



T.C.
BATMAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİ TEKNOLOJİLERİ ANABİLİM DALI

AKTARIM DERİN ÖĞRENME METOTLARI İLE
YÜZ İFADELERİNDEN DUYGU TESPİTİ

YÜKSEK LİSANS

Sadi TAŞ

Danışman
Doç. Dr. Yılmaz KAYA

Ocak-2024
BATMAN
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Sadi TAŞ tarafından hazırlanan “Aktarım derin öğrenme metotları ile yüz ifadelerinden duygu tespiti” adlı tez çalışması 17.01.2024 tarihinde aşağıdaki jüri tarafından oy birliği / oy çokluğu ile Batman Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgi Teknolojileri Anabilim Dalı’nda YÜKSEK LİSANS olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan

Doç.Dr. Melih KUNCAN

.....

Danışman

Doç.Dr. Yılmaz KAYA

.....

Üye

Dr.Öğr.Gv. Abdülkerim ÖZTEKİN

.....

Yukarıdaki sonucu onaylarım.

Dr. Öğr. Üyesi Ömer Murat ÖTER
Lisansüstü Eğitim Enstitüsü Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

İmza

Sadi TAŞ

17.01.2024

ÖZET

YÜKSEK LİSANS

AKTARIM DERİN ÖĞRENME METOTLARI İLE YÜZ İFADELERİNDEN DUYGU TESPİTİ

Sadi TAŞ

**Batman Üniversitesi Lisansüstü Eğitim Enstitüsü
Bilgi Teknolojileri Anabilim Dalı**

Danışman: Doç. Dr. Yılmaz KAYA

2024, 64 Sayfa

Jüri

Doç.Dr. Yılmaz KAYA

Doç.Dr. MELİH KUNCAN

Dr.Öğr.Gv. Abdülkerim ÖZTEKİN

Bu çalışma, yapay zekâ ve derin öğrenme tekniklerinin duygu tanıma yeteneklerini değerlendirmek amacıyla gerçekleştirilmiştir. İnsan duygularını anlama ve yüz ifadeleri aracılığıyla analiz etme ihtiyacı, teknolojinin gelişimiyle birlikte giderek artmaktadır. Bu bağlamda, Residual Network (ResNet) ve transfer derin öğrenme modelleri kullanılarak duygu analizi gerçekleştirilmiştir. Çalışmanın odak noktası, özellikle ResNet50 modelinin duygu tanıma görevindeki etkinliğini vurgulamaktır. ResNet mimarisi, derin sinir ağlarının eğitimini kolaylaştırmak ve aşırı uyum sorunlarını azaltmak amacıyla geliştirilmiştir. Modeller, transfer derin öğrenme yaklaşımıyla öğrenilen bilgileri duygu tanıma görevine aktarmayı amaçlamaktadır. Çalışma, Fer2013 veri seti üzerinde gerçekleştirilmiş ve modellerin performansı çeşitli metriklerle değerlendirilmiştir. Elde edilen sonuçlar, özellikle ResNet50 modelinin diğer modellere kıyasla daha yüksek başarı oranına sahip olduğunu göstermektedir. Bu sonuçlar, yapay zekâ temelli duygu tanıma uygulamalarının gelecekteki potansiyelini ve derin öğrenme modellerinin bu alandaki etkisini anlamamıza katkı sağlamaktadır. Çalışmanın bulguları, duygu tanıma teknolojilerinin geliştirilmesi ve iyileştirilmesi konusunda araştırmacılara rehberlik edebilir. Yapay zekâ tabanlı duygu analizi, sağlık sektöründen akıllı yaşam alanlarına kadar geniş bir yelpazede uygulamalar bulabilir ve insan-makine etkileşimini daha etkili hale getirebilir. Bu bağlamda, ResNet50 modelinin başarısı, duygu tanıma alanında derin öğrenme modellerinin kullanımının önemini vurgulamaktadır.

Anahtar Kelimeler: Derin öğrenme, Duygu Tanıma, ResNet Modeller

ABSTRACT

MS THESIS

EMOTION DETECTION FROM FACIAL EXPRESSIONS USING TRANSFER DEEP LEARNING METHODS

Sadi TAŞ

**INSTITUTE OF GRADUATE STUDIES
OF BATMAN UNIVERSITY
THE DEGREE OF MASTER OF SCIENCE
IN INFORMATION TECHNOLOGIES**

Advisor: Assoc. Prof. Yılmaz KAYA

Year, 64 Pages

Jury

Doç.Dr. Yılmaz KAYA

Doç.Dr. Melih KUNCAN

Dr.Öğr.Gv. Abdülkerim ÖZTEKİN

This study was conducted to evaluate the emotion recognition capabilities of artificial intelligence and deep learning techniques. The need to understand human emotions and analyze them through facial expressions is increasing with the advancement of technology. In this context, emotion analysis was performed using Residual Network (ResNet) and transfer learning models. The focus of the study is to emphasize the effectiveness of the ResNet50 model in emotion recognition tasks. The ResNet architecture is developed to facilitate the training of deep neural networks and reduce overfitting issues. The models aim to transfer the knowledge learned through transfer deep learning to the emotion recognition task. The study was conducted on the Fer2013 dataset, and the performance of the models was evaluated with various metrics. The results obtained indicate that the ResNet50 model has a higher success rate compared to other models. These findings contribute to our understanding of the future potential of artificial intelligence-based emotion recognition applications and the impact of deep learning models in this field. The findings of the study can guide researchers in the development and improvement of emotion recognition technologies. Artificial intelligence-based emotion analysis can find applications in a wide range of fields, from the healthcare sector to smart living spaces, and enhance human-machine interaction. In this context, the success of the ResNet50 model emphasizes the importance of using deep learning models in the field of emotion recognition.

Keywords: Deep Learning, Emotion Recognition, ResNet Models

ÖNSÖZ

Bu tezin oluşturulmasında, değerli bilgilerini benimle paylaşan, kendisine ne zaman danışsam bana kıymetli zamanını ayırıp sabırla ve büyük bir ilgiyle bana faydalı olabilmek için elinden gelenden fazlasını sunan her sorun yaşadığımda çekinmeden danışabildiğim, güler yüzünü ve samimiyetini benden esirgemeyen ve gelecekteki mesleki hayatımda da bana verdiği değerli bilgilerden faydalanacağımı düşündüğüm kıymetli ve danışman hoca statüsünü hakkıyla yerine getiren Doç. Dr. Yılmaz Kaya' ya teşekkürü bir borç biliyor ve şükranlarımı sunuyorum.

Ayrıca hiçbir zaman maddi ve manevi desteklerini benden esirgemeyen canım eşim Sıdar Taş' a ve sevgili kızım Masal Mina ile sevgili oğlum Mustafa Sungur'a teşekkür ediyor ve onlara minnettar olduğumu dile getirmek istiyorum.

Sadi TAŞ
BATMAN-2024

İÇİNDEKİLER

ÖZET	iv
ABSTRACT.....	v
ÖNSÖZ	vi
İÇİNDEKİLER	vii
SİMGELER VE KISALTMALAR	viii
ÇİZELGE LİSTESİ.....	ix
ŞEKİL DİZİNİ	x
1. GİRİŞ	1
2. LİTERATÜR ÇALIŞMALARI	6
3. MATERYAL VE METOD	13
3.1. Materyal	13
3.2. Metod	14
3.2.1. Transfer yöntemler	14
3.2.2. Evrişimsel sinir ağları (Convolutional Neural Network)	16
3.2.3. ResNET18.....	28
3.2.4. ResNET50	30
3.2.5. ResNET101	32
3.2.6. ResNET152	33
3.2.7. Performans Ölçütleri	36
4. BULGULAR.....	39
5. TARTIŞMA	47
5.1. Hedefler ve Kısıtlamalar	49
6. KAYNAKLAR	51

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış bazı simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

<u>Simgeler</u>	<u>Açıklama</u>
AI	Artificial Intelligence, Yapay Zekâ
YSA	Artificial Neural Network, Yapay Sinir Ağı
CNN	Convolutional Neural Networks, Evrişimli Sinir Ağları
DA	Sentiment Analysis, Duygu Analizi
RSA	Similarity Analysis, Benzerlik Analizi
ANOVA	Analysis Of Variance, Varyans Analizi
ReLU	Rectified Linear Unit, Rektifiye Lineer Ünite
RNN	Recurrent Neural Networks, Rekurrent Sinir Ağları
FC	Fully Connected Layer, Tam Bağlantılı Katman
ResNet	Residual Network, Artık Ağ
TP	True Positive, Doğru Pozitif
TN	True Negative, Doğru Negatif
FP	False Positive, Yanlış Pozitif
FN	False Negative, Yanlış Negatif
SGD	Stokastik Gradient Descent, Stokastik Gradyan İnişi
GPU	Graphics Processing Unit, Grafik İşleme Ünitesi
MSE	Mean Squared Error, Ortalama Kare Hatası
ROC	Receiver Operating Characteristic, Alıcı İşletim Karakteristiği
AUC	Area Under the ROC Curve, ROC Eğrisi Altındaki Alan

ÇİZELGE LİSTESİ

Çizelge 3.1. Veri setindeki görüntülerin sınıflara göre dağılımları	13
Çizelge 3.2. Karışıklık (hata) matrisi.....	36
Çizelge 4.1. ResNet model mimarileri.....	40
Çizelge 4.2. ResNet modeller hiperparametre değerleri	41
Çizelge 4.3. ResNet Modelleri ile Eğitim veri seti için performans değerleri.....	41
Çizelge 4.4. ResNet Modelleri ile Test veri seti için performans değerleri.....	41

ŞEKİL DİZİNİ

Şekil 3.1. Yüz ifadeleri	13
Şekil 3.2. Görüntülerin sınıflara dağılım grafiği.....	14
Şekil 3.3. Makine öğrenmesi ve transfer öğrenme arasındaki süreçlerin farklılıkları	15
Şekil 3.4. CNN ağ yapısı	16
Şekil 3.5. LeCun tarafından 1989 yılında ortaya atılan LeNet isimli evrimsel sinir ağı modeli	17
Şekil 3.6. Evrişim katmanı özellik çıkarımı	19
Şekil 3.7. Sigmoid aktivasyon fonksiyonu ve türevi grafiği.....	21
Şekil 3.8. Tanh aktivasyon fonksiyonu ve türevi.....	22
Şekil 3.9. ReLU aktivasyon fonksiyonu ve türevi	23
Şekil 3.10. Leaky ReLU aktivasyon fonksiyonu ve türevi	24
Şekil 3.11. 2x2 Maksimum havuzlama , ortalama havuzlama	25
Şekil 3.12. Tam bağlantı katmanı	26
Şekil 3.13. Dropout uygulanan ve uygulanmayan katmanlar	27
Şekil 3.14. Düzleştirme katmanı.....	28
Şekil 3.15. ResNet18 modeli	29
Şekil 3.16. ResNet-50'nin ağ yapısı	30
Şekil 3.17. Önerilen TCNN(ResNet-50) akış şeması	32
Şekil 3.18. Artık blok.....	32
Şekil 3.19. SE-ResNet152 model yapısı örnek diyagramı.....	34
Şekil 3.20. Residual(artık) öğrenmenin gösterimi	35
Şekil 4.1. ResNet modellere ait başarı oranları	42
Şekil 4.2. ResNet modellere ait yetim değerleri	42
Şekil 4.3. ResNet Modellerine ait ROC eğrileri	43
Şekil 4.4. ResNet modellerine ait Precision(Kesinlik) Değerleri	44
Şekil 4.5. ResNet modellerine ait recall(duyarlılık) değerleri	45
Şekil 4.6. ResNet modellerine ait F1 skor değerleri	46

1. GİRİŞ

İnsanlığın yaratılışından günümüze kadar insanlar arasında iletişim kurma ihtiyacı duyulmuştur. Çağımız gününde de bu ihtiyaç devam etmektedir ve yapay zekâ alanında gerçekleşen gelişmelerle makinelerin insan gücünü aşabileceği kaçınılmaz bir gerçek haline gelmiştir. Hatta çoğunluğu insanlardan aldığına şahit olmaktayız. Hemen hemen hayatımızın her alanında yapay zekânın kullanıldığına şahit olmaktayız. İnsanların kendi yaşamlarını sürdürmeleri ve diğer insanlarla sosyalleşmesinde duygular son derece önem arz etmektedir. İnsanlar arasındaki iletişimin, yardımlaşmanın ve sorun çözmenin temelinde duygular yer almaktadır. Aynen bunun gibi insanlar ve makineler arasında ve makineler ile makineler arasında benzer duyguların üretilmesi ve gerekli iş ve işlemlerin yapılması söz konusu olmaktadır. Duygular insan bedeninde çeşitli duruş ve anlam oluşturmaktadır. Bununla beraber yüz ifadeleri, ses tonu, bakışlar, hareketler, mimikler duygunun karşıya aktarılmasında önemlidir. Aynen bunun gibi duygu tanımının makinelerde otomatik olarak sağlanması için yapay zekâ bu anlamda ciddi çalışmalar gerçekleştirmektedir. İnsan ve makineler arasındaki iletişimi sağlamak için makinelerin insan duygularını ve hatta kendi duygularını karşılıklı olacak şekilde yorumlaması ve sahip olduğu sisteme aktarması çağımızda mümkündür. Her ne kadar makinelerin rasyonel bir düşünceye sahip olduğunu bilsek de bu düşüncelerin tamamıyla duygudan ayrık olduğunu varsayamayız (Rázuri, vd. 2013).

İletişim ihtiyacı diğer canlı varlıklar arasında da vazgeçilmez bir yere sahiptir. İnsanların kendi aralarında iletişim kurma ihtiyaçlarının yanı sıra, hayvanlarla ve hatta bitkilerle iletişim kurma gereksinimi duyulmaktadır. Bu iletişim ihtiyacı insanlar açısından son derece önemli ve gereklidir. Tıpkı insanlarda olduğu gibi diğer canlı varlıkların da insanlarla ve birbiriyle iletişim kurma ihtiyaçları her zaman elzem olmuştur. Tam da bu noktada İnsan fizyolojisi duygusal tepkiler verme açısından son derece önemlidir. Ayrıca duyguları ortaya koyma şeklimiz ve yüzümüzdeki belirtileri cinsiyet, yaş, kültür, çevre, ırka ve buna benzer sayabileceğimiz birçok etkene bağlı olarak farklılık gösterebilmektedir. Ama genel anlamıyla aynı duygulara verilen yüz ifadeleri evrensel özellikler taşımaktadır (Ananthram, vd. 2020)

Canlıların duyguları ruhlarına sirayet ettiği gibi yüz ifadelerine de yansımaktadır. Vücut bir bütün olarak kişinin o anda hissettiği duygu halini yansıtmaktadır. İnsan duygularının

sadece konuşularak tespit edilmesi çoğu durum ve zamanlarda zor olabilir. Günümüzde teknolojinin hayatımızda hızla yer edinmesi ile beraber veriler daha da karmaşık hale gelip karmaşık bir hal oluşturmaktadır. Karmaşık verilerden daha anlamlı çıkarımlarda bulunmak için yapay zekâ (AI), görüntü işleme ve duygu analizi üzerine ciddi çalışmalar yapılmaktadır. İnsan duygusunu tanımaya yönelik yapılan araştırmalar içerisinde en kapsamlı yere yüz tanıma analizi sahiptir. Bu anlamda insanoğlu çok geniş çaplı araştırmalar yapmaktadır. Özellikle duyguların otomatik olarak tespit edilmesi, sağlık sistemi, akıllı yaşam alanları, otizm, spektrum bozukluğu, şizofreni gibi sağlık alanları ve insan ile bilgisayar arasındaki etkileşimin ve hatta insan ile robot arasındaki etkileşimin gelecek yıllarda çok büyük bir öneme sahip olmasından kaynaklı ciddi manada dikkat çekmiştir. Tam da bu noktada yüz duygu tanıma çeşidi yüz ifadeleri arasındaki duygu analizini yapmakla görevlidir. Yüzdeki duygu analizinin gerçekleşmesi iki ana adımdan oluşur diyebiliriz. Özellikle yüzün algılanması, yüzdeki değişimler, kırpmalar ve tekrardan boyutlanmalar neticesinde meydana çıkan görüntüden özellik çıkarımı sağlanmaktadır. Bu anlamda yapay sinir ağı (YSA) ve diğer makine öğrenmeleri gibi derin öğrenme yöntemleri ile duygular sınıflandırılır ya da ayrıştırılır. En çok da YSA ve özellikle evrişimli sinir ağları (CNN) görüntülerden var olan öz nitelikleri çıkarma işinde kullanılırlar (Akhand, vd. 2021).

Duygu analizi için öncelikle görüntü ve videolardan elde edilen verilerin yüz tanımı sağlanmalıdır. Yüz tanımının gerçekleşmesi öncelikli olarak farklı aşamalardan geçirilerek sağlanmaktadır. Bu aşamalar 4 temel işlemde geçilerek elde edilmektedir. Öncelikle bu işlem yüz bulma ile başlar, normalleştirme, öz nitelik çıkarma ve sınıflandırma işlemlerinin yapılmasıyla son bulur. Öz nitelik çıkarma aşaması son derece önemlidir. Dolayısıyla yukarıda saydığımız diğer 3 aşamanın başarıya ulaşması öz nitelik çıkarma aşamasının başarılı bir şekilde sonuçlanmasına bağlıdır ve böylece sağlıklı veriler elde edilebilir. Bu dört adımda verilerin eğitilmesi ve test edilmesi aynı anda eş zamanlı olarak gerçekleştirilir ve resim seti üzerinden yüz görüntüleri çekilerek duygu tahminleri geliştirilir. Bu aşamada da yüz tanıma, yüz izleme ve izlenilen yüzün şekline bağlı olarak duygu analizi aşamasına geçilir (Akçelik, vd. 2021).

Bu alanda büyük teknolojik şirketlerde görüntü işleme kütüphaneleri oluşturulmaktadır. Örneğin Intel tarafından 1999 yılında görüntü işleme kütüphanesi geliştirilmiştir (Open, 2022). Daha sonraları Windows, Linux, MacOS gibi teknoloji devleri de bu alanda farklı platformlarda çalışmıştır. Özellikle görüntü işleme noktasında

yüz tanıma teknolojisi en çok yapılan çalışmalar arasındadır. Çoğu zaman ifade edilemeyen zihinsel, psikolojik ve duygusal ifadeler insan yüzündeki ve bedenindeki değişimler ile karşı tarafa aktarıldığı bilinmektedir. Bu durumda insanların söylediklerinin yanı sıra yüz ifadelerinin de iletişimde çok büyük bir rol oynadığı bilinen bir gerçektir. Özellikle çağımızda teknolojinin gelişmesi ile insanlar arası iletişimin olumsuz etkilendiği ve yüz yüze iletişimin asgari düzeyde olduğu bir gerçektir. Yüz ifadelerinin insan iletişimi üzerindeki etkilerine yönelik birçok çalışma yapılmıştır. Nitekim çağımızda güvenlik açıklarının yoğunlukta olduğu bu süreçte yüz tanıma çalışmalarının biyometrik sistemler olarak boy gösterdiği bir gerçektir. Kişiye özel onu tam anlamıyla tanımlayacak olan açıklayıcı bilgi biyometrik sistemler vasıtasıyla rahatlıkla elde edebiliyor. Bunlardan biri yüz ifadelerinin evrensel olduğunu ve doğuştan gelen birçok duygunun hayvanlarda ve insanlarda benzer özellik gösterdiğini dile getiren Charles Darwin 1872 yılında ilk çalışmalarını yapmıştır(Darwin, 1872). Yapılan bu araştırma sonraki dönemlerde birçok bilim insanının yüz ifadelerinin tespiti üzerine çalışma yapması yolunda ilham kaynağı olmuştur. Varlıklar arasındaki iletişimde konuşmadan ziyade yüz ifadelerinin çok daha büyük bir rol oynadığı yapılan çalışmalar neticesinde ortaya konulmuştur. Son yıllarda yapılan çalışmalar neticesinde yüz ifadelerinin %55 oranında etkili olduğu bununla beraber ses tonunun %38 oranında ve seçilen kelimelerin ise %7 oranında bir etkiye sahip olduğu sonucuna varılmıştır (Mehrebian, 1968).Kişinin anlık olarak hissettiği duyguların yüzüne ya da vücuduna yansması vücut dili olarak ifade edilebilir. Yüz ifadeleri anlık duygularımızın yüzümüzde oluşturduğu değişimdir. Evrensel düzeyde yüz ifadelerinden tespit edilen duygular, mutluluk, üzüntü, korku, öfke, iğrenme ve şaşkınlık olarak tespit edilmiştir. Daha sonra yapılan çalışmalarla küçümseme yüz ifadesine de vurgu yapılmıştır (Ekman ve Friesen, 1971). Günümüzde her ne kadar çok sayıda yüz ifadelerinin tespitine yönelik çalışmalar yapılıyor olsa da birkaç yıl öncesine kadar da yüz ifadelerine yönelik çalışmalar yok denecek kadar azdı. Değişik biçimlerdeki yüz ifadelerine yönelik çalışmalar az da olsa psikologların ilgi alanına girmiştir fakat buna rağmen bu alanda ciddi çalışmalar yapılmamıştır.1970'li yıllardan sonra teknolojinin de gelişmesiyle, bilgisayarın makine öğrenmesi ve hesaplamaları sonrasında yüz ifadelerine yönelik çalışmalar önem arz etmeye başlamıştır. Böylece makine öğrenmesi ve derin öğrenme alanlarında görüntü işleme çalışmaları yavaş yavaş artmıştır. Özellikle bu alanda derin öğrenmenin yeri büyüktür. Derin öğrenme bir diğer deyişle aktarım öğrenmesi, bir görevde öğrenilenlerin başka bir görevde kullanılmasıdır. Yüz ifadeleri üzerinde duygu

analizi, makine öğrenmesi ve derin öğrenme modelleri için popüler bir görevdir. Derin öğrenme yöntemleri, yüz ifadelerinden duygu analizi yapmak için etkili bir araçtır. Yüz ifadeleri, görüntü işleme teknikleriyle öncelikle çıkarılabilecek temel özelliklerle temsil edilir. Daha sonra bu özellikler, bir derin öğrenme modelini besler ve modeli, yüz ifadesindeki duyguyu sınıflandırır. Ancak, bu modelin yeterli veriyle eğitilmesi gerekir. Örneğin, duygu analizi için bilinen bir görüntü oluşturma modeli, yüz ifadelerinden duygu analizi yapmak için kullanılabilir. Ayrıca, yüz ifadelerinden duygu analizi yapmak için veri ön işleme ve veriyi artırmak kullanılabilir. Veri ön işleme, parlaklığını ve kontrastlığını ayarlar gibi önemli özellikleri ayarlamayı içerir. Veri artırmak, eğitim veri kümesini çeşitlendirmek için kullanılır. Örneğin, aynalama, döndürme, yeniden boyutlandırma ve ses ekleme işlemleri kullanılabilir. Sonuç olarak, yüz ifadelerinden duygu analizi yapmak için derin öğrenme ve elde edilen öğrenme etkili bir araçtır. Ancak, yeterli veri ve doğru veri ön işleme ve artırma ile birlikte kullanılması gerekir (Goodfellow, vd. 2016).

