

**T.C
BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI
ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ
TEZLİ YÜKSEK LİSANS PROGRAMI**

**GÖRÜNTÜ İŞLEME TABANLI YÜZ İFADELERİNDEN DUYGU
ANALİZİ**

HAZIRLAYAN

ORHAN EMRE AKSOY

YÜKSEK LİSANS TEZİ

ANKARA 2022

**T.C
BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI
ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ
TEZLİ YÜKSEK LİSANS PROGRAMI**

**GÖRÜNTÜ İŞLEME TABANLI YÜZ İFADELERİNDEN DUYGU
ANALİZİ**

HAZIRLAYAN

ORHAN EMRE AKSOY

YÜKSEK LİSANS TEZİ

TEZ DANIŞMANI

DR. ÖĞR ÜYESİ SELDA GÜNEY

ANKARA 2022

BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Elektrik-Elektronik Mühendisliği Anabilim Dalı Elektrik-Elektronik Mühendisliği Tezli Yüksek Lisans Programı çerçevesinde Orhan Emre Aksoy tarafından hazırlanan bu çalışma, aşağıdaki jüri tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 13/01/2022

Tez Adı: Görüntü İşleme Tabanlı Yüz İfadelerinden Duygu Analizi

Tez Jüri Üyeleri (Unvanı, Adı-Soyadı ve Kurumu)

İmza

Dr. Öğr. Üyesi Atilla Özgür (Başkan), Ankara Yıldırım Beyazıt Üniversitesi	
Dr. Öğr. Üyesi Selda GÜNEY (Danışman), Başkent Üniversitesi	
Prof. Dr. Hamit ERDEM, Başkent Üniversitesi	

ONAY

Prof. Dr. Ömer Faruk ELALDI

Fen Bilimleri Enstitüsü Müdürü

Tarih: .../.../.....

BAŞKENT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
YÜKSEK LİSANS TEZ ÇALIŞMASI ORJİNALLİK RAPORU

Tarih: ... / ... /

Öğrencinin Adı, Soyadı: ORHAN EMRE AKSOY

Öğrencinin Numarası: 21820154

Anabilim Dalı: Elektrik-Elektronik Mühendisliği

Programı: Elektrik-Elektronik Mühendisliği Tezli Yüksek Lisans

Danışmanın Unvanı/Adı, Soyadı: Dr. Öğr. Üyesi SELDA GÜNEY

Tez Başlığı: Görüntü İşleme Tabanlı Yüz İfadelerinden Duygu Analizi

Yukarıda başlığı belirtilen Yüksek Lisans tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam sayfalık kısmına ilişkin, ... / ... / tarihinde şahsım/tez danışmanım tarafından adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı %'dır. Uygulanan filtrelemeler:

1. Kaynakça hariç
2. Alıntılar hariç
3. Beş (5) kelimeden daha az örtüşme içeren metin kısımları hariç

“Başkent Üniversitesi Enstitüleri Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Usul ve Esaslarını” inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrenci İmzası:

ONAY

Tarih: ... / ... /

Öğrenci Danışmanı Unvan, Ad, Soyad, İmza:

.....

TEŞEKKÜRLER

Bu tez çalışmasının gerçekleştirilmesindeki desteklerinden dolayı, adı geçen kişilere tüm içtenliklerimle teşekkür ederim.

Sayın Dr. Öğr. Üyesi Selda Güney’e bilgilerini bizden esirgemeyerek çalışmanın sonuca ulaşmasında her zaman yol gösterici ve yardımcı olduğu için...

Sevgili eşime, aileme ve arkadaşlarıma her daim bana gösterdikleri destek için...



ÖZET

Orhan Emre AKSOY

Görüntü İşleme Tabanlı Yüz İfadelerinden Duygu Analizi

Başkent Üniversitesi Fen Bilimleri Enstitüsü

Elektrik-Elektronik Mühendisliği Anabilim Dalı

2022

Bu tez çalışmasında, gerçek zamanlı görüntülerde insan yüz ifadelerinin sınıflandırılma çalışması yapılmıştır. Gerçek zamanlı tanıma işlemini gerçeklemenin birtakım faydaları bulunmaktadır. Örnek vermek gerekirse; toplu fotoğraflarda ruh halinin analiz edilmesi bu bakımdan ilgi çekici bir örnektir. Bir etkinlik sırasında çekilmiş fotoğraflarda bulunan insanların yüz ifadelerinin algılanması, bu kişilerin genel anlamda ne derece eğlendiğine ilişkin niceliksel veri sağlayabilir. İçerik duyarlı görüntü erişimini bir diğer örnek olarak verebiliriz, söz gelimi bir veri tabanından, sadece şaşırان insanların olduğu fotoğraflara erişilebilir. Bu çalışma kapsamında; yüz ifadelerine ilişkin 7 farklı duygu sınıflanmış olup; bunlar "mutluluk", "üzüntü", "şaşkınlık", "iğrenme", "öfke", "korku" ve "nötr" olarak sıralanmaktadır. Python programlama diliyle yazılan uygulama ile FER2013, JAFFE ve CK+ veri setlerini kullanarak, farklı ifadelerde klasik makine öğrenme yöntemlerinden k-En Yakın Komşuluk ve Destek Vektör Makineleri kullanılırken, derin öğrenme yöntemlerinden aktarımlı öğrenme ile AlexNet, ResNet, DenseNet ve Inception mimarileri kullanılmış, detaylı analizler yapılmıştır. Ayrıca bu çalışmada; klasik makine öğrenme yöntemleri ve derin öğrenme mimarilerini kıyaslarken, 2 farklı uygulama ile gerçek zamanlı ve gerçek zamanlı olmayan uygulamalar da karşılaştırılmıştır. Bu çalışma, en uygun mimariye sahip derin öğrenme tekniklerine dayanan gerçek zamanlı ifade tanıma sistemlerinin, sadece bir yazılım ile bilgisayar donanımı aracılığıyla yüksek doğruluklarla uygulanabileceğini göstermektedir.

ANAHTAR KELİMELER: Yüz İfade Analizi, Duygu Tanıma, Öznitelik çıkartma ve Seçme, Sınıflandırma, Evrişimsel Sinir Ağları

ABSTRACT

Orhan Emre AKSOY

Sentiment Analysis from Face Expressions Based on Image Processing

Başkent University Institute of Science

Electrical and Electronics Engineering Department

2022

In this thesis, the classification study of human facial expressions in real-time images was investigated. Real-time recognition has a number of benefits, for example; analysis of mood in group photos is an interesting example in this regard. The perception of facial expressions of people in photographs taken during an event can provide quantitative data on how much fun these people have in general. Another example is context-aware image access, where, for example, only photos of people who are surprised can be accessed from a database. This scope of work; 7 different emotions related to facial expressions were classified; these are listed as "happiness", "sadness", "surprise", "disgust", "anger", "fear" and "neutral". With the application written in Python programming language, classical machine learning methods k-Nearest Neighborhood and Support Vector Machines and AlexNet, ResNet, DenseNet, Inception architectures are used from deep learning methods and detailed analyzes are made in FER2013, JAFFE and CK+ datasets. In this study, while comparing classical machine learning methods and deep learning architectures, real-time and non-real-time applications were also compared with two different applications. This study shows that real-time expression recognition systems based on deep learning techniques with the most appropriate architecture can be implemented with high accuracy via computer hardware with only one software.

KEYWORDS: Facial Expression Analysis, Mood Recognition, Feature Extraction and Selection, Classification, Convolutional Neural Networks

İÇİNDEKİLER

1. GİRİŞ.....	1
2. LİTERATÜR TARAMASI	4
3. TEMEL KAVRAMLAR.....	7
3.1. Yapay Zekâ	7
3.2. Yapay Sinir Ağı Nedir?.....	8
3.3. Görüntü İşleme Nedir?	10
3.4. Sayısal Görüntü Tanımlamaları	11
3.5. Niteliklendirme	13
3.6. Uzaysal Çözünürlük	13
3.7. Koyuluk	14
3.8. RGB ve Gri Tonları.....	14
3.9. İndekslenmiş Görüntüler	16
3.10. Histogram.....	16
4. YÜZ TANIMA VE YÜZ İFADESİ TANIMA.....	18
4.1. Yüz Bulma	18
4.2. Biyometrik Teknolojiler ve Tanıma	20
4.3. Yüz Tanıma Tarihçesi.....	21
4.4. Yüz tanımanın yaygınlaşmasındaki faktörler	21
4.5. Yüz İfadesi Tanıma	22
4.6. Yüz Bulma Algoritmaları	23
4.6.1. Bilgi Tabanlı Yöntemler	23
4.6.2. Değişmeyen Özellik Yaklaşımları	24
4.6.3. Şablon Eşlemeli Yöntem	25
4.6.4. Görünüm Tabanlı Yöntemler.....	26
4.7. Yüz İfadelerinin Değerlendirilmesi	26

4.8. Yüz İfadesi Tanıma Sistemleri	28
5. YÜZ ÖZNİTELİKLERİNİN ÇIKARTILMASI VE SEÇİLMESİ.....	30
5.1. Gabor Filtresi.....	30
5.2. Yerel İkili Örnekler Operatörü.....	31
5.3. Seyreltik Kodlama	33
5.4. Özyüz (Eigenface – TBA).....	34
5.5. Öznitelik Seçme	38
6. YÜZ İFADESİ SINIFLANDIRMA	40
6.1. Haar Dalgacıkları	40
6.2. Otomatik Nirengi Noktası Bulma	41
6.3. k-NN – En Yakın Komşuluk Algoritması	43
6.4. SVM – Destek Vektör Makineleri.....	45
6.5. AlexNet	45
6.6. ResNet.....	46
6.7. DenseNet.....	47
6.8. InceptionV3	48
6.9. Aktarımlı Öğrenme	49
7. UYGULAMA.....	51
7.1. Geliştirme Ortamı	51
7.2. Veri Seti	51
7.2.1. Veri seti – FER2013.....	52
7.2.2. Veri seti – JAFFE	53
7.2.3. Veri seti – CK +	54
7.3. Uygulamada Kullanılan Öznitelik Seçme Yöntemi.....	54
7.4. Uygulama Sonuçları	56
7.5. Uygulama Kısıtlamaları.....	65
8. SONUÇ.....	66

KAYNAKLAR.....	67
-----------------------	-----------



TABLÖLAR LİSTESİ

Tablo 1: Renklerin RGB değeri [8]	15
Tablo 2: Biyometrik tanıma tekniklerin karşılaştırma [16]	20
Tablo 3: Veri setleri karşılaştırması.....	52
Tablo 4: Veri setindeki imge türlerin sayı ve oranları.....	53
Tablo 5: Hiper parametre değeri	56
Tablo 6: Literatür Karşılaştırması.....	57
Tablo 7: Doğruluk oranları ve eğitilme süreleri	58
Tablo 8: İfadeye göre doğruluk oranları.....	63



ŞEKİLLER LİSTESİ

Şekil 3-1: Yapay sinir ağı yapısı [2].....	9
Şekil 3-2: Filtrelenmiş görüntü örneği [8].....	11
Şekil 3-3: Analog sinyalin, dijital sinyale dönüşümünde örneklerin alınması [8]	12
Şekil 3-4: Sayısal görüntü gösterimi [8].....	12
Şekil 3-5: “n” değerine göre nitelik değişimi [9]	13
Şekil 3-6: Uzaysal çözünürlük değişimi [9]	13
Şekil 3-7: Koyuluk Kavramı [8].....	14
Şekil 3-8: Aynı renkli 9 dikdörtgen [8]	14
Şekil 3-9: Renkli görüntünün oluşumu [8].....	14
Şekil 3-10: Farklı farklı gri seviyelerinin gösterimi [8].....	16
Şekil 3-11: Histogram Örnekleri [36].....	17
Şekil 4-1: Yüz oranları [13].....	19
Şekil 4-2: Erkek ve kadında alın / burun / çene uzunluğu oranları [13].....	19
Şekil 4-3: Farklı kişilere ait yüz ifadeleri [2]	27
Şekil 4-4: Farklı yüz ifadelerinin birleşimi [2].....	28
Şekil 5-1: Gabor Filtre Sistemi.....	31
Şekil 5-2: YİÖ operatörü işlemleri [7]	32
Şekil 5-3: YİÖ Histogramı	33
Şekil 5-4: Temel Bileşen Analizi Doğrusu [21]	35
Şekil 5-5: En büyük özdeğerler için elde edilen özyüzler [21]	36
Şekil 5-6: Özyüz algoritması [40]	38
Şekil 6-1: Haar filtreler ve uygulanması [38].....	40
Şekil 6-2: Adaboost basamaklandırılmış sınıflandırıcı	40
Şekil 6-3: Bezier eğrilerine örnekler [39].....	41
Şekil 6-4: (a) Bezier hacim modeli, (b) Hareket birimleri [39].....	42
Şekil 6-5: K-NN algoritması [23].....	44
Şekil 6-6: K-NN Adımları	44
Şekil 6-7: SVM algoritması [36]	45
Şekil 6-8: AlexNet mimarisi [44]	46
Şekil 6-9: ResNet mimarisi [45].....	47
Şekil 6-10: DenseNet mimarisi [47].....	48

Şekil 6-11: InceptionV3 mimarisi [49].....	48
Şekil 6-12: Aktarımlı öğrenme ve model uyarlama [56].....	50
Şekil 7-1: FER2013 veri setinden bazı örnekler [34].....	52
Şekil 7-2: JAFFE veri setinden bazı örnekler [35].....	53
Şekil 7-3: CK+ veri setinden bazı örnekler [42].....	54
Şekil 7-4: HOG örnek yüz fotoğrafı [36]	55
Şekil 7-5: Duygu tanıma sistemi akışı.....	56
Şekil 7-6: Uygulama-1 testinde kullanılan yüzler (1-5)	60
Şekil 7-7: Uygulama-1 testinde kullanılan yüzler (6-10)	61
Şekil 7-8: FER2013 – SVM sınıflandırması.....	62
Şekil 7-9: FER2013 – ResNet sınıflandırması	62
Şekil 7-10: Doğruluk grafiği.....	64

SİMGELER VE KISALTMALAR LİSTESİ

2B	2 Boyutlu
3B	3 Boyutlu
AI	Yapay Zekâ
AAM	Bağımsız Aktif Görünüm Modeli
ANN	Yapay Sinir Ağları
ART	Adaptif Rezonans Teorisi
AU	Action Unit
BRISK	Binary Robust Invariant Scalable Keypoints
BPSO	Binary Particle Swarm Optimization
CK+	Cohn-Kanade Veri Seti
CNN	Derin Konvolüsyonel Sinir Ağları
CPU	Central Processor Unit
DAA	Doğrusal Ayırtaç Analizi
DT	Dense Trajectories
DVM	Destek Vektör Makineleri
DYİÖ	Düzenli Yerel İkili Örnekler
ELM	Extreme Learning Machines
FID	Yüz İmge Tanımlayıcıları
GPU	Graphical Processor Unit
GRNN	Genel Regresyon Ağları
HMM	Hidden Markov Model
HOG-TOP	Yönlü Gradyan Histogramları
ICA	Bağımsız Bileşke Analizi
IFD	Indian Face Database
IMoFA	Incremental Mixtures of Factor Analysers
JAFFE	Japanese Female Facial Expression
KDEF	Karolinska Directed Emotional Faces
k-NN	En Yakın Komşu
LBP	Yerel İkili Desen
LL	Logarithm-Laplace
LTI	Linear Time Invariant
MLP	Multilayer Perceptron Classifier
MUTINF	Mutual Information Based
PBVD	Piecewise Bezier Volume Deformation
PNN	Probabilistik Ağlar
RaFD	Radboud Faces Database
RBF	Radyal Tabanlı Fonksiyon
RGB	Red – Green – Blue
SOM	Self Organizing Feature Maps
STTM	Spatio-Temporal Texture Map
SVM	Support Vector Machines
TAFED	Taiwanese Facial Expression Database
TBA	Temel Bileşenler Analizi
XOR	Exclusively-OR
YFN	Yerel Faz Nicelemesi
YİÖ	Yerel İkili Örnekler
YSA	Yapay Sinir Ağları

1. GİRİŞ

Yüz tabanlı, bilgisayar ile görüntü uygulamaları kapsamında karşılaşılmakta olan en büyük ve çözülmesi güç problemlerden biri yüz görüntülerindeki ifade çeşitliliğidir. Yüz ifadelerine karşı dayanıklılık konusunda bugüne dek çok sayıda sistem önerisi ileri sürülmüşse de söz konusu problem hala geçerliliğini sürdürmektedir. Erişimi kolay ve ayırt ediciliği yüksek olan yüz, uzun yıllardır araştırmacılara konu olagelmıştır. Yüz teşhis ve tanıma çalışmaları, yalnızca güvenlik gözetim sistemlerinin gelişmesine değil, nöroloji ve psikoloji bilimlerinin gelişimine de önemli katkılar sağlamıştır. Yüz görüntülerinin incelenmesine ilişkin olumsuzluklar; ağırlıklı olarak kafa duruşu, aydınlatma koşulları ve yüz ifadeleri ile alakalıdır. Bu olumsuzlukların giderilmesini hedefleyen pek çok çalışma gerçekleştirilmişse de problemlerin çözüme kavuşturulduğu söylenemez. Bugün 3B (üç boyutlu) yüz bilgisine erişim ve bu bilgiyi işlemek son derece kolay hale gelmiştir. Bu durum verinin doğası gereği değişim olmayan poz ve aydınlatmadan kaynaklı olarak çözüme önemli ölçüde katkıda bulunmuştur. Ancak yüz ifadeleri ile ilgili buna benzer bir gelişmeden bahsedilememektedir.

Yüz ifadelerinin tanınması;

- Güvenlik ve görsel gözetim ve güvenlik alanlarında (yalan algılama, akıllı ortamlar),
- Tıbbi teşhis koymada,
- Duyguya dair araştırmalarda (davranış bilimleri, nöroloji, psikiyatri),
- Yasa uygulamada,
- Eğitimde sıklıkla kullanılmaktadır.

Kişilerin duygu durumlarına, bilişsel aktivitelerine ve ilgi yönelimlerine dair önemli ipuçları sunan yüz ifadeleri; yüz yüze iletişimi etki eden başlıca unsurlar arasında yer alır. Bu da ifadeler kadar, makine-insan etkileşimini de son derece önemli kılar. Yüz ifadelerini sınıflandırma, bahsi geçen değişimler karşısında alınabilecek önlemler bakımından da oldukça önemlidir. Konu güvenlik uygulamaları bakımından ele alınacak olursa, 2B (iki boyutlu) yüz görüntülerinin barındırdığı ifadeler doğrultusunda sınıflandırılmasıyla, yüz teşhis ve tanıma sistemlerinin daha yüksek doğrulukta sonuçlar vermesi mümkün olabilecektir.

Yüz ifadeleri bireylerin hisleri, anlayış biçimleri, karakterleri ve psikolojik durumları ile ilgili somut belirtiler olarak kabul edilebilecek olup, eş zamanlı olarak insanların birbirleri ile ilişkilerinde de önemli bir araçtır. Mehrabian, dinleyicilerin yüz ifadelerinden önemli ölçüde etkilendiğini ifade etmektedir. Darwin (1872) insanların ve hayvanların doğalarında doğuştan kimi duygular barındırdığını ve söz konusu duyguların yüz ifadeleri olarak dışarıya yansıdığını ileri sürmüştür [3]. Yüz ifade analizine ilişkin bu ilk çalışmanın ardından 1971 senesinde, Ekman ve Friesen tarafından, 6 temel duygunun var olduğu ve her birinin farklı yüz ifadesi ile yansıtıldığı ortaya konmuş olup; bunlar “öfke”, “iğrenme”, “korku”, “mutluluk”, “üzüntü” ve “şaşkınlık” şeklinde sıralanmıştır [3].

Sabit görüntüler ve görüntü dizileri kullanımı, yüz ifade analizinin yalnızca psikoloji biliminin sınırlarında kalan bir araştırma konusu olmaktan çıkmasını sağlamıştır. Bunun beraberinde görüntülerde yüz bulma ve yüz tanıma gibi alanlarda kaydedilen gelişmeler, yüz ifade analizine ilişkin araştırmaların sayıca artmasına yol açmıştır. Yüz ifadeleri, biçimsel olarak yüzde üzerindeki belli bölgelerin biçim değiştirmesiyle belirlenmektedir. Özellikle yüz kaslarının kasılmasıyla yüz derisi, kaşlar, göz kapakları, burun ve dudaklar gibi yüz özniteliklerinin geçici değişimlere uğraması sonucunda oluşan, birkaç saniyelik bu geçici değişimlerin yerinin, yoğunluğunun ve dinamiğinin belirlenmesi, yüz ifadelerini tespit etme ve sınıflandırma bakımından başlıca öneme sahiptir. Söz konusu değişimlerin boyutu ve süresi kişi, yaş, ırk, cinsiyet doğrultusunda değişebilmekte olup, problemin çok katmanlaştırmakta ve çözümünü zorlaştırmaktadır. Yüz ifade analizine ilişkin çalışmalarda ortaya çıkan 3 temel problem, Pantic ve Rothkrantz tarafından belirlenmiştir;

- Görüntüde yer alan yüz bölümlerinin tespit edilmesi,
- Yüz ifadelerinin özniteliklerinin çıkarılması,
- Yüz ifadeleri sınıflandırma.

Literatürdeki çalışmaların da yukarıda yer verilen üç temel probleme çözüm getirmeye odaklandığı görülmektedir. Kotsia ve Pitas, görüntü dizileri üzerinden yüz ifade analizinde, destek vektör makinelerinden yararlanmıştır [3]. Yürüttükleri çalışma kapsamında, görüntü dizisindeki ilk görüntüye yerleştirilmiş noktalar referans alınarak, takip eden görüntülerle olan yer değiştirmeye göre ifade sınıflandırması yapılmıştır. Altı farklı yüz ifadesini sınıflandırmayı amaçlayan Otsuka, Hidden Markov Model (HMM)'ini temel alan sınıflandırıcılarından yararlanmıştır. Bartlett vd., ifade analizinde Gabor özniteliklerinin kullanıldığı çalışmaları kapsamında, kendilerinin geliştirdiği AdaSVM

metodunu önerirken; Cohn and Kanade yüz ifade veri setinde başarı oranı son derece yüksek sonuçlara ulaşmıştır [3]. Pantic ve Rothkrantz'ın tanımlamış olduğu üç temel problemten biri olan yüz ifadelerine ait özniteliklerin çıkarılmasının bir adımı olan “öznitelik seçilmesi”, ifadelerin sınıflandırmasına yönelik çalışmaların devamlı olarak yürütüldüğü bir araştırma konusudur.

Donato'nun yaklaşımında, başlarda yüz tanımaya yönelik tekniklere başvurulmuştur [3]. Yüz görüntüsünde ifadenin beraberinde kimlik bilgisi de yer almakta olup; detaylı bir ifade analizinin gerçekleştirilmesi için kimlik bilgisinin görüntüden çıkarılması faydalı olmaktadır. Burada, videodan görüntü dizilerinin alınmasıyla, fark görüntü yaratılmış ve optik akış tahminiyle ifadeler farklı uzaysal analizler doğrultusunda sınıflandırılmıştır. Benzer bir metot, yine Park'ın çalışmalarında da geliştirilmiştir [3]. Yine Pantic ve Rothkrantz'ın yaklaşımında ise görüntüye nokta-tabanlı yüz modellerinin oturtulması ile yüz hareketi birimleri (Action Units) tespit edilmiş olup; duygusal ifadelerin sınıflandırılması için bu hareketlere başvurulmuştur [3]. Colmenarez'in çalışmalarında da yüz modellemesi için öznitelik noktalarından faydalanılmıştır [3]. Öte yandan, öznitelik noktalarının doğruluğu konusunda söz konusu olan aşırı bağımlılığın, bu tür yaklaşımlar için dezavantaj teşkil ettiği söylenmelidir.

Yüz ifadesi takibi ve ifade sınıflama üzerine çok sayıda araştırma yürütülmüştür. Bu araştırmalar genel olarak yüz bulma, yüz özniteliği çıkarımı ve yüz ifadesi tanıma olmak üzere üç temel adımdan meydana gelmektedir. Yüz bölgesi belirlemede ten rengi, yüzün geometrik yapısı gibi özellikler dikkate alınmaktadır. Bilgisayarlı görüde gerçek dünyanın 3 boyutlu, görüntü işlemenin ise 2 boyutlu olması beraberinde kimi zorluklar getirmektedir. Yüz bulma problemini bir bilgisayar modeli ile gidermek son derece güçtür. Bundan dolayı, doğru ve güvenilir yüz ifadesi tanıma için yüz bulma işlemi de doğru olarak yapılmalıdır.

Yapılan çalışmalarda, seçili özniteliklerin ifade analizine ilişkin en fazla bilgi taşıyanların belirlenmesine yönelik 2 öznitelik seçme yaklaşımına başvurulmuştur [15]. Çoklu sınıflandırma probleminin, ikili sınıflandırma problemine dönüştürülmesi ile öznitelik seçiminin uygulanması ve sonrasında seçilen özniteliklerin birleştirilmesi ile sınıflandırma için kullanılması, yaklaşımların ortak özellikleri olarak sıralanabilir.

2. LİTERATÜR TARAMASI

Lucey vd., 2010, aktif şekil modelini kullanarak yüzdeki nirengi noktalarını tespit etmişlerdir [24]. Daha sonra bu noktaları benzerlik – normalize edilmiş şekilde (SPTS) ve kanonik görünüm (CAPP) özniteliklerini hesaplamada kullanmışlardır. Sınıflandırma için DVM kullanılmıştır. Deneyler CK+ veri setinde gerçekleştirilmiştir.

Arı vd., 2011, duygu problemini irdeleyip, gerçek-zamanlı olarak çalışabilen bir yöntem önermişlerdir [24]. Bu yöntemde nötr, şaşkınlık, mutluluk, üzülm, kızgınlık, korku ve iğrenme ifade tespit edilebilmektedir. Önerdikleri yöntemle, çok çözünürlüklü aktif şekil modelini temel alan takipçiyle nirengi noktalarının takip edilmesi sağlanmış; sonrasında ise, sınıflandırma, nokta konumlarından ve yüzdeki belirli bölgelerde söz konusu olan değişimler üzerinden elde edilen üst seviye öznitelik vektörleri ile duygusal ifadelerin öznitelik vektörleri arasındaki uzaklık dikkate alınarak yapılmıştır. Çalışma kapsamında gerçekleştirilen deneylere beş özne dahil edilmiştir.

Lin Zhong vd., 2012, çalışmalarında farklı duygularda yüzdeki ortak ve özel bilgileri keşfetmeyi amaçlamışlardır [24]. Sadece yüzdeki belli parçaların aktif olarak hareket etmekte olduğu görülmüştür. Bu parçaları bulabilmek için 2 seviyeli çok görevli seyrek öğrenme (MTSL) metodunu önermişlerdir. Bu iki seviyeli görevde her duyguda ayırt edici yüzdeki parçaları bulmaya çalışmışlardır. Bu öğrenilen öznitelikler ve eğitilmiş MTSL, CK+, GEMEP-FERA ve MMI veri setlerinde denemiştir.

