

Akciğer Solunum Seslerinin Spektral Öznitelikler ile Sınıflandırılması

Sezer ULUKAYA^{1,2}, Gökem SERBES^{*3}, İpek ŞEN⁴, Yasemin P. KAHYA¹

¹Boğaziçi Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, 34342, İstanbul

²Trakya Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, 22030, Edirne

³Yıldız Teknik Üniversitesi, Elektrik-Elektronik Fakültesi, Biyomedikal Mühendisliği Bölümü, 34220, İstanbul

⁴Beykent Üniversitesi, Mühendislik-Mimarlık Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, 34398, İstanbul

(Alınış / Received: 27.01.2017, Kabul / Accepted: 17.08.2017, Online Yayınlanma / Published Online: 01.11.2017)

Anahtar Kelimeler

Akciğer solunum sesleri,
Çırtırtı,
Üfürüm,
Mel frekans kepsral katsayıları,
Güç spektrum yoğunluğu,
Algısal doğrusal öngörü katsayıları

Özet: Bu çalışmada güç spektrum yoğunluğu, mel frekans kepsral katsayıları (MFKK) ve algısal doğrusal öngörü (ADÖ) yöntemleri; çırtırtı, üfürüm ve normal akciğer solunum seslerini ayırtmak amacıyla kullanılan öznitelik çıkarıcılar olarak görevlendirilmiştir. Ham özniteliklerden sekiz alt öznitelik kümesi (enerji, entropi, en küçükten, en büyüktene, ortalama, standart sapma, eğrilik ve basıklık) çıkarılarak k -en yakın komşu ve destek vektör makineleri sınıflandırıcılarına birini dışarıda bırak şeması kullanılarak beslenmiştir. Önerilen algısal doğrusal öngörü katsayıları yöntemi güç spektrum yoğunluğundan daha iyi performans sergilerken mel frekans kepsral katsayıları ile başa baş performans göstermiştir. ADÖ yönteminin üç gruplu sınıflandırma performansı var olan literatürle karşılaştırılmıştır. Çırtırtı, üfürüm ve normal sınıfları (% 94, % 95.5, % 95.5 sırasıyla) için en iyi sonuçlara ADÖ tarafından ulaşılmıştır. Diğer taraftan tüm sınıf doğruluklarının en iyi ortalama sonucuna % 91.3 ile MFKK tarafından ulaşılmıştır. MFKK ve ADÖ yöntemlerinin sınıflandırma doğruluğunun model derecesine oldukça bağlı olduğu gözlemlenmiştir.

Classification of Respiratory Sounds Using Spectral Features

Keywords

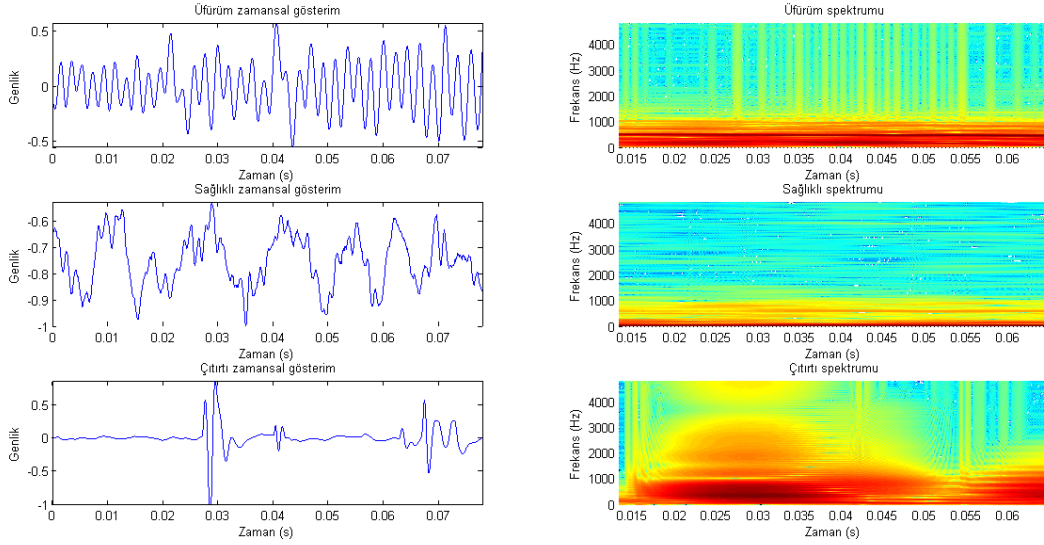
Respiratory sounds,
Crackle,
Wheeze,
Mel frequency cepstral coefficients,
Power spectral density,
Perceptual linear prediction coefficients