Çoğu alanda olduğu gibi yüz ifadelerinin tespiti noktasında da yeterli düzeyde bilgi sahibi olunmadığı için toplumda çeşitli kavramlar karıştırılmaktadır. Yüz ifadelerinin tanınması ve duyguların tanınması kavramları da bunlardan biridir ve maalesef çoğunlukla karıştırılmaktadır. Bu iki kavram birbirinden keskin hatlarla ayrılıyor olmasa da birbirinden farklı kavramlardır. Hissettiğimiz bir duygunun yüzümüzde oluşturduğu şekil yüz ifadesi olarak tanımlanıyorken, duygunun tanınması ise bundan çok daha karmaşık bir yapıdır diyebiliriz. Yüz ifadesinin tespiti yüzdeki değişimlerden anlaşılıyor, duygu tespiti için bunların çok daha fazlası gerekmektedir. Duygu tanıma insanın sesinden, bakışlarından, jestlerinden, yüz ifadelerinden kısacası bütün duruşundan yola çıkılarak, beden dilinin bir bütün olarak ifade edilmesi ve yorumlanması ile tespit edilebilir. Kişi simasında bulunan bütün uzuvların doğal yapısından farklı bir hal almasıyla meydana gelen değişiklikler yüz ifadesinin tespit edilmesi noktasında sadece fikir oluşturur. Meydana gelen yüz değişimi bir bütün olarak düşünülüp karmaşık bir yapı olarak ele alınmalıdır. Böylece ifadenin tespiti noktasında sağlıklı çıkarımlarda bulunulabilir. Yüz ifadelerinin tanımlanması yapay zekâda önemli bir yere sahiptir. Sağlık sektöründe hastalardaki ruhsal durumlar ve hastaların depresiflik ve otistik durumları tespit edilebilir. Bunun yanı sıra güvenlik alanında, adli ve idari bilgilerin toplanmasında, kişilerin sınav sistemlerindeki güvenliğinden tutun çok büyük bir kullanım alanına sahiptir. Bununla beraber yüz tanıma sistemleri, artık bireysel olarak

kullanmış olduğumuz elektronik cihazlarda ve hayatımızda daha birçok alanda kullanılmaktadır. Örneğin sürücülerin yüz ifadelerine göre yorgunluk tespiti yapıp olası kazaların önüne geçilebilir. İlegal durumlarda bulunan insanların tespit edilmesi, kişilerin ruh durumuna göre hobi edinmesi gibi birçok alanda yüz ifadelerinden faydalanılmaktadır (Kobani, vd. 2015). Bütün bu çalışmaların odağında son dönemlerde önemli bir yer edinen bir diğer yöntem ise derin öğrenme olmuştur. İsminden de anlaşıldığı gibi derin kavramı çok katmanlı bir yapıya sahip olup yapay sinir ağlarına nazaran daha katmanlı bir yapıdan oluşur. Bu çoklu ağlar vasıtasıyla görüntü işleme alanında ciddi çalışmalar yapılmıştır. Bu alana büyük katkılar sunmuştur (Garcia, vd. 2018).

Bu çalışmada yüz ifadelerinden duygu tespiti sorununu ele aldık. Özellikle altı temel duygu üzerinde durduk. Duygu tespitinin analizi için aktarım derin öğrenme metotlarından ResNet transfer öğrenme yöntemlerinden faydalandık. Bunun için yapay zekâ alanında yapay sinir ağı kullanılarak derin öğrenme yöntemleri ile büyük oranda eğitim ve test aşamalarından sonra duygu tespitini gerçekleştirmiş olduk. Bu çalışmada öncelikle çalışmanın amacı, avantajları ve gerekliliği üzerinde durulmuştur ve gerekli çalışmalar gerçekleştirilmiştir. İkinci kısımda ise çalışma ile ilgili geniş çaplı literatür çalışması yapılmış ve yapılan çalışmalar değerlendirilerek farklı çıkarımlarda bulunulmuştur. Üçüncü kısımda ise toplanan materyaller ile kullanılacak metotlar belirlenip eğitim ve test aşamaları gerçekleştirilmiştir. Nihai aşamada ise gerçekleştirilen eğitim ve test sonuçları, kullanılan aktarım derin öğrenme metotları karşılaştırılarak bu metotların olumlu ve olumsuz yönleri tespit edilip sunulmuştur. Tüm çalışmalar neticesinde başarı düzeyleri çizelge ve karışıklık matrisi olarak sunuldu. Nihayetinde ısı haritası yöntemi ile görüntülerin dağılımları değerlendirildi ve deneysel sonuçlar sunuldu.

2. LİTERATÜR ÇALIŞMALARI

İnsanlar arasında iletişimin sağlamlasında büyük bir role sahip olan yüz ifadeleri insanların kontrolü dışında gerçekleşir. Sağlıklı iletişimin gerçekleşmesinde yüz hatlarındaki fiziksel değişimler çok önemlidir. Çağlar ATILGAN bu çalışmada gerçek veriler üzerinde yapay sinir ağları kullanarak kişilere ait duyguların yalın hallerini tespit etmeye çalışmıştır. Bunu yaparken özellikle yüz bölgesinde meydana gelen değişimlerden faydalanmıştır. Ağız, çene, kaş, göz gibi bölgelerde meydana gelen değişiklikleri nirengi noktalama yöntemi işaretleme yaparak gerçekleşen değişimleri ve ilişkileri noktalar arasındaki mesafelerden tespit etmeye çalışmıştır. CohnKanadeExtendedresim veri tabanını kullanılarak, 4.000'in üzerinde resim ile YSA üzerinde eğitim gerçekleştirerek %65 doğruluk oranına erişmiştir. Toplamış olduğu resimlerden kişilerin cinsiyet, yaş, ırk ve boy gibi özelliklerinden bağımsız olarak mutlu, şaşkın, üzgün, öfkeli, sürpriz, korku gibi halleri yüksek bir oranla tespit etmiştir (Atılğan, 2019).

Bir diğer çalışma ekibi ise yüz ifadelerine yönelik çalışmalarını sağlıklı bir şekilde yürütmek için donanım özellikleri yüksek bir bilgisayar üzerinden yapmışlardır. Üzerinde çalıştıkları veri seti 720 videodan oluşan ve altı temel yüz ifadesini içeren videolardan oluşmaktadır. Bu çalışma veri seti 773 eğitim örneğinden, 383 doğrulama örneğinden ve 593 test örneğinden oluşmaktadır. Ve tüm örnekler yedi yüz ifadesine bölünmüştür. Bu videoların %81'i erkeklerden, %19'ukadınlardan toplanmıştır. Bu veri seti sürpriz, tikslenme, üzüntü, mutluluk, korku ve öfke olmak üzere altı yüz ifade şeklinden oluşmaktadır. Veri eğitim setini büyütme için çevirme, döndürme, renk bozma ve benzeri işlemlerden faydalanılmıştır. Caffé derin öğrenme çerçevesi kullanılarak bu yöntem uygulanmıştır. RML veri seti üzerindeki denemelerde geliştirilmiş AlexNet ile %72.63, VGG-Face ile %87.12 ve CNN+RNN ile %88.96 başarı elde edilmiştir (Wang, vd. 2020).

Yapılan bir diğer çalışma ise CNN derin öğrenme modeli ile duygu analizinin tespitine çalışılmasıdır. İzlenen videolar saniye saniye kayıt altına alınarak insan simalarındaki anlık değişimlerde varılan duygu analizi grafiğe dökülmüştür. Gerekli duygu yüklü materyaller toplandıktan sonra, en iyi sonucu verebilecek bir CNN derin öğrenme modeli oluşturulmuş ve en son toplanan videolardan yüz ifadelerinin analiz edilmesi sağlanmıştır. Toplam 61 adet film izlenmiş ve binlerce resim toplanmıştır. İzlenen filimler farklı duyguların yaşandığı çeşitli kategorilerde olmasına dikkat edilip bikkınlık, korku, mutluluk, sakinlik, şaşkınlık, sinirlilik ve üzünlük gibi duyguların yer

aldığı fotoğraf kareleri kırılmıştır. Yüz ifadesi tespiti kısmında Haarcascade tekniği kullanılmıştır. Tespit edilen yüzler Amazon webservisi olan FaceRecognition yardımı ile duygulara göre kategorize edilmiştir. Her ne kadar 50 bin civarı yüz resimleri toplanmış olsa da Haarcascade ile ve Amazon web ile yapılan tetkikler neticesinde 30 bine yakın resim niteliksiz, kalitesiz ve yanlış olduğundan ayıklanmıştır. Derin öğrenme ile yapılan eğitim ve sınamalar neticesinde bıkkınlık ve şaşkınlık duygu hallerinin en çok karıştırıldığı sonucuna varılmıştır. Geri kalan beş duygu resimleri ile test analizi yapıldığında ise %60 lık bir doğruluk başarısı elde edilmiştir (Arıç ve Metin, 2021). Yapılan çalışma neticesinde filmlerin içerisinde geçen duygular, duyguların yoğunlukları ve farklı zaman aralıklarında gerçekleşen duygu çeşitliliği ile orantıları tespit edilmiştir. Dolayısıyla izlenen filmlerin en çok hangi duyguya yer verdiği ve bunun neticesinde, filmler içerdiği duygu yoğunluğuna göre kategorize edilmiştir. Yayınlanan bu çalışmada yüzdeki ifadelerden duygu analizi neticesinde elde edilen başarı oranları VGG-16 ile %39.65, SVM ile %37.55, LSTM ile %38.00, HopField RNN %46.65, MLP ile %72.00 olarak gözlenmiştir.

Yüz ifadelerinin üzerine yapılan birçok çalışmadan biri de yüz felci geçirmiş hastalar üzerine yapılmıştır. Yüzdeki kasların sağlıklı olarak çalışmadığı ya da yüz felci geçirmiş hastalar üzerine yapılan bu çalışma hastaların konuşmada, çiğnemede ve duygularını ifade etmede zorluk çektikleri tespit edilmiştir. Duygu değişimlerinin yüz ifadelerine sağlıklı bir şekilde yansıyamadığı bu durum fasiyal sinir felci olarak tanımlanmıştır. Yüz felci geçirmiş veya yüzündeki sinirleri normal insanlara göre sağlıklı çalışan yaklaşık 45 hastadan yüz ifadelerini içeren resimlerden oluşan bir veri seti oluşturulmuştur. Böylece özellikle yaşamış oldukları stres ve hissettikleri altı duygu tespitine yönelik çalışma amaçlanmıştır. Bunu gerçekleştirmek için VGGNet derin öğrenme modeli kullanılmıştır. Sonuç olarak hastalara ait esas duyguların tespit edilmesi %66,58'lik bir başarıyla sonuçlanmıştır. Fakat öfke, korku ve iğrenme gibi duyguları birbirinden ayırt etme yeterli düzeyde sağlanamamıştır. Ama bunun aksine stresle alakalı olumsuz duygular %85,97 gibi bir doğruluk oranıyla sevindirici bir netice elde etmişlerdir (Xu, vd. 2022).

Ruhun Dünya'ya açılan peneceresinin yüz olduğunu dile getiren 19. yüzyıl Fransız doktoru Duchenne de Boulogne, insanları elektrik şoklarına maruz bırakarak yüz kaslarında meydana gelen değişimleri gözlemlemiştir. Böylece insanların bilinen ve bilinmeyen tüm duygularının yüzde oluşturduğu ifadeleri anlamaya çalışmıştır. William Dias ve arkadaşları da buradan yola çıkarak, temel duyguları tespit etmek için ve genel

yüz hatları var mı diye çalışmalara başlamışlardır. Bunu için evrişimli sinir ağlarını kullanmışlardır. Bu çalışmada bir grup üzerine çalışmalar yapılmıştır. Ve bu gruptan elde edilen yüz ifadelerinin diğer insanlar üzerinde de aynı etkiyi oluşturup oluşturmadığı gibi bir genellemeye gidilebilir mi diye çalışmalar yapılmıştır. Bunu da bir çapraz veri kümesi yöntemiyle yapmışlardır. Öncelikle, duyguları kategorize etmek için yüz tanımda yardımcı olacak hususi olarak bir model tasarlamışlar. Veri görüntüleme teknikleriyle temel modellerini iyileştirmişler. Ağın değişik yerlerdeki ayırt edici detayları öğrenebilmesi için yüz görüntülerinde yamalar kullanılmış. Ve bu şekilde ağlarını eğitmişler. Eğitim aşamasında CK+, MUG ve Oulu-CASIA veri kümelerini, test aşamasında ise Bosphorus, CAFE, CFEE, KDEF, NVIE ve RaFD veri kümelerini kullanmışlar. Sonuç olarak öfke duygu setindeki görüntülerin %47,93'ü nötr olarak sınıflandırıldığı görülmüştür. İğrenme duygusu için %30,21, korku duygusu için %50,72, üzüntü duygusu için %53,24 ve sürpriz duygusu için %45,92 olarak görülmüştür. Buradan yola çıkılarak bazı duyguların EMFACS kontrolünden sapabileceği görülmüştür. Dolayısıyla bazı yüz ifadelerinin geçişlerinin çok yüzeysel olduğu ve bundan dolayı da kategorize edilemediği sonucuna varılmıştır (Dias, vd. 2022).

Ann Devitt ve Khurshid Ahmad Son yıllarda, bilgi erişimi alanında hızla büyüyen bir uygulama ve araştırma alanı olarak ortaya çıkan DA (duygu analizi) üzerine çalışmalarını yürütmüşler. Otomatik DA sistemlerinin leksikal(sözcüksel) kaynaklarını seçme görevini kolaylaştırmak amacıyla bu çalışmada, dört yaygın kullanılan duygu sözlüğünün ayrıntılı bir analizini sunmaktalar. Bu analiz, her bir sözlüğün kapsamını ayrı ayrı inceleyerek sözlüklerin dildeki duygusal içeriği ne ölçüde kapsadığını ortaya koymaktadır. Ayrıca, dört kaynak arasındaki örtüşme ve tutarlılığı araştırarak farklı sözlükler arasında ortak kelimelerin ve duygusal etiketlemelerin ne kadar paylaşıldığını gözlemlemektedir. Bu analiz, farklı duygu sözlükleri arasındaki anlaşma ve tutarlılık düzeyini değerlendirerek sistem geliştiricileri için önemli bilgiler sunmaktadır. Özellikle duygu sözlüklerini seçerken veya birleştirirken bu tutarlılık ve örtüşme faktörlerini azami derecede dikkate almışlar. Üstelik makaledeki korpus analizi, duygu sözlüklerinin genel dil ve özelleşmiş dildeki leksikal içerik dağılımını incelemektedir. Bu analiz, duygu sözlüklerinin farklı alanlarda veya konularda nasıl performans gösterdiğini ve genel duygu analizi mi yoksa özelleşmiş alanlara mı daha uygun olduğunu belirlemeye yardımcı olmaktadır. Bu çalışma, geliştiricilerin DA sistemleri için uygun leksikal kaynakları seçmelerine yardımcı olmayı hedeflemektedir. Elde edilen bulgular, kapsam,

örtüşme, tutarlılık ve dil alanlarına uygunluk gibi faktörleri dikkate alarak duygu sözlüklerinin seçimi, birleştirilmesi veya özelleştirilmesi konusunda bilgi sağlayabilir.

Kısaca özetlenecek olursa Ann Devitt ve Khurshid Ahmad bu çalışmada, dört yaygın kullanılan duygusal leksikal kaynağın ayrıntılı bir analizini sunuyor. Analiz, leksikonların duygusal özellik atamaları ve leksikal içerikleri arasında tutarlılık olduğunu gösteriyor. Ancak, kaynaklar arasında yeterli farklılıklar olduğu ve bunların otomatik duygu analizi sistemleri üzerinde etkileri olduğu belirtiliyor. Ayrıca, duygusal metin içeriğinin dil çeşitlerine göre farklılık gösterdiği ve genel dilin ötesinde istatistiksel olarak ayrı bir alt küme olduğu ortaya çıkıyor. Gelecekteki çalışmaların, duygusal metinlerdeki farklılık temelinin ve dil çeşitlerine bağlı olarak nasıl değiştiğini belirlemeye odaklanması öneriliyor (Devitt & Ahmad , 2013).

Jeffrey A. Brooks ve arkadaşları yaptıkları çalışmalara göre, İnsanlar yüz hareketlerini belirli duygusal kategorilere güvenilir bir şekilde sınıflandırır. Son zamanlarda yapılan araştırmalar, duygu-kavram bilgisinin duygusal algıyı nasıl etkileyebileceğini ve yüz-duygusal algıda evrenselliğin yanı sıra değişkenliğin de olduğunu ortaya koymaktadırlar. Bu çalışmalara göre duygu algısındaki bu değişkenliği anlamak yeni yeni gelişmektedir. Ve duygu-kavram bilgisinin yapısal etkilerinin sinirsel temeli hala bilinmemektedir. Jeffrey A. Brooks ve arkadaşları bir nörogörüntüleme çalışmasında, altı duygusal kategorinin (Öfke, Tiksini, Korku, Mutluluk, Üzüntü ve Şaşkınlık) kavramsal, algısal ve nöral temsili yapıları arasındaki uyumu ölçmek için bir temsili benzerlik analizi RSA yaklaşımı kullanmışlar. Elde ettikleri sonuçlar, katılımcıların duygusal kavramsal yapılarında bireysel farklılıklar gösterdiğini ve bu farklılıkların benzersiz algısal yapılarını tahmin etmelerine yardımcı olduğunu göstermiş. Yüzleri görüntülerken, sağ fusiform girusun çoklu voxel desenlerinin temsili yapısı, katılımcının benzersiz kavramsal yapılarıyla anlamlı bir şekilde ilişkilendi ve yüzlerin kendilerindeki fiziksel benzerliklerin etkisi kontrol edildi. Ayrıca, duygu algısındaki kültürler arası farklılıkların varlığı da gözlemlendi ve bu farklılıkların bireysel kavramsal yapılarla açıklanabileceği düşünüldü. Bulguları, görsel yüz işleme bölgelerindeki duygu ifadelerinin temsili yapısının bireysel kavramsal anlayıştan etkilenebileceğini göstermektedir.

Sonuç olarak birçok kültür arasındaki farklılıkları araştıran bir çapraz-kültürel hipotezlerine göre, önceki çalışmalarda, Doğu ve Batı kültürlerinde yüksek uyarılma düzeyine sahip olumsuz yüz ifadelerinin farklı algılanmasının bulunduğunu ve Japon ve Amerikalı katılımcılarda daha az kesin bir algısal yapı oluşturduğunu (59, 61, 62)

göstermiştir. Bu çalışmada, kategorizasyon yanıtlarına atıfta bulunurken, "uyumsuzluk" terimini, amaçlanan duygu gösterimine uygun olmayan yanıtları ifade etmek için JACFEE (Facial Expressions Of Emotion) uyaran veritabanına göre, hatalı olduklarını ima etmek amacıyla kullanmışlar. Kùltürler arasında kategorizasyonlarda anlamlı düzeyde uyumsuzluk oranları bulundu. Japonlar ($n = 20$) (%18,125, $SD = 7,86$) Amerikalılardan ($n = 20$; %6,875, $SD = 7,56$) daha yüksek uyumsuzluk gösterdi [$t(38) = 4,614$, $P = 0,0000439$]. Özellikle, Japon katılımcılar, poz verilmiş öfkeli yüz ifadelerini tiksinişmiş olarak (ve tam tersi), poz verilmiş korkmuş yüz ifadelerini ise şaşırmış olarak (ve tam tersi) sıklıkla kategorize ettiler, Amerikalı katılımcılara göre.

Birincil ilgi konusu olmamakla birlikte, uyaran seti hem Kafkas hem de Japon hedef yüzlerini içerdiği için, katılımcıların kùltürü ve hedef yüzlerin kùltürü arasında olası bir etkileşim de test edildi ve uyumsuzluk oranlarına etkisi incelendi. Tekrarlanan ölçümlü bir ANOVA (Analysis of variance) analizi sonuçlarına göre, katılımcının kùltürü ve hedef yüzlerin kùltürü arasında anlamlı bir etkileşim olduğu bulundu [$F(1,38) = 4,265$, $P = 0,046$]. İleriki eşleştirilmiş örneklem t testleri, Japon katılımcıların Kafkas yüzler için daha yüksek doğruluk oranına sahip olduğunu ($M = 0,84$) Japon yüzleri için ($M = 0,79$) göre ($t(19) = 2,6$, $P = 0,017$) belirlemiştir (Brooks & Chikazoe, 2019) .

