






Sürücü yorgunluk evrelerinin EEG işaretleri ile tespiti ve analizi Detection and analysis of driver fatigue stages with EEG signals

Ahmet DEMİR¹ , Şule BEKİRYAZICI^{1*} , Oğuzhan ÇOŞKUN¹ , Recep EKEN² , Güneş YILMAZ¹ 

¹Elektrik-Elektronik Mühendisliği Bölümü, Mühendislik Fakültesi, Uludağ Üniversitesi, Bursa, Türkiye.
ahmetdemir@uludag.edu.tr, suleyapici@uludag.edu.tr, oguzhancoskun1608@gmail.com, gunesy@uludag.edu.tr

²Daiichi AR-GE Sanayi ve Ticaret Anonim Şirketi, Bursa, Türkiye.
ekenrecep12@gmail.com

Geliş Tarihi/Received: 22.10.2020
Kabul Tarihi/Accepted: 08.03.2022

Düzeltilme Tarihi/Revision: 04.03.2022

doi: 10.5505/pajes.2022.89327
Araştırma Makalesi/Research Article

Öz

Günümüzde birçok insan trafik kazalarında hayatlarını kaybetmektedir. Trafik kazalarının en önemli nedeni olarak sürücülerin uykusuzluğu ve yorgunluğu gösterilmektedir. Bu nedenle sürücü performansının analizi konusunda yapılan araştırmalar büyük önem taşımaktadır. Bu çalışmada, elektroensefalografi (EEG) verileri kullanılarak sürücü yorgunluğunu analiz edebilecek bir sistem tasarlanmıştır. Veri seti olarak National Chiao Tung Üniversitesi tarafından hazırlanan sürekli dikkat halinde sürüş deneyinden alınan EEG işaretleri kullanılmıştır. Boyutu sürücünün yorgunluk zamanlarını ve seviyesini tespit edebilmek için veri seti dört sınıfa ayrılmıştır. Sürücü yorgunluğu evrelerinde ortaya çıkan frekans aralıklarını belirleyebilmek için EEG sinyallerine filtreleme işlemleri yapılmıştır. Temel Bileşen Analizi yöntemi kullanılarak öznelik matrisinin azaltılmıştır. Böl ve Fethet algoritması ile dört sınıfın birbirinden en iyi şekilde ayrılacağı tüm kombinasyonlar belirlenmiş ve her adımda alt-sınıflandırıcılar kullanılarak sınıflandırma işlemi yapılmıştır. Alt sınıflandırıcı olarak, k-En Yakın Komşuluk, Destek Vektör Makineleri ve Doğrusal Ayırım Analizi algoritmaları kullanılmıştır. Çalışma sonucunda ortalama sınıflandırma başarıları k-En Yakın Komşuluk algoritması için %87.9, Destek Vektör Makineleri algoritması için %88.5 ve Doğrusal Ayırım Analizi için %81.6 olarak elde edilmiştir. En yüksek ortalama sınıflandırma başarısı ise 4. sınıf yorgunluk seviyesinde, sürüşün 67.5-90 dk. arasında Destek Vektör Makineleri sınıflandırıcısı ile %93.2 olarak elde edilmiştir.

Anahtar kelimeler: Sürücü yorgunluğu, Elektroensefalografi, Temel bileşen analizi, Sınıflandırma, Böl ve fethet algoritması.

Abstract

Today, many people die in traffic accidents. Sleeplessness and fatigue of drivers are shown as the most important cause of traffic accidents. For this reason, research on driver performance analysis is of great importance. In this study, a system is designed to analyze driver fatigue using electroencephalography (EEG) data. As the data set, the EEG signals from sustained-attention driving task prepared by National Chiao Tung University have been used. The data set is divided into four classes to determine the driver's fatigue times and level. In order to determine the frequency ranges that occur during driver fatigue phases, EEG signals are filtered. Principal Component Analysis method has been used to reduce the size of the features matrix. With the Divide and Conquer algorithm, all combinations in which the four classes will be separated best are determined and classification has been done at each step using sub-classifiers. As sub-classifiers, k-Nearest Neighborhood, Support Vector Machines and Linear Discrimination Analysis algorithms are used. As a result of the study, the average classification successes are 87.9% for the k-Nearest Neighborhood algorithm, 88.5% for the Support Vector Machines and 81.6% for Linear Discrimination Analysis. The highest classification success has been achieved as 93.2% with the Support Vector Machines classifier, between 67.5-90 min. of driving at the 4th grade fatigue level.

Keywords: Driver fatigue, Electroencephalography, Principal component analysis, Classification, Divide and conquer algorithm.

1 Giriş

Günümüzde trafik kazaları insan hayatı için büyük bir tehdit unsuru oluşturmaktadır. Dünya Sağlık Örgütü'ne göre, 18 ile 30 yaşları arasındaki yetişkinlerde birincil ölüm nedeni ve küresel olarak insan ölümünün dokuzuncu nedeni, trafik kazaları ile temsil edilmektedir. Ayrıca istatistiksel veriler ve araştırmalar, sürücü yorgunluğunun trafik kazalarının başlıca nedenlerinden biri olduğunu göstermektedir [1].

Sürücü yorgunluğu, nörobiyolojik süreçlerden kaynaklı uyuşukluğa ve uykuluğu neden olmaktadır [1]. Sürücülerin yorgunluğu, sürücülerin uykulu oldukları sırada algılama, tanıma ve araç kontrolünde belirgin düşüş olması nedeniyle birçok kazada nedensel bir faktör olarak gösterilmiştir. Yorgunluğun etkileri altında araba kullanmak; dikkatin azalması, uyanık kalmada güçlük, karar vermede gecikme gibi

etkilere neden olmaktadır. Uyku eksikliği, uzun sürüş süreleri, sakinleştirici ilaç kullanımı, alkol tüketimi ve özellikle monoton bir sürüş gibi bazı sürüş düzenleri uyuşukluğa veya yorgunluğa neden olabilmektedir [2].

Sürücü yorgunluğu, video kamera ile göz ve baş takibi, elektrookülogram (EOG) süresi, elektrokardiyogram (EKG) ile kalp ritmi ölçümü gibi çeşitli yöntemlerle analiz edilmektedir. Yorgunluk tespiti için mevcut yöntemler gözler, ağız, kafa hareketleri ve diğer vücut bilgilerine dayanmaktadır. Bu da zayıf uygulanabilirlik ve düşük tanıma doğruluğu ile sonuçlanmaktadır. Bu nedenle en etkili yöntemlerden biri beyin elektriksel aktivitesini ölçen Elektroensefalografi (EEG) görüntüleme yöntemidir. EEG, insan beynindeki nörofizyolojik aktiviteyi doğrudan ölçtüğü için yorgunluğu tespit etmede önemli ve güvenilir bir yöntem olarak kabul edilmektedir [1].

*Yazışılan yazar/Corresponding author

İnsan beyni, bilgilerin taşınması için milyonlarca nörondan oluşmaktadır. Bu nöronlar beyin ile vücut arasında iletişimi elektriksel uyarımlarla sağlayan bilgi taşıyıcılarıdır. EEG, kafa derisinde farklı noktalara yerleştirilen elektrotlar yardımıyla bu elektriksel aktiviteyi ölçerek değerlendiren, nörolojik aktivite bilgisini içeren ve düşük genlikli sinyallerin toplanmasını sağlayan yöntemdir [3].

Uygulanması kolay ve anlık bilgileri içerdiğinden EEG ölçümleri Beyin Bilgisayar Arayüzü uygulamalarında en çok kullanılan tekniktir. EEG sinyallerinin işlenmesi, duygu durum tespitinde, uyku durum sınıflandırmalarında, motor görüntüleri sınıflandırmalarında ve nöbet tespiti/tahmini gibi birçok alanda uygulanmaktadır [3].