Huang vd., 2012, hızlandırılmış gürbüz özniteliklerden (SURF) yararlanarak ulaştıkları öznitelik vektörlerinin normalize edilmesiyle, olasılık dağılım fonksiyonlarını hesaplamıştır [28]. Ulaşılan vektörlerin arasındaki uzaklığın Kullback-Leiber ıraksamasıyla hesaplandığı ve sınıflandırmanın ağırlıklı çoğunluk oylamasıyla gerçekleştirildiği çalışmanın deneyleri, Japanese Female Facial Expression (JAFFE) veri setinde gerçekleştirilmiştir.

Akakin ve Sankur, 2012, işaret yörüngelerinin etkin bir tespitiyle bulunduğu çalışmalarında, öznitelik çıkarımını bu yörüngelerden Bağımsız Bileşke Analizi (ICA)’ni kullanarak gerçekleştirmiştir [28]. Aynı zamanda, zaman-uzamsal yüz küpüne hizalanmış ve normalleştiren yüzlerin istiflenmesi ile yüzdeki duygunun evriminin de yakalandığı çalışmada; bu prizmadan ve ICA özniteliklerinden duygu tanımlayıcıları ortaya

ıkarılmıřtır. 6 temel duyguda sınıflandırma yapılmıř olup; sınıflandırmada birden ok yntemden yararlanılarak gerekleřtirilmiřtir. alıřmanın deneyleri CK+ veri setinde gerekleřtirilmiřtir.

Sarawagi ve Arya, 2013 tarihli arařtırmalarında, renk normalizasyonunu vurgulayan ve yz zniteliklerinin ıkarılmasında etkin znitelik ıkarma yntemi Yerel İkili Desen (LBP)'yi kullanmıřlardır [28]. Sınıflandırmanın DVM kullanılarak gerekleřtirildiėi alıřmanın deneyleri, kendilerinin oluřturduėu toplam 4 kadın 6 erkek olmak zere 10 Hintli zneden meydana gelen veri setinde gerekleřtirilmiřtir.

Hong ve Cong, 2014, alıřmalarında aktif řekil modeli (ASM) ile znitelik ıkarımı yapmıřlardır [28]. Sınıflandırma iin DVM zerinde iyileřtirme adımları ieren kme DVM adı altında bir yntem nerilmiřtir. Bunlara ek olarak, kaba kme teorisinde, elde edilen zniteliklerde seim ve sınıflandırma yapabilmek iin boyut indirgeme algoritması tanıtılmıřlardır. Alakasız ve yanlış olan znitelikler filtrelenmiřtir. Sınıflandırma DVM algoritması zerinde yapılmıřtır. Deneyler JAFFE veri setinde gerekleřtirilmiřtir.

Huang vd., 2014, kt hizalama ve byk varyasyonlarının duygu tanımda iki nemli problem olduėuna deėinmiřlerdir [3]. Bu yzden bu iki problemi zezebilmek iin kanonik korelasyon metodunu tekrar gzden geirmiřlerdir. Sebebi ise bu iki problemten kaynaklanan duygu tanıma problemini, yanlış bulmaya ve yanlış hizalamalara karřı grbz bir hale getirmektir. İlk bařta, znitelik olarak grnm tabanlı olan LBP'yi sunmıřlardır; daha sonra ise video klibin kanonik alt-uzayını oluřturmak iin geici ortogonal blgesi korunmuř izdřm nermiřlerdir. Bylece duygu tanımda hareketten kaynaklanan deėiřimleri yalayabilmek amalanmıřtır. Deneyler CK+ ve MAHNOB-HCI veri setlerinde gerekleřtirilmiřtir.

iėdem Turan ve Kim Man Lam (2014) znitelik ıkarımında, aėız ve gz blgelerinden Yerel Faz Nicelemesi (YFN)'nden ve Piramit HOG (PHOG) tanımlayıcılarından yararlanmıřtır [14]. Kanonik Korelasyon Analizine dayalı bir znitelik birleřtirme yapmıřlardır. Deneyler DVM ile Geniřletilmiř Cohn-Kanade (CK+) veri seti zerinde gerekleřtirilmiřtir.

Chen ve Chen, 2015, grnm tabanlı ve geometrik znitelikleri kullanmıřlardır [29]. Grnm tabanlı znitelikleri  Dikey Dzlem'den gelen Ynl Gradyan Histogramları

(HOG-TOP) temsil ederken, geometrik öznitelikleriye kendi içerisinde ikiye bölünerek katı olan ve olmayan değişimli öznitelikler temsil etmiştir. Bahsi geçen 2 öznitelik çeşidinin optimal kombinasyonlarının elde edilebilmesi adına Çoklu-Kernel Öğrenme yöntemine başvurulmuştur. Son olarak bu çoklu-kerneller ile DVM’de eğitilip test edilmiştir. Deneyler Genişletilmiş Cohn-Kanade (CK+) veri setinde gerçekleştirilmiştir.

Zhu vd., 2015, çalışmalarında, çapraz veri seti kullanarak duygu tanıma için transfer öğrenme önermişlerdir [30]. İki seviyeli bir eğitim süreci içerir. İlki kaynak veriden bilgiyi öğrenmek, ikincisi ise öğrenilen bilgiyi hedef veriye uyarlamaktır. Bu yaklaşım da öznitelikler gabor filtrelerinden elde edilen öznitelikler, regresyon ağaç Sınıflandırıcıları, Adaboost algoritması ve DVM-tabanlı metotlar baz alınarak oluşturulmuş olup; çalışmanın deneyleri JAFFE, FER2013 veri setleri üzerinde gerçekleştirilmiştir.

Shan vd., 2017, otomatik duygu tanıma için Derin Konvolüsyonel Sinir Ağları (CNN) üzerinde durmuşlardır [3]. Sebebi CNN’in daha derinlerdeki özniteliklere sahip bir yapıya sahip olmasıdır. Sistem dört modülden meydana gelmekte olup; bu modüller sırayla giriş, ön-işlem, tanıma ve çıkış modülleridir. Deneylerinde JAFFE ve CK+ veri setlerine başvurmuşlar; aynı zamanda karşılaştırma için en yakın komşu (K-NN) algoritmasından yararlanmışlardır.

Ding vd., 2017, ilk önce en tepe duygu çerçevesini yakalamak için çift LBP’yi önermişler [11]. Önerilen DLBP daha küçük boyutta ve başarılı bir şekilde tanıma zamanı düşürülmüştür. Bununla birlikte, LBP’deki ışık varyasyonlarını elimine edebilmek için, logarithm-laplace (LL) önerilmiştir. Böylece ışık varyasyonları karşısında sistemin daha gürbüz olmasını amaçlamışlardır. Nihayet, LBP-tabanlı Taylor öznitelik deseni (TFP) ve genişletilmiş Taylor’ı önermişlerdir. Sebebi Taylor öznitelik haritasından etkili olan yüzsel öznitelikleri elde etmektir. Deneyler FER2013 ve CK+ veri setinde gerçekleştirilmiştir.

Ruiz ve arkadaşları (2017), araştırmalarında yüz ifadesi ile ifadeye karşılık gelen görsel öznitelikleri arasındaki ilişkiyi modelleyen Yapay Sinir Ağlarını (ANN) önermişlerdir [3]. Sınıflandırma için de yine ANN kullanılmıştır. 7 temel duygu içermektedir. İmge veri setinden elde edilen öznitelik vektörleri üzerine Principle Component Analysis (PCA) uygulanarak boyut küçültme yapılmıştır. Deneyler FER2013 veri setinde gerçekleştirilmiştir.

3. TEMEL KAVRAMLAR

3.1. Yapay Zekâ

Zekânın sözlük anlamı itibariyle insanın sahip olduğu düşünebilme, kavrama, akıl yürütebilme, nesnel gerçeklikleri algılayabilme, yargılama ve sonuca varma kabiliyetlerinin tümünü kapsayan bir kavramdır. Soyutlamak, öğrenmek ve yeni durumlara uyum sağlamak gibi beceriler de yine zekanın kapsamına girmektedir. D. Lenat ve E. Feigenbaum tarafından yapılan tanımlar doğrultusunda, zekâ, “kompleks bir problemin çözümü için gereken bilgilerin toplanması ve birleştirilmesi” veya “kompleks bir problemin, çözüm arama alanının daraltılması yoluyla kısa yoldan çözülmesi” olarak ifade edilebilir. Amaç, bir problemin, etkin biçimde ve kısa yoldan çözülmesidir. Bahsi geçen becerileri taşıyan fakat organik olmayan sistemlerin sahip olduğu zekâsı ise yapay zekâ ise olarak ifade edilmektedir. Yapay zekâ; 1990’lı yılların başlarında, Luger ve Stubblefield tarafından “zeki davranışların otomasyonu ile ilişkili bilgisayar bilimleri dalı” olarak ifade edilmiştir [2].

Beynin yapısına bakıldığında zaman, yaklaşık bir buçuk kilo ağırlığında olduğu söylenebilir. Bu organ, 60 yıllık bir ömür göz önünde bulundurulduğunda, 600 birim/saniyelik hafıza kaydetme ve işleyerek programlama kapasitesi taşır; ki bu, 3.600 bit/dakika, 2.160.000 bit/saat ve 51.840.000 bit/gün büyüklüğünde bilgi anlamına gelmektedir. Yapay zekânın ne olduğu ile ilgili literatürde çeşitli tanımlar mevcuttur. Grey Walter’ın araştırmalarına göz atıldığında, insan beyni ile benzeşen bir makina oluşturulabilmesinin 300 trilyon doların üzerinde bir bütçe gerektirdiği bilgisine ulaşılmaktadır. Diğer taraftan, bu tür bir makineyi hayata geçirebilmek, 1 trilyon wattlık bir elektrik enerjisi ihtiyacı ortaya çıkarmaktadır.

Yapay zekânın tarihsel gelişimine bakıldığında zaman iki önemli görüşten bahsedilebilir. Bunlardan ilkinin, 1300’de Llull ortaya atmış olup; bir diğer görüş ise Nazilere yönelik şifre kırma makinesi olan Enigma çözümü ile Alan Turing’e aittir. Dolayısıyla, makinelerin düşünmesinin, bu görüşlerin ortak noktası olduğu ifade edilebilir.

1950 Alan Turing: Yapay zekâ ilk kez Turing tarafından temellendirilmiştir. Turing’in ürettiği çözümler “makine düşüncesi” gibi olguların gündeme gelmesine sebebiyet vermiştir. Bir makinenin bilincinin bulunup bulunmadığı, Turing’in odaklanmak istediği en önemli noktadır. Konu “bilinç” olduğunda, böyle bir makinenin duygusal tepkiler vermesi,

etrafında olanları anlamlandırması ve bu olanlar karşısında tepkiler üretmesi beklenmektedir. Dil ise, bütün bunları açığa çıkaran aracıdır. Bundan dolayı Turing, bahsi geçen testi dile odaklanarak gerçekleştirmiştir. Testin düzeneğinde; hesaplama yapan bir makine olarak bilgisayar, görülmeyecek bir noktada gönüllü bir insan ile bir arada yer alır. Testi gerçekleştiren sorgulayıcıysa, sadece soru sorar ve hangisinin bilgisayar olduğuna karar vermeye çalışır.

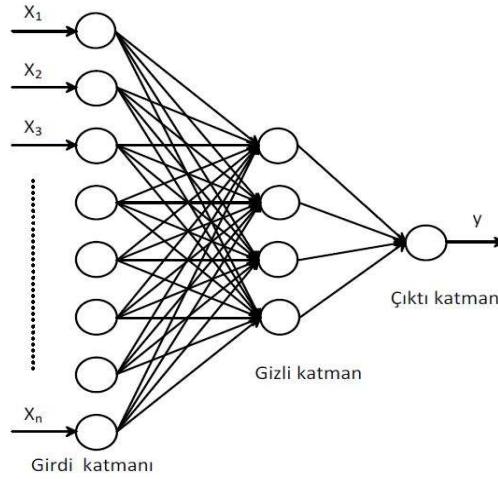
Logic Theorist: 1950'lerde geliştirilmiş olan Logic Theorist programı, düşüncenin dijitalleştirilmesinin temelini atmıştır. Hipotetik senaryolar kullanan bu programda, düşünceleri mantıksal yaklaşımlarla ele alıyordu. Moore Yasası: Moore Yasası'nın işlem kapasitesinin artırılmasıyla, IBM'in geliştirdiği yapay zekâ satranç oyunu dünya satranç şampiyonuna galip gelmişti.

Zeki yazılımların geliştirilmesini konu alan çalışmalar yıllar içerisinde daha da önemli hale gelmiştir. Yazılımların öğrenme becerisi, tecrübelerine dayalı olarak davranışlarının değişmesini mümkün kılmaktadır. Bu yazılımları içeren donanım destekli sistemlerin (öğrenebilen yazılımlar ve robotlar gibi) kullanıldığı uygulamalarla, günlük hayatta gitgide daha fazla karşılaşılmaktadır. Makinelerin öğrenmesini sağlayan en güçlü yöntemlerden biri de yapay sinir ağlarıdır.

3.2. Yapay Sinir Ağı Nedir?

Yapay sinir ağları sunulmuş örnekler aracılığıyla bağlantılı bilgileri toplayarak, bu bilgiler ışığında genellemeler yapar ve sonrasında daha önce hiç karşılaşmadığı örneklerle karşılaştırıldığında, öğrenmiş olduğu bilgilerden yola çıkarak söz konusu örneklerle kararlar verir. Bu şekilde öğrenme ve genelleme özelliklerine sahip olmaları sebebiyle, Yapay Sinir Ağları çağımızda bilimsel anlamda uygulanabilirlik imkânı bulmuştur. Diğer bir ifade ile yapay sinir ağlarını tanımlarsak; insan beyninden ilham alarak geliştirilen, çoğunlukla bağlantılar aracılığıyla birbirlerine bağlanabilen ve her birinin kendi belleği bulunan işlem elemanlarının oluşturduğu paralel ve dağıtılmış yapılardır. Kısaca, yapay sinir ağlarının, biyolojik sinir ağlarını taklit eden bilgisayar programları olduğu söylenebilir [3].

Yapay Sinir Ağları (YSA) ya da daha kısa bir ifadeyle Sinir Ağları (Neural Networks-NN) insan beyninden yola çıkılarak geliştirilmiştir. Bunlar, ağırlıklı bağlantılar aracılığıyla birbirine eklenen işlem elemanlarının oluşturduğu paralel yapılardır.



Şekil 3-1: Yapay sinir ağı yapısı [2]

YSA, insan beynine özgü bir özellik olan öğrenme yoluyla, yeni bilgi üretme ve bilgi oluşturma, keşfetme gibi becerileri yardım almaksızın otomatik bir şekilde gerçekleştirme amaçlı geliştirilmiş bilgisayar sistemleri olup; bu becerilerin konvansiyonel programlama yöntemleriyle yaratılması son derece güçlü. Dolayısıyla, YSA bilim dalı, programlamanın oldukça zor yahut imkânsız olduğu durumlara yönelik geliştirilen, uyarlanabilir bilgi işlemeyle ilişkili bilgisayar-tabanlı bir bilim dalıdır.

YSA, farklı durumlar veya olaylarla ilgili örnekleri inceleyerek ve bunlardan hareket ederek, söz konusu olay veya duruma ilişkin genellemelere varabilmekte, bilgi toplayabilmekte ve sonrasında daha önce hiç görmemiş olduğu örneklerle karşılaştığında öğrenmiş olduğu bilgilerden yararlanarak ilgili örneklerle dair karara varabilmektedir. 1990'lardan günümüze dek bilgisayar öğrenmesini sağlamakta olan YSA teknolojilerinde, önemli ölçüde ilerleme kaydedilmiştir. İnsan beynine özgü özelliklerinden yol a çıkılarak geliştirilmiş bilgisayar sistemleri olan YSA, yeni gelişmelere öncülük ederken insan beynine dair araştırma yapılan alanlara da mühim katkılarda bulunmuştur. Aynı şekilde, YSA teknolojisinin, insanın doğayı keşfetme ve taklit etme çabasının son ürünlerinden olduğu söylenebilir. YSA, basit biyolojik sinir sisteminin çalışma prensibinin benzetiminin yapılmasına yönelik tasarlanmış programlar olup, sinir hücreleri içermektedir ve bu sinir hücrelerinin farklı biçimlerde birbirine bağlanmalarıyla ağ ortaya çıkmaktadır. Bahsi geçen ağlar öğrenebilmekte, hafızaya alabilmekte ve veriler arasında mevcut olan ilişkileri ortaya koyabilmektedir. Yani, YSA, çözümünde bir insanda doğal olarak bulunan düşünme ve gözlemlene yeteneklerine ihtiyaç duyulan problemlere çözüm üretebilmektedir. Bunu

başarabilmesinin temelinde, insanda doğuştan bulunan yaşayarak ya da deneyerek öğrenme yeteneği yatmaktadır.

1890’larda YSA üzerinde düşüncelerin ilk örnekleri ortaya konmaya başlanmış, insan beynindeki yapı ve fonksiyonlar üzerine yazılar yayınlanmıştır. 20. Yüzyıl’ın birinci yarısında başlayıp, bugüne dek büyük bir ivmeyle devam etmiş olan YSA çalışmaları; 1970 öncesi ve sonrası olmak üzere iki bölümde incelenebilir. YSA için dönüm noktası olarak kabul edilen 1970’lerde aşılamaz olduğu düşünülen birçok sorunun üstesinden gelinmiştir.

1943 senesinde, sinir hekimi Warren McCulloch ve matematikçi Walter Pitts yapay sinir ağı modellerinin ilkinin geliştirmiştir. McCulloch ve Pitts, insan beyninin hesaplama becerisinden yola çıkarak, elektrik devreleri ile, karmaşık olmayan bir sinir ağı modellemiştir. Aynı yıl, yapay sinir hücrelerine dayalı hesaplama teorisi açıklanmış ve eşik değerli mantıksal devreler geliştirilmiştir.

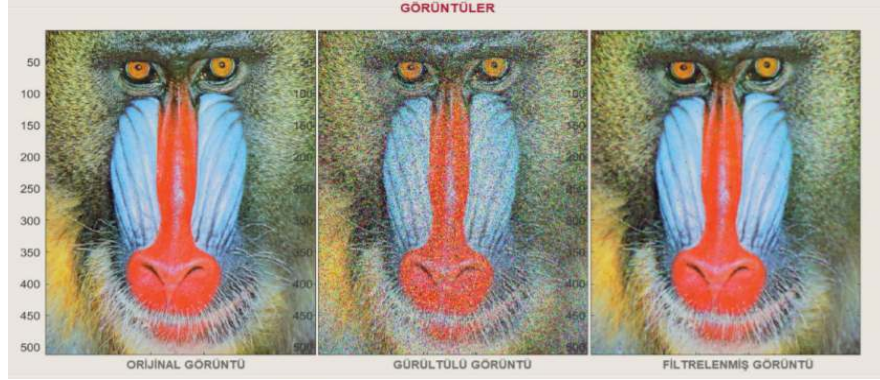
1960’lı yılların sonlarına yaklaşıldığı sırada, YSA ile ilgili çalışmalar neredeyse durmuştur. Bunun en büyük sebeplerinden biri; Minsky ve Pappert’ın yazmış olduğu Perceptrons kitabıdır. Bu kitapta; meşhur “XOR problemi” ile, doğrusal olmayan problemlerin YSA kullanılarak çözülemediği ispat edilmiş ve konu üzerine yürütülen çalışmalar adeta bıçakla kesilircesine kesilmiştir. Aynı dönemde, Grossberg, Carpenter ile Adaptif Rezonans Teorisi (ART)’ni ortaya atmıştır.

1982-1984 yılları arasında Hopfield’in yayınladığı çalışmalarla, YSA’ların geliştirilerek çözümü güç problemlerin çözümlenmesinde kullanılabileceği ortaya konmuştur. Hopfield, klasik gezgin satıcı problemine çözüm getirmiştir. Bahsi geçen çalışmaların, Hinton ve arkadaşları tarafından Boltzman Makinası’nın geliştirilmesine ön ayak olduğu bilinmektedir (1984). Çok katmanlı algılayıcıların geliştirilmesi de yine aynı dönemde olmuştur.

3.3. Görüntü İşleme Nedir?

Görüntü işleme; kaydedilmiş ya da ölçülmüş sayısal görüntülerden oluşan verilerin, dijital ortamda amaca uygun biçimde değiştirilmesini sağlayan çalışmadır. Görüntü işleme, mevcut resim ve grafiklerin değiştirilmesi, yabancılaştırılması veya iyileştirilmesi adına başvuru olan bir işlemdir. Görüntüyü işlemek için, varsa gürültünün temizlenmesi,

parlaklığın ayarlanması, koyuluğun ayarlanması, görüntünün keskinleştirilmesi ve bulanıklaştırılması, renk ayarlarının doğru yapılması gibi işlemleri kapsayan bir süreçtir.



Şekil 3-2: Filtrelenmiş görüntü örneği [8]

Görüntünün detaylarının kaybedilmesi (verim kaybı) ve görünümünün düşmesinin sebebi gürültülerdir. Çevrenin karanlık oluşu veya kameranın sallanması vb. sebeplerle gürültü söz konusu olabilir. Gürültülerin temizlenmesinde, görüntü işleme algoritmalarına başvurulmaktadır. Literatürdeki çalışmalar; farklı gürültü tipleri için farklı yöntemlerin uygulanabileceğini göstermektedir. Görüntü işleme, resimsel bilgilerin geliştirilmesi amacıyla, bir görüntünün doğasının değiştirilmesi ya da otonom makine algısına daha uygun hale getirilmesidir. Görselleştirme, görüntünün keskinleştirilmesi, görüntünün iyileştirilmesi, gürültülerin temizlenmesi, görüntünün tanınması gibi amaçlarla görüntü işleme teknolojilerinden faydalanılabilir.

Video ise birden fazla resmin ya da fotoğrafın art arda eklenmesi ile meydana gelmekte olup; bir videonun hareketli gibi algılanmasının nedeni resimlerin arka arkaya hızlı bir biçimde oynatılmasıdır. Belli bir zaman aralığında oynatılan kare sayısına kare hızı denmektedir. Bir videoda saniyede kaç resmin bulunduğu ise “frame per second” ile ifade edilmektedir. Saniyede 10 kare oynatıldığında, insan gözü bunu hareketli olarak algılayacaktır. Saniye başına düşen kare sayısı yüksek olduğunda, görüntüler doğal ve akıcı görülür. Videoların işlenmesinde, gerçekleştirilecek işlem için videodaki her bir frame çekilerek işlemlerden geçirilir. Sonrasında bu frame'ler bir araya getirilerek tekrar videoya dönüştürülür.

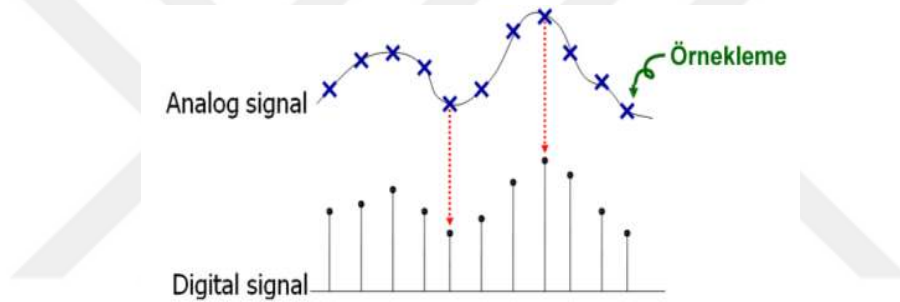
3.4. Sayısal Görüntü Tanımlamaları

Görüntü, üç boyutlu nesnelerin haritasının 2B yüzey üstüne düşürülmesi şeklinde tanımlanabilir. Bu haritalama, her noktaya ait konum ve renk bilgisini tutmaktadır. Bu

görüntülerin insan gözünün görebileceği bir biçimde gösterilmesin diye resim denebilir. Görüntünün henüz sinyal şeklindeki, insan gözünün görebileceği biçimde değil ise; bu, görüntünün henüz resme dönüşmemiş olduğu anlamına gelmektedir. Görüntü; Analog ve Dijital olmak üzere 2 kategoride değerlendirilebilir.

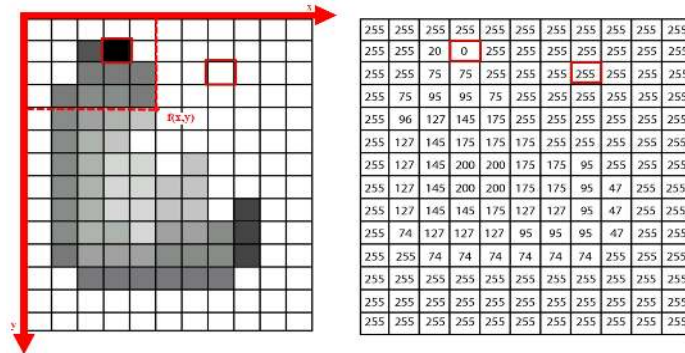
Görüntüyü meydana getiren fonksiyonun değişkenleri reel değerler alıyorsa, buna analog görüntü denmektedir. Analog görüntüde bir resme ne kadar yakından bakılırsa bakılırsa, görüntüyü oluşturan renkler halen orada bulunur. Sayısal bilgisayarlarla, sürekli parametreler işlenemez. Dolayısıyla bu fonksiyonlar, farklı yöntemlerle sayısallaştırılmalıdır.

Dijital (Sayısal) görüntüyse, $f(x,y)$ biçiminde ifade edilen sürekli görüntünün ayrık örnekler cinsinden discrete olarak gösterilmesi olup; $f[x, y]$ şeklinde gösterilmektedir.



Şekil 3-3: Analog sinyalin, dijital sinyale dönüşümünde örneklerin alınması [8]

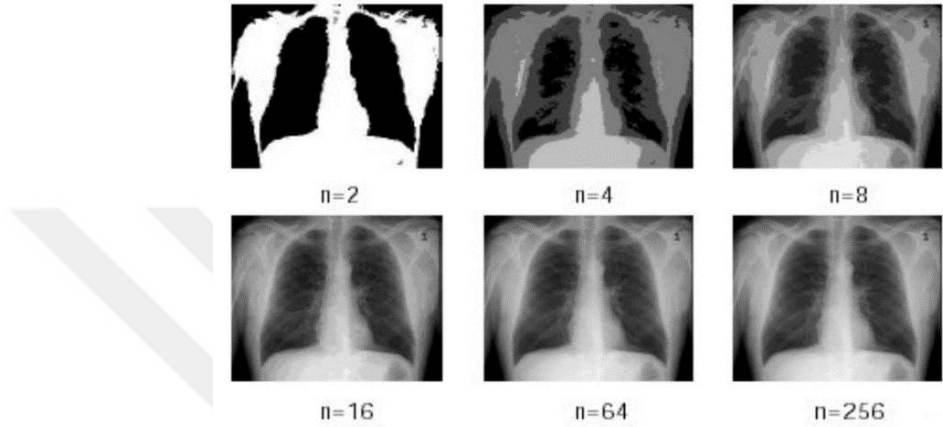
Sayısal görüntüyü sayısal değerler meydana getirmektedir. 1'lerden ve 0'lardan oluşmakta olan sayısal görüntünün yapısı $f[x,y]$, iki boyutlu dünyadan ulaşılan $f(x,y)$ fonksiyonu üzerinde örnekleme tekniği uygulanması ile elde edilir. Sayısal görüntü F ve X sayılarında satırlar ve sütunlardan oluşmaktadır. Satırlar ile sütunların kesiştiği her bir bölgeye piksel denmektedir. O pikseldeki değerse, derinlik (z), renk (λ) ve zaman (t) değişkenlerinin oluşturduğu bir fonksiyondur.