Bir diğer çalışmada ise katılımcılar, dinamik yüz ifadelerini inceleyerek nötr bir ifadeden öfkeli, mutlu veya üzgün ifadelerle doğru değişen hareketleri izlediler. Nötr bir ifadeye geçen ifadelerde bir kontrast etkisi gözlemlendi. Yani, öfkeyle başlayan nötr bir ifade hafifçe olumlu olarak değerlendirildi, ancak gülümsemeyle başlayan aynı nötr ifade olumsuz olarak değerlendirildi. İkinci deneyde, ardışık olarak sunulan durağan ifadeler de karşıt, zıt etkileri tetikledi, ancak hareketli ifadelerin ardından görülen etkiler daha zayıftı. Üçüncü deneyde, öfke ve mutluluk ifadelerinin farklı derecelerinde birçok yüz hareketi değerlendirildi. Hafifçe ifade edilen hareketlerde (25% ve 50% ifade) denge ve itme etkileri gözlemlendi. Bu tür hareketlerde algılanan ifade, hedef noktalarının ötesine geçmiş gibi algılandığı için duygu derecelendirmeleri daha yüksek oldu. Dördüncü deneyde, üzgün yüz ifadeleri değerlendirildi ve hareketli ifadelerde hem zıtlık hem de itme etkileri gözlemlendi. Bu bulgular, dinamik yüz ifadelerinde yeni ve etkili bağlamsal etkilerin olduğunu ve yüz hareketlerinin sosyal-duygusal iletişimde önemli bir rol oynadığını ortaya koymuştur.

Deney 1, University of California, Berkeley'deki 24 lisans öğrencisiyle gerçekleştirildi. Katılımcılara, öfkeli, mutlu ve nötr yüz ifadelerinin dinamik hareketlerini içeren video klipler gösterildi. Hareketin sonunda nötr bir ifadeye ulaşıldığında,

başlangıçtaki duygusal ifadenin etkisiyle nötr ifade farklı şekillerde değerlendirildi. Aynı şekilde, hareketin sonunda hafifçe ifade edilen noktalarda momentum etkileri gözlemlendi. Deney, hareketli yüz ifadelerinin bağlamsal etkilerini ve yüz hareketlerinin sosyal-duygusal iletişimdeki önemini vurgulamıştır. Deney 1'in sonuçlarına göre, katılımcılar başlangıç ve bitiş yüzlerinin duygu derecelendirmelerini etkileyebilmiştir. Öfkeli mutluya doğru hareket eden nötr bir ifade, hafif bir olumlu ifade olarak değerlendirilirken, aynı nötr ifade mutlu başladığında olumsuz olarak değerlendirilmiştir. İki nötr koşulda belirgin yanlılık etkileri gözlemlenmiştir. 100% öfkeli bir nötr ifade, başlangıçta görülen aynı nötr ifadeden daha olumlu olarak değerlendirilirken, 100% mutlu bir nötr ifade, başlangıç derecelendirmelerine göre daha olumsuz olarak değerlendirilmiştir. Deney 2'de, dinamik hareket olmaksızın ardışık olarak elde edilen kontrast, zıtlık etkilerinin derecesi değerlendirilmiştir (Marian & Shimamura, 2013).

Ray ile Chakrabarti duyguların öğrenme üzerindeki etkilerini araştırmışlar. Teknoloji Destekli Öğrenme, öğretme-öğrenme sürecini dönüştüren bir yöntemdir ve duygunun rolü genellikle ihmal edilir. Ancak duygu, insanın bilişsel sürecinde önemli bir yer tutar, bu nedenle öğrenme sürecinin tam anlamıyla başarılı olabilmesi için öğrencinin duygusal durumunu yakalamak gereklidir. Bu çalışmada, E-Öğrenme sistemlerinde duygusal yönleri yakalayan ve öğrenme faaliyetlerini, zamanlamayı ve genel öğrenme sürecini yöneten bir Duygusal Hesaplama Modülü önerilmektedir. İnsan-Bilgisayar Etkileşimi, öğrencinin duygusal durumunu tanımaya odaklanır ve insan-bilgisayar arasında etkileşimi sağlar. Bu amaçla, biyofiziksel (kalp atış hızı, deri iletkenliği, kan basıncı) ve yüz ifadesi yöntemleri kullanılmaktadır. Bu çalışma, öğrenme sürecinde duygunun nasıl dönüştüğünü ve duygu geri bildirimlerinin öğrenme deneyimlerini nasıl geliştirmekte kullanıldığını araştırmaktadır. Öğrenci duygu tespiti ve otomatik ders seçimi modeli önerilmekte olup, bu modelde iki duygu özelliği birleştirilerek etkili bir Affective E-Learning Sistemi oluşturulmuştur. Araştırma, mevcut sistemlere kıyasla olumlu sonuçlar göstermektedir ve tanıma sisteminin etkinliğini kanıtlamaktadır.

Tüm bu veri ve tespitlerden şu sonuca varılabilir ki yüz ifadeleri, duygular hakkında önemli ve keskin ipuçları sağlar (Ray ve Chakrabarti, 2016).

Gerçekleştirilen diğer bir çalışma ise duygu ve disiplinler arası etkileşimi ele almaktadır. Duygu konusundaki araştırmalar farklı disiplinler ve uzmanlık alanları arasında çeşitlilik göstermektedir. Ancak uzmanlar duyguyu tanımlama konusunda hala anlaşmaya varabilmiş değildir. Teorisyenler ve araştırmacılar, duygu terimini farklı süreçler ve anlamlar içeren şekillerde kullanmaktadır. Duyguların doğası, işlevleri, geniş

duygusal boyutlarla ilişkilendirilir. Duyguları harekete geçiren süreçler ve günlük etkinliklerimizdeki rolü konusunda hala tartışmalar devam etmektedir. Bu çalışmada, doğal türler olarak kabul edilen temel duygular ele alınmış. Duygu şemalarının doğası, duygu ve biliş ilişkilerinin duygu şemalarını nasıl geliştirdiği üzerine çalışmalar derinleştirilmiş. Ve ayırık duyguların duygusal boyutlarla ilişkisi gibi konulara odaklanmıştır. Sonuç olarak, sürekli duygu kavramının bilinci düzenlemede ve zihin ile davranış üzerinde etkili bir faktör olarak kabul edildiği yeni bir dizi işaret ediliyor. İncelenen kanıtlar, temel duygular ve duygu şemaları kavramlarına dayanan bir teorinin, daha bütünsel veya boyutsal yaklaşımlarla uyumlu bir araştırma aracı sağladığını göstermektedir.

Bu çalışmada, duyguların doğası, insan gelişimi ve bireysel-sosyal işlevlerle ilgili birçok sorunun hala yanıtlanması gerektiği belirtilmektedir. Duyguyla ilgili olgular için daha tutarlı bir isim ve sınıflandırma sisteminin geliştirilmesi önemlidir. Duygu araştırmacılarının duygusal terimlerin işlevsel açıklamaları için daha yüksek standartlar belirlemesi gerekir. Ayrıca duygu teriminin niteliksiz kullanımını engellemesi gerekmektedir. Duygusal bilim, temel duyguların işlevsel karşılıkları, duygu şemaları ve özelliklerinin işlevsel etkilerine odaklanan yeni yaklaşımların geliştirilmesiyle ilerleyebilir. Bu örnek, sürekli mevcut olan ve farklı düzeylerde etkilere sahip duyguların varlığına dayanmaktadır. Duygusal bilimcilerin, duygunun sürekli olarak algısal-bilişsel süreçlerle etkileşim halinde olduğunu test etmek için hipotezler ve yöntemler geliştirmesi gerekmektedir (Izad, 2007; Pantic, 2009).

3. MATERİYAL VE METOD

3.1. Materyal

Bu tez çalışmasında Fer2013 isimli veri seti kullanılmıştır. Bu veri seti Kaggle veri tabanından indirilmiştir. Veri seti 48x48 pikselleri renkli görüntülerden oluşmaktadır. Veri seti yüz ifadelerin duygu tanıma için kullanılan zor bir kıyaslama setidir. Veri seti 7 farklı duyguya ait yüz ifadesi belirten görüntülerden oluşmaktadır. Görüntüler zor bir veri seti olup düzgün dağılmamış görüntülerden oluşmaktadır. Toplam 35340 adet görüntü bulunmaktadır. Her duygu için görüntü sayıları Çizelge 3.1’de verilmiştir. Veri setinden rastgele görüntüler Şekil 3.1’de verilmiştir.

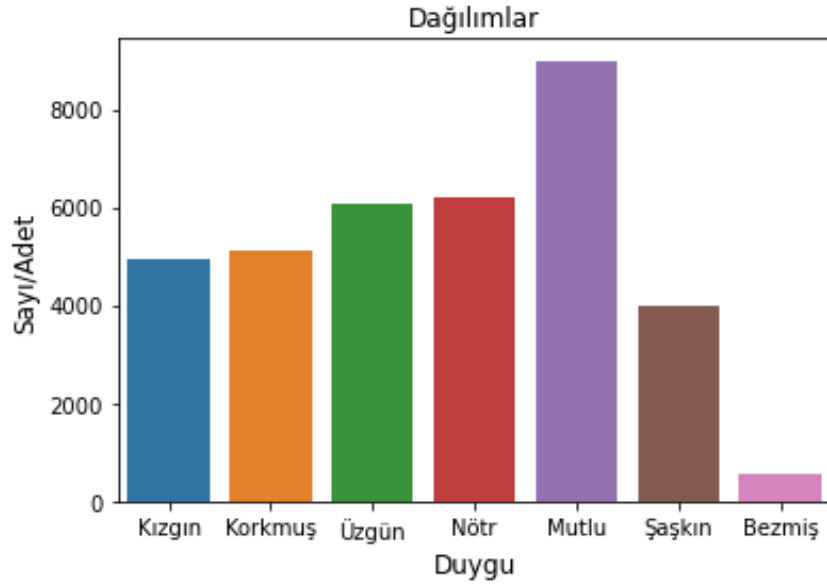
Çizelge 3.1. Veri setindeki görüntülerin sınıflara göre dağılımları

Sıra	Duygu	Görüntü Sayısı
0	Kızgın	4953
1	Korkmuş	5121
2	Üzgün	6077
3	Nötr	6198
4	Mutlu	8989
5	Şaşkın	4002
6	Bezmiş	547



Şekil 3.1. Yüz ifadeleri

Veri setindeki görüntülerin sınıflara dağılım bar grafiği şekil 3.2’de gösterilmiştir. Şekilden görüldüğü gibi en fazla görüntü “mutlu” sınıfında, en az görüntü ise “Bezmiş” sınıfında bulunmaktadır.



Şekil 3.2. Görüntülerin sınıflara dağılım grafiği

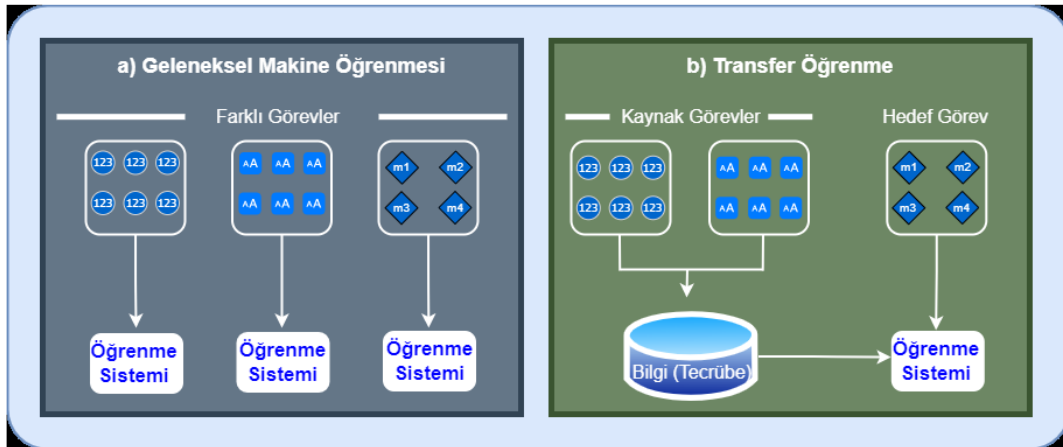
3.2. Metod

3.2.1. Transfer yöntemler

Transfer öğrenme, geleneksel öğrenmeye göre çok daha farklı bir yaklaşım izler. Geleneksel makine öğrenmesi, öğrenme gerçekleştirirken her seferinde sıfırdan öğrenmeye başlarken, transfer öğrenme daha önce öğrendiklerini yeni öğrenmede kullanarak hem öğrenme sürecini hızlandırır hem de daha etkili bir öğrenme gerçekleştirir. Başka bir deyişle transfer öğrenme, yapay öğrenme düzeneklerinin eğitildiği bir veriyle daha farklı bir konu ya da benzer bir problemi çözmek için kullanabilen bir yaklaşımdır. Tıpkı insanların kendi deneyimleriyle başka bir problemi çözmelerine benzer bir yaklaşımla, makine öğrenmesi geliştirilmektedir. Nihayetinde bir makine öğrenmiş olduğu bir konuyu aynı deneyimle başka konuları çözebilir ve doğru kararlar alabilir (Koçer, 2012). Transfer öğrenme tekniği, günümüzde derin öğrenmede sıkça kullanılmaya başlanmıştır. Eğitilmiş olan bir yapay sinir ağı modeliyle yeni ve benzer bir problem, bilgi transferi yardımıyla çözülebilmektedir. Bu görev esnasında sıfırdan bir model eğitmek yerine daha önce kullanılmış bir model göreve uyarlanır ve

görev gerçekleştirilir. Böylece eğitim süreci kısalır ve kaynak sorunu ortadan kalkar. Eğitim süreci, önceden eğitilmiş ağın desen tanıma yeteneklerini ve temel özelliklerini küçük değişikliklerle yeni göreve uyarlamakla tamamlanır. Bir görev esnasında, özellikle mevcut verinin yetersiz ve sürenin sınırlı olduğu durumlarda benzer bir görevde eğitilmiş bir modelle daha iyi sonuçlar elde edilebilir (Komodakis & Zagoruyko, 2017). Görev esnasında transfer öğrenme gerçekleştirilirken, transfer edilen modelin kullandığı veri setinin hedef veri setinden daha büyük olması gerekir ve daha çok veri çeşitliliğine sahip olması önemlidir. Bu daha etkili özellikler öğrenmeyi sağlar ve modelin öğrenmesini geneller. Bu şekilde başarılı bir öğrenme ve işin sonunda görev başarıyla gerçekleşmiş olur.

Transfer öğrenme sürecinde veri seti seçimi son derece önemlidir. Seçilen veri seti hedef veri seti ile aynı ya da benzer özelliklere sahip değilse, modelden beklenen performansın gerçekleşmesi beklenemez. Bu nedenle veri setinin benzer ve yeterli veri miktarını içermesi hedeflenen sonuç için çok önemlidir (Gao & Mosalam, 2018, s. 740-770). Geleneksel makine öğrenmesi ile transfer öğrenme arasındaki süreçlerin farklılıkları aşağıda verilen şekilde görüldüğü gibidir.



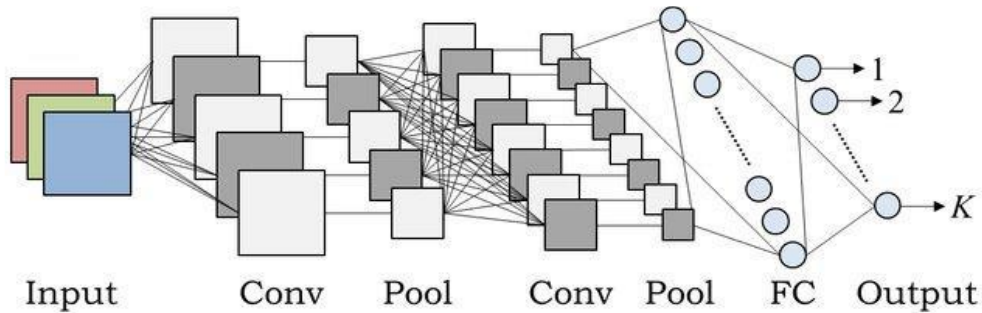
Şekil 3.3. Makine öğrenmesi ve transfer öğrenme arasındaki süreçlerin farklılıkları (Yang & Pan, 2009)

Sonuç olarak, transfer öğrenme yöntemleri, geçmişte edinilen deneyimlerin gelecekteki öğrenmeye katkı sağlamasıyla önem kazanmış bir makine öğrenme paradigmasıdır. Bu yöntemler, önceki görevlerden elde edilen bilgileri yeni ve farklı görevlere aktarabilme yeteneği ile öne çıkar. Transfer öğrenme, makine öğrenme alanında geniş bir uygulama yelpazesi sunarak, bir görevde elde edilen bilgilerin diğer görevlere etkili bir şekilde transfer edilmesini sağlar. Bu, modelin daha hızlı öğrenmesini, sınırlı veri durumlarında

daha iyi performans sergilemesini ve genel olarak daha etkili ve geniş kapsamlı bir öğrenme sürecini destekler. Transfer öğrenme, geleneksel öğrenme yöntemlerine kıyasla öğrenme süreçlerini optimize ederek, verimliliği artırır ve farklı görevler arasında bilgi paylaşımını teşvik eder. Bu bağlamda, transfer öğrenme, makine öğrenme alanındaki başarıyı artıran ve çeşitli uygulama alanlarında büyük bir kolaylık sağlayan önemli bir öğrenme paradigması olarak öne çıkar.

3.2.2. Evrimsel sinir ağı (Convolutional Neural Network)

Evrimsel sinir ağı (CNN) ağ yapısı, ileri beslemeli bir sinir ağıdır. Temel olarak iki boyutlu görüntüler üzerinden özellikler çıkarmak için tasarlanmıştır. CNN, 1980 li yıllardan itibaren temelleri atılan bir bilgisayar bilimi olan makine öğrenmesinin bir alt dalıdır. Büyük veri setleriyle öğrenme sürecinde çok iyi sonuçlar veren bir kapasiteye sahiptir (Güler, Çınar, & Cengil, 2017). Özellikle görüntü işleme ve veri sınıflandırma gibi işlemlerde kullanılan ve büyük bir başarı gösteren derin öğrenme modellerinden biridir. 1990 yıllardan sonra işlenmeye ve eğitime başlanmıştır. Ama ne yazık ki ihtiyaç duyulan sistem ve donanım eksikliğinden dolayı gerekli ilgi ve başarıyı yakalayamamıştır. 1995 yılında ise görüntü işleme uygulamalarında kullanılmaya başlanmıştır (Lawrence , Giles , & Chung , 1996). Günümüzde ise birçok alanda başarılı sonuçlar veren CNN ağının temel yapısı aşağıda görülmektedir.

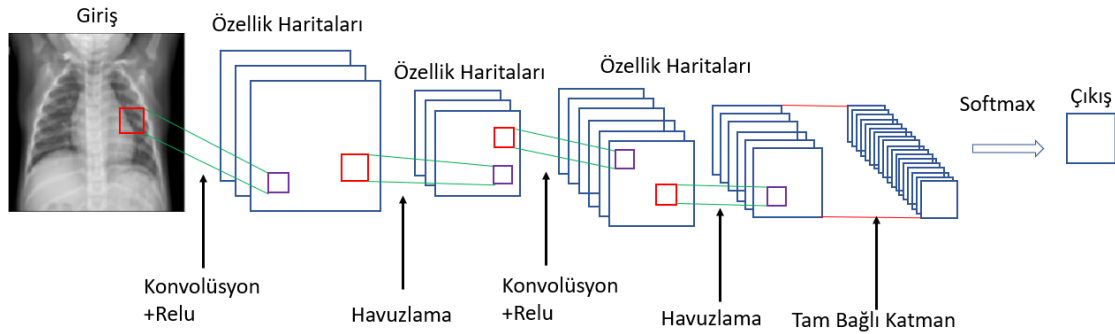


Şekil 3.4. CNN ağ yapısı (Hidaka & Akinori, 2017)

CNN'ler kendilerine has bir yapıyla verinin özelliklerini tanımlar. Böylece görsel veri işleme durumlarında sonuçlarını en iyi hale getirir. Tüm bu artı durumlarından dolayı CNN karmaşık veri analizi görevlerinde büyük bir öneme sahiptir (Szeliski, 2022). Bu model özellikle karmaşık matris verilerini işlemek için tasarlanmış ve birçok farklı alanda kullanılan bir mimaridir. Görüntü sınıflandırma, bilgisayar ve görü madenciliği, nesne

tanıma, sınıflandırma ile hastalık tespiti gibi alanlarda sıklıkla kullanılır. Temel işlemi, bir matris çarpım uygulaması olarak kabul edilebilir. Bu süreç, çekirdek matrisinin girdi matrisi üzerinde kaydırılarak ve çarpılarak evrişim sürecinin tamamlanmasıyla olur (Ian, Yoshua, & Aaron, 2018).

CNN genel anlamda görsel verileri işlemek için tasarlanmıştır. CNN öğrenme modeli, girdi görüntülerini alır ve nesnelere önem atar, bu nesneleri birbirinden ayırt edilebilir duruma getirir ve aynı zamanda nesneler arasındaki ilişkileri çıkarır. Bu işlem, evrişim, havuzlama, tam bağlantı ve sınıflandırma katmanları sayesinde gerçekleştirilir. Böylece görüntülerin ayırt edici özellikleri tespit edilir ve yapay öğrenme işlemi CNN aracılığıyla gerçekleştirilir (Jauro, vd. 2020). Şekil 3.24'te temel bir CNN yapısı gösterilmiştir



Şekil 3.5. LeCun tarafından 1989 yılında ortaya atılan LeNet isimli evrişimsel sinir ağı modeli (Le Cun, vd. 1989)

Temel çalışma prensibi özel filtreler kullanarak görüntülerin farklı özelliklerini belirlemek ve işlemektir. Filtreler genellikle kenarları, köşeleri ve diğer önemli özellikleri tanımak için eğitilir. Devamında, filtrelerden elde edilen sonuçlar katmanlar arasında iletilerek daha karmaşık sonuçların oluşturulması sağlanır. CNN, veriyi işlemek üzere art arda dizilmiş birden fazla eğitilmiş katmandan oluşur. İlk katman giriş katmanıdır ve bu katmanın giriş verileri, ağı besler. Ardından, konvolüsyon katmanında veriler çeşitli öz niteliklere göre işlenir. Bu katman, görüntüde özelliklerin tanımlanması gibi görevlerde önemli bir yere sahiptir. Daha sonra, sınıflandırma ve sonuç çıkarımı, çıkış katmanında gerçekleşir ve burada nihai sonuçlar elde edilir. Ağın genel yapısının uygun hale getirilmesi ve ağın öğrenilmesi süreci, geri yayılım (backpropagation) adı verilen bir algoritma ile gerçekleştirilir. Bu algoritma, ağın tahminlerini gerçek etiketlerle karşılaştırır, hataları hesaplar ve ardından ağıdaki parametreleri bu hataları en aza

indirmek için yeniler. Bu süreç, ağın daha iyi özellikler öğrenmesini ve daha doğru sınıflandırmalar yapmasını sağlar. Böylece CNN, görüntülerdeki özellikleri tanıma ve sınıflandırma konularında büyük bir başarı elde eder ve uygulama alanları arasında görüntü sınıflandırma, nesne tanıma, tıbbi görüntü analizi, otonom araçlar ve daha birçok özelleştirilmiş görev noktasında büyük kolaylıklar sağlar (Zhu, vd. 2017). Aynı zamanda nesne tespiti uygulamalarında, görüntülerdeki nesnelerin konumlarını belirlemek için etkili bir şekilde kullanılırlar (Gao, vd. 2017). Evrişim işlemi, bir görüntü filtresinin giriş görüntüsü üzerinde kendine has bir hesaplama yöntemi olan konvolüsyon işlemi olarak adlandırılır. Mevcut görüntülerdeki nitelik ve önemli özellikleri belirlemede çok önemli bir yere sahip olan bu yöntem görüntü işleme ve özellikle derin öğrenme alanlarında çoğunlukla kullanılmaktadır.

