

2209/A

ÜNİVERSİTE ÖĞRENCİLERİ ARAŞTIRMA PROJELERİ DESTEK  
PROGRAMI

SONUÇ RAPORU

**PROJE BAŞLIĞI: Derin Öğrenme ile Deepfake Video ve Seslerin  
Tespitinin Yapılması**

PROJE YÜRÜTÜCÜSÜNÜN ADI: Emine ŞENER

DANIŞMANININ ADI: Erdem YAVUZ

# 1. Giriş

Günümüzde dijital teknolojilerin hızla gelişmesi ve yapay zekâ alanındaki ilerlemeler, içerik üretiminde yeni imkânlar sunduğu kadar önemli güvenlik sorunlarını da beraberinde getirmektedir. Bu sorunların başında, "deepfake" adı verilen manipüle edilmiş video ve ses içerikleri gelmektedir. Deepfake teknolojisi, kişilerin yüz ifadelerini ve seslerini gerçeğinden ayırt edilemeyecek şekilde değiştirerek yanıltıcı içerikler oluşturulmasına olanak tanımaktadır.

Bu proje, derin öğrenme teknikleri kullanılarak deepfake video ve seslerin tespitini sağlayan bir sistem geliştirmeyi amaçlamaktadır. Özellikle internet ve sosyal medya platformlarında yaygınlaşan bu sahte içeriklerin tespit edilmesi, kişisel hakların korunması, özel hayatın gizliliğinin sağlanması ve yanlış bilgi yayılımının engellenmesi açısından büyük önem taşımaktadır.

Projemizde, video tespiti için DFDC (Deepfake Detection Challenge) veri seti kullanılırken, ses tespiti için Türkçe dili odaklı özel bir veri seti oluşturulmuştur. Derin öğrenme modellerimiz, bu veri setleri üzerinde eğitilerek yüksek doğruluk oranlarıyla deepfake içerikleri tespit edebilmektedir. Geliştirdiğimiz web ve mobil uygulamalar sayesinde, kullanıcılar şüpheli içerikleri kolayca analiz edebilecek ve gerçek veya sahte olduğuna dair bilgi alabileceklerdir.

Bu çalışma, mevcut deepfake tespit yöntemlerinin eksikliklerini gidererek, özellikle Türkçe içeriklere odaklanan, kullanıcı dostu ve ücretsiz bir çözüm sunmayı hedeflemektedir. Böylece toplumsal güvenliğe katkı sağlamak, kişisel veya kurumsal imajların manipülasyonuna karşı koruma geliştirmek ve dijital okuryazarlığı artırmak amaçlanmaktadır.

## 2. Rapor Dönemlerinde Yapılan Çalışmalar

### 2.1. Veri Setlerinin Hazırlanması

#### 2.1.1. Görüntü Verilerinin Hazırlanması

Proje geliştirme sürecinin en önemli adımlarından biri olan veri seti hazırlama sürecine, görüntü verileriyle başlandı. Bu süreçte hedeflenen doğrultuda DFDC veri seti kullanıldı. Oldukça kapsamlı ve yüksek boyutlu bir veri seti olan DFDC'nin, üç araştırmacının da bilgisayarlarında bellek sorunu oluşturabileceği belirlendi. Bunun üzerine, 450 GB'lık veri setinin doğrudan bilgisayara indirilmesi yerine, bu süreci daha düşük maliyetle ancak oldukça verimli şekilde tamamlayan bir algoritma geliştirildi. (Brian Dolhansky, 2020)

DFDC veri seti, her biri yaklaşık 10 GB boyutunda olan 50 adet dosyadan oluşmaktadır. Bu nedenle veri seti tek seferde indirilmek yerine, dosyalar sırayla indirildi. İndirilen videolar içerisinden, veri setinin etiket dağılımı korunarak rastgele veriler seçildi ve ardından indirilen dosya silindi. Aynı işlem bir sonraki dosya için tekrarlandı. Sonuç olarak, %49.9 fake (sahte) ve %50.1 real (gerçek) olmak üzere toplam 16 GB boyutunda bir veri seti oluşturuldu.

DFDC veri setindeki videoların özel isimlere sahip olduğu, bu bilgilerin ise ayrı bir JSON formatındaki dosyada isim ve etiketleriyle birlikte tutulduğu tespit edildi. Derin öğrenme modelinde kullanılmak üzere, video adresleri ve etiketlerini içeren bir CSV dosyası oluşturuldu.

Yapılan incelemeler sonucunda, DFDC veri setindeki sahte videoların, yine aynı veri setindeki bazı orijinal videolardan üretildiği belirlendi. Deepfake tespitinde, sahte videonun hangi orijinal videodan oluşturulduğunun projeye nasıl bir katkı sağlayabileceği üzerine kapsamlı bir araştırma yapıldı. Bu doğrultuda, SSIM (Structural Similarity Index) tekniğinin kullanılabilirliği tespit edildi. Bunun için oluşturulan veri setindeki sahte kayıtlar için ek bir alan açılarak, ilgili videonun hangi orijinal video üzerinden oluşturulduğu bilgisi eklendi. (Abhinav K. Venkataramanan, 2021)

Veri seti hazırlandıktan sonra, oldukça basit ve yalnızca birkaç katmandan oluşan bir derin öğrenme modeli oluşturularak veriler üzerinde denemeler yapıldı. Modelin oldukça basit olmasına rağmen, eğitim sürecinin uzun sürdüğü gözlemlendi. Ayrıca, 16 GB'lık veri setiyle çalışmanın zorlukları da ortaya çıktı. Yapılan araştırmalar sonucunda, videoların uygun tekniklerle sıkıştırılarak boyutlarının azaltılabileceği belirlendi. Sıkıştırma işlemi nedeniyle videoların çözünürlüğü düşmesine rağmen, geliştirilecek derin öğrenme modelinin tespit performansında belirgin bir olumsuz etki gözlemlenmediği için proje sıkıştırılmış veri setiyle devam etti.

Hazırlanan veri seti, Bursa Teknik Üniversitesi Yüksek Hesaplamalı Yapay Zekâ Laboratuvarı sunucusuna başarılı bir şekilde yüklendi. Veri setinin yüklenmesinin ardından, geliştirilecek modelin verileri daha iyi anlamlandırabilmesi için veri ön işleme aşamasına geçildi.

### 2.1.1.1 Görüntü Verilerinin Ön İşlenmesi

DFDC veri setiyle elde edilen birçok gerçek (real) ve sahte (fake) video, model tarafından etkili bir şekilde işlenebilmesi için çeşitli ön işleme adımlarından geçirildi. Farklı teknikler denendi ve nihayetinde en verimli yöntemlerden oluşan bir ön işleme sürecine karar verildi.

Farklı görüntü işleme teknikleri; orijinal videolar, deepfake videolar veya her iki türü de içeren veri setlerine ihtiyaç duyulabilir. Bu nedenle, veri seti hazırlama sürecinde oluşturulan veriler kullanılarak öncelikle yalnızca orijinal videolardan oluşan bir veri seti ile deepfake videoları ve bunların oluşturulduğu orijinal videoları içeren ek bir veri seti elde edildi. Bu verilere hızlı ve doğru bir şekilde erişimi sağlamak amacıyla üç farklı fonksiyon geliştirildi.

#### 2.1.1.1.1 Videoların Çerçvelere Ayrılması

Birçok görüntü işleme modelinde olduğu gibi, projede geliştirilmesi planlanan CNN modeli video yerine yalnızca görüntü kabul etmektedir. Bu nedenle, projenin bu aşamasında herhangi bir bilgi kaybı olmaksızın tüm videoların görüntü verilerine dönüştürülmesi hedeflendi. Her bir video, OpenCV kullanılarak okundu ve her bir görüntü çerçevesi uygun renk dönüşümü ve ölçeklendirme tekniklerinden geçirildi. İşlemin sonucunda oluşturulan çerçeveler, ait oldukları video ve etiketleriyle birlikte PyTorch Dataset formatında kaydedildi.

