

Classification of Lung Sounds with Deep Learning

 Mehmet Bilal ER 

Harran University Faculty of Engineering, Department of Computer Engineering, 63300, Haliliye/Şanlıurfa

Graphical/Tabular Abstract

Article Info:

Research article
 Received: 26/06/2020
 Revision: 05/10/2020
 Accepted: 05/10/2020

Highlights

- Lung sound.
- Data augment
- Deep learning

Keywords

Lung sounds classification
 Convolutional neural network
 Deep learning

In this study, a method based on deep learning has been proposed for the classification of lung sounds. For this purpose, the Convolutional Neural Network (CNN) has been designed. In addition, experiments are carried out using different machine learning methods based on feature extraction.

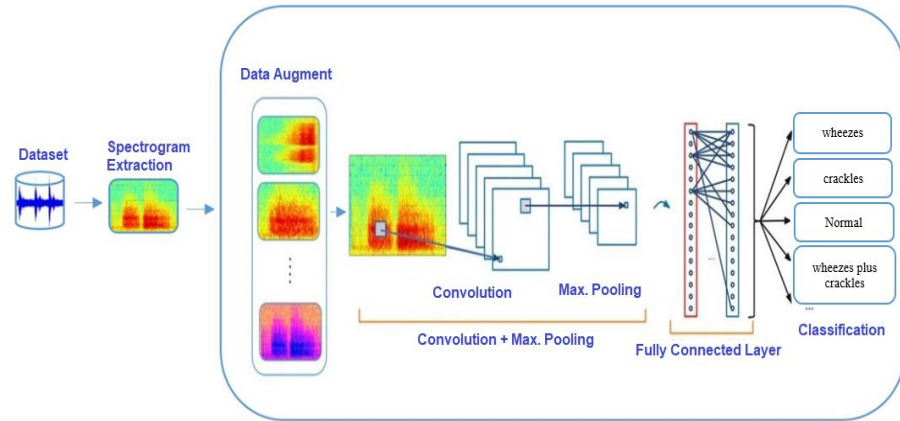


Figure A. Schema of the Proposed Method

Purpose: The focus is on automatic diagnosis of lung diseases, one of the most important issues in public health. There have been many studies on this subject in the literature, but most of these studies consist of traditional methods. The aim of this study is to increase the classification performance of lung sounds with deep learning.

Theory and Methods: The proposed method for the classification of lung sounds consists of five steps. First, lung sound signals are pre-processed, and spectrograms are obtained. After applying the data augmentation process to spectrograms, spectrogram images are given as input to the designed CNN model and the classification process is made.

Results: Experiments to evaluate the effectiveness of different methods are carried out using the ICBHI 2017 data set consisting of four classes commonly used in the literature. On average, 64.5% accuracy was obtained from the proposed method. The simulation and experimental results are presented and compared in Section 6.

Conclusion: In this study, deep learning is discussed to improve the classification performance of lung sounds. By designing a 12-layer CNN, spectrogram images are given to the first layer, and the classification process is made. Before the data augmentation, an average of 60.14% classification performance is obtained. After the data augmentation, an average classification performance of 64.50% is obtained. In addition, in order to compare the performance of the proposed method with other machine learning methods, MFCC features are extracted from the data set samples and the classification process is made with SVM and K-NN.

GU J Sci, Part C, 8(4): 830-844 (2020)

Gazi Üniversitesi

Fen Bilimleri Dergisi

PART C: TASARIM VE TEKNOLOJİ

<http://dergipark.gov.tr/gujsc>

Akciğer Seslerinin Derin Öğrenme ile Sınıflandırılması

Mehmet Bilal ER

Harran Üniversitesi Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, 63300, Haliliye/Şanlıurfa

Öz

Akciğer hastalıkları, insan sağlığını ciddi olarak tehdit eden hastalıklar arasındadır ve günümüzde birçok ölüm akciğer hastalıklarından kaynaklanmaktadır. Akciğer sesleri sayesinde akciğer hastalıkları ile ilgili önemli çıkarımlar yapılabilmektedir. Akciğer rahatsızlığı olan hastaları değerlendirmek için doktorlar genellikle oskültasyon tekniğini kullanırlar. Ancak, bu tekniğin bazı dezavantajları vardır. Örneğin, doktor iyi bir tıp eğitimi almamışsa, bu yanlış tanıya neden olabilir. Ayrıca akciğer sesleri durağan olmadığından dolayı analiz, tanıma süreci karmaşıktır. Bu nedenle otomatik tanıma sistemlerinin geliştirilmesi daha kesin ve doğru tanıları koymada yardımcı olacaktır. Akciğer hastalıklarını teşhis etmek ve uzmanlara teşhislerinde yardımcı olmak için geleneksel ses işleme rutinlerine dayanan birçok çalışma önerilmiştir. Bu çalışmada akciğer seslerinin sınıflandırılması için derin öğrenmeye dayanan bir yöntem önerilmiştir. Bu amaç doğrultusunda Evrişimsel Sinir Ağı (ESA) tasarlanmıştır. Ayrıca özellik çıkarımına dayanan farklı makine öğrenmesi yöntemleri kullanılarak da deneyler gerçekleştirilmiştir. Deneyler, farklı yöntemlerin etkinliğini değerlendirmek için literatürde yaygın olarak kullanılan, dört sınıftan oluşan ICBHI 2017 veri seti ile yapılmıştır. Ortalama olarak %64,5 doğruluk önerilen yöntemden elde edilmiştir.

Makale Bilgisi

Araştırma makalesi
Başvuru: 26/06/2020
Düzeltilme: 05/10/2020
Kabul: 05/10/2020

Anahtar Kelimeler

Akciğer seslerinin sınıflandırılması
Evrişimsel Sinir Ağı
Derin Öğrenme

Keywords

Lung sounds classification
Convolutional neural network
Deep learning

Classification of Lung Sounds with Deep Learning

Abstract

Lung diseases are among the diseases that seriously threaten human health, and many deaths today are caused by lung diseases. Thanks to the lung sounds, important inferences can be made about lung diseases. Doctors often use the auscultation technique to evaluate patients with lung conditions. However, this technique has some drawbacks. For example, this may lead to a misdiagnosis if the doctor has not received a good medical education. In addition, since the lung sounds are nonstationary, the analysis and recognition process is complex. Therefore, the development of automatic recognition systems will help in making more precise and accurate diagnoses. Many studies based on traditional sound processing routines have been proposed to diagnose lung diseases and to assist professionals in their diagnosis. In this study, a method based on deep learning has been proposed for the classification of lung sounds. For this purpose, the Convolutional Neural Network (CNN) has been designed. In addition, experiments are carried out using different machine learning methods based on feature extraction. Experiments to evaluate the effectiveness of different methods are carried out using the ICBHI 2017 data set consisting of four classes commonly used in the literature. On average, 64.5% accuracy is obtained from the proposed method. In addition, when the results obtained from the proposed method are compared with the latest methods in the literature, it is seen that it has a better performance in terms of classification success.

1. GİRİŞ (INTRODUCTION)

Modern tıpta teşhis, daha çok görsel veya işitsel bilgilerle ilişkilidir. Tıbbi bilgiler farklı yöntemlerden edinilebilir ve temel olarak bu veriler tıp uzmanına görüntü veya ses olarak gösterilir. İspirasyon ve ekspirasyon sırasında hava akımı nedeniyle akciğerler tarafından üretilen ses sinyalleri, solunum sisteminin durumu hakkında güçlü bir bilgi kaynağı oluşturur. Oskültasyon, doktorların bir stetoskop kullanarak pulmoner rahatsızlığı olan hastaları değerlendirip teşhis ettikleri bir tekniktir [1]. Bu teknik herhangi bir risk veya yan etki oluşturmaz. Bu yöntemden etkin ve doğru sonuçlar alınması, doktorun insan vücudu

