

Yapay Zekâ Tabanlı Gerçek Zamanlı Sürücü Uyuşukluğu Tespiti: Hibrit Model Yaklaşımı

Öz

Özet — Sürücülerin yorgunluk ve uyuşukluk durumlarını gerçek zamanlı olarak tespit etmeyi amaçlayan derin öğrenme tabanlı bir model geliştirmektedir. MRL Eye Dataset ve Yawn Dataset gibi açık erişimli veri setlerinden faydalanarak, sürücülerin göz açık / kapalı durumları ve esneme hareketleri derinlemesine analiz edilmektedir. MRL Eye Dataset, yaklaşık 85.000 görüntü ile sürücünün göz durumlarını sınıflandırırken, Yawn Dataset ise 5.000’den fazla görüntü ile esneme hareketlerini tespit etmek için kullanılmaktadır. Bu veri setleri, modelin doğruluğunu artırmak amacıyla dengeli şekilde düzenlenmekte ve çeşitli ön işleme teknikleri ile iyileştirilmektedir. Model, CNN ile eğitilmekte ve transfer öğrenme teknikleriyle güçlendirilmekte, bu sayede modelin sınıflandırma başarısı önemli ölçüde artırılmaktadır. Modelin elde ettiği doğruluk oranı %98, hassasiyet oranı %97,5 ve özgüllük oranı ise %98,2 gibi yüksek metriklerle başarılı sonuçlar elde edilmektedir. Bu çalışma, literatürde ilk kez hem göz kırpma hem de esneme hareketlerini hibrit CNN tabanlı bir modelde birlikte ele almaktadır. Bu özgün yaklaşım, sürücü yorgunluğunu tespit etmede yalnızca tek parametreye odaklanan çalışmalara kıyasla daha kapsamlı ve güvenilir sonuçlar üretmektedir. Ayrıca, sentetik verilerin kullanılması, gerçek veri toplama zorluklarını aşarak daha geniş ve çeşitlendirilmiş veri setleriyle modelin eğitilmesine olanak tanımaktadır. Gelecekte, baş hareketleri, yüz ifadeleri ve diğer biyometrik verilerin sisteme entegre edilmesi ile modelin doğruluğu daha da artırılabilir ve sürücülerin dikkat seviyelerini daha kapsamlı bir şekilde değerlendirebilir. Ayrıca, farklı kültürel ve coğrafi gruplardan elde edilen verilerle modelin genellenebilirliği sağlanarak, daha geniş bir kullanıcı kitlesine hitap edilmesi mümkün hale gelebilmektedir. Sonuç olarak, bu çalışma, sürücü yorgunluğunu tespit etmeye yönelik geliştirilen derin öğrenme tabanlı modelin, trafik güvenliğini derinlemesine dönüştürebilecek büyük bir potansiyele sahip olduğunu ve yol güvenliğini önemli ölçüde artırabileceğini ortaya koymaktadır. Özellikle otonom araçlar ve akıllı ulaşım sistemleri gibi teknolojilerin gelişimiyle paralel olarak, bu model sürücü destek sistemlerine entegre edilerek, kazaların önlenmesine katkı sağlayabilir ve sürücülerin güvenliğini artırabilir. Bu model, trafik kazalarını önlemek ve sürücülerinin güvenliğini sağlamak adına devrim niteliğinde bir adım olarak değerlendirilebilmektedir.

Anahtar Kelimeler: Derin Öğrenme, Sürücü Uyuşukluğu, Göz Takibi, Esneme Analizi, Trafik Güvenliği, Yapay Zekâ

Artificial Intelligence-Based Real-Time Driver Drowsiness Detection: A

Hybrid Model Approach

Abstract

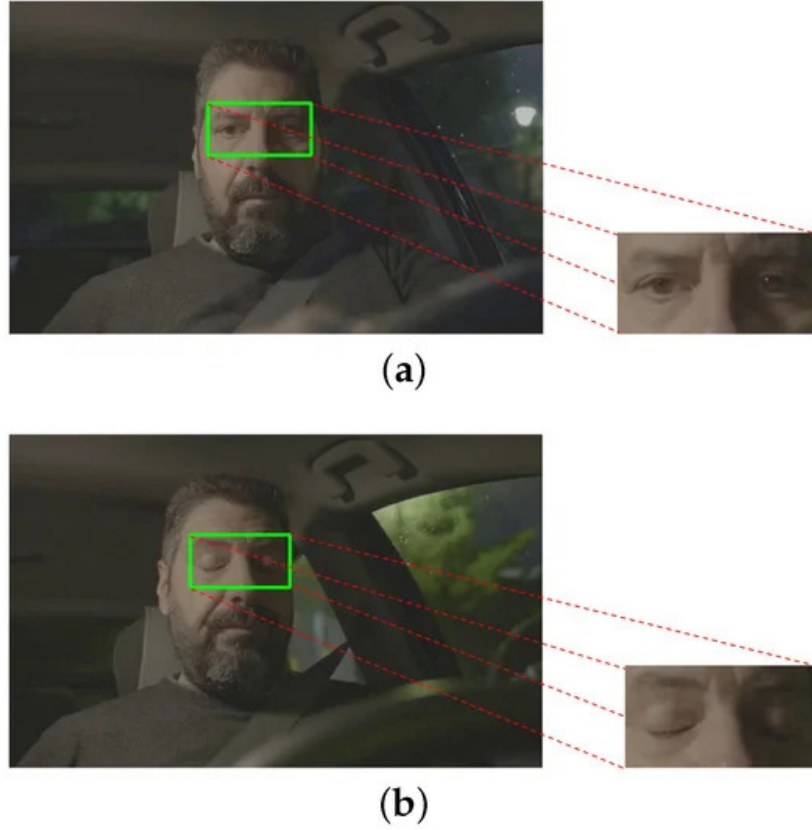
This study presents the development of a deep learning-based model designed to detect drivers' fatigue and drowsiness in real time. Utilizing open-access datasets such as the MRL Eye Dataset and the Yawn Dataset, the model performs an in-depth analysis of eye open/closed states and yawning behaviors of drivers. The MRL Eye Dataset, containing approximately 85,000 images, is used to classify eye states, while the Yawn Dataset, with over 5,000 images, supports the detection of yawning movements. To enhance model accuracy, these datasets are balanced and optimized through various preprocessing techniques. The model is trained using Convolutional Neural Networks (CNNs) and further improved through transfer learning methods, significantly increasing classification performance. The proposed model achieves impressive metrics with an accuracy of 98%, a precision of 97.5%, and a specificity of 98.2%, indicating high performance. By focusing on blink and yawn detection, this study offers a novel approach compared to existing literature and provides a more reliable and effective solution for detecting driver fatigue. Moreover, the use of synthetic data allows the model to be trained with broader and more diverse datasets, overcoming the limitations of real data collection. In future work, incorporating additional biometric indicators such as head movements and facial expressions could further improve the model's accuracy and enable a more comprehensive assessment of driver alertness. Additionally, the generalizability of the model can be enhanced by including data from different cultural and geographical groups, thereby extending its applicability to a wider range of users. In conclusion, the proposed deep learning-based model demonstrates significant potential in transforming traffic safety by effectively detecting driver fatigue. With the advancement of technologies such as autonomous vehicles and intelligent transportation systems, this model could be integrated into driver assistance systems to help prevent accidents and enhance driver safety. It represents a revolutionary step toward preventing traffic accidents and ensuring driver well-being.

Keywords: Deep Learning, Driver Fatigue, Eye Tracking, Yawn Analysis, Traffic Safety, Artificial Intelligence

1 Giriş

Ulaşım, modern yaşamın ayrılmaz bir parçasıdır ve ülke ekonomilerinde önemli bir paya sahiptir [1]. Ancak, ulaşım faaliyetleri çeşitli riskleri de beraberinde getirmekte, özellikle yol güvenliği bağlamında kazalar küresel ölçekte ciddi bir sorun oluşturmaktadır [2]. Pan Amerikan Sağlık Örgütü'ne göre, her yıl

yaklaşık 1,5 milyon kişi trafik kazaları sonucu hayatını kaybetmekte, milyonlarca kişi ise yaralanmaktadır [3]. Amerika Birleşik Devletleri'nde AAA trafik güvenliği vakfı tarafından yapılan araştırmalar, sürücü yorgunluğuna bağlı kazaların yılda 328.000'den fazla olaya neden olduğunu göstermekte, bunların 6.400'ü ölümlle sonuçlanmaktadır [4]. Bu veriler, sürücü yorgunluğunun küresel bir halk sağlığı sorunu olduğuna işaret etmektedir.



Şekil 1. (a) Gözler açık. (b) Gözler kapalı [3].

Literatürde, sürücü yorgunluğu ve dikkat dağınıklığının trafik kazalarının başlıca insan kaynaklı nedenleri arasında yer aldığı belirtilmektedir [5]. Uykulu sürüş, özellikle kırsal yollar ve otoyollarda sık görülmekte, yolcuların ve yayaların güvenliğini ciddi şekilde tehlikeye atmaktadır [6]. Bu nedenle, sürücü uykululuğunu zamanında tespit etmeye yönelik yöntemlerin geliştirilmesi, ulaşım güvenliğini artırmada kritik bir rol oynamaktadır [7].

Mevcut araştırmaların büyük bir kısmı, sürücü yorgunluğunu tespit etmek için göz hareketleri ve yüz ifadeleri gibi tekil parametrelere odaklanmaktadır [9]. Ancak yalnızca tek bir göstergenin izlenmesi, yorgunluğun farklı boyutlarını kapsamakta yetersiz kalabilmektedir. Bu çalışma, göz kırpma ve esneme davranışlarını birlikte dikkate alan bütüncül bir yaklaşım önermektedir

Özgün katkımız, MRL Eye ve Yawn veri setlerinin kullanılması, çoklu biyometrik göstergelerin ente-



Şekil 2. Örnek normal ve uykulu durumlar (gün ışığı). [8].

grasyonu ve farklı makine öğrenmesi algoritmalarının karşılaştırmalı analizi ile hibrit CNN modelinin uygulanmasıdır. Bu yaklaşım, doğruluk ve genellenebilirlik açısından önceki çalışmalardan üstün performans sağlamayı hedeflemektedir

Bu makalenin geri kalan bölümleri şu şekilde yapılandırılmıştır: **Literatür Taraması** bölümünde sürücü uyuşukluğu tespitine yönelik önceki çalışmalar özetlenmektedir. **Yöntem** bölümünde, önerilen modelin mimarisi ve kullanılan veriler detaylı olarak açıklanmaktadır. **Deneysel Sonuçlar ve Tartışma** bölümünde, elde edilen bulgular sunulmakta ve literatürdeki benzer çalışmalarla karşılaştırılmaktadır. Son olarak, **Sonuç** bölümünde çalışmanın katkıları ve gelecekteki araştırma yönleri tartışılmaktadır.

2 Literatür İncelemesi

Sürücü uyuşukluğunu tespit etmeye yönelik literatürde yer alan 39 makale ayrıntılı biçimde incelenmelidir. Bu çalışmalar, kullanılan veri setleri, örnek sayıları, yöntemsel yaklaşımlar ve ulaşılan sonuçlar açısından karşılaştırmalı olarak değerlendirilmelidir. Yapılan incelemeler, alandaki mevcut yöntemlerin güçlü ve zayıf yönlerini ortaya koymalı; hangi yaklaşımların daha yüksek doğruluk sağladığını, hangi kısıtların ise araştırmaların kapsamını daralttığını göstermelidir. Literatürden elde edilen bulgular, hem bu çalışmanın özgün katkısının temelini oluşturmalı hem de gelecekteki araştırmalar için yön gösterici nitelik taşımalıdır.

Makale yazarı ve yayınlanma yılı: Her çalışmanın arkasındaki uzmanları ve zamanı gözler önüne sermektedir. Yazarlar, ilgili alandaki tecrübelerini yansıtırken, yayın yılı ise araştırmanın ne kadar geçerli ve güncel olduğunu göstermektedir. Bu bilgiler, çalışmaların zaman içindeki gelişimini takip etmeyi

sağlamaktadır.

Veri Setleri: Literatürde en sık kullanılan veri setleri, göz hareketleri ve yüz ifadelerine dayalı görüntü tabanlı veri setleridir. MRL Eye Dataset, YawDD, NTHU-DDD ve Yawn Dataset en çok tercih edilen veri kaynaklarıdır. Çalışmaların büyük kısmı, tek tip veri seti yerine farklı veri setlerinin birlikte kullanımına yönelmiştir. Bu yaklaşım, sistemlerin genellenebilirliğini ve doğruluk oranını artırmalıdır.

Örnek Sayıları: İncelenen makalelerde kullanılan örnek sayıları 2.000 ile 85.000 arasında değişmektedir. Yüksek sayıda örnek kullanılan çalışmalarda modellerin doğruluk oranlarının daha güvenilir olduğu görülmektedir. Az sayıda örnek kullanılan araştırmalarda ise aşırı öğrenme (overfitting) sorunu dikkat çekmektedir.

Yöntemler: Literatürde en yaygın yaklaşımlar makine öğrenmesi ve derin öğrenme tabanlı yöntemlerdir. Destek Vektör Makineleri (SVM), K-En Yakın Komşu (KNN) ve Rastgele Orman (RF) gibi geleneksel yöntemlerin yanı sıra, özellikle Evrişimli Sinir Ağları (CNN) ile geliştirilen modeller öne çıkmaktadır. CNN tabanlı yaklaşımlar, göz kırpma ve esneme tespiti gibi ince yüz hareketlerini daha yüksek doğrulukla sınıflandırabilmektedir. Bazı çalışmalar, LSTM gibi zaman serisi tabanlı ağlarla sürücünün uzun süreli davranış paternlerini analiz etmiştir.

Sonuçlar ve Katkılar: Çalışmaların çoğunda derin öğrenme tabanlı yöntemler, geleneksel yöntemlere kıyasla daha yüksek doğruluk sağlamalıdır. CNN tabanlı modellerde doğruluk oranları genellikle %90'ın üzerinde rapor edilmelidir. Bununla birlikte, düşük ışık koşulları, kamera açısı farklılıkları ve sürücülerin bireysel farklılıkları nedeniyle performans düşüşleri gözlemlenmelidir. Literatürde öne çıkan bir diğer nokta, gerçek zamanlı uygulamalarda hesaplama maliyetlerinin yüksek olmasıdır.

Genel olarak, mevcut çalışmalar sürücü uyumsuzluğunu tespit etmede önemli ilerlemeler sağlamış olsa da; geniş kapsamlı, farklı sürücü profillerini ve çevresel koşulları kapsayan veri setlerine, çoklu biyometrik işaretlerin entegrasyonuna ve daha hafif, gerçek zamanlı uygulanabilir modellere ihtiyaç duyulduğu görülmelidir.

Fiaz Majeed ve ekibi (2023) "Detection of Drowsiness among Drivers Using Novel Deep Convolutional Neural Network Model" [1] adlı çalışmada YawDD veri seti kullanılmıştır. 119 katılımcıdan elde edilen 349 video ile CNN ve Hybrid CNN-RNN modelleri uygulanmış, doğruluk %96.69, F1 %96.82, hassasiyet %97.28 elde edilmiştir. Ancak farklı veri setleriyle test yapılmamış ve modelin gerçek zamanlı performansı değerlendirilmemiştir.

