YÜZ TANIMA SİSTEMLERİNDE DERİN ÖĞRENME TABANLI BİYOMETRİK SALDIRI TESPİTİ

YÜKSEK LİSANS TEZİ

Sena ÖZKARA

Danışman

Dr. Öğr. Üyesi Nevzat OLGUN

BİLGİSAYAR ANABİLİM DALI

Ocak 2026

AFYON KOCATEPE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

YÜZ TANIMA SİSTEMLERİNDE DERİN ÖĞRENME TABANLI BİYOMETRİK SALDIRI TESPİTİ

Sena ÖZKARA

Danışman
Dr. Öğr. Üyesi Nevzat OLGUN

BİLGİSAYAR ANABİLİM DALI

Ocak 2026

TEZ ONAY SAYFASI

Sena ÖZKARA tarafından hazırlanan "Yüz Tanıma Sistemlerinde Derin Öğrenme Tabanlı Biyometrik Saldırı Tespiti" adlı tez çalışması lisansüstü eğitim ve öğretim yönetmeliğinin ilgili maddeleri uyarınca 12 / 01 / 2026 tarihinde aşağıdaki jüri tarafından **oy birliği** ile Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsü **Bilgisayar Anabilim Dalı'nda YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman **:** Dr. Öğr. Üyesi Nevzat OLGUN

Başkan **:** Dr. Öğr. Üyesi Emrah ÖZKAYNAK
Karabük Üniversitesi ..... İmza .....
Bilgisayar ve Bilişim Bilimleri Fakültesi

Üye **:** Dr. Öğr. Üyesi İnayet Hakkı ÇİZMECİ
Afyon Kocatepe Üniversitesi ..... İmza .....
Mühendislik Fakültesi

Üye **:** Dr. Öğr. Üyesi Nevzat OLGUN
Afyon Kocatepe Üniversitesi ..... İmza .....
Mühendislik Fakültesi

Afyon Kocatepe Üniversitesi
Fen Bilimleri Enstitüsü Yönetim Kurulu'nun
..... /..... /..... tarih ve
……. sayılı kararıyla onaylanmıştır.

…….
Prof. Dr. Bekir YALÇIN
Enstitü Müdürü

BİLİMSEL ETİK BİLDİRİM SAYFASI

Afyon Kocatepe Üniversitesi

Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmasında;

- Tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- Görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- Başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- Atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- Kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- Ve bu tezin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

12 / 01 / 2026

Sena ÖZKARA

ÖZET

Yüksek Lisans Tezi

YÜZ TANIMA SİSTEMLERİNDE DERİN ÖĞRENME TABANLI BİYOMETRİK SALDIRI TESPİTİ

Sena ÖZKARA
Afyon Kocatepe Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Anabilim Dalı
Danışman: Dr. Öğr. Üyesi Nevzat OLGUN

Yüz tanıma sistemlerinin yaygınlaşması, fotoğraf, video yeniden oynatma ve 3D maske temelli sunum saldırılarını önemli bir güvenlik tehdidi haline getirmiştir. Bu tez çalışmasında söz konusu saldırıları donanımdan bağımsız ve gerçek zamanlı tespit edebilmek amacıyla uzamsal ve zamansal öznitelikleri birleştiren çift akışlı hibrit bir derin öğrenme mimarisi önerilmiştir. Görsel akışta transfer öğrenme ile başlatılan MobileNetV2 omurgası kullanılırken, fizyolojik canlılık belirtilerini analiz etmek için rPPG tabanlı bir LSTM ağı sisteme entegre edilmiştir. Çalışmanın temel katkısı, görsel akışın tam yüz görüntüsünü kullanması ve fizyolojik akışın anatomik segmentasyon tabanlı seçici ROI ile temiz deri bölgelerine odaklanmasıdır. Model Replay-Mobile, 3DMAD, PURE ve UBFC-RPPG veri setlerini içeren heterojen bir veri havuzunda kişi bağımsız 5-katlı çapraz doğrulama ile değerlendirilmiştir. Deneysel sonuçlar önerilen yöntemin %4,16 ACER ve %99,36 AUC değerleri ile literatürdeki karmaşık modellerle rekabet edebildiğini göstermektedir. Bulgular, görsel ve fizyolojik modalitelerin birleşiminin özellikle zor saldırı senaryolarında sistemin doğruluğunu ve kararlılığını artırdığını ortaya koymaktadır.

2026, xiii + 145 Sayfa

Anahtar Kelimeler: rPPG, Canlılık analizi, Biyometrik güvenlik, Uzaktan fotopletismografi.

ABSTRACT

M.Sc. Thesis

DEEP LEARNING-BASED BIOMETRIC ATTACK DETECTION IN FACE RECOGNITION SYSTEMS

Sena ÖZKARA

Afyon Kocatepe University

Graduate School of Natural and Applied Sciences

Department of Computer

Supervisor: Asst. Prof. Nevzat OLGUN

The widespread adoption of facial recognition systems has made presentation attacks based on photographs, video replays, and 3D masks a significant security threat. In this thesis, a dual-stream hybrid deep learning architecture that combines spatial and temporal features is proposed to detect such attacks in a hardware-independent and real-time manner. While the visual stream utilizes a MobileNetV2 backbone initiated with transfer learning, an rPPG-based LSTM network is integrated into the system to analyze physiological signs of life. The primary contribution of the study is that the visual stream utilizes the full face image, while the physiological stream focuses on clean skin regions through anatomical segmentation-based selective ROI. The model was evaluated using person-independent 5-fold cross-validation on a heterogeneous data pool including the Replay-Mobile, 3DMAD, PURE, and UBFC-RPPG datasets. Experimental results demonstrate that the proposed method competes with complex models in the literature, achieving an ACER of 4.16% and an AUC of 99.36%. The findings reveal that the combination of visual and physiological modalities increases the accuracy and stability of the system, particularly in challenging attack scenarios.

2026, xiii + 145 pages

Keywords: rPPG, Liveness analysis, Biometric security, Remote photoplethysmography.

TEŞEKKÜR

Bu araştırmanın konusu, deneysel çalışmaların yönlendirilmesi, sonuçların değerlendirilmesi ve yazımı aşamasında yapmış olduğu katkılarından dolayı tez danışmanım Sayın Dr. Öğr. Üyesi Nevzat Olgun'a teşekkür ederim.

Bu araştırma boyunca maddi ve manevi desteklerinden dolayı aileme teşekkür ederim.

Sena ÖZKARA

Afyonkarahisar 2026

İÇİNDEKİLER DİZİNİ

Sayfa

SİMGELER ve KISALTMALAR DİZİNİ

Simgeler

α	İki sinyal arasındaki ölçekleme faktörü
b	Bias (sapma) terimi
β	Kaydırma parametresi
C	Kanal sayısı
C_{in}	Giriş kanal sayısı
C_{out}	Çıkış kanal sayısı
C_t	Hücre durum vektörü
dk	Anahtar matrisinin boyutu
ε	Sayısal kararlılık sabiti (Epsilon)
ft	Unutma kapısı aktivasyon değeri
γ	Ölçeklendirme parametresi
H	Girdinin uzamsal yüksekliği
h_t	Gizli durum vektörü
i_t	Giriş kapısı aktivasyon değeri
K	Anahtar matrisi
k	Filtre boyutu
μ	Ortalama
o_t	Çıkış kapısı aktivasyon değeri
P	Projeksiyon katmanı
Q	Sorgu matrisi
s	Kayma miktarı
σ	Standart sapma veya Sigmoid fonksiyonu
$\sum$	Toplam sembolü
t	Genişleme faktörü veya zaman adımı
τ	Yumuşatma parametresi
V	Değer matrisi
W	Ağırlık matrisi
x_t	Giriş vektörü
z	Doğrusal çıktı değeri

Kısaltmalar

2B	İki boyutlu
3D	Three-Dimensional (üç boyutlu)
3D-CNN	Üç boyutlu evrişimsel sinir ağları
3DMAD	3D Mask Attack Database (3D maske saldırı veri tabanı)
ACER	Average classification error rate (ortalama sınıflandırma hata oranı)
APCER	Attack presentation classification error rate (saldırı sunumu sınıflandırma hata oranı)
AUC	Area under the curve (eğri altında kalan alan)
BPCER	Bona fide presentation classification error rate (gerçek sunum sınıflandırma hata oranı)
BPM	Beats per minute (dakika başına atım sayısı)
BSS	Blind source separation (kör kaynak ayrıştırma)
BVP	Blood volume pulse (kan hacmi nabzı)
CHROM	Chrominance-based method (krominans tabanlı yöntem)
CNN	Convolutional neural networks (evrişimsel sinir ağları)
DN	Doğru negatif
DP	Doğru pozitif
EER	Equal error rate (eşit hata oranı)
FFT	Fast fourier transform (hızlı fourier dönüşümü)
FNR	False negative rate (yanlış negatif oranı)
FPR	False positive rate (yanlış pozitif oranı)
FPS	Frames per second (saniye başına kare sayısı)
FRR	False rejection rate (yanlış reddetme oranı)
GAN	Generative adversarial networks (çekişmeli üretici ağlar)
GPU	Graphics processing unit (grafik işlem birimi)
HOG	Histogram of oriented gradients (yönlendirilmiş gradyanların histogramı)
ICA	Independent component analysis (bağımsız bileşen analizi)
IEC	International Electrotechnical Commission (uluslararası elektroteknik komisyonu)

ISO	International Organization for Standardization (uluslararası standardizasyon örgütü)
LBP	Local binary patterns (yerel ikili örüntüler)
LOOCV	Leave-One-Out cross-validation (birini dışarıda bırakarak çapraz doğrulama)
LFW	Labeled faces in the wild (doğal ortamda etiketli yüzler)
LSTM	Long short-term memory (uzun kısa-süreli bellek)
MAE	Mean absolute error (ortalama mutlak hata)
MTCNN	Multi-task cascaded convolutional networks (çok görevli basamaklı evrişimsel ağlar)
PAD	Presentation attack detection (sunum saldırısı tespiti)
PAI	Presentation attack ınstrument (sunum saldırı aracı)
PAS	Presentation attacks (sunum saldırıları)
POS	Plane-Orthogonal-to-Skin (cilt yüzeyine dik düzlem)
PPG	Photoplethysmography (fotopletismografi)
PURE	Pulse Rate Detection Dataset (nabız hızı tespiti veri seti)
RGB	Red-Green-Blue (kırmızı-yeşil-mavi)
RNN	Recurrent neural networks (tekrarlayan sinir ağları)
ROC	Receiver operating characteristic (alıcı çalışma karakteristiği)
ROI	Region of ınterest (ilgi alanı)
rPPG	Remote photoplethysmography (uzaktan fotopletismografi)
SNR	Signal-to-Noise Ratio (sinyal-gürültü oranı)
TNR	True negative rate (doğru negatif oranı)
TPR	True positive rate (doğru pozitif oranı)
UBFC-RPPG	stands for Univ. Bourgogne Franche-Comté Remote PhotoPlethysmoGraphy veri seti
ViT	Vision transformer (görüntü dönüştürücü)
YN	Yanlış negatif
YP	Yanlış pozitif

ŞEKİLLER DİZİNİ

Sayfa

ÇİZELGELER DİZİNİ

Sayfa

1. GİRİŞ

Biyometrik doğrulama sistemleri bireylerin kimliklerinin tespiti amacıyla özgün anatomik özelliklerin sayısallaştırılarak analiz edildiği otomatik tanıma teknolojileridir. Bu sistemlerde en yaygın kullanılan biyometrik öznitelikler arasında parmak izi, avuç içi geometrisi, retina ve iris desenleri, yüz morfolojisi ve ses spektrumu yer almaktadır. Biyometrik sistemler yüksek güvenlik gerektiren erişim kontrol uygulamalarından kullanıcı dostu kimlik doğrulama çözümlerine kadar geniş bir yelpazede kullanılmaktadır (Yıldırım 2024). Bu sistemler özellikle havalimanları, askeri tesisler, devlet kurumları gibi yüksek güvenlikli alanlarda yaygın olarak uygulanmaktadır (Yün 2023). Mobil cihazlar (Arslan ve Sağıroğlu 2016), e-ticaret platformları, dijital bankacılık sistemleri ve akıllı yaşam teknolojileri (Yang vd. 2021) gibi bireysel uygulama alanlarında da giderek daha fazla tercih edilmektedir. Bu sistemlerin temel avantajı kullanıcıların şifre, kart vb. taşıması ya da hatırlaması gereken bir bilgiye ihtiyaç duymadan doğrulama yapabilmesidir. Ancak her biyometrik sistem doğruluk oranı, işlem süresi, kullanıcı kabul edilebilirliği, çevresel koşullara duyarlılık ve sahtecilik direnci gibi performans kriterleri açısından farklı avantaj ve sınırlılıklar taşımaktadır (Yalçın ve Gürbüz 2015).

Yüz tanıma sistemleri bireylerin yüz anatomisine ait karakteristik özelliklerin analiz edilmesiyle kimlik doğrulama işlemini gerçekleştiren, temassız ve kullanıcı dostu biyometrik teknolojilerdir (Adjabi vd. 2020). Parmak izi veya retina tarayıcıları gibi özel donanımlar gerektirmeden yalnızca standart bir kamera aracılığıyla çalışabilmeleri bu sistemlerin yaygınlığını ve kullanım kolaylığını artıran temel etkenlerden biridir. Günümüzde yüz tanıma sistemleri güvenlik uygulamalarının yanı sıra mobil cihaz kilidi açma, kullanıcı takibi, dijital bankacılık ve eğlence teknolojileri gibi birçok farklı alanda aktif olarak kullanılmaktadır. Parmak izinden sonra en yaygın ikinci biyometrik kimlik doğrulama yöntemi haline gelmiş olan bu sistemler, fiziksel temasa ihtiyaç duymadan yüksek doğrulukla çalışabildikleri için özellikle pandemi sonrası dönemde daha da fazla ilgi görmeye başlamıştır (Gomez-Barrero vd. 2022).

Yüz tanıma sistemlerinin yaygınlaşması bu sistemlere yönelik saldırı türlerinin de çeşitlenmesine yol açmıştır (Güneş 2025). Sosyal medya platformlarında artan yüz

görüntüsü paylaşımı (Evlioğlu Gezer 2025) kötü niyetli kişilerin fotoğraf, video veya üç boyutlu maske gibi sahte materyaller kullanarak sistemleri aldatma girişimlerini kolaylaştırmıştır (Liu vd. 2021). Bu tür saldırılar yüz sahteciliği olarak tanımlanmakta ve biyometrik doğrulama sistemlerinin güvenlik açıklarını istismar etmektedir. Basit düzeydeki saldırılar basılı bir fotoğraf veya ekran görüntüsü ile gerçekleştirilebilirken ileri düzey saldırılar arasında yüksek kaliteli video oynatmaları, 3D (Three-Dimensional) maske kullanımı (Menakadevi vd. 2025) hatta derin öğrenme tabanlı sahte görüntü üretim teknikleri yer almaktadır (Patel vd. 2023). Bu bağlamda yüz tanıma sistemlerinin sahteciliğe karşı dirençli hale getirilmesi biyometrik güvenlik alanında güncel ve kritik bir araştırma problemi olarak önemini korumaktadır. Bu durum geleneksel yüz tanıma algoritmalarının yetersizliğini ortaya koymakta ve canlılık tespiti gibi ek güvenlik katmanlarının gerekliliğini açıkça göstermektedir (Shinde vd. 2025).

Canlılık tespiti biyometrik sistemlerin sahte yüz materyalleri ile aldatılmasını önlemek amacıyla geliştirilen yöntemlerin bütününü ifade etmektedir. Bu yöntemler bir yüzün gerçek bir insana mı yoksa bir taklit materyaline mi ait olduğunu analiz etmeye yönelik çeşitli yazılım ve donanım tabanlı yaklaşımları içermektedir (Pooshideh vd. 2024). Yazılım tabanlı yöntemler görüntü tabanlı sinyallerin analizi, mikro hareketleri, göz kırpma tespiti, yüz kası hareketleri veya renk değişimleri gibi fiziksel ipuçlarından yararlanmaktadır. Donanım tabanlı sistemler derinlik sensörleri, kızılötesi kameralar veya 3D yapılandırma gibi daha karmaşık teknolojilerden faydalanmaktadır (Sharma ve Selwal 2023b). Ancak son yıllarda geliştirilen deepfake teknolojileri ve 3D yazıcılarla üretilen sahte yüz yapıları gibi ileri düzey saldırılar karşısında geleneksel canlılık tespit yöntemleri yeterli güvenliği sağlayamamaktadır (Masood vd. 2023). Bu doğrultuda temassız, düşük maliyetli ve sahteciliğe karşı yüksek duyarlılık sunan yeni nesil canlılık tespiti algoritmalarının geliştirilmesi bir gereklilik hâline gelmiştir. Bu bağlamda uzaktan fotopletismografi (Remote Photoplethysmography, rPPG) tabanlı yöntemler canlılık tespitinde umut vadeden yöntemler arasında öne çıkmaktadır (Xiao vd. 2024).

Temassız bir fizyolojik ölçüm teknolojisi olan uzaktan fotopletismografi insan vücudunun özellikle yüz bölgesinden elde edilen video verileri aracılığıyla kan hacmindeki periyodik değişimleri analiz ederek kalp atış hızını temassız biçimde tahmin

edilmesine olanak tanıyan görüntü tabanlı bir ölçüm tekniğidir (Sharma ve Selwal 2023b). Bu yöntem geleneksel fotopletismografi (Photoplethysmography, PPG) tekniklerinden farklı olarak ciltle fiziksel temas gerektirmemekte ve yalnızca bir RGB kamera aracılığıyla uygulanabilmektedir. Bu yönüyle kullanıcı dostu, düşük maliyetli ve geniş bir uygulama alanına sahip bir çözüm sunmaktadır (Lee vd. 2023b, Chen vd. 2024). Ancak rPPG sinyallerinin doğruluğu aydınlatma koşulları, yüz hareketleri, cilt tonu ve kamera kalitesi gibi faktörlerden etkilenebildiği için literatürde bu zorluklara yönelik çeşitli algoritmik çözümler önerilmiştir (Lee vd. 2023a).

rPPG teknolojisinin uygulanabilirliği konusundaki öncü çalışmalar ortam ışığı kullanılarak kan hacmi değişimlerinin standart kameralarla tespit edilebileceğini ortaya koymuştur. Bu bağlamda farklı yüz bölgelerinden alınan sinyaller karşılaştırılarak en yüksek doğruluğa sahip ilgi alanları (Region of Interest, ROI) belirlenmiştir. Ancak bu yöntem temel düzeyde olduğundan dolayı aydınlatma ve hareket değişimlerine duyarlılık göstermektedir (Verkruysse vd. 2008). Bu gelişmenin ardından araştırmacılar sistemin kararlılığını artırmaya odaklanmıştır. Yüz tespiti sonrası belirlenen ROI'lerden elde edilen RGB kanal sinyalleri bağımsız bileşen analizi (Independent Component Analysis, ICA) ile ayrıştırılarak kalp atış hızı hesaplanmış ve frekans alanı analizi hızlı fourier dönüşümü (Fast Fourier Transform, FFT) ile gerçekleştirilmiştir. Bu yöntem, görüntüden sinyal çıkarımı açısından önemli bir adım olmakla birlikte farklı kamera yapılandırmalarında değişken bir başarım sergilemiştir (Poh vd. 2010).

Daha sonraki araştırmalarda eulerian video magnification yöntemi kullanılarak video kareleri arasındaki mikro düzeydeki renk değişimleri güçlendirilmiştir. Böylece çıplak gözle tespit edilemeyen fizyolojik değişiklikler görünür hale getirilmiştir. Bu yaklaşım yalnızca rPPG uygulamalarının yanı sıra çeşitli tıbbi ve biyometrik analizler için de potansiyel taşımaktadır (Wu vd. 2012).

Mevcut yöntemlerin eksikliklerini gidermek amacıyla rPPG sinyal tahminine yönelik dört aşamalı bir sistem önerilmiştir. Bu sistemdeki yüz tespiti, aydınlatma normalizasyonu, hareket telafisi ve geçici filtreleme adımları ile sistematik bir çözüm sunulmuştur. Bu yapı rPPG sinyal kalitesini artırmak için önemli bir çerçeve sağlamış olsa da her koşulda

evrensel başarıyı garanti edememektedir (Li vd. 2014). Takip eden yıllarda yüz üzerinde birden fazla ROI tanımlanarak her bölgeden elde edilen sinyaller kör kaynak ayrıştırma (Blind Source Separation, BSS) yöntemi ile analiz edilmiş ve ardından bu sinyallerin kalitesi değerlendirilerek birleştirilmiştir. Bu çok bölgeli yaklaşım sinyal kararlılığını artırmaktadır Ancak bu durum hesaplama maliyetini de yükseltmektedir (Lam ve Kuno 2015). Yüz sahteciliği tespiti amacı ile rPPG sinyallerinden faydalanılarak canlı ve cansız yüz ayrımı yapan bir sınıflandırıcı geliştirilmiştir. Üç farklı ROI'den elde edilen sinyallerin analiz edildiği bu çalışmada, sınıflandırma aşamasında yüksek doğruluk oranlarına ulaşılmıştır. Buna rağmen yöntemin 3 boyutlu maske saldırılarına karşı etkinliğinin test edilmemiş olması önemli bir kısıttır (Nowara vd. 2017). Benzer bir yaklaşımla yüz bölgesinden kalp atış sinyalinin kestirimi üzerine gerçekleştirilen bir diğer çalışmada ise lineer olmayan matris ayrıştırma yöntemi kullanılmıştır. Söz konusu yöntemin gürültüye karşı dayanıklılığı deneysel olarak ortaya konulmuştur. Elde edilen sonuçlar, modern rPPG sistemlerine kıyasla benzer ya da daha yüksek doğruluk oranları sunmaktadır (Demirezen 2022).

Son yıllarda rPPG teknolojisi temassız kalp atış hızı ölçümü ve yüz sahteciliği tespiti gibi biyometrik güvenlik uygulamalarında önemli ilerlemeler kaydetmiştir. Özellikle son birkaç yılda yapılan çalışmalar makine öğrenimi, derin öğrenme ve sinyal işleme tekniklerinin entegrasyonu ile rPPG sinyallerinin doğruluğunu ve güvenilirliğini artırmayı hedeflemiştir. Bu kapsamda Castellano Ontiveros ve arkadaşları tarafından yeni bir model önerilmiştir. Bu model, video tabanlı rPPG sinyallerini geleneksel PPG sinyalleriyle karşılaştırmaktadır. Değerlendirme sürecinde dinamik zaman eşleştirmesi ve korelasyon katsayıları gibi metrikler kullanılmıştır. Elde edilen bulgular geliştirilen modelin temas gerektiren yöntemlere yakın bir doğruluk performansı sunduğunu kanıtlamaktadır (Castellano Ontiveros vd. 2024).

Sinyal doğruluğunun yanı sıra sistemin gerçek zamanlı uygulamalardaki verimliliği de araştırmacıların odak noktası olmuştur. Bu doğrultuda Wang ve arkadaşları ME-rPPG adı verilen bellek verimli bir algoritma önermiştir. Böylece rPPG teknolojisinin gerçek zamanlı kullanılabilirliği artırılmıştır. Zamansal-mekânsal durum uzayı ayrışımı prensibini kullanan bu yöntem, kaynak kullanımında ciddi bir tasarruf sağlamaktadır.

Yöntem 3.6 MB gibi düşük bir bellek kullanımı ve 9.46 ms seviyesinde düşük gecikme sunmaktadır. Ayrıca yüksek doğrulukta sinyal elde edilmesini sağlayarak farklı veri kümeleri üzerinde yapılan deneysel değerlendirmelerde önerilen yöntemin önceki yaklaşımlara kıyasla %21,3 ile %60,2 arasında değişen performans artışları sunduğu raporlanmıştır (Wang vd. 2025) .

Gerçek dünya koşullarındaki zorlukları aşmak ve özellikle aşırı aydınlatma durumlarında kararlılığı sağlamak adına uçtan uca bir video dönüştürücü modeli geliştirilmiştir. Küresel girişim paylaşımı, arka plan referansı ve kendiliğinden denetimli ayrıştırma gibi tekniklerin kullanıldığı bu model ile dış etkenlerin etkisi azaltılmıştır. Yöntem farklı senaryolarda rekabetçi bir performans sergilenmiştir (Shao vd. 2025). Benzer bir yaklaşımla sinyallerin gürültüden arındırılması için CodePhys adı verilen yeni bir yöntem sunulmuştur. Bu yaklaşımda gürültülü rPPG özellikleri önceden oluşturulmuş bir kod kitabındaki gürültüsüz PPG özellikleri ile eşleştirilerek yüksek kaliteli sinyaller üretilmiştir. Ayrıca mekânsal dikkat mekanizması ve distilasyon kaybı kullanılarak fizyolojik açıdan aktif bölgeler vurgulanmış ve ölçüm doğruluğu iyileştirilmiştir (Chu vd. 2025). Sonuç olarak son yıllarda geliştirilen derin öğrenme tabanlı rPPG yaklaşımları hem sinyal kalitesini iyileştirmek hem de gerçek zamanlı işleme olanaklarını artırmak açısından önemli ilerlemeler kaydetmiştir. Ancak hâlen sahte yüz görüntülerine karşı yüksek doğrulukta ve düşük gecikmeli canlılık tespiti yapabilen aynı zamanda farklı veri kümelerine genellenebilir modellerin geliştirilmesi açık bir araştırma problemi olarak varlığını sürdürmektedir. Bu doğrultuda sinyal işleme teknikleri ile derin öğrenme mimarilerinin birleştirildiği rPPG sistemlerinin özellikle düşük kaliteli ve hareketli ortamlarda bile güvenilir çalışması ve sahteciliği yüksek başarı ile ayırt edebilmesi kritik bir gerekliliktir.

Literatürdeki yüz sahteciliği tespit yöntemleri incelendiğinde bazı kritik sınırlılıklar göze çarpmaktadır. Bu kısıtların başında hesaplama verimliliği gelmektedir. Mevcut algoritmaların işlem yükü, mobil platformlarda gerçek zamanlı uygulanabilirlik kapasitesini ciddi oranda düşürmektedir. Ayrıca bu yöntemler karmaşık saldırı vektörlerine karşı bütünleşik direnç oluşturmakta zorlanmaktadır. Sistemlerin fotoğraf, tekrar oynatma videoları ve 3 boyutlu maske gibi farklı saldırı türlerine karşı gösterdiği

dayanıklılık istenilen seviyeye ulaşmamıştır. Bu çalışmada doku ve hareket analizi derin öğrenme tabanlı temsil öğrenimi ve uzaktan nabız sinyali çıkarımı ile entegre edilmiştir. Bu entegrasyon sonucunda çift akışlı hibrit bir mimari önerilmiştir. Geliştirilen bu yapı sayesinde yüz sahteciliği tespitinde daha yüksek doğruluk ve dayanıklılık sağlanması hedeflenmiştir. Önerilen bu yöntemle özellikle sunum saldırılarının yaygın türleri olan fotoğraf gösterimi, video tekrar oynatma saldırıları, üç boyutlu maske saldırılarının etkin bir şekilde ayırt edilmesi ve yüzlerin canlı ya da sahte olarak sınıflandırılması amaçlanmıştır. Ayrıca model, mevcut birçok yüz sahteciliği tespiti yönteminin aksine hesaplama açısından verimli olacak şekilde tasarlanmıştır. Böylece mobil cihazlar gibi kaynak kısıtlı ortamlarda gerçek zamanlı çalışabilirlik mümkün kılınmıştır. Bu çalışma ile yüz tanıma sistemlerinin güvenilirliği artırılmakta ve özellikle mobil uygulamalarda karşılaşılan sahtecilik sorununa pratik bir çözüm sunulmaktadır.

Tez kapsamında önerilen yöntemin geliştirilmesi ve performansının değerlendirilmesi amacıyla yüz sahteciliği tespiti ve rPPG alanlarında yaygın olarak referans alınan 3DMAD (3D Mask Attack Database) (Erdogmus ve Marcel 2013), Replay-Mobile (Costa-Pazo vd. 2016), PURE (Pulse Rate Detection Dataset) (Stricker vd. 2014) ve UBFC-RPPG (Bobbia vd. 2019) halka açık veri setlerinden yararlanılmıştır. Modelin başarımı söz konusu veri setleri üzerinde kişiden bağımsız katmanlı k-katlı çapraz doğrulama (Subject-Independent Stratified K-Fold Cross-Validation) yöntemi kullanılarak değerlendirilmiştir. Farklı saldırı senaryolarını, aydınlatma koşullarını ve görüntü kalitelerini barındıran bu veri setleri önerilen modelin dayanıklılığını kapsamlı bir şekilde analiz edilmesine imkân sağlamaktadır.

Bu tez toplam sekiz bölümden oluşmaktadır. Giriş bölümünü takip eden ikinci bölümde biyometrik sistemler, yüz tanıma teknolojileri, güvenlik açıkları (sunum saldırıları) ve literatürdeki mevcut yüz sahteciliği tespiti yöntemleri ile rPPG sistemlerinin karşılaştığı zorluklar kapsamlı bir şekilde incelenmiştir. Üçüncü bölümde çalışmanın teknolojik altyapısını oluşturan derin öğrenme mimarileri hakkında teorik bilgiler sunulmuştur. Dördüncü bölümde tezin fizyolojik temelini oluşturan uzaktan fotopletismografi teknolojisinin çalışma prensipleri, sinyal çıkarım algoritmaları ve ROI seçim stratejileri detaylandırılmıştır. Beşinci bölümde önerilen modelin güvenilirliğini ölçmek için

kullanılan doğrulama stratejileri ve performans değerlendirme metrikleri açıklanmıştır. Altıncı bölümde kullanılan veri setleri, ön işleme adımları ve geliştirilen hibrit yüz sahteciliği tespiti modelinin mimari detayları ile algoritmik işleyişi sunulmuştur. Yedinci bölümde farklı rPPG algoritmaları ve ROI stratejileri ile gerçekleştirilen deneysel çalışmaların bulguları, performans analizleri ve literatür karşılaştırmaları yer almaktadır. Sekizinci ve son bölümde ise çalışmanın genel bir değerlendirmesi yapılarak elde edilen sonuçlar özetlenmiş ve gelecek çalışmalara yönelik öneriler sunulmuştur.

2. LİTERATÜR ÖZETİ

Literatürde yüz tanıma sistemlerine yönelik biyometrik saldırı tespiti temel olarak görsel doku analizi, uzamsal-zamansal yüz hareketleri ve fizyolojik sinyal çıkarımına dayalı yöntemler etrafında şekillenmektedir. Derin öğrenme tabanlı yaklaşımlar evrişimsel sinir ağları (Convolutional Neural Networks, CNN) ve tekrarlayan sinir ağları (Recurrent Neural Networks, RNN) mimarileri aracılığıyla sahte materyallerin yüz dokusunda oluşturduğu anomalileri, doğal olmayan mikro ifadeleri ve süreksiz hareket dizilerini yüksek doğrulukla ayırt edebilmektedir. Bununla birlikte rPPG analizinin de dahil olduğu fizyolojik tabanlı yöntemler gerçek yüzlerde doğal olarak bulunan kalp atımına bağlı renk değişimlerinin maske, ekran veya deepfake materyallerinde üretilememesinden yararlanarak güçlü bir canlılık göstergesi sunmaktadır. Güncel araştırmalar, ROI seçiminin fizyolojik sinyal kalitesi üzerindeki belirleyici etkisini vurgulamaktadır. Özellikle alın ve yanak gibi anatomik olarak stabil bölgeler, saldırı tespit performansını anlamlı biçimde artırmaktadır. Genel literatür taraması, tekil yöntemler yerine bütünleşik yapıların önemini ortaya koymaktadır. Çoklu ipucu ve çoklu akış yapıları hem klasik sunum saldırılarına hem de gelişmiş sahte yüz üretim tekniklerine karşı en dayanıklı çözüm olarak öne çıkmaktadır. Bu yapılar aynı zamanda en yüksek genellenebilirliği sunmaktadır.

2.1 Biyometrik Sistemlere Genel Bakış

Biyometrik sistemler bireylerin kimliklerini doğrulamak veya tanımak amacıyla fiziksel veya davranışsal özelliklerini kullanan otomatik sistemlerdir. Bu yöntemlerin kökeni binlerce yıl öncesine dayansa da modern biyometrik sistemler bilgisayar kontrollü ve otomatik işleme yeteneklerine sahiptir. Temel olarak bir kişinin parmak izi, yüzü, iris deseni, sesi, yürüyüş biçimi gibi ölçülebilir ve ayırt edici özelliklerini kullanarak kimlik tespiti yapılabilmektedir. Fiziksel biyometrik sistemler bireyin doğasında bulunan parmak izi, yüz, iris gibi sabit özelliklere dayanırken davranışsal biyometrik sistemler ise imza, yazı dinamiği vb. belirli bir zamanda sergilenen davranışlara odaklanmaktadır (Akgül 2015).

