

MFCC Öznitelikleri ve Adaboost Topluluk Öğrenme Yöntemi Kullanılarak Uyku Seslerinin Sınıflandırılması

Classification of Sleep Sounds Using MFCC Features and Adaboost Ensemble Learning Method

Ensar Arif Sağbaşı*¹ 

¹Bilişim Sistemleri Mühendisliği Bölümü, Muğla Sıtkı Koçman Üniversitesi, Muğla, Türkiye

(arifsagbas@mu.edu.tr)

Received:Aug.22,2023

Accepted:Aug.23,2023

Published:Oct.18,2023

Özetçe— Düzenli ve kaliteli bir gece uykusu insan hayatında hayati önem taşımaktadır. Uyku kalitesi, insanların ve çevrelerindeki insanların günlük yaşamları üzerinde büyük bir etkiye sahiptir. Günümüzde birçok insan uyku bozuklukları konusunda sıkıntı çekmektedir. Bu tarz rahatsızlıklar günlük hayatı etkilemekte ve akıl sağlığını bozabilmektedir. Bu çalışma uyku seslerinin otomatik olarak sınıflandırılması için topluluk öğrenme yöntemini kullanan bir yaklaşım önermektedir. Çalışmada 7 farklı uyku sesini içeren bir veri kümesinden faydalanılmıştır. Öncelikli olarak ses dosyalarından MFCC öznitelikleri çıkartılmıştır. Sonrasında çıkartılan öznitelikler ses sınıflandırılmasında sıklıkla kullanılan lojistik regresyon, destek vektör makinesi, kNN ve rastgele orman gibi bilinen yöntemlerle sınıflandırılmıştır. Sınıflandırma başarısını artırmak amacı ile bu temel sınıflandırıcılar Adaboost topluluk öğrenme yöntemi ile birlikte kullanılması yaklaşımı önerilmiştir. Önerilen yaklaşım ile sınıflandırma başarısında artış gözlemlenmiştir. En başarılı sonuç %96.439 ile Adaboost+Rastgele orman yönteminden elde edilmiştir.

Anahtar Kelimeler : Adaboost, topluluk öğrenme, rastgele orman, MFCC, uyku sesleri.

Abstract— A regular and quality night's sleep is vital in human life. Sleep quality has a great impact on the daily lives of people and those around them. Many people today suffer from sleep disorders. Such disorders affect daily life and can impair mental health. This study proposes an approach using an ensemble learning method for the automatic classification of sleep sounds. In the study, a dataset containing 7 different sleep sounds was used. First of all, MFCC features were extracted from the audio files. Afterward, the extracted features were classified by known methods such as logistic regression, support vector machine, kNN, and random forest, which are frequently used in sound classification. In order to increase classification success, the approach of using these base classifiers together with the Adaboost ensemble learning method was proposed. An increase in classification success was observed with the proposed approach. The most successful result was obtained from the Adaboost+Random forest method with 96.439%.

Keywords : Adaboost, ensemble learning, random forest, MFCC, sleep sounds.

1. Giriş

Uyku, ruh ve beden sağlığının korunması ve yaşam kalitesinin iyi olması için bilinç durumunun geçici olarak değiştiği insan davranışdır (Akbal ve Tuncer, 2021; Akyol vd., 2023). Her gün belli bir süre uyumak sağlıklı bir insan için önemli bir gerekliliktir. Fakat sindirim ve solunum rahatsızlıkları, sosyo-ekonomik durum, şehir hayatı, teknoloji kullanımı ve rüyalar gibi etkenler uyku kalitesini önemli ölçüde etkilemektedir (Ben-Israel vd., 2010; Chattu vd., 2018). Buna bağlı olarak gelişen uyku problemleri çeşitli sağlık sorunlarını da beraberinde getirmektedir (Otsuka vd., 2017). Bireyler uyku problemi ve bu problemin getirdiği sağlık sorunları ile uzun süreler yaşayabilir. Ancak, bu durum yaşam kalitesini düşürmektedir. Uyku sorunlarının tespiti, uyku tanı merkezlerinde yapılan testlerle belirlenmektedir. Ancak bu testler çok farklı donanımlara ve uzman personele ihtiyaç duymaktadır ve maliyetlidir. Ayrıca bireyler için geceyi klinikte geçirmek rahatsız edici bir yaklaşımdır (Akbal ve Tuncer, 2021). Bu noktada, ses ve görüntü kayıtlarının kullanılması daha güvenilir ve konforlu bir yöntem olarak öne çıkmaktadır (Kim vd., 2019). Bireylerin geceleri çıkardıkları sesler basit kayıt cihazlarıyla kayıt altına alınabilir ve sonrasında bu sesler yapay zekâ yaklaşımları ile işlenebilir. Bu sayede sesleri kaydetmek ve bu

