



EEG sinyallerini kullanarak Alzheimer hastalığının otomatik tespiti için bilgisayar destekli tanı sistemi

Computer aided diagnostic system for automatic detection of Alzheimer's disease using EEG signals

Zülfikar ASLAN^{1*}

¹Gaziantep Üniversitesi, Teknik Bilimler Meslek Yüksek Okulu, zulfikaraslan@gantep.edu.tr

MAKALE BİLGİLERİ

Makale Geçmişi:

Geliş 24 Mart 2022
Revizyon 13 Haziran 2022
Kabul 14 Haziran 2022
Online 28 Haziran 2022

Anahtar Kelimeler:

Dalgacık Dönüşümü, EEG,
Alzheimer Teşhisi, Sınıflandırma

ÖZ

Alzheimer beyindeki bozulmalardan kaynaklı bilişsel ve davranışsal eksiklikler gibi semptomlarla kendini gösteren önemli bir nörolojik hastalıktır. Alzheimer hastalığının kesin bir tedavi yöntemi bulunmamaktadır. Ancak hastalığın erken teşhisi ile hastalığın ilerlemesinin yavaşlatılması amaçlanmaktadır. Bu durum hastanın yaşam standartlarının korunmasında önem arz etmektedir. Ayrıca hastalığın tam olarak teşhisi deneyimli bir uzman tarafından değerlendirilecek olan maliyetli testler ve yorucu bir teşhis aşaması gerektirmektedir. Bu motivasyonla önerilen yöntemle Alzheimer hastalığının EEG sinyallerinden otomatik olarak gerçekleştirilmesini amaçlayan yeni bir bilgisayar destekli tanı sistemi sunulmaktadır. Sunulan çalışmada öncelikle ham EEG verilerine önişlem uygulanarak var olan gürültüler giderilmiştir. Sonraki aşamada ise her bir kanaldan alınan verilere dalgacık dönüşümü uygulandıktan sonra istatistiksel özellikler hesaplanmıştır. Elde edilen özelliklerin k-en yakın komşu (kNN) sınıflandırıcısı ile sınıflandırılmasıyla sağlıklı katılımcılar ile Alzheimer hastası katılımcılar 91.12% doğrulukla ayırt edilmiştir.

ARTICLE INFO

Article history:

Received 24 March 2022
Received in revised form 13 June 2022
Accepted 14 June 2022
Available online 28 June 2022

Keywords:

Wavelet Transform, EEG, Diagnosis
of Alzheimer's, Classification

ABSTRACT

Alzheimer's is an important neurological disease that manifests itself with symptoms such as cognitive and behavioral deficits caused by deterioration in the brain. There is no definitive treatment for Alzheimer's disease. However, with early diagnosis of the disease, it is aimed to slow down the progression of the disease. This is important in maintaining the patient's living standards. In addition, a full diagnosis of the disease requires costly tests and a grueling diagnostic phase, which will be evaluated by an experienced specialist. With this motivation, a new computer aided diagnosis system is presented, which aims to automatically detect Alzheimer's disease from EEG signals with the proposed method. In the presented study, first of all, the existing noises were removed by preprocessing the raw EEG data. In the next step, statistical properties were calculated after applying wavelet transform to the data received from each channel. By classifying the obtained features with the k-nearest neighbor (kNN) classifier, healthy participants and Alzheimer's patients were distinguished with an accuracy of 91.12%.

Doi: 10.24012/dumf.1092569

* Sorumlu Yazar

Giriş

Alzheimer hastalığı dünya genelinde yaygın olarak görülen ve yaşa bağlı bunamanın en yaygın şekli olan önemli bir nörolojik rahatsızlıktır. Alzheimer hastalarında beyin aktiviteleri sağlıklı insanlara göre daha yavaştır ve en önemli semptomu bilişsel fonksiyonların bozulmasıdır [1], [2]. 2018 yılında dünya genelinde 50 milyon kişi bu hastalıktan etkilenirken 2030 yılında bu rakamın 82 milyon olması ve zamana bağlı olarak artış göstermesi öngörülmektedir [3]. Alzheimer hafif, orta ve şiddetli olmak üzere üç çeşit olarak görülmektedir. Alzheimer hastalığına sahip olan bir hasta hastalığın her bir aşamasında artan şekilde bilişsel fonksiyonlarını zamanla kaybeder. Alzheimer'ın erken teşhisi hastalığın giderek hızlı bir şekilde ilerlemesini önlemek için önem arz etmektedir. Bu nedenle Alzheimer tespiti için ilgili literatürde birçok çalışma bulunmaktadır.

Elektroensefalografi (EEG) sinyalleri beyin fonksiyonlarının incelenmesinde yaygın olarak kullanılan önemli bir araçtır. EEG sinyallerinin diğer biyomedikal görüntüleme (CT- Bilgisayarlı Tomografi, MRI- Manyetik Rezonans Görüntüleme, PET- Pozitron Emisyon Tomografi, fMRI - Fonksiyonel Manyetik Rezonans Görüntüleme gibi) yöntemlerine kıyasla daha düşük maliyet ve daha düşük zaman gerektirmesinden birçok hastalığın teşhisini amaçlayan bilgisayar destekli tanı sistemlerinin geliştirilmesinde yaygın olarak tercih edilmektedir.

Makine öğrenmesi birçok alanda kullanılmış ve tatmin edici sonuçlar elde edilmiştir. Örneğin çevre alanında bir çalışmada Buyrukoğlu ve arkadaşları, tarımsal yüzey sularında Salmonella varlığının öngörülmesi için belirlenen korelasyon değerinin başarısını araştırmışlardır [25]. Spor alanında bir çalışma da ise araştırmacılar futbolcuların pozisyonlarını belirlemek için yığılmış topluluk öğrenme modeli önermişler ve 83.9% sınıflandırma başarımı elde etmişlerdir [26]. Eğitim alanında gerçekleştiren bir başka çalışmada ise akademik enstitüler tarafından toplanan büyük miktarda veri üzerinde farklı Makine Öğrenimi algoritmalarının uygulanmasına ilişkin bir literatür taraması sunulmuştur [27].

