



## Evaluation of wavelet features selected via statistical evidence from steady-state visually-evoked potentials to predict the stimulating frequency

Ebru Sayılğan<sup>1</sup> , Yılmaz Kemal Yüce<sup>2</sup> , Yalçın İşler<sup>3\*</sup> 

<sup>1</sup>Department of Biomedical Technologies, Izmir Katip Celebi University, 35620, Cigli, Izmir, Turkey

<sup>2</sup>Department of Computer Engineering, Alanya Alaaddin Keykubat University, 07425, Alanya, Antalya, Turkey

<sup>3</sup>Department of Biomedical Engineering, Izmir Katip Celebi University, 35620, Cigli, Izmir, Turkey

### Highlights:

- Eight distinct stimulation frequencies can be discriminated with the average of 36.1%.
- Reverse Biorthogonal mother wavelet gives the maximum accuracy.
- Ensemble learners give maximum accuracy among classifiers

### Keywords:

- Machine learning
- Signal processing
- BCI
- SSVEP
- ANOVA

### Graphical/Tabular Abstract

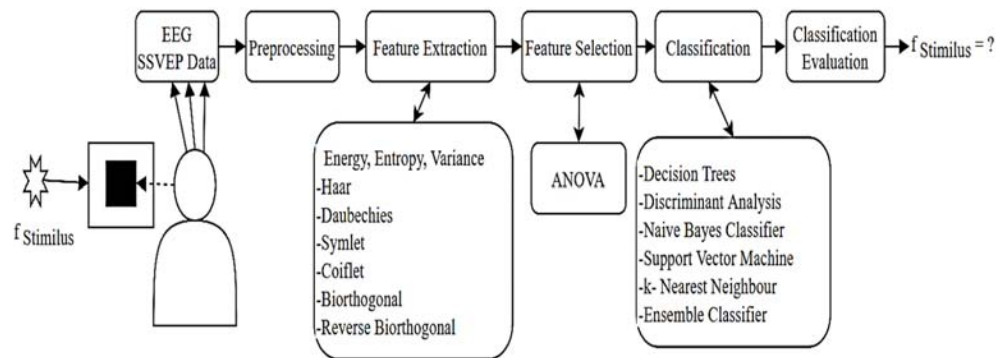


Figure A. Block diagram of the study

### Article Info:

Research Article

Received: 25.12.2019

Accepted: 27.09.2020

### DOI:

10.17341/gazimmfd.664583

### Acknowledgment:

The authors express special thanks to Adnan Vilic for answering questions about the data.

### Correspondence:

Author: Yalçın İşler  
e-mail: islerya@yahoo.com  
phone: +90 507 701 6050

**Purpose:** This study focuses on determining the stimulating frequency of a flashing image on the computer screen via electroencephalography (EEG) signals to investigate whether there is a correlation between stimulation frequency and brain activities.

### Theory and Methods:

Some researches indicate an increase in classifier performances for brain-computer interfaces (BCIs) when objects are shown as blinking on the screen. The recorded EEG signals while applying a blinking image on a screen are called steady-state visually-evoked potentials (SSVEP). Most of the studies tried to discriminate which object is gazed while recording. We tried to determine the stimulating frequency from this dataset using an open-source SSVEP dataset. Wavelet features of variance, energy, and entropy were calculated from six different mother wavelets (Haar, Daubechies, Symlet, Coiflet, Biorthogonal, Reverse Biorthogonal). All features and ANOVA-selected features were applied to distinct classifiers (Decision Trees, Discriminant Analysis, Naive Bayes, Support Vector Machines, k-Nearest Neighbors, and Ensemble Learners).

### Results:

The mother wavelet of Reverse Biorthogonal gives the maximum accuracy by using all classifiers. The ensemble learner determines the stimulating frequency correctly with a maximum accuracy of 37.9%. On the other hand, ANOVA-based feature selection decreases the classifier performances for all subjects.

### Conclusion:

Results indicate that the stimulating frequency is embedded into SSVEP signals. We concluded that SSVEP-based BCI studies for object classification or decision making should be conducted carefully since not only the location of the blinking object but also its blinking frequency is the part of EEG signals. In conclusion, it will be possible to detect gazed object among a few pictures by showing the flickering objects at the different frequencies at the same time.



## Uyartım frekansının kestiriminde istatistiksel anlamlılığa dayalı olarak seçilen durağan durum görsel uyarılmış potansiyellere ait dalgacık özneliklerinin değerlendirilmesi

Ebru Sayılğan<sup>1</sup> , Yılmaz Kemal Yüce<sup>2</sup> , Yalçın İşler<sup>3\*</sup>

<sup>1</sup>İzmir Katip Çelebi Üniversitesi, Fen Bilimleri Enstitüsü, Biyomedikal Teknolojiler ABD, 35620, Çiğli, İzmir, Türkiye

<sup>2</sup>Alanya Alaaddin Keykubat Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, 07425, Alanya, Antalya, Türkiye

<sup>3</sup>İzmir Katip Çelebi Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Biyomedikal Müh. Bölümü, 35620, Çiğli, İzmir, Türkiye

### Ö N E Ç İ K A N L A R

- DDGUP sinyallerinden sekiz farklı uyartım frekansı ortalama %36,1 doğrulukla tespit edilebilmektedir
- Reverse Biorthogonal ana dalgacığı en yüksek başarıyı vermektedir
- Topluluk Sınıflandırıcısı en yüksek başarıyı vermektedir

### Makale Bilgileri

Araştırma Makalesi  
Geliş: 25.12.2019  
Kabul: 27.09.2020

### DOI:

10.17341/gazimmfd.664583

### Anahtar Kelimeler:

Beyin-bilgisayar arayüzü,  
durağan-durum görsel-uyaran  
potansiyelleri,  
dalgacık dönüşümü,  
varyans analizi,  
öznelik seçimi,  
makine öğrenmesi

### ÖZET

Elektroensefalografi (EEG) beyin aktivitelerinin ölçümünü sağlayan ve girişimsel olmayan bir yöntemdir. Farklı EEG kayıt yöntemleri arasında, görsel bir uyarının deneye gösterilirken kayıt alınması yöntemi popüler çalışma yöntemleri arasındadır. Son zamanlarda görsel uyarıların belli bir frekansla gösterildiği durağan durum görsel uyarın potansiyeller (DDGUP) beyin bilgisayar arayüzlerinde (BBA) sıklıkla kullanılmaya başlanmıştır. Literatürde, DDGUP kayıtları analizi farklı öznelikleri kullanılmışsa da, bu sinyallerin durağan olmayan yapısı gereği dalgacık dönüşümü tabanlı özneliklerin kullanılması tercih edilmelidir. Bu çalışmada, DDGUP sinyalinden uyartım frekansının belirlenmesi için altı farklı ana dalgacık türünden çıkarılan temel öznelikler sınıflandırıcıların girişlerine uygulanarak en yüksek sınıflandırıcı performansını veren dalgacık türü ve sınıflandırıcı kombinasyonu tespit edilmeye çalışılmıştır. DDGUP sinyallerinden altı farklı dalgacık fonksiyonu kullanılarak iyi bilinen beş EEG frekans bantlarına ait enerji, varyans ve entropi öznelikleri çıkarılmıştır. Daha sonra bu öznelikler altı temel sınıflandırma algoritmasına uygulanarak performansları karşılaştırılmıştır. Çalışma hem denekler için bireysel olarak hem de tüm denekler birlikte kullanılarak tekrarlanmıştır. Elde edilen sonuçlara göre (i) ANOVA ile öznelik seçimi sınıflandırıcı performanslarını düşürmektedir, (ii) denekler tek olarak dikkate alındığında her denek için en yüksek başarıya ulaşan tek sınıflandırıcı ve ana dalgacık türü kombinasyonu bulunmamaktadır, (iii) tüm denekler birlikte değerlendirildiğinde topluluk sınıflandırma algoritması ile Reverse Biorthogonal ana dalgacığı kullanıldığında en yüksek sınıflandırıcı başarıya ulaşılmaktadır.

## Evaluation of wavelet features selected via statistical evidence from steady-state visually-evoked potentials to predict the stimulating frequency

### H I G H L I G H T S

- Eight distinct stimulation frequencies can be discriminated with the average accuracy of 36.1%
- Reverse Biorthogonal mother wavelet gives the maximum accuracy
- Ensemble learners achieve the maximum accuracy among classifiers

### Article Info

Research Article  
Received: 25.12.2019  
Accepted: 27.09.2020

### DOI:

10.17341/gazimmfd.664583

### Keywords:

Brain-computer interface,  
steady-state visually-evoked  
potentials,  
wavelet transform,  
analysis of variance,  
feature selection,  
machine learning

### ABSTRACT

Electroencephalography (EEG) is a noninvasive method to record brain activities. Among different EEG recording methods, the recording, while a visual stimulation is shown to the subject, is one of the most popular methods. Recently, steady-state visually-evoked potentials (SSVEP) where visual objects are blinking at a fixed frequency have been commonly-used method in brain-computer interfaces. Although various features extracted from SSVEP records have been used, the use of features from wavelet transform should be preferred due to the nonstationary structure of these signals. In this study, the combination of mother wavelet and classifier, which gives the highest accuracy to determine the stimulating frequency, is examined by applying common wavelet features to inputs of classifiers. Features of energy, variance, and entropy were extracted for well-known five EEG frequency bands using six different mother wavelets. Then, classifier performances of six basic classifiers were compared. This study was run for both each subjects individually and all subjects together. Results showed that (i) ANOVA-based feature selection reduces the performances, (ii) there is no unique combination of classifier and mother wavelet while evaluating each subject individually, (iii) the highest performance was achieved by combination of ensemble learner and Reverse Biorthogonal wavelet while evaluating all subjects together.

