

Stockwell Dönüşümü, ONE-R Özellik Seçme Yöntemi ve Rastgele Orman Algoritması ile Güç Kalitesi Bozulumu Sinyallerinin Sınıflandırılması

Düzgün AKMAZ^{1*}

¹ Elektrik Elektronik Mühendisliği, Mühendislik Fakültesi, Munzur Üniversitesi, Tunceli, Türkiye
^{*1} dakmaz@munzur.edu.tr

(Geliş/Received: 02/11/2021;

Kabul/Accepted: 03/02/2022)

Öz: Bu çalışmada Güç Kalitesi (GK) bozulumu sinyallerinin sınıflandırılması için bir yöntem uygulanmıştır. Yöntemde Stockwell dönüşümünün (S-dönüşümü-SD) farklı GK sinyallerine uygulanması ile 2 boyutlu Zaman-Frekans (TF Time-Frequency) matrisi elde edilmiştir. Daha sonra elde edilen matrisin ve GK sinyallerinin farklı entropi değerleri ile sınıflandırma öznelikleri elde edilmiştir. Ancak elde edilen sınıflandırma özneliklerinin tamamının kullanılması, gerçekleştiren yöntemin karmaşık hale gelmesine neden olmaktadır. Bu yüzden bu çalışmada ONE-R özellik seçme yöntemi sınıflandırmadaki etkili özelliklerin belirlenmesi, daha az öznitelik ve daha basit bir yapı ile yüksek sınıflandırma başarısının sağlanması için kullanılmıştır. Belirlenen öznelikler ve Rastgele Orman (RO) algoritması ile GK bozulmaları sınıflandırılmıştır. Benzetim sonuçları, yöntemin başarılı olduğunu göstermiştir. Aynı zamanda benzetim sonuçları gerçekleştirilen yöntemin bir kere eğitildikten sonra, 35 dB ve üzerindeki gürültü seviyelerine sahip farklı GK bozulmalarının sınıflandırılması için de etkili olduğunu göstermiştir.

Anahtar kelimeler: Güç Kalitesi, Stockwell Dönüşümü, ONE-R Özellik Seçme Yöntemi, Rastgele Orman Algoritması.

Classification of Power Quality Disturbance Signals with Stockwell Transform, ONE-R Feature Selection Method and Random Forest Algorithm

Abstract: In this study, a method was applied to classify the Power Quality (PQ) distortion signals. In the method, a 2-dimensional Time-Frequency (TF) matrix was obtained by applying the Stockwell transform (S-transform-ST) to different PQ signals. Then, classification features were obtained with different entropy values of the obtained matrix and PQ signals. However, using all of the obtained classification features causes the performing method to become complicated. Therefore, in this study, the ONE-R feature selection method was used to determine the effective features in classification, to ensure high classification success with fewer features and a simpler structure. PQ disturbances were classified with the determined features and Random Forest (RF) algorithm. The simulation results showed that the method was successful. At the same time, the simulation results showed that the applied method, once trained, is also effective in the classification of different PQ disturbances with noise levels of 35 dB and above.

Key words: Power Quality, ONE-R Feature Selection Method, Random Forest Algorithm, Stockwell Transform.

1. Giriş

Güç sistemlerindeki sinyallerin tüketicilere sürekli olarak ulaştırılması istenmektedir. Bununla birlikte son zamanlarda doğrusal olmayan yük kullanımının artması, evirici gibi anahtarlama elemanlarına sahip cihazların sistemlere dâhil edilmesi gibi bazı durumlar, şebekelerdeki sinyallerin kararlılığını ve kalitesini düşürmektedir. Aynı zamanda yıldırım ve rüzgâr gibi bazı çevresel faktörlerin eklenmesi ile güç sistemlerinde farklı arızalar oluşabilmektedir. Bu arızaların doğru tanımlanması arızalara müdahale süresini kısaltabilir. Bu yüzden literatürde birçok yöntem, farklı Güç Kalitesi (GK) arıza sinyallerinin belirlenmesi, segmentasyonu, özellik çıkarımı ve sınıflandırılması için geliştirilmiştir [1-4].

Ayrık Dalgacık Dönüşümü (ADD) ve Destek Vektör Makineleri (DVM) [5], Ampirik Kip Ayırımı-Hilbert dönüşümü ve olasılıksal sinir ağı modeli [6], Gabor-Wigner dönüşümü ve Nöral Ağlar (NA) [7], S-Dönüşümü (SD) ve NA [8] gibi birçok farklı yaklaşım GK olaylarının sınıflandırılması için uygulanmıştır. Bu çalışmalara ek olarak bazı hibrit yaklaşımlar da GK bozulumu sinyallerinin sınıflandırılması için tamamlanmıştır [9,10]. Hibrit yöntemlerde birden fazla sinyal işleme aracı veya yapay zekâ algoritması kullanılarak sınıflandırma yapılmaktadır. [9]' da ADD /hızlı Fourier dönüşümü sinyal işleme aracı ve kural bazlı bir sınıflandırma yöntemi kullanılırken, [10]' da ise ADD/Hiperbolik SD ve çok sınıflı DVM GK bozulmalarının sınıflandırılması için kullanılmıştır.