tarafından üretilen normal ve anormal sesleri tanıma yeteneğine bağlıdır. Ayrıca, akciğer sesleri durağan olmayan sinyallerdir ve buna bağlı olarak oskültasyon yöntemi sırasında seslerin analiz edilmesi zorlaşacaktır. Bu nedenle, oskültasyon yönteminin bir örüntü tanıma sistemi ile birlikte kullanımı, daha kesin ve doğru klinik tanımlar koymada yardımcı olacaktır [2], [3]. Akciğer sesleri iki kategoriye ayrılabilir: normal (vesicular) solunum sesleri ve normal olmayan (adventitious) solunum sesleri. Solunum bozuklukları olmadığında normal solunum sesleri duyulur ve solunum bozukluğu olduğunda ise ek sesler duyulur [4]. Normal solunum sesleri dinlenirken trakeal, bronşiyal ve bronkoveziküler sesler dikkate alınmalıdır. Bu sesler için inspirasyon süresi ekspirasyon süresinden daha fazladır. Normal akciğer sesi 100-1000 Hz frekans aralığında uzanır ve ses sinyalinde farklı tepe noktaları yoktur [5], [6]. Normal olmayan solunum seslerinde normal nefes seslerine eklenen ek bir solunum sesi vardır. Bunlar sürekli ve ayrık olmak üzere iki gruba ayrılırlar [7]. Sürekli sesler üfürüm (wheeze) gibiyken, ayrık sesler ise çıtırtı (crackle) şeklindedir. Bu tür seslerin varlığı genellikle akciğer bozukluğunu gösterir. Müzikal bir karakteristik yapıya sahip olan sürekli gruptaki üfürüm sesleri genellikle 100 Hz'nin üzerinde baskın frekansa sahiptir. Ayrıca 100 ms'den büyük periyodik dalga formları oluşturan ses sinyalidir. Örnek olarak Rhonchi sürekli sesler grubundadır ve yaklaşık 200 Hz veya daha düşük bir baskın frekans tarafından karakterize edilir [6]. Ayrıca üfürüm sesleri, obstrüktif akciğer hastalığının yaygın bir belirtisidir ve bu hastalıklar astım, zatürre ve bronşit olabilir. Çıtırtı sesleri ise ayrık akciğer sesleridir ve kardiyovasküler hastalıklarda sıklıkla görülür. Süreleri 20 ms'den azdır ve frekans içerikleri tipik olarak geniştir [8].

Literatürde birçok akciğer sesi kategorisi vardır. Bu çalışmada ICBHI 2017 veri setindeki sınıflar incelenmiştir. Geleneksel akciğer sesi sınıflandırma çalışmalarında genellikle akciğer sesi sinyallerinden farklı özellikler çıkarılmaktadır. Çıkarılan özellikler ile sınıf etiketleri arasındaki ilişkiyi öğrenmek için farklı makine öğrenmesi algoritmaları uygulanmaktadır. Bu çalışmada akciğer sesi sınıflandırmasında derin öğrenmeye dayalı bir yöntem kullanılmaktadır. İlk olarak ses sinyalleri ön işlemden geçirilmiştir. Daha sonra akciğer sesi sinyallerinden karışık özellikler çıkarmak yerine, bu sinyalleri görsel olarak ifade eden spektrogramlar çıkarılmıştır. Spektrogramlar tasarlanan ESA'ya girdi olarak verilerek özellik çıkarma ve sınıflandırma işlemi yapılmıştır.

Bu makalenin geri kalanı şu şekilde düzenlenmiştir. Bölüm 2'de literatürdeki akciğer seslerinin sınıflandırılması ile ilgili çalışmalar gözden geçirilmiştir ve aralarındaki farklar ortaya konulmuştur. Bölüm 3, 4 ve 5'te materyal ve önerilen yöntem tanımlanmıştır. Bölüm 6'da veri seti ve araştırmada kullanılan akciğer seslerinin sınıflandırılmasıyla ilgili deneysel uygulamalar verilmiştir. Bölüm 7'de ise araştırmanın bulguları tartışılmıştır.

2. İLGİLİ ÇALIŞMALAR (RELATED WORKS)

Akciğer seslerinin analizini ve sentezini içeren bazı dikkate değer çalışmalar bu bölümde verilecektir. Örüntü tanıma ve yapay zekanın gelişmesiyle, farklı akciğer seslerinin sınıflandırılması için otomatik sistemler geliştirmek için birçok özelliğe dayalı yaklaşım önerilmiştir.

İçer ve Gengeç (2014), normal ve anormal akciğer seslerini sınıflandırmak için güç spektral yoğunluk değerlerinin frekans oranını ve Hilbert-Huang Dönüşümü (HHD) özelliklerini kullandılar. Sınıflandırıcı olarak Destek Vektör Makineleri (DVM) kullanılmıştır ve %90'ın üzerinde doğruluk elde edilmiştir [2].

Naves ve ark. (2016), akciğer seslerinin sınıflandırılması için K-NN ve Bayes sınıflandırıcıları tercih etmişlerdir. Kullanılan veri setinde beş tip akciğer sesi bulunmaktadır. Genetik algoritmalar ve Fisher doğrusal ayırıcı özellik azaltmak için kullanılmıştır. Sınıflandırıcı doğruluğu ortalama olarak %98,1' dir [3]. Serbes ve ark. (2013), zaman frekansı ve zaman ölçeği (TS) analizini önermektedir. Özellik çıkarma adımı, zaman frekansı ve zaman ölçeği analizi kullanılmıştır. Ön işleme için, çıtırtı bilgisi içermeyen frekans bantlarını çıkarmak amacıyla Çift ağaçlı karmaşık dalgacık dönüşümü (Dual-tree complex wavelet transform - DTCWT) uygulanmıştır. Sınıflandırma adımı, çıtırtı ve çıtırtı olmayan sesleri sınıflandırmak için K-En Yakın Komşular (K-NN), DVM ve Yapay Sinir Ağları (YSA) kullanılmıştır. En iyi sınıflandırma sonucu %97,5 olarak DVM'den elde edilmiştir [9].

Jin ve Sattar (2014), solunum seslerini sınıflandırmak için DVM kullanmışlardır. Sinyallerden anlık basıklık (kurtosis), ayırma fonksiyonu (diskriminant fonksiyon) ve entropi gibi özellikler çıkarılarak optimal olarak %97,7 ile %98,8 arasında sınıflandırma doğruluğu elde edilmiştir [10].

Bahoura (2009), iki sınıftan oluşan akciğer solunum seslerinin sınıflandırılması için farklı bir yaklaşım önermiştir. Sinyallerden Mel-frekanslı cepstral katsayıları (MFCC) çıkarılmıştır ve özellik kümesi olarak

kullanılmıştır. Sınıflandırmak için ise Gauss Karışım Modeli kullanılmıştır ve %94,2 sınıflandırma başarımı elde edilmiştir [11].

Orjuela ve ark. (2014), normal, üfürüm ve çıtırtı solunum seslerini sınıflandırmak için YSA ile birlikte MFCC özelliklerini kullanmışlardır. Sınıflandırma başarımı çıtırtı solunum seslerinde %75, üfürüm solunumlarında %100 ve normal solunum seslerinde ise %80 olarak hesaplanmıştır [12].

Maruf ve ark. (2015), Gauss Karışım Modeli'ni kullanarak çıtırtı solunum seslerinin normal solunum seslerinden ayrılması ve sınıflandırılması için bir yöntem önermişlerdir. Bu çalışma dört adımdan oluşmaktadır. Bu adımlar önileme, özellik çıkarma, özellik seçme ve sınıflandırma şeklindedir. Ön işleme adımında, arka plan gürültüsü azaltımı için bir bant geçiren filtre kullanılmıştır ve daha sonra enerji ve spektrogram gibi özellikler çıkarılmıştır. Özellik seçme adımından sonra, nihai özellikler Gauss Karışım Modeli kullanılarak eğitilmiştir. Çalışmadan elde edilen genel doğruluk %97,56'dır [13].

Yeginer ve Kahya (2007), veziküler sesin çıtırtılı solunum ses sinyalinin atılması için otomatik bir yöntem önermişlerdir. İlgilenilen bir bölgeyi seçtikten sonra, o bölgedeki ham ve filtrelenmiş dalga formları arasındaki korelasyona dayalı bozulma metriği tanımlanmıştır. Filtre kesme frekansı bozulma metriğine göre tahmin edilmiştir. Hesaplama maliyetini azaltmak için, tahmini kesme frekansından yeni bir kesme frekansı öngören bir regresyon analizi de gerçekleştirilmiştir [14].

