

*Araştırma Makalesi / Research Article*

# AQDD Özelliklerine BBA Yöntemleri Uygulanarak Parkinson Hastalığının Otomatik Teşhisi

## Automatic Diagnosis of Parkinson's Disease by Applying ICA Methods to TQWT Features

Cüneyt Yücelbaş<sup>1\*</sup>, Şule Yücelbaş<sup>2</sup>

*Geliş / Received: 17/05/2019*

*Reviz / Revised: 21/06/2019*

*Kabul / Accepted: 05/07/2019*

**Ö**z- Parkinson hastalığı dopamin üreten beyin hücrelerinin kaybı sonucunda oluşan bir hastalıktır. Bu hastalığın birçok teşhis yöntemi bulunmakta olup ses sinyallerinin analizi de bunlardan birisidir. Bu çalışmada daha önceden 188 Parkinson hastası ve 64 sağlıklı olmak üzere toplam 252 kişiye ait kaydedilmiş ses sinyallerinden ayarlanabilir Q-faktör dalgacık dönüşümü (AQDD) metodu kullanılarak elde edilen özellikler kullanılmıştır. Bu özelliklere bağımsız bileşen analizi (BBA) çeşitlerinden olan hızlı BBA (HBBA), max-kurtosis BBA (KBBA) ve yeniden yapılanma BBA (YBBA) olmak üzere üç farklı özellik azaltma (boyut indirgeme) yöntemi uygulanmıştır. Bu işlemler sonucunda minimum özellik sayısı ile maksimum başarı oranı elde edilmeye çalışılmıştır. Bu amaçla, öncelikle yeni özellikler ile oluşturulan veri grubuna ayrı ayrı k-kat çapraz doğrulama yöntemi uygulanarak veriler eğitim-test olarak ayrılmıştır. Sonraki aşamada, hazırlanan veriler rastgele orman (RO) algoritması ile sınıflandırılmış ve sonuçlar çeşitli istatistiksel ölçütlerle yorumlanmıştır. Sonuçlar değerlendirildiğinde; kullanılan boyut indirgeme yöntemleri içerisinde en başarılı yöntem %82.01 sınıflandırma doğruluk oranı ve yaklaşık 0.85 ROC ve PRC değerleri ile YBBA olmuştur. Bu durum hasta ve sağlıklı sınıf ayrışımının mükemmel yaklaştığını kanıtlamıştır. Gerçek yaşam uygulamalarına uygun olan bu çalışmanın performans sonuçları ve kullanılan veri sayısının yüksek oluşu çalışmanın literatürdeki önemini ortaya koymaktadır. Ayrıca, çalışma kapsamında kullanılan özellik indirgeme yöntemlerinin analizi, bu alanda yapılabilecek çalışmalara yol gösterebilecek niteliktedir.

**Anahtar Kelimeler-** Parkinson Hastalığı, Ses Sinyali Analizi, Boyut İndirgeme Metotları, AQDD

**A**bstract- Parkinson's disease is occurred due to loss of brain cells which are releasing dopamine. There are many different techniques to diagnosis of this disease and analysis of audio signals is one of them. In this study, the features obtained by using the tunable Q-factor wavelet transform (TQWT) method from the previously recorded audio signals of 252 subjects (188 of data belongs to Parkinson's disease and the others are healthy). Three different feature reduction (size reduction) methods such as fast ICA (FICA), max-kurtosis ICA (KICA) and reconstruction ICA (RICA), which are independent component analysis (ICA), were applied to these properties. We aimed to obtain maximum success rate with minimum number of features using these operations. To do this, firstly the k-fold cross validation method was applied to the data group created with new features and was split into training and test group. The data, then, was classified with random forest (RO) algorithm and the results were evaluated with various statistical criteria. It was concluded that RICA with 82.01% classification accuracy rate and approximately 0.85 ROC and PRC values are most successful method among the size reduction methods. This result showed that patient and healthy class separation is approaching excellence. The performance results of this study, which is suitable for real life applications and the high number of data used, reveal the importance of the study in the literature. Moreover, the analysis of feature reduction methods used in this study is high enough to guide the future studies in this field.

**Keywords-** Parkinson's Disease, Audio Signal Analysis, Dimension Reduction Methods, TQWT

<sup>1\*</sup>Sorumlu yazar iletişim: cuneytyucelbas@hakkari.edu.tr (<https://orcid.org/0000-0002-4005-6557>)

Elektrik-Elektronik Mühendisliği Bölümü, Hakkari Üniversitesi, Merkez, Hakkari, Türkiye

<sup>2</sup>İletişim: suleyucelbas@hakkari.edu.tr (<https://orcid.org/0000-0002-6758-8502>)

Elektrik-Elektronik Mühendisliği Bölümü, Hakkari Üniversitesi, Merkez, Hakkari, Türkiye

## I. GİRİŞ

Parkinson hastalığı (PH), dopamin üretmekle görevli beyin hücrelerinin azalması neticesinde ortaya çıkan bir rahatsızlıktır [1, 2]. Bu hücrelerin kaybı sonucunda yeteri kadar dopamin üretilemez. Bunun sonucunda hareketlerde yavaşlama, titreme ve kaslarda sertlik gibi belirtiler ortaya çıkar [1, 2]. Erkeklerde kadınlara nazaran daha sık karşılaşılan bu rahatsızlık, çoğunlukla 40 yaş sonrasında ortaya çıkmaktadır. Bir başka faktör ise kişinin ailesine ait PH geçmiştir. Bu hastalık hareket bozukluklarına neden olup, zamanla belirtiler daha belirgin bir şekilde kendini göstermektedir. PH ile dopamin maddesi doğrudan ilişkili olduğu için bu maddenin eksikliği arttıkça rahatsızlığa ait belirtiler de artmakta ve sonuç olarak artık vücut hareketleri kontrol edilemez olmaktadır. Bu hastalığa ait belirtiler kişisel olmasına rağmen çoğu kişide vücudun tek tarafında ilk bulgular kendini göstermeye başlar ve yaş ilerledikçe vücudun diğer bölgelerinde de hareket bozuklukları ortaya çıkar [1-4]. PH, kasları büyük oranda etkilediği için kişilerin yüzlerindeki mimik hareketlerinde de değişime sebep olur. Bu durum hastalarda konuşma ve telaffuz bozukluklarını ortaya çıkarır [1-4].

