



Akciğer Hastalıklarının Dalgacık Katsayıları Kullanılarak Karar Ağaçlarına Dayalı Sınıflandırılması

Merve Çolak ^{1*}, Şerife Gengeç Benli ^{2*}, Müge Dolu ^{3*}

^{1*} Erciyes Üniversitesi, Mühendislik Fakültesi, Biyomedikal Mühendisliği Bölümü, Kayseri, Türkiye, (ORCID: 0000-0003-0840-3114), mervecolak.1938@gmail.com

^{2*} Erciyes Üniversitesi, Mühendislik Fakültesi, Biyomedikal Mühendisliği Bölümü, Kayseri, Türkiye, (ORCID: 0000-0002-5527-8574), serifegengec@erciyes.edu.tr

^{3*} (ORCID: 0000-0002-7886-3098), mugeedolu@gmail.com

(2nd International Conference on Access to Recent Advances in Engineering and Digitalization (ARACONF)-10–12 March 2021)

(DOI: 10.31590/ejosat.903249)

ATIF/REFERENCE: Çolak, M., Gengeç Benli, Ş. & Dolu, M. (2021). Akciğer Hastalıklarının Dalgacık Katsayıları Kullanılarak Karar Ağaçlarına Dayalı Sınıflandırılması. Avrupa Bilim ve Teknoloji Dergisi, (24), 463-468.

Öz

Günümüzde akciğer hastalıkları çok sık görülmektedir ve içerisinde bulunduğumuz Covid-19 salgını dolayısıyla da zatürre gibi bazı akciğer hastalıklarında yoğun artış yaşanmaktadır. Akciğer hastalıklarındaki ilk tanı oskültasyon yardımıyla konulmakta olup bu yöntem fiziksel muayene için ilk olarak tercih edilen düşük maliyetli ve etkili bir yöntemdir fakat oskültasyonda hastalık belirtileri elde edilse bile bu yöntem sesi bileşenlerine ayırıp karakterize etmez. Hastadaki sesleri dinleyen ve ön tanı koyan kişinin tecrübeli olması gerekmektedir buna rağmen ne kadar dikkatli olursa olsun insan kulağı bu sesleri her zaman eksiksiz duyacak kadar hassas bir organ değildir. Tanı sırasında doktora yardımcı olacak, küçük, hassas detayları kaçırmayacak ve sesi kolaylıkla karakterize edip ses bileşenlerin özelliklerini ve hangi hastalık grubunda yer aldığını tanımlayabilecek bir sistem gerekmektedir.

Bu çalışmada elektronik stetoskop kullanılarak kronik obstrüktif akciğer hastalığı, üst solunum yolu enfeksiyonu, pnömoni ve sağlıklı kişilerden oluşan dört gruba ait akciğer sesleri Chebyshev filtre ile filtrelenmiş ve dalgacık dönüşümü, kısa zamanlı Fourier dönüşümü, güç spektral yoğunluğu kullanılarak elde edilen öznitelikler değerlendirilmiştir. Ayrıca literatürde daha çok tercih edilen dalgacık dönüşümü katsayıları kullanılarak Waikato Environment for Knowledge Analysis (WEKA) programı üzerinden karar ağaçları ile sınıflandırılma yapılmıştır. Böylelikle akciğer seslerinin çeşitli sinyal işleme metodları ile karakteristik özelliklerinin elde edilmesi ve sonrasındaki sınıflandırma ile %95-%99 değerleri arasındaki yüksek doğruluk oranıyla hastalık teşhisi yapılabileceği görülmüştür.

Anahtar Kelimeler: Akciğer sesleri, Dalgacık dönüşümü katsayıları, Karar ağaçları

Classification of Lung Diseases Based on Decision Trees Using Wavelet Coefficients

Abstract

Today, lung diseases are very common and due to the Covid-19 epidemic we are in, there is an intense increase in some lung diseases such as pneumonia. The first diagnosis in lung diseases is made with the help of auscultation, and this method is the first preferred low-cost and effective method for physical examination, but even if disease symptoms are obtained in auscultation, this method does not separate and characterize the voice into its components. The person who listens to the sounds in the patient and makes a preliminary diagnosis must be experienced, however, no matter how carefully the human ear is not a sensitive organ that can always hear these sounds completely. A system is required to assist the doctor during the diagnosis, not to miss small, delicate details, and to easily characterize the sound and define the properties of the sound components and in which disease group it is involved.

