



T.C.
SELÇUK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

EL YAZISI RAKAMLARINDAN CİNSİYET
TAHMİNİ: DERİN ÖĞRENME
YAKLAŞIMLARININ KARŞILAŞTIRILMASI
VE ANALİZİ

Seyit Mehmet ÇİFTÇİ
YÜKSEK LİSANS TEZİ

Bilişim Teknolojileri Mühendisliği Anabilim Dalı

ŞUBAT-2025
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Seyit Mehmet ÇİFTÇİ tarafından hazırlanan “El Yazısı Rakamlarından Cinsiyet Tahmini: Derin Öğrenme Yaklaşımlarının Karşılaştırılması ve Analizi” adlı tez çalışması 11/02/2025 tarihinde aşağıdaki jüri tarafından oy birliği / oy çokluğu ile Selçuk Üniversitesi Fen Bilimleri Enstitüsü Bilişim Teknolojileri Mühendisliği Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan

Dr.Öğr.Üyesi Hazım İŞCAN

.....

Danışman

Doç.Dr. Mehmet Akif ŞAHMAN

.....

Üye

Doç.Dr. Murat KÖKLÜ

.....

Yukarıdaki sonucu onaylarım.

Prof. Dr. Ömer Faruk YÜKSEL
FBE 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

Seyit Mehmet ÇİFTÇİ

11.02.2025

ÖZET

YÜKSEK LİSANS TEZİ

EL YAZISI RAKAMLARINDAN CİNSİYET TAHMİNİ: DERİN ÖĞRENME YAKLAŞIMLARININ KARŞILAŞTIRILMASI VE ANALİZİ

Seyit Mehmet ÇİFTÇİ

Selçuk Üniversitesi Fen Bilimleri Enstitüsü
Bilişim Teknolojileri Mühendisliği Anabilim Dalı

Danışman: Doç.Dr. Mehmet Akif ŞAHMAN

2025, 49 Sayfa

Jüri

Doç.Dr. Mehmet Akif ŞAHMAN
Dr.Öğr.Üyesi Hazım İŞCAN
Doç.Dr. Murat KÖKLÜ

Uzun zamandır, çeşitli alanlardaki bilgilerin dijitalleştirilmesi ve depolanması giderek daha önemli hale gelmiştir. Özellikle basılı, el yazısı veya görüntü verilerinin düzenlenebilir bir formata dönüştürülmesine yönelik önemli bir talep vardır. Hem dijitalleştirme süreci hem de bu tür bilgilerin tanınması sırasında çeşitli sorunlar ortaya çıkmaktadır. Araştırmacılar için bu alan oldukça çekicidir. Görüntü dosyalarından metinsel verileri çıkarmak ve saklamak için çeşitli teknolojik yaklaşımlar vardır. Yaygın olarak tanınan bir teknoloji olan Optik Karakter Tanıma (Optical Character Recognizer-OCR), taranan belgelerdeki veya görüntülerdeki metni, düzenlenebilir ve depolama için kullanılabilen dijital metne dönüştürür. Bununla birlikte, karakter özellikleri, boşluk eksikliği, metindeki eksen kaymaları ve özellikle el yazısı olanlar, taranan belgelerdeki görüntü kalitesinin bir sonucu olarak, çeşitli sorunlar ve zorluklar çıkarmaktadır. Yapay zekâ tabanlı tekniklerin kullanımı karakter tanıma doğruluğunu artırmıştır. Bu teknikler arasında Evrişimsel Sinir Ağları (CNN'ler) bilgisayarlı görme problemleri için tercih edilmektedir. Bu çalışmada, Türk Milli Eğitim Sistemindeki öğrencilerden el yazısı rakamlar toplanmış ve veriler uygun bir şekilde etiketlenmiştir. Veriseti ile öğrencilerin el yazılarının ve cinsiyetinin tespit edilmesi amaçlanmıştır. İyi bilinen beş farklı CNN modeli (Özelleştirilmiş CNN, Resnet-50, InceptionV3, Vgg16, Vgg19) kullanılmıştır. Bu modellerden elde edilen sonuçlar karşılaştırılmıştır.

Anahtar Kelimeler: Cinsiyet Tahmini, Derin Öğrenme, El Yazısı Rakamları, Evrişimsel Sinir Ağları, Optik Karakter Tanıma, Sınıflandırma

ABSTRACT

MS THESIS

**GENDER ESTIMATION FROM HANDWRITTEN DIGITS: A COMPARISON
AND ANALYSIS OF DEEP LEARNING APPROACHES**

Seyit Mehmet ÇİFTÇİ

**THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCE OF
SELÇUK UNIVERSITY
THE DEGREE OF MASTER OF SCIENCE
DEPARTMENT OF INFORMATION TECHNOLOGIES ENGINEERING**

Advisor: Assoc.Prof. Dr. Mehmet Akif ŞAHMAN

2025, 49 Pages

Jury

Assoc.Prof.Dr. Mehmet Akif ŞAHMAN

Assist.Prof.Dr. Hazım İŞCAN

Assoc.Prof.Dr. Murat KÖKLÜ

For a long time, the digitization and storage of information in various fields has become increasingly important. In particular, there is a significant demand for the conversion of printed, handwritten or image data into an editable format. Various issues arise both in the digitization process and in the recognition of such information. For researchers, this area is very attractive. There are various technological approaches to extract and store textual data from image files. Optical Character Recognizer (OCR), a widely recognized technology, converts text from scanned documents or images into digital text that can be edited and used for storage. However, character features, lack of whitespace, axis shifts in text, and especially handwritten ones, pose various problems and challenges as a result of the image quality in scanned documents. The use of artificial intelligence-based techniques has improved the accuracy of character recognition. Among these techniques, Convolutional Neural Networks (CNNs) are preferred for computer vision problems. In this study, handwritten digits were collected from students in the Turkish National Education System and the data were appropriately labeled. The dataset aims to identify students' handwriting and gender. Five different well-known CNN models (Customized CNN, Resnet-50, InceptionV3, Vgg16, Vgg19) were used. The results obtained from these models are compared.

Keywords: Gender Prediction, Deep Learning, Handwritten Digits, Convolutional Neural Networks, Optical Character Recognition, Classification

ÖNSÖZ

Bu çalışmanın gerçekleşmesinde, bilgilerini benimle paylaşan ve her zaman bana doğru yolu gösteren, hep yanımda olan değerli danışman hocam Doç.Dr. MEHMET AKİF ŞAHMAN'a ve diğer kıymetli hocalarıma teşekkürlerimi sunuyorum.

Bütün eğitim hayatım boyunca bana destek veren ve her zaman yanımda olan değerli aileme teşekkür ediyorum.

Seyit Mehmet ÇİFTÇİ
KONYA-2025



İÇİNDEKİLER

ÖZET	iv
ABSTRACT.....	v
ÖNSÖZ	vi
KISALTMALAR	ix
1. GİRİŞ	1
2. LİTERATÜR ARAŞTIRMASI	3
3. MATERYAL	6
3.1. Veri Seti	6
4. METOT.....	10
4.1. Makine Öğrenmesi.....	11
4.2. Yapay Sinir Ağları	12
4.3. Derin Öğrenme	13
4.3.1. Evrişimli sinir ağları	15
4.3.1.1. Evrişim katmanı	17
4.3.1.2. Düzleştirilmiş doğrusal birim katmanı (Relu)	18
4.3.1.3. Havuzlama katmanı (Pooling layer)	18
4.3.1.4. Tam bağlantılı katman	20
4.3.1.5. Dropout katmanı	21
4.3.1.6. Aktivasyon fonksiyonu katmanı (Activation layer).....	21
4.3.1.7. Normalizasyon katmanları	21
4.3.1.8. Eğitim tur sayısı (Epoch)	22
4.3.1.9. Hiper-Parametre	22
4.3.1.10. Mini-Batch boyutu	23
4.3.1.11. Öğrenme hızı.....	23
4.3.1.12. Kayıp fonksiyonu.....	24
4.3.2. CNN Mimarileri.....	24
4.3.2.1. VGG-16	24
4.3.2.2. VGG-19	25
4.3.2.3. ResNet 50.....	26
4.3.2.4. InceptionV3	28
4.4. Değerlendirme Metrikleri	29
4.4.1. Doğruluk (Accuracy)	30
4.4.2. Kesinlik (Precision)	30
4.4.3. Duyarlılık (Recall) / tespit oranı	31
4.4.4. F1 – Skoru.....	31
4.4.5. ROC-AUC eğrisi.....	31
4.4.6. Gerçek pozitif oran (Hassasiyet).....	32
4.4.7. Yanlış pozitif oran (Özgüllük).....	32

5. DENEYSEL ÇALIŞMALAR.....	33
5.1. El Yazısı Rakamlarının Tespiti.....	33
5.2. Özelleştirilmiş CNN modelin açıklaması	36
5.3. El Yazısı Rakamları ile Cinsiyet Tespiti.....	38
5.3.1. Test doğruluk ve zamana bağlı kayıp grafiği.....	38
5.3.2. Karmaşıklık matrisleri	40
5.3.3. AUC-ROC grafikleri.....	42
5.3.4. Yeniden yapılan çalışma.....	43
5.3.5. Modellerin performansı	44
6. SONUÇ VE ÖNERİLER.....	46
KAYNAKLAR	47



KISALTMALAR

OCR	: Optik Karakter Tanıma (Optical Character Recognizer)
PCR	: Basılı Karakter Tanıma (Printed Character Recognition)
HCR	: El Yazısı Karakter Tanıma (Handwritten Character Recognition)
K-NN	: K-en yakın komşu (K-Nearest Neighbours)
SVM	: Destek Vektör Makinesi (Support Vector Machine)
RBF	: Radyal Temel Fonksiyonu (Radial Basis Function)
YSA	: Yapay Sinir Ağları
CNN	: Evrişimli Sinir Ağları (Convolutional Neural Networks)
EMNIST	: Extended Modified National Institute of Standards and Technology
ANN	: Yapay Sinir Ağları (Artificial Neural Networks)
HMM	: Gizli Markov Modelleri (Hidden Markov Model)
BP	: Geriye Yayılım (Back Propagation)
DNN	: Derin Sinir Ağı (Deep Neural Network)
DBN	: Derin İnanç Ağı (Deep Belief Network)
ReLU	: Doğrultulmuş lineer birim (Rectified Linear Unit)
ILSVRC	: ImageNet Large Scale Visual Recognition Challenge
ROC	: Alıcı İşletim Karakteristiği (The Receiver Operating Characteristic)
SDNLL Likelihood)	: Boyuta Bağlı Negatif Günlük Olasılığı (Size Dependent Negative Log Likelihood)

1. GİRİŞ

Eğitim, toplumun ilerlemesini ve gelişmesini sağlayan en temel unsurlardan biridir. Bilgi ve öğrenme, insanların potansiyelini gerçekleştirmelerini ve geleceği inşa etmelerini sağlayan güçlü araçlardır. Günümüzde teknolojinin hızla evirildiği bir dönemde, yapay zekâ ve Optik Karakter Tanıma (Optical Character Recognizer-OCR) gibi yenilikçi teknolojiler, eğitim alanında dönüştürücü etkiler yaratmaktadır. OCR süreci, basılı ve el yazısı karakterlerin tanımlanmasını ve dijital formata dönüştürülmesini, bunların düzenlenmesine ve saklanmasına olanak sağlamayı içerir. Bu dönüştürme, dijital platformlarda uygulanabilirliği ve metne erişimi kolaylaştırır. OCR, örüntü tanıma, yapay zekâ ve makine görüşü gibi alanlardaki araştırmacıların büyük ilgisini çeken zorlu bir disiplin sunmaktadır. Metin içeren bir görüntüdeki karakterleri tanımlama ve bunları dijital formata dönüştürme yeteneği, işletmeler, kütüphaneler ve devlet kurumları da dahil olmak üzere çeşitli sektörlerde, operasyonlarını kolaylaştırdığı için çok önemlidir (Liu ve ark., 2018).

OCR iki ana kategoriye ayrılır: Basılı Karakter Tanıma (Printed Character Recognition - PCR) ve El Yazısı Karakter Tanıma (Handwritten Character Recognition - HCR) (Tawde ve Kundargi, 2013). PCR'de belirli bir yazı tipinin kullanılması nedeniyle karakterlerin tanımlanması HCR'ye göre nispeten daha basittir. Bunun nedeni, HCR'de kullanılan yazı tipinin, tek tek harfleri ayırt etmeyi ve ayırmayı zorlaştıran karmaşık yapısından kaynaklanmaktadır. HCR ayrıca iki alt kategoriye ayrılmıştır: Çevrimiçi El Yazısı Karakter Tanıma ve Çevrimdışı El Yazısı Karakter Tanıma (Vinjit ve ark., 2020). Çevrimiçi tanımda, kullanıcıların kalem vuruşları yazarken takip edilebildiği için el yazısı karakterlerin tanınması daha kolay ve doğruluk oranı daha yüksektir. Ancak çevrimdışı el yazısı tanıma, karakterlerin belgelere yazılmasını içerir. Çevrimdışı el yazısı tanıma sayesinde insanlar aynı dili ve alfabeyi kullansalar bile el yazıları birbirinden farklı olacak ve bu da tanımayı araştırmacılar için ciddi bir zorluk haline getirecek. Bir kişinin el yazısındaki farklılıkların biyometrik veriler gibi dahi değerlendirilebileceği iyi bilinmektedir (Al-Shamaileh ve ark., 2019). Çevrimdışı görüntüler genellikle kameralar veya optik okuyucular kullanılarak yakalanır ve karakter tanıma genellikle görüntü işleme ve yapay zekâ tabanlı yaklaşımlar kullanılarak gerçekleştirilir.

Son on yılda teknolojinin hızla gelişmesiyle birlikte yapay zekâ tekniklerini kullanan HCR ile ilgili birçok çalışma yapılmıştır. K-En Yakın Komşu (k-Nearest Neighbours- k-NN) yaklaşımı çeşitli HCR çalışmalarında en yaygın yaklaşımlardan

biridir. Modelin bir eğitim süreci gerçekleştirmek yerine tüm eğitim verilerini tahmin amacıyla sakladığı bir tembel öğrenme yöntemidir. Basit ve etkili olması nedeniyle sınıflandırma ve regresyon işlemlerinde sıklıkla kullanılır (Babu ve Soumya, 2019). Romulus ve arkadaşları, antik Batak sembollerinin görüntüsünü Latin karakterlerine dönüştüren bir uygulama geliştirmiştir. Bu uygulamada k-NN yaklaşımı kullanılmış ve %96 oranında sınıflandırma başarısı elde edilmiştir (Romulus ve ark., 2015). Inkeaw ve arkadaşları, Lanna Dharma HCR için üç farklı özellik çıkarma yöntemi kullanmış ve sınıflandırıcı olarak k-NN'yi kullanmıştır (Inkeaw ve ark., 2015). Destek Vektör Makinesi (Support Vector Machine-SVM), yüksek boyutlu verilerle karmaşık problemleri çözmeye yönelik başarılı bir makine öğrenme algoritmasıdır. Rajakumar ve Bharathi, Tamil yazı tipi tanıma için bir algoritma önerdi ve yazı tiplerini sınıflandırmak için sınıflandırıcı olarak SVM'yi kullanmıştır. Çalışmalarında %94 gibi yüksek bir doğruluk elde etmişlerdir (Rajakumar ve Bharathi, 2012). Sandhya ve Krishnan'ın çalışmasında, Uç nokta algoritması kullanılarak bozulmuş Kannada karakterleri yeniden oluşturuldu ve Kannadaca karakter tanıma için YSA kullanıldı. Tek katmanlı ileri beslemeli YSA kullanarak yaklaşık %99 başarı elde ettiler (Sandhya ve Krishnan, 2016). Raj ve Wahi, Malayalam el yazısı tanıma için YSA'yı sınıflandırıcı olarak kullandı (Raj ve Wahi, 2013). Evrişimli Sinir Ağları (Convolutional Neural Networks-CNN), görüntü sınıflandırma, nesne algılama ve görüntü bölümlenme gibi bilgisayarlı görme uygulamalarında oldukça başarılı olan özel bir yapay sinir ağı türüdür. CNN, Çin tarihi belgelerinde karakter konumu tespiti için Yang ve diğerleri tarafından kullanılmıştır (Yang ve ark., 2018).

Bu tez çalışmasında, öncelikle öğrencilere ait el yazısı rakamlarını tanıma işlemi gerçekleştirilmiştir. Sonrasında ise diğer çalışmalardan farklı olarak bu tez çalışmasında cinsiyet tahmininin gerçekleştirilmesi üzerine çalışılmıştır. Bu konuda çalışmak için, orijinal bir veri seti hazırlanmıştır. Milli eğitim sisteminde okuyan farklı cinsiyet ve yaş gruplarından elde edilen el yazıları, belirlenen bir standart ile toplanmıştır. Toplanan veriler etiketlenerek cinsiyet ve yaş gibi farklılaşan bilgiler ile her bir karakter ve rakam uygun ebatlarda kesilerek isimlendirilmiştir. Bu çalışma için toplanan veriler ile eğitim alanındaki farklı çalışmalar için kullanabilecek iyi bir veri seti oluşturmuştur. Veriler cinsiyete bağlı olarak etiketlenmiştir. Cinsiyet tahmini için en iyi bilinen derin öğrenme yaklaşımları ve bu tez çalışmasında önerilen bir CNN modeli kullanılmıştır. Elde edilen sonuçlar temel sınıflandırma metrikleri ile kıyaslanmıştır.

2. LİTERATÜR ARAŞTIRMASI

El yazısı tanıma sistemleri birçok araştırmacının dikkatini çekmiştir ve son yıllarda makine öğrenmesi algoritmalarının gelişmesi ve çeşitlenmesi ile birçok farklı çalışma yapılmıştır. Bu çalışmalar örnekler aşağıda detaylı olarak verilmiştir.

Taşdemir yaptığı çalışmada Türkçe el yazısı tanıma sistemi kurmuştur. Bu sistem için uygun veri setleri olmadığından yapay veri seti üretmiştir. Gerçek ve yapay veri ile eğitilen sistem, 2041 kelimelik gerçek veri ile yapılan test sonucunda Türkçe karakter tanıma oranını %61'den %88'e çıkardığı görülmüştür. Alınan sonuçlara dayanarak yapay veri kullanmanın veri yetersizliği problemine bir çözüm olabileceği söylenebilir (TAŞDEMİR, 2021).

Kabakuş ve Erdoğan (Kabakus ve Erdogmus, 2021), Türkçe karakter tanımaya yönelik çalışmalarında CNN'e dayalı yeni bir el yazısı Türkçe harf tanıma modeli önerdiler. 25875 örnekten oluşan bir Türkçe el yazısı veri seti oluşturdular. Önerilen modelin doğruluğu %96,07 gibi yüksek bir değerle hesaplandı. Bu model, EMNIST (Extended Modified National Institute of Standards and Technology) veri seti üzerinden eğitilip değerlendirilmiş ve %94,61'lik bir doğruluk elde etmiştir.