Abstract: In this study, power spectral density (PSD), mel frequency cepstral coefficients (MFCC) and perceptual linear prediction (PLP) methods are employed as feature extraction methods for discriminating crackle, wheeze and normal respiratory sounds. Eight sub-feature subsets (energy, entropy, minimum, maximum, mean, standard deviation, skewness and kurtosis) that are extracted from raw features are fed into k -nearest neighbour and support vector machine classifiers in a leave-one-out cross validation scheme. It is seen that proposed method PLP shows better performance than PSD and shows break-even performance with MFCC. Three class classification performance of PLP method is compared with existing literature. Best result for crackle, wheeze and normal classes (94 %, 95.5 %, 95.5 % respectively) is reached by PLP. On the other hand, the best average accuracy is represented by MFCC with a 91.3 % value. It has been observed that the classification accuracy of MFCC and PLP is highly dependent on the model order.

1. Giriş

Analog stetoskop 120 Hz. üzerindeki frekansları sönmölendirmekte ve bu frekans aralığındaki bilginin kaybolmasına sebep olmaktadır [1]. Ayrıca, stetoskop vasıtasıyla dinlenen sesler kayıt altına

alınmamaktadır. Bu sebeplerden dolayı, günümüzde yerini disiplinler arası çalışmaların da katkısıyla elektronik stetoskoba [2] bırakmaya başlamıştır. Akciğer solunum sesleri normal (vesicular) ve ek-ses (adventitious) olmak üzere ikiye ayrılırlar. Normal sesler sağlıklı insanlarda göğüs kafesinden dinlenen



Şekil 1. Birer adet üfürüm (en üst), sağlıklı (orta) ve çıtırtı (en alt) işaretinin zaman (sol) ve frekans (sağ) alanı gösterimi

ve akciğerdeki hava akışı ile eş zamanlı duyulan seslerdir. Sağlıklı insanlarda normal sesler 200-600 Hz frekans aralığında baskındır [3]. Ek-sesler ise normal seslerin üzerine eklenen ve çeşitli akciğer hastalıklarının teşhisi için hayati bilgi taşıyan dalga çeşitleridir. Ek-sesler sürekli ve ayrık olmak üzere iki gruba ayrılırlar. Sürekli seslerin en belirgin örneği üfürüm (wheeze) iken ayrık seslerin en belirgin örneği ise çıtırtı (crackle) olarak adlandırılır. Çıtırtı sesleri, zaman alanında ani ve keskin enerji patlamalarından oluşan, genellikle 20 ms'den az süren, ve 200-2000 Hz. frekans aralığında değişen ek-seslerdir [4]. Nefes alış-verişi esnasında meydana gelen çıtırtıların sayısı hastalığın önem derecesi, zamanlaması, süresi ve tipi ise hastalığın çeşidi hakkında bilgi vermektedir [5]. Çıtırtılar kaba ve ince olmak üzere iki çeşittir, kaba çıtırtılar bronşektazi hastalarında gözlemlenirken, ince çıtırtılar ise interstisyel fibrozis ve zatürre hastalarında gözlemlenir [6]. Müzikal dalga yapısına sahip olan üfürümler ise 100 Hz.'ten büyük frekanslarda, frekans düzleminde ayrık tepeler şeklinde 80-250 ms süren yüksek enerjili sinüzoidal sinyallerdir [7]. Kullandığımız veritabanında yer alan birer adet üfürüm, sağlıklı ve çıtırtı işaretinin zaman ve frekans alanı gösterimi örnek olarak Şekil 1'de verilmiştir.