In this study, lung sounds belonging to four groups of chronic obstructive pulmonary disease, upper respiratory tract infection, pneumonia and healthy individuals were filtered with Chebyshev filter using electronic stethoscope and features obtained using wavelet transform, short-time Fourier transform, power spectral density were evaluated. Waikato Environment for Knowledge Analysis (WEKA) program was used to classify decision trees using wavelet transform coefficients, which are more preferred in the literature. Thus, it has been observed that the characteristic features of the lung sounds can be obtained with various signal processing methods and the diagnosis of the disease can be made with a high accuracy rate between 95% and 99% with the subsequent classification.

Keywords: Lung sounds, Wavelet coefficients, Decision trees.

* Sorumlu Yazar: serifegengec@erciyes.edu.tr

1. Giriş

Akciğer sesleri oskültasyon yöntemi ile elde edilmekte olup bu yöntemle elde edilen sesler hastalık teşhisi için çok önemlidir. Oskültasyon, stetoskobun cilt üzerine yerleştirilmesi ile hava akımının bronşlar ve alveollerde oluşturduğu titreşimin dinlenmesine dayalı bir muayene tekniğidir. Oskültasyon yöntemi ile akciğer seslerinin değerlendirilmesi sessiz, yeterince aydınlık bir ortamda ve oda ısısında yapılmalıdır. Stetoskobun diyafram kısmı cild ile her noktadan tam temas halindeyken belli bir sıra takip edilerek akciğerlerin dinlendiği tüm noktalar için en az bir inspiryum ve ekspiryumuna içine alacak şekilde simetrik olarak değerlendirilmelidir [1]. Bu işlem sessiz bir ortamda, stetoskop tam yerleştirilerek ve simetrik olarak uygulanmalıdır.

Akciğer sesleri 60 ve 1500 Hz aralığında bulunan seslerdir [2]. Analog stetoskoplarda elde edilen bilgi, hastadan hastaya ve sesin alındığı bölgeye göre farklılık göstermesi ve sesleri dinleyen personelin stres, yorgunluk, tecrübesizlik gibi durumlarda olmasından kaynaklı olarak yanlış teşhise neden olabilmektedir. Aynı zamanda analog stetoskoplarda dinlenen ses verileri kayıt altına alınamamakta ve 120 Hz üzerindeki ses frekansları zayıflatıldığı için bu aralıktaki ses bileşenleri kaybolmaktadır [3]. Elektronik stetoskopta ise veriler kaydedilebilmekte ve bilgisayar üzerinden görüntülenebilmektedir.

Akciğer sesleri normal ve ek sesler olarak ikiye ayrılmaktadır. Normal sesler akciğerlerde hiçbir solunum problemi olmadığında soluk alıp verme sırasında havanın göğüsten içeriye ve dışarıya hareketiyle dinlenen seslerken ek sesler normalde var olmayan patolojik durumlarda görülen seslerdir [4]. Çalışmada kronik obstrüktif akciğer hastalığı (KOA), üst solunum yolu enfeksiyonu, pnömoni ve sağlıklı kişilerden oluşan 4 gruba ait akciğer sesleri kullanılmıştır. KOAH, akciğerlerde alınan havanın kolaylıkla dışarı verilememesi olarak tanımlanırken sakatlık ve ölüme neden olabilecek bir hastalıktır. Üst solunum yolu enfeksiyonu ise genellikle üşüme ve öksürük, boğaz ağrısı, burun akıntısı, burun tıkanıklığı, balgam, baş ağrısı, ateş, kulak ağrısı gibi semptomlara yol açan bir akciğer hastalığıdır. Pnömoni ise zatürre olarak bilinen koronavirüs geçiren kişilerde de görülebilen ağır bir hastalıktır. Akciğerdeki hava keseciklerinin iltihaplı bir sıvı ile dolmasıdır. Bu hastalıklar insan yaşam kalitesini birebir ilgilendirmekle birlikte teşhisinde otomatik sistemlerin geliştirilmesi önem arz etmektedir.

Bu çalışmada yukarıda bahsedilen 4 gruptan alınan sesler öncelikle Chebyshev filtre ile filtrelenmiş ve dalgacık dönüşümü, kısa zamanlı Fourier dönüşümü, güç spektral yoğunluğu kullanılarak elde edilen öznitelikler değerlendirilmiş ve literatürde de daha çok tercih edilen dalgacık dönüşümü katsayıları kullanılarak WEKA programı üzerinden karar ağaçlarının farklı algoritmaları ile sınıflandırılma yapılmıştır.