Oluşturulan veri seti, uygun bir algoritma ile incelendi ve videodan art arda çıkarılan birkaç çerçevenin birbirine tamamen aynı olduğu tespit edildi. Bu nedenle, modelin aynı verilerle birden fazla kez eğitilmemesi gerektiğine karar verildi. Başlangıçta, her dört çerçeveden birini seçen bir algoritma yazıldı, ancak yine de benzer çerçevelerin kaldığı gözlemlendi. Sırasıyla 4, 16, 32, 64 ve 128 çerçeveyle yapılan denemelerde, 32 çerçeve üzerinden bir çerçeve seçilerek yeni bir veri seti oluşturuldu.

#### 2.1.1.1.2 Görüntü Verilerinden Yüz Verilerinin Tespit Edilmesi

Proje yönteminde, Deepfake videolarının tespit edilmesinde videodaki yüzlere odaklanılacağı belirtilmişti. Bu aşamada, görüntülerden yüz çıkarma işlemini yapan MTCNN modeli kullanıldı. MTCNN, temelde üç ayrı ağdan oluşan bir konvolüsyonel modeldir. İlk ağ, verilen görüntüyü tarar ve görüntüdeki potansiyel yüz bölgelerini işaretler. Son ağ ise işaretlenen bu bölgelerden burun, göz, ağız gibi yüz bölgelerini çıkarır. (Edgar Kaziakhmedov, 2020)

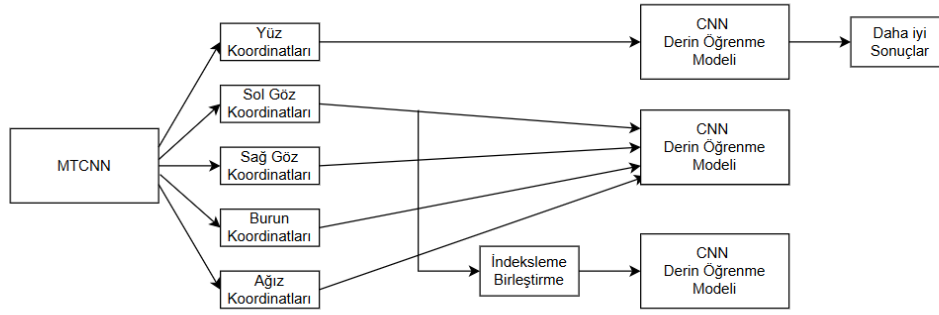
MTCNN, verilen çerçevedeki yüz koordinatlarını ve yüz bölgelerinin koordinatlarını belirler. Şekilde özetlendiği üzere, her bir çerçeve için aşağıdaki süreç izlendi ve yüz koordinatları ile göz, ağız, burun koordinatları çıkarılarak JSON dosyasına kaydedildi.

### 2.1.1.1.3 Yüz Verilerinin Ayrıştırılması

Proje kapsamında, Deepfake tespitine odaklanılacağı için videolardan ayrıştırılan görüntülerden yüz verilerinin çıkarılması hedeflendi. Böylelikle, geliştirilecek olan modelin öğrenme sürecinde dış etkenlerden ziyade yüzlere odaklanması sağlanabilir. Bu amaçla, MTCNN kullanılarak tespit edilen koordinatlar üzerinden işlem yapıldı. Geliştirilen ilk algoritma, MTCNN modelinin ürettiği, yüzü çevreleyen kutu bilgilerini kullanarak çerçevelerden yüzleri tespit etmektedir. Bu algoritma ile çerçevelerden yüz verileri ayrıştırıldı ve çıkartılan yüz verileri, geliştirilen temel bir CNN modeliyle değerlendirildi.

MTCNN modelinin çıktuları kullanılarak geliştirilen bir diğer algoritma ise, göz, burun ve ağız koordinatlarını çerçeveden çıkartıp uygun bir formatta kaydetmektedir. Bu süreçte, bir yüzden çıkarılan 4 farklı resmin birlikte ya da ayrı veri olarak verilmesi konusu tartışıldı ve temel bir CNN modeliyle denemeler yapıldı. Deneme sonuçları değerlendirildi ve sonuçlar raporlandı.

Bu aşamada, deney yapılabilmesi için hızlıca geliştirilen temel CNN modeliyle aşağıda özetlenen üç farklı deney gerçekleştirildi. Deney sonuçlarına göre, yüz verilerinin ayrıştırılmasına karar verildi.



Şekil 1: CNN Model Deneyleri

### 2.1.1.1.4 Yüz Vektörleştirme

Bu aşamada, görüntü verilerinden sayısal veriler elde edilmiştir. face\_recognition kütüphanesi kullanılarak, her 10 çerçeveden biri 128 bit uzunluğundaki verilere dönüştürülmüştür. Orijinal verilere ait vektörlerde bir tutarlılık söz konusudur; aynı kişiye ait vektörler birbirine benzer. Ancak, özellikle deepfake videolarda, deepfake yapılan kısımlarda vektörlerde değişiklikler gözlemlenmiştir. Model eğitiminde kullanılmak üzere oluşturulan vektörler, video indeksleriyle birlikte uygun formatta kaydedilmiştir. (Xinyi Wang, 2022)

### 2.1.1.1.5 Fark Haritalarının Belirlenmesi

Bu aşamada, çerçeveler arasındaki SSIM (Yapısal Benzerlik Endeksi) ölçülerek, deepfake ve orijinal video arasındaki fark haritası çıkartılmıştır. Fark haritaları, temel olarak modele hangi farkların deepfake görüntüsüne işaret ettiğini belirlemek için kullanılır. Model, deepfake ve gerçek veriler arasındaki ilişkiyi öğrenerek analiz kabiliyetini güçlendirmektedir.

### **2.1.1.1.6 Veri Seti Artırma Teknikleri**

DFDC veri seti, 50 GB büyüklüğünde bir veri setiyken, veri seti dağılımı korunarak 16 GB boyutunda bir veri seti elde edilmiştir. Buna rağmen, derin öğrenme modelinin eğitileceği veriler üzerinde veri artırma tekniklerinin uygulanması gerekmektedir. Veri artırma tekniklerinin uygulanmasının nedeni, modelin farklı büyüklükteki, çözünürlükteki ve yapıdaki veriler üzerinde de öğrenme sağlaması ve böylelikle farklı formdaki veriler üzerine de doğru sınıflandırma yapabilmesidir. 16 GB boyutundaki veri seti, teknikler ile değil de DFDC veri setinden alınacak veriler ile zenginleştirildiğinde yine diğerlerine benzeyen formlarda veriler alınacağı için bu tekniklerin uygulanmasına karar verilmiştir.

#### **2.1.1.1.6.1. Deepfake Görüntülerinin İncelenmesi ve Algoritmaların Geliştirilmesi**

Bu aşamada, öncelikle deepfake görüntülerinin nasıl oluşturulduğu incelenmiştir. Birçok deepfake videonun özellikle göz, ağız ve burun üzerinde değiştirme yapılarak üretildiği tespit edilmiştir. DFDC verilerini ön işlemden geçirerek elde edilen veriler incelendiğinde de en göze çarpan kısımların göz, burun ve ağız olduğu görülmüştür. Bu veriler modele verildiğinde model, deepfake video tespiti yaparken yalnızca bu kısımlara odaklanacak ve gelişmiş deepfake algoritmalarına karşı doğru sonuçlar üretmeyecektir. Bu sebeple modelin göz, ağız ve burun dışında diğer bölümlere de odaklanmasını sağlamak için çerçeveden burun silme, göz silme ve ağız silme gibi algoritmalar geliştirilmiştir. Bu algoritmalar, MTCNN tarafından üretilen verileri kullanarak yeni veriler üretir. Veri üretme işlemi, var olan çerçeveden belirli bir koordinattaki yüz bölgesi yerine siyah parçalar eklenmesinden ibarettir. Böylelikle 3 farklı algoritma ile veri artırma işlemi tamamlanmıştır. Bu sayede model, hem göz, ağız ve buruna odaklanabilen hem de diğer bölgelere de odaklanarak, daha gelişmiş deepfake algoritmalarıyla üretilen videoların da tespitini yapabilecektir. (Shruti Agarwal (1), 2020)

#### **2.1.1.1.6.2. Padding ve Maskelenmiş Veri Artırma**

Modelin belirli bir çerçeve üzerine öğrenme yapmaması için çerçeveler üzerinde padding değiştirmek şeklinde bir veri artırma işlemi daha yapılmıştır. Görüntü kenarlarına uygulanan farklı padding miktarlarıyla modelin anomalileri tespit etmesi kolaylaştırılmıştır.