Biyometrik sistemlerin temel işleyişi bir veri tabanında depolanan kayıtlı biyometrik verinin kimlik doğrulama veya tanıma anında yakalanan yeni biyometrik veri ile karşılaştırılmasına dayanmaktadır. Eğer karşılaştırma sonucunda yeterli eşleşme sağlanırsa sistem tarafından kullanıcıya erişim izni verilmekte veya kimliği belirlenmektedir (Arslan ve Sağıroğlu 2016). Şifre ve kart gibi geleneksel kimlik doğrulama yöntemlerinin aksine biyometrik sistemler, biyometrik verinin unutulamaz veya kolayca devredilemez olması gibi önemli avantajlar sunmaktadır. Bu özellikleri sayesinde biyometrik yöntemler güvenlik seviyelerini artırmada en etkili çözümlerden biri olarak kabul edilmektedir. Biyometrik sistemler özellikle suçların önlenmesi ve dolandırıcılıkla mücadele gibi alanlarda bireylerin doğru ve kesin bir şekilde tanımlanmasını sağlayarak önemli faydalar sunmaktadır (Yün 2023). Bununla birlikte biyometrik verilerin toplanması ve işlenmesi kişisel gizlilik ve veri güvenliği açısından ciddi riskleri de beraberinde getirmektedir. Biyometrik veriler bireylere ait son derece hassas bilgiler içerebilmekte ve yetkisiz erişim veya kötüye kullanım durumunda geri döndürülemez kimlik sorunlarına neden olabilmektedir. Bu nedenle biyometrik sistemlerin kullanımında etik ilkelere uyulması, kişilerden açık ve bilgilendirilmiş rıza onamının alınması ve verilerin güvenli bir şekilde işlenmesi büyük önem taşımaktadır (Çiçek 2024). Biyometrik verilerin korunmasına yönelik Kişisel Verilerin Korunması Kanunu (KVKK) gibi ulusal düzenlemeler ile uluslararası veri koruma mevzuatları bu alandaki hukuki ve etik hassasiyeti açık bir şekilde yansıtmaktadır. Biyometrik sistemlerin güvenliği tanıma doğruluğunun yanında veri yönetimi, gizlilik ve potansiyel saldırılara karşı direnç ile de yakından ilişkilidir.

2.2 Yüz Tanıma Sistemleri

Yüz tanıma sistemleri dijital görüntüler veya video kareleri üzerinden bireylerin kimliğini doğrulamak veya tespit etmek amacıyla kullanılan biyometrik teknolojilerdir. Bu sistemler insan yüzünün matematiksel bir haritasını çıkararak veri tabanındaki kayıtlarla karşılaştırma prensibine dayanmaktadır. Günümüzde güvenlikten ticari uygulamalara kadar geniş bir yelpazede kullanılan bu teknoloji tarihsel süreç içerisinde basit geometrik ölçümlerden karmaşık derin öğrenme mimarilerine doğru evrilmiştir.

2.2.1 Yüz Tanıma Sistemlerinin Tarihsel Gelişimi

Yüz tanıma teknolojisinin kökeni 1950'li yıllara kadar uzanmaktadır. Ancak bu alandaki ilk otomatik sistemlerin temelleri 1970'li yıllarda atılmıştır. İlk çalışmalarda yüz tanıma işlemi yüzün belirli bölgeleri arasındaki mesafelerin ölçülmesine dayandırılmıştır (Taskiran vd. 2020). Zamanla gelişen teknolojik imkânlar bu sistemlerin daha karmaşık ve güvenilir hale gelmesini sağlamıştır.

Yüz tanıma teknolojilerinin tarihsel gelişimine bakıldığında 1964 yılında kişilerin göz ve ağız konumları gibi yüz hatlarını analiz eden yarı otomatik bir sistem geliştirildiği görülmektedir. Bu sistem yaklaşık yirmi farklı yüz özelliğini analiz edebiliyordu. Yaklaşık yirmi farklı yüz özelliğini analiz edebilen bu sisteme 1977 yılında dudak genişliği ve saç rengi gibi ek değişkenlerin katılmasıyla daha gelişmiş bir yapı kazandırılmıştır (Boyer ve Boyer 1991). 1988 yılında yapay zekânın yüz tanıma alanında kullanılmaya başlanması bu teknolojinin evriminde önemli bir dönüm noktası olmuştur. 1991'de Massachusetts Institute of Technology (MIT) bünyesinde "Eigenfaces" adı verilen yöntemle ilk başarılı otomatik yüz tanıma uygulaması literatüre sunulmuştur. Bu yöntem istatistiksel bir yaklaşım olan temel bileşen analizini (Principal Component Analysis, PCA) temel almaktadır (Turk ve Pentland 1991). 1993 yılında Amerika Birleşik Devletleri Savunma Bakanlığı'na bağlı Defense Advanced Research Projects Agency tarafından başlatılan Face Recognition Technology programı kapsamında oluşturulan geniş kapsamlı veri tabanı araştırmacıların farklı yüz tanıma algoritmalarını ortak bir standartta değerlendirmelerine olanak tanımıştır. Bu veri tabanı 1199 farklı kişiye ait farklı poz, ifade ve zamanlarda çekilmiş yaklaşık 14.126 yüz görüntüsünden oluşmaktadır (Phillips vd. 2000) .

2005 yılında düzenlenen Face Recognition Grand Challenge yarışması Amerika Birleşik Devletleri'nde kullanılan yüz tanıma teknolojilerini geliştirmek ve değerlendirmek amacıyla organize edilmiştir. Bu etkinlik yüz tanıma algoritmalarının doğruluk ve güvenilirlik düzeylerinin artırılması yönünde atılan önemli bir adım olmuştur. Ayrıca farklı ışık, poz ve çözünürlük koşulları altında çalışan algoritmaların kapsamlı biçimde karşılaştırılmasına olanak tanımıştır (Phillips vd. 2007).

2014 yılında Facebook AI Research ekibi tarafından geliştirilen DeepFace adlı sistem derin öğrenme algoritmaları kullanarak %97,35 oranında doğruluk ile insan düzeyine yakın bir başarı sergilemiştir. Yüzlerin 3 boyutlu olarak hizalanıp derin sinir ağlarıyla işlendiği bu model yüz tanımanın makineler tarafından da yüksek doğrulukla yapılabileceğini kanıtlamıştır (Taigman vd. 2014).

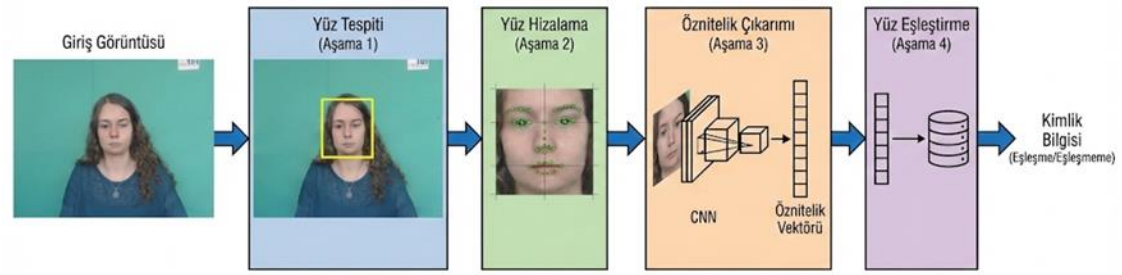
Bu gelişmeyi takiben 2015 yılında Google araştırmacıları tarafından "FaceNet" sistemi tanıtılmıştır. Önceki yöntemlerin aksine "Triplet Loss" adı verilen yeni bir kayıp fonksiyonu kullanan bu sistem yüz görüntülerini doğrudan 128 boyutlu bir Öklid uzayına haritalamayı başarmıştır. FaceNet, LFW (Labeled faces in the wild) veri seti üzerinde %99,63 gibi o dönem için rekor bir doğruluk oranına ulaşarak yüz eşleştirme ve doğrulama süreçlerinde yeni bir standart belirlemiştir (Schroff vd. 2015).

2019 yılına gelindiğinde yüz tanıma modellerinin ayırt edici gücünü artırmak amacıyla "ArcFace" (Additive Angular Margin Loss) yöntemi geliştirilmiştir. Bu çalışma derin öğrenme modellerinin eğitim sürecinde sınıflar arası ayrımı maksimize ederken sınıf içi çeşitliliği minimize eden geometrik bir yaklaşım sunmuştur. ArcFace günümüzde hala endüstriyel uygulamalarda ve akademik çalışmalarda en yaygın kullanılan referans modellerden biri olma özelliğini korumaktadır (Deng vd. 2019).

Son yıllarda ise bilgisayarlı görü alanında devrim yaratan transformer mimarileri yüz tanıma teknolojilerine entegre edilmeye başlanmıştır. 2021 ve sonrasında yapılan çalışmalarda evrişimsel sinir ağları yerine vision transformer (ViT) tabanlı modellerin özellikle büyük ölçekli veri setlerinde ve düşük çözünürlüklü görüntülerde daha yüksek genelleştirme yeteneğine sahip olduğu gösterilmiştir (Khan vd. 2023).

2.2.2 Yüz Tanıma Sistemlerinin Temel Bileşenleri ve Çalışma Prensibi

Modern bir yüz tanıma sistemi ham görüntü verisinden yüz tespiti, yüz hizalama, öznitelik çıkarımı ve yüz eşleştirme süreçlerinden oluşmaktadır. Yüz tanıma sistemlerinin genel akış şeması Şekil 2.1'de sunulmuştur.



Şekil 2.1 Yüz tanıma sisteminin genel çalışma prensibi ve temel bileşenleri

Sürecin ilk ve en kritik adımı karmaşık bir arka plana sahip görüntü içerisindeki yüz bölgesinin konumunun belirlenmesidir. Bu aşamada sistem görüntüdeki pikselleri tarayarak insan yüzü olup olmadığını etmekte ve yüzün bulunduğu bölgeyi ayıklamaktadır. Geçmişte bu işlem için Haar Cascade sınıflandırıcıları gibi yöntemler kullanılırken (Viola ve Jones 2004) günümüzde MTCNN (Multi-task Cascaded Convolutional Networks) gibi derin öğrenme tabanlı yaklaşımlar tercih edilmektedir (Yang vd. 2020). Bu yöntemler farklı ölçeklerdeki yüzleri, zorlu aydınlatma koşullarını ve kısmi kapanmaları daha yüksek başarıyla tespit edebilmektedir.

Tespit edilen yüz görüntüleri; kamera açısı, kişinin duruşu veya çevresel faktörlere bağlı olarak rotasyon, eğim ve yönelim farklılıkları içerebilmektedir. Doğru bir öznitelik çıkarımı yapabilmek için yüzün geometrik olarak standart bir forma getirilmesi gerekmektedir. Bu aşamada göz bebekleri, burun ucu ve ağız köşeleri gibi noktalar referans alınmaktadır. Tespit edilen bu noktalar doğrultusunda ham görüntü üzerinde döndürme, ölçeklendirme ve kırpma işlemleri uygulanmaktadır. Böylece tüm yüz görüntüleri kanonik bir hizaya getirilerek modelin poz değişimlerinden etkilenmesi engellenmektedir (Jin ve Tan 2017).

Öznitelik çıkarımı aşamasında ise hizalanmış yüz görüntüsü makine tarafından işlenebilir matematiksel verilere dönüştürülmektedir. Geleneksel yöntemlerde yerel ikili örüntüler (Local Binary Patterns, LBP) veya yönlendirilmiş gradyanların histogramı (Histogram of Oriented Gradients, HOG) gibi el yapımı öznitelikler kullanılırken (Xiang vd. 2018) modern sistemlerde ise CNN tabanlı modeller kullanılmaktadır. CNN modelleri yüzü 128 veya 512 boyutlu sayısal bir vektöre dönüştürmektedir. Bu vektör göz mesafesi, elmacık kemiği yapısı vb. kişinin yüzüne ait ayırt edici karakteristikleri temsil etmektedir. Ayrıca

bu vektörün aynı kişiye ait farklı fotoğraflarda benzer değerler üretmesi beklenmektedir (Ben Fredj vd. 2021).

Son aşamada elde edilen öznitelik vektörü veri tabanında kayıtlı olan vektörlerle karşılaştırılmaktadır. Bu karşılaştırma işlemi genellikle Öklid mesafesi (Euclidean Distance) veya Kosinüs Benzerliği (Cosine Similarity) gibi metrikler kullanılarak yapılmaktadır. Eğer iki vektör arasındaki benzerlik skoru belirlenen bir eşik değerin üzerindeyse sistem eşleşme kararı vermekte aksi takdirde ise eşleşmeme kararı vermektedir (Deng vd. 2019).

2.2.3 Yüz Tanıma Yöntemleri

Literatürde yüz tanıma için farklı yaklaşımlar mevcuttur. Geometrik tabanlı yöntemler yüzün şekli ve özellikleri arasındaki ilişkilere odaklanırken (Kortli vd. 2020) görünüm tabanlı yöntemler yüzün genel görünümünü analiz etmektedir (Siddiqi vd. 2022). Özellik tabanlı yöntemler burun, göz, ağız gibi yüzdeki belirgin özellikleri tespit edip bu özelliklerin arasındaki ilişkileri incelemektedir (Trigueros vd. 2018). Son olarak derin öğrenme tabanlı yöntemler ise büyük miktarda veri üzerinde eğitilmiş derin sinir ağlarını kullanarak yüzleri tanımaktadır. Derin öğrenme tabanlı yöntemler özellikle karmaşık senaryolarda ve değişken koşullar altında yüksek performans göstermesi nedeniyle günümüzde en çok tercih edilen yaklaşımdır (Guo ve Zhang 2019).

2.2.4 Yüz Tanıma Yöntemlerinin Yaygın Uygulama Alanları

Yüz tanıma teknolojisi temassız ve kullanıcı dostu yapısı sayesinde geniş bir kullanım alanına sahiptir. Kamu güvenliği, mobil cihaz kilit açma (Arslan ve Sağıroğlu 2016), bankacılık işlemleri (Yıldırım 2024), sınır kontrolleri (Yün 2023) ve sosyal medya etiketlemeleri (Taigman vd. 2014, Evlioğlu Gezer 2025) gibi birçok alanda yaygınlaştığı görülmektedir. Özellikle CNN tabanlı modeller yüz tanıma doğruluğunu önceki yöntemlere kıyasla %99'un üzerine çıkarmıştır (Tiraki vd. 2022). Ayrıca COVID-19 salgını sürecinde maske takma zorunluluğu yüzün sadece göz bölgesinden tanınmasını gerektiren adaptif modellerin geliştirilmesini hızlandırmıştır (Günay Yılmaz vd. 2023).

Ancak bu hızlı benimsenme gizlilik ve etik kaygılar ile performans değerlendirmelerinde yeni gereksinimler doğurmuştur.

2.3 Yüz Tanıma Sistemlerinin Güvenlik Açıkları: Biyometrik Saldırılar

Yüz tanıma sistemlerinin yaygınlaşması ve hassas uygulamalarda kullanılması bu sistemlerin güvenliğine yönelik tehditleri de artırmıştır. Saldırganlar bu sistemleri atlatarak yetkisiz erişim sağlamak veya kimliklerini gizlemek için çeşitli yöntemler geliştirmektedir. Bu tehditlerin en yaygın ve doğrudan olanları sunum saldırıları (Presentation Attacks, PAS) veya spoofing olarak adlandırılmaktadır (Schuckers 2016).

Sunum saldırısı, bir saldırganın biyometrik sistemin giriş sensörüne meşru bir kullanıcının yüzünü taklit eden bir sunum saldırı aracı (Presentation Attack Instrument, PAI) sunarak sistemi aldatmaya çalışması olarak tanımlanmaktadır (Kamat 2024). Bu saldırıların temel amacı sistemin canlı bir insan yerine bir artefaktı gerçek olarak algılamasını sağlamaktır (Sharma ve Selwal 2023a).

Başlıca sunum saldırıları iki boyutlu (2B) düz yüzey sunumlarından başlayarak üç boyutlu ve ileri tekniklere kadar uzanan bir spektrumda gerçekleştirilmektedir. En basit PAI türlerinden biri olan baskılı fotoğraf saldırısında yetkili kullanıcının yüksek çözünürlüklü bir görüntüsü lazer veya mürekkep püskürtmeli yazıcılarla basılarak kameraya sunulmaktadır. Buna benzer bir yöntem olan ekran fotoğrafı saldırısında ise aynı görüntüler cep telefonu, tablet veya bilgisayar ekranı üzerinden dijital olarak gösterilerek sistem aldatılmaya çalışılmaktadır (Hernandez-Ortega vd. 2019). Baskılı fotoğrafların statikliğini bir nebze hareketlendirmek amacıyla geliştirilen kesme fotoğraf saldırısında ise fotoğrafın göz, burun veya ağız bölümleri kesilip saldırganın yüzü bu boşlukların arkasına yerleştirilirken çarpık fotoğraf saldırısında basılı materyal fiziksel olarak bükülerek farklı açılardan dinamik bir yüz izlenimi yaratılmaktadır. Bir sonraki basamakta ise video tekrar oynatma saldırıları yer almaktadır. Bu yöntemde yetkili kullanıcıya ait önceden kaydedilmiş bir video döngüsel şekilde dijital bir ekranda oynatılarak göz kırpma ve baş hareketleri gibi temel canlılık kontrolleri atlatılmaya çalışılmaktadır.

İkinci boyutun ötesine geçen 3D maske saldırılarında silikon veya benzeri malzemelerden üretilmiş yüz hatlarına birebir uyumlu üç boyutlu maskeler kullanılmaktır. Bu maskelerde sıklıkla göz bölgeleri canlılık tespit mekanizmalarını aşmak üzere açık bırakılmaktadır. Makyaj saldırılarında ise profesyonel kozmetik ürünlerle saldırganın yüz hatları hedef kişinin görünümüne benzetilirken en gelişmiş tehdit unsuru sayılan deepfake saldırılarında derin öğrenme tabanlı algoritmalarla oluşturulmuş son derece gerçekçi sahte video ve görüntüler kullanılmaktadır. Bu yöntem dudak senkronizasyonu, yüz ifadeleri ve baş hareketleri gibi dinamik özellikleri taklit ederek modern tanıma sistemlerini yanıltmaktadır (Ming vd. 2020). Şekil 2.2'de bu saldırı türlerine örnek görseller yer almaktadır.

Şekil 2.2 Fotoğraf saldırısı, video saldırısı ve 3D maske saldırısı olmak üzere üç temel saldırı yaklaşımının örnek görüntüleri (Liu vd. 2016).

Başarılı biyometrik saldırılar güvenlik, finansal bütünlük ve sistem güvenilirliği açısından ciddi tehditler oluşturabilmektedir. Biyometrik sistemlere yönelik yapılan sahtecilik saldırıları yetkisiz kişilerin cihazlara, banka hesaplarına ve hassas kişisel verilere erişmesine olanak tanımaktadır (Yıldırım 2024). Bu tür ihlallerin doğrudan sonucu olarak kimlik hırsızlığı ve finansal dolandırıcılık gibi suçlar artış göstermektedir (Kurshan vd. 2024). Özellikle biyometrik verilerin değiştirilemez oluşu bu verilerin ele geçirilmesi durumunda kullanıcılar için kalıcı güvenlik açıkları yaratmaktadır. Bu duruma örnek olarak 2019 yılında meydana gelen ve biyometrik güvenliğin zafiyetlerini açıkça ortaya koyan ciddi bir veri sızıntısı verilebilir. Güney Kore merkezli Suprema adlı firmanın Biostar 2 sistemine ait 27,8 milyon kayıta gerekli güvenlik önlemleri alınmadığı için çevrimiçi olarak erişilebilmiştir. Bu veriler arasında kullanıcıların parmak izleri, yüz tanıma verileri, kullanıcı adları, şifreler ve kişisel bilgiler yer almaktadır. Üstelik bunların büyük bölümü şifrelenmemiş olarak saklanmaktaydı (Taylor 2019). Bu olay biyometrik sistemlerde veri güvenliği mimarisinin ve uçtan uca korumanın ne denli kritik olduğunu küresel ölçekte kanıtlamıştır.

Saldırıların artan sofistikasyonu ve PAI'ların daha gerçekçi hale gelmesi yüz sahteciliği tespiti araştırmalarını hızlandırmıştır. Saldırganların çevrimiçi kaynaklardan kolayca bilgi edinebilmesi ve 3D baskı gibi teknolojilerle yüksek kaliteli maskeler üretebilmesi geleneksel sahtecilik tespit yöntemlerini yetersiz kılmaktadır (Anthony vd. 2021). Bu durum yüz tanıma sistemlerinin güvenliğini sağlamak için daha gelişmiş ve genellenebilir biyometrik saldırı tespit (Presentation Attack Detection, PAD) tekniklerine acil bir ihtiyaç olduğunu göstermektedir.

Literatürdeki LBP ve HOG gibi klasik manuel özellik çıkarım yöntemleri ve optik akış temelli hareket analizleri basit baskı ve düşük kaliteli tekrar saldırılarını tespit etmekte başarılı olmaktadır. Ancak bu yöntemler yüksek çözünürlüklü baskılar veya profesyonel maskeler karşısında yetersiz kalabilmektedir (Antil ve Dhiman 2025). Derin öğrenme tabanlı sahtecilik tespit yöntemleri ham görüntülerden hem doku hem de derinlik haritası çıkarımı yaparak klasik yaklaşımlara kıyasla daha dirençli sistemler geliştirmiştir (Li vd. 2024). Yüz tanıma sistemlerinde sunum saldırı tespiti için özellikle derin öğrenme mimarileri uzamsal-zamansal özelliklerin analizinde güçlü araçlar sunmaktadır. 2025 yılında yayımlanan bir çalışmada MobileNetV3 gibi hafif bir CNN tabanlı mimari kullanılmıştır. Bu mimari sayesinde hem yüz dokusu ve yapısını içeren uzamsal özellikler hem de göz kırpma ve ışık yansımaları gibi zamansal dinamikler etkili bir şekilde analiz edilmiştir. Bu entegrasyon sistemin canlı ve sahte yüzleri ayırt etmede önemli bir başarı sergilemesini sağlamıştır (Khan vd. 2025).

Geleneksel doku ve şekil analizine dayalı yöntemler özellikle yüksek gerçekliğe sahip 3D maskeler ve deepfake videoları karşısında yetersiz kalabilmektedir. Bu durum görsel taklidin ötesine geçen fizyolojik tabanlı doğrulama yöntemlerine olan ihtiyacı artırmıştır. Bu bağlamda rPPG teknolojisi kritik bir çözüm olarak öne çıkmaktadır. Fotoğraf, maske veya ekran gibi sahte materyallerde bulunmayan deri altı kan akışı ve kardiyak nabız bilgisi bu teknikle temas gerektirmeden tespit edilebilmektedir. Bu benzersiz yetenek sayesinde rPPG, sunum saldırılarına karşı en ayırt edici ve yüksek direnç sunan yaklaşım olarak öne çıkmaktadır.

2.4 Yüz Canlılık Tespitinde Uzaktan Fotopletismografi Yaklaşımı

Uzaktan Fotopletismografi standart RGB kamera ve ortam ışığını kullanarak cilt altındaki kan hacmi değişimlerini analiz eden temassız bir fizyolojik ölçüm teknolojisidir. Bu yöntem mikroskobik renk dalgalanmalarının işlenmesi esasına dayanmaktadır. Temelleri 1937'de Hertzman'ın fotoelektrik pletismografi ile kan hacmi değişimlerini ölçmesiyle atılmıştır (Hertzman 1937). Modern rPPG ise Verkruysse ve meslektaşlarının 2008'de ortam ışığı ve kullanıcı kameraları kullanılarak kalp atış hızını uzaktan ölçülebileceğini göstermesiyle önemli bir ilerleme kaydetmiştir (Verkruysse vd. 2008). 2010'larda Eulerian video magnification gibi teknikler ciltteki ince renk değişimlerini güçlendirmek için geliştirilmiş ve bu sayede sinyal-gürültü oranı önemli ölçüde iyileştirilmiştir (Wu vd. 2012). Son yıllarda derin öğrenme tabanlı yöntemlerin entegrasyonu rPPG tekniğinin doğruluğunu ve uygulanabilirliğini daha da artırmıştır (Hernandez-Ortega vd. 2020).

Geleneksel PPG sistemleri parmak klipsi veya bileklik gibi deriyle temas gerektiren optik sensörler aracılığıyla benzer sinyalleri kaydederken (Sun ve Thakor 2016) rPPG sistemleri ise herhangi bir ek donanım olmaksızın yaygın tüketici sınıfı kameraları ile çalışabilmektedir (Verkruysse vd. 2008). Buna karşılık doğruluk, aydınlatma koşulları, cilt tonu farklılıkları ve hareket artefaktları gibi çevresel faktörlerden etkilenebilmektedir (van der Kooij ve Naber 2019). rPPG sinyalinin çıkarım süreci yüz tespiti, özellik takibi, renk kanallarının ayrıştırılması ve gürültü filtreleme adımlarını kapsamaktadır. Frekans analizi aşamasında ise Lomb–Scargle Periodogram gibi yöntemlere başvurulmaktadır (Wu vd. 2012, van der Kooij ve Naber 2019). Yeşil renk kanalı hemoglobin tarafından sağlanan güçlü emilim katsayısı nedeniyle sinyal çıkarımında en verimli kanal olarak kabul edilmektedir (Poh vd. 2010).

2.4.1 Geleneksel Sinyal İşleme Tabanlı Yöntemler

rPPG tabanlı canlılık tespiti gerçek yüzlerdeki belirgin periyodikliğin baskı, video veya 3D maske gibi sahte materyallerde ya hiç algılanamaması ya da anlamlı bir biyosinyal oluşturmaması ilkesine dayanmaktadır (de Haan ve Jeanne 2013). Sinyal kalitesini iyileştirmek için geliştirilen matematiksel ve istatistiksel yöntemler arasında CHROM

(Chrominance-based) (de Haan ve Jeanne 2013), POS (Plane-Orthogonal-to-Skin) (Wang vd. 2017a), ICA (Independent Component Analysis) (Karhunen vd. 1997) ve 2SR (Second-Order Statistical Ranking) (Wang vd. 2016) algoritmaları öncü yaklaşımlar arasında yer almaktadır. CHROM yöntemi farklı renk kanallarındaki faz bileşenlerini analiz etmektedir. Bu sayede nabız bileşeni ön plana çıkarılmaktadır (de Haan ve Jeanne 2013) Buna karşılık Wang ve arkadaşları tarafından geliştirilen POS yöntemi cildin spektral karakteristiklerini temel almaktadır. Bu algoritma RGB renk uzayında tanımlanan ortogonal bir düzlem üzerine izdüşüm yaparak yüzeydeki renk dalgalanmalarından fizyolojik sinyali ayrıştırmaktadır. (Wang vd. 2017b). Benzer şekilde ICA (Karhunen vd. 1997) yaklaşımı gözlemlenen karmaşık sinyalleri istatistiksel olarak birbirinden bağımsız bileşenlere ayırmaktadır. Böylece nabızla ilişkili bileşen izole edilmektedir (Poh vd. 2010). Bunlara ek olarak 2SR yöntemi ikincil dereceden istatistiksel sıralama kullanmaktadır. Bu sayede güvenilir bileşenler ayırt edilmektedir (Li vd. 2016). Bu yöntemlerin her biri biyometrik sinyallerin çevresel gürültüden arındırılması ve doğruluğunun artırılması noktasında literatürde temel yaklaşımlar olarak kabul edilmektedir.

2.4.2 Derin Öğrenme Tabanlı Yöntemler

Son yıllarda rPPG zaman dizilerini doğrudan işleyen Transformer tabanlı TransRPPG (Yu vd. 2021) ve CNN-LSTM hibrit modelleri canlılık tespiti ve sahtecilik sınıflandırmasında %98'in üzerinde başarı göstermiştir (Kim vd. 2022). Bu hibrit sistemler derin öğrenmenin yüksek ifade gücünü klasik sinyal işleme yaklaşımlarıyla birleştirerek rPPG temelli biyometrik doğrulama performansını önemli ölçüde artırmaktadır.

Derin öğrenme teknikleri rPPG sinyallerinin işlenmesi ve analiz edilmesi konusunda önemli ilerlemeler kaydetmiştir. Ancak derin öğrenme tabanlı rPPG yöntemleriyle elde edilen sinyallerin kalitesi özellikle zorlu aydınlatma koşulları, baş hareketleri veya kısmi yüz zorlukları gibi çevresel faktörlerden etkilenebilmektedir (Ma ve Chen 2025). Bu kalitenin artırılması amacıyla üretken çekişmeli ağlar (Generative Adversarial Networks, GAN) gibi mimariler kullanılmıştır. Kuang ve arkadaşları PhysNet tarafından çıkarılan

ham rPPG sinyallerini gürültüden arındırarak gerçek PPG sinyaline yaklaştırmak için koşullu GAN tabanlı rPPGGAN modeli önermiştir (Kuang vd. 2023a). Bu tür sinyal iyileştirme yöntemleri rPPG tabanlı biyometrik sistemlerin sağlamlığını artırma potansiyeli taşımaktadır. Ayrıca rPPG ölçümlerindeki belirsizliğin nicelleştirilmesi sistemin güvenilirliği açısından kritik bir konudur. Bu alanda Bayesian sinir ağları devreye sokulmuştur. 2022 yılında önerilen RF-BayesPhysNet (Robust Fusion Bayesian Physiological Network) modeli U-Net yapısı üzerine inşa edilerek hem aleatorik hem de epistemic belirsizlikleri modelleyebilme yeteneği kazanmıştır. Bu model hareket artefaktlarını ele almak için diferansiyel girdi işleme dalları kullanarak dinamik sahnelerdeki sağlamlığı önemli ölçüde artırmıştır (Ma ve Chen 2025). Benzer şekilde PPG sinyallerinden arteriyel kan basıncı dalga formunu tahmin etmek için geliştirilen GSW-UNet, UNet modelini GIFB, SDI ve WSB gibi yeni bloklarla geliştirerek sinyal işlemede detaylı ve anlamsal bilgileri birleştirmeyi başarmıştır (Tian vd. 2024). Bu çalışmalar rPPG sinyal kalitesini artırmaya ve daha sağlam fizyolojik çıkarımlar yapmaya yönelik önemli adımları temsil etmektedir.

Gerçek zamanlı biyometrik sistemlerde ve mobil platformlarda rPPG tabanlı uygulamaların dağıtımı için hesaplama verimliliği ve modelin hafifliği büyük önem taşımaktadır. Bu doğrultuda derin öğrenme modellerinin karmaşıklığını azaltmaya yönelik araştırmalar hız kazanmıştır. 2023 yılında literatüre kazandırılan Shuffle-rPPGNet, ShuffleNetV2 olarak adlandırılan hafif bir ağ mimarisi üzerine inşa edilmiştir. Söz konusu model 3D küresel bağlam ve 3D kanal karma tekniklerini entegre ederek kalp atış hızı değişkenliği ölçümünde yüksek hesaplama verimliliği sağlamıştır (Kuang vd. 2023b). Benzer bir yaklaşımla 3DCNN ve Transformer mimarilerini entegre ederek yerel uzamsal-zamansal özellikler ile küresel bağlamı bir araya getiren Uni-rPPGNet adında verimli bir mimari sunulmuştur (Liu vd. 2024). Benzer şekilde RTrPPG (Real-Time rPPG) çalışması ise gerçek zamanlı rPPG çıkarımı için ultra hafif bir 3DCNN modeli önermiştir. Bu model girdi boyutunu küçülterek yeni bir zaman ve frekans tabanlı kayıp fonksiyonunu kullanmaktadır. Önerilen model hem hız hem de doğruluk arasında optimum dengeyi yakalamıştır (Botina-Monsalve vd. 2022).

Transformer mimarilerinin de rPPG alanında kullanımı gözlemlenmektedir. Örneğin

2024 yılında tanıtılan CMRPPGFormer mimarisi 3 boyutlu uzamsal-zamansal evrişimsel modülasyon katmanlarını ve Transformer bloklarını birleştirmiştir. Bu mimari karmaşık görsel gürültüye ve zayıf sinyallere rağmen sağlam kalp atış hızı tahmini yapabilmektedir (Ma vd. 2024). Bu modeller hem performanstan ödün vermeden hem de kaynak kısıtlı ortamlarda çalışabilme yeteneği sunarak rPPG teknolojisinin geniş çaplı adaptasyonunu desteklemektedir. Remote photoplethysmography aynı zamanda biyometrik kimlik doğrulama süreçlerinde de kullanılmaya başlanmıştır. 2024'te sunulan bir çalışmada rPPG'ye dayalı biyometrik kimlik doğrulama tekniği önerilmiştir. Bu teknikte video verileri kalp atış hızı bilgisi ile yüzün uzamsal-zamansal dağılımını birlikte temsil eden çok ölçekli uzamsal-zamansal harita (MSTMap) biçimine dönüştürülmektedir. Bu sayede rPPG sinyalinin yüz üzerindeki dağılımına özgü ayırt edici özellikleri kimlik doğrulama amacıyla kullanılmaktadır. Özellik çıkarımı sürecinde hafif bir ViT modelinden faydalanılmış ve sınırlı eğitim verisi sorununu aşmak için maskelenmiş oto-kodlayıcılar mimariye dâhil edilmiştir (Geng vd. 2024).