* Sorumlu yazar: dakmaz@munzur.edu.tr. Yazarın ORCID Numarası: 0000-0002-4183-6424

Bazı çalışmalarda ise ayrıca farklı özellik seçme yöntemleri de GK sinyallerinin sınıflandırılması için kullanılmıştır. Özellik seçme yöntemlerinde daha az sınıflandırma özelliği ile yüksek bir sınıflandırma başarısı sağlanması amaçlanmaktadır. Özellik sayısı azaldığı için daha kısa sürede eğitim ve test verileri sınıflandırılabilir. Aynı zamanda daha az sınıflandırma özelliğinin kullanılması önerilen sınıflandırma yöntemlerinin daha basit yapılara dönüşmesini de sağlayabilmektedir. [11]'de dalgacık paket dönüşümü sinyal işleme yöntemi, genetik algoritma özellik seçme yöntemi ve bulanık k-en yakın komşu sınıflandırma algoritması kullanılmıştır. [12]'de SD sinyal işleme yöntemi, istatistiksel bir özellik seçme yöntemi ve olasılıksal sinir ağı GK sinyallerinin sınıflandırılması için uygulanmıştır. [13]'te SD/Varyasyon kip ayrıştırması sinyal işleme yöntemi, sıralı (ileri/geriye) doğru seçim-Gram-Schmidt ortogonalleşmesi özellik seçme yöntemleri ve DVM, GK bozulmalarının sınıflandırılması için uygulanmıştır.

Literatürde yukarıda belirtilen çalışmaların dışında birçok farklı yöntem geliştirilmiş olup bunlara ait bilgiler [1-3]'te özetlenmiştir. Bu çalışmalar incelendiğinde birçok farklı sinyal işleme aracının GK sinyalleri için kullanıldığı görülmektedir. Bu sinyal işleme araçlarından SD, sinyalleri zaman ve frekans düzleminde inceleyebildiği için birçok farklı çalışmada kullanılmıştır [12-17]. Bu yüzden bu çalışmada da SD sinyal işleme yöntemi kullanılmıştır. Daha önce gerçekleştirilen [12-17]'den farklı olarak bu çalışmada SD sinyal işleme aracı, ONE-R özellik seçme yöntemi ve RO algoritması ile birlikte uygulanmıştır. Benzetim sonuçları gerçekleştirilen yöntemin GK bozulmalarını sınıflandırmada başarılı olduğunu göstermiştir. Aynı zamanda benzetim sonuçları 40 dB gürültülü verilerdeki eğitim sonrası yöntemin bir daha eğitime ihtiyaç duymadan 35, 40, 45, 50,55, 60 dB gürültüye sahip GK bozulmalarını büyük bir doğruluk oranında sınıflandırabildiğini de göstermiştir.

2. Stockwell Dönüşümü (S-dönüşümü)

SD, Sürekli Dalgacık Dönüşümünün (SDD), “faz düzeltmesi” olarak görülebilir. Bir fonksiyonun SDD'si denklem (1)'deki gibi tanımlanır [17].

$$W(\tau, d) = \int_{-\infty}^{\infty} h(t)\omega(d, t - \tau)dt \quad (1)$$

Ölçek parametresi d , dalgacığın $\omega(d, t)$ “genişliğini” belirler ve böylece çözünürlüğü kontrol eder. $h(t)$ fonksiyonunun SD'si, faz faktörü ile çarpılan spesifik bir ana dalgacıklı bir SDD olarak tanımlanır [17].

$$S(\tau, f) = e^{j2\pi f\tau}W(d, \tau) \quad (2)$$

Burada ana dalgacık aşağıdaki gibi tanımlanır:

$$\omega(t, f) = \frac{|f|}{\sqrt{2\pi}} e^{-\frac{t^2 f^2}{2}} e^{-j2\pi f t} \quad (3)$$

Ölçek parametresi, f frekansının tersidir. Denklem (3)'teki dalgacık, kabul edilebilir bir dalgacık için sıfır ortalama koşulunu sağlamaz; bu nedenle, denklem (2) kesinlikle bir SDD değildir. Açıkça yazıldığında, SD denklem 4'teki gibi tanımlanır [17].

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 f^2}{2}} e^{-j2\pi f t} dt \quad (4)$$

SD ayrıca $h(t)$ 'nin Fourier spektrumu $H(f)$ üzerindeki işlemler olarak denklem 5'teki gibi de yazılabilir [17].

$$S(\tau, f) = \int_{-\infty}^{\infty} H(\alpha + f) e^{-\frac{2\pi^2 \alpha^2}{f^2}} e^{j2\pi \alpha \tau} d\alpha, \quad f \neq 0. \quad (5)$$

Güç sistemi sinyalleri $h(t)$, $h(kT)$ olarak ayrık formda ifade edilebilir. Burada $k=0, 1, \dots, N-1$ olup N toplam örnekleme sayısı iken, T ise örnekleme zaman aralığıdır. $h(kT)$ 'nin ayrık Fourier dönüşümü Denklem 6'daki gibi elde edilebilir.

$$H \left[\frac{n}{NT} \right] = \frac{1}{N} \sum_{k=1}^{N-1} h(kT). e^{(-\frac{j2\pi nk}{N})} \quad (6)$$

Burada $n=0, 1, \dots, N-1$.

Bir ayrık zaman serisinin $h(kT)$ SD'si denklem 7 ile belirlenebilir. ($\tau \rightarrow kT$ ve $f \rightarrow n/NT$)

$$S\left[kT, \frac{n}{NT}\right] = \sum_{m=0}^{N-1} \mathbf{H}\left[\frac{m+n}{NT}\right] e^{-\frac{2\pi^2 m^2}{n^2}} e^{\frac{j2\pi mk}{N}} \quad n \neq 0 \quad (7)$$

Burada $m=0, 1, \dots, N-1$ ve $n=1, \dots, N-1$ dir. Denklem 8, $n=0$ için elde edilir.

$$S[kT, 0] = \frac{1}{N} \sum_{m=0}^{N-1} h\left(\frac{m}{NT}\right) \quad (8)$$

Hızlı Fourier dönüşümü verimliliğinden ve konvolüsyon (evrişim) teoreminin avantajından yararlanarak ayrık SD hızlı bir şekilde hesaplanabilir. SD, faz spektrumunun yanı sıra genlik spektrumunu da lokalize eder.