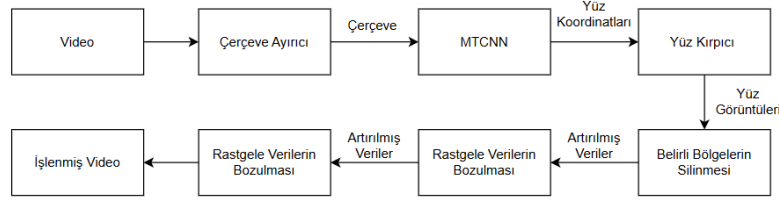
Deepfake algoritmalarının incelenmesiyle tespit edilen bir diğer durum ise, bu algoritmaların yüzün belirli bir kısmına odaklanmasıdır. Belirli bir kısmı deepfake olan verilerle eğitilen model, sınıflandırma için kullanıldığında da yine aynı şekilde belirli bir kısma odaklanabilir. Bunu engellemek için görüntülerin rastgele kısımları maskelenerek veri artırma işlemi gerçekleştirilmiştir. (Yuezun Li, 2018)

#### **2.1.1.1.6.3. Sıkıştırma ve Bulanıklaştırma**

Yapılan araştırmalar sonucu, deepfake videolarının sıkıştırma ve bulanıklaştırma içerdiği tespit edilmiştir. Bu da eğitilen modelin doğru analizi yapmasını oldukça zorlaştırmaktadır. Modelin bu tür verilere dayanıklı hale gelmesi için eğitim verileri üzerinde yeniden ölçeklendirme ile yüz bozma, bulanıklaştırma ve sıkıştırma işlemleri yapılmıştır.

### 2.1.1.1.7. Sonular ve n İleme Sreci

Tm bu deneylerin sonuları ve deney sreleri kaydedilmiřtir. Yapılan denemeler sonucunda ařağıdaki n iřleme sreci belirlenmiř ve tm verilerin iřlenmesi saėlanmıřtır.



řekil 2: Grnt Verileri n İřleme Sreci

### 2.1.2. Ses Verilerinin Hazırlanması

Deepfake ses tespiti yapabilen bir model geliřtirmek iin ncelikle deepfake ve gerek ses verilerini ieren bir veri seti elde etmek iin arařtırmalar yapıldı ve ASVspooof 2021 veri seti ile ilerlemeye karar verildi.

ASVspooof veri seti, FLAC formatında ses verileri ve deepfake/spooof olmak zere etiket bilgilerini ieren bir metadata dosyası bulundurmaktadır. Veri seti ieriėinde 3 farklı trde dosya bulunmaktadır: LA (Logical Address), PA (Physical Address) ve DF (Deepfake). Bu dosyalar zerinde yapılan kapsamlı arařtırmalar sonucu, LA dosyasında dijital olarak retilmiř veya dnřtrlmř, TTS (Text-to-Speech), VC (Voice Conversion) gibi yntemler kullanılarak oluřturulmuř seslerin bulunduėu, PA (Physical Access) dosyasındaysa dijital cihazlar yoluyla kaydedilen sahte seslerin bulunduėu tespit edilmiřtir. DF dosyasında ise Transformer, GAN tabanlı yntemler kullanılarak oluřturulmuř sahte seslerin bulunduėu belirlenmiřtir. PA ve LA dosyalarında farklı dillerde ses verileri bulunduėu iin bu dosyaların da eėitim setine dahil edilmesine karar verilmiřse de, deepfake algoritmaları kullanmadan oluřturulmuř verilerin modele dahil edilmesinin modelin derin ėrenme tespiti kabiliyetine bir katkı saėlamayacaėına karar verilerek yalnızca DF dosyası ile devam edilmiřtir. DF dosyası indirilmiř ve alıřma ortamına yklenmiřtir. Dosyanın ieriėi kontrol edildiėinde, etiket verilerinin diėer bilinen veri setlerinden oldukça farklı bir formatta olduėu gzlemlenmiřtir. Bunun zerine yazılan algoritma ile ses verileri ve etiketleri eėitim, doėrulama ve test iřlemlerinde efektif bir řekilde kullanılmak zere CSV dosyası haline getirilmiřtir. (Xuechen Liu, 2022)

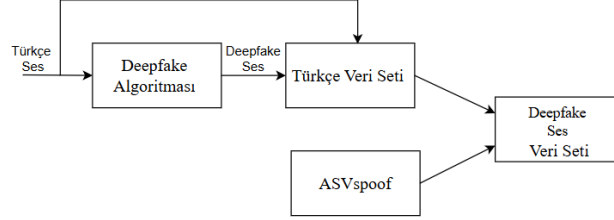
#### 2.1.2.1 Trke Verilerin Elde Edilmesi

Bu ařamada, Trke veri seti zerine arařtırmalar yapılmıřtır. Trke veri seti iin internet ortamında proje amacına ynelik uygun bir veri seti bulunmamıřtır. Bunun zerine ses verileri toplanılmaya bařlanmıřtır. Bu ařamada, ncelikle 3 arařtırmacı ve diėer gnll katılımcılardan eřitli yapılar da metinleri okuyup, sesi kaydetmeleri istenmiřtir.

Kaydedilen sesler, eřitli aık kaynak uygulamaları kullanılarak maniple edilmiř ve kaydedilen orijinal seslerden eřitli formlarda deepfake sesler elde edilmiřtir. Bunun zerine veri toplama iřlemleri zerine alıřmalara devam edilmiřtir.

Ancak, bu aşamada veri toplama işlemi her ne kadar gönüllülerin onayı ve bilinciyle gerçekleşmiş olsa da, verilerin korunması konusunda bir sıkıntı olma tehlikesine karşılık veri toplama işlemi durdurulmuştur. Yalnızca 3 araştırmacıya ait veriler kaydedilmiştir.

Devamında yapılan araştırmalar ve tartışmalar sonucu, deepfake ses tespiti olayında kullanılacak MFCC tekniğinin sesin frekansına odaklanacağına ve böylelikle dilin deepfake tespiti için görmezden gelinebilir bir anlam ifade ettiğine karar verilmiştir. Böylelikle bu aşamada yalnızca 3 araştırmacıya ait Türkçe ses verileri, indirilen ASVspoof verileriyle dahil edilmiştir. Veri seti bu şekilde hazır hale edilmiştir. Veri seti bu şekilde hazır hale getirilmiştir.



Şekil 3: Türkçe Verilerin Veri Setine Dahil Edilmesi

### 2.1.2.2 Ses Verilerinin Ön İşlenmesi

Bu aşamada, MFCC (Mel-Frequency Cepstral Coefficients) tekniği kullanılarak ses verilerinden özellik çıkarılması sağlanmıştır. Yapılan araştırmaların sonucunda, MFCC yönteminin ses işleme alanında yaygın olarak kullanılan bir öznelik çıkarım yöntemi olduğu ve özellikle ses sınıflandırma gibi görevlerde oldukça başarılı performans gösterdiği belirlenmiştir. MFCC, ses verisini birçok adımdan geçirerek derin öğrenme modeli için uygun hale getirmektedir. (Yang Gao, 2021)

#### 2.1.2.2.1 STFT ile Spektrogram Oluşturma

Deepfake sesler genellikle belirli frekans bölgelerinde doğal olmayan desenler içerdiği için, ses sinyali küçük zaman dilimlerine bölündü. Bölünen her bir parça, Fourier dönüşümü ile zaman-frekans uzayına dönüştürüldü. Fourier dönüşümü sayesinde, sentetik sesin düzensiz frekans bileşenleri ortaya çıkarıldı.

#### 2.1.2.2.2 Güç Spektrogramı

STFT ile oluşturulan spektrogramın karesi alındı ve çerçeve uzunluğuna bölündü. Böylece sinyalin güç seviyesi hesaplandı ve hangi frekans bileşenlerinin daha güçlü olduğu gözlemlendi. Elde edilen grafiklerde, gerçek seslerin daha yumuşak geçişlere sahipken, deepfake ile oluşturulmuş seslerin keskin geçişler gösterdiği gözlemlendi.

