



MFCC ve LBP Yöntemlerinin Karşılaştırılması ile Konuşmacıyı Tanıma ve Konuşmacıyı Doğrulama

Speaker Recognition and Speaker Verification by Comparison of MFCC and LBP Methods

Emrah AYDEMİR
Sakarya Üniversitesi
Yönetim Bilişim Sistemleri
Sakarya, Türkiye
emrahaydemir@sakarya.edu.tr
ORCID: 0000-0002-8380-7891

Öz

Konuşmacıyı tanıma ya da konuşmacıyı tanımlama konuşmacının ses sinyallerine ait parametrelerinin analiz edilmesi ile otomatik olarak tanınmasıdır. İnsan sesleri sahibine çok yüksek bağlılık içerir. Bu nedenle bu çalışmada Yasin Suresini okuyan 46 farklı kişiden kim olduğunu belirlemek için Youtube üzerinden veri kümesi elde edilmiştir. Elde edilen ses dosyalarından MFCC ve LBP ile öznelik çıkarımı yapılmıştır. Öznelik vektörleri çeşitli sınıflandırma algoritmaları ile sınanmış ve MFCC için %35,10 başarı elde edilirken LBP için %90,74 oranında başarılı sonuçlar elde edilmiştir. Kişi doğrulama için ise LBP'de %100 sınıflandırma başarısı elde edilmiştir.

Anahtar sözcükler: Konuşmacı tanıma, konuşmacı doğrulama, yerel ikili desen, mfcc

Abstract

Speaker recognition or speaker identification is the automatic recognition of the speaker by analyzing the parameters of the audio signals. Human voices contain a very high attachment to their owner. For this reason, in this study, a dataset was obtained from Youtube to determine who is from 46 different people who read Surah Yasin. Feature extraction was done from the obtained audio files with MFCC and LBP. Feature vectors have been tested with various classification algorithms and 35.10% success has been obtained for MFCC, while 90.74% success has been obtained for LBP. For person verification, 100% classification success was achieved in LBP.

Gönderme, düzeltme ve kabul tarihi: 06.03.2022 - 30.05.2022 - 28.06.2022

Makale türü: Araştırma

Keywords: Speaker recognition, speaker verification, local binary pattern, mfcc

1. Giriş

İnsanlar arasındaki en verimli iletişim şekli şüphesiz konuşmadır. Bu konuşma şekli aynı zamanda her bir birey için benzersiz olup biyometrik özellik de taşımaktadır. Bu nedenle kişilerin yalnızca sesleri kullanarak onları tanımlamak mümkün olabilir. Konuşma sesleri sayısal bir sinyal özelliklerine sahiptir. Bu nedenle sinyal işleme çalışmaları altında değerlendirilebilir. Konuşmacı tanıma belirli algoritmalar ile insan konuşmasının makine tarafından anlaşılması ve bir veri kümesi içinde doğrulama yapılmasıdır. Buradan hareketle konuşmacı seslerinde konuşmacı tespiti ya da konuşmacı doğrulama çalışmaları yapılabileceği gibi komut tanıma sistemlerinde de kullanılabilir. Konuşmacıyı belirleme çalışmaları konuşan kişiyi birden fazla kişiye ait konuşma sesleri arasından belirlemek için konuşmacıyı doğrulama ise bir sesin istenen kişiye ait olup olmadığının tespit edilmesidir [1, 2]. Özet olarak konuşmacı tanıma ya da konuşmacı tanımlama; konuşmacının ses sinyallerine ait parametrelerinin analiz edilmesi ile otomatik olarak tanınmasıdır [3].

1.1. Literatür Taraması

Seslerden yalnızca insanları tanımak üzerine çalışmalar yapılmamıştır. Böcekleri de çıkartıkları sesler üzerinden tanımak ve sınıflandırmak üzerine Le-Qing [4] çalışmıştır. Seslerden öznelik çıkarmak için Mel-frekans kepsral katsayıları (Mel-frequency cepstral coefficients-MFCC) yöntemi ve sınıflandırmak için ise olasılıksal sinir ağı (probabilistic neural network-PNN) yöntemi kullanılmış ve