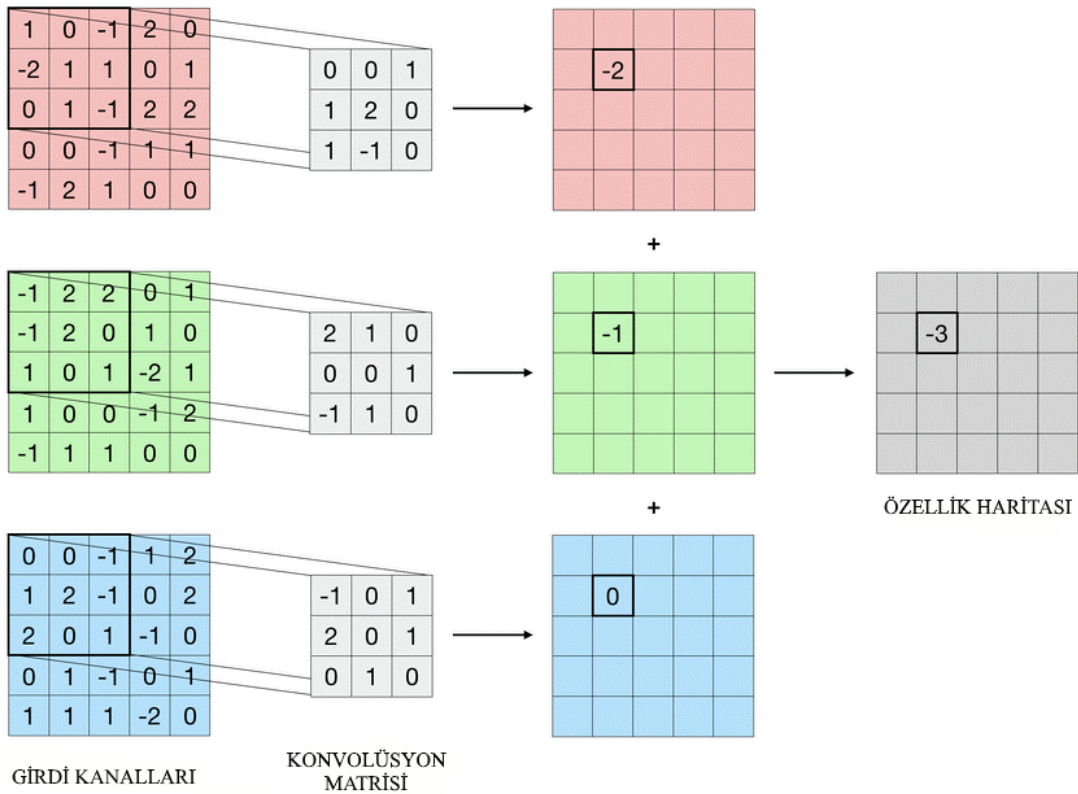
Evrişimli sinir ağları, genellikle dört temel katmandan oluşur. İlk katman, "Evrişim Katmanı", veri üzerinde özellikleri vurgular ve bu sayede örüntüleri tanımamızda kolaylık sağlar. Ardından, "Havuzlama Katmanı", verileri özgün bir şekilde özetleyerek boyutunu azaltarak işlem süresini ve maliyetini düşürmüş olur. Üçüncü olarak, "Düzeltilme Katmanı", hataları azaltmaya yöneliktir ve bu eğitimi daha düzgün ve devamlı hale getirir. Son olarak, "Tam Bağlı Katman", tüm bu özellikleri birleştirerek sonuçları çıkarır. Bu katmanlar bir araya gelerek, evrişimli sinir ağlarının çekirdeği hükmünde olan temel yapısını oluşturur. Böylelikle karmaşık görsel verilerdeki özelliklerin öğrenilmesine olanak sağlar.

Giriş Katmanı

CNN yapısında, giriş katmanı ağın başlama yeridir ve bu katmanda veri ham haliyle ağı verilir. Başka bir deyişle modelin kullanacağı verinin başlangıç noktasıdır. Modelin mimari yapısına bağlı olarak girdi değerleri özenle seçilmeli ve düzenlenmelidir, çünkü burada modelin veriyi işlemeye başlayacağı temel yerdir ve ilk adımdır. Veri setini oluşturan temel verilerin temsil edildiği bu kısım modelin eğitimi için kullanılan girdi değerlerini yapısında barındırır. Bu veriler, ağın öğrenme sürecinde kullanılır ve modelin veri üzerinde belirli işlemleri gerçekleştirmesi için kullanılır. Giriş katmanı, verinin kabul edildiği katman olup, ardından verinin ağın geri kalan kısmına aktarıldığı yerdir. Bu nedenle giriş katmanı, CNN'nin önemli bir parçasıdır ve modelin veriyle etkileşime geçtiği noktadır.

Konvolüsyon Katmanı

Konvolüsyon katmanları, Convolutional Neural Network (CNN)'nin temel yapı taşlarıdır ve özel filtreler kullanarak giriş verilerini işlerler. Bu katmanlar, başlangıçta temel özellikleri çıkarmak için girdi verilerini işler, örneğin kenarlar ve kıvrımlar gibi geometrik özellikleri tanımlar. Daha sonra, sonraki katmanlar daha karmaşık özelliklere odaklanır ve nesnelerin şekil ve boyutunu tanımak gibi daha yüksek düzeyde sonuçlara ulaşırlar. Bu sayede CNN, özellikle görüntü tanıma ve diğer görsel işleme görevlerinde oldukça başarılıdır. Konvolüsyon katmanı, verileri özel filtrelerle işleyerek özellik matrisini oluşturur. Bu matris, özelliklerin çıkarılmasına yardımcı olur ve giriş verilerindeki önemli bilgileri vurgular. Özellikle görüntü işleme ve görüntü tanıma gibi çalışmalarda, konvolüsyon katmanının ayrıntıları tespit etmede kritik bir rolü vardır. Öğrenme süreci, CNN'nin konvolüsyon katmanlarında temel olarak gerçekleşir. Bu süreçte, ağı parametreleri, özellik matrisinden elde edilen hata değerlerini minimize etmek için güncellenir. Bu, ağı daha karmaşık özellikleri öğrenmesini ve veri üzerinde daha etkili bir şekilde çalışmasını sağlar. Konvolüsyon katmanları, ağı giriş verilerindeki özellikleri hiyerarşik olarak öğrenmesine ve bu özellikleri genelleştirmesine yardımcı olarak, CNN'nin başarılı performansını sağlar.



Şekil 3.6. Evrişim katmanı özellik çıkarımı (Kuchera et al., 2019)

Yukarıdaki şekilden de anlaşıldığı üzere evrişim süreci şu şekilde gerçekleşir. İlk olarak, 5x5 boyutundaki bir girdi verisi matrisi ve 3x3 boyutlu bir filtre kullanılır. Bu filtre, girdi matrisinin üzerinde kaydırma yaparken her bir konumda çapraz çarpım işlemi yapar. Yapılan çarpımlar neticesinde elde edilen değerler toplanır, bu işlem 3x3'lük bir pencere boyunca tekrarlanır ve böylece bir özellik haritası oluşturulur. Özellik haritasının bir diğer işlevi ise görüntüdeki farklı özellikleri temsil eden ağırlık değerlerini içeriyor olmasıdır. Bu özellikler, kenar tespiti, renk değişiklikleri gibi görsel özellikleri yakalamak için kullanılır. Evrişim katmanları, bu özellikleri derin öğrenme ağlarının temel yapısı olarak kullanarak karmaşık görsel tanıma görevlerini başarıyla gerçekleştirir.

Aktivasyon Katmanı

CNN'nin temel taşlarından biri olan aktivasyon fonksiyonları, nöronların işleyeceği giriş sinyallerini düzenleyen önemli matematiksel araçlardır. Bu fonksiyonlar, negatif değerleri sıfırlayarak veya başka bir şekilde değiştirerek ağın daha verimli çalışmasını sağlarlar. Bu süreç sonucunda nöronlar, işlenmiş bu çıkış sinyalini bir sonraki katmana veya tam bağlantı katmanına ileterek ağın öğrenme sürecini sürdürürler. Aktivasyon fonksiyonları, nöronların ağ içerisindeki önemini ve bilgileri nasıl işleyeceğini belirleyen kritik bileşenlerdir. Aktivasyon fonksiyonları, sinir ağındaki veri kullanma ve bilgi işleme sürecini yönlendirmek için, nöronların etkinleştirilip etkinleştirilmediğini belirler. Bu fonksiyonlar nöronların çıkışlarını belirlerken, giriş verilerini dönüştürür ve sigmoid, ReLU (Rectified Linear Unit) veya Hiperbolik Tanjant gibi matematiksel işlemleri kullanır. Tüm bu artı özelliklerinden dolayı aktivasyon fonksiyonları, sinir ağlarının öğrenme yeteneğini etkileyen önemli bir yapıdır. (Glorot, Bordes, & Bengio, 2011). Aktivasyon fonksiyonlarının önemli diğer bir rolü ise, nöron çıktılarını belirli aralıklara sıkıştırmak ve model performansını artırmaktır. Doğrusal, ReLU, Sigmoid, Hiperbolik Tanjant, Leaky ReLU, Swish ve Softmax gibi çeşitli aktivasyon fonksiyonları bulunmaktadır. Problemin fitri yapısı seçilen aktivasyon fonksiyonunu belirler. Böylece problem, türevlenebilir, hassas, doğrusal olmayan ve belirli aralıklarda olmalıdır. Hassas bir aktivasyon fonksiyonu, matris işlemlerinin uygulanacağı görüntü işleme gibi alanlarda özellikle önemlidir. Bu durumun aksi takdirde olması öğrenme hızını azaltabilir. Aşağıda CNN ve derin öğrenme mimarisinde kullanılan aktivasyon fonksiyonlarına ait matematiksel denklemleri ve grafikleri

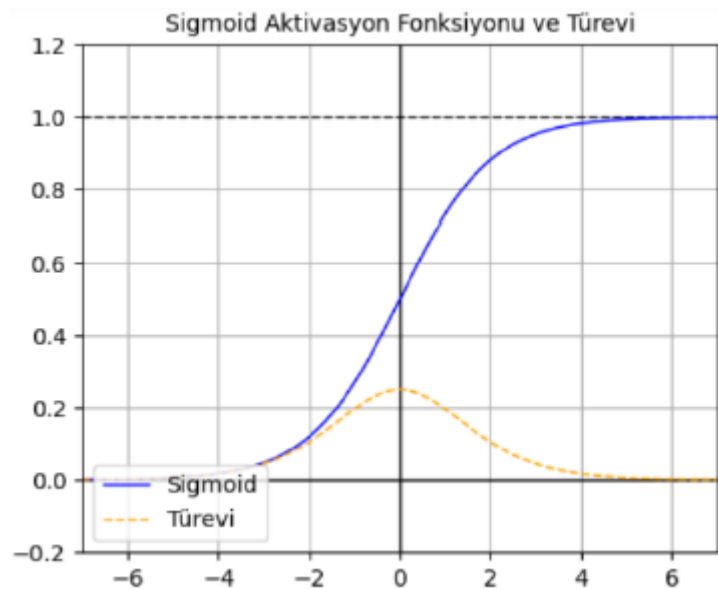
verilmiştir. Doğrusal aktivasyon fonksiyonu, sabit bir çıktı değerine sahiptir. Bu bize geri yayılım kullanımı sırasında hata oranının sabit olacağı için istenilen çıktıyı almayı ve geri yayılımın kullanımının bozulmasına sebebiyet vermektedir. Matematiksel ifadesi denklem 3.1’de gösterilmektedir. Denklemden gösterilen a değeri herhangi sabit bir sayı değeri olabilir, seçimi tercihe bağlıdır (ARSLAN, 2022, s. 39-41).

$$f(x) = ax \quad (3.1)$$

Sigmoid aktivasyon fonksiyonu, grafiksel olarak bir S harfine benzeyen bir eğri şeklindedir ve doğrusal olmayan bir aktivasyon fonksiyonudur. Yapay sinir ağlarının katmanlarında sıklıkla kullanılan bu fonksiyon, girdi değerlerini 0 ile 1 arasında sıkıştırarak, çıkışları olasılık benzeri bir formda sunar. Bu özellik, özellikle ikili sınıflandırma problemlerinde tercih edilir. Sigmoid aktivasyon fonksiyonu, ağın non-lineer özellikler eklemesine yardımcı olarak karmaşık ilişkileri öğrenme yeteneğini artırır. 0-1 aralığında çıktı değerlerine sahip olmaktadır. Aynı zamanda türevlenebilirdir (Sharma, Sharma, & Athaiya, 2017).

Şekil 3.6’da Sigmoid fonksiyonunun türevlenebilir bir aktivasyon fonksiyonu olduğu gösterilmektedir. Bu sayede öğrenme işlemini daha seri ve etkili gerçekleştirebilmektedir.

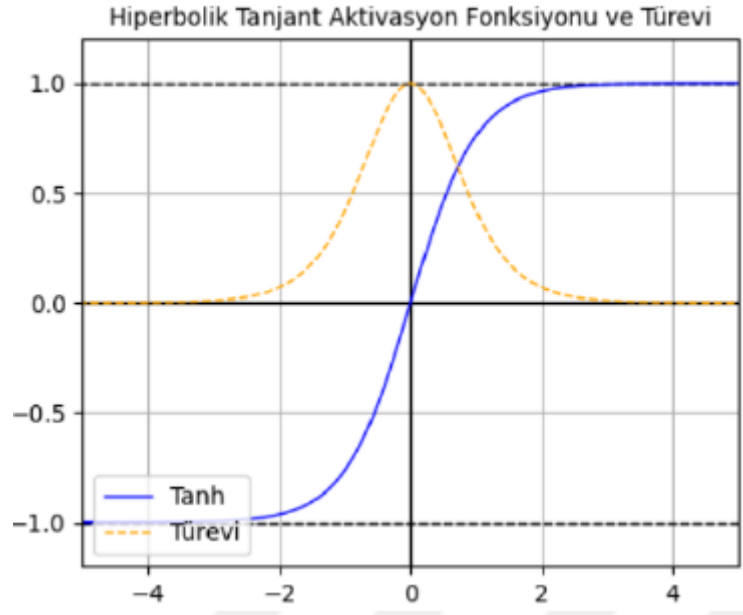
$$f(x)_{\text{sigm}} = \frac{1}{(1+e^x)} \quad (3.2)$$



Şekil 3.7. Sigmoid aktivasyon fonksiyonu ve türevi grafiği

Hiperbolik Tanjant (Tanh): Hiperbolik tanjant fonksiyonu, -1 ile 1 aralığında değerler alır ve 0 merkezinden geçen bir S şekilli eğri oluşturur. Denklemi 3.3 ile ifade edilmiştir ve her x değeri için bir çıktı üretir. Hiperbolik tanjant fonksiyonu, sigmoid fonksiyonuna göre daha simetrik bir yapıya sahiptir. Bu aktivasyon fonksiyonu $[-1, +1]$ aralığında çıktı üretir ve özellikle ikili sınıflandırma problemlerinde kullanılır. Sigmoid fonksiyonuna göre daha geniş bir değer aralığına sahip olması, sinir ağlarının daha hızlı öğrenmesine ve verimliliğin artmasına olanak sağlar. Bu özelliği sayesinde, hiperbolik tanjant aktivasyon fonksiyonu, sinir ağlarının daha geniş bir spektrumda ve daha hızlı öğrenmesini sağlayarak performansını artırır.

$$f(x)_{\tanh} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3.3)$$

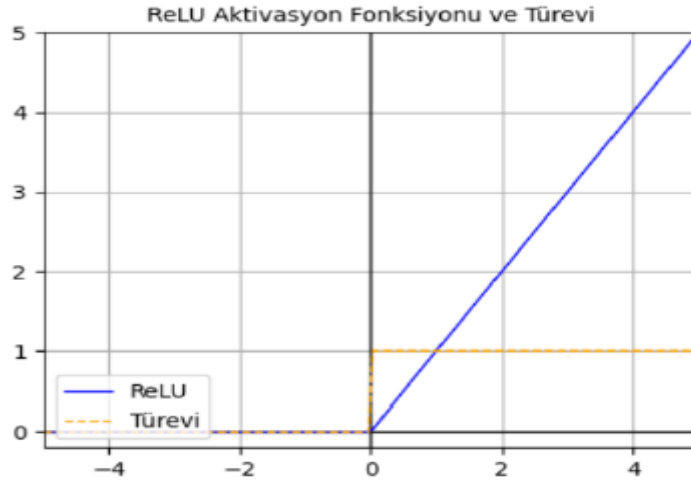


Şekil 3.8. Tanh aktivasyon fonksiyonu ve türevi

ReLU: Leaky ReLU, doğrusal olmayan ancak doğrultulmuş doğrusal birim olarak bilinen bir aktivasyon fonksiyonudur ve yapay sinir ağlarında sıkça kullanılan bir türdür. Leaky ReLU fonksiyonu, 0'dan küçük tüm giriş değerleri için 0 değerini alırken, x pozitif değeri için de x değerini alır. Matematiksel olarak ifade edildiğinde, denklem 3.4'te gösterilmiştir. Şekil 3.8'deki grafik incelendiğinde, türevin 0'dan küçük olan değerlerde 0'da sabit kaldığı görülmektedir.

Leaky ReLU, özellikle geleneksel ReLU fonksiyonunun sorunlarına bir çözüm olarak ortaya çıkar. Geleneksel ReLU fonksiyonu, giriş değeri negatif olduğunda sıfır çıktı üretir ve bu durum, öğrenme sürecinde bazı nöronların pasif hale gelmesine neden olabilir. Leaky ReLU, bu sorunu aşmak için negatif giriş değerlerinde küçük bir eğimle devam ederek, negatif değerler için de bir miktar bilgi taşımalarını sağlar. Bu özelliği sayesinde, Leaky ReLU fonksiyonu, sinir ağlarının öğrenme sürecini iyileştirmeye ve daha etkili bir şekilde bilgi işlemesine katkıda bulunur (Türkoğlu vd., 2021, s. 334-345).

$$\text{ReLU}(x) = (x)^+ = \max(0, x) \quad (3.4)$$

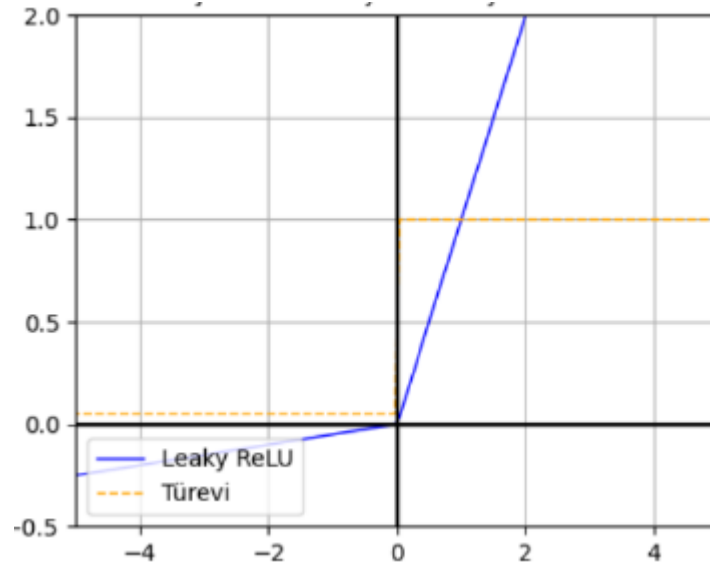


Şekil 3.9. ReLU aktivasyon fonksiyonu ve türevi

Leaky ReLU, ReLU aktivasyon fonksiyonuna göre daha geniş bir öğrenme oranına ve işlem yoğunluğuna sahiptir. Şekil 3.9'da bulunan ReLU fonksiyonunun türevidir incelendiğinde negatif değerlerde sabit bir değer olan 0'a sahiptir. Öte yandan, Leaky ReLU fonksiyonunun türevidir incelendiğinde negatif değerlere yakınsayan bir eğime sahiptir. Leaky ReLU'nun negatif değerlerde küçük bir eğime sahip olması, bu fonksiyonun ağırlık öğrenme sürecinde negatif giriş değerlerine daha iyi tepki vermesini sağlar. Bu, özellikle ReLU'nun "dying ReLU" problemi olarak adlandırılan durumu aşmada yardımcı olabilir. "Dying ReLU" problemi, ReLU fonksiyonunun negatif girişlerle sıfır çıkış üretmesi nedeniyle bazı nöronların eğitim sırasında hiç aktive olmamasına ve bu nöronların öğrenmeye katkıda bulunmamasına neden olabilir. Leaky

ReLU'nun bu soruna karşı daha esnek bir davranış sergilemesi, özellikle öğrenme sürecini daha etkili hale getirir.