Literatürde Alzheimer hastalığını ayırt etmek amacıyla yukarıda bahsedilen biyomedikal görüntüleme yöntemleri dışında EEG sinyallerinin analiz edilerek ve makine öğrenmesi algoritmasıyla sınıflandırılmasıyla gerçekleştirilen bir çok literatür çalışması bulunmaktadır [1], [4]-[9].

Kulkarni ve arkadaşları [4], çalışmada Alzheimer'ı teşhis etmek amacıyla EEG sinyallerinden farklı özellik (Spektral özellikler, dalgacık dönüşümü özellikleri, karmaşıklık özellikleri) çıkarım yöntemleri gerçekleştirmiştir. Elde ettiği özellikleri Destek Vektör Makinaları (SVM) sınıflandırıcısı ile sınıflandırarak sırasıyla spektral özellikler için 86%, dalgacık dönüşümü özellikleri için 88% ve karmaşıklık özellikleri için 96% sınıflandırma başarımı elde etmiştir. Yine benzer bir çalışmada Bairagi [1], EEG sinyallerinden spektral tabanlı özellikler ve dalgacık dönüşümü tabanlı özellikler

elde etmişlerdir. Bu özellikleri iki farklı sınıflandırıcı (SVM ve kNN) ile sınıflandırmışlardır. Çalışmalarında en yüksek sınıflandırma performansını spektral tabanlı özelliklerin SVM algoritması ile sınıflandırılmasıyla 94% olarak elde etmişlerdir. Ruiz-Gómez ve arkadaşları [5], Alzheimer tespiti için EEG sinyallerinden elde ettikleri spektral ve lineer olmayan özellikleri birleştirerek elde ettikleri özellikleri farklı makine öğrenme algoritmaları ile sınıflandırmışlardır. Çalışmalarında Alzheimer ve sağlıklı kontrolleri en yüksek Çok Katmanlı Algılayıcı (MLP) ve Kuadratik Diskriminant Analizi (QDA) algoritmaları ile 78.43% olarak elde etmişlerdir. Amezcua-Sanchez ve arkadaşları [6] Alzheimer teşhisini amaçlayan çalışmalarında EEG sinyallerine çoklu sinyal sınıflandırması and ampirik dalgacık dönüşümü (MUSIC-EWT) adını verdikleri yöntemlerini uygulamışlardır. Sonrasında ise Hurst üs ölçüleri ve Fraktal boyut özelliklerini hesaplamışlardır. Bu özellikleri Gelişmiş olasılıksal sinir ağı (EPNN) sınıflandırıcı ile sınıflandırarak 90.3% doğruluk elde etmişlerdir. Kulkarni bir çalışmada [7], EEG sinyalinden elde ettiği spektral ve karmaşıklık özelliklerini birleştirmişlerdir. Sonrasında ise bu özellikler kNN sınıflandırıcısı kullanarak sınıflandırmış ve Alzheimer teşhisini en yüksek 94% doğrulukla tespit etmişlerdir. Tzamourta ve arkadaşları [8] çalışmalarında EEG sinyallerinden hesapladıkları istatistiksel ve spektral özellikleri farklı sınıflandırma algoritmalarıyla test etmişlerdir. En yüksek sınıflandırma performansını Rastgele Orman algoritması ile 88.79% olarak elde etmişlerdir. Buyrukoğlu çalışmada Normal, Hafif Bilişsel Bozukluk ve Alzheimer hastalığı durumlarını sınıflandırmak için bir yöntem önermişler ve 91% sınıflandırma başarımı elde etmişlerdir [28]. Buyrukoğlu diğer bir çalışmada ise Normal, Hafif Bilişsel Bozukluk ve Alzheimer hastalığı durumlarını AdaBoost sınıflandırıcısı ile sınıflandırarak başarılı bir sınıflandırma performansı sergilemişlerdir [29].

EEG sinyallerinin analizi için çeşitli algoritmalar ve yeni işlem teknikleri sunulmaktadır. EEG sinyallerinden Alzheimer tespiti gerçekleştiren CAD sistemlerinin performansını etkileyen sınırlamalar bulunmaktadır. İlgili literatür çalışmalarında uygun özellik seçim yöntemlerinin belirlenememesi veya önemli bilgiler taşıyan EEG kanallarının dikkate alınmaması bu sınırlamalar için önemli unsurlardır.

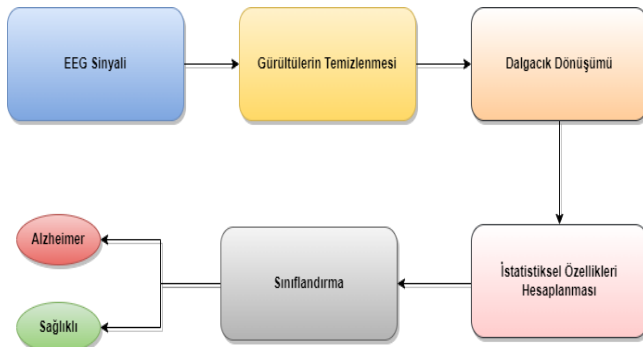
EEG tabanlı CAD uygulamalarında sinyalden doğrudan özellik çıkarımı gerçekleştirmek amacıyla çeşitli analiz yöntemleri kullanılmaktadır. Dalgacık dönüşüm yönteminin sinyalin zaman frekans uzayında önemli bilgiler içeren frekans bilgilerinin analizinde kullanılan önemli bir araçtır [10]. Diğer yandan son zamanlarda birçok EEG tabanlı teşhis gerçekleştiren CAD sistemlerinde Hjorth parametrelerinin [11] kullanımı tercih edilmiş ve yüksek sınıflandırma performansı sağlanmıştır. Bunun yanında hesaplama karmaşıklığı dalgacık dönüşümünün diğer sinyal ayrıştırma yöntemlerine ve Hjorth parametrelerinin diğer özellik çıkarım yöntemlerine göre çok daha düşüktür.

Bilinen literatür çalışmalarında EEG sinyallerinde Alzheimer tespitini dalgacık dönüşümü ve Hjorth

parametreleri özelliklerinin birlikte kullanılmasıyla gerçekleştiren bir çalışmaya rastlanmamıştır. Bu motivasyonla EEG sinyallerinden otomatik olarak Alzheimer teşhisini yüksek doğrulukla gerçekleştirebilen, sinyalin zaman-frekans ve genlik uzayında analizini sağlayan bir CAD sistemi önerilmektedir.