%96'yı aşan sınıflandırma başarısı elde edilmiştir. Wanli ve Guoxin [5] de 30 farklı bireye ait konuşmacı seslerini MFCC yöntemi ile tespit etmeye çalışmıştır. Bu tür sistemler aynı zamanda güvenlik amaçlı da kullanılmaya başlanmıştır. Özellikle telefon bankacılığı ya da cep telefonu sistemleri için konuşmacı kimliği doğrulamada kullanılmaktadır [6]. Ayrıca adli tıp çalışmalarında konuşmacıyı tanıma da yaygın olarak kullanılmaktadır [7]. Yerel ikili desen (Local Binary Pattern-LBP) yöntemi ile ortam tahmini çalışması Toffa ve diğ. [14] tarafından yapılmış ve ESC10 veri kümesi için %88,5 ve ESC50 için %64,6 oranında sınıflandırma başarısı elde edilmiştir. Er [15] ise yine LBP yöntemini kullanarak kalp atış seslerini sınıflandırmıştır ve %91,78 oranında başarı elde edilmiştir. Olay yeri sınıflandırmasında da LBP yöntemlerinin kullanıldığı ve başarılı sınıflandırma çalışmalarının olduğu farklı çalışmalar vardır [17-19]. Kalp seslerinin MFCC yöntemi ile sınıflandırılmasını Deng ve ark. [20] yapmış ve %98 başarı elde etmişlerdir. Yine stetoskop ile elde edilen ses verilerinin MFCC ile sınıflandırılmasındaki bir başka çalışmada %96,25 oranında başarılı sınıflandırmalar elde edilmiştir [21]. Kardiyak ses sınıflandırılmasında yine MFCC ile öznelik çıkarım yöntemi tercih edilmiş ve %100 oranında başarılı sınıflandırma elde edilmiştir [22]. 10 farklı hayvan seslerinden tanınmanın MFCC yöntemi ile yapıldığı bir çalışmada ise %75 oranında başarılı tanıma gerçekleştirilmiştir [23].

2. Amaç ve Önem

İnsanları diğer türlerden ayıran en önemli özellikleri arasında sözlü iletişim gelir. Bu doğal iletişim konusu uzun yıllardan beri ilgi odağı olmuş ve olmaya devam etmektedir. İnsan sesleri sahibine çok yüksek bağlılık içerir. Bu konuda yapılan çalışmalarda iki kişinin hiçbir şekilde aynı sese sahip olmadığı belirlenmiştir [8]. Burada sesteki desen konuşmanın kodlamasını ifade etmektedir. Seslerdeki bu desen çözülerek kaynakların daha verimli kullanılmasına ve maliyet etkinliğinin artmasına olanak sağlayacaktır. Özellikle güvenlik tabanlı sistemler için bu durum daha çok önemlidir [9]. Bilgisayar tabanlı birçok farklı sistemin kullanılmasında metin tabanlı bilgisayar insan iletişimine kıyasla ses ile iletişim her geçen gün yaygınlaşmaktadır. Buradaki çalışmada söz konusu bu çalışmalara katkı sunmayı amaçlamıştır.

Literatürde farklı öznelik çıkarım yöntemleri kullanılarak hem konuşmacıyı tanıma hem de konuşmacıyı doğrulama çalışmaları çok sayıda vardır. Bu çalışmada ise aynı metne ait seslerden farklı öznelik yöntemlerinden olan MFCC ve LBP ile öznelik çıkarım yöntemi ile konuşmacı tanıma ve doğrulama yapılmıştır. Literatür taraması başlığında incelenen çalışmalar irdelendiğinde farklı çalışmalarda her iki yöntemin de yüksek başarılar elde ettiği görülmektedir. Bu çalışmada bu yöntemde elde edilen sonuçlar ile her sınıflandırma sonuçları için karşılaştırılmıştır. Ayrıca bu çalışmada literatüre yeni bir veri kümesi kazandırılmış ve herkesin kullanımına açık bir ortamda paylaşılması sağlanmıştır. Bu veri kümesinin benzerlerinden farklı olan özelliği ise 46 farklı kişiye ait olmasının yanı sıra tümü için aynı metni içermesidir.