1.1 Literatür Taraması

Literatürdeki çalışmalar incelendiğinde KOAH hastalığında wheeze, pnömoni hastalığında ral ek seslerinin olduğu görülmektedir [5]. Yapılan çalışmada sesler karakteristik olarak sınıflandırılmıştır. Alınan verilerde her bir ses 15 saniye süresince olup 9600 örneğe karşılık gelmektedir. Wheeze, raller ve normal solunum sesleri için her biri 200 bölüt olmak üzere 600 bölüt ses

kullanılmıştır. Alınan seslerde öznitelik çıkarımı yapılmış ve güç spektral yoğunluğu, Mel frekansı kepsstral sayıları, algısal doğrusal öngörü yöntemleri için k en yakın komşu ve destek vektör makineleri ile sınıflandırılarak doğruluk yüzdelere ulaşılmıştır. Sonuçlar tek tek incelendiğinde en iyi sonuç algısal doğrusal öngörü yöntemleri olsa da tüm sınıfların ortalama değerine bakıldığında en iyi doğruluk değeri mel frekansı kepsstral sayılarında elde edilmiştir. Destek vektör makineleri birkaç durumda k en yakın komşu sınıflandırmasından daha düşük doğruluk oranı verse de elde ettiği sonuçlar ortalama hem daha iyi hem de daha kararlıdır [5].

Literatürdeki diğer bir çalışmada ise sağlıklı ve sağlıklı olmayan akciğer sesleri dalgacık katsayıları kullanılarak filtrelenmiştir. Çalışmada RALE'den alınan 14 sağlıklı, 20 sağlıksız olmak üzere 34 akciğer sesi kullanılmıştır. Verilerdeki seslerin süresi tek bir nefes alış-veriş döngüsü şeklinde verilirken örnekleme frekansı ise 11025 olarak belirlenmiştir. Yüksek bileşendeki gürültüyü yok etmek için 1,5 kHz kesim frekansına sahip 4. dereceden alçak geçiren filtre kullanılmıştır. Deri sürtünmesi, iç organ seslerini ortadan kaldırmak için ise 4. dereceden yüksek geçiren filtre ile bu sesler filtrelenmiştir. Çok katmanlı yapay sinir ağı kullanarak sınıflandırma yapılmış; verilerin %70'i eğitim, %30'u ise test kümesi olarak kullanılıp yapılan sınıflandırma sonucu doğruluk oranı %100 bulunmuştur. Destek vektör makinelerinde de çok katmanlı yapay sinir ağına olduğu gibi verilerin %70'i eğitim, %30'u test kümesi olarak kullanılması sonucu %100 doğruluk oranı elde edilmiştir [6].

Orjuela ve ark. (2014), wheeze, raller ve normal akciğer seslerini sınıflandırmak için yapay sinir ağlarıyla birlikte Mel frekansı kepsstral katsayılarını kullanmışlardır. Sınıflandırmanın doğruluk değeri raller için %70, wheeze için %100 ve normal seslerde %80 olarak elde edilmiştir [7]. Chambres ve ark. (2018), bir karar ağacı modeli ve daha fazla ses bileşenine sahip bir makine öğrenimi yaklaşımı önermişlerdir. Veri seti olarak ICBHI 2017 kullanılmış ve deney sonucu %50 doğruluk oranı elde edilmiştir [8]. İçer ve Gengeç (2014), sağlıklı ve sağlıksız akciğer seslerinde güç spektral yoğunluk değerlerinin frekans oranını ve Hilbert-Huang Dönüşümü özelliklerini kullanarak destek vektör makineleri ile sınıflandırarak %90'ın üzerinde doğruluk elde etmişlerdir [9]. Serbes ve ark. (2013), zaman frekansı ve zaman ölçeği analiziyle özellik çıkarmışlardır. Raller ve raller olmayan sesleri sınıflandırmak için k-en yakın komşu, destek vektör makineleri ve yapay sinir ağları kullanılmıştır. En iyi sınıflandırma sonucu %97,5 olarak destek vektör makinelerinde elde edilmiştir [10]. Jin ve Sattar (2014), solunum seslerini sınıflandırmak için destek vektör makineleri kullanmışlardır. Sinyallerden basıklık, entropi gibi özellikler çıkarılarak %97,7 ile %98,8 arasında doğruluk değeri rapor etmişlerdir [11]. Bahoura (2009), iki sınıftan oluşan akciğer solunum seslerinin sınıflandırılmasında sinyallerden Mel frekansı kepsstral katsayıları ile özellik çıkarıp Gauss karışım modeli ile sınıflandırarak %94,2 doğruluk değeri rapor etmişlerdir [12].