Swapna ve arkadaşları (Swapna ve ark., 2021), EMNIST veri setinden 112800 eğitim verisi ve 18800 test verisi olmak üzere toplam 131600 veri ile yaptıkları çalışmada, CNN kullanmışlardır. Altı katmandan oluşan bu el yazısı tanıma modeli ile %86 doğruluk oranını yakalamışlardır.

Çapar ve arkadaşları (Çapar ve ark., 2003), Türkçe büyük harfli el yazısı karakteri üzerindeki çalışmaları için 20000 karakterden oluşan bir veri seti hazırlamışlardır. Çalışmalarında, en yakın komşu, en yakın ortalama komşu, Bayes ikinci dereceden sınıflandırıcı, Parzen sınıflandırıcı, Boyuta Bağlı Negatif Günlük Olasılığı (Size Dependent Negative Log Likelihood- SDNLL) ve Yapay Sinir Ağları (Artificial Neural Networks- ANN) kullanılmıştır. En iyi sınıflandırma yüzdesi SDNLL tarafından gerçekleştirilir. Çalışmalarında elde edilen en iyi sınıflandırma sonucu %93,6'dır.

Vural ve arkadaşları (Vural ve ark., 2004), yaptıkları çalışmada Türkçe için çevrimiçi el yazı tanıma sistemi sundular. Tablet PC'ler ile kaydedilen her nokta için el yazısı verilerinden çeşitli dinamik özellikler çıkardılar ve harf ve kelime modellerini eğitmek için Gizli Markov Modelleri (Hidden Markov Model- HMM) kullandılar.

Yaptıkları çalışmada 1000 kelimelik bir sözlükten el yazısı kelimeleri tanımada %94 oranında başarı sağlamışlardır.

Bartos ve arkadaşları (Bartos ve ark., 2018), Macarca el yazısı tanıma üzerine yaptıkları çalışmaların sonucunda, veri seti için uygun özelliklerin çıkarılmasının önemli olduğunu, nokta dağılımlarına göre çıkarılan özniteliklerin kullanılması %94,4 sınıflandırma doğruluğu ile en ayırt edici öznitelikleri verirken, yalnızca dönüşümler ve seri genişletme özniteliklerinin kullanılması %58,9'a varan doğrulukları elde ettiler.

Erdoğan ve Tümer (Erdoğan ve Tümer, 2021), görüntü işleme yöntemleri kullanarak ön işleme aldıkları görüntüleri önce satırlara daha sonra kelimelere ve karakterlere bölerek elde edilen verileri CNN ile bir sınıflandırma modeli üretmişlerdir. Bu modeli, EMNIST veri seti ve belgelerden oluşturulan veri setini kullanarak eğittiler. Başarı oranı %87,81 olarak tespit ettiler.

Küçükşahin (Küçükşahin, 2019), Osmanlı harflerini tanımak için çok katmanlı CNN kullandı. Milli kütüphaneden topladığı çeşitli belgeler ile bir veri seti hazırladı. 24 sınıf için her birinde 400 resim olacak şekilde 9600 görüntü elde etti. Yaptığı çalışmaların sonuçları incelendiğinde, veri seti %80 ile %20'ye bölündüğünde CNN'nin hem eğitim hem de test için çok iyi performans gösterdiğini tespit etti.

Ahlawat ve arkadaşları (Ahlawat ve ark., 2020), önerdikleri bu çalışmalarında, CNN tabanlı el yazısı rakam tanıma için katman sayısı, adım boyutu, alıcı alan, çekirdek boyutu, dolgu ve seyreltme gibi çeşitli tasarım seçeneklerini araştırmışlardır. Kapsamlı deneyler gerçekleştirdiler ve MNIST veri seti ile %99,87'lik bir doğruluk oranına ulaştılar.

Liu ve arkadaşları (Liu ve ark., 2020), çalışmalarında MINST veri setini kullanmışlardır. YSA, SVM, CNN ve Geriye Yayılım (Back Propagation- BP) algoritmalarının el yazısı rakamlarını tanımadaki başarıları karşılaştırılmıştır. El yazısı rakam tanımada en iyi sonucun CNN algoritmasında olduğu ve %97,7 başarı sağlandığı tespit edilmiştir.

Hossain ve arkadaşları (Hossain ve Ali, 2019), el yazısı ile yazılmış rakamları tanıyabilen bir model üzerinde çalışmışlardır. Yaptıkları çeşitli deneylerde CNN'nin diğer sınıflandırıcı algoritmalarından daha iyi bir oranda başarılı olduğunun sonucuna varmışlardır. Rakam tanıma çalışmalarının, derin öğrenme tekniklerini geliştirmek için önemli avantajları olduğunu açıklamışlardır.

Ghosh ve arkadaşları (Ghosh ve Maghari, 2017), yayınladıkları yazılarında, performans ve doğruluk gibi bir çok faktör ele alınarak en iyi algoritmayı belirlemek için

karşılaştırmalar yapmışlardır. Derin sinir ağının (Deep Neural Network – DNN) doğruluk ve performans açısından en iyi olduğunu, CNN ve derin inanç ağının (Deep Belief Network – DBN) doğruluk açısından neredeyse eşit oldukları fakat DNN algoritmasının yürütme süresi olarak en başarılı algoritma olduğunu belirtmişlerdir.

Yılmaz ve arkadaşları (YILMAZ ve Çakar, 2023), okullarda kullanılması için bir uygulama geliştirmişler ve uygulamada kullandıkları sınıflandırmaların performanslarını karşılaştırmışlardır. MINST veri setinden 250 kişiden alınan 60.000 adet 0-9 arası rakamları olan el yazısı seti kullanılmıştır. YSA, CNN ve K-NN algoritmaları kullanılmış ve sırasıyla performans değerleri %98, %99,4 ve %86 olarak elde edilmiştir. En başarılı algoritma CNN olmuştur.

Görüldüğü gibi karakter tanıma ve el yazısı tanıma üzerine birçok çalışma gerçekleştirilmiştir. Fakat literatürdeki çalışmalar incelendiğinde cinsiyet tahmini üzerine yapılan bir çalışmanın olmadığı görülmüştür. Bu tez çalışmasında alandaki bu eksikliğin giderilmesi istenmektedir. Elde edilen sonuçlar, el yazısından cinsiyet tahmininin yapılabileceğini göstermiştir.

3. MATERYAL

Bu çalışmada, öncelikle Türkçe rakam ve bazı özel karakterlerin olduğu özgün bir veri seti oluşturulmuştur. Bu veri seti oluşturulurken öğrencilerden rakamlar, matematiksel işaretler ve bir pangram cümle yazması istenilmiştir. Bu araştırmadaki amaç ortaokul ve lise öğrencilerinden toplanan el yazısı rakamlarından doğru şekilde cinsiyet tanımlamasını sağlamaktır. Başlangıçta öğrencilerden gerekli bilgileri toplamak amacıyla optik formlar oluşturulmuş ve öğrencilerden bu formları doğru doldurmaları istenmiştir. Formlar öğretmenlerin gözetiminde dolduruldu, ön işleme tabi tutuldu, parçalandı ve python kodlamasıyla etiketlendi. Veri seti oluşturulduktan sonra sınıflandırma işlemi için birden fazla CNN modeli önerilmiştir. Ayrıntılar alt bölümlerde sunulmaktadır.

3.1. Veri Seti

Bu araştırmadaki veri seti için, harfler, rakamlar, matematiksel semboller ve bir pangram cümleden oluşan benzersiz bir veri kümesi oluşturulmuştur. Veri kümesi, her bir girdinin doğru veri toplama ve analiz kurallarına ve düzenlemelerine uymasını sağlamak için özenle tasarlanmıştır. Şekil 3.1'de gösterilen tasarımda toplam 8000 adet A5 formunun renkli olarak baskısı alınmıştır.

Basılan bu formlar, Konya ilinin merkez ilçelerinde bulunan okullara kayıtlı öğrencilere dağıtılmıştır. Öğrenciler, öğretmenlerinin rehberliğinde bu formları doldurmuşlardır.

Tablo 3.1 Cinsiyet ve yaşa göre öğrencilerden elde edilen temiz veri sayısı

Yaş	Erkek (0)	Kadın (1)
10(C)	128	146
11(D)	443	570
12(E)	571	596
13(F)	734	704
14(G)	536	586
15(H)	317	455
16(I)	227	416
17(J)	245	459
18(K)	157	180
Toplam	3358	4112
		7470

Ortaokul ve lise öğrencileri, cinsiyet ve yaş bilgilerini kodlamış istenen karakterleri ve pangram cümleyi yazmışlardır. Öğrencilerin doldurdıkları bu formlar, optik okuma cihazında işleme alınmıştır. Cinsiyet ve yaş bilgisi optik okuma cihazında okunmuş ve formun yüksek çözünürlükteki taraması yapılmıştır. Kaydedilen her görüntüye bir numara verilmiştir. Görüntülerin isimlendirilmesinde yaş ve cinsiyeti temsil eden karakterler eklenerek, her öğrencinin el yazısı görüntüleri ayrı ayrı kaydedilmiştir.

Konya İl Millî Eğitim Müdürlüğü ve Selçuk Üniversitesi iş birliği ile yapılan bu çalışmaya katkı sağladığınız için teşekkür ederiz. Lütfen harf ve karakterleri çerçevelere taşımadan yazınız.

CİNSİYET K YAŞ 8 9 10 11 12 13 14 15 16 17 18

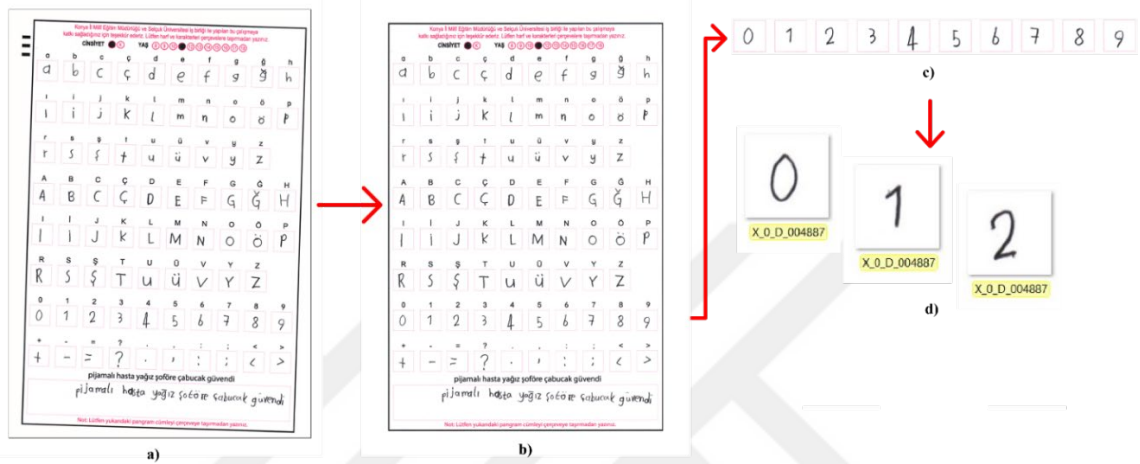
a	b	c	ç	d	e	f	g	ğ	h
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
ı	i	j	k	l	m	n	o	ö	p
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
r	s	ş	t	u	ü	v	y	z	
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
A	B	C	Ç	D	E	F	G	Ğ	H
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
I	İ	J	K	L	M	N	O	Ö	P
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
R	S	Ş	T	U	Ü	V	Y	Z	
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
0	1	2	3	4	5	6	7	8	9
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
+	-	=	?	.	,	:	;	<	>
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
pıjmalı hasta yağız şoföre çabucak güvendi									
<input type="text"/>									

Not: Lütfen yukandaki pangram cümleyi çerçeveye taşımadan yazınız.

Şekil 3.1 Veri seti toplamak için tasarlanmış optik form

Öğrencilerden cinsiyet ve yaşa bağlı olarak elde edilen temiz veri sayıları Tablo 3.1'de detaylandırılmıştır. 8000 öğrenciden elde edilen formlar detaylı bir şekilde analiz edilmiştir. Ancak bazı öğrenciler optik formların yanlış veya hatalı doldurulması nedeniyle veri tabanına dahil edilmemiştir. 3358 erkek ve 4112 kız öğrenciden toplam 7470 form elde edilmiştir.

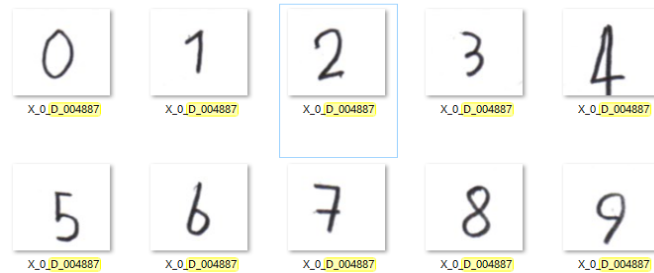
Şekil 3.2’de a aşamasında gösterildiği gibi taranmış ve etiketlenmiş form görüntülerinde taramadan kaynaklı eğiklikler görülmüştür. Hataların önüne geçmek ve karakterleri düzgün kesebilmek için tasarım sırasında eklenen siyah çerçeve, Python programında bu çalışmada yazılan yazılım ile görüntü işleme yöntemi kullanılarak yakalanmış ve düzeltilmiştir. Şekil 3.2’de b bölümünde gösterildiği gibi bütün formlar düzgün şekilde hizalanmıştır.



Şekil 3.2 Görüntü ön işleme aşamaları, a) Taranmış form, b) Düzeltilmiş kesilmiş form, c) Kesilen rakamlar, d) Yaş ve cinsiyete göre etiketlenmiş rakamlar

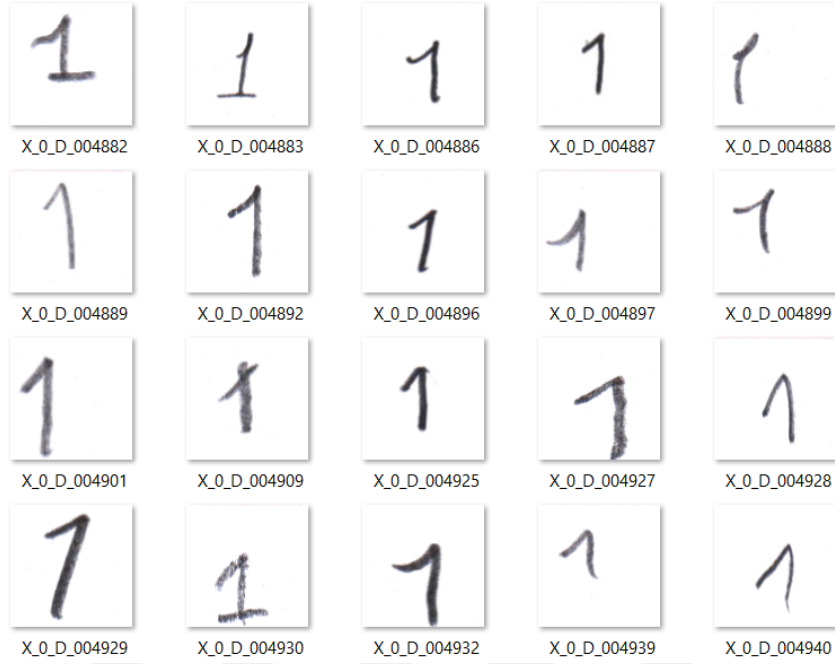
Python programında yazılan farklı bir yazılım ile bütün harfler ve karakterler uygun alanlardan kesilerek aynı dosya isimleri ile farklı klasörlerde biriktirilmiştir. 108x108 piksel ebatlarında 24 bit derinliğinde PNG resim olarak klasörlere ayrılmışlardır. Şekil 3.2’de bu aşamalar c ve d bölümünde gösterilmiştir.

Şekil 3.3’te bir öğrencinin doldurduğu bir formdan kesilen rakamlar vardır. X_0_D_004887 dosya ismi ile kayıtlı olan formda, cinsiyet erkek (0) yaş 11(D) olarak kodlanmıştır. Şekil 3.3’te gösterilen örnek rakamlar, taranmış ve kesilmiş bir öğrenci formunun, dosya ismi ile aynı isimde farklı klasörlerde biriktirilmiş görüntüleridir.

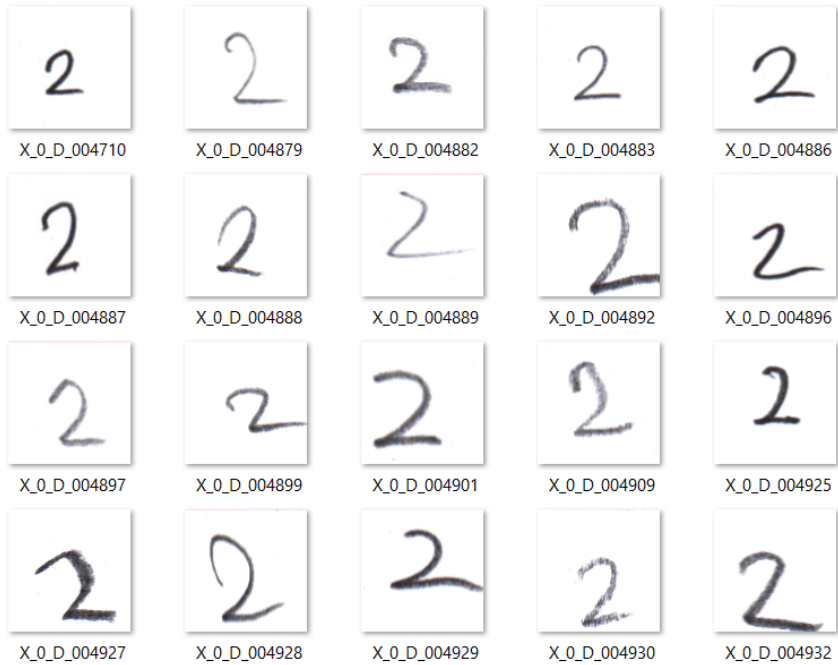


Şekil 3.3 004887 numaralı 11 yaşındaki erkek öğrencinin yazdığı rakamlar

Öğrencilerden elde edilen el yazısı rakamlarına ait örnek görseller aşağıda verilmiştir (Şekil 3.4 ve Şekil 3.5).