Şekil 3-4: Sayısal görüntü gösterimi [8]

3.5. Niteliklendirme

Görüntüye ait piksel değerlerinin belli aralıklarda olduğu durumlarda, ortaya çıkan görüntü nitelik bakımından değişir. Şekil 3-5'te görüldüğü üzere, “0” beyazı, “n_1” ise siyahı temsil ettiğinde, bu değerlerin arasındaki değerler gri tonlarına karşılık gelmektedir. Bu noktada, b değeri, $n = 2^b$ olacak şekilde görüntünün tek bir pikselini göstermek için gerekli bit sayısıdır [18].

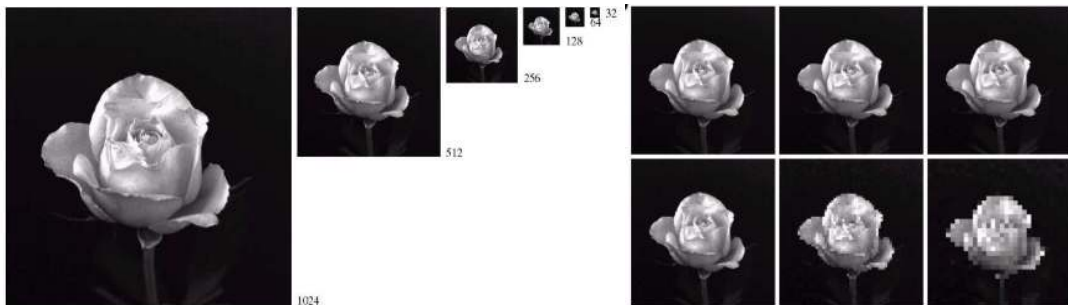


Şekil 3-5: “n” değerine göre nitelik değişimi [9]

Örnek vermek gerekirse; b=8 olması 256 adet gri tonunun mevcut olduğunu gösterir. B=1 olması ise resmin yalnızca sıfırlardan ve birlerden oluştuğunu göstermekte olup; buna ikili görüntü denmektedir.

3.6. Uzaysal Çözünürlük

Bir resmin sahip olduğu uzaysal çözünürlük, o resmin bir pikselinin fiziksel büyüklüğüne eşittir. Yani uzaysal çözünürlük, bir resmin detaylanabilir en küçük parçası olarak ifade edilebilir. Ayrıntıların fark edilebildiği en küçük çözünürlük olarak da tanımlanır. Piksel sayısı azaltıldığında, görüntünün uzaysal çözünürlüğü düşecektir ve görüntünün içindeki detaylar görülmemeye başlayacaktır.



Şekil 3-6: Uzaysal çözünürlük değişimi [9]

3.7. Koyuluk

Görüntü, lokal etkilerin etkisi ile göz tarafından farklı algılanabilir. Görüntünün oluşması ve gözün görüntüyü algılaması bakımından, kontrast önemli bir unsurdur. Şekil 3-7’de, merkezde yer alan karenin aynı olmasına karşın farklı kareler olarak algılandığı görülebilir.



Şekil 3-7: Koyuluk Kavramı [8]

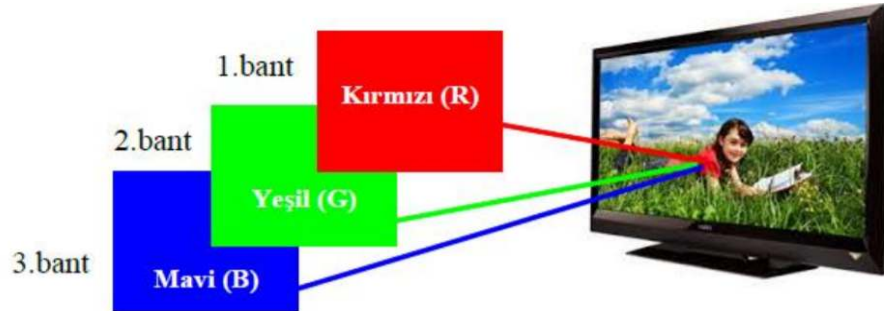
Şekil 3-8’de her bir dikdörtgensel bölgenin aynı renkte olmasına karşın, göz bunları farklı renkler olarak algılamaktadır ve 9 farklı renkte dikdörtgen varmış algısı yaratmaktadır.



Şekil 3-8: Aynı renkli 9 dikdörtgen [8]

3.8. RGB ve Gri Tonları

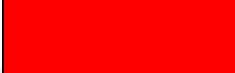
Renkli görüntüler, kırmızı (R), yeşil (G), mavi (B) kodlanmış aynı cisme ait 3 adet gri görüntünün ekranda üst üste gösterilmesiyle meydana gelir. Renkli görüntünün meydana gelmesini sağlayan bu 3 renge bant denmektedir [8].



Şekil 3-9: Renkli görüntünün oluşumu [8]

RGB görüntülerin her bir pikseli ekranında 24 bitlik veri olarak görüntülenmektedir. Şöyle ki, her bir renk 8 bitle ($2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 28 = 256$); ve dolayısıyla üç renk (RGB) $3 \times 8 = 24$ bitle kodlanacaktır [9]. Dolayısıyla, renkli olan görüntülerin her bir pikseli $28 \times 28 \times 28 = 224 = 16.777.216$ (yaklaşık 17 milyon) farklı renk değeri alabilir ve bu üç rengin birleşimi (0, 0, 0)...(255, 255, 255) değer aralığında yer alır. Aşağıda verilen tabloda, bazı örnek renkler ve değerler yer almaktadır [8].

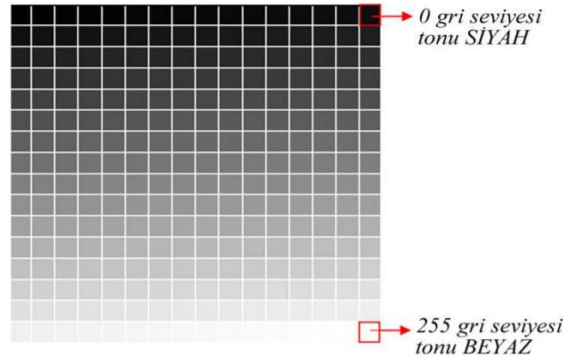
Tablo 1: Renklerin RGB değerleri [8]

	Kırmızı	Yeşil	Mavi	GÖRÜNÜM
KIRMIZI	255	0	0	
YEŞİL	0	255	0	
MAVİ	0	0	255	
BEYAZ	255	255	255	
SİYAH	0	0	0	
AÇIK GRİ	200	200	200	
KOYU GRİ	100	100	100	
SARI	255	255	0	
TURKUAZ	0	255	255	
EFLATUN	255	0	255	

Sayısallaştırma, görüntü boyutlarında ve her bir pikselde mevcut olabilecek parlaklık değerlerinin belirlenmesini gerektirir. Sayısal görüntüdeki her bir pikselin taşıdığı parlaklık değerine “gri seviyeler” denmektedir. Her bir pikselin parlaklık değerlerinin kodlandığı bit sayısı dikkate alınarak, gri seviye aralıkları belirlenmektedir. Gri seviye sınırları arasında; beyaz ve siyah olmak üzere sadece 2 renk bulunmaktadır. Bu 2 rengin arasında kodlanan görüntüler, gri tonlu görüntülerdir.

Her bir piksel 8 bit ile kodlanmış olup; bu tür görüntülerin her bir piksel, $28 = 256$ farklı gri ton karşılığı değerlerinden meydana gelmekte olup; gri değer aralığı $G = \{0, 1, 2, \dots,$

255} şeklinde ifade edilmektedir. Kural olarak; 0 gri seviyesi siyah renge, 255 gri seviyesi ise beyaz renge ve bu değerler arasındaki gri seviyeler gri tonlara karşılık gelmektedir [9].



Şekil 3-10: Farklı gri seviyelerinin gösterimi [8]

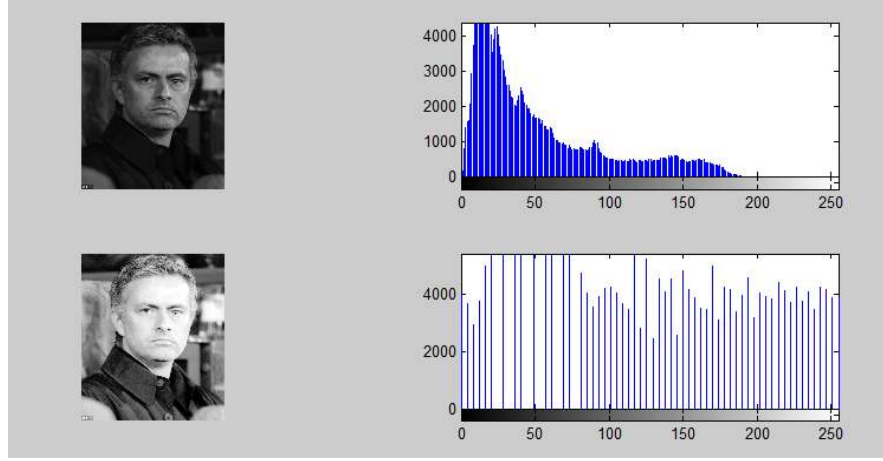
3.9. İndekslenmiş Görüntüler

Resimlerin daha iyi ve verimli saklanmasına yönelik yöntemlerden biri de indeksleme olup; indekslenmiş resimler birer renk paletine sahiptir. Renk paleti, 3 kolonlu ve her bir satırda ilgili rengin RGB değerinin yer aldığı bir matris olup; her piksel, bu tabloda bir indeks değerine sahiptir. Farklı renk paletlerinin kullanılmasıyla, bir resim farklı şekillerde ifade edilmesi mümkündür. Renkli resimlerin RGB değerlerinin ayrı ayrı tutulmasındansa, palet tutulması daha elverişli olmaktadır.

3.10. Histogram

Gruplandırılmış veri dağılımının sütun grafiği ile gösterimine histogram denmektedir. Farklı bir ifadeyle, tekrar eden sayıların oluşturduğu verilerin, uygulanan işlemlerin ardından öncelikle tablo aktarılması ve ardından bu tablodan faydalanılarak grafiğe dökülmesi, bir diğer ifadeyle veri grupları grafiğinin dikdörtgen sütunlar şeklinde gösterimidir. Sütunlar belirlenen aralıkların sınıfını belirlemekte olup; sütun yükseklikleri her bir aralıkta kaç değer bulunduğunu göstermektedir. Bölmelerin eşit boyutlara sahip ve bitişik olmaları gerekmektedir. Fazla sayıda ve dar sütunlar, histogram güvenilirliğinin artmasını sağlar. Sütun sayısının düşük olması ve sütunların geniş olması ise, ulaşılacak bilgilerle doğruluğu düşük varsayımlarda bulunulması olasılığını artıracaktır.

Görüntü üzerinde yaratılan histogramlarda, görüntüdeki her bir pikselin seviyesini, yani frekansı göstermektedir. Koyu ya da açık tonlarda olan görüntülerin bilgisine, histogram aracılığıyla ulaşılabilir.



Şekil 3-11: Histogram Örnekleri [36]

Histogramlarda görüntünün iyileştirilmesi amacıyla histogram eşitlemeye başvurulur. Bu, histogramdaki frekansla gerçekleştirilen ve doğrusal olmayan bir eşitleme işlemi olup; görüntünün frekans değeri yüksek olduğu takdirde, piksel seviyesi geniş bir alana yayılacaktır. Görüntünün düşük frekanslı olduğu durumlarda ise histogram, dar bir piksel alanına yayılacaktır. Görüntü histogramının bulunması ile kümülatif histograma ulaşılır. Değerlerin normalize edilerek, eski değerlerle ulaşılan değerler birbirine karşılık düşürülür. Böylece yeni histogram grafiği elde edilmiş olur.

4. YÜZ TANIMA VE YÜZ İFADESİ TANIMA

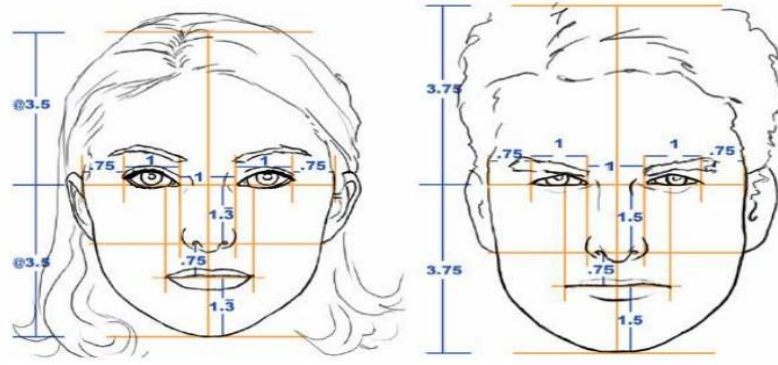
4.1. Yüz Bulma

Yüz bulma, herhangi bir görüntüde yüz olup olmadığının tespit edilmesi, varsa yüzün pozisyonunun ve kapladığı bölgenin belirlenmesidir. Yüz tanımlama sistemlerinin kullandığı başlıca ön işleyici modül, yüz bulma ve bulunan yüzün konumlandırılması işlemidir. Yüz bulma aynı zamanda yüz takibi, poz kestirimi ve ifade tanıma işleminin de ilk basamağını oluşturmaktadır [18]. Bir diğer ifadeyle, bu uygulamalarda başvuru olan çoğu yöntemde, görüntülerde yer alan yüzler tespit edilmiş ve yerleri belirlenmiş olarak varsayılır. Dolayısıyla söz konusu uygulamaların otomatik olarak gerçekleştirilebilmesinde, yüz nesnesinin bulunması ilk adım olarak kabul edilmektedir.

Yüz tespitinin ardından, yüz ifadelerinin yol açtığı yüz değişiklikleri çıkarma ve temsil etme adımı gelmektedir. İfade analizinde yüz özellik çıkarımı için temel olarak 2 farklı yaklaşım mevcuttur. Söz konusu yaklaşımlar; geometrik özellik tabanlı metotlar ve görünüm tabanlı metotlar olarak tanımlanmıştır. Geometrik yüz ifadelerinde ağız, kaşlar, göz ve burun yerleşimi söz konusudur. Yüz geometrisinin temsil edildiği bir özellik vektörü, yüz bileşenlerinden çıkarılan formudur. Görünüm tabanlı yöntemlerle görüntü filtreleri, özellik vektörü çıkarılması amacıyla, yüzün tamamına veya belli bir bölgesine için uygulanabilir [13].

Beden oranlarındaki gibi, insan yüzünde de genellikle doğru olarak kabul gören belli orantılar söz konusudur. İnsan yüzü ile ilgili altın oranların bir bölümü şu şekildedir;

- Yüzün uzunluğu/genişliği,
- Kaşların ve dudakların birleşim yeri aralığı,
- Yüzün uzunluğu
- Ağız uzunluğu ile burun genişliği,
- Burun delikleri ile burun genişliği aralığı
- Göz bebekleri aralığı ile kaşların aralığı [13].



Şekil 4-1: Yüz oranları [13]

Kadın başının oranları temelde erkek başı oranları ile aynıdır. Kadın ve erkek başlarının doğal (makyajsız) hallerinde kayda değer bir farklılık görülmemektedir. Bulunan farklılıklarsa sıklıkla kadının fiziksel yapısından ve özellikle yüz yapısına etki eden deri altı yağların fazla miktarda oluşundan ileri gelmektedir. Genellikle erkekler kadınlara kıyasla daha hareketli bir yaşantıya sahip olduğundan, kemik ve kas yapılarının daha güçlü olduğu görülmektedir. Örneğin bir erkeğin burnu çoğunlukla daha geniş ve iri olmaktadır; bunun nedeniyse erkeklerin daha güçlü solunuma ihtiyaç duyması ve dolayısıyla boğaz vb. gibi solunum organlarının daha büyük olmasıdır.

Kadın ile erkek arasında söz konusu olan farklılıklara rağmen, kimi erkekler kadınsı yüz hatlarına, kimi kadınlarsa erkeksi yüz hatlarına sahip olabilmektedir. Söz konusu farklılıklar, erişkinlik döneminde daha da belirginleşmekte fakat yaşlanma süreci ile azaldığı görülmektedir.

ERKEK					KADIN				
%	Oran	Ölçüm Değerleri (mm)		Eski Oranlar	Eski Oranlar	Ölçüm Değerleri (mm)	Oran	%	
31	1	58		1	1	51	1	30	
30	1	57		1	1	56	1.1	32	
39	1.2	72		1	1	66	1.2	38	

Şekil 4-2: Erkek ve kadında alın / burun / çene uzunluğu oranları [13]

Genellikle, kadınların yüzlerinin biraz küçük, gözlerinin biraz büyük, kaşlarının biraz yüksek ve kavisli, burun ve ağızlarının daha küçük, dudaklarının daha etli ve çenelerinin

yuvarlak olduđu gör÷lmektedir. Erkeklerin yüz hatları daha köşeli iken, kadınlar daha yumuşak ve kavisli yüz hatlarına sahiptir.

4.2. Biyometrik Teknolojiler ve Tanıma

İnsanların davranışsal ve fiziksel özellikleri ele alındığında, yürüt÷lmekte olan kimlik belirleme çalışmaları “biyometri” olarak adlandırılmaktadır. Retina ve iris görüntülerinin tanınması, el geometrisinin ve parmak izlerini konu alan araştırmalar; eldeki damar görüntüsünün, konuşma sesinin, yüz özelliklerinin, yürüyüşün ve klavye kullanımlarının incelenmesi bu çalışmalara örnek olarak gösterilebilir. Daha geniş bir çerçevede; biyometri bilişim teknolojileri alanının bir parçasıdır. Yüz bulma ve yüz tanıma da biyometrinin kapsamına girmektedir.

Organizasyon ve güvenlik, tanımlama ve doğrulama yöntemlerinin gün geçtikçe önem kazanmasıyla birlikte, günümüz koşullarında birçok alanda bir anahtar teknoloji geliştirilmeye çalışılmaktadır. Söz konusu teknolojiden, bina girişlerinin kontrolü, bilgisayarlara girişin kontrolü ya da otomatik vezne makinaları gibi önemli sahalarda yararlanılmaktadır.

Biyometrik karakteristikler davranışsal ve fiziksel olmak üzere 2 kategori altında incelenmektedir. İmza ve tiplendirme ritimleri gibi davranışlar, davranışsal biyometriklerin kapsamına girmektedir. Fiziksel biyometrik sistemlerinde ise; tanımlama amacıyla göz, parmak, el, ses ve yüz unsurları kullanılmaktadır. Yüz tanıma sistemleri, AR&GE süreçleri devam etmesine karşın; ticari, askeri, sivil olmak üzere çok sayıda kullanım alanına sahiptir.

Tablo 2: Biyometrik tanıma tekniklerin karşılaştırma [16]

Biyometrik	Doğruluk	Güvenirlik	Yaygınlık	Performans
Yüz	Yüksek	Yüksek	Yüksek	Yüksek
Parmak İzi	Orta	Yüksek	Düşük	Orta
El Geometrisi	Yüksek	Düşük	Düşük	Orta
Retina	Orta	Yüksek	Düşük	Orta
Ses Tanıma	Düşük	Düşük	Yüksek	Düşük

4.3. Yüz Tanıma Tarihçesi

İlk yüz sınıflama metodu resmi olarak 25 Mayıs 1888 tarihinde, Sir Francis Galton tarafından ortaya koyulmuştur. 1980’li yıllarda yüz tanımaya odaklanan çalışmalar sayıca fazla değildir. Yapılan çalışmaların ise zamana ayak uydurabilecek niteliğe sahip olmadığı bilinmektedir. 1990’larla birlikte dolaylı yüz tanımaya duyulan ilginin arttığı görülmekte olup; bu durumun nedenleri şöyle sıralanabilir:

- Sivil/ticari araştırma projelerine verilen önemin artması.
- YSA sınıflaması için gerçek zamanlı hesaplamaların ve adaptasyonun artan bir öneme sahip olması.
- Gerçek zamanlı donanımlara ait kullanılabilirlik düzeyi.
- Uyuşturucu trafiği, terörist faaliyetleri vb. illegal hareketlerdeki artıştan ötürü söz konusu faaliyetlerin takip edilmesi ve bu faaliyetlerle insanlar arasındaki bağlantının kurulabilmesini sağlayacak uygulamalar için duyulan gereksinimin artması.

4.4. Yüz tanımanın yaygınlaşmasındaki faktörler

Yüz tanımanın geçtiğimiz 10 yıl içinde bu derece ilerleyerek, büyük bir pazar potansiyeli sunmasının ardındaki etmenler şöyle açıklanabilir;

- Yüz tanımanın kendine özgü yanları: Yüz tanıma, kimlik tanıma kolaylığı sunan, bunu yaparken hayatın akışını bozmayan düzenekleri temel alır. Böylece kişiler, parolalarını akıllarında tutmak veya kartlarını yanlarında taşımak zorunda kalmazlar. İnsanlar için kendi yüzlerini başkalarının yüzleriyle değiştirmek olanaksızdır; ancak parolalarını, taşınabilir giriş kartlarını değiştirebilirler. Yüz tanıma için ipuçlarının unutulması ya da bir başkasınıkiyle değiştirilmesi de ihtimal dahilinde değildir.

• Altyapının tümleşik parçası: Yüz tanıma yaşamı kolaylaştırdığı ölçüde güvenlik düzeyini artıran özellikler de taşımakta olup, ticari hayatın tüm sahalarına sızmaya başlamış bir yöntemdir. Bu şekilde sağlık, savunma, finans, ulaşım, güvenlik vb. alanlardaki uygulamaların gün geçtikçe gelişmekte olduğu görülmektedir. Diğer taraftan bilgisayar ağlarının ve elektronik teknolojilerin aracılığıyla, dünya birbirine bağlı elektronik bir kente dönüşmüş durumdadır ve dünyanın bugünkü haliyle, kimlik belirleme her zamankinden çok büyük önem taşımaktadır. Fakat klasik kimlik belgelerinin bu konuda yeterli olduğu söylenemez. Örnek vermek gerekirse, giriş kontrollerinde, pasaport ve bankacılık işlemlerinde kredi kartları, arzu edilen güvenliği tek başına sağlamak için yetersizdir. Buna

benzer geleneksel belgelerin yerine geçmeseler de bugün sistemlerin yüz tanıma teknolojileriyle pekiştirilmesi bir zorunluluğa dönüşmüştür.

- Yaşam kalitesi: Yüz tanıma teknolojisine yalnızca güvenlik açısından yaklaşılması doğru değildir. Bu; yaşamın daha konforlu ve rahat bir hal almasını sağlayan yeni bir teknoloji olarak ele alınmalıdır. Mesela, bir otelin devamlı müşterisi, otele her girişinde kimlik bilgilerini ibraz etmekle yükümlüdür. Fakat yüz tanıma sayesinde, bunun gibi durumlarda kimlik bilgileri bilgisayardan dökülebilir ve müşterinin belge doldurmasına gerek kalmaz. Fabrikada, işyerinde elinde eşya veya dosya ile kapıdan geçerken kart okutmak veya parola girmek zorunda kalmak büyük zaman kayıpların sebep olarak işteki verimliliği düşürmektedir. Oysa yüz tanıma teknolojisinden yararlanılarak söz konusu problem ortadan kaldırılabilir.

4.5. Yüz İfadesi Tanıma

İnsanların duygusal durumları, yüz ifadeleriyle açığa vurulur. Yüz ifadeleri sözlü olmayan iletişim araçlarından olup, insanlar arası iletişimde başlıca rolü olan öğelerdendir. Sözle anlatılamayacak pek çok duygunun, yüz ifadeleriyle aktarılması mümkündür. Sevinç/mutluluk, kızgınlık-öfke, üzüntü, şaşırma, korku, reddetme/nefret/tiksinme vb. hislerin her birini gösteren kas hareketleri insanların tümünde aynı şekildedir.

İletişim, insanın bilerek yahut bilmeyerek vermek istediği mesajın farklı kanallarla hedefe aktarılmasıdır. Kişilerarası iletişimde mesaj, ne kadar fazla duyu organıyla algılanıyorsa, iletişim derece etkin olacaktır. Kişiler, değişik kanallar üzerinden gelen sinyalleri eş zamanlı olarak alabilir ve bir kanal tarafından alınamayan mesajın diğer kanal aracılığıyla algılanması yoluyla iletişim sağlıklı bir hal alabilir.

İletişim, genellikle sözlü iletişim ve sözsüz iletişim olmak üzere 2 başlık altında incelenir. Ses tonu ve sözcükler sözlü iletişimde, mimikler ise sözsüz iletişimde temel unsurlardır. Sözlü iletişimle düşünce, sözsüz iletişimle ise duygu aktarımı gerçekleşir. Sözle inkâr edilen ruhsal durum, sözsüz mesajlar aracılığıyla algılanabilir. Bireyler, isteyerek veya farkında olmadan yüz kaslarının farklı biçimler almasını sağlayabilir ve duygu durumlarındaki değişimleri bu şekilde aktarabilirler.

İletişim alanındaki çalışmaları ile tanınan Mehrabian, insan iletişiminde yüz ifadelerinin sadece %55, ses tonunun %38, seçilen kelimelerin ise %7 oranında önemli

olduğunu ileri sürmüştür [50]. Bilgisayarın gündelik yaşantının ayrılmaz bir parçası oluşu, bilim insanlarını insanlara daha iyi hizmet sağlayacak bilgisayarlı sistemler üretmeye itmiş, bu bağlamda, yüz ifade analizine ilişkin araştırmalara ağırlık verilmiştir. Yüz ifadesi analizine yönelik sistemler, bilimin çeşitli alanlarında, bilhassa bilgisayar teknolojileri alanında farklı amaçlarla kullanılmaktadır. Klavye-fare yerine, bakışları takip ederek verilen komutlar doğrultusunda çalışan sistemler birçok alanda büyük kolaylıklar getirmektedir. Bundan dolayı, yüz ifade analizi yapan sistemler büyük bir önem taşımaktadır. İnsansı robotların insanlar ile etkileşiminde de yüz ifade analizine başvurulmaktadır. Video konferanslarda görüntünün aktarılmasındansa, gereken noktalarda yüz ifadelerinin çözülmesi ve bilgi halinde aktarılması da tercih edilecek durumlardır. Araç kullananların yüz ifadelerinin izlenmesi, olası kazaların önüne geçilebilecek, sistemlere önceden müdahale edilmesi ve sistemlerin kontrol altına alınma sağlanabilecektir. Suçluların ruh hallerini gizlemeleri durumunda ya da hastaların kendilerini ifade etmekte zorlandığı hallerde, yüz ifade analizi sistemlerinden yararlanılması önemli avantajlar yaratacaktır.

