

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**YÜZ ANALİZİNE DAYALI DERİN ÖĞRENME
TABANLI BİR İLGİ TESPİT SİSTEMİNİN
GERÇEKLEŞTİRİLMESİ**

DOKTORA TEZİ

Gözde YOLCU ÖZTEL

**Enstitü Anabilim Dalı : BİLGİSAYAR VE BİLİŞİM
MÜHENDİSLİĞİ**
Tez Danışmanı : Dr. Öğr. Üyesi Serap KAZAN

Mayıs 2019

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**YÜZ ANALİZİNE DAYALI DERİN ÖĞRENME
TABANLI BİR İLGİ TESPİT SİSTEMİNİN
GERÇEKLEŞTİRİLMESİ**

DOKTORA TEZİ

Gözde YOLCU ÖZTEL

Enstitü Anabilim Dalı : BİLGİSAYAR VE BİLİŞİM
MÜHENDİSLİĞİ

Bu tez 17/05/2019 tarihinde aşağıdaki jüri tarafından oybirliği/oyçokluğu ile kabul edilmiştir.

Prof. Dr.
Abdullah FERİKOĞLU
Jüri Başkanı

Prof. Dr.
Cemil ÖZ
Üye

Prof. Dr.
Harun Reşit YAZGAN
Üye

Prof. Dr.
Raşit KÖKER
Üye

Dr. Öğr. Üyesi
Serap KAZAN
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.



Gözde YOLCU ÖZTEL

17.05.2019

TEŞEKKÜR

Hayatımın her anında kendilerinden aldığım sonsuz destek ile daha güçlü hissettiğim değerli ailem Cemal, Hacer, Caner YOLCU'ya, hem bu tez çalışmasındaki destekleri ile hem de hayattaki her anımda yanımda olan eşim İsmail ÖZTEL'e en içten teşekkürlerimi ve minnetlerimi sunarım.

Yüksek Lisans ve Doktora eğitimim boyunca birlikte çalıştığımız, destekleri ile yanımda olan danışman hocam Dr. Öğr. Üyesi Serap KAZAN'a, bu çalışmanın 1 yıllık kısmının yurt dışında yapılmasına vesile olan ve tez boyunca fikirlerini paylaşan hocam Prof. Dr. Cemil ÖZ'e, değerli yönlendirmeleri ve deneyimleri ile teze yön vermemi sağlayan hocalarım Dr. Filiz BUNYAK ve Dr. İlker ERSOY'a, olumlu duruşu ve destekleri ile yanımızda olan hocam Prof. Dr. Nejat YUMUŞAK'a, ihtiyacımız olduğu her an yardımlarını esirgemeyen hocam Dr. Öğr. Üyesi Veysel Harun ŞAHİN'e teşekkür ederim.

Bu çalışma Sakarya Üniversitesi Bilimsel Araştırma Projeleri Komisyonu tarafından desteklenmiştir. Proje Numarası: 2015-50-02-039.

İÇİNDEKİLER

TEŞEKKÜR.....	i
İÇİNDEKİLER.....	ii
SİMGELER VE KISALTMALAR LİSTESİ.....	v
ŞEKİLLER LİSTESİ.....	vii
TABLolar LİSTESİ.....	ix
ÖZET.....	x
SUMMARY.....	xi
BÖLÜM 1.	
GİRİŞ.....	1
BÖLÜM 2.	
KAYNAK ARAŞTIRMASI.....	8
2.1. Baş Yönelimi Tespiti İle İlgili Çalışmalar.....	8
2.1.1. Görünüm tabanlı yöntemler.....	8
2.1.2. Model tabanlı yöntemler.....	8
2.1.3. Manifold gömme yöntemi	8
2.1.4. Doğrusal olmayan regresyon yöntemi.....	9
2.2. Yüz İfadesi Tespiti İle İlgili Çalışmalar.....	9
2.2.1. Geometrik tabanlı yöntemler.....	10
2.2.2. Görünüm tabanlı yöntemler.....	10
2.2.3. Yüz ifadesinde kullanılan derin öğrenme yaklaşımları.....	11
BÖLÜM 3.	
İLGİ TESPİT SİSTEMİNDE KULLANILAN YÖNTEMLER.....	13
3.1. Viola&Jones Algoritması.....	13

3.1.1. İntegral görüntü.....	13
3.1.2. Öznitelik çıkarımı (Haar-benzeri öznitelikler).....	14
3.1.3. Adaboost algoritması.....	16
3.1.4. Aşamalı sınıflandırıcı.....	17
3.2. Kanade-Lucas-Tomasi (KLT) Algoritması.....	18
3.3. Evrişimsel Sinir Ağları (Convolutional Neural Network, CNN).....	21
3.3.1. Evrişimsel sinir ağlarını oluşturan katmanlar.....	22
3.3.1.1. Giriş katmanı (input layer).....	23
3.3.1.2. Konvolüsyon katmanı (convulation layer).....	23
3.3.1.3. Havuzlama katmanı (pooling layer).....	25
3.3.1.4. Aktivasyon katmanı.....	27
3.3.1.5. Tam bağlantılı katman (fully connected layer).....	28
3.3.1.6. Dropout katmanı.....	28
3.3.1.7. Sınıflandırma katmanı.....	29
3.3.2. Evrişimsel sinir ağlarda eğitim.....	29
3.3.3. Literatürde başarılı sonuçlar üreten derin öğrenme mimarilerinden örnekler.....	33
3.3.3.1. AlexNet.....	33
3.3.3.2. ZFNet.....	34
3.3.3.3. GoogleNet.....	34
3.3.3.4. Microsoft resNet.....	35

BÖLÜM 4.

İLGİ TESPİT SİSTEMİNİN UYGULANMASI.....	37
4.1. Uygulamada Kullanılan Veritabanı.....	37
4.2. Geliştirilen İlgi Tespit Sisteminin Mimarisi.....	40
4.3. Baş Yönelimi İçin Geliştirilen CNN (CNN-1).....	42
4.4. Yüz Bölütleme İçin Geliştirilen CNN (CNN-2).....	44
4.5. Yüz İfade Tespiti İçin Geliştirilen CNN (CNN-3).....	50
4.6. Deneysel Sonuçlar.....	52
4.7. Önerilen İki Aşamalı 5 Kanallı Yüz İfadesi Tespiti Yöntemin Literatürdeki Diğer Çalışmalar İle Karşılaştırılması.....	59

4.8. İlgı Tespit Sisteminin Çalışmasından Örnekler.....	60
BÖLÜM 5.	
TARTIŞMA VE SONUÇ.....	63
KAYNAKLAR.....	67
ÖZGEÇMİŞ.....	77



SİMGELER VE KISALTMALAR LİSTESİ

AAM	:Active Appearance Model
ADALINE	:Adaptive linear element
ANN	:Artificial Neural Network
ASM	:Active Shape Model
BDBN	:Boosted Deep Belief Network
CNN	:Convolutional Neural Network
DBN	:Deep Belief Network
DFD	:Discriminative Feature Dictionary
EEG	:Elektroensefalografi
FACS	:Facial Action Coding System
FRR-CNN	:Feature Redundancy-Reduced Convolutional Neural Network)
GLCM	:Gray-Level Cooccurrence Matrix
HMM	:Hidden Markov Model
HOG	:Histogram of Oriented Gradients
JAFFE	:Japanese Female Facial Expression Database
KLT	:Kanade-Lucas-Tomasi
KDEF	:Karolinska Directed Emotional Faces
KNN	:K-Nearest Neighbour
LBP	:Local Binary Pattern
LDA	:Linear Discriminant Analysis
LMS	:Least Mean Square
LSiBP	:Local Saliency-inspired Binary Patterns
MLP	:Multi Layer Perceptron
NNE	:Neural Network Ensemble
PCA	:Principal Component Analysis
RaFD	:Radboud Face Database

RBF	:Radial Basis Function
ReLU	:Rectified Linear Unit
RGB	:Red, Green, Blue
SDK	:Software Development Kit
SVM	:Support Vector Machines
Tanh	:Tangent hyperbolic
TFEID	: Taiwanese Facial Expression Image Database
YSA	:Yapay Sinir Ağları



ŞEKİLLER LİSTESİ

Şekil 3.1. İntegral görüntüde bir pikselin temsili.....	14
Şekil 3.2. İntegral görüntüde bir alanın, etrafındaki pikseller ile ifade edilmesi..	14
Şekil 3.3. Öznitelik çerçevelerinin yüz görüntüsü üzerinde gezdirilmesine dair bir örnek.....	15
Şekil 3.4. Aşamalı sınıflandırıcı.....	18
Şekil 3.5. ImageNet yarışmasının yıllara göre hata oranları.....	22
Şekil 3.6. Evrişimsel Sinir Ağlarını oluşturan katmanlar.....	23
Şekil 3.7. Konvolüsyon işleminin uygulanması.....	24
Şekil 3.8. Üç kanallı bir görüntünün, üç kanallı bir filtre ile konvolüsyona tabi tutulması.....	25
Şekil 3.9. Maksimum havuzlama işlemine dair bir örnek.....	26
Şekil 3.10. Ortalama havuzlama işlemine dair bir örnek.....	26
Şekil 3.11. Sigmoid aktivasyon fonksiyonunun grafiği.....	27
Şekil 3.12. Tanh aktivasyon fonksiyonunun grafiği.....	28
Şekil 3.13. Relu aktivasyon fonksiyonunun grafiği.....	28
Şekil 3.14. Dropout katmanının etkisi.....	29
Şekil 3.15. Bir sinir ağının genel yapısı.....	30
Şekil 3.16. Alex Net mimarisi.....	33
Şekil 3.17. ZF Net mimarisi.....	34
Şekil 3.18. GoogleNet mimarisi.....	35
Şekil 3.19. Microsoft ResNet mimarisi.....	36
Şekil 4.1. Kullanılan RaFD veritabanında bulunan bir kişiye ait tüm yüz ifadeleri.....	38
Şekil 4.2. Kullanılan KDEF veritabanında bulunan bir kişiye ait tüm yüz ifadeleri.....	38

Şekil 4.3. Yüz bölütleme işlemi için gereken maskelerin oluşturulması.....	39
Şekil 4.4. Evrensel yüz ifadelerinin pozitif ve negatif duygulara göre sınıflandırılma şeması.....	40
Şekil 4.5. Önerilen müşteri davranış izleme sisteminin akış şeması.....	41
Şekil 4.6. RaFD’de bulunan baş yönelimi görüntülerinden örnekler.....	42
Şekil 4.7. Önerilen sistemin CNN mimarisi.....	43
Şekil 4.8. Baş yönelimi tespiti için geliştirilen CNN (CNN-1) mimarisi.....	44
Şekil 4.9. Blok durumlarının belirlenmesi.....	45
Şekil 4.10. Yüz bölütleme için eğitim ve test görüntülerinin bloklara ayrılması..	46
Şekil 4.11. Yüz bölütleme için geliştirilen CNN (CNN-2) mimarisi.....	47
Şekil 4.12. Orijinal yüz görüntüsü.....	47
Şekil 4.13.CNN-2’ye ait ilk konvolüsyon katmanı sonuçları.....	48
Şekil 4.14. CNN-2 ikinci konvolüsyon katmanı çıktıları.....	49
Şekil 4.15. CNN-2 üçüncü konvolüsyon katmanı çıktıları.....	49
Şekil 4.16. CNN-2 son konvolüsyon katmanı çıktıları.....	50
Şekil 4.17. Yüz ifade sınıflandırması için kullanılan CNN (CNN-3) mimarisi...	51
Şekil 4.18. Önerilen sistemin yanlış bulduğu yüz ifadesi görüntülerinden örnekler.....	59
Şekil 4.19. KLT algoritması ile yüz anahtar noktalarının tespit edilip, yüzün etiketlenmesi.....	61
Şekil 4.20. Sistemin çalışma zamanından örnekler.....	61
Şekil 4.21. Sistemin iki kişiye ait yüz görüntüsünü etiketleyip, takip etmesi ve yüz ifadelerini tespit etmesinden bir örnek.....	62

TABLolar LİSTESİ

Tablo 3.1. Viola Jones yönteminde öznitelik çıkarıcı kareler.....	15
Tablo 4.1. Önerilen CNN sisteminin detaylı katman bilgileri.....	52
Tablo 4.2. RaFD için baş yönelimi tespiti başarımlar oranları.....	53
Tablo 4.3. KDEF için baş yönelimi sonuçları.....	54
Tablo 4.4. Yüz bölütleme aşamasında eğitimde kullanılan blok sayıları.....	54
Tablo 4.5. 5 kanallı giriş ile eğitilen yüz ifadesi tespiti sisteminin RaFD için başarımlar oranları.....	55
Tablo 4.6. RaFD'ye ait sadece orijinal yüz görüntüleri ile CNN-3 sisteminin eğitilmesi sonucunda elde edilen başarımlar oranları.....	56
Tablo 4.7. RaFD'ye ait sadece bölütlenmiş yüz görüntüleri ile CNN-3 sisteminin eğitilmesi sonucunda elde edilen başarımlar oranları.....	57
Tablo 4.8. RaFD'ye ait özet yüz görüntüleri ve orijinal yüz görüntülerinin birlikte giriş olarak verilmesi ile eğitilen yüz ifadesi tespiti sisteminin başarımlar oranları.....	57
Tablo 4.9. Pozitif ve negatif duygular olarak gerçekleştirilen iki sınıflı eğitimin RaFD için karmaşıklık matrisi.....	57
Tablo 4.10. 5 kanallı giriş ile eğitilen yüz ifadesi tespiti sisteminin KDEF için başarımlar oranları.....	58
Tablo 4.11. Önerilen yüz ifadesi tespiti sisteminin literatürdeki diğer çalışmalar ile karşılaştırılması.....	60

ÖZET

Anahtar kelimeler: Yüz ifadesi tespiti, baş yönelimi tespiti, yüz analizi, müşteri izleme, evrimsel sinir ağları

Pazarlama alanında, en heyecan verici, yenilikçi ve gelecek vaat eden konulardan biri müşteri ilgisinin ölçülmesidir. Müşteri ilgisini ölçmek için geleneksel bir yaklaşım olan müşteri memnuniyet anketleri, günümüzde müşteriye rahatsız edici bir yöntem olarak değerlendirilmektedir. Diğer bir müşteri ilgisi ölçme yöntemi de bir insan gözlemcinin müşteri davranışlarını izleyip kaydetmesi şeklinde olabilir ancak bu da deneyimli ve yetenekli insan gerektirir. Ayrıca her gözlemci, insan davranışlarını farklı yorumlayabileceğinden, sonuçlar tarafsız olamayabilir. Bu nedenle müşteri davranışlarını izlemek için, rahatsız edici olmayan, nicel, tarafsız ve otomatik sonuçlar üretebilen sistemlere ihtiyaç vardır.

Bu tez çalışması ile müşteri davranışının bilgisayar aracılığı ile izlenmesi ve bir ürüne ya da reklama ilgi duyan müşterilerin belirlenmesi için derin öğrenme tabanlı bir sistem önerilmektedir. Bu sistem ilk olarak müşterinin dikkatini baş yönelimi tahminiyle ölçer. Baş pozisyonları reklama veya ilgilenilen ürüne yönelik olan müşteriler için, sistem yüz ifadelerini analiz eder ve yüz ifadesine dayalı olarak müşterilerin ürünlere veya reklamlara olan ilgisini tahmin eder. Sistem ön yüz görüntülerinin algılanmasıyla çalışmaya başlar, ardından yüz ifadesi tespiti için önemli olan ağız, göz ve kaş bileşenleri tespit edilip yüz üzerinde bölütlenir ve bölütlenmiş bir yüz görüntüsü oluşturulur. Son olarak, ham yüz görüntüleri ile birlikte, elde edilen bölütlenmiş yüz görüntülerine ait güven değerleri kullanılarak yüz ifadeleri tespit edilir. İki aşamalı olan bu yüz ifadesi tespit yöntemi, parça tabanlı özellikler ile bütünsel yüz özelliklerini birleştirerek daha güçlü bir yüz ifadesi sistemi sunar. Sistemde ayrıca müşteri yüzleri etiketlenerek video çerçevesi boyunca takip edilir. Her müşteriye ait yüz ifadeleri belirli bir süre boyunca depolanır ve bu süre sonunda müşterinin ürüne ilgili olup olmadığı ile ilgili sonuç bildirilir.

Önerilen sistem müşteri davranışlarının izlenmesine ek olarak, farklı odak gruplarının çeşitli fikirlere, resimlere, seslere, kelimelere ve diğer uyaranlara duygusal tepkisini izlemek için de kullanılabilir.

DEVELOPMENT OF AN INTEREST DETECTION SYSTEM BASED ON FACIAL ANALYSIS USING DEEP LEARNING

SUMMARY

Keywords: Facial expression recognition; head pose estimation; facial analysis; customer monitoring; convolutional neural network

In the marketing research, one of the most exciting, innovative, and promising trends is quantification of customer interest. The customer satisfaction survey, which is a traditional approach to quantify customer interest, has come to be considered as an invasive method in recent years. Recording customer interest by a salesperson who observes customers' behavior during the advertisement watching or shopping phase is another approach. However, this task requires specific skills for every salesperson, and each observer may interpret customer behaviors differently. Consequently, there is a critical need to develop non-invasive, objective, and quantitative tools for monitoring customer interest.

This study presents a deep learning-based system for monitoring customer behavior specifically for detection of interest. The proposed system first measures customer attention through head pose estimation. For those customers whose heads are oriented toward the advertisement or the product of interest, the system further analyzes the facial expressions and reports customers' interest. The proposed system starts by detecting frontal face poses; facial components important for facial expression recognition are then segmented and an iconized face image is generated; finally, facial expressions are analyzed using the confidence values of obtained iconized face image combined with the raw facial images. This approach fuses local part-based features with holistic facial information for robust facial expression recognition. The system is also tracked human faces along the video frame by labeling the faces. The facial expressions of each customer are stored for a certain period of time; at the end of this period, the result of whether the customer is related to the product or advertisement is notified.

With the proposed processing pipeline, using a basic imaging device, such as a webcam, head pose estimation and facial expression recognition is possible. The proposed pipeline can be used to monitor emotional response of focus groups to various ideas, pictures, sounds, words, and other stimuli.

BÖLÜM 1. GİRİŞ

İnsanın görsel ilgisinin bilgisayar sistemleri aracılığı ile tespit edilmesi çeşitli alanlarda kullanılabilecek, gelecek vaat eden bir alandır. Teknolojik gelişmeler, kameraların ve güçlü işlemcilerin gelişmesi ve bilgisayar görmesi alanındaki gelişmeler ilgi tespitinin otomatikleştirilmesini mümkün kılmıştır. Bu tez çalışması kapsamında müşteri ilgisi tespiti senaryosu kurgulanarak, bu senaryoya dayalı olarak bir ilgi tespit sistemi geliştirilmiştir.

Müşterilerin ürünlere ve reklamlara olan ilgisinin belirlenmesi pazarlama ve reklamcılık alanının önemli konularındandır [1]. Bu amaçla yapılan ve geleneksel bir yaklaşım olan müşteri anketleri, son zamanlarda müşteriyi rahatsız edici bir yöntem olarak değerlendirilmektedir [1]. Ayrıca müşterilerin bu anketlere cevap verme oranı da gün geçtikçe azalmaktadır. Bu nedenle pazarlama alanında müşterilerin ürünlere ve reklamlara olan ilgisini tespit etmek için yeni yöntemler bulmak önemli bir problem haline gelmiştir [1].

Müşterilerin ilgilerinin belirlenmesinde diğer bir yaklaşım, müşteri bir reklam izlerken veya bir ürünün başındayken bir insan gözlemci tarafından davranışlarının incelenip kayıt altına alınması şeklinde olabilir. Ancak bu yaklaşım, deneyimli ve özel becerilere sahip insanlar gerektirir [2]. Ayrıca her gözlemci müşteri davranışlarını farklı yorumlayabilir. Missaglia ve ark. [3] çalışmalarında, öznel duygusal algılara dayalı yöntemlerin, kişisel duygusal durumu yakalamak için her zaman doğru şekilde çalışmayabileceğini ileri sürmüştür. Buna karşılık, otomatik ölçümler daha doğru ve güvenilir bir çıktı sağlar [3]. Bu bilgilere dayanarak, müşteri ilgisinin izlenmesi için müşteriyi rahatsız etmeyen, objektif ve nicel sonuçlar üretebilen, otomatik sistemlere kritik bir ihtiyaç vardır.

Son zamanlarda pazarlama alanında insan davranışlarını incelemek için, bu görevi otomatik olarak yerine getirebilen araçlar sıklıkla kullanılmaktadır. Bunlara örnek olarak; beyin görüntüleri [4], EEG (Elektroensefalografi, beyin akım grafiği) [5], [6], göz izleme [7]–[9], kalp atış hızı kaydı [10], yüz analizi çalışmaları [11], [12] vb. verilebilir.

Ürünlere veya reklamlara olan müşteri ilgisini tespit etmek için, ilk olarak görsel ilgi odağı tespit edilmelidir. Literatürde birçok görsel ilgi odağı tespiti çalışması, baş yönelimine dayalı olarak çalışılmıştır [13]–[15]. Baş yönelimi tespiti; robotik [16], trafik [17]–[19], eğitim [20], sağlık [21] gibi birçok amaçla çalışılmaktadır.

Müşterinin duygusal modunun belirlenmesi de, pazarlama alanı için önemli bir konudur [22]. Çünkü müşterinin duygusal modu, satın alma kararını etkilemektedir [22]. İnsanlar olumlu modda iken, çevrelerini daha güvenli hissederler ve satın almaya, yeni ürünler denemeye, yeni teklifleri değerlendirmeye daha eğilimlidirler. Negatif modda olduklarında ise, çevrelerinde sürekli bir tehdit varmış hissine kapılıp, dikkatli ve uyanık olmaları gerektiğini düşünürler [23]. Bu nedenle satın almaya daha kapalıdır. Müşteri duyguları ve modlarının anlaşılması satıcı için de önemlidir [24]. Bu psikolojik durumların doğru anlaşılması, pazarlamacının olumlu duygulara yol açan uyaranlar tasarlamasına, marketlerinde ürünlerini buna göre yerleştirmelerine yol gösterebilir [24].

Son zamanlarda geliştirilen uygulamalarda, müşteri duygularını anlamak için başvurulan yöntemlerden biri olarak yüz ifadesi tespiti ile karşılaşılmaktadır [11], [12]. Yüz ifadeleri analiz edilerek, müşteri duyguları belirlenip, pazarlamacının satışları artıracak doğru kararlar alması mümkün olabilir. Mehrabian’a göre, yüz yüze iletişimde beden dili, sözlerden daha etkilidir. Mehrabian çalışmasında; yüz yüze iletişim esnasında sözlerin %7, ses tonunun %38, beden dilinin ise %55 oranında etkili olduğunu vurgulamıştır. Yüz ifadeleri de önemli beden dili araçlarından. Ünlü psikoloji bilimcileri Ekman ve Friesen ise çalışmalarında [25], iğrenme, mutluluk, korku, üzüntü, kızgınlık ve şaşkınlık olmak üzere, altı yüz ifadesinin evrensel olduğunu tespit etmişlerdir. Ayrıca Ekman ve ark. [26]; kızgınlık, korku ve iğrenme

duygularını olumsuz duygu olarak, mutluluk duygusunu ise olumlu duygu olarak kategorize etmişlerdir. Şaşkınlık duygusunun ise olumlu ya da olumsuz duygu olarak sınıflandırılmayacağını ileri sürmüşlerdir. Laros ve ark. [27] ise mutlu müşterilerin daha iyimser, umutlu, coşkulu, cesaretlendirilmiş, memnun, neşeli, rahatlamış ve heyecanlı hissettiklerini gözlemlemiştir.

Yüz ifadelerinin sistemler aracılığı ile tespit edilmesi, bilgisayar görmesinin önemli konularındandır. Yüz ifadesi tespiti, avatar animasyonu [28], medikal [29], robotik [30], trafik [8][31], akıllı çevreler [32], insan-bilgisayar etkileşimi [33][34] gibi alanlarda kullanılan aktif bir çalışma alanıdır. Literatürdeki yüz ifadesi tespiti çalışmaları genellikle Ekman ve Friesen [25] tarafından belirlenen altı evrensel duygu tabanlı çalışılmaktadır.

Bu tez çalışmasında müşterilerin yüzlerini inceleyerek, pazarlamacıya potansiyel müşterilerin o anki duygu modları ve ürünlere olan müşteri ilgisi ile ilgili bilgi veren bir sistem geliştirilmiştir. Müşteri davranışlarını incelemek için müşterilerin görsel ilgi odağına ve yüz ifadelerine odaklanılmıştır. Görsel ilgi odağı, baş yönelimi ile belirlenmiştir. Baş yönelimi tespiti ve yüz ifadesi tespiti için derin öğrenme tabanlı algoritmalar geliştirilmiştir.

Önerilen sistemde ilk olarak kameranın görüş açısına giren insan yüzleri Viola&Jones algoritması kullanılarak tespit edilir. Ardından insanların görsel ilgi odağını tespit edebilmek amacıyla baş yönelimleri incelenir. Bu aşamada, insan davranışlarını inceleyen kameranın, ilgi duyulması beklenen ürün ya da reklamın üzerinde olduğu varsayılmıştır. Bu nedenle müşterinin baş yöneliminin kameraya doğru olması, görsel ilgi odağının ilgili üründe olduğu anlamına gelmektedir. Eğer baş yönelimi ilgili ürün ya da reklama doğru değil ise sistem bir sonraki aşamaya geçmez ve yeniden ortamdaki yüzleri tespit etmeye başlar. Eğer baş yönelimi, ilgili ürün ya da reklama doğru ise sistem yüz ifadelerini tespit etmeye başlar. Sistem, müşteri yüz ifadelerini uzman tarafından belirlenen spesifik bir süre boyunca izler ve depolar. Bu sürenin sonunda ise, sistem süre boyunca depolanan tüm yüz ifadelerini analiz ederek müşterinin olumlu ya da olumsuz modda olduğu ile ilgili karar verir.