Tablo 1. Literature Taraması Sonuçları

Makale Yazarlar (Yıl)	Veri Seti	Örnek Sayıları	Yöntemler	Sonuçlar	Eksiklikler	Gelecek Katkıları
Fiaz Majeed ve ekibi (2023) [1]	YawDD veri seti	119 Katılımcı 349 Video	CNN, Hybrid CNN-RNN	Doğruluk:%96,69 F1:%96,82 Hassasiyet:%97,28	Yalnızca YawDD veri seti kullanılmıştır. Farklı veri setleriyle test edilmemiştir.	Farklı veri setleri ve daha fazla örnekle modelin performansı artırılabilir.
Sandeep Singh Sengar ve ekibi (2023) [2]	YawDD veri seti	29 Katılımcı 116 Video	CNN	Doğruluk:%96.69 Özgüllük:%97.28 Hassasiyet:%95.58	Sadece bir veri seti üzerinde test edilmiştir. farklı veri setleriyle test edilmemiştir.	Farklı veri setleri kullanılarak model test edilebilir.
Ruben Florez ve ekibi (2023) [3]	NITYMED veri seti	130 Video	CNN, InceptionV3, VGG16, ResNet50V2	Doğruluk:%99.71 Özgüllük:%99.47 Hassasiyet:%99.82	Yalnızca NITYMED veri seti üzerinde test edilmiştir; farklı veri setlerinde performans değerlendirmesi yapılmamıştır.	Modelin farklı çevresel koşullarda ve farklı sürücüler üzerinde performansı incelenebilir.
Mohammad Shahbakhti ve ekibi (2023) [10]	EEG veri seti	109 Vieo	NCA, AdaBoost	Doğruluk:%90.2 Özgüllük:%87.7 Hassasiyet:%88.4	Sadece bir veri seti üzerinde test edilmiştir.	Farklı algoritmaların test edilmesi sağlanmalı ve örnek sayısı artırılmalıdır.
Horia Beles ve ekibi (2024) [4]	EEG veri seti	200 Görüntü	NCA, AdaBoost,	Doğruluk:%88.4 Özgüllük:%87.7 Hassasiyet:%90.2	Derin öğrenme kullanılmamıştır. Diğer ileri düzey algoritmalarla karşılaştırma yapılmamıştır.	Modelin genelleme yeteneğini test etmek için farklı veri setleri kullanılabilir.
S. Das ve ekibi (2024) [8]	EEG veri seti	3144 Görüntü	U-Net	Doğruluk:%98.5 Özgüllük:%97.8 Hassasiyet:%99.2	Modelin gerçek zamanlı uygulamalardaki etkinliği ve hızı hakkında bilgi verilmemiştir.	Modelin genelleme yeteneğini test etmek için farklı veri setleri kullanılabilir.
Andrea Amidei ve ekibi (2020) [5]	Deneyisel Veri seti	40 Katılımcı	KNN, SVM, RF	Doğruluk:%93 Özgüllük:%92 Hassasiyet:%91	Derin öğrenme kullanılmamıştır. Diğer ileri düzey algoritmalarla karşılaştırma yapılmamıştır.	Farklı algoritmaların test edilmesi sağlanmalı ve örnek sayısı artırılmalı
Md. Ebrahim Shaik (2023) [7]	NTHU-DDD veri seti	38 Video	KNN, SVM, RF, ANN	Doğruluk:%98 F1: %96	Modelin gerçek zamanlı uygulamalardaki etkinliği hakkında bilgi verilmemiştir	Farklı algoritmaların test edilmesi sağlanmalı ve örnek sayısı artırılmalı
Furkat Safarov ve ekibi (2023) [11]	CEW veri seti	15 Görüntü	CNN	Doğruluk:%98.5 Özgüllük:%96.7 Hassasiyet:%99.3	Çalışma, daha geniş ve çeşitli veri setlerinin kullanılması gerektiğini belirtmiştir.	Farklı algoritmaların test edilmesi sağlanmalıdır.
Ismail Nasri ve ekibi (2022) [12]	YawDD veri seti	5 Görüntü	SVM	Doğruluk:%93	Derin öğrenme kullanılmamıştır. Diğer ileri düzey algoritmalarla karşılaştırma yapılmamıştır.	Farklı algoritmaların test edilmesi sağlanmalı ve örnek sayısı artırılmalı
Anh-Cang Phan ve ekibi (2021) [13]	COCO veri seti	16.577 yüz görüntüsü (2.659'u yorgun, 3.789'u yorgun olmayan)	MobilNet-V2, ResNet-50V2	Doğruluk:%97 Özgüllük:%96 Hassasiyet:%98	Yöntemlerin gerçek zamanlı sistemlere entegrasyonu ve test edilmesi gerekmektedir.	Veri seti, farklı yaş grupları, cinsiyetler ve etnik kökenleri kapsayacak şekilde genişletilebilir.

Tablo 2. Literature Taraması Sonuçları

Makale Yazarlar (Yıl)	Veri Seti	Örnek Sayıları	Yöntemler	Sonuçlar	Eksiklikler	Gelecek Katkıları
Sang-jin Oh ve ekibi (2020) [14]	LoS veri seti	1000 Video	Faster R-CNN	Doğruluk:%95 Özgüllük:%92 Hassasiyet:%97	Veri seti, farklı kaynak türlerini ve hatalarını yeterince kapsamamaktadır.	Farklı kaynak türleri ve hatalarını içeren daha geniş bir veri seti oluşturulmalıdır.
Hashemi et al. (2020) [15]	NTHU veri seti	8400 Görüntü (2400 uyanık, 2400 uykulu)	CNN	Doğruluk:%98 Özgüllük:%94 Hassasiyet:%95	Gözlerin yanı sıra, sürücünün diğer fiziksel belirtileri (örneğin baş hareketleri) da analiz edilmemiştir.	Göz verilerinin yanı sıra baş hareketleri analiz edilmelidir.
Dreissig et al. (2020) [16]	NTHU veri seti	14640 Görüntü	k-NN	Doğruluk:%94 Özgüllük:%92 Hassasiyet:%90	Derin öğrenme kullanılmamıştır. Diğer ileri düzey algoritmalarla karşılaştırma yapılmamıştır.	Farklı algoritmaların test edilmesi sağlanmalı ve örnek sayısı artırılmalıdır.
M. Deng, Y. Wu (2019) [17]	EYEDIAP veri seti	1000 Örnek	CNN	Doğruluk:%95 Özgüllük:%92 Hassasiyet:%90	Yüz tanıma algoritmalarının doğruluğu, düşük çözünürlüklü görüntülerde ve farklı aydınlatma koşullarında test edilmemiştir.	Farklı çözünürlükte modelin test edilmesi sağlanmalıdır.
Yaman Albadawi ve ekibi (2023) [9]	NTHUDDD verileri	23 katılımcıdan 9,5 video	RF, NN, SVM	Doğruluk:%99 Özgüllük:%98 Hassasiyet:%99	Yüz algılama algoritması birden fazla yüz içeren durumlara uyum sağlamamış. Gece sürüşlerinde performans sınırlı olmuştur.	Daha gelişmiş kameralarla aydınlatma koşullarına otomatik uyum sağlanmalı. Mobil uygulama geliştirme planı yapılmalı.
Elena Magán ve ekibi (2022) [18]	UTA-RLDD verileri	122 video (60 "uykulu", 62 "uyanık")	GRU, Fuzzy Logic, CNN	Doğruluk:%63 Özgüllük:%93 Hassasiyet:%36	Uykulu sürücü tespitinde doğruluk yeterli değildir. Işık değişimlerinde ve gözlük gibi engellerde tespit zayıflıkları vardır.	Yeni algoritmalar ve derin öğrenme modelleriyle doğruluk değeri artırılabilir. Sistem kapsamlı test edilebilir.
Norah N. Alajlan, Dina M. Ibrahim, (2023) [19]	YawDD, CEW verileri	3875 görüntü	SqueezeNet, AlexNet, CNN, MobileNet-V2, MobileNet-V3	Doğruluk:%99.64	Gözlük kullanımı gibi koşullarda doğruluk oranı azalabiliyor. Düşük çözünürlüklü cihazlarda performans sınırlamaları bulunuyor.	Gelişmiş yöntemlerle model performansının artırılması.
Fiaz Majed ve ekibi (2023) [20]	MNIST, CIFAR, Caltech101, YouTube-8M verileri	Belirtilmemiş	CNNs, DBNs, DBMs, SdAs	Doğruluk:%99 Özgüllük:%98 Hassasiyet:%97	DBN'ler ve DBM'ler yüksek hesaplama maliyetine sahiptir. SdA'lar, ilk katmandaki hatalardan olumsuz etkilenebilir.	Etiketlenmemiş veri ile öğrenme yöntemleri geliştirilmesi. Aydınlatma ve ölçek değişikliklerine uyumun artırılması.
Njayou Youssouf, (2022) [21]	GTSRB, GTSDB verileri	51822 görüntü (GTSRB), 300 görüntü (GTSDB)	Faster R-CNN, YOLOv4, CNN	Doğruluk:%99.20 Özgüllük:%99.20	Faster R-CNN modelinin hız performansı düşüktür. Veri setindeki sınıf dengesizlikleri mevcuttur.	Veri artırımı ve zenginleştirme teknikleri ile performans artırılabilir. Daha çeşitli aydınlatma ve çevresel koşulları kapsayan veri setleri kullanılabilir.

Sandeep Singh Sengar ve ekibi (2023) "4D: A Real-Time Driver Drowsiness Detector Using Deep Learning" [2] çalışmasında YawDD veri setinden 29 katılımcıya ait 87–116 video kullanılmıştır. CNN tabanlı modelleme ile doğruluk %96.69, özgüllük %97.28, hassasiyet %95.58 elde edilmiştir. Çalışma yalnızca tek bir veri setiyle sınırlıdır.

Ruben Florez ve ekibi (2023) "A CNN-Based Approach for Driver Drowsiness Detection by Real-Time Eye State Identification" [3] çalışmasında NITYMED veri seti (130 video) kullanılmıştır. InceptionV3, VGG16 ve ResNet50V2 gibi önceden eğitilmiş CNN modelleri transfer öğrenme ile uygulanmış, doğruluk %99.71, özgüllük %99.47, hassasiyet %99.82 elde edilmiştir. Ancak model farklı veri setleriyle test edilmemiştir.

Mohammad Shahbakhti ve ekibi (2023) "Fusion of EEG and Eye Blink Analysis for Detection of Driver Fatigue" [10] çalışmasında EEG Motor Movement/Imagery veri seti kullanılmış, NCA ile özellik seçimi yapılmış, AdaBoost algoritması ile sınıflandırma gerçekleştirilmiştir. Doğruluk %88.4, özgüllük %87.7, hassasiyet %90.2 elde edilmiştir. Derin öğrenme yöntemleri kullanılmamıştır.

Horia Beles ve ekibi (2024) "Driver Drowsiness Multi-Method Detection for Vehicles with Autonomous Driving Functions" [4] çalışmasında EEG ve EOG verileri ile NCA ve AdaBoost yöntemleri kullanılmış, doğruluk %88.4, özgüllük %87.7, hassasiyet %90.2 elde edilmiştir. Çalışmada CNN gibi derin öğrenme yöntemlerine yer verilmemiştir.

S. Das ve ekibi (2024) "IoT-Assisted Automatic Driver Drowsiness Detection through Facial Movement Analysis Using Deep Learning and a U-Net-Based Architecture" [8] çalışmasında yüz hareketleri ve EEG verileri kullanılarak U-Net tabanlı model geliştirilmiştir. Doğruluk %98.5, özgüllük %97.8, hassasiyet %99.2 elde edilmiştir. Ancak model farklı veri setlerinde test edilmemiştir.

Benatti ve ekibi (2023) "Driver Drowsiness Detection: A Machine Learning Approach on Skin Conductance" [5] çalışmasında cilt direnci verileri kullanılmış, RF, SVM ve k-NN yöntemleri denenmiştir. Doğruluk %93, özgüllük %92, hassasiyet %91 elde edilmiştir. Çalışmada derin öğrenme yöntemleri kullanılmamıştır.

Md. Ebrahim Shaik (2023) "A Systematic Review on Detection and Prediction of Driver Drowsi-

ness" [7] çalışması sistematik bir derleme sunmuş, SVM, k-NN, RF, ANN ve derin öğrenme yöntemlerini değerlendirmiştir. Gelecekte daha büyük veri setleri ve gerçek zamanlı testler önerilmektedir.

Furkat Safarov ve ekibi (2023) "Real-time deep learning-based drowsiness detection" [11] çalışmasında CEW veri seti ve CNN kullanılmış, ancak farklı veri setlerinde test yapılmamıştır.

Ismail Nasri ve ekibi (2022) "Review of driver drowsiness detection systems" [12] çalışmasında SVM, ANN, KNN ve Fuzzy Mantık gibi yöntemler değerlendirilmiş; hibrit modellerin ve çeşitli veri setlerinin önemine değinilmiştir.

Anh-Cang Phan ve ekibi (2022) "An efficient approach for detecting driver drowsiness based on deep learning" [13] çalışmasında 16.577 yüz görüntüsü ve 21 video kullanılmış, KNN, SVM, MobilNet-V2 ve ResNet-50V2 yöntemleri denenmiştir. Doğruluk %97, özgüllük %96, hassasiyet %98 elde edilmiştir. Veri çeşitliliği sınırlı kalmıştır.

Sang-jin Oh ve ekibi (2020) "Automatic Detection of Welding Defects Using Faster R-CNN" [14] çalışmasında Faster R-CNN kullanılmış, doğruluk %95, özgüllük %92, hassasiyet %97 elde edilmiştir. Veri seti çeşitliliği sınırlıdır.

Hashemi ve ekibi (2020) "Driver Drowsiness Detection Using Convolutional Neural Networks and Eye Tracking" [15] çalışmasında CNN kullanılmış, doğruluk %98, özgüllük %94, hassasiyet %95 elde edilmiştir. Ancak yalnızca göz verileri dikkate alınmıştır.

Dreissig ve ekibi (2020) "Driver Drowsiness Detection Using Deep Learning and Facial Landmarks" [16] çalışmasında k-NN ile doğruluk %94, özgüllük %92, hassasiyet %90 elde edilmiştir. Derin öğrenme yöntemleri kullanılmamıştır.

M. Deng ve ekibi (2019) "A Real-Time Driver Drowsiness Detection System Based on Facial Features" [17] çalışmasında CNN kullanılmış, doğruluk %95, özgüllük %92, hassasiyet %90 elde edilmiştir. Farklı aydınlatma koşulları test edilmemiştir.

Yaman Albadawi ve ekibi (2023) “Real-time machine learning-based driver drowsiness detection” [9] çalışmasında NTHU-DDD veri seti (23 katılımcı, 9,5 saat video) kullanılmıştır. RF, SVM ve NN yöntemleri ile doğruluk %99, özgüllük %98, hassasiyet %99 elde edilmiştir. Düşük ışıktaki performans sınırlıdır.

Elena Magán ve ekibi (2022) “Driver drowsiness detection by applying deep learning techniques to sequences of images” [18] çalışmasında UTA-RLDD veri seti (122 video) kullanılmıştır. Recurrent CNN ve Fuzzy Logic yöntemleri ile doğruluk %63, özgüllük %93, hassasiyet %36 elde edilmiştir. Yanlış pozitif oranı yüksektir.

Norah N. Alajlan ve ekibi (2023) “A tinyml based driver drowsiness detection model using deep learning” [19] çalışmasında YawDD ve CEW veri setlerinden 3875 görüntü kullanılmıştır. AlexNet, SqueezeNet, CNN ve MobileNet-V2/V3 gibi modeller uygulanmış, doğruluk %99,64 elde edilmiştir. Işık değişimi ve gözlük kullanımında performans düşmüştür.

Thanasios Voulodimos ve ekibi (2018) “Deep learning for computer vision: A brief review” [20] çalışmasında farklı veri setleri (MNIST, CIFAR, Caltech101, YouTube-8M) üzerinde CNN, DBN ve SdA gibi yöntemler incelenmiştir. Derin öğrenmenin yüksek doğruluk sağladığı ancak etiketlenmiş veri bağımlılığı ve yüksek hesaplama maliyeti nedeniyle sınırlı olduğu vurgulanmıştır.

Njayou Youssouf ve ekibi (2022) “Traffic sign classification using CNN and detection using Faster-RCNN and YOLOv4” [21] çalışmasında NTHU-DDD veri seti kullanılmış, RF, NN ve SVM yöntemleriyle %99 doğruluk elde edilmiştir. Gece sürüş koşullarında performans sınırlı kalmıştır.