Üfürümlerin varlığı genellikle astım ve kronik obstrüktif akciğer hastalığı (KOAH) göstergesidir. Üfürümlerin baskın frekans değerleri ve süreleri akciğer içindeki hava yollarının tıkanıklık derecesi ile doğru orantılıdır [7]. Hem zamanda hem de frekans bölgesinde normal, çıtırtı ve üfürüm sinyalleri arasında gözle görülmüş bir örtüşme olması problemin çözümünü zorlaştırmakla birlikte, bu seslerin hastalıklı-sağlıklı ayırımında çok önemli olmaları bizi bu sesleri en yüksek başarıyla ayırmaya çalışan yöntemleri aramaya itmiştir. Literatürde üç grup sınıflandırma problemi için güç spektrum yoğunluğu [8,11] ve Mel frekans kepsstral katsayıları (MFKK) [9,10] yöntemlerini öneren birkaç çalışmaya

rastlanılmıştır. [8]'de, 96 öznenin yarısı eğitim diğer yarısı ise test amaçlı kullanılmıştır. Bu çalışmada elde edilen güç spektrum yoğunluğu özellikleri, genetik algoritma ile seçilmiş ve ilgili özellikler çok katmanlı algılayıcı yapay sinir ağları (Multilayer perceptron neural network) ile sınıflandırarak % 91.7 doğruluk oranına ulaşmıştır. [9]'da veri kümesinden MFKK öz-nitelikleri çıkarılmış ve saklı Markov model (hidden Markov model) ile modellenerek en-büyük olabilirlik (maximum likelihood) tabanlı bir sınıflandırıcı yardımıyla % 83 doğruluk oranına ulaşılmıştır. MFKK öz-niteliği kullanılan bir başka çalışmada [10], Gauss karışım modelleri (Gaussian mixture model) tabanlı sınıflandırıcı kullanılarak eğitimde % 98.75, testte ise % 52.5 doğruluk elde edilmiştir. Güç spektrum yoğunluğu (Power spectral density) özellikleri kullanılarak çıkarılan verinin doğrudan çok katmanlı algılayıcı yapay sinir ağları ile sınıflandırıldığı çalışmada [11] ise % 88.7 doğruluk oranına ulaşılmıştır. Önerilen çalışmada ise akciğer solunum seslerini ayırt edebilmek amacıyla; güç spektrum yoğunluğu ve MFKK gibi Fourier dönüşümü tabanlı, klasik yöntemlere ek olarak, ilk defa kişiden bağımsız konuşma tanıma amacıyla önerilmiş [12], algısal doğrusal öngörü (Perceptual linear prediction) özellikleri kullanılmıştır. Önerilen yöntemin başarısı literatürdeki diğer yöntemlerin referans veri tabanımız üzerindeki başarı oranları ile karşılaştırılmıştır. Solunum sesleri alanında kullanılan veriler genellikle farklı özelliklerdeki cihazlardan ve farklı akciğer bölgelerinden alınmaktadır. Ayrıca kullanılan cihazların örnekleme frekansları da aynı olmamaktadır. Bu sebeplerden dolayı literatürde karşılaştırmalı sonuçlara nadir olarak rastlanmaktadır. Bu çalışmada, aynı veri kümesi ve aynı sınıflandırıcılar kullanılarak farklı özellik çıkarma yöntemleri ile literatürle karşılaştırmalı bir çalışma yapılmıştır. Ayrıca, elde edilen özelliklerin enerji (energy), entropi (entropy), standart sapma (standard deviation),

ortalama (mean), en küçük (minimum), en büyük (maximum), eğrilik (skewness) ve basıklık (kurtosis) gibi istatistiksel alt kümeleri kullanılarak yapılan sınıflandırma sonucunda daha az öznelikle, en yüksek başarıma ulaşılmaya çalışılmıştır.

Çalışmada sırasıyla 2. bölümde kullanılan yöntemler, 3. bölümde deneysel sonuçlar, 4. bölümde ise vargılar ve tartışmaya yer verilmiştir.

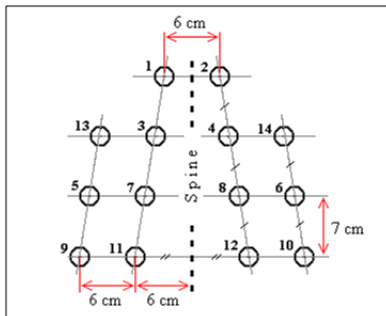
2. Materyal ve Metot

Bu bölümde önce kullanılan veri tabanı tanıtılmış, daha sonra ise kullanılan öznelilik çıkarma yöntemleri ve deneysel kurulum hakkında bilgi verilmiştir.