Serbes ve ark. (2017), çeşitli akciğer seslerini sınıflandırmak için radyal temelli fonksiyon çekirdeğe sahip DVM kullanmışlardır. Veri seti olarak 126 kişiden elde edilen 920 kayıttan oluşan ICBHI 2017 kullanılmıştır. Seslerin ayırt edici özelliklerini temsil eden spektral özellikler çıkarılarak dört sınıftan oluşan veri seti sınıflandırılmaya çalışılmıştır. %57,88 doğruluk elde edilmiştir [15].

Chambres ve ark. (2018), güçlendirilmiş bir karar ağacı modeli ve daha fazla ses özelliğine sahip bir makine öğrenimi yaklaşımı önermişlerdir. Veri seti olarak ICBHI 2017 kullanılmıştır. Elde edilen deney sonuçlarına göre %49,63 sınıflandırma başarımına ulaşılmıştır [16]. ESA'lar, geniş ölçekli nesne sınıflandırması için oldukça iyi sonuçlar vermektedir.

Bardou ve ark. (2018), akciğer seslerinin sınıflandırılması için üç makine öğrenme yaklaşımını karşılaştırmışlardır. İlk iki yaklaşım, manuel özellik (handcrafted features) çıkarımı ve DVM, K-NN ve Gauss Karışım Modelinin eğitimine dayanmaktadır. Üçüncü yaklaşım, ESA tasarımına dayanmaktadır. İlk yaklaşımda, ses dosyalarından 12 MFCC çıkarılmıştır. İkinci yaklaşımda ses dosyalarını görsel olarak temsil eden spektrogramlar çıkarılmıştır. ESA'nın nihai doğruluğunu arttırmak için spektrogramlara veri seti büyütme teknikleri uygulanmıştır. Sonuçlar, ESA'nın manuel çıkarılan özelliklere dayalı sınıflandırıcılardan daha iyi performans gösterdiğini göstermektedir [17].

Jácome ve ark. (2019), akciğer ses kayıtlarından solunum fazı tespiti için derin öğrenmeye dayalı bir algoritma oluşturmuşlardır. Algoritma da evrişimli bir sinir ağı kullanılmıştır ve özellik olarak ise spektrogramlar tercih edilmiştir. Literatürde var olan ve üç alt kümesi bulunan veri seti kullanılarak değerlendirilmiştir ve ortalama olarak %84 başarımla elde edilmiştir. [18].

Demir ve ark. (2019), akciğer ses sinyalleri başlangıçta zaman-frekans yöntemi kullanılarak spektrogram görüntülerine dönüştürmüştür. Sınıflandırma için iki derin öğrenme tabanlı yaklaşım kullanılmıştır. İlk yaklaşımda, özelliklerin çıkarılması için önceden eğitilmiş bir derin ESA modeli ve akciğer seslerinin sınıflandırılması için ise DVM sınıflandırıcısı kullanılmıştır. İkinci yaklaşımda, önceden eğitilmiş ESA modeli kullanılarak transfer öğrenme işlemi uygulanmıştır. Önerilen yöntemlerin doğrulukları on kat çapraz doğrulama kullanılarak test edilmiştir. Önerilen birinci ve ikinci yöntemlerden elde edilen doğruluklar sırasıyla %65,5 ve %63,09'dur [1].

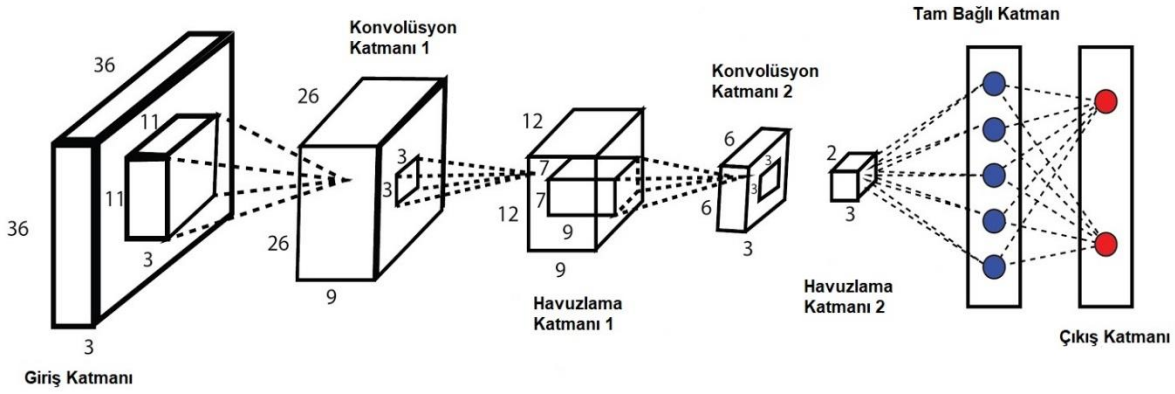
Kochetov ve ark (2018), solunum seslerinin sınıflandırılması için gürültü maskeleyme tekniği ve tekrarlayan sinir ağı kullanarak yeni bir model önermişlerdir. Önerilen model, solunum benzeri sesleri gürültü olmadan çıkarmaya dayanmaktadır. Çıkarılan özellikler dört sınıftan oluşan veri seti örneklerini sınıflandırmak için kullanılmıştır. Önerilen model, açık erişimli veri seti üzerinde test edilmiştir ve %65,7 başarımla elde edilmiştir [19].

Bu çalışmanın ana katkıları aşağıdaki gibidir:

- Sınıflandırma başarımını arttırmak için veri artırma işlemi uygulanarak veriler eğitilmiştir.
- Çoğu akciğer sesi sınıflandırma çalışmalarında ya derin öğrenme ya da makine öğrenimi tek başına kullanılmıştır. Bu çalışmadaki deneylerde derin öğrenme ve makine öğreniminin sınıflandırma sonuçları birlikte verilmiştir.

3. EVRİŞİMSSEL SİNİR AĞLARI (CONVOLUTIONAL NEURAL NETWORK)

ESA, yapay sinir ağlarının özel bir mimarisidir. Ana görevi, derin öğrenme teknolojisinde görüntülerin etkili bir şekilde tanınmasıdır. ESA mimarisi, insan beyindeki nöronların bağlantı modeline benzer ve görsel korteksin yapısından esinlenilerek tasarlanmıştır [20]. Diğer görüntü sınıflandırma algoritmalarına kıyasla, bu mimariler minimum ön işleme kullanır. ESA'lar, bilgisayarlı görme alanında en etkili yeniliklerden biri olmuşlardır ve geleneksel yaklaşımlardan çok daha iyi performans göstermişlerdir [21]. Görüntü sınıflandırma, nesne algılama, segmentasyon, yüz tanıma gibi birçok çalışmada başarılı sonuçlar vermişlerdir. Bununla birlikte, ESA'lar görüntü tanıma ile sınırlı değildir. Bu mimariler ses işleme, metin işleme gibi çalışmalarda da etkin sonuç vermektedir. Şekil 1'de ESA genel mimarisi yer almaktadır. Teknik olarak genel bir ESA modelinde konvolüsyon, havuzlama, tam bağlı katmanlar ve sınıflandırma katmanı bulunur [22].



