

**BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ**

**TEXT TO IMAGE AI TEKNOLOJİSİ İLE OLUŞTURULMUŞ RESİMLERİN TESPİTİ**

**BİTİRME PROJESİ**

Bilgisayar Mühendisliği Bölümü

**PROJE YAZARI**

Tuğrahan Güler

170421021

**DANIŞMAN**

Dr. Öğr. Üyesi Gözde Karataş Baydoğmuş

İSTANBUL, 2025

**MARMARA ÜNİVERSİTESİ**

**TEKNOLOJİ FAKÜLTESİ**

**BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ**

Marmara Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Öğrencisi Tuğrahan Güler tarafından “**TEXT TO IMAGE AI TEKNOLOJİSİ İLE OLUŞTURULMUŞ RESİMLERİN TESPİTİ**” başlıklı proje çalışması, 19.06.2025 tarihinde savunulmuş ve jüri üyeleri tarafından başarılı bulunmuştur.

**Jüri Üyeleri**

|  |  |  |
| --- | --- | --- |
| Dr. Öğr. Üyesi Gözde Karataş BAYDOĞMUŞ  Marmara Üniversitesi | **(Danışman)** | (İMZA)………… |
| Prof. Dr. Serhat ÖZEKEŞ  Marmara Üniversitesi | (Üye) | (İMZA)………… |
|  |  |  |
| Dr. Öğr. Üyesi Timur İNAN  Marmara Üniversitesi | (Üye) | (İMZA)………… |
|  |  |  |

**ÖNSÖZ**

Proje çalışmamız süresince karşılaştığım bütün problemlerde, sabırla yardım ve bilgilerini esirgemeyen, tüm desteğini sonuna kadar yanımda hissettiğim değerli hocam, sayın Dr. Öğr. Üyesi Gözde Karataş Baydoğmuş’a en içten teşekkürlerimi sunarım.

Bu proje çalışması fikrinin oluşması ve ortaya çıkmasındaki önerisi ve desteğinden dolayı değerli hocam Dr. Öğr. Üyesi Gözde Karataş Baydoğmuş’a teşekkür ederim.

**İÇİNDEKİLER**

[1. GİRİŞ 1](#_Toc201096706)

[1.1. Proje Çalışmasının Amacı ve Önemi 1](#_Toc201096707)

[1.2. Literatür Taraması 2](#_Toc201096708)

[1.2.1 Derin Öğrenme Modellerinin Görüntü İşleme Alanındaki Kullanımı 3](#_Toc201096709)

[1.2.2. ResNet'in Önemi ve Sahte Görsel Tespiti 3](#_Toc201096710)

[1.2.3. Diğer Derin Öğrenme Modelleri 3](#_Toc201096711)

[1.2.4. GAN'lar ve Sahte Görsel Tespiti 4](#_Toc201096712)

[1.2.5. Sahte Görsel Tespiti ve Transfer Öğrenme 4](#_Toc201096713)

[2. YÖNTEM 5](#_Toc201096714)

[2.1. Veri Seti ve Veri Hazırlığı 5](#_Toc201096715)

[2.1.1. Veri Seti Kaynakları 5](#_Toc201096716)

[2.1.2. Veri Ön İşleme 6](#_Toc201096717)

[2.1.3. Veri Temizleme ve Dengeleme 6](#_Toc201096718)

[2.1.4. Veri Artırma (Augmentasyon) ve Dönüşümleri 6](#_Toc201096719)

[2.2. Model Tasarımı ve Mimarisi 7](#_Toc201096720)

[2.3. Eğitim Süreci 8](#_Toc201096721)

[2.3.1. Eğitim Parametrelerinin Belirlenmesi 9](#_Toc201096722)

[2.3.2. Eğitimde Kullanılan Optimizasyon Yöntemleri 9](#_Toc201096723)

[2.3.3. Eğitim Sürecinin Uygulanması ve Early Stopping 9](#_Toc201096724)

[2.3.4. Eğitim ve Model Sonuçlarının Değerlendirilerek Kaydedilmesi 10](#_Toc201096725)

[2.4. Web Arayüzü ve Uygulama Katmanı 11](#_Toc201096726)

[2.4.1. Arayüz Mimarisi 11](#_Toc201096727)

[2.4.2. Görsel Yükleme ve İşleme Akışı 12](#_Toc201096728)

[2.4.3. Frontend-Backend İletişimi ve Sonuçlar 13](#_Toc201096729)

[3. BULGULAR VE TARTIŞMA 14](#_Toc201096730)

[3.1. Model Performans Analizi 15](#_Toc201096731)

[3.2. Çalışmanın Önemi ve Gelecekteki Uygulamaları 18](#_Toc201096732)

[3.3. Projenin Katkıları ve Geliştirilebilir Alanlar 19](#_Toc201096733)

[4. SONUÇLAR 21](#_Toc201096734)

[5. KAYNAKLAR 24](#_Toc201096735)

**ÖZET**

**TEXT TO IMAGE AI TEKNOLOJİSİ İLE OLUŞTURULMUŞ RESİMLERİN TESPİTİ**

Günümüzde yapay zekâ tabanlı görsel üretim araçlarının hızla gelişmesiyle birlikte, sahte görsellerin tespiti önemli bir hale gelmiştir. Bu projede, yapay zekâ tarafından üretilmiş görsellerin, gerçek görsellerden ayrılmasını sağlayabilen bir derin öğrenme modeli geliştirilmiştir. Model, 60.000 adet gerçek ve yapay görselden oluşan geniş bir veri seti ile eğitilmiştir. Kullanılan veri seti, Stable Diffusion, MidJourney ve DALL-E gibi popüler yapay zekâ araçlarıyla üretilen 30.000 sahte görsel ve Pexels, Unsplash gibi kaynaklardan alınan 30.000 gerçek görseli içermektedir. Model, ResNet50 mimarisi kullanılarak geliştirilmiş ve test seti üzerinde %92,5 doğruluk oranına ulaşmıştır. Ayrıca, kullanıcıların görsellerini yükleyerek analiz yapabileceği bir web arayüzü oluşturulmuştur. Bu sayede, görselin gerçek mi yoksa yapay mı olduğu hakkında anlık geri bildirim sağlanarak, görsel doğruluğunun hızlı bir şekilde analiz edilmesi mümkün hale gelmiştir. Proje, dijital güvenlik, medya, eğitim gibi birçok alanda kullanılabilecek potansiyele sahip olup, sahte görsellerin tespiti için pratik ve etkili bir çözüm sunmaktadır. Eğitim sürecinde kullanılan PyTorch, OpenCV gibi kütüphaneler, modelin verimli bir şekilde eğitilmesine ve görsel analizlerinin hızlıca yapılmasına olanak sağlamaktadır. Bu çalışma, sahte görsellerin tespitine yönelik var olan araştırmalara yenilikçi bir yaklaşım sunmayı hedeflemektedir.

**Anahtar Kelimeler:** Yapay Zekâ, Görsel Analiz, Derin Öğrenme, ResNet50, Sahte Görsel Tespiti, Web Arayüzü, Veri Seti

**Haziran, 2025 Öğrenciler**

**Tuğrahan Güler – 170421021**

**ABSTRACT**

**DETECTION OF IMAGES CREATED WITH TEXT TO IMAGE AI TECHNOLOGY**

Nowadays, with the rapid development of artificial intelligence-based image generation tools, the detection of fake images has become an important issue. In this project, a deep learning model has been developed to distinguish AI-generated images from real images. The model is trained on a large dataset of 60,000 real and artificial images. The dataset includes 30,000 fake images generated by popular AI tools such as Stable Diffusion, MidJourney and DALL-E and 30,000 real images from sources such as Pexels and Unsplash. The model was developed using ResNet50 architecture and achieved 92.5% accuracy on the test set. In addition, a web interface was created where users can upload their images and analyze them. In this way, it became possible to quickly analyze the accuracy of the image by providing instant feedback on whether the image is real or artificial. The project has the potential to be used in many areas such as digital security, media and education, and offers a practical and effective solution for detecting fake images. Libraries such as PyTorch and OpenCV used in the training pfess allow the model to be trained efficiently and visual analysis to be performed quickly. This study aims to provide an innovative approach to the existing research on fake image detection.

