için, en yakın eşleşme olan yüzü seçmek amacıyla elde edilen alan referans veri tabanı ile karşılaştırılır. Geleneksel yüz tanıma yöntemlerini kullanarak “düğüm noktaları” olarak adlandırılan yüzün 80’den fazla farklı bölümünü incelemek mümkündür. Bu özellikler arasında gözler arasındaki mesafe, burnun uzunluğu gibi geometrik oranlar yer almaktadır. Yüz tanıma yazılımı, benzersiz görsel ve geometrik özelliklere göre farklı bireyleri ayırt edebildiğinde, daha iyi performans göstermektedir. En yeni buluşsal

teknolojiler, daha kesin bulgular sağlamak için cildin yapısını, şeklini ve rengini değerlendirmeye öncelik vermektedir. ABD Savunma Bakanlığı 1990'ların başında bireyleri tanımlayabilecek bir teknoloji aramaya başladığında, yüz tanıma teknolojisi halk arasında popülerlik kazanmıştır. Yüz tanıma ilk olarak 2001 yılının başlarında Florida, Tampa'daki 25. Super Bowl'da halka gösterilmiştir. Daha sonra polis güçleri tarafından kalabalığın arasına gizlenmiş suçluları ve teröristleri aramak için kullanılmıştır. Son olarak, sistemler ülke çapında ek yüksek riskli bölgelere yerleştirilmiştir. Uluslararası Biyometri Derneği'ne göre, biyometrik tanımlama açısından yüz tanıma artık ilk üç sırada yer bulan ve en hızlı büyüyen biyometrik teknolojilerden biridir. Hükümetler ve yetkililer, özellikle video gözetim sistemleriyle eşleştirildiğinde, gözetim ve güvenlik nedenleriyle bu teknolojinin kullanımında bir artış görmektedirler [67]. Yüz tanıma teknolojisi Fransa, Birleşik Krallık ve Amerika Birleşik Devletleri'nde inanılmaz bir şekilde kullanılmaya başlanmıştır. Örneğin, Nice şehri Karnaval şenliklerinin bir parçası olarak yüz tanıma teknolojisini denerken, ABD hükümeti bunu havaalanlarında vize süresi dolmuş kişileri tespit etmek için kullanmaktadır. Sağlık, pazarlama ve turizm sektörleri, yüz tanıma teknolojisini kullanan alanlardan sadece birkaçıdır. Bunlara ek olarak, genel nüfusa sunulan bir dizi hizmet ve ürün mevcuttur. Apple şirketi 2017 yılında müşterilerin ön kameraya basit bir bakışla telefonlarının kilidini açmasını sağlayan Face ID teknolojisine sahip iPhone X model cep telefonunu piyasaya sürmüştür. Bir 3D tarayıcı, kimliğini doğrulamak için kullanıcının 30.000'den fazla özgün özelliğini değerlendirmektedir. Apple Pay işlemleri, iPhone 7 model cep telefonunun Face ID işlevi kullanılarak da doğrulanabilir. Facebook şirketi, sosyal ağda paylaşılan görüntülerdeki kişilerin yüzde 97'sini başarıyla tanımlayabilen Deep-Face teknolojisini geliştirirken, Google şirketi ise bir yüz tanıma sistemi geliştirmektedir. Bir Facebook kullanıcısı bir resimde 'etiketlendiğinde', sistem kişinin yüz özelliklerini kaydeder ve saklar. Öte yandan, yüz ifadesi tanıma, bir kişinin duygularının en belirgin temsilini belirlemeye odaklanır. Son yirmi yılda, bilgisayarla görme ve örüntü tanıma araştırmacıları, yüz ifadesi analizi ve tanımaya büyük ilgi göstermişlerdir. Bilişsel bilim insanlarının elde ettiği sonuçlar, bilgisayarla görme ve bilimsel araştırmacıları, diğer şeylerin yanı sıra, hareketli fotoğraflarda veya hareketsiz fotoğraflarda oyuncuların yüz duygularını ayırt edebilecek sistemler tasarlamaya teşvik etmiştir [68].

Yüz ifadelerini sınıflandırmak için bu sistemlerin çoğu, daha kesin sınıflandırmalardan ziyade sevinç, üzüntü veya gazap gibi geniş duygusal kategorilere dayanmaktadır. Yüz yüze iletişim, çok çeşitli sözsüz duygu ve niyetleri iletmek için yüz ifadeleriyle iletilen sözsüz

sinyallere dayanır. Bir kişinin faaliyetleri, amaçları, kişiliği ve psikopatolojisi, yüz ifadelerinden çıkarılabilir. Yüz yüze insan teması sırasında sözsüz iletişim sinyalleri, yüz ifadeleri ve diğer bedensel hareketlerle iletilir. Bu göstergeler, konuşmacının mesajını yorumlamada dinleyiciye fayda sağlayabilir. Sağlıklı ilişkiler sürdürülebilir yeteneği, etkili iletişim kapasitesine dayanmaktadır. Bireylerin yüz ifadeleri, dikkatleri, kişilikleri, amaçları ve psikolojik sağlıkları dahil olmak üzere onlar hakkında çok sayıda bilgi aktarabilir.



Şekil 3.7. Yüz simgelerine göre yüz özelliklerinin sınıflandırılması [69].

Çevre ve yakın çevremizdeki insanlarla olan tüm etkileşimlerimiz, aldığımız etkileşimli sinyallerden etkilenir. Bu atama, çeşitli müşterilere hitap etmek için çeşitli şekillerde kullanılabilir. Otizm spektrum bozukluğu olan çocuklar, sosyal çevrelerini daha iyi anlamaktan fayda görebilirler. Ayrıca, konu e-öğrenme materyallerini ve kamu hizmetlerini değerlendirmeye geldiğinde bu teknoloji daha etkili, doğru ve gerçek zamanlı olacaktır [70]. İnsan Destek Robotu sektörü, etkileşimlerini kullanım bağlamlarının duygusal koşullarına uyarlayabilen robotlar üretecektir. Bilgisayarla görme söz konusu olduğunda, en zorlu görevlerden biri, görünmeyen kişiler için gerçek zamanlı duygu tanımadır. Bu en zor çabalardan biridir. Literatürde yüz tanıma dört zorluk derecesine ayrılmaktadır. Bunların her biri, ifade edilen duygunun (Spontane ve Pozlanmış) ve onu ifade eden kişinin (Spontane ve Pozlanmış) (kendisiyle aynı kişi veya farklı bir kişi) ayırt edilmesinde farklı bir zorluk derecesi ile ilişkilidir. Sonuç olarak, her insan belirli bir duyguyu kendine özgü bir şekilde gösterdiğinden, spontane duyguları tanımlamak zordur. Sınıflar arası örnek müdahaleler, iki farklı duygu sınıfının aynı genel görünüme sahip iki farklı resimle temsil edildiği sık görülen bir olgudur. Burada kişinin spontane duygularını not etmesi çok önemlidir. Kişiler olup

bitenlerden tamamen habersiz olmalıdırlar ancak kişilerin görüntülerinin ve filmlerinin edinilmesini ve kullanılmasını kabul etmeleri gerektiği göz önüne alındığında bu durum pek mümkün görünmemektedir. Sonuç olarak, sağlanan etiketlerin gözlemlenen olaylardan herhangi biriyle eşleşip eşleşmediğini belirlemek yerine, kendiliğinden duygusal deneyimleri tanımlamaya ve sınıflandırmaya yönelik çok az araştırma yapılmıştır. Uzmanlara göre, gerçek insanların yüz ifadelerinin fotoğraflarını içeren veri tabanlarının çoğu, internette gezinmek ve belirli duygusal terimlerle ilişkili resimleri depolamak suretiyle oluşturulmaktadır. Araştırmacılar, FER ve AFEW gibi veri tabanlarını sıklıkla kullanırlar. Katılımcılara dersler boyunca yüz duygularını değiştirmeyi öğretmek yerine, poz ifadeleri, onlara pozlanan ifadeler boyunca sabit bir ifadeyi sürdürmeleri talimatı verilerek geliştirilir. Genellikle sanatçı oldukları için yüz ifadeleri yüz tanıma yazılımı kullanılarak değerlendirilebilir [71]. Yüz ifadelerini başarılı bir şekilde tanımlamanın ikinci engeli, bireye bağlı olmayan bir şekilde verilen eğitim oturumu sırasında hazır bulunmayan bireylerin yaptığı gözlemlerin anlaşılabilmesidir. Bu tez, FER olgusunu incelemektedir.

3.4.3 Bütünsel Açıklama

Bütünsel olarak yapılan yüz tanıma, bir kişinin yüzüyle ilgili mevcut tüm bilgileri kullanır ve bu nedenle doğruluğu daha yüksek bilgi sağlar. Yüzler hakkındaki genel bilgilerin çoğu, yüz görüntülerinin piksel bilgilerinden kolayca çıkarılabilen sınırlı sayıda niteliklerden elde edilir. Bireyler, bu ayırt edici özelliklerin az bir kısmı ile tanımlanabilir. Bütünsel yaklaşımlarda yüzün tamamı bir girdi aracı olarak kullanılır. Her yüz görüntüsünü bir matris kullanarak temsil etmek yerine, her bir yüz görüntüsünü temsil etmek için tüm pikselleri tek bir vektör olarak listeleyen yüksek boyutlu bir sinyal vektörü kullanılır. Yüz tanıma literatüründe, dalgacık alt bantları, Gabor filtreleri, optimal matris çarpanlarına ayırma, yönlendirilebilir piramit dönüşümü, bağımsız bileşen analizi (ICA) yöntemi, Zernike momentleri, küresel Gabor-Zernike özellik tanımlayıcıları ve temel bileşen analizi (PCA) tabanlı Fisherface yöntemi dahil olmak üzere çok sayıda yöntem önerilmiştir. Akademik araştırmalarda, bütünsel yaklaşımların özellikler arasındaki benzerlikleri çıkarmada ve hesaplamada etkili olduğu gösterilmiştir. Bir yüz görüntüsünden genel bilgileri çıkaran 2 seviyeli dalgacık örneğinde gösterildiği gibi, bazı faydaları olmasına rağmen bazı dezavantajları da vardır.

3.4.4 Bölgesel Açıklama

Yerel (Bölgesel) İkili Kalıplar (LBP), çeşitli konumlarda görülebilen bir kalıp türüdür. Bölgesel iki bitlik (binary) modeller geliştirilmiş olup bu araştırmada ilk bölgesel tanımlayıcı olarak değinilmiştir. LBP, doku sınıflandırmasının yanı sıra, yüz görüntülerinin özellik yamalarını hesaplama konusundaki büyük potansiyeli sayesinde, özellikle yüz tanıma olmak üzere farklı pek çok sınıflandırma işlemi için kullanılmıştır. Her görüntü pikseli için bir etiket oluşturmak amacıyla LBP tanımlayıcısı, değerini üç komşu 3×3 alt blok komşusunun değerleriyle karşılaştırır, burada kesinlikle negatif değerler 0 ve pozitif pozitif değerler 1 olarak kodlanır ve LBP tanımlayıcısı uygulanır. Sol alt köşeden başlayarak, sekiz binary değer, piksel etiketi tarafından ondalık biçimde temsil edilen sekiz bitlik kodu yapmak için saat yönünde birleştirilir.