EEG sinyallerinin işlenmesinde beynin fonksiyonel durumu hakkında bilgiler farklı frekans bandlarında ayırt edilmektedir. Bu frekans bandları, Delta (0.5-4 Hz), Teta (4-8 Hz), Alfa (8-12 Hz), Beta (12-30 Hz) ve Gama (30-50 Hz) frekans aralıklarıdır. Delta beyin dalgaları, derin uyku halinde; teta beyin dalgaları, vücut uykuda/zihin uyanık durumda; alfa beyin dalgaları, uyanık, sakin ve rahatlama halinde; beta beyin dalgaları, uyanık, odaklanma ve dikkat halinde; gama beyin dalgaları, beynin karar verme gibi aşırı zihinsel aktivitelerinde ortaya çıkmaktadır [4].

Yüksek frekanslar genellikle uyarılma, uyanıklık ve odaklanma dikkat ile birlikte görülmektedir. Bu nedenle, uyanık ve dikkat halinde bir kişinin EEG sinyali yüksek frekanslarda salınımlar içerebilirken; aynı kişinin derin gevşeme ve uyusukluk sırasındaki EEG sinyali düşük frekans, yüksek genlik aktivitesi göstermektedir [5]. Laboratuvar tabanlı birçok simülasyon kurulumunda, uyanıklık ve uyarı durumundan uykuluğa geçiş sırasında delta, teta ve alfa aktivitelerinde tutarlı artışlar gözlemlenmiştir [6].

Literatür incelendiğinde, EEG tabanlı sürücü yorgunluğu ve duygu durum tespitinde farklı sınıflandırma yöntemleri kullanılarak yapılan birçok çalışma vardır. Yapılan bir çalışmada, derin kovaryans öğrenme yöntemleri kullanılarak EEG sinyallerinden sürücü yorgunluk tahmini yapılmıştır. Çalışmada üç tip derin kovaryans öğrenme yöntemi (SPDNet, CNN ve DNN) sunulmuştur. Çalışma sonucunda en yüksek ortalama sınıflandırma başarısı CNN (Evrişimsel Sinir Ağları) yöntemiyle %86 olarak elde edilmiştir [7].

Bir diğer çalışmada ise açık kaynak bir veri seti kullanarak farklı sınıflandırma yöntemleriyle sürücü yorgunluğu tahmini yapılmıştır. Çalışma sonucunda en yüksek başarı oranı DVM (Destek Vektör Makinesi) sınıflandırıcısı ile %83 bulunmuştur [8].

Literatürde yapılan bir diğer çalışmada, yapay sinir ağı mimarisi kullanılarak sürücü yorgunluğunu tespit edebilecek bir sistem tasarlanmıştır. 20 profesyonel ve 35 profesyonel olmayan kamyon sürücüsünden alınan EEG sinyalleri, alt frekans bandlarına (Delta, teta, alfa, beta, gama) ayrılarak iki sınıflı (yorgun-yorgun değil) sınıflandırma işlemi gerçekleştirilmiştir. Çalışma sonucunda profesyonel sürücülere ait ortalama sınıflandırma başarısı %81.49, profesyonel olmayan sürücülere ait ortalama sınıflandırma başarısı %83.06 olarak elde edilmiştir [9].

Bashar ve Bhuiyan'ın yaptığı bir çalışmada, kullanılan veri seti, farklı sınıflandırıcılar ile sınıflandırma yapılmış ve performansları karşılaştırılmıştır. Buna göre sınıflandırma başarıları; DVM algoritması için %86.43, kNN algoritması için %85.71 ve DAA (LDA) algoritması için %85.71 olarak elde

edilmiştir. Farklı öznelik çıkarma ve boyut azaltma yöntemleri ve farklı sınıflandırıcılar ile sınıflandırma başarısı mevcut çalışmalar arasında değişkenlik gösterebilmektedir [10].

Bu çalışmada ise National Chiao Tung Üniversitesi tarafından hazırlanmış olan sürücü performansı açık veri seti kullanılarak sürücü yorgunluk seviyesi tespiti yapılmıştır. EEG sinyalleri dört farklı sınıfa ayrılarak, farklı yorgunluk seviyeleri oluşturulmuştur. Bu sınıflar Tablo 1'de detaylandırılmıştır. EEG sinyalleri band geçiren sonlu darbe cevaplı filtreden geçirilerek alt frekans bandlarına ayrılmıştır.

Elde edilen sinyalleri en iyi şekilde temsil edecek frekans tabanlı öznelikler çıkarılmıştır. Öznelik matrisinin boyutunu azaltmak için Temel Bileşen Analizi yöntemi kullanılmıştır. Boyutu azaltılmış öznelik matrisi, Böl ve Fethet algoritması (alt sınıflandırıcılar: k-en yakın komşuluk algoritması, destek vektör makineleri, doğrusal ayırım analizi) kullanılarak sınıflandırılmıştır.

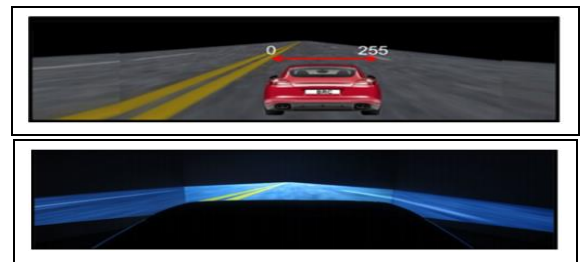
2 Materyal ve metot

2.1 Kullanılan veri seti

Bu çalışmada veri kümesi olarak, National Chiao Tung Üniversitesi tarafından hazırlanmış olan sürücü performansı EEG kayıtları kullanılmıştır. Kayıtlar üniversite öğrencileri ve görevlilerinden oluşan 27 gönüllü (yaş aralığı: 22-27 yaş) katılımcı ile gerçekleştirilmiştir. Deneklerin tümünün görme yetisinde kayıp, uyku düzeninde bozukluk veya yoksunluk ve uyuşturucu bağımlılığı gibi problemlerin olmadığı deney öncesinde raporlanmıştır. Ayrıca tüm katılımcıların ehliyetinin olması zorunlu hale getirilmiştir.

Deneyden önceki bir hafta boyunca katılımcılardan, gece en geç 23:00'da uykuda olmaları ve düzenli olarak yaklaşık sekiz saat uyumaları istenmiştir. Buna ek olarak katılımcılardan deneyden bir gün önce alkol veya kafeyli içecekler tüketmemeleri ve yoğun egzersizlere katılmamaları istenmiştir. Tüm katılımcılar, deneylere katılmadan önce bilgilendirilmiş ve çalışmada gönüllü olarak EEG kaydı vereceğine dair onaylı bir form imzalamışlardır [11].

Deneyler, dinamik bir sürüş simülasyonuna sahip sanal gerçeklik sürüş ortamında gerçekleştirilmiştir. Gerçek dünyadaki sürüş durumlarında ortaya çıkan somatik hislerin beyinde üretilebilmesi için hareket platformuna sahip VR tabanlı bir sürüş simülasyonu kullanılmıştır. Sürüş simülasyonu altı serbestlik derecesine sahip Stewart hareket platformu üzerine monte edilmiştir. Deneklerin sürüş sahnesini takip edebilmesi için otoyol; 0°, 42°, 84°, 180°, 276° ve 318° görüş açılarında ekranlara yansıtılmıştır. Şekil 1'de ilgili deney platformu gösterilmiştir [11].



Şekil 1. Sanal gerçeklik dinamik sürüş simülasyonu deney düzeneği.

Figure 1. Virtual reality dynamic driving simulator experiment setup.

