**Solunum Sesi Sınıflandırması: Veri Toplama ve Ön İşleme**

İbrahim Buğra SAN

Bilişim Sistemleri Mühendisliği

Kocaeli Üniversitesi  
ibugrasan@gmail.com

**ÖZET:**

*Bu makale, beş farklı solunum durumunu (bronşit, krup, normal nefes, zatürre ve astım) tanımlamak için tasarlanmış bir makine öğrenmesi sınıflandırma sistemi için solunum sesi verilerinin toplanması ve ön işlenmesine yönelik kapsamlı bir yaklaşım sunmaktadır. Veri toplama süreci, önceden tanımlanmış parametrelere dayalı olarak çevrimiçi video kaynaklarından ilgili ses segmentlerini çıkarmak için özel bir crawler geliştirmeyi içermiştir. Python ve çeşitli ses işleme kütüphaneleri kullanılarak, solunum sesi örneklerini toplama, segmentlere ayırma ve düzenleme için sağlam bir pipeline oluşturulmuştur. Toplanan veri seti, solunum durumu sınıflandırmasına göre düzenlenmiş, 5 saniyelik sürelere standardize edilmiş ses segmentlerinden oluşmaktadır. Bu makale, özellikle veri edinme ve ön işleme metodolojisine odaklanarak, bu özelleştirilmiş ses veri setinin geliştirilmesi sırasında karşılaşılan zorluklar ve uygulanan çözümler hakkında bilgiler sunmaktadır.*

**I. GİRİŞ**

*Solunum hastalıkları, dünya genelinde önemli bir sağlık sorunu olup, erken ve doğru teşhis, etkili tedavi için kritik öneme sahiptir . Geleneksel teşhis yöntemleri genellikle pahalı ekipman ve uzman personel gerektirmektedir. Son yıllarda, makine öğrenmesi teknolojilerindeki ilerlemeler, solunum seslerinin analizi yoluyla hastalıkların otomatik sınıflandırılması için yeni fırsatlar sunmaktadır.*

*Literatürde, farklı solunum patolojilerinin kendine özgü akustik özellikler sergilediği gösterilmiştir. Çeşitli araştırmacılar, bronşit, astım ve zatürre gibi hastalıklarda normal solunum seslerine kıyasla belirgin spektral farklılıklar tespit etmiştir. Solunum seslerinin* *sınıflandırılması için Destek Vektör Makineleri, Rastgele Ormanlar ve Derin Öğrenme modelleri gibi çeşitli yaklaşımlar önerilmiştir. Ancak, gerçek dünya koşullarında kaydedilmiş, çeşitli solunum patolojilerini içeren geniş veri setlerinin eksikliği, bu alandaki araştırmaların klinik uygulamalara dönüşmesini sınırlamaktadır.*

*Bu çalışmanın amacı, çevrimiçi kaynaklardan toplanan solunum sesi örneklerini kullanarak, beş farklı solunum durumunu (bronşit, krup, normal nefes, zatürre ve astım) sınıflandırmak için bir veri toplama ve ön işleme metodolojisi sunmaktır. Çalışmamızın katkıları: (1) otomatik bir crawler sistemi geliştirmek, (2) standardize edilmiş bir solunum sesi veri seti oluşturmak, (3) etkili bir veri temizleme metodolojisi önermek ve (4) veri kalitesinin değerlendirilmesi için metrikler sunmaktır.*

*Bu makalenin geri kalanı şu şekilde düzenlenmiştir: Bölüm II'de veri toplama metodolojisi, Bölüm III'te veri temizleme ve ön işleme teknikleri, Bölüm IV'te verilerin analizi ve görselleştirmeleri, Bölüm V'te ise sonuçlar ve gelecekteki çalışmalar için öneriler sunulmaktadır.*

**II. VERİ TOPLAMA**

**A. Veri Kaynakları**

*Bu proje için birincil veri kaynağı, çeşitli tıbbi durumlardan solunum seslerini içeren çevrimiçi videolardan oluşmaktadır. Bu videolar, ilgililik, ses kalitesi ve hedef solunum durumlarının doğru temsili temelinde dikkatle seçilmiştir. Her video URL'si, sınıflandırma kategorisi ve ilgili zaman damgası bilgileri dahil olmak üzere meta verilerle birlikte yapılandırılmış bir Excel dosyasında belgelenmiştir.*