Çalışmada, müşterinin reklam panosunun önündeyken veya bir ürünü incelerken ortamdaki ayrılması ya da kamera açısına göre yer değiştirmesi, ayrıca kamera açısına başka müşteriler girmesi olasılıkları da göz önüne alınmıştır. Bu durumda sistem her bir müşteri yüzünü etiketleyerek, yüz ifadelerini kişi bazlı olarak ayrı ayrı analiz eder ve kaydeder. Her bir kişinin modunu kişi etiketiyle birlikte ekranda gösterir. Video çerçeveleri boyunca insan yüzlerini etiketleyip takip etmek için Kanade-Lucas-Tomasi (KLT) algoritmasından yararlanılmıştır.

Önerilen sistem sayesinde, reklam panosunun üzerine veya ürünün üzerine basit bir kamera yerleştirilerek veya online reklamlar için webcam kullanılarak, baş yönelimi tespiti ve yüz ifadesi tespiti sistemi mümkün olabilecektir. Önerilen sistem insanı rahatsız etmeyen, nicel sonuçlar üretebilen, düşük maliyetli bir insan davranışı izleme sistemidir. Bu tez çalışmasının motivasyonu müşteri davranışlarını incelemek olmasına rağmen, önerilen sistem aynı zamanda sağlık, güvenlik, sanal gerçeklik, trafik gibi birçok alanda, ilgi tespiti konulu çalışmalarda kullanılma olanağı bulabilir.

Çalışmada yüz ifadesi tespiti kısmı iki aşamalı olarak ele alınmıştır. Bunun için birbirine seri bağlı olan iki Evrimsel Sinir Ağı (Convolutional Neural Network, CNN) mimarisi geliştirilmiştir. İlk olarak; bölütleme amacı ile geliştirilen CNN mimarisi ile yüz ifadesi tespiti için önemli rol oynayan kaş, göz ve ağız bileşenleri yüz görüntülerinde bölütlenmiştir. Böylece bölütlenmiş, özet yüz görüntüleri elde edilmiştir. Ardından bu bölütlenmiş özet yüz görüntülerine ait olasılıksal değerler (confidence value), ham yüz görüntüleri ile bir matriste birleştirilerek, yüz ifadesi tespiti için geliştirilen diğer CNN'ye verilmiştir.

Önerilen iki aşamalı yüz ifadesi tespiti yönteminin, sisteminin başarı oranına etkisini gösterebilmek için, yüz ifadesi tespiti için geliştirilen CNN mimarisi sadece veritabanındaki ham yüz görüntüleri ile ve sadece ikinci CNN'den elde edilen bölütlenmiş özet yüz görüntüleri ile de eğitilmiştir. Son olarak yüz ifadesi tespiti için geliştirilen CNN mimarisi; bölütlenmiş, özet yüz görüntülerinin olasılıksal değerleri yerine, görüntülerin bizzat kendisinin ham görüntüler ile birleştirildiği matris ile

eğitilmiştir. Elde edilen tüm sonuçlar karşılaştırıldığında, ikinci CNN'den elde edilen olasılıksal değerler ile birlikte ham yüz görüntüleri kullanmanın çalışmaya özgünlük katmanının yanında başarımı da önemli ölçüde etkilediği sonucuna varılmıştır.

Bu çalışmanın bilime katkıları aşağıdaki gibi dört madde ile özetlenebilir.

- Rahatsız edici olmayan, nicel sonuç üreten, tarafsız ve düşük maliyetli bir müşteri davranışı izleme sistemi önerilmiştir.
- Önerilen sistem müşteri görüntülerini kaydetmez, böylece kişisel gizliliği korur. Sistem sadece çalışma zamanında kamera karşısında olan insan görüntüsü üzerinde işlem yapar ve işlem yapmak için görüntülerin kaydedilmesine ihtiyaç duymaz. Eğer arşivleme amacıyla insan yüz ifadelerine ait bilgilerin kaydedilmesi gerekiyorsa da, bölütleme ile elde edilen özet yüz görüntüleri arşivleme amacıyla kaydedilebilir. Bu görüntüler kişisel hiçbir veri içermediği gibi, anlamlı yüz ifadesi bilgilerini barındırır.
- Önerilen algoritma, çok görevli (multi-task) ve üçlü ardışıl (3-cascade) CNN kullanır. Üçüncü CNN yapısı, giriş olarak özet yüz görüntüsünün olasılıksal değerleri ile birlikte ham yüz görüntüsünü birlikte kullanır. Özet yüz görüntüsünün olasılıksal değerleri parçalı bilgiler içerirken, ham görüntüler bütünsel bilgiler içerir. Parça tabanlı ve bütünsel bilgilerin birlikte kullanılması yüz ifadesi sistemini güçlendirir.
- Yüz bölütleme aşamasında; sistem, yüz ifadesi tespiti için önemli rol oynayan yüz bölgelerini bölütler ve bir özet yüz görüntüsü oluşturur. Yüz ifadesi tespiti aşamasında, bu özet görüntünün olasılıksal değerleri giriş olarak kullanılır. Böylece yüz ifadesi tespiti yapılırken; üçüncü CNN, ifade tespiti için önemli olan bu bilgileri ilk katmanlardan itibaren öğrenmek üzere eğitilir. Dolayısıyla kılavuzlu öğrenme gerçekleştirilmiş olur. Bu da sistem performansını artırır.

Bu tez kapsamında literatürde yayımlanan ve basılma aşamasında olan bilimsel çalışmaların listesi ise aşağıda verilmiştir:

- Basılan makaleler:

1. Yolcu G, Oztel I, Kazan S, Oz C, Bunyak F (2019) Deep Learning-based Face Analysis System for Monitoring Customer Interest. Ambient Intelligence and Humanized Computing.
2. Oztel I, Yolcu G, Oz C, Kazan S, Bunyak F (2018) iFER: facial expression recognition using automatically selected geometric eye and eyebrow features. J Electron Imaging 27:1.
3. Yolcu G, Oztel I, Kazan S, Oz C, Bunyak F (2018) Facial component segmentation using convolutional neural network. Online J Sci Technol 8:84–88.
4. Oztel I, Yolcu G, Oz C, Kazan S (2018) Facial expression recognition with robust feature selection. Online J Sci Technol 8:83–91.
 - Değerlendirme aşamasında olan makaleler:
5. Yolcu G, Oztel I, Kazan S, Oz C, Palaniappan K, Lever TE, Bunyak F (2019) Facial Expression Recognition for Monitoring Neurological Disorders based on Convolutional Neural Network. Multimedia Tools and Applications.
 - Yayımlanan 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 kitapçığının geri kalan bölümlerinin planı şu şekildedir: Bölüm 2’de çalışmanın iki önemli konusunu oluşturan baş yönelimi tespiti ve yüz ifadesi tespiti ile ilgili literatürde kullanılan yöntemler özetlenmiştir. Bölüm 3’te uygulama kapsamında kullanılan yüz tespiti, baş yönelimi tespiti, yüz ifadesi tespiti ile yüz takibi ve etiketleme yöntemleri detayları ile açıklanmıştır. Bölüm 4’te uygulamada kullanılan veritabanları hakkında bilgi verilmiş, geliştirilen uygulamalar ile bunların deneysel sonuçları açıklanmış ve bu sonuçların literatüre göre karşılaştırması detaylı olarak sunulmuştur. Son olarak, Bölüm 5’te çalışmanın sonuçları tartışılmış ve gelecek çalışmalar için öneriler sunulmuştur.

BÖLÜM 2. KAYNAK ARAŞTIRMASI

2.1. Baş Yönelimi Tespiti İle İlgili Çalışmalar

Baş yönelimi tespiti çalışmaları 4 temel başlıkta toplanabilir [35]. Bunlar görünüm tabanlı yöntemler, model tabanlı yöntemler, manifold gömme ve lineer olmayan regresyon metodlarıdır.

2.1.1. Görünüm tabanlı yöntemler

Görünüm tabanlı yöntemlerde, bir baş görüntüsü, daha önce oluşturulan arşivdeki diğer baş görüntüleri ile karşılaştırılır ve en benzer olan yönelimdeki sınıfına dahil edilir. Görünüm tabanlı yöntemlerin dezavantajı; limitli pozlara çözüm üretebilmesidir. Ayrıca bazı karşılaştırmalar işlem yükünü çok artırabilir [36].

2.1.2. Model tabanlı yöntemler

Model tabanlı yaklaşımlarda, geometrik bilgiler ve yüz anahtar noktalarının pozisyonları kullanılarak baş yönelimi tespiti yapılır. Model tabanlı yaklaşımların doğruluk oranı, görüntüden elde edilen geometrik ipuçlarının miktarı ve kalitesine bağlıdır [35].

2.1.3. Manifold gömme yöntemi

Manifold gömme yaklaşımları PCA (temel bileşen analizi, principle component analysis) gibi boyut azaltma yaklaşımlarına bağlıdır [36]. Baş pozisyonu tahmini, görüntüyü bir PCA alt uzayına yansıtmak suretiyle uygulanır ve sonuçlar bir dizi gömülü şablonla karşılaştırılır [37]. Manifold gömme yöntemlerinin zayıf yönleri;

görünüm değişiklikleri, aydınlatma ve poz gibi diğer faktörlerden etkilenebilmesidir [36].

2.1.4. Doğrusal olmayan regresyon yöntemi

Doğrusal olmayan regresyon yöntemlerinde, görüntülerden pozlara doğrusal olmayan bir haritalama oluşturmak için etiketli bir eğitim seti kullanılır. Doğrusal olmayan regresyon yöntemlerinde, parametreleri eğitmek için tutarlı bir veri kümesi gerekir [35].

CNN doğrusal olmayan regresyon yöntemine dahil edilir [35]. CNN, literatürde baş yönelimi tespiti çalışmalarında sıklıkla kullanılmakta ve yüksek başarılı sonuçlar göstermektedir. Mukherjee ve ark. [38], baş yönelimi tespiti için CNN tabanlı sınıflandırma ve regresyon yöntemi geliştirmiştir. Xu ve ark. [39], baş yönelimini tespit etmek ve yüz anahtar noktalarını yerleştirmek için CNN'den elde edilen global ve lokal öznitelikleri kullanmışlardır. Patacchiola ve ark. [35] ise baş yönelimi tespiti için CNN ve adaptive gradient yöntemi kullanmışlardır.

2.2. Yüz İfadesi Tespiti İle İlgili Çalışmalar

Yüz ifadesi çalışmaları genel olarak aşağıda belirtildiği gibi iki aşamada gerçekleştirilir.

- Yüz özniteliklerinin çıkarılması
- Çıkarılan özniteliklerin sınıflandırılması

Yüz özniteliklerinin çıkarılmasında öznitelik çıkarım yöntemleri temel olarak geometrik tabanlı (geometric-based) ve görünüm tabanlı (appearance-based) olmak üzere iki ana yaklaşımda incelenebilir [40]. Bu yaklaşımlar aşağıda detaylı olarak anlatılmıştır. Ayrıca literatürde geometrik ve görünüm tabanlı yöntemlerin birlikte kullanıldığı hibrit yöntemler de yer almaktadır [41][42].

2.2.1. Geometrik tabanlı yöntemler

Geometrik tabanlı yöntemler, ifade esnasında ortaya çıkan, yüz bileşenleri arasındaki geometrik ilişkileri dikkate alır [40]. Ağız, gözler ve kaş gibi yüz bileşenlerinin yer ve şekil bilgileri kullanılarak, yüz ifadesi tespitinde anlamlı öznitelikler çıkarılmaya çalışılır. Geometrik tabanlı yöntemlere örnek olarak; Aktif Şekil Modeli (Active Shape Model, ASM) [43] ve Optik Akış Yöntemi [44] verilebilir.

2.2.2. Görünüm tabanlı yöntemler

Görünüm tabanlı yöntemler, ifade ortaya çıkarken etkilenen yüz dokusunu inceler [45]. Görünüm tabanlı yöntemlere, Local Binary Pattern (LBP) [46], Histogram of Oriented Gradients (HOG) [47], PCA [48] örnek olarak verilebilir.

Literatürde yüz ifadesi tespiti çalışmalarında çeşitli özellik çıkarımı ve sınıflandırma algoritmalarından yararlanılmıştır. Luo ve ark. [49] yüz ifadesi tespiti için PCA ve LBP'den oluşan hibrit bir yaklaşım sunmuşlardır. Çalışmalarında ilk olarak eight eyes segmentation yöntemi ile görüntüdeki efektif bölgeler çıkarılmıştır. Ardından PCA ile tüm resimdeki gri seviyeli global öznitelikler çıkarılmış, aynı zamanda veri boyutu küçültülmüştür. LBP ile sadece ağız bölgesindeki local öznitelikler çıkarılmıştır. Bu aşamadan sonra local ve global öznitelik birleşimi SVM ile sınıflandırılmıştır. Yapılan deneyler, sunulan yöntemin geleneksel yöntemlere göre daha güçlü sonuçlar ürettiğini göstermiştir. Matamoros ve ark. [50] çalışmalarında yüz bölgesini alın-gözler ve ağız olarak 2'ye bölerek yüz ifadesi tespiti yapmışlardır. Boyut azaltma için PCA kullanmışlardır. Bulanık mantık ile yüksek başarımlı sonuçlar üretmişlerdir. Boruah ve ark. [51] ise çalışmalarında haar-like features ve Adaboost algoritmasını kullanarak yüzün göz ve ağız bölgelerini tespit etmişlerdir. Ardından bu bölgelerden çeşitli uzaklık parametreleri ile özellik çıkarılmıştır. Son olarak bu özellikler Saklı Markov Modeli (Hidden Markov Model, HMM) ile sınıflandırılmıştır. Seyed ve ark. [52] yüz ifadesi tespiti için Gabor filtresi kullanarak bütün yüzden ve yüzün ağız ve göz bölgelerinden özellik çıkarımı yapmışlardır. Naive Bayesian sınıflandırma yöntemi ile 6 yüz ifadesini sınıflandırmışlardır.

2.2.3. Yüz ifadesinde kullanılan derin öğrenme yaklaşımları

Son zamanlarda yüz ifadesi tespitinde derin öğrenme algoritmaları sıklıkla kullanılmaktadır. Yüz ifadelerinde derin öğrenme yaklaşımlarının kullanılması ile öznitelik seçimi aşaması otomatikleştirilmiştir. Lopes ve ark. [53], yüz ifade tespiti için yalnızca belirli özellikleri çıkarmak ve altı temel ifadeyi incelemek için görüntü döndürme, yüz kırpma ve yoğunluk normalleştirme gibi bazı görüntü ön işleme teknikleriyle birlikte bir CNN ağı kullanmışlardır. Pitaloka ve ark. [54], CNN ile altı yüz ifadesini sınıflandırmaya çalışmışlardır. Çalışmalarında yeniden boyutlandırma, yüz algılama, kırpma ve veri normalleştirme, yerel normalleştirme, global kontrast normalleştirme ve histogram eşitleme gibi ön işleme teknikleri uygulanmıştır. Matsugu ve ark. [55], CNN kullanarak; gülümseme algılama ve yüz tespiti için kural tabanlı bir algoritma geliştirmişlerdir. Xie ve ark. [56] yüz ifadesi tespiti için FRR-CNN (Feature redundancy-reduced convolutional neural network) geliştirmişlerdir. Bu yöntemin geleneksel CNN'den farkı, FRR-CNN'nin konvolüsyon çekirdekleri, aynı katmanın özellik haritaları arasındaki ayrımları daha çok analiz ederek farklılaşırlar. Böylece gereksiz özellikler azaltılır ve görüntünün daha kısa ve öz hali elde edilir. Liu ve ark. [57] ise yüz ifadesi tespitinde kullanılmak üzere Boosted Deep Belief Network (BDBN) geliştirmişlerdir. BDBN çerçevesi, birbirine bağlı iki öğrenme sürecinden oluşur. Bunlardan biri girdinin hiyerarşik özellik gösterimlerini öğrenen denetimsiz öğrenme, diğeri ise öğrenilen özellikleri birleştiren denetimli öğrenmedir. Khorrani ve ark. [58] yüz ifadesi tespiti için yüz aksiyon ünitelerini (Facial Action Units) derin öğrenme ile belirleyen bir CNN ağı geliştirmişlerdir. Kim ve ark. [59] güçlü yüz özniteliklerini çıkarabilmek için bir DBN (Deep Belief Network) mimarisi geliştirmişlerdir. Zhao ve ark. [60] da yüz ifadesi tespiti için DBN ve MLP (Multi Layer Perceptron) tabanlı bir yöntem geliştirmişlerdir. Önerilen yöntem, DBN'lerin denetimsiz özellik öğreniminin avantajını MLP'nin sınıflandırma avantajı ile birleştirmektedir.

Son zamanlarda farklı alanlarda yüz ifadesi tespitine dayalı sistemlerin gerçekleştirimi oldukça artmıştır. Çin'de kameralar aracılığı ile öğrencilerin derse olan ilgisini takip

edebilmek için yüz ifadesi tespitine dayalı dikkat analizi çalışması yapılmaktadır [61], [62]. Bu çalışmada öğretmene yardımcı bir dikkat analizi sistemi ile eğitimde kalitenin artırılması planlanmaktadır. Ayrıca ABD'nin ünlü marketler zinciri Walmart, mutsuz ve sinirli müşterilerini kameralar aracılığı ile tespit edip, onlara yardımcı elemanlar yönlendirmek üzere bir yüz analizi çalışması yapmaktadır [63]–[65]. Bu şekilde müşteri sorunlarının daha verimli bir şekilde çözülmesi planlanmaktadır. Geliştirilen son teknoloji otomobillerde, sürücülerin yorgunluk ve dikkatsizlik durumlarının tespit edilip, dikkatsizliğin algılandığı durumlarda onları çeşitli şekillerde uyaran yüz analizi sistemleri bulunmaktadır [17], [18], [66]. Sağlık alanında geliştirilen bir yüz analizi uygulamasında ise Parkinson hastalarının yüz ifadelerini tanıyarak, bakıcıları ile iletişimini kolaylaştıran bir araç geliştirilmiştir [67]. Ayrıca; nörolojik hastalığa sahip ve şizofreni hastalarının duygusal modunun tespitinde [29], [68], [69], güvenlikte şüpheli davranış gösteren kişilerin tespitinde [70], bebek ve yaşlılarda ağrı tespitinde [71], kişinin moduna göre davranan akıllı televizyonların programlanmasında [72], akıllı bilgisayarlarda [73], yüz ifadelerini kullanan çalışmalar literatürde kendilerine yer bulmuştur.

BÖLÜM 3. İLGİ TESPİT SİSTEMİNDE KULLANILAN YÖNTEMLER

Bu çalışmada yüz tespiti için literatürde başarılı sonuçları ile öne çıkan Viola&Jones algoritması kullanılmıştır. Video çerçevelerindeki yüzlerin izlenebilmesi ve etiketlemesi için KLT algoritmasından yararlanılmıştır. Baş yönelimi tespiti ve yüz ifadesi tespiti aşamaları için derin öğrenme yöntemlerinden CNN ile sınıflandırma işlemleri gerçekleştirilmiştir.

3.1. Viola&Jones Algoritması

Viola&Jones algoritması [74] literatürde yüz algılamada oldukça başarılı sonuçlar üretmektedir. Algoritma üç temel fikre dayanmaktadır. Bunlar: integral görüntü (integral image), Adaboost algoritması ile zayıf sınıflandırıcılardan güçlü sınıflandırıcı üretmek ve sınıflandırıcıları ardışıl bir yapıda birleştirmektir.

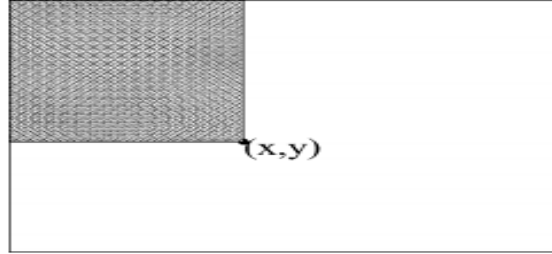
3.1.1. İntegral görüntü

İntegral görüntü hızlı özellik değerlendirmeleri sağlar ve özellik tespiti karmaşıklığını azaltır. İntegral görüntüde $P(x, y)$ noktasındaki bir piksel, kendisinin solundaki ve üstündeki piksellerin toplamı ile ifade edilir [74]. Denklem 3.1 ve Denklem 3.2’de integral görüntü oluşturma formülleri verilmiştir.

$$s(x, y) = s(x, y - 1) + i(x, y) \quad (3.1)$$

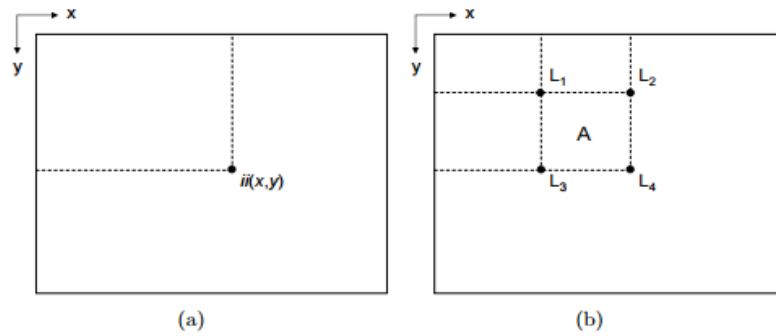
$$ii(x, y) = ii(x - 1, y) + s(x, y) \quad (3.2)$$

Denklem 3.1 ve Denklem 3.2’de $s(x, y)$ kümülatif satır toplamı, $s(x, -1) = 0$, ve $ii(-1, y) = 0$ ’dır [74]. İntegral görüntünün oluşturulması ayrıca Şekil 3.1.’de de gösterilmiştir. $P(x, y)$ noktasındaki piksel değeri koyu renkli gösterilen çerçevedeki tüm piksellerin toplamını içerir.



Şekil 3.1. İntegral görüntüde bir pikselin temsili [74].

Bu şekilde normal görüntüde bir alandaki pikseller toplanmak istenirse, integral görüntüde bu alanın dört köşesinin referansını kullanarak sadece 4 referans ile istenilen boyuttaki dikdörtgenin pikselleri toplanabilir. Şekil 3.2.’de (a) görüntüsü integral görüntü ve (b) görüntüsü de bir bölgenin piksel toplamının hesaplanması işlemini gösterir. Burada Şekil 3.2.(b)’deki A bölgesine ait piksel toplamı $L_4 + L_1 - L_2 - L_3$ ifadesinin sonucudur.



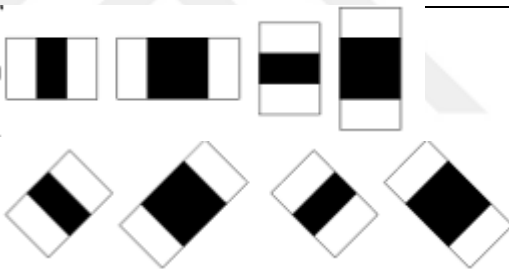
Şekil 3.2. İntegral görüntüde bir alanın, etrafındaki pikseller ile ifade edilmesi [74].

3.1.2.Öznitelik çıkarımı (haar-benzeri öznitelikler)

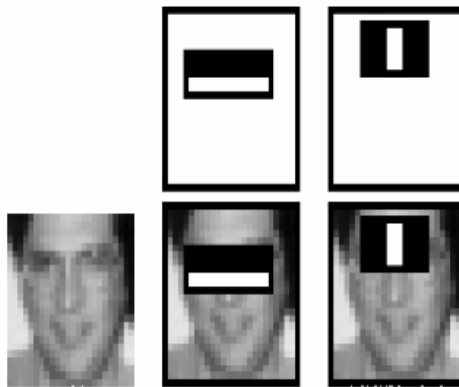
Viola&Jones algoritmasında eğitim yapılırken, içinde tespit edilecek nesnenin yer aldığı ve yer almadığı örnek görüntülere ihtiyaç duyulur. Yüz tespiti yapılırken Tablo 3.1.’de gösterilen çerçeveler, yüz görüntüleri üzerinde gezdirilir. Ardından

görüntüdeki beyaz ve siyah alanlarda kalan bölgelerin piksel değerleri ayrı ayrı toplanır ve toplamının farkı alınarak özellikler belirlenir. Elde edilen değer, belli bir eşik değerinden büyükse nesne algılandı, küçükse nesne algılanmadı şeklinde sonuç üretilir. Yani bu çerçeveler hem öznitelik çıkarıcı, hem de sınıflandırıcı özelliği gösterir. Ancak tek başına güçlü sınıflandırmalar yapamaz. Bu nedenle zayıf sınıflandırıcı (weak classifier) olarak adlandırılır.

Tablo 3.1. Viola Jones yönteminde öznitelik çıkarıcı kareler [74]

Kenar Özellikleri	
Merkez etrafını çevreleme	
Özel Köşegen hat özellikleri	
Hat özellikleri	

Öznitelik çerçevelerinin yüz görüntüsü üzerinde gezdirilmesine ait bir örnek Şekil 3.3.'de gösterilmiştir.



Şekil 3.3. Öznitelik çerçevelerinin yüz görüntüsü üzerinde gezdirilmesine dair bir örnek [74].

3.1.3. Adaboost algoritması

Adaboost algoritması bir veri indirgeme tekniği olup, aynı zamanda hem öznitelik seçici hem de sınıflandırıcı bir algoritmadır. Bu algoritma ilk olarak Freund ve Schapire tarafından önerilmiştir [75]. Haar özelliklerinden herbirini görüntüye uyguladıktan sonra elde edilen değer, belirli bir eşikten (threshold) büyükse sınıflandırmanın sonucu pozitif, değilse negatif olur. Bu işlem zayıf sınıflandırıcı olarak anlandırılır. Sonuç olarak bu haar özelliklerinin her biri birer zayıf sınıflandırıcıdır. Bu özellikler uygulandıktan sonra hata hesaplanır. Daha sonra hata oranı en düşük zayıf sınıflandırıcılar kullanılarak güçlü bir sınıflandırıcı oluşturulur. Güçlü sınıflandırıcı içerisinde yer almayan zayıf sınıflandırıcılara ilişkin öznitelikler elenmiş olur [76] [77]. Algoritmanın adımları aşağıda gösterilmiştir [76].

1. Eğitim örnekleri $(x_1, y_1), \dots, (x_n, y_n)$ olarak ayarlanır.

Burada y_i negatif örnekler için 0, pozitif örnekler için 1 olarak kabul edilir.