#### 2.1.2.2.3 Mel Filtre Bankası Uygulama

Yapılan araştırmalar sonucu, Deepfake seslerin genellikle yüksek frekans bileşenlerinde doğal olmayan değişimler gösterdiği saptandı. Bunun için ses verilerine bir mel filtresi uygulandı ve veriler mel ölçeğine uygun hale getirildi. Filtre sayesinde düşük frekansların daha ayrıntılı işlenmesi sağlanarak, deepfake seslerdeki doğal olmayan bileşenler belirgin hale getirildi.

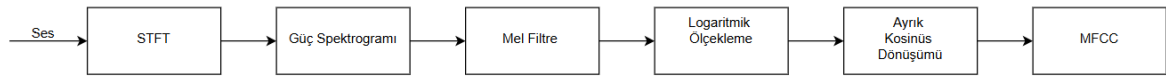
#### 2.1.2.2.4 Logaritmik Ölçekleme

İncelenen Deepfake seslerde, belirli frekans aralıklarında ani değişimler gözlemlendiği için, bu ani değişimlerin kolayca tespit edilebilmesi amacıyla ses verilerinin gücü logaritmik olarak ölçeklendirildi.

#### 2.1.2.2.5 DCT (Ayrık Kosinüs Dönüşümü) ile MFCC Çıkarma

Önceki aşamalarda elde edilen spektrogramdaki en önemli öznitelikler seçilerek bir öznitelik vektörü oluşturuldu. Böylelikle derin öğrenme modeli için anlamlı öznitelikler çıkarılmış oldu. (Jun Xue, 2022)

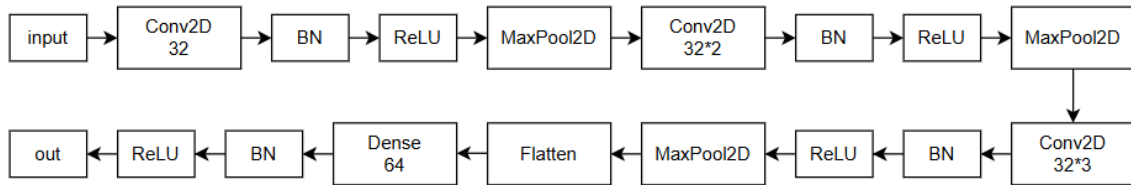
Tüm bu adımlar sırasıyla birbiri ardına eklenerek aşağıda özetlenen bir özellik çıkarım fonksiyonu elde edildi.



Şekil 4: MFCC ile Ses Verilerinden Özellik Çıkarmı

## 2.2. Model Eğitimi

Bu aşamada, CNN mimarisi temel alınarak biri video üzerine diğeri de ses üzerine yoğunlaşmış iki farklı model eğitimi gerçekleştirildi. Her bir model için temel CNN yapısında şekilde görülen mimaride bir derin öğrenme modeli tanımlandı. Temel CNN yapısıyla tasarlanan bu modelin video ve ses deepfake tespiti konusunda yetersiz kaldığı gözlemlendiği için, model eğitim süreçlerinde önceden eğitilmiş ve daha gelişmiş olan modellerin transfer learning tekniğiyle yeniden eğitilmesine karar verildi.



Şekil 5:Deneyler için Geliştirilen CNN Mimarisi

## 2.2.1. Video Tespit Modelinin Geliştirilmesi

Konvolüsyon ve Pooling katmanlarının birbiri ardına eklenerek kurulan bir CNN mimarisinin, hedeflenen karmaşık amaç için oldukça yetersiz kaldığı tespit edilince, projede belirtilen yöntemlerden sapmadan sorunu çözebilecek bir yöntem için kapsamlı araştırmalar yapıldı. Yapılan kapsamlı araştırmalar sonucunda, Deepfake video tespitinde CNN tabanlı EfficientNet – B7 modelinin kullanılmasına karar verildi. EfficientNet, konvolüsyon katmanlarının genişlik, derinlik ve çözünürlük olmak üzere katmanların yeniden ölçeklendirilmesiyle geliştirilmiş bir model olup, Deepfake tespiti gibi karmaşık görüntü analizi uygulamalarının kapsamlı özellik çıkarımı ve bu özelliklerin verimli bir şekilde işlenmesi ihtiyacını, sahip olduğu ölçeklendirilmiş mimariyle karşılamaktadır. ( *Davide Cocomini, 2022*)

### 2.2.1.1 Model Mimarisinin Oluşturulması

EfficientNet model ailesinde, EfficientNet B2 modelinden EfficientNet B7 modeline kadar toplam 6 farklı model bulunmaktadır. Bu modeller, tespit ettikleri özellik sayısına göre birbirinden ayrılmaktadır. B2 modeli 1408 farklı özelliği tespit edebilirken, B7 modeli 2560 farklı özelliği tespit edebilmektedir. Deepfake tespitinde, verilen görüntüden olabildiğince fazla özelliğin çıkarılması gerektiğinden, bu aşamada EfficientNet B7 modelinin kullanılması tercih edilmiştir. (*Mingxing Tan, 2020*)

EfficientNet B7 modeli, ImageNet veri kümesiyle eğitilmiş olup, temel haliyle deepfake tespiti yapamayan bir modeldir. Transfer öğrenme yöntemi uygulanmadan önce, modelin deepfake tespiti yapabilir hale getirilmesi için çeşitli denemeler ve araştırmalar gerçekleştirilmiştir. EfficientNet B7 modeli, timm kütüphanesinden alınmış ve önceden eğitilmiş haliyle kullanılmıştır. Deepfake tespitine uygun hale getirilmesi amacıyla, öncelikle son katmanları çıkarılmış ve yalnızca EfficientNet B7 modelinin encoder yapısı mimariye dahil edilmiştir.

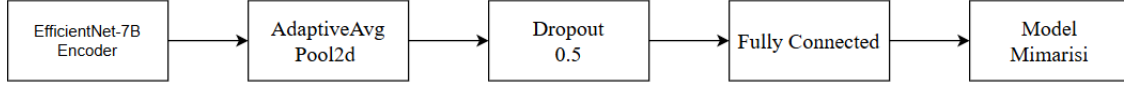
CNN modellerinde boyut küçültme, gereksiz detayları eleme, önemli özellikleri koruma ve overfitting'i önleme amacıyla Pooling katmanları üzerinde çalışmalar gerçekleştirilmiştir. İlk aşamada, model mimarisinde AvgPool2d katmanı kullanılarak bir deneme yapılmıştır.

AvgPool2d katmanında kernel (filtre) boyutu ve stride (adım) değeri model tasarım sürecinde sabit olarak belirlenmektedir. Bu nedenle, katmanın çıkış boyutu giriş boyutuna bağlı olarak değişmekte ve sabit kernel ile çalışmaktadır. Ancak, bu durum projenin istenen esnekliğe ulaşmasını engellediğinden, yapılan araştırmalar sonucunda AdaptiveAvgPool2d katmanının kullanılması tercih edilmiştir.

AdaptiveAvgPool2d sayesinde kernel ve stride değerleri otomatik olarak hesaplanmakta, böylece giriş boyutu değişse dahi çıkış boyutu model mimarisinde belirlenen değere bağlı kalmaktadır. Bu yapı, modelin farklı giriş boyutlarına uyum sağlamasına olanak tanıyarak esneklik kazandırmıştır.

Veri seti hazırlama aşamasında, verilerin aşırı öğrenmeye oldukça müsait olduğu tespit edilmiş ve bu durumu önlemek amacıyla modele 0.5 oranına sahip bir Dropout katmanı eklenmiştir.

EfficientNet B7 modeli, 2560 özellik ile çalışan bir modeldir. Bu modelin, deepfake ya da gerçek olmak üzere ikili sınıflandırma yapabilmesi için son katmanın tek bir nörondan oluşması gerekmektedir. Bu nedenle, son olarak sinir ağına Fully Connected katmanı eklenmiştir.



Şekil 6: Deepfake Video Tespiti Model Mimarisi

### 2.2.1.2 Scheduler(Öğrenme Oranı Planlayıcı) Oluşturulması

Proje için gerekli olan deepfake video tespiti yapan modelin eğitimi açısından en önemli kavramlardan biri de öğrenme oranı (learning rate) belirlemektir. Bu aşamada ulaşılmaması gereken iki farklı hedef bulunmaktadır: yüksek değerli bir öğrenme oranı ile hızlı bir eğitim süreci ve düşük değerli bir öğrenme oranı ile daha kararlı bir eğitim süreci. Birbiriyle çakışan bu iki hedef doğrultusunda, başlangıç aşamasında küçük bir öğrenme oranı seçilerek kararlı ve iyi sonuçlar üreten bir model elde edilmesi amaçlanmıştır.

Düşük değerli bir öğrenme oranının, EfficientNet-B7 karmaşıklığında bir model ve DFDC gibi kapsamlı bir eğitim seti ile başlatılan eğitimin oldukça uzun bir süreç gerektirdiği gözlemlenmiştir. Eğitim süreci devam ederken yapılan araştırmalar sonucunda, model eğitiminin başlangıç aşamalarında düşük değerli bir öğrenme oranının model için pek anlam ifade etmediği, ancak eğitim ilerledikçe bu orana daha fazla ihtiyaç duyulduğu tespit edilmiştir. Bunun sonucunda, eğitim sürecinde değişen bir öğrenme oranı için bir zamanlayıcı oluşturulmuştur.

LRStepScheduler (Adımlı LR Zamanlayıcısı), PolyLR (Polinomsal LR Zamanlayıcısı) ve ExponentialLR Scheduler (Üstel LR Zamanlayıcısı) olmak üzere üç farklı zamanlayıcı ile denemeler gerçekleştirilmiştir. Polinomsal LR Zamanlayıcısının iterasyon ilerledikçe öğrenme oranını kademeli olarak azalttığı, Üstel LR Zamanlayıcısının ise öğrenme oranını her epoch'ta belirli bir çarpan (gamma) ile düşürdüğü gözlemlenmiştir. Bu süreçte, diğerlerine kıyasla daha yönetilebilir bir seçenek olan Adımlı LR Zamanlayıcısı tercih edilerek, belirtilen epoch'lara bağlı olarak öğrenme oranının değiştirilmesi sağlanmıştır. (Defazio, 2024)

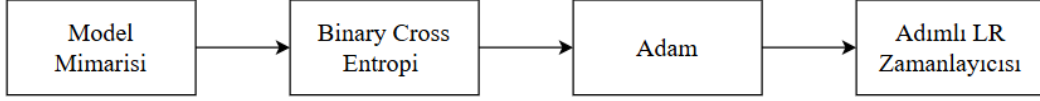
### 2.2.1.3 Kayıp Fonksiyonlarının Belirlenmesi

Model mimarisi belirlendikten sonra, eğitim sürecine hazırlık aşamasında ilk olarak kayıp fonksiyonları (loss) üzerine araştırmalar yapılmıştır. İkili sınıflandırma modelleri için en uygun kayıp fonksiyonunun, PyTorch kütüphanesi içerisinde bulunan Binary Cross Entropy fonksiyonu olduğu tespit edilerek bu fonksiyonun kullanılması kararlaştırılmıştır.

#### 2.2.1.4 Optimizasyon Algoritmasının Belirlenmesi

Kayıp fonksiyonu belirlendikten sonra, modelin ağırlıklarını kayıp fonksiyonunun belirlediği değere göre doğru bir şekilde güncelleyebilmesi için bir optimizasyon algoritmasına ihtiyaç duyulmuştur. PyTorch içerisinde birçok optimizasyon algoritması bulunmaktadır.

SGD, Adam ve RMSprop algoritmaları ile yapılan denemeler sonucunda en etkili sonucun Adam optimizasyon algoritması ile elde edildiği tespit edilerek, eğitim sürecine Adam fonksiyonu ile devam edilmiştir.



Şekil 7:Eğitim Konfigürasyonları

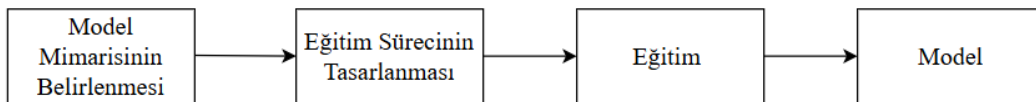
#### 2.2.1.5 Transfer Öğrenme Yoluyla Modelin Eğitilmesi

Derin öğrenme ile deepfake video tespiti yapmak için yalnızca CNN tabanlı bir model geliştirmek yetersiz kalacağından, önceden eğitilmiş CNN tabanlı EfficientNet B7 modeli transfer öğrenme ile yeniden eğitilerek proje amacına uygun hale getirilmiştir. Bu doğrultuda model mimarisi oluşturulmuş ve eğitim için gerekli ayarlamalar yapılmıştır. Bu aşamada ise eğitim süreci başlatılmıştır.

Eğitim sürecini başlatmadan önce, veri seti hazırlama aşamasında oluşturulan veri seti; eğitim, doğrulama ve test olmak üzere üç bölüme ayrılmıştır. Eğitim sürecinde doğrulama ve eğitim verileri kullanılmıştır.

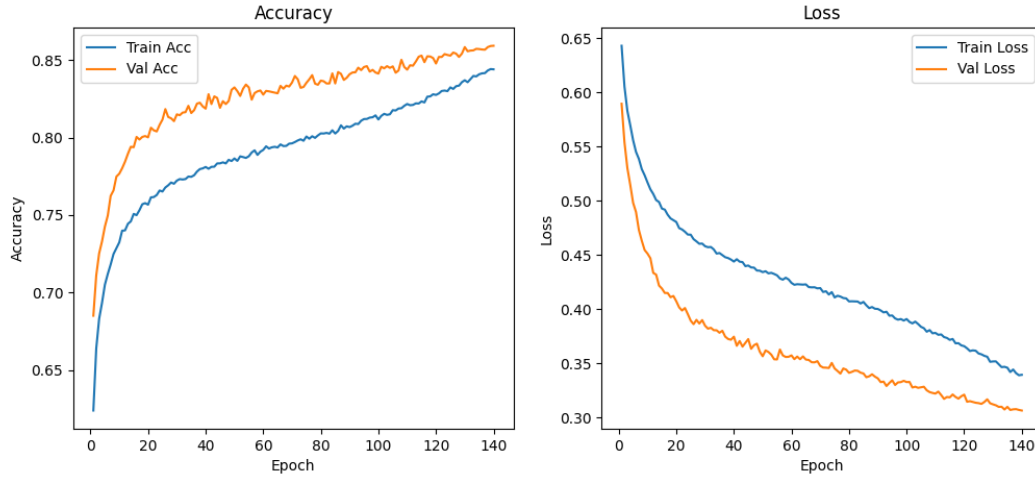
Modelin daha hızlı ve verimli çalışması için FP16 desteği eklenmiş, böylece yarım hassasiyet ile daha hızlı eğitim sağlanırken istenen doğruluk korunmuştur.

Eğitim sürecinin başlangıçta 150 epoch olarak belirlenmesine rağmen, bu değer tahmini bir değer olduğu için gereksiz epoch'ları önlemek adına EarlyStopping mekanizması eklenmiştir. Ayrıca, eğitim sürecinde kötüleşme yaşanması durumunda en iyi modele erişebilmek için ModelCheckpoint tekniği kullanılmıştır.



Şekil 8:Deepfake Tespit Modelinin Geliştirme Süreci

Tasarlanan eğitim sürecinin tamamlanmasıyla aşağıdaki grafikler elde edilerek, deepfake video tespitinde kullanılabilir bir model başarıyla geliştirilmiş oldu.



Şekil 9: Eğitim Grafikleri

## 2.2.2 Ses Tespit Modelinin Geliştirilmesi

Deepfake videoların tespit edilmesi için oldukça karmaşık bir model olan EfficientNet-B7, transfer learning yöntemiyle özelleştirilmiştir. Projede temel amaç, hem ses modelinin hem de video modelinin bir web/mobil ortamında birlikte çalışarak nihai sonucu üretmesidir. Bu nedenle, video modeli oldukça karmaşık olduğu için ses modelinde herhangi bir hazır model kullanılmayarak karmaşıklığın dengelenmesi amaçlanmıştır.

Yapılan kapsamlı araştırmalar sonucunda, ses modeli için LCNN tabanlı bir model geliştirilmesine karar verilmiştir. LCNN, ağırlıkları bir sözlük ve katsayılarla ifade ederek hem modelin hafıza gereksinimini azalttığı hem de konvolüsyon işlemlerini hızlandırdığı için bu problem için uygun bir model olarak seçilmiştir. (Jiangyan Yi, 2023)

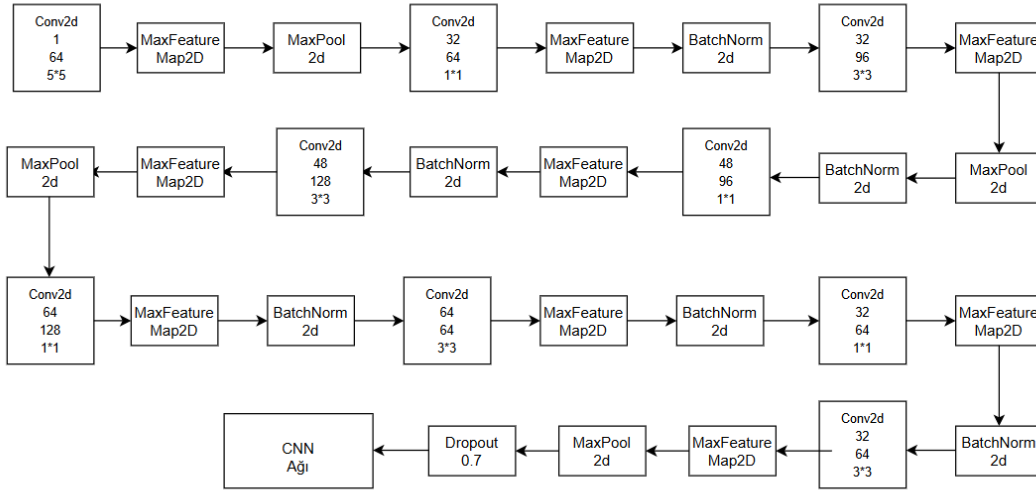
### 2.2.2.1 Model Mimarisinin Oluşturulması

Bu aşamada, karmaşık bir hazır model kullanılmadığı için istenen hedefe ulaşabilmek adına farklı katmanlarla kapsamlı bir ağ oluşturuldu. Güçlü bir temsil gücü elde etmek için 5 bloktan oluşan bir CNN sinir ağı tasarlandı.

İlk blok, temel özellik çıkarımını sağlamak amacıyla  $5 \times 5$  çekirdek boyutuna ve 64 filtreye sahip bir Conv2D katmanı ile başlar. Elde edilen özellik haritasındaki en güçlü aktivasyonları seçebilmek için MaxFeatureMap2D katmanı eklenmiş ve bunu boyut indirgeme sağlayan MaxPool2D katmanı takip etmiştir.

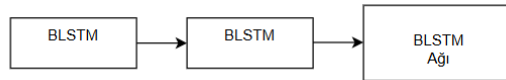
Daha derin özellikler öğrenebilmek için Conv2D, MaxFeatureMap2D, BatchNorm2D, Conv2D, MaxFeatureMap2D, MaxPool2D ve BatchNorm2D katmanlarından oluşan ek bir blok tasarlanmıştır. Conv2D katmanlarında çekirdek boyutu ve kanal sayısı farklılık göstermektedir. Modelin hızlı ve dengeli öğrenmesini sağlamak için BatchNorm2D katmanları da eklenmiştir.

Ayrıca, aşırı öğrenmeyi önlemek amacıyla son katmana 0.7 dropout oranına sahip bir Dropout katmanı eklenmiştir. Böylece aşağıda özetlenen şekilde kapsamlı ve dengeli bir ağ tasarlanmıştır.



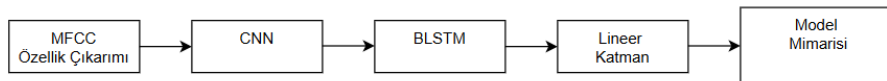
Şekil 10: Deepfake Ses Tespit Model Mimarisi CNN Katmanları

CNN katmanlarından sonra zaman serisi özelliklerini işlemek amacıyla çift yönlü Uzun Kısa Süreli Bellek (BLSTM) katmanları eklenmiştir. Bu yapıda, ardışık olarak iki adet BLSTM katmanı tanımlanarak modelin zaman bağımlı desenleri daha iyi öğrenmesi sağlanmıştır.



Şekil 11 : Deepfake Ses Tespit Model Mimarisi BLSTM Katmanları

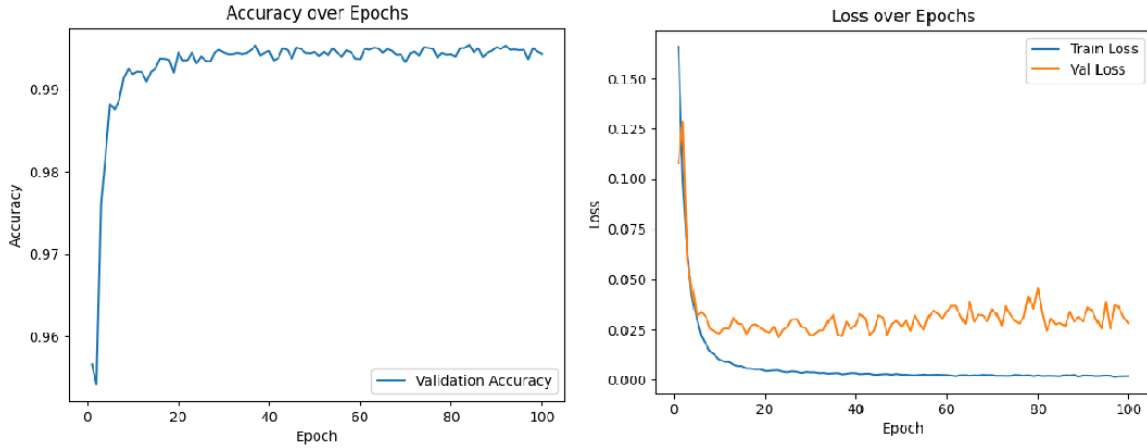
Ara katmanlar tanımlandıktan sonra, ön işleme adımında oluşturulan MFCC ile özellik çıkarımı yapan fonksiyonun çalıştığı ön işleme katmanı, CNN ağı, BLSTM ağı ve özelliklerin tek bir değere dönüşmesini sağlayan Lineer katman sırasıyla eklenerek model mimarisi oluşturulmuştur.



Şekil 12: Deepfake Ses Tespit Model Mimarisi

### 2.2.2.2 Model Eğitimi

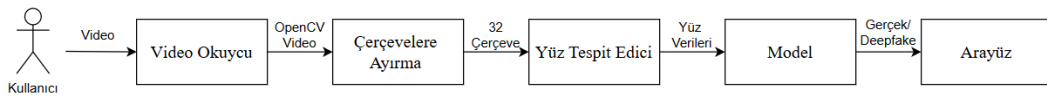
Oluşturulan mimari, video tespit modelinin geliştirilme aşamasında belirlenen öğrenme oranı planlayıcısı, optimizasyon algoritması ve kayıp fonksiyonu ile birlikte 150 epoch boyunca eğitilmiştir. Aşağıda eğitim sürecini temsil eden grafikler gösterilmektedir.



Şekil 13: Ses Modeli Eğitim Grafikleri

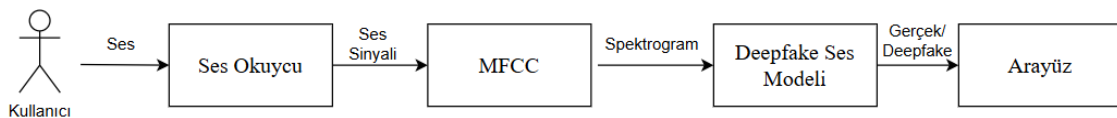
## 2.3. Web ve Mobil Uygulama Geliştirilmesi

Bu süreçte geliştirilen modellerin kullanılması için algoritmalar oluşturulmuştur. Deepfake video tespiti için geliştirilen algoritma, öncelikle modelin yüklenmesini sağlar. Model arayüzüne yüklenen video, özellik çıkarma sürecinde olduğu gibi OpenCV ile okunur ve her videodan 32 adet çerçeve elde edilir. Veri seti ön işleme sürecine benzer bir süreç izlenerek, elde edilen çerçeveler içerisinde MTCNN modeli kullanılarak yüzler ayıklanır. Ayıklanan veriler, yüklenen modele verilerle modelin cevap üretmesi sağlanır.