4.6. Yüz Bulma Algoritmaları

Literatürdeki çalışmalarda yüz bulma yöntemleri genel olarak; bilgi tabanlı yöntemler, değişmeyen özellik yaklaşımları, şablon eşleştirme yöntemleri ve görünüm tabanlı yöntemler şeklinde sıralanmaktadır.

Bilgi tabanlı yöntemler üç temel başlık altında incelenmekte olup; bunlar, şablon eşleştirme yöntemleri, değişmeyen özellik yaklaşımları ve görünüm tabanlı yöntemlerdir. Bilgi tabanlı yöntemler ve değişmeyen özellik yaklaşımları çoğu zaman yüz lokalizasyonu problemiyle uğraşmaktadır. Yüz lokalizasyonu, içinde tek bir yüz nesnesinin bulunduğu resimde, yüz pozisyonunu belirlenme işlemi olup; söz konusu problem yüz bulma işleminin basitleştirilmiş halidir. Şablon eşleştirme yöntemi ve görünüm tabanlı yöntemlerse yüz bulma ve lokalizasyonu problemleriyle eş zamanlı olarak uğraşmaktadır [17].

4.6.1. Bilgi Tabanlı Yöntemler

Bilgi tabanlı yöntemler, bir insan yüzünü oluşturan tipik yapı parçalarının belirli kurallar çerçevesinde çözümlenmeleri prensibine dayalı olarak çalışır. Belirlenmiş olan kurallar, yüzü oluşturan parçaların birbirlerine ilişkilerini tanımlar. Bu parçaların ilişkileri, birbirlerine olan uzaklıkları ve farklı pozisyonları üzerinden tanımlanabilmektedir. Yüz

nesnesini tespit etmede, öncelikli olarak eldeki resimde bahsi geçen yapısal parçalar aranır ve sonrasında kurallarla uyumlu aday bölgelerin tespiti yapılır. Tespitin hatalı olmasını önlemek adına, belirlenen aday bölgeler yeniden incelenerek üzerlerinde doğrulama işlemi gerçekleştirilir.

Bu yöntemle ilişkin en belirgin problem, kuralları uygun biçimde tanımlanmanın zor olmasıdır. Kurallar aşırı fazla katı olduğu takdirde, tüm kuralları geçilemeyecek ve de dolayısıyla yüz tespit edilemeyecektir. Öte yandan kurallar genel bir çerçevede tanımlandığında, hatalı tespitlerin sayısı oldukça fazla olabilir. Bunun yanı sıra, resimde bulunan yüzlerin kafa yönelimlerinin farklı olması da karşılaşılan problemlerden bir diğeridir. Bu durumda söz konusu olduğunda yapısal parçaların arasındaki ilişkiler değişmekte ve potansiyel tüm durumların tanımlanması olanaklı olmamaktadır. Bundan ötürü, bu yöntemin ön profilden ve arka planı karmaşık olmayan resimlerde genellikle daha başarılı sonuçlara ulaştığı görülmektedir.

Bahsi geçen problemlerin bir bölümünün aşılması adına Yang ve Huang tarafından hiyerarşik bir model geliştirilmiştir. Test edilecek resmin farklı düşük çözünürlükte kopyaları kullanılarak, her birine farklı kurallar uygulanmakta olup; çözünürlüğü en düşük resimlere bakıldığı zaman, daha genel kurallar uygulandığı görülürken, bu kurallardan ise geçen tüm aday bölgelerin daha yüksek çözünürlükteki versiyonları için daha ayrıntılı kurallar uygulandı [1]. Başvurulan hiyerarşik yaklaşım gereken hesaplamaların sayısının azaltılmasını sağlamaktadır. Yöntemin sunduğu tespit oranları çok başarılı olmasa da sonraki yüz tespiti çalışmalarında hiyerarşik yaklaşımın kullanılması bakımından önemlidir.

4.6.2. Değişmeyen Özellik Yaklaşımları

Kafa yönelimi veya ışıklandırma gibi koşullar değişse dahi kimi yapısal özelliklerle yüz tespit edilebilmektedir. Kaşlar, gözler, ağız, burun, saç hattı vb. yapısal özelliklerin konumları çoğunlukla kenar sezicilerden yararlanılarak belirlenir. Söz konusu yapısal parçalar arasında mevcut olan ilişkilerin tanımlandığı istatistiki bir model hazırlanarak, yüz tespiti bu modelle gerçekleştirilmektedir. Aynı zamanda deri rengi, yüz biçimi ya da yüz ve saç dokusu vb. özelliklerden de yüz tespitinde yararlanılabilmektedir. Global nitelikte bu özellikler resimde aday bölgelerin belirlenmesinde kullanılmakta olup; sonrasında kaş, göz, ağız gibi yerel, ayrıntılı özelliklerle ilgili bölgeler kontrol edilmektedir.

Aydınlatma, tespit edilecek yüzün bir bölümünün başka bir nesne tarafından örtülmesi gibi durumların problem algoritmalarının başarısına aşırı etki etmesi, bu yönetime dair en belirgin problemidir. Böyle bir durumda oluşan gölge, gerçekte olmayan ayrıtların ortaya çıkmasına yol açabilmekte yahut ayrıt sınırlarının belirsizleşmesiyle metot kullanışsız bir hal alabilmektedir [1].

Sirohey, yüzü karmaşık bir arka plandan ayırıştıran bir lokalizasyon metodu geliştirmiştir. Canny kenar sezicisiyle belirlenen ayrıtlar gruplanmakta ve yalnızca yüze ait olanların tespiti yapılmaktadır. Bunu takiben, baş ve arka planın arasında bulunan sınıra uygun bir elips oturtulur.

4.6.3. Şablon Eşlemeli Yöntem

Bu yöntem ile; genellikle ön taraftan yansıtılan standart bir yüz modeli elle veya bir fonksiyon aracılığıyla tanımlanır. Verilen deneme resmi ile model arasındaki ilinti yüz hattı, gözler, ağız ve burun için ayrı ele alınıp hesaplama sürecinden geçmektedir. İlintinin yüksek olması durumunda yüz bulunmuş olmaktadır. Gerçekleştirmenin basit olması avantaj sağlamakla birlikte boyut, kafa yönelimi ve şekil değişimlerinden fazla etkilenmesi yöntemin önemli bir dezavantajıdır. Dezavantajları incelemek amacıyla çok çözünürlüklü şablonlar, çok ölçekli şablonlar, alt şablonlar ve deforme edilebilen şablonlar tanımlanmıştır. Buna rağmen bu yöntem, yüz algılamada beklenen sonuçları verememektedir [1].

Ön profilden görünen yüzleri bulunmaya yönelik bir metot geliştiren Sakai; gözler, burun, ağız ve yüzün çizgi detayları üzerinden tanımlandığı alt şablonlar hazırlamıştır. İki aşamalı bu yöntemde, aday bölgeler işlenmiş resim ile kontur şablonunun ilintisi doğrultusunda belirlenir. Sonrasında, yüz sezimi diğer alt şablonların eşleştirilmesi ve ayrıntılı biçimde analiz edilmesi ile tamamlanır.

Craw lokalizasyon probleminin çözümlenmesi amacıyla, şekilsel şablondan yararlanılarak ön cepheden yüz modeli geliştirmiştir. Teknik, deneme resimlerinden Sobel süzgeçleri aracılığıyla tüm ayrıtları çıkararak yüz hatlarını sabitledikten hemen sonra; gözler, kaşlar vb. özniteliklerin lokalizasyonuna ulaşabilmek için peş peşe uygulamaya konmaktadır.

Samal, ana bileşenler analizi yöntemi ile, bir dizi yüz silueti elde etmektedir. Elde edilen yüz ifadeleri genelleştirilmiş Hough dönüşümü ile kullanıldıktan sonra yüz lokalizasyonu gerçekleştirmek için deneme resimlerine uygulanmaktadır [1].

4.6.4. Görünüm Tabanlı Yöntemler

Şablonun, şablon eşleştirmeden farklı olarak şablon elle hazırlandığı bu yöntemde, şablon bir grup örnek seti üzerinden elde edilir. Nesne içeren ve içermeyen görüntülerle alakalı karakteristiklerin istatistiksel dağılımları veya ayırtaç fonksiyonları ile ifade edilmektedir. Hesaplama kazanç sağlamak ve hata oranlarını en aza indirmek için genel itibari ile boyut düşürülmektedir.

Yüz bulma uygulamaları için, sinir ağlarına yaygın biçimde başvurulmaktadır. YSA genellikle matematiksel modeli oluşturulamayan ya da modellemenin güç olduğu uygulamalarda daha başarılı veriler doğuran, doğrusal olmayan giriş çıkış ilişkilerini veri setlerinden yararlanarak kuran ve insan beyninden ilham alınarak tasarlanmış yapılardır. Zor koşullarda olasılık dağılımlarına uygun kalınarak sistem eğitimi verilmesine imkân tanınması yöntemin çok önemli bir avantajı olarak yer almaktadır; buna rağmen tatmin edici sonuçlar elde edilebilmesi için katman sayısı, kullanılan düğüm sayısı ve öğrenme oranı parametrelerinin uygun şekilde ayarlanmış olması gerekmektedir. Yapılmış çalışmalardan en dikkat çekici olanları Rowley tarafından gerçekleştirilmiştir [1].

Görünüm tabanlı çalışmalar açısından önem taşıyan bir diğer araştırmayı ise Viola ve Jones gerçekleştirmiştir. Çok sayıda nesne sezimi yönteminde, görüntünün her bir noktasında uygulandığından oldukça fazla sayıda hesaplama ihtiyacı söz konusudur. Hedefteki konumların birçoğunda yüz görüntüsü bulunmadığından, her yerde yüz bulunması düşük olasılık taşımaktadır. Bundan ötürü yüz barındırmayan görüntü bloklarının birçoğu, basit sınıflandırıcılarından yararlanılarak elenebilmektedir.

4.7. Yüz İfadelerinin Değerlendirilmesi

Yüz ifadelerinin değerlendirilmesinde, kişinin yüz fotoğrafları incelenerek duyguları ve heyecanlarına ilişkin bilgi edinilmeye çalışılır.

- Mutluluk ifadesi, hoşnutluk olarak değerlendirilmekte olup; dudaklar yukarı kıvrılır.
- Korku ve öfke hoşnutsuzluk olarak değerlendirilir ve dudaklar aşağıya kıvrılır.

- Sevinç durumu yaşandığında gözlerin yukarı doğru, keder ifadesindeyse aşağı doğru baktığı görülür.

- Dikkat durumunda en belirgin yüz ifadesi olarak; gözler büyürken, burun delikleri genişler ve ağız açılır.

- Reddetmede ifadesinde gözler, burun delikleri ve dudaklar uyarımı dışarıda bırakacak şekilde, sıkı bir biçimde kapanmaktadır [2].



Şekil 4-3: Farklı kişilere ait yüz ifadeleri [2]

Kişiler, duyguları ve heyecanları yüz ifadeleriyle yansıtmaktadır. Esasen bu yüz ifadesi şekilleri, kültürden kültüre değişebilmektedir. Bunun yanı sıra farklı kültürlerle mensup kişilerin, diğer insanların ifadelerini algılama bakımından benzerliklere sahip olduğu bilinmektedir. Söz konusu benzerlik, doğuştan bir özellik sayesinde, bir heyecan duygusu yaşanırken çok sayıda kas hareketinin aktifleşmesi ile açıklanabilir.

Yüz ifadeleri üzerinden ulaşılan bulgular, yüz ifadelerinin yalnızca heyecanların yansımından ibaret olmadığını, aynı zamanda duygusal tecrübelerin belirlenmesine de destek olduğunu ortaya koymaktadır. Yüz ifadelerini sınıflandırmada başvurulan tekniklerden birisi de özyüz yöntemi olup; yüz ifadeleri öfke, mutluluk, iğrenme, üzüntü, korku, şaşırma ve nötr şeklinde belirtilmiştir. Nötr durum, yüzde herhangi bir belirgin ifadenin bulunmadığı durumdur. Bir ifadenin derecesi ise nötr ifadeden olan uzaklığına bakılarak belirlenmektedir [2].

Yüz ifadesinin algılanması çözümü güç bir problem olup, bölge ve kültür bakımından yüz ifadelerinde farklılıklarla karşılaşılabilceğinden bunun bilgisayarla çözülmesi kolay değildir.

	KORKU	ÖFKE	REDDETME	SEVİNÇ	NÖTR	ÜZÜNTÜ	ŞAŞIRMA
KORKU							
ÖFKE							
REDDETME							
SEVİNÇ							
NÖTR							
ÜZÜNTÜ							
ŞAŞIRMA							

Şekil 4-4: Farklı yüz ifadelerinin birleşimi [2]

Yüz ifade analizi sistemlerinde temelde, algılama kapasitesi insanınki ile aynı olan; yaş, cinsiyet, ırk, kültür gibi farklar gözetilmeksizin çalışabilen otomatik bir sistemin geliştirilmesi amaçlanmaktadır. Geliştirilmiş olan sistemlerde kaynak veriler, yorumlanan bilgi ve inceleme yöntemleri bakımından farklar görülebilmektedir.

4.8. Yüz İfadesi Tanıma Sistemleri

Yüz ifadesi tanıma sistemleri 3 temel adımda gerçekleşmekte olup; bunlar yüz edinimi, yüz öznitelikleri çıkarımı ve ifadelerin sınıflandırılması olarak sıralanmaktadır.

- Yüz edinimi, yüzün konumunun belirlendiği aşamadır. Seçilecek olan öznitelik çıkarma yöntemi doğrultusunda, yüzün tamamı veya belli bir bölgesi belirlenir. Üzerinde

işlem yapılacak olan verinin yapısı dikkate alınarak yüz normalleştirme ön işlemleri uygulanır ve başarı oranının artması sağlanır [11]. Bir boyuta sabitleyerek yeniden boyutlandırma, aydınlatma adımlarına başvurulabilir [2]. Arka planların ayrılması veya yüz özniteliklerinin ayrılması da ihtiyaç doğrultusunda uygulanmakta olan adımlardır.

- Yüz özniteliklerinin çıkarılması, tek bir imgeyle veya imge dizileriyle ifade tespitinde kullanılacak bilgilerin çıkarımıdır. Tek bir görüntü tabanlı çalışmada yüz özellikleri ve özniteliklerini tespit edilir. Yüz öznitelik çıkarma işleminde yüz şekli bölgesel veya bütün olarak kullanılabilir. Bölgesel kullanımda yüzdeki temel bölgeler üzerinde incelemeler yapılır. Bütünsel kullanımda ise yüz bir bütün olarak değerlendirilir [2]. Yüz öznitelikleri, temelde kalıcı ve geçici yüz öznitelikleri olmak üzere iki grupta incelenir.

- Kalıcı öznitelikler, ifade değişiminin sonucunda oluşmayan, yüzün halihazırda sahip olduğu nitelikler olup; ağız, göz kapağı, kaş gibi temel bölgelerin yanı sıra, ileri yaşlarda karşılaşılan kırışıklık, sakal ve bıyık vb. nitelikler kalıcı özniteliklerdendir. Bu öznitelikler, ifade değişimleriyle deformasyona uğrayabilir [2].

- Geçici öznitelikler ise, ifade değişim esnasında oluşan ve hemen kaybolan niteliklerdir. Gözü ve ağzı çevreleyen bölgelerde meydana gelen kasılmalar, tendeki değişimler geçici öznitelikler olarak sayılabilir [2].

Yüz öznitelik çıkarma işleminde, temelde; geometrik özellik tabanlı yaklaşım, görünüm tabanlı yaklaşım ve karma yaklaşım olmak üzere üç yaklaşım benimsenmektedir:

- Geometrik özellik tabanlı yaklaşımla ağız, burun gibi yüz bileşenlerine ait geometrik özellikler ve yüzün nirengi noktalarının konumlarına dayanılarak çözüm üretilmeye çalışılır.

- Görünüm tabanlı yaklaşımlar, geometrik değişikliklerden ziyade yüz dokusu üzerindeki kırışıklıkları ve esnemeleri temel alır. İfade değişiminde yüz dokusunda oluşan değişiklikler incelenmeye çalışılır.

- Yüz ifadesi tanıma sistemindeki son aşama, sınıflandırmadır. Yüz edinim ve yüz öznitelik çıkarımı aşamaları sırasında ulaşılan öznitelik verileri sınıflandırılır, yorumlanır ve çıktıya ulaşılır. Mesaj veya işaret yorumlama yaklaşımları doğrultusunda çıktı tipi değişebilir. İşaret yorumlamada çıktıyı hareket birimleri ve bileşimleri meydana getirirken, mesaj yorumlamada çıktıyı algılanan ifade oluşturmaktadır.

5. YÜZ ÖZNİTELİKLERİNİN ÇIKARTILMASI VE SEÇİLMESİ

5.1. Gabor Filtresi

Görüntülerin analiz edilmesinde yararlanılan başlıca yöntemler arasında yer alan Gabor filtresi, bir görüntüde belirli yönde uzanan ayrıtları belirlenebilmesini sağlar. Bu sayede; plaka, parmak, iris veya yüz tanıma gibi görüntü işleme yöntemlerini temel alan yaklaşımlar adına önemli bir adım olma niteliği taşımaktadır.

Yüz görüntülerinin özniteliklerinin elde edilmesi, diğer yöntemlerle kıyaslandığında daha üstün performans gösterdiği görülen Gabor filtrelerinin kullanılmasıyla gerçekleştirilebilir. Bunun sebebi, Gabor filtrelemenin neticesinde, görüntünün homojen bir ışık dağılımına sahip olmayışının yol açtığı olumsuz etkilerdir. Bunun yanı sıra Gabor filtrelerinin, düşük seviyedeki yer değiştirmeler ve deformasyonlar karşısında daha az duyarlı olduğu bilinmektedir.

Gabor filtresi, görüntü işlemede ve bilgisayarla görmede, bilhassa öznitelik çıkarımı için oldukça yaygın bir tekniktir. İki boyutlu karmaşık Gabor fonksiyonunun matematiksel ifadesi aşağıdaki gibidir;

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[-\frac{1}{2} \left(\frac{\tilde{x}^2}{\sigma_x^2} + \frac{\tilde{y}^2}{\sigma_y^2} \right) \right] \exp [j2\pi f \tilde{x}] \quad (5.1)$$

$$\tilde{x} = x \cos \theta + y \sin \theta \quad (5.2)$$

$$\tilde{y} = -x \sin \theta + y \cos \theta \quad (5.3)$$

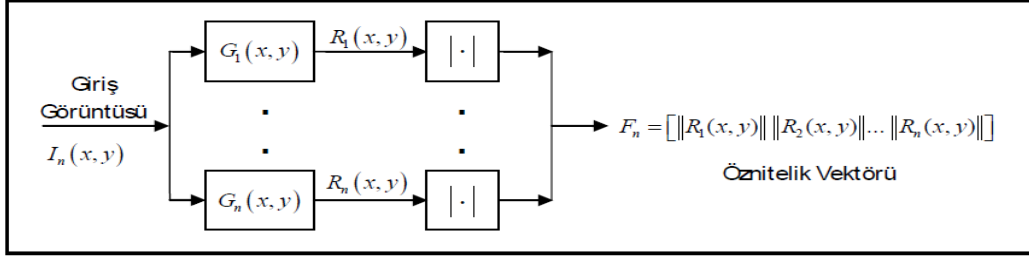
Bu fonksiyonda, σ_x ve σ_y sırayla x ve y eksenlerinde, Gauss standart sapmalarını temsil etmekte olup; f sinüzoidal dalga frekansını temsil etmektedir. Gabor filtrelerinin oluşturduğu bir filtre bankasının yaratılmasıyla, farklı açılar ve farklı frekanslar ile farklı niteliklerdeki özniteliklere ulaşılabilmesi mümkün olmaktadır.

Filtreye ait çekirdek matrisin hesaplanmasında fonksiyon (5.1) kullanılmaktadır. Bu fonksiyonda bir kompleks sayı döndürülür. Çarpımın gerçel ve imajinal bileşenlerinin ayrı ayrı hesaplanması mümkündür. Aşağıda yer alan ifade dönüş değerinin gerçel kısmı olup, bir sonrasındaki imajinal kısımıdır. Her iki fonksiyondan da yararlanılarak, reel ve imajinal değerlerin hesaplanması yapılır ve bu değerlerin kareleri toplamının kareköklerinin

alınmasıyla ulaşılan sonuçtan yararlanılabilir. Fakat genel itibariyle yalnızca reel değerleri hesaplamak yeterlidir.

Bir görüntünün, Gabor filtresi ile işlenmesi durumunda, ulaşılan çıkış, giriş görüntüsü $I(x, y)$ ve Gabor fonksiyonu $G_n(x, y)$ 'nin konvolüsyonudur.

$$R_n(x, y) = I_n(x, y) * G_n(x, y) \quad (5.4)$$



Şekil 5-1: Gabor Filtre Sistemi

Rastgele bir giriş görüntüsü $I_n(x, y)$ ve her kompleks Gabor fonksiyonu $G_n(x, y)$ 'nin konvolüsyonundan ulaşılan sonuç $R_n(x, y)$ olup, reel ve sanal şeklinde iki kısımdan meydana gelir. Bu durumda öznitelik vektörüne ulaşılması adına, bu sonucun genliği $||R_n(x, y)||$;

$$||R_n(x, y)|| = \sqrt{\Re^2\{R_n(x, y)\} + \Im^2\{R_n(x, y)\}} \quad (5.5)$$

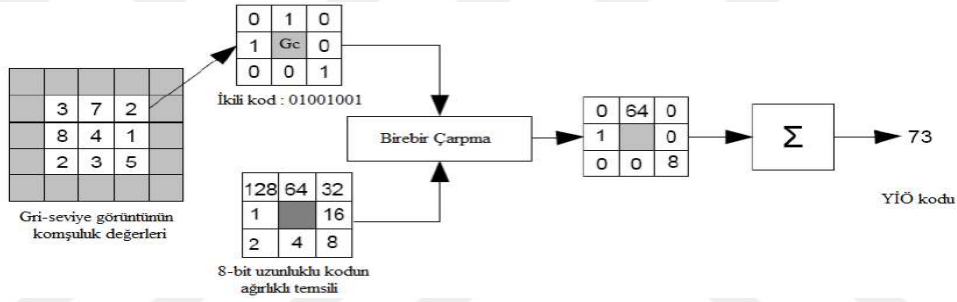
olarak elde edilir.

5.2. Yerel İkili Örnekler Operatörü

Yerel komşuluk değerlerinin temel alındığı bir öznitelik çıkarma yöntemi olan Yerel İkili Örnekler Operatörü, gri seviyeden bağımsız bir doku ölçüm yöntemi olup, Ojala vd. tarafından bulunmuştur. Görüntüdeki her piksel için bir etiket yaratılır. Etiketler, merkez piksel ile 3x3 komşuluktaki piksellerin karşılaştırılmasıyla ulaşılan ikili sayılardır. YİÖ operatörüyle farklı ölçeklerdeki dokular analiz edilebilmektedir.

YİÖ operatörüyle, özellik vektörü hesaplamak daha hızlı ve kolay bir işlem haline gelmektedir. Gri seviye üzerinde uygulanan bu yöntem, ışığa daha az duyarlı olduğundan tanıma uygulamalarında sıklıkla kullanılır. YİÖ operatöründen; yüz tanıma/yüz bulma, arka plan modelleme, endüstriyel kontrol, 3 boyutlu doku yüzeyleri sınıflandırma, yüz ifadelerini tanıma, yaş belirleme, hareketli nesne tespiti gibi bilgisayar tabanlı uygulamalarda yararlanılmış ve elde edilen sonucun yüksek performansa sahip olduğu görülmüştür.

Uygulamalar hayata geçirilirken çoğunlukla komşuluk değerleri göz önünde bulundurularak çeşitli operatörlere başvurulmaktadır. Bunlar; P merkez piksel etrafında alınacak komşu ve R örnek sayılarının temsil edildiği; YİÖP,R ile ifade edilmektedir. Çalışma kapsamında, YİÖ8,1 operatöründen yararlanılmıştır. 3x3 boyutlu bir çekirdek olan YİÖ operatörü, görüntülerdeki her pikseli, merkez piksel kabul etmekte ve etiketleme işlemini bu şekilde gerçekleştirmektedir. Merkez değer eşik değer olarak alınır ve etrafındaki 8 adet komşu değeri, bu değer ile karşılaştırılır. Bu karşılaştırma, ikili yöntem dikkate ile yapılır. Sonuç olarak ortaya çıkan 3x3lük operatör, dairesel olarak incelenir ve 8 bit uzunluğuna sahip ikili veri koduna ulaşılır. Bu ikili verinin onluk değeri, ilgili merkez pikselin yerel değeridir. Şekil 5-2’de; YİÖ operatörünün elde edilmesine ve bu bitlerin ağırlık katsayılarına yer verilmektedir [7].



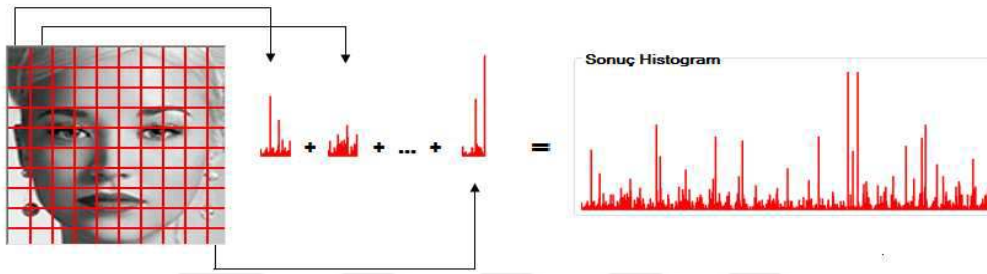
Şekil 5-2:YİÖ operatörü işlemleri [7]

YİÖ, merkez pikselin komşularına karşı düzenini olduğu gibi alırken, komşularının düzenliliğini olduğu gibi almadığından, YİÖ kodunun düzenli olup olmadığı kararı bit geçişleri arası düzenden yola çıkılarak verilir; bu karar ölçütü 0-1 ya da 1-0 geçişinin en fazla iki tane olmasıdır. 256 YİÖ kodundan 58'i düzenli olup, diğer 198'i düzensizdir. Düzenli YİÖ kodlarının bulunduğu durumlar, aşağıda yer verilen fonksiyonlarla gösterilebilmektedir. Fonksiyondaki U_c ifadesi, merkez pikselin düzenlilik ölçüsüdür, s değeri de özel dönüşüm fonksiyonu için kullanılmaktadır [7].

$$U_c = \sum_{i=0}^{P-1} |s(g_i - g_c) - s(g_{i-1} - g_c)| \quad (5.6)$$

$$YİÖ_{cj} = \begin{cases} \sum_{i=0}^{P-1} |s(g_i - g_c)|, & \text{eğer } U_c \leq j, j = 2 \\ P + 1, & \text{aksitaktirde} \end{cases} \quad (5.7)$$

Görüntü YİÖ ile etiklendikten sonra, histogram ile görüntüdeki verilerin dağılımı belirlenir. Tanımlama işlemi için, görüntü üzerinde genel bir histogram yerine yerel dağılımlar doğrultusunda bir histogramın oluşturulması daha başarılı bir sonuç elde edilmesini sağlayacaktır. Yerel histogram, görüntünün belli boyutlardaki (örneğin 10x10) bloklara ayrılması ve ardından her birinin histogramının çıkarılmasıyla oluşturulur. DYİÖ'lerin önemi dolayısıyla, her blok için 58 düzenli örnek dikkate alınmaktadır. Düzensiz örneklerde ise tek bir değerin tutulması yeterli olmaktadır; bu 10x10'luk bloklar için 59x100 özellik vektörünün bulunacağı anlamına gelmektedir. Alt bloklar ile elde edilen histogramların uç uca eklenmesiyle, YİÖ yerel histogramı elde edilecektir. (Şekil 5-3)



Şekil 5-3: YİÖ Histogramı

5.3. Seyrelik Kodlama

Seyrelik kodlamanın altındaki temel çıkış noktası toplam sembolüyle ifade edilebilir. Gözlem kümesinde yer alan çok boyutlu v vektörleri, W özniteliklerinin, yani taban vektörlerinin doğrusal üst düşümleri olarak modellenir ve n -boyutlu v görüntüsü olarak ifade edilir:

$$v = Wh = \sum_{i=1}^r W_i h_i \quad (5.8)$$

W 'nin r tane sütunu taban görüntüleridir [12]. Her biri n boyutlu olup; her taban görüntüsünün v 'ye katkısını ifade eden parametreleri gözlem kümesindeki her bir elemanına göre değişebilir. Kapsamlı bir taban küme bulunması durumunda, v 'nin doğru ve seyrek biçimde temsil edilmesi olanaklıdır.

