Avrupa Bilim ve Teknoloji Dergisi
Özel Sayı, 34, S. 492-495, Mart 2022
© Telif hakkı EJOSAT'a aittir
Araştırma Makalesi

www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
Special Issue 34, pp. 492-495, March 2022
Copyright © 2022 EJOSAT
Research Article

Öksürük Sesi Kayıtlarından Spektral Özellikler ile Otomatik COVID-19 Tespiti

Semiye Demircan

Bilgisayar Mühendisliği/Konya Teknik Üniversitesi, Türkiye (ORCID: 0000-0001-8570-2412), sdemircan@ktun.edu.tr

(2nd International Conference on Applied Engineering and Natural Sciences ICAENS 2022, March 10-13, 2022)

(**DOI:** 10.31590/ejosat.1083052)

ATIF/REFERENCE: Demircan, S. (2022) Öksürük Sesi Kayıtlarından Spektral Özellikler ile Otomatik COVID-19 Tespiti. Avrupa Bilim ve Teknoloji Dergisi, (34), 492-495.

Öz

COVID-19 pandemisi son iki yıldır dünyada hızla yayılmış ve bu alanda yapılan çalışmalar da artmıştır. COVID-19 olan hastaların, hasta olmayanlardan ayırt edilmesi de pandemideki en önemli sorunlardan bir tanesidir. Gerek hastalığın erken teşhisi gerekse hasta olmayanlara bulaşma riski açısından COVID-19'un otomatik tespiti oldukça önem arz etmektedir. Hastalığın teşhisinde farklı semptomların görülebilmesi ve hatta hiç semptom görülmeden bile oluşabilmesi teşhisi çok daha zor hale getirmiştir. Bu durum hastalığın teşhisi konusunda yapılan çalışmaları arttırmıştır. Öksürük ses kayıtları gibi solunum kayıtlarında var olan önemli özellikler kullanılarak teşhis yapılabilmesi de bu uygulamalardan bir tanesidir. Bu çalışmada öksürük ses kayıtları kullanılarak otomatik COVID-19 hastalık tespiti yapılmıştır. "COVID-19 Positive and Negative Patients' Cough Recordings" (HIMANSHU) veri seti kullanılarak gerçekleştirilen çalışmada ilk olarak ses dosyalarından Mel-Frekansı Kepstrum Katsayıları (MFCC) çıkarılmıştır. Farklı sayıda olan MFCC öznitelikleri istatistiksel değerler kullanılarak eşit boyutlu hale getirilmiştir. MFCC yöntemi ile elde edilen spektral özellikler 8, 16, 32, 64 tane olacak şekilde dört farklı uzunlukta katsayılar çıkarılmıştır. Son olarak makine öğrenmesi algoritmalarından Yapay Sinir Ağları (YSA), Naive Bayes (NB), K-en Yakın Komşu Algoritması (kNN), Rastgele Orman (RO) algoritmaları kullanılarak hastalık teşhisi yapılmıştır. Yapılan çalışmada COVID veya COVID-DEGİL şeklinde 2 sınıf kullanılmıştır. Uygulama on çapraz doğrulama yöntemi ile çalıştırılmıştır. Çalışma sonunda en yüksek sınıflandırma başarası kNN algoritması ile % 99.39 olarak gerçekleştirilmiştir.

Anahtar Kelimeler: Covid-19 Tespiti, Öksürük Sesi, MFCC, YSA, Spektral Özellikler

Automatic COVID-19 Detection with Spectral Features from Cough Sound Recordings

Abstract

The COVID-19 pandemic has spread rapidly around the world in the last two years, and studies in this area have also increased. Distinguishing patients COVID-19 or not is one of the most important problems in the pandemic. Automatic detection of COVID-19 is very important in terms of both early diagnosis of the disease and the risk of transmission to non-patients. The fact that different symptoms can be seen in the diagnosis of the disease and even occur without any symptoms has made the diagnosis much more difficult. So, the studies have focused on the diagnosis of the disease. Diagnosis can be made by using the important features of respiratory recordings, such as cough sound recordings. In this study, automatic COVID-19 disease detection was performed using cough voice recordings. In the study carried out using the "COVID-19 Positive and Negative Patients' Cough Recordings" (HIMANSHU) dataset, firstly, Mel-Frequency Cepstrum Coefficients (MFCC) were extracted from audio files. Different numbers of MFCC features were made equal in size using statistical values. The spectral features obtained by the MFCC method are 8, 16, 32, 64, and coefficients of four different lengths have been extracted. Finally, the disease diagnosis was made using Artificial Neural Networks (ANN), Naive Bayes (NB), K-Nearest Neighbor Algorithm (kNN), Random Forest (RF) algorithms from machine learning algorithms. In the study, 2 classes were used as COVID or NOT-COVID. The application was run with ten cross validation methods. Finally, the highest classification success was achieved with the kNN algorithm as 99.39%.