2. Materyal ve Metot

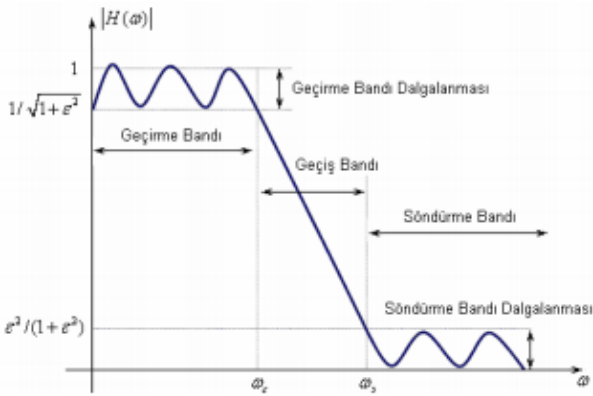
2.1 Verilerin Alınması

Bu çalışmada analiz edilen veriler ICBHI Challenge sitesinde yayınlanan hazır veri setinden [13] alınmıştır. Çalışmada kullanılmak istenen hastalık durumlarına sahip 6 sağlıklı ve 6

pnömoni, 6 üst solunum yolu enfeksiyonu ile 6 KOAH hastalığı taşıyan kişilerden elde edilen 20 saniye uzunluğuna sahip 24 ses verisi 4'e bölünerek sesin alt çeyrek, üst çeyrek, aralık ve ortalamasına göre 4 adet veri setine bölünmüştür. Böylelikle 24 kişiden oluşan gruba ait 96 ses sinyali içeren veri seti elde edilmiştir. Veri setimizde 12 kadın 12 erkek bulunmaktadır. Bu çalışmada bahsedilen verilere uygulanan metodlar aşağıdaki bölümlerde sunulmaktadır.

2.2 Chebyshev Filtre

Chebyshev filtreleri bir çeşit yüksek-Q filtreleridir. Bu filtreler; söndürme bandında dik iniş istenildiği zaman, geçiş bandının düz olmasının gerekli olmadığı durumlarda kullanılır. Bu filtre cevabında, geçiş bandı dalgalanmasına izin verilir. Butterworth cevabına oranla söndürme bandındaki başlangıç inişleri daha keskindir [14]. Bu karşılaştırma Şekil 1'de yer almaktadır. Şekilde eğriler n=3 derecesindeki filtreler içindir. Chebyshev filtresi, geçiş bandında 3 dB'lik dalgalanma yapar. Butterworth filtresinden 10 dB kadar söndürme bandında daha fazla zayıflama yapar [15].



Şekil 1: Chebyshev Filtre Karakteristiği [14]

Chebyshev filtresi için zayıflama;

$$\alpha_n = 10 \log[1 + \varepsilon^2 + C_n^2(w)] \text{ dB}$$

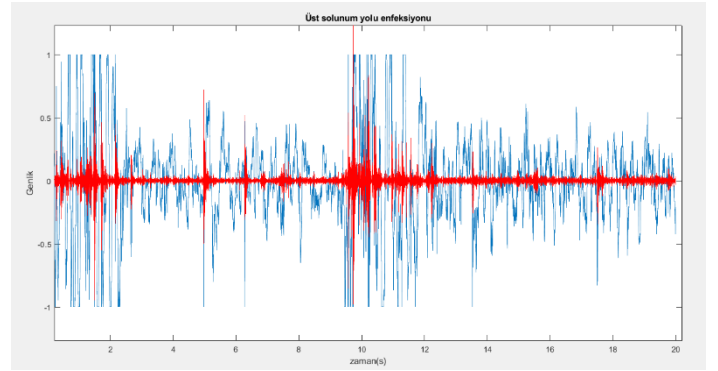
n. dereceden Chebyshev polinomu;

$$C_n(w) = \cosh n \cosh^{-1}(w)$$

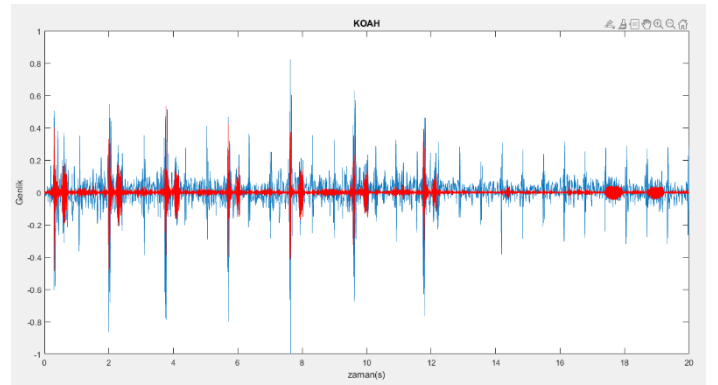
$$|H(jw)| = \frac{1}{\sqrt{1 + \varepsilon^2 C_n^2(w)}}$$

Akciğer sesleri 60 Hz-1,5 kHz aralığındadır buna karşın stetoskoplardan alınan verilerde başta kalp sesleri, iç organ sesleri, deri sürtünmeleri ve kas sesleri etkin bir gürültü oluşturmaktadır [1]. Bu çalışmada sesi gürültüden arındırmak için 60-1500 Hz arasındaki frekanslarda bulunan sesin kullanılması kalan gürültülerin ise filtrenmesi gerekliliğinden 9. Dereceden 60Hz ve 1,5kHz kesim frekanslarına sahip bir Chebyshev band geçiren filtre tasarlanmıştır.