N tane yüz yönteminin K adet kategoriye ayrılacağı varsayılacak olursa; her örnek için v vektörleri yöntemlerin tek tek taranmasıyla oluşturulacaktır. Bu vektörler, doğrudan veya boyutları indirgenerek ve her biri bir sütun olacak şekilde W matrisini veya farklı bir ifadeyle sözlük oluştururlar. Burada, test görüntüsü y için h vektörünün bulunması problemi (ε eşik değeri olacak şekilde) şu şekli alır:

$$\tilde{h} = \underset{h}{\operatorname{argmin}} \|h\|_1 \|Wh - y\| < \varepsilon \quad (5.9)$$

En iyi h vektörünün hesaplanmasının ardından y görüntüsü K adet sınıftan birine atanır. Bunun için iki yol mevcuttur;

- Minimum Kalan Hatası: Test görüntüsü, y sırayla her bir sınıf için tek tek yaklaşılanır. Bu, h 'de bulunan sadece bir sınıfa ait girişler kullanılarak yapılır. Yaklaşıklaşma sonuçları doğrultusunda görüntü, kalan hatanın minimuma indirildiği sınıfa atanmaktadır.

$$I(y) = \operatorname{argmin}_k r_k(y) = ||y - \tilde{y}_k||_2 \quad (5.10)$$

- Maksimum Sınıf-içi Katsayı Toplamı: Sınıflandırmanın en yüksek katsayı toplamına ait sınıf dikkate alarak yapıldığı yaklaşığıdır. h vektörünün her sınıfa ait girdileri toplanarak, görüntünün en yüksek toplam elden edilen sınıfa atanması sağlanır.

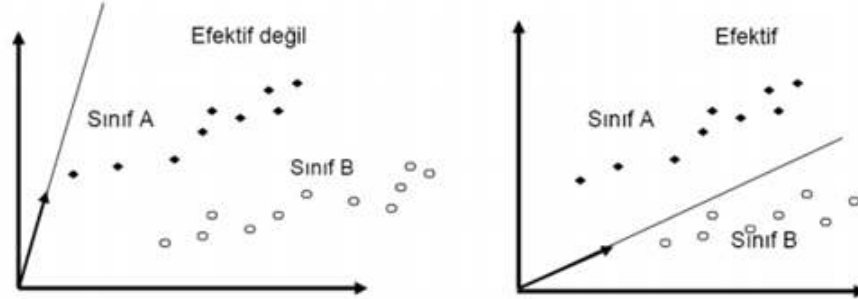
5.4. Özyüz (Eigenface – TBA)

Özyüz yönteminde yüz tanımaya yönelik Temel Bileşenler Analizi (TBA)'ne başvurulmuştur. Matematiksel olarak yüz dağıtımı ana bileşenleri ya da yüz görüntülerinin kovaryans matrisinin öz vektörleri, oldukça yüksek boyuttaki bir uzayda görüntünün nokta halinde ele alınmasıyla incelenir [21]. Düzenlenen öz vektörlerinin her biri görüntünün arasında farklı bir değişikliğe karşılık gelmektedir. Bunların, yüz görüntüleri arasındaki varyasyonun karakterize edildiği bir özellikler seti şeklinde ele alınması mümkündür. Her bir görüntünün konumu yaklaşık olarak her özvektöre katkı sağlamakta, söz konusu özvektörlerin “özyüz” adındaki, bir anlamda hayalet yüzler halinde görüntülenmesini sağlanmaktadır.

Temel Bileşenler Analizi (TBA) yöntemi, boyutun düşürülmesi ve görüntüler arası temel farklılıkların çıkarılması amacıyla kullanılır. TBA karşılıklı ilişkili bulunan verinin boyutunun, N -boyuttaki uzaydansa daha düşük boyuttaki uzayda önemli bilgi kaybı olmaksızın azaltılması için en elverişli yöntemdir. Azaltılmış alt uzay kullanılarak, veri tabanı üzerinde daha verimli hesaplamaların yapılması mümkün olmaktadır. Ortaya çıkarılan betimlemelere genellikle öz uzaylar adı verilmekte olup, yüz tanımaya yönelik uygulamalarda bunlara öz yüzler denmektedir.

Öz vektörler, özellik vektörleri oluşturmada kullanılmakta olup; sonrasında bu özellik vektörlerinin tanınması işlemlerinde de kullanılan öz vektörler gelen örnek görüntü ile en

fazla benzerlik gösteren eğitim seti ögesi tespit edilir. Tanımanın doğru şekilde gerçekleştirilmesini sağlayacak doğrultuların bulunmasında TBA'dan yararlanılır. Şekil 5-4'de bu sınıflandırma gösterilmektedir.



Şekil 5-4: Temel Bileşen Analizi Doğrusu [21]

TBA aynı zamanda sınıflar arası dağılımın maksimize edilmesini sağlamaktadır.

Özyüz yaklaşımında, her biri $N \times N$ boyuttaki görüntüler N^2 ’lik vektörlere dönüştürülmektedir.

$$\boxed{\text{GÖRÜNTÜ}} = \begin{bmatrix} I1_{11} & \dots & \dots & \dots & I1_{1N} \\ I1_{21} & \dots & \dots & \dots & I1_{2N} \\ \vdots & & & & \vdots \\ I1_{N1} & \dots & \dots & \dots & I1_{NN} \end{bmatrix} \quad II = \begin{bmatrix} I1_{11} \\ \vdots \\ I1_{1N} \\ I1_{21} \\ \vdots \\ I1_{2N} \\ \vdots \\ I1_{N1} \\ \vdots \\ I1_{NN} \end{bmatrix} \quad (5.11)$$

Yüz veri tabanı görüntülerini temsilen I kullanırken; piksel sayısı için N kullanılmıştır. Ardından veri tabanındaki görüntülerin tamamından, ortalama görüntü vektörüne ulaşılmaktadır (Eşitlik 5.11)

$$\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n \quad (5.12)$$

Burada Γ görüntü vektörü ve Ψ ortalama görüntü vektörü iken; M toplam görüntü sayısını ifade etmektedir. Bu fonksiyonda; I yüz veri tabanındaki görüntüleri ifade ederken, piksel sayısını temsilen N ve toplam görüntü sayısını temsilen M kullanılmıştır. Kolon vektörüne dönüştürülen görüntülerin pikselleri toplanarak görüntü sayısının toplamına bölünür. Ortalama görüntüyü görüntülemek için, ortaya çıkan kolon matris $N \times N$ ’lik matrise dönüştürülür.

Ortalama görüntü vektörünün, her bir görüntü vektöründen çıkarılmasıyla 0 ortalamalı bir veri tabanı matrisi (A) elde edilir.

$$\phi_i = \Gamma_i - \Psi \quad (5.13)$$

Eşitlik 5.12’de Φ ortalama görüntüden farkı alınmış görüntü vektörünü, $i=1, 2, \dots; M$ ise görüntü sayısını temsil etmektedir.

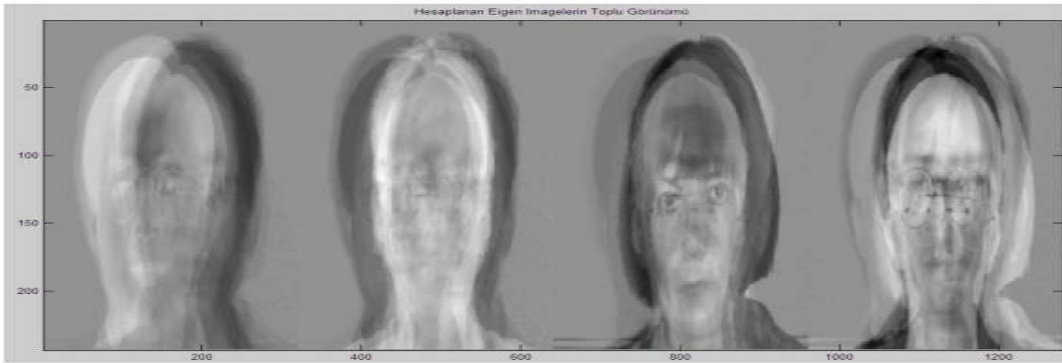
$$A = [\phi_1 \phi_2 \phi_3 \dots \phi_M] \quad (5.14)$$

Eşitlik 5.13’te A : $N^2 \times M$ boyutlu görüntü veri tabanı matrisini ifade etmektedir.

Ortaya çıkan görüntü veri tabanından özyüzlere ulaşılabilmesi için A matrisi üzerinden kovaryans işlemi yapılmalıdır.

$$C = \frac{1}{M} \sum_{n=1}^M \phi_n \phi_n^T = AA^T \quad (5.15)$$

Eşitlik 5.14’te C , kovaryans matrisini ifade etmektedir. Bu eşitlikle ulaşılan kovaryans matrisi boyutu, görüntü vektörü uzunluğuyla ilişkilidir. N^2 ’lik görüntü vektörleri üzerinden elde edilecek kovaryans matrisi $N^2 \times N^2$ boyuttadır. Resim vektör boyutunun, toplam görüntü sayısından fazla olması durumunda ($N^2 \gg M$), bu büyüklükte bir matris ortaya çıkması, özdeğer ve özvektörlerin elde edilmesi son derece zordur. $N^2 \times N^2$ ’lik kovaryans matrisi ile ulaşılacak N^2 tane özvektörlerden yalnızca $M-1$ tanesi anlamlı olacaktır. Anlamlı olan özvektör sayısının veri tabanındaki görüntülerin sayısı ile ilişkilidir. Oluşturulan diğer özvektörler, sıfır değerine sahip özdeğerlere karşılık gelmektedir. Dolayısıyla, $N^2 \times N^2$ boyutlu bir matrisin çözülmesinden $M \times M$ boyutta bir matris elde etmek ve bunu işlemek daha anlamlı olacaktır [5].



Şekil 5-5: En büyük özdeğerler için elde edilen özyüzler [21]

Elde edilen özvektörler, özyüz; özyüzlerden meydana gelen matris ise görüntü uzayı şeklinde ifade edilir. Tanımlama yapabilmek, her bir görüntünün, görüntü uzayındaki izdüşümünü bulmayı ve ağırlık vektörünü oluşturmayı gerektirir (Eşitlik 5.16).

$$\omega_k = V_k^T A \quad (5.16)$$

Eşitlik 5.16’da ω ağırlık vektörü ($1 \times M$), $k=1,2,\dots,M$ görülmektedir.

Her bir göüntü için, ağırlık vektörleri veri tabanında, Ω matrisinde saklanır (Eşitlik 5.17). Ω matrisi, tanımlamaya ilişkin esas ağırlık matrisidir.

$$\Omega = [\omega_1^T \omega_2^T \dots \omega_M^T] \quad (5.17)$$

Eşitlik 5.16’de Ω ağırlık matrisi (MxM) görülmektedir. Bilinmeyen yüzün tanınmasındaysa yüze ait görüntü eğitim setindeki gibi $N^2 \times 1$ uzunluktaki vektöre dönüştürülmektedir. Şekil 5-9’da örnek görüntüye yer verilmiştir. Görüntünün ortalama görüntüden farkı alınır (Eşitlik 5.18) her bir özyüze düşürülmesiyle (Eşitlik 5.19) ağırlık vektörüne ulaşılır (Eşitlik 5.20).

$$\Phi_T = \Gamma_T - \Psi \quad (5.18)$$

Eşitlik 4.18’de Γ_T test görüntüsü, Ψ ortalama görüntü vektörünü temsil etmektedir.

$$[\bar{t}_m] = \begin{bmatrix} t_1 & - & m_1 \\ t_2 & - & m_2 \\ \vdots & & \\ t_{N^2} & - & m_{N^2} \end{bmatrix} \quad (5.19)$$

$$\omega_k = V_k^T \Phi_T = V_k^T (\Gamma_T - \Psi) \quad (5.20)$$

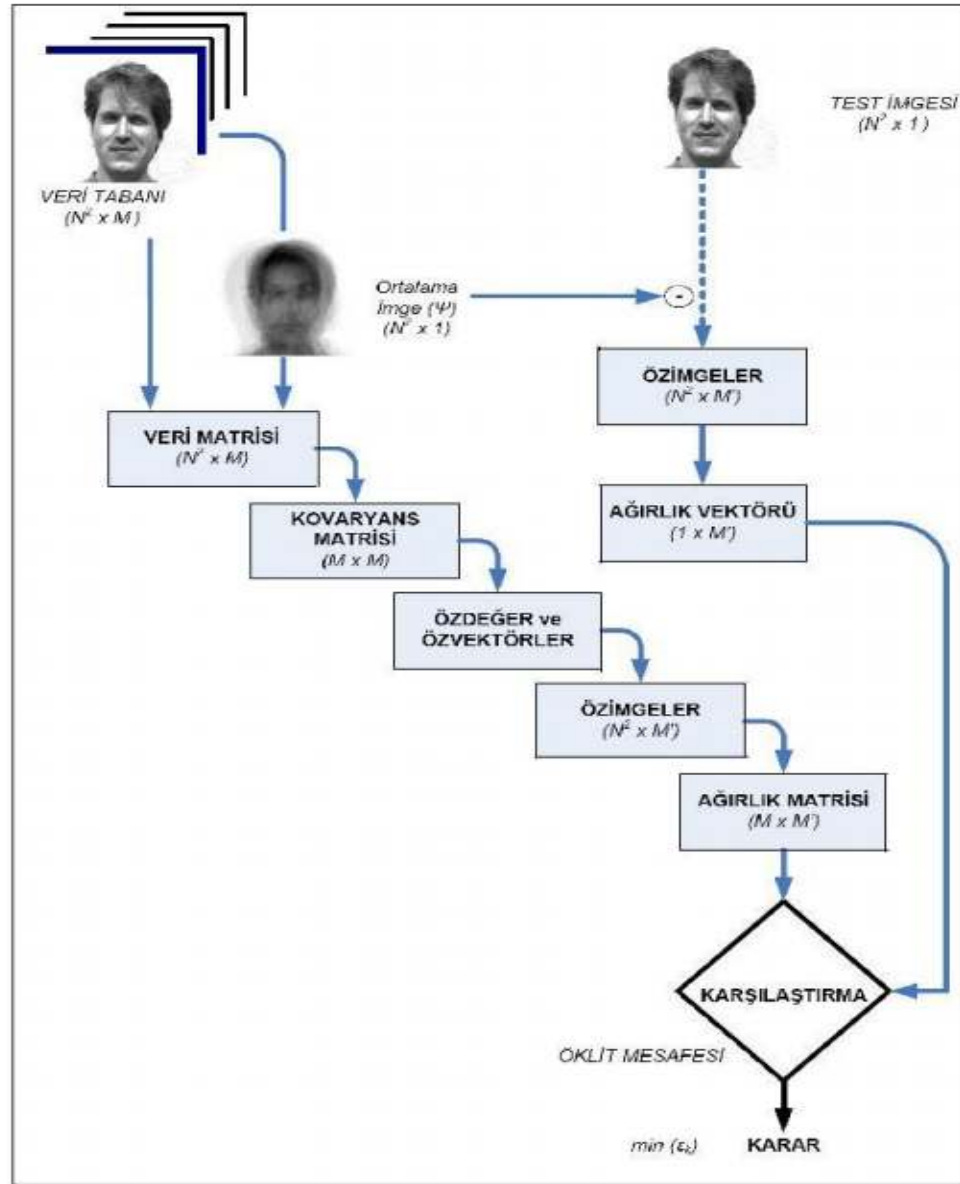
Eşitlik 5.20’da, ω her bir özyüze ait ağırlıklar, $k=1,2,\dots,M$ ’dir.

$$\Omega_T = [\omega_1 \omega_2 \dots \omega_M] \quad (5.21)$$

Eşitlik 5.21’de Ω_T test görüntüsüne ait ağırlık vektörü (1 x M)’dir. Tanımlama, görüntü uzayına düşürülen her bir görüntünün, ağırlık matrisine hangi uzaklıkta olduğuna göre gerçekleştirilir. Yakınlık ölçütü, Öklid mesafesiyle tanımlanır (Eşitlik 5.22).

$$\varepsilon_k^2 = P(\Omega - \Omega_k)P^2 \quad (5.22)$$

Ağırlık matrisine göre Öklid mesafesi elde edilen test görüntüsünün, en küçük Öklid değerine karşılık gelen görüntü olduğu kabul edilerek tanımlanması gerçekleştirilmektedir.



Şekil 5-6: Özyüz algoritması [40]

5.5. Öznitelik Seçme

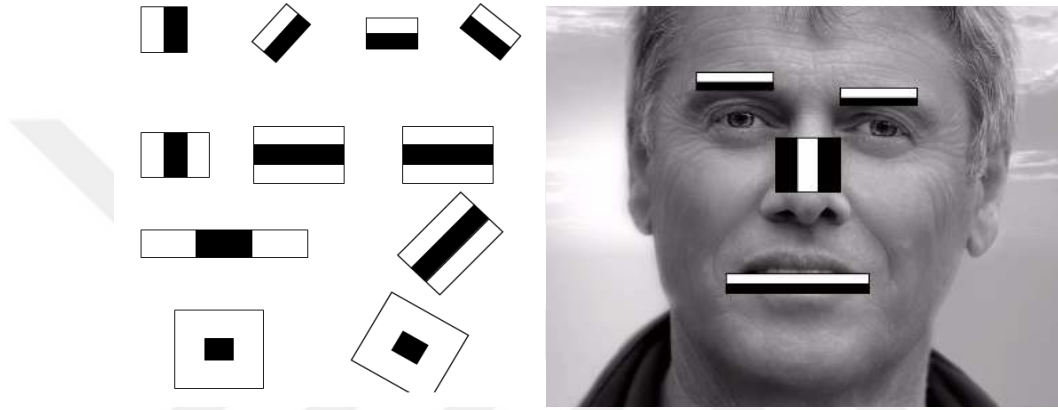
Gabor filtrelemenin neticesinde ulaşılan tüm öznitelikler, yüz ifade analizinde gerekli olmayan, ifade sınıflandırmada önemli ve faydalı olmayan bilgileri içerir. Bundan dolayı, önemli özniteliklerin seçilmesi, kullanılan sınıflandırma algoritmalarının çalışma hızına, zamandan tasarrufuna ve bellek kullanım verimine etki edecektir. Özniteliğin seçilmesiyle boyutu kısalan veri ile, ayırım gücü ve güvenilirliği yüksek esas bilgi korunmalıdır. Boyutun küçülmesi, literatürde “curse of dimensionality” adıyla tanınan “boyutun laneti” problemini de hafifletebilir. Böylece, yüksek sınıflandırma performansına sahip daha az kompleks ve daha anlaşılır sınıflandırma modelleri elde edilir.

Filtre Yöntemleri ve Sarmalama Yöntemleri olmak üzere 2 başlık denetimli öznitelik seçme yöntemlerini içermektedir. Sarmalama yöntemi algoritmalarında, seçme yöntemi bir spesifik tahmin edicinin performansının direk bir şekilde optimize edilmesini amaçlar. Her basamakta seçilen öznitelikler için tahmin edicinin genel performansı ölçülür. Filtre yöntemi algoritmalarında seçim; spesifik tahmin edici ile elde edilen performansın doğrudan optimal yapılması hedeflenmeksizin, bir önışlem gibi gerçekleştirilir. Burada genellikle özniteliklerin hedef ile ilişkisinin güçlü olması beklenir. Bu doğrultuda en çok doğrusal ilişki içinde olan öznitelikler seçilmelidir. Bir özneliğin ilgili problemde taşıdığı önem belirlenerek, önemsiz öznitelikler atılır ve öznelik sayısı böylece azaltılır [3]. Örüntü tanıma uygulamalarının birçoğunda sarmalama yöntemi algoritmalarının, filtre yöntemi algoritmaları doğrultusunda üst düzey performans gösterdiği görülmüştür. Ancak filtre yöntemi algoritmaları ile karşılaştırıldıkları takdirde, sarmalama yöntemi algoritmaları hesaplama yönünden son derece zahmetli olabilmektedir.

6. YÜZ İFADESİ SINIFLANDIRMA

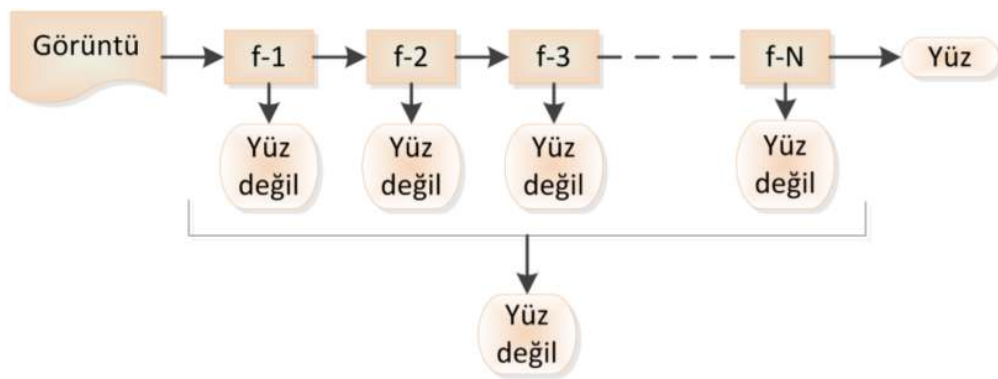
6.1. Haar Dalgacıkları

Yüz ifadesinin bulunabilmesi adına ilk önce yüz bölgesinin tespit edilmesi gerekir. Nesne tanıma amacıyla kullanılan sayısal görüntü özellikleri olan Haar filtrelerine; Şekil 6.1’de yer verilmektedir. Herhangi bir Haar özelliğinin bulunduğu, açık bölgenin piksel ortalaması koyu olan bölgenin piksel ortalaması değerinden çıkarılır.



Şekil 6-1: Haar filtreler ve uygulanması [38]

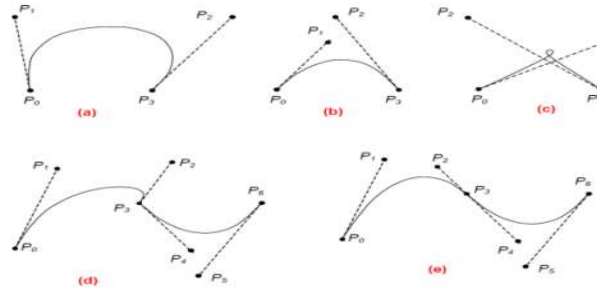
Haar özelliklerinin kullanılmasında ve eşik değerlerinin seçilmesinde, Viola-Jones makine öğrenme tekniklerinden AdaBoost’a başvurmaktadır. Bu teknikte, üçlü bir sınıflandırıcının oluşturulması çok sayıda zayıf sınıflandırıcının birleşimiyle gerçekleştirilir.



Şekil 6-2: Adaboost basamaklandırılmış sınıflandırıcı

Şekil 6-2, AdaBoost sınıflandırıcısını görselleştirmektedir. Bir bölgenin yüz bölgesi olabilmesi, tüm sınıflandırıcılarından geçmesine bağlıdır.

Bezier eğrileri, 1962 senesinde, otomobil gövdesi tasarımına yönelik olarak Fransız mühendis Bezier tarafından ortaya atılmıştır. Günümüzdeyse, yazı karakteri üretimi, modelleme, yüz ifadesi tanıma gibi çok sayıda farklı kullanım alanı bulunmaktadır. Bezier eğrisi, kendisini tanımlayacak poligonun köşeleriyle doğrudan ilişkili olup; poligonun sadece ilk ve son köşeleri eğri üzerinde bulunur. Diğer köşelerse, eğrinin derecesinin ve şeklinin belirlenmesinde katkı sağlar [10].



Şekil 6-3: Bezier eğrilerine örnekler [39]

Şekil 6-3'te farklı noktalar için Bezier eğrileri gösterilmektedir. Detaylı Bezier eğrilerinin tanımlanması, birden fazla sayıda kontrol noktasının kullanılmasıyla mümkün olmaktadır. Ancak Bezier eğrilerinde kapsamlı bir kontrolün söz konusu olması ve polinom derecesinin artmasıyla cebirsel hesaplamaların daha da karmaşık hale gelmesi sebebiyle fazla sayıda kontrol noktasının bulunması istenmez. 4 kontrol noktası bulunan Bezier eğrilerinin art arda birleştirilmesi yoluyla daha ayrıntılı eğrilerin oluşturulması mümkün olabilmektedir. Bundan ötürü, Kübik Bezier eğrilerinin (4 kontrol noktası bulunan Bezier eğrileri) kullanımının diğer dereceden Bezier eğrilerine kıyasla daha yaygın olduğu görülmektedir. Kübik Bezier eğrisinin denklemlerle ifadesi eşitlik 7.1'de verilmiştir [10].

$$P(t) = \sum_{k=0}^n P_k B_k^n(t) = (1-t)^3 P_0 + 3t(1-t)^2 P_1 + 3t^2(1-t) P_2 + t^3 P_3 \quad (6.1)$$

6.2. Otomatik Nirengi Noktası Bulma

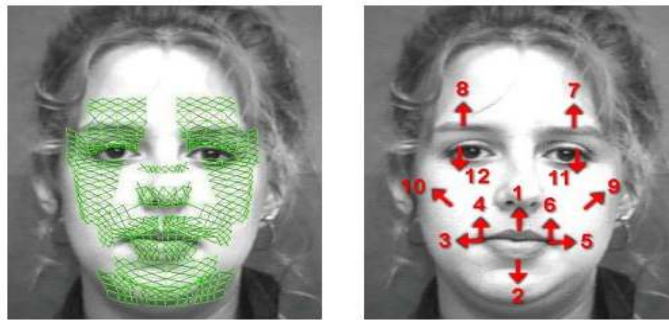
Ağız, göz kenarları, burun ucu şeklinde sıralanan 7 özneteliğin yerini belirlemek için, yüzün bulunmasının ardından nirengi noktasını bulma modülüne başvurulur. Tespit edilen nirengi noktaları Bezier hacim deformasyon izleyicisi başlatmada kullanılmaktadır. İzleyici tarafından varsayılan yaklaşık yüz biçimi, ince levha kıvrımı (TPS) algoritmasıyla tespit edilen nirengi noktalarının tanımladığı şekle bükülür. Ardından optik akış, tabanlı izleyici yüz üzerindeki nokta hareketlerini takip eder. Nihayet, ifade belirlemek için, bazı noktalarda

tanımlanan hareket birimlerinin davranışlarından yararlanılarak Bayes sınıflandırıcısına başvurulur.