Literatürdeki bu gelişmeler rPPG teknolojisinin çok yönlü ve genişletilebilir bir yapıya sahip olduğunu göstermektedir. Bu yöntem canlılık tespiti başarısının ötesine geçmektedir. Bireylerin benzersiz fizyolojik imzalarını kullanarak kimlik doğrulama süreçlerinde de etkin bir çözüm sunmaktadır. Tablo 2.1'de görsel doku özelliklerinden bağımsız olarak yalnızca yüzeyden yansıyan kan hacmi değişimlerine ve bu sinyallerin zamansal dinamiklerine odaklanan güncel çalışmalar kullanılan temel metodoloji, ön işleme teknikleri ve elde edilen performans metrikleri açısından özetlenmiştir.

Tablo 2.1 Sadece rPPG tabanlı yaklaşımları kullanan modellerin karşılaştırılması

Kaynak	Temel Metodoloji	Ön İşleme / Hizalama Yöntemi	rPPG Yöntemi	Öznitelik Füzyon Mekanizması	Kullanılan Veri Setleri	Doğrulama Şeması	Başarı oranı (%)
(Gomez vd. 2023)	Transfer Öğrenimi ve Alan Uyarlaması	Yüz Tespiti, ROI Ayarlama, Stabilizasyon, Boyutlandır-ma	CAN, Hareket Modeli, Görünüm Modeli	Alanlar Arası Transfer Öğrenim	Veridas DB (11 saldırı türü içeren özel veri seti)	Çift Katmanlı Ayırma	ACER: 14,61 BPCER: 18,28 APCER: 10,94 DeepPhys tabanlı model ACER sonucu: 43,38

Tablo 2.1 (Devam) Sadece rPPG tabanlı yaklaşımları kullanan modellerin karşılaştırılması

Kaynak	Temel Metodoloji	Ön İşleme / Hizalama Yöntemi	rPPG Yöntemi	Öznitelik Füzyon Mekanizması	Kullanılan Veri Setleri	Doğrulama Şeması	Başarı oranı (%)
(Liu vd. 2022)	Yerel rPPG Zamansal Benzerlik ve STConvNet; Denetimli	Yerel rPPG Bölgeleri için Landmark Tespiti, ROI Seçimi, Görünüm Bastırma	Öğrenilebi-lir Hafif Spatio temporal Konvolüs-yon Ağı	Skor Seviyesi Füzyon	3DMAD, HKBU-MARS V1/V2, CSMAD, HQ-WMCA-RGB, Replay Attack	LOOCV (20 Tur), Kümeler Arası, LOVO	Replay Attack: AUC: 88,9 EER: 18,6
(Savic ve Zhao 2025))	Güçlü Öz-Denetimli Öğrenme, Temel ağ mimarisi olarak Swin-Unet kullanılır.	Yüz Tespiti ve Hizalama, ROI Seçimi, Harita Oluşturma (MSTmap), Filtreleme	Derin Öğrenme Tabanlı Tahmin (Swin-Unet), Pozitif Örnekleme (Tmap)	Tmap içinde Kanal Düzeyinde Birleştirme	UBFC-RPPG, COHFACE, PURE, VIPL-HR, MR-NIRP Car	Eğitim - Doğrulama ayrımı	UBFC-RPPG (Oturum İçi): MAE: 0,62 bpm RMSE: 0,99 bpm R:0,99 PURE (Oturum İçi): MAE: 0,34 bpm RMSE: 0,58 bpm R: 99
(Huang vd. 2025))	Denetimsiz Girişimden Arındırılmış rPPG Tahmin Dalı (3DLDC) ve Girişim Tahmin Dalı	Yüz Tespiti, Arka Plan Seçimi, 30s Kesişmeyen Kliplere Bölme Veri Artırma	3DLDC tabanlı 3DCNN ve Denetimsiz Contrastive Learning	İki Aşamalı Çıkarma Yoluyla Temizleme	UBFC-RPPG, MMSE-HR, PURE, VIPL-HR, BH	Alan İçi Test, Çapraz Alan Testi	UBFC-RPPG (Oturum İçi): MAE: 0,19 bpm RMSE: 0,22 bpm R: 99 PURE (Oturum İçi): MAE: 0,10 bpm RMSE: 0,18 bpm R: 100
(Sun vd. 2024)	Morfolojik Geliştirme için Hibrit rPPG-cPPG Eğitimi; İki Aşamalı Denetimli	Yüz Videolarının Kimliksizleş-tirilmesi, Kırpma, Alt Örnekleme, Piksel Permütasyonu Harita Oluşturma	Morfolojik Geliştirme ile rPPG Modeli (CP2D (Contrast-Phys-2D) ve PPG-Morph Model.)	Hibrit rPPG-cPPG Eğitimi (Eğitimde Hibrit Dal Füzyonu / Çıkarımda Zamansal Ortalama.)	Özel Kimliksizleştir ilmiş Videolar + Harici cPPG	Oturum İçi ve Oturumlar Arası Test	PURE (Oturumlar Arası Test-Hareketli): AUC: 93,70 EER: 9,59
(Geng vd. 2024)	MSTMap üzerine Hafif ViT; MAE Ön Eğitimi	MSTMap'a Video Ön İşleme	rPPG İlişkili STMap Çıkarımı	Veri Seviyesinde Bütünleşik Temsil (MSTMap)	UBFC-RPPG, UBFC-PHYS	Eğitim/Test Ayrımı ve OVR	Accuracy: 97,64 FAR:4,33 FRR: 4,34
(Lee vd. 2025)	ROI/Çıkarım Değişkenleri + İstatistiksel Öznitelikler ile HR Örüntü Analizi; Denetimli	ROI Tespiti + RGB Hesaplama + Öznitelik Çıkarımı	Değişken Faktörler ile rPPG Çıkarımı	İstatistiksel Öznitelik Mühendisliği	Özel 164 Canlı/Sahte Video	İstatistiksel Doğrulama (164 Video)	feat – 1: t-value:5.216 df:125.958 p: 5.53E-07

Tablo 2.1 incelendiğinde fizyolojik tabanlı saldırı tespit sistemlerinin gelişiminde özellikle derin öğrenme mimarilerinin, transfer öğrenimi stratejilerinin ve Transformer tabanlı yapıların ağırlık kazandığı görülmektedir. Çalışmalar kullanılan veri setlerinin çeşitliliğine ve doğrulama şemalarına göre farklı performans karakteristikleri sergilemektedir. Gomez ve arkadaşları transfer öğrenimi ve alan uyarlama tekniklerini kullanmıştır. Araştırmacılar farklı saldırı türlerini içeren Veridas DB üzerinde %14,61 ACER ve %10,94 APCER değerlerine ulaşmıştır. Bu çalışmada kullanılan CAN (Convolutional Attention Network) mimarisinin temel DeepPhys modeline kıyasla %43,38 ACER hata oranını yaklaşık üçte bir oranına düşürerek belirgin bir iyileşme sağladığı görülmektedir. Benzer şekilde hafif mimarilere odaklanan Liu ve arkadaşları yerel rPPG bölgelerindeki zamansal benzerlikleri işleyen STConvNet mimarisini önermiştir. Bu mimari ile Replay Attack veri setinde %88,9 AUC ve %18,6 EER performansı elde edilmiştir. Son dönemlerde yapılan çalışmalarda Transformer ve öz-denetimli öğrenme yaklaşımlarının etkisinin arttığı görülmektedir. Savic ve Zhao Swin-Unet mimarisini güçlü bir öz-denetimli öğrenme stratejisiyle birleştirmiştir. UBFC-RPPG veri setinde 0,62 bpm MAE ve 0,99 korelasyon katsayısı gibi yüksek doğruluk değerlerine ulaşılmıştır. Geng ve arkadaşları ise rPPG'nin uzay-zamansal dağılımını hafif bir Vision Transformer ile işleyerek %97,64 doğruluk oranı ve %4,33 FAR ile başarılı bir sınıflandırma performansı raporlamıştır.

Yöntemsel çeşitlilik açısından Sun ve arkadaşları morfolojik geliştirme ve hibrit rPPG-cPPG eğitimi ile kimliksizleştirilmiş videolar üzerinde çalışmışlardır. Bu yaklaşım ile PURE veri setinde %93,70 AUC değerine ulaşılmıştır. Huang ve arkadaşları denetimsiz karşıtlık öğrenimi ve 3DCNN kullanarak rPPG sinyalini girişimlerden arındırmaya odaklanmıştır. Lee ve arkadaşları istatistiksel öznitelik mühendisliği ve HR örüntü analizi ile daha geleneksel bir yaklaşım sergilemiştir. Söz konusu bulgular rPPG sinyallerinin canlılık tespiti için kritik bir belirteç olduğunu açıkça ortaya koymaktadır. Ancak bu yöntemler yüksek kaliteli maske veya video saldırılarına karşı tek başına yeterli ayırt ediciliği sağlamakta zorlanabilmektedir. Bu nedenle tez çalışmasında önerilen görsel ve fizyolojik özellikleri birleştiren yaklaşım tek modaliteli yöntemlerin kısıtlarını aşmayı hedeflemektedir.

2.5 rPPG Sistemlerinde Karşılaşılan Zorluklar ve Kritik Faktörler

rPPG tabanlı canlılık tespit sistemlerinin başarısı elde edilen kan hacmi sinyallerinin kalitesine ve sinyal-gürültü oranının yüksekliğine doğrudan bağlı olmaktadır. Ancak optik prensiplere ve uzaktan görüntülemeye dayanan bu teknoloji dış faktörlere karşı oldukça hassas olabilmektedir. Kontrolsüz aydınlatma koşulları, kamera sensör gürültüleri ve özellikle deneğe ait biyolojik değişkenler performansı sınırlayan temel etkenler arasındadır. Sistemin gerçek dünya senaryolarında güvenilir bir şekilde çalışabilmesi için sinyal kalitesini bozan demografik etkenlerin ve fiziksel hareketlerin doğru analiz ve elimine edilmesi gerekmektedir. Bu bölümde rPPG performansını sınırlayan temel zorluklar ve literatürde bu sorunlara yönelik geliştirilen çözüm yaklaşımları ele alınmıştır.

2.5.1 Cilt Tonunun rPPG Sinyal Kalitesi ve PAD Üzerindeki Etkisi

Cilt rengi rPPG sinyalinin sağlamlığını etkileyen en kritik faktörlerden biridir (Chen vd. 2024). rPPG teknolojisi, kalp atışına bağlı olarak cilt yüzeyinde oluşan mikroskobik renk değişimlerinin ve ışık yansımasının analizine dayanmaktadır. Ancak koyu cilt tonlarında bulunan yüksek melanin içeriği ışık emilimini artırarak kameraya geri dönen yansıma miktarını azaltmaktadır. Işık yoğunluğundaki bu azalma fotopletismografi sinyalinin genliğini zayıflatmakta ve sinyalin gürültüye karşı daha savunmasız hale gelmesine neden olmaktadır (Kim vd. 2021). Bu fiziksel kısıt cilt tonu, aydınlatma ve kamera sensörü arasındaki etkileşimle birleştiğinde sinyal çıkarım sürecine önemli miktarda gürültü eklenmesine yol açmaktadır (Chen vd. 2024). rPPG teknolojisinin kan akışından kaynaklanan mikroskobik renk değişimlerini tespitine dayanması nedeniyle bu durum temel bir zorluk teşkil etmektedir (Kim vd. 2021).

Modellerin ağırlıklı olarak açık cilt tonlarına sahip verilerle eğitilmesi koyu cilt tonları üzerindeki performansın doğal olarak düşmesine neden olmaktadır. Bu sonuç gerçek dünya uygulamalarında erişim veya doğruluk noktasında eşitsizliklere yol açmaktadır. Dolayısıyla cilt tonu transferi gibi hedefe yönelik çözüm stratejilerinin geliştirilmesi bir zorunluluk haline gelmiştir.

Demografik Yanlılık ve Etkileri

Mevcut rPPG veri kümeleri özellikle daha koyu cilt tonları açısından çeşitlilikten yoksundur. Bu eksiklik mevcut yaklaşımların performansında yanlılığa yol açmaktadır. Literatürdeki çalışmalar özellikle Fitzpatrick Tip V ve VI olarak sınıflandırılan en koyu cilt tonlarına sahip denekler için kalp atış hızı tahmin hatalarının belirgin şekilde arttığını raporlamaktadır. Comas ve arkadaşları en zorlu cilt tipleri için bazı geleneksel rPPG yöntemlerinin derin öğrenme modellerinden daha iyi veya karşılaştırılabilir performans gösterebildiğini raporlamıştır. Bu demografik yanlılık rPPG tabanlı sistemlerin gerçek dünyadaki çeşitli popülasyonlarda adilliği ve genellenebilirlik için önemli sonuçlar doğurmaktadır (Comas vd. 2024).

Cilt tonuna bağlı farklılıkları azaltmaya yönelik çeşitli ileri düzey derin öğrenme teknikleri geliştirilmektedir. Bu kapsamda Comas ve arkadaşları uzaktan kalp atış hızı tahmininde cilt tonu çeşitliliğini artırmayı hedefleyen koşullu normalleştirici akışlara dayalı yeni bir yöntem olan PhysFlow mimarisini önermişlerdir. PhysFlow mimarisi denetimli rPPG yaklaşımlarının hem orijinal hem de üretilen veriler üzerinde eş zamanlı eğitimini sağlayan uçtan uca eğitim optimizasyonunu benimsemektedir. Modeller manuel Fitzpatrick cilt tonu etiketlerine ihtiyaç duymaksızın CIELAB renk uzayında cilt özellikleri üzerinden doğrudan yüz videolarından koşullandırılmaktadır. Bu yaklaşım özellikle koyu cilt tonları için hatayı önemli ölçüde azaltmıştır. Yönteme bağlı olarak ortalama mutlak hata değerinde 1 ila 5 BPM'lik bir düşüş sağlandığı raporlanmıştır (Comas vd. 2024). Elde edilen bulgular performans düşüşlerinin sadece model mimarisinden kaynaklanmadığını bu durumun eğitim veri kümelerindeki çeşitlilik eksikliğiyle doğrudan ilişkili olduğunu göstermektedir. PhysFlow örneğinde olduğu gibi sentetik veri üretimi yoluyla çeşitliliğin artırılması veri kümesi sınırlamalarını aşmak için güçlü bir strateji sunmaktadır. Bu bağlamda adil biyometrik sistemlerin geliştirilmesi için sentetik veri üretim teknikleri vazgeçilmez bir çözüm olarak öne çıkmaktadır.

Literatürde derin öğrenme tabanlı yaklaşımlara alternatif olarak renk uzayı dönüşümleri ve filtreleme yöntemleri de önerilmektedir. Chen ve arkadaşları geleneksel RGB renk uzayı yerine YCbCr renk uzayının kullanılmasını önermektedir. YCbCr renk uzayı daha

ince cilt rengi varyasyonlarının temsil edilmesine olanak tanıyarak BVP sinyal çıkarım performansını iyileştirebilmektedir. Ayrıca aydınlatma varyasyonlarını bastırmak ve telafi etmek amacıyla normalize edilmiş en küçük kareler (NLMS) adaptif filtrelerinin kullanılabileceği belirtilmektedir. Ancak bu yöntemin düzeltici referans sinyaline ihtiyaç duyması uygulanabilirliğini zorlaştırmaktadır. Benzer şekilde Distance-PPG yöntemi farklı yüz bölgelerindeki renk değişimlerini ağırlıklandırarak yüksek gürültü önleme performansı sunmaktadır. Ancak algoritmanın hesaplama karmaşıklığı ve işlem süresi gerçek zamanlı uygulamalar için bir kısıt oluşturmaktadır (Chen vd. 2024).

2.5.2 rPPG Tabanlı PAD'de Mikro Hareket ve Doku Analizi

Fotopletismografi sinyalleri temel olarak yüzdeki renk değişimlerine odaklansa da yüksek kaliteli 3D maskeler veya deepfake videoları gibi sofistike sunum saldırılarını tespit etmek için yalnızca renk bilgisini kullanmak her zaman yeterli olmayabilmektedir. Gerçek bir insan yüzü ile cansız bir nesne veya dijital bir ekran arasındaki fark çoğu zaman yüzey dokusundaki mikroskobik detaylarda ve fizyolojik kaynaklı istemsiz mikro hareketlerde gizlenmektedir. Bu nedenle rPPG sinyallerini tamamlayıcı nitelikteki doku ve yapı gibi uzamsal özniteliklerin yanı sıra hareket ve dinamiklere dayalı zamansal özniteliklerin analizi bütüncül bir canlılık tespiti için kritik önem taşımaktadır. Aşağıdaki alt başlıklarda bu dinamik özelliklerin fizyolojik temelleri ve sahtecilik tespitindeki rolleri incelenmiştir.

İnce Yüz Mikro Hareketlerinin Fizyolojik İpuçları Olarak Rolü

Kardiyovasküler nabız, yüzde oluşan mikro hareketlerin veya renk varyasyonlarının zamansal sinyallerinin analizi yoluyla tahmin edilebilmektedir (Chen vd. 2024). Gerçek yüz ifadeleri ağız hareketleri ve göz kırpma gibi ince düzeydeki yüz hareketlerini barındırmaktadır. Bu tür hareketler çoğu sunum saldırısında kullanılan maskelerin malzeme sertliği nedeniyle taklit edilmesi zor olan unsur arasında yer almaktadır (Jiang vd. 2025). Ayrıca söz konusu mikro hareketler canlılık tespiti sürecinde ayırt edici fizyolojik ipuçları olarak kullanılabilmektedir (Lee vd. 2025).

Yüz hareketindeki farklılıkları yakalamak için optik akış özellikleri yaygın olarak kullanılmaktadır. Bu kapsamdaki çalışmalar evrimsel özellik kanallarını kullanarak optik akışı tahmin etmek ve dinamik doku özelliklerini öğrenmek suretiyle ince yüz hareketlerini incelemektedir (Tsitiridis vd. 2019). Hareket rPPG sinyal kalitesini düşüren önemli gürültü kaynağı olarak tanımlanmaktadır. Bu durum sofistike hareket telafi tekniklerine ihtiyaç duyulduğunu ortaya koymaktadır (Lee vd. 2025). Ancak Jiang ve arkadaşları göz kırpma gibi ince mikro hareketlerin, sahtecilik saldırılarının taklit etmekte zorlandığı en güçlü canlılık kanıtları olduğunu savunmaktadır (Jiang vd. 2025). Bu durum hassas bir denge gerektirmektedir. Makro hareketler zararlı gürültüyken mikro hareketler değerli bir sinyal kaynağıdır. Dolayısıyla etkili rPPG tabanlı PAD sistemleri yalnızca istenmeyen hareket artefaktlarını bastırmakla kalmamalı aynı zamanda gerçek ve ince fizyolojik hareketleri akılcı biçimde çıkarıp analiz etmelidir. Bu gereklilik hareket dinamiklerinin derinlemesine anlaşılmasını zorunlu kılmaktadır.

Hareket Telafisi Teknikleri ve Önemi

Hareket artefaktları rPPG için video kayıtlarındaki en yaygın parazit kaynaklarından biri olarak değerlendirilmektedir (Chen vd. 2024). Güvenilir ve doğru rPPG sinyal çıkarımını sağlamak için hareket telafisi tekniklerinin her kareyi bir referans yüzle hizalaması kritik öneme sahiptir (Seibold vd. 2025).

Hareket artefaktlarını azaltmak veya ortadan kaldırmak için alt-bant rPPG, sürekli dalgacık dönüşümü, sınırlı Kalman filtresi tekniği ve Hareket İndeksi (Motion Index, MI) göstergeleri gibi yöntemler geliştirilmiştir (Chen vd. 2024). Örneğin MediaPipe tabanlı yaklaşımlar yüz dönüm noktalarını tespit ederek 2B bir ağ yapısı oluşturmaktadır. Bu ağ yapısı daha sonra her bir üçgenin karşılık gelen referans konumuna dönüştürmek için kullanılmakta ve hareket telafili bir görüntü dizisi oluşturmaktadır (Seibold vd. 2025). Düzgün bir şekilde tespit edilmemiş yüze sahip karelerin analizden dışlanması kısa süreli analizlerde genel performansı artırabilmektedir (Hernandez-Ortega vd. 2018). Her ne kadar hareket çoğu senaryoda bir gürültü kaynağı olsa da bazı yaklaşımlar uygun kamera yapılandırmasıyla güvenilir kardiyak hareket modelleri çıkarmayı amaçlamaktadır (Li vd. 2023).

Doku Analizi Tabanlı Ayrıştırma

Sunum saldırıları canlılık, yansıma, doku, kalite ve spektral bilgi gibi görüntü özelliklerindeki anormalliklerden tespit edilebilmektedir (Tsitiridis vd. 2019). Doku tabanlı yöntemler yüzlerin dokusunu, yapısını ve genel şekil bilgilerini incelemektedir. 3D maskeler insan yüzünün hacimsel yapısını taklit edebilse de üretim teknolojisi nedeniyle genellikle renk dağılımı ve doku kusurları içermektedir. Bu doku farklılıklarını keşfetmek amacıyla LBP öznitelikleri genellikle destek vektör makineleri sınıflandırıcılarıyla birleştirilmektedir.

Mikro doku farklılıklarını daha hassas biçimde yakalayabilmek için yerel görüntü yamalarından derin özniteliklerin öğrenilmesine dayalı yaklaşımlar da önerilmektedir. Yüksek kaliteli silikon maskelerde dokuya özgü ince yapıları ayırt edebilmek amacıyla çok seviyeli derin sözlük tabanlı temsil yöntemleri kullanılmaktadır (Jiang vd. 2025). rPPG temelli yaklaşımlar renk değişikliklerine odaklanırken temel cilt dokusu ve varyasyonları ise PAD sistemleri için tamamlayıcı bilgi sağlayabilmektedir. rPPG güçlü bir canlılık göstergesi olsa da gelişmiş videolarında kalp atış hızı örüntülerinin yapay olarak taklit edilebilmesi bu yöntemin tek başına etkinliğini sınırlayabilmektedir (Seibold vd. 2025). Doku analizi ve mikro hareket analizi gibi tekil PAD yöntemleri incelendiğinde saldırıların artan gerçekçilik düzeyi karşısında yalnızca tek bir modaliteye dayalı savunma mekanizmalarının yetersiz kalabileceği açıkça görülmektedir (Jiang vd. 2025). Bu kısıtları aşmak adına rPPG sinyallerinin mikro-doku varyasyonları gibi yapısal özellikler ve optik akış temelli hareket ipuçları ile entegre edilmesi gerekmektedir. Bu entegrasyon çok modlu ve katmanlı bir savunma mekanizması oluşturmaktadır. Mevcut literatürde elde edilen bulgular gelecekte geliştirilecek sağlam ve genellenebilir PAD sistemlerinin başarısının tekil modalitelerin temsil kapasitesini aşacak biçimde fizyolojik ve görsel bilgi kaynaklarının bütünleşik olarak modellenmesine bağlı olduğunu ortaya koymaktadır.

3. DERİN ÖĞRENME MİMARİLERİ

Derin öğrenme çok katmanlı yapay sinir ağlarının yüksek boyutlu veri temsillerini hiyerarşik olarak öğrenebilmesini sağlayan bir hesaplama paradigması olarak tanımlanmaktadır. İnsan beynindeki nöronal bilgi işleme prensiplerinden esinlenen bu yöntem özellikle görüntü, ses, metin ve zaman serisi gibi karmaşık veri dağılımlarının temsil edilmesi ve sınıflandırılması için modern yapay zekâ uygulamalarında standart bir yaklaşım haline gelmiştir. Derin öğrenme mimarileri geleneksel makine öğrenmesi yöntemlerinden temel bir farkla ayrılmaktadır. Geleneksel yaklaşımlarda öznitelik çıkarımı süreçleri manuel özellik mühendisliği gerektirmektedir. Buna karşılık derin öğrenme modelleri veri üzerinden uçtan uca ve otomatik öznitelik öğrenimi gerçekleştirmektedir. Bu yetenek ham veriden soyut ve üst düzey temsillerin doğrudan öğrenilmesine olanak tanımaktadır.

Bu mimarilerin temelinde doğrusal olmayan aktivasyon fonksiyonlarıyla birleşen çok katmanlı dönüşümler yer almaktadır. Bu yapı sayesinde modeller giriş uzayındaki karmaşık ve doğrusal olmayan ilişkileri öğrenebilme yeteneği kazanabilmektedir. Bir x giriş verisi ardışık doğrusal dönüşümler ve aktivasyon fonksiyonlarından geçirilerek yüksek seviyeli soyut özellikler hâline dönüştürülmektedir. Bir gizli katmanın çıktısı genel olarak Denklem 3.1'de gösterildiği şekilde ifade edilmektedir (Goodfellow vd. 2016).

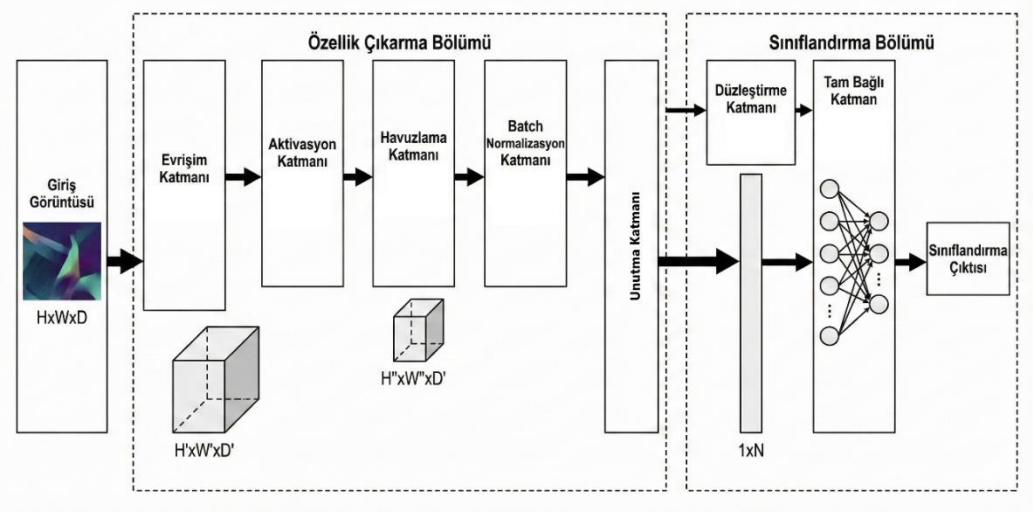
$$h^{(l)} = \sigma\left(W^{(l)}h^{(l-1)} + b^{(l)}\right) \tag{3.1}$$

Denklem 3.1'de $h^{(l)}$ ifadesi ağın l. katmanındaki çıktı aktivasyonlarını temsil etmektedir. Bu çıktı bir önceki katmandan gelen $h^{(l-1)}$ girdilerinin $W^{(l)}$ ağırlık matrisi ile çarpılması ve buna $b^{(l)}$ bias teriminin eklenmesiyle elde edilen doğrusal birleşimine, doğrusal olmayan bir aktivasyon fonksiyonu olan σ fonksiyonunun uygulanması sonucu hesaplanmaktadır. Burada $W^{(l)}$ ilgili katmandaki öğrenilebilir ağırlıkları, $b^{(l)}$ ise bias terimini ifade etmektedir. σ aktivasyon fonksiyonu ise ağın karmaşık ve doğrusal olmayan ilişkileri öğrenebilmesini sağlamaktadır.

3.1 Evrişimsel Sinir Ağları

Derin öğrenme, verilerdeki karmaşık yapıları ve örüntüleri öğrenmek amacıyla çok katmanlı yapay sinir ağlarının kullanıldığı bir makine öğrenmesi alt dalıdır (Akın ve Şahin 2024). Özellikle görsel verilerin işlenmesinde biyolojik görme sisteminden ilham alınarak geliştirilen evrişimsel sinir ağları standart yaklaşım haline gelmiştir (Lecun vd. 1998). CNN mimarileri ham piksel verisinden kenar, doku, şekil ve anlamsal bütünlük gibi hiyerarşik öznitelikler çıkarma yeteneğine sahiptir (Lecun vd. 1998). Bu hiyerarşi ile erken katmanlarda kenar ve köşe gibi düşük seviyeli öznitelikler öğrenilirken derin katmanlara doğru nesne parçaları veya bütünsel nesneler gibi soyut kavramlara geçiş yapılmaktadır.

Temel bir CNN mimarisi evrişim katmanı, havuzlama katmanı ve tam bağlı katmandan (Lecun vd. 1998) oluşmaktadır. Modern CNN mimarilerinde bu temel katmanlara ek olarak aktivasyon katmanları, düzleştirme katmanı, parti normalizasyonu ve unutma katmanı gibi yardımcı katmanlar da sıklıkla kullanılmaktadır. Bu katmanlar modelin doğrusal olmayan öğrenme yeteneğini artırırken eğitim stabilitesini sağlamaya ve modelin ezberlemesini önlemeye yardımcı olabilmektedir. Bu katmanlar girdi görüntüsünün işlenmesini adım adım gerçekleştirerek modelin verimli ve genellenebilir olmasını sağlamaktadır. Aşağıdaki alt başlıklarda CNN mimarisini oluşturan her bir katmanın çalışma prensibi ayrıntılı olarak açıklanmaktadır. CNN mimarisinin temel işleyişi Şekil 3.1'de yer almaktadır.



Şekil 3.1 CNN mimarisinin temel blok diyagramı

3.1.1 Evrişim Katmanı

Evrişimsel sinir ağlarının temel yapı taşı olan evrişim katmanı, girdi görüntüsü üzerinde sistematik bir biçimde dolaştırılan filtreler aracılığıyla yerel özniteliklerin tespit edildiği ana bileşendir. Bu süreçte her bir filtre, görüntünün belirli bir bölgesindeki piksel değerlerini ağırlıklandırarak bir öznitelik haritası üretmektedir. Evrişim işlemi, filtrelerin kaydırma penceresi tekniğiyle görüntü üzerinde ilerletilmesi ve her bir konumda nokta çarpımı hesaplanarak sonuçların birleştirilmesiyle gerçekleştirilmektedir. Öznitelik haritasının niteliği birkaç temel hiper-parametre tarafından kontrol edilmektedir. Filtre boyutu yerel bağlamı yakalamak ve hesaplama maliyetini düşük tutmak amacıyla genellikle 3×3 veya 5×5 gibi küçük boyutlarda tercih edilmektedir. Adım boyutu ise filtrenin her kaydırma adımında ne kadar ilerleyeceğini belirleyen temel bir parametredir. Adım boyutunun bir olarak seçilmesi tüm piksellerin işlenmesini sağlarken iki olarak belirlenmesi çıktının uzamsal boyutlarını yaklaşık olarak yarıya indirmektedir. Görüntü kenarlarına sıfır değerler eklenmesini ifade eden dolgu işlemi ise filtrelerin kenar piksellerini de işleme dâhil etmesine olanak tanıyarak çıktı boyutunun korunmasına yardımcı olmaktadır. Bir evrişim katmanında genellikle 32 veya 64 adet olmak üzere birden fazla filtre kullanılmaktadır. Bu çoklu filtre yapısı kenar ve doku gibi farklı özniteliklerin paralel bir şekilde çıkarılmasını sağlamaktadır. İşlem sonucunda çıktı tensörünün kanal sayısı artırılmaktadır. Böylece $Y \times Z \times C$ formatındaki bir girdi ile $Y' \times Z' \times F$ formatında bir çıktı elde edilmektedir. Bu işlem ağırlık paylaşımı mekanizması sayesinde parametre sayısını önemli ölçüde azaltmaktadır. Aynı zamanda modele ötelemeye karşı değişmezlik yeteneği kazandırmaktadır. Ağırlık paylaşımı prensibi aynı filtrenin görüntünün farklı bölgelerinde yeniden kullanılması esasına dayanmaktadır. Bu yaklaşım ile modelin konumdan bağımsız öznitelik öğrenme kapasitesi güçlendirilmektedir.

Girdi görüntüsü $X \in \mathbb{R}^{H \times W \times C}$ için tek bir evrişim katmanında uygulanan evrişimsel işlem, çekirdek $K \in \mathbb{R}^{k_h \times k_w \times C}$ ile nokta çarpımların kaydırmalı toplanması şeklinde formüle edilmektedir. Bir çıktı öznitelik haritasının tek bir konumundaki değeri Denklem 3.2'de gösterilmektedir (Gonzalez ve Woods 2018).

$$\mathrm{y(i,j)} = \sum_{u=1}^{kh} \sum_{v=1}^{kw} \sum_{c=1}^{C} K\,(\mathrm{u,v,c})\ \mathrm{X}\ (\mathrm{i+u-1, j+v-1, c}) + \mathrm{b} \qquad (3.2)$$

Denklemde 3.2'de $y(i,j)$ ifadesi çıktı özellik haritasının (i, j) konumundaki değerini temsil etmektedir. Bu değer giriş verisi üzerinde uygulanan evrişim işlemi sonucunda elde edilmektedir. Evrişim işlemi sırasında yüksekliği k_h ve genişliği k_w olan bir çekirdek, giriş verisi üzerinde kaydırılarak uygulanmaktadır. Her bir (i, j) konumu için çekirdeğin u ve v indisleri boyunca ve giriş verisinin tüm C kanal sayısı üzerinden çekirdek ağırlıkları $K(u, v, c)$ ile giriş verisinin ilgili piksel değerleri $X(i + u - 1, j + v - 1, c)$ çarpılarak toplanmaktadır. Bu toplama işlemi sonucuna bias terimi olan b eklenerek çıktı değeri elde edilmektedir. Böylece evrişim işlemi çekirdeğin tüm elemanları ve girişin tüm kanalları dikkate alınarak gerçekleştirilmiş olmaktadır.