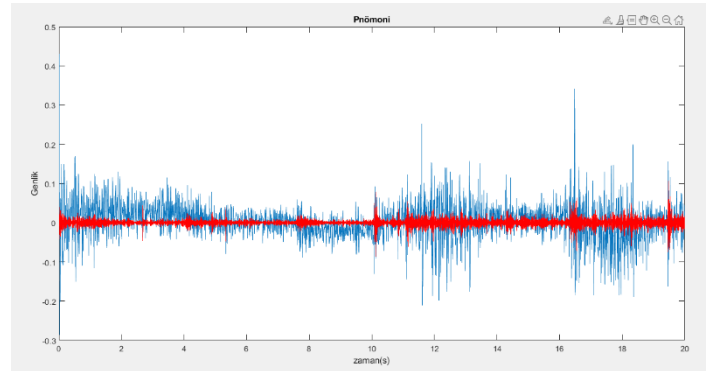
Aşağıdaki şekillerde (Şekil-2, 3, 4, 5) seslerin ait olduğu hastalıkların filtresiz (mavi) ve filtreli (kırmızı) sinyalleri sunulmuş olup uygulanan filtrenin yüksek oranda gürültüyü elimine etmekteki başarısı görülmektedir.



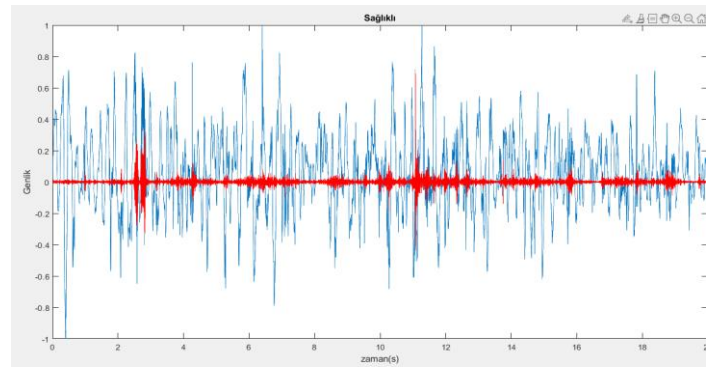
Şekil 2: Üst Solunum Yolu Enfeksiyonu Hastalığı Akciğer Ses Sinyali



Şekil 3: KOAH Hastalığı Akciğer Ses Sinyali



Şekil 4: Pnömoni Hastalığı Akciğer Ses Sinyali



Şekil 5: Sağlıklı Akciğer Ses Sinyali

2.3 Öznitelik Çıkarma

Tüm kişilere ait akciğer sesleri filtrelendikten sonra kısa zamanlı Fourier dönüşümü alınmıştır. Bu dönüşümdeki amaç, durağan olmayan işareti durağan olan küçük zaman parçalarına ayırmak ve bu zaman parçalarındaki frekansları belirlemek için bahsedilen aralıkların Fourier analizini yapmaktır [16]. Kısa zamanlı Fourier dönüşümü zaman-frekans eksenlerine ayrılarak kullanılmıştır. Daha sonra güç spektral yoğunluğu alınmıştır ve bu metod elimizdeki ses verisinin enerjisinin frekansla nasıl dağıtıldığını açıklar. Son olarak ise dalgacık dönüşümü uygulanmıştır. Dalgacık dönüşümü solunum sesindeki gibi durağan olmayan yani dinamik sinyallerin analizinde kullanılan bir yöntemdir.

Dalgacık analizi düşük frekans bilgisinin daha önemli olduğu durumlar için büyük zaman aralıklarının, yüksek frekans bilgisinin daha önemli olduğu durumlar için de daha küçük zaman aralıklarının kullanımına izin veren değişik boyutlarda bölgelere sahip bir pencereleme tekniğidir [17]. Bu yöntemde yaklaşım katsayıları (cA) ve detay katsayıları (cD) olmak üzere iki kısım elde edilmiştir. Yaklaşım katsayıları (cA), sinyalin yüksek ölçekli alçak frekans bileşenlerini ve detay katsayıları (cD) ise sinyalin düşük ölçekli yüksek frekans bileşenlerini göstermektedir [18].

Dinamik sinyallerde Fourier dönüşümü sinyal ile ilgili istenilen bilgiyi sağlayamaz. Fourier dönüşümü sinyalin sadece frekans bileşenlerini gösterse de zaman eksenindeki bileşenler hakkında bilgi vermemektedir.