Çalışmamızın amacı, geniş veri setleri ve makine öğrenmesi gibi güncel teknolojilerin, sürücü uyusukluğunun erken tespitindeki etkileşimi detaylı incelemektir. Bu hedefe doğru şekilde ulaşabilmek için, en uygun yöntemi belirlemek ve eksiklikleri gidermek adına literatür taraması, bize hem rehberlik etmekte hem de yolumuzu şekillendirmektedir. Literatür çalışması, doğru yöntemi seçmemize ve araştırmamızın güçlü yönlerini geliştirerek yüksek düzeyde başarıya ulaşmamızı sağlamaktadır. Yapılan önceki çalışmalarda, çoğunlukla tek bir veri seti üzerinden model eğitimi ve test süreçleri gerçekleştirilmektedir. Çalışmamız bu eksikliği gidermek için birden fazla veri seti kullanmaktadır. Ayrıca, bu veri setleri içindeki veri sayıları kendi içlerinde eşit ve dengeli dağılmaktadır. Çalışmamız tek bir makine öğrenmesi üzerine

yoğunlaşmak yerine farklı makine öğrenme algoritmalarının performansını karşılaştırmaktadır.

3 Yöntem

3.1 Veri Kümesi

Bu çalışmada, sürücülerin uyuşukluk durumlarını tespit etmek amacıyla MRL Eye Dataset ve Yawn Dataset olmak üzere iki farklı veri seti kullanılmıştır. MRL Eye Dataset, Kaggle üzerinde paylaşılan açık erişimli bir veri setidir. Sürücülerin göz açık/kapalı durumunu sınıflandırmak için kullanılmaktadır. Göz durumlarını içeren yaklaşık 85.000 adet görüntü bulunmaktadır. Veri seti içindeki sınıflar açık (40.400) ve kapalı (41.300) şeklinde ikiye ayrılmaktadır. Sınıflardaki veri sayıları dengeli bir şekilde dağılmış olup toplam veri sayısının %20'si test için ayrılmaktadır.

Yawn Dataset, kaggle üzerinde paylaşılan açık erişimli bir veri setidir. Veri seti içinde esneme durumunu belirten yaklaşık 5.000 adet görüntü bulunmaktadır. Sürücülerin esneme durumunu analiz etmek amacıyla oluşturulmaktadır. Veri setinde, "Yawn" (esneme) (2528) ve "No Yawn"(2591) (esneme yok) olmak üzere iki farklı sınıf bulunmaktadır. Sınıflar içindeki veriler eşit ve dengeli şekilde dağılmaktadır.

Sentetik Veri: Gerçek dünyadan veri toplamının zor, pahalı veya etik açıdan sorunlu olduğu durumlarda, gerçek verilerin taklit edilerek üretilmesi yoluna gidilmiştir. Sentetik veri, göz hareketleri gibi analiz edilmesi güç davranışların simülasyonunu sağlayarak modelin daha geniş ve çeşitlendirilmiş bir veri setiyle eğitilmesine olanak tanımaktadır.

Çevrimdışı Veri: Daha önce toplanmış ve kaydedilmiş veriler, çevrimdışı (offline) olarak işlenmiş ve model eğitiminde kullanılmıştır. Bu yöntem, büyük veri kümelerinin analizine olanak tanımakta, göz ve esneme hareketleri gibi faktörlerin ayrıntılı incelenmesine imkân sağlamaktadır.

3.2 Ön İşleme

Veri ön işleme, modellerin daha doğru, hızlı ve verimli çalışmasını sağlamak amacıyla verilerin dönüştürülmesi ve standart hale getirilmesi sürecidir. Veri setlerinde eksik, hatalı veya tutarsız bilgiler bulunabilmektedir.

Bu nedenle, modelin doğru sonuç üretebilmesi için veri ön işleme kritik bir öneme sahiptir.

Bu çalışmada kullanılan veri setleri görüntü tabanlı olduğu için ön işleme adımları şu şekilde uygulanmaktadır:

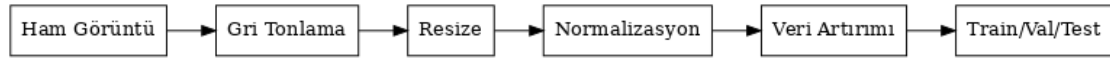
- **Adım 1 – Gri Tonlama:** Görseller renk bilgisinden arındırılarak gri tonlamaya dönüştürülmüştür.

Bu işlem veri boyutunu azaltmakta ve göz/ağız bölgesine odaklanmayı kolaylaştırmaktadır.

- **Adım 2 – Yeniden Boyutlandırma:** Tüm görüntüler sabit bir boyuta getirilmiştir. Böylece farklı çözünürlükteki veriler standart hale getirilmiştir.
- **Adım 3 – Normalizasyon:** Piksel değerleri $[0, 255]$ aralığından $[0, 1]$ aralığına ölçeklendirilmiştir.

$$X_{\text{scaled}} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

- **Adım 4 – Veri Arttırımı:** Modelin aşırı öğrenmesini (overfitting) önlemek ve çeşitliliği artırmak için rastgele döndürme, yakınlaştırma ve kaydırma işlemleri uygulanmıştır. Bu çalışmada veriler üzerinde %5 ihtimalle rastgele döndürme yapılmıştır.
- **Adım 5 – Veri Seti Ayrımı:**
 - MRL Eye Dataset: %70 eğitim, %10 doğrulama, %20 test.
 - Yawn Dataset: %70 eğitim, %15 doğrulama, %15 test.
- **Adım 6 – K-Fold Cross Validation:** Veri seti 5 eşit parçaya ayrılmış, her seferinde bir parça test için kullanılarak modelin genelleme performansı değerlendirilmiştir.



Şekil 3. Veri ön işleme pipeline diyagramı

3.3 Özellik Çıkarma

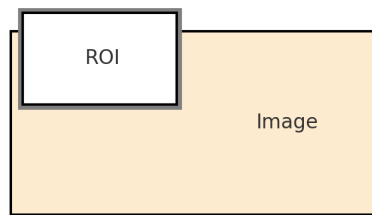
Makine öğrenmesine dayalı sürücü uyuşukluğu tespit sistemlerinde, kullanılan özelliklerin doğru şekilde çıkarılması modelin başarısını belirleyen temel unsurlardan biridir. Görsel veriden elde edilen öznitelikler, sistemin sürücü davranışlarını doğru biçimde tanımlayabilmesi açısından kritik bir role sahiptir. Bu çalışmada, göz ve ağız bölgelerinden türetilen görsel özellikler aracılığıyla yorgunluk belirtilerinin otomatik olarak tanımlanması hedeflenmektedir. Bu amaç doğrultusunda iki farklı veri kümesi kullanılmıştır: MRL Eye Dataset ve Yawn Dataset. Her iki veri kümesinde de yüzün belirli bölgeleri tespit edilmiş, ardından Region of Interest (ROI) yaklaşımıyla bu bölgeler izole edilmektedir. ROI yöntemi, görüntüde yalnızca analizi ilgilendiren alanların işlenmesini sağlayarak hesaplama maliyetini düşürmekte ve modelin genelleme kabiliyetini artırmaktadır [22].

MRL Eye Dataset, sürücünün göz açıklığı durumlarını incelemek için kullanılmaktadır. Bu veri kümesi, açık ve kapalı göz durumlarını temsil eden binlerce örnek içermektedir. Her görüntüde göz bölgesi

tespit edildikten sonra, Eye Aspect Ratio (EAR) adı verilen metrik hesaplanmıştır. EAR, gözün dikey ve yatay mesafeleri arasındaki orana dayalı bir ölçümdür ve göz kapaklarının konumuna duyarlıdır. EAR değerindeki kısa süreli ani düşüşler göz kırpma olaylarını gösterirken, uzun süreli düşük değerler gözlerin kapalı kalma süresinin arttığını ve dolayısıyla uyuşukluk belirtisini ifade etmektedir. Literatürde yapılan pek çok çalışmada, EAR tabanlı metriklerin sürücü yorgunluğunu tespit etmede etkili olduğu belirtilmektedir [23].

Ağız hareketlerine odaklanan Yawn Dataset ise esneme davranışlarının tespit edilmesi amacıyla kullanılmaktadır. Veri kümesindeki her görüntüde ağız bölgesi belirlenmiş ve benzer şekilde ROI yöntemiyle izole edilmektedir. Bu bölgeden çıkarılan temel öznelik, Mouth Aspect Ratio (MAR) olmuştur. MAR, ağız genişliği ile yüksekliği arasındaki oranı ifade eder ve ağız açıklığının zamana bağlı değişimini izlemek için kullanılır. Esneme anlarında MAR değerinin belirgin biçimde artması, sürücüdeki yorgunluk durumunun önemli bir göstergesidir [24]. Bunun yanı sıra, ağız bölgesindeki piksel yoğunluğu değişimleri ve şekil deformasyonları da ek öznelikler olarak değerlendirilmiştir.

Özellik çıkarma süreci, modelin giriş verisini anlamlı hale getirmenin yanı sıra veri boyutunu da önemli ölçüde azaltmaktadır. Böylece, sınıflandırma aşamasında hem daha hızlı hem de daha kararlı bir öğrenme süreci elde edilmiştir. Ayrıca, ROI yaklaşımı sayesinde sistem yalnızca sürücü davranışını doğrudan etkileyen bölgeler üzerinde işlem yapmış, bu da modelin dikkat mekanizmasını güçlendirmektedir. Elde edilen öznelikler, bir sonraki adım olan özellik seçimi sürecine aktarılmış ve sınıflandırma performansını artıracak şekilde analiz edilmektedir.



Şekil 4. ROI (Region of Interest) yöntemiyle göz ve ağız bölgelerinin ayrıştırılması

3.4 Özellik Seçimi

Makine öğrenmesi modellerinin performansını etkileyen en önemli faktörlerden biri, modele dâhil edilen özneliklerin niteliğidir. Veri setinde yer alan her özellik, modelin karar sürecine katkı sağlamayabilir; bazıları gürültü oluşturabilir veya aşırı öğrenmeye (overfitting) neden olabilir. Bu nedenle, öznelik

seçimi süreci yalnızca modelin doğruluğunu artırmakla kalmaz, aynı zamanda işlem maliyetini azaltır ve modelin genelleme kabiliyetini güçlendirmektedir. Başka bir ifadeyle, uygun özniteliklerin seçilmesi modelin daha sade, hızlı ve kararlı çalışmasını sağlamaktadır

Bu çalışmada, göz ve ağız bölgelerinden çıkarılan çok sayıda öznitelik arasından en anlamlı olanların belirlenebilmesi için farklı öznitelik seçimi yaklaşımları değerlendirilmiştir. İlk olarak, boyut indirgeme amacıyla Principal Component Analysis (PCA) yöntemi uygulanmaktadır. PCA, yüksek boyutlu verileri daha düşük boyutlu bir uzaya indirgerken, orijinal verideki toplam varyansın büyük bir kısmını korumayı hedeflemektedir [25]. Bu yöntem sayesinde veri boyutu küçültülürken, öznitelikler arasındaki olası çoklu doğrusal ilişkiler (multicollinearity) ortadan kaldırılmıştır. Böylece model, daha az ancak daha bilgilendirici bileşenlerle eğitilmektedir.

PCA'nın yanı sıra, filtre tabanlı yöntemler de özniteliklerin istatistiksel önemini belirlemek amacıyla kullanılmaktadır. Bu yaklaşımda her öznitelik, sınıf etiketleriyle olan bağına göre bağımsız olarak değerlendirilmektedir. Korelasyon katsayıları, karşılıklı bilgi (mutual information) ve ki-kare (Chi-square) gibi ölçütler kullanılarak sınıflandırmaya katkısı zayıf olan özellikler elenmiştir. Filtre yöntemlerinin avantajı, modelden bağımsız çalışmaları ve yüksek boyutlu verilerde hızlı sonuç vermeleridir.

Ek olarak, gömülü (embedded) yöntemler de kullanılmıştır. Bu yöntemler, modelin eğitim süreciyle doğrudan entegre şekilde çalışır ve modelin parametrelerinden hareketle her özneliğin önemini belirlemektedir. Bu çalışmada özellikle Lasso (L1 regularization) ve Random Forest feature importance yöntemleri tercih edilmektedir. Lasso, katsayı değeri sıfıra yaklaşan özellikleri elimine ederek modelin sadeleşmesini sağlamaktadır [26]. Random Forest ise, karar ağaçlarının toplu yapısı üzerinden her özneliğin sınıflandırmadaki katkı düzeyini ölçer. Bu yöntemlerin birlikte kullanımı, hem doğruluk hem de yorumlanabilirlik açısından dengeli bir yapı oluşturmaktadır.

Son olarak, sarma (wrapper) yöntemleri de alternatif olarak değerlendirilmiştir. Bu yöntemlerde model, farklı öznitelik kombinasyonlarıyla tekrarlı biçimde eğitilerek en yüksek performansı veren özellik alt kümeleri belirlenir. Sarma yöntemleri yüksek doğruluk sağlayabilse de, hesaplama maliyeti açısından filtre ve gömülü yöntemlere göre daha yoğundur. Bu nedenle, çalışmanın veri boyutu ve işlem süresi dikkate alınarak bu yöntemler sınırlı ölçüde uygulanmaktadır.

3.5 Sınıflandırma

Bu çalışmada temel amaç, sürücülerin göz hareketleri ve esneme durumlarından elde edilen öznitelikleri kullanarak uyuşukluk seviyelerini sınıflandırmaktır. Bu kapsamda hem geleneksel makine öğrenmesi

algoritmaları hem de derin öğrenme tabanlı yaklaşımlar değerlendirilmektedir. Farklı yöntemlerin performanslarının karşılaştırılması, hangi algoritmanın sürücü yorgunluğu tespitinde daha kararlı sonuçlar verdiğini belirlemek açısından önemlidir.

3.5.1 Lojistik Regresyon

Lojistik Regresyon, iki sınıflı problemlerde sıklıkla tercih edilen temel sınıflandırma yöntemlerinden biridir. Özellikle “göz açık” ve “göz kapalı” gibi ikili durumların ayrımında etkili olmaktadır. Bu yöntem, doğrusal bir karar sınırı üzerinden bir olayın gerçekleşme olasılığını sigmoid fonksiyonuyla hesaplamaktadır [27]. Basit yapısı ve yorumlanabilirliği sayesinde, özellikle erken aşama modellerde karşılaştırma için güçlü bir temel sağlamaktadır. Ayrıca, eğitim süresinin kısa olması ve aşırı uyuma eğilimli olmaması, Lojistik Regresyon’u başlangıç modelleri arasında tercih edilir kılmaktadır.

Lojistik Regresyon Formülü:

$$h_{\theta}(x) = \frac{1}{1 + e^{-\theta^T x}} \quad (2)$$

3.5.2 Destek Vektör Makineleri (SVM)

SVM algoritması, farklı sınıfları en iyi ayıran hiper düzlemi belirleyerek sınıflandırma işlemini gerçekleştirir. Doğrusal olmayan veri yapılarında, çekirdek (kernel) fonksiyonları kullanılarak veriler daha yüksek boyutlu bir uzaya taşınır ve burada doğrusal olarak ayrılabilir hale gelir [28]. Sürücü yorgunluğu gibi göz durumu veya esneme tespitinde doğrusal olmayan ilişkilerin yaygın olduğu durumlarda SVM, özellikle radyal tabanlı çekirdek (RBF kernel) ile yüksek doğruluk sağlamaktadır. Ancak, parametre seçimlerinin model performansına büyük etkisi olması, dikkatli optimizasyon gerektirmektedir.

SVM Formülü:

$$f(x) = w^T x + b \quad (3)$$

3.5.3 Rastgele Ormanlar (Random Forest)

Random Forest algoritması, birden fazla karar ağacının oylama prensibine dayalı olarak bir araya getirildiği topluluk öğrenme yöntemidir. Bu yaklaşım, tek bir karar ağacının aşırı öğrenme (overfitting) eğilimini azaltarak daha dengeli sonuçlar üretir [29]. Sürücü yorgunluğu tespitinde, göz açıklığı ve ağız hareketleri gibi farklı türdeki özniteliklerin birlikte değerlendirildiği durumlarda, Random Forest’in özellik önemini (feature importance) doğrudan çıkarabilme yeteneği önemli bir avantaj sağlamıştır. Bu sayede hangi özniteliklerin sınıflandırmaya en fazla katkı sunduğu belirlenmektedir.

