

## Evrişimsel Sinir Ağları ile Türkçe Videolarda Geçen Küfür Seslerinin Sansürlenmesi

*Muhammed Mustafa TAYİZ<sup>1\*</sup>, Murat CANAYAZ<sup>2</sup>*

<sup>1</sup>Yapay Zeka ve Robotik Bölümü, Yüzüncü Yıl Üniversitesi, Van, Türkiye, ORCID: 0000-0001-5697-4101

<sup>2</sup>Bilgisayar Mühendisliği Bölümü, Yüzüncü Yıl Üniversitesi, Van, Türkiye, ORCID: 0000-0001-8120-5101

Sorumlu yazar: [m.mustafa.tayiz@gmail.com](mailto:m.mustafa.tayiz@gmail.com)

Geliş tarihi:13.12.2021

Kabul tarihi:29.12.2021

### Özet

İnternet ortamında ve özellikle sosyal mecralardaki video ve ses dosyalarında toplum ahlakını olumsuz yönde etkileyebilecek çok sayıda içerik bulunmaktadır. Ayrıca bu içeriklerin sayısı her geçen dakika artmaktadır. Bu kadar hızlı artan içerik sayısının kontrolü ve incelenmesi geleneksel yöntemler ile malesef mümkün olmamaktadır. Bu sebeple yapay zeka ve ses işleme yöntemleri kullanılarak bir oto-sansür uygulaması geliştirilmiştir. Çalışmada gelişmiş bir derin öğrenme modeli olan evrişimsel sinir ağı mimarisi kullanılmıştır. Bunun yanında tahmin aşamasında kullanılmak üzere, ses özellikleri çıkarmak için insan işitsel sistemine benzerliği sebebiyle mel-frekansı kepsral katsayıları algoritması tercih edilmiştir.

**Anahtar Kelimeler:** EVRİŞİMSSEL SİNİR AĞLARI, MFCC ALGORİTMASI, TÜRKÇE KÜFÜR SANSÜRLEME

### Censoring the Profanity Sounds in Turkish Videos with Convolutional Neural Networks

#### Abstract

There is a large amount of content on the Internet, especially in video and audio files on social media, that can negatively affect public morality. In addition, the number of these contents is increasing every minute. Unfortunately, it is not possible to control and examine the number of content that increases so rapidly with traditional methods. For this reason, a self-censorship prototype was developed using artificial intelligence and voice processing methods. Convolutional neural network architecture, which is an advanced deep learning model, was used in the study. In addition, the mel-frequency kepsral coefficients algorithm was preferred because of its similarity to the human auditory system to extract sound features to be used in the estimation phase.

**Keywords:** CONVOLUTIONAL NEURAL NETWORKS, MFCC ALGORITHM, TURKISH CURSING CENSORSHIP

## 1. GİRİŞ

Günümüzde sosyal mecralardaki içeriklerin internete yüklenirken ya da ilgili ortamlarda kullanıcılara sunulurken kontrol edilmesi, incelenmesi ve uygun olmayan kısımların sansürlenmesi manual yöntemler ile yapılmaktadır. İçeriklerin kontrol edilmesi ve sansürlenmesi insan müdahalesi ile olduğu için hataya çok açık bir süreçtir. Çünkü uzun zaman almakta ve dikkatli bir şekilde incelenmesi gerekmektedir. Dolayısıyla içerik artışı bu şekildeki bir kontrol mekanizmasını çok kolay bir şekilde aşabilmektedir. Bu sorun, özellikle gençler ve çocuklardan başlayarak toplumsal ahlakı ciddi bir şekilde etkilemektedir. Çünkü çocuklar sosyal mecralarda duydukları ve gördükleri içeriklerden hemen etkilenmekte ve sorgulamadan günlük hayatlarında kullanmaktadır. Karşılaşılan argo ve şiddet içerikli kelimelerin günlük hayatta kullanılması zamanla insanların birbirine olan saygısını ve dolayısıyla insan ilişkilerini olumsuz bir şekilde etkilemektedir. Sorun bu şekilde tanımlandıktan sonra çözüm için araştırma ve incelemeler yapılmıştır. Yapay zeka teknolojisinin her geçen gün artan yetenekleri ve başarı oranı bu problem için otomatize edilebilecek bir kontrol sistemi geliştirilebileceği fikrini ortaya çıkardı. Bu kapsamda yapılan yerli ve yabancı çalışmalar, tezler ve projeler incelenmiştir. Türkçe literatürde benzer bir çalışmaya rastlanmamıştır. Literatürdeki eksiklik ve dahada önemlisi uzun vadede toplumsal ahlakın korunması için bir adım atmak sureti ile çalışmalar başlatılmıştır.