## 1. GİRİŞ (INTRODUCTION)

Beyin Bilgisayar Arayüzü (BBA), kişilerin sadece beyin aktivitesini (motor hareket hayali, duyu durumu, odaklanılan herhangi bir görsel veya işitsel uyarın vb.) yorumlayarak insanların dış dünya ile iletişim kurabilmelerini ve çeşitli elektronik cihazları kontrol edebilmelerini sağlayan bir sistemdir. Günümüzde üzerinde yoğun bir şekilde çalışılan araştırma konuları arasında yer almaktadır. BBA sayesinde ALS veya benzeri hastalıklar nedeniyle, merkezi sinir sistemindeki kasların kontrolü ile ilişkili nöronları hasar görmüş fakat zihinsel fonksiyonları ve bazı duyu organları sağlıklı olan bireylerin dış dünya ile iletişim kurabilmeleri ve böylece robot kolu, tekerlekli sandalye veya bilgisayar klavye ve faresi gibi cihazları düşünce gücüyle kontrol etmeleri mümkün olmaktadır [1]. Beyindeki bazı nörofizyolojik olguların algılanmasına dayanan BBA sistemlerinin temel yapı taşlarını, biyolojik sinyal ölçümüne yönelik yazılım ve donanımları, çeşitli sinyal işleme yöntemleri ve makine öğrenme algoritmaları oluşturmaktadır.

Elektroensefalografi (EEG) deneğin beyin aktivitelerinin ölçümünü sağlayan bir yöntemdir. EEG sinyallerinin, girişimsel olmayışı, maliyetinin nispeten az olması, yüksek zaman çözünürlüğü, taşınabilirliği ve uygulama kolaylığı sebepleri ile BBA çalışmalarında en yaygın şekilde kullanılan sinyal çeşididir [2,3]. EEG tabanlı BBA sistemleri kontrol sinyali üretmek için kullandıkları yöntemlere göre yedi gruba ayrılır: görsel uyarılmış potansiyeller, yavaş kortikal potansiyeller (SCP), uyarılmış potansiyellerin P300 bileşeni, sinir hücresi aktivitesi (SCN), zihni görev tepkisi, çoklu nöromekanizma ve duyuusal motor ritimleri (SMR'ler) [4].

Ayrıca, deneğe bazı görsel uyarınlar (resim, video, ışık) uygulanması esnasında kayıt edilen EEG sinyallerine görsel potansiyeller denir. Eğer görsel uyarının kendisi 6 Hz'den daha yüksek bir frekansta modüle edilerek uygulanırsa elde edilen görsel uyarılmış potansiyeller ise durağan durum görsel uyarılmış potansiyeller (DDGUP) adını alır [5]. DDGUP tabanlı BBA sistemi, daha yüksek sinyal / gürültü oranı (SNR) sağlaması, daha az eğitim süresi gerektirmesi ve daha hızlı bilgi transfer oranına (ITR) ulaşması gibi avantajları nedeniyle, diğer BBA sistemlerine göre son zamanlarda yaygın kullanılmaya başlayan bir araştırma konusu haline gelmiştir [6]. DDGUP tabanlı BBA sisteminde, bireylere bilgisayar ekranı veya harici ışık kaynağı tarafından farklı frekanslarda yanıp sönen görsel uyarınlar gösterilir. Bireyler bu uyarınlardan birine odaklandığında, beyinin görme ile ilgili bölgesi olan oksipital bölgede, odaklanılan uyarın ile aynı frekansta salınımlar gösteren örüntüler gözlemlenebilmektedir [7]. Bu örüntüler (yani farklı frekanstaki uyarınlar) çeşitli sinyal işleme teknikleri ve makine öğrenmesi algoritmaları kullanılarak ayırt edildiğinde, bir bilgisayar arayüzü kullanılarak uyarın ile ilişkili komut (çeşitli elektronik sistemlerin kontrolü gibi) çalıştırılabilmektedir [8]. DDGUP ile BBA sistemi tasarımı ile ilgili literatürde, genellikle en fazla dört farklı frekans kullanılarak görsel uyarın çeşitlerinin (şekil, boyut, ışık

kaynağı, renk vb.) aralarındaki farkları ve gösterilen uyarına dikkatle odaklanma durumunu inceleyen çeşitli yöntemlerle yapılmış çalışmalar bulunmaktadır. Bu çalışmalarda çoğunlukla sadece iki uyartım frekansı ile ilişkilendirilen komutlar sınıflandırılmaya çalışılmıştır. Örneğin, Zhang vd. [9] sürekli dalgacık dönüşümü (SDD) temelli yeni bir öznitelik önermiş ve öznitelikleri destek vektör makineleri (DVM) sınıflandırıcısı ile yaklaşık %95'e yakın başarı elde etmişlerdir. Başka bir çalışmada, Müller-Putz vd. [10] hızlı Fourier dönüşümü ile çıkardığı öznitelikleri maksimum ve minimum değerlere dayalı bir sınıflandırma yöntemi ile %44-88 aralıklarında başarı elde etmişlerdir. Diez vd. [11] ise kısa zamanlı Fourier dönüşümünden elde edilen öznitelikleri de ekleyerek oluşturduğu modelden %65-100 arasında başarı elde etmişlerdir. İstatistikte sık kullanılan t testi ile öznitelik seçimi de yapılan başka bir çalışmada, Heidari ve Einalou [12] dalgacık dönüşümü (DD) özniteliklerini kullanarak Karar ağaçları (KA), Bayes ve DVM sınıflandırıcıları ile oluşturduğu modellerden %83,32 başarı elde etmişlerdir. Yeh vd. [13] tarafından DDGUPE ait faz ve genlik değerleri de öznitelikler olarak alınıp DVM sınıflandırıcısı ile %89,88 başarıya ulaşmışlardır.

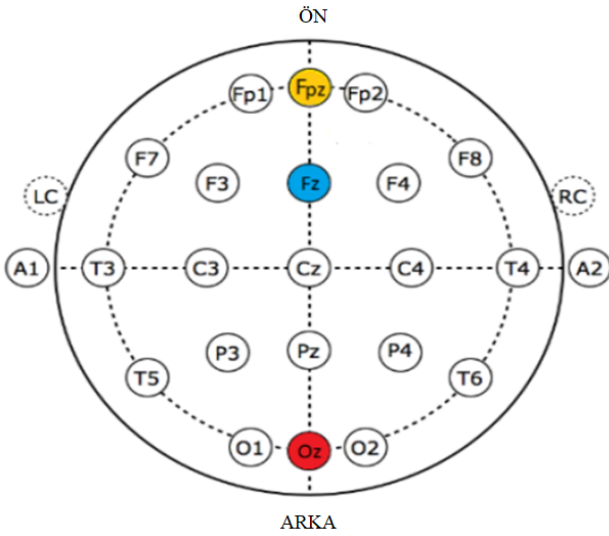
Sadece iki farklı frekansın kullanıldığı çalışmaların yanında, DDGUP sinyallerinden aynı anda daha fazla uyartım frekansının tahmin edilmeye çalışıldığı çalışmalar oldukça azdır. Yakın tarihli bir çalışmada, Sayılğan vd. [14] DDGUP sinyallerinden beş farklı kutudan hangisinin seçildiğinin tahmininde özbağlanımlı (AR) parametreler, Hjorth ve güç spektral yoğunluğu olmak üzere çıkarılan üç farklı öznitelik seti kullanılarak Naive Bayes, aşırı öğrenme makinesi (ELM) ve DVM algoritmalarını kullanmışlardır. Bu çalışmada en yüksek sınıflandırıcı başarımları olarak %83,42 değerine güç spektral yoğunluk özniteliği ve aşırı öğrenme makinesi algoritması çifti ile ulaşmışlardır. Fakat bu çalışmada uyartım frekansı sabit olup beş adet kutu rengine ve konumuna göre DDGUP sinyalinden hangi kutunun seçildiği tahmin edilmeye çalışılmıştır. Aynı ekip tarafından yürütülen başka bir çalışmada ise DDGUP sinyallerinden elde edilen zaman alanı öznitelikleri ile 6 temel sınıflandırıcının 24 farklı algoritması kullanılarak uyarın frekansı tahmin edilmeye çalışılmıştır. Bireysel değerlendirmede deneklerin DDGUP sinyallerinden uyartım frekansları %22,20-42,90 arasında doğrulukla tahmin edilmiştir. Aynı çalışmada tüm denekler birlikte değerlendirildiğinde ise %20,00 ortalama performans elde edilebilmiştir [15]. Alternatif olarak, DDGUP sinyalleri ile uygulanan görselin frekansının tahmininde sınıflandırıcılar yerine Kanonik Korelasyon Analizi (CCA) yönteminin kullanıldığı çalışmalar da mevcuttur. Yakın tarihli örnek bir çalışmada, Oojha vd. [16] DDGUP sinyallerinden dört farklı uyartım frekansının tahmin edilmesi için CCA kullanmışlardır. Bu yöntemle %42,65 ile %54,45 arasında değişen ortalama %47,25'lik bir başarıya ulaşmışlardır. Buna göre daha fazla sayıda uyartım frekansını aynı anda daha yüksek başarımla tahmin edebilen yöntemlere ihtiyaç duyulduğu açıkça görülmektedir. Bu çalışmada, EEG sinyalleri üzerinden farklı ana dalgacık türlerine göre yaygın

dalgacık özniteliklerine literatürde yaygın kullanılan ve farklı parametrelerin denendiği 6 farklı sınıflandırıcı ile deneye uygulanan sekiz farklı uyarım frekansı aynı anda tahmin edilmeye çalışılmıştır. Bunun için internetteki “www.setzner.com” web adresinden açık erişim yoluyla elde edilen veri seti [17] kullanılarak sekiz farklı frekansta (6-12 Hz arasında) yanıp sönen bir LCD monitöre baktıkları sırada kayıt edilen EEG sinyalleri kullanılmıştır. Ayrıca varyans analizi (ANOVA) ile aralarında istatistiksel olarak fark bulunan öznitelikler kullanılarak çalışma tekrarlanmış ve bu öznitelik seçiminin etkisinin olup olmadığı da irdelenmiştir.