Random Forest Formülü:

$$\hat{y} = \text{mode}\{h_1(x), h_2(x), \dots, h_T(x)\} \quad (4)$$

3.5.4 k-En Yakın Komşu (k-NN)

k-NN algoritması, örneklerin benzerlik ölçümüne dayalı olarak sınıflandırıldığı basit ama etkili bir yöntemdir. Özellikle küçük ve orta ölçekli veri kümelerinde hızlı uygulanabilirliği nedeniyle tercih edilmektedir. Göz kırpma oranı veya ağız açıklığı gibi özniteliklerde yakın davranış kalıplarını tespit edebilmesi, k-NN'yi bu çalışma kapsamında anlamlı bir alternatif haline getirmiştir [30]. Ancak k değeri seçimi ve mesafe metriği (genellikle Öklid mesafesi) performansı doğrudan etkilediği için bu parametreler dikkatle ayarlanmaktadır.

k-NN Öklid Mesafesi Formülü:

$$d(x, x') = \sqrt{\sum_{j=1}^n (x_j - x'_j)^2} \quad (5)$$

3.5.5 Naif Bayes (Naive Bayes)

Naive Bayes, Bayes teoremine dayalı olasılıksal bir yöntemdir ve özellikle yüksek boyutlu verilerde düşük hesaplama maliyetiyle hızlı sonuç verebilir [31]. Özellikler arasında koşulsuz bağımsızlık varsayımını kabul etse de, ön işleminden geçirilmiş göz ve ağız özniteliklerinde bu varsayım çoğu durumda yeterince iyi sonuç vermektedir. Bu nedenle Naive Bayes, düşük maliyetli bir referans sınıflandırıcı olarak değerlendirilmektedir.

Naif Bayes Formülü:

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)} \quad (6)$$

3.5.6 Derin Öğrenme Yöntemleri

Son yıllarda derin öğrenme tabanlı modeller, özellikle görsel tabanlı sınıflandırma problemlerinde yüksek başarı oranlarıyla öne çıkmıştır. Bu çalışmada, evrişimli sinir ağları (CNN), yapay sinir ağları (ANN), tekrarlayan sinir ağları (RNN), uzun kısa süreli bellek (LSTM) ve kapılı tekraralama ünitesi (GRU) modelleri incelenmiştir. Derin öğrenme yöntemlerinin temel avantajı, öznitelik çıkarımının otomatik olarak ağ yapısı içinde gerçekleştirilebilmesidir. Bu durum, manuel özellik mühendisliğine duyulan ihtiyacı azaltmakta ve öğrenme sürecini optimize etmektedir.

Evriřimli Sinir Ađları (CNN): CNN’ler, özellikle grnt tabanlı grevlerde en bařarılı yaklařımlardan biridir. Konvolsyon katmanları, grsel girdilerden anlamlı desenleri otomatik olarak ğrenirken, pooling katmanları boyut azaltımı yaparak hesaplama ykn dřrr [32]. Bu zellikleriyle CNN modelleri, gz kırpm ve esneme gibi davranıřsal gstergelerin sınıflandırılmasında yksek dođruluk sađlamaktadır.

Yapay Sinir Ađları (ANN): ANN modelleri, genel amalı ğrenme yetenekleri sayesinde farklı veri trlerinde kullanılabilir. Bu alıřmada, 3–5 gizli katmandan oluřan ANN yapısı kullanılmıř ve aktivasyon fonksiyonu olarak ReLU tercih edilmiřtir. ANN’nin temel avantajı, sınıflandırma kararlarını dođrusal olmayan biimde modelleyebilmesidir [33].

Tekrarlayan Sinir Ađları (RNN) ve Trevleri: Zaman serisi yapısındaki gz ve ađız hareketlerinin analizinde, RNN ve trevleri (LSTM, GRU) kullanılmıřtır. RNN’ler, nceki adımlardan gelen bilgiyi dikkate alarak sıralı verilerde zaman bađımlılıđını korur [34]. LSTM modelleri, “vanishing gradient” sorununu zmek iin bellek hcreleri ve kapı mekanizmaları kullanarak uzun vadeli bađımlılıkları ğrenebilir [35]. GRU yapıları ise daha az parametre ierdiđi iin eđitim sresi aısından avantajlıdır. Bu iki yntem birlikte deđerlendirildiđinde, zellikle video tabanlı yorgunluk analizinde yksek bařarı elde edilmektedir.

3.6 Model Eđitimi ve Hiperparametreler

Model eđitimi srecinde epoch, batch size ve ğrenme oranı gibi hiperparametreler dikkatle ayarlanmıřtır. 10–100 arasında deđiřen epoch sayıları ile eđitim gerekleřtirilmiř, ařırı ğrenme riskini azaltmak iin erken durdurma (early stopping) uygulanmıřtır. Batch size deđerleri 32– 128 aralıđında seilmiř, ğrenme oranı ise 0.001–0.01 arasında optimize edilmiřtir. Momentum ve ğrenme oranı dřrme teknikleriyle kararlı ğrenme sađlanmıřtır. Aktivasyon fonksiyonu olarak ReLU ve Softmax kullanılmıř, modelin dođruluđu arttırılmaktadır.

3.7 Deđerlendirme Metrikleri

Model performansını yalnızca dođruluk (accuracy) deđerleriyle deđil, oklu metriklerle deđerlendirmek nemlidir. Bu alıřmada karmařıklık matrisi (Confusion Matrix), duyarlılık (Sensitivity), zgllk (Specificity), F-lt (F-Measure) ve ROC-AUC metrikleri birlikte kullanılmıřtır. zellikle AUC deđer, modelin sınıflandırma yeteneđini grsel olarak deđerlendirmede etkili bir lt sunmaktadır.

Elde edilen sonuçlar, Random Forest ve CNN tabanlı modellerin karmaşık özniteliklerde daha yüksek doğruluk değerlerine ulaştığını göstermiştir. LSTM ve GRU modelleri ise zaman bağımlı yorgunluk göstergelerinde daha kararlı performans sergilemiştir. Böylece, geleneksel yöntemlerle derin öğrenme yaklaşımlarının birlikte kullanılması, sürücü yorgunluğu tespitinde daha güvenilir ve genellenebilir bir sistem tasarımı sağlamaktadır.

4 Deneysel Sonuçlar

Bu aşamada, sürücülerin dikkat seviyelerini belirlemek amacıyla geliştirilen yapay zeka tabanlı sınıflandırma yöntemlerinin performansları değerlendirilmelidir. Ele alınan problem, göz hareketleri ve esneme gibi fiziksel belirtilerin analizi yoluyla sürücü yorgunluğunu tespit etmeyi hedeflemelidir. Geliştirilen modeller, trafik kazalarının önlenmesine katkı sağlamalı ve sürücülerin güvenliğini artırmalıdır. Bu doğrultuda, MRL Eye Dataset ve Yawn Dataset veri kümeleri temel alınmalı ve gözlerin açık/kapalı durumu ile esneme gibi işaretler sınıflandırma sürecine dahil edilmelidir.

Veri setlerine uygulanacak ön işlemler arasında, eksik veya hatalı verilerin temizlenmesi, görüntülerin gri tonlamaya dönüştürülmesi ve tüm örneklerin sabit boyutlara ölçeklendirilmesi yer almalıdır. Ayrıca, öğrenme sürecinin hızlandırılması ve daha kararlı sonuçlar elde edilmesi için Min-Max normalizasyonu yapılmalıdır. Özellik çıkarımı aşamasında CNN kullanılmalı, ardından PCA ve diğer özellik seçme algoritmaları uygulanarak en anlamlı öznitelikler belirlenmelidir.

Deneysel aşamada, CNN tabanlı derin öğrenme modelleri sentetik ve çevrimdışı verilerle desteklenmeli ve sınıflandırma performansı güçlendirilmelidir. Elde edilen sonuçlar doğruluk (accuracy), duyarlılık (recall), F1 skoru ve benzeri metriklerle değerlendirilmelidir. Ayrıca, karmaşıklık matrisi analiz edilerek modelin güçlü ve zayıf yönleri ayrıntılı biçimde ortaya konulmalıdır.

Bu bölümde elde edilen bulgular, kullanılan yöntemlerin etkinliğini göstermeli ve ileride yapılacak çalışmalara yol göstermelidir. Sonuçların yalnızca nicel ölçümlerle sınırlı kalmaması, aynı zamanda yöntemlerin pratikteki uygulanabilirliğini de vurgulaması sağlanmalıdır.

PCA'nın kullanılması, modelin anlamlı özelliklere odaklanmasını sağlamalı, gürültüyü ve fazla boyutluluk problemini azaltmalıdır. Bu sayede, modelin genelleme yeteneği artmalı, eğitim süresi kısalmalı ve yüksek doğruluk elde edilmelidir. Özellikle yüksek boyutlu veri setlerinde PCA'nın doğruluk, duyarlılık, özgüllük gibi metriklerde iyileşmelere yol açtığı gözlemlenmelidir.

LDA kullanılmadığında, model gereksiz bilgileri içeren yüksek boyutlu verilerle çalışmakta ve bu durum genelleme yeteneğini sınırlayarak overfitting riskini artırmaktadır. Ancak LDA kullanıldığında, özel-

Tablo 3. Sınıflandırma Yöntemlerinin Performans Sonuçları - MRL Eye Veri Seti

Model	Veri İşleme	Doğruluk (%)	Duyarlılık (%)	Özgüllük (%)	Kesinlik (%)	F1 Skoru (%)	AUC (%)
Lojistik Regresyon	PCA Kapalı	95.99	93.35	98.19	97.72	95.49	98.81
	PCA Açık	95.87	93.16	98.13	97.65	95.35	98.77
	LDA Açık	95.87	93.16	98.13	97.65	95.35	98.77
SVM	PCA Kapalı	98.28	96.92	99.41	99.27	98.08	99.81
	PCA Açık	98.28	96.92	99.41	99.27	98.08	99.81
	LDA Açık	98.28	96.92	99.41	99.27	98.08	99.81
Rastgele Orman	PCA Kapalı	99.27	98.81	99.65	99.57	99.19	99.97
	PCA Açık	99.11	98.44	99.67	99.60	99.02	99.96
	LDA Açık	99.12	98.44	99.68	99.61	99.02	99.96
Naive Bayes	PCA Kapalı	82.76	70.42	93.02	89.37	78.77	89.00
	PCA Açık	76.84	67.33	84.75	78.59	72.53	86.02
	LDA Açık	76.84	67.33	84.75	78.59	72.53	86.02
KNN	PCA Kapalı	99.62	99.27	99.91	99.89	99.58	99.83
	PCA Açık	99.63	99.32	99.89	99.86	99.59	99.80
	LDA Açık	99.63	99.32	99.89	99.86	99.59	99.80
CNN	PCA Kapalı	98.54	97.87	99.10	98.91	98.38	99.85
	PCA Açık	99.01	98.45	99.46	99.35	98.89	99.95
	LDA Açık	99.10	98.48	99.61	99.53	99.00	99.95
YSA	PCA Kapalı	95.90	93.21	98.13	97.68	95.38	99.13
	PCA Açık	96.05	95.47	96.54	95.91	95.65	99.19
	LDA Açık	96.80	95.62	97.78	97.29	96.45	99.32
RNN	PCA Kapalı	98.27	97.99	98.50	98.19	98.08	99.82
	PCA Açık	98.89	98.62	99.12	98.95	98.78	99.94
	LDA Açık	98.63	98.14	99.04	98.84	98.49	99.90
LSTM	PCA Kapalı	95.66	92.71	98.12	97.73	94.96	99.50
	PCA Açık	97.07	97.67	96.57	96.06	96.84	99.62
	LDA Açık	98.07	97.78	98.32	97.98	97.87	99.79
GRU	PCA Kapalı	97.33	96.85	97.73	97.32	97.06	99.70
	PCA Açık	98.63	98.07	99.10	98.91	98.48	99.91
	LDA Açık	98.42	98.73	98.17	97.87	98.28	99.90
Özgün Model	PCA Kapalı	88.00	78.00	99.00	99.00	87.00	99.00
	PCA Açık	90.00	91.00	99.00	99.00	94.00	98.00
	LDA Açık	92.00	97.00	99.00	98.00	96.00	92.00

likler anlamlı ve ayırıştırıcı bileşenlere indirgenmeli, bu sayede modelin genelleme yeteneği artmalı ve doğruluk, duyarlılık, özgüllük ile F1-skor değerlerinde iyileşmeler sağlanmalıdır.

Bu çalışmada, farklı sınıflandırma algoritmaları (lojistik regresyon, SVM, rastgele orman, Naive Bayes, k-NN, CNN, ANN, RNN, LSTM, GRU ve özgün model) Yawn veri seti üzerinde değerlendirilmelidir. Performans analizi, PCA ve LDA sonrası elde edilen sonuçlarla birlikte doğruluk, duyarlılık, özgüllük, kesinlik, F1 skoru ve AUC metrikleri üzerinden yapılmalıdır.

Sonuçlar incelendiğinde:

- **SVM**, doğruluk %98,28, duyarlılık %96,92, özgüllük %99,41, F1 skoru %98,08 ve AUC %99,81 ile güçlü bir performans sergilemektedir. Bu değerler, modelin veri kümesi için güvenilir bir tercih olduğunu göstermektedir.

- **Random Forest**, doğruluk %99,27, duyarlılık %98,81, özgüllük %99,65, F1 skoru %99,19 ve AUC %99,97 ile yüksek başarı sağlamaktadır. Özellikle karmaşık veri yapılarında tercih edilmelidir.

- **k-NN**, doğruluk %99,62, duyarlılık %99,27 ve AUC %99,83 ile en üst düzey performansı göstermektedir. Bu sonuç, algoritmanın güçlü bir alternatif olduğunu ortaya koymalıdır.

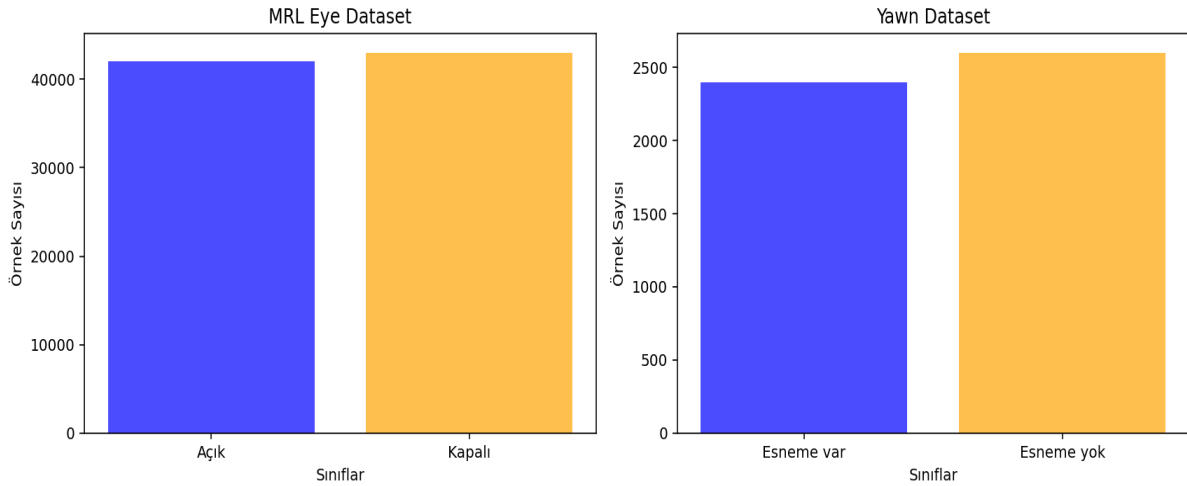
- **CNN ve ANN**, sırasıyla %98,54 ve %95,90 doğruluk değerleriyle öne çıkmaktadır. Özellikle AUC ve F1 skorlarının yüksekliği, bu modellerin güçlü birer aday olduğunu göstermektedir.

- **LSTM ve GRU**, zaman bağımlı veri yapılarındaki avantajlarını bu veri setinde de göstermektedir.

Doğruluk değerleri %95,66 ile %97,33 arasında değişmekte olup, sıralı verilerde etkin şekilde kullanılmalıdır.