Şekil 3.4 Veri setindeki 1 rakamının bazı örnekleri

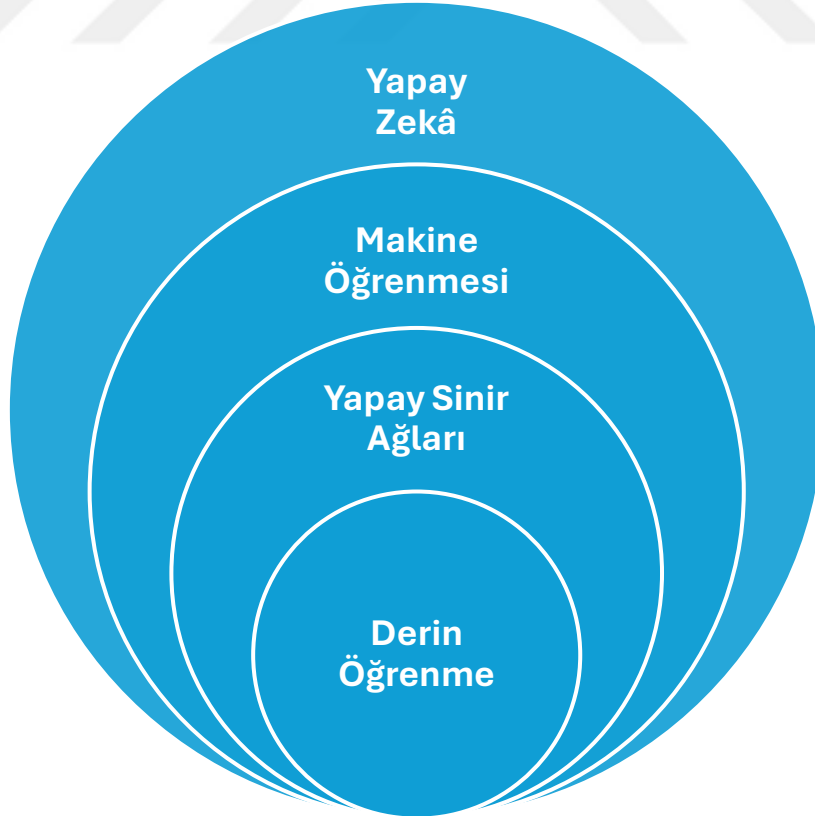


Şekil 3.5 Veri setindeki 2 rakamının bazı örnekleri

4. METOT

Yapay zekâ alanı, makinelerin insan zekasını taklit eden davranışlar sergileme kapasitesini kapsar. Bu yetenekler problem çözüme, öğrenme, karar verme, dil işleme ve görsel algıyı kapsamaktadır. Yapay zekanın genel amacı, insan zekasını kavrayarak ve taklit ederek makinelerin karmaşık görevleri özerk bir şekilde yerine getirmesini sağlamaktır (Şekil 4.1).

Makine öğrenimi, yapay zekanın bir alt dalını oluşturur ve yapay zekâ sistemlerinin temel bileşenlerinden biri olarak kabul edilir. Makine öğrenimi, karar verme veya tahmine dayalı modelleme hakkında bilgi vermek için veri çıkarma gibi süreçler yoluyla sistemlerin insan müdahalesine gerek kalmadan bilgi edinmesini ve gelişmesini sağlar. Makine öğrenimi, yapay zekanın çok sayıda uygulamasını mümkün kılan temel mekanizmayı sağladığı için yapay zekâ alanında büyük önem taşımaktadır. Makine öğrenimi yöntemleri, dil işleme, görüntü tanıma, tahmin modelleri ve oyun dahil olmak üzere çeşitli yapay zeka alanlarında uygulama alanı bulmaktadır.



Şekil 4.1 Yapay zekânın kapsadığı alanlar

4.1. Makine Öğrenmesi

Makine öğrenimi, yapay zekanın, görüntü kategorizasyonu, veri analizi ve tahmin gibi yalnızca insanların yapabileceği görevleri makinelerin gerçekleştirmesini sağlayan modeller oluşturmak için veri kümeleri üzerinde eğitilmiş algoritmalar kullanan bir alt alanıdır.

Görüntülerden bilgi çıkarma süreci, sınıflandırma ve kümeleme de dahil olmak üzere bir dizi görüntü işleme tekniği ve yapay zekâ uygulamasını içerir. Görüntü sınıflandırma, bu alanda sıklıkla kullanılan makine öğrenimi algoritmaları ile uzun yıllardır bir araştırma konusu olmuştur. Derin öğrenme, makine öğrenmesi alanının alt dalıdır, kanıtlanmış üstün performansı nedeniyle son zamanlarda bu alanda öncelikli bir yöntem olarak ortaya çıkmıştır (Çetinkaya ve Turhal, 2024).

Makine öğrenmesi, veri madenciliği alanına dâhil olan akademik bir disiplindir. Veri madenciliği, geniş veri kümelerinden geleceğe yönelik tahminler yapmayı sağlayan örüntü ve kuralların analiz edilmesini hedefler. Bu alanda kullanılan algoritmaların büyük bir kısmı, makine öğrenimi araştırmalarından kaynaklanmaktadır (Pekel, 2018).

Makine öğrenmesi sürecinin sorunları çözmek için nasıl işlediğine dair genel bir bakış aşağıda sunulmuştur:

- **Verileri toplama ve hazırlama:** Veri kaynakları tanımlandıktan sonra, mevcut veriler bir araya getirilir. Verilerin yapısı, tercih edilecek makine öğrenimi algoritmalarının seçiminde önemli bir rol oynar. Veri analizi sırasında, anomaliler belirlenir, veri yapısı iyileştirilir ve veri bütünlüğü ile ilgili problemler çözülür.
- **Modeli eğitme:** Hazırlanan veriler, eğitim seti ve test seti olmak üzere iki gruba ayrılır. Eğitim seti, veri kümesinin büyük bir kısmını kapsar. Bu veriler, makine öğrenimi modellerinin doğru şekilde çalışması ve en yüksek doğruluğa ulaşması için gerekli optimizasyonları yapmak amacıyla kullanılır.
- **Modeli doğrulama:** Son veri modelinizi belirlediğinizde, modelin performansını ve doğruluğunu ölçmek için test seti kullanılır.
- **Sonuçları yorumlama:** İçgörülerini elde etmek, sonuçlar çıkarmak ve gelecekteki tahminleri yapmak için elde edilen sonuçları dikkatlice analiz edilir.

Makine öğrenmesi ile aşağıdaki işlemler yapılabilir.

Değerleri Öngörme: Regresyon algoritmaları, değişkenler arasındaki neden-sonuç ilişkisini ortaya koymada etkilidir. Verilerden bir model oluşturur ve bu modeli kullanarak tahminlerde bulunurlar. Regresyon analizi, gelecekteki sonuçları tahmin etme konusunda güçlü bir araçtır. Bu özellik, ürün talebi, satış rakamları ve kampanya sonuçlarının tahmin edilmesine olanak tanır.

Olağandışı Oluşumları Tanımlama: Anomali tespit algoritmaları, normalden sapmalar gösteren verileri ayırt ederek olası riskleri tespit etmek için yaygın olarak kullanılır. Ekipman arızaları, yapısal hasarlar, metin hataları ve dolandırıcılık gibi sorunlar, makine öğrenimi teknikleri ile analiz edilerek çözülebilir.

Yapı bulma: Makine öğrenimindeki ilk adım, genellikle bir veri kümesindeki temel yapıyı belirlemek için kümeleme algoritmalarının uygulanmasıdır. Kümeleme algoritmaları benzer öğelerin gruplandırılması, pazar segmentasyonu için kullanılır ve bu, fiyatlandırma stratejilerini oluşturmak ve müşteri tercihlerinin tahmin edilmesine yardımcı olmak için içgörüler sağlar.

Kategorileri tahmin etme: Sınıflandırma süreci, verilerin doğru kategorilere ayrılmasında etkili bir yöntemdir. Kümeleme sürecine benzer, ancak sınıflandırma, önceden belirlenmiş etiketlerin veriye atanmasıyla yapılan denetimli öğrenme uygulamalarını içerir.

4.2. Yapay Sinir Ağları

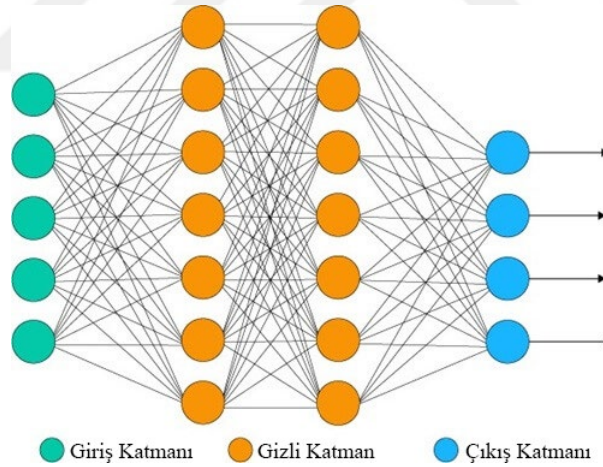
Son teknolojik gelişmeler, Yapay Sinir Ağları alanında önemli ilerlemeler göstermiş ve bununla beraber bu alana artan ilgiyi hızlandırmıştır (Aydın ve ark.). YSA'lar, biyolojik sinir hücrelerinin beynimizdeki işlevlerini taklit ederek problemlere çözüm bulmaya çalışan bir tekniktir. Ancak, insan beyninin karmaşıklığı nedeniyle, YSA'lar beynin gerçekleştirdiği tüm işlemleri tam anlamıyla taklit edemez. Yine de talimatlar yoluyla, YSA zamanla beynin işleyişini daha doğru bir şekilde taklit etmeyi öğrenebilir. (Aytekin, 2021)

Yapay Sinir Ağı yapısı, beş temel işlem elemanından oluşur. Girdi değerleri, ağırlıklar, toplama fonksiyonu, aktivasyon fonksiyonu ve çıktı değerleri (Şekil 4.2). Girdi değerleri, ağın dış kaynaklardan aldığı verileri temsil eder ve ağın öğrenme sürecinde kullandığı verilerdir. Ağırlıklar, bu girdilerin hücre üzerindeki etkisini anlamayı sağlar. Birleştirme fonksiyonu, bir hücredeki net girdiyi hesaplar ve bunun için farklı fonksiyonlar kullanılır. Aktivasyon fonksiyonu, net giriş sinyallerini çıkış sinyallerine

dönüştürmekten sorumludur; bu işlem, çıktı katmanına aktarılan değerler üzerinde gerçekleşir (Odabaşı ve Toklu, 2023).

Yapay Sinir Ağı, insan sinir sistemi ve bilişsel yeteneklerinden ilham alan bir makine öğrenme tekniğidir. YSA'lar, gelişmiş öğrenme yetenekleri sayesinde, geleneksel makine öğrenme tekniklerinin çözmekte zorlandığı karmaşık problemleri çözebilir. Yeni senaryolara genelleme yapabilme kapasitesi, YSA'nın öğrenme yeteneği sayesinde, karar verme süreçlerinde geçmiş bilgileri kullanarak kolaylaştırılır. Yapay sinir ağlarının yalnızca sayısal verilerle çalıştığını unutmamak önemlidir. Geçmiş verilerden öğrenerek, henüz gerçekleşmemiş olaylar hakkında tahminlerde bulunabilir. Bu özellik, YSA'ları mühendislik, sağlık hizmetleri ve endüstriyel süreçler gibi birçok alanda çok yönlü bir araç haline getirmektedir. Örneğin, sinir ağları sayesinde bir sistemde veya cihazda olası arızalar önceden tahmin edilebilir. YSA, genellikle tahmin, sınıflandırma ve veri kümeleme gibi işlemler için etkili bir şekilde kullanılmaktadır (Pekel, 2018).

Yapay sinir ağlarının temel olarak, doğrusal olmama, paralel çalışma, öğrenme, genelleme, hata toleransı ve esneklik, eksik verilerle çalışma, çok sayıda değişken ve parametre kullanma ve uyarlanabilirlik gibi nitelikleri vardır (Aytekin, 2021).



Şekil 4.2 Derin Öğrenme Katman yapısı

4.3. Derin Öğrenme

Derin öğrenme, insan beyninden esinlenen bir makine öğrenimi yöntemidir. En belirgin özelliği, ham verilerden otomatik olarak özellik çıkarabilme kapasitesidir. Derin öğrenme yöntemleri, sinir ağlarını, aşamalı olasılık modellerini, denetimli ve denetimsiz özellik öğrenme algoritmalarını içerir ve zamanla daha da gelişmektedir (Önün, 2019).

Derin öğrenme, özelliklerin çıkarılması ve dönüştürülmesi için doğrusal olmayan işlem birimlerinin çoklu katmanlarını kullanan bir makine öğrenimi alt dalıdır. Derin

öğrenme mimarisi, her katmanın bir önceki katmanın çıktısını girdi olarak kullanarak bilginin katmanlar boyunca özyinelemeli bir şekilde iletilmesini sağlar. Derin öğrenme algoritmaları, iki ana türe ayrılabilir: denetimli öğrenme, sınıflandırma gibi görevleri içerirken; denetimsiz öğrenme, örüntü analizi gibi süreçleri kapsar (Şeker ve ark., 2017).

Tablo 4.1 Derin Öğrenme tarihindeki önemli olaylar

Yıllar	Gerçekleşen Olaylar
1943	Elektronik Beyin
1957	Perceptron – Tek katmanlı algılayıcı
1960	Adaline
1965	Multi Layer Perceptron – Çok katmanlı algılayıcı
1979	Neocognitron
1986	Backpropagation
1991	Destek Vektör Makineleri (SVM-Support Vector Machine)
2006	Deep Neural Networks – Derin Sinir Ağları

Bu alanın tarihi, yapay zekâ ve makine öğrenimi gibi ilgili konuların gelişimiyle yakından bağlantılıdır. Derin öğrenmenin modern versiyonunun temelleri, 1940'larda yapay sinir ağlarının ortaya çıkışına dayanırken, daha net bir zaman çizelgesi 20. yüzyılın sonlarından 21. yüzyılın başlarına kadar izlenebilir (Tablo 4.1).

1943 yılında Warren McCulloch ve Walter Pitts, insan beynindeki sinir hücrelerini taklit etmeye yönelik 'Yapay Sinir Ağı' adı verilen ilkel bir model geliştirdiler. Bu model, sinir hücrelerinin işleyişini taklit etme potansiyeli nedeniyle yapay zekâ araştırmalarında büyük ilgi uyandırmış ve popülerlik kazanmıştır. (McCulloch ve Pitts, 1943)

2010'lar ve sonrasındaki yıllarda, derin öğrenme birçok alanda öncü bir teknoloji olarak öne çıkmıştır. Derin sinir ağları, konvolüsyonel sinir ağları (CNN'ler) ve tekrarlayan sinir ağları (RNN'ler) gibi pek çok mimari ve algoritma geliştirilmiştir. Günümüzde derin öğrenme, çeşitli endüstrilerde geniş bir uygulama alanına sahip olmuş ve sürekli olarak gelişmeye devam etmiştir. Özellikle otomatik sürüş, tıbbi görüntüleme, finansal analiz ve oyun endüstrisi gibi alanlardaki kullanımları büyük ilgi uyandırmıştır (Pichler ve Hartig, 2022).

En iyi bilinen ve en yaygın kullanılan üç derin öğrenme algoritmaları Derin Sinir Ağı (DNN), Evrişimli Sinir Ağı (CNN) ve Tekrarlayan Sinir Ağı (RNN) algoritmalarıdır. Bu tez çalışmasında CNN algoritmalarının en popüler olan modelleri de incelenmektedir.

4.3.1. Evrişimli sinir ađları

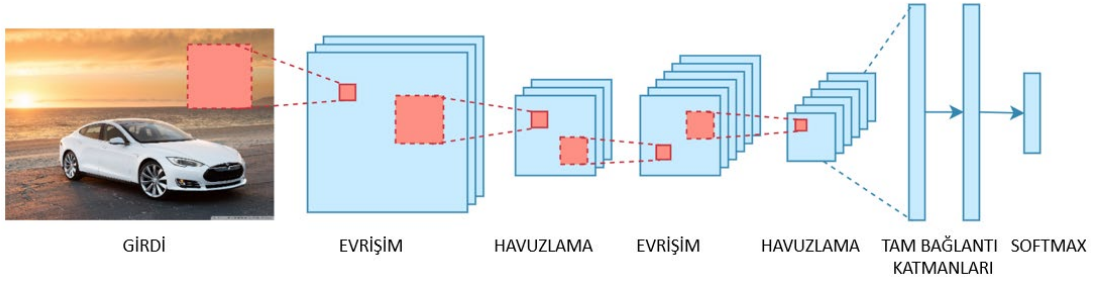
Evrişimli sinir ađları (CNN'ler), biyolojik sinir ađlarından esinlenilmiş ve özellikle görüntü ve uzamsal verileri işlemek için optimize edilmiş bir yapay sinir ađı türüdür. Klasik yapay sinir ađlarının aksine, CNN'ler görüntülerdeki yerel özellikleri tanımak ve bu özellikleri soyut seviyelere çıkarmak için özel olarak tasarlanmış katmanlar içerir.

Evrişimli Sinir Ađları (CNN'ler), özellikle görüntü işleme alanında yaygın olarak kullanılan yapay sinir ađlarıdır. CNN'lerin temel bileşenlerinden biri, giriş verilerine uygulanan öğrenilebilir filtrelerdir. Bu filtreler, görüntülerdeki yerel örüntüleri ve özellikleri tanımlamak amacıyla kullanılır. CNN'ler, genellikle bir dizi konvolüsyonel katman kullanır. Konvolüsyon, giriş görüntülerindeki özellik haritalarını anlamlı ve uzamsal bir şekilde indirgemeyi amaçlayan bir süreçtir. Bu işlem, ađın temsil gücünü artırırken aynı zamanda hesaplama karmaşıklığını azaltır.

Görüntüdeki bu özellik haritaları daha sonra ađın eğitim seviyesini artırmak amacıyla doğrusal olmayan aktivasyon fonksiyonlarına tabi tutulur. Ađın ilk katmanları genellikle basit ilişkileri ayırt edebilirken, sonraki katmanlar daha karmaşık ve soyut ilişkileri tespit edebilir. Bu hiyerarşik yapısı, CNN'lerin daha yüksek doğrulukla öğrenmesini sağlar.

CNN'ler, geri yayılım algoritması kullanılarak eğitilen ileri beslemeli ađlardır ve genellikle çok katmanlı algılayıcılar olarak adlandırılır. CNN'lerin girdileri, katmanlar aracılığıyla işlenir ve her katman, ađın öğrenme yeteneğini artıran farklı bir işlemi gerçekleştirir (Şekil 4.3). Bu çok katmanlı yapı sayesinde, ađ görüntülerdeki soyut ve yüksek düzeydeki özellikleri etkili bir şekilde öğrenebilir.