3. ONE-R Özellik seçme yöntemi

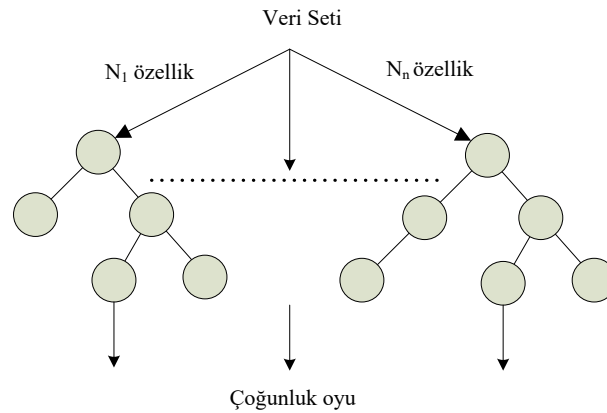
One-R, Holte [18] tarafından önerilen basit bir algoritmadır. Eğitim verisindeki her öznelik için bir kural oluşturulur ve ardından en küçük hataya sahip kural seçilir. Sayısal olarak değerli tüm özellikleri sürekli olarak ele alır ve değer aralığını birkaç ayrık aralığa bölmek için basit bir yöntem kullanır. Yalnızca bir özelliğe dayalı basit kurallar üretir. Minimal bir sınıflandırıcı biçimi olmasına rağmen, diğer öğrenme şemaları için bir kıyaslama olarak temel bir performansı belirlemek için yararlı olabilir [18].

One-R algoritma adımları kısaca aşağıdaki gibi ifade edilebilir [19]:

- *Her bir özellik f için, f 'nin etki alanındaki her v değeri için, v değerine sahip f özellik örnek setini seç,
- * c = seçilen sette en sık rastlanan sınıf değeri, her bir f özelliği için “ f özelliği v değerini alıyorsa sınıfı c ” kuralını ekle,
- *En yüksek sınıflandırma oranına sahip kuralı çıktı olarak üret.

4. Rastgele Orman Algoritması

Bu algoritmada ilk olarak orijinal eğitim veri setinden birden çok önyüklemeli eğitim veri seti oluşturur. Daha sonra her bir önyüklemeli eğitim veri seti için bir ağaç üretilir. Son olarak bu ağaçların çoğunluk oyunu kullanarak sınıflandırma yapılır. Önyükleme, rastgele ormandaki her bir karar ağacının farklı olmasını sağlar. Aynı zamanda bu algoritmanın bir sınıflandırıcı yerine birden fazla sınıflandırıcı kullanması ve bu sınıflandırıcıların çoğunluk oylarını kullanması RO algoritmasını etkili bir sınıflandırma algoritması haline getirir. Tüm bu özellikler RO yönteminin aşırı öğrenme durumunu minimize eder ve birçok farklı algoritmaya göre başarılı olmasını sağlar [20,21]. Şekil 1, bu algoritmanın akış şemasını göstermektedir.



Şekil 1. Rastgele orman algoritması

5. Benzetim Sonuçları

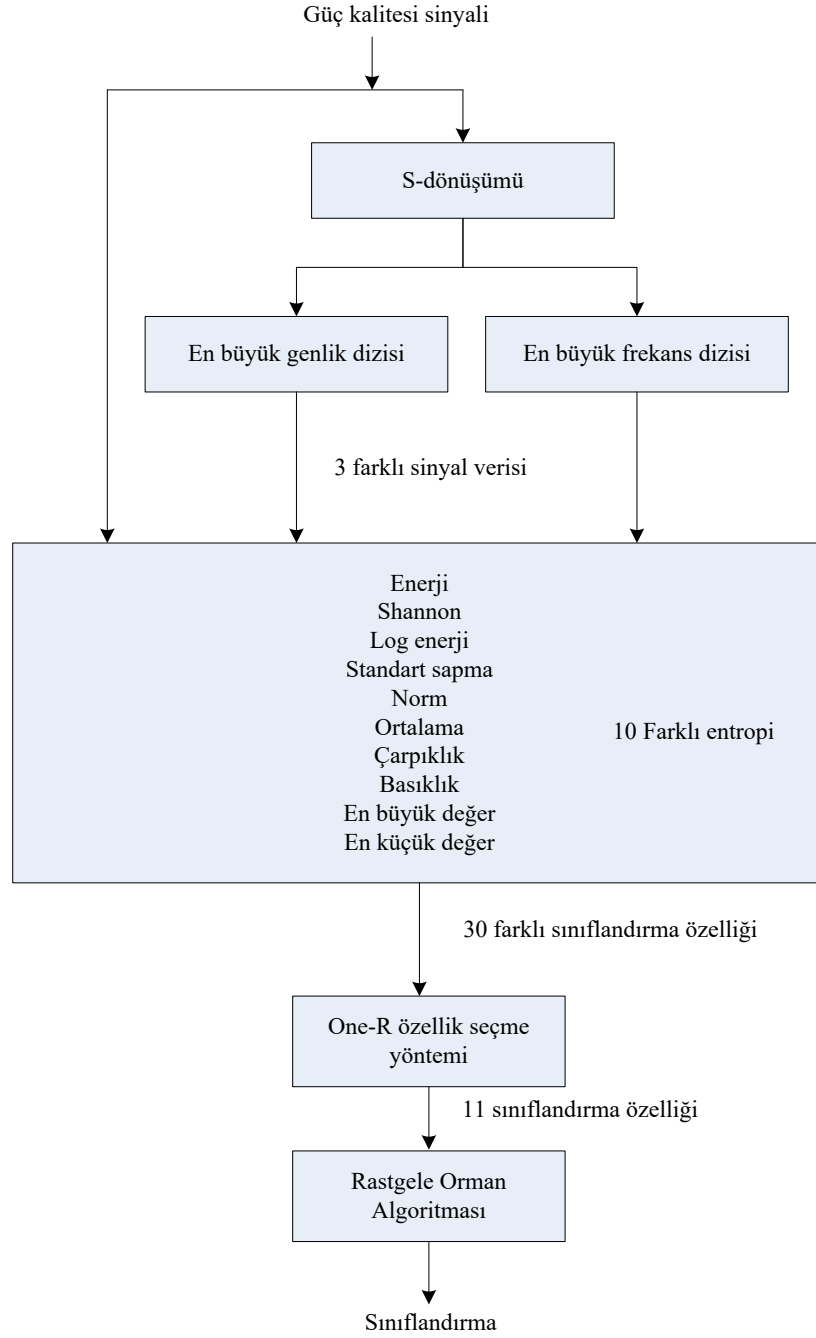
Bu çalışmada sinüs, gerilim çukuru (sag), gerilim tepesi (swell), harmonik (harmonic), geçici rejim gerilim bileşenleri (transient), gerilim kesintisi (interruption), gerilim çukuru ile harmonik (sag with harmonic), gerilim tepesi ile harmonik (swell with harmonic) ve gerilim kırışması (flicker) sinyalleri incelenmiştir.