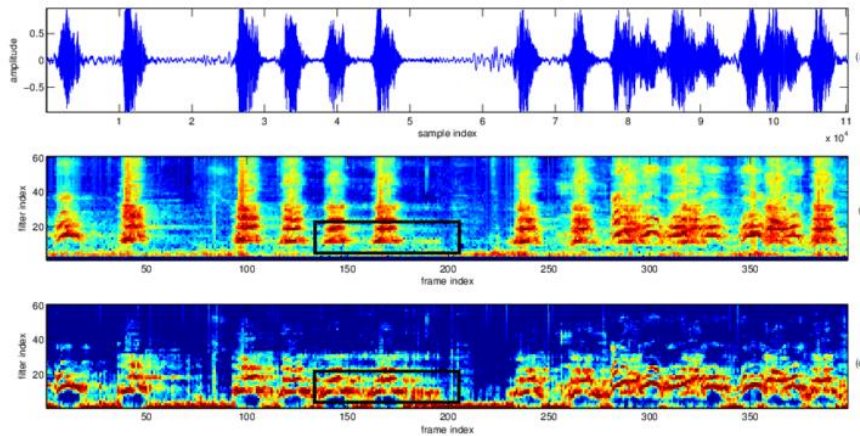
Ses sınıflandırma sistemleri incelendiğinde en önemli iki faktörün; çeşitli, kaliteli ve çok sayıda örnekten oluşan bir veriseti ile doğru ses özellik çıkarım algoritmasının tespit edilmesi olarak görülmüştür. İnternette erişime açık çok fazla veriseti bulunmaktadır. Ancak bu verisetlerinin büyük çoğunluğu yabancı diller özellikle ingilizce çalışmalar için hazırlanmıştır. Uygun bir ses veritabanı bulunamadığı için çalışmanın ilk aşaması olarak sıfırdan bir veriseti oluşturulmuştur. Çalışma sonunda literatüre yapay zeka alanında Türkçe bir çalışma kazandırmanın yanı sıra

birde Türkçe küfürlerden oluşan bir veriseti kazandırılacaktır. Çalışma süreci ve literatürdeki durum göz önüne alındığında Türkçe'deki bütün küfürlerin sansürlenebileceği bir çalışmanın yapılması çok gerçekçi görünmemektedir. Bu sebeple çok kullanılan 3 tane küfür seçilmiş, sadece bu 3 küfürün sansürlenebileceği bir prototip hazırlanması sonucunda çalışmanın amacına ulaştığı kabul edilmiştir.

Küfürlü seslerin sansürlendiği yabancı dildeki bir çalışmada evrişimsel sinir ağları kullanılmıştır. Gelişmiş CNN (Evrişimsel Sinir Ağı) modellerinden olan; Resnet50, GoogleNet, AlexNet, Vgg16 gibi derin öğrenme modelleri kıyaslanmıştır. Çalışma kapsamında 9 tane küfür sözcüğü belirlenmiş ve bu küfür sesleri toplanarak çeşitli ve çok sayıda örnekten oluşan bir veriseti oluşturulmuştur. 5100 tane günlük konuşma sesi ve 3105 tane küfür örneği ile toplamda 8 bin küsür ses örneği toplanmıştır. Belirlenen her küfür sesi için farklı lokasyonlardan 345 tane örnek alınmış ve tahmin için hazırlanmıştır. İlk denemelerde 9 tane küfürün her biri ayrı sınıflar ve birde günlük konuşma sesleri bir sınıf olarak belirlenmiş ve derin öğrenme modelleri 10 sınıfa göre çalıştırılmıştır. Sonra 9 tane küfürün hepsi bir sınıf birde günlük konuşma sesleri farklı bir sınıf olmak üzere 2 sınıf olarak denemeler yapılmıştır. Sonuçta Resnet50 modelinin diğer modellere göre daha iyi sonuçlar verdiği bildirilmiştir [1]. Başka bir ses sınıflandırma çalışmasında; 5 katmanlı bir CNN modeli ile evcil kedi seslerinin sınıflandırması başarılmıştır. Gerçek olarak elde edilen kedi seslerinin yetersiz olması sebebi ile mevcut sesler sentetik yöntemler ile çoğaltılmıştır. Sonuçta yaklaşık %87 oranında bir doğruluk elde edilmiştir [2]. Akciğer seslerinin sınıflandırılması ve bazı hastalıkların tahmin edilmesi amacıyla yapılan bir çalışmada çok sayıda tescilli örnek toplanmış ve bu ses örnekleri için özellik çıkarma işlemi MFCC algoritması ile sağlanmıştır [3]. Konuşma tanıma, çevresel seslerin sınıflandırılması ve müzik seslerinin sınıflandırılması kapsamında CNN, GAN (Çekişmeli Üretici Ağ) ve RNN (Tekrarlayan Sinir Ağı) gibi gelişmiş derin öğrenme modelleri ile çeşitli ses özellik çıkartma algoritmaları kıyaslanmıştır [4].

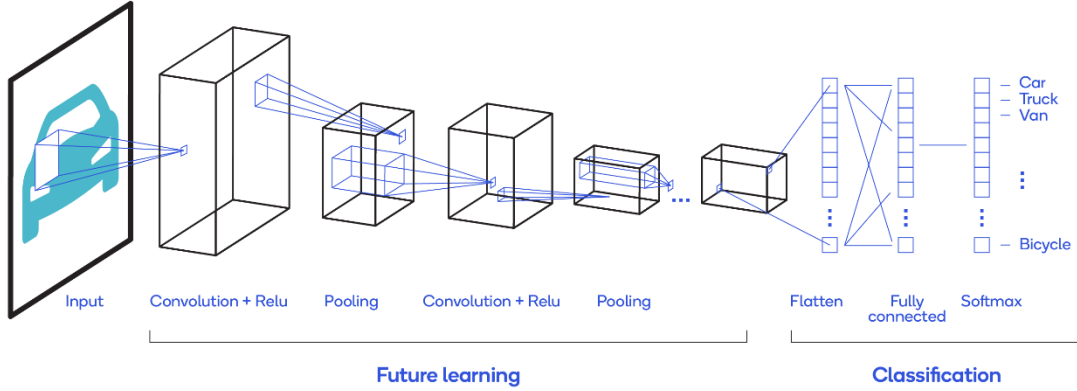
### 1.1. Evrişimsel Sinir Ağları (CNN)

