

T.C.  
SAKARYA ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

**KİSMİ VE TAM YÜZ GÖRÜNTÜLERİ ÜZERİNDE  
MAKİNE ÖĞRENMESİ YÖNTEMLERİYLE  
YÜZ İFADESİ TESPİTİ**

**DOKTORA TEZİ**

**İsmail ÖZTEL**

**Enstitü Anabilim Dalı : BİLGİSAYAR VE BİLİŞİM  
MÜHENDİSLİĞİ**  
**Tez Danışmanı : Prof. Dr. Cemil ÖZ**

**Ekim 2018**

T.C.  
SAKARYA ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

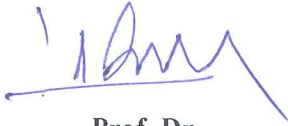
KISMİ VE TAM YÜZ GÖRÜNTÜLERİ ÜZERİNDE  
MAKİNE ÖĞRENMESİ YÖNTEMLERİYLE  
YÜZ İFADESİ TESPİTİ

DOKTORA TEZİ

İsmail ÖZTEL

Enstitü Anabilim Dalı : BİLGİSAYAR VE BİLİŞİM  
MÜHENDİSLİĞİ

Bu tez 22/10/2018 tarihinde aşağıdaki jüri tarafından oybirliği/oyçokluğu ile kabul edilmiştir.



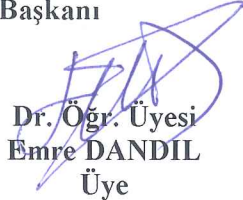
Prof. Dr.  
İbrahim ÇİL  
Jüri Başkanı



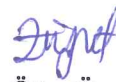
Prof. Dr.  
Cemil ÖZ  
Üye



Dr. Öğr. Üyesi  
Ali GÜLBAĞ  
Üye



Dr. Öğr. Üyesi  
Emre DANDİL  
Üye



Dr. Öğr. Üyesi  
Ziynet PAMUK  
Üye

## BEYAN

Tez içindeki tüm verilerin akademik kurallar çerçevesinde tarafımdan elde edildiğini, görsel ve yazılı tüm bilgi ve sonuçların akademik ve etik kurallara uygun şekilde sunulduğunu, kullanılan verilerde herhangi bir tahrifat yapılmadığını, başkalarının eserlerinden yararlanılması durumunda bilimsel normlara uygun olarak atıfta bulunulduğunu, tezde yer alan verilerin bu üniversite veya başka bir üniversitede herhangi bir tez çalışmasında kullanılmadığını beyan ederim.

İsmail ÖZTEL

22.10.2018

## TEŐEKKÜR

Tüm hayatım boyunca bana her türlü desteęi saęlayan aileme ve doktora süresince çalışmalarımı kararlılık ve azimle birlikte yürüttüğümüz eşim Gözde Öztel'e teşekkür ederim.

Ayrıca danışman hocam Prof. Dr. Cemil Öz'e; bilgi ve deneyimlerini benimle paylaşan Dr. Filiz Bunyak ve Dr. İlker Ersoy'a; tez çalışmalarım süresince desteęini esirgemeyen Prof. Dr. Nejat Yumuşak'a, Dr. Serap Kazan'a, Dr. Veysel Harun Şahin'e teşekkür ederim.

Bu tez çalışması TÜBİTAK tarafından BİDEB-2214A Yurt Dışı Doktora Sırası Araştırma Burs Programı çerçevesinde 1 yıl süre ile desteklenmiştir.

# İÇİNDEKİLER

TEŞEKKÜR.....	i
İÇİNDEKİLER .....	ii
SİMGELER VE KISALTMALAR LİSTESİ .....	v
ŞEKİLLER LİSTESİ .....	vii
TABLolar LİSTESİ .....	x
ÖZET .....	xii
SUMMARY .....	xiii
BÖLÜM 1.	
GİRİŞ .....	1
BÖLÜM 2.	
LİTERATÜR İNCELEMESİ .....	7
2.1. Yüz İfadesi Tespiti İçin Literatürde Önerilen Yöntemler .....	7
2.1.1. Klasik makine öğrenmesi kullanan çalışmalar .....	7
2.1.2. Derin öğrenme kullanan yöntemler .....	9
2.2. Uygulamalarda Kullanılan Veritabanları .....	10
2.2.1. Radboud Face Database (RaFD) .....	10
2.2.2. Japanese Female Facial Expression Database (JAFFE) .....	10
2.2.3. Karolinska Directed Emotional Faces (KDEF) .....	10
2.2.4. Cohn–Kanade Database (CK) .....	11
BÖLÜM 3.	
YÜZ İFADESİ TESPİTİ YÖNTEMLERİ .....	12
3.1. Çalışmalar Kapsamında Kullanılan Klasik Makine Öğrenmesi Yöntemleri .....	12

3.1.1. Yüz ifadesi tespiti için gereken adımlar .....	12
3.1.2. Görüntülerde yüz tespiti .....	13
3.1.3. Öznitelik çıkarımı .....	15
3.1.4. Öznitelik seçimi .....	18
3.1.4.1. Sequential Forward Selection – SFS .....	18
3.1.4.2. Sequential Floating Forward Selection – SFFS .....	20
3.1.4.3. ReliefF .....	21
3.1.5. Sınıflandırma .....	21
3.1.5.1. K-en yakın komşu algoritması (K-nearest neighbors) -KNN .....	22
3.1.5.2. Destek vektör makineleri (Support vector machines) -SVM .....	23
3.2. Çalışmalar Kapsamında Kullanılan Derin Öğrenme Yöntemi .....	26
3.2.1. Yapay sinir ağları .....	27
3.2.1.1. Yapay sinir hücresi yapısı .....	28
3.2.1.2. Aktivasyon fonksiyonları .....	29
3.2.1.3. Tek katmanlı algılayıcı ağ yapısı .....	31
3.2.1.4. Çok katmanlı algılayıcı ağ yapısı .....	32
3.2.1.5. Geriye yayılım algoritması .....	32
3.2.2. Evrimsel sinir ağları (Convolutional Neural Network - CNN) .....	36
3.2.2.1. Giriş .....	36
3.2.2.2. Konvolüsyon katmanı (convolution layer) .....	37
3.2.2.3. Havuzlama katmanı (pooling layer) .....	42
3.2.2.4. Dropout katmanı .....	43
3.2.2.5. Tam bağlantılı katman (fully connected layer) ve Softmax .....	43
 BÖLÜM 4.	
DENEYSEL SONUÇLAR .....	45
4.1. Giriş .....	45

4.2. Kaş ve Göz Bölgelerinden Klasik Makine Öğrenmesi Yöntemleriyle Yüz İfadesi Tespiti .....	45
4.2.1. Öznitelik çıkarımı .....	47
4.2.2. Gürbüz özniteliklerin seçimi .....	48
4.2.3. Seçilen gürbüz özniteliklerin sınıflandırılması .....	50
4.2.4. Önerilen sistemin genelleme yeteneği .....	52
4.2.5. Hata durumlarının incelenmesi .....	55
4.3. Derin Öğrenme ile Tüm Yüz Görüntüleri Kullanılarak Yüz İfadesi Tespiti .....	56
4.3.1. Kaş , göz ve ağız bölgelerinin yüz görüntülerinde bölütlenmesi .....	57
4.3.1.1. İki sınıflı bölütleme .....	60
4.3.1.2. Dört sınıflı bölütleme .....	61
4.3.2. Sınıflandırma çalışmaları .....	64
4.3.2.1. Uygulama 1 - orjinal yüz görüntülerinin sınıflandırılması .....	64
4.3.2.2. Uygulama 2 – iki sınıflı bölütleme ile elde edilen bölütlenmiş görüntülerin sınıflandırılması .....	65
4.3.2.3. Uygulama 3 – dört sınıflı bölütleme ile elde edilen bölütlenmiş görüntülerin sınıflandırılması .....	67
4.3.2.4. Uygulama 4 - orjinal görüntüler ile iki sınıflı bölütlenmiş görüntülerin birleşiminin sınıflandırılması	68
4.3.2.5. Uygulama 5 – orjinal görüntüler ile dört sınıflı bölütlenmiş görüntülerin birleşiminin sınıflandırılması	69
4.3.3. Yapılan sınıflandırmaların literatür ile karşılaştırılması .....	71

## BÖLÜM 5.

SONUÇ VE ÖNERİLER .....	73
-------------------------	----

KAYNAKLAR .....	76
-----------------	----

ÖZGEÇMİŞ .....	88
----------------	----

## SİMGELER VE KISALTMALAR LİSTESİ

AAM	: Active Appearance Model
ANN	: Artificial Neural Network
ASM	: Active Shape Model
CK	: Cohn–Kanade Database
CNN	: Convolutional Neural Network
FACS	: Facial Action Coding System
FAR	: False Accept Rate
FRR	: False Reject Rate
HOG	: Histogram of Gradient
HSV	: Hue, Saturation, Value
JAFFE	: Japanese Female Facial Expression Database
KDEF	: Karolinska Directed Emotional Faces
KNN	: K-Nearest Neighbour
LDA	: Linear Discriminant Analysis
LBP	: Local Binary Pattern
LMS	: Least Mean Square
LTP	: Local Ternary Pattern
MLP	: Multi Layer Perceptron
NN	: Neural Networks
ÖY	: Önerilen Yöntem
PCA	: Principal Component Analysis
RaFD	: Radboud Face Database
RBF	: Radyal Temelli Fonksiyon
ReLU	: Rectified Linear Unit
RGB	: Red, Green, Blue
R-CNN	: Regional Convolutional Neural Network



SDK	: Software Development Kit
SFFS	: Sequential Floating Forward Selection
SFS	: Sequential Forward Selection
SRC	: Sparse Representation Based Classification
SVM	: Support Vector Machines
YSA	: Yapay Sinir Ağları

## ŞEKİLLER LİSTESİ

Şekil 3.1. Klasik makine öğrenmesi yöntemleriyle kısmi yüz görüntüleri üzerinde yapılan yüz ifadesi tespiti uygulamasının akış diyagramı...	13
Şekil 3.2. Yüz ifadesi tespiti blok diyagramı.....	13
Şekil 3.3. Face++ tarafından bulunan 83 adet yüz anahtar noktası.....	16
Şekil 3.4. Sınıflandırma algoritmalarının genel blok diyagramı.....	22
Şekil 3.5. SVM tarafından oluşturulan hiperdüzlem.....	23
Şekil 3.6. İyi ve kötü ayırım yapan iki hiperdüzlem örneği, (a) kötü ayırım, (b) iyi ayırımı gösterir.....	24
Şekil 3.7. Doğrusal olarak ayıramayan verinin girdi uzayından daha yüksek boyuttaki uzaya aktarılması.....	25
Şekil 3.8. Derin öğrenme, makine öğrenmesi ve yapay zeka arasındaki ilişki.....	27
Şekil 3.9. Biyolojik sinir hücresi.....	28
Şekil 3.10. Tekli yapay sinir ağı modeli.....	29
Şekil 3.11. ReLU aktivasyon fonksiyonu grafiği.....	30
Şekil 3.12. Sigmoid aktivasyon fonksiyonu grafiği.....	30
Şekil 3.13. Tanh aktivasyon fonksiyonu grafiği.....	31
Şekil 3.14. Tek katmanlı algılayıcı örneği.....	31
Şekil 3.15. Çok katmanlı algılayıcı örneği.....	32
Şekil 3.16. Geriye yayılım algoritmasını anlamak için örnek bir ağ yapısı.....	33
Şekil 3.17. LeNet mimarisi.....	36
Şekil 3.18. AlexNet mimarisi.....	37
Şekil 3.19. GoogleNet mimarisi.....	38
Şekil 3.20. Konvolüsyon işleminin şekilsel gösterimi.....	39
Şekil 3.21. Konvolüsyon işlemi örneği.....	40
Şekil 3.22. Konvolüsyon işlemi sonuçlarını görmek için örnek bir girdi görüntüsü.....	40

Şekil 3.23. Örnek girdiye konvolüsyon çekirdeklerinin uygulanması sonucu oluşan katman çıktıları.....	41
Şekil 3.24. Konvolüsyon işlemi sonucu ilgili çekirdeğin kırmızı renge duyarlı olduğunu gösteren örnek konvolüsyon çıktısı.....	41
Şekil 3.25. Havuzlama işlemi örneği.....	42
Şekil 3.26. Dropout etkisi.....	43
Şekil 4.1. Önerilen sistem akışı.....	47
Şekil 4.2. Farklı öznelik seçim metotlarının sınıflandırma performansına etkileri.....	48
Şekil 4.3. Açık özneliklerinden elde edilen gürbüz öznelikler: (a) 17 ve 34. noktaların birleşmesiyle oluşan doğru ile 4 ve 30. noktaların birleşmesiyle oluşan doğrunun kesişimi ile oluşan açık; (b) 10 ve 17. noktaların birleşmesiyle oluşan doğru ile 10 ve 25. doğruların birleşmesiyle oluşan doğrunun kesişimi ile oluşan açık; (c) açık kümesinden seçilen tüm öznelikler.....	50
Şekil 4.4. (a) Yüz anahtar noktaları hatalı işaretlenmiş bir durum, (b) aynı kişiye ait farklı duyguda anahtar noktaların doğru işaretlendiği bir durum.....	56
Şekil 4.5. İki sınıflı bölütleme işleminde eğitim işlemi için bölütlenmiş görüntü oluşturma aşamaları.....	58
Şekil 4.6. Dört sınıflı bölütleme işleminde eğitim işlemi için bölütlenmiş görüntü oluşturma aşamaları.....	58
Şekil 4.7. Bölütleme işlemi için eğitim bloklarını oluşturma aşamaları. a) orijinal RGB görüntü, b) eğitim için oluşturulan maske, c) eğitime girecek blokların etiketlenmesi.....	59
Şekil 4.8. İki sınıflı bölütleme için oluşturulan CNN yapısı.....	61
Şekil 4.9. Dört sınıflı bölütleme için oluşturulan CNN yapısı.....	62
Şekil 4.10. Maske oluşturma aşamaları: Her bir bileşen türünü içeren bölütlenmiş görüntüler kendi içlerinde toplanır, ortaya çıkan ara maskelere bir eşikleme işlemi uygulanır ve gürültü giderme amaçlı maskeler oluşturulur. Ara maskeler üzerinde düşük piksel değerine sahip alanlar açık gri renkli bölgelerde koaylıkla gözlemlenebilmektedir.....	63

Şekil 4.11. Bölütlenmiş görüntünün son halinin oluşturulması.....	64
Şekil 4.12. Üç kanal orijinal görüntülerin sınıflandırılması için oluşturulan CNN yapısı.....	65
Şekil 4.13. İki sınıflı bölütlenmiş görüntünün sınıflandırılması.....	66
Şekil 4.14. Dört sınıflı bölütlenmiş görüntünün sınıflandırılması.....	67
Şekil 4.15. İki sınıflı bölütlenmiş görüntü ile orijinal ham görüntünün birleşiminin sınıflandırılması.....	69
Şekil 4.16. Dört sınıflı bölütlenmiş görüntü ile orijinal ham görüntünün birleşiminin sınıflandırılması.....	70

## TABLolar LİSTESİ

Tablo 3.1. Hesaplanan öznitelik kümelerinin eleman sayıları ve formülleri.....	18
Tablo 3.2. Uzaklık hesaplama yöntemleri ve formülleri.....	23
Tablo 4.1. Kaş ve göz bölgelerinde seçilen yüz anahtar noktaları.....	46
Tablo 4.2. Öznitelik kümeleri, seçilen gürbüz öznitelikler ve sınıflandırma sonuçları.....	49
Tablo 4.3. Açık öznitelik kümesinden seçilen gürbüz özniteliklerin KNN ile sınıflandırılmasıyla oluşan karmaşıklık matrisi.....	50
Tablo 4.4. Açık öznitelik kümesinden seçilen gürbüz özniteliklerin SVM ile sınıflandırılmasıyla oluşan karmaşıklık matrisi.....	51
Tablo 4.5. Önerilen sistem için FAR ve FRR sonuçları.....	51
Tablo 4.6. Tüm yüz özniteliklerine SFS uygulanması ile elde edilen gürbüz öznitelikler.....	52
Tablo 4.7. Önerilen yöntemin literatürdeki diğer çalışmalarla karşılaştırılması: Önerilen sistem beş duygu üzerinde çalışmaktadır, karşılaştırılan diğer çalışmalarda daha fazla ifade çalışılmış olması durumunda ya da daha az duygu çalışılmış olması durumunda başarımlar çalışmalarda verilmiş olan karmaşıklık matrislerine göre yeniden hesaplanmıştır.....	53
Tablo 4.8. Farklı veri tabanlarında eğitim ve test sonuçlarının performans analizi	53
Tablo 4.9. Önerilen sistemin genelleme kapasitesi için kullanılan veri tabanları ve özellikleri.....	55
Tablo 4.10. Bölütleme amaçlı kullanılan CNN'in konvolüsyon katmanlarına ait bilgiler.....	60
Tablo 4.11. Sınıflandırma amaçlı kullanılan CNN'in konvolüsyon katmanlarına ait bilgiler.....	64

Tablo 4.12. Orijinal yüz görüntülerinin sınıflandırılması ile oluşan karmaşıklık matrisi.....	65
Tablo 4.13. İki sınıflı bölütlenmiş görüntülerin sınıflandırılması ile oluşan karmaşıklık matrisi.....	66
Tablo 4.14. Dört sınıflı bölütlenmiş görüntülerin sınıflandırılması ile oluşan karmaşıklık matrisi.....	68
Tablo 4.15. İki sınıflı bölütlenmiş görüntü ile orijinal ham görüntünün birleşiminin sınıflandırılması ile oluşan karmaşıklık matrisi.....	69
Tablo 4.16. Dört sınıflı bölütlenmiş görüntü ile orijinal ham görüntünün birleşiminin sınıflandırılması ile oluşan karmaşıklık matrisi.....	71
Tablo 4.17. Önerilen yöntemlerin literatürdeki çalışmalar ile karşılaştırılması.....	72

## ÖZET

Anahtar kelimeler: Yüz ifadesi tespiti, derin öğrenme, makine öğrenmesi, bilgisayar görmesi, CNN

Yüz ifadeleri insanlar arası iletişimin önemli bir parçası olduğu gibi insan makine etkileşiminde de önemli rol oynamaktadır. Suçlu tespiti, sürücü dikkatinin izlenmesi, hasta takibi gibi önemli konularda karar vermede yüz ifadesi tespiti kullanılmaktadır. Bu sebeple, yüz ifadelerinin sistemler aracılığı ile otomatik tespiti popüler bir makine öğrenmesi çalışma alanıdır. Bu tez çalışmasında yüz ifadesi sınıflandırma çalışmaları yapılmıştır. Yapılan yüz ifadesi tespiti uygulamaları genel olarak iki başlık altında toplanabilir. Bunlardan ilki kısmi yüz görüntülerinin klasik makine öğrenmesi yöntemleriyle analizi ve ikincisi ise tüm yüz görüntülerinin derin öğrenme yöntemleri ile analiz edilmesidir.

Geliştirilen ilk uygulamada, yüz görüntülerinden duygu tespiti için literatürdeki çalışmalardan farklı olarak sadece göz ve kaşların bulunduğu bölgeler kullanılarak sınıflandırma yapılmış ve yüksek başarımlar elde edilmiştir. Önerilen bu yöntem sayesinde yüz ifadesi tespitleri alt yüz kapanmalarından veya ağız hareketlerinden etkilenmeyecek, gürbüz özniteliklerin seçimi ile daha az öznitelikle sınırlı kaynaklara sahip cihazlarda çalışabilecek niteliktedir. Ayrıca önerilen sistemin genelleme yeteneğinin yüksek olduğu karşılaştırmalı olarak deneysel çalışmalarla ortaya konulmuştur.

Tez kapsamında yapılan diğer yüz ifadesi sınıflandırma çalışmaları tüm yüz görüntüleri kullanılarak derin öğrenme yöntemleri ile gerçekleştirilmiştir. Önerilen yaklaşımlardan birisi yüz bölütleme çalışmasıdır. Bu çalışmalar ile elde edilen bölütlenmiş görüntüde yüz ifadesi ile ilgili öznitelikler korunmakta, kişisel herhangi bir veri saklanmamakta ve böylece kişisel gizlilik de korunmuş olmaktadır. Ayrıca bölütlenmiş görüntü ile orijinal yüz görüntüsünün birleşimi; yüz ifadesi için önemli olan kaş, göz ve ağız bölgelerine odaklanılarak yüz ifadelerinin tanınma başarımının artırılması sağlanmıştır.

# **FACIAL EXPRESSION RECOGNITION ON PARTIAL AND WHOLE FACE IMAGES WITH MACHINE LEARNING METHODS**

## **SUMMARY**

Keywords: Facial expression recognition, deep learning, machine learning, computer vision, CNN

Facial expressions are important for interpersonal communication also play an important role in human machine interaction. Facial expressions are used in many areas such as criminal detection, driver attention monitoring, patient monitoring. Therefore, automatic facial expression recognition systems are a popular machine learning problem. In this thesis study, facial expression recognition studies are performed. In general, the applications of facial expression recognition can be grouped under two topic in this thesis: analysis of partial facial images with classical machine learning methods and analysis of whole facial images with deep learning methods.

In the first application, classification of the facial expressions from facial images was performed using only eye and eyebrows regions. This approach is different from the studies which are studied facial expression recognition in the literature and high success rate was achieved. With this approach, proposed system is more robust for under facial occlusions and mouth motion during speech. Further, according to our experiments, the generalization ability of the proposed system is high.

In this thesis, the rest of the facial expression recognition applications was developed with whole face images using deep learning techniques. One of the proposed methods is segmentation of facial parts with CNN. After segmentation process, facial segmented images were obtained. With this segmented images, personal privacy is protected because the segmented images don't include any personal information. Also, the success rate of the classification was increased with combining original raw image and segmented image. Because; eyes, eyebrows and mouth are crucial for facial expression recognition and segmented images have these areas. Therefore, the proposed CNN architecture for classification forces the earlier layers of the CNN system to learn to detect and localize the facial regions, thus providing decoupled and guided training.



## BÖLÜM 1. GİRİŞ

Teknolojik ilerlemenin hızla değişim gösterdiği günümüzde donanımların gelişmesi, teknolojik ürünlerin hayatımızın birçok alanında kendine yer bulmasını tetiklemektedir. Görüntü alma sistemlerinin kalitesinin artması, bu görüntüleri işleyecek güçlü işlemcilerin yaygınlaşması, bilgisayar görmesi alanındaki ilerlemelere de katkı sağlamaktadır. Bu alanın alt dallarından biri ise yüz görüntüleri üzerinde yüz ifadesi tespittir. Yüz ifadeleri insan davranışları hakkında pek çok bilgi içermektedir [1]. Araştırmalara göre yüz yüze iletişim esnasında mesajı iletmede yüz ifadesi, kullanılan kelimelerden daha etkilidir. Mehrabian bir çalışmasında yüz yüze iletişimde mesajın iletilmesinde kelimelerin %7, ses tonunun %38, beden dilinin ise %55 oranında etkili olduğunu vurgulamıştır [2]. Beden dilinin içinde göz, el ve baş hareketleri, görünüm, duruş ve yüz ifadeleri sayılabilir.

Yüz ifadesi tespiti insan beyninin zorlanmadan yaptığı günlük aktivitelerden biri olmasına rağmen makine öğrenmesi yolu ile tespit edilmeye çalışıldığında literatürdeki zor problemlerden biri haline gelmektedir. Aynı kişinin aynı yüz ifadeli görüntüsü farklı parlaklık, arkaplan ve açılarda analiz açısından çeşitlilik gösterir [3]. Bu nedenle yüz ifadesi tespiti bilgisayar görmesinin ilgi çekici ve zorlayıcı konularından biridir. Yüz ifadesi tespiti; robotik [4]–[8], sağlık [9], [10], bilgisayar oyunları [11], trafik [12], [13], medya [14], [15], sanal gerçeklik [16]–[18], vb. alanlarda eğlence, konfor, güvenlik ve hayatı kolaylaştırma gibi birçok nedenle kullanılmaktadır.

İnsanlar arası iletişimin temel taşlarından biri olan yüz ifadeleri, yüz derisi altında bulunan kasların farklı kombinasyonları ile ortaya çıkmakta olup, sözsüz iletişimin varyasyonlarından biridir [19]. Yüz derisi altında bulunan 20 kas grubu farklı kombinasyonlarla kasılıp gevşeyerek yaklaşık 250.000 adet ifade ortaya

çıkarmaktadır [20]. Yüz üzerinde, ifadelerin en yoğun olduğu yerler göz çevresi kaslarının olduğu bölgelerdir hissedilir [9]. Yapılan araştırmalara göre yüz üzerinde beliren ifadelerin büyük çoğunluğu yalnızca gözlerin bulunduğu bölgeye bakılarak anlaşılabilir [21]. Örneğin kızgınlık, şaşkınlık, mutluluk gibi duygular sadece gözlere bakılarak anlaşılabilir [21]. Kaşlar da gözlerdeki ifadeleri destekler ve güçlendirir [21].

Bu çalışmada makine öğrenmesi yöntemleri kullanılarak yüz görüntüleri üzerinde yüz ifadesi tespiti uygulamaları yapılmıştır. İlk deneysel çalışmada kısmi yüz görüntüleri üzerinde geometrik öznitelikler kullanılarak, Sequential Forward Selection (SFS) ile öznitelik seçimi ve Support Vector Machines (SVM – Destek Vektör Makineleri) ile sınıflandırma işlemi gerçekleştirilmiştir. Diğer uygulamalarda ise tam yüz görüntüleri kullanılarak Convolutional Neural Network (CNN – Evrimsel Sinir Ağları) yapıları kullanılarak yüz ifadesi tespiti gerçekleştirilmiştir.

Deneysel olarak yapılan ilk uygulamada yüz görüntülerinden kaş ve göz bölgelerine odaklanılmıştır. Kaşlar ve gözlerdeki anahtar noktalar tespit edilip, bu noktaların her birinin diğerlerine olan uzaklıkları, her iki noktanın birleşmesi ile oluşan doğruların eğimleri, oluşan her bir doğru parçasının diğerleri ile oranları ve her iki doğru parçasının kendilerini ya da uzantılarının kesişmeleri ile oluşan açılar öznitelik olarak hesaplanmıştır. Ardından bu özniteliklerden yüz ifadelerine en güçlü şekilde etki edenleri öznitelik indirilmesiyle seçilmiştir. Son olarak seçilen gürbüz öznitelikler kullanılarak yüz ifadeleri sınıflandırılmıştır.

Kısmi yüz görüntüleri üzerinde yapılan çalışmanın literatüre katkıları aşağıdaki gibi verilebilir:

- Tüm yüze odaklanan yüz ifadesi tespiti çalışmalarında alt yüz bölgesinin atkı vb. nesnelere ile kapanması veya konuşma esnasında ağız hareketleri duygu tanımayı zorlaştırabilmektedir. Bu çalışma kapsamında yüz görüntülerinde sadece kaş ve göz bölgelerine bakıldığından önerilen sistem, alt yüz bölgelerinde meydana gelebilecek ifade tanımayı zorlaştıracak etkenlere karşı güçlüdür.

- Uygulama kapsamında çıkarılan çok sayıdaki öznitelik, sonuca olumlu anlamda etki edenlerin seçilmesiyle daha az öznitelige düşürülmüştür. Böylece az sayıdaki öznitelik ile yüksek başarı oranında yüz ifadesi tespiti mümkün olmuştur. Ayrıca az sayıda öznitelik kullanımı ile işlem maliyeti düşmüş, uygulama zamanı kısalmış ve uygulamanın gerçek zamanlı çalışma kabiliyeti artmıştır. Bu nedenle önerilen yöntem, sınırlı hafıza ve işlem gücü kaynaklarına sahip mobil platform vb. sistemler üzerinde de kolaylıkla gerçekleştirilebilir.
- Öznitelik seçim yöntemi ile seçilen gürbüz öznitelikler sayesinde önerilen yöntem başka yüz veritabanları kullanıldığında da yüksek genelleme oranına ulaşmaktadır.

Bu tez kapsamında yukarıda anlatılan geometrik öznitelikler ile sadece göz bölgesinden yüz ifadesi tespitine ek olarak derin öğrenme yöntemleri ile tüm yüz görüntülerinden yüz ifadesi tespiti üzerinde deneysel çalışmalar da yapılmıştır. Ayrıca yüz ifadesi başarımını arttırmak amacıyla yüz görüntülerinde kaş, göz ve ağız gibi yüz ifadesi tespitine etki eden yüz bileşenlerini öne çıkarmak için yüz bölütleme işlemleri yapılmıştır. Ardışık yapay ağlar, yüz bölütleme ve yüz ifadesi sınıflandırma amaçlı kullanılmıştır. Ham yüz görüntüleri, bölütlenmiş yüz görüntüleri ve bu iki görüntünün birleştirilmesi ile elde edilen veriler farklı ağ yapıları üzerinde yüz ifadelerini sınıflandırmak için kullanılmıştır.

Tüm yüz görüntüleri üzerinde derin öğrenme yöntemleri ile yapılan çalışmanın literatüre katkıları aşağıdaki gibi verilebilir:

- Bu bölümde geliştirilen uygulamalardan biri de yüz ifadesi tespiti için yüz görüntülerinde bölütleme işlemidir. Yüz görüntülerinden kaş, göz ve ağız bölgelerinin bölütlenmesi ile kişisel bilgiler içermeyen fakat bu yüz bileşenlerini içeren özet bir görüntü elde edilmiştir. Böylece yüz ifadelerine dair bilgiler depolanmak istendiğinde bu bölütlenmiş görüntüler depolanabileceği gibi, kişisel bir veri içermediğinden kişisel gizlilik de korunmuş olacaktır.

- Bölütlenmiş görüntüler kaş, göz ve ağız bilgilerini parçalı olarak içerirken orijinal yüz görüntüleri bütünsel bilgiler içermektedir. Bölütlenmiş görüntü ve orijinal yüz görüntüsünün birleştirilmesi ile elde edilen görüntü sistemin başarı oranını arttırmıştır.
- Kaş, göz ve ağız bileşenlerinin koordinat ve şekil bilgilerini içeren bölütlenmiş görüntüler orijinal yüz görüntüleri ile birleştirilirler. Oluşan bu yeni veriler üzerinde çalışan yapay ağız yüz ifadesi tespiti için önemli olan kaş, göz ve ağız bölgelerine yoğunlaşmaya zorlanır ve böylece rehberli bir eğitim mümkün olur.
- Önerilen yöntem ile müdahalesiz, nesnel ve nicel sonuçlar üreten bir yüz ifadesi tespiti sistemi geliştirilmiştir. Bu sistem ile medikal, pazarlama, güvenlik, sanal gerçeklik, medya gibi birçok alanda uygulama gerçekleştirilebilir.