seslerden çıkarımlarda bulunmak daha kolay olacaktır. Uyku sesleri sayesinde hem hastalığın teşhisi konulabilir hem de uyku kalitesinin yükseltilmesi için gerekli adımlar atılmasına imkân sağlanır (Akyol vd., 2023).

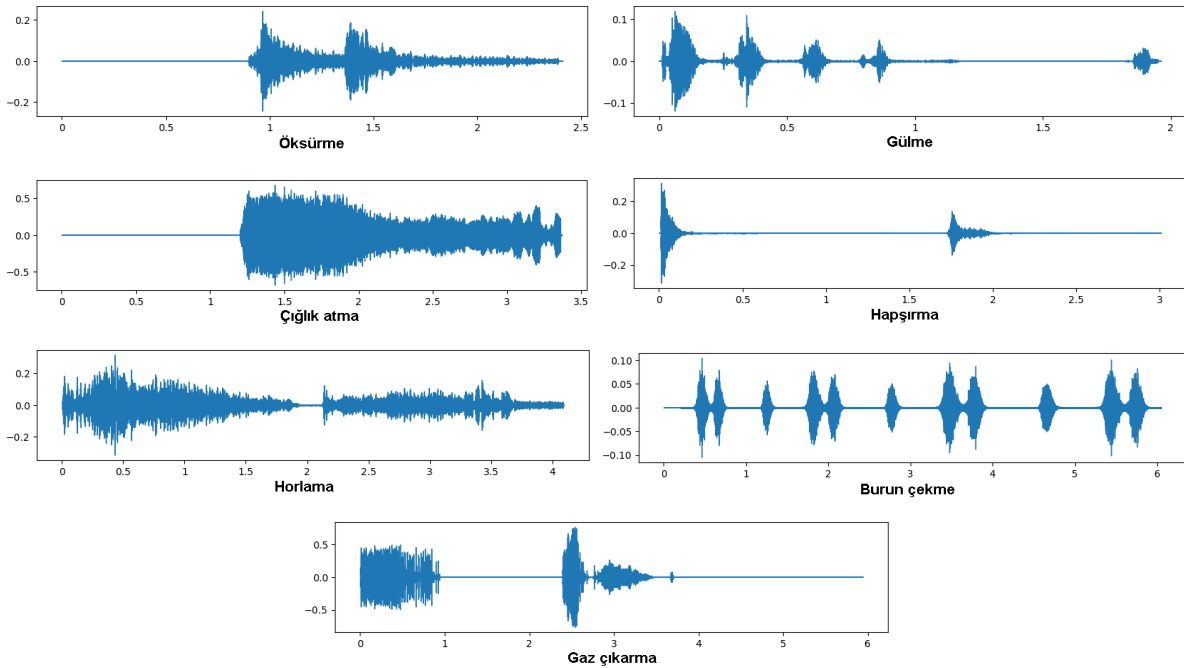
Geçmiş çalışmalar incelendiğinde, uyku seslerinin yapay zekâ yaklaşımları ile sınıflandırıldığı çeşitli araştırmalara rastlamak mümkündür. Akbal ve Tuncer (2021) farklı ses özneliklerini birlikte çıkartarak uyku seslerini yüksek başarı ile sınıflandırmışlardır. Akyol vd. (2023) uyku seslerini daha başarılı sınıflandırmak amacıyla farklı ses özneliklerini kombine etmişler ve en verimli sonucu elde etmek için metasezgisel tabanlı bir model önermişlerdir. Christofferson vd. (2022) zayıf veya düzensiz uykuya ilişkili sesleri ayırt etmek için aktif gürültü önleyici kulaklıklarda bulunan dâhili ve harici mikrofonlardan yararlanarak evde sürekli uyku sesi izlemeyi mümkün kılmışlardır. Adesuyi vd. (2022) çoklu öznelik çıkarımı kullanarak ses sınıflandırması için yeni bir evrimsel sinir ağı modeli önermiştir. Sillaparaya vd. (2022) obstrüktif uyku apnesini horlama seslerine dayalı olarak sınıflandırmak için derin bir öğrenme modeli önermiştir.

