

T.C.
İSTANBUL KÜLTÜR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

ÇEKİŞMELİ ÜRETİCİ AĞLAR İLE YÜZ İFADELERİNDEN DUYGU
ETİKETLEME

YÜKSEK LİSANS TEZİ
Gizem ÖZGÜL
1900004867

Anabilim Dalı: Bilgisayar Mühendisliği
Programı: Bilgisayar Mühendisliği

Tez Danışmanı: Dr.Öğr.Üyesi Fatma Patlar AKBULUT

HAZİRAN 2022

T.C.
İSTANBUL KÜLTÜR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

ÇEKİŞMELİ ÜRETİCİ AĞLAR İLE YÜZ İFADELERİNDEN DUYGU
ETİKETLEME

YÜKSEK LİSANS TEZİ
Gizem ÖZGÜL
1900004867

Anabilim Dalı: Bilgisayar Mühendisliği
Programı: Bilgisayar Mühendisliği

Tez Danışmanı: Dr.Öğr. Üyesi Fatma Patlar AKBULUT
Jüri Üyeleri: Doç.Dr. Akhan AKBULUT
Prof.Dr.Özgür Koray ŞAHİNGÖZ

HAZİRAN 2022

ÖNSÖZ

“ÇEKİŞMELİ ÜRETİCİ AĞLAR İLE YÜZ İFADELERİNDEN DUYGU ETİKETLEME”

konulu yüksek lisans tez çalışmam süresince bilgi ve deneyimi ile çalışmalarımı yönlendiren ve desteğini esirgemeyen değerli tez danışmanım Dr. Fatma PATLAR AKBULUT’a, teşekkürü bir borç bilirim.

Hayatım boyunca desteğini esirgemeyen ve varlığı ile sırtımda bir dağ olan babam Ali ÖZGÜL’e başta olmak üzere, başarıma inancı tam olan annem Fatma Nezihe ÖZGÜL’e ve her türlü zorlukla mücadele ederken yanımda olan ablam Merve ÖZGÜL’e tüm kalbimle teşekkür ederim. Sizi gururlandırmaya bir ömür devam edeceğim.



İÇİNDEKİLER

ÖNSÖZ.....	III
ŞEKİL LİSTESİ	VI
TABLO LİSTESİ	VII
KISALTMALAR	VIII
ÖZET	IX
ABSTRACT	X
1. GİRİŞ	iii
1.1. Problem Tanımı	5
1.2. Tezin Amacı.....	5
1.3. Tezin Organizasyonu	5
2. ÖN BİLGİLER ve LİTERATÜR ARAŞTIRMASI	6
2.1. Duygu Tanımı ve Literatür Taraması.....	6
2.1.1. Yüz Görüntülerinden Duygu Tanıma Kullanım Alanları.....	9
2.1.2. Yüz Görüntülerinden Duygu Analizi Yapmanın Zorlukları	10
2.2. Görüntü İşleme.....	10
2.3. Derin Öğrenme İle İlgili Ön Bilgiler	11
2.3.1. Makine Öğrenmesi Genel Bakış	11
2.3.1.1. Gözetimli (Danışmanlı) Öğrenme.....	13
2.3.1.2. Gözetimsiz (Danışmansız) Öğrenme	13
2.3.1.3. Yarı Gözetimli Öğrenme	13
2.3.2. Veri Ön İşleme.....	14
2.3.3. Yapay Sinir Ağı.....	15
2.3.4. Evrişimli Sinir Ağları (CNN)	19
2.3.5. Çekişmeli Üretici Ağlar (GAN).....	22
2.3.6. Yaygın Kullanılan Veri Setleri.....	23
2.3.6.1. FER2013 (Facial Expression Recognition) Veri Seti	23
2.3.6.2. JAFFE (The Japanese Female Facial Recognition) Veri Seti	25
2.3.6.3. KDEF (Karolinska Directed Emotional Faces) Veri Seti.....	26
2.3.6.4. AFFECNET Veri Seti	26
2.3.6.5. Veri Setlerinin Özelliklerinin Kıyaslanması.....	27
2.4. Veri Setleri ve Derin Öğrenme Algoritmalarının Literatür Taraması	27
3.YÖNTEM	27

3.1. Veri Ön İşleme (Data Preprocessing).....	30
3.2. Normalizasyon İşlemi	30
3.3. Light Weight GAN ile Sahte Görüntü Üretme.....	31
3.4. ResNET (Artık Ağlar)	32
3.5. Inception-v3.....	34
3.6. Karışıklık Matrisi (Confusion Matrix)	35
4. ÖNERİLEN MODEL	37
4.1. Kullanılan Veri Seti	37
4.2. Modelin Akış Diyagramı	37
4.3. Modelde Kullanılan Teknoloji	42
5. TEST SONUÇLARI VE DEĞERLENDİRMELER.....	44
6. SONUÇLAR.....	50
7. SINIRLAMALAR VE GELECEK ÇALIŞMA.....	50
KAYNAKÇA.....	52

ŞEKİL LİSTESİ

Şekil 1.1- Yapay Zekâ İlişki Kümesi	12
Şekil 2.1- Kümeleme	13
Şekil 2.2- Makine Öğrenmesi Yöntemleri	14
Şekil 2.3- Nöron Ağ Yapısı	16
Şekil 2.4- YSA Çalışma Prensibi	17
Şekil 2.5- Çok Katmanlı Algılayıcılar Çalışma Prensibi	18
Şekil 2.6-CNN Çalışma Modeli	19
Şekil 2.7- Evrişim İşlemleri	20
Şekil 2.8- Ortalama Havuzlama İşlemi	20
Şekil 2.9- Maximum Havuzlama İşlemi	21
Şekil 2.10- Minimum Havuzlama İşlemi	21
Şekil 2.11- Düzleştirme İşlemi.....	22
Şekil 2.12- GAN Sistem Mimarisi	23
Şekil 2.13- FER2013 Veri Seti	25
Şekil 2.14- JAFFE Veri Seti	25
Şekil 2.15- KDEF Veri Seti	26
Şekil 3.1- Görsel-Etiket Bilgisi	30
Şekil 3.2-LightWeight Gan İle Elde Edilen Yüz Görselleri.....	32
Şekil 3.3- ResNET Katmanları	33
Şekil 3.4- ResNET Mimarisi	34
Şekil 3.5- Inception V3 Mimarisi	35
Şekil 3.6- Karışıklık Matrisi Modeli	36
Şekil 4.1- ResNet152 Modeli Akış Diyagramı	41
Şekil 4.2- Inception v3 Modeli Akış Diyagramı	42
Şekil 5.1-ResNet152 Eğitim Doğruluk Sonuçlar	44
Şekil 5.2-ResNet152 Kayıp Değerler.....	45
Şekil 5.3-ResNet152 Karışıklık Matrisi Eğitim Sonuçları	45
Şekil 5.4- ResNet152 Karışıklık Matrisi Test Doğrulama Sonuçları	46
Şekil 5.5-Inception v3 Eğitim Doğruluk Sonuçları.....	47
Şekil 5.6-Inception v3 Kayıp Değerler	48
Şekil 5.7-Inception v3 Karışıklık Matrisi Eğitim Sonuçları	48
Şekil 5.8-Inception v3 Karışıklık Matrisi Test Doğrulama Sonuçları	49

TABLO LİSTESİ

Tablo 1.1- YSA-Biyolojik Sistem Elemanları	15
Tablo 2.1-FER2013 Veri Seti Duygu Dağılımı	24
Tablo 2.2-AFFECNET Veri Seti Duygu Dağılımı	26
Tablo 2.3- Veri Setlerinin Özellikleri	27
Tablo 3.1-Akış Diyagramı	39
Tablo 4.1-ResNet152 Eğitim Verisi Doğruluk Sonuçları	46
Tablo 4.2- ResNet152 Test Verisi Doğruluk Sonuçları	47
Tablo 4.3- Inception v3 Eğitim Verisi Doğruluk Sonuçları	49
Tablo 4.4- Inception v3 Test Verisi Doğruluk Sonuçları	50



KISALTMALAR

GAN	: Generative Adversarial Networks
CNN	: Convolution Neutral Networks
CM	: Karışıklık Matrisi
YSA	: Yapay Sinir Ağları
EDA	: ElektroDermal Aktivite
ECG	: ElektroCardioGrafı
EEG	: ElektroEnsefaloGrafı
DÖ	: Derin Öğrenme
ILSVRC	: ImageNet Large Scale Visual Recognition Challenge
JAFPE	: The Japanese Female Facial Recognition
FER2013	: Facial Expression Recognition
KDEF	: Karolinska Directed Emotional Faces
CK	: Cohn Kanade

Üniversite	:	T.C. İstanbul Kültür Üniversitesi
Enstitüsü	:	Lisansüstü Eğitim Enstitüsü
Anabilim Dalı	:	Bilgisayar Mühendisliği
Program	:	Bilgisayar Mühendisliği
Tez Danışmanı	:	Dr. Fatma Patlar AKBULUT
Tez Türü ve Tarihi	:	Yüksek Lisans – Haziran 2022

ÖZET

ÇEKİŞMELİ ÜRETİCİ AĞLAR İLE YÜZ İFADELERİNDEN DUYGU ETİKETLEME

Günümüzün trendi haline gelen insan-bilgisayar etkileşimi ile ilgili çalışmalar giderek artan temalardan biri haline gelmiştir. Özellikle teknolojinin hızla artması ile birlikte pek çok alanda duygu analizi konusuna olan ilginin de arttığı görülmektedir. Bunlara; psikolojik hastalıklar, adli işlemler, eğitim, tıp ve akıllı sistemler örnek gösterilebilir.

İnsan yüz ifadeleri yardımıyla duygu analizi yapılabilmesi oldukça zor, karmaşık ve maliyetli problemlerden biridir. Zor ve karmaşık olmasının en önemli sebebi yüz görüntüleri içerisinde jest ve mimiksel ifadeler barındırmasıdır. Maliyetli olmasının sebebi ise çalışmalarda başarıyı artırılabilmesi için defalarca kez işlemlerin tekrarlanması gerekmektedir. Bu çalışmada insanların yüz görüntüleri üzerinden duygu analizi yapılmaya çalışılmıştır.

Çalışmada AffectNet veri kümesindeki yedi farklı duygusal durum incelenmiştir: kızgın, iğrenme, korkma, mutlu, nötr, üzgün ve şaşırılmış. Bu çalışmada gerçek görüntüler ile Light Weight GAN algoritması kullanılarak elde edilen sahte görüntülerden oluşan yeni veri kümesi elde edilmiştir.

CNN mimarileri yardımıyla duygu analizi yapılmıştır. Önerilen modelin ResNET152 ve Inception V3 mimarileri üzerinde başarı oranı test edilmiştir.

Anahtar Kelimeler: Duygu Analizi, Derin Öğrenme, Görüntü İşleme, CNN, GAN, ResNET, Inception

University	: T.C. İstanbul Kültür University
Institute	: Institute of Graduate Studies
Department	: Computer Engineering
Program	: Computer Engineering
Thesis Advisor	: Dr. Fatma Patlar AKBULUT
Degree	: MA – June 2022
Awarded And Date	

ABSTRACT

EMOTION LABELING FROM FACIAL EXPRESSIONS WITH CONTACT MANUFACTURER NETWORKS

Studies on human-computer interaction, which has become a trend today, have become one of the increasingly popular themes. Especially with the rapid increase in technology, it is seen that the interest in sentiment analysis has increased in many areas. To these, psychological diseases, forensic procedures, education, medicine and intelligent systems can be cited as examples.

Emotion analysis with the help of human facial expressions is one of the most difficult, complex and costly problems. The most important reason why it is difficult, and complex is that it contains gestures and mimic expressions in facial images. The reason why it is costly is that the processes need to be repeated many times in order to increase the performance in the studies. In this study, emotion analysis was tried to be done on people's facial images.

The study examined seven different emotional states in the AffectNet dataset: angry, disgusted, afraid, happy, neutral, sad, and surprised. In this research, a new dataset consisting of real images and fake images obtained using the Light Weight GAN algorithm was obtained. Sentiment analysis was performed with the help of CNN architectures. The performance rate of the proposed model has been tested on the ResNET152 and Inception V3 architectures.

Keywords: Sentiment Analysis, Deep Learning, Image Processing, CNN, GAN, ResNET, Inception

1. GİRİŞ

Çağımızın trendi ve gerekliliği haline gelen insan-makine, insan-robot etkileşimi günümüzdeki teknolojik gelişimin ve yönünün adeta göstergesi gibidir. Birçok amaç için yapımı süren insan bilgisayar etkileşimli teknolojik altyapı ve bunun yanında insanla etkileşim içindeki robotların geliştirilmesindeki hız göze çarpmaktadır. Özellikle yaşlılar [1], çocuklar [2] ve belirli hastalıkların[3] terapisinde de yardımcı olarak kullanılacak olan insan bilgisayar etkileşimli cihazlar günlük hayatın bir parçası haline gelmiştir. Bu da birtakım gereklilikleri de beraberinde getirmiştir; bunların başında etkili bir iletişim gelir ki etkili iletişimin vazgeçilmez unsuru ise duyguların anlaşılabilmesidir [4]. Bu kapsamda geliştirilecek olan sistemlerin insan duygularını anlayabilecek aşamaya gelebilmesi şüphesiz birçok kullanım alanında[5] verimliliği arttıracaktır.

Bu araştırmalar içinde sıklıkla çalışılan duygu tanıma tekniklerinden birisi yüz görüntülerinden duygunun tanınmasıdır. İnsanlar iletişim kurarken duygularını aktarmak için birçok jest ve mimik hareketinden faydalanır. Sevinçli olduğumuz anda gülmemiz, çok mutsuz olduğumuz bir anda ağlamamız buna verilebilecek en basit örnekler arasında yer alır. Dolayısıyla insan duygularını anlayan bir sistem, kişi ile daha rahat bağlantı kurabilecek ve böylece daha iyi bir fizyoterapi, eğitim veya yardım sunabilecektir.

İnsanların sosyal etkileşimlerinde duygularını anlamak oldukça önemli bir konudur. Öyle ki araştırmalar, insan yaşamının yalnızca %10 unun tamamen duygusuz [6] olarak geçebileceği göstermektedir. Teknolojinin hızlı gelişimi: fizyolojik, işitsel ve görsel bilgilerin anlık ve kesintisiz olarak kaydedilebilmesi ve işlenmesindeki ilerlemeler duyguların anlaşılması üzerine yapılan çalışmaları da arttırmıştır. Bunlar ElektroEnsefaloGrafı (EEG) [7], ElektroCardioGrafı (ECG) [8], ElektroDermal Aktivite(EDA) [8] gibi fizyolojik sinyallerden [9], ses sinyalinden [10] yazılı içeriklerden veya yüz görüntülerinden [11] yararlanılarak yapılabilir. Özellikle yüz ifadelerindeki duyguların tanınmasıyla ilgili araştırmalar oldukça zengindir. İnsan beyni gördüğü veya duyduğu ifade ile kişinin ruh haline ilişkin bir, çıkarım yapar [12] ve belirli duygularla eşleştirir [13]. Bu ,çoğu zaman başarılı sonuç üreten bir nevi başarılı bir tahmin algoritması çalıştırmaya benzer. Yüz ifadeleriyle iletilen duygular, insan etkileşimlerini anlamanın önemli bir anahtarıdır.

Mortensen'e [14] göre, diğer insanların duygular yalnızca üçte birini sözler ve ses tonları vasıtasıyla anlıyoruz; geriye kalan üçte ikilik kısım insanların yüzlerinden geliyor. Bu sebeple duygu tanıma veya etiketleme problemleri bu bakış açısıyla, veri toplama kısmına odaklanılır.

Duygusalılığı bir sisteme/makinaya dâhil etmeye çalıştığımızda, duyguları bilgisayar bilimleri açısından da değerlendirmemiz gerekir. Burada, genellikle birlikte kullanılmalarına rağmen tamamen ayrı iki önemli konsept vardır. Bunlardan birincisi insan duygularının tespiti ve etiketlenmesi, ikincisi ise bir algoritma [15] ile duyguların modellenmesidir. Biz bu, çalışma kapsamında birinci kısma odaklanarak duyguların etiketlenmesi için bir çalışma yürüttük, doğrulama yapabilmek için modelleme yöntemi de uygulanmıştır. Duygular, yüz üzerindeki bileşenlerin olası pozisyonları ile hareketlerinden anlaşılır. Örneğin, kaşları, çatmak, alnı kırışmak ve dudak kenarların aşağı dönmesi öfke belirtisi [16] olarak tanımlanır. Bu konuda yapılmış ,çalışmaların öncülüğünü Paul Ekman yapmaktadır. Ekman [17] 10000'den fazla yüz ifadesini sınıflandırarak 3000 civarı duygu ile ilişkilendirmiştir. Ekman'a göre yedi temel duygu bulunmaktadır bunlar; mutluluk, sürpriz, öfke, üzüntü, korku, iğrenme ve nötr 'dür.

Yüzden duygu tanıma alanındaki ilerlemeler her ne kadar cesaret verici olsa da fotoğraftaki aydınlık [18], fotoğrafın çözünürlüğü [19] ,bakış açısı[20], ifadenin kompleksitesi [21], kullanılan aksesuarlar [22] ve benzeri etmenlerin başarıyı önemli ölçüde etkilediği bilinmektedir. Yüzdeki ifadeyi belirli bir duygunun temsilcisi olarak tanımlamak insanlar için dahi zorlu bir süreçtir.

Araştırmalar, farklı insanların aynı yüz ifadesine farklı anlamlar yüklediğini [23] göstermektedir. Bu sebeptendir ki mevcut duygu tanıma [24] , çözümlerinin başarımına dair tartışmalar hala sürmektedir. Belirtmiş olduğumuz pek çok faktör duygu tanıma işlemini zorlaştırmaktadır. Bu faktörleri teknik ve psikolojik olarak ikiye ayırabiliriz. Psikolojik olarak ele alındığında duyguların insan zihninin bir göstergesi [25] olarak kabul edilebilir.

İnsanların duyguları yapmış oldukları konuşmalarda, anlık mesajlaşmalarında vb. gibi pek çok sebepten yüzüne jest veya mimiksel olarak yansıdığı bilinmektedir [26]. Burada meydana gelen psikolojik çıkarımlardan teknik bulgular elde edebilmek için akıllı sistemler [27] günden güne hızla gelişmektedir. Günlük hayatımızın önemli bir parçasını oluşturan mesajlara, epostalara içerisinde bulunan durum bilgisini anlatan emoji ifadelerini [28] üreterek karşı tarafa duygularımızı daha doğru bir şekilde aktarma ihtiyacı duyduk. Böylece artık bütün duyguların bilgisayarlar gibi akıllı sistemlerce öğrenilmesini ve bu öğrendiği bilgiye göre hareket etmesi beklemektedir.

Kişisel bilgisayarlar haline gelen akıllı telefonlar ile pek çok duygu tanıma özelliği getirildi. Örneğin; ses ile aranmak istenen kişinin aranması emrinin verilmesi [29], yüz tanıma sistemi ile ekran kilidinin açılması, parmak izi tanıma[30] ile telefona giriş yapılabilmesi gibi.

İnsanlara ait bu özelliklerin kişisel makinelerde tutulması ile bir nevi makineler ile kişiler hakkında bir bilgi ağı kurulmasına sebep olmaktadır. Örneğin; dinlediğin müziğe göre müziklerin listelenmesi gibi. Bütün bu iletişimin gerçekleşebilmesi için duyguların iyi bir şekilde tanınması ve makinenin bunları algılayabilmesi gerekmektedir. Duygu tahminleme işleminin yapılabilmesi için kullanılabilecek belli başlı yazılımlar vardır. Bunlar makine öğrenmesi algoritmaları ile derin öğrenme algoritmalarıdır.

Makine öğrenimi algoritmaları ile eğitim yapabilmek için büyük ölçekli veriye ihtiyaç bulunmaktadır [31]. Nicelik önemli bir faktörken, nitelikte eşdeğer etkiye sahiptir; aydınlık, fotoğrafın çözünürlüğü, bakış açısı, hatta cinsiyet, yaş ve ırk gibi demografik [32] özellikleri yine sonucu doğrudan etkilemektedir. Bunlar bazen birkaç veri kümesini birleştirerek karma bir veri seti oluşturularak çözülebilmektedir. Düşük hacimli veri depolamayı kolaylaştırırsa da modelin öğrenmesini zorlaştırmaktadır. Bu sebepten dolayı yüz tanıma çalışmaları için ortaya çıkarılan veri setlerinde geliştirmeler yapılırken veri boyutlarına da dikkat edilmektedir [33]. Özetlemek gerekirse duygu analizi pek çok bileşenden oluşmuş bir yapay zekâ problemidir. Bu problemten etkin ve efektif bir sonuç elde edilmesi için girdilerin ve uygulanacak olan algoritmalarının iyi bir şekilde harmanlanması gerekmektedir.

Yapılan çalışmada yüz görüntülerine ait veri setleri incelenmiş ve karma bir veri seti oluşturulmuştur. Veri setinde ne kadar fazla örnek olursa algoritmaların daha iyi sonuç verdiği görülmüştür. Çalışma kapsamında limitli sınıf dağılımına sahip olan veri setlerinden AffectNet kullanılmıştır. Dağılımda düşük örneğe sahip sınıfları dengeli bir yapıya taşımak için derin öğrenme algoritmalarından “Üretken Çekişmeli Ağlar” (GAN) ile sentetik yüz görüntülerinden üretilmiş ve veri seti bu görüntülerle zenginleştirilmiştir. Bu kapsamda literatüre yüz görüntülerinden duygu tanıma çalışması yapabilecek kadar zengin veri kümesi katılması hedeflenmiştir. Bir sonraki aşama olarak herhangi bir etiketi olmayan yüz görüntülerini etiketlemek için bir derin öğrenme mimarisi geliştirilmiştir.