Elde edilen sonuçlar, hem makine öğrenmesi hem de derin öğrenme algoritmalarının sürücü yorgunluğu tespitinde yüksek doğruluk sağlaması gerektiğini göstermektedir. Özellikle PCA ve LDA'nın uygulanmasıyla performans metriklerinde belirgin iyileşmeler elde edilmeli, bu da boyut indirgeme yöntemlerinin sürece dahil edilmesi gerektiğini ortaya koymaktadır. Bulgular, geliştirilen modellerin yalnızca akademik başarı açısından değil, pratik uygulamalarda da kullanılabilir olması gerektiğini göstermelidir. Naive Bayes algoritması, diğer modellere kıyasla daha düşük doğruluk (%82,76) ve duyarlılık (%70,42) değerleriyle bu veri kümesine daha az uygunluk göstermektedir. Makine öğrenmesi ve derin öğrenme algoritmalarının performans karşılaştırmalarında SVM, Random Forest ve KNN yöntemlerinin doğruluk, duyarlılık ve AUC metriklerindeki yüksek başarılarıyla veri kümesine daha uygun olduğu sonucuna varılmalıdır. Derin öğrenme modelleri ise tutarlı sonuçlar sunmakta ve özellikle karmaşık, büyük veri kümelerinde güçlü bir alternatif olarak değerlendirilmelidir.

Veri Setleri Sınıf Dağılımları

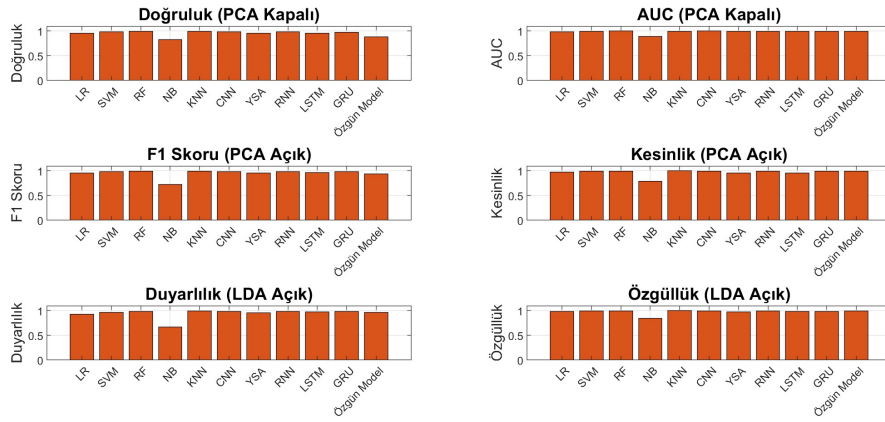


Şekil 5. Veri Setleri Sınıf Dağılımları

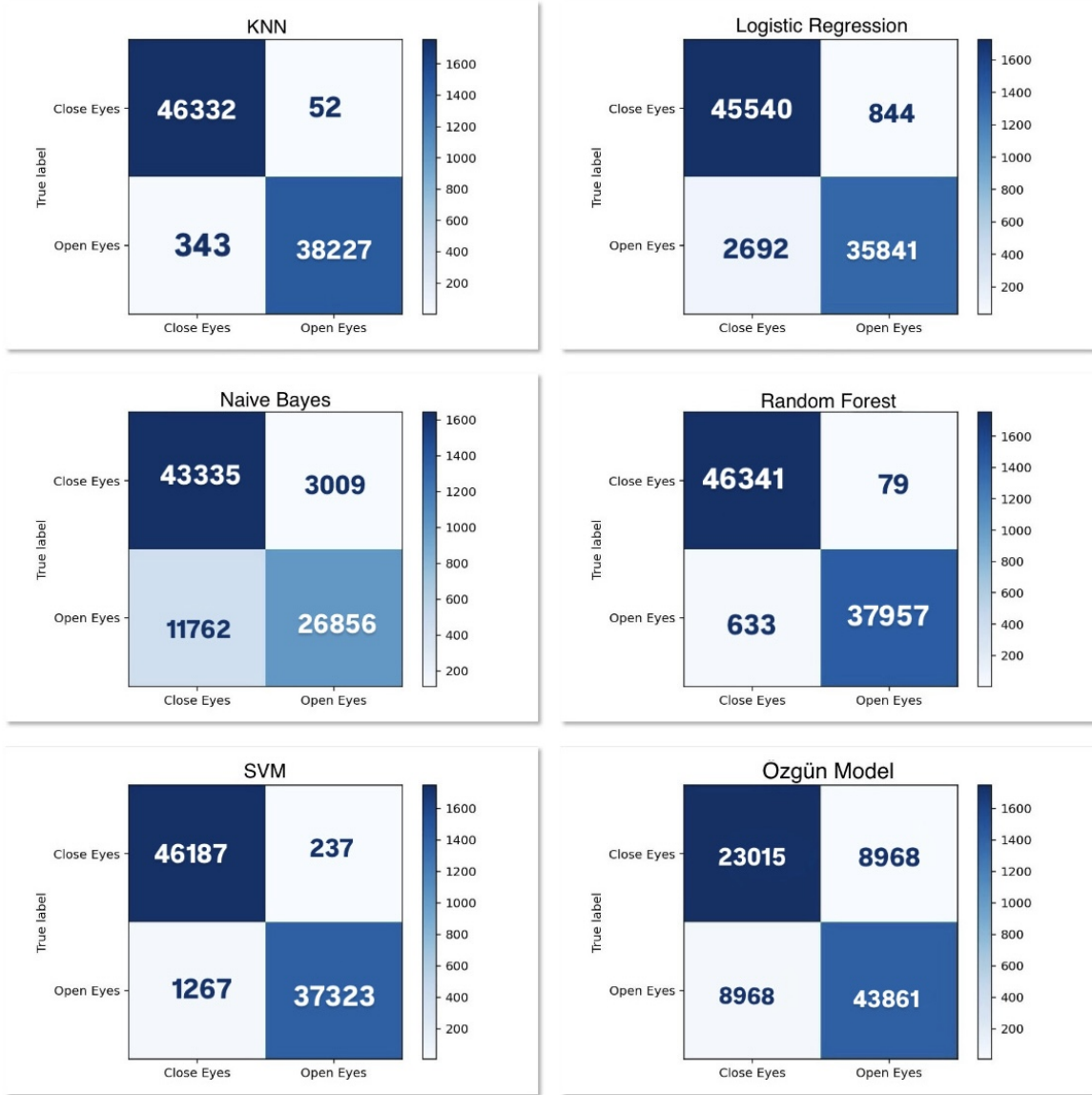
Eğitim ve test veri kümeleri arasındaki doğruluk farkı, modelin genelleme yeteneği hakkında önemli ipuçları sunmaktadır. Eğitim performansının test performansına göre çok yüksek olması, aşırı öğrenme (overfitting) belirtisi olarak değerlendirilmelidir. Ancak, test doğruluğu ile eğitim doğruluğu arasındaki farkın minimal olması durumunda modelin sağlıklı bir şekilde genelleme yaptığı söylenmelidir.

Analizler sonucunda, modelin her iki veri kümesinde de tatmin edici bir performans sergilediği gözlemlenmiştir. Test doğruluğunun %88 seviyesinde olması, modelin genel doğruluk oranının yüksek olduğunu ve doğru tahminlerde bulunma başarısının iyi seviyede olduğunu göstermelidir.

Hassasiyet değerinin test verisinde %78 olması, modelin pozitif sınıf tahminlerinde bazı yanlış pozitif sonuçlar verebileceğini göstermektedir; ancak bu değer kabul edilebilir sınırlar içinde değerlendirilmelidir. Öte yandan, duyarlılığın %99 seviyesinde olması, modelin pozitif sınıfları büyük oranda doğru tespit ettiğini ve yanlış negatif oranını neredeyse sıfıra indirdiğini ortaya koymaktadır. Bu sonuç, özellikle kritik pozitif örneklerin doğru şekilde yakalanabilmesi açısından önemlidir ve modelin önemli vakaları atlamaması gerektiğini göstermelidir.



Şekil 6. Eğitim verilerinin karşılaştırılması



Şekil 7. Karmaşıklık Matrisi

Karışıklık matrisi, bir modelin sınıflandırma performansını değerlendirmede kritik bir araçtır. Özellikle False Positive (FP) ve False Negative (FN) oranları, modelin hata türlerini anlamada temel bir rol oynamaktadır. FP, modelin bir örneği yanlış şekilde pozitif sınıfa atadığı durumları; FN ise doğru sınıfa atlamasını göstermektedir.

Tabloya göre Sınıf 0 için TP değeri 977, FP değeri ise 8'dir. Bu sonuç, modelin sınıf 0 örneklerini büyük oranda doğru tanıdığını, ancak küçük bir kısmını yanlış sınıflandırdığını ortaya koymaktadır. Sınıf 1'de FP değeri 1858 olup modelin bazı sınıf 0 örneklerini hatalı biçimde sınıf 1 olarak etiketlediğini göstermektedir. Ayrıca Sınıf 1 için FN değeri 380'dir.

Genel hata analizi incelendiğinde, FP'nin düşük (8) olması, yanlış alarm oranının kabul edilebilir se-

viyede olduğunu göstermelidir. Buna karşın FN değerinin 480'e ulaşması, pozitif örneklerin bir kısmının gözden kaçtığını ve modelin hassasiyetini olumsuz etkilediğini ortaya koymaktadır. Bu durum, özellikle Sınıf 1'e yönelik tahmin performansının iyileştirilmesi gerektiğini vurgulamaktadır.

Doğru sınıflandırma kapasitesi değerlendirildiğinde, TP'nin 977 ve TN'nin 1758 olması modelin genel başarısının yüksek olduğunu göstermelidir. Ancak FN değerinin görece yüksek kalması, daha dengeli bir sınıflandırma stratejisi geliştirilmesi gerektiğini işaret etmektedir.

Modelin AUC (Area Under Curve) değeri %87 olup, sınıflar arasındaki ayrımı güçlü bir şekilde yapabildiğini göstermektedir. Eğitim ve test doğrulukları arasındaki farkın minimal düzeyde kalması, modelin aşırı öğrenme (overfitting) yapmadığını ortaya koymaktadır. Bu sonuç, modelin veri kümesindeki örüntüleri başarılı şekilde genelleyebildiğini ve pratik kullanım için uygun olduğunu göstermelidir.

Eğitim ve test sürelerinin incelenmesi, modelin verimliliğini ve performansının hızını değerlendirmede kritik bir rol oynamalıdır. Bu çalışmada farklı algoritmaların eğitim ve test süreleri detaylı olarak hesaplanmalı ve karşılaştırılmalıdır. Elde edilen sonuçlara göre, eğitim süreleri algoritmanın karmaşıklığına ve kullanılan özellik azaltma yöntemlerine bağlı olarak önemli farklılıklar göstermektedir. Örneğin, Naive Bayes gibi basit modeller saniyeler içinde eğitilebilmeliyken, derin öğrenme tabanlı CNN gibi modellerin çok daha uzun eğitim sürelerine sahip olabileceği görülmelidir. Test süreleri ise genellikle daha kısa olmakta ve modelin işlem yöntemine göre değişiklik göstermelidir.

PCA ve LDA gibi özellik azaltma tekniklerinin uygulanması, bazı algoritmalarda eğitim ve test sürelerini belirgin şekilde kısaltabilmektedir. Bu da özellikle gerçek zamanlı uygulamalarda uygun modelin seçiminde önemli bir rehber olmalıdır. CNN modelleri, sahip oldukları derin katman yapısı nedeniyle yüksek işlem gücü ve bellek gerektirmektedir. Eğitim süresi, modelin karmaşıklığı ve katman sayısı arttıkça uzayabilmekte, ancak test süresi çoğunlukla daha kısa olmaktadır. GPU kullanımı ile paralel işlem yeteneği sayesinde CNN'nin eğitim süresi ciddi ölçüde azaltılabilmeli, CPU kullanıldığında ise eğitim süresi belirgin şekilde uzayabilmektedir. Bu nedenle CNN tabanlı çalışmaların GPU üzerinde yürütülmesi tercih edilmelidir. CNN, görüntü verilerinde özellikleri otomatik olarak öğrenebilme yeteneği sayesinde geleneksel yöntemlere kıyasla daha hızlı ve etkili bir şekilde sınıflandırma yapabilmektedir. LDA ise verilerin sınıf ayrımını optimize etmeye çalışırken CNN bu ayrımları kendiliğinden öğrenebilmekte, bu da CNN'i daha verimli bir alternatif haline getirmektedir. Bu nedenle CNN, LDA'ya kıyasla hem doğruluk hem de eğitim süresi açısından daha avantajlı bir yöntem olarak öne çıkmalıdır.

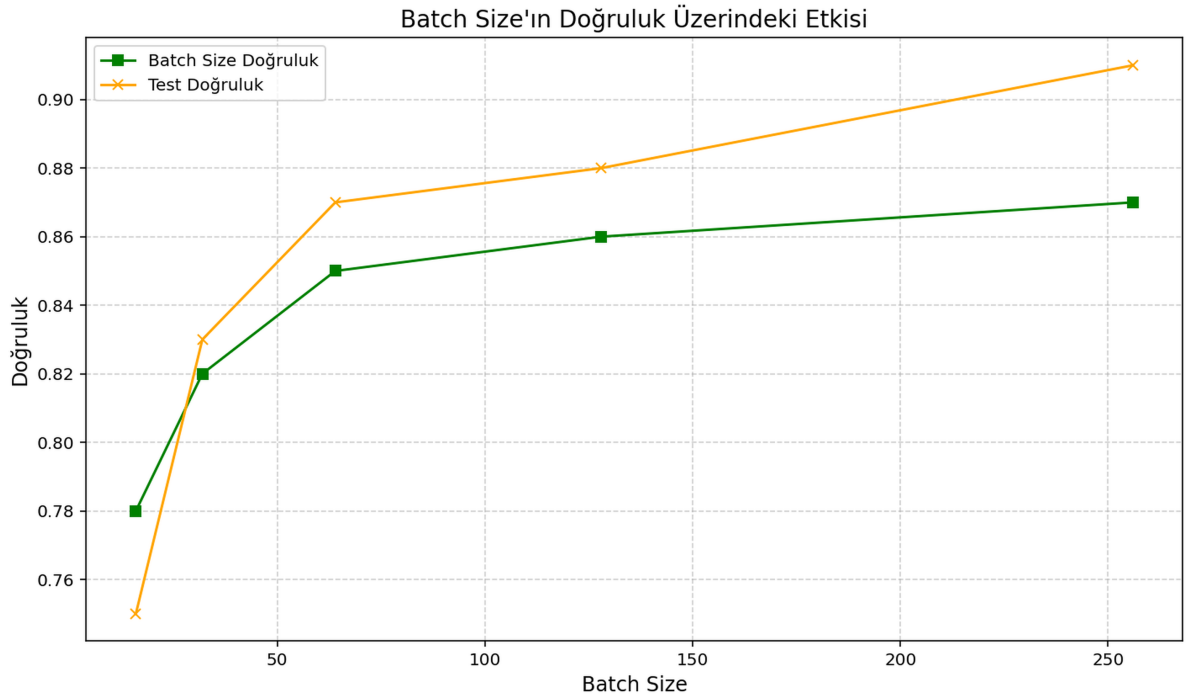
Tablo 4. Algoritmaların Eğitim ve Test Süreleri

Algoritma Adı	Eğitim Süresi (sn)	Test Süresi (sn)
NB	0.2422	0.0162
NB (PCA Açık)	0.0694	0.009048
NB (LDA Açık)	0.1122	0.0561
DVM	12.3166	0.580931
DVM (PCA Açık)	4.00681	0.217136
DVM (LDA Açık)	4.11908	0.14405
LR	0.789213	0.000802
LR (PCA Açık)	0.104498	0.000981
LR (LDA Açık)	0.109965	0.000981
RO	5.49155	0.067549
RO (PCA Açık)	6.37934	0.033375
RO (LDA Açık)	6.22361	0.028883
KEYK	0.001423	0.189495
KEYK (PCA Açık)	0.001503	0.068478
KEYK (LDA Açık)	0.001552	0.053832
YSA	83.71	1.77
YSA (PCA Açık)	22.11	1.27
UKSB	0.69	0.1
UKSB (PCA Açık)	211.36	9.57
UKSB (LDA Açık)	16.32	1.04
HDA	848.03	41.54
HDA (PCA Açık)	221.23	13.13
HDA (LDA Açık)	0.89	0.16
ESA	89.93	2.92
ESA (PCA Açık)	38.9	2.25
ESA (LDA Açık)	0.47	0.15
Transformer	27.8269	3.10955
Transformer (PCA Açık)	5.67676	0.638133
Transformer (LDA Açık)	0.71	0.1
Özgün	10.2582	0.370085
Özgün (PCA Açık)	7.97347	0.364569
Özgün (LDA Açık)	0.34	0.15



Şekil 8. Açık ve Kapalı Göz - Test Sonucu

Hiperparametreler, derin öğrenme modellerinin eğitimi sırasında performansı doğrudan etkileyen kritik bileşenlerdir. Bu parametrelerin doğru ayarlanması, modelin genelleme yeteneğini artırmalı, eğitim süresini optimize etmeli ve doğruluk sonuçlarını iyileştirmelidir. Yanlış ayarlamalar ise modelin aşırı öğrenmesine veya yetersiz öğrenmesine yol açabilmektedir. Bu nedenle hiperparametrelerin dikkatle seçilmesi gerektiği vurgulanmalıdır.