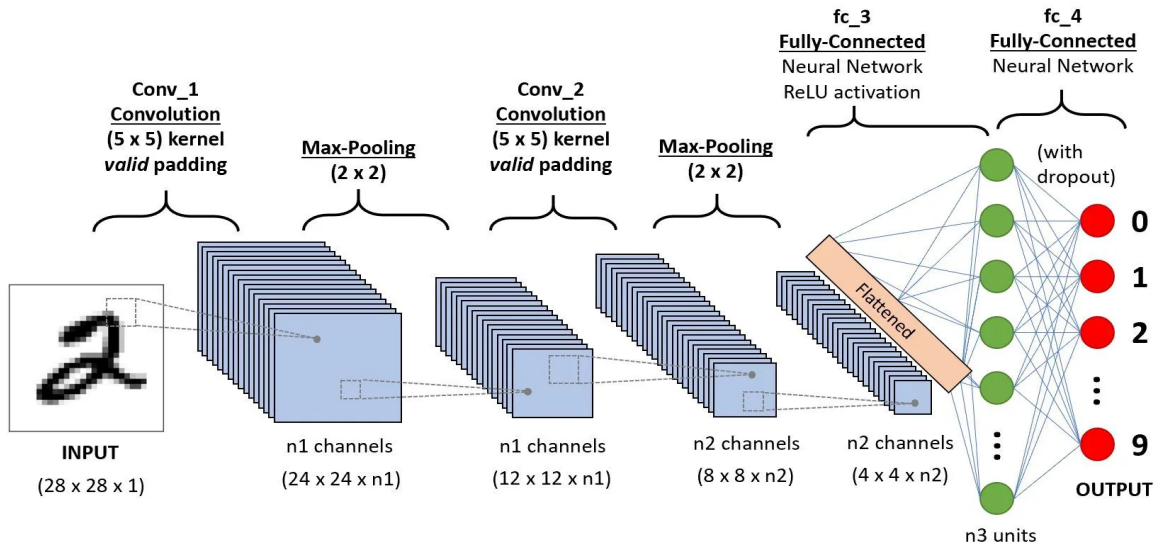
Evrişimli Sinir Ađı (CNN) mimarisinde, ilk katman genellikle konvolüsyonel katmandır. Bunu takip eden katmanlar arasında doğrusal olmayan aktivasyon fonksiyonu olan Doğrultulmuş Lineer Birim (ReLU), havuzlama katmanı, normalleştirme (düzleştirme) katmanı ve tam bađlı katmanlar yer alabilir. Bu katmanlar, ađın görüntü verilerindeki önemli özellikleri öğrenmesini sağlayarak, daha karmaşık yapıları tanımlama kapasitesini artırır (Küçükşahin, 2019).



Şekil 4.3 CNN katman mimarisi

Çok katmanlı sinir ağlarında bir dizi farklı katman bulunur; bunlar arasında birden fazla konvolüsyon katmanı, tam bağlı katmanlar, aktivasyon katmanları, sınıflandırıcı katmanlar, havuzlama katmanları ve çeşitli ek katmanlar yer alır. Her bir katman, belirli bir işlevi yerine getirerek ağın öğrenme sürecine katkıda bulunur ve nihayetinde elde edilen çıktı, sınıflandırıcı katmanda oluşturulur. Konvolüsyonel Sinir Ağları (CNN'ler), derin öğrenme yapılarında en yaygın kullanılan sinir ağı türünü oluşturur. Bu ağlar, özellikle sınıflandırma ve algılama gibi görevler için tercih edilmektedir (Doğan ve Türkoğlu, 2019).

Imagenet tarafından 2012 ve 2014 yıllarında düzenlenen ve büyük ölçekli görsel tanıma yarışması olarak bilinen Büyük Ölçekli Görsel Tanıma yarışmasında, en yüksek başarıyı Konvolüsyonel Sinir Ağları (CNN'ler) elde etmiştir. Bu sonuçlar, CNN'lerin görsel tanıma alanındaki etkinliğini ve gücünü bir kez daha gözler önüne sermiştir (Doğan ve Türkoğlu, 2019). CNN'in Temel bileşenlerinden aşağıda bahsedilmiştir (Şekil 4.4).



Şekil 4.4 CNN Katman yapısı

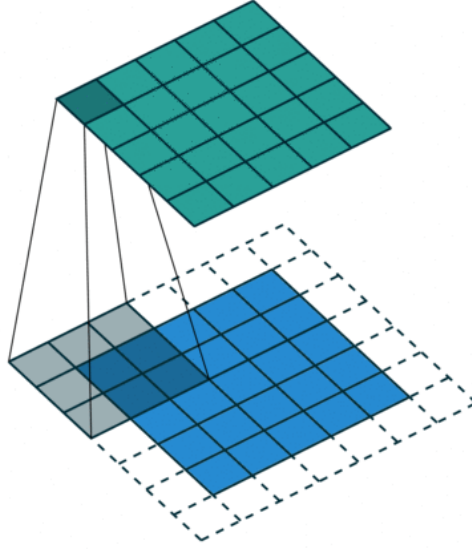
4.3.1.1. Evrişim katmanı

Konvolüsyon işlemi de CNN'lerin temel bir bileşenidir. Genellikle çekirdek veya filtre olarak adlandırılan özellik dedektörü tarafından giriş görüntüsüne uygulanır. Filtreler 3x3 veya 5x5 gibi küçük boyutlara sahiptir ve derinlikleri, girdinin siyah-beyaz veya renkli görüntü olmasına bağlı olarak farklılık gösterir. Başlangıçta, filtre giriş matrisinin sol üst kısmına yerleştirilir, ardından eşleşen hücrelerin sayısı sayılır. Bu sayı, özellik haritası adı verilen yeni bir matrise eklenir. Özellik dedektörü sağa kaydırıldıkça ve giriş matrisinin sağ alt sınırına ulaşana kadar aşağı doğru hareket ettikçe, eşleşen hücrelerin sayısı sırasıyla sayılır ve özellik haritasına eklenir. Özellik dedektörleri, özellik haritasının boyutunu küçültmek için kullanışlıdır.

Giriş görüntüsüne birden fazla evrişim işlemi uygulanmasının, ağın derinliğini artırarak daha hassas ve doğru sonuçların elde edilmesini sağladığı literatürde ortaya konulmuştur. Her bir evrişim katmanında, filtreler önceki katmanlardan elde edilen görüntü verilerine uygulanmaktadır. Bu süreç sonucunda giriş görüntüsünün en-boy oranı azalırken, derinlik seviyesi artmakta ve nihai olarak görüntünün özellik haritası elde edilmektedir. Dikkat çekici bir şekilde, alt seviyedeki katmanlardan birleştirilen özelliklerin ayırt edicilik kapasitesi düşükkken, daha yüksek seviyedeki katmanların birleştirilmesiyle oluşan özelliklerin ayırt ediciliğın arttığı gözlemlenmiştir (Bayram, 2020). Evrişim (Convolution) Katmanının Çalışma Mantiğı aşağıda verilmiştir.

Evrişim katmanı, bir görüntü üzerinde belirli bir özelliği (örneğin kenarları, köşeleri veya dokuları) algılamak için küçük bir filtreyi (kernel veya weight matrix) görüntü üzerinde kaydırır ve her pozisyonda matematiksel bir işlem gerçekleştirir.

- **Girdi (Input):** Genellikle bir görüntü matrisidir (örneğin, 2D bir matris olarak ifade edilen gri tonlamalı bir görüntü veya 3D bir tensor olan RGB görüntü).
Örneğin, 28x28 boyutunda bir görüntü giriş olarak alınabilir.
- **Filtre (Kernel):** Küçük bir ağırlık matrisi (örneğin 3x3 veya 5x5) giriş üzerinde kaydırılarak belirli bir özelliği çıkarır.
Her filtre farklı bir özelliği (örneğin, yatay kenar, dikey kenar, dokular) öğrenir. Filtreler öğrenilir, yani model eğitimi sırasında optimize edilir.
- **Filtreyi Uygulama (Sliding Window):** Filtre, giriş görüntüsü üzerinde belirli bir adım (stride) aralığıyla kaydırılır. Her pozisyonda filtre ve görüntü parçacığı (patch) arasında nokta çarpımı (dot product) yapılır (Şekil 4.5).



Şekil 4.5 CNN Filtre Uygulaması

4.3.1.2. Düzleştirilmiş doğrusal birim katmanı (Relu)

Düzleştirilmiş Doğrusal Birim Katmanı (Rectified Linear Units Layer(ReLU)), konvolüsyon katmanlarının ardından yer alan bu katman, ESA nöronlarının çıktılarına en sık aktive edilen doğrultucu birim olarak tanımlanmış ve literatürde aktivasyon katmanı olarak adlandırılmıştır. Temel işlevi, negatif değerleri sıfıra indirgemek suretiyle giriş verilerini dönüştürmektir.

Bu katmandan önce gelen konvolüsyon katmanında belirli matematiksel işlemlerin gerçekleştirilmesi, ağıın doğrusal bir yapı sergilemesine yol açmaktadır. Ancak, aktivasyon katmanının uygulanması, ağıın doğrusal olmayan bir yapıya dönüşmesini sağlar. Bu yapısal dönüşüm, derin öğrenme ağıının daha hızlı bir öğrenme süreci sergilemesine önemli ölçüde katkıda bulunmaktadır (İnik ve Ülker, 2017).

4.3.1.3. Havuzlama katmanı (Pooling layer)

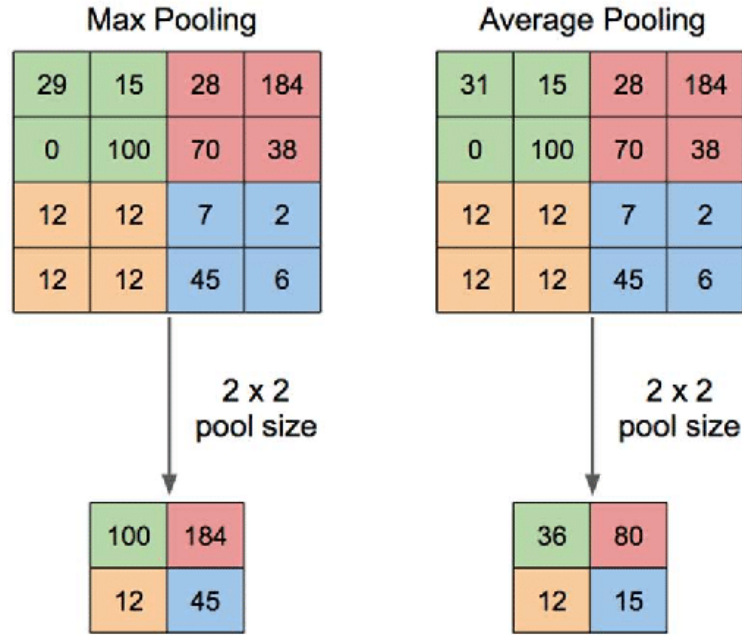
Havuzlama işlemi, ağa görüntü içindeki belirli örüntüleri tespit etme yeteneği sağlar. Evrişimsel katmandan elde edilen özellikler yüksek sayıda olabilir ve çok fazla özellik üzerinde eğitim aşırı uyuma neden olabilir. Bu nedenle, boyutları azaltmak için genellikle bir havuzlama işlemi uygulanır. Ortalama havuzlama, toplam havuzlama ve en yaygın kullanılan operatör olan maksimum havuzlama gibi farklı havuzlama operatörleri vardır.

Maksimum havuzlamada, havuzlama operatörü bir özellik dedektörüne benzer şekilde tüm giriş görüntüsü boyunca kaydırılır, her kaydırmada alt bölgenin en büyük piksel değerini alır.

Havuzlama işlemi, genellikle ReLU katmanından sonra konumlandırılır ve temel amacı, bir sonraki konvolüsyon katmanı için giriş boyutlarını (Genişlik x Yükseklik) azaltmaktır. Bu işlem literatürde "Aşağı Örnekleme" olarak adlandırılmaktadır. Havuzlama katmanında gerçekleştirilen bu boyut azaltma işlemi, bilgi kaybına yol açsa da ağıın performansı açısından iki önemli avantaj sunar. İlk olarak, sonraki katmanlar için hesaplama maliyetini azaltarak işlem verimliliğini artırır. İkinci olarak, sistemin ezberleme eğilimini azaltır ve daha genelleştirilebilir bir öğrenme sağlar (Şekil 4.6).

Konvolüsyon işleminde olduğu gibi, havuzlama katmanında da belirli filtreler tanımlanır. Bu filtreler, önceden belirlenmiş bir adımlama değerine göre görüntü üzerinde kaydırılarak uygulanır. İşlem sırasında, görüntü piksellerinin maksimum değerleri (maksimum havuzlama) veya değerlerinin ortalaması (ortalama havuzlama) alınır. Maksimum havuzlama genellikle daha iyi performans sergilediği için tercih edilmektedir. Havuzlama işlemi, konvolüsyon katmanından elde edilen her bir filtre çıktısı için ayrı ayrı gerçekleştirilir (İnik ve Ülker, 2017).

- **Maksimum havuzlama (Max pooling)** : Maksimum havuzlama tekniği, her bir havuzlama bölgesindeki en büyük elemanı belirlemek amacıyla uygulanmaktadır. Ancak, bu teknikle ilgili temel bir sınırlama, havuzlama bölgesindeki diğer tüm unsurların tamamen göz ardı edilmesi ve yalnızca en büyük unsura odaklanılmasıdır. Bu yaklaşım, hesaplama maliyetini azaltma açısından avantaj sağlasa da probleme özgü kritik ayırt edici özelliklerin tanımlanmasını riske atabilir. Dolayısıyla, maksimum havuzlama kullanımı, bilgi kaybının etkilerini minimize edecek şekilde dikkatle değerlendirilmelidir (Zafar ve ark., 2022).
- **Ortalama havuzlama (Average pooling)** : Ortalama havuzlamanın görüntüye uygulanması görüntünün yumuşamasına neden olur ve bu da bu özel havuzlama yöntemi kullanıldığında keskin özelliklerin tespit edilmesini tehlikeye atabilir. Ortalama havuzlama, her havuzlama bölgesindeki piksellerin ortalama değerinin hesaplanmasını içerir ve görüntüdeki ince ayrıntıların kaybolmasına neden olur. Birden fazla pikselin tek bir pikselde birleştirilmesi kaçınılmaz olarak orijinal görüntüde bulunan ince değişikliklerin ve kenar bilgilerinin kaybolmasına neden olur. Ek olarak, ortalama havuzlama uzamsal olarak değişmez bir işlemdir, yani ortalama havuzlama sonucu, özelliğin görüntüdeki konumundan bağımsız olarak aynı olacaktır. Bu durum, nesne algılama görevleri gibi konumsal bilginin önemli olduğu bazı durumlarda arzu edilmeyebilir.

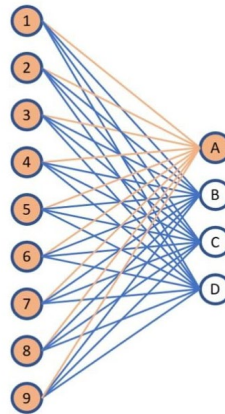


Şekil 4.6 Maksimum havuzlama ve ortalama havuzlama

4.3.1.4. Tam bağlantılı katman

Sinir ağırları, bir dizi bağımlı doğrusal olmayan fonksiyon olarak tanımlanır ve her fonksiyon bir nöron içerir. Tam bağlantılı katmanlar bağlamında, nöronlar giriş vektörüne doğrusal bir dönüşüm uygulamaktan sorumludur ve bu işlem için bir ağırlık matrisi kullanılır (Şekil 4.7).

Bu özellik, "tam bağlantılı" veya bazen "yoğun şekilde bağlı" terimlerinin ortaya çıkmasına yol açmaktadır. Katmandan katmana tüm olası bağlantıların mevcut olduğuna dikkat edilmelidir; yani, giriş vektörünün her elemanı, çıktı vektörünün her elemanını etkiler. Ancak, önemli bir nokta, tüm ağırlıkların tüm çıktılar üzerinde etkili olmadığıdır.



Şekil 4.7 Tam Bağlantılı Katman Yapısı

4.3.1.5. Dropout katmanı

Bu katman, aynı zamanda bir seyreltme katmanı olarak da kabul edilmektedir ve birincil işlevi aşırı öğrenmeyi(overfitting) önlemektir. Bu katman, sinir ağındaki nöronları rastgele devre dışı bırakmak için kullanılır ve kullanıcı tarafından çıkarılacak nöronların yüzdesi belirlenir. Bu yöntem, ağın aşırı öğrenmeyi engelleyerek genelleştirme kabiliyetini artırır ve performansını iyileştirir. Seyreltme katmanı, modern sinir ağı mimarilerinin temel bileşenlerinden biri olarak kabul edilir ve ağın etkinliğini ve bütünlüğünü sağlamak amacıyla tam bağlantılı katmanlarla birlikte yaygın olarak kullanılır.

4.3.1.6. Aktivasyon fonksiyonu katmanı (Activation layer)

Aktivasyon fonksiyonları, doğrusal olmayan gerçek dünya özelliklerini yapay sinir ağlarına entegre etmek için kritik bir role sahiptir. Yapay sinir ağlarında Rectified Linear Unit (ReLU), Tanh ve Sigmoid gibi birçok aktivasyon fonksiyonu kullanılmaktadır.

Çıktı katmanında tercih edilen aktivasyon fonksiyonu, modelin çözmeye çalıştığı problemin özelliklerine bağlıdır. Örneğin, kedi ve köpek görüntülerini ayırt etmek gibi ikili sınıflandırma problemlerinde, sigmoid aktivasyon fonksiyonunun çıktı katmanında kullanılması önerilir. Buna karşın, çoklu sınıflandırma problemlerinde, softmax aktivasyon fonksiyonu genellikle çıktı katmanında tercih edilmektedir.

4.3.1.7. Normalizasyon katmanları

Adından da anlaşılacağı üzere, toplu normalizasyonun eğitim sırasında toplu işlemlerle bir ilgisi olması gerektiğini tahmin edebilirsiniz. Basitçe söylemek gerekirse, toplu normalizasyon tek bir toplu iş boyunca bir katmanın girdisini standartlaştırır.

Belirli bir katmanda ortalama ve varyansın hesaplanarak ardından normalizasyon yapılmasının daha verimli bir yöntem olabileceği öne sürülebilir. Ancak, eğitim süreci boyunca model parametrelerinin değişime uğraması bu yaklaşımı zorlaştırmaktadır. Bu durum, ara katmanlardaki aktivasyonların sürekli dalgalanmasına neden olur. Aktivasyonların her iterasyonda değiştiği göz önüne alındığında, her iterasyon için tüm eğitim seti boyunca ortalama ve varyansı hesaplamak hem zaman alıcı hem de gereksiz bir çaba olabilir. Bu bağlamda, toplu normalizasyon (batch normalization) kavramı, bu zorlukların üstesinden gelmek için kritik bir öneme sahiptir.

Toplu normalleştirme (Batch Normalization, BN), girdilerin dağılımını ortalaması 0 ve varyansı 1 olan standart bir normal dağılıma dönüştüren bir süreçtir. Bu dönüşüm, dağılımın aktivasyon fonksiyonunun hassas aralığında kalmasını sağlar ve böylece girdideki küçük değişikliklerin önemli tepkiler üretmesine olanak tanır. Bu durum, kayıp fonksiyonunda belirgin değişikliklere yol açabilir. Ayrıca, toplu normalizasyon, gradyan büyüklüğünü artırarak gradyan dağılımı sorununu ele almayı amaçlar. Daha büyük gradyan değerleri, derin sinir ağlarının yakınsamasını kolaylaştırır ve böylece eğitim sürecini önemli ölçüde hızlandırır (Ioffe, 2015).

4.3.1.8. Eğitim tur sayısı (Epoch)

Eğitim turlarının sayısı, tüm veri setinin algorithmadan kaç kez geçirileceğini ifade eden bir parametredir. Model eğitimi sırasında veriler tümüyle ve tek seferde değil, parçalar halinde işlenir. Bu yaklaşım, verilerin eğitim sürecine kademeli olarak entegre edilmesini sağlar ve her bir veri parçası modelin gelişimine belirli bir sırayla katkıda bulunur. Eğitim süreci şu şekilde işler: İlk olarak model, başlangıçtaki ağırlıklarla eğitilir ve performansı değerlendirilir. Ardından, geriye yayılım algoritması kullanılarak ağırlıklar performans ölçütlerine göre güncellenir. Bu adım sonrasında model, yeni eğitim verileriyle tekrar eğitilir ve ağırlıklar yeniden ayarlanır. Bu döngüsel süreç, modelin en uygun ağırlık değerlerine ulaşması için her eğitim turunda tekrarlanır. Her bir eğitim turu, literatürde "epoch" olarak adlandırılır.