1.1. Problem Tanımı

Makineler çevrelerindeki olan olayları doğrudan algılama yeteneğine sahip değildirler. İnsan ile robotun arasındaki en büyük farklardan biride nesnelerin tanınma yetisidir. İnsanlar bilindiği üzere çevresinde bulunan her bir nesneyi gözleri vasıtasıyla görebilmekte ve bunları beyinde ilişkilendirebilmektedir. İnsan dışındaki makinelerin (robotların) bu ilişkiyi kurabilmeleri için onları yapay zekâ teknikleri [34] ile sistemlerini akıllı hale getirilmesi gerekmektedir. Bu süreçte sistemlerin öncelik ihtiyacı verilerdir. Sistemin gönderilen verileri algılaması ve paylaşılan bilgilere ait özellikleri çıkarılabilmesi beklenmektedir. Eğitim sonunda sistemin gelen verileri doğru bir şekilde sınıflandırma işlemi yapabilmesi gerekmektedir. Duygu analizi yapmak bir örüntü problemini de içinde barındırdığı için oldukça karmaşık bir süreçtir. Bu nedenle problem üzerinde daha verimli sonuçlar alabilmek için araştırmacılar eski yöntemleri kullanmayı bırakıp daha etkili olacağını düşündükleri yapay zekâ algoritmalarına yönelmiştir. Yapılan çalışmalarda [35] görüntü tanıma işlemlerinde makine öğrenmesi ve derin öğrenme yöntemlerinin etkin sonuçlar verdiği görülmüştür. Yüz ifadesi tanıma ve bu konuda doğru çıkarım yapabilmek oldukça karmaşık bir süreçtir. Sistemin doğru sonuçlar üretebilmesi, örneklem sayısı yüksek bir veri seti ve etkili bir algoritma bileşimini, ek olarak verilerin depolanabilmesi ve etkin bir şekilde işlenip çalıştırılabilmesi içinde güçlü bir donanım sistemi gerekmektedir. Derin öğrenme algoritmalarının örneklem sayısı fazla olduğunda daha başarılı eğitim sonuçları ürettiği bilinmektedir. Dolayısıyla burada birtakım problemler ortaya çıkmaktadır.

- Problem 1: Derin öğrenme algoritmasının etkin çalışabilmesi için bu kadar çok sayıda veri nereden ve nasıl bulunacaktır?
- Problem 2: Mevcut olan verilerin dengesiz (inbalance) olması durumu verilerin doğru etiketlenmesini nasıl etkileyecektir?
- Problem 3: Başarım oranını artırılabilmesi için çalışmada hangi evrişimli sinir ağı mimarilerinden yararlanılacaktır?

Tez çalışmasında belirtilen problemleri ele alıp verimli ve etkin çözümler aranmıştır.

1.2. Tezin Amacı

İnsanoğlu doğduğu günden bu yana birbirleriyle iletişim halindedir. İletişim sadece sözlü olarak değil jest ve mimiksel olarak da sağlanmaktadır. İnsanlar birbirlerinin yüz ifadelerinden sözsüz iletişimin çıkarımı yapabilmektedirler. Gelişen teknoloji ve dijitalleşen ortamda ise bu tahminlemelerin makineler tarafından yapılması birçok alanda fark yaratacaktır. Bu bakımdan iletişimin iyi ayırt edilmesi ve gelişen dünyada insan-makine, insan-robot etkileşimine duyguların aktarılması onları daha akıllı hale getirilmesi gerekliliği ortaya çıkmıştır. Bununla birlikte yapay zekâ, robotik, makine öğrenmesi, derin öğrenme gibi konulara olan ilgi artmıştır.

Duygu analizi alanında bir modelle gerçek görüntü sayısını artırma işlemi yapılarak veri setindeki örneklem sayısının ve dağılımının önemi gösterilmek istenmiş ek olarak derin öğrenme algoritması yöntemleri kullanılarak sentetik görüntüler üreterek örneklem dağılımı dengelenmiş bir veri seti oluşturulmuştur. Buna ek olarak etiketi bulunmayan veri setleri için hassasiyeti yüksek bir duygu etiketi üretimine olanak sağlayan derin öğrenme mimarisi geliştirilmiştir açabilmektedir.

1.3. Tezin Organizasyonu

- Bu tez çalışması şu şekilde organize edilmiştir.
- Birinci bölümde problem tanımı anlatılmış, yapılan çalışma tanıtılmış, amacı ve önemi anlatılmış ve literatüre katkısından söz edilmiştir.
- İkinci bölümde CNN ve GAN derin öğrenme algoritmalarından ve kullanılan veri setlerinden bahsedilerek literatür taraması yapılmıştır.
- Üçüncü bölümde, çalışmada kullanılan yöntemler, çalışmada kullanılan teoriler, yaklaşımlardan ve bunların nasıl uygulandığından, amaçlarından bahsedilmiştir.
- Dördüncü bölümde yapılan çalışma ve önerilen yöntemle ilgili modellere ve yöntemin akış diyagramından detaylı bir şekilde bahsedilmiştir.
- Beşinci bölümde ise yapılan testler ve sonuçları detaylı bir şekilde belirtilmiştir.
- Altıncı bölümde kaynakça bölümüne ayrılmıştır.

2. ÖN BİLGİLER ve LİTERATÜR ARAŞTIRMASI

Bu bölümde yapılan araştırma çalışmasının temel tanımları ve araştırma yapılan konuya ait yapılan literatür taraması sunulmuştur. İlk olarak duygu kavramının ne olduğu ve duygu ile ilgili literatür taraması anlatılmıştır. Bu kısmın altında yüz duygu sisteminin kullanım alanları ve yüz üzerinden duygu analizi yapmanın zorluklarından bahsedilmiştir. İkinci kısımda görüntü işleme kavramına yer verilmiştir. Üçüncü kısımda derin öğrenme ve makine öğrenmesi kavramları özetlenmiştir. Dördüncü kısımda CNN algoritmasına yer verilmiştir. Son kısımda tezde kullanılan derin öğrenme algoritması GAN özetlenmiştir.

2.1. Duygu Tanımı ve Literatür Taraması

Duygular, insan davranışının tetikleyicisi olarak kişilerin iletişimde inanılmaz derecede güçlü bir etkiye sahiptir. Yapılan çalışmalar incelendiğinde duyguların: oluşma biçimleri, özellikleri ve çevre faktörlerinden etkileşimi gibi pek çok farklı çalışma [36], [37], [38], [39] yapıldığı görülmüştür.

Tarih boyunca bilim insanları için duygunun anlaşılması kavramı merak ve ilgi çekici bir araştırma konusu olmuştur. Literatüre bakıldığında duygunun pek çok farklı tanımı olduğu görülmektedir. Izard [40] duygunun davranışsal ve öznelikten oluştuğunu duygu tanımına ek olarak duyguların genel kişilik kümesinin alt kümesi olduğunu belirtmiştir. Greenberg [41] ‘ye göre duygu ve düşünce kavramı birbirinden oldukça farklıdır, fizyolojik ve nörokimyasal etkiler duyguların oluşmasının yapıtaşı olduğu tanımında bulunmuştur. Ekman ve Cordara [42] “duyguların evrensel paylaşımlara, kültüre ve bireye özgü olaylara verilen ve ayırt edici özellikleri olan otomatik tepkiler” olarak tanımlamaktadırlar. Literatür incelendiğinde duyguların sınıflandırılmasında 2 ana görüş bulunmaktadır [43]. Bunlardan birincisi, duyguların çeşitli olması yani öfke, korku, mutluluk, öğrenme gibi birbirinden farklı olan duyguları temsil etmektedir. Araştırmalar sonucunda bu görüşleri pek çok kuramcının benimsediği görülmüştür. Bunların başında Darwin [44] gelmektedir. Darwin’e göre duygular sekiz gruba ayrılmıştır bunlar; şaşırma, korkma, öğrenme, aşağılama, öfkelenme, mutluluk, üzüntü ve merak etme olarak gösterilebilir. Bunun yanında; Ekman ve Friesen [45], Aebi [46] Aldinger ve arkadaşları [47], Demetriooff, Porter ve Baker [48]’da bu sınıflandırma yaklaşımına benzer yaklaşımlar kullanmışlardır. Sınıflandırmanın ikinci kolu ise iyi/orta ve düşük / yüksek yoğunluğa göre boyutsal bir yaklaşımı ele alan aşamadır. Bu sınıflandırmayı

başta Wundt, olmak üzere pek çok kuramcı; Plutchik [49] ,Russell ve Fernandez – Dols [50] benzer şekillerde kullanmıştır. Bu yaklaşıma göre örneğin mutluluk için mutlu olmanın şiddeti ve düzeyi farklılık göstermektedir. 1994 yılında Averil [51] her duygunun temel duygu olarak görülemeyeceğini bazı duyguların temel olduğu görüşünü ortaya atmıştır. Modelsel sınıflandırmanın duygulara ilişkin genel bilgiler verdiğini her şekilde kullanılmanın doğru tahminler yapamayacağını ve sınıflandırma yapılırken hatalara neden olabileceğini ifade etmiştir. Plutric [52] ise temel duyguların en yalın şekilde belirtilmesi gerektiğini ve bu bilgiler ışığında organizasyonel bir tanımda bulunarak organizasyonun elemanlarını biyolojik, kültürel(toplumsal) ve psikolojik(ruhsal) açıdan ele alınması gerektiğini belirtmiştir. Bu aşamaların ilki evrim olarak kabul edilmiş, kültürel ve ruhsal boyutta içsel durum olarak birleştirilmiştir. Bu çalışmalar incelendiğinde duyguların sınıflandırılmasında temel olarak iki organizasyon olsa da daha alt seviyelerde pek çok farklı yaklaşım mevcut olduğu görülmüştür. Ekman' a göre her bir duygunun bir başka duyguya göre sözel, jest ve mimiksel ve fizyolojik olarak ayırt edilebilir özelliklere sahiptir. Bu durum ele alındığında daha karışık duygu durumlarının meydana gelebilmesi sağlanmaktadır. Bu karmaşık yapıyı anlayabilmek için öncelikle yüz bölgesinde beliren duyguların ne ifade ettiği ele alınmıştır Ekman ve Friesen [53] .Yüz ifadeleri karmaşık bir yapıya sahip olduğu için içinde sadece duyguları barındırmaz. Ekman [54] 'a göre yüz ifadelerindeki hareketler pek çok farklı duygunun işlevini temsil etmektedir. Örneğin; kaşları çatarak öfkeli görünümünün belirtilmek istenilmesi gibi. Bu örnekten de anlaşılacağı üzere insan yüzünde çok sayıda kritik bilgi bulunur ve bu kritik bilgileri doğru olarak anlamlandırmak oldukça önem kazanmaktadır. Yapılan araştırmalara göre yüz ifadelerinin temsil ettikleri duyguların evrensel niteliğe sahip olduğu kabul edilmiştir Batty ve Taylor[55] , Ekman ve Friesen [56] Elfenbein ve Ambady [57], Keltner ve Ekman [58] , Matsumoto ve Ekman [59],Matsumoto ve Willingham [60].

Yüz ifadelerinin evrensel oluşunun keşfedilmiş olduğu süreçte kültürlerin özellikleri ile ilgili pek çok sayıda çalışma yapılmış ve sonuç olarak da kültürlerin yüzden duygu çıkarımı yapılırken benzer sonuçları işaret ettikleri görülmüştür Ekman ve Friesen [61] vb. Sonuç olarak farklı kültürler üzerine yapılmış olan pek çok çalışma insan yüz ifadelerinin aynı olduğunu ve bu yüz ifadelerinin evrensel olduğunu desteklemektedir. Bu bilgilerin doğru olmadığını düşünen bir grup araştırmacı duyguların tespitinde evrenselliği kabul etmenin eksik sonuçlar doğurduğuna dair açıklamalarda bulunmuşlardır. Crivelli, Jarillo, Russell ve Fernandez [62] . Kültürlerin doğuda ve batıda aynı olmadığından ötürü bazı duyguların da karşılığının

olmadığını ya da eksik olarak temsil edilmiş olduğunu göstermek için bir takım katılımcılar üzerinde çalışmalar yapılmıştır.

Çalışmalarda batı kültüründeki katılımcılardan alınan iğrenme ve korku duyguları, doğu kültüründeki katılımcılardan da istenmiş sonuçlar incelendiğinde doğu kültüründe örneği alınan duyguların batı kültürüne oranla eksik olduğu tespit edilmiştir Jack ve ark. 2009 [63] .

Bahsedilen tespitten sonra insan yüz ifadelerindeki duyguların evrensel olduğu kabul edilse bile kültürel özelliklerin de duygu analizindeki yeri anlaşılmıştır. Ekman'ın Nöro-kültür teorisinde de belirtildiği üzere insanların yüzünde bulunan duyguların evrensel oluşu kabul edilirken, ifadelerin yüzdeki durma süresinin değişebileceği belirtilmiştir. Yüzde duyguların görünme sürelerine göre makro ve mikro ifadeler olmak üzere ikiye ayrılırlar. Makro ifadeler; yüzdeki ifadelerin görünme süreleri saniye bazında incelendiğinde beşte bir (1/5) ile dört (4) saniye aralığında değişim içinde olduğu belirtilmektedir [64]. Mikro ifadeler ise; yüzdeki görünme süreleri makro ifadelere göre daha kısa süreden oluşmaktadır. Saniyenin otuzda biri (1/30)'nden daha kısa sürede oluşurlar Matsumoto ve Hwang [65] .Yüzde oluşan bu ifadeler iletişim sırasında görülmektedir.

Yüz ifadelerini tanımayla ilişkili olabileceği düşünülen bir diğer faktör ise ruh bilimcilerinin (psikologlar vb.) terapilerde onlara tedavi için gelen hastalarının sözel veya görünüşsel (ifadesel) duygularının iyi bir şekilde anlaşılması gerekmektedir. Terapiye gelenler için sadece söylemlerden çıkarım yapmak tedavinin seyrini negatif yönde etkilemesi mümkündür. Bu konuda Ekman ve Friesen [66] uzmanların, tedavi sürecinde sadece sözel ifadelere güvenilmemesi gerektiğini bunun yanında jest ve mimiksel olarak duygularında incelenmesi gerekeceğini ifade etmiştir. Temel olarak bahsedilmek istenen şeyde yüz üzerinde beliren bir takım duygular o kişiyle iletişimin ana ögesidir.

Teknolojinin giderek mühendislik çerçevesinde ilerlemesi tıp alanının da dolaylı bir şekilde etkilenecek gelişmesine sebep olmuştur. Görüntü işleme sistemlerinin yaygınlaşması ile sağlık alanında bir takım sorunların çözülmesine yardımcı olabileceği görülmüştür. Bu alanda yapılmış olan pek çok çalışma mevcuttur. Örneğin; yaşlı bakımı, otistik hastaları tedavisi, çocuk terapisi ve şizofreni tedavisinde yüz analizi sistemlerinin yardımı ile kişilerin ruhsal durumlarının nasıl olduğu çıkarımı ile ilgili çalışmalar literatürde görülmüştür.

2.1.1. Yüz Görüntülerinden Duygu Tanıma Kullanım Alanları

Günümüzde insan – robot etkileşimlerinin giderek hızlanması üzerine artan robot (makine) sayısı hayatımızın her alanında kullanım alanları artmaktadır. Artış ile birlikte bu konudaki çalışmaların sayısı günden güne çoğaldığı görülmektedir. Yüzden duygu analizi amacıyla yapılan çalışmalarda üç farklı basamak bulunmaktadır. Bu basamaklar, görüntü içerisinde olan mevcut yüz kısımlarının belirlenmesi, buna göre özellik çıkarılması ve bu özelliklere bakılarak duyguların sınıflandırılabilmesi işlemidir. Yüz tanıma sistemlerinde öncelikli olarak yüz bölgesinin doğru bir şekilde belirlenmesi gerekmektedir. Görüntü belirlenirken etraftaki(arka plan) verilerden yüzün net bir şekilde çıkarılması gerekmektedir. Bu ayırım yapıldıktan sonra video, içerik tabanlı erişim, çoklu alan takibi gibi pek çok alanda uygulama imkânı bulabilecektir [67] .

İlk basamakta karşımıza yaygın olarak kullanılan Haar - Cascade sınıflandırma algoritması çıkmaktadır. Bu algoritma ile resimlerdeki ve videolarda istenen nesnelere ait yüz ve yüzdeki kısımların belirlenmesi sağlanabilmektedir. Haar sınıflandırma yöntemlerinden sonra görüntü verilerine ait özellikler çıkarılmış olmaktadır. Haar sınıflandırma yönteminin dışında, doğrusal ikili model (LBP) destek vektör makinesi (SVN) yöntemleri de yaygın olarak kullanılmaktadır [68] . İkinci basamağı duygulardan özellik çıkarmak oluşturmaktadır. Bu aşamada yapılan çalışmalar geometrik ve görünüm özelliğine göre ayrılmaktadır [69]. Son kısma gelindiğinde ise mevcut sistemin manuel olmadan (otomatik) bir şekilde sınıflandırma yapılması beklenmektedir.

Bu aşamada devreye literatür incelendiğinde yapay sinir ağları [70] ve derin öğrenme algoritmaları karşımıza çıkmaktadır. Yüz tanıma sistemlerinin otomatikleşmesinden ötürü yapılan çalışmalar ile yüz görüntüleri üzerindeki duyguların daha kolay işlenmesiyle pek çok uygulama alanı ortaya çıkmıştır. Uygulama alanları şu şekilde sıralanabilir: akıllı sistemlerde [71] , belirli hastalıkların tedavisinde [72] ,görsel inceleme uygulamalarında [73] ,güvenlik sistemlerinde [74], okullarda[75] , adli tıp araştırmalarında [76], görme engelli insanlara yardım gibi günlük hayatın pek çok önemli kısmında karşımıza çıkmaktadır. Son zamanlarda büyük teknolojik firmalarında yüz tanıma sistemleri üzerine pek çok çalışma yaptıkları görülmektedir. Örneğin; bilişim dünyasının önde gelen firmalarından Microsoft Firması 2018 yılında geliştirmiş olduğu yüz tanıma uygulaması [77] aracılığıyla fotoğrafların aynı kişiye ait olup-olmadığını gösteren sistemler geliştirmişlerdir. Bu sistemler fotoğraftaki kişilerin yaşlarını ve cinsiyetlerini tahmin etmeye çalışmışlardır. Ayrıca literatürde yapılmış olan pek çok çalışma cinsiyetin [78] duygu tanımadaki etkili bir faktör olduğunu ve kadınların yüzlerindeki yedi temel

duygunun erkeklere göre daha belirgin olarak temsil edildiği [79] belirtilmektedir. Bu görüş kimi araştırmacılar tarafından kabul edilmemiş olup erkekler ile kadınların yüz analizinin eşit yetkinliğe sahip olduğu görüşü savunulmuştur [80].

2.1.2. Yüz Görüntülerinden Duygu Analizi Yapmanın Zorlukları

İnsan – robot etkileşiminin artması ile birlikte akıllı sistemlerin gelişimi hızlı bir şekilde ilerlemiştir. Sistemlerin gelişmesi ile birlikte, robotlar tarafından belirli kapsamlarda izlenebilir ve insanların yüz ifadelerine bakarak duygu durumları çıkarılabilir hale gelmiştir. Fakat bu çalışmaların yapılabilmesi oldukça kompleks bir probleme dönüşmüştür [81]. Duygu analizi düşünüldüğünde akla ilk gelen şey yüz ifadeleri kullanılarak yapılmış çalışmalar olsa da bunun yanında ses, video ve fizyolojik öğelerde içermektedir. Yüz tabanlı yapılan çalışmalarda belli oranlarda başarımlar elde edilmesine rağmen halen tam anlamıyla birçok domainde başarı ile çalışacak sistem ve algoritmalar bulunamamıştır. Başarım oranının tatmin edici olmamasının pek çok sebebi bulunmaktadır. Bunlardan bazıları yazılımsal bazıları donanımsal problemlere dayanmaktadır. Yüz analizinde yapılan çalışmalarda veri setlerinde bulunan verilerin miktarı doğru tahminleme çıkarımını sağlayabilmek için önemli bir yere sahiptir. Çalışmalarda kullanılmış olan mevcut veri setleri örneklem bakımından yetersiz olduğundan bazı duyguların tahminlemelerinin doğru yapılamamasına sebep olmaktadır. Günümüzde derin öğrenme yöntemleri ile görüntüler üzerinde oldukça iyi sonuçlar çıkarmaktadır fakat matrislerin çok boyutlu olması ve vektör çarpımlarının büyüklüğü sebebiyle yüksek seviyede donanımsal altyapı gerektirmektedir. Bu anlamdaki yetersizlikler de görüntü verilerin analizini zorlaştırmakta ve verilerin işleme sürelerini de yavaşlatmaktadır[82]. Buda görüntü işleme çalışmalarının yapılmasını teknik anlamda zorlaştırmakta fazla iş yükü meydana getirmektedir.

2.2. Görüntü İşleme

Görüntü işleme; görüntüleri sayısal veri formatına çevirerek üzerinde çeşitli işlemler yapılabilmesini ve yeniden anlamlandırılarak yeni veriler elde edilmesini sağlamaktadır. Görüntü işleme teknikleri temel olarak çözünürlük kalitesini iyileştirme, görüntü içerisinde bulunan öğelerin algılanması gibi pek çok işleve sahiptir [83]. Herhangi bir görüntünün makine tarafından anlamlandırılabilmesi için onun dijitalleştirilmesi yani sayısallaştırılması gerekmektedir.

Sayısallaştırılma işlemi sayesinde görüntü üzerinde bulunan renkler de sayısallaştırılmış olmaktadır [84] Teknolojinin ilerlemesi ile birlikte akıllı sistemlerin yüz, iris, parmak izi görüntülerinden çıkarımlar yapabildikleri görülmüştür. Bu bilgiler ışığında yüzdeki duyguların anlamlandırılabilmesi bir görüntü işleme problemi ifade etmektedir. Görüntü işleme çalışmaların yaygın bir şekilde kullanılmasından dolayı internet üzerinde açık kaynaklı kütüphaneler bulunmaktadır. Bunlardan OpenCV Kütüphanesi pek çok farklı dilde kullanıldığından yaygın olarak kullanılmaktadır [85].