3. Materyal ve Yöntem

3.1 Veri Kümesi

Bu çalışmada 46 farklı kişiye ait ses kayıtları toplanmıştır. Youtube üzerinde Yasin Süresini okuyan kişilerin ses kayıtları toplanmıştır. Her bir kişi farklı sürelerde okumayı tamamlamıştır. Tüm kayıtlar ortalama 16 dakika 11 saniye sürmüştür. Veriler başka araştırmacılar tarafından kullanılabilmesi için Kaggle veri kümesi deposunda <https://www.kaggle.com/emrahaydemr/speaker-audio-files-for-same-text> bağlantısı ile saklanmıştır. Tüm kullanıcıların erişimine açık hale getirilmiştir. Her bir ses dosyası 10 saniyelik parçalara bölünmüştür. Böylece toplam 4428 örneklem elde edilmiştir ve her bir ses dosyasına yönelik ortalama 96 örneklem oluşturulmuştur. 10 saniyeden daha küçük olan ses parçaları örneklem dışında tutulmuştur.

Çizelge-1: Veri Kümesi Ayrıntıları

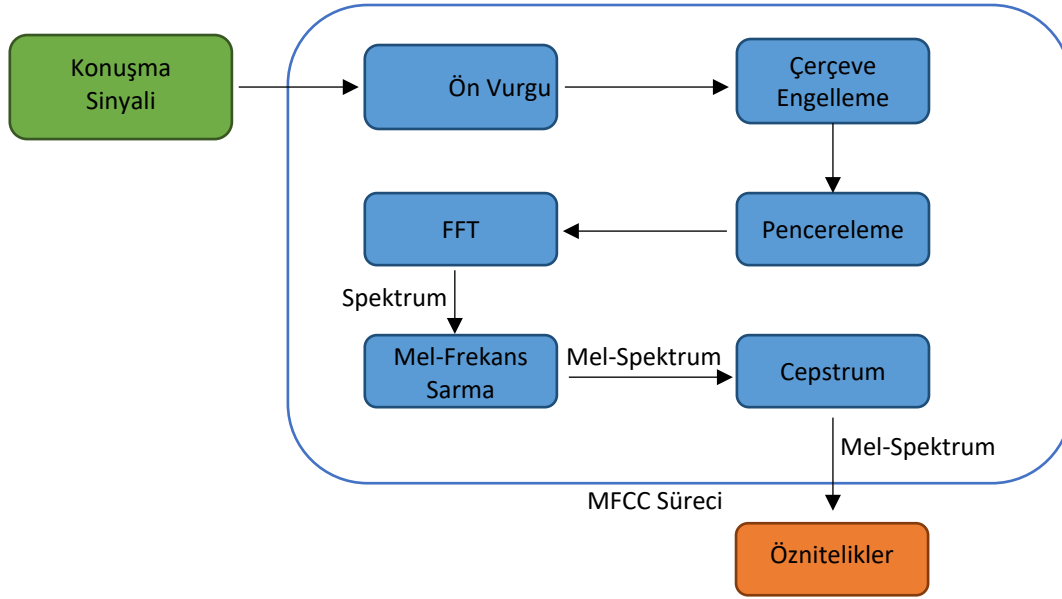
Konuşmacı	Süre	Örneklem	Konuşmacı	Süre	Örneklem
Abdulaziz az Zahrani	19:09	114	M. Sıddık Minşevi	32:05	192
AbdulAziz Bandar Balila	18:17	109	Mustafa Kaya	07:12	43
Abdulbasit Abdussamed	32:50	197	Mustafa Uyar	14:29	86
Abdullah Altun	14:45	88	M. Özcan Güneşdoğdu	38:39	231
Abdullah Gürbüz	14:29	86	Nasser al Qatami	20:33	123
Adem Kemaneci	15:21	92	Qadr Al Kurdi	08:58	53
Ahmed Khedr	17:09	102	Saad Al Ghamdi	13:12	79
Ahmet el acemi	13:32	81	Salah Bukhatir	08:26	50
Ali Tel	18:43	112	Salem al Ruwaili	16:29	98
Ayasofya Açılışı	18:57	113	Salman al Utaybi	15:23	92
A. Mahmut Ünlü	14:02	84	Salih Polat	07:00	42
Davut Kaya	10:15	61	Selman Okumuş	11:48	70
Ebubekir Şatırı	14:12	85	Sudeys	09:10	55
Erhan Mete	16:17	97	Suud Eş-Şureym	10:17	61
Fatih Hoca	10:04	60	Tayyar Altrıkulaç	16:14	97
Fatih Koca	16:39	99	Yasser al Dosari	15:20	92
Fatih Çollak	21:30	129	İdris Abkar	19:01	114
Hazza al Balushi	15:16	91	İhsan Atasoy	15:41	94
Mahir al-Mu'ayqali	07:44	46	İlhan Tok	15:35	93
Mansur al Salimi	17:37	105	İshak Danış	17:03	102
Mehmet Emin AY	14:36	87	İslam Subhi	16:24	98
M. Rashid Al Afasy	17:45	106	İsmail Biçer	21:41	130
Muhammad al Kurdi	21:25	128	İsmail Damar	10:15	61