Deneyden önce katılımcıların talimatları anlaması ve simülasyon ortamını tanıması amacıyla bir ön test oturumu gerçekleştirilmiştir. Deneyler trafik olmayan, dört şerite bölünmüş monoton bir yolda, gece sürüşünü içermektedir. Katılımcılar bu otoyolda sürekli dikkat halinde doksan dakika boyunca sürüş görevini gerçekleştirmiştir. Görev boyunca deneklerden tek bir şeridin merkezinde hareket etmesi istenmiş ve EEG kayıtları alınmıştır. Ayrıca şerit, 0-255 arasında değerlere ayrılarak aracın şerit sapması da sürekli olarak kaydedilmiştir. Senaryo çerçevesinde, aracın hızı ortalama 100 km/s şeklinde ayarlanmıştır [11].

Deneyde kaydedilen EEG işaretleri Compumedics firmasına ait Scan SynAmps2 Express kayıt cihazı kullanılarak toplanmıştır. Otuz adet Ag/AgCl EEG elektrotu ve iki referans elektrot kullanılarak otuz iki kanallı EEG kayıtları oluşturulmuştur. Elektrotlar uluslararası 10-20 elektrot yerleşim sistemine göre deneklerin kafa derisine yerleştirilmiştir. Tüm elektrotlar ve cilt arasındaki yüzey empedansı 5 k Ω ' un altında tutulmuştur. Her katılımcıdan birkaç defa olmak üzere toplamda 62 EEG verisi toplanmıştır [11].

Bu çalışmada veri seti içerisinde bulunan 62 EEG verisinden, boyut farklılığından dolayı 52 tanesi kullanılmıştır. Farklı yorgunluk seviyelerinin belirlenebilmesi için, sürüş esnasında toplanan 52 EEG verisi dört farklı sınıfa ayrılmıştır. Böylece 90 dakikalık sürüş verileri 22.5 dk.'lık bölümlere ayrılmıştır. Veri setinin detayları Tablo 1'de verilmiştir.

Tablo 1. Kullanılan veri seti.

Table 1. Data set used.

Yorgunluk Seviyesi	Süre/dk.	Sınıf Bilgisi
Kişi uyanıktır.	0-22.5	1 No.lu sınıf
Kişi ne uyanık, ne de uykulu bir durumdadır. Ara değerdir.	22.5-45	2 No.lu sınıf
Kişide artık bazı uyuşukluk belirtileri gözlemlenmektedir.	45-67.5	3 No.lu sınıf
Kişi uykulu-uyanık kalmak için bazı çabalar göstermektedir.	67.5-90	4 No.lu sınıf

2.2 Filtreleme yöntemi

EEG sinyalleri mikro voltlar mertebesinde çok küçük genliklere ve farklı frekans bileşenlerine sahip elektriksel işaretlerdir. Frekans aralığı 0.5 Hz-50 Hz arasında olan EEG sinyallerini farklı frekans bantlarında incelemek ve sürücü yorgunluğu evrelerinde ortaya çıkan frekans aralıklarını belirleyebilmek için filtreleme işlemleri yapılmaktadır. Filtreleme sonucunda; Delta (0.5-4 Hz), Teta (4-8 Hz), Alfa (9-13 Hz), Beta (14-31 Hz) ve Gama (32-50 Hz) frekans bantları oluşmaktadır [12].

Çalışma kapsamında EEG sinyalleri band geçiren sonlu darbe cevaplı filtreden geçirilmiştir. Filtre kesim frekansları her bir alt bantı filtrelemek için sırasıyla 0.5-4 Hz, 4-8 Hz, 8-13 Hz, 14-31 Hz ve 32-50 Hz olarak seçilmiştir.

2.3 Öznitelik çıkarma yöntemleri

Bu çalışmada, EEG işaretleri içerisinde var olan bilgiyi en iyi şekilde ifade edebilecek yedi farklı öznitelik kullanılmıştır. Belirtilen öznitelikler EEG sinyalinin tüm frekans bantlarına ve her bir kanalına ayrı ayrı uygulanmıştır. Uygulanan öznitelikler Tablo 2'de gösterilmiştir.

Dalgacık dönüşümü ve Fourier Dönüşümü gibi sinyal analiz işlemleri, EEG işaretlerini temsil etmek için kullanılan etkili

yöntemlerden biridir. Bu dönüşümler ile işaretin içerdiği farklı frekans bileşenleri ve bu frekanslara ait genlik büyüklükleri elde edilmektedir.

Tablo 2. EEG işaretlerine uygulanan öznitelikler.

Table 2. Attributes applied to EEG signals.

Uygulanan Öznitelikler	
1	Fourier dönüşümü alınmış işaretin standart sapması
2	Fourier dönüşümü alınmış işaretin ortalaması-normu
3	Fourier dönüşümü alınmış işaretin varyansı-normu
4	Dalgacık dönüşümü alınmış işaretin ortalaması
5	Dalgacık dönüşümü alınmış işaretin standart sapması
6	Dalgacık dönüşümü alınmış işaretin basıklık katsayısı
7	Dalgacık dönüşümü alınmış işaretin çarpıklık katsayısı

Fourier Dönüşümü, zaman domeninde olan işareti frekans domenine çeviren ve işaret içerisindeki tüm frekans bilgisini ortaya çıkaran matematiksel bir araçtır. Fourier Dönüşümünün matematiksel ifadesi ise Denklem (1)'de verilmiştir.

$$X(w) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt \quad (1)$$

Bu denklemde; $x(t)$ sinyalin zamana bağlı ifadesini ve $X(w)$ Fourier Dönüşümü alınmış işareti temsil etmektedir. Fourier Dönüşümü alınmış olan işareti standart sapması, ortalamasının normu ve varyansının normu alınarak öznitelik olarak kullanılmıştır. Burada, ortalama Denklem (2), varyans Denklem (3) ve Standart sapma da Denklem (4) kullanılarak hesaplanmıştır. Bu denklemlerde x_i i. örnek değerini temsil etmektedir.

$$\bar{x} = \frac{\sum_{i=1}^N x_i}{N} \quad (2)$$

$$\sigma^2 = \frac{\sum_{i=1}^N (\bar{x} - x_i)^2}{N} \quad (3)$$

$$\sigma = \sqrt{\sigma^2} \quad (4)$$

Dalgacık Dönüşümü, düşük frekanslar için geniş, yüksek frekanslar için dar olacak şekilde değişen pencere boyutları ile odağı ayarlanabilen matematiksel araçtır. Bu yöntem işaret içerisindeki tüm frekans aralıklarında optimum zaman-frekans çözünürlüğü sağlayabilmektedir. Bu sayede analiz edilen işaret hem frekans bölgesinde hem de zaman bölgesinde incelenebilmektedir.

Dalgacık Dönüşümü genel olarak daraltılıp, genişletilebilen bir dalgacık fonksiyonu $\psi(t)$ ile analiz edilecek bir $x(t)$ işaretinin evrişimi olarak tanımlanmaktadır. Dalgacık dönüşümünün matematiksel ifadesi Denklem (5)'te verilmektedir.

$$DWT[n] = \frac{1}{\sqrt{a^j}} * \varphi\left(\frac{n}{a^j}\right) \quad (5)$$

Bu denklemde; a ölçekleme parametresini, b dönüşüm parametresini, φ ana dalgacık fonksiyonunu ve $CWT(a, b)$ işaretin sürekli dalgacık dönüşümünü temsil etmektedir. Çalışmada ana dalgacık fonksiyonu olarak Morlet fonksiyonu kullanılmıştır. Dalgacık dönüşümü yapılmış olan sinyalin ortalaması, standart sapması, basıklık ve çarpıklık katsayısı hesaplanarak öznitelik olarak kullanılmıştır. Ortalama ve standart sapma için Denklem (2) ve Denklem (4) kullanılmıştır.

