



**T.C.
HARRAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

**DERİN AĞ VE MFCC ÖZELLİKLERİNE DAYALI AKUT SOLUNUM
YOLU HASTALIKLARININ SINIFLANDIRILMASI**

İSMAİL BALIK

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

**Şanlıurfa
2025**



**T.C.
HARRAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

**DERİN AĞ VE MFCC ÖZELLİKLERİNE DAYALI AKUT SOLUNUM
YOLU HASTALIKLARININ SINIFLANDIRILMASI**

İSMAİL BALIK

**BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
Tez Danışmanı: Dr. Öğr. Üyesi MEHMET BİLAL ER**

**Şanlıurfa
2025**

TEŞEKKÜR

Bana yüksek lisans eğitimim süresince danışmanlık yaptığından, bilgilerini ve deneyimlerini aktardığından, tez konusunun seçimi ve yürütülmesi konusundaki yardımları, tezle ilgili olan çalışmalarına katkıları ve yakın ilgisinden dolayı tez danışmanım **Sayın Dr. Öğr. Üyesi Mehmet Bilal ER'e**, tez jürimde görev alan hocalarıma, ders döneminde ders aldığım tüm hocalarıma ve bana her zaman desteklerini esirgemeyen ailemdeki herkese ayrı ayrı teşekkür ederim.

İÇİNDEKİLER

ÖZET	i
ABSTRACT	ii
ŞEKİLLER DİZİNİ	iii
ÇİZELGELER DİZİNİ	iv
KISALTMALAR	v
1. GİRİŞ	1
2. ÖNCEKİ ÇALIŞMALAR	14
3. GEREÇ VE YÖNTEM	26
3.1. Veri Seti	26
3.2. Model Mimarisi	28
3.3. Mel Frekans Kepstrum Katsayıları (MFCC) ile Özellik Çıkarımı	30
3.4. Evrişimli Sinir Ağları (Convolutional Neural Networks - CNN)	32
3.5. Tek Boyutlu Evrişimsel Sinir Ağları (One Dimensional Convolutional Neural Networks)	34
3.6. LSTM ile Özellik Çıkarımı	36
3.7. Değerlendirme metrikleri	39
4. BULGULAR	41
5. TARTIŞMA	53
6. SONUÇLAR	55
7. ÖNERİLER	56
KAYNAKLAR	57
ÖZGEÇMİŞ	67

ÖZET

YÜKSEK LİSANS TEZİ

DERİN AĞ VE MFCC ÖZELLİKLERİNE DAYALI AKUT SOLUNUM YOLU HASTALIKLARININ SINIFLANDIRILMASI

İSMAİL BALIK

HARRAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

Tez Danışman: Dr. Öğr. Üyesi MEHMET BİLAL ER
Yıl: 2025, Sayfa : 67

Solunum yolu hastalıklarının erken teşhisi, hastalık yönetimi açısından kritik öneme sahiptir. Bu çalışmada, solunum seslerinin otomatik analizi için derin öğrenme tabanlı bir sınıflandırma sistemi geliştirilmiştir. Ses verilerinden Mel Frekans Kepstrum Katsayıları (MFCC) gibi spektral özellikler çıkarılmış ve bu özellikler, hibrit model CNN+LSTM katmanları içeren bir derin sinir ağı modeli ile işlenmiştir. Sınıf dengesizliğini azaltmaya yönelik teknikler kullanılarak modelin genellenebilirliği artırılmıştır. Yapılan deneysel çalışmalar sonucunda, geliştirilen model test verisi üzerinde %86 doğruluk oranına ulaşmış ve solunum seslerinden hastalık tespitinde etkili ve güvenilir bir performans sergilemiştir. Elde edilen bulgular, ses temelli analizlerin akut solunum yolu hastalıklarının erken tanısına katkı sağlayabileceğini göstermektedir.

ANAHTAR KELİMELER: Ses sınıflandırma, Solunum yolu hastalıkları, Derin öğrenme, CnnLstm

ABSTRACT

MASTER THESIS

CLASSIFICATION OF ACUTE RESPIRATORY DISEASES BASED ON DEEP NETWORK AND MFCC FEATURES

İSMAİL BALIK

**HARRAN UNIVERSITY
INSTITUTE OF GRADUATE EDUCATION
DEPARTMENT OF COMPUTER ENGINEERING**

**Thesis Supervisor: Assist. Prof. Dr. MEHMET BİLAL ER
Year: 2025, Page : 67**

Early diagnosis of respiratory diseases is critical for disease management. In this study, a deep learning-based classification system was developed for the automatic analysis of respiratory sounds. Spectral features such as Mel Frequency Cepstrum Coefficients (MFCC) were extracted from the audio data and processed using a deep neural network model consisting of hybrid CNN + LSTM layers. Techniques to reduce class imbalance were used to increase the generalizability of the model. Experimental studies have shown that the developed model achieved 86% accuracy on test data and demonstrated effective and reliable performance in disease detection from respiratory sounds. The findings demonstrate that audio-based analysis can contribute to the early diagnosis of acute respiratory diseases.

KEYWORDS: Sound classification, Respiratory diseases, Deep learning, CnnLstm

ŞEKİLLER DİZİNİ

Şekil 3.6.	Bronşit, Koah, Pnömoni, Sağlıklı ve Astım Orijinal Sinyalleri.	29
Şekil 3.7.	MFCC ve CNN+LSTM ile Önerilen Modelin Mimarisi	29
Şekil 4.2.	CNN Algoritmasının Hastalık Bazlı Sınıflandırma Performansı	41
Şekil 4.2.	CNN Algoritmasının Genel Performans Değerleri	42
Şekil 4.2.	CNN Modelin Karışıklık Matrisi	43
Şekil 4.2.	CNN Sınıf Bazlı Doğru Ve Yanlış Tahmin Grafiği	44
Şekil 4.2.	CNN için Eğitim/doğrulama verilerindeki doğruluk/kayıp grafikleri	45
Şekil 4.2.	LSTM Algoritmasının Hastalık Bazlı Sınıflandırma Performansı	45
Şekil 4.2.	LSTM Algoritmasının Genel Performans Değerleri	46
Şekil 4.2.	LSTM Model için Karışıklık Matrisi	47
Şekil 4.2.	LSTM Model için Sınıf Bazlı Doğru Ve Yanlış Tahmin Grafiği	48
Şekil 4.2.	LSTM Model için Eğitim/doğrulama verilerindeki doğruluk/kayıp grafikleri	48
Şekil 4.2.	Hibrit Model CNN+LSTM Algoritmasının Hastalık Bazlı Sınıflandırma Performansı	49
Şekil 4.2.	Hibrit Model CNN+LSTM için Karışıklık Matrisi	50
Şekil 4.2.	Hibrit Model CNN+LSTM için Sınıf Bazlı Doğru Ve Yanlış Tahmin Grafiği	51
Şekil 4.2.	Hibrit Model CNN+LSTM için Eğitim/doğrulama verilerindeki doğruluk/kayıp grafikleri	51

ÇİZELGELER DİZİNİ

Çizelge 3.1.	Veri setindeki ses sayısı	26
Çizelge 3.1.	Önerilen Çerçevenin Sahte Kodu	27
Çizelge 3.1.	Evrişimli sinir ağının tipik mimarisi	33
Çizelge 3.1.	CNN Önerilen Derin Model	33
Çizelge 3.1.	Conv1D max pooling	35
Çizelge 3.1.	Conv1D Önerilen Derin Model Adımları	36
Çizelge 3.2.	Uzun kısa süreli belleğin iç yapısı.	37
Çizelge 3.3.	LSTM Modelin Adımları-Akış Şeması	38
Çizelge 3.4.	Önerilen CNN-LSTM otomatik kodlayıcı ağı	39
Çizelge 5.1.	Mevcut yöntemlerle performans karşılaştırması.	53

KISALTMALAR

BiLSTM	Bidirectional Long Short-Term Memory (Çift Yönlü Uzun Kısa Süreli Bellek)
CNN	Convolutional Neural Network (Evrişimsel Sinir Ağları)
Conv1D	Tek Boyutlu Evrişimsel Sinir Ağları
ConvLSTM	Evrişimsel Uzun Kısa Süreli Bellek (Convolutional Long Short-Term Memory)
LSTM	Long Short Term Memory (Uzun Kısa Süreli Bellek)
MFCC	Mel-frekans kepstrumu
RNN	Recurrent Neural Network (Yinelemeli Sinir Ağı)

1. GİRİŞ

Akut solunum yolu hastalıkları, her yıl milyonlarca insanı etkileyen önemli bir küresel sağlık sorunudur (Lee vd., 2025; Park vd., 2025; Singh vd., 2025). Pnömoni, astım, kronik obstrüktif akciğer hastalığı (KOAH) ve son dönemde COVID-19 gibi hastalıkları kapsayan bu durumlar, yaşlılar, çocuklar ve bağışıklık sistemi zayıf bireyler gibi hassas gruplar için özellikle tehlikeli olabilir. Vaka sayılarının yüksekliği ve gecikmiş teşhisin ciddi sonuçları, doğru, zamanında ve erişilebilir teşhis yöntemlerine duyulan ihtiyacı vurgulamaktadır. Geleneksel teşhis yöntemleri klinik muayeneler, laboratuvar testleri, görüntüleme teknikleri ve bazen invazif prosedürlere dayanmaktadır 2 başarılı sorgu Uygulama olmadan tekrar deneyin (Alyaa vd., 2025; Taloba vd., 2025). Ancak bu yöntemler maliyetli, zaman alıcı olabilir ve kaynakları kısıtlı bölgelerde ulaşılabilir değildir. Kırsal veya dezavantajlı bölgelerde tıbbi ekipman ve uzman eksikliği, teşhiste gecikmelere yol açabilmektedir.

Bu zorluklar ışığında, invaziv olmayan ve düşük maliyetli alternatif teşhis yöntemlerine ilgi artmaktadır (Park vd., 2025) Öksürük seslerinin analizi, umut vaat eden bir yaklaşımdır. Solunum sisteminin doğal bir refleksi olan öksürük, altta yatan fizyolojik veya patolojik durumun bir göstergesi olabilir. Öksürük sesleri, hastalığın türü, şiddeti ve ilerleyişi gibi önemli akustik bilgiler taşır (Le vd., 2025; Sajid vd., 2025; Xiao vd., 2025). Son araştırmalar, öksürük seslerinin uygun şekilde analiz edildiğinde solunum hastalıklarının teşhisinde anlamlı ipuçları sağlayabileceğini göstermiştir (Archana vd., 2025; Miah vd., 2025; Neili vd., 2025).

Ancak öksürük seslerinin karmaşık yapısı, çevresel gürültü, bireysel ses farklılıkları ve zaman-frekans özellikleri gibi faktörlerden etkilenebilir. Bu durum, farklı hastalıkların ayırt edilmesini zorlaştırır. Geleneksel sinyal işleme ve özellik çıkarımı yöntemleri, verideki bu incelikleri yakalamada yetersiz kalabilir (Dubey vd., 2025; Malviya vd., 2025).

Günümüzde hastalıkların tanısında yaygın olarak başvurulanan yöntemlerden biri oskültasyondur. Oskültasyon, organların fonksiyonel durumunu değerlendirmek amacıyla stetoskop yardımıyla vücut içinden gelen seslerin dinlenmesi işlemidir. Bu dinleme yöntemi; dolaşım, solunum ve sindirim sistemlerinin değerlendirilmesinde önemli bir yere sahiptir. Fizik muayenesinin temel bileşenlerinden biri olan oskültasyon, elde ettiği bilgilerle patolojik bulguların tespitini kolaylaştırır. Ancak geleneksel stetoskoplara yapılan değerlendirmelerde, duyuma dayalı yorumlar sebebiyle yanlış teşhisler meydana gelebilir. Bu süreç, sağlık personelinin işitme

yetisi, deneyimi, dikkat düzeyi ve psikolojik durumu gibi değişkenlere bağlı olduğundan; yorgunluk, stres ya da dikkatsizlik gibi durumlarda hata payı artar. Ayrıca, klasik stetoskopların 120 Hz üzerindeki frekans bileşenlerini bastırması da tanıda yetersizliğe neden olabilmektedir. Bu sınırlamaların aşılması amacıyla geliştirilen elektronik stetoskoplar, iç organlardan gelen seslerin dijital olarak kaydedilmesini ve bilgisayar ortamına aktarılmasını mümkün kılmaktadır. Böylece kaydedilen sesler tekrar dinlenebilir, analiz edilebilir ve uzmanlar arasında paylaşılabilir hale gelir. Verilerin dijital ortama aktarılması, bu sinyaller üzerinde gelişmiş sinyal işleme yöntemlerinin uygulanmasına olanak tanır. Bu teknikler, otomatik hastalık tanı sistemlerinin geliştirilmesine zemin hazırlayarak teşhis sürecine destek sağlar. Özellikle akciğer sesleri, hem eğitim hem de bilimsel araştırmalar kapsamında sıklıkla kaydedilmekte ve analiz edilmektedir. Bu çalışmaların temel hedeflerinden biri de, akciğer seslerinden elde edilen sinyaller aracılığıyla fizyolojik veya patolojik durumlara dair niteliksel ya da niceliksel bilgilerin çıkarılmasıdır. Bu tür seslerin kolaylıkla ve invaziv olmayan yöntemlerle kaydedilebiliyor olması, bu yaklaşımı daha da değerli hale getirmektedir (Kraman, Wodicka, Pressler ve Pasterkamp, 2006). Akciğer ve kalp gibi biyolojik kaynaklı ses sinyallerinin kaydı ve analiziyle ilgili ilk sistematik çalışmalar 1950’li yıllarda McKusick ve çalışma arkadaşları tarafından gerçekleştirilmiştir (VA vd.,1955). Kaydedilen bu ses sinyallerinden çeşitli öznitelikler (özellikler) çıkarılarak sınıflandırma işlemleri yapılmaktadır. Bu öznitelikler genellikle zaman, frekans ve zaman-frekans alanlarında analiz edilerek elde edilir. En yaygın kullanılan öznitelik çıkarma yöntemleri arasında Otoregresif Modeller (AR), Hızlı Fourier Dönüşümü (HFD), Kısa Zamanlı Fourier Dönüşümü (KZFD), Mel Frekanslı Kepstrum Katsayıları (MFCC) ve Dalgacık Dönüşümü yer almaktadır (KA vd., 2004). Literatürde, akciğer seslerinin sınıflandırılması ve analizi üzerine çok sayıda çalışma bulunmaktadır. Bu tez kapsamında, yalnızca akciğer seslerine odaklanan araştırmalar incelenmiş ve amaca uygun olanlar detaylı şekilde değerlendirilmiştir. Söz konusu çalışmalar, çeşitli filtreleme teknikleri ve analiz yöntemleriyle akciğere ait ses bileşenlerini ayırtırmayı ve tanımlamayı hedeflemiştir.

Li ve çalışma arkadaşları, wheeze (hırıltı) türündeki anormal solunum sesleri üzerine odaklanmışlardır. Bu çalışmada, 150-1000 Hz aralığında kesme frekansına sahip analog bir filtre kullanılmış ve öznitelik çıkarımı için sinyal segmentasyonu, Güç Spektrum Yoğunluğu (Power Spectral Density, PSD) ve Hızlı Fourier Dönüşümü (Fast Fourier Transform, FFT) yöntemlerinden yararlanılmıştır. Sınıflandırma aşamasında ise Fisher Diskriminant Analizi tercih edilmiştir (Li, Lin, Tsai, Yang ve Lin, 2017). İçer ve Gengeç ise çalışmalarını crackle (ince çıtırtı) ve

ronküs (gürültülü hırıltı) sesleri üzerine yürütmüşlerdir. 150-1800 Hz bant aralığında kesme frekansına sahip IIR (Sonsuz Darbeli Tepki) tipi Butterworth bant geçiren filtre (Bandpass Filter, BPF) kullanmışlardır. Öznelik çıkarımı için PSD ve Welch yöntemi uygulanmış, sınıflandırmada ise Destek Vektör Makineleri (Support Vector Machines, SVM) kullanılmıştır (İçer vd., 2014). Serbes ve ekibi crackle sesleri üzerinde gerçekleştirdikleri çalışmada, 80-4000 Hz frekans aralığında Butterworth bant geçiren filtre ve Bessel yüksek geçiren filtre (Highpass Filter, HPF) kullanmışlardır. Öznelikler, zaman-frekans analizi ve zaman-ölçekleme analizleri yoluyla elde edilmiştir. Sınıflandırma işlemi için ise SVM, K-En Yakın Komşu (K-Nearest Neighbors, KNN) ve Çok Katmanlı Algılayıcı (Multi-Layer Perceptron, MLP) algoritmaları kullanılmıştır (Serbes vd., 2013). Rielle ve Nahoma, wheeze sesleri üzerine yaptıkları analizde 20-4000 Hz aralığında çalışan FIR (Sonlu Darbeli Tepki) düşük geçiren filtre (Lowpass Filter, LPF) ve IIR Butterworth yüksek geçiren filtre uygulamışlardır. Öznelik çıkartımı için spektral izdüşüm (spectral projection) yöntemini kullanmışlar; sınıflandırma ise Yapay Sinir Ağları (Artificial Neural Networks, ANN) ile gerçekleştirilmiştir (Riella vd., 2009). Güler ve arkadaşları hem wheeze hem de crackle ses bileşenlerini incelemişlerdir. Çalışmada 80-2000 Hz kesme frekansına sahip Bessel tipi HPF ve Butterworth LPF filtrelerinden faydalanılmıştır. Öznelikler PSD yöntemiyle çıkarılmış; sınıflandırma ise Çok Katmanlı Algılayıcı (MLP) ile yapılmıştır (Güler vd., 2005). Piirilab ise ral (ıslak ve derin karakterli) seslerin frekans yoğunluğu seviyelerinde belirgin farklılıkların bulunduğunu ortaya koymuştur (Piirila vd., 1995).

Günümüzde pnömoninin yaygınlığı, COVID-19 salgınının da etkisiyle artmıştır ve salgınla daha iyi mücadele edebilmek için erken teşhis önemlidir. Ancak, pnömoninin COVID-19 salgını kaynaklı olup olmadığını gözle ayırt etmek, zor olduğu için teşhis koyma işlemini geciktirebilmektedir. Çünkü COVID-19 kaynaklı pnömoni ile lobar pnömoni örüntü olarak birbirine çok benzemektedir. Doktorlar, bu iki hastalığın ayırt edilebilmesi için diğer görüntüleme tiplerine göre daha ayrıntılı sonuçlar veren bilgisayarlı tomografi (BT) görüntülerini kullanılmaktadır. BT, diğer görüntüleme tiplerine göre daha ayrıntılı sonuçlar verir. Ancak BT işlemi, çok miktarda radyasyon içerir (Frush, 2004). Röntgen (X-Ray) cihazlarını ise BT ile kıyasladığımızda daha az radyasyon içerir. Akciğer grafisi görüntülerinden hastalıkların gözle tespiti ne kadar zor olsa da derin öğrenme teknikleriyle başarılı sonuçlar almak mümkündür. Çünkü derin öğrenmede piksel düzeyinde işlemler yapılmaktadır. Bu sayede pnömoni ve coronanın akciğerde sebep olduğu farklı tutulumlar ayırt edilebilir. Bu tutulumların en yaygın olanları interstisyel ve lobar pnömonidir (Lobar pnömoni daha çok bakteriyel, interstisyel pnömoni ise viral

kökenli olarak izlenmektedir). Bu tutulum tiplerini birbirinden ayıran anatomik yerleşimlerdeki farklarıdır. Lobar tutulum, akciğerin farklı loblarında oluşan genel bir tutulum olarak tanımlanabilir. İnterstisyel tutulum deseni buzlu cam opaklaşması (GOP) olarak tanımlanır. Literatürdeki çalışmalarda COVID-19 interstisyel tutulumuna benzer bir radyolojik desene neden olmaktadır ve hastaların akciğerlerinde alt loblarda daha sık tutulduğu saptanmıştır (Haseli vd., 2020).

Mart 2020'den bu yana dünya bir COVID-19 salgını içinde (Dünya Sağlık Örgütü, 2020) ve Temmuz 2021 itibarıyla 4 milyondan fazla insanın ölümüne neden oldu. Bilim insanları ve araştırma enstitüleri, virüsün dünya çapında yayılmasını izlemek ve engellemek için insanları test etmenin etkili yollarını bulmaya çalışıyor. Dünya çapındaki araştırmacılar, insanlarda COVID-19'u test etmek için öksürük veya ses tabanlı modeller kullanarak biyobelirteçler bularak bu sorunu çözmeyi hedefliyor. Bu hızlı test yöntemi temassız ve non-invazivdir ve bir hastaneye gitmeye veya herhangi bir tıbbi cihaza ihtiyaç duymaz. Öksürüğe dayalı AI çözümü, ön saflardaki sağlık çalışanlarımızın yükünü önemli ölçüde hafifletmeye yardımcı olur ve asgari veya hiç tıbbi tesis olmayan bölgelere ulaşabilir. Tez çalışmam, okullara, şirketlere, hastanelere vb. yardımcı olmak için öksürük tabanlı bir COVID-19 tarama aracını iyileştirmeye yönelik araştırmaları hızlandırmaya bilimsel olarak katkıda bulunacaktır. Sinyal ediniminin kolaylığı ve güvenilir teşhis sonuçları nedeniyle tıp ve AI araştırma alanında konuşma/sesle ilgili dijital biyobelirteçlere son zamanlarda bir ilgi duyulmaktadır (Vaishya vd., 2020). Önerilen tez çalışması, COVID-19 tanısını iyileştirmek için hem konuşma işlemeyi hem de son teknoloji derin öğrenme yaklaşımını bir araya getiriyor. Amaçlar iki yönlüdür: Bir ses akışındaki öksürük olaylarını tespit etmek ve yerleştirmek ve zaman damgasını çıkarmak için bir öksürük dedektörü modeli oluşturun. Öksürük tabanlı özellikler kullanarak COVID-19'u tahmin etmek için bir sınıflandırıcı eğitin.

Yapay zekanın birçok alanda çok yaygın olarak kullanıldığı bir çağda yaşıyoruz. Yaşam biçimimizi etkiledi ve geçtiğimiz on yılda yaşadığımız dünyayı kökten değiştirdi. Yapay zeka, hayatlarımızı iyiye doğru devrim niteliğinde değiştirme potansiyeline sahiptir. Tıp, devrim niteliğinde değişiklikler geçiren alanlardan biridir ancak kendi alanında yapay zekayı kullanma potansiyeli çok daha fazladır. Yapay zeka, özellikle Evrişimli Sinir Ağları'nın tanıtılmasından sonra, bu verileri kullanarak görüntü verileri üzerinde yararlı tahminler yapma konusunda çok büyük bir potansiyele sahiptir. Bir hastalığın semptomlarını bulmak için kullanılabilecek ortamlardan biri de öksürüktür. Öksürük sesleri üzerinde araştırma yapan birçok doktor, öksürüğün en azından altta yatan hastalığı ön teşhis etmek için

değerli bilgiler taşıdığına inanmaktadır. Temel varsayım, altta yatan hastalığın öksürüğün fiziksel karakterini belirlemesidir (Korpáš vd., 1996). Dünya Sağlık Örgütü'ne göre Boğmaca, yalnızca 2008 yılında 89.000 kişinin hayatına mal olurken, dünya çapında 151.074 kişiyi enfekte eden hastalıklardan biridir. Öncül, akciğerler etkilenirse öksürük sesinin de etkilenmesi gerektiğidir, ancak insan kulağı çeşitli hastalıkların öksürüklerini ayırt edemeyebilir. İşte yapay zekanın, bir sinyal/öksürükteki karmaşık örüntüleri veya gizli özellikleri veya ayırt edici özellikleri tespit etmek için devreye girdiği yer burasıdır. Bu çalışmada sırasıyla 37, 96, 35, 247 ve 130 boyutlarında Astım, Bronşit, Bronşiolit, Normal ve Boğmaca olmak üzere beş tür öksürük sesi kullanılmıştır. Aşağıdaki bölümlerde açıklanan tespit ve tanı için RESNET ve XGBoost modelleri kullanılmıştır. Bu veri kümesindeki modelin hassasiyeti ve geri çağırması yüksektir ve modelin büyük verilerle en azından aynı, hatta daha iyi performans gösterdiğini varsayarak tıbbi amaçlar için yararlıdır. Bu model daha sonra bir Uygulama üzerinde kullanılabilir ve akıllı telefonu olan herkesin öksürüğü tespit edip teşhis edebilmesi için erişilebilir hale getirilebilir.