2. m ve l sırasıyla pozitif ve negatif örnek sayısını göstermek üzere, ağırlıklar

$$w_{1,i} = \frac{1}{2m}, \frac{1}{2l} \text{ olarak her } y_i \in \{0,1\} \text{ için ilklendir.}$$

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

– Ağırlıklar Denklem 3.3'deki gibi normalize edilir.

$$w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}} \quad (3.3)$$

– Her bir j özniteliği için, sadece bu j özniteliğini kullanan bir h_j sınıflandırıcısı eğitilir. Hata, Denklem 3.4'te gösterildiği gibi ölçülür.

$$\epsilon_j = \sum_i w_i |h_j(x_i) - y_i| \quad (3.4)$$

– En az ϵ_t hatasına sahip h_t sınıflayıcısı seçilir.

– Ağırlıklar Denklem 3.5'teki gibi güncellenir.

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i} \quad (3.5)$$

Burada x_i doğru olarak sınıflandırıldıysa $e_i = 0$, aksi halde $e_i = 1$ olur. β_t ise Denklem 3.6'daki gibi hesaplanır.

$$\beta_t = \frac{\epsilon_t}{1 - \epsilon_t} \quad (3.6)$$

4. Sonunda oluşan sınıflandırıcı Denklem 3.7'deki gibidir.

$$h(x) = \begin{cases} 1, \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0, \text{diğer} \end{cases} \quad (3.7)$$

Burada α_t Denklem 3.8'deki gibidir.

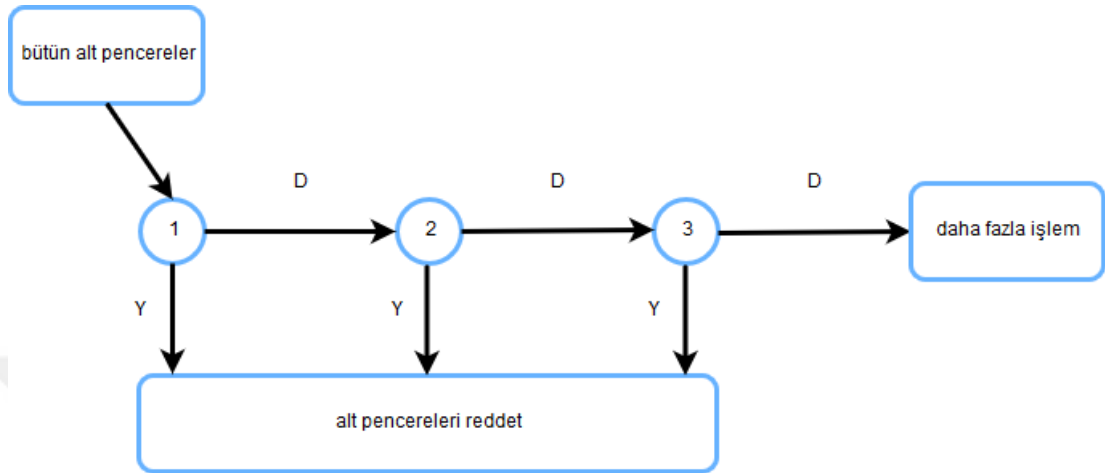
$$\alpha_t = \log \frac{1}{1 - \beta_t} \quad (3.8)$$

Adaboost algoritması çalışmaya, her bir eğitim örneği için eşit bir dağılımla başlar. Her adımda sınıflama performansına bağlı olarak en güçlü zayıf sınıflandırıcı bulunur. Ardından ağırlıklar güncellenerek bir olasılık dağılım fonksiyonu elde edilir. Sonraki adımda bu işlemler tekrarlanır, belirli sayıda iterasyon sonucunda en güçlü zayıf sınıflandırıcılar bir araya getirilerek güçlü bir sınıflandırıcı oluşturulur [77].

3.1.4. Aşamalı sınıflandırıcı

Bir görüntüde onlarca yüz bölgesi bulunmasına rağmen; yüz bulunmayan alt pencere sayısı binlerce. Bu nedenle aşamalı sınıflandırıcı fikri ortaya çıkmıştır. Bu şekilde algılayıcı bir defada tüm sınıflandırıcıları uygulamak yerine kademeli bir şekilde

çalışır ve bunların ilk aşaması da yüz olma ihtimali az olan yerleri direkt reddedecek şekilde seçilir. Bu işlem Şekil 3.4.'te gösterilmiştir. Böylece daha ilk aşamada birçok alt pencere elendiği için sınıflandırıcılar daha hızlı çalışır. [76].



Şekil 3.4. Aşamalı sınıflandırıcı [76]. D:doğru sınıflandırmayı, Y:yanlış sınıflandırmayı temsil eder.

3.2. Kanade-Lucas-Tomasi (KLT) algoritması

KLT algoritması [78], video çerçeveleri boyunca bir dizi özellik noktasını izler. KLT algoritması, nesnelerin yer değiştirme takibinde kullanılır ve pratikteki uygulamalarda oldukça başarılı sonuçlar üretir [79]. Algoritma temeli 1981 yılında Lucas ve Kanade'nin çalışması [80] ile atılmıştır. Bu çalışmada bir görüntü bölgesini takip etme yöntemi ile ilgili araştırma yapılmıştır. 1991 yılında Tomasi ve Kanade [81] ise görüntüyü takip etmek için en iyi öznitelikleri seçme konusunda araştırma yapmıştır. Bu çalışmaya göre görüntü takibinde öznitelik seçerken, pürüzsüz bölgelerden ve kenarlardan kaçmak gerekir. Bunun için görüntünün özdeğerlerinden yararlanılır.

Algoritma dört temel adımdan oluşur. Bunlar [78] :

1. Referans video çerçevesinde, izlenecek olan alanın (ilgi alanının) belirlenmesi,
2. İlgi merkezinde belirteç noktaların tespiti,
3. Belirteç noktaların izlenmesi,
4. Yanlış belirteç noktaların kaldırılması.

KLT algoritmasında video çerçevesindeki ilgi merkezinin sonraki video çerçevesindeki konumu Denklem 3.9 ve Denklem 3.10 ile hesaplanır [78].

$$J(Ax + d) = I(x) \quad (3.9)$$

$$A = 1 + D \quad (3.10)$$

Denklemde 3.9 ve Denklem 3.10'da I , referans çerçevenin x noktasındaki piksel yoğunluğunu; J , sonraki çerçevenin x noktasındaki piksel yoğunluğunu; D 2x2 deformasyon matrisini; 1, birim matrisi; d ise yer değiştirmeyi ifade etmektedir. Deformasyon matrisinin tahminindeki zorluk, genellikle yer değiştirmenin hatalı hesaplanmasına yol açmaktadır. Bu nedenle, genellikle sadece ötelemenin olduğu modeller tercih edilip, deformasyon matrisi sıfır kabul edilir [78].

İlgi alanının yer değiştirmesi $d = [x' y']$ ise, hata fonksiyonu minimize edilerek hesaplanır. Bu işlem Denklem 3.11'de gösterilmiştir [78].

$$\varepsilon = \int \int_w [J(Ax + d) - I(x)]^2 w(x) dx \quad (3.11)$$

Denklemde w ilgi alanını, $w(x)$ ise piksel noktalarına vurgu yapan ağırlık fonksiyonunu ifade eder. Ağırlık fonksiyonu genel durumda 1 alınır [78].

Hata fonksiyonu minimum değeri için, yer değiştirmeye göre bir kez türetilip sıfıra eşitlenir. $J(Ax + d)$ 'nin Taylor açılımı Denklem 3.12'de verilmiştir [78].

$$J(Ax + d) = J(x) + g^T d \quad (3.12)$$

Bu denklem hata fonksiyonunda yerine konulursa Denklem 3.13 elde edilir [78].

$$\int \int_w g(g^T d) w dx dy + \int \int_w [J(x) - I(x)] g \quad (3.13)$$

Burada;

$$g^T = [g_x g_y] = \left[\frac{\partial j}{\partial x}, \frac{\partial j}{\partial y} \right], \text{dir [78].} \quad (3.14)$$

Denklem 3.14'ün çözümü ile yer değiştirmeler hesaplanabilir ancak bu denklem doğrusal olduğundan hata oranı fazladır. Bu nedenle Newton-Raphson gibi iteratif yöntemlerden yararlanılır. İlgili alanının deformasyonu göz ardı edilip, sadece yer değiştirmesi ile ilgileniliyorsa, bu denklemin daha basit formu olan Denklem 3.15, Denklem 3.16 ve Denklem 3.17 kullanılabilir [78]:

$$Zd = c \quad (3.15)$$

$$Z = gg^T = \begin{bmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{bmatrix} \quad (3.16)$$

$$c = [I(x) - J(x)] \quad (3.17)$$

Belirteç noktaları, görüntünün parlaklığının, ışığının değişiminden, ilgili alanının dönmesi gibi faktörlerden etkilenmeyen noktalardır [78]. Bu noktaların tespiti için çeşitli yöntemler kullanılabilir. Bu çalışmada bu amaç için [78]'de anlatılan yöntemden yararlanılmıştır. Buna göre ∂_1, ∂_2 Z 'ye ait iki özdeğer ise

$$\min(\partial_1, \partial_2) > \partial \quad (3.18)$$

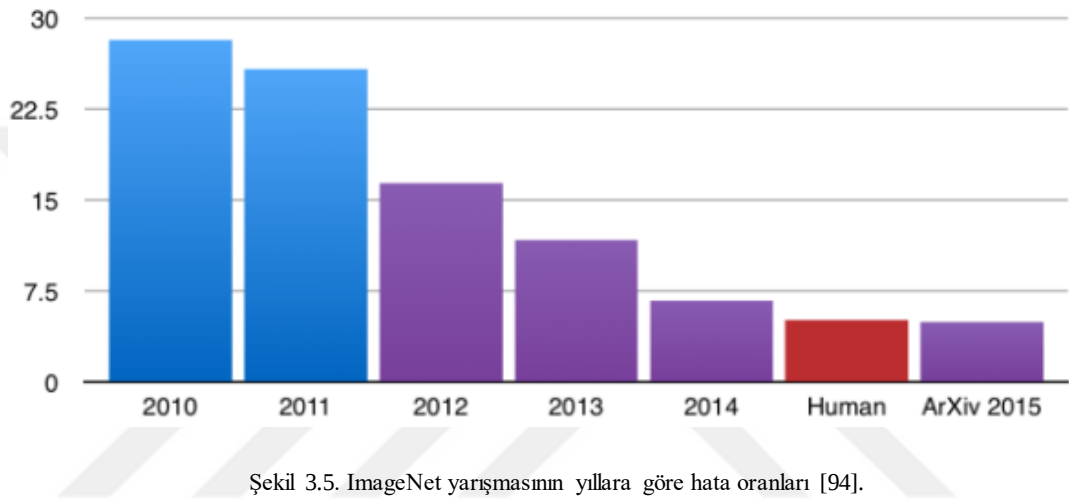
olması durumunda tatmin edici köşeler elde edilir [78]. Denklemde ∂ önceden belirlenmiş eşik değerini ifade etmektedir.

3.3. Evrişimsel Sinir Ağları (Convolutional Neural Network, CNN)

Klasik makine öğrenme yöntemlerinde gereken öznitelik çıkarımı aşaması zorlayıcı bir işlemdir. Çünkü bir problemin çözümüne etki eden özniteliklerin belirlenmesi için alanında uzman kişilere ihtiyaç duyulmaktadır [82]. Bu durum da oldukça zaman alıcıdır. Derin öğrenme kavramında ise öznitelik seçimi gerekmez, öğrenme işlemi ham veri üzerinde yapılmaktadır. Bu nedenle derin öğrenme, öznitelik çıkarmadaki problemleri ortadan kaldırmıştır. Derin öğrenme, bir eğitim süresi boyunca ham girdi ve çıktı verileri arasındaki karmaşık doğrusal olmayan ilişkilerin etkin öğrenimini sağlar [83]. Derin öğrenmenin temelleri eskiye dayanmasına rağmen popülerliği son yıllarda artmıştır. Bu durumun nedenleri; geçmişte yeterli veri elde edilememesi ve güçlü donanımların olmamasıdır [82]. Günümüzde veri kaynakları artmış ve büyük veri (big data) adı verilen yeni bir alan ortaya çıkmıştır. Ayrıca günümüzde paralel işlemcilerin kullanımı ile birden çok sayıda gizli katmana sahip ağların eğitimi ve geriye yayılım algoritmasındaki karmaşık hesaplamalar daha kolay yapılabilir hale gelmiştir [82].

İnsan karar verme yetisini modelleyen sistemler üretmek yapay zekanın önemli konularındandır. Bu amaçla ilk defa McCulloch ve ark. [84] insan sinir sisteminden esinlenerek, beyin fonksiyonlarının işleyişini mantıksal olarak hesaplayan bir model geliştirmiştir. Bu model yapay sinir ağlarının temeli olarak kabul edilmektedir. Devam eden çalışmalarda ise Adaptive linear element (ADALINE) [85] ve Perceptron [86] modelleri üretilmiştir. Ancak bu modeller sadece doğrusal problemlere çözüm üretip, XOR gibi doğrusal olmayan problemlere çözüm üretememiştir [87]. Bu durum, o yıllarda, yapay sinir ağlarına olan ilgiyi azaltmıştır [82]. Ardından, 1980'lerde ortaya atılan paralel dağıtık işlem yaklaşımları [88], [89] ile Yapay Sinir ağları (YSA, Artificial Neural Network, ANN) yeniden popülerlik kazanmıştır. 2006 yılında Hinton ve ark. [90] derin sinir ağları ile etkin öğrenme sağlanacağını göstermiştir. Ardından yapılan çalışmalar [91] [92] ile de daha performanslı çözümler üretmek için daha derin ağların tasarlanması gerektiği dikkat çekmiş ve derin öğrenme teriminin kullanımı yaygınlaşmıştır.

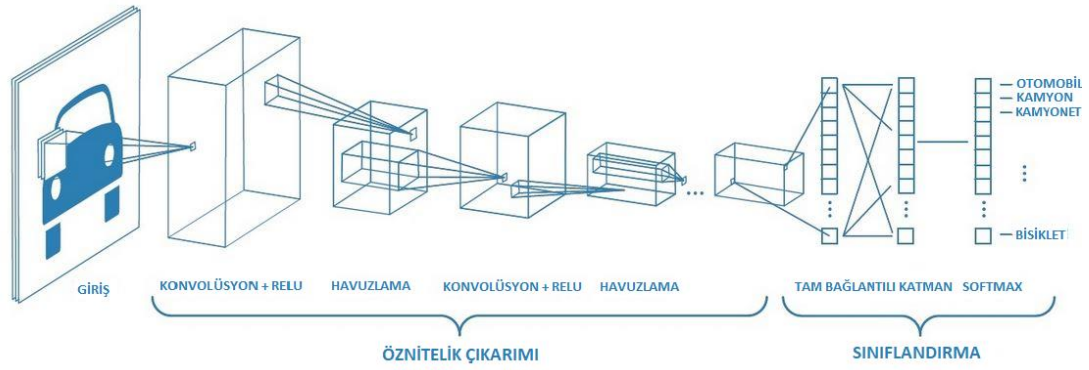
Nesne tanıma alanında önemli yarışmalardan olan ImageNet (Büyük ölçekli görsel tanıma) yarışması 2012 yılında derin öğrenmede temel mimari olan CNN ile kazanılmıştır [93]. Bu sonuç büyük etki yaratmıştır. Çünkü yarışmada, %26,1 olan hata oranı, CNN ile %15,3'e düşürülmüştür. Bu durum derin öğrenmenin popülerliğini yüksek oranda artırmıştır. Daha sonraki yıllarda derin öğrenmedeki ilerlemeler ile hata oranı %3,6'ya kadar düşürülmüştür. Hata oranının yıllara göre durumu Şekil 3.5.'te gösterilmiştir.



Derin öğrenme modellerinden CNN, literatürde görüntü analizinde sıklıkla kullanılan ve özellikle büyük boyutlu verilerde başarılı sonuçlar üreten popüler bir yaklaşımdır [54], [95], [96]. CNN'de 2 boyutlu çekirdek çerçevelerin kullanımı ile hiyerarşik olarak öğrenme gerçekleşir. CNN ile genel olarak, ardışık yerleştirilmiş eğitim bölümlerinin sonuna bir sınıflandırıcı yerleştirilmesi ile sınıflandırma işlemi gerçekleştirilebilir. CNN sınıflandırma amacıyla kullanılabildiği gibi, regresyon işlemlerini de başarıyla gerçekleştirir.

3.3.1. Evrimsel sinir ağlarını oluşturan katmanlar

CNN; Şekil 3.6.'da gösterildiği gibi giriş katmanı, konvolüsyon katmanı, havuzlama katmanı, aktivasyon katmanı, dropout katmanı, tam bağlantılı katman ve sınıflandırma gibi katmanların değişik biçimde sıralanması ile oluşur.



Şekil 3.6. Evrimsel Sinir Ağlarını oluşturan katmanlar [97].

3.3.1.1. Giriş katmanı (input layer)

CNN'nin ilk katmanıdır. Bu katmanda ağı verilecek verinin boyutu sistemin başarımını etkilemektedir. Giriş görüntüsünün büyük boyutta olması hem yüksek bellek ihtiyacı gerektirir hem de eğitim ve test süresini artırır. Buna rağmen eğitim başarısına da olumlu yansıması olabilir. Giriş görüntüsünün küçük boyutta olması ise, eğitim başarısını olumsuz etkileyebileceği gibi eğitim ve test süresini azaltır [82].

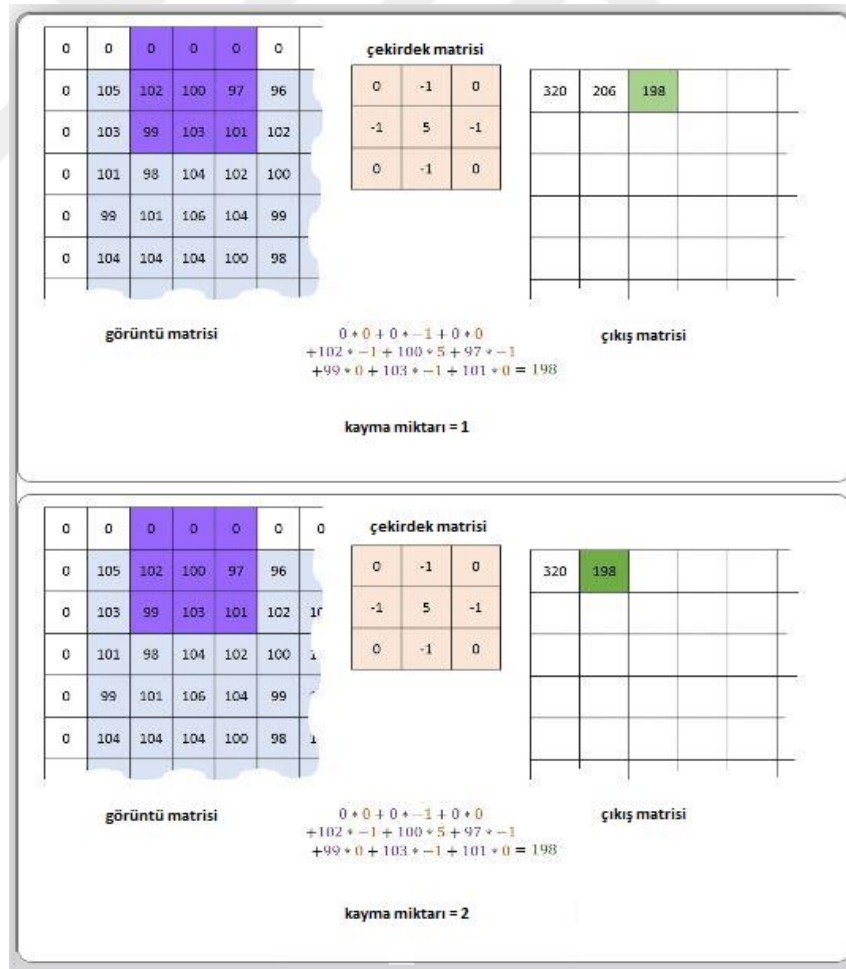
3.3.1.2. Konvolüsyon katmanı (convolution layer)

Bu katmanda belirli filtreler görüntü üzerinde dolaşarak konvolüsyon işlemini gerçekleştirir. Konvolüsyon işlemleri özellik haritalarını oluşturur. Özellik haritalarında her bir filtrenin ortaya çıkardığı özellikler haritalanır. Konvolüsyonlar genellikle birden fazla eksen üzerinde kullanılır. Örneğin, giriş olarak iki boyutlu bir görüntü, iki boyutlu bir çekirdek (kernel, filtre) ile kullanılır [98]. Konvolüsyon işleminin matematiksel ifadesi Denklem 3.19'da gösterilmiştir [98].

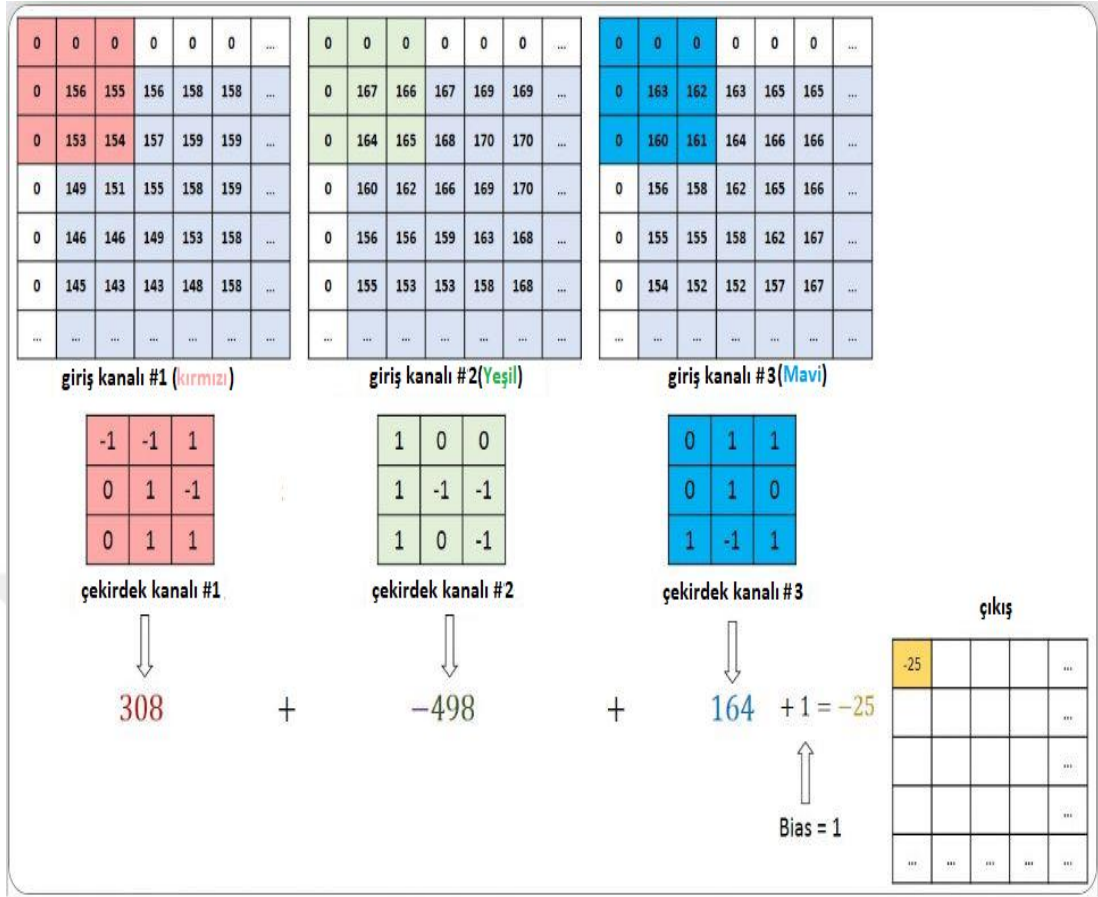
$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i + m, j + n) K(m, n) \quad (3.19)$$

Denklem 3.19'da I , giriş görüntüsünü; K , çekirdeği (kernel); S ise konvolüsyon işleminden sonra oluşan çıktıyı ifade etmektedir. Konvolüsyon işlemi literatürde '*' işareti ile sembolize edilmektedir.

Konvolüsyon işleminde; bir çekirdek çerçevesi, bir giriş görüntüsünün üzerinde sağa veya sola doğru dolaşarak görüntünün piksel değerleri ile çarpılır, ardından bunların toplamı alınarak değer çıktı verisinin ilgili piksel alanına yazılır. Sağa veya sola doğru ilerlerken, matris sınırına gelindiğinde ise filtre bir birim aşağı kaydırılarak işleme devam edilir. Bu işlem Şekil 3.7.'de bir örnekle gösterilmiştir. Renkli görüntülere konvolüsyon işlemi uygulanırken, filtreler her bir renk kanalına ayrı ayrı uygulanır ardından bunların toplamı alınır [82]. Üçünün toplamı aktivasyon haritasını oluşturur. Çekirdek matrisi görüntü üzerinde dolaşırken sağa veya sola doğru ne kadar kaydırılarak ilerleneceği bilgisi stride (kayma miktarı) adı verilen parametre ile belirlenir. Şekil 3.8.'de çok katmanlı bir giriş görüntüsü için uygulanan konvolüsyon işlemi gösterilmiştir. Burada giriş görüntüsünün her bir kanalı, çekirdek fonksiyonunun her bir kanalı ile konvolüsyon işlemine tabi tutularak, sonuçlar toplanır ve böylece çıkış görüntüsü elde edilir.



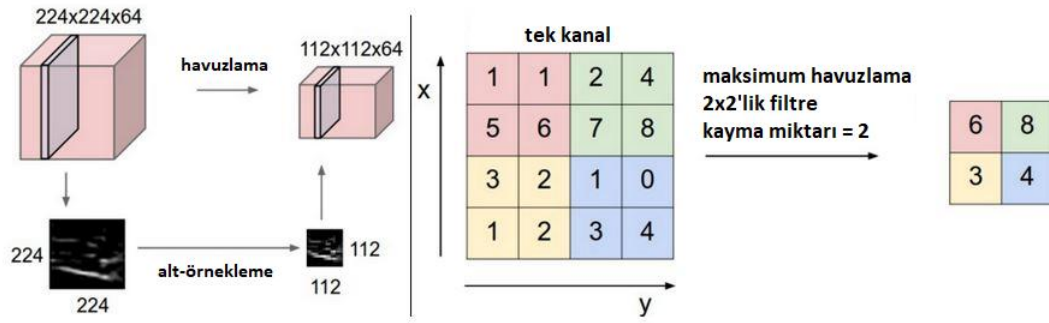
Şekil 3.7. Konvolüsyon işleminin uygulanması (kayma miktarı: 1 ve kayma miktarı:2 için) [99].