## 2. MATERYAL VE YÖNTEM (MATERIAL AND METHOD)

### 2.1. Veri Seti Tanıtımı (Data Set Description)

Bu çalışmada Adnan Vilic tarafından kaydedilen DDGUP sinyallerini içeren veri seti (AVI SSVEP Dataset) kullanılmıştır [17]. Veri seti internet üzerinden ücretsiz erişime açık olarak sunulmaktadır. Veri setinde, DDGUP sinyallerinin farklı frekanslardaki yanıtlarını tetiklemek için yanıp sönen hedefe bakan sağlıklı bireylerin EEG ölçümleri mevcuttur. Tüm veriler üç elektrot (Oz, Fpz, Fz) kullanılarak kaydedilmiştir. Elektrot yerleşim düzeni için standart uluslararası 10-20 sistemi dikkate alınarak Oz sinyal elektrotu, Fz referans elektrot ve Fpz ise toprak elektrotu olacak şekilde konumlandırılmıştır (Şekil 1).



**Şekil 1.** Uluslararası 10-20 elektrot yerleşim düzeni ve çalışmada kullanılan elektrotların konumları (International standard 10-20 electrode layout and electrode locations used in the study)

Denekler 120 Hz yenileme hızına ve 1680x1050 piksel çözünürlüğüne sahip bir LCD bilgisayar ekranı karşısına oturmuşlardır. Kontrast ve parlaklık maksimum olacak şekilde ayarlanmıştır. Deneklere sunulan hedefler ekranda 2,89 cm<sup>2</sup>’lik bir alana sahip olacak şekilde düzenlenmiştir. Görsel uyarının deneklere gösterilmesi Microsoft Silverlight’ta bir uygulama geliştirilmiş ve bu uygulama Windows 8 tabanlı bir bilgisayarda çalıştırılmıştır. Elde

edilen verilere şebeke frekansında (50Hz) bir analog çentik filtre uygulanmıştır [17].

Bu deneyde, bireyler, rengi siyahtan beyaza hızla değişen tek bir yanıp sönen hedefe baktıkları bir monitörden 60 cm uzağa oturmuşlardır. Deneyin uyarımı monitörde sunulan 8 farklı frekansta (6 - 6,5 - 7 - 7,5 - 8,2 - 9,3 - 10 - 12 Hz) yanıp sönen bir kutudur. Veri seti dört farklı katılımcı ile yapılan dört seanstan (oturumdan) oluşmaktadır. Bir oturumdaki her deneme 30 saniye sürmekte ve katılımcılar denemeler arasında kısa bir mola vermektedir. Her bir frekans için denemeler en az 3 kez tekrarlanmıştır. Tablo 1’de deneye katılan sağlıklı bireylerin bazı fizyolojik bilgilerinin (cinsiyetini ve yaşını) belirten bir liste sunulmuştur.

**Tablo 1.** Tek hedef uyarımı DDGUP deneyine katılanların listesi

(List of participants for single target flickering)

Denek	1	2	3	4
Cinsiyet	Erkek	Erkek	Erkek	Kadın
Yaş	32	27	27	31

### 2.2. Öznitelik Çıkarma (Feature Extraction)

DDGUP tabanlı BBA sistemlerinde öznitelik vektörleri oluşturmak için genellikle EEG işaretlerinin zaman ve frekans bilgilerini içeren farklı yöntemlerden faydalanılır [18, 19]. Öznitelik çıkarma, DDGUP sinyallerinin içinde gizli olan bilginin yorumlama ve tanımlama amaçlı olarak elde edilmesi işlemidir [20].

Bu çalışmada, öznitelik vektörleri olarak, düşük frekanslardaki sinyaller için geniş, yüksek frekanslardaki sinyaller için dar olacak şekilde değişebilen pencere boyutlarına sahip olduğundan bütün frekans aralıklarında uygun zaman-frekans çözünürlüğü sağlayan ayrık dalgacık dönüşümü yöntemi [21] kullanılarak öznitelikler oluşturulmuştur. EEG sinyali durağan olmayan [22] bir sinyal olduğundan dolayı ve sinyalin yalnızca spektrum analizlerini değil, aynı zamanda sinyalin zamanda spektral davranışlarını da inceleyebilmek adına bu yöntem kullanılmıştır. İlk olarak, işaretteki yüksek frekanslı değişimleri analiz etmek için işaret yüksek geçiren filtreler serisinden, alçak frekanslı değişimleri analiz etmek için ise alçak geçiren filtreler serisinden geçirilir. Her filtre veya dönüşüm dalgacık dönüşümü ayrışma seviyesi (k) olarak adlandırılır. Alçak geçiren filtre çıkışındaki işaretin alt-örneklenmesiyle elde edilen işaret yaklaşıklık katsayıları (A<sub>k</sub>) ve yüksek geçiren filtre çıkışındaki ise detay katsayıları (D<sub>k</sub>) olarak adlandırılır. Farklı Ayrık Dalgacık Dönüşümü (ADD) fonksiyonları (Haar, Daubechies, Symlet, Coiflet, Biorthogonal, Reverse Biorthogonal) kullanılarak EEG sinyalleri frekans alt bantlarına ayrıştırılmış (delta, teta, alfa, beta, gama) ve her bir bantın enerji, entropi ve varyans [23, 24, 25] değerleri hesaplanmıştır. Böylece frekans bantlarını temsil edilen bir dizi öznitelik elde edilmiştir.

Her ayrışma seviyesindeki enerji, aşağıdaki denklemler kullanılarak hesaplanır:

$$E_j = C_j^2 \quad (1)$$

Denklemdaki  $C_j$ , j-inci dalgacık katsayısını ifade etmektedir [25, 26].  $C_j$ , EEG alt bantları (delta, teta, alfa, beta, gama) için ayrı ayrı hesaplanmıştır. Toplam enerji olarak ifade edilirse:

$$E_{toplam}(m) = \frac{1}{N} \sum_{j=1}^N E_j \quad (2)$$

Burada,  $m$  dalgacık seviyesi, aynı zamanda karşılık geldiği EEG alt bandı,  $N$  EEG frekans bandını oluşturan dalgacık katsayılarının toplam sayısıdır [25, 26].

Bir diğer öznitelik olan entropi ( $WS_f$ ) ise aşağıdaki gibi hesaplanmaktadır:

$WS_f = -\sum_{j \in f} (p_j \log_2(p_j))$  (3) Burada ilgilenilen frekans bandındaki tüm  $f$  frekans değerlerinin enerjilerinin olasılığı  $p_j$  olarak hesaplanmaktadır [25] ve  $p_j$  ise ilgilenilen frekansın enerjisinin, toplam enerjiye bölünmesiyle ( $\frac{E_j}{E_{toplam}}$ ) elde edilen değerdir. Dalgacık katsayıları bilindikten sonra, her dalgacık ölçeği seviyesinin varyansı şu şekilde hesaplanabilir:

$$Dm = \sum_{k=1}^{N_m} (C_k - \bar{C}_k)^2 \quad (4)$$

Burada  $N_m$ ,  $m$  dalgacık seviyesine karşılık gelen katsayıların toplamının sayısıdır,  $C_k$ ,  $m$  seviyesindeki dalgacık katsayılarını ve  $\bar{C}_k$  ise  $m$  seviyesindeki tüm dalgacık katsayılarının ortalama değeridir [26].

Farklı kombinasyonlardan oluşan bu öznitelikler giriş vektörü olarak kullanılmıştır. Tablo 2'de dalgacık katsayılarının bileşenleri ve frekans aralıkları sunulmuştur.

**Tablo 2.** 512 Hz örneklem frekansı için alt bantların ayrışımı  
(Decomposition of subbands for sampling frequency of 512 Hz)

Seviye	Bileşen (A ve D)	Frekans Aralığı (Hz)	Merkez Frekans (Hz)
9	D1	0-1	0,5
8	D2	1-2	1,5
7	D3	2-4	3
6	D4	4-8	6
5	D5	8-16	12
4	D6	16-32	24
3	D7	32-64	48
2	D8	64-128	96
1	D9	128-256	192
0	A9	256-512	384

### 2.3. Öznitelik Seçimi (Feature Selection)

Öznitelik seçimi, öznitelik çıkarımı sonrasında veri setimizde var olan öznitelik vektörleri arasından sınıflandırıcı performansını arttıracak ve hesap yükümüzü azaltacak önemli özniteliklerin seçilimi veya azaltılması

olarak adlandırılmaktadır [18, 27, 28]. Öznitelik seçimi yapıldığında veri boyutu küçüldüğünden dolayı hesaplama yükü oranını düşürmeyi garantilerken tahmin oranını arttırmayı garantilememektedir [29].