Ayrıca, yapay zekâ yaklaşımlarından topluluk öğrenme yöntemleri kullanılarak ses sınıflandırma alanında çeşitli çalışmalar gerçekleştirilmiştir. Vankara ve Devi (2020) kalp atış seslerini kullanarak minimum yanlış alarmla kalp hastalığı tahmin doğruluğunda yüksek kesinlik sağlayan yeni bir topluluk öğrenme stratejisini tasvir etmeye çalışmıştır. Zhao vd. (2019) çevresel ses veri seti üzerinde kısıtlama puanına dayalı geliştirilmiş bir yöntem ve çok modellenli topluluk öznelik seçim yöntemleri kullanmıştır. Chen vd. (2022) kalp seslerini sınıflandırmak için istifleme yöntemini kullanmışlardır. Mohammed vd. (2021) kitle kaynaklı öksürük sesi verilerinden bir COVID-19 ön tarama modeli için topluluk öğrenme tabanlı güçlü bir sınıflandırıcı geliştirmiştir.

Bu çalışmada ise Akbal ve Tuncer'den (2021) sağlanan 7 farklı uyku sesinin sınıflandırılmasında topluluk öğrenme yöntemlerinden olan Adaboost yönteminin performansı değerlendirilmiştir. Öncelikli olarak ses dosyaları üzerinden Mel-frekans kepstrum katsayısı yöntemi ile öznelik çıkartımı işlemi gerçekleştirilmiştir. Geçmiş çalışmalarda sıklıkla kullanılan lojistik regresyon (Hajipour vd., 2020), destek vektör makinesi (Kılıç ve Erdamar, 2020), kNN (Akbal ve Tuncer, 2021; Akyol vd., 2023) ve topluluk öğrenme altında yer alan rastgele orman (Hajipour vd., 2020) yöntemleri temel sınıflandırıcılar olarak seçilmiştir. Rastgele orman yöntemi ile umut verici sonuçlar elde edilmiştir.

2. Veri Kümesi

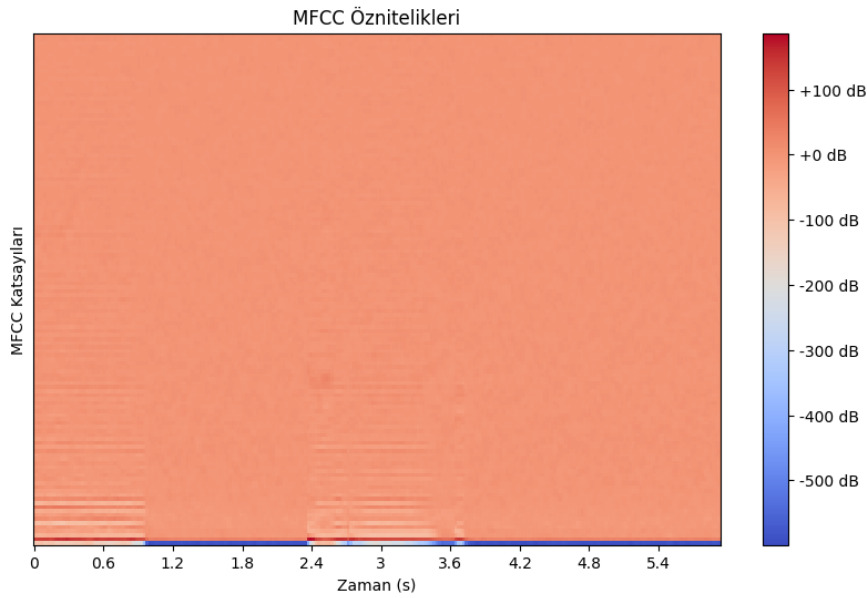
Bu çalışmada Akbal ve Tuncer (2021) tarafından elde edilen veri kümesi kullanılmıştır. Veri seti 7 kategoride 700 ses verisi içermektedir. Kategoriler gülme, hapsirme, öksürme, horlama, gaz çıkarma, burnunu çekme ve çılgık atmayı içermektedir. Her kategoride 20 ila 40 değişken videodan 100 adet ses dosyası bulunmaktadır. Ses dosyalarının örnekleme hızının frekansı 48 kHz ve boyutu 4–8 saniyedir. Her sınıftaki ses sinyallerine ait örnekler Şekil 1'de sunulmuştur.