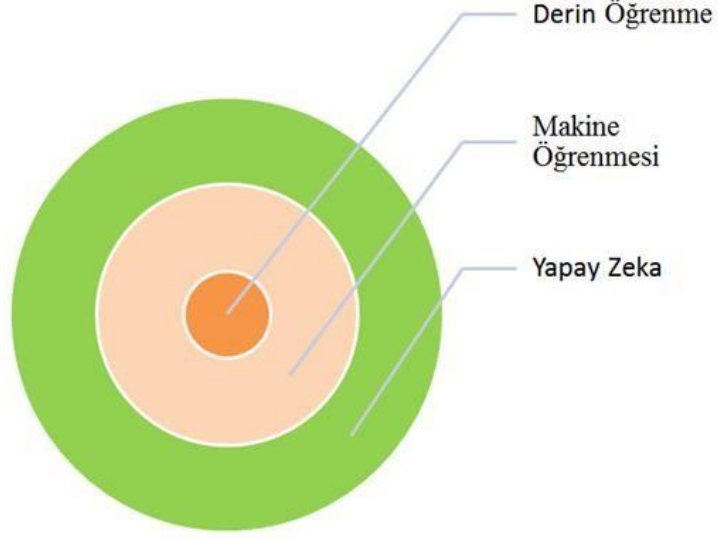
2.3. Derin Öğrenme İle İlgili Ön Bilgiler

Derin öğrenme kavramı yıllar boyunca birçok farklı süreçten geçtiği için halen belirsiz bir terim olarak karşımıza çıkmaktadır. Goodfellow'ın “Derin Öğrenme” [86] kitabında modern bir tanım bulunabilir. Kitaba göre derin öğrenme; dünyadaki kavramları anlayabilen ve deneyimleyebilen bir araç olarak kabul edilmektedir. Skansi'nin kitabına göre [87] derin öğrenme kavramı makine öğrenmesinin bir alt kümesi olarak ifade edilmektedir. İnsanların bir şeyleri öğrenme şekli kazandıkları deneyimlere dayanmaktadır. Kazanılan deneyimler artıkça öğrenme kalitesinde de bir artış görülecektir. Derin Öğrenme (DÖ) bir tür makine öğrenme yöntemi olduğundan insanlarda olduğu gibi deneyimlere dayalı öğrenmeyi sağlar [88].

Bu bölümün başında makine öğrenmesi kavramı kısa bir genel bakış içerisinde değinilmiştir. Ardından temel (vanilya) sinir ağları tartışılarak bunların eğitimi nasıl olduğundan bahsedilecektir. Bu bölümün sonunda derin öğrenme algoritmalarından evrimsel sinir ağları (CNN) ve çekişmeli üretici ağlardan (GAN) detaylı bir şekilde bahsedilecektir.

2.3.1. Makine Öğrenmesi Genel Bakış

Makine öğrenmesi bir problemin elinizdeki mevcut verilere göre en iyi başarımla oranı verecek şekilde modellenmesidir. Temel olarak yapay zekânın alt kümesini makine öğrenmesi, makine öğrenmesinin de alt kümesinin derin öğrenme oluşturmaktadır. Temsili yapay zekâ ve ilişki kümesi Şekil 1.1 'de gösterilmiştir. Makine öğrenmesi algoritmaları mevcut probleme göre üç gruba ayrılır, bunlar; gözetimli (danışmanla) öğrenme, gözetimsiz (danışmansız) öğrenme ve yarı gözetimli öğrenme şeklindedir.



Şekil 1.1- Yapay Zekâ İlişki Kümesi

2.3.1.1. Gözetimli (Danışmanla) Öğrenme

Bu öğrenme yönteminde bir uzman tarafından sisteme bilgi ve kurallar koyması ile oluşturulmuş olan öğrenme yöntemidir. Bir başka deyişle sistem önceden veriler ile eğitilmektedir. $F(x) = y$ fonksiyonel ilişkisi mevcuttur yani her x değerine karşılık gelen bir y çıktı değeri bulunmaktadır. Gözetimli öğrenme algoritmaları şunlardır; Karar Ağaçları, Lojistik Regresyon, Doğrusal Regresyon, Yapay Sinir Ağları (YSA), k-en yakın komşu Algoritması ve Destek Vektör Makinaları (SVN)'dir [89]. Gözetimli öğrenme ile gerçekleştirilen iki temel görev vardır bunlar:

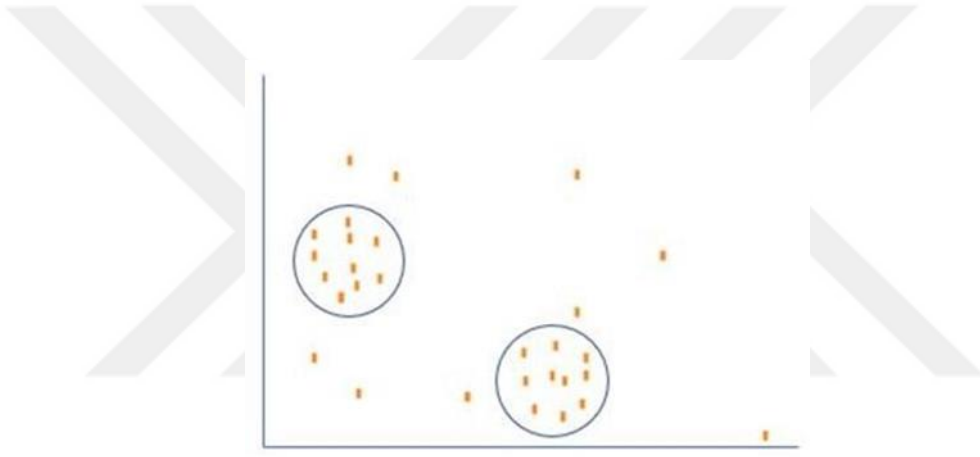
Sınıflandırma: Temel olarak sınıflandırma bir eşleme işlemine yaklaşımdır. Fonksiyonellik olarak incelendiğinde x girdisinden y çıktısı elde edilir. Çıktı değeri genellikle etiketler ya da kategoriler olarak düşünülebilir. Örneğin; sisteme güzel, kaliteli, özel gibi girdiler verilip sonucun pozitif çıkması, sisteme iğrenç, kötü, ucuz gibi girdilerin verilip sonucunun negatif çıkması gibi değerlendirmeler önceden verilir ve daha sonra gönderilen verilerin bu değerlendirmeleri yapması beklenmesidir.

Regresyon: Bir yaklaşıma yaklaşma görevi olarak adlandırılabilir. Sınıflandırmada y çıktısı sözel bir kavramı temsil ederken burada ise miktarlar, boyutlar gibi matematiksel ifadeler mevcuttur.

2.3.1.2. Gözetimsiz (Danışmansız) Öğrenme

Bu öğrenme yönteminde sistemde bir uzman bilgisi mevcut değildir. Dolayısıyla sistem herhangi bir bilgiye sahip değildir. Gönderilen verilere ait herhangi bir şey bilmediğinden dolayı sistem kendi kendine öğrenir ve çıkarım yapması beklenir. Gözetimsiz öğrenme algoritmaları şunlardır; Kümeleme, Birliktelik Kuralları, Temel Bileşen Analizidir.

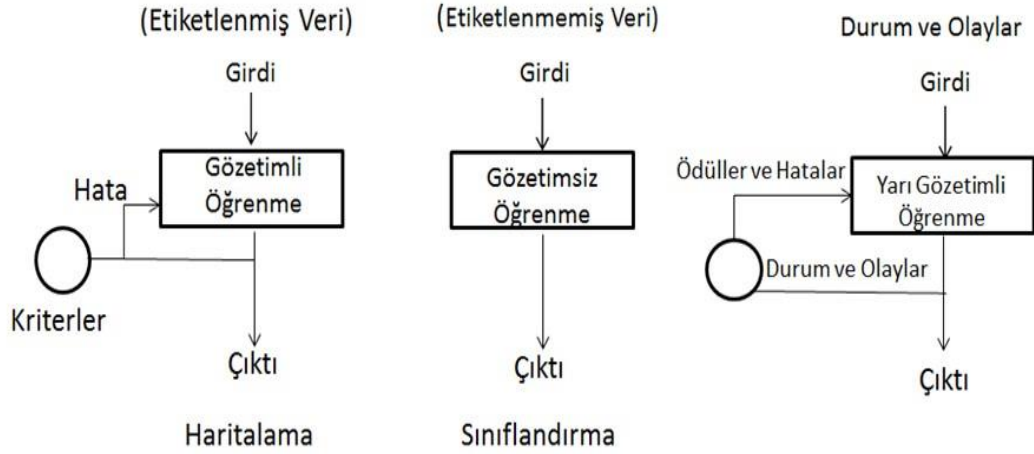
Kümeleme: Matematikteki kümeler mantığı ile ilişkilendirilerek düşünüldüğünde ortak özelliklere sahip verileri gruplandırma işlemi yapılmasıdır. Örneğin bir otele gelen müşterilerin en çok hangi ayda oteli ziyaret ettiği gibi bir gruplandırma işlemi yapılması isteniyor. Öyleyse burada bir kümeleme problemi söz konusu. Burada yapılması gereken aynı ayda gelen müşteri verilerin gruplama yapmaktır. Şekil 2.1 de Kümeleme örneği gösterilmiştir.



Şekil 2.1- Kümeleme

2.3.1.3. Yarı Gözetimli Öğrenme

Bu öğrenme yöntemi ise gözetimli ve gözetimsiz öğrenme yöntemlerinin karma bir hali olduğu düşünülebilir. Gözetimli öğrenme için kullandığımız $F(x) = y$ formülündeki her x değeri için bir y sonucu üretiliyordu, burada ise bazı x değerlerinin çıktılarının etiketleri üretilmiş olduğu bilinmektedir. Her x değerinin çıktısı olan y değerinin etiketinin bilinmesi oldukça pahalı bir işlemdir. Elimizde etiketi olmayan verinin tutulması ve artırılması kolaydır. Bu öğrenme yönteminin avantajı verinin etiketleri ile ilgili tahminleme yapılacağında gözetimli öğrenme şeklinde davranması, sınıfları ve içeriği öğrenmek istediğimizde gözetimsiz şeklinde davranabildiğidir [90]. Üç farklı MÖ yöntemleri Şekil 2.2’de gösterilmiştir.



Şekil 2.2- Makine Öğrenmesi Yöntemleri

2.3.2. Veri Ön İşleme

Makinenin öğrenim şekillerinin öğrenme sürecini etkileyen pek çok etmen vardır. Bunların başında makinenin boyutu gelmektedir. Ayrıca kullanılan veri setleri ve veri ön işleme teknikleri de öğrenim hızını etkileyen faktörlerdendir [91]. Ağ üzerinde veri besleme eğitimi işlemi olmadan önce ağı gelen veriler işlenir. Veri ön işlemeye için oldukça yaygın olan form öncelikle veri setinden ortalamanın çıkarılmasıdır. Veriler her boyutta orijin (0,0) etrafında merkezlenerek bir geometri oluştururlar. RGB(red-green-blue) görüntüler içi tüm görüntüler eklenir ve her kanal için ortalama değer hesaplanır, ardından ilgili ortalama o kanaldaki bütün piksellerden çıkarılır. Çok farklı özellik aralıklarıyla başa çıkmak için uygulamanın başka bir yöntem ise örnekleri farklı boyutlarda ölçeklendirmektir. Denklem 2.1, 2.2 ve 2.3 veri ön işlemede için kullanılan bazı yöntemleri gösterir. Denklemdeki x girdi değerini, u değeri giriş örneğinin ortalaması, σ değeri ise standart sapmayı ifade etmektedir.

$$\text{Min} - \text{Max} \text{ Ölçeklendirme } xi = \frac{xi - \min(Xi)}{\max(Xi) - \min(Xi)} \quad (2.1)$$

$$\text{Normalizasyon } xi = \frac{xi}{||xi||} \quad (2.2)$$

$$\text{Standartlaşma } xi = \frac{xi - \mu x}{\sigma x} \quad (2.3)$$

2.3.3. Yapay Sinir Ağı

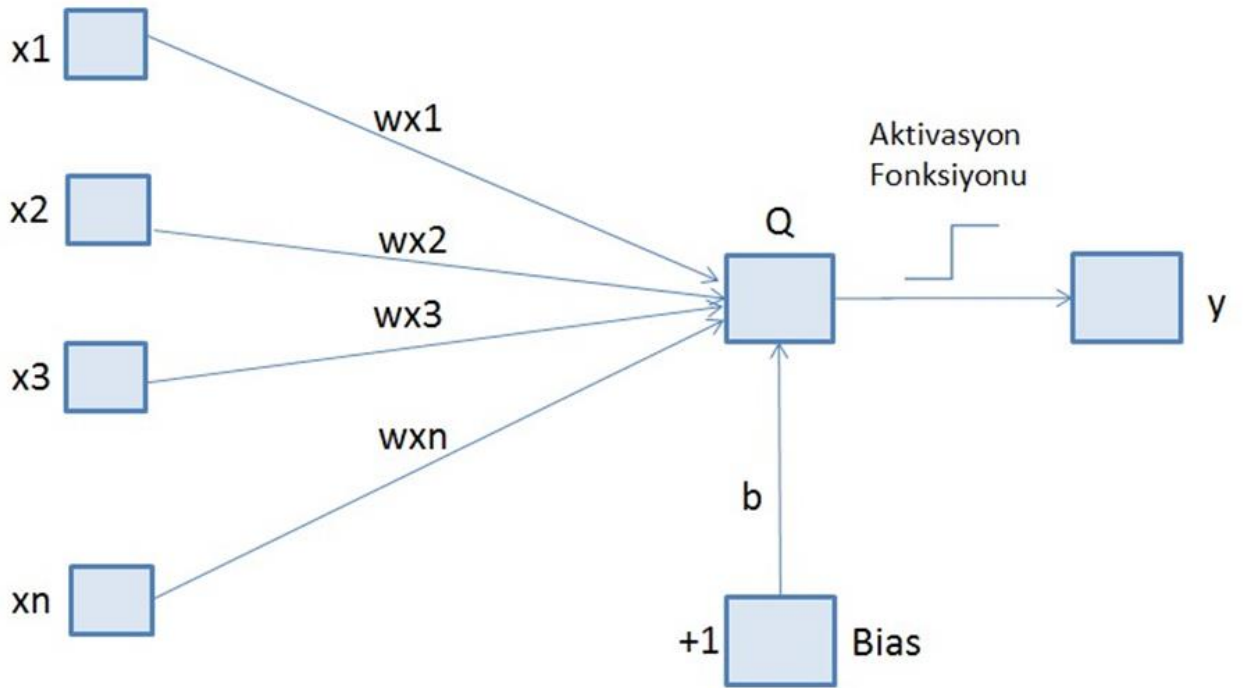
1943 yılında, McCulloch-Pitts tarafından yazılan bir makalede insandaki sinir sisteminin çalışma prensibinden esinlenerek mantıksal hesaplama yapabilen bir model ortaya konulmuştur [92]. Derin öğrenme temel olarak YSA'nın daha kompleks bir halidir. Derin öğrenmeyi iyi bir şekilde anlayabilmek için YSA'nın çalışma şeklinin iyi bir şekilde analiz edilmesi gerekmektedir. YSA'lar temelde bir insan vücudundaki sinir ağının makinelerin anlayabileceği ölçüde nasıl matematiksel olarak modellenebileceğidir. İnsan vücudu çok sayıda sinir hücresinden meydana gelmektedir. Sinir hücreleri yani nöronlar sayesinde düşünme, konuşma gibi pek çok eylem insanlar tarafından yapılabilmektedir. Nöronlarda bulunan sensorlar etrafındaki verileri toplar bu bilgileri bir başka nörona iletirler [93]. Her bir nöron hücresi bir başka nöronun girdisi olacak şekilde karmaşık bir düzen mevcuttur. Bu sebepten dolayı YSA'larındaki algılayıcılar insan beynindeki nöronları taklit etmektedir. YSA bileşenlerinin biyolojik sinir sistemi karşılıkları aşağıdaki Tablo 1.1 de gösterilmiştir;

Yapay Sinir Ağı Elemanları	Biyolojik Sinir Ağı Elemanları
İşlemci	Nöron
Toplama İşlemi	Dentrit
Taşıma İşlemi	Hücre Gövdesi (Çekirdek)
Yapay Nöronların Çıkış Kısmı	Akson
Ağırlıklar	Sinapsis

Tablo1.1 YSA-Biyolojik Sistem Elemanları

Şekil 2.5'de her bir $X = \{x_1, x_2, x_3, \dots, x_n\}$ değeri input değerlerini, $W = \{w_1, w_2, w_3, \dots, w_n\}$ ağırlıkları temsil etmektedir. Hücre gövdesi olan Q değeri, input değerleri ve o inputa ait olan

ağırlık değerlerinin çarpımı ile b (bias) değerinin toplamından oluşmaktadır. Bias değeri skaler bir değerden oluşmaktadır yani 2 boyutlu sistemler üzerinde kaydırma yapmaya sağlanmaktadır. Matematiksel olarak formalize edilirse $Q = (X_n \cdot W_n + b)$ 'dır. Q çıktı değeri aktivasyon eşiği (genellikle sıfır) aşarsa algılayıcı etkinleşir ve bir y çıktı değeri göndermektedir. Çıktı değeri y Şekil 2.3 de belirtildiği üzere Q değerinden elde edilen sonucun aktivasyon fonksiyonu ile değerlendirilip elde edilmesinden oluşmaktadır. YSA modelinden anlaşıldığı üzere derin öğrenme çalışmaları verilerin üst düzey soyutlanabilmesi için ortaya bir model çıkarılmaya çalışılmasıdır[94].



1, ise $Q(x \cdot w + b) \geq$ eşik değeri

-1, diğer

Şekil 2.3- Nöron Ağ Yapısı

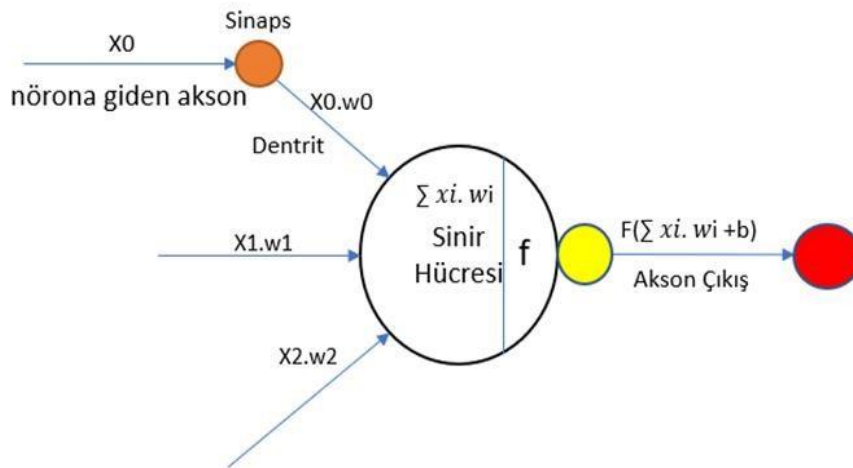
Girdiler: Sinir ağının bilgileri topladığı kısımdır.

Ağırlıklar: Yapay sinir ağında giriş değerlerinin çıkış üzerinde ne kadar etkiye sahip olduğunu göstermektedir. Örneğin; elimizde 4 giriş değerine sahip bir sistem olsun giriş değerleri x_1, x_2, x_3, x_4 olarak düşünölsün. Döğüm 3'ten döğüm 4 'e kadar olan ağırlık daha büyük olduğunu varsayalım o zaman toplam fonksiyonunda elde edilecek değeri x_3 için daha büyük olacaktır. Bir başka deyişle nöron açısından bakıldığında nöron 3 'ün nöron 4'e göre sistemde daha büyük bir değeri ve etkiye sahip olduğu anlamına gelmektedir eğer ağırlık değeri pozitif değilse negatif değeri içeriyorsa giriş değeri artması çıktı değeri düşmesine neden olacaktır.

Toplam Fonksiyonu Σ : Girdi ve ağırlık değeri matematiksel olarak hesaplandığı yerdir. Formöl incelendiğinde; $\Sigma x_i \cdot w_i$ n tane girdi ve onların her birine ait olan ağırlıklarının toplamıdır.

Aktivasyon Fonksiyonu: Sistemindeki nöronların çıkış değeri belirlenen aralıklar içinde sınırlandırabilmek için kullanılmaktadır. Fonksiyona bağılı olarak (0- 1) ya da (-1 1) arasında değeri alabilecek şekilde seçim yapılır. Bu sebepten ötürü aktivasyon fonksiyonu lineer (doğrusal) değildir. Aktivasyon fonksiyonu çeşitleri mevcuttur bunlardan bazıları; Basamak (Step) fonksiyonu Doğrusal (Lineer) fonksiyon, Sigmoid Fonksiyon 'dur.

Çıktı: Aktivasyon fonksiyonu sonucu ortaya çıkan değeri. Detaylı olarak YSA çalışma prensibi aşağıdaki Şekil 2.4 de gösterilmiştir.



Şekil 2.4 -YSA Çalışma Prensibi

Derin öğrenme sistemlerinde alt soylardan üst soylara özelliklerin geçmesi alt-üst arasında bir hiyerarşik yapısal düzen oluşturmaktadır. Böylece özellikler, soyutlamanın farklı düzeylerine karşılık gelen birden fazla özelliğin seviyesini öğrenmiş olmaktadır. Bir görüntü için özellik düşünüldüğünde; piksellerin boyutu, yoğunluk değerleri, kenar kümeleri gibi nicelikler düşünülebilmektedir. Belirtilen özelliklerden bazıları veriyi daha iyi anlamamızı sağlayabilmektedir [95]. Verilerdeki özellikleri elle tek tek çıkarmak çok uzun ve zaman alan bir iştir. Derin öğrenme algoritmaları ile veriye ait hiyerarşik özelliklerin çıkarımı işlemleri etkin ve hızlı bir şekilde yapılabilmektedir [96].

Algoritmaların çalışma prensibinde genellikle veri seti üzerinde iki başlık altında çalışılır bunlar; Eğitim veri seti (Training) ve test veri seti (Testing)'dir. Eğitim veri setinin amacı sinir ağı ağırlıkları üzerindeki matematiksel hesaplamalar yapmak, test veri setinin amacı ise hesaplanan ağırlık değerlerinin doğruluğunu gizli tutulan veriler ile doğru olup olmadığını kontrol etmektir [97]. Çok katmanlı algılayıcılar Şekil 2.5 de gösterilmiştir.