**Keywords:** Artifical Intelligence, Visual Analytics, Deep Learning, ResNet50, Fake Image Detection, Web Interface, Dataset

**June, 2025 Students**

**Tugrahan Guler – 170421021**

**SEMBOLLER**

**KISALTMALAR**

**GAN** : Generative Adversaria Network (Çekişmeli Üretici Ağ)

**AI** : Artificial Intelligence (Yapay Zekâ)

**VR** : Virtual Reality (Sanal Gerçeklik)

**AR** : Augmented Reality (Artırılmış Gerçeklik)

**GPU** : Graphics Processing Unit (Grafik İşlem Birimi)

**RAM** : Random Access Memory (Rastgele Erişim Belleği)

**CNN** : Convolutional Neural Network (Evrişimsel Sinir Ağı)

**LR** : Learning Rate (Öğrenme Oranı)

**API** : Application Programming Interface (Uygulama Programlama Arayüzü)

**ResNet** : Residual Networks

**OpenCV** : Open Source Computer Vision Library (Açık Kaynak Bilgisayar Görüsü Kütüphanesi)

**DL** : Deep Learning (Derin Öğrenme)

**HTML** : HyperText Markup Language (Hiper Metin İşaretleme Dili)

**CSS** : Cascading Style Sheet (Basamaklı Stil Sayfaları)

**JS** : Javascript

**SGD** : Stochastic Gradient Descent (Stokastik Gradyan İnişi)

**AJAX** : Asynchronous JavaScript and XML (Eşzamansız JavaScript ve XML)

**JSON** : Javascript Object Notation

**SMOTE** : Synthetic Minority Over-sampling Technique (Sentetik Azınlık Aşırı Örnekleme Tekniği)

**ŞEKİL LİSTESİ**

Şekil 2.1 Veri Setindeki Gerçek Fotoğrafların Bir Parçası 6

Şekil 2.2 Proje Ön Yüzü 12

Şekil 2.3 Örnek Sonuç Gösterimi 14

Şekil 3.1 Elde Edilen Modelin Training Curves Tablosu Sonuçları 16

Şekil 3.2 Elde Edilen Modelin Confusion Matrix Görseli 17

**TABLO LİSTESİ**

[Tablo 3.1 Modelin 5 Epoch Başına Kaydedilen Değerleri 15](#_Toc201098718)

# GİRİŞ

Yapay zekâ teknolojilerinin hızla gelişmesi, dijital görsellerin üretimi ve manipülasyonu konusunda büyük bir dönüşüm yaratmıştır. Günümüzde, yapay zekâ ile üretilen görseller, gerçekçi olmaları nedeniyle sahte görsellerin yayılmasına olanak sağlamaktadır. Bu gelişim, medya, dijital güvenlik ve eğitim gibi alanlarda önemli bir sorun teşkil etmektedir. Özellikle derin öğrenme algoritmalarının gücüyle üretilen sahte görseller hem bireylerin hem de toplumların güvenliğini tehdit edebilecek kadar gerçekçi hale gelmiştir. Bu sahte görsellerin tespit edilmesi, dijital güvenliğin sağlanması ve doğru bilgiye erişimin sağlanması açısından kritik bir öneme sahiptir. Görsel sahteciliği, özellikle medya manipülasyonu ve dijital kimlik sahteciliği gibi alanlarda büyük bir tehdit oluştururken, aynı zamanda e-ticaret ve reklamcılık gibi sektörlerde de önemli sorunlara yol açmaktadır [1]. Bu nedenle, yapay zekâ tabanlı sahte görsellerin doğru tespiti, yalnızca dijital güvenlik açısından değil, aynı zamanda toplumsal doğruluğun korunması açısından da hayati bir öneme sahiptir.

Yapay zekâ destekli sahte görsellerin tanımlanması için geliştirilen modeller, çeşitli derin öğrenme tekniklerini kullanarak bu soruna çözüm üretmektedir. Bu projede, sahte ve gerçek görselleri ayırt edebilen bir derin öğrenme modelinin geliştirilmesi amaçlanmaktadır. ResNet50 gibi güçlü derin öğrenme mimarileri, bu tür görsel sınıflandırma problemleri için başarıyla kullanılmış ve oldukça etkili sonuçlar elde edilmiştir [2]. Bu proje, görsel güvenliğini sağlayacak bir araç geliştirmeyi hedeflerken, aynı zamanda dijital dünyada karşılaşılan sahte görsel problemlerine karşı önemli bir çözüm sunmayı amaçlamaktadır. Projenin geliştirilmesinde kullanılacak olan model, eğitim verileri ile sahte görsellerin tanımlanması için optimize edilecektir [3]. Böylece, dijital ortamda karşılaşılan bu güvenlik açığının kapatılması sağlanacaktır.

## Proje Çalışmasının Amacı ve Önemi

Bu projenin temel amacı, yapay zekâ destekli sahte görsellerin tespiti için derin öğrenme tabanlı bir model geliştirmektir. Dijital ortamda sahte görsellerin yayılması, güvenlik tehditlerine ve toplumsal sorunlara yol açmaktadır. Yapay zekâ ile üretilen görseller, giderek daha gerçekçi hale gelmesiyle birlikte, bu görsellerin ayırt edilmesi giderek zorlaşmaktadır. Özellikle medya manipülasyonu, sahte haberler ve dijital kimlik sahteciliği gibi alanlarda sahte görsellerin kullanımı, bireyler ve toplumlar için büyük tehditler oluşturmuştur [4]. Bu projenin amacı, derin öğrenme tekniklerini kullanarak sahte ve gerçek görselleri ayırt edebilen bir model geliştirmektir. Böylece, dijital güvenlikteki açıkların kapatılması ve doğru bilgiye erişimin sağlanması hedeflenmektedir [5].

Sahte görsellerin tespitinde kullanılacak model, güçlü derin öğrenme mimarileri olan ResNet50 ve benzeri ağlar kullanılarak geliştirilecektir. Bu tür derin öğrenme teknikleri, görsel sınıflandırma ve doğrulama konularında oldukça etkili sonuçlar vermektedir. Örneğin, ResNet gibi derin sinir ağları, görsel tanıma ve doğrulama işlemlerinde yüksek başarıya ulaşmıştır [3]. Proje kapsamında, sahte görsellerin yanı sıra, gerçek görseller de modelin eğitim verisi olarak kullanılacaktır. Eğitim verisi, yapay zekâ ile üretilen görsellerin yanı sıra, güvenilir kaynaklardan alınan gerçek görselleri içerecektir. Bu çeşitlendirilmiş veri seti, modelin doğruluğunu artırmak için optimize edilecektir [6]. Veri seti içerisinde yer alan sahte görsellerin çoğu, Stable Diffusion, MidJourney ve DALL-E gibi gelişmiş yapay zekâ araçlarıyla üretilmiş olacaktır [1].

Modelin doğruluğu, test verileri üzerinde yapılan denemelerle değerlendirilecek ve çeşitli performans metrikleri, doğruluk oranları ve karışıklık matrisleri ile ölçülecektir. Derin öğrenme modellerinin doğruluğu, genellikle büyük veri setleri üzerinde yapılan testlerle belirlenmektedir [7]. Bu sayede, sahte görsellerin tespiti için etkin ve güvenilir bir çözüm önerilecektir. Ayrıca, kullanıcı dostu bir arayüz ile modelin entegre edilmesi sağlanacaktır. Bu sayede, kullanıcılar görsellerini yükleyerek, anında analiz yapabilecek ve görselin gerçek mi yoksa yapay mı olduğunu öğrenebilecektir [2].

Projede geliştirilen model, dijital güvenlik, medya, e-ticaret ve eğitim gibi çeşitli alanlarda kullanılabilecek ve bu sektörlerde doğruluğun artırılmasına önemli katkılar sağlayacaktır. Dijital güvenlikteki en büyük tehditlerden biri olan sahte görsellerin tespiti, bu tür projelerle güvenliğin artırılması açısından kritik öneme sahiptir [5]. Bu proje, dijital ortamda doğru bilgiye ulaşmayı ve güvenliği sağlamayı hedefleyen bir çözüm sunmaktadır.

## Literatür Taraması

Bu bölümde, derin öğrenme modellerinin sahte görsellerin tespitindeki rolü, ResNet50 modelinin tercih edilme nedenleri ve diğer alternatif modellerle karşılaştırmaları üzerinde durulacaktır. Sahte görsellerin tespiti, günümüzde medya manipülasyonu, sahte haberler ve dijital kimlik sahteciliği gibi önemli dijital güvenlik sorunlarını çözmeye yönelik kritik bir alan haline gelmiştir. Sahte görsellerin giderek daha gerçekçi hale gelmesi, bu tür tehditlerin tespit edilmesini zorlaştırmaktadır. Bu nedenle, bu alandaki ilerlemeler, yapay zekâ ve derin öğrenme alanındaki gelişmelere dayanarak hız kazanmıştır [8].

### 1.2.1 Derin Öğrenme Modellerinin Görüntü İşleme Alanındaki Kullanımı

Derin öğrenme (DL) modelleri, özellikle görüntü işleme ve bilgisayarla görü alanlarında önemli bir devrim yaratmıştır. Bu modeller, büyük miktarda veriden anlamlı özellikler çıkarabilme yetenekleri ile geleneksel makine öğrenme yöntemlerinin çok ötesindedir. Bu sebeple, sahte görsellerin tespitinde en yaygın kullanılan yöntemler arasında yer almaktadır. Özellikle CNN, GAN ve ResNet gibi modeller, bu tür görevlerde büyük başarılar elde etmiştir.