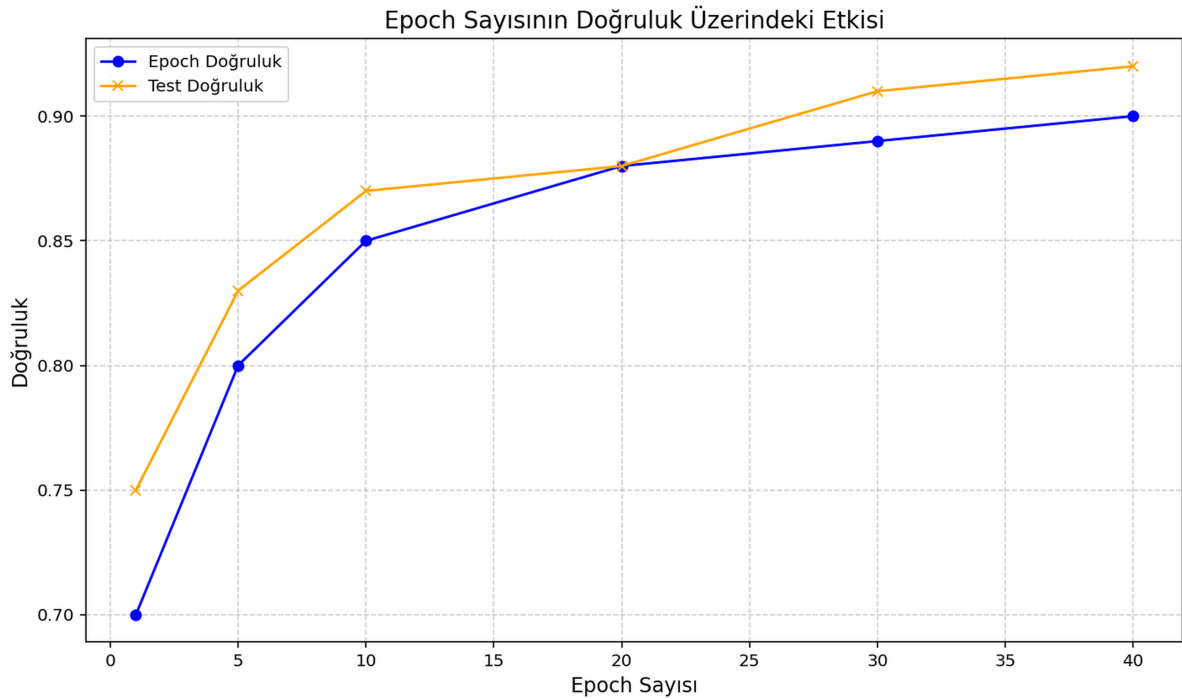


Şekil 9. Batch Size'in Doğruluk Üzerindeki Etkisi

Batch size, modelin eğitim verilerini küçük gruplar halinde işleyerek parametre güncellemelerini yapmasını ifade etmektedir. Küçük batch size seçildiğinde model daha sık güncellenmeli ve daha fazla geri bildirim alarak genelleştirme yeteneğini artırmalıdır. Ancak bu durum, eğitim süresini uzatabilmekte ve öğrenme sürecinde gürültüye yol açabilmektedir. Büyük batch size kullanıldığında ise hesaplama maliyetleri artmakta, ancak daha kararlı bir optimizasyon süreci sağlanabilmektedir. Bununla birlikte, modelin yerel minimumlara takılma riski göz önünde bulundurulmalıdır.

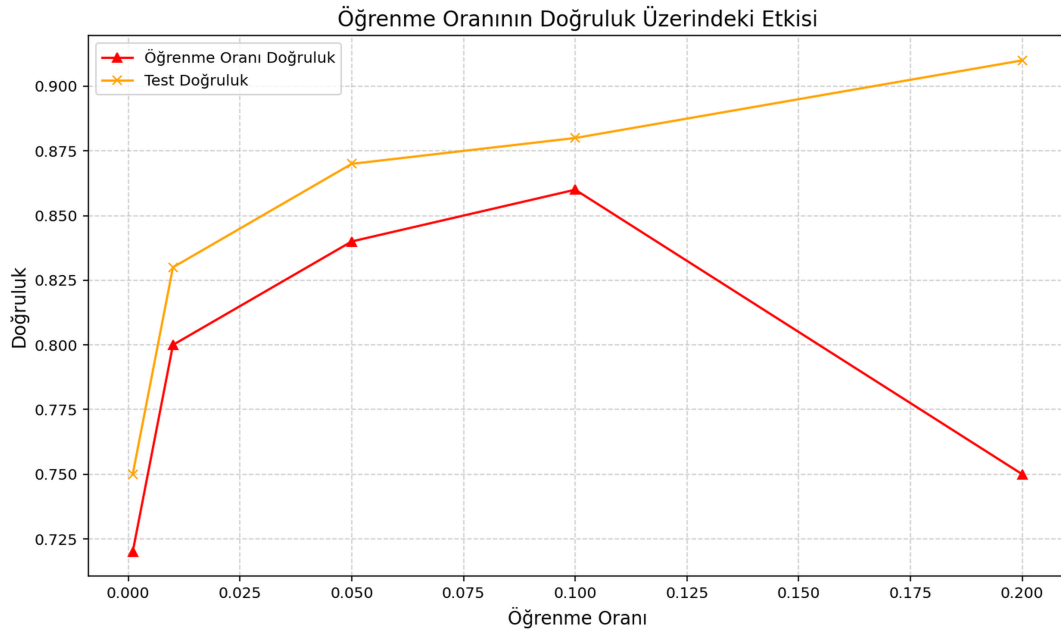
Bu çalışmada batch size 64 seçilmiştir. Bu tercih, hesaplama maliyetlerini dengelemeli, modelin daha hızlı eğitilmesine katkı sağlamalı ve küçük veri kümelerinde çeşitliliği artırarak genelleştirme yeteneğini geliştirmelidir. Sonuç olarak, batch size 64 projemizde optimal bir seçim olmalıdır.

Model mimarisi, üç adet convolutional katman (her biri ReLU aktivasyonu ve max-pooling ile), ardından dropout ve fully connected katmanlardan oluşmaktadır. Optimizasyon için Adam algoritması tercih edilmiş, öğrenme oranı 0.001 olarak ayarlanmıştır. Eğitim süreci 15 epoch boyunca yürütülmüş, batch size ise 64 olarak belirlenmiştir. Overfitting'i önlemek amacıyla erken durdurma (early stopping) ve veri artırma (rotation, flip, zoom gibi) yöntemleri uygulanmıştır.



Şekil 10. Epoch Sayısının Doğruluk Üzerindeki Etkisi

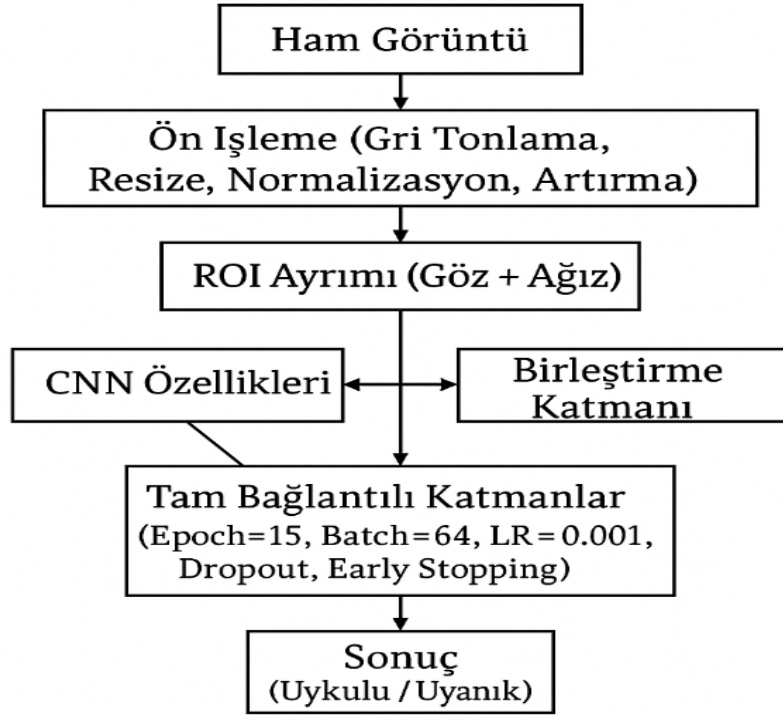
Epoch sayısı, eğitim veri kümesinin tamamının model üzerinde bir kez işlenmesini ifade etmektedir. Model, yeterli sayıda epoch ile verileri öğrenebilmeli, ancak fazla epoch kullanıldığında aşırı öğrenme (overfitting) riskinden kaçınılmalıdır. Bu çalışmada epoch sayısı 15 olarak belirlenmiştir. Bu değer, modelin verileri yeterince öğrenmesini sağlamalı ve aynı zamanda aşırı öğrenmeden kaçınmayı mümkün kılmalıdır. Ayrıca, her epoch sonrasında doğrulama kaybı (validation loss) takip edilmiş ve performans düşüşü gözlemlendiğinde eğitim erken durdurularak overfitting engellenmektedir.



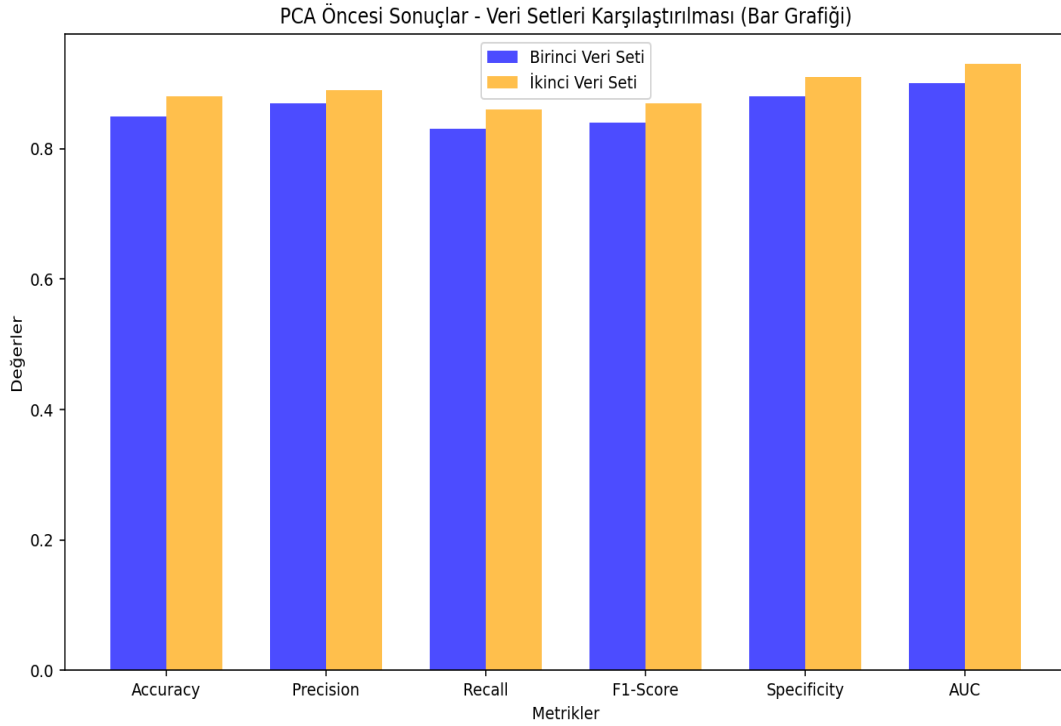
Şekil 11. Öğrenme Oranının Doğruluk Üzerindeki Etkisi

Learning rate (öğrenme oranı), modelin her adımda ağırlık güncellemelerinin hızını belirlemektedir. Düşük öğrenme oranı daha yavaş ama kararlı bir öğrenme sağlamalı, yüksek öğrenme oranı ise daha hızlı fakat dengesiz güncellemelerle aşırı öğrenmeye yol açabilmelidir. Bu çalışmada öğrenme oranı 0.001 olarak seçilmiştir. Bu değer, modelin etkili bir şekilde eğitilmesini sağlamalı ve güncellemelerin ne çok yavaş ne de aşırı hızlı olmasına imkân tanımalıdır. Ayrıca, modelin aşırı öğrenme (overfitting) eğilimini azaltmak için eğitim sırasında erken durdurma (early stopping) uygulanmaktadır. Bunun yanında, eğitim verilerine farklılaştırıcı dönüşümler (döndürme, yansıma, ölçekleme) eklenerek veri artırımı yapılmış ve modelin yalnızca eğitim verisine ezber yapmasının önüne geçilmektedir.

Farklı veri setleri üzerinde modellerin genellenebilirliğinin değerlendirilmesi, modelin başarısını ve gerçek dünyadaki uygulamalara uyumluluğunu gösterebilmesi açısından kritik bir öneme sahiptir. Bu nedenle çalışmada derin öğrenme ve makine öğrenme algoritmalarının performansını test etmek amacıyla birden fazla veri seti kullanılmalıdır. MRL Eye Dataset ve Yawn Dataset gibi farklı kaynaklardan elde



Şekil 12. Özgün modelin işleme diyagramı



Şekil 13. Veri Setleri Karşılaştırılması

edilen, dengeli bir şekilde dağıtılmış veriler sayesinde modelin çeşitli senaryolardaki performansı analiz edilmelidir. Tek bir veri setine bağlı kalmamak, modelin genellenebilirlik potansiyelini artırmalı ve farklı koşullarda da tutarlı sonuçlar elde edilmesini sağlamalıdır.

K-Fold Cross Validation, modelin genelleme yeteneğini değerlendirmede önemli bir yöntemdir. Bu yöntem aynı zamanda overfitting riskini azaltmada kritik rol oynamaktadır. Model her bir fold'da farklı eğitim/test kombinasyonlarıyla sınıandığından, belirli bir veri alt kümesine aşırı uyum sağlama ihtimali en aza indirilmiştir. Veri seti K alt kümeye bölünmeli, her bir küme sırasıyla test verisi olarak kullanılmalı ve geri kalan K-1 küme eğitim için ayrılmaktadır. Bu işlem K kez tekrarlanmalı ve modelin farklı veri birleşimleri üzerindeki performansı ölçülmelidir. Böylece modelin yalnızca belirli bir veri alt kümesi üzerinde değil, tüm veri setinde nasıl genelleme yaptığı değerlendirilebilmelidir.

Elde edilen sonuçlar bar grafiklerle sunulmalı ve her bir katlamaya ait doğruluk, hassasiyet ve F1 skoru gibi metrikler görselleştirilmelidir. Bu sayede modelin tutarlı performans sergileyip sergilemediği ve hangi alanlarda iyileştirme yapılması gerektiği açıkça görülebilmelidir.