Bu çalışmada, sekiz farklı frekansın ayırt edilmesi problemi için dalgacık öznitelikleri oluşturulmuş ve öznitelikler arasında anlamlı farklılıklar olup olmadığını görmek için istatistiksel analizde kullanılan IBM SPSS 24 [30] programı kullanılmıştır. Tek yönlü varyans analizi (ANOVA) uygulanarak uyarım frekanslarına ait öznitelikler arasında istatistiksel olarak anlamlı bir fark olup olmadığına bakılmıştır. Bu testin uygulanabilmesi için öncelikle veriler içerisinde aykırı değer olup olmadığı incelenmiş ve normal dağılıma uygun olup olmadığı araştırılmıştır. Veriler üzerinde gözle yapılan inceleme sırasında aykırı veriye rastlanmamıştır. Daha sonra veriler üzerinde SPSS yazılımı ile sırasıyla normallik testi, varyans homojenliği testi ve tek yönlü ANOVA testi uygulanmıştır. Normallik testi sonucunda eğer verilerin normal dağılıma uygun olmadığı tespit edilseydi, çeşitli dönüşümler uygulanarak verilerin normalleştirilmesi sağlanmaya çalışılırdı. Levene testi ile varyans homojenliği test edilmiştir. Aslında bu test sadece ANOVA testi sonuçlarına göre uygulanacak post-hoc testlerin seçiminde kullanılacak olmasına rağmen, nu çalışmada varyans eşitliği de test edilmiştir. Son olarak tek yönlü ANOVA testi ile gruplar (uyarım frekansları) arasında istatistiksel olarak anlamlı bir fark olup olmadığı test edilmiştir. ANOVA testi gruplar arasında fark olduğunu gösterebilse de, hangi grupların diğerlerin ayrıştığını gösterme kapasitesine sahip değildir. Fakat bu çalışmada amaçlanan ANOVA testi ve post-hoc testi ile ayrıştırma yapmak olmadığından, post-hoc testleri yapılmamıştır. Seçilen istatistiksel anlamlılık düzeyine ( $p$ ) göre, uyarım frekanslarına ait ölçümler arasında istatistiksel olarak anlamlı bir fark tespit edilen öznitelikler seçilen öznitelikler havuzuna eklenmiştir. Bu çalışmada istatistiksel anlamlılık düzeyi ( $\alpha$ ) 0,05 olarak seçilmiştir.

### 2.4. Sınıflandırma (Classification)

DDGUP sinyalini tanımlamak ve bu sinyali komuta dönüştürebilmek için sinyalden çıkarılan özniteliklerin sınıflandırılması gerekmektedir [18]. Sınıflandırma için, hangi sınıfa ait olduğu önceden bilinen belirli sayıda öznitelik vektörünün oluşturduğu "veri kümeleri" kullanılarak eğitilir. Bu eğitim sonrasında algoritmasının gerektirdiği parametreleri belirlenmiş olan sınıflandırıcılar, eğitimde kullanılmayan başka öznitelik vektörleri uygulanarak test edilirler [31].

Bu çalışmada DDGUP sinyalinden çıkarılan öznitelik vektörleri altı temel sınıflandırıcı ile denenmiştir. Sınıflandırma işlemi için MATLAB yazılımında mevcut olan "Classifier Learner" uygulaması [32] kullanılmıştır ve uygulamada yer alan bütün sınıflandırıcıların ve alt parametrelerinin veri üzerindeki performansları incelenmiştir. Bu sınıflandırıcılar şu algoritmalarla oluşmaktadır: Karar Ağaçları (*Fine, Medium, Coarse*), Ayırıcı Analizi (*Linear, Quadratic*), Naive Bayes (*Gaussian*,

*Kernel-Based*), Destek Vektör Makineleri (*Linear, Quadratic, Cubic, Fine Gaussian, Medium Gaussian, Coarse Gaussian*), k-En Yakın Komşuluk (*Fine, Medium, Coarse, Cubic, Cosine, Weighted*) ve Topluluk Sınıflandırıcıları (*Boosted, Bagged, Subspace discriminant, Subspace KNN, RUSBoosted Trees*).

Takip eden alt bölümlerde bu algoritmalarla ilgili temel bilgiler verilmeye çalışılmıştır. Daha detaylı bilgi için bu çalışmanın kaynaklarında belirtilen kitap, makale ve yardım dokümanları kullanılabilir [27, 28, 29, 30, 31, 32].

#### 2.4.1. Karar Ağaçları (Decision Trees)

Karar ağaçları veride bulunan bağımsız değişkenleri bilgi kazancına göre düğümlere ayıran ve tahmin esnasında ilgili aralıktan bir değer sorulduğunda cevap olarak aralıktaki (eğitim esnasında öğrendiği) ortalamayı söyleyen bir makine öğrenme algoritmasıdır [27]. Karar ağacı algoritmalarında düğümlere bölünmenin nasıl gerçekleşeceği ağacın doğruluğunu etkileyen faktörlerdendir. Bu nedenle karar ağaçları, bir düğümü iki veya daha fazla alt düğümde bölmeye karar vermek için birden fazla algoritma kullanır [28]. Algoritma seçimi, hedef değişkenin türüne bağlı olmakla birlikte en sık kullanılan algoritmalar; kategorik değişkenler için Entropi, Gini, Sınıflandırma Hatası yöntemi şeklindedir. Bu çalışmada entropi kullanılmıştır. Entropi, verilerimizle ilgili belirsizliğin bir ölçüsüdür. Karar ağaçları bulunan durumun entropi değerini (rastgelelik derecesini) düşüren seçimler yaparak bilgi kazancını (information gain) maksimize etmeye çalışır. Bunun için de her düğümde hata fonksiyonunu tekrar hesaplayıp en düşük hataya sahip soruyu/durumu seçer.

#### 2.4.2. Ayırıcı analizi (Discriminant analysis)

Ayırıcı analiz sınıflandırıcısının amacı, verideki bağımsız değişkenleri homojen gruplara doğru bir şekilde bölmektir. Doğrusal Ayırım Analizi (LDA), grup elemanlarını belirler ve her eleman için farklı gruplara ait olma olasılığını hesaplar. Eleman daha sonra olasılık puanı en yüksek olan gruba atanır. LDA, öngörücülerin normal olarak dağıtıldığını (Gauss dağılımı) ve farklı sınıfların sınıfa özgü elemanlara ve eşit varyans / kovaryansa sahip olduğunu varsayan bir doğrusal ayırıcı fonksiyon oluşturur. Quadratic Diskriminant Analizi (QDA) ise, varyans / kovaryans eşitliğini kabul etmez. Yani, QDA için kovaryans matrisi her sınıf için farklı olabilir. Dolayısıyla ayırıcı fonksiyonu ikinci dereceden olacak şekilde yapılındırır [27, 28, 31].

#### 2.4.3. Naive bayes sınıflandırıcısı (Naive bayes classifier)

Naive Bayes sınıflandırıcısı Thomas Bayes tarafından 1812 yılında bulunan koşullu olasılık hesaplama formülüne dayanan bir makine öğrenme algoritmasıdır [28, 32]. Bayes teorimi olarak adlandırılan bu teorem şu şekilde hesaplanmaktadır:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (5)$$

Burada her bir terim sırasıyla şu şekilde tanımlanmaktadır [27]:

- $P(A \setminus B) = B$  olayı gerçekleştiğinde A olayının gerçekleşme olasılığı,
- $P(A) = A$  olayının gerçekleşme olasılığı,
- $P(B \setminus A) = A$  olayı gerçekleştiğinde B olayının gerçekleşme olasılığı,
- $P(B) = B$  olayının gerçekleşme olasılığı.

#### 2.4.4. Destek Vektör makineleri (Support vector machines)

Destek Vektör Makineleri (DVM), istatistiksel öğrenme teorisine dayalı bir gözetimli makine öğrenmesi algoritmasıdır [18]. DVM, temel olarak iki veya daha fazla sınıfa ait verileri birbirinden en uygun şekilde ayırmak için hiper düzlemler kullanır. DVM, veri setinin doğrusal olarak ayrılabilir ve ayrılamama durumuna göre ikiye ayrılmaktadır. Doğrusal olarak ayrılabilen verilerde, karar doğrusunun yeni katılacak olan veriye karşı dayanıklı olabilmesi için sınır çizgisinin, iki sınıfın sınır çizgilerine en yakın uzaklıkta olması gerekmektedir. Bu sınır çizgisine en yakın noktalar, destek noktaları olarak adlandırılmaktadır. Veri kümesinin doğrusal olarak sınıflandırılması mümkün olmayan durumlarda ise, her bir verinin üst özellik uzayıyla eşlenmesi ve yine bu yeni uzayda bir hiper düzlem yardımıyla sınıflandırılması “çekirdek hilesi” kullanılarak gerçekleştirilmektedir. Bu hile verileri, çekirdek işlevi  $K(x, y)$  (Kernel fonksiyonu) kullanarak, genellikle çok daha yüksek boyutlara sahip başka bir boşluğa dolaylı olarak eşlemekle sağlanır. EEG temelli BBA araştırmalarında genellikle kullanılan çekirdekler Gauss veya Radyal Temel Fonksiyon (RBF) çekirdeğidir [31].

#### 2.4.5. k-en yakın komşuluk (k-nearest neighbor)

k-En Yakın Komşuluk algoritması, 1967 yılında T. M. Cover ve P. E. Hart tarafından önerilmiştir [27, 28]. Algoritma, sınıfları belli olan bir örnek kümesindeki verilerden yararlanılarak çalışır. Örnek veri setine katılacak olan yeni elemanın, mevcut elemanlara göre uzaklığı Euclidean, Manhattan gibi uzaklık fonksiyonları ile hesaplanıp, “k” sayıda yakın komşuluğuna bakılır. Bu komşulardan hangisi fazla ise yeni verinin etiketi o sınıfa ait kararını oluşturur.