3.2 Yöntem

3.2.1 Mel-Frekans Cepstrum Katsayıları (MFCC)

Mel-Frekans Cepstrum Katsayıları (MFCC) ses sinyallerinden öznelik çıkarımı için en iyi bilinen, sağlam, doğru ve bilinen

yöntemlerden biridir. Bu yöntem insan işitme sisteminin algısını dikkate alarak cepstral katsayısını kullanır. Bu yöntem her bir ses dosyası için 1723 öznelik çıkarmaktadır. Şekil-1 MFCC yöntemini sürecini anlatmaktadır [10].



Şekil-1: MFCC ile Öznelik Çıkarımı

- Ön vurgu aşaması konuşma sinyalindeki frekanstan daha düzgün bir spektral biçimi elde etmeyi amaçlar. Böylece ses yakalamada gürültü azaltılarak bir filtreleme işlemi yapılmış olur. Öznelik çıkarma sürecinde örnekleme işlemi sonrası ön vurgu filtresi gereklidir.
- Çerçeve engelleme aşamasında ses sinyali üst üste binmiş çok sayıda çerçeveye bölünür. Böylece sinyal silinmesi azalır. Bu durum her bir sinyalin en az bir çerçeve içinde yer almasına kadar sürer.
- Pencereleme aşaması özellikle uzun bir konuşma sinyalini temsil eden bir kesit almayı kapsar. Böylece Sonlu Darbe Tepkisi (FIR) dijital filtre yaklaşımı ile sinyal parçalarının süreksizliği nedeniyle örtüşme sinyali ortadan kalkmış olur. Bu süreksizlikler çerçevenin süreci bloke etmesi sonucu ortaya çıkar.
- Fourier dönüşümleri (FFT) aşaması ses sinyalinin sınırlı alanını bir frekans spektrumuna dönüştürmek için kullanılır. Bu yöntem zaman alanından N örnek içeren her bir karenin frekans alanına dönüştürmek amacıyla yararlı olan ve DFT'deki tekrarlanabilir çarpmayı azaltan hızlı bir Ayırık Fourier Dönüşümü (DFT) algoritmasıdır.
- Mel-Frekans Sarma aşamasında insan konuşmasının algısını yansıtan frekans eksenindeki birim olan Mel ölçeğinin kullanılmasını içerir. Frekansın düşmesi aralığı daraltırken frekansın yükselmesi aralığı genişletecektir. Bu durum insanların düşük frekanstaki sesleri daha iyi anladığına yorumlanabilir.
- Cepstrum aşamasında ise mel spektrumu Ayırık Kosinüs Dönüşümü (DCT) kullanılarak zaman alanına dönüştürülür

ve mel-frekans cepstrum katsayısı (MFCC) olarak adlandırılan sonucu üretir.

