



T.C.

BARTIN ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLİŞİM SİSTEMLERİ VE TEKNOLOJİLERİ ANABİLİM DALI

YÜKSEK LİSANS TEZİ

**GERÇEK ZAMANLI DUYGU DURUMU ANALİZİ: DERİN
ÖĞRENME TABANLI AKILLI SİSTEM TASARIMI**

TUĞBA GÜLER

DANIŞMAN

DR. ÖĞR. ÜYESİ AHMET BERK ÜSTÜN

BARTIN-2022



T.C.

BARTIN ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİLİŞİM SİSTEMLERİ VE TEKNOLOJİLERİ ANABİLİM DALI

**GERÇEK ZAMANLI DUYGU DURUMU ANALİZİ: DERİN ÖĞRENME
TABANLI AKILLI SİSTEM TASARIMI**

YÜKSEK LİSANS TEZİ

TUĞBA GÜLER

BARTIN-2022

BEYANNAME

Bartın Üniversitesi Lisansüstü Eğitim Enstitüsü tez yazım kılavuzuna göre Dr. Öğr. Üyesi Ahmet Berk ÜSTÜN danışmanlığında hazırlamış olduğum “Gerçek Zamanlı Duygu Durumu Analizi: Derin Öğrenme Tabanlı Akıllı Sistem Tasarımı” başlıklı yüksek lisans tezimin bilimsel etik değerlere ve kurallara uygun, özgün bir çalışma olduğunu, aksinin tespit edilmesi halinde her türlü yasal yaptırımını kabul edeceğimi beyan ederim.

06.07.2022

Tuğba GÜLER



ÖNSÖZ

Bu tez çalışmasında, canlı kamera görüntüsünden yüz duygu tanıma, cinsiyet ve yaş aralığı belirleme ve aynı zamanda duygu durumuna bağlı olarak da kişinin stres oranı saptanmıştır. Bu işlemlerin yapılabilmesi için günümüzde çok hızlı şekilde gelişen çeşitli derin öğrenme tekniklerine başvurulmuş ve farklı modeller oluşturulmuştur. Yapılan çalışmanın eğitim, ticaret, sağlık ve güvenlik gibi farklı alanlarda fayda sağlayacağı düşünülmektedir.

Öncelikle tez çalışmamın seçim, planlama ve yürütülme aşamalarında bana her konuda yardımcı olup destekte bulunan tez danışmanım Sayın Dr. Öğr. Üyesi Ahmet Berk ÜSTÜN'e katkılarından dolayı sonsuz teşekkürlerimi sunarım. Hayatın her aşamasında olduğu gibi bu aşamada da manevi desteğini üzerimden hiç esirgemeyen annem, babam ve kardeşlerime, minnetimi özellikle belirtmek istiyorum. Son olarak her konuda yardım ve desteğini her anlamda hissettiren değerli eşim Evrim GÜLER ve sevgili oğlum Arda GÜLER'e sonsuz teşekkürler.

Tuğba GÜLER

ÖZET

Yüksek Lisans Tezi

GERÇEK ZAMANLI DUYGU DURUMU ANALİZİ: DERİN ÖĞRENME TABANLI AKILLI SİSTEM TASARIMI

Tuğba GÜLER

Bartın Üniversitesi

Lisansüstü Eğitim Enstitüsü

Bilişim Sistemleri ve Teknolojileri Anabilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi Ahmet Berk ÜSTÜN

Bartın-2022, sayfa: 72

Duygu tanıma ve davranış yargısı gibi duyguları anlamamanın anahtarı görevini gören yüz ifadelerini anlama, insanların günlük hayatında oldukça değerlidir. Anlık olarak sistemlerde algılanan görüntü ve videodan duyguları tanımak, her ne kadar insanların gözleri için kolay ve önemsiz olsa da makineler için çok doğru bir şekilde algılanması zorlayıcı olup uygulanmasında birçok görüntü işleme tekniği gerektirmektedir.

Günümüzde önemli bir yere sahip olan yüz duygu durumu sistemleri, her geçen gün daha çeşitli alanlarda uygulanıp ilgi çekmeye devam etmektedir. Yüz duygu durumu ile birlikte yaş ve cinsiyetin canlı şekilde kameradan belirlenmesinin, birçok potansiyel amaca hizmet edeceği tahmin edildiğinden bu çalışmada canlı kamera görüntüsündeki yüzlerin duygu durumu, cinsiyeti, yaş aralığı ve duygu durumuna bağlı olarak da stres oranı hesaplanmaktadır. Çalışmada veri seti olarak FER2013, FER+ ve UTKFace kullanılmış olup, CNN modellerinden ise özelleştirilmiş CNN modeli, VGG-16, ResNet-50 ve ResNet-152 mimarileri tercih edilmiştir. Elde edilen sonuçlarda en iyi performansı ResNet-50 modeli sağlamıştır. Tüm tahmin edilen anlık durumlar için gerçek hayatta eğitim, sağlık, iş güvenliği, suç tespitleri, ticaret vb. gibi farklı uygulamalarda oluşturulan sistemin kullanımı

önceden uyarıcılar oluşturmaya da yardımcı olabilecektir. Hali hazırda uygulanan birçok sistem günlük hayatta mevcut olsa da bu sistemleri aynı anda tek bir çatı altında toplayarak bir adım daha ileriye götürmeyi amaçlayıp bu çalışmayı daha önce literatürde olmayan bir şekilde tamamlamış bulunuyoruz.

Anahtar Kelimeler: Derin Öğrenme, Görüntü İşleme, Yapay Zeka, Duygu Analizi



ABSTRACT

M. Sc. Thesis

REAL-TIME EMOTION ANALYSIS: INTELLIGENT SYSTEM DESIGN BASED ON DEEP LEARNING

Tuğba GÜLER

Bartın University

Graduate School

Department of Information System and Technology

Thesis Advisor: Assist. Prof. Dr. Ahmet Berk ÜSTÜN

Bartın-2022, pp: 72

Understanding facial expressions, which serve as the key to understanding emotions such as emotion recognition and behavioral judgment, is very valuable in people's daily lives. Although it is easy and insignificant for people's eyes, it is difficult for machines to perceive it accurately and requires many image processing techniques to be applied.

Facial emotion recognition systems, which have an important place today, are being applied in various areas and continue to attract attention. Since the prediction of facial emotion with age and gender in live camera captures will serve many potential purposes, in this study, the stress rate is calculated based on the emotional perspectives of the faces in the live camera image. In the study, FER2013, FER+ and UTKFace were used as data sets, and customized CNN model, VGG-16, ResNet-50, and ResNet-152 architectures were preferred among CNN models. In the obtained results, the ResNet-50 model provided the best accuracy performance. All predicted instant situations in real life are able to help in various application areas such as education, health, job security, crime detection, trade, etc. Although many systems that are currently applied exist in daily life, we have completed this study in a way that has not been in the literature before, aiming to take these systems one step further by gathering these systems under one roof at the same time.

Keywords: Deep Learning, Image Processing, Artificial Intelligence, Emotion Analysis



İÇİNDEKİLER

BEYANNAME	ii
ÖNSÖZ	iii
ÖZET	iv
ABSTRACT	vi
İÇİNDEKİLER.....	viii
ŞEKİLLER DİZİNİ.....	xi
SİMGELER VE KISALTMALAR DİZİNİ.....	xiii
1. GİRİŞ.....	1
1.1 Problem.....	1
1.2 Amaç	2
1.3 Önem	3
1.4 Sınırlılıklar	4
2. KURAMSAL TEMELLER ve KAYNAK ARAŞTIRMASI	5
2.1. Yüz Duygu Durumu ve Önemi	6
2.2. Yüz Duygu Tanıma.....	7
2.3. Yüz Duygu Tanıma Kullanım Alanları	7
2.4. Yüz Duygu Tanımda Risk, Sorun ve Eksiklikler.....	9
2.5. Yapay Zeka (Artificial Intelligence – AI)	9
2.5.1. Zayıf Yapay Zeka (ANI)	11
2.5.2. Güçlü Yapay Zeka (AGI)	11
2.5.3. Süper Yapay Zeka (ASI)	12
2.5.3.1. Makine Muhakemesi (Machine Reasoning - MR):	12
2.5.3.2. Robotik:.....	12
2.5.3.3. Makine Öğrenimi (Machine Learning - ML):	12
2.5.3.4. Derin Öğrenme (Deep Learning - DL):	12
2.6. Makine Öğrenimi (ML) Sistemlerinin Yapısı	12
2.6.1. Veri Kümeleri:.....	14
2.6.2. Özellikler:.....	14
2.6.3. Algoritma:	14
2.6.4. Denetimli Öğrenme	15
2.6.5. Denetimsiz Öğrenme.....	16
2.6.6. Yarı Denetimli Öğrenme	17

2.6.7. Pekiştirmeli Öğrenme (Reinforcement Learning – RL).....	17
2.7. Derin Öğrenme (Deep Learning – DL)	18
3. YÖNTEM	22
3.1. Anlık Duygu Analizi, Stres, Cinsiyet ve Yaş Belirleme Mimarisi	22
3.2. Dijital Görüntü İşleme (Digital Image Processing - DIP) Mimarisi	23
3.3. Dijital Görüntü İşleme Uygulamaları	25
3.3.1. Tıbbi Görüntü Alma	25
3.3.2. Trafik Algılama Teknolojileri.....	25
3.3.3. Görüntü Yeniden Yapılandırma.....	26
3.3.4. Yüz Tanıma	26
3.4. Dijital Görüntü İşlemenin Faydaları	27
3.5. Evrişimli Sinir Ağları (Convolutional Neural Networks – CNN).....	28
3.6. Evrişimli Sinir Ağlarının İlham Kaynakları	28
3.7. CNN Mimarisi	29
3.7.1. CNN Katmanları	30
3.7.1.1 Evrişimsel Katman	30
3.7.1.2. Havuzlama Katmanı	34
3.7.1.3. Tam Bağlantılı Katman	35
3.7.2. Ağ Eğitimi	40
3.7.3. Kayıp Fonksiyonu	41
3.7.4. Gradyan İnişi.....	41
3.8. Veri ve Temel Doğruluk Etiketleri.....	42
3.8. Aşırı Uyum Gösterme (Overfitting)	43
3.9. Evrişimli Sinir Ağları Uygulama Alanları.....	44
3.10. Yüz Duygu Tanıma (Facial Emotion Recognition – FER) Modeli.....	44
3.11. Duygu Durumu Analizi Bölümü.....	45
3.11.1. Duygu Durumu Analizi İçin Kullanılan Modeller	45
3.11.1.1. Özelleştirilmiş CNN Modeli.....	46
3.11.1.2. VGG-16.....	46
3.11.1.3. ResNet-50	47
3.11.1.4. ResNet-152	48
3.12. Cinsiyet ve Yaş Analizi	48
3.12.1. Yaş Analizi Özelleştirilmiş Modeli	48

3.12.2. Cinsiyet Analizi Özelleştirilmiş Modeli.....	49
3.13. Yüz Duygu Durumlarından Stres Analizi	49
4. DENEYSEL KURULUM VE SONUÇLAR	50
4.1. Duygu Durumu Analizi İçin Kullanılan Veri Kümeleri.....	50
4.1.1. FER2013 Veri Kümesi.....	50
4.1.2. FER+ Veri Kümesi.....	50
4.1.3. Cinsiyet ve Yaş Analizi İçin Kullanılan Veri Seti	51
4.2. Veri Kümeleri Üzerinde Eğitim ve Test Modelleri.....	51
4.2.1. Uygulanan Özelleştirilmiş CNN Modeli.....	54
4.2.2. Uygulanan VGG-16 Modeli.....	54
4.2.3. Uygulanan ResNet-50 ve ResNet-152 Modelleri.....	54
5. TARTIŞMA VE GELECEK ÇALIŞMALAR	60
6. SONUÇ	63
KAYNAKLAR.....	65

ŞEKİLLER DİZİNİ

Şekil	Sayfa
No	No
2.1. Yapay Zeka Sistemleri.....	10
2.2. Yapay Zeka Sistemlerinin İlişkisel Yapıları.....	13
2.3. Makine Öğrenmesi Bileşenleri.....	13
2.4. Makine Öğrenmesi Yöntemlerinin Çalışma Mimarisi.....	15
2.5. Makine Öğrenmesi Kategorileri.....	16
2.6. Makine Öğrenimi ve Derin Öğrenme Mimarisi.	18
2.7. Sinir Ağları Katman Mimarisi.	19
2.8. Derin Öğrenme Modellerinin Uygulama Alanları.....	20
2.9. Görüntü İşleme Mimarisi.....	21
3.1. Dijital Görüntü İşleme Adımları.....	23
3.2. Görüntü İşleme ile Kesişen Alanlar.....	27
3.3. CNN Mimarisi.....	30
3.4. Evrişim İşlemi Örneği (Yamashita, Nishio, Do ve Togashi, 2018)	32
3.5. Sinir ağlarına yaygın olarak uygulanan etkinleştirme işlevleri: a) doğrultulmuş doğrusal birim (ReLU), b) sigmoid, ve c) hiperbolik tanjant (tanh).....	34
3.6. Maksimum Havuzlama Örneği.....	35
3.7. Küresel Ortalama Havuzlama Örneği.....	35
3.8. ReLU Aktivasyon İşlevi Girdi ve Çıktı Değerleri.....	38
3.9. Çıktı Katmanında Aktivasyon Fonksiyonu Seçimi.....	39
3.10. Lineer Aktivasyon Fonksiyonu.....	39
3.11. Sigmoid Aktivasyon Fonksiyonu.....	40
3.12. CNN modelinde Ağ Eğitim Modeli.....	41
3.13. Veri Modelleri ve Etiketleri.....	43
3.14. Yüz Duygu Tanıma İş Akışı.....	45
3.15. VGG-16 Mimarisi.....	46
3.16. ResNet-50 Mimarisi.....	48
4.1. Duygu Durumu Tanıma (Facial Emotion Recognition – FER) Veri Kümesi.....	50
4.2. Duygu Durumu Tanıma+ (Facial Emotion Recognition – FER+) Veri Kümesi.....	51
4.3. UTKFace Veri Seti.....	52
4.4. Özelleştirilmiş CNN Modeli.....	53

4.5. VGG-16 Modeli.....	55
4.6. ResNet-50 ve ResNet-152 Modelleri.....	56
4.7. Özelleştirilmiş CNN Modeli Sonuçları.....	57
4.8. VGG-16 Modeli Sonuçları.....	58
4.9. ResNet-50 Modeli Sonuçları.....	58
4.10. ResNet-152 Modeli Sonuçları.....	58
4.11. Özel CNN Modeli Cinsiyet ve Yaş Belirleme Sonuçları.....	59



SİMGELER VE KISALTMALAR DİZİNİ

$\max$	: Maksimum değer alma
e	: Euler sabit sayısı
Σ	: Toplam işareti
w	: Ağırlık ölçütü
α	: Değişim katsayısı
∂	: Türev
$\lceil \rceil$	: Tavan fonksiyonu

KISALTMALAR

ER	: Duygu Tanıma (Emotion Recognition)
FER	: Yüz Duygu Tanıma (Facial Emotion Recognition)
AI	: Yapay Zeka (Artificial Intelligence)
ANI	: Zayıf Yapay Zeka (Artificial Narrow Intelligence)
AGI	: Güçlü Yapay Zeka (Artificial General Intelligence)
ASI	: Süper Yapay Zeka (Artificial Super Intelligence)
MR	: Makine Muhakemesi (Machine Reasoning)
ML	: Makine Öğrenimi (Machine Learning)
DL	: Derin Öğrenme (Deep Learning)
TSVM	: Transdüktif Destek Vektör Makinesi
DNN	: Derin Sinir Ağları (Deep Neural Network)
LSTM	: Uzun-Kısa Dönemli Bellek (Long Short Term Memory)
HCI	: İnsan Bilgisayar Etkileşimi (Human Computer Interaction)
DIP	: Dijital Görüntü İşleme (Digital Image Processing)
VIPS	: Video Görüntü İşleme Sistemi (Video Image Processing System)
CNN	: Evrişimli Sinir Ağı (Convolutional Neural Network)
ANN	: Yapay Sinir Ağları (Artificial Neural Network)
ILSVR	: ImageNet Büyük Ölçekli Görsel Tanıma
FC	: Tam Bağlantılı Katman (Fully Connected Layer)
2D	: İki Boyutlu (Two Dimensional)
ReLU	: Doğrultulmuş Lineer Birim (Rectified Linear Unit)
1D	: Tek Boyutlu (1-Dimensional)

MLP : Çoklu Katman Algılayıcıları (MultiLayer Perceptron)
ICML : Uluslararası Makine Öğrenimi Konferansı
VGG : Görsel Geometri Grubu (Visual Geometry Group)



1. GİRİŞ

Bu bölümde araştırmanın problem durumu, araştırmanın amacı, önemi ve sınırlılıkları bulunmaktadır.

1.1 Problem

Yüzler, insanların sosyal etkileşimlerinin vazgeçilmez unsurlarıdır. Bir insanın yüzü, kişinin başka kişiler tarafından olumlu veya uygun olarak değerlendirilmeye olan ilgisine göre oluşturduğu bir ifade belirtir (Lawrence vd. , 2015). Yüzlerin tek görevi karşıdaki bireyin yaşını, cinsiyetini, kimliğini ve aşinalığını öğrenmek değildir, aynı zamanda karşıdaki kişinin duygularını da anlamamıza yardımcı olmaktır. Ayrıca, başka kişilerin sahip oldukları veya kendilerinin bilincinde oldukları ve diğer kişilerin de farkında oldukları saygıyı veya güveni yansıtan bir kişinin sosyal temsilidir (Azcarate vd., 2005). Yüz ifadelerini algılayıp analiz etmek, sosyal hayatımızda karşımıza çıkabilecek olumsuzlukları düzeltmemize de katkı sağlamaktadır (Mellouk ve Handouzi, 2020).

Yüz ifadelerinden altısı (mutluluk, üzüntü, korku, şaşkınlık, iğrenme ve öfke), yürütülen psikolojik araştırmalar sonucu dünya çapında kabul edilmiştir (Khairuddin ve Chen, 2021). Bu altı yüz duygu durumunun analizi ile insanların yaşlanma etkileri arasındaki bağıntıya yönelik ilgi oldukça fazladır (Orgeta ve Phillips, 2008). Sosyo-duygusal seçicilik teorisine göre artan yaşlarıyla birlikte kişilerin dikkati duygusal açıdan yoğun olaylara yönelmektedir. Böylece her geçen gün insanlar duygu durumlarını yüzlerinde daha fazla yansıtmaktadırlar (Carstensen vd., 2003). Dolayısıyla yüz duygu durumu belirleme, yaş ile ilgili bilgi almada da yardım sağlamaktadır.

Psikoloji alanında olduğu kadar tarihsel açıdan da uzun yıllar tartışma konusu olan cinsiyet farklılıkları matematik, bilim ve sosyal alanlarda bir cinsiyetin diğer cinsiyete olan üstünlüğü gibi konularda öne çıkmıştır (Hoffmann vd., 2010). Günümüzde hala bu tip tartışmalar söz konusudur. Bunların dışında, güvenlik konusunda cinsiyet analizi sosyal, asayiş ve askeri açıdan artan gereksinimler arasındadır (Dehghan vd., 2017). Kişinin cinsiyet analizinin önemi işte bu noktalarda değer kazanmaktadır. Farklı cinsiyetlerde mevcut olan farklı yüz oranları gibi ölçümler, cinsiyet belirlenmesinde önemli rol oynamaktadır (Hall, 1978). Bu sebeple yüzün incelenip analiz edilmesi cinsiyet belirlemeyi kolaylaştırmaktadır.

İnsanların bulundukları ruh hallerinin farklılaşması, vücutlarındaki birçok sistemde de fizyolojik açıdan değişikliğe neden olmaktadır (Wei, 2013). Bu değişiklikleri de vücut duruşu, yüz ifadeleri ve ses gibi yöntemlerle göstermektedirler (He vd., 2009). Kişilerin duygu durumlarına bağlı olarak dışa vurduğu durumlardan biri olan stres, insan hayatında önemli bir yere sahiptir (Hartling vd., 2019). Psikoloji alanında yapılan araştırmaların bulgularına göre, orta seviyedeki stres kişilerin bazı sorunlarını çözmelerinde fayda sağlarken aşırı stres ise beyne zarar vermeye kadar giden çeşitli sağlık sorunları doğurabilmektedir (Wei, 2013). Dolayısıyla yüzden belirlenen duygu durumlarına bağlı olarak kişinin stres oranı tespit edilip, ruh sağlığına katkı sağlanabilmektedir.

Son yıllarda tahmin edilmesi zor olan duygu durumları için güçlü bir analiz sistemi olarak adından sıkça bahsedilen ve yüz tanıma alanında öğrenme yöntemleri genel olarak sınıfların yeniden örneklemeyi veya sürece duyarlı öğrenmeyi benimseyen oldukça güçlü bir yapıya sahip olan derin öğrenme, yüz duygu durumu, yaş ve cinsiyet belirleme konularında ilgi çekmektedir (Dehghan vd., 2017). İnsan- bilgisayar etkileşim alanına çokça katkıda bulunan bu sistemler, altyapısında bulunan farklı modellerle desteklenip en iyi sonuçları elde etme konusunda veri bilimcilere ışık olmaktadır.

Yüz ifadelerinden duygu, yaş ve cinsiyet analizi insanlar için önemsiz ve basit bir görev gibi gözükse de makineler için ayrıntı ve bilgi gerektiren bir süreçtir (Mellouk ve Handouzi, 2020). Kişiselleştirilmiş, güvenilir, etik yapıdaki otomatik yüz tanıma sistemlerine olan gereksinim arttıkça insan-bilgisayar etkileşimi alanı gün geçtikçe daha fazla öneme sahip olmaktadır (Wei, 2013). Günümüzde hali hazırda bulunan yüz duygu durumu, yaş, cinsiyet ve stres analiz sistemleri vardır ve farklı alanlarda uygulanmaktadır. Fakat bu sistemlerin güvenilirliği, kullanım alanları ve anlık durumu belirtmesi gibi özellikleri çeşitlilik göstermektedir. Aynı anda hem duygu, yaş ve cinsiyet belirleyen hem de stres oranı ölçen iyi performansa sahip bir sistem de henüz literatürde bulunmamaktadır.

1.2 Amaç

Yüz duygu durumu belirleme, geçtiğimiz on yıl boyunca öne çıkan güncel konulardan biridir (Azcarate vd., 2005). Yaş ve cinsiyetin yüzlerden belirlenebilmesi de aynı oranda ilgi görmüş çeşitli araştırmalara konu olmuştur (Dehghan vd., 2017). Duygu durumuna bağlı olarak stres analizi de son zamanlarda öne çıkan başka bir başlık olmaktadır (Wei, 2013).

Bu çalışmada, yüz analizinin gereksinim duyulduğu alanlarda kamera karşısındaki yüzlerin duygu durumları, yaş ve cinsiyet özellikleri dışında, kişilerin o andaki stres oranlarını ölçerek bulundukları duruma yönelik bilgiler elde edilip gerekli değerlendirmelerin yapılabilmesi ve literatüre katkı sağlanabilmesi amaçlanmaktadır. Çalışmanın cevap vermeyi hedeflediği sorular aşağıda belirtilmiştir:

1. Kamera aracılığıyla kişilerin duygu durumları gerçek zamanlı olarak hangi sınıflandırmalara dahil olmaktadır?
2. Kamera aracılığıyla kişilerin yaşları gerçek zamanlı olarak hangi aralıklara dahil olmaktadır?
3. Kamera aracılığıyla kişilerin cinsiyetleri gerçek zamanlı olarak hangi sınıflara dahil olmaktadır?
4. Kamera aracılığıyla kişilerin duygu durumlarına bağlı stres oranları gerçek zamanlı olarak hangi gruba dahil olmaktadır?
5. Potansiyel olarak çalışmanın katkı sağlayabileceği alanlar hangileridir?

1.3 Önem

İnsan yüzleri, sadece nefes alma, duyma, ve görme gibi duyulara hizmet etme gibi faydalar sağlamaz, aynı zamanda kişiler hakkında detaylı bilgiler edinmemize de yardımcı olur. Yüz kaslarının ya da organlarının bulunduğu pozisyon ve ölçümlere dayalı olarak duygu durumları, yaş aralıkları, cinsiyetleri ve stres oranları hesaplanabilmektedir (Jain vd., 2019).

Yüz ifadeleri, insanların ilişkilerinin geliştirilip düzenlenmesi dışında bilgi aktarımında da etkili bir iletişim aracıdır (Azcarate vd., 2005). Yüz ifadeleri, insanlar arasında sağlıklı iletişimi sağlamak için duygusal durumların aktarımı ve kişilerin durum ve amaçlarını iletmek için kullanılır (Nitisha, 2018). Örnek olarak, sinirli bir kişinin yüzü, içinde bulunduğu durumu kabul etmek istemediğini açıkça belli ederek o durumu kendinden uzaklaştırabilir. Böylece duyguların yüz ifadelerini öğrenme durumu, günlük yaşamdaki sosyal etkileşimlerde fazlasıyla önem taşır (Reddy vd., 2022).

Yüz duygu durumu analizlerinin kullanılabileceği alan çok geniş olmakla birlikte, her geçen gün ihtiyaç ve gereksinimleri değişip gelişmeye devam etmektedir (Josephine Julina ve Sharmila, 2019). Yüz ifadesinden çıkarılabilen duygular, müşteri memnuniyeti için oldukça yararlı olabilmektedir (Lawrence vd., 2015). Örneğin, kameradan elde edilen görüntülerden müşterinin duygu durumu anlaşılıp, almayı düşündüğü ürün hakkında alınan bilgiler o ürünün seçilebilmesi açısından satıcılara fikir verebilmektedir (Mellouk ve Handouzi, 2020).

Bunların dışında, yüz duygu durumu kameradan canlı olarak öğrenilen bir kamyon sürücüsünün, yoldaki emniyeti sağlanıp olası kazaların önüne geçilebilir.

Yüzden elde edilen yaş ve cinsiyet bilgileri de kişi hakkında ayrıntılı bilgiye sahip olup herhangi bir durum için gerekli kararlar vermeye yardımcı olmaktadır (He vd., 2009). Bu durum daha çok suçlu ve kaçak durumda olan kişilerin daha kolay ve hızlı bir şekilde bulunabilmesine hizmet etmektedir (Wei, 2013). Ayrıca, belli bir topluluğa ait sayıca fazla kişinin verilerini depolarken sadece görüntüye dayanarak iş yükünün azaltılıp, zaman açısından azalmayı da sağlayabilmektedir.

Kişinin duygu durumuna bağlı olarak elde edilecek stres bilgisi, o kişinin özel bir durumda nasıl bir davranış göstereceğinin tahmini ve olası sağlık sorunlarının engellenmesi gibi önemli konulara ışık tutarak gelecek adımlarında yönlendirmede bulunabilir.

Gerçekleştirilen çalışmada aynı anda duygu durumu, yaş, cinsiyet ve stres analizini gerçek zamanlı olarak gerçekleştirmek, daha önce yapılan çalışmalardan farklı olmakla birlikte elde edilen performans yönünden de benzerlerine göre gayet iyi bir durum sergilemektedir. Çalışmanın yukarıda belirtilen durum ve alanlar dışında da farklı birçok alanda ve projede hizmet verebilmesi amaçlanmaktadır.

1.4 Sınırlılıklar

Bu çalışma,

- FER2013, FER+ ve UTKFace veri setleri kullanılarak eğitilip,
- Özelleştirilmiş CNN modeli, VGG16, ResNet-50 ve ResNet-152 mimarileri kullanılarak gerçekleştirilmiştir.

Başka araştırmalarda farklı veri seti ve modeller de kullanılabileceğinden yukarıda verilenler sınırlılık başlığı altında verilmiştir.