Derin öğrenme bağlamında, optimum ağırlık değerleri iteratif olarak hesaplanır; bu durum, ilk epoklarda modelin performansının optimal olmamasına yol açar. Ancak, epok sayısının artmasıyla birlikte performansta belirgin bir iyileşme gözlemlenir. Bununla birlikte, belirli bir noktadan sonra modelin öğrenme kapasitesinde önemli bir azalma meydana gelebilir.

Bu tür modellerin eğitimi genellikle oldukça zaman alıcıdır ve bazı durumlarda günler, hatta aylar sürebilir. Bu durum, derin öğrenme süreçlerinde yaygın bir zorluk olarak karşımıza çıkar. Bu süreci hızlandırmak ve daha verimli hale getirmek için, diğer hiper parametrelerin dikkatli bir şekilde ayarlanmasıyla eğitim süresi minimize edilmeye çalışılır.

4.3.1.9. Hiper-Parametre

Makine öğrenimi modellerinin tasarım sürecinde, tasarımcılar modelin performansını etkileyen algoritma veya tekniklere ilişkin hiper parametreleri belirlemek

zorundadır. Modeller genellikle, yüksek performansa katkıda bulunan birden fazla hiper parametre grubuna sahiptir ve bu grupların model tasarımında kullanılmasında herhangi bir sakınca bulunmamaktadır. Ancak, en uygun hiper parametre grubunun seçimi, model geliştirme sürecindeki en önemli zorluklardan biri olarak kalmaktadır.

Hiper parametre seçimi, genellikle tasarımcının sezgilerine, benzer problemlerle ilgili önceki deneyimlere, farklı alanlardan elde edilen yöntemlerin uygulanabilirliğine, mevcut eğilimlere ve model içindeki tasarım bağımlılıklarına dayanmaktadır. Bununla birlikte, son dönemdeki ilerlemeler, en uygun hiper parametre grubunun seçimi için çeşitli tekniklerin geliştirilmesine yol açmıştır. Bu teknikler, problem çözme sürecinde daha etkin ve sistematik bir hiper parametre optimizasyonu sağlama amacı taşımaktadır.

4.3.1.10. Mini-Batch boyutu

Derin Öğrenme çalışmalarında veri setinin hepsini aynı anda işleme olarak öğrenmeye çalışmak hem kaynaklar açısından hemde zaman açısından maliyetlidir. Öğrenme sürecinin her iterasyonunda, ağ parametrelerinin gradyanını hesaplamak için geri yayılım algoritmasının kullanıldığı açıktır. Bu hesaplama geriye dönük olarak yapılır ve sonuç olarak ağırlık değerleri güncellenir. Bu sorunu çözmek için, öğrenme sürecini kolaylaştırmak amacıyla veri seti daha küçük gruplara ayrılır. Bu süreç “mini-batch” olarak adlandırılır. Model tasarlanırken, mini-batch parametresi modelin eşzamanlı olarak işleyeceği veri öğelerinin sayısını gösterir. Önceki çalışmalar, verilerin toplu işlenmesinin kullanılmasının kayıp değerinde artışa yol açabileceğini ve aynı zamanda işleme için gereken süreyi azaltabileceğini göstermiştir.

4.3.1.11. Öğrenme hızı

Derin öğrenme modellerinde parametre güncellemeleri, geriyayılım (backpropagation) algoritması aracılığıyla gerçekleştirilir. Bu algoritma, “chain rule” olarak bilinen yöntemle geriye doğru türev alınarak hata farkının hesaplanmasını içerir. Hesaplanan hata farkı, öğrenme oranı (learning rate) parametresi ile çarpılır ve elde edilen sonuç, mevcut ağırlık değerlerinden çıkarılarak yeni ağırlıklar hesaplanır. Öğrenme oranı parametresi sabit bir değer olarak belirlenebileceği gibi, eğitim süreci boyunca adım adım artan veya azalan bir değer olarak da tanımlanabilir. Ayrıca, momentum terimi ile ilişkilendirilebilir veya adaptif algoritmalar tarafından dinamik olarak öğrenilebilir. Genellikle varsayılan öğrenme oranı değeri 0,01 olup, belirli bir epoch sayısından sonra

0,001'e düşürülmektedir. Momentum için yaygın olarak kullanılan beta katsayısı 0,9'dur ve uygun aralık genellikle 0,8 ile 0,99 arasında değişmektedir.

4.3.1.12. Kayıp fonksiyonu

Makine öğreniminde, kayıp fonksiyonları modellerin ne kadar yanlış olduğunu belirlemesine ve bu yanlışlığa göre kendini iyileştirmesine yardımcı olur. Bunlar, bir makine öğrenimi modelinde tahmin edilen ve gerçek değerler arasındaki farkı nicelleştiren matematiksel fonksiyonlardır, ancak yaptıkları tek şey bu değildir.

Kayıp fonksiyonuna maliyet fonksiyonu veya hata fonksiyonu da denir. Kayıp fonksiyonu, bir modelin eğitim aşamasında ne kadar iyi performans gösterdiğini niceliksel olarak belirler. Kayıp fonksiyonu, bir algoritmanın verilen verileri ne kadar iyi modellediğini değerlendirmek için kullanılır.

4.3.2. CNN Mimarileri

Günümüzde en çok bilinen CNN mimarileri aşağıda bahsedilmiştir.

4.3.2.1. VGG-16

VGG16, Oxford Üniversitesi Visual Geometry Group (VGG) tarafından geliştirilmiş bir evrişimli sinir ağı (CNN) mimarisidir. İlk olarak Karen Simonyan ve Andrew Zisserman tarafından 2014 yılında yayımlanan "Very Deep Convolutional Networks for Large-Scale Image Recognition" başlıklı akademik makalede tanıtılmıştır (Gezer ve ark.). Evrişimli katmanlar, giriş görüntülerinden öznitelikleri çıkarmak için kullanılırken, tam bağlantılı katmanlar sınıflandırma için kullanılır. İlk blok 64 filtreye sahipken, kalan bloklar sırasıyla 128, 256, 512 ve 512 filtreye sahiptir. VGG16 mimarisinin en önemli özelliklerinden biri, evrişimli katmanlarda küçük 3x3 filtreler kullanmasıdır, bu da onun uzamsal bilgileri daha büyük filtrelerden daha etkili bir şekilde yakalamasına olanak tanır (Tablo 4.2). VGG-16, derin öğrenme modelleri arasında oldukça basit bir yapıya sahiptir. Model, aşağıdaki gibi düzenlenmiştir:

- **Giriş Katmanı:** Giriş olarak 224x224 boyutunda bir RGB görüntü alır. Bu boyuta uygun giriş verisi sağlamak için görüntüler önceden yeniden boyutlandırılır.
- **Konvolüsyonel Bloklar:** Model, 13 konvolüsyon katmanını ardışık olarak içerir. Her konvolüsyon katmanı, **3x3 filtreler** kullanır ve padding ile boyutların korunmasını sağlar (örneğin, "same padding"). ReLU aktivasyon fonksiyonu kullanılarak doğrusal olmayanlık eklenir.

- **Havuzlama Katmanları:** Her birkaç konvolüsyon katmanından sonra bir **Max Pooling** katmanı yerleştirilir. Bu katmanlar, boyutu 2x2 filtrelerle küçültür.
- **Tam Bağlantılı Katmanlar (Fully Connected):** En son katmanlarda, 4096, 4096 ve 1000 nöronlu üç tam bağlantılı katman bulunur. 1000 nöronlu son katman, ImageNet veri setindeki 1000 sınıfa ait olasılıkları tahmin eder.
- **Softmax Çıkışı:** Çıktı, sınıfların olasılık dağılımını üretmek için softmax fonksiyonundan geçirilir.

Tablo 4.2 Vgg16 Katman yapısı

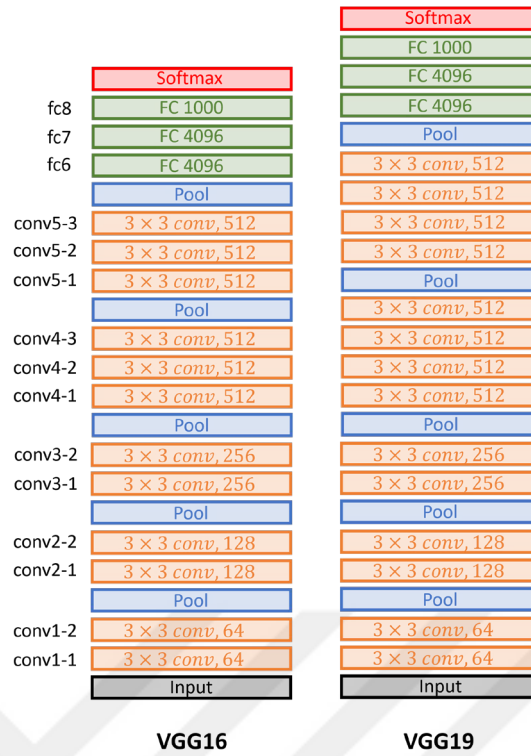
Katman Türü	Filtre Sayısı	Filtre Boyutu	Çıkış Boyutu	Aktivasyon
Giriş (Input)	-	-	224x224x3	-
Conv1-1	64	3x3	224x224x64	ReLU
Conv1-2	64	3x3	224x224x64	ReLU
MaxPooling	-	2x2	112x112x64	-
Conv2-1	128	3x3	112x112x128	ReLU
Conv2-2	128	3x3	112x112x128	ReLU
MaxPooling	-	2x2	56x56x128	-
Conv3-1, Conv3-2...	256	3x3	28x28x256	ReLU
...				
Son Katman	-	-	1000 sınıf	Softmax

4.3.2.2. VGG-19

VGG-19, bilgisayarla görü (computer vision) alanında yaygın olarak kullanılan, derin bir Konvolüsyonel Sinir Ağı (CNN) mimarisidir. 2014 yılında Oxford Üniversitesi'ndeki Visual Geometry Group (VGG) tarafından ImageNet Large Scale Visual Recognition Challenge (ILSVRC) için geliştirilmiştir. 19 ağırlıklı katmandan oluştuğu için "VGG-19" adını almıştır.

VGG-19, görüntü sınıflandırma, nesne algılama ve diğer görsel görevlerde güçlü performans göstermiştir ve derin öğrenme topluluğunda popüler bir model olarak kabul edilir.

VGG-19, VGG-16'ya göre üç ek konvolüsyon katmanı içerir (Şekil 4.8). Bu, modelin daha derin ve özellik öğreniminde daha güçlü olmasını sağlar. Modelde toplam 16 konvolüsyon katmanı bulunur.



Şekil 4.8 Vgg16 ve Vgg19 katman yapıları

4.3.2.3. ResNet 50

ResNet-50, derin öğrenme dünyasında yaygın olarak kullanılan, Residual Network (Artık Ağ) mimarisi tabanlı bir derin sinir ağıdır. Microsoft Research tarafından 2015 yılında geliştirilmiştir ve ImageNet Large Scale Visual Recognition Challenge (ILSVRC) karşılaşmasında çok başarılı olmuştur. ResNet-50'nin adı, toplamda 50 katmanlı bir model olmasından gelir.

ResNet mimarisi, özellikle derin ağların eğitimi sırasında karşılaşılan "kaybolan gradyan" problemini çözmek için geliştirilmiştir. Bunun için ağ, residual (artık) bağlantılar kullanır. Derin ağlarda, ağın derinleştikçe öğrenme performansı bir noktadan sonra düşebilir. Bu sorun, modelin katmanlar arasında gradyanı doğru bir şekilde yayamamasından kaynaklanır. ResNet, bu sorunu çözmek için skip connections (atlama bağlantıları) kullanır (Tablo 4.3).

Gradyan dağılımının iyi bir şekilde çözüldüğü söylenebilir. Ağ derinleşmesiyle ilgili bir diğer sorun ise bozulmadır, yani derinlik arttıkça ağın performansı kötüleşir. Deneyimlerden yola çıkarak, ağın derinliği modelin performansı için çok önemlidir. Ağ katmanlarının sayısı artırıldığında, ağ daha karmaşık özellik deseni çıkarma işlemi gerçekleştirebilir, bu nedenle model daha derin olduğunda teorik olarak daha iyi sonuçlar elde edilebilir. Ancak deney, derin ağın dejenerasyona uğradığını tespit etti.

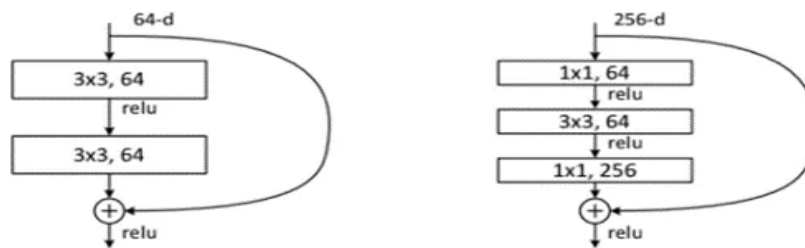
Ağ derinliğinin artmasıyla, ağın doğruluğu doyunluğa ulaşmaya veya hatta azalmaya meyillidir. Eğitim setinin doğruluğunda bir azalma vardır. Bunun aşırı uyumdan kaynaklanmadığı belirlenebilir. Çünkü aşırı uyum durumunda eğitim setinin doğruluğu yüksek olmalıdır. ResNet'teki artık ağ bu sorunu çözmek için tasarlanmıştır ve bu sorunu çözdükten sonra, ağın derinliği birkaç büyüklük sırası kadar artar (Li ve Lima, 2021).

Tablo 4.3 Vgg19 Katman yapısı

Katman Türü	Filtre Sayısı	Filtre Boyutu	Çıkış Boyutu
Giriş (Input)	-	-	224x224x3
İlk Konvolüsyon	64	7x7	112x112x64
İlk MaxPooling	-	3x3	56x56x64
Residual Blok-1	64, 256	1x1, 3x3, 1x1	56x56x256
Residual Blok-2	128, 512	1x1, 3x3, 1x1	28x28x512
Residual Blok-3	256, 1024	1x1, 3x3, 1x1	14x14x1024
Residual Blok-4	512, 2048	1x1, 3x3, 1x1	7x7x2048
Global Average Pooling	-	-	1x1x2048
Tam Bağlantılı (FC)	-	-	1000 sınıf

Model, toplamda 50 katmandan oluşur. Modelin daha derin olmasını sağlar ve gradyanların katmanlar arasında düzgün bir şekilde yayılmasına yardımcı olur. Yaklaşık 25 milyon parametre içerir, bu da daha düşük hesaplama gücü gerektirir (VGG-19 ile kıyaslandığında). Derin ağların daha kolay eğitilmesine olanak tanır. ResNet-50, görüntü sınıflandırma gibi görevlerde yüksek doğruluk sağlar. Önceden eğitilmiş ağırlıkları, diğer görüntü işleme görevlerinde kolayca kullanılabilir.

ResNet-50, 34 katmanlı ağdaki her iki katmanlı bloğun, üç katmanlı darboğaz bloklarıyla değiştirilmesiyle oluşturulmuştur (Şekil 4.9). Her bir artık fonksiyon F için, iki katman yerine üç katmandan oluşan bir yapı kullanılır. Bu üç katman, sırasıyla 1×1 , 3×3 ve 1×1 boyutlarına sahiptir. 1×1 boyutundaki katmanlar, boyutları küçültme ve ardından geri yükleme işlemlerinden sorumlu olup, 3×3 boyutundaki katmanlar daha küçük girdi/çıkış boyutlarına sahip bir darboğaz işlevi görür.



Şekil 4.9 ResNet-34 ve Resnet-50 yapı taşı

Daha fazla üç katmanlı blok eklenerek 101 katmanlı ve 152 katmanlı ResNet mimarileri oluşturulmuştur. 50, 101 ve 152 katmanlı ResNet modelleri, 34 katmanlı modele kıyasla anlamlı ölçüde daha yüksek doğruluk sunmaktadır. Artık ağların kullanımı, derin ağlarda sıklıkla karşılaşılan bozulma problemlerini önlemiş ve ağ derinliğinin artmasıyla elde edilen doğruluk kazanımlarını önemli ölçüde artırmıştır.

4.3.2.4. InceptionV3

InceptionV3 "Rethinking the Inception Architecture for Computer Vision" adlı makalede önerilen bir modüldür (Szegedy ve ark., 2016). Inception v2 ile karşılaştırıldığında, temel fark yalnızca evrişim katmanları değil, aynı zamanda yardımcı sınıflayıcı olarak küme normalizasyonu (batch normalization) ve tam bağlı (fully connected, FC) katmanların eklenmiş olmasıdır. Inception v2 bu şekilde düzenlendiğinde, bu yapı Inception v3 olarak adlandırılmaktadır. İki model arasındaki karşılaştırma aşağıda sunulmaktadır.

GoogleNet (Inception-v1) ile karşılaştırıldığında, Inception-v3 modeli nesne tanımada üstün performansa sahiptir. Özellikle, Inception-v3 modeli üç ana bileşenden oluşmaktadır: temel evrişimsel blok, sınıflandırıcı ve geliştirilmiş inception modülü. Evrişimsel ile maksimum havuzlama katmanlarını dönüşümlü olarak kullanan temel evrişimsel blok, özellik çıkarma için kullanılır. Geliştirilmiş Inception modülü Network-In-Network temel alınarak tasarlanmıştır (Lin ve ark., 2019).

Inception-v3, derin öğrenme tabanlı bir Konvolüsyonel Sinir Ağı (CNN) mimarisidir ve özellikle bilgisayarla görü (computer vision) görevlerinde yaygın olarak kullanılır. Google Research tarafından 2015 yılında geliştirilmiş ve ImageNet Large Scale Visual Recognition Challenge (ILSVRC) yarışmasında yüksek doğruluk oranlarıyla öne çıkmıştır. Bu model, Inception mimarisinin üçüncü sürümüdür ve öncekilere göre daha optimize edilmiş ve performansı artırılmıştır.

Inception-v3, derin ve geniş ağlar oluşturma sorunlarını çözmek için geliştirilmiştir. Geleneksel derin ağların aksine, Inception mimarisi hesaplama maliyetlerini azaltırken doğruluğu artırmayı hedefler. Inception-v3, temel olarak Inception modülleri üzerine kurulmuştur. Bu modüller, birden fazla konvolüsyon ve havuzlama işlemini aynı anda gerçekleştirerek farklı ölçekte özelliklerin öğrenilmesine olanak tanır. Inception modülleri, giriş verisini aynı anda: 1x1, 3x3 ve 5x5 filtrelerle işler. Max Pooling ve Average Pooling uygular. Bu sayede, farklı seviyelerdeki özellikler

paralel olarak çıkarılır ve hesaplama maliyeti düşürülür. Inception-v3, büyük boyutlu konvolüsyonları daha küçük boyutlu konvolüsyonlara böler (örneğin, bir 5x5 konvolüsyonu iki ardışık 3x3 konvolüsyona ayırır). Bu hem hesaplama verimliliğini artırır hem de parametre sayısını azaltır. 1x1 konvolüsyonlar, giriş verisinin boyutunu azaltmak için kullanılır (Tablo 4.4). Bu, hesaplama yükünü önemli ölçüde düşürürken bilgiyi kaybetmemeyi sağlar. Eğitim sırasında, modelin erken katmanlarına yerleştirilen bir yardımcı sınıflandırıcı kullanılır. Bu, gradyanların derin katmanlara daha kolay ulaşmasını sağlar ve aşırı uyuma (overfitting) karşı direnci artırır.