Dinamik sinyallerde frekans bileşeni zamanda değişkenlik gösterir bu yüzden zaman-frekans bilgisi gereklidir. Dalgacık dönüşümü ise sinyalin süre ve frekanstaki bileşenlerini aynı anda gösterebildiğinden bu yöntemle sınıflandırma yapılmıştır.

2.3.1 Fourier Dönüşümü

Bir sinyalin fonksiyonları zamanla değişmiyorsa, bunlara "durağan sinyal" denir. Durağan bir sinyalde de beklenmeyen olaylar görülebilir, ancak bu olayların olasılığı istatistik olarak tahmin edilebilir bilgilerdir. Durağan sinyalleri incelemek için Fourier dönüşümü kullanılır. Fourier dönüşümü gürültüye karşı olan dayanıklılıklığı ve öteleme bağımsızlığı sebebiyle çalışmalarda sıklıkla tercih edilmektedir.

Fourier dönüşümü ile bu sinyaller, sinüs ve kosinüs dalgalarının lineer birleşime ayrışır [19]. Herhangi bir $f(t)$ fonksiyonunun Fourier dönüşümü,

$$X(f) = \int_{-\infty}^{\infty} f(t)e^{-i2\pi ft} dt$$

formülü ile ifade edilir.

2.3.2 Güç Spektral Yoğunluğu

Solunum sesi penceresi bir zaman serisi $x(k)$ olarak düşünülürse, güç spektrum yoğunluğu aşağıda belirtilen şekilde hesaplanır;

$$P(f) = \frac{1}{K} |X(f)|^2$$

$i=0, \dots, M-1$ olmak üzere, $X(f)$, $x(k)$ nın Fourier dönüşümüne karşılık gelmektedir. Welch yöntemi güç spektrum yoğunluğunun $w(k)$ pencereleri ile çarpılmasına ve veri bölütlerinin örtüşmesine izin verdiğinden dolayı kullanılabilirlik sağlamaktadır [20]. Literatürde durağan sinyallerde çoğunlukla Fourier dönüşümü ve güç spektral yoğunluğu, dinamik sinyallerde ise dalgacık dönüşümü metoduna bağlı özellikler kullanılarak analiz ve sınıflandırma tercih edilmiştir [6].

2.3.3 Dalgacık Dönüşümü

Dalgacık dönüşümü akciğer ses sinyalleri gibi durağan olmayan sinyallerin analizinde kullanılacak olan yararlı bir araçtır. Bu durum dalgacık dönüşümünün farklı ölçeklerde yeniden çözüm özelliğinden kaynaklanır. Sinyaller çoğu durum için frekans tanım alanında analiz edilirler. Fourier dönüşümü bir sinyalin ihtiva ettiği frekansları gösterir fakat hangi zaman değerlerinde hangi frekansların mevcut olduğunu bildirmez.

Dalgacık analizi düşük frekans bilgisinin daha değerli olduğu durumlar için büyük zaman aralıklarının, yüksek frekans bilgisinin daha değerli olduğu durumlar için de daha küçük zaman aralıklarının kullanımına izin veren değişik boyutlarda bölgelere sahip bir pencereleme tekniğidir [6]. Bu özelliğinden dolayı akciğer sesi gibi dinamik özellik taşıyan sinyallerin analizinde yüksek başarı göstermektedir.

3. Araştırma Sonuçları ve Tartışma

3.1 Karar Ağacı (Decision Tree) ile Sınıflandırma

Veri madenciliği tekniklerinden biri olan karar ağaçları, sınıflandırma ve tahmin yapmak için kullanılmaktadır. Yapay sinir ağları, destek vektör makineleri gibi sınıflandırmalara göre daha kolay anlaşılabilir bir sınıflandırma yöntemidir. Karar ağacında kullanılan veri seti, öğrenme ve sınıflandırma olmak üzere iki bölüme ayrılıp öğrenmeden sonra veriler sınıflandırılarak karar ağacı şeklinde gösterilir. Sınıflandırma aşamasındaysa test verisi sınıflandırmanın doğruluğunu ve sınıflandırmadaki kuralları belirlemek için kullanılır [21]. Bulunan doğruluk istenilen seviyede ise bu kurallar sonraki verilerin de sınıflandırılması için kullanılır.

Bu çalışmada akciğer seslerinin dalgacık katsayılarına karar ağacı algoritmalarından olan random tree, reptree, random forest ve m5p uygulanmıştır.

Random Forest algoritmasında eğitsel verilerdeki örneklerin rastgele bir şekilde seçimi yapılır ve oluşturulan budanmamış ham sınıflandırma ve regresyon ağaçlarından oluşan modellerdir. Özellikle büyük miktarda verilerin sınıflandırılmasında iyi performans göstermektedir. Her bir bağın parçalanmasında kullanılacak özelliklerin rastgele seçimi, algoritmanın Adaboost ile yarışacak sonuçlar doğurmasına ve gürültülü değerlere karşı daha dayanıklı olmasına yol açmaktadır [22].