$$f(x)_{LeakyReLU(x)} = f(x) \begin{cases} x, & x \geq 0 \\ mx, & x \leq 0 \end{cases} \quad (3.5)$$



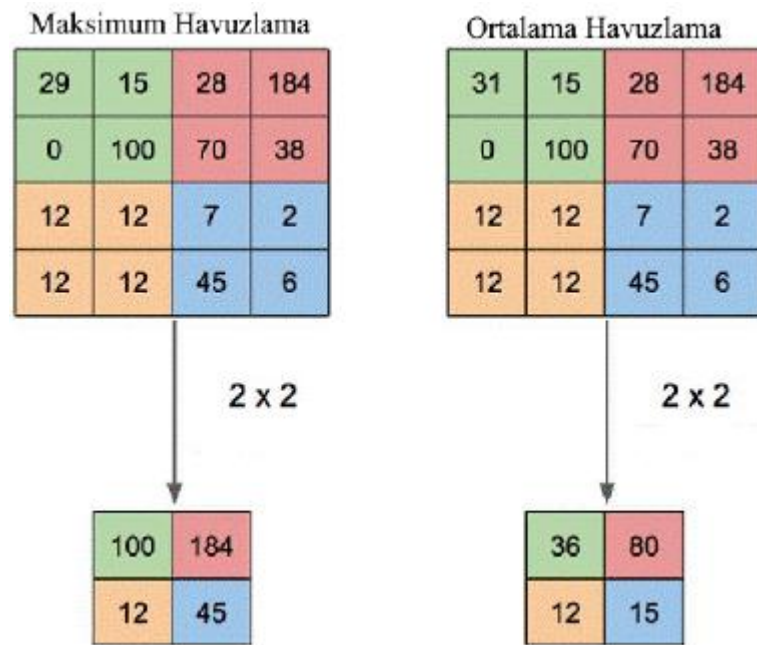
Şekil 3.10. Leaky ReLU aktivasyon fonksiyonu ve türevi

Swish (A Self-Gated): Swish aktivasyon fonksiyonu, negatif bölgelerde de değer alabilen ve bu bölgelerde doğrusal olmayan bir davranış sergileyen bir fonksiyondur. Diğer birçok aktivasyon fonksiyonu genellikle tek düze bir davranış gösterirken, Swish fonksiyonu negatif bölgelerde de değer alarak daha esnek bir özellik sergiler. Swish, özellikle derin öğrenme modellerinde kullanılmak üzere tasarlanmıştır ve giriş arttıkça çıktının düşmemesi gibi bir özelliğe sahiptir. Swish fonksiyonunun eğrisi incelendiğinde, girdi arttıkça çıktının düşmemesi ve negatif değerlerde de değer alması görülebilir. Bu özellik, Swish'in öğrenme sürecinde daha etkili olmasını sağlar. Swish aktivasyon fonksiyonu, özellikle rekurrent sinir ağları (RNN) gibi modellerde iyi performans gösterebilir ve girişin geniş bir aralıkta değişebileceği durumlarda tercih edilebilir.

$$f(x)_{Swish} = 2x\sigma(\beta x) = \begin{cases} \beta = 0 \text{ için } f(x) = x \\ \beta \rightarrow \infty \text{ için } f(x) = 2 \max(0, x) \end{cases} \quad (3.6)$$

Havuzlama (pooling) katmanı

Havuzlama katmanı, büyük veri kümesini en uygun hale getirerek ağırlık sayısını azaltır. Böylece öğrenme sürecini daha hızlı hale getirir ve maliyeti düşürerek işlevsellik kazandırır. Bu katmanda herhangi bir eğitim gerçekleşmez. Burada özelliklerin korunması sürdürülür ve eğitim işlemi evrişim katmanından sonra gerçekleşir. Havuzlama katmanı sayesinde eğitim süresi kısılır, bu da zaman ve kaynak tasarrufu sağlar. Havuzlama işlemi, iki farklı yöntemle gerçekleştirilebilir. İlk yöntem, uygulanan filtrenin değerlerinin ortalamasının alındığı ortalama havuzlamadır. Diğer bir yöntem ise, filtrenin içindeki en büyük değerin seçilip kullanıldığı maksimum havuzlamadır. Bu iki yöntem, veri işleme ve sinyal analizi gibi birçok alanda kullanılır. 2x2 lik boyutta maksimum ve minimum havuzlama katmanları, görüntü matrislerinde maksimum veya ortalama değerlerin alınarak boyutun küçülmesini ve işlem maliyetinin azaltılmasını sağlayan önemli bir bileşendir. Bu katmanda kaydırma (stride) sayısı, girdi verisi üzerinde uygulanacak filtrenin kaç birim kaydırılacağını belirler. Ayrıca havuzlama katmanının boyutunu ayarlamak, bu işlemi hassas bir şekilde kontrol etmemize yardımcı olur. Özellikle havuzlama katmanı, derin öğrenme ağlarının işlem hızını artırırken aynı zamanda öğrenme gerçekleştirilerek verinin önemli özelliklerini özetler (Ali & Hanbay, 2018, s. 1395-1408).

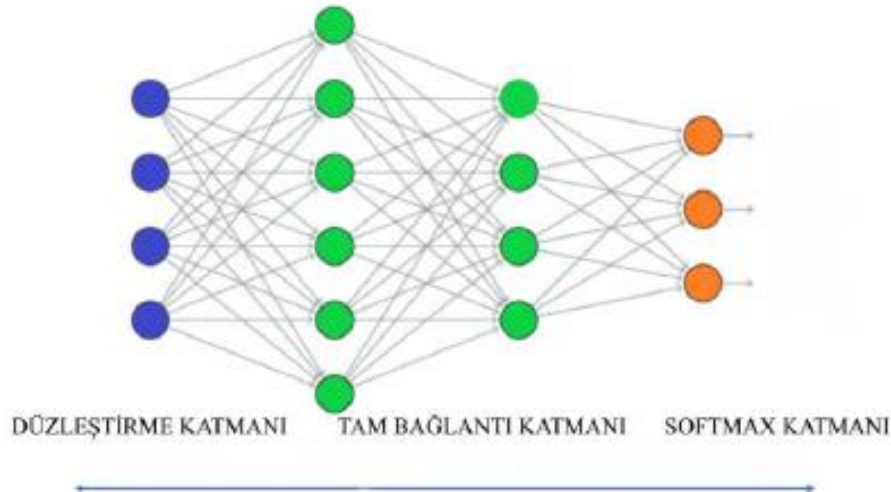


Şekil 3.11. 2x2 Maksimum havuzlama, ortalama havuzlama (Yani, Irawan, & Setianingsih, 2019)

Şekil 3.10'da maksimum havuzlama ve ortalama havuzlama yöntemleri gösterilmektedir. Veri matrisinde bulunan indisler üzerinde 2×2 havuzlama uygulandığında matris verisinde bulunan maksimum değerleri alınarak boyutu indirgemekte ve işlem yükünü azaltmaktadır. Ortalama havuzlama yönteminde ise yeni bir matris oluşturmak üzere 2×2 boyutunda matris uygulanan girdi verisinde bulunan değerlerin ortalaması alınmaktadır.

Tam bağlantılı katman (FC)

Tam bağlantılı katman, önceki evrişim ve havuzlama katmanlarının çıktılarını bir matris olarak alır ve bu matrisi düzleştirme işlemiyle vektör haline getirir. Şekil 3.11'de tam bağlantılı katmana ait görsel bulunmaktadır.

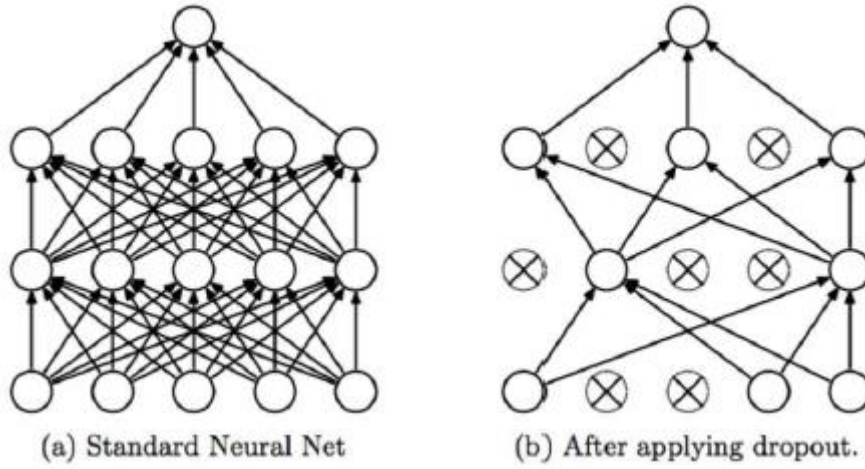


Şekil 3.12. Tam bağlantı katmanı

Havuzlama katmanının ardından yer alan tam bağlantılı katman, sınıflandırma işlemi için kullanılan son katmandır. Bu katmanda aktivasyon fonksiyonu kullanılmadan verilerin son halini alır.

Dropout katmanı

Dropout katmanı, eğitim sırasında rastgele düğümleri devre dışı bırakarak aşırı öğrenmeyi engellemeye yardımcı olan bir teknik olarak kullanılır. Özellikle sınırlı veriye sahip olduğunda veya uzun eğitim süreleri gerektiğinde, modelin eğitim verilerini ezberlemesini önler. Bu sayede, ağ genellemeyi öğrenir ve daha iyi sonuçlar elde eder.

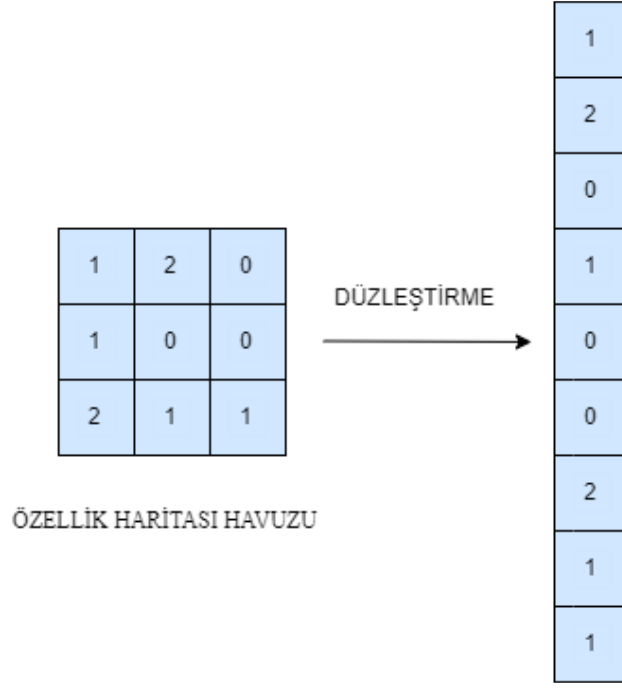


Şekil 3.13. Dropout uygulanan ve uygulanmayan katmanlar(Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014)

Şekil 3.12’de standart bir sinir ağı ve rastgele bazı düğümleri silinen ağa ait görseller bulunmaktadır. Dropout katmanı, aşırı öğrenmeyi engellemek için son derece önemlidir. Bu katman, modelin belirli nöronlarını rastgele devre dışı bırakarak çeşitli ağırlık ayarlarının öğrenilmesini teşvik eder. Sonuç olarak, modelin genel performansı artar ve daha sağlam bir öğrenme süreci elde edilir.

Düzleştirme katmanı

CNN, işleme adımlarından biri olarak son havuzlama katmanını kullanır. Bu katmandan sonra özellik haritaları elde eder. Bu özellik haritaları daha sonra birleştirilir ve düzleştirilme işlemi uygulanarak vektör haline getirilir. Bu sayede, görüntü verileri daha işlenebilir bir şekilde getirilir ve daha sonraki sinir ağı katmanlarına iletilir (Doğan & Türkoğlu, 2018, s. 10-21).



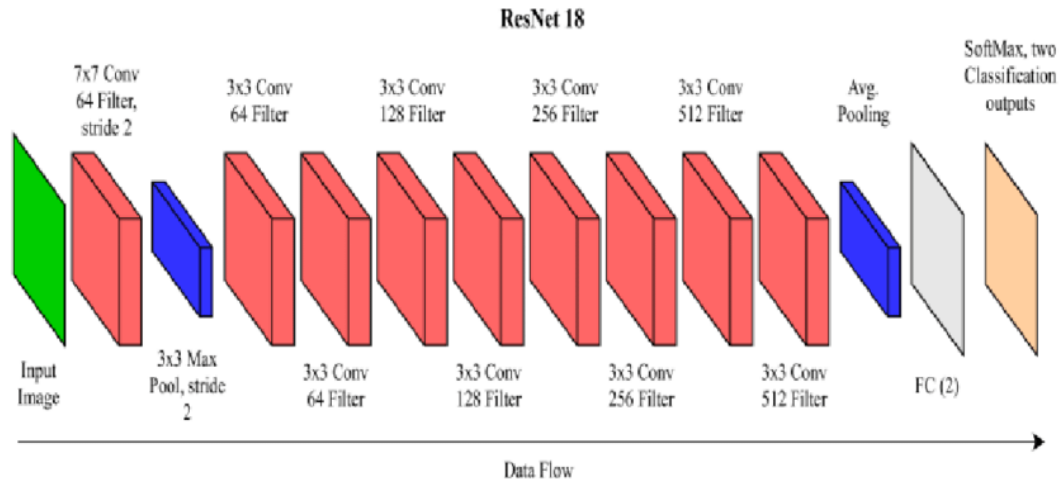
Şekil 3.14. Düzleştirme katmanı

Şekil 3.13'te düzleştirilmiş matris tek boyuta indirgenmiştir. Matris, tam bağlı bir katman kullanılarak işlenir ve istenilen boyuta dönüştürülür. Sonuç olarak, çıktı boyutu sınıf sayısı ile uyumlu hale getirilir.

3.2.3. ResNET18

He ve arkadaşları tarafından 2016 yılında geliştirilen ResNet (Residual Network) derin öğrenme modellerinden biridir. Derin öğrenme eğitiminde ortaya çıkan zorluklar He ve arkadaşlarını yeni modeller üzerinde çalışmalar yapmaya zorlamıştır. Karşılarına çıkan bu zorlukların üstesinden gelmek amacıyla ResNet mimarisini oluşturdular. Derin öğrenme eğitimi genel anlamda çok uzun süren bir süreçtir. Kullanılan katman sayıları sınırlı ve yetersiz olabilir. ResNet tam da bu noktada devreye girerek atlama bağlantısı veya kısa yol gibi uygulama çözümleri getirerek işin daha hızlı ve sağlıklı gerçekleşmesini sağlamaktadır. Ayrıca ResNet modeli, mimari daha derinleştikçe performansın azalmaması gibi bir özelliğiyle diğer mimari modellere göre çok daha büyük avantaj sağlamaktadır. Bununla beraber ağırları eğitme yeteneği daha iyi olmakla, hesaplama işlemlerini de daha hafif yapmaktadır. ResNet modeli, mimariler arasında ReLU ve toplu normalizasyon içeren iki veya üç katmandan oluşan atlama bağlantıları ile uygulanır. Uygulamalar esnasında varılan diğer bir gerçek ise, ResNet modelinin daha

önce kullanılan diğer modellere göre görüntü sınıflandırmasında daha iyi sonuçlar vermesidir. Sonuç olarak ResNet tarafından elde edilen görüntü özelliklerinin iyi olduğu görülmektedir. ResNet' de, bir giriş katmanı ve bir çıkış katmanı arasına birden fazla katmanın uygulanmasıyla artıklı öğrenme tercih edilir. ResNet'deki artıklı blok, giriş veri boyutları çıkış veri boyutlarına eşit olduğunda başarılı olur. Ayrıca her ResNet bloğu, iki katmandan (ResNet-18 ve ResNet-34 ağırları için) veya üç katmandan (ResNet-50 ve ResNet-101 ağırları için) oluşur (Sarwindaa, Paradisaa, & Alhadi , 2021, s. 424-427). Şekil 3.14. te ResNet18 modelinin genel yapısı görülmektedir.



Şekil 3.15. ResNet18 modeli

ResNet-18, derin öğrenme ve CNN alanında oldukça önemli bir modeldir. 2015 yılında Microsoft Research tarafından geliştirilip kullanılmaya başlanılan ResNet-18 özellikle ImageNet kümesindeki görüntü sınıflandırma ve ayrıştırma görevlerinde gözle görülür büyük sonuçlar ortaya çıkarmıştır. ResNet, kavramına dayanır. Geleneksel sinir ağırları, her katmandan geçen verileri bir sonraki katmana ileterek işlerler. Bu durum her ne kadar işlevsel olsa da derin ağlarda bu yöntem, gradyan kaybı gibi sorunlara yol açabilir. ResNet, farklı bir yaklaşımla bu sorunu çözmeye yönelmiştir. Katmanlardan geçen verileri bir sonraki katmana iletmek yerine "atlam bağlantıları" veya "rezidüel bağlantılar" olarak adlandırılan özel bağlantıları tanıtır ve gelmiş olduğu katmana giriş verisini doğrudan ileterek bu sorunu çözer. ResNet-18, bu rezidüel denilen özel bağlantıları içeren bir CNN modelidir. Ve toplam olarak 18 katmandan oluşur. Her katman bir dizi rezidüel bağlantı içerir ve ayrıca iki veya üç konvolüsyon katmanına sahiptir. Sahip olduğu rezidüel bağlantıları, gradyanların daha düzgün bir şekilde geriye yayılmasına olanak tanır ve daha derin ağların eğitilmesini mümkün kılar.

ResNet-18, birçok öğrenme işlevselliğine sahiptir fakat genellikle transfer öğrenme ve çeşitli görüntü işleme görevlerinde kullanılır. Hem hızlı eğitilebilir olması hem de iyi performans gösteriyor olması, bu modelin birçok uygulama için popüler olmasını sağlamıştır. ResNet-18, daha büyük ResNet varyantlarının temelini atmıştır. Ve derin öğrenme alanının ilerlemesinde önemli bir ivme görevi üstlenmiştir.

3.2.4. ResNET50

ResNet, residual öğrenme konsepti altında önerilen bir derin sinir ağıdır. ResNet geleneksel derin sinir ağlarından farklı olarak, konvolüsyonel katmanlar arasında atlama bağlantılar ekleyerek, bilgiyi çoklu örtülü katmanlar boyunca yayılmasına izin verir. Bu da geleneksel derin sinir ağlarında ortaya çıkan gradyan kaybolma ve ağ bozulma sorunlarını asgari düzeye indirir ve mevcut olan ağı çok daha fazla diyebileceğimiz onlarca hatta yüzlerce katmanlı hale getirir. Günümüzde ResNet-50, birçok alanda olmakla birlikte yüz tanıma ve otonom sürüş gibi çeşitli mühendislik alanlarında kayda değer bir oranda uygulanmaktadır. Özellikle geleneksel derin sinir ağlarına nazaran parametre sayısı ve modelin performansı göz önüne alındığında, yapılan çoğu çalışmada alt model ResNet-50 tercih edilmeye başlanmıştır (Wu, Ying, & Zhou, 2023, s. 1-10).

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

Şekil 3.16. ResNet-50'nin ağ yapısı

Şekil 3.15' te görüldüğü gibi, ResNet-50 mimarisi aşağıdaki öğeleri içerir.

- 7 * 7 çekirdek boyutuna ve 64 farklı çekirdeğe sahip bir kıvrım ve tümü 2 boyutunda adım 1 katman verir.
- Ardından, adım boyutu 2 olan maksimum havuzlama görünüyor.

- Bir sonraki evrişimde $1 * 1,64$ çekirdek, ardından $3 * 3,64$ çekirdek ve son olarak $1 * 1,256$ çekirdek vardır. Bu üç katman toplam 3 kez tekrarlanır ve bu adımda 9 katman verir.
- Daha sonra $1 * 1,128$ çekirdeğini, ardından $3 * 3,128$ çekirdeğini ve son olarak $1 * 1,512$ çekirdeği görölüyor, bu adım 4 kez tekrarlanır, bu da bu adımda 12 katman verir.
- Bundan sonra $1*1,256$ 'lık bir çekirdek ve $3*3,256$ ve $1*1,1024$ 'lük iki çekirdek daha gelir ve bu 6 kez tekrarlanarak toplam 18 katman verir.
- Ve sonra yine $1*1,512$ 'lik bir çekirdek, iki tane daha $3*3,512$ ve $1*1,2048$ ve bu 3 kez tekrarlanarak toplam 9 katman verir.
- Bundan sonra ortalama bir havuz yapılı ve bunu 1000 düğüm içeren tamamen bağlı bir katman ve sonunda bir softmax işlevi ile sonlandırılır, bu da 1 katman verir. Gerçekte aktivasyon fonksiyonları ve maksimum/ortalama havuzlama katmanları sayılmaz. Yani bunun toplamı $1 + 9 + 12 + 18 + 9 + 1 = 50$ katmanlı Derin Evrişimli bir ağ verir.

Bu çalışmada, ResNet-50 derin öğrenme modeli, Squeeze-and-Excitation (SE) dikkat modülü ile güçlendirilerek kanallar arasındaki özelliklerin füzyonunu ve önemini ayarlamak üzere yeniden kalibre edilmiştir. Ayrıca, modelin performansını artırmak amacıyla aktivasyon fonksiyonları Swish ile değiştirilmiş ve eğitim için yeni bir Ranger optimizyer kullanılmıştır. Transfer öğrenme ile desteklenen TCNN(ResNet-50), hata tespiti için eğitilmiş ImageNet ResNet-50'yi özellik çıkarıcısı olarak kullanmaktadır. Çalışmanın başlangıcında, zaman etkisi sinyalleri RGB görüntü formatına dönüştürmek için özel bir sinyal-görüntü yöntemi geliştirilmiş ve bu format ResNet-50'nin girdi veri tipi olarak kullanılmıştır. Daha sonra, önerilen yeni TCNN(ResNet-50) mimarisi, KAT veri merkezinin sağladığı yatak hasarı, Case Western Reserve Üniversitesi'nin motor yatak ve kendinden emişli santrifüj pompa veri setleri üzerinde test edilmiştir. Elde edilen sonuçlar, TCNN(ResNet-50)'nin diğer derin öğrenme modelleri ve geleneksel yöntemlere göre daha yüksek tahmin doğruluklarına sahip olduğunu göstermektedir. Özellikle, TCNN(ResNet-50)'nin tahmin doğrulukları sırasıyla $\%98.95 \pm 0.0074$, $\%99.99 \pm 0$ ve $\%99.20 \pm 0$ seviyelerindedir, bu da modelin üstün performansını doğrulamaktadır (Wen, Li, & Gao, 2020).