Şekil 1. Evrişimsel Sinir Ağı

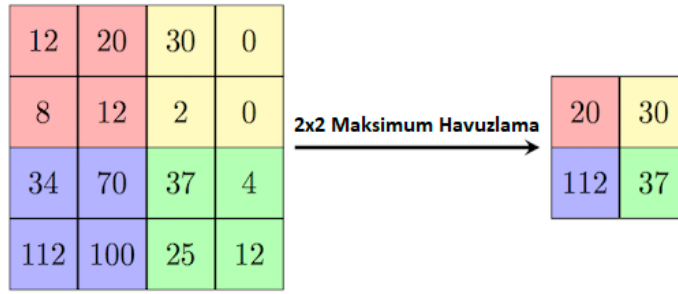
Konvolüsyon, bir giriş görüntüsünden özellikleri ayıklayan ilk katmandır. Evrişim, giriş görüntüsünden özellik haritası çıkarmak için giriş görüntüsüne uygulanan filtredir. Görüntü matrisi ve filtre gibi iki giriş alan bir matematik işlemidir. Kenar algılama, bulanıklaştırma ve keskinleştirme gibi filtreler uygulayarak konvolüsyonu işlemi gerçekleştirilir. Filtrelerin yüksekliği ve ağırlığı giriş hacminden daha küçüktür. Konvolüsyon işlemi için gerekli formül denklem (1)'de verilmiştir. Giriş görüntüsü f ile, çekirdeğimiz h ile gösterilir. Sonuç matrisinin satır ve sütun indeksleri sırasıyla m ve n ile gösterilir.

$$G[m, n] = (f * h)[m, n] = \sum_j \sum_k h[j, k] f[m - j, n - k] \quad 1$$

Her konvolüsyondan sonra aktivasyon fonksiyonu kullanılır. Derin ağlarda aktivasyon fonksiyonu olarak genellikle Düzleştirilmiş Doğrusal Birim (ReLU) kullanılır. ReLU'nun amacı modelin doğrusal olmayan bir yapıya gelmesini sağlamaktır. ReLU fonksiyonu denklem (2)'de verilmiştir.

$$ReLU(x) = \max(x, 0) \quad 2$$

Denklem 2'ye göre ReLU çıkışı, sıfır ile giriş değeri arasındaki maksimum değerdir. Giriş değeri negatif olduğunda çıkış değeri sıfıra, giriş pozitif olduğunda çıkış değeri giriş değerine eşittir. ESA'larda ayrıca havuzlama (pooling) katmanı bulunur. Havuzlama katmanı, evrişimden sonra girdi görüntüsünün uzaysal hacmini azaltmak için ve hesaplamaları hızlandırmak için kullanılırlar. Şekil 2'de maksimum havuzlama (max pooling) örneği verilmektedir. Örnekte 4×4 boyut girişinin 2×2 boyuta azaldığı görülmektedir.



Şekil2. Maksimum Havuzlama

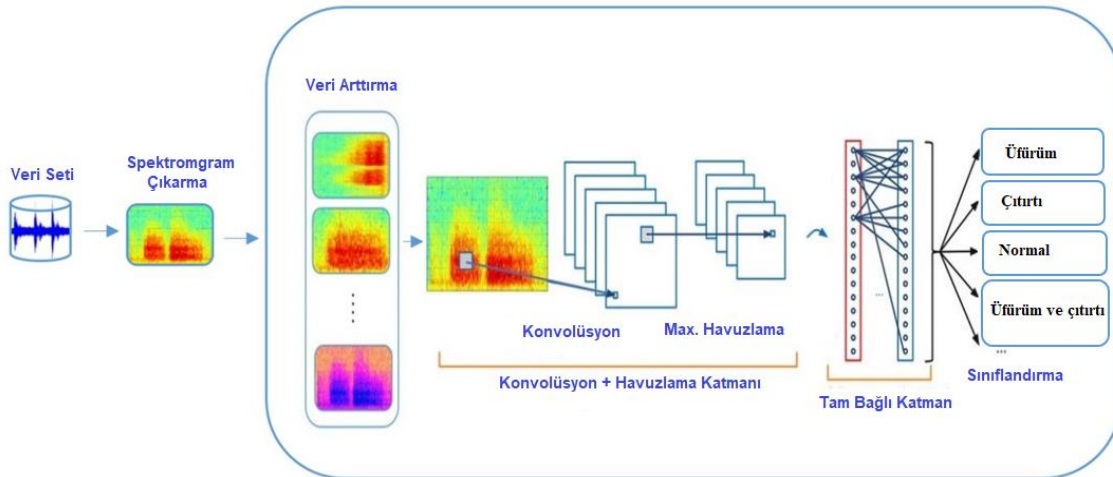
Tamamen bağlı katmanlar, görüntüleri tanıma ve sınıflandırmada çok başarılı olduğu kanıtlanmış derin ağların önemli bir bileşenidir. ESA süreci, evrişim ve havuzlama, görüntüyü unsurlara ayırma ve bağımsız olarak analiz etme ile başlar. Bu sürecin sonucu, nihai sınıflandırma kararını yönlendiren tamamen bağlantılı bir sinir ağı yapısına bağlanır. Tamamen bağlı giriş katmanı, önceki katmanların çıktısını alıp düzleştirir ve sonraki aşama için girdi olabilecek tek bir vektöre dönüştürür. Tamamen bağlı giriş katmanından sonra özellik girdileri alınır ve doğru etiketi tahmin etmek için ağırlıklar güncellenir. Son katman ise çıkış katmanıdır ve her etiket için olasılıklar tahmin edilir. Bu katmanda genellikle Softmax tercih edilir. Softmax formülü denklem (3)'de verilmiştir.

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

3

4. ÖNERİLEN YÖNTEM (THE PROPOSED METHOD)

Akciğer seslerinin sınıflandırılması için önerilen yöntem beş adımdan oluşmaktadır. İlk olarak akciğer sesi sinyalleri ön işlemden geçirilerek spektrogramlar çıkarılmıştır. Spektrogramlara veri artırma (data augment) işlemi uygulandıktan sonra spektrogram görüntüleri tasarlanan ESA modeline girdi olarak verilerek sınıflandırma işlemi yapılmıştır. Her adımın detayı aşağıdaki alt bölümlerde açıklanmıştır. Önerilen yöntem Şekil 3'de verilmiştir.



Şekil 3. Önerilen Yöntem

4.1. Sinyal Ön İşleme ve Spektrogram Çıkarma (Signal Pre-Processing and Spectrogram Extraction)

Akciğer ses sinyalleri ilk olarak 150-250 Hz frekans aralığında bant geçiren filtre uygulanarak ön işleme tabi tutulmuştur. Filtreden geçirilen akciğer sinyallerinden spektrogram çıkarabilmek için sinyal 20 ms

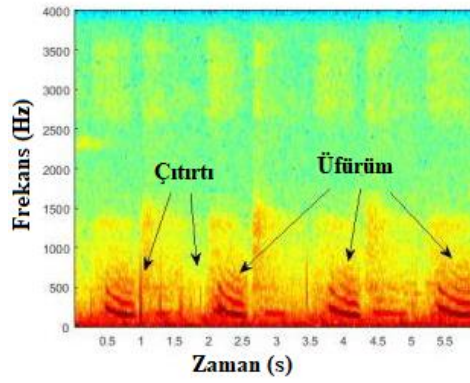
uzunluklarda çerçevelere ayrılmıştır. Birbirini takip eden çerçevelerin örtüşme oranı %50 olarak ayarlanmıştır. Bir sinyale çerçeveleme uygulandıktan sonra pencereleme (window) işlemi uygulanır. Bu araştırmada “Hamming penceresi” tercih edilmiştir. Hamming pencereleme formülü denklem (4)’de verilmiştir.

$$k[n] = 0.54 - 0.46 \cos\left(2\pi \frac{n}{N}\right), \quad N: \text{Pencere uzunluğu} \quad 4$$

$$n: 0, 1, 2 \dots \dots N - 1$$

Hamming penceresinden sonra sinyale Fourier dönüşümü uygulanmıştır. Bu araştırmada Hızlı Fourier Dönüşümü (FFT) kullanılmıştır. FFT dönüşümü denklem (5)’te verilmiştir. Son aşamada ise Fourier dönüşümü uygulanan sinyallerin güç spektrumları çıkarılmıştır. Örnek bir akciğer ses sinyali spektrum görüntüsü Şekil 4’te verilmiştir.

$$X_n = \sum_{k=0}^{N-1} x_k e^{-2\pi jkn / N}, \quad n = 0, 1, 2, \dots, N - 1 \quad 5$$



Şekil 4. Akciğer Ses Sinyali Spektrumunun Gösterimi

4.2. Veri Arttırma (Data Augmentation)

Veri arttırma, orijinal verileri değiştirerek yeni veri noktaları oluşturma işlemi içerir. Bu süreçte, yeni veri toplamaya gerek kalmadan derin öğrenme alanındaki eğitim modelleri için mevcut verilerin çeşitliliği artırılır [23]. Ayrıca genel olarak, derin öğrenme modellerinin performansını arttırmada etkin bir işlemdir. Veri arttırma işlemi görüntüye veya sese uygulanabilir. Sese uygulanan veri arttırma işlemleri arka plana gürültü eklenmesi, sinyalin kaydırılması ve sinyalin uzatılması şeklinde olabilir. Görüntüye uygulanan veri arttırma işlemleri döndürme, yeniden boyutlandırma, kırpmaya ve farklı şekillerde olabilir [24]. Bu araştırmada veri arttırma işlemi orijinal veriler olan ses sinyalleri yerine, ses sinyallerinden elde edilen spektrum görüntülerine uygulanmıştır. Sese uygulanan tekniklerin kullanılmamasının sebebi veri setindeki ses sinyallerinin süresinin kısa olmasıdır. Ayrıca sesin arka planına gürültü eklenmesi gibi işlemler normal akciğer sesleri ile çıtırtılı ve üfürümlü akciğer seslerinin birbirine karışmasına neden olacaktır. Bu sebeplerden dolayı görüntüye uygulanan 5 farklı veri arttırma tekniği bu çalışmada kullanılmıştır. Bunlar, rastgele döndürme, kesme (shearing), renk kaydırma (color shifting), Gauss filtresi ve Tuz ve Biber Gürültüsü (Salt and Pepper Noise) şeklindedir. Bu işlem sonucunda bir dizi deformasyona uğratılmış ek eğitim verileri oluşturulmuştur. Fakat orijinal spektrum görüntüleri veri setinde korunmuştur.