**B. Araçlar ve Ortam**

*Veri toplama sistemi aşağıdaki araçlar ve ortam kullanılarak uygulanmıştır:*  
*• Programlama Dili: Python*   
*• Geliştirme Ortamı: Anaconda ve Jupyter Notebook*  
*• VS Code*  
*• Versiyon Kontrolü: Git*

**C. Kullanılan Kütüphaneler**

*Crawler uygulaması, birkaç özelleştirilmiş Python kütüphanesinden yararlanmıştır:*  
*• pandas: Video meta verilerini içeren Excel dosyasını okumak ve işlemek için*  
*• yt-dlp: Çevrimiçi platformlardan video indirmek için*  
*• ffmpeg-python: Ses çıkarma ve segmentasyon için*  
*• os: Dosya sistemi işlemleri ve dizin yönetimi için*  
*• re: Dosya adı işlemede düzenli ifade deseni eşleştirmesi için*  
*• glob: Dosya yolu deseni eşleştirmesi için*

**D. Veri Toplama Süreci**

*Veri toplama iş akışı şu sıralı adımları izlemiştir:*

*1. Meta Veri Hazırlama: Dört sütunlu bir Excel dosyası hazırlanmıştır:*  
*• link: Video kaynağının URL'si*  
*• sınıf: Sınıflandırma etiketi (bronşit, krup, normal, zatürre, astım)*  
*• başlangıç: İlgili sesin başlangıcını işaretleyen zaman damgası (HH:MM:SS formatında)*  
*• bitiş: İlgili sesin sonunu işaretleyen zaman damgası (HH:MM:SS formatında)*

*2. Video Edinimi: Crawler, Excel dosyasındaki her satırı sırayla işlemiştir:*  
*• Optimize edilmiş parametrelerle yt-dlp kullanarak videoyu indirme*  
*• Henüz mevcut değilse sınıfa özgü dizinler oluşturma*  
*• Başarısız indirmeler için yeniden deneme mekanizmaları ile hata yönetimi uygulama*

*3. Ses Çıkarma: Başarıyla indirilen her video için:*  
*• ffmpeg kullanarak ses parçasını çıkarma*  
*• Sesi belirtilen başlangıç ve bitiş zaman damgalarına göre kırpma*  
*• Kırpılmış sesi standartlaştırılmış 5 saniyelik kliplere bölme*  
*• Tüm sesleri tutarlı örnekleme hızı ve bit hızı ile MP3 formatına dönüştürme*

*4. Organizasyon: Çıkarılan ses segmentleri:*  
*• Kaynaklarına ve zaman damgası bilgilerine göre adlandırılmış*  
*• Model eğitimi sırasında kolay erişim için sınıfa özgü dizinlerde saklanmış*

**E. Veri Seti İstatistikleri**

*Veri toplama süreci, ses örneklerinin aşağıdaki dağılımıyla sonuçlanmıştır:*

*• Bronşit: 89 örnek (7.6 dakika)*  
*• Krup: 80 örnek (7.5 dakika)*  
*• Normal: 72 örnek (7.1 dakika)*  
*• Zatürre: 77 örnek (7.3 dakika)*  
*• Astım: 82 örnek (7.5 dakika)*

**F. Zorluklar ve Çözümler**

*Veri toplama süreci sırasında birkaç zorlukla karşılaşılmıştır:*

*1. Video Kullanılabilirliği: Bazı videolar toplama süreci sırasında kullanılamaz hale gelmiştir.*  
 *Çözüm: Başarısız indirmeleri kaydeden ve kalan videolarla devam eden sağlam bir hata yönetim sistemi uygulanmıştır.*

*2. Tutarsız Ses Kalitesi: Videolarda değişen ses kalitesi ve arka plan gürültü seviyeleri vardı.*  
 *Çözüm: Daha sonraki ön işleme kararlarını bilgilendirmek için her kaynak için ses kalitesi metrikleri belgelenmiştir.*

*3. Ağ Kesintileri: Aralıklı ağ sorunları indirme başarısızlıklarına neden olmuştur.*  
 *Çözüm: Üstel geri çekilme ile otomatik yeniden deneme mekanizması uygulanmıştır.*

*4. Zaman Damgası Doğruluğu: Sağlanan bazı zaman damgaları solunum sesleriyle tam olarak hizalanmamıştır.*  
 *Çözüm: Kritik örnekler için zaman damgalarını manuel olarak kontrol etmek ve ayarlamak için bir doğrulama adımı geliştirilmiştir.*