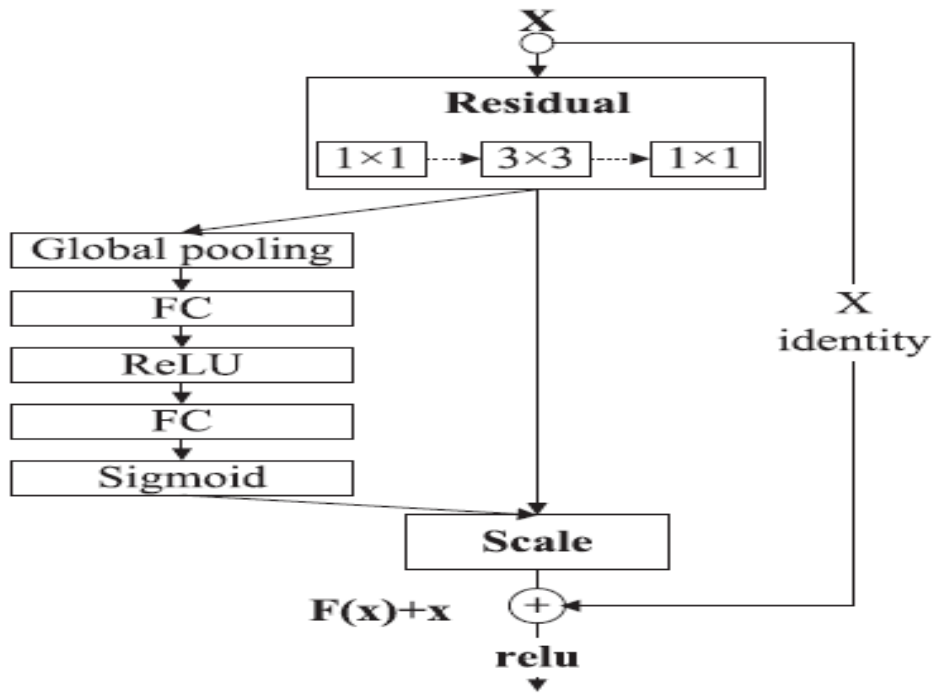
ResBlock katmanları, artık değerleri sonraki katmanlara iletmekle görevli katmanlardır. ResBlock katmanları, ResNet mimarisinde önceki katmanlarda öğrenilemeyen bilgiyi ve verileri yeni katmanlara taşır ve böylece bilgi kaybını en aza indirir. Her iki katmanın etkinleştirilmesinde, ağırlık katmanları ile Relu etkinleştirme kodu arasındaki atlama ile eklenen bu değerler, sistem hesabını değiştirir. İki katmandan oluşan Artıklık Bloğu, non-lineer Relu işlevini temsilen görev alır ve bu yapı, girdinin işlenmesi neticesinde elde edilen ikinci Relu değeri ile zenginleştirilerek çıktıyı elde eder (Aksoy & Salman, 2022, s. 2-4).

3.2.6. ResNET152

ResNet152, bir derin öğrenme modeli olup derin öğrenme alanında çok büyük yere ve öneme sahiptir. "Residual Network" veya "ResNet" adıyla bilinen derin öğrenme modeli familyasına ait bir öğrenme modelidir. Daha önce bahsi geçen derin öğrenme modellerinin yeterli olmadığı, daha derin ağ yapılarının eğitilmesinde aşılamayan güçlükleri gidermek için tasarlanan ve özellikle veri madenciliği ve görüntü işleme alanında büyük başarılar sağlamış bir modeldir. ResNet152 derin öğrenme modeli 152 katmandan oluşmaktadır. "Residual Network" model familyasının en derin alternatif öğrenme modellerindendir. İsmini de içerdiği katman sayısından almaktadır. Öncelikle "rezidüel bağlantılar" kavramını anlamakla başlayarak niye "rezidüel" olduğunu anlamaya çalışmak gerekir. "Rezidüel bağlantılar" işlenmekte olan katmanlardan yola çıkan veri ve çıktıların sekteye uğramadan doğrudan ardından gelen daha sonraki katmanlara eklenmesiyle meydana gelir. Bu bağlantılar olmadan normal şartlarda girdi verisi eklenme yapmadan bir katmandan diğerine geçer. Fakat rezidüel bağlantılarda bir katmandan çıkan veri diğer katmanın verisine yani çıktısına eklenir. Böylece hem ağ daha derin bir hal almış olur hem de öğrenme sürecinin uzamasının aksine kısalmasına imkân sağlar. Ayrıca ağ bağlantısının daha iyi gerçekleşmesini ve daha iyi ve sağlıklı sonuçlar üretmesini sağlar. Bununla beraber rezidüel bağlantılar, ağdaki gradyanların nizamlı ve sıralı bir düzen içinde geriye yayılmasına olanak sağlar ve gereksiz fazla öğrenmeyi minimize etmeye yardımcı olur. ResNet152, yukarıda anlatıldığı gibi işlenen rezidüel bağlantılarını kullanan ve toplamda 152 katmandan oluşan bir derin ağ yapısıdır. Bu yapı, katman sayısının artırılmasıyla modelin daha kompleks özellikleri öğrenmesini sağlar. Böylece çok daha karmaşık ve yoğun görüntü işleme durumlarında daha iyi performans ortaya koymasına ve akabinde daha iyi neticeler elde etmesine yardımcı olur.

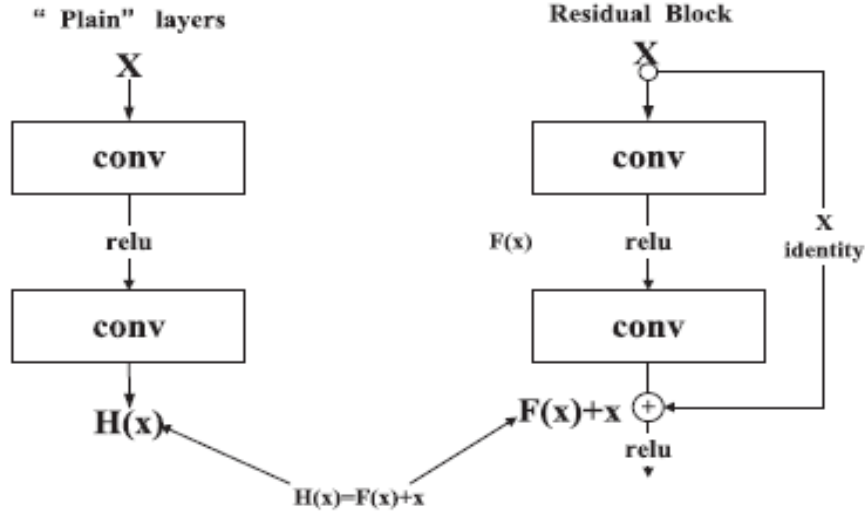
Sonuç olarak denilebilir ki ResNet152, derin sinir ağlarının daha derin hale getirilmesi ve daha etkili ve verimli bir şekilde eğitilmesi amacıyla geliştirilmiş bir modeldir. Görüntü sınıflandırma, nesne tanıma ve diğer görsel işleme görevlerinde üstün performans gösterme potansiyeline sahiptir (He, Zhang, Ren, & Sun, 2016).

Diğer bir derin sinir ağı SE-ResNet152 ise, Residual Neural Network (ResNet) mimarisine SE bloklarının birleştirilip uyum içinde çalışmalarını sağlayan bütünleşik devrelerden oluşan bir derin öğrenme modelidir. ResNet, derin sinir ağlarındaki kaybolan gradyanlar ve ağ bozulma sorunlarını önlemek için rezidüel bağlantılar kullanarak ağı derinliğini artıran ve ağı yeni bir açığa kazandıran bir yapıya sahiptir. Özetle SE-ResNet152, ResNet152 mimarisine SE bloklarının eklenmesiyle oluşturulur denilebilir.



Şekil 3.19. SE-ResNet152 model yapısı örnek diyagramı

SE blokları, farklı özellik sahibi olan tüm kanalların önemini belirleyen düzenekler kullanır. Özellik tanımlamada kullanılan bu bloklar, tüm özellik kanalının önemini yeniden düzenler ve kanalize eder. Daha sonra önemli bilgi içeren özellikleri vurgular ve daha az önemli olan özellikleri baskılar. Böylece kullanılan model, önemli bilgileri vurgulayarak ve az önemli bilgileri baskılayarak daha etkili bir şekilde özellikleri temsil eder. Sahip olduğu bu özelliği sayesinde SE-ResNet152, daha derin ve daha genelleyebilen özellikler elde eder. Bu da daha karmaşık veri kümelerinde daha iyi başarımlar ve sonuç sağlar.



Şekil 3.20. Residual(artık) öğrenmenin gösterimi

Residual öğrenme yönteminin temel çalışma prensibi Şekil 3.19'da gösterilmektedir. Residual fonksiyonun formülü ise aşağıdaki gibi ifade edilir.

$$F(x) = H(x) - x \quad (3.9)$$

Burada $F(x)$ residual fonksiyonu, x girdiyi ve $H(x)$ doğrudan eğitimle öğrenilen parametre fonksiyonunu temsil eder. Denklem 3.9 da sol tarafta "düz" katmanların potansiyel bir kimlik eşleme fonksiyonunu doğrudan uydurması, derin ağların eğitimini zorlaştırır. Ancak sağ tarafta gösterildiği gibi $H(x) = F(x) + x$ dönüşümü, residual ağın daha kolay optimize edilmesini ve daha hızlı yakınsamasını sağlar. Ağın belirli bir derinliğe ulaşması, daha güçlü ifade yeteneğine sahip özelliklerin çıkarılmasını sağlar (Xu a., Li, & Duan, 2021).

Özet olarak açıklanacak olursa Resnet152, toplamda 23 milyondan fazla parametreye sahiptir. Adaptive Resnet152, 152 katmanlı olup önceden eğitilmiş bir modeldir. Bu model, artıklı öğrenme yöntemini kullanarak, genellikle bu katmanın girişinde özellik çıkarma için rasgele erişim bağlantıları kullanır. Artıklı öğrenmenin model eğitim performansını önemli ölçüde artırdığı bilinmektedir. (Raju & Rao, 2022, s. 195-197).

3.2.7. Performans Ölçütleri

Yapay zekâ, veri madenciliği, derin öğrenme ve yapay sinir ağları gibi yaklaşımlar kullanılarak çözümlenmeye çalışılan problemlerin performansı ve başarısını ölçmek üzere literatürde kullanılabilecek birden fazla ölçüt bulunmaktadır. Derin öğrenme modellerinin sınıflandırıcı performanslarının değerlendirilmesi için performans ölçütleri kullanılmaktadır. Birçok değerlendirme metriği bulunmaktadır. Bu metriklerdeki performans değerleri bazılarında güçlü, bazılarında ise düşük değerde çıkabilmektedir. Bununla beraber denilebilir ki kullanılan bir modelin başarısı bu kriterlerin haricinde ayrıca eğitim ve test sınıflarına atanan verilerin sayılarına da bağlıdır. Modeller bu metrikler sayesinde tahmin edilir ve gerçek değerleri karşılaştırılır. Bu tez çalışmasında faydalanan performans metrikleri aşağıda bahsedildiği gibidir.

Karışıklık Matrisi

Karışıklık matrisi, derin öğrenme ve makine öğrenmesinde uygulanan doğruluk değerlerinin bilindiği test veri seti üzerinde, modellerin performansının değerlendirilmesi için hedef ve tahmin değerlerinin karşılaştırılmasında sıklıkla kullanılmaktadır. Kullanılan modelin başarı performansı, karışıklık matrisi (confusion matrix) tablosu ile ifade edilir. Çizelge 3.19.'da karışıklık matrisi gösterilmektedir.

Çizelge 3.2. Karışıklık (hata) matrisi

		GERÇEK	
		Pozitif	Negatif
TAHMİN	Pozitif	TP	FP
	Negatif	FN	TN

Doğru pozitif (true positive, TP) ile doğru şekilde sınıflandırılmış pozitif kayıtların sayısı, doğru negatif (true negative, TN) ile sınıflandırıcı tarafından doğru sınıflandırılmış negatif kayıt sayısı ifade edilir. Yanlış pozitif (False Positive, FP) ile yanlış

sınıflandırılan pozitif etiket sayısı, yanlış negatif (false negative, FN) ile yanlış sınıflandırılan negatif kayıt sayısı ifade edilir Çizelge 3.2’de verilen hata matrisi, sınıflandırma modelinin, gerçek değerlerin bilindiği belirli bir test kümesinde nasıl performans gösterdiği ölçmek amacıyla kullanılan bir yöntemdir (Powers, 2020).

$$\text{TOPLAM} = TP + FP + TN + FN \quad (3.10)$$

$$\text{Gerçek Pozitif Değerler} = TP + FN \quad (3.11)$$

$$\text{Gerçek Negatif Değerler} = TN + FP \quad (3.12)$$

Doğruluk Oranı (Accuracy Rate)

Özellikle sınıflandırma başarısının ve performansının ölçülmesinde kullanılan en yaygın yöntem doğruluktur. Doğruluk; doğru şekilde ayrıştırılmış kayıtların sayısının toplam kayıtların sayısına bölünmesiyle elde edilir. Yani karışıklık matrisinde doğru pozitif veya doğru negatif olarak sınıflandırılan örnek sayılarının toplamının toplam örnek sayısına (TP+FP +FN +TN) bölünmesiyle bulunur. Doğruluk formülü denklemi 3.13’te ifade edilmiştir.

$$\text{Doğruluk} = (TP+TN)/(TP+FP+FN+TN) \quad (3.13)$$

Duyarlılık (Recall)

Duyarlılık, sınıflandırıcının hangi oranda gerçek pozitif ya da gerçek negatif değeri doğru tahminde bulunduğunun ölçütüdür. İsabet oranı veya hatırlama oranı olarak da isimlendirilmektedir. Bu değerlerin mümkün olduğunca yüksek çıkması beklenir. Yapılan farklı tahminler üzerinden kullanılan modelin tutarlılığını ölçmede duyarlılık ölçeği kullanılır. Duyarlılık doğru şekilde sınıflandırılmış pozitif kayıt sayısının toplam pozitif kayıt sayısına oranlanmasıyla hesaplanır. Duyarlılık formülü denklem 3.14’te gösterilen şekilde ifade edilir.

$$\text{Duyarlılık} = TP/(TP+FP) \quad (3.14)$$

Hassasiyet (Precision)

Hassasiyet tüm sınıflar içerisinde yapılan doğru pozitif tahminlerin, toplam pozitif tahminlere bölünmesiyle elde edilir. Hassasiyet ile doğru olarak ne kadar tahmin yapıldığı ölçülmektedir. Hassasiyet aynı zamanda pozitif tahmin edici değer olarak da isimlendirilmektedir. Doğru şekilde sınıflandırılmış olan kayıtların oranı ise hassaslık ile ifade edilir. Hassaslık denklem 3.15'teki formül yardımıyla hesaplanır (Yıldız & Karadeniz, 2019, s. 452-462).

$$Hassaslık = TP / (TP + FN) \quad (3.15)$$

Kesinlik (Recall)

Kesinlik doğru pozitif tahmin sayısının, doğru pozitif tahmin sayısı ile doğru negatif tahmin sayısına bölünmesiyle hesaplanan bir ölçüdür. Kesinlik denklemi 3.16'da gösterildiği şekilde hesaplanmaktadır.

$$Kesinlik = TP / (TP + TN) \quad (3.16)$$

F-1 Puanı (F-1-Score)

F puanı, gerçek pozitif değerlerin oranını ifade eden kesinlik (recall) ve hassasiyetin harmonik ortalamasıdır. F puanı bir sınıflandırıcının gösterdiği performansın ne kadar iyi olduğunun bir ölçüsüdür. F-1score literatürde sınıflandırıcıların başarısını karşılaştırmak için sıklıkla kullanılmaktadır. F-1 puanı denklem 3.17'de gösterildiği şekilde hesaplanmaktadır

$$f-1 \text{ puanı} = 2 * (hassasiyet * duyarlılık) / (hassasiyet + duyarlılık) \quad (3.17)$$

4. BULGULAR

Canlıların duygusal durumları, sadece sözel ifadelerle değil, aynı zamanda yüz ifadeleri aracılığıyla da belirginleşir. Bir kişinin duygu durumu, vücut genelinde o anki duygu halini yansıtarak ortaya çıkar. İnsan duygularını sadece konuşma yoluyla anlamak genellikle zorlu ve zaman alıcı bir süreçtir. Günümüzde, teknolojinin hızla gelişmesiyle birlikte, veriler giderek karmaşık hale gelmiş ve bu karmaşıklık, yapay zekâ (AI), görüntü işleme ve duygu analizi üzerinde yoğun çalışmalara neden olmuştur. İnsan duygularını tanımak için yapılan araştırmalarda, yüz tanıma analizi en kapsamlı alanlardan birine sahiptir. İnsanoğlu, özellikle duyguların otomatik olarak tespit edilmesi konusunda, sağlık sektöründen akıllı yaşam alanlarına, otizm ve spektrum bozukluğundan şizofreniye kadar geniş bir yelpazede araştırmalar yapmaktadır. Bu alanda, özellikle yüz duygu tanıma çeşidi, yüz ifadeleri arasındaki duygu analizini gerçekleştirmekle sorumludur. Yüzdeki duygu analizi, yüzün algılanması, değişiklikler, kırpmalar ve boyut değişimleri gibi faktörlere dayanarak gerçekleşir. Bu tez çalışmasında yüz görüntü ifadelerinden duygu analizi gerçekleştirilmiştir. Bunun için Residual Network (ResNet), transfer derin öğrenme metotları kullanılmıştır. Transfer derin öğrenme, bir görevde öğrenilen bilgilerin başka bir göreve transfer edilmesini amaçlayan bir öğrenme yaklaşımıdır. Model, bir görevdeki deneyimini başka bir görevde kullanarak genelleme yeteneğini artırmaya çalışır. ResNet derin sinir ağlarının eğitimini kolaylaştırmak ve aşırı uyum sorunlarını azaltmak amacıyla geliştirilmiş bir derin öğrenme mimarisidir. ResNet modelleri, kendi içinde bloklar halinde düzenlenmiş ve her bir blok, giriş ve çıkış arasında bir "artıklık bağlantısı" içerir. Bu bağlantılar, modelin daha derin ve etkili öğrenme yapabilmesini sağlar. ResNet modellerinin avantajları arasında eğitim kolaylığı, daha derin ağların başarıyla eğitilebilmesi ve aşırı uyum sorunlarının azaltılması bulunur. Mimarileri genellikle 18, 34, 50, 101 ve 152 katmanlı olmak üzere çeşitli derinliklerde gelir. ResNet modelleri, ImageNet gibi büyük veri setlerinde olağanüstü performans sergileyerek derin öğrenme alanında önemli bir etki yaratmıştır. Bu çalışma ResNet 50,101,152 ve bu modellere ait V2'ler kullanılmıştır. Modellerin kullanım mimarisi Çizelge 4.1'de verilmiştir. Öncelikle ResNet 50,101,152 ve bunlara ait V2 modeller alınmıştır. Model çıktısı bir global havuzlama katmanına verilmiştir. Global havuzlama katmanından sonra modellerin overfitting(Ezberleme) problemlerini engellemek için DropOut katmanı kullanılmıştır. Daha sonra 128 nöronlu oluşan bir Dense katmanı kullanılmıştır. Son

katmanda girilen bir görüntüyü ilgili sınıfa aitliğini olasılıksal olarak hesaplayan softmax aktivasyon fonksiyonlu, 7 çıkışlı bir Dense katmanı kullanılmıştır.

Çizelge 4.1. ResNet model mimarileri

```
x = base_model.output (ResNet50, ResNet50V2, ...)
x = GlobalAveragePooling2D()(x)
x = Dropout(0.3)(x)
x = Dense(128, activation='relu')(x)
tahminler = Dense(7, activation='softmax')(x)
```

ResNet modelleri aynı hiperparametre değerleri ile kullanılmıştır. Bu değerlere denemeler sonucunda karar verilmiştir. Denemeler sonucunda en iyi optimizasyon yöntemi olarak “adam” olarak bulunmuştur. Optimizasyon, bir makine öğrenimi modelinin parametrelerini (ağırlıklarını) güncellemek ve eğitim verisine daha iyi uyan bir model elde etmek için kullanılan bir optimizasyon algoritmasıdır. Bu algoritmalar, kayıp fonksiyonu tarafından belirlenen hatayı minimize etmeye çalışır. Gradient Descent, en yaygın kullanılan optimizasyon algoritmalarından biridir. Temelde, modelin parametrelerini belirli bir adım boyutu (learning rate) kullanarak kayıp fonksiyonunun negatif gradyanına doğru günceller. Stokastik Gradient Descent (SGD), her iterasyonda yalnızca bir örnek kullanarak güncellemeler yapar, mini-batch SGD belirli bir örnek grubunu kullanır ve Batch Gradient Descent tüm eğitim verisini kullanır. Farklı optimizasyon algoritmaları, momentum, RMSprop, ve Adam gibi teknikleri içerebilir, bu da daha hızlı ve stabil bir optimizasyon süreci sağlar. Optimize edici seçimi, eğitim performansını büyük ölçüde etkileyebilir. batch_size=64 alınmıştır. Batch size, bir makine öğrenimi modelinin eğitim sırasında kullanılan örnek veri gruplarının sayısını belirtir. Batch size'ın değiştirilmesi, eğitim sürecinde hız, bellek kullanımı, stabilite, genelleme yeteneği, GPU kullanımı ve optimizasyon performansı gibi faktörlere etki eder. Büyük batch size'lar genellikle hızlı eğitim ve stabil gradyan tahminleri sağlar, ancak genelleme yeteneğini azaltabilir. Küçük batch size'lar daha az bellek kullanır, stokastik gradyan tahminleri ile genelleme yeteneğini artırabilir, ancak eğitim süresini uzatabilir. Optimal batch size, görev ve veri kümesine bağlı olarak değişir, denemelerle belirlenmesi önerilir. Kayıp(loss) fonksiyonu olarak “CategoricalCrossentropy” kullanılmıştır. Loss fonksiyonu, bir makine öğrenimi modelinin tahminleri ile gerçek değerleri arasındaki farkı ölçer ve bu farkı minimize ederek modelin doğruluğunu artırmayı amaçlayan bir matematiksel fonksiyondur. Regresyon problemleri için kullanılan Mean Squared Error (MSE), sınıflandırma problemleri için kullanılan Cross-

Entropy Loss gibi farklı türleri bulunur. Eğitim süreci boyunca bu loss fonksiyonu, modelin parametrelerini güncellemek için optimizasyon algoritmaları tarafından kullanılır. Amaç, modelin gerçek verilere daha iyi uymasını sağlamaktır. Öğrenme oranı “0.001” olarak alınmıştır. Modellerin performanslarını değerlendirmek için 'accuracy', 'loss', 'precision', 'recall', 'AUC', 'f1_score' metrikleri kullanılmıştır. Modellere ait hiperparametreler Çizelge 4.2’de gösterilmiştir.