Yukarıda belirtilen sinyaller sentetik olarak üretilmiş olup sinyallerin parametreleri [22]'de belirtilen aralıklar içinde rastgele olarak seçilmiştir. Tablo 1'de bu sentetik sinyallerin parametreleri verilmiştir [22].

Tablo 1. Güç kalitesi olaylarının matematiksel denklemleri.

Sınıf	Güç kalitesi sinyali	Güç kalitesi sinyallerinin matematiksel denklemleri	Parametreler
C ₁	Sinüs	$v(t)=\sin(\omega t)$	Frekans=50Hz
C ₂	Gerilim çukuru (Sag)	$v(t)=(1-\alpha(u(t-t_1)-u(t-t_2)))\sin(\omega t)$	$0.1 \leq \alpha \leq 0.9$ $T \leq t_2 - t_1 \leq 9T$
C ₃	Gerilim tepesi (Swell)	$v(t)=(1+\alpha(u(t-t_1)-u(t-t_2)))\sin(\omega t)$	$0.1 \leq \alpha \leq 0.8$ $T \leq t_2 - t_1 \leq 9T$
C ₄	Gerilim kesintisi (Interruption)	$v(t)=(1-\alpha(u(t-t_1)-u(t-t_2)))\sin(\omega t)$	$0.9 \leq \alpha \leq 1$ $T \leq t_2 - t_1 \leq 9T$
C ₅	Harmonik (Harmonic)	$v(t)=(\alpha_1 \sin(\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t) + \alpha_7 \sin(7\omega t))$	$0.05 \leq \alpha_3 \leq 0.15$ $0.05 \leq \alpha_5 \leq 0.15$ $0.05 \leq \alpha_7 \leq 0.15$ $\sum \alpha_i^2 = 1$
C ₆	Geçici rejim (Transient)	$v(t)=(\sin(\omega t) + \alpha_{osc} \exp(-(t-t_1)/\tau_{osc}))\sin(\omega_{nosc}(t-t_1))$	$\tau_{osc}=0.008-0.04$ sn $\omega_{nosc}=100-400$ Hz
C ₇	Gerilim çukuru ile harmonik (Sag with harmonic)	$v(t)=(1-\alpha(u(t-t_1)-u(t-t_2))) (\alpha_1 \sin(\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t) + \alpha_7 \sin(7\omega t))$	$0.1 \leq \alpha \leq 0.9$ $T \leq t_2 - t_1 \leq 9T$ $0.05 \leq \alpha_3 \leq 0.15$ $0.05 \leq \alpha_5 \leq 0.15$ $0.05 \leq \alpha_7 \leq 0.15$ $\sum \alpha_i^2 = 1$
C ₈	Gerilim yükselmesi ile harmonik (Swell with harmonic)	$v(t)=(1+\alpha(u(t-t_1)-u(t-t_2))) (\alpha_1 \sin(\omega t) + \alpha_3 \sin(3\omega t) + \alpha_5 \sin(5\omega t) + \alpha_7 \sin(7\omega t))$	$0.1 \leq \alpha \leq 0.8$ $T \leq t_2 - t_1 \leq 9T$ $0.05 \leq \alpha_3 \leq 0.15$ $0.05 \leq \alpha_5 \leq 0.15$ $0.05 \leq \alpha_7 \leq 0.15$ $\sum \alpha_i^2 = 1$
C ₉	Gerilim kırışması (Flicker)	$v(t)=(1+\alpha_f \sin(\beta_f \omega t))\sin(\omega t)$	$\alpha_f=0.1-0.2$ $\beta_f=5-10$ Hz

SD sinyal işlemini kullanan bazı çalışmalarda sinyallerin örnekleme frekansı 3.2 kHz olarak seçilmiştir [15, 16]. Bu yüzden bu çalışmada benzer şekilde 3.2 kHz örnekleme frekansı kullanılmıştır. Bu çalışmadaki sinyal üretme ve özellik çıkarma adımları için MATLAB programı kullanılmıştır. ONE-R ve RO algoritması için ise WEKA programı kullanılmıştır. WEKA programı, veri madenciliğinde ve farklı güç kalite çalışmalarında güncel olarak kullanılan bir programdır [23, 24]. Çalışmada kullanılan sınıflandırma yönteminin genel algoritması Şekil 2'de gösterilmiştir. Aynı zamanda Şekil 2'de sınıflandırma özelliklerinin nasıl elde edildiği de gösterilmiştir.