Bu hastalığın teşhisinde ilk olarak kişinin aile geçmişine, daha sonra ise tıbbi ve fiziki muayene bulgularına bakılır [5, 6]. Tedavide asıl amaç kişinin, başkalarının desteğine gerek kalmadan aktif olarak kendi kendine günlük işlerini yapabilmesini sağlamaktır. Bunun için hastalarda görülen belirtilerin medikal destek ile kontrol altına alınması hedeflenmektedir [5, 6]. Uygulanan başlangıç tedavi yöntemleri arasında denge egzersizlerinin yanında konuşma ve dil terapistleri eşliğinde gerekli uygulamaların düzenli olarak gerçekleştirilmesi yer almaktadır. Hastalığın tedavisinde cerrahi seçenek her zaman son başvurulacak durumdur [5-7]. Tüm bu medikal destek ve yöntemlere rağmen hastada herhangi bir ilerleme meydana gelmemesi durumunda cerrahi işlem uygulanabilmektedir [7].

Tıp alanında gerçekleştirilen çalışmalar genellikle medikal uygulama, cerrahi işlem ve bir takım yöntemlerin hastalar üzerindeki etkisinin incelenmesi üzerinedir. PH' nin yapay zekânın aktif şekilde kullanılarak uzman sistemler tarafından otomatik tespitine çalışıldığı alan ise mühendislik olmuştur. Mühendislik alanında bu amaç için gerçekleştirilen çalışmaların büyük bir bölümünde kişilere ait ses ve yürüme kayıtlarından yararlanılmıştır. Bu rahatsızlığa sahip kişilerin yaklaşık %90' ında hastalığın en önemli erken evre belirtilerinden biri olan konuşma bozuklukları görülmektedir [8, 9]. Bu sebepten dolayı, kişilere ait seslerin kaydedilmesi ve bu kayıtların kullanılması uzman sistemler için önem arz etmektedir. Mühendislik alanında gerçekleştirilen çalışmalarda genellikle Parkinson hastalarının sağlıklı kişilerden otomatik olarak ayırt edilmesi amaçlanmıştır [9-14].

Bu alanda gerçekleştirilen bir çalışmada PH' nin otomatik tespiti amacıyla gerekli veriler kaydedilmiş, bunlardan ilgili özellikler çıkarılarak destek vektör makinesi (DVM) sınıflayıcısına verilerek sonuçlar değerlendirilmiştir [10]. Farklı bir veri setinin kullanıldığı bir diğer çalışmada ise, dört farklı sınıflayıcı sistem aynı amaç için kullanılmış ve en iyi sonuç %92.9 ile yapay sinir ağları (YSA)'nda elde edilmiştir [15]. Buna benzer birçok araştırma çalışmasında YSA, DVM, çok katmanlı YSA, genel regresyon YSA ve en küçük kareler DVM (EKK-DVM) gibi sınıflandırma sistemleri kullanılmıştır [16, 17]. Sonuç olarak en iyi performansın EKK-DVM' de elde edildiği görülmüştür. Bir başka çalışmada, hastalığın teşhisi amacıyla çıkarılan vokal özelliklerin etkinliği analiz edilmiş ve en iyi sonuç %96.4 ile DVM' de elde edilmiştir [18]. Benzer bir çalışmada ise, yine aynı şekilde otomatik teşhis sistemi için farklı hazır veri setleri üzerinde araştırmalar yapılmıştır [19]. Mühendislik alanındaki bu çalışmalarda kullanılan verilerin kaydedildiği kişi sayılarının [9, 17, 18] için 42, [11-13, 15-16] için 31, [10] için 32, [14] için 40 ve [19] için ise 70 olduğu tespit edilmiştir.

Bu tür çalışmalarda ilk olarak hasta ve sağlıklı kişilerden uzman eşliğinde veri alımı gerçekleştirilir. Daha sonra bu veriler çeşitli yöntemlerle işlenir ve gerekli olabilecek özellikler çıkarılarak ihtiyaca göre özellik seçme işlemi uygulanır. Sonraki aşama, hazırlanan veri setlerinin uzman sisteme girdi olarak verilmesidir. Son olarak ise, uzman sistemler eşliğinde sınıflandırma işlemi gerçekleştirilerek elde edilen sonuçlar yorumlanır. Çalışmaların tamamında ortak amaç, yüksek sınıflandırma sonuçlarının elde edilmesidir.

Bu çalışmada, İstanbul Üniversitesi Cerrahpaşa Tıp Fakültesi Nöroloji Bölümünde kaydedilmiş 188' i Parkinson hastası olmak üzere toplam 252 kişiye ait kaydedilmiş ses sinyallerinden AQDD yöntemi ile elde edilen özellikler kullanılmıştır [20]. Bu özelliklere çalışma kapsamında üç farklı bağımsız bileşen analizi (BBA) yöntemi (HBBA, KBBA, YBBA) uygulanarak bunların otomatik Parkinson teşhisi üzerindeki etkisi araştırılmıştır. Çalışmada minimum özellik ile maksimum sınıflandırma performansı elde edilmesi amaçlanmıştır. Bu amaçla eğitim ve test veri gruplarını oluşturmak için veriler üzerinde 10-kat çapraz doğrulama yöntemi uygulanmıştır. Daha sonra hazırlanan bu veri grupları hem hızlı hem de yüksek doğrulukta performans sergileyen rastgele orman (RO) algoritması ile sınıflandırılmış ve sonuçlar bazı istatistiksel ölçütlerle (doğru pozitif oranı-DPO, yanlış pozitif oranı-YPO, Hassasiyet, F puanı-F, ROC alanı, PRC (hassasiyet-DPO) alanı ve

