

Fourier tabanlı Senkrosıkıştırma Dönüşümü Kullanarak LSTM ile Solunum Seslerinin Sınıflanması

Yücel KOÇYİĞİT^{1*}

¹Elektrik-Elektronik Müh. Bölümü /Manisa Celal Bayar Üniversitesi, Türkiye

*(yucel.kocyigit@cbu.edu.tr)

Özet – Bilgisayarlı solunum sesi analiz sistemleri, akciğerin mevcut durumu hakkında hayati bilgiler sağlar. Hekimlerin çeşitli solunum yolu hastalıklarının teşhisinde kullandıkları bu sistemler solunum seslerinin sınıflandırılmasına yardımcı olur. Hekimlerin farklı derecelerde bilgi ve deneyime sahip olması tanı ve dolayısıyla tedavide farklılıklara neden olabilmektedir. İyi kalibre edilmiş makine öğrenimi araçları, hekimlerin daha bilinçli kararlar almasına yardımcı olabilir. Bu amaçla, sağlıklı bireylerden ve hastalardan solunum seslerini sınıflandırmak için farklı makine öğrenme sınıflandırıcıları ve öznitelik çıkarım modelleri geliştirilmiştir. Bu çalışmada öznitelik çıkarımı için Fourier tabanlı Senkrosıkıştırma Dönüşümü, sınıflandırma için ise Derin Öğrenme algoritmalarında Zaman serileri için kullanılan Uzun Kısa Süreli Bellek (LSTM) yöntemi kullanılmış ve Doğruluk %84,67 olarak elde edilmiştir.

Anahtar Kelimeler – Solunum sesleri, Senkrosıkıştırma Dönüşümü, FSST, LSTM, Sınıflama

I. GİRİŞ

Solunum yolu hastalıkları dünyada milyonlarca erken ölüme neden olmaktadır [1]. Bu nedenle solunum yolu hastalıklarının erken teşhisi çok önemli bir tıbbi araştırma alanıdır. Akciğer oskültasyonu solunum yolu hastalıklarının tanısında kullanılan en basit ve en ucuz yöntemdir [3]. Klasik stetoskoplar aracılığıyla gerçekleştirilen yöntem, hekimlere solunum yolu hastalıklarının teşhisinde kritik bilgiler sağlamaktadır. Bu özelliklerine rağmen, hekimin işitme yeteneğine, deneyimine ve çeşitli ses kalıplarını ayırt etme becerisine bağlı olarak birçok sınırlaması vardır [2].

Birçok çalışma, dijital dinlemeyi veri ve algoritmalar açısından güçlendirmeyi amaçlamaktadır. Stetoskopun hafızasına kaydedilen sesler Bilgisayarlı Solunum Sesi Analiz (CORSA) sistemleri ile analiz edilir. CORSA sistemlerinin mevcudiyeti, akciğer sesleri alanında araştırmaların artmasına neden olmuştur.

Spesifik hastalıkla ilişkili gelişigüzel solunum sesleri birçok çalışmada normal solunum sesleriyle

karşılaştırılmıştır. Bu amaçla literatürde solunum seslerinden öznitelik çıkartmak için dalgacık analizi, spektral yöntemler, Mel Frekans Cepstral, ampirik mod ayrışımı gibi çok farklı yöntemler denenmiştir. Sınıflama yöntemleri olarak Yapay Sinir Ağları, Destek Vektör Makineleri, kNN gibi yöntemlerin kullanıldığı görülmektedir.[1-5]

Bu çalışma, önceki çalışmalara benzer şekilde normal ve hasta solunum seslerini teşhis etmeyi amaçlamaktadır. Bunun için yeni bir yaklaşım sergileyerek önce Fourier tabanlı Senkrosıkıştırma Dönüşümü ile zaman-frekans gösterimi elde edilmiş ve bu gösterimden bazı istatistiksel momentler öznitelik olarak bulunmuş; sonra Uzun Kısa Süreli Bellek (LSTM) derin öğrenme ağıyla da sınıflama yapılmıştır.