### 1.2.2. ResNet'in Önemi ve Sahte Görsel Tespiti

ResNet, 2015 yılında Microsoft Research tarafından geliştirilen bir derin öğrenme mimarisidir. ResNet, residual learning (artık öğrenme) tekniklerini kullanarak derin ağlarda karşılaşılan *vanishing gradient* problemini çözmeyi hedefler. Bu ağın temel yeniliği, katmanların derinliği arttıkça modelin eğitiminde karşılaşılan zorlukları aşabilmesidir. ResNet50 gibi daha hafif versiyonları, hesaplama gereksinimlerini karşılayarak yüksek doğruluklu sınıflandırmalar yapabilir. Bu nedenle, ResNet, sahte görsel tespiti gibi görevlerde yaygın olarak kullanılır [9].

Sahte görsellerin tespitinde ResNet'in tercih edilmesinin başlıca nedeni, yüksek doğruluk oranları ve verimli hesaplama sağlamasıdır. ResNet50, 50 katmandan oluşarak, daha derin yapılar sunmakla birlikte modelin hesaplama verimliliğini de artırır [3]. Diğer derin öğrenme ağlarına göre daha düşük hesaplama maliyetine sahip olan ResNet50, transfer öğrenme için de uygundur ve daha küçük veri setleriyle bile yüksek doğruluk elde edebilir [6].

### 1.2.3. Diğer Derin Öğrenme Modelleri

ResNet dışında, sahte görsellerin tespiti için birkaç önemli derin öğrenme modeli daha kullanılmaktadır. Bunlar arasında Inception, DenseNet, VGGNet ve EfficientNet yer alır.

1. **Inception**: Inception, daha derin ağlar kullanarak çok farklı ölçeklerdeki özellikleri öğrenebilir. Ancak, bu modellerin daha karmaşık yapıları, eğitim sürecini yavaşlatabilir. ResNet, daha verimli residual bloklarıyla benzer sonuçlar sunarak daha hızlı eğitilebilir [10].
2. **VGGNet**: VGGNet, çok derin ağlar kullanarak görsel sınıflandırma görevlerinde başarılı sonuçlar elde etmiştir. Ancak, hesaplama açısından çok pahalıdır ve derin ağlar ile performans kaybı yaşanabilir. ResNet'in daha verimli yapısı, VGGNet ile karşılaştırıldığında daha hızlı sonuçlar sunar [11].
3. **EfficientNet**: EfficientNet, daha az hesaplama kaynağı ile yüksek performans gösteren bir başka derin öğrenme modelidir. Bu modelin avantajı, ağın parametrelerini optimize ederek yüksek doğruluk elde edebilmesidir. Ancak, ResNet50 gibi sağlam temellere dayanan modellerin uygulamada daha fazla tercih edilmesinin nedeni, daha geniş veri setlerinde daha tutarlı sonuçlar sunmasıdır [12].
4. **DenseNet**: DenseNet, her katman için çıktıyı tüm önceki katmanlarla birleştirerek derin ağlarda daha etkili öğrenmeyi sağlar. Ancak, bu katmanlar arası bağlantılar, ağın hesaplama verimliliğini azaltabilir. ResNet ise bu tür bağlantıları minimalize ederek daha verimli hesaplamalar sağlar [13].

### 1.2.4. GAN'lar ve Sahte Görsel Tespiti

Sahte görsellerin tespitinde başka bir önemli model ailesi GenerativeAdversarialNetworks(GAN)'dır. GAN'lar, bir generator ve discriminator ağından oluşan iki aşamalı bir yapıya sahiptir. Bu ağlar, sahte görseller üreterek ve bunları tespit ederek birbirlerine karşı eğitilirler. GAN'lar, özellikle deepfake videoları ve sahte görsellerin üretiminde yaygın olarak kullanılmaktadır [14]. Ancak, GAN'lar çok büyük hesaplama kaynakları gerektirdiği için, genellikle ResNet gibi daha küçük ve verimli modellerle birleştirilerek daha hızlı ve doğru sonuçlar elde edilebilir.

### 1.2.5. Sahte Görsel Tespiti ve Transfer Öğrenme

Sahte görsel tespiti için transfer öğrenme, büyük veri setlerinde önceden eğitilmiş modelleri kullanarak daha küçük veri setlerinde bile başarılı sonuçlar elde edilmesine olanak tanır. ResNet50, ImageNet gibi geniş veri setlerinde eğitilmiş ve transfer öğrenme için oldukça uygundur. Bu sayede, sınırlı sayıda etiketli veri ile bile doğru sonuçlar elde edilebilir. Transfer öğrenme ile, ResNet50 gibi modellerin önceden öğrenilen özellikleri, sahte görsellerin tespiti için yeniden kullanılabilir [15].

# YÖNTEM

## Veri Seti ve Veri Hazırlığı

### 2.1.1. Veri Seti Kaynakları

Bu projede kullanılan veri seti, sahte görsellerin tespitini sağlamak amacıyla, hem yapay zekâ tabanlı araçlarla üretilen sahte görsellerden hem de güvenilir kaynaklardan alınan gerçek görsellerden oluşmaktadır. Veri seti, farklı kaynaklardan temin edilen 60.000 adet görselden oluşmaktadır:

* **Sahte Görseller**: Sahte görseller, Stable Diffusion, MidJourney ve DALL-E gibi popüler yapay zekâ araçları tarafından üretilmiştir. Bu araçlar, metin tabanlı açıklamalardan görsel üretme yeteneğine sahip olduğundan, dijital manipülasyonlar konusunda yüksek doğrulukta sahte görseller üretmektedir.
* **Gerçek Görseller**: Gerçek görseller ise Pexels ve Unsplash gibi güvenilir, telif hakkı olmayan görsel kaynaklarından temin edilmiştir. Bu kaynaklar, doğru etiketlenmiş ve yüksek kaliteli görseller sağlar. Bu görseller, sahte görsellerle karşılaştırılacak ve modelin eğitiminde kullanılacaktır.

metin, ekran görüntüsü, multimedya yazılımı, yazılım içeren bir resim

Yapay zeka tarafından oluşturulmuş içerik yanlış olabilir.

Şekil 2.1 Veri Setindeki Gerçek Fotoğrafların Bir Parçası

### 2.1.2. Veri Ön İşleme

Veri hazırlama süreci, modelin doğru şekilde öğrenebilmesi için oldukça kritik bir adımdır. Veriler üzerinde yapılan ön işleme adımları şu şekilde özetlenebilir:

* **Boyutlandırma**: Görseller, modelin giriş katmanına uygun hale gelmesi için 224x224 piksel boyutlarına yeniden boyutlandırılmıştır. Bu, ResNet50 modelinin eğitilmesi için yaygın olarak kullanılan bir boyutlandırmadır [3].
* **Normalize Etme**: Görsellerin renk değerleri (RGB), modelin daha hızlı ve etkili öğrenmesi için normalleştirilmiştir. Görsellerin her bir kanalındaki (Kırmızı, Yeşil, Mavi) piksel değerleri, [0, 1] aralığına çekilmiştir. ResNet50'nin daha önce eğitildiği ImageNet veri setine benzer bir şekilde, her kanal için ortalama ve standart sapma değerleri kullanılarak normalize edilmiştir [16].

### 2.1.3. Veri Temizleme ve Dengeleme

Veri temizleme, modelin eğitim sürecinde doğru sonuçlar alabilmesi için oldukça önemlidir. Aşağıdaki adımlar veri temizleme sürecinde uygulanmıştır:

* **Eksik ve Hatalı Etiketler**: Görsellerin doğru etiketlendiğinden emin olmak için her görselin etiketleri kontrol edilmiştir. Eksik veya hatalı etiketler düzeltilmiş, veri setinin doğruluğu artırılmıştır.
* **Dengeleme**: Veri setindeki sınıfların (gerçek ve sahte görseller) dengeli olmasına dikkat edilmiştir. Eğer veri setinde dengesizlik bulunuyorsa, SMOTE gibi tekniklerle veri çoğaltma yapılabilir. Ancak, bu projede veri seti zaten dengeli olduğundan, ekstra bir işlem yapılmamıştır.

### 2.1.4. Veri Artırma (Augmentasyon) ve Dönüşümleri

Modelin daha sağlam ve genelleme yeteneği yüksek sonuçlar verebilmesi için veri artırma yöntemleri kullanılmıştır. Bu yöntemler, eğitim verisinin çeşitlendirilmesi ve modelin daha fazla farklı koşulda eğitilmesini sağlamaktadır. Uygulanan veri artırma teknikleri şunlardır:

* **Random Horizontal Flip**: Görsellerin rastgele yatay olarak çevrilmesi, modelin sağa ve sola yönelimli nesneleri tanımasını sağlar.
* **Random Rotation**: Görsellerin rastgele belirli bir açıyla döndürülmesi, modelin döner nesneleri tanıma yeteneğini artırır.
* **Color Jitter**: Görsellerin parlaklık, kontrast ve doygunluk gibi renk özelliklerinin rastgele değiştirilmesi, modelin ışık değişimlerine karşı daha sağlam hale gelmesini sağlar.
* **Random Crop ve Resize**: Görsellerin rastgele kırpılması ve ardından belirli bir boyuta yeniden boyutlandırılması, modelin farklı boyutlardaki nesneleri tanımasını sağlar.