Bu tez kapsamında çıkarılan ve değerlendirme aşamasında olan çalışmaların listesi ise aşağıda verilmiştir:

- Basılan makaleler:
  1. Oztel I, Yolcu G, Oz C, Kazan S, Bunyak F (2018) iFER: facial expression recognition using automatically selected geometric eye and eyebrow features. *Journal of Electronic Imaging* 27:1.
  2. Oztel I, Yolcu G, Oz C, Kazan S (2018) Facial expression recognition with robust feature selection. *The Online Journal of Science and Technology* 8:83–91.
  3. Yolcu G, Oztel I, Kazan S, Oz C, Bunyak F (2018) Facial component segmentation using convolutional neural network. *The Online Journal of Science and Technology* 8:84–88.

– Değerlendirme aşamasında olan makaleler:

4. Yolcu G, Oztel I, Kazan S, Oz C, Bunyak F (2018) Deep Learning-based Face Analysis System for Monitoring Customer Interest.

5. Yolcu G, Oztel I, Kazan S, Oz C, Palaniappan K, Lever TE, Bunyak F (2018) Facial Expression Recognition for Monitoring Neurological Disorders based on Convolutional Neural Network

– Yayınlanan bildiriler:

6. Oztel I, Yolcu G, Oz C, Kazan S (2017) Facial expression recognition with robust feature selection. International Science and Technology Conference, Cambridge, MA, USA.

7. Yolcu G, Oztel I, Kazan S, Oz C, Bunyak F (2017) Facial component segmentation using convolutional neural network. International Science and Technology Conference, Cambridge, MA, USA.

8. Yolcu G, Oztel I, Kazan S, Oz C, Palaniappan K, Lever TE, Bunyak F (2017) Deep learning-based facial expression recognition for monitoring neurological disorders. In: 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas, MO, USA.

– Sunulan posterler:

9. Yolcu G, Oztel I, Kazan S, Oz C, Palaniappan K, Lever TE, Bunyak F (2018) Facial Expression Recognition using Deep Convolutional Neural Network for Monitoring Neurological Disorders. In: Life sciences week. Columbia, MO, USA

Tez kitapçığında devamında yer alan bölümler şu şekilde organize edilmiştir: Bölüm 2’de yüz ifadesi tespiti çalışmaları için literatür incelemesi klasik makine öğrenmesi yöntemleri ve derin öğrenme yöntemleri olmak üzere iki başlık altında verilmiştir. Bölüm 3’te geliştirilen uygulamalarda kullanılan öznelik seçimi, sınıflandırma ve derin öğrenme yöntemlerine dair bilgiler aktarılmıştır. Bölüm 4’te geliştirilen yüz ifadesi tespiti uygulamalarının başarımları ve benzer çalışmalar ile

karşılaştırılmalarına yer verilmiştir. Bölüm 5’te sonuçlar ve gelecek çalışmalara dair önerilere yer verilmiştir.

## **BÖLÜM 2. LİTERATÜR İNCELEMESİ**

Bilgisayar destekli yüz ifadesi tanıma çalışmaları genel olarak üç adımda gerçekleştirilir. İlk olarak görüntülerde yüz tespiti yapılır. İkinci aşamada tespit edilen yüz görüntülerinden öznitelik çıkarımı gerçekleştirilir. Son olarak, elde edilen öznitelikler sınıflandırılarak tanıma yapılır. Bu adımlar kullanılan yöntemlere bağlı olarak doğrudan görüntülere ve görüntülerden elde edilen ara çıktılara uygulandığı gibi; ön işleme yöntemleri de girdi görüntülerine, ara çıktılara ve/veya son çıktılara uygulanabilir.

Yüz görüntüleri üzerinden duygu analizi yöntemleri genel olarak 2 kategori altında incelenmektedir [22], [23]. Bunlar; geometrik tabanlı ve görünüm tabanlı yaklaşımlardır. Geometrik tabanlı yöntemde, yüze ait şekilsel bilgiler kullanılmaktadır. Bu yaklaşıma Active Shape Model (ASM) örnek olarak verilebilir [24]. Görünüm tabanlı yaklaşımda ise yüzün dokusuna ait özellikler ön plana çıkmaktadır. Bu yaklaşımda Local Binary Pattern (LBP) [25] vb. yöntemler yaygın olarak kullanılmaktadır. Active Appearance Model (AAM) [26] ise hem geometrik tabanlı yaklaşımı hem de görünüm tabanlı yaklaşımı kullanan hibrit bir yöntemdir.

### **2.1. Yüz İfadesi Tespiti İçin Literatürde Önerilen Yöntemler**

#### **2.1.1. Klasik makine öğrenmesi kullanan çalışmalar**

De ve arkadaşları Eigenface yöntemi ile duygu tespiti yapmışlardır. Çalışmalarında [19] HSV (Hue, Saturation, Value) renk uzayında ten rengi tespiti ile yüz ve boyun bölgelerini belirlemişlerdir. İstenmeyen ten rengi bölgelerini yok etmek için segmentasyon yapılarak yüz bölgesi elde edilmiştir. Elde edilen görüntüler üzerinde Temel Bileşen Analizi (Principal Component Analysis - PCA) ile boyut azaltma

yapılmıştır. Ardından test verileri Öklid uzaklıklarına göre sınıflandırılmıştır. Mahersia ve ekibinin çalışmasında [27] ise öznitelik çıkarmada istatistiksel özellik çıkarma yaklaşımı olan Steerable Pyramid Decomposition yöntemi kullanılmış, sınıflandırma ise Neural Networks (NN) ile Bayesian yöntemlerinin hibrit olarak kullanılmasıyla gerçekleştirilmiştir. Ouyang ve ekibince sunulan bir çalışmada [28] ise yüz ifadesi tespiti için Histogram of Gradient (HOG) + Sparse Representation Based Classification (SRC) ve LBP + SRC yöntemleri kullanılmıştır. HOG ve LBP yöntemleri birbirinin tamamlayıcısı olarak kurgulanmıştır. Çünkü HOG yöntemi daha çok şekil bilgisi çıkarmada LBP yöntemi ise gri seviyeli fotoğraflarda doku bilgisi çıkarmada kullanılır. SRC yöntemi yüz ifadesi tespitinde çok kullanılan bir yöntem olmakla beraber, hangi özelliklerin SRC için uygun olabileceğini kestirmek zordur. Ouyang ve ekibinin daha önceki çalışmalarında da SRC için özniteliklerin rastgele seçildiği gözlenmiştir. Bu çalışmalarında ise hangi özelliklerin seçilmesi gerektiği analiz edilmiştir. Ayrıca geliştirdikleri sistem gürültülü fotoğraflarda da iyi sonuç vermektir. Sönmez ve Albayrak çalışmalarında [29] yüzleri; gözler ve ağızlar gibi alt bloklara bölmüştür. Öznitelik çıkarımı için LBP, PCA ve Linear Discriminant Analysis (LDA) yöntemlerini, sınıflandırma içinse SRC ve NN yöntemlerini kullanarak performans karşılaştırmaları yapmışlardır. Çalışmalarının sonucuna göre en iyi performansı blok tabanlı SRC sergilemiştir.

Yüz ifadesi çalışmalarında, SVM yöntemi verdiği başarılı sonuçlar nedeni ile sınıflandırma amaçlı sıklıkla kullanılmaktadır. Liao ve ekibi farklı duyguların farklı yüz bölgelerindeki ağırlıklarını hesaplayarak yüz ifadesi tespiti yapmıştır [30]. Çalışmalarında nötr fotoğraf ile duygu ifadesi olan fotoğraf arasındaki optik akış incelenmiş olup, sınıflandırma için SVM kullanılmış ve ortalama %92,5 başarımla elde edilmiştir. Sümer'in 2014 yılında tamamladığı yüksek lisans tezinde [31] öznitelik çıkarma yöntemi olarak LBP, Local Ternary Pattern (LTP) ve Gabor filtrelerini, sınıflandırma yöntemi olarak da SVM'yi kullanmıştır. Karşılaştırmalı sonuçlara göre en iyi sonuç öfke, mutluluk ve şaşkınlık ifadelerinde elde edilmiştir. Bu ifadeler için başarımlar sırasıyla %97,78, %100, %97,59'dur. Zavaschi ve arkadaşlarının 2011 yılındaki çalışmalarında [32] öznitelik çıkarma yöntemi olarak LBP ve Gabor filtreleri karşılaştırmalı olarak kullanılmış, SVM yöntemiyle yüz ifadesi tespiti için



sınıflandırılma yapılmıştır. Valstar ve arkadaşları 2011 yılındaki çalışmalarında [33] öznitelik çıkarımı için LBP ve PCA yöntemlerini ve sınıflandırma için de SVM sınıflandırıcısını kullanmışlardır. Çalışmalarında; öfke, korku, sevinç, rahatlama ve üzüntü duygularını ortalama %83,5 başarıyla belirlemişlerdir. Bartlett ve arkadaşları ise 2005 yılındaki çalışmalarında [34] Gabor filtrelerini öznitelik çıkarımı için, AdaBoost ve SVM yöntemlerini de sınıflandırma için kullanmışlardır.

### 2.1.2. Derin öğrenme kullanan yöntemler

Son yıllarda derin öğrenme algoritmaları birçok alanda olduğu gibi yüz ifadesi tespiti alanında da başarılı sonuçlar ortaya koymuştur. Lopes ve arkadaşları derin öğrenme yöntemlerinden evrimsel sinir ağlarını CNN kullanarak altı farklı yüz ifadesini tanıyabilen bir sistem geliştirmiştir [35]. Geliştirdikleri uygulamada, görüntüleri CNN'e göndermeden önce; görüntülere yoğunluk normalizasyonu, yüz kırpma, döndürme gibi ön işlemler uygulamışlardır. Pitaloka ve arkadaşları da yine altı farklı yüz ifadesi tanıma üzerine çalışmış olup; görüntülere yüz tespiti, kırpma işlemi, histogram denkleştirme ve veri normalizasyonu gibi ön işlemler uygulamışlardır [36]. Son adımda işe ön işlenen görüntüler CNN ile sınıflandırılmıştır. Matsugu ve arkadaşları ise yüz ifadesi çalışmasında CNN kullanarak yüz tespiti ve gülümseme tespiti işlemlerini gerçekleştirmişlerdir [37]. Ayrıca CNN kural tabanlı bir sistem ile desteklenmiştir. Li ve arkadaşları 2017'de sundukları bir çalışmada [38] özelleştirilmiş bir CNN ile yüz ifadesi tespiti yapmışlardır. Eğitilebilir konvolüsyon çekirdeği ile öznitelikler çıkarıp boyut indirgemesi yapılmış ve Faster Regions with Convolutional Neural Network Features (Faster R-CNN) ile sınıflandırma işlemi gerçekleştirilmiştir. Fathallah ve arkadaşları 2017'deki çalışmalarında [39] altı duyguyu sınıflandırmak için bir CNN yapısı önermişlerdir. Önerilen sistem, içlerinde Radboud Face Database'inde (RaFD) bulunduğu farklı veri tabanları ile test edilmiştir. RaFD ile %94,16 başarımla elde edilmiştir. Bilgiç ve arkadaşlarının 2017 deki çalışmalarında [40] dört farklı yüz ifadesini sınıflandırmak için CNN ile derin öznitelikler çıkarmış, PCA ile bu öznitelikler üzerinde boyut indirgenmesi gerçekleştirilmiş ve Multi Layer Perceptron (MLP) ile sınıflandırılmıştır. Sundukları çalışmada %100 başarımla ulaştıklarını bildirmişlerdir.

## 2.2. Uygulamalarda Kullanılan Veritabanları

### 2.2.1. Radboud Face Database (RaFD)

Çalışmalar kapsamında sıklıkla RaFD [41] yüz veri tabanı kullanılmıştır. RaFD ücretsiz bir yüz veri tabanıdır. Veri tabanında Kafkas etnik kökenli insanlara ait yüz görüntüleri bulunmaktadır. Bu yüz görüntüleri farklı bakış açılara (0, 45, 90, 135, 180 derece) sahip 5 kamera kullanılarak elde edilmiştir. Bu veri tabanının içeriği sayesinde araştırmacılar duygu analizi, bakış yönü tespiti ve/veya baş oryantasyonu konularında çalışmalar yapmaktadır. Veri tabanında 8 farklı yüz ifadesi (nötr, sinirli, üzgün, korku, iğrenme, şaşkın, mutlu ve aşağılama) ve 3 farklı bakış yönü (tam karşıya, sağa ve sola) bulunmaktadır. Tüm fotoğraflar Facial Action Coding System (FACS) [42] kurallarını temel alarak hazırlanmıştır. Veri tabanı 39 yetişkin ve 10 çocuk olmak üzere 49 insana ait görüntü içermektedir; çocukların 6'sı, yetişkinlerin ise 19'u kadındır.

### 2.2.2. Japanese Female Facial Expression Database (JAFFE)

Bir diğer ücretsiz yüz ifadesi veri tabanı olan JAFFE'de [43] 10 Japon kadın modellik etmiştir. Bu görüntülerde her bir kişiye ait 7 farklı duygu bulunmaktadır, bu duygular: nötr, sinirli, üzgün, korku, iğrenme, şaşkın, mutlu olarak belirlenmiştir. Görüntüler kaydedilirken saçların yüzden uzak olmasına ve ortam ışığının aynı olmasına dikkat edilmiştir. Veritabanı yalnızca ilgili modellerin ön cepheden çekilmiş yüz görüntülerini içerir.

### 2.2.3. Karolinska Directed Emotional Faces (KDEF)

KDEF [44] yüz veri tabanı 70 insanın katılımıyla oluşturulmuştur. Veri tabanında görüntüleri bulunan 35 kadın ve 35 erkeğin yaş aralığı ise 20 ile 30 arasında değişmektedir. Katılımcıların görüntüleri 5 farklı kamera ile farklı 5 açıdan çekilmiştir. Veritabanında her bir kişiye ait nötr, sinirli, üzgün, korku, iğrenme, şaşkın

ve mutlu olmak üzere 7 farklı duygu bulunmaktadır. Görüntülerde bıyık, sakal, görülebilir makyaj ve gözlük olmamasına dikkat edilmiştir.

#### **2.2.4. Cohn–Kanade Database (CK)**

Çalışmada kullanılan diğer bir yüz veri tabanı olan CK [45], FACS tanımlamalarına göre oluşturulmuştur. 210 kişinin katılımıyla oluşan CK veri tabanında katılımcıların yaşları 18-50 aralığında değişmektedir. Katılımcıların %69'u kadın, %31'i erkek olup, veri tabanı hem beyaz hem de siyah tenli kişilerin yüz görüntülerini barındırmaktadır. Katılımcılar bir odada bir sandalyeye oturup iki farklı kamera ile görüntülenmiştir. Kameralardan biri ön yüz görüntülerini alırken, diğeri ise 30 derecelik açı ile kişiyi sol cepheden kaydetmiştir. Ayrıca her kayıt için sabit ışıklandırma koşulları oluşturulmuştur.

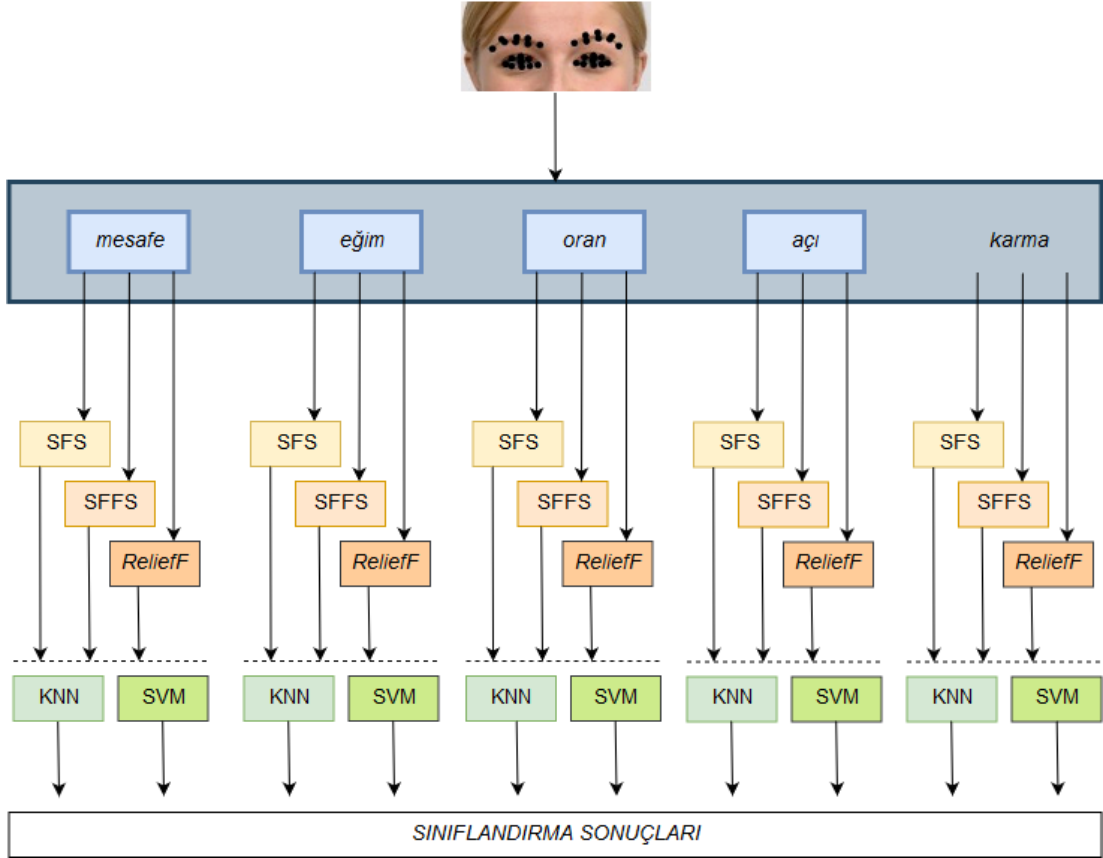
## **BÖLÜM 3. YÜZ İFADESİ TESPİTİ YÖNTEMLERİ**

### **3.1. Çalışmalar Kapsamında Kullanılan Klasik Makine Öğrenmesi Yöntemleri**

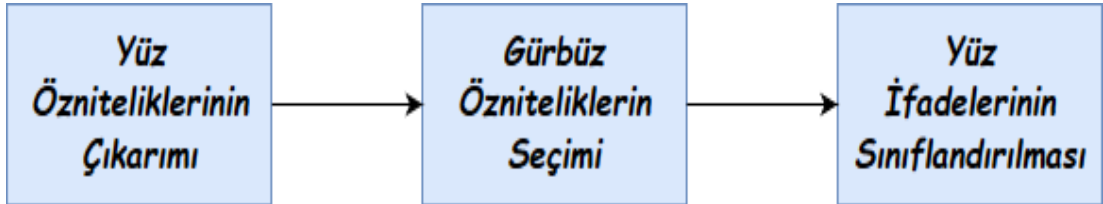
Bu tez çalışması kapsamında geliştirilen ilk uygulamada öznitelik çıkarımı aşamasında; yüz anahtar noktaları geometrik tabanlı işlemlere tabi tutulmuş ve mesafe, oran, açı, eğim ile karma olmak üzere 5 kategoride öznitelik kümeleri elde edilmiştir. Elde edilen her bir öznitelik kategorisi çok fazla sayıda veri içerdiğinden, iş yükünü azaltmak ve başarımlarını arttırmak amacıyla bu öznitelikler SFS, Sequential Floating Forward Selection (SFFS) ve ReliefF öznitelik seçim algoritmaları ile analiz edilmiştir. Böylece yüz ifadesi tespitine etki eden en gürbüz öznitelikler seçilmiştir, başarımları olumsuz etkileyen öznitelikler sınıflandırma işlemine dahil edilmemiştir. Seçilen gürbüz öznitelikler; SVM ve K-en Yakın Komşu (k-Nearest Neighbour- KNN) algoritmaları ile sınıflandırılmıştır. Bu işlemleri anlatan akış şeması Şekil 3.1.'de gösterilmiştir.

#### **3.1.1. Yüz ifadesi tespiti için gereken adımlar**

Yüz ifadesi tespiti, yüz görüntüleri üzerinde genel olarak üç ana adımda gerçekleştirilir. Bunlar: öznitelik çıkarımı, öznitelik seçimi ve sınıflandırma aşamalarıdır (Şekil 3.2.). Öznitelik çıkarma aşamasında sınıflandırma için kullanılmak üzere görüntülerden veriler elde edilir. Bu veriler üzerinde sonuca olumlu anlamda etki edenleri ön plana çıkarmak için öznitelik seçim aşaması gerçekleşir. Bu aşamadan sonra elde edilen gürbüz öznitelikler sınıflandırma yöntemleri ile sınıflandırılır. Görüntülerin yalnızca yüz içermediği durumlarda ise ayrıca yüz bulma algoritmalarından faydalanılır. Böylece yüz görüntüleri, girdi görüntülerinden seçilerek kırılır.



Şekil 3.1. Klasik makine öğrenmesi yöntemleriyle kısmi yüz görüntüleri üzerinde yapılan yüz ifadesi tespiti uygulamasının akış diyagramı.



Şekil 3.2. Yüz ifadesi tespiti blok diyagramı.

### 3.1.2. Görüntülerde yüz tespiti

Yüz algılama literatürde sıklıkla çalışılan bir konu olmuştur. Yüz tespiti algoritmaları dört ana kategoride incelenebilir [46]. Bunlar bilgi tabanlı yöntemler, sabit özellik yaklaşımları, şablon eşleme yöntemleri ve görünüm tabanlı yöntemlerdir. Bilgi tabanlı yöntemler, yüz tespiti için, önceden belirlenmiş insan bilgisine dayalı olan yöntemleri kullanır. Sabit özellik yaklaşımları, farklı pozlama ve aydınlatma varyasyonları için gürbüz yüz özellikleri bulmayı amaçlar. Şablon eşleme yöntemleri, bir görüntünün yüz

olup olmadığını anlamak için önceden depolanmış şablonları kullanır. Görünüm tabanlı yöntemler ise bir dizi temsilci eğitim verileri ile yüz modelleri öğrenirler. Görünüm tabanlı yöntemler, üstün hesaplama gücü sayesinde diğerlerine göre daha iyi performans gösterir [46].

Viola&Jones algoritması ile yüz tespitinde son yıllarda önemli aşamalar kaydedilmiştir. Viola&Jones algoritması dijital kameralarda ve fotoğraf yazılımları gibi gerçek dünya uygulamalarında yüz bulmayı pratik olarak mümkün hale getirmiştir.

Bu çalışmada, öznitelik çıkarımı öncesinde, yüz görüntülerinden arka planı silerek sadece yüz bölgelerine odaklanma sağlanmıştır. Bu amaçla yüzler tespit edilip mümkün oldukça ortamdaki bağımsız hale getirilmiştir. Yüz tespiti için literatürde oldukça sık kullanılıp başarılı sonuçlar üreten Viola&Jones [47] yönteminden yararlanılmıştır.

Viola&Jones yöntemi üç temel fikire dayalı olarak ortaya çıkmıştır. Bunlar;

- İntegral görüntü,
- Adaboost algoritması yardımıyla seçilen az sayıdaki öznitelikten yeni bir sınıflandırıcı oluşturma,
- Ardışık bir yapıda sınıflandırıcıları kombinlemedir.

İntegral görüntü, çok hızlı olarak öznitelik çıkarmaya yarayan yeni bir görüntü temsil etme biçimidir. İntegral görüntüde; bir piksel, kendisinin solundaki ve üstündeki piksellerin toplamı olarak ifade edilir. Denklem 3.1’de bu işlem gösterilmiştir.

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

Denklemden  $ii(x, y)$  ifadesi integral görüntüyü,  $i(x', y')$  ifadesi ise ham görüntüyü temsil eder.

Viola&Jones yönteminde öznitelikler Adaboost algoritması yardımıyla seçilir. Her bir zayıf sınıflandırıcı tek bir özellik ile sınırlandırılmıştır. Zayıf sınıflandırıcılar pozitif ve negatif örnekleri ideal olarak ayıran tek dikdörtgen özelliğini seçer. Zayıf sınıflandırıcı tanımı Denklem 3.2’de gösterilmiştir.

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

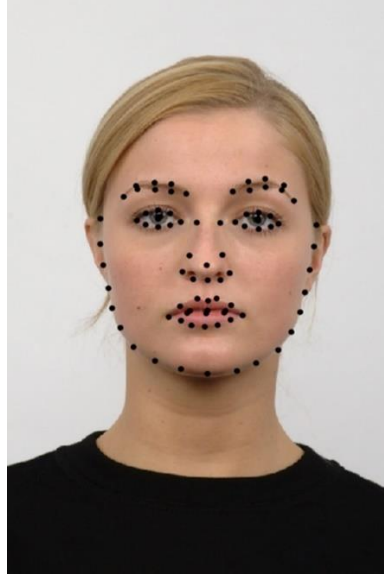
Denklemde  $h_j(x)$  bir zayıf sınıflandırıcıyı,  $f_j$  bir özelliği,  $\theta_j$  eşik değerini ve  $p_j$  verilen eşitliğin yönünü sembolize eder.

Son olarak sınıflandırıcılar ardışık olarak birbirine bağlanır. Ardışık yapı, önemli yüz bölgelerine odaklanma sağlayarak hız açısından performansı artırır.

### 3.1.3. Öznitelik çıkarımı

Öznitelik çıkarımı, problemin çözümüne yönelik olarak, orijinal veriyi çok iyi temsil edebilen özelliklerin elde edilmesi işlemidir. Yüz ifadesi tespiti için literatürde sıklıkla kullanılan ve başarılı sonuçlar üreten öznitelik çıkarım algoritmalarına örnek olarak PCA [48], LBP [49], AAM [50], Active Shape Model (ASM) [51], HOG [52] vb. verilebilir.

Yüz ifadesi tespiti için uygulanan geometrik yaklaşımda, ilk olarak ifade tespiti için önemli olan yüz anahtar noktalarının bulunması gerekmektedir. Yüz anahtar noktalarını tespit eden çeşitli kütüphaneler ve araçlar bulunmaktadır. Bu çalışmada Face++ SDK’sı (Software Development Kit) [53] kullanılarak insan yüzüne ait anahtar noktalar tespit edilmiştir. Face++ yüz görüntülerinde, yüz hatlarını temsil edecek şekilde 83 adet anahtar nokta bulabilmektedir. Face++, arka planında güçlü bir CNN yapısı kullanmaktadır. Yüz anahtar noktaları Face++ tarafından işaretlenmiş örnek bir görüntü Şekil 3.3.’de verilmiştir. Bu şekil üzerinden de görüleceği gibi, kaşlar, gözler, burun, ağız ve yüz çevresi başarılı bir şekilde tespit edilebilmektedir.



Şekil 3.3. Face++ tarafından bulunan 83 adet yüz anahtar noktası.

Geliştirilen ilk uygulamada, geometrik öznitelikler ile yüz ifadesi tespiti aşamasında sadece göz bölgesinden ifade tespiti yapılmıştır. Bu nedenle göz ve kaşlar için 34 adet yüz anahtar noktası üzerinden öznitelik çıkarımı yapılmıştır. Elde edilen anahtar noktaların her birinin diğerlerine olan mesafesi, her mesafenin oluşturduğu doğrunun diğer doğrulara olan oranı, her iki noktanın oluşturduğu doğruların eğimi ve bu doğruların her birinin diğer doğrularla arasında kalan açılar hesaplanmıştır. Böylece mesafe, oran, eğim ve açı olmak üzere 4 kategoride öznitelik kümeleri elde edilmiştir. En son tüm kümeler birleştirilerek bir öznitelik kümesi daha elde edilmiştir, bu öznitelik kategorisi ise karma olarak isimlendirilmiştir. Kümelerdeki eleman sayıları Denklem 3.3'e göre hesaplanır.

$$C(n, r) = \frac{n!}{r!(n-r)!} \quad (3.3)$$

Denklem 3.3'e göre 34 adet yüz anahtar noktasının her birinin diğerine olan mesafelerinin sayısı;  $n=34$   $r=2$  olduğundan 561 adet olarak bulunur.  $p_i(x_i, y_i)$  ve  $p_j(x_j, y_j)$  noktalarının arasındaki mesafe Denklem 3.4 yardımı ile bulunur.

$$|L_{ij}| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (3.4)$$



Denklemden  $|L_{ij}|$  ifadesi mesafeyi belirtmektedir.

Sonraki adımda bu mesafelerin oluşturduğu doğruların birbirine olan oranları hesaplanmıştır. Denklem 3.3'e göre  $n=561$   $r=2$  olduğundan toplam 157.080 adet oran verisi elde edilmiştir. Bir sonraki aşamada ise bulunan anahtar noktalardan geçen doğrular arası eğimler hesaplanmıştır. Eğim hesaplamak için Denklem 3.5 kullanılmıştır.

$$m_{ij} = \frac{y_{pj} - y_{pi}}{x_{pj} - x_{pi}} \quad (3.5)$$

Denklem 3.5'te  $(x_{pi}, y_{pi})$  ve  $(x_{pj}, y_{pj})$  ifadeleri herhangi iki noktanın koordinatlarını,  $m_{ij}$  ifadesi ise bu iki noktanın birleşmesiyle oluşan doğrunun eğimini sembolize etmektedir.

Denklem 3.3'e göre 34 adet yüz anahtar noktasının her ikisini birleştiren doğruların eğim sayısı hesaplanırken;  $n=34$   $r=2$  olduğundan 561 adet eğim verisi elde edilmiştir. Son aşamada ise anahtar noktalardan geçen doğrular arasında kalan açının tanjant değerleri hesaplanmıştır. Açının tanjant değerlerini hesaplamak için Denklem 3.6'dan yararlanılmıştır.

$$\tan(\alpha) = \frac{m_{ij} - m_{kp}}{1 + m_{ij} \cdot m_{kp}} \quad (3.6)$$

Denklem 3.6'da  $m_{ij}$  ve  $m_{kp}$  ifadeleri herhangi iki doğrunun eğimini sembolize etmektedir. Denklem 3.3'e göre  $n=561$   $r=2$  olduğundan toplam 157.080 adet tanjant değeri elde edilmiştir. Ayrıca elde edilen tüm özneliklerin toplanması ile karma olarak isimlendirilen bir öznelik kategorisi daha elde edilmiştir. Bu kategorideki öznelik sayısı ise 315.282 adettir. Öznelik kategorilerinin formülleri ve sayıları Tablo 3.1.'de gösterilmiştir.