4.3. ESA Tasarımı (Design of CNN)

Tasarlanan ESA'nın topolojisi, 1 giriş katmanı, 4 konvolüsyon katmanı, 4 havuzlama katmanı, iki tam bağlantılı katmanı ve 1 çıkış katmanından oluşur. Tasarlanan mimaride toplam 12 katman bulunmaktadır. Birinci konvolüsyon katmanında 7×7 boyutunda 64 adet konvolüsyon filtresi, ikinci konvolüsyon katmanında 5×5 boyutunda 128 adet konvolüsyon filtresi, üçüncü konvolüsyon katmanında 3×3 boyutunda 256 adet konvolüsyon filtresi, dördüncü konvolüsyon katmanında ise 3×3 boyutunda 384 adet konvolüsyon

filtresi bulunmaktadır. Farklı sayılarda ve boyutlarda konvolüsyon katmanı kullanımının amacı spektrogramlardan etkin derin özellik çıkarımını sağlamak ve hesaplama maliyetini azaltmaktır. Aktivasyon fonksiyonu olarak ReLU kullanılmıştır. Tüm konvolüsyon katmanlarından sonra ReLU katmanı (aktivasyon katmanı) uygulanmıştır. Amaç, doğrusal olmayışı önerilen sisteme uyarlamaktır. Ayrıca her konvolüsyon ve ReLU işleminden sonra [2 2] boyutunda kaydırma adımıyla 3×3 maksimum havuzlama yapılmıştır. Tasarlanan mimari Tablo 1’de verilmiştir. Öğrenme oranı, epok sayısı ve mini-batch size ESA’nın performansını etkileyen en önemli hiper parametrelerdir. Tasarlanan mimaride mini-batch boyutu 64, maksimum epok sayısı 20 ve öğrenme oranı 1e-4 olarak ayarlanmıştır.

Tablo 1. Tasarlanan ESA mimarisi

No	Katman İsmi	Açıklama	Özellikler
1	Giriş	Görüntü girişi	Spektrogram
2	'conv1'	Konvolüsyon	7×7 boyutunda 64 adet konvolüsyon filtresi
3	'ReLU1'	ReLU	
4	'pool1'	Maksimum Havuzlama	[2 2] boyutunda kaydırma adımıyla 3×3 havuzlama
5	'conv2'	Konvolüsyon	5×5 boyutunda 128 adet konvolüsyon filtresi
6	'ReLU2'	ReLU	
7	'pool2'	Maksimum Havuzlama	[2 2] boyutunda kaydırma adımıyla 3×3 havuzlama
8	'conv3'	Konvolüsyon	3×3 boyutunda 256 adet konvolüsyon filtresi
9	'ReLU3'	ReLU	
10	'pool3'	Maksimum Havuzlama	[2 2] boyutunda kaydırma adımıyla 3×3 havuzlama
11	'conv4'	Konvolüsyon	3×3 boyutunda 384 adet konvolüsyon filtresi
12	'ReLU4'	ReLU	
13	'pool4'	Maksimum Havuzlama	[2 2] boyutunda kaydırma adımıyla 3×3 havuzlama
14	'fc1'	Tam Bağlı Katman	
15	'ReLU7'	ReLU	
16	'drop7'	Dropout	50% dropout
17	'fc1'	Tam Bağlı Katman	
18	Çıkış	Softmax	

5. MAKİNE ÖĞRENİMİNE DAYALI SINIFLANDIRMA (CLASSIFICATION BASED ON MACHINE LEARNING)

Özellik tabanlı yaklaşım ile ESA arasında performans karşılaştırması yapmak için veri setinden çıkarılan özellikler ile makine öğrenimine dayalı sınıflandırma yapılmıştır. Makine öğrenmesine dayalı sınıflandırma işleminde özellik çıkarma ve daha sonrasında ise sınıflandırma işlemi yapılmıştır. İlk olarak MFCC özellik olarak çıkarılmıştır. MFCC, otomatik ses tanımda en sık kullanılan öznelik çıkarma yöntemidir. MFCC özelliklerini çıkarmak için gerekli adımlar sırası ile aşağıda verilmiştir;

- **Çerçeveleme:** Sinyal 25 ms uzunluklarda çerçevelere bölünmüştür ve her çerçeve birbirini %50 oranında örtecek şekildedir.
- **Pencereleme:** Hamming pencereleme tekniği uygulanmıştır.
- **FFT:** Hızlı Fourier Dönüşümü uygulanmıştır.
- **Mel Filtre:** Kulağın duyma hassasiyetini modellemek için sinyal mel filtre bankasından geçirilmiştir. Bu filtre bankaları üçgen şeklide birbirini %25 oranında örtecek şekilde tasarlanmıştır.
- **Kepstrum Çıkarımı:** Son adım, kepstrum katsayılarının elde edilmesine yönelik bir çalışmadır. Tüm filtre bankalarının enerjilerinin logaritmasını alınır ve Ayrık Kosinüs dönüşümü uygulanarak Mel Frekans Kepstrum katsayıları elde edilir.

Bu işlemlerin ardından 12 MFCC özelliği elde edilmiştir. Son olarak 12-MFCC katsayılarından 6 istatistik değer hesaplanarak 72 boyutlu bir özellik vektörüne elde edilmiştir. Hesaplanan istatistiksel değerler ortalama, standart sapma, minimum, maksimum, mutlak farkın ortalaması ve mutlak farkın standart sapması şeklindedir. Sınıflandırma için ise DVM ve K-NN kullanılmıştır.

DVM: Destek vektör makinesi sınıflandırıcısı, hem doğrusal hem de doğrusal olmayan verilerin sınıflandırılması için iyi sonuçlar veren bir yöntemdir [25]. DVM, konuşma tanıma, metin sınıflandırması ve nesne sınıflandırması gibi birçok alandaki sınıflandırma problemlerinde kullanılmaktadır.

K-NN: K-En Yakın Komşular sınıflandırıcı örüntü tanımada kullanılan en basit ve etkili yöntemlerden biridir. Genellikle verimli performans sağlar ve bazı durumlarda doğruluğu, modern sınıflandırıcılardan daha yüksektir. Bu algoritma ile yeni bir nokta verisini sınıflandırmak için bu nokta verisine en yakın K tane eğitim noktası bulunur. Sınıflandırma işlemi, komşularının oy çoğunluğuyla yapılır; sınıflandırma işlemi yapılacak bir eleman, bir uzaklık fonksiyonuyla ölçülen en yakın komşular arasında en yakın olan sınıfa dağıtılır [26].