II. MATERYAL VE YÖNTEM

A. Veri seti

Bu çalışmada Solunum Sesleri (RS) Manisa Celal Bayar Üniversitesi Hafsa Sultan Hastanesi'nde uzman hekimler tarafından kaydedilmiştir. Tüm

kayıtlar, hastanenin solunum hastalıkları kliniğinde tedavi gören 25 sağlıklı ve 25 hasta gönüllüden Littmann 3200 Elektronik stetoskop kullanılarak elde edildi [6].

B. Önışleme

Solunum sesleri kaydedilirken kas ve kalp sesleri gibi düşük frekanslı seslerden ve ani hareketlerden kaynaklanan yüksek frekanslı seslerden etkilenirler. Bu etkilerden kaçınmak için solunum seslerinin baskın frekans aralığı dikkate alınarak ön filtreleme yapıldı. Bunun için Eliptik filtrenin frekans aralığı 100-1500 Hz alındı.

C. Senkrosıkıştırma yöntemi

Fourier tabanlı Senkrosıkıştırma Dönüşümü (Fourier-Based Synchrosqueezing Transform, FSST), zaman-frekans düzleminde çok bileşenli sinyallerin temsilini sağlamak için etkili bir yaklaşım olan bir zaman-frekans dönüşümüdür [7]

Sinyal bileşenlerini zaman içinde yerleştirme ve ardından davranışını frekansta analiz etme yeteneğine sahiptir. Bu dönüşüm, katsayıları ölçek veya frekans olarak yeniden atayarak kısa süreli Fourier dönüşümünün (STFT) zaman-frekans alanında çalışır. Bu yer değiştirme işlemi, seyrek ve keskinleştirilmiş zaman-frekans temsilleri verir ve alınan çeşitli dalga biçimleri arasında ayrım yapmayı kolaylaştırır.

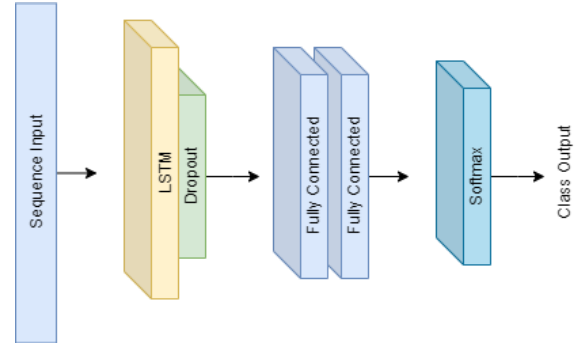
FSST, çok bileşenli bir $f(t)$ sinyalini K salınımlı bileşenlere ayırır[7].

$$f(t) = \sum_{k=1}^K f_k(t) = \sum_{k=1}^K A_k(t)e^{j2\pi\phi_k(t)}$$

burada $A_k(t)$ k bileşeninin anlık genliği ve $\phi_k(t)$ anlık frekansdır. SSD yöntemi her bir k değeri için anlık genliği $A_k(t)$ ve anlık frekansı $\phi_k(t)$ elde etmeyi amaçlar.

D. Uzun Kısa Süreli Bellek (Long Short Time Memory, LSTM)

Derin bir sinir ağı olarak derin öğrenme, makine öğreniminde yeni bir araştırma alanıdır. Uzun Süreli Kısa Süreli Bellek (LSTM) LSTM ağları, tekrarlayan sinir ağı (RNN) yapılarıdır [8] ve lineer olmayan zaman serilerinin etkili bir şekilde işlenmesini gerçekleştirebilir. Şekil 1, bu çalışmada kullanılan LSTM yapısını göstermektedir.