Çizelge 4.2. ResNet modeller hiperparametre değerleri

Hiperparametre	Değer
batch_size	64
epoch	100
Metrikler	'accuracy', 'loss', 'precision', 'recall', 'AUC', 'f1_score'
optimizer	Adam
learning_rate	0.001
losses fonksiyonu	CategoricalCrossentropy

ResNet modelleri %70-30 eğitme-test oranları ile kullanılmıştır. Modelleri test etmek için Fer2013 isimli veri seti kullanılmıştır. Bu veri seti Kaggle veri tabanından indirilmiştir. Veri seti 48x48 pikselleri renkli görüntülerden oluşmaktadır. Veri seti yüz ifadelerin duygu tanıma için kullanılan zor bir kıyaslama setidir. Eğitim seti için elde edilen performans değerleri Çizelge 4.3’te test seti için elde edilen değerler ise Çizelge 4.4’te verilmiştir.

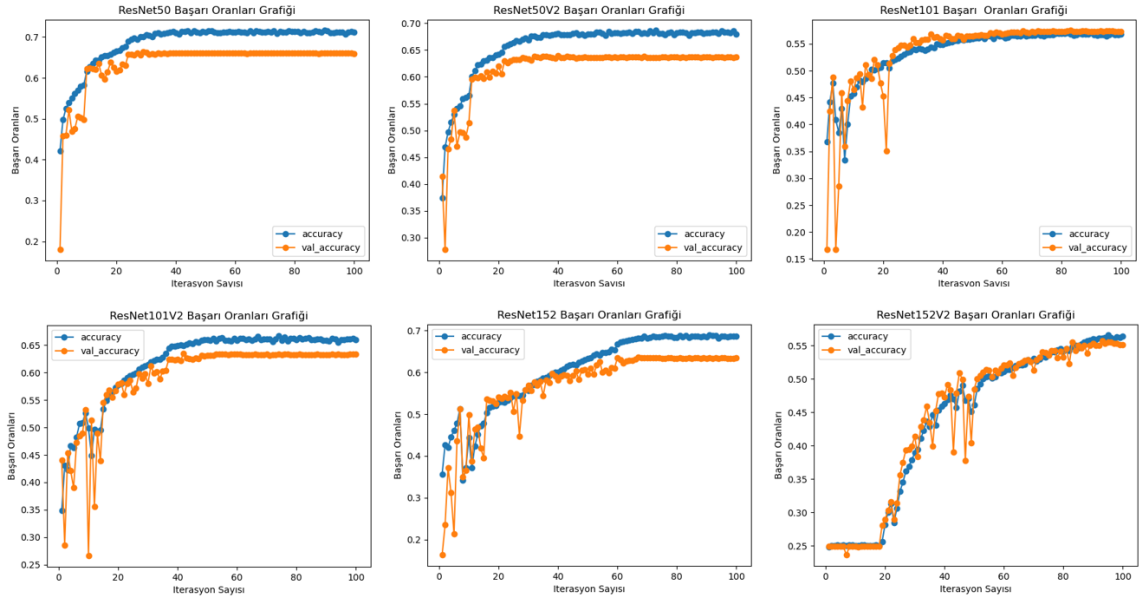
Çizelge 4.3. ResNet Modelleri ile Eğitim veri seti için performans değerleri

Model	Başarı	Kesinlik	Duyarlılık	ROC	F1-Ölçütü
ResNet50	0.715978	0.801933	0.628653	0.951398	0.703103
ResNet50V2	0.68644	0.789642	0.575151	0.940771	0.664381
ResNet101	0.571006	0.757551	0.385593	0.892909	0.509106
ResNet101V2	0.666655	0.779763	0.543279	0.932732	0.639391
ResNet152	0.689122	0.790212	0.582326	0.941685	0.6688
ResNet152V2	0.565432	0.786559	0.372078	0.889189	0.498799

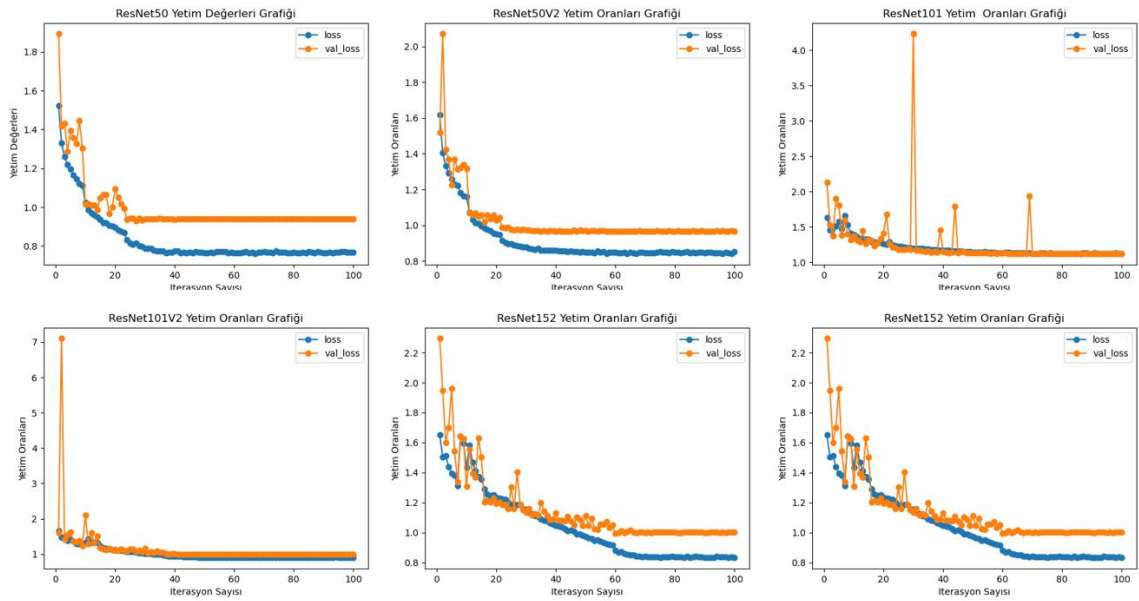
Çizelge 4.4. ResNet Modelleri ile Test veri seti için performans değerleri

Model	Başarı	Kesinlik	Duyarlılık	ROC	F1-Ölçütü
ResNet50	0.664113	0.750251	0.588047	0.928684	0.649823
ResNet50V2	0.639175	0.770026	0.544302	0.922707	0.624436
ResNet101	0.574812	1	0.408192	0.892905	0.519496
ResNet101V2	0.635135	0.76031	0.528281	0.918669	0.61516
ResNet152	0.636389	0.815857	0.546252	0.918597	0.621442
ResNet152V2	0.556701	0.925	0.376567	0.882451	0.496903

Çizelge 4.3 ve 4.4’e bakıldığında eğitim seti için en yüksek başarı oranı “**0.715978**” olarak ResNet50 modeli ile gözlenmiştir. Yine aynı model ile test veri seti için en yüksek başarı “0.664113” olarak elde edilmiştir. ResNet modellerin başarı grafiklerine bakıldığında modellerin eğitim sürecinin stabil olduğu ancak test veri seti için ResNet50 ve ResNet50V2 modellerin daha stabil bir öğrenme süreci gerçekleştirdiği görülmektedir. Çizelgelere bakıldığında en başarısız modelin ResNet101 olduğu görülmektedir. Modellere ait başarı grafikleri iterasyon değerlerine göre Şekil 4.1’de verilmiştir. Bunun yanında modellerin kayıp(hata) değerlerine ait grafikler Şekil 4.2’de gösterilmiştir.

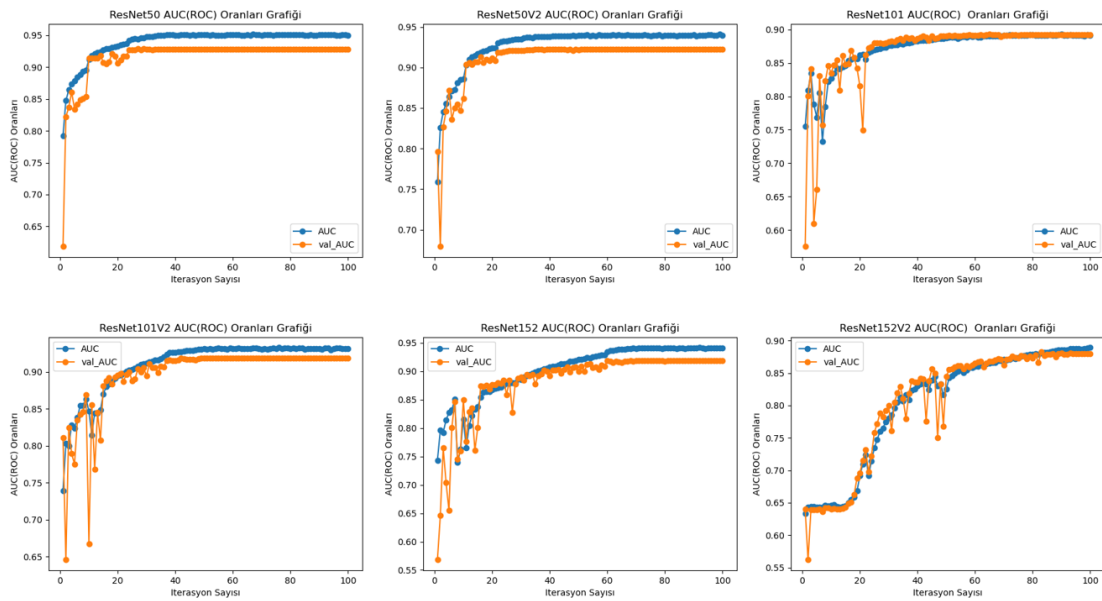


Şekil 4.1. ResNet modellere ait başarı oranları



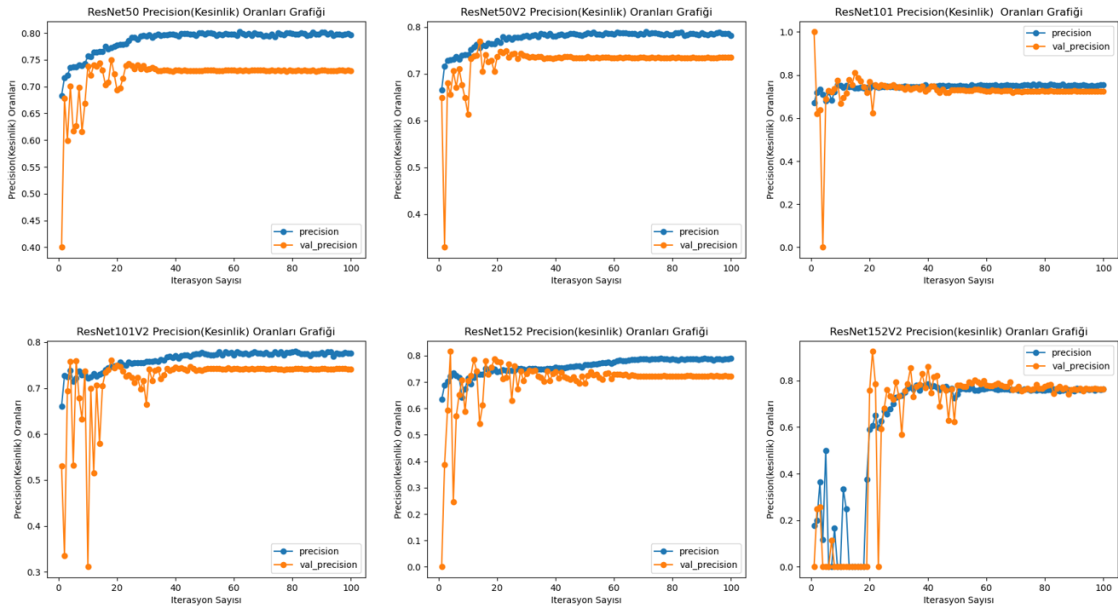
Şekil 4.2. ResNet modellere ait yetim değerleri

Receiver Operating Characteristic (ROC) eğrisi, sınıflandırma modellerinin performansını ayrıntılı bir şekilde değerlendirmek amacıyla kullanılan bir görselleştirme aracıdır. Bu eğri, bir modelin gerçek pozitif oranı (hassasiyet) ile gerçek negatif oranı (özgüllük) arasındaki dengeyi inceleyerek, modelin farklı sınıflandırma eşiklerinde nasıl performans gösterdiğini gösterir. ROC eğrisi, genellikle bir modelin sınıflandırma doğruluğunu değerlendirmenin yanı sıra, özellikle dengesiz sınıf dağılımlarına sahip veri setlerinde modelin davranışını daha iyi anlamak için kullanılır. ROC eğrisi, bir modelin her bir eşik değeri altında elde ettiği hassasiyet ve özgüllük değerlerini gösteren bir eğri çizer. Eğri, sol üst köşeye ne kadar yaklaşırsa, modelin performansının o kadar iyi olduğunu gösterir. Bu durumda, hem gerçek pozitifler (doğru pozitif tahminler) hem de gerçek negatifler (doğru negatif tahminler) yüksektir. Eğri sol alt köşeye yaklaştıkça, modelin performansı azalır, çünkü özgüllük artarken hassasiyet düşer. AUC (Area Under the ROC Curve) değeri, ROC eğrisinin altındaki alanı ifade eder ve genellikle tek bir sayı ile modelin sınıflandırma performansını ölçmek için kullanılır. AUC değeri 1'e ne kadar yakınsa, modelin performansı o kadar iyidir. ROC analizi, özellikle tıbbi teşhis, makineler arası öğrenme ve dolandırıcılık tespiti gibi alanlarda kullanılarak, bir modelin doğruluk, hassasiyet ve özgüllük gibi performans metriklerini kapsamlı bir şekilde değerlendirme imkânı sağlar. Modellere ait ROC grafikleri Şekil 4.3'te verilmiştir. Grafiklere bakıldığında hem eğitim hem de test veri setleri için model en iyi modelin ResNet50 olduğu görülmektedir. En kötü model ise ResNet101 olarak görülmektedir.



Şekil 4.3. ResNet Modellerine ait ROC eğrileri

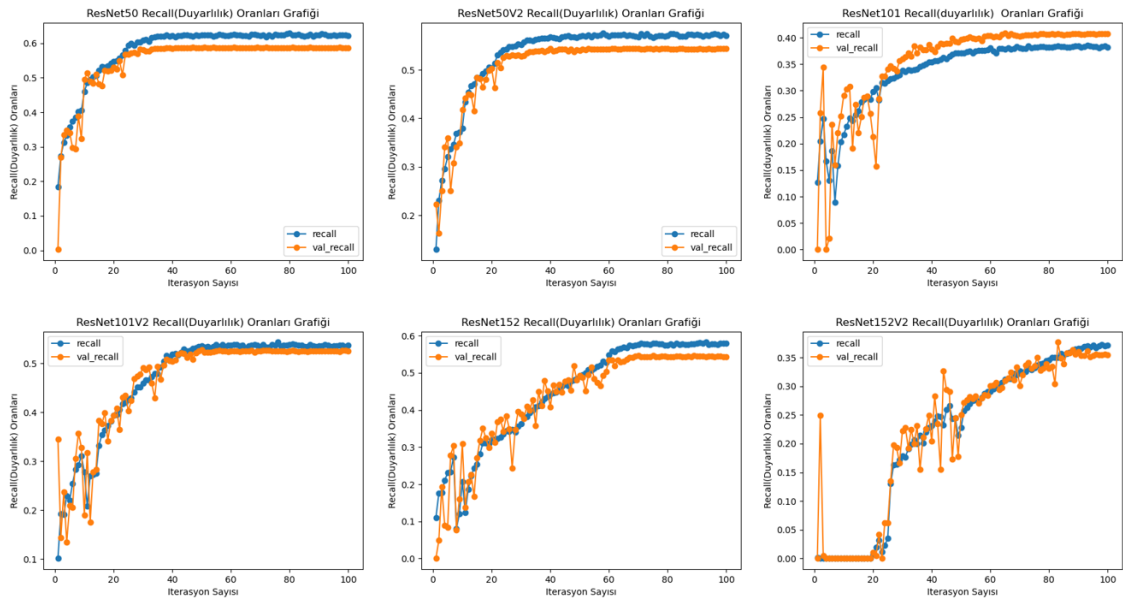
Precision (Kesinlik), sınıflandırma modellerinin performansını değerlendiren kritik bir metrik olarak tanımlanır ve genellikle pozitif olarak tahmin edilen örneklerin ne kadarının gerçekten pozitif olduğunu ölçer. Bu metrik, modelin yanlış pozitif tahminleri (False Positives) minimize etme amacını taşır ve bu nedenle özellikle yanlış alarmın maliyetli olduğu veya istenmeyen durumların olduğu senaryolarda önemlidir. Precision, True Positives'in, toplam pozitif tahminlerin içindeki oranını temsil eder ve yüksek bir precision değeri, modelin pozitif olarak tahmin ettiği örneklerin gerçekte çoğunlukla pozitif olduğunu gösterir. Bu durum, modelin yanlış pozitif oranının düşük olduğu anlamına gelir, dolayısıyla modelin güvenilir bir şekilde pozitif tahmin yapma yeteneğini vurgular. Ancak, precision değeri tek başına yeterli bir performans ölçütü olmayabilir; bu nedenle genellikle recall, F1-score ve ROC eğrisi gibi diğer metriklerle birlikte değerlendirilir. Yüksek precision, modelin sınıfları doğru bir şekilde tanıma eğiliminde olduğunu ancak bu değer diğer metriklerle dengelenerek yorumlanması gerektiğini ifade eder. Modellere ait kesinlik grafikleri Şekil 4.4'te verilmiştir. Grafiklere bakıldığında hem eğitim hem de test veri setleri için model en iyi modelin ResNet50 olduğu görülmektedir. En kötü model ise ResNet101 olarak görülmektedir.



Şekil 4.4. ResNet modellerine ait Precision(Kesinlik) Değerleri

Recall (duyarlılık), sınıflandırma modellerinin performansını değerlendiren önemli bir metrik olarak kabul edilir ve genellikle gerçek pozitif örneklerin, toplam pozitif örnekler içindeki oranını ölçer. Recall, aynı zamanda "duyarlılık" olarak da adlandırılır ve

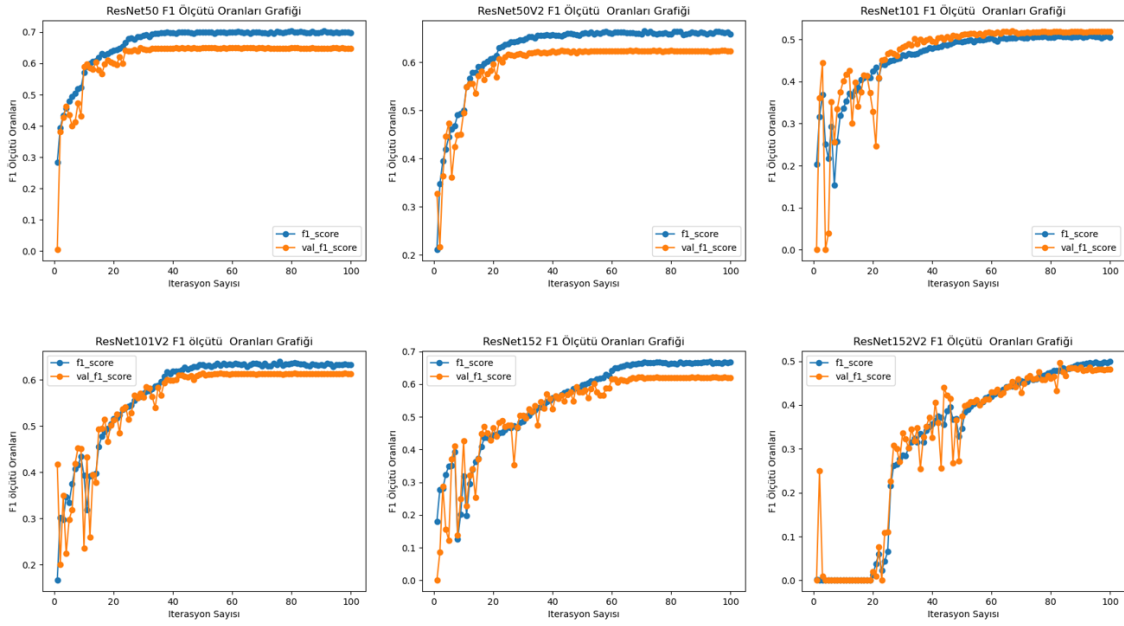
aşağıdaki formülle hesaplanır: $\text{Recall} = \text{True Positives} / (\text{True Positives} + \text{False Negatives})$. Bu formülde, "True Positives" (TP) doğru pozitif tahminleri, "False Negatives" (FN) yanlış negatif tahminleri temsil eder. Yüksek bir recall değeri, modelin gerçek pozitifleri doğru bir şekilde tespit etme yeteneğini gösterir ve bu, özellikle modelin pozitif vakaları kaçırmamak, yani düşük yanlış negatif oranına sahip olması durumunda önemlidir. Ancak, recall'ın artması genellikle precision'ın düşmesine neden olur, bu nedenle bu iki metrik birbirine karşı bir ticaret-off ilişkisi içindedir. Recall, genellikle tıbbi teşhis ve acil durumlar gibi hassaslık gerektiren uygulamalarda önemlidir ve diğer performans ölçütleri ile birlikte değerlendirilerek bir modelin genel etkinliğini anlamak için kullanılır. Modellere ait duyarlılık grafikleri Şekil 4.5'te verilmiştir. Grafiklere bakıldığında hem eğitim hem de test veri setleri için model en iyi modelin ResNet50 olduğu görülmektedir. En kötü model ise ResNet101 olarak görülmektedir.