Tablo 4.4 InceptionV3 Katman yapısı

Katman Türü	Filtre Sayısı	Filtre Boyutu	Çıkış Boyutu
Giriş	-	-	299x299x3
İlk Konvolüsyon	32	3x3, stride 2	149x149x32
Max Pooling	-	3x3, stride 2	73x73x32
Inception Modül-1	-	-	35x35x256
Inception Modül-2	-	-	17x17x768
Inception Modül-3	-	-	8x8x2048
Global Average Pooling	-	-	1x1x2048
Tam Bağlantılı Katman	-	-	1000 sınıf

Inception-v3, hesaplama verimliliği ve doğruluk arasındaki dengeyi sağlayan bir modeldir. Görüntü işleme görevlerinde güçlü bir performans sergiler ve transfer öğrenme için geniş bir kullanım alanına sahiptir. Modern bilgisayarla görü uygulamalarında hala popüler bir mimari olarak kullanılmaktadır.

4.4. Değerlendirme Metrikleri

Sınıflandırma sürecinde farklı metrikler kullanılmaktadır ve bu metrikler temelde karmaşıklık matrisinden elde edilen değerler kullanılarak bulunmaktadır. Karmaşıklık Matrisi, makine öğrenimi sınıflandırma problemlerinde kullanılan bir performans değerlendirme aracıdır. Gerçek ve öngörülen değerlerin dört farklı kombinasyonunu içeren bir tablo sunar. (Şekil 4.10). Özellikle dengesiz veri setlerinde modelin performansını daha detaylı anlamak için kritik bir araçtır. Bu sayede, modelin hangi sınıflarda daha başarılı veya başarısız olduğu tespit edilerek, gerekli iyileştirmeler yapılabilir. Karmaşıklık matrisinde dikkate alınması gereken bazı temel terimler

bulunmaktadır. Bu terimler, ilerleyen aşamalarda çeşitli metriklerin hesaplanmasında kullanılabilir ve şu şekilde özetlenebilir.

- **True Positive (TP):** Modelin olumlu olarak doğru sınıflandırdığı örnekler.
- **True Negative (Romulus ve ark.):** Modelin olumsuz olarak doğru sınıflandırdığı örnekler.
- **False Positive (FP):** Modelin yanlışlıkla olumlu olarak sınıflandırdığı olumsuz örnekler.
- **False Negative (FN):** Modelin yanlışlıkla olumsuz olarak sınıflandırdığı olumlu örnekler.

		Gerçek Değerler	
		Pozitif (1)	Negatif (0)
Tahmin Değerleri	Pozitif (1)	True Positive	False Positive
	Negatif (0)	False Negative	True Negative

Şekil 4.10 Karmaşıklık Matrisi

4.4.1. Doğruluk (Accuracy)

Toplam doğru tahminlerin tüm tahminlere oranı. Denklem 4.1' de gösterilmiştir. Veri dengeli olduğunda iyi bir metriktir. Ancak dengesiz veri setlerinde yanıltıcı olabilir.

$$\text{Doğruluk} = \frac{TP+TN}{TP+FN+FP+TN} \quad (4.1)$$

4.4.2. Kesinlik (Precision)

Pozitif olarak tahmin edilen örneklerin, ne kadarının gerçek pozitif olduğunu ölçer. Denklem 4.2' de gösterilmiştir. Yanlış pozitiflerin maliyeti yüksek olduğunda önemlidir.

$$Kesinlik = \frac{TP}{TP+FP} \quad (4.2)$$

4.4.3. Duyarlılık (Recall) / tespit oranı

Gerçek pozitiflerin ne kadarının doğru tahmin edildiğini ölçer. Denklem 4.3' de gösterilmiştir. Yanlış negatiflerin maliyeti yüksek olduğunda önemlidir.

$$Duyarlılık = \frac{TP}{TP+FN} \quad (4.3)$$

4.4.4. F1 – Skoru

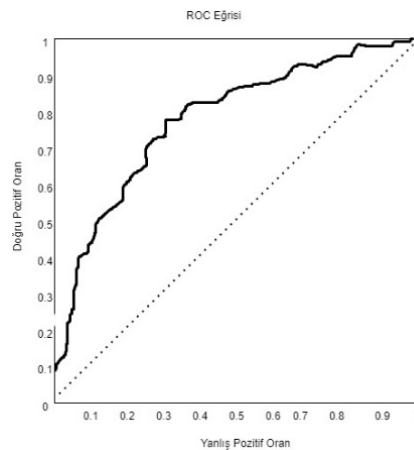
F1 skoru, bir testin doğruluğunu değerlendiren bir metriktir ve kesinlik ile duyarlılığın harmonik ortalamasını temsil eder. Denklem 4.4'te gösterildiği gibi, maksimum değeri 1 (mükemmel kesinlik ve duyarlılık) ve minimum değeri 0 olabilir. Bu metrik, modelin genel kesinliğini ve performans dengesini ölçmek için kullanılır

$$F1 - Skor = 2 * \frac{Kesinlik * Duyarlilik}{Kesinlik + Duyarlilik} \quad (4.4)$$

4.4.5. ROC-AUC eğrisi

ROC-AUC eğrisi, farklı eşik değerlerinde sınıflandırma problemleri için kullanılan bir performans ölçütüdür. ROC, bir olasılık eğrisi olup, AUC ise sınıflar arasındaki ayrılabilirlik derecesini temsil eder. AUC, ROC eğrisinin altında kalan alanı ifade eder ve bir modelin sınıflar arasındaki ayrımı ne kadar iyi yapabildiğini gösterir (Şekil 4.11). AUC değeri arttıkça, modelin tahmin performansı da o oranda artar. Örneğin, yüksek bir AUC değeri, modelin hastalığı olan ve olmayan bireyleri daha iyi ayırt edebildiğini ifade eder. Denklem 4.5'te gösterilmiştir.

$$AUC = \frac{\sum_{i=1}^{n^+} \sum_{j=1}^{n^-} 1_{f(x_i^+) > f(x_j^-)}}{mn} \quad (4.5)$$



Şekil 4.11 ROC eğrisi örneği

4.4.6. Gerçek pozitif oran (Hassasiyet)

Gerçek Pozitif Oran (TPR), tüm pozitif veri noktaları içinde doğru bir şekilde pozitif olarak sınıflandırılan veri noktalarının oranını ifade eder. Denklem 4.6'da gösterilmiştir.

$$\text{Gerçek Pozitif Oran} = \frac{TP}{FN+TP} \quad (4.6)$$

4.4.7. Yanlış pozitif oran (Özgüllük)

Yanlış Pozitif Oran (FPR), tüm negatif veri noktaları içinde yanlış bir şekilde pozitif olarak sınıflandırılan veri noktalarının oranını ifade eder. Denklem 4.7'de gösterilmiştir.

$$\text{Yanlış Pozitif Oran} = \frac{FP}{FP+TN} \quad (4.7)$$

Sonuç olarak, pozitifleri doğru bir şekilde tanımlamak öncelikliyse daha yüksek duyarlılığa (sensitivity) sahip bir model tercih edilmelidir. Ancak, negatifleri doğru bir şekilde tanımlamak daha önemli olduğunda, ölçüm metriği olarak özgüllük (specificity) kullanılmalıdır.

5. DENEYSEL ÇALIŞMALAR

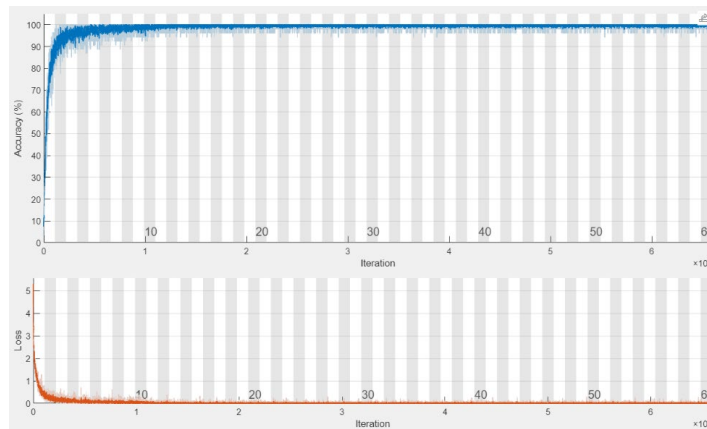
Bu tez çalışmasında, önceki bölümlerde detaylandırıldığı üzere, öncelikle veri kümesi elde edilmiş (mevcut sınıflar içerisindeki tamamen hatalı örnekler silinerek normalde 74700 olması gereken toplam örnek sayısı 74226'ya düşürülmüştür) ve ardından derin öğrenme yaklaşımları kullanılmıştır.

Bu tez çalışmasında, öncelikle el yazısı rakamlarının sınıflandırılması yapılmıştır. Bu çalışmayı yapmanın motivasyonu, derin öğrenme algoritmalarının mevcut verisetinde bulunan el yazısı rakamları ile sınıflandırma yapıp yapamayacağını ve hangi düzeyde sınıflandırma başarısı elde edileceğinin görülmek istenmesi olmuştur. Daha sonrasında el yazısı rakamlarına ait veriler için cinsiyete göre yeniden düzenleme yapılarak en iyi bilenen derin öğrenme algoritmaları ile deneysel çalışmalar yapılmıştır.

5.1. El Yazısı Rakamlarının Tespiti

Bu bölümde el yazısı rakamlarının tespiti ve sınıflandırması için bir CNN modeli kullanılmıştır. CNN modeli için eğitim ve test oranları sırasıyla %80 ve %20 olarak belirlenmiştir. CNN modeli 60 epok boyunca çalıştırılmıştır. CNN modeli doğruluk ve kayıp grafiği Şekil 5.1'de verilmiştir.

Yapılan deneysel çalışmalarda el yazısı rakamlarını tespit edebilmek için CNN modelinin son katmanında, oluşturulan veri setinde kullanılan rakamları, her birinin bir sınıf puanına karşılık geldiği 10 boyutlu bir vektörle sonuçlanan sınıf puanları hesaplanacaktır. Rakamları sınıflandırmak için kullanılacak eğitilmiş modeli hazırlayacaktır. Böylece görüntülerde bulunan rakamları (0,1,2,3,4,5,6,7,8,9) sınıflandırabilir.



Şekil 5.1: Eğitim ve Kayıp grafikleri

CNN ağı, eğitim sürecini yüksek bir başarı oranıyla tamamladıktan sonra, Şekil 5.2’de bulunan, karmaşıklık matrisi elde edilir, bu matris ağın test verilerindeki sınıflandırma başarısını göstermektedir. Karmaşıklık matrisi incelendiğinde, neredeyse tüm sınıflar için test başarı oranının %99 olduğu görülmektedir. Bu sonuçlar ağın başarılı bir şekilde eğitildiğini göstermektedir.

0	1484						3			1	99.7%	0.3%
1		1481		1			1	5			99.5%	0.5%
2	2	5	1462	2				3			99.2%	0.8%
3		1	1	1468		6		1	3	2	99.1%	0.9%
4		5	1		1477		1	1			99.5%	0.5%
5				2		1477		2		2	99.6%	0.4%
6	2	1			3	1	1478	1			99.5%	0.5%
7		9		3	3	1		1465		1	98.9%	1.1%
8	1		1	1		1	4	1	1475	5	99.1%	0.9%
9	3			4	1	5		3	2	1469	98.8%	1.2%

99.5%	98.6%	99.8%	99.1%	99.5%	99.1%	99.4%	98.9%	99.7%	99.3%
0.5%	1.4%	0.2%	0.9%	0.5%	0.9%	0.6%	1.1%	0.3%	0.7%
0	1	2	3	4	5	6	7	8	9

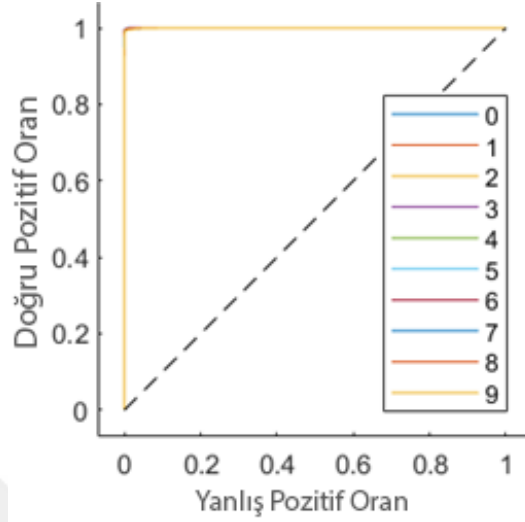
Şekil 5.2: Karmaşıklık Matrisi

Test süreci sonucunda temel değerlendirme metriklerinden doğruluk, hassasiyet, özgüllük, kesinlik, duyarlılık ve F1-skoru için sonuçlar Tablo 5.1’de gösterilmiştir. Verilen değerler tek bir sınıf için değil tüm sınıflar için ortalama değerleri göstermektedir. Test sonucunda tüm metriklerde %99’un üzerinde sonuçlar elde edildiği görülmektedir.

Tablo 5.1 Elde edilen ortalama değerler

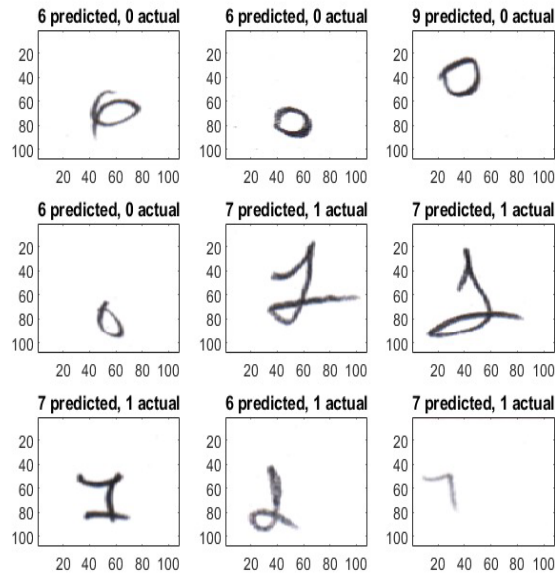
Doğruluk	Hassasiyet	Özgüllük	Kesinlik	Duyarlılık	F1-Skoru
0.99854	0.99272	0.99919	0.99274	0.99272	0.99273

Alıcı İşletim Karakteristiği (The Receiver Operating Characteristic- ROC) eğrisi, bir sınıflandırma sonucunu değerlendirmek için kullanılan, duyarlılık ve özgüllük arasındaki ilişkiyi gösteren önemli bir grafikdir. ROC eğrisi (Şekil 5.3) analiz edildiğinde, ideal olarak tüm sınıflarda istenen sonuca ulaşıldığı görülür.



Şekil 5.3: ROC Eğrisi

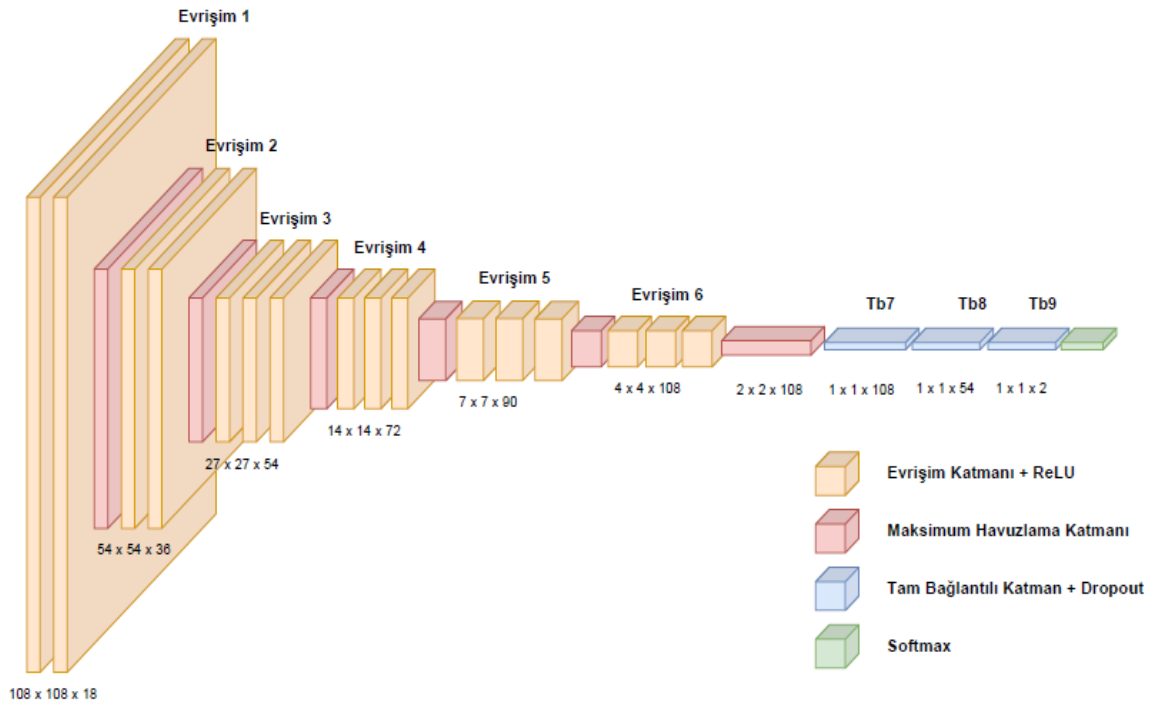
Şekil 5.2'de gösterilen karışıklık matrisinden de görülebileceği gibi, eğitilen CNN ağının test sırasında örnekleri yanlış tahmin etme oranı oldukça düşüktür. Şekil 5.4, insan gözüyle doğru tahmin edilmesi zor olan bazı yanlış tahmin örnekleri göstermektedir.



Şekil 5.4: Hatalı tahmin edilen bazı örnekler.

Bu çalışmada önerilen CNN modeli deneysel çalışmalarda %99 doğruluk elde etmiştir. Umut vaat eden bu çalışma sonrası bu algoritma ile son tam bağı katman değiştirilerek iki vektörlü sınıflama ile cinsiyet tespit edebilmek için çalışmalar yapılmıştır.

Özelleştirilmiş CNN ağ modeli aşağıdaki şekilde detaylandırılmıştır (Şekil 5.5).