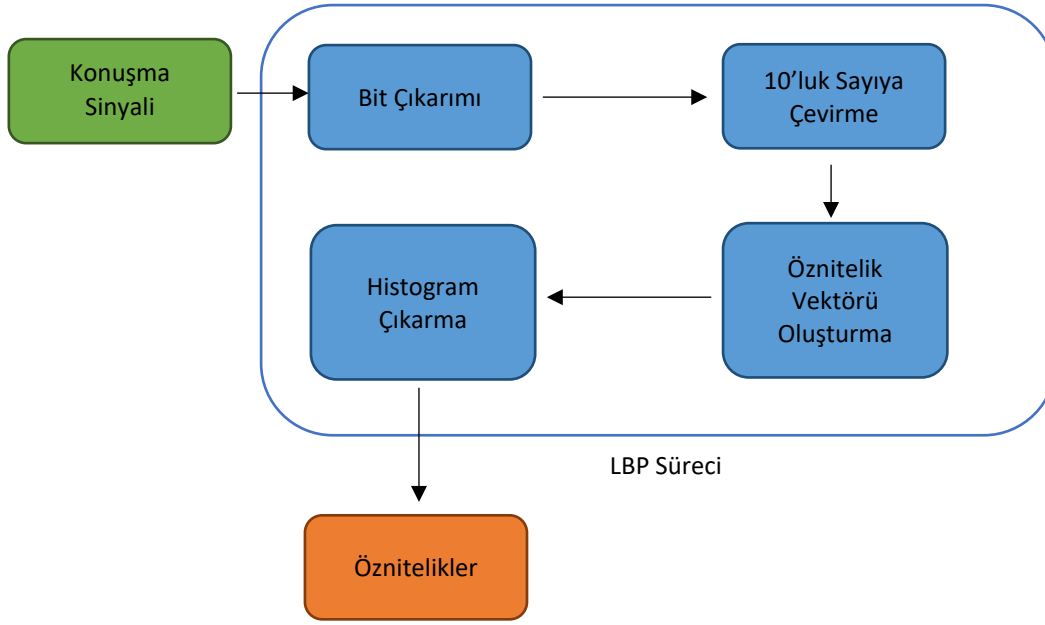
3.2.2 Yerel İkili Desen (LBP)

Yerel İkili Desen (Local Binary Pattern-LBP) yöntemi bir sayı grubunu dokuz parça halinde alır. Bu grubun orta değeri ile diğer tüm sayıları karşılaştırır. Ortadaki sayıdan büyük olması durumunda bit olarak 1 sayısını ve küçük ya da eşit olması durumunda bit olarak 0 sayısını elde eder. Bu grup için elde edilen bitler yan yana yazılarak 10'luk sayı tabanına çevrilir. Bu işlem tüm sayı dizisi için sırasıyla tekrar eder ve bir öznelik vektörü oluşturulur. 0-255 aralığında değerlerden oluşan sayıların histogramı çıkartılarak öznelikler elde edilir. Sonuç olarak bu yöntem 256 öznelik çıkarmaktadır. Bu işlemin aşamaları Şekil-2'de görsel olarak anlatılmıştır.

4. Bulgular

Bu çalışmada Python 3.x programlama dili kullanılmıştır. Elde edilen özneliklerden yapılan sınıflandırma işlemlerinde ise sklearn kütüphanesinden yararlanılmıştır. Bu kütüphane altında farklı sınıflandırma algoritmaları vardır ve birçok sınıflandırma algoritması ile öznelikler makineye öğretilmiştir. Elde edilen sonuçların değerlendirilmesi için doğruluk oranı, bulma, tutturma ve F-Ölçüsü dikkate alınmıştır.

Yukarıda verilen formülleri içindeki dp değeri doğru pozitifleri dn değeri ise doğru negatifleri ifade ederken, yp değeri yanlış pozitifleri ve yn değeri ise yanlış negatifleri ifade etmektedir. Verilerin eğitim ve test için parçalanmasında rastgele seçilmiş %30'u test için ve %70'i de eğitim için kullanılmıştır. Başarı ölçütlerindeki değerler dikkate alınarak



Şekil-2: LBP ile Öznitelik Çıkarımı

seslerden kişi tanıma için LBP ile çıkartılan öznitelikler için Çizelge-2'deki sonuçlar elde edilmiştir.