Şekil 2. Sınıflandırma yöntemi

Yöntemin başarısını test etmek için farklı şartlar altında eğitim ve test verileri üretilmiştir. Eğitim verilerinde her bir GK bozulumu için 100 farklı benzetim gerçekleştirilmiştir. Toplamda 9 farklı GK sinyali için 900 farklı arıza meydana gelmiştir.

Üretilen eğitim verilerine 40 dB gürültü eklenmiştir. Benzer şekilde rastgele üretilmiş 900 tane test verisi için de GK sinyallerine 40 dB gürültü eklenmiştir.

Sınıflandırma özelliklerini elde etmek için SD, 40 dB gürültü içeren GK sinyallerine uygulanmıştır. SD sonrası her bir sinyalin 2 boyutlu Zaman-Frekans (TF Time-Frequency) matrisi elde edilmiştir. Bu matris, karmaşık sayılar içeren NxM boyutunda bir matristir. Matrisin sütunları zamanla ve satırları frekansla ilgilidir.

TF konturunun/matrisinin sütun değerlerinde en büyük değer aranarak zaman - en büyük genlik dizisi (TmA-Time-maximum amplitude), satır değerlerinde en büyük değer aranarak ise frekans - en büyük genlik dizisi (FmA-Frequency-maximum amplitude) elde edilmiştir. Daha sonra GK sinyali, zaman en büyük genlik dizisi ve frekans en büyük genlik dizisinin 10 farklı entropi (enerji, shannon, log enerji, standart sapma, norm, ortalama, çarpıklık, basıklık, en büyük değer, en küçük değer entropi) değerleri alınarak toplamda 30 (3 sinyal x10 entropi) farklı sınıflandırma özelliği elde edilmiştir. Daha sonra 30 farklı sınıflandırma özelliği ve RO algoritması kullanılması ile GK bozulmaları sınıflandırılmıştır. Tablo 2, 30 farklı sınıflandırma özelliği ve RO algoritması ile elde edilen benzetim sonuçlarını göstermektedir. Burada hem eğitim hem de test verilerindeki GK sinyallerinde 40 dB gürültü bulunmaktadır.

Tablo 2. 30 farklı sınıflandırma özelliği ve RO yöntemi ile elde edilen benzetim sonuçları (eğitim ve test 40 dB gürültülü).

	Eğitim	Test
Başarı %	100	99.11
Süre (sn)	0.12	0.1

Tablo 2' den görüldüğü gibi eğitim sonrası yapılan test verilerindeki başarı %99.11 olmuştur. WEKA program sonuçları, modeli eğitim verileri üzerinde test etmek için geçen sürenin 0.12 sn olduğunu gösterirken, verilen test setinde modeli test etmek için geçen sürenin ise 0.1 sn olduğunu göstermiştir. Bunlara ek olarak WEKA programı eğitim verilerinde modeli oluşturmak için geçen sürenin 0.3 sn olduğunu göstermiştir. Bu tablodan görüldüğü gibi 30 farklı sınıflandırma özelliği ile yüksek bir sınıflandırma başarısı elde edilmiştir. Fakat özellik seçimi yöntemi ile daha az sınıflandırma özelliği ile benzer şekilde yüksek sınıflandırma başarısı elde edilebilmektedir. Bu çalışmada, daha az sınıflandırma verisi ile yüksek bir sınıflandırma başarısı sağlamak ve benzetimlerdeki süreyi kısaltmak için ONE-R özellik seçimi yöntemi ve RO algoritması kullanılmıştır. İlk olarak ONE-R yöntemi ile sınıflandırma için en etkili özellikler belirlenmiş ve bu özellikler sıralanmıştır. Daha sonra bu özellikler sıra ile eklenerek RO algoritmasının eğitim ve test başarısı incelenmiştir. Burada en az sınıflandırma özelliği ile en yüksek sınıflandırma başarısı sağlanmaya çalışılmıştır.

Tablo 3, ONE-R algoritmasının seçtiği özellikler ve RO yönteminin sınıflandırma başarısını göstermektedir. Tablo 3'ten görüldüğü gibi yöntemin eğitim başarısı değişmemektedir. Ancak test başarısı değişmektedir. 11 veri kullanılarak en yüksek test başarısı elde edilmiştir. 11 özellikten sonra eklenen özellikler test sınıflandırma başarısını düşürmektedir. Bu yüzden belirlenen 11 özellik, sınıflandırma için kullanılmıştır. Bu yaklaşım, bazı çalışmalarda kullanılmıştır [5, 10]. Benzer şekilde, bu çalışmalardan görüldüğü gibi en az veri ile elde edilen en yüksek sınıflandırma başarısından sonra eklenen özellikler sınıflandırma başarısını düşürmektedir [5].

Tablo 3. ONE-R yönteminin belirlediği özellikler ve RO algoritmasının sınıflandırma başarısı.

Farklı veri sayısı	Eğitim sınıflandırma başarısı %	Test sınıflandırma başarısı %	Belirlenen özellikler (KO)
1	100	58.55	12
2	100	82.11	12,20
3	100	95.44	12,20,3
4	100	96	12,20,3,11
5	100	96	12,20,3,11,15
6	100	95.77	12,20,3,11,15,16
7	100	96.88	12,20,3,11,15,16,5
8	100	97	12,20,3,11,15,16,5,13
9	100	97.77	12,20,3,11,15,16,5,13,2
10	100	97.33	12,20,3,11,15,16,5,13,2,24
11	100	98.1	12,20,3,11,15,16,5,13,2,24,8
12	100	97.88	12,20,3,11,15,16,5,13,2,24,8,26
13	100	97.55	12,20,3,11,15,16,5,13,2,24,8,26,25
14	100	97.98	12,20,3,11,15,16,5,13,2,24,8,26,25,19

Tablo 4'te ise belirlenen özelliklerin tanımları gösterilmiştir. Tablo 4 incelendiğinde farklı sinyallerin farklı entropi değerlerinin sınıflandırma için seçildiği görülmektedir.