Genel olarak resimleri-görüntüleri sınıflandırmak ve nesne tespit etmek için kullanılan, çok katmanlı, gelişmiş derin öğrenme modelleridir. Çoğu teknolojik gelişim gibi doğadan ve canlılardaki biyolojik mekanizmalardan esinlenerek ortaya çıkarılmış bir yapay sinir ağı çeşididir. Hayvanların görme sistemlerinden yola çıkılarak, 1988'li yıllarda ilk CNN ağı LeNet mimarisi geliştirilmeye başlanmıştır [5]. Örüntü tanıma gibi karmaşık problemlerin çözümünde de kullanılmaktadır. İlk çıkış zamanında daha çok kenar tespiti, basit şekillerin tanımlanması gibi alanlarda kullanılırken zamanla insan yüzü tanıma, video analizi gibi çok daha zorlu problemlerin çözümünde kullanılmaya başlamıştır [6]. Ses sınıflandırma problemlerinde ise sinyallerinin görsel temsili olan spektrogram özellikleri kullanılmaktadır.



Şekil 1. Ham ses sinyal grafiği ve MFCC spektrogram örnekleri [7].

Girdi olarak alınan görüntü çeşitli katmanlardan geçirilerek görüntünün özellikleri öğrenilir. Bu katmanlar sırasıyla evrişim katmanı, aktivasyon katmanı (non-linearity), havuzlama (pooling) şeklinde olacağı gibi katman sayısı artırılarak öğrenme süreci bu kapsamda değerlendirilebilir. Sonraki aşamada ise sınıflandırma yapılır. Bu aşamada önceki katmanlardan gelen matrisler tek boyutlu diziye çevrilir ve tahminler elde edilir. Flattening layer ve tam bağlantı katmanı (fully connected layer) bu aşamada görev alır [8].

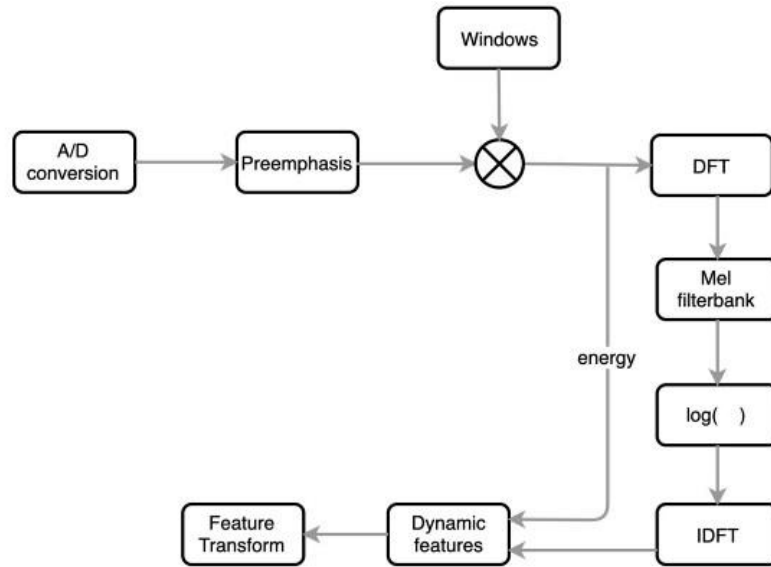


Şekil 2. Örnek bir Evrişimsel Sinir Ağı Modeli [9].

## 1.2. Mel-Frekans Cepstral Katsayıları (MFCC) Algoritması

Ses sınıflandırma modellerinde en önemli aşamalardan biri ses dosyalarının özelliklerinin çıkarılmasıdır. Ses özellik çıkarımı; bir ses dosyasının diğer seslerden ayrıştırılabilmesi amacıyla benzersiz matematiksel değerlerin hesaplanması anlamına gelir. Literatürde çok sayıda özellik çıkarım algoritması kullanılmıştır. Ancak insan işitsel sistemine benzerliği sebebi ile birçok ses sınıflandırma probleminde MFCC algoritması kullanılmıştır.

Aşağıdaki diyagram üzerinden MFCC algoritmasının çalışma mantığı incelenebilir.



Şekil 3. MFCC tekniğinin yol haritası [10].

### A/D Dönüşümü

Bu adımda analog olarak bulunan ses, belirlenen örnekleme frekansına göre dijital bir forma çevirilerek bilgisayar ortamına aktarılır.

### Ön Vurgu (Preemphasis)

Yüksek geçiren filtre kullanılarak frekans enerjisinin büyüklüğü artırılır. Bu sayede anlaşılabilirlik ve performans artırılır.

### Pencereleme (windows)

Bu aşamada ses dosyası Hamming pencereleri kullanılarak segmentlere ayrılır.