Tablo 3.1. Hesaplanan öznitelik kümelerinin eleman sayıları ve formülleri.

Öznitelik kategorisi	Gösterim formülleri	Öznitelik Sayısı
Mesafe	$ L_{ij} $	561
Eğim	$m_{ij} = \frac{y_{p_j} - y_{p_i}}{x_{p_j} - x_{p_i}}$	561
Oran	$ L_{ij}  /  L_{kp} $	157.080
Açı	$\tan(\alpha) = \frac{m_{ij} - m_{kp}}{1 + m_{ij} \cdot m_{kp}}$	157.080
Karma	$\{ L_{ij} \} \cup \{m_{ij}\} \cup \{ L_{ij}  /  L_{kp} \}$ $\cup \{\tan(\theta)\}$	315.282

### 3.1.4. Öznitelik seçimi

Öznitelik seçimi, çok fazla verinin olduğu durumlarda, işlem yükünü azaltmak, zaman tasarrufu sağlamak ve sonuca olumsuz etki eden öznitelikleri eleyip başarımı arttırmak için başvurulan bir yöntemdir. Literatürde çeşitli öznitelik seçim algoritmaları bulunmaktadır. Bu algoritmalara örnek olarak SFS, SFFS ve ReliefF verilebilir. Bu algoritmalar genel olarak, birçok veri içinden problemin çözümüne en çok katkı sağlayan öznitelikleri bulmayı hedefler. Böylece problemin çözümüne az katkı sağlayan ya da hiç katkı sağlamayan gereksiz öznitelikler gözardı edilerek başarımları yüksek, daha güçlü çözümler sağlanmış olur.

#### 3.1.4.1. Sequential Forward Selection - SFS

SFS algoritması Whitney [54] tarafından geliştirilen iteratif bir öznitelik seçim algoritmasıdır. SFS algoritması boş bir öznitelik kümesi ile çalışmaya başlar. İçerisinde bulunan sınıflandırıcı ile sonuca en iyi derecede etki eden özniteliği bulur. Bir sonraki iterasyonda ise ilk seçilen öznitelik ile birlikte sonuca en iyi derecede etki eden ikinci özniteliği tespit eder. Bu iterasyonlar, önceden belirlenen başarımları oranına

ulaşılıncaya kadar ya da yeni seçilen öznitelikler başarımlarında hiçbir iyileşme sağlamayınca kadar devam eder.

SFS algoritması, içerisinde bir sınıflandırıcıyla birlikte çalıştığından çalışma zamanını olumsuz etkilenir [54]. Ancak öznitelikler arasındaki ilişkileri dikkate alarak seçim yaptığından başarımlar oranı yüksektir. Bir öznitelik tek başına problemin çözümüne yüksek oranda katkı sağlamasına rağmen, diğer öznitelikler ile birlikte çalıştığında, başarımlar oranı düşebilir. SFS algoritması bu noktada başarımlar oranı artacak şekilde seçimler yapar.

SFS algoritmasının sözde kodu aşağıda verilmiştir [55]:

1. Boş bir öznitelik kümesi oluştur:  $F_0 = \{ \}$
2. En iyi sonucu üreten özniteliği seç:  $Z^l = \arg \max [J(F_k + Z)] \& Z^l \notin F_k$
3. Seçilen özniteliği gümbüz öznitelik kümesine ekle  $F_{k+1} = F_k + Z^l \& k++$
4. Adım II'ye git.

SFS algoritması çalışma prensibi bir örnekle aşağıdaki gibi açıklanabilir [56].

$J(x) = -2x_1x_2 + 3x_1 + 5x_2 - 2x_1x_2x_3 + 7x_3 + 4x_4 - 2x_1x_2x_3x_4$  amaç fonksiyonu için  $x_1$ ,  $x_2$ ,  $x_3$  ve  $x_4$  değişkenlerinin öznitelikleri temsil ettiği varsayalım.  $x_1$  özniteliğinin amaç fonksiyonuna etkisi araştırılırken,  $x_1 = 1$ , diğer özniteliklerin ise etkisiz yani  $x_2 = x_3 = x_4 = 0$  olduğu düşünülmelidir. Bütün özniteliklerin amaç fonksiyonuna etkisi araştırıldığında aşağıdaki sonuçlar elde edilir:

$$J(x_1) = 3 \quad J(x_2) = 5 \quad J(x_3) = 7 \quad J(x_4) = 4$$

Yukarıda görüldüğü gibi amaç fonksiyonunu en yüksek değere ulaştıran öznitelik  $x_3$ 'tür. Bu nedenle ilk iterasyonda  $x_3$  özniteliği seçilerek, ikinci iterasyonda  $x_3$  ile birlikte amaç fonksiyonunu en yüksek değere ulaştıran diğer öznitelikler aranır.

$$J(x_3x_1) = 10 \quad J(x_3x_2) = 12 \quad J(x_3x_4) = 11$$

Yukarıdaki sonuçlara göre  $x_2$  özneliği de seçilir. Sıradaki iterasyonda  $x_2$  ve  $x_3$  ile birlikte sonuca etki eden öznelikler incelenir.

$$J(x_3x_2x_1) = 11 \quad J(x_3x_2x_4) = 16$$

Buna göre  $x_4$  özneliği seçilir. Sonuca tüm özneliklerin birlikte etkisi ise;

$$J(x_3x_2x_4x_1) = 13 \text{ olarak hesaplanır.}$$

### 3.1.4.2. Sequential Floating Forward Selection - SFFS

SFFS algoritması Pudil [57] tarafından 1994 yılında geliştirilmiştir. SFFS algoritması da SFS algoritmasına benzer şekilde boş bir öznelik kümesi ile öznelik seçim işlemine başlar. Her bir adımda, sonuca en çok etki eden öznelik bulunur. SFFS algoritmasının SFS algoritmasından farkı, seçilen öznelik kümesinden bir öznelik çıkarıldığında sonucun nasıl değişeceğinin analizinin de yapılıyor olmasıdır. Böylece varsa en zayıf öznelik de kümeden çıkarılır. Bu şekilde istenilen değere ulaşıncaya kadar öznelik kümesinde dinamik olarak öznelik ekleme ve çıkarma işlemleri yürütülür.

SFFS algoritmasının sözde kodu aşağıda verilmiştir [56]:

1. Boş bir öznelik kümesi oluştur:  $F_0 = \{ \}$
2. En iyi sonucu üreten özneliği seç:  $Z^l = \arg \max [J(F_k + Z)] \& Z^l \notin F_k$
3. Seçilen özneliği öznelik kümesine ekle  $F_{k+1} = F_k + Z^l \& k++$
4. En düşük sonucu üreten özneliği seç:  $x = \arg \max [J(F_k - x)] \& x \in F_k$
5. Eğer  $J(F_k - x) > J(F_k)$  seçilen özneliği kümeden çıkar.  $F_{k+1} = F_k - x \& k++$   
adım III'e git

Değilse adım II'ye git.

### 3.1.4.3. ReliefF

ReliefF algoritması Kira ve Rendel [58] tarafından önerilmiştir. ReliefF algoritmasında her bir öznitelik  $[-1 \ 1]$  aralığında bir ağırlık olarak değerlendirilir. Her bir özneliğin ağırlığı iteratif olarak güncellenir. Verilerden bir örnek alınır ve bu örnekle aynı sınıfta olan en yakın bir örnek ve farklı sınıfta olan en yakın bir örnek seçilir. Ardından her bir özneliğe ait ağırlıklar, bu veriler kullanılarak güncellenir.

ReliefF algoritması üç temel maddede aşağıdaki gibi özetlenebilir [59]:

1. Aynı sınıfa ait ve farklı sınıfa ait örneklerden ilgili öznitelik değerini bul.
2. İlgili özelliğin ağırlığının hesapla.
3. Özellikleri ağırlıklarına göre sırala ve belirli bir eşik değeri veren özellikleri seç.

İlgili özelliğin ağırlığı Denklem 3.7 ile hesaplanır [59].

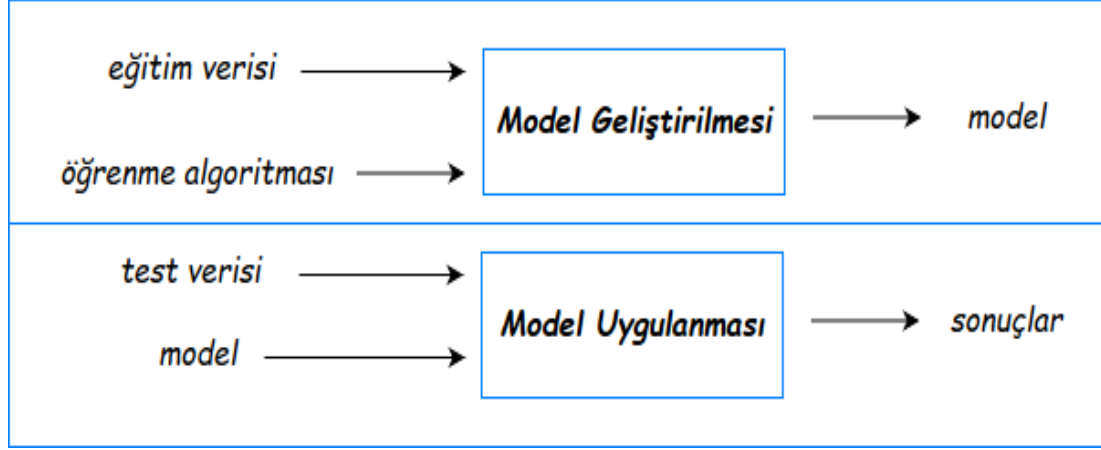
$$w_i = w_{i-1} - (x_i - nearHit_i)^2 + (x_i - nearMiss_i)^2 \quad (3.7)$$

Denklemden  $w_i$   $i$ . özneliğin ağırlığını,  $nearHit_i$  aynı sınıfa sahip en yakın örnekteki ilgili öznitelik değerini ve  $nearMiss_i$  farklı sınıfa ait en yakın örnekteki ilgili öznitelik değerini gösterir.

### 3.1.5. Sınıflandırma

Sınıflandırma, veri kümesindeki verilerin hangi sınıfa ait olduğunun belirlenmesi işlemidir, eğitici bir öğrenme çeşididir. Önceden hangi sınıfa ait olduğu bilinen verilerden, eğitim ve test kümeleri oluşturulur. Daha sonra eğitim kümesindeki verilerle sınıflandırma algoritması tarafından bir model oluşturulur. Ardından test kümesindeki veriler oluşturulan bu modele giriş olarak verilir ve bu giriş verilerine ait

sınıf değerleri model tarafından üretilir. Elde edilen cevaplar gerçek sınıfları ile karşılaştırılır ve modelin başarısı ölçülür. Bu işlem Şekil 3.4.'de gösterilmiştir.



Şekil 3.4. Sınıflandırma algoritmalarının genel blok diyagramı.

Literatürde kullanılan, başarılı ve verimli sonuçlar veren sınıflandırma algoritmalarına örnek olarak SVM [60] , Naive Bayes Classifier [61], KNN [62], karar ağaçları (decision tree) [63], yapay sinir ağları (Artificial Neural Network - ANN) [64] vb. verilebilir. Tezin bu bölümündeki sınıflandırma için Destek Vektör Makineleri ve k-en Yakın Komşu algoritması kullanılmıştır.

### 3.1.5.1. K-en yakın komşu algoritması (K-nearest neighbors) - KNN

Cover ve arkadaşları tarafından geliştirilen k-en yakın komşu algoritması [65] istatistiksel ve denetimli bir algoritmadır. Bu algortmada verinin k adet en yakın komşusunun sınıfsal konumları göz önüne alınır. Bu komşular, hangi sınıf etiketi altında daha çok toplanmış ise veri o sınıfa dahil edilir. Veriler arası uzaklığın hesaplanmasında literatürde çeşitli yöntemler bulunmaktadır. Chebychev, Manhattan, Öklid gibi çeşitli uzaklık hesaplama yöntemleri veriler arası mesafeyi hesaplamak için kullanılabilir [66]. Koordinat düzleminde yer alan  $p_1(x_1, y_1)$  ve  $p_2(x_2, y_2)$  noktaları arasındaki mesafeyi bu yöntemlerle hesaplamak için Tablo 3.2.'de verilen formüller kullanılır.

Tablo 3.2. Uzaklık hesaplama yöntemleri ve formülleri.

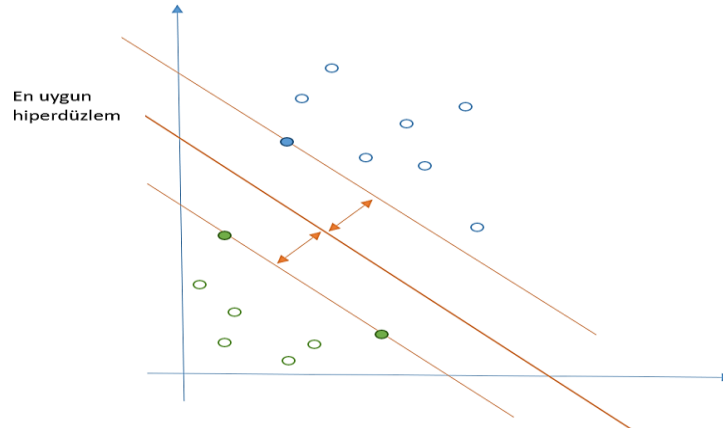
Uzaklık Hesaplama Yöntemi	Formül
Chebyshev	$d = \max( x_1 - x_2 ,  y_1 - y_2 )$
Manhattan	$d =  x_1 - x_2  +  y_1 - y_2 $
Öklid	$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$

K-en yakın komşu algoritmasının sözde kodu şu şekildedir [65]:

- I. *k* değeri belirlenir.
- II. Diğer nesnelere hedef nesneye olan uzaklıklar hesaplanır.
- III. Uzaklıklar sıralanır, en yakın komşular bulunur.
- IV. En yakın komşu kategorileri toplanır.
- V. *k* ye bağlı olarak hedef nesnenin sınıfı bulunur.

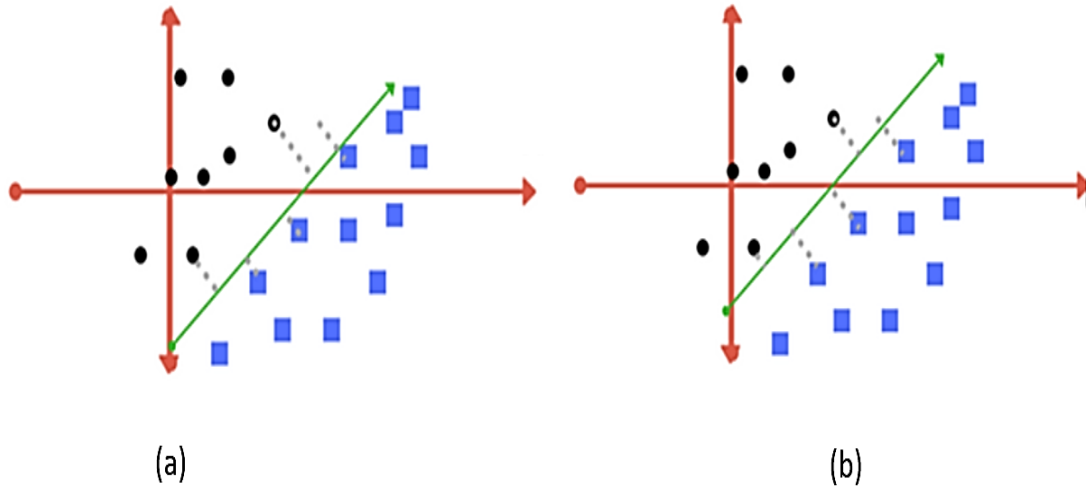
### 3.1.5.2. Destek vektör makineleri (Support vector machines) - SVM

SVM Vapnik ve ekibi [67] tarafından önerilen, istatistiksel öğrenme teorisine dayanan ve farklı sınıfları ayırmak için, sınıflar arası en büyük mesafeyi hesaplayan bir sınıflandırma algoritmasıdır [68]. SVM eğitim setinde, farklı sınıf verilerinin uzaklığını maksimize edecek şekilde ve aynı sınıftaki verilerin ise mümkün olan en büyük bölümünü kapsayacak şekilde bir hiperdüzlem çizer. Bu işlem Şekil 3.5.'te gösterilmiştir [69].



Şekil 3.5. SVM tarafından oluşturulan hiperdüzlem [69].

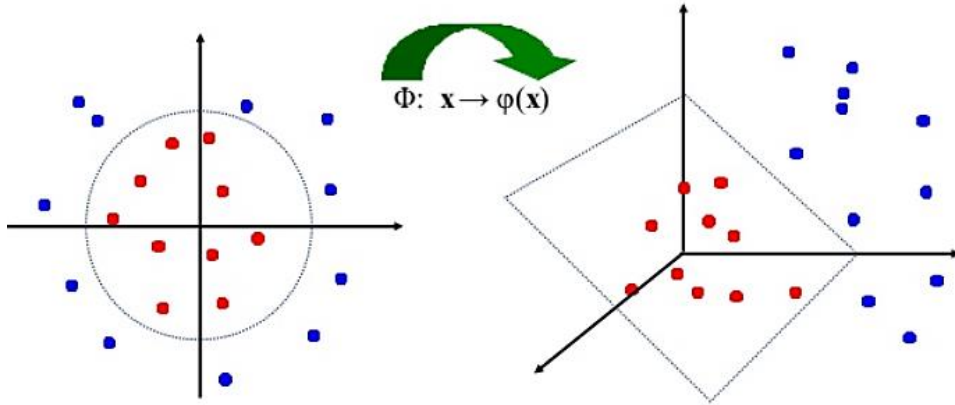
Şekil 3.6.'da SVM'nin sınıflar arasında iyi ve kötü ayırım yaptığı iki örnek grafik üzerinde gösterilmiştir. Şekil 3.6.(b)'de hiperdüzlem iki sınıfa da mümkün olduğunca eşit uzaklıktadır. Fakat Şekil 3.6.(a)'da mavi renkli sınıfa daha yakındır, bu nedenle kötü ayırım yapılmıştır.



Şekil 3.6. İyi ve kötü ayırım yapan iki hiperdüzlem örneği, (a) kötü ayırım, (b) iyi ayrımı gösterir [70].

SVM yüksek doğrulukta çözümler üretebilir ve karmaşık karar sınırlarını modelleyebilir. Çok sayıda bağımsız değişkenle çalışabilme ve doğrusal olarak ayrılan ve ayrılamayan verilere uygulanabilme özelliğine sahiptir. Yüz tanıma [71]–[73], yüz ifadesi tanıma [74]–[76], el yazısı tanıma [77]–[79], biyomedikal veri analizi [80]–[82] gibi konularda sıklıkla kullanılmakta ve başarılı sonuçlar üretmektedir. Doğrusal SVM'ler Hard margin ve Soft margin olmak üzere iki yaklaşımda incelenir. Bunlardan Hard margin yaklaşımı, verilerin tamamen doğrusal olarak ayrılabilirdiği durumlarda kullanılır. Soft margin yaklaşımı ise az sayıda noktadan (gözlemden) dolayı verilerin doğrusal olarak ayrılamadığı durumlarda başvurulan yaklaşımdır. Veriler doğrusal olarak ayrılamadığında ise doğrusal olmayan haritalama ( $\Phi$ ) ile veri, orijinal girdi uzayından daha yüksek boyuttaki bir uzaya aktarılır. Ardından yeni boyutta veriyi en iyi ayıracak düzlem araştırılır. Bu işlem Şekil 3.7.'de gösterilmiştir.





Şekil 3.7. Doğrusal olarak ayrılamayan verinin girdi uzayından daha yüksek boyuttaki uzaya aktarılması [83].

Doğrusal olmayan SVM’de çekirdek düzenlemesi yapılarak  $\Phi(x)$  vektörü yerine girdi uzayındaki verilerden oluşan çekirdek fonksiyonlar kullanılır. SVM’nin kullandığı çeşitli çekirdek fonksiyonlar vardır [84]. Bu fonksiyonlar aşağıda verilmiştir.

- Doğrusal
- Polinomiyal
- Sigmoid
- Radyal Temelli Fonksiyon (RBF)

Bu çalışmada RBF çekirdek fonksiyonu kullanılmıştır. RBF çekirdek fonksiyonu SVM kullanılan sınıflandırma çalışmalarında ağırlıklı olarak tercih edilmektedir. Çeşitli veri kümeleri üzerinde performans incelemeleri yapıldığında en başarılı sonuçlara RBF çekirdeği ile ulaşılmakla birlikte [85] her veri tipi için bu durum geçerli olmayabilir.

SVM iki sınıflı problemlerin çözümünde kullanılmak üzere geliştirilmiştir. Ancak çok sınıflı problemlerde de kullanılmak üzere iki yaklaşım mevcuttur. Bunlar bire-bir (one-vs-one) ve bire-çok (one-vs-all) yaklaşımlarıdır.

Bire-bir yaklaşımında, problem ikili gruplara indirgenerek çözüm bulunur. Herbir ikili grup için ayrı ayrı SVM eğitimi yapılır. Örneğin bir veri A, B, C sınıflarından birine

dahil edilmek istendiğinde; ilk olarak A ve B arasında hiperdüzlem çizilir, buna göre veri örneğin A sınıfında olsun; ardından B ve C arasında hiperdüzlem çizilir, buna göre de verinin B sınıfına dahil olduğu sonucunun çıktığı varsayalım. Son olarak A ve C arasında hiperdüzlem çizilir, buna göre de veri A sınıfında olsun. Tüm bu işlemler sonucunda 3 sınıflandırmada en çok A sınıfı sonucu ortaya çıktığı için veri A kümesine dahil edilir. Bu sınıflandırma yaklaşımında N adet sınıf varsa;  $\frac{N * (N - 1)}{2}$  adet eğitim yapılıır.

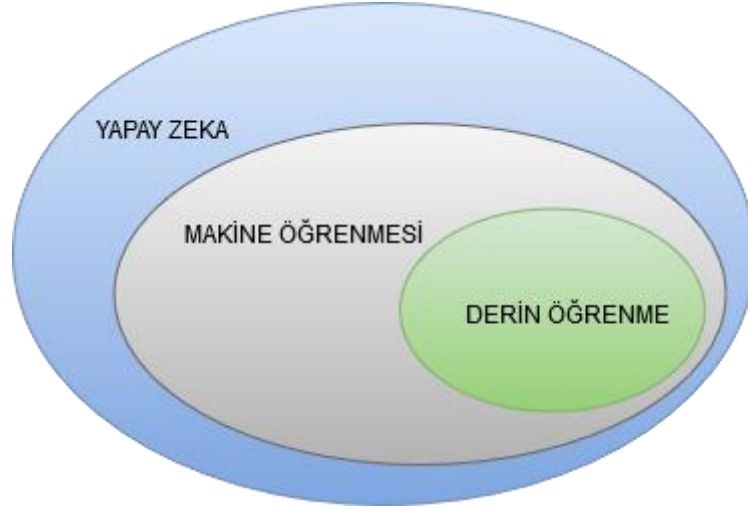
Bire-çok yaklaşımında, verinin bir sınıfa dahil olduğu varsayılarak, ilgili sınıf pozitif, diğer sınıflar negatif kabul edilir ve eğitim yapılıır. İteratif olarak her bir sınıf pozitif kabul edilerek diğer sınıflarla karşılaştırılarak bir sonuca varılır. Buna göre N kadar sınıf için N adet eğitim yapılarak sonuca ulaşılır. Bu nedenle hesaplama açısından bire-bir yaklaşımına göre daha az işlem yükü gerektiren bir yöntemdir.

### 3.2. Çalışmalar Kapsamında Kullanılan Derin Öğrenme Yöntemi

Derin öğrenme, bünyesinde yapay sinir ağı (YSA / ANN) yapısal mimarisini kullanan bir makine öğrenmesi yöntemidir (Şekil 3.8.). YSA yapısında 3 katman içerir: giriş katmanı, gizli katman ve çıkış katmanı. YSA'da bulunan gizli katmanın birden fazla katman içermesi durumlarda ortaya çıkan ağ derin ağ, öğrenme ise derin öğrenme olarak adlandırılır [86]. Derin öğrenme son yıllarda popüleritesini arttırmış olsa da aslında tarihi on yıllar öncesine dayanmaktadır. Yakın zamanda gelişen teknoloji ile grafik işlemcilerinin işlem gücünün artması, yazılımsal gelişmeler ve veri setlerinin hacim olarak büyümesi yeniden derin öğrenme yöntemlerine olan ilgiyi arttırmıştır [87], [88].

Derin öğrenme yaklaşımlarına aşağıdaki yöntemler örnek olarak verilebilir [90]:

- Convolutional Neural Network



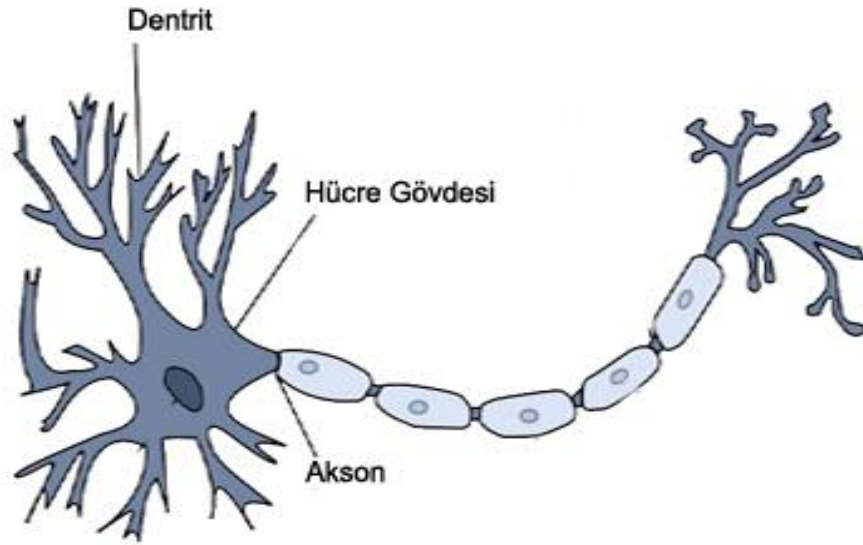
Şekil 3.8. Derin öğrenme, makine öğrenmesi ve yapay zeka arasındaki ilişki [86], [89].

- Deep Neural Network
- Deep Belief Network
- Convolutional Deep Belief Network
- Deep Boltzmann Machines
- Stacked Auto-Encoders
- Deep Stacking Networks
- Tensor Deep Stacking Networks
- Spike-and-Slap RBMs
- Compound Hierarchical-Deep Models
- Deep Coding Network
- Deep Kernel Machines
- Deep Q-Network

### 3.2.1. Yapay sinir ağları

CNN yapısını anlayabilmek için YSA yapısı ve içerisinde barındırdığı yöntemleri iyi anlamak gerekir. Çünkü, daha önce de bahsedildiği üzere CNN'ler YSA mimarisini kullanan yöntemlerdir. YSA insan beyinde bulunan sinir hücrelerinden (nöron) ilham alınarak geliştirilmiştir. Şekil 3.9.'da biyolojik bir sinir hücresi verilmiştir. Bir insanda bulunan çok sayıda sinir hücresi birbirleri ile kimyasal ve elektriksel yollarla iletişim

kurar. Bir sinir hücresi: hücre gövdesi, dentrit ve akson olmak üzere 3 farklı kısımdan oluşur. Dentritler üzerinden gelen sinyaller hücre gövdesinde transer işleme tabi tutulur ve aksonlara yönlendirilir. Hücre gövdesinden gelen sinyaller aksonlar aracılığı ile bir sonraki sinir hücresine iletilirler.



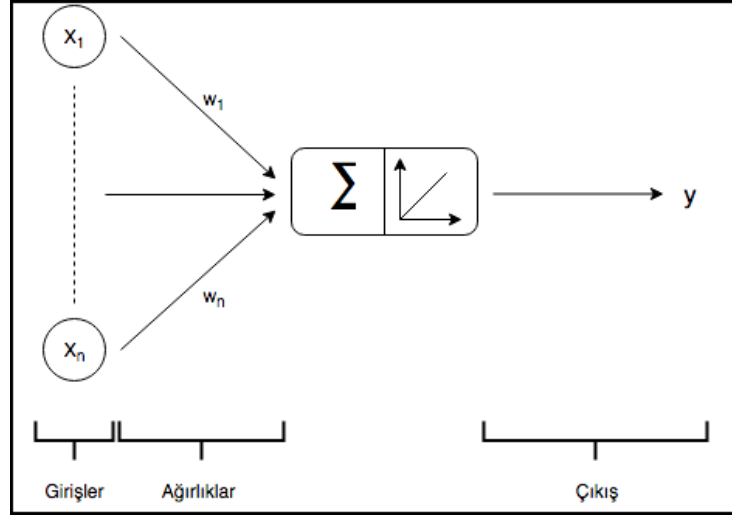
Şekil 3.9. Biyolojik sinir hücresi [91].

Bir YSA modelini aşağıdaki başlıklara göre karakterize etmek mümkündür [92]:

- Ağ topolojisi
- Aktivasyon fonksiyonu
- Öğrenme stratejisi
- Öğrenme kuralı

### 3.2.1.1. Yapay sinir hücresi yapısı

Biyolojik sinir hücresinden esinlenilerek geliştirilen yapay sinir hücresinin sembolik gösterimi Şekil 3.10.'da gösterilmiştir. Şekilde görüleceği üzere yapay sinirin girdileri aldığı bir giriş katmanı, girişlerin değerlendirildiği bir gövdesi ve sonuçların elde edildiği bir çıktı katmanı vardır.



Şekil 3.10. Tekli yapay sinir ağı modeli.

Denklem 3.8’de tekli sinir ağına gelen girdilerin ağırlıklar ile işleme alınıp, yapay sinir hücresine ait bias değerinin eklenmesi formülize edilmiştir. Toplam değer elde edildikten sonra bu sonuç aktivasyon fonksiyonuna gönderilir. Aktivasyon fonksiyonunda sonuç belirlenerek ağın çıkışına yansıtılır.