Bu artırma teknikleri sayesinde, modelin overfitting (aşırı öğrenme) yapma olasılığı azaltılmıştır ve model daha genel bir öğrenme kapasitesine sahip olmuştur.

## Model Tasarımı ve Mimarisi

Bu projede, sahte ve gerçek görsellerin tespiti için derin öğrenme tabanlı ResNet50 mimarisi kullanılmıştır. ResNet50, Residual Networks (ResNet) ailesinin bir parçasıdır ve derin ağlarda karşılaşılan en büyük sorunlardan biri olan vanishing gradient problemini çözmeyi amaçlayan bir yapıdır. Bu problem, ağ derinleştikçe gradyanların küçülmesi ve dolayısıyla ağın öğrenme kapasitesinin azalmasıdır. ResNet, her katman arasına eklenen shortcut bağlantıları sayesinde bu sorunu aşmayı başarmıştır. Shortcut bağlantıları, öğrenilen temsillerin daha derin katmanlara aktarılmadan önce bir sonraki katmana doğrudan geçmesini sağlar, bu sayede derin ağların eğitiminde etkili sonuçlar alınabilir. ResNet50, bu yapısıyla derin öğrenme modellerinde yaygın olarak kullanılmaktadır.

ResNet50 mimarisi, daha önce büyük veri setleriyle eğitilmiş ve çok iyi sonuçlar elde etmiştir. Bu, modelin transfer learning (transfer öğrenme) için ideal bir seçenek olmasını sağlamaktadır. Transfer öğrenme, önceden eğitilmiş bir modelin ağırlıklarını yeni bir görev için uyarlamayı ifade eder. Bu projede, ResNet50 modelinin önceden eğitilmiş ağırlıkları kullanılarak modelin eğitilme süresi büyük ölçüde kısaltılmış ve modelin doğruluğu artırılmıştır. Bu yaklaşım, derin öğrenme modellerinin daha verimli bir şekilde çalışmasını sağlayan güçlü bir yöntemdir.

ResNet50, her katmanda özellik çıkarımını derinlemesine gerçekleştiren ve fully connected (tam bağlantılı) katmanlardan oluşan bir yapıya sahiptir. Bu modelin son katmanlarında yapılan değişiklikler, projenin hedeflerine uygun hale getirilmiştir. ResNet50'nin son katmanları, genellikle çok sayıda sınıfı tanıyabilmek için softmax aktivasyon fonksiyonu kullanır. Ancak, bu projede yalnızca iki sınıf (gerçek ve sahte görseller) bulunduğu için modelin son katmanı binary classification (ikili sınıflandırma) amacıyla yeniden yapılandırılmıştır. Bu yapılandırma, modelin yalnızca iki sınıf arasında bir seçim yapabilmesini sağlar.

Modelin son katmanında, 2048 nöronlu bir dense (yoğun) katman yer almaktadır [16]. Bu katmandan çıkan sonuçlar, ikinci bir katmana geçerek sınıflandırma yapılır. İkinci katman, 2 nöron içerir ve bu nöronlar, görsellerin gerçek veya sahte olduğunu belirler. Bu yapı, modelin sınıflandırma sırasında sahte ve gerçek görselleri doğru bir şekilde ayırt edebilmesini sağlar. Sonuçlar, softmax fonksiyonu kullanılarak elde edilen olasılıklarla sınıflandırılır ve modelin hangi sınıfa ait olduğunu belirler.

Modelin doğru çalışabilmesi için, eğitim sırasında dropout ve batch normalization gibi teknikler de kullanılmıştır. Dropout, modelin aşırı öğrenmesini (overfitting) engellemek için kullanılan bir tekniktir. Eğitim sırasında, belirli nöronlar rastgele devre dışı bırakılarak ağın genelleme kapasitesi artırılır. Bu yöntem, özellikle derin ağlarda overfitting sorununu azaltmak için yaygın olarak kullanılmaktadır. Batch normalization ise her katmandaki girişlerin normalleştirilmesini sağlar ve böylece modelin daha hızlı öğrenmesine yardımcı olur.

Sonuç olarak, ResNet50 mimarisi, bu projede sahte görsellerin tespitinde etkili bir şekilde kullanılmıştır. Modelin mimarisi, görsel sınıflandırma problemlerine uygun şekilde yapılandırılmış ve son katmanları, yalnızca iki sınıfı doğru şekilde ayırt edecek şekilde tasarlanmıştır. Bu yapı, derin öğrenme alanındaki başarılı uygulamalara dayanarak, projede yüksek doğruluk oranları elde edilmesini sağlamaktadır.

## Eğitim Süreci

Modelin eğitim süreci, doğru sonuçların elde edilmesi için kritik bir aşamadır. Bu süreç, doğru hiperparametre ayarları, optimizasyon algoritmalarının kullanımı ve uygulanan eğitim stratejileriyle birlikte modelin doğruluğunu ve genel öğrenme kapasitesini en üst seviyeye çıkarmayı amaçlamaktadır. Eğitim sırasında kullanılan donanım, GPU tarafında bir 1650Ti ekran kartı ve 16GB RAM ile sınırlıydı. Bu donanım, günümüzün daha güçlü donanımlarına kıyasla daha sınırlı bir kapasiteye sahip olsa da modelin eğitiminde oldukça başarılı bir performans sergiledi. Daha güçlü bir donanım ile eğitim süreci hızlandırılabilir ve modelin doğruluğu daha da iyileştirilebilir. Ancak mevcut donanım ile yapılan eğitim süreci, modelin yeterli doğrulukta sonuçlar elde etmesini sağlamış ve çalışma amacına ulaşmıştır.

Veri setinin %80'i (48.000 görsel) modelin eğitiminde kullanılmıştır. Geriye kalan %20'si (12.000 görsel) ise test veri seti olarak ayrılmıştır. Bu test verisi, 6.000 gerçek görsel ve 6.000 yapay zekâ tarafından üretilen sahte görselden oluşmaktadır.

### 2.3.1. Eğitim Parametrelerinin Belirlenmesi

Eğitim sırasında kullanılan hiperparametreler, modelin öğrenme sürecinde önemli bir rol oynamaktadır. Bu projede kullanılan temel parametreler şunlardır:

* **Batch Size**: Her iterasyonda modelin üzerinde çalışacağı örnek sayısını belirler. Bu projede 32 olarak seçilmiştir, böylece model her iterasyonda 32 görselle eğitim yapar.
* **Epoch Sayısı**: Eğitim verisi üzerinde modelin kaç kez geçiş yapacağı belirlenir. Bu projede toplamda 30 epoch belirlenmiştir, bu da modelin tüm eğitim verisi üzerinde yeterli sayıda geçiş yapabilmesini sağlar.
* **Öğrenme Oranı (Learning Rate)**: Modelin ağırlıklarını ne kadar hızlı güncelleyeceğini belirler. Başlangıçta 0.001 olarak belirlenmiştir. Bu değer, eğitim sırasında modelin daha etkili öğrenmesini sağlayacak şekilde optimize edilmiştir.

### 2.3.2. Eğitimde Kullanılan Optimizasyon Yöntemleri

Bu projede, “Adam Optimizer” kullanılmıştır. Adam,SGD algoritmasının bir iyileştirmesi olarak geliştirilmiştir ve özellikle büyük veri setlerinde ve karmaşık modellerde yüksek performans gösterir.

Adam, her parametre için öğrenme oranını adaptif olarak ayarlayarak, her iterasyonda daha hızlı ve verimli öğrenmeyi sağlar. Adam optimizasyon algoritması, gradyanların ilk momentlerini (ortalamasını) ve ikinci momentlerini (karelerin ortalamasını) kullanarak, daha hızlı ve düzgün bir şekilde öğrenme süreci gerçekleştirilmesine olanak tanır. Bu optimizasyon, modelin eğitim sürecini hızlandırarak, daha kısa sürede daha iyi sonuçlar elde edilmesini sağlar [17].

### 2.3.3. Eğitim Sürecinin Uygulanması ve Early Stopping

Eğitim sürecinde, modelin train ve test veri kümeleri arasında ayrım yapılarak eğitim gerçekleştirilmiştir. Modelin parametreleri, yalnızca train veri kümesi üzerinde güncellenmiş ve modelin başarısı, test veri kümesi ile değerlendirilmiştir. Eğitim sürecinin her epoch'unda, modelin loss (kayıp) ve accuracy (doğruluk) değerleri takip edilmiştir. Bu metrikler, modelin ne kadar iyi performans gösterdiğini değerlendirmek için kullanılmıştır.

Eğitim süreci sırasında, her batch (veri kümesi parçası) işlendiğinde ilerleme çubuğu güncellenmiştir. Bu, modelin eğitim sürecinin takip edilmesini sağlamış ve her epoch'un ne kadar ilerlediği konusunda görsel bir geribildirim sunmuştur. Ayrıca, batch başına kayıp ve doğruluk gibi metrikler de anlık olarak hesaplanarak ekrana yazdırılmıştır. Bu işlem, eğitim sürecindeki ilerlemeyi daha şeffaf hale getirmiştir.