### DFT (Ayrık Fourier Dönüşümü)

Zaman alanında olan ses sinyali, frekans alanına çevrilerek daha kolay analiz yapılması için hazırlanır.

### Mel-Filtre Bankası (Mel Filterbank)

Bu aşamada ses sinyalinin gerçek frekansını insanların algılayabileceği frekansa haritalamak için mel ölçeği kullanılır.

$$mel(f) = 1127 \ln \left( 1 + \frac{f}{700} \right) \quad (1)$$

### Log

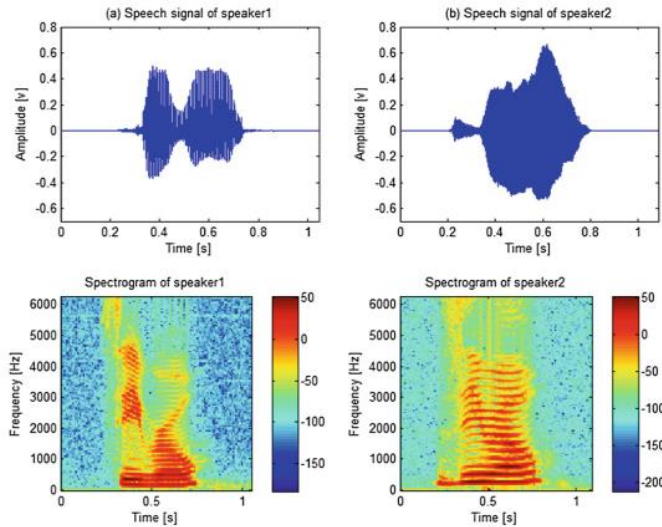
İnsan işitsel sistemini taklit etmek amacı ile Mel-filtre çıkışına  $\log()$  fonksiyonu uygulanır. Çünkü insan işitme sistemi yüksek sinyal enerjisindeki değişime, düşük sinyal enerjisine kıyasla daha duyarlıdır. Log fonksiyonunda bu duruma benzer bir özelliğe sahiptir.

### IDFT (Ters Ayrık Fourier Dönüşümü)

Log aşamasında elde edilen çıktının ters dönüşümü alınır.

### Dinamik Özellikler (Dynamic Features)

Ses sinyallerinin arasındaki katsayıların farkları ile türev hesaplanır [10].



Şekil 4. Konuşmacı ses örneği ile spektrogram grafiği [11].

## 2. MATERYAL VE METOD

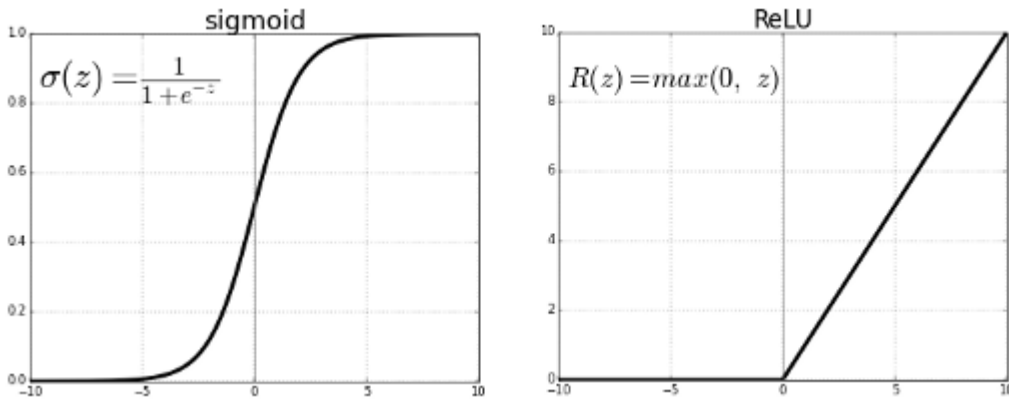
### 2.1. Kullanılan Sinir Ağı Modeli

Çalışma kapsamında gelişmiş bir sinir ağı modeli olan evrışimsel sinir ağları kullanılmıştır. İlgili yapay sinir ağı modeli tasarlandıktan sonra bütün ses dosyaları için; MFCC yöntemi ile elde edilen spectrogram çıktıları ile model eğitilmiştir. Model tasarlanırken python programlama dili ile numpy, pandas, SpeechRecognition, TensorFlow, Keras ve SkLearn gibi yazılım kütüphaneleri kullanılmıştır.

**Tablo 1.** Evrışimsel Sinir Ağı Katman Ayrıntıları

<u>10-Katman</u>	<b>Dense</b>	<b>10</b>	<b>Softmax</b>
<u>9-Katman</u>	Dense	64	ReLU
<u>8-Katman</u>	Dense (FC)	128	ReLU
<u>7-Katman</u>	Flatten		
<u>6-Katman</u>	Dropout	(0, 5)	
<u>5-Katman</u>	MaxPool2D	(2, 2)	
<u>4-Katman</u>	Conv2D	128	ReLU
<u>3-Katman</u>	Conv2D	64	ReLU
<u>2-Katman</u>	Conv2D	32	ReLU
<u>1-Katman</u>	Conv2D	16	ReLU

Tablo 1’de görüldüğü üzere 10 katmandan oluşan bir model tasarlanmıştır. Aktivasyon fonksiyonu olarak ReLu kullanılmıştır. ReLU aktivasyon fonksiyonunun derin öğrenmede ve özellikle evrışimsel sinir ağlarında çok kullanıldığı bilinmektedir. Bu fonksiyona ait grafik Şekil 5’te verilmiştir.