Materyal Metot

Bu bölümde EEG veri setinin tanımı, ön işlem aşamaları ve sunulan analiz yönteminin adımları sunulmaktadır. Şekil 1' de EEG sinyallerinden otomatik AD tespiti için önerilen yöntemin akış diyagramı sunulmaktadır. Bu çalışma şu aşamaları içermektedir: (i) EEG sinyalinden gürültülerin Çok Ölçekli Temel Bileşen Analizi (MSPCA) yöntemiyle temizlenmesi. (ii) Her bir EEG kanalından alınan verilerin Dalgacık Dönüşümü yönteminin uygulanması. (iii) İstatistiksel özelliklerin hesaplanması. (iv) Önerilen yöntemin doğruluğunun kNN algoritması ile sınıflandırılarak test edilmesi. (vi) Dalgacık dönüşümü ile elde edilen frekans bantlarının Alzheimer hastalığının ayırt edilmesindeki sınıflandırma performansının test edilmesi. Çalışmanın tüm deneysel aşamaları Matlab yazılımı kullanılarak gerçekleştirilmiştir.



Şekil 1. Önerilen yöntemin aşamalarını içeren akış diyagramı

Veri Seti

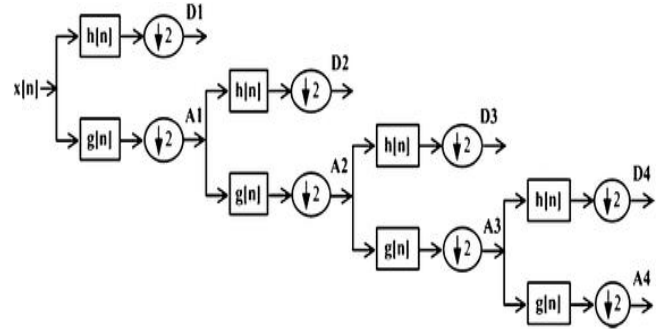
Bu çalışmada kullanılan EEG veri seti Florida Eyalet Üniversitesi araştırmacıları tarafından uluslararası 10-20 sistemine göre Biologic Systems Brain Atlas III Plus workstation kullanan 19 elektrota sahip bir kaydedici ile kaydedilmiştir. Kayıtlar dört grubu (A, B, C, D) içeren kayıtlar iki dinlenme durumu altında gerçekleştirilmiştir. A ve C gruplarının gözleri görsel olarak sabitlenerek gözleri açık olarak kaydedilmiştir. B ve D grupları ise gözleri kapalı olarak kaydedilmiştir. A ve B grupları 24 sağlıklı yaşlıdan oluşan (yaşları 61-83 aralığında ve ortalama yaşı 72) ve herhangi bir nörolojik rahatsızlık geçirmeyen deneklerden oluşmaktadır. C ve D grupları ise Ulusal Nörolojik ve İletişimsel Bozukluklar ve İnme Enstitüsü (NINCDS) ve Alzheimer Hastalığı ve İlişkili Bozukluklar Derneği (ADRDA) tarafından teşhis edilen 24 olası AD hastasından (yaşları 53-85 yaş aralığında ve ortalama yaş 69) oluşmaktadır. EEG kayıtları 1-30Hz frekans aralığında sınırlandırılmış ve 8 saniye süreli kayıtlardan oluşmaktadır. Kayıtların örnekleme frekansı 128Hz alınarak gerçekleştirilmiştir. Hastanın uyanıklık durumunun kontrol edilmesi amacıyla kayıt süresince bir

teknisyen görevlendirilmiştir. Bu EEG veri seti açık kaynak olarak erişime sunulmuştur [12].

MSPCA, Temel Bileşen Analizi (PCA) ve Dalgacık Dönüşümü yöntemlerinin önemli özelliklerini bir araya getiren bir analiz yöntemidir. MSPCA değişen sinyalin analizi veya gürültülerin temizlenmesi amacıyla çok ölçekli yapısı nedeniyle tercih edilmektedir [13]. EEG kaydı süresince kas, göz hareketi, elektrik hattı ve diğer cihazlarla etkileşimden kaynaklanan gürültüler oluşabilmektedir. Önerilen çalışmada EEG üzerindeki gürültülerin giderilmesi amacıyla her bir elektrottan alınan EEG verilerine MSPCA yöntemi uygulanmıştır. MSPCA yöntemi uygulanırken dalgacık dönüşümü fonksiyonu 'sym4' ve seviye 5 alınarak parametreler belirlenmiştir.

Dalgacık Dönüşümü

Dalgacık dönüşümü, biyomedikal sinyallerin [10], [14] zaman-frekans analizi için, özellikle durağan olmayan özelliklerinden dolayı EEG sinyal analizinde yaygın olarak kullanılmaktadır. Dalgacık dönüşümü, düşük frekanslar için geniş zaman pencereleri ve daha yüksek frekanslar için kısa zaman pencereleri kullanır, bu da iyi bir zaman-frekans analizi sağlar. Bir sinyalin dalgacık dönüşüm ayrıştırması, zaman serisinin ve 2 aşağı örnekleyicinin ardışık yüksek geçiş ve düşük geçiş filtrelemesini kullanır. Yüksek geçiş filtresi $g(n)$ ayrık ana dalgacıktır ve düşük geçiş filtresi $h(n)$ onun aynasıdır. İlk yüksek geçiş ve düşük geçiş filtrelerinin çıktısı, sırasıyla A1 ve D1 ile temsil edilen yaklaşıklık ve ayrıntılı katsayılar olarak adlandırılır. A1 daha da parçalanır ve prosedür belirtilen sayıda ayrışma seviyesine ulaşılan kadar tekrarlanır (Bakınız Şekil 2) [14], [15].



Şekil 2. Dalgacık dönüşümü ayrıştırma süreci [15]

Genişletme işlevi $\phi_j, k(n)$ düşük geçiş filtresine bağlıdır ve dalgacık işlevi $\psi_j, k(n)$ aşağıdaki gibi gösterilen yüksek geçiş filtresini takip eder.

$$\phi_j, k(n) = 2^{j/2} h(2^j n - k) \quad (1)$$

$$\psi_j, k(n) = 2^{j/2} g(2^j n - k) \quad (2)$$

burada $n = 0, 1, 2, \dots, M-1$; $j = 0, 1, 2, \dots, J-1$; $k = 0, 1, 2, \dots, 2^j - 1$; $J = \log_2(M)$; ve M , sinyalin uzunluğudur [16].