Son yıllarda ses temelli tanı sistemleri, tıbbi alanda giderek artan bir ilgi görmektedir. Solunum yolu hastalıklarının teşhisinde; hastanın anamnezi ve mevcut semptomlarının yanı sıra, oskültasyon bulguları ve bilgisayarlı görüntüleme tekniklerinden de yararlanılmaktadır. Akciğer seslerinin analizi, klinik uygulamalarda hastalıkların değerlendirilmesinde önemli bir yer tutmaktadır. Solunum sesleri, akciğerlerin fiziksel muayenesinde temel bir bileşen olup, vücut içi seslerin stetoskop aracılığıyla dinlenmesini içeren oskültasyon süreci, akciğer bozukluklarının ve patolojik durumların tespitinde etkili bir yöntem olarak kullanılmaktadır (Earis, 1992). Akciğer seslerinin sınıflandırılmasında, sinyalin karakteristik özelliklerini belirleme süreci olan öznitelik çıkarımı kritik bir rol oynamaktadır (Bobrowski vd., 2009). Bu öznitelikler; zaman, frekans ve zaman-frekans alanlarında gerçekleştirilen analizlerle elde edilmektedir. Literatürde bilgisayar destekli analizlerde en sık kullanılan öznitelik çıkarma yöntemleri arasında; Otoregresif Model (AR), Mel Frekansı Kepstral Katsayıları (MFCC), Hızlı Fourier Dönüşümü (HFD), Kısa Zamanlı Fourier Dönüşümü (KZFD), enerji ve entropi tabanlı ölçütler, spektral özellikler ve Dalgacık Dönüşümü yer almaktadır (Kandaswamy vd., 2004). Bu tez çalışmasının temel amacı; sinyal işleme tekniklerinden yararlanarak, astım, bronşektazi, kronik obstrüktif akciğer hastalığı (KOA) ve idiyopatik pulmoner fibrozis gibi hastalıkların teşhisinde hekimlere yardımcı olabilecek bir sistem geliştirmektir. Oskültasyon, göğüs ve karın hastalıklarının değerlendirilmesinde fizik muayenenin vazgeçilmez bir parçası

olmasına rağmen; geleneksel stetoskoplar 120 Hz üzerindeki ses bileşenlerini zayıflatmakta ve çevresel gürültülerden kolayca etkilenmektedir. Ayrıca, oskültasyonun tamamıyla klinisyenin işitsel algısı ve deneyimine bağlı olması, tanı sürecinde subjektifliğe ve hatalara yol açabilmektedir. Bu çalışmada, klasik stetoskopların sınırlılıklarını aşmak amacıyla elektronik stetoskoplar aracılığıyla kaydedilen akciğer sesleri üzerinde bilgisayar destekli analizler gerçekleştirilerek, hastalıkların otomatik teşhisine yönelik bir yaklaşım geliştirilmiştir. Çalışma kapsamında; sağlıklı bireylerden ve akciğer hastalığı tanısı konmuş hastalardan elektronik stetoskop kullanılarak akciğer sesleri kaydedilmiş, bu sesler üzerinde sinyal işleme teknikleri uygulanarak anlamlı öznitelikler çıkarılmış ve elde edilen veriler sınıflandırılmıştır. Bu sınıflama sonucunda, akciğer hastalıklarının tespiti hedeflenmiştir. Çalışmada kullanılan ses kayıtları, İstanbul Üniversitesi Çapa Tıp Fakültesi Göğüs Hastalıkları polikliniği ve servisinde yatmakta olan hastalardan alınmıştır. Veri toplama sürecinde Littmann 3200 model elektronik stetoskop kullanılmıştır. Bu stetoskop 4000 Hz örnekleme frekansına sahiptir. Kayıtlar, hastaların sırt bölgelerinde yer alan oskültasyon noktalarından, sesin en baskın olduğu bölgelerden alınmıştır. Astım, bronşektazi, KOAH ve idiyopatik pulmoner fibrozis tanısı almış 20 hastadan, her biri için 10'ar adet (çalışmada 5'i kullanılmıştır) solunum örneği kaydedilmiştir. Kayıt sonrası sesler, daha uygun analiz koşulları sağlamak amacıyla 11025 Hz frekansında yeniden örneklenmiştir. Kayıt sırasında oluşabilecek deri sürtünmeleri, kas kasılması sesleri ve ortam gürültüleri gibi parazitleri azaltmak amacıyla 9. dereceden Chebyshev bant geçiren filtre uygulanmış; ayrıca, dalgalık tabanlı Shrinkage yöntemi ile gürültü giderme işlemi gerçekleştirilmiştir. Akciğer sesleri, birer nefes döngüsü olan inspirasyon (soluk alma) ve ekspirasyon (soluk verme) fazlarına göre bölütlenmiştir. Elde edilen temiz ses verileri üzerinde öznitelik çıkarımı için; Dalgacık Dönüşümü, Kısa Zamanlı Fourier Dönüşümü (KZFD), Doğrusal Öngörü Katsayıları (LPC), Mel Frekansı Kepstral Katsayıları (MFCC) ve istatistiksel öznitelikler kullanılmıştır. Ral, ronküs ve hırıltı gibi akciğer ek sesleri, çeşitli hastalıkların tanısında önemli bilgiler içermektedir. Bu seslerden elde edilen öznitelikler, ilk aşamada sağlıklı ve sağlıklı olmayan akciğer seslerinin ayrımında; ikinci aşamada ise sağlıklı seslerin ral, ronküs ve hırıltı olarak alt sınıflara ayrılmasında kullanılmıştır. Sınıflandırma işlemleri için Destek Vektör Makineleri (SVM) ve Çok Katmanlı Algılayıcı (Multi-Layer Perceptron, MLP) yöntemleri uygulanmış ve sinyal işleme tekniklerinin doğruluk oranları karşılaştırılmıştır. Bu çalışma, akciğer seslerinin iki aşamalı sınıflandırılmasıyla daha önceki çalışmalardan farklı olarak daha yüksek doğruluk oranları elde etmiş ve literatürdeki birçok çalışmaya kıyasla daha kapsamlı bir veri kümesiyle gerçekleştirilmiştir (Uysal, 2014).

Dünya Sağlık Örgütü (WHO) tarafından yayınlanan tablolara göre, kronik obstrüktif akciğer hastalığı (KOAH), alt solunum yolu enfeksiyonu ve akciğer kanserleri gibi kronik solunum yolu hastalıkları (CRD), dünya çapında önde gelen ölüm nedenleri arasında yer almakta ve milyonlarca insanın ölümüne neden olmaktadır. (Dünya Sağlık Örgütü, 2020) 2019'dan bu yana CoViD-19, dikkatleri güçlü bir şekilde akciğer sistemini hedef alan enfeksiyonlara çevirmiştir. Gelişmekte olan ülkelerdeki ve gelişmiş ülkelerin yoksul bölgelerindeki nüfusun, tedaviye yetersiz erişim, kaynak eksikliği ve genellikle kalifiye personel eksikliğinden kaynaklanan bir fenomen olan CRD'ye daha duyarlı olduğu bulunmuştur (Ait-Khaled vd., 2005.). Önlenebilir CRD söz konusu olduğunda nihai hedef, solunum sisteminde kalıcı hasarı önlemek için erken teşhistir. Solunum sistemindeki anormallikleri tespit etmede uzmanlara yardımcı olan başlıca araç, hala stetoskopla oskültasyondur. Oskültasyon, doktorların bir kişinin akciğerlerini, kalbini ve bağırsaklarını dinlediği en eski ve en geleneksel muayenedir. Her ne kadar temel, ucuz ve invazif olmayan bir yöntem olsa da ciddi dezavantajları da beraberinde getirmektedir. Akciğer seslerinin yorumlanması büyük ölçüde doktorun deneyimine ve işitme durumuna, ayrıca muayenenin yapıldığı ortama (yani gürültü seviyelerine) bağlıdır. Açıkçası, oskültasyon doğruluğu tartışmalı olabilecek manuel ve subjektif bir yöntem olmaya devam etmektedir. Hem solunum yolu hastalıklarının önlenmesi hem de oskültasyonun verimsizliği, akciğer seslerindeki düzensizliklerin otomatik olarak tespit edilmesinin yükselişiyle desteklenecektir. Solunum seslerinin otomatik analizi, manuel oskültasyonun dezavantajlarını azaltabileceği ve bize daha hızlı ve daha güvenilir tanı sağlayabileceği için büyük bir araştırma alanı haline gelmiştir. Akustik sesi elektronik sinyallere dönüştüren dijital stetoskoplar, daha az vasıflı hekimlere ihtiyaç duyacak veya belki de ortalama bir ev halkı için yararlı olacak bir teknolojiye yol açacak bir adım olabilir (Swarup vd., 2018).

Bu, dünya çapında kalifiye sağlık çalışanlarının azlığı sorununu çözecektir. Dijital stetoskoplar, verileri yapay sinir ağları ve derin öğrenme algoritmalarının anormallikleri tespit edebileceği bir formata dönüştürmek için dijital sinyal işleme yöntemlerini kullanmamıza olanak sağlayacaktır. Bireylerin evde basitçe kullanabilecekleri ve akciğer seslerinde bir düzensizlik olduğuna dair ilk teşhisi koyabilecekleri ve doktorlarından randevu almaları için onları bilgilendirecek bir teknoloji, CRD'nin önlenmesini artırırken gereksiz doktor ziyaretlerinin sayısını da azaltacaktır. Bu tezin amacı, akciğer sesleri üzerindeki maceracı seslerin otomatik olarak tespit edilmesinin nasıl çalıştığını araştırmak ve oluşturulan algoritmaların gerçek zamanlı bir sistem üzerinde uygulanma olasılıklarını kısaca tartışmaktır. Akciğer seslerinin kayıtlarını içeren açık bir veritabanı olan ICBHI 2017 Challenge

Veritabanını kullanarak, sağlıklı solunum seslerini sınıflandırmak için farklı temel derin öğrenme modellerini keşfediyor ve sonuçları en gelişmiş yaklaşımların sonuçlarıyla karşılaştırıyoruz. Algoritmik bölümün ardından, gelecekte gerçek zamanlı bir sistem inşa edilecekse gerekli olacak olan, uygulanan modellerin nicelleştirilmesini araştırıyoruz. Bu yöntem, ses sinyali işleme, derin öğrenme modelleri, algoritmalar ve kuantize sinir ağları hakkında kapsamlı uzmanlık kazanılmasını sağladı. Ayrıca PyTorch ve özellikle sinir ağı modülleri ve librosa gibi ses işleme modülleri gibi faydalı teknolojilerle ilgili deneyim kazanıldı (Charalampos, 2022).

COVID-19, ilk olarak 31 Aralık 2019 tarihinde Çin'in Hubei eyaletine bağlı Wuhan şehrinde ortaya çıkmış ve kısa sürede küresel bir salgın hâline gelmiştir. SARS-CoV-2 olarak adlandırılan bu yeni tip koronavirüs, Dünya Sağlık Örgütü (DSÖ) tarafından COVID-19 olarak tanımlanmıştır. Zoonotik kökenli olan bu virüs, hayvanlardan insanlara bulaşmış ve temas yoluyla insandan insana yayılmıştır. Çin'in diğer bölgelerine yayılması yalnızca 30 gün sürmüştür, 30 Ocak 2020 tarihinde DSÖ tarafından Uluslararası Öne Sahip Halk Sağlığı Acil Durumu, 11 Mart 2020 tarihinde ise pandemi ilan edilmiştir. Ocak 2021 itibarıyla dünya genelinde 100 milyondan fazla doğrulanmış vaka ve 2 milyondan fazla ölüm rapor edilmiştir (Muhammad vd., 2021). COVID-19'un başlıca belirtileri arasında ateş, kuru öksürük, yorgunluk ve nefes darlığı yer almakta olup; ağır vakalarda ölümcül pnömoni görülmektedir. Bu nedenle, hastalığın hızlı ve doğru şekilde teşhis edilmesi büyük önem arz etmektedir. Derin öğrenme yöntemlerinden biri olan Evrişimli Sinir Ağları (Convolutional Neural Networks – CNN), hem sınıflandırma hem de segmentasyon problemleri için sıklıkla kullanılan yapay zekâ temelli bir yaklaşımdır. Özellikle X-ışını, Bilgisayarlı Tomografi (BT) ve Akciğer Ultrasonografi (LUS) gibi medikal görüntüler üzerinde oldukça başarılı sonuçlar vermektedir. Pandemi sürecinde, COVID-19'un tanılanması için CNN temelli birçok model geliştirilmiş ve yüksek başarımlar elde edilmiştir (Huang vd., 2021; Polsinelli vd., 2020). Ancak bu çalışmaların büyük bir kısmı, klinik karar destek sistemlerinin geliştirilmesinde yetersiz kalmıştır. Yapay zekâ, makine öğrenmesi (Machine Learning – ML) ile doğrudan ilişkilidir. Makine öğrenmesi, bir bilgisayarın açık talimatlar olmaksızın verilerdeki örüntüleri öğrenerek sınıflandırma veya regresyon gibi görevleri gerçekleştirmesini sağlar. AI ve ML yöntemleri, yüksek boyutlu veriler üzerinde tahmin ve sınıflandırma gerçekleştirmesi bakımından geleneksel istatistiksel yöntemlerden ayrılır. Bu sistemlerin başarısı büyük ölçüde kaliteli ve yeterli veri setlerine bağlıdır. Derin öğrenme, makine öğrenmesinin daha karmaşık yapay sinir ağı modelleriyle genişletilmiş hâlidir. Görüntü işleme ve konuşma tanıma gibi

alanlarda önemli ilerlemelere yol açmış ve diğer birçok alanda da yaygın uygulama alanı bulmuştur (Jackson, 2019; Ertel, 2018). Makine öğrenmesi, verilerdeki örüntüleri keşfetmeyi, tahmine dayalı analizler yapmayı ve bu analizleri anlaşılır modeller hâline dönüştürmeyi amaçlar. Bu modeller, hastalıkların risk faktörlerine göre sınıflandırılmasında ve etkilerinin analiz edilmesinde kullanılabilir. Eğitim ve test verileri olarak ayrılan veri kümeleriyle geliştirilen modeller, sağlık hizmetlerinde teşhis ve tedavi kararlarını desteklemek amacıyla entegre edilebilmektedir (Yağın, 2020; Sarker, 2021; BalıkcıCicek vd., 2020). Makine öğrenmesi uygulamaları günümüzde; arama motorları, sosyal medya içerik filtreleme, öneri sistemleri ve tüketici elektroniği gibi birçok alanda yaygınlaşmıştır. Bununla birlikte, geleneksel ML yöntemleri ham veri üzerinde sınırlı performans göstermekte, dolayısıyla verilerin özellik çıkarımı için uzman müdahalesi gerektirmektedir (Shorten vd., 2021; Janiesch vd., 2021). Bu noktada CNN mimarisi devreye girmektedir. CNN, yüksek boyutlu veriler üzerinde otomatik öznitelik çıkarımı gerçekleştirebilen, denetimli öğrenmeye dayalı bir sinir ağıdır. Evrişim, havuzlama ve tam bağlantılı katmanlardan oluşur. Evrişim katmanları, girdilerden yerel özellikleri çıkarırken, havuzlama katmanları bu bilgileri özetler ve ağız parametre sayısını azaltır. Bu katmanlar sonunda tam bağlantılı katmanlar aracılığıyla sınıflandırma işlemi gerçekleştirilir. Özellikle görüntü sınıflandırmada yaygın olarak kullanılan CNN'ler, sağlık alanında da tanı koyma süreçlerine entegre edilmektedir (Yan vd., 2015). Buna ek olarak, sıralı veriler üzerinde etkili bir şekilde çalışabilen Uzun Kısa Vadeli Hafıza (Long Short-Term Memory – LSTM) ağları da COVID-19 teşhisinde kullanılabilecek modeller arasında yer almaktadır. LSTM, geri beslemeli sinir ağı (RNN) mimarisine dayalıdır ve zaman bağımlılığı içeren veri dizileri (örneğin, konuşma, video, biyo-sinyaller) üzerinde başarılı performans sergilemektedir. Giriş, gizli ve çıkış katmanlarından oluşan yapısı sayesinde zaman içerisindeki uzun süreli bağımlılıkları öğrenebilme yeteneğine sahiptir. Bu yönüyle, el yazısı tanıma, konuşma analizi ve saldırı tespit sistemlerinde olduğu gibi COVID-19 gibi zamanla değişen semptom ve sinyallere sahip hastalıkların analizinde de kullanılabilir (Qing vd., 2019; Sherstinsky, 2020).

Günümüzde kronik hastalıkların görülme sıklığındaki artış ve bu hastalıkların bireylerin yaşam kalitesi üzerindeki olumsuz etkileri, halk sağlığı politikalarının öncelikli gündemlerinden biri haline gelmiştir. Kronik hastalıklar, uzun süre devam eden ve çoğunlukla geri döndürülemez yapıda olan sağlık problemleri olarak tanımlanmakta olup, tedavi süreçleri semptomların kontrol altına alınmasına odaklanmakta, ancak genellikle tam bir iyileşme sağlanamamaktadır (Altuntaş vd., 2015). Bu hastalıkların erken teşhisi ve etkin şekilde yönetilmesi, bireylerin yaşam

kalitesini artırmakla birlikte, sağlık hizmetlerine erişimdeki bölgesel ve sosyoekonomik eşitsizlikler nedeniyle bu süreç çoğu zaman sekteye uğramaktadır (Küçükberber vd., 2011).

Bu doğrultuda, özellikle son yıllarda gelişen makine öğrenmesi (ML) ve yapay zekâ (YZ) teknikleri, büyük hacimli ve yapısal olarak karmaşık tıbbi verilerin analizinde çığır açan çözümler sunmaktadır (Sönmez vd., 2023). Bu teknolojiler, veri içerisindeki örüntüleri keşfederek hastalıkların erken teşhisini mümkün kılmakta ve kişiselleştirilmiş sağlık hizmetlerinin geliştirilmesine katkı sağlamaktadır. Yapay Sinir Ağları (YSA), Evrişimli Sinir Ağları (CNN), Uzun Kısa Süreli Bellek (LSTM), Destek Vektör Makineleri (SVM), Rastgele Orman (Random Forest), Gradyan Artırma (Gradient Boosting) ve daha yakın dönemde geliştirilen TabNet gibi modeller, sağlık verilerinin sınıflandırılması ve yorumlanması süreçlerinde yaygın olarak kullanılmaktadır.

Topluluk öğrenme yöntemlerinden biri olan Rastgele Orman algoritması, özellikle yüksek boyutlu veri kümelerinde tahmin doğruluğunu artırma ve aşırı öğrenmeyi önleme kapasitesiyle dikkat çekmektedir. Sağlık alanında risk değerlendirmesi, hasta sınıflandırması ve tanısal modelleme gibi çok sayıda uygulamada etkili sonuçlar verdiği bildirilmektedir (Luo vd., 2018). Benzer şekilde, YSA'lar da karmaşık doğrusal olmayan ilişkileri modelleyebilme yetenekleri sayesinde, hastalık teşhisi ve prognozunda tercih edilen modellerden biri hâline gelmiştir (Ahsan vd., 2023). LSTM'lerin ise zaman serisi biçimindeki sağlık verileriyle çalışırken gösterdiği üstün performans, uzun vadeli bağımlılıkların tanınmasına dayanmaktadır. Öte yandan, CNN'ler özellikle görüntü ve ses gibi yapılandırılmamış veriler üzerinde güçlü sonuçlar üretmekte, duygu analizi ve erken teşhis gibi uygulamalarda yüksek doğruluk oranları sunmaktadır (Almutairi vd., 2022).

Sağlık alanında kullanılan diğer bir önemli model olan SVM, özellikle yüksek boyutlu verilerle çalışırken elde ettiği istikrarlı başarı ile dikkat çekerken, genetik analiz ve kanser teşhisi gibi birçok klinik uygulamada yaygın olarak yer bulmaktadır (Guido vd., 2024). Yakın dönemde önerilen TabNet modeli ise sıralı dikkat mekanizması kullanarak özellik seçimini öğrenmeye entegre etmekte, bu sayede hem yorumlanabilirliği artırmakta hem de özellikle tablo biçimindeki verilerle yapılan sınıflandırma görevlerinde yüksek performans sunmaktadır (Arik vd., 2021).

Makine öğrenimi algoritmalarının sağlık uygulamalarında ses verilerini analiz

etme kapasitesi de giderek artmaktadır. Özellikle öksürük, solunum ve konuşma sesleri, son yıllarda biyobelirteç olarak değerlendirilmeye başlanmıştır. Konuşma verilerinde duyguların tanınması, temel ses sinyali özelliklerinin çıkarılmasıyla mümkün olmaktadır. Bu süreçte, geleneksel akustik özelliklerin yanı sıra CNN katmanlarından türetilen derin özellikler de sıklıkla kullanılmaktadır (Badshah vd., 2019).

Öksürük sesi, geçmişte genel bir semptom olarak değerlendirilse de, günümüzde yapay zekâ tabanlı çalışmalarla spesifik hastalıkların tanısında kullanılabilir hâle gelmiştir. Özellikle pediatrik hastalarda yapılan çalışmalarda, öksürük sesi analizine dayalı sistemler, yüksek doğruluk oranlarıyla zatürre gibi solunum yolu hastalıklarının teşhisinde etkili sonuçlar üretmiştir (Kosasih vd., 2014). Örneğin, MFCC gibi akustik özelliklerin kullanıldığı HMM tabanlı bir sistemde %95 duyarlılık ve %72 özgüllük elde edilmiştir. Ancak, yetişkin bireylerde öksürük sesine dayalı zatürre tanısına ilişkin literatürde ciddi bir boşluk bulunmaktadır (Botha, 2018).

Bu eksiklikten hareketle yapılan çalışmalar, özellikle öksürük seslerinin zamansal ve spektral özelliklerini analiz ederek, bunların belirli hastalıklarla ilişkili olup olmadığını belirlemeye odaklanmaktadır. Geliştirilen YZ tabanlı sistemler, hem teşhis doğruluğunu artırmakta hem de uzaktan hasta takibi ve erken müdahale fırsatlarını sunmaktadır. Bu bağlamda, akustik veriler üzerinden yapılan analizler, geleneksel fiziksel muayenelere ek olarak maliyeti düşük ve erişilebilir tanı araçları olarak öne çıkmaktadır (Rao vd., 2024; Kim vd., 2020).

Yalnızca öksürük değil, solunum ve konuşma seslerinin de hastalık tanısında önemli biyobelirteçler olduğu, çeşitli araştırmalarla ortaya konmuştur. Solunum esnasında duyulan hırıltı, wheezing veya stridor gibi ses değişimleri, alt solunum yolu hastalıklarının erken belirtileri olabilirken; konuşmadaki tonlama, perde ve şiddet değişimleri, ses teli patolojileri veya laringeal hastalıkların göstergesi olabilir (Anvarjon vd., 2020; Milton vd., 2013; Badshah vd., 2019)

COVID-19 pandemisiyle birlikte bu alandaki araştırmalar hız kazanmış, çeşitli makine öğrenimi modelleri, öksürük ve solunum seslerinden COVID-19'u yüksek doğrulukla tanıyan sistemler geliştirmiştir. CNN ve Otomatik Kodlayıcılar gibi modellerin bu süreçte sıklıkla kullanıldığı görülmüştür (Lella vd., 2021; Sezgin vd., 2001). Ayrıca, Dalgacık Dönüşümü (Wavelet Transform), Gauss Karışım Modeli (GMM), EMD ve Topluluk Deneysel Mod Ayırıştırma gibi teknikler,

solunum seslerinin daha etkili bir şekilde analiz edilmesini sağlamıştır. Tüm bu gelişmeler, yapay zekânın sağlık alanında ses temelli biyobelirteçlerle çalışarak hastalık tanı süreçlerinde hız, doğruluk ve maliyet avantajı sunduğunu ortaya koymaktadır. Ancak, hâlen karşılaşılan veri çeşitliliği, cihaz farklılıkları ve model genellenebilirliği gibi sınırlamalar, bu teknolojilerin klinik uygulamalara entegrasyonu açısından çözülmesi gereken önemli sorunlar arasında yer almaktadır. Bu bağlamda, bu çalışma hem mevcut literatürü kapsamlı biçimde taramakta, hem de özellikle yetişkin bireylerde zatürre tanısı için geliştirilen bir YZ tabanlı tanı modelini değerlendirerek alana katkı sunmayı amaçlamaktadır (Lozano vd., 2016; Maruf vd., 2015).

Son yıllarda yapay zekâ (YZ), makine öğrenmesi (ML) ve derin öğrenme (DL) teknikleri, sağlık alanında karmaşık ve yüksek hacimli tıbbi verilerin analizinde önemli bir rol oynamaya başlamıştır. Özellikle veri çeşitliliği ve çok boyutluluğu göz önüne alındığında, bu teknolojiler; hastalıkların erken teşhisi, sınıflandırılması ve bireyselleştirilmiş tedavi yaklaşımlarının geliştirilmesinde yüksek doğruluk oranlarıyla dikkat çekmektedir (Lundberg vd., 2018). Sağlık hizmetlerinde dijitalleşmenin artmasıyla birlikte, çok kaynaklı ve dinamik veri setlerinin analizi giderek daha kritik hale gelmiş; makine öğrenmesi tabanlı sistemlerin, tıbbi karar destek süreçlerine katkı sunma potansiyeli de artmıştır.

Sağlık alanında ses verilerinin incelenmesi, özellikle son yıllarda önem kazanan araştırma konularından biri haline gelmiştir. Ses, konuşma ve nefes alma gibi biyolojik işaretler, solunum yolları ve ses telleriyle doğrudan ilişkilidir ve bu nedenle bazı hastalıkların tanısında önemli biyobelirteçler olarak kullanılabilir. Literatürdeki çalışmalar, toplumun yaklaşık %3 ila %9'unun, çalışan nüfusun ise yaklaşık %10'unun hayatlarının bir döneminde ses kısıklığına bağlı sorunlar yaşadığını ortaya koymaktadır. Ses teli hastalıklarının tanısı çoğunlukla kulak burun boğaz (KBB) uzmanları tarafından gerçekleştirilen fiziki muayene, endoskopik görüntüleme ve akustik analiz yöntemlerine dayanmaktadır. Ancak sadece ses kısıklığı semptomu üzerinden kesin tanıya ulaşmak çoğu zaman güçtür. Bunun başlıca nedenleri arasında, insan kulağının sınırlı ayırt edebilme kapasitesi ve ses sinyallerinin bireyden bireye farklılık göstermesi yer almaktadır. Ayrıca, endoskopik muayenelerin maliyetli olması ve özel uzmanlık gerektirmesi, bu tür tanı yöntemlerinin birinci basamak sağlık hizmetlerinde yaygın şekilde uygulanmasını zorlaştırmaktadır. Bu noktada, ses verilerini analiz edebilen ve klinik karar süreçlerinde hekime destek sunabilecek makine öğrenimi tabanlı algoritmalar, tanı doğruluğunu artırma potansiyeline sahiptir. Bu sistemler, ses verilerinden çıkarılan

akustik ve spektral özellikleri kullanarak, olası patolojileri tespit edebilir ve hekimlere ön değerlendirme konusunda yardımcı olabilir. Böylelikle, hem yanlış tanı oranı azaltılabilir hem de ileri tetkik gerektiren vakalar daha etkin şekilde yönlendirilebilir. Ayrıca, bu tür sistemler, ses tabanlı teşhisin uzaktan izlenebilirliğini mümkün kılarak, özellikle kırsal ve sağlık hizmetlerine erişimin kısıtlı olduğu bölgelerde önemli bir boşluğu doldurabilir. Sonuç olarak, ses tabanlı biyobelirteçlerin makine öğrenmesi algoritmalarıyla bütünleşik şekilde kullanımı, hem klinik hem de teknolojik açıdan yenilikçi bir alan olarak öne çıkmaktadır. Bu bağlamda yapılacak çalışmalar, sadece mevcut tanı yöntemlerini tamamlamakla kalmayıp, aynı zamanda sağlık hizmetlerinin erişilebilirliğini ve kalitesini artırma yönünde de önemli katkılar sağlayacaktır (Verdolini vd., 2001).

Çalışmanın temel katkısı, ses temelli analizlerin tanı süreçlerinde kullanılabilecek etkili bir biyobelirteç niteliği taşıyabileceğini ortaya koymaktır. Bu kapsamda, astım, bronşit, KOAH, zatürre (pnömoni) ve sağlıklı bireyler olmak üzere beş farklı kategorideki solunum ses örnekleri incelenmiştir. Elde edilen bulgular, ses analizinin yalnızca yardımcı bir unsur değil, aynı zamanda potansiyel bir tanı aracı olarak değerlendirilmesini mümkün kılmaktadır. Bu bulguların, gelecekte mobil sağlık uygulamaları, giyilebilir cihazlar ve uzaktan teşhis sistemlerinde erken uyarı mekanizmalarının geliştirilmesine katkı sağlayacağı değerlendirilmektedir. Bu yönüyle çalışma, hem klinik hem de klinik dışı ortamlarda kullanılabilecek yenilikçi bir yaklaşım sunmaktadır.

Solunum yolu hastalıklarının tanısında erken teşhis, hastalığın seyri ve tedavi başarısı açısından büyük önem arz etmektedir. Tanı ve tedavi süreçlerinde uzman hekimlerin klinik karar verme süreci zaman alabilmekte ve bu durum, erken müdahaleyi zorlaştırabilmektedir. Bu nedenle, mevcut tanı yöntemlerinin zaman açısından verimsizliği, erken teşhis ve müdahale süreçlerinde belirgin zorluklar teşkil etmektedir. Son yıllarda, yapay zekâ ve özellikle derin öğrenme tabanlı yöntemlerin sağlık alanındaki kullanımı hızla yaygınlaşmakta olup, bu teknolojiler erken teşhis süreçlerinde önemli katkılar sunmaktadır. Klinik karar destek sistemleri olarak işlev gören bu yapay zekâ tabanlı uygulamalar, uzmanların tanıya ulaşma sürecini hızlandırmakta ve böylece daha hızlı, doğru ve etkili bir şekilde teşhis konulmasına olanak sağlamaktadır.

2. ÖNCEKİ ÇALIŞMALAR

Spesifik olmayan özellikler ve okuyucular arası değişkenlik gibi zorluklara odaklanarak göğüs röntgenlerinden ARDS tespiti için yapay zeka ve makine öğrenimi yaklaşımlarını araştırdılar. Yaygın alveolar hasarın bulut benzeri görünümünü yakalamak için tasarlanmış Directionality Measure adlı yeni bir özellik çıkarma yöntemi önerdiler. Directionality Measure, AdaBoost'u kullanarak %78 doğruluk ve %74 AUC elde ederek histogram, GLCM ve ResNet-50'ye dayalı özellikleri geride bıraktı. Tüm özellik kümelerinin birleştirilmesi performansı daha da artırarak %83 doğruluk ve %79 AUC'ye ulaştı. Bu bulgular, ARDS'nin klinik tanısını desteklemek için el yapımı ve derin özellikleri birleştirmenin potansiyelini vurgulamaktadır (Reamaroon vd., 2021).

ARDS'yi tespit etmek için ayrıcalıklı bilgi (LUPI) paradigmasını kullanarak öğrenme altında ayrıcalıklı lojistik regresyon (PLR) modelleri geliştirildi. Mekanik ventilasyon değişkenleri veya göğüs röntgeni özellikleri ayrıcalıklı bilgi olarak kullanılırken, elektronik sağlık kayıtları temel alanı oluşturdu. PLR, bilgi transferini geliştirmek ve yakınsama oranlarını iyileştirmek için eğitim sırasında ayrıcalıklı verileri dahil etti. Sonuçlar, PLR'nin kısmi ayrıcalıklı bilgilerle bile standart lojistik regresyonu geride bıraktığını gösterdi. Önerilen modeller klinik uygulamalar için güçlü performans, yorumlanabilirlik ve açıklanabilirlik sunar (Gao vd., 2024).

Göğüs röntgen görüntülerinden otomatik COVID-19 tespiti için yeni bir yapay zeka tabanlı derin öğrenme modeli önerildi. WavStaCovNet-19 modeli, dalgacık dönüşümlerini yığılmış derin öğrenme mimarisiyle (ResNet50, VGG19, Xception ve DarkNet19) birleştirir. İki genel kullanıma açık veri kümesi üzerinde test edildiğinde, sırasıyla 4 sınıf ve 3 sınıf sınıflandırma için %94,24 ve %96,10 doğruluk oranları elde etti. Sonuçlar, bu modelin daha düşük zaman ve maliyetle daha hızlı ve daha doğru COVID-19 tespitine yardımcı olabileceğini göstermektedir. Bu çalışma, sağlık uygulamaları için önemli bir potansiyel taşımaktadır (Soundrapandiyan vd., 2023).

Göğüs BT görüntülerini COVID pozitif ve COVID negatif olarak sınıflandırmak için Maske Evrişimli Sinir Ağları'nın (CNN) kullanılması önerildi ve COVID pnömonisini tipik pnömoniden ayırt etmedeki zorluklar ele alındı. Model, COVID ile ilişkili akciğer rahatsızlıklarının şiddet düzeylerini değerlendirmek için temel sınıflandırıcı olarak ResNet-50 ile meta öğrenme birkaç atışlı öğrenme tekniğini birleştirir. İlk sınıflandırma yöntemi, %98,60 doğrulukla %91,52'lik ortalama hassasiyet (mAP) elde etti. Akut akciğer rahatsızlıklarını tespit etmek için

kullanılan ikinci yöntem, %95,4'lük ortalama doğruluk, %94,02'lik G-ortalaması ve %93,27'lik AUC gösterdi. Bu sonuçlar, sağlık uzmanlarının uygun tedavi için şiddeti belirlemesine yardımcı olabilir (D vd., 2023).

Çocuklarda akut solunum yolu hastalıklarını teşhis etmek için öksürük seslerini analiz etmede makine öğrenimi algoritmalarının tahmin yeteneğini değerlendirmek üzere sistematik bir inceleme yürüttüler. Scopus, Medline ve Embase'de bir arama yaparak altı makale belirlediler. İnceleme, lojistik regresyon gibi geleneksel yöntemlerden CNN gibi derin öğrenme tekniklerine kadar algoritmalar tarafından kullanılan özelliklerde değişkenlik buldu. Bronşiolit, krup, boğmaca ve zatürre için sınıflandırma doğruluğu %82 ile %96 arasında değişiyordu. Sınırlı çalışmalara rağmen sonuçlar, pediatrik solunum yolu hastalığı tanısında öksürük sesi analizinin kullanılması için umut verici sonuçlar göstermektedir (Sharan vd., 2023).

Ön kalibrasyon olmaksızın solunum seslerinden boyutsuz solunum hava akışını tahmin etmek için derin öğrenme modeli önerilmiştir. Son solunum hava akışı dalga formu için çıktıları birleştirmek üzere bir topluluk yaklaşımı kullanarak, özellikleri çıkarmak ve zamansal bağımlılıkları modellemek için hibrit CNN + LSTM modelleri geliştirdiler. Modeller, 50 hastadan alınan solunum sesi ve EIT verileriyle değerlendirilmiş ve karmaşık değerlendirme görevlerinde 0,134'lük ortalama mutlak hata ve 0,770'lik bir Pearson korelasyonu ile umut verici sonuçlar elde edilmiştir. Solunum Sesi Veritabanı'nı kullanan harici testler, eğitim setinden farklı cihazlar kullanıldığında performans düşüşü göstermiştir. Çalışma, solunum hava akışını tahmin etmek için derin öğrenmenin potansiyelini göstermektedir (Pessoa vd., 2024).

Covid-19 pnömonisi ile Covid-19 olmayan interstisyel pnömonilerin göğüs röntgen görüntüleri arasında ayırım yapmak için bulanık mantık tabanlı bir derin öğrenme yaklaşımı olan CovNNet önerilmiştir. Model, bulanık kenar algılama algoritması tarafından oluşturulan CXR görüntülerinden ve bulanık görüntülerden özellikler çıkarmaktadır. Deneysel sonuçlar, kıyaslama derin öğrenme modellerini geride bırakarak %81'e kadar bir doğruluk oranı gösterdi. Yaklaşım ek veri kümeleriyle doğrulandı ve akut ortamlarda hastaların triyajı için potansiyel sunuyor. Yöntem, gelişmiş tanı için mevcut klinik karar destek sistemlerine entegre edilebilir (Ieracitano vd., 2022).

Tıbbi görüntüleme verilerini klinik bilgilerle birleştirerek erken akciğer

hastalığı tespiti için çok modlu bir füzyon yaklaşımı önerildi. Geç ve ara füzyona dayalı iki çok modlu ağ mimarisi ve uyarlanabilir bir toplu boyut yöntemi tanıtıldı. Deneysel sonuçlar, ara füzyon modelinin geç füzyon modelinden daha iyi performans gösterdiğini ve hem doğrudan hem de uyarlanabilir toplu büyüklüklerle daha iyi performans elde ettiğini göstermiştir. Çalışma, daha güvenilir erken evre akciğer hastalığı teşhisi için birden fazla veri kaynağının birleştirilmesinin potansiyelini vurgulamaktadır (Kumar vd., 2023).

Üç adımlı bir işlemle göğüs röntgen görüntüleri kullanılarak COVID-19 ve pnömoniye tespit etmek için bir yöntem önerilmiştir. İlk olarak, ham X-ışını görüntüleri koşullu üretken bir düşmanca ağ (C-GAN) kullanılarak segmentlere ayrılır. Ardından, anahtar nokta çıkarma yöntemlerini derin sinir ağları (DNN) ile birleştirilerek ayırt edici özellikler çıkarılır. Son olarak, çeşitli makine öğrenimi modelleri akciğer görüntülerini COVID-19, pnömoni veya normal olarak sınıflandırır. Yöntem, BRISK algoritmasıyla VGG-19 modelini kullanarak %96,6'lık en yüksek sınıflandırma doğruluğunu elde ederek COVID-19 taraması için verimli bir araç sunar (Bhattacharyya vd., 2022).

Alt düzey görüntü kalitesi ve büyük hacimli CXR'ler nedeniyle ön saflardaki hekimlerin karşılaştığı zorlukları ele alarak taşınabilir göğüs röntgenlerinden COVID-19 akciğer enfeksiyonlarını doğru bir şekilde tespit etmek için AI tabanlı bir analiz aracı geliştirildi. Kamuya açık veri kümelerini kullanarak, çalışma hem hiperparametre optimizasyonu ile hem de hiperparametre optimizasyonu olmadan ResNet101 ile derin öğrenmeyi kullandı. KNN ve SVM gibi makine öğrenimi algoritmalarıyla birleştirilen optimize edilmiş ResNet101 modeli, 3 sınıflı sınıflandırma için sırasıyla %99,86 ve %99,46 doğruluk elde etti. Önerilen yöntem, COVID-19 için sağlık sistemlerinde tanı verimliliğini ve doğruluğunu iyileştirme konusunda umut vad ediyor (Nawaz vd., 2023).

Klasik CNN ve kuantum sınıflandırıcılarını birleştiren solunum yolu akciğer hastalığı tespiti için bir hibrit çerçeve önerildi. Model, kuantum varyasyonel devrelere dayalı MMS ve MSMS kuantum makine öğrenimi algoritmalarıyla birlikte özellik çıkarma için özel bir CNN kullanır. COVID-19 Radyografi Veri Seti (CRD) üzerinde test edilen model, %98,9 eğitim doğruluğu ve %98,1 test doğruluğu elde ederek geleneksel derin öğrenme modellerini geride bıraktı. Kuantum sınıflandırıcılar, yüksek verimlilik ve düşük hesaplama maliyeti gösteren IBM Q-QASM gerçek zamanlı kuantum bilgisayarı kullanılarak doğrulandı (Rao vd., 2024).

İnterstisyel akciğer hastalıklarının erken teşhisini sağlamayı amaçlayan, bağ dokusu hastalıkları olan hastalardan gelen akciğer seslerini sınıflandırmak için derin öğrenme tekniklerinin kullanımı araştırıldı. Çalışma, gürültü giderme ve veri artırma için yeni bir ön işleme hattı tanıttı. Yaklaşım, yüksek çözünürlüklü bilgisayarlı tomografiye göre doğrulanan akciğer seslerini sınıflandırmada %91'e kadar bir doğruluk elde etti. Bu yöntem, yaşlı bireylerde interstisyel akciğer hastalıklarının yaygın olarak taranması için invaziv olmayan, uygun maliyetli bir çözüm sunarak pahalı görüntüleme tekniklerine olan bağımlılığı azalttı (Dianat vd.,2023).

Otomatik solunum sesi sınıflandırma sistemlerine derin öğrenmeye dayalı bir ses geliştirme ön işleme adımının entegrasyonunu araştırdılar. Çalışmada çeşitli ses geliştirme mimarileri test edildi ve bunların iki veri kümesinde etkililiği değerlendirildi: ICBHI ve Formosa Archive of Breath Sound. Sonuçlar gürültülü ortamlarda sırasıyla %21,88 ve %4,1'lik önemli bir performans artışı gösterdi. Bir hekim doğrulama çalışması ayrıca %11,61 oranında iyileştirilmiş tanı duyarlılığını gösterdi. Bulgular, ses geliştirmenin dahil edilmesinin sağlamlığı ve klinik uygulanabilirliği geliştirdiğini, tıp uzmanları arasında güveni artırdığını vurgulamaktadır (Tzeng vd., 2025).

Kardiyopulmoner ses analizindeki zorlukların üstesinden gelmek için ortak dikkat mekanizması ve yarı-denetimli öğrenme kullanan bir kalp-akciğer sesi ayırma ağı olan JASSNet önerildi. Model, evrimsel modülleri, küresel ve yerel dikkat mekanizmalarını ve özellik çıkarma için bir geçit mekanizmasını birleştirir. Klinik deneyler, kalp atış hızı (HR) ve solunum hızı (RR) doğruluğunun sırasıyla %10 ve %15 oranında artmasıyla önemli bir gelişme gösterdi. Çalışma, modelin sağlık izleme sistemleri için ön işlemedeki potansiyelini ve klinik uygulama beklentilerini vurgulamaktadır (Zhang vd., 2025).

10.448 adet 15 saniyelik trakeal ses kaydı içeren HF_Tracheal_V1 adlı bir trakeal ses veritabanı geliştirildi ve solunum sesi analizi için derin öğrenme modeli eğitmek amacıyla akciğer sesi veritabanı HF_Lung_V2 ile birlikte kullanıldı. Çalışmada, karma set eğitimi, alan uyarlaması ve tek ses eğitimi dahil olmak üzere eğitim yöntemleri karşılaştırıldı. Sonuçlar, yalnızca akciğer sesleriyle yapılan eğitimin trakeal ses analizi için etkisiz olduğunu ve bunun tersinin de geçerli olduğunu gösterdi. Ancak, karma set veya alan uyarlamasıyla yapılan eğitim, hem akciğer hem de trakeal seslerde inhalasyon, ekshalasyon ve tesadüfi sesleri (CAS) tespit etme performansını iyileştirdi. Karma set modeli özellikle çok yönlüydü ve hem akciğer hem de trakeal ses analizini etkili bir şekilde destekliyordu (Hsu vd.,

2023).

Astımın otomatik tespiti için Attention ResNet18 tabanlı bir model önerdiler. 767 astım ve 722 sağlıklı segment dahil olmak üzere 203 denekten alınan akciğer sesleri, modeli eğitmek için kullanıldı. Astım seslerini normal seslerden ayıran ısı haritaları oluşturmak için Grad-CAM kullanıldı. Model, 10 katlı çapraz doğrulama ile SVM kullanarak %99,73'lük bir doğruluk elde etti. Bu model, gerçek dünya senaryolarında astım tespiti için önemli bir potansiyele sahiptir (Topaloglu vd., 2023).

Varyasyonel mod ayrıştırma (VMD) tekniğini kullanarak gürültü gidermeyi içeren solunum sesi sınıflandırması için yeni bir ön işleme tekniği önerin. Temizlenen sinyaller daha sonra Gammatonegram görüntüleri biçiminde zaman-frekans dağılımları oluşturmak için bir gamatone filtre bankasından işlenir. Bu görüntüler, AlexNet, GoogLeNet, ResNet-50 ve Inceptionv3 dahil olmak üzere transfer öğrenmesi yoluyla çeşitli derin CNN mimarileri kullanılarak sınıflandırılır. Yöntem, akciğer seslerini %98,8 doğruluk, %97,7 hassasiyet, %100 duyarlılık ve %97,6 özgüllük ile üç sınıfa ayırarak yüksek sınıflandırma performansı elde eder ve bu da onu solunum yolu hastalığı teşhisi için etkili ve verimli bir araç haline getirir (Gupta vd., 2021).

Akciğer seslerinin (LS) otomatik tanısı için evrimsel katmanları ve LSTM katmanlarını birleştiren yeni bir derin öğrenme mimarisi olan ConvLSNet sunulmaktadır. ConvLSNet, çok kanallı LS kayıtlarının mekansal ve zamansal özelliklerini işlemek için ConvLSTM katmanını kullanarak, ağır ön işleme gerek kalmadan son derece doğru sınıflandırma sunar. Model, üç akciğer rahatsızlığı için etkileyici bir %97,4 sınıflandırma doğruluğu elde eder: astım, KOAH ve sağlıklı durum. Yalnızca CNN veya LSTM katmanlarını veya 2D CNN ve LSTM katmanlarının bir basamağını kullanan diğer modellerle karşılaştırıldığında, ConvLSNet, parametreler, eğitim süresi ve öğrenme oranı açısından daha düşük bir hesaplama maliyetini korurken sınıflandırma doğruluğunda onları geride bırakmaktadır (Majzoobi vd., 2024).

Hematolojik hastalıklar alanında, mikroskopik görüntülerden lösemi alt türlerini belirlemek için CNN'leri kullanmıştır. Geliştirdikleri model %88,25'lik bir doğruluk oranına ulaşarak derin öğrenme tekniklerinin karmaşık hastalıklar için teşhis doğruluğunu artırma potansiyeline işaret etmiştir. Bu durum, lösemi teşhisi için CNN'lerle derin transfer öğrenmesi kullanan ve modelin kan hücresi

görüntülerinden ilgili özellikleri etkili bir şekilde çıkarma yeteneğini vurgulayan Loey vd. (2020) tarafından da desteklenmektedir (Loey vd., 2020).

Yaptıkları bir çalışmada, tüp bebek tedavisinde pozitif gebelik sonucunu tahmin etmek amacıyla yapay zekâ tabanlı bir klinik karar destek sistemi geliştirilmişlerdir. Üniversitesi Çalışmada, Yeditepe Hastanesi'nden alınan 1154 tedavi siklusuna ait veriler kullanılarak, beş farklı sınıflandırma yöntemi (DVM, Çok Katmanlı Algılayıcı, Rastgele Orman, XGBoost ve LightGBM) karşılaştırmalı olarak test edilmiştir. En yüksek sınıflandırma performansı, Destek Vektör Makineleri yöntemi ile elde edilmiştir, AUC değeri 0.70 olarak bulunmuş ve karar eşik değerinin optimizasyonu ile gebelik sonucunun %71.7 Doğru Pozitif ve %59.4 Doğru Negatif oranıyla tahmin edilmesi sağlanmıştır (Pacci vd., 2021).

Yapmış oldukları bir çalışmada, Alzheimer hastalığının (AD) ilerleyişini, beyin manyetik rezonans görüntüleme (MRI) verileri kullanılarak çeşitli makine öğrenimi algoritmaları ile sınıflandırmış ve tahmin etmişlerdir. ADNI veri tabanından alınan 560 katılımcı, kognitif normal (CN), erken hafif bilişsel bozukluk (EMCI), geç hafif bilişsel bozukluk (LMCI) ve Alzheimer (AD) olmak üzere dört gruba ayrılmıştır. Rastgele Orman, DVM ve Karar Ağacı algoritmaları kullanılarak bu grupların hastalık ilerleyişi sınıflandırılmıştır. Rastgele Orman algoritması en yüksek doğruluğa (CN-AD için %96.14) ve en yüksek AUC değerine (0.92) ulaşarak diğer modellerden daha başarılı olmuştur. Sınıflandırmada kullanılan MRI özellikleri ile Rastgele Orman modeli, hastalığın erken teşhisinde yardımcı bir araç olarak önerilmiştir (Tang vd., 2021).

Çalışmasında, Kaggle veri tabanından alınan kalp hastalığı veri seti kullanılarak Python aracılığıyla 7 sınıflandırma algoritması (Destek Vektör Makineleri, Gaussian Naive Bayes, Gradient Boosting ve Rastgele Orman) performanslarının karşılaştırılması yapılmıştır. En iyi performans gösteren algoritma, %89.7 F1 skoru ve %90.2 doğruluk ile Rastgele Orman olmuştur. Açlık kan şekeri özelliğinin önem sıralamasında en altta yer alması, sınıflandırma performansı üzerinde minimal etkisi olduğunu göstermiştir. Bu bulgular, kalp hastalığı tahmininde daha etkili ve doğru sistemlerin geliştirilmesine katkı sağlayabilir (Nawaz vd., 2023).

Yapmış oldukları bir çalışmada, deri kanseri tespiti için CNN tabanlı yeni bir model geliştirmişlerdir. Model, Autoencoder, MobileNetV2 ve Spiking Neural Networks (SNN) bileşenlerini kullanarak, iyi huylu ve kötü huylu tümörleri

sınıflandırmışlardır. ISIC veri seti kullanılarak yapılan deneylerde, MobileNetV2 modeli ve spiking ağırları ile %95,27'lik bir doğruluk oranı elde edilmiştir. Bu sonuçlar, Autoencoder ve SNN'nin MobileNetV2 modelinin performansını artırmada etkili olduğunu göstermiştir. Çalışma, deri kanseri tespiti için yüksek hassasiyetli ve tamamen otomatik bir karar destek aracı sunmaktadır (Nawaz vd., 2023).

Konvolüsyonel sinir ağlarını (CNN), LSTM ile birleştiren çok görevli bir öğrenme çerçevesi kullanarak kronik hastalık tahminini araştırmıştır. Yaklaşımları, birden fazla kronik hastalığı aynı anda tahmin etmede %94,3'lük bir doğruluk elde etmiştir. Bu çalışma, LSTM'nin kronik hastalık tahmini için gerekli olan zaman serisi verilerindeki zamansal ilişkileri modelleme yeteneğini vurgulamaktadır (Kraman vd., 2006).

Çalışma, pozitif vakaları tahmin etmek için yaygın olarak kullanılan 10 önceden eğitilmiş CNN modelini kullandı. En iyi performans, pozitif COVID-19 vakalarını diğer atipik ve viral pnömoni hastalıklarından %99,51 doğruluk oranıyla başarıyla ayıran ResNet-101 kullanılarak elde edildi. ResNet-101'in kullanımı ayrıca düşük maliyetli bir çözüm ve yüksek hassasiyet sağladı (Fraïwan vd., 2021).

Çalışma, karar ağacı (DT) sınıflandırıcılarına ve CNN'lere dayalı bir model önerdi. 3 DT modeli, her DT'nin karşılık gelen CNN tarafından eğitildiği 3 CNNS ile birlikte kullanılır. İlk DT görüntüleri normal veya anormal olarak sınıflandırır, ikinci DT tüberkülozu belirler ve üçüncü DT COVID-19'u belirler. İlk DT, ikinci DT ve üçüncü DT sırasıyla %98, %80 ve %95 doğruluk oranlarına ulaştı (Alqudah vd., 2022)

Çalışma, göğüs röntgeni görüntülerinden özellik çıkarmada kullanılmak üzere ImageNet kullanılarak eğitilen farklı CNN mimarileri kullandı. Bundan sonra, CNN'ler çeşitli makine öğrenimi algoritmalarıyla entegre edildi. En yüksek doğruluk oranı, %98,5'e ulaşan doğrusal bir çekirdek SVM'si olan MobileNet'ten elde edildi. Derin öğrenme modelleri, özellikle önceden eğitilmiş modeller, CovidNet ve ResNet 101, yaklaşık %100'lük bir doğruluk oranına ulaştı. Hibrit yöntemlerin veya makine öğrenimine dayalı yöntemlerin başarı oranları, bu derin öğrenmeye dayalı yöntemlerin başarı oranları kadar yüksek değildir. Bu paragrafta incelenen çalışmalar arasında, en iyi doğruluk oranı, DT kullanılarak elde edilen %98'dir. Çalışmalar, daha yüksek doğruluk elde etmek için derin öğrenme modellerinin kullanılmasının makul bir neden olacağını göstermiştir (Hassan vd., 2024).

Berlin (EMO-DB) veri kümesi, konuşma seslerinin sınıflandırılması için seçilir. Veri kümesindeki yedi farklı duyguyu sınıflandırmak için 3 aşamalı SVM sınıflandırıcısı kullanılır. MFCC özellikleri, veri kümesinde bulunan 535 kaydın tamamından çıkarılır ve özellikler üzerinde dokuz ayrı istatistiksel değerlendirme yapılır. Eğitim aşaması ve verilerin test aşaması için 10 katlı çapraz doğrulama yapılır. Performans analizi, karışıklık matrisi kullanılarak yapılır ve %68'lik bir doğruluk oranı elde edilir (Nawaz vd., 2023).

MFCC aranır ve bu özellikler Doğrusal Ayrımıcı Analiz (LDA) ile sınıflandırılır. Ayrıca, makalede, yapay duygusal Marathi konuşmasına ait bir veritabanı kullanılır. Veri örnekleri 5 Marathi filminde oynayan aktör ve aktrislerden toplanır ve günlük iletişimde kullanılabilen ve tüm duygularda yeniden ifade edilebilen Marathi ifadeleri üreten duygular simüle edilir. Veri örnekleri Mutlu, Üzgün, Öfke, Korkmuş ve Sürpriz olmak üzere 5 temel kategoriye ayrılır (Orkweha, Phapatanaburi, Pathonsuwan, Jumphoo, Rattanasak, Anchuen, Pinthurat, Uthansakul, ve Uthansakul, 2024).

Dört duygusal durum, konuşmadaki duyguları sınıflandırmak için ele alınmıştır. Bu amaçla, duygusal konuşmanın ses özellikleri olan Doğrusal Frekans Cepstral Katsayıları (LFCC) ve (MFCC) kullanılarak özellikler çıkarılır. Ek olarak, bu özellikler Gizli Markov Modeli (HMM) ve SVM kullanılarak sınıflandırılır (Chenchah vd., 2015).

MFCC ve özellik vektörünün bileşenleri olarak enerjiyle ilişkili özellikler kullanarak bir duygu tanıma sisteminin performansını iyileştirmek için çalıştılar. Duygudan etkilenen frekans aralığını belirledikten sonra, dinamik zaman bükme çok katmanlı algılayıcı hibrit modeli kullanılarak normalleştirme gerçekleştirilir. Bu çalışmada, özellik sayısını azaltmak için hızlı korelasyona dayalı filtre ve varyasyon analizi yöntemleri kullanıldı. Duygusal durumların tanınması Gauss karışım modeli kullanılarak gerçekleştirilir (Gharavian vd., 2013).

MFCC, kroma gram, mel ölçekli spektrogram, Tonnetz gösterimi ve spektral kontrast özellikleri ses dosyalarından çıkarılır ve 1D-CNN için girdi olarak kullanılır (Issa vd., 2020).

Konuşma duygularını tanımak için derin duygusal özellikler araştırılmıştır. Konuşma ve log-mel spektrogramından duygusal özellikleri öğrenmek için CNN ve Uzun-kısa-dönemli bellek (LSTM) oluşturulmuştur. Bunlar tek boyutlu CNN-LSTM ağı ve iki boyutlu CNN-LSTM ağıdır. DeneySEL çıktılar, tasarlanan ağların konuşma

duygularını tanıma görevi için iyi performans gösterdiğini göstermektedir. Özellikle 2DCNN-LSTM ağı, geleneksel yaklaşımlar olan Derin İnanç Ağı (DBN) ve CNN'den daha iyi performans göstermektedir (Zhao vd., 2019).

Konuşma seslerinin sınıflandırılması için Berlin (EMO-DB) veri kümesi seçilmiştir. Veri kümesindeki yedi farklı duyguyu sınıflandırmak için 3 aşamalı bir SVM sınıflandırması kullanılmıştır. MFCC özellikleri, veri kümesinde bulunan 535 kaydın tamamından çıkarılmış ve bu özellikler üzerinde dokuz ayrı istatistiksel değerlendirme yapılmıştır. Eğitim aşaması ve verilerin test aşaması için 10 katlı çapraz doğrulama yapılmıştır. Performans analizi, karışıklık matrisi kullanılarak yapılmış ve %68 doğruluk oranına ulaşılmıştır (Milton vd., 2013)

Yeni bir model sunulmuştur. Uçtan uca eğitilmiş model, ham sinyalden özellikler alan 2 katmanlı Uzun Kısa Süreli Bellek (LSTM) içeren bir CNN'den oluşur. Sistem, konuşma sinyalinin bağlamını tamamen analiz etmiş ve tutarlılık ve tanıma katsayısı açısından diğer sistemlerden daha iyi performans göstermiştir (Tzirakis vd., 2018).

MFCC, kromagram, mel ölçekli spektrogram, Tonnetz gösterimi ve spektral kontrast özellikleri ses dosyalarından çıkarılmış ve 1D-CNN için girdi olarak kullanılmıştır (Tzirakis vd., 2018).

Kısa süreli Fourier dönüşümü (STFT), Mel-frekans cepstral katsayıları (MFCC), spektrogram görüntüleri kullanarak akciğer sesi (LS) anormalliklerini belirlemek için her bir CNN ve SVM sınıflandırıcısı için dört veri seti oluşturdular ve her iki sınıflandırıcıda da %86 doğruluk elde ettiler. [115]'te Damir ve arkadaşları, spektrogram özellikleri oluşturmak için ICBHI veri setini kullandılar ve onları SVM ve CNN sınıflandırıcıları üzerinde eğittiler. En iyi doğruluk oranı olan %65,5, kübik bir SVM sınıflandırıcısı kullanılarak elde edildi. İkinci yöntemde, transfer öğreniminin ince ayarı uygulandı ve %63,09 doğruluk elde edildi (Aykanat vd., 2017).

Naqvi ve arkadaşları ICBHI veritabanından zaman, cepstral ve spektral alan özellikleri çıkarıldı ve bunlar Karar Ağacı (DT), Lojistik Regresyon (LR), Destek Vektör Makinesi (SVM), K-En Yakın Komşu (KNN), İkinci Dereceden Ayırıcı (QD) gibi farklı sınıflandırıcılara girdi olarak verilmek üzere birleştirildi (Naqvi vd., 2020).

ICBHI veri setini kullanarak 22 kHz'de yeniden örneklediler ve dolgu teknikleriyle ön işleme tabi tuttular. Log Mel spektrogramlarından üst düzey

özellikler, CNN kullanılarak çıkarıldı ve çoklu sınıflandırma için %78,4 ve ikili sınıflandırma için %83,7 doğruluk elde edildi (Nguyen vd., 2020).

Sengupta ve arkadaşları, üç farklı kamuya açık kaynak kullanmış, bunları 4 kHz'de indirgemiş ve FFT tabanlı güç spektrumlarını ve MFCC'lerini hesaplamıştır. Çok Katmanlı Algılayıcı (MLP) özellikli YSA, çıkarılan özellikler üzerinde eğitilmiş ve hırıltılar için %95,17 doğruluk elde etmiştir (Sengupta vd., 2016).

Demir ve arkadaşları, ICBHI üzerinde eğitilmiş bir CNN modeli kullanmıştır. Derin özellikler, oluşturulan spektrogramdan çıkarılmış ve daha sonra Doğrusal Ayırım Analizi (LDA) sınıflandırıcısına ve VGG-16'ya beslenmiştir (Demir vd., 2020).

Araştırmacılar, farklı çevrimiçi veri kümelerini kullandılar ve genlik normalizasyonu, DC bileşen kaldırma uyguladılar ve Ayırık Dalgacık Dönüşümü (DWT) ve Dalgacık Paket Ayırıştırma (WPD) ile LS sinyali ayrıştırdılar (Rizal vd., 2019).

Akciğer tüberkülozu tanısı için 38 çocukta öksürük seslerinin MFCC'lerini analiz etmek amacıyla vücut ısısı, vücut kitle indeksi (VKİ) ve kalp atış hızını kullanmıştır. HMM tabanlı bir sınıflandırma sistemi kullanılarak %95 duyarlılık ve %72 özgüllük sonuçları elde edilmiştir. Bu önceki çalışmalar, pediatrik hastalarda solunum yolu hastalıklarını teşhis etmek için öksürük sesi sinyallerini ve yapay zeka teknolojilerini birleştirerek öncü rol oynamıştır. Ancak, yetişkinlerde zatürre tanısı için öksürük seslerinin analizi henüz rapor edilmemiştir (Botha vd., 2018).

Clifford ve arkadaşlarının çalışmasında, kısa ve tek kanallı EKG kayıtlarından atriyal fibrilasyon (AF) sınıflandırmasını otomatik olarak gerçekleştirebilecek bir algoritma geliştirilmesi hedeflenmiştir. EKG sinyalleri dört kategoriye ayrılmış ve toplamda 12,186 EKG kaydı kullanılmıştır. Bu kayıtlardan 8,528'i eğitim seti, 3,658'i ise test seti olarak ayrılmıştır. Çalışmada Random Forest, CNN, RNN gibi yöntemler kullanılmış ve 75 yarışmacı katılmıştır. Sonuç olarak, 0.83 F1 puanı ile dört yarışmacı yarışmayı kazanmış ve en başarılı algoritmanın AF'yi tanımlamada en yüksek doğruluk ve güvenilirliği sağladığı belirlenmiştir. Bu çalışma, EKG analiz sistemlerinin geliştirilmesine önemli katkılar sağlamıştır (Clifford vd., 2017).

Benzer bir yaklaşımla, Liang ve arkadaşları, kalp atışını doğru ve verimli bir şekilde sınıflandırmak için iki yöntem kullanmışlardır. İlk yöntemde genetik algoritma (GA) ve makine öğrenimini birleştiren evrimsel bir sinir sistemi yaklaşımı

uygulanırken, ikinci yöntemde Evrişimsel Sinir Ağı (CNN) ile Çift Yönlü Uzun Kısa Süreli Bellek (BiLSTM) ağı (CNN-BiLSTM) birleştirilmiştir. İlk yöntem, normal sinüs ritmi, AF ve supraventriküler taşiaritmi gibi durumları saptamada üstün performans gösterirken, ikinci yöntem PAC, atriyal flutter ve ventriküler bigemini gibi durumları saptamada üstünlük sağlamıştır. Farklı veri kümeleri üzerinde %80, %82.6 ve %85 doğruluk oranlarına ulaşmıştır (Liang vd., 2018).

Bu çalışmaları takip eden Śmigiel ve arkadaşları, PTB-XL veritabanını kullanarak birincil EKG sinyallerinin otomatik sınıflandırılması için üç farklı ağ mimarisi geliştirmişlerdir. Konvolüsyonel ağ, SincNet ağı ve entropi tabanlı özelliklere sahip bir konvolüsyonel ağ gibi farklı mimariler kullanılmıştır. Elde edilen sonuçlar, entropi özellikli konvolüsyonel ağın iki sınıf için %89.82 ve beş sınıf için %76.5 doğrulukla diğer ağlardan daha iyi performans gösterdiğini ortaya koymuştur (Śmigiel vd., 2021).

Benzer bir şekilde, Er'in çalışmasında, EKG sinyallerinin sınıflandırılması için PASCAL ve PhysioNet 2016 veri setleri kullanılmış ve 1D Yerel İkili Örüntü (1D-LBP) ile 1D Yerel Üçlü Örüntü (1D-LTP) teknikleriyle özellik çıkarımı yapılmıştır. ReliefF algoritmasıyla seçilen en önemli özellikler, Tek Boyutlu Evrişimsel Sinir Ağı (1D-CNN) modelinde sınıflandırılmış ve PASCAL veri kümesinde %91.66, PhysioNet 2016 veri kümesinde ise %91.78 doğruluk elde edilmiştir (Er, 2021).

Sattar ve arkadaşları, sayısallaştırılmış CPEIC kardiyak veri seti üzerinde makine öğrenimi ve derin öğrenme teknikleri kullanarak EKG sınıflandırması yapmışlardır. Bu çalışmada, EKG görüntülerinin zaman serisi sinyallerine dönüştürülmesi ve kardiyak sınıfların doğru sınıflandırılması için CNN, LSTM ağı ve oto kodlayıcılarla desteklenen bir SSL tabanlı model kullanılmıştır. CNN modeli, yaklaşık %92 doğruluk oranı ile en yüksek performansı göstermiştir (Sattar vd., 2024).

Solunum sesi analizi, yapay zekâ destekli yöntemlerin gelişimiyle birlikte son yıllarda solunum hastalıklarının erken teşhisi ve dolayısıyla başarılı bir tedavi süreci için giderek daha fazla önem kazanan bir araştırma alanı hâline gelmiştir. Bu alandaki çalışmalar, ham ses sinyallerinden anlamlı öznitelikler çıkarmak, bu öznitelikleri işleyerek sınıflandırmak ve sonuçlar üzerinden hastalık tahmini yapmak gibi çeşitli aşamaları içermektedir. Özellikle COVID-19 pandemisi sürecinde, ses temelli tanı sistemlerine olan ilgi önemli ölçüde artmış ve bu bağlamda birçok

yenilikçi yöntem önerilmiştir. Uzun süredir kullanılan oskültasyon yöntemiyle toplanan akciğer seslerini otomatik olarak sınıflandırmak amacıyla yarı gözetimli bir derin öğrenme algoritması geliştirilmiştir. Veriler, cihaza özel geliştirilen bir mobil uygulama ve uygun maliyetli bir elektronik stetoskop aracılığıyla toplanmıştır. Etiketleme ihtiyacını minimuma indiren bu yöntem, en yaygın akciğer seslerinden olan hırıltı (wheeze) ve çıtırtı (crackle) sınıflarında sırasıyla 0.86 ve 0.74 AUC değerlerine sahip ROC eğrileri elde ederek başarılı bir performans sergilemiştir (Chamberlain vd., 2016).

Bu doğrultuda, ses sinyallerindeki gürültüyü azaltmak amacıyla z-puanı normalizasyonu (z-score normalization) ve dalgacık yumuşatma (wavelet denoising) gibi ön işleme teknikleri uygulanmıştır. Ardından, farklı akciğer seslerini ayırt edebilmek için Evrişimli Sinir Ağı (CNN) mimarisi, Mel Frekans Kepstral Katsayıları (MFCC) özellikleriyle birlikte hibrit bir ağ modeli olarak eğitilmiş; bu model ile %95 eğitim doğruluğu ve %85 test doğruluğu elde edilmiştir. Bu sonuçlar, geliştirilen sistemin klinik ortamlarda uygulanabilirliğini desteklemektedir (Koravanavar vd., 2023).

Öte yandan, farklı bir çalışmada ön işleme adımlarında ham ses kayıtları sinyal hâline getirilerek sırasıyla sinyal işleme, Fourier dönüşümü ile frekans analizi, spektral gürültü artırımı, doku analizi, öznitelik çıkarımı ve veri artırımı işlemleri gerçekleştirilmiştir. Bu işlemlerin ardından AlexNet, VGG-B1, VGG-B3, VGG-D1, ResNet ve InceptionNet gibi farklı derin öğrenme mimarileriyle sınıflandırma yapılmış ve %95 ile %100 arasında doğruluk oranları elde edilmiştir. Bu sonuçlar, derin öğrenme yöntemlerinin solunum hastalıklarının tespitinde yüksek başarı sağladığını göstermektedir (Zulfiqar vd., 2021).

3. GEREÇ VE YÖNTEM

3.1. Veri Seti

Bu çalışmada, çeşitli solunum yolu hastalıklarının sınıflandırılmasını kolaylaştırmak için elde edilen stetoskopik akciğer sesi verileri entegre edildi. Bu veri setleri literatürde yaygın olarak referans kaynakları olarak kabul edilir ve sıklıkla farklı metodolojilerin karşılaştırmalı performansını değerlendirmek için kullanılır. Veri seti, kamuya açık ICBHI 2017 Solunum Sesi Veritabanı (Minaee vd., 2020), 2017 yılında düzenlenen Uluslararası Biyomedikal ve Sağlık Bilişim Konferansı kapsamında geliştirilmiştir. Solunum yolu enfeksiyonları, astım, bronşit, sağlıklı, zatürre ve koah dahil olmak üzere çeşitli solunum bozukluklarını kapsayan, farklı yaş gruplarını (çocuklar, yetişkinler ve yaşlılar) kapsayan 404 kişiden toplanan 1.211 kayıttan oluşmaktadır. Bu kayıtlar benzer şekilde 4 kHz örnekleme frekansında çeşitli elektronik stetoskop sistemleri kullanılarak elde edilmiştir. Her kaydın süresi beş saniyedir ve en az bir tam solunum döngüsünü yakalamayı amaçlamaktadır. Ayrıca, tüm veri kaynaklarında tutarlılığı sağlamak için tüm ses kayıtları 4000 Hz'lik tekdüze bir örnekleme frekansına yeniden örneklenmiştir (Tawfik vd., 2022). Veri setini gösterir ses sayısı aşağıda gösterilmiştir.

Çizelge 3.1. Veri setindeki ses sayısı

	Kategori	Ses Sayısı
1	Astım	288
2	Bronşit	104
3	KOAH	401
4	Sağlıklı	133
5	Zatürre	285

Bu çalışmanın amacı, öksürük seslerinden çıkarılan veriler kullanılarak akut solunum yolu hastalıklarının doğru ve erken tespiti için yenilikçi bir analiz yöntemi geliştirmektir. Bu bağlamda, bu çalışmada geliştirilen model, yalnızca Conv1D tabanlı bir CNN+LSTM mimarisi temelinde tasarlanmıştır. Modelin amacı, MFCC temsilleri üzerinden çok sınıflı solunum hastalıklarını sınıflandırmaktır. Bu yaklaşım, hem sinyal işleme, hem de derin öğrenme tekniklerini entegre eden çok aşamalı bir

sistem sunarak literatüre katkıda bulunmayı amaçlamaktadır. Akut Solunum Yolu Hastalıklarının Tespiti için MFCC öznitelik çıkarımı ve 1D-CNN+LSTM tabanlı önerilen mimarinin sözde kodunu göstermektedir. Önerilen çerçevenin bir akış şema, mimarisi sunmaktadır.

Çizelge 3.1. Önerilen Çerçevenin Sahte Kodu

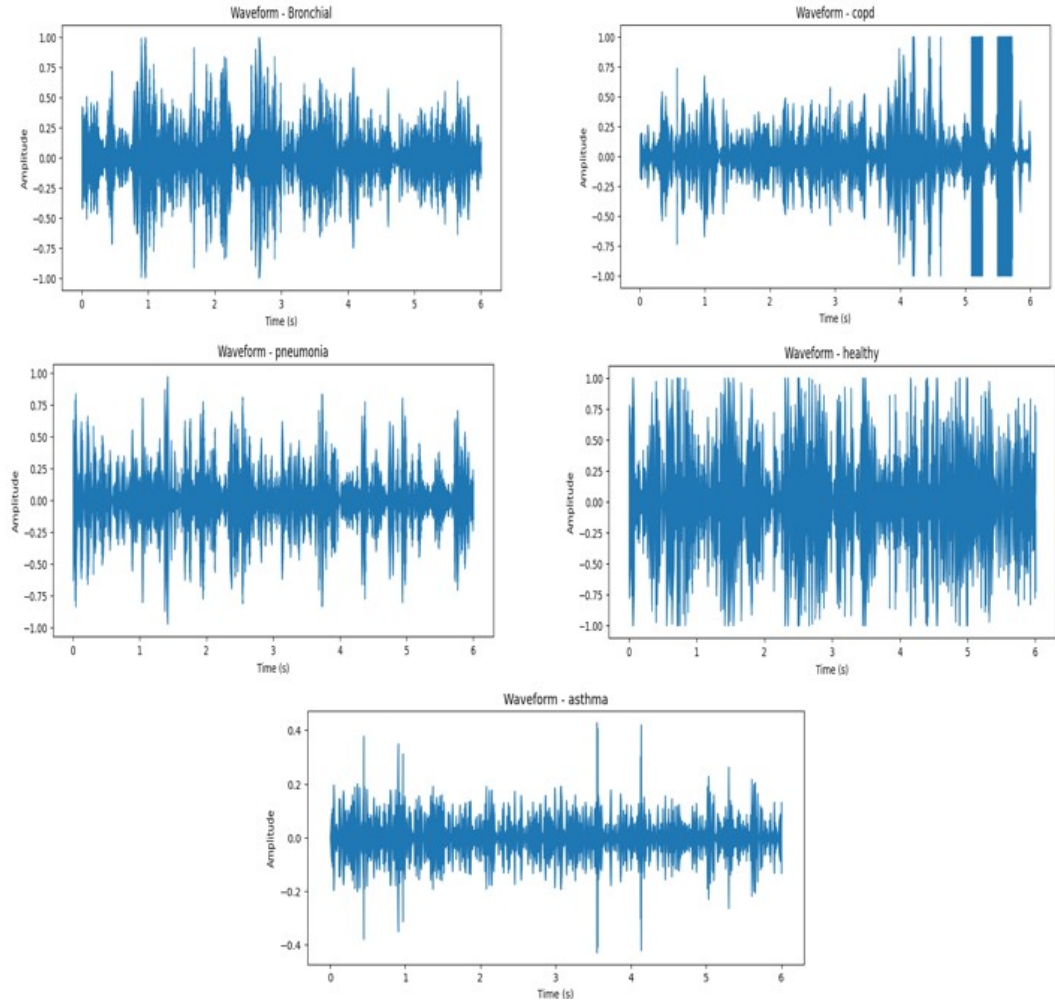
Girdi: Öksürük ses sinyalleri $S = \{s_1(t), s_2(t), \dots, s_n(t)\}$, her sinyal 44.1–50 kHz aralığında örneklenmiştir. Çıktı: $r = [\text{dogruluk}, \text{kesinlik}, \text{duyarlılık}, \text{f1-skoru}]$ (rastım, (rbronşit), (rkoah), (rsağlıklı), (rpnömoni),]	
1:	Başlat Boş özellik kümesini başlat $F=$, etiket kümesi $Y=$
2:	Her bir sinyal $s \in S$ için:
3:	$\text{kategori_sayımı} \leftarrow$ her bir category için .wav dosya sayısını hesapla
	$\text{max_count} \leftarrow$ tüm sınıflardaki en büyük ses dosyası sayısı her bir category için: eksik sayıda örnek varsa: random ses dosyalarını seç
4:	MFCC Özellik Çıkartımı
	veri kümesindeki her ses için : $y, sr \leftarrow \text{yükle}(\text{audio})$ $\text{mfcc} \leftarrow \text{mfcc_özellik_çıkart}(y, sr, n_mfcc=13)$ $\text{özellikler.ekleme}(\text{ortalama}(\text{mfcc}, \text{zaman_ekseninde}))$
5:	Etiket Kodlama (Label Encoding)
	$\text{label_encoder.fit}(y_labels)$ $y_encoded \leftarrow$ etiketleri sayıya çevir
6:	Veriler %80 eğitim ve %20 test olacak şekilde bölünür.
7:	$X_train, X_test, y_train, y_test \leftarrow \text{train_test_split}(X, y_encoded, \text{test_size}=0.2)$
	Conv1D+LSTM Tabanlı Derin Öğrenme Modelinin Oluşturulması Modelin Derlenmesi ve Eğitilmesi
8:	$\text{model.compile}(\text{optimizer}='adam', \text{loss}='sparse_categorical_crossentropy', \text{metrics}=['accuracy'])$ $\text{model.fit}(X_train, y_train, \text{epochs}=50, \text{batch_size}=32, \text{validation_split}=0.2)$
9:	Modelin Test Verisi Üzerinde Değerlendirilmesi
	$y_pred \leftarrow \text{model.predict}(X_test)$ $y_pred_labels \leftarrow \text{argmax}(y_pred)$ hesapla: $rC=\text{Doğruluk } C, \text{Kesinlik } C, \text{Duyarlılık } C, \text{F1 skoru } CC$
10:	Değerlendirme metriklerini hesapla: Doğruluk C, Kesinlik C, Duyarlılık C, F1 skoru C
11:	Performansın Görselleştirilmesi çiz: $\text{confusion_matrix}(y_test, y_pred_labels)$ çiz: eğitim vs doğrulama accuracy/loss grafikleri

3.2. Model Mimarisi

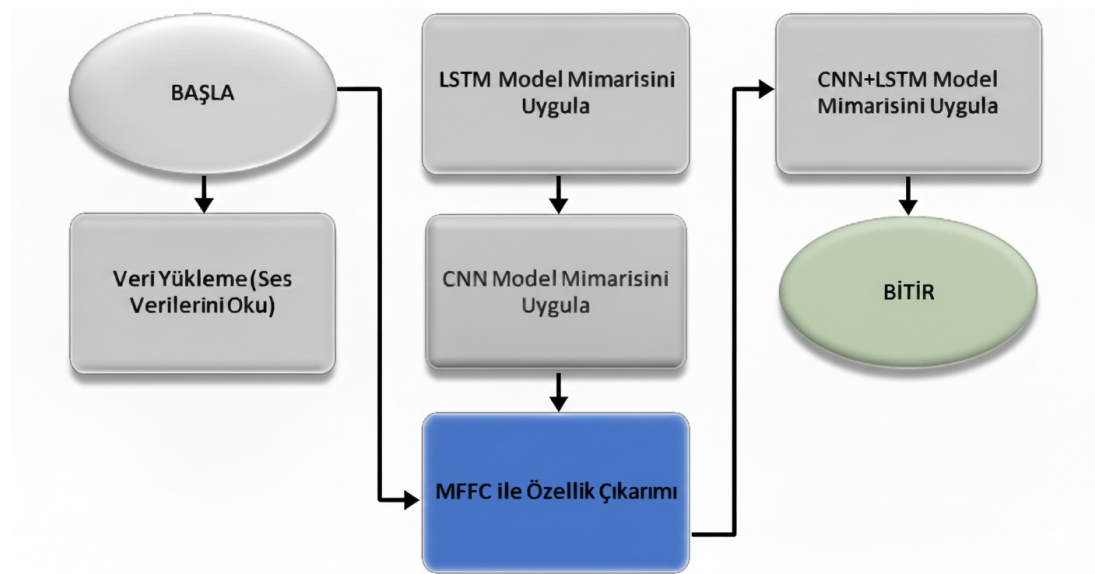
Bu çalışmada geliştirilen model, Conv1D tabanlı bir CNN+LSTM mimarisi temelinde tasarlanmıştır. Modelin amacı, öksürük seslerinin sınıflandırılmasıdır. Bu doğrultuda, ses sinyallerinden Mel Frekans Kepstrum Katsayıları (MFCC) yöntemi ile öznelik haritaları elde edilmiştir. ICBH veri setinden çıkarılan bu öznelikler, önerilen derin öğrenme mimarisi kullanılarak sınıflandırılmıştır. Önerilen modelde, ICBH veri setindeki çevresel sesleri sınıflandırmak amacıyla 1D-CNN mimarisi tercih edilmiştir.

Model, giriş olarak 13 boyutlu MFCC öznelik vektörlerini almaktadır. İlk aşamada, 32 filtreli ve ReLU aktivasyon fonksiyonuna sahip bir Conv1D katmanı kullanılmakta, ardından 64 filtreli ikinci bir Conv1D katmanı gelmektedir. Bu katmanların ardından, havuzlama işlemi için havuz boyutu 2 olan bir MaxPooling1D katmanı uygulanmıştır. Elde edilen öznelik haritaları, Flatten katmanı aracılığıyla tek boyutlu bir vektöre dönüştürülmüş ve sonrasında 64 nöronlu bir Dense katmanına aktarılmıştır. Çıkış katmanında ise softmax aktivasyon fonksiyonuna sahip bir Dense katmanı yer almakta olup, modelin çıktısı beş sınıfa ayrılmaktadır.

Modelin eğitimi, Adam optimizasyon algoritması ile gerçekleştirilmiş; kayıp fonksiyonu olarak ise etiketlerin tamsayı formatında sunulduğu çok sınıflı sınıflandırma problemlerine uygun olan "Sparse Categorical Crossentropy" fonksiyonu tercih edilmiştir. Ayrıca Astım, Bronşit, KOAH, Sağlıklı ve Zatürre olmak üzere beş kategoride sınıflandırılan öksürük seslerine ait ses sinyali örneği sunulmuştur.



Şekil 3.6. Bronşit, Koah, Pnömoni, Sağlıklı ve Astım Orijinal Sinyalleri.



Şekil 3.7. MFCC ve CNN+LSTM ile Önerilen Modelin Mimarisi

3.3. Mel Frekans Kepstrum Katsayıları (MFCC) ile Özellik Çıkarımı

Öksürük sinyalleri, zamanla değişen fraktal özellikler gösterebilen karmaşık, biyolojik kaynaklı sinyallerdir. Bu nedenle, geleneksel tek ölçekli analiz yöntemleri bu sinyallerin tüm karmaşıklığını yakalamak için yetersiz olabilir. MFCC, bu tür sinyallerin çok ölçekli yapısını ve uzun menzilli korelasyonlarını ortaya çıkarmak için kullanılan güçlü bir araçtır (Loey vd., 2020). Bu çalışmada, MFCC yöntemi öksürük seslerinden özellik çıkarmak için uygulandı. Her ses kaydından 13 boyutlu Mel Frekans Kepstrum Katsayıları (MFCC) çıkarılmıştır. MFCC, sesin zaman ve frekans bileşenlerini insan işitme sistemine benzer şekilde modelleyerek, ses sınıflandırma görevlerinde yaygın olarak kullanılmaktadır. MFCC altı ana adımdan oluşur. İlk olarak, Önileme (Pre-emphasis) aşağıdaki gibi hesaplanır:

Yüksek frekanslı bileşenleri güçlendirmek için:

$$y[n] = x[n] - x[n - 1], \text{burada } 0.97 \quad (3.1)$$

Böylece sinyaldeki yüksek frekans enerjisi daha belirgin hale gelir.

Çerçeveleme ve Pencereleme: Ses sinyali kısa zamanlı özellikler içerdiğinden sabit uzunlukta ve genellikle %50 bindirme (overlap) içeren pencerelere bölünür. Her çerçeveye Hamming veya Hanning penceresi uygulanarak kenar etkileri azaltılır.

$$w[n] = 0.54 - 0.46 \cos(2\pi n / (N - 1)) \quad (3.1)$$

Hızlı Fourier Dönüşümü (FFT): Her pencerenin frekans içeriği hesaplanır:

$$P[k] = |X[k]| \quad (3.1)$$

Mel Frekans Ölçeği Dönüşümü: Lineer frekans eksenini Mel ölçeğine dönüştürülür.

$$m = 2595 \log(1 + f/700) \quad (3.2)$$

Logaritmik Dönüşüm: Mel bant enerjileri logaritmaya alınır

$$E = \log(P[k]H[k]) \quad (3.3)$$

Ayrık Kosinüs Dönüşümü (DCT): Log enerjileri sıkıştırmak ve korelasyonu azaltmak için DCT uygulanır. Burada genellikle ilk 12–13 katsayı alınır (0. katsayı, log-enerji olarak ayrı tutulabilir).

$$c = E \cos[n/M(m - 0.5)] \quad (3.4)$$

MFCC, otomatik konuşma ve konuşmacı tanıma için yaygın olarak kullanılan oldukça etkili bir özellik çıkarma algoritmasıdır (Sunil vd., 2015). Bunun nedeni, MFCC'nin insan kulağının frekans seçiciliğini taklit ederek konuşmacıları yüksek doğrulukla ayırt edebilmesidir. Ayrıca, MFCC'ler genellikle değişiklik ve ses dalgası yapısından çok daha az etkilendikleri için tercih edilirler.

Son yıllarda, MFCC yöntemi birçok çalışmada CORSA sistemlerinde kullanılmış ve umut verici sonuçlar elde edilmiştir (Palaniappan vd., 2013; Aykanat vd., 2017). MFCC, ses sinyalinin düşük frekanslı bileşenlerinin yüksek frekanslı bileşenlerden daha kritik bilgi taşıdığını varsayarak doğrusal olmayan bir ölçekleme gerçekleştirir. MFCC analizi, frekans sarmalama dışında sepstral analize benzer. MFCC analizinde, frekans Mel-Scale'e göre sarılır. Literatürde MFCC'yi hesaplamak için çeşitli yöntemler mevcuttur. Hızlı Fourier Dönüşümü (FFT) tabanlı yöntem, bunlar arasında en yaygın kullanılan yöntemlerden biridir.

Ses verisinde öznelik çıkarımı, ses sinyallerinden anlamlı özelliklerin çıkarılması işlemidir. Bu özellikler, ses verisini sayısal bir formata dönüştürerek makine öğrenimi modellerinin anlayabileceği bir şekle getirir. Sesin özelliklerini elde etmemizde kullanılan bazı ses özellikleri vardır. Örneğin ses dosyaları genellikle dalga formunda gelir ve bu dalga formu, örnekleme frekansı ve bit derinliği gibi özelliklere sahiptir. Çalışmada ses verisi elde edildikten sonra öznelik çıkarımı için Mel-Frekans Öznelik Katsayıları (MFCC) yöntemi kullanılmıştır. MFCC sesin frekans içeriğini temsil etmek için birçok konuşma tanıma sisteminde yaygın olarak kullanılan bir özelliktir (Young, 1996). MFCC, bir ses sinyalinin özelliklerini temsil etmek için kullanılan bir dizi katsayıdır ve bu katsayılar sesin frekans ve zaman

özelliklerini temsil eder. MFCC, insan kulağı ölçeğine dayanan Mel ölçeğini kullanarak frekans alanında temel alan konuşma tanıma sistemlerinde kullanılan en popüler özellik çıkarım tekniklerinden biridir (Shen vd., 2011). Mel-Frekans Öznitelik Katsayıları ses sinyalleri nin zaman-frekans alanındaki özelliklerini yakalamak için ses sinyalinin kesilmesi, pencereleme, Fourier dönüşümü, Mel-Frekans Filtre Bankası Uygulaması, logaritmik dönüşüm olmak üzere bir dizi adımdan geçer. Bu aşamaların sonunda her çerçeve için bir MFCC vektörü elde edilir. Bu vektörler, ses sinyalinin zaman içindeki değişimini temsil eder ve makine öğrenimi modelleri için girdi olarak kullanılabilir.

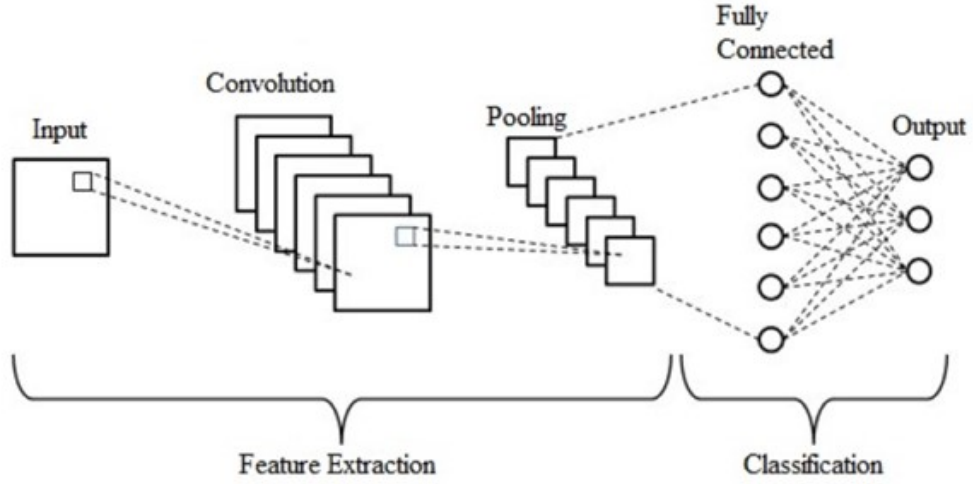
Her ses örneğinden elde edilen MFCC'ler zaman boyunca ortalanarak sabit uzunluklu öznitelik vektörleri oluşturulmuştur. MFCC özellik çıkarma adımları şu şekildedir: Sinyal örnekleme: Sinyalin örnekleme frekansı belirlenir ve sinyal bu frekansta örneklenir. Hamming penceresi: Sinyal, her bir örneği birbirinden ayrılmış kısa sürelerle işlenir. Bu işlem, "hamming penceresi" olarak adlandırılır. FFT: Hamming penceresi boyunca sinyal işlendikten sonra, FFT (Fast Fourier Transform) uygulanır ve sinyal frekans bileşenlerine ayrılır. Mel Filtre Bankası: Frekans bileşenleri, Mel ölçeğinde ölçeklendirilir. Mel ölçeği, insan işitme sisteminin duyarlılığına benzer bir şekilde ölçeklendirilmiştir. Logaritmik dönüşüm: Mel filtre bankası çıktısı üzerinde logaritmik dönüşüm uygulanır. Cepstral katsayıları: Logaritmik dönüşüm sonrası çıktı, cepstral katsayıları olarak adlandırılır.

3.4. Evrişimli Sinir Ağları (Convolutional Neural Networks - CNN)

CNN'ler, yerel özellikleri çıkarabilme yeteneği sayesinde özellikle görsel ve zamansal verilerin analizinde yüksek başarı sağlayan derin öğrenme mimarileridir. Evrişim (convolution), aktivasyon ve havuzlama (pooling) katmanlarından oluşan bu yapılar, çok katmanlı yapıları ile düşük düzeyden yüksek düzeye doğru karmaşık örüntüleri öğrenebilir. CNN'ler, ses sinyallerinin zaman-frekans temsilleri üzerinde (örneğin spektrogram ya da MFCC çıktıları) örüntü tanıma görevlerinde etkin şekilde kullanılmaktadır. CNN, uzamsal veya zamansal örüntüleri otomatik öğrenmek için

tasarlanmış bir derin öğrenme mimarisidir.

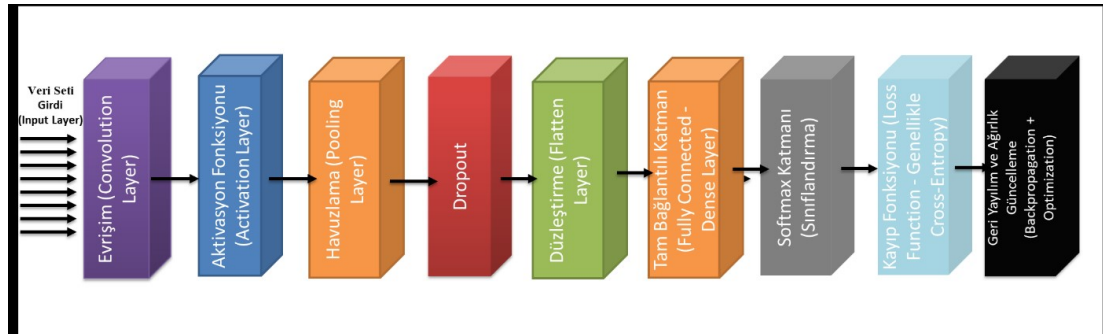
Çizelge 3.1. Evrişimli sinir ağının tipik mimarisi



CNN'ler, parametre paylaşımı ve hiyerarşik öznelik öğrenimi sayesinde görüntü tanıma, nesne tespiti ve zaman serisi analizinde (1D-CNN) yüksek performans gösterir. LeNet, AlexNet, ResNet gibi mimarilerle derin öğrenme devrimini başlatmıştır.

CNN, özellikle görüntü işleme ve sınıflandırma gibi alanlarda yaygın olarak kullanılan bir derin öğrenme modelidir. CNN'ler, verilerin mekânsal ve zamansal ilişkilerini anlamak ve yakalamak için katmanlar aracılığıyla özellik çıkarımı yapmaktadırlar. Görüntü sınıflandırma, nesne tanıma ve ses analizi gibi birçok sıralı ve yapısal veriyle çalışabilirler (Maggiori vd., 2017).

Çizelge 3.1. CNN Önerilen Derin Model



CNN, genellikle üç ana katman tipinden oluşmaktadır. Bunlardan ilki, evrişim katmanıdır (convolutional layer). Bu katman, girdi verisine küçük filtreler (kernels)

uygular. Bu filtreler, veri üzerinde kaydırılarak (stride) özellikler çıkarılır. Çıkan özellik haritaları, girişin mekânsal özelliklerini koruyarak daha derin bir seviyede temsil edilmesini sağlar. İkincisi, havuzlama katmanıdır (pooling layer). Bu katman, evrişim katmanından gelen özellik haritalarının boyutunu azaltmak ve işlem maliyetini düşürmek için kullanılır. Genellikle “Max Pooling” veya “Average Pooling” gibi yöntemler kullanılır. Max Pooling, belirli bir alan içindeki en yüksek değeri alarak boyutu küçültür. Sonuncu ana katman ise tam bağlantılı katmandır (fully connected layer). Bu katman, özellik haritalarını düzleştirerek (flatten) klasik yapay sinir ağına benzer şekilde tüm nöronları birbirine bağlar. Son olarak, çıkış katmanında softmax veya sigmoid aktivasyon fonksiyonu kullanılarak sınıflandırma yapılır (Hsu vd., 2023).

Bir CNN, sınıflandırma işlemlerinde belirli adımları takip ederek çalışır. İlk olarak, model bir girdi alır; bu genellikle bir görüntü veya sıralı veri şeklindedir. Ardından, evrişim ve havuzlama katmanları devreye girerek girdi verisinden özellikler çıkarır. Bu katmanlar, verinin mekânsal ve yapısal ilişkilerini kullanarak özellikleri giderek daha soyut bir şekilde temsil eder. Daha sonra, bu özellikler tam bağlantılı katmanlara iletilir ve burada işlenerek sınıf olasılıkları hesaplanır. Son adımda ise, çıktı katmanında her bir sınıf için olasılıklar verilir ve en yüksek olasılığa sahip sınıf, modelin tahmini olarak belirlenir. Bu süreç, CNN’nin veriyi işleyip doğru sınıflandırmayı yapmasını sağlar (Dubey vd., 2025).

3.5. Tek Boyutlu Evrişimsel Sinir Ağları (One Dimensional Convolutional Neural Networks)

CNN, derin öğrenme teknolojisindeki en etkili yöntemlerden biridir ve 1D, 2D ve 3D CNN dahil olmak üzere farklı ağ yapısı türleri vardır. Tek Boyutlu Evrişimsel Sinir Ağı (1D-CNN) genellikle ses sinyalleri gibi tek boyutlu sinyaller için kullanılır (LeCun vd., 2015). Teknik olarak, genel bir CNN modeli evrişim, havuzlama, tam bağlı katmanlar ve sınıflandırma katmanını içerir. Evrişim, giriş görüntüsünden bir özellik haritası çıkarmak için giriş görüntüsüne uygulanan filtredir. Görüntü matrisi ve filtre gibi iki girdi alan matematiksel bir işlemdir (Er, 2020). Evrişim işlemi, kenar algılama ve bulanıklaştırma gibi filtreler uygulanarak gerçekleştirilir. Filtrelerin yüksekliği ve ağırlığı giriş hacminden daha küçüktür. Evrişim işleminin formülü denklem (8)’de verilmiştir. Giriş görüntüsü f ile gösterilir ve çekirdeğimiz h ile gösterilir. Sonuç matrisinin satır ve sütun dizinleri sırasıyla m ve n ile gösterilir.

$$G[m, n] = (f * h)[m, n] = \sum_j \sum_k h[j, k] f[m - j, n - k]$$

(3.5)

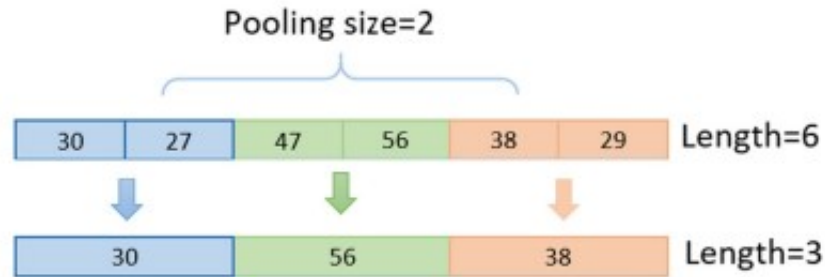
Aktivasyon fonksiyonu her bir evrişimden sonra kullanılır. Rektifiye Doğrusal Birim (ReLU) genellikle derin ağlarda aktivasyon fonksiyonu olarak kullanılır. ReLU'nun amacı modeli doğrusal olmayan bir yapıya dönüştürmektir. ReLU fonksiyonu denklem (9)'da verilmiştir.

$$ReLU(x) = \max(x, 0) \quad (3.6)$$

Denklem 9'a göre, ReLU çıktısı sıfır ile giriş değeri arasındaki maksimum değerdir. Giriş değeri negatif olduğunda, çıkış değeri sıfıra eşit olur, ancak giriş pozitif olduğunda, çıkış değeri giriş değerine eşit olur. Ek olarak, CNN'lerin bir havuzlama katmanı vardır. Havuzlama katmanı, evrişimden sonra giriş görüntüsünün mekansal hacmini azaltmak ve hesaplamaları hızlandırmak için kullanılır. 1D maksimum havuzlamanın bir örneğini göstermektedir. Tamamen bağlı katmanlar, görüntüleri tanıma ve sınıflandırmada çok başarılı olduğu kanıtlanmış derin ağların önemli bir bileşenidir. Bu sürecin sonucu, nihai sınıflandırma kararını yönlendiren tamamen bağlı bir sinir ağı yapısına bağlıdır. Tamamen bağlı giriş katmanı, giriş olarak önceki katmanların çıktılarını alır ve bunları düzleştirerek bir sonraki katman için giriş olabilecek tek bir vektöre dönüştürür. Özellik girişleri, tamamen bağlı giriş katmanından sonra alınır ve ağırlıklar doğru etiketi tahmin etmek için güncellenir. Son katman çıktı katmanıdır ve her etiket için olasılıklar tahmin edilir. Bu katmanda genellikle Softmax tercih edilir. Softmax formülü denklem (10)'da verilmiştir.

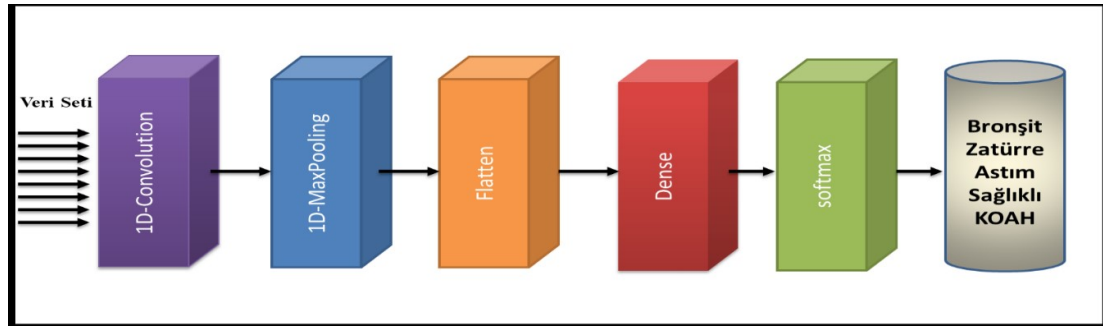
$$Softmax(x)_j = (e^{x_j}) / (\sum_{n=1}^N e^{x_n}), j = 1, \dots, N \quad (3.7)$$

Çizelge 3.1. Conv1D max pooling



Conv1D, yani 1-boyutlu evrişimsel sinir ağı katmanı (1D Convolutional Layer- Evrişimsel Katman), derin öğrenmede özellikle zaman serileri, ses verisi veya metin gibi ardışık (sıralı) veriler üzerinde çalışan bir yapay sinir ağı katmanıdır. "1D", yani "bir boyutlu" ifadesi, evrişim işleminin tek ekseninde – genellikle zaman ekseninde – uygulandığını belirtir. Conv1D, giriş verisinin zaman içindeki örüntülerini (pattern) öğrenmek için kullanılır. Örneğin: Ses sinyalinde ton değişimleri, metin verisinde kelime dizileri, biyomedikal sinyallerde (ör. EKG) nabız dalgaları gibi örüntüleri yakalamaya çalışır. Diyelim ki elinde bir ses sinyalinden çıkarılmış MFCC matrisi var: Boyutu 100 zaman adımı $\times$ 13 öznitelik. Conv1D katmanı burada 13 öznitelik boyunca değil, zaman eksenini boyunca kayan küçük pencereler (filtreler) kullanarak verideki örüntüleri tanımlar. Zaman içindeki ilişkileri yakalamada başarılıdır. Veri boyutunu azaltmadan anlamlı öznitelikler çıkarır. Daha az parametreyle eğitilebildiği için hızlı ve hesaplaması kolaydır.

Çizelge 3.1. Conv1D Önerilen Derin Model Adımları



Bu çalışmada geliştirilen model, yalnızca Conv1D tabanlı bir CNN mimarisi temelinde tasarlanmıştır. Modelin amacı, MFCC temsilleri üzerinden çok sınıflı solunum hastalıklarını sınıflandırmaktır. Girdi olarak 13 boyutlu MFCC öznitelik vektörlerini alan model, ilk aşamada 32 filtreli ve ReLU aktivasyon fonksiyonlu bir Conv1D katmanı, ardından 64 filtreli ikinci bir Conv1D katmanı ile devam etmektedir. Bu katmanları takiben, havuzlama işlemi için MaxPooling1D katmanı (pool size=2) kullanılmıştır. Elde edilen öznitelik haritaları, Flatten katmanı ile tek boyutlu hâle getirilmiş ve ardından 64 nöronlu bir Dense katmanına aktarılmıştır. Son katmanda ise softmax aktivasyon fonksiyonuna sahip Dense katmanı yer almakta olup, verileri beş sınıfa ayırmaktadır.

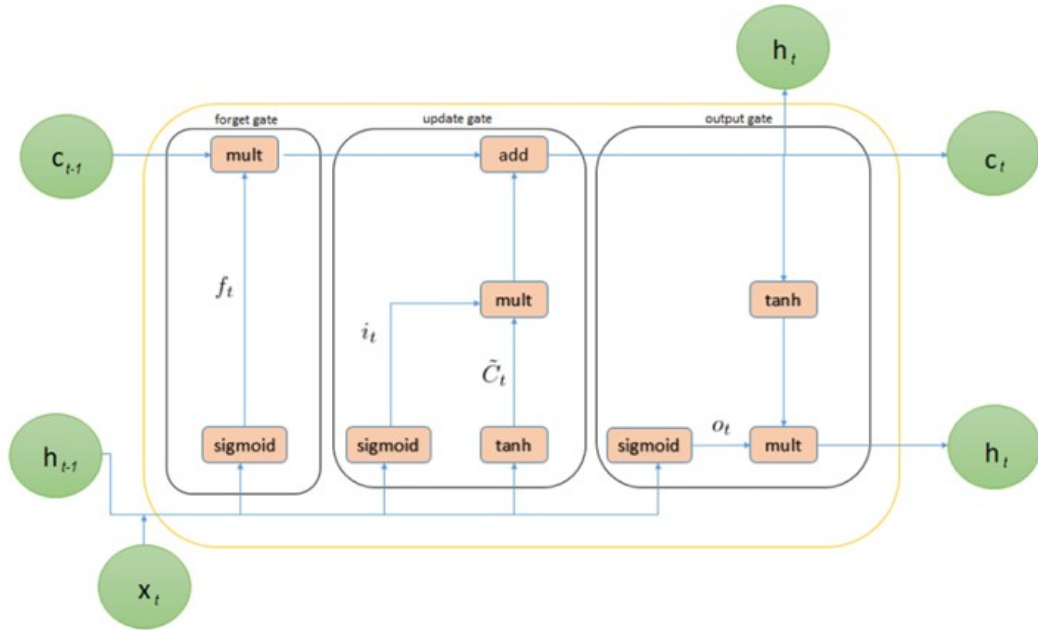
3.6. LSTM ile Özellik Çıkarımı

Öksürük sesi sinyalleri zamanla değişen dinamik bir yapıya sahip olduğundan, bu tür ardışık verilerden zamana bağlı özellikler çıkarmak için LSTM

ağları kullanıldı (Yu vd., 2019). LSTM, klasik RNN'lerin uzun vadeli bağımlılıkları öğrenmedeki sınırlamalarını ele almak için geliştirilen bir tür tekrarlayan sinir ağıdır (RNN). Bir LSTM hücresi aşağıdaki gibi ifade edilir:

$$f_t = (W_f[h_{t-1}, x_t] + b_f) i_t = (W_i[h_{t-1}, x_t] + b_i), C_t = \tanh(W_C[h_{t-1}, x_t] + b_C) \quad (3.8)$$

Çizelge 3.2. Uzun kısa süreli belleğin iç yapısı.

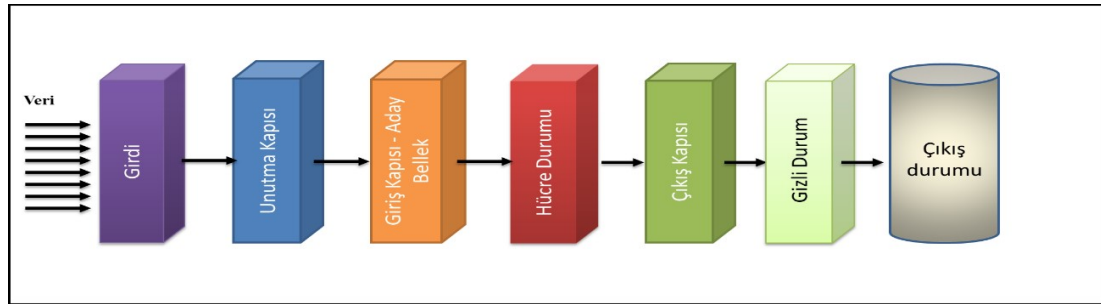


Burada, x_t t anındaki giriş vektörünü, h_t t LSTM'nin çıktısını, C_t hücre durumunu, σ aktivasyon fonksiyonunu ve $\square$, W ve b sırasıyla eleman bazlı çarpmayı, ağırlığı ve önyargıyı ifade eder. Ham ses sinyalinden zaman tabanlı özellikleri çıkarmak için tasarlanan LSTM ağı ağı gösterilmiştir. Bu mimaride, ses sinyali önceden işlenmeden doğrudan modele beslenir. Sinyal, sabit uzunluktaki örneklerin bir dizisi olarak normalleştirilir. Model, zaman içinde bağımlılıkları öğrenmek için bir LSTM katmanı ve ardından sabit boyutlu bir özellik vektörü elde etmek için bir Küresel Ortalama Havuzlama katmanı içerir.

LSTM, geleneksel RNN'lerden farklı olarak, uzun vadeli bağımlılıkları daha iyi öğrenebilmek için hücre durumu (cell state) ve çeşitli kapılar (gates) kullanmaktadır. Bu kapılar, hücre durumunun ne kadarının güncellenip ne kadarının tutulacağını belirlemektedir. Giriş kapısı, yeni gelen bilginin hücre durumuna ne kadar ekleneceğini kontrol etmektedir. Unutma kapısı, hücre durumundaki bilginin

ne kadarının unutulacağını belirlemektedir. Çıkış kapısı ise hücre durumunun hangi kısmının çıkışa aktarılacağını kontrol etmektedir (Dai vd., 2019). Bu kapılar, LSTM'nin sıralı verilere dayanarak kararlar almasını sağlar ve sınıflandırma görevlerinde verinin zamansal özelliklerini dikkate alarak tahminler yaparlar. LSTM'nin sınıflandırma işlemlerinde çalışma şekline bakıldığında, öncelikle LSTM katmanına, sıralı veri girilir. Her veri noktası, belirli bir zamanda (time step) olan veriyi temsil etmektedir. Ardından, LSTM katmanları, zaman içindeki bağımlılıkları ve kalıpları öğrenir. LSTM'den elde edilen özellikler, sınıflandırma katmanına (genellikle dense veya fully connected layer) aktarılır ve son olarak softmax veya sigmoid aktivasyon fonksiyonu ile sınıf olasılıkları hesaplanır (Jongjaraunsuk vd., 2024).