#### 2.4.6. Topluluk öğrenimi sınıflandırıcısı (Ensemble learning classifier)

Topluluk öğrenmesi, varyansı azaltmak (bagging), sapmayı (boosting) ve/veya tahminleri iyileştirmek (stacking) için birden fazla makine öğrenme tekniklerini tek bir tahmin modelinde (sınıflandırıcıda) birleştiren meta algoritmalarıdır [32]. Bu algoritmaya göre verilerdeki olası gürültü, çakışan veri dağılımları ve aykırı değerler nedeniyle genellikle tek sınıflandırıcılar belirli ve kesin bir sınıflandırma doğruluğu elde edemezler. Yani her sınıflandırma problemi için en iyi şekilde çalışan tek bir model (sınıflandırıcı) yoktur. Bu nedenle topluluk öğrenmesi yöntemlerine ihtiyaç duyulmaktadır.

## 2.5. Sınıflandırıcı Performanslarını Değerlendirme (Evaluating Classifier Performances)

Bu çalışmada kullanılan sınıflandırma algoritmalarının performanslarını değerlendirmek için  $k$ -katlamalı çapraz doğrulama ve hata matrisi değerlendirme kriterleri kullanılmıştır.

### 2.5.1. $k$ - katlamalı çapraz doğrulama ( $k$ -fold cross validation)

Örüntü tanıma (veya makine öğrenmesi) çalışmalarında çıkarılan özneliklerden oluşan veri seti eğitim ve test kümeleri olmak üzere ikiye ayrılır. Eğitim kümesi ile kullanılan sınıflandırma algoritmasının kendisine özgü parametrelerinin düzenlenmesi sağlanır. Test kümesi ile de bu parametrelere göre algoritmanın vereceği cevaplardaki hatalar tespit edilir. Eğitim ve test kümesi olarak veri setini ayırmada en sık kullanılan yöntemlerden birisi  $k$ -katlamalı çapraz doğrulama yöntemidir. Bu yöntemde veri seti rasgele  $k$  parçaya bölünür. Her seferinde  $k-1$  parça algoritmanın eğitimi için kullanılırken, kalan 1 parça algoritmanın testi için kullanılır. Bu işlem tüm parçalar test amacıyla kullanılıncaya kadar tekrarlanır. Her seferinde test hataları kayıt edilerek son parçadan sonra hataların ortalaması hesaplanır. Kullanılan sınıflandırıcı algoritmasının performansı bu şekilde değerlendirilir [33]. Bu çalışmada veri seti 10 eşit parçaya bölünmüştür [34, 35]. Her tekrarda, mavi boyalı alan test setine ayrılırken diğer renkli parçalar eğitim için kullanılmıştır (Şekil 2).

### 2.5.2. Hata matrisi (Confusion matrix)

Sınıflandırıcı başarımını değerlendirmek için öncelikle hata matrisi hesaplanmaktadır. Sınıflandırma algoritmasının test kümesine verdiği cevapların veri setindeki gerçek değerlerle karşılaştırılması ile oluşturulur. İki durumlu problemlerde dört farklı durumdan oluşan bir tablo şeklindedir [27, 28]:

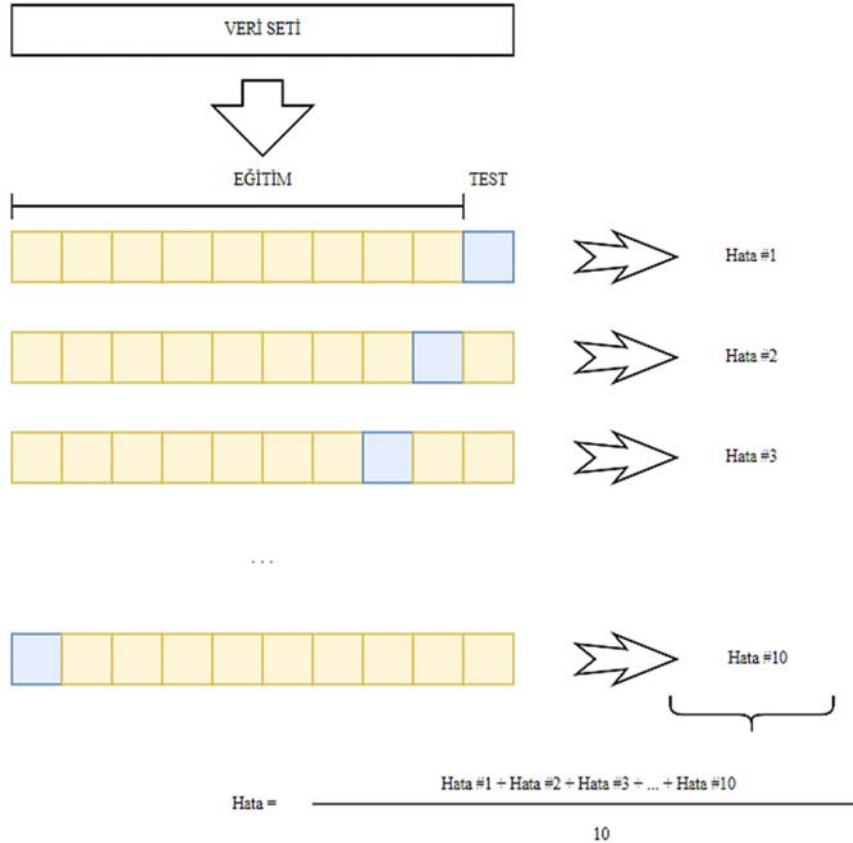
- Doğruya doğru (True Positive - TP)
- Doğruya yanlış (True Negative - TN)
- Yanlışta doğru (False Positive - FP)
- Yanlışta yanlış (False Negative - FN)

Bu değerlere bağlı olarak sınıflandırıcı performansı olarak doğruluk değeri hesaplanır [27, 28]:

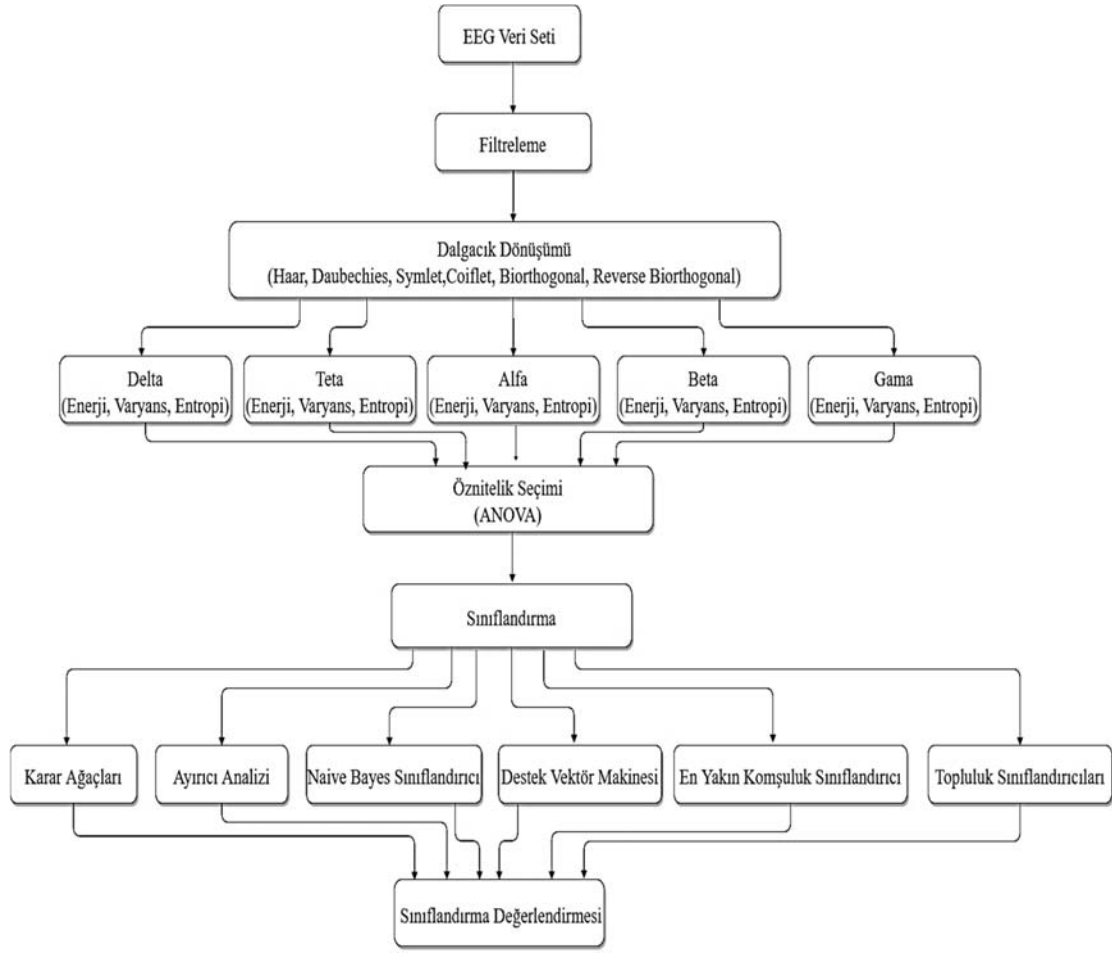
$$\text{Doğruluk (ACC)} = \frac{TP+TN}{TP+FN+FP+TN} \quad (6)$$

### 2.5. Uygulama Detayları (Implementation Details)

Yukarıda açıklanan yöntemler birleştirilerek aşağıdaki akış şeması şeklinde özetlenmiştir (Şekil 3). Buna göre, DDGUP verilerini içeren EEG verisetinden altı farklı ana dalgacık türü kullanılarak dalgacık katsayıları hesaplanır. Algoritmanın geri kalanı bu altı farklı dalgacık dönüşümü için tekrarlanmaktadır. Bu dalgacık katsayıları beş farklı EEG frekans bandına karşılık gelecek şekilde gruplanır. Bu şekilde, her bir frekans bandı için elde edilmiş olan dalgacık



Şekil 2. Sınıflandırmada kullanılan 10-kat çapraz doğrulama modeli (10-fold cross validation model used in classification)



Şekil 3. Önerilen çalışmanın akış şeması (Flowchart of the proposed study)

katsayılarından enerji, varyans ve entropi değerleri hesaplanarak öznitelik vektörü oluşturulmuş olur. SPSS yazılım paketi kullanılarak bu öznitelikler üzerinde varyans analizi (ANOVA) ile istatistiksel anlamlı farklılık gösteren öznitelikler seçilir. Algoritmanın bu aşamadan sonrasında yer alan sınıflandırma algoritmalarının girişine önce tüm öznitelikler ve daha sonra sadece ANOVA ile seçilen öznitelikler uygulanır. Altı temel sınıflandırıcı algoritmasının farklı alt türleri dahil edildiğinde yirmi dört farklı sınıflandırma işlemi sırasıyla çalıştırılmıştır. Böylece 6 farklı dalgacık türü, öznitelik seçiminin uygulanma durumu, 6 temel sınıflandırıcının toplam 24 farklı algoritması ile sınıflandırma işleminin koşturulduğu ve hem 4 denek için ayrı hem de 4 denek birlikte çalıştırıldığı düşünüldüğünde, bu çalışma kapsamında toplam 1440 kez sınıflandırma algoritması çalıştırılmıştır.