İfade analizinde, yüzdeki nirengi noktalarını doğru tespit etmek büyük öneme sahiptir. Literatüre göre; nirengi noktalarının bulunması için otomatik algoritmanın bulunmadığı durumlar için, ilk hizalama elle yapılmakta olup, bu durum da gerçek uygulamalarda, tavsiye edilen sistemlerden yararlanılmasını zorlaştırmaktadır. Geniş çapta değişikliklerin söz konusu olmadığı devamlı görüntülerde, yüzün izlenmesi her kare için yeniden nirengi noktalarını bulmaya çalışmaktan daha kolay bir işlemdir. Bundan dolayı sistemimiz, otomatik nirengilemeyi yalnızca izleyici başlatılır iken yapmaktadır.

Yüzdeki nirengi noktalarının bulunmasında özniteliğe has buluşsallardan sık sık yararlanılmaktadır. Bir diğer yaklaşım nirengi noktaları arasındaki yapısal ilişkiler ile yerel öznitelik kısıtlamalarının bir arada eniyilenmesidir. Bu tip metotlar, sayıca yeterli yapısal kısıtlamanın oluşturulabilmesi adına fazla sayıda nirengi noktası (30-40) kullanmaktadır. Üçüncü bir yaklaşım, her nirengi noktası için Viola-Jones algoritmasını kullanmaktadır [6]. Kaba ölçekte algılamının böylece sağlanmasıyla, Aktif Şekil Modeli'ne başvurularak detaylı arama gerçekleştirilebilir. Nirengi noktalarının bağımsız olarak bulunması önerilmiş olup; nirengi noktası bulma yaklaşımı Viola-Jones algoritması ile yüzün tespit edilmesine dayanmaktadır.

Eğitim sürecinde her nirengi noktası sınıfı için yalnızca gerçek nirengi noktası konumlarından elde edilen doğru örnekler kullanılmaktadır. Elde edilen özniteliklerin istatistiki dağılımı artımlı faktör analizi karışımı (IMoFA-incremental mixtures of factor analysers) algoritmasıyla modellenmiş olup; söz konusu yöntemde nirengi noktalarının her biri ve Gabor dalgacıklarının her bir kanalı için ayrı karışım hesaplanır. Modelde kullanılan dağılımların sayısı ve bu dağılımların her birinde kaç faktör kullanılacağı belirlenir [4].



Şekil 6-4: (a) Bezier hacim modeli, (b) Hareket birimleri [39]

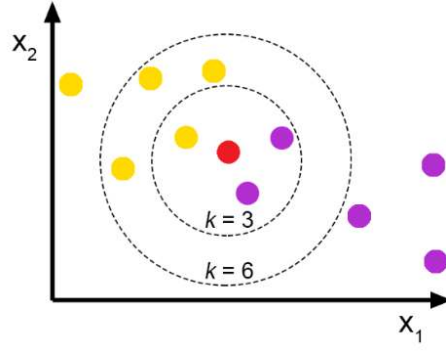
Yakın geçmişte, ifade analizine odaklanan çalışmaların sayısının arttığı görülmektedir. Bilhassa, Ekman tarafından tanımlanmış olan yüz aksiyon birimlerinin (AU-Action Unit) bulunması hedefleyen ve ifadeleri bu aksiyon birimleri ile ilişkilendiren metotlar yaratılmıştır. Söz konusu yöntemler, eğitim için, ilgili aksiyon birimlerinin uzmanlarca işaretlenmesini gerektirmektedir. Bezier hacmi kontrol noktalarının değiştirilmesiyle, yüz görüntüsüne ait şeklin değiştirilmesi mümkün olabilir.

Valenti tarafından önerilen sistem, Parçalı Bezier Hacim (PBVD) izleyicisini eğitim sürecinde öğrenilen ortalama bir yüz modeli ile başlatır. Yüz ifadesi sınıflandırmada bu yöntem taban olarak kullanılmakta olup; yüz modeli, nirengi noktalarının hizalanmasıyla iyileştirilir. Yüz modeli, TPS algoritması aracılığıyla tespit edilen nirengi noktalarına bükülür. Model üzerindeki noktalar, bulunan nirengi noktaları ile tam anlamıyla karşılaştırılır ve diğer noktaların deforme edilmesi sağlanır. İfade sınıflandırmada Bayes yöntemine başvurulmuştur. Yüz modeli üzerindeki belirli noktalardan çıkarılan nicemlenmiş değişim vektörleri sınıflandırıcıya girdi olarak sağlanır [4]. Hareket birimleri ile aksiyon birimleri arasındaki farklılık, hareket birimlerinin temel ifadelere uyarlanmış olmasından ileri gelmektedir (Şekil 6-4.b).

6.3. k-NN – En Yakın Komşuluk Algoritması

k en yakın komşuluk algoritması, gözetimli öğrenme algoritmalarından birisidir. Hem regresyon hem de sınıflandırma problemlerinin çözümünde de kullanılmaktadır. k-NN algoritması, 1960'lı yılların sonunda Cover, T. ve Hart, P tarafından ortaya çıkarılmıştır. Algoritma, sınıfları daha önceden belli olan bir veri kümesindeki verilerden faydalanmaktadır. Örnek veri setine katılacak olan yeni verinin, mevcut verilere olan uzaklığını hesaplayıp, k sayıda yakın komşuluğuna bakılarak bulunur. Uzaklık hesaplamaları için ise gelen olarak, aşağıdaki 3 tip uzaklık fonksiyonu tercih edilmektedir.

- Euclidean
- Manhattan
- Minkowski



Şekil 6-5: K-NN algoritması [23]

k-NN; en popüler makine öğrenme algoritması olarak bilinmektedir, bunun sebebi olarak ise eski, basit ve gürültülü eğitim verilerine karşı dirençli olması sebep olarak gösterilmektedir. Birçok avantajının yanında, dezavantajı da bulunmaktadır. Örneğin, herhangi bir uzaklık hesabı yapılırken, büyük kümeler kullanıldığında çok sayıda bellek alanına gereksinim duymaktadır.



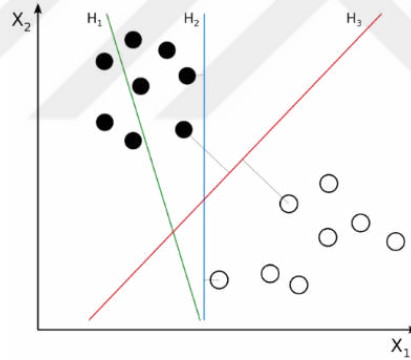
Şekil 6-6: K-NN Adımları

- Birinci adım olarak k parametresi seçilir ve bu parametre verilen bir noktaya en yakın komşuların değerini alır. Bu durumda en yakın k değerindeki komşuya göre sınıflandırma işlemi uygulanacaktır.
- Veri setine katılacak olan yeni gelen bir verinin, mevcut diğer verilere olan uzaklık değeri tek tek hesaplanır.
- İlgili uzaklıklardan da en yakın komşu k değeri seçilir. Seçilen öznelilik değerlerine göre k komşu sınıflarına ataması yapılır.
- Belirlenen bu sınıf, beklenen gözlem değerinin bir sınıfı olarak ifade edilir, böylelikle veri etiketlenmiş olur ve sınıflandırma işlemi tamamlanmış olur.

6.4. SVM – Destek Vektör Makineleri

Destek Vektör Makinesi (SVM), sınıflandırma ve regresyon zorlukları için kullanılabilen denetimli bir makine öğrenme algoritmalarındandır. SVM algoritmasında, her veri ögesini, her özelliğin değeri belirli bir koordinatın değeri olacak şekilde n-boyutlu uzayda bir nokta olarak çizilir ve daha sonra iki sınıfı çok iyi ayıran hiper düzlemi bularak sınıflandırma yapılır [37].

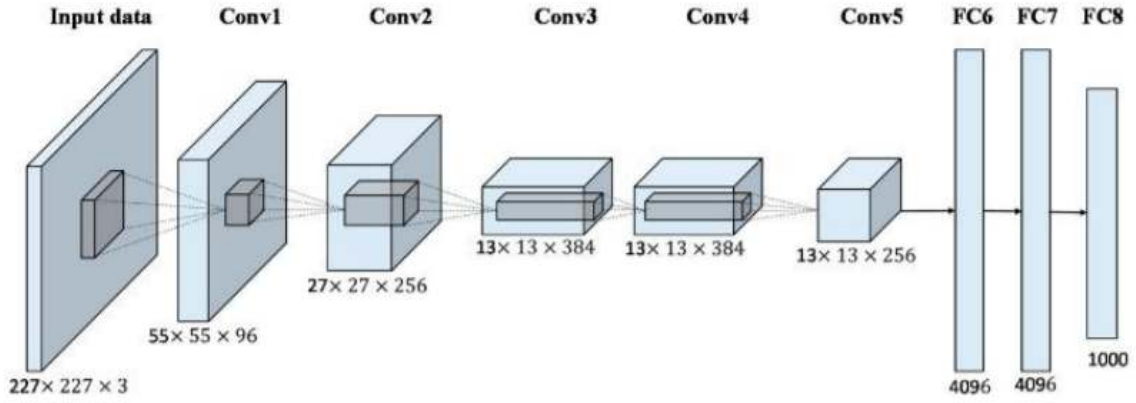
SVM, veri sınıflandırması için kullanışlı bir tekniktir. Bu sınıflandırma çoğunlukla, veri örneklerinden oluşan test ve eğitim verilerini içermektedir. Eğitim kümesindeki her örnek, bir hedef değer ve birkaç öznelik içerir. SVM'nin amacı, yalnızca öznelikleri verilen test kümesindeki veri örneklerinin hedef değerini tahmin eden bir model yaratmaktır. SVM'de sınıflandırma, denetimli öğrenmeye bir örnektir. Bilinen etiketler, sistemin doğru bir şekilde çalışıp çalışmadığını göstermeye yardımcı olur. Bu bilgi istenen bir yanıtı işaret eder, sistemin doğruluğunu onaylar veya sistemin doğru hareket etmeyi öğrenmesine yardımcı olmak için kullanılır.



Şekil 6-7: SVM algoritması [36]

6.5. AlexNet

Makine öğrenimi ve yapay zekâ, son yıllarda çok büyük ölçüde gelişim göstermiştir. Yüksek performans ve kesin sonuçlar üretmekte oldukça başarılılardır. AlexNet, ImageNet sınıflandırmasında son derece başarılı olan ilk büyük ölçekli evrimsel sinir ağı mimarisidir. AlexNet 2012'de sunulduğu ImageNet ILSVRC yarışmasında, %15,3 hata oranı ile oldukça başarılı bir performans göstermiştir [43].



Şekil 6-8: AlexNet mimarisi [44]

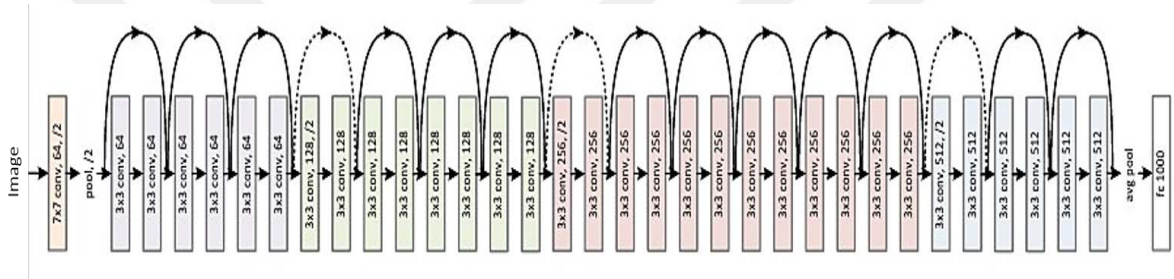
AlexNet mimarisi, 5 tane “convolutional layer” ve 3 tane “fully connected” katmandan oluşan bir evrimsel sinir ağıdır. 7 gizli katman ve bir çıkış katmanı içeren 8 katmanı bulunmaktadır. Alex Krizhevsky ve araştırmacı arkadaşları, neredeyse iki hafta süren, neredeyse bin sınıf içeren 1,2 milyon görüntü üzerinde bir sinir ağı eğitmişlerdir. AlexNet; 60.000.000 parametre, 650.000 nöron, 630.000.000 bağlantı içermektedir. [44]. AlexNet, 1,3 milyon nesne görüntüsü üzerinde eğitilmiştir ve görüntüleri binlerce farklı nesne kategorisine göre sınıflandırılmıştır. Sinir ağları insan beyninin mimarisine oldukça benzer çalıştığı için yüzlerce farklı kategoriye öğrenme yeteneğine sahiptir.

6.6. ResNet

ResNet, 2015 yılında “Deep Residual Learning for Image Recognition” makalesinde tanıtılan bir sinir ağı türüdür. Shaoqing Ren, Xiangyu Zhang, Shaoqing Ren, Jian Sun ve Kaiming He tarafından daha derin olan ağların eğitimini kolaylaştırmak için geliştirilmiştir. Derin evrimsel sinir ağları ile ImageNet sınıflandırması için birçok gelişim yaşanmıştır. ResNet ile görüntü sınıflandırma ve görüntü tanıma gibi durumlarda çok başarılı sonuçlar elde edilmiştir. Daha fazla katman eklenerek, daha kompleks sorunları çözümü için daha derin bir hale getirilmiştir.

ResNet'in ana temel ögesi “residual” bloklardır. Çok sayıda katman içeren ağda daha derine indikçe, hesaplama daha karmaşık hale gelir. Bu katmanlar birbirinin üzerine konulur ve her katman istenen işlevin bazı temel haritalarını öğrenmeye çalışır. Residual sinir ağları, atlama bağlantılarını veya bazı katmanların üzerinden atlamak için kısayolları kullanarak bunu yapar. Tipik ResNet modelleri, doğrusal olmayanlar ve arada toplu normalleştirme içeren çift veya üç katmanlı atlamalarla uygulanır. Atlama bağlantıları eklemenin iki ana

nedeni vardır; kaybolan gradyanlar sorununu önlemek ve uygun şekilde derin bir modele daha fazla katman eklemektir. ResNet kısa yol bağlantıları ile bazı katmanları atlayarak daha derin bir ağda optimizasyon/bozulma sorunu ile başa çıkarak daha derin ağlarda modelin performansını arttırmayı başarmıştır [37]. En basit durumda, yukarı akış katmanı için açık ağırlıklar olmaksızın yalnızca bitişik katmanın bağlantısı için ağırlıklar uyarlanır. Bu, tek bir doğrusal olmayan katmanın üzerine basıldığında veya ara katmanların tümü doğrusal olduğunda en iyi sonucu verir. Atlama, ilk eğitim aşamalarında daha az katman kullanarak ağı etkili bir şekilde basitleştirir. Bu, yayılmak için daha az katman olduğundan, kaybolan gradyanların etkisini azaltarak öğrenmeyi hızlandırır. Ardından ağ, özellik alanını öğrenirken atlanan katmanları kademeli olarak geri yükler. Eğitimin sonuna doğru, tüm katmanlar genişletildiğinde daha hızlı öğrenme gerçekleşir.

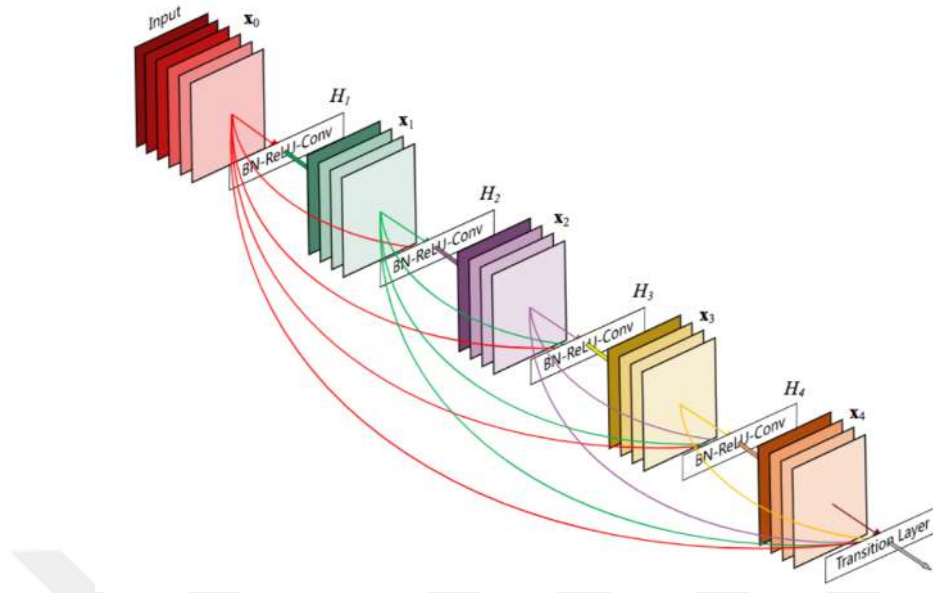


Şekil 6-9: ResNet mimarisi [45]

ResNet, ImageNet veri setinde, VGG ağlarından daha derin, ancak yine de daha düşük karmaşıklığa sahip 152 katmana kadar derinliğe sahip artık ağların değerlendirilmesi sonucunda %3,57 hata oranı ILSVRC yarışmasında birinci olmuştur [46].

6.7. DenseNet

DenseNet, yoğun bloklardan oluşan bir sinir ağıdır. Bloklar içindeki, katmanlar birbirine yoğun bir şekilde bağlıdır. Her katman, önceki katmanların çıktı özelliği haritalarından girdi alır. Residual ağların yeniden kullanımı derin bir denetim yaratır çünkü her katman bir önceki katmandan daha fazla denetim alır. Bundan dolayı DenseNet ağının diğer ağlara göre daha güçlü bir yapısı vardır.

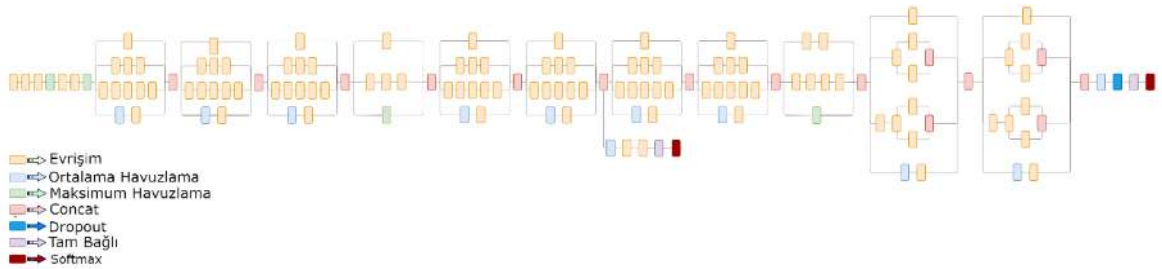


Şekil 6-10: DenseNet mimarisi [47]

Temel olarak, DenseNet ağının evrimi daha az sayıda özellik haritası oluşturur. DenseNet, geniş katmanlara daha az ihtiyaç duyar çünkü katmanlar yoğun bir şekilde bağlı olduğundan öğrenilen özelliklerde çok az fazlalık vardır. Bir katmanın çıktı özelliği haritalarının sayısı, büyüme oranı olarak tanımlanır. Sonuç olarak, büyüme oranı, her katmanın küresel olarak ne kadar yeni bilgiye katkıda bulunduğunu kontrol eder.

6.8. InceptionV3

InceptionV3 yaygın olarak kullanılan bir görüntü tanıma modelidir. 2014 yılındaki ILSVRC görsel tanıma yarışmasında %6,67 hata oranı ile birinciliği elde etmiştir [48]. Model, yıllar boyunca birden fazla araştırmacı tarafından geliştirilen birçok fikrin birleşimidir. Model, evrimler, ortalama ve maksimum havuzlama ve tamamen bağlı katmanlar dahil olmak üzere simetrik ve simetrik olmayan yapılardan oluşur. Batchnorm, model boyunca yaygın olarak kullanılır ve aktivasyon girdilerine uygulanır ve kayıp ise softmax üzerinden hesaplanır.



Şekil 6-11: InceptionV3 mimarisi [49]

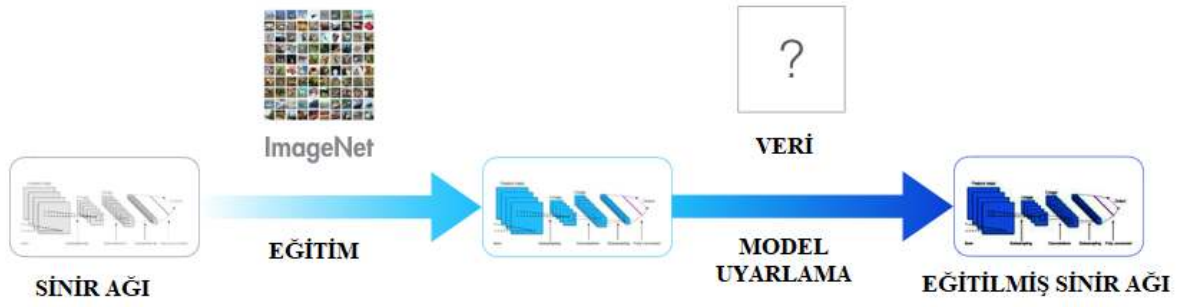
6.9. Aktarımlı Öğrenme

Önceden eğitilmiş bir modelin yeni bir problem üzerinde yeniden kullanılması, makine öğreniminde aktarımlı öğrenme olarak tanımlanır. Aktarımlı öğrenme, bir ağın başka bir ağın performansını geliştirmek için başka bir veri setinde öğrendiği bilgilerin yeniden kullanılmasıdır [54].

Bir makine, aktarım öğreniminde yeni bir görevle ilgili öngörüü artırmak için önceki bir atamadan öğrenilen bilgiyi kullanır. Bir görüntünün herhangi bir şey içerip içermediğini tahmin etmek için bir sınıflandırıcıyı eğitirken içecekleri ayırt etmek için eğitim sırasında edinilen bilgileri kullanılabilir. Halihazırda eğitilmiş bir makine öğrenimi modelinin bilgisi, aktarım öğrenimi boyunca farklı ancak yakından bağlantılı bir soruna aktarılır. Bir görüntünün gülen insan yüzü içerip içermediğini tahmin etmek için basit bir sınıflandırıcı eğitildiyse, diğer yüz ifadelerini tanımlamak için de modelin eğitim bilgisi kullanılabilir.

Aktarımlı öğrenme genel olarak duygu analizi gibi doğal dil işleme görevlerinde sıkça uygulanır. Duygu analizinde, sinir ağları tipik olarak ilk katmandaki kenarları, orta katmandaki formları ve sonraki katmanlardaki göreve özgü özellikleri algılamayı amaçlar. Erken ve merkezi katmanlar transfer öğrenmede kullanılır ve sonraki katmanlar yalnızca yeniden eğitilir. Oluşturulan yeni model, yüz ifadelerini daha önceki seviyelerde tanımak üzere eğitildiğinden, gülen ifadeyi diğer ifadelerden neyin ayırt ettiğini anlamak için sonraki katmanları yeniden eğitilir.

Derin bir ağ tarafından öğrenilen bu bilgiler başka bir ağın performansını geliştirmek ve de eğitim setine uygun olmayan bir ağı geliştirmek için model uyarlama yapılarak kullanılmaktadır. Bir veri seti için eğitilen model, daha sonra farklı bir veri seti için biraz daha eğitip kullanılan durumlar model uyarlama olarak adlandırılmaktadır. Bu yöntemde amaç; konuyu iyi öğrenmiş belirli seviyedeki modellerin üzerine yeni bir model oluşturarak önceki modelin bilgisi kullanılmaktadır [55].



Şekil 6-12: Aktarımlı öğrenme ve model uyarlama [56]

Bilgisayarlı görme uygulamaları için sıkça kullanılan ve de oldukça başarılı olan model uyarlama; bu çalışmada AlexNet, ResNet, DenseNet ve Inception evrimsel sinir ağ modelleri eğitilirken transfer öğrenmesinde kullanılmıştır. Bu modellere transfer öğrenmesinin uygulanması sırasında model uyarlaması da yapılarak, parametreler değiştirilmiş ve daha yüksek başarı oranı hedeflenmiştir.

7. UYGULAMA

Gerçek zamanlı yüz tanıma sistemi, yüz tanıma yönteminden oluşan bütünleşmiş bir sistemdir. Bölüm 2'de ki literatür taraması bölümünde sunulduğu gibi, en iyi performans sergileyen yüz tanıma yöntemleri hakkında tarama yapılarak, güçlü yanları ve dezavantajlarını özetlendi. Ardından, sistemin doğru bir şekilde çalışabilmesi için hangi yöntem ve tekniklerin kullanıldığından bahsedildi. Bu bölümde, sistemin çalıştırılmasının ardından gerçekleştirilen uygulamaların sonuçları sunuldu. Gerçekleştirilen sistemin daha iyi değerlendirilebilmesi için test veri seti gerçek zamanlı olarak uygulandı.

7.1. Geliştirme Ortamı

Bu tez çalışmasında amaç, Python programlama dili ile gerçek zamanlı bir yüz tanıma sistemi oluşturmaktır. Veri tabanı ile çalışılacağı için yazılımı koşturma üzere güçlü bir donanıma sahip dizüstü bilgisayar seçildi. En iyi sonuçları elde edebilmek adına, uygulama ve farklı parametreler ile testleri 8 çekirdekli Intel Core i7 işlemci HP dizüstü bilgisayar üzerinde Python dilini kullanarak Spyder ve Anaconda geliştirme ortamı üzerinde birkaç kez tekrarlandı.

Python, nesne yönelimli, yorumlamalı, birimsel (modüler) ve etkileşimli yüksek seviyeli bir programlama dilidir. Girintilere dayalı basit söz dizimi, dilin öğrenilmesini ve akılda kalmasını kolaylaştırır. Bu da ona söz diziminin ayrıntıları ile vakit yitirmeden programlama yapılmaya başlanabilen bir dil olma özelliği kazandırır [27].

Yüz tanıma aşaması hariç diğer tüm aşamalar sadece CPU çekirdeklerini kullanılmaktadır. Dizüstü bilgisayarın GPU'su kullanılmadı çünkü yukarıda da belirttiği gibi amaç çok az hesaplama kaynağı ile çalışabilen bir sistem geliştirmektir. Ayrıca, sistemi standart dizüstü bilgisayarda test etmek için görüntüleri ve videoları çekmek için dizüstü bilgisayardan harici olarak Logitech C270 720p web kamerası kullanılmaktadır. Kamera tam karşıda olacak şekilde ayarlanır ve de aydınlık bir ortamda bu program çalıştırılır ise uygulama daha doğru sonuçlar elde edecektir.