Modelin fazla öğrenmesini engellemek amacıyla, “earlystopping” stratejisi kullanılmıştır. Eğer doğrulama kaybı (validation loss) birkaç epoch boyunca iyileşmemişse, eğitim erken sonlandırılmıştır. Bu, modelin test veri kümesine fazla uyum sağlamadan genelleme yapmasını sağlar ve overfitting (aşırı öğrenme) problemini engeller.

Modelin doğruluğu her epoch sonunda kontrol edilmiştir. En iyi doğruluk değerine sahip modelin ağırlıkları kaydedilmiştir. Eğitim sürecinde doğrulama seti üzerinde en iyi performansı gösteren modelin ağırlıkları, “models” dizinine kaydedilmiş ve böylece gelecekte kullanılmak üzere en iyi model saklanmıştır.

### 2.3.4. Eğitim ve Model Sonuçlarının Değerlendirilerek Kaydedilmesi

Eğitim ve doğrulama sırasında elde edilen kayıp ve doğruluk değerleri, logs dosyasına kaydedilmiştir. Bu log dosyası, modelin eğitim sürecindeki tüm önemli verileri içerir ve gelecekte yapılacak iyileştirmeler veya analizler için faydalı olacaktır. Ayrıca, eğitim tamamlandığında testveri kümesi üzerinde modelin genel başarısı değerlendirilmiş ve test doğruluğu kaydedilmiştir.

Eğitim süreci sonrasında, modelin performansı test verisi üzerinde değerlendirilmiştir. Bu değerlendirme sırasında accuracy, precision, recall, f1-score gibi metrikler kullanılarak modelin doğruluğu ve genelleme kapasitesi ölçülmüştür. Ayrıca, confusion matrix (karışıklık matrisi) kullanılarak modelin hangi sınıflarda hata yaptığı görsel olarak incelenmiştir. Modelin başarısı bu metriklerle değerlendirilmiştir.

## Web Arayüzü ve Uygulama Katmanı

Bu proje, kullanıcılara yapay zekâ tarafından oluşturulmuş görselleri tespit etme imkânı sunan bir platform geliştirmektedir. Uygulama, iki ana katmandan oluşmaktadır: Frontend (kullanıcı arayüzü) ve Backend (sunucu ve model katmanı). Frontend, kullanıcıların görsel yüklemesi, görselin işlenmesi ve sonuçların görsel/metin olarak kullanıcıya sunulmasını sağlar. Backend ise, yüklenen görseli alır, model aracılığıyla analiz eder ve sonucu kullanıcıya ileterek, tüm işlemin gerçekleştirilmesini sağlar.

### 2.4.1. Arayüz Mimarisi

Web arayüzü, HTML, CSS ve JavaScript kullanılarak geliştirilmiştir. Kullanıcı dostu ve interaktif bir tasarım hedeflenerek, görsel yükleme, ön izleme ve analiz etme işlemleri kolaylıkla yapılabilecek şekilde düzenlenmiştir. Web sitesi, kullanıcıların görseli hızlı ve kolay bir şekilde yükleyebileceği, ardından modelin analizi gerçekleştireceği adımları takip eder.

**HTML/CSS/JS Yapısı:**

* **HTML**: Temel yapıyı sağlar, sayfa düzeni ve içerik düzeni burada tanımlanır.
* **CSS**: Arayüzün stilini belirler. Kullanıcı dostu ve şık bir görünüm için modern tasarım öğeleri kullanılmıştır.
* **JavaScript (JS)**: Kullanıcı etkileşimlerini yönetir, görsel yükleme, görselin işlenmesi ve sonuçların gösterilmesi gibi işlemleri sağlar. Ayrıca, backend ile iletişimi sağlayan kodları içerir.

Arayüzde, görsel yükleme alanı, analiz etme butonu ve sonuçların gösterileceği alan gibi ana bölümler bulunmaktadır. Kullanıcı, görseli sürükleyip bırakabilir ya da dosya gezgini aracılığıyla seçebilir. Yüklenen görsel, kullanıcıya küçük bir önizleme olarak gösterilir ve ardından model analizi için "Analiz Et" butonuna tıklanarak işlem başlatılır.

Web arayüzü, mobil cihazlarda da düzgün çalışacak şekilde optimize edilmiştir. Mobil cihazlarda menü, responsive (duyarlı) tasarım ile otomatik olarak uyumlu hale gelir. Kullanıcılar, mobil cihazlarıyla da kolayca görsel yükleyip analiz yapabilirler. Frontend tasarımında, flexbox ve CSS grid gibi modern CSS özellikleri kullanılarak tasarımın tüm ekran boyutlarına uyum sağlaması amaçlanmıştır. Ayrıca, mobil cihazlar için özel bir menü butonu (hamburger menüsü) eklenmiştir. Bu buton, menü öğelerinin mobil cihazlarda düzenli bir şekilde görüntülenmesini sağlar.

Web arayüzü, kullanıcı dostu ve işlevsel bir platform sunmayı hedeflemektedir. Görsel yükleme, ön izleme, analiz etme ve sonuçları görüntüleme işlemleri, adım adım kullanıcıya rehberlik edecek şekilde düzenlenmiştir. Eğitimli modelin yüksek doğrulukla çalıştığı ve kullanıcıların görselleri hızlı bir şekilde analiz edebileceği bir platform oluşturulmuştur. Yüklenen görsellerin doğru şekilde işlenmesi ve sonuçların görsel/metinsel olarak sunulması, kullanıcı deneyimini arttırmaktadır. Bu sistem, kullanıcıların herhangi bir teknik bilgiye ihtiyaç duymadan yapay zekâ tarafından üretilen görselleri kolaylıkla tespit etmelerine yardımcı olur.

metin, ekran görüntüsü, yazılım, web sayfası içeren bir resim

Yapay zeka tarafından oluşturulmuş içerik yanlış olabilir.

Şekil 2.2 Proje Ön Yüzü

### 2.4.2. Görsel Yükleme ve İşleme Akışı

Kullanıcı, yüklediği görseli backend'e gönderir. Backend, görseli alır, modele gönderir ve sonuçları geri döner. Backend kısmında Flask framework'ü kullanılmakta olup, bu framework HTTP isteklerine cevap verir. Görselin analizi sırasında, model ResNet50 ile eğitilmiş olup, sahte ve gerçek görselleri sınıflandırmaktadır.

Kullanıcı, görseli yüklemek için "Dosya Seç" butonuna tıklayarak ya da görseli sürükleyip bırakarak yükleme işlemini başlatır. HTML5'in dosya API'si kullanılarak, yüklenen dosya işlenir ve ön izleme kısmında kullanıcıya gösterilir. Bu işlem, frontend tarafından yönetilir ve görselin doğru şekilde yüklenmesi sağlanır. Görsel, base64 formatına dönüştürülerek backend'e gönderilir. Bu format, görselin metin bazlı bir şekilde iletilmesine imkân tanır ve verinin daha verimli bir şekilde taşınmasını sağlar.

Yüklenen görsel, Flask sunucusunda çalışan backend kısmına iletilir. Backend'deki model, gelen görseli alır, ön işleme adımlarından geçirir ve tahmin işlemini başlatır. Model, görselin yapay zekâ tarafından üretilip üretilmediğini analiz eder ve bir sınıf tahmini (gerçek ya da sahte) döndürür. Ayrıca, tahmin edilen sınıfın güven oranını da hesaplar. Bu oran, kullanıcıya görsel ve metin olarak sunulur. Modelin verdiği sonuçlar ve tahmin oranları, frontend tarafında kullanıcılara görsel olarak aktarılır.

### 2.4.3. Frontend-Backend İletişimi ve Sonuçlar

Frontend ile backend arasındaki iletişim AJAX ve fetch API ile sağlanmaktadır. Kullanıcı, görseli yükledikten sonra, görsel verisi frontend tarafından base64 formatında backend'e gönderilir. Backend tarafında, Flask API istekleri kabul eder ve modelle yapılan tahminin sonucunu JSON formatında frontend'e iletir. Bu işlem asenkron bir şekilde gerçekleşir ve kullanıcı herhangi bir gecikme hissetmeden sonuçları alır. Ayrıca, frontend tarafındaki analyzeBtn (analiz et butonu) etkinleştirildiğinde, kullanıcı bu butona tıkladığında, işlemin başladığını göstermek için animasyonlu bir yükleme simgesi de görüntülenir.

Kullanıcı, görselin analiz edilmesinin ardından sonuçları görsel ve metin olarak sunulmaktadır. Sonuçlar, modelin tahmin ettiği sınıf ve tahminin doğruluk oranı ile kullanıcıya sunulmaktadır. Görselin, "gerçek" veya "yapay" olduğu bilgisi metin olarak kullanıcıya bildirilir. Sonuçlar, ayrıca görsel üzerinde bir gösterim ile kullanıcıya sunulur. Görselin doğruluğu, ilgili "probability bars" (probabilite çubukları) kullanılarak görselleştirilir ve her sınıf için elde edilen doğruluk oranları kullanıcıya sunulur.