Çizelge-2: LBP ile sınıflandırma sonuçları

Yöntem	Başarı Oranı	Bulma	Tutturma	F1
discriminant_analysis.LinearDiscriminantAnalysis	0,9074	0,9074	0,9074	0,9074
ensemble.ExtraTreesClassifier	0,8984	0,8984	0,8984	0,8984
ensemble.RandomForestClassifier	0,8973	0,8973	0,8973	0,8973
ensemble.BaggingClassifier	0,8826	0,8826	0,8826	0,8826
linear_model.LogisticRegressionCV	0,8826	0,8826	0,8826	0,8826
ensemble.GradientBoostingClassifier	0,8736	0,8736	0,8736	0,8736
linear_model.RidgeClassifier	0,7844	0,7844	0,7844	0,7844
linear_model.RidgeClassifierCV	0,7844	0,7844	0,7844	0,7844
neighbors.KNeighborsClassifier	0,7336	0,7336	0,7336	0,7336
ensemble.VotingClassifier	0,7156	0,7156	0,7156	0,7156
tree.ExtraTreeClassifier	0,6930	0,6930	0,6930	0,6930
neural_network.MLPClassifier	0,6569	0,6569	0,6569	0,6569
linear_model.LogisticRegression	0,6422	0,6422	0,6422	0,6422
naive_bayes.MultinomialNB	0,6366	0,6366	0,6366	0,6366
svm.LinearSVC	0,6219	0,6219	0,6219	0,6219
neighbors.NearestCentroid	0,4334	0,4334	0,4334	0,4334
naive_bayes.BernoulliNB	0,3837	0,3837	0,3837	0,3837
svm.SVC	0,3036	0,3036	0,3036	0,3036
tree.DecisionTreeClassifier	0,2777	0,2777	0,2777	0,2777
linear_model.SGDClassifier	0,2246	0,2246	0,2246	0,2246
linear_model.Perceptron	0,1930	0,1930	0,1930	0,1930
linear_model.PassiveAggressiveClassifier	0,1761	0,1761	0,1761	0,1761
ensemble.AdaBoostClassifier	0,1524	0,1524	0,1524	0,1524
discriminant_analysis.QuadraticDiscriminantAnalysis	0,1332	0,1332	0,1332	0,1332
naive_bayes.ComplementNB	0,1106	0,1106	0,1106	0,1106

LBP ile elde edilen özniteliklerin sınıflandırılmasında LinearDiscriminantAnalysis algoritması ile %90,74 oranında başarılı bir sınıflandırma elde edilmiştir. Aynı veri kümesi için MFCC ile öznitelikler çıkarılmış ve yine farklı sınıflandırma algoritmaları ile makineye öğretilmiştir. Elde edilen sonuçlar Tablo 3'te verilmiştir.

Çizelge-3: MFCC ile sınıflandırma sonuçları

Yöntem	Başarı Oranı	Bulma	Tutturma	F1
ensemble.ExtraTreesClassifier	0,3510	0,3510	0,3510	0,3510
ensemble.RandomForestClassifier	0,3273	0,3273	0,3273	0,3273
svm.SVC	0,2856	0,2856	0,2856	0,2856
ensemble.GradientBoostingClassifier	0,2844	0,2844	0,2844	0,2844
ensemble.VotingClassifier	0,2765	0,2765	0,2765	0,2765
ensemble.BaggingClassifier	0,2381	0,2381	0,2381	0,2381
neighbors.KNeighborsClassifier	0,2111	0,2111	0,2111	0,2111
tree.ExtraTreeClassifier	0,1603	0,1603	0,1603	0,1603
neighbors.NearestCentroid	0,1309	0,1309	0,1309	0,1309
ensemble.AdaBoostClassifier	0,1174	0,1174	0,1174	0,1174
tree.DecisionTreeClassifier	0,0937	0,0937	0,0937	0,0937
linear_model.LogisticRegressionCV	0,0813	0,0813	0,0813	0,0813
discriminant_analysis.LinearDiscriminantAnalysis	0,0767	0,0767	0,0767	0,0767
svm.LinearSVC	0,0722	0,0722	0,0722	0,0722
linear_model.LogisticRegression	0,0576	0,0576	0,0576	0,0576
naive_bayes.BernoulliNB	0,0564	0,0564	0,0564	0,0564
neural_network.MLPClassifier	0,0564	0,0564	0,0564	0,0564
linear_model.Perceptron	0,0474	0,0474	0,0474	0,0474
linear_model.RidgeClassifier	0,0418	0,0418	0,0418	0,0418
linear_model.RidgeClassifierCV	0,0418	0,0418	0,0418	0,0418

linear_model.SGDClassifier	0,0316	0,0316	0,0316	0,0316
linear_model.PassiveAggressi veClassifier	0,0248	0,0248	0,0248	0,0248
discriminant_analysis.Quadrat icDiscriminantAnalysis	0,0192	0,0192	0,0192	0,0192

MFCC ile çıkartılan özniteliklerin sınıflandırılmasında Extra Trees Classifier algoritması ile en yüksek başarı oranı olarak %35,10 elde edilmiştir. Diğer algoritmalar ise daha düşük başarı oranlarına sahiptir. LBP ile kişi tanıma çalışmasında kullanılan Linear Discriminant Analysis algoritması aynı zamanda kişi doğrulama için de kullanılmıştır. Yalnızca 1 numaralı kişi için %97,85 oranında başarı elde edilmiştir. Bunun dışındaki tüm kişilerin doğrulanmasında LBP ile %100 sınıflandırma başarıları elde edilmiştir.