2. KURAMSAL TEMELLER ve KAYNAK ARAŞTIRMASI

İnsanların yüzleri, görsel algının ekolojik olarak en önemli uyaranlarından biridir. Son yıllarda, gerçekleştirilen algısal ve bilişsel çalışmalar, insanların cinsiyet, yaş, kimlik ve yüz ifadeleri gibi yüz bilgilerini tanımada oldukça iyi olduğunu göstermektedir (Lawrence vd., 2015). Yüz ifadeleri, görsel konuşma için ağız hareketlerinin yanı sıra çeşitli karmaşık şekillerde hızlı bir şekilde ve sürekli olarak değişen yüzdeki tek bilgiyi oluşturdukları için özeldir. Ancak kısa bir bakışla yüzlerde oluşan farklı ifadeler kolayca ayırt edilebilir ve farklı yüz ifadelerinin oluşturduğu doğru anlamı çıkarmak için kişinin mutlaka bireysel olarak tanınması gerekmez. Diğer bir ifadeyle, yüz ifadesinin doğru olarak algılanması yüz kimliğine bağlı olarak değişkenlik göstermemektedir (Marini vd., 2021). Yalnızca algısal ve bilişsel bilimlerdeki uygulamalarla değil, aynı zamanda duyuşsal hesaplama ve bilgisayar animasyonlarındaki uygulamalarla birlikte, yüz ifadesi araştırmalarının son yıllarda çok fazla ilgi görmesi şaşırtıcı olmamaktadır (Azcarate vd., 2005). Diğer türlerle karşılaştırıldığında, insanlar sosyal etkileşimler için oldukça karmaşık iletişim sistemleri geliştirmiştir. 1970 yılında Bridghistell, tipik bir iletişim sırasında sözlü bileşenlerin üçte birini ve sözel olmayan bileşenlerin üçte ikisini sosyal anlamın ilettiğini göstermektedir (Safira vd., 2020). Vücut ifadelerine ek olarak, yüz ifadeleri de sözlü olmayan kişiler arası iletişimde temel bilgi kanallarından biridir. Sözsüz iletişim için önemleri göz önüne alındığında, yüz ifadelerinin duygusal anlam taşıma ve iletişimsel bir bilgi kanalı olarak hizmet etme olarak iki yönü bulunmaktadır (Nitisha, 2018). Bu durum, Darwin tarafından yüz ifadeleri üzerine yaptığı ufuk açıcı çalışmasında da belirtilmektedir (Kaulard vd., 2012).

Yüz ifadelerinin güçlü iletişimsel yönüne rağmen, yüz ifadelerini yalnızca duygusal ifadelerle eşitleme eğilimi yaygın olarak kullanılmaktadır. Karmaşık doğalarını vurgulamaya yardımcı olarak farklı yüz ifadesi sınıflarını üç farklı kategoride: (i) tamamen dönüşümlü yüz duruşları ve hareketleri, (ii) görünüşte duygusal ifadeler ve (iii) dil ötesi veya iletişimsel yüz ifadeleri olarak inceleyebiliriz (Reddy vd., 2022). Bunlara örnek olarak, insanların onaylayan bir baş sallama hareketi de vücut hareketlerindeki ifadeler ile konuşmaya veya iletişim kurma ile ilgilidir (H. Zhang vd., 2019). Yüz ifadeleri, gönderenin gerçek zihinsel durumuna göre de ayırt edilebilir. Bilişsel bir zihinsel duruma işaret eden yüz ifadeleri (örneğin, düşünme, karar verme veya planlama) ve duygusal bir zihinsel duruma işaret eden yüz ifadeleri insanlar tarafından sıklıkla kullanılmaktadır. Bu nedenle, yüz duygu durumu insanların iletişimde önemli bir yere sahiptir (Bull, 2001).

2.1. Yüz Duygu Durumu ve Önemi

İnsan duyguları, günlük yaşamda ve insan iletişiminde son derece önemli bir yere sahiptir. Genellikle, insanın günlük iletişim sürecinde, diğer insanların duygularını anlama, davranış yargısı, duygusal ve zihinsel durumları kişilerin yüz ifadeleriyle uyum içinde olmaktadır. Dolayısıyla insanların yüz ifadelerini anlamak aynı zamanda duyguları anlamanın da anahtarıdır. Duygu Tanıma (Emotion Recognition - ER), kişiyi gözlemleyerek belirli bir duygusal durumda, hedef hakkında geçerli bilgileri analiz etme ve tanımlama, ayrıca duygunun ne anlama geldiğini çıkarma sürecidir (Josephine Julina ve Sharmila, 2019). İnsanlar sosyal yaşamlarında işitsel ve görsel olarak iki ana kanal aracılığıyla iletişim kurarlar. Yüz ifadeleri, güçlü bir görsel kanal olarak yeri doldurulamaz bir rol oynamaktadır (Karbauskaitė vd., 2020). Daha önce yapılan psikolog araştırmalarına göre, insan iletişimi sürecinde bilginin sadece %7'lik kısmı saf dil ifadesinden, %38'lik oranı sestten, %55'lik gibi büyük oranı ise yüz ifadeleri ve vücut duruşu gibi görsellerden anlaşılmaktadır (Cai vd., 2021). Sonuç olarak, insan iletişiminde yüz ifadelerinin doğru tanınması, bilgiyi anlama açısından kritik öneme sahiptir.

Yüz ifadesi ile ilgili çalışmalar sadece sosyal ve klinik psikoloji veya psikodilbilimin ilgi alanı ile sınırlı olmamaktadır. Günümüzde, bilgisayarlar artık günlük hayatımızın çoğu yönünü kaplamaktadır ve bilgisayar biliminin ana amaçlarından biri bu insan-bilgisayar etkileşimini gerçekleştirerek uygulama alanlarını optimize etmektir. İnsan-bilgisayar arayüzlerinin tasarımı, önemli iletişimsel yönleri hesaba kattığı için bilgisayar merkezli tasarımlardan insan merkezli tasarımlara dönüşmüştür. Bu fikir, bilgisayarlarla etkileşimde dikkate alınabilecek insan-insan etkileşiminin tipik özelliklerini ilk kez özetleyen Nickerson'a kadar uzanır (Nickerson, 1976). O zamanlar amaç, bir insan sözsüz iletişim sinyali ile ilgili bilgisayar sinyali arasında analoglar oluşturmaktaydı. Günümüzde, insan-bilgisayar arayüzlerinin önemli bir amacı, yüz yüze konuşmalarda insanlarla aynı iletişim becerilerini sergileyen somutlaşmış sistemler gibi insanların iletişimlerinde kullanacakları yüz ifadelerini veya duygularını tanımaya yarayan araçlar meydana getirmektir (Adolphs, 2002).

2.2. Yüz Duygu Tanıma

Yüz Duygu Tanıma (Facial Emotion Recognition - FER), resimler ve videolar gibi farklı kaynaklar yardımıyla duyguları analiz etmek için kullanılan bir teknolojidir. Genellikle 'duygusal hesaplama' olarak adlandırılan teknolojiler ailesine aittir. Ayrıca, bilgisayarların insan duygularını ve duygusal durumlarını tanıma ve yorumlama yetenekleri üzerine çok disiplinli bir araştırma alanıdır. Yüz ifadeleri, insan duyguları için ipuçları sağlayan sözsüz iletişim biçimleridir. Bu tür duygusal ifadelerin çözülmesi, psikoloji alanında (Harms vd., 2010; Kohler vd., 2003) ve aynı zamanda İnsan-Bilgisayar Etkileşimi alanında (Bani vd., 2021; Harms vd., 2010; Ko, 2018) ilgiyi artırmaktadır. Son zamanlarda, kameraların yüksek difüzyonu ve biyometrik analiz, makine öğrenimi ve örüntü tanımadaki teknolojik gelişmeler FER (Facial Emotion Recognition) teknolojisinin gelişmesinde önemli bir rol oynamıştır. NEC veya Google gibi teknoloji devlerinden Affectiva veya Eyeris gibi daha küçük şirketlere kadar birçok şirket, artan önemi nedeniyle bu teknolojiye yatırım yapmaktadır. Ayrıca teknolojinin kullanımını araştıran Horizon2020 girişimi gibi birkaç Avrupa Birliği araştırma ve yenilik programı da bulunmaktadır ve kullanım alanlarını genişletmeye çalışmaktadır.

2.3. Yüz Duygu Tanıma Kullanım Alanları

Yüz duygu tanıma sistemleri günlük hayatta ve gelecek yıllarda çeşitli durumlar için kullanılabilir (Y. Wang ve Zheng, 2018). Bunların başında insan-bilgisayar etkileşimi açısından insan duygularını belirlemek, makinelerin insanlarla etkileşimde daha uygun, doğru, etkili ve doğal olmasını sağlaması gelmektedir. Ayrıca, güvenlik alanında, suç niyeti taşıyan şüphelileri etkili bir şekilde tespit etmek için kullanılmaktadır. Bunların dışında ulaşım yönünden, yüz duygu tanıma sistemleri sürücülerin yüz ifadelerini tanıyarak yorgun düşen sürücünün yola devam edip edemeyeceğine daha iyi karar verebilmesine imkan vermektedir (Gavrilescu ve Vizireanu, 2019; Hu vd., 2019). Yukarıdaki ifade edilen durumlar dışındaki iletişim alanları, pazarlama, otomasyon, reklam sektörü gibi daha birçok alanda yüz duygu tanıma sistemlerinin kullanımına rastlamak mümkündür (Oh vd., 2021).

FER ayrıca biyometrik tanımlama ile birleştirilebilir. Doğruluğu, ses, metin veya sensörler aracılığıyla elde edilen sağlık verileri veya görüntüden çıkarılan kan akışı modelleri gibi farklı türdeki kaynakları analiz eden teknolojilerle geliştirilebilir. FER'nin potansiyel kullanımları, uygulama alanlarına göre gruplar halinde aşağıdaki gibi listelenen geniş bir uygulama yelpazesine sahiptir.

Kişiselleştirilmiş hizmetlerin sağlanması:

- akıllı ortamlarda kişiselleştirilmiş mesajları görüntülemek için duyguları analiz etmek
- kişiselleştirilmiş öneriler sağlamak (örneğin müzik seçimi veya kültürel materyal seçimi vb.)
- filmlere verilen bireysel tepkileri tahmin etmek için yüz ifadelerini analiz etmek

Müşteri davranış analizi ve reklam:

- alışveriş yaparken ya mallara ya da mağaza içindeki düzenlemelerine odaklanarak müşterilerin duygularını analiz etmek
- pazarlama amacıyla tanıma ve yüz izleme sistemi kullanan bir tren istasyonunda reklam tabelaları kullanmak

Sağlık hizmeti:

- otizm veya nörodejeneratif hastalıkları tespit etmek
- yardıma ihtiyacı olan kullanıcıları belirlemek için psikotik bozuklukları veya depresyonu tahmin etmek
- intiharı önlemek
- yaşlılarda depresyonu tespit etmek
- tedavi sırasında hastaların koşullarını gözlemlemek

İş süreçleri:

- işe alım görevlilerinin karar vermesine yardımcı olmak
- bir iş görüşmesinde ilgisiz adayları belirlemek
- çalışanların ruh hallerini ve dikkatlerini izlemek

Kamu güvenliği:

- yalan dedektörleri ve akıllı sınır kontrolü
- kamusal alanların tahmine dayalı olarak taranması
- potansiyel terör tehdidini tetikleyen duyguları belirlemek
- bir suçtaki olası nedenleri belirtmek için suç mahallindeki görüntüleri analiz etmek

Suç tespiti:

- hileli sigorta taleplerini tespit etmek ve azaltmak
- dolandırıcılığı önleme stratejileri uygulamak
- mağaza hırsızlarını tespit etmek

Eğitim Faaliyetleri:

- öğrencilerin dikkatini izlemek
- eğitici bir programa kullanıcıların duygusal tepkisini tespit etmek ve öğrenme yolunu uyarlamak
- etkili öğretim sistemi tasarlamak
- çevrimiçi öğrenmeye katılımı tespit etmek

2.4. Yüz Duygu Tanımda Risk, Sorun ve Eksiklikler

Biyometrik veri ve Yapay Zeka (Artificial Intelligence – AI) teknolojilerini kullanması nedeniyle FER, yüz tanıma ve yapay zeka kullanımının bazı risklerini paylaşmaktadır. Bununla birlikte, bu teknolojinin kendine özgü riskleri de vardır. Tanımlamayı amaçlayan bir biyometri teknolojisi olması birincil hedef olarak görülmemekle birlikte, duygu yorumlama doğruluğu ve uygulaması ile ilgili riskleri de bulunmaktadır. Gelişen her teknoloji gibi duygu tanıma da mükemmel olmamakla birlikte eksiklik ve zorlukları vardır. Zorluklardan biri, veri kümelerinin insanlar tarafından etiketlenmesi ve farklı kişilerin duyguları farklı şekillerde okuyup yorum yapabilmesidir. Ayrıca, çatılmış kaşlar gibi bazı görsel ipuçları, öfke dışında başka duygular anlamına gelebilir ve aynı zamanda diğer ipuçları da bariz olmasa bile öfkeyi gösteren ince ipuçları sayılabilir (Geetha vd., 2021). Bu teknolojinin karşılaştığı bir diğer sorun da farklı renklerden insanların duygularını algılamaktır. Siyahi insanlarda öfke tespit eden modeller daha fazladır (Lawrence vd., 2015). Bu durum, eğitim setlerinin daha çeşitli olması gerektiği anlamına gelmektedir. Duygu tanıma, yaşamın birçok kurumuna ve yönüne fayda sağlamaktadır. Güvenlik ve sağlık açısından faydalı ve önemlidir. Ayrıca, belirli bir anda insan duygularını, onlara sormadan kolay ve basit bir şekilde tespit etmek için çok önemlidir. Bundan dolayı, temel olarak yüz tanıma sistemi yapay zeka teknolojileri üzerine kuruludur.

2.5. Yapay Zeka (Artificial Intelligence – AI)

Bilgisayarın bir görevi herhangi bir işlem yapmadan tamamlamasını sağlayan ve kurulum sırasında tüm temel talimatların programlandığı algoritma ve programları temsil eden yapay zeka, başlangıcından bu yana büyük bir başarı ile farklı bir çok alandaki çeşitli uygulamaların önünü açmış ve hepsinin temelini oluşturmuştur (Vaishya vd., 2020). Bunlarla birlikte de bilgisayar alanında karmaşık karar verme ve bilgisayar öğrenimi konularında da yardımcı olmuştur (Haenlein ve Kaplan, 2019). Bu, oluşturulan uygulamada önemli miktarda zaman ve enerji tasarrufu sağlamaktadır.

İnsanlar, teknolojinin benimsenmesinin başlangıcından beri otomasyona ilgilidir. Yapay zeka, makinelerin herhangi bir insan müdahalesi olmadan düşünmesini sağlar. Bilgisayar biliminin günümüzdeki en geniş alanlarından biridir.

Yapay zeka ile bağlantılı teknolojilere örnek olarak sayısız sistem bulunmaktadır (Saravanan ve Sujatha, 2018). Bunların birkaçı:

- yol tarifi alma,
- e-postaları okuma,
- müzik çalma,
- kendi kendine giden otomobiller,
- film önerileri alma,
- kendi kendine park eden araçlar,
- sohbet botlarıyla konuşmak vb.

Yapay zeka teknolojileri, hayatımızı kolaylaştırıp insanlara daha iyi hizmet verdikçe günlük hayatımızda daha fazla yere sahip olmaya devam edecektir ve farklı sistemler ile kullanılmaya devam edecektir. Yapay zeka sistemleri Şekil 1’de gösterildiği gibi üç ana başlık altında incelenmektedir:



Şekil 2.1. Yapay Zeka Sistemleri

Zayıf Yapay Zeka (Artificial Narrow Intelligence - ANI): Hedefe yönelik ve tek bir görevi gerçekleştirmek üzere programlanmış yapay zeka altyapısıdır.

Güçlü Yapay Zeka (Artificial General Intelligence - AGI): Makinelerin belirli bir durumda insanlardan ayırt edilemez bir şekilde öğrenmesini, anlamasını ve hareket etmesini sağlayan yapay zeka altyapısıdır.

Süper Yapay Zeka (Artificial Super Intelligence - ASI): Makinelerin en parlak insanları geride bırakan zeka sergileyebildiği varsayımsal bir yapay zeka sistematiğidir.

Yapay zeka'nın bu farklı versiyonları hedeflere ulaşmaya çalışmaktadır ve bunları daha iyi anlamak için karşılaştırmak mantıklı olacaktır. Güçlü yapay zeka, olağanüstü bilinç de dahil olmak üzere sahip olduğumuz tüm zihinsel güçlere sahip makinelerden oluşan yapay insanlar yaratmayı amaçlamaktadır (Grudin ve Jacques, 2019). Zayıf yapay zeka ise, insanların zihinlerine yapı olarak en yakın görünen yapay zekâ makineleri oluşturmaktır (Chalmers ve Chalmers, 2004).

2.5.1. Zayıf Yapay Zeka (ANI)

Bu yapay zeka türü, spesifik görevleri yerine getirmede uygundur, ancak kendi imkânlarının dışında herhangi bir alanda yeterli olmayacaktır. Bu sisteme örnek olarak, en iyi yapay zeka örneklerinden biri de satrançta dünya satranç şampiyonu olan (1996 - Garry Kasparov) bir insanı yenen ilk bilgisayar olan Deep Blue'dur (Lihui Wang, 2019). Deep Blue yapay zekâ örneği, nerdeyse 200 milyon satranç pozisyonunu bir saniyede üretip değerlendirebilmektedir. Deep Blue'nun tam anlamıyla yapay zeka olmadığı yada zayıf yapay zekanın ilk örneklerinden biri olduğu iddia edilmiştir. Buna ek olarak, yapay zekanın oyunlarda insanlardan üstün gelmesinin farklı bir örneği de AlphaGo oyunudur. Bu oyunda, en zor derecede nasıl oynaması gerektiğini öğrenmiştir ve karşısına çıkabilecek tüm pozisyonları hesaplamakla kalmayıp oyunu kazanmıştır (Bory, 2019).

Günümüzde, yapay dar zeka bilim, iş ve sağlık alanında çok ünlüdür. Buna örnek olarak, DOMO şirketi 2017 yılında Mr. Roboto isimli bir robot üretmiştir (Chandrasekhar, 2020). Sistem, iş geliştirmeye yönelik öneriler ve öngörüler üzerine güçlü analitik yapılar bulundurmaktadır. Google ve Amazon gibi firmalarda da benzer programlar mevcuttur ve gelişimine önemli miktarlarda yatırım yapmaktadırlar.

2.5.2. Güçlü Yapay Zeka (AGI)

Bu tür, makinelerin insan benzeri hale geleceği noktadır. Güçlü yapay zekalı sistemler kararlarını kendileri vermektedirler ve insan müdahalesi olmadan öğrenmektedirler. Mantıksal görevlerini yerine getirirler fakat duyguları tam olarak oluşmamaktadır (Haenlein ve Kaplan, 2019).

2.5.3. Süper Yapay Zeka (ASI)

Bu yapay zeka türü ise, makinelerin insanlardan çok ileride olduğu durumu tanımlamaktadır. Süper yapay zeka sistemlerinde akıllı, bilge, yaratıcı, sosyal yönden çok gelişmiş makinelerden bahsedilir. İnsanların hayatlarını daha iyi hale getirmeyi ya da yok etmeyi görev edinebilirler.

Üzerinde durulan görevler şu şekildedir (Vaishya vd., 2020):

2.5.3.1. Makine Muhakemesi (Machine Reasoning - MR): Makine Muhakemesi (MR), veri tabanı gibi bilgilerden meydana gelir. Tümdengelim ve tümevarım tekniklerini kullanıp elde edilen verilere göre bazı değerleri hesaplayarak analiz edebilecek formülleri oluşturmaktadır. Planlama, veri temsili, optimizasyon ve arama içerebilmektedirler (Ying, 2019).

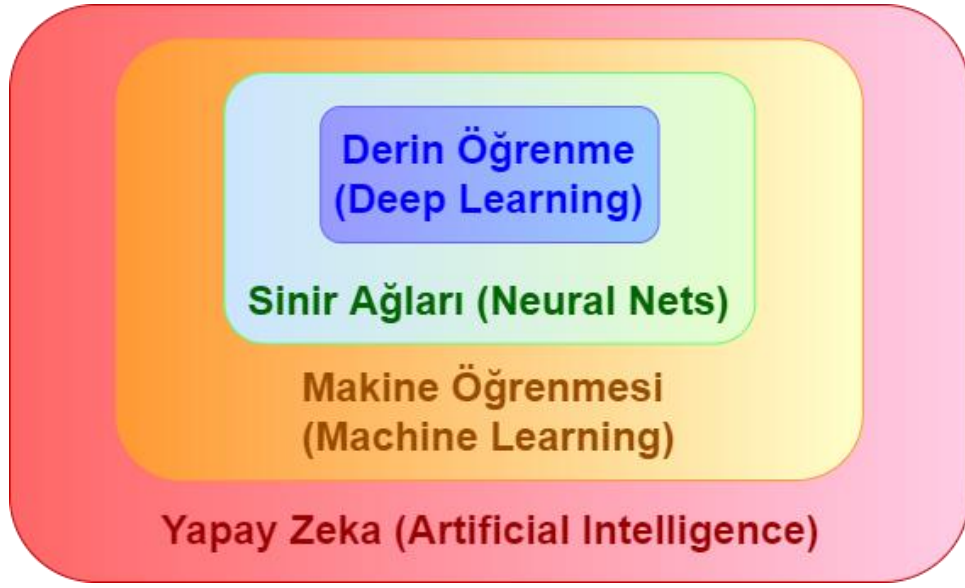
2.5.3.2. Robotik: Robotik, günümüzde insanların işlerinde kullandıkları makinelerden (örneğin, oda temizleyici veya düzenleyici) akıllı robotlar oluşturmaya, geliştirmeye ve kontrol etmeye odaklanmaktadır (Alzubaidi vd., 2021).

2.5.3.3. Makine Öğrenimi (Machine Learning - ML): Verilen görevi tamamlamak amacıyla bilgisayar modellerinin ve algoritmalarının incelenip kullanılmasıdır. Sistemden gerçek verilerin alınarak makinelerin geleceğe yönelik hesaplamalar yapabilmesi için formüller oluşturup, yeni bir durum verisi için tahminde bulunmalarını sağlamak amacıyla kurgulanmaktadır (Haenlein ve Kaplan, 2019).

2.5.3.4. Derin Öğrenme (Deep Learning - DL): Makine öğreniminin, insan beyninin yapısından ilham alarak oluşturulmuş halidir. Çok sayıda sinir ağının kullanımıyla uygulanır. Şekil 2’de gösterildiği gibi Derin Öğrenme, yapay zekanın kapsamakta olduğu makine öğreniminin özelleştirilmiş bir alt kümesidir. Yapay zeka alanı, makine öğrenimini ve makine öğrenimi de derin öğrenme alanını kapsamaktadır (Kim ve Lee, 2016).

2.6. Makine Öğrenimi (ML) Sistemlerinin Yapısı

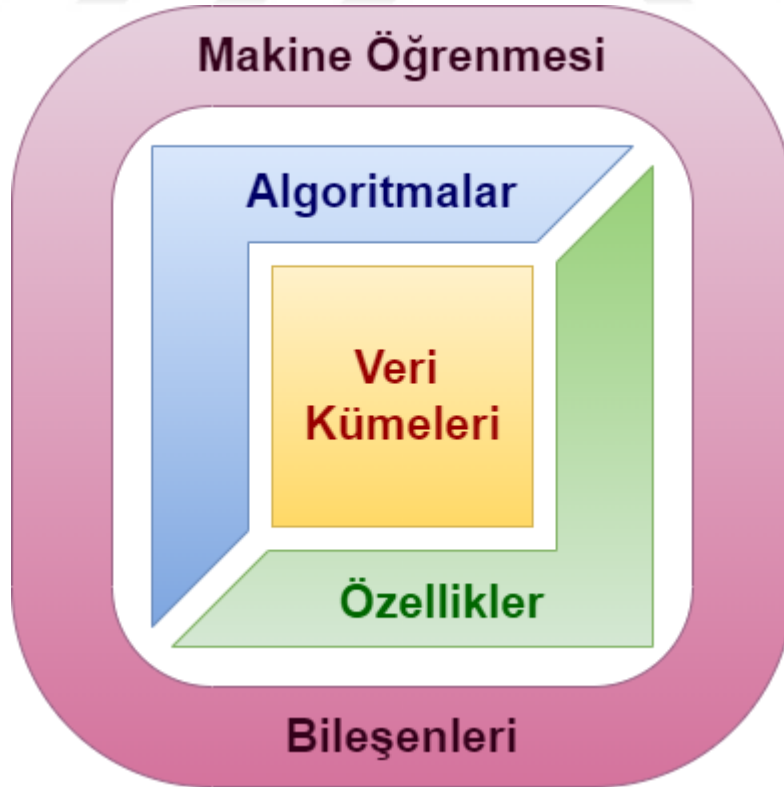
Makine Öğrenimi (ML), akıllı sistemler oluşturmak için istatistiksel öğrenme algoritmalarını kullanan bir yapay zeka alt kümesidir. Makine öğrenimi sistemleri açıkça programlanmadan otomatik olarak öğrenebilir ve geliştirilebilir.



Şekil 2.2. Yapay Zeka Sistemlerinin İlişkisel Yapıları

Makine öğrenimi, programlanmaya ihtiyaç duymadan öğretmeye dayalı ayrıntılı bir yapay zeka yapısıdır. Makine öğrenimi mantıksal açıdan eldeki verilerin incelenmesiyle öğrenen ve tahminler yapan algoritmalar kullanılmaktadır.

Makineyi eğitiminde Şekil 3'te gösterilmekte olan 3 bileşene ihtiyaç vardır:



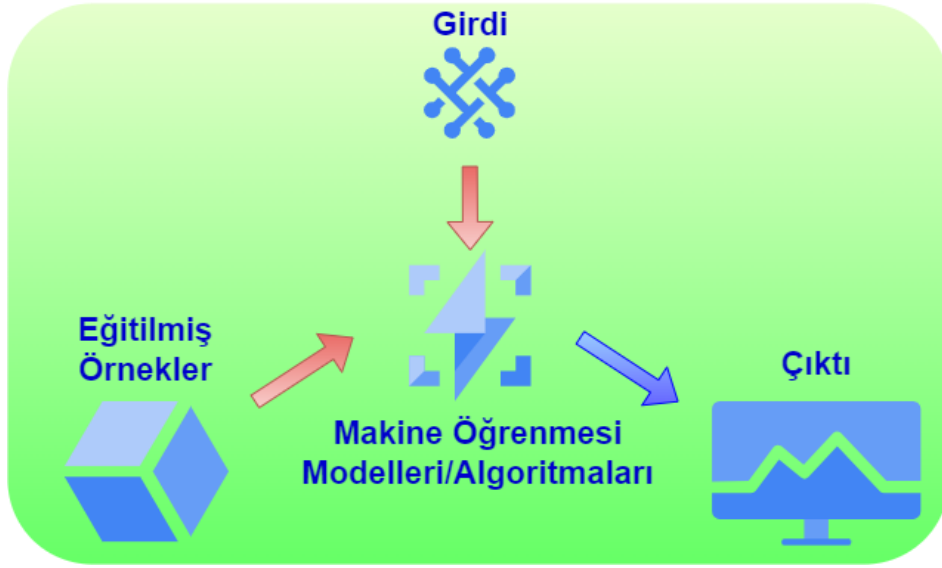
Şekil 2.3. Makine Öğrenmesi Bileşenleri

2.6.1. Veri Kümeleri: Makine öğrenimi sistemlerinin ilk olarak gereksinim duydukları yapılar veri kümeleri olarak adlandırılan belirli özellikleri barındıran örnek değerleri kapsamaktadır. Gerçek olarak oluşturulmuş veya ölçümlenmiş bu örnekler, sayısal, resim, metin veya herhangi başka türden oluşan veriyi içerebilir. Tüm bu veri kümelerini iyi ve anlamlı bir şekilde oluşturmak için araştırmacıların genel olarak çok zaman ve çaba göstermeleri gerekmektedir.

2.6.2. Özellikler: Veri kümelerinde bulunan verilerin anlamlı halde oluşturulacak görevlerde çözüm üretmek için anahtar olarak çalışan yapılardır. Veriler üzerinde dikkat edilmesi gereken anlamsal yapıları makineye belirtmektedirler. Bu sebeple, özellikleri oluşturmak veya seçmek için farklı yöntemler bulunmaktadır ve oluşturulacak yapının talebine göre değişiklik gösterebilmektedir. Özellikler, birkaç farklı verinin bir arada oluşturulmuş kombinasyonlarına göre de kullanılmaktadır. Örneğin, bir apartman dairesinin boy uzunluğu ve genişliğine bakılarak fiyat belirlemeyi iterasyonel bir yapıyla çözmek zor olacaktır. Bu nedenle, dairenin uzunluğu ve genişliğine bağlı olarak oluşturulacak yeni oransal özellik ile fiyat belirlemesi yapılabilir ve dairenin birim maliyetinin ne olabileceği daha kolay hesaplanabilir.

2.6.3. Algoritma: Çözüm bulunması istenilen problemleri, elde edilen veri kümeleri ve özelliklere bağlı olarak farklı yöntemler veya algoritmalar kullanarak çözebiliriz (Şekil 4). Farklı algoritmalar farklı modeller üzerinde çözüm bulmaya çalışırken, alınan sonuçların doğruluğu veya çözüm hızlarında farklılıklar olabilmektedir. Ayrıca, zor problemlerin çözümlerinde en iyi performansa sahip olma amacıyla çeşitli algoritmalar veya farklı algoritmaların belirli bölümleri birleştirilebilir.