Evrişimsel yapının yerel bağlantı düzeni ve paylaşılmış ağırlık mekanizması hem parametrik verimliliği artırmakta hem de mekânsal yapı bilgisinin korunmasını sağlamaktadır. Bu özellikler görüntü sınıflandırma, nesne tespiti ve yüz analizi gibi görevlerde öğrenilen temsillerin etkinliğini artırmaktadır. CNN'lerin temel başarı gerekçeleri ve uygulama örnekleri literatürde sistematik biçimde ortaya konulmuştur. Bununla birlikte ağ derinliğinin artırılmasıyla temsil kapasitesi yükselirken optimizasyon güçlükleri ve gradyan sönmesi gibi problemler daha belirgin hâle gelmiştir. Bu problemlere yönelik geliştirilen mimari çözümler ve normalizasyon yöntemleri ile model eğitiminin kararlılığı sağlanmıştır. Bu bağlamda evrişim katmanlarının derinleştirilmesine dayalı mimariler modern görsel analiz problemlerinde temel bir uygulama paradigması hâline gelmiştir.

3.1.2 Aktivasyon Katmanı

Evrişim katmanından sonra genellikle doğrusal olmayan bir aktivasyon fonksiyonu uygulanmaktadır. Bu katman modelin doğrusal olmayan dönüşümleri öğrenebilmesini sağlamakta ve negatif değerleri sıfırlayarak hesaplama hızını ve verimliliğini artırmaktadır (Nair ve Hinton 2010a). Literatürde yaygın olarak kullanılan ReLU (Rectified Linear Unit) aktivasyon fonksiyonunun matematiksel ifadesi Denklem 3.3'te sunulmuştur (Nair ve Hinton 2010b).

$$f(x) = \max(0, x) \tag{3.3}$$

Denklemde 3.3'te $f(x)$ aktivasyon katmanının çıktısını temsil ederken x katmana giren değeri ifade etmektedir. ReLU aktivasyon fonksiyonu giriş değeri x ile sıfır değerini karşılaştırarak büyük olanı seçmektedir. Yani $f(x) = \max(0, x)$ şeklinde tanımlanmaktadır. Bu sayede negatif girişler bastırılırken pozitif değerler aynen aktarılmakta ve ağın daha etkin bir şekilde öğrenmesi sağlanmaktadır. Bu işlem gradyan sönmesi problemini önemli ölçüde azaltarak derin ağların daha hızlı ve kararlı biçimde eğitilmesine olanak tanımaktadır (Nair ve Hinton 2010a). Standart ReLU dışında Leaky ReLU ve ELU (Exponential Linear Unit) gibi varyantlar da bulunmaktadır. Bu aktivasyon fonksiyonlarının negatif değerlerdeki bilgi kaybını azalttığı ve öğrenme performansını iyileştirdiği raporlanmıştır (Maniatopoulos ve Mitianoudis 2021).

3.1.3 Havuzlama Katmanı

Havuzlama katmanı öznitelik haritalarının boyutsal çözünürlüğünü düşürerek hesaplama yükünü azaltmak ve modelin gürültüye karşı direncini artırmak amacıyla kullanılmaktadır. Bu katman evrişim katmanlarından elde edilen özellik haritalarını boyut indirgeme yoluyla küçültmektedir. Bu işlem modelin girdideki küçük konum kaymalarına karşı belirli ölçüde değişmezlik kazanmasını sağlamaktadır. Bu özellik nesnenin konumundan bağımsız olarak tanınabilmesi açısından kritik öneme sahiptir.

Maksimum havuzlama işlemi evrişim katmanından elde edilen özellik haritalarının boyutunu azaltmak ve hesaplama maliyetini düşürmek amacıyla kullanılmaktadır. Bu yöntemde $k \times k$ olarak belirlenen bir pencere boyutu içerisindeki en yüksek değer seçilerek çıktı oluşturulmaktadır. Böylece güçlü kenar veya köşe bilgileri gibi ayırt edici özellikler korunmaktadır. Özellik haritasının uzamsal boyutu genellikle kayma miktarı $s = 2$ olacak şekilde yaklaşık olarak yarıya indirilmektedir. Maksimum havuzlama işleminin matematiksel ifadesi Denklem 3.4'te verilmiştir (Scherer vd. 2010).

$$y(i,j) = max_{u=0}^{k-1}\, max_{v=0}^{k-1} X(i \cdot s + u, j \cdot s + v) \tag{3.4}$$

Denklem 3.4'te $y(i,j)$ havuzlama işlemi sonucunda elde edilen çıktı özellik haritasının i. satır ve j. sütunundaki elemanını temsil etmektedir. Bu değer havuzlama katmanına giren giriş özellik haritası X üzerinde boyutu k olan bir havuzlama penceresinin kayma miktarı s ile gezdirilmesi sonucu elde edilmektedir. Her bir konumda pencere içerisindeki u ve v indeksleri boyunca yer alan değerler arasından maksimum olanı seçilmektedir. Burada k havuzlama penceresinin boyutunu, s ise pencerenin giriş üzerinde kaç piksel kaydırılacağını ifade etmektedir. Bu işlem verilen bölgedeki en yüksek değeri seçerek bilgi kaybını en aza indirirken boyut indirgeme sağlamaktadır.

Maksimum havuzlamaya alternatif olarak kullanılan bir diğer yöntem ise ortalama havuzlama işlemidir. Ortalama havuzlama pencere içerisindeki tüm değerlerin aritmetik ortalamasını alarak çıktı üretmektedir. Bu yöntem gürültüyü daha fazla yumuşatabilmesine rağmen keskin ve ayırt edici özniteliklerin zayıflamasına neden olabilmektedir. Ortalama havuzlama işleminin matematiksel ifadesi Denklem 3.5'te gösterilmektedir (Goodfellow vd. 2016).

$$y(i,j) = \frac{1}{k^2} \sum_{u=0}^{k-1} \sum_{v=0}^{k-1} X(i \cdot s + u, j \cdot s + v) \tag{3.5}$$

Denklem 3.5'te $y(i,j)$ ortalama havuzlama işlemi sonucunda elde edilen çıktı değerinin i. satır ve j. sütunundaki elemanını ifade etmektedir. $X(i \cdot s + u, j \cdot s + v)$ havuzlama katmanına giren giriş özellik haritasındaki ilgili değeri temsil ederken k havuzlama penceresinin boyutunu s ise kayma miktarını belirtmektedir. $\frac{1}{k^2}$ katsayısı ise pencere içerisindeki tüm değerlerin ortalamasını almak amacıyla uygulanmaktadır.

3.1.4 Yığın Normalizasyon Katmanı

Yığın (Batch) normalizasyon katmanı her bir mini grup için giriş verilerinin ortalama ve varyansını normalize ederek ağın eğitim sürecini hızlandırmaktadır. Bu katman aynı zamanda gradyan temelli optimizasyon sorunların azaltılmasına katkı sağlamaktadır. Özellikle derin sinir ağlarında karşılaşılan iç kovaryant problemini azaltarak daha kararlı ve hızlı bir öğrenme süreci sunmaktadır. Yığın normalizasyon işleminin temel

normalizasyon adımı Denklem 3.6'da gösterilmektedir (Ioffe ve Szegedy 2015).

$$\hat{x} = \frac{x - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \tag{3.6}$$

Denklem 3.6'da $\hat{x}$ normalize edilmiş çıktı değerini, x normalizasyon uygulanacak giriş değerini ifade etmektedir. μ_B ilgili mini gruba ait giriş değerlerinin ortalamasını, σ_B^2 ise bu mini gruba ait varyansı temsil etmektedir. ϵ terimi bölme işlemi sırasında sıfıra bölme hatasını önlemek ve sayısal kararlılığı artırmak amacıyla eklenen çok küçük bir sabit değerdir.

Normalizasyon adımının ardından ağın temsil gücünü koruyabilmesi için ölçekleme ve kaydırma işlemleri uygulanmaktadır. Bu işlem Denklem 3.7'de verilen doğrusal dönüşüm ile gerçekleştirilmektedir (Ioffe ve Szegedy 2015).

$$y = \gamma \hat{x} + \beta \tag{3.7}$$

Denklem 3.7'de y yığın normalizasyon işlemi sonrası elde edilen son çıktıyı temsil etmektedir. $\hat{x}$ ise Denklem 3.6'da elde edilen normalize edilmiş değeri ifade etmektedir. γ ölçeklendirme parametresini, β ise kaydırma parametresini temsil etmektedir. Her iki parametre de eğitim sürecinde öğrenilebilmektedir. Bu parametreler sayesinde ağ gerekirse normalizasyon sonrası veriyi yeniden ölçekleyerek uygun bir dağılıma dönüştürebilmektedir. Sonuç olarak μ_B ve σ_B^2 mini grup istatistiklerini, γ ve β ise öğrenilebilir parametreleri temsil etmektedir. Yığın normalizasyon katmanı daha yüksek öğrenme oranlarının güvenli biçimde kullanılmasına olanak tanıyarak eğitim sürecini stabilize etmekte ve derin ağların daha etkin ve hızlı bir şekilde eğitilmesini sağlamaktadır.

3.1.5 Unutma Katmanı

Unutma (Dropout) katmanı derin sinir ağlarında aşırı öğrenmeyi önlemek amacıyla kullanılan etkili bir düzenleme yöntemidir. Bu katmanda eğitim süreci boyunca belirli bir

oranla nöronlar rastgele olarak geçici biçimde devre dışı bırakılmaktadır. Diğer bir ifadeyle bu nöronların çıktıları sıfırlanmaktadır. Bu yaklaşım modelin belirli nöronlara veya bağlantılara aşırı bağımlı hâle gelmesini engelleyerek genelleme yeteneğini artırmaktadır. Unutma oranı genellikle 0,5 olarak seçilmektedir. Bu oran her bir nöronun eğitim sırasında pasif hale getirilme olasılığını ifade etmektedir. Eğitim aşamasında uygulanan Unutma işlemi test aşamasında devre dışı bırakılmaktadır. Test sürecinde tüm nöronlar aktif hâle getirilmekte ve çıktılar uygun şekilde ölçeklendirilmektedir. Bu teknik ağın farklı alt ağ kombinasyonlarıyla öğrenmesini teşvik etmektedir. Böylece model topluluğu etkisi yaratılarak daha sağlam, genellenebilir ve aşırı öğrenmeye karşı dirençli bir öğrenme süreci sağlanmaktadır.

3.1.6 Düzleştirme Katmanı

Düzleştirme katmanı, evrişim ve havuzlama katmanlarından elde edilen çok boyutlu öznitelik haritalarını tek boyutlu bir vektöre dönüştürmek amacıyla kullanılmaktadır. Bu dönüşüm tam bağlı katmanlara geçişi mümkün kılmaktadır. Örneğin boyutu H'×W'×F olan bir öznitelik tensörü düzleştirme işlemi sonucunda H'·W'·F uzunluğunda bir vektöre dönüştürülmektedir. Bu işlem uzamsal konum bilgisinin kaybolmasına neden olsa da sınıflandırma gibi yüksek seviyeli görevler için gerekli olan küresel özniteliklerin bir araya getirilmesini sağlamaktadır.

3.1.7 Tam Bağlı Katman

Tam bağlı katman evrişim ve havuzlama katmanlarından elde edilen yüksek seviyeli özniteliklerin sınıflandırma kararına dönüştürüldüğü son aşamayı oluşturmaktadır. Bu katman geleneksel çok katmanlı algılayıcı yapısına benzer şekilde çalışmaktadır. Düzleştirilmiş giriş vektöründeki tüm elemanlar bu katmandaki her bir nöron ile tam bağlantılıdır. Bu aşamada öncelikle özellik haritaları tek boyutlu bir vektöre dönüştürülmektedir. Ardından her bir nöron bu vektör ile ağırlık matrisi W arasında nokta çarpımı gerçekleştirmektedir. Bu doğrusal işlem Denklem 3.8'de gösterilmektedir (Goodfellow vd. 2016).

$$z = W \cdot x + b \tag{3.8}$$

Denklem 3.8'de z nöronun doğrusal kombinasyon sonucu ürettiği değeri ifade etmektedir. W ağırlık matrisini, x düzleştirilmiş giriş vektörünü ve b bias terimini ifade etmektedir. Doğrusal dönüşüm sonucunda elde edilen z değerlerine bir aktivasyon fonksiyonu uygulanarak nihai çıktı elde edilmektedir. Çok sınıflı sınıflandırma problemlerinde bu amaçla sıklıkla softmax aktivasyon fonksiyonu tercih edilmektedir. Softmax fonksiyonunun matematiksel ifadesi Denklem 3.9'da verilmiştir (Goodfellow vd. 2016).

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{3.9}$$

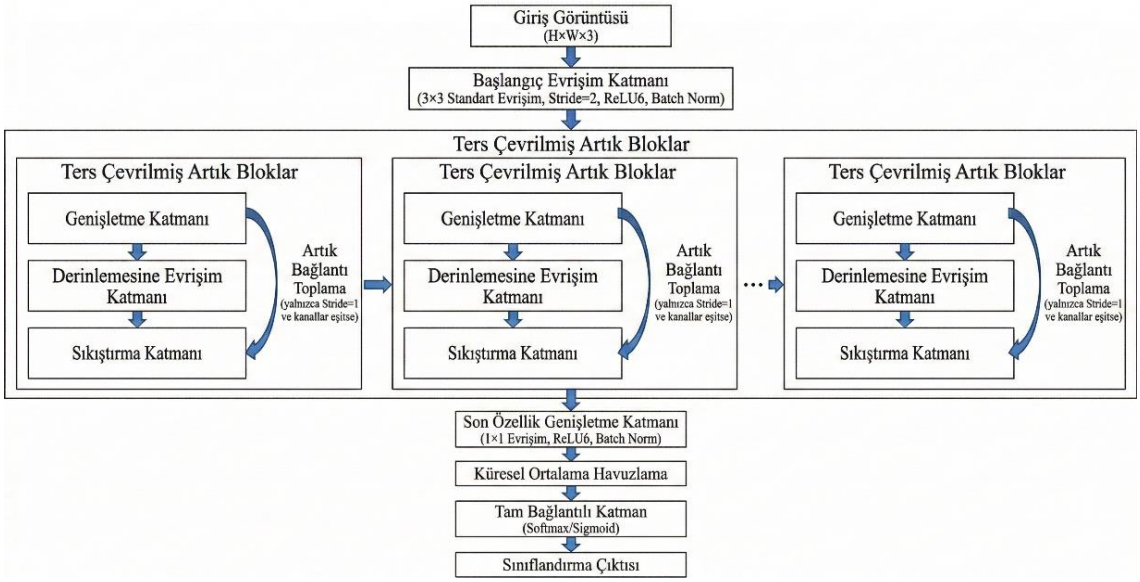
Denklem 3.9'da $\sigma(z_i)$ ifadesi i. sınıf için hesaplanan olasılığını göstermektedir. z_i ilgili sınıfa ait doğrusal çıktıyı ve $\sum_{j=1}^{K} e^{z_j}$ tüm sınıflara ait üstel değerlerin toplamını temsil etmektedir. Bu işlem sayesinde çıktı değerleri 0 ile 1 arasında normalize edilerek olasılık dağılımı hâline getirilmektedir.

Tam bağlı katmanlar uzamsal bilgiyi doğrudan korumamakla birlikte sınıflandırma için gerekli olan küresel temsillerin üretilmesinde kritik bir rol oynamaktadır. Bu katmanlarda aşırı öğrenmeyi önlemek amacıyla unutma gibi düzenleme yöntemleri sıklıkla uygulanmaktadır. Bununla birlikte modern CNN mimarilerinde tam bağlı katmanların sayısı azaltılarak parametre verimliliği artırılmakta ve model karmaşıklığı düşürülmektedir. Girdi görüntüsü $X \in \mathbb{R}^{H \times W \times C}$ için bu katmanlar zinciri ham veriyi soyut temsillere dönüştürmektedir. Evrişimsel sinir ağlarında katman sayısının artırılarak mimarinin derinleştirilmesi ağın temsil kapasitesinin artırılmasıyla birlikte optimizasyon güçlükleri ve gradyan sönmesi gibi sorunları daha belirgin hale gelmiştir.

3.2 MobileNetV2 Mimarisi

Bu çalışmada, mobil ve gömülü sistemlerde gerçek zamanlı çalışabilme gereksinimi göz önünde bulundurularak Google tarafından geliştirilen MobileNetV2 mimarisi tercih edilmiştir. MobileNetV2 mimarisi hesaplama maliyetini minimize ederken model başarımını en üst düzeyde tutmayı hedeflemektedir. Bu mimari, tersine artık bloklar ve

doğrusal darboğazlar üzerine kurulu yenilikçi bir yapıya sahiptir. Bu yapı özellikle düşük boyutlu temsil uzaylarında bilgi kaybını önlerken parametre ve hesaplama verimliliğini önemli ölçüde artırmaktadır. Toplamda 17 temel bloktan oluşan mimari, başlangıç aşamasında 32 filtre ile evrişim işlemini gerçekleştirmektedir. İlerleyen katmanlarda kanal sayısı kademeli olarak artırılmaktadır. Öznitelik çıkarma sürecinin ardından küresel ortalama havuzlama katmanı uygulanmaktadır. Sınıflandırma işlemi ise tam bağlı katman aracılığı ile gerçekleştirilmektedir. MobileNetV2 mimarisinin blok diyagramı Şekil 3.2'de sunulmuştur.



Şekil 3.2 MobileNetV2 mimarisinin blok diyagramı

3.2.1 Derinlemesine Ayrılabilir Evrişim

MobileNet mimarisinin temel bileşeni olan derinlemesine ayrılabilir evrişim katmanı, hesaplama maliyetini ve parametre sayısını azaltmak amacıyla geleneksel evrişim işlemini derinlemesine evrişim ve noktasal evrişim olmak üzere iki ardışık aşamaya ayırmaktadır. İlk aşama olan derinlemesine evrişim işleminde, standart evrişimden farklı olarak her bir giriş kanalı için yalnızca kendisine ait ve bağımsız bir filtre uygulanmaktadır. Böylece kanallar arasında herhangi bir etkileşim olmaksızın uzamsal öznitelikler kanal bazında çıkarılmaktadır. Standart evrişim işlemlerinde her çıktı kanalı tüm giriş kanalları üzerinden hesaplama yaparken derinlemesine evrişimde her giriş

kanalı yalnızca kendisine ait tek bir filtre ile işlenmektedir. Bu yaklaşım hesaplama maliyetini önemli ölçüde azaltmaktadır. Derinlemesine evrişim işleminin hesaplama maliyeti kayan nokta işlem sayısı cinsinden Denklem 3.10'da ifade edilmektedir (Sandler vd. 2018).

$$FLOP_{s_{dw}} = k^2 \cdot C_{in} \cdot H \cdot W \tag{3.10}$$

Denklem 3.10'da k filtre boyutunu, C_{in} giriş kanal sayısını, H ve W ise girişin uzamsal boyutlarını temsil etmektedir.

İkinci aşama olan noktasal evrişim işleminde ise derinlemesine evrişimden elde edilen çıktılar 1×1 boyutunda filtreler kullanılarak kanallar arası doğrusal kombinasyona dönüştürülmektedir. Bu aşamada çıktı kanal sayısı C_{out} olarak belirlenmekte ve kanal boyutunun kontrolü sağlanmaktadır. Noktasal evrişim işlemi girdinin uzamsal boyutları değiştirmeden kanal boyutunu artırma veya azaltma işlevi görmektedir. Noktasal evrişim işleminin hesaplama maliyeti Denklem 3.11'de verilmektedir (Wu vd. 2019).

$$FLOP_{s_{pw}} = C_{in} \cdot C_{out} \cdot H \cdot W \tag{3.11}$$

Denklem 3.11'de C_{out} hedeflenen çıktı kanal sayısını göstermektedir. $H \cdot W$ terimi ise girdinin uzamsal boyutlarını temsil etmektedir. Bu aşama modelin kanal boyutunu kontrol ettiği ve bilgi entegrasyonunu sistematik biçimde gerçekleştirdiği süreci ifade etmektedir. Sonuç olarak derinlemesine ayrılabilir evrişim işlemi hem kanal bazlı uzamsal öznitelik çıkarımı hem de kanallar arası doğrusal kombinasyon ile hesaplama verimliliğini standart evrişime kıyasla belirgin şekilde iyileştirmektedir.

3.2.2 Tersine Artık Yapı ve Doğrusal Darboğazlar

Geleneksel evrişimli sinir ağlarında artık bloklar geniş katmanları birbirine bağlarken darboğaz katmanları aracılığıyla kanal sayısı önce azaltılıp sonra tekrar artırılarak bir geniş - dar - geniş yapısı izlenmektedir. MobileNetV2 mimarisinde ise bunun tersine bir yaklaşım benimsenmiş ve tersine artık yapı önerilmiştir. Bu yapıda düşük boyutlu giriş

tensörü önce genişletilmekte ardından derinlemesine ayrılabilir evrişim uygulanmakta ve son olarak tekrar düşük boyuta indirgenmektedir. Bu dönüşüm süreci ve artık bağlantı mekanizması Denklem 3.12 ile ifade edilmektedir (Sandler vd. 2018).

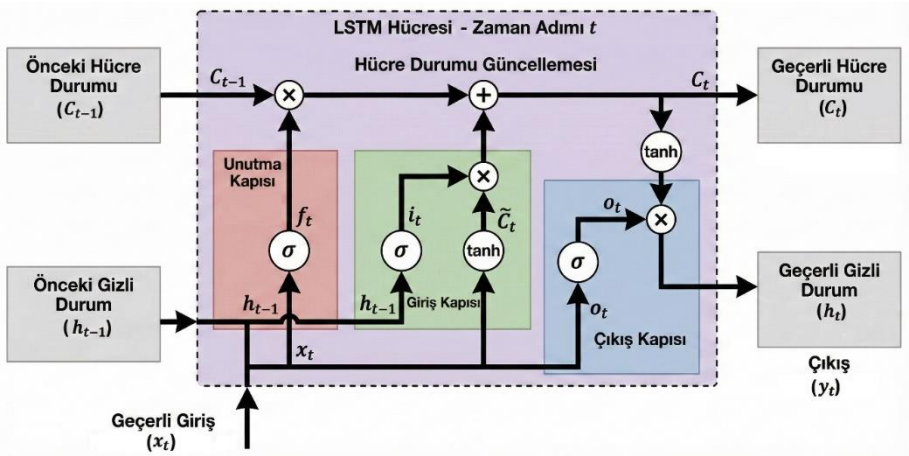
$$y = [\mathcal{P} \circ \mathcal{D} \circ \mathcal{E}](x) + x \tag{3.12}$$

Denklem 3.12'de x giriş tensörünü, y ise işlem sonucunda elde edilen çıkış tensörünü temsil etmektedir. Blok içerisindeki dönüşüm süreci fonksiyonel bileşke ile ifade edilen üç aşamalı bir katman dizisinden oluşmaktadır. İlk aşamada genişletme katmanı $\mathcal{E}$ devreye girerek 1×1 evrişim yoluyla giriş tensörünü daha yüksek boyutlu bir öznitelik uzayına taşımaktadır. Ardından derinlemesine evrişim katmanı $\mathcal{D}$, mekânsal filtreleme işlemini gerçekleştirmektedir. Son aşamada ise projeksiyon katmanı $\mathcal{P}$, kanal sayısını tekrar azaltarak veriyi doğrusal bir darboğaz yapısına indirgemektedir. Denklemdeki $+ x$ terimi giriş ve çıkış boyutlarının eşleşmesi durumunda devreye giren artık bağlantıyı ifade etmektedir. Bu mekanizma ham giriş verisinin doğrudan çıkışa eklenmesini sağlayarak derin ağlarda gradyan akışını iyileştirmekte ve eğitim sürecini stabilize etmektedir. Özetle bu mimari yaklaşım genişletme katmanlarında çok daha zengin öznitelik temsillerinin çıkarılmasını sağlamaktadır. Aynı zamanda darboğaz katmanlarında verinin kompakt ve hesaplama açısından verimli bir biçimde temsil edilmesini güvence altına almaktadır.

Modelin performansını ve bilgi kapasitesini korumak amacıyla projeksiyon katmanının sonunda ReLU yerine doğrusal aktivasyon kullanılmıştır. ReLU fonksiyonu negatif değerleri sıfıra eşitlediği için düşük boyutlu darboğaz katmanlarında önemli bilgi kayıplarına yol açabilmektedir. Doğrusal aktivasyon kullanımı bu bilgi kaybını minimize ederek sinyal iletimini optimize etmektedir. Bu mimarinin en kritik özelliği düşük boyutlu darboğaz katmanlarında bilgi kaybını önlemek amacıyla doğrusal aktivasyon kullanılmasıdır. Projeksiyon aşamasında ReLU gibi doğrusal olmayan fonksiyonların kullanılmaması sayesinde kanal boyutu daraltılırken verinin yapısal bütünlüğünü oluşturan temel bilgi korunmakta ve sinyal iletimi optimize edilmektedir.

3.3 Uzun Kısa Süreli Bellek Ağları

Yüz canlılık tespiti ve rPPG sinyal analizi gibi zaman serisi problemlerinde veriler arasındaki zamansal bağımlılıkların modellenmesi kritik öneme sahiptir. Geleneksel tekrarlayan sinir ağları uzun dizilerde ortaya çıkan gradyan kaybolması problemi nedeniyle geçmiş bilgiyi taşımakta yetersiz kalmaktadır. Bu sorunu çözmek amacıyla Hochreiter ve Schmidhuber tarafından uzun kısa süreli bellek (Long Short-Term Memory, LSTM) mimarisi önerilmiştir (Hochreiter ve Schmidhuber 1997). LSTM, zaman adımları boyunca sıralı verileri işlemektedir. Ağ, her bir zaman adımında o anki girdiyi (x_t), bir önceki adımın gizli durumunu (h_{t-1}) ve önceki hücre durumunu (C_{t-1}) kullanarak kendi iç durumunu güncellenmektedir. LSTM mimarisinin temel işleyişi Şekil 3.3'te yer almaktadır.



Şekil 3.3 LSTM temel işleyiş şeması

LSTM mimarisi, zaman serileri üzerinde uzun dönemli bağımlılıkları etkili biçimde modelleyebilmek amacıyla tasarlanmış özel bir yinelemeli sinir ağı yapısıdır. LSTM'in temelini bilgiyi uzun süreler boyunca taşıyabilen bir hücre durumu ile bu bilginin akışını denetleyen kapı mekanizmaları oluşturmaktadır. Hücre durumu konveyör bant benzeri doğrusal bir yol üzerinden zaman adımları boyunca bilgiyi iletirken kapılar sigmoid aktivasyon fonksiyonu aracılığıyla 0 ile 1 arasında değerler üreterek hangi bilginin korunacağına, güncelleneceğine veya silineceğine karar vermektedir. Bir LSTM hücresi unutma kapısı, giriş kapısı ve çıkış kapısı olmak üzere üç temel kapıdan oluşmaktadır.

Unutma kapısı önceki zaman adımına ait hücre durumunda yer alan bilginin ne kadarının korunacağına karar veren mekanizmadır. Bu kapı bir önceki gizli durum vektörü ile mevcut zaman adımındaki giriş vektörünün birleştirilmesiyle elde edilen girdiyi sigmoid aktivasyon fonksiyonundan geçirerek 0 ile 1 arasında bir aktivasyon değeri üretmektedir. Üretilen bu aktivasyon değerinin önceki hücre durumu ile eleman bazlı olarak çarpılmasıyla geçmişe ait hangi bilgilerin korunacağı ve hangilerinin unutulacağı belirlenmektedir. Unutma kapısının matematiksel ifadesi Denklem 3.14'te verilmiştir (Gers vd. 2000).

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{3.14}$$

Bu denklemde f_t ilgili zaman adımındaki unutma kapısı aktivasyon değerini ifade etmektedir. Ayrıca σ sigmoid aktivasyon fonksiyonunu, W_f unutma kapısına ait öğrenilebilir ağırlık matrisini, h_{t-1} bir önceki zaman adımının gizli durum vektörünü, x_t mevcut giriş vektörünü ve b_f ise unutma kapısına ait sapma terimini temsil etmektedir.

Giriş kapısı hücre durumuna hangi yeni bilginin ekleneceğini belirleyen mekanizmadır ve iki aşamalı bir hesaplama süreci içermektedir. İlk aşamada giriş kapısı sigmoid aktivasyon fonksiyonu aracılığıyla hücre durumunun hangi bileşenlerinin güncelleneceğini belirleyen bir maske üretmektedir. İkinci aşamada ise hiperbolik tanjant fonksiyonu kullanılarak hücre durumuna eklenebilecek yeni aday değerler vektörü oluşturulmaktadır. Bu işlemler sırasıyla Denklem 3.15 ve Denklem 3.16 ile ifade edilmektedir (Hochreiter ve Schmidhuber 1997).

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3.15}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3.16}$$

Bu denklemlerde i_t giriş kapısının aktivasyon değerini ve dolayısıyla güncellenecek bilginin seçimini ifade etmektedir. $\tilde{C}_t$ hücre durumuna eklenmesi aday olan yeni bilgileri içeren vektördür. W_i ve W_C sırasıyla giriş kapısı ve aday değer katmanlarına ait ağırlık matrislerini b_i ve b_C ise bu katmanlara ait sapma terimlerini göstermektedir. Hiperbolik

tanjant fonksiyonu aday değerleri −1 ile 1 aralığına sıkıştırarak sayısal kararlılığı sağlamaktadır. Yeni hücre durumu C_t önceki hücre durumunun unutma kapısı aracılığıyla ölçeklenmesi ve giriş kapısından geçen yeni bilginin eklenmesiyle güncellenmektedir. Bu güncelleme işlemi Denklem 3.16'da gösterilmiştir (Gers vd. 2000).

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{3.17}$$

Denklem 3.17'de $*$ işlemi literatürde Hadamard çarpımı olarak da bilinen eleman bazlı çarpım işlemini ifade etmektedir. Denklemde yer alan bu toplamsal yapı, LSTM mimarisinin temel avantajlarından birini oluşturmaktadır. Hücre durumu üzerinden geçen bu doğrusal bilgi yolu, gradyanın zaman adımları boyunca sönümlenmeden geriye doğru akmasına olanak tanımaktadır. Geleneksel yinelemeli sinir ağlarında ardışık matris çarpımları nedeniyle ortaya çıkan üstel gradyan sönümlenmesi problemi LSTM mimarisindeki bu yapı sayesinde büyük ölçüde ortadan kaldırılmaktadır. Bu ise ağın uzun zaman aralıklarındaki bağımlılıkları kararlı bir şekilde öğrenmesi mümkün hâle getirmektedir.

Çıkış kapısı güncellenmiş hücre durumuna dayanarak mevcut zaman adımındaki gizli durum vektörünün hesaplanmasını sağlamaktadır. Elde edilen bu gizli durum hem bir sonraki zaman adımına aktarılmakta hem de sınıflandırma veya regresyon gibi çıktı katmanlarına girdi olarak sunulmaktadır. Çıkış kapısının aktivasyon değeri sigmoid fonksiyonu ile hesaplanırken hücre durumu hiperbolik tanjant fonksiyonundan geçirilerek bu aktivasyon değeri ile eleman bazlı olarak çarpılmaktadır. Bu işlemler sırasıyla Denklem 3.18 ve Denklem 3.19'da sunulmuştur (Hochreiter ve Schmidhuber 1997).

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{3.18}$$

$$h_t = o_t * \tanh(C_t) \tag{3.19}$$

Bu denklemlerde o_t çıkış kapısının aktivasyonunu, h_t mevcut zaman adımına ait gizli durum vektörünü, W_o çıkış kapısına ait ağırlık matrisini ve b_o sapma terimini temsil etmektedir. Bu kapı mekanizmaları ağ içerisindeki gradyan akışını dengeli hâle getirerek öğrenme sürecinin kararlılığını artırmaktadır.

Bu çalışmada, LSTM mimarisi MobileNetV2 ağından elde edilen görsel özniteliklerin zamansal sürekliliğini modellemek ve eş zamanlı olarak uzaktan fotopletismografi temelli sinyallerdeki periyodik nabız dalgalanmalarını temsil edebilmek amacıyla iki ayrı akışta kullanılmıştır. Bu sayede hem uzamsal-zamansal görsel bilgilerin hem de fizyolojik sinyal örüntülerinin birlikte değerlendirilerek daha güçlü ve ayırt edici bir öznitelik temsilinin elde edilmesi hedeflenmiştir.