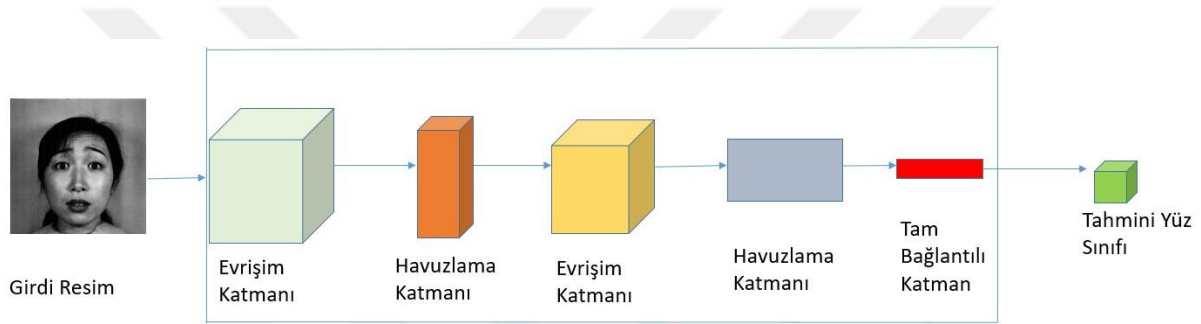


Şekil 2.5- Çok Katmanlı Algılayıcılar Çalışma Prensibi

Buradaki geri yayılım algoritması öncelikle değişkenler eğitim ağına verilir. Değişkenlerin üzerinde etkileşim halinde olan nöronların atanmış olan ağırlık değerleri ile çarpılıp toplandıktan sonra bias (b) değerinin eklenmesiyle gizli katmanda(hidden) bulunan nöronların aktivasyon fonksiyonu ile bir sonraki katmandaki nöronlara iletilir. Elde edilen değerler, çıkış katmanındaki nöronlar için girdi özelliğine sahiptirler. Son olarak çıkış katmanında da aynı işlemler yapılarak bir çıktı elde edilir ve elde edilen çıktı gerçek değer ile eşleştirilir. Bu işlemten sonra hata değeri tespit edilir [98].

2.3.4. Evrişimli Sinir Ağları (CNN)

Derin sinir ağlarının çeşitlerinden biri olan evrişimli sinir ağları (CNN) detaylı bir şekilde açıklanmıştır. Konvolüsyonel sinir ağları olarak da adlandırılan CNN çok katmanlı algılayıcıların bir çeşididir. İlk olarak Hubel ve Wiesel tarafından hayvanların görme merkezindeki yapı ele alınarak ortaya bir model konulmuştur. 1990’larda Lecun ve arkadaşları LeNet-5 adında elle yazılmış rakamları sınıflandıran çok katmanlı bir yapı geliştirmişlerdir [99]. Doğal dil işleme, görüntü ve ses işleme gibi pek çok alanda kullanılmaktadır. Yapısal olarak çok sayıda katmandan oluşan CNN’ler her bir katmanda o probleme ait bir özellik öğrenir ve bu özellikleri bir sonraki katmana çıktı olarak gönderirler. Şekil 2.6 da CNN çalışma modeli gösterilmiştir.



Şekil 2.6- CNN Çalışma Modeli

İşlevsel olarak bir filtrenin (matrisin) görüntünün pikselleri üzerinde belli oranlarda kaydırılması ve görüntü üzerindeki piksellerin bitinceye kadar devam eder. CNN yapıları nanlineer (doğrusal olmayan) fonksiyonları içerirler. Evrişim katmandaki filtre yapısı öğrenebilir şekildedir bununla birlikte bu katmanda filtre boyutu, filtre sayısı ve adım aralığı önemli özelliklerdendir [100]. Şekil 2.6 da gösterildiği üzere CNN ‘de her katmanda farklı filtreler (matrisler) kullanılarak farklı özellikler öğrenilmiş olur. Bu sayede görüntünün her pikselindeki özellikler elde edilmiş olur [101].

CNN’lerin yapısını; evrişim katmanı (Convolution Layer), havuzlama katmanı (pooling layer), aktivasyon fonksiyonları (aktivation functions), tam bağlantı katmanı (fully connected layer), kayıp fonksiyonları (loss functions), düzenleme (regularization) ve optimizasyon (optimization) oluşturmuştur. Örnek bir evrişim işlemi Şekil 2.7 de evrişim süreci gösterilmiştir.

Burada A, iki boyutlu görüntüyü, B i X j boyutundandık filtre özelliğini ifade etmektedir. Filtreler kare matris olacak şekilde istenilen boyutlarda seçilebilir. Örneğin; 2X2,3X3,6X6,gibi.

A = 7X7

0	1	1	1	1	1	1
0	0	1	1	1	0	0
0	0	0	1	1	1	0
0	0	0	1	1	0	0
0	0	1	1	0	0	0
0	1	1	0	0	0	0
0	1	1	0	0	0	1

B = 3X3

1	0	1
0	1	0
1	0	1

×

=

A*B = 5X5

1	4	4	5	3
1	2	4	3	3
1	2	3	4	1
1	3	3	1	1
3	3	2	1	1

Şekil 2.7- Evrişim İşlemleri

Derin ağ yapısında maliyet oldukça önemli bir kavramdır. Bu sebepten ötürü havuzlama katmanı önemlidir. Ağda meydana gelebilecek olan değişimlere ağı güçlendiren bazı işlemler yapılmaktadır. Bunlar çoğunlukla ortalama, maksimum, minimum işlemleridir. Şekil 2.8,2.9 ve 2.10 da sırasıyla ortalama, maksimum ve minimum havuzlamaya örnek olacak işlemler gösterilmiştir.

4*4

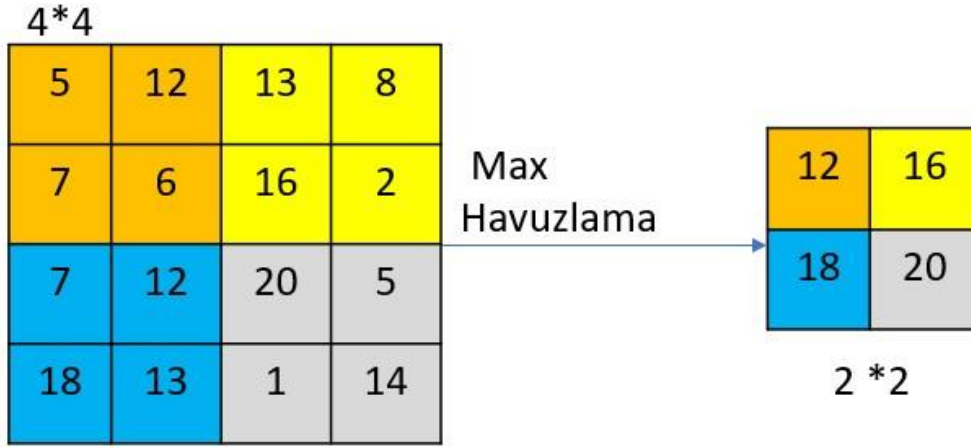
3	12	11	8
7	6	16	9
7	12	20	15
16	13	11	14

Ortalama
Havuzlama

7	11
12	15

2*2

Şekil 2.8- Ortalama Havuzlama İşlemi

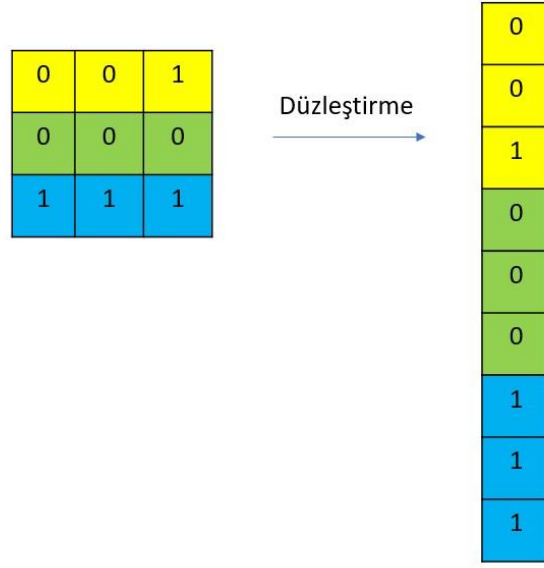


Şekil 2.9- Maximum Havuzlama İşlemi



Şekil 2.10- Minimum Havuzlama İşlemi

CNN'in performansını etkileyen önemli bir elemanda aktivasyon fonksiyonudur. Yaygın olarak kullanılan fonksiyon türleri; Rectified linear unit (ReLU), exponential linear unit (ELU) ve sigmoid'dir. Bir başka katman ise tam bağlantılı katmandır. Bu katmanda adından anlaşılacağı üzere ağdaki sinirler önceki katmandakilerle bağlantılı bir yapıya sahiptir. Üzerinde çalışılan veriler bu aşamaya gelmeden önce pek çok aşamadan geçerler. Bu işleme düzleştirme (flattening) denilmektedir. Örnek bir düzleştirme işlemi Şekil 2.11 de gösterilmiştir.



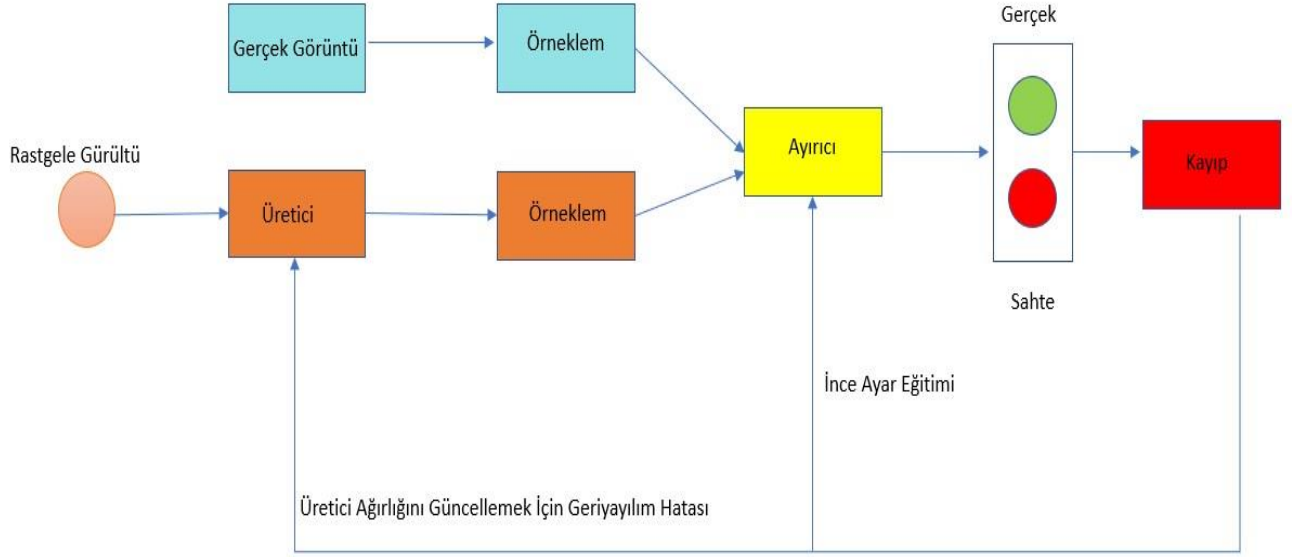
Şekil 2.11- Düzleştirme İşlemi

CNN için bahsedeceğimiz son katmanda kayıp katman (Loss Layer) 'dır. Bu katmanın işlevi eğitim aşamalarında tahmin edilmiş veriler ile gerçek verilerin etiketleri arasında oluşan sapmanın incelemesini yapmaktır. Yapılan araştırmalarda yaygın olarak Softmax kayıp fonksiyonu kullanılmaktadır.

2.3.5. Çekişmeli Üretici Ağlar (GAN)

İlk olarak Goodfellow ve arkadaşları [102] tarafından 2014 yılında ortaya atılmıştır. Bu ağların geliştirilmesi derin öğrenme üzerinde yapılan görüntü işleme çalışmalarına pozitif yönde ivme kazandırmıştır. GAN, gerçek görüntüler ile ayırt etmenin imkânsız olduğu sentetik görüntü üretmeyi sağlayan modellerdir. Çekişmeli üretken ağların çalışma prensibini iki algoritma meydana getirmiştir. Bu algoritmalar üretici(G) ve ayırcı (D) olarak ayrılmaktadır. Üretici, herhangi bir gürültü ile belli başlı işlemlerden yeni görüntüler üreterek bu görüntüler gerçek görüntüymüş gibi ayırcı ağı kandırabilecek yapay görüntüler üretir. Ayırt edici ise gerçek görüntü ile sahte görüntüyü ayırmaya çalışmaktadır. Sistem çalışırken ayırcı üreticinin

gönderdiği görüntülere not vermektedir. Bu notlara bakılarak üretici ağıdan gelen sahte görüntüler gerçek görüntülere benzemeye çalışır bu işlemler bazen defalarca tekrar (epoch) gerektirmektedir [103].GAN sistem mimarisi Şekil 2.12’de gösterilmiştir.



Şekil 2.12 – GAN Sistem Mimarisi

2.3.6. Yaygın Kullanılan Veri Setleri

Derin öğrenme ile görüntü işleme konusuna olan ilginin giderek ilerlemesi üzerine pek çok veri seti ortaya çıkarılmıştır. Görüntü işleme çalışmalarında yaygın olarak kullanılan veri setlerinden bazıları şunlardır; CK (Cohn Kanade), FER2013(facial emotion recognition), JAFFE, KDEF (Karolinska Directed Emotional Faces)’dır. Çalışmamızda AffecNet veri setinden yararlanılmıştır.

2.3.6.1. FER2013 (Facial Expression Recognition) Veri Seti

FER2013 veri seti, insanların yüz ifadesindeki duygunun iyi bir şekilde nasıl tahmin edildiğini ortaya çıkarmak için düzenlenmiş olan bir yarışmada tanıtılmıştır. Bu veri seti Pierre

Luc Carier ve Aaron Courville tarafından oluşturulmuştur [104]. Veri seti toplamda 35887 yüz ifadesi içermektedir. Verilerin eğitim için 28709'u, test için 3589'u ve doğrulama için 3589'u kullanılmıştır. Ayrıca veriler piksel bakımından 48 X 48 ölçekli ve gri tonlu yüz ifadelerinden meydana gelmiştir. Çalışmalar Google Görsel Aramalar (Google Image Search) API 'sinden bulunmuştur[105]. Arama işleminde kullanılan duygu içeren 184 referans sözcüğün (mutlu, üzgün, şaşırılmış, nötr vb), yaş, cinsiyet, ırk gibi yan sözcüklerle birleştirilmesiyle 600'e yakın sözcük dizisi meydana gelmiştir. Google Görsel Aramalar üzerinde yapılan her bir sorgudan dönen ilk 1000 yüz görüntüsü bir sonraki adım için depolanmıştır [106]. Toplanan veriler açık kaynak kodlu görsel resim kütüphanesi olan (OpenCV) üzerinde yüz görüntülerinin sınırları belirlenmiştir. Bu işlemlerden sonra yanlış etiketlenmiş görüntülerin olduğu fark edilmiştir; örneğin üzgün görselleri içerisinde mutlu görselinin bulunması gibi. Bir grup araştırmacı tarafından bu görseller ayrıştırılmış, kırpma yöntemi ile düzeltilmiştir bir nevi filtreleme işlemi yapılmıştır. Bahsedilen veri seti 7 temel duyguyu içermektedir. Etiketleri sırasıyla; 0= Kızgın, 1 =İğrenme, 2 = Korku,3= Mutlu,4=Üzgün,5= Sürpriz,6 = Nötr 'dür. Son olarak yarışmada Toronto Yüz Veri seti ile ortaya atılan yüz görüntüleri eşleştirilmiştir. Çalışmanın sonunda 35887 yüz görüntüsünün dağılımı ve yüzdesi Tablo 2.1'de gösterilmiştir.

Etiket-Anlamı	FER2013 Veri Setinde Bulunma Sayısı	%Yüzdesi
0-Kızgın	4953	%13.8
1-İğrenme	547	%1.5
2-Korku	5121	%14.2
3-Mutlu	8989	%25.4
4-Üzgün	6077	%16.9
5-Sürpriz	4002	%11.1
6-Nötr	6198	%17.2

Tablo 2.1 FER2013 Veri Seti Duygu Dağılımı

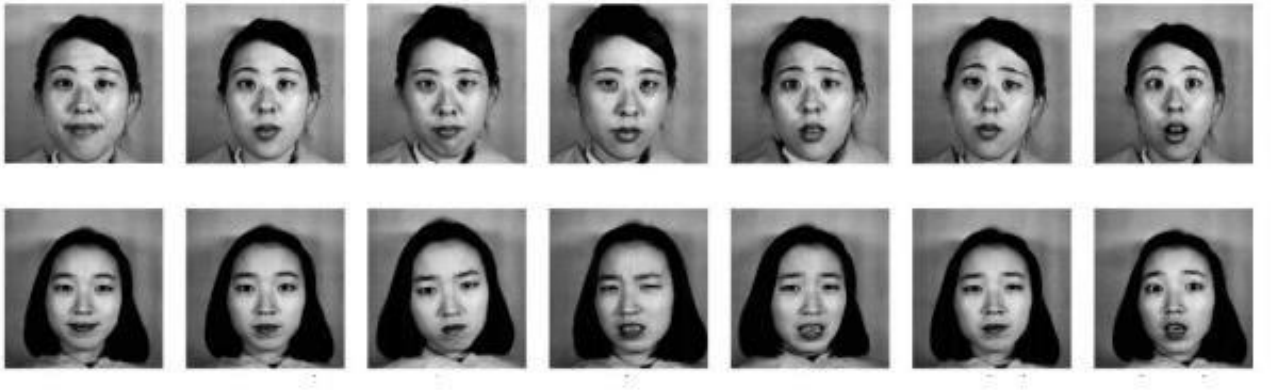
FER2013'ün elde edilen sonuçlara bakılarak insan duygu analizinde $\%65 \pm 5$ başarı gösterdiği görülmüştür [107]. FER2013 veri setine ait örnek görüntüler Şekil 2.13 'de gösterilmiştir.



Şekil 2.13 – FER2013 Veri Seti

2.3.6.2. JAFFE (The Japanese Female Facial Recognition) Veri Seti

JAFFE veri seti, 1997 yılında Japonya'nın önde gelen üniversitelerinden biri olan Kyushu üniversitesinin psikoloji bölümünde hazırlanmıştır [108]. Erişime açık olan veri seti 10 kadın Japon modelden alınan 213 farklı gri tonlamalı yüz ifadesinden oluşmuştur. Modellerin her birinden 7 (6 temel duygu + 1 nötr) farklı duygu ifadesinin temsil edilmesi beklenmiştir. Her bir görüntünün çözünürlüğü 256x256 olarak ayarlanmıştır. [109] Aşağıda JAFFE veri setinden bazı görüntüler Şekil 2.14 de gösterilmiştir.



Şekil 2.14 – JAFFE Veri Seti Duygu

2.3.6.3. KDEF (Karolinska Directed Emotional Faces) Veri Seti

KDEF veri seti Lundqvist ve ekibi tarafından 1998 yılında ortaya çıkarılmıştır [110]. Çalışmaya model olarak 70 denek katılmış olup deneklerin yarısı erkek, yarısı kadından oluşmuştur. Veri setinde 7 temel duyguyu içeren 562 X 762 piksel çözünürlüklü 4900 JPEG formatında görüntüsü mevcuttur [111].

Bu veri setine ait örnek görseller Şekil 2.15 'de gösterilmiştir.



Şekil 2.15 – KDEF Veri Seti Duygu

2.3.6.4. AFFECNET Veri Seti

AffectNet veri seti; üç büyük arama motoru (Google, Yahoo, Bing) üzerinde altı farklı dilde (İngilizce, İspanyolca, Portekizce, Almanca, Arapça ve Farsça) 1250 duygu ile anahtar kelimeler kullanılarak arama motorlarından sorgulanarak elde edilmiştir [112].

Duygu	Yüz Görüntü Sayısı
Nötr (Neutral)	75374
Mutlu (Happy)	134915
Üzgün(Sad)	25959
Şaşırmış(Suprise)	14590
Korkmuş(Fear)	6878
İğrenme(Disgust)	4303
Kızmak(Anger)	25382
Hor görmek (Contempt)	4250
İfadesiz (None)	33588
Belirsiz(Uncertain)	12145
Hiçbir yüz (Non-face)	82915
Toplam	420299

Tablo 2.2 AFFECNET Veri Seti Duygu Dağılımı

Toplamda 1.000.000 dan fazla yüz görüntüsü içermektedir. On iki uzman tarafından elde edilen veri kümesinin yaklaşık yarısını yedi farklı yüz ifadesi cinsinden, değerliği ve uyarılma yoğunluğu bilgileri manuel olarak etiketlendi. AffectNet gerek kategorik modelde görüntüleri sınıflandırmak gerekse uyarılma yoğunluğunu açısından açık ara en büyük yüz görüntülerini içeren veri kümesidir. Kategorik modeldeki görüntüleri sınıflandırmak ve değerlik ile uyarılma yoğunluğunu tahmin etmek için iki temel derin sinir ağı kullanılır. Yüz ifadesi analizi çalışmalarında yapılmış olan çeşitli değerlendirme kriterleri derin sinir ağının makine öğrenmesi yöntemlerine göre daha iyi bir performans gösterebilmektedir.

2.3.6.5. Veri Setlerinin Özelliklerinin Kıyaslanması

Görüntü işleme problemlerinin daha iyi çözülebilmesi için ortaya pek çok veri seti çıkarılmıştır. Ortaya çıkarılmış olan her veri seti birbirinden farklı özelliklere sahiptir. Bu kapsamda ortaya en iyi sonuçları çıkarabilmek için farklı özelliklerde veriler kullanılmış bunlara en uygun olabilecek model üretme çalışmaları yapılmıştır. Tablo 2.3 de bahsedilmiş olan veri setlerinin özellikleri gösterilmiştir.

Veri Seti	Katılımcı Cinsiyeti	Veri Sayısı	Çözünürlük	Format
FER2013	Karma (E+K)	35887	48X48	JPEG
JAFFE	Kadın(K)	213	256X256	TFF
KDEF	Karma (E+K)	4900	562X762	JPEG
AFFECNET	Karma (E+K)	420299	224X224	numpy

Tablo 2.3 Veri Setlerinin Özellikleri

2.4. Veri Setleri ve Derin Öğrenme Algoritmalarının Literatür Taraması

Bu tez çalışmasında derin öğrenme yöntemi ile insan yüz görüntülerinden duygu tanıma işlemlerine ait literatür taraması gerçekleştirilmiştir. Duygu analizi çalışmaları bireylerin biyokimyasal yani içsel dünyasında meydana gelen tepkinin dışa yansımasıdır [113]. Bu sebepten ötürü duyguların anlatmak istediği ifadeleri anlamak oldukça zordur. Çalışma alanının büyüklüğünden ötürü, literatürde duygu analizi ile ilgili oldukça fazla olduğu tespit edilmiştir. Ruh hastalıkları, biyotittik işlemler, akıllı ev sistemleri ve güvenlik gibi pek çok farklı alanda

alıřmalar yapıldığı grlmřtr. Duygu tanıma sistemlerini temel olarak iki ana paraya ayırabiliriz. Bunlar dinamik(hareketli) ve statik (duraėan) kavramlardır. Dinamik kavramlar video gibi hareketli ierik ieren veri setleridir. Tez alıřmamda statik kavramlar yani yz resimleri veri seti zerinden alıřmalar saėlanmıřtır. Arařtırmalar yapılırken sık kullanılan imge tabanlı veri setlerinden bazıları řunlardır; Toronto Face Database (TFD) [114], Facial Expression Recognition (FER2013) [115], The Static Facial Expressions in the Wild (SFEW) [116], The Radboud Faces Dataset (RaFD) [117], Karolinska Directed Emotional Faces (KDEF) [118] ve The Japanese Female Facial Expression (JAFPE) [119]'dir.