5. Tartışma ve Sonuç

MFCC ile çıkartılan özniteliklerin sınıflandırma başarıları ile LBP ile çıkartılan özniteliklerin başarıları arasında çok büyük farklılıklar olduğu bu çalışmada görülmüştür. Bu durum LBP ile seslerden kişi tanımanın daha başarılı olduğunu göstermektedir. %90,74 oranındaki başarılı sınıflandırma yüksek bir değerdir. Bu çalışmada öznitelik çıkarımı sırasında farklı bir filtre ya da gürültü azaltma gibi önleme uygulanmamıştır. Tiwari [11] ise MFCC ile kişi tanımda farklı filtreler uygulayarak en yüksek %85 başarıya ulaşabilmiştir. Buradaki çalışmada MFCC algoritması için varsayılan parametreler kullanılmıştır. Farklı parametreler ile MFCC'nin başarısını arttıran çalışmalar vardır [12]. Yutai ve diğ. [13] tarafından farklı filtreler ile MFCC'nin başarıları %5-6 oranında daha yüksek başarı elde edilmiştir. Dhingra, Nijhawan ve Pandit yaptıkları çalışmalarında yine MFCC ile sestan konuşmacı tanıma için %83 ve ses tanıma için %60 başarı elde edilmiştir. Söz konusu çalışmalar buradaki yöntemle kıyasla çok düşük sonuçlar elde etmiştir. Sengupta ve diğ. [15] tarafından yapılan akciğer seslerinin sınıflandırılması çalışmasında da LBP yönteminin MFCC yöntemine kıyasla daha yüksek başarı gösterdiği bulunmuştur. Bebek ağlama seslerinden tanımanın yapıldığı bir çalışmada MFCC öznitelik çıkarma yöntemi ile Linear Frequency Cepstral Coefficient (LFCC) yöntemi kullanılarak sınıflandırma yapılmış ve LFCC yönteminin MFCC yöntemine kıyasla daha yüksek başarı elde ettiği görülmüştür [24]. Bu anlamda bu çalışmanın sonuçları buradaki sonuçlarla benzerlik göstermektedir. Buradaki veriler içinde gürültüler vardır. MFCC yöntemi gürültülü verilerde sağlıklı değildir [25]. Bu nedenle de MFCC yönteminin daha düşük başarı göstermesine neden olduğu söylenebilir. Bu çalışma ile LBP yönteminin MFCC yöntemine kıyasla gürültülü ortamlarda dahi olsa daha yüksek başarı gösterdiği söylenebilir.