3.4 Üç Boyutlu Evrişimsel Ağlar

Üç boyutlu evrişimsel sinir ağları (3D Convolutional Neural Network, 3D-CNN) geleneksel iki boyutlu evrişimsel sinir ağlarının video verilerine uyarlanmış bir uzantısıdır. Bu mimari, uzamsal (x, y) ve zamansal (t) boyutları eş zamanlı olarak modelleyebilmektedir. İki boyutlu evrişimsel ağlarda filtreler yalnızca görüntü düzlemi üzerinde hareket ederek tekil karelerden uzamsal öznitelikler çıkarmaktadır. Buna karşılık üç boyutlu evrişimsel ağlarda filtreler zaman ekseni boyunca da kaydırılmaktadır. Bu sayede ardışık kareler arasındaki zamansal ilişkiler doğrudan öğrenilebilmektedir. Söz konusu yapı özellikle video tabanlı sahtekârlık tespiti uygulamalarında önemli bir avantaj sağlamaktadır. Bu mimari ile yüz hareketleri, parlaklık değişimleri ve çerçeveler arası zamansal doku varyasyonları başarılı bir şekilde modellenebilmektedir.

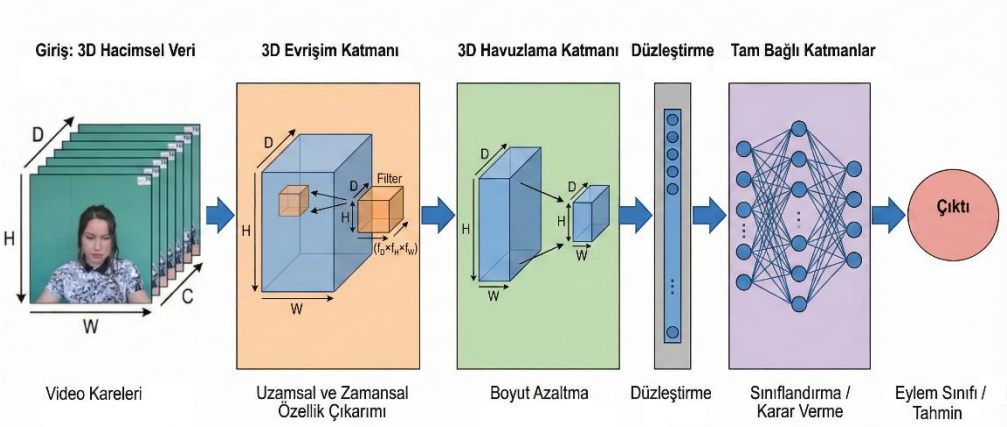
Üç boyutlu evrişim işlemi video verisini bir hacim olarak ele almakta ve bu hacim üzerinde üç eksenli öznitelik çıkarımı gerçekleştirmektedir. Böylece ağ yalnızca tek bir kareye ait uzamsal bilgileri işlemenin yanı sıra zaman içerisinde meydana gelen dinamik değişimleri de temsil edebilmektedir. Bu işlem Denklem 3.20'de verilen matematiksel ifade ile tanımlanmaktadır (Ji vd. 2013). 3D-CNN mimarisinin temel işleyişi ise Şekil 3.4'te gösterilmektedir.

$$v_{ij}^{xyz} = \tanh\left(b_{ij} + \sum_m \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} \sum_{r=0}^{R_i-1} w_{ijm}^{pqr}\, v_{(i-1)m}^{(x+p)(y+q)(z+r)}\right) \qquad (3.20)$$

Denklem 3.20'de v_{ij}^{xyz}, i. katmandaki j. özellik haritasının (x, y, z) koordinatındaki aktivasyon değerini ifade etmektedir. Bu değer bir önceki katmandaki tüm özellik haritaları üzerinden yapılan evrişimsel işlemlerin sonucunda elde edilmektedir.

Denklemde yer alan b_{ij} ilgili özellik haritasına ait sapma terimini temsil ederken m indeksi bir önceki katmandaki özellik haritalarını göstermektedir. Filtrenin uzamsal boyutları P_i ve Q_i ile ifade edilmektedir. R_i parametresi filtrenin zamansal derinliğini yani kapsadığı çerçeve sayısını belirtmektedir. w_{ijm}^{pqr} ise filtrenin (p, q, r) konumundaki öğrenilebilir ağırlık değerini temsil etmektedir. Elde edilen toplam, doğrusal olmayan bir dönüşüm sağlamak amacıyla hiperbolik tanjant aktivasyon fonksiyonundan geçirilerek nihai çıktı oluşturulmaktadır.

Bu yapı sayesinde 3D-CNN'ler video dizilerindeki kısa ve orta vadeli zamansal bağımlılıkları doğrudan evrişimsel filtreler aracılığıyla öğrenebilmekte ve böylece hareket temelli ipuçlarını etkili biçimde yakalayabilmektedir. Özellikle yüz tabanlı biyometrik güvenlik sistemlerinde canlı yüz hareketleri ile sunum saldırıları arasındaki ince zamansal farkların ayırt edilmesinde üç boyutlu evrişimsel ağlar güçlü bir temsil yeteneği sunmaktadır.



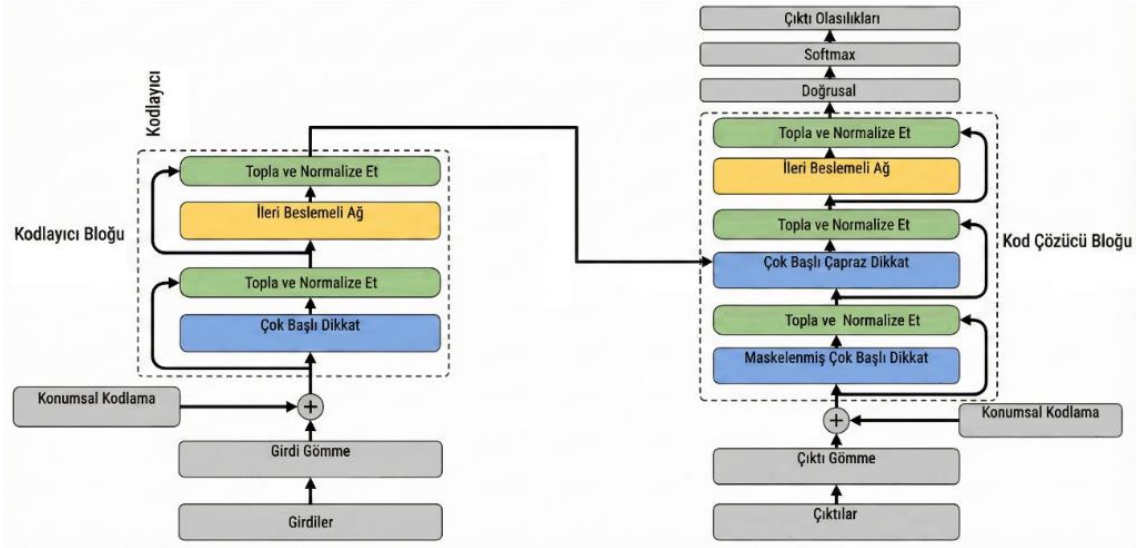
Şekil 3.4 3D-CNN temel işleyiş şeması

3.5 Transformer Tabanlı Mimariler

Son yıllarda transformer tabanlı mimariler özellikle çok başlı öz-dikkat mekanizması sayesinde hem görüntü işleme hem de zaman serisi analizi alanlarında güçlü bir alternatif olarak öne çıkmaktadır. Bu mimariler ardışık verilerde uzun mesafeli bağımlılıkları doğrudan modelleyebilme yeteneğine sahiptir. Bu sayede geleneksel evrişimsel veya yinelemeli yapılara kıyasla daha esnek bir temsil sunmaktadır. Bu esneklik özellikle uzaktan fotopletismografi tabanlı uygulamalarda önemli bir avantaj sağlamaktadır. rPPG

sinyalleri genellikle düşük genlikli ve zamana yayılmış nabız dalgalanmalarından oluşmaktadır. Ayrıca sunum saldırılarında yüz dokusunda ince sahtecilik artefaktları meydana gelmektedir. Mimaride yer alan öz-dikkat mekanizması sayesinde hem bu zayıf fizyolojik sinyaller hem de dokusal bozulmalar çok daha belirgin bir biçimde yakalanabilmektedir.

Transformer mimarilerinin temel avantajlarından biri evrişimsel katmanların doğası gereği sahip olduğu yerel algı alanı sınırlamasını aşarak tüm giriş dizisi üzerinde küresel bağlamı dikkate alabilmesidir. Bu sayede video çerçeveleri arasındaki uzun vadeli tutarsızlıklar, zamansal kopukluklar ve doğal olmayan desenler daha etkili bir şekilde tespit edilebilmektedir. Bu özellik yüz tabanlı biyometrik sahtekârlık tespiti gibi hem uzamsal hem de zamansal ipuçlarının kritik olduğu uygulamalarda önemli bir üstünlük sağlamaktadır. Transformer tabanlı mimarilerin genel işleyişi Şekil 3.5'te şematik olarak sunulmaktadır.



Şekil 3.5 Transformer tabanlı mimarilerin temel işleyiş şeması

Bir öz-dikkat katmanının temel çalışma prensibi her bir giriş öğesinin dizideki diğer tüm öğelerle olan ilişkisini ölçerek bağlamsal olarak en anlamlı bilgiyi ağırlıklandırılmış biçimde birleştirmeye dayanmaktadır. Bu işlem sorgu (Q), anahtar (K) ve değer (V) olarak adlandırılan üç farklı temsilden yararlanılarak gerçekleştirilmektedir. Öz-dikkat mekanizmasının matematiksel ifadesi Denklem 3.21'de verilmiştir (Vaswani vd. 2017).

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{3.21}$$

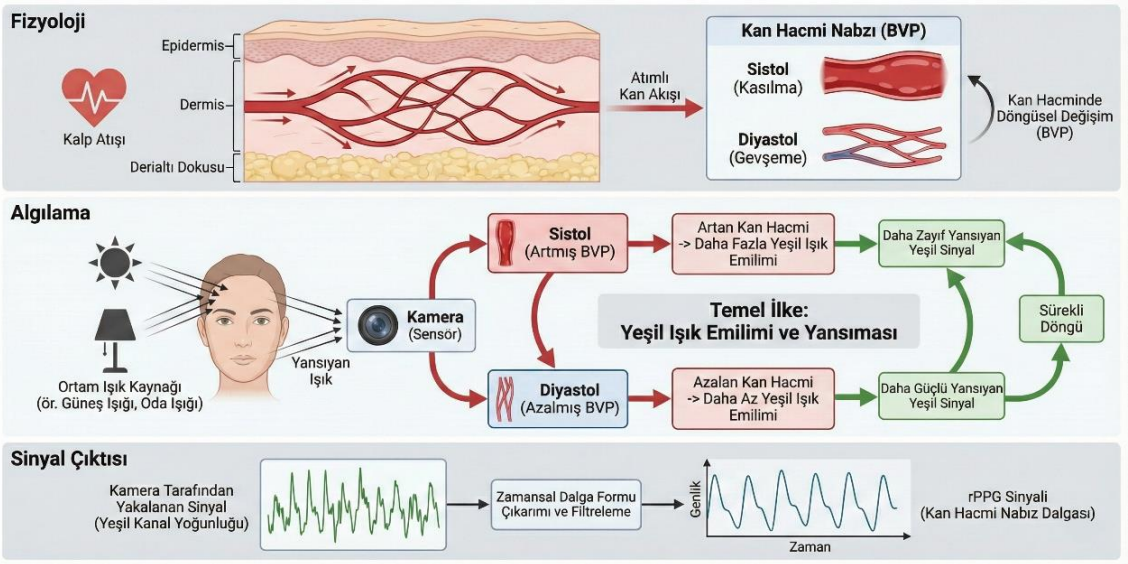
Bu denklemde Q her bir öğenin dizideki diğer öğelerle olan ilişkisini sorgulayan sorgu matrisini temsil etmektedir. K dizideki öğelerin ayırt edici özelliklerini içeren anahtar matrisini, V ise bu ilişkiler doğrultusunda aktarılacak asıl bilgiyi barındıran değer matrisini ifade etmektedir. Q ve K matrislerinin çarpımı ile elde edilen benzerlik skorları anahtar matrisinin boyutu olan d_k ile ölçeklendirilerek sayısal kararlılık sağlanmakta ve ardından softmax fonksiyonu aracılığıyla 0 ile 1 aralığında normalize edilmektedir. Elde edilen bu dikkat ağırlıkları değer matrisi ile çarpılarak bağlamsal olarak zenginleştirilmiş çıktı temsili oluşturulmaktadır. Bu yapı sayesinde transformer tabanlı modeller, video dizilerinde veya fizyolojik zaman serilerinde hem kısa hem de uzun vadeli bağımlılıkları eş zamanlı olarak öğrenebilmektedir. Ayrıca çerçeveler arası doğal olmayan geçişleri ve sahtecilik kaynaklı zamansal tutarsızlıkları da yüksek doğrulukla ayırt edebilmektedir. Bu nedenle transformer mimarileri modern biyometrik güvenlik ve canlılık tespiti sistemlerinde giderek daha fazla tercih edilen bir yaklaşım hâline gelmiştir.

4. UZAKTAN FOTOPLETİSMOGRAFİ

Fotopletismografi doku içerisindeki kan hacmi değişimlerini (Blood Volume Pulse, BVP) optik yöntemlerle ölçen bir tekniktir. Fotopletismografi sinyali kalbin kasılma ve gevşeme evrelerinde meydana gelen kan hacmi değişimlerine bağlı olarak oluşmaktadır. Bu süreçte damardaki kan miktarının periyodik olarak artıp azalması dokudan geri yansıyan veya dokudan geçen ışığın yoğunluğunda ölçülebilir dalgalanmalar yaratmaktadır (Allen 2007). Bu değişim oksi-hemoglobin ve deoksi-hemoglobinin farklı dalga boylarındaki ışığı soğurma özelliklerinden kaynaklanmaktadır. Hemoglobin molekülleri özellikle 520-580 nm spektrum aralığındaki yeşil ışığı yüksek oranda absorbe etmektedir. Bu güçlü emilim özelliği sayesinde yeşil dalga boyundan elde edilen sinyaller dokudaki mikroskobik kan hacmi değişimlerini diğer dalga boylarına kıyasla daha belirgin bir nabız dalgası olarak yansıtmaktadır. Kırmızı ve mavi dalga boylarında da absorpsiyon farkları bulunmaktadır. Ancak bu dalga boylarında nabız bileşeni genellikle daha zayıf gözlenmektedir. Geleneksel fotopletismografi sensörleri söz konusu değişimleri ölçebilmek için ışık kaynağının ve dedektörün cilde doğrudan fiziksel temasını gerektirmektedir. Ancak teknolojinin gelişimiyle birlikte bu ölçümün temassız bir şekilde yapılması mümkün hale gelmiştir. Uzaktan fotopletismografi (Remote Photoplethysmography, rPPG) olarak adlandırılan bu yöntem standart bir video kamera kullanılarak ortam ışığı altındaki yüzey yansımalarının analiz edilmesine dayanmaktadır (Verkruysse vd. 2008). Bu optik yaklaşım sayesinde kardiyovasküler sinyallerin tamamen temassız, donanımdan bağımsız ve esnek bir şekilde elde edilmesi mümkün hâle gelmiştir.

Uzaktan fotopletismografinin temel çalışma prensibi, kalp atışlarıyla senkronize olarak değişen kan hacminin dokudaki optik soğurma ve saçılma özelliklerinde yarattığı varyasyonlara dayanmaktadır. Kan hacmindeki artış özellikle yeşil dalga boyundaki ışığın emilimini artırırken geri yansıyan ışık miktarında mikroskobik azalmalara neden olmaktadır. İnsan gözüyle algılanamayan bu renk değişimleri video karelerindeki piksel yoğunluklarının zaman serisi analizi ile tespit edilebilmektedir. Uzaktan fotopletismografi yüzey yansımalarından fizyolojik sinyallerin çıkarılmasına dayalı non-invaziv bir yaklaşımdır. Bu özelliği sayesinde yüz canlılık tespiti gibi uygulamalarda

kritik rol oynamaktadır. rPPG sinyalinin çıkarım süreci Şekil 4.1'de şematik olarak özetlenmiştir. Uzaktan fotopletismografinin gelişimi optik sinyal işleme ve makine öğrenmesi algoritmalarının entegrasyonuyla büyük bir ivme kazanmıştır. Ancak kontrolsüz ortamlarda sinyal kalitesi düşebilmektedir. Bu nedenle güncel araştırmalar özellikle hareket artefaktları ve aydınlatma varyasyonları gibi dışsal bozuculara karşı daha dayanıklı yöntemlerin geliştirilmesi üzerine odaklanmıştır.



Şekil 4.1 rPPG sinyalinin çıkarım süreci

Video sekanslarından nabız sinyali elde etmek için literatürde çeşitli algoritmalar geliştirilmiştir. Bu yöntemler genellikle hareket gürültüsü ve aydınlatma değişimleri gibi bozucu etkenlere karşı sağladıkları dayanıklılık seviyelerine göre kategorize edilmektedir. Literatürdeki geleneksel yaklaşımlar büyük ölçüde kör kaynak ayrıştırma tekniklerine ve fizyolojik modellere dayanmaktadır. Ancak son yıllarda kaydedilen ilerlemelerle birlikte bu yöntemlerin yerini yüksek temsil yeteneğine sahip derin öğrenme tabanlı mimariler almaya başlamıştır (Chen vd. 2024). Bu bağlamda literatürde öne çıkan ve rPPG sinyal çıkarımında standart kabul edilen temel yöntemler aşağıda detaylandırılmıştır.

4.1 rPPG Sinyal Çıkarım Yöntemleri ve Algoritmalar

Bu çalışmada, fizyolojik sinyalin kalitesini ve elde edilen sinyalin derin öğrenme tabanlı modeller üzerindeki etkisini analiz etmek amacıyla literatürde yaygın olarak kabul gören yeşil kanal, CHROM ve POS rPPG algoritmaları uygulanmış ve karşılaştırmalı olarak değerlendirilmiştir.

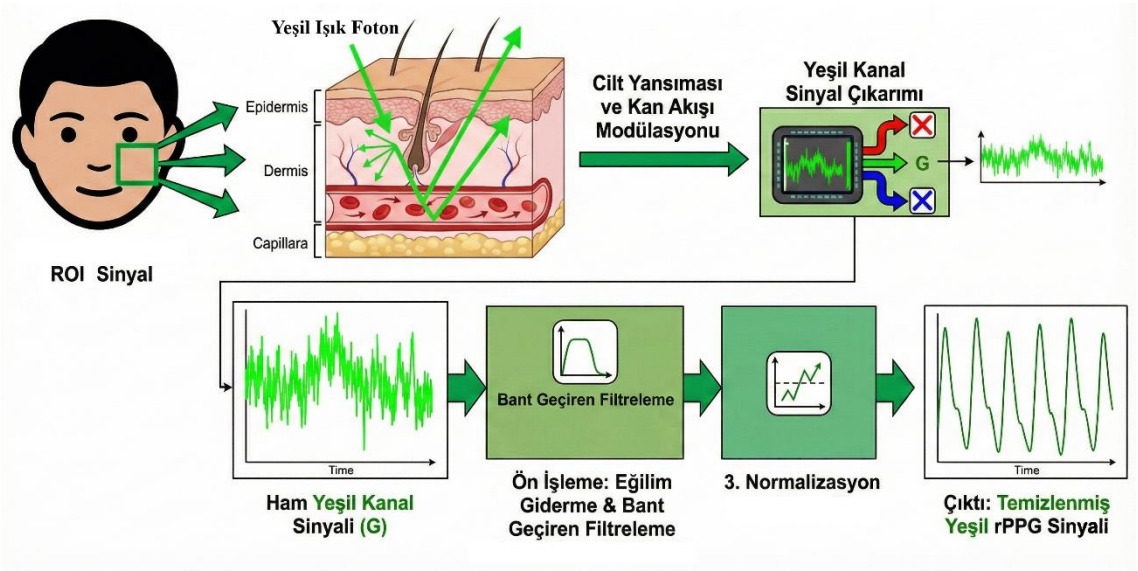
4.1.1 Yeşil Kanal Yöntemi

Literatürde yer alan en temel rPPG yaklaşımlarından biri olan Yeşil Kanal (Green) yöntemi, RGB video sinyallerinden yalnızca yeşil bileşenin analiz edilmesine dayanmaktadır. Bu yöntem hemoglobinin yaklaşık olarak 520 - 580 nm aralığındaki yeşil dalga boyunda maksimum absorpsiyon özelliğine sahip olması prensibine dayanmaktadır. Yöntem, bu biyofiziksel özellikten yararlanarak ciltteki kan hacmi değişimlerini en belirgin şekilde yeşil kanalın yoğunluk değişimleri üzerinden yakalamayı hedeflemektedir (Chen vd. 2024). Uygulama aşamasında öncelikle yüz bölgesinden belirlenen ilgi alanı içerisindeki pikseller işleme alınmaktadır. Uzamsal ortalama tekniği kullanılarak ilgi alanı içerisindeki tüm piksellere ait yeşil kanal değerlerinin ortalaması her bir video karesi için hesaplanmaktadır. Bu işlem sonucunda tek boyutlu bir zaman serisi elde edilmektedir. Bu zaman serisi ciltteki kan hacmi değişimlerine karşılık gelen ham fotopletismografi sinyalini diğer bir ifadeyle kan hacmi dalga formunu temsil etmektedir. t zaman adımındaki ham BVP sinyalinin matematiksel ifadesi Denklem 4.1'de gösterilmiştir.

$$S_{Yeşil\ Kanal}(t) = \frac{1}{N}\sum_{i=1}^{N} G_i\,(t) \tag{4. 1}$$

Bu denklemde $S_{Yeşil\ Kanal}(t)$, t anına karşılık gelen ortalama yeşil kanal yoğunluğunu ve dolayısıyla ham BVP sinyalini ifade etmektedir. N, seçilen ilgi alanı içerisinde yer alan toplam piksel sayısını temsil etmektedir. $G_i(t)$ ifadesi ise ROI içerisindeki $i.$ pikselin ilgili zaman adımındaki yeşil kanal değerini göstermektedir. Toplam sembolü ilgi alanı içerisindeki tüm piksellerin yeşil kanal bileşenlerinin kümülatif olarak işlendiğini belirtmektedir.

Ontiveros ve arkadaşları tarafından web kamerası tabanlı temassız nabız ölçümü üzerine yapılan deneysel çalışmalarda RGB renk kanalları ayrı ayrı analiz edilmiştir. Çalışma sonucunda hemoglobin absorpsiyon spektrumuna en uygun kanalın yeşil kanal olduğu ve kalp atım hızı kestiriminde en yüksek doğruluğu bu kanalın sağladığı raporlanmıştır (Ontiveros vd. 2023). Yeşil Kanal yönteminin rPPG sinyal çıkarma sürecini detaylandıran akış şeması Şekil 4.2'de yer almaktadır.



Şekil 4.2 rPPG Yeşil Kanal yönteminin sinyal çıkarma akış şeması

Yeşil Kanal yöntemi, hesaplama karmaşıklığının düşük olması nedeniyle özellikle mobil uygulamalarda ve gerçek zamanlı sistemlerde tercih edilmektedir. Yöntem deneklerin hareketsiz olduğu ortamlarda yüksek SNR değerleri üretmektedir. Ancak hareket artefaktlarına ve ortamdaki ani aydınlatma değişimlerine karşı duyarlılığının yüksek olması daha karmaşık senaryolarda daha gelişmiş rPPG yöntemlerine geçişi zorunlu kılmaktadır.

4.1.2 Kör Kaynak Ayrıştırma Tabanlı Yöntemler

İlk nesil rPPG yöntemleri RGB kanallarındaki sinyallerin nabız, hareket, gürültü gibi bağımsız kaynakların doğrusal bir karışımı olduğu varsayımına dayanmaktadır. Bağımsız bileşen analizi ve temel bileşen analizi gibi istatistiksel yöntemler kullanılarak bu karışım içerisindeki en periyodik bileşenin nabız sinyali olduğu kabul edilmiştir (Poh vd. 2010,

2011). ICA sinyallerin istatistiksel bağımsızlıklarını maksimize ederek ayrıştırma yapmaktadır. PCA ise varyans bazlı projeksiyon uygulamaktadır. Her iki yöntem de FastICA veya JADE gibi algoritmalarla uygulanmaktadır (Debnath ve Kim 2025). Ancak bu yöntemler sinyal kaynaklarının istatistiksel olarak bağımsız olmasını gerektirdiği için hareketli ve değişken ışıklı ortamlarda kararsızlık gösterebilmektedir. Literatürde, BSS tabanlı yaklaşımların SNR değerlerinin hareket artefaktlarından etkilendiği rapor edilmiştir. Bu sınırlılık daha gelişmiş fizyolojik modellemelere geçişi tetiklemiştir (Maity vd. 2022).

4.1.3 Krominans Tabanlı Yöntem

Krominans Tabanlı Yöntem (CHROM) yöntemi, cilt yüzeyinden yansıyan ışığı speküler ve difüz saçılma bileşenleri üzerinden modelleyerek hareket gürültüsünü minimize etmeyi amaçlamaktadır. Bu yaklaşım standart RGB sensör verileri üzerinde doğrusal bir kombinasyon tanımlayarak renk farkı sinyalleri aracılığıyla nabız bileşenini ayrıştırmaktadır (de Haan ve Jeanne 2013). Yöntemin temel varsayımı hareket kaynaklı gürültünün tüm renk kanallarını benzer biçimde etkilemesine karşın, nabız bilgisinin farklı renk kanallarında ayırt edilebilir varyasyonlar oluşturduğu yönündedir. Algoritmanın temel işleyişi sırasıyla şu adımları takip etmektedir. İlk aşamada ham RGB kanal sinyalleri zamansal olarak normalize edilmektedir. Ardından cilt yüzeyindeki speküler yansıma etkisini elimine etmek amacıyla birbirinden bağımsız iki ortogonal krominans sinyali ($X\ ve\ Y$) oluşturulmaktadır. Bu sinyallerin hesaplanması Denklem 4.2 ve Denklem 4.3'te verilen doğrusal kombinasyonlar ile gerçekleştirilmektedir.

$$X = 3R_n - 2G_n \tag{4.2}$$

$$Y = 1.5R_n + G_n - 1.5B_n \tag{4.3}$$

Denklem 4.2 ve Denklem 4.3'te yer alan R_n, $G_n$ve B_n ifadeleri sırasıyla normalize edilmiş kırmızı, yeşil ve mavi renk kanallarını temsil etmektedir. R, G ve B ise ilgili zaman adımındaki ham piksel yoğunluk değerlerini ifade etmektedir. Elde edilen X bileşeni kırmızı ve yeşil kanallar arasındaki farkı vurgulayan birinci krominans sinyalini ifade

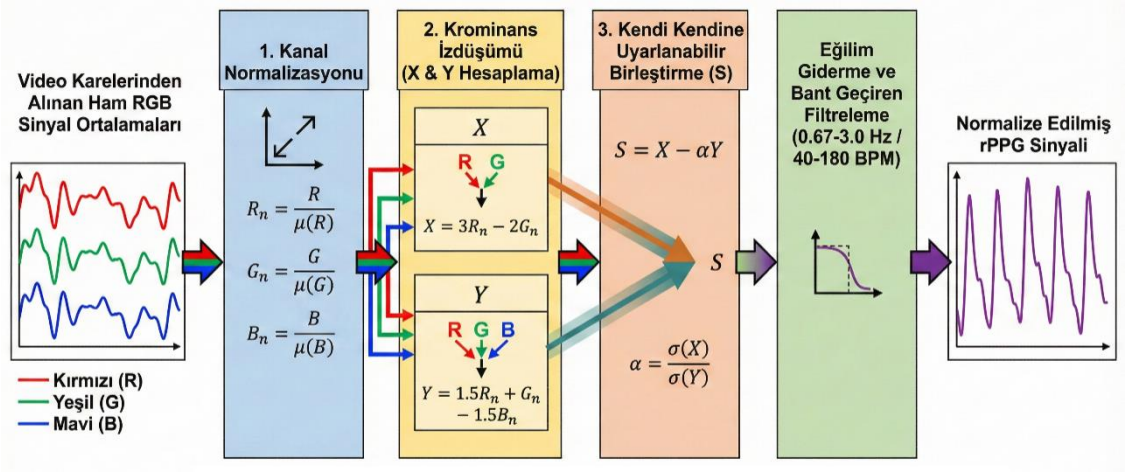
etmektedir. Y bileşeni ise kırmızı, yeşil ve mavi kanalların belirli katsayılarla birleştirilmesiyle elde edilen ve harekete karşı tamamlayıcı özellik taşıyan ikinci krominans sinyalini temsil etmektedir. Denklemlerde yer alan katsayılar ise cilt yansıma özellikleri ve speküler yansıma modellerine dayalı olarak türetilmiştir ve literatürde yaygın biçimde kullanılan sabit katsayılardır.

Nihai rPPG sinyali elde edilen bu iki krominans bileşeninin istatistiksel özellikleri dikkate alınarak oluşturulmaktadır. Bu amaçla X ve Y sinyallerinin standart sapmaları hesaplanmakta ve bu değerler aracılığıyla ölçeklendirme faktörü belirlenmektedir. Nihai kan hacmi dalga formu Denklem 4.4'te gösterildiği üzere X bileşeninden ölçeklendirilmiş Y bileşeninin çıkarılmasıyla elde edilmektedir.

$$S = X - \alpha Y \tag{4.4}$$

Bu denklemde S, hareket gürültüsünden büyük ölçüde arındırılmış nihai rPPG sinyalini diğer bir ifadeyle kan hacmi dalga formu sinyalini ifade etmektedir. α parametresi X ve Y sinyallerinin standart sapmalarının oranı olarak tanımlanan ölçekleme faktörünü ifade etmektedir. $\sigma(\cdot)$ işleci ise ilgili sinyalin standart sapmasını temsil etmektedir. Bu ağırlıklandırma işlemi hareket kaynaklı bozulmaların baskılanmasını ve nabız bileşeninin daha baskın hâle gelmesini sağlamaktadır.

Bu formülasyon özellikle yüzün hareket ettiği veya ışık açısının değiştiği durumlarda gürültü bileşenlerini elimine ederek yüksek Sinyal Kalite İndeksi'ne (SQI) sahip bir BVP sinyali üretmektedir (de Haan ve Jeanne 2013). Deneysel çalışmalarda CHROM yönteminin BSS temelli renk uzayı ayrıştırmalarına kıyasla daha yüksek kararlılık sergilediği ve hata oranlarında belirgin bir iyileşme sağladığı raporlanmıştır (Debnath ve Kim 2025). CHROM yönteminin rPPG sinyal çıkarma sürecini detaylı olarak gösteren akış şeması Şekil 4.3'te yer almaktadır.



Şekil 4.3 rPPG CHROM yönteminin sinyal çıkarma akış şeması

4.1.4 Düzlem-Ortogonal-Deri Yöntemi

Wang ve arkadaşları tarafından önerilen Düzlem-Ortogonal-Deri Yöntemi (Plane-Orthogonal-to-Skin, POS), rPPG sinyalini cilt tonu uzayında tanımlanan bir düzleme dik izdüşüm yaparak çıkarmayı hedeflemektedir (Wang vd. 2017a). Bu yöntem özellikle farklı cilt tonlarına sahip bireylerde aydınlatma değişimlerine karşı kararlı bir performans sunmaktadır (Wang vd. 2017a). Bu yöntem CHROM yöntemine benzer bir matematiksel alt yapıya sahip olmakla birlikte parametre ayarı gerektirmemesi ve zamansal filtrelemeye ihtiyaç duymadan anlık projeksiyon yapabilmesi ile öne çıkmaktadır. Algoritmada ilk olarak ilgi alanı içerisindeki kırmızı, yeşil ve mavi renk kanallarına ait zaman serileri normalize edilmektedir. Normalize edilmiş renk kanalları kullanılarak cilt tonu düzlemine dik olacak şekilde tanımlanan iki adet projeksiyon sinyali elde edilmektedir. Bu projeksiyon bileşenleri Denklem 4.5 ve Denklem 4.6 ile ifade edilmektedir.

$$X = G_n - B_n \tag{4.5}$$

$$Y = G_n + B_n - 2R_n \tag{4.6}$$

Denklem 4.5 ve 4.6'daki ifadelerde yer alan R_n, $G_n$ve B_n sırasıyla normalize edilmiş kırmızı, yeşil ve mavi renk kanalı değerlerini temsil etmektedir. X bileşeni, yeşil ve mavi

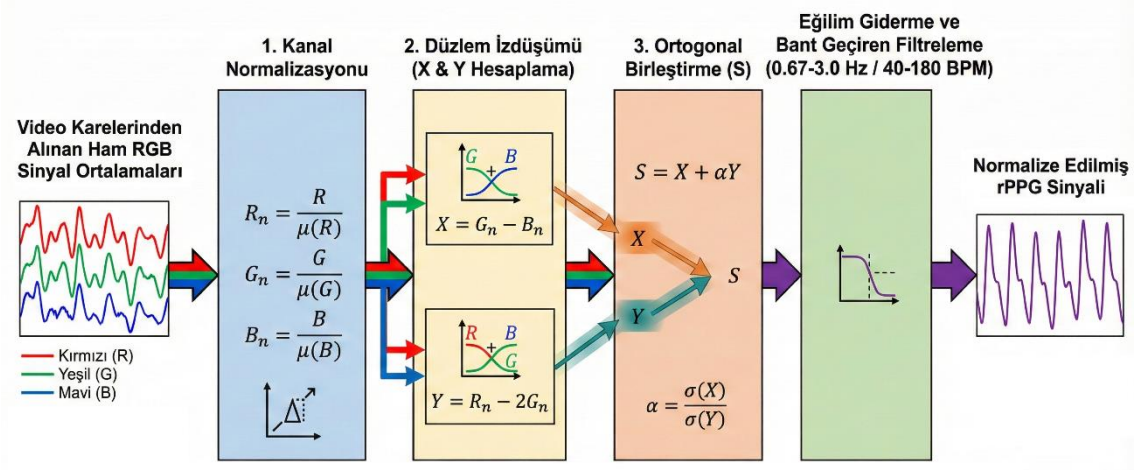
kanallar arasındaki farkı ifade eden birinci projeksiyon sinyalini tanımlamaktadır. Y bileşeni ise yeşil, mavi ve kırmızı kanalların ağırlıklı toplamı ile elde edilen ikinci projeksiyon sinyalini temsil etmektedir. Denklemde yer alan 2 katsayısı, kırmızı kanalın cilt yansıması üzerindeki baskın etkisini dengelemek amacıyla kullanılan matematiksel bir sabittir.