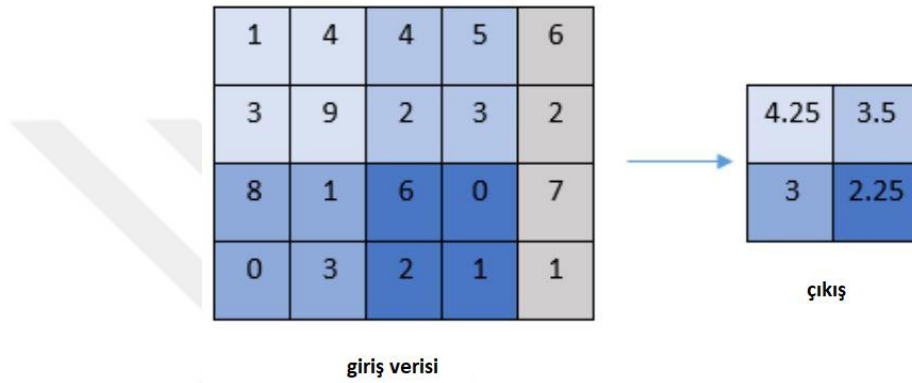
Şekil 3.8. Üç kanallı bir görüntünün, üç kanallı bir filtre ile konvolüsyona tabi tutulması [99].

3.3.1.3. Havuzlama katmanı (pooling layer)

Havuzlama katmanında temel amaç giriş boyutunu azaltmaktır. Havuzlama katmanı, giriş boyutunu genişlik ve yükseklik olarak azaltırken derinlik boyutunda bir azalma veya artma gözlenmez. Havuzlama katmanı zorunlu bir katman değildir. Tasarıma bağlı olarak ağ mimarisinde kullanılabilir ya da kullanılmayabilir. Havuzlama katmanı kendisinden sonraki katmanlar için daha az hesaplama yükü oluşturur [100]. Havuzlama işlemi iki şekilde gerçekleşir. Bunlar maksimum havuzlama ve ortalama havuzlama olarak isimlendirilmiştir. Maksimum havuzlamada; filtre içinde kalan görüntü bölümünde bulunan piksellerdeki maksimum değer, çıkış piksel değeridir. Ortalama havuzlamada ise filtre içindeki piksellerin ortalaması alınarak çıkış piksel değeri hesaplanır. Maksimum ve ortalama havuzlama işlemlerinin nasıl uygulandığını açıklayan örnekler Şekil 3.9. ve Şekil 3.10.'da gösterilmiştir.



Şekil 3.9. Maksimum havuzlama işlemine dair bir örnek [101].



Şekil 3.10. Ortalama havuzlama işlemine dair bir örnek [102].

Maksimum havuzlama genellikle daha iyi performans sağladığından literatürde daha sık tercih edilmektedir. Havuzlama işleminin ardından elde edilen çıkış görüntü boyutu filtrenin boyutuna, kayma miktarına ve giriş görüntüsünün boyutuna bağlıdır. Havuzlama işlemi ardından oluşan çıkış görüntüsünün boyutunun hesaplanma formülü Denklem 3.20, Denklem 3.21 ve Denklem 3.22’de verilmiştir [82].

$$G_2 = \frac{G_1 - F}{A} + 1 \quad (3.20)$$

$$Y_2 = \frac{H_1 - F}{A} + 1 \quad (3.21)$$

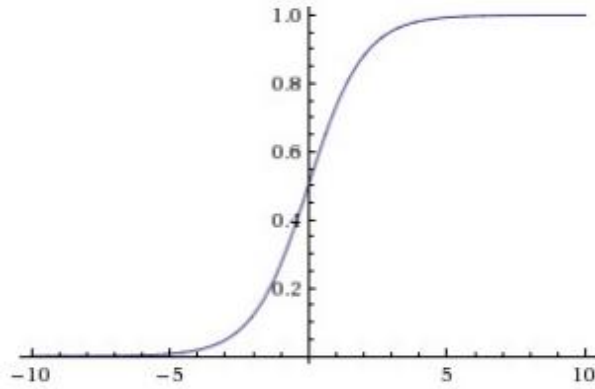
$$D_2 = D_1 \quad (3.22)$$

Denklemlerde G_1 giriş görüntüsünün genişlik değerini, G_2 çıkış görüntüsünün genişlik değerini, Y_1 giriş görüntüsünün yükseklik değerini, Y_2 çıkış görüntüsünün yükseklik değerini, F filtre boyutunu, A adım sayısını, D_1 giriş görüntüsünün derinlik boyutunu, D_2 çıkış görüntüsünün derinlik boyutunu ifade etmektedir. Literatürde sıklıkla $F = 2$ ve $A = 2$ olarak seçilir [82].

3.3.1.4. Aktivasyon katmanı

Nair ve Hinton'un çalışmasında [103] doğrultulmuş doğrusal birim (Rectified Linear Units-ReLU), tanh veya sigmoid gibi doğrusal olmayan bir aktivasyon fonksiyonu kullanan CNN'lerin daha hızlı bir şekilde eğitildiği kanıtlanmıştır. Bu fonksiyonların grafikleri ve formülleri ise aşağıda gösterilmiştir.

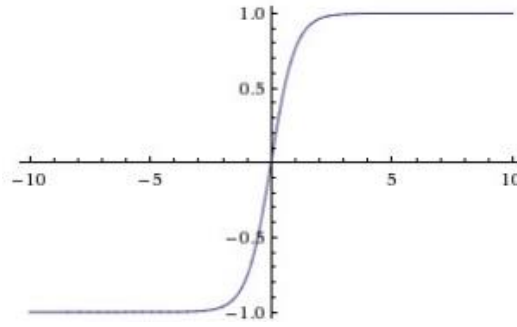
- Sigmoid aktivasyon fonksiyonu grafiği ve denklemi



Şekil 3.11. Sigmoid aktivasyon fonksiyonunun grafiği [104].

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (3.23)$$

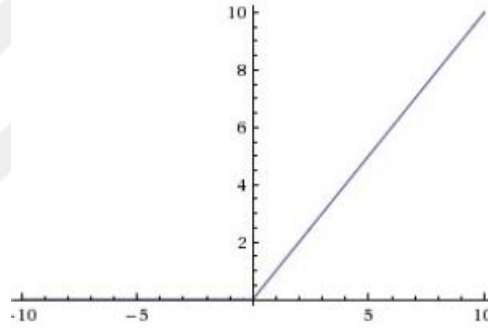
- Tanh aktivasyon fonksiyonu grafiği ve denklemi



Şekil 3.12. Tanh aktivasyon fonksiyonunun grafiği [105].

$$\tanh(x) = 2\sigma(2x) - 1 \quad (3.24)$$

– Relu aktivasyon fonksiyonu grafiği ve formülü



Şekil 3.13. Relu aktivasyon fonksiyonunun grafiği [98].

$$f(x) = \max(0, x) \quad (3.25)$$

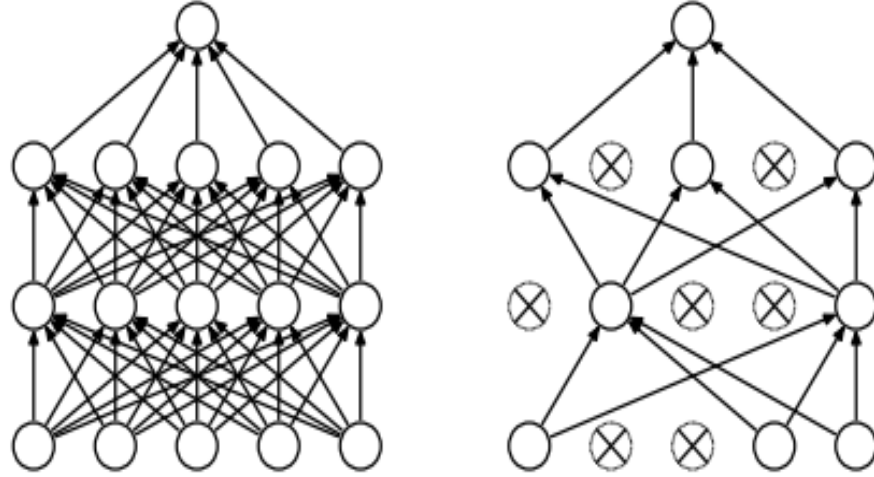
3.3.1.5. Tam bağlantılı katman (fully connected layer)

Tam bağlantılı katman kendinden önceki katmanın tüm alanlarına bağlıdır. Sınıflandırma veya regresyon katmanından önceki katmandır.

3.3.1.6. Dropout katmanı

Dropout katmanı, evrimsel sinir ağlarda ezberlemeyi önler [106]. Dropout katmanında ağı belirli bir nörona bağımlılığını önlemek için ağı bazı düğümleri kaldırılır. Dropout sayesinde ağı, belirli bilgilerin yokluğunda bile doğru bir şekilde

öğrenmeye zorlanabilir. Ağın standart durumu ve dropout katmanından sonraki durumu Şekil 3.14.'te gösterilmiştir [107].



Şekil 3.14. Dropout katmanının etkisi [107].

3.3.1.7. Sınıflandırma katmanı

Sınıflandırma işleminin yapıldığı son katmandır. Tam bağlantılı katmanın ardından gelir. Çıkış değeri sınıf sayısı kadardır. Örneğin problem 10 adet sınıfa bölünecekse sınıflandırma katmanı 10 adet çıkış üretir. Bu çıkışlarda 0-1 aralığında değerler üretilir. 1'e en yakın çıkış, ağı tahmin ettiği sınıftır. Sınıflandırma katmanında farklı sınıflandırıcılar kullanılmaktadır. Ancak yüksek başarımlar göstermesi nedeniyle literatürde sıklıkla softmax sınıflandırıcısı kullanılmaktadır [82].

3.3.2. Evrimsel sinir ağlarda eğitim

CNN'de eğitim aşağıda 4 madde ile özetlenmiştir:

1. CNN modeli tasarlandıktan sonra, filtrelerin çekirdek çerçevelerinin boyutları, sayıları ve kayma miktarı tanımlanır ve filtreler için rastgele değerler atanır.
2. Ağ ileri beslemeli olarak eğitilir; ileri beslemeli bir ağda, giriş değeri ve her bir katmanda üretilen çıktı değerler kullanılarak, bilgi ağ boyunca ileri doğru ilerler.

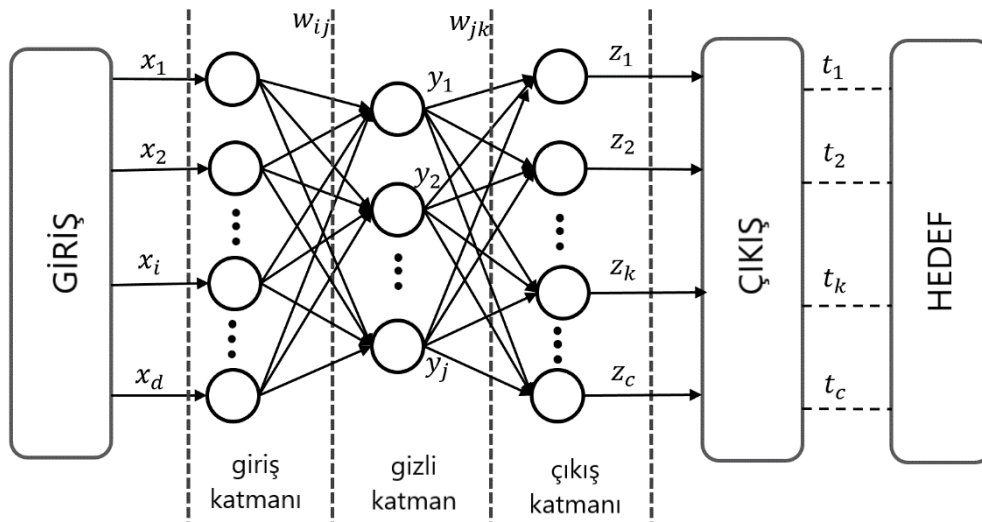
3. Ardından ağın ürettiği sonuç ile hedef sonuç bilgilerine göre hata değeri hesaplanır. Hata değerinin hesaplanmasında kullanılan yöntemlerden biri en küçük ortalamaların karesi (Least Mean Square - LMS) hatasıdır. (Denklem 3.26).

$$J(w) = \frac{1}{2} \sum_{k=1}^c (t_k - z_k)^2 \quad (3.26)$$

Denklemde z ağın ürettiği sonuç vektörünü, t ise hedef sonuç vektörünü göstermektedir. c vektör uzunluğunu, w ise ağıdaki ağırlıkları ifade eder.

4. Hata değerine göre ağın güncellenmesi için geriye yayılım (backpropagation) algoritması kullanılır. Geriye yayılım algoritmasında, hatanın hesaplanması ile elde edilen bilgiler ağ üzerinden geriye doğru ilerler. Geriye yayılım algoritması zincir kuralına göre çalışır [98]. Şekil 3.15.'teki ağ örneği için geriye yayılım algoritması çalışma prensibi aşağıda anlatılmıştır.

Ağın ağırlıkları, hatayı indirmek için Gradient Descent yöntemine göre güncellenirler. Ağırlıklardaki değişim miktarı öğrenme oranı (learning rate) olarak ifade edilir ve η ile sembolize edilir (Denklem 3.27).



Şekil 3.15. Bir sinir ağının genel yapısı [108].

$$\Delta w_{mn} = -\eta \frac{\partial J}{\partial w_{mn}} \quad (3.27)$$

Giriş katmanı ve gizli katman ağırlıkları, ağırlıkların bağılılık durumuna göre güncellenir. Hatanın gizli katmanın çıkış ağırlıklarına doğrudan bağı yoksa zincir kuralında (chain rule) türevsel genişleme kullanılmalıdır. Bu işlem Denklem 3.28'de gösterilmiştir.

$$\frac{\partial J}{\partial w_{kj}} = \frac{\partial J}{\partial net_k} \frac{\partial net_k}{\partial w_{kj}} \quad (3.28)$$

Denklemde w_{kj} gizli katman ve çıkış katmanı arasındaki ağırlıkları göstermektedir. Çıkış katmanının girişi net_k olan ağın çıkış değerlerinin duyarlılığı δ_k 'nin hesaplanması Denklem 3.29 ile gösterilmiştir.

$$\delta_k = -\frac{\partial J}{\partial net_k} \quad (3.29)$$

Denklem 3.27'nin türevi alındığında;

$$-\frac{\partial J}{\partial net_k} = \frac{\partial J}{\partial z_k} \frac{\partial z_k}{\partial net_k} = (t_k - z_k) f'(net_k) \quad (3.30)$$

$$\frac{\partial net_k}{\partial w_{kj}} = y_j \quad (3.31)$$

Denklem 3.31'in türevi alındığında ağırlık değişimi bulunur (Denklem 3.32).

$$\Delta w_{kj} = \eta \delta_k y_j = \eta (t_k - z_k) f'(net_k) y_j \quad (3.32)$$

Giriş katmanı ile gizli katman arasındaki ağırlıkların güncellenmesi de aynı işlemler ile gerçekleştirilir:

$$\frac{\partial J}{\partial w_{ji}} = \frac{\partial J}{\partial y_j} \frac{\partial y_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}} \quad (3.33)$$

$$\frac{\partial J}{\partial y_j} = \frac{\partial J}{\partial y_j} \left[\frac{1}{2} \sum_{k=1}^c (t_k - z_k)^2 \right] \quad (3.34)$$

$$= - \sum_{k=1}^c (t_k - z_k) \frac{\partial z_k}{\partial y_j} \quad (3.35)$$

$$= - \sum_{k=1}^c (t_k - z_k) \frac{\partial z_k}{\partial net_k} \frac{\partial net_k}{\partial y_j} \quad (3.36)$$

$$= - \sum_{k=1}^c (t_k - z_k) f'(net_k) w_{jk} \quad (3.37)$$

Çıkış sinyalinin, hata değerine etkisini bulabilmek için Denklem 3.38'deki denklem kullanılır.

$$\delta_j = f'(net_j) \sum_{k=1}^c w_{kj} \delta_k \quad (3.38)$$

Gizli katman ile giriş katmanı arasındaki ağırlıklar da Denklem 3.39'daki eşitlikler ile güncellenir. Denklemdeki δ_j ve δ_k gizli katman ve çıkış katmanının duyarlılık değeridir.

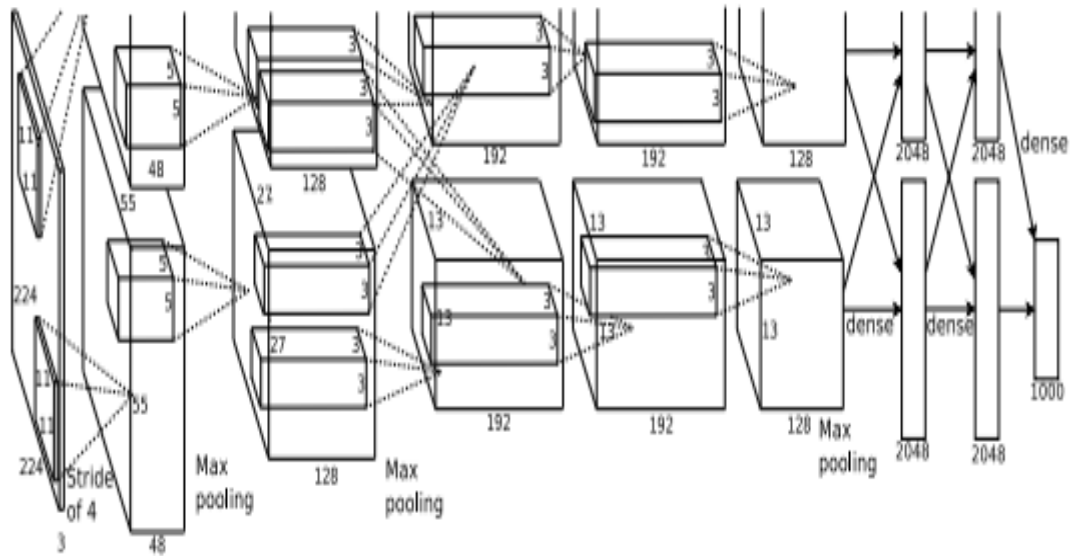
$$\Delta w_{ji} = \eta x_i \delta_j = \eta x_i f'(net_j) \sum_{k=1}^c w_{kj} \delta_k \quad (3.39)$$

3.3.3. Literatürde başarılı sonuçlar üreten derin öğrenme mimarilerinden örnekler

Bu bölümde, son yıllarda literatürde sıklıkla kullanılıp başarılı sonuçlar üreten, farklı mimarilere sahip derin öğrenme yapılarından örnekler verilmiştir. Bu derin öğrenme yapıları nesne tanıma amacıyla geliştirilmiş olup, literatürde farklı sınıflandırma amaçları ile de kullanılmaktadır. Bunu yapabilmek için ilk katmanlardaki temel öznelikleri çıkarabilen ağırlıklar değiştirilmeyip, son katmanlardaki ağırlıklar sıfırlanarak yeniden eğitim işlemi gerçekleştirilir. Literatürde transfer öğrenme (transfer learning) olarak adlandırılan bu yöntem, çoğunlukla başarılı sonuçlar üretmektedir.

3.3.3.1. AlexNet

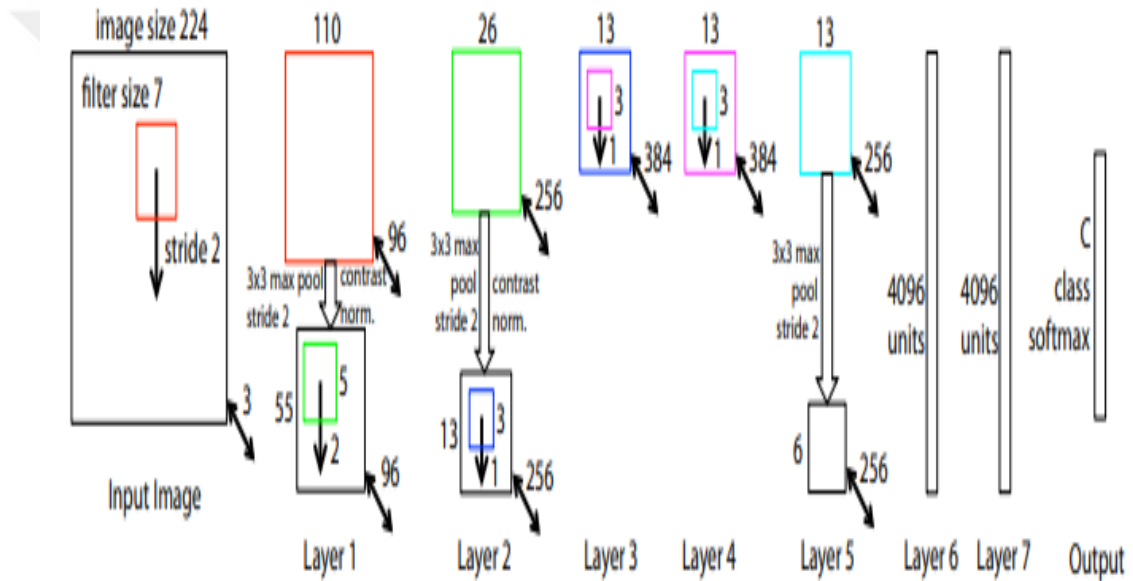
Derin öğrenmenin temeli daha eskiye dayansa da, dünya çapında duyulması 2012 yılında ImageNet yarışmasını kazanan AlexNet modeli ile olmuştur [93]. AlexNet mimarisi 1000 nesneyi sınıflandırabilmektedir. 5 konvolüsyon, 3 tam bağlantılı katmandan oluşur. Filtreler 11x11, kayma miktarı 4 olarak tasarlanmıştır. AlexNet mimarisi şekil 3.16.'da gösterilmiştir [93].



Şekil 3.16. Alex Net mimarisi [93].

3.3.3.2. ZFNet

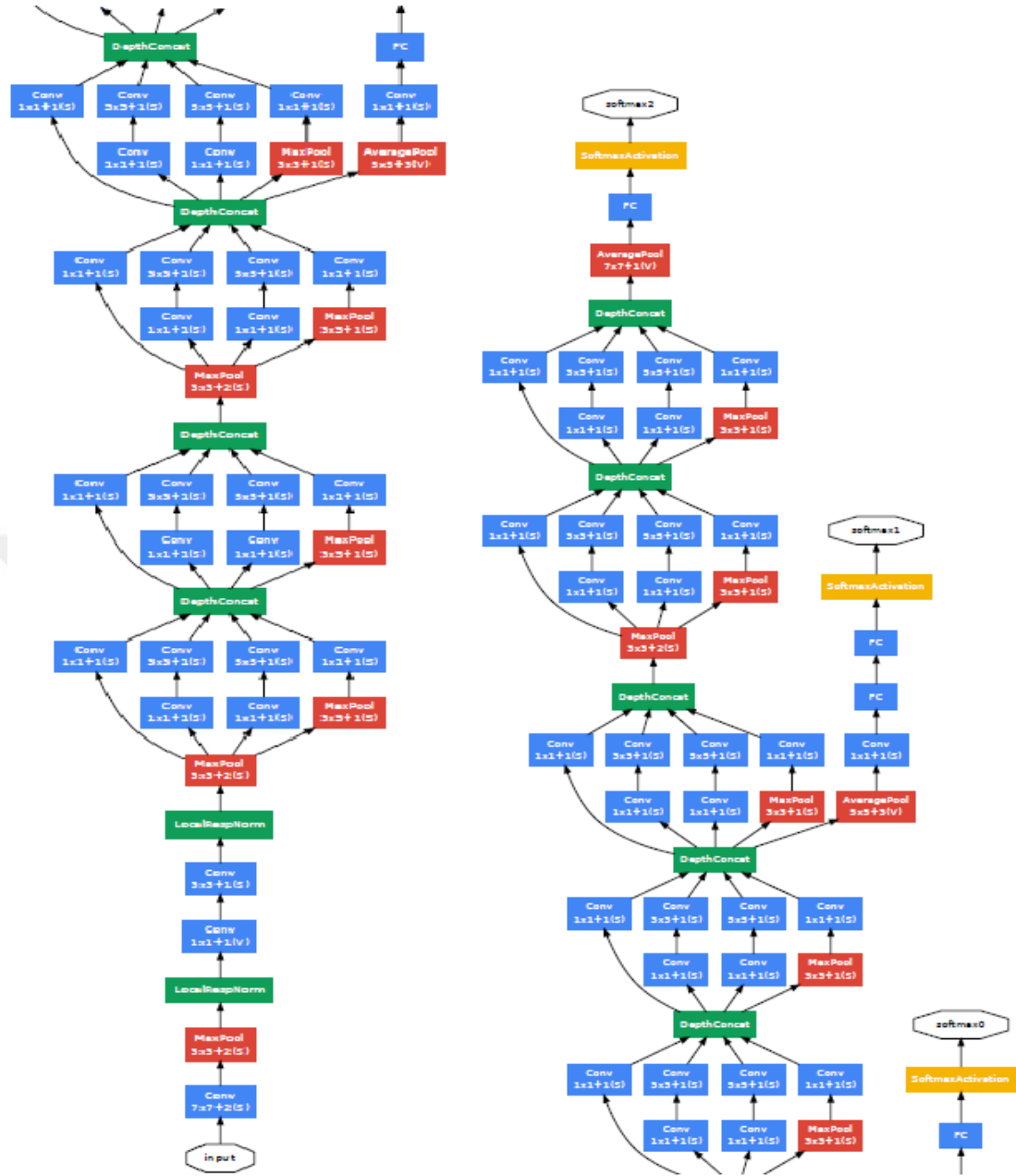
ImageNet yarışmasının 2013 yılında kazananı ZF Net mimarisi olmuştur. Bu mimari Matthew Zeiler ve Rob Fergus [109] tarafından tasarlanmıştır. Geliştirilen bu mimari ile ImageNet yarışmasında nesne tanıma hata oranı %11,2'ye düşürülmüştür. AlexNet'in geliştirilmiş hali olup, farklı olarak 7x7 boyutlu çekirdek çerçeveleri kullanılmış ve havuzlama katmanında kayma miktarı 2 olarak tasarlanmıştır. Aktivasyon katmanı olarak ReLU, hata fonksiyonu olarak Cross-Entropy Loss kullanılmıştır. ZF Net mimarisi Şekil 3.17.'de gösterilmiştir.