Şekil 1. Her sınıftaki ses sinyallerine ait örnekler

3. Öznitelik Çıkarımı

Bu çalışmada uyku seslerinin özelliklerini çıkarmak için Librosa (McFee vd., 2015) kütüphanesinden faydalanılmıştır. Uyku seslerinin özniteliklerini elde etmek için Mel-frekans kepstrum katsayısı (MFCC) yöntemi kullanılmıştır. MFCC, ses sinyallerinden öznitelikleri çıkarmak için literatürde yaygın olarak kullanılmaktadır (Akbal ve Tuncer, 2021; Yıldırım, 2022; Ayvaz vd., 2022). MFCC yöntemi ilk olarak Davis ve Mermelstein (1980) tarafından kullanılmıştır. MFCC, 1Khz üzerindeki frekansları tanıyamayan insan işitme davranışına dayalı bir tekniktir. MFCC, insan kulağının ayırt edebileceği frekans farkını temel alır. Sinyal MEL ölçeğinde ifade edilir, bu ölçek, perdelerin gözlemciler tarafından değerlendirilen eşit aralıklı aralıklarla algılanmasına dayanır. Bu ölçek, 1000 Hz'in altındaki frekanslarda doğrusal olarak ve 1000Hz'in üzerindeki logaritmik aralıklarda aralıklı bir filtre kullanır (Martines vd., 2012). Bu çalışmada her bir ses dosyasından Librosa kütüphanesi kullanılarak MFCC öznitelikleri çıkartılmıştır. n_mfcc parametresi 128 olarak ayarlanmıştır. Bu durumda 128 satırlık bir matris elde edilmektedir. Örnek bir uyku sesine ait çıkarılan MFCC özniteliği görseli Şekil 2'de sunulmaktadır. Elde edilen matrisin satırlarının ortalaması alınarak çalışmada kullanılacak olan MFCC öznitelik vektörü elde edilmiştir. Sonrasında elde edilen bu öznitelikler önerilen topluluk öğrenme yaklaşımı ile sınıflandırılmış ve sonuçları karşılaştırılmıştır.



Şekil 2. Örnek bir uyku sesine ait MFCC özniteliğinin görselleştirilmesi

4. Adaboost Algoritması

Freund ve Schapire (1997) tarafından tanıtılan AdaBoost algoritması, daha önceki güçlendirme algoritmalarının birçok pratik zorluklarını çözmüştür. Algoritma girdi olarak bir eğitim seti $(x_1, y_1), \dots, (x_m, y_m)$ alır. Burada her x_i , X etki alanına veya örnek uzayına aittir ve her y_i etiketi, Y etiket kümesindedir. AdaBoost, belirli bir zayıf veya temel öğrenme algoritmasını bir dizi turda tekrar tekrar çağırır $t = 1, \dots, T$. Algoritmanın ana fikirlerinden biri, eğitim seti üzerinde bir ağırlık dağılımı veya seti sağlamaktır. t turundaki eğitim örneği i'deki bu dağılımın ağırlığı $D_t(i)$ ile gösterilir. Başlangıçta tüm ağırlıklar eşit olarak ayarlanır, ancak her turda yanlış sınıflandırılan örneklerin ağırlıkları artırılır, böylece zayıf öğrenci eğitim setindeki zor örnekler odaklanmak zorunda kalır (Freund vd., 1999).

5. Deneysel Bulgular

Bu çalışmada, 7 farklı uyku sesinin sınıflandırılması için temel sınıflandırıcılar olarak Lojistik regresyon (LR), destek vektör makinesi (DVM), kNN ve rastgele orman (RO) yöntemleri seçilmiştir. Gerçekleştirilen ön testler sonucunda kNN yönteminde k değerinin 1, destek vektör makinesi yönteminde çekirdek fonksiyonunun polinom, rastgele orman yönteminde ise tur sayısının 300 olarak belirlenmesine karar verilmiştir. Bütün deneyler 10 katmanlı çapraz doğrulama ile gerçekleştirilmiştir. Bu yöntemlere ait sınıflandırma doğruluk oranı, precision, recall ve f-measure değerlerinin ortalama sonuçları Tablo 1'de sunulmaktadır.