ortalama sınıflandırma doğruluğu-OSD) yorumlanmıştır. Sonuç olarak; en başarılı BBA yönteminin %82.01 OSD ve yaklaşık 0.85 ROC ve PRC değerleri ile YBBA olduğu anlaşılmıştır. Gerçek yaşam uygulamalarına uygun olan bu çalışmanın performans sonuçları ve kullanılan veri sayısının yüksek oluşu çalışmanın literatürdeki önemini ortaya koymaktadır. Ayrıca, çalışma kapsamında kullanılan özellik indirgeme yöntemlerinin analizi, bu alanda yapılabilecek çalışmalara yol gösterebilecek niteliktedir. Çalışmanın bir başka amacı ve özgün yanı ise çeşitli yöntemlerin kullanıldığı otomatik teşhis sistemi sayesinde uzmanlar arasındaki görüş farklılığının en aza indirgenmesi olmuştur.

## II. MATERYAL VE METOD

PH, kasları etkileyen nörolojik bir rahatsızlık olduğu için beynin konuşmayı kontrol eden kasları da bundan etkilenmektedir. Bu sebeple hastaların yüz kasları sertleşir ve bazı seslerin söylenmesi zorlaşır [21]. Hastalığın bu belirtisinden dolayı kişilerin ses kayıtları kullanılarak otomatik bir şekilde hasta-sağlıklı ayrımı yapılabilir. Bu amaç doğrultusunda çalışma kapsamında kullanılan veriler [20] uzman doktor eşliğinde her deneğe /a/ harfi 3 kere tekrarlatılarak 44.1 KHz frekansa ayarlanmış mikrofon aracılığı ile kayıt altına alınmıştır [20]. Verilerin alındığı kaynaktan [20] araştırmacılar en iyi sonucu AQDD yönteminde elde ettiklerinden dolayı bu çalışmada da aynı veriler üzerinde analizler yapılmıştır. AQDD, Q (Q-faktör), j (seviye sayısı) ve r (fazlalık) olmak üzere 3 temel parametreden oluşan yeni bir ayırık dalgacık yöntemidir [22].

Bu çalışmada, hasta ve sağlıklı toplam 252 kişinin ses kayıtlarına uygulanan AQDD yönteminin uygulanmasıyla elde edilen 36 alt-bandın her birinden çıkarılan 12 farklı özellik grubu kullanılmıştır. Bu alt-bantların her birinden ayrı ayrı enerji, Shannon entropi, log enerji entropi, Teager-Kaiser enerji operatörü (TKEO) ortalaması, TKEO standart sapması, medyan, ortalama, standart sapma, minimum, maksimum, skewness (çarpıklık) ve kurtosis (basıklık) değerleri olmak üzere 12 farklı özellik çıkarılmıştır [26] [36-39]. Böylelikle 432 (12x36) adet parametre her bir denek için elde edilmiştir. Ayrıca her bireye ilgili sesler üç kez tekrarlatılarak kayıt altına alındığı için üzerinde çalışılan veri sayısı 756 (252x3) olarak ele alınmıştır. Tüm bu işlemlerin ardından sistem dahilindeki veri grubu 756x432 olarak elde edilmiştir. Bu veri grubu üzerinde üç farklı BBA yöntemi (HBBA, KBBA, YBBA) uygulandıktan sonra boyutu değiştirilmiş veriler 10-kat çapraz doğrulama yöntemi aracılığıyla eğitim-test gruplarına ayrılmıştır. Son aşama olarak ise, elde edilen bu veriler RO sınıflandırıcısına sunularak sonuçlar analiz edilmiştir.

### A. Kullanılan BBA Yöntemleri

BBA ses işleme, görüntü işleme gibi alanlarda verilerden özellik çıkarma veya çıkarılan özelliklerin boyutlarını değiştirme amacıyla kullanılan istatistiksel bir sinyal işleme yöntemidir [23]. Sinyal işleme alanında sıklıkla kullanılan bu yöntem ilk kez kör kaynak ayırma tekniği olarak ortaya konmuştur [24, 25]. Bu yöntemde karışık grup sinyalleri içerisinde istatistiksel olarak bağımsız olan kaynak sinyali ayırmak asıl amaçtır [25, 26]. BBA'ya ait matematiksel model aşağıdaki gibi tanımlanabilir [24]:

$$X_{m \times 1} = H_{m \times n} \times S_{n \times 1} \quad (1)$$

burada,  $S_{n \times 1}$  bilinmeyen kaynak sinyali ifade etmektedir.  $H_{m \times n}$  karışık matrisi, son olarak  $X_{m \times 1}$  ise gözlenen sinyalden oluşan sütun vektörünü temsil etmektedir. BBA yönteminde hedef,  $W_{n \times m}$  ayırık matrisinin hesaplanmasıdır. Bu şekilde ayrılmış olan hedefteki orijinal sinyal  $Y_{n \times 1}$  tahmin edilebilir.

$$Y_{n \times 1} = W_{n \times m} \times X_{m \times 1} \quad (2)$$

Gerçekleştirilen çalışma kapsamında HBBA, KBBA ve YBBA yöntemleri özellikleri çıkarılmış veriler üzerinde uygulanarak sonuçlar analiz edilmiştir.

Kullanılan ilk yöntemlerden HBBA, gözlem sinyallerinin ön hazırlığı ve bağımsız bileşenlerin çıkarılması olmak üzere iki kısımdan meydana gelmektedir. Bu kısımlardan ön hazırlık aşaması veri merkezleme ve beyazlaştırma işlemlerinden oluşmaktadır. İlk adım olan veri merkezlemede gözlem sinyalinin ortalaması hesaplanır ve bu ortalama değer gözlenen veriden çıkarılarak sıfır ortalama işlemi gerçekleştirilmiş olur. Beyazlatma aşamasında ise veri merkezleme sonucunda oluşturulan ortalanmış verinin kovaryans matrisi hesaplanır. Daha sonra bu matrisin öz-değer ayrışımı gerçekleştirilir. Bu aşamanın amacı, bağımsız bileşenlerin çıkarımını basitleştirmektir. Bu işlemlerden sonra gerekli olan bileşenler çıkarılır [23-26].