Şekil 5.5 Çalışmada kullanılan Özel CNN Modeli

5.2. Özelleştirilmiş CNN modelin açıklaması

Basit bir evrişimli sinir ağı, bir dizi katmandan oluşur ve her katman, türevlenebilir bir fonksiyon aracılığıyla bir aktivasyon hacmini diğerine dönüştürür. Ağı oluşturmak için üç ana katman türü kullanılır. Bunlar evrişim katmanları, havuzlama katmanları ve tam bağı katmanlardır. Bu katmanlar istifleyerek ağ mimarisi oluşturulur.

Model genel olarak 9 aşamadan oluşmaktadır. İlk altı aşama evrişim katmanları ve havuzlama katmanları, son üç aşamada ise tam bağı katmanlar vardır.

Şekil 5.5'de, bu çalışmada önerilen CNN modelinin mimarisi gösterilmektedir. İlk olarak görüntüler üzerinde görüntülerin yeniden boyutlandırılması, piksel değerlerinin normalleştirilmesi vb. gibi ön işlemlere ihtiyaç vardır. Gerekli ön işlemlerden sonra veriler CNN modelini beslemeye hazır hale gelir.

CNN mimarisinin ilk evrişimli katmanı conv1'dir. ReLu aktivasyon fonksiyonuna sahip evrişimli bir katmandan oluşur. Bu katman, önceden işlenmiş görüntüyü $n*n=108*108$ boyutunda girdi olarak alır. Evrişim filtresi boyutu ($f*f$) $3*3$ 'tür; dolgu (padding – p) 0, adım (stride-s) 2 ve filtre sayısı 18'dir. Bu evrişim işleminden sonra, $18*108*108$ boyutunda özellik haritaları elde eder, daha sonra batchNormalizationLayer katmanı ile toplu normalleştirme yapılır. Böylece eş zamanlı öğrenme sağlanır. Eğitim süresi azalarak modelin performansı artar.

Sonra her özellik haritasında ReLU aktivasyonu yapılır. Ardından maksimum havuzlama katmanı uygulanır. Havuzlama boyutu $2*2$ 'dir. Dolgu 0 ve adım sayısı 2'dir. Bu işlem sonucunda $54*54*36$ boyutunda özellik haritaları elde edilir.

Şekil 6 da gösterildiği gibi art arda evrişim işlemleri, Conv1 işlemine benzer şekilde beş defa daha tekrarlanır. Conv6, işleminden sonra maksimum havuzlama katmanı uygulanarak $1*1*108$ boyutunda özellik haritası elde edilir.

Fc7 işlemi, Tam bağlı katman (Fully-Connected Layer) işlemi hemen başlar. Bu katman son ve en önemli katmandır. Verileri düzleştirme (Flattening) işlemine alır ve tek boyutlu $1*1*108$ bir vektör üretir. Ardından dropoutLayer katmanı kullanılır. Böylece, modelde aşırı öğrenmeyi (Overfitting) önleme sağlanır.

Hemen ardından fc8 işlemi tekrar kullanılarak $1*1*54$ boyutunda bir vektör elde edilir. Bu vektörde de dropoutLayer kullanılarak aşırı öğrenme engellenir. Son olarak fc9 işlemi ile tekrar tam bağlı katman kullanılarak $1*1*2$ boyutlarında vektör elde edilir. Bu katman, oluşturulan veri setindeki kullanılan rakamları, her birinin bir sınıf puanına karşılık geldiği 2 boyutlu bir vektörle sonuçlanan sınıf puanlarını hesaplayacaktır. Nihai çıkışlar için softmax aktivasyonu fonksiyonu kullanılır.

Bu şekilde CNN, orijinal görüntüyü katmanlar halinde orijinal piksel değerlerinden son sınıf puanlarına dönüştürür. Bazı katmanlar parametreler içerir, bazıları ise içermemektedir. Özellikle, evrişim / tamamen bağlı katmanlar, yalnızca giriş hacmindeki aktivasyonların değil aynı zamanda parametrelerin (nöronların ağırlıkları ve eğilimleri) bir fonksiyonu olan dönüşümleri gerçekleştirir.

Öte yandan, ReLu / havuzlama katmanları sabit bir işlevi uygulayacaktır. Algoritma, test verilerinde bulunan rakamları sınıflandırmak için kullanılacak eğitilmiş modeli hazırlar. Böylece görüntülerde bulunan rakamları kimlerin yazdığını tahmin ederek, Erkek veya Kız olarak sınıflandırılabilir.

5.3. El Yazısı Rakamları ile Cinsiyet Tespiti

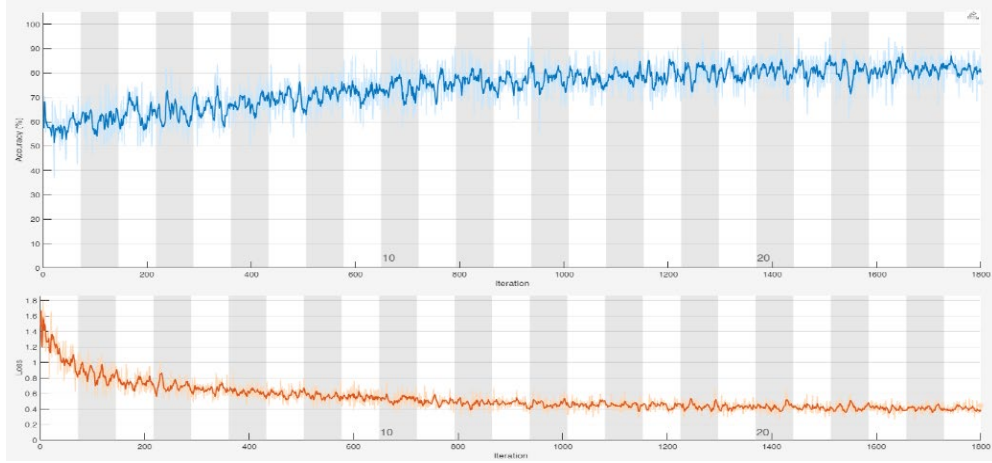
El yazısı rakamlarının Özelleştirilmiş CNN ile yüksek başarı oranları elde etmesi sonucunda tez çalışmasının esas amacı olan el yazısı rakamlarından cinsiyet tespiti işlemine geçilmiştir. El yazısı rakamlarından cinsiyet tespiti için veri kümesi elde düzenlenmiş ve ardından CNN ağ modelleri sırasıyla aynı bilgisayar sistemi üzerinde kullanılmıştır. CNN modelleri için eğitim ve test oranları sırasıyla %80 ve %20 olarak belirlenmiştir. Modellerin katman sayıları ve parametre sayısı Tablo 5.2’de gösterilmiştir. CNN modeli 25 epok boyunca çalıştırılmıştır. CNN modellerinin doğruluk ve kayıp grafikleri Bölüm 5.2.1. de verilmiştir.

Tablo 5.2 Derin öğrenme modelleri ve Katman Sayıları

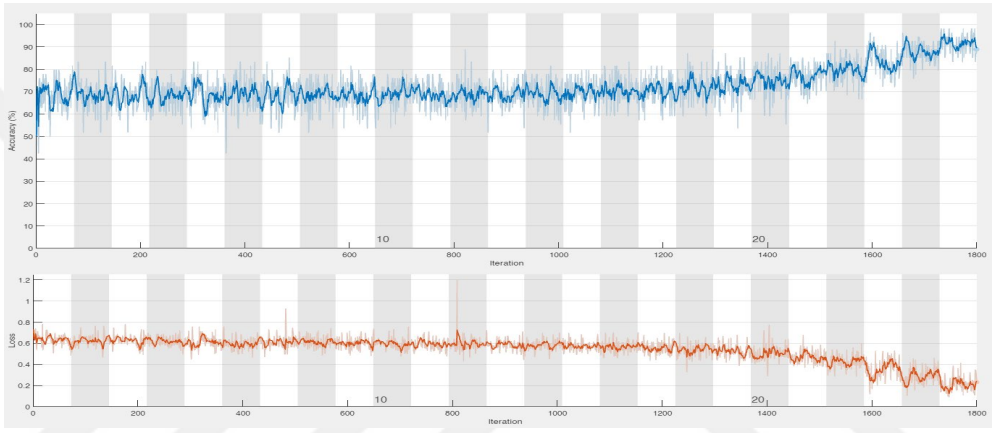
Derin Öğrenme Modeli	Katman Sayıları	Parametre Sayısı (milyon)
Özelleştirilmiş CNN	9	0,72
VGG-16	16	14.7
VGG-19	19	20.01
ResNet-50	50	23.75
Inception V3	48	21.8

5.3.1. Test doğruluk ve zamana bağlı kayıp grafiği

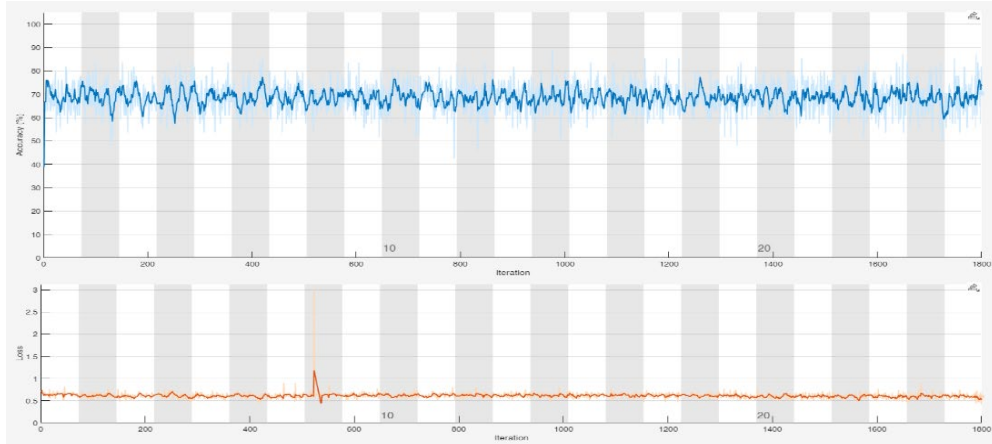
Test doğruluğu, bir modelin daha önce eğitilmediği bir test veri kümesi üzerindeki tahminlerinin kesinliği olarak tanımlanır. Modelin genelleme kabiliyetini değerlendirmek için kullanılır ve model karşılaştırması için çok önemli bir metriktir. Farklı derin öğrenme mimarilerinin test doğruluğu üzerindeki etkisi model derinliği, parametre sayısı, öğrenme kapasitesi ve veri kümesi boyutu gibi faktörlere bağlıdır. Detaylı sonuçlar Şekil 5.6-5.10'da gösterilmektedir.



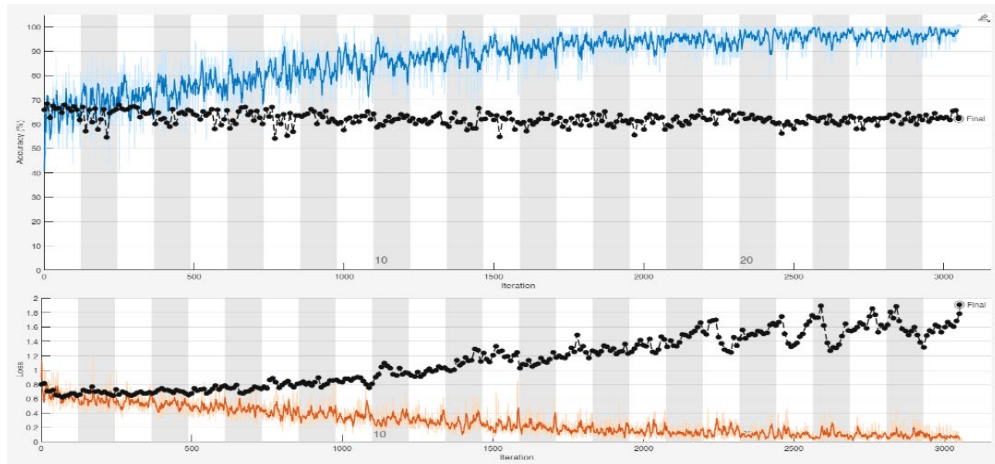
Şekil 5.6: Özelleştirilmiş CNN Test Doğruluğu ve Zamana Bağlı Kayıp Grafiği



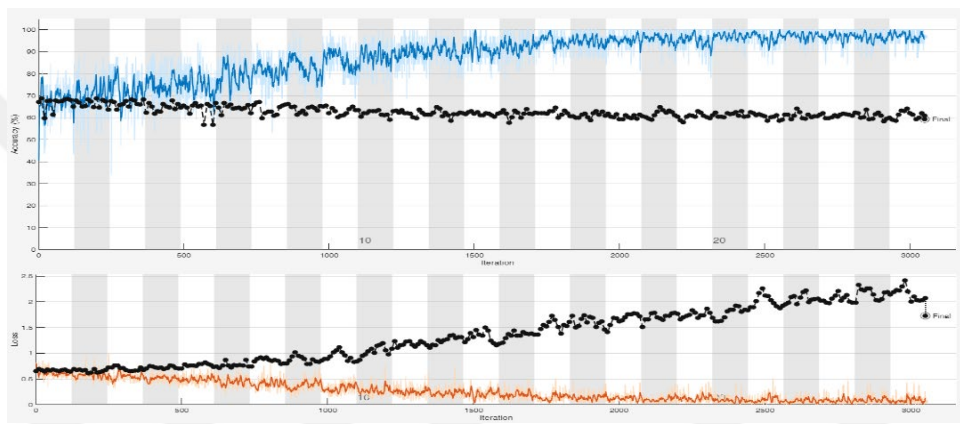
Şekil 5.7: VGG16 Test Doğruluğu ve Zamana Bağlı Kayıp Grafiği



Şekil 5.8: VGG19 Test Doğruluğu ve Zamana Bağlı Kayıp Grafiği



Şekil 5.9: ResNet-50 Test Doğruluğu ve Zamana Bağlı Kayıp Grafiği



Şekil 5.10: InceptionV3 Test Doğruluğu ve Zamana Bağlı Kayıp Grafiği

Yukarıda verilen test doğruluğu (Accuracy) ve zamana bağlı kayıp (Loss) grafikleri incelendiğinde, Özelleştirilmiş CNN modelinin diğer modellere (VGG16, VGG19, ResNet-50 ve InceptionV3) kıyasla daha iyi bir eğitim süreci sürdürdüğü söylenebilir ve bu modelin test doğruluğu grafiği, diğer modellere kıyasla daha yüksek bir doğruluk oranına ulaştığını göstermektedir. Ayrıca Özelleştirilmiş CNN, hem daha yüksek bir test doğruluğu elde etmiş hem de kayıp değerlerini daha hızlı ve tutarlı bir şekilde azaltmıştır. Özellikle VGG16 ve VGG19 gibi daha derin modeller, özelleştirilmiş CNN'den daha düşük bir doğruluk oranı sergilemiş olduğu görülmüştür.

5.3.2. Karmaşıklık matrisleri

CNN modelleri eğitim süreçleri başarı ile tamamladıktan sonra, Şekil 5.11-Şekil 5.15 de bulunan karmaşıklık matrisleri, test doğruluğu ve zamana bağlı kayıp grafikleri alınmıştır. Matrisler incelendiğinde, diğerlerinden daha başarılı olan Özelleştirilmiş CNN modelinin %65,2 olduğu görülmektedir.

		Karmaşıklık Matrisi		
Çıkış Sınıfı	E	Hedef Sınıf		
		↺	↕	
	E	565 87.5%	233 23.7%	70.8% 29.2%
	K	109 11.1%	76 7.7%	41.1% 58.9%
		83.8% 16.2%	24.6% 75.4%	65.2% 34.8%

Şekil 5.11 Özelleştirilmiş CNN matrisi

		Karmaşıklık Matrisi		
Çıkış Sınıfı	E	Hedef Sınıf		
		↺	↕	
	E	411 41.8%	165 16.8%	71.4% 28.6%
	K	263 26.8%	144 14.6%	35.4% 64.6%
		61.0% 39.0%	46.6% 53.4%	56.5% 43.5%

Şekil 5.12 VGG-16'nın matrisi

		Karmaşıklık Matrisi		
Çıkış Sınıfı	E	Hedef Sınıf		
		↺	↕	
	E	500 59.9%	195 19.8%	71.9% 28.1%
	K	174 17.7%	114 11.6%	60.4% 39.6%
		74.3% 25.8%	63.1% 36.9%	62.4% 37.6%

Şekil 5.13 VGG-19'un matrisi

		Karmaşıklık Matrisi		
Çıkış Sınıfı	E	Hedef Sınıf		
		↺	↕	
	E	506 51.6%	203 20.7%	71.4% 28.6%
	K	168 17.1%	106 10.8%	38.7% 61.3%
		75.1% 24.9%	34.3% 65.7%	62.3% 37.7%

Şekil 5.14 Resnet-50'nin matrisi

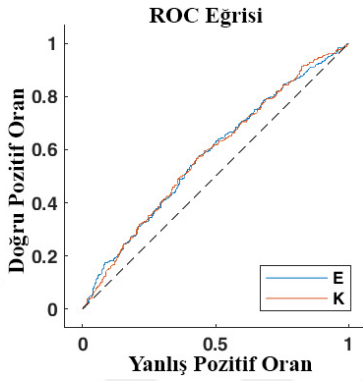
		Karmaşıklık Matrisi		
Çıkış Sınıfı	E	Hedef Sınıf		
		↺	↕	
	E	479 48.7%	206 21.0%	69.9% 30.1%
	K	195 19.8%	103 10.5%	34.6% 65.4%
		71.1% 28.9%	33.3% 66.7%	59.2% 40.8%

Şekil 5.15 InceptionV3'ün matrisi

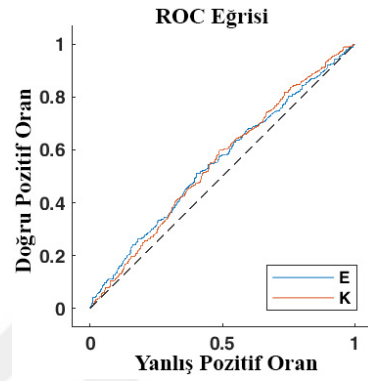
Karmaşıklık matrisleri incelendiğinde, Özelleştirilmiş CNN modelinin doğruluk oranı %65,2 olarak bulunduğu görülmektedir. Genel olarak incelendiğinde ise Özelleştirilmiş CNN modelinin diğer modellere göre daha iyi sonuçlar elde ettiği söylenebilir.

5.3.3. AUC-ROC grafikleri

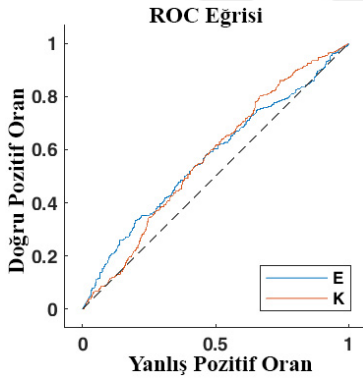
Alıcı İşletim Karakteristiği (The Receiver Operating Characteristic- ROC) eğrisi, bir sınıflandırma sonucunu değerlendirmek için kullanılan, duyarlılık ve özgüllük arasındaki ilişkiyi gösteren önemli bir grafikdir. ROC eğrisi analiz edildiğinde, Özelleştirilmiş CNN verisi daha başarılıdır (Şekil 5.16- Şekil 5.20).