Tablo 1. Temel sınıflandırıcılardan elde edilen sonuçlar

Yöntem	Doğruluk (%)	Precision	Recall	F-measure
Lojistik regresyon	93.143	0.933	0.931	0.931
Destek vektör makinesi	93.286	0.938	0.933	0.933
kNN	89.571	0.908	0.896	0.894
Rastgele orman	95.000	0.951	0.950	0.950

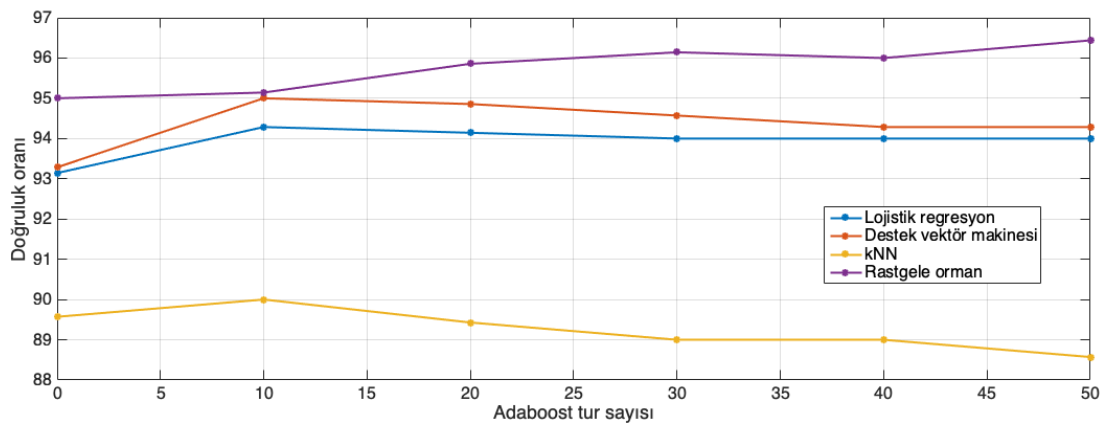
Temel sınıflandırıcıların performansı değerlendirildiğinde en başarılı yöntemin topluluk öğrenme yaklaşımı olan rastgele orman olduğu görülmektedir. Bu yöntem ile %95 sınıflandırma başarısı yakalanmıştır. Bu yöntemi destek vektör makinesi ve lojistik regresyon yöntemleri takip etmiştir. Bu dört yöntem arasında en düşük sınıflandırma başarısına sahip yöntem ise %89.571 ile kNN olmuştur.

Bu çalışmada, temel sınıflandırma yöntemlerinin başarısını artırmaya yönelik, bu sınıflandırıcıları Adaboost topluluk öğrenme yöntemi ile birlikte kullanılması önerilmiştir. Adaboost algoritmasında ise tur sayısı 10, 20, 30, 40 ve 50 olacak şekilde 5 farklı deney gerçekleştirilmiştir. Yönteme ait yeniden örnekleme özelliğinin ise kullanılmasına gerçekleştirilen ön testler sonucunda karar verilmiştir. Adaboost topluluğu ile gerçekleştirilen sınıflandırmalara ait sayısal bilgiler Tablo 2’de sunulmaktadır.

Tablo 2. Adaboost topluluğu ile elde edilen sınıflandırma doğruluk oranları

Yöntem / Tur sayısı	10	20	30	40	50
Adaboost+LR	94.286	94.143	94.000	94.000	94.000
Adaboost+DVM	95.000	94.857	94.571	94.286	94.286
Adaboost+kNN	90.000	89.429	89.000	89.000	88.571
Adaboost+RO	95.143	95.857	96.143	96.000	96.439

Adaboost yöntemi ile elde edilen sonuçlar incelendiğinde ise LR, DVM ve kNN yöntemlerinde tur sayısı 10 olarak seçildiğinde en başarılı sonuçların sağlandığı görülmektedir. Bu yöntemlerde tur sayısı arttıkça sınıflandırma doğruluğunda bir artış gözlemlenmemiştir. Fakat LR ve DVM yöntemleri için tekil sınıflandırmalarından daha başarılı sonuç elde edilmiştir. kNN yönteminde ise tur sayısı 10 olarak belirlendiğinde daha başarılı sonuçlar elde edilirken, tur sayısı arttıkça tekil sınıflandırma başarısının altına düşmüştür. Rastgele orman yönteminde ise en başarılı sonuç tur sayısı 50 olarak belirlendiğinde elde edilmiştir. Bu sınıflandırmadan %96.439 oranında sınıflandırma doğruluğu sağlanmıştır. Adaboost algoritmasına ait tur sayısı değişkenine göre yöntemlerden elde edilen sınıflandırma doğruluk oranlarına ait grafik Şekil 3’te sunulmaktadır.