*5. Format Uyumluluğu: İndirilen bazı videolar ffmpeg işleme için uyumsuz formatlara sahipti.*  
 *Çözüm: Uyumluluğu sağlamak için format algılama ve dönüştürme adımları eklenmiştir.*

**III. VERİ TEMİZLEME VE ÖN İŞLEME**

**A. Ses Ön İşleme Teknikleri**

*Veri kalitesini iyileştirmek için aşağıdaki ön işleme teknikleri uygulanmıştır:*

*1. Gürültü Azaltma: Solunum sesi özelliklerini korurken arka plan gürültüsünü azaltmak için spektral gürültü geçidi uygulanmıştır.*

*2. Genlik Normalizasyonu: Tutarlı ses seviyeleri sağlamak için tüm örneklerde ses genliği normalize edilmiştir.*

*3. Sessizlik Kaldırma: Ses segmentlerinden öndeki ve arkadaki sessizlikler tespit edilip kırpılmıştır.*

*4. Örnekleme Hızı Standardizasyonu: Tüm ses örnekleri tutarlı bir 22.05 kHz örnekleme hızına dönüştürülmüştür.*

*5. Kanal Dönüşümü: Stereo varyasyonları ortadan kaldırmak için tüm sesler mono kanala dönüştürülmüştür.*

*6. Ses Formatı Standardizasyonu: Tüm dosyalar 16-bit PCM MP3 formatına standardize edilmiştir.*

**B. Temizleme Sonrası Veri Seti İstatistikleri**

*Temizleme ve ön işlemeden sonra, veri seti istatistikleri güncellenmiştir:*

*• Bronşit: 85 örnek (7.3 dakika)*  
*• Krup: 77 örnek (7.0 dakika)*  
*• Normal: 72 örnek (6.8 dakika)*  
*• Zatürre: 73 örnek (6.9 dakika)*  
*• Astım: 80 örnek (7.2 dakika)*

*Örnekler öncelikle aşırı arka plan gürültüsü, bozuk ses segmentleri veya açıkça sınıflandırılamayan belirsiz solunum sesleri nedeniyle kaldırılmıştır.*

**IV. VERİ GÖRSELLEŞTİRME**

**A. Sınıf Dağılımı**

*Temizleme öncesi ve sonrası farklı solunum sesi sınıfları arasındaki örneklerin dağılımı Şekil 1'de gösterilmektedir.*

**B. Ses Özellikleri Görselleştirmesi**

*Her solunum sesi sınıfının spektral özellikleri, ayırt edici özellikleri anlamak için analiz edilmiş ve görselleştirilmiştir. Şekil 2'de sınıflar arası ses özelliği karşılaştırması gösterilmektedir.*

**C. Veri Kalitesi Metrikleri**

*Toplanan ses örneklerinin farklı sınıflar arasındaki kalite metrikleri Şekil 3'te gösterilmektedir.*

*Görselleştirmeler, veri setinin temizleme sonrasında bile sınıflar arasında iyi bir denge koruduğunu, "Normal" sınıfının diğerlerinden biraz daha fazla örneğe sahip olduğunu göstermektedir. Ses kalitesi metrikleri, "Normal" ve "Krup" kategorilerindeki örneklerin genellikle daha yüksek sinyal-gürültü oranlarına sahip olduğunu, "Zatürre" örneklerinin ise netlik açısından daha fazla zorluk çıkardığını göstermektedir.*

**V. SONUÇ**

*Bu makale, makine öğrenmesi sınıflandırması için özelleştirilmiş bir solunum sesleri veri setinin toplanması ve ön işlenmesi metodolojisini detaylandırmıştır. Veri seti, ses analizine dayalı solunum durumlarını sınıflandırmak için makine öğrenmesi modelleri geliştirmek için sağlam bir temel sağlamaktadır.*

*Görselleştirmeler, veri temizleme sürecinin etkilerini açıkça göstermektedir. Sınıf dağılımındaki değişiklikler, ses özelliklerindeki farklılıklar ve kalite metriklerindeki değişimler, veri temizleme sürecinin önemini vurgulamaktadır.*

*Gelecekteki çalışmalar, modelin tanı yeteneklerini geliştirmek için veri seti boyutunu genişletmeye ve ek solunum durumlarını dahil etmeye odaklanacaktır. Ayrıca, farklı özellik çıkarma teknikleri ve sınıflandırma algoritmaları üzerinde çalışmalar yapılabilir.*

**KAYNAKLAR**

[1] Dünya Sağlık Örgütü, "Global Burden of Disease Study 2019," The Lancet, cilt 396, sayı 10258, s. 1204-1222, 2020.