Nihai nabız sinyali, elde edilen bu iki projeksiyon bileşeninin adaptif biçimde birleştirilmesiyle hesaplanmaktadır. Bu birleşim işlemi Denklem 4.7'de gösterildiği üzere X bileşeninden ölçeklendirilmiş Y bileşeninin çıkarılmasıyla gerçekleştirilmektedir.

$$S = X - \alpha Y \tag{4.7}$$

Denklem 4.7'de S, hareket gürültüsünden büyük ölçüde arındırılmış nihai rPPG sinyalini, diğer bir ifadeyle kan hacmi dalga formu sinyalini ifade etmektedir. α parametresi, iki projeksiyon sinyali arasındaki ölçekleme faktörünü temsil etmektedir. X ve Y sinyallerinin standart sapmalarının oranı $\frac{\sigma(X)}{\sigma(Y)}$ olarak tanımlanmaktadır. $\sigma(\cdot)$ işleci ise ilgili sinyalin standart sapmasını ifade etmektedir. Bu ölçekleme sinyaller arasındaki varyans dengesini sağlayarak hareket ve aydınlatma kaynaklı bozulmaların etkisini azaltmaktadır.

POS yöntemi, sunduğu matematiksel sadelik ve karmaşık parametre ayarı gerektirmeyen yapısı ile dikkat çekmektedir. Özellikle farklı cilt tonları ve değişken aydınlatma koşullarında sergilediği yüksek kararlılık yöntemi diğer yaklaşımlardan ayıran en önemli faktördür. Sahip olduğu bu avantajlar, POS algoritmasının literatürde yaygın tercih edilen rPPG yöntemlerinden biri hâline gelmesini sağlamıştır. Yöntemin rPPG sinyal çıkarım sürecini ayrıntılı olarak gösteren akış şeması Şekil 4.4'te sunulmaktadır.



Şekil 4.4 rPPG POS yönteminin sinyal çıkarma akış şeması

Ancak yapılan deneysel çalışmalarda, CHROM yönteminin özellikle 3D maske saldırıları sırasında ortaya çıkan renk sapmalarını tespit etme noktasında yüksek bir ayırt ediciliğe sahip olduğu gözlemlenmiştir. Buna karşın POS yöntemi spektral varyasyonlara karşı gösterdiği direnç ile öne çıkmaktadır. Zhang ve diğerleri tarafından yapılan çalışmada SNR iyileştirmesi ve genel doğruluk performansı açısından POS yönteminin CHROM yöntemi ile benzer düzeyde sonuçlar ürettiği raporlanmıştır (Zhang vd. 2023).

4.2 rPPG Sistemlerinde İlgi Alanı Seçimi

İlgi alanı seçimi yüz görüntüsü tabanlı rPPG algoritmalarında verimli sinyal çıkarımı için kritik bir rol oynamaktadır. ROI seçimi, elde edilen fizyolojik sinyalin kalitesini ve sistemin hesaplama yükünü doğrudan etkilemektedir. Bu nedenle literatürde ROI belirleme stratejisi rPPG sistemlerinin genel başarımında en az sinyal çıkarma algoritması kadar belirleyici bir faktör olarak kabul edilmektedir (Kim vd. 2021). Deepfake tespiti amacıyla gerçekleştirilen kalp atış hızı ölçümlerinde de analiz için seçilen yüz bölgesinin performans üzerinde doğrudan etkili olduğu raporlanmıştır (Lee vd. 2025). Bu bağlamda ROI seçimi yalnızca görüntü üzerinden bir yüz alanı belirleme işlemi olarak değerlendirilmemelidir. Aksine bu süreç cilt kalınlığı gibi anatomik faktörlerin sistemin hesaplama verimliliğinin (Kim vd. 2021) ve elde edilen sinyal kalitesinin (Kossack vd. 2021) bir arada ele alındığı çok yönlü bir optimizasyon problemi niteliği taşımaktadır. Daha ince cilt katmanına sahip bölgelerin daha kaliteli rPPG sinyalleri sağladığı yönündeki bulgular ROI seçimi için doğrudan fizyolojik bir zemin oluşturmaktadır (Kim

vd. 2021). Bununla birlikte literatürdeki eğilim birden fazla ROI bölgesinin kullanımı, güvenilirlik metriklerinin entegrasyonu (Kossack vd. 2021) ve istatistiksel doğrulama yöntemlerinin benimsenmesi yönündedir (Lee vd. 2025). Bu gelişmeler basit bir tam yüz analizi yaklaşımlarından daha incelikli ve veri odaklı bir yaklaşıma geçişi işaret etmektedir. Sonuç olarak etkili bir ROI seçimi sürecin fizyolojik temellerinin, sinyal işleme tekniklerinin ve performans metriklerinin bütünsel bir bakış açısıyla değerlendirilmesini zorunlu kılmaktadır.

4.2.1 Farklı ROI Stratejilerinin Kullanımı

rPPG tabanlı sistemlerde ROI seçimi hem sinyal-gürültü oranını hem de nabız bileşeninin zamansal kararlılığını doğrudan etkileyen temel bir tasarım parametresidir. Literatürdeki erken dönem çalışmalarda genellikle tam yüzün ROI olarak kullanılması yaygın bir yaklaşım olmuştur. Ancak bu yaklaşım göz kırpma, ağız hareketi, konuşma ve yüz kas aktivitesinden kaynaklanan yüksek frekanslı parazitlere karşı duyarlıdır. Bu durum özellikle yüksek zamansal kararlılık gerektiren nabız çıkarım süreçlerinde ciddi doğruluk kayıplarına yol açmaktadır. Bu nedenle sonraki çalışmalarda ROI seçimi anatomik olarak daha stabil bölgelere indirgenmiştir. Özellikle alın ve yanak bölgelerinin daha homojen cilt dokusuna sahip olmaları ve daha düşük kas aktivitesi içermeleri nedeniyle sinyal kalitesi açısından üstün performans sunduğu gösterilmiştir. Yapılan analizlerde, bu bölgelerin hem AC bileşeninin daha güçlü olduğu hem de spektral bantta nabız tepe frekansını daha yüksek belirginlikle taşıdığı rapor edilmiştir. Buna karşın göz ve ağız çevresi gibi bölgeler hareket kaynaklı distorsiyonlar ve yansıma değişimleri nedeniyle çoğu çalışmada analiz dışı tutulmaktadır. Güncel araştırmalar, sabit bölgeler yerine adaptif ROI belirleme ve yüz hareketine duyarlı dinamik ROI stratejilerine odaklanmaktadır. Bu yöntemlerin özellikle serbest baş hareketi içeren zorlu ortamlarda sinyal kararlılığını artırdığı gözlemlenmiştir. Genel bir değerlendirme yapıldığında literatür ROI seçiminin rPPG sistemlerinin doğruluğu, stabilitesi ve biyometrik saldırı tespit performansı üzerinde belirleyici bir rol oynadığını ortaya koymaktadır. Mevcut bulgular anatomik özelliklere dayalı bölgesel ROI stratejilerinin en yüksek fizyolojik sinyal kalitesini sağladığını göstermektedir.

4.2.1.1 Bütünsel Yüz ROI

Erken dönem rPPG çalışmalarında ve hesaplama maliyeti düşük temel yaklaşımlarda genellikle Viola-Jones algoritması aracılığıyla tespit edilen yüz çerçevesinin tamamı ROI olarak kullanılmıştır (Kim vd. 2021). Ancak dikdörtgen formundaki bu tespit çerçevesi saç, boyun ve yüz dışı arka plan piksellerini de analize dahil etmektedir. Fizyolojik bilgi içermeyen bu gürültü kaynaklarının sürece katılması sinyal kalitesini düşürmekte ve optimal sonuçların elde edilmesini zorlaştırmaktadır (Kim vd. 2021). Bazı araştırmacılar, ROI seçimi ve takibini göz ardı ederek tüm video karesi üzerinden kalp hızı çıkarımı yapmayı önermiştir. Ancak bu yöntem hareket gürültüsüne karşı oldukça duyarlı olduğundan yalnızca uyku izleme gibi arka planın ve deneğin sabit olduğu kontrollü senaryolar için uygunluk göstermektedir (Ontiveros vd. 2023).

4.2.1.2 Anatomik Olarak Tanımlanmış Yüz Bölgeleri

Literatürde rPPG sinyal çıkarımı için genel kabul gören yaklaşım analizin alın, yanaklar, burun, ağız ve çene gibi temel anatomik bölgeler üzerine yoğunlaştırılmasıdır (Kim vd. 2021). Ancak araştırmacılar sinyal kalitesini optimize etmek amacıyla bu bölgeleri farklı segmentasyon stratejileriyle ele almışlardır. Kossack ve diğerleri tam yüz, alın, sağ yanak, sol yanak ve burun olmak üzere beş farklı ROI önermektedir. Çalışmada görünür cilt dokusunun eksikliği nedeniyle gözler, düşük sinyal kalitesi ve hareket artefaktları nedeniyle ise ağız ve çene bölgeleri analiz dışı bırakılmıştır. Ayrıca saçla kapanan alın, gözlük kullanımı veya başın yana dönmesi gibi durumlara karşı dayanıklılık sağlamak amacıyla simetrik segmentlerin kullanımı benimsenmiştir (Kossack vd. 2021). Benzer şekilde Kim ve diğerleri yüzü anatomik olarak 39 alt bölgeye ayırarak detaylı bir analiz gerçekleştirmiştir. Göz çevresindeki alanların birleştirilmesi ve nazolabial kıvrımlar gibi simetrik kısımların sol ve sağ olarak ayrılmasıyla toplamda 31 önerilen ROI bölgesi tanımlanmıştır. (Kim vd. 2021). Daha güncel çalışmalarda ise öznitelik noktası ve ızgara tabanlı yaklaşımlar öne çıkmaktadır. Bu kapsamda Ontiveros ve diğerleri, MediaPipe kütüphanesini kullanarak 468 yüz dönüm noktasını tespit etmiştir. Çalışmalarında ROI'leri sağ yanak, sol yanak ve alından elde edilen noktaların kombinasyonu olarak belirlemişlerdir (Ontiveros vd. 2023). Wu ve diğerleri, akrabalık doğrulaması amacıyla

yüz bölgesinden 100 farklı dikdörtgen ROI çıkarmış ve farklı cilt bölgelerinin farklı renk yansımalarına sahip olduğunu ortaya koymuştur (Wu vd. 2024). Lee ve diğerleri ise deepfake tespiti için yüz alanını 3×3'lük bir matrise ayırmış ve bu bölgelerin ortalama sinyal özelliklerini incelemiştir. Yapılan analizde, yüzün merkezi bölgesinin en ayırt edici performansı sergilediği rapor edilmiştir (Lee vd. 2025). İlerleyen yıllarda, Seibold ve diğerleri sinyal işlemeyi tüm yüz bölgesine uygularken gürültü karakteristiğini modellemek için arka planda seçilen iki homojen kare bölgeyi referans olarak kullanmıştır (Seibold vd. 2025).

rPPG'nin temel çalışma prensibi, speküler ve dağınık yansıma bileşenleri arasındaki optik kontrasta dayanmaktadır. Dağınık yansıma, cilt kalınlığıyla doğrudan ilişkili olan kan damarlarının derinliğine duyarlıdır. Kim ve diğerleri tarafından yapılan deneylerde, 39 anatomik yüz bölgesinin dermis ve epidermis kalınlıkları dikkate alınarak sinyal kalitesi analiz edilmiştir. Çalışmanın temel bulgusu, cilt ne kadar ince ve seçilen bölge ne kadar geniş olursa rPPG başarımının o ölçüde arttığı yönündedir. Analiz sonuçlarına göre en yüksek performansı gösteren 5 bölge (sağ yanak, iki kaş arası, alt orta alın, üst orta alın ve sol yanak) ortalama 1191.11 μm cilt kalınlığına sahiptir. En düşük performanslı 5 bölgenin ise ortalama 1581.39 μm cilt kalınlığına sahip olduğu raporlanmıştır. Bu bulgular, optimal ROI seçimi için anatomik özelliklerin ve doku kalınlığının kritik bir öneme sahip olduğunu göstermektedir (Kim vd. 2021).

5. DOĞRULAMA PROTOKOLLERİ ve PERFORMANS METRİKLERİ

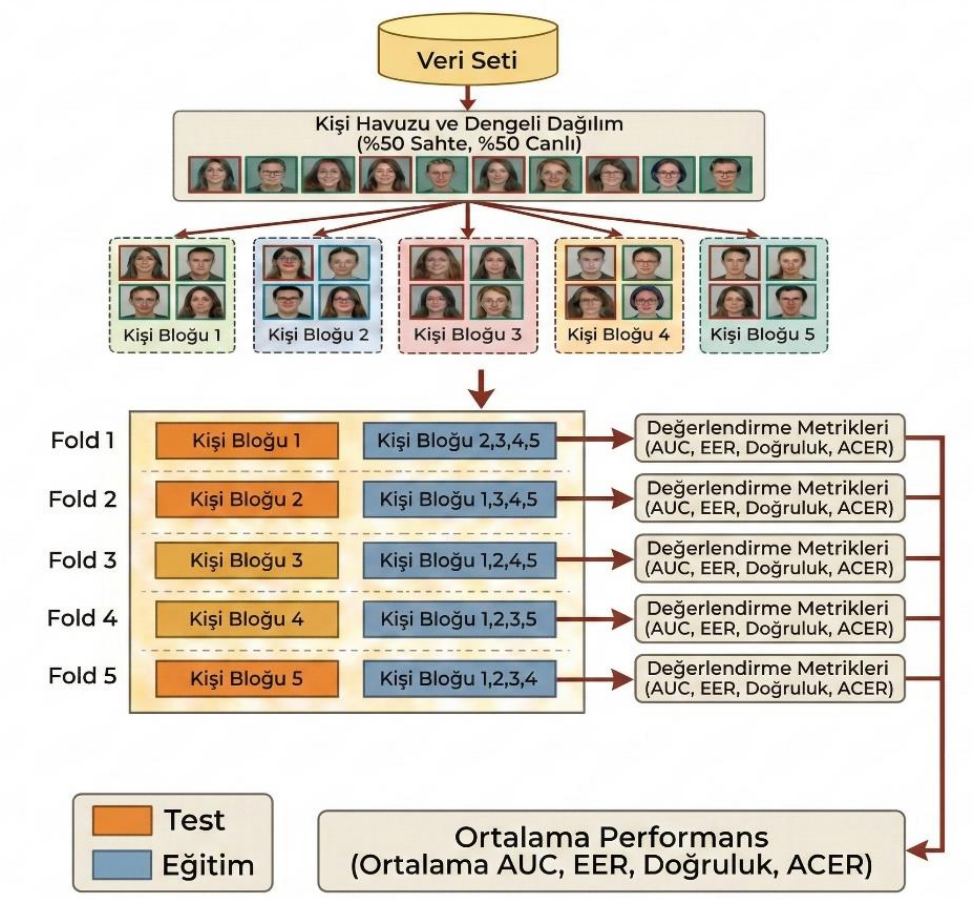
Bu bölümde geliştirilen rPPG tabanlı biyometrik canlılık tespiti sisteminin performansını nesnel, güvenilir ve tekrarlanabilir bir biçimde değerlendirmek amacıyla uygulanan doğrulama stratejileri ve değerlendirme metodolojileri detaylandırılmıştır. Deneysel sonuçların güvenilirliğini sağlamak ve modelin farklı bireyler, saldırı türleri ve veri dağılımları karşısındaki genelleme yeteneğini ortaya koymak amacıyla kişiden bağımsız doğrulama protokolleri esas alınmıştır. Bu kapsamda veri bölme stratejileri, çapraz doğrulama yaklaşımları ve sınıf dengesinin korunmasına yönelik yöntemler sistematik bir çerçevede sunulmuştur.

5.1 Doğrulama Protokolleri

Geliştirilen derin öğrenme modelinin performansını yansız ve istatistiksel olarak güvenilir bir şekilde değerlendirmek biyometrik güvenlik sistemlerinde kritik öneme sahiptir. Makine öğrenmesi modelleri ve özellikle yüksek kapasiteli derin sinir ağları eğitim verisindeki kestirme yolları ezberleme eğilimindedir. Biyometrik veri analizinde bu sorun modelin asıl odaklanması gereken canlılık özniteliklerini öğrenmek yerine kişiye özgü kimlik bilgilerini veya çevresel faktörleri ezberlemesi şeklinde kendini göstermektedir. Ezberlenen bu yanıltıcı unsurlar arasında özellikle arka plan dokusu ve kamera sensör gürültüsü gibi ortama özgü değişkenler yer almaktadır. Doğrulama stratejisinin söz konusu ezberlemeyi tespit edecek ve engelleyecek biçimde kurgulanmaması durumunda elde edilen test sonuçları yanıltıcı derecede yüksek çıkabilmektedir. Yapay başarı olarak nitelendirilen bu durum modelin gerçek dünya koşullarında eğitim setinde yer almayan yeni bir bireyle veya saldırı türü ile karşılaştığında başarısız olmasına yol açmaktadır.

Bu çalışmada, modelin genelleme yeteneğini ölçmek ve veri setine özgü yanlılıkları minimize etmek amacıyla literatürde kabul gören kişi tabanlı katmanlı 5-katlı çapraz doğrulama (Subject-based Stratified 5-Fold Cross-Validation) protokolü uygulanmıştır (Kohavi 1995). Bu protokol veri setini rastgele karıştırmak yerine deneklerin kimliklerine göre mantıksal bloklara ayırmakta ve her bir deneğin yalnızca tek bir katmanda yer

almasını garanti etmektedir. Bu stratejinin teorik temelleri, veri sızıntısını önleme mekanizması ve istatistiksel sağlamlığı Şekil 5.1'de detaylı şekilde sunulmuştur.



Şekil 5.1 Kişi tabanlı katmanlı 5-katlı çapraz doğrulama protokolünün şematik gösterimi

5.1.1 Kişi Bağımsız Doğrulama Protokolü

Biyometrik sahtekarlık tespiti problemlerinde karşılaşılan en büyük metodolojik hatalardan biri eğitim ve test kümeleri arasında veri sızıntısı yaşanmasıdır. Eğer aynı kişiye ait görüntüler hem eğitim hem de test setinde yer alırsa derin öğrenme modelleri canlılık özniteliklerini öğrenmek yerine kişinin kimlik bilgisini veya o kişiye özgü ortam koşullarını ezberleme eğilimi göstermektedir (Chingovska vd. 2012). Bu durum literatürde kimlik yanlılığı olarak tanımlanmaktadır. Chingovska ve diğerleri tarafından yapılan öncü çalışmalarda kişi çakışması bulunan test protokollerinde hata oranlarının %0'a yakın çıktığı ancak aynı modelin yalnızca görülmemiş bireylerden oluşan bir test setinde değerlendirildiğinde hata oranlarının %50 seviyelerine kadar gerilediği ortaya

konulmuştur (Chingovska vd. 2012). Bu performans düşüşünün temel nedeni modelin yüzey dokusundaki mikro-desenleri analiz etmek yerine veri setindeki kişileri doğrudan ezberleyerek kimlik bilgisi ile etiket eşleştirmesi yapmasıdır.

Bu çalışmada söz konusu metodolojik hatayı engellemek amacıyla veri seti bölünürken rastgele karıştırma yerine kişi tabanlı gruplama yöntemi esas alınmıştır. Bu yöntemde veri setindeki toplam N adet denek $S = \{s_1, s_2, \dots, s_N\}$ kümesi olarak tanımlanmaktadır. Çapraz doğrulamanın her bir k adımında, test kümesi (S_{test}) ile eğitim kümesi ($S_{eğitim}$) arasındaki kesişim kümesinin boş olması ($(S_{test}) \cap (S_{eğitim}) = \emptyset$) garanti altına alınmıştır. Söz konusu protokol eğitim ve test kümelerini birbirinden kesin sınırlarla ayırmakta ve veri sızıntısını ortadan kaldırmaktadır. Model, eğitim aşamasında bir deneğe ait yüz verisini gördüyse test aşamasında aynı deneğe ait ne canlı ne de sahte hiçbir örnekle karşılaşmamaktadır. Bu yaklaşım, modelin eğitim sırasında hiç görmediği yeni yüzler üzerindeki performansını ölçerek sistemin gerçek dünya koşullarındaki başarısını daha doğru biçimde simüle etmektedir (Marcel vd. 2019).

rPPG sinyalleri doğası gereği kişiye özgü biyometrik karakteristikler barındırmaktadır. Bu nedenle uygulanan ayrım stratejisi modelin bireysel fizyolojik imzaları ezberlemesini engelleyerek bunun yerine periyodiklik ve BVP morfolojisi gibi evrensel nabız örüntülerine odaklanmasını sağlamaktadır. Bu metodolojik yaklaşım sistemin genellenebilirliği ve gerçek dünya koşullarındaki uygulanabilirliği için zorunlu bir ön koşul niteliği taşımaktadır.

5.1.2 Katmanlı Veri Bölme Yaklaşımı ile Sınıf Dengesinin Sağlanması

Kullanılan veri setlerinde canlı ve sahte saldırı örneklerinin sayıları çoğu zaman eşit olmamaktadır. Veri toplama süreçlerinde gerçek kullanıcılara ait veri elde etmek görece kolayken yüksek kaliteli 3D maskeler veya özel senaryolu saldırı videoları üretmek maliyetli ve zaman alıcı bir süreçtir. Bu durum, veri setinde sınıf dengesizliği problemine yol açmaktadır. Veri setindeki bu dengesizlik modelin eğitim sürecinde çoğunluk sınıfına eğilimli olmasına neden olabilmektedir. Örneğin eğitim setinde %90 oranında gerçek ve %10 oranında sahte örnek bulunduğu bir senaryoda model tüm girdileri gerçek olarak

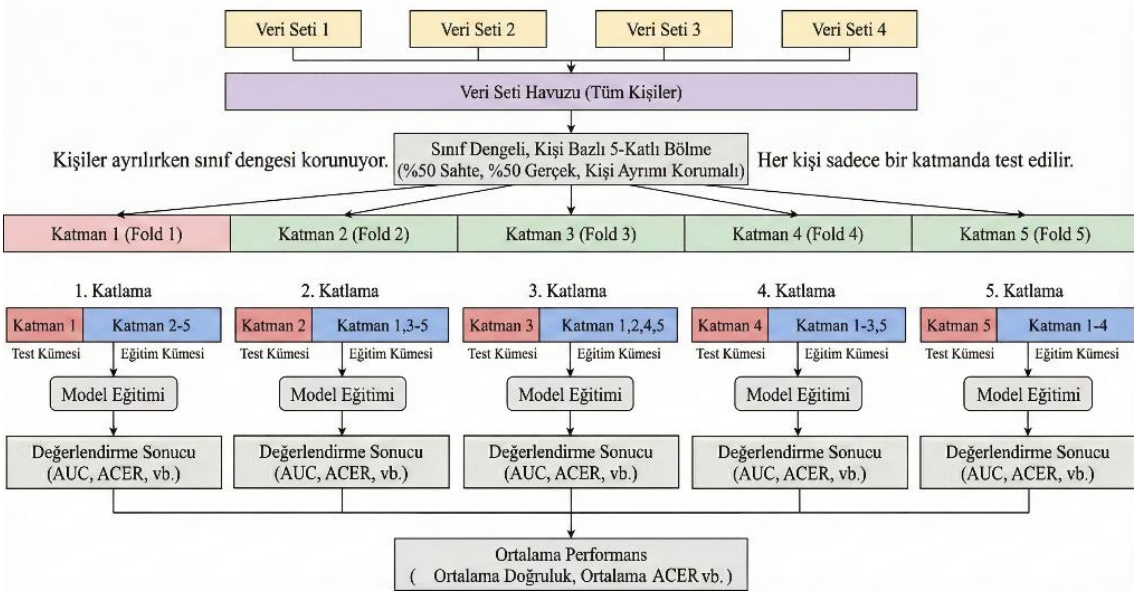
etiketleyerek %90 başarı elde edebileceğini öğrenmektedir. Bu optimizasyon hatası, modelin ayırt edici öznitelikleri öğrenme motivasyonunu kaybetmesine yol açmaktadır. Bu durumda model veri setine karşı yanlılık geliştirmekte ve nadir sınıf olan saldırı örneklerini tespit etmekte başarısız olmaktadır. Bu problemi minimize etmek ve adil bir değerlendirme sağlamak amacıyla bu çalışmada katmanlı örnekleme yöntemi kullanılmıştır. Bu yöntem, veri seti parçalanırken orijinal dağılımdaki sınıf oranlarının her bir alt kümede korunmasını esas almaktadır. Matematiksel olarak tüm veri setindeki pozitif örnek oranı α ise oluşturulan her bir k eğitim ve test bloğunda da bu oranın yaklaşık olarak α seviyesinde tutulması hedeflenmektedir. Bu sayede modelin her eğitim ve test döngüsünde veri setinin genel karakteristiğini temsil eden dengeli bir örneklem uzayı ile karşılaşması sağlanmaktadır. Bu sayede değerlendirme metriklerinin istatistiksel sapması minimize edilmektedir (Forman ve Scholz 2010). Katmanlı yapı varyansı düşürerek farklı katmanlardan elde edilen sonuçların birbirine daha yakın ve tutarlı olmasını sağlamaktadır.

5.1.3 K-Fold Çapraz Doğrulama Tekniği

Literatürde sıkça kullanılan tekil eğitim-test ayrımı seçilen test kümesinin şans eseri çok kolay veya çok zor örneklerden oluşması riskini taşımaktadır. Bu durum modelin performansının yanlış değerlendirilmesine ve yanıltıcı sonuçlar üretmesine neden olabilmektedir. Kohavi tarafından gerçekleştirilen temel çalışmalar, k değerinin seçiminin çapraz doğrulama sonuçları üzerinde belirleyici bir etkiye sahip olduğunu ortaya koymuştur. Düşük k değerlerinin tercih edilmesi durumunda, her bir katlamada modelin eğitildiği veri kümesinin boyutu azalmaktadır. Bu durum modelin veri dağılımına ait karakteristikleri yeterince öğrenememesine dolayısıyla yüksek yanlılık sergilemesine neden olmaktadır. Bunun bir sonucu olarak elde edilen performans ölçütleri modelin gerçek genelleme yeteneğini yansıtmakta yetersiz kalmakta ve model başarımının olduğundan daha düşük değerlendirilmesine yol açmaktadır. Buna karşılık çok yüksek k değerlerinin seçilmesi eğitim kümelerinin büyük ölçüde birbirine benzemesine neden olmaktadır. Bu durum her bir örneğin tek başına test kümesi olarak kullanıldığı bırak-bir-çıkar çapraz doğrulama (Leave-One-Out Cross-Validation, LOOCV) yaklaşımında açık biçimde gözlemlenmektedir. Bu durum farklı katlamalarda

eğitilen modeller arasındaki korelasyonu artırarak tahmin sonuçlarının varyansını yükseltmektedir. Aynı zamanda hesaplama maliyetinin aşırı derecede artmasına neden olmaktadır. Bu nedenle k değerinin seçimi yanlılık-varyans dengesi ile hesaplama verimliliğinin birlikte gözetildiği kritik bir optimizasyon problemi olarak değerlendirilmektedir.

Literatürde en uygun denge noktası olarak k değerinin 5 veya 10 olarak seçilmesi önerilmektedir. Bu aralıkta hem yanlılığın hem de varyansın minimize edildiği modelin genelleme yeteneğinin en kararlı şekilde ölçüldüğü kabul edilmektedir (Kohavi 1995). Bu doğrultuda bu çalışmada $k = 5$ seçilerek veri seti 5 eşit parçaya bölünmüştür. Her döngüde verinin %80'i eğitim %20'si test için kullanılmıştır. Bu strateji Replay-Mobile, 3DMAD, PURE ve UBFC-RPPG gibi farklı karakteristiklere sahip veri setlerinin birleşimi olan heterojen bir havuzda her bir örneğin hem eğitim hem de test setinde en az bir kez yer almasını garanti etmektedir. Sonuç olarak, raporlanan performans modelin sadece belirli bir veri alt kümesindeki başarısını yanı sıra tüm veri uzayındaki genelleştirilmiş başarısını yansıtmaktadır. K- katlı çapraz doğrulama mimarisinin şematik gösterimi Şekil 5.2'de sunulmuştur. Bu protokol ISO/IEC 30107-3 standartlarında önerilen değerlendirme kriterleri ile uyumludur.



Şekil 5.2 Kişi tabanlı 5- katlı çapraz doğrulama mimarisi

5.2 Değerlendirme Metrikleri

Önerilen modelin performansı biyometrik doğrulama ve yüz canlılık tespiti alanında yaygın biçimde kullanılan standart değerlendirme metrikleri kullanılarak analiz edilmiştir. Biyometrik sistemlerin performans değerlendirmesi yalnızca algoritmik bir sınıflandırma doğruluğu problemi olarak ele alınamaz. Aksine bu süreç operasyonel risklerin, kullanıcı deneyiminin ve olası güvenlik açıklarının istatistiksel olarak yönetildiği çok katmanlı bir analiz sürecidir. Bu bağlamda çalışmada benimsenen değerlendirme metodolojisi uluslararası standartlara özellikle ISO/IEC 30107-3:2017 (Information technology - Biometric Presentation Attack Detection) standardına tam uyumluluk gösterecek şekilde yapılandırılmıştır (Busch 2023). Yüz canlılık tespiti (Face Anti-Spoofing, FAS), doğası gereği asimetrik maliyetlere sahip bir problemdir. Bir saldırı girişiminin sistem tarafından fark edilemeyerek onaylanması ile yetkili bir kullanıcının hatalı bir şekilde reddedilmesi durumlarının doğuracağı sonuçlar eşdeğer bir risk profili oluşturmamaktadır. Söz konusu hata türlerinin tolere edilebilirliği, uygulamanın operasyonel senaryosuna göre dramatik farklılıklar göstermektedir. Örneğin ulusal sınır güvenliği gibi kritik uygulamalarda öncelik güvenlik olduğundan yanlış kabul riski minimize edilmeye çalışılmaktadır. Günlük mobil cihaz erişimi gibi senaryolarda öncelik kullanıcı konforu olduğundan yanlış ret oranının düşük olması hedeflenmektedir (Busch 2023). Bu nedenle tek bir skaler değer yerine sistemin karakteristiklerini farklı çalışma noktalarında analiz eden çok boyutlu bir metrik seti kullanılmıştır.

5.2.1 Doğruluk

Doğruluk, modelin genel sınıflandırma başarımını ifade eden en temel performans göstergelerinden biridir. Doğru sınıflandırılan örneklerin toplam örnek sayısına oranı olarak tanımlanmaktadır. Bu ilişki Denklem 5.1'de sunulmuştur.

$$Doğruluk = \frac{DP+DN}{DP+DN+YP+YN} \tag{5.1}$$

Denklem 5.1'de yer alan DP (Doğru Pozitif) gerçek canlı örneklerin doğru bir şekilde canlı olarak sınıflandırıldığı durumları, DN (Doğru Negatif) ise sahte örneklerin doğru

şekilde sahte olarak sınıflandırıldığı durumları göstermektedir. YP (Yanlış Pozitif) ise gerçekte sahte olan örneklerin yanlışlıkla canlı olarak sınıflandırıldığı durumları göstermektedir. Bu durumda model sahte bir örneği canlı sınıfına atayarak güvenlik açısından kritik bir hata yapmaktadır. YN (Yanlış Negatif) ise gerçekte canlı olan örneklerin yanlışlıkla sahte olarak sınıflandırıldığı durumları ifade etmektedir. Bu durum modelin geçerli ve gerçek bir kullanıcıyı reddederek hata yapması anlamına gelmektedir.