Keywords: Detection COVID-19, Cough Sound, MFCC, Spectral Features, ANN

[1] Sorumlu Yazar: sdemircan@ktun.edu.tr

1. Giriş

Covid-19, şiddetli akut solunum sendromu Coranavirus2 (SARS-CoV-2) yeni bir patajeni olan Corona Virüsü 2019 (COVID-19) Kasım 2019'un sonudna ortaya çıktı ve Dünya Sağlık Örgütü (World Health Organization - WHO) tarafından 11 Mart 2020'de pandemi ilan edildi (*Coronavirus disease 2019 (covid-19): situation report 162*, 2020; Dong, Du, & Gardner, 2020). Pandemi başlangıcından bu yana 430.257.564 kişiyi etkileyen Coronavirus 27 Şubat 2022 tarihi itibariyle 5.922.047 insanın ölümüne yol açtı ("WHO Coronavirus (COVID-19) Dashboard,"). Bir milyona yaklaşan ölüm sayısı ile Amerika Birleşik Devletleri (ABD) en çok etkilenen ülkeler arasında ilk sıradadır ("WHO Coronavirus (COVID-19) Dashboard,"). Hastalığın ilerlemesindeki en büyük faktör hastalığın farklı semptomlar göstermesinin yanı sıra bazı hastaların hiç semptom göstermemesidir. Ateş, inatçı öksürük, karın ağrısı, göğüs ağrısı ve ses kısıklığı görülen semptomlar arasındadır(Menni et al., 2020). Hastalığın teşhisinde kullanılan en yaygın yöntem PCR testidir. PCR testinin hızlı sonuç vermesi bir avantaj olsa da test sayısındaki sınırlamalar, maliyet ve hastaların hastaneye gitmek zorunda olmaları gibi dezavantajları da vardır. Bu dezavantajlardan dolayı COVID-19 tespiti için farklı yollar araştırılmıştır. Bunlardan biri de öksürük sesinden hastalık teşhisidir.

Öksürük genellikle kapalı bir nefes borusunun ağzına karşı zorunlu bir manevrayla oluşan karakteristik bir sestir (Morice et al., 2004). COVID-19'un da belirtileri arasında yer alan öksürük, bilim insanlarını hastalığın teşhisine dair araştırmalara konu olmuştur.

İslam ve ark. (Islam, Abdel-Raheem, & Tarique, 2022) yaptıkları çalışmada öksürük sesi örnekleri ve derin bir sinir ağı kullanarak COVID-19'un otomatik teşhisi için bir algoritma geliştirmeyi içeren bir çalışma sunmuşlardır. Önerilen algoritmada öksürük sesinden akustik özellikler çıkarılarak bir öznitelik vektörü oluşturulmuştur. Daha sonra Derin Sinir Ağı (Deep Neural Network -DNN) algoritması kullanılarak sınıflandırma işlemi yapılmış ve %97.5 sınıflandırma başarısı elde edilmiştir.

Tena ve ark. (Tena, Clarià, & Solsona, 2022) COVID-19'un otomatik tespiti için ücretsiz olarak erişilebilen, hızlı ve verimli bir metodoloji tasarlamayı amaçlamışlardır. Öksürük sesinden zaman-frekans özelliklerini çıkararak elde ettikleri öznitelik kümesine Rastgele Orman (RO) Algoritması uygulayarak sınıflandırma işlemi yapmışlardır. Öksürük sesinden otomatik COVID-19 teşhis etmeyi amaçlayan bu çalışma da %90'a yakın başarı elde etmişlerdir.

Melek (Mesut, 2021) çalışmasında, öksürük sesini analiz ederek COVID-19 olan ve COVID-19 olmayan hastaların ayırımı yapmıştır. Mel-Frekansı Kepstrum Katsayıları (MFCC) yöntemi kullanılarak öznitelikler elde etmiştir. Tek-çıkışlı çapraz doğrulama (leave-one-out cross-validation, LOO-CV) stratejisi uygulanarak k-en yakın komşuluk (kNN) algoritması ile sınıflandırmıştır. Sıralı ileri seçim (SFS) yöntemi uygulanarak yapılan çalışmada %0.98 sınıflandırma başarısı elde etmiştir.

Mohanad ve ark. (Alkhodari & Khandoker, 2022) çalışmalarında, 120 tane Covid-19 ve 120 tane Covid-19 olmayan sağlıklı kişiler tarafından kaydedilen nefes sesi kayıtlarının bulunduğu, Caswara veri tabanını kullandılar. MFCC katsayıları çıkarıldıktan sonra tanıma için CNN-BİLSTM derin öğrenme yöntemini önermişlerdir. Önerilen yöntemle %94.58 sınıflandırma başarısı elde etmişlerdir.

Literatürde de özetlendiği gibi COVID-19 tespitinde öksürük sesinin önemi açıktır. Bu çalışmada da "COVID-19 Cough Recordings COVID-19 Positive and Negative Patients' Cough Recordings" başlıklı HIMANSHU tarafından oluşturulan ve en son 9 ay önce update edilen "Kaggle" veri seti (*COVID-19 Cough Recordings / COVID-19 Positive and Negative Patients' Cough Recordings*, 2021) kullanılmıştır. Toplam 166 veri bulunan bu veri setinden MFCC yöntemi ile çıkarılan özelliklere istatistik değerleri uygulanmıştır. Elde edilen veri seti Yapay Sinir Ağları (YSA), Naive Bayes (NB), K-en Yakın Komşu Algoritması (kNN), Rastgele Orman (RO) Algoritmaları ile sınıflandırılmıştır. Sonuçların kabul edilebilir AUC değerine sahip olduğu gözlemlenmiştir.

Bu çalışma şu şekilde planlanmıştır: Kullanılan materyal ve yöntemler ikinci bölümde, gerçekleştirilen performans bulgular ve tartışma üçüncü bölümde sonuç ise dördüncü bölümde sunulmuştur.

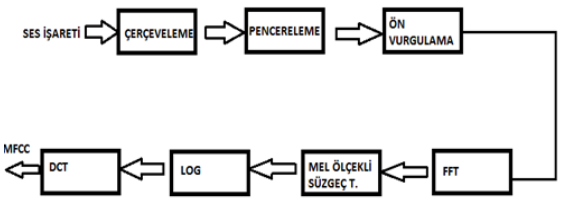
2. Materyal ve Yöntem

2.1. Veri Seti

Bu çalışmada HIMANSHU tarafından oluşturulan açık erişimli COVID-19 öksürük kayıtları (*COVID-19 Cough Recordings / COVID-19 Positive and Negative Patients' Cough Recordings*, 2021) kullanılmıştır. Bu veri setinde 19 tane COVID-19 hastası ve 147 tane COVID-19 olmayan kişilere ait öksürük ses kayıtları bulunmaktadır. Veri tabanı 2021 de oluşturulmuş ve en son 9 ay önce güncellenmiştir. Veri tabanında bulunan "trial_covid" ses kayıtları kullanılarak elle özellikler çıkarılmıştır.

2.2. Mel -Frekansı Kepstrum Katsayıları (MFCC)

Ses sinyallerinden özellik elde edilirken, en çok kullanılan spektral özelliklerin başında Mel Frekansı Kepstrum Katsayıları (MFCC) gelir. MFCC özellikleri insan kulağını modellemektedir (Becchetti & Prina-Ricotti, 2004). MFCC yöntemi öksürük, nefes ses kayıtlarından COVID-19 teşhisi yapan uygulamalarda da oldukça başarılı sonuçlar vermektedir(Bansal, Pahwa, & Kannan, 2020; Zhou et al., 2021). MFCC katsayılarını elde ederken uygulanan işlem adımları sırasıyla: Çerçeveleme, Pencereleme, Ön Vurgulama, Hızlı Fourier Dönüşümü (FFT), Mel Ölçekli Süzgeç Takımı, Logaritma Alma (LOG), Ayrık Kosinüs Dönüşümü (DCT) işlemleridir (Becchetti & Prina-Ricotti, 2004). (Şekil 1)



Şekil 1. MFCC işlem adımları

Çerçevelemede giriş işareti, M örnekten oluşan kısımları örtüşen N örnek uzunluğunda konuşma parçalarına bölünür (M<N). Çerçeveleme işleminden sonraki adım olan pencereleme işleminde amaç sinyalin başındaki ve sonundaki süreksiz

kısımları azaltmak, dolayısıyla sinyalin başındaki ve sonundaki bilgi içermeyen bölümleri bastırarak spektral bozulmayı engellemektir. Hamming pencere fonksiyonu kullanılarak pencereleme işlemi gerçekleştirilir. Hamming pencere fonksiyonu Denklem 1 de gösterilmektedir. (Becchetti & Prina-Ricotti, 2004)

$$w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right), \quad 1 \le n \le N \tag{1}$$

Ön vurgulama işleminde giriş işareti birinci dereceden bir Sonlu Dürtü Yanıtı Filtesi-FIR süzgeç girişine uygulanır. Amaç sinyalin yüksek frekans bileşenlerini daha baskın hale getirmektir. N örnekten oluşan konuşma parçasını zaman domeninden, frekans domenine çevirmek için Hızlı Fourier Dönüşümü (FFT) uygulanır. Konuşma sinyallerinin frekans ortamındaki içeriklerinin doğrusal ölçekli olmadığı akustik çalışmalar sonucunda ortaya konulmuştur 1 kHZ frekanslı bir sesin, insan kulağının algısal duyma eşiğinin 40 dB yukarısı 1000 mel kHz olarak tanımlanır. Diğer değerler referans sese göre ayarlanır Mel ölçeği 1 kHz'e kadar doğrusal, 1 kHz'den sonra ise logaritmik olarak değişen aralıklarla ifade edilen bir ölçektir. Verilen bir f (Hz) frekansını mel frekansı ölçeğinde ifade etmek için Denklem 2 kullanılır: (Becchetti & Prina-Ricotti, 2004)

$$mel(f) = 2595\, log\left(1 + \frac{f}{700}\right) \tag{2}$$

2.3. Sınıflandırıcılar

Bu çalışmada hastalık teşhisi için dört farklı sınıflandırıcı kullanılmıştır. Bunlardan birincisi Yapay Sinir Ağları (YSA)'dır. YSA insan beynin çalışna mantığından esinlenerek geliştirilmiş, ağırlıklı bağlantılar aracılığıyla birbirine bağlanan ve her biri kendi belleğine sahip işlem elemanlarından oluşan paralel ve dağıtılmış bilgi işleme yapılarıdır (Elmas, 2003).

İkinci olarak K-en Yakın Komşu Algoritması (kNN) kullanılmıştır. kNN algoritması bir örnek tabanlı öğrenme biçimidir, yani yeni gözlemler sınıflanırken hafızada tutulan bir veri kümesi kullanılır. Bu hafızadaki veri kümesinde hedef değişkenin alacağı değerler bellidir. (Akküçük, 2011)

Üçüncü olarak en iyi bilinen Bayes öğrenme metodu olan Sade Bayes (NB) öğrenme algoritması kullanılmıştır.

Son sınıflandırma yöntemi olan Rastgele Orman (RO) algoritması denetimli bir sınıflandırma algoritmasıdır.

3. Uygulama- Bulgular ve Tartışma

Bu çalışmada HIMANSHU tarafından oluşturulan açık erişimli COVID-19 öksürük kayıtları (*COVID-19 Cough Recordings / COVID-19 Positive and Negative Patients' Cough Recordings*, 2021) kullanıldı. Özellik veri seti MFCC yöntemi kullanılarak elde edildi. Farklı boyutlarda MFCC katsayıları oluşturuldu.

Uygulamada ilk olarak verilerden MFCC öznitelikleri çıkartılmıştır. MFCC özellikleri 8, 16, 32, 64 adet olacak şekilde dört farklı uzunlukta oluşturulmuştur. Her bir veriden 8 Mel Katsayısı çıkarılarak elde edilen veri seti Veri-Seti-1, 16 Mel Katsayısı çıkarılarak elde edilen veri seti Veri-Seti-2, 32 Mel Katsayısı çıkarılarak elde edilen veri seti Veri-Seti-3, 64 Mel Katsayısı çıkarılarak elde edilen veri seti Veri-Seti-4 olarak isimlendirilmiştir. Elde edilen katsayılar farklı boyuttadır. Bu durumda katsayılar sınıflandırmaya giriş olarak verilirken sorun oluşturacağından ikinci adım olarak her bir katsayı için istatistiksel değerler kullanılarak boyut eşitleme işlemi yapılmıştır. Kullanılan istatistiksel özellikler maksimum, minimum, ortalama, standart sapma, skewness, kurtosis ve median değerleridir. (Semiye Demircan, 2020)(S. Demircan & Ornek, 2020)

Oluşturulan veri setleri dört farklı sınıflandırıcı ile sınıflandırılmıştır. Uygulamalarda 10 çapraz doğrulama yöntemi kullanılmıştır. Her bir sınıflandırıcı sonucunda elde edilen Sınıflandırma Doğruluğu (SD), Eğrinin Altında Kalan Alan (ing. Area Under the Curve- AUC) ve Kappa İstatistiği değeri Tablo-1'de verilmiştir.

Doğru sınıflandırılmış pozitif örneklerin yanlış pozitiflere oranı ROC (ing. Receiver operating Characteristic) eğrisi olarak tanımlanır. AUC ise bu eğrinin altında kalan alan hesabı ile ifade edilir. 0-1 arasında bir değerdir. AUC değeri ne kadar büyük ise kullanılan modelin farklı sınıflara ayırmadaki başarısı da o kadar büyüktür.

Tablo 1. Öksürük se kayıtlarında elde edilen sınıflandırma Doğruluğu

Veri Kümesi	Sınıflandırıcı	SD (%)	AUC (%)	Kappa ist.
Veri-Seti-1	YSA	97.59	0.99	0.88
	NB	96.98	0.98	0.85
	kNN	99.39	0.99	0.97
	RO	95.18	0.88	0.76
Veri-Seti-2	YSA	98.19	0.99	0.91
	NB	96.98	0.99	0.84
	kNN	99.39	0.99	0.97
	RO	98.79	0.99	0.93
Veri-Seti-3	YSA	98.19	0.99	0.91
	NB	92.17	0.89	0.65
	kNN	99.39	0.99	0.97
	RO	96.38	0.99	0.79
Veri-Seti-4	YSA	98.19	0.99	0.91
	NB	96.39	0.92	0.80
	kNN	99.39	0.99	0.97
	RO	95.18	0.88	0.75

Kappa katsayısı 0-1 değerleri arasında değişen istatistiksel bir değerdir. Buna göre kappa = 1 tam uyumu, kappa = 0 ise zayıf uyumu temsil etmektedir. Kappa değeri 0.4'ün üzerinde ise kappa istatistiği kabul edilebilir bir anlaşmayı ifade eder. (Landis & Koch, 1977)

Tablo 1'de görüldüğü gibi en yüksek SD değeri %99.39 ile kNN sınıflandırıcı ile elde edilmiştir. Elde edilen sonuçlara göre YSA ile %98.19, NB ile %96.39 RO ile % 98.79 en iyi SD değerleri olarak gözlemlenmiştir. AUC değerlerine baktığımızda performans sonuçlarındaki ROC eğrisi altıda kalan alanın 0.88 ve üstü olduğu görülmektedir. Bu da modelin farklı sınıflara ayırmadaki başarısını ortaya koymaktadır. Kappa istatistiği değerlerinin 0.4'ün üzerinde olduğundan kabul edilebilirdir.

Gerçekleştirilen çalışmada öksürük ses kayıtlarından spektral özelliklerden biri olan MFCC özellikleri çıkarılarak, COVID-19 olan ve COVID-19 olmayan hastaların ayrımının yapılabileceği gösterilmiştir.