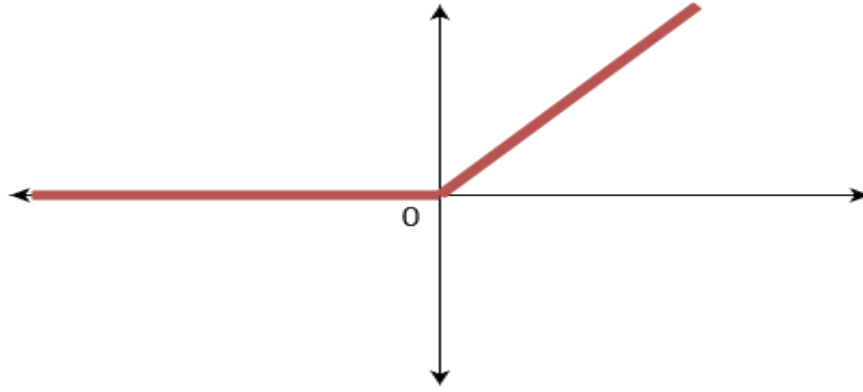
$$y = f\left(b + \sum_{i=0}^n x_i w_i\right) \quad (3.8)$$

### 3.2.1.2. Aktivasyon fonksiyonları

YSA, yapısında eşikleme amaçlı aktivasyon fonksiyonlarını kullanır. Bu fonksiyonlardan Rectified Linear Unit (ReLU) verdiği başarılı sonuçlardan ötürü doğrusal olmayan gerçek hayat problemlerinde sıklıkla kullanılır [88]. YSA içinde, Sigmoid ve Tanh fonksiyonları yine aynı amaçla kullanılan diğer aktivasyon fonksiyonu yöntemlerindedir. Aktivasyon fonksiyonu olarak genellikle doğrusal olmayan fonksiyonlar kullanılır. Bunun sebebi, mühendislik problemlerinin genelde doğrusal problemler olmamasıdır. Aktivasyon fonksiyonu seçilirken dikkat edilen bir diğer parametre ise türevi kolay alınabilen bir fonksiyonun tercih edilmesidir. Böylece ağ geri besleme ile güncellenirken aktivasyon fonksiyonunun türevi de kolaylıkla alınabilir.

*Rectified Linear Unit (ReLU)*: İyileştirilmiş Doğrusal Birim, ReLU, Denklem 3.9'da verilen formüle göre çıkış üretmektedir. ReLU'nun giriş parametresi olarak aldığı değer negatif veya sıfır ise sıfır, değilse kendisini döndürür. ReLU'yu diğerlerinden ayıran en önemli özellik, diğerlerine göre daha hızlı çalışmasıdır. ReLU aktivasyon fonksiyonu grafiği Şekil 3.11.'de verilmiştir.

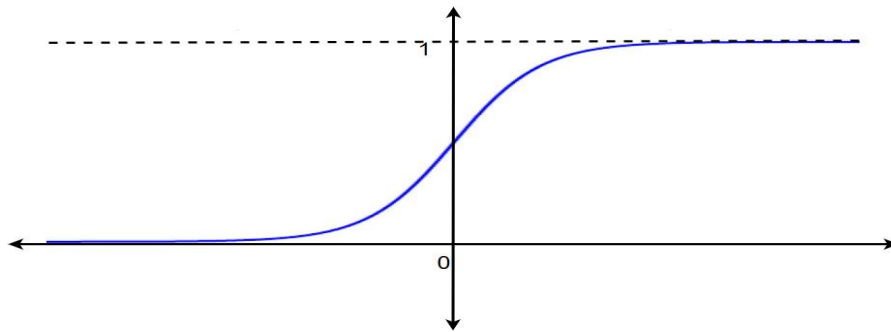
$$G(z) = \max\{0, z\} \quad (3.9)$$



Şekil 3.11. ReLU aktivasyon fonksiyonu grafiği.

*Logistic sigmoid*: Sigmoid aktivasyon fonksiyonu eşikleme amaçlı kullanılan diğer bir aktivasyon fonksiyonudur. Denklem 3.10'a göre çıkış değeri üreten sigmoid fonksiyonunun grafiği Şekil 3.12.'de verilmiştir.

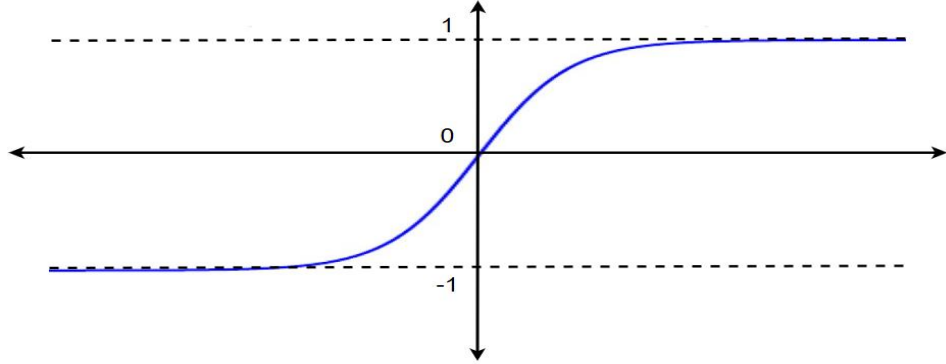
$$G(z) = \sigma(z) = \frac{1}{1 + e^{-z}} \quad (3.10)$$



Şekil 3.12. Sigmoid aktivasyon fonksiyonu grafiği.

*Hyperbolic Tangent*: Tanh aktivasyon fonksiyonu Denklem 3.11'e göre -1 ve +1 aralığında çıkış üretir. Şekil 3.13.'te tanh fonksiyonunun grafiği verilmiştir.

$$G(z) = \tanh(z) = \frac{2}{1 + e^{-2z}} - 1 \quad (3.11)$$

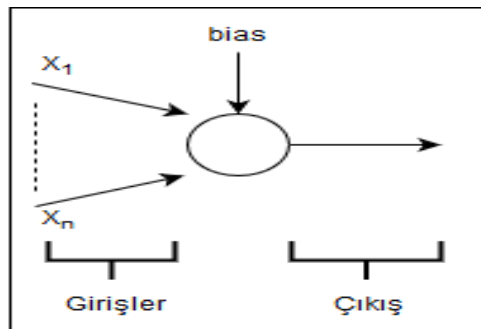


Şekil 3.13. Tanh aktivasyon fonksiyonu grafiği.

Sigmoid ve tanh aktivasyon fonksiyonları birbirine oldukça yakındır ve aralarında  $\tanh(z) = 2\sigma(2z) - 1$  şeklinde bir bağıntı vardır.

### 3.2.1.3. Tek katmanlı algılayıcı ağ yapısı

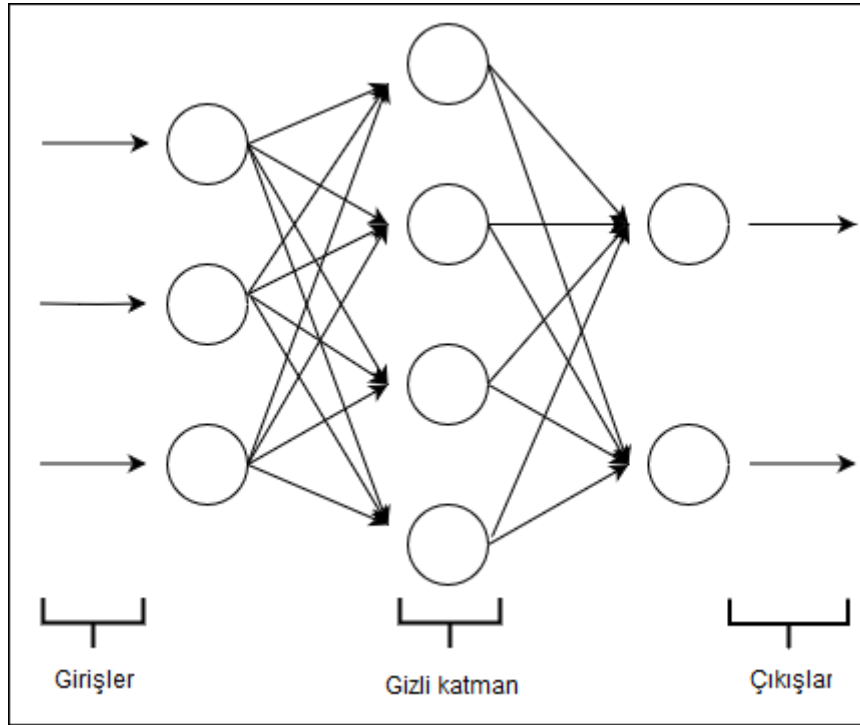
Bu tip ağ modelleri yapısı gereği veri uzayını bir doğru veya düzlem ile sınıflara böler ve doğrusal olmayan problemlere çözüm üretemezler [92]. Yalnızca giriş ve çıkış katmanları vardır. Şekil 3.14.'te tek katmanlı algılayıcıya bir örnek verilmiştir.



Şekil 3.14. Tek katmanlı algılayıcı örneği.

### 3.2.1.4. Çok katmanlı algılayıcı ağ yapısı

Daha önce de bahsedildiği üzere tekli yapay sinir ağ modeli ile doğrusal olmayan problemlere çözüm üretilemiyordu. Bu sorunun üstesinden gelebilmek için çok katmanlı algılayıcılar geliştirildi. Çok katmanlı algılayıcılar bünyesinde en az bir adet gizli katman barındırır. Bu ağ XOR mantıksal işlemi çözebilmek için geliştirilmiş olsa da sonraki süreçte birçok mühendislik problemine sonuç üretmiştir. Şekil 3.15.'te çok katmanlı algılayıcıya bir örnek verilmiştir. Şekilde de görüleceği üzere, giriş katmanındaki her yapay sinir hücresi ara katmandaki her yapay sinir hücresine bağlıdır. Ara katmandaki her yapay sinir hücresi de çıkış katmanındaki her bir yapay sinir hücresine bağlıdır.



Şekil 3.15. Çok katmanlı algılayıcı örneği.

### 3.2.1.5. Geriye yayılım algoritması

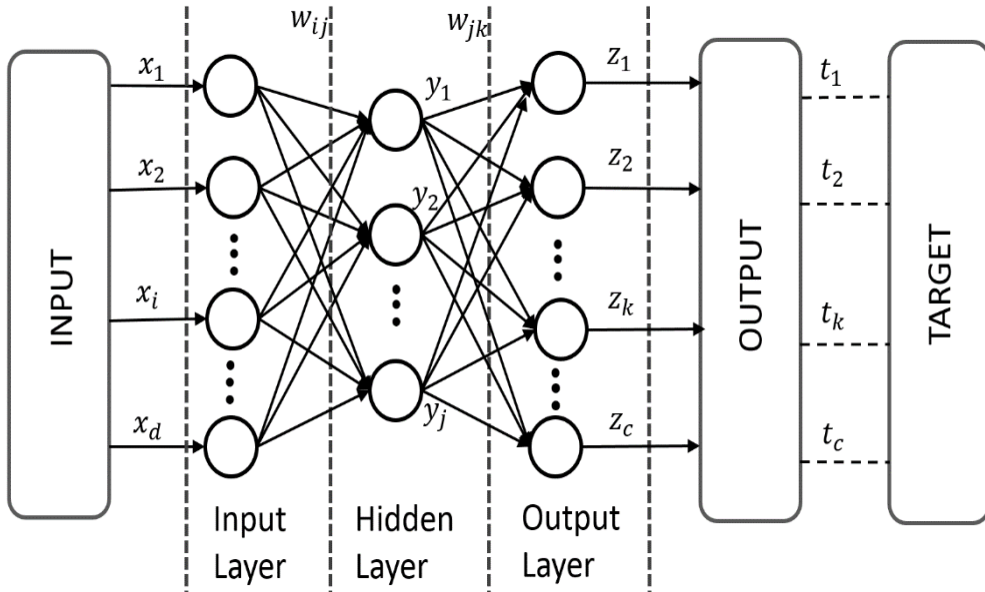
Yapay sinir ağlarında sıklıkla kullanılan eğitim yöntemlerinden geriye yayılım algoritması, geri yayılım yaklaşımı ile aşağıda verilen güncellemeleri ağ üzerinde gerçekleştirir. Bu güncellemeler ağın çıktısı ile beklenen çıktı arasındaki hatayı



minimize etmeyi amaçlar. Bu hata değerleri en küçük ortalamaların karesi (Least Mean Square - LMS) hatası ile hesaplanır. Bu kısım geriye yayılım algoritmasının ana bileşenidir. LMS hata değeri, ağda tahmin edilen sonuçlar ile olması gereken sonuç değerlerinin arasındaki farkların kareleri toplamına eşittir. Bu işlem Denklem 3.12’de verilmiştir.

$$J(w) = \frac{1}{2} \sum_{k=1}^c (t_k - z_k)^2 \quad (3.12)$$

Denklemden  $z$  tahmin edilen sonuç vektörünü,  $t$  ise olması gereken sınıflandırma sonuç vektörünü göstermektedir.  $c$  vektör uzunluğunu,  $w$  ise ağdaki ağırlıkları sembolize etmektedir. Şekil 3.16., denklemler ve denklemlerde kullanılacak parametreleri daha iyi takip edebilmek için aşağıda verilmiştir.



Şekil 3.16. Geriye yayılım algoritmasını anlamak için örnek bir ağ yapısı [93].

Başlangıçta atanan ağırlıklar *gradient descent* yöntemine göre hatayı indirmek için güncellenirler. Bu ağırlıklardaki değişim miktarı *öğrenme oranı* (*learning rate*) olarak tanımlanır ve  $\eta$  ile gösterilir.

$$\Delta w_{mn} = -\eta \frac{\partial J}{\partial w_{mn}} \quad (3.13)$$

Giriş katmanı ve gizli katman ağırlıklarını güncellerken ağırlıkların bağımlılıkları göz önüne alınmalıdır. Eğer hata, gizli katmanın çıkış ağırlıklarına açıkça bağlı değilse zincir kuralında (*chain rule*) türevsel genişleme kullanılmalıdır. Örneğin bir gizli katmanı bulunan bir ağı, gizli katman ile çıkış katmanı arasındaki ağırlıklar için öğrenme kuralı aşağıdaki gibidir. Bu işlem LMS hatasının ( $J(w)$ ) parçalı türevini bulmak için zincir kuralı uygulanarak yapılır.

$$\frac{\partial J}{\partial w_{kj}} = \frac{\partial J}{\partial net_k} \frac{\partial net_k}{\partial w_{kj}} \quad (3.14)$$

$w_{k,y}$  gizli katman ve çıkış katmanı arasındaki ağırlıkları temsil eder.  $\delta_k$  çıkış değerlerinin duyarlılığını verir ve aşağıdaki Denklemler 3.15 ile gösterilir.

$$\delta_k = -\frac{\partial J}{\partial net_k} \quad (3.15)$$

$net_k$  çıkış katmanının girişini temsil eder. Denklem 3.13'ün türevi alındığında;

$$-\frac{\partial J}{\partial net_k} = \frac{\partial J}{\partial z_k} \frac{\partial z_k}{\partial net_k} = (t_k - z_k) f'(net_k) \quad (3.16)$$

$$\frac{\partial net_k}{\partial w_{kj}} = y_j \quad (3.17)$$

Denklemler 3.17'nin türevi alındığında ağırlık değişimi olan  $\Delta w_{k,j}$  bulunur.

$$\Delta w_{kj} = \eta \delta_k y_j = \eta (t_k - z_k) f'(net_k) y_j \quad (3.18)$$

Giriş katmanı ile gizli katman arasındaki ağırlıklar için de aynı işlemler yapılır. Zincir kuralını bir kez daha işleterek

$$\frac{\partial J}{\partial w_{ji}} = \frac{\partial J}{\partial y_j} \frac{\partial y_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}} \quad (3.19)$$

$$\frac{\partial J}{\partial y_j} = \frac{\partial}{\partial y_j} \left[ \frac{1}{2} \sum_{k=1}^c (t_k - z_k)^2 \right] \quad (3.20)$$

$$= - \sum_{k=1}^c (t_k - z_k) \frac{\partial z_k}{\partial y_j} \quad (3.21)$$

$$= - \sum_{k=1}^c (t_k - z_k) \frac{\partial z_k}{\partial net_k} \frac{\partial net_k}{\partial y_j} \quad (3.22)$$

$$= - \sum_{k=1}^c (t_k - z_k) f'(net_k) w_{jk} \quad (3.23)$$

Denklem 3.23, çıkış sinyalinin LMS hata değerine etkisini verir.

$$\delta_j = f'(net_j) \sum_{k=1}^c w_{kj} \delta_k \quad (3.24)$$

Gizli katman ile giriş katmanı arasındaki ağırlıkların güncellenmesi Denklem 3.25'te verilen denklem ile yapılabilir.

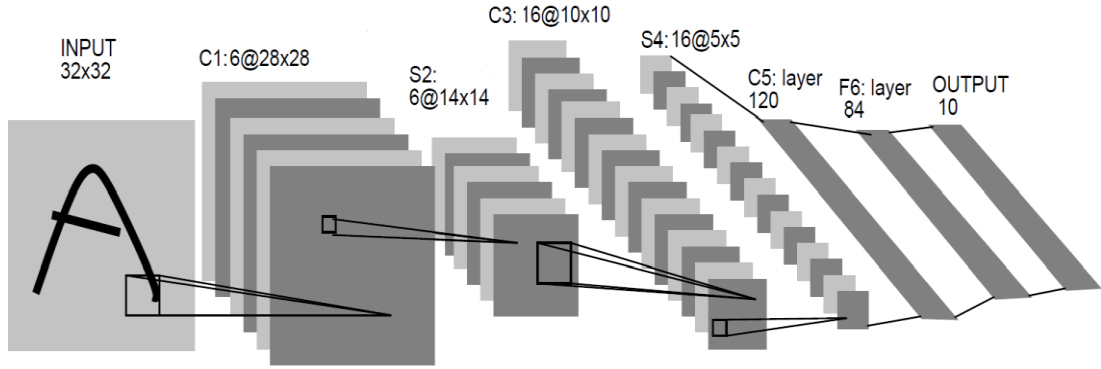
$$\Delta w_{ji} = \eta x_i \delta_j = \eta x_i f'(net_j) \sum_{k=1}^c w_{kj} \delta_k \quad (3.25)$$

$\delta_j$  ve  $\delta_k$  gizli katman ve çıkış katmanının duyarlılık değerleridir.

### 3.2.2. Evrişimsel sinir ağları (Convolutional Neural Network - CNN)

#### 3.2.2.1. Giriş

Evrişimsel sinir ağları bilgisayar görmesi alanında oldukça başarılı sonuçlar üreten akademik ve ticari olarak son zamanlarda sıklıkla kullanılan bir derin öğrenme yöntemidir. CNN'in tarihi Hubel ve Wiesel'in 1968 yılında yaptığı ve "memeli canlılarda görme işlemi nasıl gerçekleşir" ana temalı çalışmalarına [94] dayanmaktadır [95]. CNN'in ilk başarılı uygulaması [96] 1998 yılında Lecun ve ark. tarafından optik karakter tanıma yapma amacıyla geliştirilmiştir. Bu çalışmada geliştirdikleri ağ (*LeNet*) Şekil 3.17.'de verilmiştir.



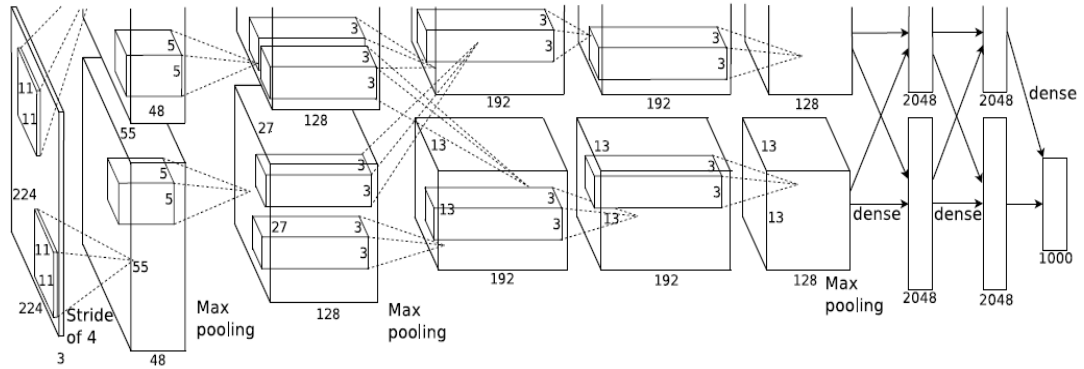
Şekil 3.17. LeNet mimarisi [96].

CNN, çok katmanlı algılayıcıların gelişmiş bir versiyonudur. CNN adındaki "convolution" kelimesi konvolüsyon işlemi ifade etmektedir. Yukarıdaki şekilden de izlenebileceği gibi CNN, temel olarak iki kısımdan oluşmaktadır. Birinci kısım konvolüsyon (*convolution*) ve havuzlama (*pooling*) katmanlarını içerip derin özniteliklerin çıkarımı ile ilgilidir. İkinci bölüm ise tam bağlantılı (*fully connected*) katmanlarda sınıflandırma işlemlerini icra eder.

Görüntülerin sınıflandırılması probleminde öncelikle uygun özniteliklerin çıkarılması gerekmektedir. Bu aşama bilgisayar görmesi problemlerinin zorlayıcı bir aşamasıdır. CNN bu işlemi kendi içerisinde otomatik olarak gerçekleştirerek öznitelik haritalarını elde eder. Ardından tam bağlantılı katmanlarda sınıflandırma işlemini gerçekleştirir.

Derin öğrenme yöntemlerinden olan CNN, görüntü verileri üzerinde yüksek performanslı uygulamalar ortaya koymakta olup, çok çeşitli alanlarda başarılı bir şekilde uygulanmaktadır. Bu çalışma alanlarına biyomedikal görüntü işleme [97]–[99], yüz ifadesi tespiti [10], [100], [101], yaş tespiti [102]–[104], nesne tespiti [105]–[107], yüz bileşenleri bölütleme [108], [109], yüz tespiti [110]–[112] ses tanıma [113]–[115] ve cinsiyet tespiti [116], [117] örnek olarak verilebilir.

CNN üzerine çalışan bazı gruplar, geliştirdikleri ağları büyük hacimli veri kümeleri ile eğitip paylaşma açmışlardır. Bu şekilde eğitilmiş ağların araştırmacılarla paylaşılması, bu ağların başka alanlara da uyarlanabilmesinin önünü açmıştır. Eğitilmiş ağın sonunda bulunan sınıflandırma katmanları yeniden oluşturulup yeni verilerle eğitildiğinde başarılı sonuçlar alınabilmektedir. Bu işlem literatürde “*transfer learning*” olarak geçmektedir. Şekil 3.18. ve Şekil 3.19.’da bu alanda bilinen ve popüler olan iki CNN mimarisi verilmiştir.



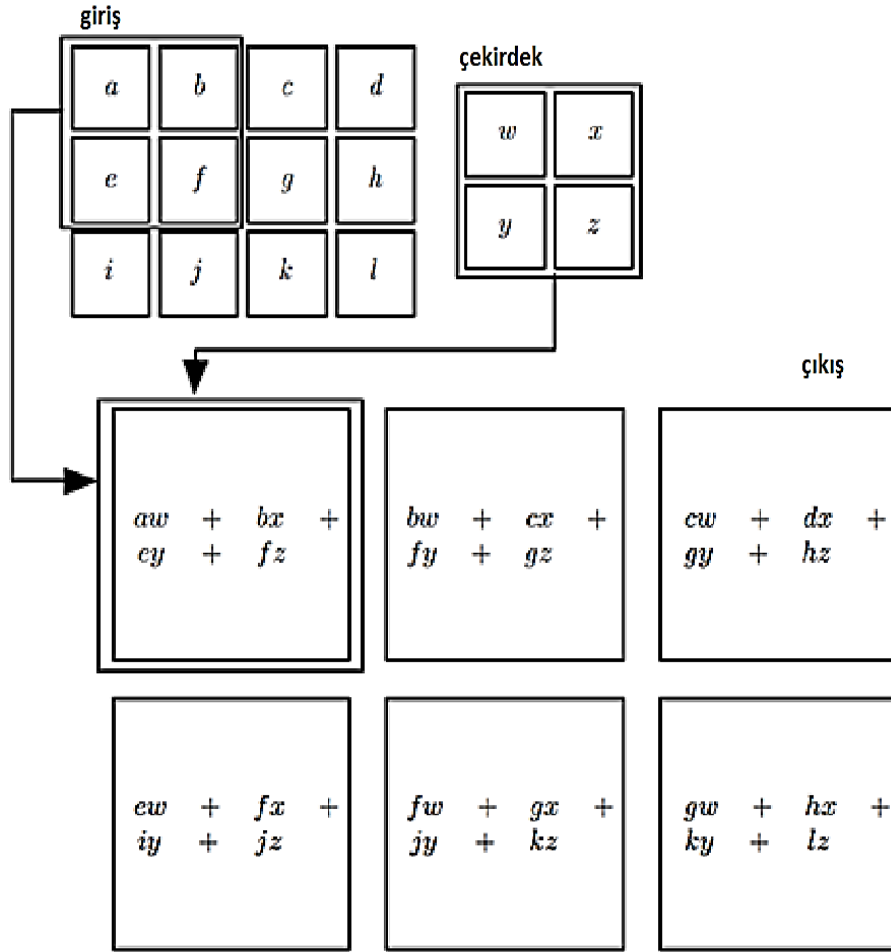
Şekil 3.18. AlexNet mimarisi [118].

### 3.2.2.2. Konvolüsyon katmanı (convolution layer)

Konvolüsyon, özelleştirilmiş bir çeşit lineer işlemidir. Bu işlem girdi olarak genellikle iki boyutlu veri dizilerini ve bu diziler üzerinde dolaşabilecek daha küçük iki boyutlu çekirdek dizilerini parametre olarak alır. Konvolüsyon işlemi “ $*$ ” ile gösterilir ve matematiksel ifadesi Denklem 3.26.’da verilmiştir.

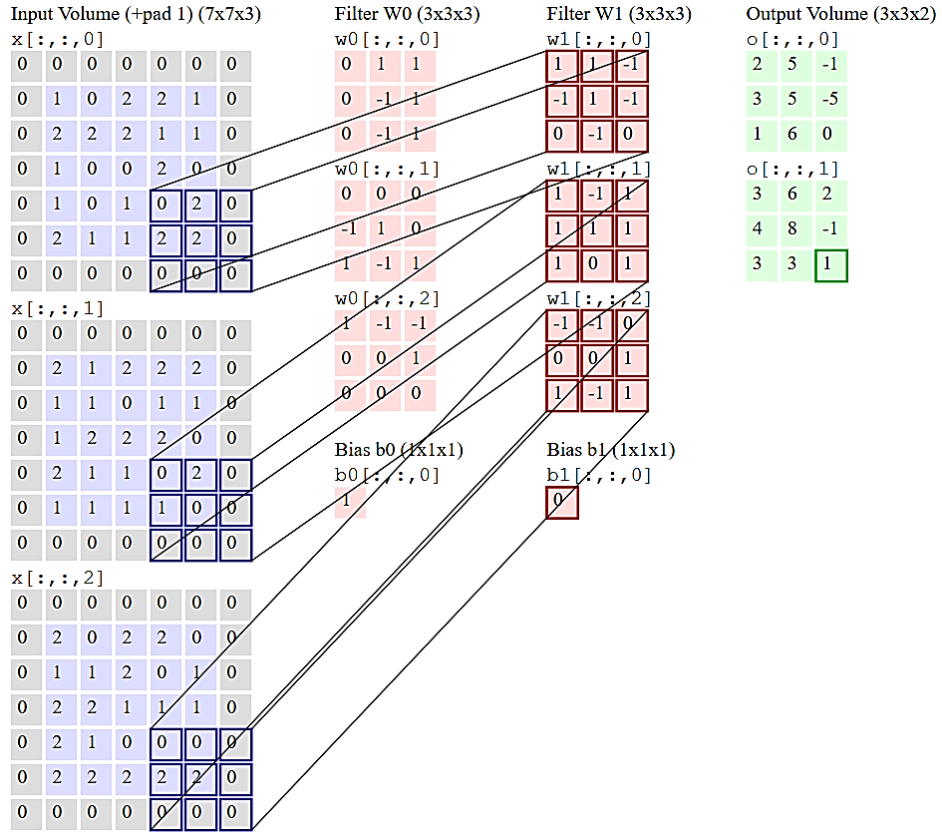


Şekilden de görüleceği üzere çekirdek penceresinin kayma oranı (*stride*) 1 pikseldir ve görüntü kenarlarına ekleme (*padding*) yapılmamıştır.



Şekil 3.20. Konvolüsyon işleminin şekilsel gösterimi [88].

Şekil 3.21.'de 5x5x3 lük bir girdi matrisi 3x3x3'lük iki adet çekirdek matrisi ile konvolüsyon işlemine tabi tutulmuştur. Görüntünün kenarlarına padding olarak birer sıra sıfır eklenip, veri matrisi 7x7x3'lük boyuta getirilmiştir. Çekirdek çerçeveleri ikişer ikişer veri matrisi üzerinde kaydırılarak işlemler gerçekleştirilmiştir. Kaydırma işlemi (*stride*) çıkış katmanının boyutlarına etki eder. Örneğin kaydırma değerinin artması ile çıktı boyutu küçülür. Ayrıca, kaydırma işlemi ile işlenen bölgelerin üst üste binmesi engellenebilir. Örneğin 3x3'lük bir çekirdek pencere için kaydırma değeri 3 olarak seçilirse, çerçeveler her adımda üçer piksel kayacağı için veri matrisi üzerinde işlenen bölgeler örtüşmemiş olur. Padding işlemi ise görüntü kenarlarındaki bilgilerin kaybını engeller. Bu durumda performansta artışına imkan sağlar. Ayrıca padding kullanılmayan durumlarda hacimsel azalma gözlemlenir.



Şekil 3.21. Konvolüsyon işlemi örneği [120].

Konvolüsyon işlemleri sonrasında CNN, ilgili katmanda bulunan çekirdek sayısı kadar çıktı üretir. Bu çıktılar matrislerdeki sayısal değerler olduklarından görüntü olarak gösterilebilirler. Şekil 3.22.'de 96 adet çekirdek içeren bir konvolüsyon katmanına çıktı olarak verilen bir görüntü verilmiştir. Şekil 3.23.'te ise konvolüsyon katmanı çıktıları gösterilmiştir.



Şekil 3.22. Konvolüsyon işlemi sonuçlarını görmek için örnek bir girdi görüntüsü [121].





Şekil 3.23. Örnek girdiye konvolüsyon çekirdeklerinin uygulanması sonucu oluşan katman çıktıları [121].