2014 yılında Devries ve arkadaşları yz simgelerinin ve ifadesinin ok grevli ėrenimi ile ilgili yapmıř oldukları alıřmada TFD (Toronto yz veri tabanı) kullanarak yzdeki iřaretlere ve konum řekillerine gre grntleri elde etmeye alıřmıřlardır. TFD veri setini kullanırken evriřimli sinir aėı mimarisinden yararlanmıřlardır. Veri seti zerinde yapılmıř olan test sonuları %85,13 oranında bařarı gstermiřtir [120].2015 yılında Khorrami ve arkadaşları nerdikleri sıfır nyargılı CNN modelde TFD veri setini kullanarak %88,6'lık son teknoloji performans bařarımını elde etmiřlerdir [121]. 2017 yılında Ding ve arkadaşları ince ayarlı CNN kullanarak son teknoloji bařarısını geerek TFD veri setindeki bařarım oranını %88,9 oranına ıkarmıřtır. Yaygın olarak kullanıla bir diėer veri seti FER2013'dr. Fer 2013 veri seti 7 farklı duygudan oluřmaktadır. Bunlar; kızgın, mutlu, zgn, řařırmıř, iėrenme, srpriz ve ntr 'dr. Bu veri seti zerinde yapılan CNN tabanlı alıřmalar dan bazıları; Kim ve ekibi[122], Guo ve ekibi [123],Zhang ve ekibi[124], Tang [125], tarafından gerekleřtirilmiřtir. Yukarıdaki alıřmalar detaylı olarak incelendiėinde en ok bařarı eldesi Pramerdorfer ve Kampel [126] tarafından %75.2 olarak elde edilmiřtir.

Hollanda Radboud niversitesi Nijmegen Davranıř Bilimleri Enstits tarafından hazırlanıř bir diėer veri seti ise RaFD' dir. Bu veri seti Kafkas 67 modelin yz verilerini ieren bir veri setidir. Bu veri setinde kadın, erkek, kız ocuėu ve erkek ocuėu modeller kullanılmıřtır.8 farklı duyguya sahip olan veri setindeki duygular; fke, ařaėılama, tikslenme, korku, mutluluk, ntr, znt ve řařkınlıktır. Ayrıca grntleri ok ynl incelemiřtir. Bunlar n taraf, sol taraf ve saė taraf olacak řekilde 45 derecelik aılar ile 0'dan 180 derece kadar bakıř aıları 5 farklı kamera ile tespit edilmeye alıřılmıřtır. RaFD veri seti zerinde yapılan alıřmalarda Sun ve arkadaşları klasik metotlar kullanarak %93,4 bařarıma varmıřlardır [127].

Bir diėer veri seti ise JAFPE veri setidir. JAFPE veri seti 10 farklı japon kadın modelden alınan 213 yz duygusu iermektedir [128]. Lui ve ekibi [129] tarafından derin inan aėları (Deep

belief networks) ve parçacık sürü optimizasyonuna dayanan bir yüz tanıma modeli çalışmalarında %91,8 oranında başarı elde etmişlerdir Hamester ve arkadaşları[130] CNN ve yığınlı evrişimli otomatik kodlayıcılar (stacked convolutional autoencoders) çalışmalarında JAFFE veri seti üzerinden %95,8 oranında başarı elde etmişlerdir.

3. YÖNTEM

Yapılan çalışma birçok aşamadan meydana gelmektedir. Bunlardan birincisi çalışmada kullanılacak olan veri setinin belirlenmesi ve verilerin toplanması adımıdır. AffectNet veri kümesi daha önceki bölümlerde belirtildiği üzere Amerika'daki Denver Üniversitesinde; Elektrik ve Bilgisayar Mühendisliği bölümünde Profesör olarak görev yapmakta olan Mohammad H. Mahoor tarafından hazırlanmış olan veri seti özel izinle ve tez çalışmasında belirtmek üzere alınmıştır. Veri kümesi seçilirken; içermiş olduğu verilerin tutarlı olması, veri sayısı miktarı ve derin öğrenme yöntemleri ile yüz görüntülerinden duygu analizine yardımcı olabileceği göz önünde bulunmuştur. Büyük AffectNet veri kümesinin 122GB çok büyük verileri olmasına rağmen bizim kullandığımız veri kümesinin manuel olarak etiketlenmiş olan yaklaşık 250bin(4GB) veri içeren küçük bir bölümüdür. Özel izinle elde edilmiş olan veri kümesi; One drive üzerinde train_set ve val_set olmak üzere iki klasör bulunmakta olup veriler bu şekilde indirilmiştir. Zip formda olan dosyalar dışarı çıkarıldıktan sonra klasörün içinde görüntüleri ve ilgili etiketler elde edilmiştir.

İkinci adım olarak indirilmiş bu veriler bir araya getirilerek standart normalizasyon işlemine tabii tutulmuştur. Görüntü verilerinin normalleştirilmesi RGB formatından gri tonlamaya getirilerek veriler sağlanmıştır. Normalize edilmiş olan veriler: Veri ön işleme ile derlemeye hazır hale getirilen görüntü verileri Hafif Ağırlıklı Gan (Light Weight Gan) yöntemi kullanılarak eksik olan sınıflardaki görüntü miktarının artırılması sağlanmıştır.

Daha önceki bölümlerde bahsettiğimiz gibi sahte veriler üretmek için derin öğrenme algoritmalarından biri olan GAN algoritmasından yararlanılmıştır. Elde edilen verilerin işlenebilmesi ve doğru sonuçların elde edilmesinde evrişimli sinir ağı model yapısının iyi bir şekilde oluşturulmasına bağlıdır. Model yapısı oluşturulurken katman miktarı, sıralaması ve değişkenlerin değerlerinin kombinasyonları oldukça önem taşımaktadır. Eğitim veri seti için aşağıdaki veri ön işleme ve veri artırma tekniklerini uygulanmıştır.

- Görüntüler yatay olarak rastgele çevrildi.
- Görüntüler yatay olarak 30 dereceye kadar döndürülmüştür.

- Duygu analizi çalışmasında CNN mimarilerinden yaygın olarak kullanılan RESNET152 ve InceptionNetv3 model geliştirme de kullanılmışlardır. ResNet152 için 224x224 piksel olarak, InceptionNetv3 için 229x229 piksel olarak yeniden boyutlandırılmıştır.
- Piksel değerlerini [-1,1] aralığında normalize edilmiştir.

Elde edilen sonuçlar daha sonra test için ayrılan veriler ile karşılaştırılarak Karşılık Matrisi üzerinden kontrol edilmiştir. Tüm bu hesaplamalar bittikten sonra elde edilen değerler hangi mimarinin AffectNet veri seti üzerinde yüz görüntülerinden duygu analizi yaparken daha iyi bir sonuç verdiğini gösterilmiş olacaktır.

3.1. Veri Ön İşleme (Data Preprocessing)

Veri ön işleminin yapılmasının temel sebebi; veri kümesi üzerindeki parametrik olarak gelen değerlerin bütünlüğünü, gereksiz tekrarlanmasını, bozulma durumlarının engellenmesini ve gürültülü olan verilerin ayıklanması gibi önemli işlemlerin gerçekleştirilmesidir. Veri ön işleme sayesinde gürültülü ve eksik olan veriler sistemden ayıklana bilinmektedir. Bu sayede üzerinde çalışılmak istenen veri kümesinde bütünlüğün artması ve dolayısıyla performansın artması beklenmektedir. Çalışmamızda kullandığımız AffectNet veri kümesinde ki Annotations klasörü içinde bulunan etiketli görüntüler doğrudan CSV formatında değildir, görüntüler numpy formattadır. Bu sebepten ötürü çalışmaya öncelikli olarak görüntülerin CSV forma getirilmesi ile başlanmıştır. Şekil 3.1 'de görsel-etiket bilgileri gösterilmiştir.

1	Image	class
2	259156.jpg	1
3	298047.jpg	1
4	383421.jpg	1
5	334492.jpg	7
6	55904.jpg	1
7	339091.jpg	1
8	133985.jpg	1
9	372593.jpg	0
10	22321.jpg	3
11	244470.jpg	0
12	304528.jpg	1

Şekil 3.1 – Görsel-Etiket Bilgisi

3.2. Normalizasyon İşlemi

Normalizasyon işleminin yapılmasının sebebi; Veri setindeki parametreler arasında veri bütünlüğünü kontrol etmek, verilerin bozulmasını engellenmek ve gereksiz veri tekrarlarını ortadan kaldırmak böylelikle de yapılacak olan işlemlerdeki performansın iyileştirmesini sağlanmış olacaktır. Çalışmada yapılan normalize işlemi denklem 3.1 de gösterilmiştir.

$$X_{yeni} = \frac{(X - X_{min})}{(X_{max} - X_{min})} \quad (3.1)$$

Formülde hesaplama mantığı şu şekildedir; X değeri mevcut olan parametrenin değerini temsil etmektedir X_{max} değeri parametrenin en büyük değerini, X_{min} değeri ise parametredeki en küçük değerini ifade etmektedir. X'in güncel değerinden minimum değer ile farkının hesaplanıp, maksimum ve minimum değer farkına oranı ile yeni değer elde edilmesi sağlanır. X_{yeni} değeri denklemin sonucunda ortaya çıkacak olan normalize edilmiş yeni değerdir.

3.3. Light Weight GAN ile Sahte Görüntü Üretme

Son yıllarda derin öğrenme algoritmalarına olan ilginin artması ile güçlü ve üretken çalışmalar ortaya konulmuştur. Bunun yanında modellerin çıktılarının gerçek görüntüler ile ayırt edilemeyecek derecede başarılı olması da büyük beğeni toplamıştır. “Generative Adversarial Networks (GANs)” [131] çalışma yapısından diğer bölümlerde incelenmiştir. AffectNet veri setindeki sınıfların eleman sayıları eşit miktarda değildir, bu durum bize veri kümesinin dengesiz olduğunu göstermektedir. Gan algoritması kullanılarak eleman sayısı eksik olan sınıflar için sahte görüntüler üretilmeye ve veri kümesini dengeli bir hale getirmeye çalışıldı. AffectNet veri kümesindeki öğrenme ve korkma verileri üzerinde hafif ağırlıklı çekişmeli üretken ağlar'dan yararlanıldı.

GAN Eğitim Algoritması Uygulaması Üretici Eğitimi

GAN eğitim algoritmalarının çalışma şekli aşağıdaki gibi özetlenmiştir. Üretici (G), girdi olarak rastgele gürültü (z) değerini alır ve G(z) olarak yeni görüntü üretir. G 'nin eğitimini denklem 1 'e bakarak elde edilir. Üretici ağı yani G, ayırıcı yani D'nin hatasını maksimum seviyeye çıkarabilmek için aynı benzerlikte görüntüler üretir. Denklem 3.2' deki m değerleri ağı girdi olarak gelen görüntü sayısıdır.

Ayrıcı Eğitimi: Giriş değeri olarak ayrııcıya gelen X gerçek görüntü ve $G(z)$ yeni üretilmiş görüntüyü alır. Denklem 3.3’ deki m değerleri ağa girdi olarak gelen görüntü sayısıdır.

$$\nabla \theta g = \frac{1}{m} \sum_{i=1}^m \log(1 - D(G(z^i))) \quad (3.2)$$

$$\nabla \theta g = \frac{1}{m} \sum_{i=1}^m [\log D(x^i) + \log(1 - D(G(z^i)))] \quad (3.3)$$



Şekil 3.2 – Light Weight Gan İle Elde Edilen Yüz Görselleri

3.4. ResNET (Artık Ağlar)

ResNET mimarisi; 2015 yılında ILSVRC (ImageNet Large Scale Visual Recognition Challenge) yarışmasını %3.57 lik hata oranı ile kazanmış olan yapay sinir ağı mimarisidir [132] [133]. ResNet mimarisinin ortaya çıkma nedenleri arasında CNN’lerin performans düşüklüğü sorununu çözmek yatmaktadır. ResNET geliştirilirken kaybolan gradyan probleminin çözülebilmesi için oldukça derin ağ yapısı geliştirmiştir [134]. ResNet 18,34,50,101,152,1202

olmak üzere pek çok çeşiti bulunmaktadır. ResNet 18 ve 34 2 katman, ResNet 50,101,152 ise 3 katman derinliğe sahiptir [135].

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Şekil 3.3 – ResNet Katmanları

Evrişimli sinir ağları, yapay sinir ağlarında olduğu gibi modelde derinlik arttıkça, elde edilen sonuçların arttığı görülmüştür. Fakat derinliğin fazlalaşması ile ağların doymun olması sebebi ile ilerleyen katmanlarda giderek bozulmalar meydana gelir ve eğitim hatalarının oluşmasına neden olur.

Giderek bozulma eğiliminde olan sorun ResNet 'kullanımı ile çözülmektedir [136]. Katmanlar arası geçiş işlemi denklem (3.4,3.5 ve 3.6) ile hesaplanabilmektedir.

$$y = F(x) \quad (3.4)$$

$$y' = y + x \quad (3.5)$$

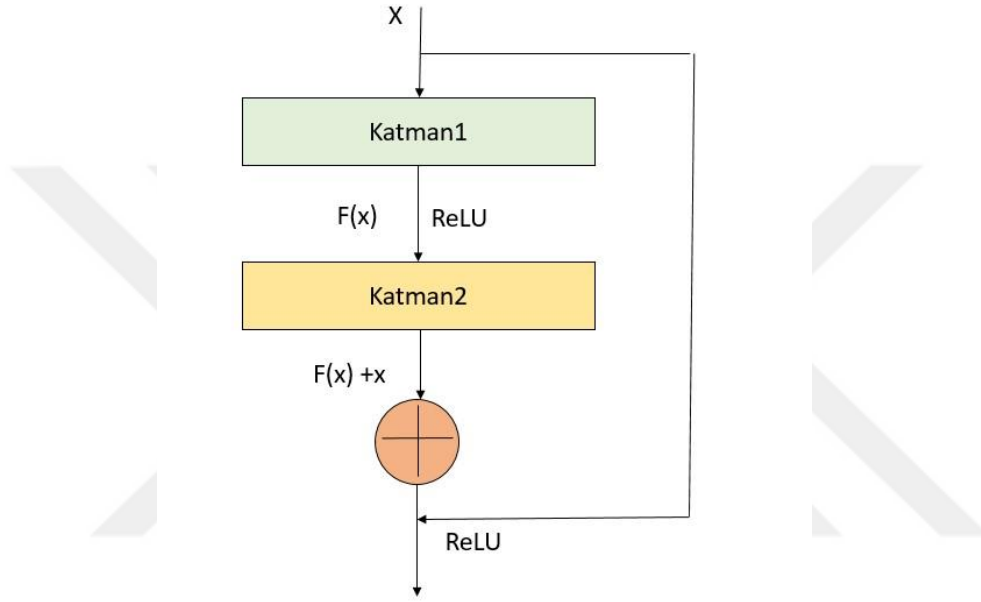
$$y' = F(x) + x \quad (3.6)$$

Denklemdaki x değeri girdi değeri olarak kabul edilip y değeri ise elde edilen çıktıyı temsil etmektedir. Girdi değerine eklenen çıktı değerleri sonucu yeni bir çıktı değerinin oluşmasını

sağlamaktadır. ReLU aktivasyon fonksiyonu $F(x) + x$ olarak kabul edilmektedir. Aktivasyon fonksiyonun uygulanması ile bir önceki katmana ait çıktı daha doğru bir şekilde bir sonraki katmana aktarılmış olur. Artık ağırlar yapısal işlevsellik olarak Yinelene Sinir Ağlarına oldukça benzemektedir. ResNet mimarisi $224 \times 224 \times 3$ boyutlarındaki görüntüler ile ilgilenmektedir.

ResNet mimarisinin yapısı

Şekil 3.3’de gösterilmiştir.

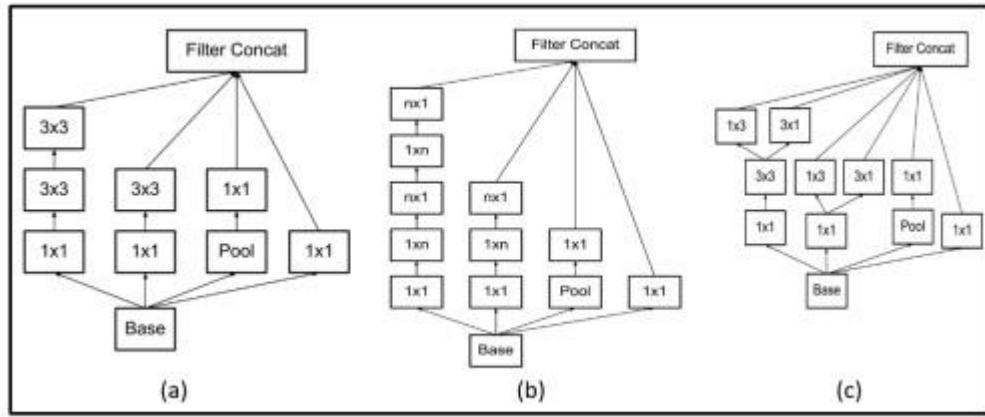


Şekil 3.4 – ResNet Mimarisi

3.5. Inception-v3

GoogLeNet mimarisinin daha derine giderek çalışma isteği Inception isimindeki modüllerin ortaya çıkmasına sebep olmuştur. Geliştirilen her bir modül bir öncekindeki eksiklikleri gidermeyi hedeflemiştir. Bu modüller; Inception v1, Inception v2, Inception v3 ve Inception v4’dür. Inception v3 ILSVR 2012 de üzerinde yapılmış olan testlerde oldukça başarılı sonuçlar elde etmiştir [137]. Mimari $229 \times 220 \times 3$ boyutundaki görsellerde çalışmaktadır. Geliştirilmiş olan mimaride filtrelerde bir takım değişiklikler yapılmıştır. Klasik olarak 7×7 özelliğindeki evrişim filtresi, 3×3 ‘lük üç filtreye çevrilmiştir. Başlangıç modülün de ki 5×5 boyutundaki ilk üç filtrede 3×3 boyutlarında olacak şekilde iki adet filtreye çevrilmiştir (Şekil

7.a) . Bu işlemlerin çıktısı olarak 17x17 boyutunda 768 tane özellik elde edilmiştir. İşlemlerin devamında evrişimsel filtreler (7x7) yerine asimetrik filtreler (1x7,7x1) kullanılmıştır (Şekil 7.b).Bu işlemlerin sonucunda 1280 tane 8x8 boyutunda özellik haritası elde edilmektedir. Mimarinin son kısmında 3x3 boyutundaki evrişimsel filtrelerde asimetrik olarak değiştirilerek 1x3 ve 3x1 haline, 1x1 boyutunda olan filtre ise 3x1 ve 1x3 olarak değiştirilmiştir (Şekil 3.4).Tüm bu işlemlerden 2048 adet 8x8 boyutunda öznetelik çıkarımı yapılmıştır. Inception v3, 2048 adet nörondan oluşan tam bağlantılı (Fully Connected) blok Softmax ‘den önce gelmektedir.



Şekil 3.5 – Inception V3 Mimarisi

3.6. Karışıklık Matrisi (Confusion Matrix)

Karışıklık matrisi, önceden belirlenmiş olan veri setlerinden oluşturulan modellerin performansını test etmek için kullanılan yaygın yöntemlerdendir [138]. Matrisin satır ve sütun sayısı mevcut sınıf sayısını gösteren bir kare matristir. Örneğin; iki sınıfı bulunması onun 2x2 özelliğinde olduğunu göstermektedir. Şekil 3.5’de gösterilmiştir. Matriste bulunan negatif ve pozitif veriler değerlendirme sonuçlarını ifade etmektedir. Matrisi koordinat sistemine göre ifade edilirse x eksen (yatay) tahmin edilen değerlere, y eksen (dikey) Çalışmada karışıklık matrisi kullanılarak elde edilen sonuçların analizi yapılabilmektedir. Matriste gösterilen Yanlış Pozitif (FP) alanların: gerçek değerleri negatif olmasına rağmen pozitif sınıflandırma yapılmış olduğunu ifade etmekte olup yanlış sınıflandırmayı temsil etmektedir. Gerçek Negatif (TN) : belirtilen alanlarında gerçek değerleri negatif olup negatif olarak doğru bir şekilde sınıflandırma

yapılmıştır. Gerçek Pozitif (TP) : doğru sınıflandırma yapıldığını ifade etmektedir yani test verilerindeki değer ile model üzerindeki sınıflarının aynı olduğunu ifade etmektedir. Yanlış Negatif (FN): yanlış sınıflandırılma yapıldığını ifade etmektedir yani test verilerindeki değer ile modelin çıktısı farklı sınıftır.