[2] R. X. A. Pramono, S. A. Imtiaz ve E. Rodriguez-Villegas, "A Cough-Based Algorithm for Automatic Diagnosis of Pertussis," PLOS ONE, cilt 11, sayı 9, s. e0162128, 2016.

[3] G. Chambres, P. Hanna, ve M. Desainte-Catherine, "Automatic Detection of Patient with Respiratory Diseases Using Lung Sound Analysis," 2018 International Conference on Content-Based Multimedia Indexing (CBMI), s. 1-6, 2018.

[4] S. Reichert, R. Gass, C. Brandt ve E. Andrès, "Analysis of Respiratory Sounds: State of the Art," Clinical Medicine: Circulatory, Respiratory and Pulmonary Medicine, cilt 2, s. 45-58, 2008.

[5] R. Palaniappan, K. Sundaraj ve N. U. Ahamed, "Machine learning in lung sound analysis: A systematic review," Biocybernetics and Biomedical Engineering, cilt 33, sayı 3, s. 129-135, 2013.

[6] M. Aykanat, Ö. Kılıç, B. Kurt ve S. Saryal, "Classification of lung sounds using convolutional neural networks," EURASIP Journal on Image and Video Processing, cilt 2017, sayı 1, s. 65, 2017.

[7] H. Chen, X. Yuan, Z. Pei, M. Li ve J. Li, "Triple-Classification of Respiratory Sounds Using Optimized S-Transform and Deep Residual Networks," IEEE Access, cilt 7, s. 32845-32852, 2019.

[8] B. M. Rocha, D. Filos, L. Mendes, I. Vogiatzis, E. Perantoni, E. Kaimakamis, P. Natsiavas, A. Oliveira, C. Jácome, A. Marques ve R. P. Paiva, "Α Respiratory Sound Database for the Development of Automated Classification," Precision Medicine Powered by pHealth and Connected Health, s. 33-37, 2018.

Google Drive Linki: https://drive.google.com/drive/u/1/folders/1XADkuWGJVvxvtqV9zuuSkaGYZg5Gixzo