Tablo 4. ONE-R algoritmasının belirlediği sınıflandırma özelliklerinin tanımı.

	Özellik	Sinyal En büyük genlik dizisi En büyük frekans dizisi	Entropi
1	12	En büyük genlik dizisi	Shannon
2	20	En büyük genlik dizisi	En küçük değer
3	3	Sinyal	Log enerji
4	11	En büyük genlik dizisi	Enerji
5	15	En büyük genlik dizisi	Norm
6	16	En büyük genlik dizisi	Ortalama
7	5	Sinyal	Norm
8	13	En büyük genlik dizisi	Log enerji
9	2	Sinyal	Shannon
10	24	En büyük frekans dizisi	Standart sapma
11	8	Sinyal	Basıklık

Tablo 4'teki seçilen sınıflandırma özelliklerine göre eğitim ve test verisinde belirlenen 11 sınıflandırma özelliği kullanılmıştır. Diğer sınıflandırma özellikleri veri tabanından kaldırılmıştır. Tekrardan WEKA programı kullanılarak benzetim sonuçları incelenmiştir.

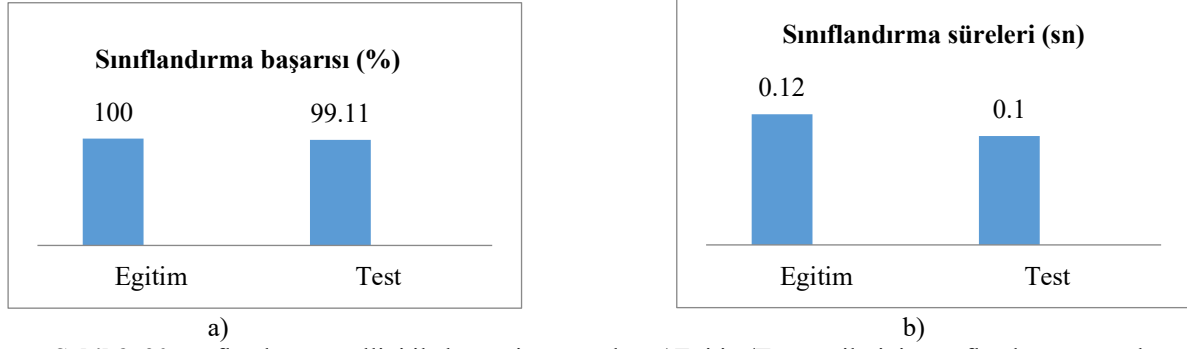
Tablo 5'te ise belirlenen 11 farklı sınıflandırma özelliği ile elde edilen benzetim sonuçlarını göstermektedir.

Tablo 5. ONE-R özellik seçme yöntemi ile belirlenen 11 farklı sınıflandırma özelliği ve RO algoritması ile elde edilen benzetim sonuçları (eğitim ve test 40 dB gürültülü).

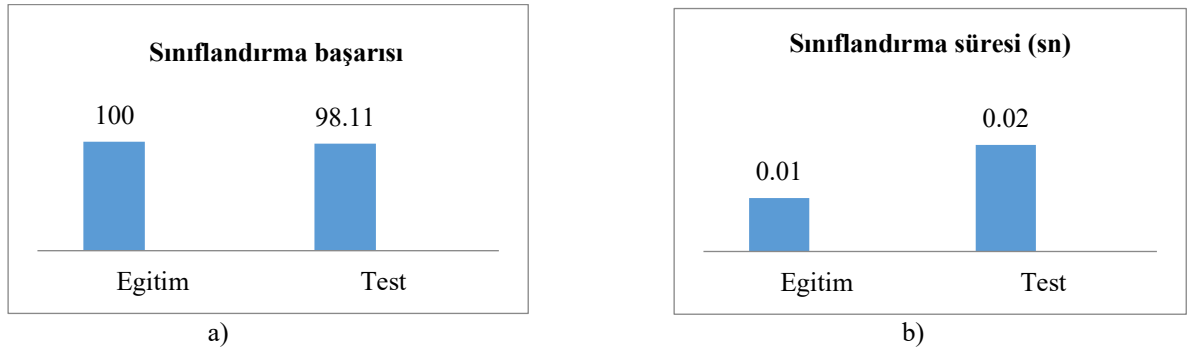
	Eğitim	Test
Başarı %	100	98.11
Süre (sn)	0.01	0.02

Tablo 5'ten görüldüğü gibi seçilen 11 özellik ile test verisinde %98.11'lik bir sınıflandırma başarısı elde edilmiştir. Tablo 2'de ise 30 sınıflandırma özelliği kullanılması durumunda test verilerinde %99.11'lik bir sınıflandırma başarısının elde edildiği görülmüştür. Elde edilen sonuçlar iki yöntemin de sınıflandırmada başarılı olduğunu göstermiştir. Bununla birlikte Tablo 2 ve Tablo 5'te görüldüğü gibi belirlenen 11 sınıflandırma özelliği ile çok daha kısa sürede eğitim ve test gerçekleştirilmiştir. Tablo 2'de 30 sınıflandırma verisi kullanıldığında test için gerekli sürenin 0.1 sn olduğu sonucuna varılırken, Tablo 5'te 11 sınıflandırma özelliği ile bu sürenin 0.02 sn olduğu görülmüştür. Böylece daha kısa sürede yüksek bir sınıflandırma başarısı elde edilmiştir. Şekil 3 ve Şekil 4'te elde edilen sonuçlar grafiksel olarak da verilmiştir. Bu şekiller incelendiğinde belirlenen 11 özellik ile başarılı bir sınıflandırma yapıldığı görülmektedir. Ancak, Şekil 3 ve Şekil 4'ten görüldüğü gibi belirlenen 11 özellik ile çok daha kısa sürede eğitim ve test verilerinin sınıflandırması yapılmıştır.