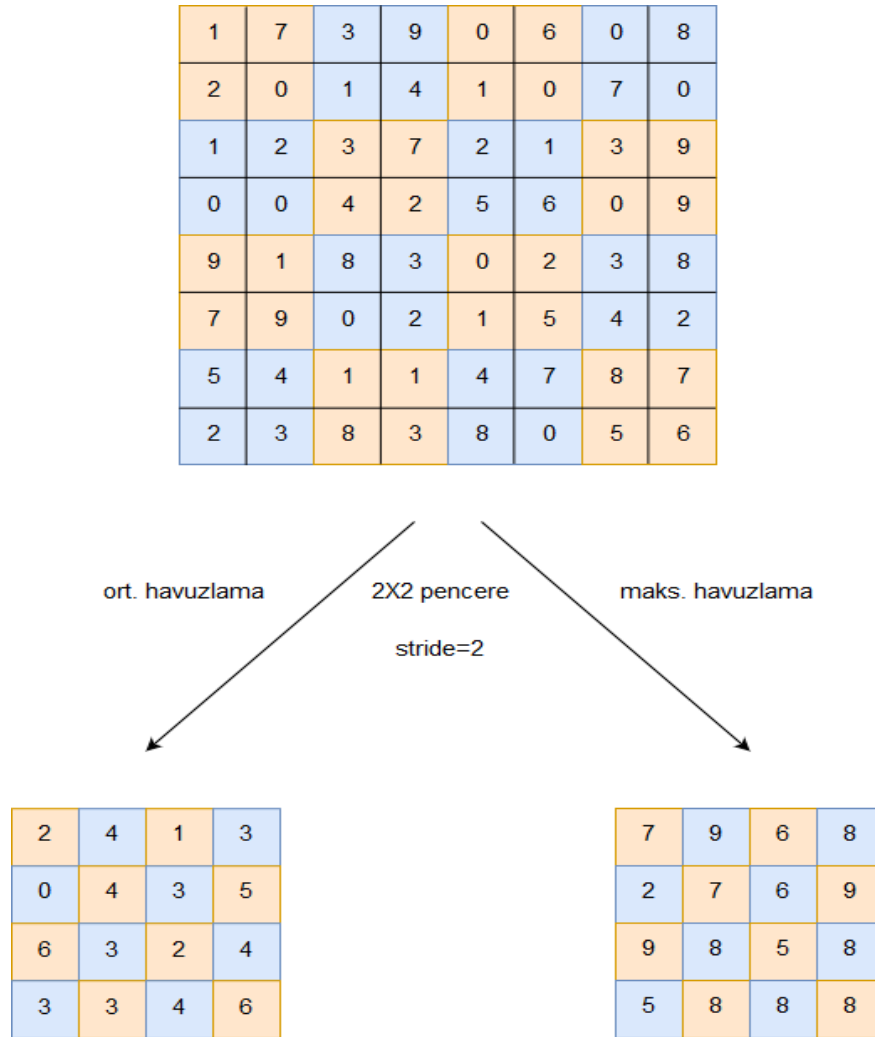
96 adet görüntüden de görüleceği üzere her bir çekirdek farklı bir özneliğe yönelir, ilgili özneliği yakalamaya çalışır. Şekil 3.24.'te kırmızı piksellere odaklanan bir çekirdek ile konvolüsyon işlemi sonrası oluşan çıktı görülmektedir. Çıktı görüntüsü incelendiğinde; orijinal görüntüde piksel görüntüsü kırmızıya ne derece yakın ise üretilen çıktıda da ilgili piksel beyaza o derece yakındır.



Şekil 3.24. Konvolüsyon işlemi sonucu ilgili çekirdeğin kırmızı renge duyarlı olduğunu gösteren örnek konvolüsyon çıktısı [121].

### 3.2.2.3. Havuzlama katmanı (pooling layer)

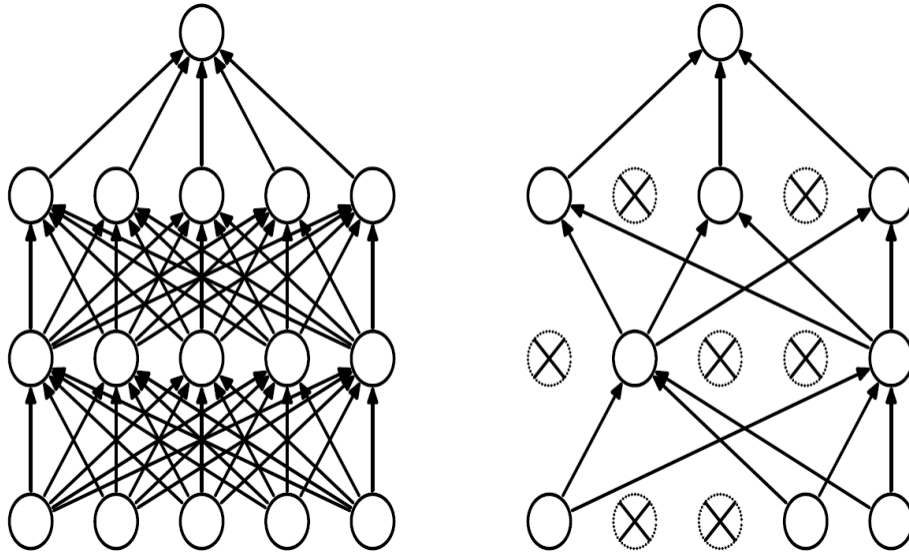
Havuzlama katmanı, giriş verisinin konvolüsyon işlemine tabi tutulmasının ardından oluşan öznitelik haritalarını alt bölgelere ayırarak seçilen yönteme göre alt bölgeyi temsilen bir değer belirler. Bu yöntemler ortalama ve maksimum havuzlama (average pooling, maximum pooling) olmak üzere iki yaklaşımla çalışır. Ortalama havuzlama yönteminde, belirtilen çerçeve boyutları ile girdi verisi üzerinde dolaşılırken çerçeve içine denk gelen pixel değerlerinin ortalaması alınır iken, maksimum havuzlamada ise en yüksek değer alınır. Şekil 3.25.'te havuzlama işlemine ait bir örnek verilmiştir. Örnekte havuzlama işlemine giren verinin üzerinde dolaşmak üzere 2x2'lik bir pencere belirlenmiştir ve kayma miktarı 2 olarak seçilmiştir.



Şekil 3.25. Havuzlama işlemi örneği.

### 3.2.2.4. Dropout katmanı

Dropout katmanı, ağlarda aşırı uyumsuzluğun (*overfitting*) önüne geçmek için kullanılan bir tekniktir. Bu teknikte, belirli bir olasılık değeri ile gizli katmandaki düğüm noktalarında elemeler yapılır. Şekil 3.26.'da sağ tarafta yer alan ağ, sol taraftaki ağa dropout uygulandığındaki etkisini göstermektedir.



Şekil 3.26. Dropout etkisi [122].

### 3.2.2.5. Tam bağlantılı katman (fully connected layer) ve Softmax

Konvolüsyon ve havuzlama katmanlarını içeren ağın son katmanları; sınıflandırma işlevlerini yerine getiren tam bağlantılı ve softmax katmanlarıdır. Tam bağlantılı katman adından da anlaşılacağı üzere, kendinden önceki katmanın tüm elemanlarıyla bağlantılıdır ve sınıflandırmadan önceki aşama için, sınıf etiketlerine dair olasılıksal değerler içerir. Bu katmandan elde edilen sayısal değerlerin hangi sınıf ile ilişkili olduğuna softmax katmanı karar verir ve bu olasılıksal değerleri  $[0,1]$  aralığına çeker. Çıktılar her bir sınıfa ait olasılıkları ifade ettiği için toplamları 1 olmak zorundadır. Değerler incelendiğinde; en büyük değer hangi sınıfa ait ise, ağ o sınıf etiketini çıktı olarak üretir. Denklem 3.27.'de softmax fonksiyonunun formülü verilmiştir.

Denklemdede;  $x_i$  tam bağlantılı katmanın çıkış vektörünü,  $i$  örnek sayısını,  $j$  ise sınıf sayısını temsil etmektedir.

$$\text{softmax}(x)_i = \frac{\exp(x_i)}{\sum_{j=1}^n \exp(x_j)} \quad (3.27)$$

## **BÖLÜM 4. DENEYSEL SONUÇLAR**

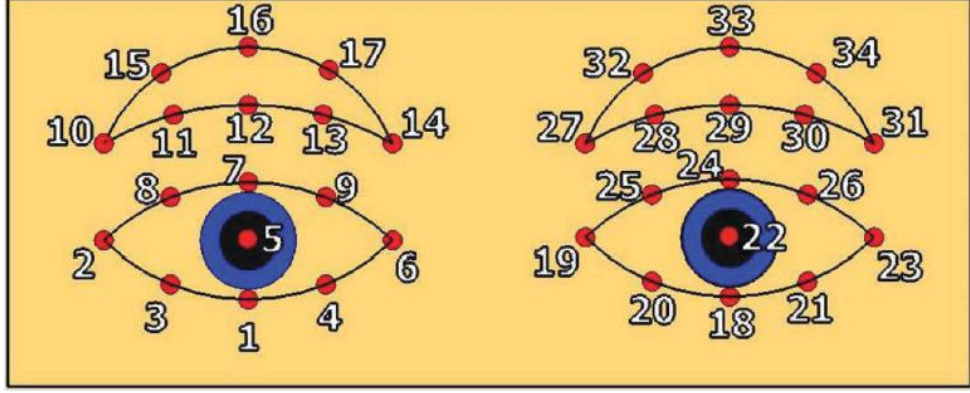
### **4.1. Giriş**

Bu tez kapsamında kısmi yüz görüntüleri ve tam yüz görüntüleri kullanılarak iki farklı yaklaşımla, yüz ifadesi tespiti uygulamaları geliştirilmiştir. İlk yaklaşımla geliştirilen uygulamada yüz görüntülerindeki kaş ve göz bölgelerinden anahtar noktalar bulunmuştur. Bu noktaların koordinat verileri kullanılarak, geometrik işlemlerle öznitelik kategorileri elde edilmiştir. Elde edilen bu kategorilerdeki geometrik yüz öznitelikleri SFS, SFFS ve ReliefF yöntemi ile öznitelik seçimi işlemine tabi tutulmuş ve gürbüz öznitelikler seçilmiştir. Her kategoriden seçilen gürbüz öznitelikler klasik makine öğrenmesi yöntemlerinden SVM ve KNN ile sınıflandırılıp yüz ifadesi tespiti yapılmıştır. Diğer yaklaşımla geliştirilen uygulamalarda ise tüm yüz görüntüleri, derin öğrenme yöntemlerinden CNN ile bölütleme ve sınıflandırma işlemlerine tabi tutulmuştur. Bölütleme işleminde yüz görüntülerinde kaşlar, gözler ve ağız bölgeleri; görüntülerdeki geriye kalan bölgelerden ayırt edilmiştir. Son olarak; orijinal görüntüler, bölütlenmiş görüntüler ve bu iki görüntü tipinin birleştirilmesiyle oluşan verilerden CNN kullanılarak yüz ifadesi tespiti yapıp sonuçlar karşılaştırılmıştır.

### **4.2. Kaş ve Göz Bölgelerinden Klasik Makine Öğrenmesi Yöntemleriyle Yüz İfadesi Tespiti**

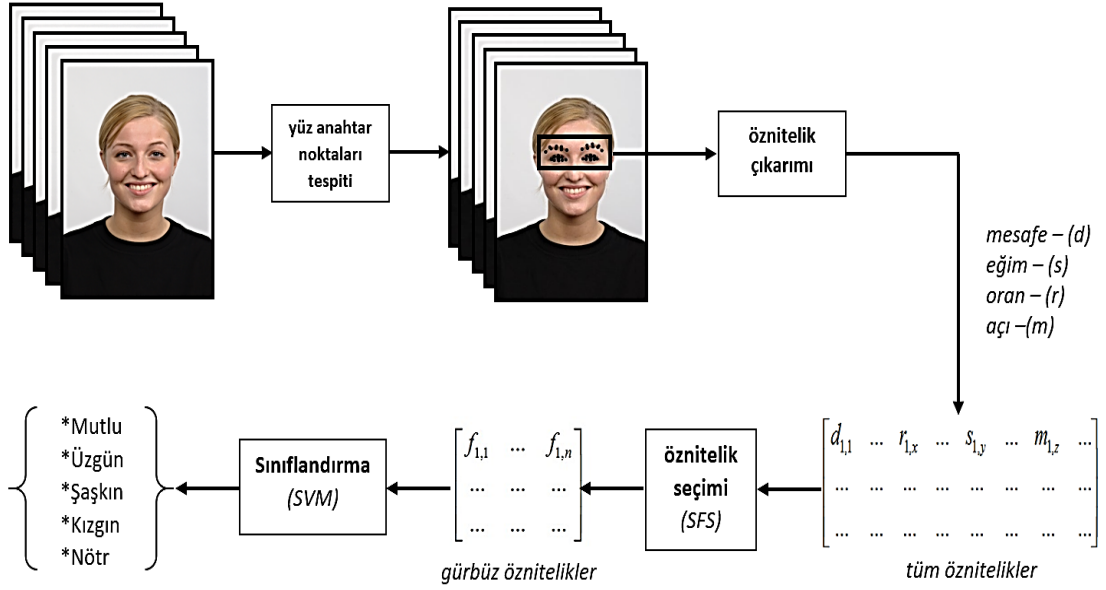
Bu tez kapsamında yapılan ilk uygulamada; yüz görüntüleri üzerinde 83 adet anahtar nokta bulabilen Face++ SDK [53] kullanılmıştır. Bu çalışmada kaş ve göz bölgeleri kullanıldığı için bu bölgelere ait 34 adet anahtar nokta elde edilmiştir. Bu noktalar Tablo 4.1.'de verilmiştir.

Tablo 4.1. Kaş ve göz bölgelerinde seçilen yüz anahtar noktaları.



1	Sağ göz, alt	18	Sol göz, alt
2	Sağ göz, sağ köşe	19	Sol göz, sağ köşe
3	Sağ göz, sağ alt çeyrek	20	Sol göz, sağ alt çeyrek
4	Sağ göz, sol alt çeyrek	21	Sol göz, sol alt çeyrek
5	Sağ göz, göz bebeği	22	Sol göz, göz bebeği
6	Sağ göz, sol köşe	23	Sol göz, sol köşe
7	Sağ göz, üst	24	Sol göz, üst
8	Sağ göz, sağ üst çeyrek	25	Sol göz, sağ üst çeyrek
9	Sağ göz, sol üst çeyrek	26	Sol göz, sol üst çeyrek
10	Sağ kaş, sağ köşe	27	Sol kaş, sağ köşe
11	Sağ kaş, sağ alt çeyrek	28	Sol kaş, sağ alt çeyrek
12	Sağ kaş, alt	29	Sol kaş, alt
13	Sağ kaş, sol alt çeyrek	30	Sol kaş, sol alt çeyrek
14	Sağ kaş, sol köşe	31	Sol kaş, sol köşe
15	Sağ kaş, sağ üst çeyrek	32	Sol kaş, sağ üst çeyrek
16	Sağ kaş, sağ üst	33	Sol kaş, sağ üst
17	Sağ kaş, sol üst çeyrek	34	Sol kaş, sol üst çeyrek

Elde edilen bu noktaların koordinat bilgileri kullanılarak; var olan her iki nokta arasındaki mesafeler, her iki noktanın birleşimi ile oluşan doğru parçası uzunlukları, her doğrunun eğimi ve her iki doğrunun veya uzantılarının kesişimi ile ortaya çıkan açılarının tanjant değerleri öznitelik olarak hesaplanmıştır. Çok sayıdaki bu özniteliklerden gürbüz olanlarını seçmek için SFS yöntemi kullanılmıştır. Elde edilen gürbüz öznitelikler SVM ve K-NN ile sınıflandırılmıştır. Sistemin akış diyagramı Şekil 4.1.'de verilmiştir.



Şekil 4.1. Önerilen sistem akışı.

Bu çalışmada mutlu, üzgün, şaşkın kızgın ve nötr olmak üzere kaş ve göz bölgelerine bakılarak tespit edilebilen 5 farklı yüz ifadesi sınıflandırılmıştır. Yüz ifadesi çalışmalarında sıklıkla kullanılan iğrenme gibi bazı duygular ağız bölgesinde yoğun olarak gözlemlendiği için bu uygulamaya dahil edilmemiştir. Çünkü bu çalışmada amaç yüzün üst bölgesinin analizi ile yüz ifadesi tespittir.

Uyulama Matlab ortamında gerçekleştirilmiş olup RaFD, JAFFE, KDEF ve CK veritabanları eğitim ve test amaçlı kullanılmıştır. Temel eğitim ve test işlemi RaFD üzerinde yapılmış olup, diğer veri tabanları sistemin genelleme kapasitesini test etme amacıyla kullanılmıştır. Ayrıca, önerilen sistem yalnızca ön yüz görüntüleri ile çalışmaktadır; çünkü yüz ifadesi tespit edilmeye çalışılan yüzde, kaşlardan veya gözlerden biri veya her ikisinin de görülmemesi sistemin sağlıklı çalışmasını engellemektedir. Eğitim ve test işlemleri 2.6 GHz Intel Core i7 işlemcili, 32 GB RAM ve 6 GB ekran kartına sahip bir bilgisayar üzerinde yapılmıştır.

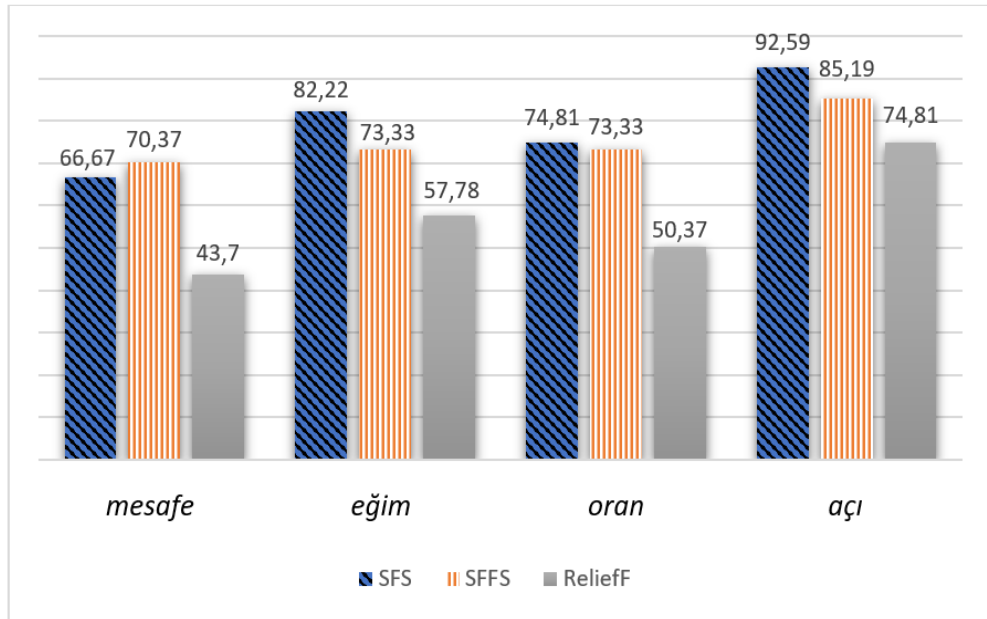
#### 4.2.1. Öznitelik çıkarımı

Bu çalışmanın ilk adımında, kaş ve göz bölgelerinde Face++ SDK ile bulunan yüz anahtar noktalarının koordinat bilgileri ( $p_i(x_i, y_i)$ ) kullanılarak öznitelik kümeleri

oluşturulmuştur. Bir yüz görüntüsü üzerinde bulunan 34 adet anahtar noktadan, her iki nokta arası uzaklıkları temsil eden 561 adet mesafe ( $|L_{ij}|$ ) verisi elde edilmiştir. Yine aynı şekilde her iki noktanın birleşimi ile oluşan doğruların eğim ( $m_{ij}$ ) verileri 561 eleman içeren bir başka öznelik kümesini oluşturmuştur. 561 adet doğru parçasının her birinin diğerlerine olan oranları ( $|L_{ij}|/|L_{kp}|$ ) ile 157.080 adet öznelik içeren üçüncü bir öznelik kümesi elde edilmiştir. Yine 561 adet doğru parçasının herhangi ikisinin ya da herhangi ikisinin uzantılarının kesişimi ile elde edilen açının tanjant değerleri ( $\tan(\alpha)$ ) 157.080 adet öznelik içeren dördüncü bir öznelik kümesini oluşturmuştur. Bu dört özneliğin birleşimi ile 315.282 adet veri içeren karma öznelik kümesi elde edilmiştir.

#### 4.2.2. Gürbüz özneliklerin seçimi

Çıkarılan özneliklerin hepsi ile çalışmak hem zaman açısından hem de iş yükü açısından verimsiz olacağından bu özneliklerden sınıflandırma sonucuna direk etki edecek olan gürbüz öznelikler seçilmelidir. Bu amaç için SFS, SFFS ve ReliefF algortimaları kullanılmıştır. Şekil 4.2.'de görüleceği üzere en yüksek başarımları SFS özneliği ile açı özneliğinde gözlemlenmiştir.



Şekil 4.2. Farklı öznelik seçim metodlarının sınıflandırma performansına etkileri.

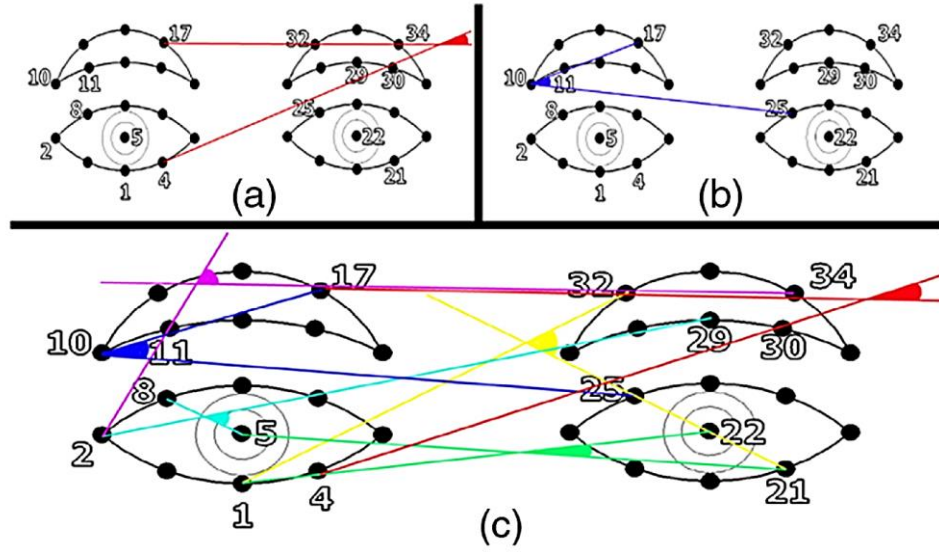


Şekil 4.2.'den de görüleceği üzere; ilk öznitelik grubu olan mesafe öznitelik kümesine ait gürbüz öznitelik seçimleri hariç, diğer üç grupta SFS ile seçilen özniteliklerin sınıflandırma sonuçları daha yüksek başarıma ulaşmıştır. Seçilen öznitelik grupları, toplam öznitelik sayıları ve SFS sonrası seçilen gürbüz öznitelik sayıları Tablo 4.2.'de verilmiştir.

Tablo 4.2.Öznitelik kümeleri, seçilen gürbüz öznitelikler ve sınıflandırma sonuçları.

Öznitelik tipi	Toplam öznitelik sayısı	SFS		Gürbüz öznitelikler	SVM başarımları oranları (%)
		sonrası gürbüz öznitelik sayısı			
uzunluk	561	4		$ L_{21,29} ,  L_{21,25} ,  L_{24,31} ,  L_{28,29} $	66,67
eğim	561	5		$m_{21,26}, m_{18,22}, m_{18,24}, m_{12,34}, m_{20,32}$	82,22
oran	157.080	4		$ L_{5,31}  /  L_{21,32} ,  L_{1,22}  /  L_{18,30} ,  L_{5,24}  /  L_{22,25} ,  L_{2,18}  /  L_{8,11} $	74,81
açı	157.080 (açı) 315.282(karma)	6		$\frac{m_{10,17}-m_{10,25}}{1+m_{10,17}.m_{10,25}}, \frac{m_{4,30}-m_{17,34}}{1+m_{4,30}.m_{17,34}}, \frac{m_{2,11}-m_{32,34}}{1+m_{32}.m_{34}}, \frac{m_{2,29}-m_{5,8}}{1+m_{2,29}.m_{5,8}}, \frac{m_{1,22}-m_{5,21}}{1+m_{1,22}.m_{5,21}}, \frac{m_{1,32}-m_{21,25}}{1+m_{1,32}.m_{21,25}}$	92,59

Tablo 4.2.'de  $|L_{ij}|$ ,  $p_i$  ve  $p_j$  noktaları arasında oluşan doğru parçasının uzunluğunu;  $|L_{ij}| / |L_{kp}|$  iki doğru parçasının uzunluklarının birbirine olan oranını;  $m_{ij}$ ,  $L_{ij}$  doğrusunun eğimini ve  $\frac{m_{ij}-m_{kp}}{1+m_{ij}.m_{kp}}$  ise  $L_{ij}$  ve  $L_{kp}$  doğrularının kesişimleri ile oluşan açının tanjant değerini göstermektedir. Tabloda ayrıca her bir öznitelik kümesinden SFS sonrası elde edilen gürbüz özniteliklerin SVM algoritması ile sınıflandırılması sonucunda ortaya çıkan başarımların yüzdeleri de verilmiştir. Sonuçlar incelendiğinde en yüksek başarımların açı öznitelikleri ile elde edildiği gözlemlenmiştir. Şekil 4.3.'de açı kümesinden seçilen tüm gürbüz öznitelikler verilmiştir.



Şekil 4.3. Açık özelliklerinden elde edilen gürbüz özellikler: (a) 17 ve 34. noktaların birleşmesiyle oluşan doğru ile 4 ve 30. noktaların birleşmesiyle oluşan doğrunun kesişimi ile oluşan açı; (b) 10 ve 17. noktaların birleşmesiyle oluşan doğru ile 10 ve 25. doğruların birleşmesiyle oluşan doğrunun kesişimi ile oluşan açı; (c) açı kümesinden seçilen tüm özellikler [74].

#### 4.2.3. Seçilen gürbüz özelliklerin sınıflandırılması

Göz ve kaş bölgelerinde bulunan anahtar noktaların koordinat bilgileri kullanılarak elde edilen özelliklerden yüz ifadesi tespiti sonucuna olumlu etki edecek olanları seçme işlemini gerçekleştirdikten sonraki adım sınıflandırma aşamasıdır. Seçilen açı kümesine ait gürbüz özellikler KNN ve SVM algoritmaları ile sınıflandırılmıştır. Bu sınıflandırma sonuçlarına ait karmaşıklık matrisleri Tablo 4.3. ve Tablo 4.4.'te verilmiştir.

Tablo 4.3. Açık özellik kümesinden seçilen gürbüz özelliklerin KNN ile sınıflandırılmasıyla oluşan karmaşıklık matrisi.

gerçek	tahmin				
	kızgın	mutlu	nötr	üzgün	şaşkın
kızgın	96,30	0	3,70	0	0
mutlu	14,81	81,48	3,70	0	0
nötr	3,70	3,70	74,07	14,81	3,70
üzgün	11,11	0	3,70	85,19	0
Şaşkın	0	0	7,41	0	92,59
<b>ortalama: %85,93</b>					

Tablo 4.4. Açık öznitelik kümesinden seçilen gürbüz özniteliklerin SVM ile sınıflandırılmasıyla oluşan karmaşıklık matrisi.

gerçek	tahmin				
	kızgın	mutlu	nötr	üzgün	şaşkın
kızgın	96,30	0	3,70	0	0
mutlu	7,41	92,59	0	0	0
nötr	3,70	3,70	88,89	0	3,70
üzgün	11,11	0	0	88,89	0
şaşkın	0	0	3,70	0	96,30
<b>ortalama: %92,59</b>					

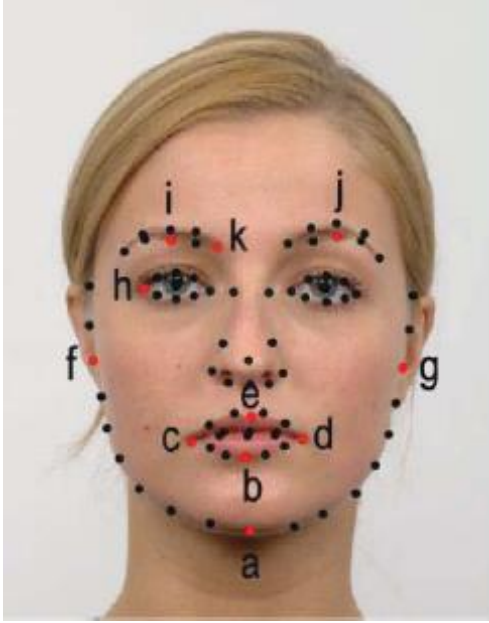
SVM sonuçlarına göre yanlış kabul oranı (False Accept Rate - FAR) ve yanlış ret oranı (false reject rate – FRR) ise Tablo 4.5.’de verilmiştir.

Tablo 4.5. Önerilen sistem için FAR ve FRR sonuçları.

Pozitif sınıf	Negatif sınıf	FAR	FRR
kızgın	mutlu, nötr, üzgün, şaşkın	0,056	0,037
mutlu	kızgın, nötr, üzgün, şaşkın	0,009	0,074
nötr	kızgın, mutlu, üzgün, şaşkın	0,019	0,111
üzgün	kızgın, mutlu, nötr, şaşkın	0,000	0,111
şaşkın	kızgın, mutlu, nötr, üzgün	0,009	0,037

Önerilen yöntem karşılaştırma amaçlı tüm yüz görüntülerine de uygulanmıştır. Bu amaçla yüz görüntülerinde, 83 adet anahtar nokta Face++ SDK kullanılarak tespit edilmiştir. Bu noktaların koordinat bilgileri kullanılarak geometrik işlemler yardımı ile öznitelikler çıkarılmıştır. Elde edilen mesafe, eğim, oran ve açı öznitelik kümelerinden SFS algoritması ile gürbüz öznitelikler seçilmiştir. Seçilen gürbüz özniteliklerin SVM ile sınıflandırılmasıyla, önerilen yöntem ile tüm yüz görüntülerinde yüz ifadesi tespiti yapılmıştır. Tüm yüz görüntülerine ait seçilen gürbüz öznitelikler Tablo 4.6.’da verilmiştir. SFS ile seçilen gürbüz öznitelikler kaş ve göz bölgelerindeki çalışmalarda olduğu gibi yine açı öznitelik kümesinden elde edilmiştir.