6. DENEYSEL UYGULAMALAR (EXPERIMENTAL APPLICATIONS)

6.1 Veri Seti (Dataset)

Bu çalışmada ICBHI 2017 veri seti kullanılmıştır. Bu veri tabanı, farklı kayıt araçlarıyla farklı araştırma ekiplerinin solunum sesleri kayıtlarını toplamasıyla hazırlanmıştır [27]. Veri setinde 126 kişiden alınan 920 ses örneği vardır. Sınıf etiketlerine göre, bu ses dosyalarının her biri, döngülere bölünür. Örnek bir akciğerin sesinin etiketleme detayları Tablo 2'de verilmiştir. Tablo 2 'ye göre, verilen kayıta yer alan her döngü için, zaman içindeki konumları verilir ve anormal ses içeren veya içermeyen bilgiler ikili bir puan olarak verilir. Çıtır ve üfürüm değerleri sırasıyla 1-0 ise sınıf etiketi “çıtır”, 0-1 ise “üfürüm”, 0-0 ise “normal” ve 1-1 ise “üfürüm ve çıtır” şeklindedir. Çıtır sınıfında 1864, üfürüm sınıfında 886, normal sınıfında 3642, üfürüm ve çıtır sınıfında ise 502 kayıt bulunmaktadır.

Tablo 2. Örnek ses dosyası için döngü bilgisi

Döngü	Başlangıç	Son	çıtır (crackle)	üfürüm (wheeze)
Döngü 1	0,541	3,612	1	0
Döngü 2	3,612	5,234	1	0
Döngü 3	5,234	7,426	1	1
Döngü 4	7,426	11,845	1	0
Döngü 5	11,845	14,157	0	1
Döngü 6	14,157	16,124	1	0
Döngü 7	16,124	17,271	1	0
Döngü 8	17,271	19,772	1	0

6.2. Deneysel Sonuçlar (Experimental Results)

Bu çalışmada sınıflandırma işlemi için akciğer ses kayıtlarından elde edilen spektrogramlar ve tasarlanan ESA mimarisi kullanılmıştır. Önerilen yöntemin başarımları kriterleri doğruluk, kesinlik ve F-skoru oranlarına dayalı olarak yapılmıştır. Sınıflandırma problemlerindeki değerlendirme ölçütleri, karışıklık matrisi adı verilen her sınıf için doğru ve yanlış sınıflandırılmış örnek sayılarının bulunduğu bir matris kullanılarak yapılır [28]. Doğruluk ölçümü denklem (6)'da verilmiştir.

- Yanlış pozitifler (YP): negatif sınıftan olan, pozitif olarak tahmin edilen örnekler.
- Yanlış negatifler (YN): gerçek sınıfı pozitif olan negatif olarak tahmin edilen örnekler.
- Doğru pozitifler (DP): pozitif sınıfa ait doğru tahmin edilen örnekler.
- Doğru negatifler (DN): negatif sınıfa ait olarak doğru tahmin edilen örnekler.

$$Doğruluk = \frac{|DN|+|DP|}{|YN|+|YP|+|DN|+|DP|}$$

Kesinlik (P), pozitif bir tahminin doğru olma olasılığını tahmin eden bir ölçüdür. Kesinlik ölçümü denklem (7)'de verilmiştir.

$$Kesinlik(P) = \frac{|DP|}{|DP|+|YP|} \quad 7$$

F-skoru, pozitif kestirim oranı ve duyarlılık ölçülerinin uyumlu bir ortalaması olup denklem (8)'de gösterildiği gibi hesaplanır.

$$F - skor = \frac{2*|DP|}{2|DP|+|YP|+|YN|} \quad 8$$

Eğitim ve test için ayrılan verilerin boyutunun ağın performansı üzerindeki etkisini belirlemek amacıyla, bu veriler 2 farklı şekilde bölünmüştür. İlk deneyde, verilerin %70'i eğitim için ve %30'ı test için kullanılır; ikinci deneyde, verilerin %80'i eğitim için ve %20'si test için kullanılmıştır. Tablo 3'de tasarlanan ESA mimarisi kullanılarak veri arttırma işleminden önce elde edilen sınıflandırma sonuçları verilmiştir. Tablo 3'deki verilere göre verilerin eğitim ve test için sırasıyla %70-%30 bölünmesiyle en yüksek doğruluk %62,47 olarak "üfürüm ve çıtırı" sınıfından elde edilmiştir. Ayrıca ortalama sınıflandırma başarımı %59,57, ortalama kesinlik değeri %54 ve ortalama F-skoru değeri ise %57,84' tür. Verilerin eğitim ve test için sırasıyla %80-%20 bölünmesiyle en yüksek doğruluk %63,14 olarak "üfürüm ve çıtırı" sınıfından elde edilmiştir. Ayrıca ortalama sınıflandırma başarımı %60,14, ortalama kesinlik değeri %56,66 ve ortalama F-skoru değeri ise %53,11'dir.

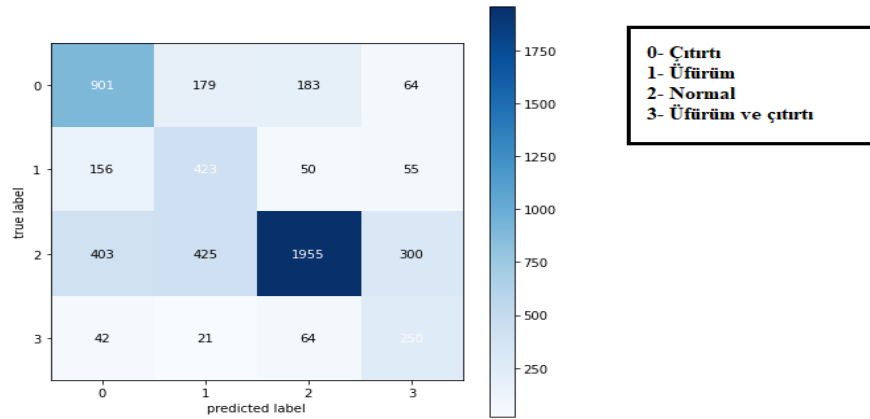
Tablo 3. Veri Arttırma İşlemi Uygulamadan Önce Elde Edilen Sınıflandırma Sonuçları

Verinin Eğitim ve Test için Farklı oranlarda Bölünmesi	Sınıf	Doğruluk %	Kesinlik %	F-skoru %
%70- %30	Çıtırı	58,45	68,21	68,36
	Üfürüm	60,22	38,45	33,41
	Normal	57,14	62,27	72,36
	Üfürüm ve çıtırı	62,47	46,85	57,25
	Ortalama Başarım	59,57	54,00	57,84
%80- %20	Çıtırı	61,26	55,36	56,46
	Üfürüm	61,87	50,27	42,35
	Normal	54,29	69,67	75,36
	Üfürüm ve çıtırı	63,14	51,34	38,28
	Ortalama Başarım	60,14	56,66	53,11

Tablo 4'de tasarlanan ESA mimarisi kullanılarak veri arttırma işlemi sonrası elde edilen sınıflandırma sonuçları verilmiştir. Verilerin eğitim ve test için sırasıyla %70-%30 bölünmesiyle en yüksek doğruluk %65,28 olarak "çıtırı sınıfından elde edilmiştir. Ayrıca ortalama sınıflandırma başarımı %62,10, ortalama kesinlik değeri %53,87 ve ortalama F- skoru değeri %50,29'dur. Verilerin eğitim ve test için sırasıyla %80-%20 bölünmesiyle en yüksek doğruluk %67,89 olarak "çıtırı" sınıfından elde edilmiştir. Ayrıca ortalama sınıflandırma başarımı %64,50, ortalama kesinlik değeri %65,70 ve ortalama F- skoru değeri %59,16'dır. Verilerin eğitim ve test için sırasıyla %80-%20 bölünmesiyle elde edilen sonucuna ait karışıklık matrisi şekil 5'te verilmiştir. Tablo 3 ve Tablo 4'teki sonuçlara göre, eğitim verisinin arttırılması ile birlikte sınıflandırma doğruluğu da arttığı gözlemlenmiştir. Bu durum beklenildiği gibidir çünkü görüntülerin farklı varyasyonları (ek eğitim verileri) ile derin ağların eğitilmesi derin modellerin performansını önemli ölçüde etkilemektedir.