BBA için rastgele değişkenin gauss olmayan tahmini için kullanılan klasik ölçümlerden biri de kurtosis [27]. BBA yönteminin hedef fonksiyonları genel olarak gaussluğu olmayan (kurtosis ve negentropi), karşılıklı bilginin minimizasyonu, infomax metodu ve maksimum olabilirlik tahminini içermektedir. Bu çalışmada KBBA ele alınmıştır. KBBA, gauss olmayan bir ölçü kurtosisidir. Sıfır ortalamalı ve kurtosisli rastgele bir değişken olarak ifade edilebilir [28]:

$$kurt(\bar{y}(t)) = E\{(\bar{y}(t))^4\} - 3(E\{(\bar{y}(t))^2\})^2 \quad (3)$$

Buna göre; eğer rastgele gauss değişkenleri olur ise  $y^*(t)$ 'nin kurtosisi 0' a eşit olur. Oysaki gauss olmayan bir değişkenin kurtosisi genelde sıfıra eşit değildir. O halde,  $y^*(t)$ 'nin kurtosisi sıfırdan büyük olduğu durumlarda  $y^*(t)$  süper bir gauss değişkeni gösterirken, küçük olduğunda ise bir gauss alt değişkeni ifade etmektedir.  $y^*(t)$ 'nin gauss olmaması durumu büyüdükçe kurtosisin mutlak değeri de büyümektedir.

$Kurt(y^*(t)) > 0$  olduğunda,  $y^*(t)$  bir süper Gauss değişkeni gösterirken,  $(y^*(t)) < 0$ ,  $y^*(t)$  kurtosisi ise bir Gauss alt değişkenini gösterir. Gaussianlık  $y^*(t)$  arttıkça, kurtosisin mutlak değeri de artar. Böylelikle, BBA problemi mutlak değer maksimumunu veya kaynak sinyal bileşenine ait kurtosisinin karesinin hesaplanmasına dönüştürülebilir [27-29].

Çalışma kapsamında son olarak YBBA yöntemi kullanılmıştır. Bu metot geleneksel BBA oluşumunun yeniden tasarlanması sonucunda ortaya çıkmıştır. Aşağıda optimizasyon problemi olarak tipik bir BBA formülasyonu görülmektedir [30]:

$$\min_w \sum_{i=1}^m \sum_{j=1}^k g(W_j x^{(i)}), \quad WW^T = I \quad (4)$$

burada;  $g$  doğrusal olmayan bir dışbükey fonksiyondur.  $W$  ağırlık matrisi olup;  $k$  ise özellik sayısıdır.  $W_j$ , ağırlık matrisi içerisinde bir satır özellik vektörüdür.  $WW^T = I$  ise ortonormallik kısıtı olarak kullanılmıştır.

Denklemin 4'ün revize edilmesi sonucunda aşağıdaki YBBA meydana gelmiştir [31]:

$$\min_w \frac{\lambda}{m} \sum_{i=1}^m \|W^T W x^{(i)} - x^{(i)}\|_2^2 + \sum_{i=1}^m \sum_{j=1}^k g(W_j x^{(i)}) \quad (5)$$

Yapılan çalışmalar göstermiştir ki; YBBA metodu geleneksel BBA ile karşılaştırıldığında pratikte daha hızlı ve iyi çalışmaktadır [31]. Bu çalışma neticesinde elde edilen sonuçlara göre bu durumun doğruluğu anlaşılmaktadır.

### B. Sınıflandırma İşlemi ve Performans Ölçütleri

Yapay zekâ algoritmalarının kullanılarak oluşturulduğu uzman sistemler bir hekim gibi hasta-sağlıklı ayrışımını kolay ve yüksek doğruluklarda yapabilmektedir. Bu işlem genelde uzman sistem bünyesindeki bir sınıflandırıcı aracılığıyla gerçekleştirilmektedir. Ayrıca bu sistemler sayesinde elde edilen sonuçlar ayrıntılı bir şekilde analiz edilebilmektedir.

Bu amaç doğrultusunda çalışma kapsamında birçok sınıflayıcı (Yapay Sinir Ağları: YSA, Sıralı Minimal Optimizasyon temelli Destek Vektör Makineleri: SMO-DVM, karar ağacı J48: J48, RO) denenmiş ancak en iyi sonucun elde edildiği RO algoritması seçilmiştir [32]. Bu algoritmanın çalıştırılması ve performans değerlendirmesi Weka 3.8 yazılımı [33] bünyesinde gerçekleştirilmiştir. Üzerinde çalışılan veriler öncelikle 10-kat çapraz doğrulama metodu ile eğitim-test veri gruplarına ayrılmıştır. Veriler bünyesinde hastalar 0, sağlıklılar ise 1 olarak etiketlenmiştir. Sistem dahilinde kullanılan RO sınıflayıcısının öğrenme parametreleri varsayılan (BagSizePercent:100, BatchSize:100, MaxDepth:0, NumDecimalPlaces:2, NumExecutionSlots:1, NumFeatures:0, NumIterations:100 ve Seed:1) olarak seçilmiş ve işlemler bu değerler altında gerçekleştirilmiştir.

Çalışma kapsamında kullanılan performans ölçütleri otomatik tespit sistemi tarafından doğru ve yanlış olarak belirlenen veri sayılarına göre oluşturulan karmaşıklık matrisine göre yapılmıştır. Bu matristen yararlanılarak DPO, YPO, Hassasiyet, F, ROC alanı, PRC alanı ve OSD ölçütleri elde edilmiştir [34-38].