Makine öğrenimi mimarisi ile entegre olmuş herhangi bir yazılım, manuel olarak kodlanmış yazılımların problemleri çözme kabiliyetlerine göre daha bağımsız kararlar alabilmektedir ve sistemin belirli kalıpları tanıyarak güçlü tahminler yapması sağlanmaktadır. Makine öğrenimi sistemini destekleyen veri kümesinin kalitesi yüksek ve özelliklerin tutarlı olması, çözüm bulunması gereken problemleri insanlardan daha hızlı ve efektif olarak yanıtlamayı sağlayacaktır (Carleo vd., 2019).



Şekil 2.4. Makine Öğrenmesi Yöntemlerinin Çalışma Mimarisi

Makine öğrenimi/öğrenmesi mimarileri, öğrenim yaptıkları modellere bağlı olarak Şekil 5’te gösterildiği gibi denetimli, denetimsiz, yarı denetimli ve pekiştirmeli öğrenme olmak üzere 4 kategoriye ayrılır.

2.6.4. Denetimli Öğrenme

Denetimli sözcüğü kelime anlamı olarak yardım ile birlikte işlemlerin yürütmesi anlamına gelmektedir. Denetimli öğrenme ise, bir eğitmenin öğrenme süreci boyunca sistemsal olarak yardımcı olduğunu göstermektedir ve bu kapsamda eğitim setlerinde verilerin etiketli tanımlanmasının gerektiğini bizlere belirtmektedir (Saravanan ve Sujatha, 2018). Örnek olarak, bir bilgisayar sisteminden farklı türdeki verilerin kategorisel olarak ayrılması talep edildiğinde, sisteme verilerin ilk olarak hangi kategoriye ait olduklarının belirtilmesi gerekmektedir. Bu işlemten sonra, oluşturulan bilgisayar yazılımının öğrendiği operasyonun doğruluğunu test etmek amacıyla oluşturulan bir doğrulama kümesinde çalıştırılması beklenmektedir. Eğer yazılımın çıkarımlarında eksiklikler mevcut ise yazılım içerisinde kullanılan yöntemler tekrar gözden geçirilerek düzenlenmelidir. Yöntemlerin veya modellerin eğitim süreçleri istenen doğruluk düzeyine erişene kadar tekrarlı olarak sürdürülmelidir ve bu tür öğrenme sistemleri genel olarak sınıflandırma ve regresyon gerektiren yapılarda kullanılmaktadır. Kullanılan modellere veya yöntemlere örnek olarak verilecek bilenen yapılar Naive Bayes, K-En Yakın Komşular, Karar Ağacı, Destek Vektör Makinesi, Lojistik Regresyon, Polinom ve Doğrusal Regresyonlar, vb. olarak verilebilir ve bu yapılar genel olarak bilgisayarla görme, dil algılama, istenmeyen/gereksiz posta

filtreleme, arama ve sınıflandırma alanlarında kullanılmaktadır.



Şekil 2.5. Makine Öğrenmesi Kategorileri

2.6.5. Denetimsiz Öğrenme

Denetimli öğrenme modelinden farklı olarak denetimsiz öğrenme yapısında bilgisayar yazılımı oluşturulacak veri kümeleri üzerinde bağımsız olarak arama yaparak herhangi bir etiketlenmiş özellik bulundurulmadan sistemin belirli yöntemleri kullanarak veri kümelerini kategorisel olarak ayırt etmesini sağlamaktadır (Caron vd., 2020). Bu sayede, sezgisel olarak verilerin analitik anlamlarını belirlemek için denetimsiz öğrenme kullanılabilir ve verileri genellikle benzerliklerine göre farklı gruplara ayırt edebilmek için uygun bir yaklaşımdır. Sezgisel yaklaşımların faydası, denetimsiz öğrenme ile birlikte büyük veriyi işlerken insanların gözünden kaçabilecek farklı kalıpları algılayarak belirleyebilmesidir (Zhuang vd., 2019). Örnek olarak, oluşturulacak sistemler, firmaların satışlarını ve ürün indirimlerini önceden tahmin etmek, müşterilerin geçmiş aramalarını inceleyerek gelecek tercihlerini analiz etmek ve sahte işlemleri veri kümelerinin içerisinde algılayabilmek gibi farklı işlemlerde kullanılabilmektedir. Ayrıca, yazılımcılar tam olarak hangi kategoriye veya veriyi bulmaları gerektiğini bilemiyorsa sistemler belirli kalıpları kendi analiz yeteneği ile oluşturarak yazılımcılara ön değerlendirme altyapısı oluşturabilmektedir. Denetimsiz öğrenme modelinde genel olarak kullanılan yöntemler K-Ortalama Kümelemesi, Ortalama

Kaydırma, Tekil Değer Ayırışımı, Temel Bileşen Analizi, Gizli Semantik Analiz, vb. olarak göz önünde bulundurulmaktadır ve anomali tespiti, öneri sistemleri, risk yönetimi veya sahte görüntü analizi gibi alanlarda sıklıkla kullanılmaktadır.

2.6.6. Yarı Denetimli Öğrenme

Bu öğrenme, denetimsiz ve denetimli öğrenme modellerinin birleşimi halinde etiketlenmiş ve etiketlenmemiş verilerin oluşturduğu bir karışık örneklem havuzudur (Ouali vd., 2020). Bu havuzda, genellikle etiketlenmiş verilerin miktarı etiketlenmemiş verilerin miktarından daha az olmaktadır. Yazılımcılar, bu yapıdaki öğrenme modelinde istenen bir tahmin sonucunu aklında bulundurur, ancak model, verileri yapılandırmak ve tahminleri kendisi yapmak için kalıplar bulmalıdır. Üretken modellerle yarı denetimli öğrenme, denetimli öğrenmenin bir uzantısı veya denetimsiz öğrenmenin bir uzantısı olarak görülebilir. Bu nedenle, denetimli ve denetimsiz öğrenme modellerindeki aynı tür algoritmalar kullanılabilir. Bu yöntemlerden farklı olarak da Transdüktif Destek Vektör Makinesi (TSVM) örnek olarak verilebilir.

2.6.7. Pekiştirmeli Öğrenme (Reinforcement Learning – RL)

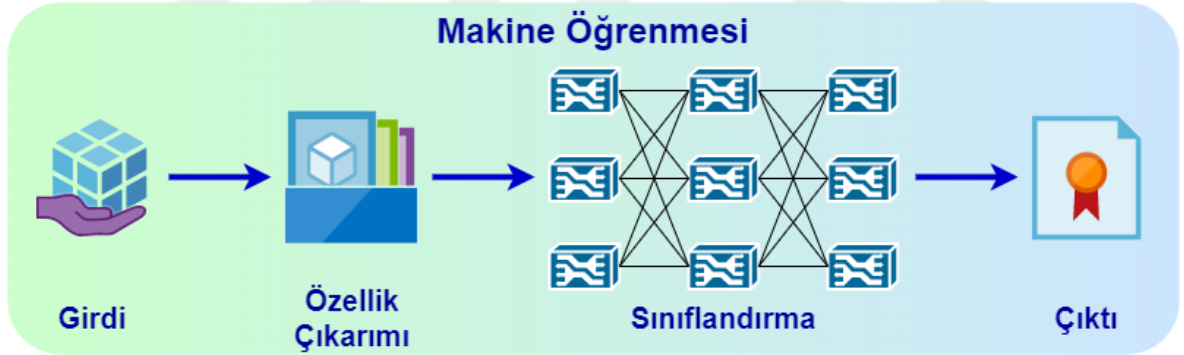
Diğer öğrenme modellerinden farklı olarak pekiştirmeli öğrenme modelinde deneme yoluyla sistemin öğrenmesi mantığıyla çalışmaktadır. Denetimli öğrenmeden farklı olarak, pekiştirmeli öğrenme modelinde insanların sürekli denetime ihtiyacı yoktur. Pekiştirmeli öğrenme modelinde, yapılan eylemlere karşı olarak pozitif veya negatif pekiştirme sinyalleri yardımıyla etkili olarak öğrenme işlevi yapılmaktadır (Cobbe vd., 2019; Lu Wang vd., 2018). Örnek olarak, kötü bir durum karşısında insanların çözüm olarak pozitif düşünceleri verilebilir ya da iyi yolu tercih etmeleri sağlanmalıdır. Eğer bir çocuk sıcak bir tavaya temas ettiğinde acı hissettikten sonra artık sıcak tavaya dokunmamayı öğrenmiş olacaktır.

Pekiştirmeli öğrenme modelinde statik veri kümeleri üzerinde çalışmamaya izin vermesi en olumlu yaklaşımlardan biridir. Bunun yerine dinamik olarak hatalı verilere de sahip çevresel etkilerden çıkarımlar yaparak öğrenme işlemini gerçekleştirir. Pekiştirmeli öğrenme için bilgisayar oyunları veri bakımından zengin ve ideal bir ortam oluşturmalarından dolayı araştırmalarda faydalı olmaktadır. Ödülle dayalı faaliyetler modellerin eğitimleri için puanlama sistemi ideal öğrenme sinyallerini oluşturmaktadır (Cobbe vd., 2019). Pekiştirmeli öğrenme modellerinde Q-Öğrenme, Genetik Algoritmalar, SARSA, vb gibi algoritmalar sürücüsüz arabalar, oyunlar, robotlar veya kaynak yönetimi alanlarında kullanılmaktadır.

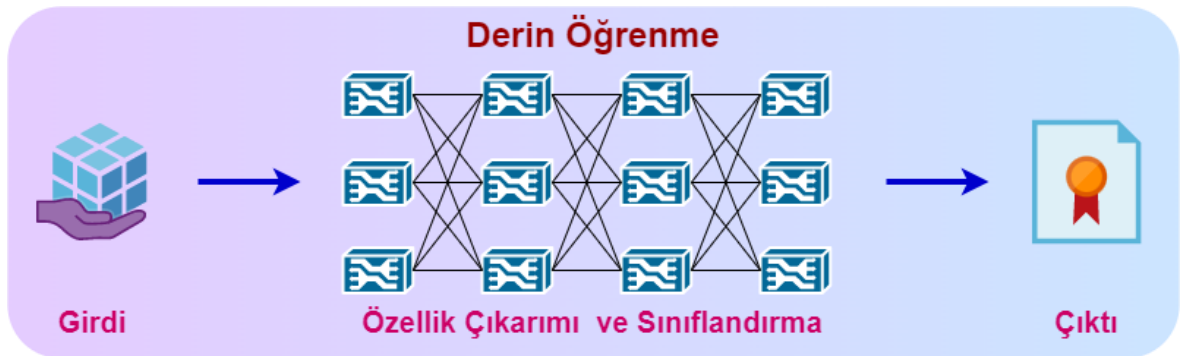
2.7. Derin Öğrenme (Deep Learning – DL)

Diğer makine öğrenimi yöntemlerinden farklı olarak insan beyninin çalışma yapısından yararlanılarak oluşturulan öğrenme modeline Derin Öğrenme (DL) denilmektedir. Şekil 6’da bahsedildiği üzere derin öğrenme, makine öğreniminin aksine veri kümelerinden girdileri doğrusal olmayan dönüşümlerin soyutsal yapıda farklı seviyelerde artış gösteren karmaşık çok katmanlı sinir ağıları modellerini kullanmaktadır (Kamilaris ve Prenafeta-Boldú, 2018). Girdi olarak kullanılan veri kümelerinden kademeli şekilde daha verimli ve seviye olarak yüksek özellikler elde etmek için derin öğrenme birden fazla katman kullanmaktadır. Bu duruma örnek olarak, derin öğrenmenin kullanıldığı alanlardan biri olan görüntü işleme operasyonlarında alt katmanların görevleri görüntü kenar noktalarını tanımlamak iken, daha yüksek mertebeden katmanlar harfler, rakamlar veya yüz yapıları gibi farklı ve anlamlı insan ile alakalı kavramları tanımlayabilir (Johnson vd., 2020).

Derin öğrenme, girdiden kademeli şekilde seviyesi daha yüksek özellikler elde edebilmek amacıyla çoklu katman kullanmaktadır. Buna örnek olarak, alt katmanlar kenarları belirlerken, yüksek katmanlar yüzler, harfler, rakamlar gibi daha spesifik kavramları algılar.



(a) Makine Öğrenmesi Katmanları

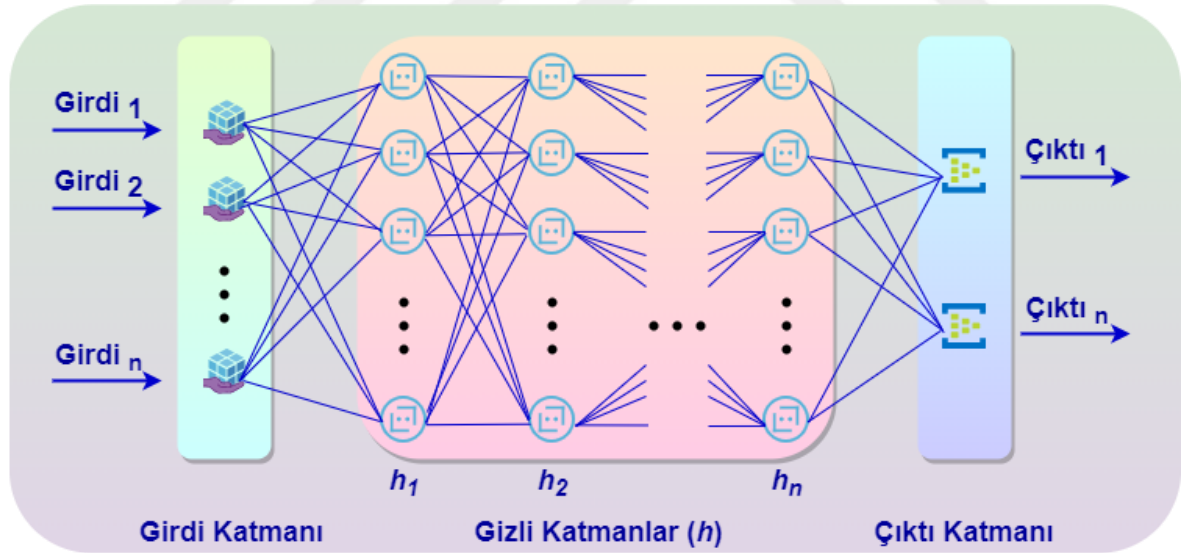


(b) Derin Öğrenme Katmanları

Şekil 2.6. Makine Öğrenimi ve Derin Öğrenme Mimarisi

Şekil 6’da gösterilen çoklu mimaride, bir sinir ağında işlenecek olan bir bilgi karmaşık bağlantılı kanallar üzerinden bir katmandan diğer katmana Şekil 7’de gösterildiği gibi aktarılmaktadır. Bu bağlantı kanalları kendilerine ait bir değer barındırdıkları için ağırlıklı kanallar olarak adlandırılmaktadır. Oluşturulan bu sinir ağı modellerinde, nöron olarak adlandırılan her bir girdi bağlantısı sapma adı verilen benzersiz bir değeri tutmaktadır. Bu sapma değeri nöronların aldıkları girdilerin sahip olduğu değerlerin ağırlıklı toplamına eklenir ve aktivasyon fonksiyonu uygulanarak işleme alınır. Yapılan işlemler sonucunda nöronun aktive olup olmayacağı belirlenmektedir. Eğer nöron aktive olmuş ise, bilgiyi bir sonraki katmana iletebilmektedir ve son katmana kadar devam eder. En son olarak da çıktı katmanında da oluşturulan model yazılımı için çıktılar üretilir.

Veri bilimlerinde Şekil 7’de anlatılan sinir ağı mimarisini eğitebilmek için oldukça büyük miktarda eğitim yapılabilir veriye gereksinim duyulmaktadır. Büyük miktarda veri kullanımı sistemin çözümü bulma aşamasında gerekli olan çok sayıda farklı parametreye ihtiyaç duyması ve kullanmasıdır. Derin öğrenme model yapısı kullanılarak oluşturulan Derin Sinir Ağları (Deep Neural Network – DNN) yöntemleri farklı uygulama alanlarına Şekil 8’de gösterildiği gibi sahiptir.



Şekil 2.7. Sinir Ağları Katman Mimarisi

DNN, yapısı gereği nesnelerin katmanlı bir bileşimi olarak karmaşık doğrusal olmayan ilişkilerin modellerini oluşturabilir. Alt katmanlarda oluşturulan özelliklerin birleşimi ekstra katman kullanımı ile daha zayıf bir ağ modelindeki karmaşık verileri yüksek performans göstererek modellemektedir (Chen vd., 2020). Örnek olarak, seyrek çok değişkenli polinomların DNN'lerle tahmin edilmesinin sık ağlara göre katlanarak daha kolay olduğu

kanıtlanmıştır ve farklı uygulama alanlarında yukarıdaki şekilde belirtildiği gibi kullanılabilir.



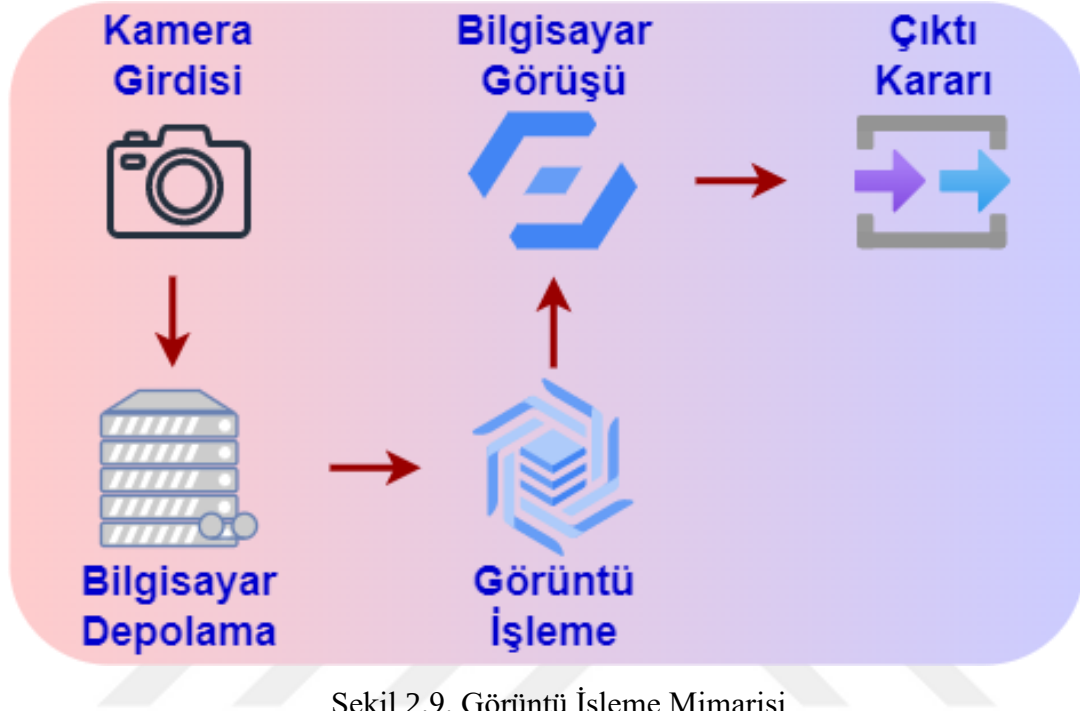
Şekil 2.8. Derin Öğrenme Modellerinin Uygulama Alanları

Derin öğrenmenin yaygın olarak kullanıldığı pratik uygulamalara örnek olarak Google Asistan veya Amazon Alexa benzeri insanların konuşmalarını tanıyan sistemleri verebiliriz. Burada algılanan ses dalgalarının özelliği olarak bilinen frekansların durumsal olarak belirlenmesi ile oluşan spektogramları tanımlanır ve kullanılan Uzun-Kısa Dönemli Bellek (Long Short Term Memory – LSTM) benzeri yöntemler ile öğrendiği spektogram desteği yardımıyla kelimeleri eşleyebilir, tanıyabilir veya işlenebilir.

Bunlara ek olarak derin öğrenme mimarileri bu çalışmada da yapılmakta olan dijital görüntü işleme, bilgisayar ile görme, doğal dil işleme, konuşma tanıma ve biyoinformatik gibi farklı alanlara da uygulanmaktadır. Şekil 9’da gösterilen derin öğrenme modelleri ile destekli olarak tasarlanmış bilgisayarlı görme veya görüntü işleme mimarisinde girdi olarak kameralardan alınan veriler ve bu verilerin işleme adımları kısaca özetlenmektedir. Ayrıca, derin öğrenme ile gelişen dijital görüntü işleme altyapıları gün geçtikçe daha popüler bir hal almaktadır.

Bu çalışmanın ana konusu olarak anlatılmak istenilen dijital görüntü işleme derin öğrenme modellerinin son zamanlarda güçlü sistemlerin kurulması ile gelişmeye devam etmektedir. Bu sektörde parlayan en sıcak konulardan biri olan Bilgisayarlı Görme, bilgisayarların

görüntüleri ve videoları kendi başlarına anlama yeteneği olarak tanımlanabilir. Kendi kendini süren arabalar, biyometri ve bu çalışmanın ana temasını oluşturan yüz tanımanın modellerinin tümü, çalışmak için bilgisayar vizyonuna dayanmaktadır ve bu yüzden bilgisayarlı görmenin merkezinde dijital görüntü işleme vardır.



Şekil 2.9. Görüntü İşleme Mimarisi

İnsanların yüzlerindeki duyguların anlık olarak analizi ve yaş gruplarına veya cinsiyetlerine göre analiz edilmesiyle birlikte anlık stres durumlarının derin öğrenme modelleriyle incelenmesi ana tema olarak çalışılmaktadır. Yüz duygu analizinde insanların en temel olarak belirledikleri mutluluk, üzüntü, vb. duygularının anlık görüntülerden alınarak analiz edilip raporlanması sağlanmaktadır. Bu analizlerin yapılmasında kullanılacak modeller dijital görüntü işleme alanının kapsadığı derin öğrenme yöntemleriyle birlikte yüksek doğruluk/performans sağlayarak doğru tahminleri oluşturmayı amaçlamaktadır.

Bu çalışmada sonraki bölümlerde sırasıyla çalışmada kullanılan görüntü işleme modelleri, bu modellerin yazılım olarak oluşturulması, oluşturulan yazılımlar ile model eğitimlerindeki deneysel sonuçların analizi yapılacaktır. Son bölümde de çalışmanın sonuçlandırılması ile birlikte gelecek geliştirme imkânlarından da bahsedilecektir.

3. YÖNTEM

3.1. Anlık Duygu Analizi, Stres, Cinsiyet ve Yaş Belirleme Mimarisi

Günümüzde makinelerin insan duygularını tanıyabilmesi ve sentezleyebilmesini sağlamak amacıyla duygusal hesaplama yapıları kullanılmaktadır (Bashar, 2019). İnsanlar arasındaki iletişimin temellerinden biri olarak bilinen bireylerin duygularının değişimi, insanların eylemlerdeki motivasyonunu veya duygusal durumlarının anlamlarını tanımlayabilir (Zhang vd., 2020). İnsan Bilgisayar Etkileşimi (Human Computer Interaction – HCI), bireylerin duygusal durumlarını göz ardı eder ve etkileşim sürecinde bilginin büyük bir kısmını kaçırabilir. Bu sebeple duyguya duyarlı insan bilgisayar etkileşim sistemleri daha güçlü bir yapıya sahip olurlar ve eğlence, ticaret, psikoloji, vb. uygulama alanlarında artan gereksinimleri ile duygusal hesaplama ile ilgili ilgi artmaktadır.

Görsel sinyaller, duygusal hesaplamanın en önemli parçalarıdır. Etkileyici yüzümüz, sözlerimizin içerdiği ötesinde duygu durumumuzun yarısından fazlasını içerir. Bu nedenle, çok sayıda araştırma, duygusal hesaplamanın görsel özelliklerine odaklanmaktadır. Yüz ifadelerindeki duygusal ifadelerin hesaplanmasında kullanılan mevcut algoritmalar ve sistemler esas olarak orijinal duyguların kasıtlı olarak görüntülenmesi ve abartılan ifadelerine dayanır (Domínguez-Jiménez vd., 2020). Kasıtlı davranışlar ile doğal olarak ortaya çıkan davranışlar arasındaki farkı tüm yönleriyle (görselleştirme, ses ve zaman) görmezden gelir. Bu sorunu çözmek için son zamanlarda kendiliğinden oluşan duygusal davranışları işlemek için birçok algoritma önerilmiştir (Gao vd., 2020). Ayrıca, özellik düzeyinde füzyon, karar düzeyinde füzyon ve multimodal füzyon gibi insan duygularını analiz etmek için füzyon yöntemlerine yönelik artan araştırmalar önerilmektedir. Ayrıca yüz ifadeleri, baş hareketleri ve vücut hareketleri ile ilgili bilgilerin birleştirilmesi için füzyon yöntemleri de önerilmiştir.

Bu çalışmada duygusal ifadelerin analiz edilmesinin cinsiyet, yaş ve stres hesaplamaları ile hesaplanması için farklı öğrenme yöntemlerine odaklanmaktadır. Bildiğimiz gibi, farklı öğrenme yöntemleri farklı tahmin problemlerine uygundur. Bu çalışmada, genel durumlarda en iyi yöntemlerden birkaç öğrenme yöntemini uyguladık ve bu tahmin sistemlerinin öznetelik kümesinin özelliklerine dayanarak zorluklara en uygun tahmin sistemini basit birleştirme yapıları ile uygulamaktır.

3.2. Dijital Görüntü İşleme (Digital Image Processing - DIP) Mimarisi

Dijital Görüntü İşleme (DIP), sayısal görüntünün sayısal bir bilgisayar aracılığıyla işlenmesi anlamına gelmektedir. Bazı yararlı bilgileri çıkarmak için ve gelişmiş görüntü elde etmek için bilgisayar algoritmalarının bir kullanımı olduğu da söylenebilir. Nihai analizden önce ham olarak elde edilen bir görüntüde istenen sinyalleri izole etmeyi ve/veya vurgulamayı amaçlayan hayati bir ara adımdır (Aaron ve Chew, 2021). Görüntü işlemeden elde edilen sonuçlar, görüntüyü büyük ölçüde etkileyebilir.

Dijital görüntü işleme genel olarak üç adımdan oluşmaktadır:

- Elde edilen görüntülerin sisteme aktarılması,
- İçeri aktarılan görüntülerin analiz ve manipüle edilmesi,
- Elde edilecek görüntü sonucunun analizi veya değiştirilebileceği çıktıyı içeren rapor oluşturulması.

Aynı zamanda, dijital görüntü işleme mimarisi sinyal işleme sistemlerinin özellikle görüntülere odaklanan bir alt alanıdır. Genel olarak elde edilen görüntü ile ilgili hesaplamalar yapabilen bir yazılım geliştirmeyi amaçlamaktadır. Farklı yöntemler veya algoritmalar ile yazılım girdi olarak aldığı dijital görüntüyü işleyerek anlamlı bir çıktı elde eder. Günümüzde genel olarak kullanımı yaygın olan yazılım Adobe Photoshop uygulaması manuel olarak görüntü işlemede kullanılmaktadır.

Şekil 10'da belirtildiği üzere, bir kamera ile çekilmiş görüntü dijital bir yazılıma gereksinim duyulmayan detaylarının kaldırılması için gönderilmektedir. Yapılacak olan odaklanma işleminin sonucunda görüntüde bulunan su damlasının yakınlaştırılmış hali oluşturulmaktadır.



Şekil 3.1. Dijital Görüntü İşleme Adımları

Dijital görüntü işlemede en önemli eleman tartışmasız görüntüdür. Görüntüler iki boyutlu bir fonksiyon da kullanılan uzamsal koordinat düzlemi belirtilerek $F(x,y)$ olarak tanımlanarak herhangi bir (x,y) koordinat çiftinin belirlediği noktalar arasındaki yoğunluğunu belirler. Eğer F fonksiyonundaki x ve y değerleri ile birlikte genlik değeri de sonlu olursa, bu görüntü yapısını dijital görüntü olarak tanımlarız. Dijital görüntülerin her

biri belirlenen bölgelerde bilinen sonlu bir değer sahip olarak yine sonlu sayıda resim, görüntü öğeleri ve piksellerden oluşmaktadır. Ayrıca, genel olarak dijital görüntünün öğelerini belirlemek için piksel kullanılmaktadır. Piksel, görüntü üzerinde belirli bir gölge, opaklık veya renk alan bir noktadır. Genellikle aşağıdakilerden birinde temsil edilir:

Gri Tonlamalı- Piksel, 0 ile 255 arasında (0 tamamen siyah ve 255 tamamen beyaz) bir tamsayıdır.

RGB- Bir piksel, 0 ile 255 arasında 3 tam sayıdan oluşur (tam sayılar kırmızı, yeşil ve mavinin yoğunluğunu temsil eder).

RGBA- Görüntünün opaklığını temsil eden alfa alanı eklenmiş bir RGB uzantısıdır.