		Tahmin Edilen Değerler	
		Yok	Var
Gerçek Değerler	Yok	TN	FP
	Var	FN	TP

Şekil 3.6 – Karışıklık Matrisi Modeli

Karışıklık matrisi kullanılarak veri setlerinin; kesinlik (precision), doğruluk (accuracy), ve duyarlılık (recall) gibi parametrelerinin değerleri hesaplanmıştır. Kesinlik değeri denklemi 3.7’de, doğruluk değeri denklemi 3.8’de, duyarlılık değeri denklemi 3.9 ‘deki formüller ile hesaplanması yapılmıştır.

$$Kesinlik (Precision) = TP / (TP + FP) \quad (3.7)$$

$$Doğruluk (Accuracy) = (TP + TN) / (TP + TN + FP + FN) \quad (3.8)$$

$$Duyarluluk (Recal) = TP/(TP + FN) \quad (3.9)$$

4. ÖNERİLEN MODEL

Yapılan çalışmada, daha önceden geliştirilen çalışmalara göre daha iyi sonuçların sağlanması ve tutarlılık amaçlanmıştır. Bu amaca hizmet edebilmek için hangi veri setleri kullanıldığı daha önce 2.Bölümde anlatılmıştır. Hangi GAN modelinin kullanıldığı ve hangi metodolojilerin kullanıldığından bu bölümde bahsedilmiştir. Yapılan tez çalışması ile pozitif yönde dönüşümler elde edilecektir. Bu bölümde sistemin genel yapısı, kullanılan özellikler ve çalışmanın genel işleyişi anlatılacaktır. Çalışmasında geliştirilen modelde kullanılan ide ve kütüphanelere ait kısa bilgiler verilmiştir.

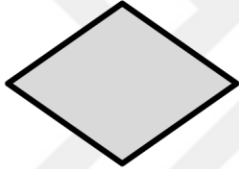
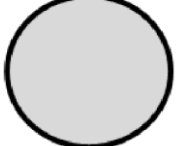
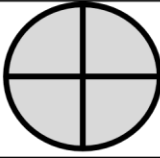
4.1. Kullanılan Veri Seti

Veri kümesi bölümünde anlatıldığı gibi; AFFECTNET veri kümesi kullanılmıştır. AFFECTNET veri kümesi bu alandaki araştırmacılar tarafından kullanılabilecek veri miktarına sahip olmasından dolayı derin öğrenme çalışmaları için idealdir.

4.2. Modelin Akış Diyagramı

Bilgisayar dünyasında algoritmalar; problemlerin çözümünün açık ve ardışık biçimde belirtilme şeklidir. Problemlerin çözümleri için geliştirilmiş olan algoritmaları metinsel, kaba kodlar (sözde) ve akış diyagramları ile ifade edilebilmektedir. Akış diyagramları algoritmaların belirli simgesel görsellerle ifade edilebilmesi işlemidir. İster karmaşık olsun ister yalın pek çok süreçte akış diyagramları sistemlerin anlaşılabilmesini artırma özelliğine sahiptir. Bu sebepten ötürü yaygın bir şekilde karşımıza çıkmaktadır.

Akış diyagramında herkesin açık bir şekilde anlayabileceği şekiller bulunmakta olup bu şekiller genel geçer ifadelerdir. Şekillerin herkes tarafından anlaşılabilir olması başka yazılım geliştiriciler tarafından da inceleme yapıldığında rahatlıkla sistemin anlaşılabilmesini sağlayacaktır. Tablo 3.1 de akış diyagramlarında kullanılan semboller gösterilmektedir.

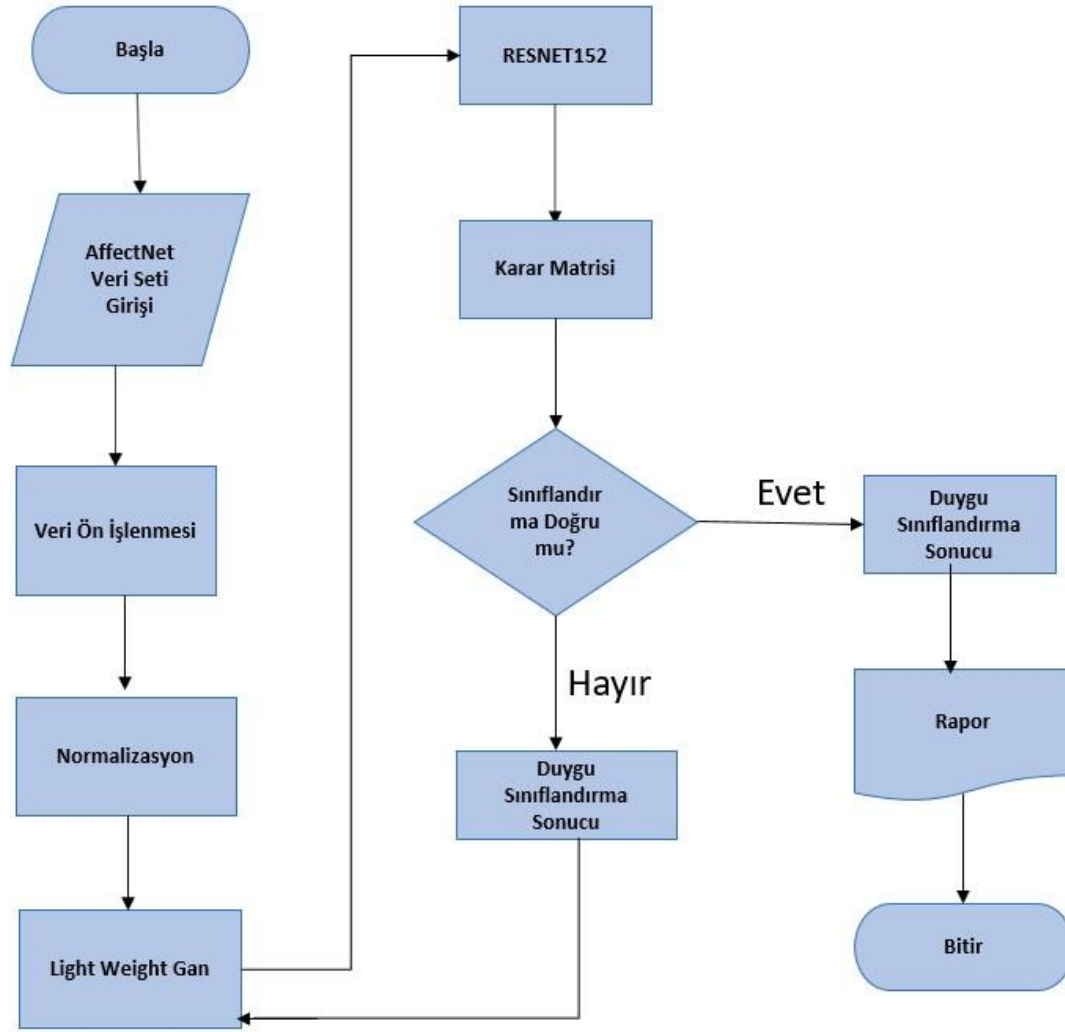
	Bağlantı Sembolü: Akış yönü ve semboller arasındaki ilişkiyi göstermektedir
	Başlangıç / Bitiş Sembolü: Sonlandırıcı, işlemin nerede başladığını veya nerede bittiğini göstermektedir
	Eylem veya Süreç Sembolü: Bir işlemi, süreci veya aktiviteyi ifade etmektedir.
	Belge Sembolü: Basılı bir belge veya raporu.
	Birden Fazla Belge Sembolü: İşlemden birden çok belgeyi temsil etmektedir.
	Karar Sembolü: Bir karar veya dallanma noktasını temsil etmektedir.
	Giriş / Çıkış Sembolü: Giriş – çıkış yani sisteme giren veya sistemden gönderilen materyal, bilgileri temsil etmektedir.
	Manuel Giriş Sembolü: Bilgilerin sisteme manuel olarak girmesi istendiği bir adımı temsil etmektedir.
	Hazırlık Sembolü: Süreçte başka bir adımın ön hazırlık veya kurulumunu temsil eder. Devam etmeden önce bir işlemin yapılması, ayarlanması veya değiştirilmesi gerektiğini gösterir.
	Bağlayıcı Sembolü: Akışın, eşleşen bir sembolün yerleştirildiği yerde devam ettiğini göstermektedir.
	Veya Sembolü: İşlem akışının ikiden fazla dalda devam ettiğini göstermektedir.

	Alt Yordam Sembolü: Daha büyük bir süreç içinde gömülü belirli bir görevi gerçekleştiren eylemler dizisini belirtmektedir.
	Manuel Döngü Sembolü: Manuel olarak durdurulana kadar tekrarlamaya devam edecek bir dizi komut dizisini belirtmektedir.
	Döngü Sınırı Sembolü: Bir döngünün durması gereken noktayı belirtmektedir.
	Veri Depolama – Saklama Sembolü: Verilerin saklandığı bir adımı belirtmektedir.
	Veritabanı Sembolü: Arama ve sıralama için izin veren standart bir yapıya sahip bilgi listesini göstermektedir.
	Gecikme Sembolü: İşlemden gecikme olduğunu göstermektedir.
	Dahili Depolama Sembolü: Yazılım tasarımı akış şemalarında kullanılan bir program sırasında bilgilerin bellekte saklandığını belirtmektedir.

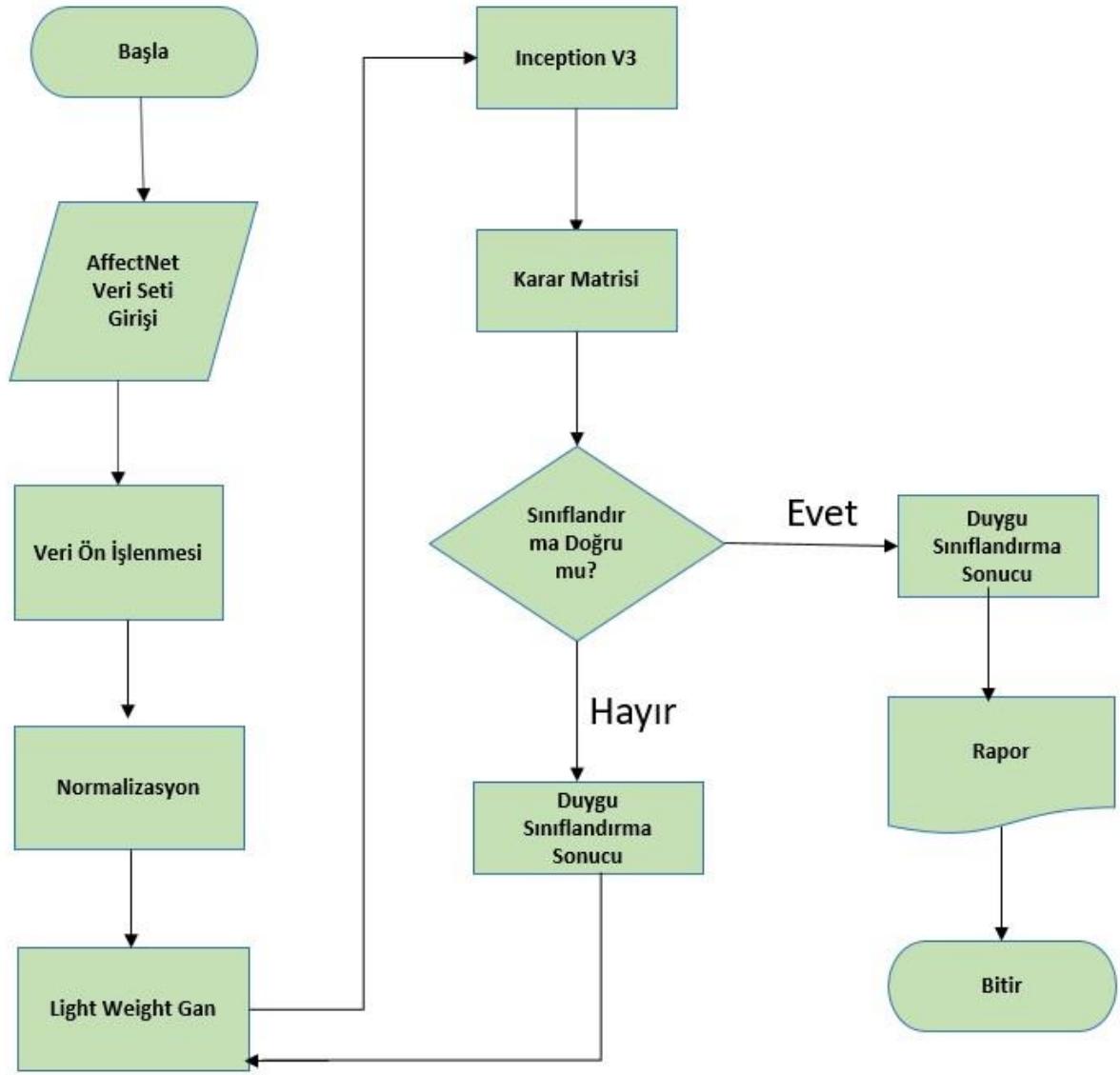
Tablo 3.1 Akış Diyagramı Elemanları

Yapılan çalışmada kullanılan akış diyagramı Şekil 4.1 ve Şekil 4.2’de gösterilmektedir. Şemanın adımları belirtilmiştir.

1. Adım: Başla işlemi ile akış başlatılır.
2. Adım: Sisteme AffectNet veri seti girişi sağlanır.
3. Adım: Sisteme girişi yapılan veriler işlenir.
4. Adım: Verilerin normalizasyon işlemi yapılır.
5. Adım: Okunmuş olan veriler Hafif ağırlıklı GAN derin öğrenme algoritması ile eğitilir.
6. Adım: Üretilmiş olan sahte görüntüler CNN mimarileri RESNET152 /Inception v3 ile eğitilir.
7. Adım: Eğitim sonucu karışıklık matrisleri ile gösterilir ve izlenir.
8. Adım: Yapılan işlemlerde “Sınıflandırma Doğru Yapıldı mı?” sorusu sorulur.
9. Adım: 8. Adımda sorulan sorunun cevabı hayır ise yanlış sınıflandırma sonucu GAN içerisine tekrar gönderilir. Eğer sorunun cevabı evet ise 10.adıma atlanır.
- 10.Adım: Doğru sınıflandırma analizi yapılır.
- 11.Adım: Doğru ve Yanlış olarak işlenen sonuçlar karşılaştırılır.
12. Adım: Bitir işlemi ile akış sona erdirilir.



Şekil 4.1 – ResNet152 Modeli Akış Diyagramı



Şekil 4.2 – Inception v3 Modeli Akış Diyagramı

4.3. Modelde Kullanılan Teknoloji

Görüntü verilerini işlenmesi aşamasında GPU ve CPU desteği sunan Google Colab Pro üzerinden çalışma yürütülmüştür. Veri kümesi, Google drive 'da yüklü olduğu için iki ortam birbirleri ile senkronize haldedir. Colab'ın notebook özelliği sayesinde; metni, resimleri ve kodu içiçe barındırması bu platformda çalışma sebeplerimizdendir. Tez çalışmasında programlama dili olarak Python kullanılmıştır. Python 'un tercih edilmesinin temel sebebi veri analizi konularında diğer dillere göre daha kullanışlı ve hızlı olmasıdır. Python görüntü işleme

ve yapay zekâ konularında oldukça fazla sayıda kütüphaneye (numpy, dlib vb.) sahip olmasından dolayı bahsedilen konularda yazılım sürecinin pozitif yönde ivmelenmesini sağlamaktadır.

Kullanılan kütüphaneler:

- İşletim sistemi üzerinden yapılabilecek işlemler için os modülü kullanıldı.
- Dizi işlemleri için numpy,
- Veriyi görselleştirmek için matplotlib-pyplot,
- PIL (Python Image Library) ile resim dosyaların birleştirilir, RGB formatına ayırma işlemi
- Cv2 kütüphanesi sayesinde görüntü işleme çalışmaları yapılmaktadır.
- Sys python'ın standart kütüphanelerindendir.

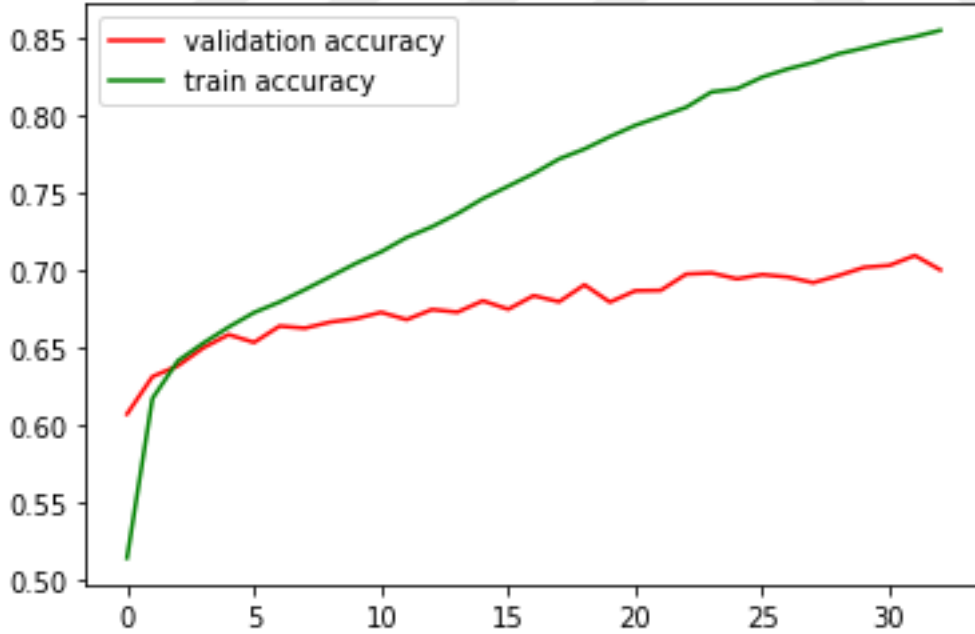
5. TEST SONUÇLARI VE DEĞERLENDİRMELER

Bu bölümde AffectNet veri kümesindeki etiket sayısı eksik olan sınıflardaki görüntülerin GAN ile artırılması sonucunda elde edilen yüz görüntü verilerinin ResNet152 ve Inception v3 mimarileri ile modellenmesi sonucunda duygu analizi yapılması sağlanmıştır. Çalışmada yapılmış olan modellemelere göre Inception v3'ün daha başarılı sonuçlar verdiği gözlemlenmiştir. Önerilen modellere ait tüm sonuçlar detaylı bir şekilde açıklanmakta ve grafiklerde gösterilmektedir.

Sonuç – 1

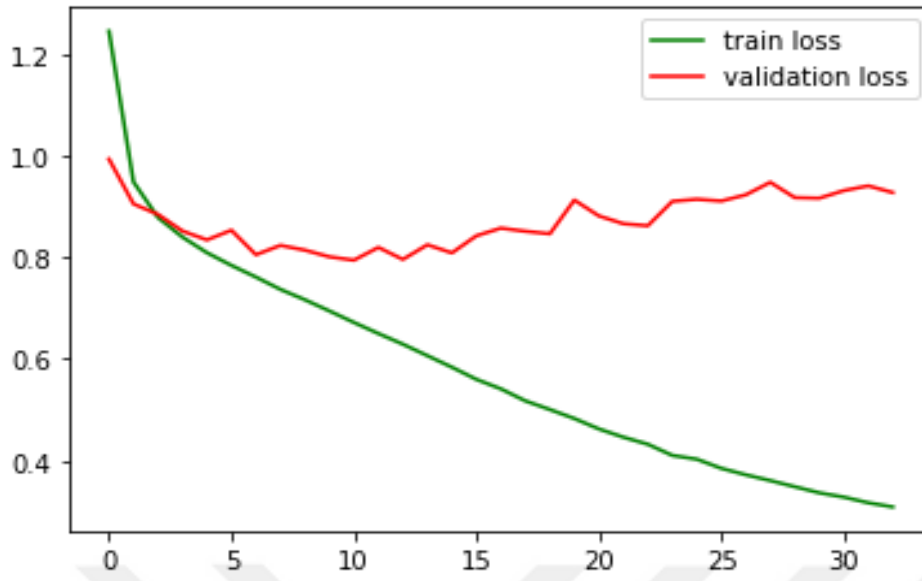
ResNet152 mimarisi kullanılarak doğruluk ve test sonuçlarına ait grafikler gösterilmiştir. Oluşan grafiğe ait sonuçlar Tablo 4.1 ve Tablo 4.2 de gösterilmektedir.