2.1. Veri toplama

Bu çalışmada, Boğaziçi Üniversitesi Solunum Sesleri Laboratuvarı 'nda tasarlanan 14 kanallı veri toplama sistemi [13] ile kayıt edilen veri tabanı kullanılmıştır. Solunum sesi kaydı yapan mikrofonların göğüs kafesi üzerindeki konumları Şekil 2'de verilmiştir. Her bir veri alımı 15 saniye sürmüş ve örnekleme oranı saniyede 9600 örneğe karşılık gelmiştir. Kayıtlar yapılmadan önce her kişiden bilgilendirilmiş olur onayı alınmıştır. Ayrık ve sürekli hastalık ses örnekleri İstanbul Yedikule Göğüs Hastalıkları ve Göğüs Cerrahisi Eğitim ve Araştırma Hastanesi'nde tedavi altında olan hastalardan alınmıştır. Veri tabanı on dördü erkek, altısı kadın olmak üzere yirmi kişiden oluşmaktadır.

Normal sesler ise daha önce herhangi bir akciğer rahatsızlığı geçirmemiş on kişiden alınmıştır. Çırtırtı, üfürüm ve normal sınıfları için sınıf başına 200 bölüt (segment) ses kullanılmıştır, toplamda ise 600 bölüt ses örneği kullanılmıştır. Elde edilen ses örnekleri uzman hekim tarafından hem dinlenilerek hem de görsel olarak doğrulanmış ve sınıfına göre çırtırtı, üfürüm veya normal olarak etiketlenmiştir. Bölütlerin uzunlukları literatürde ilgili sınıfa ait en kısa zaman dilimiyle uyumlu olacak şekilde değişkendir, bölütlerdeki çırtırtı sayısı birden fazla olabilmektedir. Bölütlerde, önerilen sistemin gerçekliğini arttırmak adına, zaman-frekans özelliklerine göre tek ve çok sesli üfürümler kullanılmıştır.



Şekil 2. Solunum sesi kaydı yapan mikrofonların göğüs kafesi üzerindeki konumları. Şekil, [13] den alınmıştır.

2.2. Güç spektrum yoğunluğu

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

$$P(f) = \frac{1}{K} \left| \sum_{k=0}^{K-1} x_i(k)w(k)e^{-j2\pi f k} \right|^2 = \frac{1}{K} |X(f)|^2 \quad (1)$$

$i = 0, \dots, M - 1$ 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 kullanım kolaylığı sağlamaktadır [14] ve literatürdeki yöntemlerce tercih edilmiştir [8,11]. Welch yöntemi her bir pencerelenmiş bölütün güç spektrum yoğunluklarının ortalamasına denk gelmektedir ve ortalaması alınmış güç spektrum yoğunluğu olarak adlandırılabilir. Solunum ses bölütleri 256 lık Hamming pencereleri (Hamming window) ile % 50 örtüşme (overlap) olacak şekilde çarpılmış ve daha sonra 256 noktalı hızlı Fourier dönüşümü [8] alınmıştır. Elde edilen 129 noktalı spektrum katsayılarının logaritması (log spectrum) alınmış ve enerji, entropi, standart sapma, ortalama, en küçük, en büyük, eğrilik ve basıklık alt kümeleri çıkarılıp sınıflandırıcılara verilmiştir.

2.3. Mel frekans kepsral katsayıları

MFKK otomatik konuşma tanıma için önerilmiş daha sonra akciğer akustiğine de uygulanmıştır. MFKK yöntemi genel olarak insanın işitme sistemini modelleyip doğrusal olmayan mel ölçeğine (Mel scale) göre spektral olarak benzetmeye çalışmaktadır. Denklem 1 de gösterildiği gibi solunum ses bölütü frekans bölgesine geçirilir, daha sonra her bir bölütün genlik spektrumu (amplitude spectrum) bulunur. Mel pencereleri düşük frekanslarda sık, yüksek frekanslarda seyrek dağılım gösteren üçgen süzgeçlerden (triangular filters) oluşmaktadır ve $f(mel) = 1125 * \ln(1 + f/700)$ ile mel frekans ölçeğine çevrilirler. Süzgeç öbeği (Filter group), güç spektrum yoğunluğuna uygulandıktan sonra her bir süzgecin enerjisi hesaplanır ve teorik olarak bu mel süzgecinin frekans tepkisi yani art arda bant geçiren süzgeçlerin frekans tepkisini vermektedir. Son olarak Ayrık Kosinüs Dönüşümü (Discrete Cosine Transform) uygulanarak MFKK bulunur;

$$MFKK = \sum_{k=1}^N P(k) \cos \left(n \left(k - \frac{1}{2} \right) \frac{\pi}{K} \right) \quad (2)$$

öyle ki $n \geq 0$ ve $n < K$, K ise toplam katsayıların sayısıdır, $P(k)$ ise süzgeçlerin çıktısıdır.