Basıklık katsayısı Denklem (6), çarpıklık katsayısı ise Denklem (7) ile hesaplanmıştır.

$$BK = \frac{1}{N} \sum_{i=1}^N \left[\frac{x_i - \bar{x}}{\sigma} \right]^3 \quad (6)$$

$$\zeta K = \frac{1}{N} \sum_{i=1}^N \left[\frac{x_i - \bar{x}}{\sigma} \right]^4 \quad (7)$$

Belirtilen öznelikler, EEG sinyalinin her bir kanalına ayrı ayrı uygulanmıştır. Bu işlem sonucunda 52x224 (7 öznelik x 32 kanal) boyutlu öznelik matrisi oluşmuştur. EEG sinyali içerisinde bulunan her bir frekans bandı (Delta, Teta, Alfa, Beta, Gama) için bu işlem tekrarlanmıştır.

2.4 Temel bileşen analizi

Temel Bileşen Analizi, aralarında ilinti bulunan çok değişkenli verileri, aralarında ilinti bulunmayan yeni bir koordinat sistemine dönüştüren istatistiksel bir veri dönüşüm yöntemidir. Temel Bileşen Analizinin temel ilkesi, multispektral vektör uzayında, verileri korelasyonsuz olarak ifade edebilen diğer bir deyişle yeni sistemde kovaryans matrisi köşegen olan, yeni bir koordinat sisteminin araştırılmasıdır. Ayrıca Temel Bileşen Analizi yüksek boyutlu verileri daha düşük boyutlu verilere indirgemeye yarayan güçlü bir veri analiz tekniğidir [13].

Fourier ve Dalgacık Dönüşümü kullanılarak elde edilen özneliklerin birleştirilmesi sonucu öznelik matrisi oldukça büyük boyutlara ulaşmaktadır. Yüksek boyutlu veriler sınıflandırıcıların işlem yükünü arttırdıklarından, sistem performansını olumsuz yönde etkilemektedir. Bu nedenle çalışmada, öznelik matrisinin boyutunu azaltmak için Temel Bileşen Analizi algoritması kullanılmıştır. Algoritmada kullanılan objektif fonksiyon Denklem (8) ile ifade edilebilir.

$$\text{Objektif Fonksiyonu} = \text{argmax} (w^t C w) \quad (8)$$

$$C = \frac{1}{N} \sum_{i=1}^N x_i x_i^T \quad (9)$$

Bu denklemde; C veriye ait varyans-kovaryans matrisini, w varyans-kovaryans matrisinin özvektörlerini ve x orjinal veriyi temsil etmektedir. C ifadesi Denklem (9) kullanılarak elde edilmektedir.

2.5 Sınıflandırma yöntemleri

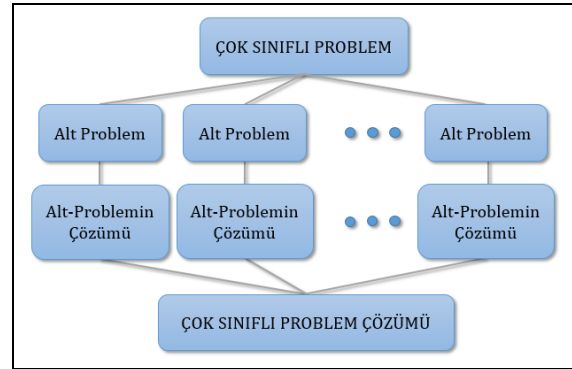
Sınıflandırma, ham EEG işaretlerinden elde edilen öznelik vektöründeki verilerin belirlenen bir sınıfa atanması işlemidir. Sınıflandırma algoritmalarının kendi içlerinde güçlü ve zayıf yönleri bulunmaktadır. Bir sınıflandırma algoritması belirli bir veri seti üzerinde başarılı olurken farklı veriler üzerinde başarı oranı azalabilmektedir. Bu nedenle farklı sınıflandırma algoritmaları denenerek, kullanılan veri için optimum sonucun elde edildiği algoritma belirlenmektedir [14]. Bu çalışmada Mathworks tarafından geliştirilen MATLAB programı kullanılarak, sınıflandırıcı algoritmalarından biri olan Böl ve Fethet yöntemi kullanılmıştır.

Böl ve Fethet algoritması karar ağacı yapısına benzer bir yaklaşımdır. Bu yöntemde karar ağacı yapısını oluşturmak için veri seti alt kümelere ayrılmaktadır. Böl ve Fethet algoritmasının mantığı, sınıfları her bir alt sette sadece tek bir

sınıf kalıncaya kadar alt kümelere ayırmaktır. Bu yöntemde sınıf sayısına göre alt küme sayısı değişmektedir.

Karar ağacı yapısı, veri seti içerisindeki tüm olası sınıf kombinasyonlarının oluşturulması ile elde edilmektedir. Bu kombinasyonlar bire-karşı-hepsi ve bire-karşı-biri şeklinde alt kümelere ayrılmaktadır. Sınıflandırma başarısı en yüksek olan grup karar ağacı yapısından ayrılırken, kalan gruplar ile aynı şekilde sınıflandırma işlemine devam edilmektedir [15].

Bu çalışmada, Böl ve Fethet algoritması içerisinde alt sınıflandırıcı olarak, literatürde sıklıkla kullanılan k-En Yakın Komşuluk Algoritması, Destek Vektör Makineleri ve Doğrusal Ayırım Algoritması kullanılmıştır. Böl ve Fethet algoritmasına ait temel yapı Şekil 2'deki gibidir.



Şekil 2. Böl ve Fethet algoritma yapısı.

Figure 2. Divide and Conquer algorithm structure.

k-En Yakın Komşuluk (k-EYK) Algoritmasında, test verilerinin seçilen k sayısı kadar eğitim verisine uzaklıkları belirli mesafe ölçüm yöntemleri kullanılarak hesaplanır ve test verisine en uygun sınıf ataması yapılmaktadır. Bu algoritmanın performansına etkiyen en önemli iki parametre mevcuttur. Bunlardan birincisi, test örneğine en yakın eğitim örneğinin bulunmasında kullanılan uzaklık hesaplama yöntemidir. İkincisi ise belirlenecek komşu sayısı (k) değeridir [16]. Bu çalışmada mesafe ölçümü için öklit uzaklığı yöntemi kullanılmıştır.

Destek Vektör Makineleri (DVM) Algoritmasında, çekirdek fonksiyonları kullanılarak farklı sınıfları birbirinden en iyi şekilde ayıracak hiper düzlemler oluşturulmaktadır. Özellikle iki sınıflı sınıflandırma problemlerinde, DVM iki sınıf arasındaki optimal ayırt etme yüzeyini belirlemektedir. DVM yöntemiyle sınıflandırma problemi karesel optimizasyon problemine dönüştürülerek çözülür. Böylece eğitim aşamasında işlem sayısı azalmakta ve sonuca daha kısa sürede ulaşılmaktadır [14].

Doğrusal Ayırım Analizi (DAA) Algoritması ise, önceden belirlenmiş iki veya daha fazla sınıfın ortalama niteliklerinin istatistiksel olarak anlamlı bir farklılık gösterip göstermediğini ve grupların farklılıklarını ayırt etmede en fazla katkıyı hangi değişkenlerin yaptığını test etmekte kullanılan bir sınıflandırma tekniğidir. Bu yöntemin çalışma mantığı, mevcut verilerin bileşkesi olan yeni bir değişken üretmesi şeklindedir [17].