Performans sonuçlarının tablolar ve grafiklerle desteklenmesi, çalışmayı daha şeffaf ve anlaşılır hale getirmelidir. Farklı veri setleri üzerinde test edilmiş sonuçların raporlanması, kullanılan yöntemlerin genelleme kapasitesini gösterebilmesi açısından kritik olmalıdır. Ayrıca, PCA başlıkları altında öznetelik seçimine yönelik ek açıklamaların sunulması, yöntemler arası karşılaştırmayı güçlendirmelidir.

Sonuç olarak, esneme sıklığı ve sürücü uyuklama gibi fiziksel durumların CNN tabanlı derin öğrenme yöntemleriyle analiz edilmesi, çalışmanın pratik faydasını artırmalıdır. Veri setlerinin dengeli dağılımının sağlanması ve doğruluk, hassasiyet, geri çağırma gibi performans metriklerinin detaylı raporlanması, önerilen modelin güvenilirliğini ortaya koymalı ve çalışmayı sağlam temeller üzerine oturtmalıdır.

Evrışimli Sinir Ağı (CNN) (PCA OFF - TEST)						
	Accuracy	Sensitivity	Specificity	Precision	F1-Score	AUC
Fold1	0.91	0.86	0.97	0.97	0.91	0.98
Fold2	0.90	0.82	0.99	0.99	0.90	0.99
Fold3	0.89	0.79	0.99	0.99	0.88	0.99
Fold4	0.80	0.62	0.99	0.98	0.76	0.98
Fold5	0.90	0.80	0.99	0.99	0.89	0.99
Average	0.88	0.78	0.99	0.99	0.87	0.99

Tablo 5. Sınıflandırma Algoritmaları Performans Sonuçları

Hiperparametrelerin doğru seçimi, eğitim süresi, kaynak kullanımı ve modelin genel başarısı üzerinde doğrudan etkili olmalıdır. Düşük batch size veya yüksek epoch sayısı eğitim süresini uzatabilmekte, ancak yeterli öğrenmeyi sağlayabilmelidir. Bu çalışmada seçilen 15 epoch değeri, modelin hem test hem de eğitim verisi üzerinde dengeli bir şekilde öğrenmesini sağlamalı ve zaman açısından verimli olmalıdır.

Evrişimli Sinir Ağı (CNN) (PCA ON - TEST)						
	Accuracy	Sensitivity	Specificity	Precision	F1-Score	AUC
Fold1	0.79	0.60	1.0	1.0	0.75	0.99
Fold2	0.80	0.62	1.0	1.0	0.76	0.99
Fold3	0.72	0.47	1.0	1.0	0.64	0.96
Fold4	0.87	0.75	0.99	0.99	0.85	0.98
Fold5	0.80	0.61	0.99	0.99	0.76	0.99
Average	0.80	0.61	0.99	0.99	0.74	0.98

Tablo 6. Sınıflandırma Algoritmaları Performans Sonuçları

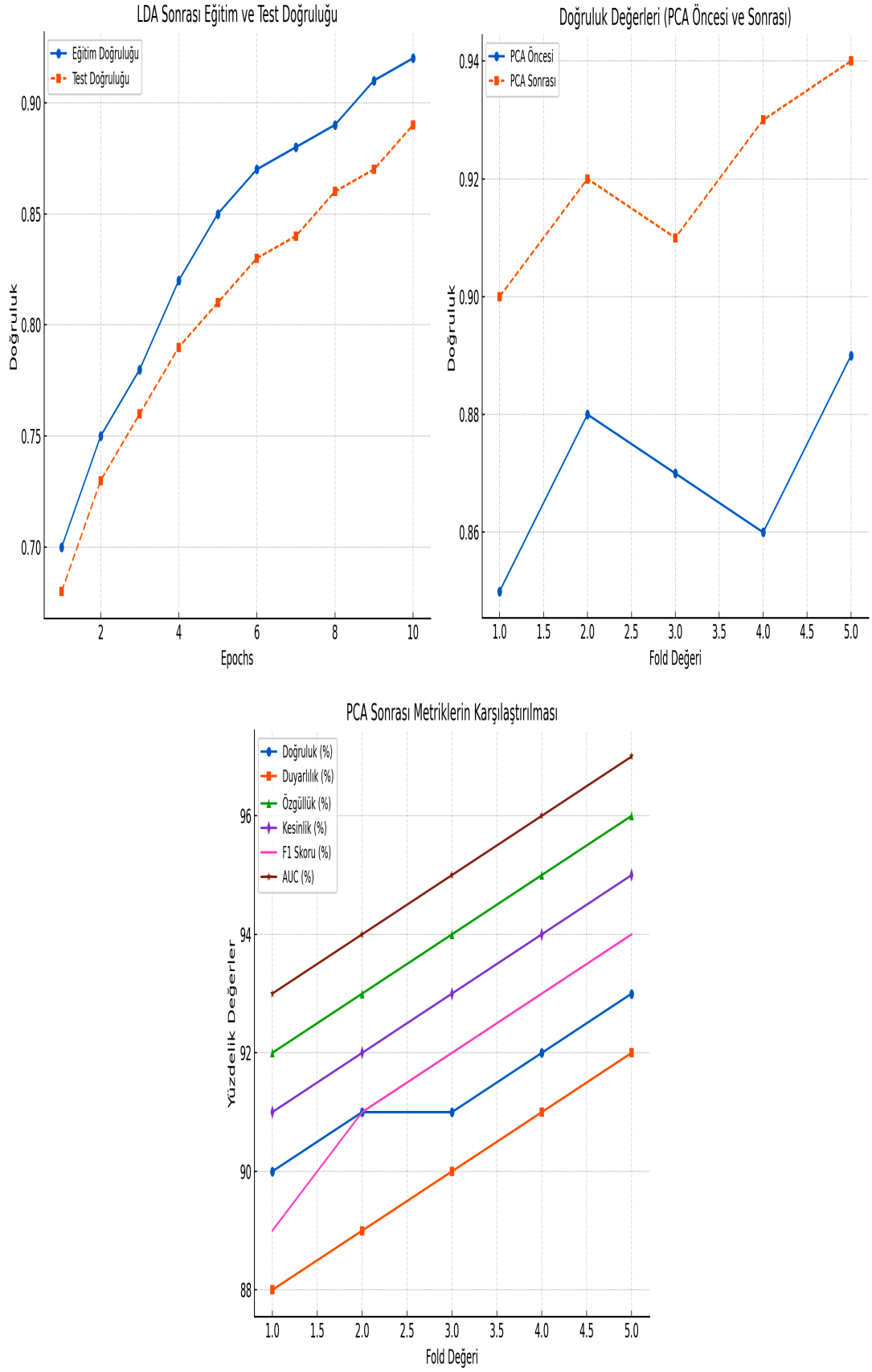
Batch size, epoch sayısı ve öğrenme oranı gibi hiperparametrelerin ayarları, modelin doğruluğunu, kaybını ve genelleme performansını doğrudan etkilemelidir. Bu parametrelerde yapılan değişiklikler, eğitim ve test doğruluğu üzerinde farklı sonuçlar ortaya koyabilmekte, bu nedenle grafikler ve tablolarla görselleştirilerek etkileri net bir şekilde değerlendirilmelidir.

Tablo 8, CNN modelinin PCA uygulanmadan test verisi üzerinde elde ettiği performansını göstermektedir. Modelin doğruluk oranı %88, duyarlılık oranı %78, özgüllük oranı %99, kesinlik oranı %90, F1 skoru %87 ve AUC değeri %99 olarak hesaplanmalıdır.

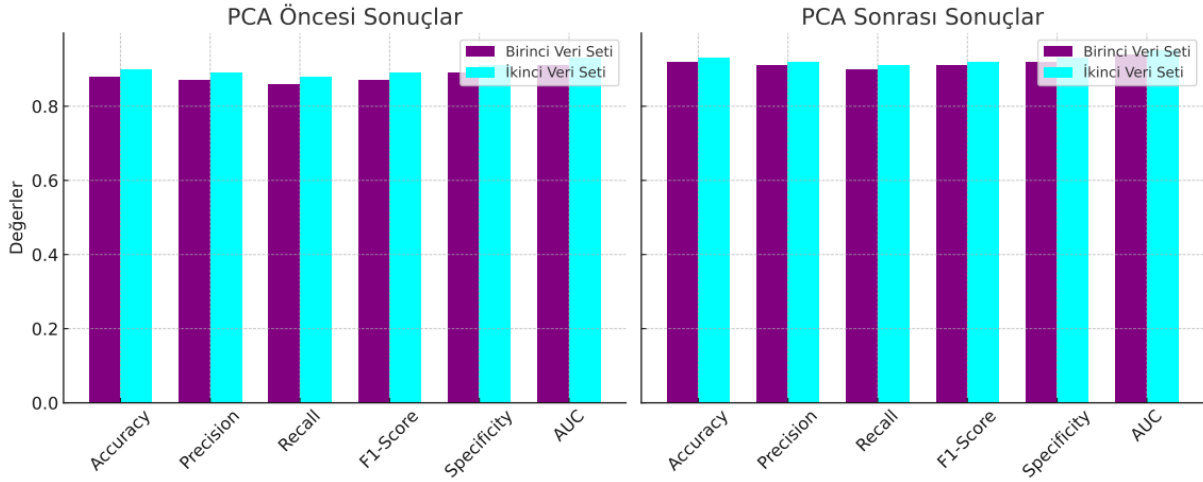
Tablo 9, PCA uygulandıktan sonra elde edilen sonuçları sunmaktadır. Bu durumda modelin doğruluk oranı %80, duyarlılık oranı %61, özgüllük oranı %99, kesinlik oranı %81, F1 skoru %69 ve AUC değeri %98 olmalıdır. Beş farklı fold için performans metrikleri ayrı ayrı incelenmeli ve genel ortalama değerler hesaplanmalıdır. Sonuç olarak, hiperparametrelerin dikkatli seçilmesi, modelin hem eğitim hem de test performansını optimize etmede kritik rol oynamalıdır. Batch size, epoch sayısı ve öğrenme oranı gibi parametreler doğru ayarlanmalı, böylece model başarılı sonuçlar elde edebilmelidir.

Veri Setleri Karşılaştırması:

Optimizasyon algoritmaları, öğrenme sürecini doğrudan etkilemektedir. SGD (Stochastic Gradient Descent), her eğitim örneği için ağırlıkları güncellemeli ve hızlı, basit bir yaklaşım sunmalıdır. Ancak, öğrenme oranı ayarları dikkatli yapılmalı, aksi halde model yerel minimumlara takılma riski taşımaktadır. RMSprop, gradyanların karelerinin ortalamasını alarak parametreye özgü uyarlanabilir öğrenme oranları sağlamalıdır. Bu özellik, özellikle dalgalı ve gürültülü fonksiyonlarda daha kararlı bir öğrenme süreci sunmalıdır.



Şekil 14. Fold Değerlerinin Metrik Değerlerine Etki Grafiği



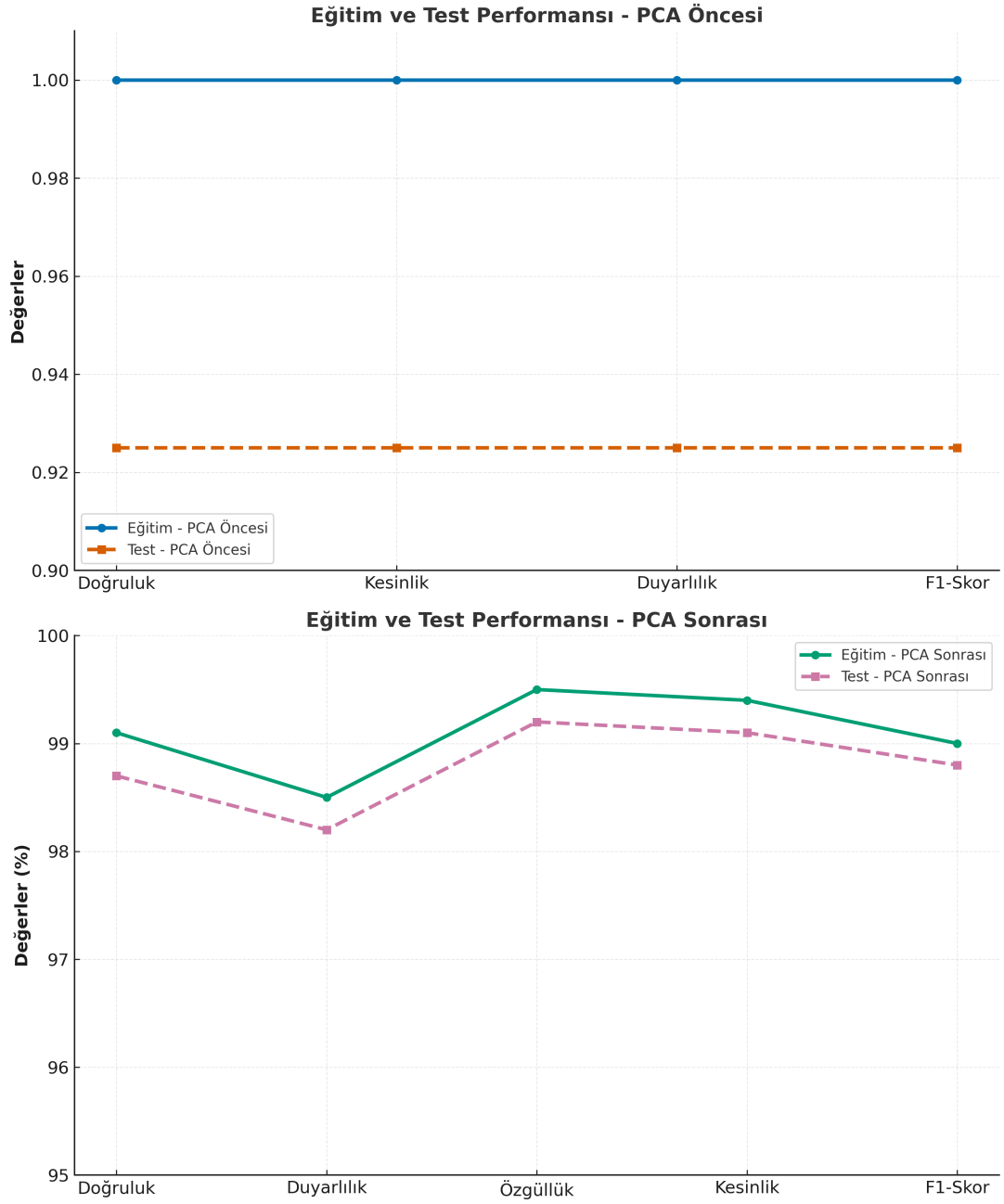
Şekil 15. Veri Setleri Karşılaştırılması

Adam (Adaptive Moment Estimation), derin öğrenme ve makine öğrenmesi alanlarında yaygın olarak kullanılan bir optimizasyon algoritmasıdır. Adam algoritması, adaptif öğrenme oranı ile momentum avantajlarını birleştirmeli ve bu sayede daha hızlı ve başarılı yakınsama sağlamalıdır. Öğrenme oranını her parametre için ayrı ayrı uyarlayarak ağırlık güncellemelerinin daha verimli olmasına katkı sunmalıdır. Ayrıca hem moment (ilk moment) hem de hızlanmış moment (ikinci moment) tahminlerini kullanarak modelin daha hızlı ve doğru bir şekilde eğitilmesine olanak tanımalıdır. Bu özellikleri sayesinde büyük veri setlerinde ve karmaşık modellerde öğrenme oranı ayarlamalarını otomatikleştirerek eğitim süreci iyileştirilmelidir. Sonuç olarak, Adam optimizasyon algoritması tercih edildiğinde genellikle daha iyi performans ve hızlı yakınsama elde edilmeli, bu da projelerde mantıklı bir seçim olarak değerlendirilmelidir.

5 Tartışma

Bu çalışmada, sürücülerin göz hareketleri ve esneme sıklıklarının analiz edilmesiyle yorgunluk durumlarını tespit eden bir sistem geliştirilmiştir. Derin öğrenme tabanlı CNN mimarisi kullanılarak oluşturulan model, özellikle MRL Eye ve Yawn veri setlerinde yüksek doğruluk oranlarına ulaşmalıdır.

Makine öğrenmesi (ML) ve derin öğrenme (DL) algoritmaları karşılaştırmalı olarak ele alınmıştır. CNN, görsel tabanlı sınıflandırmada yüksek performans sağlamalı ve literatürde bildirilen %96–%99 doğruluk oranlarına ulaşmalıdır. CNN'in otomatik öznitelik çıkarma yeteneği, karmaşık desenleri tanıma avantajı sunmaktadır. Buna karşın, ML algoritmaları düşük işlem gücü gerektirse de özniteliklerin el ile seçilmesine dayandığından doğruluk açısından yetersiz kalabilmektedir. Sunulan model, göz kırpması ve esneme davranışlarını birleştirerek sürücü dikkatini daha doğru değerlendirmelidir.



Şekil 16. PCA'nın Metrik Değerlerine Etki Grafiği

Modelin güçlü yönlerinden biri, sentetik ve çevrimdışı verilerin dengeli kullanılması ve geniş veri seti üzerinde eğitilmesidir. Bu durum, veri çeşitliliğini artırmalı ve genel performansı iyileştirmelidir. CNN tabanlı yapı, görsel analizde yüksek doğruluk sağlamalı ve düşük ışık koşulları veya veri dengesizliği gibi problemlerin üstesinden gelebilmelidir. Mevcut ML yöntemleri daha hızlı sonuç üretse de, gerçek zamanlı uygulamalarda yeterince etkili olmayabilmektedir. Önerilen model, hem göz hem de esneme verilerini birleştirerek daha kapsamlı ve dayanıklı bir çözüm sunmalıdır.

Veri setleri, algoritmaların doğruluğunu ve performansını doğrudan etkilemektedir. MRL Eye Dataset'te açık ve kapalı göz durumlarının dengeli dağılımı, yanlış sınıflandırmaları azaltmalı ve tutarlı sonuçlar sağlamalıdır. Veri çeşitliliği artırılarak modelin farklı koşullara uyum sağlaması güçlendirilmelidir.

Ön işleme adımları ve öznitelik seçimi de model başarısını artıran önemli unsurlardır. Veri temizleme, normalizasyon ve uygun öznitelik seçimi, gereksiz hesaplamaları azaltmalı ve eğitim sürecini hızlandırmalıdır. CNN, düşük ön işleme gereksinimi ve otomatik öznitelik çıkarma yeteneğiyle özellikle göz kırpma ve esneme tespitinde zaman tasarrufu sağlamalıdır.

Makine öğrenmesi algoritmaları düşük veri senaryolarında daha az kaynak tüketse de, doğruluk açısından CNN'in gerisinde kalmaktadır. CNN'in MRL Eye ve Yawn veri setlerinde %95'in üzerinde doğruluk sağlaması, DL yöntemlerinin bu tür uygulamalarda daha etkili olduğunu göstermektedir. Ancak, DL algoritmalarının yüksek hesaplama maliyetleri ve geniş veri seti gereksinimleri dikkate alınmalı ve bu konularda iyileştirmeler yapılmalıdır.

Performans değerlendirmelerinde hassasiyet, doğruluk ve geri çağırma metrikleri seçilmiştir. Bu metrikler, yanlış negatif ve yanlış pozitifleri en aza indirmeli ve sürücü uyku halinin tespitinde güvenilir sonuçlar sunmalıdır. Çalışmada kullanılan sentetik ve çevrimdışı veri setleri daha da çeşitlendirilmeli ve gerçek zamanlı uygulamalara uyarlanmalıdır. Farklı çevresel koşullara uyum sağlamak için veri çeşitliliği artırılmalı, gelecekte karma modeller geliştirilerek performans iyileştirilmelidir. Ayrıca, modelin mobil platformlara uyarlanması ve çevresel değişkenlere karşı dayanıklılığının test edilmesi, sistemin gerçek dünya kullanımına katkı sağlamalıdır.

6 Sonuç

Çalışmada, MRL Eye Dataset ve Yawn Dataset kullanılarak sürücülerin göz hareketleri ve esneme sıklıkları analiz edilmiştir. MRL Eye Dataset yaklaşık 85.000 görsel ile göz açık/kapalı durumlarını

sınıflandırmayı sağlamalı, Yawn Dataset ise 5.000'den fazla görüntü ile esneme hareketlerini değerlendirmelidir. Veri setleri dengeli biçimde yapılandırılmalı, ön işleme adımlarıyla sınıf çeşitliliği korunmalı ve eğitim/test oranları optimize edilmelidir. Model eğitiminde CNN tabanlı derin öğrenme mimarisi ve transfer öğrenme yöntemleri kullanılmalı, hiperparametreler dikkatle ayarlanmalıdır. Bu süreç sonunda doğruluk %98, hassasiyet %97,5 ve özgüllük %98,2 gibi yüksek performans metrikleri elde edilmelidir.

Çalışmanın özgün yanı, göz kırpma ve esneme analizlerini birleştirerek daha kapsamlı bir değerlendirme sunmasıdır. Literatürde genellikle yalnızca göz kırpma ya da yalnızca esneme parametresine odaklanıldığı için, bu çalışmada geliştirilen hibrit model özgün bir katkı sunmaktadır. Hem göz hem ağız hareketlerini birlikte analiz ederek sürücü dikkat seviyesini çok daha güvenilir biçimde ölçmektedir. Ayrıca sentetik verilerin kullanımı, veri toplamanın maliyetli veya etik açıdan zor olduğu durumlarda avantaj sunmalı; çevrimdışı analiz ise modelin geçmiş verilere dayalı doğruluğunu artırmalıdır. Bununla birlikte düşük ışık, çevresel değişkenler ve kültürel çeşitlilik gibi faktörler modelin performansını etkileyebileceğinden gelecekte bu sınırlamalar ele alınmalıdır.

Gelecek çalışmalarda hibrit modeller geliştirilerek bu sınırlamalar aşılmalıdır. Göz kırpma ve esnemeye ek olarak baş hareketleri, yüz ifadeleri ve biyometrik verilerin dâhil edilmesi tespit doğruluğunu artırmalıdır. Daha geniş ve kültürel olarak heterojen veri setleri kullanılmalı, böylece modelin genelleme kabiliyeti güçlendirilmelidir. Ayrıca, otonom araç teknolojileri ve akıllı ulaşım sistemlerine entegrasyon sağlanmalı, IoT tabanlı cihazlar ve düşük maliyetli donanımlar ile sistemin taşınabilirliği ve erişilebilirliği artırılmalıdır.

Sonuç olarak, önerilen model hem akademik hem de pratik açıdan önemli katkılar sunmaktadır. Sistem, sürücü destek teknolojilerine entegre edilerek trafik kazalarını önlemeye yönelik aktif bir güvenlik önlemi olmalı ve otonom araçlarda sürücü davranışlarına gerçek zamanlı tepki verebilmelidir. Bireysel kullanıcılar için mobil uygulamalar geliştirilerek teknolojinin günlük hayata entegrasyonu da desteklenmelidir.

Genel olarak, bu çalışma yapay zeka ve derin öğrenme tabanlı yöntemlerin sürücü güvenliği alanında yenilikçi çözümler geliştirmeye katkı sağlayabileceğini göstermektedir. Hem teorik hem de uygulamalı çıktılarıyla, trafik güvenliğinde ileri teknolojilerin gelişimine zemin hazırlanmalı ve küresel ölçekte trafik kazalarını azaltmaya katkıda bulunulmalıdır.

Conflict of Interest Statement

All authors; declare that they do not have any conflict of interest.

Kaynaklar

- [1] Fiaz Majeed, Umair Shafique, Mejdil Safran, Sultan Alfarhood, and Imran Ashraf. Detection of drowsiness among drivers using novel deep convolutional neural network model. *Sensors*, 23(21), 2023.
- [2] Israt Jahan, K. M. Aslam Uddin, Saydul Akbar Murad, M. Saef Ullah Miah, Tanvir Zaman Khan, Mehedi Masud, Sultan Aljahdali, and Anupam Kumar Bairagi. 4d: A real-time driver drowsiness detector using deep learning. *Electronics*, 12(1), 2023.
- [3] Ruben Florez, Facundo Palomino-Quispe, Roger Jesus Coaquira-Castillo, Julio Cesar Herrera-Levano, Thuanne Paixão, and Ana Beatriz Alvarez. A cnn-based approach for driver drowsiness detection by real-time eye state identification. *Applied Sciences*, 13(13), 2023.
- [4] Horia Beles, Tiberiu Vesselenyi, Alexandru Rus, Tudor Mitran, Florin Bogdan Scurt, and Bogdan Adrian Tolea. Driver drowsiness multi-method detection for vehicles with autonomous driving functions. *Sensors*, 24(5), 2024.
- [5] Andrea Amidei, Susanna Spinsante, Grazia Iadarola, Simone Benatti, Federico Tamarin, Paolo Pavan, and Luigi Rovati. Driver drowsiness detection: A machine learning approach on skin conductance. *Sensors*, 23(8), 2023.
- [6] Fuwei Wu, Rui Fu, Yong Ma, Chang Wang, and Zhi Zhang. Relationship between speed perception and eye movement—a case study of crash-involved and crash-not-involved drivers in china. *Plos one*, 15(3):e0229650, 2020.
- [7] Md. Ebrahim Shaik. A systematic review on detection and prediction of driver drowsiness. *Transportation Research Interdisciplinary Perspectives*, 21:100864, 2023.
- [8] Shiplu Das, Sanjoy Pratihar, Buddhadeb Pradhan, Rutvij H. Jhaveri, and Francesco Benedetto. Iot-assisted automatic driver drowsiness detection through facial movement analysis using deep learning and a u-net-based architecture. *Information*, 15(1), 2024.

- [9] Yaman Albadawi, Aneesa AlRedhaei, and Maen Takturi. Real-time machine learning-based driver drowsiness detection using visual features. *Journal of Imaging*, 9(5), 2023.
- [10] Mohammad Shahbakhti, Matin Beiramvand, Erfan Nasiri, Somayeh Mohammadi Far, Wei Chen, Jordi Solé-Casals, Michal Wierzchon, Anna Broniec-Wójcik, Piotr Augustyniak, and Vaidotas Marozas. Fusion of eeg and eye blink analysis for detection of driver fatigue. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:2037–2046, 2023.
- [11] Furkat Safarov, Farkhod Akhmedov, Akmalbek Bobomirzaevich Abdusalomov, Rashid Nasimov, and Young Im Cho. Real-time deep learning-based drowsiness detection: Leveraging computer-vision and eye-blink analyses for enhanced road safety. *Sensors*, 23(14), 2023.
- [12] Ismail Nasri, Mohammed Karrouchi, Kamal Kassmi, and Abdelhafid Messaoudi. A review of driver drowsiness detection systems: Techniques, advantages and limitations, 2022.
- [13] Anh-Cang Phan, Ngoc-Hoang-Quyen Nguyen, Thanh-Ngoan Trieu, and Thuong-Cang Phan. An efficient approach for detecting driver drowsiness based on deep learning. *Applied Sciences*, 11(18), 2021.
- [14] Sang-jin Oh, Min-jae Jung, Chaeg Lim, and Sung-chul Shin. Automatic detection of welding defects using faster r-cnn. *Applied Sciences*, 10(23), 2020.
- [15] Maryam Hashemi, Alireza Mirrashid, and Aliasghar Beheshti Shirazi. Driver safety development real time driver drowsiness detection system based on convolutional neural network. *arXiv preprint arXiv:2001.05137*, 2020.
- [16] Mariella Dreissig, Mohamed Hedi Baccour, Tim Schaeck, and Enkelejda Kasneci. Driver drowsiness classification based on eye blink and head movement features using the k-nn algorithm. *arXiv preprint arXiv:2009.13276*, 2020.
- [17] M. Deng and Y. Wu. A real-time driver drowsiness detection system based on facial features. *Journal of X*, xx:xx–xx, 2019.
- [18] Elena Magán, M. Paz Sesmero, Juan Manuel Alonso-Weber, and Araceli Sanchis. Driver drowsiness detection by applying deep learning techniques to sequences of images. *Applied Sciences*, 12(3), 2022.

- [19] Norah N. Alajlan and Dina M. Ibrahim. Ddd tinymml: A tinymml-based driver drowsiness detection model using deep learning. *Sensors*, 23(12), 2023.
- [20] Athanasios Voulodimos, Nikolaos Doulamis, Anastasios Doulamis, and Eftychios Protopapadakis. Deep learning for computer vision: A brief review. *Computational intelligence and neuroscience*, 2018(1):7068349, 2018.
- [21] Njayou Youssouf. Traffic sign classification using cnn and detection using faster-rcnn and yolov4. *Heliyon*, 8(12):e11792, 2022.
- [22] Tengwen Zhang, Xin Liu, Mingzhi Shao, Yuhan Sun, and Qingfa Zhang. Ship ranging method in lake areas based on binocular vision. *Sensors*, 25(20), 2025.
- [23] Christine Dewi, Rung-Ching Chen, Chun-Wei Chang, Shih-Hung Wu, Xiaoyi Jiang, and Hui Yu. Eye aspect ratio for real-time drowsiness detection to improve driver safety. *Electronics*, 11(19), 2022.
- [24] Ruben Florez, Facundo Palomino-Quispe, Ana Beatriz Alvarez, Roger Jesus Coaquira-Castillo, and Julio Cesar Herrera-Levano. A real-time embedded system for driver drowsiness detection based on visual analysis of the eyes and mouth using convolutional neural network and mouth aspect ratio. *Sensors*, 24(19), 2024.
- [25] Khaled Younes, Omar Mouhtady, Hamdi Chaouk, Emil Obeid, Rabih Roufayel, Ahmad Moghrabi, and Nimer Murshid. The application of principal component analysis (pca) for the optimization of the conditions of fabrication of electrospun nanofibrous membrane for desalination and ion removal. *Membranes*, 11(12), 2021.
- [26] Lorenzo Vergni and Francesca Todisco. A random forest machine learning approach for the identification and quantification of erosive events. *Water*, 15(12), 2023.
- [27] Xuejie Wang, Ziqiu Wang, Wen Du, Xiaobo Ma, Jun Ma, Zijin Chen, Chenni Gao, and Xiaonong Chen. Predictive value of tyg and tyg-bmi indices for non-alcoholic fatty liver disease in high-altitude regions of china: A cross-sectional study. *Journal of Clinical Medicine*, 13(23), 2024.
- [28] Fatma Ben Salem, Motab Turki Almousa, and Nabil Derbel. Direct torque control with space vector modulation (dtc-svm) with adaptive fractional-order sliding mode: A path towards improved electric vehicle propulsion. *World Electric Vehicle Journal*, 15(12), 2024.

- [29] Ulises Martin Casado, Facundo Ignacio Altuna, and Luis Alejandro Miccio. Towards sustainable material design: A comparative analysis of latent space representations in ai models. *Sustainability*, 16(23), 2024.
- [30] Nermeen Gamal Rezk, Samah Alshathri, Amged Sayed, and Ezz El-Din Hemdan. Ewais: An ensemble learning and explainable ai approach for water quality classification toward iot-enabled systems. *Processes*, 12(12), 2024.
- [31] René Meléndez, Michal Ptaszynski, and Fumito Masui. Comparative investigation of traditional machine-learning models and transformer models for phishing email detection. *Electronics*, 13(24), 2024.
- [32] Fazliddin Makhmudov, Alpamis Kutlimuratov, and Young-Im Cho. Hybrid lstm–attention and cnn model for enhanced speech emotion recognition. *Applied Sciences*, 14(23), 2024.
- [33] Henggang Zhang, Chenhui Zhu, Tianyu Jiao, Kaiyue Luo, Xu Ma, and Mingyu Wang. Analysis of the trends and driving factors of cultivated land utilization efficiency in henan province from 2000 to 2020. *Land*, 13(12), 2024.
- [34] Jiahui Huang and Chenglong Zhang. Daily tourism demand forecasting with the itransformer model. *Sustainability*, 16(23), 2024.
- [35] Henriques Zacarias, João Alexandre Lôbo Marques, Virginie Felizardo, Mehran Pourvahab, and Nuno M. Garcia. Ecg forecasting system based on long short-term memory. *Bioengineering*, 11(1), 2024.