LBP tanımlayıcısı çeşitli uygulamalara uygulandığında, LBP'ye benzer yaklaşımların önemli ölçüde artmasına ve tamamen yeni yaklaşımların meydana gelmesine yol açmıştır. Genel LBP benzeri yaklaşımlar önemli ölçüde popülerlik kazanırken, orijinal LBP tekniğinin ayırt edici gücünü, sağlamlığını ve kırılmasını elde etmek ve geliştirmek için literatürde çok sayıda LBP modifikasyonu önerilmiştir. LTP'nin yüz tanıma uygulamalarında Tan ve Triggs'in ilk operatöründen (LBP) daha etkili olduğu kanıtlanmıştır. Bir eşik kullanıldığında, merkez piksel ve komşu pikseller arasındaki fark, üç sayı (1, 0 veya -1) kullanılarak temsil edilmektedir.

3.5 Sınıflandırma

Denetimli sınıflandırmanın temel amacı, nesnelerin onları farklılaştıran niteliksel veya niceliksel özelliklere dayalı olarak sınıflandırılmasına imkan veren kurallar geliştirmektir. Nicel Y değişkenleri genellikle regresyon işlemlerinde kullanılır. Aşağıda en yaygın kullanılan yüz analizi algoritmaları incelenecektir.

3.5.1 En yakın komşu

“Sonraki komşu araması” terimi, bu araştırmada kesin ve tahmini sonuçlar sağlayan sonraki komşuları keşfetmeye yönelik metodolojilere atıfta bulunmak için yorumlanmalıdır. Herhangi bir yanlış anlaşılmayı veya yanlış yorumlamayı önlemek ve çeşitli arama türleri arasında ayırım yapmak için ‘kesin’ ve ‘tahmini’ sözcükleri kullanılacaktır. Bunun dışında,

bu çalışma boyunca “doğrusal arama” ve “kaba kuvvet araması” terimleri birbirinin yerine kullanılmaktadır.

1. Yakın veya yan komşunun belirlenmesi

Düşük boyutlu görselleştirmeler indirgenmiş bir boyuta sahip olduğundan, belirli bir nokta koleksiyonunda o noktaya en yakın k noktasını tespit etmek amacıyla belirli bir nokta olan sonraki komşu k 'nın doğruluğu daha kolay bir şekilde gösterilebilir.

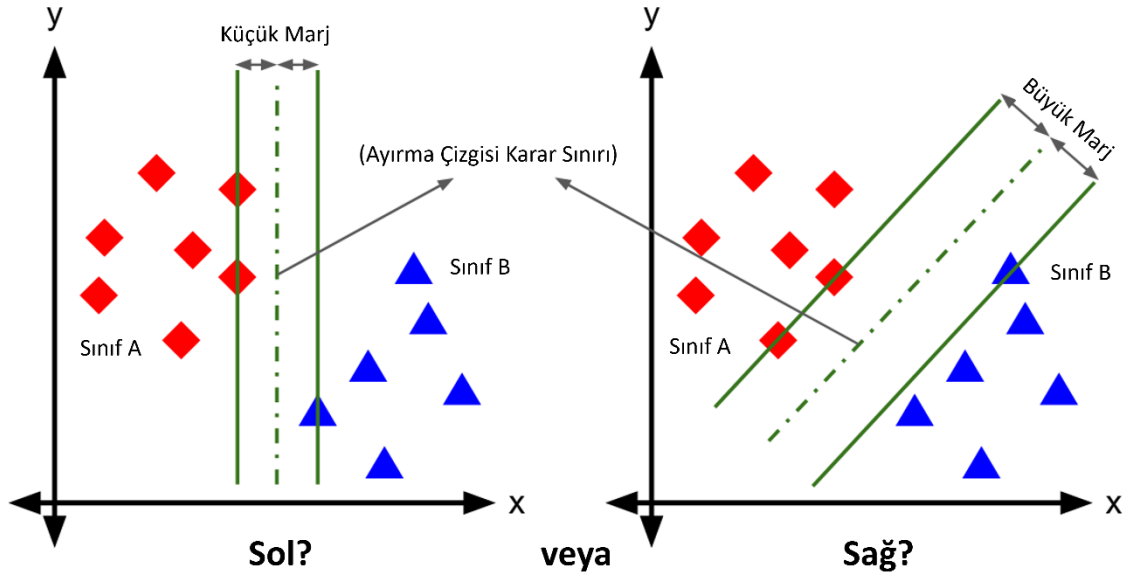
Sıralı ve uzaysal bölümlenme algoritmalarını birleştirerek, bir aramada kesin en yakın komşuları elde etmek mümkündür. Sorgu noktası ve tüm veri noktaları, aralarında en kısa mesafeye sahip olanları belirlemek için sırayla incelenir. Birçok arama uzayında, arama uzayını bir k -d ağacı kullanarak bölmek, çok daha hızlı aramalarla sonuçlanabilecek yaygın bir uzay bölümlenme tekniğidir. Düşük boyutlu bir ortamda, uzay bölümlenme yaklaşımları sıralı arama algoritmalarından daha iyi performans gösterebilir, ancak sıralı arama teknikleri genellikle ortalama olarak daha iyi performans gösterir.

2. k 'ye en yakın komşuların arama yapılarak bulunması

Standart algoritmaların gerektiği gibi performans göstermediği durumlarda bu yöntem, mevcut yaklaşımlara uygun bir alternatif olabilir. K -en yakın komşu sorununun tahmini olan en yakın komşu varyantı bu senaryoda geçerlidir.

3.5.2 Destek Vektör Makineleri

Veri madenciliği söz konusu olduğunda, SVM'leri kullanmanın çeşitli avantajları vardır ve her biri özgündür. SVM ile bölgesel minimumlarla ilgili herhangi bir sorun yoktur (bu durum Sınır Ağları ile ilgili önemli bir sorundur) ve diğer yaklaşımların aksine, SVM'nin etkili bir şekilde çalışması için yalnızca birkaç parametre gereklidir: temel olarak kapasite ve kullanılacak çekirdek (ve çekirdeğin gerektirdiği tüm parametreler). Çekirdek işlevselliği, SVM'lerin performansı üzerinde önemli bir etkiye sahiptir.

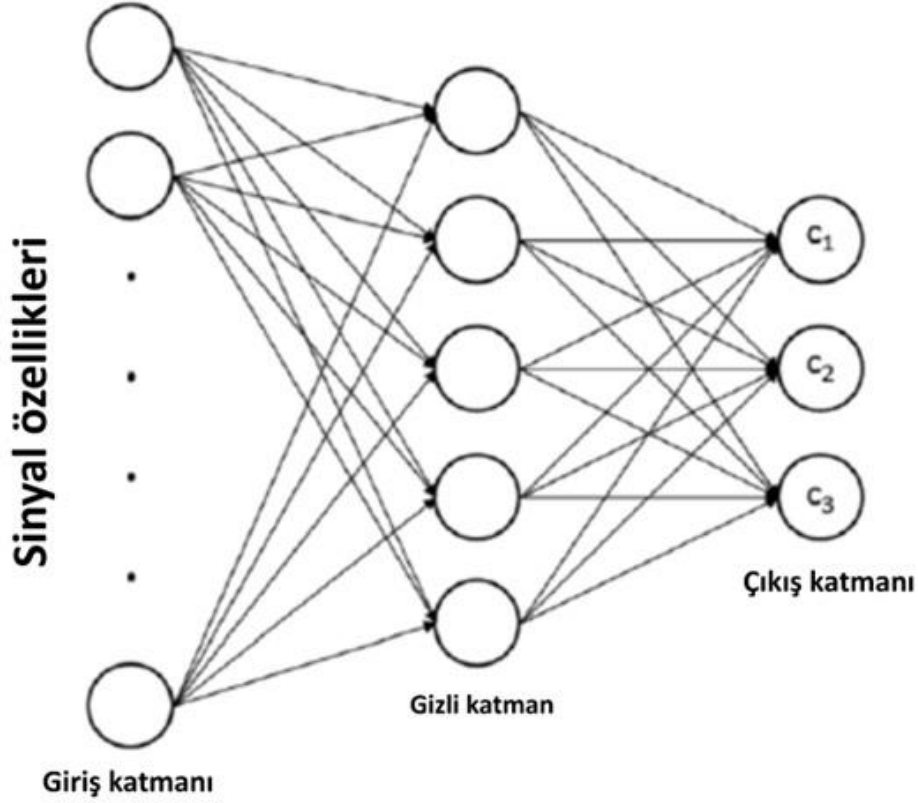


Şekil 3.8. Destek vektör makinesi kullanılarak sınıflandırmanın yapılması.

SVM'ler yüzlerce özelliği işleyebilir, genellikle son derece hızlıdır ve oldukça çok yönlüdür. Çekirdek çok güçlü bir fikirdir çünkü yalnızca uğraştığımız veri türünü değerlendirmemizi gerekli kılar. Koşullara çok uygun, yüksek kaliteli bir Çekirdek ile uğraşırken çok sayıda ön işleme yöntemi gereksizdir. Bunun yerine, mümkün olduğunca "ham forma yakın" verilerle çalışmak önemlidir. Tıpkı gerçek hayatta olduğu gibi, verilerin sınıflandırılmasına odaklanılmalıdır. Bazen veri madenciliği alanında ortaya çıkan tüm matematik ve istatistik konularının ele alınması gerekli değildir. Bunun yerine gerçeklere odaklanılmalıdır. Konu hakkında bilgi sahibi olunması durumunda kolayca bir Kernel geliştirilebilir ve probleme uygulanabilir. Böylelikle düzelme kendiliğinden sağlanacaktır.

3.5.3 Nöral Ağlar

Sınıflandırma ve regresyon uygulamaları için sinir ağlarını kullanmak gelenek bir yöntemdir ve bu yaklaşım popülerlik kazanmaktadır.



Şekil 3.9. Yapay sinir ağlarının yapısı.

SVM tabanlı sınıflandırıcılarda olduğu gibi bir sinir ağı tabanlı sınıflandırıcı, araştırma görüntülerinin etiketlerini tahmin eden bir model üretmek için eğitilmelidir. Sinirsel sınıflandırıcının yapısı üç katmanlıdır: girdi katmanı, çıktı katmanı ve bir veya daha fazla ara katman (gizli katmanlar olarak da bilinir). Bir katmandaki her nöron, üst katmandaki her nöronla etkileşime girer. Sinir ağının son katmanındaki her sınıf etiketi için, görüntüden oluşturulan özellik vektörü ile aynı olan tek bir nöron atanır. Otomatik kodlayıcı sınıflandırıcı, girdiler ve etiketleri arasındaki doğrusal olmayan etkileşimleri tanımlamak için sınıflandırma sürecinde gizli bir katman kullanır. Tek bir gizli katman kullanmak yerine birçok gizli katman kullanıldığında, sürekli olmayan bir ayırt etme fonksiyonunun modellenmesi daha basit hale gelir. Otomatik kodlayıcı sınıflandırıcı eğitimi sırasında her giriş özellik vektörüne bir etiket atamak için nöronlar arasındaki bağlantıların ağırlıklarını tahmin etmek önemlidir. Daha fazla gizli katman eklendikçe, olası ağırlık kombinasyonlarının sayısı hızla artmaktadır.