### 3. SONUÇLAR VE TARTIŞMALAR (RESULTS AND DISCUSSIONS)

Bu çalışmada kullanılan DDGUP verileri “www.setzner.com” [17] sitesinden açık erişim yoluyla, veri seti sahibinin izni dahilinde alınmıştır ve analizler için oluşturulan algoritmalar MATLAB programı [32] ve IBM SPSS 24 paket programı [30] kullanılarak test edilmiş ve 600

sınıflandırılma performans ölçütleri ise yine MATLAB programıyla incelenmiştir. Her bir sınıflandırma algoritmasının performansı doğruluk kıstası kullanılarak değerlendirilmiştir. Sekiz farklı frekansı sınıflandırmak için EEG sinyalinin alt bantlarındaki (delta, teta, alfa, beta, gama) enerji, varyans ve entropi gibi önemli üç öznitelik, altı farklı dalgacık fonksiyonu (Haar, Daubechies, Symlet, Coiflet, Biorthogonal, Reverse Biorthogonal) kullanılarak hesaplanmıştır. Bu çalışmada öznitelik seçimi için tek yönlü ANOVA testi ile yapılmıştır. Bu testin uygulanabilmesi için verilerin normal dağılmış olması ön koşulu vardır [30]. Bunun için SPSS yazılımı kullanılarak normallik testi uygulanmıştır ve elde edilen istatistiksel anlamlılık değerleri (p) kayıt edilmiştir (Tablo 3). Bu tablodaki tüm değerler seçilen 0,05 anlamlılık düzeyinden büyük olduğu için, tüm özniteliklerin normal dağılıma uydukları kabul edilmektedir. Tek yönlü ANOVA testi uygulanırken ikinci adımda varyans homojenliği testi uygulanmaktadır. Bu amaçla Levene istatistiği tercih edilmektedir [30]. Genellikle varyans homojenliği testi sonuçları post-hoc testlerinde uygulanmaktadır ve anlamlılık değeri 0,05’ten büyük veya eşit olması ( $p \geq 0,05$ ) halinde varyans homojen olarak değilse heterojen olarak kabul edilmektedir. Bu sonuca göre ANOVA testi sonrası uygulanacak post-hoc testinin seçimi yapılır. Bu çalışmadaki öznitelikler üzerinde yapılan test



**Tablo 3.** DDGUP verilerinden çıkarılan özniteliklerin normallik testi sonuçları  
(Normality test results of features extracted from SSVEP data)

Öznitelikler	Normallik Testi Anlamlılık Değerleri (p)				
	Delta	Teta	Alfa	Beta	Gama
Haar Enerji	0,200	0,258	0,179	0,191	0,166
Haar Varyans	0,127	0,256	0,153	0,123	0,145
Haar Entropi	0,117	0,070	0,085	0,061	0,097
Db4 Enerji	0,200	0,468	0,356	0,345	0,258
Db4 Varyans	0,082	0,143	0,124	0,263	0,066
Db4 Entropi	0,200	0,259	0,149	0,179	0,109
Sym4 Enerji	0,140	0,091	0,119	0,413	0,234
Sym4 Varyans	0,117	0,070	0,142	0,256	0,080
Sym4 Entropi	0,200	0,468	0,154	0,374	0,099
Coif1 Enerji	0,082	0,143	0,175	0,219	0,123
Coif1 Varyans	0,166	0,092	0,200	0,110	0,254
Coif1 Entropi	0,154	0,155	0,147	0,201	0,125
Rbio2.8 Enerji	0,126	0,080	0,200	0,119	0,096
Rbio2.8 Varyans	0,200	0,138	0,169	0,104	0,200
Rbio2.8 Entropi	0,069	0,094	0,086	0,073	0,051
Bior3.5 Enerji	0,093	0,071	0,076	0,087	0,120
Bior3.5 Varyans	0,062	0,052	0,089	0,060	0,052
Bior3.5 Entropi	0,100	0,069	0,117	0,090	0,109

sonucunda anlamlılık değeri 0,62 olarak bulunduğu gruplarda varyans homojenliğinden söz edilebilir (Tablo 4).

**Tablo 4.** Normal dağılıma sahip verilerin varyans homojenliği testi  
(Variance homogeneity test of normally distributed data)

Varyans Homojenliği Testi			
Levene İstatistiği	df1	df2	Anlamlılık (p)
2,508	7	87	0,62

Öznitelik seçiminde son adım olarak tek yönlü ANOVA testi ile gruplar arasındaki farklılıklar, anlamlılık değeri ( $p \leq 0,05$ ) kıstasına göre tespit edilmiştir (Tablo 5). Tabloda istatistiksel olarak anlamlı bulunan öznitelikler ait oldukları EEG frekans bantlarına göre “✓” ile gösterilmiştir. Buna göre Haar ve Bio3.5 ana dalgacıklarına ait özniteliklerin seçilmedikleri görülmektedir. Ayrıca Beta ve Gama bantlarındaki özniteliklerin de seçilmedikleri görülmektedir. ANOVA testi gruplar arasında anlamlı bir fark olduğunu göstermekle birlikte hangi grupların birbirlerinden ayrıldığı vermemektedir ve bu amaçla post-hoc testlerinin yapılması tavsiye edilmektedir. Fakat bu çalışmada amaçlanan ANOVA testi ile grupları birbirinden ayırmak olmadığı ve sadece anlamlı bir farkın olduğuna dair bir tespitin yeterli olmasından dolayı, bu çalışmada pot-hoc testleri uygulanmamıştır. Sınıflandırma aşamasında, öncelikle öznitelik seçimi yapılmadan bütün öznitelikler kullanılarak sınıflandırma algoritmaları çalıştırılarak performansları kayıt edilmiştir (Tablo 6). Birinci denek için Bior3.5 ana dalgacığı kullanılan öznitelikler ve LDA sınıflandırıcısı ile %55,6, ikinci denek için Bior3.5 ana dalgacığı kullanılan öznitelikler ve Ensemble Learner sınıflandırıcısı ile %26,9, üçüncü denek için Haar ve DB4 ana dalgacığı kullanılan öznitelikler ve Naive Bayes sınıflandırıcısı ile %33,3, ve

dördüncü denek için Bior3.5, DB4 ve Sym4 ana dalgacığı kullanılan öznitelikler ve Naive Bayes - Ensemble Learner sınıflandırıcısı ile %28,6 sınıflandırıcı başarımları elde edilmiştir. Bu tabloya göre, ortalama %36,1 başarımlar elde edildiği ve Bior3.5 ana dalgacığı genel olarak iyi sonuç vermekle birlikte tüm denekler için ortak bir dalgacık türü ve sınıflandırıcı belirlenemediği görülmektedir. Daha sonra, sadece seçilen özniteliklerin dört farklı katılımcı için sınıflandırma algoritmaları çalıştırılarak elde edilen en yüksek sınıflandırıcı performansları kayıt edilmiştir (Tablo 7). Birinci denek için Coif1 ve Sym4 ana dalgacığı kullanılan öznitelikler ve KNN - Ensemble Learner sınıflandırıcısı ile %25,9, ikinci denek için tüm ana dalgacık türlerindeki öznitelikler ve KNN - SVM - Ensemble Learner sınıflandırıcısı ile %23,1, üçüncü denek için Sym4 ana dalgacığı kullanılan öznitelikler ve LDA sınıflandırıcısı ile %38,1, ve dördüncü denek için RBio2.8 ana dalgacığı kullanılan öznitelikler ve Naive Bayes sınıflandırıcısı ile %38,1 sınıflandırıcı başarımları elde edilmiştir. Bu tabloya göre, ortalama %31,3 başarımlar elde edildiği ve genel olarak tüm denekler için iyi sonuç veren ortak bir dalgacık türü ve sınıflandırıcı belirlenemediği görülmektedir. Son olarak, bütün katılımcıların EEG kayıtlarının birlikte değerlendirilmesi için sınıflandırma algoritmaları hem öznitelik seçimi yapılmadan tüm öznitelikler için hem de öznitelik seçimi yapıldıktan sonra çalıştırılarak sınıflandırma performansları kayıt edilmiştir (Tablo 8). Bu sonuçlara göre tüm denekler birlikte değerlendirildiğinde en yüksek ortalama sınıflandırıcı başarımlarına tüm sınıflandırıcılarda Rbio2.8 ana dalgacık kullanıldığı zaman ulaşılmaktadır. Tabloda, öznitelik seçimi yapıldığı zamanki sınıflandırıcı başarımları (\*) işareti ile gösterilmiştir. Buna göre, öznitelik seçimi yapılarak tüm deneklerin verilerinin birlikte çalıştırıldığı sınıflandırıcılar her durumda daha kötü bir sınıflandırıcı başarımları vermektedir.