Tablo 4. Veri Arttırma İşlemi Uygulandıktan Sonra Elde edilen Sınıflandırma Sonuçları

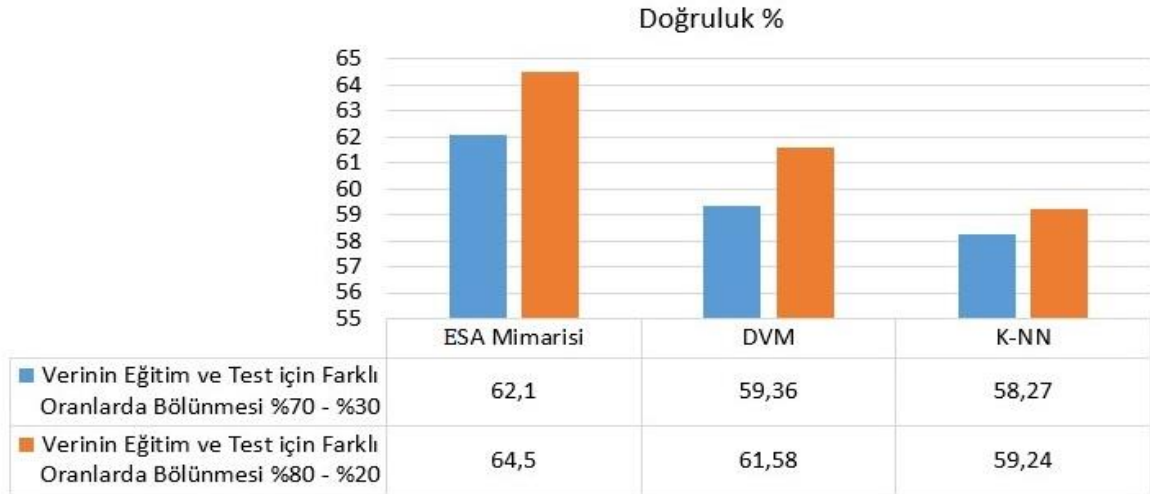
Verinin Eğitim ve Test için Farklı oranlarda Bölünmesi	Sınıf	Doğruluk %	Kesinlik %	F-skoru %
%70- %30	Çıtırıtı	65,28	68,76	68,42
	Üfürüm	60,45	38,24	33,26
	Normal	58,67	62,33	72,17
	Üfürüm ve çıtırıtı	64,00	46,15	27,32
	Ortalama Başarım	62,10	53,87	50,29
%80- %20	Çıtırıtı	67,89	68,47	65,14
	Üfürüm	61,84	62,67	49,38
	Normal	63,41	60,45	71,00
	Üfürüm ve çıtırıtı	66,31	71,23	51,15
	Ortalama Başarım	64,50	65,70	59,16

**Şekil 5.** En İyi Sınıflandırma Sonucuna Ait Karışıklık Matrisi

Tablo 5’ de MFCC özellikleri kullanılarak DVM ve K-NN’den elde edilen sonuçlar verilmiştir. K-NN için k sayısı 5 olarak seçilmiştir. DVM ise polinom çekirdeği kullanılarak eğitilmiştir. Verilerin eğitim ve test için sırasıyla %70-%30 bölünmesiyle en yüksek doğruluk %59,36, en yüksek kesinlik değeri %61 ve en yüksek F- skoru değeri %59 olarak DVM’den elde edilmiştir. Verilerin eğitim ve test için sırasıyla %80-%20 bölünmesiyle ise en yüksek doğruluk %61,58, en yüksek kesinlik değeri %61 ve en yüksek F- skoru değeri %62 olarak yine DVM’den elde edilmiştir. Şekil 6’da makine öğrenimi ve derin öğrenmeye dayalı sonuçların karşılaştırılması verilmiştir. Bu bilgiler ışığında tasarlanan ESA mimarisi özellik tabanlı makine öğrenmesi yöntemlerinden daha iyi performans göstermiştir.

Tablo 5. Makine öğrenimine dayalı sınıflandırma sonuçları

Model	Verinin Eğitim ve Test için Farklı oranlarda Bölünmesi	Doğruluk %	Kesinlik %	F-skoru %
K-NN	%70- %30	58,27	0,57	0,54
K-NN	%80- %20	59,24	0,58	0,56
DVM (Polinom)	%70- %30	59,36	0,61	0,59
DVM (Polinom)	%80- %20	61,58	0,61	0,62



Şekil 6. Derin Öğrenme ve Makine Öğrenmesinin Performans Karşılaştırılması

Bu çalışmada önerilen yöntemin performansını daha iyi değerlendirebilmek için, elde ettiğimiz bulgularla literatürde kullanılan diğer yöntemlerden elde edilen sonuçlar karşılaştırılmıştır. Tablo 6'da akciğer seslerinin sınıflandırılmasıyla ilgili önemli çalışmalar özetlenmektedir.

Tablo 6. ICBHI 2017 veri seti için Literatürdeki Çalışmaların Performans Karşılaştırılması

Metot	Kullanılan Özellikler	Sınıflandırıcı	Doğruluk %
Kochetov ve ark. (2018) [29]	MFCC	Tekrarlayan Sinir Ağları (RNN)	65,70
Demir ve ark. (2019) [1]	ESA modelinden Çıkarılan Derin Özellikler	DVM	65,50
Serbes ve ark. (2017) [30]	Kısa Süreli Fourier Dönüşümü + Dalgacık (Wavelet)	DVM	57,88
Chambres ve ark. (2018) [31]	Çoklu Özellikler	Karar Ağaçları	49,63
Önerilen Yöntem	Spektrogram	ESA	64,50

Tablo 6'de görüldüğü gibi önerilen yöntem, literatürde kullanılan diğer iki yöntemden daha iyi bir performansa göstermiştir. Bu bilgiler ışığında, önerilen mimari, akciğer ses sinyallerinin sınıflandırma görevi için uygun ve kabul edilebilir bir yöntemdir.

7. SONUÇLAR (CONCLUSION)

Bu çalışma, halk sağlığında en önemli konulardan biri olan akciğer hastalıklarının otomatik teşhisine odaklanmaktadır. Literatürde bu konuda birçok çalışma yapılmıştır, ancak bu çalışmaların çoğu geleneksel yöntemlerden oluşmaktadır. Bu çalışmada akciğer seslerinin tanınması probleminde sınıflandırma performansını artırmak için derin öğrenme ele alınmaktadır. Önerilen yöntem, literatürde popüler olan ICBHI 2017 veri seti üzerinde değerlendirilmiştir. İlk olarak akciğer ses sinyalleri ön işlemden geçirilerek spektrogramlar çıkarılmıştır. Spektrogram görüntülerine ek eğitim veri örnekleri oluşturmak için 5 farklı veri arttırma işlemi uygulanmıştır. 12 katmanlı ESA tasarlanarak ilk katmana spektrogram görüntüleri verilerek sınıflandırma işlemi yapılmıştır. Veri arttırma öncesi ortalama %60,14 sınıflandırma başarıımı elde edilmiştir. Veri arttırma sonrası ise ortalama %64,50 sınıflandırma başarıımı elde edilmiştir. Beklendiği gibi, veri arttırma işleminin uygulanmasından sonra doğruluk oranında arttığı deney sonuçlarından gözlenmektedir. Ayrıca önerilen yöntemin performansını diğer makine öğrenmesi yöntemleri ile

karşılaştırmak için veri seti örneklerinden MFCC özellikleri çıkarılarak DVM ve K-NN ile sınıflandırma işlemi yapılmıştır. MFCC özellikler ile en yüksek sınıflandırma başarımı %61,58 olarak DVM'den elde edilmiştir. Önerilen yöntem, özellik çıkarımına dayanan çalışmalarından daha iyi performans göstermiştir. İleriki çalışmalarda, tekrarlayan sinir ağlarının ESA ile kombinasyonlarını kullanması önerilmektedir.