Şekil 3.17. ZF Net mimarisi [109].

3.3.3.3. GoogleNet

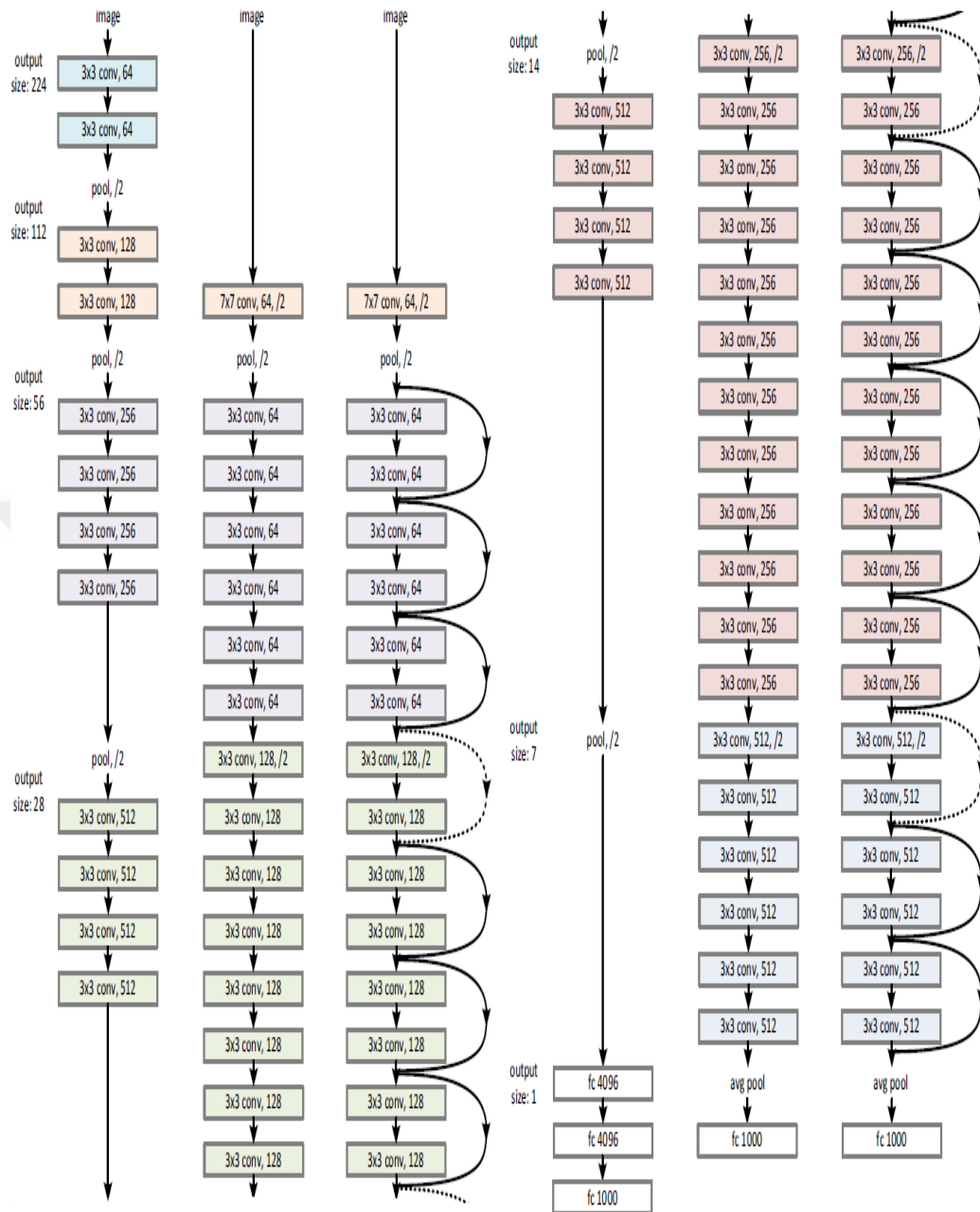
ImageNet yarışmasının 2014 yılındaki kazananı GoogleNet mimarisi olmuştur. Bu yapı Szegedy ve ark. [110] tarafından tasarlanmış 22 katmanlı bir mimariye sahiptir. Bu mimari ile nesne tanıma hata oranını %5,7'ye kadar düşürmüştür. Bu mimari incelendiğinde konvolüsyon ve havuzlama katmanlarının ardışık olarak kullanılmasının yanı sıra, farklı olarak paralel yapılar da gözlemlenmektedir. GoogleNet mimarisi Şekil 3.18.'de verilmiştir [110].



Şekil 3.18. GoogleNet mimarisi [110].

3.3.3.4. Microsoft resNet

ImageNet yarışmasının 2015 yılındaki kazananı, Zhang ve ark. [111] tarafından geliştirilen Microsoft ResNet mimarisi olmuştur. Bu mimarinin %3,6 hata oranı ile nesne tanımadaki başarısı, insan kabiliyetinin de üstüne çıkmıştır. İnsanların nesne tanımadaki hata oranı %5-%10 arasındadır. 152 katmandan oluşan Microsoft Resnet mimarisi Şekil 3.19.'da gösterilmiştir [112].



Şekil 3.19. Microsoft ResNet mimarisi [112].

BÖLÜM 4. İLGİ TESPİT SİSTEMİNİN UYGULANMASI

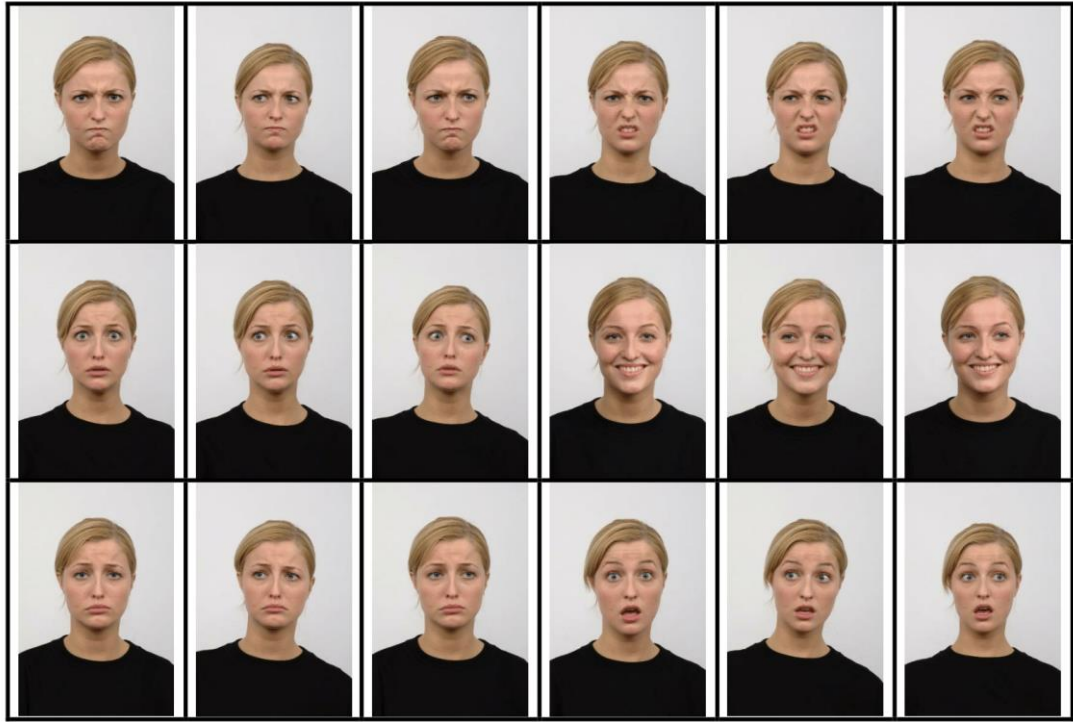
4.1. Uygulamada Kullanılan Veritabanı

Bu tez çalışmasının uygulama bölümü üç öğrenme aşamasından oluşur. Bu aşamalar:

1. Baş yönelimi tespiti
2. Yüz bölütleme
3. Yüz ifadesi tespitidir.

Tüm aşamalarda eğitim ve test işlemleri için Radboud Face Database (RaFD) [113] kullanılmıştır. RaFD ücretsiz bir yüz veri tabanıdır. Veri tabanı 49 insana ait yüz görüntüleri içermektedir. Bu kişilerin 39'u yetişkin ve 10'u çocuk olmak üzere; 24'ü erkek 25'i kadındır. Veritabanı içerisinde 5 farklı bakış açısına (0, 45, 90, 135, 180 derece) sahip 5 kamera kullanılarak elde edilen yüz görüntüleri barındırır. Ayrıca veri tabanında her kişiye ait 8 farklı yüz ifadesi (nötr, sinirlilik, üzünlük, korku, iğrenmişlik, şaşkınlık, mutluluk ve aşağılama) bulunur. Tüm fotoğraflar Facial Action Coding System (FACS) [114] örneklerini temel alarak hazırlanmıştır. Veritabanından, bir kişiye ait bakış yönü ve yüz ifadelerini içeren örnek, Şekil 4.1.'de gösterilmiştir. Bu veri tabanının içeriği sayesinde araştırmacılar duygu analizi, bakış yönü tespiti veya baş yönelimi tespiti konularında çalışmalar yapabilmektedir.

Sistemin başarımını test etmek için ayrıca Karolinska Directed Emotional Face (KDEF) [115] veritabanı kullanılmıştır. KDEF hem baş açısı etiketleri hem de yüz ifadesi etiketleri içerdiğinden, sistem performansını ölçmek için uygun bir veritabanıdır. KDEF her modelin 5 açılı 7 yüz ifadeli görüntüsü olmak üzere toplam 4900 yüz görüntüsü içerir. Araştırmacıların ücretsiz olarak kullanabildiği bir veritabanıdır. Veritabanından, bir kişiye ait yüz ifadelerini içeren örnek Şekil 4.2.'de gösterilmiştir.



Şekil 4.1. Kullanılan RaFD veritabanında bulunan bir kişiye ait tüm yüz ifadeleri.



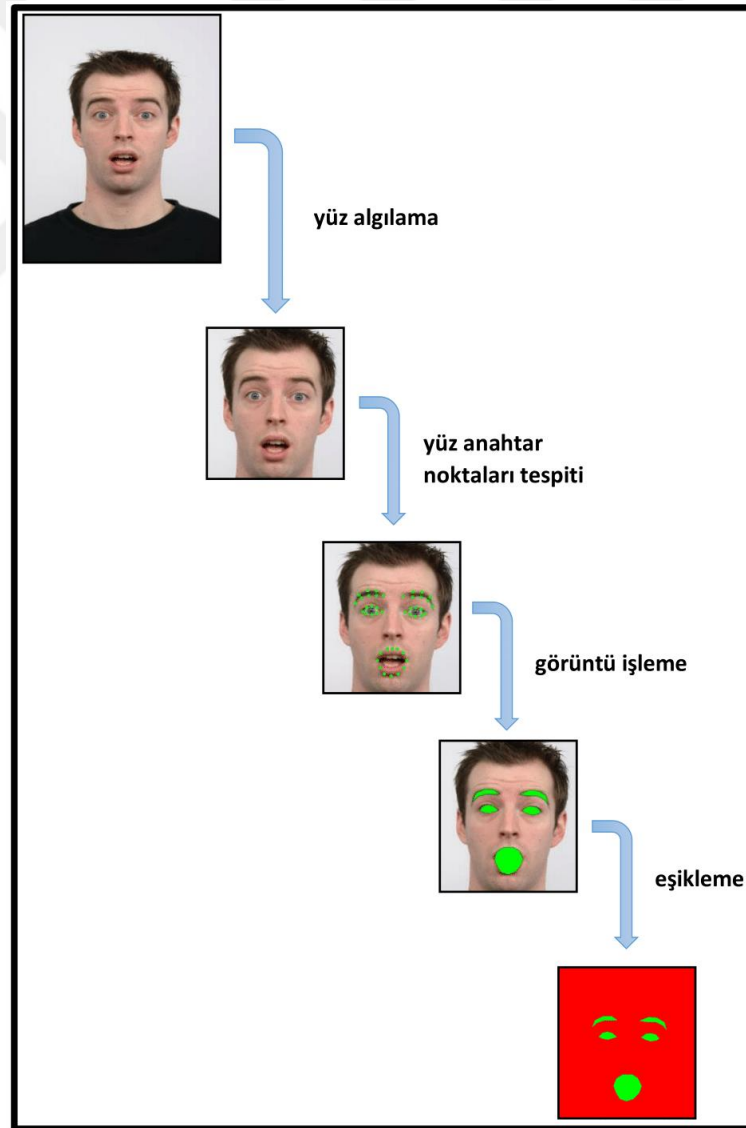
Şekil 4.2. Kullanılan KDEF veritabanında bulunan bir kişiye ait tüm yüz ifadeleri.

Yüz bölütleme işlemi için gerekli olan eğitim maskelerinin oluşturulması işlemi aşağıda açıklanmıştır.

1. Eğitim setindeki her bir görüntü, ilk olarak sadece yüz bölgesine odaklanılacak şekilde kesilir.

2. Ardından Face++ [116] aracı kullanılarak yüz anahtar noktaları tespit edilir. Bu araç insan yüzündeki 83 adet yüz anahtar noktasını tespit edebilmektedir. Ancak bu çalışmada kullanılan eğitim maskelerinde sadece kaşlar, gözler ve ağız bölgelerine odaklanılacağından, bu bölgeleri çevreleyen 45 adet yüz anahtar noktası kullanılmıştır.
3. Yüz noktaları kapalı bir şekil oluşturacak şekilde birbirine bağlanır.
4. 3. maddede oluşturulan kapalı şekillerin içi renkle doldurulur.
5. Son olarak eşikleme uygulanarak kaş göz ve ağız bölgeleri bir sınıf, geri kalan bölgeler ise farklı bir sınıf olacak şekilde maskeler oluşturulur.

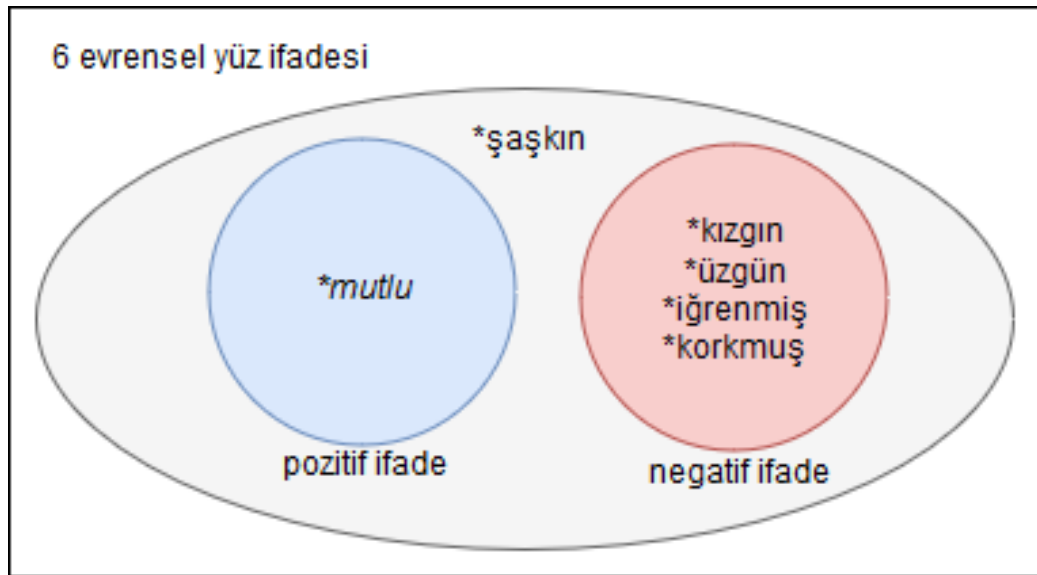
Bu işlem ayrıca Şekil 4.3.'te gösterilmiştir.



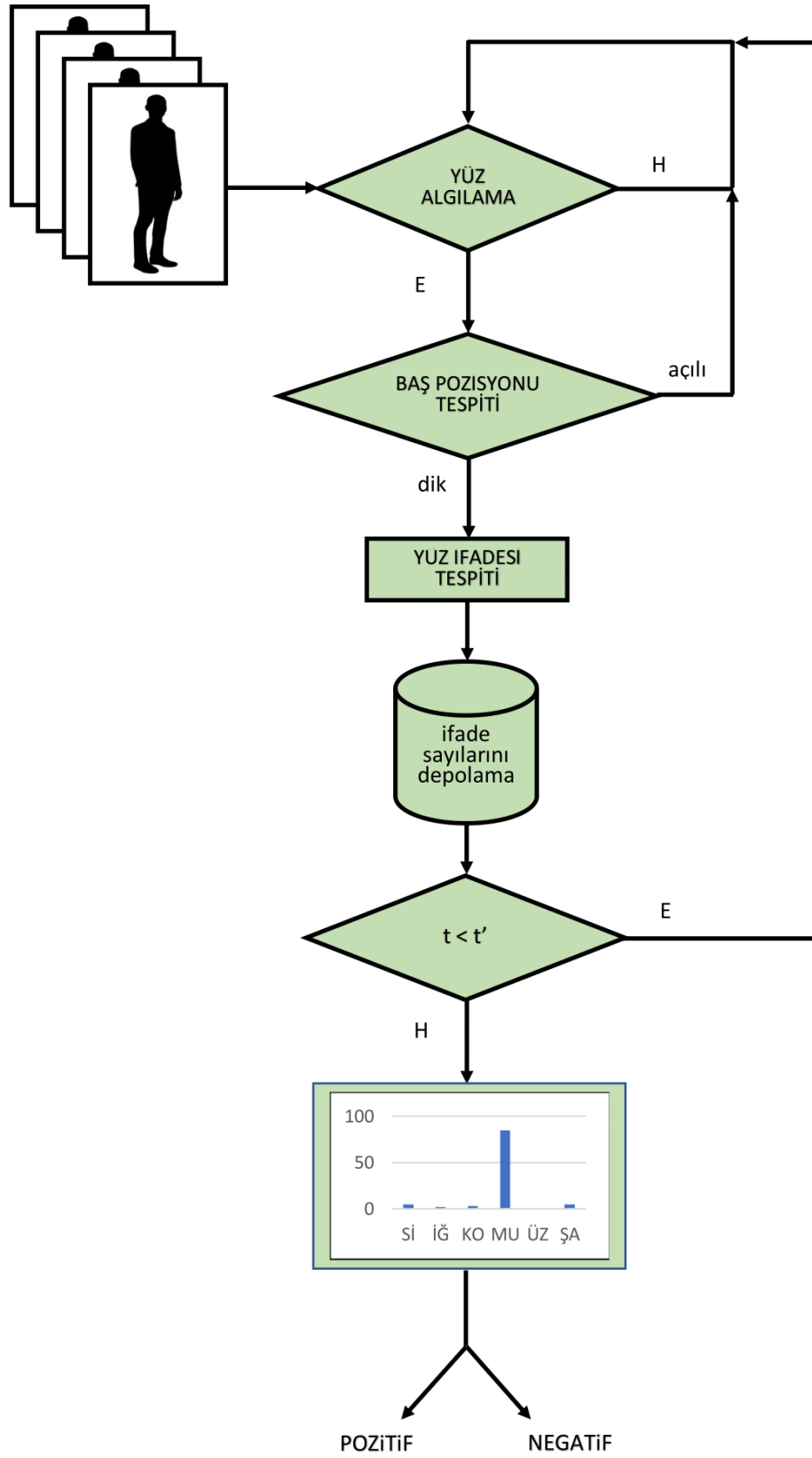
Şekil 4.3. Yüz bölütleme işlemi için gereken maskelerin oluşturulması.

4.2. Geliştirilen İlgil Tespit Sisteminin Mimarisi

Müşteri ilgisinin incelenmesi amacıyla geliştirilen bu sistemde ilk olarak kameranın görüş açısına giren müşteri yüzleri tespit edilir, ardından tespit edilen yüzlerin ilgili ürüne ya da reklama bakıp bakmadığını belirlemek üzere baş yönelimi analiz edilir. Bu aşamada, kameranın ilgili ürün ya da reklamın üzerinde olduğu varsayılarak işlem yapılmıştır. Böylece kişi ilgili ürün ya da reklama bakıyorsa yüzü kameraya dönüktür. Bu nedenle baş yönelimi tespiti için ön yüz veya ön yüz değil şeklinde ikili sınıflandırma yapılması yeterlidir. Eğer baş yönelimi kameraya doğru değilse müşterinin görsel ilgi odağı, ilgili ürün ya da reklamda değildir. Bu durumda sonraki aşamalara devam edilmeyip yeniden yüz tespiti aşamasına geçilir. Eğer müşterinin baş yönelimi kameraya doğruysa yüz bölütleme işlemi ve yüz ifadesi tespiti aşamasına geçilir. Bu şekilde belirli bir süre boyunca müşterinin yüz ifadeleri analiz edilerek depolanır. Ardından bu süre boyunca müşterinin gösterdiği duygular pozitif ve negatif duygular olarak sınıflandırılır ve müşterinin duygusal modu hakkında bilgi verilir. Pozitif ve negatif duyguların sınıflandırılmasında Ekman ve ark. [26] çalışması referans alınmıştır. Bu duygular Şekil 4.4.'te şematik olarak gösterilmiştir. Önerilen sistemi özetleyen akış diyagramı ise Şekil 4.5.'te gösterilmiştir.



Şekil 4.4. Evrensel yüz ifadelerinin pozitif ve negatif duygulara göre sınıflandırılma şeması.



Şekil 4.5. Önerilen müşteri davranış izleme sisteminin akış şeması.

Önerilen sistem müşteri davranışlarını incelemek için üç öğrenme aşaması gerçekleştirir. Öncelikle ilk ağ ile (CNN-1) müşterinin baş yönelimi tespit edilir ve ilgili ürüne ya da reklama bakan müşterilere odaklanır. İlgili ürün ya da reklama bakmayan yüzler ile ilgilenilmez ve sonraki aşamalara devam edilmez. İkinci ağ (CNN-2) müşterinin yüz bölgesini bölütler. Bu aşamada bölütlemedeki amaç, yüz ifadesi tespitinde önemli rol oynayan ağız, göz ve kaş bölgelerini tespit etmektir. Böylece bölütlenmiş özet bir yüz görüntüsü oluşturulur. Üçüncü CNN (CNN-3) de bu bölütlenmiş yüz görüntüsünü oluşturan ağın tam bağlantılı katmanına ait olasılıksal değerler ile ham yüz görüntüsünü birlikte giriş verisi olarak kullanarak yüz ifadelerini sınıflandırır. Önerilen sistemin mimarisi Şekil 4.7.'de verilmiştir.

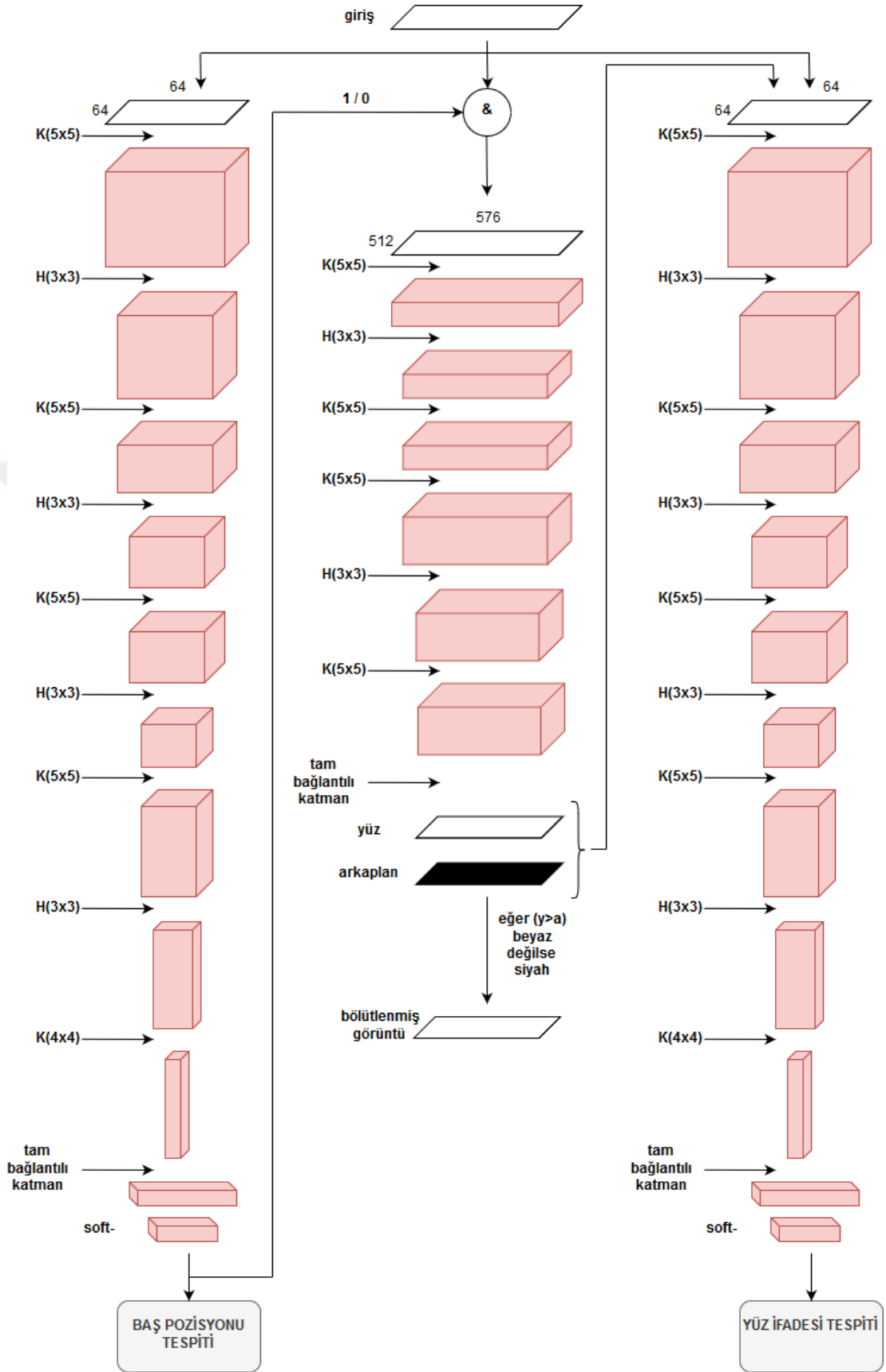
4.3. Baş Yönelimi İçin Geliştirilen CNN (CNN-1)

Müşterinin görsel ilgi odağının ilgili ürün ya da reklamda olup olmadığının belirlenmesi için baş yönelimi tespiti gerçekleştirilir. Önerilen sistem için, başın karşıya bakıp bakmadığının tespiti yeterlidir ve bu nedenle iki sınıflı baş pozisyonu tespiti yapılabilir. Ancak kullanılan veritabanında 5 açılı olarak baş yönelimleri yer aldığından, sistemde 5 sınıflı baş yönelimi tespiti yapılmıştır. Veritabanında yer alan baş açılarından örnekler Şekil 4.6.'da gösterilmiştir.