Şekil 1. Eğitimde kullanılan LSTM yapısı

E. Sınıflandırıcıların Performans Değerlendirmesi

Bu çalışmada olduğu gibi, ikili sınıflandırma iki ayrı sonuç verir. Şekil 2'de gösterildiği gibi ikili bir sınıflandırmada dört olası sonuç vardır; Pozitif bir örnek doğru şekilde sınıflandırılırsa, gerçek pozitif (TP), aksi takdirde yanlış negatif (FN) olarak sayılır, negatif örnek doğru şekilde sınıflandırılırsa, gerçek negatif (TN), aksi takdirde yanlış pozitif olarak sayılır. (FP).

		Tahmin Edilen Sınıf	
		Tahmin Edilen Pozitifler	Tahmin Edilen Negatifler
Gerçek Sınıf	Gerçek Pozitifler	TP	FN
	Gerçek Negatifler	FP	TN

Şekil 2. Karışıklık matrisi

Doğruluk kriteri genellikle sınıflandırıcıların performansını değerlendirmek için kullanılır. Ancak, sınıflandırıcıların performansını değerlendirmek için kullanılması gereken tek başına gerçekçi bir ölçüt değildir. Doğruluğun yanı sıra, duyarlılık ve özgüllük bu nedenle bu tür sınıflandırıcıların performansını daha objektif bir şekilde değerlendirmek için kullanılmalıdır.

$$\text{Doğruluk} = (TP+TN)/(TP+TN+FP+FN)$$

$$\text{Duyarlılık} = TP/(TP+FN)$$

$$\text{Özgüllük} = TN/(TN+FP)$$

burada TP, TN, FP ve FN sırasıyla gerçek pozitifleri, gerçek negatifleri, yanlış pozitifleri ve yanlış negatifleri gösterir.

III. BULGULAR

Bu çalışmada solunum seslerinin normal ve hasta olmak üzere Derin Öğrenme algoritması ile sınıflandırılması amaçlanmıştır. Bu amaçla örnekleme frekansı 11025 Hz olan 250 normal ve 250 hasta solunum sesi kullanılmıştır. Solunum sesleri önce 100-1500 Hz frekans aralığında Eliptik tipi 4.dereceden bant geçiren filtreden geçirilmiştir.

Seslere ait özneliklerin bulunması için Fourier Tabanlı Senkrosıkıştırma Dönüşümü sinyallere uygulanarak Zaman-Frekans Düzleminde spektral değişim katsayıları elde edilmiştir. Bu katsayıların ortalaması ve standart sapması hesap edilip normalleştirme işlemi gerçekleştirilmiştir. Frekans ekseninde istatistiksel momentlerden kurtosis ve eğrilik/skewness) hesaplamaları yapılmış ve ayrıca minimum ve maksimum değerleri de bulunmuştur.

Her bir sese ait bu dört öznelik derin öğrenme algoritmalarından zaman serileri için kullanılan LSTM ağına uygulanmıştır. Tablo 1’de LSTM yapısına ait parametreler verilmektedir. Eğitimde normal ve hasta verilerinden oluşan toplam 200 adet veri kullanılırken; ayrıca, 150 adet veri validasyon, 150 adet veri test için ayrılmıştır (Şekil 3). Ağın eğitimde verileri öğrenme başarısına ait karışıklık matrisi (confusion matrix) Şekil 4’te görülmektedir. Eğitim sonunda LSTM ağına test verileri girilerek sınıflama başarısı Şekil 5’teki karışıklık matrisinde verilmektedir. Karışıklık matrisindeki rakamlara göre elde edilen Duyarlılık, Özgüllük ve Doğruluk değerleri hem eğitim hem de

test verileri için sırasıyla Tablo 2 ve Tablo 3’te verilmektedir.