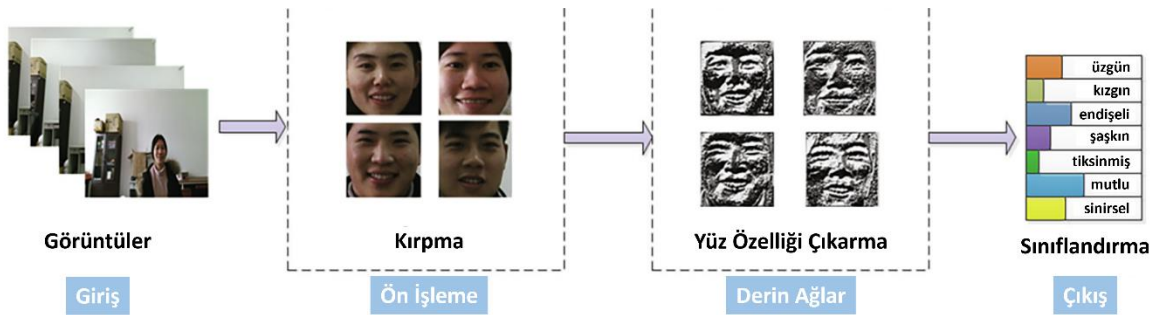
4. ÖNERİLEN YÖNTEM

4.1 Sistemin Ana Hatları

Mevcut çalışma, kayıtlardan elde edilen verilerin analiz süresini azaltmak için görüntü işleme teknikleri, yüz tanıma ve yüz ifadesi tanıma gibi sinir ağlarının kullanımı yoluyla kullanılabilirlik testine otomatik bir yaklaşım geliştirmeyi amaçlamaktadır. Bu şekilde, kullanıcının bir arayüz kullanırken yaptığı ifadelerin nasıl sınıflandırılacağı ve bu ifadelerin kullanıcının sistemle belirli bir anda etkileşimi ile nasıl ilişkilendirileceği analiz edilecektir.

4.1.1 Önerilen CNN Algoritması

Resim tanıma ve sınıflandırma açısından en başarılı yapı konvolüsyonlu sinir ağıdır (CNN). CNN algoritmasıyla çalışmak için yalnızca CNN'in insanları, nesnelere, trafik işaretlerini ve otonom arabaları ayırt etme yeteneğini inceleyeceğiz. Konvolüsyon matematiksel bir yöntem olup “evrişimli” kelimesi bu işlemin bir ağ tarafından kullanılmasını ifade etmektedir. Bir ağ “evrişimli” olarak nitelendiriliyorsa, konvolüsyon yöntemini kullanmaktadır. Konvolüsyona benzer doğrusal bir işlem yoktur. Evrişimli ağlar, konvolüsyonun katmanlarından en az birinde matris çarpımının yerini aldığı bir tür sinir ağıdır. İki ayrı bölüme ayrılırlar. Üç ana rengi doğru bir şekilde tasvir etmek için, gri tonlamalı veya renkli bir görüntü, iki boyutlu veya üç boyutlu bir piksel matrisi (kırmızı, yeşil ve mavi) içermelidir. Şekil 4.1’de bir CNN’nin konvolüsyonlu kısım olarak bilinen bölümü gösterilmektedir.



Şekil 4.1. Önerilen sistemin ana adımları.

“Özellik haritaları” olarak da bilinen konvolüsyon haritaları, bir görüntünün “konvolüsyon çekirdekleri” olarak bilinen bir dizi filtreden geçirilmesiyle oluşturulur. Birkaç ara filtrenin kullanılması ve yerel bir maksimum işlem sayesinde, resmin çözünürlüğü azalır. CNN kodu olarak bilinen bir özellik vektörü elde etmek için haritaları düzleştirme ve onları evrişimli

bir sinir ağına bağlama süreci gereklidir. Çok katmanlı algılayıcılar olarak da bilinen MLP'ler, ağın tüm nöronlarının özelliklerini entegre ederek bir görüntüyü kategorize etmek için kullanılır. Bir MLP'nin çıktı katmanında her sınıf için bir nöron bulunur ve görüntünün sınıf kategorilerinin olasılık dağılımını yansıtır [50].

4.1.2 Evrişimli (Konvolüsyonlu) Sinir Ağlarının Mimarisi

Daha önce söylenenlere göre çok katmanlı algılayıcı, evrişimli nöron(sinir) ağlarının (MLP) temelini sağlar. CNN mimarisi, farklı işlem seviyelerine sahip birkaç yığından oluşur. CNN'in web sitesinde üç farklı CNN seviyesi vardır ve her seviye farklı bir amaca hizmet eder.

- Çeşitli alıcı alanların örtüşme derecesini kontrol eder.
- Çıkış hacmini artırmak, alıcı alanların daha düşük perde seviyelerinde örtüşen alıcı alanların olasılığını artırır.
- Giriş hacminin sınırına sıfır koymak “sıfır doldurma” olarak bilinir ve bazı ayarlarda yöntem “marj (sıfırda)” olarak da bilinir.

İlk evrişimli katmanının (CONV) çıkış hacminin boyutu, aşağıdaki hiper parametreler tarafından belirlenir:

- Derinlik
- Eğim
- Kenar boşluğu
- İnç kare başına konvolüsyon çekirdek yoğunluğu (veya aynı alıcı alanla ilişkili nöron sayısı)

Sağlanması gereken son hiper parametre, sıfır boyutlu dolgudur. Bu marj üretilen ürün miktarını belirler. Özel olarak, hacmin orijinal yüzey alanının korunması istenilebilir [48]. Evrişimli ağların terminolojisinde, konvolüsyonun ilk parametresi genellikle girdi olarak adlandırılırken, ikinci parametre genellikle çekirdek olarak adlandırılır. Uzmanlar genellikle tamamlanmış ürünü “özellik haritası” olarak adlandırırlar. Temel bilgilere bağlı kalma eğilimine rağmen erişilebilecek sayısız karmaşıklığın farkında olmak çok önemlidir.

Konvolüsyon çekirdeği olan bir görüntüyü “geçersiz kılmak” istenirse, esasen konvolüsyonlar arasında bir çekirdek kayması olan dolgu kullanılabilir. Bu durum, konvolüsyon çekirdeği kullanıldığında mümkündür. Genişletici konvolüsyon, daha geniş bir

çekirdek kullanan konvolüsyonla karşılaştırılabilir bir tekniktir (örneğin, konvolüsyonu hesaplamak için iki pikselden biri alınır). Hesaplama göz ardı edilecek piksel sayısını belirleyen genişleme oranı başka bir seçenektir.

Ayrılabilir konvolüsyonlar, genellikle yer değiştirmiş konvolüsyonlar olarak anılırlar. Daha küçük konvolüsyonlara bölünebilen konvolüsyon türleridir. Ayrılabilir konvolüsyonun sonucunda, görüntü yatay veya dikey olarak ters çevrilmiş gibidir. Bu seviyede suyun biriktiği görülebilir.

Bir CNN mimarisinde havuzlama kullanmanın faydalarından biri, bazen sonraki iki konvolüsyonlu katmanı yerleştirerek aşırı yüklemeyi önlemesidir. Havuzlama, küçük çevirilere karşı değişmezliğini korurken, temsillerin ölçeğini küçülterek görüntü alt örneklemedir. Havuzlama, “örnekleme” olarak bilinen resim tekniğine bir örnektir. Ayrıca, havuzlama yönteminin doğrudan bir sonucu olarak öteleme değişmezliği tanımlanmıştır. Havuzlama, gelen verilerin her bir derinlik diliminde bağımsız olarak çalışan ve onu yalnızca yüzey düzeyine ölçeklendiren bir katmandır. Filtreler için en temel kullanım, girdinin maksimum değeri çıktı olarak hizmet ederken 2×2 'lik bir genişlik ve yüksekliktir. Bu durumda “Max-Pool 2×2 ” ifadesi kullanılacaktır. Bir havuzlama stratejisinin kullanılması, mevcut işlem gücünde önemli kazanımlar sağlayabilir. Buna karşılık, günümüzde yaygın yaklaşım çok ince filtreler kullanmaktır. Bu durum, temsil boyutunun agresif bir şekilde küçültülmesinin bir sonucu olarak ortaya çıkar ve bilgi kaybına (2×2 tipi) neden olur. Havuzlama katmanını çıkarmak mümkün olsa da bunu yapmak kılıfın çok sıkı olma olasılığını artırır.

Havuzlama için birkaç farklı yöntem vardır:

- Havuzdan en uygun değerini çıkarılmasını ifade eden “maksimum havuzlama” yöntemi, hesaplamaların hızlı ve verimli bir şekilde yapılmasıyla zor problemlerin basitleştirilebildiği bir yöntem olduğu için en çok kullanılan havuzlama yöntemidir.

–Tüm bireysel değerlerin toplamının, toplam bireysel değerlerin sayısına bölünmesi ile elde edilen “ortalama havuzlama” yönteminde piksel koleksiyonunu doğru bir şekilde temsil etmek için bir ara değere ihtiyaç söz konusudur.

–Ayrıştırılabilir konvolüsyonların iki türünden biri, içerdiği veri miktarıyla çarpılmayan bir ortalama olan “toplam havuzlama” yönteminde sadece değerlerin toplamı hesaplanır.

Tablo 4.1 Önerilen algoritmanın parametreleri

Model İçeriği	Detaylar
İlk konvolüsyon katmanı	3×3 boyutunda 32 filtre, ReLU, giriş boyutu 48×48
İlk maksimum havuzlama katmanı	Havuzlama boyutu 2×2
İkinci konvolüsyon katmanı	3×3 boyutunda 64 filtre, ReLU
İkinci maksimum havuzlama katmanı	Havuzlama boyutu 2×2
Üçüncü konvolüsyon katmanı	3×3 boyutunda 128 filtre, ReLU
Üçüncü maksimum havuzlama katmanı	Havuzlama boyutu 2×2
İlk tam bağlantılı katman	750 düğüm, ReLU
Çıkarma katmanı	Nöronların %50'sini rastgele hariç tutar
İkinci tam bağlantılı katman	850 düğüm, ReLU
Çıkarma katmanı	Nöronların %50'sini rastgele hariç tutar
Üçüncü tam bağlantılı katman	850 düğüm, ReLU
Çıkarma katmanı	Nöronların %50'sini rastgele hariç tutar
İleri tam bağlı katman	750 düğüm, ReLU
Çıkarma katmanı	Nöronların %50'sini rastgele hariç tutar
Çıktı katmanı	7 sınıf için 7 düğüm, Softmax
Optimizasyon işlevi	Stokastik gradyan inişi (SGD)
Öğrenme oranı	0,01
Geri alma	Erken durdurma, LROn Platosunu azaltma, Model kontrol noktası, Tensör panosu

Verilerin özellikleri çıkarıldıktan sonra ağa bir MLP veya çok katmanlı algılayıcı eklenir. Böylece ağın tasarımı tamamlanır. Toplanan özellikler girdi olarak kullanılarak, öğelerinin her biri söz konusu nesnenin belirli bir kategoriye ait olma şansını gösteren N boyutlu bir vektör oluşturulur. Sınıflar bir arada bulunmadığı zaman, her olasılığı belirlemek için softmax fonksiyonu kullanılır [49]. Bir katmanın “tamamen bağlantılı” olduğu söylenildiğinde, altındaki katmandaki her nöronun, üstündeki katmandaki her nörona bağlı olduğu kastedilmektedir.