2.6 Sınıflandırma başarımlar ölçütleri

Sınıflandırma işlemlerinden sonra, sisteme ait başarımlar sonuçlarının verilmesi için literatürde sıklıkla; toplam sınıflandırma başarısı, duyarlılık, seçicilik, kappa katsayısı, ayırma gücü, hata matrisi, ROC eğrisi gibi birçok farklı ölçüt

kullanılmaktadır. Bunlardan sınıflandırma başarısı ve hata matrisi çalışmalarında en çok tercih edilen ölçütlerdir. Özellikle çok sınıflı problemlerde, hata matrisi sistem performansını en iyi şekilde ifade edebilmektedir.

Bu çalışmada, sınıflandırma işlemi dört sınıflı bir problem olduğundan sınıflandırma ölçütü olarak toplam sınıflandırma başarısı ve hata matrisi değerleri kullanılmıştır. Dört sınıflı bir problemde, sınıflandırıcı algoritması sonuçlarında on altı farklı olası durum meydana gelmektedir. Şekil 3'te bu olası durumlar verilmektedir.

		Gerçek Sınıflar			
		G_1	G_2	G_3	G_4
Tahmini Sınıflar	T_1	TG_{11}	TG_{12}	TG_{13}	TG_{14}
	T_2	TG_{21}	TG_{22}	TG_{23}	TG_{24}
	T_3	TG_{31}	TG_{32}	TG_{33}	TG_{34}
	T_4	TG_{41}	TG_{42}	TG_{43}	TG_{44}

Şekil 3. Dört sınıflı problem için hata matrisi.

Figure 3. Error matrix for four classes of problems.

Tabloda verilen T_1, T_2, T_3 ve T_4 ifadeleri sınıflandırıcı çıkışından elde edilen tahmini sınıf bilgileridir. G_1, G_2, G_3 ve G_4 ifadeleri ise verilere ait gerçek sınıf bilgisidir. Hata matrisi içerisinde bu sınıfların kesişmesi ile $TG_{11}, TG_{12}, TG_{13}, TG_{14}, TG_{21}, \dots, TG_{43}$ ve TG_{44} ile gösterilen on altı durum oluşmaktadır. Bu kesişim verilerinin bazıları aşağıdaki gibi açıklanmaktadır.

- ✚ TG_{11} : Gerçekte S_1 sınıfına ait olan ve sınıflandırıcının T_1 sınıfına atadığı veri kümesi,
- ✚ TG_{12} : Gerçekte S_2 sınıfına ait olan ve sınıflandırıcının T_1 sınıfına atadığı veri kümesi,
- ✚ TG_{22} : Gerçekte S_2 sınıfına ait olan ve sınıflandırıcının T_2 , sınıfı atadığı veri kümesi,
- ✚ TG_{34} : Gerçekte S_4 sınıfına ait olan ve sınıflandırıcının T_3 sınıfına atadığı veri kümesi,
- ✚ TG_{42} : Gerçekte S_2 sınıfına ait olan ve sınıflandırıcının T_4 sınıfına atadığı veri kümesi,
- ✚ TG_{44} : Gerçekte S_4 sınıfına ait olan ve sınıflandırıcının T_4 sınıfına atadığı veri kümesi.

Hata marisinde i. Örnek için köşegen değeri Denklem (10) ile hesaplanmaktadır.

$$\text{Hassasiyet(precision)} = \frac{TG_{ii}}{TG_{i1} + TG_{i2} + TG_{i3} + TG_{i4}} \quad (1)$$

Tahmin edilen tüm sınıfların, gerçek sınıflar ile aynı olması durumunda ($T_1 = G_1, T_2 = G_2, T_3 = G_3, T_4 = G_4$), sınıflandırma işlemi %100 doğruluk ile gerçekleştirilmektedir. Hata matrisinden elde edilen bulgular kullanılarak toplam sınıflandırma başarısı Denklem (11) ile elde edilmektedir.

$$\text{TSB} = \frac{TG_{11} + TG_{22} + TG_{33} + TG_{44}}{TG_{11} + TG_{12} + \dots + TG_{42} + TG_{43} + TG_{44}} \quad (2)$$

Bu çalışmada, "Bulgular" sınıflandırıcı sonuç tablolarında TSB, $\%TG_{11}, \%TG_{22}, \%TG_{33}$ ve TG_{44} değerleri verilmektedir.

3 Bulgular

Bu çalışmada ilk olarak EEG dalgalarına ait her bir frekans bandı (Delta, teta, alfa, beta ve gama bandı) öznitelikleri üç farklı sınıflandırıcı algoritması ile sınıflandırılmıştır. Her bir frekans bandındaki öznitelik matrislerinin boyutu 52×224 olarak elde edilmiştir. Dört sınıfa ait öznitelik matrislerin birleştirilmesi sonucu öznitelik matrisi boyutu 208×224 olarak elde edilmiştir. Tüm sınıflandırıcı işlemlerinde 10-kat çapraz doğrulama yöntemi kullanılmıştır.

Tablo 3'te farklı komşuluk sayısı (k) değerleri için k-EYK algoritması ortalama sınıflandırma başarıları ve hata matrisi köşegen yüzde değerleri verilmiştir. Bu işlem sonucu delta frekans bandında en yüksek başarı oranı %86.2 (k=7), teta frekans bandında en yüksek başarı oranı %87.2 (k=7), alfa frekans bandında en yüksek başarı oranı %83.3 (k=9), beta frekans bandında en yüksek başarı oranı %75.1 (k=9) ve gama frekans bandında en yüksek başarı oranı %70.8 (k=3) olarak elde edilmiştir.

Tablo 3. Farklı frekans bandlarına ait k-EYK algoritması ortalama sınıflandırma başarıları.

Table 3. k-NN algorithm average classification successes of different frequency bands.

Frekans Bandı	Farklı Komşuluk (k) Değerleri				
	k=1	k=3	k=5	k=7	k=9
Delta 208x224	%78.3	%75.5	%82.4	%86.2	%83.4
	[75-80-77-81]	[73-76-71-78]	[75-81-79-87]	[79-88-82-91]	[78-85-83-90]
Teta 208x224	%78.4	%78.4	%82.9	%87.2	%80.3
	[72-79-75-81]	[71-80-75-81]	[78-84-75-86]	[79-88-85-90]	[78-80-75-82]
Alfa 208x224	%77.9	%76.4	%80.2	%72.6	%83.3
	[70-77-75-79]	[71-78-74-78]	[75-81-78-83]	[72-75-70-78]	[78-85-80-88]
Beta 208x224	%69.8	%70.8	%68.8	%67.9	%75.1
	[59-72-61-75]	[62-69-60-73]	[61-69-61-72]	[52-68-59-71]	[68-77-70-80]
Gama 208x224	%68.4	%70.8	%65.3	%62.4	%67.5
	[57-69-62-72]	[60-68-65-75]	[52-67-55-70]	[50-65-53-68]	[58-71-65-77]

Tablo 4'te farklı çekirdek fonksiyonları için DVM algoritması ortalama sınıflandırma başarıları ve hata matrisi köşegen yüzde değerleri verilmiştir. Bu işlem sonucu delta frekans bandında en yüksek başarı oranı %87.8 (çekirdek fonksiyon: doğrusal), teta frekans bandında en yüksek başarı oranı %81.4 (çekirdek fonksiyon: karesel), alfa frekans bandında en yüksek başarı oranı %80.4 (çekirdek fonksiyon: doğrusal), beta frekans bandında en yüksek başarı oranı %72.3 (çekirdek fonksiyon: kübik) ve gama frekans bandında en yüksek başarı oranı %71.3 (çekirdek fonksiyon: karesel) olarak elde edilmiştir. Tablo 5' da farklı çekirdek fonksiyonları için DAA algoritması ortalama sınıflandırma başarıları ve hata matrisi köşegen yüzde değerleri verilmiştir. Bu işlem sonucu; delta frekans bandında en yüksek başarı oranı %87.7 (çekirdek fonksiyon: doğrusal), teta frekans bandında en yüksek başarı oranı %86.7 (çekirdek fonksiyon: doğrusal), alfa frekans bandında en yüksek başarı oranı %80.9 (çekirdek fonksiyon: doğrusal), beta frekans bandında en yüksek başarı oranı %75.7 (çekirdek fonksiyon: doğrusal) ve gama frekans bandında en yüksek başarı oranı %67.0 (çekirdek fonksiyon: doğrusal) olarak elde edilmiştir.

Tablo 4. Farklı frekans bandlarına ait DVM algoritması ortalama sınıflandırma başarıları.

Table 4. SVM algorithm average classification successes of different frequency bands.

Frekans Bandı	Çekirdek Fonksiyonları		
	Doğrusal	Karesel	Kübik
Delta	%87.8	%84.9	%83.4
208x224	[78-85-80-89]	[76-85-78-86]	[76-84-80-85]
Teta	%76.9	%81.4	%78.9
208x224	[73-77-71-81]	[78-83-80-85]	[72-80-75-83]
Alfa	%80.4	%76.9	%75.5
208x224	[81-82-78-85]	[75-77-72-78]	[70-77-74-80]
Beta	%65.7	%68.4	%72.3
208x224	[51-68-63-70]	[53-70-62-72]	[63-72-68-75]
Gama	%68.8	%71.3	%67.9
208x224	[63-70-65-72]	[63-70-65-75]	[66-70-63-75]

Tablo 5. Farklı frekans bandlarına ait DAA algoritması ortalama sınıflandırma başarıları.

Table 5. LDA algorithm average classification successes of different frequency bands.

Frekans Bandı	Çekirdek Fonksiyonları	
	Doğrusal	Karesel
Delta	%87.7	%84.3
208x224	[81-88-84-90]	[78-85-80-87]
Teta	%86.7	%82.2
208x224	[78-85-81-90]	[75-82-78-85]
Alfa	%80.9	%75.2
208x224	[71-78-75-83]	[70-77-72-79]
Beta	%75.7	%69.4
208x224	[70-76-72-80]	[58-67-61-72]
Gama	%67.0	%61.3
208x224	[54-68-62-70]	[50-62-58-68]

Literatürde yapılan birçok çalışmada, delta (1-3 Hz), teta (4-7 Hz) ve alfa (8-12 Hz) frekans bandlarının yorgunluk, uyusukluk, uykululuk ve iş performansında düşüş anları ile ilişkili olduğu gösterilmiştir. Bu çalışmada da sınıflandırıcı başarı sonuçları incelendiğinde, tüm sınıflandırıcılar için en yüksek başarı oranları delta, teta ve alfa frekans bandlarında elde edilmiştir. Bu nedenle çalışmanın ikinci kısmında bu frekans bandlarına ait öznelikler kullanılarak, toplam öznelik matrisi oluşturulmuştur.

Yeni öznelik matrisinin boyutu 208x672 olarak elde edilmiştir. Öznelik matrisinin yüksek boyutlu olması sınıflandırma başarısını ve işlem süresini olumsuz yönde etkilemektedir. Bu nedenle bir sonraki adımda temel bileşen analizi algoritması kullanılarak öznelik matrisinin boyutu 208x256 olarak azaltılmıştır. Düşük boyutlu öznelik matrisi, Böl ve Fethet algoritması kullanılarak sınıflandırılmış ve ortalama sınıflandırma başarıları değerlendirilmiştir.

Böl ve Fethet algoritmasında veri seti içerisinde bulunan sınıflar iki parçaya ayrılmakta ve sonucunda yedi olası durum meydana gelmektedir. Bu durumlar 1'e karşı 2-3-4, 2'ye karşı 1-3-4, 3'e karşı 1-2-4, 4'e karşı 1-2-3, 1-2'ye karşı 3-4, 1-3'e karşı 2-4 ve 1-4'e karşı 2-3 olarak sıralanmaktadır. Böylece sınıf bilgisi karıştırılarak sınıflandırma problemi iki sınıfa indirgenmektedir. Algoritma içerisinde alt sınıflandırıcı olarak

k-EYK, DVM ve DAA sınıflandırıcıları kullanılmıştır. Sınıf sayısının ikiye düşmesi ile hata matrisi 2x2 boyutuna düşmektedir. Tablo 6'da belirtilen yedi duruma ait ortalama sınıflandırma başarıları ve hata matrisi köşegen yüzde değerleri verilmektedir.

Tablo 6. Böl ve Fethet algoritması ile ilk sınıfın ayrılması.

Table 6. Separation of the first class by Divide and Conquer algorithm.

Gruplar	Sınıflandırıcı Algoritmaları		
	k-EYK [k]	DVM [Çekirdek Fonk.]	DAA [Çekirdek Fonk.]
1-2,3,4	%76.7	%78.2	%75.3
	[k=3]	[Doğrusal]	[Doğrusal]
2-1,3,4	71-82	73-83	70-79
	%80.2	%82.4	%79.8
3-1,2,4	[k=5]	[Karesel]	[Doğrusal]
	80-83	82-85	78-84
4-1,2,3	%73.2	%71.3	%70.2
	[k=3]	[Karesel]	[Karesel]
1,2,3,4	70-75	68-73	63-72
	%92.8	%93.2	%84.4
1,3-2,4	[k=7]	[Karesel]	[Karesel]
	96-91	97-90	89-81
1,2-3,4	%80.5	%80.4	%80.4
	[k=11]	[Doğrusal]	[Doğrusal]
1,3-2,4	78-85	76-82	75-82
	%84.3	%88.5	%81.3
1,4-2,3	[k=3]	[Doğrusal]	[Doğrusal]
	81-87	83-90	77-85
1,4-2,3	%75.8	%77.9	%76.3
	[k=5]	[Kübik]	[Karesel]
	74-80	76-82	75-81

Tablo 6'da verilen ortalama sınıflandırma başarıları incelendiğinde, 4 No.lu sınıf diğer sınıflardan (1-2-3) DVM sınıflandırıcısı kullanılarak en başarılı şekilde ayrılmıştır. Böylece sınıflandırmanın ikinci kısmında, 1-2-3 No.lu sınıflar ayırt edilmek için iki sınıflı problemler haline getirilmiştir. Bu işlem sonucu üç olası durum oluşmaktadır. Bu durumlar 1'e karşı 2-3, 2'ye karşı 1-3 ve 3'e karşı 1-2 olarak sıralanmaktadır. Tablo 7'de belirtilen üç duruma ait ortalama sınıflandırma başarıları ve hata matrisi köşegen yüzde değerleri verilmektedir.

Tablo 7. Böl ve Fethet algoritması ile ikinci sınıfın ayrılması.

Table 7. Separation of the second class by Divide and Conquer algorithm.

Gruplar	Sınıflandırıcı Algoritmaları		
	k-EYK [k]	DVM [Çekirdek Fonk.]	DAA [Çekirdek Fonk.]
1-2,3	%84.5	%80.3	%75.6
	[k=7]	[Doğrusal]	[Doğrusal]
2-1,3	75-86	71-82	68-78
	%88.3	%91.2	%81.2
3-1,2	[k=11]	[Karesel]	[Karesel]
	85-91	87-95	77-86
1-2,3	%78.2	%79.4	%73.7
	[k=3]	[Doğrusal]	[Doğrusal]
	70-82	73-83	69-77

Tablo 7’ de verilen ortalama sınıflandırma başarıları incelendiğinde, 2 No.lu sınıf diğer sınıflardan (1-3) DVM sınıflandırıcısı kullanılarak en başarılı şekilde ayrılmıştır. Bu nedenle sınıflandırmanın son kısmında, 1 No.lu ve 3 No.lu sınıflar birbirinden ayırt edilmeye çalışılmıştır. Tablo 8’de belirtilen duruma ait ortalama sınıflandırma başarıları hata matrisi köşegen yüzde değerleri verilmektedir.