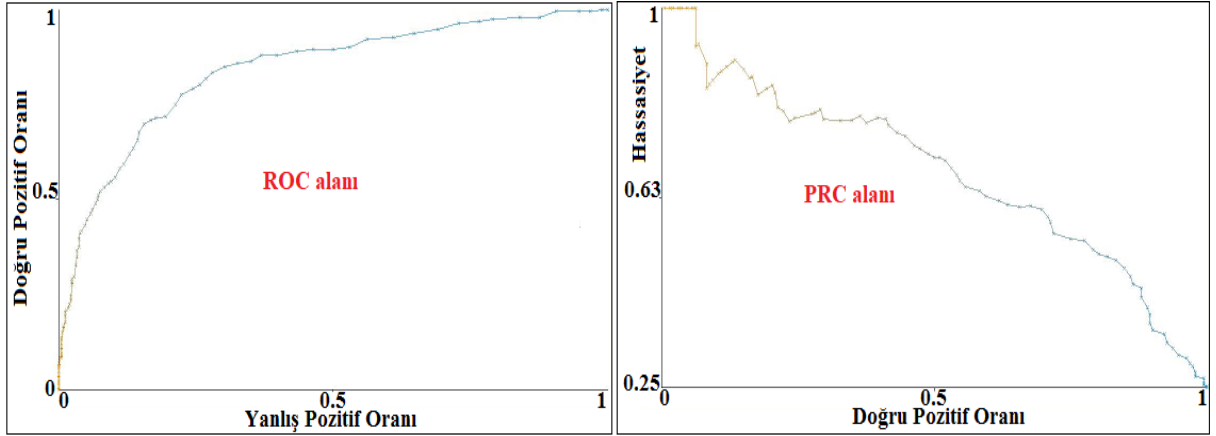
## III. DENEYSEL SONUÇLAR

Bu çalışmada daha önceden kaydedilmiş 188' i Parkinson hastası toplam 252 denekten kaydedilmiş verilerden çıkarılmış özellik grupları arasından en iyi sonucun elde edildiği AQDD metoduna ait özellik grubu kullanılmıştır [26]. Bu özellik grubu üzerinde gerekli düzenlemeler yapıldıktan sonra 756x432 boyutunda yeni bir veri kümesi elde edilmiştir. Daha sonra 3 farklı BBA yöntemi çeşitli boyutlar için çalıştırılmış ve her yöntem için farklı boyutlara sahip yeni veriler oluşturulmuştur. Bu kapsamda her BBA çeşidi 10, 15, 20, 25, 30, 35 ve 40 boyutları için çalıştırılmıştır. Elde edilen yeni boyutlu veriler üzerinde ayrı ayrı 10-kat çapraz doğrulama yöntemi ile eğitim-test grupları elde edilmiş ve bu veriler RO sınıflayıcısına verilmiştir. Sınıflayıcı çıkışında elde edilen sonuçlar Tablo 1' de ayrıntılı olarak görülmektedir.

**Tablo 1.** RO sınıflayıcısı tarafından BBA çeşitlerine göre elde edilen istatistiksel sonuçlar  
(Toplam Örnek Sayısı: TÖS - Doğru Sınıflandırılan Örnek Sayısı: DSÖS)

İstatistiksel parametreler	TÖS	DSÖS	DPO	YPO	Hassasiyet	F	ROC	PRC	OSD (%)
HBBA 10	756	557	0.737	0.718	0.651	0.652	0.533	0.651	73.677
HBBA 15	756	566	0.749	0.694	0.701	0.671	0.566	0.675	74.867
HBBA 20	756	557	0.737	0.746	0.583	0.635	0.538	0.650	73.677
HBBA 25	756	565	0.747	0.735	0.711	0.645	0.553	0.650	74.735
HBBA 30	756	558	0.738	0.738	0.620	0.641	0.578	0.669	73.809
HBBA 35	756	561	0.742	0.744	0.607	0.638	0.612	0.691	74.206
HBBA 40	756	563	0.745	0.740	0.659	0.642	0.617	0.695	74.470
KBBA 10	756	567	0.750	0.521	0.725	0.731	0.673	0.741	75.000
KBBA 15	756	592	0.783	0.483	0.764	0.764	0.715	0.781	78.306
KBBA 20	756	595	0.787	0.478	0.769	0.768	0.711	0.780	78.703
KBBA 25	756	599	0.792	0.466	0.776	0.774	0.726	0.792	79.232
KBBA 30	756	605	0.800	0.460	0.786	0.782	0.737	0.792	80.026
KBBA 35	756	600	0.794	0.462	0.777	0.776	0.740	0.798	79.365
KBBA 40	756	608	0.804	0.465	0.791	0.784	0.751	0.809	80.423
YBBA 10	756	620	0.820	0.436	0.812	0.802	0.840	0.861	82.010
YBBA 15	756	596	0.788	0.567	0.781	0.746	0.789	0.817	78.836
YBBA 20	756	597	0.790	0.580	0.792	0.743	0.769	0.816	78.968
YBBA 25	756	596	0.788	0.594	0.798	0.737	0.749	0.798	78.836
YBBA 30	756	579	0.766	0.664	0.764	0.694	0.744	0.781	76.587
YBBA 35	756	579	0.766	0.674	0.780	0.689	0.762	0.802	76.587
YBBA 40	756	572	0.757	0.705	0.764	0.668	0.712	0.773	75.661