2.4. Algısal doğrusal öngörü öznelilikleri

Bu öznelilik çıkarma yöntemi daha önce sağlıklı ve hasta insanları 15 saniyelik kanalları ayırmadan ham öznelilikleri kullanarak ikili sınıflandırma amacıyla önerilmişti [15]. Bu çalışmada ise geçmişteki

çalışmadan farklı olarak hastalıklara özel anormalliklerin (çıtırtı ve üfürüm vb.) olduğu bölütler üzerinden çıkarılan özniteliklerin çeşitli alt kümelerinin üçlü sınıflandırma (normal, çıtırtı ve üfürüm sınıfları) üzerindeki başarımları başka yöntemlerle (Mel frekans kepsstral katsayıları ve güç spektrum yoğunluğu) de karşılaştırılarak incelenmiştir. Bu öznitelik çıkarma yöntemi ilk olarak kişiden bağımsız otomatik konuşma tanıma amacıyla önerilmiştir [12]. Bu çalışmada, ADÖ yönteminin kişiden bağımsız hastalıkların ayırmsayıcı (discriminative) özelliklerini ortaya çıkarmasını ve başarımları iyileştirmesini beklemekteyiz. Bu yöntem, spektrumu kişiye özel (subject specific) özellikleri en küçülterek bükmeyi (warping) ve en önemli ayırmsayıcı öznitelikleri bulmayı amaçlamaktadır.

Solunum sesi penceresi zaman serisi $x(k)$ 'nın Hamming penceresi ile % 50 örtüşme olacak şekilde yumuşatılması sonrası 256 lık hızlı Fourier dönüşümü Denklem 1 deki gibi hesaplanır. Her bir bölütün genlik spektrumu bulunur, $\tau(f) = 6 \ln(w/1200\pi + (w/1200\pi)^2 + 1)^{0.5}$ ile Bark ölçeğine (Bark scale) dönüştürülen frekanslar, hızlı Fourier dönüşümü sonucu elde edilen frekans kutuları (frequency bins) kritik frekans gruplaması ile Bark kutucuklarında birleştirilir [12]. Bu işlem sonucunda seyreltilen frekans kutucuklarında farklı sınıflara ait tepkilerin farklı yoğunlukta ve sayıda olması beklenir. İnsan işitmesine benzetebilmek amacıyla sırasıyla eşit ses şiddeti önvurgulaması (pre-emphasis) ve kübik kök genlik sıkıştırması uygulanır [12]. Ters hızlı Fourier dönüşümü alınan sinyal n. dereceden tüm kutup modeli (all pole) ile Yule-Walker denklemleri çözülürölerek özbağlanımlı (autoregressive) katsayılarla ulaşılır;

$$x(n) = \sum_{k=0}^N a_k y(n-k) \quad (3)$$

öyle ki N model derecesi, $y(n-k)$ geçmiş çıkışlar, $x(n)$ ise giriştir.

2.5. Deneyisel kurulum

Güç spektrum yoğunluğunu kullanan yöntemler ya ham öznitelikleri [11] (129'luk vektör) ya da öznitelik seçimi yaparak [8] elde ettikleri bilgiyi yapay sinir ağlarına sınıflandırma amacıyla vermişlerdir. Bu çalışmada güç spektrum yoğunluğunu Welch yöntemi ile 256 noktalı hızlı Fourier dönüşümü kullanarak hesapladık. Öncesinde %50 örtüşmeli 256 elemanlı Hamming pencereleri seçildi ki bu hem literatür ile uyumlu [8,11] hem de çıtırtıların 10 ms. civarında sürdüğü düşünülürse bir pencere içinde en az bir çıtırtı olmasını sağlamaya dönüktür.

MFKK ve ADÖ için de yukarıda anlatılan benzer ön işleme süreci izlenmiş olup kullanılan yöntemin derecesi literatürde önerilen derecelere (Ref. [6] için 6., [9] için 5., [10] için 13. derece) göre 3 den 13 e kadar deneyisel olarak doğrulanarak seçilmiştir.