Tablo 8. Böl ve Fethet algoritması ile üçüncü sınıfın ayrılması.

Table 8. Separation of third class with Divide and Conquer algorithm.

Gruplar	Sınıflandırıcı Algoritmaları		
	k-EYK [k]	DVM [Çekirdek Fonk.]	DAA [Çekirdek Fonk.]
1-3	%88.3	%88.5	%81.6
	[k=9] 78-85	[Karesel] 83-79	[Doğrusal] 74-82

Çalışmada 1 No.lu ve 3 No.lu sınıflar birbirinden ayrıştırılarak Böl ve Fethet algoritması tamamlanmıştır. Algoritma sonucu ortalama sınıflandırma başarıları k-EYK algoritması için %88.3, DVM algoritması için %88.5 ve DAA için %81.6 olarak elde edilmiştir.

4 Sonuçlar

Bu çalışmada, EEG sinyalleri kullanılarak sürücü yorgunluk evrelerini dinamik olarak tespit edebilecek bir sistem tasarlanmıştır. Veri seti içerisinde bulunan EEG sinyalleri dört farklı sınıfa ayrılmış ve yorgunluk evreleri oluşturulmuştur. Çalışmanın ilk kısmında EEG sinyalleri alt frekans bandlarına ayrılmış ve her frekans bandı için Fourier Dönüşümü ve Dalgacık Dönüşümü tabanlı öznelikler çıkarılmıştır. Bu işlem sonucu boyutu 208x224 olan öznelik matrisi, üç farklı sınıflandırıcı algoritması ile sınıflandırılmıştır.

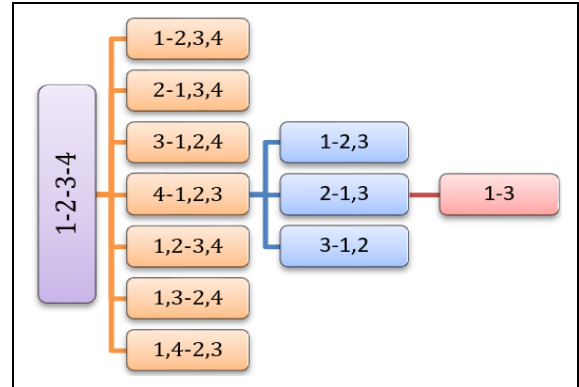
Ortalama sınıflandırma başarıları k-EYK Algoritması için; delta frekans bandında en yüksek %86.2 (k=7), teta frekans bandında en yüksek %87.2 (k=7), alfa frekans bandında en yüksek %83.3 (k=9), beta frekans bandında en yüksek %75.1 (k=9) ve gama frekans bandında en yüksek %70.8 (k=3) olarak elde edilmiştir. Kullanılan bir diğer sınıflandırıcı, DVM algoritması için ortalama sınıflandırma başarıları; delta frekans bandında en yüksek %87.8 (Doğrusal), teta frekans bandında en yüksek %81.4 (Karesel), alfa frekans bandında en yüksek %80.4 (Doğrusal), beta frekans bandında en yüksek %72.3 (Küçük) ve gama frekans bandında en yüksek %71.3 (Doğrusal) olarak elde edilmiştir. Çalışmada kullanılan son sınıflandırıcı, DAA algoritması için ortalama sınıflandırma başarıları; delta frekans bandında en yüksek %87.7 (Doğrusal), teta frekans bandında en yüksek %86.7 (Doğrusal), alfa frekans bandında en yüksek %80.9 (Doğrusal), beta frekans bandında en yüksek %75.7 (Doğrusal) ve gama frekans bandında en yüksek %67.0 (Doğrusal) olarak elde edilmiştir.

Ortalama sınıflandırma başarıları analiz edildiğinde, tüm sınıflandırıcı algoritmaları için delta, teta ve alfa frekans bandları en yüksek ortalama sınıflandırma başarısına sahiptir. Bu frekans bandlarında gerçekleşen beyin aktiviteleri incelenecek olursa, delta frekans bandı NREM derin uyku hali, teta frekans bandı hafif uyku hali ve alfa frekans bandı gözler kapalı, sakin uyanıklık hali ile ilişkilendirilmektedir. Çalışmanın ilk kısmında elde edilen ortalama sınıflandırma başarımları ile bu beyin aktiviteleri tutarlılık göstermektedir. Bu nedenle çalışmanın ikinci kısmında temel öznelik matrisi

oluşturulurken bu frekans bandlarına ait öznelikler birleştirilmiştir.

Çalışmanın ikinci kısmında, boyutu 208x672 olan öznelik matrisi, sistem performansını ve ortalama sınıflandırma başarısını olumsuz etkileyeceği için Temel Bileşen Analizi algoritması ile boyutu 208x256’ya düşürülmüştür. Yeni öznelik matrisi Böl ve Fethet algoritması kullanılarak sınıflandırılmıştır. Sınıflandırıcı algoritması içerisinde alt sınıflandırıcı olarak kullanılmıştır. Sınıflandırma işleminin ilk adımında, var olan dört sınıfın olası tüm kombinasyonları oluşturularak sistem iki sınıflı sınıflandırma problemi haline getirilmiştir. Bu işlem sonucu 4 No.lu sınıf diğer sınıflardan DVM algoritması kullanılarak %93.2 ortalama sınıflandırma başarısı ile en iyi şekilde ayrılmıştır.

Kalan sınıflar (1-2-3) için tüm kombinasyonlar oluşturularak işlem tekrarlanmıştır. Bu işlem sonucu 2 No.lu sınıf diğer sınıflardan DVM algoritması kullanılarak %91.2 ortalama sınıflandırma başarısı ile en iyi şekilde ayrılmıştır. Son olarak 1 No.lu ve 3 No.lu sınıflar k-EYK algoritması kullanılarak %83.8 ortalama sınıflandırma başarısı ile birbirinden ayrılmıştır. Algoritmaya ait ortalama sınıflandırma başarıları ise k-En Yakın Komşuluk algoritması için %88.3, Destek Vektör Makineleri algoritması için %88.5 ve Doğrusal Ayırım Analizi için %81.6 olarak elde edilmiştir. Böl ve fethet yöntemine ait temel ağaç yapısı Şekil 4’te verilmektedir.



Şekil 4. Böl ve Fethet algoritması temel ağaç yapısı.

Figure 4. Divide and Conquer algorithm basic tree structure.

Sonuç olarak 4 No.lu sınıfın en yüksek ortalama sınıflandırma başarısına sahip olması; 27 sürücünün %93.2 doğruluk oranı ile 67.5-90 dk. arasında 4. sınıf yorgunluk seviyesine sahip olduğunu göstermektedir. Araştırmalar ve istatistiksel veriler, sürücü kaynaklı kazaların, en çok varış noktasına yaklaştığı sürelerde olduğunu belirtmektedir. Çalışmamız sonucunda da kazaların başlıca nedeni olan yorgunluk, sürüşün son çeyreğinde gerçekleşmiştir. Daha uzun süreli sürüş yorgunluk seviyesini arttıracığı için, sürüş esnasında alınan veriler tasarladığımız sistemde analiz edildiğinde, doğruluk seviyesi artacaktır.