Tablo 4.6. Tüm yüz özneliklerine SFS uygulanması ile elde edilen gürbüz öznelikler.

83 anahtar nokta	SFS ile elde edilen gürbüz öznelikler
	$\frac{m_{f,d} - m_{c,e}}{1 + m_{f,d} \cdot m_{c,e}}$
	$\frac{m_{a,k} - m_{h,i}}{1 + m_{a,k} \cdot m_{h,i}}$
	$\frac{m_{a,c} - m_{k,j}}{1 + m_{a,c} \cdot m_{k,j}}$
	$\frac{m_{a,g} - m_{b,d}}{1 + m_{a,g} \cdot m_{b,d}}$

Önerilen sistemin sonuçları literatürdeki tüm yüz ve parçalı yüz görüntüleri üzerine çalışmış diğer yöntemler ile Tablo 4.7’de karşılaştırılmıştır. Tablodaki sonuçlar dikkate alındığında; yüz görüntülerinde yalnızca kaş ve göz bölgelerine bakılarak yüksek başarımla duygu tespitinin yapılabilirliği görülmektedir. Yine tabloda görüleceği üzere önerilen sistem kısmi yüz görüntüleri kullanan literatürdeki diğer çalışmalardan daha yüksek bir başarımla göstermiştir. [123]’de önerilen yöntem %93,33 başarıma ağız, gözler ve kaşları inceleyerek ulaşmıştır. Bu çalışmada önerilen sistem sadece kaş ve gözlere bakarak %92,59 başarıma ulaşmıştır. RaFD veritabanı üzerinde tüm yüz ile çalışmış çalışmalara da yakın ve daha iyi sonuçlar üretmiştir.

#### 4.2.4. Önerilen sistemin genelleme yeteneği

Önerilen, kaş ve göz bölgeleri incelenerek yüz ifadesi tespiti yapan sistemin genelleme kabiliyetini ölçmek için RaFD üzerinde tespit edilen gürbüz öznelikler, diğer yüz veri tabanlarından da elde edilerek SVM üzerinde eğitim ve test işlemleri farklı veri tabanları üzerinde yapılmıştır. Bu veri tabanları RaFD, KDEF, JAFFE ve CK olup, sonuçlar Tablo 4.8.’de verilmiştir.

Tablo 4.7. Önerilen yöntemin literatürdeki diğer çalışmalarla karşılaştırılması: Önerilen sistem beş duygu üzerinde çalışmaktadır, karşılaştırılan diğer çalışmalarda daha fazla ifade çalışılmış olması durumunda ya da daha az duygu çalışılmış olması durumunda başarımlar çalışmalarda verilmiş olan karmaşıklık matrislerine göre yeniden hesaplanmıştır.

Yöntemler	Veritabanı	Yüz bileşenleri	Ortalama başarımlar (%)
HOG+NNE [124]	RaFD, TFEID, JAFFE	tüm yüz	94,83
LVPZM+FUZZY [125]	RaFD	tüm yüz	93,54
Surf boosting [126]	RaFD	tüm yüz	90,25
Viola&Jones+AAM+ANN [127]	RaFD	tüm yüz	91,46
Fuzzy classification [128]	JAFFE	ağız+kaşlar+gözler	83,75
PCA+LDA+LSH [129]	CK+	ağız+gözler	88,00
Haar like features+HMM [123]	MMI	ağız+kaşlar+gözler	93,33
Haar like features+HMM [123]	Self database	ağız+kaşlar+gözler	86,66
CHD+OD+WL [130]	CK	üst yüz bölgesi	75,78
Gabor+RFLD+KNN [131]	CK	üst yüz bölgesi	73,69
Önerilen yöntem	RaFD	tüm yüz kaşlar+gözler	94,07 92,59

Tablo 4.8. Farklı veri tabanlarında eğitim ve test sonuçlarının performans analizi.

Araştırmacılar	Eğitim	Test	Başarımlar (%)	En iyi sonuç (%)
Farajzadeh ve ark. [132]	CK	JAFFE	45,19	51,13
	CK	TFEID	51,13	
Gu ve ark. [131]	JAFFE	CK	54,05	55,87
	CK	JAFFE	55,87	

Tablo 4.8. (Devam)

Araştırmacılar	Eğitim	Test	Başarım (%)	En iyi sonuç (%)
Abd el Meguid ve Levine [133]	BU3DFE	KDEF	46,17	55,03
	BU3DFE	RAFD	55,03	
	BU3DFE	JAFFE	41,96	
	BU3DFE	PICS	35,81	
	BU3DFE	SFEW	20,57	
Da Silva ve Pedrini [134]	CK+	JAFFE	42,30	57,60
	CK+	MUG	47,80	
	CK+	BOSPHORUS	43,00	
	JAFFE	CK+	48,20	
	JAFFE	MUG	32,90	
	JAFFE	BOSPHORUS	30,00	
	MUG	CK+	45,60	
	MUG	JAFFE	32,40	
	MUG	BOSPHORUS	53,80	
	BOSPHORUS	CK+	57,60	
	BOSPHORUS	JAFFE	36,20	
Ali ve ark. [124]	RaFD	JAFFE	48,67	63,57
	RaFD	TFEID	63,57	
	JAFFE	RaFD	52,15	
	JAFFE	TFEID	52,76	
	TFEID	RaFD	55,85	
	TFEID	JAFFE	47,33	
Önerilen sistem	RaFD	KDEF	65,71	75,52
	RaFD	JAFFE	45,33	
	RaFD	CK	48,00	

Tablo 4.8. (Devam)

Araştırmacılar	Eğitim	Test	Başarım (%)	En iyi sonuç (%)
Önerilen sistem	KDEF	RaFD	75,52	
	JAFFE	RaFD	40,30	
	CK	RaFD	66,87	
	KDEF	KDEF	69,00	
	JAFFE	JAFFE	67,35	
	CK	CK	68,33	
	RaFD	RaFD	92,59	

Tablo 4.8.'de girilen diğer çalışmalara ait değerler [124]'de atıf olarak verilen çalışmadan alınmıştır. Tablo incelendiğinde, yöntemlerin genelleme performansları genel olarak oldukça düşüktür. [124]'de belirtilen sonuçlarda RaFD-JAFFE ve JAFFE-RAFD'ye ait başarımlar önerilen sistemden yüksek görünse de; [124]'de tüm yüz görüntüsünden elde edilen 100 adet özniteliğin, bu tez çalışmasında ise sadece 6 adet özniteliğin kullanıldığı göz önüne alınmalıdır.

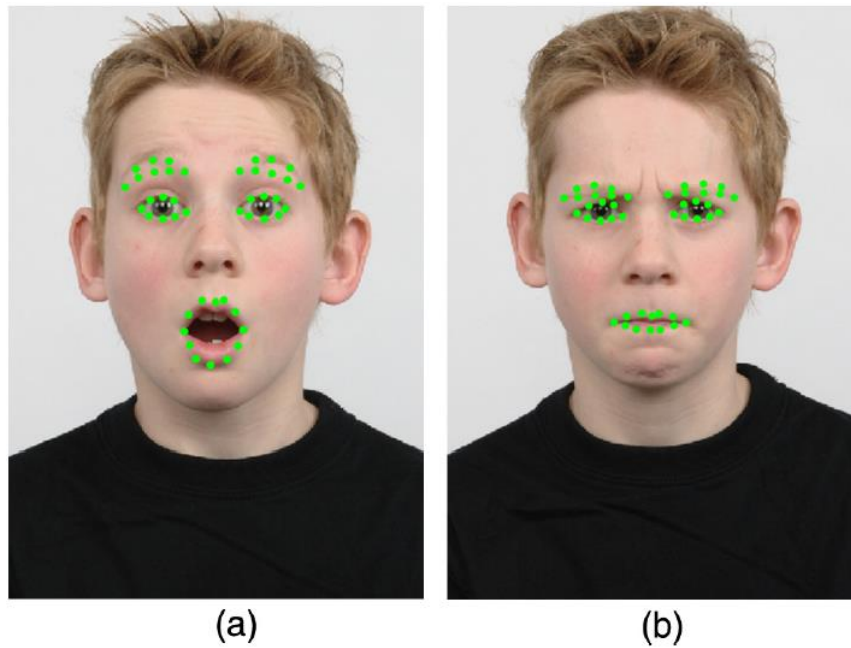
Tablo 4.9. Önerilen sistemin genelleme kapasitesi için kullanılan veri tabanları ve özellikleri.

Veri tabanı	Model sayısı	Görüntü sayısı	Model özellikleri
RaFD	67	335	Çocuk ve yetişkinlerin hem erkek hem kadın içerdiği topluluk
JAFFE	10	150	Kadın topluluk
KDEF	70	350	20-30 yaş aralığında 35 kadın, 35 erkek model
CK	100	500	18-50 yaş aralığında 350 kadın, 150 erkek görüntüsü

#### 4.2.5. Hata durumlarının incelenmesi

Yüz ifadesi tespiti karmaşıklık matrisi analiz edildiğinde hatalı sonuç üretmiş olan görüntülerde öznitelik seçimi ya da sınıflandırma hatalarından ziyade anahtar

noktaların tespitinde problem yaşanabileceği görülmüştür. Şekil 4.4. (a)'daki görüntüde bir çocuğa ait şaşkın ifadeye anahtar noktaların saptanmış hali gösterilmektedir. Bu durumda kaş bölgesindeki kontrastın düşük olması, anahtar nokta belirleme işleminde hataya neden olmuştur. Kaşlar olması gereken yerin daha altında işaretlendiğinden bu durumda da şaşkınlığın önemli bir ayırt edici özelliği ortadan kalkmıştır. Şekil 4.4. (b)'de ise aynı çocuğa ait kızgın görüntüde anahtar nokta tespiti başarılı olarak gerçekleşmiştir. Bu çalışmada kaş ve göz bölgeleri üzerinde çalışıldığından ağız üzerindeki noktalar incelenmemiştir.



Şekil 4.4. (a) Yüz anahtar noktaları hatalı işaretlenmiş bir durum, (b) aynı kişiye ait farklı duyguda anahtar noktaların doğru işaretlendiği bir durum [74].

### 4.3. Derin Öğrenme ile Tüm Yüz Görüntüleri Kullanılarak Yüz İfadesi Tespiti

Bu tez kapsamında yapılan diğer uygulamalarda; yüz görüntülerinde yüz ifadesi tespitine etki eden özniteliklerle ilgilenip, arka plan bilgilerini büyük oranda azaltmak için Viola&Jones [47] yöntemi kullanılmıştır. Arka plan bilgileri büyük oranda azaltıldıktan sonra, RaFD veri tabanından elde edilen yüz görüntüleri; oluşturulan tekli CNN ağlarda ve seri/paralel bağlanmış CNN ağlarda girdi olarak kullanılmıştır. CNN uygulamaları için MatConvNet [135] kütüphanesinden faydalanılmıştır. MatConvNet,



Matlab için geliştirilmiş, bilgisayar görmesi uygulamalarında kullanılan bir CNN kütüphanesidir.

Bu çalışmada derin öğrenme kullanılarak geliştirilen beş farklı algoritma ile yüz ifadesi tespiti yapılmıştır. Bu algoritmalar aşağıda özetlenmiştir.

- Orijinal yüz görüntüleri, tekli CNN ağında, ifade tespiti için sınıflandırılmıştır.
- Yüz görüntüleri kaş, göz ve ağız yüz bileşenleri olarak bir sınıf, geriye kalan pikseller ise arka plan olmak üzere ikinci bir sınıf olarak özel bir CNN ağında bölütlenmiştir. Daha sonra bölütlenmiş görüntüler seri bağlı tekli CNN’de sınıflandırılmıştır.
- Üç kanal RGB (Red, Green, Blue) yüz görüntülerine, eşleşen bölütlenmiş tek kanal görüntülerin, tek bir matriste birleştirilmesi ile elde edilen dört kanaldan oluşan veri matrisi, CNN ile sınıflandırılmıştır.
- Bölütleme işlemi kaş, göz, ağız ve arka plan olmak üzere dört sınıf olarak yapılmış ve oluşan tek kanal bölütlenmiş görüntü CNN ile sınıflandırılmıştır.
- Tek kanal dört sınıflı bölütleme sonucu ile üç kanal orijinal görüntünün birleşimi ile oluşan dört kanallı veri matrisi tasarlanan CNN ile sınıflandırılmıştır.

Literatürdeki yüz ifadesi tespiti çalışmaları genel olarak Ekman tarafından tanımlanan altı evrensel yüz ifadesi [136] üzerinden yapılmaktadır. Bu tez çalışmasında da kızgınlık, iğrenme, korku, mutluluk, üzümlük ve şaşkınlık olmak üzere bu altı yüz ifadesi kullanılmıştır. Uygulama Matlab ortamında gerçekleştirilmiş olup, RaFD yüz veri tabanındaki 1206 görüntü eğitim ve test amaçlı kullanılmıştır. Uygulamalar 2.6 GHz Intel Core i7 işlemcili, 32 GB RAM ve 6 GB ekran kartına sahip bir bilgisayar üzerinde yapılmıştır.

#### **4.3.1. Kaş, göz ve ağız bölgelerinin yüz görüntülerinde bölütlenmesi**

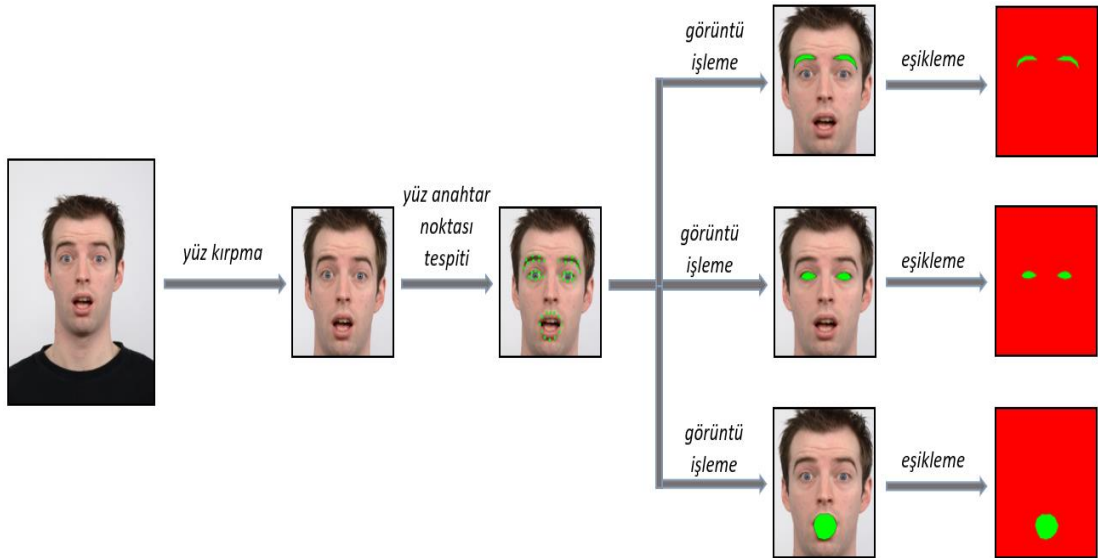
Yüz görüntülerinin Viola&Jones yöntemiyle elde edilmesinin ardından; yüzler, yüz ifadesi tespiti uygulamalarında kullanılmak üzere Bölüm 3’de verilmiş olan, bölütleme

amaçlı geliştirilen CNN ile bölütlenmiştir. Bölütleme, yukarıda bahsedildiği üzere iki sınıflı ve dört sınıflı olarak iki farklı şekilde gerçekleştirilmiştir. Sonuç olarak her iki yaklaşımda da siyah-beyaz ikili (binary) görüntüler elde edilmiştir.

Bölütlenmiş eğitim görüntülerini oluşturmak için önce Face++ SDK ile yüz anahtar noktaları elde edilmiştir. Bu anahtar noktalardan faydalanılarak; kaşlar, gözler ve ağız üzerinde bulunan yüz anahtar noktaları kendi içlerinde birleştirilerek bu bileşenlerin sınırları belirlenmiştir. Daha sonra bu kapalı şekillerin içindeki pikseller yeşil, geri kalan piksel değerleri ise kırmızı renk ile doldurulmuştur. Eğitim için RaFD yüz veri tabanında yer alan görüntülerin ilk yarısı kullanılmıştır. İki sınıflı bölütleme eğitimi için bölütlenmiş görüntü oluşturma aşaması Şekil 4.5.'de, dört sınıflı bölütleme eğitimi için bölütlenmiş görüntü oluşturma işlemi ise Şekil 4.6.'da verilmiştir.



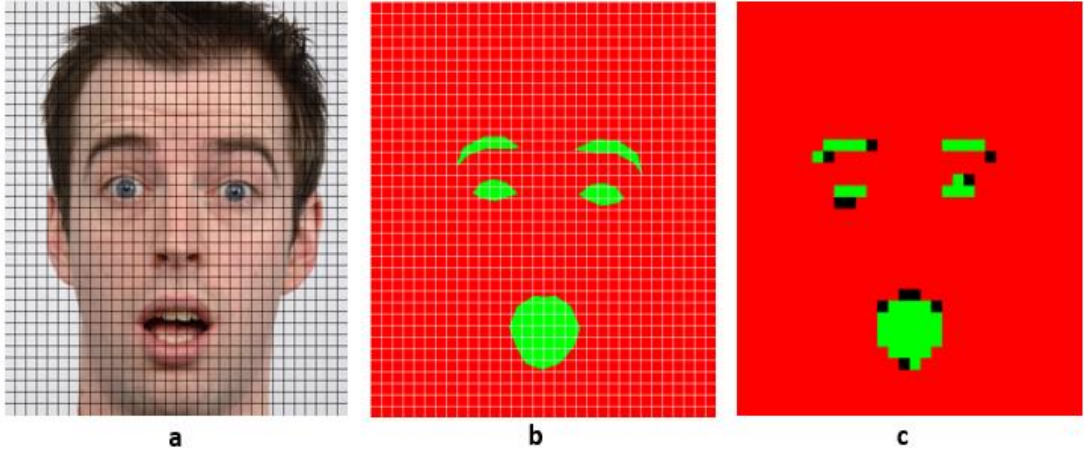
Şekil 4.5. İki sınıflı bölütleme işleminde eğitim işlemi için bölütlenmiş görüntü oluşturma aşamaları.



Şekil 4.6. Dört sınıflı bölütleme işleminde eğitim işlemi için bölütlenmiş görüntü oluşturma aşamaları.

Yüz ifadesi tespiti için önemli öznitelikler barındıran kaşlar, gözler ve ağız yüz bileşenlerine ait bölgeler, yukarıdaki yöntemle arka plan bölgelerinden ayrılmıştır. Yüz

bileşenleri segmentasyonu, tasarlanan CNN için bir ikili sınıflandırma problemidir. Bu sınıflandırma için eğitim CNN içerisinde 16x16 pikselik boklar ile yapılmaktadır. Veritabanında yer alan her orijinal RGB görüntü ile ona denk düşen yukarıda eğitim için oluşturulan maskeler 16x16'lık bloklara ayrılır. Maske üzerindeki her bloğun pixellerine bakılır. Eğer bir bloktaki maskenin %80'inden fazlası yüz bileşenlerinden birine aitse (yeşil renkli ise), orijinal görüntüde ona denk düşen bloğun sınıf etiketi 1 olur. Eğer blokta %80'den fazla piksel arkaplan verisi içeriyorsa (kırmızı), orijinal görüntüdeki bloğun etiketi 0 olur. %80'den fazla yüz bileşeni veya arkaplan görüntüsü içermeyen bloklar eğitime dahil edilmez. Bölütleme amaçlı oluşturulan CNN için eğitim bloklarının oluşturulması Şekil 4.7.'de verilmiştir.



Şekil 4.7. Bölütleme işlemi için eğitim bloklarını oluşturma aşamaları. a) orijinal RGB görüntü, b) eğitim için oluşturulan maske, c) eğitime girecek blokların etiketlenmesi.

Şekil 4.7. (a)'da 16x 16'lık bloklara ayrılan orijinal görüntü, (b)'de ise Face++ SDK ile oluşturulan maske verilmiştir. Yukarıda anlatılan %80 eşik değerine göre de sınıf etiketleri belirlenmiştir. (c)'deki görüntüde yeşil pikseller yüz bileşenleri sınıfını, kırmızı pikseller bu bileşenlerin dışında kalan alanları ve siyah pikseller de %80 değerine ulaşmamış ve eğitime katılmayacak olan pikselleri temsil etmektedir. Buna göre örneğin; (a) görüntüsünün sol en üstündeki blok eğitime girerken, bu bloğa denk gelen (c) görüntüsündeki blok kırmızı olduğundan, sınıf değeri olarak CNN'e 0 değeri verilir. Aynı şekilde (a) görüntüsünde dudakların içinde kalan herhangi bir blok için, (c) görüntüsünde ona karşılık gelen blok yeşil olduğundan, CNN'e 1 etiketi gönderilir.

Eđitim 16x16'lık bloklar ile yapılıp tamamlandığında test işlemeine geçilir. Test işleminde görüntüler Şekil 4.8'de de görüleceđi üzere bütün halde verilir. Test işlemleri Shelhamer ve ark. tarafından önerilen kayan pencere (sliding window) yöntemi [137] ile yapılmaktadır. Buna göre kayan çerçevelerin veri matrisi üzerinde taradığı bölgeler üst üste denk gelmeden network içinde analiz edilir. CNN havuzlama katmanlarından dolayı orijinal görüntüden boyut olarak daha küçük ama tam bir bölütlenmiş görüntü oluşur. Bölütleme amaçlı kullanılan CNN'in konvolüsyon katmanlarına ait çekirdek boyutları, filtre sayıları ve çıkış hacimleri Tablo 4.10.'da verilmiştir.

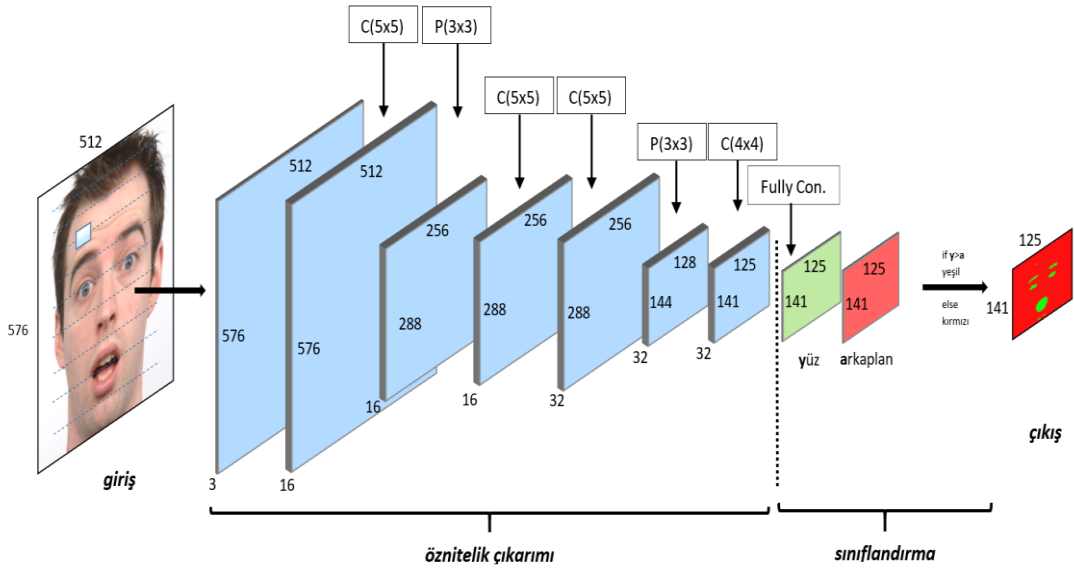
Tablo 4.10. Bölütleme amaçlı kullanılan CNN'in konvolüsyon katmanlarına ait bilgiler.

Katman	Çekirdek	Filtre sayısı	Çıkış
Konvolüsyon1	5x5	16	16x16x16
Konvolüsyon2	5x5	16	8x8x16
Konvolüsyon3	5x5	32	4x4x32
Konvolüsyon4	4x4	32	1x1x32

Eđitim aşamasında iki adet dropout kullanılmıştır. Dropout değeri her ikisi için de 0,5'tir. Ağ içerisinde dört adet ReLU katmanı bulunmaktadır.

#### 4.3.1.1. İki sınıflı bölütleme

Bu aşamada yüz görüntüleri iki sınıflı olarak bölütlenmiştir. Sınıflardan biri kaşlar, gözler ve ağız yüz bileşenlerini içerirken, geriye kalan görüntü bölgeleri de arka plan olarak diğer sınıfa dahil edilmiştir. CNN aldığı üç kanal orijinal görüntüyü işleyerek yine üç kanallı bölütlenmiş bir görüntü oluşturur. Şekil 4.8.'de bölütleme amaçlı oluşturulan CNN yapısı verilmiştir.



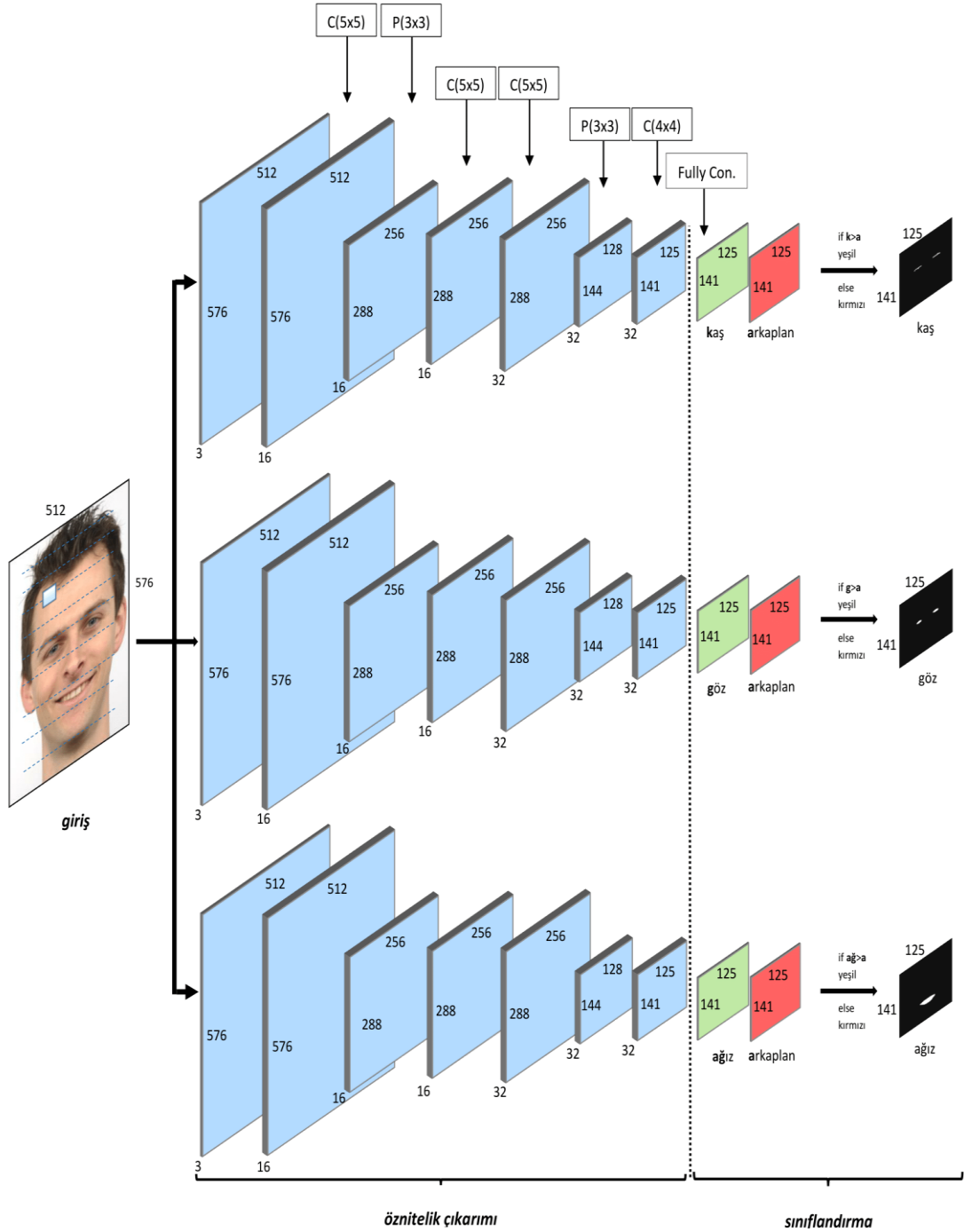
Şekil 4.8. İki sınıflı bölütleme için oluşturulan CNN yapısı.