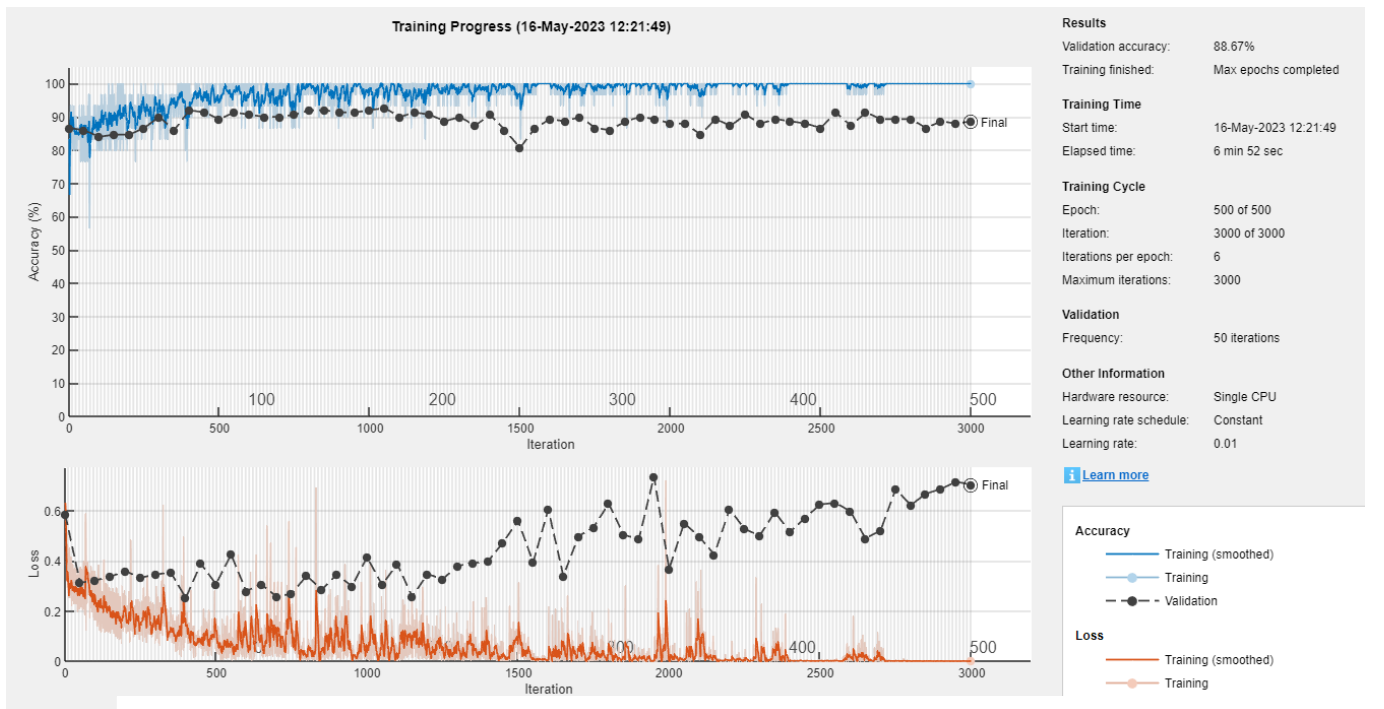
Tüm bu işlemler İ7-7700 HQ işlemci, Nvidia GTX 1060 ekran kartı ve 16GB RAM donanıma sahip dizüstü bilgisayarda MATLAB ortamında gerçekleştirilmiştir

Tablo 1. LSTM Yapısının Parametreleri

Parametreler	Değer
Eğitim Verisi	200
Test Verisi	150
Validasyon Verisi	150
Giriş	4
Batch Boyutu	50
Eğitim Sayısı (Epoch)	500
Öğrenme Hızı	0,01
Seyreltme Değeri (dropout)	0,2
Katman	1
Çıkış	1