Bu adımda, confidence level (güven oranı) ve class probabilities (sınıf olasılıkları) gibi ek bilgilerle kullanıcıya modelin güven seviyeleri hakkında detaylı bilgi verilir. Görselin yanında, tahmin edilen sınıfın ikonu ve renkli bir işaret yer alır, böylece kullanıcı daha hızlı bir şekilde sonuçları anlayabilir.

Sonuçlar, kullanıcıya görsel ve metin olarak sunulur. Frontend kısmında, modelin tahmin ettiği sınıf, güven oranı ve sınıf olasılıkları, özel bir panelde görselin yanında yer alır. Kullanıcıya hem görsel hem de yazılı bilgi sağlanır. Bu sonuçlar, görselin hemen altında bir result-container içinde gösterilir ve kullanıcıya sonuç hakkında açık bilgi verilir. Sonuç ekranında ayrıca, olasılıkları gösteren çubuklar yer alır ve her sınıf için elde edilen olasılık değeri kullanıcıya görsel olarak sunulur. Bu, görselin analizine dayalı daha fazla bilgi edinmesini sağlar.

metin, ekran görüntüsü, web sitesi, web sayfası içeren bir resim

Yapay zeka tarafından oluşturulmuş içerik yanlış olabilir.

Şekil 2.3 Örnek Sonuç Gösterimi

# BULGULAR VE TARTIŞMA

Model eğitimi tamamlandıktan sonra, doğrulama verisi üzerinde yapılan değerlendirmeler doğrultusunda çeşitli analizler gerçekleştirilmiştir. Aşağıda sunulan bulgular modelin genel başarısını yansıtmaktadır.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Epoch | Train\_Loss | Train\_Accuracy | Validation\_Loss | Validation\_Accuracy | LR |
| 1 | 0.48077 | 0.775937 | 0.504268 | 0.7725 | 0.001 |
| 5 | |  | | --- | |  |  |  | | --- | | 0.359554 | | 0.842812 | 0.390943 | 0.836083 | 0.001 |
| 10 | 0.318003 | 0.864688 | 0.321764 | 0.868833 | 0.001 |
| 15 | 0.288883 | 0.879271 | 0.304654 | 0.869167 | 0.001 |
| 20 | 0.262578 | 0.891458 | 0.314453 | 0.864583 | 0.001 |
| 25 | 0.189283 | 0.922771 | 0.1947 | 0.924167 | 0.001 |
| 30 | 0.174663 | 0.930958 | 0.191523 | 0.9255 | 0.001 |

Tablo 3.1 Modelin 5 Epoch Başına Kaydedilen Değerleri

## 3.1. Model Performans Analizi

İlk birkaç epoch boyunca, modelin eğitim ve doğrulama doğruluğu sürekli olarak arttı. Örneğin, eğitim doğruluğu 1. epoch'ta %77,59 iken, 5. epoch'ta %84,28'e yükseldi. Aynı şekilde, doğrulama doğruluğu da %77,25'ten %83,60'a kadar arttı.

Eğitim süreci boyunca, eğitim ve doğrulama kayıplarının dikkate değer şekilde azaldığı görüldü. Eğitim kaybı 0.48'den 0.36'ya, doğrulama kaybı ise 0.50'den 0.39'a düştü. Bu, modelin verilerden etkili bir şekilde öğrenip genelleme yeteneğinin arttığını gösteriyor.

Ancak, eğitim devam ettikçe özellikle 10. epoch'tan sonra doğrulama doğruluğunda bazı dalgalanmalar gözlemlendi. Eğitim doğruluğu artmaya devam ederken, doğrulama doğruluğu sabit kalmaya başladı. Bu, modelin aşırı uyum sağladığını ve eğitim verisi üzerinde çok iyi performans gösterirken doğrulama seti üzerinde daha az genel hale geldiğini gösteriyor.

metin, çizgi, diyagram, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Yapay zeka tarafından oluşturulmuş içerik yanlış olabilir.

Şekil 3.1 Elde Edilen Modelin Training Curves Tablosu Sonuçları

Confusion matrix (karışıklık matrisi) sonuçlarına bakıldığında, modelin genel olarak oldukça iyi performans gösterdiği görülmektedir. Modelin doğru sınıflandırmalar yaptığı örneklerin sayısı oldukça yüksektir. Ancak, her modelde olduğu gibi, bu sonuçların bazı iyileştirilmesi gereken yönleri de bulunmaktadır. Karışıklık matrisi bize doğru ve yanlış sınıflandırmalar hakkında önemli bilgiler sunar. Bu sonuçlar, modelin sahte ve gerçek görselleri ayırt etme yeteneğini daha iyi anlamamıza yardımcı olmaktadır.

metin, ekran görüntüsü, diyagram, dikdörtgen içeren bir resim

Yapay zeka tarafından oluşturulmuş içerik yanlış olabilir.

Şekil 3.2 Elde Edilen Modelin Confusion Matrix Görseli

Özellikle, modelin 488 yanlış pozitif (false positive) ve 406 yanlış negatif (false negative) sınıflandırma yaptığı gözlemlenmiştir. Yanlış pozitif sayısı, modelin sahte görselleri yanlışlıkla gerçek olarak sınıflandırdığı durumları belirtir. Bu değer, modelin genellikle gerçek görselleri doğru şekilde tanıdığını gösterse de sahte görsellerin bazı özelliklerinin gerçek görsellerle benzer olabileceğini de ima etmektedir. Sahte görsellerin bazen yüksek kaliteli, insan yapımı içerikler veya AI ile üretilmiş görüntüler gibi bazı detaylara sahip olması, modelin bu görselleri yanlışlıkla gerçek olarak tanımasına sebep olabilir.

Diğer taraftan, modelin 406 yanlış negatif (false negative) sınıflandırması yaptığı gözlemlenmiştir. Bu durum, modelin gerçek görselleri yanlışlıkla sahte olarak sınıflandırdığı ve bu görselleri doğru şekilde tanımlamada zorlandığı anlamına gelir. Yanlış negatiflerin sayısının 488 olan yanlış pozitiflerden daha düşük olması, modelin gerçek görselleri çoğunlukla doğru şekilde sınıflandırmaya daha yakın olduğunu gösteriyor. Ancak, bu yanlış negatifler, modelin bazı durumlarda gerçek görsellerin belirli özelliklerini kaçırdığı veya görselin içerdiği ince ayrıntıları analiz etmekte zorluk yaşadığı anlamına gelebilir. Özellikle karmaşık sahnelerde veya belirli pozisyon ve açıdaki gerçek görsellerde bu tür hataların yapılması olasıdır.

Bununla birlikte, modelin precision (kesinlik) ve recall (duyarlılık) değerleri oldukça iyi sonuçlar göstermektedir. Precision %92, recall ise %93 olarak hesaplanmıştır. Precision, modelin gerçek olarak sınıflandırdığı görsellerin ne kadarının doğru olduğunu belirtirken, recall, modelin doğru sınıflandırdığı gerçek görsellerin oranını ifade eder. Yüksek precision ve recall değerleri, modelin çoğunlukla doğru sınıflandırmalar yaptığını ve her iki sınıfın da doğru şekilde tanımlanmasında başarılı olduğunu göstermektedir. Özellikle recall değerinin yüksek olması, modelin çoğu gerçek görseli doğru şekilde tanıdığını ve hata oranının düşük olduğunu gösteriyor.

Bu yüksek değerler, modelin genellikle sahte ve gerçek görseller arasında doğru sınıflandırmalar yapma yeteneğini gösteriyor. Precision ve recall'daki yüksek oranlar, modelin güvenilir olduğunu ve görsel analizde sağlam bir temel oluşturduğunu gösteriyor. Yanlış pozitiflerin ve yanlış negatiflerin sayısının birbirine yakın olması, modelin doğru ve yanlış sınıflandırmalar arasında dengesiz bir eğilim göstermediğini ve her iki sınıf için de belirli bir genelleme yapma kapasitesine sahip olduğunu ortaya koyuyor.

Modelin başarısını artırmak için yapılabilecek bazı iyileştirmeler ise doğrulama seti üzerindeki performansı daha da artırmak için veri çeşitliliğini artırmak ve modelin aşırı uyum yapmasını engellemek için düzenleme tekniklerini daha fazla kullanmak olabilir. Ayrıca, modelin daha geniş ve farklı veri setleriyle eğitilmesi, modelin daha iyi genelleme yapmasını sağlayabilir ve sahte görsellerin tanınmasında daha iyi performans elde edilmesine olanak tanıyabilir. Bu, özellikle sahte görsellerin üretiminde kullanılan farklı tekniklerin zamanla gelişmesi ve çeşitlenmesiyle birlikte daha da önemli hale gelecektir.