4.1.3 Aktivasyon Fonksiyonları

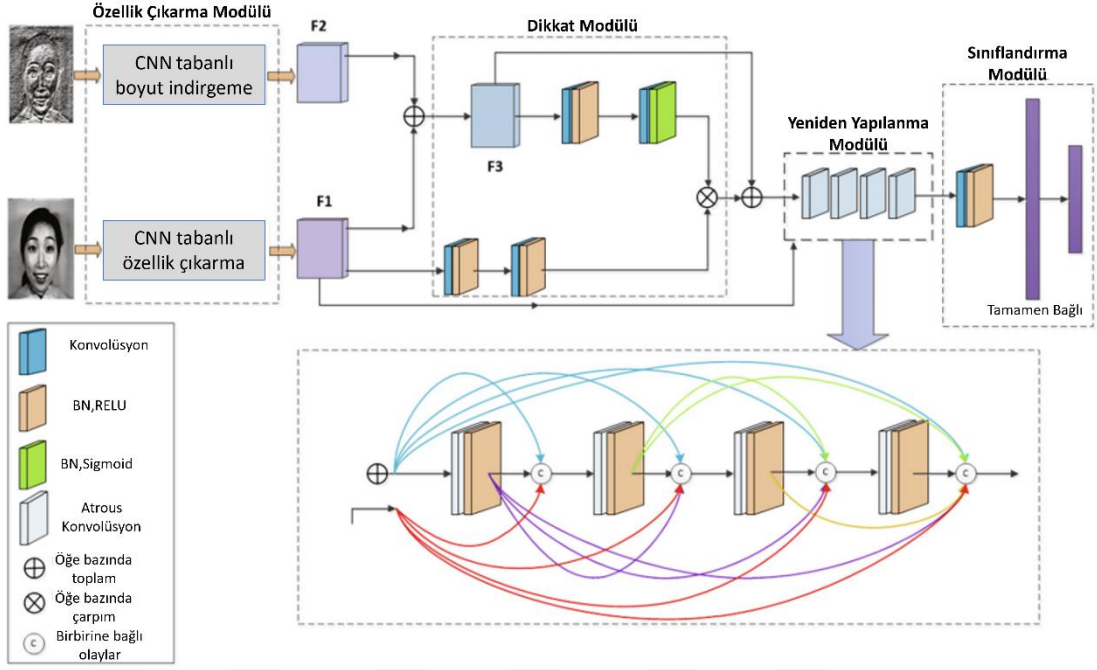
Yapay bir nöronun çıkış sinyali, matematiksel bir fonksiyon olan bir aktivasyon fonksiyonuna tabidir. Bir uyarının, beyin reaksiyonunu indükleyen bir eşiği aşmasına “aktivasyon potansiyeli” denir. Ortak aktivasyon fonksiyonları doğrusal değildir. Gizli bir katmanın ağırlıklarının sadece iki katına çıkarılmasının doğrusal bir dönüşümle sonuçlandığı göz önüne alındığında, amaç sinir ağlarının temel doğrusal regresyondan daha karmaşık işlevleri öğrenmesini sağlamaktır. Etkin bir fonksiyon örneği, “düzeltilmiş doğrusal birimler” (ReLU) olarak bilinir. Her konvolüsyon işleminin tamamlanmasından sonra tüm negatif piksel değerleri sıfırlanır.

4.1.4 CNN’de Önerilen Değişiklikler

Gradyan iniş yöntemi, yapay sinir ağları ve lojistik regresyon gibi makine öğrenimi teknikleri tarafından kullanılan yaygın bir optimizasyon aracıdır. Bu işlem, katsayıları veya ağırlıkları belirlemek için yapılır. Eğitim verilerine dayanarak tahminler üretmek mümkündür ve bu tahminlerden kaynaklanan hatalar, modeli iyileştirmek ve hatalı tahminler üretme sayısını azaltmak için kullanılır. Bu yaklaşım, eğitim veri kümesine uygulandığında tahmin modeline en az hata derecesi sağlayan model parametrelerini (katsayılar veya ağırlıklar gibi) belirlemek için kullanılmalıdır. Modeli ideal bir hata değerine yaklaştırmak için hata eğimleri kullanılır. Gradyan iniş yöntemi üç şekilde gerçekleştirilebilir:

- Sürekli yükselen toplu gradyanı
- Eğitim inişi için stokastik yöntemlerin uygulanması
- Küçük parçalarda gradyan iniş

Bu çalışma, birçok sinir ağı türünün ve bunların türleri nasıl işlediklerinin bir açıklamasını sunmaktadır. Tespitin hassasiyetini artırmak için, lineer aktivasyon fonksiyonları ekleyerek ve CNN’in yerleşik softmax katmanını ek sınıflandırıcılarla değiştirerek CNN yönteminde hassas ayar yapıyoruz. Böylece CNN daha doğru hale getirilebilir. Daha hızlı bir iş akışını kolaylaştırmak için, güncellenen CNN algoritması, Conv-2 katmanının yardımıyla özniteliklerin çıkarılmasına ek olarak, özellik şablonlarından ve Haar benzeri özelliklerden yararlanmaktadır. Buradan çıkarılabilecek en önemli nokta budur. Bir sonraki önemli adım, birden çok katmandan aynı ağırlıklara sahip Haar benzeri özellikleri, hesaplama süresini azaltacak ve işlemeyi hızlandıracak tek bir özellikte birleştirmektir.



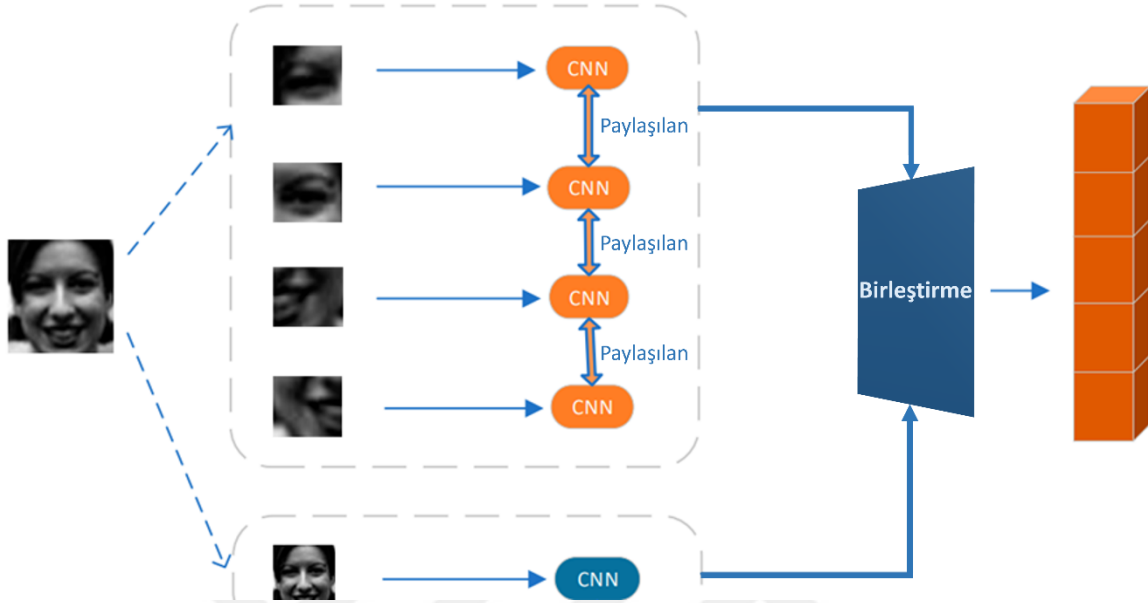
Şekil 4.2. CNN algoritması için önerilen optimizasyon yöntemleri [72].

Daha sonra, CNN evrişimli sinir ağlarını ve yapılarını ve farklı katmanlarını açıklamaya çalışılmıştır. CNN evrişimli doğrusal olmayan fonksiyon (ReLU), havuzlama ve tam bağlantılı katman olmak üzere dört ana işlemi vardır. İlk işlem, girdi görüntüsünden özellik çıkarımı için konvolüsyondur. İkinci işlem, tüm negatif piksel değerlerini sıfırla değiştirmek için doğrusal olmayan işlemdir (ReLU) [73]. Üçüncü işlem, düzeltilmiş özellik haritasının boyutunu kademeli olarak azaltmak için havuzlamadır. Son olarak, sınıflandırma için tamamen bağlantılı bir katmandır. Tüm bu işlemlerin sonunda, aralarında Xception ve VGGnet'in de bulunduğu bazı mimari örnekleri sunulmuştur.

4.1.5 Özellik Füzyonu

Evrişimli sinir ağı performanslarının çoğu, çok küçük veri kümeleriyle aşırı uyum göstermenin getirdiği ifadeden bağımsız sınıf içi eşitsizlikler nedeniyle önemli zorluklarla karşı karşıyadır. İstenen sonuçları elde etmek için bu endişeleri gidermenin daha etkili bir yolu olarak CNN katmanlarına bir Çift yönlü özellik füzyonu yerleştirilmesini öneriyoruz. Başlangıçta, yüzlerin ana kısımlarını ayırmak ve yüzün geri kalanını göz ardı ederek bu bölgelerden bilgi çıkarmak için segmentasyon kullanılır. Bu işlem, yüzün geri kalanını göz ardı ederek yapılır ve her sınıfta görülen çeşitlilik miktarını etkili bir şekilde daha düşük bir düzeye indirmektedir. İkinci olarak, derin ağ problem olan aşırı uyum sorunu, küresel resmin yanı sıra parçalanmış yerel görüntü bölgelerinden oluşan eğitim verileriyle entegre bir çift

yollu model sağlayarak etkili bir şekilde hafifletilebilir. Bu durum, modelin görüntünün yapısını daha doğru bir şekilde öğrenmesini sağlayacaktır. Sonuç olarak model, fotoğrafların içeriğini çok daha iyi yorumlayabilecektir.

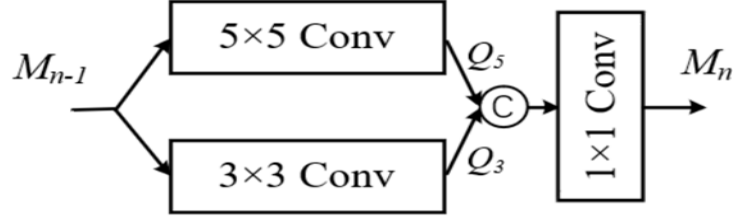


Şekil 4.3. Paylaşılan CNN katmanlarında özelliklerin birleştirilmesi.

Bu düşünce, iki farklı bileşenden oluşmaktadır. Çift yönlü özellik birleştirme modülünün yardımıyla, giriş yüz görüntüsü eş zamanlı olarak segmentasyon, özellik çıkarma ve füzyon işlemlerinden geçebilir. İkinci olasılık, her kanalın genel sınıflandırma kararına yaptığı katkılara göreli önem vermek için özellik haritasının farklı kanallarını kullanan SA modülüdür. İlk modülün tamamlanmasıyla ilgili adımlar aşağıda bütünüyle özetlenmiştir. Model, önce bölgesel alanı temel görüntüden çıkararak ağ için girdi hazırlama sürecini başlatır. Bundan sonra hem bölgesel alan görüntüsünü hem de temel resmin boyutunu sıkıştırır ve ölçeklendirir. Çift yönlü özellik füzyon ağı, dört yerel bölgeden çıkarılan özellikleri alır ve bunları tam yüz resminden çıkarılan özelliklerle birleştirir. Bu sürecin sonucunda, birleştirilmiş özellik haritasının oluşturulur. SA modülünün ana sorumluluğu, her katmanda bulunan özellik haritalarına öncelik vermek ve gereksiz verileri sıkıştırmaktır. Doğrusal bağlantı katmanı nihayet bize uzun zamandır gelen ifadeleri etiketleme kapasitesini sağlamaktadır.

Yakındaki ağ modüllerinde bulunan ve aynı boyuta sahip yüz özelliklerini entegre etmek için, eleman bazında toplama ve birleştirme teknikleri rutin olarak kullanılır. Bu teknikler oldukça sık kullanılmaktadır. Bu yöntemler, tüm ağ modüllerinden gelen özellikleri aynı

oranda entegre ettiğinden, farklı perspektiflerden ve derinlik modüllerinden gelen özelliklerin sınıflandırma görevine önemli katkılar sağladığı gerçeğini göremezler.



Şekil 4.4. Convnet mimarisinde iki özelliği birleştirme.

Bu modüllerin farklı bakış açıları sağladığı düşünülebilir.

Tablo 4.2 Füzyondan önceki ve sonraki özellik boyutları

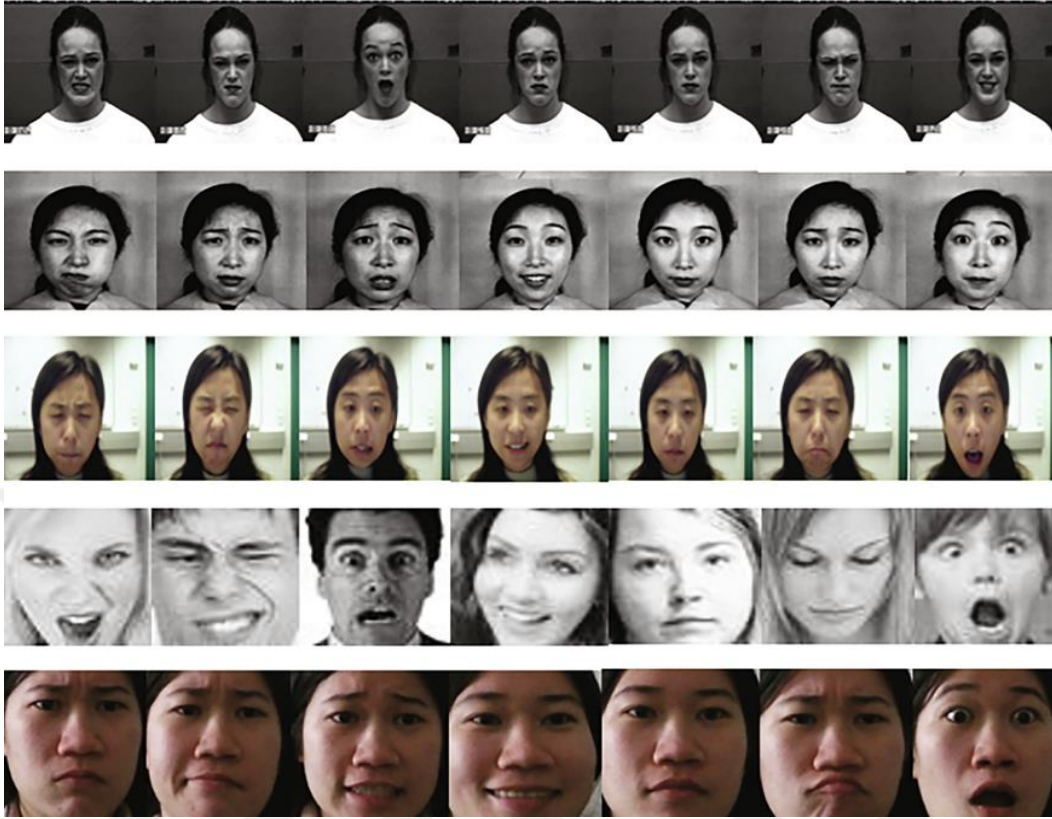
Özellik adı	örnek sayısı	Düşük düzeyli boyut	Orta düzeyli boyut	Yüksek düzeyli boyut	Füzyon faktörü	Birleştirilmiş özellik boyutu
Göz	20	15×15	30×30	60×60	0,5	25×30
Burun	7	15×15	30×30	60×60	0,5	25×30
Ağız	7	15×15	30×30	60×60	0,5	25×30
Kaş	25	15×15	30×30	60×60	0,5	25×30
Yanaklar	32	15×15	30×30	60×60	0,5	25×30

Programda hem Gri hem de Renkli Görüntüleri desteklemekte olup, Alfa Faktörü, her görüntünün karıştırma oranını değiştirmek için değiştirilebilir. Alfa Faktörü = 0,5 ile iki görüntü eşit olarak karıştırılır. Alfa Faktörü < 0,5 ile arka plan görüntüsünün katkısı daha fazla olacaktır. Alfa Faktörü > 0,5 ile ön plan görüntüsünün katkısı daha fazla olacaktır.

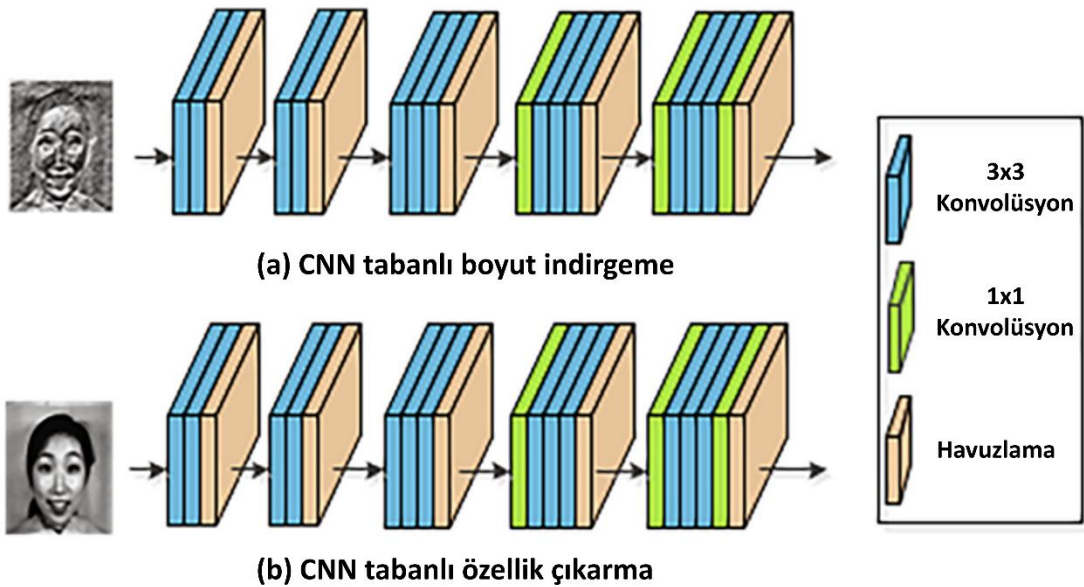
4.2 Veri Tabanı

Yüz ifadesi tanıma için kullanılan veri tabanı, Şekil 4.5'te görülebileceği gibi, neşe, üzüntü, şaşkınlık, öfke, tiksinti ve korku olmak üzere 6 temel ifadenin 213 görüntüsünü ve ayrıca 10 Japon kadın model tarafından oluşturulan doğal bir ifadeyi sağlayan JAFFE görüntü bankasıydı. Veri tabanının sağladığı ifadeler arasında, bu tür ifadeleri bilgisayar sistemlerinin kullanımında oluşabilecek problemlerle ilişkilendirmenin zorluğundan dolayı üzüntü ve korku duygularına ait görseller çıkarılmış, diğer ifadelerle 150 görsel bırakılmıştır [74]. Burada görüntüler başlangıçta 64×64 piksel olarak yeniden boyutlandırılmıştır.

Sınıflandırmanın düşük görüntü sayısından etkilenmemesi için, orijinal görüntülerin 48×48 piksellik kesimleri, Şekil 4.5'te görüldüğü gibi, her görüntü için toplam 16 kesim seçilmiştir.



Şekil 4.5. JAFFE veri tabanından örnek ifadeler; soldan sağa tiksinti, şaşkınlık, mutluluk, üzüntü ve korku ifadeleri [75-77].



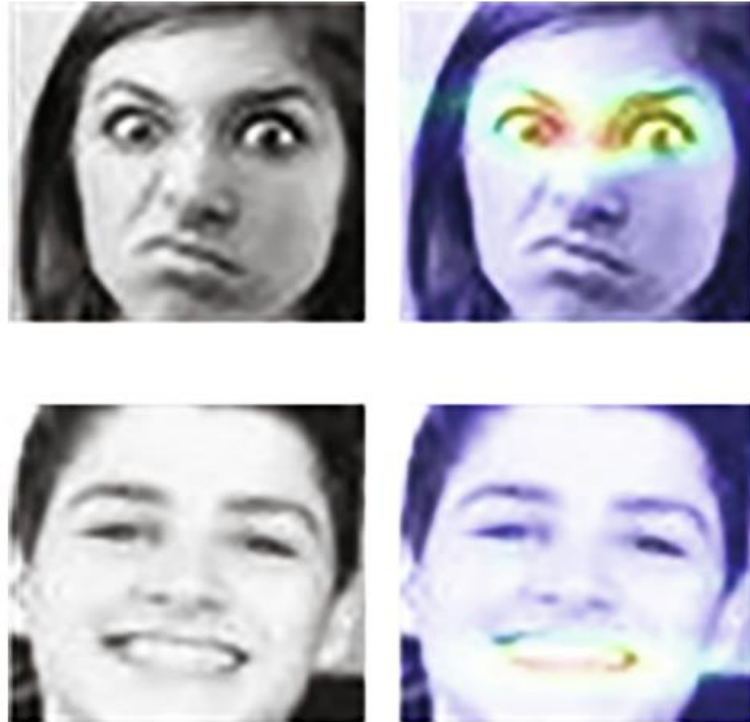
Şekil 4.6. CNN'ye dayalı Haar özellik çıkarma yöntemi.

Bu teknik, sınıflandırma için verileri artırmak ve ağır performansını iyileştirmek için, tabandaki görüntü sayısını Şekil 4.6’te görüldüğü gibi 2400 dağıtılmış görüntüye çıkarmak için kullanılmıştır.

4.3 Video Çekim Modülü

Bu adımda, kullanıcıların ifadelerinin kayıt altına alınabilmesi ve bununla birlikte OpenCV kullanılarak yüzün görüntülerinin çıkarılabilmesi için bir video çekimi modülü geliştirilmiştir. Başlangıçta C/C++’da geliştirilmiş; nesne tanıma, yüz algılama ve diğer görüntü ile ilgili etkinlikler için yaygın olarak kullanılan açık kaynaklı bir kütüphanedir. OpenCV kütüphanesini kullanmanın avantajlarından biri, Şekil 4.7’te gösterildiği gibi Haar Özellikleri tabanlı kademeli sınıflandırıcıya sahip olmasıdır.

Her özellik, siyah dikdörtgenin altındaki piksellerin toplamından beyaz dikdörtgenin altındaki piksellerin toplamı çıkarılarak elde edilen benzersiz bir değerdir. OpenCV kütüphanesini de destekleyen Python programlama dilinde geliştirilen modül, ana video kayıt cihazından ve çalıştığı bilgisayar ekranından kare görüntüler yakalar ve yüzü algılamak için kademeli sınıflandırıcıyı kullanarak görüntünün bir dilimini oluşturur. Daha sonra bu görüntü dilimini, ifadeyi sınıflandırabilen bir sinir ağına girdi olarak gönderir.



Şekil 4.7. Video yakalama aşamasında algılanan yüzler.

Sınıflandırmadan sonra modül, resimde tanımlanan duyguyu geri dönüş olarak alır ve Şekil 4.8’te görüldüğü gibi video çekiminde algılanan duygular elde edilir.



Şekil 4.8. Video çekiminde algılanan duygular.

4.4 İfadelerin Sınıflandırılması

Bölüm 4.2’de açıklanan nihai veri setinin görüntülerinin sınıflandırılma adımı için bir Evrişimli Sinir Ağı kullanılmıştır. İki boyutlu yapıları göz önünde bulundurularak filtrelemeye izin verdikleri için, esas olarak görüntülere uygulanan görsel bilgilerin işlenmesinin önemi göz önüne alındığında, CNN bu aşama için en iyi seçenek olduğunu kanıtlamıştır. Her bir kategorideki eğitim için az sayıda örneklem verildiğinde, en iyi sonucu veren bulmak için, eğitimin kaç kez tekrarlanacağını tanımlayan az sayıda döngü ile CNN mimarisinin bazı konfigürasyonları kullanılarak deneyler yapılmıştır.

En büyük potansiyeli gösteren CNN, LeNET’e benzer bir mimari olmakla beraber, aşırı uyum sorununu azaltmaya yardımcı olmak için yoklama katmanlarından önce, ağın öğrenme aşamasında daha az ilgili nöronların devre dışı bırakılmasından sorumlu olan ağa bırakma katmanları uygulanmıştır. Kullanılan CNN mimarisi, başlangıçta serpiştirilmiş konvolüsyon ve aktivasyon katmanlarından oluşan 18 katmandan, ardından bir maksimum havuzlama katmanından ve bir bırakma katmanından oluşmaktadır. Model, maksimum havuzlama ve bırakmadan sonra daha fazla konvolüsyon ve etkinleştirme serpiştirilmesiyle (ReLU) tekrar eder. Kısa bir süre sonra, yoğun bir katman, bir aktivasyon katmanı ve Tablo 4.3’te

gösterildiği gibi başka bir yoğun katman ve bir aktivasyon katmanı ile biten üçüncü bir çıkarma katmanı gelir.

Tablo 4.3 Önerilen sistemin karışıklık matrisi

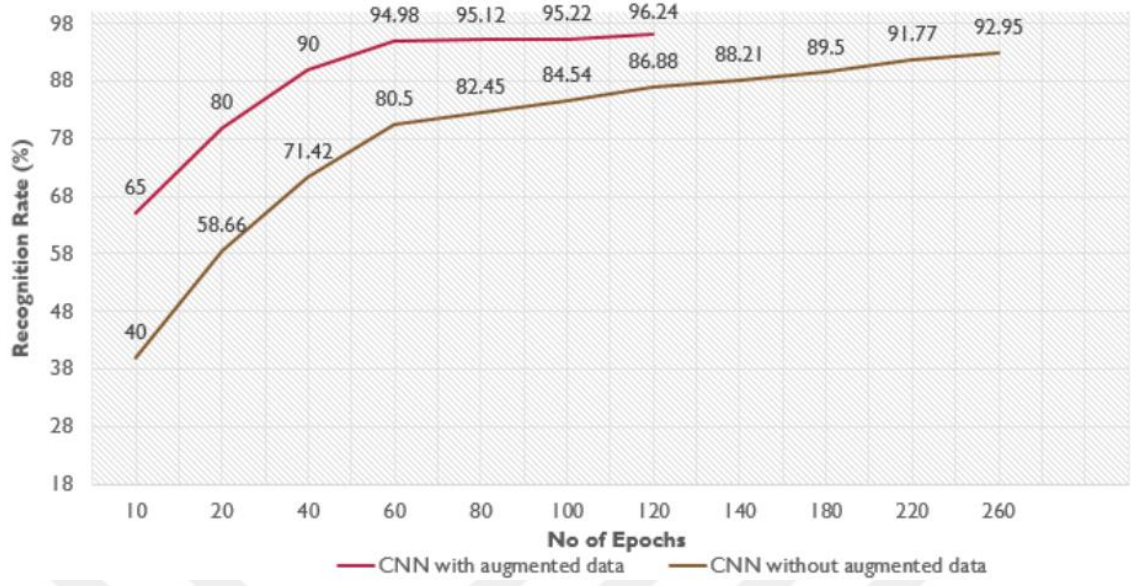
Cf matrisi ×							
Sinirli	0,94	0,01	0,02	0,0	0,0	0,02	0,0
İğrenme	0,02	0,96	0,0	0,0	0,0	0,01	0,0
Korku	0,0	0,0	0,96	0,0	0,0	0,0	0,02
Mutlu	0,0	0,0	0,04	0,90	0,01	0,04	0,0
Doğal	0,0	0,02	0,03	0,0	0,92	0,02	0,01
Üzgün	0,02	0,0	0,0	0,0	0,04	0,91	0,01
Şaşırmış	0,0	0,0	0,02	0,0	0,0	0,0	0,97
	Sinirli	Tiksinti	Korku	Mutlu	Doğal	Üzgün	Şaşırmış
Öngörülen etiket							

4.5 Sonuçlar

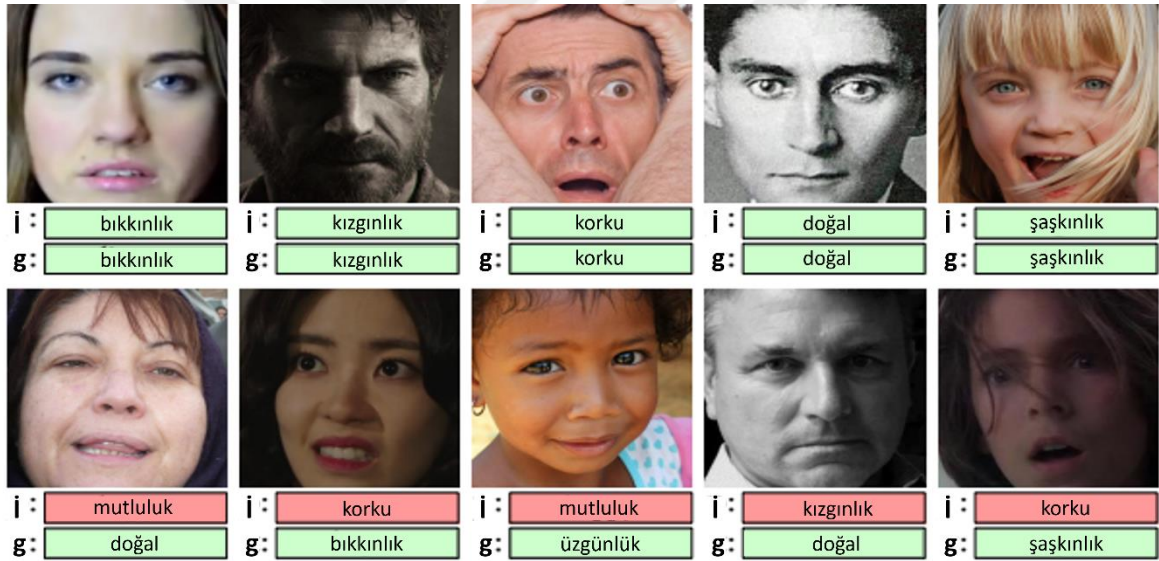
Bölüm 4.2’de açıklanan veri seti; eğitim seti için %80 ve doğrulama seti için %20 oranında iki alt gruba ayrılmıştır. Eğitim için toplam 1920 örnek, sevinç için 393 örnek, tiksinti için 371 örnek, öfke, şaşırma ve doğal için 384 örnek olarak sınıflandırılmış ve doğrulama için 480 örnek, sevinç için 100 örnek, tiksinti için 93 örnek, öfke, şaşırma ve doğal için 96 örnek değerlendirilmiştir.

Deneyler sırasında, dönem sayısı ve her katmandaki nöron sayısı gibi çeşitli parametreler keyfi olarak değiştirilmiştir. Öncelikle 100 döngü sayısı ile, her yoğun katmandaki nöronların sayısı analiz edilmiştir. İlk yoğun katmanda 512 ve 1024, ikinci yoğun katmanda 5 (sınıf sayısı) değerleri uygulanmıştır. İlk yoğun katmanın 1024’e göre daha iyi olduğu kanıtlanmıştır. Her ikisinin de yaklaşık %94’lük bir başlangıç doğruluğuna ulaşmasına rağmen 512 kazanç ve 100 döngü sağlanmıştır. Ağ bir seferde 50 periyot çalışmıştır. Böylece, Şekil 4.9’da görüldüğü gibi, en yüksek geçerlilik doğruluğu 500 iterasyondan sonra elde edilmiştir. Bununla birlikte, daha sonraki zamanlarda ağın doğruluğu önemli ölçüde gelişmemiş ve 600 iterasyona kadar çıkmıştır. Fakat bu noktada aşırı uyumdan mustarip olmaya başlamıştır.

JAFFE tabanının artırılmış veri setini kullanan CNN, 3 kademeli çıkarma katmanının desteklediği nihai mimariyle, yüz ifadelerinin sınıflandırılmasının karmaşıklığı nedeniyle görüntüler üzerinde herhangi bir ön işlem yapılmadığı ve eğitim için kullanılan görüntü sayısının az olduğu düşünüldüğünde, doğru sınıflandırma sayısının doğrulama için toplam örnek sayısına bölünmesiyle hesaplanan %97’lik bir doğruluk elde etmektedir.



Şekil 4.9. Model doğruluğu.



Şekil 4.10. Yanlış ifade sınıflandırması ve doğru ifade sınıflandırması örnekleri.

Şekil 4.10’da CNN tarafından yapılan bir sınıflandırılma gösterilmektedir. Burada ‘i’ harfi ile algılanan ifade, ‘g’ harfiyle gerçekte olan durum belirtilmektedir. Sınıflandırıcıların üst sırasında sırasıyla bıkkınlık, kızgınlık, korku, doğal ve şaşkınlık ifadeleri doğru bir şekilde değerlendirilmiştir. Sınıflandırıcıların alt sırası incelendiğinde, mutlu ifadenin yanlış yollardan biri ile derecelendirildiği iki görüntüden biri doğal, diğeri üzünlük ifadesini göstermektedir. Korku ifadesinin yanlış yollardan biri ile derecelendirildiği iki görüntüden

biri bıkkınlık, diğeri şaşkınlık ifadesini göstermektedir. Kızgınlık ifadesinin yanlış değerlendirildiği görüntüde ise gerçekte doğal ifade gösterilmiştir.

Bu yöntemin doğrulaması, farklı veri kümeleri kullanılarak yapılan 10 deney sonucunda elde edilmiştir. Elde edilen bulgular Tablo 4.4 ve Tablo 4.5'te verilmiştir.

Tablo 4.4 JAFFE 1M veri setindeki döneme göre doğruluk

Yineleme	Dönem	Doğruluk	Veri Kümesi
1	4	%80,4	JAFFE 1M
2	12	%88,5	JAFFE 1M
3	15	%93,2	JAFFE 1M
4	26	%91,5	JAFFE 1M
5	29	%94,6	JAFFE 1M
6	33	%91,3	JAFFE 1M
7	36	%88,7	JAFFE 1M
8	39	%85,2	JAFFE 1M
9	40	%82,8	JAFFE 1M
10	43	%81,6	JAFFE 1M

Tablo 4.5 CNN'e dayalı farklı algılama algoritmalarının doğruluğunun karşılaştırılması

Yötem	Veri seti	Doğruluk
CNN	Digi-Face 1M	%91,4
VG16	VGG Face2	%94,1
FACENET	UMDFaces	%90,9
U-NET	CelebA	%93,15
RESNET	UMDFaces	%92,6
PROPOSED	JAFFE 1M	%94,6

5. SONUÇ

Algılama doğruluğunu artırmak için değiştirilmiş CNN algoritması, yalnızca Conv-2 katmanını kullanarak özellikleri tespit etmekle kalmaz, aynı zamanda iş akışını hızlandırmak için özellik şablonları ve Haar benzeri özellikler kullanır. Bu modifikasyonlar, CNN'nin bölgesel softmax katmanının diğer sınıflandırıcılarla değiştirilmesiyle mümkün olmuştur. Önerilen CNN'de kullanılan özellik birleştirme yöntemi, söz konusu özellikleri sınıflandırmak için gereken süreyi optimize etmemizi sağlamıştır.

Az sayıda görüntü içeren bir veri tabanı ve görüntülerde ön işleme yapılamaması gibi elverişsiz koşullarda bile CNN'in ifadelerin sınıflandırılması için mükemmel bir seçim olduğunu kanıtlayabildiği görülmektedir. Yakalama ve sınıflandırma modülü, bir bilgisayar sisteminin kullanıcı üzerinde meydana getirebileceği duygusal etkinin bilinmesini ve bu bilgilerle kullanıcı etkileşimini iyileştirmek için önlemler alınmasını mümkün kılmaktadır. Böylece, kullanılabilirlik testlerine uygulanan ifadeleri sınıflandırmak için CNN kullanmanın mümkün olduğu ve hatta testi yapanlara gerçek zamanlı yanıtlar verebileceği sonucuna varılmaktadır. Son on yılda bilgisayarla görme araştırmaları alanında büyük ilerlemeler kaydedilmiştir. Son yıllarda, yüz ifadelerini analiz eden bilgisayarlı görme alanında çok büyük bir ilerleme kaydedilmiştir. İstihbarat sistemi girdi olarak bir yüz görüntüsü aldığından yüz ifadesi tanıma (FER) gereklidir. Derin öğrenme araştırması alanındaki son gelişmelere bağlı olarak CNN'ler, olağanüstü görüntü tanıma yeteneği göstermiştir. CNN tabanlı FER yaklaşımları, daha geleneksel FER stratejilerinden çok daha iyi performans göstermiştir. Buna karşılık, gerçek dünya uygulamaları için bir FER modeli oluştururken üstesinden gelinmesi gereken birkaç engel vardır. Gerçek dünya uygulamaları için FER, yalnızca vahşi ortamda çekilen ve açıklama eklenmiş fotoğraf koleksiyonları olan vahşi FER veri kümelerini gerektirmektedir. Cinsiyet, yaş, aydınlatma koşulları, sosyoekonomik geçmiş ve atış kalitesi dahil olmak üzere çok sayıda kriter, vahşi hayvanların geniş ölçekli veri tabanlarında bireyleri ayırt eder. Sonuç olarak, vahşi FER veri kümeleri iki engelle karşı karşıyadır. Sınıflar arasında önemli miktarda tutarlılık olmasına rağmen, her sınıf içinde büyük bir çeşitlilik vardır. Derin öğrenme yöntemlerinin tipik uygulamaları, performansı beklentilerin altında kalan tahmin modelleri üretmektedir. Derin Metrik Öğrenmeden ilham alan modeller ve süreçler kullanan bu tez ile, yukarıda bahsedilen iki sorun çözüme kavuşacaktır. Bu modeller ve stratejiler, zorlu ortamlarda mevcut emsallerinden daha iyi performans gösterecektir. Genellikle CNN kaybı olarak bilinen

diskriminant dağıtım-agnostik kaybı, yaygın olarak kullanılan softmax kaybının bir uzantısıdır. CNN kaybı, bir sınıfın derin özelliklerini diğer sınıfların özelliklerinden uzaklaştırırken aynı anda onları birbirine yaklaştırma etkisine sahiptir. Bu durum hem sınıflar arası ayrımı hem de sınıf içi kompaktlığı sağlar. CNN kaybı, derin özellikleri ayırt etmek için etkili bir yöntemdir ve hem çoğunluk hem de azınlıklar için kullanılabilir. Stokastik gradyan iniş (SGD) yönteminin temel bir anlayışıyla, CNN kaybı büyük ölçüde azaltılabilir. Mevcut teknolojik gelişmelerden herhangi birinden bir CNN mimarisi, herhangi bir görsel tanıma problemini ele alabilir. Önerilen CNN kayıp metodolojisinin performansını temel tekniklere (softmax kaybı ve merkez kaybı) göre değerlendirmek için simüle edilmiş bir vahşi veri seti kullanılır. Bu çalışma ile CNN kaybı için önerilen stratejimizin emsallerinden daha üstün olduğunu göstermek için, iyi bilinen iki vahşi FER veri seti üzerinde kapsamlı testler yaparak, sorunu çözmek amacıyla yeni bir yaklaşım denenmektedir. Sonuçlarımıza göre standart DML teknikleri, aşırı uyum sorunu ve yetersiz genelleme ile sonuçlanır. RELU'nun kapılı bir varyantı olan RELU ve g-RELU, özellikle doğal ortamlarda kullanım için tasarlanmış iki anti-FER ilacıdır. Derin bir özelliği gömme alanı boyutları boyunca ayırt etmek için hem g-RELU yöntemi hem de geleneksel RELU yöntemi, seyrek merkez kaybı olarak bilinen bir teknik kullanır. Bu seyrek merkez kaybı, geleneksel SGD yaklaşımı kullanılarak eğitilmiştir. Bu seyrek merkez kaybı daha sonra olası maksimum performans seviyesini elde etmek için softmax kaybı ile birleştirilir. Ayrıca, modüler bir dikkat mekanizması kullanılarak seyrek bir merkezin eksikliğini telafi edebilen, tekrarlayan bir sinir ağı (CNN) oluşturulmaktadır. Bu yöntem tüm CNN mimarilerine uygulanabilir. Dikkat ağırlıkları, girdiyi dikkate alarak derin özelliğin her bir boyutundaki katkısını değerlendirmektedir. RELU stratejisinin ana bileşenleri, seyrek bir merkez kaybı ve bir dikkat mekanizmasıdır. Kapsamlı testler, RELU'nun düzenli olarak kullanılan iki vahşi FER veri kümesine (softmax kaybı ve merkez kaybı) uygulandığında hem en gelişmiş yaklaşımlardan hem de temel yöntemlerden daha iyi performans gösterdiğini ortaya koymaktadır. Performans söz konusu olduğunda, g-RELU'nun saf RELU ile karşılaştırılabilir sonuçlar verdiği de görülmektedir. Diğer DML hedef işlevleri, göreceli kolaylıkla RELU'ya eklenebilir. Temeldeki tahmin modeli, önerilen işlemlere dahil edilebilecek yüksek düzeyde ayırt edici özellikler üretebilir. Tüm modeller uçtan uca bir öğrenme çerçevesi içinde eğitildiğinde, bir dizi pratik fayda sağlanır. Bu paradigmayı kullanırken avantajlardan biri, her modülün kendisinden önce gelen modüller üzerine inşa edilmesidir. Böylece gereken kod miktarını azaltılmış olur. Tahmin modelleri, GPU'ların eşzamanlı hesaplamaları kullanılarak paralel olarak eğitilebildiğinden, kısa sürede belirli bir

hedefe göre hızlı bir şekilde özelleştirilebilirler. RELU, eğitim sürecini basitleştiren hiper parametrelerin değiştirilmesine ihtiyaç duymadığı için CNN kaybına göre önemli bir avantaja sahiptir. RELU, tüm sınıflarda artan tanıma doğruluğunu vurguladığı için CNN kaybına göre ek bir avantaja sahiptir, CNN kaybı ise yeterince temsil edilmeyen sınıflara daha fazla vurgu yapar. RELU'nun entegre modülü ise öğretilmesi gereken iş miktarını artırır. Sonuç olarak, belirli bir uygulama için en uygun yöntemi oluştururken bir dengeleme göz önünde bulundurulmalıdır. Ayrıca, FER'in bazı durumlarda engeller çıkarmaya devam ettiği bilinmekte, ancak bu durum kabul edilmektedir. FER veri kümelerine kesin olarak açıklama eklemek zordur. Devasa bir veri kümesine açıklama eklemek gerektiğinde, bir araştırma ekibinin etiketleme hatalarını önlemesi zordur. Eğitim veri kümelerinde büyük miktarda etiket gürültüsü olduğunda, öğrenme stratejisi zarar görür. Az miktarda arka plan gürültüsü bile genellemeyi artırabilir. Açıklamalı yüz ifadelerinin yorumlanması, bunun doğrudan bir sonucu olarak çok öznelidir.

5.1 Öneriler

Bu çalışma, gelecekteki birçok çalışma için alan açmakla beraber, farklı cinsiyet, yaş ve etnik kökene sahip çok sayıda insan görüntüsü ile kontrollü bir ortamda yüz ifadelerinden oluşan bir veri tabanının oluşturulmasının CNN'den gelen doğruluğu önemli ölçüde artırabileceğini vurgulayabiliriz. Ağın doğruluğunu artırmak için, elde edilen görüntüleri kaydetmek ve tabanı daha fazla görüntü ile beslemek amacıyla kullanmak için gerçek kullanılabilirlik testlerinde modülün kullanılması da mümkündür. Ancak bunun için manuel olarak doğrulama yapılması gerekecektir. Yeni görüntülerle yeniden eğitim yapmadan önce ağın her tahmini, ayrıca bu doğrulamayı gerçekleştirmenin otomatik yollarını bulmak amacıyla araştırmaya alan açmaktadır. Gelecekte yapılacak çalışmalarda üretilen verilerin analiz aşamasında, yardımcı olacak daha fazla raporun elde edilmesiyle modül daha da geliştirilebilir. Bu nedenle bilgisayar sistemlerinin kullanımını daha keyifli hale getirmek amacıyla kullanılabilirlik testi, duyuşsal bilgi işlem ve kullanıcının duygularının tanımlanması işlemleri de sürekli olarak geliştirilerek uygulanmalıdır.

KAYNAKLAR

- [1]. Ekman, P. and Friesen, W. V., 1978, Facial action coding system consulting psychologists press. Palo Alto, CA
- [2]. Russell, J. A., 1980, A circumplex model of affect. *Journal of personality and social psychology*, 39(6), 1161.
- [3]. Alshamsi, H., Kepuska, V., & Meng, H. (2017, October). Real time automated facial expression recognition app development on smart phones. In 2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON) (pp. 384-392). IEEE.
- [4]. Zhong, L., Liu, Q., Yang, P., Liu, B., Huang, J. and Metaxas, D. N., 2012, June, Learning active facial patches for expression analysis. In 2012 IEEE Conference on Computer Vision and Pattern Recognition, 2562-2569.
- [5]. Goodfellow, I. J., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B. and Bengio, Y., 2013, Challenges in representation learning: A report on three machine learning contests. In *International conference on neural information processing*. Springer, 117-124.
- [6]. Dhall, A., Ramana Murthy, O. V., Goecke, R., Joshi, J. and Gedeon, T., 2015, November). Video and image based emotion recognition challenges in the wild: EmotiW 2015. In *Proceedings of the 2015 ACM on international conference on multimodal interaction*, 423-426.
- [7]. Krizhevsky, A., Sutskever, I. and Hinton, G. E., 2012, Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25., 1097-1105.
- [8]. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A., 2015, Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1-9.
- [9]. Schubert, E., 1999, Measuring emotion continuously: Validity and reliability of the two-dimensional emotion-space. *Australian Journal of Psychology*, 51(3), 154-165.
- [10]. Ekman, P. and Friesen, W. V., 1971, Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2), 124.
- [11]. Shan, C., Gong, S. and McOwan, P. W., 2009, Facial expression recognition based on local binary patterns: A comprehensive study. *Image and vision Computing*, 27(6), 803-816.
- [12]. Zhao, G. and Pietikainen, M., 2007, Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE transactions on pattern analysis and machine intelligence*, 29(6), 915-928.
- [13]. Zhi, R., Flierl, M., Ruan, Q. and Kleijn, W. B., 2010, Graph-preserving sparse nonnegative matrix factorization with application to facial expression recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 41(1), 38-52.
- [14]. Mavadati, S. M., Mahoor, M. H., Bartlett, K., Trinh, P. and Cohn, J. F., 2013, Disfa: A spontaneous facial action intensity database. *IEEE Transactions on Affective Computing*, 4(2), 151-160.
- [15]. LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., 1998, Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.