Görüntü işleme, bir görüntünün her pikselinde gerçekleştirilen sabit işlem dizilerini gerektirir. Görüntü işlemcisi, görüntü üzerinde ilk işlem sırasını piksel piksel gerçekleştirir. Bu tamamen yapıldığında, ikinci işlemi gerçekleştirmeye başlayacak ve bu şekilde devam edecektir. Bu işlemlerin çıktı değeri görüntünün herhangi bir pikselinde hesaplanabilir.

Dijital görüntü işlemenin temel adımlarını aşağıdaki gibi sıralayabiliriz (Alzubaidi vd., 2021):

Görüntü edinme: Görüntünün elde edilmesi, dijital görüntü işlemenin başlangıç adımıdır. Bu adım, görüntü işlemede ön işleme olarak da bilinir. Görüntünün bir kaynaktan, genellikle donanım tabanlı bir kaynaktan alınmasını içerir.

Görüntü Geliştirme: Görüntü iyileştirme, gizlenmiş bir görüntüdeki belirli ilgi çekici özellikleri ortaya çıkarma ve vurgulama işlemidir. Bu, parlaklığı, kontrastı vb. değiştirmeyi içerebilir.

Görüntü Restorasyonu: Görüntü restorasyonu, bir görüntünün görünümünü iyileştirme işlemidir. Ancak, görüntü iyileştirmeden farklı olarak, görüntü restorasyonu belirli matematiksel veya olasılıksal modeller kullanılarak yapılır.

Renkli Görüntü İşleme: Renkli görüntü işleme, dijital bir alanda bir dizi renk modelleme tekniğini içerir. Bu adım, dijital görüntülerin internet üzerinden önemli ölçüde kullanılması nedeniyle önem kazanmıştır.

Dalgacıklar ve Çoklu Çözünürlüklü İşleme: Dalgacıklar, görüntüleri çeşitli çözünürlük derecelerinde temsil etmek için kullanılır. Görüntüler, veri sıkıştırma ve piramidal gösterim için dalgacıklara veya daha küçük bölgelere bölünür.

Sıkıştırma: Sıkıştırma, bir görüntüyü kaydetmek için gereken depolamayı veya onu iletmek

için gereken bant genişliğini azaltmak için kullanılan bir işlemdir. Bu, özellikle görüntü internette kullanım için olduğunda yapılır.

Morfolojik İşleme: Morfolojik işleme, görüntüleri şekillerine göre dönüştürmek için bir dizi işleme işlemidir.

Segmentasyon: Segmentasyon, görüntü işlemenin en zor adımlarından biridir. Bir görüntünün kendisini oluşturan parçalara veya nesnelere bölünmesini içerir.

Temsil ve Açıklama: Segmentasyon işleminde bir görüntü bölgelere bölündükten sonra, her bölge daha sonraki bilgisayar işlemleri için uygun bir biçimde temsil edilir ve tanımlanır. Temsil, görüntünün özellikleri ve bölgesel özellikleriyle ilgilenir. Açıklama, bir nesne sınıfını diğerinden ayırt etmeye yardımcı olan nicel bilgilerin çıkarılmasıyla ilgilenir.

Tanıma: Tanıma, açıklamasına dayalı olarak bir nesneye bir etiket atar.

Beş ana dijital görüntü işleme türü bulunmaktadır:

1. Göselleştirme- Resimde görünmeyen nesneleri bulma
2. Tanıma- Görüntüdeki nesneleri ayırt etme veya algılama
3. Keskinleştirme ve geri yükleme- Orijinal görüntüden gelişmiş bir görüntü oluşturma
4. Desen tanıma- Görüntüdeki nesnelerin etrafındaki çeşitli desenleri ölçme
5. Alma- Orijinal görüntüye benzeyen geniş bir dijital görüntü veri tabanındaki görüntülere göz atma ve bu görüntüleri arama

Günlük hayatta dijital görüntü işleme mimarisi birçok alanda kullanılmaktadır.

3.3. Dijital Görüntü İşleme Uygulamaları

3.3.1. Tıbbi Görüntü Alma

Görüntü işleme, tıbbi araştırmalarda yaygın olarak kullanılmaktadır ve daha verimli ve doğru tedavi planlarını mümkün kılmıştır. Örneğin, göğüs taramalarında gelişmiş bir nodül tespit algoritması kullanılarak göğüs kanserinin erken tespiti için kullanılabilir. Tıbbi kullanım, yüksek eğitimli görüntü işlemcileri gerektirdiğinden, bu uygulamalar, kullanım için kabul edilmeden önce önemli uygulama ve değerlendirme gerektirir (Alzubaidi vd., 2021).

3.3.2. Trafik Algılama Teknolojileri

Trafik sensörleri durumunda, bir video görüntü işleme sistemi (Video Image Processing

System - VIPS) kullanılabilir. Bu sistem, bir görüntü yakalama sistemi, bir telekomünikasyon sistemi ve bir görüntü işleme sisteminden oluşmaktadır. Video çekerken, bir VIPS, bir araç bölgeye girdiğinde bir "açık" sinyali veren ve ardından araç algılama bölgesinden çıktığında bir "kapalı" sinyali veren birkaç algılama bölgesine sahiptir. Bu algılama bölgeleri birden fazla şerit için ayarlanabilir ve belirli bir istasyondaki trafiği algılamak için kullanılabilir. Bunun yanı sıra, aracın plakasını otomatik olarak kaydedebilir, araç tipini ayırt edebilir, sürücünün otoyoldaki hızını izleyebilir ve çok daha fazlasını yapabilir (Liang, 2010).

3.3.3. Görüntü Yeniden Yapılandırma

Görüntü işleme, bir görüntünün eksik veya bozuk kısımlarını kurtarmak ve doldurmak için kullanılabilir. Bu, eski ve hasarlı fotoğrafların daha yeni sürümlerini oluşturmak için mevcut fotoğraf veri kümeleriyle kapsamlı bir şekilde eğitilmiş görüntü işleme sistemlerinin kullanılmasını içerir (Wei, 2013).

3.3.4. Yüz Tanıma

Bugün kullandığımız en yaygın görüntü işleme uygulamalarından biri yüz tanımadır ve bu çalışmanın odaklandığı görüntü işleme grubudur. Makinenin yüz şekli, gözler arasındaki mesafe gibi insan yüzlerinin belirli özellikleriyle ilk kez eğitildiği derin öğrenme algoritmalarını takip etmektedir (Serengil ve Ozpınar, 2020). Makineye bu insan yüzü özelliklerini öğrettikten sonra, insan yüzüne benzeyen bir görüntüdeki tüm nesneleri kabul etmeye başlayacaktır. Yüz algılama, bugünlerde çoğu sosyal medya uygulamasında bulunan güvenlik, biyometri ve hatta filtrelerde kullanılan hayati bir araçtır.

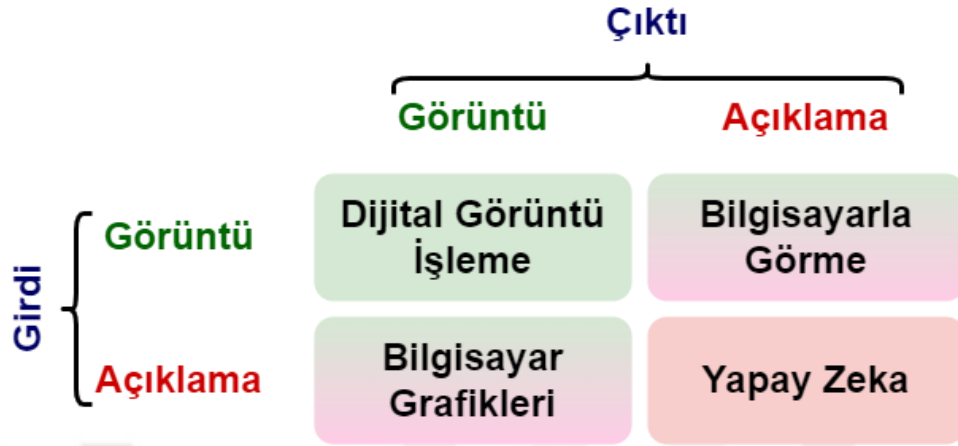
Dijital Görüntü İşleme ve diğer alanlar görevleri bakımından ortak özelliklere sahiptir. Bu alanları Sekil 11'i inceleyerek daha iyi anlayabiliriz.

Blok 1'e göre, oluşturulan sistemin girdisi bir görüntü ise ve sistem görüntü üzerindeki işlemlerden sonra yine görüntü olarak çıktı veriyorsa, bu işleme dijital görüntü işleme denilmektedir.

Blok 2'ye göre, oluşturulan sistemin girdisi bir görüntü ise ve sistem görüntü üzerinde uyguladığı yöntemler sonucunda bir bilgi veya tanım oluşturuyorsa, bu işleme bilgisayarlı görme denilmektedir.

Blok 3'e göre, işlem yapacak sistemin girdisi bir tür açıklama veya kod ve sistem işlemler sonucunda görüntü çıktısı veriyorsa, bu işleme bilgisayar grafik yapısı denilmektedir.

Blok 4'e göre, işlem yapacak sisteminde girdisi bir tür açıklama veya kod ve sistem işlemler sonucunda bir açıklama veya oluşturduğu anahtar kelimeleri veriyorsa, bu sistemlere yapay zeka denilmektedir.



Şekil 3.2. Görüntü İşleme ile Kesişen Alanlar

Tüm bu farklı işlemlerin ışığında, dijital görüntü işlemenin çeşitli alanlarda birçok faydası bulunmaktadır.

3.4. Dijital Görüntü İşlemenin Faydaları

Dijital Görüntü İşleme tekniklerinin uygulanması, birçok teknoloji kuruluşu üzerinde büyük bir etki yaratmıştır. İşlem alanından bağımsız olarak görüntü işlemenin en iyi faydalarından bazıları şunlardır (Aaron ve Chew, 2021):

- Dijital görüntü istenilen herhangi bir formatta (geliştirilmiş görüntü, X-Ray, fotoğraf negatifi, vb.) olabilir,
- İnsan yorumlaması için görüntülerin geliştirilmesine yardımcı olabilir,
- Bilgi, makine yorumu için işlenebilir ve görüntülerden çıkarılabilir,
- Görüntüdeki pikseller, istenilen yoğunluk ve kontrastta manipüle edilebilir,
- Görüntüler kolayca saklanabilir ve alınabilir,
- Görüntülerin üçüncü taraf sağlayıcılara kolayca elektronik olarak iletilmesini sağlayabilir.

Günümüzde teknolojik altyapıların gelişmesi ile birlikte yaygın olarak kullanılan Evrişimli Sinir Ağı (Convolutional Neural Network – CNN) dijital görüntü işleme yapılarında yüksek performansa sahip derin öğrenme yöntemlerinden biridir. Güçlü görüntü işleme ve yapay zeka mimarisine ihtiyaç duyulan doğal dil işleme, video veya görüntü tanımlama gibi yapılarda derin öğrenme yapısı olan CNN hem üretim hem de tanımlama alanlarındaki işlemleri gerçekleştirebilmektedir.

3.5. Evriřimli Sinir Ađları (Convolutional Neural Networks – CNN)

Çeřitli derin öğrenme yöntemleri, farklı nesne sınıflarının sınıflandırılması gibi çeřitli makine öğrenimi görevlerini yapmak için sinir ađı algoritmalarını eğitmek için verileri kullanır. Evriřimli Sinir Ađları, görüntülerin analizi için çok güçlü olan derin öğrenme algoritmalarıdır (Oh vd., 2021).

Evriřimsel sinir ađları, dijital görüntü işleme için güçlü bir algoritmadır. Bu algoritmalar, son dönemde görüntülerin otomatik olarak işlenmesi için sahip olunan en iyi algoritmalarıdır. Birçok alanda, bir görüntüdeki nesneleri tanımlamak gibi görevleri yapmak için bu algoritmalarından yararlanılmaktadır.

Evriřimsel sinir ađları (CNN) beynin mimarisinden ilham alır. Tıpkı beyindeki bir nöronun tüm vücutta bilgiyi işleme ve iletmesi gibi, CNN'lerdeki yapay nöronlar veya düğümler girdileri alır, işler ve sonucu çıktı olarak gönderir (Roy vd., 2020). Görüntü girdi olarak beslenir. Giriř katmanı, görüntü piksellerini diziler biçiminde girdi olarak kabul eder. CNN'lerde, hesaplamalar yaparak görüntüden özellik çıkarımı yapan birden fazla gizli katman olabilir.

Bilgisayar sistemlerinde sinir ađı modeli oluşturulması için insanların beyin yapısında bulunan nöronların sistematik çalışma mantıkları gözlemlenerek donanımsal ve yazılımsal olarak modelleme yapılmaktadır. Genel olarak dijital görüntü işleme mimarileri için Yapay Sinir Ađları (Artificial Neural Network – ANN) yapıları görüntüleri düşük çözünürlüklü bölümlere yaparak incelediğinden ideal bir yöntem olmamaktadır (Abiodun vd., 2019). CNN yapısında modellenen sinir ađları, insanların veya hayvanların görüntü algılama sistemlerinden sorumlu olarak yapılarında bulunan ön lobundaki sinir yapılarına benzemektedir. Bu sayede, kısmi olarak parça parça görüntüleri işleme almaktansa tüm görsel alanı kaplayacak durumda nöron katmanlarını modelleyerek çalışmaktadır (Rahman vd., 2020).

3.6. Evriřimli Sinir Ađlarının İlham Kaynakları

1959'da Hubel ve Wiesel, beynin görsel korteksinin görsel bilgiyi nasıl işlediğini anlamak için bir deney yapmışlardır (Hubel ve Wiesel, 1959). Bu deney için, bir kedinin önünde parlak bir çizgiyi hareket ettirirken kedinin görsel korteksindeki nöronların aktivitesini kaydetmişlerdir. Parlak çizgi belirli bir açıda ve belirli bir yerde gösterildiğinde bazı basit

hücrelerin ateşlendiğini fark etmişlerdir. Diğer nöronlar, açı/konum ne olursa olsun parlak çizgi gösterildiğinde ateşlenmiş ve hareketi karmaşık hücreler algılamıştır. Karmaşık hücrelerin birden fazla basit hücreden girdi aldığı ve hiyerarşik bir yapıya sahip olduğu görülmüştür. Hubel ve Wiesel, 1981'de bulguları için Nobel ödülünü kazanmışlardır.

1980 yılında, karmaşık ve basit hücrelerin hiyerarşik yapısından ilham alan Fukushima, el yazısı Japonca karakter tanıma için kullanılan hiyerarşik bir sinir ağı olan Neocognitron'u (Fukushima 1988) önermiştir. Neocognitron ilk CNN olmuştur ve kendi eğitim algoritmasına sahiptir. 1989 yılında LeCun vd., geri yayılım algoritması ile eğitilebilecek bir CNN önermiştir (LeCun vd., 1989). CNN, ImageNet Büyük Ölçekli Görsel Tanıma Mücadelesinde (ILSVR) diğer modellerden daha iyi performans gösterdiğinde muazzam bir popülerlik kazanmıştır. ILSVRC, yüzlerce nesne kategorisi ve milyonlarca görüntü üzerinde nesne sınıflandırma ve algılama ile ilgili bir yarışmadır. Yarışma 2010'dan günümüze kadar her yıl yürütülmekte ve elliden fazla kurumun katılımıyla gerçekleştirilmektedir (Russakovsky vd., 2015). ILSVRC'yi kazanan önemli CNN mimarileri, 2012'de AlexNet (Krizhevsky, Sutskever ve Hinton, 2012), 2013'te ZFNet (Wan vd., 2013), 2014'te GoogLeNet ve VGG (Goodfellow vd., 2015; Jaderberg vd., 2015) ve 2015'te ResNet'tir.

Yukarıda belirtilen olay ve durumlar, CNN'in doğup gelişmesinde katkıda bulunmuştur.

3.7. CNN Mimarisi

Tipik bir sinir ağı, bir girdi katmanına, gizli katmanlara ve bir çıktı katmanına sahip olacaktır. Bu, evrişim, havuzlama, doğrultulmuş doğrusal birimler ve tamamen bağlı katmanları içerebilir. Evrişim, bir girdi görüntüsünden özellik çıkarımı yapan ilk katmandır. Tam bağlantılı katman, nesneyi sınıflandırır ve çıktı katmanında tanımlar. CNN'ler, tıpkı yapay sinir ağlarının biyolojik olarak ilham alması gibi bilgi akışının girdilerinden çıktılarına kadar yalnızca bir yönde gerçekleştiği ileri beslemeli ağlardır (Razavian vd., 2014). CNN mimarileri çeşitli varyasyonlarda gelir; bununla birlikte, genel olarak, modüller halinde gruplandırılmış evrişim ve havuzlama (veya alt örnekleme) katmanlarından oluşurlar.



Şekil 3.3. CNN Mimarisi

CNN mimarisinde oluşturulan her bir yeni katman dijital görüntünün büyük miktardaki segmentasyonlarını belirleyerek karmaşıklığını artırmaktadır. Başta oluşturulan katmanlar genellikle daha basit özellikler olan renkler veya kenarlara odaklanmaktadır. Girdi olarak alınan dijital görüntü, oluşturulan CNN mimarisindeki katmanlar üzerindeki işlemler sırasıyla uygulandıkça, ulaşılmaya çalışılan nesneyi anlaşılır bir hale getirine kadar daha büyük nesneleri ve yapıları tanımaya başlamaktadır (Chua ve Roska, 1993).

3.7.1. CNN Katmanları

Evrişimli Sinir Ağlarını (CNN) diğer sinir ağları sistemlerinden ayıran en önemli özelliği görüntü, konuşma veya ses sinyali girdileri üzerindeki işlemlerde sahip olduğu yüksek performansdır (Alzubaidi vd., 2021). Temel olarak üç ana katmandan oluşmaktadır:

- Evrişimsel Katman (Convolutional Layer)
- Havuzlama Katmanı (Pooling Layer)
- Tam Bağlantılı Katman (Fully Connected Layer - FC)

3.7.1.1 Evrişimsel Katman

Evrişim katmanı, bir CNN'nin temel yapı taşıdır ve hesaplamaların çoğunluğunun gerçekleştiği yerdir. Bu katmanda oluşturulan girdi verileri, veriler içerisindeki gereksinim duyulmayan verilerin filtrelenmesi ve gerekli özellik yapı haritası gibi kullanılacak öğeler oluşturulmaktadır. Girilen görüntü verilerinin, üç boyutlu piksel matrisi olarak renklendirilmiş dijital bir görüntü olacağını varsaydığımızda, girdinin içerisindeki RGB'ye denk olarak üretilen yükseklik, genişlik ve derinlik olarak üç boyuta sahip olması gerekmektedir. Ayrıca, reseptör olarak belirlenen yapılar tarafından görüntünün hareket ederek belirlenen özellik yapısının var olup olmadığını kontrol eden, çekirdek veya filtre olarak da bilinen bir özellik algılayıcıya da sahip oluruz. Bu süreç konvolüsyon olarak bilinir.

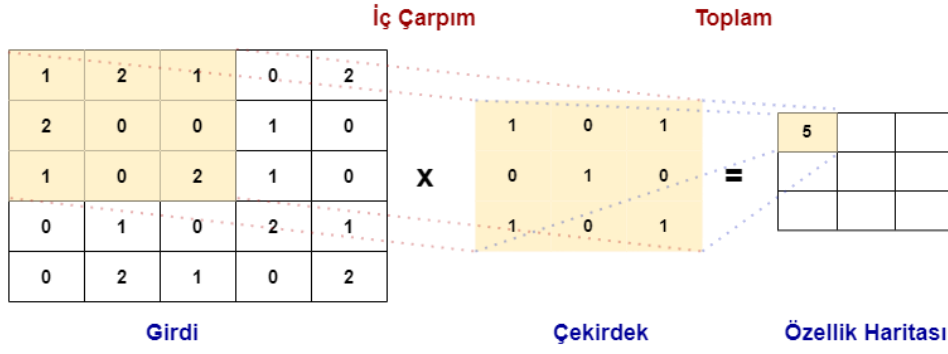
Evrişim ise tensör olarak adlandırılan bir sayı dizisi yapısındaki girdi boyutunda çekirdek adında daha küçük sayı dizisinin özellik çıkarım için kullanılan özelleştirilmiş doğrusal operasyon türüdür. Özellik algılayıcı, görüntünün bir bölümünü temsil eden iki boyutlu (Two Dimensional - 2D) bir ağırlık dizisidir. Boyutları değişebilse de filtre ya da çekirdek boyutu tipik olarak 3×3 , ancak bazen 5×5 veya 7×7 'lik bir matristir; bu aynı zamanda alıcı alanın boyutunu da belirler. Birden çok çekirdek, yatay kenar algılayıcı (üst), dikey kenar algılayıcı (orta) ve ana hat algılayıcı (alt) gibi farklı özellik çıkarıcılar olarak çalışır. Filtre daha sonra görüntünün bir alanına uygulanır ve giriş pikselleri ile filtre arasında bir nokta çarpım hesaplanır. Bu nokta çarpımı daha sonra bir çıktı dizisine beslenir. Daha sonra, filtre bir adım öteye geçer ve çekirdek tüm görüntü boyunca dolaşana kadar bu işlemi tekrarlar. Girdi ve filtreden elde edilen nokta ürünleri serisinin son çıktısı, özellik haritası, etkinleştirme haritası veya kıvrımlı özellik olarak tanımlanır.

Şekil 13'te gösterildiği gibi, giriş tensörüne bir çekirdek uygulanıp çekirdeğin her ögesi ile giriş tensörü arasındaki öge bazında bir ürün her konumda hesaplanır ve çıkış tensörünün karşılık gelen konumunda çıkış değerini elde etmek için toplanır, buna özellik haritası denir.

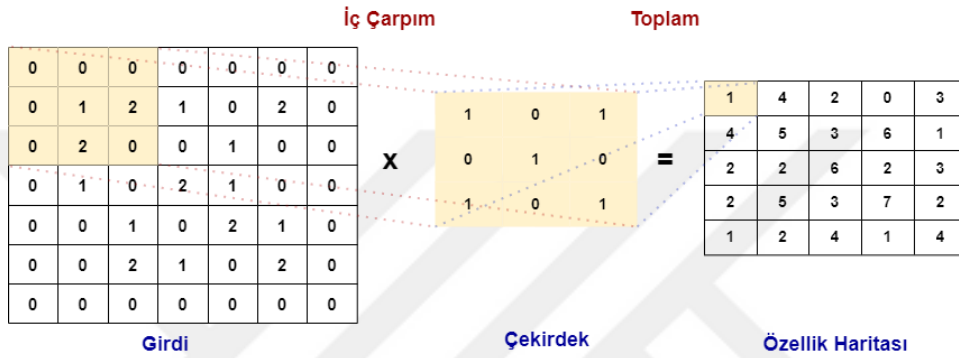
Çıktı olan özellik haritasında 5×5 'lik bir girdi boyutunun tutulduğu görülmektedir. Bu örnekte, sıfır dolgu kullanılıp bir çekirdek boyutu ve adım, sırasıyla 3×3 ve 1 olarak ayarlanmıştır.

Aşağıdaki şekillerde de görebileceğiniz gibi, özellik haritasındaki her bir çıkış değeri, giriş görüntüsündeki her bir piksel değerine bağlanmak zorunda değildir. Yalnızca filtrenin uygulandığı alıcı alana bağlanması gerekir. Çıkış dizisinin her giriş değerine doğrudan eşlenmesi gerekmediğinden, evrişimli (ve havuzlama) katmanlara genellikle "kısmen bağlı" katmanlar denir. Ancak bu özellik yerel bağlantı olarak da tanımlanabilir.

Parametre paylaşımı olarak da bilinen, görüntü boyunca hareket ederken özellik algılayıcıdaki ağırlıkların sabit kalması önemlidir. Ağırlık değerleri gibi bazı parametreler, eğitim sırasında geri yayılım ve gradyan inişi süreci boyunca ayarlanır. Ağırlıkların bağımsız olduğu yapay sinir ağlarından farklı olarak, CNN'nin ağırlıkları görüntünün her bölümündeki özellikleri çıkarmak için komşu piksellere eklenir.



(a) Çekirdek (kernel) boyutu 3×3 olan, dolgunsuz (padding) ve 1 adımlı (stride) bir evrişim işlemi örneği



(b) Düzlem içi boyutları korumak için sıfır dolgulu bir evrişim işlemi

Şekil 3.4. Evrişim İşlemi Örneği (Yamashita vd., 2018)

Yapay sinir ağının eğitimi başlamadan önce ayarlanması gereken çıktının hacim boyutunu etkileyen üç hiper parametre vardır. Bunlar şunları içerir:

1. Filtre sayısı çıktının derinliğini etkiler. Örneğin, üç farklı filtre, üç derinlik oluşturarak üç farklı özellik haritası verir. Ayrıca, tüm giriş piksellerinin çıkış pikselinin etrafında olması güzel geometrik özelliğinden dolayı, tek boyutlu filtreler çift boyutlu filtrelere tercih edilir.
2. Adım, (Stride) çekirdeğin girdi matrisi üzerinde hareket ettiği mesafe veya piksel sayısıdır. İki veya daha büyük adım değerleri nadir olmakla birlikte, daha büyük bir adım daha küçük bir çıktı verir.
3. Sıfır dolgu (Zero-padding) genellikle filtreler giriş görüntüsüne uymadığında kullanılır. Bu, girdi matrisinin dışında kalan tüm öğeleri sıfıra ayarlar ve daha büyük veya eşit boyutlu bir çıktı üretir. Üç tip dolgu vardır:

Geçerli dolgu: Bu aynı zamanda dolgu yok olarak da bilinir. Bu durumda, boyutlar hizalanmazsa son evrişim atlanır.

Aynı dolgu: Bu dolgu, çıktı katmanının giriş katmanı ile aynı boyutta olmasını sağlar.

Tam dolgu: Bu dolgu türü, girdinin sınırına sıfır ekleyerek çıktının boyutunu artırır.

Evrişim işleminde temel olarak yapılan çekirdeklerin tüm görüntü konumlarında ağırlık paylaşımı yapmasıdır. Evrişim işlemleri ile ağırlık paylaşımı aşağıdaki özellikleri üretir:

- 1) çekirdekler tüm görüntü konumları boyunca seyahat ederken ve öğrenilen yerel kalıpları tespit ederken, çekirdek çevirisi tarafından çıkarılan yerel öznitelik modellerinin değişmez olması sağlanır,
- 2) alt örnekleme yoluyla özellik modellerinin uzamsal hiyerarşilerini öğrenmek giderek daha geniş bir görüş alanının yakalanmasıyla sonuçlanan bir havuzlama işlemi ile bağlantılıdır,
- 3) tam bağlantılı sinir ağlarına kıyasla öğrenilecek parametre sayısını azaltarak model verimliliğini artırmaktadır.

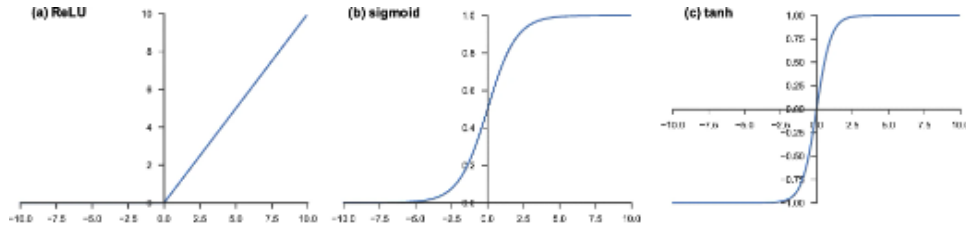
Bir CNN modelinin evrişim katmanına göre eğitilmesi süreci, belirli bir eğitim veri kümesine dayalı olarak belirli bir görev için en iyi çalışan çekirdekleri belirlemektir. Çekirdekler, evrişim katmanındaki eğitim süreci sırasında otomatik olarak öğrenilen tek parametredir. Öte yandan, çekirdeklerin boyutu, çekirdek sayısı, dolgu ve adım, eğitim süreci başlamadan önce ayarlanması gereken hiper parametrelerdir.

Evrişim gibi doğrusal bir işlemin çıktıları daha sonra doğrusal olmayan bir etkinleştirme işlevinden geçirilir. Sigmoid veya hiperbolik tanjant ($\tanh$) işlevi gibi düzgün doğrusal olmayan işlevler, biyolojik bir nöron davranışının matematiksel temsilleri oldukları için daha önce kullanılmış olsa da şu anda kullanılan en yaygın doğrusal olmayan etkinleştirme işlevi, basitçe hesaplayan doğrultulmuş doğrusal birimdir.

$$(\text{Rectified Linear Unit} - \text{ReLU}): f(x) = \max(0, x)$$

Her evrişimli ve maksimum havuzlama işleminden sonra Doğrultulmuş Lineer Birim (Rectified Linear Unit – ReLU) uygulanabilir. ReLU işlevi, $x > 0$ değerleri için doğrusal

olmama durumunu tanıtmak için "yeterince büyük bir uyarın" üzerindeki nöron aktivasyonlarını taklit eder ve koşulu karşılamıyorsa 0 döndürür (Ide ve Kurita, 2017). Bu yöntem, azalan gradyanları çözmek için etkili olmuştur. Çok küçük olan ağırlıklar ReLU aktivasyon fonksiyonu sonrasında 0 olarak kalacaktır. Ayrıca, Şekil 14'te belirtilen sigmoid ve tanh aktivasyon operatörleri de kullanılabilir.