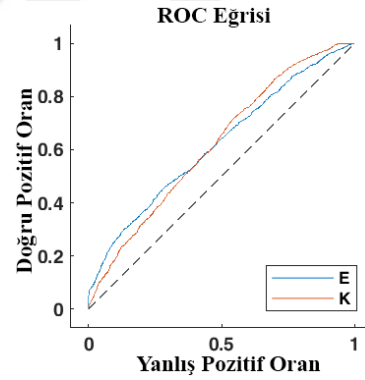
Şekil 5.16. Özelleştirilmiş CNN'nin grafiği



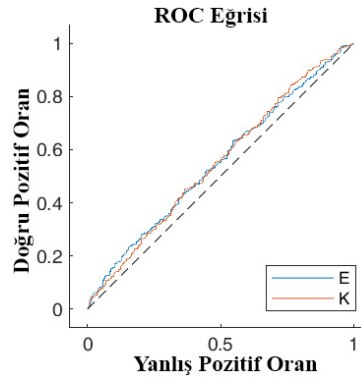
Şekil 5.17 VGG-16'nın grafiği



Şekil 5.18 VGG-19'un grafiği



Şekil 5.19 Resnet-50'nin grafiği

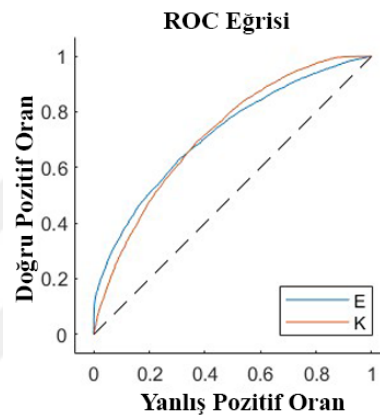


Şekil 5.20 InceptionV3'ün grafiği

ROC eğrileri incelendiğinde de benzer şekilde Özelleştirilmiş CNN modelinin diğer modellere kıyasla daha başarılı olduğu görülmektedir. Özelleştirilmiş CNN modeli ile en yüksek AUC değerine ulaşılmıştır, bu da duyarlılık ve özgüllük arasındaki en iyi dengenin sağlandığını göstermiştir.

5.3.4. Yeniden yapılan çalışma

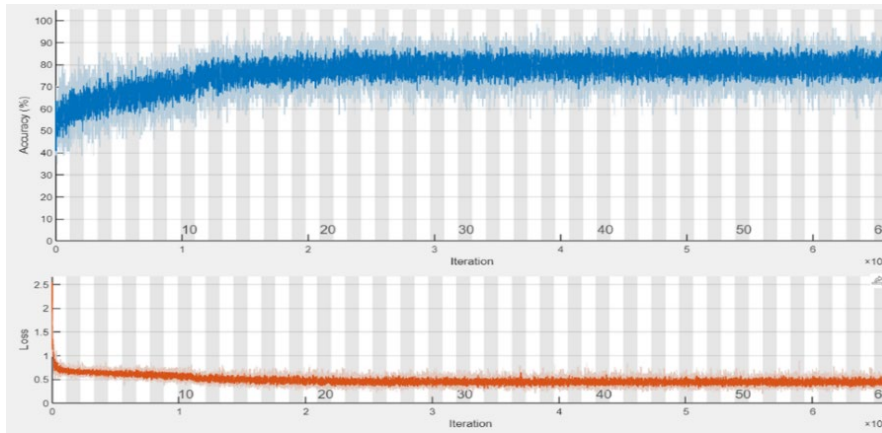
Yapılan çalışmalar ve karşılaştırmalarda en umut verici çalışmanın Özelleştirilmiş CNN algoritmasında olduğu görülmüştür. Bu sebeple Özelleştirilmiş CNN algoritması 25 epok ile yapılan deneysel çalışma, 60 epok ile yeniden uygulanmıştır. Elde edilen sonuçlar aşağıda verilmiştir (Şekil 5.21- Şekil 5.23).



Şekil 5.21 60 Epok uygulanmış Özelleştirilmiş CNN'nin AUC-ROC grafiği

		Karmaşıklık Matrisi		
		E	K	Oranlar
Cılas Sınıfı	E	3426 23.1%	1695 11.4%	66.9% 33.1%
	K	3232 21.8%	6492 43.7%	66.8% 33.2%
		ε	+	Oranlar
		51.5% 48.5%	79.3% 20.7%	66.8% 33.2%

Şekil 5.22 60 Epok uygulanmış Özelleştirilmiş CNN'nin karmaşıklık matrisi



Şekil 5.23: 60 epok uygulanmış Özelleştirilmiş CNN'nin test doğruluğu ve zamana bağlı kayıp grafiği

5.3.5. Modellerin performansı

Bu bölümde, el yazısı rakamlarından cinsiyet tahmini yapan farklı yapay zeka modellerinin performans karşılaştırmaları detaylı bir şekilde sunulmuştur. Deneylerde kullanılan modeller Inception, Özelleştirilmiş CNN, ResNet-50, VGG16, VGG19 ve Özelleştirilmiş 60 Epok uygulanan CNN'dir.

Test süreci sonucunda temel değerlendirme metriklerinden Doğruluk (Accuracy), Kesinlik (Precision), Duyarlılık (Recall), F1 Skor, Gerçek Pozitif Oran, Yanlış Pozitif Oran için sonuçlar Tablo 5.3'te gösterilmiştir. Verilen değerler tek bir sınıf için değil tüm iki sınıf için ortalama değerleri göstermektedir.

Tablo 5.3 Derin öğrenme modellerinin metrik ortalama değerleri (%)

	Doğruluk	Kesinlik	Duyarlılık	F1 Skoru	Gerçek Pozitif Oranı	Yanlış Pozitif Oranı
Inception	59,20	69,92	71,06	70,49	71,06	66,66
Özelleştirilmiş CNN	65,20	70,80	83,82	76,76	83,82	75,40
ResNet-50	62,25	71,36	75,07	73,17	75,07	65,69
VGG16	56,45	71,35	60,97	65,76	60,97	53,39
VGG19	62,46	71,94	74,18	73,04	74,18	63,10
Özelleştirilmiş CNN (60 Epok)	66,81	66,76	79,29	72,49	79,29	48,54

Özelleştirilmiş 60 Epok uygulanan CNN modeli, %66,81 doğruluk oranıyla en iyi performansı göstermiştir. Bu sonuç, modelin büyük bir veri setiyle daha uzun süreli eğitimden geçtiği için genel performansının iyileştiğini göstermektedir. En düşük doğruluk oranı VGG16 modeline (%56,45) aittir. Bu durum, modelin derinliği ve

karmaşıklığının, el yazısı rakamlarından cinsiyet tahmini gibi bir problemde yeterince etkili olmadığını göstermektedir. Kesinlik, pozitif tahminlerin gerçekte ne kadarının doğru olduğunu ifade eder. Yanlış pozitif tahminlerin minimize edilmesi gereken senaryolarda bu metrik öne çıkar. VGG19 modeli, %71,94 kesinlik oranı ile en yüksek değeri elde etmiştir. Bu durum, bu modellerin yanlış pozitif tahminlerini minimize ettiğini göstermektedir. Duyarlılık, pozitif sınıfın ne kadar doğru bir şekilde tahmin edildiğini gösterir. Özelleştirilmiş CNN, %83,82 duyarlılık oranıyla en yüksek performansa sahiptir. Bu, pozitif sınıfın büyük bir kısmının başarıyla yakalandığı anlamına gelir. VGG16, duyarlılıkta %60,97 ile en düşük performansa sahip olmuştur. Bu durum, modelin pozitif sınıfların bir kısmını yakalayamadığını göstermektedir. F1 skoru, kesinlik ve duyarlılığı dengeleyen bir metriktir. Özelleştirilmiş CNN, %76,76 ile en yüksek F1 skoruna ulaşmıştır. Bu, modelin hem duyarlılıkta hem de kesinlikte diğer modellere göre dengeli bir performans sergilediğini gösterir. VGG16, %65,76 F1 skoru ile en düşük performansı sergilemiştir. Pozitif sınıfların ne kadarının doğru tespit edildiğini ifade eder ve duyarlılıkla aynı değerleri ifade eder. Özelleştirilmiş CNN (%83,82) en yüksek performansa sahiptir. Yanlış Pozitif Oran, modellerin negatif sınıfları yanlış bir şekilde pozitif olarak tahmin etme oranını ifade eder. Özelleştirilmiş CNN (60 Epok) (%48,54), yanlış pozitif oranını minimize ederek bu metrikte en başarılı model olmuştur.

6. SONUÇ VE ÖNERİLER

Basılı, el yazısı ve görüntü dosyalarındaki karakterlerin doğru bir şekilde tanımlanması ve düzenlenebilir bir biçimde dijitalleştirilmesi ve kaydedilmesi son derece önemlidir. El yazısı karakterleri tanımanın zorluklarının üstesinden gelmek için çeşitli yaklaşımlar kullanılmış ve bu alanda derin öğrenme algoritmaları kayda değer başarılar elde edilmiştir. Bu çalışmada, el yazıları 10-18 yaş arası öğrencilerden düzenli bir biçimde, bir form aracılığıyla toplanmıştır. Daha sonra bu formlar veri kümeleri oluşturmak için tarama ve ön işleme tabi tutulmuştur. Bu çalışmada, görüntü işlemedeki başarısı ile bilinen derin öğrenme algoritmaları, el yazısı rakamlarından cinsiyet belirlemek için bir sınıflandırıcı olarak kullanılmıştır.

Bu çalışma, el yazısı rakamlarından cinsiyet tahmini yapmaya yönelik farklı modellerin performanslarını detaylı bir şekilde incelemiş ve her bir modelin öne çıkan yönlerini ortaya koymuştur.

Özelleştirilmiş CNN (60 Epok) modeli, genel doğruluk ve duyarlılık oranları ile veri setindeki en başarılı modeldir. Pozitif sınıfları yakalama oranıyla (Recall) kritik uygulamalarda tercih edilebilir. Uygulama amacına bağlı olarak, pozitif sınıfı yakalama oranı önemli olan senaryolarda Özelleştirilmiş CNN (60 Epok), dengeli performans gerektiren durumlarda ise VGG19 modelleri tercih edilebilir.

Gelecek çalışmalarda, veri seti yeniden düzenlenip çeşitlendirilerek derin öğrenme yaklaşımlarının genelleme ve sınıflandırma başarıları artırılabilir. Ayrıca, bu veri seti kullanılarak yaş, eğitim seviyesi gibi daha ayrıntılı sınıflandırmalar yapılabilir.

KAYNAKLAR

- Ahlawat, S., Choudhary, A., Nayyar, A., Singh, S. ve Yoon, B., 2020, Improved handwritten digit recognition using convolutional neural networks (CNN), *Sensors*, 20 (12), 3344.
- Al-Shamaileh, M. Z., Hassanat, A. B., Tarawneh, A. S., Rahman, M. S., Celik, C. ve Jawthari, M., 2019, New online/offline text-dependent Arabic handwriting dataset for writer authentication and identification, *2019 10th International Conference on Information and Communication Systems (ICICS)*, 116-121.
- Aydin, S. E., Kilic, S. S., Aytekin, C., Kumar, A., Porras, O., Kainulainen, L., Kostyuchenko, L., Genel, F., Kütükcüler, N. ve Karaca, N., 2015, DOCK8 deficiency: clinical and immunological phenotype and treatment options-a review of 136 patients, *Journal of clinical immunology*, 35, 189-198.
- Aytekin, N. M., 2021, Yapay Sinir Ağları ve Regresyon Yöntemleri ile Hisse Senedi Getirilerinin Tahmini: Bist-30 Üzerine Bir Uygulama.
- Babu, N. ve Soumya, A., 2019, Character recognition in historical handwritten documents—a survey, *2019 international conference on communication and signal processing (ICCSP)*, 0299-0304.
- Bartos, G. E., Hajnal, É. ve Hoşcan, Y., 2018, Comparison of Feature Extraction Techniques for Handwriting Recognition, *2018 IEEE 12th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, 000405-000410.
- Bayram, F., 2020, Derin öğrenme tabanlı otomatik plaka tanıma, *Politeknik Dergisi*, 23 (4), 955-960.
- Çapar, A., Taşdemir, K., Kılıç, Ö. ve Gökmen, M., 2003, A Turkish handprint character recognition system, *Computer and Information Sciences-ISCIS 2003: 18th International Symposium, Antalya, Turkey, November 3-5, 2003. Proceedings 18*, 447-456.
- Çetinkaya, M. ve Turhal, Ü. Ç., 2024, BİLGİSAYAR DESTEKLİ TANI YAKLAŞIMLARININ PERFORMANS DEĞERLENDİRMESİ: GELENEKSEL MAKİNE ÖĞRENMESİ VE DERİN ÖĞRENME, 1. *Uluslararası Üsküdar Bilimsel Araştırmalar Kongresi*.
- Doğan, F. ve Türkoğlu, İ., 2019, Derin öğrenme modelleri ve uygulama alanlarına ilişkin bir derleme, *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, 10 (2), 409-445.
- Erdoğan, A. A. ve Tümer, A. E., 2021, Deep Learning Method for Handwriting Recognition, *MANAS Journal of Engineering*, 9 (1), 85-92.
- Gezer, A., Yılmaz, T. ve Özdemir, D., VGG-16'da Araç Hasar Tespiti, *Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi*, 5 (2), 1-7.
- Ghosh, M. M. A. ve Maghari, A. Y., 2017, A comparative study on handwriting digit recognition using neural networks, *2017 international conference on promising electronic technologies (ICPET)*, 77-81.
- Hossain, M. A. ve Ali, M. M., 2019, Recognition of handwritten digit using convolutional neural network (CNN), *Global Journal of Computer Science and Technology*, 19 (2), 27-33.
- İnik, Ö. ve Ülker, E., 2017, Derin öğrenme ve görüntü analizinde kullanılan derin öğrenme modelleri, *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, 6 (3), 85-104.
- Inkeaw, P., Chueaphun, C., Chaijaruwanch, J., Klomsae, A. ve Marukatat, S., 2015, Lanna Dharma handwritten character recognition on palm leaves manuscript based on Wavelet transform, *2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, 253-258.
- Ioffe, S., 2015, Batch normalization: Accelerating deep network training by reducing internal covariate shift, *arXiv preprint arXiv:1502.03167*.
- Kabakus, A. T. ve Erdogmus, P., 2021, A novel handwritten Turkish letter recognition model based on convolutional neural network, *Concurrency and Computation: Practice and Experience*, 33 (21), e6429.

- Küçükşahin, N., 2019, Design of an Offline Ottoman Character Recognition System for Translating Printed Documents to Modern Turkish, *Izmir Institute of Technology (Turkey)*.
- Li, B. ve Lima, D., 2021, Facial expression recognition via ResNet-50, *International Journal of Cognitive Computing in Engineering*, 2, 57-64.
- Lin, C., Li, L., Luo, W., Wang, K. C. ve Guo, J., 2019, Transfer learning based traffic sign recognition using inception-v3 model, *Periodica Polytechnica Transportation Engineering*, 47 (3), 242-250.
- Liu, C.-L., Fink, G. A., Govindaraju, V. ve Jin, L., 2018, Special issue on deep learning for document analysis and recognition, Springer. 21: 159-160.
- Liu, W., Wei, J. ve Meng, Q., 2020, Comparisons on KNN, SVM, BP and the CNN for Handwritten Digit Recognition, *2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, 587-590.
- McCulloch, W. S. ve Pitts, W., 1943, A logical calculus of the ideas immanent in nervous activity, *The bulletin of mathematical biophysics*, 5, 115-133.
- Odabaşı, M. B. ve Toklu, M. C., 2023, Yapay sinir ağları ve derin öğrenme algoritmalarının kripto para fiyat tahmininde karşılaştırmalı analizi, *Journal of Intelligent Systems: Theory and Applications*, 6 (2), 96-107.
- Önün, F., 2019, Derin öğrenme ile yüksek çözünürlüklü hava görüntülerinde yolların tespit edilmesi, *Fen Bilimleri Enstitüsü*.
- Pekel, E., 2018, Farklı makine öğrenmesi algoritmalarının karşılaştırılması, *Fen Bilimleri Enstitüsü*.
- Raj, P. ve Wahı, A., 2013, Zone based Method to Classify Isolated Malayalam Handwritten Characters using Hu-Invariant Moments and Neural Networks, *International Journal of Computer Applications*, 975, 8887.
- Rajakumar, S. ve Bharathi, V. S., 2012, Eighth century Tamil consonants recognition from stone inscriptions, *2012 International Conference on Recent Trends in Information Technology*, 40-43.
- Romulus, P., Maraden, Y., Purnamasari, P. D. ve Ratna, A. A. P., 2015, An analysis of optical character recognition implementation for ancient Batak characters using K-nearest neighbors principle, *2015 International Conference on Quality in Research (QiR)*, 47-50.
- Sandhya, N. ve Krishnan, R., 2016, Broken Kannada character recognition—a neural network based approach, *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, 2047-2050.
- Swapna, M., Sowmya, M. ve Saraswathi, G., 2021, Handwrite Recognition using Convolutional Neural Network, *Turkish Online Journal of Qualitative Inquiry*, 12 (7).
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. ve Wojna, Z., 2016, Rethinking the inception architecture for computer vision, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2818-2826.
- Şeker, A., Diri, B. ve Balık, H. H., 2017, Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme, *Gazi Mühendislik Bilimleri Dergisi*, 3 (3), 47-64.
- TAŞDEMİR, E. F. B., 2021, Online Turkish handwriting recognition using synthetic data, *Avrupa Bilim ve Teknoloji Dergisi* (32), 649-656.
- Tawde, G. Y. ve Kundargi, J., 2013, An overview of feature extraction techniques in ocr for indian scripts focused on offline handwriting, *International Journal of Engineering Research and Applications*, 3 (1), 919-926.
- Vinjit, B., Bhojak, M. K., Kumar, S. ve Chalak, G., 2020, A review on handwritten character recognition methods and techniques, *2020 International Conference on Communication and Signal Processing (ICCSP)*, 1224-1228.
- Vural, E., Erdogan, H., Oflazer, K. ve Yanikoglu, B., 2004, An online handwriting recognition system for Turkish, *Proceedings of the IEEE 12th Signal Processing and Communications Applications Conference, 2004.*, 607-610.

- Yang, H., Jin, L., Huang, W., Yang, Z., Lai, S. ve Sun, J., 2018, Dense and tight detection of chinese characters in historical documents: Datasets and a recognition guided detector, *IEEE Access*, 6, 30174-30183.
- YILMAZ, S. M. ve Çakar, S., 2023, Yeni Bir Dilimleme Yöntemi Kullanılarak El Yazısı Rakamlarının Tanınması ve Performans Değerlendirmesi, *Sivas Cumhuriyet Üniversitesi Mühendislik Fakültesi Dergisi*, 1 (1), 38-47.
- Zafar, A., Aamir, M., Mohd Nawî, N., Arshad, A., Riaz, S., Alruban, A., Dutta, A. K. ve Almotairi, S., 2022, A comparison of pooling methods for convolutional neural networks, *Applied Sciences*, 12 (17), 8643.