#### 4.3.1.2. Dört sınıflı bölütleme

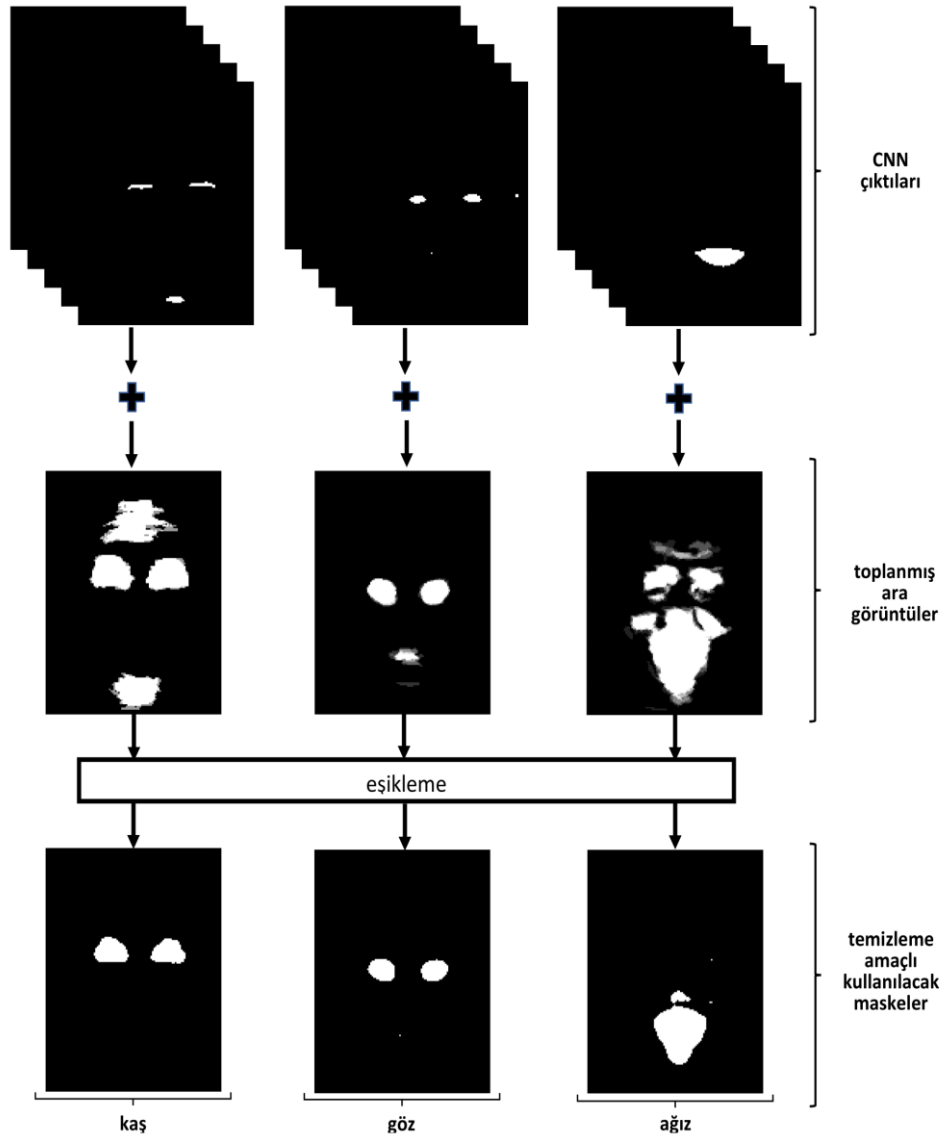
Dört sınıflı bölütlemeye yüz görüntülerinden kaşlar, gözler, ağız ve geriye kalan bölgeler olmak üzere dört farklı sınıf elde edilmektedir. Bu bölütleme işlemi için Şekil 4.9.'da verilen birbirine paralel bağlanmış 3 CNN yapısı kullanılmıştır. Her bir CNN, ilgili yüz bileşeni ve arkaplan görüntüsü olmak üzere ikili sınıflandırma yapmaktadır. İlk CNN kaşlar ve arkaplan, ikinci CNN gözler ve arkaplan, üçüncü CNN ağız ve arkaplan olmak üzere toplamda dört farklı sınıf üretmiş olur.

CNN bölütleme sonuçları incelendiğinde; çıkan bölütlenmiş görüntülerde odaklanılan bölgeler haricinde gürültüler olduğu tespit edilmiştir. Örneğin gözler ve arka planı bölütleyen CNN çıktılarında ağız bölgelerinde ağıza sınıfına ait olmayan piksel çıktılarının bu sınıfa ait olarak etiketlendiği gözlemlenmiştir. Bu gürültüleri giderebilmek için elde edilen görüntülerden bir maske oluşturup, tüm bölütleme sonuçlarını bu maskelerle işleme sokup, gürültüden arındırılmış bir bölütlenmiş görüntü elde edilmeye çalışılmıştır. Bu işlemi yapabilmek için her bir bileşen türünü içeren bölütlenmiş görüntüler kendi içlerinde toplama işlemine tabi tutulmuştur. Ortaya çıkan ara görüntülerdeki gürültülü piksel değerlerini gidermek için eşikleme noktası belirlenmiştir. Bu eşikleme noktası deneysel çalışmalar çerçevesinde 128 olarak belirlenmiştir. Her bir bileşen türünün toplamı ile elde edilen ara maskelerin her

piksel değeri 128'den küçükse 0, değilse 1 değerine sahip olacak şekilde sınımlanmıştır. Böylece temizleme işlemini yerine getirecek olan ara maskeler elde edilmiş olur. Maske elde etme işleminin adımları Şekil 4.10.'da verilmiştir.

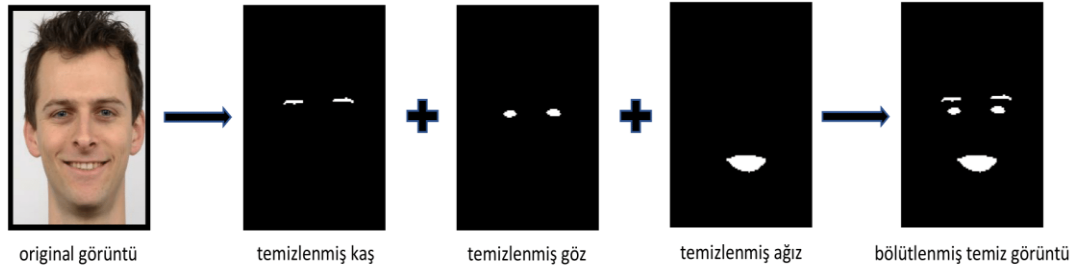


Şekil 4.9. Dört sınıflı bölütleme için oluşturulan CNN yapısı.



Şekil 4.10. Maske oluşturma aşamaları: Her bir bileşen türünü içeren bölütlenmiş görüntüler kendi içlerinde toplanır, ortaya çıkan ara maskelere bir eşikleme işlemi uygulanır ve gürültü giderme amaçlı maskeler oluşturulur. Ara maskeler üzerinde düşük piksel değerine sahip alanlar açık gri renkli bölgelerde koaylıkla gözlemlenebilmektedir.

Kaşlar, gözler ve ağız için elde edilen maskeler kendi türündeki sınıf görüntüleri ile mantıksal VE işlemine tabi tutulurlar. Bu şekilde son bölütlenmiş görüntüde bir piksel değerinin '1' olabilmesi için o pixelin hem ilgili bölütlenmiş görüntüde hem de ona ait maskede '1' olması gerekmektedir. Maskede '1' olmayan tüm değerler, sınıflandırma sonucuna olumlu yönde etki eden gürültü giderme işlemi gerçekleştirir. Son olarak, gürültüden arındırılmış kaş, göz ve ağız görüntüleri tek bir bölütlenmiş görüntü elde etmek için mantıksal VEYA işlemine tabi tutulurlar. Bu işlem Şekil 4.11.'de verilmiştir.



Şekil 4.11. Bölütlenmiş görüntünün son halinin oluşturulması.

### 4.3.2. Sınıflandırma çalışmaları

Duygu sınıflandırmada kullanılan beş farklı derin öğrenme uygulaması ve sonuçları aşağıdaki alt başlıklarda verilmiştir. İlk bölümde orijinal görüntülerin, ikinci ve üçüncü bölümde bölütlenmiş görüntülerin, dördüncü bölümde orijinal ve tek kanal bölütlenmiş görüntülerin birleşiminin ve son bölümde ise orijinal ve dört sınıflı tek kanal bölütlenmiş görüntülerin birleşiminin sınıflandırılması ve sonuçları verilmiştir.

Sınıflandırma çalışmalarında kullanılan ağda iki adet dropout ve 4 adet ReLU katmanı kullanılmıştır. Dropout değeri 0,5 olup, Tablo 4.11.'de konvolüsyon katmanlarına ait özellikler verilmiştir.

Tablo 4.11. Sınıflandırma amaçlı kullanılan CNN'in konvolüsyon katmanlarına ait bilgiler.

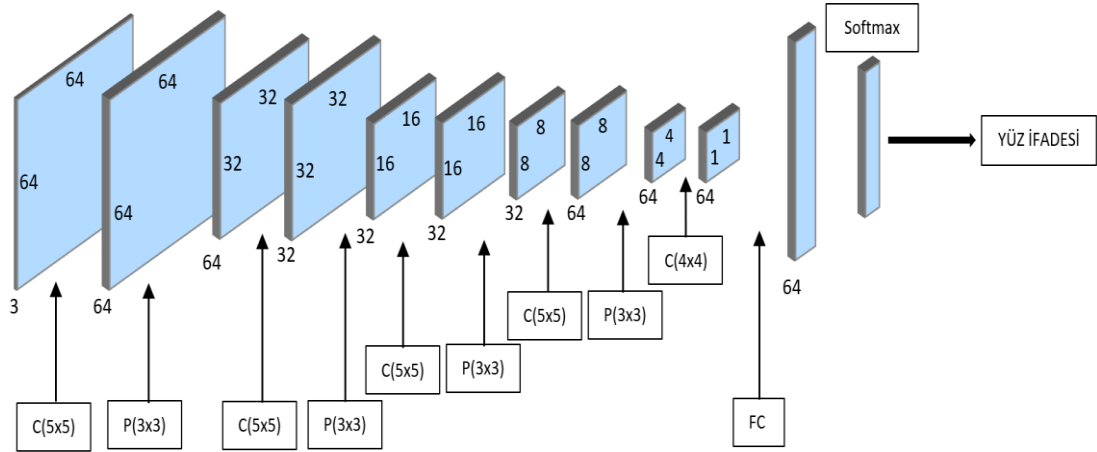
Katman	Çekirdek	Filtre sayısı	Çıkış
Konvolüsyon1	5x5	64	64x64x64
Konvolüsyon2	5x5	32	32x32x32
Konvolüsyon3	5x5	32	16x16x32
Konvolüsyon4	5x5	64	8x8x64
Konvolüsyon5	4x4	64	1x1x64

#### 4.3.2.1. Uygulama 1 - orijinal yüz görüntülerinin sınıflandırılması

Derin öğrenme çalışmalarında ilk olarak üç kanal RGB ham görüntüler, CNN ile sınıflandırılmıştır. CNN girdi olarak 64x64 boyutunda orijinal yüz görüntülerini alır ve 6 duygu sınıfından hangisine ait olduğu ile ilgili bir çıktı üretir. Bu amaç için



oluşturulan CNN yapısı Şekil 4.12.'de verilmiştir. Sınıflandırma sonucuna ait karmaşıklık matrisi ise Tablo 4.12.'de gösterilmiştir.



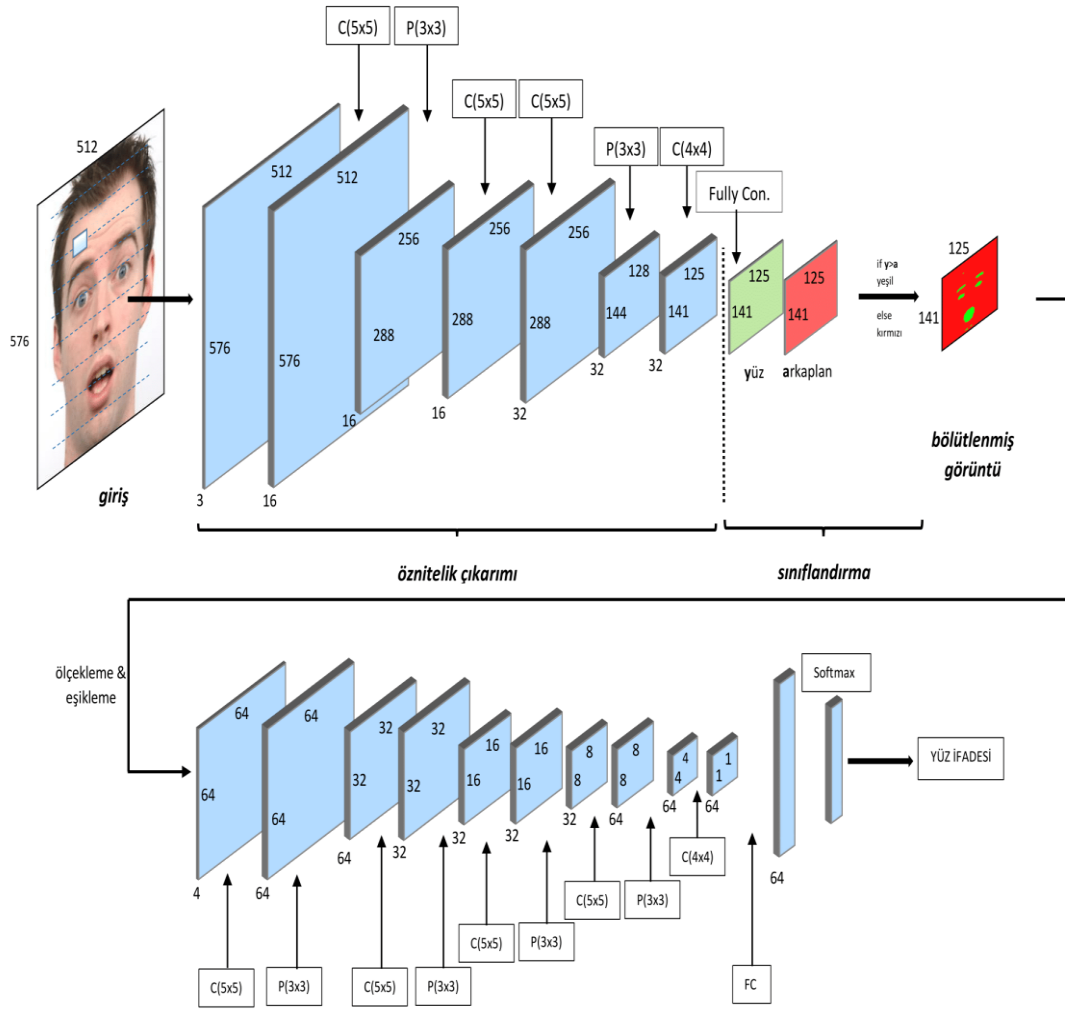
Şekil 4.12.Üç kanal orijinal görüntülerin sınıflandırılması için oluşturulan CNN yapısı.

Tablo 4.12.Orijinal yüz görüntülerinin sınıflandırılması ile oluşan karmaşıklık matrisi.

gerçek/tahmin	kızgınlık	iğrenme	korku	mutluluk	üzgünlük	şaşkınlık
kızgınlık	88,89	3,03	0	0	8,08	0
iğrenme	7,07	90,91	0	2,02	0	0
korku	1,01	2,02	84,85	1,01	8,08	3,03
mutluluk	9,09	3,03	14,14	73,74	0	0
üzgünlük	9,09	2,02	19,19	0	69,70	0
şaşkınlık	0	0	8,08	0	0	91,92
<b>ortalama: %83,33</b>						

#### 4.3.2.2. Uygulama 2 – iki sınıflı bölütleme ile elde edilen bölütlenmiş görüntülerin sınıflandırılması

Bu uygulamada ardışıl iki CNN kullanılmıştır. İlk CNN iki sınıflı bölütleme yaparak kaş, göz ve ağız bölgelerini bir sınıf, arkaplanı bir sınıf olarak belirlerken sonraki CNN ise yüz ifadesi sınıflandırma işlemi yapmaktadır. Bu işlemlerin akış diyagramı Şekil 4.13.'de verilmiştir. Bu ağ yapısı ile elde edilen bölütlenmiş görüntülerin sınıflandırılmasıyla elde edilen karmaşıklık matrisi Tablo 4.13.'de verilmiştir.



Şekil 4.13. İki sınıflı bölütlenmiş görüntünün sınıflandırılması.

Tablo 4.13. İki sınıflı bölütlenmiş görüntülerin sınıflandırılması ile oluşan karmaşıklık matrisi.

gerçek/tahmin	kızgınlık	iğrenme	korku	mutluluk	üzgünlük	şaşkınlık
kızgınlık	98,67	0	0	1,3	0	0
iğrenme	0	89,34	0	9,33	0	1,33
korku	4	2,67	76	0	8	9,33
mutluluk	0	1,33	1,33	97,34	0	0
üzgünlük	13,33	5,33	5,34	0	76	0
şaşkınlık	0	2,67	0	0	0	97,33
<b>ortalama: %89,11</b>						

### 4.3.2.3. Uygulama 3 – dört sınıflı bölütleme ile elde edilen bölütlenmiş görüntülerin sınıflandırılması

Dört sınıflı bölütleme için özelleştirilmiş üç CNN yapısı birbirine paralel bağlanmıştır. Bu yapıya bağlı dördüncü bir CNN ile de sınıflandırma işlemi yapılmıştır. Bu işlemlerin akış diyagramı Şekil 4.14.'de verilmiştir. Bu ağ yapısı ile elde edilen bölütlenmiş görüntülerin sınıflandırılmasıyla elde edilen karmaşıklık matrisi Tablo 4.14.'de verilmiştir.



Şekil 4.14. Dört sınıflı bölütlenmiş görüntünün sınıflandırılması.

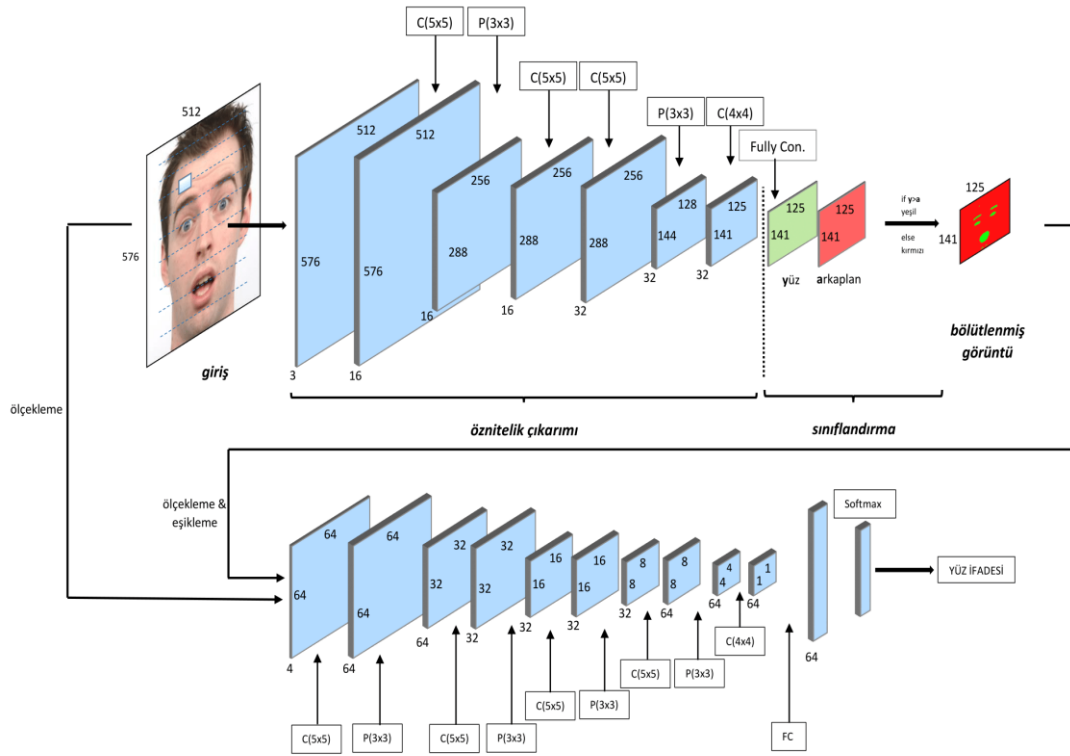
Tablo 4.14.Dört sınıflı bölütlenmiş görüntülerin sınıflandırılması ile oluşan karmaşıklık matrisi.

gerçek/tahmin	kızgınlık	iğrenme	korku	mutluluk	üzgünlük	şaşkınlık
kızgınlık	87,88	0	0	1,01	11,11	0
iğrenme	1,01	97,98	0	1,01	0	0
korku	0	3,03	77,78	0	14,14	5,05
mutluluk	3,03	1,01	2,02	91,92	2,02	0
üzgünlük	5,05	0	1,01	0	93,94	0
şaşkınlık	0	8,08	1,01	0	0	90,91
<b>ortalama:90,07</b>						

#### 4.3.2.4. Uygulama 4 - orijinal görüntüler ile iki sınıflı bölütlenmiş görüntülerin birleşiminin sınıflandırılması

Bu uygulamada iki sınıflı bölütleme için özelleştirilmiş tek bir CNN yapısı ve buna seri ve paralel bağlı ikinci bir CNN yapısı ile sınıflandırma işlemi yapılmıştır. Bu yapıda orijinal ham görüntü bölütleme için özelleştirilmiş CNN için girdi olarak kullanılmıştır. Bölütlenmiş görüntü ile orijinal ham görüntü 4 kanallı bir matris birleştirilerek elde edilen veri matrisi sınıflandırma işleminin yapılacağı CNN yapısında giriş olarak kullanılmıştır. Bu işlemlerin akış diyagramı Şekil 4.15.'de verilmiştir. Bu ağ yapısı ile elde edilen girdilerin sınıflandırılmasıyla ortaya çıkan karmaşıklık matrisi Tablo 4.15.'de verilmiştir.

Bölütleme amaçlı oluşturulan CNN yapısında orijinal görüntüler konvolüsyon ve havuzlama katmanlarından geçtikten sonra elde edilen olasılıksal iki kanallı değerler sinama işlemlerine tabi tutulup her bir piksel için bir sınıf belirlenir. İki kanaldan biri yüz bileşenleri için, diğeri de arkaplan için olasılıksal veri içerir. Karşılıklı piksellerde hangi değer daha büyükse o pikselin bulunduğu kanala ait sınıf değeri çıkışa yansıtılır. Üç kanallı bölütlenmiş görüntü siyah beyaza dönüştürülüp 64x64 boyutuna ölçeklenir. Orijinal görüntü ile birleştirilen bölütlenmiş tek kanal görüntü ikinci CNN'de sınıflandırılır.



Şekil 4.15. İki sınıflı bölütlenmiş görüntü ile orijinal ham görüntünün birleşiminin sınıflandırılması.

Tablo 4.15. İki sınıflı bölütlenmiş görüntü ile orijinal ham görüntünün birleşiminin sınıflandırılması ile oluşan karmaşıklık matrisi.

gerçek/tahmin	kızgınlık	iğrenme	korku	mutluluk	üzgünlük	şaşkınlık
kızgınlık	92,93	0	0	0	7,07	0
iğrenme	1,01	98,99	0	0	0	0
korku	0	0	87,88	0	10,10	2,2
mutluluk	1,01	4,04	1,01	92,93	1,01	0
üzgünlük	8,08	0	3,03	0	88,89	0
şaşkınlık	0	0	1,01	0	0	98,99
<b>ortalama: %93,43</b>						

#### 4.3.2.5. Uygulama 5 – orijinal görüntüler ile dört sınıflı bölütlenmiş görüntülerin birleşiminin sınıflandırılması

Dört sınıflı bölütleme için özelleştirilmiş üç CNN yapısı birbirine paralel bağlanmıştır. Bu yapıya bağlı dördüncü bir CNN ile de sınıflandırma işlemi gerçekleştirilmiştir. Bu



Tablo 4.16. Dört sınıflı bölütlenmiş görüntü ile orijinal ham görüntünün birleşiminin sınıflandırılması ile oluşan karmaşıklık matrisi.

gerçek/tahmin	kızgınlık	iğrenme	koru	mutluluk	üzgünlük	şaşkınlık
kızgınlık	90,91	3,03	0	0	6,06	0
iğrenme	0	98,99	1,01	0	0	0
koru	0	1,01	89,90	0	7,07	2,02
mutluluk	0	2,02	2,02	95,96	0	0
üzgünlük	3,03	4,04	0	0	92,93	0
şaşkınlık	0	2,02	1,01	0	1,01	95,96
<b>Ortalama: %94,11</b>						

#### 4.3.3. Yapılan sınıflandırmaların literatür ile karşılaştırılması

Derin öğrenme yöntemlerinden olan CNN'in kullanıldığı bu çalışmalarda yüz ifadesi tespiti yapılmıştır. Yapılan uygulamalarda yalnızca orijinal görüntüler sınıflandırıldığı gibi; iki farklı şekilde bölütlenmiş görüntüler ve orijinal görüntüler ile bölütlenmiş görüntülerin birleştirilmesi ile oluşan veri matrisleri de sınıflandırılmıştır. Eğitim ve test işlemleri için RaFD yüz veritabanı kullanılmıştır. Literatürde RaFD kullanan yüz ifadesi tespiti çalışmaları ile bu tez kapsamında yapılan uygulamaların sonuçları Tablo 4.17.'de verilmiştir.

Tablo 4.17.'den de görüleceği üzere önerilen yöntemlerden en başarılı sonuç orijinal ham görüntü ile dört sınıflı bölütlenmiş görüntünün birleştirilmesiyle oluşan veri matrisinden elde edilmiştir. Tek olarak orijinal ham görüntünün sınıflandırılmasına göre yaklaşık %10'luk bir artış sağlanmıştır. Yüz ifadesi tespitinde kaş, göz ve ağız bölgeleri önemli bir rol oynamaktadır. Kaş, göz ve ağız bilgilerinin yer aldığı bölütlenmiş görüntü ve orijinal ham görüntünün birleşimi sınıflandırma amaçlı oluşturulan CNN'e girdi olarak verilmiş; böylece ağız, önemli yüz bileşenlerine odaklanması sağlanarak başarımlar artırılmıştır. Ayrıca, yalnızca bölütlenmiş görüntülerin kullanılmasıyla da 90,07'lik bir başarımlar elde edilmiştir. Yüz ifadelerinin önemli olduğu uygulamalarda, bu ifadelerin kişiden bağımsız saklanması gerekirse bölütlenmiş görüntüler kişisel gizliliği koruduğundan sorunsuzca saklanabilir.

Tablo 4.17.Önerilen yöntemlerin literatürdeki çalışmalar ile karşılaştırılması. (ÖY: önerilen yöntem).

Çalışmalar	Veri tabanı	Kullanılan ifade sayısı	Başarım (%)
HOG+NNE [124]	RaFD, TFEID, JAFFE	5	93,75
Facial component detection + KNN [138]	RaFD	6	75,61
Gabor filter + GLCM [139]	RaFD	7	88,41
Viola&Jones + AAM + ANN [127]	RaFD	7	89,55
Surf boosting [126]	RaFD	6	90,64
Facial component detection + Fuzzy [138]	RaFD	6	93,96
ÖY: orijinal ham görüntü			83,33
ÖY: iki sınıflı bölütlenmiş görüntü			89,11
ÖY: dört sınıflı bölütlenmiş görüntü			90,07
ÖY: orijinal ham görüntü + iki sınıflı bölütlenmiş görüntü	RaFD	6	93,43
ÖY: orijinal ham görüntü + dört sınıflı bölütlenmiş görüntü			94,11



## BÖLÜM 5. SONUÇ VE ÖNERİLER

Bu tez çalışmasında kısmi ve tam yüz görüntüleri kullanılarak makine öğrenmesi yöntemleri ile yüz ifadesi tespiti gerçekleştirilmiştir. Gerek kısmi yüz görüntüleri gerekse de tam yüz görüntüleri kullanılarak yapılan çalışmalarda farklı modeller ve makine öğrenmesi yöntemleri kullanılmıştır. Geliştirilen ilk uygulamada; öznelik çıkarımı, öznelik seçimi ve sınıflandırma aşamaları olmak üzere klasik makine öğrenmesi metotları kullanılmıştır. Diğer uygulamalarda ise bu işlemleri tek bir yapı içerisinde gerçekleştirebilen CNN yapıları kullanılarak yüz ifadesi tespiti yapılmıştır.

Geliştirilen ilk uygulamada RaFD yüz veritabanında bulunan yüz görüntülerinde yalnızca göz bölgesine odaklanılacağı için Face++ SDK tarafından gözler ve kaşları işaretleyecek şekilde 34 adet anahtar nokta tespit edilmiştir. Her bir görüntü için elde edilen bu 34 adet anahtar noktanın koordinat bilgileri kullanılarak mesafe, eğim, oran, açı ve karma öznelik kümeleri elde edilmiştir. Bu öznelik kümelerindeki yüzlerce öznelik ile çalışmak, işlem yükünden dolayı çalışma zamanını arttıracığından ve sonuca olumsuz etki eden özneliklerin başarımı düşüreceğinden dolayı SFS, SFFS ve ReliFF öznelik seçim algoritmaları ile gürbüz öznelikler seçilmiştir. Bu seçim sonucunda en yüksek başarımı sağlayan altı adet açı özneliği elde edilmiştir. En yüksek başarımı veren bu gürbüz öznelikler SVM ve KNN sınıflandırma algoritmaları ile sınıflandırılarak yüz ifadesi tespiti sağlanmıştır. Sadece kaşlar ve göz bölgeleri kullanılarak yapılan bu çalışmalarda %92,59 başarı oranı elde edilmiştir.

Önerilen bu yöntem ile yüzün yaklaşık sadece üçte birlik bir bölümüne bakılarak yüksek başarımla yüz ifadesi tespiti işlemi yapılmıştır. Ayrıca bu yaklaşım, tüm yüz görüntüleri üzerinde çalışan yaklaşımlarda oluşabilecek ağız bölgesinin atkısı vb. nedenlerle kapanması ya da konuşma esnasında ağız hareketlerinin olması gibi duygu tanımayı zorlaştıran durumlardan etkilenmeyecektir. Bununla birlikte öznelik seçme yöntemi ile seçilen az sayıdaki öznelik gerçek zamanlı çalışmaya olumlu yönde etki

ederken; önerilen yöntemin zayıf bellek ve işlemci gücüne sahip cihazlarda kullanım şansını da arttırmaktadır. Öznitelik seçim işlemi ile sonuca olumsuz etki eden öznitelikler elenerek başarımlar arttırılmıştır. Ayrıca sistemin genelleme yeteneğinin tespiti için RaFD, KDEF, JAFFE ve CK veritabanları ile çapraz eğitim ve test işlemleri gerçekleştirilmiş, literatüre göre yüksek bir başarımlar oranı elde edilmiştir.

Tez kapsamında geliştirilen derin öğrenme kullanılan uygulamalarda, RaFD yüz veritabanında bulunan yüz görüntüleri kullanılmıştır. İlk olarak, Viola&Jones algoritması ile yüz bölgeleri veritabanındaki görüntülerden kırılarak yalnızca yüz özniteliklerine odaklanılmaya çalışılmıştır. Elde edilen yüz verileri ile CNN üzerinde bölütleme ve sınıflandırma işlemleri yapılmıştır.

Yüz bölütleme için iki farklı yaklaşım gerçekleştirilmiştir. Bu yaklaşımların ilkinde bölütlenmek istenen kaşlar, gözler ve ağız bir sınıf, geri kalan yüz bölgeleri ikinci bir sınıf olarak değerlendirilmiş, yani iki sınıflı bölütleme gerçekleştirilmiştir. İkinci yüz bölütleme yaklaşımında ise kaşlar bir sınıf, gözler ikinci sınıf, ağız üçüncü sınıf ve geri kalan yüz bölgeleri dördüncü sınıf olarak ele alınmış; dolayısıyla dört sınıflı yüz bölütleme yapılmıştır. İkinci yaklaşımda her bir kişiye ait; ağız, kaşlar ve gözler olmak üzere 3 adet bölütlenmiş görüntü elde edilmiştir. Bu bölütlenmiş görüntüler, görüntü işleme yöntemleri ile temizlenip gürültüden arındırılmış, ardından birleştirilerek her bir kişiye ait tek bir bölütlenmiş görüntü elde edilmiştir.