Şekil 4.5. ResNet modellerine ait recall(duyarlılık) değerleri

F1-score, bir sınıflandırma modelinin performansını değerlendirmek için kullanılan bir metrik olup, precision (kesinlik) ve recall (duyarlılık) metriklerinin harmonik ortalamasını temsil eder. Bu metrik, modelin hem yanlış pozitifleri minimize etme yeteneğini hem de yanlış negatifleri en aza indirmeye yeteneğini dengeler. F1-score, özellikle dengesiz sınıf dağılımlarına sahip veri setlerinde, modelin hem hassas hem de odaklı çalışmasını sağlamak adına önemlidir. Yüksek bir F1-score değeri, modelin hem precision hem de recall'ın etkin bir şekilde optimize edildiği durumları gösterir, bu da modelin genel performansının dengeli olduğunu ifade eder. F1-score, özellikle tıbbi

teşhis, bilgi güvenliği ve dolandırıcılık tespiti gibi alanlarda kullanılarak, modelin pozitif ve negatif sınıfları başarıyla tanıma yeteneğini değerlendirmek için tercih edilir. Modellere ait duyarlılık grafikleri Şekil 4.6’da verilmiştir. Grafiklere bakıldığında hem eğitim hem de test veri setleri için model en iyi modelin ResNet50 olduğu görülmektedir. En kötü model ise ResNet101 olarak görülmektedir.



Şekil 4.6. ResNet modellerine ait F1 skor değerleri

5. TARTIŞMA

Günümüzde iletişim, insanlar arasında temel bir ihtiyaç olarak varlığını sürdürmekte ve teknolojik gelişmelerle birlikte bu iletişim dinamikleri önemli ölçüde evrim geçirmektedir. Yapay zekâ alanındaki ilerlemeler, makinelerin insanlarla daha etkili bir şekilde iletişim kurabilme potansiyelini artırmıştır. Bu bağlamda, duygusal zekâ ve duygu tanıma teknolojileri, insan-makine etkileşimini daha zengin ve derinleştirebilecek önemli bir rol oynamaktadır. İnsanların duygusal ifadelerle iletişim kurma ihtiyacı, insanlar arasındaki etkileşimin temelini oluşturur. Ancak, günümüzde yapay zekâ algoritmalarının duygu tanıma yetenekleri gelişmiş durumdadır. Bu algoritmalar, yüz ifadeleri, ses tonu, metin analizi gibi çeşitli modalitelerde duyguları anlamak ve yorumlamak için eğitilebilmektedir. Bu sayede makineler, insanlarla benzer duygusal bağlar kurma kapasitesine sahip hale gelmiştir (Rázuri, vd. 2013). Duygusal zekâ, makinelerin sadece görevleri yerine getirme yeteneğini aşarak, insanlarla duygusal bir bağ kurabilme yeteneği olarak tanımlanabilir. Bu noktada, derin öğrenme modelleri, özellikle de transfer öğrenme metotları, duygu tanıma algoritmalarının eğitiminde önemli bir rol oynamaktadır. ResNet gibi transfer öğrenme modelleri, geniş veri setlerinden öğrenilen bilgileri benzer görevlerde kullanarak duygu tanıma performansını artırabilir (Akhand, vd. 2021). Bu tez çalışması, duygu tanıma konusundaki başarı değerlerini ele alarak, özellikle altı temel duygu üzerinde odaklanmıştır. Bu duygular arasında mutluluk, üzüntü, korku, öfke, iğrenme ve şaşkınlık bulunmaktadır. Elde edilen başarı değerleri, kullanılan transfer öğrenme yöntemleri ile sağlanmış ve bu yöntemlerin duygu tanıma performansındaki etkinliği değerlendirilmiştir. Yapılan geniş çaplı literatür taraması, duygu tanıma teknolojisinin sağlık sektöründen güvenlik uygulamalarına, eğitimden insan-makine etkileşimine kadar pek çok alanda kullanılabileceğini ortaya koymuştur. Özellikle sağlık sektöründe, hastaların duygusal durumlarının izlenmesi ve anksiyete, depresyon gibi durumların erken tespiti konusunda duygu tanıma sistemleri önemli bir rol oynayabilir. Eğitim ve test aşamalarında kullanılan aktarım derin öğrenme metotları, ResNet gibi modeller üzerinden değerlendirilmiş ve elde edilen sonuçlar çizelgeler ve karışıklık matrisi aracılığıyla sunulmuştur. Bu sonuçlar, transfer öğrenme metotlarının duygu tanıma performansını artırdığını göstermiş, özellikle geniş veri setlerinden öğrenilen bilgilerin benzer görevlerde kullanılmasının etkili bir strateji olduğunu ortaya koymuştur. Sonuç olarak, bu çalışma, yapay zekânın duygu tanıma yeteneklerini ele alarak, gelecekteki potansiyelini vurgulamıştır. Yapay zekânın duygusal zekâ

konusundaki ilerlemeleri, insan-makine etkileşimini daha zengin ve anlamlı hale getirebilir. Ancak, bu teknolojilerin etik ve gizlilik konuları da göz önünde bulundurularak dikkatlice yönetilmesi gerekmektedir.

Duygusal durumların anlaşılması, sadece sözel ifadelerle değil, aynı zamanda yüz ifadeleri aracılığıyla da belirginleşir. Bu gerçekten hareketle, duyguları tanıma ve anlama konusundaki araştırmalar, yapay zekâ (AI), görüntü işleme ve duygu analizi alanında büyük bir ilgi odağı haline gelmiştir. İnsanların duygusal ifadelerini anlamak, genellikle karmaşık ve zaman alıcı bir süreçtir, ancak teknolojik gelişmeler sayesinde bu süreç daha etkili ve hızlı hale gelmiştir (Akçelik, vd. 2021).

Gelişen teknoloji, özellikle yapay zekâ ve derin öğrenme alanındaki yenilikler, duygusal zekâ konusunda çeşitli uygulamaların ortaya çıkmasını sağlamıştır. Bu çalışmada, duygu tanıma amacıyla Residual Network (ResNet) ve transfer derin öğrenme modelleri kullanılmıştır. ResNet, derin sinir ağlarının eğitimini kolaylaştırmak ve aşırı uyum sorunlarını azaltmak amacıyla geliştirilmiş bir derin öğrenme mimarisidir (Aksoy & Salman, 2022). Bu çalışmanın temel amacı, ResNet modellerinin duygu tanıma performansını değerlendirmektir. Modellerin eğitim ve test aşamalarında kullanılan hiperparametreler, dikkatlice seçilmiş ve denemeler sonucunda belirlenmiştir. Optimizasyon algoritması olarak "Adam" kullanılmış, öğrenme oranı "0.001" olarak belirlenmiş ve Categorical Crossentropy loss fonksiyonu tercih edilmiştir (Çizelge 4.1). Eğitim ve test aşamalarında kullanılan veri seti, duygu tanıma için zorlu bir kıyaslama seti olan Fer2013'ten alınmıştır. Eğitim seti için elde edilen performans değerleri, özellikle başarı, kesinlik, duyarlılık, ROC (Receiver Operating Characteristic) eğrisi ve F1-Ölçütü olmak üzere çeşitli metriklerle değerlendirilmiştir (Çizelge 4.2). Eğitim seti için en yüksek başarı oranı "0.715978" ile ResNet50 modelinde gözlemlenmiştir. Benzer şekilde, test seti için en yüksek başarı "0.664113" ile yine ResNet50 modelinde elde edilmiştir (Çizelge 4.3 ve Çizelge 4.4). Bu sonuçlar, ResNet50'nin duygu tanıma görevinde diğer modellere kıyasla daha başarılı olduğunu göstermektedir. Performans metriklerinin yanı sıra ROC eğrisi analizi, modellerin sınıflandırma performansını detaylı bir şekilde değerlendirmek için kullanılmıştır. ROC eğrilerine bakıldığında, hem eğitim hem de test veri setleri için en iyi performansın ResNet50 modelinde olduğu görülmüştür. En düşük performans ise ResNet101 modelinde gözlemlenmiştir. Precision, recall ve F1-score gibi metrikler de incelendiğinde, yine ResNet50'nin en yüksek değerlere sahip

olduğu görülmüştür. Bu metrikler, modelin yanlış pozitif ve yanlış negatif oranlarını dengeleyerek genel performansını değerlendirmekte önemli bir rol oynamaktadır.

Sonuç olarak, bu çalışma, duygu tanıma alanında ResNet modellerinin performansını değerlendirmiş ve özellikle ResNet50'nin diğer modellere kıyasla daha etkili olduğunu ortaya koymuştur. Bu bulgular, duygu tanıma teknolojilerinin gelecekteki gelişmelerine ışık tutabilir ve bu alandaki araştırmalara yeni yönler sağlayabilir. Ancak, model seçiminde ve performans değerlendirmesinde dikkatli bir metodoloji takip edilmesi önemlidir.

5.1. Hedefler ve Kısıtlamalar

Bu çalışmanın temel hedefi, yapay zekâ ve derin öğrenme teknikleri kullanarak duygu tanıma yeteneklerini değerlendirmektir. Duygusal ifadelerin anlaşılması, insan-makine etkileşiminde önemli bir rol oynamakta ve bu alandaki teknolojik gelişmeler, duygu tanıma uygulamalarının etkisini artırmaktadır. Çalışmada, özellikle Residual Network (ResNet) ve transfer derin öğrenme modelleri kullanılarak duygu analizi gerçekleştirilmiştir. Bu modeller, önceden eğitilmiş bilgilerin başka bir göreve aktarılmasını amaçlayan transfer derin öğrenme yaklaşımını benimsemektedir. ResNet mimarisi, derin sinir ağlarının eğitimini kolaylaştırmak, aşırı uyum sorunlarını azaltmak ve daha etkili öğrenme sağlamak amacıyla geliştirilmiştir. Çalışma, bu modellerin duygu tanıma performansını değerlendirerek, özellikle ResNet50 modelinin bu görevdeki etkinliğini vurgulamayı amaçlamaktadır. Eğitim ve test aşamalarında kullanılan veri seti, duygu tanıma için zorlu bir kıyaslama seti olan Fer2013'ten alınmıştır. Modellerin performansı, başarı, kesinlik, duyarlılık, ROC eğrisi ve F1-Ölçütü gibi çeşitli metriklerle değerlendirilmiştir. Elde edilen sonuçlar, özellikle ResNet50 modelinin diğer modellere kıyasla daha yüksek bir başarı oranına sahip olduğunu göstermektedir. Bu çalışmanın sonuçları, duygu tanıma teknolojilerinin gelecekteki potansiyelini ve derin öğrenme modellerinin bu alandaki etkisini anlamamıza katkı sağlayabilir. Ayrıca, yapay zekâ tabanlı duygu tanıma uygulamalarının geliştirilmesi ve iyileştirilmesi konusunda araştırmacılara rehberlik edebilir.

Bu çalışmanın kısıtlı taraflarına bakıldığında veri setinde her duygu için görüntü dağılımları çok dengesiz. Bu da derin öğrenme modellerinin veri setine uyumunu zorlaştırmaktadır. Bir diğer kısıtlayıcı durum veri setindeki görüntülerin 48x48 pikselden

ve görüntülerin renksiz olmasıdır. Bu durum da görüntülerin duyguları ifade etmek için yeterli bilgi barındırmadığını göstermektedir. Sonraki çalışmalarda daha dengeli veriler üzerinde denemeler gerçekleştirilecektir.

6. KAYNAKLAR

- Akçelik, C., Okatan, A. ve Çetinkaya, A. (2021). Bilgisayarla Görme Yöntemlerine Dayalı Derin Öğrenme Kullanılarak Yüz ve Duygu Tanıma. *algoritmalar*, 11, 12.
- Akhand, MAH, Roy, S., Siddique, N., Kamal, MAS ve Shimamura, T. (2021). Derin CNN'de transfer öğrenmeyi kullanarak yüz duygu tanıma. *Elektronik*, 10 (9), 1036.
- AKSOY, B., & SALMAN, O. K. M. (2022). Prediction of Covid-19 disease with Resnet-101 deep learning architecture using Computerized Tomography images. *Türk Doğa ve Fen Dergisi*, 11(2), 36-42.
- Ali, A. R. I. ve Hanbay, D. (2018). Bölgesel evrimsel sinir ağları tabanlı mr görüntülerinde tümör tespiti. Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 34(3), 1395-1408.
- Ananthram, A., Saravanakumar, K. K., Huynh, J., & Beigi, H. (2020). Multi-modal emotion detection with transfer learning. *arXivpreprint arXiv:2011.07065*.
- ARSLAN, N. N. (2022). Analysis of Deep Transfer Learning Methods for Early Diagnosis of the Covid-19 Disease with Chest X-ray Images. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 10(2), 628-640.
- Atilgan, Ç., 2018 Yüz ifadelerine göre otomatik duygu tespiti, Yüksek lisans Tezi, *Fırat Üniversitesi Fen Bilimleri Enstitüsü*, Elazığ 1-13.
- Brooks, JA, Chikazoe, J., Sadato, N. ve Freeman, JB (2019). Yüz-duygu kategorilerinin nöral temsili kavramsal yapıyı yansıtır. *Ulusal Bilimler Akademisi Tutanakları*, 116 (32), 15861-15870.
- Cengil, E., Çinar, A., & Güler, Z. (2017, September). A GPU-based convolutional neural network approach for image classification. In 2017 International Artificial Intelligence and Data Processing Symposium (IDAP) (pp. 1-6). IEEE.
- Darwin, C. (1872). İnsanın iniş ve cinsiyete göre seçim (Cilt 2). D. Appleton, *London*, 7-8.
- Devitt, A., & Ahmad, K. (2013). Is there a language of sentiment? An analysis of lexical resources for sentiment analysis. *Language resources and evaluation*, 47, 475-511.
- Dias, W., Andaló, F., Padilha, R., Bertocco, G., Almeida, W., Costa, P., & Rocha, A. (2022). Cross dataset emotion recognition from facial expressions through convolutional neural networks. *Journal of Visual Communication and Image Representation*, 82, 103395., 2-16.
- Doğan, F., ve Türkoğlu, İ. (2018). Derin öğrenme algoritmalarının yaprak sınıflandırma başarımlarının karşılaştırılması. *Sakarya University Journal of Computer and Information Sciences*, 1(1), 10-21.

- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2), 124.
- Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Martinez Gonzalez, P., & Garcia-Rodriguez, J. (2018). A survey on deep learning techniques for image and video semantic segmentation. *Applied Soft Computing*, 70, 41-65.
- Gao, Y., & Mosalam, K. M. (2018). Deep transfer learning for image- based structural damage recognition. *Computer- Aided Civil and Infrastructure Engineering*, 33(9), 748-768.
- Gao, X., Li, W., Loomes, M., & Wang, L. (2017). A fused deep learning architecture for viewpoint classification of echocardiography. *Information Fusion*, 36, 103-113.
- Glorot, X., Bordes, A., & Bengio, Y. (2011, June). Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics* (pp. 315-323). JMLR Workshop and Conference Proceedings.
- Goodfellow, I., Bengio, Y. ve Courville, A. (2016). *Derin Öğrenme*. MIT Basın.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- Ian, G., Yoshua, B., & Aaron, C. (2018). *Derin Öğrenme*. Ita, K. (2022). Chapter 4 - Anatomy and physiology of the eye. In K. Ita (Ed.), *Microneedles* (pp. 73-92): Academic Press.
- İzard, CE (2007). Temel duygular, doğal türler, duygu şemaları ve yeni bir paradigma. *Psikoloji bilimine bakış açıları*, 2 (3), 260-280.
- Jauro, F., Chiroma, H., Gital, A. Y., Almutairi, M., Shafi'i, M. A., & Abawajy, J. H. (2020). Deep learning architectures in emerging cloud computing architectures: Recent development, challenges and next research trend. *Applied Soft Computing*, 96, 106582.,106582.
- Koçer, B., 2012 *Transfer Öğrenmede Yeni Yaklaşımlar*, Doktora Tezi, Selçuk Üniversitesi *Fen Bilimleri Enstitüsü*, Konya.
- Komodakis, N. ve Zagoruyko, S. (2017, Haziran). Dikkate daha fazla dikkat etmek: Dikkat aktarımı yoluyla evrişimli sinir ağlarının performansının artırılması. ICLR'de.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90.
- Kuchera, MP, Ramanujan, R., Taylor, JZ, Strauss, RR, Bazin, D., Bradt, J. ve Chen, R. (2019). AT-TPC'de parça sınıflandırması için makine öğrenimi yöntemleri. *Fizik*

Araştırmalarında Nükleer Cihazlar ve Yöntemler Bölüm A: Hızlandırıcılar, Spektrometreler, Dedektörler ve İlgili Ekipmanlar, 940, 156-167.

Maraş, A. *EEG alt bandlarının tekil spektrumu ile duygu durumları arasındaki ilişki* (Master's thesis, Fen Bilimleri Enstitüsü).

Mehrabian, A. (1968). Relationship of attitude to seated posture, orientation, and distance. *Journal of personality and social psychology*, 10(1), 26.

Newmark, C. (2022). Charles Darwin: the expression of the emotions in man and animals. In *Schlüsselwerke der Emotionssoziologie* (pp. 111-115). Springer VS, Wiesbaden.

Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345-1359.

Pantic, M. Machine analysis of facial behaviour: naturalistic and dynamic. Marian, DE ve Shimamura, AP (2013). Dinamik yüz ifadeleri üzerindeki bağlamsal etkiler. *Amerikan Psikoloji Dergisi*, 126 (1), 53-66.

Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv preprint arXiv:2010.16061.

Raju, M. S. N., & Rao, B. S. (2022). Colorectal multi-class image classification using deep learning models. *Bulletin of Electrical Engineering and Informatics*, 11(1), 195-200.

Ray, A. ve Chakrabarti, A. (2016). Teknolojinin tasarımı ve uygulanması, biyo-fiziksel ve yüz ifadesinin füzyonunu kullanarak etkili öğrenmeyi mümkün kıldı. *Journal of Education Technology&Society*, 19 (4), 112-125.

Rázuri, JG, Sundgren, D., Rahmani, R., &Cardenas, AM (2013, Kasım). Yapay sinir ağına dayalı birleştirilmiş görüntülerde yüz ifadesi analizi yoluyla otomatik duygu tanıma. 2013 yılında 12. Meksika Uluslararası Yapay Zekâ Konferansı (s. 85-96). IEEE.

S. Lawrence, C. L. Giles and Ah Chung Tsoi, "Convolutional neural networks for face recognition," *Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, USA, 1996, pp. 217-222

Sharma, S., Sharma, S. ve Athaiya, A. (2017). Sinir ağlarında aktivasyon fonksiyonları. *Veri Bilimine Doğru*, 6 (12), 310-316.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1), 1929-1958.

Szeliski R., (2022). *Computer Vision: Algorithms and Applications* 2nd Edition Springer ISBN-13: 978-3030343712.

- Turan, M., & Emre, A. R. I. Ğ. (2021). Video Duygu Analizi. *Avrupa Bilim ve Teknoloji Dergisi*, (Ejosat Ek Özel Sayı (HORA)), 34-41.
- Tursyngaliyeva, A. (2019). Deep Learning for medical image segmentation: Pneumonia Detection by Assem Tursyngaliyeva. August).
- Türkoğlu, M., Hanbay, K., Sivrikaya, I. S., ve Hanbay, D. (2021). Derin evrişimsel sinir ağı kullanılarak kayısı hastalıklarının sınıflandırılması. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 9(1), 334-345.
- Wang, X., Chen, X., &Cao, C. (2020). Human emotion recognition by optimally fusing facial expression and speech feature. *Signal Processing: Image Communication*, 84, 115831.
- Wen, L., Li, X., & Gao, L. (2020). A transfer convolutional neural network for fault diagnosis based on ResNet-50. *Neural Computing and Applications*, 32, 6111-6124.
- Wu, D., Ying, Y., Zhou, M., Pan, J., & Cui, D. (2023). Improved ResNet-50 deep learning algorithm for identifying chicken gender. *Computers and Electronics in Agriculture*, 205, 107622., 1-4.
- Xu, C., Yan, C., Jiang, M., Alenezi, F., Alhudhaif, A., Alnaim, N., ... &Wu, W. (2022). A novel facial emotion recognition method for stress inference of facial nerve paralysis patients. *ExpertSystems with Applications*, 197, 116705.
- Xu, X., Li, W., & Duan, Q. (2021). Transfer learning and SE-ResNet152 networks-based for small-scale unbalanced fish species identification. *Computers and Electronics in Agriculture*, 180, 105878., 1-3.
- Yani, M., Irawan, S., & Setianingsih, C. (2019). Application of Transfer Learning Using Convolutional Neural Network Method for Early Detection of Terry's Nail. *of Physics: Conference Series*, 1201, 012052. doi:10.1088/1742-6596/1201/1/012052
- Yıldız, İ., ve Karadeniz, A. T. (2019). Enhancement of breast cancer diagnosis accuracy with deep learning. *Avrupa Bilim ve Teknoloji Dergisi*, 452-462.
- Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks?. *Advances in neural information processing systems*, 27.
- Zhu A., vd. (2017, 5-8 Eylül). Wind power prediction based on a convolutional neural network [Sözlü bildiri]. 2017 International Conference on Circuits, Devices and Systems, IEEE, Chengdu.