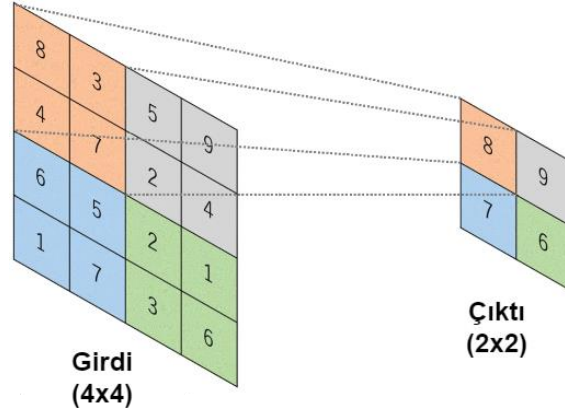
Şekil 3.5. Sinir ağlarına yaygın olarak uygulanan etkinleştirme işlevleri: a) doğrultulmuş doğrusal birim (ReLU), b) sigmoid, ve c) hiperbolik tanjant (tanh)

3.7.1.2. Havuzlama Katmanı

Alt örnekleme olarak da bilinen havuzlama katmanları, girdideki parametre sayısını azaltarak boyutsallık azaltmayı gerçekleştirir. Bu katmanda yapılan operasyonel işlemler evrişim katmanındakine benzer olarak bir filtreyi tüm girdi boyunca süpürürken, uygulanacak filtrenin bir ağırlığa sahip olmamasıyla evrişim katmanından ayrılır (Gong, Wang, Guo ve Lazebnik, 2014). Filtrelerin ağırlıkları yerine çekirdek alıcı alan olarak belirlenen değerlere çıktı dizisini oluşturacak şekilde toplama operatörünü uygulamaktadır. Temel olarak iki havuzlama türü mevcuttur:

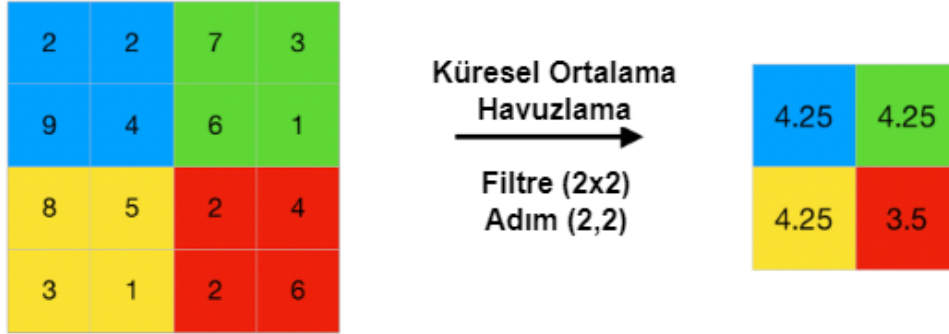
Maksimum havuzlama: Filtre giriş boyunca hareket ederken, çıkış dizisine gönderilecek maksimum değere sahip pikseli seçer. Bu yaklaşım ortalama havuzlamaya kıyasla daha sık kullanılma eğilimindedir. Boyutlar arasındaki farkındalık sağlayacak yükseklik ve genişlik yapılarından değişik olarak derinlik boyutu oluşturulan özellik haritalarında herhangi bir değişime uğramadan kalır.

Şekil 15'te verilen örnekte, 2×2 filtre boyutu, dolgu olmayan ve giriş tensörlerinden 2×2 parçaları çıkaran, her parçadaki maksimum değeri veren ve diğer tüm değerleri atan 2 adımlı bir maksimum havuzlama işlemi örneği gösterilmektedir.



Şekil 3.6. Maksimum Havuzlama Örneği

Küresel Ortalama Havuzlama: Filtre giriş boyunca hareket ettikçe, çıkış dizisine gönderilecek alıcı alan içindeki ortalama değeri hesaplar ve tamamen bağlı katmanlardan önce bu işlem bir kez uygulanmaktadır. Havuzlama katmanında birçok bilgi kaybolmasına rağmen, CNN'ye birtakım faydaları da vardır. Karmaşıklığı azaltmaya, verimliliği artırmaya ve fazla uyum riskini sınırlamaya yardımcı olurlar.



Şekil 3.7. Küresel Ortalama Havuzlama Örneği

Şekil 16'da verilen örnekte, 2×2 filtre boyutu, dolgu olmayan ve giriş tensörlerinden 2×2 yamaları çıkaran, her parçanın ortalamasını alan ve diğer tüm değerleri atan 2 adımlı bir küresel ortalama havuzlama işlemi örneği gösterilmektedir.

3.7.1.3. Tam Bağlantılı Katman

Tam bağlantılı katmanın adı kendini uygun bir şekilde tanımlamaktadır. Daha önce bahsedildiği gibi, kısmen bağlantılı katmanlarda girdi görüntüsünün piksel değerleri doğrudan çıktı katmanına bağlı değildir. Ancak, tam bağlantılı katmanda, çıktı katmanındaki her bir düğüm, bir önceki katmandaki bir düğüme doğrudan bağlanır.

Bu katman, önceki katmanlardan ve bunların farklı filtrelerinden çıkarılan özniteliklere dayalı olarak sınıflandırma görevini gerçekleştirir. Evrişim ve havuzlama katmanları ReLU işlevlerini kullanma eğilimindeyken, Tam bağlantı katmanları genellikle girdileri uygun şekilde sınıflandırmak için bir softmax etkinleştirme fonksiyonundan yararlanarak 0 ile 1 arasında bir olasılık üretir (Lin, RoyChowdhury ve Maji, 2015).

Son evrişim veya havuzlama katmanının özellik harita çıktıları tipik olarak düzleştirilir, yani tek boyutlu (1-Dimensional - 1D) sayı dizisine (veya vektöre) dönüştürülür ve yoğun katmanlar olarak da bilinen bir veya daha fazla her girdinin her çıktıya öğrenilebilir bir ağırlıkla tam olarak bağlandığı katmana bağlanır (Prajapati vd., 2019). Evrişim ve havuz katmanları tarafından oluşturularak alt örneklenen yapısal özellikler oluşturulduktan sonra, sınıflandırma görevlerindeki her bir sınıf için olan olasılıklar gibi, ağırlık çıktılarına tam bağlantılı katmanın bir alt kümesi tarafından eşlenir. Meydana getirilen tam bağlantılı katman, genel olarak çıktı sayısını oluşturulan sınıf sayısı ile eşit olacak şekilde düğüme sahiptir (Cimpoi vd., 2016). Yukarıda da açıklandığı gibi ReLU gibi doğrusal olmayan bir fonksiyonu her tam bağlantılı katman takip eder.

Son tam bağlantılı katmana uygulanan aktivasyon fonksiyonu genellikle diğerlerinden farklıdır. Her göreve göre uygun bir aktivasyon fonksiyonu seçilmelidir. Çok sınıflı sınıflandırma görevine uygulanan bir aktivasyon fonksiyonu, her bir değerin 0 ile 1 arasında değiştiği ve tüm değerlerin toplamının 1 olduğu, son tam bağlı katmandan çıktı gerçek değerlerini hedef sınıf olasılıklarına normalleştiren bir softmax fonksiyonudur.

Aktivasyon Fonksiyonları

Aktivasyon fonksiyonları, bir yapay sinir ağı tasarımının kritik bir parçasıdır. Gizli katmandaki aktivasyon fonksiyonunun seçimi, ağ modelinin eğitim veri setini ne kadar iyi öğrendiğini kontrol edecektir, ayrıca yapay sinir ağının kapasitesi ve performansı üzerinde büyük bir etkiye sahiptir (Soe vd., 2019). Çıktı katmanındaki aktivasyon fonksiyonunun seçimi ise, modelin yapabileceği tahminlerin türünü tanımlayacaktır.

Bir sinir ağı yapısındaki aktivasyon fonksiyonu, girdinin ağırlıklı toplamının bir düğümden veya ağ katmanındaki düğümlerden bir çıktıya nasıl dönüştürüleceğini tanımlar (Thiele vd., 2019). Bazen aktivasyon fonksiyonuna "Aktarım Fonksiyonu" da denir. Birçok aktivasyon fonksiyonu doğrusal değildir. Teknik olarak, ağlar bir katmandaki tüm düğümler

için aynı aktivasyon fonksiyonunu kullanmak üzere tasarlanmasına rağmen, aktivasyon fonksiyonu ağıdaki her düğümün dahili işlenmesi içinde veya sonrasında kullanılır.

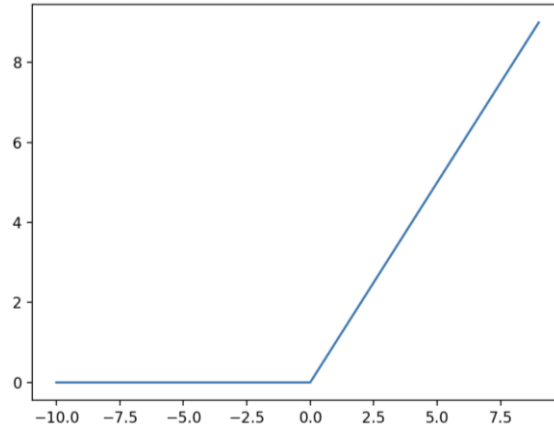
Gizli Katmanlar için Aktivasyon

Bir sinir ağındaki gizli katman, başka bir katmandan (başka bir gizli katman veya bir girdi katmanı gibi) girdi alan ve başka bir katmana (başka bir gizli katman veya bir çıktı katmanı gibi) çıktı sağlayan bir katmandır. Gizli bir katman, bir model için doğrudan girdi verileriyle temas kurmaz veya çıktılar üretmez. Bir sinir ağı modeli, sıfır veya daha fazla gizli katmana sahip olabilir.

Tüm gizli katmanlar tipik olarak aynı aktivasyon fonksiyonunu kullanır. Çıktı katmanı tipik olarak gizli katmanlardan farklı bir aktivasyon fonksiyonu kullanır ve modelin gerektirdiği tahmin türüne bağlıdır. Tipik olarak, bir sinir ağının gizli katmanlarında türevlenebilir bir doğrusal olmayan aktivasyon işlevi kullanılır. Bu, modelin doğrusal bir etkinleştirme işlevi kullanılarak eğitilmiş bir ağdan daha karmaşık işlevleri öğrenmesini sağlar (Pedamonti, 2018).

Gizli katmanlarda kullanılan aktivasyon fonksiyonu, tipik olarak sinir ağı mimarisinin türüne göre seçilir. Çoklu katman algılayıcıları (Multilayer perceptron – MLP) ve CNN gibi ortak mimarilere sahip modern sinir ağı modelleri, ReLU etkinleştirme işlevini veya uzantılarını kullanacaktır. En yaygın şekilde kullanılan ReLU fonksiyonu sigmoid veya tanh gibi etkinleştirme fonksiyonlarına göre sınırlamaları daha basit ve etkili olarak aşabilmektedir. Ayrıca, derin öğrenme modellerinin eğitim aşamalarında engellemeye sebebiyet verebilecek kaybolan gradyanlara daha az duyarlı iken, doymuş birimlere karşıda hassasiyet gösterebilmektedir (Stursa ve Dolezel, 2019).

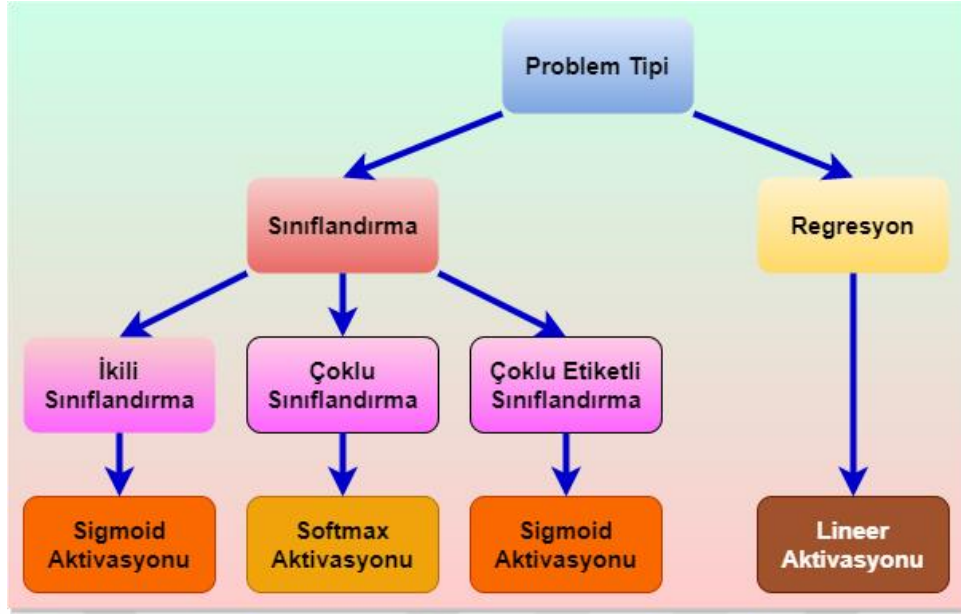
ReLU aktivasyon fonksiyonu $f(x) = \max(0.0, x)$ olarak hesaplanır. Eğer x değeri pozitif bir değer olarak girdiden gelmiyorsa, Şekil 17’de belirtildiği gibi 0.0 değeri geri döndürülürken, pozitif değerler için kendi özgün değerleri döndürülecektir.



Şekil 3.8. ReLU Aktivasyon İşlevi Girdi ve Çıktı Değerleri

Çıktı Katmanları için Aktivasyon

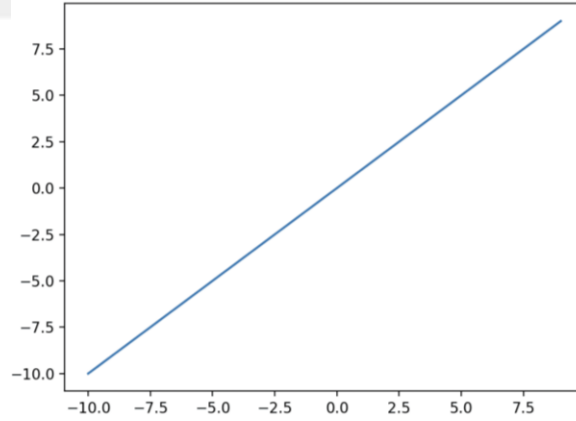
Çıktı katmanı, doğrudan bir tahmin çıktısı veren bir sinir ağı modelindeki katmandır. Tüm ileri beslemeli sinir ağları modellerinde bir çıktı katmanı bulunur. Çıktı katmanında kullanmak için göz önünde bulundurulmuş doğrusal, lojistik (sigmoid) ve softmax aktivasyon fonksiyonları mevcuttur. Çıktı aktivasyon fonksiyonu tahmin probleminin türüne göre çıktı katmanı için aktivasyon fonksiyonunu seçilir ve Şekil 18’de türleri gösterilmektedir. Eğer çözülmesi gereken problem bir regresyon sorunuysa, doğrusal bir etkinleştirme işlevi kullanılmalıdır (Nwankpa vd., 2018). Sorun bir sınıflandırma sorunuysa, üç ana sınıflandırma sorunu türü vardır ve her biri farklı bir etkinleştirme işlevi kullanabilir. Bir olasılığı tahmin etmek bir regresyon problemi yerine sınıflandırma problemi olarak işleme alınır. Tüm sınıflandırma durumlarında, model yuvarlama (sigmoid için) veya argmax (softmax için) yoluyla net bir sınıf etiketine dönüştürebilecek sınıf üyeliği olasılığı (örneğin, bir örneğin her sınıfa ait olma olasılığı) tahmin edilecektir. Eğer birbirini dışlayan iki sınıf (ikili sınıflandırma) varsa, çıktı katmanı bir düğüme sahip olacak ve bir sigmoid aktivasyon fonksiyonu kullanılmalıdır. İki den fazla birbirini dışlayan sınıf varsa (çok sınıflı sınıflandırma), çıktı katmanı sınıf başına bir düğüme sahip olacaktır ve bir softmax aktivasyonu kullanılmalıdır. İki veya daha fazla karşılıklı olarak kapsayıcı sınıf (çok etiketli sınıflandırma) varsa, çıktı katmanında her sınıf için bir düğüm bulunur ve bir sigmoid aktivasyon fonksiyonu kullanılır.



Şekil 3.9. Çıktı Katmanında Aktivasyon Fonksiyonu Seçimi

Lineer Çıktı Katman Aktivasyon Fonksiyonu

Doğrusal aktivasyon fonksiyonu aynı zamanda "kimlik" (1.0 ile çarpılır) veya "etkinleştirme yok" olarak da adlandırılır. Bunun nedeni, doğrusal etkinleştirme işlevinin girdinin ağırlıklı toplamını hiçbir şekilde değiştirmemesi ve bunun yerine değeri doğrudan döndürmesidir (Eckle ve Schmidt-Hieber, 2019).



Şekil 3.10. Lineer Aktivasyon Fonksiyonu

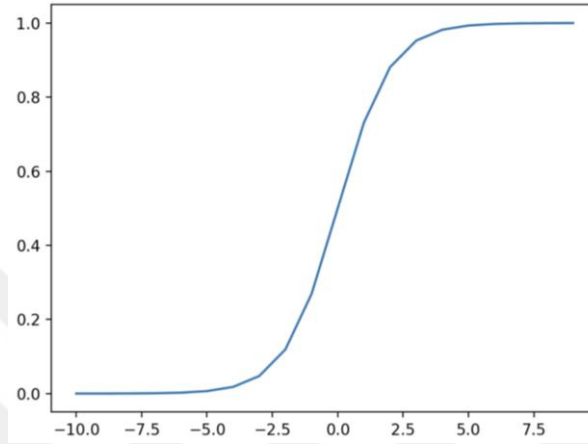
Çıktı katmanında doğrusal bir aktivasyon fonksiyonuna sahip bir modeli eğitmek için kullanılan hedef değerler, normalleştirme veya standardizasyon dönüşümleri kullanılarak modellenmeden önce tipik olarak ölçeklenir ve Şekil 19'da örnek bir birim bazlı ölçeklendirme gösterilmektedir.

Sigmoid Çıktı Katman Aktivasyon Fonksiyonu

Bu fonksiyonun diğer ismi Lojistik Fonksiyon'dur. Bir gerçek değer alınır ve 0 ile 1 aralığındaki değerleri elde edilir. Girdinin büyüklüğü kadar, çıktı değeri 1.0'a o kadar yakın olur, oysa girdinin küçüklüğü kadar, çıktı 0.0 'a o derece yakın olur ve Şekil 20'de örnek olarak gösterilmektedir. Sigmoid aktivasyon fonksiyonu aşağıdaki gibi hesaplanır:

$$1.0/(1.0 + e^{-x})$$

Burada e, doğal logaritmanın temeli olan matematiksel bir sabittir.



Şekil 3.11. Sigmoid Aktivasyon Fonksiyonu

Softmax Çıktı Katman Aktivasyon Fonksiyonu

Softmax fonksiyonu, sınıf üyeliği olasılıkları olarak yorumlanabilen, toplamı 1.0 olan bir değerler vektörü verir. Tüm seçenekler için 0 ve seçilen seçenek için 1 çıktısı veren argmax fonksiyonuyla ilgilidir. Softmax, kazanan hepsini alır fonksiyona benzeyen argmax fonksiyonun farklı bir çeşididir. Bu nedenle, fonksiyonun girdisi gerçek değerlerin bir vektörüdür ve çıktı, 1.0 benzeri olasılıkları toplayan değerlerle aynı uzunlukta bir vektördür. Softmax Fonksiyonu aşağıdaki gibi hesaplanır:

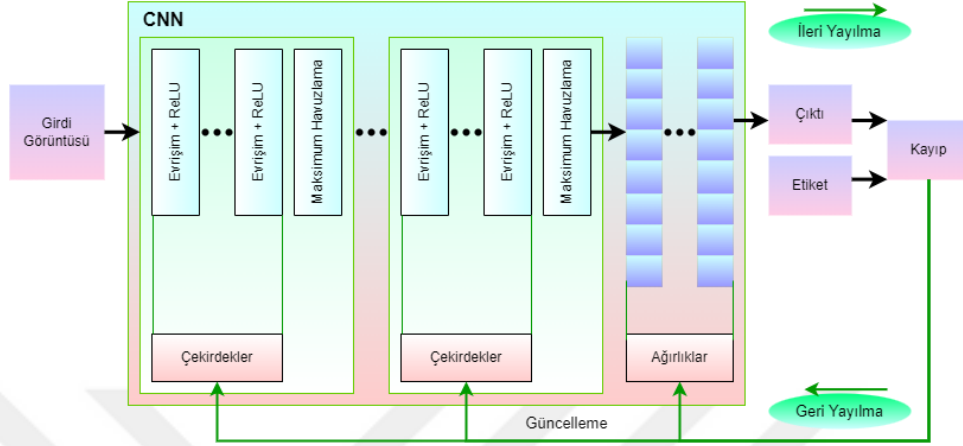
$$\frac{e^x}{\sum(e^x)}$$

Burada x , çıktıların bir vektörü ve e , doğal logaritmanın temeli olan matematiksel bir sabittir.

3.7.2. Ağ Eğitimi

Bir ağı eğitmek, eğitim veri setindeki gerçek etiketler ve elde edilen çıktıların tahminlerin farklarını azaltabildiği kadar azaltan, tam bağlantılı katmanlardaki evrişim katmanlarında ve ağırlıklarda çekirdek bulma durumudur (Yaghoubi vd., 2020). Geri yayılım algoritması ise, sınır ağlarını eğitmek amacıyla kullanılan önemli bir işlemdir. Oluşturulan bir modelin

performansı Şekil 21’de gösterildiği gibi, bir eğitim veri kümesi üzerinde ileriye doğru yayılım yoluyla ve kayıp fonksiyonu ile eğitim veri seti üstünde ileri yayılım vasıtasıyla ve bir kayıp fonksiyonu ile hesaplanıp gradyan inişi gibi parametrelerle kayıp değeri göz önünde bulundurularak değiştirilir (Frankle vd., 2020).



Şekil 3.12. CNN modelinde Ağ Eğitim Modeli

3.7.3. Kayıp Fonksiyonu

Maliyet işlevi olarak da adlandırılan bir kayıp işlevi, ileriye doğru yayılım yoluyla ağıın çıktı tahminleri ile verilen temel doğruluk etiketleri arasındaki uyumluluğu ölçer (Cheng vd., 2016). Çok sınıflı sınıflandırma için yaygın olarak kullanılan kayıp fonksiyonu çapraz entropidir (cross entropy), oysa ortalama kare hatası tipik olarak sürekli değerlere gerilemeye uygulanır. Bu tür kayıp fonksiyonu hiper parametrelerden biridir ve verilen görevlere göre belirlenmesi gerekir.

3.7.4. Gradyan İnişi

Gradyan inişi, maliyeti indirebileceği en alt seviyeye ulaştırmak için ağıın öğrenilebilir parametrelerini, yani çekirdekleri ve ağırlıkları yinelemeli olarak güncelleyen ve yaygın olarak kullanılan bir optimizasyon algoritmasıdır (Du vd., 2019). Kayıp fonksiyon gradyanı, en dik şekilde artış hızının olduğu tarafı verir ve öğrenilebilir her parametre, öğrenme hızı adı verilen bir hiper parametreye dayalı olarak belirlenen rastgele bir adım boyu ile gradyanın negatif yönünde olacak şekilde güncellenir. Gradyan, öğrenilebilir her parametre açısından kaybın kısmi türevi olup matematiksel olarak aşağıdaki şekilde gösterilir:

$$w = w - \alpha * \frac{\partial L}{\partial w}$$

w her öğrenilebilir parametreyi, α bir öğrenme oranını ve L bir kayıp fonksiyonunu temsil eder. Pratikte bir öğrenme oranının eğitim başlamadan önce ayarlanması gereken en önemli hiper parametrelerden biri olduğu unutulmamalıdır. Pratikte, hafıza sınırlamaları gibi nedenlerle, parametrelere göre kayıp fonksiyonunun gradyanları, mini-batch adı verilen eğitim veri kümesinin bir alt kümesi kullanılarak hesaplanır ve parametre güncellemelerine uygulanır.

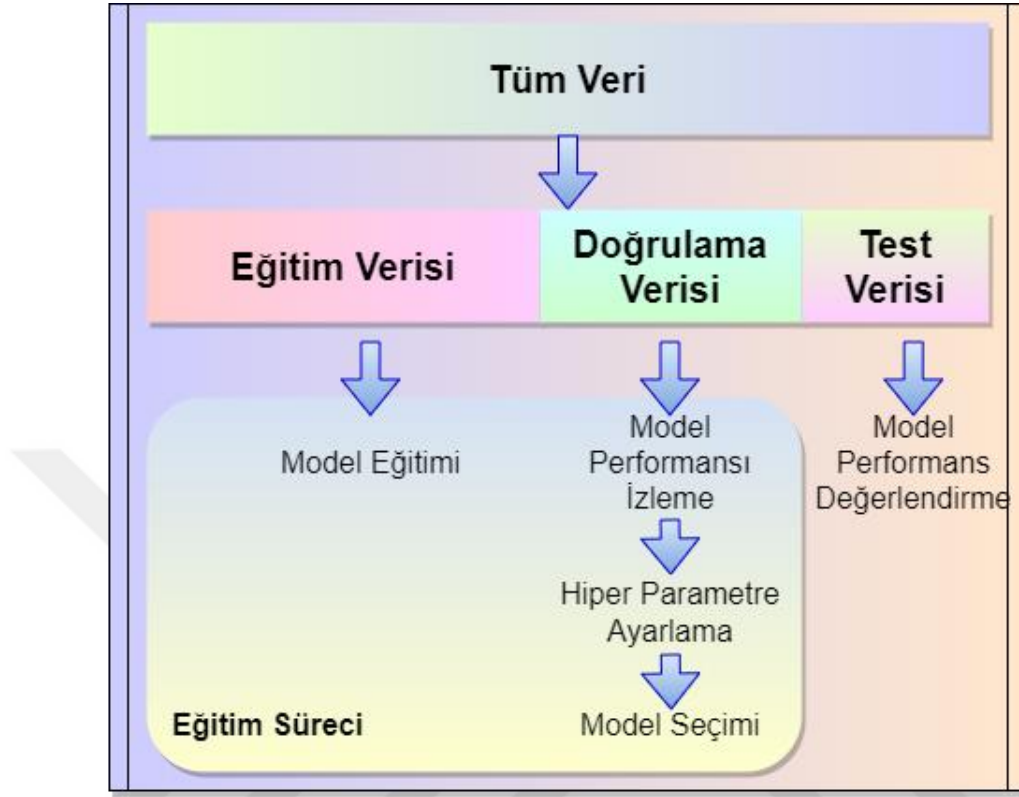
Gradyan iniş, bir maliyet fonksiyonunu en aza indiren bir fonksiyonun parametrelerinin (katsayılarının) değerlerini bulmak için parametreleri sürekli olarak yenileyen bir çeşit optimizasyon algoritmasıdır (Yang ve Yang, 2018).

3.8. Veri ve Temel Doğruluk Etiketleri

Veri ve temel doğruluk etiketleri, derin öğrenme veya diğer makine öğrenimi yöntemlerini uygulayan araştırmalarda en önemli bileşenlerdir. Başarılı bir derin öğrenme projesi için bir modeli eğitmek ve test etmek için veri ve kesin bilgi etiketlerinin dikkatli bir şekilde toplanması zorunludur, ancak yüksek kaliteli etiketli veriler elde etmek maliyetli ve zaman alıcı olabilmektedir (Lecun vd., 2015). Mevcut veriler tipik olarak üç gruba ayrılır: eğitim, doğrulama ve test seti (Şekil 22), ancak çapraz doğrulama gibi bazı değişkenler de vardır. Bir ağı eğitmek için bir eğitim seti kullanılır, burada kayıp değerleri ileri yayılım yoluyla hesaplanır ve öğrenilebilir parametreler geri yayılım yoluyla güncellenir. Eğitim süreci sırasında modeli değerlendirmek, hiper parametrelere ince ayar yapmak ve model seçimini gerçekleştirmek için bir doğrulama seti kullanılır. Bir test seti, eğitim ve doğrulama setleri ile eğitim sürecinde ince ayarlanmış ve seçilen modelin performansını değerlendirmek için ideal olarak projenin en sonunda yalnızca bir kez kullanılır.