Doğruluk metriği makine öğrenmesi problemlerinde en temel performans göstergesi olarak kabul edilmektedir. Ancak biyometrik sahtekarlık tespiti gibi sınıf dengesizliğinin yaygın olduğu alanlarda yanıltıcı sonuçlar doğurabilme potansiyeline sahiptir. Gerçek hayat senaryolarında sisteme yapılan canlı giriş denemeleri saldırı denemelerinden sayıca çok daha fazladır. Literatürde doğruluk paradoksu olarak adlandırılan bu durumda, bir model tüm girdileri canlı olarak sınıflandırsa dahi test kümesinin %99'unun canlı veriden oluşması hâlinde %99 doğruluk oranına ulaşabilmektedir. İstatistiksel olarak başarılı görünen bu modelin saldırıları hiç tespit edememesi nedeniyle güvenlik değeri pratikte sıfıra yakındır. Bu nedenle çalışmamızda doğruluk metriği, yalnızca modelin genel öğrenme kapasitesine dair ilk bakış sağlayan bir gösterge olarak sunulmuştur. Sistemin asıl güvenilirliği hataya dayalı metrikler üzerinden değerlendirilmiştir. Özellikle Replay-Mobile ve 3DMAD gibi heterojen veri setlerinin birleştirildiği bu çalışmada, doğruluk değerinin veri setleri arasındaki dağılım farklılıklarından etkilenebileceği göz önünde bulundurulmalı ve yorumlanırken dikkatli olunmalıdır.

5.2.2 Saldırı Sunumu Sınıflandırma Hata Oranı

Saldırı sunumu sınıflandırma hata oranı (Attack Presentation Classification Error Rate, APCER) ISO/IEC 30107-3:2017 standardına göre sahte sunumların yanlışlıkla canlı olarak sınıflandırılma oranını ifade etmektedir. ISO/IEC standardı, biyometrik sistemleri değerlendirirken sunum saldırısı aracı kavramını merkeze almaktadır. PAI biyometrik sensörü aldatmak amacıyla kullanılan fotoğraf, maske, video ekranı veya silikon yüz gibi yapay nesneleri ifade etmektedir. Bu bağlamda APCER, sistemin güvenlik duvarının ne kadar geçirgen olduğunu ölçen en kritik parametrelerden biridir.
Bir biyometrik güvenlik sistemi için yanlış kabul, tolere edilmesi en güç hata türüdür.

Çünkü bu durum yetkisiz bir kişinin sisteme erişim sağlaması anlamına gelmektedir. Düşük APCER değeri sistemin sahte saldırılara karşı yüksek dayanıklılık gösterdiğini ifade etmektedir. APCER metriği Denklem 5.2'de gösterildiği şekilde hesaplanmaktadır.

$$APCER = \frac{FP}{YP+DN} \tag{5.2}$$

Denklemde yer alan YP sahte bir sunumun sistem tarafından hatalı bir şekilde canlı olarak etiketlendiği güvenlik ihlali durumlarını temsil etmektedir. DN ise saldırıların sistem tarafından başarıyla engellendiği durumları temsil etmektedir. ISO/IEC 30107-3 standardı, APCER hesaplanırken farklı saldırı türlerinin ayrı ayrı ele alınmasını önermektedir. Bu kapsamda sistemin en zayıf performans gösterdiği saldırı türüne ait hata oranının en kötü durum APCER olarak raporlanmasını önermektedir.

Bu çalışmada kullanılan veri setleri basit 2B baskı saldırılarından derinlik ve doku bilgisi içeren sofistike 3D maske saldırılarına kadar geniş bir saldırı spektrumunu kapsamaktadır. Önerilen hibrit modelin rPPG akışı ve görsel doku akışı birlikte değerlendirildiğinde sistemin video tekrar saldırılarında ekranda doğal olarak oluşmayan fizyolojik kan akışı imzalarını tespit edebildiği görülmektedir. Aynı zamanda ekran yüzeyine özgü kenar yapıları ve yansıma artefaktlarını analiz ederek APCER değerini azaltmayı hedeflediği görülmektedir.

Bununla birlikte model 3DMAD veri setindeki yüksek kaliteli 3D maske saldırılarını ayırt ederken hem maske materyallerine özgü dokusal tutarsızlıkları hem de gerçek yüzlerde gözlemlenen fizyolojik sinyal bileşenlerinin yokluğunu birlikte kullanarak güvenliği artırmaktadır. Dolayısıyla raporlanan APCER değeri önerilen hibrit yönteminin birden fazla saldırı vektörüne karşı sağladığı ortalama dayanıklılığı temsil etmektedir. Modelin hem fotopletismografik hem de görsel temelli sahtecilik göstergelerini bütüncül biçimde değerlendirebildiğini ortaya koymaktadır.

5.2.3 Gerçek Sunum Sınıflandırma Hata Oranı

Gerçek Sunum Sınıflandırma Hata Oranı (Bona Fide Presentation Classification Error Rate, BPCER), ISO/IEC 30107-3 standardına göre gerçek sunumların yanlışlıkla sahte olarak sınıflandırılma oranını göstermektedir. Bu metrik, biyometrik sistemin kullanıcı konforunu ve sürtünmesiz erişim kapasitesini ölçmektedir. BPCER, doğrulanmış bir kullanıcının sisteme erişiminin hatalı biçimde reddedildiği durumları ifade etmektedir. Yüksek BPCER değerleri, kullanıcıların sistemi birden fazla kez denemek zorunda kalmasına, zaman kaybına ve nihayetinde sisteme duyulan güvenin azalmasına yol açmaktadır. Özellikle bankacılık veya havaalanı geçiş kontrolü gibi yüksek hacimli ticari uygulamalarda %5'in üzerindeki bir BPCER oranı müşteri memnuniyetsizliği nedeniyle sistemin kullanımdan kaldırılmasına sebep olabilmektedir. Düşük BPCER değeri, sistemin gerçek kullanıcıları reddetme olasılığının az olduğunu göstermektedir. BPCER metriği Denklem 5.3'te gösterildiği şekilde hesaplanmaktadır.

$$BPCER = \frac{FN}{YN+DP} \tag{5.3}$$

Denklemde yer alan YN, canlı bir yüzün sistem tarafından hatalı bir şekilde sahte olarak etiketlendiği durumları temsil etmektedir. DP ise gerçek kullanıcıların sistem tarafından başarıyla tanındığı ve doğrulandığı durumları ifade etmektedir. Bu çalışmada kullanılan PURE ve UBFC-RPPG veri setleri farklı kafa hareketleri ve ışık koşulları altında kaydedilmiş gerçek kullanıcı videolarını içermektedir. Bu bağlamda BPCER metriği, modelin çevresel gürültülere ve doğal varyasyonlara karşı dayanıklılığını test etmek için uygun bir gösterge niteliği taşımaktadır. Ayrıca biyometrik sistemlerde sıkça tartışılan demografik ön yargı konusu da BPCER ile doğrudan ilişkilidir. Literatürdeki çalışmalar, koyu tenli bireylerde veya belirli yaş gruplarında melanin yoğunluğunun ışık emilimini etkilemesi nedeniyle rPPG sinyal kalitesinin düşebileceğini göstermektedir. Bu durum, söz konusu gruplarda BPCER oranlarını artmasına neden olabilmektedir. Çalışmada elde edilen düşük ortalama BPCER değeri modelin kararlı ve dengeli yapısını doğrulamaktadır. Bu performans ideal aydınlatma koşullarının ötesine geçerek farklı kullanıcı profillerini ve cilt tonlarını da kapsamaktadır. Sonuçlar önerilen yöntemin geniş bir demografik yelpazede güvenilir ve kapsayıcı bir başarım sergilediğine işaret etmektedir.

5.2.4 Ortalama Sınıflandırma Hata Oranı

Ortalama Sınıflandırma Hata Oranı (Average Classification Error Rate, ACER), APCER ve BPCER metriklerinin aritmetik ortalaması alınarak sistemin genel performansını özetleyen bütüncül bir değerlendirme ölçütüdür. Bu metrik, biyometrik sistemlerin iki temel bileşeni olan güvenlik ve kullanılabilirlik arasındaki dengeyi temsil etmektedir. Genellikle sistemde tek bir eşik değeri belirlendiğinde APCER ve BPCER değerlerinin ters orantılı biçimde değiştiği gözlemlenmektedir. Karar eşiği yükseltildiğinde sistem daha katı hale gelmekte ve güvenlik seviyesi artmaktadır. Bu durum APCER değerini düşürmektedir. Ancak bu katı durum eş zamanlı olarak BPCER oranının artmasına neden olmaktadır. ACER metriği, bu ödünleşimin merkez noktasındaki genel hatayı ifade etmektedir. Dolayısıyla, bu metrik sistemin hem saldırı sunumlarına hem de gerçek sunumlara karşı gösterdiği dengeli başarımı ölçmek amacıyla kullanılmaktadır. ACER metriği Denklem 5.4'te gösterildiği şekilde tanımlanmaktadır.

$$ACER = \frac{APCER+BPCER}{2} \tag{5.4}$$

Bu metriğin aritmetik ortalama kullanması matematiksel olarak güvenlik hatasının ve kullanım hatasının eşit ağırlıkta cezalandırıldığı varsayımına dayanmaktadır. Ancak gerçek dünya uygulamalarında bu ağırlıklar senaryoya göre değişkenlik gösterebilmektedir. Güvenlik ihlallerinin tolere edilemeyeceği askeri tesisler gibi alanlarda APCER'in asgari düzeyde tutulması hayati önem taşımaktadır. Kullanıcı memnuniyetinin ve etkileşiminin odak noktası olduğu sosyal platformlarda ise geçerli kullanıcıyı reddetmemek adına BPCER değerinin düşük olması daha çok tercih edilmektedir. Buna rağmen akademik kıyaslamalarda ACER en tarafsız performans göstergesi olarak kabul edilmektedir. Çünkü bu metrik modelin herhangi bir sınıfa aşırı eğilim göstermediğini doğrulamaktadır.

5.2.5 Eşit Hata Oranı

Eşit hata oranı (Equal Error Rate, EER), sistemin yanlış kabul (False Acceptance Rate, FAR) ve yanlış reddetme (False Rejection Rate, FRR) oranlarının birbirine eşitlendiği

noktadaki hata değerini ifade etmektedir. Biyometrik sistemlerde karar verici mekanizma, genellikle 0-1 aralığında bir olasılık skoru üretmektedir. Elde edilen bu skor belirlenen bir karar eşiği (τ) ile karşılaştırılarak nihai sınıflandırma yapılmaktadır. EER metriği, spesifik bir eşik değerinin seçiminden bağımsızdır. Bu nedenle sistemin doğal ayrım gücünü yansıtan en nesnel performans göstergelerinden biri olarak kabul edilmektedir. Grafik üzerinde APCER ve BPCER eğrilerinin kesişim noktasını temsil eden EER değerinin düşük olması, sistemin genel hata oranının asgari düzeyde olduğunu ifade etmektedir. Bu durum, sistemin daha geniş ve güvenli çalışma aralığına sahip olduğunu göstermektedir. EER değeri düşük olan sistemler, genel olarak daha dengeli ve güvenilir performans sergilemektedir. EER metriği matematiksel olarak Denklem 5.5'te gösterilen koşulun sağlandığı noktadaki hata değeri olarak tanımlanmaktadır.

$$EER = APCER(\tau_{EER}) = \text{BPCER}(\tau_{EER}) \tag{5.5}$$

Denklemde yer alan τ_{EER} iki hata oranını eşitleyen optimal eşik değerini temsil etmektedir.

5.2.6 ROC

ROC (Receiver Operating Characteristic) eğrisi, bir ikili sınıflandırma modelinin karar eşiğinin farklı değerleri altında sergilediği performansı görselleştiren temel bir değerlendirme aracıdır. Bu analiz yöntemi, doğruluk gibi sınıf dağılımına duyarlı metriklerin aksine sınıflar arasındaki dağılımın dengesiz olduğu veri kümelerinde daha nesnel, güvenilir ve karşılaştırılabilir bir değerlendirme çerçevesi sunmaktadır ROC analizi sınıf oranlarından bağımsız bir çerçeve sunduğu için biyometrik güvenlik ve sunum saldırısı tespiti gibi uygulamalarda literatürde yaygın biçimde tercih edilmektedir (Davis ve Goadrich 2006).

ROC eğrisi, karar eşiği değiştirildikçe modelin iki temel performans ölçütündeki değişimi görselleştirmektedir. Bu ölçütlerden ilki olan gerçek pozitif oranı (True Positive Rate, TPR), doğru biçimde pozitif olarak sınıflandırılan örneklerin tüm gerçek pozitif örnekler içindeki oranını ifade etmektedir ve literatürde duyarlılık olarak da adlandırılmaktadır. TPR değeri, Denklem 5.6'da gösterildiği üzere gerçek pozitiflerin toplam pozitif

örneklere oranı şeklinde tanımlanmaktadır. Ayrıca yanlış negatif oranının (False Negative Rate, FNR) tümleyeni olarak hesaplanmaktadır.

$$TPR = \frac{TP}{TP+FN} = 1 - FNR \tag{5.6}$$

Bu denklemde DP gerçek pozitifleri, YN ise yanlış negatifleri temsil etmektedir. ROC analizinde kullanılan ikinci temel ölçüt olan yanlış pozitif oranı (False Positive Rate, FPR) ise gerçekte negatif olan ancak yanlışlıkla pozitif olarak sınıflandırılan örneklerin tüm gerçek negatifler içindeki oranını ifade etmektedir. FPR değeri, Denklem 5.7'de gösterildiği üzere yanlış pozitiflerin toplam negatif örneklere oranı şeklinde tanımlanmakta ve gerçek negatif oranının (True Negative Rate, TNR) tümleyeni olarak hesaplanmaktadır.

$$FPR = \frac{FP}{FP+DN} = 1 - TNR \tag{5.7}$$

Burada YP yanlış pozitifleri, DN ise gerçek negatifleri ifade etmektedir. ROC uzayı yatay eksende yanlış pozitif oranının dikey eksende ise gerçek pozitif oranının yer aldığı [0, 1] × [0, 1] aralığında tanımlanmaktadır. Bu uzayda rastgele tahmin yapan bir sınıflandırıcı, diyagonal doğru (y = x) üzerinde konumlanmakta ve bu duruma karşılık gelen eğri altındaki alan (Area Under Curve, AUC) değeri 0,5 olmaktadır. Buna karşılık ideal bir sınıflandırıcıda ROC eğrisi sol üst köşeye yani FPR = 0 ve TPR = 1 noktasına yaklaşmakta ve bu durumda AUC değeri 1,0'a ulaşmaktadır.

ROC eğrisinin oluşturulması sürecinde parametrik olmayan bir yaklaşım izlenmektedir. Model tarafından üretilen tahmin skorları kümesi $S = \{s_1, s_2, \dots, s_n\}$ sıralanmakta ve bu küme içerisindeki her bir skor potansiyel bir karar eşiği olarak ele alınmaktadır. Algoritmik açıdan bakıldığında süreç ilk olarak $O(n \log n)$ karmaşıklığında bir sıralama işlemi ile başlamakta ardından $O(n)$ karmaşıklığında gerçekleştirilen doğrusal tarama ile karışıklık matrisleri güncellenerek TPR ve FPR noktaları hesaplanarak iki boyutlu uzayda eğri oluşturulmaktadır. Bu hesaplama verimliliği, ROC analizinin hem teorik çalışmalarda hem de gerçek zamanlı sistemlerde yaygın biçimde kullanılmasını sağlamaktadır.

5.2.7 AUC

AUC (Area Under Curve), ROC eğrisinin altında kalan alanın hesaplanması yoluyla bir sınıflandırma modelinin genel ayırt edicilik kapasitesini tek bir skaler değerle ifade eden temel performans ölçütüdür. Eşik değerinden bağımsız olması ve sınıf dengesizliğine karşı doğruluk metriğine kıyasla daha dirençli yapısı sayesinde AUC metriği ikili sınıflandırma problemlerinin değerlendirilmesinde literatürde en yaygın kullanılan ölçütlerden biri olmasını sağlamıştır. Matematiksel olarak AUC, yanlış pozitif oranının sıfırdan bire değişimi boyunca gerçek pozitif oranının integrali olarak tanımlanmaktadır. Bu ilişki Denklem 5.8'de gösterilmiştir.

$$AUC = \int_0^1 TPR(FPR^{-1}(x))\, dx \tag{5.8}$$

Bu ifadede $TPR(t)$ duyarlılığı, $FPR(t)$ yanlış pozitif oranını, $FPR^{-1}(x)$ ise FPR değerinin x olduğu karar eşik değerini veren ters fonksiyonu temsil etmektedir. İntegral ifadesi, FPR değerinin 0'dan 1'e ilerlemesi sırasında ROC eğrisinin altında kalan toplam alanın hesaplanmasını ifade etmektedir. Teorik olarak sürekli bir fonksiyon gerektirse de pratikte ROC eğrisi sonlu sayıda veri noktasından oluşmaktadır. Bu nedenle integral işlemi çoğunlukla sayısal yöntemler kullanılarak yaklaşık olarak hesaplanmaktadır. En yaygın kullanılan yöntemlerden biri trapez kuralıdır ve Denklem 5.9'da gösterildiği gibi matematiksel olarak ifade edilmektedir.

$$AUC \approx \sum_{i=1}^{m-1} \frac{(y_i+y_{i+1})(x_{i+1}-x_i)}{2} \tag{5.9}$$

Bu denklemde x_i ROC eğrisi üzerindeki i'inci yanlış pozitif oranı noktasını, y_i ise karşılık gelen gerçek pozitif oranı değerini ifade etmektedir. m ROC eğrisi üzerinde elde edilen toplam nokta sayısını temsil etmektedir. $(x_{i+1}-x_i)$ ardışık iki FPR noktası arasındaki yatay mesafeyi temsil etmektedir. $\frac{(y_i+y_{i+1})}{2}$ ise ilgili aralıktaki ortalama TPR değerini yani trapezin yüksekliğini göstermektedir. Bu yöntem ROC eğrisi altındaki alanın etkin ve hesaplama açısından verimli bir biçimde tahmin edilmesini sağlamaktadır.

İstatistiksel açıdan bakıldığında AUC metriği, Wilcoxon–Mann–Whitney U istatistiğinin normalize edilmiş bir biçimiyle eş değerdir. Bu durum AUC'ye olasılıksal bir yorum kazandırmaktadır (Hanley ve McNeil 1982). Bu yorum Denklem 5.10'da ifade edildiği üzere rastgele seçilen bir pozitif örneğin tahmin skorunun rastgele seçilen bir negatif örneğin skorundan daha yüksek olma olasılığına karşılık gelmektedir.

$$AUC = P\left(\hat{y}_{pos} > \hat{y}_{neg}\right) \tag{5.10}$$

Burada $\hat{y}_{pos}$ rastgele bir pozitif örneğin tahmin skorunu, $\hat{y}_{neg}$ ise rastgele bir negatif örneğin skorunu temsil etmektedir. $P(\cdot)$ ise olasılık fonksiyonunu temsil etmektedir. $P\left(\hat{y}_{pos} > \hat{y}_{neg}\right)$ ise rastgele seçilen bir pozitif örneğin skorunun rastgele seçilen bir negatif örneğin skorundan daha yüksek olma olasılığını ifade etmektedir. Bu olasılıksal yorum AUC'yi bir sıralama metriği haline getirmekte ve modelin kalibrasyonundan bağımsız bir değerlendirme sunmaktadır. Hesaplanan AUC değerinin istatistiksel güvenilirliğini göstermek amacıyla AUC'nin standart hatası Denklem 5.11 gösterildiği şekilde hesaplanmaktadır.

$$SE(AUC) = \sqrt{\frac{AUC(1-AUC)+(n_p-1)(Q_1-AUC^2)+(n_n-1)(Q_2-AUC^2)}{n_p n_n}} \tag{5.11}$$

Bu denklemde n_p canlı örnek sayısını, n_n saldırı örnek sayısını ifade etmektedir. $Q_1 = \frac{\mathrm{AUC}}{2-\mathrm{AUC}}$ ve $Q_2 = \frac{2\mathrm{AUC}^2}{1+\mathrm{AUC}}$ yardımcı terimler olarak tanımlanmaktadır. Söz konusu istatistiksel çerçeve AUC değerleri için güven aralıklarının oluşturulmasına ve farklı modellerin performanslarının istatistiksel anlamlılık düzeyinde karşılaştırılmasına olanak tanımaktadır.

AUC metriğinin temel avantajları arasında eşik değerinden bağımsız olması ve sınıf dengesizliğine karşı görece dayanıklı bir yapı sunması öne çıkmaktadır. Bununla birlikte özellikle aşırı dengesiz veri kümelerinde ROC eğrisinin erken sıralama hatalarını maskeleyebilmesi nedeniyle Kesinlik-Duyarlılık (Precision–Recall, PR) eğrisi ve bu eğrinin altındaki alanı temsil eden Ortalama Kesinlik (Average Precision, AP) metriği daha bilgilendirici sonuçlar sağlayabilmektedir. Ayrıca AUC metriği, skor tabanlı

modeller için daha uygun olmasına rağmen kural tabanlı sistemlere doğrudan genellenmesi sınırlı kalabilmektedir.

5.2.8 Karmaşıklık Matrisi

Karmaşıklık matrisi, ikili sınıflandırma problemlerinde bir modelin tahmin performansını ayrıntılı biçimde analiz edebilmek amacıyla kullanılan temel değerlendirme araçlarından biridir. Özellikle biyometrik canlılık tespiti ve sahtekârlık algılama sistemlerinde modelin gerçek sunumlar ile saldırı sunumlarını ne ölçüde doğru ayırt edebildiğini nicel olarak ortaya koymak açısından kritik bir rol oynamaktadır. ROC eğrisinin temelini oluşturan gerçek pozitif oranı ve yanlış pozitif oranı gibi metrikler doğrudan karmaşıklık matrisinde yer alan dört temel bileşene dayanmaktadır.

Doğru pozitifler sistem tarafından canlı olarak doğru biçimde sınıflandırılan gerçek kullanıcı örneklerinin sayısını ifade etmektedir. Bu değer modelin gerçek kullanıcıları doğru tanıma yeteneğini yansıttığı için özellikle kullanıcı deneyimi ve sistemin kullanılabilirliği açısından önem taşımaktadır. DP değerinin yüksek olması canlı sunumların hatasız biçimde kabul edildiğini göstermektedir.

Yanlış negatifler, gerçekte canlı olan ancak sistem tarafından sahte olarak yanlış sınıflandırılan örnekleri temsil etmektedir. Bu tür hatalar canlı kullanıcıların sistem tarafından reddedilmesine yol açtığından biyometrik güvenlik sistemlerinde kritik bir sorun teşkil etmektedir. YN değerinin artması, sistemin aşırı katı davrandığını ve gerçek sunumları yeterince tolere edemediğini göstermektedir.

Yanlış pozitifler, gerçekte sahte olan ancak sistem tarafından canlı olarak yanlış biçimde sınıflandırılan örneklerin sayısını ifade etmektedir. Bu hata türü sahtekârlık tespiti açısından en riskli durumlardan biridir. Zira saldırı sunumlarının sistem tarafından kabul edilmesi doğrudan güvenlik açığına neden olmaktadır. YP değerinin düşük tutulması sistemin saldırılara karşı dayanıklılığını ve güvenilirliğini göstermektedir.

Doğru negatifler ise sahte örneklerin sistem tarafından doğru biçimde sahte olarak

sınıflandırıldığı durumları ifade etmektedir. Bu bileşen sistemin sahte sunumları başarıyla engelleme kapasitesini temsil etmektedir. DN değerinin yüksek olması saldırıların etkili biçimde tespit edildiğini ve reddedildiğini göstermektedir.

Bu dört temel bileşen modelin hata profiline ilişkin kapsamlı bir bakış sunmakta ve doğruluk, duyarlılık, özgüllük, TPR, FPR, EER ve ACER gibi birçok değerlendirme metriğinin hesaplanmasında doğrudan kullanılmaktadır. Dolayısıyla karmaşıklık matrisi ROC eğrisi analizi ve genel performans değerlendirmesinin temel yapı taşını oluşturmaktadır.

6. MATERYAL ve METOT

Bu çalışmada, yüz tanıma sistemlerine yönelik gelişmiş sunum saldırılarını tespit etmek amacıyla fizyolojik ve görsel ipuçlarını bir arada değerlendiren hibrit bir derin öğrenme mimarisi önerilmiştir. Geliştirilen sistem, yüzün dinamik doku özelliklerini analiz eden görsel bir akış ile canlılık belirtilerini ölçen fizyolojik bir akıştan oluşan çift akışlı bir yapıya sahiptir. Bu bölümde önerilen mimarinin teknik detayları, kullanılan sinyal işleme algoritmaları ve veri ön işleme stratejileri ayrıntılı olarak açıklanmıştır.

6.1 Veri Setleri

Modelin farklı saldırı türleri ve çekim koşulları altındaki genellenebilirliğini değerlendirmek amacıyla bu çalışmada halka açık dört farklı veri seti kullanılmıştır. Seçilen veri kümeleri, basit fotoğraf baskılarından yüksek çözünürlüklü video tekrar oynatma saldırılarına ve karmaşık 3D maske senaryolarına kadar geniş bir PAI spektrumunu kapsamaktadır. Ayrıca veri setlerinin içerdiği değişken aydınlatma seviyeleri ve farklı kamera sensör özellikleri sistemin kontrollü laboratuvar ortamlarının ötesindeki performansını ölçmeyi mümkün kılmaktadır. Bu çeşitlilik sayesinde önerilen rPPG tabanlı canlılık tespiti yaklaşımının eğitim setinde yer almayan kişiler üzerindeki başarımı ve gerçek dünya koşullarındaki güvenilirliği kapsamlı bir biçimde analiz edilmiştir.

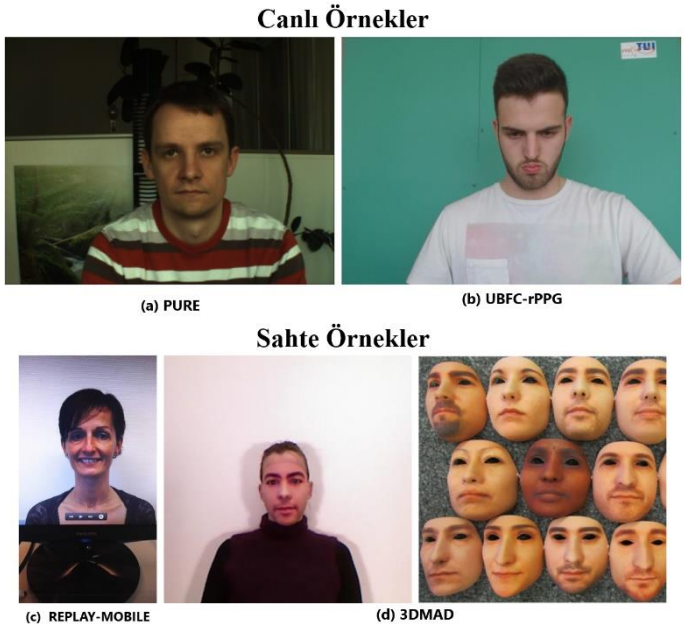
Bu veri setlerinden ilki olan Replay-Mobile, özellikle mobil cihazlar aracılığıyla gerçekleştirilen sunum saldırılarını hedef almaktadır. Veri seti 40 farklı denekten oluşmaktadır. Veri seti hem iç hem de dış mekanlarda farklı aydınlatma koşulları altında kaydedilmiş toplam 1.190 video içermektedir. Çekimler, LG-G4 ve iPad Air gibi dönemin yaygın akıllı telefon ve tablet kameraları kullanılarak gerçekleştirilmiştir. Saldırı senaryoları, deneklerin yüksek çözünürlüklü fotoğraflarının basılı kopyalarını ve ekranlarda oynatılan video kayıtlarını içermektedir. Replay-Mobile, mobil platformlardaki yüz tanıma sistemlerinin farklı ortam koşulları ve çeşitli 2B saldırı türleri karşısındaki dayanıklılığını ölçmek için tasarlanmış kapsamlı bir referans veri setidir (Costa-Pazo vd. 2016).

2B saldırılara ek olarak modelin gelişmiş sahtekârlık türlerine karşı performansını ölçmek amacıyla yüksek kaliteli 3D maske saldırılarını içeren 3DMAD (3D Mask Attack Database) veri seti de kullanılmıştır. Bu veri seti, özellikle yüksek kaliteli ve gerçekçi 3D maske saldırılarını içeren halka açık veri tabanlarından biridir. Çalışma, 17 farklı denekten toplanan video kayıtlarından oluşmaktadır. Her denek için hem gerçek erişim hem de kişiye özel olarak üretilmiş 3D maske kullanılarak gerçekleştirilen saldırı senaryoları kaydedilmiştir. Saldırı amacıyla kullanılan maskeler deneklerin ön ve yan profilden çekilmiş fotoğraflarının 3D modellemeye dönüştürülmesi ve sert reçine kompozit materyal kullanılarak renkli 3D yazıcı teknolojisiyle üretilmesi yoluyla elde edilmiştir. Bu üretim tekniği, maskelerin cilt dokusunu ve renk dağılımını yüksek gerçeklikle taklit etmesini sağlamaktadır. Videolar Microsoft Kinect v2 sensörü kullanılarak 640×480 çözünürlükte ve 30 FPS kare hızında kontrollü bir stüdyo ortamında kaydedilmiştir. Veri seti, modelin özellikle doku ve şekil olarak canlı bir yüze çok benzeyen gelişmiş sahtekarlık türlerini ne ölçüde ayırt edebildiğini test etmek açısından kritik bir öneme sahiptir (Erdogmus ve Marcel 2013).

Canlı sınıfı örneklerini zenginleştirmek ve modelin gürültüye karşı dayanıklılığını test etmek amacıyla uzaktan fotopletismografi algoritmaları için oluşturulmuş PURE (Pulse Rate Estimation) veri setinden de faydalanılmıştır. Bu veri seti, uzaktan fotopletismografi algoritmalarının doğruluğunu değerlendirmek amacıyla oluşturulmuştur. Ancak canlılık tespiti çalışmalarında canlı sınıfı örneklerini zenginleştirmek için de literatürde yaygın olarak kullanılmaktadır. Veri seti 8 erkek 2 kadın olmak üzere toplam 10 denekten toplanan 60 adet birer dakikalık video kaydından oluşmaktadır. Her denek için altı farklı senaryo kaydedilmiştir. Bunlar sabit durma, konuşma, yavaş baş çevirme, küçük ve büyük baş hareketleri ile hızlı baş çevirme senaryolarıdır. Videolar, endüstriyel bir RGB kamera ile sıkıştırılmamış formatta 640×480 çözünürlükte ve 30 FPS kare hızında kaydedilmiştir. Eş zamanlı olarak deneklerin parmağından bir nabız oksimetresi (CMS-50E) kullanılarak referans kalp atış hızı verileri de toplanmıştır. PURE veri seti, özellikle baş hareketlerinin rPPG sinyali üzerindeki etkilerini analiz etmek ve bu tür gürültülere dayanıklı modeller geliştirmek için değerli bir kaynak niteliğindedir (Stricker vd. 2014).

Benzer şekilde rPPG araştırmaları için tasarlanmış bir diğer önemli kaynak olan UBFC-

RPPG (University of Bourgogne Franche-Comté rPPG) veri seti de kullanılmıştır. Bu veri seti, rPPG sinyallerinin çıkarılmasına yönelik araştırmalar için tasarlanmış referans veri kümelerinden biridir. Veri seti 42 denekten oluşmaktadır ve her denekten yaklaşık bir dakika uzunluğunda video kaydedilmiştir. Çekimler, Logitech C920 HD Pro web kamerası ile 640 × 480 çözünürlükte ve 30 FPS kare hızında iç mekân aydınlatması altında gerçekleştirilmiştir. Deneklerden, kameranın yaklaşık 1 metre uzağında oturarak kalp atış hızlarını doğal bir şekilde artırmayı ve tipik bir insan-bilgisayar etkileşimi senaryosunu taklit etmeyi amaçlayan zamana duyarlı bir matematik oyunu oynamaları istenmiştir. Bu yaklaşım sadece dinlenme halindeki durumlarının yanı sıra aynı zamanda hafif zihinsel stres ve odaklanma anlarını da içeren daha dinamik ve pratik koşulları yansıtmaktadır. Nabız verileri, deneklerin bileğine takılan bir nabız oksimetresi (CONTEC CMS50E) aracılığıyla videolarla senkronize biçimde toplanmıştır. UBFC-RPPG veri seti, standart bir web kamerası ve doğal aydınlatma koşulları altında rPPG tabanlı sistemlerin performansını değerlendirmek için sıklıkla kullanılan bir referans veri setidir (Bobbia vd. 2019). Çalışmada kullanılan veri setlerinden örnek görsellere Şekil 6.1'de yer verilmiştir. Kullanılan tüm veri setlerinin içerik, sınıf dağılımı ve barındırdıkları saldırı türlerine ilişkin detaylı bilgiler Tablo 6.1'de sunulmuştur.