Sınıflandırıcı olarak destek vektör makineleri ve k-en yakın komşu sınıflandırıcıları kullanılmıştır. Destek vektör makineleri, doğrusal ve radyal temelli fonksiyon çekirdekleri için, çeşitli C ve gamma (g) değerleri denenerek en iyilenmiştir. k-en yakın komşu sınıflandırıcısında ise komşuluk (k) değerleri 1'den 10'a kadar denenmiştir. Sonuçların nesnel olabilmesi amacıyla birini dışarıda bırak çapraz geçerleme (Leave-one-out cross validation) şeması kullanılmıştır, sınıf bazlı doğruluk değerlerinin yanı sıra tüm sınıfların doğruluklarının ortalaması da bulunmuştur.

3. Bulgular

Tablo 1 de çıtırtı, üfürüm ve normal sesler için güç spektrum yoğunluğu, mel frekans kepsstral katsayıları ve algısal doğrusal öngörü katsayıları için k-en yakın komşu ve destek vektör makineleri doğruluk sonuçları verilmiştir. Ayrıca, üç sınıfın doğruluk sonuçları üzerinden ortalama doğruluk da hesaplanmıştır. Bulunan sonuçların sınıflandırıcılar arası istatistiki testleri eşleştirilmiş iki grup t-test (paired samples t-test) ile yapılmış olup, p-değerleri (en büyüğü 0.0021 olmak üzere) $p < 0.05$ koşulunu sağlayarak anlamlı çıkmıştır.

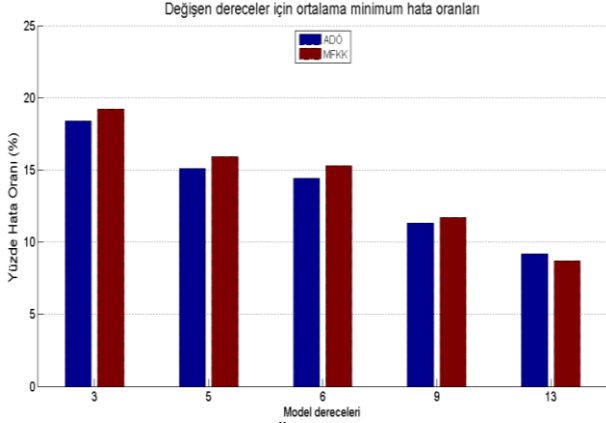
Tablo 1 deki sonuçlara göre ADÖ ve MFKK bariz bir şekilde güç spektrum yoğunluğundan daha yüksek başarıma ulaşmıştır. Bunun sebebi, MFKK ve ADÖ nün düşük frekanslarda zaman çözünürlüğü iyi olmayan güç spektrum yoğunluğundan daha iyi frekans çözünürlüğüne sahip olması ve sabit zaman-frekans çözünürlüğü kullanmasıdır. ADÖ ve MFKK kendi aralarında ise birbirine çok benzer yöntemler olduklarından üç alt öznitelik grubunda (enerji, standart sapma ve en büyülen) ADÖ, diğerlerinde ise MFKK en yüksek ortalama başarıma ulaşmıştır.

ADÖ ve MFKK bant geçiren süzgeçleri sırasıyla oval (çan) ve üçgen tipte olduğundan bazı durumlarda üçgen bazı durumlarda da oval tipteki süzgeçler daha ayırmsayıcı frekans tepkilerini ortaya çıkarmışlardır. Ortalama ve eğrilik alt grup sonuçları ise diğerlerinden daha düşük çıktığından tabloya konulmamıştır.

MFKK ve ADÖ sonuçları incelendiğinde en iyi sonuçların en küçülen, enerji ve en büyülen alt grupları ile elde edildiği bulunmuştur. Bunun sebebi, bant geçiren süzgeçlerin enerji dağılımlarının en büyük ve en küçük değerlerinin ayırıcı özellikleri ön plana çıkarabilmesidir. Örneğin üfürüm belirli frekanslarda yoğun enerjiye sahiptir ve o bant yeterince iyi lokalize edilebilirse ayırıcı özelliği ön plana çıkacaktır, aynı şekilde normal seslerin enerjisi çıtırtıya göre yüksek frekanslarda düşük kalmaktadır [6].

k-en yakın komşu sınıflandırıcısı için en iyi sonuçlar 3. komşu ile alınmıştır, k-en yakın komşu sınıflandırıcısı DVM sınıflandırıcısına göre ortalama

doğruluk değerlerinde birkaç senaryoda daha iyi sonuç verse de istatistiksel öz-nitelik alt kümelerinin çoğunda DVM sınıflandırıcısı daha iyi sonuç vermiştir. Hatta bazı durumlarda k -en yakın komşu sınıflandırıcısı yüzde bazında 15 puan geride kalmıştır (MFKK-enerji), bu da DVM in bu problem türü için daha gürbüz ve istikrarlı sonuçlar verebildiğini ortaya koymaktadır.