		Tahmin Edilen Sınıf	
		Hasta	Normal
Gerçek Sınıf	Hasta	94	0
	Normal	1	105

Şekil 4. Eğitime ait karışıklık matrisi



Şekil 3. LSTM ağının eğitim grafiği

		Tahmin Edilen Sınıf	
		Hasta	Normal
Gerçek Sınıf	Hasta	66	9
	Normal	14	61

Şekil 5. Teste ait karışıklık matrisi

Tablo 2. LSTM ağının Eğitim sonuçları

	Duyarlılık	Özgüllük	Doğruluk
LSTM	%100	%99,1	%99,5

Tablo 3. LSTM ağının Test Sonuçları

	Duyarlılık	Özgüllük	Doğruluk
LSTM	%88	%81,3	%84,67

IV. TARTIŞMA

Fourier tabanlı Senkrosıkıştırma dönüşümü seyrek ve keskinleştirilmiş zaman-frekans temsilleri verdiği ve alınan çeşitli dalga biçimleri arasında ayırım yapmayı kolaylaştırdığından dolayı sınıflama probleminde oldukça etkin olacağı düşünülmüştür. Nitekim eğitim aşamasında verilerin ağ tarafından çok iyi öğrenildiği (doğruluk oranı %99.5) Şekil 4 ve Tablo 1'den anlaşılmaktadır. Buna karşılık test aşamasında Doğruluk oranının bekleneni vermediği ancak Normal ve Hasta verileri (sınıfları) arasında dengeli sınıflama yaptığı Şekil 5 ve Tablo 2'den görülmektedir. Literatürde yapılan makine öğrenmesi yöntemleriyle kıyas edildiğinde başarısının düşük olduğu söylenebilir.

V. SONUÇLAR

Bu çalışmada solunum seslerinin zaman-frekans düzleminde elde edilen spektral katsayılarını kullanarak derin öğrenme algoritması olan LSTM ağıyla sınıflanması amaçlanmıştır. Modelin eğitilmesinde kullanılan veri kümesi %40 eğitim, %30 Validasyon ve %30 test olarak bölümlenmiştir. LSTM ağının eğitiminde %99.5, testinde ise %84.67 doğruluk elde edilmiştir. Bulunan sonuçlar klasik yöntemlerle yapılan sonuçların epey gerisinde kalmakla birlikte derin öğrenme algoritmalarında daha fazla sayıda veriye

ihtiyaç duyması başarının daha yüksek olmasına engel teşkil ettiği düşünülmektedir.

TEŞEKKÜR

Çalışmaya destek olan Manisa Celal Bayar Üniversitesi Hafsa Sultan Hastanesi Göğüs Hastalıkları Anabilim Dalı üyelerine teşekkür ederiz.

KAYNAKLAR

- [1] Sezgin, M. C., et al. Classification of respiratory sounds by using an artificial neural network. In: 2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2001. p. 697-699.
- [2] Maruf SO, Azhar MU, Khawaja SG, Akram MU. Crackle separation and classification from normal respiratory sounds using Gaussian Mixture Model. 2015 IEEE 10th International Conference on Industrial and Information Systems, Sri Lanka, 2015, pp. 267-271
- [3] Lozano, M., Fiz, J. A., & Jané, R. "Performance evaluation of the Hilbert–Huang transform for respiratory sound analysis and its application to continuous adventitious sound characterization", *Signal Processing*, vol. 120, pp. 99-116, 2016.
- [4] Palaniappan, R., Sundaraj, K., & Sundaraj, S., "A comparative study of the svm and k-nn machine learning algorithms for the diagnosis of respiratory pathologies using pulmonary acoustic signals" *BMC bioinformatics*, vol.15, pp 1-8, 2014.
- [5] Sengupta, N., Sahidullah, M., & Saha, G., "Lung sound classification using cepstral-based statistical features", *Computers in Biology and Medicine*, vol. 75, pp. 118-129, 2016.
- [6] Demirci, B. A., Koçyiğit, Y., Kızırlırmak, D., & Havlucu, Y., "Adventitious and Normal Respiratory Sound Analysis with Machine Learning Methods", *Celal Bayar University Journal of Science*, vol. 18(2), pp. 169-180, 2022.
- [7] Oberlin, T., Meignen, S., & Perrier, V., "The Fourier-based synchrosqueezing transform", in *2014 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, 2014, pp. 315-319.
- [8] Hochreiter, S., & Schmidhuber, J., "Long short-term memory", *Neural computation*, vol. 9(8), pp. 1735-1780, 1997.