RandomTree algoritması sonucunda oluşan ağaç, olası ağaç kümeleri içinden rastgele olarak seçilmekte olup bu ağaçlar düzensiz dağılım göstermektedir. Bu rastgele ağaçlar, etkili bir şekilde oluşturulmakta ve birçok rastgele oluşturulmuş ağaç modeller genellikle yüksek doğruluk oranına sahiptir [23].

Reptree algoritması en hızlı karar ağaç sınıflandırma algoritmalarından biridir. Bu algoritma regresyon ya da karar ağaçlarının oluşmasında bilgi kazancı ölçütünü kullanır ve oluşan bu ağaç, azaltılmış hata budaması metodolojisine dayalı olarak budama işlemine tabi tutulmaktadır. Bu algoritmada, yalnızca sayısal niteliklerin sıralanması işlemi söz konusudur. M5P ise son düğümlerin sürekli sayısal öznelikler üretebilen doğrusal regresyon fonksiyonlarının olduğu ikili bir regresyon ağacı modelidir. Lineer regresyon modellerini birleştirmektedir [24].

Akciğer seslerinin filtrelenmiş sinyaline uygulanan dalgacık dönüşümlerinden elde edilen katsayılara karar ağaçlarının dört farklı algoritması uygulanarak elde edilen yüksek başarılı sınıflandırma sonuçları aşağıda sunulmaktadır.

4. Sonuç

Bu çalışmada WEKA programı ile karar ağaçlarının Random Forest, Random Tree, M5P ve Reptree algoritmaları kullanılarak sınıflandırılma yapılmıştır.

Yaklaşım katsayıları(cA) ve detay katsayılarına(cD) ayrı sınıflandırma yapılmış, sınıflandırıcıları eğitmek için %75 eğitim, %25 öğrenme metodu ile 10 kat çapraz doğrulama ve birini dışarıda çapraz doğrulama (leave-one-out) metodu kullanılmıştır ve iki yöntemin sonuçlarının yakın olduğu gözlenmiştir. 10 Kat çapraz doğrulama ve birini dışarıda çapraz doğrulama sınıflandırma metodu ile yaklaşım katsayıları ve detay katsayıları kullanılarak yapılan sınıflandırma işlemlerinin doğruluk değerleri sırasıyla Tablo 1 ve Tablo 2’de verilmiştir.

Tablo 1: Yaklaşım Katsayıları(cA) Sınıflandırmasında Doğruluk Değerleri

| Sınıflandırma Algoritmaları | Doğruluk Değeri |
|-----------------------------|-----------------|
| Reptree | %98 |
| Random Tree | %94 |
| Random Forest | %97 |
| M5P | %99 |

Tablo 2: Detay Katsayıları(cD) Sınıflandırmasında Doğruluk Değerleri

| Sınıflandırma Algoritmaları | Doğruluk Değeri |
|-----------------------------|-----------------|
| Reptree | %95 |
| Random Tree | %91 |
| Random Forest | %95 |
| M5P | %94 |

Elde edilen sonuçlara göre yaklaşım katsayılarında M5P algoritması daha yüksek doğruluk değeri verirken detay katsayılarında Reptree ve Random Forest algoritmaları en yüksek doğruluk değerini vermiştir. Her iki sınıflandırma için de en düşük doğruluk değeri Random Forest algoritmasındadır. Bu çalışma sonucunda akciğerdeki solunum sesleri elektronik stetoskop ile alınıp bilgisayar ortamında işlendiğinde akciğer seslerinin sınıflandırmasının yüksek doğrulukta olduğu böylelikle hastalık tayininin başarılı bir şekilde yapılabileceği görülmektedir.