- [16]. Deng, J., Dong, W., Socher, R., Li, L. J., Li, K. and Fei-Fei, L., 2009, Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition, 248-255.
- [17]. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S. and Fei-Fei, L., 2015, Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3), 211-252.
- [18]. Rumelhart, D. E., Hinton, G. E. and Williams, R. J., 1986, Learning representations by back-propagating errors. *nature*, 323(6088), 533-536.
- [19]. He, K., Zhang, X., Ren, S. and Sun, J., 2015, Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In Proceedings of the IEEE international conference on computer vision, 1026-1034.
- [20]. Krogh, A. and Hertz, J., 1991, A simple weight decay can improve generalization. *Advances in neural information processing systems*, 4., 950-957.
- [21]. Sutskever, I., Martens, J., Dahl, G. and Hinton, G., 2013, On the importance of initialization and momentum in deep learning. In International conference on machine learning. PMLR., 1139-1147.
- [22]. Glorot, X., Bordes, A. and Bengio, Y., 2011, Deep sparse rectifier neural networks. In Proceedings of the fourteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 315-323
- [23]. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G. and Chintala, S., 2019, Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 8026-8037
- [24]. Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z. and Matthews, I., 2010, The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In 2010 IEEE computer society conference on computer vision and pattern recognition-workshops, 94-101.
- [25]. Pantic, M., Valstar, M., Rademaker, R. and Maat, L., 2005, Web-based database for facial expression analysis. In 2005 IEEE international conference on multimedia and Expo, 5, 6-13.
- [26]. Lyons, M., Akamatsu, S., Kamachi, M. and Gyoba, J., 1998, Coding facial expressions with gabor wavelets. In Proceedings Third IEEE international conference on automatic face and gesture recognition, 200-205.
- [27]. Taini, M., Zhao, G., Li, S. Z. and Pietikainen, M., 2008, Facial expression recognition from near-infrared video sequences. In 2008 19th International Conference on Pattern Recognition, 1-4.
- [28]. Li, S., Deng, W. and Du, J., 2017, Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2852-2861.
- [29]. Mollahosseini, A., Hasani, B. and Mahoor, M. H., 2017, Affectnet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing*, 10(1), 18-31.
- [30]. Xie, L., Wang, J., Wei, Z., Wang, M. and Tian, Q., 2016, Disturblabel: Regularizing cnn on the loss layer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4753-4762.
- [31]. Krizhevsky, A. and Hinton, G., 2009, Learning multiple layers of features from tiny images.
- [32]. He, H. and Garcia, E. A., 2009, Learning from imbalanced data. *IEEE Transactions on knowledge and data engineering*, 21(9), 1263-1284.

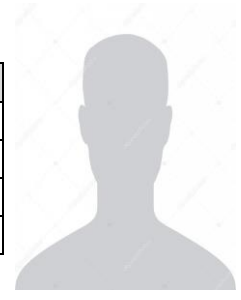
- [33]. Hadsell, R., Chopra, S. and LeCun, Y., 2006, Dimensionality reduction by learning an invariant mapping. In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), 2, 1735-1742.
- [34]. Schroff, F., Kalenichenko, D. and Philbin, J., 2015, Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE conference on computer vision and pattern recognition, 815-823.
- [35]. Meng, Z., Liu, P., Cai, J., Han, S. and Tong, Y., 2017, Identity-aware convolutional neural network for facial expression recognition. In 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), 558-565.
- [36]. Dhall, A., Goecke, R., Lucey, S. and Gedeon, T., 2011, Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark. In 2011 IEEE international conference on computer vision workshops (ICCV workshops), 2106-2112.
- [37]. Guo, Y., Tao, D., Yu, J., Xiong, H., Li, Y. and Tao, D., 2016, Deep neural networks with relativity learning for facial expression recognition. In 2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 1-6.
- [38]. Liu, X., Vijaya Kumar, B. V. K., You, J. and Jia, P., 2017, Adaptive deep metric learning for identity-aware facial expression recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 20-29.
- [39]. Sohn, K., 2016, Improved deep metric learning with multi-class n-pair loss objective. *Advances in neural information processing systems*, 29., 1857-1865.
- [40]. Liu, H., Tian, Y., Yang, Y., Pang, L. and Huang, T., 2016, Deep relative distance learning: Tell the difference between similar vehicles. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2167-2175.
- [41]. Wen, Y., Zhang, K., Li, Z. and Qiao, Y., 2016, A discriminative feature learning approach for deep face recognition. In European conference on computer vision Springer, pp. 499-515.
- [42]. Li, S., Deng, W. and Du, J., 2017, Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. In Proceedings of the IEEE conference on computer vision and pattern recognition, 28, 2852-2861.
- [43]. Cai, J., Meng, Z., Khan, A. S., Li, Z., O'Reilly, J. and Tong, Y., 2018, Island loss for learning discriminative features in facial expression recognition. In 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition, 302-309.
- [44]. Li, Y., Lu, Y., Li, J. and Lu, G., 2019, Separate loss for basic and compound facial expression recognition in the wild. In Asian Conference on Machine Learning, 101, 897-911.
- [45]. Li, Z., Wu, S. and Xiao, G., 2018, Facial expression recognition by multi-scale cnn with regularized center loss. In 2018 24th International Conference on Pattern Recognition, 3384-3389.
- [46]. Ren, S., Cao, X., Wei, Y. and Sun, J., 2014, Face alignment at 3000 fps via regressing local binary features. In Proceedings of the IEEE conference on computer vision and pattern recognition, 1685-1692.
- [47]. Whitehill, J., Wu, T. F., Bergsma, J., Movellan, J. and Ruvolo, P., 2009, Whose vote should count more: Optimal integration of labels from labelers of unknown expertise. *Advances in neural information processing systems*, 22, 2035-2043.
- [48]. Simonyan, K. and Zisserman, A., 2014, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

- [49]. Li, Y., Zeng, J., Shan, S. and Chen, X., 2018, Patch-gated CNN for occlusion-aware facial expression recognition. In 2018 24th International Conference on Pattern Recognition, 2209-2214.
- [50]. Li, Y., Zeng, J., Shan, S. and Chen, X., 2018, Patch-gated CNN for occlusion-aware facial expression recognition. In 2018 24th International Conference on Pattern Recognition, 81, 2209-2214.
- [51]. Li, Y., Zeng, J., Shan, S. and Chen, X., 2018, Occlusion aware facial expression recognition using CNN with attention mechanism. *IEEE Transactions on Image Processing*, 28(5), 2439-2450.
- [52]. Zhao, S., Cai, H., Liu, H., Zhang, J. and Chen, S., 2018, Feature Selection Mechanism in CNNs for Facial Expression Recognition. In *BMVC*, 317.
- [53]. Florea, C., Florea, L., Badea, M. S., Vertan, C. and Racoviteanu, A., 2019, September). Annealed Label Transfer for Face Expression Recognition. In *BMVC*, 104.
- [54]. Barsoum, E., Zhang, C., Ferrer, C. C. and Zhang, Z., 2016, Training deep networks for facial expression recognition with crowd-sourced label distribution. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, 279-283.
- [55]. Kazemi, V. and Sullivan, J. 2014. One millisecond face alignment with an ensemble of regression trees. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1867-1874.
- [56]. Zeng, J., Shan, S. and Chen, X., 2018, Facial expression recognition with inconsistently annotated datasets. In *Proceedings of the European conference on computer vision*, 222-237.
- [57]. Lee, J., Kim, S., Kim, S., Park, J. and Sohn, K., 2019, Context-aware emotion recognition networks. In *Proceedings of the IEEE/CVF international conference on computer vision*, 10143-10152.
- [58]. Dhall, A., Goecke, R., Lucey, S. and Gedeon, T., 2011, Acted facial expressions in the wild database. Australian National University, Canberra, Australia, Technical Report TR-CS-11, 2, 1.
- [59]. He, H. and Garcia, E. A., 2009, Learning from imbalanced data. *IEEE Transactions on knowledge and data engineering*, 21(9), 1263-1284.
- [60]. Liu, W., Wen, Y., Yu, Z. and Yang, M., 2016, Large-margin softmax loss for convolutional neural networks. *arXiv preprint arXiv:1612.02295*.
- [61]. Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B. and Song, L., 2017, Spheroface: Deep hypersphere embedding for face recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 212-220.
- [62]. Wan, W., Zhong, Y., Li, T. and Chen, J., 2018, Rethinking feature distribution for loss functions in image classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 9117-9126.
- [63]. Hayat, M., Khan, S., Zamir, S. W., Shen, J. and Shao, L., 2019, Gaussian affinity for max-margin class imbalanced learning. In *Proceedings of the IEEE/CVF international conference on computer vision*, 6469-6479.
- [64]. He, K., Zhang, X., Ren, S. and Sun, J., 2016, Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770-778.
- [65]. He, K. and Sun, J., 2015, Convolutional neural networks at constrained time cost. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 5353-5360.

- [66]. Ioffe, S. and Szegedy, C., 2015, Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International conference on machine learning, 448-456.
- [67]. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D. and Rabinovich, A., 2015, Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, 1-9.
- [68]. He, K., Girshick, R. and Dollár, P., 2019, Rethinking imagenet pre-training. In Proceedings of the IEEE/CVF International Conference on Computer Vision, 4918-4927.
- [69]. Li, S., Deng, W. and Du, J., 2017, Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2852-2861.
- [70]. Kollias, D. and Zafeiriou, S., 2019, Expression, affect, action unit recognition: Aff-wild2, multi-task learning and arcface. arXiv preprint arXiv:1910.04855.
- [71]. Wang, C., Wang, S. and Liang, G., 2019, Identity-and pose-robust facial expression recognition through adversarial feature learning. In Proceedings of the 27th ACM international conference on multimedia, 238-246.
- [72]. Dauphin, Y. N., Fan, A., Auli, M. and Grangier, D., 2017, Language modeling with gated convolutional networks. In International conference on machine learning, 933-941.
- [73]. Guo, Y., Zhang, L., Hu, Y., He, X. and Gao, J., 2016, Ms-celeb-1m: A dataset and benchmark for large-scale face recognition. In European conference on computer vision, Springer, 87-102.
- [74]. Martinez, A. M., 2019, Context may reveal how you feel. Proceedings of the National Academy of Sciences, 116(15), 7169-7171.
- [75]. Zhao, G., Huang, X., Taini, M., Li, S. Z. and Pietikäinen, M. 2011. Facial expression recognition from near-infrared videos. Image and vision computing, 29(9): 607-619.
- [76]. Lyons, M., Akamatsu, S., Kamachi, M. and Gyoba, J. 1998. Coding facial expressions with gabor wavelets. In Proceedings Third IEEE international conference on automatic face and gesture recognition, 2, 200-205.
- [77]. Zhao, G., Huang, X., Taini, M., Li, S. Z. and Pietikäinen, M. 2011. Facial expression recognition from near-infrared videos. Image and vision computing, 29(9), 607-619.

ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı Soyadı	Salah Haraj Meshal MESHAL
Doğum Yeri	
Doğum Tarihi	
Uyruğu	<input type="checkbox"/> T.C. <input checked="" type="checkbox"/> Diğer



Eğitim Bilgileri	
Lisans	
Üniversite	Al Anbar Üniversitesi
Fakülte	Bilgisayar Fakültesi
Bölümü	Bilim ve Bilgi Teknolojisi
Mezuniyet Yılı	2007

Yüksek Lisans	
Üniversite	Kırşehir Ahi Evran Üniversitesi
Enstitü Adı	Fen Bilimleri Enstitüsü
Anabilim Dalı	İleri Teknolojiler
Programı	Tezli Yüksek Lisans
Mezuniyet Tarihi	