Şekil 4.6. RaFD'de bulunan baş yönelimi görüntülerinden örnekler [113].

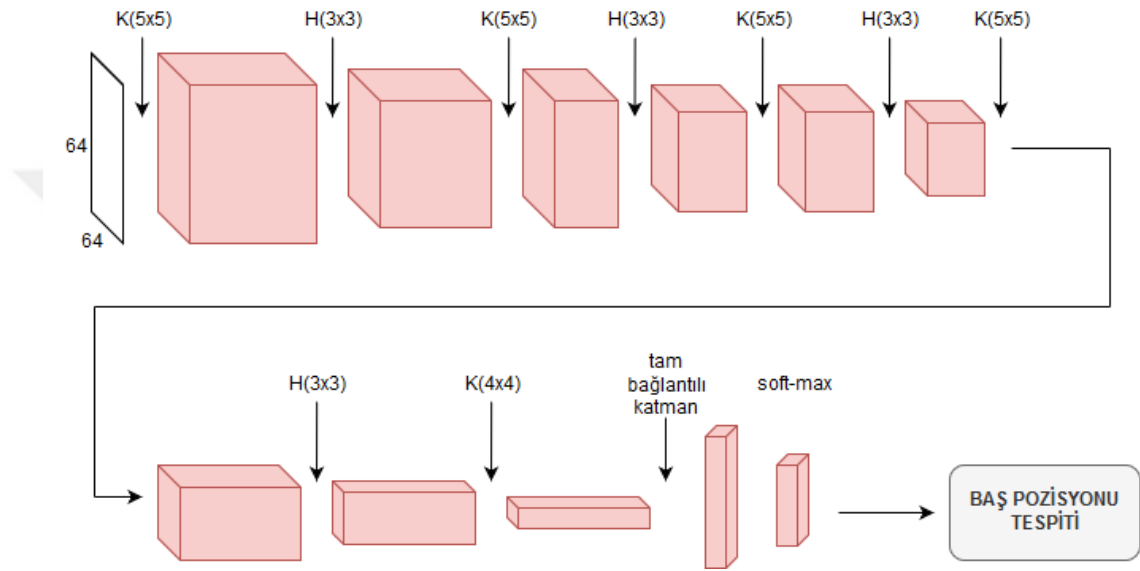
CNN-1 'yaw' hareketi açısından kafa poz açısını (0, 45, 90, 135, 180 derece) tahmin etmek üzere eğitilmiştir. CNN-1 mimarisi 64x64 piksel olarak yeniden boyutlandırılmış bir tam yüz görüntüsü üzerinde çalışır. CNN-1 yapısı 10 katman içerir:



Şekil 4.7. Önerilen sistemin CNN mimarisi.

1. 5 adet konvolüsyon katmanı (64 5x5'lik 1 katman, 32 5x5'i olan 2 katman, 64 5x5'i olan 1 katman ve 64 4x4 filtreli 1 katman),
2. 4 adet havuzlama katmanı,
3. 1 adet tam bağlantılı katman.

CNN-1'in mimarisi Şekil 4.8.'de gösterilmiştir.



Şekil 4.8. Baş yönelimi tespiti için geliştirilen CNN (CNN-1) mimarisi.

4.4. Yüz Bölütleme İçin Geliştirilen CNN (CNN-2)

CNN-2'nin amacı, görüntünün ağız, göz ve kaş bölgelerini yüzün diğer bölümlerinden ayırmaktır. Yüz bölütleme, ikili bir sınıflandırma problemi olarak tasarlanmıştır. Buradaki sınıflar ise; yüz bileşeni (ağız, göz ve kaşlar) bir sınıf, geri kalan yüz bölgeleri (arka plan) ise bir sınıf olarak değerlendirilmiştir.

Eğitim aşamasından önce, orijinal ham görüntüler ve karşılık gelen eğitim maskeleri 16x16 boyutunda örtüşmeyen bloklara ayrılır. %80 veya daha fazla oranda yüz bileşen pikseli içeren bloklara yüz bileşeni etiket olarak atanır. %80 veya daha fazla oranda arka plan pikseline sahip olan bloklara ise arkaplan etiketi atanır. Kalan karışık sınıf blokları ise eğitimde başarıyı düşürmemesi amacıyla eğitim sırasında göz ardı edilir.

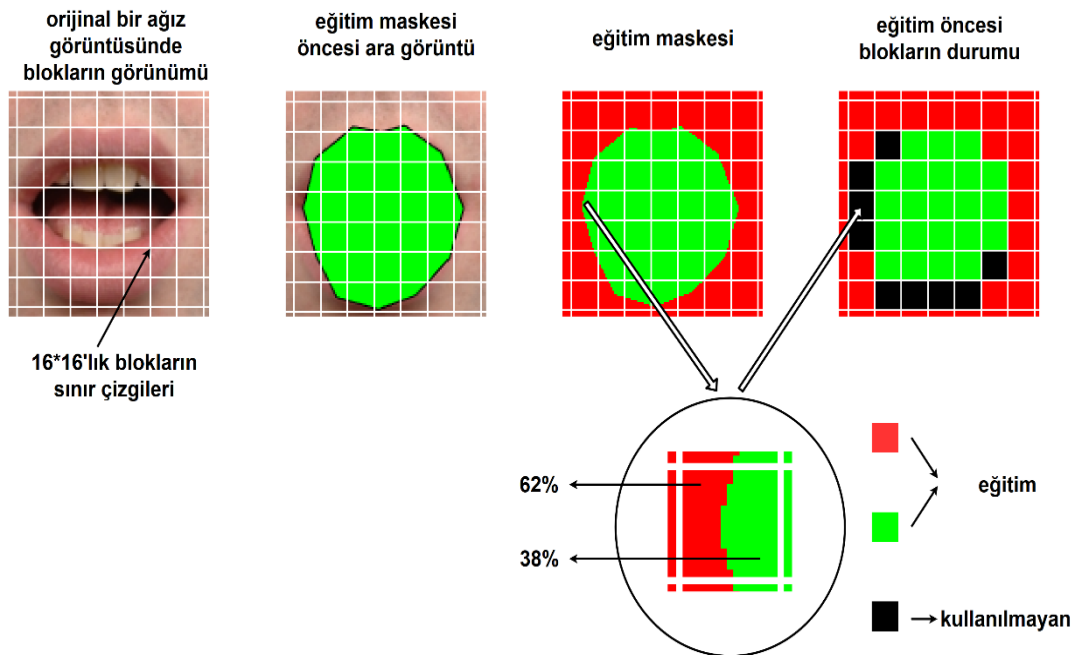
Yukarıda bahsedilen %80 eşik değeri ise deneysel çalışmalarımıza göre belirlenmiştir. Bu işlemler Denklem 4.1 ve Denklem 4.2’de de belirtilmiştir.

$$Eger \frac{YPS}{TPS} \geq 80\% \rightarrow blok_yuz_bilesenidir \quad (4.1)$$

$$Eger \frac{KPS}{TPS} \geq 80\% \rightarrow blok_arkaplandir \quad (4.2)$$

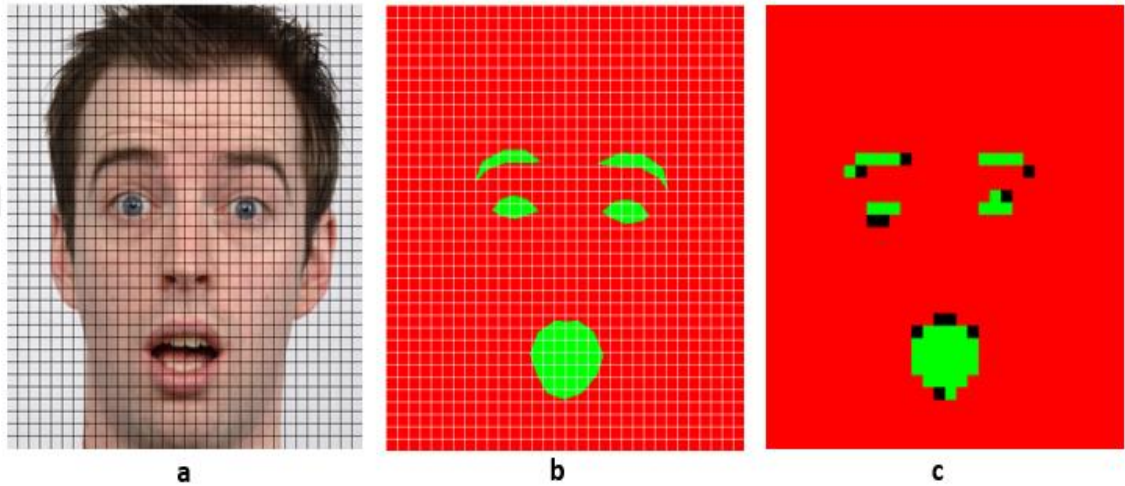
Denklem 4.1 ve Denklem 4.2’de *YPS* bloğa ait yeşil piksel sayısını, *TPS* bloğa ait toplam piksel sayısını, *KPS* ise bloğa ait kırmızı piksel sayısını ifade etmektedir.

Şekil 4.9.’da bir yüz görüntüsünün ağız bölgesine ait blok durumlarının belirlenmesi örnek olarak verilmiştir. İlk olarak orjinal görüntü ve eğitim maskesi 16x16 piksellik bloklara ayrılır. Eğitim maskesinde her bloktaki piksellerin dağılımlarına bakılır. Şekilde örnek olarak alınan blok %62 oranında kırmızı piksel, %38 oranında yeşil piksel içermektedir. Bu bloktaki toplam kırmızı veya toplam yeşil piksel sayıları %80 eşik değerini aşamadıkları için bu blok ve bu bloğa denk gelen orijinal görüntüdeki blok eğitime dahil edilmez.



Şekil 4.9. Blok durumlarının belirlenmesi.

Şekil 4.10., bir görüntünün bloklarına etiket atama adımlarını göstermektedir. Şekil 4.10.'da a görüntüsü orijinal ham görüntü üzerindeki blokları, b görüntüsü eğitim maskesinin bloklarını, c görüntüsü ise b görüntüsü kullanılarak elde edilen etiketlenmiş blokları göstermektedir; c görüntüsündeki siyah renkli bloklar karışık sınıf bloklarını içerir. Yukarıda bahsedildiği gibi siyah piksellerin bulunduğu bloklar eğitime alınmaz.



Şekil 4.10. Yüz bölütleme için eğitim ve test görüntülerinin bloklara ayrılması.

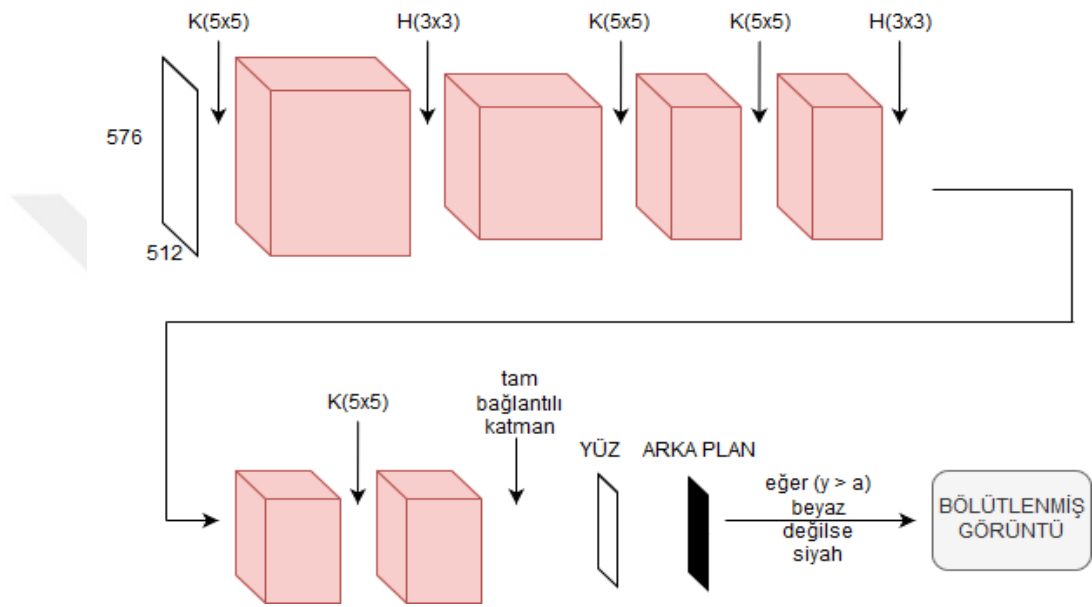
CNN-2'ye ait tam bağlantılı katmanın çıkışı iki kanaldır ve bu kanallardan biri yüz bileşenleri olasılıksal değerlerini, diğer kanal ise arka plan olasılıksal değerlerini içerir. Tam bağlantılı katmandaki olasılıksal değerlere göre bölütlenmiş özet yüz görüntüleri elde edilir. Bir piksel için hangi sınıfa ait olasılıksal değer daha yüksekse, sonuç görüntüsündeki ilgili piksel değeri o sınıf kabul edilir. Sonuç görüntüsündeki bir piksele, piksel arkaplana ait ise kırmızı, yüz bileşeni sınıfına ait ise yeşil renkli olarak değer atanır.

Ayrıca tam bağlantılı katmandaki iki kanalın olasılıksal değerleri, yüz ifadesi için geliştirilen CNN-3'e aktarılır. Böylece kılavuzlu görüntü eğitimi ve daha güçlü yüz ifadesi tespiti gerçekleştirilmiş olur. CNN-2 nin test işlemi ise [117]'de açıklanan sürgülü pencere yaklaşımı (sliding window) kullanılarak gerçekleştirilir.

CNN-2 mimarisi 7 katman içerir. Bunlar:

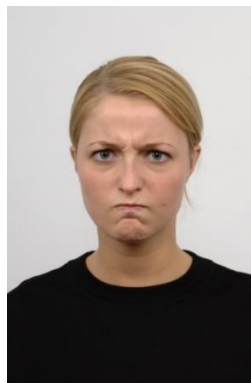
1. 4 adet konvolüsyon katmanı (16 5x5x3'lük 1 katman, 16 5x5'lik 1 katman, 32 5x5'lik 1 katman ve 32 4x4 filtreli 1 katman),
2. 2 adet havuzlama katmanı,
3. 1 adet tam bağlantılı katman.

CNN-2'nin mimarisi Şekil 4.11'de gösterilmiştir.



Şekil 4.11. Yüz bölütleme için geliştirilen CNN (CNN-2) mimarisi.

CNN-2'nin eğitimi sonucunda, ilk konvolüsyon katmanında öğrenilen filtreler ile Şekil 4.12.'de gösterilen orijinal görüntünün konvolüsyona tabi tutulması sonucunda oluşan çıkış görüntüleri Şekil 4.13.'te gösterilmiştir.



Şekil 4.12. Orijinal yüz görüntüsü.



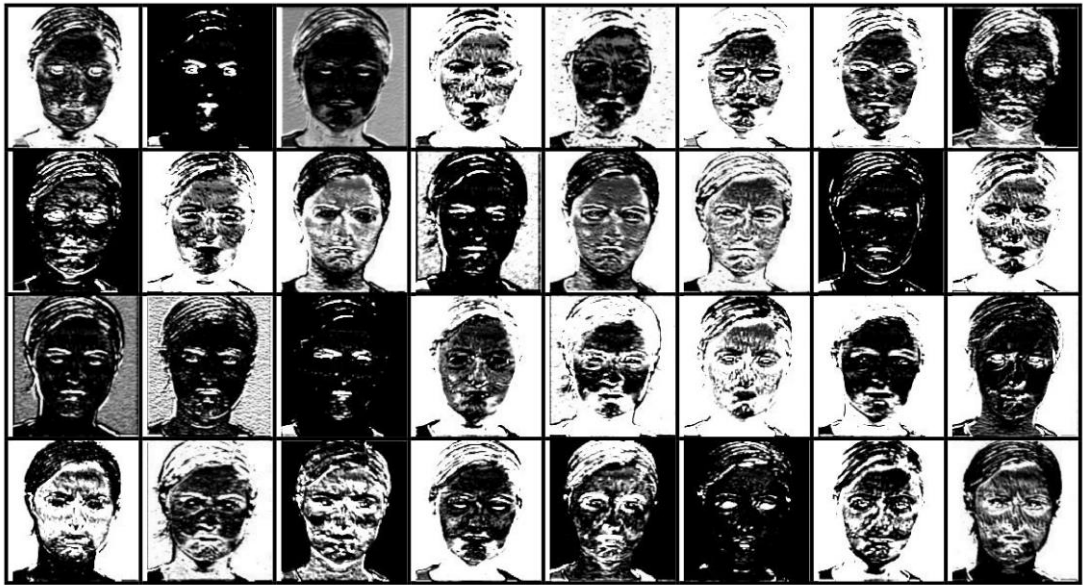
Şekil 4.13.CNN-2'ye ait ilk konvolüsyon katmanı sonuçları.

Şekil 4.13. incelendiğinde siyah ve beyaz arası gri ton renk geçişlerinin piksellerde yer aldığı görülmektedir. Bu dağılımda beyaz pikseller çekirdek çerçevelerinin öğrendiği özniteliklerin ön plana çıkarıldığı yerlerdir. Örnek olarak şekildeki son satırda bulunan ilk iki görüntüye ait çekirdek çerçeveleri göz ve kaş özniteliklerini öğrenmeye başlamıştır. Şekil 4.12.'nin giriş olarak kullanıldığı ağda, ikinci konvolüsyon katmanında öğrenilen filtrelerin bir önceki katmandan gelen girdiler ile konvolüsyona tabi tutulması sonucunda oluşan çıkış görüntüleri de Şekil 4.14.'te gösterilmiştir.



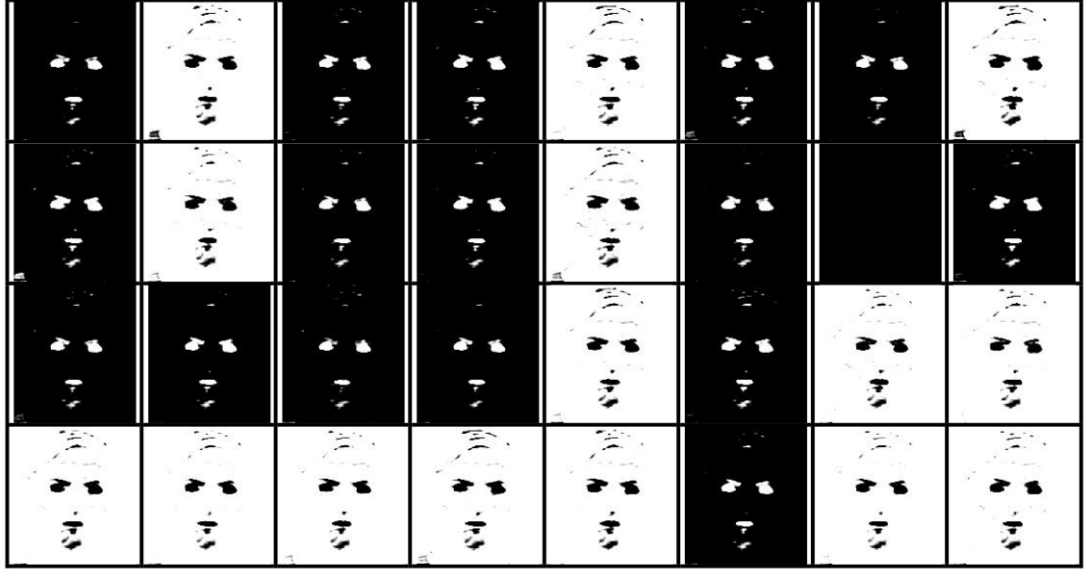
Şekil 4.14. CNN-2 ikinci konvolüsyon katmanı çıktıları.

Üçüncü konvolüsyon katmanında öğrenilen çekirdek çerçeveleri bir önceki katmandan gelen girdilere uygulandığında ise Şekil 4.15.'teki görüntüler elde edilir.



Şekil 4.15. CNN-2 üçüncü konvolüsyon katmanı çıktıları.

Son konvolüsyon katmanı çıktıları ise Şekil 4.16.'da verilmiştir. Bu görüntülerden anlaşılacağı gibi son konvolüsyon katmanında ağ; gözler, kaşlar ve ağız bölgelerini öğrenmeye odaklanmıştır.



Şekil 4.16. CNN-2 son konvolüsyon katmanı çıktıları.

4.5. Yüz İfade Tespiti İçin Geliştirilen CNN (CNN-3)

Seri bağlı CNN mimarisinde, üçüncü ve son CNN yüz ifadelerini sınıflandırmak üzere eğitilmiştir. Yüz ifadesi sınıflandırılırken, giriş olarak CNN-2'ye ait tam bağlantılı katmandaki yüz bileşen katmanı ile arkaplan katmanındaki (2-kanal) olasılıksal değerler ile orijinal ham yüz görüntüsü (3-kanal) birlikte kullanılır. Bu şekilde CNN-2'ye ait değerlerin kullanılmasının avantajları aşağıda sıralanmıştır, bu avantajlar sayesinde daha yüksek başarıma sahip yüz ifadesi tespiti sistemi geliştirilmiştir.

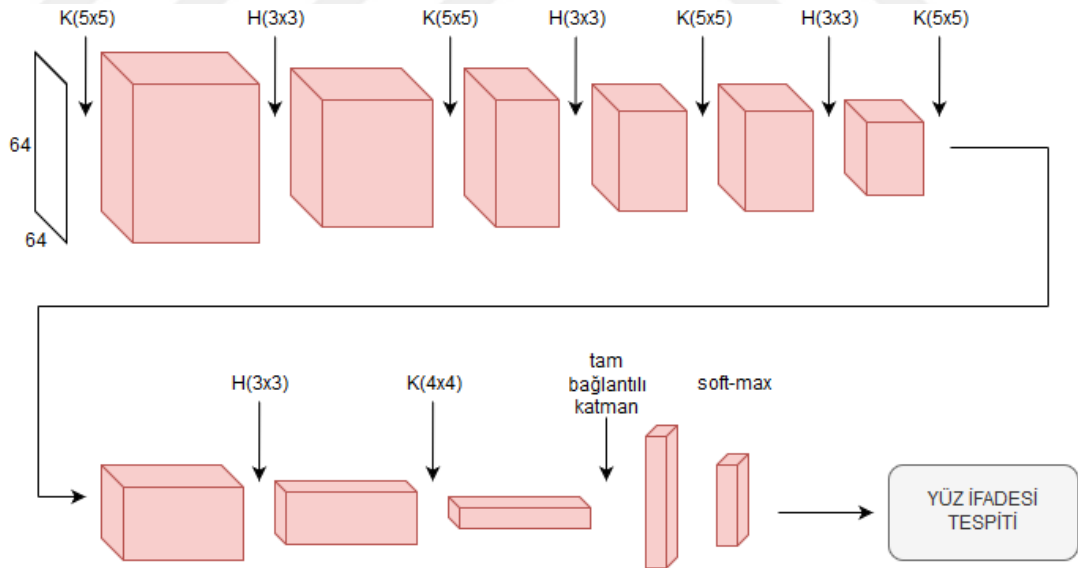
1. Kılavuzlu öğrenme: CNN-2'nin tam bağlantılı katmanına ait olasılıksal değerler kullanılarak, ağı daha ilk katmanlardan itibaren yüz ifadesi tespiti için önemli olan yüz bileşenlerini kullanması sağlanmıştır, bu kılavuzlu öğrenme (guided image training) sayesinde başarımlar artmıştır.
2. Parça tabanlı ve bütünsel yüz bilgilerinin birlikte kullanımı: CNN-2'den elde edilen parça tabanlı bilgiler ile birlikte CNN-3'te orijinal yüz görüntülerinin birlikte kullanımı sistemi daha güçlü kılmıştır.

3. Kişisel gizliliği sağlama: CNN-2'den elde edilen bölütlenmiş özet yüz görüntüsü, hiçbir kişisel bilgi içermemesine rağmen, yüz ifadesi ile ilgili önemli bilgiler içerir. Eğer yüz görüntülerinin arşivlenmesi gereken özel bir durum gerekirse, bu görüntüler arşivlenebilir. Böylece kişisel bilgiler gizlenirken, yüz ifadesi bilgileri arşivlenmiş olur. Bu durum, özellikle medikal uygulamalarda hastanın modunu gösterme biçimi kaydedilirken işe yarayabilir.

CNN-3 yapısı, CNN-1'e benzer şekilde 64x64 olarak yeniden boyutlandırılmış tam yüz görüntüsü üzerinde çalışır. 10 katman olarak tasarlanmıştır. Bunlar:

1. 5 adet konvolüsyon katmanı (64 5x5x3 boyutunda 1 katman, 32 5x5 ile 2 katman, 64 5x5 ile 1 katman ve 64 4x4 filtreli 1 katman),
2. 4 adet havuzlama katmanı,
3. 1 adet tam bağlantılı katman.

CNN-3'ün mimarisi Şekil 4.17.'de gösterilmiştir.



Şekil 4.17. Yüz ifade sınıflandırması için kullanılan CNN (CNN-3) mimarisi.