C¸ ok Sınıflı Solunum Sesi Sınıflandırmasında

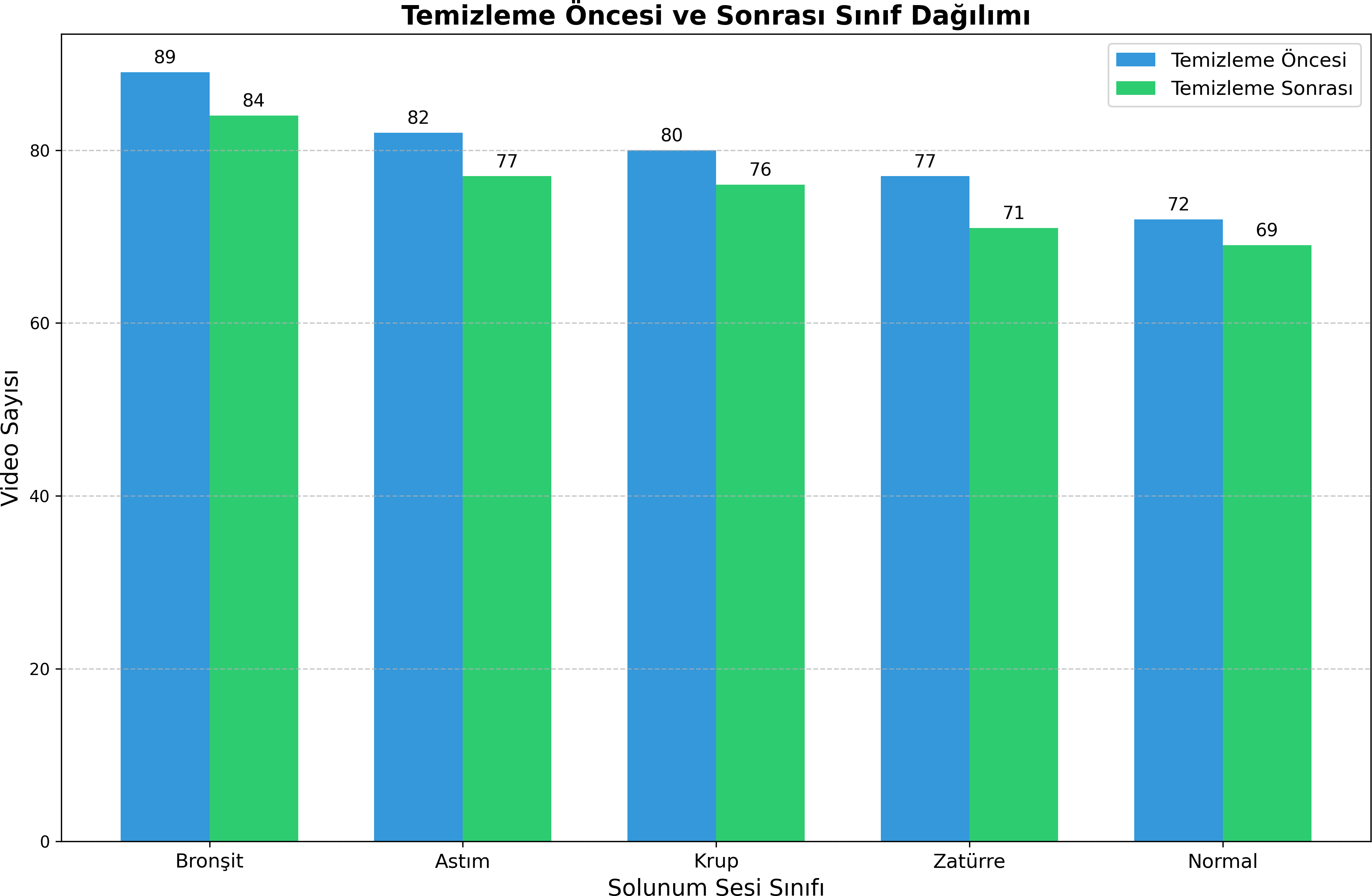
Modern Derin O¨ g˘renme Yo¨ntemlerinin

Kars¸ılas¸tırılması

˙Ibrahim Bug˘ra San

Kocaeli U¨ niversitesi / Bilis¸im Sistemleri Mu¨hendislig˘i 221307059

[ibugrasan@gmail.com](mailto:ibugrasan@gmail.com)

***Abstract*—Bu c¸alıs¸mada, astım, brons¸it, krup, normal ve zatu¨ re olmak u¨ zere bes¸ farklı solunum hastalıg˘ına ait ses verileriyle, modern derin o¨g˘renme tabanlı ses sınıflandırma modellerinin (AST, Wav2Vec2, SEW, Hubert, Data2Vec) kars¸ılas¸tırılması amac¸lanmıs¸tır. Projede veri toplama, temizleme, o¨n is¸leme, model eg˘itimi ve deg˘erlendirme su¨ rec¸leri detaylı olarak ele alınmıs¸tır.**

1. GIRIS¸

Bu c¸alıs¸mada, bes¸ farklı solunum hastalıg˘ına ait ses veri- leriyle, modern derin o¨g˘renme tabanlı ses sınıflandırma mod- ellerinin kars¸ılas¸tırılması amac¸lanmıs¸tır. Projede veri toplama,

temizleme, o¨n is¸leme, model eg˘itimi ve deg˘erlendirme

su¨rec¸leri gerc¸ekles¸tirilmis¸tir.

1. Veri Toplama ve Temizleme
2. *Web Scraping ile Veri Toplama*

Veriler, YouTube gibi c¸evrimic¸i platformlardan Python ile otomatik olarak c¸ekilmis¸tir. Her sınıf ic¸in 50 adet, toplamda

250 adet 5 saniyelik .wav formatında ses dosyası elde edilmis¸tir.