Sınıflandırma işlemlerinde beş farklı CNN mimarisi kullanılmıştır. İlk mimaride tekli bir CNN ile orijinal RGB görüntüler sınıflandırılmıştır. İkinci mimaride ise kaşlar, gözler ve ağız bir sınıfı, geriye kalan alanların ise başka bir sınıfı temsil ettiği bölütlenmiş görüntüler tekli bir CNN yapısına verilmiştir. Bu CNN'in, orijinal görüntüler için kullanılan CNN'den tek farkı girdi olarak tek kanallı görüntü almasıdır. Orijinal görüntüler için kullanılan CNN ise RGB görüntüyü giriş olarak aldığından, girişi 3 kanalıdır. Diğer bir mimaride ise bölütleme amaçlı yukarıda bahsedilen üçlü paralel CNN yapısı ve ön işlem yöntemleri kullanılarak elde edilen tek kanallı bölütlenmiş görüntü, bu üçlü paralel CNN'e seri bağlı bir CNN ile sınıflandırılmıştır. Dördüncü mimaride ise orijinal görüntü ile iki sınıflı olarak bölütlenmiş görüntünün 4

kanal üzerinde birleştirilmesi ile elde edilen veri matrisi CNN ile sınıflandırılmıştır. Beşinci ve son mimaride ise 4 sınıflı olarak bölütlenmiş görüntülerin orijinal görüntülerle birleşimi ile oluşturulan dört kanal veri matrisi CNN ile sınıflandırılmıştır. Sınıflandırma sonuçları incelendiğinde, özellikle son mimarideki önerilen yöntemin, literatürde RaFD yüz veri tabanını kullanan çalışmalardan daha yüksek bir başarımla çalıştığı görülmüştür.

Tüm yüz görüntüleri kullanılarak önerilen sistemlerin CNN ile sınıflandırılmasının literatüre katkısı şu şekilde özetlenebilir. Bölütleme işlemi ile elde edilen, bölütlenmiş görüntüler kişisel bilgi içermediğinden, yüz ifadelerinin saklanması gerektiren yüz ifadesi tespiti çalışmalarında kişisel gizliliği koruyarak depolama amaçlı kullanılabilir. Bölütlenmiş görüntülerdeki kaşlar, gözler ve ağız yüz ifadesi için önemli bilgiler içerdiğinden, orijinal görüntüler ile birleştirilmesi; CNN'in kaş, göz ve ağız bölgelerine odaklanmasını sağlayarak başarımları arttırmıştır.

Bu tez kapsamında yapılan derin öğrenme tabanlı çalışmalarda, ham görüntüler ve bölütlenmiş görüntülerin birleştirilmesi ile oluşturulan veri yerine, orijinal görüntü ile bölütlenmiş görüntünün sınıfsal olasılık değerlerini gösteren iki kanallı veriler birleştirilerek, 5 kanallı veriler oluşturulup sınıflandırılabilir. Bu şekilde olasılıksal verilen sınıflandırma amaçlı oluşturulan CNN'deki etkileri incelenebilir. Ayrıca, önerilen yöntemler kullanılarak; düşük maliyetli, ilgi alanındaki kişiyi rahatsız etmeyen, nesnel ve nicel sonuçlar üreten sistemler geliştirilebilir. Bu sistemler ile medikal, pazarlama, sağlık gibi birçok alanda başarılı uygulamalar gerçekleştirilebilir.

## KAYNAKLAR

- [1] C. S. Reddy and T. Srinivas, "Improving the Classification Accuracy of Emotion Recognition using Facial Expressions," *Int. J. Appl. Eng. Res.*, vol. 11, no. 1, pp. 650–655, 2016.
- [2] A. Mehrabian, "Some referents and measures of nonverbal behavior," *Behav. Res. Methods Instrum.*, vol. 1, no. 6, pp. 203–207, Jan. 1968.
- [3] A. T. Lopes, E. de Aguiar, A. F. De Souza, and T. Oliveira-Santos, "Facial Expression Recognition with Convolutional Neural Networks: Coping with Few Data and the Training Sample Order," *Pattern Recognit.*, 2016.
- [4] L. Marco, M. Del Coco, P. Carcagni, and C. Distanto, "Automatic Emotion Recognition in Robot-Children Interaction for ASD Treatment," in *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, 2015, pp. 537–545.
- [5] L. Zhang, K. Mistry, M. Jiang, S. Chin Neoh, and M. A. Hossain, "Adaptive facial point detection and emotion recognition for a humanoid robot," *Comput. Vis. Image Underst.*, vol. 140, pp. 93–114, Nov. 2015.
- [6] Shuzhi Sam Ge, H. A. Samani, Yin Hao Janus Ong, and Chang Chieh Hang, "Active affective facial analysis for human-robot interaction," in *RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication*, 2008, pp. 83–88.
- [7] R. C. Luo, Y. C. Wu, and P. H. Lin, "Multimodal information fusion for human-robot interaction," in *2015 IEEE 10th Jubilee International Symposium on Applied Computational Intelligence and Informatics*, 2015, pp. 535–540.
- [8] D. Johnson, M. Malmir, D. Forster, M. Alac, and J. Movellan, "Design and early evaluation of the RUBI-5 sociable robots," in *2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL)*, 2012, pp. 1–2.
- [9] L. Lo Presti and M. La Cascia, "Using Hankel matrices for dynamics-based facial emotion recognition and pain detection," in *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2015, vol. 2015–Octob, pp. 26–33.

- [10] G. Yolcu *et al.*, “Deep learning-based facial expression recognition for monitoring neurological disorders,” in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017, pp. 1652–1657.
- [11] K. Osone, G. Weiyi, and T. Onisawa, “Human-agent communication in poker game,” in *2010 IEEE International Conference on Systems, Man and Cybernetics*, 2010, pp. 857–864.
- [12] I. Teyeb, O. Jemai, M. Zaied, and C. Ben Amar, “Vigilance measurement system through analysis of visual and emotional driver’s signs using wavelet networks,” in *2015 15th International Conference on Intelligent Systems Design and Applications (ISDA)*, 2015, pp. 140–147.
- [13] Y. Zhang and C. Hua, “Driver fatigue recognition based on facial expression analysis using local binary patterns,” *Opt. - Int. J. Light Electron Opt.*, vol. 126, no. 23, pp. 4501–4505, Dec. 2015.
- [14] T. De Pessemier, D. Verlee, and L. Martens, “Enhancing Recommender Systems for TV by Face Recognition,” in *Proceedings of the 12th International Conference on Web Information Systems and Technologies*, 2016, pp. 243–250.
- [15] K. H. An and M.-J. Chung, “Cognitive face analysis system for future interactive TV,” *Consum. Electron. IEEE Trans.*, vol. 55, no. 4, pp. 2271–2279, 2009.
- [16] N. M. Thalmann, P. Kalra, and M. Escher, “Face to virtual face,” *Proc. IEEE*, vol. 86, no. 5, pp. 870–883, May 1998.
- [17] M. Kakarla and G. R. Mohana Reddy, “A real time facial emotion recognition using depth sensor and interfacing with Second Life based Virtual 3D avatar,” *Recent Adv. Innov. Eng. (ICRAIE)*, 2014, pp. 1–7, 2014.
- [18] P. Fechteler, W. Paier, A. Hilsmann, and P. Eisert, “Real-time avatar animation with dynamic face texturing,” in *2016 IEEE International Conference on Image Processing (ICIP)*, 2016, pp. 355–359.
- [19] A. De, A. Saha, and M. C. Pal, “A Human Facial Expression Recognition Model Based on Eigen Face Approach,” *Procedia Comput. Sci.*, vol. 45, pp. 282–289, 2015.
- [20] E. Altintas and D. Camur, *Beden Dili : Sozsuz Iletisim*. Alfa Aktuel Yayinlari, 2005.
- [21] E. Kaşıkçı, *Doğrucu Beden Dili*. Hayat Yayıncılık, 2009.
- [22] C.-T. Hsu, S.-C. Hsu, and C.-L. Huang, “Facial expression recognition using Hough forest,” in *2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, 2013, pp. 1–9.

- [23] Lin Zhong, Qingshan Liu, Peng Yang, Bo Liu, Junzhou Huang, and D. N. Metaxas, "Learning active facial patches for expression analysis," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 2562–2569.
- [24] A. B. Oktay, "Machine Learning Based Object Localization And Contour Extraction In Medical Images With Prior Information," Gebze Technical University, 2011.
- [25] C. Hung, M. Pham, and S. Arasteh, "Image texture classification using texture spectrum and local binary pattern," *Geosci. Remote Sens. Symp. 2006. IGARSS 2006. IEEE Int. Conf. on. IEEE*, pp. 2739–2742, 2006.
- [26] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," 1998, pp. 484–498.
- [27] H. Mahersia and K. Hamrouni, "Using multiple steerable filters and Bayesian regularization for facial expression recognition," *Eng. Appl. Artif. Intell.*, vol. 38, pp. 190–202, Feb. 2015.
- [28] Y. Ouyang, N. Sang, and R. Huang, "Accurate and robust facial expressions recognition by fusing multiple sparse representation based classifiers," *Neurocomputing*, vol. 149, pp. 71–78, Feb. 2015.
- [29] E. Sonmez and S. Albayrak, "A facial component-based system for emotion classification," *TURKISH J. Electr. Eng. Comput. Sci.*, vol. 24, pp. 1663–1673, 2016.
- [30] C.-T. Liao, H.-J. Chuang, C.-H. Duan, and S.-H. Lai, "Learning spatial weighting for facial expression analysis via constrained quadratic programming," *Pattern Recognit.*, vol. 46, no. 11, pp. 3103–3116, Nov. 2013.
- [31] O. Sumer, "An Embedded Design And Implementation Of A Facial Expression Recognition System," İstanbul Technical University, 2014.
- [32] T. H. H. Zavaschi, A. L. Koerich, and L. E. S. Oliveira, "Facial expression recognition using ensemble of classifiers," in *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 1489–1492.
- [33] M. F. Valstar, B. Jiang, M. Mehu, M. Pantic, and K. Scherer, "The first facial expression recognition and analysis challenge," in *Face and Gesture 2011*, 2011, pp. 921–926.

- [34] M. S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, "Recognizing Facial Expression: Machine Learning and Application to Spontaneous Behavior," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, 2005, vol. 2, pp. 568–573.
- [35] A. T. Lopes, E. de Aguiar, A. F. De Souza, and T. Oliveira-Santos, "Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order," *Pattern Recognit.*, vol. 61, pp. 610–628, Jan. 2017.
- [36] D. A. Pitaloka, A. Wulandari, T. Basaruddin, and D. Y. Liliana, "Enhancing CNN with Preprocessing Stage in Automatic Emotion Recognition," *Procedia Comput. Sci.*, vol. 116, pp. 523–529, 2017.
- [37] M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda, "Subject independent facial expression recognition with robust face detection using a convolutional neural network," *Neural Networks*, vol. 16, no. 5–6, pp. 555–559, Jun. 2003.
- [38] J. Li *et al.*, "Facial Expression Recognition with Faster R-CNN," *Procedia Comput. Sci.*, vol. 107, pp. 135–140, 2017.
- [39] A. Fathallah, L. Abdi, and A. Douik, "Facial Expression Recognition via Deep Learning," in *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*, 2017, pp. 745–750.
- [40] A. Bilgic, O. C. Kurban, and T. Yildirim, "Face recognition classifier based on dimension reduction in deep learning properties," in *2017 25th Signal Processing and Communications Applications Conference (SIU)*, 2017, pp. 1–4.
- [41] O. Langner, R. Dotsch, G. Bijlstra, D. H. J. Wigboldus, S. T. Hawk, and A. van Knippenberg, "Presentation and validation of the Radboud Faces Database," *Cogn. Emot.*, vol. 24, no. 8, pp. 1377–1388, Dec. 2010.
- [42] P. Ekman and W. V. Friesen, "Facial action coding system: a technique for the measurement of facial movement," 1978.
- [43] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding facial expressions with Gabor wavelets," in *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 200–205.
- [44] D. Lundqvist, A. Flykt, and A. Öhman, "The Karolinska Directed Emotional Faces (KDEF). CD ROM from Department of Clinical Neuroscience. Psychology section, Karolinska Institutet; 1998."

- [45] T. Kanade, Y. Tian, and J. F. Cohn, "Comprehensive Database for Facial Expression Analysis," in *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000*, 2000, p. 46--.
- [46] Ming-Hsuan Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: a survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 1, pp. 34–58, 2002.
- [47] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, May 2004.
- [48] M. R. Mohammadi, E. Fatemizadeh, and M. H. Mahoor, "PCA-based dictionary building for accurate facial expression recognition via sparse representation," *J. Vis. Commun. Image Represent.*, vol. 25, no. 5, pp. 1082–1092, Jul. 2014.
- [49] C. Shan, S. Gong, and P. W. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study," *Image Vis. Comput.*, vol. 27, no. 6, pp. 803–816, May 2009.
- [50] Likun Lei and Zhenjiang Miao, "Individual facial expression transferring using Active Appearance Model and FAP in MPEG-4," in *2010 International Conference on Image Analysis and Signal Processing*, 2010, pp. 150–154.
- [51] C.-C. Hsieh and M.-K. Jiang, "A Facial Expression Classification System Based on Active Shape Model and Support Vector Machine," in *2011 International Symposium on Computer Science and Society*, 2011, pp. 311–314.
- [52] P. Kumar, S. L. Happy, and A. Routray, "A real-time robust facial expression recognition system using HOG features," in *2016 International Conference on Computing, Analytics and Security Trends (CAST)*, 2016, pp. 289–293.
- [53] "Face++ Cognitive Services." [Online]. <http://www.faceplusplus.com.>, Eriřim Tarihi: 08.01.2017.
- [54] A. W. Whitney, "A Direct Method of Nonparametric Measurement Selection," *IEEE Trans. Comput.*, vol. C-20, no. 9, pp. 1100–1103, Sep. 1971.
- [55] A. Smith, O. Mendoza-Schrock, S. Kangas, M. Dierking, and A. Shaw, "An end-to-end vehicle classification pipeline using vibrometry data," 2014, p. 907900.
- [56] R. Gutierrez-Osuna, "CSCE 666: Pattern Analysis," 2013. [Online]. [http://courses.cs.tamu.edu/rgutier/csce666\\_f13/](http://courses.cs.tamu.edu/rgutier/csce666_f13/).
- [57] P. Pudil, J. Novovičová, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognit. Lett.*, vol. 15, no. 11, pp. 1119–1125, Nov. 1994.



- [58] K. Kira and L. A. Rendell, "A Practical Approach to Feature Selection," in *Proceedings of the Ninth International Workshop on Machine Learning*, 1992, pp. 249–256.
- [59] H. BUDAK, "Özellik Seçim Yöntemleri ve Yeni Bir Yaklaşım," *Süleyman Demirel Üniversitesi Fen Bilim. Enstitüsü Derg.*, p. 10, May 2018.
- [60] I. Kotsia and I. Pitas, "Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 172–187, Jan. 2007.
- [61] N. Sebe, M. S. Lew, I. Cohen, A. Garg, and T. S. Huang, "Emotion recognition using a Cauchy Naive Bayes classifier," in *Object recognition supported by user interaction for service robots*, vol. 1, pp. 17–20.
- [62] Q. Wang, K. Jia, and P. Liu, "Design and Implementation of Remote Facial Expression Recognition Surveillance System Based on PCA and KNN Algorithms," in *2015 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)*, 2015, pp. 314–317.
- [63] F. Z. Salmam, A. Madani, and M. Kissi, "Facial Expression Recognition Using Decision Trees," in *2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV)*, 2016, pp. 125–130.
- [64] S. Borah and S. Konwar, "ANN based human facial expression recognition in color images," in *2014 International Conference on High Performance Computing and Applications (ICHPCA)*, 2014, pp. 1–6.
- [65] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, Jan. 1967.
- [66] P. Mulak and N. Talhar, "Analysis of Distance Measures Using K-Nearest Neighbor Algorithm on KDD Dataset," *Int. J. Sci. Res.*, vol. 4, no. 7, pp. 2101–2104, 2015.
- [67] V. N. Vapnik, "An overview of statistical learning theory," in *IEEE transactions on neural networks*, 1999, vol. 10, no. 5, pp. 988–999.
- [68] C. Cortes and V. Vapnik, "Support-Vector Networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
- [69] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification (2Nd Edition)*. Wiley-Interscience, 2000.
- [70] S. Patel, "Chapter 2 : SVM (Support Vector Machine) — Theory." [Online]. <https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72>., Erişim Tarihi: 30.08.2018.

- [71] Guodong Guo, S. Z. Li, and Kapluk Chan, "Face recognition by support vector machines," in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, pp. 196–201.
- [72] Jianhua Dong, Jason Gu, Xin Ma, and Yibin Li, "Face recognition based on KPCA and SVM," in *Proceeding of the 11th World Congress on Intelligent Control and Automation*, 2014, pp. 1439–1444.
- [73] Zhao Lihong, Song Ying, Zhu Yushi, Zhang Cheng, and Zheng Yi, "Face recognition based on multi-class SVM," in *2009 Chinese Control and Decision Conference*, 2009, pp. 5871–5873.
- [74] I. Oztel, G. Yolcu, C. Oz, S. Kazan, and F. Bunyak, "iFER: facial expression recognition using automatically selected geometric eye and eyebrow features," *J. Electron. Imaging*, vol. 27, no. 02, p. 1, Mar. 2018.
- [75] Y. Zilu and Z. Guoyi, "Facial Expression Recognition Based on NMF and SVM," in *2009 International Forum on Information Technology and Applications*, 2009, pp. 612–615.
- [76] O. F. Soylemez, B. Ergen, and N. H. Soylemez, "A 3D facial expression recognition system based on SVM classifier using distance based features," in *2017 25th Signal Processing and Communications Applications Conference (SIU)*, 2017, pp. 1–3.
- [77] S.-W. Lee and H.-C. Wu, "Effective multiple-features extraction for off-line SVM-based handwritten numeral recognition," in *2012 International Conference on Information Security and Intelligent Control*, 2012, pp. 194–197.
- [78] B. El qacimy, M. Ait kerroum, and A. Hammouch, "Handwritten digit recognition based on DCT features and SVM classifier," in *2014 Second World Conference on Complex Systems (WCCS)*, 2014, pp. 13–16.
- [79] G. Katiyar and S. Mehfuz, "SVM based off-line handwritten digit recognition," in *2015 Annual IEEE India Conference (INDICON)*, 2015, pp. 1–5.
- [80] E. I. Zacharaki *et al.*, "MRI-based classification of brain tumor type and grade using SVM-RFE," in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2009, pp. 1035–1038.
- [81] D. Fernandez-Duque and S. E. Black, "Impaired recognition of negative facial emotions in patients with frontotemporal dementia," *Neuropsychologia*, vol. 43, no. 11, pp. 1673–1687, Jan. 2005.
- [82] D. Lin, L. Sun, K.-A. Toh, J. B. Zhang, and Z. Lin, "Biomedical image classification based on a cascade of an SVM with a reject option and subspace analysis," *Comput. Biol. Med.*, vol. 96, pp. 128–140, May 2018.

- [83] M. Tan, “Support Vector Machine & Its Applications.” [Online]. [www.iro.umontreal.ca/~pift6080/H09/documents/papers/svm\\_tutorial.ppt.](http://www.iro.umontreal.ca/~pift6080/H09/documents/papers/svm_tutorial.ppt.), Erişim Tarihi: 30.08.2018.
- [84] B. Lee, J. Chun, and P. Park, “Classification of facial expression using SVM for emotion care service system,” in *9th ACIS Int. Conf. Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, 2008, pp. 8–12.
- [85] S. Ayhan and S. Erdoğan, “Destek Vektor Makineleriyle Sınıflandırma Problemlerinin Çözümü İçin Çekirdek Fonksiyonu Seçimi,” *Eskisehir Osmangazi Univ. İİBF J.*, vol. 9, no. 1, pp. 175–198, 2014.
- [86] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT Press, 2016.
- [87] C. Dong, C. C. Loy, K. He, and X. Tang, “Learning a Deep Convolutional Network for Image Super-Resolution,” 2014, pp. 184–199.
- [88] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
- [89] A. S. Jackson, M. Valstar, and G. Tzimiropoulos, “What’s the Difference Between Artificial Intelligence, Machine Learning, and Deep Learning?,” *arXiv*, 2016.
- [90] S. McClure, “Data Science Ontology.” [Online]. <https://www.thoughtworks.com/insights/blog/data-science-ontology.>, Erişim Tarihi: 02.09.2018.
- [91] D. Kriesel, *A Brief Introduction to Neural Networks*. 2007.
- [92] E. Öztemel, *Yapay sinir ağları*, 3. baskı. İstanbul: Papatya Yayıncılık, 2012.
- [93] P. U. Hatipoğlu, “Time Series Classification using Deep Learning,” Middle East Technical University, 2016.
- [94] D. H. Hubel and T. N. Wiesel, “Receptive fields and functional architecture of monkey striate cortex.,” *J. Physiol.*, vol. 195, no. 1, pp. 215–43, Mar. 1968.
- [95] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [96] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

- [97] I. Oztel, G. Yolcu, I. Ersoy, T. White, and F. Bunyak, "Mitochondria segmentation in electron microscopy volumes using deep convolutional neural network," in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017, pp. 1195–1200.
- [98] Z. Al-Milaji, I. Ersoy, A. Hafiane, K. Palaniappan, and F. Bunyak, "Integrating segmentation with deep learning for enhanced classification of epithelial and stromal tissues in H&E images," *Pattern Recognit. Lett.*, Sep. 2017.
- [99] A. Hamad, F. Bunyak, and I. Ersoy, "Nucleus Classification in Colon Cancer H&E Images using Deep Learning," *Microsc. Microanal.*, vol. 23, no. S1, pp. 1376–1377, Jul. 2017.
- [100] E. P. Ijjina and C. K. Mohan, "Facial Expression Recognition Using Kinect Depth Sensor and Convolutional Neural Networks," in *2014 13th International Conference on Machine Learning and Applications*, 2014, pp. 392–396.
- [101] X. Chen, X. Yang, M. Wang, and J. Zou, "Convolution neural network for automatic facial expression recognition," in *2017 International Conference on Applied System Innovation (ICASI)*, 2017, pp. 814–817.
- [102] S. Chen, C. Zhang, M. Dong, J. Le, and M. Rao, "Using Ranking-CNN for Age Estimation," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 742–751.
- [103] C. Shang and H. Ai, "Cluster convolutional neural networks for facial age estimation," in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 1817–1821.
- [104] F. Gurpinar, H. Kaya, H. Dibeklioglu, and A. A. Salah, "Kernel ELM and CNN Based Facial Age Estimation," in *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2016, pp. 785–791.
- [105] C. Lee, H. J. Kim, and K. W. Oh, "Comparison of faster R-CNN models for object detection," in *2016 16th International Conference on Control, Automation and Systems (ICCAS)*, 2016, pp. 107–110.
- [106] H. Yanagisawa, T. Yamashita, and H. Watanabe, "A study on object detection method from manga images using CNN," in *2018 International Workshop on Advanced Image Technology (IWAIT)*, 2018, pp. 1–4.
- [107] J. T. Lee and Y. Chung, "Deep Learning-Based Vehicle Classification Using an Ensemble of Local Expert and Global Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017, pp. 920–925.

- [108] G. Yolcu, I. Oztel, S. Kazan, C. Oz, and F. Bunyak, “Facial component segmentation using convolutional neural network,” *Online J. Sci. Technol.*, vol. 8, no. 2, pp. 84–88, 2018.
- [109] A. S. Jackson, M. Valstar, and G. Tzimiropoulos, “A CNN cascade for landmark guided semantic part segmentation,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9915 LNCS, pp. 143–155, 2016.
- [110] A. Raj, S. Gupta, and N. K. Verma, “Face detection and recognition based on skin segmentation and CNN,” in *2016 11th International Conference on Industrial and Information Systems (ICIIS)*, 2016, pp. 54–59.
- [111] K. Wang, Y. Dong, H. Bai, Y. Zhao, and K. Hu, “Use fast R-CNN and cascade structure for face detection,” in *2016 Visual Communications and Image Processing (VCIP)*, 2016, pp. 1–4.
- [112] X. Hou, J. Zhu, K. Sun, L. Shen, and G. Qiu, “Object specific deep feature and its application to face detection,” in *2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA)*, 2017, pp. 173–176.
- [113] E. Tatulli and T. Hueber, “Feature extraction using multimodal convolutional neural networks for visual speech recognition,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 2971–2975.
- [114] S. Park, Y. Jeong, and H. S. Kim, “Multiresolution CNN for reverberant speech recognition,” in *2017 20th Conference of the Oriental Chapter of the International Coordinating Committee on Speech Databases and Speech I/O Systems and Assessment (O-COCOSDA)*, 2017, pp. 1–4.
- [115] S. Park, Y. Jeong, M. S. Kim, and H. S. Kim, “Linear prediction-based dereverberation with very deep convolutional neural networks for reverberant speech recognition,” in *2018 International Conference on Electronics, Information, and Communication (ICEIC)*, 2018, pp. 1–2.
- [116] M. J. Darugar and L. C. Kiong, “Gender estimation based on supervised HOG, Action Units and unsupervised CNN feature extraction,” in *2017 Artificial Intelligence and Robotics (IRANOPEN)*, 2017, pp. 23–27.
- [117] J.-H. Lee, Y.-M. Chan, T.-Y. Chen, and C.-S. Chen, “Joint Estimation of Age and Gender from Unconstrained Face Images Using Lightweight Multi-Task CNN for Mobile Applications,” in *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, 2018, pp. 162–165.
- [118] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Adv. Neural Inf. Process. Syst.*, pp. 1–9, 2012.

- [119] C. Szegedy *et al.*, “Going deeper with convolutions,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.
- [120] CS231n Course Materials, “CS231n Convolutional Neural Networks for Visual Recognition.” [Online]. <https://skymind.ai/wiki/convolutional-network>., Eriřim Tarihi: 06.09.2018.
- [121] Mathworks, “Visualize Activations of a Convolutional Neural Network.” [Online]. <https://ww2.mathworks.cn/help/nnet/examples/visualize-activations-of-a-convolutional-neural-network.html>., Eriřim Tarihi: 06.09.2018.
- [122] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, 2014.
- [123] D. Boruah, K. K. Sarma, and A. K. Talukdar, “Different face regions detection based facial expression recognition,” in *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, 2015, pp. 459–464.
- [124] G. Ali, M. A. Iqbal, and T.-S. Choi, “Boosted NNE collections for multicultural facial expression recognition,” *Pattern Recognit.*, vol. 55, pp. 14–27, Jul. 2016.
- [125] M. Ahmady, R. Ghasemi, and H. Rashidy kanan, “Local weighted Pseudo Zernike Moments and fuzzy classification for facial expression recognition,” in *2013 13th Iranian Conference on Fuzzy Systems (IFSC)*, 2013, pp. 1–4.
- [126] Q. Rao, X. Qu, Q. Mao, and Y. Zhan, “Multi-pose facial expression recognition based on SURF boosting,” in *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2015, pp. 630–635.
- [127] G. Bijlstra and R. Dotsch, “FaceReader 4 emotion classification performance on images from the Radboud Faces Database,” 2011.
- [128] A. Nicolai and A. Choi, “Facial Emotion Recognition Using Fuzzy Systems,” in *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 2015, pp. 2216–2221.
- [129] M. Ghayoumi and A. K. Bansal, “An Integrated Approach for Efficient Analysis of Facial Expressions,” in *Proceedings of the 11th International Conference on Signal Processing and Multimedia Applications*, 2014, pp. 211–219.
- [130] X. Huang, G. Zhao, W. Zheng, and M. Pietik?inen, “Towards a dynamic expression recognition system under facial occlusion,” *Pattern Recognit. Lett.*, vol. 33, no. 16, pp. 2181–2191, Dec. 2012.

- [131] W. Gu, C. Xiang, Y. V. Venkatesh, D. Huang, and H. Lin, "Facial expression recognition using radial encoding of local Gabor features and classifier synthesis," *Pattern Recognit.*, vol. 45, no. 1, pp. 80–91, Jan. 2012.
- [132] N. Farajzadeh, G. Pan, and Z. Wu, "Facial expression recognition based on meta probability codes," *Pattern Anal. Appl.*, vol. 17, no. 4, pp. 763–781, Nov. 2014.
- [133] M. K. Abd El Meguid and M. D. Levine, "Fully automated recognition of spontaneous facial expressions in videos using random forest classifiers," *IEEE Trans. Affect. Comput.*, vol. 5, no. 2, pp. 141–154, Apr. 2014.
- [134] F. A. M. da Silva and H. Pedrini, "Effects of cultural characteristics on building an emotion classifier through facial expression analysis," *J. Electron. Imaging*, vol. 24, no. 2, p. 023015, Mar. 2015.
- [135] A. Vedaldi and K. Lenc, "MatConvNet," in *Proceedings of the 23rd ACM international conference on Multimedia - MM '15*, 2015, pp. 689–692.
- [136] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion.," *J. Pers. Soc. Psychol.*, vol. 17, no. 2, pp. 124–129, 1971.
- [137] E. Shelhamer, J. Long, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," May 2016.
- [138] M. Ilbeygi and H. Shah-Hosseini, "A novel fuzzy facial expression recognition system based on facial feature extraction from color face images," *Eng. Appl. Artif. Intell.*, vol. 25, no. 1, pp. 130–146, Feb. 2012.
- [139] R. Li, P. Liu, K. Jia, and Q. Wu, "Facial Expression Recognition under Partial Occlusion Based on Gabor Filter and Gray-Level Cooccurrence Matrix," in *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, 2015, pp. 347–351.

## ÖZGEÇMİŞ

İsmail Öztel, 1988'de Trabzon'da doğdu. İlk, orta ve lise eğitimini Samsun'da tamamladı. 2007 yılında başladığı Sakarya Üniversitesi Bilgisayar Mühendisliği Bölümü'nü 2011 yılında bitirdi. 2012 yılında Sakarya Üniversitesi Bilgisayar Mühendisliği Bölümü'nde yüksek lisans eğitimine başladı. Aynı yılda Sakarya Üniversitesi Bilgisayar Mühendisliği Bölümü'nde araştırma görevlisi olarak çalışmaya başladı. 2014 yılında Sakarya Üniversitesi'nde Bilgisayar ve Bilişim Bilimleri Anabilim Dalı'nda doktora eğitime başladı. 2017-2018 yılları arasında bir yıl süreyle University of Missouri, Columbia, MO, USA'da misafir öğretim üyesi olarak bulundu. Halen Sakarya Üniversitesi Bilgisayar Mühendisliği Bölümü'nde Araştırma Görevlisi olarak görev yapmaktadır.