- Eğitim sonuçları- Doğruluk Eğrileri



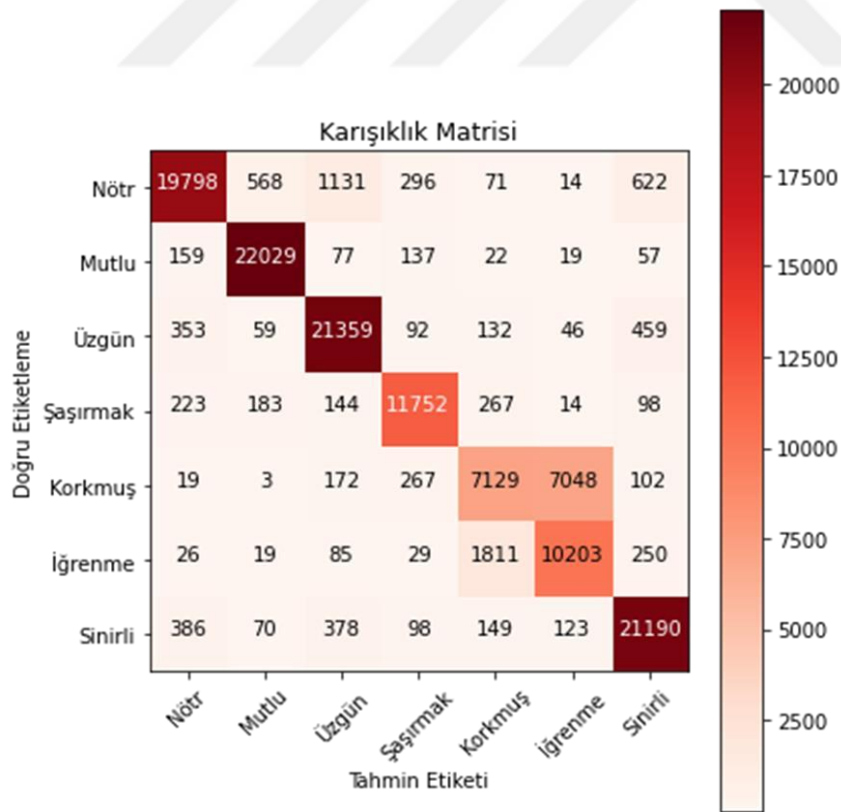
Şekil 5.1 – ResNet152 Eğitim Doğruluk Sonuçları

- Kayıp Değerler



Şekil 5.2 – ResNet152 Kayıp Değerler

- Karışıklık Matrisi Eğitim Sonuçları:**



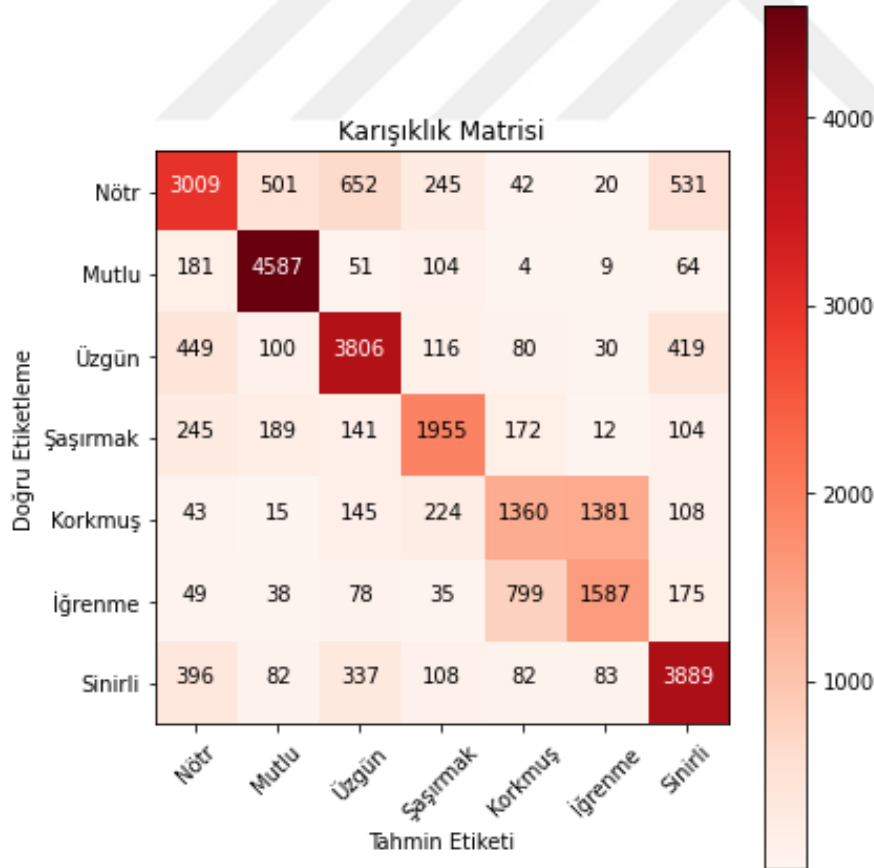
Şekil 5.3 – ResNet152 Karışıklık Matrisi Eğitim Sonuçları

Sınıf	Nötr	Mutlu	Üzgün	Şaşırmak	Korkmuş	İğrenme	Sinirli
Kesinlik	0.94	0.96	0.91	0.92	0.74	0.58	0.92
Recall	0.90	0.97	0.93	0.92	0.58	0.68	0.93
F1 Skor	0.87	0.98	0.95	0.92	0.48	0.82	0.94

Tablo 4.1 ResNet152 Eğitim Verisi Doğruluk Sonuçları

Geliştirilmiş olan modelin sonuçları incelendiğinde ResNet152 ile %87 başarımla elde edildiği görülmüştür.

- Karışıklık Matrisi Test Doğrulama Sonuçları**



Şekil 5.4 – ResNet152 Karışıklık Matrisi Test Doğrulama Sonuçları

Sınıf	Nötr	Mutlu	Üzgün	Şaşırmak	Korkmuş	İğrenme	Sinirli
Kesinlik	0.68	0.83	0.73	0.70	0.53	0.50	0.73
Recall	0.64	0.87	0.74	0.69	0.46	0.53	0.75
F1 Skor	0.60	0.91	0.76	0.69	0.41	0.57	0.78

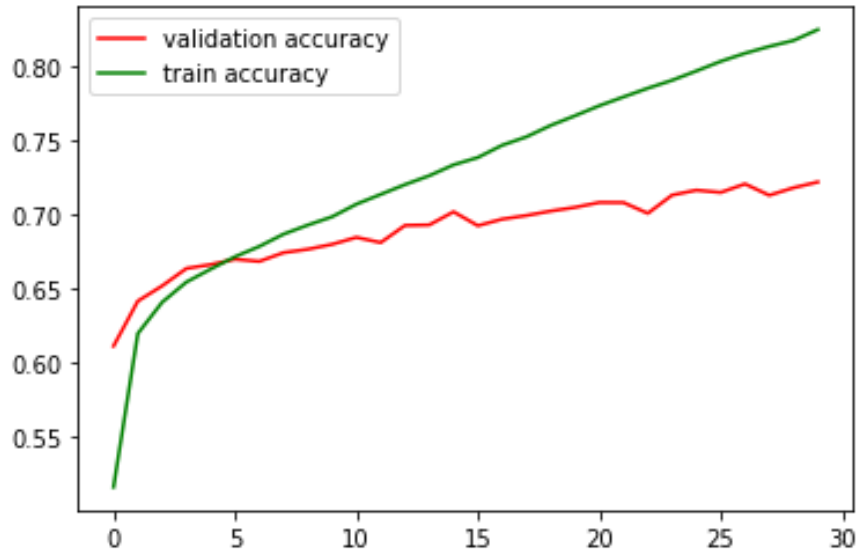
Tablo 4.2 ResNet152 Test Verisi Doğruluk Sonuçları

Geliştirilmiş olan modelin sonuçları incelendiğinde ResNet152 ile %70 test doğruluğu elde edildiği görülmüştür.

Sonuç – 2

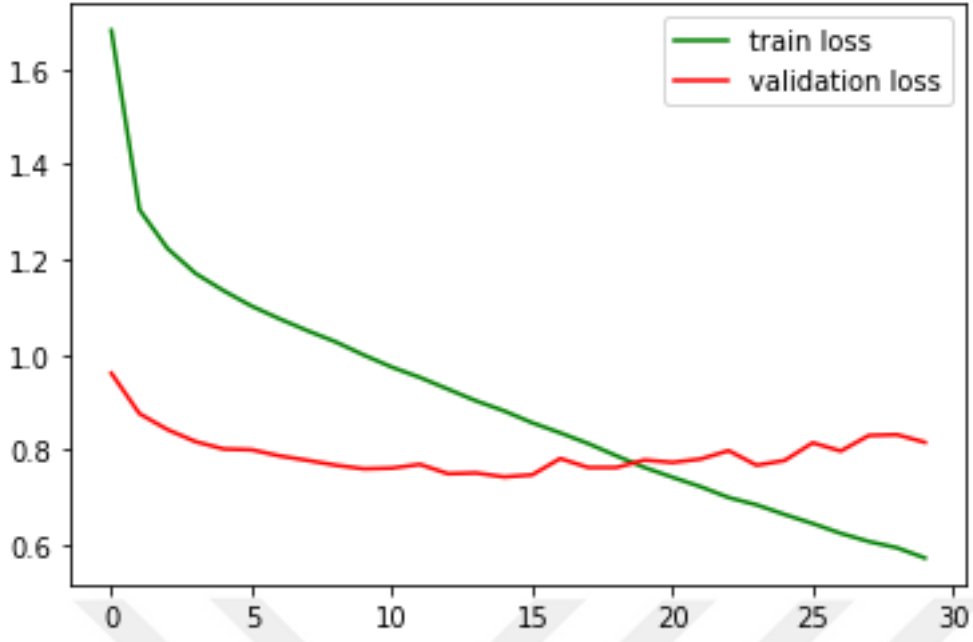
Inception v3 mimarisi kullanılarak doğruluk sonuçlarının grafiği Şekil 5.4 Doğruluk Sonuçları, Şekil 5.5 Kayıp Değerleri ve Şekil 5.6 Karışıklık matrisine ait sonuçları göstermektedir. Oluşan grafiğe ait sonuçlar Tablo 4.3 ve Tablo 4.4’de gösterilmektedir.

- Eğitim sonuçları- Doğruluk Eğrileri



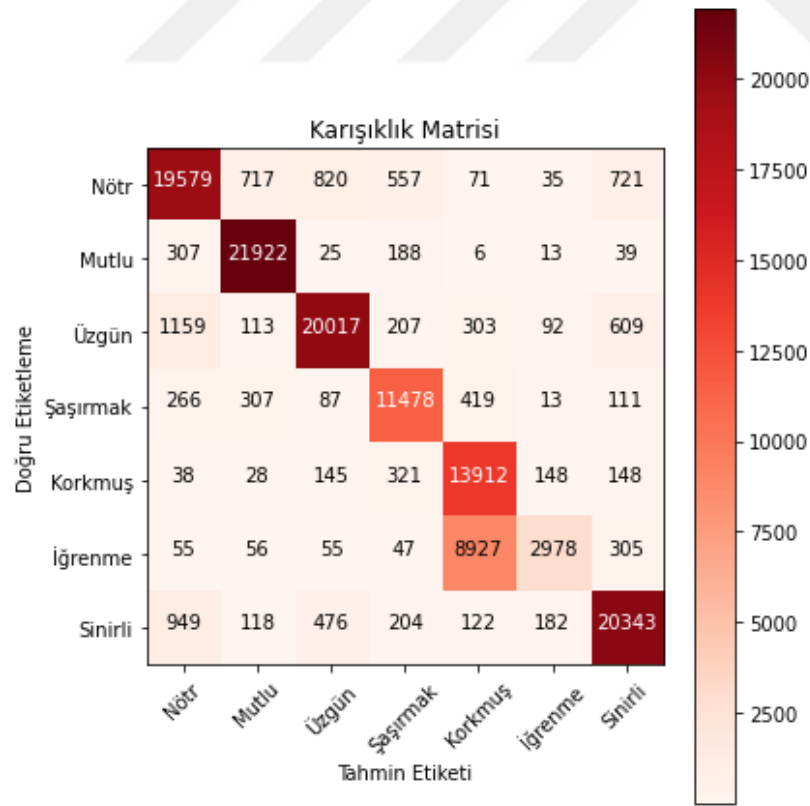
Şekil 5.5 – Doğruluk Sonuçları

- Kayıp Değerler



Şekil 5.6 – Kayıp Değerler

- **Karışıklık Matrisi Eğitim Sonuçları**



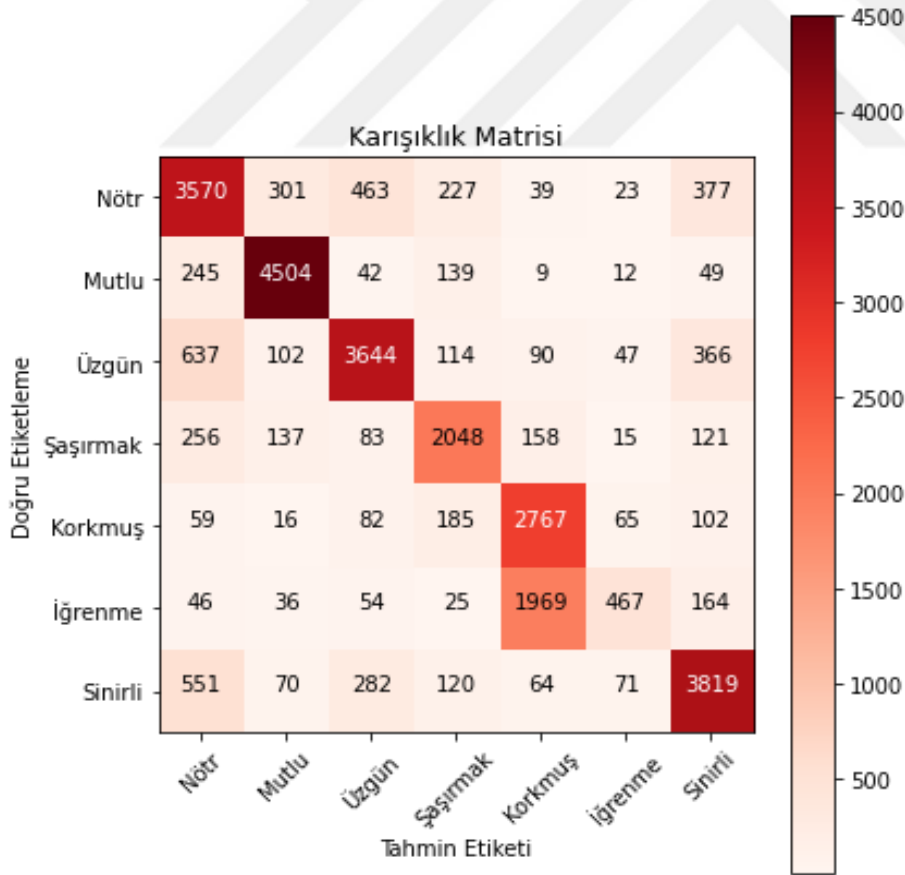
Şekil 5.7– Karışıklık Matrisi Eğitim Sonuçları

Sınıf	Nötr	Mutlu	Üzgün	Şaşırmak	Korkmuş	İğrenme	Sinirli
Kesinlik	0.87	0.94	0.92	0.88	0.58	0.86	0.91
Recall	0.87	0.97	0.89	0.90	0.94	0.24	0.90
F1 Skor	0.87	0.95	0.90	0.89	0.72	0.37	0.90

Tablo 4.3 Inception v3 Eğitim Verisi Doğruluk Sonuçları

Geliştirilmiş olan modelin sonuçları incelendiğinde Inception v3 ile %85 doğruluk elde edildiği görülmüştür.

- Karışıklık Matrisi Test Doğrulama Sonuçları**



Şekil 5.8 – Karışıklık Matrisi Test Doğrulama Sonuçları

Sınıf	Nötr	Mutlu	Üzgün	Şaşırmak	Korkmuş	İğrenme	Sinirli
Kesinlik	0.66	0.87	0.78	0.71	0.54	0.66	0.76
Recall	0.68	0.88	0.75	0.72	0.66	0.26	0.76
F1 Skor	0.71	0.90	0.72	0.72	0.84	0.16	0.76

Tablo 4.4 Inception v3 Test Verisi Doğruluk Sonuçları

Geliştirilmiş olan modelin sonuçları incelendiğinde Inception v3 ile %72 test doğruluğu elde edildiği görülmüştür.

6. SONUÇLAR

Duygu analizi kavramı, Bilgisayar Mühendisliği alanındaki en popüler süreçlerden biridir. Doğru veya minimum hata ile duygu analizi yapılabilmesi için veri sayısı hayati öneme sahiptir. Eksik olan verilerin GAN algoritması kullanılarak artırılması ile yüz çalışmasında bulunacak kişilerin daha fazla sayıda örnekleme sahip olmasına yardımcı olacaktır. Görüntü işleme işlemleri oldukça uzun süren aşamalardan meydana geldiğinden burada maliyetlerde kritik önem taşımaktadır. Örnek sayısı arttıkça işlemek uzun bir zaman ve işlemci gücüne sahip olma sorununu ortaya çıkarır. Doğru modellerin geliştirilmesi ile verimliliğinde artması sağlanacaktır.

Literatürde bu kritik işlemler için farklı yaklaşımlar kullanılmaktadır. Yüz görüntülerinden duygu analizi için pek çok CNN mimarileri kullanıldığı görülür. CNN mimarilerinden Inception v3 ve ResNet152 tercih edilerek model geliştirme yapılmıştır. AffectNet veri kümesi kullanılmıştır. Karşılaştırmalı çalışmalar kullanılan mimarilerden Inception v3 AffectNet veri kümesi için ResNet152 ye göre daha yüksek doğruluk oranına sahip olduğu tespit edilmiştir.

7. SINIRLAMALAR VE GELECEK ÇALIŞMA

Mevcut çalışmanın bir takım sınırlamaları mevcuttur. Bu sınırlamaların birincisi veri setlerinin inbalance durumda olmasından kaynaklı olmasıdır. Sınıfların içerisinde bulunan veri sayısının aşırı derecede birbirinden farklı olması veri işleme aşamalarında geçen eforu negatif

yönde etkilemektedir. Örneğin; bir sınıfta iki yüz bin görüntü varken, bir sınıfta 3700 görüntü mevcuttur. Burada yukarı doğru yuvarlama işlemi yapılmak istenildiğinde rakamlar 7 farklı duygu için bir milyonun üzerine çıkacaktır, bu da kontrolü imkânsıza yakın bir veri kümesi oluşturulması demektir. İkinci bir sınırlama ise; mevcutta bulunan izinsiz kullanılabilecek olan veri kümeleri yetersiz olup, yeterli olanlar ile çalışma yapılmak istenildiğinde özel izin alınması gerekmektedir. Bu izinlerde çalışma süresini negatif yönde etkilemektedir. Üçüncü bir sınırlama ise GAN gibi algoritmalarda epoch sayısı başarımı artırdığı için ne kadar çok kez tekrarlanırsa sonuçlar daha iyi olacaktır. Fakat bunun içinde işlemcinin oldukça iyi olması gerekmekte olup normal bir kişisel bilgisayarın lokalinde bu kadar çok veriyi işlemek oldukça zordur. Son bir sınırlama ise ücretli ortamlardan(Google Collab Pro) yararlanılarak çalışmanın yapılması maliyet açısından negatif etkilemektedir. Daha iyi sonuçlar elde edilebilmesi için gelecekte; zaman sınırlaması olmadan yüz görüntülerindeki epoch sayısını 2000 seviyelerinde tekrar tekrar eğitilirse %75'in üzerinde görüntüler elde edilebilecektir.

KAYNAKÇA

- [1] Jaiswal, S. ve Nandi, GC (2020). CNN mimarisini kullanan sağlam gerçek zamanlı duygu algılama sistemi. Sinirsel Hesaplama ve Uygulamaları, 32 (15), 11253-11262.
- [2] Patil, S., Joshi, P., Deore, G., Taley, A., & Sawant, P. (2020, May). Emotion Detection Using Facial Recognition. In 2nd International Conference on Communication & Information Processing (ICCIP).
- [3] Ng, HW, Nguyen, VD, Vonikakis, V., & Winkler, S. (2015, Kasım). Aktarım öğrenimini kullanarak küçük veri kümelerinde duygu tanıma için derin öğrenme. 2015 ACM Bildirilerinde, multimodal etkileşim üzerine uluslararası konferansta (s. 443-449).
- [4] Verma, G., & Verma, H. (2020). Hybrid-deep learning model for emotion recognition using facial expressions. The Review of Socionetwork Strategies, 14(2), 171-180.
- [5] Akçay, C., & Çoruk, A. (2012). Çalışma yaşamında duygular ve yönetimi: Kavramsal bir inceleme. Eğitimde Politika Analizi, 1(1), 3-25.
- [6] Ramakrishnan, S. ve El Emary, IM (2013). İnsan bilgisayar etkileşiminde konuşma duygu tanıma yaklaşımları. Telekomünikasyon Sistemleri, 52 (3), 1467-1478 [6]
- [7] Gonzalez, HA, Muzaffar, S. Yoo, J. ve Elfadel, IM (2020). BioCNN: EEG tabanlı duygu tespiti için bir donanım çıkarım motoru. IEEE Erişimi, 8, 140896-140914.
- [8] Sarkar, P. ve Etemad, A. (2020). Duygu tanıma için kendi kendini denetleyen ekg temsili öğrenimi. Duyuşsal Hesaplama IEEE İşlemleri.
- [9] Yu, D. ve Sun, S. (2020). EDA tabanlı duygu tanıma için derin sinir ağlarının sistematik bir araştırması. Bilgi, 11 (4), 212.
- [10] Er, M. B. (2019). Makine öğrenmesi ile Türk müziğinde duygu analizi (Master's thesis, Maltepe Üniversitesi, Fen Bilimleri Enstitüsü).
- [11] Yar, HIKMAT, Jan, TAUSEEF, Hussain, ALTAF, & Din, SU Gerçek Zamanlı Yüz Duygu Tanıma ve CNN Kullanarak İnsan Robot Etkileşimi için Cinsiyet Sınıflandırması.
- [12] Gu, D., Zhu, M., Zhang, L., & Serikawa, S. Face Recognition with Age, Gender and Emotion Estimations.

[13], M. E., Erkmen, B., & Batur, O. Z. (2021, October). Neuromorphic Logic Gates Design. In] 2021 Innovations in Intelligent Systems and Applications Conference (ASYU) (pp. 1-4). IEEE.

[14] Mortensen, CD (Ed.). (2017). İletişim teorisi . Routledge.
Budden, A. (2009).

[15] The role of shame in posttraumatic stress disorder: A proposal for a socio-emotional model for DSM-V. Social Science & Medicine, 69(7), 1032-1039.

[16] Akpınar, R. B. (2011). Duygusal yüz ifadelerini anlama becerisini geliştirmeye yönelik örnek bir öğretim programı. Journal of Human Sciences, 8(1).

[17] Ekman, P. (1999). Temel duygular. Biliş ve duygu el kitabı, 98 (45-60), 16.

[18] Ozturk, A. O. Halici, U. Ulusoy, I. & Akagunduz, E. (2008, April). 3D face reconstruction using stereo images and structured light. In 2008 IEEE 16th Signal Processing, Communication and Applications Conference (pp. 1-4). IEEE.

[19] ŞENALP, F. M., & CEYLAN, M. Düşük Çözünürlüklü Termal Yüz Görüntü Çözünürlüğünün Derin Öğrenme İle Artırılması. Avrupa Bilim ve Teknoloji Dergisi, 131-135.

[20] KARASULU, B. (2016). Görüntülerde insan kulağı tespit ve bölütlemesini temel alan biyometrik yetkilendirme üzerine bir inceleme. Bilişim Teknolojileri Dergisi, 9(2), 97.

[21] Hatunoğlu, E. (2014). Sınıf III hastalarda erken ve geç dönemlerde uygulanan farklı yüz maskesi protokollerinin sert ve yumuşak dokulardaki etkilerinin sefalometrik ve 3 boyutlu yüz taraması kullanılarak değerlendirilmesi.

[22] Yavuz, H. S. Çevikalp, H., & Edizkan, R. (2013, April). Automatic face recognition from frontal images. In 2013 21st Signal Processing and Communications Applications Conference (SIU) (pp. 1-4). IEEE.

[23] Akpınar, R. B. (2011). Duygusal yüz ifadelerini anlama becerisini geliştirmeye yönelik örnek bir öğretim programı. Journal of Human Sciences, 8(1).

[24] KILIÇ-TÜLÜ, B., & Ergül, C. (2016). Öğrenme güçlüğü olan çocukların duyguları tanıma becerileri. Ankara Üniversitesi Eğitim Bilimleri Fakültesi Özel Eğitim Dergisi, 17(03), 207-229.