Tablo 4.1., önerilen sisteme ait CNN-1, CNN-2, CNN-3 yapılarının detaylı katman bilgisini göstermektedir.

Tablo 4.1. Önerilen CNN sisteminin detaylı katman bilgileri

Mimari	Katman	Çekirdek boyutu	Filtre sayısı	Çıkış boyutu
CNN-1 (Baş yönelimi tespiti için giriş: 64x64)	Konv1	5x5	64	64x64x64
	Konv2	5x5	32	32x32x32
	Konv3	5x5	32	16x16x32
	Konv4	5x5	64	8x8x64
	Konv5	4x4	64	1x1x64
CNN-2 (Yüz bölütleme için giriş: 16x16)	Konv1	5x5	16	16x16x16
	Konv2	5x5	16	8x8x16
	Konv3	5x5	32	4x4x32
	Konv4	4x4	32	1x1x32
CNN-3 (Yüz ifadesi tespiti için giriş: 64x64)	Konv1	5x5	64	64x64x64
	Konv2	5x5	32	32x32x32
	Konv3	5x5	32	16x16x32
	Konv4	5x5	64	8x8x64
	Konv5	4x4	64	1x1x64

4.6. Deneysel Sonuçlar

Yukarıda anlatılan CNN mimarisi MatConvNet kütüphanesi kullanılarak geliştirilmiştir. Sistemin baş yönelimi tespiti, yüz bölütleme ve yüz ifadesi tespiti ile ilgili analiz sonuçları bu bölümde anlatılmıştır.

Baş yönelimi tespiti için RaFD'ye ait 8040 görüntü, eğitim ve test seti olmak üzere 2 eşit parçaya bölünmüştür. Baş yönelimi tespitinde, sistem orijinal yüz görüntüleri kullanılarak eğitilmiştir. Sistemin sadece baş kameraya dönük ya da değil şeklinde sınıflandırılması yeterli olmasına rağmen, veritabanında 5 açılı baş yönelimi görüntüleri yer aldığından, 5 açılı sınıflandırma yapılmıştır.

Baş yönelimi tespitinin RaFD için sonuçları Tablo 4.2.'de verilmiştir. Tablodaki 90 derecelik açı başın kameraya baktığı andır. 0 derece basın tam sağa, 180 derece ise

başın tam olarak sola dönük olduğu anları temsil etmektedir. Tablodan da görülebileceği gibi önerilen yöntem 5 sınıflı baş yönelimini %99,9 oranında doğru olarak tespit etmektedir. Daha da dikkat çekici olan ise bu tezin konusu gereği önemli olan başın kameraya baktığı pozisyonlar %100 oranında doğru olarak tespit edilmiştir. Kameranın ilgi analizi yapılan ürün ya da reklam üzerinde olduğu varsayıldığından müşterinin görsel ilgi odağı kameraya doğru olduğunda, sistem diğer adımlara geçebilmektedir.

Tablo 4.2. RaFD için baş yönelimi tespiti başarımları (%)

Gerçek \ Tahmin					
	0	45	90	135	180
0	99,88	0,012	0	0	0
45	0	100	0	0	0
90	0	0	100	0	0
135	0	0	0	100	0
180	0	0	0	0,037	99,63
Ortalama Başarımları: % 99,90					

Baş yönelimi için KDEF'e ait 4900 görüntü kullanılmıştır. Veri sayısı az olduğundan bu veritabanına ait görüntüler %90 eğitim, %10 test olacak şekilde bölünmüştür. KDEF için baş yönelimine ait karmaşıklık matrisi Tablo 4.3.'te verilmiştir. Tablodan da görülebileceği gibi başın tam kameraya karşı olduğu (90 derecelik) açı yine %100 oranında doğru tespit edilmiştir. KDEF için ortalama baş yönelimi başarımları ise %99,39'dur.

Yüz bölütleme için RaFD'ye ait 1206 görüntülük set, eğitim ve test seti olmak üzere 2'ye bölünmüştür. Önerilen sistem, yüz bölütleme aşamasında direk görüntüleri kullanmak yerine bu görüntüleri 16x16 bloklara ayırarak eğitim yapar. Blokların her birine etiket atama işlemleri yukarıda anlatılmıştır. Bu şekilde toplam 691.200 blok elde edilmiştir. Bu blokların 6.035 tanesi yukarıda anlatılan %80 eşik değerini geçemediği için karma sınıf olarak değerlendirilmiş ve gürbüz bilgileri içermediğinden

eğitime dahil edilmemiştir. Blokların 14.935 tanesi yüz bileşen sınıfı ve 670.230 tanesi ise arka plan sınıfı olarak etiketlenmiştir. Tablo 4.4. toplam blok sayısı ile eğitimde kullanılan ve kullanılmayan blok sayılarını göstermektedir.

Tablo 4.3. KDEF için baş yönelimi sonuçları

Gerçek \ Tahmin	Tahmin				
	0	45	90	135	180
0	97,96	2,04	0	0	0
45	0	100	0	0	0
90	0	0	100	0	0
135	0	0	0	100	0
180	0	0	0	1,02	98,98
Ortalama Başarım: % 99,39					

Tablo 4.4. Yüz bölütleme aşamasında eğitimde kullanılan blok sayıları

Sınıf tipi	Blok sayısı
Yüz bileşeni sınıfı	14.935
Arka-plan sınıfı	670.230
Karma sınıf	6.035
Toplam blok	691.200

Yüz ifadesi tespiti için RaFD'ye ait 1206 ön yüz görüntüsü kullanılmıştır. Bu görüntü seti; eğitim ve test seti için iki eşit parçaya bölünmüştür. Bu dağılımda eğitim ve test setinde aynı kişiye ait aynı yüz ifadeli ya da farklı yüz ifadeli yüz görüntüleri bulunmamaktadır.

Yüz ifadesi tespiti aşamasında CNN-3, 5 kanallı giriş verisi ile eğitilmiştir. Bu kanalların 3'ünü orijinal yüz görüntüsünün, red (kırmızı), green (yeşil), blue (mavi) (RGB) kanalları oluşturur; son 2 kanal ise CNN-2'den elde edilen tam bağlantılı katmanın olasılıksal değerleridir. Tablo 4.5. RaFD için yüz ifadesi tespiti sonuçlarının karmaşıklık matrisini gösterir. Karmaşıklık matrisinden görülebileceği gibi en yüksek sonucu %100 başarımla oranı ile şaşkınlık duygusu, en düşük sonucu ise %86,87

başarım oranı ile korku duygusu vermiştir. Kızgınlık duygusu %97,98 başarım oranında, iğrenme %96,97 oranında, üzgünlük ise %90,91 oranında doğru tespit edilmiştir. Bu tabloda kırmızı ile işaretli hücreler negatif duyguları, yeşil ile işaretli hücreler pozitif duyguları gösterir. Bu karmaşıklık matrisinden pozitif ve negatif duyguların başarım oranları yeniden hesaplandığında, pozitif duygular %94,95 oranında doğru tespit edilmişken, negatif duygular ise %93,18 oranında doğru tespit edilmiştir.

Tablo 4.5. 5 kanallı giriş ile eğitilen yüz ifadesi tespiti sisteminin RaFD için başarım oranları (%)

Tahmin \ Gerçek	Kızgınlık	İğrenmişlik	Korku	Mutluluk	Üzgünlük	Şaşkınlık
Kızgınlık	97,98	0	0	0	2,02	0
İğrenmişlik	1,01	96,97	0	0	0	2,02
Korku	0	0	86,87	0	11,11	2,02
Mutluluk	0	4,04	0	94,95	1,01	0
Üzgünlük	4,04	2,02	3,03	0	90,91	0
Şaşkınlık	0	0	0	0	0	100
Ortalama Başarım: % 94,61						
Olumlu Duygular		Olumsuz Duygular				
94,95		93,18				

Önerilen 5 kanal girişli yüz ifadesi sisteminin etkilerini daha iyi gösterebilmek için, tasarlanan yüz ifadesi tespiti CNN yapısı (CNN-3) RaFD'ye ait;

- sadece orijinal ham yüz görüntüleri ile,
- sadece CNN-2'den elde edilen bölütlenmiş özet yüz görüntüleri ile ve
- CNN-2'den elde edilen özet yüz görüntüleri ile birlikte ham görüntülerin kombinlenmesi ile eğitilmiştir.

Bu eğitimlerin test edilmesi ile elde edilen karmaşıklık matrisleri sırasıyla Tablo 4.6., Tablo 4.7. ve Tablo 4.8.'de sunulmuştur. Tablolar analiz edildiğinde görülebileceği gibi, orijinal yüz görüntüleri ile eğitim sonucunda kızgınlık duygusu %88,89; iğrenme

duygusu %90,91; korku duygusu %84,85; mutluluk duygusu %73,74; üzgünlük duygusu %69,70; şaşkınlık duygusu ise %91,92 oranında doğru tespit edilmiştir. Sistem sadece bölütlenmiş yüz görüntüleri ile eğitildiğinde ise başarımları kızgınlık için %98,67; iğrenme için %89,34; korku için %76, mutluluk için %97,34; üzgünlük için %76; şaşkınlık için %97,33'tür. Sistem özet yüz görüntüleri ile birlikte 3 kanallı ham görüntülerin kombinlenmiş hali ile eğitildiğinde ise genel başarımları biraz daha yükselmiştir (%93,43). Bu durumda kızgınlık duygusu %92,93; iğrenme duygusu %98,99; korku duygusu %87,88; mutluluk duygusu %92,93; üzgünlük duygusu %88,89; şaşkınlık duygusu ise %98,99 oranında doğru tespit edilmiştir. Ancak bahsedilen üç eğitimde de, önerilen 5 kanallı giriş ile elde edilen yüz ifadesi tespiti başarımları yakalanamamıştır.

Tablo 4.6. RaFD'ye ait sadece orijinal yüz görüntüleri ile CNN-3 sisteminin eğitilmesi sonucunda elde edilen başarımlar (%)

[illegible]

Tablo 4.7. Tablo 4.7. RaFD'ye ait sadece bölütlenmiş yüz görüntüleri ile CNN-3 sisteminin eğitilmesi sonucunda elde edilen başarımlar oranları (%)

[illegible]

Tablo 4.8. RaFD'ye ait özet yüz görüntüleri ve orijinal yüz görüntülerinin birlikte giriş olarak verilmesi ile eğitilen yüz ifadesi tespiti sisteminin başarımları (%)

Gerçek/Tahmin	Kızgınlık	İğrenmişlik	Korku	Mutluluk	Üzgünlük	Şaşkınlık
Kızgınlık	92,93	0	0	0	7,07	0
İğrenme	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
Ortalama Başarımlar: %93,43						

Ayrıca CNN-2; 6 sınıflı sınıflandırma yapmak yerine ifadelerin pozitif ve negatif olarak 2 sınıflı sınıflandırılması için RaFD'ye ait görüntüler ile yeniden eğitilmiştir. Bunun için veritabanındaki mutluluk duygusu pozitif sınıf; kızgınlık, iğrenmişlik, korkmuşluk ve üzüntü duyguları negatif sınıf olacak şekilde veritabanı yeniden düzenlenmiştir. Şaşkınlık ifadesi referans çalışmada olumlu ya da olumsuz bir duygu olarak sınıflandırılmadığından bu çalışmada da bu iki sınıfa dahil edilmemiştir. Yeni oluşturulan veritabanındaki 510 görüntü eğitim, 495 görüntü test seti olacak şekilde ayarlanmıştır. Bu eğitim sonucunda %92,48 genel başarımlar elde edilmiştir. Sonucun 6 sınıflı eğitime göre daha düşük sonuç vermesinin nedeni; negatif sınıftaki görüntüler bir çok duyguyu içerdiğinden, öğrenmenin spesifik olarak ifadelerin öğrenilmesine göre daha zor olması şeklinde yorumlanabilir. 2 sınıflı eğitim sonucunda Tablo 4.9.'da gösterilen karmaşıklık matrisi elde edilmiştir.

Tablo 4.9. Pozitif ve negatif duygular olarak gerçekleştirilen iki sınıflı eğitimin RaFD için karmaşıklık matrisi

Gerçek/Tahmin	Olumlu Duygular	Olumsuz Duygular
Olumlu Duygular	84,85	15,15
Olumsuz Duygular	0	100
Ortalama Başarımlar: 92,48		

Son olarak; CNN-3, KDEF veritabanına ait 5 kanallı veriler ile eğitilmiş ve test edilmiştir. Bunun için KDEF'e ait baş açısı karşıya doğru olan 840 yüz görüntüsü kullanılmıştır. Veri sayısı az olduğundan bu görüntülerin %90'ı eğitim, %10'u test verisi olacak şekilde düzenlenmiştir. Bu eğitimin test edilmesi ile Tablo 4.10.'daki karmaşıklık matrisi elde edilmiştir. Karmaşıklık matrisinden görülebileceği gibi en

yüksek sonucu %100 başarımlı oranı ile mutluluk ve şaşkınlık duygusu, en düşük sonucu ise %85,72 başarımlı oranı ile korku ve kızgınlık duygusu vermiştir. İğrenme duygusu %92,86 oranında doğru tespit edilmiştir. Bu tabloda kırmızı ile işaretli hücreler negatif duyguları, yeşil ile işaretli hücreler pozitif duyguları gösterir. Bu karmaşıklık matrisinden pozitif ve negatif duyguların başarımlı oranları yeniden hesaplandığında, pozitif duygular %100 oranında, negatif duygular ise %89,29 oranında doğru tespit edilmiştir.

Tablo 4.10. 5 kanallı giriş ile eğitilen yüz ifadesi tespiti sisteminin KDEF için başarımlı oranları (%)

Tahmin \ Gerçek	Kızgınlık	İğrenmişlik	Korku	Mutluluk	Üzgünlük	Şaşkınlık
Kızgınlık	85,72	7,14	0	0	7,14	0
İğrenmişlik	0	92,86	0	0	7,14	0
Korku	0	0	85,72	0	7,14	7,14
Mutluluk	0	0	0	100	0	0
Üzgünlük	7,14	0	0	0	92,86	0
Şaşkınlık	0	0	0	0	0	100
Ortalama Başarımlı: % 92,86						
Olumlu Duygular		Olumsuz Duygular				
100		89,29				

Sistemin verdiği hatalar analiz edildiğinde, sistemin başarımlı olduğunu düşüren önemli bir etkenin kullanılan veritabanının doğal duygulardan oluşmaması olduğu söylenebilir. Literatürde kullanılan bir çok veritabanı gibi RaFD ve KDEF de laboratuvar ortamında, kişilerin duyguları taklit etmesi sonucu çekilen fotoğrafların oluşturduğu veritabanlarıdır. Bu nedenle bu veritabanlarında yer alan bazı fotoğrafların ifade ettiği duygular hakkında insanın bile yanılma payı bulunabilir. Sistemin yanlış bulduğu bazı görsel örnekler Şekil 4.18.'de sunulmuştur. Sistemin gerçek duygular ile eğitilmesi ve test edilmesi ile başarımlı oranının daha da yükseleceği öngörülmektedir.

tahmin gerçek	kızgınlık	iğrenmişlik	korku	mutluluk	üzgünlük	şaşkınlık
kızgınlık	✓					
iğrenmişlik		✓				
korku			✓			
mutluluk				✓		
üzgünlük					✓	
şaşkınlık						✓

Şekil 4.18. Önerilen sistemin yanlış bulduğu yüz ifadesi görüntülerinden örnekler.

4.7. Önerilen İki Aşamalı 5 Kanallı Yüz İfadesi Tespiti Yöntemin Literatürdeki Diğer Çalışmalar İle Karşılaştırılması

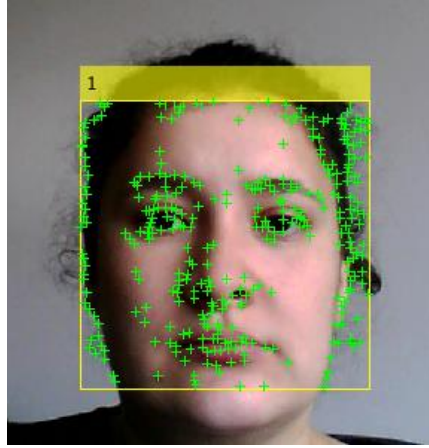
Önerilen iki aşamalı 5 kanal girişli yüz ifadesi tespit sisteminin başarımlar oranı, literatürdeki aynı veritabanlarını kullanan (RaFD ve KDEF) diğer yüz ifadesi tespiti çalışmaları ile karşılaştırılmıştır. Tablo 4.11. bu karşılaştırma sonuçlarını göstermektedir. Tabloda, CNN-3'ün 3 kanallı (RGB) orijinal ham yüz görüntüsü, 1 kanallı bölütlenmiş özet yüz görüntüsü ve 4 kanallı orijinal ham yüz görüntüsü ile birleştirilmiş bölütlenmiş özet yüz görüntüsü verileri ile eğitilmesi sonucunda elde edilen başarımlar oranları da sıralanmıştır. Tablo 4.11. analiz edildiğinde görülebileceği gibi, bu tez çalışmasında önerilen 5 kanal girişli yüz ifadesi sistemi, literatüre göre daha yüksek başarımlar göstermiştir (%94,61). Sistemde sadece 1 kanallı bölütlenmiş yüz görüntülerinin kullanımı literatüre göre eş veya biraz daha düşük (%89,11) sonuç verse de, yukarıda da bahsedildiği gibi bu görüntülerin yüz ifadesi ile ilgili temsili görüntüyü barındırıp kişisel hiç bir veri içermemesi, yüz ifadelerinin arşivlenmesini gerektirebilecek medikal uygulamalar gibi bir çok alanda tercih sebebi olabilir.

Tablo 4.11. Önerilen yüz ifadesi tespiti sisteminin literatürdeki diğer çalışmalar ile karşılaştırılması (*: 7 sınıfla çalışılan çalışmaları gösterir)

Yöntemler	Veritabanı	Başarım Oranı (%)
HOG + NNE [118]	RaFD, TFEID, JAFFE	93,75
Surf Boosting [119]	RaFD	90,64
Gabor F. + GLCM* [120]	RaFD	88,41
Facial Components Detection + Fuzzy [121]	RaFD	93,96
Facial Components Detection + KNN [121]	RaFD	75,61
Viola & Jones + AAM + ANN* [122]	RaFD	89,55
DFD+KNN* [123]	KDEF	82.24
LSiBP+SVM* [124]	KDEF	84.07
HOG+SVM* [125]	KDEF	80.20
HOG+AdaBoost* [125]	KDEF	75.20
HOG+AdaBoost+SVM* [126]	KDEF	87.20
Önerilen Yöntem	RaFD (3-kanallı ham görüntü)	83,33
	RaFD (1-kanallı bölütlenmiş özet görüntü)	89,11
	RaFD (3-kanallı ham görüntü ile kombinlenmiş 1-kanallı bölütlenmiş özet görüntü)	93,43
	RaFD (2-kanallı CNN-2 olasılık değerleri ile kombinlenmiş 3-kanallı ham görüntü)	94,61
	KDEF(2-kanallı CNN-2 olasılık değerleri ile kombinlenmiş 3-kanallı ham görüntü)	92,86

4.8. Müşteri Davranış Analizi Sisteminin Çalışmasından Örnekler

Şekil 4.19.'da KLT algortiması ile yüz anahtar noktalarının tespit edilip yüzün etiketlenmesine ait bir örnek sunulmuştur.



Şekil 4.19. KLT algoritması ile yüz anahtar noktalarının tespit edilip, yüzün etiketlenmesi.

Sistemin çalışmasından bir örnek ise Şekil 4.20.'de sunulmuştur. Şekilde de görüldüğü gibi kişinin başı kameraya dönükse (kameranın ilgili ürün üzerinde olduğu varsayıldığından, baş ilgili ürüne dönükse) kişinin görsel ilgi odağı ilgili ürün ya da reklamdır. Bu nedenle yüz ifadesi aşamasına geçilir. Yüz ifadeleri belli bir süre boyunca depolanır ve bu süre sonunda en çok sayıda gösterdiği duygular analiz edilerek kişinin modu hakkında bilgi verilir. Eğer kişinin baş yönelimi kameraya doğru değilse, kişinin görsel dikkati üründe olmadığından ürünle ilgilenmediği gerekçesi ile yüz ifadesi tespiti aşamasına geçilmez.



Şekil 4.20. Sistemin çalışma zamanından örnekler.

Şekil 4.21.'de sistemin birden fazla kişiye ait yüzleri etiketlemesi ve bu kişilerin yüz ifadelerini bulmasına ait örnekler gösterilmiştir. Kişiler yer değiştirse bile sistem

kişileri takip edebilmekte ve her kişinin etiket numarası ile birlikte yüz ifadelerini ayrı ayrı depolamaktadır. Kamera alanına yeni kişiler girmesi ile bu kişiler de etiketlenmekte ve kamera alanından kişilerin çıkması durumunda etiketler silinmektedir.



Şekil 4.21. Sistemin iki kişiye ait yüz görüntüsünü etiketleyip, takip etmesi ve yüz ifadelerini tespit etmesinden bir örnek.

BÖLÜM 5. TARTIŞMA VE SONUÇ

Müşteri davranışlarının incelenip, ilgi duyulan reklam ve ürünlerin tespit edilmesi pazarlamacılık alanının önemli konularındandır. Bu amaçla yapılan geleneksel anketler günümüzde müşteriye rahatsız edici bir yöntem olarak değerlendirilmektedir. İnsan gözlemcilerin, müşteri davranışlarını incelemesi ve karar vermesi ise hem deneyim gerektiren hem de göreceli olarak değişebilen bir durumdur. Bu nedenle müşteri davranışlarını inceleyip ilgi duyulan ürünleri belirleyen, satın alma eğilimlerini değerlendiren sistemlerin geliştirilmesine ihtiyaç duyulmaktadır. Böylece bu alanda tarafsız, güvenilir ve nicel sonuçlar üreten sistemler etkili olarak satıcıya destek verebilecek ve çözüm üretmesine katkı sağlayabilecektir.

Bu çalışmada önerilen sistem, müşterinin görsel ilgi odağına ve duygu moduna göre satın alma eğilimi hakkında satıcıya destek vermeyi hedeflemektedir. Bunu gerçekleştirmek için ilk olarak ortamda bulunan insan yüzleri tespit edilmektedir. İnsan yüzleri tespit edilip ortamdan bağımsız hale getirildikten sonra bu yüzlerin baş yönelimleri tespit edilmektedir. Bu aşamada kameranın, ilgi duyulması beklenen ürün ya da reklamın üzerinde olduğu varsayılmıştır. Dolayısıyla eğer müşterinin baş yönelimi kameraya doğru ise görsel ilgi odağı ilgili üründe olabilir. Müşterinin görsel ilgi odağı üründe değilse zaman kaybetmemek için yüz ifadesi tespiti aşamasına geçilmez ve yeniden ortamdaki yüzler tespit edilerek analizler yeniden gerçekleşir. Eğer müşteri görsel odağı ilgili ürüne doğru ise bundan sonra sıradaki aşama müşterinin duygu modunun belirlenmesidir. Müşteri duygu modunu belirlemek için yüz ifadelerinden yararlanılmıştır. Psikoloji alanındaki araştırmalara göre kızgınlık, üzüntü, iğrenme, korku duyguları negatif duygu iken mutluluk duygusu pozitif duygudur. Bu çalışmada da insan duygu modunu belirlemek amacıyla, yüz ifadeleri pozitif ve negatif olarak sınıflandırılmıştır. Sistem belli bir süre boyunca müşterinin yüz ifadelerini analiz eder ve depolar. Belirlenen süre sonunda müşterinin gösterdiği

duygu grafiğine göre müşterinin olumlu ya da olumsuz modda olduğu ile ilgili sonucu ekranda gösterir.

Belirlenen süre boyunca izlenen müşterinin kamera görüş açısının değişmesi yani kişilerin yer değiştirmesi, kamera görüş açısından ayrılması veya kamera görüş açısına yeni müşteri / müşteriler girmesi durumları da göz önünde bulundurulmuştur. Sistem her bir insan yüzünü etiketleyerek, kişileri birbirinden bağımsız olarak inceler ve takip eder. Her bir kişi için analiz sonucunu kişi etiketi ile birlikte ekranda gösterir.

Bu tez çalışmasında önerilen sistemi gerçekleyebilmek için bilgisayar görmesi ve derin öğrenme yaklaşımlarından yararlanılmıştır. Ortamdaki insan yüzleri Viola&Jones yöntemi ile tespit edilmektedir. Görüntü işleme yöntemleri ile ortamdan bağımsız hale getirilen insan yüzleri, baş yönelimi tespiti için geliştirilen CNN-1 mimarisinde giriş olarak kullanılır ve baş yönelimi tespit edilir. Baş yöneliminin kameraya doğru olması durumunda; yüz görüntüsü, yüz bölütleme için geliştirilen CNN-2 mimarisinde kullanılır. Böylece yüz ifade tespitinde önemli olan kaşlar, gözler ve ağız bölgelerinden oluşan özet yüz görüntüsü elde edilir. Ardından CNN-2'deki tam bağlantılı katmanda elde edilen arkaplan ve yüz bileşeni olasılıksal değerleri ile birlikte ham yüz görüntüsü 5 kanallı bir matriste birleştirilir. Bu matris yüz ifadesi tespiti için geliştirilen CNN-3'e giriş olarak verilir ve yüz ifadesi bilgisi elde edilir. İnsan yüzlerinin etiketlenip takip edilmesi ise KLT algoritması ile gerçekleştirilmiştir.

Çalışmanın sonuçları analiz edildiğinde RaFD üzerinde 5 sınıflı baş yöneliminin %99,90 oranında, 6 sınıflı yüz ifadesinin ise %94,61 oranında doğru sonuç verdiği görülmüştür. Önerilen sistem için başın kameraya dönük olması ya da olmaması şeklindeki iki sınıflı baş yönelimi yeterli olmasına rağmen, kullanılan veritabanında 5 açılı baş yönelimleri yer aldığından sınıflandırma buna göre yapılmıştır. Test aşamasında başın kameraya doğru olduğu pozisyonları sistem %100 doğruluk oranı ile tespit edebilmiştir.

