

Alkolik ve Normal EEG Sinyallerinin Zaman-Alan Tanımlayıcı Analizi Tabanlı Otomatik Sınıflandırılması

Berna ARI^{1*}

¹ Elektrik-Elektronik Teknolojileri, Teknoloji Fakültesi, Fırat Üniversitesi, Elâzığ, Türkiye
^{*1} bernagurler.a@gmail.com

(Geliş/Received: 21/12/2022;

Kabul/Accepted: 07/01/2023)

Öz: Alkolizm, beyin problemlerine ve buna bağlı bilişsel, duygusal ve davranışsal bozukluklara yol açan ciddi bir hastalıktır. Alkolizmi tespit etmek için öne çıkan kaynaklardan biri, Elektroensefalogram (EEG) sinyallerini analiz etmektir. Fakat alkolik EEG sinyallerinin sınıflandırılması, alkolik kişilerin beyin hastalıklarının tanı ve tedavisine yönelik biyomedikal araştırmalarda zorlu bir süreçtir. Bu çalışmada, alkolik EEG sinyallerinden zaman-alan tanımlayıcılarına ve topluluk öğrenmesine dayalı otomatik olarak tanımlayan yeni bir yöntem sunulmaktadır. Önerilen yöntem, tek kanallı EEG sinyallerinin dalgacık paket ayrıştırma ile farklı frekans alt bantlarına ayrılması, zaman-alan tanımlayıcıları ile öznelik çıkarımı ve topluluk torbalama ağaçları ile sınıflandırma aşamalarından oluşmaktadır. Tek kanallı EEG veri seti ile yapılan deneysel çalışmalarda %97,50 başarımla sağlanmıştır. Deneysel sonuçlar önerilen yöntemin, son teknoloji yöntemlere kıyasla daha iyi bir performansa sahip olduğunu göstermektedir. Bu yöntem alkolik bireylerin otomatik tespitinde uzmanlara yardımcı olabilecektir.

Anahtar kelimeler: EEG, dalgacık paket ayrıştırma, zaman-alan tanımlayıcıları, sınıflandırma.

Automatic Classification of Alcoholic and Normal EEG Signals Based on Time-Domain Descriptor Analysis

Abstract: Alcoholism is a serious disease that causes brain problems and related cognitive, emotional, and behavioral disorders. One of the prominent resources for detecting alcoholism is analyzing Electroencephalogram (EEG) signals. However, the classification of alcoholic EEG signals is a challenging process in biomedical research for the diagnosis and treatment of brain diseases in alcoholics. In this study, a new method that automatically identifies alcoholic EEG signals based on time-domain descriptors (TDD) and ensemble learning (EBT) is presented. The proposed method consists of the separation of single-channel EEG signals into different frequency subbands with Wavelet Packet Decomposition (WPD), feature extraction with TDD, and classification with EBT. In experimental studies with a single-channel EEG data set, 97.50% success was achieved. Experimental results show that the proposed method has a better performance compared to the state-of-the-art methods. This method can help experts in the automatic detection of alcoholic individuals.

Key words: EEG, wavelet packet decomposition, time-domain descriptors, classification

1. Giriş

Alkolizm, obsesif-kompulsif (sürekli tekrar eden düşüncelere sahip olunması ve bu düşüncelerinin kişinin kendisini rahatsız etmesi) bozukluğa benzer şekilde, içmeye bağımlı olunan fiziksel bir hastalıktır [1,2]. Alkol bağımlısı hastalarda yaygın görülen olumsuz etkiler; ülser, özofagus kanaması, mide kanseri, akut ve kronik pankreas iltihabı gibi sindirim sistemi hastalıkları ve zihinsel engellilik, Alzheimer, inme gibi sinir sistemi bozukluklarıdır [3]. Ayrıca aşırı alkol tüketiminin yüksek tansiyona ve gut hastalığına neden olabileceği belirtilmektedir [2]. Dünya Sağlık Örgütü'nün (WHO) 2018 yılında yayımlanan birinci basamak sağlık hizmetlerinde kardiyovasküler raporuna göre alkolizm, hastalıklara neden olması bakımından üçüncü en yüksek risk faktörü olarak kabul edilmektedir [4]. Yine aynı rapora göre her yıl yaklaşık 3,3 milyon ölümün aşırı alkol tüketiminden kaynaklandığını ve bunun dünyadaki genel ölümlerin yaklaşık %5,9'unu oluşturduğunu belirtilmektedir [5].

Alkolizmin erken teşhisi; bireylerin durumlarını anlamalarına, sakatlık gibi kalıcı hasarların ve ölüm gibi hayati sonuçların önlenmesine yardımcı olabilir [6]. Geleneksel alkolizm tespiti anketlere, nefes testlerine ve kan testlerinin yanı sıra fizyolojik incelemelere dayanmaktadır [2]. Fakat bu geleneksel tespit yöntemleri, her birinin ayrı sınırlıklarının olması nedeni ile pek tercih edilmezler [5]. Bu nedenle araştırmacılar ve uzmanlar alkolizm, epilepsi, Alzheimer gibi karmaşık beyin bozukluklarını kapsayan hastalıkların tespitinde beynin çeşitli fizyolojik

* Sorumlu yazar: bernagurler.a@gmail.com. Yazarların ORCID Numarası: ¹ 0000-0003-1000-2619

ve patolojik aktivitelerin elektriksel resmini temsil eden elektroensefalografi (EEG) verilerini daha çok tercih etmektedirler [7]. Çünkü EEG verilerinin elde edilmesi düşük maliyetli, sistemin taşınabilir olması, cerrahi işlem gerektirmemesi, yüksek zamansal çözünürlük sağlaması ve beyin modellerini analiz etmek için etkili bir yaklaşımdır.

EEG sinyallerinden alkolizmin erken tespiti için literatürde birçok yöntem önerilmiştir. Bu yöntemler; zaman ve frekans alanlarında doğrusal ve doğrusal olmayan özellik tabanlı yöntemler, sinyal ayrışımına dayalı yöntemler ve bunların yanı sıra makine öğrenmesi ve derin öğrenme tabanlı yaklaşımları kapsamaktadır. İlk çalışmalar genellikle sinyal ayrışımına dayalı gerçekleştirilmiştir. Gökşen ve diğerleri [8], alkolizm sınıflandırmak için zaman alanı korelasyonunun karşılıklı bilgisi ile ölçülen fonksiyonel bağlantıyı vurgulamış ve %82,3 doğruluk elde etmiştir. Mümtaz ve diğerleri [9], EEG sinyallerinden alkolizmi tespit etmek için frekans alanının güç tutarlılığı fonksiyonel bağlantısını önererek ve %89,3 doğrulukta alkolizm tespiti yapmışlardır.

Sonrasında makine öğrenmesi yöntemleri, sınıflandırma çalışmalarında yaygın olarak tercih edilmiştir. Geleneksel istatistiksel sınıflandırma yöntemleriyle karşılaştırıldığında, makine öğrenimi yöntemleri daha doğru sınıflandırma sonuçları sağlayabildiği görülmüştür. Bajaj ve ark. [10], alkollü ve kontrol grubu EEG sinyallerini otomatik olarak sınıflandırmak için zaman-frekans görüntüleri, doku görüntüsü öznelik çıkarımı ve negatif olmayan en küçük kareler sınıflandırıcı (NNLS) temelli yeni bir hibrit yöntem geliştirdiler. Zaman-Frekans (Z-F) analizinde, kısa süreli Fourier dönüşümünün spektrogramı dikkate alarak Z-F görüntüleri elde ettiler. Elde edilen Z-F görüntüleri daha sonra 8 bitlik gri tonlamalı görüntülere dönüştürülerek yönlendirilmiş gradyanların (CoHOG) ve Eig(Hess)-CoHOG özelliklerinin histogramlarını çıkardılar. Son olarak, elde edilen özellikleri alkollü ve kontrol grubunun EEG sinyallerini NNLS sınıflandırıcısına uygulayarak sınıflandırdılar. Deneysel çalışmalarda %95,83 doğruluk oranına ulaşılar. Feyyaz ve ark. [11], kullanılan Destek Vektör Makinesi (Support Vector Machine-DVM) ve tepe (peak) görselleştirme yöntemiyle Uzun Kısa Süreli Bellek (Long-Short Time Memory-LSTM) modelinde %90,97 sonucuna ulaşılar. Agarwal ve Zubair [12] alkolik denekleri saptamak için, kayan tekil spektrumu analizi, bağımsız bileşen analizi ve XGBoost sınıflandırıcı kombinasyonundan oluşan bir yöntemi referans aldılar ve yapılan deneyle %98,97'lik bir doğruluk elde ettiler. Patidar ve ark. [13], EEG sinyallerinden gelen özneliklere dayanan ayarlanabilir Q dalgacık dönüşümü (Tunable Q-Factor Wavelet Transform-TQWT) ve en küçük kareler SVM yöntemlerinin kombinasyonunu kullanarak alkolizmin tespitinde yeni bir yaklaşım önerdiler. Yapılan deneysel çalışmalarda, %97,02 doğrulukta alkolizmi tespitini gerçekleştirdiler.

Son olarak manuel özellik çıkarımı yerine otomatik özellik çıkarımını içeren derin öğrenme yöntemleri kapsamında bazı çalışmalar yapılmıştır. Shen vd. [2] alkolik EEG sinyallerinden tüm beyin bağlantı analizi ve derin öğrenme yöntemleri aracılığıyla sınıflandırılmasına odaklandılar. EEG sinyalinden seçilen her frekans bandında Z-F alanı bilgisini çıkarmak için sürekli Dalgacık dönüşümü uygulandıktan sonra 2B ve 3B evrişimli sinir ağıları (Convolution Neural Networks-CNN), alkolik denekler ve sağlık kontrolü deneklerini sınıflandırdılar. UCI Alkollü EEG veri seti ile önerilen yöntemi test ederek, %96,25 doğrulukla sınıflandırma gerçekleştirdiler. Farsi ve vd. [14], derin öğrenme yöntemlerinin LSTM algoritmasının EEG alkolizm sinyalini doğrudan sınıflandırabildiğini ve deneysel çalışmalarda %93'lük doğruluk oranına ulaştıklarını belirtmektedirler.

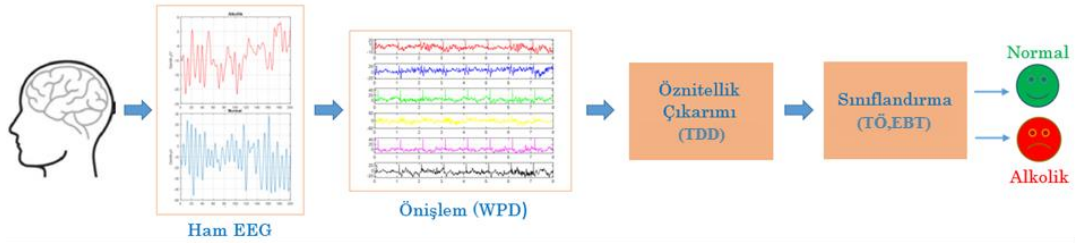
Bu çalışmada alkolizmin tespiti Zaman Alan Tanımlayıcı (Time-Domain Descriptors-TDD) ve Topluluk Torbalama Ağaçları (Ensemble Bagged Tree-EBT) öğrenmesi temelli bir model önerilmiştir. Bu modelde tek kanallı EEG sinyalleri Dalgacık Paket Ayrıştırması (Wavelet Packet Decomposition-WPD) yöntemini kullanarak EEG sinyalleri frekans alt bantlarına ayrıldıktan sonra öznelik çıkarımında TDD uygulanmıştır. Elde edilen öznelikler alkolizmin tespiti için EBT sınıflandırıcıya uygulanmıştır. Çalışmanın literatüre katkıları şunlardır;

- Alkollü EEG sinyallerinin otomatik olarak tanımlanması için yeni bir model önerildi.
- WPD ile tek kanallı EEG sinyalleri frekans alt bantlarına ayrılarak sınıflandırma performansını arttıracak daha fazla öznelik elde edilmesi sağlandı.
- Mevcut çalışmalara kıyasla daha yüksek başarımlar sağlanması uzmanlara karar vermede destek sağlayacaktır.

2. Materyal ve Metodoloji

Bu çalışmada alkolik ve normal olarak etiketlenen iki sınıflı EEG sinyallerini sınıflandırmak için zaman-alan tanımlayıcılarına dayalı bir model önerilmektedir. Önerilen yöntem; ön işlem, öznelik çıkarımı ve bu özneliklerin sınıflandırılması olmak üzere üç aşamadan oluşmaktadır. İlk olarak, ön işlem aşamasında WPD ile tek kanallı EEG sinyalleri dört farklı frekans alt bantlarına ayrıldı. Ardından tüm bu sinyallerden TDD yöntemi ile öznelik çıkarımı yapıldı. Son aşamada ise öznelikler alkolizmin tespiti için EBT sınıflandırıcıya

gönderilmiştir. Öznitelik çıkarımı ve sınıflandırıcı parametrelerinin belirlenmesi için kapsamlı deneysel çalışmalar yapılmıştır. Önerilen yöntemin genel yapısı Şekil 1’de verilmiştir.

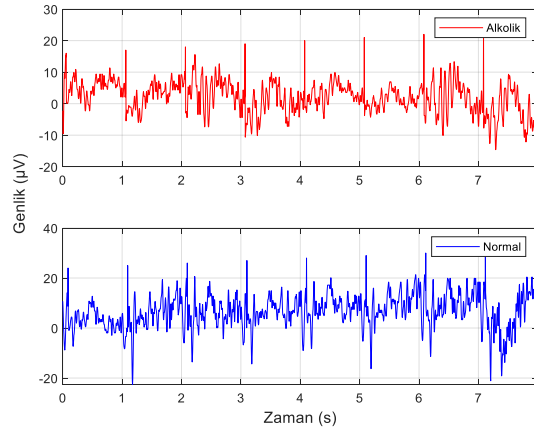


Şekil 1. Önerilen yöntemin genel yapısı.

2.1. Veri seti

Bu çalışmada, California Üniversitesi, Irvine (UCI) tarafından sağlanan açık veri seti kullanılmıştır [15]. Veri seti, EEG sinyalleri ile alkolizme genetik yatkınlık arasındaki ilişkiyi araştırmak için oluşturulmuştur. Veri seti, Amerikan Elektroensefalografi Derneği, 1990 Standart Elektrot Konum Adlandırmasına uygun olarak elde edilen 122 normal ve alkolik hastanın EEG kayıtlarını içermektedir. Bu deneklerin her biri, sırasıyla [16]’de açıklanan görüntülerden seçilen nesnelere 90 görüntüsünden oluşan farklı uyaranlar için 120 denemeye tabi tutulmuştur.

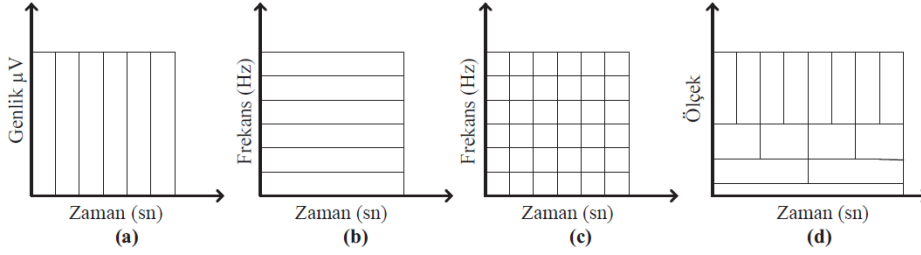
EEG sinyalleri, 61 elektrotlu başlık (Electrocap International, ECI) ile 32 saniye süreyle kaydedilmiş olup 256 Hz ve 12 bit çözünürlükte örneklendirilmiştir. Veri setinin hazırlanmasında, istenmeyen göz ve vücut hareketlerini içeren EEG sinyalleri hariç tutulmuş ve her normal ve alkolik sınıf için 30 kayıt dikkate alınmıştır. Bu çalışma, her normal ve alkolik sınıftan 120 EEG sinyali kullanmaktadır. Bu EEG sinyallerinin her biri 8 saniye süre ile kaydedilmiş ve 2048 örnek içermektedir. Şekil 1’de, sırasıyla alkolik ve normal EEG sinyallerinin örnekleri görülmektedir.



Şekil 2. Veri setinde sınıflara ait EEG sinyalleri.

2.2. Dalgacık dönüşümü

Dalgacık dönüşümü (Wavelet Transform-WT), Mallat [17] tarafından geliştirilen çok çözünürlüklü bir analiz yöntemidir. Fourier Dönüşümü (Fourier Transform-FT) gibi diğer klasik yöntemler bir sinyali yalnızca frekans alanında analiz ederken, WT’nin sinyali hem zaman hem de frekans alanında açıklayan üstün yönleri vardır. EEG sinyalleri, beyin aktiviteleri sırasında üretilen elektriksel bir işarettir. Bu sinyaller aynı zamanda durağan olmayan ve zamanla değişen kayıtlardır. Bu nedenle WT, EEG analizi için önemli bir yaklaşımdır [18]. WT, bir sinyalin düşük ve yüksek frekanslı bileşenleri için üstün zaman çözünürlüğü sunar ve bu nedenle, kısa süreli FT (Short Time Fourier Transform-STFT) gibi sabit bir zaman-frekans probleminin üstesinden gelebilmektedir. Zaman-frekans düzleminin genel bir çerçevesi Şekil 3’te gösterilmektedir.



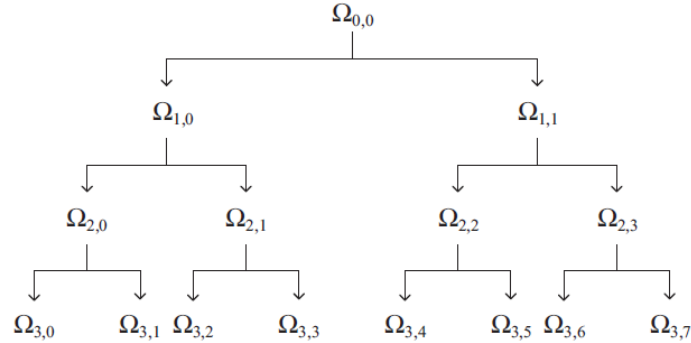
Şekil 3. (a) Pencereyi zaman-genlik düzlemi, (b) pencereyi Fourier dönüşümü frekans-genlik spektrumu, (c) STFT ile frekans-zaman düzlemi ve (d) WT ile pencereyi teknik [18].

WT'nin en önemli özelliği, farklı ayrıştırma seviyelerine bağlı olarak pencere genişliğinin değişmesidir. Böyle bir dalgacık analizinde ölçekleme fonksiyonu ve ana dalgacık dikkate alınması gereken en önemli fonksiyonlardır. Bir ana dalgacık $\psi(t)$ sıfır ortalamalı bir fonksiyondur ve aşağıdaki denklemle ifade edilir:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (1)$$

WPT, bir sinyalin daha yüksek frekans alanlarını analiz etmek ve daha yüksek ayrıştırma elde etmek için özel yeteneklere sahip kullanışlı bir araçtır. Dolayısıyla bu araç, sinyal işleme için WT'den daha uygun ve etkilidir.

WPT'yi anlamak için alt uzay ağacını dikkate almak gerekir. Ağaçta, düğümleri adlandırmak için $\Omega_{j,k}$ gösterimi kullanılır [19,21]. Burada j , skalayı ve k mevcut skalaya bağlı alt bantları göstermektedir. $\Omega_{0,0}$ kök düğüm ve EEG sinyalimizi temsil etmektedir. Şekil 4'de örnek bir WPT ağacı görülmektedir.



Şekil 4. WPD ağacının gösterimi.

Kök ağacında ilk olarak kök düğümünden, yaklaşım uzayına $\Omega_{j,k} \rightarrow \Omega_{j+1,2k}$ ve detay uzayına $\Omega_{j,k} \rightarrow \Omega_{j+1,2k+1}$ ayrıştırır. Bu işlemler, $\Omega_{j,k}$ 'nin $\{\phi_j(t - 2^{j+1}k)\}_{k \in \mathbb{Z}}$ yeni ortogonal tabanlara $\Omega_{j+1,2k}$ 'nin $\{\phi_{j+1}(t - 2^{j+1}k)\}_{k \in \mathbb{Z}}$ ve $\Omega_{j+1,2k+1}$ 'nin $\{\psi_{j+1}(t - 2^{j+1}k)\}_{k \in \mathbb{Z}}$ 'ye ayırarak uygulanır. Burada $\Phi_{j,k}(t)$ ölçek fonksiyonları olup $\Psi_{j,k}(t)$ [19,21] tarafından verilen dalgacık fonksiyonlarını temsil etmektedir.

$$\phi_{j,k}(t) = \frac{1}{\sqrt{|2^j|}} \phi\left(\frac{t - 2^j k}{2^j}\right) \quad (2)$$

$$\psi_{j,k}(t) = \frac{1}{\sqrt{|2^j|}} \psi\left(\frac{t - 2^j k}{2^j}\right) \quad (3)$$

Denklem 2 ve 3'teki 2^j , ölçekleme derecesini ölçen genişleme parametresi ve $2^j k$ dalgacığın zaman konumunu belirleyen öteleme parametresini tanımlamaktadır.

2.3. Zaman alan tanımlayıcıları

TDD, aslında EMG sinyallerinin sınıflandırılması amacıyla geliştirilen bir yöntemdir [22]. Bununla birlikte EEG sinyalleri de EMG sinyalleri gibi insan vücudunun belirli bölgelerindeki aktiviteler sonucu oluşan elektriksel işaretler olması bakımında benzerlikler göstermektedir [23]. TDD, detayları aşağıda verilen ve f_1, f_2, \dots, f_6 şeklinde tanımlanan altı ana özellikten oluşmaktadır [24].

- a) Kök Kare Sıfır Mertebe Momenti; bu özellik frekans alanındaki toplam gücü gösterir ve aşağıdaki gibi tanımlanabilir.

$$\bar{m}_0 = \sqrt{\sum_{j=0}^{N-1} x[j]^2} \quad (4)$$

burada $x[i]$, f_s örnekleme ile örneklenmiş ve $j = 1, 2, \dots, N$ uzunluğunda EEG sinyalini göstermektedir.

- b) Kök Kare İkinci ve Dördüncü Mertebe Momentler: ayrıık Fourier dönüşümü kullanılarak elde edilen $X[k]$ sinyali referans alınarak ikinci ve dördüncü derece momentler aşağıdaki gibi yazılabilir.

$$\bar{m}_2 = \sqrt{\sum_{j=0}^{N-1} k^2 P[k]} = \sqrt{\frac{1}{N} \sum_{j=0}^{N-1} (kX[k])^2} = \sqrt{\frac{1}{N} \sum_{j=0}^{N-1} (\Delta x[i])^2} \quad (5)$$

$$\bar{m}_4 = \sqrt{\sum_{j=0}^{N-1} k^4 P[k]} = \sqrt{\frac{1}{N} \sum_{j=0}^{N-1} (\Delta^2 x[i])^2} \quad (6)$$

burada $kX[k]$ frekans fonksiyonuna karşılık gelen $k^2 P[k]$ notasyonu ile gösterilen, değişken spektrumlu bir güç olarak tanımlanmaktadır [25]. Ayrıca, zaman fonksiyonunun birinci türevi k ile çarpılan bir frekans fonksiyonunun zaman eşdeğeri olarak m_0, m_2 ve m_4 momentleri aşağıdaki gibi tanımlanabilir.

$$m_0 = \frac{\bar{m}_0^\lambda}{\lambda}, m_4 = \frac{\bar{m}_4^\lambda}{\lambda}, m_8 = \frac{\bar{m}_8^\lambda}{\lambda} \quad (7)$$

Böylece ilk üç özelliği Denklem 7'yi referans alarak aşağıdaki gibi yazılabilir.

$$f_1 = \log(m_0), f_2 = \log(m_0 - m_2), f_3 = \log(m_0 - m_4) \quad (8)$$

- c) Seyreklik (SP): bu özellik bir vektörün ne kadar enerjisinin yalnızca birkaç bileşene sıkıştırıldığını ölçer ve aşağıdaki tanımlanabilir.

$$f_4 = \log \left(\frac{m_0}{\sqrt{m_0 - m_2} - \sqrt{m_0 - m_4}} \right) \quad (9)$$

- d) Düzensizlik faktörü (irregularity faktör-IF): tepe sayısına bölünen yukarı doğru sıfır geçiş sayısının oranını temsil eden bir ölçüdür. Bu özellik aşağıdaki şekilde yazılabilir.

$$f_5 = \log \left(\frac{ZC}{NP} \right) = \log \left(\frac{\sqrt{\frac{m_4}{m_0}}}{\sqrt{\frac{m_8}{m_4}}} \right) = \log \left(\frac{m_4}{\sqrt{m_0 m_8}} \right) \quad (10)$$

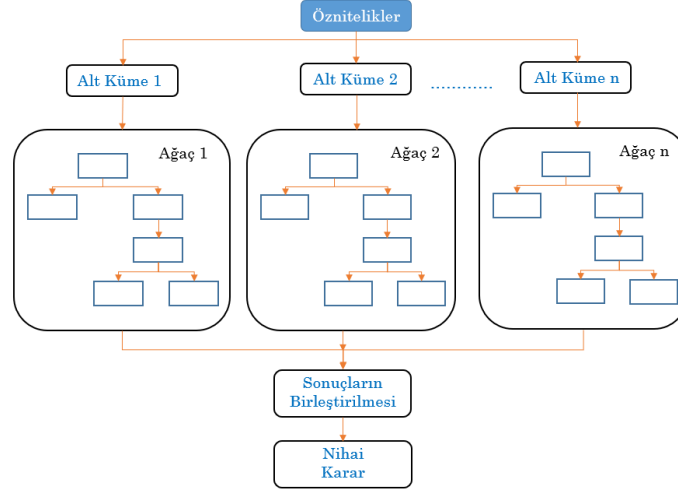
burada ZC sıfır geçişlerini ve NP ise tepe noktalarının sayısını tanımlamaktadır.

- e) Dalga boyu oranı (wavelength ratio-WL): Birinci türevin dalga biçimi uzunluğunun ikinci türevin dalga biçimi uzunluğuna oranı olarak tanımlarız. Bu özellik aşağıdaki gibi tanımlanabilir.

$$f_6 = \log \left(\frac{\sum_{j=0}^{N-1} |\Delta^2 x|}{\sum_{j=0}^{N-1} |\Delta^4 x|} \right) \quad (11)$$

2.4. Topluluk Torbalama Ağaçları

Topluluk Öğrenimi (TÖ), denetimli öğrenmeye dayalı bir makine öğrenimi tekniğidir [26]. Topluluk öğrenim sistemleri, herhangi bir sorunun birkaç uzmanın görüşü aranarak ve uygulanarak kolayca çözülebileceğini düşünen insan davranışından ilham almıştır. Bu farklı görüşler doğrultusunda karar verilir. Topluluk öğrenim sistemleri, tek bir sınıflandırıcı kullanmaya kıyasla daha iyi performans sağlar. Topluluk öğrenme tekniği tipik olarak iki alt kategoriye ayrılır. İlki torbalama olarak bilinir ve Breiman [27] tarafından geliştirilmiştir, ikincisi artırma olarak bilinir ve Freund ve Schapire tarafından önerilmiştir [28]. Bunlar arasında torbalama, en verimli ve yüksek doğruluğa sahip topluluk algoritmasıdır [29]. EBT’de “torbalama”, eğitim veri kümelerinin sürekli olarak rastgele örnekler çekerek değiştirildiği ön yükleme toplaması anlamına gelir. Karar ağaçları (Decision Tree-DT), eğitim modeline duyarlı oldukları için zayıf öğrenenler kategorisine aittir ve bu nedenle, bireysel bir karar ağacı normalde belirli eğitim modeline aşırı uyum sağlar [30]. Torbalı DT, karar ağaçlarının performansını iyileştirmek için kullanılabilir, çünkü kararlarının çoğunluğunu alarak birçok bireysel karar ağacının sonuçlarını toplar, sonuç olarak aşırı uyum sorununu çözer ve bireysel karar ağaçlarının performansını iyileştirir. Torbalama eğitim aşamaları Şekil 5’de görülmektedir. Şekil 5’de görüldüğü üzere öznelik veri setini n sayıda öznelik veri setine rastgele bir kombinasyonunun üretilmesiyle başlar. Bu, orijinal öznelik veri setinin bu alt kümeleri üzerinde farklı sınıflandırıcıların eğitimi gerçekleşir. Sonunda, bireysel modellerin çoğunluk oylarına dayalı olarak bir model geliştirilir. Daha sonra nihai modelin kararına dayalı olarak bir tahmin yapılır [31]. Belirli bir veri kümesinde, önyüklemeli alt örnekler belirlenir. Önyükleme uygulanan her örnekte bir DT oluşturulur. Her DT’nin sonucu, en güçlü ve en doğru sonuçları elde etmek için toplanır.



Şekil 5. Torbalama eğitim aşamaları [31].

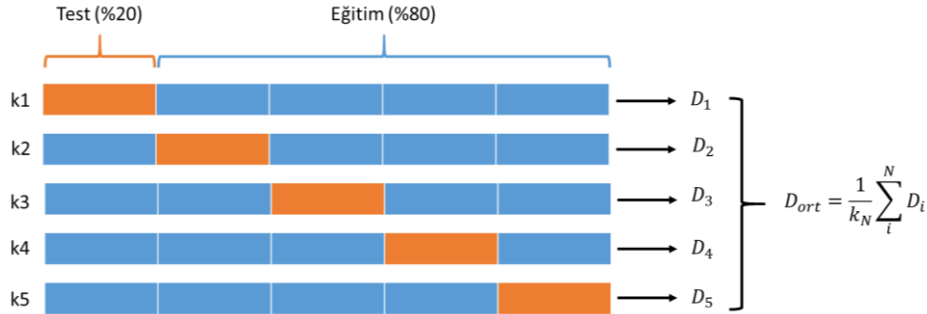
2.5. Performans değerlendirme kriterleri

Genellikle sınıflandırma modellerinde, önerilen sınıflandırıcı başarısının ölçülmesi sınıflandırıcı tarafından tahmin edilen sınıf değerleri ile gerçek sınıf değeri arasındaki ilişkilere göre belirlenir. Bunun için karmaşıklık

matrislerinden faydalanılmıştır. Önerilen yöntemin sonuçlarını analiz etmek için doğruluk, duyarlılık, özgüllük, kesinlik ve F1-skor parametreleri kullanılmıştır [32].

3. Deneysel Çalışma ve Sonuçlar

Deneysel çalışmalar, MATLAB kullanılarak, 32 GB RAM, NVIDIA P2200 Quadro GPU ve Intel Xeon Silver 3.3 GHz işlemciye sahip bilgisayarda gerçekleştirildi. Deney düzeneğinin ilk aşamasında tek kanallı her biri 1×2048 boyutundaki alkolik ve normal EEG sinyalleri WPD ile dört farklı frekans alt bantlarına ayrılarak 4×512 boyutunda EEG sinyalleri elde edilir. Elde edilen bu EEG sinyallerinden öznelik çıkarımı için zaman alan tanımlayıcıya uygulanır. TDD her bir alt bant için 4×6 boyutunda veri ve 1×2048 boyutundaki bir EEG sinyali için 4×24 boyutunda bir öznelik matrisi form edilir. Bu matris düzeltme işlemi ile 1×96 boyutunda öznelik vektörüne dönüştürülür. Bu işlem sonucunda toplamda bütün veri seti için 240×96 boyutunda öznelik matrisi elde edilir. Alkolizmin tespiti için öznelikler matrisi EBT sınıflandırıcıya girdi olarak uygulanır. Önerilen EBT sınıflandırıcıyı test etmek için yüksek düzeyde güvenilirliğe sahip olan 5-katlı çapraz doğrulama yöntemi kullanılmıştır [33]. Şekil 6'da 5-kat çapraz doğrulama yönteminin genel prensibi gösterilmektedir.



Şekil 6. 5 kat çapraz doğrulama modeli.

Şekil 6'de görüldüğü üzere D_i , ($i=1, 2, \dots, 5$) her bir kat için sistem performansını temsil etmekte ve D_{ort} genel performansı temsil etmektedir.

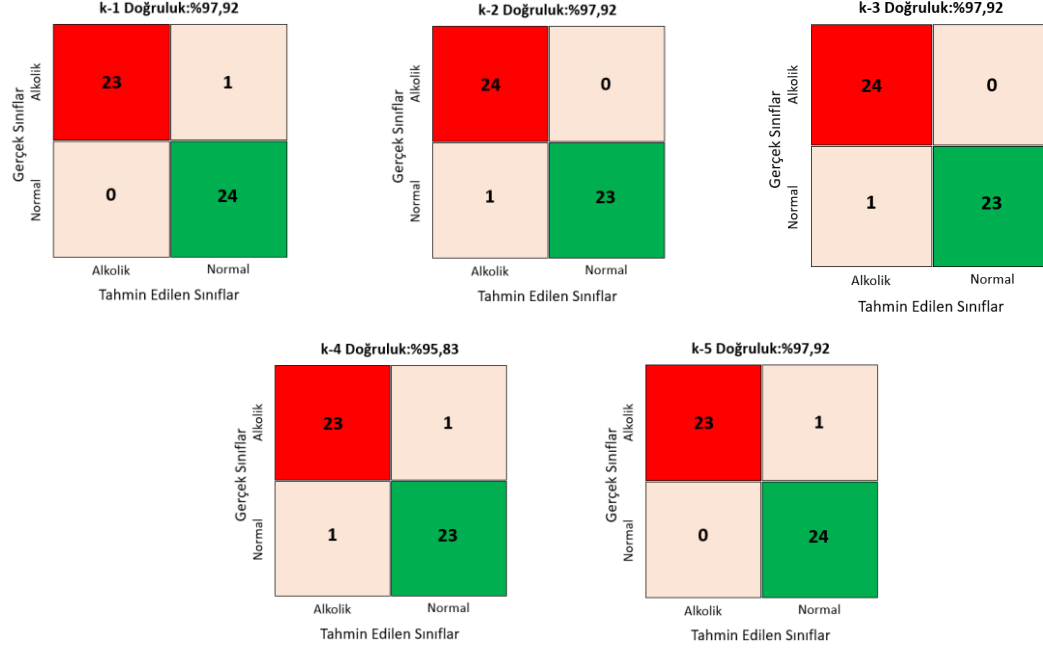
Veri seti ile yapılan deneysel çalışmaların her bir k-kat için sınıflandırma performansları, doğruluk, duyarlılık, özgüllük, kesinlik ve F1 skor sonuçları Tablo 1'de detaylı olarak verilmiştir.

Tablo 1. EBT sınıflandırıcının her katman için sınıflandırma performansı (%)

k-Kat	Doğruluk	Duyarlılık	Özgüllük	Keskinlik	F1 Skor
k-1	97,92	95,83	100,00	100,00	97,87
k-2	97,92	100,00	96,00	95,83	97,96
k-3	97,92	100,00	96,00	95,83	97,96
k-4	95,83	95,83	95,83	95,83	95,83
k-5	97,92	95,83	100,00	100,00	97,87
Genel Ortalama	97,50	97,49	97,56	97,49	97,49

Tablo 1'deki sonuçlarda görüldüğü üzere en yüksek doğruluk oranına k-1, k-2, k-3 ve k-5 katlı doğrulamalarda ulaşılmıştır. Ayrıca diğer performans değerlendirme ölçütleri olan duyarlılıkta k-2, k-3, özgüllük, kesinlik ve F1-skorda k-1, k-5 katlı doğrulamalarda sırasıyla %100, %100 ve %97,87 başarı oranına ulaşılmıştır. Tüm katmanların genel ortalamasında, %97,50 doğrulukta EEG sinyalleri alkolik ve normal olarak sınıflandırılmıştır. Genel ortalama %97,56 özgüllük dikkat çekmektedir.

Önerilen yöntemin genel başarımlarını değerlendirme ile birlikte veri setindeki her bir sınıfa ait daha ayrıntılı performans analizi önerilen yöntemin performans başarımları görebilmek için karmaşıklık matrisleri kullanılmaktadır. Şekil 7’de her bir katmana ait karmaşıklık matrisleri verilmiştir.



Şekil 7. Her k-kat'a ait karmaşıklık matrisleri.

Şekil 7’de EBT sınıflandırıcının katmanlara göre sınıflandırma performanslarına ait karmaşıklık matrislerinde görüldüğü gibi k-1 ve k-5 katlı doğrulamalarda test veri setindeki 24 normal sınıf etkili verinin tamamı doğru olarak tahmin edilmiştir. Aynı şekilde k-2 ve k-3 doğrulamalarında test veri setindeki 24 alkolik sınıf etkili verinin tamamı doğru olarak tahmin edilmiştir.

Önerilen yöntemin daha detaylı performans değerlendirmesi için, aynı veri setlerinin kullanıldığı bazı son teknolojik yöntemlerin performans karşılaştırmaları Tablo 2’de verilmiştir. Tablo 2’deki çalışmalar, doğrusal ve

Tablo 2. Bazı yöntemlerin performans karşılaştırmaları.

Yazar	Model	Doğruluk (%)
Ekaputri vd. (2020) [34]	DVM	77,80
Goksen vd. (2017) [7]	Karşılıklı bilgi işlevsel bağlantı+K-NN	82,33
Malar vd. (2020) [35]	WPT+Aşırı Öğrenme Makinesi	87,60
Khan vd. (2021) [6]	CNN	87,85±4,64
Mumtaz vd. (2017) [8]	Tutarlılık işlevsel bağlantı+Lojistik regresyon	89,30
Kannathal vd. (2005) [36]	En büyük Lyapunov üssü	90,00
Kumari vd. (2022) [37]	CNN	92,70
Farsi vd. (2020) [13]	LSTM	93,00
Bajaj vd. (2017) [9]	Yönlendirilmiş gradyanların histogramları+NNLS	95,83
Shen vd. (2023) [2]	CNN	96,25±3,11
Önerilen Yöntem	WPT+TDD+EBT	97,50

doğrusal olmayan özellik tabanlı yöntemler, sinyal ayrışımına dayalı yöntemler, makine öğrenmesi ve derin öğrenme tabanlı yaklaşımları içermektedir. Önerilen yöntemin mevcut diğer yöntemlerden daha iyi doğruluk performansına sahip olduğu görülebilmektedir. Tablo 2'deki çalışmalara ait sınıflandırma sonuçlarında en yüksek sınıflandırma doğruluğu elde edilmesi önerilen yöntemin başarısını göstermektedir.

4. Sonuçlar

Bu çalışmada, EEG sinyallerinden alkolikleri otomatik olarak belirlemek için WPT yaklaşımı ile tek kanallı EEG sinyalleri frekans alt bantlarına ayrıştırılır. Ardından alt bantlara ayrılan EEG sinyalleri TDD kullanılarak zaman-alan tanımlayıcıları belirlenir ve bu tanımlayıcılar öznitelik olarak kullanılmaktadır. Son olarak alkolik ve normal EEG verilerini içeren sinyallerine ait öznitelikler topluluk öğrenim sınıflandırıcısına girdi olarak uygulanarak sınıflandırılmaktadır. Deneysel çalışmalarda tek kanallı EEG verileri kullanılmıştır. Deneysel çalışmalarda %97,50 doğruluk oranına ulaşılmış olması önerilen yöntemin etkinliğini göstermektedir. Ayrıca önerilen çalışmaya ait deneysel sonuçlar son teknolojik çalışmalar ile karşılaştırıldı. Gelecekte önerilen yöntem alan uzmanların daha iyi tedavi için beynin fizyolojik durumlarını belirlemesine yardımcı olabilmek için farklı EEG veri setlerinde test edilecektir.

Kaynaklar

- [1] Oscar-Berman, M., & Marinković, K. (2007). Alcohol: effects on neurobehavioral functions and the brain. *Neuropsychology review*, 17(3), 239-257.
- [2] Shen, M., Wen, P., Song, B., & Li, Y. (2023). Detection of alcoholic EEG signals based on whole brain connectivity and convolution neural networks. *Biomedical Signal Processing and Control*, 79, 104242.
- [3] Das, D., Zhou, S., & Lee, J. D. (2012). Differentiating alcohol-induced driving behavior using steering wheel signals. *IEEE Transactions on Intelligent Transportation Systems*, 13(3), 1355-1368.
- [4] World Health Organization. (2018). What Quantitative and Qualitative Methods Have Been Developed to Measure Community Empowerment at a National Level? (Vol. 59). World Health Organization.
- [5] Sadiq, M. T., Akbari, H., Siuly, S., Li, Y., & Wen, P. (2022). Alcoholic EEG signals recognition based on phase space dynamic and geometrical features. *Chaos, Solitons & Fractals*, 158, 112036.
- [6] Khan, D. M., Yahya, N., Kamel, N., & Faye, I. (2021). Effective connectivity in default mode network for alcoholism diagnosis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29, 796-808.
- [7] Demir, F., Sengur, A., Ari, A., Siddique, K., & Alswaitti, M. (2021). Feature Mapping and Deep Long Short Term Memory Network-Based Efficient Approach for Parkinson's Disease Diagnosis. *IEEE Access*, 9, 149456-149464.
- [8] Gökşen, N., & Arıca, S. (2017). A simple approach to detect alcoholics using electroencephalographic signals. In *EMBEC & NBC 2017* (pp. 1101-1104). Springer, Singapore.
- [9] Mumtaz, W., Vuong, P. L., Xia, L., Malik, A. S., & Rashid, R. B. A. (2017). An EEG-based machine learning method to screen alcohol use disorder. *Cognitive neurodynamics*, 11(2), 161-171.
- [10] Bajaj, V., Guo, Y., Sengur, A., Siuly, S., & Alcin, O. F. (2017). A hybrid method based on time-frequency images for classification of alcohol and control EEG signals. *Neural Computing and Applications*, 28(12), 3717-3723.
- [11] Fayyaz, A., Maqbool, M., & Saeed, M. (2019, August). Classifying alcoholics and control patients using deep learning and peak visualization method. In *Proceedings of the 3rd International Conference on Vision, Image and Signal Processing* (pp. 1-6).
- [12] Agarwal, S., & Zubair, M. (2021). Classification of Alcoholic and Non-Alcoholic EEG Signals Based on Sliding-SSA and Independent Component Analysis. *IEEE Sensors Journal*, 21(23), 26198-26206.
- [13] Dong, H., Li, T., Ding, R., & Sun, J. (2018). A novel hybrid genetic algorithm with granular information for feature selection and optimization. *Applied Soft Computing*, 65, 33-46.
- [14] Farsi, L., Siuly, S., Kabir, E., & Wang, H. (2020). Classification of alcoholic EEG signals using a deep learning method. *IEEE Sensors Journal*, 21(3), 3552-3560.
- [15] American Electroencephalographic Association 1990 (2007) Standard electrode position nomenclature, <http://kdd.ics.uci.edu/databases/eeg/eeg.data.html/>
- [16] Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of experimental psychology: Human learning and memory*, 6(2), 174.
- [17] Mallat, S. (1989). A Theory for Multiresolution Approximations and Wavelet Orthonormal Bases of l_2 (r). *IEEE Trans. Pattern Recognition and Machine Intelligent*, 11, 674-693.
- [18] Alçin, Ö. F., Budak, Ü., Aslan, M., Akbulut, Y., Cömert, Z., Akpınar, M. H., & Şengür, A. (2020). Classification of physical actions from surface EMG signals using the wavelet packet transform and local binary patterns. In *Modelling*

- and Analysis of Active Biopotential Signals in Healthcare, Volume 1. IOP Publishing.
- [19] Hu, Y., Wong, Y., Wei, W., Du, Y., Kankanhalli, M., & Geng, W. (2018). A novel attention-based hybrid CNN-RNN architecture for sEMG-based gesture recognition. *PloS one*, 13(10), e0206049.
- [20] Wei, W., Wong, Y., Du, Y., Hu, Y., Kankanhalli, M., & Geng, W. (2019). A multi-stream convolutional neural network for sEMG-based gesture recognition in muscle-computer interface. *Pattern Recognition Letters*, 119, 131-138.
- [21] Ari, A. (2020). Analysis of EEG signal for seizure detection based on WPT. *Electronics Letters*, 56(25), 1381-1383.
- [22] Khushaba, R. N., Al-Ani, A., Al-Timemy, A., & Al-Jumaily, A. (2016, December). A fusion of time-domain descriptors for improved myoelectric hand control. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1-6). IEEE.
- [23] Ari, A., Ayaz, F. & Hanbay, D. (2019). EMG sinyallerinin kısa zamanlı fourier dönüşüm özellikleri kullanılarak yapay sinir ağları ile sınıflandırılması. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 31(2), 443-451.
- [24] Al-Timemy, A. H. (2017, October). An investigation of feature combinations of time-domain power spectral descriptors feature extraction for myoelectric control of hand prostheses. In *2017 Fourth International Conference on Advances in Biomedical Engineering (ICABME)* (pp. 1-4). IEEE.
- [25] Hjorth, B. (1970). EEG analysis based on time domain properties. *Electroencephalography and clinical neurophysiology*, 29(3), 306-310.
- [26] Aslan, M., & Zurel, E. N. (2022). An efficient hybrid model for appliances classification based on time series features. *Energy and Buildings*, 266, 112087.
- [27] Breiman, L. (1996). Bagging predictors. *Machine learning*, 24(2), 123-140.
- [28] Freund, Y. R., Schapire, (1995), A decision theoretic generalization of online learning and application to boosting. In *European Conference on Computational Learning Theory* (pp. 23-37).
- [29] PAZOKI, M. A Novel Fault Classification Scheme for Series Capacitor Compensated Transmission Line Based on Bagged Tree Ensemble Classifier.
- [30] Chairatanasongporn, N., & Jaiyen, S. (2015, October). A hybrid ensemble of machine and statistical learning using confidence-based boosting. In *2015 7th International Conference on Information Technology and Electrical Engineering (ICITEE)* (pp. 41-45). IEEE.
- [31] Saeed, M. S., Mustafa, M. W., Sheikh, U. U., Jumani, T. A., & Mirjat, N. H. (2019). Ensemble bagged tree based classification for reducing non-technical losses in multan electric power company of Pakistan. *Electronics*, 8(8), 860.
- [32] ASLAN, M. Derin Öğrenme Tabanlı Otomatik Beyin Tümör Tespiti. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 34(1), 399-407.
- [33] Uzen, H., Turkoglu, M., & Hanbay, D. (2021). Texture defect classification with multiple pooling and filter ensemble based on deep neural network. *Expert Systems with Applications*, 175, 114838.
- [34] Ekaputri, C., Widadi, R., & Rizal, A. (2020, June). EEG signal classification for alcoholic and non-alcoholic person using multilevel wavelet packet entropy and support vector machine. In *2020 8th International Conference on Information and Communication Technology (ICoICT)* (pp. 1-4). IEEE.
- [35] Malar, E., & Gauthaam, M. (2020). Wavelet analysis of EEG for the identification of alcoholics using probabilistic classifiers and neural networks. *International Journal of Intelligence and Sustainable Computing*, 1(1), 3-18.
- [36] Kannathal, N., Acharya, U. R., Lim, C. M., & Sadasivan, P. K. (2005). Characterization of EEG—a comparative study. *Computer methods and Programs in Biomedicine*, 80(1), 17-23.
- [37] Kumari, N., Anwar, S., & Bhattacharjee, V. (2022). A Deep Learning-Based Approach for Accurate Diagnosis of Alcohol Usage Severity Using EEG Signals. *IETE Journal of Research*, 1-15.