Bu çalışmadaki yöntemin başarısını göstermek için aynı zamanda farklı özellik seçme yöntemlerinin de test edilmesi gereklidir. Bu yüzden bu çalışmada ayrıca korelasyon öznitelik değerlendirmesi ve kazanç oranı özellik seçme yöntemleri de kullanılmıştır. Bu özellik seçme yöntemleri ile ilgili ayrıntılı bilgi [25, 26]'da bulunabilir. Tablo 6'da farklı özellik seçme yöntemleri ile belirlenen 11 sınıflandırma özelliğinin kullanılması ile elde edilen test verisi sınıflandırma sonuçları gösterilmiştir. Bu tablodan görüldüğü gibi en yüksek test sınıflandırma başarısı, ONE-R özellik seçme yöntemi ile elde edilmiştir. Bu yüzden çalışmada ONE-R özellik seçme yöntemi kullanılmıştır.



Şekil 3. 30 sınıflandırma özelliği ile benzetim sonuçları a) Eğitim/Test verilerinin sınıflandırma sonuçları b) Eğitim/Test verilerinin sınıflandırma süreleri



Şekil 4. 11 sınıflandırma özelliği ile benzetim sonuçları a) Eğitim/Test verilerinin sınıflandırma sonuçları b) Eğitim/Test verilerinin sınıflandırma süreleri

Tablo 6. Farklı özellik seçme yöntemi ile belirlenen 11 farklı sınıflandırma özelliği ve RO algoritması ile elde edilen benzetim sonuçları (eğitim ve test 40 dB gürültülü).

Özellik seçme algoritması	Test
ONE-R	98.11
Kazanç oranı	97.88
Korelasyon öznelik değerlendirmesi	97

Burada diğer bir önemli durum ise 11 seçilen sınıflandırma özelliğinin farklı gürültü seviyelerinde de başarılı olduğunun test edilmesidir. Bunu gerçekleştirmek için farklı gürültüye sahip test verileri kullanılmıştır. Test verilerine 30, 35, 40, 45, 50, 55 ve 60 dB gürültü eklenmiştir. Eğitim verisinde ise 40 dB gürültü kullanılmıştır. Tablo 7’de farklı gürültü seviyeleri için gerçekleştirilen yöntemin sınıflandırma başarıları verilmiştir.

Tablo 7. ONE-R özellik seçme yöntemi ile belirlenen 11 farklı sınıflandırma özelliği ve RO yöntemi ile farklı gürültüye sahip test verilerindeki sınıflandırma başarıları (eğitim 40 dB gürültülü).

Test verisi gürültü seviyesi (dB)	30	35	40	45	50	55	60
Başarı (%)	90.66	95.66	98.1	97.66	98	98.11	98

Bu çalışmada 40 dB gürültü değerine sahip eğitim verileri için bir yöntem geliştirilmiştir. Burada GK olaylarının parametreleri Tablo 1’de belirtilen aralıklarda rastgele seçilmiştir. Bu nedenlerden dolayı Tablo 7’den görüldüğü gibi 40 ile 60 dB gürültü aralığında yöntemin başarıları küçük değişiklikler göstermiştir.

Ancak genel olarak bütün benzetim sonuçları yöntemin, bir kere eğitildikten sonra 35 dB üzerindeki gürültü seviyelerindeki GK sinyallerinin sınıflandırılmasında etkili olduğunu göstermiştir.

Tablo 8’de ise bu çalışmadaki yöntem sonuçları ile [11]’de tamamlanan çalışma sonuçları karşılaştırılmıştır. Bu tablodan görüldüğü gibi tamamlanan yöntemde daha az sınıflandırma özelliği ile 35 ve 40 dB gürültü seviyeleri için daha yüksek bir sınıflandırma başarıları elde edilmiştir.

Tablo 8. Farklı gürültü seviyeleri için önerilen yöntemin, [11] ile karşılaştırılması.

Referans	Güç kalitesi sayısı	Özellik sayısı	Başarı (35 dB)	Başarı (40 dB)
[11]	10	16	93.96	95.96
Önerilen	9	11	95.66	98.1

Elde edilen tüm sonuçlar önerilen yöntemin GK sinyallerinin sınıflandırılmasında başarılı olduğunu göstermiştir.

6. Sonuçlar

Günümüzde güç sistemlerinde oluşabilecek arızalara daha doğru bir şekilde müdahale edebilmek için güç kalitesi sinyallerinin sınıflandırılması gereklidir. Bu yüzden birçok farklı yöntem, güç kalite sinyallerinin sınıflandırılması için geliştirilmiştir. Bu yöntemlerin olabildiğince etkili ve basit olması istenmektedir. Bu çalışmada S dönüşümü ve Rastgele Orman algoritmasına dayalı bir yöntem, sınıflandırma için uygulanmıştır. Uygulamada 30 farklı sınıflandırma verisi kullanılmıştır. Bu yöntem her ne kadar sınıflandırma için başarılı olsa da verilerin arttığı günümüzde daha az sınıflandırma verisi ile yüksek sınıflandırma başarısının elde edilmesi beklenmektedir. Bu yüzden çalışmada ayrıca ONE-R özellik seçme yöntemi uygulanmıştır. Benzetim sonuçları test verilerinde, 30 veri yerine belirlenen 11 farklı sınıflandırma verisi ile yüksek sınıflandırma başarısı elde edildiğini göstermektedir. Aynı zamanda benzetim sonuçları incelendiğinde 30 veri yerine belirlenen 11 veri kullanılmasının sınıflandırma yöntemini daha basit bir yapıya dönüştürdüğünü ve daha kısa sürede farklı güç kalitesi sinyallerini sınıflandırabildiğini de göstermiştir.