**Şekil 3.** Adaboost algoritması tur sayısına göre yöntemlerden elde edilen sınıflandırma doğrulukları (%)

Şekil 3 incelendiğinde Rastgele orman yönteminin tekil başarısının, karşılaştırılan diğer yöntemlerden daha yüksek olduğu görülmektedir. LR, DVM ve kNN yöntemlerinde tur sayısı arttıkça başarı düşerken RO yönteminde ise artış görülmüştür. Böylece tur sayısı arttıkça yöntemler arasındaki fark açılmıştır. En yüksek sınıflandırma başarısına sahip Adaboost+RO yönteminden elde edilen karışıklık matrisi Tablo 3’te sunulmuştur.

Tablo 3. Adaboost+RO yönteminden elde edilen karışıklık matrisi

	a	b	c	d	e	f	g
a = Hapşırma	97	0	2	0	1	0	0
b = Öksürme	1	99	0	0	0	0	0
c = Gülme	4	2	89	0	5	0	0
d = Horlama	0	0	0	100	0	0	0
e = Çığlık atma	5	0	3	0	92	0	0
f = Gaz çıkarma	0	0	1	0	0	99	0
g = Burun çekme	1	0	0	0	0	0	99

Tablo 3 incelendiğinde gülme ve çığlık atma sınıflarında, diğer sınıflara göre daha fazla hatalı sınıflandırma gerçekleştirildiği görülmektedir. 100 çığlık atma sesi örneğinin 5 tanesi hapşırma, 3 tanesi gülme olarak sınıflandırılmıştır. Gülme sesleri ise en çok çığlık atma hapşırma sesleri ile karışmıştır. Horlama sesi ise diğer hiçbir uyku sesi ile karışmamış, %100 doğruluk ile sınıflandırılmıştır.

6. Sonuç

Uyku seslerinin sınıflandırılmasında topluluk öğrenme yöntemlerinden olan Adaboost yönteminin performansı değerlendirilmiştir. 7 farklı uyku sesini içeren 700 adet ses dosyasından MFCC yöntemi ile 128 adet öznitelik çıkartılmış ve sonrasında LR, DVM, kNN ve RO yöntemleri ile sınıflandırılmıştır. En yüksek sınıflandırma doğruluğu %96.439 ile Adaboost+RO yönteminden elde edilmiştir. Bu yaklaşımda, Adaboost algoritmasının tur sayısı arttıkça sınıflandırma başarısının arttığı gözlemlenmiştir. RO yönteminin ve Adaboost algoritmasının parametrelerinin optimizasyonu üzerine durulursa daha başarılı sonuçların elde edilebileceği düşünülmektedir. Ayrıca bu çalışmada öznitelik çıkarımında sadece MFCC yönteminden faydalandığı göz önünde bulundurulmalıdır. Gelecek çalışmalarda farklı ve etkili öznitelik çıkarımı ve öznitelik seçimi yaklaşımları ile bu yaklaşımın daha da geliştirilmesi planlanmaktadır.