Bir modelin eğitimi her zaman hiper parametrelerinin ince ayarını yapmayı ve model seçimi gerçekleştirmeyi içerdiğinden ayrı doğrulama ve test setleri gereklidir. Bu işlem doğrulama kümesindeki performansa dayalı olarak gerçekleştirildiğinden, bu doğrulama kümesiyle ilgili bazı bilgiler modelin kendisine sızar, yani model hiçbir zaman doğrudan öğrenilebilir parametreler için eğitilmemiş olsa bile doğrulama kümesine aşırı uyum sağlar (Ng vd., 2015). Bu nedenle, doğrulama setinde ince ayarlanmış hiper parametrelere sahip modelin aynı doğrulama setinde iyi performans göstereceği garanti edilir. Bu nedenle, model performansının uygun bir şekilde değerlendirilmesi için tamamen görünmeyen bir veri seti, yani ayrı bir test seti gereklidir, çünkü önemli olan daha önce hiç görülmemiş veriler

üzerindeki model performansının değerlendirilmesidir. Makine öğreniminde "doğrulama" terimi genellikle eğitim sürecinde modellerin ince ayarını yapmak ve modelleri seçmek için bir adımı ifade eder.



Şekil 3.13. Veri Modelleri ve Etiketleri

3.8. Aşırı Uyum Gösterme (Overfitting)

Aşırı uyum, bir modelin eğitim kümesine özgü istatistiksel düzenlilikleri öğrendiği, yani sinyali öğrenmek yerine alakasız gürültüyü ezberlediği ve bu nedenle sonraki yeni bir veri kümesinde daha az iyi performans gösterdiği bir durumu ifade eder (Ying, 2019). Bu, makine öğrenimindeki ana zorluklardan biridir, çünkü fazla donatılmış bir model, daha önce hiç görülmemiş verilere genelleştirilemez (Poggio vd., 2017). Bu anlamda, bir test seti, makine öğrenimi modellerinin uygun performans değerlendirmesinde çok önemli bir rol oynar. Eğitim verilerine aşırı uymayı tanımak için rutin bir kontrol, eğitim ve doğrulama setlerindeki kayıp ve doğruluğu izlemektir.

Model, doğrulama setine kıyasla eğitim setinde iyi performans gösteriyorsa, model büyük olasılıkla eğitim verilerine aşırı uyuyordur. Aşırı uyumu göstermeyi azaltmak için en iyi çözüm, daha fazla eğitim verisi elde etmektir. Diğer çözümler, bırakma veya ağırlık azalması ile düzenleme, toplu normalleştirme ve veri artırmanın yanı sıra mimari karmaşıklığı

azaltmayı içerir. Bırakma, modelin ağıdaki belirli ağırlıklara daha az duyarlı hale gelmesi için eğitim sırasında rastgele seçilen aktivasyonların 0'a ayarlandığı, yakın zamanda tanıtılan bir düzenleme tekniğidir (Gal ve Ghahramani, 2016). L2 düzenlemesi olarak da adlandırılan ağırlık azalması, modelin ağırlıklarını cezalandırarak, ağırlıkların yalnızca küçük değerler alması için aşırı uyumu azaltır. Fazla uyumu tanımak için rutin bir kontrol, eğitim yinelenmesi sırasında eğitim ve doğrulama setlerindeki kaybı izlemektir. Model, doğrulama setine kıyasla eğitim setinde iyi performans gösteriyorsa, model eğitim verilerine aşırı uyuyor demektir. Model hem eğitim hem de doğrulama setlerinde düşük performans gösteriyorsa, model verilere yetersiz uyum sağlamıştır. Bir ağı ne kadar uzun süre eğitilirse, eğitim setinde o kadar iyi performans gösterir, bir noktada ağı eğitim verilerine çok iyi uyum sağlar ve genelleme yeteneğini kaybeder.

3.9. Evrişimli Sinir Ağları Uygulama Alanları

Görüntü ve video tanıma uygulamaları için uzmanlaşmış Evrişimli Sinir Ağları (CNN) esas olarak görüntü tanıma, nesne algılama ve segmentasyon gibi görüntü analizi görevlerinde kullanılır. Bunların dışında ise doğal dil işleme, anomali tespiti, ilaç keşfi, sağlık riski değerlendirmesi ve yaşlanma keşfinin biyobelirteçleri, Dama, Go oyunu, zaman serisi tahmini ve kültürel miras ve 3B veri kümeleri gibi alanlarda da kullanılmaktadır (Li vd., 2021). Bu çalışmada CNN modellerini yüz duygu tanıma, cinsiyet ve yaş, ve stres belirleme yapıları için kullanılmaktadır.

3.10. Yüz Duygu Tanıma (Facial Emotion Recognition – FER) Modeli

Yüz Duygu Tanıma (FER) analizi Şekil 22’de de belirtilen bazı temel adımlardan oluşur:

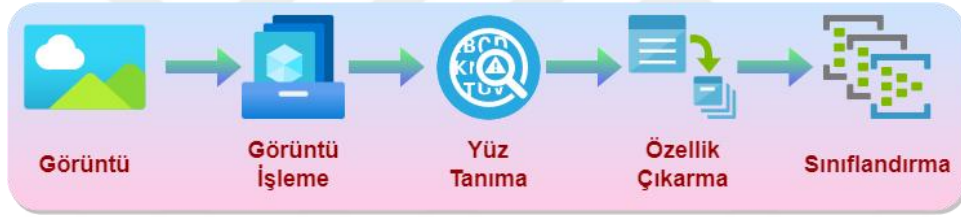
Görüntü Sayısallaştırma: Bu aşamada görüntü, bilgisayar tarafından daha iyi algılanabilmesi için 2 boyutlu bir matrise dönüştürülür.

Görüntü Ön İşleme: Görüntü ön işleme aşaması, görüntüyü FER’e göndermeden önce uygulanan değişiklikler olarak tanımlanabilir. Bir ön işleme aşaması bunun dışında görüntüyü yeniden boyutlandırma, rotasyon düzeltilmesi gibi farklı yöntemlerle giriş verilerinin kalitesini ve özelliklerini iyileştirmek gibi amaçları hedeflemektedir. Böylece, bozuk olan ögeler de düzeltilmiş olmaktadır (Prakash vd., 2018).

Yüz Tespiti: Görüntü ön işleme aşamasından sonra yüzü algılamak için görüntüdeki kas, göz, burun ağız gibi bazı temel noktalar belirlenir.

Özellik (Yüz İfadesi) Çıkarımı: Temel noktalar belirlenerek yüz algılandıktan sonra, yüz hatlarındaki özelliklerden yararlanarak ve analiz edilerek ifade çıkarım işlemi gerçekleştirilir.

Yüz Duygu Durumu Sınıflandırması: Bu aşama, yüz duygu durumu belirlenmesindeki en son aşamadır. Modeldeki duygusal ifadeler denenerek aralarındaki benzerliklere göre en yakın duygu durumu sınıflandırmanın sonucu olarak belirlenir. Farklı model ve algoritmalar da kullanılabilir. Algoritmaya bağlı olarak, yüz ifadeleri temel duygular (örneğin öfke, iğrenme, korku, sevinç, üzüntü ve şaşkınlık) veya bileşik duygular (örneğin mutlu üzgün, mutlu bir şekilde şaşırılmış, mutlu bir şekilde iğrenmiş) olarak sınıflandırılabilir. FER algoritmalarına girdi olarak hizmet veren görüntülerin veya videoların kaynakları, güvenlik kameraları, mağazalardaki reklam ekranları, sosyal medya ve akış servisleri veya kişisel cihazlara yerleştirilmiş kameralar gibi farklı sistemlerden oluşabilir (Jain vd., 2019).



Şekil 3.14. Yüz Duygu Tanıma İş Akışı

3.11. Duygu Durumu Analizi Bölümü

Bu bölümde tezin duygu durumu analizi kısmıyla ilgili model ve veri seti bilgileri bulunmaktadır.

3.11.1. Duygu Durumu Analizi İçin Kullanılan Modeller

Duygu durumu analizi için kod içerisinde 4 farklı CNN modeli kullanılmış ve aralarında karşılaştırmalar yapılmıştır. Bunlar:

- Özelleştirilmiş CNN Modeli
- VGG-16
- ResNet-50
- ResNet-152

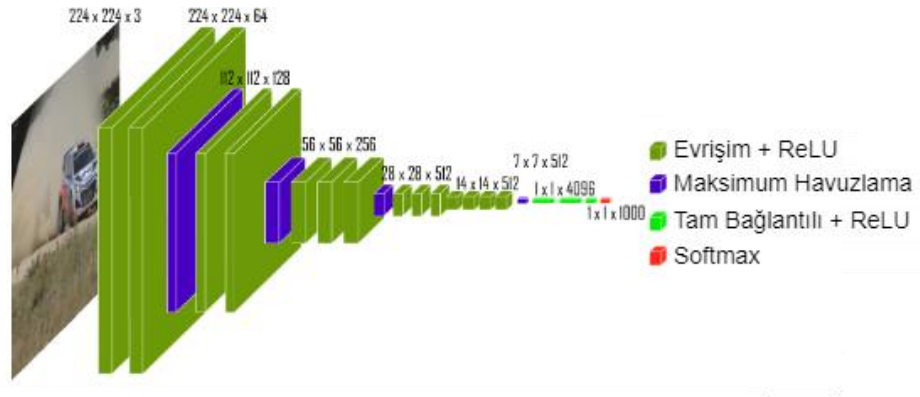
modelleridir.

3.11.1.1. Özelleştirilmiş CNN Modeli

Tezdeki amacımız farklı modeller deneyerek en iyi performans ve sonuç sergileyen modeli bulup uygulamak, aynı zamanda diğer modeller arasında karşılaştırma yapmak olduğu için temel bir yapı ile oluşturduğumuz özelleştirilmiş bir CNN modeli de kullandık. 4 katlamalı katmandan oluşan bu modelde, fazla uyumu azaltmak ve model zamanını ortalamak için her bir katman bir bırakma fonksiyonuna sahiptir. Düzleştirme işlemi ve buna ek olarak 4 tam bağlantılı katmandan oluşmaktadır. Tam bağlantılı katmanlarda aktivasyon fonksiyonu olarak “ReLU” ve son olarak da “softmax” kullanılmıştır. İyileştirici olarak ise “Adam” tercih edilmiştir. Bu model, 1.743.560 eğitilebilir parametre ile sonuçlanmaktadır.

3.11.1.2. VGG-16

Görsel Geometri Grubu (Visual Geometry Group – VGG) ya da VGGNet, sınıflandırma ve yerelleştirme için tasarlanmış özel bir evrişimsel ağ mimarisidir. Derin bir sinir ağı olarak geliştirilen VGG mimarisi, çığır açan nesne tanıma modellerinin temeli sayılarak en popüler görüntü tanıma mimarilerinden biri olarak ün kazanmıştır (Sitaula ve Hossain, 2021). VGG mimarilerinden olan VGG-16’daki 16 (bkz., Şekil 24) ve VGG-19’daki 19 sayıları, katlamalı katmanlardan oluşan katman sayısını ifade etmektedir.



Şekil 3.15. VGG-16 Mimarisi

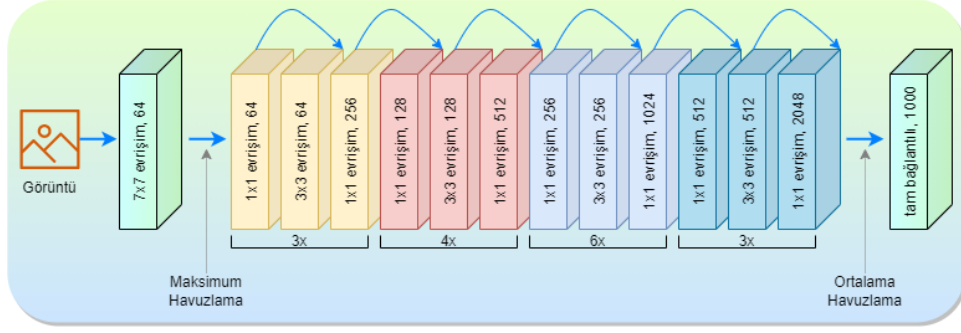
VGG16 modeli, Oxford Üniversitesi'nden A. Zisserman ve K. Simonyan tarafından önerilen bir evrişimsel sinir ağı modelidir (Guan vd., 2019). Ayrıca The ImageNet Large Scale Visual Recognition Challenge (ILSVRC)-2014'e sunulan en popüler modellerden biriydi. Büyük çekirdek boyutundaki filtreleri birbiri ardına birkaç 3x3 çekirdek boyutunda filtreyle değiştirerek AlexNet üzerinde önemli iyileştirmeler yapmaktadır (Tammima, 2019).

VGGNet-16 görüntüleri hayvanlar, kalem, defter vb. dahil olmak üzere 1000 nesne kategorisinde sınıflandırılabilir. VGG-16, 13 evrişim katmandan ve üç tam bağlantılı katmandan oluşur. Ayrıca modelin görüntü giriş boyutu 224'e 224'tür ve her bloğun 2D evrişim ayrıca yüksekliği ve genişliği azaltan max pooling katmanlarından oluştuğu bloklardan meydana gelir. Kullanabilecek filtre sayısı önce 128'e, sonra 256'ya iki katına çıkarılabilecek 64 civarında filtre kullanılmaktadır. Son katmanlarda 512 filtre kullanılabilir.

3.11.1.3. ResNet-50

Artık ağ anlamına gelen ResNet, ilk olarak Kaiming He, Xiangyu Zhang, Shaoqing Ren ve Jian Sun tarafından 2015 yılında tanıtılan yenilikçi bir sinir ağıdır (Xiao vd., 2021). ResNet, bir ağ oluşturmak için artık blokları üst üste istifleyen türden bir yapay sinir ağıdır. ResNet'in aynı konsept üzerinde çalışan ancak farklı sayıda katmana sahip birçok çeşidi vardır. Görüntü işleme ve bilgisayarla görme alanlarında derin evrişimli sinir ağlarıyla çalışırken en doğru sonuçları elde etmek için daha fazla katman ile çalışmak karmaşık sorunların daha verimli bir şekilde çözülmesine yardımcı olmaktadır (Theckedath ve Sedamkar, 2020). Fakat bu durum, modelin özelliklerini zenginleştirebilirken, daha derin bir ağ bozulma sorununu oluşturabilir. Kısaca, katman sayısı arttıkça sinir ağının doğruluk seviyeleri doygun hale gelip bir noktadan sonra yavaş yavaş düşebilir. Böylece, modelin performansı eğitim ve test verileri üzerinde bozulmaktadır. Bu bozulma, aşırı yığılmanın bir sonucu değildir. Bunun yerine, ağın başlatılmasından, optimizasyon işlevinden veya daha da önemlisi, yok olma veya gradyanların patlaması sorunundan kaynaklanabilir (Mukti ve Biswas, 2019).

ResNet, tam da bu sorunun üstesinden gelmek amacıyla oluşturulmuş olup, modellerin doğruluğunu artırmak için artık bloklardan yararlanmaktadır. Artık blokların merkezinde yer alan "bağlantıyı atla" işlemleri düşüşün geçmesi için alternatif bir kısayol ayarlayarak düşme sorununu hafifletirler ve modelin kimlik fonksiyonlarını öğrenmesini sağlarlar (Chu vd., 2020). Bu durum, modelin üst katmanlarının alt katmanlardan daha kötü performans sergilemesini engeller. Artık bloklar, katmanların kimlik fonksiyonlarını öğrenmesini önemli ölçüde kolaylaştırır. Sonuç olarak ResNet, hata yüzdesini en aza indirirken daha fazla sinir katmanı ile derin sinir ağlarının verimliliğini artırır. Başka bir deyişle, atlama bağlantıları, önceki katmanlardan gelen çıktıları yığın katmanlarının çıktılarına ekleyerek, daha önce mümkün olandan çok daha fazla derin ağın eğitilmesini mümkün kılar (Tian ve Chen, 2019).



Şekil 3.16. ResNet-50 Mimarisi

ResNet-50 modeli (bkz., Şekil 25), 50 katman derinliğine sahip bir evrişimsel sinir ağıdır. Bu model, ILSVRC 2015 sınıflandırma yarışmasında sadece %3,57'lik bir hata ile en üst sırayı alıp, 2015 ILSVRC & COCO yarışmalarında ImageNet tespiti, ImageNet lokalizasyonu, COCO tespiti ve COCO segmentasyonunda da birinci olmuştur (Walvekar ve Shinde, 2020). ResNet50'nin mimarisi 4 aşamaya sahiptir: giriş görüntüsü boyutları, yükseklik ve genişlik 32'nin, kanal genişliği ise 3'un katları olarak alınabilir. Açıklama amacıyla, girdi boyutu $224 \times 224 \times 3$ olarak ele alınacaktır. Her ResNet mimarisi, sırasıyla 7×7 ve 3×3 çekirdek boyutlarını kullanarak ilk evrişim ve maksimum havuzlamayı gerçekleştirir. Daha sonra ağın 1. aşaması başlar ve her biri 3 katman içeren 3 artık bloğa sahiptir. Aşama 1 bloğunun tüm 3 katmanında evrişim işlemini gerçekleştirmek için kullanılan çekirdeklerin boyutları sırasıyla 64, 64 ve 128'dir. Son olarak, ağda bir ortalama havuzlama katmanı ve ardından ImageNet sınıf çıktısını sağlayan tam bağlantılı 1000 nörona sahip bir katman bulunur.

3.11.1.4. ResNet-152

ResNet-152, temel olarak ResNet-50 ile aynı yapıda olup katman sayısı 152'dir. ResNet-50 mimarisine daha fazla 3 katmanlı bloklar eklenerek oluşturulmuştur (Bharati vd., 2021).

3.12. Cinsiyet ve Yaş Analizi

Bu bölümde, çalışmamızdaki cinsiyet ve yaş analizi kısmıyla ilgili model ve veri seti bilgileri bulunmaktadır. Cinsiyet ve yaş analizi için ayrı ayrı iki adet özelleştirilmiş model kullanılmıştır.

3.12.1. Yaş Analizi Özelleştirilmiş Modeli

Bu model, 4 katlamalı katmandan sonra düzleştirme işlemini takip eden bir tam bağlantılı

katmandan oluşmaktadır. Her bir katmandan sonra “Maxpooling” fonksiyonu kullanmış olup, aktivasyon fonksiyonu olarak “ReLU” ve “Lineer” kullanılmıştır. İyileştirici olarak ise “Adam” tercih edilmiştir.

3.12.2. Cinsiyet Analizi Özelleştirilmiş Modeli

Cinsiyet analizi özelleştirilmiş modeli 5 katlamalı katmandan oluşup düzleştirme işleminden sonra bir tam bağlantılı katman ile meydana gelmektedir. Her bir katmandan sonra “Maxpooling” fonksiyonu kullanmıştır ve aktivasyon fonksiyonu olarak “ReLU” ve “Sigmoid” tercih edilmiştir. Bu modelde iyileştirici fonksiyon olarak “Adam” kullanılmıştır.

3.13. Yüz Duygu Durumlarından Stres Analizi

Stres, vücudun dış ve iç uyaranlara karşı doğal tepkisidir. Doğal bir şey olmasına rağmen, stres faktörlerine uzun süre maruz kalmak ciddi sağlık sorunlarına katkıda bulunabilir (Liguori vd., 2018). Bu tepkiler sadece fizyolojik olarak değil, aynı zamanda psikolojik olarak da yansır, duygulara ve yüz ifadelerine dönüşür (Pascoe vd., 2020). Yüz ifadelerini sınıflandırabilen bir evrişimsel sinir ağı ve bu modeli kullanıcının yüzünün gerçek zamanlı görüntülerini sınıflandırmak ve böylece stres belirtilerinin varlığını değerlendirmek için analiz modeli oluşturulmaktadır. Sınıflandırma modelinin oluşturulması için ince ayar ile birlikte transfer öğrenme kullanılmıştır. Bu şekilde, eldeki sorunu çözmek için önceden eğitilmiş VGG16 veya ResNet gibi ağlardan yararlanılmaktadır. Transfer öğrenme süreci için ise VGG16'nın evrişimsel katmana dayalı bir sınıflandırıcı ile birlikte stresli duyguları sınıflandırmada kullanılmak için oluşturulmaktadır. Bu oluşturulan model ile duygu durumlarına göre insanları stresli veya stressiz olarak iki farklı sınıf içerisinde değerlendirmeye almaktayız.

4. DENEYSEL KURULUM VE SONUÇLAR

4.1. Duygu Durumu Analizi İçin Kullanılan Veri Kümeleri

Bu tezin “Duygu Durumu” bölümü için aşağıda gösterildiği gibi 2 veri seti kullanılmıştır:

4.1.1. FER2013 Veri Kümesi

Genel olarak yüz tanıma sistemlerinde kullanılan, 2013 Yüz Duygu/İfade Tanıma (Facial Emotion Recognition – FER) veri seti olarak da bilinen FER2013, 2013 yılında Uluslararası Makine Öğrenimi Konferansı'nda (ICML) Pierre-Luc Carrier ve Aaron Courvill tarafından tanıtılan ve Kaggle tarafından sağlanan bir veri setidir (Goodfellow, Erhan, vd., 2015). Bu veri seti, CSV dosyası formatında olup yüz duygu kategorilerine göre sınıflandırılmış 7 farklı sınıftan oluşan (0=Kızgın (anger), 1=İğrenme (disgust), 2=Korku (fear), 3=Mutlu (Happiness), 4=Üzgün (Sadness), 5=Şaşkın (Surprise), 6=Nötr (Neutral)) gri tonlamalı ve her görüntü için 48-piksel x 48-piksel boyutunda olan 35.887 farklı görüntüden oluşmaktadır ve Şekil 26'da kısmi olarak örneklem gösterilmektedir. Bu veri setinde yüzler otomatik olarak kaydedilir, böylece yüz yaklaşık olarak ortalanır ve hemen hemen her görüntüde aynı yeri kaplamaktadır. Eğitim seti 28.709 örnekten, genel test seti 3.589 örnekten oluşmaktadır.



Şekil 4.1. Duygu Durumu Tanıma (Facial Emotion Recognition – FER) Veri Kümesi

4.1.2. FER+ Veri Kümesi

FER+ veri seti, dizin ve satırlar yönünden FER2013 veri kümesiyle tamamen aynıdır. Bunlardan farklı olarak etiketleme açısından hor görme, NF (yüz değil) ve bilinmeyen olmak üzere 3 ek duygu etiketi sunmaktadır ve FER2013 veri setinin bir uzantısı görevi görmektedir. Ayrıca, FER+'da her görüntü, orijinal FER etiketlerinden daha iyi kalitede

temel gerçeği sağlayan 10 kitle kaynaklı etiketleyici tarafından etiketlenmiştir. Her görüntü için 10 etiketleyiciye sahip olmak, araştırmacıların yüz başına bir duygu olasılık dağılımını tahmin etmelerini sağlar ve Şekil 27’de kısmi olarak örneklem gösterilmektedir. Böylece, FER2013 veri setindeki gibi tek etiketli çıktı yerine istatistiksel dağılımlar veya çok etiketli çıktılardan sonuçlar elde edilir.



Şekil 4.2. Duygu Durumu Tanıma+ (Facial Emotion Recognition – FER+) Veri Kümesi

4.1.3. Cinsiyet ve Yaş Analizi İçin Kullanılan Veri Seti

Cinsiyet ve yaş belirlemesi yapabilmek için kullanılan, UTKFace veri seti, 20.000'den fazla yüz görüntüsünden oluşan uzun bir yaş aralığına sahip (0 ila 116 yaş aralığında) bir yüz veri setidir ve Şekil 28’de kısmi olarak örneklem gösterilmektedir. Bu veri seti, yaş, cinsiyet ve etnik köken etiketlerinden oluşmaktadır. Görüntüler, poz, yüz ifadesi, aydınlatma, oklüzyon, çözünürlük vb. konularda çok çeşitlilik gösterir ve uygun şekilde hizalanmış ve kırpılmış yüzleri kapsar. Veri seti yüz algılama, yaş tahmini, yaş ilerlemesi/gerilemesi, önemli nokta lokalizasyonu vb. gibi farklı görevlerde kullanılabilir (Das, Dantcheva ve Bremond, 2019).

4.2. Veri Kümeleri Üzerinde Eğitim ve Test Modelleri

Duygu durumu, cinsiyet, yaş ve stres testleri için önceki bölümde anlatmış olduğumuz modelleri Python 3.9 programlama dili kullanılarak simülasyonları i7 2.2GHz işlemcili, 16 GB RAM ve MacOS işletim sistemine sahip bilgisayar sisteminde gerçekleştirildi.

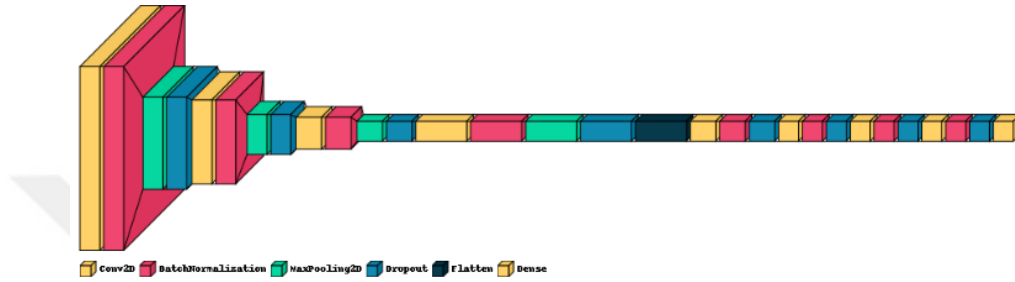


Şekil 4.3. UTKFace Veri Seti

Tüm modellerin çalıştırılmasından önce gereksinim duyulan Keras, Tensorflow, Opencv, Pandas, Numpy, ImageDataGenerator kütüphaneleri sisteme eklenerek görüntüler üzerinde ön işleme adımları tamamlanmaktadır (Github ve Group, 2017). ImageDataGenerator, daha sonra yerleştirmek üzere eğitim ve doğrulama verileri arasında bölme yapmak için çok kullanışlıdır. Ek olarak, yerleştirme aşamasında verilerin gerçek zamanlı olarak artırılmasına olanak tanır. Burada, veri artırma için birkaç hiperparametreyle birlikte, eğitim ve doğrulama verilerinin 75 ve 25'lik bir oranla dağılımını belirlemekteyiz: Ayarlanabilir parlaklık aralığı, 30 derece sağa ve sola dönüş, 0,2 faktörü ile görüntülerin kırılması, Kanal değerlerinin kaydırılması, Görüntü kesme, Görüntü yakınlaştırma. Burada örneğin veri miktarına göre ayarlanabilen toplu iş boyutu ve numuneye dayalı olarak `batch_size = 128` ve `steps_per_epoch`'u belirlemekteyiz. “steps_per_epoch” her bir örneklem miktarının `batch_size`'a oranına göre tamsayı olacak şekilde belirlenmektedir. Kolayca ayarlanabilmeleri için özel modellerin bir listesini işlevsel biçimde tanımlamaktayız. Ek olarak, oluşturulan modellere ek katmanlar eklemek veya genişletmek gerekirse, onu ayrı ayrı kod olarak da ekleyebilmekteyiz. Ancak, her yeni ekleme yapılan katman için sistemin yeniden derlenmesi gerekmektedir ve bu süre kullanılan veri kümeleri için oldukça yüksektir.

Bu bölümde, özel bir model CNN modeli ve ayrıca Görsel Geometri Grubundan (VGG) VGG16, Resnet50 ve Resnet152 gibi bazı modelleri sistemin çalışma yapısını anlayabilmek için hazırlanmaktadır. Bu sınıflandırma görevi için 8 sınıflık bir çıktıya ihtiyaç

duyulmaktadır. Önceden eğitilmiş modeller, mimarilerinde varsayılan olarak farklı çıktı sınıflarına sahiptir. Varsayılan olarak, VGG16, 2014'teki ImageNet Büyük Ölçekli Görsel Tanıma Mücadelesine (ILSVRC2014) dayalı olarak 1000 sınıf verir. Önceden eğitilmiş modelleri sınıflandırma oluşturulacak modele entegre edebilmek için son katmanın değiştirilerek tüm yoğun bağlı katmanlara ve çıktı katmanına bağlanması gerekmektedir. Tüm modellerin en iyileştirilmesi için Adam yöntemi doğruluk oranlarını belirlemede kullanılmaktadır.