**Şekil 5.** ReLu ve Sigmoid aktivasyon fonksiyonlarının grafikleri [12].

## 2.2. Kullanılan Veriseti

Çalışma için ilk aşamada üç tane çok kullanılan Türkçe küfür belirlenmiştir. Yapılan ön araştırma aşamasında “ş\*\*\*\*, a\*\*\*\*k\*\*\*\* ve p\*\*\*\*” küfürleri ile çalışılabileceği kararlaştırılmıştır. Türkçe ses veritabanları incelendiğinde bu kapsamda herhangi bir verisetine rastlanmamıştır. Bu sebeple sıfırdan bir veriseti oluşturma durumu ortaya çıkmıştır. Tahminleme sürecinde başarılı olabilmek için çok sayıda ve çeşitli bir ses kütüphanesine sahip olmak gerektiği yapılan literatür taraması sonucunda kesin olarak anlaşılmıştır.

Küfür seslerini toplamak için birçok farklı yol ve yöntem denenmiştir. Bunlardan birkaçı; youtube videoları, film ve diziler, yazıyı sese dönüştüren uygulamalar ve whatsapp gibi iletişim kanalları olarak sıralanabilmektedir. Sesler toplandıktan sonra bazı sentetik ses artırma yöntemleri uygulanarak verisetinin çeşitliliği arttırılmıştır. Son durumda aşağıdaki istatistiklere sahip bir veri seti oluşturulmuştur.

**Tablo 2.** Veriseti İstatistikleri

	<b>İlk Küfür (p****)</b>	<b>İkinci Küfür (ş****)</b>	<b>Üçüncü Küfür (a****k****)</b>	<b>Arka plan ve küfür olmayan konuşma sesleri</b>
<b>Toplam (adet)</b>	93	90	117	100
<b>Kadın Sesi</b>	14	11	12	-
<b>Erkek Sesi</b>	79	79	105	-
<b>En kısa örnek (sn)</b>	0,37	0,56	0,46	0,06
<b>En uzun örnek (sn)</b>	1,08	2,61	2,98	3,23
<b>Toplam süre (sn)</b>	40,46	71,41	99,63	171,11

Veri seti oluşturulduktan sonra sesler için çıkarılacak özelliklerin daha belirleyici olabilmesi için çeşitli ön işlemler uygulanmıştır. Bunlardan bazıları; sinyaller belirli bir eşik değerine göre filtrelenmiş, arkaplan gürültüleri temizlenmiş, ses artırma ve parazit giderme gibi işlemler uygulanmıştır. Son olarak ses sinyalleri etiketlenerek bir csv dosyası oluşturulmuştur.

**Tablo 3.** Etiketlenmiş ses örnekleri

<b>Ses</b>	<b>Etiket</b>
speed_k1	kufur_1
noise_k3	kufur_3
ses_12	diger_ses

## 2.3. Küfür Sansür Sisteminin Genel Çalışma Aşamaları

Model oluşturulduktan ve eğitildikten sonra aşağıdaki aşamalar ile ilerlenmiştir.

- Geçici seslerin ve önceki çıktıların olduğu klasörler temizlenir.
- Girdi olarak bir video alınır ve video ile ses ayrıştırılır.
- Ayrıştırılan ses dosyaları belirlenen süre parametresine göre parçalanır. (örnek: 0.7 sn)
- Parçalanmış sesler sırayı koruyacak şekilde isimlendirilir ve bir klasörde tutulur.

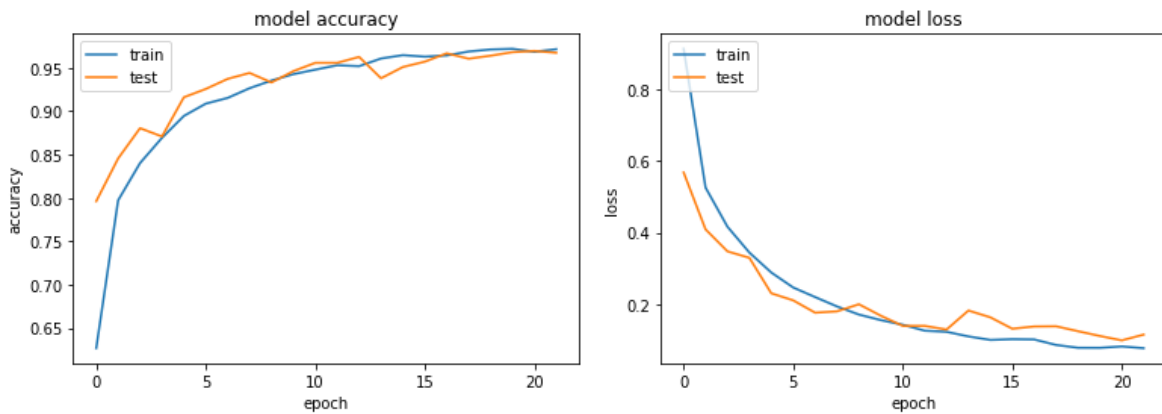
- Bütün bu ses dosyaları belirli ön işlemlerden geçirilerek farklı bir klasöre kaydedilir. (Tahmin için kullanılacak)
- Bütün bu sesler “test” etiketi ile etiketlenir.
- Eğitilmiş model dosyaları yüklenir. (Model eğitildikten sonra kaydedilmişti.)
- Her ses dosyası sırasıyla evrişimsel sinir ağı modeline girdi olarak verilir ve belirlenen doğruluk yüzde eşliğinin üstündekiler ayrıştırılır.
- Ayrıştırılan bu ses dosyaları ilgili metotlara gönderilerek zaman bilgisine dokunulmadan ses kısma işlemi uygulanır.
- Varsa kısılan sesler ile tahmin sonucu ilgili eşığı geçmeyen sesler aynı sıra ile birleştirilir.
- Son olarak sansürlenmiş ses ile sesi ayrıştırılmış video birleştirilerek çıktı klasörüne taşınır.