Şekil 3. MFKK ve ADÖ yöntemlerinin farklı model derecelerindeki en iyi ortalama hata oranları.

Tabloya tüm sonuçları koymak mümkün olmadığından, MFKK ve ADÖ yöntemlerinin başarılarının, kullanılan model derecesine göre nasıl değiştiğini açıklamak adına model derecesi 3 den 13 e doğru (literatürde rapor edilen dereceleri de içerecek şekilde) artırılmış ve en düşük ortalama hata oranları Şekil 3’de verilmiştir. Model derecesi arttıkça hata oranlarının düştüğü ve bir süre sonra anlamlı artışların olmadığı bulunmuştur. Bu sebeple en-iyi model derecesi olarak 13 seçilmiş ve analiz sonuçları verilmiştir.

4. Tartışma ve Sonuç

MFKK ve ADÖ gibi konuşma tanıma amacıyla önerilen sistemlerin solunum sesi problemi için de umut

vadettiği görülmüştür. Bireysel sınıf bazında doğruluk için en iyi sonuçlar ADÖ yöntemi ile elde edilirken, tüm sınıfların ortalama doğruluk performansı MFKK yöntemi için daha yüksek çıkmıştır.

Literatürde, var olan güç spektrum yoğunluğu [8,11] ve MFKK [9,10] tabanlı yöntemler kendi veritabanlarında sadece ilgili yöntemi denemişlerdir, bu çalışmada ise hem daha önce tarafımızdan sağlıklı hastalıklı tanıma [15] amacıyla önerilen ADÖ yöntemi farklı bir problem olan üçlü sınıflandırmaya uygulanmıştır hem de literatürde üçlü sınıflandırma problemi üzerine önerilen yöntemlerle karşılaştırmalı sonuçlar elde edilmiştir. Ayrıca, ham öz-nitelikler veya öz-nitelik seçimi yerine en az sayıda istatistiksel öz-nitelik alt kümesi kullanılarak, [8] de olduğu gibi genetik algoritması tabanlı öz-nitelik seçimine ihtiyaç kalmadan daha yüksek başarıma ulaşılmıştır.

ADÖ oval, MFCC ise üçgen biçiminde süzgeçlere sahiptir, bu süzgeçlerin hangi alt öz-nitelik grubuna nasıl tepki verdiğinin (ayırmsayıcı bant geçiren süzgeç davranışı) işitsel modelleme alanında akıllı stetoskop tasarımında göz önünde bulundurulması faydalı olabilir.

DVM birkaç durumda k -en yakın komşu sınıflandırıcısının gerisinde kalsa da elde ettiği sonuçlar ortalama hem daha iyi hem de daha karardır. Fourier dönüşümü temelli oldukları için MFKK ve ADÖ zamansal çözünürlükleri (sabit zaman-frekans çözünürlüğü) ayırmsayıcı performansını bir noktaya kadar taşıyabilmektedir. Geleceğe yönelik olarak zamansal çözünürlüğü yüksek olan dalgacık dönüşümü temelli yöntemler daha yüksek başarımlı sağlayabilirler. Çıtır ve üfürümün zaman ve frekans alanında üst üste bindiği durumlar da gelecekte uğraşılması düşünülen problemlerdendir.

Tablo 1. Güç spektrum yoğunluğu (en üstte), Mel frekans kepsral katsayıları (ortada) ve algısal doğrusal öngörü (en altta) yöntemleri için k en yakın komşu ve destek vektör makineleri (DVM) doğruluk yüzdeleri