KAYNAKLAR (REFERENCES)

- [1] F. Demir, A. Sengur, and V. Bajaj, "Convolutional neural networks based efficient approach for classification of lung diseases," *Heal. Inf. Sci. Syst.*, vol. 8, no. 1, p. 4, Dec. 2019, doi: 10.1007/s13755-019-0091-3.
- [2] S. İçer and Ş. Gengeç, "Classification and analysis of non-stationary characteristics of crackle and rhonchus lung adventitious sounds," *Digit. Signal Process.*, vol. 28, pp. 18–27, 2014, doi: 10.1016/j.dsp.2014.02.001.
- [3] R. Naves, B. H. G. Barbosa, and D. D. Ferreira, "Classification of lung sounds using higher-order statistics: A divide-and-conquer approach," *Comput. Methods Programs Biomed.*, vol. 129, pp. 12–20, 2016, doi: 10.1016/j.cmpb.2016.02.013.
- [4] G.-C. Chang and Y.-P. Cheng, "Investigation of noise effect on lung sound recognition," *2008 International Conference on Machine Learning and Cybernetics*. IEEE, 2008, doi: 10.1109/icmlc.2008.4620605.
- [5] S. Reichert, R. Gass, C. Brandt, and E. Andrès, "Analysis of respiratory sounds: state of the art," *Clin. Med. Circ. Respirat. Pulm. Med.*, vol. 2, pp. 45–58, May 2008, doi: 10.4137/ccrpm.s530.
- [6] A. Kandaswamy, C. S. Kumar, R. P. Ramanathan, S. Jayaraman, and N. Malmurugan, "Neural classification of lung sounds using wavelet coefficients," *Comput. Biol. Med.*, vol. 34, no. 6, pp. 523–537, 2004, doi: https://doi.org/10.1016/S0010-4825(03)00092-1.
- [7] S. ULUKAYA, G. SERBES, İ. ŞEN, and Y. P. KAHYA, "Akciğer Solunum Seslerinin Spektral Öznitelikler ile Sınıflandırılması," *Süleyman Demirel Üniversitesi Fen Bilim. Enstitüsü Derg.*, vol. 22, no. 2, p. 711, 2017, doi: 10.19113/sdufbed.84471.
- [8] A. Sovijärvi *et al.*, "Characteristic of breath sounds and adventitious respiratory sounds," *Charact. Breath Sounds Adventitious Respir. Sounds*, vol. 10, pp. 591–596, Jan. 2000.
- [9] G. Serbes, C. O. Sakar, Y. P. Kahya, and N. Aydin, "Pulmonary crackle detection using time–frequency and time–scale analysis," *Digit. Signal Process.*, vol. 23, no. 3, pp. 1012–1021, 2013, doi: https://doi.org/10.1016/j.dsp.2012.12.009.
- [10] F. Jin, F. Sattar, and D. Y. T. Goh, "New approaches for spectro-temporal feature extraction with applications to respiratory sound classification," *Neurocomputing*, vol. 123, pp. 362–371, 2014, doi: 10.1016/j.neucom.2013.07.033.
- [11] M. Bahoura, "Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes.," *Comput. Biol. Med.*, vol. 39, no. 9, pp. 824–843, Sep. 2009, doi: 10.1016/j.compbiomed.2009.06.011.
- [12] A. D. Orjuela-Cañón, D. F. Gómez-Cajas, and R. Jiménez-Moreno, "Artificial Neural Networks for Acoustic Lung Signals Classification," *Advanced Information Systems Engineering*. Springer Berlin Heidelberg, pp. 214–221, 2014, doi: 10.1007/978-3-319-12568-8_27.
- [13] S. O. Maruf, M. U. Azhar, S. G. Khawaja, and M. U. Akram, "Crackle separation and classification from normal Respiratory sounds using Gaussian Mixture Model," *2015 IEEE 10th International Conference on Industrial and Information Systems (ICIIS)*. IEEE, 2015, doi: 10.1109/iciinfs.2015.7399022.

- [14] M. Yeginer and Y. P. Kahya, "Elimination of vesicular sounds from pulmonary crackle waveforms," *Comput. Methods Programs Biomed.*, vol. 89, no. 1, pp. 1–13, 2008, doi: 10.1016/j.cmpb.2007.10.002.
- [15] G. Serbes, S. Ulukaya, and Y. P. Kahya, "An Automated Lung Sound Preprocessing and Classification System Based On Spectral Analysis Methods," *Precision Medicine Powered by pHealth and Connected Health*. Springer Singapore, pp. 45–49, 2017, doi: 10.1007/978-981-10-7419-6_8.
- [16] G. Chambres, P. Hanna, and M. Desainte-Catherine, "Automatic Detection of Patient with Respiratory Diseases Using Lung Sound Analysis," *2018 International Conference on Content-Based Multimedia Indexing (CBMI)*. IEEE, 2018, doi: 10.1109/cbmi.2018.8516489.
- [17] D. Bardou, K. Zhang, and S. M. Ahmad, "Lung sounds classification using convolutional neural networks," *Artif. Intell. Med.*, vol. 88, pp. 58–69, 2018, doi: https://doi.org/10.1016/j.artmed.2018.04.008.
- [18] C. Jácome, J. Ravn, E. Holsbø, J. C. Aviles-Solis, H. Melbye, and L. Ailo Bongo, "Convolutional Neural Network for Breathing Phase Detection in Lung Sounds," *Sensors (Basel)*, vol. 19, no. 8, p. 1798, Apr. 2019, doi: 10.3390/s19081798.
- [19] K. Kochetov, E. Putin, M. Balashov, A. Filchenkov, and A. Shalyto, "Noise Masking Recurrent Neural Network for Respiratory Sound Classification," *Artificial Neural Networks and Machine Learning – ICANN 2018*. Springer International Publishing, pp. 208–217, 2018, doi: 10.1007/978-3-030-01424-7_21.
- [20] Y. Fu and C. Aldrich, "Flotation froth image recognition with convolutional neural networks," *Miner. Eng.*, vol. 132, pp. 183–190, 2019, doi: 10.1016/j.mineng.2018.12.011.
- [21] A. GÜLCÜ and Z. KUŞ, "Konvolüsyonel Sinir Ağlarında Hiper-Parametre Optimizasyonu Yöntemlerinin İncelenmesi," *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji*, vol. 7. Gazi Üniversitesi, pp. 503–522, 2019, doi: 10.29109/gujsc.514483.
- [22] D. C. Cireşunđinedan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, "Flexible, High Performance Convolutional Neural Networks for Image Classification," in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Two*, 2011, pp. 1237–1242.
- [23] M. Bilal Er and I. B. Aydilek, "Music emotion recognition by using chroma spectrogram and deep visual features," *Int. J. Comput. Intell. Syst.*, vol. 12, no. 2, 2019, doi: 10.2991/ijcis.d.191216.001.
- [24] J. Salamon and J. P. Bello, "Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification," *IEEE Signal Process. Lett.*, vol. 24, no. 3, pp. 279–283, 2017, doi: 10.1109/LSP.2017.2657381.
- [25] "Support Vector Machines, 1992; Boser, Guyon, Vapnik," in *SpringerReference*, Springer-Verlag.
- [26] C. Li *et al.*, "Using the K-Nearest Neighbor Algorithm for the Classification of Lymph Node Metastasis in Gastric Cancer," *Comput. Math. Methods Med.*, vol. 2012, pp. 1–11, 2012, doi: 10.1155/2012/876545.
- [27] B. M. Rocha *et al.*, "An open access database for the evaluation of respiratory sound classification algorithms," *Physiol. Meas.*, vol. 40, no. 3, p. 35001, 2019, doi: 10.1088/1361-6579/ab03ea.
- [28] M. B. ER and H. ÇİĞ, "Türk Müziđi Uyarınları Kullanılarak İnsan Duygularının Makine Öğrenmesi Yöntemi İle Tanınması," *Gazi Üniversitesi Fen Bilim. Derg. Part C Tasarım ve Teknol.*, vol. 8, no. 2, pp. 458–474, 2020, doi: 10.29109/gujsc.687199.
- [29] K. Kochetov, E. Putin, M. Balashov, A. Filchenkov, and A. Shalyto, "Noise Masking Recurrent Neural Network for Respiratory Sound Classification," *Artificial Neural Networks and Machine*

- Learning – ICANN 2018*. Springer International Publishing, pp. 208–217, 2018, doi: 10.1007/978-3-030-01424-7_21.
- [30] G. Serbes, S. Ulukaya, and Y. P. Kahya, “An Automated Lung Sound Preprocessing and Classification System Based On Spectral Analysis Methods,” *Precision Medicine Powered by pHealth and Connected Health*. Springer Singapore, pp. 45–49, 2017, doi: 10.1007/978-981-10-7419-6_8.
- [31] G. Chambres, P. Hanna, and M. Desainte-Catherine, “Automatic Detection of Patient with Respiratory Diseases Using Lung Sound Analysis,” *2018 International Conference on Content-Based Multimedia Indexing (CBMI)*. IEEE, 2018, doi: 10.1109/cbmi.2018.8516489.