[25] Altunel, Ö. Demirdöğen, G., Dural, U., & Kuşçu, M. K. (2008). Şizofrenide duygu algılama ve tanıma süreçleri. Klinik Psikiyatri, 11(Suppl 4), 3-11.

[26] Jadhav, R., Bhuke, J., & Patil, N. (2019). Facial Emotion Detection using Convolutional Neural Network. International Research Journal of Engineering and Technology, e-ISSN, 2395-0056.

[27] Kadak, M. T., Demir, T., & Doğangün, B. (2013). Otizmde Yüz ve Duyguların Yüz İfadelerini Tanıma. Psikiyatride Güncel Yaklaşımlar, 5(1), 15-29.

[28] Zeybek, I. (2018). Dijitalde Duygu Dili: Emojilerin Sihirli Dünyası.

- [29] Dede, G. & Sazlı, M. H. (2010). Biyometrik sistemlerin örüntü tanıma perspektifinden incelenmesi ve ses tanıma modülü simülasyonu. EEBM Ulusal Kongresi, 1-5.
- [30] Sönmez, E. B., Özbek, N. Ö., & Özbek, Ö. (2007). Avuç izi ve parmak izine dayalı bir biyometrik tanıma sistemi. Akademik Bilişim Konferansları.
- [31] Bilgin, M. (2017). Gerçek veri setlerinde klasik makine öğrenmesi yöntemlerinin performans analizi. Breast, 2(9), 683.
- [32] Burcu, A. K. I. N. & ŞİMŞEK, U. T. G. (2018). Sosyal medya analitiği ile değer yaratma: duygu analizi ile geleceğe yönelim. Mehmet Akif Ersoy Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 5(3), 797-811.
- [33] YAZAR, İ., YAVUZ, H. S., & ÇAY, M. A. (2009). Temel bileşen analizi yönteminin ve bazı klasik ve robust uyarlamalarının yüz tanıma uygulamaları. Eskişehir Osmangazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi, 22(1), 49-63.
- [34] Uğur, A., & Kınacı, A. C. (2006). Yapay zekâ teknikleri ve yapay sinir ağları kullanılarak web sayfalarının sınıflandırılması. XI. Türkiye'de İnternet Konferansı (inet-tr'06), Ankara, 1(4).
- [35] Karakoç, M. (2011). Görüntü işleme teknikleri ve yapay zekâ yöntemleri kullanarak görüntü içinde görüntü arama (Master's thesis).
- [36] Wang, K., Peng, X., Yang, J., Meng, D., & Qiao, Y. (2020). Region attention networks for pose and occlusion robust facial expression recognition. IEEE Transactions on Image Processing, 29, 4057-4069.
- [37] Wang, K., Peng, X., Yang, J., Lu, S., & Qiao, Y. (2020). Suppressing uncertainties for largescale facial expression recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 6897-6906).
- [38] Zhang, Z, Luo, P., Loy, C. C., & Tang, X. (2014, September). Facial landmark detection by deep multi-task learning. In European conference on computer vision (pp. 94-108). Springer, Cham.
- [39] Wu, Y. & Ji, Q. (2019). Facial landmark detection: A literature survey. International Journal of Computer Vision, 127(2), 115-142.
- [40] İzard, CE (1992). Temel duygular, duygular arası ilişkiler ve duygu-biliş ilişkileri

- [41] Kam, CM, Greenberg, MT ve Kusche, CA (2004). PATHS müfredatının özel eğitimdeki çocukların sosyal ve psikolojik uyumu üzerindeki kalıcı etkileri. *Duygusal ve davranışsal bozukluklar dergisi*, 12 (2), 66-78.
- [42] Ekman, P. ve Cordaro, D. (2011). Duyguları temel olarak adlandırmakla ne kastedilmektedir. *Duygu incelemesi*, 3 (4), 364-370.
- [43] Ekman, P. (2016). What Scientists Who Study Emotion Agree About. *Perspectives on Psychological Science*. <https://doi.org/10.1177/1745691615596992>
- [44] Ekman, P. ve Friesen, W. V. (1982). False, felt, and miserable smiles. *Journal of Nonverbal Behavior*, 6(4), 238–258. Retrieved from http://scholar.google.com/scholar?q=related:-8pw6P4UvJcJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5.
- [45] Ekman, P. Sorenson, E. R. ve Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science*, 164(3875), 86-88. doi: 10.1126/science.164.3875.86
- [46] Adolphs, R., Damasio, H., Tranel, D., ve Damasio, A. R. (1996). Cortical systems for the recognition of emotion in facial expressions. *Journal of Neuroscience*, 16(23), 7678-7687. <https://doi.org/10.1523/JNEUROSCI.16-23-07678.1996>.
- [47] Aldinger, M., Stopsack, M., Barnow, S., Rambau, S., Spitzer, C., Schnell, K., ve Ulrich, I. (2013). The association between depressive symptoms and emotion recognition is moderated by emotion regulation. *Psychiatry Research*, 205(1), 59-66. <https://doi.org/10.1016/j.psychres.2012.08.032>
- [48] Demetriooff, S., Porter, S., ve Baker, A. (2017). I know how you feel: the influence of psychopathic traits on the ability to identify micro-expressions. *Psychology, Crime & Law*, 23(3), 274-290. doi: <https://doi.org/10.1080/1068316X.2016.1247159>.
- [49] Plutchik, R. (1962). *The Emotions: Facts. Theories and a New Model*, New York.
- [50] Fridlund, A. J. (1997). The new ethology of human facial expressions. Russell, J.A. ve Fernandez – Dols, J.M. (Ed.), *The psychology of facial expression*, (s. 103- 129). Cambridge University Press.

- [51] Averill, J. R. (1994). In the eyes of the beholder. Ekman, P. ve Davidson, R. J. (Ed.), The nature of emotion: Fundamental questions. (s. 7-14) içinde. Oxford University Press
- [52] Plutchik, R. (1962). The Emotions: Facts. Theories and a New Model, New York
- [53] Ekman, P., ve Friesen, W. V. (1983). EMFACS facial coding manual. Human Interaction Laboratory, San Francisco
- [54] Ekman, P., ve Rosenberg, E. L. (2005). What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). (2.ed.). Oxford University Press, USA
- [55] Batty, M., ve Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research*, 17(3), 613-620.
[https://doi.org/10.1016/S09266410\(03\)00174-5](https://doi.org/10.1016/S09266410(03)00174-5)
- [56] Ekman, P. (1971). Universals and cultural differences in facial expressions of emotion. Nebraska symposium on motivation içinde. University of Nebraska Press.
- [57] Elfenbein, H. A., Marsh, A. A., ve Ambady, N. (2002). Emotional intelligence and the recognition of emotion from facial expressions. *The wisdom in feeling: Psychological processes in emotional intelligence*, 37-59.
- [58] Keltner, D., ve Ekman, P. (2003). Expression of emotion. *Handbook of Affective Sciences*. Oxford University Press, New York, 411-414
- [59] Matsumoto, D., ve Ekman, P. (2004). The relationship among expressions, labels, and depiction of contempt. *Journal of Personality and Social Psychology*, 87(4), 529.
<http://dx.doi.org/10.1037/0022-3514.87.4.529>
- [60] Matsumoto, D. (1989). Cultural influences on the perception of emotion. *Journal of Cross-Cultural Psychology*, 20(1), 92-105. doi:<https://doi.org/10.1177/0022022189201006>
- [61] Ekman, P., ve Friesen, W. V. (1986). A new pan-cultural facial expression of emotion. *Motivation and Emotion*, 10(2), 159-168.

- [62] Crivelli, C., Jarillo, S., Russell, J. A., ve Fernández-Dols, J.-M. (2016). Reading emotions from faces in two indigenous societies. *Journal of Experimental Psychology: General*.
<https://doi.org/10.1037/xge0000172>.
- [63] Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., ve Caldara, R. (2009). Cultural Confusions Show that Facial Expressions Are Not Universal. *Current Biology*.
<https://doi.org/10.1016/j.cub.2009.07.051>.
- [64] Ekman, P. (2003a). *Emotions revealed* (2nd ed.). New York: Times Books.
- [65] Hall, J. A., ve Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion*, 4(2), 201. <http://dx.doi.org/10.1037/1528-3542.4.2.201>
- [66] Ekman, P., ve Friesen, W. V. (1978). *Facial Action Coding System: Investigator's Guide*. Consulting Psychologists Press.
- [67] Bütün, G. (2014). Renk ve doku öznitelikleri kullanılarak içerik tabanlı görüntü erişimi.
- [68] Aktaş, S. (2020). Sürücü performans takip sistemi (Master's thesis, Fen Bilimleri Enstitüsü).
- [69] B. C. Ko, "A brief review of facial emotion recognition based on visual information," *Sensors* (Basel, Switzerland), vol. 18, no. 2, p. 401, Jan 2018.
- [70] Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, 18(7), 1527-1554.
- [71] Çetinkaya, H. H., & Akçay, M. (2012). Yüz tanıma sistemleri ve uygulama alanları. *Akademik Bilişim'12*.
- [72] İŞLEYEN, F., GÜLKESEN, K. H., SAMUR, M. K., ZAYİM, N., & CİNEMRE, B. (2012). Yüz İfadelerini Öğreniyorum: Şizofreni Hastalarına Yüz İfadelerini Öğretmek İçin Geliştirilmiş Oyun Web Sitesi.
- [73] KEKÜL, H., BİRCAN, H., & ARSLAN, H. (2018). Yüz Tanıma Uygulamalarında Özyüzler Ve Yapay Sinir Ağlarının Karşılaştırılması. *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi*, 2(1), 51-59.

- [74] Filiz, S. (2018). Siber gvenlikte biyometrik sistemler Ve yz tanıma (Master's thesis, Biliřim Enstits).
- [75] DEMİR, C. G., & YILMAZ, H. (2018). Sınıf dıřı eēitim faaliyetlerinin ērencilerin bilim ve teknolojiye ynelik tutumlarına etkisi ve duygu analizi. İnsan ve Toplum Bilimleri Arařtırmaları Dergisi, 7(5), 101-116.
- [76] Alagney, A. H. (2015). Fotoantropometri Ynteminin Adli Yz Karřılařtırmalarında Kullanımı. Antropoloji, (29), 117-130.
- [77] Yılmaz, B. (2021). Bulut biliřim tabanlı yz tanıma teknolojisi ile akıllı duyuru sistemi tasarımı (Master's thesis, Kocaeli niversitesi, Fen Bilimleri Enstits).
- [78] Yılmaz, A. Yz ifadelerinden cinsiyet, duygu ve yař tahmini iin yeni yaklařımlar (Master's thesis, Fen Bilimleri Enstits).
- [79] ABUT, S. (2015). Yz grntlerinden cinsiyet tahmini/Yz grntlerinden cinsiyet tahmini.
- [80] Hall, J. A., and Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion* 4, 201–206. doi: 10.1037/1528-3542.4.2.201
- [81] Gneř, T. (2009). Grnt dizilerinden insan yz ifade analizi (Master's thesis, Fen Bilimleri Enstits).
- [82] Elfatimi, L. (2018). Farklı sayısal grnt iřleme teknikleri ile biyometrik grnt analizi (Master's thesis, Pamukkale niversitesi Fen Bilimleri Enstits).
- [83] YAMAN, K., SARUCAN, A., Mehmet, A. T. A. K., & AKTRK, N. (2001). Dinamik izelgeleme iin grnt iřleme ve arıma modelleri yardımıyla veri hazırlama. Gazi niversitesi Mhendislik Mimarlık Fakltesi Dergisi, 16(1), 19-40.
- [84] Samtař, G., & Glesin, M. (2011). Sayısal grnt iřleme ve farklı alanlardaki uygulamaları. *Ejovoc (Electronic Journal of Vocational Colleges)*, 2(1), 85-97.
- [85] Eriřti E. 2010. Grnt iřlemede Yeni Bir Soluk, OPENCV, Akademik Biliřim'10 - XII.

Akademik Bilişim Konferansı Bildirileri, 10-12 Şubat, Muğla.

- [86] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 104 2016.
- [87] S. Skansi, Introduction to deep learning: From Logical Calculus to Artificial Intelligence vol.114 no. 6. 2018.
- [88] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nat. Methods, vol. 13, no. 1, p. 35, 2015.
- [89] Onan, A., & Korukoğlu, S. (2016). Makine öğrenmesi yöntemlerinin görüş madenciliğinde kullanılması üzerine bir literatür araştırması. Pamukkale University Journal of Engineering Sciences, 22(2).
- [90] Çolakoğlu, A. A. (2020). Makine öğrenmesi algoritmaları ile Avrupa havalimanları analizi (Master's thesis, Pamukkale Üniversitesi Sosyal Bilimleri Enstitüsü).
- [91] N. M. Nawi, W. H. Atomi, and M. Z. Rehman, "The Effect of Data Preprocessing on Optimized Training of Artificial Neural Networks," Procedia Technol., vol.11, pp. 32–39, Jan. 2013.
- [92] Zhang, L. ve Zhang, B. (1999). McCulloch-Pitts sinir modeli ve uygulamalarının geometrik bir temsili. Sinir Ağlarında IEEE İşlemleri, 10 (4), 925-929.
- [93] Ayşe, A. R. I., & Berberler, M. E. (2017). Yapay sinir ağları ile tahmin ve sınıflandırma problemlerinin çözümü için arayüz tasarımı. Acta Infologica, 1(2), 55-73.
- [94] Yavuz, S., & Deveci, M. (2012). İstatiksel Normalizasyon Tekniklerinin Yapay Sinir Ağı Performansına Etkisi. Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, (40), 167187.
- [95] BAŞARAN, E., CÖMERT, Z., & CELİK, Y. (2021). Timpanik Membran Görüntü Özellikleri Kullanılarak Sınıflandırılması. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 33(2), 441-453.
- [96] Gündüz, G., & Cedimoğlu, İ. H. (2019). Derin öğrenme algoritmalarını kullanarak görüntüden cinsiyet tahmini. Sakarya University Journal of Computer and Information Sciences, 2(1), 9-17.
- [97] Kavzoğlu, T., & Çölkesen, İ. (2010). Karar ağaçları ile uydu görüntülerinin sınıflandırılması. Harita Teknolojileri Elektronik Dergisi, 2(1), 36-45.

[98] Kavzoğlu, T., & Çölkesen, İ. (2010). Karar ağaçları ile uydu görüntülerinin sınıflandırılması. *Harita Teknolojileri Elektronik Dergisi*, 2(1), 36-45.

[99] Bruna, J., Sprechmann, P., & LeCun, Y. (2015). Super-resolution with deep convolutional sufficient statistics. *arXiv preprint arXiv:1511.05666*.

[100] Üzen, H., Yeroğlu, C., & Hanbay, D. (2019, September). Development of CNN architecture for honey bees disease condition. In *2019 International Artificial Intelligence and Data Processing Symposium (IDAP)* (pp. 1-5). IEEE.

[101] Hüseyin, Ü. Z. E. N., & HANBAY, K. (2020). Yaya özellik tanıma için LM filtre temelli derin evrimsel sinir ağı. *Politeknik Dergisi*, 23(3), 605-613.

[18] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.

[19] Li, S., Jang, S. ve Sung, Y. (2019). Gelişmiş GAN kullanarak otomatik melodi kompozisyonu. *Matematik*, 7 (10), 883.

[20] Kahou, SE, Pal, C., Bouthillier, X., Froumenty, P., Gülçehre, Ç., Memisevic, R., ... & Wu, Z.

(2013, Aralık). Videoda duygu tanıma için modaliteye özgü derin sinir ağlarını birleştirmek. 15.

ACM Bildiriler Kitabı'nda, *multimodal etkileşim üzerine uluslararası konferansta* (s. 543-550). [105] Giannopoulos, P., Perikos, I. ve Hatzilygeroudis, I. (2018).

Yüzdeki duygu tanıma için derin

öğrenme yaklaşımları: FER-2013 üzerine bir vaka çalışması. *Akıllı yöntemlerin hibridizasyonundaki gelişmeler* (s. 1-16). Springer, Şam.

[106] Tümen, V., Söylemez, Ö. F., & Ergen, B. (2017, Eylül). Evrimsel sinir ağı kullanan bir veri kümesinde yüz duygu tanıma. 2017 yılında *Uluslararası Yapay Zeka ve Veri İşleme Sempozyumu (IDAP)* (s. 1-5). IEEE.

- [107] Tümen, V., Söylemez, Ö. F., & Ergen, B. (2017, September). Facial emotion recognition on a dataset using convolutional neural network. In *2017 International Artificial Intelligence and Data Processing Symposium (IDAP)* (pp. 1-5). IEEE.
- [108] Favara, B. E., Feller, A. C., Pauli, M., Jaffe, E. S., Weiss, L. M., Arico, M., ... & Pritchard, J. (1997). Contemporary classification of histiocytic disorders. *Medical and Pediatric Oncology: The Official Journal of SIOP—International Society of Pediatric Oncology (Société Internationale d'Oncologie Pédiatrique, 29(3), 157-166.*
- [109] Shih, FY, Chuang, CF ve Wang, PS (2008). JAFFE veri tabanında yüz ifadesi tanımanın performans karşılaştırmaları. *Uluslararası Örüntü Tanıma ve Yapay Zeka Dergisi , 22 (03), 445459.*
- [110] Lundqvist, D., Flykt, A., & Öhman, A. (1998). Karolinska directed emotional faces. *Cognition and Emotion.*
- [111] Calvo, M. G., Fernández-Martín, A., Gutiérrez-García, A., & Lundqvist, D. (2018). Selective eye fixations on diagnostic face regions of dynamic emotional expressions: KDEF-dyn database. *Scientific reports, 8(1), 1-10.*
- [112] Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2017). Affectnet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing, 10(1), 18-31.*
- [113] Özcan, T. (2020). Derin öğrenme ile insan edimlerinin tanınması.
- [114] Susskind, J.M., Anderson, A.K., and Hinton, G.E., 2010. The Toronto face database, Technical report, University of Toronto, Department of Computer Science, ON, Canada, vol. 3.

- [115] Goodfellow, I.J., Erhan, D., Carrier, P.L., Courville, A., Mirza, M., Hamner, B., Cukierski, W., Tang, Y., Thaler, D., Lee, D.H., Zhou, Y., Ramaiah, C., Feng, F., Li, R., Wang, X., Athanasakis, D., Shawe-Taylor, J., Milakov, M., Park, J., Ionescu, R., Popescu, M., Grozea, C., Bergstra, J., Xie, J., Romaszko, L., Xu, B., Chuang, Z., and Bengio, Y., 2015. Challenges in representation learning: A report on three machine learning contests, *Neural Networks*, 64:59 – 63
- [116] Dhall, A., Goecke, R., Lucey, S., and Gedeon, T., 2011. Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark, 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), 2106–2112.
- [117] Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D.H.J., Hawk, S.T., and van Knippenberg, A., 2010. Presentation and validation of the Radboud Faces Database, *Cognition and Emotion*, 24(8):1377–1388
- [118] Lundqvist, D., Flykt, A., and Ohman, A., 1998. The Karolinska Directed Emotional Faces – KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section.
- [119] Lyons, M., Akamatsu, S., Kamachi, M., and Gyoba, J., 1998. Coding facial expressions with Gabor wavelets, *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, 200–205.
- [120] Devries, T., Biswaranjan, K., and Taylor, G.W., 2014. Multi-task learning of facial landmarks and expression, *Canadian Conference on Computer and Robot Vision*, 98–103, IEEE.
- [121] Ding, H., Zhou, S.K., and Chellappa, R., 2017. Facenet2expnet: Regularizing a deep face recognition net for expression recognition, 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), 118–126, IEEE

- [122] Kim, B., Dong, S., Roh, J., Kim, G., and Lee, S., 2016. Fusing Aligned and Non-aligned Face Information for Automatic Affect Recognition in the Wild: A Deep Learning Approach, 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 1499–1508.
- [123] Guo, Y., Tao, D., Yu, J., Xiong, H., Li, Y., and Tao, D., 2016. Deep neural networks with relativity learning for facial expression recognition, 2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 1–6, IEEE.
- [124] Zhang, Z., Luo, P., Loy, C.C., and Tang, X., 2015. Learning social relation traits from face images, Proceedings of the IEEE International Conference on Computer Vision, 3631–3639.
- [125] Tang, Y., 2013. Deep Learning using Support Vector Machines, CoRR, 1306.0239:1–5.
- [126] Pramerdorfer, C. and Kampel, M., 2016. Facial Expression Recognition using Convolutional Neural Networks: State of the Art, CoRR, 1612.02903:1–6.
- [127] Venkatraman, S., Balasubramanian, S., and Gera, D., 2017. Multiple face-component analysis: A unified approach towards facial recognition tasks, 2017 2nd International Conference on Man and Machine Interfacing (MAMI), 1–6, IEEE
- [128] Jaffe, A. B., & Palmer, K. (1997). Environmental regulation and innovation: a panel data study. *Review of economics and statistics*, 79(4), 610-619.
- [129] Liu, P., Han, S., Meng, Z., and Tong, Y., 2014. Facial Expression Recognition via a Boosted Deep Belief Network, 2014 IEEE Conference on Computer Vision and Pattern Recognition, 1805– 1812.

- [130] Hamster, D., Barros, P., and Wermter, S., 2015. Face expression recognition with a 2-channel Convolutional Neural Network, 2015 International Joint Conference on Neural Networks (IJCNN), 1–8.
- [131] Goodfellow et al., “Generative Adversarial Nets,” ArXiv, Jun. 2014
- [132] İ.N. İ.K Ozkan and E. “ Ulker, “Derin “ Öğrenme Ve Göörüntü Analizinde Kul- “ lanılan Derin Öğrenme Modelleri,” “ Gaziosmanpaşa Bilimsel Araştırma Dergisi, vol. 3, no. 6, pp. 85– 104, 2017.
- [133] Z. Alom, “Improved Deep Convolutional Neural Networks (DCNN) Approaches for Computer Vision and Bio-Medical Imaging,” Ph.D. dissertation, Dept. Electrical and Computer Engineering, University of Dayton, USA, 2018
- [134] H. Kaiming, “Deep residual learning for image recognition,” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016
- [135] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778)
- [136] He, K., Zhang, X., Ren, S., ve Sun, J., 2016, Deep residual learning for image recognition. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern Recognition.
- [137] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).
- [138] Batista, GE, Prati, RC ve Monard, MC (2004). Makine öğrenimi eğitim verilerini dengelemek için çeşitli yöntemlerin davranışı üzerine bir çalışma. *ACM SIGKDD keşif bülteni*, 6 (1), 20-29.