### 3. BULGULAR

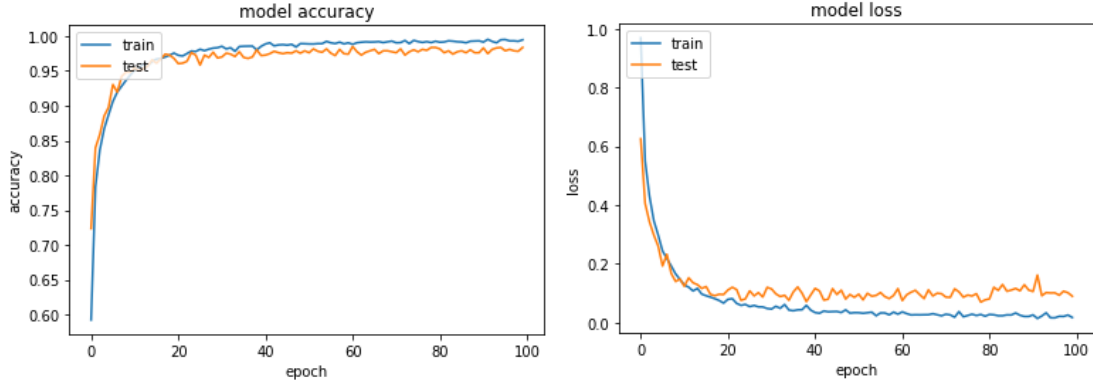
Oluşturulan veri setinde 3 tane küfür ve birde küfür içermeyen diğer seslerden oluşan bir sınıf olmak üzere toplam 4 farklı sınıf bulunmaktadır. Evrişimsel sinir ağı tasarlanırken farklı parametreler kullanılarak çok kez çalıştırılmıştır. Farklı epoch (eğitim tur) değerleri, farklı katman sayısı ve dizilimleri, dropout (seyreltme) gibi hiperparametreler kullanılmıştır.

Hiperparametre: derin öğrenme ya da diğer sınıflandırma sistemlerinin çalışma mantığı bellidir, ancak her problem aynı olmadığı için model tasarımcısının müdahale etmesi ve değiştirmesi gereken bazı parametreler bulunmaktadır. Verisetine ve probleme göre değişiklik gösteren parametrelere hiper-parametre denir. Epoch (eğitim tur) ise, derin öğrenme modelleri eğitim aşamasında iken; veriler gruplar halinde ağıya verilmektedir. Her seferinde ilgili veri grubu ile eğitim tamamlanmakta, elde edilen sonuç geriye-yayılım algoritmasına girdi olarak verilerek ağırlık değerleri güncellenmektedir. Tekrar eden bu süreçteki her bir aşamaya, ya da her bir tura epoch denir. Modeller çalıştırılmadan önce epoch değeri belirlenmelidir.

Model en son halini aldıktan sonra farklı verisetleri kullanılarak çeşitli testler yapılmış ve başarı oranının iyi olduğu görülmüştür. Elimizdeki verisetine sentetik veri artırma teknikleri uygulamadan önce aynı modeli kullanmamıza rağmen beklenen doğruluk oranı elde edilememiştir. Veriseti üzerinde bir süre daha çalışıldıktan sonra Tablo 2’de verilen istatistiklere sahip bir set elde edilmiştir. Aynı model üzerinde yapılan denemelerde başarı oranının veriseti kalitesi ile doğru orantılı olduğu açık bir şekilde görülmüştür. Son olarak daha iyi sonuçlar elde etmek için verisetindeki sınıflara literatürde görülen teknikler uygulanmıştır. İlk denemelerde 3 küfür, 1 diğer sesler olmak üzere 4 sınıf olarak düzenlenen sistem küfürler 1 sınıf ve diğer sesler 1 sınıf olarak güncellenmiştir. Bu güncelleme sonrasındaki denemelerde (denemeler sadece epoch değeri değiştirilerek yapılmıştır) başarı oranının arttığı görülmüştür. Örnek olarak kullanılan videolarda geçen küfürlerin yüksek doğruluk oranları ile tespit edilmesi sağlanmıştır. Eğitim sırasında elde edilen doğruluk ve kayıp fonksiyonu grafikleri şekil 6 ve 7’de verilmiştir.