Tip	Öznitelik											
	Enerji		Entropi		Standart sap.		En küçülten		En büyülten		Basıklık	
	K -en yakın	DVM	K -en yakın	DVM	K -en yakın	DVM	K -en yakın	DVM	K -en yakın	DVM	K -en yakın	DVM
Çıtır	86.0	90.5	48.5	51.5	60.0	72.5	68.5	76.0	81.0	82.0	67.5	71.0
Üfürüm	46.5	76.0	39.0	48.7	49.0	67.0	42.0	39.5	61.0	73.0	63.5	47.0
Normal	45.0	19.5	54.0	51.3	39.5	21.0	43.0	56.0	48.0	62.0	45.5	70.5
Ortalama	59.2	62.0	47.2	50.5	49.5	53.5	51.2	57.2	63.3	72.3	58.8	62.8
Çıtır	55.0	79.5	54.5	74.0	85.5	88.5	88.0	89.5	75.0	75.0	32.5	42.5
Üfürüm	86.5	95.0	82.5	84.0	91.0	44.5	93.5	87.0	87.0	77.5	76.0	78.0
Normal	81.5	92.5	54.0	47.0	40.5	87.5	92.5	92.5	87.0	87.5	57.5	68.5
Ortalama	74.3	89.0	63.7	68.3	72.3	73.5	91.3	89.7	83.0	80	55.3	63.0
Çıtır	90.5	94.0	50.5	53.5	84.5	89.5	75.5	92.5	71.5	87.0	53.0	59.5
Üfürüm	90.0	79.0	46.5	41.0	95.5	72.0	92.0	86.0	91.0	87.5	79.0	81.0
Normal	92.0	95.5	86.5	94.5	40.0	76.0	74.0	89.5	77.0	89.0	36.5	47.5
Ortalama	90.8	89.5	61.2	63.0	73.3	79.2	80.5	89.3	79.8	87.8	56.2	62.7

Teşekkür

Bu çalışma, Boğaziçi Üniversitesi Bilimsel Araştırma Projeleri Fonu tarafından desteklenmiştir (Proje kodu:16A02D2). S. Ulukaya' nın çalışması Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK) BİDEB 2211 programı çerçevesinde desteklenmiştir.

Kaynakça

- [1] Gavriely, N., Cugell, D. W. 1995. Breath sounds methodology, CRC Press.
- [2] Leng, S., Tan, R. S., Chai, K. T. C., Wang, C., Ghista, D., Zhong, L. 2015. The electronic stethoscope, Biomedical Engineering Online, 1-37.
- [3] Oliveira, A., Marques, A. 2014. Respiratory sounds in healthy people: a systematic review, Respiratory medicine 108, 4(2014), 550-570.
- [4] Reichert, S., Gass, R., Brandt, C., Andres, E. 2008. Analysis of respiratory sounds: state of the art, Clinical medicine. Circulatory, respiratory and pulmonary medicine, 2(2008), 45-58.
- [5] Piirila, P., Sovijarvi, A. R. 1995. Crackles: recording, analysis and clinical significance, European Respiratory Journal, 8(1995), 2139-2148.
- [6] Sankur, B., Güler, E. C., Kahya, Y. P. 1996. Multiresolution biological transient extraction applied to respiratory crackles, Computers in Biology and Medicine, 26(1996), 25-39.
- [7] Meslier, N., Charbonneau, G., Racineux, J. L. 1995. Wheezes, European Respiratory Journal, 8(1995), 1942-1948.
- [8] Güler, I., Polat, H., Ergün, U. 2005. Combining neural network and genetic algorithm for prediction of lung sounds, Journal of Medical Systems, 29(2005), 217-231.
- [9] Matsunaga, S., Yamauchi, K., Yamashita, M., Miyahara, S. 2009. Classification between normal and abnormal respiratory sounds based on maximum likelihood approach. in: Acoustics, Speech and Signal Processing (ICASSP), IEEE International Conference on, 517-520.
- [10] Mayorga, P., Druzgalski, C., Morelos, R. L., Gonzalez, O. H., Vidales, J. 2010. Acoustics based assessment of respiratory diseases using GMM classification, in: Engineering in Medicine and Biology Society (EMBC), 32nd Annual International Conference of the IEEE, 6312-6316.
- [11] Abbas, A., Fahim, A. 2010. An automated computerized auscultation and diagnostic system for pulmonary diseases, Journal of Medical Systems, 34(2010), 1149-1155.
- [12] Hermansky, H. 1990. Perceptual linear predictive PLP analysis of speech, The Journal of the Acoustical Society of America, 87(1990), 1738-1752.
- [13] Sen, I., Kahya, Y. P. 2005. A multi-channel device for respiratory sound data acquisition and transient detection, in: Engineering in Medicine and Biology Society (EMBC), 27th Annual International Conference of the IEEE, 6658-6661.
- [14] 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.
- [15] Ulukaya, S., Kahya, Y. P. 2014. Respiratory sound classification using perceptual linear prediction features for healthy-pathological diagnosis. In 18th National Biomedical Engineering Meeting (BIYOMUT)