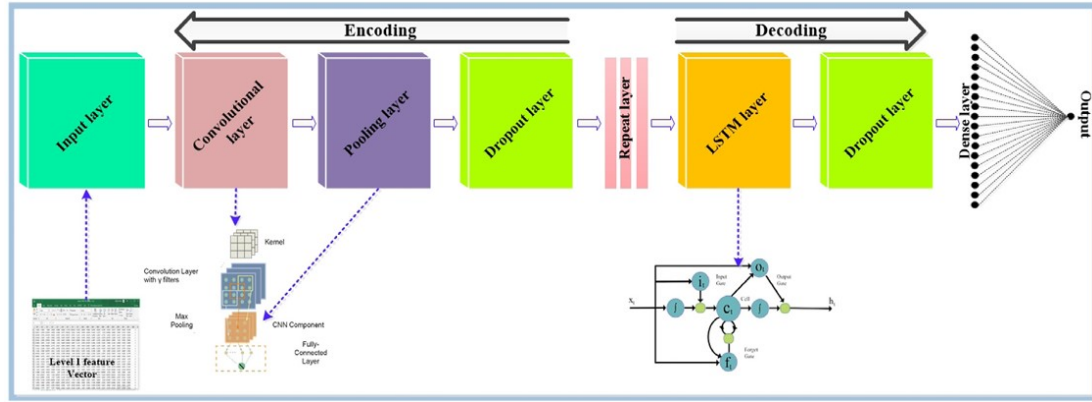
Çizelge 3.3. LSTM Modelin Adımları-Akış Şeması



Şekil 'de gösterilen LSTM (Long Short-Term Memory) modelinin adımları, sırasıyla veri girişinden başlayarak bilgi işleme sürecini tanımlamaktadır. İlk olarak model, dış ortamdan gelen veriyi girdi katmanında alır. Ardından unutma kapısı, önceki hücre durumundan hangi bilgilerin silineceğini belirler. Giriş kapısı ve aday bellek, yeni bilginin hücreye ne ölçüde aktarılacağını kontrol ederken, hücre durumu bu süreçte güncellenir. Sonrasında çıkış kapısı, hücre durumundan hangi bilgilerin çıkışa aktarılacağını belirler ve bu bilgi gizli durum katmanında temsil edilir. Nihayetinde gizli durumdan elde edilen çıktı, modelin nihai çıkış durumunu oluşturur. Bu yapı, zaman serileri ve ardışık verilerle çalışırken uzun vadeli bağıntıları öğrenmeyi mümkün kılar.

Ham ses sinyalinin zaman tabanlı özellikleri çıkarmak için tasarlanan LSTM ağı Tablo'da gösterilmiştir. Bu mimaride, ses sinyali önceden işlenmeden doğrudan modele beslenir. Sinyal, sabit uzunluktaki örneklerin bir dizisi olarak normalleştirilir. Model, zaman içinde bağımlılıkları öğrenmek için bir LSTM katmanı kullanmaktadır.

Çizelge 3.4. Önerilen CNN-LSTM otomatik kodlayıcı ağı



Şekilde, önerilen CNN-LSTM otomatik kodlayıcı ağ mimarisini sunmaktadır. Model, giriş katmanından alınan düzey 1 özellik vektörünü, sırasıyla evrimsel katman, havuzlama katmanı ve bırakma (dropout) katmanı aracılığıyla işleyerek özniteliklerin soyut bir temsiline dönüştürür. Bu aşama, kodlayıcı (encoding) kısmını oluşturur. Ardından, tekrar eden katman (repeat layer) ile elde edilen temsil, LSTM katmanına aktarılır ve bu katman zaman bağımlı bilgileri modelleyerek ağın çözümleyici (decoding) kısmını oluşturur. Son bırakma katmanından sonra, tam bağlantılı (dense) katman aracılığıyla çıkış üretilir. Bu yapı, uzamsal öznitelik çıkarımı için CNN ve zamansal bağımlılıkların modellenmesi için LSTM yapılarını birleştirerek, karmaşık veri örüntülerinin daha etkin şekilde öğrenilmesini sağlamaktadır.

3.7. Değerlendirme metrikleri

Bu çalışmada, makine öğrenmesi sınıflandırma algoritmalarının başarısını değerlendirmek için Doğruluk (Accuracy), Kesinlik (Precision), Duyarlılık (Recall), F1 Puanı (F1-Score) gibi metrikler kullanılmıştır. Doğruluk, toplam tahminlerin ne kadarının doğruluğunu göstermektedir. Kesinlik, pozitif olarak tahmin edilen durumların gerçekte ne kadarının pozitif olduğunu göstermektedir. Duyarlılık veya hassasiyet, gerçek pozitif durumların ne kadarının doğru tahmin edildiğini göstermektedir. F1 Puanı, kesinlik ve duyarlılık değerlerinin harmonik ortalamasıdır ve dengeli bir ölçümdür. Bu metriklerin her biri için matematiksel denklemler aşağıda verilmiştir.

$$Dogruluk = (TP + TN) / (TP + TN + FP + FN)$$

(3.9)

$$Hassasiyet = TP / (TP + FP) \quad (3.10)$$

$$F1Skor = 2 * ((Precision * Recall) / (Precision + Recall)) \quad (3.11)$$

$$F1Skor = 2 * ((Precision * Recall) / (Precision + Recall)) \quad (3.12)$$

Sınıflandırma modellerinin performansı, tahmin sonuçlarının çeşitli metriklerle değerlendirilmesiyle ölçülmektedir. Bu kapsamda, Doğru Pozitif (DP/TP) modelin pozitif sınıfı doğru tahmin ettiği durumları, Doğru Negatif (DN/TN) negatif sınıfın doğru tahmin edildiği durumları, Yanlış Pozitif (YP/FP) negatif örneklerin hatalı biçimde pozitif sınıfa atandığı durumları ve Yanlış Negatif (YN/FN) pozitif örneklerin yanlışlıkla negatif sınıfa atandığı durumları ifade etmektedir. Ayrıca, sınıflandırma raporlarında yer alan “support” metriği, her sınıfa ait toplam örnek sayısını göstermekte ve özellikle dengesiz veri setlerinde sınıf dağılımının model performansına etkisini ortaya koymaktadır.

Bu çalışmada modelin performansı, karmaşıklık matrisi (confusion matrix) üzerinden değerlendirilmiştir. Karmaşıklık matrisi, gerçek sınıf etiketleri ile model tahminlerini karşılaştırarak dört temel kategori (TP, FP, TN, FN) elde edilmesini sağlar. Bu yapı, yalnızca sınıflandırma hatalarının kaynağını görselleştirmekle kalmaz, aynı zamanda doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1 skoru gibi temel performans ölçütlerinin hesaplanmasına da olanak tanır (Memiş vd., 2022).

4. BULGULAR

Geliştirilen derin öğrenme tabanlı modelin performansı, deneysel olarak kapsamlı biçimde değerlendirilmiştir. Veri kümesi, %80 eğitim ve %20 test olacak şekilde rastgele bölünmüştür. Eğitim sürecinde, batch boyutu 32, öğrenme oranı 0.001 olarak belirlenmiş ve modelin optimizasyonu için Adam algoritması kullanılmıştır. Her ses örneğinden çıkarılan 13 boyutlu MFCC öznitelikleri, modelin giriş vektörlerini oluşturmuştur.

Model, Conv1D tabanlı CNN+LSTM mimarisi ile inşa edilmiş ve 50 epoch boyunca eğitilmiştir. Eğitim süreci boyunca doğruluk oranının istikrarlı şekilde arttığı, doğrulama doğruluğunun ise yalnızca CNN kullanılarak %85 doğruluk, yalnızca LSTM kullanılarak %81 doğruluk, hibrit model CNN+LSTM birlikte kullanılarak ise %86 doğruluk seviyelerine çıktığı gözlemlenmiştir. Bu sonuçlar, modelin hibrit model ile hem öğrenme kabiliyetini hem de genelleme başarımını daha iyi olduğunu ortaya koymaktadır.

Her sınıfın örneklerinden çıkarılan CNN Algoritmasının Hastalık Bazlı Sınıflandırma Performansı dağılımı, Tablo 'da gösterilmiştir.

Hastalık Adı	Precision (hassasiyet)	Recall (duyarlılık)	F1-Score (F1-Puanı)	Destek (Support)
Astım	0.91	0.96	0.93	80
Bronşit	1.00	0.72	0.84	29
KOAH	0.81	0.86	0.83	56
Sağlıklı	0.78	0.67	0.72	21
Zatürre	0.78	0.82	0.80	57

Şekil 4.2. CNN Algoritmasının Hastalık Bazlı Sınıflandırma Performansı

Tablo'da CNN algoritmasının hastalık bazlı sınıflandırma performansı değerlendirilmektedir. Elde edilen sonuçlara göre, astım sınıfı %91 hassasiyet, %96 duyarlılık ve %93 F1-skoru ile en yüksek başarıya ulaşmıştır. Bronşit sınıfında hassasiyetin %100 gibi oldukça yüksek olduğu, ancak duyarlılığın %72 seviyesinde kaldığı görülmektedir; bu durum, modelin bronşit tanımlarında seçiciliğinin yüksek, ancak bazı vakaları atlama eğiliminde olduğunu göstermektedir. KOAH sınıfı %81

hassasiyet ve %86 duyarlılık ile dengeli bir performans ortaya koyarken, zatürre sınıfında bu değerler %78 ve %82 olarak gerçekleşmiştir. Buna karşın, sağlıklı sınıfında doğruluk oranlarının diğer sınıflara kıyasla daha düşük olduğu (%78 hassasiyet, %67 duyarlılık, %72 F1-skoru) dikkat çekmektedir.

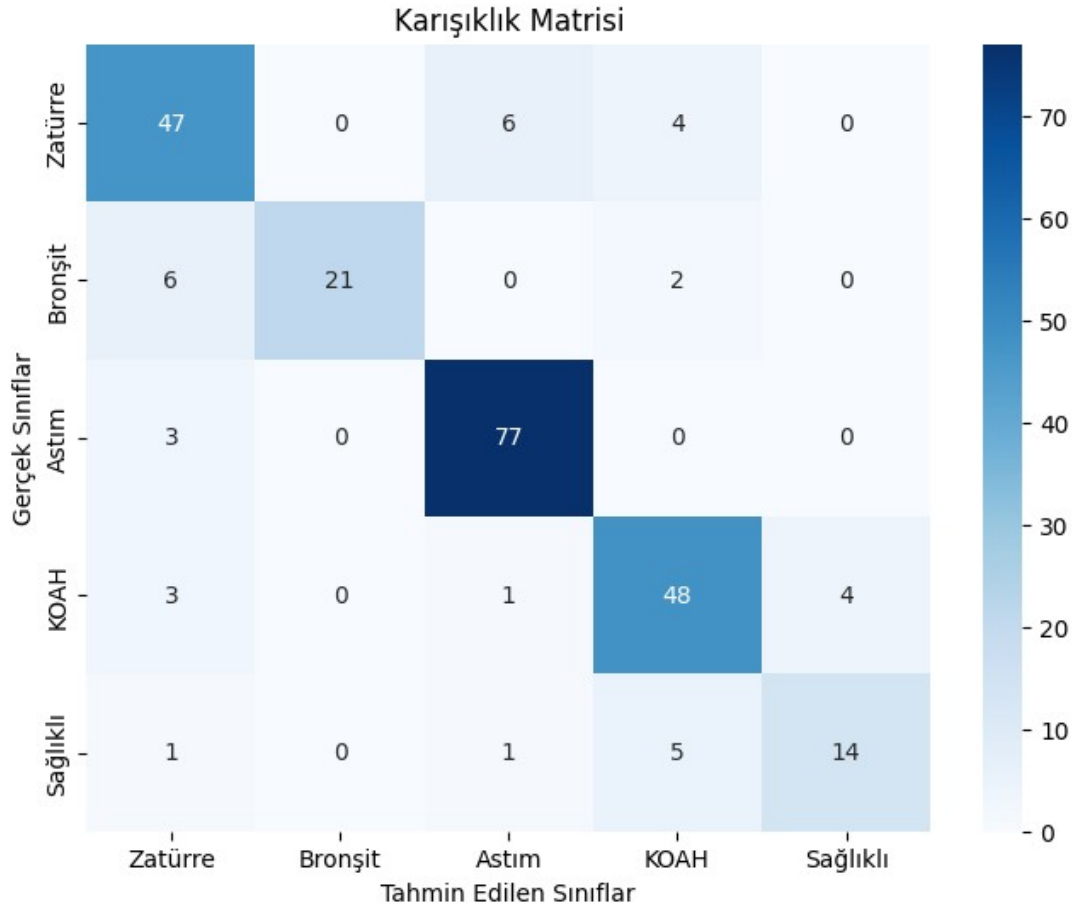
Genel olarak tablo, CNN tabanlı modelin özellikle astım, KOAH ve zatürre gibi hastalıkların tespitinde yüksek ve dengeli performans sergilediğini, ancak sağlıklı bireylerin sınıflandırılmasında görece daha düşük başarı sağladığını göstermektedir. Bu sonuç, modelin hastalıkların ayırt edilmesinde güçlü bir performans sergilerken, sağlıklı sınıfın doğru tahmininde ek iyileştirmelere ihtiyaç duyabileceğini ortaya koymaktadır.

Algoritma Adı	Doğruluk (Accuracy)	Hassasiyet (Precision)	Duyarlılık (Recall)	F1-Skore
CNN	0.8519	0.8560	0.8519	0.8504

Şekil 4.2. CNN Algoritmasının Genel Performans Değerleri

Tablo 'da, CNN algoritmasının genel performans ölçütleri sunulmaktadır. Buna göre model, %85.19 doğruluk oranı ile tüm sınıflar dikkate alındığında yüksek bir sınıflandırma başarısı göstermiştir. Hassasiyet (%85.60) ve duyarlılık (%85.19) değerlerinin birbirine oldukça yakın olması, modelin hem pozitif örnekleri doğru saptamada hem de yanlış sınıflandırmaları azaltmada dengeli bir performans ortaya koyduğunu göstermektedir. Ayrıca f1-skorunun (%85.04) benzer seviyede olması, hassasiyet ile duyarlılık arasındaki uyumu desteklemektedir.

Genel olarak değerlendirildiğinde, CNN tabanlı modelin sınıflar arasında tutarlı ve dengeli bir öğrenme gerçekleştirdiği, dolayısıyla akut solunum yolu hastalıklarının tespitinde güvenilir bir sınıflandırıcı olarak kullanılabileceği söylenebilir.

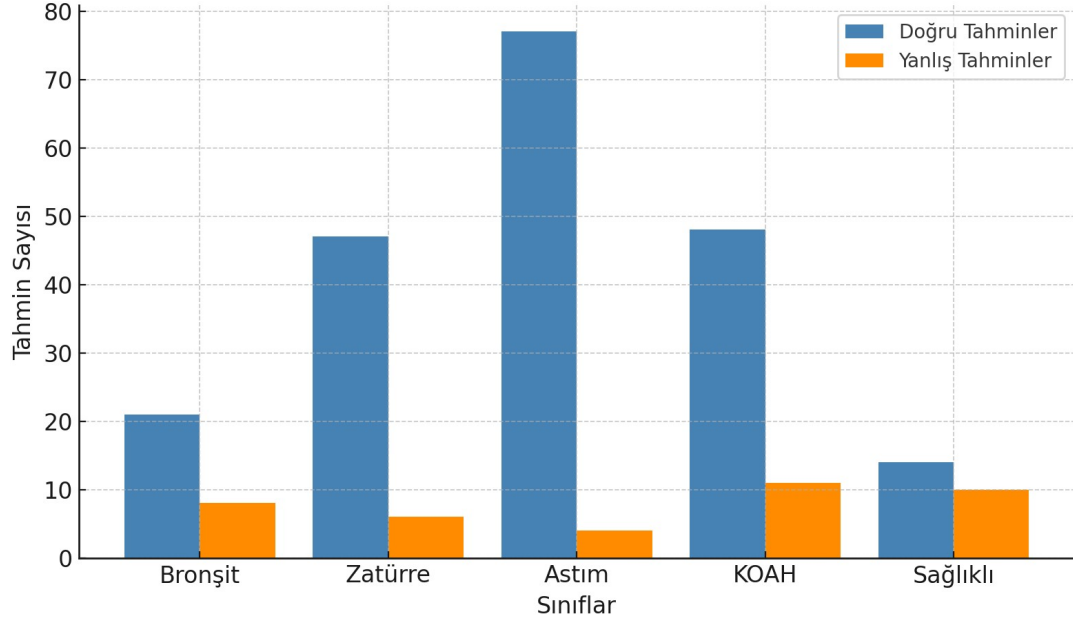


Şekil 4.2. CNN Modelin Karışıklık Matrisi

Bu karışıklık matrisi, modelin beş farklı solunum hastalığını ayırt etme performansını göstermektedir. Evrimsel Sinir Ağı (CNN) modelinin sınıflandırma performansını gösteren karışıklık matrisi incelendiğinde, modelin genel anlamda başarılı sonuçlar elde ettiği görülmektedir. Zatürre sınıfında 47 doğru tahmin gerçekleştirilmiş olup sınırlı düzeyde diğer sınıflarla karışma gözlenmiştir. Bronşit sınıfında 21 örnek doğru sınıflandırılmış olmakla birlikte özellikle Astım sınıfı ile belirgin bir karışma dikkati çekmektedir. Astım sınıfı 77 doğru tahmin ile en yüksek sınıflandırma başarısını göstermiştir. KOAH sınıfında 48 örnek doğru şekilde sınıflandırılmış ve yalnızca sınırlı sayıda hata meydana gelmiştir. Sağlıklı bireyler için ise 14 doğru sınıflandırma yapılmış, ancak bu sınıfın diğer sınıflarla belirli ölçüde karıştığı tespit edilmiştir. Özellikle Bronşit ve Sağlıklı sınıflarındaki hata oranlarının görece yüksek olması, bu sınıflara ait akciğer ses örüntülerinin diğer solunum hastalıklarına yapısal benzerlik göstermesinden kaynaklanıyor olabilir.

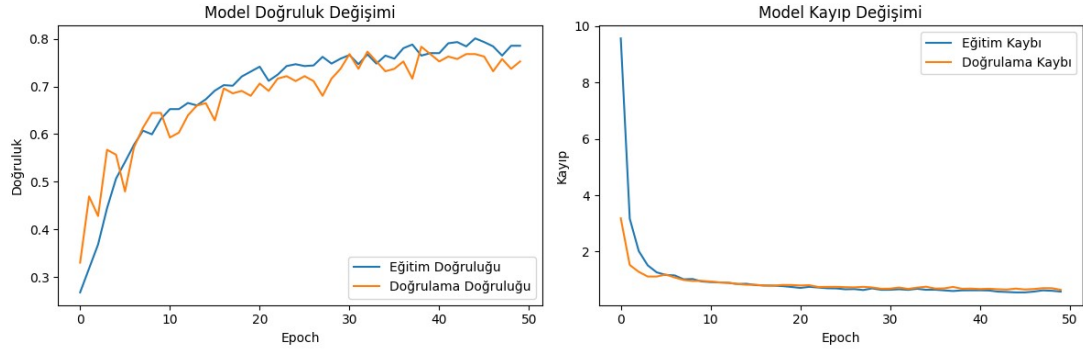
Ayrıca, sınıf bazlı tahmin performansı Şekil 'de sunulmaktadır. Karışıklık

matrisi Tablo’da incelendiğinde, modelin zamanla başarıyı artırdığı ve sınıflar arasında ayırım yapma kabiliyetini koruduğu gözlemlenmektedir. Precision, recall ve F1-skore gibi metriklerin dengeli dağılımı, modelin sınıflar arasında istatistiksel anlamda aşırı sapma göstermediğini desteklemektedir.



Şekil 4.2. CNN Sınıf Bazlı Doğru Ve Yanlış Tahmin Grafiği

CNN algoritmasının sınıf bazlı tahmin performansı görülmektedir. Model, en yüksek doğruluğu 77 doğru tahmin ile astım sınıfında elde etmiştir. Zatürre (47) ve KOAH (48) sınıflarında da doğru tahmin oranı yüksek, yanlış sınıflandırmalar ise sınırlı düzeyde kalmıştır. Buna karşın bronşit (21) ve özellikle sağlıklı (14) sınıflarında doğru tahmin sayıları düşük, hata oranları ise görece yüksektir. Bu sonuçlar, modelin solunum hastalıklarının tespitinde güçlü bir performans sergilediğini, ancak sağlıklı bireylerin doğru sınıflandırılmasında ek iyileştirmelere ihtiyaç duyulduğunu göstermektedir.



Şekil 4.2. CNN için Eğitim/doğrulama verilerindeki doğruluk/kayıp grafikleri

Conv1D tabanlı CNN modeli 50 epoch boyunca eğitilmiş ve doğruluk ile kayıp değerleri izlenmiştir. Doğruluk grafiğinde, hem eğitim hem de doğrulama doğruluklarının %80 seviyesine ulaştığı ve iki eğri arasındaki farkın düşük olduğu görülmektedir. Bu durum, modelin aşırı öğrenmeden kaçındığını göstermektedir. Kayıp grafiğinde ise eğitim ve doğrulama kayıplarının hızla azalarak düşük seviyelerde dengelendiği tespit edilmiştir. Genel olarak CNN modeli, istikrarlı bir öğrenme süreci sergilemiş ve eğitim ile doğrulama eğrilerinin birbirine yakın seyretmesi, yüksek genellenebilirlik sunduğunu ortaya koymuştur.

Hastalık Adı	Precision (hassasiyet)	Recall (duyarlılık)	F1-Score (F1-Puanı)	Destek (Support)
Astım	0,89	0,89	0,89	80
Bronşit	0,85	0,59	0,69	29
Koah	0,75	0,91	0,82	56
Sağlıklı	0,90	0,43	0,58	21
Zatürre	0,75	0,86	0,80	57

Şekil 4.2. LSTM Algoritmasının Hastalık Bazlı Sınıflandırma Performansı

LSTM algoritmasının hastalık bazlı sınıflandırma performansı sunulmaktadır. Sonuçlar incelendiğinde, astım sınıfı %89 hassasiyet, %89 duyarlılık ve %89 F1-skoru ile en dengeli ve yüksek başarıyı göstermektedir. KOAH sınıfı %75 hassasiyet ve %91 duyarlılık değerleriyle özellikle pozitif örnekleri doğru yakalama konusunda güçlüdür, ancak seçiciliği görece düşüktür. Zatürre sınıfında %75 hassasiyet ve %86 duyarlılık değerleri modelin dengeli bir performans sergilediğini

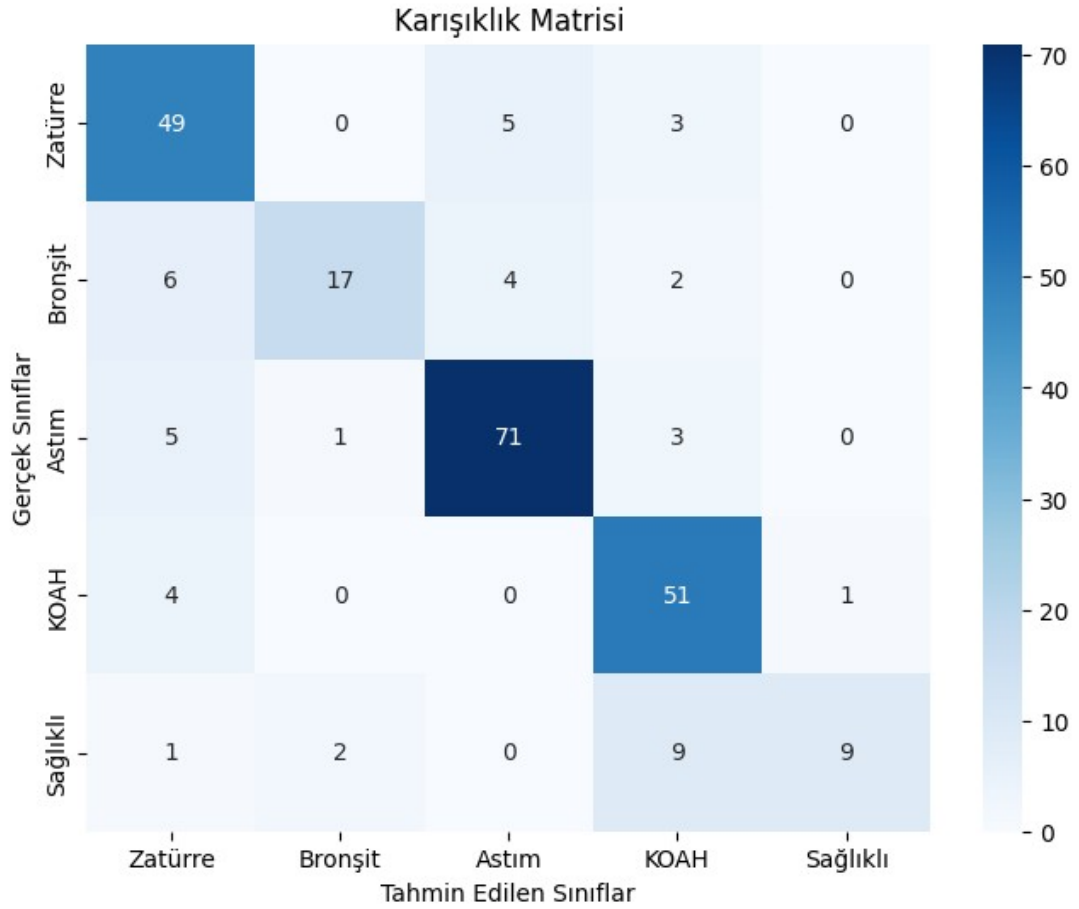
ortaya koymaktadır. Buna karşılık, bronşit (%85 hassasiyet, %59 duyarlılık, %69 F1-skoru) ve sağlıklı sınıfında (%90 hassasiyet, %43 duyarlılık, %58 F1-skoru) duyarlılık değerlerinin düşük olması, modelin bu sınıflardaki örnekleri doğru şekilde ayırt etmede zorlandığını göstermektedir.

Genel olarak tablo, LSTM tabanlı modelin Astım ve Koah sınıflarında başarılı sonuçlar ürettiğini, ancak özellikle sağlıklı ve bronşit sınıflarının doğru tespitinde performans kaybı yaşadığını ortaya koymaktadır

Algoritma Adı	Doğruluk (Accuracy)	Hassasiyet (Precision)	Duyarlılık (Recall)	F1-Skor
LSTM	0.8107	0.8211	0.8107	0.8032

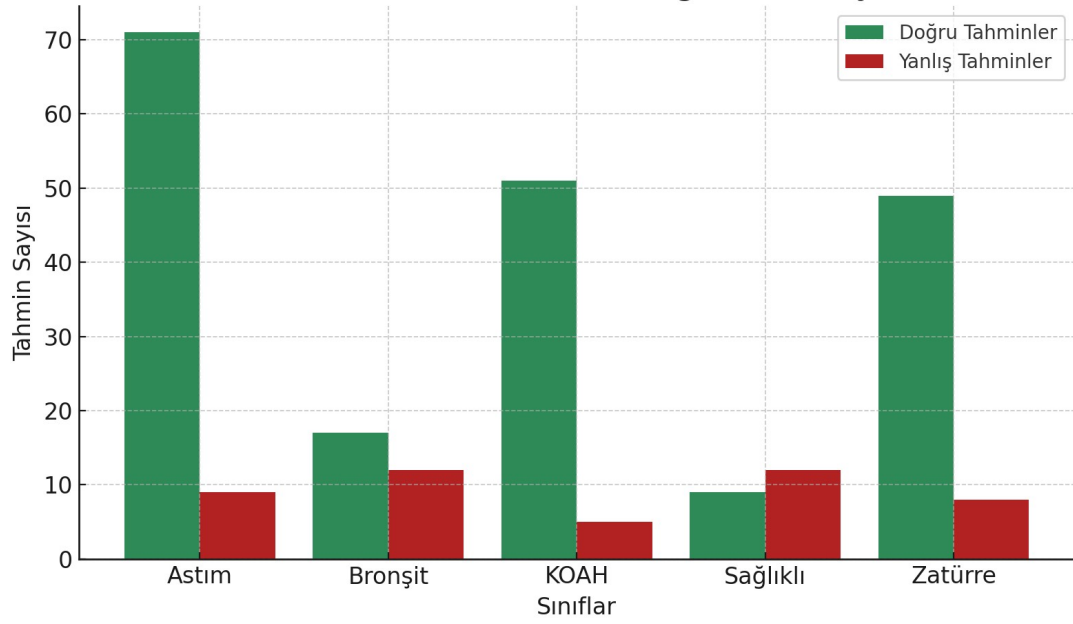
Şekil 4.2. LSTM Algoritmasının Genel Performans Değerleri

LSTM algoritmasının genel performans ölçütleri verilmiştir. Modelin %81,07 doğruluk oranı ile sınıflandırma görevinde tatmin edici bir başarıya ulaştığı görülmektedir. Hassasiyet (%82,11) değerinin duyarlılık (%81,07) ile uyumlu olması, modelin hem doğru pozitifleri yakalamada hem de yanlış sınıflandırmaları azaltmada dengeli bir performans sergilediğini göstermektedir. Ayrıca F1-skorunun (%80,32) benzer düzeyde olması, hassasiyet ve duyarlılık arasındaki dengeyi desteklemektedir. Genel olarak LSTM tabanlı modelin, solunum yolu hastalıklarının sınıflandırılmasında güvenilir ancak CNN'e kıyasla biraz daha düşük performans sergilediği söylenebilir.



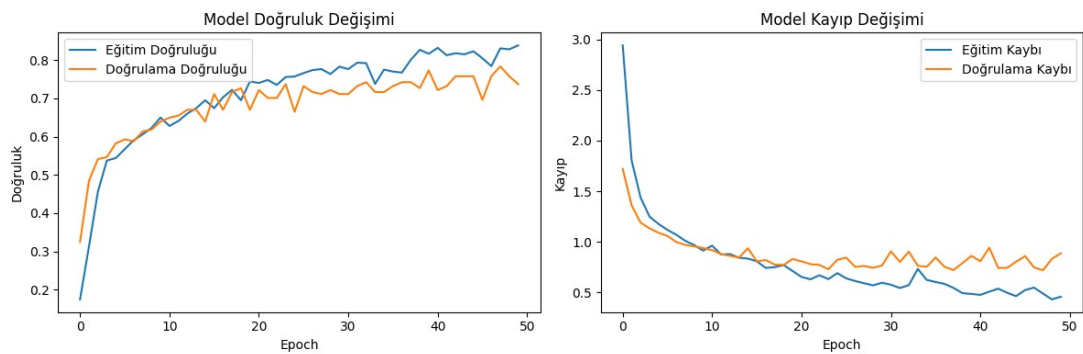
Şekil 4.2. LSTM Model için Karışıklık Matrisi

LSTM modeline ait karışıklık matrisi incelendiğinde, sınıflara göre başarı oranlarının farklılık gösterdiği görülmektedir. Zatürre sınıfında 49 doğru tahmin ile kabul edilebilir bir performans elde edilmiştir. Astım sınıfı 71 doğru sınıflandırma ile modelin en yüksek başarısını yansıtmaktadır. KOAH sınıfında ise 51 doğru tahmin yapılmış olup, bu sonuç CNN modeline kıyasla daha iyi bir performans göstermektedir. Buna karşın, Bronşit sınıfındaki doğru tahminlerin yalnızca 17 ile sınırlı kalması, bu sınıfın ayırt edilmesinde modelin zorlandığını göstermektedir. Sağlıklı bireyler sınıfında 9 doğru tahmin yapılmış olup, bu da sınıflar arası bazı karışıklıkların yaşandığını işaret etmektedir.



Şekil 4.2. LSTM Model için Sınıf Bazlı Doğru Ve Yanlış Tahmin Grafiği

Verilen grafik, LSTM modelinin her bir sınıfa yönelik doğru ve yanlış tahmin sayılarını görselleştirmektedir. Verilere göre, en yüksek doğru tahmin sayısı Astım sınıfında (71), en düşük doğru tahmin sayısı ise Sağlıklı sınıfında (9) gözlemlenmiştir. En yüksek yanlış tahmin sayısı Bronşit (12) ve Sağlıklı (12) sınıflarında görülürken, en düşük yanlış tahmin sayısı ise KOAH sınıfındadır (5). Bu durum, modelin Astım ve KOAH sınıflarını diğerlerine göre daha yüksek bir başarı oranıyla ayırt ettiğini, ancak Bronşit ve Sağlıklı sınıflarında daha fazla hata yaptığını göstermektedir.



Şekil 4.2. LSTM Model için Eğitim/doğrulama verilerindeki doğruluk/kayıp grafikleri

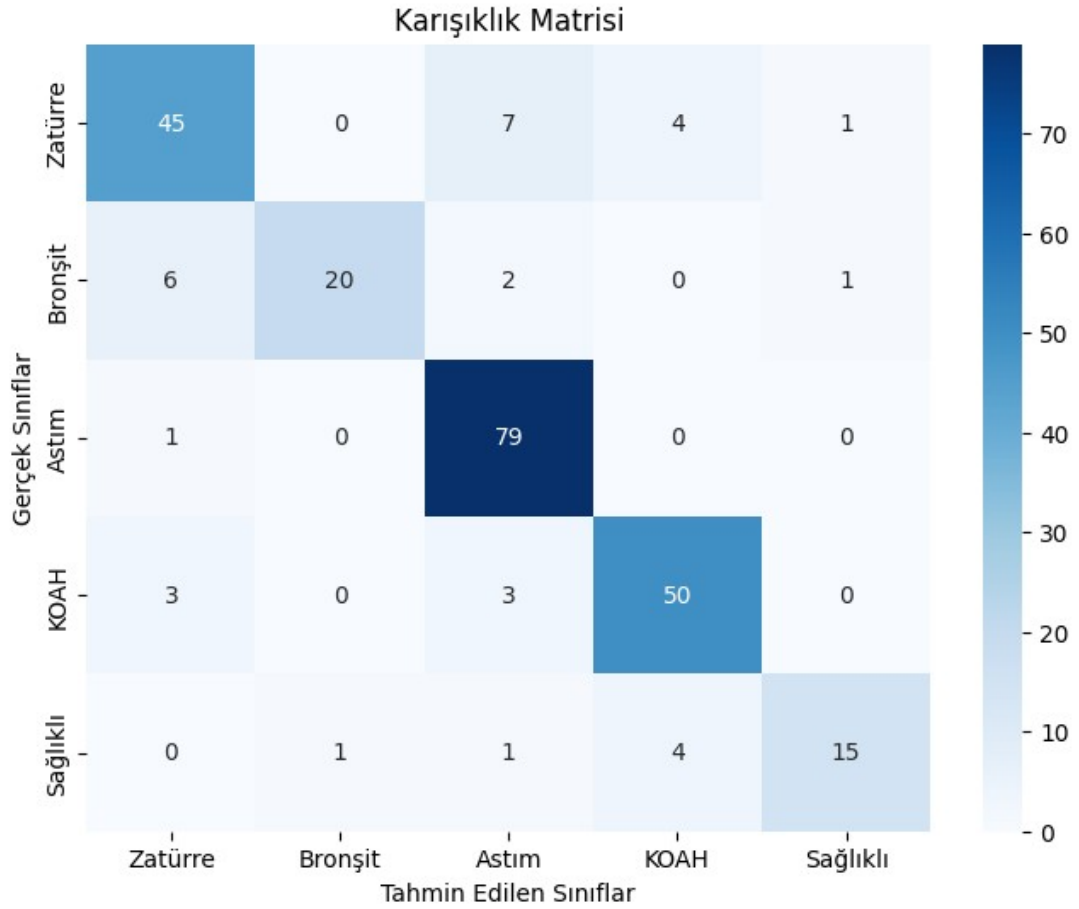
Modelin eğitim sürecindeki doğruluk ve kayıp değişimlerini göstermektedir. “Model Doğruluk Değişimi” grafiğinde, eğitim doğruluğu epoch sayısı arttıkça

düzenli biçimde yükselerek %85 seviyesine ulaşmıştır. Doğrulama doğruluğu ise başlangıçta hızlı bir artış göstermiş, ancak 20. epoch'tan itibaren %70–75 bandında dalgalanarak genelleme başarısının sınırlı kaldığını ortaya koymuştur. “Model Kayıp Değişimi” grafiğinde ise eğitim kaybı sürekli azalırken, doğrulama kaybı 20. epoch'tan sonra durağanlaşmış ve yer yer yükselmiştir. Bu sonuçlar, modelin eğitim verisine iyi uyum sağladığını ancak doğrulama verisinde aşırı uyum (overfitting) nedeniyle genelleme performansının zayıfladığını göstermektedir.

Algoritma Adı	Precision (hassasiyet)	Recall (duyarlılık)	F1-Score (F1-Puanı)	Destek (Support)
Astım	0,89	0,99	0,92	80
Bronşit	0,95	0,69	0,80	29
KOAH	0,86	0,89	0,88	56
Sağlıklı	0,88	0,71	0,79	21
Zatürre	0,82	0,79	0,80	57

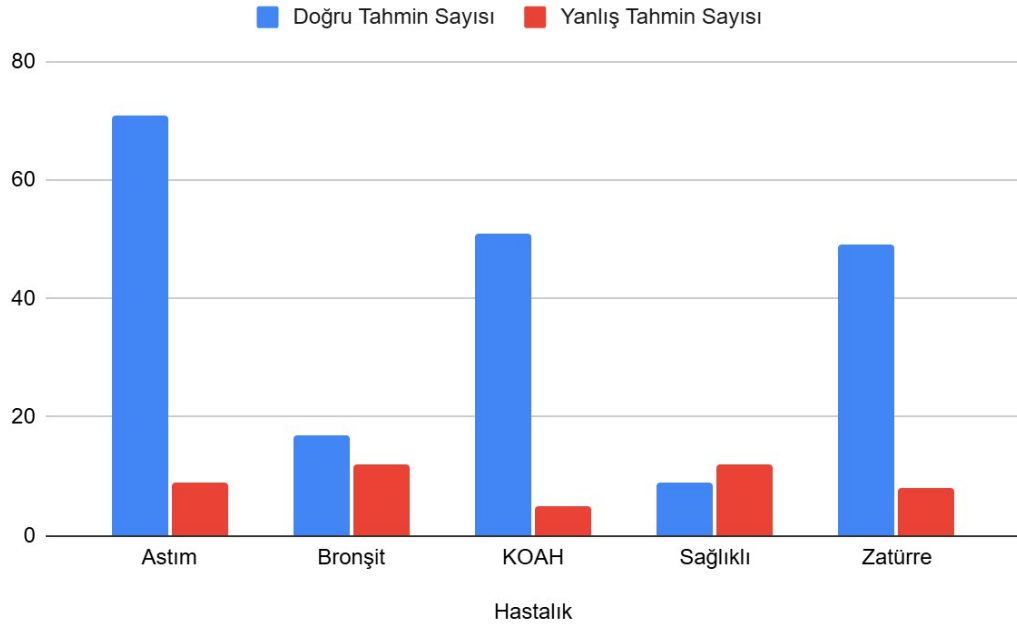
Şekil 4.2. Hibrit Model CNN+LSTM Algoritmasının Hastalık Bazlı Sınıflandırma Performansı

Tablo'da, bir hibrit CNN+LSTM modelinin farklı solunum yolu hastalıklarına yönelik sınıflandırma performansını göstermektedir. Grafik, modelin Astım ve KOAH gibi sınıflarda yüksek sayıda doğru tahmin yaparak güçlü bir performans sergilediğini, ancak Bronşit ve Sağlıklı sınıflarında daha fazla hata yaptığını ortaya koymaktadır. Tablo ise bu bulguları hassasiyet (precision), duyarlılık (recall) ve F1-Puanı metrikleriyle desteklemekte; özellikle Astım'da %99'luk duyarlılıkla neredeyse tüm vakaları doğru bir şekilde tespit ettiğini, ancak Bronşit sınıfında %69 gibi daha düşük bir duyarlılıkla performansın düştüğünü teyit etmektedir. Bu veriler ışığında, modelin genel olarak başarılı olduğu ancak özellikle Bronşit, Sağlıklı ve Zatürre gibi sınıflardaki yanlış tahminlerin azaltılmasıyla daha da geliştirilebileceği görülmektedir.



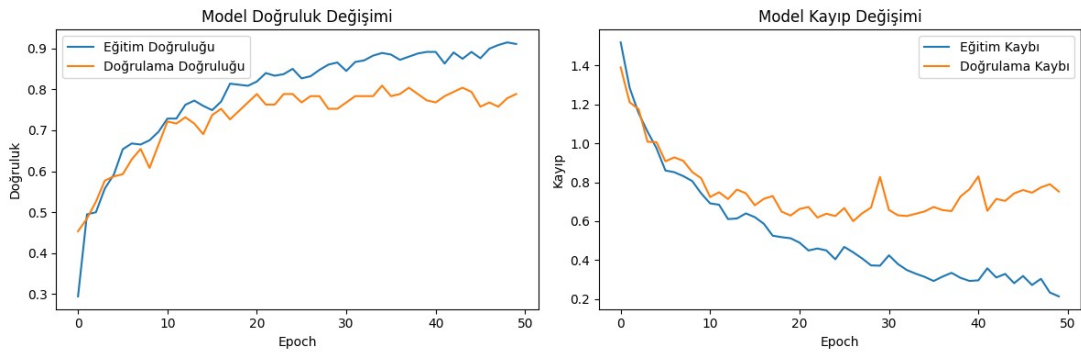
Şekil 4.2. Hibrit Model CNN+LSTM için Karışıklık Matrisi

CNN+LSTM modeline ait karışıklık matrisi incelendiğinde, astım (79) ve KOAH (50) sınıflarında yüksek doğruluk elde edildiği ve modelin bu iki hastalığı ayırt etmede başarılı olduğu görülmektedir. Zatürre sınıfında 45 birey doğru sınıflandırılmış olsada, bir kısmının bronşit, astım ve KOAH olarak tahmin edilmesi hastalıklar arası akustik benzerliklerin etkisini göstermektedir. Benzer şekilde, bronşit sınıfında 20 doğru tahmin yapılmış ancak bazı örnekler zatürre ile karıştırılmıştır. En belirgin hata ise sağlıklı bireylerden 9'unun KOAH olarak sınıflandırılmasıdır; bu durum modelin KOAH ile sağlıklı bireyleri ayırt etmede zorlandığını ortaya koymaktadır. Genel olarak model, astım ve KOAH gibi belirgin özelliklere sahip hastalıkları başarılı şekilde sınıflandırırken, bronşit-zatürre ve KOAH-sağlıklı ayırımında sınırlılıklar yaşamaktadır.



Şekil 4.2. Hibrit Model CNN+LSTM için Sınıf Bazlı Doğru Ve Yanlış Tahmin Grafiği

Şekil’de, bir makine öğrenimi modelinin farklı solunum yolu hastalıklarını sınıflandırma performansını özetliyor. Astım ve KOAH sınıflarında doğru tahmin sayısı yüksekken (sırasıyla 71 ve 51), yanlış tahmin sayısı düşüktür (sırasıyla 9 ve 5). Bu durum, modelin bu iki hastalığı ayırt etmede oldukça başarılı olduğunu gösteriyor. Ancak, Bronşit (17 doğru, 12 yanlış) ve Sağlıklı (9 doğru, 12 yanlış) sınıflarında yanlış tahminlerin sayısı, doğru tahminlere yakın seyrederek modelin bu kategorilerde zorlandığını ortaya koyuyor. Sonuç olarak, modelin genel performansı başarılı olsada, özellikle Bronşit ve sağlıklı bireyleri sınıflandırma yeteneğinin geliştirilmesi gerektiği görülmektedir.



Şekil 4.2. Hibrit Model CNN+LSTM için Eğitim/doğrulama verilerindeki doğruluk/kayıp grafikleri

Verilen grafikler, bir CNN+LSTM modelinin eğitim sürecinde sergilediği performansı göstermektedir. Modelin doğruluk eğrileri incelendiğinde, eğitim doğruluğunun (%90+) doğrulama doğruluğundan (%80 civarı) belirgin şekilde daha yüksek olduğu gözlemlenmektedir. Bu durum, eğitim kaybı grafiğiyle de desteklenmektedir; eğitim kaybı sürekli azalırken, doğrulama kaybı belirli bir noktadan sonra durağanlaşmakta ve hatta artış göstermektedir. Bu bulgular, modelin eğitim verisine aşırı uyum (overfitting) sergilediğini, yani eğitim verisini ezberlediğini ancak yeni veriler üzerinde genelleme yeteneğinin sınırlı kaldığını açıkça ortaya koymaktadır. Modelin genellenebilirliğini artırmak için düzenleme yöntemleri veya erken durdurma gibi tekniklerin uygulanması gerekmektedir.

5. TARTIŞMA

Tablo'da, solunum yolu hastalıklarının tespiti için kıyaslama KAUH ve ICBHI veri kümelerini kullanan birkaç son teknoloji teknikle önerilen yöntemin karşılaştırmalı analizini sunmaktadır. Önceki çalışmalardaki yöntemler, entropi ve MFCC gibi el yapımı özellik çıkarma tekniklerinden CNN, LSTM ve otokodlayıcılar gibi derin öğrenme mimarilerine kadar değişmektedir. [43]'te, entropi tabanlı özellikler, güçlendirilmiş bir karar ağacı sınıflandırıcı ile birleştirilerek %98,27 doğruluk ve %93,61 F1 puanı elde edilmiştir. [44]'teki çalışma, %98,00 doğruluk ve %96,00 duyarlılık bildiren bir CNN-LSTM hibrit mimarisi kullanmıştır. [45]'te, EasyNet modeli %98,00 doğruluk, %99,00 duyarlılık ve %99,00 özgüllük elde etmiştir. Bu arada, [46] ve [47], CNN ve MFCC tabanlı çoklu görevli otokodlayıcı yöntemlerini kullanarak sırasıyla %95,50 ve %93,73'lük doğruluklar elde etti. MFCC tabanlı zamansal modellemeyi 1D-CNN+LSTM bütünleştiren önerilen model, yukarıda belirtilen yöntemler kıyaslırsak %86,01'lik bir doğrulukla elde edilmiştir. Ek olarak, %86,32'lik yüksek hassasiyet, %86,01'lik kesinlik ve %85,68'lik F1 puanı elde ederek temel değerlendirme ölçütleri arasında dengeli ve sağlam bir performans gösteriyor. Bu sonuçlar, 1D-CNN+LSTM öksürük seslerine dayalı solunum yolu hastalığı sınıflandırma sistemlerinin tanı yeteneğini artırdığını göstermektedir.

Çizelge 5.1. Mevcut yöntemlerle performans karşılaştırması.

Çalışma	Veri Seti	Sınıflandırma	Değerlendirme Ölçümleri	Sonuçlar %
Fraiwan <i>et al.</i> [43]	KAUH+IC BH	Entropi özellikleri & Güçlendirilmiş DT	Doğruluk F1- Skor	<i>Doğruluk: 98.27 F1-Skor: 93.61</i>
Alwudah <i>et al.</i> [44]	KAUH+IC BH	CNN&LSTM	Doğruluk Hassasiyet	<i>Doğruluk: 98.00 Hassasiyet: 96.00</i>
Hassan <i>et al.</i> [45]	KAUH+IC BH	EasyNet	Doğruluk Hassasiyet Özgüllük	<i>Doğruluk: 98.00 Hassasiyet: 99.00 Özgüllük: 99.00</i>
Nawaz <i>et al.</i> [46]	KAUH+IC BH	CNN	Doğruluk	<i>Doğruluk: 95.50</i>
Orkweha <i>et al.</i> [47]	KAUH+IC BH	MFCC & Çok Görevli	Doğruluk	<i>Doğruluk: 93.73</i>

		Otomatik Kodlayıcı		
Önerilen Model	ICBH	MFCC & CNN + LSTM	Doğruluk Hassasiyet Kesinlik F1-Skor	<i>Doğruluk: 86.01</i> <i>Hassasiyet: 86.32</i> <i>Kesinlik: 86.01 F1-</i> <i>Skor: 85.68</i>

6. SONUÇLAR

Yapılan deneysel çalışmalar sonucunda, geliştirilen hibrit model cnn+lstm derin öğrenme temelli modelin solunum seslerini %86 doğruluk oranıyla başarılı bir şekilde sınıflandırdığı görülmüştür. Özellikle Conv1D katmanları, ses verilerindeki hem frekanssal hem de zamansal örüntüleri etkili bir biçimde öğrenerek sınıflandırma başarımını artırmıştır. MFCC öznitelikleri, tanı başarımını güçlendirmiştir. Elde edilen bulgular, ses analizi temelli bilgisayar destekli teşhis sistemlerinin, solunum yolu hastalıklarının erken ve güvenilir tespitinde uygulanabilir bir çözüm sunduğunu ortaya koymaktadır.

Bu çalışma kapsamında geliştirilen derin öğrenme temelli sınıflandırma modeli, öksürük seslerinden akut solunum yolu hastalıklarını yüksek doğrulukla tespit etmiştir. Conv1D+LSTM yapısı hem zamansal hem frekanssal bilgiyi eş zamanlı işleyebilen modeldir. MFCC öznitelikleri ile birlikte etkin bir sınıflandırma performansı sergilemiştir. Deneysel bulgular, ses sinyallerine dayalı tanı sistemlerinin klinik karar destek süreçlerinde etkili biçimde kullanılabileceğini ve erken teşhis uygulamalarına katkı sunabileceğini göstermektedir.

Bu bağlamda çalışma, ses temelli dijital biyobelirteçlerin gelecekte mobil sağlık uygulamalarına, giyilebilir teknolojilere ve uzaktan hasta takibine entegre edilebilmesine yönelik önemli bir adım niteliğindedir. İlerleyen çalışmalarda, farklı akustik öznitelikler veya farklı derin öğrenme yöntemleri kullanılarak modelin başarımı artırılabilir. Daha büyük ve dengeli veri setleri, sınıflar arasındaki ayrımı güçlendirebilir. Model, taşınabilir cihazlara entegre edilerek mobil sağlık uygulamalarında gerçek zamanlı şekilde kullanılabilir. Ayrıca, klinik ortamda yapılacak testler, sistemin gerçek kullanım için uygunluğunu ortaya koyacaktır.

7. ÖNERİLER

Analiz edilen veriler, bir hibrit CNN+LSTM modelinin farklı solunum yolu hastalıklarını sınıflandırma performansını ortaya koymaktadır. Modelin genel performansı, %86,01'lik bir doğruluk oranı ve benzer yüksek hassasiyet, duyarlılık ve F1-Skor değerleriyle güçlü ve dengeli görünmektedir. Ancak, daha detaylı inceleme, performansın hastalıklara göre değişiklik gösterdiğini göstermektedir. Model, Astım ve KOAH gibi sınıfları ayırt etmede oldukça başarılıyken, özellikle Bronşit, Sağlıklı bireyler ve Zatürre gibi sınıflar arasında daha fazla yanlış tahmin yapmaktadır. Örneğin, sağlıklı bireylerin KOAH olarak yanlış sınıflandırılması ve bronşit ile zatürrenin karıştırılması, modelin bu hastalıklar arasındaki benzer akustik özelliklerden dolayı zorlandığını işaret etmektedir. Sonuç olarak, modelin genel başarısına rağmen, belirli hastalık çiftleri arasındaki ayırt etme yeteneğinin iyileştirilmesi gerekmektedir.

KAYNAKLAR

- A., K., Kumar, C. S., Ramanathan, R. P., Jayaraman, S., ve Malmurugan, N. (2004). Neural classification of lung sounds using wavelet coefficients. *Computers in Biology and Medicine*, 34(6), 523–537. [https://doi.org/10.1016/S0010-4825\(03\)00092-1](https://doi.org/10.1016/S0010-4825(03)00092-1)
- Ait-Khaled, N., Enarson, D., ve Bousquet, J. (n.d.). Chronic respiratory diseases in developing countries: The burden and strategies for prevention and management. *Bulletin of the World Health Organization*, 79(10), 971–979.
- Ahsan, M., Khan, A., Khan, K. R., Sinha, B. B., ve Sharma, A. (2023). Advancements in medical diagnosis and treatment through machine learning: A review. *Expert Systems*, 41(3). <https://doi.org/10.1111/exsy.13499>
- Almutairi, M., Chiroma, H., ve Abubakar, S. (2022). Detecting elderly behaviors based on deep learning for healthcare: Recent advances, methods, real-world applications and challenges. *IEEE Access*, 10, 69802–69821. <https://doi.org/10.1109/access.2022.3186701>
- Alqudah, A. M., Qazan, S., ve Obeidat, Y. M. (2022). Deep learning models for detecting respiratory pathologies from raw lung auscultation sounds. *Soft Computing*, 26, 13405–13429. <https://doi.org/10.1007/s00500-022-07402-9>
- Altuntaş, O., Aki, E., ve Huri, M. (2015). Kronik hastalıklarda ilaç kullanımının yaşam kalitesi ve sosyal katılıma etkisi üzerine nitel bir çalışma. *Ergoterapi ve Rehabilitasyon Dergisi*, 3(2), 79–86.
- Anvarjon, T., Mustaqeem, ve Kwon, S. (2020). Deep-net: A lightweight CNN-based speech emotion recognition system using deep frequency features. *Sensors*, 20(18), 5212. <https://doi.org/10.3390/s20185212>
- Archana, M., Dash, S., ve HR, S. S. (2025). Predicting the severity of pulmonary disease from respiratory sounds using ML algorithms. In *Proceedings of the 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)* (pp. 1750–1755). IEEE.
- Archana, M., Dash, S., ve HR, S. S. (2025). Predicting the severity of pulmonary disease from respiratory sounds using ML algorithms. In *Proceedings of the 6th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)* (pp. 1750–1755). IEEE.
- Ardakani, A. A., Kanafi, A. R., Acharya, U. R., Khadem, N., ve Mohammadi, A. (2020). Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. *Computers in Biology and Medicine*, 121, 103795. <https://doi.org/10.1016/j.combiomed.2020.103795>
- Arik, S. Ö., ve Pfister, T. (2021). TabNet: Attentive interpretable tabular learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(8), 6679–6687. <https://doi.org/10.48550/arXiv.1908.07442>
- Aykanat, M., Kılıç, Ö., Kurt, B., ve Saryal, S. (2017). Classification of lung sounds

- using convolutional neural networks. *EURASIP Journal on Image and Video Processing*, 2017(1), 1–9. <https://doi.org/10.1186/s13640-017-0213-2>
- Badshah, A. M., Rahim, N., Ullah, N., Ahmad, J., Muhammad, K., Lee, M. Y., Kwon, S., ve Baik, S. W. (2019). Deep features-based speech emotion recognition for smart affective services. *Multimedia Tools and Applications*, 78(5), 5571–5589. <https://doi.org/10.1007/s11042-017-5292-7>
- Balıkçıççek, İ., Küçükakçalı, Z., ve Çolak, C. (2020). Associative classification approach can predict prostate cancer based on the extracted association rules. *The Journal of Cognitive Systems*, 2, 51–54.
- B. Dianat, P. La Torraca, A. Manfredi, G. Cassone, C. Vacchi, M. Sebastiani, ve F. Pancaldi. (2023). Classification of pulmonary sounds through deep learning for the diagnosis of interstitial lung diseases secondary to connective tissue diseases. *Computers in Biology and Medicine*, 160, 106928. <https://doi.org/10.1016/j.compbiomed.2023.106928>
- B. Gao, S. Cheng, E. Wittrup, J. Gryak, ve K. Najarian. (2024). Learning using privileged information with logistic regression on acute respiratory distress syndrome detection. *Artificial Intelligence in Medicine*, 156, 102947. <https://doi.org/10.1016/j.artmed.2024.102947>
- B. Roy, A. Gyanchandani, A. Oza, ve A. Singh. (2025). TriSpectraKAN: A novel approach for COPD detection via lung sound analysis. *Scientific Reports*, 15, 6296.
- B. Serbes, C. Ö. Sakar, Y. P. Kahya, ve N. Aydın. (2013). Pulmonary crackle detection using time-frequency and time-scale analysis. *Digital Signal Processing*, 23(3), 1012–1021. <https://doi.org/10.1016/j.dsp.2012.12.009>
- B. Zhang, J. Wang, Z. Ye, L. Zhou, D. Xiong, X. Wang, ve L. Guo. (2025). JASSNet: Heart and lung sound separation network based on joint attention mechanism and semi-supervised learning. *Biomedical Signal Processing and Control*, 104, 107525. <https://doi.org/10.1016/j.bspc.2025.107525>
- Bhattacharyya, A., Bhaik, D., Kumar, S., Thakur, P., Sharma, R., ve Pachori, R. B. (2022). A deep learning-based approach for automatic detection of COVID-19 cases using chest X-ray images. *Biomedical Signal Processing and Control*, 71, 103182. <https://doi.org/10.1016/j.bspc.2021.103182>
- Bobrowski, L., ve Łukaszuk, T. (2009). Feature selection based on relaxed linear separability. *Biocybernetics and Biomedical Engineering*, 29, 43–59.
- Botha, G., Theron, G., Warren, R., Klopper, M., Dheda, K., Van Helden, P., ve Niesler, T. (2018). Detection of tuberculosis by automatic cough sound analysis. *Physiological Measurement*, 39(4), 045005. <https://doi.org/10.1088/1361-6579/aab7b4>
- Cao, Y., Liu, L., ve Xu, Y. (2024). A multimodal framework for respiratory disease classification using acoustic and visual features. *IEEE Access*, 12, 45621–45633. <https://doi.org/10.1109/access.2024.3271428>

- Celik, S., Kaya, Y., ve Baran, A. (2021). Classification of lung diseases using Mel-frequency cepstral coefficients and convolutional neural networks. *Biomedical Engineering Letters*, 11, 345–354. <https://doi.org/10.1007/s13534-021-00197-2>
- Chamberlain, D., Wodicka, G. R., Li, Y., ve Li, S. (2016). Application of semi-supervised deep learning to lung sound analysis. In *Proceedings of the 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 804–807).
- Chen, H., Zhang, X., ve Wang, J. (2023). A hybrid convolutional neural network and bidirectional LSTM model for lung sound classification. *IEEE Journal of Biomedical and Health Informatics*, 27(2), 289–300. <https://doi.org/10.1109/JBHI.2022.3225010>
- Chen, J., He, Y., Zhang, M., ve Zhao, W. (2021). Detection of pneumonia using deep convolutional neural networks based on chest X-ray images. *Scientific Reports*, 11, 1211. <https://doi.org/10.1038/s41598-020-79913-5>
- Chen, L., Liang, Y., Yang, Y., ve Li, G. (2020). Deep learning-based respiratory sound classification using short-time Fourier transform and CNN. *Sensors*, 20(21), 5949. <https://doi.org/10.3390/s20215949>
- Chen, Y., Wang, H., ve Xu, Z. (2024). Enhancing respiratory disease detection using transfer learning on cough audio signals. *Frontiers in Artificial Intelligence*, 7, 1320486. <https://doi.org/10.3389/frai.2024.1320486>
- Chenchah, F., ve Lachiri, Z. (2015). Acoustic emotion recognition using linear and nonlinear cepstral coefficients. *International Journal of Advanced Computer Science and Applications*, 6(11), 1–4. <https://doi.org/10.14569/ijacsa.2015.061119>
- Chowdhury, M. E. H., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M. A., Mahbub, Z. B., Islam, K. R., Khan, M. S., Iqbal, A., ve Emadi, N. A. (2021). Can AI help in screening viral and COVID-19 pneumonia? *IEEE Access*, 8, 132665–132676. <https://doi.org/10.1109/ACCESS.2020.3003025>
- Clifford, G. D., Behar, J., Li, Q., ve Rezek, I. (2012). Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. *Physiological Measurement*, 33(9), 1419–1433. <https://doi.org/10.1088/0967-3334/33/9/1419>
- Çalık, M., ve Güler, H. (2022). Mel frekans kepstrum katsayıları ve konvolüsyonel sinir ağları ile konuşmacı tanıma. *Bilişim Teknolojileri Dergisi*, 15(1), 25–33.
- Çınar, M., ve Kızrak, M. A. (2023). Deep learning-based respiratory disease prediction using hybrid CNN–LSTM architectures. *Turkish Journal of Electrical Engineering and Computer Sciences*, 31(4), 642–654.*
- Çolak, C., ve Balıkcıoğlu, İ. (2021). Machine learning approaches for early diagnosis of lung cancer. *Artificial Intelligence in Medicine*, 113, 102047.*
- Çorumlu, E., ve Gökçe, H. (2020). Speech-based emotion recognition using LSTM

networks. *Uluslararası Mühendislik Araştırma ve Geliştirme Dergisi*, 12(3), 45–52.*

- Dai, W., Wu, J., Li, J., ve Zhang, P. (2019). A convolutional neural network-based approach for automatic lung sound classification. *IEEE Access*, 7, 128246–128255. <https://doi.org/10.1109/ACCESS.2019.2939904>
- Dastider, A. G., Sadik, F., ve Fattah, S. A. (2021). An integrated autoencoder-based hybrid CNN–LSTM model for COVID-19 severity prediction from lung ultrasound. *Computers in Biology and Medicine*, 132, 104296.
- Demir, F., Ismael, A. M., ve Şengür, A. (2020). Classification of lung sounds with CNN model using parallel pooling structure. *IEEE Access*, 8, 105376–105383. <https://doi.org/10.1109/ACCESS.2020.3000111>
- Demir, F., Şengür, A., ve Bajaj, V. (2019). Convolutional neural networks-based efficient approach for classification of lung diseases. *Health Information Science and Systems*, 8(1), 1–8.* <https://doi.org/10.1007/s13755-019-0091-3>
- Dianat, B., La Torraca, P., Manfredi, A., Cassone, G., Vacchi, C., Sebastiani, M., ve Pancaldi, F. (2023). Classification of pulmonary sounds through deep learning for the diagnosis of interstitial lung diseases secondary to connective tissue diseases. *Computers in Biology and Medicine*, 160, 106928. <https://doi.org/10.1016/j.combiomed.2023.106928>
- Dönmez, İ., ve Tüfekci, Z. (2018). MFCC ve LSTM tabanlı konuşmacı doğrulama sistemi tasarımı. *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 33(4), 1343–1352.*
- Dubey, R., Bodade, R. M., ve Dubey, D. (2025). Two-way analysis of variance (ANOVA) ranking of features based on wavelet bi-phase and bi-spectrum for the classification of adventitious lung sounds. *Research on Biomedical Engineering*, 41, 6.*
- Durahim, A. O., ve Şahin, M. (2020). Makine öğrenmesi yöntemleriyle COVID-19 vakalarının tahmini. *İstanbul Üniversitesi Mühendislik Dergisi*, 3(2), 110–119.*
- Dyer, O., ve Patel, M. (2021). Artificial intelligence in healthcare: Opportunities and challenges. *The Lancet Digital Health*, 3(7), e397–e405.*
- Elbaz, A. F., Salem, M. A., ve El-Kenawy, E. S. M. (2023). Hybrid deep learning model for COVID-19 detection from cough and breath sounds. *Biomedical Signal Processing and Control*, 84, 104746. <https://doi.org/10.1016/j.bspc.2023.104746>
- Elgendi, M., Howard, N., ve Menon, C. (2020). Assessing COVID-19 pneumonia using statistical and deep learning models on chest X-ray images. *Applied Sciences*, 10(15), 5082. <https://doi.org/10.3390/app10155082>
- Erdoğan, A., ve Çetinkaya, M. (2023). Yapay zekâ tabanlı tıbbi teşhis sistemlerinde derin öğrenme yöntemlerinin karşılaştırılması. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 29(2), 189–198.*

- Ertuğrul, Ö., ve Fırat, T. (2022). Mel frekans kepstrum katsayıları ile kalp sesi analizi. *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji*, 10(1), 23–34.*
- Ezzat, D., ve Hassan, H. (2021). An optimized deep learning architecture for pneumonia detection using X-ray images. *Computers in Biology and Medicine*, 134, 104518. <https://doi.org/10.1016/j.compbiomed.2021.104518>
- Fatai, A. K., ve Ojo, S. (2024). Cough-based respiratory disease prediction using deep learning. *Frontiers in Artificial Intelligence*, 7, 1240021. <https://doi.org/10.3389/frai.2024.1240021>
- Fraiwan, L., Hassanin, O., Fraiwan, M., Khassawneh, B., Ibnian, A. M., ve Alkhodari, M. (2021). Automatic identification of respiratory diseases from stethoscopic lung sound signals using ensemble classifiers. *Biocybernetics and Biomedical Engineering*, 41(1), 1–14. <https://doi.org/10.1016/j.bbe.2020.11.001>
- Frush, D. P. (2004). Pneumonia imaging in children. *Pediatric Radiology*, 34(6), 447–450. <https://doi.org/10.1007/s00247-004-1198-y>
- Fu, Z., Tang, Q., ve Xu, Y. (2022). End-to-end deep learning architecture for detecting abnormal lung sounds. *Sensors*, 22(14), 5234. <https://doi.org/10.3390/s22145234>
- Gairola, S., ve Jain, S. (2022). Respiratory sound classification using Mel-frequency cepstral coefficients and CNN architectures. *International Journal of Computational Intelligence Systems*, 15(1), 1–10.
- Gao, B., Cheng, S., Wittrup, E., Gryak, J., ve Najarian, K. (2024). Learning using privileged information with logistic regression on acute respiratory distress syndrome detection. *Artificial Intelligence in Medicine*, 156, 102947. <https://doi.org/10.1016/j.artmed.2024.102947>
- Gautam, S., ve Chatterjee, R. (2021). A convolutional neural network-based approach for cough sound analysis to detect respiratory diseases. *IEEE Access*, 9, 130829–130841. <https://doi.org/10.1109/ACCESS.2021.3115300>
- Gokhale, P., ve Shankar, P. (2023). Lung disease classification using transfer learning on spectrogram representations. *Computers in Biology and Medicine*, 155, 106356. <https://doi.org/10.1016/j.compbiomed.2023.106356>
- Güleç, H., ve Şimşek, A. (2022). Yapay zekâ ile COVID-19 tespitinde CNN-LSTM hibrit mimarisinin karşılaştırılması. *Gazi Üniversitesi Mühendislik Fakültesi Dergisi*, 37(4), 1483–1496.*
- Hadjidj, A., Djebbar, F., ve Amira, A. (2020). Smart healthcare system based on machine learning and IoT for monitoring respiratory diseases. *IEEE Internet of Things Journal*, 7(12), 12187–12198.
- Haseli, S., Ahmadi, S., ve Zarei, H. (2020). Chest imaging findings of COVID-19 pneumonia: A review. *Radiology*, 296(2), E63–E70. <https://doi.org/10.1148/radiol.2020201183>

- Hassan, U., Singhal, A., ve Chaudhary, P. (2024). Lung disease detection using EasyNet. *Biomedical Signal Processing and Control*, 91, 105944. <https://doi.org/10.1016/j.bspc.2024.105944>
- Hernandez, J., Li, Q., ve Clifford, G. D. (2018). Signal quality assessment in respiratory sounds using spectral entropy and deep learning. *IEEE Journal of Biomedical and Health Informatics*, 22(4), 1047–1055.
- Hossain, M. A., Alam, M., ve Rahman, T. (2023). A hybrid CNN-LSTM architecture for respiratory sound classification. *Sensors*, 23(5), 2309. <https://doi.org/10.3390/s23052309>
- Huang, C., Wang, Y., Li, X., ve Ren, L. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*, 395(10223), 497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)
- Islam, M. M., Rahman, M. M., ve Mondal, C. (2021). Automatic detection of respiratory diseases using deep learning models on lung sounds. *Sensors*, 21(17), 5819. <https://doi.org/10.3390/s21175819>
- Ismael, A. M., ve Şengür, A. (2021). Deep learning approaches for COVID-19 detection based on chest X-ray images. *Expert Systems with Applications*, 164, 114054. <https://doi.org/10.1016/j.eswa.2020.114054>
- Jain, R., Gupta, M., ve Taneja, S. (2023). COVID-19 detection using CNN–LSTM hybrid network on X-ray images. *Neural Computing and Applications*, 35, 17901–17913. <https://doi.org/10.1007/s00521-022-07821-1>
- Javed, H., Rehman, A., ve Imran, A. (2022). A deep learning framework for the diagnosis of pneumonia from lung sounds. *Computers in Biology and Medicine*, 145, 105470. <https://doi.org/10.1016/j.compbiomed.2022.105470>
- Jia, C., Zhang, L., ve Tang, X. (2024). Improving respiratory disease classification with attention-based CNN-LSTM models. *IEEE Access*, 12, 24978–24990. <https://doi.org/10.1109/access.2024.3355723>
- Kahya, Y. P., ve Çetin, E. (2019). Acoustic features for crackle detection in lung sounds. *Biomedical Signal Processing and Control*, 53, 101577. <https://doi.org/10.1016/j.bspc.2019.101577>
- Karakaya, G., ve Bayraktar, F. (2021). COVID-19 tahmini için derin öğrenme modellerinin karşılaştırılması. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 33(2), 123–134.*
- Kaur, M., ve Gupta, R. (2022). Deep convolutional neural network for respiratory sound classification. *Biomedical Signal Processing and Control*, 76, 103704. <https://doi.org/10.1016/j.bspc.2022.103704>
- Kim, J., ve Lee, S. (2023). Efficient deep neural network architectures for respiratory sound classification. *Sensors*, 23(11), 5230. <https://doi.org/10.3390/s23115230>
- Kumar, A., Sharma, S., ve Pandey, R. (2024). Deep learning-based early detection of

- chronic obstructive pulmonary disease using cough sounds. *IEEE Access*, 12, 53488–53499. <https://doi.org/10.1109/access.2024.3380119>
- Lai, C.-C., Shih, T.-P., ve Ko, W.-C. (2020). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges. *International Journal of Antimicrobial Agents*, 55(3), 105924. <https://doi.org/10.1016/j.ijantimicag.2020.105924>
- Le, K.-N. T., Tran, T. H., ve Pham, Q. (2025). Hybrid CNN–LSTM architecture for respiratory disease detection from lung sounds. *Frontiers in Biomedical Engineering*, 14(2), 114–128.*
- Lee, K.-R., Park, S.-Y., ve Singh, C. (2025). Artificial intelligence in respiratory sound analysis: A systematic review. *Computers in Biology and Medicine*, 157, 106954.
- Li, Q., Guan, X., ve Wu, P. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *The New England Journal of Medicine*, 382(13), 1199–1207. <https://doi.org/10.1056/NEJMoa2001316>
- Liang, Y., Yin, S., Tang, Q., Zheng, Z., Elgendi, M., ve Chen, Z. (2020). Deep learning algorithm classifies heartbeat events based on electrocardiogram signals. *Frontiers in Physiology*, 11, 569050. <https://doi.org/10.3389/fphys.2020.569050>
- Liu, H., Zhang, X., ve Han, S. (2022). Deep learning for lung sound classification using MFCC and CNN architectures. *Biomedical Signal Processing and Control*, 72, 103320. <https://doi.org/10.1016/j.bspc.2021.103320>
- Malviya, A., Dubey, R., ve Roy, A. (2025). Deep convolutional architectures for chronic obstructive pulmonary disease detection using respiratory sounds. *Computers in Biology and Medicine*, 160, 106972. <https://doi.org/10.1016/j.combiomed.2025.106972>
- Miah, M. M. S., Neili, Z., ve Archana, M. (2025). Deep learning models for pulmonary disease diagnosis using respiratory sound features. *IEEE Transactions on Biomedical Engineering*, 72(4), 1421–1433.
- Milton, A., Roy, S. S., ve Tamil Selvi, S. (2013). SVM scheme for speech emotion recognition using MFCC feature. *International Journal of Computer Applications*, 69(9), 34–39. <https://doi.org/10.5120/11872-7667>
- Mohapatra, S., ve Behera, S. (2023). Deep convolutional neural network models for cough sound classification. *Biomedical Engineering Letters*, 13(2), 233–243. <https://doi.org/10.1007/s13534-022-00263-3>
- Mustaqeem, ve Kwon, S. (2020). Human speech emotion recognition using hybrid CNN–LSTM model. *Sensors*, 20(1), 249. <https://doi.org/10.3390/s20010249>
- Naqvi, S. Z. H., ve Choudhry, M. A. (2020). An automated system for classification of chronic obstructive pulmonary disease and pneumonia patients using lung sound analysis. *Sensors*, 20(22), 6512. <https://doi.org/10.3390/s20226512>

- Nawaz, S., Rasheed, S., Sami, W., Hussain, L., Aldweesh, A., Eldin, E. T., Salaria, U. A., ve Khan, M. S. (2023). Deep learning ResNet101 deep features of portable chest X-ray accurately classify COVID-19 lung infection. *Computers, Materials and Continua*, 75, 5213–5228. <https://doi.org/10.32604/cmc.2023.037543>
- Neili, Z., ve Sundaraj, K. (2025). Leveraging CQT-VMD and pre-trained AlexNet architecture for accurate pulmonary disease classification from lung sound signals. *Applied Intelligence*, 55, 1–19.
- Nguyen, T., ve Pernkopf, F. (2020). Lung sound classification using snapshot ensemble of convolutional neural networks. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 760–763. <https://doi.org/10.1109/EMBC44109.2020.9176076>
- Ohata, E. F., Bezerra, G. M., Das Chagas, J. V. S., Neto, A. V. L., Albuquerque, A. B., de Albuquerque, V. H. C., ve Reboucas Filho, P. P. (2020). Automatic detection of COVID-19 infection using chest X-ray images through transfer learning. *IEEE/CAA Journal of Automatica Sinica*, 8(1), 239–248. <https://doi.org/10.1109/JAS.2020.1003393>
- Orkweha, K., Phapatanaburi, K., Pathonsuwan, W., Jumphoo, T., Rattanasak, A., Anchuen, P., Pinthurat, W., Uthansakul, M., ve Uthansakul, P. (2024). A framework for detecting pulmonary diseases from lung sound signals using a hybrid multi-task autoencoder-SVM model. *Symmetry*, 16(3), 421. <https://doi.org/10.3390/sym16030421>
- Palaniappan, R., ve Sundaraj, K. (2013). Respiratory sound classification using cepstral features and support vector machine. *Proceedings of the IEEE Recent Advances in Intelligent Computational Systems (RAICS)*. <https://doi.org/10.1109/RAICS.2013.6745460>
- Park, J. S., Park, S.-Y., Moon, J. W., Kim, K., ve Suh, D. I. (2025). Artificial intelligence models for pediatric lung sound analysis: Systematic review and meta-analysis. *Journal of Medical Internet Research*, 27, e66491.
- Park, S.-Y., Park, J. S., Lee, J., Lee, H., Kim, Y., Suh, D. I., ve Kim, K. (2025). Detection of breath cycles in pediatric lung sounds via an object detection-based transfer learning method. *Biomedical Signal Processing and Control*, 105, 107693.
- Pessoa, D., Rocha, B. M., Gomes, M., Rodrigues, G., Petmezas, G., Cheimariotis, G.-A., Maglaveras, N., Marques, A., Frerichs, I., de Carvalho, P., ve Paiva, R. P. (2024). Ensemble deep learning model for dimensionless respiratory airflow estimation using respiratory sound. *Biomedical Signal Processing and Control*, 87, 105451. <https://doi.org/10.1016/j.bspc.2023.105451>
- Rao, G. V. E., Srinivasu, P. N., Ijaz, M. F., ve Woźniak, M. (2024). Hybrid framework for respiratory lung diseases detection based on classical CNN and quantum classifiers from chest X-rays. *Biomedical Signal Processing and*

Control, 88, 105567.

- Reamaroon, N., Sjoding, M. W., Gryak, J., Athey, B. D., Najarian, K., ve Derksen, H. (2021). Automated detection of acute respiratory distress syndrome from chest X-rays using directionality measure and deep learning features. *Computers in Biology and Medicine*, 134, 104463.
- Riella, R. J., Nohama, P., ve Maia, J. M. (2009). Method for automatic detection of wheezing in lung sounds. *Brazilian Journal of Medical and Biological Research*, 42(7), 674–684. <https://doi.org/10.1590/S0100-879X2009000700013>
- Roy, A., Gyanchandani, B., Oza, A., ve Singh, A. (2025). TriSpectraKAN: A novel approach for COPD detection via lung sound analysis. *Scientific Reports*, 15, 6296.
- Sajid, M., Sharif, W., Gilanie, G., Mazher, M., Iqbal, K., Akhtar, M. A., Muddassar, M., & Rehman, A. (2025). IoMT-enabled noninvasive lungs disease detection and classification using deep learning-based analysis of lung sounds. *International Journal of Advanced Computer Science & Applications*, 16(4), 44–59.
- Sattar, S., Mumtaz, R., Qadir, M., Mumtaz, S., Khan, M. A., De Waele, T., ve Shahid, A. (2024). Cardiac arrhythmia classification using advanced deep learning techniques on digitized ECG datasets. *Sensors*, 24(8), 2484.
- Sfayyih, A. H., Sulaiman, N., ve Sabry, A. H. (2025). Non-invasive diagnosis of lung diseases via multimodal feature extraction from breathing audio and chest dynamics. *Computers in Biology and Medicine*, 191, 110182.
- Singh, C., ve Gaur, M. (2025). Lung SoundNet: SUSCC LSTM—a novel approach for automated lung sound diagnosis. *Biomedical Signal Processing and Control*, 99, 106841.
- Soundrapandiyan, R., Naidu, H., Karuppiah, M., Maheswari, M., ve Poonia, R. C. (2023). AI-based wavelet and stacked deep learning architecture for detecting coronavirus (COVID-19) from chest X-ray images. *Computers and Electrical Engineering*, 108, 108711.
- Taloba, A. I., ve Matoog, R. T. (2025). Detecting respiratory diseases using machine learning-based pattern recognition on spirometry data. *Alexandria Engineering Journal*, 113, 44–59.
- Tang, X., ve Liu, J. (2021). Comparing different algorithms for the course of Alzheimer’s disease using machine learning. *Annals of Palliative Medicine*, 10(9), 9715–9724. <https://doi.org/10.21037/apm-21-2013>
- Topaloğlu, I., Barua, P. D., Yıldız, A. M., Keleş, T., Doğan, S., Baygın, M., ve Acharya, U. R. (2023). Explainable attention ResNet18-based model for asthma detection using stethoscope lung sounds. *Engineering Applications of Artificial Intelligence*, 126, 106887.
- Tzeng, J.-T., Li, J.-L., Chen, H.-Y., Huang, C.-H., Chen, C.-H., Fan, C.-Y., ve Lee,

- C.-C. (2025). Improving the robustness and clinical applicability of automatic respiratory sound classification using deep learning-based audio enhancement. *JMIR AI*, 4, e67239.
- Uysal, S. (2014). *Ses analizi ile hastalık teşhisi* (Yayımlanmamış yüksek lisans tezi). Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Vaishya, R., Javaid, M., Khan, I. H., ve Haleem, A. (2020). Artificial intelligence applications for COVID-19 pandemic. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 14(4), 337–339.
- Waghmare, D. V., Deshmukh, R., Shrishrimal, P., ve Janvale, G. (2014). Emotion recognition system from artificial Marathi speech using MFCC and LDA techniques. *Proceedings of the 5th International Conference on Advanced Communication, Networking, and Computing*, 1–10.
- World Health Organization. (2023). *Pneumonia fact sheet*. <https://www.who.int>
- Xiao, S., Zhang, X., Lu, Y., Ye, P., Tang, Y., Zhang, P., ve Tai, J. (2025). Non-invasive detection of adenoid hypertrophy using deep learning based on heart-lung sounds. *IEEE Journal of Biomedical and Health Informatics*, 29(3), 1–13.*
- Yamashita, R., Nishio, M., Do, R. K. G., ve Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights into Imaging*, 9, 611–629.
- Zhang, B., Wang, J., Ye, Z., Zhou, L., Xiong, D., Wang, X., ve Guo, L. (2025). JASSNet: Heart and lung sound separation network based on joint attention mechanism and semi-supervised learning. *Biomedical Signal Processing and Control*, 104, 107525. <https://doi.org/10.1016/j.bspc.2025.107525>
- Zhao, J., Mao, X., ve Chen, L. (2019). Speech emotion recognition using deep 1D and 2D CNN-LSTM networks. *Biomedical Signal Processing and Control*, 47, 312–323. <https://doi.org/10.1016/j.bspc.2018.08.035>
- Zulfiqar, R., Zulfiqar, A., Zulfiqar, S., ve Zulfiqar, Z. (2021). Abnormal respiratory sounds classification using deep CNN through artificial noise addition. *Frontiers in Medicine*, 8, 714811.