Şekil 6. [22 Epoch] için doğruluk (accuracy) ve kayıp (loss) grafikleri

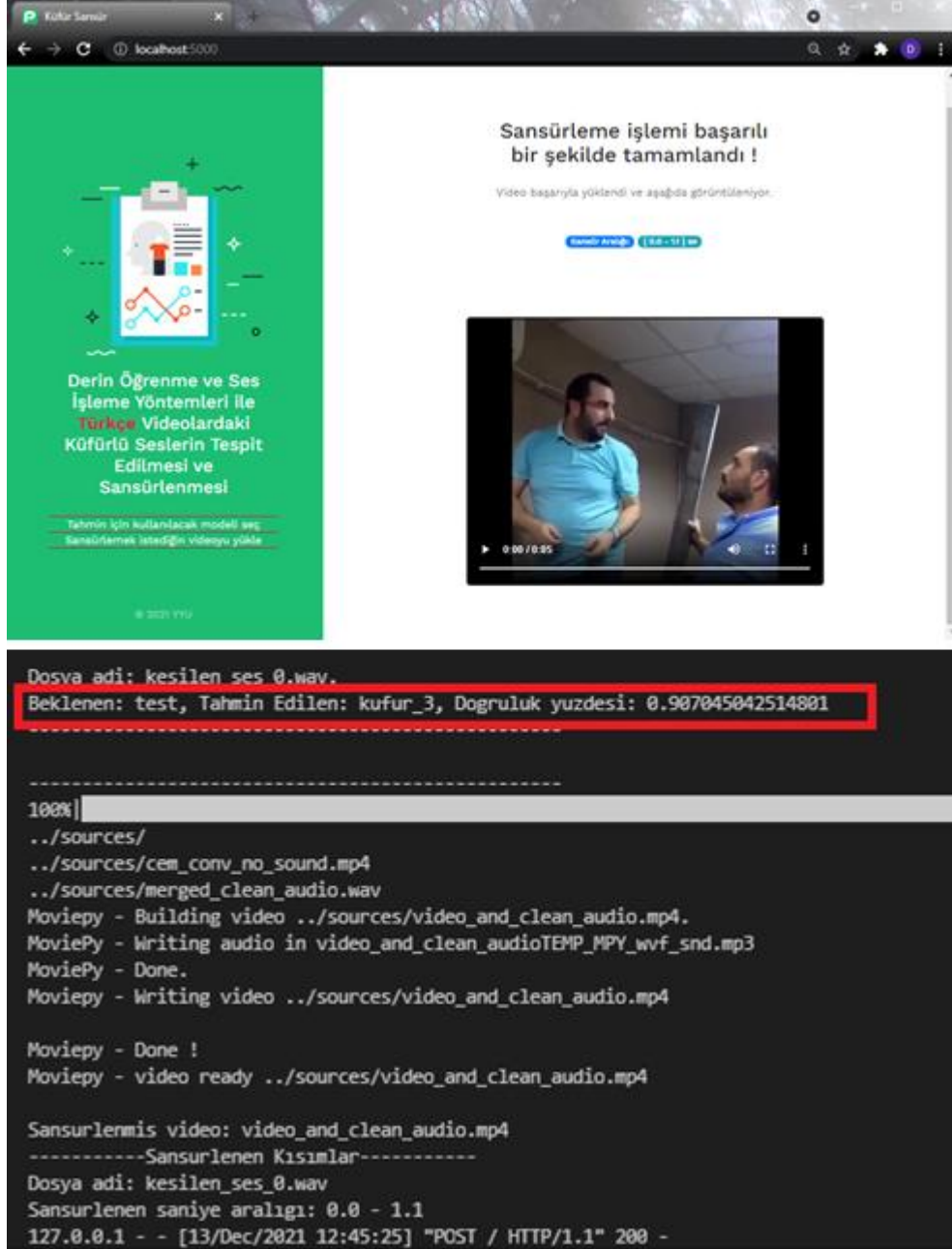


Şekil 7. [100 Epoch] için doğruluk (accuracy) ve kayıp (loss) grafikleri

### Kullanılan Örnek Videolar

Derin öğrenme modeli verisetinin son hali kullanılarak eğitildikten sonra eğitilmiş model kaydedilmiştir. Kardeş Payı dizisi, çeşitli vinelar ve Youtube'dan alınan video kesitlerinden ilgili küfürleri içeren kısımlar test için kullanılmak üzere video düzenleme programları ile manuel olarak kırılmıştır. Çalışma kapsamındaki küfürleri içeren bu video örnekleri farklı süreler ile bölünerek elde edilen evrişimsel sinir ağı modeline girdi olarak verilmiştir. Belirlenen eşik değerinin üzerinde doğruluk ile küfür olarak tespit edilen kısımlar sansürlenerek oluşturulan yazılımın çıktısı olarak alınmıştır. Kullanılan örnek videolar ile yapılan testler sonucunda algoritmanın doğruluk eşik değeri %90 olarak ayarlanmıştır. Sistemde bu eşik değeri kullanıldığında istenilen performansta çalıştığı gözlemlenmiştir. Test olarak kullanılan videoların birinde toplam 45 parçaya bölünen ses dosyasında geçen 5 tane küfür ile küfür olmayan bazı sesler sansürlenmiştir. Başka bir örnekte ise 15 küfür olmayan kelime ile  $a^{***}k^{***}$  ve  $p^{**}$  küfürleri geçen 7 saniyelik bir videoda sadece küfürler tespit edilip sansürlenmiştir.