**Tablo 5.** EEG frekans bantlarına göre seçilen öznitelikler. ANOVA anlamlılık değerlerine ( $p \leq 0,05$ ) göre seçilen öznitelikler “✓” simgesi ile belirtilmiştir.(Selected features from EEG frequency bands. The selected features via significance values ( $p \leq 0,05$ ) obtained by ANOVA are indicated by “✓”)

Öznitelikler	EEG Frekans Bantları				
	Delta	Teta	Alfa	Beta	Gama
Haar_Enerji					
Haar_Varyans					
Haar_Entropi					
Db4_Enerji	✓	✓			
Db4_Varyans	✓	✓			
Db4_Entropi	✓	✓			
Sym4_Enerji	✓	✓	✓		
Sym4_Varyans	✓	✓	✓		
Sym4_Entropi	✓	✓	✓		
Coif1_Enerji		✓			
Coif1_Varyans	✓	✓			
Coif1_Entropi					
Rbio2.8_Enerji	✓	✓	✓		
Rbio2.8_Varyans	✓	✓	✓		
Rbio2.8_Entropi	✓	✓	✓		
Bior3.5_Enerji					
Bior3.5_Varyans					
Bior3.5_Entropi					

**Tablo 6.** Bütün özniteliklerin sınıflandırma performansları (Classification performance of all features)

Katılımcılar	Dalgacık	En yüksek başarımlar (%)	Sınıflandırıcılar
1	Haar	37,0	SVM (Cubic)
	Coif1	25,9	Ensemble (Bagged Trees)
	Db4	29,6	Tree (Fine, Medium, Coarse)
	Bior3.5	<b>55,6</b>	LDA
	Rbio2.8	22,2	Naive Bayes, SVM (Coarse Gaussian), KNN (Coarse)
	Sym4	25,9	KNN (Coarse), SVM (Coarse Gaussian), Ensemble (Boosted Trees)
2	Haar	23,1	SVM (Fine Gaussian), SVM (Coarse Gaussian), KNN (Coarse), Ensemble (Boosted Trees)
	Coif1	23,1	SVM (Coarse Gaussian), KNN (Coarse), Ensemble (Boosted Trees)
	Db4	23,1	SVM (Coarse Gaussian), KNN (Coarse), Ensemble (Boosted Trees)
	Bior3.5	<b>26,9</b>	Ensemble (Subspace KNN)
	Rbio2.8	23,1	SVM (Coarse Gaussian), KNN (Coarse), Ensemble (Boosted Trees)
	Sym4	23,1	SVM (Coarse Gaussian), KNN (Coarse), Ensemble (Boosted Trees)
3	Haar	<b>33,3</b>	Naive Bayes (Kernel), Ensemble (Subspace KNN)
	Coif1	23,8	SVM (Quadratic)
	Db4	<b>33,3</b>	Naive Bayes (Kernel)
	Bior3.5	19,0	KNN (Weighted), Ensemble (Subspace KNN)
	Rbio2.8	19,2	SVM (Coarse Gaussian), KNN (Coarse), Ensemble (Boosted Trees)
	Sym4	19,0	LDA
4	Haar	19,0	KNN (Cubic)
	Coif1	23,8	Naive Bayes (Kernel)
	Db4	<b>28,6</b>	Ensemble (RUSBoosted Trees)
	Bior3.5	<b>28,6</b>	Naive Bayes (Kernel)
	Rbio2.8	19,0	Ensemble (Subspace Discriminant)
	Sym4	<b>28,6</b>	Ensemble (Subspace Discriminant)

**Tablo 7.** Seçilen özneliklerin sınıflandırma performansları (Classification performance of selected features)

Katılımcılar	Dalgacık	En yüksek başarı (%)	Sınıflandırıcılar
1	Coif1	<b>25,9</b>	Ensemble (Boosted Trees)
	Db4	22,2	SVM (Coarse Gaussian), KNN (Coarse)
	Rbio2.8	22,2	SVM (Coarse Gaussian), KNN (Coarse), Ensemble (Boosted Trees)
	Sym4	<b>25,9</b>	KNN (Coarse), Ensemble (Boosted Trees)
2	Coif1	<b>23,1</b>	Ensemble (Boosted Trees)
	Db4	<b>23,1</b>	SVM (Coarse Gaussian), KNN (Coarse)
	Rbio2.8	<b>23,1</b>	SVM (Fine Gaussian), SVM (Coarse Gaussian), KNN (Coarse), Ensemble (Boosted Trees)
	Sym4	<b>23,1</b>	KNN (Coarse), Ensemble (Boosted Trees)
3	Coif1	9,5	Ensemble (RUSBoosted Trees)
	Db4	33,3	Tree (Fine, Medium, Coarse)
	Rbio2.8	28,6	Ensemble (Subspace KNN)
	Sym4	<b>38,1</b>	LDA
4	Coif1	19,0	Ensemble (RUSBoosted Trees)
	Db4	19,0	Naive Bayes (Gaussian)
	Rbio2.8	<b>38,1</b>	Naive Bayes (Gaussian)
	Sym4	23,8	LDA

**Tablo 8.** Öznelik seçimi öncesinde ve sonrasında (\*) sınıflandırıcıların ortalama başarıları (Average success of classifiers before and after feature selection (\*))

Sınıflandırıcılar	Dalgacıklar					
	Haar	Db4	Sym4	Coif1	Bior3.5	Rbio2.8
Karar Ağaçları	15,8	15,8	18,9	20,0	22,1	<b>26,3</b>
Ayrırcı Analizi	18,9	14,7*	14,7*	11,6*	15,8	<b>21,1*</b>
		10,5	13,7	18,9		<b>21,1</b>
Naive Bayes	15,8	8,4*	14,7*	13,7*	12,6	<b>16,8*</b>
		12,6	12,6	14,7		<b>16,8</b>
Destek Vektör Makinesi	20,0	15,8*	<b>16,8*</b>	15,8*	14,7	<b>16,8*</b>
		14,7*	<b>24,2*</b>	17,9*		<b>28,4</b>
En Yakın Komşuluk Sınıflandırıcısı	21,1	14,7	13,7	14,7	17,9	<b>24,2*</b>
		15,8*	15,8*	18,9*		<b>34,7</b>
Topluluk Sınıflandırıcıları	23,2	12,6	15,8	21,1	23,2	<b>29,5*</b>
		16,8*	16,8*	17,9*		<b>37,9</b>
						<b>33,7*</b>

#### 4. SONUÇLAR (CONCLUSIONS)

Bu çalışmada açık erişim yoluyla elde DDGUP verileri üzerinden ayrık dalgacık dönüşümü aracılığıyla enerji, varyans ve entropi gibi öznelikler hesaplanarak deneklere sunulan görsel uyarının frekansının beyin oksipital lobunda da aynı frekansla tespit edilip edilemediği çoklu sınıflandırma yöntemi ile değerlendirilmeye çalışılmıştır. Bu amaçla, hem altı farklı dalgacık fonksiyonu (Haar, Daubechies, Symlet, Coiflet, Biorthogonal, Reverse Biorthogonal) hem de varyans analizi (ANOVA) tabanlı öznelik seçiminin etkileri incelenmiştir. Altı temel sınıflandırma algoritmasının farklı parametreleri ile sekiz uyarım frekansı tespit edilmeye çalışılmıştır. Her denek için ayrı ayrı yapılan sınıflandırma işlemleri sonucunda öznelik seçimi yapılmadan tüm öznelikler kullanıldığında ortalama %36,1 (%26,9-55,6) (Tablo 6) ve sadece seçilen özneliklerin kullanılması durumunda ise %31,3 (%23,1-38,1) (Tablo 7) arasında doğrulukla uyarım frekansının

tahmin edilebildiği görülmüştür. Ayrıca tüm deneklerin verileri birlikte kullanıldığında ise öznelik seçimi yapılmadan tüm öznelikler kullanıldığında %37,9 (Tablo 8) ve sadece seçilen özneliklerin kullanılması durumunda ise %33,7 (Tablo 8) ortalama doğrulukla uyarım frekansının tahmin edilebildiği görülmüştür. Buna göre, deneklerin verileri ayrı ayrı değerlendirildiğinde herhangi bir ana dalgacık türünün ve sınıflandırıcı algoritmasının baskın olmadığı görülmüşüne rağmen deneklerin verileri birlikte kullanıldığında topluluk sınıflandırıcılarının ve Reverse Biorthogonal (Rbio2.8) ana dalgacığının diğer sınıflandırıcı ve ana dalgacık kombinasyonlarına göre bariz bir şekilde üstün olduğu anlaşılmaktadır. Bununla birlikte tek yönlü ANOVA gibi istatistiksel tabanlı bir öznelik seçiminin DDGUP ile BBA kontrolünde daha kötü bir sonuç verdiği görülmüştür. Bu nedenle literatürde mevcut olan diğer öznelik seçimi yöntemlerinden birisinin kullanılması daha yüksek sonuçlar verebilir. Literatürde mevcut olan diğer çalışmalar, farklı uyarım frekanslarında ekranda gösterilen

resimlerin (kutuların) hangisine odaklanıldığının bulunması üzerine yoğunlaşmışlardır. Bu nedenle yüksek sınıflandırıcı başarımları elde etmişlerdir. Sadece Sayilgan vd. [16] tarafından yapılan çalışmada aynı DDGUP veri seti [17] kullanılarak elde edilen zaman alanı öznelikleri ile uyaran frekansı tespit edilmeye çalışılmıştır. Bireysel değerlendirmede deneklerin DDGUP sinyallerinden uyartım frekansları %22,20-42,90 arasında doğrulukla ve tüm denekler birlikte değerlendirildiğinde ise %20,00 performans elde edilmiştir [16].