7.2. Veri Seti

Bu çalışmada 3 farklı veri seti kullanılmış ve her veri setindeki sonuçlar uygulama kısmında açıklanmıştır. JAFFE – FER2013 ve CK+ veri tabanlarının kullanılmasının nedeni

ise görüntü sayısı, çözünürlük ve renk gibi özelliklerin uygulamanın doğruluk oranına etkisinin gözlemlenebilmesi içindir. Tablo 3'te bu veri setlerinin özellikleri ve karşılaştırılması verilmiştir.

Tablo 3: Veri setleri karşılaştırması

	JAFPE	FER2013	CK+
Denek Sayısı	10	-	123
Görüntü Sayısı	213	35.887	593
Renk	Gri	Gri	Çoğunlukla Gri
Çözünürlük	256 x 256	48 x 48	640 x 480
Kaynak	Açık	Açık	Açık
Tür	Pozlanmış	Hem pozlanmış Hem rastlantısal	Hem pozlanmış Hem rastlantısal

7.2.1. Veri seti – FER2013

Bu tez çalışmasında modeli oluşturmak için kullanılan FER2013 veri seti, 35.887 adet tamamen rastlantısal olarak bir araya getirilmiş 48 x 48 piksel boyutlarında, gri formatta yüz fotoğrafları içermektedir. Bu fotoğrafların etiket kısımları bu veri seti içinde yer almaktadır. Daha önceki bölümlerde tanımlanmış olan, 7 adet farklı duyguları (mutluluk, üzüntü, şaşkınlık, iğrenme, öfke, korku ve nötr) içeren fotoğraflar internet ortamında herkese açık olarak kullanılabilen bir veri seti olma özelliği taşımaktadır. FER2013 veri setinde kullanılan bazı fotoğraflar Şekil 7-1'de yer almaktadır.



Şekil 7-1: FER2013 veri setinden bazı örnekler [34]

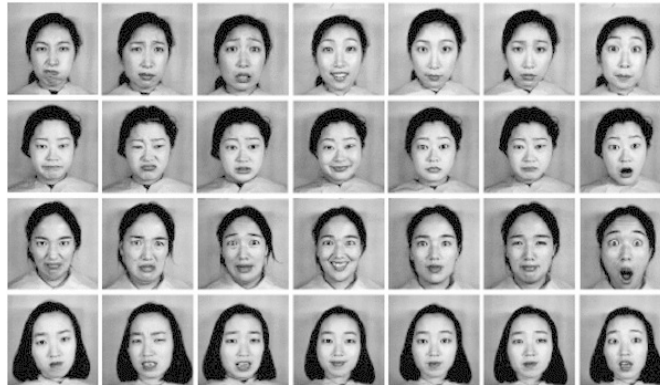
Bu veri setinde bulunan her 10 fotoğraftan 8'i eğitim, 1'i test, 1'i de doğrulama amacı ile kullanılmaktadır. Aşağıdaki tabloda FER2013 veri setinde kullanılan yüz ifadelerinin sayıları ve de toplam sayılara oranı verilmiştir.

Tablo 4: Veri setindeki imge türlerin sayı ve oranları

Yüz İfadesi	Veri Sayısı	Eğitim için kullanılan veri sayısı	Test için kullanılan veri sayısı	Doğrulama için kullanılan veri sayısı	Toplam Sayıya Oranı
Mutluluk	8.989	7.191	899	899	0,250480675
Üzüntü	6.077	4.862	608	608	0,169337086
Şaşkınlık	4.002	3.202	400	400	0,111516705
İğrenme	547	438	55	55	0,015242288
Öfke	4.953	3.962	495	495	0,138016552
Korku	5.121	4.097	512	512	0,142697913
Nötr	6.198	4.958	620	620	0,17270878
Toplam	35.887	28.710	3.589	3.589	1

7.2.2. Veri seti – JAFFE

JAFFE veri tabanı Japon Kadın Yüz İfadesi (JAFFE) olarak isimlendirilen, 10 Japon kadınının 213 yüz ifadesi görüntüsünden oluşan kamuya açık bir veri setidir. Her denek 7 temel duygu sergilemiştir ve de her ifade denek başına 3 ila 4 görüntü içeriyordur. 30 kızgın, 29 iğrenme, 33 korku, 30 mutluluk, 31 üzgün, 30 sürpriz ve 30 nötr ifadeleri içeren bu veri tabanı gri tonlamalı çözünürlüğe sahiptir ve de çözünürlüğü 256 x 256 pikseldir. Tüm yüz görüntüleri, oldukça kontrollü benzer aydınlatma koşulları altında, saç veya gözlük gibi herhangi bir çevresel faktör olmadan oluşturulmuştur.



Şekil 7-2: JAFFE veri setinden bazı örnekler [35]

7.2.3. Veri seti – CK +

Geniřletilmiř Cohn-Kanade (CK+) veri seti, eřitli cinsiyet ve ırklara sahip 18 ila 50 yařları arasında deėiřen toplam 123 farklı denekten 593 grnt ierir. Her grnt, 640x480 piksel znrlkte ve oėunlukla gri formattadır. Bu grntlerden 327'si yedi ifade “fke, ntr, iėrenme, korku, mutluluk, znt ve řařkınlık” sınıfından biriyle etiketlenmiřtir. CK+ veri tabanı, mevcut en yaygın olarak kullanılan laboratuvar kontroll yz ifadesi sınıflandırma veri tabanı olarak kabul edilir ve yz ifadesi sınıflandırma yntemlerinin oėunda kullanılır [42].

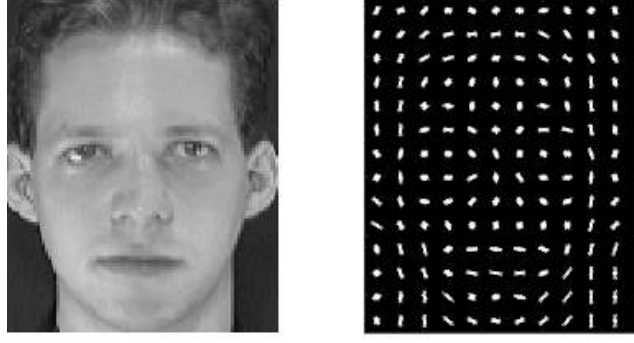


řekil 7-3: CK+ veri setinden bazı rnekler [42]

7.3. Uygulamada Kullanılan znitelik Seme Yntemi

Bu tez alıřmasında gerek zamanlı bir yz ifadesi tanıma yazılımı zerine alıřılmaktadır; bu sistemleri temelde 7 adet duyguyu "mutluluk", "znt", "řařkınlık", "iėrenme", "fke", "korku" ve "ntr" tanıma zerine kurulmaktadır. Bu yazılım, yz ifadelerinin belirlenip, karar mekanizmaları aracılıėı ile sınıflandırılmaları esasına dayanmaktadır. Yz ifadelerinin sınıflandırılmaları iin duyguları ifade eden bařlıca znitelikler ıkarılmalıdır, bu alıřmada ynlendirilmiř gradyanların histogramı (HOG - Histograms of Oriented Gradients) isimli yntemi ile bu iřlemler gerekleřtirilmiřtir.

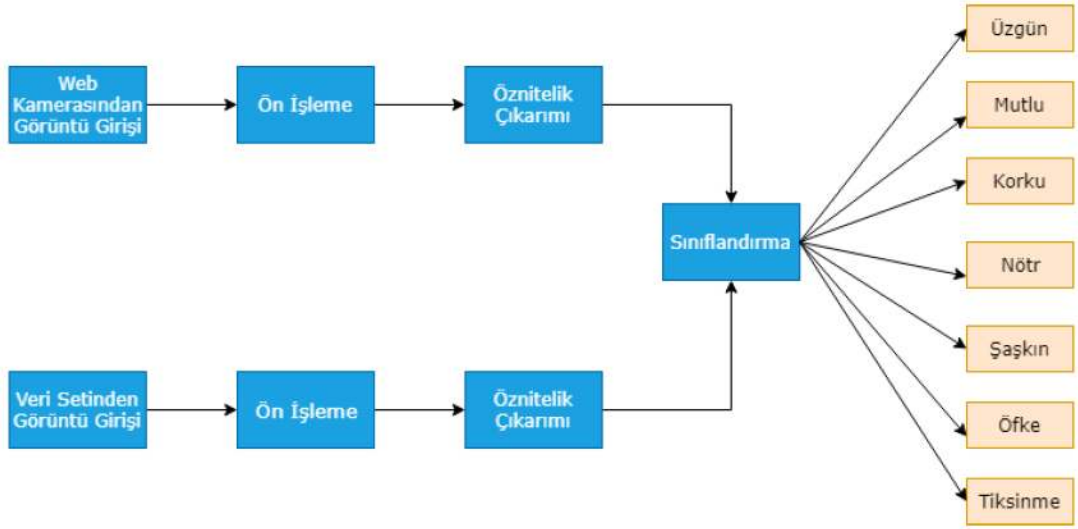
HOG, nesne algılayabilmek iin, grnt iřlemede kullanılan bir znitelik seme yntemidir, genellikle nesne tespiti iin kullanılan bir zellik tanımlayıcıdır. Yoėunluk gradyanlarının veya kenar ynlerinin daėılımına sahip olmak iin bir grnt iindeki nesnelerin zelliliėine dayanır. Gradyanlar blok bařına bir grnt iinde hesaplanır. Bir blok, blok iindeki piksel yoėunluklarındaki deėiřimin byklė ve ynnden gradyanların oluřturulduėu bir piksel ızgarası olarak kabul edilir.



Şekil 7-4: HOG örnek yüz fotoğrafı [36]

İnsan yüzlerinin genel en boy oranıyla eşleştiği için yüz görüntüleri için 128 x 128'lik bir pencere boyutu kullanılır. Tanımlayıcılar, 8 x 8 boyutlu piksel blokları üzerinden hesaplanır. 8 x 8 blok üzerindeki her piksel için bu tanımlayıcı değerler, her bir bölmenin yönlü bir gradyan açısını temsil ettiği ve aynı açığa sahip tüm piksellerin büyüklüklerinin toplamı olan bu bölmedeki değeri temsil ettiği 9 bölmeye nicelenir. Ayrıca, histogram daha sonra 16 x 16 blok boyutunda normalleştirilir, bu da ışık koşullarını en aza indirmek için 8 x 8'lik dört bloğun birlikte normalleştirilmesi anlamına gelir. Bu mekanizma, ışıktaki bir değişiklik nedeniyle doğruluk düşüşünü azaltır.

Bu tez çalışmasında, Python programlama dilinin derin öğrenme kütüphanelerinden biri olan Keras- TensorFlow kullanılmıştır. Program çalıştırıldığında web kamerasından yüz görüntülerini alır ve ön işleme (preprocessing) adı verilen bir ön işlemde geçiriliyor. Kameradan gelen görüntü ön işleme ile, gelen renkli görüntü ilk olarak siyah beyaz (gray scale) görüntüye dönüştürülür. RGB görüntülerde her pikselin 3 farklı değeri ve 2 farklı konum değeri (x,y) olacağı için gri tonlamalı görüntüler kullanılır. Gri tonlamalı görüntüler için piksel değeri, pikselin parlaklığını temsil eden tek bir sayıdır. Genel piksel formatı, bu sayının 0 ile 255 arasında değerler aldığı, 8 bitlik bir tam sayı olarak saklandığı değerdir. Genel olarak 0 (siyah), 255 (beyaz) olarak kullanılır. Renkli görüntüleri temsil etmek için, her piksel için ayrı kırmızı, yeşil ve mavi bileşenlerin belirtilmesi gerekir ve bu nedenle piksel "değeri" aslında üç sayıdan oluşan bir vektördür. Genellikle üç farklı bileşen, renk düzlemleri olarak bilinen ve görüntüleme veya işleme sırasında yeniden birleştirilmesi gereken üç ayrı "gri tonlamalı" görüntü olarak saklanır. Daha sonrasında saklanan bu görüntünün boyutu (48 x 48 ya da 256 x 256) değiştirilerek uygun boyuttaki çözünürlüğe getirilir. Çıkarılan öznitelikler sayesinde mutluluk, üzüntü, öfke, tikslenme, şaşkınlık, nötr ve korku olmak üzere yedi evrensel duygu sınıflandırılması yapılabilmektedir.



Şekil 7-5: Duygu tanıma sistemi akışı

7.4. Uygulama Sonuçları

Bu çalışmada uygulanan yönetimin, literatürdeki yöntemler ile farkının gözlemlenebilmesi için yüz ifadesi veri tabanı olarak JAFFE, FER2013 ve CK+ veri tabanı, yüz okuma ve yüz bulma işlemleri için de OpenCV kütüphanesi kullanılmıştır. Metot olarak ise k-NN ve SVM sınıflandırıcıları ve AlexNet, ResNet, DenseNet ve Inception evrişimsel sinir ağları mimarileri ile çalışmalar yürütülmüştür.

Uygulamada kullanılan hiper parametre değerleri Tablo 5'te yer almaktadır. Bu değerler uygulamada yer alan tüm modellerde kullanılmıştır. Hiper parametre, bir öğrenme algoritması eğitilmeden önce seçilen bir makine öğrenme parametresidir. Eğitim başlamadan önce doğru parametreleri seçmek önemlidir çünkü bu tür değişkenler, ortaya çıkan makine öğrenimi modelinin performansı üzerinde doğrudan bir etkiye sahiptir [57].

Tablo 5: Hiper parametre değerleri

Hiper Parametre	Değer
Eğitim Tur Sayısı (Epoch)	100
Parça Boyutu (Mini- Batch Size)	32
Momentum Katsayısı	0,9
Optimizasyon Yöntemi	ADAM
Öğrenme Hızı (Learning Rate)	0,01

Eğitim tur sayısı; bir modelin eğitilmesi sırasında, modelin ağırlıklarını güncellemek için tüm eğitim setini kullandığı bir yinelemeyi ifade etmek için kullanılan bir terimdir [49].

Eğitim aşamasında, ağırlıkların güncellenmesi bazen hesaplama karmaşıklıklarına sebep olmaktadır. Bu hesaplama işleminde veri sayısı ne kadar fazla ise hesaplama da aynı şekilde uzun sürmektedir. Bu sorunu çözebilmek için, veri setleri küçük gruplara ayrılır ve de öğrenme işlemi seçilen bu küçük gruplar üzerinde uygulanır. Bu şekilde birden fazla girdinin parçalar halinde işlenmesi “mini-batch” olarak adlandırılmaktadır [58].

Çok fazla salınım oluşturan, gürültü üreten yöntemler ağırlık gibi teknikler ile normalize edilip salınımları azaltılmaktadır, yapılan bu işlemler de momentum katsayıları ile düzenlenmektedir. Böylece gürültü ve grafikteki salınımlar azaltılarak daha hızlı bir yöntem oluşturulur [58].

Doğrusal olmayan problemlerin çözümünde en optimum değeri bulmak için optimizasyon yöntemleri kullanılır. Derin öğrenme uygulamalarında en yaygın olarak kullanılan algoritmalarından biri de “adam” optimizasyon algoritmasıdır. [58]. Optimizasyon yöntemleri adım adım işletilen bir süreçtir. Bu süreçte işletilen, adım miktarına öğrenme katsayısı denilmektedir. Öğrenme katsayısı belirlenirken en uygun değerin belirlenmesine dikkat etmek gerekmektedir. Öğrenme katsayısının çok küçük seçilmesi durumunda süreç çok uzamaktadır, büyük seçilmesi durumunda ise minimum noktanın gözden kaçırılmasına sebep olabilmektedir [59].

Literatürde yapılan çalışmalarda, hangi veri setinin, hangi öznitelik çıkarma ve sınıflandırma yöntemi ile yapıldığının karşılaştırılması Tablo 6’da yer almaktadır. Bu çalışmalarda elde edilen sonuçlarda tabloya eklenmiştir.

Tablo 6: Literatür Karşılaştırması

Çalışma Adı	Kullanılan Veri Seti	Öznitelik Çıkarma Yöntemi	Sınıflandırma Yöntemi	Gerçek Zamanlı	Doğruluk (%)
[31]	JAFFE	MLP	ANN	x	73
[31]	JAFFE	Haar	ADABOOST	x	92
[31]	GWI	Spectral Embedding	GABOR	x	62.3
[32]	CK	-	SVM	x	79

[33]	-	Optical Flow	K-NN	✓	86
[49]	JAFFE	Viola Jones	DenseNet	✓	90,18
[49]	JAFFE	Viola Jones	InceptionV3	✓	87,50
[49]	JAFFE	Viola Jones	ResNet	✓	91,52
[52]	FER2013	BPSO	ResNet	x	83,40
[53]	CK+	LBP	ResNet	✓	93,21
[53]	CK+	LBP	InceptionV3	✓	85,77

Bu tez çalışmasında, 2 farklı uygulamanın doğruluk ve süre analizi gerçekleştirilmiştir. Uygulama-1’de gerçek zamanlı olmayan 10 farklı kişiye ait 7 farklı yüz ifade gösterilmiştir. Uygulama-2’de ise gerçek zamanlı olarak, yüz ifadesinden duygu tanıma işlevi gerçekleştirilemeye çalışılmıştır.

Tablo 7’de bu uygulamalar sırasında elde edilen doğruluk oranları ve makine / derin öğrenme metotlarında her bir eğitim tur sayısında geçen süreler saniye cinsinden verilmiştir. Bütün transfer öğrenmeleri en iyi sonucu alabilmek için 100 tur sayısında gerçekleştirilmiştir. Doğruluk oranlarında ise derin öğrenme ağlarının oldukça yüksek doğruluk oranına sahip olduğu gözlemlenmektedir.

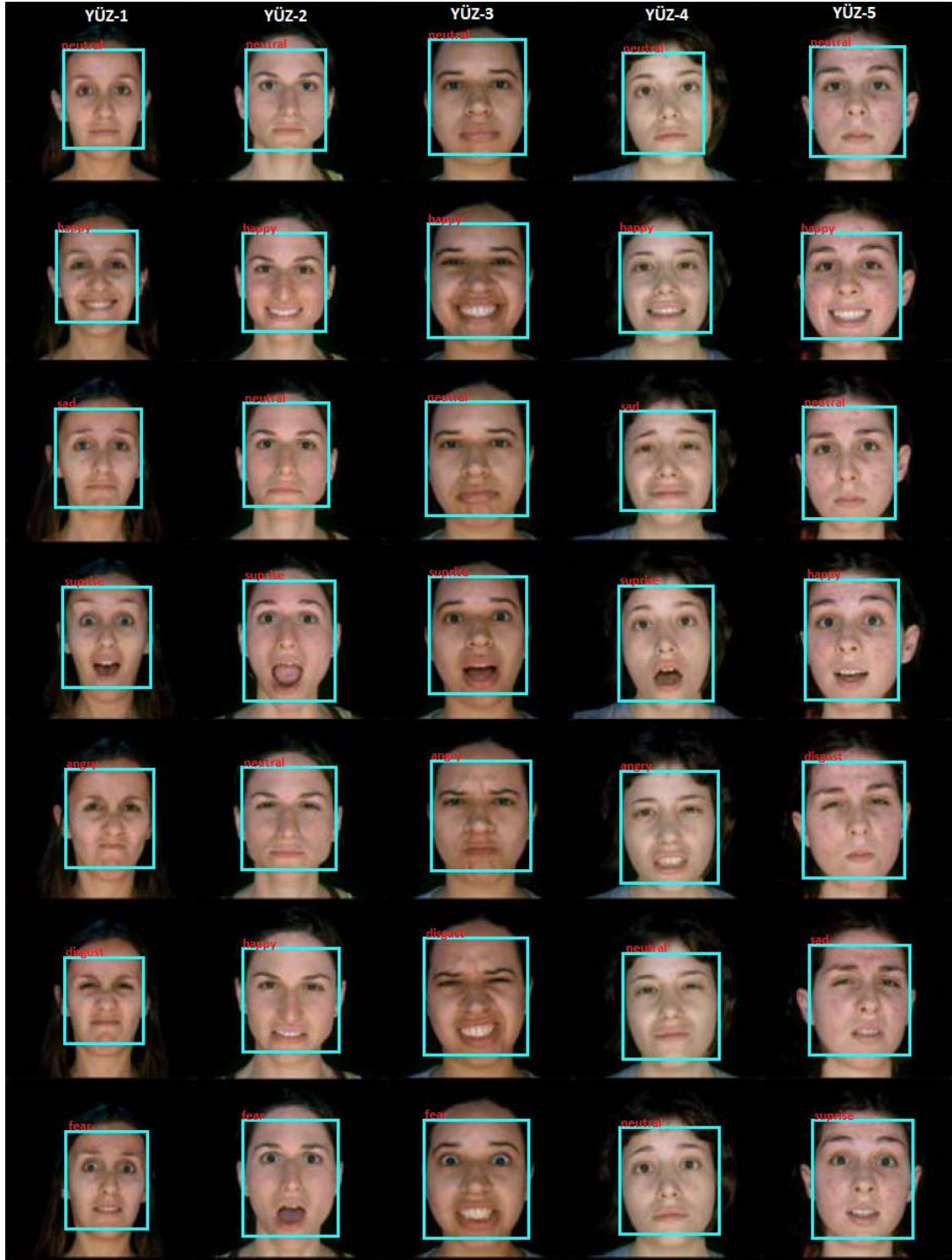
Tablo 7: Doğruluk oranları ve eğitim süreleri

Metot	Veri Tabanı	Uygulama-1	Uygulama-2	Eğitim Süresi (sn)
		Doğruluk Oranı	Doğruluk Oranı	
SVM	JAFFE	73,08%	74,80%	11
	FER2013	71,88%	73,22%	12
	CK+	72,74%	74,45%	11
KNN	JAFFE	72,58%	74,98%	12
	FER2013	77,39%	78,92%	14
	CK+	73,94%	76,71%	13

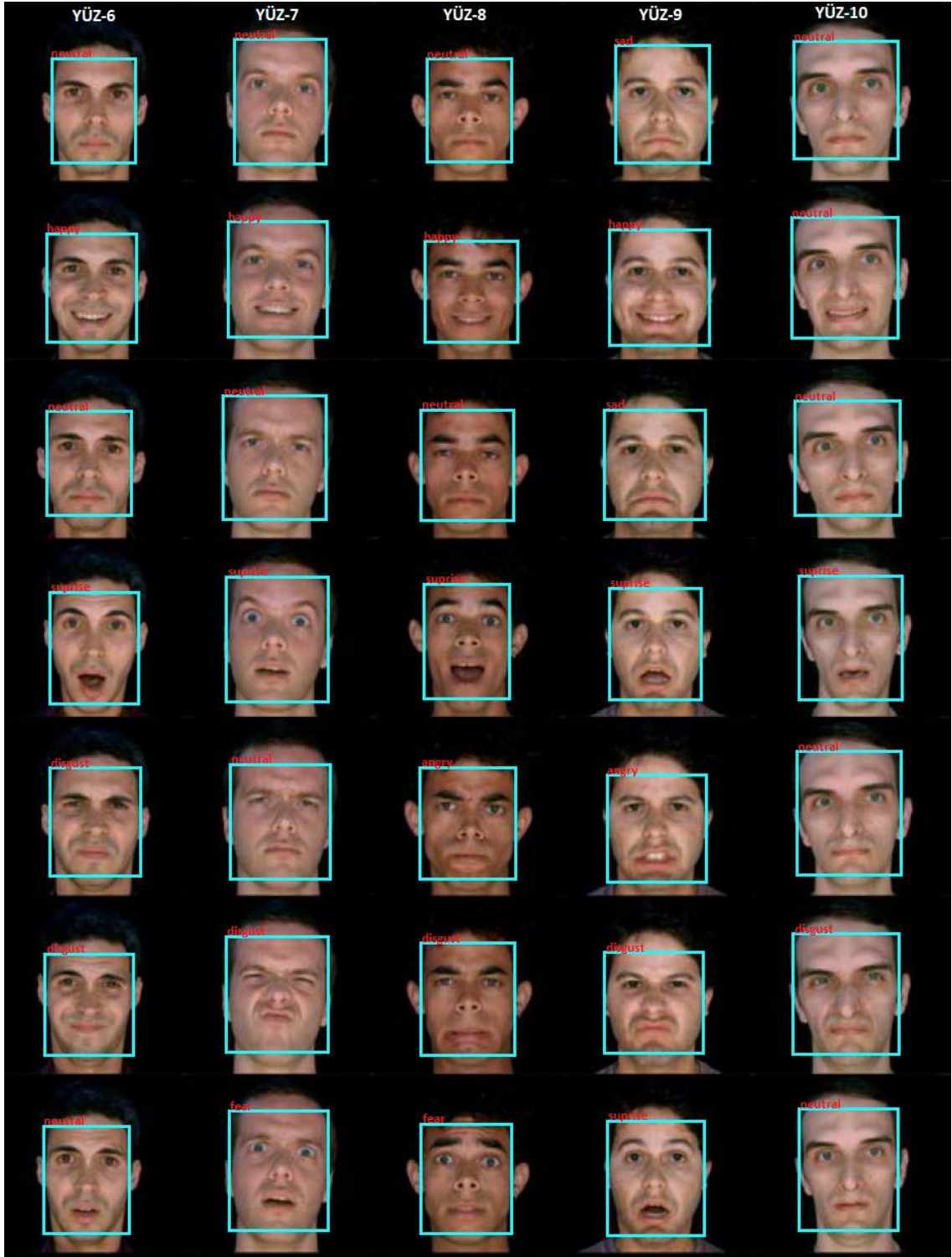
AlexNet	JAFFE	81,67%	83,48%	30
	FER2013	83,42%	84,70%	34
	CK+	82,26%	84,11%	32
ResNet50	JAFFE	90,34%	91,64%	22
	FER2013	93,85%	94,51%	26
	CK+	92,11%	93,12%	25
DenseNet121	JAFFE	89,82%	90,26%	28
	FER2013	92,04%	92,30%	31
	CK+	90,79%	91,54%	30
InceptionV3	JAFFE	84,58%	87,52%	16
	FER2013	86,73%	89,30%	19
	CK+	85,07%	87,85%	17

Tablo 7’deki değerler incelendiğinde ResNet mimarisinin görüntü tabanlı ifade tanıma sistemlerinde farklı veri setleri için en başarılı yöntem olarak elde edilmiştir. ResNet mimarisi, çok daha derinlerde düşük bir hata oranını korumaya yardımcı oldukları için en verimli sinir ağı mimarilerinden biridir. ResNet mimarisini kullanmak, daha fazla katmana sahip sinir ağlarının performansını önemli ölçüde arttırmıştır. Düz katmanlara sahip sinir ağları ile karşılaştırıldığında hata yüzdesi çok daha düşüktür. Bu nedenle, özellik çıkarma gibi derin sinir ağlarının gerekli olduğu yerlerde gerçekten iyi performans göstermektedir [51]. Tablo 6’da belirtilen literatür çalışmaları ve tablo 7’de ise bu tez çalışmasında elde edilen sonuçlar kıyaslandığında tüm veri setlerinde literatüre kıyasla yüksek başarılar elde edilmiştir. Bu tez çalışması kapsamında; gri tona çevirme, uygun boyuttaki çözünürlüğe getirme ve öznitelik çıkarma ön işleme yöntemlerinden sonra, ResNet50 mimarisi kullanıldığında başarının literatür sonuçlarına göre daha arttığı gözlenmektedir.