Şekil 8. Geliştirilen yazılım ile sansür işlemi uygulanmış örnek bir video ve ilgili çıktılar.

#### 4. SONUÇLAR

Çalışmada belirlenen hedefe ulaşılmış ve 3 tane Türkçe küfürlü sesin yapay zeka yöntemleri ile sansürlenebileceğini gösteren bir prototip hazırlanmıştır. Süreç boyunca birçok zorluk ile karşılaşmıştır. Özellikle veriseti oluşturulurken projenin amacının anlatılmasına rağmen insanlardan ses kaydı talebi çoğu zaman olumsuz sonuçlanmıştır. Bu sebeple sentetik veri artırma teknikleri kullanılarak veriseti zenginleştirilmiştir. Karşılaşılan başka bir problem ise sansür uygulamak için parçalanmış ana ses dosyasının ilgili kısmı küfürün sadece bir kısmını içerebilmesidir. Örneğin ana ses dosyası 0.4 saniyelik parçalara ayrıldığında, ilgili küfür 3. parçada başlamış ve 4. parçada bitmiş ise bu gibi durumlarda sistem beklenen çıktıyı verememektedir. İlgili sorun başlı başına bir araştırma-geliştirme projesi olarak çalışılabilecek bir konudur.

Literatür taramasında veriseti kalitesinin, derin öğrenme modellerinin başarımını doğrudan etkilediği görülmüştü. Ancak çalışma sürecinde verisetinin düzen, çeşitlilik ve sınıf ayrımı gibi parametrelerinin öğrenme sürecindeki etkisinin ne kadar ciddi olduğu deneysel çalışmalarla açıkça anlaşılmıştır.

Türkçe literatürde daha önce küfür sansürü ile ilgili bir çalışma olmadığı bilinmektedir. Dolayısıyla yapılan bu çalışma ile literatüre yapay sinir ağları kapsamında Türkçe küfür sansürü yapan bir çalışma kazandırılmıştır. Bunun yanında sıfırdan oluşturulan Türkçe küfür seslerinden oluşan bir veriseti yayınlanacak, yeni fikirler ve çalışmalar için kullanıma sunulacaktır.

#### KAYNAKÇA

1. A. . S. B. Wazir, H. A. Karim, M. H. L. Abdullah, S. Mansor, N. AlDahoul, M. . F. A. Fauzi ve J. See, SPECTROGRAM-BASED CLASSIFICATION OF SPOKEN FOUL LANGUAGE USING DEEP CNN, 2020.
2. Y. R. Pandeya ve J. Lee, Domestic Cat Sound Classification Using Transfer Learning, International Journal of Fuzzy Logic and Intelligent Systems, pp. 154-160, 2018.
3. N. Sengupta, M. Sahidullah ve G. Saha, Lung sound classification using cepstral-based statistical features, Computers in Biology and Medicine, 2016.
4. H. Purwins, B. Li, T. Virtanen, J. Schlüter, S.-y. Chang ve T. Sainath, Deep Learning for Audio Signal Processing, JOURNAL OF SELECTED TOPICS OF SIGNAL PROCESSING, pp. 206-219, 2019.
5. A. Şeker, B. Diri ve H. H. Balık, Derin Öğrenme Yöntemleri ve Uygulamaları Hakkında Bir İnceleme, Gazi Mühendislik Bilimleri Dergisi, pp. 47-64, 2017.
6. S. Albawi, T. A. Mohammed ve S. Al-Zawi, Understanding of a convolutional neural network, IEEE, Antalya, 2017.
7. R. N. Tak, D. Agrawal ve H. Patil, Novel Phase Encoded Mel Filterbank Energies for Environmental Sound Classification, International Conference on Pattern Recognition and Machine Intelligence, 2017.
8. Prabhu, Understanding of Convolutional Neural Network (CNN) — Deep Learning. Available: <https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148>. [Erişildi: 2021].
9. Q. Developer, Deep Learning and Convolutional Neural Networks for Computer Vision, Qualcomm. Available: <https://developer.qualcomm.com/software/qualcomm-neural-processing-sdk/learning-resources/cnn-architectures/deep-learning-convolutional-neural-networks-computer-vision>. [Erişildi: 2021].
10. U. Kiran, MFCC Technique for Speech Recognition, Analytics Vidhya. Available: <https://www.analyticsvidhya.com/blog/2021/06/mfcc-technique-for-speech-recognition/>. [Erişildi: 2021].
11. M. Jenhi, A. Roukhe ve L. Hlou, Analysis of Speaker's Voice in Cepstral Domain Using MFCC Based Feature Extraction and VQ Technique for Speaker Identification System, pp. 857-868, 2019.
12. S. SHARMA, Activation Functions in Neural Networks. Available: <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>. [Erişildi: 2021].
13. S. Albawi, T. A. Mohammed ve S. Al-Zawi, Understanding of a convolutional neural network, International Conference on Engineering and Technology (ICET), Antalya, 2017.