5. Kaynakça

- [1] Aydemir, Y.(2011). Türkiye Solunum Araştırmaları Derneği, Göğüs Hastalıkları Erişim adresi: <https://www.solunum.org.tr/TusadData/Book/853/1412021102014-1072020165943bolum02.pdf>.
- [2] Reichert, S., Gass, R., Brandt, C., ve Andrès, E. (2008), Analysis of Respiratory Sounds: State of the Art, Clin Med Circ Respirat Pulm Med. 2008; 2: 45–58. doi: 10.4137/ccrpm.s530.
- [3] Gavriely, N., Cugell, D. W. (1995). Breath sounds methodology, CRC Press.
- [4] Şengül Emeksiz, Z, Bostancı, İ. (2018). Akciğer: Sesimi Duyan Var Mı?. Güncel Pediatri, 16 (3), 79-84. Erişim adresi: <https://dergipark.org.tr/tr/pub/pediatri/issue/40492/485522>.
- [5] Serbes, G, Ulukaya, S, Şen, İ, Kahya, Y. (2018). Akciğer Solunum Seslerinin Spektral Öznelikler ile Sınıflandırılması. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 22 (2), 711-716. Erişim adresi: <https://dergipark.org.tr/tr/pub/sdufenbed/issue/38975/456380>.
- [6] Uysal, S., Uysal, H., Bolat, B., & Yıldırım, T., (2013). Sağlıklı ve Sağlıksız Akciğer Seslerinin Dalgacık Katsayıları Kullanılarak Sınıflandırılması. IEEE 22. Sinyal İşleme ve Uygulamaları Kurultayı, Turkey.
- [7] Orjuela-Cañón, A. D., Gómez-Cajas, D. F., ve Jiménez-Moreno, R. (2014), “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.
- [8] Chambres, G., Hanna, P., ve Desainte-Catherine, M. (2018), “Sound Analysis,” 2018 International Conference on ContentBased Multimedia Indexing (CBMI). IEEE, 2018, doi: 10.1109/cbmi.2018.8516489.
- [9] İçer, S., ve Gengeç, Ş. (2014), “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.
- [10] Serbes, G., Sakar, C. O., Kahya, Y. P., ve Aydin, N. (2013), “Pulmonary crackle detection using time–frequency and time–scale analysis,” Digit. Signal Process., vol. 23, no. 3, pp. 1012–1021, 2013. Doi:10.1016/j.dsp.2012.12.009.
- [11] Jin, F., Sattar, F., ve Goh, D. Y. T. (2014), “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.
- [12] Bahoura, M. (2009), “Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes.” Comput. Biol. Med., vol. 39, no. 9, pp. 824–843, Eylül 2009, doi: 10.1016/j.combiomed.2009.06.011.classes.
- [13] ICBHI 2017 Challenge Respiratory Sound Database Erişim adresi: https://bhichallenge.med.auth.gr/ICBHI_2017_Challenge.
- [14] Çekiç, Y. (2004), Durağan Olmayan İşaretler İçin Zaman-Frekans Analizi (Doktora tezi), İstanbul Üniversitesi, Fen Bilimleri Enstitüsü, Elektronik Mühendisliği Ana Bilim Dalı
- [15] Valkenburg, M.E.V(1982)., Analog Filter Design: Cbs College College Publishing Company, New York, 1982.
- [16] Çekiç, Y. (2004), Durağan Olmayan İşaretler İçin Zaman-Frekans Analizi (Doktora tezi), İstanbul Üniversitesi, Fen Bilimleri Enstitüsü, Elektronik Mühendisliği Ana Bilim Dalı

- [17] Shi Y, Li Y, Cai M, Zhang XD. (2019), A Lung Sound Category Recognition Method Based on Wavelet Decomposition and BP Neural Network. *Int J Biol Sci* 2019; 15(1):195-207. doi:10.7150/ijbs.29863. Erişim adresi: <https://www.ijbs.com/v15p0195.html>.
- [18] Tepe, C., Sezgin, H., EKG Sinyallerinde Gürültü Gidermede Ayrık Dalgacık Dönüşümünde Farklı Ana Dalgacıkların ve Ayrıştırma Seviyelerinin Karşılaştırılması, Elektrik Elektronik Mühendisliği Bölümü, Ondokuz Mayıs Üniversitesi, Samsun. Erişim adresi: https://www.emo.org.tr/ekler/48132c7802e7bad_ek.pdf.
- [19] Ayaz, E. (1997), “Dalgacıklar ve Elektrik Mühendisliğindeki Uygulamaları”, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enstitüsü, 53s, İstanbul.
- [20] Welch, P. D. (1967), The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, *IEEE Transactions on Audio and Electroacoustics* 15, 2(1967), 70-73.
- [21] Breiman, L. (2001), Random Forests, *Machine Learning*, 45 (1): 5-32.
- [22] Priyama, A., Abhijeeta, Gupta, R., Ratheeb, A., ve Srivastava, S. (2013), Comparative Analysis of Decision Tree Classification Algorithms, *International Journal of Current Engineering and Technology*, Vol.3, No.2
- [23] Nor Haizan, W., Mohamed, W. (2012), A Comparative Study of Reduced Error Pruning Method in Decision Tree Algorithms, 2012 IEEE International Conference on Control System, Computing and Engineering.
- [24] Chaitanya Kumar, S., Deepak Chowdary, E., Venkatramaphanikumar S, Krishna Kishore, K.V. (2016), M5P Model Tree in Predicting Student Performance: A Case Study, 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT).