



Development of ensemble machine learning model for single-channel EEG-based biometric identification system

Furkan Balci^{1,2*}, Hakkı Alparslan Ilgın³

¹Department of Electrical and Electronics Engineering, Faculty of Technology, Gazi University, 06560, Ankara, Türkiye

²Department of Electrical and Electronics Engineering, Graduate School of Natural and Applied Sciences, Ankara University, 06110, Ankara, Türkiye.

³Department of Electrical and Electronics Engineering, Faculty of Engineering, Ankara University, 06830, Ankara, Türkiye

Highlights:

- High accuracy is achieved using data collected from only a single channel
- Feature extraction is achieved at low computational cost without the need for complex transformations
- Identification is performed by recording without any stimulus or task definition

Keywords:

- Biometric
- Machine Learning
- Ensemble Classifier
- Identification
- EEG

Article Info:

Research Article

Received: 14.04.2025

Accepted: 12.12.2025

DOI:

10.17341/gazimmfd.1676219

Correspondence:

Author: Furkan Balci

e-mail:

furkanbalci@gazi.edu.tr

phone: +90 507 263 7744

Graphical/Tabular Abstract

In this study, a community classifier model is proposed to identify the identities of users within a closed system with low computational cost and high accuracy. For efficient and practical application in daily life, the number of channels used should be low, the accuracy rate high, and the computational cost minimized. Low computational cost directly impacts training and testing times. Therefore, single-channel EEG data collection reduces both application and computational costs. The flow diagram of the proposed model is presented in Figure A.

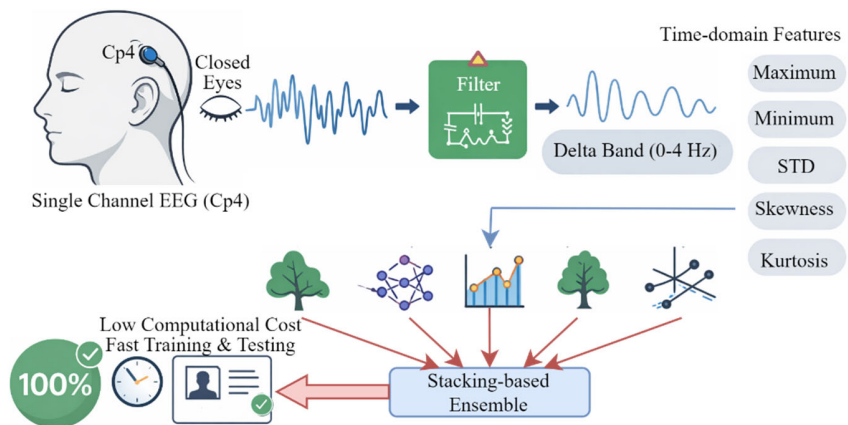


Figure A. Proposed single-channel EEG-based ensemble machine learning model

Purpose:

This study aims to achieve high-accuracy user identification in a closed system using single-channel EEG data.

Theory and Methods:

In the initial stage, data belonging to participants with recordings collected under various conditions were selected from a dataset comprising 109 participants for closed-system modeling. In the subsequent stage, the selected EEG data were decomposed into frequency subcomponents using a Butterworth filter, and the most suitable frequency component was identified by means of the ANOVA F-test. The extracted low-complexity features were then provided as inputs to an ensemble classifier consisting of random forest, multilayer neural network, gradient boosting, decision tree, and support vector machine algorithms.

Results:

In the single-channel system consisting of preprocessing, feature extraction, and classification steps, 100% classification accuracy was achieved using the Cp4 channel's delta frequency band and an ensemble algorithm composed of random forests, multi-layer neural network, gradient boosting, decision trees, and support vector machine algorithms on eyes-closed data. Additionally, compared to 64-channel experiments, an average reduction of 68 times in training time and 38 times in testing time was achieved.

Conclusion:

In conclusion, the findings indicate that a highly reliable single-channel biometric identification system is feasible and can be further developed, despite reducing the number of EEG channels through channel optimization.



Topluluk makine öğrenmesi modeli oluşturularak tek kanallı EEG tabanlı biyometrik kimlik tanımlama sistemi

Furkan Balcı^{1,2*}, Hakkı Alparslan İlgin³

¹Gazi Üniversitesi, Teknoloji Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, 06560, Ankara, Türkiye

²Ankara Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik-Elektronik Mühendisliği Anabilim Dalı, 06110, Ankara, Türkiye

³Ankara Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği Bölümü, 06830, Ankara, Türkiye

Ö N E Ç İ K A N L A R

- Sadece tek kanal üzerinden toplanan verilerle yüksek doğruluğa ulaşılmaktadır
- Karmaşık dönüşümlere gerek duymadan, düşük maliyetli öznetelik çıkarımı sağlanır
- Herhangi bir uyarın veya görev tanımlı olmadan kayıt yapılarak tanımlama gerçekleştirilmektedir

Makale Bilgileri

Araştırma Makalesi

Geliş: 14.04.2025

Kabul: 12.12.2025

DOI:

10.17341/gazimmfd.1676219

Anahtar Kelimeler:

Biyometrik,
makine Öğrenmesi,
topluluk sınıflandırıcısı,
kimlik tanımlama,
EEG

ÖZ

Klinik uygulamalarda yapay zekâya olan ilginin artmasıyla birlikte, bireylere özgü desenler içermesi ve sahte biyometrik verilere karşı yüksek güvenlik sağlaması nedeniyle elektroensefalografi (EEG), biyometrik tabanlı güvenlik sistemlerinde giderek daha fazla araştırılan bir yöntem haline gelmiştir. Bu çalışmada, düşük maliyetli bir yaklaşım hedeflenerek kapalı bir sistemde, 109 katılımcıdan 64 kanal üzerinden elde edilen EEG veri seti kullanılmıştır. Rastgele seçilen sınırlı sayıda deneğe ait EEG verileriyle, topluluk makine öğrenmesi tabanlı tek kanallı kimlik tespiti yaklaşımının başarımı incelenmiştir. Karar ağaçları tabanlı algoritmalarda kullanılan Gini önem katsayısı yardımıyla, gözler açık ve gözler kapalı koşullarında kaydedilen veriler arasında en ayırt edici kanal seçilmiştir. Seçilen kanaldan çıkarılan özneteliklerle eğitilen tek kanallı makine öğrenmesi modelleri; 64 kanallı modellerle eğitim süresi, test süresi ve sınıflandırma performansı açısından karşılaştırılmıştır. Ön işleme, öznetelik çıkarımı ve sınıflandırma adımlarından oluşan sistemde, gözler kapalı veriler için Cp4 kanalının delta frekans bandı kullanılmış; rastgele ormanlar, çok katmanlı sinir ağı, gradyan artırma, karar ağaçları ve destek vektörleri makinesinden oluşan topluluk modeliyle sınıflandırma gerçekleştirilmiştir. Çapraz doğrulama sonuçlarında %100 doğruluk elde edilirken, 64 kanallı sistemlere kıyasla eğitim ve test sürelerinde belirgin iyileşmeler sağlanmıştır. Bu sonuçlar, kanal sayısı azaltılsa dahi güvenilir tek kanallı EEG tabanlı biyometrik kimlik tespitinin mümkün olduğunu göstermektedir.

Development of ensemble machine learning model for single-channel EEG-based biometric identification system

H I G H L I G H T S

- High accuracy is achieved using data collected from only a single channel
- Feature extraction is achieved at low computational cost without the need for complex transformations
- Identification is performed by recording without any stimulus or task definition

Article Info

Research Article

Received: 14.04.2025

Accepted: 12.12.2025

DOI:

10.17341/gazimmfd.1676219

Keywords:

Biometric,
machine learning,
ensemble classifier,
identification, EEG

ABSTRACT

With the growing interest in artificial intelligence in clinical applications, electroencephalography (EEG) has become an increasingly researched technique in biometric-based security systems due to its ability to capture individual-specific patterns and provide higher resistance to spoofing compared to other biometric modalities. In this study, a low-cost approach was targeted by utilizing an EEG dataset acquired from 109 participants using 64 channels within a closed system. The performance of an ensemble machine learning-based single-channel identity recognition approach was investigated using EEG data from a randomly selected limited number of subjects. Using the Gini importance coefficient employed in decision tree-based algorithms, the most discriminative channel was selected from datasets recorded under both eyes-open and eyes-closed conditions. Single-channel machine learning models trained with features extracted from the selected channel were compared with 64-channel models in terms of training time, testing time, and classification performance. In the system consisting of preprocessing, feature extraction, and classification stages, the delta frequency band of the Cp4 channel was used for eyes-closed data, and classification was performed using an ensemble model comprising random forests, multilayer perceptron, gradient boosting, decision trees, and support vector machine algorithms. Cross-validation results achieved 100% classification accuracy, while significant reductions in training and testing times were obtained compared to 64-channel systems. These findings demonstrate that a reliable single-channel EEG-based biometric identification system is feasible even with a reduced number of channels.

1. Giriş (Introduction)

Biyometri, bireylerden toplanan davranışsal ve fizyolojik özellikleri araştıran ve bu özellikleri topluluk içindeki bireylerin kimlik tespiti amacıyla kullanan bir alandır. İnsan davranışına bağlı davranışsal (ses, yürüyüş gibi) ve insan fizyolojisine bağlı fizyolojik (DNA, parmak izi gibi) olarak ikiye ayrılan biyometrik veriler, bireylerden toplanan benzersiz verilerdir [1]. Biyometrik veriler kullanılarak tasarlanan kimlik tanıma sistemlerinin en önemli sınırlamalarından biri, sahte verilerin üretilmesidir. Özellikle içsel sinyaller olan elektrokardiyogram (EKG), elektromiyogram (EMG) ve elektroensefalogram (EEG) gibi biyolojik sinyallerin kopyalanması daha zor olduğu için, bu veriler kullanılarak tasarlanan deneysel çalışmaların sayısı son yıllarda artmıştır [2]. Her insan özelliği biyometrik olarak kullanılmadığı gibi, biyometrik kimlik tanıma sistemlerinde kullanılan veriler; kullanıcıya özgünlük, benzersizlik, sağlık ve evrensellik gibi bazı kriterleri karşılamalıdır. Bu kriterler göz önünde bulundurularak beynin elektriksel aktivitesini temsil eden EEG sinyallerinin kimlik doğrulama sistemlerinde kullanımının mümkün olabileceği kanıtlanmıştır [3-6]. EEG tabanlı yöntemler her ne kadar bu kriterleri karşılıyor olsa da parmak izi ve iris gibi geleneksel biyometriklere göre kayıt ortamının oluşturulması daha zor ve kayıt cihazlarının maliyeti kanal sayısına bağlı olarak artmaktadır. Ayrıca EEG, gürültü ve artefaktlardan kolayca etkilenmektedir. Bu yüzden doğrudan karakteristik bir temsil olarak kullanılması zor iken, beyin zihinsel görevlere göre karakteristik bir sinyal olması, taklit edilmesini son derece zorlaştırmaktadır. EEG, beyin aktiviteleriyle ilişkilidir ve beyin aktiviteleri, bireyin zihinsel durumu ile bağlantılıdır. Bu yüzden güvenli bir tanımlama için belirli koşulların (gözler açık, gözler kapalı, dış uyaran etkisi, dinlenme durumu kaydı gibi) sağlanması gerekmektedir. Stres ve baskı gibi durumlarda oluşan EEG verileri, istenilen koşullardaki verilerden farklıdır. Bu farklılık, EEG tabanlı sistemlerde başka kişiler tarafından zorla kimlik tanımlama uygulamasının önüne geçmektedir.

EEG sinyali, içerdiği anlamlı koşullara dayalı olarak frekansına göre; delta (0-4 Hz), teta (4-8 Hz), alfa (8-16 Hz), beta (16-28 Hz) ve gama (28-32 Hz) olmak üzere beş dalgayı ayırır [7]. Bazı uygulamalarda bu dalgalarla örtüşen frekanslara sahip mu dalgası (8-13 Hz) da kullanılmaktadır [8]. EEG tabanlı biyometrik kimlik tanımlama sistemindeki temel amaç, kişiden alınan verilerin mevcut veri tabanındaki verilerle karşılaştırılmasıdır. Kimlik tanımlama sistemlerinde birkaç temel adım bulunmaktadır. Genellikle ilk adımda, kişilerin istenilen özelliklerinin toplandığı kayıt aşaması yer alır. Bazı sistemler sadece tek bir biyometrik veri kullanırken, bazı sistemler birden fazla biyometrik veri kullanarak çok modlu bir çalışma sunmaktadır. Daha sonra, elde edilen bu özellikler için ön işlemler ve öznitelik çıkarımı uygulanır. Çıkarılan öznitelikler sayesinde ham verilerden daha anlamlı veriler elde edilmektedir. Önerilen sistemin son aşamasında, çıkarılan öznitelikleri kullanarak sistemdeki veri tabanında yer alan özniteliklerle karşılaştırma yaparak kimlik tespitinde bulunan algoritma yer almaktadır. Günümüzde yapay zekânın gelişmesiyle birlikte, bu karşılaştırma işleminde çeşitli makine öğrenmesi algoritmaları kullanılarak kimlik tanımlamada yüksek sınıflandırma doğruluklarına ulaşılmaktadır. Henüz sağlıklı bireyler üzerinde uygulanan ve laboratuvar aşamasında olan EEG tabanlı biyometrik sistemlerden elde edilen sonuçlar incelendiğinde, EEG verisinin biyometrik olarak kullanılmasının mümkün olduğu görülmektedir. Yapılan son çalışmalarda, EEG tabanlı biyometrik sistemlerde genellikle yüksek doğruluğun sağlanabilmesi için çok sayıda kanaldan elde edilen EEG verileri kullanılmaktadır [9, 10]. Çok sayıda kanalın kimlik tanımda kullanılması, hem hesaplama maliyetini hem de kullanılacak kayıt cihazının maliyetini artırmaktadır. Çoğu çalışmada kullanılan 32 veya 64 kanallı kayıt

cihazlarının önerilen sistemlerdeki varlığı, bu tekniklerin uygulanabilirliğini imkânsız hale getirmektedir. EEG sinyallerinin kullanıldığı biyometrik kimlik tanımlama çalışmaları için istatistiksel olarak anlamlı sonuçların elde edilmesi adına daha geniş kapsamlı verilerle gerçekleştirilen daha büyük ölçekli uygulamaların analizine ihtiyaç duyulmaktadır.

Bu çalışmanın temel amacı, EEG tabanlı bir kimlik tanımlama sistemini hem hesaplama hem de kanal sayısı maliyetlerinden arındırarak, daha kullanılabilir ve düşük maliyetli bir tek kanallı EEG tabanlı kimlik tanımlama sisteminin tasarlanmasıdır. Bu sayede EEG desenlerinin bireye özgü özellikleri, geniş ölçekli insan gruplarına uygun, kompakt, kullanıcı dostu ve maliyet açısından verimli bir sistem tasarımı ile değerlendirilecektir. Önerilen sistem, literatürde yer alan EEG tabanlı biyometrik kimlik tanımlama çalışmalarından birkaç önemli açıdan ayrılmaktadır. Öncelikle, çok kanallı EEG verileri yerine yalnızca tek kanal (Cp4) versiyonunun delta frekans bandı kullanılarak yüksek doğruluk elde edilmesi, sistemin hem donanım hem de veri işleme maliyetini ciddi oranda azaltmaktadır. Bu durum, literatürde çoğunlukla 32 veya 64 kanallı EEG sistemleriyle çalışılarak elde edilen yüksek doğruluk oranlarının, daha az karmaşık ve düşük maliyetli bir yapı ile de sağlanabileceğini göstermektedir. Kanal seçiminde Gini önem katsayısı yaklaşımı kullanılmaktadır. Hem gözler açık hem de gözler kapalı olarak kaydedilen veri setindeki sınırlı sayıdaki kullanıcıdan alınan verilerle, seçilen kanal ile eğitilen sınıflandırıcının 64 kanallı veri setiyle eğitilen sınıflandırıcılara göre performans karşılaştırması yapılmıştır. Değerlendirme aşamasında doğruluk, duyarlılık, kesinlik ve F1-skoru gibi performans parametreleri açısından karşılaştırmalar yapılmıştır. İnsanlardan toplanan biyometrik veriler zamanla değişiklik gösterebilmektedir. Bu nedenle biyometrik sistemlerde belirli zaman aralıklarında biyometrik veri tabanının güncellenmesi gerekmektedir. Bu yüzden tasarlanan sistemin eğitim sürecinin güncellemeye izin verebilir olması ve bunu daha hızlı ve efektif gerçekleştirmesi için eğitim sürecinin incelenmesi de önemlidir. Literatürde sıklıkla kullanılan doğruluk, kesinlik, hassasiyet ve F1 skoru gibi karşılaştırma metriklerinin yanında, eğitim ve test süreleri de incelenmiştir. Ön işleme, öznitelik çıkarımı ve sınıflandırma adımlarından oluşan tek kanallı sistemde gözler kapalı veri setiyle, Cp4 kanalının delta frekans bandı ile rastgele ormanlar (RO, Random forest:RF), çok katmanlı sinir ağı (ÇKSA, Multi-layer perceptron: MLP), gradyan artırma (GA, Gradient boosting classifier: GBC), karar ağaçları (KA, Decision tree: DT) ve destek vektörleri makinesi (DVM, Support vector machine: SVM) algoritmalarından oluşan topluluk algoritması kullanılarak %100 sınıflandırma doğruluğuna ulaşılmıştır. Önerilen model karmaşık derin öğrenme mimarileri yerine düşük hesaplama maliyetine sahip geleneksel makine öğrenmesi algoritmalarının topluluk yapısını kullanarak hem yüksek doğruluk hem de eğitim ve test sürelerinde çok ciddi iyileşme sağlamaktadır. En yakın doğruluğa sahip 64 kanallı sisteme göre yaklaşık olarak eğitim süresi 42 kat ve test süresi 6 kat kısalmıştır. Hem gözler açık hem de gözler kapalı durumlar için alt frekans bantlarının karşılaştırılmasıyla, farklı senaryolarda en uygun frekans bileşeninin belirlenmesine yönelik deneysel bir analiz sunulmaktadır. 64 kanallı veri setinden seçilen tek elektrod EEG verisinden çıkarılan özniteliklerle hesaplama maliyeti düşük, yüksek doğruluk sağlayan bir kimlik doğrulama sistemi geliştirilmesini mümkün olduğunu göstermek kullanılan biyometrik veri açısından yenilikçi ve piyasada uygulanabilir bir sistemin tasarlanabileceğini kanıtlamaktadır. Bu yönleriyle önerilen sistem, hem akademik araştırmalara hem de gerçek dünya uygulamalarına uygun, maliyet-etkin ve yüksek performanslı bir biyometrik kimlik doğrulama çözümü ortaya koymaktadır. Bu makalenin ana katkıları şu şekilde özetlenebilir:

- Kapalı ortamlar için tesis ve ekipmanların erişiminin kontrolünün sağlanması için biyometrik kimlik tanımlaması yapan topluluk sınıflandırıcısı önermek.
- Karmaşık ve yoğun katmanlı derin öğrenme mimarilerini kullanmadan geleneksel ve hızlı eğitim-test sürelerine sahip modelleri verimli bir şekilde eğitmek için düşük hesaplama maliyetli zaman bölgesi özniteliklerini çıkaran ön işlem basamağı kullanmak.
- Önerilen topluluk sınıflandırıcısı yaklaşımını doğruluk, keskinlik, hassasiyet, F1 skoru, eğitim ve test süresi gibi metriklerle çeşitli algoritmalarla karşılaştırmalı değerlendirmek.
- Gözler açık veya kapalı kaydedilen verilerin istatistiksel özniteliklerini inceleyerek hangi veride hangi alt frekans bandının kullanılacağı konusunda avantajlar ve dezavantajları belirlemek.

1.1. Motivasyon (Motivation)

Bu makalenin temel motivasyonu, kopyalanması zor içsel biyometriklerden biri olan EEG verisini kullanan, hem düşük hesaplama maliyetine sahip hem de tek kanal kullanarak donanım ve veri toplama maliyetini önemli ölçüde azaltan yüksek doğruluklu bir kapalı ortam biyometrik kimlik tanımlama sisteminin oluşturulmasıdır. Çok kanallı EEG sistemleri yüksek donanım maliyeti, karmaşık kablolu, uzun kurulum süresi ve yüksek veri işleme yükü gerektirirken; tek kanal kullanımı bu dezavantajları ortadan kaldırarak sistemin taşınabilirliğini, kurulabilirliğini ve erişilebilirliğini artırmaktadır. Böylece, düşük bütçeli kurumlar veya portatif güvenlik çözümleri gerektiren senaryolarda uygulanabilirlik sağlanmaktadır. Güvenirlikleri düşük olan parmak izi, ses gibi çeşitli biyometriklere karşı EEG tabanlı yaklaşımların hem güncel deneysel bir araştırma konusu olması hem de yüksek güvenilirlik sunması bu çalışmayı önemli hale getirmektedir. Kopyalama sorununa karşı EEG sinyalinin çoklu kişi ve çoklu görev senaryolarında bile güvenlik ve gizlilik yönetimde kullanılabilir olduğu literatürdeki çalışmalarla kanıtlanmıştır. EEG sinyalleri kişiden kişiye farklılık gösterirken, zorlama durumunda oluşan baskı ve stres tepkilerinde de değişiklik göstermesi bu biyometrik verinin zorla tanımlama girişimlerine karşı da güvenilir olduğunu kanıtlamaktadır. Kapalı ortam modellemesi, bu çalışmaların evrensel olabilmesi için tümevarım yaklaşımının bir örneğini oluşturmaktadır. Henüz deneysel aşamalarda olan bu kimlik tanımlama yaklaşımı, yüksek güvenilirlikli ortamlarda erişim kontrolü için yüksek koruma potansiyeline sahiptir.

Bu makalenin ikinci bölümü, biyometrik sinyal olarak kullanılan EEG sinyali üzerinde yapılan önceki çalışmaları tartışarak genel bir bakış sunmaktadır. Üçüncü bölümde yöntemler açıklanmaktadır. Dördüncü bölümde yapılan testlerden elde edilen sonuçlar paylaşılarak üzerinde tartışılmaktadır. Son bölüm ise sonuçlar üzerine değerlendirme yapılmaktadır.

2. Biyometrik Olarak EEG ve İlgili Çalışmalar (EEG as Biometric and Related Studies)

Biyometrik sistemler, bireylerden elde edilen davranışsal veya fizyolojik biyometriklerin kullanılarak kişilerin tanımlandığı sistemlerdir. Elde edilen bu biyometrikler, sistem içerisinde ham veri halinden çeşitli özniteliklerin çıkarıldığı forma dönüştürülerek, veri tabanlarında daha sonra karşılaştırılmak üzere saklanır. Kimlik karşılaştırma algoritmaları, sistemden sisteme farklılık göstermektedir. Günümüz teknolojisinde, bu karşılaştırıcı algoritmalar için özellikle makine öğrenmesi gibi yapay zekâ tekniklerine sıklıkla başvurulmaktadır [12]. Biyometrik sistemler sayesinde, geleneksel kullanıcı adı ve parola tekniğine göre, kullanıcı deneyimleri iyileştirilerek güvenlik düzeyinde önemli bir iyileştirme sağlanabilmektedir [13]. Bireylerden toplanan her özellik biyometrik veri olarak kullanılmaz. Bir biyolojik özelliğin biyometrik olabilmesi

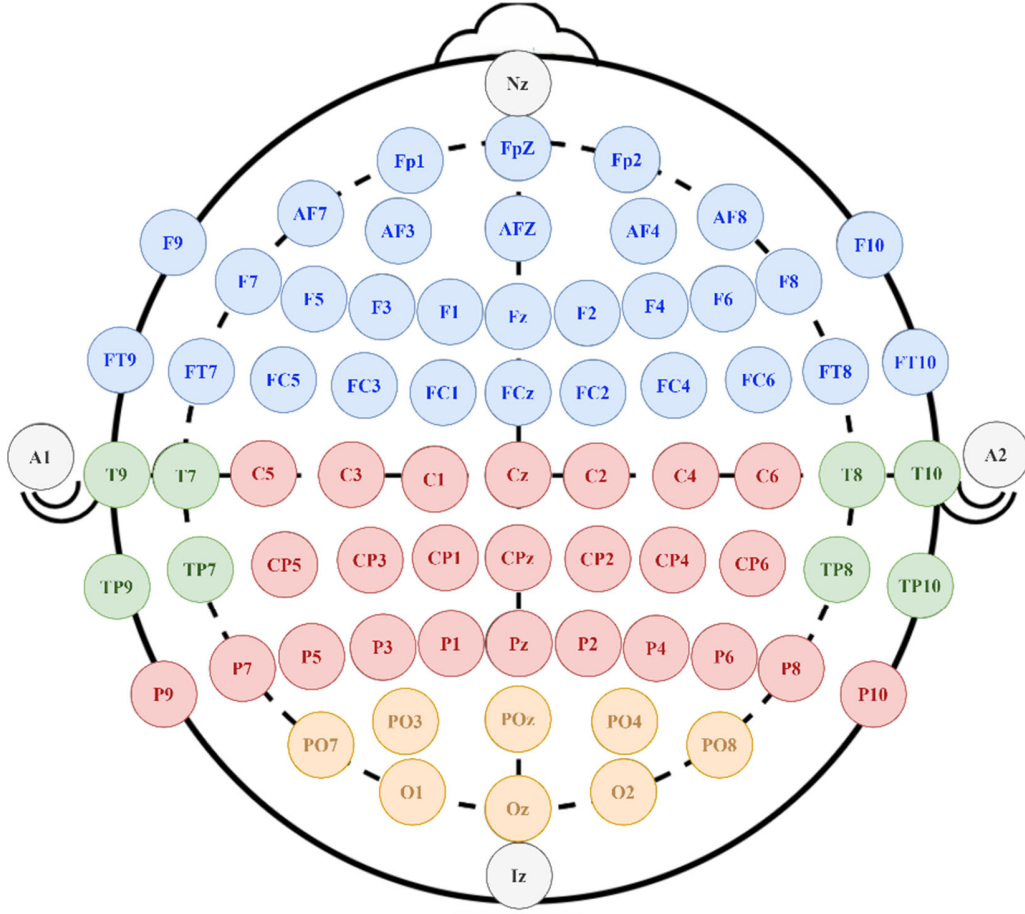
için, ilgili özelliklerin ölçülebilir ve bireye özgü olması gerekmektedir. Özellikle son yıllarda yapılan araştırmalar, hem uyarıcı tabanlı kayıtlarla hem de dinlenme durumundaki kayıtlarla elde edilen bulgular ışığında, EEG sinyallerinin güvenli bir biyometrik sistem gereksinimlerinin çoğunu karşıladığı göstermektedir. Temel olarak, EEG tabanlı bir kimlik tanımlama sisteminde yüksek doğruluklara ulaşılması için veri toplama protokolü, öznitelik seçimi teknikleri ve sınıflandırma algoritmaları gibi faktörlerin araştırılması gerekmektedir.

EEG sinyalleri genellikle uluslararası bir standart haline gelen 10/20 elektrot düzenine uygun şekilde kaydedilmektedir. İnsan beyin yapısı anatomik olarak incelendiğinde, frontal lob, parietal lob, oksipital lob ve temporal lob olmak üzere dört ana bölgeden oluşmaktadır. Her bölgenin farklı görevler altında farklı sinyaller ürettiği tespit edilmiştir [13]. Beynin ön kısmında bulunan frontal lob, motor kontrol, yüksek bilişsel işlevler, duygusal kontrol ve konuşma gibi aktivitelerde etkin rol oynar. Başın üst ve orta kısmında yer alan parietal lob, dokunma duyusu, el-göz koordinasyonu ve mekânsal farkındalık gibi aktivitelerde etkin rol oynar. Beynin arka kısmında yer alan oksipital lob, görsel işleme, renk algısı, hareket algısı ve şekil algısı gibi aktivitelerde etkin rol oynar. Beynin yan kısımlarında yer alan temporal lob ise işitsel aktiviteler, dil işleme, hafıza ve duygusal işlemler gibi aktivitelerde etkin rol oynar. Bazı çalışmalarda, kimlik tanımlama doğruluğunun artırılması için çeşitli zihinsel ve bilişsel aktiviteler altında kayıtlar toplanmaktadır. Herff vd., konuşma durumundaki EEG sinyallerini incelediklerinde, telaffuz edilen kelime ile o anki EEG arasında bir ilinti tespit etmişlerdir [14]. Bu durum, konuşma gibi aktivitelerin biyometrik tanımlamada kullanılabileceğini göstermektedir. Bu durumda, ilgili beyin bölgesine denk gelen kanallardaki veriler daha etkin şekilde işlenmektedir.

Şekil 1'de (Mavi renkle gösterilen kanallar frontal lobu, kırmızı renkle gösterilen kanallar parietal lobu, yeşil renkle gösterilen kanallar temporal lobu ve sarı renkle gösterilen kanallar ise oksipital lobu temsil etmektedir), 10/20 elektrot düzenine göre 64 kanallı bir EEG kayıt cihazının kanallarının hangi bölgeden verileri topladığı gösterilmektedir. Zihinsel görevlere ihtiyaç duyulması, kayıt esnasında ek karmaşıklığa sebep olduğu için bazı çalışmalar herhangi bir zihinsel aktivite olmadan EEG özelliklerini incelemektedir. Örneğin, Poulos vd., O2 kanalından topladıkları sinyaller üzerinde yaptıkları çalışmayla dinlenme durumunda alfa bandının baskınlığını vurgulamıştır [15]. Ayrıca, beyin fonksiyonel bağlantısallık kavramıyla beraber beyin hiyerarşik modülerlik göstermesi, EEG tabanlı biyometriklerde özevte merkezilik yaklaşımının da kullanılabilir olmasına olanak sağlamaktadır [16].

Zihinsel aktivitelerin yanı sıra, bu makalenin en önemli araştırma konusu olan kanal sayısı da biyometrik sistemlerde EEG sinyallerinin kullanılabilirliğini önemli derecede etkilemektedir. Dan vd., tek kanallı bir EEG kayıt cihazı olan Neurosky Mindset kullanılarak 13 kişilik bir kapalı sistem modellemesine uygun deney grubunda %87 doğrulukla kimliklendirme yapmıştır [17]. Bu sonuç her ne kadar yeterli gibi gözükse de her bir kişiden 18 dakikalık kayda ihtiyaç duyması, özellikle kayıt süresi konusunda bazı olumsuzlukları ortaya çıkarmaktadır. Benzer şekilde, Su vd., HXD-I model bir EEG kayıt cihazı ile %97.5 gibi yüksek bir doğruluğa ulaşmıştır [18]. Ancak, kayıt süresinin 60 dakika olması bu sistemden pratikliği konusunda şüphe uyandırmaktadır. Bu çalışmalar genel olarak incelendiğinde, sistemde kullanılan kanal sayısının ve kayıt sürelerinin hem algoritmanın eğitimi hem de testi için minimum düzeyde olması gerektiği anlaşılmaktadır.

İşlenecek olan EEG verilerinin daha anlamlı bilgiler içerdiğini göstermek için çeşitli öznitelik çıkarımı yaklaşımlarına ihtiyaç



Şekil 1. 10/20 elektrot yerleşim sistemine göre kanalların beyin loblarına göre dağılımı
(Channel locations mapped to corresponding brain lobes according to the 10/20 electrode placement system)

duyulmaktadır. Ham EEG verisinden ayırt edici desenlerin çıkarılması oldukça zordur. Kişiyeye özgü bu anlamlı bilgilerin çıkarılması, özellikle doğruluk seviyesinin artırılması ve sistemin istikrarlı bir şekilde çalışabilmesi için önemli bir adımdır. Çıkarılan bu öznitelikler, beyin aktivitesinin daha iyi anlaşılmasına olanak sağlarken algoritmalar için ayırt ediciliği artırmaktadır. Bu ayırt edici öznitelikler; frekans, zaman ve zaman-mekânsal bölge olmak üzere üç temel bölgeden elde edilmektedir. Zaman bölgesi öznitelikleri, ham EEG sinyallerinden sinyalin zamanda genel dinamiğinin nasıl değiştiğini doğrudan göstermek için çıkarılmaktadır. Zaman bölgesi öznitelikleri diğer özniteliklere göre daha az karmaşıktır. Ortalama, standart sapma, tepe değer, güç, enerji, entropi, Skewness ve Kurtosis gibi öznitelikler EEG tabanlı çalışmalarda sıklıkla kullanılan zaman bölgesi özniteliklerinden bazılarıdır. Frekans bölgesi öznitelikleri ise EEG sinyallerini frekans bileşenlerine ayırarak anlamlı öznitelikler çıkarmaktadır. Bu öznitelikler genellikle Fourier veya dalgacık dönüşümü gibi yöntemlerle elde edilmektedir. Frekans bölgesi öznitelikleri sıklıkla kullanılan özniteliklerden bazıları; güç spektrumu yoğunluğu, bant güç oranları, dominant frekans ve ortalamaya dayalı bant gücüdür. Bu gibi özniteliklerle beynin bilişsel durumu hakkında detaylı bilgilere ulaşılabilmektedir. Hem zaman hem de frekans bilgilerine ulaşabilmeye olanak sağlayan zaman-mekânsal öznitelikler ise beynin farklı bölgeleri arasında zamanla değişen etkileşimleri incelemektedir. Koherens, Granger nedenselliği, faz kilitlenme değeri ve kapsamlı bağıllık gibi tekniklerle zaman-mekânsal öznitelikler çıkarılabilmektedir. Literatürde yer alan çalışmalar ve bu çalışmaların kullandığı parametreler (kanal sayısı,

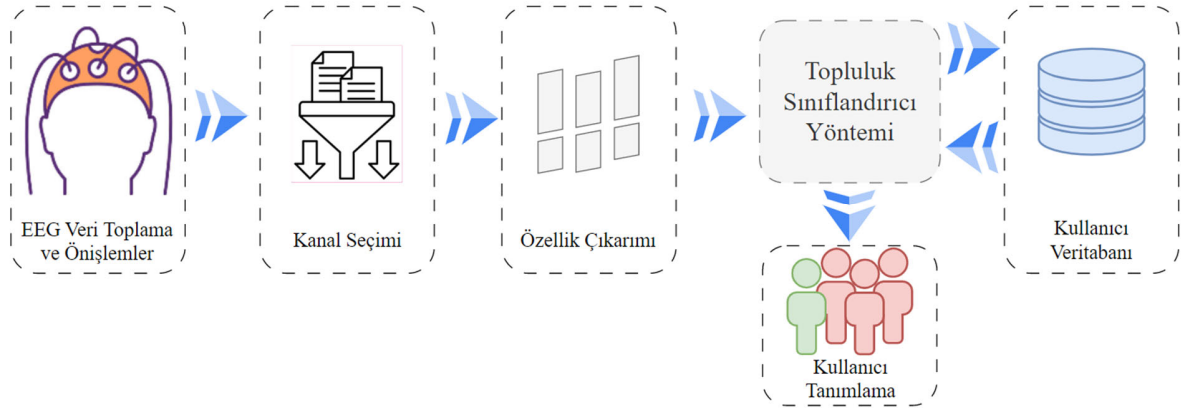
öznitelik çıkarımı, denek sayısı, karşılaştırma algoritması gibi) Tablo 1'de paylaşılmaktadır.

Suppiah ve Vinod, yaptıkları çalışmada tek kanallı bir EEG kimlik doğrulama sistemi tasarlamıştır [19]. Kimlik doğrulama yaklaşımı, kimlik tanımlamaya benzer yapıda olup kişinin veri tabanında var ya da yok olduğunu belirtmektedir. Bu yüzden, bu çalışmalar ikili sınıf mimarisinde oldukları için daha yüksek doğruluklara ulaşmaktadır. Hem gözlerin açık hem de gözlerin kapalı olduğu durumlarda kaydedilen 109 denegin verileriyle yapılan doğrulama, özellikle algoritmaların çok daha hızlı olduğunu vurgulamıştır. Palaniappan vd., yaptıkları deneysel bir çalışmada, üç denekle tek kanallı bir EEG tabanlı kimliklendirme sistemi tasarlamıştır [20]. Bu çalışmada, diğer çalışmalardan farklı olarak EEG verisiyle beraber kullanıcıların tanımladığı bir kişisel kimlik numarası (PIN) da kullanılmıştır. En yüksek doğruluğa Cz kanalıyla ulaşılmıştır. Nakanishi vd., 23 kişinin Fp1 kanalından elde edilen alfa bandından iç büyüklük ve dış büyüklük öznitelikleri ile kimlik tanımlama modeli önermiştir [21]. Kanal sayısının azaltılması maliyeti düşürürken her çalışmada tek kanallı bir yaklaşım hedeflenmemektedir. Örneğin, Keshishzadeh vd., üç kanallı bir kimlik tanımlama sistemi geliştirmiştir [22]. Çalışmada bu üç kanaldan Özünlü Kesirli Fourier Dönüşümü (DFrFT) kullanarak öznitelikler çıkarılmış ve %0,22 Eşit Hata Oranına (EER) ulaşılmıştır. EEG tabanlı kimlik tanımlama ve doğrulama çalışmaları hâlâ açık bir araştırma alanı olup geliştirmeye ihtiyaç duymaktadır. Bu çalışmada özellikle kullanılan kanal sayısının tek olabildiği ve çıkarılan özniteliklerin de daha az işlem maliyetine sahip olması gibi konulara değinilmektedir.

Tablo 1. Literatürde yer alan ilgili çalışmaların temel parametrelerinin özeti
(Summary of the key parameters reported in the relevant studies in the literature)

Çalışma	Kanal Sayısı	Denek Sayısı	Çalışma Amacı	Doğruluk
[23]	6	7	Ayrık dalgacık dönüşümü, log enerji entropisi, örnek entropisi ve oto-regresif katsayı öznitelikleri, farklı görevler (dinlenme hali, matematik problemi çözme, görsel sayma, geometrik şekil döndürme) sonucunda toplanan EEG verilerine SVM ve Bayes ağı kullanarak uygulandı.	%95
[24]	19	20	Güç Spektral Yoğunluğu öznitelik çıkarımı olarak kullanarak görsel uyarıcıların kimlik tespiti üzerindeki etkisi kNN ve LDA sınıflandırıcılarıyla araştırıldı.	%94,95
[25]	4, 16, 32, 64	109	Kimlik tespiti, motor hareketleri sonucunda oluşan EEG sinyali kullanılarak 1D Evrimsel LSTM mimarisi ile gerçekleştirildi.	%99,58
[26]	14	8	LSTM mimarisi kullanarak düşük maliyetli bir sistem tasarlamak amaçlanmıştır.	%89,5
[27]	1	109	Optimum sınıflandırma sonucunu bulmak için SVM ve LDA sınıflandırıcısı kullanarak tek kanal performansı analiz edilmiştir.	%82,58
[28]	14	109	Dinlenme durumu EEG kayıtları kullanarak düşük maliyetli, CNN tabanlı bir model tasarlandı. Kanal sayısı çok düşük tutularak istenen optimum doğruluk elde edildi.	%99,32

kNN: k en yakın komşuluk algoritması (k-nearest neighbors), LDA:Doğrusal ayırt etme analizi (Linear discriminant analysis), LSTM: Uzun kısa süreli bellek (Long-short term memory)



Şekil 2. Önerilen tek kanallı EEG tabanlı topluluk makine öğrenmesi modelinin işlem akışını gösteren diyagram
(Flowchart illustrating the processing pipeline of the proposed single-channel EEG-based ensemble machine learning model)

Suppiah ve Vinod, yaptıkları çalışmada tek kanallı bir EEG kimlik doğrulama sistemi tasarlamıştır [19]. Kimlik doğrulama yaklaşımı, kimlik tanımlamaya benzer yapıda olup kişinin veri tabanında var ya da yok olduğunu belirtmektedir. Bu yüzden, bu çalışmalar ikili sınıflandırma mimarisinde oldukları için daha yüksek doğruluklara ulaşmaktadır. Hem gözlerin açık hem de gözlerin kapalı olduğu durumlarda kaydedilen 109 denegın verileriyle yapılan doğrulama, özellikle algoritmaların çok daha hızlı olduğunu vurgulamıştır. Palaniappan vd., yaptıkları deneysel bir çalışmada, üç denekle tek kanallı bir EEG tabanlı kimliklendirme sistemi tasarlamıştır [20]. Bu çalışmada, diğer çalışmalardan farklı olarak EEG verisiyle beraber kullanıcıların tanımladığı bir kişisel kimlik numarası (PIN) da kullanılmıştır. En yüksek doğruluğa Cz kanalıyla ulaşılmıştır. Nakanishi vd., 23 kişinin Fp1 kanalından elde edilen alfa bandından iç büyüklük ve dış büyüklük öznitelikleri ile kimlik tanımlama modeli önermiştir [21]. Kanal sayısının azaltılması maliyeti düşürürken her çalışmada tek kanallı bir yaklaşım hedeflenmemektedir. Örneğin, Keshishzadeh vd., üç kanallı bir kimlik tanımlama sistemi geliştirmiştir [22]. Çalışmada bu üç kanaldan Özünlü Kesirli Fourier Dönüşümü (DFrFT) kullanarak öznitelikler çıkarılmış ve %0,22 Eşit Hata Oranına (EER) ulaşılmıştır. EEG tabanlı kimlik tanımlama ve doğrulama çalışmaları hâlâ açık bir araştırma alanı olup geliştirmeye ihtiyaç duymaktadır. Bu çalışmada özellikle kullanılan kanal sayısının tek olabilmesi ve çıkarılan özniteliklerin de daha az işlem maliyetine sahip olması gibi konulara değinilmektedir.

3. Yöntemler (Methods)

Bu bölümde, makalede önerilen tek kanallı EEG kimlik tanımlama sisteminde kullanılan veri seti, öznitelik çıkarma yöntemleri ve sınıflandırma yöntemleri hakkında bilgiler paylaşılmaktadır. Şekil 2’de, önerilen tek kanallı EEG kimlik tanımlama sisteminin akış diyagramı yer almaktadır. Önerilen sistemde, EEG verilerinin toplanması aşamasını takip eden aşama önışlem aşamasıdır. Önışlem aşamasında EEG sinyali alt frekans bileşenlerine ayrılmaktadır. Uygun alt frekans bileşeninin belirlenmesiyle birlikte zaman bölgesinde öznitelik çıkarılmaktadır. Sonraki aşamada, belirlenen makine öğrenmesi algoritmalarından oluşan topluluk makine öğrenmesi sınıflandırıcısı önce eğitilmekte, ardından test edilmektedir.

3.1. EEG Veri Seti (EEG Dataset)

Bu makalede, PhysioNet’te paylaşılan EEG motor hareketi/hayali veri seti kullanılmaktadır [29]. Bu veri setinde, 109 denekten toplanan 64 kanallı veriler bulunmaktadır. Farklı görevler altında toplanan veri setinden yalnızca loş ışıklı bir ortamda, gözler açık ve gözler kapalı olarak dinlenme durumunda kaydedilen veri setleri kullanılmaktadır. EEG verileri, 64 kanala sahip olan BCI2000 kayıt cihazıyla saniyede 160 örnek toplanarak kaydedilmiştir. Ortalama olarak, her denekten 1 dakikalık veriler kaydedilmiştir. Ancak, deneklerin kayıtlarının

çoğunda kayıt başlangıcında ve bitişinde çeşitli gürültüler bulunmaktadır. Gözler açık olan kayıta göz kaslarının gürültüsünün etkisini azaltmak için deneklerden gözlerini kırpmamaları istenmiştir. Bu sayede sinyal kontaminasyonunun önüne geçilmeye çalışılmıştır. Çalışmada, önerilen yaklaşımın performansını farklı ölçeklerde test edebilmek amacıyla 8, 16, 24, 50 ve 109 denek üzerinde deneyler gerçekleştirilmiştir. Deneklerin seçiminde herhangi bir gürültü kriteri gözetilmemiş, seçim süreci tamamen rastgele yapılmıştır. Her denekten, gözler kapalı olarak toplanan veriler yaklaşık 9500, gözler açık olarak toplanan veriler ise 1600 örneğe sahiptir.

3.2. Sinyal Önışleme ve Kanal Seçimi (Signal Preprocessing and Channel Selection)

Rastgele seçilen deneklerin her kanalından 7000 adet örnek alınmıştır. Bu sayede, kayıt başlangıcı ve bitişindeki artefaktlardan arındırılmış temiz veriler elde edilmiştir. Veri dağılımının eşit olması için rastgele seçilen deneklerin verilerinin uzunluk kontrolü yapılmıştır. Uygun alt frekans bileşenine ayırmak için geçiş bandında maksimum seviyede düz bir frekans yanıtına sahip olan 5. dereceden Butterworth bant geçiren filtre kullanılmıştır. Bant geçiren filtrelerde bir alt kesme frekansı ve bir üst kesme frekansı bulunmaktadır. Eş. 1'de, bant geçiren Butterworth filtresinin frekans yanıtının genlik karesi gösterilmektedir. Burada ω açısal frekansı, ω_L alt kesme açısal frekansını, ω_H üst kesme açısal frekansını, n filtre derecesini, ve $H(j\omega)$ frekans yanıtını temsil etmektedir.

$$|H(j\omega)|^2 = \frac{1}{1 + \left(\frac{\omega^2 - \omega_L \omega_H}{\omega_H - \omega_L}\right)^{2n}} \quad (1)$$

EEG alt frekans bileşenlerinin sınırlarına göre delta (0-4 Hz), teta (4-8 Hz), alfa (8-16 Hz), beta (16-28 Hz) ve gama (28-32 Hz) dalgaları ayrıştırılarak Gini önem katsayılarına göre en yüksek doğruluğa sahip olan kanallar tespit edilmektedir. Gini önem katsayısı, karar ağaçları gibi topluluk makine öğrenmesi yöntemlerinde doğruluğa en yüksek etki eden özneliğin tespitinde kullanılan bir ölçektir. Temel olarak, Gini önem katsayısı Gini safsızlığı kullanarak düğümün saf olmayan bir veri kümesine sahip olup olmadığını tespit eder. Bir öznelik kullanarak bir düğüm bölündüğünde, Gini safsızlığının ne kadar iyileştiği incelenir. Eş. 2'de Gini safsızlığı gösterilmektedir. Burada C sınıf sayısını ve p_i i 'inci sınıfın olasılığını temsil etmektedir.

$$Gini = 1 - \sum_{i=1}^C p_i^2 \quad (2)$$

Genellikle öznelik seçimi olarak kullanılan Gini önem katsayısı, kanal seçiminde de kullanılarak boyut azaltması yapılmaktadır. Karar ağaçları tabanlı yaklaşımlarda bir düğümdeki safsızlığı azaltmak için veri kümesi, Gini saflaştırmasındaki azalma hesabına göre düğümde bir bölme yapılır. Eş. 3'te Gini saflaştırmasındaki azalma hesabı için kullanılan denklem gösterilmektedir. Burada $Gini_{orj}$ düğümün orijinal safsızlığını, $Gini_{sol}$ sol düğümün safsızlığını, $Gini_{sağ}$ sağ düğümün safsızlığını, N_{sol} ve $N_{sağ}$ sol-sağ düğümdeki örnek sayısını ve N toplam örnek sayısını temsil etmektedir.

$$\Delta Gini = Gini_{orj} - \left(\frac{N_{sol}}{N} Gini_{sol} + \frac{N_{sağ}}{N} Gini_{sağ} \right) \quad (3)$$

Safsızlıktaki azalma ne kadar büyükse, o öznelik daha önemlidir. Gini önem katsayısında her özneliğin önem katsayısını hesaplamak bir özneliğin model boyunca tüm düğümlerde sağladığı toplam saflaştırma toplamıdır.

3.3. Öznelik Çıkarımı (Feature Extraction)

Öznelik çıkarmanın temel amacı, sınıflandırıcının performansını iyileştirerek kimlik tanımlama için benzersiz desenler oluşturmaktır.

Butterworth fitresinden sonra ayrıştırılan frekans alt bileşenlerinden seçilen en yüksek Gini önem katsayısına sahip kanal, gözler kapalı dinlenme durumu verisi için veriler 140 örneklilik (0,875 saniye) pencerelere, gözler açık dinlenme durumu verisi için veriler 100 örneklilik (0,625 saniye) pencerelere bölünerek kanal öznelik çıkarımına uygun hale getirilmiştir. Öznelik çıkarımı sonrasında veri $x = [x_1, x_2, \dots, x_n]$ formunda öznelik vektörüne dönüştürülür. Öznelik vektöründe tekrar eden veriler kaldırılarak boyut indirgeme yapılabilir. Phinyomark'a göre biyosinyallerden çıkarılan öznelikler, zaman bölgesi, frekans bölgesi ve zaman-mekân bölgesi öznelikleri olarak üç türe ayrılır [30]. Zaman bölgesi öznelikleri, genellikle daha kolay ve daha hızlı öznelik çıkarma tekniklerine sahip olmasından dolayı biyosinyallerde sıklıkla kullanılmaktadır. Frekans öznelikleri gibi dönüşüme ihtiyaç duymaması, sinyalin genlik bilgisinin yeterli olması sağlar. Bu çalışmada maksimum, minimum, standart sapma, Skewness ve Kurtosis öznelikleri kullanılmaktadır. Eş. 4-Eş. 8 arasında bu özneliklerin denklemleri gösterilmektedir. Burada A_{ij} zaman serisi olarak EEG parçasını, \bar{A} EEG parçasının genlik ortalamasını ve N EEG parçasındaki örnek sayısını temsil etmektedir.

$$EEG_{maksimum} = maksimum(A_{ij}) \quad (4)$$

$$EEG_{minimum} = minimum(A_{ij}) \quad (5)$$

$$EEG_{standart\ sapma} = \sqrt{\frac{1}{N} \sum_{j=1}^N (x_j - \bar{x})^2} \quad (6)$$

$$EEG_{skewness} = \frac{\sum_{i=1}^N (A_{ij} - \bar{A}_{ij})^3 / N}{(EEG_{standart\ sapma}(A_{ij}))^3} \quad (7)$$

$$EEG_{kurtosis} = \frac{\sum_{i=1}^N (A_{ij} - \bar{A}_{ij})^4 / N}{(EEG_{standart\ sapma}(A_{ij}))^4} \quad (8)$$

3.4. Sınıflandırıcı Yapısı (Classifier Structure)

Tek bir makine öğrenmesi yöntemini kullanıldığı çalışmalarda, genellikle kullanılan algoritmaya bağlı olarak o algoritmanın yetersizlikleri hatalı sınıflandırmaya ve algoritmanın performansının azalmasına sebep olmaktadır. Bu sorunun önüne geçilebilmesi için, birden fazla makine öğrenmesi modelinin beraber kullanılarak ve en yoğun sonucun sınıflandırma sonucu olarak belirlendiği topluluk makine öğrenmesi yöntemi kullanılabilir. Rastgele ormanlar gibi, kendi başına topluluk makine öğrenmesi yöntemler bulunmaktadır. Bunlara ek olarak, farklı makine öğrenmesi yöntemlerini bir grup haline getirerek de topluluk makine öğrenmesi algoritması oluşturulabilir. Bu sayede daha güçlü ve güvenilir bir model oluşturulmaktadır. Farklı modeller, diğer modellerin yetersizliklerini gidererek sınıflandırma performansını arttırmaktadır. Topluluk yöntemleri, kullanılan algoritmaların çeşidine göre *bagging*, *boosting* ve *stacking* olmak üzere üç türe ayrılmaktadır. Rastgele ormanlar gibi, aynı algoritmanın farklı versiyonlarının birleştirildiği yöntemlere *bagging* denir. Gradyan artırma gibi, düşük performansla çalışan algoritmanın sonuçlarını arttırmak için aşamalı olarak birden fazla modelin eğitildiği yöntemlere *boosting* denir. Bu çalışmada olduğu gibi, farklı algoritmaların birlikte kullanıldığı ve sonuçların iyileştirildiği yöntemlere *stacking* denir. Bu çalışmada, rastgele ormanlar, çok katmanlı sinir ağı, gradyan artırma, karar ağaçları ve destek vektörleri makinesi algoritmalarından oluşan bir topluluk yöntemi önerilmektedir. Bu yöntemle, aşırı uyum, düşük doğruluk ve esneklik gibi sorunlara çözüm üretilmektedir. Önerilen bu *stacking* tabanlı topluluk makine öğrenmesi algoritması, baz model ve meta-model olmak üzere iki seviyeli bir yapıya sahiptir. Burada baz model olarak adlandırılan yapı sınıflandırıcı mimarilerinin bulunduğu seviyedir. Baz model kısmında tanımlı N sayıda sınıflandırıcı modelinin tanımlandığı varsayılırsa $X \in \mathbb{R}^{n \times m}$ eğitim verisini ve $y \in$

\mathbb{R}^n etiketleri, n veri sayısı, m öznitelik sayısını temsil ederse, Eş. 9'da aynı verilerle ayrı ayrı eğitilen bir baz modeli temsil edilmektedir.

$$\hat{y}_i = f_i(X), \quad i = 1, 2, \dots, N \quad (9)$$

Burada \hat{y} ile temsil edilen her modelin bir tahmini bulunmaktadır. Bu tahminler $\hat{y} \in \mathbb{R}^n$ şeklinde olup, test verileri sıralı bir şekilde modelleri beslemektedir. Her \hat{y}_i baz model çıktıları, ikinci seviyedeki meta-modelin girdisini oluşturmaktadır. Meta-modeller, baz modellerin çıktısını kullanarak tahmin sonuçlarından iyileştirilmiş nihai tahmini yapar. Eş. 10'da g ile temsil edilen meta-model, her baz modelin o anki test verisine karşılık tahmin sonuçlarından en sık tekrar eden tahmin sonucunu kullanarak nihai tahmini hesaplar:

$$\hat{y}_{meta} = g(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_N) \quad (10)$$

Bu sayede daha esnek bir algoritma elde edilerek, modellerin güçlü ve zayıf yönleri dengelenebilmektedir. Fakat, her modelin tek tek eğitilip test edilmesi, hesaplama maliyeti gibi sorunlara sebep olabilmektedir. Karar Ağaçları, veriyi dallara ayırarak sınıflandırma yapan hiyerarşik yapılardır. Eş. 11'de görüldüğü gibi bölünme noktaları bilgi kazancı (*Information Gain*) veya Gini katsayısı gibi ölçütlerle belirlenir:

$$Gini(p) = 1 - \sum_{k=1}^K p_k^2 \quad (11)$$

Burada p_k , düğümdeki k -inci sınıfa ait örneklerin oranıdır. EEG tabanlı kimlik tanımlamada karar ağaçlarının en büyük avantajı, yorumlanabilirlik sağlamasıdır. Hangi frekans bandı ya da özneliğin sınıflandırmaya katkı sağladığını doğrudan gösterebilmesi, biyometrik güvenlik sistemlerinde şeffaflık ve güvenilirlik açısından değerlidir. Rastgele Ormanlar, birden fazla karar ağacının birlikte çalıştığı topluluk tabanlı bir algoritmadır. Eğitim sırasında farklı alt örneklemeler (*bootstrap*) ve öznitelik alt kümeleri kullanılarak bağımsız karar ağaçları oluşturulur. Eş. 12'de gösterildiği gibi tahmin sürecinde bu ağaçların çıktıları çoğunluk oylaması veya ortalama ile birleştirilir:

$$\hat{y}_{RO} = mod(h_1(x), h_2(x), \dots, h_T(x)) \quad (12)$$

Burada $h_t(x)$ t -inci karar ağacının çıktısını, \hat{y}_{RO} ise nihai sınıf tahminini göstermektedir. EEG tabanlı kimlik doğrulamada, farklı deneklerin beyin sinyallerindeki varyasyonu tolere edebilmesi ve gürlütlü veriler karşısında dayanıklı sonuçlar üretebilmesi, bu yöntemin en önemli avantajları arasındadır. Çok Katmanlı Sinir Ağları giriş-gizli katmanlar-çıkış yapısından oluşur ve doğrusal dönüşüm ile aktivasyon fonksiyonlarının birleşimiyle Eş. 13'te ifade edildiği gibi çalışır:

$$a_j = f\left(\sum_{i=1}^n w_{ij}x_i + b_j\right) \quad (13)$$

Burada x_i , giriş öznitelikleri, w_{ij} , bağlantı ağırlıkları, b_j , sapma (*bias*) terimi ve $f(\cdot)$ aktivasyon fonksiyonunu temsil etmektedir. Çok katmanlı sinir ağları, doğrusal olmayan ilişkileri modellemede güçlüdür. EEG tabanlı kimlik tanımlamada, özellikle beynin karmaşık ve doğrusal olmayan dinamiklerini öğrenebilmesi sayesinde yüksek doğruluk oranlarına ulaşmakta ve bireyler arası ayırt edici özellikleri etkin şekilde yakalayabilmektedir. Gradyan Artırma, zayıf öğrenicilerin (genellikle karar ağaçlarının) ardışık olarak eğitilmesi ve hataların aşamalı olarak azaltılması mantığına dayanır. Güncelleme kuralı Eş. 14'te ifade edilmiştir:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \quad (14)$$

Burada $h_m(x)$ yeni eklenen zayıf öğrenici, γ_m öğrenme oranıdır. Bu yaklaşım sapmayı azaltarak tahmin performansını artırır. EEG tabanlı kimlik doğrulamada gradyan artırmanın avantajı, küçük fakat anlamlı

farklılıkları yakalayarak bireylerin ayırt edici frekans bileşenlerini öne çıkarması ve yüksek genelleme kabiliyeti sağlamasıdır. Destek Vektör Makineleri, sınıflar arasındaki ayrımı en iyi yapan hiperdüzlemi bulmayı hedefler. İki sınıf için optimizasyon problemi Eş. 15'te ifade edilmiştir:

$$\min_{w,b} = \frac{1}{2} \|w\|^2 \quad \text{öyle ki} \quad y_i(w \cdot x_i + b) \geq 1, \quad \forall i \quad (15)$$

Burada w ağırlık vektörü, b sapma, x_i giriş vektörleri ve y_i sınıf etiketleridir. Çekirdek fonksiyonları sayesinde doğrusal olmayan ayrımlar da yapılabilmektedir. EEG tabanlı kimlik tanımlamada destek vektör makinelerinin avantajı, özellikle az sayıda kanal veya sınırlı öznitelik kullanıldığında bile yüksek doğruluk sağlaması ve küçük veri kümelerinde genelleme başarısının yüksek olmasıdır.

4. Sonuçlar ve Tartışmalar (Results and Discussions)

Makalenin bu bölümünde gerçekleştirilen deneylerin sonuçları ve bu sonuçların analizleri karşılaştırmalı olarak paylaşılacaktır. Hem gözler açık hem de gözler kapalı olarak toplanan dinlenme durumu verileri kullanılarak gerçekleştirilen testlerde, özellikle eğitim-test süresi, doğruluk, kesinlik, hassasiyet ve F1 skoru gibi temel sınıflandırma metrikleri kullanılarak yaklaşımların performansları sistematik ve aşamalı olarak analiz edilecektir. EEG sinyali, içerdiği anlamlı koşullara dayalı olarak frekansına göre delta, teta, alfa, beta ve gama olmak üzere beş alt frekans bölgesine ayrılır. Deneyler öncesinde hangi frekans bandının daha yüksek doğruluğa ulaşabileceği ANOVA F testi kullanılarak tespit edilmiştir. ANOVA F testi, istatistiksel olarak yapılan incelemelerle birden fazla grubun ortalamalarını karşılaştırarak bu gruplar arasındaki farkların anlamlılığını kontrol eden bir tekniktir. Temel olarak grupların varyansları ve ortalamalarıyla ilgilenen bu yöntemin kullandığı iki hipotez bulunmaktadır. Sıfır hipotezi (null hypothesis), grupların ortalamalarının eşit olduğu durumda gruplar arasında anlamlı bir fark olmadığını savunurken, alternatif hipotezinde ortalamalar arasında fark olduğu durumda gruplar arasında anlamlı bir fark olduğunu savunur. ANOVA F testi, gruplar arası varyansın grup içi varyansa oranıyla hesaplanır. Bu oran, F değeri olarak adlandırılır. F değerinin ne kadar yüksek olması, gruplar arasındaki farkın o kadar büyük olduğunu gösterir. Tablo 2'de incelenen tüm frekans bantlarının F-değerleri gösterilmektedir. Gözler kapalı kayıt verilerinde, F-değerinin en yüksek olduğu derin uyku ve bilinçsiz durumlarla ilişkili delta bandının denekler arasında daha belirgin farklılıklara sahip olduğu görülmektedir. Gözler kapalı kayıt verilerinde en düşük F-değerine sahip bandın, aktif düşünme, problem çözme, stresli durumlarla ilişkili beta bandı olarak tespit edilmiştir. Gözler açık kayıt verilerinde de F-değerinin en yüksek olduğu delta bandının denekler arasında daha belirgin farklılıklara sahip olduğu görülmektedir. Gözler açık kayıt verilerinde de en düşük F-değerine sahip bant beta bandı olarak tespit edilmiştir. Delta, teta ve alfa bantları, yüksek F değerleri ile denekler arasında önemli farklılıklar oluşturmaktadır. Bu yüzden yapılacak deneyler ve testlerde bu frekans bantları, deneklerin biyometrik olarak ayırt edilebilmesinde güçlü bir sinyal kaynağı olabilir. Beta ve Gama bantları için düşük çıkan F-değerleri, bu bantlardaki varyansın çok düşük olduğu veya veri ile ilgili bir problem olduğunu göstermektedir. Bu nedenle, bu frekans bantlarının biyometrik sınıflandırma için uygun olmadığını göstermektedir. Biyometrik sınıflandırmada, hem ayırt edicilik gücü yüksek olan hem de anlamlı sonuçlar veren frekans bandını seçmek önemlidir. Bu yüzden bu aşamadan sonra gerçekleştirilen tüm deneylerde gözler açık ve gözler kapalı olarak yapılan testlerde delta bandı kullanılacaktır. Biyometrik sınıflandırmada doğru frekans bandının seçilmesi sistem performansı için kritik olduğundan, bu analizler deneyimizin temelini oluşturmaktadır. Ayrıca, Tablo 2'de yer alan ANOVA F testi sonuçları, gözler kapalı veri setinin gözler açık veri setine kıyasla daha yüksek F-değerlerine sahip olması nedeniyle,

Tablo 2. 109 denek için uygun frekans bandının belirlenmesine yönelik ANOVA F testi sonuçları
(ANOVA F-test results for determining the appropriate frequency band for 109 subjects)

Gözler Açık Kaydedilen Veri Seti		Gözler Kapalı Kaydedilen Veri Seti	
Frekans Bandı	F-değeri	Frekans Bandı	F-değeri
Delta Bandı (0,5 – 4 Hz)	327,584	Delta Bandı (0,5 – 4 Hz)	1712,436
Teta Bandı (4 – 8 Hz)	288,193	Teta Bandı (4 – 8 Hz)	1335,274
Alfa Bandı (8 – 12 Hz)	267,845	Alfa Bandı (8 – 12 Hz)	1578,621
Beta Bandı (12 – 30 Hz)	201,427	Beta Bandı (12 – 30 Hz)	952,318
Gama Bandı (30 – 40 Hz)	278,906	Gama Bandı (30 – 40 Hz)	1147,532

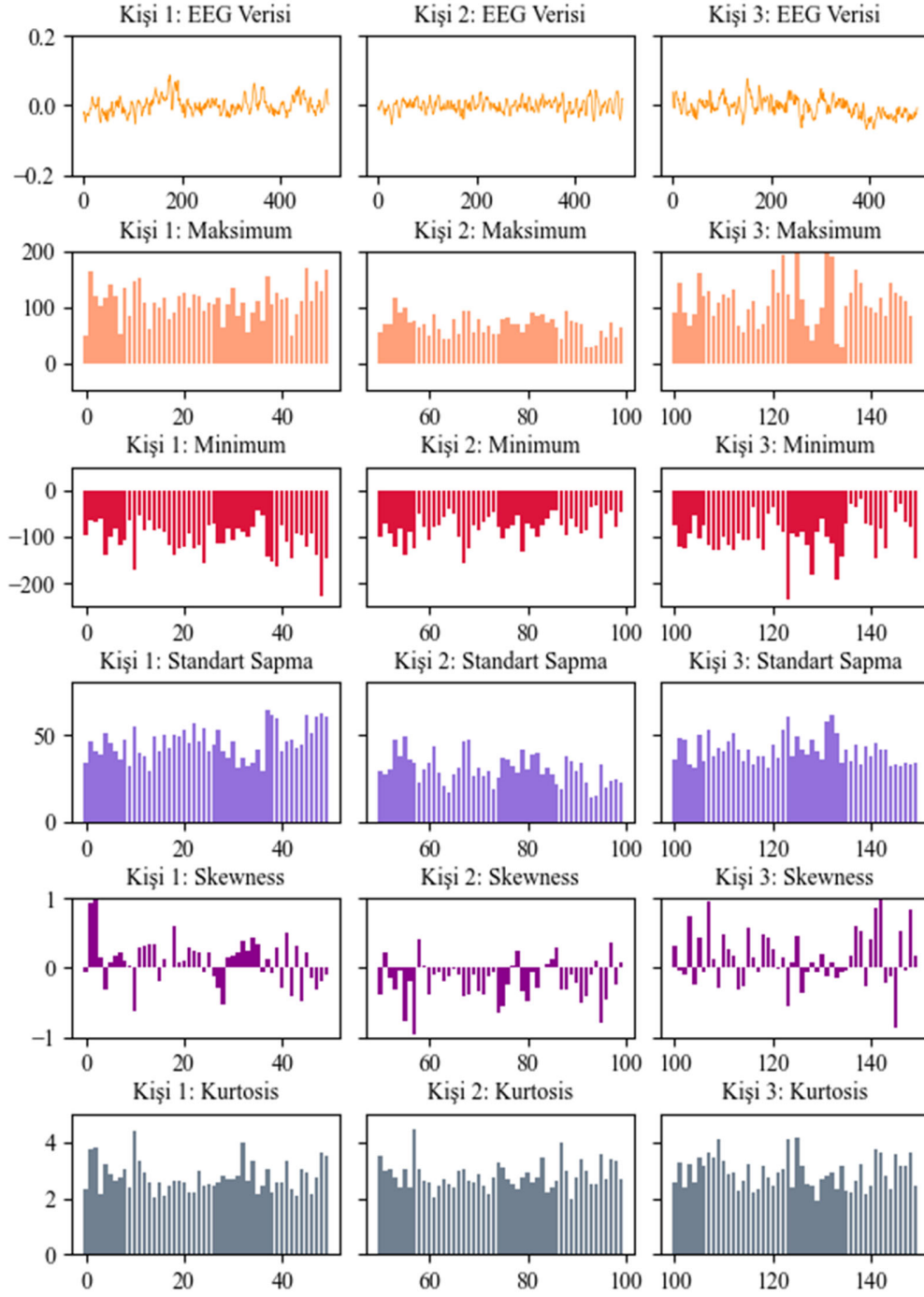
gözler kapalı koşulların biyometrik ayrımcılık için daha elverişli olduğunu göstermektedir. Bu durum, gözler açık veri setindeki dış uyaranlar ve artefaktların sinyal kalitesini düşürmesiyle açıklanabilir. Deneysel çalışmalarda kullanılan EEG motor hareketi/hayali veri seti, bu çalışmada olduğu gibi literatürde biyometrik kimlik tanımlama çalışmalarında sıklıkla kullanılmaktadır. Çoğu çalışmada tüm veri setinin kullanıldığı görülmektedir. Fakat bazı verilerde, özellikle kayıt başlangıcında ve bitişinde yüksek gürültü seviyeleri tespit edilmiştir. Bu yüzden, bu çalışmada bu veri setinde yer alan görece daha az gürültüye sahip olan 8 denekten alınan veriler kullanılarak bir tek kanallı EEG tabanlı biyometrik kimlik tanıma sistemi araştırılmaktadır. Gerçekleştirilen ilk deneyde, öncelikle Gini önem katsayısının EEG tabanlı biyometrik sistemler üzerindeki kanal seçimindeki performansının geçerli olup olmadığını ölçen testler gerçekleştirilmiştir. Şekil 4'te hem gözler açık hem de gözler kapalı olarak dinlenen durumda kaydedilen verilerin önem katsayılarının bar grafiği paylaşılmaktadır. Bu grafik incelendiğinde, gözler açık olarak kaydedilen veri setinde en yüksek önem katsayısına sahip olan kanallar Ft7, Cp4 ve C2'dir. Gözler kapalı olarak kaydedilen veri setinde en yüksek önem katsayısına sahip olan kanallar Cp4, C2 ve T9'dur. Beyin lobları, beynin farklı bölgelerinde farklı işlevleri yerine getirmek için özelleşmiş bölgelerdir ve bu kanalların işlevleri, sınıflandırmada neden önemli olduklarını anlamamıza yardımcı olacaktır. Bu kanallardan Ft7 kanalı, motor fonksiyonlar, karar verme, dikkat ve duygusal tepkilerle ilişkili olan frontal lobda; T9 kanalı, işitsel uyarılar, hafıza süreçleri ve duygusal tepkilerle ilişkili; Cp4 ve C2 kanalları ise, somatosensoryel (dokunsal) işlemler ve mekânsal farkındalık ile ilişkili parietal lobda yer almaktadır. Biyometrik kimlik tanımlama sistemi için, bu kanalların farklı kişiler arasındaki ayırt edici nöral aktiviteleri en iyi yakaladığı Gini önem katsayıları ile kanıtlanmaktadır. Bu kanalların doğrulukları, önem katsayılarıyla doğru orantılı olarak değişmektedir.

Sonraki deneyde, tüm kanallar kullanılarak gerçekleştirilen sınıflandırma sonuçları incelenmiştir. Hem gözler kapalı hem de gözler açık olarak kaydedilen veri setlerinde, doğruluklar incelendiğinde, gözler kapalı olarak gerçekleştirilen testlerde daha yüksek doğruluklara ulaşıldığı görülmektedir. Ancak, bu deneyde 64 kanalın tamamının kullanılması hem maliyet hem de pratik kullanımı açısından yüksek maliyetli olacaktır. Algoritmalar, belirli oranlarda eğitim ve test verilerine ihtiyaç duyar. Gerçekleştirilen deneylerde, deneklerden toplanan her verinin %80'i eğitim ve %20'si test verisi olarak bölünmektedir. Bölünen bu veriler, kayıt cihazının örnekleme frekansı (yani, saniye başına veri yakalama hızı) olan 160 Hz'e ve veri sayısının tam bölünmesine uygun olmak üzere, gözler açık veri seti için 140 örnekleme pencerele (her biri belirli süreyi kapsayan veri parçaları), gözler kapalı veri seti için 100 örnekleme pencerele ayrıştırılarak, öznitelik çıkarma katmanına gönderilmektedir. Ayrıştırılan pencerelerden öznitelik çıkarımı yapılmadan önce, ham verinin kısa bir örneği Şekil 3'te gösterilmiştir. Bu örnekte, üç farklı deneğe ait EEG sinyalleri pencerelenmiş şekilde sunulmuştur. Ardından, her pencere için seçilen istatistiksel öznitelikler olan minimum, maksimum, skewness, standart sapma ve kurtosis değerleri hesaplanmış ve aynı şekilde Şekil 3 içerisinde bar grafiklerle

görselleştirilmiştir. Böylece, öznitelik çıkarımı sonrası her pencerenin sayısal olarak temsil edilebilmesi ve bireyler arası farklılıkların gözlemlenebilmesi sağlanmıştır.

Sınıflandırma doğruluğu, testlerde doğru tanımlanan denek sayısının toplam test sayısına oranı olarak hesaplanmıştır. İşlem karmaşıklığının azaltılması için zaman bölgesinde çıkarılan özniteliklerin etkisini değerlendirmek amacıyla Tablo 3'te, 64 kanallı EEG veri seti kullanılarak 5-fold çapraz geçerlilik yöntemiyle elde edilen deneysel sonuçlar sunulmaktadır. Elde edilen bulgular, gözler açık ve gözler kapalı koşulları arasında belirgin performans farklılıkları olduğunu ortaya koymaktadır. Gözler açık durumda, göz kırpması ve oküler kas hareketleri sinyallere gürültü olarak yansımakta, bu da sınıflandırma başarımını sınırlamaktadır. Buna karşılık, gözler kapalı durumda kas kaynaklı artefaktların azalması daha temiz EEG sinyalleri elde edilmesine imkân tanımakta ve özellikle güçlü öğrenme algoritmalarında sınıflandırma doğruluğunu artırmaktadır. Algoritmalar özelinde incelendiğinde, en yüksek doğruluk oranına gözler kapalı veri setinde Çok Katmanlı Sinir Ağı (MLP) algoritması ile %98,9 ± 0,1 değerine ulaşılmıştır. MLP, gözler açık verilerde de %98,6 ± 0,2 doğrulukla en başarılı algoritma olmuştur. Rastgele Ormanlar her iki durumda da yüksek ve tutarlı sonuçlar göstermiş; gözler açıkta %97,8 ± 0,3, gözler kapalıda ise %98,3 ± 0,2 doğruluk elde ederek MLP'nin ardından en başarılı algoritma olmuştur. SVM, Karar Ağaçları ve kNN algoritmaları %86–89 aralığında doğruluk değerleriyle orta düzeyde performans sergilemiş, gözler açık ve kapalı durumlar arasında kayda değer bir değişim göstermemiştir. Öte yandan, Gradyan Artırma (GBC) algoritması gözler açıkta %94,1 ± 0,4 doğruluk elde etmişken, gözler kapalıda ise %86,5 ± 0,6'ya gerileyerek performansında belirgin bir düşüş yaşamıştır. Lojistik Regresyon ve Bayes tabanlı sınıflandırıcılar ise her iki koşulda da düşük doğruluk oranlarıyla dikkat çekmiş, karmaşık EEG verilerini ayırt etmede yetersiz kalmıştır. Genel olarak sonuçlar, gözler kapalı kayıt koşullarının artefaktların azalması sayesinde daha temiz sinyal sağladığını ve bu durumun özellikle MLP ve Rastgele Ormanlar gibi gelişmiş yöntemlerde sınıflandırma başarımını anlamlı ölçüde artırdığını göstermektedir. Bununla birlikte, her algoritma bu avantajdan eşit şekilde yararlanamamış; özellikle GBC örneğinde olduğu gibi, daha düşük varyanslı veri koşullarında genelleme kapasitesinin sınırlı kaldığı görülmüştür. Bu bulgular, EEG tabanlı sınıflandırma çalışmalarında kayıt koşullarının algoritma performansı üzerindeki belirleyici rolünü ortaya koymaktadır.

Önerilen topluluk sınıflandırma yaklaşımı, *stacking* yöntemine dayalı olarak tasarlanmıştır. Bu yöntemde, farklı makine öğrenmesi algoritmalarının sınıflandırma performanslarındaki çeşitlilikten yararlanılarak daha güçlü ve genelleme kabiliyeti yüksek bir model elde edilmesi amaçlanmaktadır. Önerilen modelde temel sınıflandırıcılar olarak Rastgele Ormanlar, Çok Katmanlı Sinir Ağı, Gradyan Artırma, Karar Ağaçları ve Destek Vektör Makineleri kullanılmıştır. Bu modellerin her biri farklı öğrenme stratejilerine, parametrik yapılarına ve karar mekanizmalarına sahip olduğundan, birlikte değerlendirilmeleri sınıflandırıcı çeşitliliğini artırmakta ve daha kararlı sonuçların elde edilmesini sağlamaktadır.



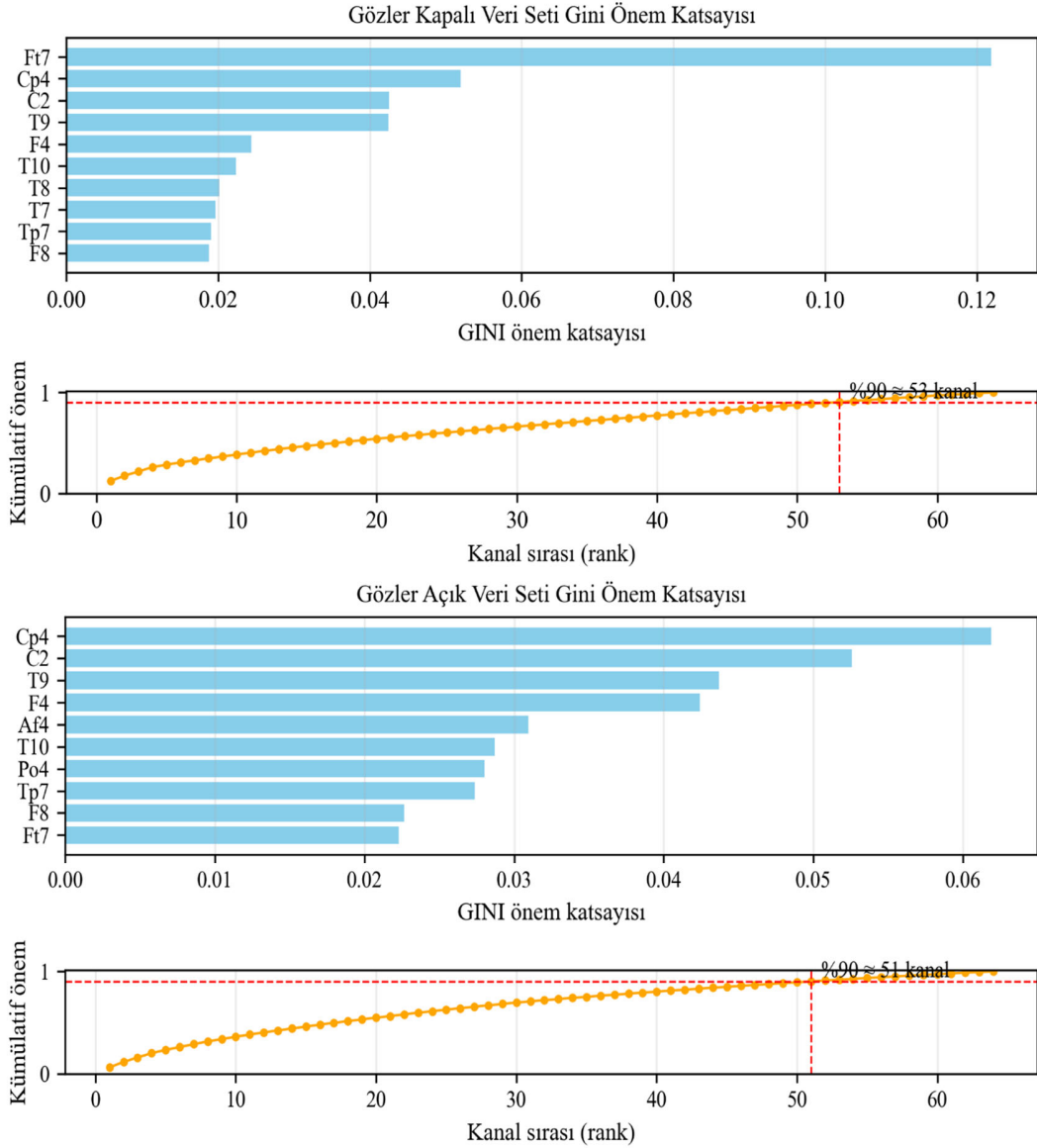
Şekil 3. Üç farklı denekten alınan EEG sinyali ham veri kesiti ve bu kişiden elde edilen istatistiksel öznitelikler
(Raw EEG signal segment from three different subjects and the statistical features extracted from them)

Stacking yaklaşımı, birinci aşamada temel sınıflandırıcıların bağımsız olarak eğitilmesi ve her birinin ürettiği tahminlerin kaydedilmesi üzerine kuruludur. İkinci aşamada ise bu tahminler, meta-öğrenici (meta-learner) olarak adlandırılan üst seviye bir modelin girdisi haline getirilmekte ve nihai sınıflandırma bu model aracılığıyla yapılmaktadır. Böylelikle temel sınıflandırıcıların güçlü yönleri birleştirilirken, zayıf yönlerinin etkisi azaltılmakta ve genel sınıflandırma başarımı iyileştirilmektedir.

Temel sınıflandırıcıların etkinliğini en üst düzeye çıkarmak amacıyla hiperparametre optimizasyonu süreci gerçekleştirilmiştir. Hiperparametre optimizasyonu, her bir algoritmanın model karmaşıklığı, öğrenme hızı, düzenleme katsayıları, ağaç derinliği veya katman boyutları gibi kritik parametrelerinin sistematik biçimde ayarlanması sürecini içermektedir. Bu sayede, modellerin yalnızca eğitim verisi üzerindeki doğruluğu değil, aynı zamanda bilinmeyen veriler üzerindeki genelleme performansı da artırılmıştır.

Tablo 3. 64 kanallı EEG veri seti ile gerçekleştirilen deneysel sonuçlar (Experimental results obtained using the 64-channel EEG dataset)

Gözler Açık Veri Seti Deneşleri					Gözler Kapalı Veri Seti Deneşleri				
Algoritma	Doğr.	Kesk.	Hass.	F1 Skoru	Algoritma	Doğr.	Kesk.	Hass.	F1 Skoru
SVM	%88,2±0,5	%88,4±0,4	%88,1±0,5	%88,2±0,4	SVM	%89,0±0,5	%89,1±0,4	%88,9±0,5	%89,0±0,4
Rastgele Orm.	%97,8±0,3	%98,5±0,2	%97,9±0,3	%97,9±0,2	Rastgele Orm.	%98,3±0,2	%98,4±0,2	%98,3±0,2	%98,3±0,2
MLP	%98,6±0,2	%98,7±0,2	%98,6±0,2	%98,6±0,2	MLP	%98,9±0,1	%98,9±0,1	%98,9±0,1	%98,9±0,1
GBC	%94,1±0,4	%94,3±0,4	%94,1±0,4	%94,2±0,4	GBC	%86,5±0,6	%87,0±0,5	%86,4±0,6	%86,6±0,5
Log. Reg. Sın.	%21,0±1,1	%23,1±1,0	%20,9±1,1	%21,2±0,9	Log. Reg. Sın.	%16,0±0,8	%15,5±0,9	%15,7±0,7	%15,9±0,6
Bayes Sın.	%14,3±0,8	%18,9±0,9	%14,2±0,8	%11,2±0,7	Bayes Sın.	%12,5±0,6	%15,2±0,7	%12,4±0,6	%8,1±0,5
Karar Ağacı	%89,0±0,5	%89,2±0,4	%88,9±0,5	%89,0±0,4	Karar Ağacı	%88,2±0,5	%88,3±0,4	%88,1±0,5	%88,2±0,4
kNN	%87,2±0,6	%88,5±0,5	%87,1±0,6	%87,4±0,5	kNN	%86,1±0,6	%88,2±0,5	%86,0±0,6	%86,5±0,5

**Şekil 4.** En yüksek Gini önem katsayısına sahip 10 kanalın grafikleri (Plots of the 10 channels with the highest Gini importance coefficients)

Tablo 4'te kullanılan temel algoritmalar için oluşturulan örnek hiperparametre havuzu sunulmaktadır. Bu havuz, hiperparametre optimizasyon sürecinde model yapılandırmasının esnekliğini

sağlamak ve farklı senaryolarda en uygun parametre kombinasyonlarının belirlenmesine imkân tanımak amacıyla hazırlanmıştır.

Tablo 4. Stacking yaklaşımında temel algoritmalar için belirlenen hiperparametre havuzu ve seçilen parametreler (Hyperparameter Pool and Selected Parameters for Base Algorithms in the Stacking Approach)

Algoritma	Hiperparametre	Denenen Değer Aralıkları	Seçilen Parametre
Rastgele Ormanlar	Ağaç sayısı	100, 200, 500	200
	Maksimum derinlik	5, 10, 20, None	10
	Minimum yaprak sayısı	1, 2, 5	5
Çok Katmanlı Sinir Ağı	Öğrenme oranı	0.001, 0.01, 0.05	
	Gizli katman boyutları	(50.), (100.), (100, 50), (100,50,100), (200, 250, 250, 200), (200, 350, 320, 200), (300, 400, 400, 300)	(200, 350, 320, 200)
	Aktivasyon fonksiyonu	ReLU, tanh, logistic	ReLU
Gradyan Artırma	Ağaç sayısı	100, 200, 300	100
	Öğrenme oranı	0.01, 0.05, 0.1	0.1
	Maksimum derinlik	3, 5, 10	3
Karar Ağaçları	Maksimum derinlik	5, 10, 20, None	10
	Bölünme kriteri	Gini, entropy	Entropy
	Minimum yaprak sayısı	1, 2, 5	5
Destek Vektör Makineleri	Kernel fonksiyonu	Linear, Poly, RBF	Linear
	Düzenleme katsayısı	0.1, 1, 10, 100	10
	Gamma	Scale, auto, 0.01, 0.001	Auto

Tablo 5. Tek kanal EEG veri seti ile gerçekleştirilen deneysel sonuçlar (Experimental results obtained using the single-channel EEG dataset)

Gözler Açık Veri Seti Deneyleri					Gözler Kapalı Veri Seti Deneyleri				
Algoritma	Doğr.	Kesk.	Hass.	F1 Skoru	Algoritma	Doğr.	Kesk.	Hass.	F1 Skoru
SVM	%87,6±0,5	%81,3±0,6	%87,4±0,4	%84,2±0,4	SVM	%93,6±0,7	%95,1±0,6	%93,7±0,8	%94,4±0,6
Rastg. Orm.	%83,5±0,6	%83,6±0,5	%83,4±0,5	%83,5±0,4	Rastg. Orm.	%96,3±0,5	%96,8±0,5	%96,2±0,6	%96,5±0,5
MLP	%79,2±0,8	%79,1±0,7	%79,3±0,6	%79,2±0,6	MLP	%98,8±0,3	%99,0±0,3	%98,7±0,4	%98,9±0,3
GBC	%83,4±0,5	%83,5±0,5	%83,3±0,4	%83,4±0,4	GBC	%96,2±0,6	%96,7±0,5	%96,3±0,6	%96,5±0,5
Log. Regr.	%79,3±0,7	%77,5±0,8	%79,4±0,6	%78,4±0,5	Log. Regr.	%83,9±1,0	%88,0±0,9	%83,7±1,1	%85,8±0,8
Bayes Sin.	%37,8±1,0	%16,6±0,7	%37,6±1,1	%22,6±0,8	Bayes Sin.	%65,2±1,5	%65,6±1,2	%65,1±1,3	%64,0±1,0
Karar Ağacı	%83,4±0,6	%83,3±0,6	%83,5±0,5	%83,4±0,5	Karar Ağacı	%87,6±0,9	%93,0±0,8	%87,5±1,0	%90,1±0,8
kNN	%75,2±0,9	%71,8±0,8	%75,1±0,7	%73,4±0,6	kNN	%77,6±1,1	%82,3±1,0	%77,4±1,2	%79,7±0,9
Önerilen	%84,1±0,5	%83,7±0,6	%84,4±0,5	%84,0±0,5	Önerilen	%100±0,0	%100±0,0	%100±0,0	%100±0,0

Gözler açık ve kapalı olarak kaydedilen verilerde, seçilen tek kanal ile gerçekleştirilen test sonuçları Tablo 5'te sunulmaktadır. Gözler açık kaydedilen verilerde, önerilen sistem dışında kullanılan geleneksel makine öğrenmesi algoritmalarında sınıflandırma doğruluğu %37,8 ile %87,6 arasında değişirken, önerilen sistem ile %84,1 ± 0,5 doğruluğa ulaşılmıştır. Bu sonuç, önerilen topluluk sınıflandırıcısının bazı algoritmalar (örneğin kNN ve Bayes) üzerinde belirgin bir üstünlük sağladığını, ancak destek vektör makineleri algoritmasının ulaştığı %87,6 ± 0,5'lik ortalama doğruluk değerine kıyasla daha düşük kaldığını göstermektedir. Gözler kapalı kaydedilen verilerde ise, önerilen sistem dışında kullanılan geleneksel makine öğrenmesi algoritmalarında sınıflandırma doğruluğu %65,2 ile %98,8 arasında değişmiştir. Bu durumda, Rastgele Ormanlar (%96,3 ± 0,5), Gradyan Artırma (%96,2 ± 0,6) ve Çok Katmanlı Sinir Ağı (%98,8 ± 0,3) oldukça yüksek performans göstermiştir. Ancak önerilen topluluk sınıflandırıcısı, tüm bu yöntemlerden daha üstün bir performans sergileyerek %100 ± 0,0 doğruluk elde etmiştir. Genel olarak sonuçlar, gözler kapalı verilerde artefaktların azalması sayesinde hem geleneksel algoritmaların hem de önerilen sistemin daha yüksek doğruluk değerlerine ulaştığını; özellikle önerilen sistemin, %100 sınıflandırma başarımı ile tek kanal verisinde dahi güvenilir bir biyometrik kimlik doğrulama aracı olabileceğini ortaya koymaktadır.

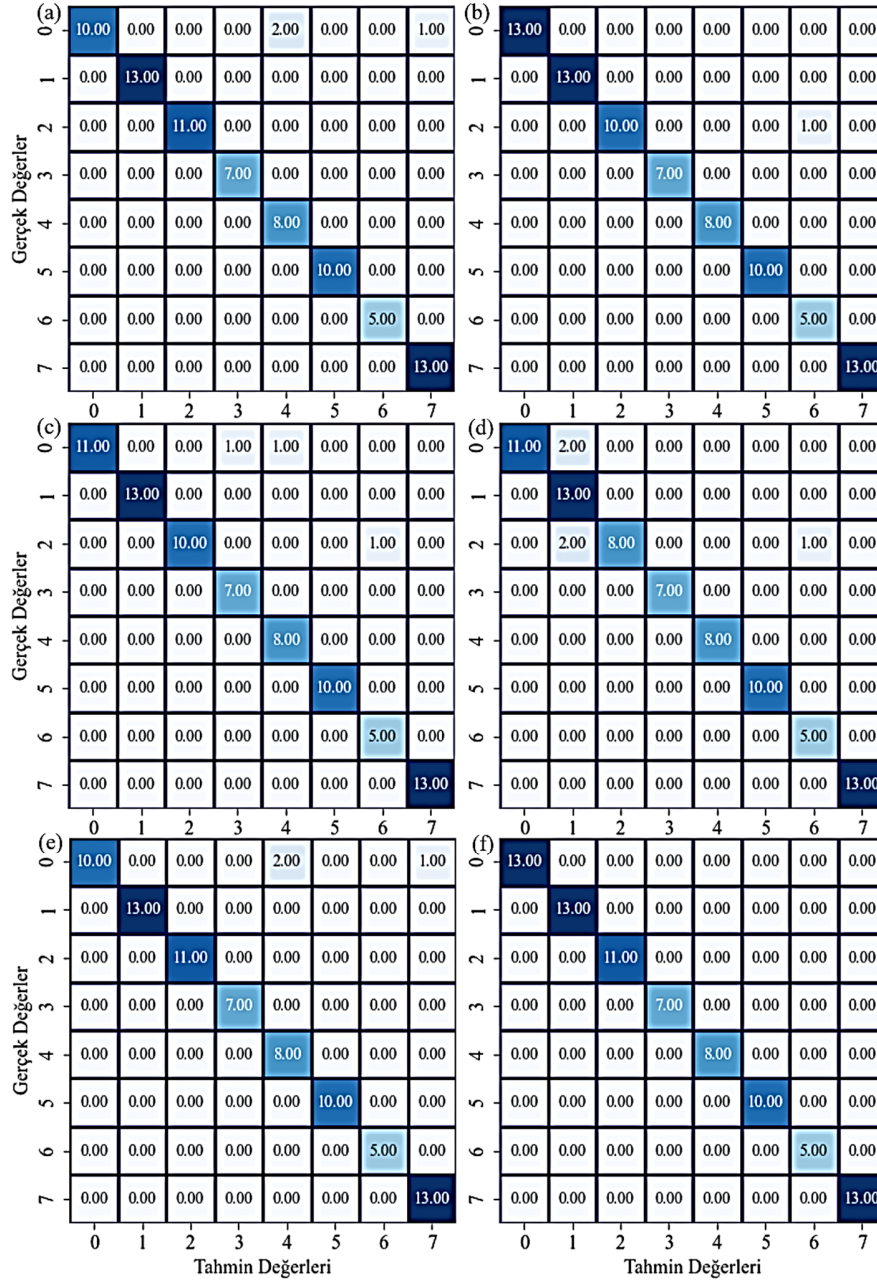
Test sonuçlarının daha kapsamlı değerlendirilebilmesi amacıyla, sınıflandırma performansını ayrıntılı olarak ortaya koyan karmaşıklık matrisleri Şekil 5'te sunulmaktadır. Bu matrisler, modelin %80 eğitim ve %20 test verisi ayrımla gerçekleştirilen deneyler sonucunda elde edilmiştir. Karmaşıklık matrisleri aracılığıyla yalnızca genel doğruluk oranı değil, aynı zamanda doğru ve yanlış sınıflandırmaların dağılımı da incelenebilmekte; böylelikle modelin hangi sınıflarda güçlü performans sergilediği ve hangi sınıflarda hata yapma eğiliminde olduğu ayrıntılı bir biçimde ortaya konmaktadır. Bu sonuçlarla birlikte, Tablo 6'da 64 kanallı ve tek kanallı modellerin eğitim ve test performansları ve süreleri karşılaştırılmalı olarak verilmiştir. Rastgele

Ormanlar, Çok Katmanlı Sinir Ağı, Gradyan Artırma, Karar Ağaçları ve Destek Vektör Makineleri modellerinden oluşan önerilen topluluk sınıflandırıcısı, tek kanallı geleneksel modellere göre hem eğitim hem de test süresi bakımından daha yavaş olmasına rağmen, 5 parçalı çapraz geçerlilik yöntemiyle elde edilen ortalama sonuçlara göre en yakın sınıflandırıcıya kıyasla %1,2 daha yüksek doğruluk sağlamıştır. Yapılan bu deneyler sonucunda, önerilen sistemin araştırmacılar açısından umut verici bulgular sunduğu söylenebilir. Çalışmadaki en büyük sorunlardan biri olan, daha yüksek sınıflandırma doğruluğu için 64 kanallı veriye ihtiyaç duyulması, önerilen bu sistem ile aşılmıştır. Bu yaklaşım sayesinde hem maliyet hem de pratiklik açısından önemli bir iyileştirme sağlanmış, doğruluk yüzdeleri ile eğitim-test sürelerindeki karşılaştırmalar bu durumu açıkça ortaya koymuştur.

Eğitim ve test aşamalarında geçen sürelerin yanında, bu süreçlerde kullanılan verilerin uzunluğu da sistemin performansını etkileyen önemli bir ölçüttür. Genel olarak, biyometrik sistemlerin kullanıcı tanımlaması uygulamalarında, sistemin pratikliğinin savunulabilmesi için tanımlamanın birkaç saniye içinde gerçekleştirilmesi beklenmektedir. Artık veriler oluşmaması için, pencerede yer alan örnek sayıları hem eğitim hem de test verilerini tam bölecek şekilde belirlenmiştir. Gözler kapalı olarak kaydedilen verilerde, her kullanıcı için 7000'er örnek alınmış olup, bu verilerin %80'i sistemdeki algoritmaların eğitiminde, %20'si ise test için kullanılmaktadır. Kayıt cihazının örnekleme frekansının 160 Hz olduğu ve her pencere için 140 örnek alındığı göz önünde bulundurulursa, her kullanıcıdan 35 saniyelik eğitim verisi ve 8,75 saniyelik test verisi kullanılarak bu doğruluklara ulaşılması, sistemin gözler kapalı kullanımının pratikliğini kanıtlamaktadır. Aynı oranların kullanıldığı gözler açık olarak kaydedilen verilerde, her denekten toplanan 1500 örnek için 7,5 saniyelik eğitim verisi ve 1,875 saniyelik test verisinin kullanılması, bu sistemin de pratikliğini kanıtlamaktadır.

Tablo 6. 64 kanallı ve tek kanallı EEG veri seti kullanılan deney eğitim ve test süreleri (saniye)
(Training and testing times for experiments using 64-channel and single-channel EEG datasets)

Algoritma	Tek Kanal Gözler Açık		Tek Kanal Gözler Kapalı		64 Kanal Gözler Açık		64 Kanal Gözler Kapalı	
	Eğitim (s)	Test (s)	Eğitim (s)	Test (s)	Eğitim (s)	Test (s)	Eğitim (s)	Test (s)
SVM	0,00201	0,00001	0,03356	0,00101	298,728	2,72541	379,427	3,75486
Rastgele Orm.	0,12123	0,00789	0,17746	0,01020	26,2796	0,08503	172,262	0,36341
MLP	0,47703	0,00099	1,54007	0,00198	27,8219	0,02326	91,4836	0,08928
GBC	0,23400	0,00101	0,41674	0,00120	97,3134	0,02884	525,344	0,08726
Log. Reg. Sm.	0,01490	0,00110	0,03366	0,00100	0,42677	0,00226	1,60048	0,00189
Bayes Sm.	0,00101	0,00103	0,00099	0,00100	0,00847	0,00405	0,05198	0,00902
Karar Ağaçları	0,00199	0,00001	0,00200	0,00001	1,57607	0,00153	9,44690	0,00511
kNN	0,00189	0,00211	0,00201	0,00198	0,00465	0,28102	5,74782	0,01656
<i>Önerilen Topl.</i>	<i>0,83626</i>	<i>0,00991</i>	<i>2,16983</i>	<i>0,01440</i>	-	-	-	-

**Şekil 5.** Önerilen topluluk sınıflandırıcısında kullanılan algoritmalara ait karmaşıklık matrisleri: a) RF, b) MLP, c) GBC, d) SVM, e) DT ve f) Önerilen topluluk sınıflandırıcısı (Confusion matrices of the algorithms employed in the proposed ensemble classifier: a) RF, b) MLP, c) GBC, d) SVM, e) DT, and f) Proposed ensemble classifier)

Rocca vd., beyindeki fonksiyonel bağlantısallık gibi karmaşık bir yapı üzerinde, farklı senaryolarda %95-99 arasında doğruluk seviyesine ulaşmıştır [31]. Fonksiyonel bağlantısallık gibi araştırmalarda genellikle çok sayıda elektrod yerleştirilmesi gerekmektedir. Bu durum, sistemin genel tasarımını ve maliyetini önemli ölçüde etkilemektedir. Benzer bir yapıya sahip çalışmada, Frascini vd. faz gecikmeli bir yaklaşımla beyindeki bağlantısallık yapısını kullanarak, gama ve beta bandında kimlik tanımlamada dinlenme durumu için 64 kanallı veri setiyle en yüksek doğruluğa ulaşıldığını kanıtlamaktadır [16].

Benzer çalışmalara göre, özellikle hesaplama maliyetleri ve kullanılan kanal sayısı açısından önerilen sistemin yüksek bir avantajı bulunmaktadır. Daha az sayıda kanal kullanılarak gerçekleştirilen çalışmalarda farklı görevler kullanılabilir. Hema vd., EEG kaydı sırasında deneklere uyguladıkları çarpma işlemi ve metin okuma görevleriyle, iki kanal kullanarak %91,6-98,3 arasında doğruluk seviyesine ulaşmıştır [32]. Bu çalışmada, görevler göz önünde bulundurularak beta bandında daha yüksek doğruluğa ulaşılabileceği savunulmuştur. Bu çalışma ve önerilen sistem karşılaştırıldığında, özellikle tanımlanan bu görevlerin sistemin başlatıcısı olarak çalışması, yine bir hesaplama karmaşıklığına sebep olmaktadır. Bir başka karşılaştırma metriği olan eğitim-test verisi süresi konusunda, literatürde bu alanda en yüksek başarıya ulaşmış olan çalışmada, Safont vd. test verileri için 20. saniyeden sonra %0'a yakın eşleşme hatası oranına (EER) ulaşmaktadır [33]. İki test grubu için 8,75 saniyelik ve 1,875 saniyelik test verisi süreleriyle daha verimli ve pratik uygulama potansiyeli sonuçlarına ulaşılmaktadır. Ayrıca, aynı veri setini kullanan farklı çalışmalarda, %99,5 gibi yüksek doğruluğa ulaşılırken motor hareketleri kullanılmaktadır [34]. Ancak, bu durum fiziksel olarak engelli bireyler için dezavantaj

oluşturabileceği bilindiği için, benzer doğruluklara dinlenme durumu verileriyle ulaşılmaya çalışılmıştır. Tablo 7'de, önerilen sistemin farklı deney grupları üzerinde (8, 16, 24, 50 ve 109 denek) elde edilen sınıflandırma performansları özetlenmektedir. Denek sayısındaki artış, sistemin genelleme kabiliyetini ve istatistiksel güvenilirliğini değerlendirmek açısından kritik bir ölçüt olarak ele alınmıştır. Tek kanallı sistemde küçük ölçekli grup (8 denek) ile başlatılan analizler, daha geniş katılımcı gruplarına doğru genişletilmiş ve farklı grup büyüklüklerinde doğruluk oranlarının istikrarlı biçimde korunabildiği gözlemlenmiştir. Özellikle denek sayısının artmasıyla birlikte, sistemin hesaplama maliyetinde belirgin bir artış olmadan yüksek performansın sürdürülebilmesi, önerilen yaklaşımın ölçeklenebilirliğini ve pratik uygulanabilirliğini desteklemektedir. Bu bağlamda, elde edilen sonuçlar hem küçük hem de geniş örneklem gruplarında önerilen sistemin etkinliğini ortaya koymakta ve literatürdeki benzer çalışmalara kıyasla güçlü bir avantaj sunduğunu göstermektedir.

Tablo 8'te yapay zekâ tabanlı biyometrik kimlik doğrulama tekniklerine ilişkin güncel çalışmalar karşılaştırmalı olarak sunulmaktadır. Literatürde farklı yöntemlerin kullanıldığı ve doğruluk oranlarının oldukça yüksek seviyelere ulaştığı görülmektedir. Örneğin, Rastgele Orman tabanlı öznelik seçimi ile kanal optimizasyonu yapan çalışmada gözler kapalı durumda %99,96, gözler açık durumda ise %99,70 doğruluk elde edilmiştir [35]. Birleştirilmiş Evrişimsel Sinir Ağı ve Destek Vektör Makineleri tabanlı yaklaşımla %98,49 doğruluk sağlanmıştır [36]. Yalnızca iki kanal ve 5 saniyelik kısa bir zaman penceresi kullanılarak CNN tabanlı sınıflandırıcı ile %99 doğruluk oranına ulaşılmıştır [37]. Fonksiyonel bağlantısallık temelli özellik çıkarımı ve Çok Akışlı Grafik Evrişimsel Sinir Ağı sınıflandırıcısı ile gerçekleştirilen

Tablo 7. Farklı denek sayılarında (8, 16, 24, 50 ve 109) önerilen tek kanallı sistemin sınıflandırma performans sonuçları
(Classification performance results of the proposed single-channel system across different subject sizes (8, 16, 24, 50, and 109))

Tek Kanal Gözler Kapalı Veri Seti					
Metrik	8 Denek	16 Denek	24 Denek	50 Denek	109 Denek
Doğruluk	%100,0±0,0	%99,459±0,04	%99,293±0,03	%98,832±0,06	%98,264±0,03
Keskinlik	%100,0±0,0	%99,438±0,05	%99,262±0,05	%98,810±0,08	%98,218±0,10
Hassasiyet	%100,0±0,0	%99,460±0,06	%99,268±0,04	%98,794±0,08	%98,190±0,08
Tek Kanal Gözler Açık Veri Seti					
Metrik	8 Denek	16 Denek	24 Denek	50 Denek	109 Denek
Doğruluk	%84,145±0,46	%83,145±0,52	%81,653±0,48	%80,742±0,44	%80,025±0,41
Keskinlik	%83,747±0,58	%83,026±0,55	%81,428±0,47	%80,742±0,43	%79,864±0,40
Hassasiyet	%84,439±0,49	%83,245±0,50	%81,625±0,46	%80,742±0,42	%80,225±0,39

Tablo 8. Yapay zekâ tabanlı biyometrik kimlik tanımlama tekniklerinin karşılaştırması
(Comparison of artificial intelligence-based biometric identification techniques)

Çalışma	Yıl	Denek sayısı	Yöntem	Performans
[35]	2024	8	Önerilen sistem, en yüksek doğruluk için optimum kanal sayısını belirlemede Rastgele Orman tabanlı öznelik seçimi kullanılmaktadır.	Gözler kapalı doğruluk: %99,96. Gözler açık doğruluk: %99,70
[36]	2022	109	Birleştirilmiş evrişimsel sinir ağı ve destek vektör makineleri tabanlı sınıflandırıcı ile yüksek doğruluğa ulaşmak amaçlanmıştır.	Doğruluk: %98,49
[37]	2023	109	Sistem, yalnızca iki EEG kanalı ve 5 saniyelik kısa bir zaman penceresi üzerinden ölçülen sinyali, sınıflandırıcı olarak CNN kullanarak işlemektedir.	Doğruluk: %99
[38]	2025	109	Özellik çıkarımı için fonksiyonel bağlantısallık yöntemi kullanılmış, sınıflandırıcı olarak ise Çok Akışlı Grafik Evrişimsel Sinir Ağı tercih edilmiştir.	Doğruluk: %98,05
Önerilen model	2025	8, 16, 24, 50, 109	EEG verilerinden kanal optimizasyonu yapılarak tek kanallı topluluk makine öğrenmesi tabanlı bir biyometrik kimlik doğrulama sistemi geliştirilmiş ve 64 kanallı modellere kıyasla çok daha düşük işlem maliyetiyle %100 doğruluk elde edilebileceğini göstermiştir.	Doğruluk 8 denek: %100 Doğruluk 16 denek: %99,5 Doğruluk 24 denek: %99,3 Doğruluk 50 denek: %98,9 Doğruluk 109 denek: %98,3

çalışmada ise doğruluk %98,05 olarak rapor edilmiştir [38]. Önerilen model, farklı denek sayıları (8, 16, 24, 50 ve 109) üzerinde test edilmiş ve tek kanallı topluluk makine öğrenmesi yaklaşımı ile, 64 kanallı modellere kıyasla çok daha düşük işlem maliyetiyle yüksek doğruluk değerlerine ulaşılabileceğini göstermiştir. Özellikle 8 denek üzerinde %100 doğruluk elde edilirken, denek sayısı arttıkça doğruluk oranında küçük düşüşler gözlenmiş; 109 denek için dahi %98,3 doğruluk elde edilerek sistemin ölçeklenebilirliği ve güvenilirliği ortaya konulmuştur.