Bizim çalışmamızda da görsel uyaranda sunulan frekans bilgisinin tespitine çalışılmış ve literatürde sekiz uyartım frekansının aynı anda kullanıldığı tek çalışmadan çok daha iyi bir sonuç elde edilmiştir. Üstelik son dönemlerde umut vaat eden CCA algoritmasının kullanıldığı ve sadece dört frekansın kullanıldığı çalışmalarla da kıyaslanabilir bir performans elde edilmiştir. Elde edilen sonuçlara göre, sadece görsel uyarının gösterilip gösterilmediği değil, aynı zamanda hangi sıklıkla gösterildiği de tespit edilebilmektedir. Böylece mevcut çalışmalara ilave olarak ekranda aynı anda tek resmin gösterilmesi ve buna kişinin odaklanıp odaklanmadığı değil, birden fazla resmin farklı sıklıklarda aynı anda gösterilmesi ve bunlardan hangisine deneğin odaklandığının tespit edilmesinin yolu açılmaktadır. Ayrıca literatürdeki çoğu çalışmadan farklı olarak sekiz sınıflı bir veri setinin çoklu sınıflandırması gerçekleştirilmiştir. Dolayısıyla DDGUP sinyalleri üzerine yapılan çalışmalarda şimdiye kadar genellikle 2 veya 3 sınıfla yüksek sınıflandırma performansları elde edilirken bizim çalışmamızda 8 sınıfla elde edilen sonuçlar nispeten iyi olduğunu göstermektedir.

Yine de bu çalışmada elde edilen sınıflandırıcı başarımlarının tatmin edici düzeyde olmaması, verilerin elde edilmesi sırasında seçilen uyartım frekansları arasındaki frekans farklarının eşit olmamasından kaynaklanıyor olabilir. Bazı birbirini takip eden frekanslar arasında 0,5 Hz fark varken, bazıları arasında ise 2 Hz fark bulunmaktadır (6 - 6,5 - 7 - 7,5 - 8,2 - 9,3 - 10 - 12 Hz). Bu nedenle gelecek çalışmalarda, bu sonuçları daha da iyileştirmek adına kullanılan yöntemlerde parametre optimizasyonu hedef alınarak sonuçların iyileştirilmesi hedeflenmektedir.

#### TEŞEKKÜR (ACKNOWLEDGEMENT)

Yazarlar olarak, veri setini sağlayan ve veri seti ile ilgili her türlü sorumuzu cevaplayan Adnan Vilic'e teşekkür ederiz.

#### KAYNAKLAR (REFERENCES)

1. Wolpaw J.R., Birbaumer N., Heetderks W.J., McFarland D.J., Peckham P.H., Schalk G., Donchin E., Quatrano L.A., Robinson C.J., Vaughan T.M., Brain-computer interface technology: a review of the first international meeting, *IEEE Trans Rehabil Eng*, 8, 164-173, 2000.
2. Vaughan T.M., Heetderks W.J., Trejo L.J., Rymer W.Z., Weinrich M., Moore M.M., Kubler A., Dobkin B.H., Birbaumer N., Donchin E., Wolpaw E.W., Wolpaw J.R., Brain-computer interface technology: a review of the second international meeting, *IEEE Trans Neural Syst Rehabil Eng*, 11 (2), 94-109, 2003.
3. Gao X., Xu D., Cheng M., Gao S., A BCI-based environmental controller for the motion-disabled, *IEEE Trans Neural Syst Rehabil Eng*, 11 (2), 137-140, 2003.
4. Bashashati A., Fatourehchi M., A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals, *J Neural Eng*, 4, 32-57, 2007.
5. Wang Y.J., Wang R.P., Gao X.R., et al., A practical VEP-based brain-computer interface, *IEEE Trans Neural Syst Rehabil Eng*, 14, 234-239, 2006.
6. Bin G.Y., Gao X.R., Yan Z., Hong B., Gao S.K., An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method, *J Neural Eng*, 6, 046002, 2009.
7. Regan D., Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine, *J Clin Neurophysiol*, 7, 450-451, 1990.
8. Kahya E., Özkan N.F., Ulutaş B., Evaluation of brain computer interface usage in terms of cognitive load: A pilot study, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 34 (2), 647-662, 2019.
9. Zhang Z., Li X., Deng Z., A CWT-based SSVEP classification method for brain-computer interface system, *International Conference on Intelligent Control and Information Processing*, Dalian, China, 43-48, 13-15 August, 2010.
10. Muller-Putz G.R., Pfurtscheller G., Control of an electrical prosthesis with an SSVEP-based BCI, *IEEE Trans Biomed Eng*, 55 (1), 361-364, 2008.
11. Diez P.F., Mut V.A., Avila Perona E.M., Laciari Leber E., Asynchronous BCI control using high-frequency SSVEP, *Journal of neuroengineering and rehabilitation*, *J Neuroeng Rehabil*, 8, 39, 2011.
12. Heidari H., Einalou Z., SSVEP extraction applying wavelet transform and decision tree with Bayes classification, *International Clinical Neuroscience Journal (ICNSJ)*, 4 (3), 91-97, 2017.
13. Yeh C., Lee P., Chen W., Improvement of classification accuracy in a phase-tagged steady-state visual evoked potential-based brain computer interface using multiclass support vector machine, *BioMed Eng OnLine*, 12, 46, 2013.
14. Sayilgan E., Yuçe Y.K., Isler Y., Determining gaze information from steady-state visually-evoked potentials, *Karaelmas Fen ve Mühendislik Dergisi*, Kabul edildi, 2020.
15. Sayilgan E., Yuçe Y.K., Isler Y., Prediction of evoking frequency from steady-state visual evoked frequency, *Natural and Engineering Sciences*, 4 (3), 91-99, 2019.
16. Oojha M. K., Rajotiya R. N., Tyagi N., Frequency detection of single channel steady state visual evoked potential using canonical correlation analysis approach, *ICRAMSTEL-2019At: JIMS Engineering Management Technical Campus*, Greater Noida, 2019.

17. Vilic A., AVI Steady-State Visual Evoked Potential (SSVEP) Signals Dataset 2013, on line: <https://www.setzner.com/avi-ssvep-dataset/>, 2013.
18. Lotte F., Congedo M., Lecuyer A., Lamarche F. Arnaldi B., A review of classification algorithms for EEG-based brain-computer interfaces, *J Neural Eng*, 4 (2), R1-R13, 2007.
19. Iscan Z., Dokur Z., Demiralp T., Classification of electroencephalogram signals with combined time and frequency features. *Expert Systems with Applications*, 38 (8), 10499-10505, 2011.
20. Iscan Z., Development of classification methods for electroencephalogram based brain computer interfaces, Doktora Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul, 2012.
21. Mohamed E.A., Yusoff M.Z.K., Selman N.K., Malik A.S., Enhancing EEG signals in brain computer interface using wavelet transform, *International Journal of Information and Electronics Engineering*, 4 (3), 234-238, 2014.
22. Shiliang S.J.Z., A review of adaptive feature extraction and classification methods for EEG-based brain-computer interfaces, *IEEE Joint Conference (IJCNN) on Neural Networks International*, 1746-1753, 2014.
23. Rosso O.A., Blanco S., Yordanova J., Kolev V., Figliola A., Başar E, Wavelet entropy: a new tool for analysis of short duration brain electrical signals, *J Neurosci Methods*, 105, 65-75, 2001.
24. Tapan G., Bijaya P., Sneh A., A comparative study of wavelet families for EEG signal classification, *Neurocomputing*, 74, 3051-3057, 2011.
25. Isler Y., Kuntalp M., Combining classical HRV indices with wavelet entropy measures improves to performance in diagnosing congestive heart failure, *Comput Biol Med*, 37 (10), 1502-1510, 2007.
26. İşler Y., A detailed analysis of the effects of various combinations of heart rate variability indices in congestive heart failure, Doktora Tezi, Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü, İzmir, 2009.
27. Duda R. O., Hart P. E., Stork D. G., *Pattern Classification*, 2nd edition, Wiley Interscience, 2001.
28. Alpaydin E., *Introduction to Machine Learning*, MIT Press, 2004.
29. Liuand H., Motoda H., *Feature Selection for Knowledge Discovery and Data Mining*, Kluwer Academic Publishers, London, GB, 1998.
30. IBM SPSS Statistics. <https://www.ibm.com/tr-tr/products/spss-statistics>.
31. Lotte F., Baugrain L., Cichocki A., Clerc M., Congedo M., Rakotomamonjy A., Yger F., A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update, *J Neural Eng*, 15 (031005), 2018.
32. *MATLAB and Machine Learning and Deep Learning Toolbox Release 2019a*, The MathWorks, Inc., Natick, Massachusetts, United States.
33. Jung Y., Hu J., A k-fold averaging cross-validation procedure, *J Nonparametr Stat*, 27, 1-13, 2015.
34. Narin A., İşler Y., Özer M., Konjestif kalp yetmezliği teşhisinde kullanılan çapraz doğrulama yöntemlerinin sınıflandırıcı performanslarının belirlenmesine olan etkilerinin karşılaştırılması, *DEÜ Mühendislik Fakültesi Mühendislik Bilimleri Dergisi*, 16 (48), 1-8, 2014.
35. Isler Y., Narin A., Ozer M., Comparison of the effects of cross-validation methods on determining performances of classifiers used in diagnosing congestive heart failure, *Meas Sci Rev*, 15 (4), 196-201, 2015.