Kaynaklar

- Adesuyi, T. A., Kim, B. M., & Kim, J. (2022). Snoring sound classification using 1D-CNN model based on multi-feature extraction. *International Journal of Fuzzy Logic and Intelligent Systems*, 22(1), 1-10.
- Akbal, E., & Tuncer, T. (2021). FusedTSNet: an automated nocturnal sleep sound classification method based on a fused textural and statistical feature generation network. *Applied Acoustics*, 171, 107559.
- Akyol, S., Yildirim, M., & Alatas, B. (2023). Multi-feature fusion and improved BO and IGWO metaheuristics based models for automatically diagnosing the sleep disorders from sleep sounds. *Computers in Biology and Medicine*, 157, 106768.
- Ayvaz, U., Gürüler, H., Khan, F., Ahmed, N., Whangbo, T., & Bobomirzaevich, A. (2022). Automatic speaker recognition using mel-frequency cepstral coefficients through machine learning. *CMC-Computers Materials & Continua*, 71(3).
- Ben-Israel, N., Tarasiuk, A., & Zigel, Y. (2010, August). Nocturnal sound analysis for the diagnosis of obstructive sleep apnea. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, (pp. 6146-6149). IEEE.
- Chattu, V. K., Manzar, M. D., Kumary, S., Burman, D., Spence, D. W., & Pandi-Perumal, S. R. (2018, December). The global problem of insufficient sleep and its serious public health implications. In *Healthcare*, (Vol. 7, No. 1, p. 1). MDPI.
- Chen, J., Dang, X., & Li, M. (2022, April). Heart Sound Classification Method based on Ensemble Learning. In *2022 7th International Conference on Intelligent Computing and Signal Processing*, (pp. 8-13). IEEE.
- Christofferson, K., Chen, X., Wang, Z., Mariakakis, A., & Wang, Y. (2022, March). Sleep Sound Classification Using ANC-Enabled Earbuds. In *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events*, (pp. 397-402). IEEE.
- Davis, S., & Mermelstein, P. (1980). Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE transactions on acoustics, speech, and signal processing*, 28(4), 357-366.
- Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1), 119-139.

- Freund, Y., Schapire, R., & Abe, N. (1999). A short introduction to boosting. *Journal-Japanese Society For Artificial Intelligence*, 14(771-780), 1612.
- Hajipour, F., Jozani, M. J., & Moussavi, Z. (2020). A comparison of regularized logistic regression and random forest machine learning models for daytime diagnosis of obstructive sleep apnea. *Medical & Biological Engineering & Computing*, 58, 2517-2529.
- Kılıç, E., & Erdamar, A. (2020). Destek vektör makineleri kullanarak uyku seslerinin çoklu sınıflandırılması. *Journal of the Institute of Science and Technology*, 10(4), 2474-2485.
- Kim, J. W., Cho, S. W., Lee, K., & Shin, J. Y. (2019). Correlation analysis of sleep breathing sound and polysomnographic features. *ERJ Open Research*, 5 : Suppl. 3, P55
- Martinez, J., Perez, H., Escamilla, E., & Suzuki, M. M. (2012, February). Speaker recognition using Mel frequency Cepstral Coefficients (MFCC) and Vector quantization (VQ) techniques. *In Conielectomp 2012, 22nd International conference on electrical communications and computers* (pp. 248-251). IEEE.
- McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., & Nieto, O. (2015, July). librosa: Audio and music signal analysis in python. *In Proceedings of the 14th python in science conference* (Vol. 8, pp. 18-25).
- Mohammed, E. A., Keyhani, M., Sanati-Nezhad, A., Hejazi, S. H., & Far, B. H. (2021). An ensemble learning approach to digital corona virus preliminary screening from cough sounds. *Scientific Reports*, 11(1), 1-11.
- Otsuka, Y., Kaneita, Y., Itani, O., Nakagome, S., Jike, M., & Ohida, T. (2017). Relationship between stress coping and sleep disorders among the general Japanese population: a nationwide representative survey. *Sleep medicine*, 37, 38-45.
- Sillaparaya, A., Bhatranand, A., Sudthongkong, C., Chamnongthai, K., & Jiraraksopakun, Y. (2022, November). Obstructive Sleep Apnea Classification Using Snore Sounds Based on Deep Learning. *In 2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference* (pp. 1152-1155). IEEE.
- Vankara, J., Lavanya Devi, G. PAELC: Predictive Analysis by Ensemble Learning and Classification heart disease detection using beat sound. *Int J Speech Technol*, 23, 31-43 (2020). <https://doi.org/10.1007/s10772-020-09670-6>
- Yıldırım, M. (2022). MFCC Yöntemi ve Önerilen Derin Model ile Çevresel Seslerin Otomatik Olarak Sınıflandırılması. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 34(1), 449-457.
- Zhao, S., Zhang, Y., Xu, H., & Han, T. (2019). Ensemble classification based on feature selection for environmental sound recognition. *Mathematical Problems in Engineering*, 2019.