Şekil 14: Video Tespiti

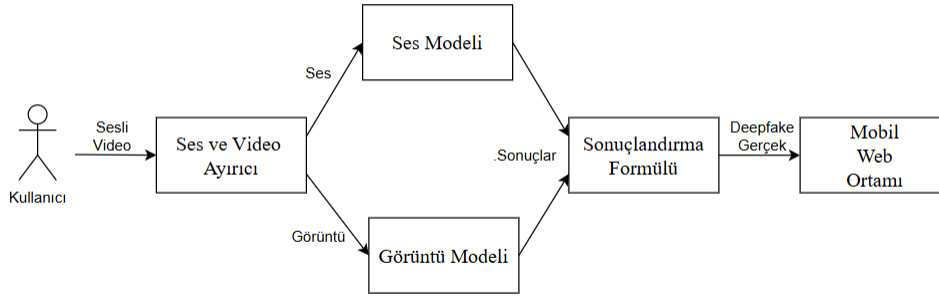
Oldukça benzer bir süreç, Deepfake ses tespiti için de geçerlidir. Kullanıcıdan alınan ses verisi uygun sinyale dönüştürülür. Ardından, MFCC ile özellik çıkarımı yapılarak, eğitilen modelin tahmin algoritması kullanılarak sonuç alınır.



Şekil 15: Ses Tespiti

Kullanıcının sesle birleşmiş bir şekilde deepfake ya da orijinal içerik yükleyebileceği ve sonuçları gözlemleyebileceği kullanıcı dostu bir web ortamı, Django teknolojisiyle geliştirilmiştir. Web ortamına oldukça benzer bir mobil platform da geliştirilerek, projenin her iki ortamda da hizmet verebilmesi sağlanmıştır.

Her iki platformda da kullanıcı, bir sesli video girdiğinde ses ve görüntü verilerinin ayrıştırılması için açık kaynak kütüphaneler kullanılarak bir algoritma tasarlanmıştır. Ayrıştırılan veriler, eğitilen modellerin tahmin etme fonksiyonlarına verilmiştir. Elde edilen sonuç, deneme yanılma yoluyla belirlenen formüle verilerle nihai sonuç elde edilmiştir ve sonuçlar kullanıcıya gösterilmiştir.



Şekil 16: Geliştirilen Platformun Mimarisi

## 4. Dağıtım

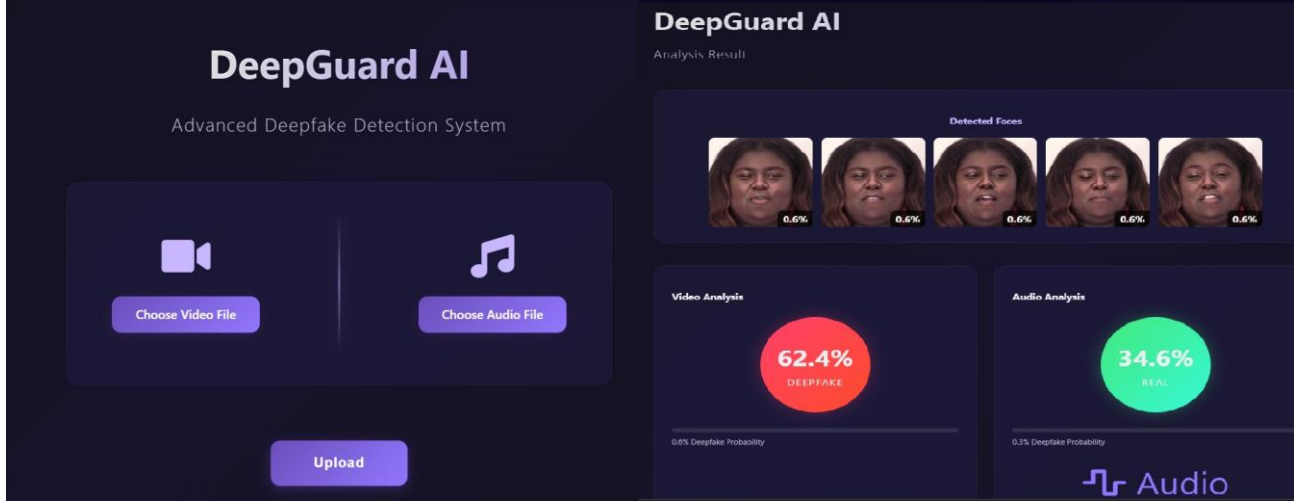
Bu aşamada hazırlanan mobil ve web uygulamaları dağıtımına uygun hale getirildi. Çeşitli web sunucularına uyarlanabilir hale getirilen web uygulaması için sunucu araştırması yapıldı. AWS, Azure gibi bilinen ortamlarda çalışabilir hale getirildi. Ancak bu ortamların ücretli olması sebebiyle modelin yerel sunucuda dağıtımını sağlandı. Bu aşamada proje desteğinin kullanılmama sebebi, AWS ve Azure gibi ortamların TÜBİTAK'ın istediği fatura formatına uygun fatura vermemesi ve ayrıca her ay sonunda yeniden faturalandırma yapmamasıydı. Yüklenen projenin AWS ortamında çalışabilmesi için düzenli olarak desteğe ihtiyaç duyulacağı tespit edildiği için yerel makinede kalındı.

Mobil uygulamanın Google Play Store ortamına aktarılması için kullanıcı hesabına 25 Dolar ödeme yapılması gerekmektedir. Bu ödeme, kullanıcının kendi hesabına özel olduğu için ve TÜBİTAK'ın bu harcamayı kabul etmeyeceği öngörüldüğü için mobil uygulamanın da dağıtımını yerel makinede sağlandı.

Proje, ngrok kullanılarak dağıtılmış ve internet üzerinden erişilebilir hale getirilmiştir. Ngrok, yerel sunucuyu güvenli bir şekilde internete açarak, projenin test edilmesini ve uzaktan erişim sağlanmasını mümkün kılmıştır. Ngrok, ters proxy mekanizması ile çalışarak, yerel makinede çalışan Django tabanlı sunucuyu güvenli bir HTTPS adresi üzerinden erişilebilir hale getirmiştir. Böylece, sabit bir IP adresine ihtiyaç duyulmadan, sistemin dış dünyaya açılması sağlanmıştır.

Ngrok'un kullanımı sayesinde, güvenlik duvarı ve ağ yapılandırması değiştirilmeden proje internet üzerinden erişime açılmıştır. Dinamik olarak oluşturulan alt alan adı, her bağlantı açıldığında güncellenmiş ve proje dış erişime hazır hale getirilmiştir. Ngrok'un erişim istatistikleri ve HTTP isteği kayıtları, sistemin kullanım verilerini analiz etmek ve hata ayıklama süreçlerini kolaylaştırmak için kullanılmıştır.

Proje kapsamında geliştirilen sistem, kullanıcıların video ve ses dosyalarını yükleyerek derin sahtecilik tespiti yapmalarına olanak tanımaktadır. Ngrok'un sağladığı güvenli tünelleme altyapısı, bu sürecin internet üzerinden erişilebilir olmasını sağlarken, geliştirme ve test aşamalarında uzaktan erişimi mümkün kılmıştır.