Tablo 1' de de görüldüğü üzere 3 farklı BBA çeşidi arasından en iyi performans sonuçları YBBA' nın 10 boyutlu veri grubunda elde edilmiştir. Bu veri grubunda toplam 756 adet farklı girdinin 620' si doğru bir şekilde sınıflandırılarak %82.010 OSD oranına ulaşılmıştır. Ayrıca DPO, Hassasiyet ve F değerlendirme kıstasları da 1' e yaklaşarak elde edilen OSD oranını istatistiksel olarak desteklemişlerdir. Bunun yanında YPO oranının diğer yöntemler ve boyutlara nazaran düşük olması da YBBA-10 ' nun başarısını destekleyen bir diğer parametre olarak ön plana çıkmıştır. YBBA-10 yöntemi ve veri grubunu %80.423 OSD oranı ile 40 boyutlu KBBA metodu takip etmiştir. Diğer değerlendirme kıstasları ise aynı şekilde OSD sonucunu KBBA-40 için desteklemektedir. Ancak bu metotta en iyi sonuçlar 40 boyutunda elde edildiği için boyut fazlalığından dolayı sistem ağırlaşmaktadır. Asıl amaç en az boyut ile en yüksek doğruluk oranının elde edilmesiydiğundan YBBA yönteminin üstünlüğü bir kez daha ortaya çıkmıştır. En düşük performansı HBBA yöntemi göstermiş olup boyut farklılığının sonuca etkisinin olmadığı açıkça görülmektedir. YBBA-10 yöntemine ait sonuçlar ROC ve PRC ölçütleri açısından değerlendirildiğinde bu oranların Şekil 1' de de görüldüğü üzere yaklaşık 0.85 bant aralığında olması başlangıçtaki sınıfların iyi bir şekilde birbirinden ayrılmış olmasını kanıtlamaktadır.



Şekil 1. YBBA-10 yöntemi sonucunda elde edilen ROC ve PRC alan grafikleri

Çalışma kapsamında RO algoritmasının haricinde YSA, SMO-DVM ve J48 sınıflayıcıları da en iyi sonucun elde edildiği YBBA-10 yöntemi için çalıştırılarak elde edilen sonuçlar karşılaştırmalı olarak Tablo 2' de verilmiştir.

Tablo 2. YBBA-10 yöntemi için YSA, SMO-DVM, J48 ve RO sınıflayıcılarına ait karşılaştırmalı sonuçlar

Diğer Sınıflayıcılar	İstatistiksel parametreler								
	TÖS	DSÖS	DPO	YPO	Hassasiyet	F	ROC	PRC	OSD (%)
YSA	756	596	0.788	0.385	0.781	0.784	0.775	0.815	78.836
SMO-DVM	756	600	0.794	0.558	0.791	0.753	0.618	0.691	79.365
J48	756	597	0.790	0.474	0.772	0.771	0.658	0.714	78.968
RO	756	620	0.820	0.436	0.812	0.802	0.840	0.861	82.010

Tablo 2' de de görüldüğü üzere en iyi performans sonuçları RO algoritmasında elde edilmiştir. RO sınıflayıcısını %79.365 OSD oranı ile SMO-DVM sistemi takip etmiştir. Diğer sınıflayıcılarda ise birbirine yakın sonuçlara ulaşıldığı görülmüştür. Ancak ROC ve PRC parametreleri açısından sınıflayıcılar değerlendirildiğinde ise RO' dan sonra en iyi performans 0.775 ve 0.815 değerleri ile YSA' da elde edilmiştir.

#### IV. SONUÇLAR VE TARTIŞMA

PH, genelde 40 yaşından sonra görülen nörolojik bir beyin hastalığıdır. Günümüzde insanların yaşamını zorlaştıran birçok ağır hastalığın erken teşhisi kişilerin yaşam kalitesi açısından önem arz etmektedir. Parkinson da bu hastalıklardan biri olarak literatürde yerini almıştır. Teknolojinin ve bilimin gelişmesi ile hastalıkların herhangi bir uzmana gerek kalmadan otomatik olarak teşhis edilmesine yönelik araştırmalar da artmıştır. PH teşhisinde kişilere ait ses sinyallerinin analizi uzmanlar tarafından başvuru başlıca yöntemlerden birisidir [39]. Eğer bu sinyaller kullanılarak yüksek doğrulukta otomatik PH teşhis sistemi gerçekleştirilebilirse, bu uzman sistemler hekimlerin kararını destekleyebilir. Bu duruma ek olarak uzman yetersizliği bulunan bölgelerde ön araştırma amacıyla da kullanılabilir. Daha da ilerletilen çalışmalar neticesinde kişiler evlerinde çevrimiçi olarak ses kayıtlarını sisteme yüklemek suretiyle bu hastalığın kendilerinde bulunma ihtimali hakkında ön bilgi sahibi olabilirler. Bu sebepler doğrultusunda bu çalışmada hızlı ve yüksek performans sonuçlarına sahip otomatik bir PH tespit sistemi sunulmuştur. Çalışma kapsamında 252 deneğe ait ses kayıtlarından çıkarılmış etkin özellikler üzerinde 3 farklı BBA boyut indirgeme yönteminin sistem üzerindeki etkisi araştırılmış ve bu doğrultuda uzman bir sistem oluşturulmuştur. Bu sistem içerisinde karar mekanizması olarak RO algoritması kullanılmış ve sonuçlar bazı istatistik ölçütlerle değerlendirilmiştir. Analizi yapılan BBA türlerinden en iyi sonuçlar ise YBBA yönteminin 10 boyutu kullanıldığında elde edilmiştir. Bu kapsamda 756 adet girdinin 620 tanesi doğru olarak tespit edilmiş ve %82.010 sınıflandırma sonucuna ulaşılmıştır. Bunun dışında DPO, Hassasiyet ve F ölçütleri sırasıyla 0.820, 0.812 ve 0.802 değerleri ile OSD oranını onaylamıştır. Ayrıca elde edilen 0.840 ile ROC, 0.861 ile PRC ve 0.436 ile YPO sonuçları da kullanılan YBBA yönteminin performans sonuçlarını desteklemiştir. Çalışma kapsamında ayrıca en iyi sonucun elde edildiği YBBA-10 yöntemi için RO algoritmasının dışında YSA,

SMO-DVM ve J48 sınıflayıcıları da çalıştırılmıştır. Elde edilen sonuçlar değerlendirildiğinde RO' dan sonra en iyi performansa OSD açısından SMO-DVM' de, ROC ve PRC açısından ise YSA' da ulaşılmıştır.