Bu çalışmada gerçekleştirilen gerçek zamanlı ifade tanıma sisteminin ne kadar başarılı olduğunu görmek için ilk uygulamada 10 farklı kişiye ait toplamda 70 görüntü ile bütün metotlar ve de bunların üç farklı veri tabanındaki etkileri (JAFFE – FER2013 – CK+) test edilmiştir. Şekil 7-6 ve Şekil 7-7’de gösterilen ifadeler, uygulama tarafından tespit edilen duyguları göstermektedir.



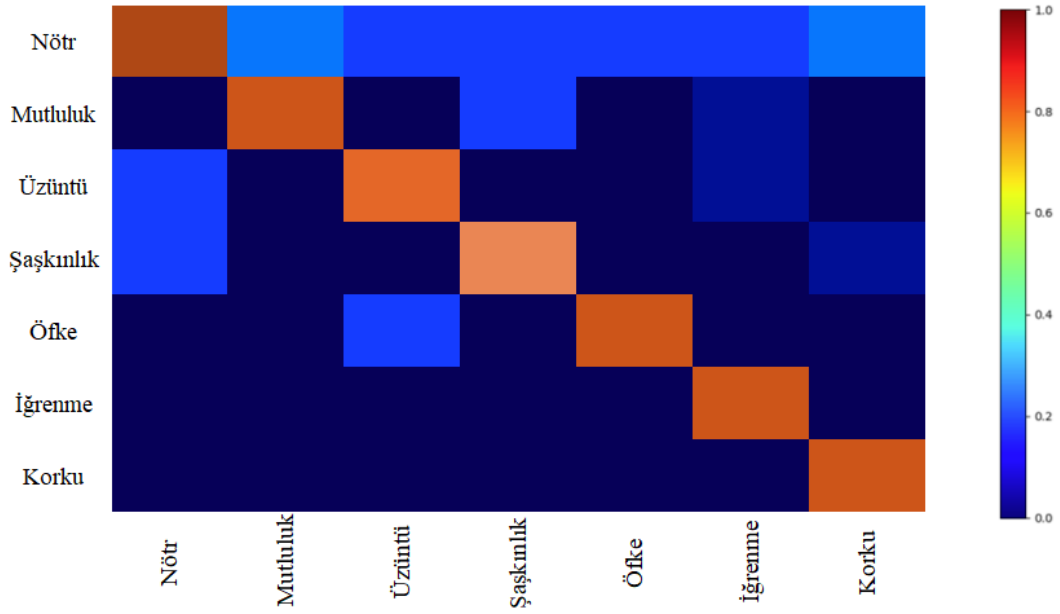
Şekil 7-6: Uygulama-1 testinde kullanılan yüzler (1-5)



Şekil 7-7: Uygulama-1 testinde kullanılan yüzler (6-10)

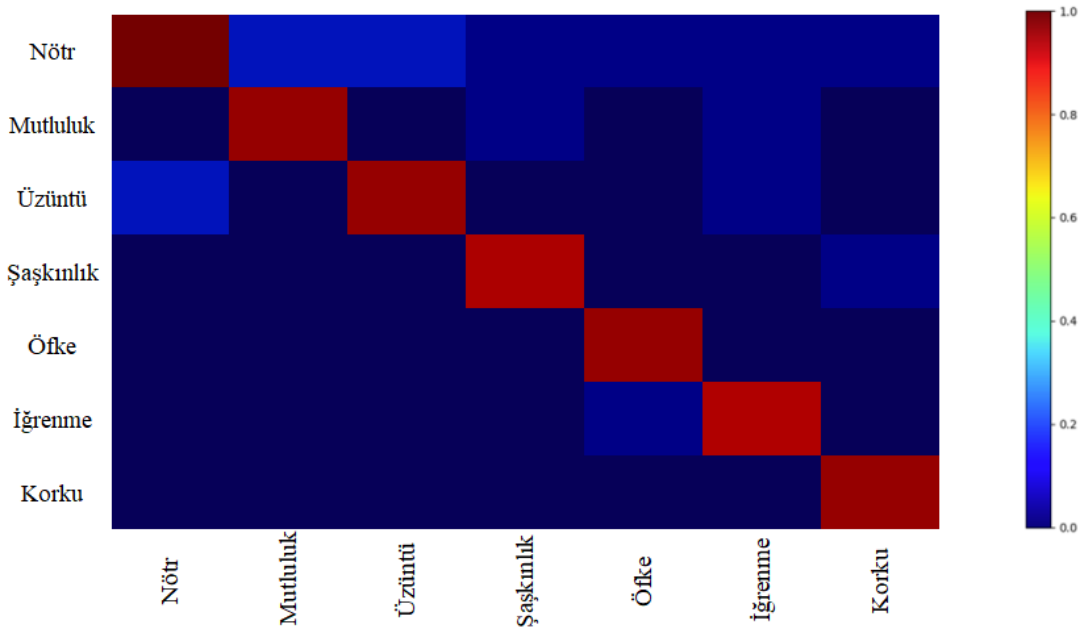
Önerilen bu model, FER2013 veri tabanı ve SVM ile sınıflandırıldığında Şekil 7-8’de gösterilen karışıklık matrisi elde edilmiştir. Model, “nötr, mutluluk ve şaşkınlık” yüz ifadesini tanımada, en yüksek başarıyı elde etmiştir. En düşük doğruluk oranı ise “üzüntü”,

“iğrenme” ve “korku” yüz ifadelerinde kaydedilmiştir. “Üzüntü”, “iğrenme” ve “korku” yüz ifadeleri test adımında en yüksek hata oranını gözlemlenmiştir.



Şekil 7-8: FER2013 – SVM sınıflandırması

Diğer bir modelde ise; FER2013 veri tabanı transfer öğrenmesi metotlarından olan ResNet ile eğitildiğinde en yüksek doğruluk oranı elde edilmiştir. Bu eğitim verisinde ise en düşük doğruluk oranı “iğrenme ve üzüntü” ifadelerine aittir, en yüksek doğruluk oranı ise “nötr” ifadesinde gözlemlenmiştir. Şekil 7-9’da da bu çalışmanın karışıklık matrisine yer verilmiştir.



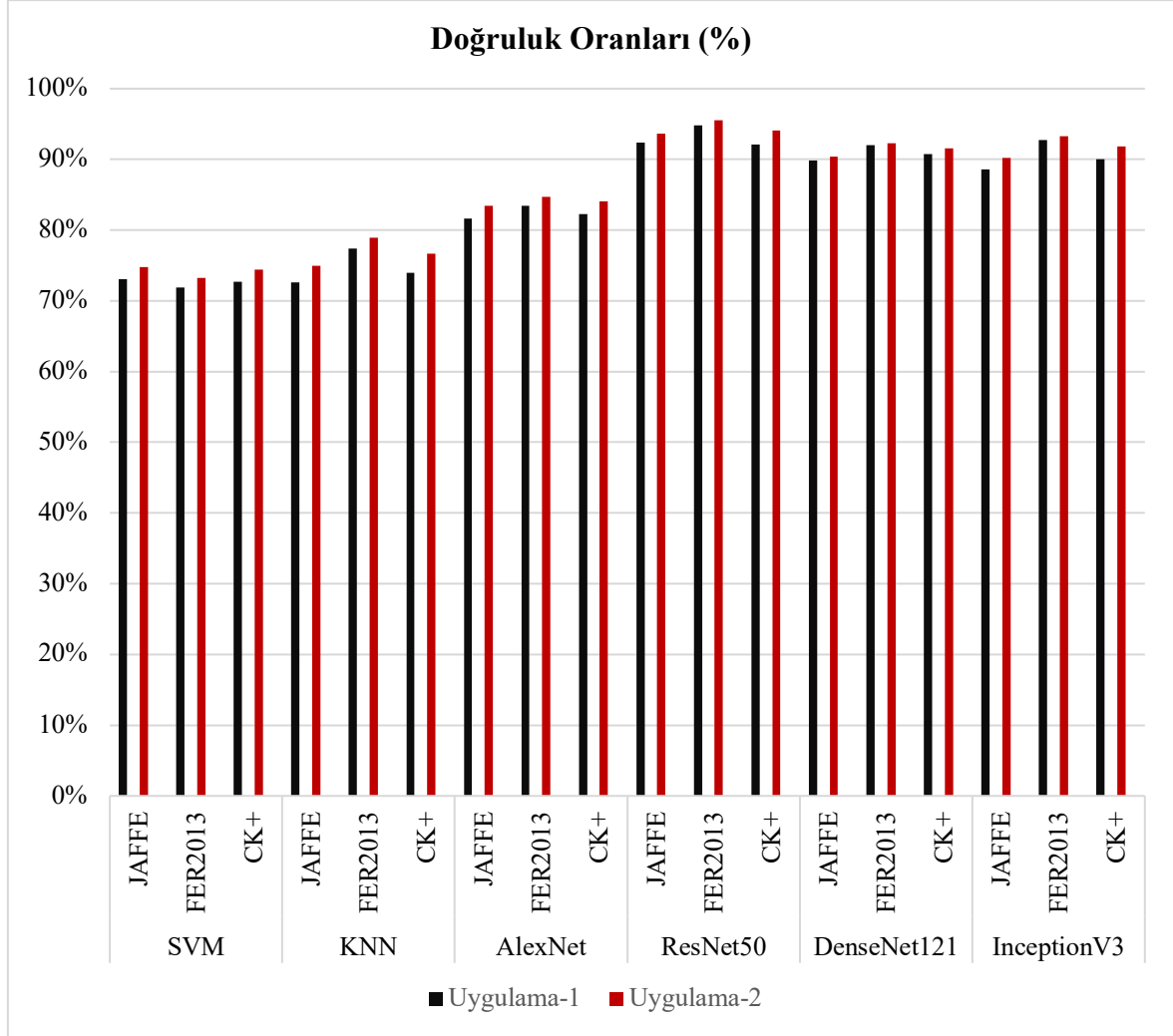
Şekil 7-9: FER2013 – ResNet sınıflandırması

Temel olarak ele alınan 7 farklı ifadeye göre doğruluk oranlarının yer aldığı karışıklık matrisi Tablo 8’de verilmiştir. Nötr ifadesi çoğunlukla en yüksek doğruluk oranına sahip ifade olarak gözlemlenmiştir. Bunun temel sebebinin ise, nötr yüz ifadesinin yüz üzerinde herhangi yüz kasını aktive etmeden oluşturulabilmesidir. Diğer ifadelerde yüzün birçok kasının aynı anda aktif olması ve kasılması gerekmektedir, bu kasılmalar her yüzde farklı şekilde olmaktadır, bu farklılıklarda veri seti büyüdükçe ve farklılaştıkça daha iyi tespit edilebilmektedir. Böylece veri seti büyüdükçe ifade tespit oranı da yükselmektedir.

Tablo 8: İfadeye göre doğruluk oranları

	Veri Tabanı	Nötr	Mutluluk	Üzüntü	Şaşkınlık	Öfke	İğrenme	Korku
SVM	JAFPE	73,63%	73,87%	71,24%	71,92%	72,94%	71,63%	73,36%
	FER2013	73,45%	71,39%	70,75%	69,83%	72,22%	72,98%	72,46%
	CK+	73,15%	73,48%	71,13%	71,26%	71,91%	71,56%	72,54%
KNN	JAFPE	73,46%	73,43%	72,91%	71,82%	72,67%	71,83%	71,92%
	FER2013	78,64%	77,84%	76,97%	76,58%	77,42%	76,24%	76,43%
	CK+	74,65%	74,18%	74,81%	72,26%	75,27%	72,74%	73,36%
AlexNet	JAFPE	82,29%	82,45%	80,91%	79,83%	80,19%	80,13%	81,59%
	FER2013	83,86%	83,89%	83,56%	81,76%	83,47%	83,61%	83,72%
	CK+	82,93%	82,79%	81,65%	81,26%	82,73%	81,88%	82,46%
ResNet50	JAFPE	90,83%	90,61%	90,14%	90,50%	89,73%	89,84%	90,89%
	FER2013	94,83%	94,42%	92,84%	92,76%	94,51%	93,79%	93,23%
	CK+	92,61%	92,17%	91,68%	91,13%	92,09%	91,49%	91,88%
DenseNet121	JAFPE	89,42%	89,35%	89,07%	88,59%	89,29%	89,12%	89,20%
	FER2013	92,18%	91,28%	91,07%	90,69%	91,68%	90,83%	92,05%
	CK+	90,12%	90,25%	90,92%	90,87%	91,26%	90,45%	91,62%
InceptionV3	JAFPE	86,17%	85,97%	85,57%	84,74%	85,21%	85,15%	85,23%
	FER2013	88,59%	88,21%	86,65%	86,41%	86,26%	85,42%	85,33%
	CK+	86,13%	85,70%	85,42%	84,69%	84,57%	84,12%	84,24%

Uygulama 2’de ise gerçek zamanlı olarak testler gerçekleştirilmiştir, kameranın çektiği video üzerinde bulunan ifadeler sonuç olarak kayıt altına alınmıştır. Uygulamamızın sonuçları göstermektedir ki, gerçek zamanlı uygulamada testler daha başarılı olmuştur. Bunun nedeni olarak ön işleme ve öz nitelik çıkarma işlemlerinin uygun koşullar yaratıldığında gerçek zamanlı uygulamalarda başarılı olarak sonuçlanmasıdır. Bu uygulamanın doğruluk oranları da Şekil 7-10’da verilmiştir.



Şekil 7-10: Doğruluk grafiği

Şekil 7-10’da yer alan grafikte; hangi metodun, hangi uygulamada daha iyi sonuç verdiği ve de hangi yöntemin daha yüksek doğruluk oranı elde ettiği ayrıntılı olarak yer almıştır. ResNet50, FER2013 veri setinde %94,51’lik doğruluk oranı ile gözlemlenmiştir. Eğitim doğruluğu, bu veri setinin eğitim için ayrılan görüntüler yüzde kaçının doğru olarak etiketlenmiş olduğunu göstermektedir.

Klasik makine öğrenme yöntemlerinin bu veri seti için hızları yüksek olsa da başarılarının derin öğrenme yöntemlerine kıyasla çok düşük olduğu gözlenmektedir. Bu tez çalışması kapsamında önerilen ön işleme aşamalarından sonra hem gerçek zamanlı uygulamada hem de gerçek zamanlı olmayan uygulamada ResNet50 mimarisi ile oldukça iyi sonuçlar elde edilmiştir.

7.5. Uygulama Kısıtlamaları

Yazılımın daha yüksek doğruluk oranını elde edebilmesi ve de daha iyi bir performans ile çalışabilmesi için, sistemde karşılaşılan bazı kısıtlama basamakları aşağıda liste halinde belirtilmiştir.

- 720p ve üzeri boyuttaki görüntülerde gerçek zamanlı uçtan uca tanıma oldukça zorlaşmaktadır. Bunun nedeni, yüz konumu adımını yürütmek için geçen sürenin platformda da önemli ölçüde uzamış olmasıdır. Bu nedenle, sistemimizin gerçek zamanlı olarak tanınması, 180p'den 720p'ye kadar olan boyutlardaki görüntülerle sınırlıdır.

- Sistem, web kamerası ile yaklaşık 50 ila 100 cm arasındaki yüzleri algılayabilir ve tanıyabilir. Bu aralığın dışında sistem düşük performans gösterir.

- Sistemde ekstra flaş veya aydınlatma sistemi kullanılmadığı için, kameranın açısı ve ortamdaki ışık miktarı da sistemin performansını oldukça etkilemektedir, en iyi performans için tepeden aydınlatma ya da gün ışığı yansıma yapmayacak şekilde kamera yerleştirilmelidir.

8. SONUÇ

Bilgisayar tabanlı ifade tanıma modelleri, gündelik hayatta geniş bir kullanım alanına sahiptir. Bilgisayar tabanlı modellerin amacına ulaşması, öngörülen koşullarda problemin çözülebilmesi adına yeterli verinin bulunmasına bağlıdır. Model başarısız olduğu takdirde, net bir sonuca varılması güç olacaktır. Programın hazırlandığı süreçte model iyi eğitilememesi veya birtakım sınırlandırıcı varsayımlarının hesaba katılmaması olasıdır. Deneyler iyi tasarlandığı takdirde, bilgisayar modelleriyle, sistemlerde kullanılabilecek önemli sonuçlara ulaşılabilir.

Biyolojik sistemler üzerinde detaylı bir incelemenin gerçekleştirilmesi, kurulacak ifade tanıma sistemlerinin daha başarılı olmasını mümkün kılabilir. Karmaşık bir sistemin ortaya konmasıyla yüz ifadelerinin tanınması probleminin bir anda çözüme kavuşturulması, doğru ara gösterimlerin seçilmesi ve sistemin aşamalı bir şekilde eğitilmesi tavsiye edilmektedir.

Bu çalışmada, yüz mimiklerine dayalı ifade durumlarını gerçek zamanlı olarak tespit edebilmek için yüz ifade tanıma sistemi geliştirilmiş ve veri tabanlarındaki görüntüler ile farklı sınıflandırma yöntemleri, derin evrimsel sinir ağlarının doğruluk oranları ve eğitilme zamanlarının karşılaştırılması yapılmıştır. Bu çalışma ortamı için, Spyder yazılımı, Keras, Tensorflow ve OpenCv kütüphaneleri bilgisayar ortamına kurulmuştur. Klasik sınıflandırma ve derin evrimsel sinir ağı yöntemlerinden k-NN, SVM yöntemleri ile, AlexNet, ResNet, DenseNet ve Inception mimarilerinin kullanılması tercih edilmiştir. Ek olarak, web kamerası ile yüz tespiti yapabilmek için Haar Cascade yöntemi kullanılmıştır. Veri tabanı olarak ise JAFFE, FER2013 ve CK+ tercih edilmiştir, bu veri setlerinin en büyük tercih sebebi açık kaynak olarak kullanılabilmesidir. Bu farklı yöntem ve veri tabanlarında yapılan uygulamalardaki sonuçlara göre, en yüksek doğruluk oranı FER2013 veri seti ile ResNet derin evrimsel ağı kullanılarak yapılan modelde gözlenmiştir. Literatür çalışmaları ve bu tez çalışmasında elde edilen sonuçlar kıyaslandığında tüm veri setlerinde literatüre kıyasla yüksek başarılar elde edilmiştir.

Bu çalışma doğrultusunda gerçekleştirilen doğruluk analizi ve performans sonuçları, daha fazla derin evrimsel sinir ağlarını, farklı veri setleri üzerinde denenerek geliştirilebilir kılınmaktadır.

KAYNAKLAR

- [1] Ayhan, S., Erdoğan, Ş. (2014) “Destek Vektör Makineleriyle Sınıflandırma Problemlerinin Çözümü İçin Çekirdek Fonksiyonu Seçimi”
- [2] Uğur, A., Kınacı, A.C., 2006, “Yapay Zekâ Teknikleri ve Yapay Sinir Ağları Kullanılarak Web Sayfalarının Sınıflandırılması”
- [3] Karaboyacı, C., 2009, Geometrik Özellik Temelli Yüz İfadesi Analizi, Yüksek Lisans Tezi, İTÜ Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı
- [4] Dibeklioglu, H., Salah, A.A., Gevers, T. (2010) "Automatic landmarking and alignment for facial expression analysis" IEEE 18th Signal Processing and Communications Applications Conference
- [5] Kaynar, O., Arslan, H., Görmez, Y., Işık, Y.E. (2018) “Makine Öğrenmesi ve Öznitelik Seçim Yöntemleriyle Saldırı Tespiti”
- [6] Özdemir, M.F. (2018) “Gömülü Sistemlerde Yüz Tanıma Algoritmalarının Karşılaştırılması”
- [7] Nabiye, V., Koçak, F. (2011) “Yerel İkili Örneklerle Yüz Güzelliği Kararı”
- [8] http://www.ibrahimcayiroglu.com/Sayfalar/Ders_GoruntuIsleme.aspx
- [9] Ekersular, M.N. (2019) “Bilgisayarlı Tomografi Görüntülerinde Özelliklerinin Belirlenmesi”
- [10] Şimşek, B. (2016) “Bezier Eğrileri ile Deforme El Modelinin Geliştirilmesi”
- [11] https://stringfixer.com/tr/Facial_Action_Coding_System
- [12] https://personel.omu.edu.tr/docs/ders_dokumanlari/5229_55322_1298.pdf
- [13] Özmen, G., Kandemir, R. (2012) “Facial Expression Detection With Haar Wavelet and Cubic Bezier Curves”

- [14] Yao, A., Shao, J., Ma, N., & Chen, Y. (2015). Capturing au-aware facial features and their latent relations for emotion recognition in the wild
- [15] Güneş, T., Polat, Ediz. (2009) "Yüz ifade analizinde öz nitelik seçimi ve çoklu SVM sınıflandırıcılarına etkisi"
- [16] Yapay Zekanın Tarihi Gelişimi- <https://teknofesor.com/yapay-zekanin-tarihi-gelisimi>
- [17] Bayrakdar, S., Akgün, D., Yücedağ, İ. (2015) "Yüz ifadelerinin otomatik analizi üzerine bir literatür çalışması"
- [18] <https://avesis.erciyes.edu.tr/resume/downloadfile/tubac?key=b2f6a348-b7a5-480e-9584-0387c7a56a0e>
- [19] Kamarol, S.K, Jaward, M.H, Parthiban, R. (2016) Spatio-Temporal Feature Extraction for Facial Expression Recognition
- [20] Alan, A., Karabatak, M. (2020) "Veri Seti- Sınıflandırma İlişkisinde Performansa Etki Eden Faktörlerin Değerlendirilmesi"
- [21] Holat, R., Kulaç, S., (2014) "ID identification by using face detection and recognition systems", 22nd Signal Processing and Communications Applications Conference (SIU)
- [22] <https://avys.omu.edu.tr/storage/app/public/selim.aras/69939/1.Hafta-Giri%C5%9F-Biyomedikal%20G%C3%B6r%C3%BCnt%C3%BCleme%20Teknikleri.pdf>
- [23] Mahmud, F., Hossain, A. (2018) Facial Region Segmentation Based Emotion Recognition Using K-Nearest Neighbors
- [24] Arı, I., Alsaran, F.O., and Akarun, L. (2011) "Vision-based real-time emotion recognition", IEEE 19th Signal Processing and Communications Applications Conference (SIU)
- [25] İnci, G. (2021) "Galton'dan Günümüze Zekâ ve Zekâ Kuramları"
- [26] Erdoğan, N., Dugelay, J.L. (2011) "Classification of facial expressions by sparse coding", IEEE 19th Signal Processing and Communications Applications Conference (SIU)

- [27] [https://tr.wikipedia.org/wiki/Python_\(programlama_dili\)](https://tr.wikipedia.org/wiki/Python_(programlama_dili))
- [28] Ekmekçi, K., Özbay, S., (2021) Emotion Analysis using Facial Expressions in Video. European Journal of Science and Technology
- [29] Wen, G., Chen, C., He, X. (2015) “Improving face recognition with domain adaptation”
- [30] Ou, W., Zhang, P., Zhu, Z. (2015) “Robust face recognition via occlusion dictionary learning”
- [31] Saxena, A., Khanna, A., Gupta, D. (2020) "Emotion Recognition and Detection Methods"
- [32] Chen, Y., Wu, H. (2018) "A Comparison of Methods of Facial Expression Recognition"
- [33] Sebe, N., Cohen, I., Huang, T.S (2004) "Multimodal Emotion Recognition"
- [34] Verma, G., Verma, H. (2020) “Hybrid-Deep Learning Model for Emotion Recognition Using Facial Expressions”
- [35] Akinrotimi, A, O. (2018) “Facial Emotion Recognition Using Principal Component Analysis and Support Vector Machine”
- [36] Islam, K., Al-Murad, A. (2017) “Performance of SVM, CNN, and ANN with BoW, HOG, and Image Pixels in Face Recognition”
- [37] <https://towardsdatascience.com/when-do-support-vector-machines-fail-3f23295ebef2>
- [38] Shetty, A, B., Roberio J. (2021) “Facial recognition using Haar cascade and LBP classifiers”
- [39] Pal, S., Biswas, P., Abraham, A. (2004) “Face Recognition Using Interpolated Bezier Curve Based Representation”
- [40] Kaya, Z., Güneş, A. (2016) “Biyometrik Güvenlik Sistemleri Ve Yüz Tanıma Dayalı Çevrimiçi Sınav Sistemi”

- [41] https://en.wikipedia.org/wiki/List_of_facial_expression_databases
- [42] <https://paperswithcode.com/dataset/ck>
- [43] <https://www.ahmetcevahircinar.com.tr/2017/08/11/imagenet-classification-with-deep-convolutional-neural-networks/>
- [44] Naeem, M., Aldabbas, H. (2020) “Facial Emotion Recognition Using Transfer Learning”
- [45] <https://becominghuman.ai/resnet-convolution-neural-network-e10921245d3d>
- [46] <https://suhedacilek.medium.com/resnet-residual-network-nedir-49105e642566>
- [47] <https://towardsdatascience.com/architecture-comparison-of-alexnet-vggnet-resnet-inception-densenet-beb8b116866d>
- [48] <https://medium.com/coinmonks/paper-review-of-googlenet-inception-v1-winner-of-ilsvlc-2014-image-classification-c2b3565a64e7>
- [49] Devrim, M., O. (2019) “Farklı Derin Evrişimsel Sinir Ağlarının Yüz İfadesi Tanıma İşleminde Karşılaştırılması”
- [50] https://en.wikipedia.org/wiki/Albert_Mehrabian
- [51] <https://www.mygreatlearning.com/blog/resnet/>
- [52] Özdemir, M., A., Elagöz, B., Alaybeyoğlu, A., Akan, A. (2019) “Real Time Emotion Recognition from Facial Expressions Using CNN Architecture”
- [53] Engin, D. (2017) “Facial Expression Pair Matching”
- [54] <https://www.analyticsvidhya.com/blog/2021/10/understanding-transfer-learning-for-deep-learning/>
- [55] Çarkacı, N. (2017) “Derin Öğrenme Uygulamalarında Başarım İyileştirme Yöntemleri”

- [56] <https://medium.com/novaresearchlab/%C3%B6%C4%9Frenme-aktar%C4%B1m%C4%B1-transfer-learning-c0b8126965c4>
- [57] <https://towardsdatascience.com/what-are-hyperparameters-and-how-to-tune-the-hyperparameters-in-a-deep-neural-network-d0604917584a>
- [58] <https://medium.com/deep-learning-turkiye/derin-ogrenme-uygulamalarinda-en-sik-kullanilan-hiper-parametreler-ece8e9125c4>
- [59] Seyyarer, E., Ayata, F., Uçkan, T., Karcı, A. (2020) “Derin Öğrenmede Kullanılan Optimizasyon Algoritmalarının Uygulanması ve Kıyaslanması”
- [60] <https://analyticsindiamag.com/why-resnets-are-a-major-breakthrough-in-image-processing/>