Maksimum ayrışma seviyesi, verilen sinyaldeki temel frekans bileşenlerine bağlı olarak belirlenir. Dalgacık Dönüşümünün katsayıları, orijinal zaman serisinin nokta çarpımı ve belirlenmiş temel fonksiyonları olarak adlandırılır. Yaklaşık katsayılar A_i (Denklem 3) ve i .

seviyedeki ayrıntılı katsayılar D_i (Denklem 4) olarak gösterilir:

$$A_i = \frac{1}{\sqrt{M}} \sum_n x(n) \times \varphi_j, k(n) \quad (3)$$

$$D_i = \frac{1}{\sqrt{M}} \sum_n x(n) \times \psi_j, k(n) \quad (4)$$

Burada $k = 0, 1, 2, \dots, 2j-1$ ve M , ayrıık noktalarda EEG zaman serisinin uzunluğudur [10]. Sunulan çalışmada EEG sinyallerine daubechies (db8) ana dalgacık fonksiyonu uygulanarak dalgacık dönüşümü işlemi gerçekleştirilmiştir.

İstatistiksel Özelliklerin Hesaplanması

Hjorth parametreleri, 1970 yılında zaman alanındaki EEG sinyalini tanımlamak için üç parametreden oluşan bir set sundu [11]: Aktivite, sinyal gücüdür (geniş bant filtreli); Hareketlilik, ortalama frekanstır ve Karmaşıklık, frekanstaki değişimdir. Dalgacık dönüşümü uygulanan her bir EEG kanalından alınan sinyaller üzerinden Hjorth parametreleri Alzheimer ve sağlıklı deneklerin ayırt edilmesinde özellik olarak kullanılmıştır.

Aktivite: Aktivite parametresi, bir zaman fonksiyonunun sinyal gücünü ve varyansını gösteren parametredir. Frekans alanındaki güç spektrumunun yüzeyini gösterebilir. Aşağıdaki Denklem (5) aktiviteyi temsil eder:

$$Activity = var(x(t)) \quad (5)$$

Hareketlilik: Spektrum gücünün ortalama frekansı veya standart sapma oranı, hareketlilik parametresi [15] ile temsil edilir. Hareketlilik, $x(t)$ sinyalinin birinci türevinin varyansının karekökünün, $x(t)$ sinyalinin varyansına oranı olarak tanımlanır ve denklem (6)'da sunulur:

$$Mobility = \sqrt{var(dx(t)/dt)^{-1} var(x(t))^{-1}} \quad (6)$$

Karmaşıklık: Karmaşıklık parametresi frekanstaki değişimi temsil eder. Sinyalin ilk türevinin hareketliliğinin sinyal hareketliliğine oranına karmaşıklık denir. Denklem (7)'de verilen:

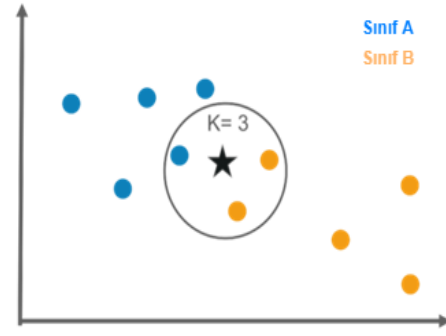
$$Complexity = (\sigma_s'' (\sigma_s')^{-1}) (\sigma_s' (\sigma_s)^{-1})^{-1} \quad (7)$$

k-En Yakın Komşu Algoritması

k-En Yakın Komşu sınıflandırıcısı [17], belirli bir veri noktasını komşularının çoğuna göre sınıflandıran parametrik olmayan bir yaklaşımdır. KNN algoritması, yürütmesini iki adımda tamamlar, ilk önce en yakın komşuların sayısını bulur ve ikincisi, birinci adımı kullanarak veri noktasını belirli bir sınıfa sınıflandırır. Komşuyu bulmak için denklem 8'de verilen Öklid mesafesi gibi uzaklık ölçütlerini kullanır.

$$Distance(x, y) = \sqrt{\sum_i (x_i - y_i)^2} \quad (8)$$

Eğitim setinden en yakın k örneğini seçer, daha sonra belirsizliği önlemek için k 'nin tek bir sayı olması gereken sınıflarının çoğunluk oyu alır. Şekil 3, kNN sınıflandırıcısının mimarisini göstermektedir. Sınıf A ve sınıf B olmak üzere sınıflar vardır. Mavi daireler sınıf A'ya ve sarı daireler sınıf B'yi gösterir. k değeri 3 olarak seçilmiştir ve en yakın 3 komşu arasından 1 örnek sınıf A'ya ve 2 örnek sınıf B'ye aittir. kNN sınıflandırıcısı, tanımlanan k 'da oyların çoğunluğu ile sınıfa yeni bir örnek verme ilkesine göre çalışır. Böylece yeni test örneği, sınıf 1'e atanır.



Şekil 3. kNN sınıflandırıcısının mimarisini

Bu bölümde önerilen yöntemle hesaplanan özelliklerin kNN öğrenme algoritması ile sınıflandırılması işlemi gerçekleştirilmiştir.

Sınıflandırma için kullanılan k değerleri deneysel olarak ($k=1 \dots 50$ arası değerler için) test edilmiş ve en yüksek sınıflandırma doğruluğu $k=3$ değeri için 91.12% olarak elde edilmiştir. Sınıflandırma sürecinde çapraz doğrulama 10 alınarak gerçekleştirilmiştir. En yakın komşu arama algoritması olarak doğrusal arama algoritması kullanılmış, mesafe fonksiyonu olarak da Öklid mesafe fonksiyonu kullanılmıştır.

Önerilen yöntemde kNN sınıflandırıcısı, uygulaması basit ve anlaşılması sezgisel olması, sınıflandırma için eğitim süresi gerektirmemesi, k değerinin ayarlanabilir olması, sürekli olarak yeni verilerle gelişmeye açık olması, tek bir hiper parametreye sahip olması, mesafe fonksiyonunun seçilebilir olması yönleriyle tercih sebebi olmuştur.