4. Sonuç

COVID-19 hastalığının otomatik tespitinde yapılan çalışmalara bakıldığında az bir maliyetle ve hastalığın yayılmasını önlemek adına teması en aza indiren hızlı bir yöntem olarak öksürük ses kayıtlarından hastalık teşhisi oldukça önemlidir. Bu çalışmada HIMANSHU tarafından oluşturulan COVID-19 veri tabanındaki öksürük sesi kayıtları kullanılarak spektral özellikler elde edilmiştir. Spektral özellik olarak MFCC öznitelikleri seçilmiştir. Farklı sayıda Mel katsayıları elde edilerek sunulan çalışmada Mel katsayı sayısının sınıflandırma doğruluğuna etkisi incelenmiştir. Dört farklı sınıflandırıcı kullanılmıştır. Bu veri tabanı üzerinde elle çıkarılan MFCC öznitelikleriyle yapılan ilk çalışmadır. Sonraki çalışmalarda spektral özelliklerin yanında prosodi özelliklerin ve frekans değerlerinin de sınıflandırmaya dahil edilerek hastalık teşhisine etkisi incelenecektir.

Kaynakça

Akküçük, U. (2011). *Veri Madencilği - Kümeleme ve Sınıflandırma Algoritamaları* (Vol. 1). İstanbul: Yalın Yayıncılık.

Alkhodari, M., & Khandoker, A. H. (2022). Detection of COVID-19 in smartphone-based breathing recordings: A pre-screening deep learning tool. *PLOS ONE, 17*(1), e0262448. doi:10.1371/journal.pone.0262448

Bansal, V., Pahwa, G., & Kannan, N. (2020, 2-4 Oct. 2020). *Cough Classification for COVID-19 based on audio mfcc features using Convolutional Neural Networks.* Paper presented at the 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON).

Becchetti, C., & Prina-Ricotti, L. (2004). *Speech Recognition;theory an C++ Implementation*: John Wiley &Sons.

Coronavirus disease 2019 (covid-19): situation report 162. (2020). Retrieved from

COVID-19 Cough Recordings / COVID-19 Positive and Negative Patients' Cough Recordings. (2021). Retrieved from: https://www.kaggle.com/himanshu007121/coughclassifier-trial

Demircan, S. (2020). *Duygu Sınıflandırma Problemine Etmen Tabanlı Yaklaşım.* (PHD), Konya Teknik Üniversitesi , Lsiasüstü Eğitim Enstitüsü.

Demircan, S., & Ornek, H. K. (2020). Comparison of the Effects of Mel Coefficients and Spectrogram Images via Deep Learning in Emotion Classification. *Traitement Du Signal, 37*(1), 51-57. doi:10.18280/ts.370107

Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis, 20*(5), 533-534. doi:10.1016/S1473-3099(20)30120-1

Elmas, Ç. (2003). *Yapay Sinir Ağları (Kuram, Mimari, Eğitim, Uygulama)* (Vol. 1). Ankara: Seçkin Yayıncılık.

Islam, R., Abdel-Raheem, E., & Tarique, M. (2022). A study of using cough sounds and deep neural networks for the early detection of Covid-19. *Biomedical Engineering Advances, 3*, 100025. doi:https://doi.org/10.1016/j.bea.2022.100025

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics, 33*(1), 159-174. doi:10.2307/2529310

Menni, C., Valdes, A. M., Freidin, M. B., Sudre, C. H., Nguyen, L. H., Drew, D. A., . . . Spector, T. D. (2020). Real-time tracking of self-reported symptoms to predict potential COVID-19. *Nature Medicine, 26*(7), 1037-1040. doi:10.1038/s41591-020-0916-2

Mesut, M. (2021). Diagnosis of COVID-19 and non-COVID-19 patients by classifying only a single cough sound. *Neural Computing & Applications, 33*(24), 17621-17632.

Morice, A. H., Fontana, G. A., Sovijarvi, A. R. A., Pistolesi, M., Chung, K. F., Widdicombe, J., . . . Kastelik, J. (2004). The diagnosis and management of chronic cough. *European Respiratory Journal, 24*(3), 481-492. doi:10.1183/09031936.04.00027804

Tena, A., Clarià, F., & Solsona, F. (2022). Automated detection of COVID-19 cough. *Biomedical Signal Processing and Control, 71*, 103175. doi:https://doi.org/10.1016/j.bspc.2021.103175

WHO Coronavirus (COVID-19) Dashboard. Retrieved from https://covid19.who.int/

Zhou, Q., Shan, J. H., Ding, W. L., Wang, C. Y., Yuan, S., Sun, F. C., . . . Fang, B. (2021). Cough Recognition Based on Mel-Spectrogram and Convolutional Neural Network. *Frontiers in Robotics and Ai, 8*. doi:ARTN 58008010.3389/frobt.2021.580080