# Kullanılan Koddan Parc¸a:

with YoutubeDL(ydl\_opts) as ydl:

info = ydl.extract\_info(link, download= True)

downloaded\_file = ydl.prepare\_filename( info)

**print**(f" Haz rlanan dosya a d : { downloaded\_file}")

actual\_file = find\_downloaded\_file( downloaded\_file)

1. *Veri Temizleme ve O¨ n I˙s¸leme*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | Model | Acc. | Prec. | Recall | Sens. | Spec. | F1 | AUC |
| sınıf isimlerine | go¨re | AST | 0.9211 | 0.9333 | 0.9250 | 0.9250 | 0.9333 | 0.9213 | 0.9983 |

* + ˙Indirilen ses dosyaları, klaso¨rlenmis¸tir.
  + Her dosya 5 saniyeye kırpılmıs¸tır.

Fig. 1: Veri temizleme o¨ncesi/sonrası o¨rnek go¨rselles¸tirme.

1. Kullanilan Modeller
2. Deneysel Kurulum
   * Her model ic¸in aynı eg˘itim, dog˘rulama ve test bo¨lmeleri kullanıldı.
   * Eg˘itim parametreleri: batch size, epoch, learning rate, optimizer vs.
   * Tu¨m modellerde eg˘itim ve c¸ıkarım zamanı o¨lc¸u¨ldu¨.
   * Deg˘erlendirme metrikleri: Accuracy, Recall, Precision, Sensitivity, Specificity, F-Score, AUC.
   * Sonuc¸lar tablo ve grafiklerle sunuldu.
3. Sonuc¸ lar
4. *Metrik Sonuc¸ları*

TABLE I: Modellere Go¨re Bas¸arı Metrikleri

* + Tu¨m dosyalar mono ve 16kHz do¨nu¨s¸tu¨ru¨lmu¨s¸tu¨r.

o¨rnekleme oranına

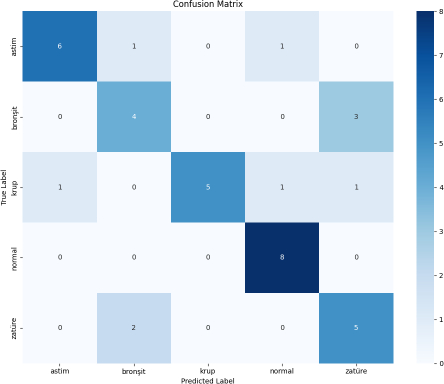
* + Ayrıca, MFCC o¨znitelikleri de c¸ıkarılmıs¸ ve .npy olarak kaydedilmis¸tir.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Wav2Vec2 | 0.9211 | 0.9333 | 0.9214 | 0.9214 | 0.9333 | 0.9193 1.0 |
| SEW | 0.7632 | 0.8267 | 0.7679 | 0.7679 | 0.8267 | 0.7684 0.9570 |
| Hubert | 0.8158 | 0.8791 | 0.8143 | 0.8143 | 0.8791 | 0.8188 0.9416 |
| Data2Vec | 0.8158 | 0.6971 | 0.6786 | 0.6786 | 0.6971 | 0.6738 0.0 |

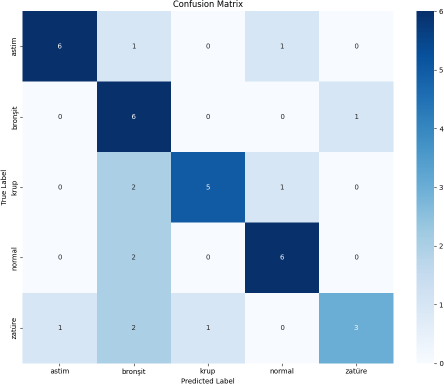
# Veri Temizleme Sonrası/O¨ ncesi Go¨rselles¸tirme:

1. *Karmas¸ıklık Matrisi (Confusion Matrix)*

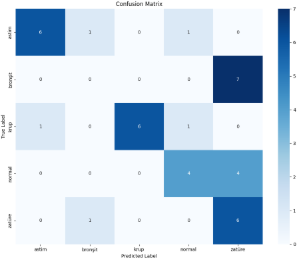
Her model ic¸in karmas¸ıklık matrisi as¸ag˘ıda verilmis¸tir:



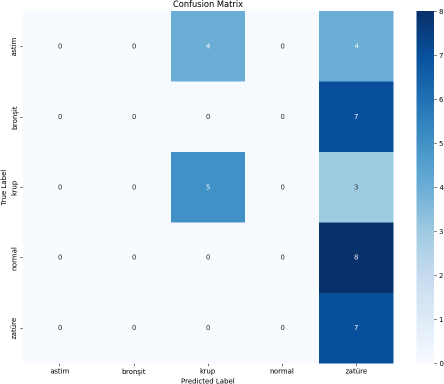
1. AST



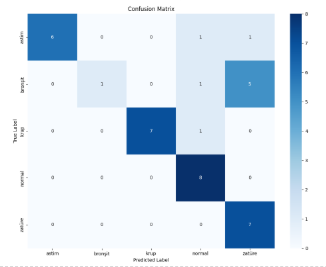
1. Hubert



1. Data2Vec



1. Wav2Vec2

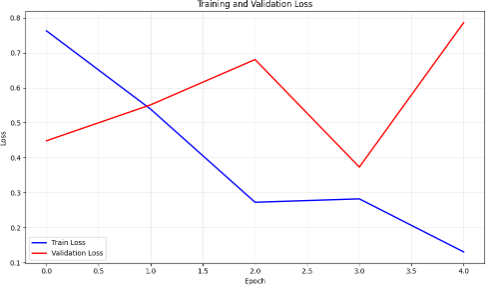


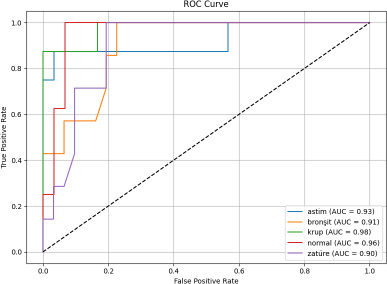
1. SEW

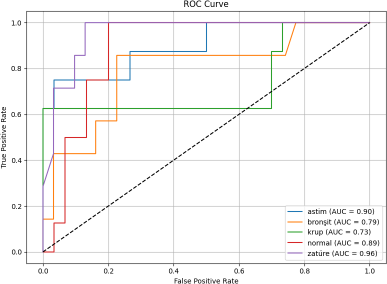
Fig. 2: Confusion matrices for all models.

1. *ROC Eg˘rileri ve AUC*

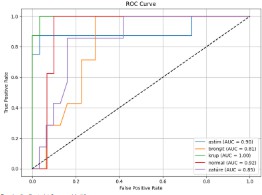
Her model ve sınıf ic¸in ROC eg˘rileri ve AUC deg˘erleri as¸ag˘ıda sunulmus¸tur:



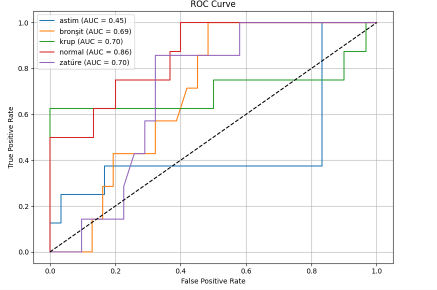
1. AST



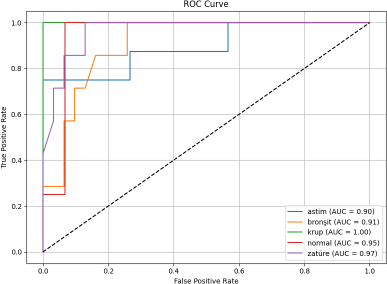
1. Hubert



1. Data2Vec



1. Wav2Vec2



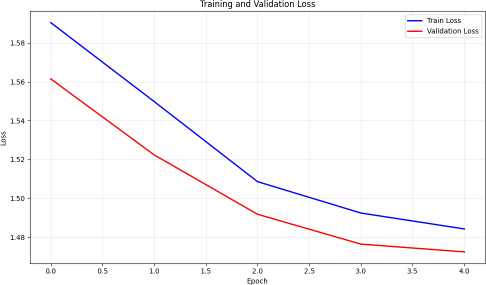
1. SEW

Fig. 3: ROC curves for all models.