Yüz ifadesi tespiti yukarıda bahsedildiği gibi yüz bölütleme ve bölütlenmiş görüntülere ait bilgileri kullanarak yüz ifadesi sınıflandırma olmak üzere iki aşamalı

olarak gerçekleştirilmiştir. İki aşamalı sistemin etkinliğini gösterebilmek için yüz ifadesi aşamasında farklı deneyler de yapılmıştır. İlk olarak CNN-3 sadece ham yüz görüntüleri ile eğitilmiş ve bu eğitimin başarısı %83,33 olarak tespit edilmiştir. İkinci olarak CNN-3 sadece bölütlenmiş özet yüz görüntüleri ile eğitilmiş ve bu eğitimden %89,11 başarı oranı elde edilmiştir. Ardından CNN-3; ham yüz görüntüleri ile birlikte bölütlenmiş özet yüz görüntülerinin kombinlenmiş hali ile eğitilmiş ve bu başarımdan da %93,43 sonuç elde edilmiştir. Bu deneylerin sonuçlarından anlaşılabileceği gibi önerilen yöntemde CNN-2'ye ait tam bağlantılı katmanın olasılıksal değerlerinin kullanılması çalışmanın başarımını artırmıştır. Ayrıca bu yöntem, hem parçalı hem bütünsel yüz özneteliklerini birlikte kullanması bakımından da çalışmaya özgünlük katmıştır. Çalışma literürdeki diğer yüz ifadesi tespiti çalışmaları ile karşılaştırıldığında ise başarı oranıyla da literatüre katkı sağlamaktadır.

Test aşamasında sistemin hatalı bulduğu sonuçlar analiz edildiğinde, hatalı bulunan bazı yüz görüntüleri için insanın bile ifade tespitinde zorlandığı gözlenmiştir. Kullanılan veritabanlarında yer alan görüntüler, modellerin yüz ifadelerini taklit etmesi ile oluşturulan, doğal olmayan ifadeler içerir. Eğitim ve test bu görüntüler ile yapıldığından, sistemde hatalar oluşabilmektedir. Eğitim ve testin doğal ifadeler içeren veritabanı ile yapılması ile daha başarılı sonuçlar elde edilebilir.

Önerilen sistem; satıcıları, müşteri davranışlarını takip edip, ilgi çeken ürünlerin tespit edilmesi konusunda destekler. Satın alma potansiyelinde olan pozitif moddaki müşterilerin belirlenmesinde yardımcı bir araçtır. Nicel sonuçlar ürettiği ve tarafsız olduğu için insan gözlemcilerle göre daha güvenilir olabilir. Bu sistem sayesinde satıcılar ilgi çeken ürünleri tespit edebileceği için satış politikalarını ve market düzenlerini belirlerken de daha bilinçli davranabilir. Ayrıca pazarlamacı tarafından bir grup insana ürün tanıtımı yapılırken de sistemin ürettiği sonuca göre pazarlamacı ürüne ilgi duyan müşterilere odaklanarak satışlarını artırabilir ve zaman kazanabilir.

Geliştirilen müşteri davranış analizi sistemine gelecek çalışmalarda cinsiyet tespiti, yaş tespiti gibi modüller eklenebilir. Böylece sistem, sadece belirli cinsiyetteki veya

belirli yaştaki müşterilere hitap eden ürünlere olan ilgi tespitinde kullanılmak üzere özelleştirilebilir. Ayrıca herhangi bir ürünün hangi yaştaki veya hangi cinsiyetteki kitleye daha çok hitap ettiği de bu şekilde tespit edilebilir. Bu tez çalışmasının motivasyonu müşteri davranışlarını izleme olsa da çalışmada sunulan iki aşamalı yüz ifadesi tespiti yöntemi veya üç aşamalı ilgi tespiti sistemi, öğrenci ilgisinin izlenmesi, hasta dikkatinin ve modunun izlenmesi gibi başka motivasyonlar için de kullanılabilir.



KAYNAKLAR

- [1] A. M. Barreto, Application of facial expression studies on the field of marketing. Porto: FEELab Science Books, 2017.
- [2] K. Menon and L. Dubé, Ensuring greater satisfaction by engineering salesperson response to customer emotions, *J. Retail.*, vol. 76, no. 3, pp. 285–307, Jul. 2000.
- [3] A. L. Missaglia, A. Oppo, M. Mauri, B. Ghiringhelli, A. Ciceri, and V. Russo, The impact of emotions on recall: An empirical study on social ads, *J. Consum. Behav.*, vol. 16, no. 5, pp. 424–433, Sep. 2017.
- [4] D. D. Langleben, J. W. Loughhead, K. Ruparel, J. G. Hakun, S. Busch-Winokur, M. B. Holloway, A. A. Strasser, J. N. Cappella, and C. Lerman, Reduced prefrontal and temporal processing and recall of high ‘sensation value’ ads, *Neuroimage*, vol. 46, no. 1, pp. 219–225, May 2009.
- [5] R. Ohme, D. Reykowska, D. Wiener, and A. Choromanska, Analysis of neurophysiological reactions to advertising stimuli by means of EEG and galvanic skin response measures., *J. Neurosci. Psychol. Econ.*, vol. 2, no. 1, pp. 21–31, 2009.
- [6] I. A. Cook, C. Warren, S. K. Pajot, D. Schairer, and A. F. Leuchter, Regional brain activation with advertising images., *J. Neurosci. Psychol. Econ.*, vol. 4, no. 3, pp. 147–160, 2011.
- [7] M. Wedel and R. Pieters, Eye fixations on advertisements and memory for brands: a model and findings, *Mark. Sci.*, vol. 19, no. 4, pp. 297–312, Nov. 2000.
- [8] 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, 2015.
- [9] F. Ungureanu, R. G. Lupu, A. Cadar, and A. Prodan, Neuromarketing and visual attention study using eye tracking techniques, in 2017 21st International Conference on System Theory, Control and Computing (ICSTCC), 2017, pp. 553–557.

- [10] A. C. Micu and J. T. Plummer, Measurable emotions: how television ads really work, *J. Advert. Res.*, vol. 50, no. 2, pp. 137–153, Jun. 2010.
- [11] T. Teixeira, M. Wedel, and R. Pieters, Emotion-induced engagement in internet video advertisements, *J. Mark. Res.*, vol. 49, no. 2, pp. 144–159, Apr. 2012.
- [12] Z. Kasiran and S. Yahya, Facial expression as an implicit customers' feedback and the challenges, in *Computer Graphics, Imaging and Visualisation*, 2007, pp. 377–381.
- [13] A. Riener and A. Sippl, Head-pose-based attention recognition on large public displays, *IEEE Comput. Graph. Appl.*, vol. 34, no. 1, pp. 32–41, Jan. 2014.
- [14] S. Sheikhi and J.-M. Odobez, Combining dynamic head pose–gaze mapping with the robot conversational state for attention recognition in human–robot interactions, *Pattern Recognit. Lett.*, vol. 66, pp. 81–90, Nov. 2015.
- [15] S. O. Ba and J. Odobez, Multiperson visual focus of attention from head pose and meeting contextual cues, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 1, pp. 101–116, Jan. 2011.
- [16] A. Gaschler, S. Jentzsch, M. Giuliani, K. Huth, J. de Ruiter, and A. Knoll, Social behavior recognition using body posture and head pose for human-robot interaction, in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 2128–2133.
- [17] N. Alioua, A. Amine, A. Rogozan, A. Bensrhair, and M. Rziza, Driver head pose estimation using efficient descriptor fusion, *EURASIP J. Image Video Process.*, vol. 2016, no. 1, p. 2, Dec. 2016.
- [18] E. Murphy-Chutorian, A. Doshi, and M. M. Trivedi, Head pose estimation for driver assistance systems: a robust algorithm and experimental evaluation, in *IEEE Intelligent Transportation Systems Conference*, 2007, pp. 709–714.
- [19] E. Murphy-Chutorian and M. M. Trivedi, “Head Pose Estimation and Augmented Reality Tracking: An Integrated System and Evaluation for Monitoring Driver Awareness,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 2, pp. 300–311, Jun. 2010.
- [20] S. Asteriadis, P. Tzouveli, K. Karpouzis, and S. Kollias, Estimation of behavioral user state based on eye gaze and head pose---application in an e-learning environment, *Multimed. Tools Appl.*, vol. 41, no. 3, pp. 469–493, 2009.
- [21] S. Alghowinem, R. Goecke, M. Wagner, G. Parkerx, and M. Breakspear, Head pose and movement analysis as an indicator of depression, in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013, pp. 283–288.

- [22] I. Karu-salo, The effect of universal emotions on customer behaviour, Estonian Business School, 2013.
- [23] D. Kahneman, Thinking fast and slow. Allen Lane, 2011.
- [24] S. Sahney, Module-6 consumer behaviour, Vinod Gupta School of Management, 2011, pp. 1–24.
- [25] 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.
- [26] P. Ekman and W. V. Friesen, *Unmasking the face: a guide to recognising emotions from facial clues*, New Jersey:Prentice Hall [Palo Alto, CA]: Consulting Psychologists Press, 1975.
- [27] F. J. M. Laros and J.-B. E. M. Steenkamp, Emotions in consumer behavior: a hierarchical approach, *J. Bus. Res.*, vol. 58, no. 10, pp. 1437–1445, Oct. 2005.
- [28] L. Zalewski and Shaogang Gong, 2D statistical models of facial expressions for realistic 3D avatar animation, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005, vol. 2, pp. 217–222.
- [29] G. Yolcu, I. Oztel, S. Kazan, C. Oz, K. Palaniappan, T. E. Lever, and F. Bunyak, Deep learning-based facial expression recognition for monitoring neurological disorders, in *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017, pp. 1652–1657.
- [30] G. Littlewort, M. S. Bartlett, I. Fasel, J. Chenu, T. Kanda, H. Ishiguro, and J. R. Movellan, *Towards social robots: Automatic evaluation of human-robot interaction by face detection and expression classification*, 2003.
- [31] I. Shaykha, A. Menkara, M. Nahas, and M. Ghantous, FEER: Non-intrusive facial expression and emotional recognition for driver's vigilance monitoring, in *57th International Symposium ELMAR*, 2015, pp. 233–237.
- [32] Jung-Bae Kim, Youngkyoo Hwang, Won-Chul Bang, Heesae Lee, J. D. K. Kim, and ChangYeong Kim, Real-time realistic 3D facial expression cloning for smart TV, in *IEEE International Conference on Consumer Electronics (ICCE)*, 2013, pp. 240–241.
- [33] E. Niforatos and E. Karapanos, EmoSnaps: a mobile application for emotion recall from facial expressions, *Pers. Ubiquitous Comput.*, vol. 19, no. 2, pp. 425–444, Feb. 2015.
- [34] V. Terzis, C. N. Moridis, and A. A. Economides, Measuring instant emotions based on facial expressions during computer-based assessment, *Pers. Ubiquitous Comput.*, vol. 17, no. 1, pp. 43–52, Jan. 2013.

- [35] M. Patacchiola and A. Cangelosi, Head pose estimation in the wild using convolutional neural networks and adaptive gradient methods, *Pattern Recognit.*, vol. 71, pp. 132–143, Nov. 2017.
- [36] E. Murphy-Chutorian and M. M. Trivedi, Head pose estimation in computer vision: a survey, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 4, pp. 607–626, Apr. 2009.
- [37] S. . McKenna and S. Gong, Real-time face pose estimation, *Real-Time Imaging*, vol. 4, no. 5, pp. 333–347, Jan. 1998.
- [38] S. S. Mukherjee and N. M. Robertson, Deep head pose: gaze-direction estimation in multimodal video, *IEEE Trans. Multimed.*, vol. 17, no. 11, pp. 2094–2107, Nov. 2015.
- [39] X. Xu and I. A. Kakadiaris, Joint head pose estimation and face alignment framework using global and local CNN features, in *12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 2017, pp. 642–649.
- [40] G. Pons and D. Masip, Supervised committee of convolutional neural networks in automated facial expression analysis, *IEEE Trans. Affect. Comput.*, pp. 1–1, 2017.
- [41] S. K. Gupta, S. Agrwal, Y. K. Meena, and N. Nain, A hybrid method of feature extraction for facial expression recognition, in *2011 Seventh International Conference on Signal Image Technology & Internet-Based Systems*, 2011, pp. 422–425.
- [42] L. Zhang, S. Chen, T. Wang, and Z. Liu, Automatic facial expression recognition based on hybrid features, *Energy Procedia*, vol. 17, pp. 1817–1823, 2012.
- [43] T. Cootes, E. Baldock, and J. Graham, An introduction to active shape models, *Image Process. Anal.*, pp. 223–248, 2000.
- [44] Y. Yacoob and L. S. Davis, Recognizing human facial expressions from long image sequences using optical flow, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 6, pp. 636–642, Jun. 1996.
- [45] D. Ghimire and J. Lee, Geometric feature-based facial expression recognition in image sequences using multi-class adaboost and support vector machines, *Sensors*, vol. 13, no. 6, pp. 7714–7734, Jun. 2013.
- [46] G. Zhao and M. Pietikainen, Dynamic texture recognition using local binary patterns with an application to facial expressions, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 915–928, Jun. 2007.

- [47] D. Ghimire and J. Lee, Histogram of orientation gradient feature-based facial expression classification using bagging with extreme learning machine, *Adv. Sci. Lett.*, vol. 17, no. 1, pp. 156–161, Oct. 2012.
- [48] M. C. Sobia, V. Brindha, and A. Abudhahir, Facial expression recognition using PCA based interface for wheelchair, in *International Conference on Electronics and Communication Systems (ICECS)*, 2014, pp. 1–6.
- [49] Y. Luo, C. Wu, and Y. Zhang, Facial expression feature extraction using hybrid PCA and LBP, *J. China Univ. Posts Telecommun.*, vol. 20, no. 2, pp. 120–124, 2013.
- [50] A. Hernandez-Matamoros, A. Bonarini, E. Escamilla-Hernandez, M. Nakano-Miyatake, and H. Perez-Meana, Facial expression recognition with automatic segmentation of face regions using a fuzzy based classification approach, *Knowledge-Based Syst.*, vol. 0, pp. 1–14, 2016.
- [51] D. Boruah, K. K. Sarma, and A. K. Talukdar, Different face regions detection based facial expression recognition, *2nd Int. Conf. Signal Process. Integr. Networks, SPIN 2015*, pp. 459–464, 2015.
- [52] S. M. Lajevardi and Z. M. Hussain, Feature extraction for facial expression recognition based on hybrid face regions, vol. 9, no. 3, pp. 63–67, 2009.
- [53] 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.
- [54] 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.
- [55] 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.
- [56] S. Xie and H. Hu, Facial expression recognition with FRR-CNN, *Electron. Lett.*, vol. 53, no. 4, pp. 235–237, Feb. 2017.
- [57] P. Liu, S. Han, Z. Meng, and Y. Tong, Facial Expression Recognition via a boosted deep belief network, in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
- [58] P. Khorrami, T. Le Paine, and T. S. Huang, Do deep neural networks learn facial action units when doing expression recognition?, in *The IEEE International Conference on Computer Vision (ICCV) Workshops*, 2015.

- [59] Y. Kim, H. Lee, and E. M. Provost, Deep learning for robust feature generation in audiovisual emotion recognition, in 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp. 3687–3691.
- [60] X. Zhao, X. Shi, and S. Zhang, Facial expression recognition via deep learning, IETE Tech. Rev., vol. 32, no. 5, pp. 347–355, Sep. 2015
- [61] <https://mashable.com/2018/05/18/chinese-facial-recognition-class/#9Ejjgt63GQqb>, Erişim Tarihi: 01.04.2019.
- [62] <https://www.telegraph.co.uk/news/2018/05/17/chinese-school-uses-facial-recognition-monitor-student-attention/>, Erişim Tarihi: 01.04.2019.
- [63] <https://bgr.com/2017/07/19/walmart-stores-facial-recognition-angry-customers> . , Erişim Tarihi: 01.04.2019.
- [64] <https://www.forbes.com/sites/retailwire/2017/07/27/walmarts-facial-recognition-tech-would-overstep-boundaries/#6f8f266745f8le>., Erişim Tarihi: 01.04.2019.
- [65] <https://www.businessinsider.com/walmart-is-developing-a-robot-that-identifies-unhappy-shoppers-2017-7>., Erişim Tarihi: 01.04.2019.
- [66] 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.
- [67] R. C. Arkin, M. Scheutz, and L. Tickle-Degnen, Preserving dignity in patient caregiver relationships using moral emotions and robots, in 2014 IEEE International Symposium on Ethics in Science, Technology and Engineering, 2014, pp. 1–5.
- [68] W. Gaebel and W. Wölwer, Facial expression and emotional face recognition in schizophrenia and depression, Eur. Arch. Psychiatry Clin. Neurosci., vol. 242, no. 1, pp. 46–52, Sep. 1992.
- [69] C. Alvino, C. Kohler, F. Barrett, R. E. Gur, R. C. Gur, and R. Verma, Computerized measurement of facial expression of emotions in schizophrenia, J. Neurosci. Methods, vol. 163, no. 2, pp. 350–361, Jul. 2007.
- [70] A. Al-Modwahi, O. Sebetela, L. Nehemiah Batleng, B. Parhizkar, and A. Habibi Lashkari, Facial expression recognition intelligent security system for real time surveillance, 2012.
- [71] Guanming Lu, Xiaonan Li, and Haibo Li, Facial expression recognition for neonatal pain assessment, in 2008 International Conference on Neural Networks and Signal Processing, 2008, pp. 456–460.

- [72] 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.
- [73] A. Benmohamed, M. Neji, M. Ramdani, A. Wali, and A. M. Alimi, Feast: face and emotion analysis system for smart tablets, *Multimed. Tools Appl.*, vol. 74, no. 21, pp. 9297–9322, Nov. 2015.
- [74] P. Viola, O. M. Way, and M. J. Jones, Robust real-time face detection, vol. 57, no. 2, pp. 137–154, 2004.
- [75] Y. Freund and R. E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, *J. Comput. Syst. Sci.*, vol. 55, no. 1, pp. 119–139, 1997.
- [76] P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features, *Comput. Vis. Pattern Recognit.*, vol. 1, p. I-511--I-518, 2001.
- [77] Y. Tetik and B. Bolat, “Gürültülü ortamlarda konuşma tespiti için yeni bir öz nitelik çıkarım yöntemi,” in *Elektrik-Elektronik ve Bilgisayar Sempozyumu*, 2011.
- [78] Jianbo Shi and Tomasi, Good features to track, in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition CVPR-94*, 1994, pp. 593–600.
- [79] Y. S. Erdoğan, Sayısal Görüntü işleme ile titreşim yer değiştirmelerinin ölçümü ve modal parametre tayini, *Deu Muhendis. Fak. Fen ve Muhendis.*, vol. 20, no. 59, pp. 400–417, Jan. 2018.
- [80] B. D. Lucas and T. Kanade, “An Iterative Image Registration Technique with an Application to Stereo Vision,” in *Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2*, 1981, pp. 674–679.
- [81] C. Tomasi and T. Kanade, “Detection and Tracking of Point Features,” 1991.
- [82] Ö. İnik and E. Ülker, Derin öğrenme ve görüntü analizinde kullanılan derin öğrenme modelleri, *Gaziosmanpaşa J. Sci. Res.*, vol. 6, no. 3, pp. 85–104, 2017.
- [83] J.-S. Choi, W.-H. Lee, J.-H. Lee, J.-H. Lee, and S.-C. Kim, Deep learning based NLOS identification with commodity WLAN devices, *IEEE Trans. Veh. Technol.*, pp. 1–1, 2017.
- [84] W. S. McCulloch and W. Pitts, A logical calculus of the ideas immanent in nervous activity, *Bull. Math. Biophys.*, vol. 5, no. 4, pp. 115–133, 1943.
- [85] B. Widrow and M. E. Hoff, Adaptive switching circuits, 1960 IRE WESCON Conv. Rec., pp. 96–104, 1960.

- [86] F. Rosenblatt, The perceptron: a probabilistic model for information storage and organization in the brain, *Psychol. Rev.*, pp. 65–386, 1958.
- [87] M. L. Minsky and S. A. Papert, *Perceptrons*, MIT Press, Cambridge, 1969.
- [88] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Learning representations by back-propagating errors, *Nature*, vol. 323, p. 533, Oct. 1986.
- [89] D. E. Rumelhart, J. McClelland, and L. James, *Parallel distributed processing: explorations in the microstructure of cognition. Volume 1. Foundations*. 1986.
- [90] G. E. Hinton, S. Osindero, and Y.-W. Teh, A fast learning algorithm for deep belief nets, *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.
- [91] Y. Bengio and Y. Lecun, Scaling learning algorithms towards AI, in *large-scale kernel machines*, L. Bottou, O. Chapelle, D. DeCoste, and J. Weston, Eds. MIT Press, 2007.
- [92] O. Delalleau and Y. Bengio, Shallow vs. deep sum-product networks, in *advances in neural information processing systems 24*, J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2011, pp. 666–674.
- [93] A. Krizhevsky, I. Sutskever, and G. E. Hinton, Imagenet classification with deep convolutional neural networks, in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.
- [94] <https://devblogs.nvidia.com/mocha-jl-deep-learning-julia/>, Erişim Tarihi: 01.04.2019.
- [95] I. Oztel, G. Yolcu, I. Ersoy, T. White, and F. Bunyak, Mitochondria segmentation in electron microscopy volumes using deep convolutional neural network, in *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2017, pp. 1195–1200.
- [96] 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.
- [97] <https://www.mathworks.com/solutions/deep-learning/convolutional-neural-net-work.html>, Erişim Tarihi: 01.04.2019.
- [98] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT Press, 2016.
- [99] http://machinelearningguru.com/computer_vision/basics/convolution/convolution_layer.html, Erişim Tarihi: 01.04.2019.

- [100] S. Srinivas, R. K. Sarvadevabhatla, K. R. Mopuri, N. Prabhu, S. S. S. Kruthiventi, and R. V. Babu, A taxonomy of deep convolutional neural nets for computer vision, *Front. Robot. AI*, Jan. 2016.
- [101] https://leonardoaraujosantos.gitbooks.io/artificial-intelligence/content/pooling_layer.html., Eriřim Tarihi: 01.04.2019.
- [102] https://software.intel.com/sites/products/documentation/doclib/daal/daal-user-and-reference-guides/daal_prog_guide/GUID-9B434D4F-C723-4191-9A88-69148C75A3F1.htm., Eriřim Tarihi: 01.04.2019.
- [103] V. Nair and G. E. Hinton, Rectified linear units improve restricted boltzmann machines, in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, 2010, pp. 807–814.
- [104] H. K. Verma and S. Pal, Modified sigmoid function based gray scale image contrast enhancement using particle swarm optimization, *J. Inst. Eng. Ser. B*, vol. 97, no. 2, pp. 243–251, 2016.
- [105] P. Bhattacharya, A. Sarkar, , IndranilSarkar, and S. Chatterjee, An ANN based call handoff management scheme for mobile cellular network, *Int. J. Wirel. Mob. Networks*, vol. 5, 2014.
- [106] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, Improving neural networks by preventing co-adaptation of feature detectors, *arXiv*, Jul. 2012.
- [107] 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.
- [108] P. Hatipoglu, Time series classification using deep learning, Middle East Technical University, 2016.
- [109] M. D. Zeiler and R. Fergus, Visualizing and understanding convolutional networks, Nov. 2013.
- [110] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, Going deeper with convolutions, Sep. 2014.
- [111] R. Zhang, P. Isola, and A. A. Efros, Colorful image colorization, Mar. 2016.
- [112] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, Dec. 2015.
- [113] 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.

- [114] P. Ekman and W. V. Friesen, Facial action coding system: a technique for the measurement of facial movement, Consulting Psychologists Press, 1978.
- [115] D. Lundqvist, A. Flykt, and A. Öhman, The karolinska directed emotional faces (KDEF), Department of Clinical Neuroscience. Psychology section, Karolinska Institutet, 1998.
- [116] <https://www.faceplusplus.com/> , Erişim Tarihi: 01.04.2019.
- [117] E. Shelhamer, J. Long, and T. Darrell, Fully convolutional networks for semantic segmentation, *Cogn. Emot.*, vol. 24, no. 8, pp. 1377–1388, May 2016.
- [118] 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.
- [119] 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.
- [120] 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 *International Conference on Computational Intelligence and Communication Networks (CICN)*, 2015, pp. 347–351.
- [121] 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.
- [122] G. Bijlstra and R. Dotsch, FaceReader 4 emotion classification performance on images from the Radboud Face Database, 2011.
- [123] Z. Sun, Z.-P. Hu, M. Wang, and S.-H. Zhao, Discriminative feature learning-based pixel difference representation for facial expression recognition, *IET Comput. Vis.*, vol. 11, no. 8, pp. 675–682, Dec. 2017.
- [124] B. Santra and D. P. Mukherjee, Local saliency-inspired binary patterns for automatic recognition of multi-view facial expression, in *IEEE International Conference on Image Processing (ICIP)*, 2016, pp. 624–628.
- [125] C. F. Liew and T. Yairi, Facial Expression Recognition and Analysis: A Comparison Study of Feature Descriptors, *IPSJ Trans. Comput. Vis. Appl.*, vol. 7, pp. 104–120, 2015.
- [126] C. F. Liew and T. Yairi, A comparison study of feature spaces and classification methods for facial expression recognition, in *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2013, pp. 1294–1299.

ÖZGEÇMİŞ

Gözde Yolcu Öztel, 31.01.1988'de Zonguldak'ta doğdu. İlk, orta ve lise eğitimini Zonguldak'ta tamamladı. 2006 yılında Zonguldak Fen Lisesi'nden mezun oldu. 2007 yılında başladığı Sakarya Üniversitesi Bilgisayar Mühendisliği Bölümü'nü 2011 yılında bitirdi. 2012 yılında Sakarya Üniversitesi Bilgisayar ve Bilişim Mühendisliği Bölümü'nde yüksek lisans eğitimine başladı. 2014 yılında yüksek lisans eğitimini tamamlayarak aynı yıl Sakarya Üniversitesi'nde Bilgisayar ve Bilişim Mühendisliği Bölümü'nde doktora eğitime başladı. 2017-2018 yılları arasında, 1 yıl süre ile ABD'de bulunan University of Missouri-Columbia'da Misafir Araştırmacı olarak çalıştı. 2012 yılından beri Sakarya Üniversitesi Bilgisayar Mühendisliği Bölümü'nde Araştırma Görevlisi olarak görev yapmaktadır.