Kaynakça

- [1] Abdul Z. K., "Kurdish speaker identification based on one dimensional convolutional neural network," *Computational Methods for Differential Equations*, vol. 7, no. 4 (Special Issue), pp. 566-572, 2019.
- [2] Patel K. ve Prasad R., "Speech recognition and verification using MFCC & VQ," *Int. J. Emerg. Sci. Eng.(IJESE)*, vol. 1, no. 7, pp. 137-140, 2013.
- [3] Kumar C. S. ve Rao P. M., "Design of an automatic speaker recognition system using MFCC, vector quantization and LBG algorithm," *International Journal on Computer Science and Engineering*, vol. 3, no. 8, p. 2942, 2011.
- [4] Le-Qing L., "Insect sound recognition based on mfcc and pnn," in *2011 International Conference on Multimedia and Signal Processing*, 2011, vol. 2: IEEE, pp. 42-46.
- [5] Wanli Z. ve L. Guoxin L., "The research of feature extraction based on MFCC for speaker recognition," in *Proceedings of 2013 3rd International Conference on Computer Science and Network Technology*, 2013: IEEE, pp. 1074-1077.
- [6] Bimbot F. et al., "A tutorial on text-independent speaker verification," *EURASIP Journal on Advances in Signal Processing*, vol. 2004, no. 4, pp. 1-22, 2004.
- [7] Singh S., "Forensic and Automatic Speaker Recognition System," *International Journal of Electrical & Computer Engineering (2088-8708)*, vol. 8, no. 5, 2018.
- [8] Kinnunen T. ve Li H., "An overview of text-independent speaker recognition: From features to supervectors," *Speech communication*, vol. 52, no. 1, pp. 12-40, 2010.
- [9] Larcher A, Lee K. A., Ma B., ve Li H., "Text-dependent speaker verification: Classifiers, databases and RSR2015," *Speech Communication*, vol. 60, pp. 56-77, 2014.
- [10] Sanjaya M. ve Salleh Z., "Implementasi Pengenal Pola Suara Menggunakan Mel-Frequency Cepstrum Coefficients (Mfcc) Dan Adaptive Neuro-Fuzzy Inferense System (Anfis) Sebagai Kontrol Lampu Otomatis," *ALHAZEN Journal of Physics*, vol. 1, no. 1, pp. 43-54, 2014.
- [11] Tiwari V., "MFCC and its applications in speaker recognition," *International journal on emerging technologies*, vol. 1, no. 1, pp. 19-22, 2010.
- [12] Bansal P., Imam S. A., ve Bharti R., "Speaker recognition using MFCC, shifted MFCC with vector quantization and fuzzy," in *2015 International Conference on Soft Computing Techniques and Implementations (ICSCTI)*, 2015: IEEE, pp. 41-44.
- [13] Yutai W., Bo L., Xiaoqing J., Feng L., ve Lihao W., "Speaker recognition based on dynamic MFCC parameters," in *2009 International Conference on Image Analysis and Signal Processing: IEEE*, pp. 406-409, 2009.
- [14] Ohini Kafui T., ve Mignotte M., "Environmental sound classification using local binary pattern and audio features collaboration." *IEEE Transactions on Multimedia*, 2020, vol. 23: pp. 3978-3985.
- [15] Sengupta, N., Sahidullah, M., & Saha, G. 2017. "Lung sound classification using local binary pattern". arXiv preprint arXiv:1710.01703.
- [16] ER, M.B. "Heart sounds classification using convolutional neural network with 1D-local binary pattern and 1D-local ternary pattern features". *Applied Acoustics*, 2021, vol. 180: 108152.
- [17] Yang, W., Krishnan, S. "Combining temporal features by local binary pattern for acoustic scene classification." *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2017, vol. 25, no. 6: 1315-1321.
- [18] Abidin, S., Togneri, R., & Sohel, F. "Spectrotemporal analysis using local binary pattern variants for acoustic scene classification." *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 11, 2112-2121, 2018.
- [19] Yang L., Chen X. ve Tao L., "Acoustic scene classification using multi-scale features", *Proc. Detection Classification Acoustic Scenes Events (DCASE)*, pp. 29-33, 2018.

- [20] Deng, M., Meng, T., Cao, J., Wang, S., Zhang, J., & Fan, H. "Heart sound classification based on improved MFCC features and convolutional recurrent neural networks." *Neural Networks*, 130, 22-32, 2020.
- [21] Chauhan, S., Wang, P., Lim, C. S., & Anantharaman, V. "A computer-aided MFCC-based HMM system for automatic auscultation." *Computers in biology and medicine*, 38(2), 221-233, 2008.
- [22] Rahmandani, M., Nugroho, H. A., & Setiawan, N. A. Cardiac sound classification using Mel-frequency cepstral coefficients (MFCC) and artificial neural network (ANN). In 2018 3rd International Conference on Information Technology, Information System and Electrical Engineering (ICITISEE), pp. 22-26, IEEE, 2018.
- [23] Şaşmaz E., ve Tek, F. B. "Animal sound classification using a convolutional neural network." In 2018 3rd International Conference on Computer Science and Engineering (UBMK) (pp. 625-629). IEEE, 2018.
- [24] Dewi S. P., Prasasti A. L., ve Irawan B. "The study of baby crying analysis using MFCC and LFCC in different classification methods." In 2019 IEEE International Conference on Signals and Systems (ICSigSys) (pp. 18-23). IEEE, 2019.
- [25] Leena R., Mehta S.P., Mahajan A.S.. "Dabhade Comparative Study Of MFCC And LPC For Marathi Isolated Word Recognition System." *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering* vol. 2, no. 6, p. 2133-2139, 2013.