Literatürde bu alanda yapılmış birçok çalışma [10, 15-19] mevcuttur. Bu çalışmalarda genel olarak elde edilen verilerden bazı özellikler çıkarıldıktan sonra DVM, NN, EKK-DVM gibi çeşitli sistemler aracılığı ile sınıflandırma işleminin yapılması ortak nokta olmuştur. Ancak bu çalışmaların [10, 15-19] hiçbirinde çıkarılan özellikler üzerinde sonuçlara pozitif etkisi olabilecek özellik indirgeme yöntemleri uygulanmamıştır. Bu çalışmada ise literatürden farklı olarak çeşitli BBA yöntemlerinin otomatik PH tespit sistemi üzerindeki etkinlikleri araştırılmıştır. Bu yöntemlerin kullanılması sonucunda mevcut özellik sayısının boyutuyla birlikte daha az sayıya indirgenmiş olması ve buna bağlı olarak sistem performansının yükselmiş olması çalışmanın literatürdeki önemini arttırmıştır. Bunun yanında kullanılan veri sayısı bu alandaki çalışmalara [9-19] göre oldukça yüksek seçilmiştir. Bu durum sistemlerin güvenilirliği açısından önemli bir durumdur. Çünkü bir sistem ne kadar fazla girdi için çalıştırılırsa güvenilirliği de o derecede artmaktadır. Bu açıdan değerlendirildiğinde, çalışmanın önemi bir kez daha ortaya çıkmaktadır.

Tıp alanında geliştirilen otomatik teşhis sistemlerinin ilgili alandaki uzmanlara birçok faydası bulunmaktadır. Parkinson gibi teşhisinin son derece önemli olduğu hastalıkların gerekli veriler ışığında tespit edilmesi çoğu zaman birçok uzman tarafından karıştırılmaktadır. Yani; bir uzman, kişi hakkında hastalık teşhisinde bulunurken bir başka uzman bazı verileri gözden kaçırma gibi nedenlerden dolayı aksini söyleyebilmektedir. Bu gibi durumlarda geliştirilen otomatik teşhis sistemleri sayesinde hekimlerin karar verme karmaşıklığının önüne geçilebilir. Bu çalışmanın bir başka özgül yanı ise bu gibi durumlarda ortaya çıkmakta ve geliştirilen sistemlerin uzmanların karar aşamasında sorgulayıcı rol oynamasını amaçlamaktadır. Bu çalışmanın bir diğer önemli sayılabilecek katkısı ise otomatik PH tespit sisteminde çeşitli sınıflayıcıların karşılaştırmalı olarak kullanılması ve BBA çeşitlerinin aynı sistem üzerindeki etkisinin ayrıntılı analiz edilmesi olmuştur. Kısaca özetlemek gerekirse çevrimiçi sürümlere uyarlanabilecek ve tıp uzmanları tarafından kabul görebilecek BBA temelli yüksek performanslı efektif ve etkili bir otomatik PH tespit sistemi sunulmuştur.

Bu alanda çalışan veya çalışmayı düşünen araştırmacılar farklı veri setlerine farklı yöntemleri daha fazla sayıda veriye uygulamak suretiyle yüksek performans sonuçlarına ulaşmayı hedefleyebilirler. Ayrıca geliştirilen uzman sistemlerin internet üzerinden herkesin ulaşmasına imkân veren sürümleri de ortaya konabilir. Böylelikle daha farklı ve fazla sayıda denek gruplarına ulaşmak mümkün olabilir.

#### KAYNAKLAR

- [1] K. Rana. (2014, 26 December 2018). *Parkinson Hastalığı* [Online]. Available: <http://www.noroloji.org.tr/TNDDData/Uploads/files/parkinson%20hastal%C4%B1%C4%9F%C4%B1.pdf>.
- [2] S. Özekmekçi, H. Apaydın, S. Oğuz, & İ. Zileli. (2013). *Parkinson Hastalığı Hasta ve Yakınları İçin El Kitabı*. İstanbul, Turkey: Bayçınar Tıbbi Yayıncılık ve Reklam Hiz. Tic. Ltd. Şti, p. 98.
- [3] J. W. Langston. (2002). Parkinson's disease: current and future challenges. *Neurotoxicology*, vol. 23, no. 4-5, pp. 443-450.
- [4] J. Parkinson. (1817). An essay on the shaking palsy (Printed by Whittingham and Rowland for Sherwood, Neely, and Jones), ed: London.
- [5] J. Jankovic. (2008). Parkinson's disease: clinical features and diagnosis. *Journal of neurology, neurosurgery & psychiatry*, vol. 79, no. 4, pp. 368-376.
- [6] H. Gümüş, Z. Akpınar, & O. Demir. (2013). Erken evre Parkinson hastalığında motor olmayan semptomların değerlendirilmesi. *Türk Nöroloji Dergisi*, vol. 19, no. 3, pp. 97-103.
- [7] Y. Akgün & S. Peker. (2010). Tremor tedavisinde cerrahi girişimler. *Acıbadem Üniversitesi Sağlık Bilimleri Dergisi*, vol. 1 (3), no. 3, pp. 123-127.
- [8] B. Harel, M. Cannizzaro, & P. J. Snyder. (2004). Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease: A longitudinal case study. *Brain and cognition*, vol. 56, no. 1, pp. 24-29.
- [9] A. Tsanas, M. A. Little, P. E. McSharry, & L. O. Ramig. (2010). Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests. *IEEE transactions on Biomedical Engineering*, vol. 57, no. 4, pp. 884-893.