Şekil 17: Projenin Arayüzü

### 3. Sonuç

İP No	İş Paketlerinin Adı ve Hedefleri	Kim(ler) Tarafından Gerçekleştirileceği	Zaman Aralığı (... Ay)	Tamamlanma Durumu
1	<b>Veri Setlerinin Hazırlanması</b> Video ve ses tespiti için kullanılacak veri setleri bu aşamada hazırlanacaktır.	Ali Darı, Emine Şener, Mustafa Yüksel Yılmaz	2 Ay	Geliştirilecek olan Deepfake görüntü tespiti yapan modelin eğitiminde kullanılacak olan DFDC (DeepFake Detection Challenge) veri seti elde edildi, işlendi ve kapsamlı özellik çıkarma işlemleriyle model eğitime uygun hale getirildi. Bu veri seti, Deepfake video tespitine yönelik modelin başarıyla eğitilmesi için kritik bir kaynak olarak kullanıldı. Deepfake ses tespiti için ise ASVspoof veri seti elde edildi ve işlendi. Türkçe ses verileri kaydedildi, ardından Deepfake algoritmalarını kullanarak manipüle edilerek veri setine dahil edildi. Bu işlem, ses verilerinin daha geniş ve çeşitli bir eğitim kümesi oluşturmasına olanak sağladı. Uygun özellik çıkarımı yöntemleriyle ses verileri, model eğitime hazırlanarak eğitim sürecine uygun hale getirildi.
2	<b>Model Eğitimi</b> Seçilen derin öğrenme modelinin eğitimi bu aşamada yapılacaktır.	Ali Darı, Emine Şener, Mustafa Yüksel Yılmaz	4 Ay	Bu aşamada, CNN tabanlı iki ayrı model başarıyla geliştirildi. Görüntü tespit modeli, EfficientNet-B7 modelinin transfer öğrenme yöntemiyle özelleştirilmesiyle elde edildi. Bu model, önceden eğitilmiş olan EfficientNet-B7 modelinden faydalanarak, Deepfake video tespiti için yüksek doğruluk elde edilmesini sağladı. Ses modeli ise oluşturulan karmaşık CNN mimarisinin sıfırdan eğitilmesiyle başarıyla elde edildi. Bu model, sesli içeriklerdeki manipülasyonları tespit edebilmek için derin öğrenme tekniklerinden yararlanarak etkili bir şekilde eğitim aldı ve sesli veriler üzerinde başarılı sonuçlar üretmesi sağlandı.
3	<b>Web ve Mobil Uygulama Geliştirilmesi</b> Bu aşamada modelimiz internet ortamı sayesinde herkesin kullanımına sunulacaktır.	Ali Darı, Emine Şener, Mustafa Yüksel Yılmaz	3 Ay	Bu aşamada, kullanıcıdan veri alıp sonuçları etkili bir şekilde gösteren mobil ve web ortamları kuruldu. Ortamlar, eğitilen modellerle analiz yaparak sonuçları kullanıcıya başarıyla sunuyor. Kullanıcı dostu arayüzler sayesinde, kullanıcılar sesli ve görüntülü içeriklerini kolayca yükleyip deepfake tespiti yapabiliyor. Sonuçlar net ve anlaşılır bir şekilde gösterildi.
4	<b>Dağıtım</b> Bu kısımda geliştirdiğimiz projemizin dağıtımı olacaktır.	Ali Darı, Emine Şener, Mustafa Yüksel Yılmaz	1 Ay	Bu aşamada, geliştirilen mobil ve web ortamlarının çeşitli sunucu ortamlarında esnek bir şekilde çalışması sağlandı. Uygulamanın yerel ortamda dağıtımı başarıyla gerçekleştirildi, böylece uygulama farklı platformlarda da etkin bir şekilde çalışabilir hale geldi. Sunucu ortamlarına uyumlu hale getirilen bu sistem, farklı ihtiyaçlara ve kullanım durumlarına göre esneklik sağladı.

#### RİSK YÖNETİMİ TABLOSU\*

İP No	En Önemli Riskler	Risk Durumu
1	Yüksek İşlem Gücü ve Altyapı Sorunları Riski, Derin Öğrenme modelinin eğitimi ve videoların işlenmesi için gereken yüksek işlem gücü ve veri tabanı altyapısı gereksinimi	Bursa Teknik Üniversitesi Yüksek Hesaplamalı Yapay Zeka Laboratuvarı desteği sayesinde gerekli işlem gücü ve veri tabanı elde edildi.
2	Deepfake detection modelinin yanlış pozitif veya yanlış negatif sonuçlar vermesi	Modelin ürettiği sonuçlar, uygulama ve kullanım amacına göre tolere edilebilir düzeydedir. Hatalar minimal olup, sistemin genel verimliliğini etkilemeyecek seviyededir.

## 4. Çıktı

#### ARAŞTIRMA ÖNERİSİNDEN BEKLENEN YAYGIN ETKİ TABLOSU

Yaygın Etki Türleri	Önerilen Araştırmadan Beklenen Çıktı, Sonuç ve Etkiler	Tamamlanma Durumu
<b>Bilimsel/Akademik</b> (Makale, Bildiri, Kitap Bölümü, Kitap)	En az 1 adet bilimsel konferans bildirisi hazırlanacaktır.	Proje önerisinin kabul edilmesinin ardından, deepfake tespiti konusunda yayımlanan makaleler takip edilerek, pek çok konunun daha önce ele alındığı gözlemlendi. Bu nedenle, bildirinin yayımlanması kararından vazgeçildi.
<b>Ekonomik/Ticari/Sosyal</b> (Ürün, Prototip, Patent, Faydalı Model, Üretim İzni, Çeşit Tescili, Spin-off/Start-up Şirket, Görsel/İşitsel Arşiv, Envanter/Veri Tabanı/Belgeleme Üretimi, Telif Konu Olan Eser, Medyada Yer Alma, Fuar, Proje Pazarı, Çalıştay, Eğitim vb. Bilimsel Etkinlik, Proje Sonuçlarını Kullanacak Kurum/Kuruluş, vb. diğer yaygın etkiler)	Bu projeyi teknik olarak gerçekleştirebilirsek uygulamamızı sosyal medya platformlarına entegre ederek, platformlar üzerinde paylaşılan içeriklerin güvenliğini artırabilir ve toplumsal güvenliğe katkı sağlayabiliriz. Ayrıca uygulamanın API'lerini geliştirici topluluklarına veya diğer firmalara açarak, daha geniş bir ekosistem oluşturulabilir.	Proje teknik olarak başarıyla gerçekleştirildi. Üzerinde yapılacak ek çalışmalarla, kolaylıkla daha geniş bir ekosisteme yayılabilir.
<b>Araştırmacı Yetiştirilmesi ve Yeni Proje(ler) Oluşturma</b> (Yüksek Lisans/Doktora Tezi, Ulusal/Uluslararası Yeni Proje)	Bu projemizi başarıyla tamamlayabilirsek ileride biraz daha kapsamlı hale getirilerek bir yüksek lisans konusu olarak değerlendirilebilir ve daha büyük bütçeli bir proje kapsamında lisansüstü öğrencilere burs desteği sağlanabilir.	Yüksek lisansta aynı konu, büyük dil modeli desteği ve Agent teknolojisiyle yeniden çalışmaya uygundur.

# Kaynakça

- [1]Abhinav K. Venkataramanan, C. W. (2021). A Hitchhiker's Guide to Structural Similarity.
- [2]Brian Dolhansky, J. B. (2020). The DeepFake Detection Challenge (DFDC) Dataset.
- [3]Davide Cocomini, N. M. (2022). Combining EfficientNet and Vision Transformers for Video Deepfake Detection.
- [4]Defazio, A. (2024). Optimal Linear Decay Learning Rate Schedules and Further Refinements.
- [5]Edgar Kaziakhmedov, K. K. (2020). Real-world adversarial attack on MTCNN face detection system.
- [6]Jiangyan Yi, C. W. (2023). Audio Deepfake Detection: A Survey.
- [7]Jun Xue, C. F. (2022). Audio Deepfake Detection Based on a Combination of F0 Information and Real Plus Imaginary Spectrogram Features.
- [8]Mingxing Tan, Q. V. (2020). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.
- [9]Shruti Agarwal (1), T. E.-G.-N. (2020). Detecting Deep-Fake Videos from Appearance and Behavior.
- [10]Xinyi Wang, J. P. (2022). A Survey of Face Recognition.
- [11]Xuechen Liu, X. W. (2022). ASVspoof 2021: Towards Spoofed and Deepfake Speech Detection in the Wild.
- [12]Yang Gao, T. V. (2021). Generalized Spoofing Detection Inspired from Audio Generation Artifacts.
- [13]Yuezun Li, S. L. (2018). Exposing DeepFake Videos By Detecting Face Warping Artifacts.