Özetle, modelin confusion matrix sonuçları, modelin genel performansının oldukça başarılı olduğunu ancak bazı iyileştirme alanlarının olduğunu göstermektedir. Yanlış pozitif ve yanlış negatiflerin sayılarının düşürülmesi, modelin genelleme kapasitesinin artırılması ve eğitim verilerinin çeşitlendirilmesi ile modelin doğruluğu daha da geliştirilebilir.

Eğitim sürecinin sonunda, modelin eğitim doğruluğu %93'e ulaşırken, doğrulama doğruluğu da %92,5 seviyelerine çıkmıştır. Bu, modelin eğitim verisi üzerinde olumlu sonuçlar elde ettiğini ve aynı zamanda daha önce görmediği doğrulama verileri üzerinde de güçlü bir genelleme kapasitesine sahip olduğunu göstermektedir. Eğitim sürecinin başlarında, eğitim doğruluğu ve doğrulama doğruluğu arasında belirli bir fark olsa da son epochlar itibarıyla doğrulama doğruluğundaki dalgalanmalar azalmış ve modelin genelleme yeteneği artmıştır. Özellikle son epoch (30. epoch) ile, eğitim doğruluğundaki sürekli artış, modelin öğrenme sürecinin oldukça verimli olduğunu ve modelin, daha önceki epochlarda öğrendiği bilgileri başarıyla test verisi üzerinde uyguladığını göstermektedir. Bu sonuçlar, modelin eğitim sırasında gösterdiği başarının doğrulama verisi üzerinde de sürdürülebilir olduğunu ortaya koymakta ve modelin potansiyel olarak gerçek uygulamalarda da yüksek başarı gösterebileceğini işaret etmektedir.

## 3.2. Çalışmanın Önemi ve Gelecekteki Uygulamaları

Bu çalışma, yapay zekâ tarafından üretilen görsellerin tespiti alanında önemli bir katkı sağlamaktadır. Günümüz dijital çağında, sahte görsellerin hızlı bir şekilde üretilebilmesi ve dağıtılabilmesi, özellikle sosyal medya, haber kaynakları ve dijital kimlik doğrulama süreçlerinde ciddi güvenlik ve güvenilirlik problemleri yaratmaktadır. Yapay zekâ ve derin öğrenme teknikleri kullanılarak yapılan bu çalışma, görsellerin gerçek mi yoksa yapay zekâ tarafından mı üretildiğini tespit etmek için geliştirilmiş güçlü bir model sunmaktadır. Çalışma, yalnızca akademik bir araştırma olarak kalmayıp, aynı zamanda pratik uygulamalara da taşınabilir.

Bu modelin en önemli katkılarından biri, sahte görsellerin tespiti ve bu görsellerin potansiyel tehlikelerinin önceden belirlenmesine olanak sağlamasıdır. Bu, özellikle medya, güvenlik, dijital kimlik yönetimi ve hukuk gibi alanlarda çok kritik bir öneme sahiptir. Medya manipülasyonu, sahte haberler ve dijital kimlik sahteciliği gibi tehditlerin önüne geçilmesi, toplumu bu tür tehditlerden korumak için oldukça önemlidir. Bu tür bir modelin gelişmesi, kullanıcıların daha güvenilir bilgiye erişimini sağlayarak toplumsal güveni artırabilir.

Gelecekte, bu modelin daha geniş çapta uygulamalara uyarlanması mümkün olacaktır. Örneğin, medya ve haber platformlarında, gerçek ve sahte görsellerin otomatik olarak tespit edilmesi sağlanabilir. Ayrıca, dijital kimlik doğrulama süreçlerinde sahte belgeler veya kimliklerin tespiti için bütünleşmiş çözümler sunulabilir. Bu teknolojinin, özellikle otomatik görüntü işleme ve güvenlik sistemlerinde de kullanılabileceği öngörülmektedir. Ayrıca, sanal gerçeklik ve artırılmış gerçeklik uygulamalarında, sanal ve gerçek dünyayı ayırt etmek için bu tür algoritmaların entegrasyonu, kullanıcı deneyimini iyileştirebilir.

Son olarak, bu çalışma, daha fazla görsel veri ve farklı tekniklerin entegrasyonu ile geliştirilerek, daha geniş veri setleri üzerinde daha fazla çeşitliliğe sahip sonuçlar elde edebilir. Modelin doğruluğu, farklı görsel türlerine ve kaynaklarına uyarlanarak, daha genel ve kapsamlı bir uygulama alanına taşınabilir. Bu sayede, güvenlik, medya doğruluğu ve dijital güvenlik gibi alanlarda önemli bir ilerleme sağlanabilir.

## 3.3. Projenin Katkıları ve Geliştirilebilir Alanlar

Bu projenin en önemli katkılarından biri, yapay zekâ tarafından üretilen görsellerin tespiti konusunda sağladığı güçlü ve güvenilir bir modelin geliştirilmesidir. Çalışma, görsel doğruluğunun ve güvenliğinin sağlanması noktasında büyük bir ihtiyacı karşılamaktadır. Dijital medyanın giderek daha fazla sahte içerik ve manipülasyon ile şekillendirildiği günümüzde, bu tür tespit yöntemlerinin etkinliği, dijital güvenliği ve doğruluğu artırma açısından kritik öneme sahiptir. Bu bağlamda, projede geliştirilen model, sadece akademik bir çalışma olarak kalmayıp, gerçek dünya uygulamalarında da büyük bir potansiyele sahiptir.

Proje, özellikle medya ve dijital güvenlik alanlarında önemli bir katkı sunmaktadır. Sahte görsellerin, videoların ve haberlerin hızla yayıldığı bir çağda, bu tür görsellerin doğru bir şekilde tespit edilmesi, toplumsal güvenin sağlanmasında önemli bir rol oynamaktadır. Medya manipülasyonu, sahte haberler ve kimlik sahteciliği gibi tehditler, yalnızca bireylerin güvenliğini tehlikeye atmakla kalmaz, aynı zamanda toplumsal huzursuzluklara da yol açabilir. Bu proje, sahte görselleri tespit etmek için kullanılan teknolojileri geliştirerek bu tehditlerin önüne geçilmesine yardımcı olabilecek önemli bir araç sunmaktadır. Özellikle medya sektöründe, görsel doğrulama ve güvenlik protokollerinin otomatikleştirilmesi, sahte haberlerin yayılmasının engellenmesi için büyük bir potansiyel taşımaktadır.

Ayrıca, projenin bir diğer katkısı, insan ve nesne etkileşimlerini göz önünde bulunduran daha sofistike bir modelin geliştirilmesidir. Görselin yalnızca gerçek mi sahte mi olduğu değil, aynı zamanda bu görselin içindeki etkileşimleri de doğru şekilde analiz etme kabiliyetine sahip olmak, bu tür bir modelin çok daha geniş bir uygulama alanına hitap etmesini sağlamaktadır. Örneğin, VR ve AR gibi alanlarda, bu modelin insan-nesne etkileşimlerini doğru bir şekilde modellemesi, daha gerçekçi ve güvenli sanal ortamların oluşturulmasına olanak tanıyacaktır.

Projenin geliştirilebilir alanlarına bakıldığında, modelin doğruluğu ve genelleme yeteneği üzerine yapılabilecek bir dizi geliştirme bulunmaktadır. İlk olarak, farklı veri setleri ve daha fazla çeşitlilik sunan görseller kullanılarak modelin eğitimi genişletilebilir. Mevcut model, belirli bir veri kümesine dayalı olarak eğitildiği için, yeni ve farklı veri setlerine karşı performansı sınırlı olabilir. Bu nedenle, modelin doğruluğunu artırmak için veri setinin çeşitlendirilmesi ve daha büyük çaplı, etiketlenmiş veri kümeleri ile modelin yeniden eğitilmesi gereklidir.

Bir diğer geliştirme alanı ise, modelin sahte görsel ve gerçek görsel arasında daha ayrıntılı bir ayrım yapabilme kapasitesine sahip olmasıdır. Şu anki model yalnızca basit bir sınıflandırma yapmaktadır; ancak, daha karmaşık sınıflandırmalar ve farklı sahtecilik türlerini tespit edebilecek bir model, daha etkili olacaktır. Örneğin, "deepfake" teknolojisiyle oluşturulmuş görsellerin daha iyi tespit edilebilmesi için özel olarak bu tür görseller üzerinde yapılan eğitimlerle modelin güçlendirilmesi sağlanabilir.

Proje kapsamında kullanılan derin öğrenme tekniklerinin daha da iyileştirilmesi ve optimize edilmesi de bir diğer geliştirme alanıdır. Özellikle eğitim sürecindeki erken durdurma (early stopping) tekniklerinin ve öğrenme oranı ayarlamalarının daha hassas bir şekilde yapılması, modelin daha hızlı ve doğru sonuçlar elde etmesini sağlayabilir. Ayrıca, modelin hızını artırarak daha verimli hale getirebilmek için farklı derin öğrenme teknikleri ve optimize edici algoritmalar üzerinde çalışmalar yapılabilir.