Değerlendirme Metrikleri

Sunulan çalışmada sınıflandırıcı performansının değerlendirilmesi amacıyla gerçek pozitifler (TP rate), yanlış pozitifler (FP rate), hassasiyet (precision), geri çağırma (recall), f-ölçütü, matthews korelasyon katsayısı (MCC), ROC alanı, doğruluk ve karmaşıklık matrisi olmak üzere toplam dokuz değerlendirme metriği hesaplanmıştır. Sınıflandırma performansının görselleştirilmesi amacıyla duyarlılık ve özgüllük değerleri arasındaki ROC alanı çizdirilmiştir. ROC alanının yüksek olması sınıflandırıcısının sınıfları doğru olarak tahmin ettiğini göstermektedir. Yani bir sınıflandırıcının hasta ve sağlıklı bireyin ayırt edilmesinde yüksek bir doğruluk sağlaması ROC eğrisi altındaki alanın yüksek olmasını gerektirir. Bu nedenle çalışmada önerilen yöntemin doğruluğunu açıkça ortaya koymak adına ROC eğrisi çizdirilerek ROC altındaki alan (AUC-ROC) değerlendirme metrikleri hesaplanmıştır. Karmaşıklık matrisinde bulunan doğru negatifler (TN), gerçek pozitifler (TP), yanlış pozitifler (FP) ve yanlış negatifler (FN) parametreleri kullanılarak değerlendirme metrikleri hesaplanabilmektedir. Gerçek sınıflara karşılık doğru veya yanlış tahminlerin sayısal olarak ifade edilmesi amacıyla karmaşıklık matrisi kullanılmaktadır. Önerilen yöntemde kullanılan değerlendirme metriklerine ait hesaplamalar aşağıdaki denklemlerde gösterilmektedir.

$$Doğruluk = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (9)$$

$$Geri \text{ Çağırma} = \frac{TP}{TP+FN} \times 100 \quad (10)$$

$$\text{Özgüllük} = \frac{TN}{TN+FP} \times 100 \quad (11)$$

$$Hassasiyet = \frac{TP}{TP+FP} \times 100 \quad (12)$$

$$F1 \text{ Ölçütü} = 2 * \frac{\text{hassasiyet} * \text{geri çağırma}}{\text{hassasiyet} + \text{geri çağırma}} \times 100 \quad (13)$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP) + (TN + FN)}} \quad (14)$$

Karmaşıklık matrisi (Bakınız Şekil 4) hata matrisi olarak bilinmektedir. Matrisin her satırı tahmin edilen bir sınıftaki örnekleri temsil ederken her sütun ise gerçek bir sınıftaki örnekleri temsil etmektedir. Sınıflandırma işlemlerinde yoğun olarak kullanılan karmaşıklık matrisi, sınıflandırmanın doğruluğunu test etmek amacıyla kullanılan bir performans değerlendirme metriğidir.



Şekil 4. Karmaşıklık matrisi

Deneysel Sonuçlar

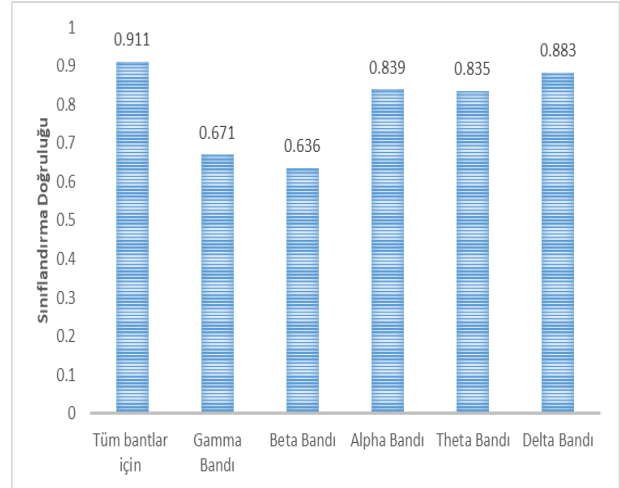
Her bir EEG kanalından alınan verilere öncelikle MSPCA yöntemi uygulanarak gürültüler temizlenmiştir. Sonrasında dalgacık dönüşümü yöntemiyle ile sinyal alt bantlarına ayrılmıştır. Sonrasında her bir alt banttan istatistiksel özellik çıkarımı gerçekleştirilmiştir. Çalışmanın ilk aşamasında her bir alt banda ait özellikler ayrı ayrı değerlendirilmiştir. Bu doğrultuda her deneğe ait her bir kanaldan alınan veri için toplam 3 özellik elde edilmiştir. Her bir Alzheimer hastası ve sağlıklı kontrol için (19 kanal sayısı x 3 her bir kanal için elde edilen özellik sayısı) uzunluklu bir vektör elde edilmiştir. Sonuçta tüm denekler için toplam uzunluğu (48 denek sayısı x 19 kanal sayısı) olmak üzere 912x3 boyutuna sahip bir matris elde edilmiştir. Tüm frekans bantlarından elde edilen bu özelliklerin kNN algoritması ile sınıflandırılmasıyla 91.12% sınıflandırma doğruluğu elde edilmiştir.

Tablo 1'de tüm alt bantlara ait özelliklerin ve her bir frekans bandına ait özelliklerin kNN algoritması ile sınıflandırılmasıyla elde edilen değerlendirme metrikleri gösterilmektedir.

Tablo 1. kNN algoritması ile gerçekleştirilen sınıflandırma sonucunda elde edilen değerlendirme metrikleri

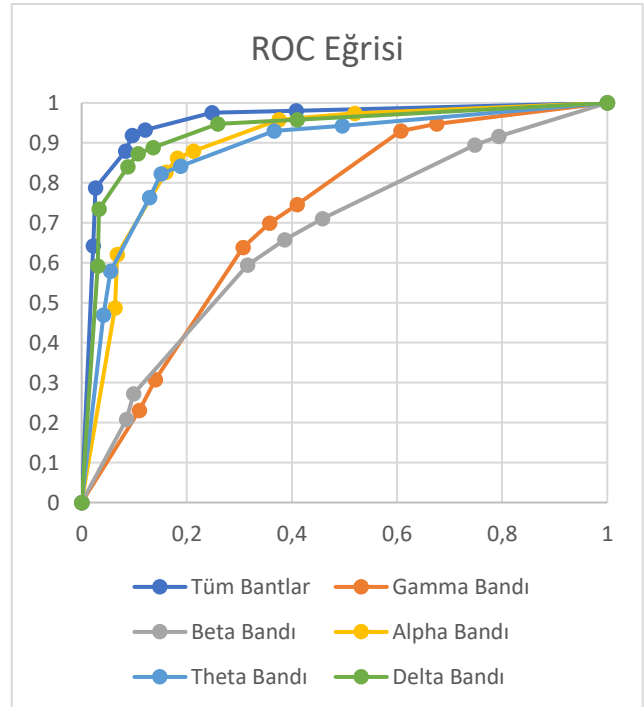
	TP Oran	FP Oran	Hassasiyet	Geri Çağırma	F-Ölçütü	MCC	ROC Alanı	Doğruluk
Tüm bantlar için	0.911	0.089	0.911	0.911	0.911	0.822	0.956	0.911
Gamma Bandı	0.671	0.329	0.672	0.671	0.671	0.343	0.717	0.671

Beta Bandı	0.636	0.364	0.636	0.636	0.636	0.272	0.671	0.636
Alpha Bandı	0.840	0.160	0.841	0.840	0.840	0.680	0.893	0.839
Theta Bandı	0.836	0.164	0.836	0.836	0.835	0.671	0.884	0.835
Delta Bandı	0.883	0.117	0.883	0.883	0.883	0.765	0.930	0.883



Şekil 5. Frekans bantlarından ve tüm frekans bileşenlerine ait özelliklerden elde edilen doğruluk değerleri

Şekil 5'de önerilen yöntemle, tüm frekans bileşenlerine ait özelliklerin ve frekans bantlarına ait özelliklerin kNN sınıflandırıcı kullanılarak sınıflandırılmasıyla elde edilen sınıflandırma doğruluklarının grafiği gösterilmektedir.



Şekil 6. Sınıflandırma işlemi sonucu elde edilen ROC eğrisi

Literatürde sınıflandırma performansını değerlendirmek için kullanılan en yaygın metriklerden biride ROC eğrisidir. Bu çalışmada önerilen yöntemin sınıflandırma doğruluğunu daha net olarak ifade edebilmek adına Şekil 6'da önerilen yöntem ile farklı frekans bantları için elde edilen ROC eğrisi oluşturulmuştur. Şekil 6'daki ROC eğrisi kNN sınıflandırıcısı ile sınıflandırma sonucu elde edilen metrikler göre oluşturulmuştur. Şekil 6 incelendiğinde tüm bantlardan alınan özelliklerin sınıflandırılmasıyla elde edilen sınıflandırma performansı frekans bantları düzeyinde elde edilen sınıflandırma performansına göre daha yüksek bir AUC-ROC değerine (95.6%) sahip olduğu görülmektedir.

		Gerçek Değerler	
		Sağlıklı	Alzheimer
Tahmin Değerler	Sağlıklı	419 (91.88%)	37 (8.12%)
	Alzheimer	44 (9.65%)	412 (90.35%)

Şekil 7. Önerilen yöntemle elde edilen karmaşıklık matrisi

Çalışmada ayrıca önerilen yöntemle elde edilen karmaşıklık matrisi de değerlendirilmiştir ve sonuçlar Şekil 7'de gösterilmektedir. Önerilen yöntemle sağlıklı kontrollerin doğru tahmini 91.88% iken Alzheimer hastalarının doğru tahmini 90.35% olarak elde edilmiştir. Aynı şekilde sağlıklı kontrollerin yanlış tahminin 8.12% iken Alzheimer hastalarının yanlış tahmini 9.65% olarak hesaplanmıştır. Bu durumda gerçekleştirilen deneysel çalışmalarda tüm değerlendirme metriklerinin oldukça yüksek bir değere sahip olması önerilen yöntemin Alzheimer hastaları ve sağlıklı kontrollerin ayırt edilmesinde başarılı olduğunu ortaya koymaktadır. Frekans bantlarına göre ise sınıflandırma başarımları incelendiğinde en yüksek sınıflandırma performansının delta bandı ile elde edildiği görülmektedir. Bu durum Alzheimer hastası ve sağlıklı bireylerin ayırt edilmesinde düşük frekans bileşenlerinin daha anlamlı bilgiler taşıdığını açıkça göstermektedir.

Tartışma

Sağlık sektöründeki olumlu gelişmeler dünya nüfusunun daha sağlıklı ve uzun bir hayat sürmesine imkân sağlamaktadır. Ancak demans gibi istenmeyen hastalıklarının sayısının artması istenmeyen sonuçlara sebep olmaktadır [5], [6], [18]–[22]. Alzheimer hastalığının henüz tam kesin bir tedavisi olmamakla beraber hastalığın erken teşhisi süreci yavaşlatarak yaşam kalitesinin artırılması sağlanabilir [23]. Hastalığın teşhisi nörolojik ve psikiyatrik muayene, kan analizi, bilgisayarlı tomografi gibi bir çok tıbbi işlem gerektirmesinin yanında bu işlemlerden elde edilen sonuçların uzman tarafından kapsamlı olarak değerlendirilmesi gerekmektedir [21]. Bütün bu zorlu teşhis süreçlerine ek olarak deneyimli

uzmanların bile vakaların 10-15% 'inde başarısız olduğu belirtilmektedir [22], [24]. Tüm bu nedenlerden dolayı sadece biyomedikal sinyallerin veya görüntüler kullanılarak CAD sistemlerin geliştirilmesi önem arz etmektedir. Biyomedikal görüntüleme (CT, fMRI, PET gibi) yöntemlerin yüksek maliyet ve zaman gerektirmesinin yanında EEG sinyallerinin daha düşük maliyet ve zaman gerektirmesi çalışmamızda tercih sebebi olmuştur.

Tablo 2'de EEG sinyalleri kullanarak Alzheimer teşhisi gerçekleştiren ilgili literatür çalışmaları gösterilmektedir.

Kulkarni ve arkadaşları [4], sundukları çalışmada Alzheimer hastalığını teşhis etmek amacıyla EEG sinyallerinden farklı özellik değerleri hesaplayarak elde ettikleri özellikleri Destek Vektör Makinaları (SVM) sınıflandırıcısı ile sınıflandırmışlardır. Sundukları çalışmada en yüksek sınıflandırma karmaşıklık özellikleri ile 96% olarak elde etmişlerdir. Bir başka çalışmada ise Bairagi [1], farklı özellik değerlerini SVM ve KNN sınıflandırıcısı ile sınıflandırmıştır. Çalışmalarında en yüksek performansı 94% olarak hesapladıkları spektral özellikleri SVM sınıflandırıcısı ile sınıflandırarak elde etmişlerdir. Ruiz-Gómez ve arkadaşları [5], önerdikleri çalışmada farklı özellik değerlerini birleştirerek bunları farklı sınıflandırıcılar ile test etmişlerdir. Önerdikleri yöntemle MLP ve QDA algoritmaları ile 78.43% sınıflandırma başarımları gerçekleştirmişlerdir. Bir diğer çalışmada ise Amezcua-Sanchez ve arkadaşları [6], sundukları MUSIC-EWT isimli yöntemleriyle gerçekleştirdikleri sinyal analizi sonucunda Hurst üs ölçüleri ve Fraktal boyut özelliklerini hesaplayarak hastalığı en yüksek EPNN sınıflandırıcı ile 90.2% doğrulukla tespit etmişlerdir. Kulkarni bir diğer çalışmasında [7], EEG sinyalinden hesapladıkları spektral ve karmaşıklık özellikleri birleştirerek bunları kNN algoritması ile sınıflandırarak Alzheimer hastası ve sağlıklı kontrolleri 94% doğrulukla ayırt edebilmiştir. Tzamourta ve arkadaşları [8], sundukları çalışmalarında en yüksek sınıflandırma başarımlarını hesapladıkları istatistiksel ve spektral özellikleri Rastgele Orman algoritması ile sınıflandırarak 88.79% olarak elde etmişlerdir.

Yukarıda bahsedilen makine öğrenme tabanlı Alzheimer teşhisi gerçekleştiren çalışmaların tamamında fazla istatistiksel özellikler ve yüksek hesaplama karmaşıklığına sahip yöntemler önerilmiştir. Ancak bu çalışmada sadece üç istatistiksel özelliğin (Hjorth Parametreleri) kNN sınıflandırıcısı (basit ağ yapısına sahip olması ve eğitim gerektirmemesinden dolayı) ile sınıflandırılmasıyla oldukça tatmin edici sınıflandırma başarımları elde edilmiştir.

Çalışmanın sonraki çalışmalarda telafi edilebileceği sınırlamalarında en önemlisi farklı ensemble algoritmaların tercih edilerek sınıflandırma performansının daha fazla artırılabilir olmasıdır. Ancak çalışmada özellikle kNN sınıflandırıcısının tercih edilmesinin başlıca nedeni eğitim aşamasına ihtiyaç duymaması ve daha az hesaplama karmaşıklığına sahip olmasıdır.

Dalgacık dönüşümü literatürde üstün sinyal analiz yeteneklerinden dolayı EEG sinyallerinin zaman-frekans analizinde sıkça kullanılmaktadır. Bu çalışmada EEG sinyallerinden Alzheimer hastası ve sağlıklı kontrollerin yüksek doğrulukla ayırt edilebilecek, dalgacık dönüşümü ve Hjorth parametrelerini birlikte kullanan bir CAD sistemi sunulmaktadır. Çalışmada kullanılan veri seti ile henüz makine öğrenme tabanlı bir çalışma sunulmamıştır. Bu yönüyle çalışma başka bir çalışma ile karşılaştırılmamaktadır. Önerilen yöntemle yüksek sınıflandırma başarımı elde edilmesinin yanında hesaplama karmaşıklığı olarak da oldukça yüksek bir performansa sahiptir.

Tablo 2. Önerilen yöntemin diğer literatür çalışmalarıyla karşılaştırılması

Referans	Sınıflandırıcı	EEG Veri seti	Özellik Çıkarım	Doğruluk
Kulkarni ve ark. [4]	SVM	50 sağlıklı +50 Alzheimer	Karmaşık özellikleri	96%
Bairagi [1]	SVM	50 sağlıklı +50 Alzheimer	Spektral özellikler	94%
Ruiz-Gómez ve ark. [5]	MLP	37 sağlıklı +37 hafif bilişsel bozukluk + 37 Alzheimer	Spektral ve lineer olmayan özellikler	78.43%
Amezquita-sanchez ve ark.[6]	EPPN	37 hafif bilişsel bozukluk + 37 Alzheimer	Hurst üs ölçüleri ve fraktal boyut	90.3%
Kulkarni [7]	kNN	50 sağlıklı +50 Alzheimer	Spektral ve karmaşık özellikleri	94%
Tzamourta ve ark.[8]	Rastgele Orman	10 sağlıklı + 8 hafif düzey Alzheimer +6 orta düzey Alzheimer	İstatistiksel ve spektral özellikler	88.79%
Önerilen Yöntem	kNN	24 sağlıklı + 24 AD	Dalgacık Dönüşümü ve Hjorth parametreleri	91.12%

Sonuçlar

Bu çalışma ile EEG kayıtlarından Alzheimer hastaları ve sağlıklı kontrollerin sağlam bir şekilde ayırt edilmesine imkân sağlayan dalgacık dönüşümü ve Hjorth parametrelerinin hibrit olarak kullanıldığı yaklaşıma dayalı bir CAD sistemi sunulmaktadır. Ayrıca çalışmada farklı frekans bantları ile elde edilen sonuçlar değerlendirilmiştir. Hastalığın teşhisi maliyetli testler ve bu testlerin bir uzman tarafından değerlendirilmesi gibi yorucu ve zaman alan bir süreç içermektedir. Ayrıca bu testlerin değerlendirilmesi mesleki deneyime önemli ölçüde bağlı olduğu için bazı durumlarda tutarsız değerlendirmelere yol açabilir. Bu nedenle önerilen

çalışma ile hastalığın makine öğrenme algoritmalarıyla tespitini gerçekleştiren bir model geliştirilmiştir. Sunulan yöntemle elde edilen yüksek sınıflandırma başarımı ile Alzheimer ve sağlıklı kontrollerin ayırt edilmesinde uzman görüşüne destek sağlayacağı düşünülen oldukça tatmin edici sonuçlar elde edilmiştir.

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.

Kaynaklar

- [1] V. Bairagi, "EEG signal analysis for early diagnosis of Alzheimer disease using spectral and wavelet based features," *Int. J. Inf. Technol.*, vol. 10, no. 3, pp. 403–412, 2018.
- [2] L. R. Trambaiolli, N. Spolaôr, A. C. Lorena, R. Anghinah, and J. R. Sato, "Feature selection before EEG classification supports the diagnosis of Alzheimer's disease," *Clin. Neurophysiol.*, vol. 128, no. 10, pp. 2058–2067, 2017.
- [3] R. H. Blank, "Alzheimer's Disease and Other Dementias: An Introduction," in *Social & Public Policy of Alzheimer's Disease in the United States*, Springer, 2019, pp. 1–26.
- [4] N. N. Kulkarni and V. K. Bairagi, "Extracting salient features for EEG-based diagnosis of Alzheimer's disease using support vector machine classifier," *IETE J. Res.*, vol. 63, no. 1, pp. 11–22, 2017.
- [5] S. J. Ruiz-Gómez et al., "Automated multiclass classification of spontaneous EEG activity in Alzheimer's disease and mild cognitive impairment," *Entropy*, vol. 20, no. 1, p. 35, 2018.
- [6] J. P. Amezcua-Sanchez, N. Mammone, F. C. Morabito, S. Marino, and H. Adeli, "A novel methodology for automated differential diagnosis of mild cognitive impairment and the Alzheimer's disease using EEG signals," *J. Neurosci. Methods*, vol. 322, pp. 88–95, 2019.
- [7] N. Kulkarni, "Use of complexity based features in diagnosis of mild Alzheimer disease using EEG signals," *Int. J. Inf. Technol.*, vol. 10, no. 1, pp. 59–64, 2018.
- [8] K. D. Tzamourta et al., "EEG window length evaluation for the detection of Alzheimer's disease over different brain regions," *Brain Sci.*, vol. 9, no. 4, p. 81, 2019.
- [9] M. S. Safi and S. M. M. Safi, "Early detection of Alzheimer's disease from EEG signals using Hjorth parameters," *Biomed. Signal Process. Control*, vol. 65, p. 102338, 2021.
- [10] U. Orhan, M. Hekim, and M. Ozer, "EEG signals classification using the K-means clustering and a

- multilayer perceptron neural network model,” *Expert Syst. Appl.*, vol. 38, no. 10, pp. 13475–13481, 2011.
- [11] B. Hjorth, “EEG analysis based on time domain properties,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 29, no. 3, pp. 306–310, 1970.
- [12] A. M. Pineda, F. M. Ramos, L. E. Betting, and A. S. L. O. Campanharo, “Quantile graphs for EEG-based diagnosis of Alzheimer’s disease,” *PLoS One*, vol. 15, no. 6, p. e0231169, 2020.
- [13] B. R. Bakshi, “Multiscale PCA with application to multivariate statistical process monitoring,” *AIChE J.*, vol. 44, no. 7, pp. 1596–1610, 1998.
- [14] P. Jahankhani, V. Kodogiannis, and K. Revett, “EEG signal classification using wavelet feature extraction and neural networks,” in *IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing (JVA’06)*, 2006, pp. 120–124.
- [15] H. U. Amin et al., “Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques,” *Australas. Phys. & Eng. Sci. Med.*, vol. 38, no. 1, pp. 139–149, 2015.
- [16] R. C. Gonzalez, *Digital image processing*. Pearson education india, 2009.
- [17] N. S. Altman, “An introduction to kernel and nearest-neighbor nonparametric regression,” *Am. Stat.*, vol. 46, no. 3, pp. 175–185, 1992.
- [18] W. H. Organization and others, *Dementia: a public health priority*. World Health Organization, 2012.
- [19] J. Jeong, “EEG dynamics in patients with Alzheimer’s disease,” *Clin. Neurophysiol.*, vol. 115, no. 7, pp. 1490–1505, 2004.
- [20] C. Patterson and others, “World alzheimer report 2018,” 2018.
- [21] J. Dauwels, F. Vialatte, and A. Cichocki, “Diagnosis of Alzheimer’s disease from EEG signals: where are we standing?,” *Curr. Alzheimer Res.*, vol. 7, no. 6, pp. 487–505, 2010.
- [22] A. Alberdi, A. Aztiria, and A. Basarab, “On the early diagnosis of Alzheimer’s Disease from multimodal signals: A survey,” *Artif. Intell. Med.*, vol. 71, pp. 1–29, 2016.
- [23] R. Cassani, M. Estarellas, R. San-Martin, F. J. Fraga, and T. H. Falk, “Systematic review on resting-state EEG for Alzheimer’s disease diagnosis and progression assessment,” *Dis. Markers*, vol. 2018, 2018.
- [24] B. Oltu, M. F. Ak\csahin, and S. Kibaro\uglu, “A novel electroencephalography based approach for Alzheimer’s disease and mild cognitive impairment detection,” *Biomed. Signal Process. Control*, vol. 63, p. 102223, 2021.
- [25] S. Buyrukoğlu, Y. Yılmaz , & Z. Topalcengiz, “Correlation value determined to increase Salmonella prediction success of deep neural network for agricultural waters,” *Environmental Monitoring and Assessment*, 194(5), 1-12, 2022
- [26] S. Buyrukoğlu, & S. Savaş, “Stacked-Based Ensemble Machine Learning Model for Positioning Footballer”. *Arabian Journal for Science and Engineering*, 1-13, 2022.
- [27] R. Halde, “Application of Machine Learning algorithms for betterment in education system”. In *2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT)* (pp. 1110-1114). IEEE, 2016.
- [28] S. BUYRUKOĞLU, “Early detection of alzheimer’s disease using data mining: comparison of ensemble feature selection approaches”, *Konya Mühendislik Bilimleri Dergisi*, 9(1), 50-61,2021.
- [29] S. Buyrukoğlu, “Improvement of Machine Learning Models’ Performances based on Ensemble Learning for the detection of Alzheimer Disease”, In *2021 6th International Conference on Computer Science and Engineering (UBMK)* (pp. 102-106). IEEE, 2021.