Genel olarak, benzer veri tabanlarıyla gerçekleştirilen çalışmalarla kıyaslandığında, önerilen sisteminin kullanıcı sayısındaki dezavantajına rağmen bazı belirgin avantajları şunlardır:

- Sadece tek kanal üzerinden toplanan verilerle yüksek doğruluğa ulaşılmaktadır.
- Frekans öznitelikleri ve zaman-frekans öznitelikleri gibi karmaşık dönüşümlere sahip özniteliklere ihtiyaç duyulmadan, düşük hesaplama maliyetiyle öznitelik çıkarılmaktadır.
- Herhangi bir uyarın veya görev tanımı olmadan kayıt yapılarak tanımlama gerçekleştirilmektedir.

Bu çalışmadan elde edilen sonuçlar, dinlenme durumunda farklı bireylerin gözlerinin kapalı olarak kaydedilen EEG verileriyle, tek kanal üzerinden düşük maliyetli bir EEG tabanlı kimlik tanımlama sisteminin tasarlanabileceğini göstermektedir.

5. Sonuçlar (Conclusions)

Fizyolojik biyometriklere dayalı sistemler, yüksek güvenlik seviyelerine sahiptir. Ayrıca, bu sistemlerin taklit edilmesi oldukça zordur. Günümüzde gelişen teknolojiler ile sensörlerin kullanımı ve yapay zekâ tekniklerinin gelişmesi gibi nedenler, bu biyometriklerin gerçek hayatta kullanılabilirliği üzerinde çalışmaların artmasına yol açmıştır. Bu makalede yapılan çalışmada, tek kanal kullanılarak EEG tabanlı bir kimlik tanımlama yöntemini düşük maliyet ve hızlı işlem kapasitesi gibi büyük uygulama sorunlarını giderebilecek bir yaklaşım geliştirmeye odaklanılmaktadır. EEG tabanlı biyometrik kimlik tanımlama sistemlerinin uygulanabilirliğinin artırılabilir olması için hem daha az sayıda kanala hem de daha yüksek doğruluğa ulaşılması gerekmektedir. Önerilen tanımlama sisteminde, tek kanallı topluluk makine öğrenmesi yöntemi ile 8 kişiden oluşan bir kapalı sistem modellemesinde %100 doğruluk seviyesine 35 saniye eğitim ve 8,75 saniye test verisi süresinde ulaşılmıştır. Mevcut veri tabanında, kullanıcılara özgü yüksek benzersizliğe sahip kanallı seçerken Gini önem katsayısından faydalanılmıştır. Kanal sayısının azaltılması, sistem karmaşıklığını da azaltmaktadır. Ayrıca, parietal ve frontal loblara yerleştirilen elektrotların, diğer loblara yerleştirilenlerden daha yüksek Gini önem değerlerine sahip olduğu ve beyin merkezine yakın elektrotların daha yüksek Gini önem değerine sahip olması, bu kanallar kullanılarak tasarlanan sistemlerin daha yüksek doğruluğa sahip olduğu test sonuçlarıyla kanıtlanmıştır. Gözler açık olarak kaydedilen veri setinde, frontal lobunda yer alan Ft7 kanalı, gözler kapalı olarak kaydedilen veri setinde, parietal lobunda yer alan Cp4 kanalı ile yapılan testlerde en yüksek doğruluğa ulaşılmıştır. Gözleri kapalı olarak kaydedilen verilerde daha yüksek doğruluğa ulaşılmasındaki temel neden olarak, ortamda etkilenebilecek materyallerin varlığının EEG gibi hassas sinyallerde bozucu etkiye sebep olması olarak yorumlanabilir. Önerilen sistemin test sonuçlarıyla birlikte duyarlılığının, tüm kanalların kullanıldığı yaklaşımlara göre daha iyi ve kabul edilebilir olduğu, parietal ve frontal lob elektrotlarının etkisinin yüksek olduğu gözlemlenmiştir.

Araştırma sonucu, sınırlı sayıda kullanıcıyla oluşturulan sistemde veri toplamanın tek bir kanala indirgenerek kullanılabilirliği kanıtlanmaktadır. Zaman bölgesinde çıkarılan öznitelikler, sistemin

daha hızlı öznitelik çıkarmasına olanak sağlayarak ayırt edici benzersiz verileri yakalamada tutarlılık sağlamıştır. Günümüzde kullanılan biyometriklere göre, EEG sinyalinin biyometrik olarak kullanılabilir olduğu sonucuna varılması için henüz erken bir aşamadır. Gelecek çalışmalar için, EEG sinyallerinin zamanla genliğinin değişmesi, klinik sorunlu bireylerin katılımı, geniş denek sayısı, kimlik doğrulama modu kullanımı, ticari uygulamaya uygunluğu gibi sorunların çözülmesi için daha fazla çalışma yapılması gerekmektedir.

Kaynaklar (References)

1. Gui Q., Jin Z., Xu W., Ruiz-Blondet M.V., Laszlo S., Multichannel EEG-based biometric using improved RBF neural networks, 2015 IEEE signal processing in medicine and biology symposium, Philadelphia-USA, 1-6, 12 Aralık, 2015.
2. Sohankar J., Sadeghi K., Banerjee A., Gupta S.K., E-bias: A pervasive EEG-based identification and authentication system, the 11th ACM Symposium on QoS and Security for Wireless and Mobile Networks, Cancun-Meksika, 165-172, Kasım, 2015.
3. Poulos M., Rangoussi M., Evangelou A., Person identification based on parametric processing on the EEG, the 6th Int. Conf. on Electronics, Circuits and Systems, USA, 283-286, Aralık, 1999.
4. Paranjape R.B., Mahovsky J., Benedicenti L., The electroencephalogram as a biometric, the Canadian Conf. On Electrical and Computer Engineering, Kanada, 1363-1366, Mayıs, 2001.
5. Palaniappan R., Mandic D.P., Biometrics from brain electrical activity: a machine learning approach, IEEE Trans. Pattern Anal. Mach. Intell., 29 (4), 738-742, 2007.
6. Palaniappan R., Raveendran P., Individual identification technique using visual evoked potential signals, Electron. Lett., 38 (25), 1634-1635, 2002.
7. Ficici C., Telatar Z., Kocak O., Eroglu O., Identification of TLE Focus from EEG Signals by Using Deep Learning Approach, Diagnostics, 13 (13), 2261, 2023.
8. Masahiro H., Hirokazu M., Noriyuki M., Fumitaka U., Hirokazu T., Measurement of Brain Activity on Force Adjustment Skill Acquisition by using EEG, International Conference on Knowledge Based and Intelligent Information and Engineering Systems, Marsilya-Fransa, 6-8 Ekim, 2017.
9. Singh B., Mishra S., Tiwary U. S., EEG based biometric identification with reduced number of channels, International Conference on Advanced Communication Technology, Ağustos, 2015.
10. Wijayanto I., Hadiyoso S., Sekarningrum F. A., Biometric Identification Based on EEG Signal with Photo Stimuli using Hjorth Descriptor, 8th International Conference on Information and Communication Technology, Yogyakarta-Endonezya, 24-26 Haziran, 2020.
11. Kim D., Kim K., Resting State EEG-Based Biometric System Using Concatenation of Quadrantal Functional Networks, IEEE Access, 7, 65745-65756, 2019.
12. Prabhakar S., Pankanti S., Jain A.K., Biometric recognition: security and privacy concerns, IEEE Secur. Priv., 99 (2), 33-42, 2003.
13. Constant L., Sabourdin N., The EEG signal: a window on the cortical brain activity, Pediatric Anesthesia, 22 (6), 539-552, 2012.
14. Herff C., Heger D., de Pestors A., Brain-to-text: decoding spoken phrases from phone representations in the brain, Front. Neurosci., 9 (217), 2015.
15. Poulos M., Rangoussi M., Alexandris N., Neural network based person identification using EEG features, IEEE Int. Conf. Acoustics, Speech, Signal Processing, Phoenix-USA, 1117-1120, 15-19 Mart, 1999.
16. Fraschini M., Hillebrand A., Demuru M., An EEG-Based Biometric System Using Eigenvector Centrality in Resting State Brain Networks, IEEE Signal Process. Lett., 22 (6), 666-670, 2015.
17. Dan Z., Xifeng Z., Qiangang G., An identification system based on portable EEG acquisition equipment, Third Int. Conf. on Intelligent System Design and Engineering Applications (ISDEA), Hong Kong-Çin, 281-284, 16-18 Ocak, 2013.
18. Su F., Xia L., Cai A., EEG-based personal identification: from proof of concept to a practical system, 20th Int. Conf. on Pattern Recognition (ICPR), İstanbul-Türkiye, 3728-3731, 23-26 Ağustos, 2010.
19. Suppiah R., Vinod A. P., Biometric identification using single channel EEG during relaxed resting state, IET Biometrics, 7 (4), 342-348, 2018.

20. Ramaswamy P., Jenish G., Kenneth R., Andrews S., PIN Generation Using Single Channel EEG Biometric, *Advances in Computing and Communications ACC 2011 Communications in Computer and Information Science*, Berlin-Almanya, 378–385, 2011.
21. Isao N., Sadanao B., Chisei M., EEG Based Biometric Authentication Using New Spectral Features, 2009 IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS2009), Kanazawa-Japonya, 651-654, 7-9 Ocak, 2009.
22. Sarineh K., Ali F., Saeid R., Electroencephalogram Based Biometrics: A Fractional Fourier Transform Approach, 2nd International Conference on Biometric Engineering and Applications ICBEA '18, Amsterdam-Hollanda, 1-5, 16-18 Mayıs, 2018.
23. Chin T.Z., Saidatul A., Ibrahim Z., Exploring EEG based authentication for imaginary and nonimaginary tasks using power spectral density method, The 1st International Conference on Mechanical Electronic and Biosystem Engineering (MEBSE 2018), Bogor-Endonezya, 15–16 Aralık, 2018.
24. Valsaraj A., Madala I., Garg N., Patil M., Baths V., Motor imagery based multimodal biometric user authentication system using EEG, In 2020 International Conference on Cyberworlds (CW), Caen-Fransa, 272-279, 29 Ekim, 2020.
25. Sun Y., Lo F.P.W., Lo B., EEG-based user identification system using 1D-convolutional long short term memory neural networks, *Expert Systems with Applications*, 125, 259-267, 2019.
26. Balcı F., Oralhan Z., LSTM ile EEG Tabanlı Kimliklendirme Sistemi Tasarımı, *Avrupa Bilim ve Teknoloji Dergisi*, Ejosat Special Issue 2020 (HORA), 135-141, 2020.
27. Hendrawan M.A., Saputra P.Y., Rahmad C., Identification of optimum segment in single channel EEG biometric system, *Indonesian Journal of Electrical Engineering and Computer Science*, 23 (3), 1847-1854, 2021.
28. Fan Y., Shi X., Li Q., CNN-based personal identification system using resting state electroencephalography, *Computational Intel. and Neuro.*, 2021 (1), 2021.
29. Schalk G., McFarland D.J., Hinterberger T., Birbaumer N., Wolpaw J.R., BCI2000: A general-purpose brain-computer interface (BCI) system, *IEEE Transactions on Biomedical Engineering*, 51 (6), 1034-1043, 2004.
30. Phinyomark A., EMG Feature evaluation for improving myoelectric pattern recognition robustness, *Expert Syst. Appl.*, 40 (12), 832–840, 2013.
31. Rocca D.L., Campisi P., Vegso B., Human brain distinctiveness based on EEG spectral coherence connectivity, *IEEE Trans. Biomed. Eng.*, 61 (9), 2406–2412, 2014.
32. Hema C.R., Paulraj M.P., Kaur H., Brain signatures: a modality for biometric authentication, *Int. Conf. on Electronic Design*, Penang-Malezya, 1–3 Aralık, 2008.
33. Safont G., Salazar A., Soriano A., Combination of multiple detectors for EEG based biometric identification and authentication, *IEEE Int. Carnahan Conf. on Security Technology (ICCST)*, Boston-USA, 230–236, 15-18 Ekim, 2012.
34. Yang S., Deravi F., Novel HHT-based features for biometric identification using EEG signals, 22nd Int. Conf. on Pattern Recognition, Stockholm-İsveç, 24-28 Ağustos, 2014.
35. Balcı F., DM-EEGID: EEG-Based Biometric Authentication System Using Hybrid Attention-Based LSTM and MLP Algorithm, *Trait. Du Signal*, 40, 1–14, 2024.
36. Lai C.Q., Ibrahim H., Abdullah M.Z., Suandi S.A., EEG-Based Biometric Close-Set Identification Using CNN-ECOC-SVM, *Artificial Intelligence in Data and Big Data Processing*, Springer Nature: Berlin, Almanya, 723–732, 2022.
37. Alsumari W., Hussain M., Alshehri L., Aboalsamh H.A., EEG-Based Person Identification and Authentication Using Deep Convolutional Neural Network, *Axioms*, 12, 74, 2023.
38. Tian W., Li M., Hu D., Multi-band Functional Connectivity Features Fusion Using Multi-stream GCN for EEG Biometric Identification, 2025 International Conference on Autonomous Unmanned Systems (ICAUS 2025), Xi'an-Çin, 23–25 Eylül, 2025.