Son olarak, bu modelin daha fazla endüstri uygulamalarına entegrasyonu için, kullanıcı dostu web tabanlı platformların geliştirilmesi önemlidir. Kullanıcıların modelin sonuçlarını anlık olarak görebilmesi ve işlem yapabilmesi, bu teknolojinin daha geniş bir kitleye hitap etmesini sağlayacaktır. Web arayüzü üzerinden yapılan bu tür analizlerin, kullanıcı etkileşimi açısından daha verimli ve kullanıcı dostu bir şekilde sunulması, projenin etki alanını önemli ölçüde artıracaktır.

Sonuç olarak, bu proje, sahte görsellerin tespiti alanında önemli bir adım atmıştır ve birçok farklı sektörde kullanılabilecek potansiyele sahiptir. Bununla birlikte, modelin geliştirilmesi ve çeşitlendirilmesi, doğruluk oranlarının daha da artırılması ve yeni alanlarda uygulanabilirliğinin test edilmesi, projenin gelecekteki etkisini daha da güçlendirecektir.

# SONUÇLAR

Bu proje, yapay zekâ ve derin öğrenme teknikleri kullanarak sahte görsellerin tespit edilmesine yönelik geliştirilen bir modelin eğitim, değerlendirme ve uygulama süreçlerini kapsamaktadır. Eğitim süreci boyunca elde edilen sonuçlar, modelin sahte görselleri tespit etme yeteneğinin yüksek olduğunu ve derin öğrenme modellerinin görsel sınıflandırma alanında ne kadar etkili bir araç sunduğunu açıkça ortaya koymuştur. Eğitim sürecinde, modelin başlangıçtaki doğruluğu %77 civarlarında başlasa da zamanla bu değer %92'ye kadar yükselmiştir. Bu da modelin eğitimi sırasında veri setindeki desenleri öğrenme kapasitesinin arttığını ve zaman içinde daha doğru sınıflandırmalar yapabildiğini göstermektedir.

Eğitim ve doğrulama doğruluğunun yanı sıra kayıplar da dikkatle incelenmiştir. Eğitim kaybı başlangıçta 0.48'den 0.36'ya düşerken, doğrulama kaybı da aynı şekilde 0.50'den 0.39'a gerilemiştir. Bu, modelin eğitim süreci boyunca giderek daha etkili bir şekilde verileri analiz ettiğini ve öğrenme kapasitesinin arttığını göstermektedir. Ancak, doğrulama doğruluğunda gözlemlenen küçük dalgalanmalar, modelin aşırı uyum yapmaya başladığının ve eğitim verisinde yüksek performans gösterirken doğrulama verisinde daha az genel bir performans sergilediğinin bir göstergesidir. Bu tür dalgalanmalar, modelin sadece eğitim verisine değil, aynı zamanda genel veriye de uyum sağlaması için daha fazla optimizasyon gerektirdiğini gösteren önemli bir bulgudur.

Confusion matrix (karışıklık matrisi) sonuçları da modelin genel başarısını doğrulamaktadır. Yanlış pozitif (false positive) sınıflamalar 488, yanlış negatif (false negative) sınıflamalar ise 406 olarak hesaplanmıştır. Bu düşük sayılar, modelin çoğu zaman doğru sınıflandırmalar yaptığına işaret etmektedir. Precision ve recall değerleri sırasıyla %92 ve %93 olarak hesaplanmıştır. Precision değerinin yüksekliği, modelin sahte görselleri doğru bir şekilde tanıma kapasitesinin yüksek olduğunu, recall değerinin yüksekliği ise modelin doğru pozitif sınıflamaları tanıma konusunda etkili olduğunu göstermektedir. Bu başarılar, sahte görsellerin doğru tespit edilmesinin yanı sıra, gerçek görsellerin doğru şekilde ayırt edilmesini sağlamak adına da oldukça önemli bir bulgudur.

Eğitim sürecindeki bulgular, sadece doğruluk ve kayıp değerlerinin izlenmesiyle sınırlı kalmamış, aynı zamanda modelin performansını görselleştirerek eğitim ve doğrulama sürecinde elde edilen kayıpların ve doğrulukların nasıl geliştiği de görsel olarak incelenmiştir. Eğitim ve doğrulama doğruluğu ve kayıplarının grafiksel olarak analiz edilmesi, modelin iyileşme sürecini daha iyi anlamamıza olanak sağlamıştır. Eğitim doğruluğunun %92'ye kadar yükselmesi, modelin başarı seviyesini oldukça iyileştirdiğini gösterirken, doğrulama doğruluğundaki artışlar ve dalgalanmalar modelin aşırı uyum yapma eğilimlerini gözler önüne sermektedir. Bu bağlamda, modelin doğrulama verisine daha genel uyum sağlaması için ek geliştirmeler yapılması gerektiği aşikârdır. Bunun için eğitim süreci boyunca öğrenme oranı optimizasyonu, erken durdurma ve diğer hiperparametre ayarları gibi stratejiler uygulanarak modelin genel doğruluğu arttırılabilir.

Bu projenin bir diğer önemli katkısı, modelin performansını geliştirmek ve daha fazla genel veri üzerinde test etmek için kullanılabilecek bir altyapı sunmasıdır. Sahte görsellerin tespit edilmesi, medya sektöründe önemli bir yere sahiptir ve bu teknoloji sahte haberlerin, manipüle edilmiş fotoğrafların ve videoların yayılmasını engellemeye yönelik büyük bir adım olabilir. Ayrıca, bu tür teknolojilerin eğitim, güvenlik, hukuk ve dijital güvenlik alanlarında da önemli bir uygulama alanı bulacağı öngörülmektedir. Eğitim sürecinde elde edilen sonuçlar, modelin gelecekte daha büyük ve çeşitlendirilmiş veri setlerinde de başarılı olma potansiyeline sahip olduğunu göstermektedir. Sahte görsellerin tanınması, sadece görsel analizde değil, aynı zamanda video analizi, yüz tanıma sistemleri ve sosyal medya güvenliği gibi alanlarda da önemli bir gereksinim haline gelmiştir.

Sonuç olarak, bu proje sahte görsellerin tespit edilmesine yönelik geliştirilmiş derin öğrenme tabanlı modelin başarısını ortaya koymuş ve gelecekte bu alanda yapılacak çalışmalara öncülük etmiştir. Modelin başarısı, yalnızca eğitim sürecindeki kayıpların ve doğrulukların izlenmesiyle değil, aynı zamanda elde edilen performans metriklerinin sahte görselleri doğru bir şekilde tespit etmek için ne kadar etkili olduğunu da kanıtlamıştır. Gelecekte bu modelin doğruluğu daha da iyileştirilebilir ve genişletilmiş veri setlerinde daha fazla test edilerek sahte görsellerin tespiti konusunda daha yüksek başarı elde edilebilir.

# KAYNAKLAR

[1] Zhang, T., & Zhang, X. (2020). Detecting Deepfake Videos using Convolutional Neural Networks. Journal of Artificial Intelligence, 27(3), 45-58.

[2] Zhou, Y., & Wang, J. (2020). Deep Learning for Image Authentication: Detecting Fake Images in Digital Media. Journal of Artificial Intelligence and Data Science, 18(4), 112-121.

[3] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[4] Matern, J., & Anastasopoulos, A. (2021). A Comprehensive Review of Deep Learning Models for Fake Image Detection. IEEE Access, 9, 34978-34995.

[5] Suwajanakorn, S., Seitz, S. M., & Matusik, W. (2017). Synthesizing Obama: Learning Lip Sync from Audio. ACM Transactions on Graphics, 36(4), Article 121.

[6] Zhao, L., & Chen, X. (2023). Fake Face Detection with GANs: A Comparative Study and Future Directions. Computer Vision and Image Understanding, 214, 103284.

[7] Nguyen, T. T., & Yi, M. (2022). Fake Image Detection Using Convolutional Neural Networks: A Survey. Journal of Machine Learning and Applications, 15(1), 62-80.

[8] Zhang, H., & Lee, A. (2021). Deepfake Detection with Deep Learning: A Comprehensive Review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. Retrieved from

[9] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Identity Mappings in Deep Residual Networks. Proceedings of the European Conference on Computer Vision (ECCV).

[10] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. Proceedings of the International Conference on Learning Representations (ICLR).

[11] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems (NeurIPS).

[12] Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. Proceedings of the International Conference on Machine Learning (ICML).

[13] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely Connected Convolutional Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[14] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS).

[15] Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How Transferable Are Features in Deep Neural Networks? Advances in Neural Information Processing Systems (NeurIPS).

[16] PyTorch. (N.d.). ResNet50 — Torchvision 0.22 documentation. Advances in Neural Information Processing Systems (NeurIPS). Retrieved from

https://pytorch.org/vision/stable/models/generated/torchvision.models.quantization.resnet50.html

[17] Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. In Proceedings of the International Conference on Learning Representations (ICLR).