5 Conclusion

In this study, a system that can dynamically detect driver fatigue phases using EEG data is designed. EEG data in the data set were divided into four different classes and fatigue phases were created. In the first part of the study, EEG data were divided into sub-frequency bands and features based on Fourier Transform and Wavelet Transform were obtained for each frequency band. As a result of this process, the feature

matrix, whose size is 208x224, is classified by three different classifier algorithms.

Average classification success for k-EYK Algorithm; The highest 86.2% (k=7) in the delta frequency band, the highest 87.2% (k = 7) in the theta frequency band, the highest 83.3% (k=9) in the alpha frequency band, the highest 75.1% (k=9) in the beta frequency band and the highest 70.8% (k=3) in the gamma frequency band. Another classifier used is average classification successes for the DVM algorithm; The highest 87.8% (Linear) in the delta frequency band, the highest 81.4% (Quadratic) in the theta frequency band, the highest 80.4% (Linear) in the alpha frequency band, the highest 72.3% (Cubic) in the beta frequency band and in the gamma frequency band, the highest 71.3% (linear) was obtained. The last classifier used in the study, average classification successes for DAA algorithm; The highest 87.7% (Linear) in the delta frequency band, the highest 86.7% (Linear) in the theta frequency band, the highest 80.9% (Linear) in the alpha frequency band, the highest 75.7% (Linear) in the beta frequency band and in the gamma frequency band, the highest 67.0% (Linear) was obtained

When the average classification successes are analyzed, the delta, theta and alpha frequency bands have the highest average classification success for all classifier algorithms. If the brain activities in these frequency bands are examined, the delta frequency band NREM is associated with deep sleep, theta frequency band mild sleepiness, and the alpha frequency band eyes closed, calm wakefulness. These brain activities are consistent with the classification achievements obtained in the first part of the study. As a result of this process, in the second part of the study, the features of these frequency bands are combined while creating the basic feature matrix.

In the second part of the study, the size of the feature matrix, which is 208x672, has been reduced to 208x256 with Principal Component Analysis detection since it will negatively affect the system performance and classification performance. The new feature matrix is classified using the Divide and Conquer algorithm. K-EYK, SVM and DAA algorithms are used as sub-classifiers in the classifier algorithm. In the first step of the classification process, the system has been turned into a two-class classification problem by creating all possible combinations of the four existing classes. As a result of this process, class number four is distinguished from other classes with 93.2% average classification success using the DVM algorithm.

The process was repeated by creating all combinations for the remaining classes (1-2-3). As a result of this process, the second class is distinguished from other classes with 91.2% average classification success using the DVM algorithm. Finally, the number one and number three classes were distinguished from each other with 83.8% average classification success using the k-EYK algorithm. The average classification success of the algorithm was obtained as 88.3% for k-Nearest Neighborhood algorithm, 88.5% for Support Vector Machines algorithm and 81.6% for Linear Discrimination Analysis. The basic tree structure of the divide and conquer method is given in Figure 4.

As a result, class number four has the highest average classification success; It shows that 27 drivers had a grade 4 fatigue level between 67.5 and 90 min. with 93.2% accuracy. Studies and statistical data indicate that driver-related accidents occur mostly at the time they approach the destination. As a result of our study, fatigue, which is the main

cause of the accidents, occurred in the last quarter of the driving. Since longer driving will increase the level of fatigue, the level of accuracy will increase when the data received during driving are analyzed in the system we designed.

6 Yazar katkı beyanı

Bu çalışmada Şule BEKİRYAZICI araştırma sürecinde kullanılacak uygun veri seti bulunması tarafında katkıda bulunmuş olup Ahmet DEMİR sınıflandırma ve boyut azaltma algoritmaları üzerinde çalışmalar yapmıştır. Oğuzhan COŞKUN Matlab programında gerekli özneliklerin çıkarılması ve ilgili yazılım geliştirme sürecinin oluşturulması kısmında destekçi olup Recep EKEN makalenin kaynak araştırması ve bulguların derlenmesinde görev almıştır. Güneş YILMAZ çalışma süresince çok yönlü katkıda bulunmuş ve bilimsel altyapının oluşmasında destek olmuştur.

7 Etik kurul onayı ve çıkar çatışması beyanı

Hazırlanan makalede etik kurul izni alınmasına gerek yoktur. Hazırlanan makalede herhangi bir kişi/kurum ile çıkar çatışması bulunmamaktadır.

8 Kaynaklar

- [1] Eoh HJ, Chung MK, Kim S. "Electroencephalographic study of drowsiness in simulated driving with sleep deprivation". *International Journal of Industrial Ergonomics*, 35(4), 307-320, 2005.
- [2] Lin CT, Wu RC, Liang SF, Chao WH, Chen WH, Jung TP. "EEG-Based drowsiness estimation for safety driving using independent component analysis". *IEEE Transactions on Circuit and Systems I*, 52(3), 2726-2738, 2005.
- [3] Alakuş TB. EEG Verilerinden İşaret İşleme ve Sınıflandırma Teknikleri Kullanılarak Duygu Tahmini. Yüksek Lisans Tezi, Fırat Üniversitesi, Elazığ, Türkiye, 2018.
- [4] Turkey A, Fathi ES, Saleh AA, Ishtiaq A. "A review of channel selection algorithms for EEG signal processing". *EURASIP Journal on Advances in Signal Processing*, 66, 1-21, 2015.
- [5] Scott M, Tzyy-Ping J, Terrence S. "Awareness during drowsiness: dynamics and electrophysiological correlates". *Canadian Journal of Experimental Psychology*, 54(4), 266-273, 2000.
- [6] Lin CT, Chuang CH, Huang CS, Tsai SF, Lu SW, Chen YH, Ko LW. "Wireless and wearable EEG system for evaluating driver vigilance". *IEEE Transactions on Biomedical Circuits and Systems*, 8(2), 165-176, 2014.
- [7] Mehdi H, Jianqiu MZ, Yufei H. "Driver's fatigue prediction by deep covariance learning from EEG". *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Banff, Canada, 5-8 October 2017.
- [8] Chuang CH, Huang CS, Ko LW, Lin CT. "An EEG-Based perceptual function integration network for application to drowsy driving". *Knowledge-Based Systems*, 80, 143-152, 2015.
- [9] King LM, Nguyen HT, Lal SKL. "Early driver fatigue detection from electroencephalography signals using artificial neural networks". *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, New York, USA, 30 August-3 September. 2006.

- [10] Bashar SK, Bhuiyan MIH. "Classification of motor imagery movements using multivariate empirical mode decomposition and short time Fourier transform based hybrid method". *Engineering Science and Technology an International Journal*, 19, 1457-1464, 2016.
- [11] Cao Z, Chuang C, King J, Lin C. "Multi-Channel EEG recording during a sustained-attention driving task". *Scientific Data, Nature*, 6, 9-16, 2018.
- [12] Kumar JS, Bhuvaneswari P. "Analysis of EEG signals and its categorization-a study". *Procedia Engineering*, 38, 2525-2536, 2016.
- [13] Smith LI. "A tutorial on Principal Components Analysis". Department of Computer Science, Computer Science Technical Report No. OUCS-2002-12, University of Otago, Otago, New Zealand, 2002.
- [14] Kiang MY. "A comparative assessment of classification methods". *Decision Support Systems*, 35, 441-454, 2003.
- [15] Smith DR. "Applications of a strategy for designing divide-and-conquer algorithms". *Science of Computer Programming*, 8, 213-229, 1987.
- [16] Wang J, Neskovic P, Cooper, LN. "Improving nearest neighbor rule with a simple adaptive distance measure". *Pattern Recognition Letters*, 28, 207-213, 2007.
- [17] Coyle D, Prasa G, McGinnity TM. "A time-series prediction approach for feature extraction in a brain-computer interface". *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 13(4), 461-467, 2005.