(a) CNN Modeli

```

model = Sequential()

# 1st Convolution Layer
model.add(Conv2D(64,(3,3), input_shape=(48, 48,3)))
model.add(BatchNormalization())
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))

# 2nd Convolution Layer
model.add(Conv2D(128,kernel_size=(3,3), activation='relu'))
model.add(BatchNormalization())
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))

# 3rd Convolution Layer
model.add(Conv2D(256,kernel_size=(3,3), activation='relu'))
model.add(BatchNormalization())
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))

# 4th Convolution Layer
model.add(Conv2D(512,kernel_size=(3,3), activation='relu'))
model.add(BatchNormalization())
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))

# Flattening
model.add(Flatten())

# Fully connected layer 1st layer
model.add(Dense(256, activation='relu'))
model.add(BatchNormalization())
model.add(Dropout(0.25))

# Fully connected layer 2nd layer
model.add(Dense(128, activation='relu'))
model.add(BatchNormalization())
model.add(Dropout(0.25))

# Fully connected layer 3rd layer
model.add(Dense(128, activation='relu'))
model.add(BatchNormalization())
model.add(Dropout(0.25))

# Fully connected layer 4th layer
model.add(Dense(64, activation='relu'))
model.add(BatchNormalization())
model.add(Dropout(0.25))

model.add(Dense(num_classes, activation='softmax'))

model.compile(loss='categorical_crossentropy',
              optimizer='adam',
              metrics=['accuracy'])

```

(b) CNN Python Kod Bloğu

Şekil 4.4. Özelleştirilmiş CNN Modeli

4.2.1. Uygulanan Özelleştirilmiş CNN Modeli

Özelleştirilmiş CNN modelini oluştururken, düzleştirmeden ve 4 ek tam bağlantılı katman eklemekten önce 4 evrişim katmanından oluşan temel bir model oluşturulmaktadır. Her katman, aşırı uyum göstermeyi (overfitting) azaltmak ve modelin zaman içinde ortalamasını almak için bir bırakma veya azaltma (dropout) işlevine sahiptir. Bu modelin oluşturulması, 1.743.560 eğitilebilir parametre ile sonuçlanabilmektedir.

CNN modelinde her evrişim katmanının aktivasyon işlevi olarak ReLU aktivasyon işlemi, çekirdek boyutu 3x3, maksimum havuzlama ölçeklendirmesi 2x2, azaltma oranı 0,25, ve sırasıyla katmanların çekirdek boyutları 64, 128, 256 ve 512 olarak Şekil 29'da gösterildiği gibi belirlenmektedir.

4.2.2. Uygulanan VGG-16 Modeli

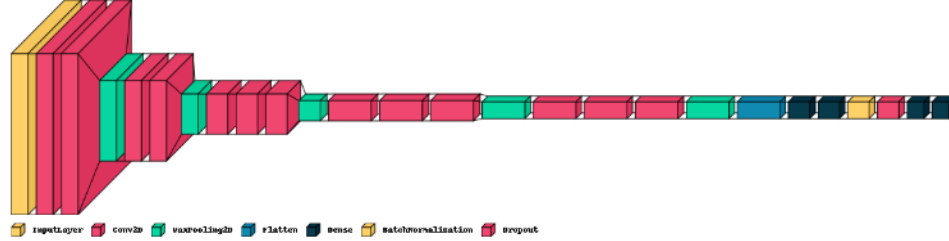
Sistemin çalışabilirliğini anlayabilmek için oluşturulan VGG-16 modeli, görüntü sınıflandırma amaçları için en popüler önceden eğitilmiş modellerden biridir. Oluşturulan VGG-16 modeli, 5 evrişim katmanından, ardından 3 tam bağlantılı katmandan ve ardından 1 çıktı katmanından oluşur. Modelin yüz analizi için gereksinim duyduğu çalışma mimarisini oluşturmak için belirttiğimiz son katmanı kaldırılarak bağlı katmanları, sınıflandırma sayısına uygun olacak şekilde çıktı katmanıyla Şekil 30'da gösterildiği gibi birleştirilmektedir. VGG-16 modeli, 14.946.248 eğitilebilir parametreyle sonuçlanır.

VGG-16 modelinde, önceden ImageNet veri kümesi ile eğitilmiş ağırlıklar alınarak her bir inputun boyutu 48x48x3 olarak belirlenmektedir. Belirlenen her input için tam yoğun bağlantılı katmanlarda çekirdek boyutu 256 ve ReLU aktivasyon işlemi uygulanmaktadır. Ayrıca, toplu halde verilerin hesaplanmasında normalizasyon uygulanarak aralıkların belirli sınırlar içerisinde tutulması sağlanmaktadır. Girdiler üzerinde katmanlardaki azaltma oranı ise 0,4 olarak belirlenmektedir ve en son bağlanan tam bağlantılı katmanda 128 çekirdek boyutu olarak belirlenmiştir. Çıktı katmanına bağlanan katmanda ise aktivasyon işlevi olarak Softmax kullanılmaktadır.

4.2.3. Uygulanan ResNet-50 ve ResNet-152 Modelleri

Sistemde verilerin analizi ve test edilmesi için kurgulanan ResNet-50 ve ResNet-152 modelleri 50 ve 152 katmandan oluşur ve ayrıca ImageNet veritabanındaki 1 milyondan fazla görüntü üzerinde ağırlıkların başta belirlenmesi için eğitilir. ResNet-152 modeli daha

derin öğrenme gereksinimi duyulan ağlarda iyi sonuçlara erişebilmek için modellenmektedir ve Şekil 31’de gösterilmektedir.



(a) VGG-16 Modeli

```
vgg_model = VGG16(weights='imagenet', include_top=False, input_shape=(48,48,3))
add_model = Flatten()(vgg_model.layers[-1].output)
add_model = Dense(256, activation='relu')(add_model)
add_model = Dense(256, activation='relu')(add_model)
add_model = BatchNormalization()(add_model)
add_model = Dropout(0.4)(add_model)
add_model = Dense(128, activation='relu')(add_model)

add_model = Dense(num_classes, activation='softmax')(add_model)

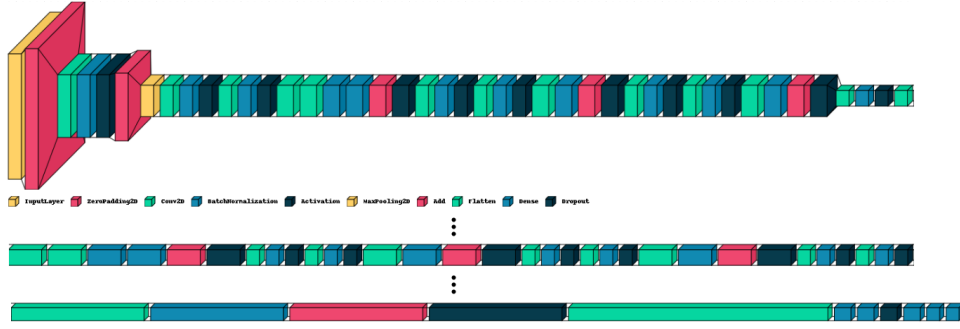
final_model = Model(inputs=vgg_model.inputs, outputs=add_model)

final_model.compile(loss='categorical_crossentropy',
                    optimizer='adam',
                    metrics=['accuracy'])
```

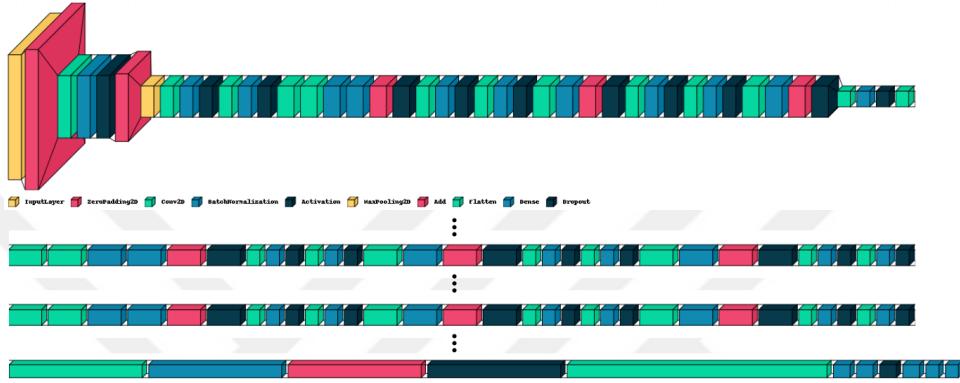
(b) VGG-16 Python Kod Bloğu

Şekil 4.5. VGG-16 Modeli

Sistemde modellerin katmanlarında tam bağlantılı katmanların çekirdek boyutları 256, 256, ve 128 olarak belirlenmiş ve tam bağlantılı katmanlarda çıktı katmanından önceki katman haricindekilerde aktivasyon işlemi olarak ReLU tanımlanmaktadır. Çıktı katmanından önceki en son kullanılan yoğun bağlantılı katmanda ise Softmax aktivasyon işlemi kullanılmaktadır. Veriler üzerinde azaltma katmanı operasyonunda 0,3'lük oran kullanılmaktadır.



(a) ResNet-50 Modeli



(b) ResNet-152 Modeli

```
res_model = ResNet50(weights='imagenet', include_top=False, input_shape=(48,48,3))
add_model = Flatten()(res_model.layers[-1].output)
add_model = Dense(256, activation='relu')(add_model)
add_model = BatchNormalization()(add_model)
add_model = Dropout(0.3)(add_model)
add_model = Dense(256, activation='relu')(add_model)
add_model = Dense(128, activation='relu')(add_model)
add_model = Dense(num_classes, activation='softmax')(add_model)

final_model = Model(inputs=res_model.inputs, outputs=add_model)

final_model.compile(loss='categorical_crossentropy',
                    optimizer='adam',
                    metrics=['accuracy'])
```

(c) ResNet-50 Python Kod Bloğu

```
res_model = ResNet152(weights='imagenet', include_top=False, input_shape=(48,48,3))
add_model = Flatten()(res_model.layers[-1].output)
add_model = Dense(256, activation='relu')(add_model)
add_model = BatchNormalization()(add_model)
add_model = Dropout(0.3)(add_model)
add_model = Dense(256, activation='relu')(add_model)
add_model = Dense(128, activation='relu')(add_model)
add_model = Dense(num_classes, activation='softmax')(add_model)

final_model = Model(inputs=res_model.inputs, outputs=add_model)

final_model.compile(loss='categorical_crossentropy',
                    optimizer='adam',
                    metrics=['accuracy'])
```

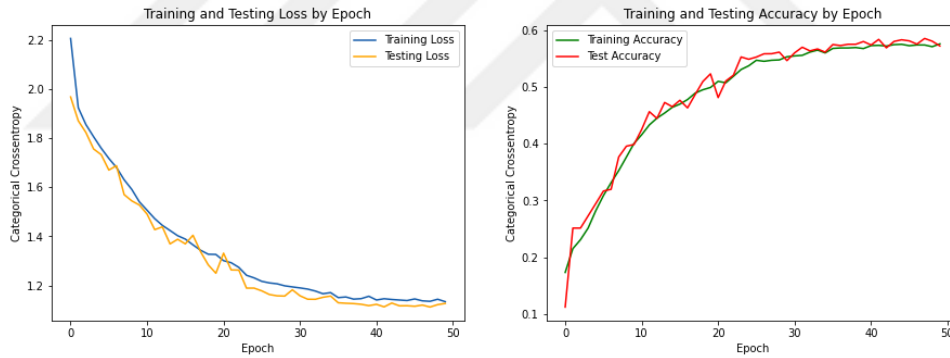
(d) ResNet-152 Python Kod Bloğu

Şekil 4.6. ResNet-50 ve ResNet-152 Modelleri

Tüm modelleri sırasıyla işleme koyarken geriye yönelik arama yapılarını da aktif ederek farklı doğrulama noktalarında her model için modellerin onaylanma doğruluklarını bulmamıza yardım etmektedir. Her model için, 90 ila 150 arasında değişen çeşitli devir (epoch) sayıları ile tekrar tekrar çalıştırma yapılarak genellikle 20'inci veya 25'inci epoch'ta geliştirilmiş performans platosuna (modeline) sahip olduğu bulunmaktadır. Bu sebepten, her bir modelin çalışma devirlerini veya dönemlerini 50 ile sınırlayarak sonuçların daha kısa sürede alınması sağlanmaktadır. Batch_size ise 128'e ayarlanmıştır ancak hesaplama gücüne bağlı olarak farklı değerlerde de ayarlanabilir. Benzer şekilde, step_per_epoch da daha önce belirtmiş olduğumuz batch_size oranına göre aşağıdaki formül ile ayarlanabilir:

$$steps_per_epoch = \left\lceil \frac{Toplam\ Örneklem\ Sayısı}{batch_size} \right\rceil$$

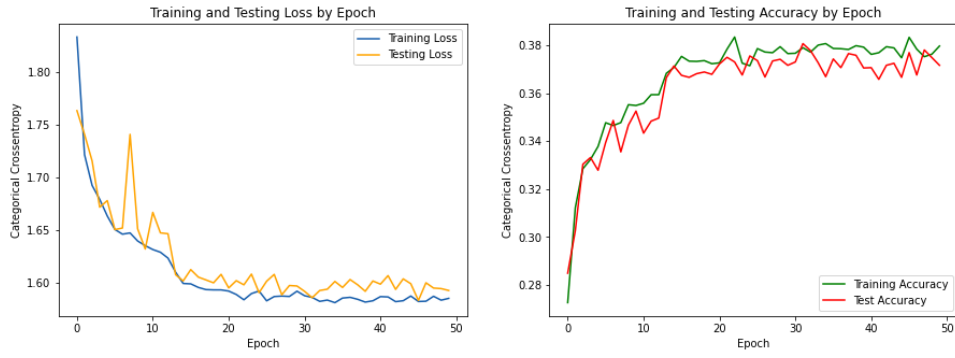
Oluşturulan CNN özel modeli, simülasyon sonunda Şekil 32'de de gösterildiği gibi, 0,74'lük bir kayıpla, doğrulama setine karşı %74'lük bir doğruluk puanı elde etmektedir. Buradan çıkarılacak sonuç olarak, eğitim ve doğrulama puanları arasında çok fazla fark olmadığı görülmektedir. Bu model başlangıç için oldukça iyi bir performans sergilemektedir ve 20'inci epoch'tan sonra kayıpların minimize olduğu gözlemlenmektedir.



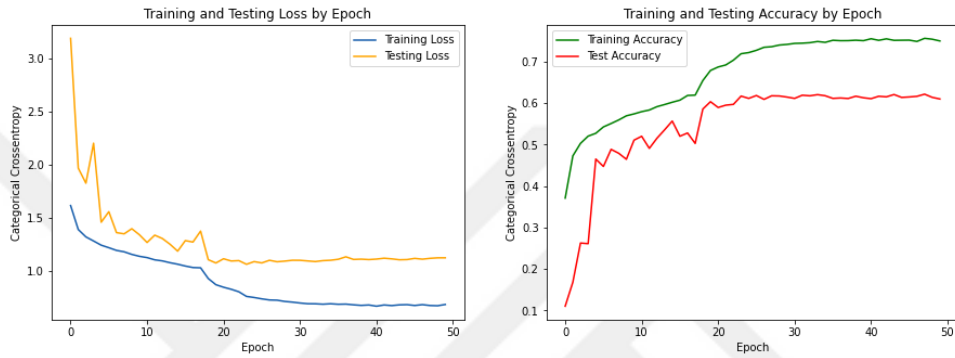
Şekil 4.7. Özelleştirilmiş CNN Modeli Sonuçları

Şaşırtıcı bir şekilde, VGG16 modeli Şekil 33'te belirtildiği üzere özel CNN modelinden biraz daha düşük performans göstermektedir ve 0,81 kayıpla %72'lik bir doğrulama oranına ulaşmaktadır. 30'uncu epoch'tan sonra kayıp ve doğruluk analizi birbirleriyle uyumlu olarak ilerleme göstermektedir.

ResNet-50 modeli, 0,76'lık bir kayıpla %78 ile şimdiye kadarki en yüksek doğrulama doğruluğunu elde etmektedir (bkz. Şekil 34). Bununla birlikte, eğitim ve doğrulama seti arasındaki doğruluk ve kayıptaki fark, biraz fazla uyum olduğunu göstermektedir.

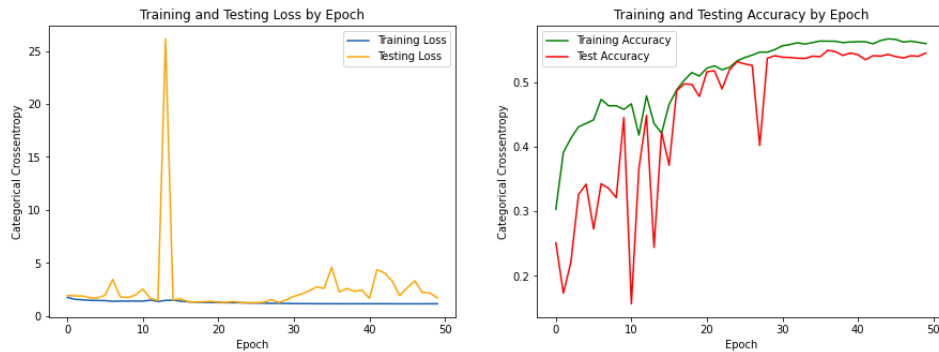


Şekil 4.8. VGG-16 Modeli Sonuçları



Şekil 4.9. ResNet-50 Modeli Sonuçları

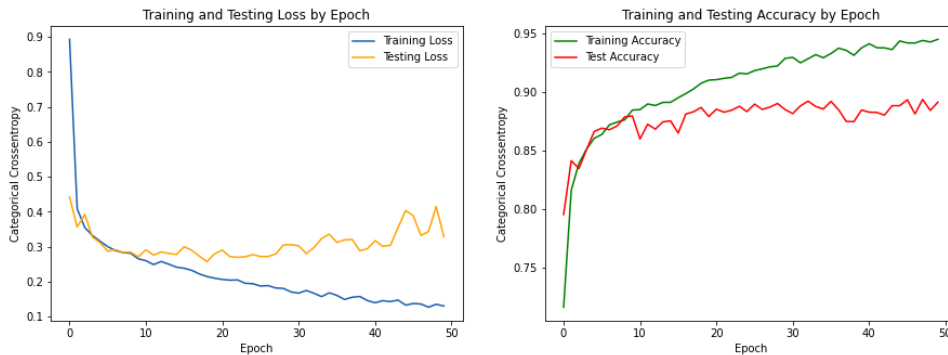
Daha derin veya çoklu katmanların doğruluk puanını iyileştirmeye yardımcı olup olmayacağını görmek için analizimizde Resnet152 modeli garip bir şekilde Resnet50'ye göre daha iyi bir doğruluğa sahip olamamaktadır (bkz. Şekil 35). Oluşturulan çok katmanlı model, 0,81'lik bir kayıpla %71'lik bir doğrulama doğruluğuna ulaşmaktadır. Doğrulama setine karşı kayıp ve doğrulukta da ani artışlar bulunmaktadır ve daha fazla katman ile birlikte çalışma zamanları da oldukça artmaktadır.



Şekil 4.10. ResNet-152 Modeli Sonuçları

Modellerimizi tekrar ve yeniden eğitmek için gereken süreyi azaltmak için, önceden eğitilmiş ve kontrol noktaları sayesinde dışa aktarılmış kayıtlı '.h5' modellerini kullanmaktayız. Bazı duyguların analizinde farklılıklar meydana gelebiliyor ve bunlara örnek olarak öfke duygusunun ilk yanlış sınıflandırılan görüntüsünü alarak, yanlış sınıflandırılan etiketlerin çoğunun mutluluk olduğuna dikkat etmekteyiz. Diğer modeller bu görüntüyü nötr, hüznün veya sürpriz olarak sınıflandırmaktadır. Öfke için ikinci yanlış sınıflandırılmış etiket ise nötr ya da mutluluk olarak gözlemlenmektedir. Bunun nedeni olarak, modellerin öfkeyi mutlulukla karıştırabilecek yüz algısından gelen dişlerin açığa çıkması olabilir. Aşağılama duygusu, nötr bir duyguya benzediği için anlaşılması çok zor bir duygudur. Bu duygu için modellerimizde nötrün yanlış sınıflandırılmasına büyük olasılıkla herhangi bir gülümseme veya üzüntü olmadan dudaklardan kaynaklanan bir analiz yapılmasından dolayı denk gelmekteyiz. Yanlış sınıflandırmalar, görüntünün çoğunlukla mutluluk veya öfke olarak yanlış sınıflandırılmasından kaynaklanmaktadır. Bunun nedeni, nötr yüz yapılarının bunun yerine bir duygu olarak algılanması olabilir. Bunlardan dolayı, stres analizlerinde de bazı hatalı sonuçlar meydana gelmektedir.

Cinsiyet ve yaş tespiti için insanların yüzlerinin tespiti için kullanılan özelleştirilmiş CNN modelinin 0,69'luk kayıpla %86'luk doğrulama oranına sahip olduğu tespit edilmektedir (bkz. Şekil 36). Bu cinsiyet ve yaş tespiti için oluşturulan modelin yeterli ölçüde analizi doğru yapabildiğini göstermektedir. Bu ölçümler modellerin çalışma zamanından dolayı gelecek çalışmalarda farklı modeller ile de desteklenerek analiz edilmeye açıktırlar.



Şekil 4.11. Özel CNN Modeli Cinsiyet ve Yaş Belirleme Sonuçları

5. TARTIŞMA VE GELECEK ÇALIŞMALAR

Bu çalışmada, canlı video görüntüsünde bulunan yüzlerin duygu durumları, yaşları, cinsiyetleri ve duygu durumlarındaki sınıflandırmalara bağlı olarak stres oranları hesaplanmıştır. Stres oranları, “korkmuş” ve “kızgın” sınıflandırmalarına göre ve insanların kaşlarının durumlarının normalize edilmiş hesaplamalarındaki oranları ile değerlendirilmiştir. Strese maruz kalmanın duygu tanıma geliştirme veya duyguların stres analizi üzerindeki etkisini araştırmayı amaçlamıştır. Spesifik olarak, yapılan çalışma genel duygu tanıma ile akut psikososyal strese maruz kalmanın (Gamsakhurdashvili vd., 2021) etkileyip etkilemediğini araştırmıştır ve ayrıca bu geliştirme modelinin duygular arasında birleşip birleşmediğini veya bireye yönelik potansiyel tehdit sinyalleriyle ilişkili yüz ifadelerine farklılaşıp farklılaşmadığını araştırmıştır.

Fizyolojik sinyallerin ön işleminin karmaşıklığı, fizyolojik sinyaller aracılığıyla duygu tespiti için büyük bir zorluktur. Bu oldukça zor bir araştırma alanıdır. Şimdiye kadar, yalnızca yedi temel duygu başarıyla tanımlandı. Yediden fazla duyguyu tanımlamak için farklı araştırmalar yapılmalıdır. Literatürde elektrokardiyogram sinyalleri aracılığıyla duygu tespiti ve kas hareket sinyallerini kullanan cilt sıcaklıkları ve elektromiyografi sinyalleri gibi özellikler hala ortaya çıkmaktadır (Saxena vd., 2020). Bu yöntemlerin gücünü kontrol etmek için detaylı araştırmalar yapılabilir. Ancak bu çalışmalar bireyler üzerinde anlık ölçümler ve oldukça yüksek maliyetler ortaya çıkarmaktadır. Hala doğru kombine ve hibrit modellerin yetersizliği mevcuttur. İnsan duygularının daha iyi tahmin edilmesi için daha etkili hibrit ve kombine yöntemler geliştirilmelidir (Katsigiannis ve Ramzan, 2017).

Yüz duygu ve stres tanımlama çalışmaları geleneksel mimariden ayrılarak oluşturulan sistemler aracılığıyla kendi kendine öğrenme yetkinliğine de sahip olmaktadır. Bu sayede, yazılımsal olarak oluşturulmuş modeller, yüz, stres, yaş ve cinsiyet tespitleri ile Parkinson hastalığı, bipolar bozukluğu, öfke patlamaları, şizofreni, klinik depresyon hastalıklarının önceden belirlenebilmesi ve tedavi sürecinin ilerleme tespiti içinde kullanılabilecektir (Argaud vd., 2018).

Bir başka ilginç nokta, araştırma gruplarının belirlenen düzensizlik durumlarının karşılaştırmalarını mümkün kılmak için test tasarımlarının standartlaştırılması esastır. Ayrıca, yüz etkisi testleri üzerinde doğru bir şekilde performans göstermek, dikkat, yürütme

ve duygusal yeteneklerin bir derlemesini gerektirir ve bu da temel eksikliğin aslında yüz ifadelerinin işlenmesi olup olmadığını belirlemeyi zorlaştırır (Gamsakhurdashvili vd., 2021). Ayrıca, bazı araştırmalar küçük örneklemeleri değerlendirdiğinde, bu da sonuçların uygulama yapılacak alanlardaki genelleştirilmesini zorlaştırmıştır.

Buna benzer olarak, eğitim alanında da insanların duygularının büyük önem taşıdığı gözlemlenmiştir. Bireylerin toplumun içerisinde yaşadıkları bölgelere, etnik kökenlerine, kültürlerine ve kullandıkları dil farklılıkları gösterse de duygunun küresel olarak bir değerlendirme kriteri olduğu kabul edilmektedir. Yapılan araştırmalarda, devlet üniversitesinde eğitim gören öğrencilerin derslerindeki duygu analizleri ve buna bağlı olarak ders etkinliklerinin üç ayrı bölümde incelenmesi uygun görülmüştür. Duygulardaki değişimlerin derslerin ilk bölümünde daha çok öfke, korku ve şaşkınlık olarak yüksek belirlendiği, mutluluk, üzüntü ve iğrenme duygularının da azaldığı gözlemlenmiştir. Genel bir bulgu analizi yapıldığında eğitim kalitesinin ders bölümlerindeki durumlara göre etkileşimli olarak desteklenmesiyle artacağı öngörülmektedir (Tonguç ve Ozaydın Ozkara, 2020). Bununla birlikte, yapılan bu çalışmanın eğitim alanında uygulanması ile birlikte öğrencilerin anlık stres analizi de yapılarak derslerdeki algının artırılması da mümkün olabilecektir.

Tüm bunları bilgisayar ortamında gerçekleştirebilmek amacıyla, yüz duygu tanıma alanı için en çok kullanılan ve en iyi sonuçları veren Evrişimli Sinir Ağları (CNN) mimarisi çalışmada kullanılmış olup VGG16, ResNET50 ve ResNET152 farklı CNN mimarilerinin yanı sıra özelleştirilmiş bir CNN modeli de kullanılmıştır. Ayrıca modeller içerisinde en iyi sonuçları elde edebilmek amacıyla bazı parametreler değiştirilerek ince ayar yapılmıştır. Veri seti olarak duygu durumu analizi başlığında FER2013 ve FER+ veri setleri, cinsiyet ve yaş kısmı için de UTKFace veri seti uygulanmıştır.

Bu çalışma, modelin farklı performanslarının incelenebilmesi açısından

- başka veri setleri üzerine uygulanabilir,
- duygu durum sınıfları genişletilebilir,
- yaş sınıf aralıkları küçültülerek sayıları artırılabilir,
- stresi belirleyen duygu durumları değiştirilebilir,

- stres oranı yerine yüz duygu durumuna dayalı farklı başlıklara ait oranlar hesaplanabilir,
- farklı modelleme teknikleri uygulanabilir,
- veri seti düzen sayıları değiştirilebilir ve
- farklı durumlara ait duygu durumları belirlenip psikoloji gibi çeşitli alanlara yarar sağlayabilir.

Bu araştırmalar ile birlikte bireylerin canlı kamera görüntüsündeki yüzlerin üzerinde bulunan çerçevelerde belirtilerek tespit edildi. Ayrıca gelecekte yapılacak olan çalışmalarda yaş ve cinsiyete bağlı olarak etnik köken çalışması da yapılması planlanmaktadır.



6. SONUÇ

İnsaların yüzlerinde oluşan duygu, iletişimde başkalarının düşüncelerini veya niyetlerini tahmin etmemizde bizlere oldukça yardımcı olan önemli faktörlerdir. Genel olarak, iletişim için insanların yaygın olarak kullandığı anlamsal yapıları yüz ifadeleri ve ses tonu algılama ile, neşe, üzüntü veya öfke durumlarını çıkartabiliriz. Çalışmamızda belirttiğimiz farklı araştırmalara göre, yaygın olarak insanların birbirleri iletişimlerinde sözsüz bileşenler sözlü ifadelerden daha çok anlam aktarmaktadır. Pek çok sözel olarak ifade edilmeyen bileşenler arasında, yüz ifadeleri duygusal bilgilendirme taşıyarak yaygın bilgi aktarım kanallarından biridir. Bu sebeple, son yıllarda sadece algısal veya bilişsel olarak yapılmış araştırmalar gibi, yüz duygularının araştırması gelişen teknolojik altyapı ile birlikte duyuşsal hesaplama ve bilgisayar animasyonları ile yeni uygulamalar daha çok ilgi görmektedir.

Bu kapsamda, çalışmamızda önerilen birleşik yapıdak oluşturulmuş yüzdeki duygu durumu, cinsiyet, yaş ve stres oran tespit modelimizde ilk olarak verilen görüntü bilgisayarın daha iyi anlayıp işlem yapabilmesi için sayısallaştırıldı, model tarafından kullanılan OpenCV kütüphanesi ile yüzler belirlendi, ilgili belirleyici bölgeler algılanıp her bir sınıflandırma grubu için kritik bölgelerin taranmasıyla birlikte yüzdeki duygu durumu, yaş, cinsiyet ve stres oranları sınıflandırma sonuçları bulundu. Uygulanan modeller, bir görüntüdeki birden fazla yüzü algılayıp her biri için aynı işlemleri gerçekleştirebilmektedir.

Görüntülerdeki yüzler her ne kadar net olsa da ayırt edilmesi insanlar için bile zor olan bazı duygu durumlarında (örneğin, iğrenme ve hor görme vb.), modeller tam olarak karar veremeyip farklı sonuçlar elde edebilmektedir. Bu durum için eğitim verilerinin artırılarak benzer farklı çözümler denenebilir.

Duygu durumu analizi için kullandığımız FER2013 ve FER+ veri setlerindeki görüntüler gri tonlamalı olmasına rağmen canlı yapılan duygu durumu analizinde kameradaki görüntü renklidir ve model yine kolaylıkla analizini doğru şekilde gerçekleştirebilmektedir. Elde edilen sonuçlara göre özel CNN modeli VGG16 mimarisine göre daha iyi bir sonuç vermiştir. Fakat, kullanılan modeller arasında en iyi sonucu veren model ResNet50 olmuştur.

Model, işlevi açısından bu alanda daha önce uygulanmış, incelenmiş ve en iyi sonuçlar alınmış sistemleri barındırdığından elde edilen sonuçlar da bu doğrultuda daha önce yapılan

modellemelerle yarışabilecek cinstendir. Bunlara ek olarak yüz ifadeleriyle bir arada hesaplanan duygu, cinsiyet, yaş ve stres analizi farklı alanlarda kullanım olanaklarıyla uygulandığı disiplinlerde yarar sağlayacaktır.

Ayrıca, bu çalışma günümüzde yaygın olarak kamera sistemlerinin kullanıldığı eğitim, sağlık, iş güvenliği, ticaret, ve suç tespiti gibi farklı alanlarda insanların duygu durumlarını, streslerini, yaş ve cinsiyetlerini hesaplayarak kişiler hakkında anlık olarak öngöründe bulunma fırsatı da oluşturacaktır. Bu sistemlerin birleşik olarak uygulanması, çalışmamızda belirtmiş olduğumuz bazı hastalıkların önceden tespiti veya tedavi süreçlerinin olumlu halde ilerlemesi, eğitim süreçlerindeki motivasyonların duygu analizi ve strese bağlı olarak aktivitelerdeki devamlılığın ve sürdürülebilirliğin sağlanması, vb. gibi uygulamalarda akıllı sistemlerin kullanımında yaygınlaştırıcı örnek olarak görülebilecektir.

KAYNAKLAR

- Aaron, J. ve Chew, T. L. (2021). A guide to accurate reporting in digital image processing – Can anyone reproduce your quantitative analysis? *Journal of Cell Science*, 134(6). doi:10.1242/jcs.254151
- Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Umar, A. M., Linus, O. U., ... Kiru, M. U. (2019). Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition. *IEEE Access*, 7(February 2017), 158820–158846. doi:10.1109/ACCESS.2019.2945545
- Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and cognitive neuroscience reviews*, 1(1), 21–62. doi:10.1177/1534582302001001003
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data* (C. 8). Springer International Publishing. doi:10.1186/s40537-021-00444-8
- Argaud, S., Vérin, M., Sauleau, P. ve Grandjean, D. (2018). Facial emotion recognition in Parkinson's disease: A review and new hypotheses. *Movement Disorders*, 33(4), 554–567. doi:10.1002/mds.27305
- Azcarate, A., Hageloh, F., Sande, K. Van De ve Valenti, R. (2005). Azcarate et al. - 2005 - Automatic facial emotion recognition.pdf, (June).
- Bani, M., Russo, S., Ardenghi, S., Rampoldi, G., Wickline, V., Nowicki, S. ve Strepparava, M. G. (2021). Behind the Mask: Emotion Recognition in Healthcare Students. *Medical Science Educator*, 31(4), 1273–1277. doi:10.1007/s40670-021-01317-8
- Bashar, D. A. (2019). Survey on Evolving Deep Learning Neural Network Architectures. *Journal of Artificial Intelligence and Capsule Networks*, 2019(2), 73–82. doi:10.36548/jaicn.2019.2.003
- Bharati, S., Podder, P., Mondal, M. R. H. ve Prasath, V. B. S. (2021). CO-ResNet: Optimized ResNet model for COVID-19 diagnosis from X-ray images. *International Journal of Hybrid Intelligent Systems*, 17(1–2), 71–85. doi:10.3233/his-210008
- Bory, P. (2019). Deep new: The shifting narratives of artificial intelligence from Deep Blue to AlphaGo. *Convergence*, 25(4), 627–642. doi:10.1177/1354856519829679
- Bull, P. (2001). State of the art: Nonverbal communication. *Psychologist*, 14(12), 644–647.
- Cai, W., Gao, M., Liu, R. ve Mao, J. (2021). MIFAD-Net: Multi-Layer Interactive Feature Fusion Network With Angular Distance Loss for Face Emotion Recognition. *Frontiers in Psychology*, 12(October), 1–11. doi:10.3389/fpsyg.2021.762795
- Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., ... Zdeborová, L. (2019). Machine learning and the physical sciences. *Reviews of Modern Physics*, 91(4), 45002. doi:10.1103/RevModPhys.91.045002
- Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P. ve Joulin, A. (2020). Unsupervised learning of visual features by contrasting cluster assignments. *Advances in Neural Information Processing Systems, 2020-Decem(NeurIPS)*, 1–13.
- Carstensen, L. L., Fung, H. H. ve Charles, S. T. (2003). Socioemotional selectivity theory and the

- regulation of emotion in the second half of life. *Motivation and Emotion*, 27(2), 103–123. doi:10.1023/A:1024569803230
- Chalmers, D. J. ve Chalmers, D. J. (2004). Consciousness & the Philosophers, 1–13.
- Chandrasekhar, S. S. (2020). Domo Arigato, Mister Roboto. *Otolaryngologic Clinics of North America*, 53(6), xvii–xix.
- Chen, Y., Xie, Y., Song, L., Chen, F. ve Tang, T. (2020). A survey of accelerator architectures for deep neural networks. *Engineering*, 6(3), 264–274.
- Cheng, D., Gong, Y., Zhou, S., Wang, J. ve Zheng, N. (2016). Person re-identification by multi-channel parts-based CNN with improved triplet loss function. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem*, 1335–1344. doi:10.1109/CVPR.2016.149
- Chu, Y., Yue, X., Yu, L., Sergei, M. ve Wang, Z. (2020). Automatic Image Captioning Based on ResNet50 and LSTM with Soft Attention. *Wireless Communications and Mobile Computing, 2020*. doi:10.1155/2020/8909458
- Chua, L. O. ve Roska, T. (1993). The CNN paradigm. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 40(3), 147–156.
- Cimpoi, M., Maji, S., Kokkinos, I. ve Vedaldi, A. (2016). Deep Filter Banks for Texture Recognition, Description, and Segmentation. *International Journal of Computer Vision*, 118(1), 65–94. doi:10.1007/s11263-015-0872-3
- Cobbe, K., Klimov, O., Hesse, C., Kim, T. ve Schulman, J. (2019). Quantifying generalization in reinforcement learning. *36th International Conference on Machine Learning, ICML 2019, 2019-June*, 2280–2293.
- Das, A., Dantcheva, A. ve Bremond, F. (2019). Mitigating bias in gender, age and ethnicity classification: A multi-task convolution neural network approach. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11129 LNCS, 573–585. doi:10.1007/978-3-030-11009-3_35
- Dehghan, A., Ortiz, E. G., Shu, G. ve Masood, S. Z. (2017). DAGER: Deep Age, Gender and Emotion Recognition Using Convolutional Neural Network. <http://arxiv.org/abs/1702.04280> adresinden erişildi.
- Domínguez-Jiménez, J. A., Campo-Landines, K. C., Martínez-Santos, J. C., Delahoz, E. J. ve Contreras-Ortiz, S. H. (2020). A machine learning model for emotion recognition from physiological signals. *Biomedical Signal Processing and Control*, 55, 101646. doi:10.1016/j.bspc.2019.101646
- Du, S. S., Lee, J. D., Li, H., Wang, L. ve Zhai, X. (2019). Gradient descent finds global minima of deep neural networks. *36th International Conference on Machine Learning, ICML 2019, 2019-June*, 3003–3048.
- Eckle, K. ve Schmidt-Hieber, J. (2019). A comparison of deep networks with ReLU activation function and linear spline-type methods. *Neural Networks*, 110, 232–242. doi:10.1016/j.neunet.2018.11.005
- Frankle, J., Schwab, D. J. ve Morcos, A. S. (2020). The Early Phase of Neural Network Training, 1–20. <http://arxiv.org/abs/2002.10365> adresinden erişildi.

- Gal, Y. ve Ghahramani, Z. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. *international conference on machine learning* içinde (ss. 1050–1059).
- Gamsakhurdashvili, D., Antov, M. I. ve Stockhorst, U. (2021). Facial Emotion Recognition and Emotional Memory From the Ovarian-Hormone Perspective: A Systematic Review. *Frontiers in Psychology*, 12. doi:10.3389/fpsyg.2021.641250
- Gao, Q., Wang, C. han, Wang, Z., Song, X. lin, Dong, E. zeng ve Song, Y. (2020). EEG based emotion recognition using fusion feature extraction method. *Multimedia Tools and Applications*, 79(37–38), 27057–27074. doi:10.1007/s11042-020-09354-y
- Gavrilescu, M. ve Vizireanu, N. (2019). *Predicting depression, anxiety, and stress levels from videos using the facial action coding system. Sensors (Switzerland)* (C. 19). doi:10.3390/s19173693
- Geetha, M., Latha, R. S., Nivetha, S. K., Hariprasath, S., Gowtham, S. ve Deepak, C. S. (2021). Design of face detection and recognition system to monitor students during online examinations using Machine Learning algorithms. *2021 International Conference on Computer Communication and Informatics, ICCCI 2021*, 27–30. doi:10.1109/ICCCI50826.2021.9402553
- Github, A. ve Group, D. G. (2017). The Keras Blog Building powerful image classification models using very little data, 1–12.
- Gong, Y., Wang, L., Guo, R. ve Lazebnik, S. (2014). Multi-scale orderless pooling of deep convolutional activation features. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8695 LNCS(PART 7), 392–407. doi:10.1007/978-3-319-10584-0_26
- Goodfellow, I. J., Erhan, D., Luc Carrier, P., Courville, A., Mirza, M., Hamner, B., ... Bengio, Y. (2015). Challenges in representation learning: A report on three machine learning contests. *Neural Networks*, 64, 59–63. doi:10.1016/j.neunet.2014.09.005
- Goodfellow, I. J., Shlens, J. ve Szegedy, C. (2015). Explaining and harnessing adversarial examples. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–11.
- Grudin, J. ve Jacques, R. (2019). Chatbots, humbots, and the quest for artificial general intelligence. *Conference on Human Factors in Computing Systems - Proceedings*, 1–11. doi:10.1145/3290605.3300439
- Guan, Q., Wang, Y., Ping, B., Li, D., Du, J., Qin, Y., ... Xiang, J. (2019). Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: A pilot study. *Journal of Cancer*, 10(20), 4876–4882. doi:10.7150/jca.28769
- Haenlein, M. ve Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), 5–14. doi:10.1177/0008125619864925
- Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, 85(4), 845–857. doi:10.1037/0033-2909.85.4.845
- Harms, M. B., Martin, A. ve Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review*, 20(3), 290–322. doi:10.1007/s11065-010-9138-6

- Hartling, C., Fan, Y., Weigand, A., Trilla, I., Gärtner, M., Bajbouj, M., ... Grimm, S. (2019). Interaction of HPA axis genetics and early life stress shapes emotion recognition in healthy adults. *Psychoneuroendocrinology*, 99, 28–37. doi:10.1016/j.psyneuen.2018.08.030
- He, L., Lech, M., Maddage, N. ve Allen, N. (2009). Stress and emotion recognition using log-Gabor filter analysis of speech spectrograms. *Proceedings - 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009*, 0–5. doi:10.1109/ACII.2009.5349454
- Hoffmann, H., Kessler, H., Eppel, T., Rukavina, S. ve Traue, H. C. (2010). Expression intensity, gender and facial emotion recognition: Women recognize only subtle facial emotions better than men. *Acta Psychologica*, 135(3), 278–283. doi:10.1016/j.actpsy.2010.07.012
- Hu, M., Wang, H., Wang, X., Yang, J. ve Wang, R. (2019). Video facial emotion recognition based on local enhanced motion history image and CNN-CTSLSTM networks. *Journal of Visual Communication and Image Representation*, 59, 176–185. doi:10.1016/j.jvcir.2018.12.039
- Hubel, D. ve Wiesel, T. (1959). RECEPTIVE FIELDS OF SINGLE NEURONES IN THE CAT ' S STRIATE CORTEX By D . H . HUBEL * AND T . N . WIESEL * From the Wilmer Institute , The Johns Hopkins Hospital and In the central nervous system the visual pathway from retina to striate cortex provides an. *The Journal of Physiology*, 148(3), 57(12), 574–591. <http://doi.wiley.com/10.1113/jphysiol.2009.174151> adresinden erişildi.
- Ide, H. ve Kurita, T. (2017). Improvement of learning for CNN with ReLU activation by sparse regularization. *Proceedings of the International Joint Conference on Neural Networks, 2017-May*, 2684–2691. doi:10.1109/IJCNN.2017.7966185
- Jaderberg, M., Simonyan, K., Zisserman, A. ve others. (2015). Spatial transformer networks. *Advances in neural information processing systems*, 28.
- Jain, D. K., Shamsolmoali, P. ve Sehdev, P. (2019). Extended deep neural network for facial emotion recognition. *Pattern Recognition Letters*, 120, 69–74. doi:10.1016/j.patrec.2019.01.008
- Johnson, J., Ravi, N., Reizenstein, J., Novotny, D., Tulsiani, S., Lassner, C. ve Branson, S. (2020). Accelerating 3D deep learning with PyTorch3D, 1–1. doi:10.1145/3415263.3419160
- Josephine Julina, J. K. ve Sharmila, T. S. (2019). Facial Emotion Recognition in Videos using HOG and LBP. *2019 4th IEEE International Conference on Recent Trends on Electronics, Information, Communication and Technology, RTEICT 2019 - Proceedings*, 56–60. doi:10.1109/RTEICT46194.2019.9016766
- Kamilaris, A. ve Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147(February), 70–90. doi:10.1016/j.compag.2018.02.016
- Karbauskaitė, R., Sakalauskas, L. ve Dzemyda, G. (2020). Kriging Predictor for Facial Emotion Recognition Using Numerical Proximities of Human Emotions. *Informatica (Netherlands)*, 31(2), 249–275. doi:10.15388/20-INFOR419
- Katsigiannis, S. ve Ramzan, N. (2017). DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. *IEEE journal of biomedical and health informatics*, 22(1), 98–107.
- Kaulard, K., Cunningham, D. W., Bühlhoff, H. H. ve Wallraven, C. (2012). The MPI facial expression database - a validated database of emotional and conversational facial expressions. *PLoS ONE*, 7(3). doi:10.1371/journal.pone.0032321

- Khairuddin, Y. ve Chen, Z. (2021). Facial Emotion Recognition: State of the Art Performance on FER2013. <http://arxiv.org/abs/2105.03588> adresinden erişildi.
- Kim, G. ve Lee, C. (2016). Convolutional Neural Network Using Convolutional Neural Network. *Springer*, 2644(2), 747–749. https://link.springer.com/chapter/10.1007/978-1-4842-2845-6_6 adresinden erişildi.
- Ko, B. C. (2018). A brief review of facial emotion recognition based on visual information. *Sensors (Switzerland)*, 18(2). doi:10.3390/s18020401
- Kohler, C. G., Turner, T. H., Bilker, W. B., Brensinger, C. M., Siegel, S. J., Kanes, S. J., ... Gur, R. C. (2003). Facial emotion recognition in schizophrenia: Intensity effects and error pattern. *American Journal of Psychiatry*, 160(10), 1768–1774. doi:10.1176/appi.ajp.160.10.1768
- Krizhevsky, A., Sutskever, I. ve Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- Lawrence, K., Campbell, R. ve Skuse, D. (2015). Age, gender, and puberty influence the development of facial emotion recognition. *Frontiers in Psychology*, 6(June), 1–14. doi:10.3389/fpsyg.2015.00761
- Lecun, Y., Bengio, Y. ve Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. doi:10.1038/nature14539
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. ve Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541–551.
- Li, Z., Liu, F., Yang, W., Peng, S. ve Zhou, J. (2021). A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *IEEE Transactions on Neural Networks and Learning Systems*, 1–21. doi:10.1109/tnnls.2021.3084827
- Liang, Q. (2010). Biologically Inspired Target Recognition in Radar Sensor Networks. *EURASIP Journal on Wireless Communications and Networking*, 2010, 1–9. doi:10.1155/2010/523435
- Liguori, I., Russo, G., Curcio, F., Bulli, G., Aran, L., Della-Morte, D., ... Abete, P. (2018). Oxidative stress, aging, and diseases. *Clinical Interventions in Aging*, 13, 757–772. doi:10.2147/CIA.S158513
- Lin, T.-Y., RoyChowdhury, A. ve Maji, S. (2015). Bilinear CNN models for fine-grained visual recognition. *Proceedings of the IEEE international conference on computer vision* içinde (ss. 1449–1457).
- Marini, M., Ansani, A., Paglieri, F., Caruana, F. ve Viola, M. (2021). The impact of facemasks on emotion recognition, trust attribution and re-identification. *Scientific Reports*, 11(1), 1–14. doi:10.1038/s41598-021-84806-5
- Meena Prakash, R., Saraswathy, G. P., Ramalakshmi, G., Mangaleswari, K. H. ve Kaviya, T. (2018). Detection of leaf diseases and classification using digital image processing. *Proceedings of 2017 International Conference on Innovations in Information, Embedded and Communication Systems, ICIIECS 2017, 2018-Janua*, 1–4. doi:10.1109/ICIIECS.2017.8275915
- Mellouk, W. ve Handouzi, W. (2020). Facial emotion recognition using deep learning: Review and insights. *Procedia Computer Science*, 175, 689–694. doi:10.1016/j.procs.2020.07.101
- Mukti, I. Z. ve Biswas, D. (2019). Transfer Learning Based Plant Diseases Detection Using

ResNet50. *2019 4th International Conference on Electrical Information and Communication Technology, EICT 2019*, (December), 1–6. doi:10.1109/EICT48899.2019.9068805

- Ng, H. W., Nguyen, V. D., Vonikakis, V. ve Winkler, S. (2015). Deep learning for emotion recognition on small datasets using transfer learning. *ICMI 2015 - Proceedings of the 2015 ACM International Conference on Multimodal Interaction*, 443–449. doi:10.1145/2818346.2830593
- Nickerson, R. S. (1976). On Conversational Interaction with Computers. *Proceedings of the ACM/SIGGRAPH Workshop on User-Oriented Design of Interactive Graphics Systems içinde*, UODIGS '76 (ss. 101–113). New York, NY, USA: Association for Computing Machinery. doi:10.1145/1024273.1024286
- Nitisha, R. J. S. U. (2018). Facial Emotion Recognition Using Machine Learning Algorithms, 632(May), 43–47. https://scholarworks.sjsu.edu/etd_projects/632 adresinden erişildi.
- Nwankpa, C., Ijomah, W., Gachagan, A. ve Marshall, S. (2018). Activation Functions: Comparison of trends in Practice and Research for Deep Learning, 1–20. <http://arxiv.org/abs/1811.03378> adresinden erişildi.
- Oh, G., Ryu, J., Jeong, E., Yang, J. H., Hwang, S., Lee, S. ve Lim, S. (2021). Drer: Deep learning–based driver’s real emotion recognizer. *Sensors*, 21(6), 1–29. doi:10.3390/s21062166
- Orgeta, V. ve Phillips, L. H. (2008). Effects of age and emotional intensity on the recognition of facial emotion. *Experimental Aging Research*, 34(1), 63–79. doi:10.1080/03610730701762047
- Ouali, Y., Hudelot, C. ve Tami, M. (2020). An Overview of Deep Semi-Supervised Learning, 1–43. <http://arxiv.org/abs/2006.05278> adresinden erişildi.
- Pascoe, M. C., Hetrick, S. E. ve Parker, A. G. (2020). The impact of stress on students in secondary school and higher education. *International Journal of Adolescence and Youth*, 25(1), 104–112. doi:10.1080/02673843.2019.1596823
- Pedamonti, D. (2018). Comparison of non-linear activation functions for deep neural networks on MNIST classification task, (3). <http://arxiv.org/abs/1804.02763> adresinden erişildi.
- Poggio, T., Kawaguchi, K., Liao, Q., Miranda, B., Rosasco, L., Boix, X., ... Mhaskar, H. (2017). Theory of Deep Learning III: explaining the non-overfitting puzzle, (073). <http://arxiv.org/abs/1801.00173> adresinden erişildi.
- Prajapati, S. A., Nagaraj, R. ve Mitra, S. (2019). Proceedings - 6th International Symposium on Computational and Business Intelligence, ISCBi 2018. *Proceedings - 6th International Symposium on Computational and Business Intelligence, ISCBi 2018*.
- Rahman, T., X-ray, C., Chowdhury, M. E. H. ve Khandakar, A. (2020). applied sciences Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using. *MDPI, J. app Sci.*, 3233, 1–17.
- Reddy, A. H., Kolli, K. ve Kiran, Y. L. (2022). Deep cross feature adaptive network for facial emotion classification. *Signal, Image and Video Processing*, 16(2), 369–376. doi:10.1007/s11760-021-01941-2
- Roy, S. K., Krishna, G., Dubey, S. R. ve Chaudhuri, B. B. (2020). HybridSN: Exploring 3-D-2-D CNN Feature Hierarchy for Hyperspectral Image Classification. *IEEE Geoscience and Remote Sensing Letters*, 17(2), 277–281. doi:10.1109/LGRS.2019.2918719

- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, 115(3), 211–252. doi:10.1007/s11263-015-0816-y
- Safira, I., Rangkuti, R., Nasution, E. H. ve Harefa, Y. (2020). Non-Verbal Communication by Autistic Children. *ELS Journal on Interdisciplinary Studies in Humanities*, 3(4), 492–505. doi:10.34050/elsjish.v3i4.8065
- Saravanan, R. ve Sujatha, P. (2018). Algorithms : A Perspective of Supervised Learning Approaches in Data Classification. *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, (Iciccs), 945–949. <https://ieeexplore.ieee.org/abstract/document/8663155> adresinden erişildi.
- Saxena, A., Khanna, A. ve Gupta, D. (2020). Emotion Recognition and Detection Methods: A Comprehensive Survey. *Journal of Artificial Intelligence and Systems*, 2(1), 53–79. doi:10.33969/ais.2020.21005
- Serengil, S. I. ve Ozpinar, A. (2020). LightFace: A Hybrid Deep Face Recognition Framework. *Proceedings - 2020 Innovations in Intelligent Systems and Applications Conference, ASYU 2020*. doi:10.1109/ASYU50717.2020.9259802
- Sharif Razavian, A., Azizpour, H., Sullivan, J. ve Carlsson, S. (2014). CNN features off-the-shelf: an astounding baseline for recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* içinde (ss. 806–813).
- Sitaula, C. ve Hossain, M. B. (2021). Attention-based VGG-16 model for COVID-19 chest X-ray image classification. *Applied Intelligence*, 51(5), 2850–2863.
- Soe, Y. N., Santosa, P. I. ve Hartanto, R. (2019). DDoS Attack Detection Based on Simple ANN with SMOTE for IoT Environment. *Proceedings of 2019 4th International Conference on Informatics and Computing, ICIC 2019*, 0–4. doi:10.1109/ICIC47613.2019.8985853
- Stursa, D. ve Dolezel, P. (2019). Comparison of ReLU and linear saturated activation functions in neural network for universal approximation. *Proceedings of the 2019 22nd International Conference on Process Control, PC 2019*, 146–151. doi:10.1109/PC.2019.8815057
- Tammina, S. (2019). Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images. *International Journal of Scientific and Research Publications (IJSRP)*, 9(10), p9420. doi:10.29322/ijsrp.9.10.2019.p9420
- Theckedath, D. ve Sedamkar, R. R. (2020). Detecting Affect States Using VGG16, ResNet50 and SE-ResNet50 Networks. *SN Computer Science*, 1(2), 1–7. doi:10.1007/s42979-020-0114-9
- Thiele, J. C., Bichler, O. ve Dupret, A. (2019). SpikeGrad: An ANN-equivalent Computation Model for Implementing Backpropagation with Spikes. <http://arxiv.org/abs/1906.00851> adresinden erişildi.
- Tian, X. ve Chen, C. (2019). Modulation Pattern Recognition Based on Resnet50 Neural Network. *2019 2nd IEEE International Conference on Information Communication and Signal Processing, ICICSP 2019*, 34–38. doi:10.1109/ICICSP48821.2019.8958555
- Tonguç, G. ve Ozaydın Ozkara, B. (2020). Automatic recognition of student emotions from facial expressions during a lecture. *Computers & Education*, 148, 103797. doi:https://doi.org/10.1016/j.compedu.2019.103797
- Vaishya, R., Javaid, M., Khan, I. H. ve Haleem, A. (2020). Artificial Intelligence (AI) applications

- for COVID-19 pandemic. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, 14(4), 337–339. doi:10.1016/j.dsx.2020.04.012
- Walvekar, S. ve Shinde, S. (2020). International Conference on Communication and Information Processing Detection of COVID-19 from CT Images Using resnet50. <https://ssrn.com/abstract=3648863> adresinden erişildi.
- Wan, L., Zeiler, M., Zhang, S., Le Cun, Y. ve Fergus, R. (2013). Regularization of neural networks using dropconnect. *International conference on machine learning* içinde (ss. 1058–1066).
- Wang, Lihui. (2019). From Intelligence Science to Intelligent Manufacturing. *Engineering*, 5(4), 615–618. doi:10.1016/j.eng.2019.04.011
- Wang, Lu, He, X., Zhang, W. ve Zha, H. (2018). Supervised reinforcement learning with recurrent neural network for dynamic treatment recommendation. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2447–2456. doi:10.1145/3219819.3219961
- Wang, Y. ve Zheng, J. (2018). Real-time face detection based on YOLO. *1st IEEE International Conference on Knowledge Innovation and Invention, ICKII 2018*, 2, 221–224. doi:10.1109/ICKII.2018.8569109
- Wei, C. Z. (2013). Stress emotion recognition based on RSP and EMG signals. *Advanced Materials Research*, 709, 827–831. doi:10.4028/www.scientific.net/AMR.709.827
- Xiao, C., Liu, Z., Zhang, T. ve Zhang, X. (2021). Deep learning method for fault detection of wind turbine converter. *Applied Sciences (Switzerland)*, 11(3), 1–22. doi:10.3390/app11031280
- Yaghoubi, S., Fainekos, G. ve Sankaranarayanan, S. (2020). Training Neural Network Controllers Using Control Barrier Functions in the Presence of Disturbances. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems, ITSC 2020*. doi:10.1109/ITSC45102.2020.9294485
- Yamashita, R., Nishio, M., Do, R. K. G. ve Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into imaging*, 9(4), 611–629.
- Yang, J. ve Yang, G. (2018). Modified convolutional neural network based on dropout and the stochastic gradient descent optimizer. *Algorithms*, 11(3), 1–15. doi:10.3390/a11030028
- Ying, X. (2019). An Overview of Overfitting and its Solutions. *Journal of Physics: Conference Series*, 1168(2). doi:10.1088/1742-6596/1168/2/022022
- Zhang, H., Jolfaei, A. ve Alazab, M. (2019). A Face Emotion Recognition Method Using Convolutional Neural Network and Image Edge Computing. *IEEE Access*, 7, 159081–159089. doi:10.1109/ACCESS.2019.2949741
- Zhang, J., Yin, Z., Chen, P. ve Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. *Information Fusion*, 59(January), 103–126. doi:10.1016/j.inffus.2020.01.011
- Zhuang, C., Zhai, A. ve Yamins, D. (2019). Local aggregation for unsupervised learning of visual embeddings. *Proceedings of the IEEE International Conference on Computer Vision, 2019-Octob*, 6001–6011. doi:10.1109/ICCV.2019.00610