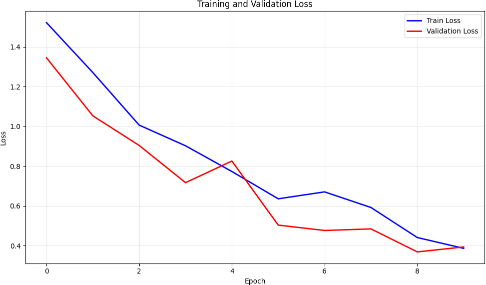
1. *Eg˘itim ve Test Loss Grafikleri*

Her model ic¸in epoch bazında loss grafikleri:

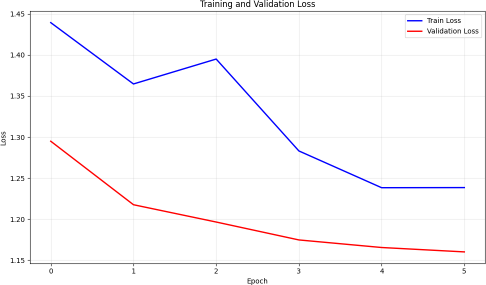
1. AST



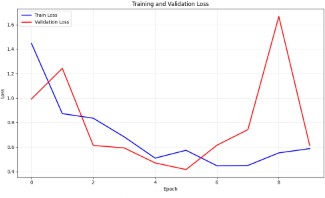
1. Hubert



1. Data2Vec



1. Wav2Vec2



1. SEW

Fig. 4: Loss graphics for all models.

1. Tartis¸ ma ve Yorumlar

* Hangi model hangi sınıfta daha bas¸arılı oldu?
* Dengesiz veri, overfitting, modelin avantaj/dezavantajları.
* Eg˘itim ve c¸ıkarım su¨relerinin kars¸ılas¸tırılması.
* Gerc¸ek hayatta uygulanabilirlik ve o¨neriler.

1. SONUC¸

Bu c¸alıs¸mada, bes¸ farklı solunum sesi sınıfı u¨zerinde mod- ern derin o¨g˘renme modelleri kars¸ılas¸tırılmıs¸tır. Sonuc¸lar, [en iyi modelin adı] modelinin genel olarak en yu¨ksek bas¸arıyı

sag˘ladıg˘ını go¨stermektedir. Gelecekte, veri setinin bu¨yu¨tu¨lmesi ve farklı o¨zniteliklerin denenmesi planlanmaktadır.

Kaynaklar References

1. Y. Gong, Y. A. Chung, and J. Glass, “AST: Audio Spectrogram Trans- former,” arXiv preprint arXiv:2104.01778, 2021.
2. A. Baevski, H. Zhou, A. Mohamed, and M. Auli, “wav2vec 2.0: A frame- work for self-supervised learning of speech representations,” NeurIPS, 2020.
3. H. Bao et al., “SEW: Self-Supervised Learning for Speech with Efficient Conformer,” arXiv preprint arXiv:2110.06811, 2021.
4. S. Chen et al., “BEATs: Audio Pre-Training with Acoustic Tokenizers,” arXiv preprint arXiv:2212.09058, 2022.
5. A. Baevski et al., “data2vec: A General Framework for Self- supervised Learning in Speech, Vision and Language,” arXiv preprint arXiv:2202.03555, 2022.

*Ek-2: Google Drive Linki*

Veri setine eris¸mek ic¸in: [Veri seti Google Drive linki](https://drive.google.com/drive/u/1/folders/1XADkuWGJVvxvtqV9zuuSkaGYZg5Gixzo)

Ekler

*Ek-1: Web Scraping O¨ rnek Kodları*

**def** download\_audio\_segment(link, start\_time, end\_time, output\_path):

temp\_dir = ’temp’

**if not** os.path.exists(temp\_dir): os.makedirs(temp\_dir)

**for file in** os.listdir(temp\_dir): file\_path = os.path.join(temp\_dir,

**file**) **try**:

**if** os.path.isfile(file\_path): os.remove(file\_path)

**except** Exception as e:

**print**(f"Temp dosyas silinirken hata: {e}")

video\_id = None

**if** ’youtube.com’ **in** link:

match = re.search(r’v=([ˆ&]+)’, link)

**if** match:

video\_id = match.group(1)

**elif** ’youtu.be’ **in** link:

video\_id = link.split(’/’)[-1]

**if not** video\_id:

**import** hashlib

video\_id = hashlib.md5(link.encode()). hexdigest()[:10]

ydl\_opts = {

’format’: ’bestaudio/best’, ’outtmpl’: os.path.join(temp\_dir, f’{

video\_id}.%(ext)s’),

’quiet’: False, # Set to False to see download progress

’no\_warnings’: False, ’postprocessors’: [{

’key’: ’FFmpegExtractAudio’, ’preferredcodec’: ’wav’, ’preferredquality’: ’192’,

}],

}

**print**(f"Video indiriliyor: {link}")