- [10] C. O. Sakar & O. Kursun. (2010). Telediagnosis of Parkinson's disease using measurements of dysphonia. *Journal of medical systems*, vol. 34, no. 4, pp. 591-599.
- [11] H. Gürüler. (2017). A novel diagnosis system for Parkinson's disease using complex-valued artificial neural network with k-means clustering feature weighting method. *Neural Computing and Applications*, vol. 28, no. 7, pp. 1657-1666.
- [12] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, & L. O. Ramig. (2009). Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. *IEEE transactions on biomedical engineering*, vol. 56, no. 4, pp. 1015-1022.
- [13] M. Peker, B. Sen, & D. Delen. (2015). Computer-aided diagnosis of Parkinson's disease using complex-valued neural networks and mRMR feature selection algorithm. *Journal of healthcare engineering*, vol. 6, no. 3, pp. 281-302.
- [14] B. E. Sakar ve ark. (2013). Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 4, pp. 828-834.
- [15] R. Das. 2010). A comparison of multiple classification methods for diagnosis of Parkinson disease. *Expert Systems with Applications*, vol. 37, no. 2, pp. 1568-1572.
- [16] F. Åström & R. Koker. 2011). A parallel neural network approach to prediction of Parkinson's Disease. *Expert systems with applications*, vol. 38, no. 10, pp. 12470-12474.
- [17] Ö. Eskidere, F. Ertaş, & C. Haniççi. (2012). A comparison of regression methods for remote tracking of Parkinson's disease progression. *Expert Systems with Applications*, vol. 39, no. 5, pp. 5523-5528.
- [18] B. E. Sakar, G. Serbes, & C. O. Sakar. (2017). Analyzing the effectiveness of vocal features in early telediagnosis of Parkinson's disease. *PloS one*, vol. 12, no. 8, p. e0182428.
- [19] D. Braga, A. M. Madureira, L. Coelho, & R. Ajith. (2019). Automatic detection of Parkinson's disease based on acoustic analysis of speech. *Engineering Applications of Artificial Intelligence*, vol. 77, pp. 148-158.
- [20] C. O. Sakar ve ark. (2019). A comparative analysis of speech signal processing algorithms for Parkinson's disease classification and the use of the tunable Q-factor wavelet transform. *Applied Soft Computing*, vol. 74, pp. 255-263.
- [21] Ö. Eskidere. (2012). A Comparison Of Feature Selection Methods For Diagnosis Of Parkinson's Disease From Vocal Measurements. *Sigma*, vol. 30, pp. 402-414.
- [22] I. W. Selesnick. (2011). Wavelet transform with tunable Q-factor. *IEEE transactions on signal processing*, vol. 59, no. 8, pp. 3560-3575.
- [23] D. P. Acharya, G. Panda, & Y. Lakshmi. (2010). Effects of finite register length on fast ICA, bacterial foraging optimization based ICA and constrained genetic algorithm based ICA algorithm. *Digital Signal Processing*, vol. 20, no. 3, pp. 964-975.
- [24] S. Jiang, P. Lin, Y. Chen, C. Tian, & Y. Li. (2019). Mixed-signal extraction and recognition of wind turbine blade multiple-area damage based on improved Fast-ICA. *Optik*, vol. 179, pp. 1152-1159.
- [25] V. Zarzoso & P. Comon. (2010). Robust independent component analysis by iterative maximization of the kurtosis contrast with algebraic optimal step size. *IEEE Transactions on Neural Networks*, vol. 21, no. 2, pp. 248-261.
- [26] R. Xixi & Z. Qun. (2011). Power quality harmonic detection based on Fast-ICA. in *2011 IEEE Power Engineering and Automation Conference*, vol. 3: IEEE, pp. 26-29.
- [27] T. Ahmad & M. Ghanbari. (2011). A review of independent component analysis (ica) based on kurtosis contrast function. *Australian Journal of Basic and Applied Sciences*, vol. 5, no. 9, pp. 1747-1755.



- [28] J. Wang, C. Wang, T. Zhang, & B. Zhong. (2016). Comparison of different independent component analysis algorithms for output-only modal analysis. *Shock and Vibration*, vol. 2016.
- [29] H. Li & T. Adali. (2008). A class of complex ICA algorithms based on the kurtosis cost function. *IEEE Transactions on Neural Networks*, vol. 19, no. 3, pp. 408-420.
- [30] A. Hyvärinen, J. Karhunen, & E. Oja. (2004). *Independent component analysis*. John Wiley & Sons.
- [31] Q. V. Le, A. Karpenko, J. Ngiam, & A. Y. Ng. (2011). ICA with reconstruction cost for efficient overcomplete feature learning. in *Advances in neural information processing systems*, pp. 1017-1025.
- [32] L. Breiman. (2001). Random forests. *Machine learning*, vol. 45, no. 1, pp. 5-32.
- [33] I. H. Witten, E. Frank, L. E. Trigg, M. A. Hall, G. Holmes, & S. J. Cunningham. (1999). Weka: Practical machine learning tools and techniques with Java implementations.
- [34] Y. Ma, L. Guo, & B. Cukic. (2006). A statistical framework for the prediction of fault-proneness. *Advances in Machine Learning Application in Software Engineering, Idea Group Inc*, pp. 237-265.
- [35] D. M. Powers. (2011). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.
- [36] N. Nicolov. (2012, 10 April). *Machine Learning with Applications in Categorization, Popularity and Sequence Labeling: 57th and 58nd slides*. [Online]. Available: [http://www.slideshare.net/Nicolas\\_Nicolov/machine-learning-14528792](http://www.slideshare.net/Nicolas_Nicolov/machine-learning-14528792).
- [37] Ş. Yücelbaş, C. Yücelbaş, G. Tezel, S. Özşen, & Ş. Yosunkaya. (2018). Automatic sleep staging based on SVD, VMD, HHT and morphological features of single-lead ECG signal. *Expert Systems with Applications*, vol. 102, pp. 193-206.
- [38] T. Saito & M. Rehmsmeier. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS one*, vol. 10, no. 3, p. e0118432.
- [39] T. J. Wroge, Y. Özkanca, C. Demiroglu, D. Si, D. C. Atkins, & R. H. Ghomi. (2018). Parkinson's Disease Diagnosis Using Machine Learning and Voice. in *2018 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*: IEEE, pp. 1-7.