Kaynaklar

- [1] Choong F, Reaz MBI, Mohd-Yasin F. Advances in signal processing and artificial intelligence technologies in the classification of power quality events: a survey. *Electr. Power Compon. Syst* 2005; 33(12): 1333-1349.
- [2] Mishra M. Power quality disturbance detection and classification using signal processing and soft computing techniques: A comprehensive review. *Int. Trans. Electr. Energy Syst.* 2019; 29(8): e12008.
- [3] Khetarpal P, Tripathi MM. A critical and comprehensive review on power quality disturbance detection and classification. *Sustainable Comput. Inf. Syst* 2020; 100417.
- [4] Akmaz D, Approximate-derivative-based signal-processing method to segment power-quality disturbances. *IET Gener. Transm. Distrib* 2020;14(21): 4835-4846.
- [5] Erişti, H., Uçar, A., & Demir, Y. (2010). Wavelet-based feature extraction and selection for classification of power system disturbances using support vector machines. *Electric power systems research*, 80(7), 743-752.
- [6] Shukla S, Mishra S, Singh B. Empirical-mode decomposition with Hilbert transform for power-quality assessment. *IEEE Trans. Power Delivery* 2009;24(4): 2159-2165.
- [7] Cho SH, Jang G, Kwon SH. Time-frequency analysis of power-quality disturbances via the Gabor–Wigner transform. *IEEE Trans. Power Delivery* 2009;25(1): 494-499.
- [8] Lee IW, Dash PK. S-transform-based intelligent system for classification of power quality disturbance signals. *IEEE Trans. Ind. Electron* 2003; 50(4): 800-805.
- [9] Deokar, S. A., & Waghmare, L. M. (2014). Integrated DWT–FFT approach for detection and classification of power quality disturbances. *International Journal of Electrical Power & Energy Systems*, 61, 594-605.
- [10] Hajian, M., & Foroud, A. A. (2014). A new hybrid pattern recognition scheme for automatic discrimination of power quality disturbances. *Measurement*, 51, 265-280.
- [11] Panigrahi BK, Pandi VR. Optimal feature selection for classification of power quality disturbances using wavelet packet-based fuzzy k-nearest neighbour algorithm. *IET Gener. Transm. Distrib* 2009;3(3): 296-306.
- [12] Huang N, Xu D, Liu X, Lin L. Power quality disturbances classification based on S-transform and probabilistic neural network. *Neurocomputing* 2012; 98: 12-23.
- [13] Abdoos AA, Mianaei P K, Ghadikolaei M R. Combined VMD-SVM based feature selection method for classification of power quality events. *Appl. Soft Comput* 2016;38: 637-646.
- [14] Dash PK, Panigrahi BK, Panda G. Power quality analysis using S-transform. *IEEE Trans. Power Delivery* 2003; 18(2): 406-411.
- [15] Bhende CN, Mishra S, Panigrahi BK. Detection and classification of power quality disturbances using S-transform and modular neural network. *Electr. Power Syst. Res.* 2008;78(1): 122-128.
- [16] Mishra S, Bhende CN, Panigrahi BK. Detection and classification of power quality disturbances using S-transform and probabilistic neural network. *IEEE Trans. Power Delivery* 2007; 23(1): 280-287.

- [17] Zhao F, Yang R. Power quality disturbance recognition using S-transform. In 2006 IEEE Power Engineering Society General Meeting 2006; 7-pp.
- [18] Yildirim P. Filter based feature selection methods for prediction of risks in hepatitis disease. *International Journal of Machine Learning and Computing* 2015;5(4): 258.
- [19] Budak H. Özellik seçim yöntemleri ve yeni bir yaklaşım. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi* 2018;22: 21-31.
- [20] Misra S, Li H, He J. Noninvasive fracture characterization based on the classification of sonic wave travel times. In *Machine Learning for Subsurface Characterization* 2020;243-287.
- [21] Kiranmai SA, Laxmi AJ. Data mining for classification of power quality problems using WEKA and the effect of attributes on classification accuracy. *Protection and Control of Modern Power Systems* 2018; 3(1):1-12.
- [22] Moravej Z, Abdoos AA, Pazoki MJEPC. Detection and classification of power quality disturbances using wavelet transform and support vector machines. *Electr. Power Compon. Syst* 2009;38(2):182-196.
- [23] Vinayagam A, Veerasamy V, Radhakrishnan P, Sepperumal M, Ramaiyan, K. An ensemble approach of classification model for detection and classification of power quality disturbances in PV integrated microgrid network. *Applied Soft Computing* 2021; 106; 107294.
- [24] Kiranmai SA, Laxmi AJ. Data mining for classification of power quality problems using WEKA and the effect of attributes on classification accuracy. *Protection and Control of Modern Power Systems* 2018; 3(1); 1-12.
- [25] Budak H. Özellik seçim yöntemleri ve yeni bir yaklaşım. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi* 2018; 22;21-31.
- [26] Yildirim P. Filter based feature selection methods for prediction of risks in hepatitis disease. *International Journal of Machine Learning and Computing* 2018; 5(4); 258.