Şekil 6.1 Çalışmada kullanılan veri setlerinden örnek görüntüler a) Gerçek görüntü b) Gerçek görüntü c) Video saldırısı d) Maske Saldırısı

Tablo 6.1 Çalışmada kullanılan veri setlerinin genel özellikleri

Veri Seti	Denek Sayısı	Video Sayısı	Saldırı Türleri	Kullanılan Cihazlar	Çözünürlük ve FPS	Ortam / Senaryolar
Replay-Mobile (Costa-Pazo vd. 2016)	40	1.190	Basılı fotoğraf ve video tekrar oynatma saldırıları	LG-G4, iPad Air	1920x1080, 30 FPS	İç ve dış mekân, farklı ışık koşulları
3DMAD (Erdogmus ve Marcel 2013)	17	255	Kişiye özel üretilmiş 3D maske saldırıları	Microsoft Kinect v2 (RGB-D)	640x480, 30 FPS	Kontrollü stüdyo ortamı
PURE (Stricker vd. 2014)	10	60	Saldırı yok yalnızca gerçek erişim senaryoları	Endüstriyel RGB kamera ve PPG oksimetre	640x480, 30 FPS	Normal iç mekân. 6 farklı senaryo: konuşma, sabit durma, yavaş/küçük/büyük/ hızlı hareket
UBFC-RPPG (Bobbia vd. 2019)	42	42	Saldırı yok yalnızca gerçek erişim	Logitech C920 web kamera ve PPG oksimetre	640x480, 30 FPS	Normal iç mekân aydınlatması, matematik oyunu ile çeşitli stres seviyeleri

Önerilen modelin başarımını, tekil veri setlerinin sunduğu sınırlı varyasyonların ötesine taşıyarak daha heterojen ve zorlu bir test ortamında doğrulamak amacıyla bahsi geçen dört halka açık veri seti bütünleşik bir yapı altında birleştirilmiştir. Bu kapsamda modelin hem ekran tekrar oynatma gibi yaygın 2B saldırı senaryolarına hem de tespiti zor yüksek kaliteli 3D maske saldırılarına karşı direncinin ölçülmesi hedeflenmiştir. Ayrıca farklı sensör ve donanım özelliklerinin model üzerindeki etkisini analiz edebilmek amacıyla akıllı telefon kameraları, standart web kameraları ve yüksek çözünürlüklü sensörler ile kaydedilmiş videolar bir arada kullanılmıştır. Kontrollü ortam çeşitliliğine ek olarak PURE veri setindeki farklı yoğunluktaki baş hareketleri ve UBFC-RPPG veri setindeki doğal insan-bilgisayar etkileşimi senaryoları sürece dahil edilmiştir. Bu sayede modelin

stüdyo koşullarının dışına çıkarak gerçek dünya kullanımında kaçınılmaz olan hareket ve aydınlatma değişimleri altındaki performansı değerlendirilmiştir. Sonuç olarak oluşturulan bu heterojen veri havuzu modelin farklı saldırı türleri, sensör kaliteleri ve çevresel değişkenler karşısındaki genelleme kabiliyetinin bütünsel bir bakış açısıyla analiz edilmesini sağlamıştır.

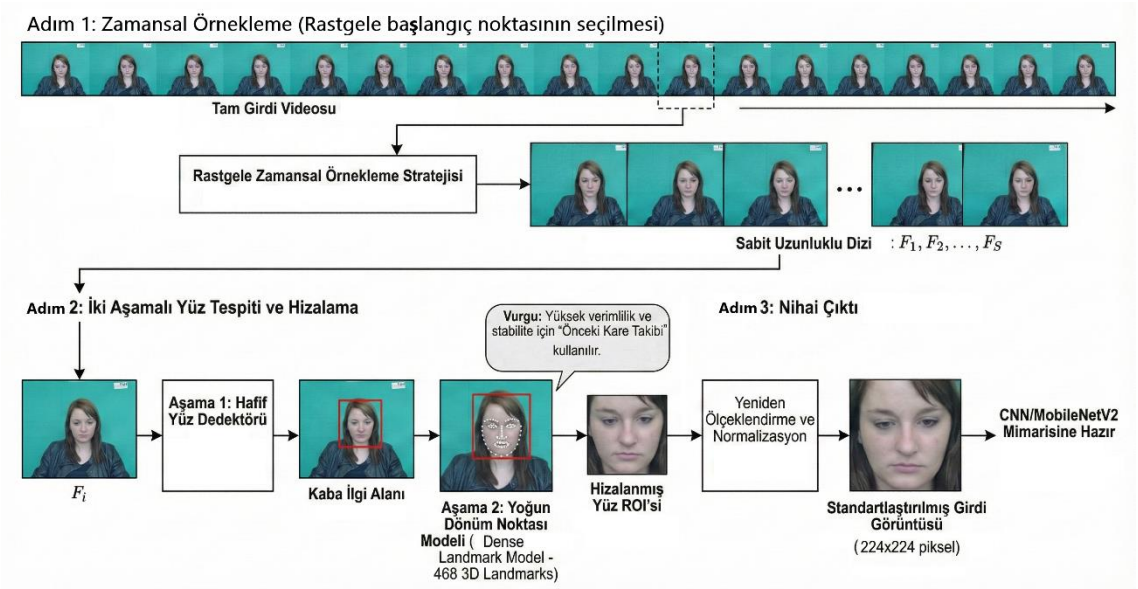
Çalışmada kullanılan tüm video örnekleri, deneysel sürece dahil edilmeden önce "saldırı" ve "canlı" olmak üzere iki ana sınıfa ayrılmıştır. Sınıflandırma sürecinde saldırı sunumları 0 ve canlı sunumlar 1 olacak şekilde ikili bir kodlama yapısı benimsenmiş ve bu etiketleme şeması tüm deneysel protokol boyunca korunmuştur. Sınıf ayrımı veri setlerinin orijinal protokol yapıları dikkate alınarak gerçekleştirilmiştir. Replay-Attack ve 3DMAD veri kümelerindeki basılı fotoğraf, video yeniden oynatma ve maske içerikli kayıtlar saldırı sınıfına, gerçek yüz kayıtları ise canlı sınıfına atanmıştır. PURE ve UBFC-RPPG veri kümeleri yalnızca canlı kayıtlar içerdiğinden bu veri kümelerindeki tüm örnekler canlı sınıfına dahil edilmiştir. Belirlenen etiketler hem görüntü tabanlı akışta hem de fizyolojik akışta ortak hedef değişken olarak modele sunulmuştur. Veri etiketleme adımının tamamlanmasının ardından modelin eğitim setinde görmediği kişilere karşı başarısını değerlendirebilmek amacıyla kişi tabanlı katmanlı 5-katlı çapraz doğrulama protokolü uygulanmıştır. Bu yöntemde, aynı kişiye ait video örneklerinin hem eğitim hem test kümelerine dağılmasını engellemek için örnekler öncelikle kişi kimliklerine göre gruplandırılmıştır. Ardından her bir gruptaki saldırı-canlı dağılım oranı istatistiksel olarak korunarak veri seti beş eşit parçaya ayrılmıştır. Böylece her katta hem sınıf dengesi korunmuş hem de kişi temelli tam ayrışma sağlanmıştır. Bu yaklaşım modelin aynı yüzü farklı koşullarda tekrar görerek yapay bir başarım artışı elde etmesini engellemekte ve sahne farklılıklarına karşı genelleme yeteneğini artırmaktadır.

Eğitim süreci her iterasyonda yalnızca ilgili eğitim seti üzerinde gerçekleştirilmiştir. Model, her bir örneğin sınıfını öğrenirken görüntü akışı ve fizyolojik akış hibrit bir yapıda değerlendirmiştir. Karar aşamasında ise görsel akıştan elde edilen uzamsal-zamansal özellikler ile fizyolojik akıştan elde edilen özellikler füzyon katmanında birleştirilerek nihai canlılık kararı üretilmiştir. Fold dışındaki test setinde yapılan değerlendirmeler ile modelin tüm kişi grupları üzerindeki performansı objektif bir biçimde ölçülmüştür.

Uygulanan bu bütünleşik deneysel yapı çalışmanın deneysel tutarlılığını, yeniden üretilebilirliğini ve farklı veri kaynaklarından gelen yüzlere karşı genelleme başarımını güvence altına almaktadır.

6.2 Ön İşleme Adımları

Önerilen çift akışlı mimarinin hem görsel hem de fizyolojik analiz kollarında ihtiyaç duyduğu veri kalitesini sağlamak amacıyla bütünleşik ve çok katmanlı bir ön işleme hattı tasarlanmıştır. Modelin çevresel gürültülere karşı dayanıklılığını artırmak ve sinyal-gürültü oranını optimize etmek amacı ile dört temel adımdan oluşan bir önişleme süreci gerçekleştirilmiştir. Bu süreçte ardışık olarak videolardan karelerin örneklenmesi, yüz tespitinin yapılması, hizalanması, normalize edilmesi ve özelleştirilmiş ilgi alanı çıkarılması işlemleri yapılmaktadır. Ön işleme sürecinin genel akış şeması Şekil 6.2'de sunulmuştur.



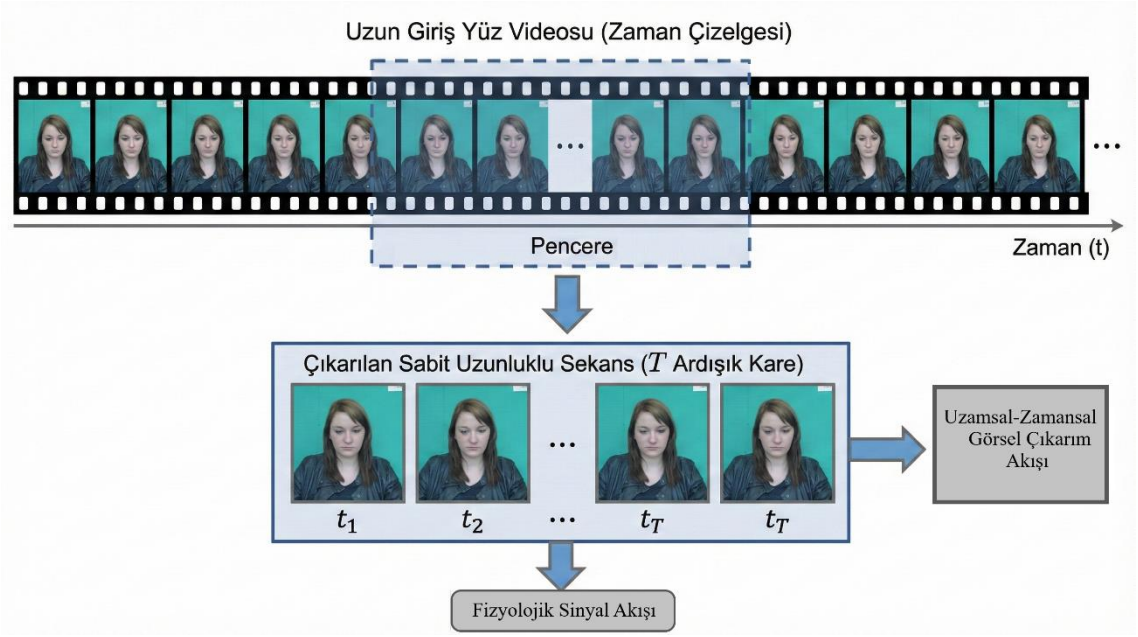
Şekil 6.2 Önerilen hibrit model için uygulanan ön işleme akış şeması

6.2.1 Zamansal Örnekleme ve Kare Seçimi

Kullanılan veri setlerinde yer alan videoların süreleri değişkenlik göstermektedir. Geliştirilen derin öğrenme modelinin standartlaştırılmış bir girdi tensörü ile çalışabilmesi

için her bir videodan sabit uzunlukta (T) ardışık kare dizilerinin örneklenmesi gerekmektedir. Önerilen modelin hem görsel akıştaki mikro-hareket desenlerini hem de fizyolojik akıştaki rPPG sinyalinin yaklaşık 1 ila 3 kalp atım döngüsünü kapsayan periyodik yapısını öğrenebilmesi amacıyla bu sabit dizi uzunluğu 90 kare olarak optimize edilmiştir. 30 FPS kayıt hızında bu değer 3 saniyelik kesintisiz bir zamansal pencereye karşılık gelmektedir.

Eğitim aşamasında veri çeşitliliğini artırmak ve modelin videonun belirli bir zaman aralığına aşırı uyum sağlamasını engellemek amacıyla rastgele zamansal örnekleme stratejisi uygulanmıştır. Bu yöntemde, her eğitim iterasyonunda video içerisinden rastgele seçilen bir başlangıç karesinden itibaren ardışık 90 karelik bir segment alınmakta ve modele girdi verisi olarak sunulmaktadır. Bu yaklaşım, modelin videonun tamamı boyunca farklı zamansal öznitelikleri öğrenmesini sağlayarak genelleme yeteneğini güçlendirmektedir. Uygulanan bu adımlar Şekil 6.3'te gösterilmiştir. Bu yaklaşım, modelin videonun tamamı boyunca farklı zamansal öznitelikleri öğrenmesini sağlayarak genelleme yeteneğini güçlendirmektedir. Uygulanan bu adımlar Şekil 6.3'te gösterilmiştir.



Şekil 6.3 Rastgele zamansal örnekleme ve sabit boyutlu pencereleme stratejisi

6.2.2 Yüzey Geometrisinin Modellenmesi

Bu çalışmada yüz tespiti, yüz hizalama ve yoğun nirengi noktalarının çıkarılması amacıyla Google MediaPipe Face Mesh kütüphanesi kullanılmıştır. Bu sistem, tek bir RGB görüntüsü üzerinden eş zamanlı olarak hem yüz konumunu hem de yaklaşık 3D yüzey geometrisini tahmin edebilen hafif ve gerçek zamanlı bir makine öğrenmesi tabanlı yüz modelleme yaklaşımıdır. Model, derin sinir ağı tabanlı iki aşamalı bir işlem hattı üzerinden çalışmaktadır. Her video karesi için yüz anatomisini yüksek doğrulukla temsil eden 468 adet yoğun yüz işaretleyicisini tahmin etmektedir.

Sistemin çalışma prensibi, hesaplama verimliliğini maksimize edecek şekilde tasarlanmış ardışık bir yapıdan oluşmaktadır. İlk aşamada, görüntü üzerinde yüzün genel konumunu belirleyen hafif mimarili bir yüz tespit ağı çalışmaktadır. Mobil GPU çıkarımı için optimize edilmiş bu ağ, yüzün sınır kutusunu ve temel oryantasyonunu tahmin etmenin yanı sıra sonraki aşama için referans çerçeveyi de oluşturmaktadır. Bu erken tespit mekanizması, özellikle video işleme uygulamalarında her karede tüm görüntünün taranması ihtiyacını ortadan kaldırarak hesaplama yükünü önemli ölçüde azaltmaktadır.

İkinci aşamada, tespit edilen ilgili bölge daha karmaşık bir derin sinir ağına yönlendirilmektedir. Dikkat mekanizması ile güçlendirilmiş bu ağ, yüz yüzeyini yüksek çözünürlüklü bir nirengi ağı şeklinde temsil eden 468 yoğun nirengi noktasını regresyon yöntemiyle tahmin etmektedir. Her bir nirengi noktası, iki boyutlu piksel koordinatlarına ek olarak, yüzün üç boyutlu yapısını temsil eden ve referans noktaya göre normalize edilmiş bir derinlik bileşeni ile modellenmektedir. Bu üç boyutlu modelleme yeteneği sayesinde yüzün geometrik yapısı, mimik değişimleri ve mikro ifadeler dâhil olmak üzere birçok ince detay yüksek hassasiyetle yakalanabilmektedir.

Önerilen iki aşamalı mimari, standart donanımlarda dahi gerçek zamanlı olarak çalışabilmektedir. Bu mimari aynı zamanda yüksek geometrik hassasiyete sahip bir yüz temsili sağlamaktadır. Bununla birlikte, modelin başarımı için en kritik faktörlerden biri tahmin edilen nirengi noktalarının sergilediği zamansal kararlılığıdır. MediaPipe Face Mesh'in sunduğu düşük titreme oranı bu çalışmada önerilen görsel ve fizyolojik akış

stratejilerinin tutarlı ve senkronize bir biçimde uygulanmasına olanak tanımıştır. Özellikle baş rotasyonları, mimik değişimleri veya değişken aydınlatma gibi dinamik koşullar altında dahi nirengi noktalarının konumsal kararlılığını koruması hareket kaynaklı gürültülerin minimize edilmesi ve rPPG sinyal çıkarımının güvenilirliği açısından hayati bir rol oynamaktadır.

6.2.3 Yüz Tespiti ve Geometrik Hizalama

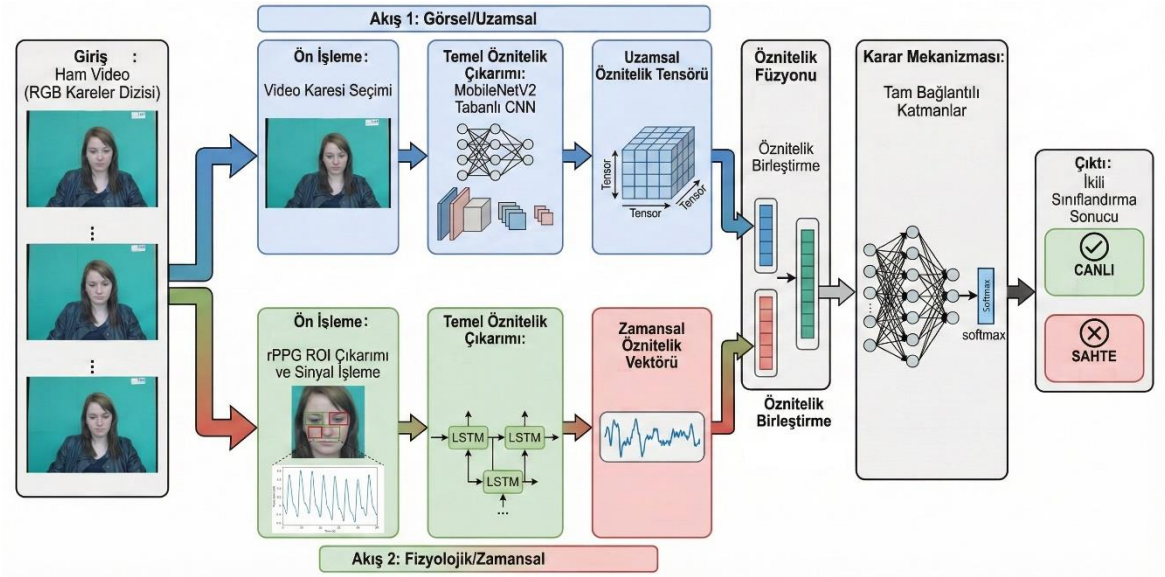
Çalışmada önerilen uzamsal-zamansal görsel öznitelik ve rPPG tabanlı fizyolojik sinyal akışlarında, ilk aşama olarak veri ön işleme adımı gerçekleştirilmiştir. Bu kapsamda, ham video kareleri üzerinden yüz bölgesinin tespiti ve standardizasyonunu sağlamak amacıyla iki kademeli bir mimari kullanılmıştır. İlk aşamada, görüntü içerisindeki yüzün kaba konumu ve sınırları tespit edilmektedir. Ancak yüzün kamera açısına göre eğik olması veya kişinin baş hareketleri rPPG sinyalinde hareket kaynaklı gürültülere neden olabilmektedir. Bu problemi minimize etmek ve elde edilen özellikleri standardize etmek amacıyla ikinci aşamada tespit edilen yüz bölgesine geometrik hizalama işlemi uygulanmıştır.

Hizalama süreci, yüz üzerindeki göz bebekleri, ağız köşeleri gibi belirgin anatomik nirengi noktalarını referans alan bir afin dönüşüm matrisi kullanılarak gerçekleştirilmektedir. Tespit edilen yüz bölgesi, bu referans noktaları temel alınarak kanonik bir düzleme oturtulmakta ve modelin giriş boyutuna uygun olacak şekilde 224×224 piksel çözünürlüğüne yeniden ölçeklendirilmektedir. Uygulanan bu standardizasyon işlemi, hem CNN tabanlı modelin görsel öznitelikleri daha kararlı ve tutarlı bir biçimde öğrenmesini sağlamakta hem de rPPG sinyalinin hareket kaynaklı bozunumlardan arındırılmasına katkı sunarak sinyal-gürültü oranının artırılmasına yardımcı olmaktadır.

6.3 Hibrit Model Mimarisi

Bu tez kapsamında önerilen model, yüz canlılık tespiti problemini hem dokusal hem de fizyolojik boyutlarıyla ele alan, tamamlayıcı bilgi kaynaklarını paralel olarak işleyebilen

çift akışlı hibrit bir derin öğrenme mimarisi üzerine inşa edilmiştir. Geliştirilen bu bütünleşik yapı, sunum saldırılarının neden olduğu anomalileri görsel doku bozulmaları ve zaman içerisinde ortaya çıkan fizyolojik tutarsızlıklar üzerinden analiz edebilmektedir. Bu doğrultuda mimari, Şekil 6.4'te gösterildiği üzere iki ana alt akış ve bu akışlardan elde edilen öznitelikleri birleştiren bir füzyon katmanından oluşmaktadır.

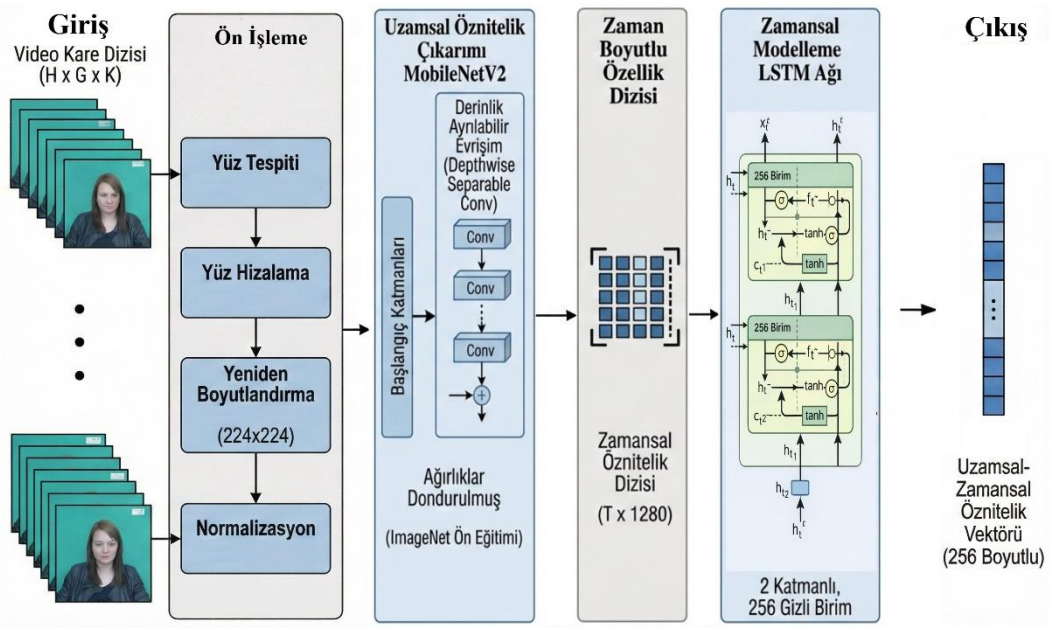


Şekil 6.4 Çift akışlı hibrit derin öğrenme mimarisi

Uzamsal-zamansal görsel akış video karelerindeki doku deformasyonlarını, yüzey yansımalarını ve yapay kenar bilgilerini uzamsal boyutta analiz etmektedir. Aynı zamanda bu görsel ipuçlarının zaman içerisindeki değişimlerini modelleyerek fiziksel kökenli anomalilerin tespit edilmesini sağlamaktadır. Fizyolojik sinyal akışı ise doğrudan yüz derisindeki mikroskobik renk dalgalanmalarına odaklanarak kalp atış döngüsüne bağlı kan hacmi değişimlerini analiz etmekte ve canlı dokuya özgü nabız bilgisini ayrıştırmaktadır. Bu iki akışın bütünleşik biçimde kullanılması, modelin hem görsel sahneye ait yapısal ipuçlarını hem de bireye özgü fizyolojik biyometrik sinyalleri eş zamanlı olarak değerlendirmesine olanak tanımaktadır. Her iki akıştan elde edilen derin öznitelik vektörleri öznitelik füzyonu katmanında birleştirilmekte ve nihai Canlı / Saldırı kararı tam bağlı bir sınıflandırıcı katmanı tarafından üretilmektedir.

6.3.1 Uzamsal-Zamansal Görsel Çıkarım Akışı

Uzamsal-zamansal görsel çıkarım akışı sunum saldırısı araçlarında sıklıkla gözlemlenen ve insan gözüyle algılanması güç olan dokusal bozulmaları, anormal yüzey yansımalarını ve doğal olmayan mikro-hareketleri tespit etmek üzere tasarlanmıştır. Bu akışa ait mimari yapı ve veri işleme aşamaları Şekil 6.5'te ayrıntılı olarak gösterilmektedir.



Şekil 6.5 Uzamsal-zamansal görsel çıkarım akışı

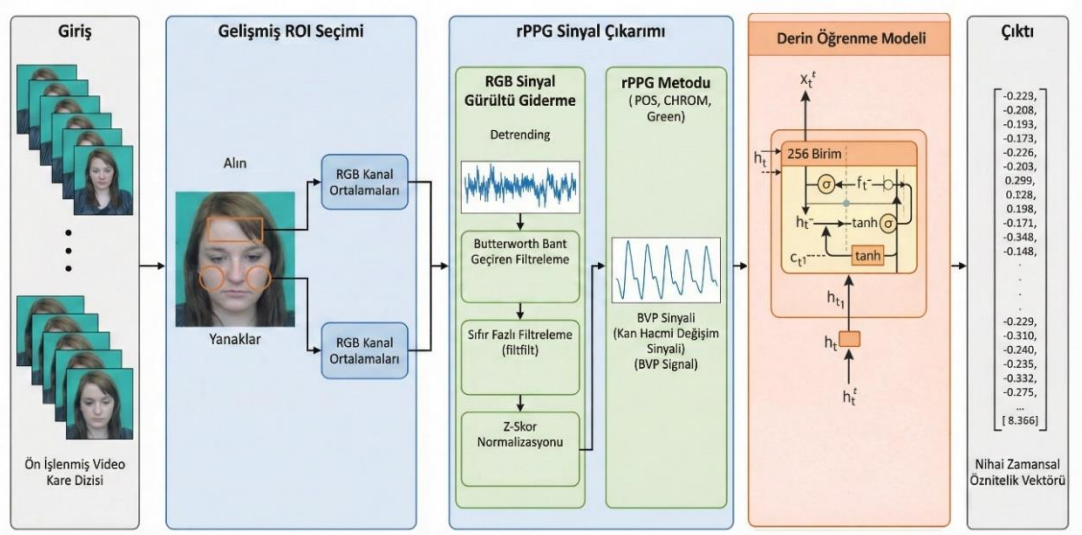
Sistem, ön işleme aşamasında yüz tespit algoritması kullanılarak video karelerinden izole edilen yüz bölgesini temel almaktadır. Derin öğrenme mimarisinin giriş tensör gereksinimlerini karşılamak amacıyla tespit edilen ilgili yüz alanı 224×224 piksel boyutlarına yeniden ölçeklendirilmiş ve normalize edilmiştir. Görsel özniteliklerin çıkarılması sürecinde hesaplama maliyeti ile sınıflandırma başarımı arasındaki denge gözetilerek MobileNetV2 mimarisi omurga ağ olarak tercih edilmiştir. Bu ağ, ImageNet veri seti üzerinde önceden eğitilmiş ağırlıklar ile başlatılmış ve transfer öğrenme metodolojisi yaklaşımı benimsenmiştir.

Modelin genelleme yeteneğini korumak ve aşırı öğrenmeyi önlemek amacıyla ağın evrişim katmanları dondurulmuş, mimari yalnızca bir öznitelik çıkarıcı olarak kullanılmıştır. Her bir video karesi için sınıflandırıcı katmanından önceki son global

havuzlama katmanından elde edilen 1280 boyutlu derin öznitelik vektörü çıkarılmıştır. Ancak statik karelerdeki uzamsal doku bilgisinin yanı sıra sunum saldırılarının zaman içindeki hareket dinamiklerini de öğrenilmesi kritik öneme sahiptir. Bu bağlamda ardışık video karelerinden elde edilen CNN öznitelik dizisi zamansal bağlamı modellemek üzere LSTM ağına aktarılmıştır. Bu çalışmada kullanılan LSTM modülü 2 katmanlı ve 256 gizli birimli bir yapıya sahiptir. Bu yapı sayesinde video boyunca değişen uzamsal öznitelikler arasındaki zamansal korelasyon öğrenilmekte ve sunum saldırılarına özgü hareket temelli anomaliler etkin bir biçimde ayrıştırılmaktadır.

6.3.2 Fizyolojik Sinyal Akışı ve Karşılaştırmalı ROI Stratejisi

Yüz canlılık tespitinde cansız sunum saldırısı araçlarının otonom sinir sistemi tarafından yönetilen dinamik bir nabız sinyali üretemeyeceği gerçeğinden hareketle uzaktan fotopletismografi analizi sisteme tamamlayıcı bir güvenlik katmanı olarak entegre edilmiştir. Önerilen modelin deneysel tasarım aşamasında görsel akış ve fizyolojik akış için kullanılan ilgi alanları farklı stratejilerle ele alınmıştır. Görsel öznitelik çıkarım akışında sahte doku detaylarının, maske kenar hatalarının ve cansızlık anomalilerinin yakalanabilmesi kritik öneme sahiptir. Bu nedenle fizyolojik akışta maskeleme yöntemleri kullanılsa dahi görsel akışta veri kaybını önlemek amacıyla tüm yüz bölgesi sabit bir girdi olarak kullanılmış ve ROI stratejisi bu kolda değiştirilmemiştir. Buna karşılık fizyolojik akışta seçilen yüz bölgesi rPPG sinyalinin kalitesini doğrudan etkilemektedir. Yüzün farklı anatomik bölgeleri hem kılcal damar yoğunluğu hem de mimik kaslarının yarattığı hareket gürültüsü açısından heterojen bir yapı sergilemektedir. Ancak rPPG sinyalinin başarımı yalnızca seçilen ROI stratejisine ile sınırlı kalmamaktadır. Aynı zamanda bu bölgeden nabız sinyalini ayrıştırmak için kullanılan sinyal işleme algoritmasına da doğrudan bağlıdır. Bu bağlamda fizyolojik sinyal akışının performansını optimize etmek ve modelin kararlılığını ölçmek amacıyla çok katmanlı bir karşılaştırma stratejisi izlenmiştir. Söz konusu fizyolojik sinyal akışının genel mimarisi ve maskeleme tabanlı ROI seçimi Şekil 6.6'da şematik olarak sunulmuştur.



Şekil 6.6 Fizyolojik sinyal akışı ve anatomik segmentasyon tabanlı seçici ROI seçimi

6.3.2.1 Bütünsel Yüz ROI Yaklaşımı

Çalışmanın ilk aşamasında, fizyolojik sinyal çıkarımı için literatürdeki en temel yöntem olan bütünsel yüz ROI stratejisi uygulanmıştır. Bu yaklaşımda, yüz tespit algoritması tarafından belirlenen sınır kutusu içerisindeki alın, yanaklar, göz çevresi, ağız ve burun bölgesinin tamamı analize dahil edilmektedir. İzole edilen bölge içerisindeki tüm yüz piksellerinin uzamsal ortalaması alınarak ham RGB zaman serileri elde edilmektedir.

Bu yöntemin görsel akış açısından en belirgin üstünlüğü yüzün tamamına ait ince dokusal ayrıntıların eksiksiz biçimde korunmasıdır. Yüzün tamamının analize dahil edilmesi sayesinde model yüksek çözünürlüklü doku desenlerini, gözenek yapısını, cilt geometrisini, cilt yüzeyindeki yansımaları ve farklı materyallere özgü mikro-yapısal tutarsızlıkları daha etkin biçimde öğrenebilmektedir. Bu durum, derin öğrenme mimarisinin görsel öznitelik uzayını önemli ölçüde zenginleştirmektedir. Ancak aynı yaklaşım fizyolojik sinyal çıkarımı açısından ele alındığında önemli kısıtlar barındırmaktadır. Bütünsel ROI yönteminde, kan akışı bilgisi barındırmayan göz küresi, dişler ve dudak boşluğu gibi anatomik bölgeler de sinyal ortalamasına zorunlu olarak dahil edilmektedir. Özellikle göz kırpma, dudak ve çene hareketleri ile mimik değişiklikleri gibi fizyolojik olmayan hareketler rPPG sinyali üzerinde ani genlik dalgalanmalarına neden olmaktadır. Bu tür hareket artefaktları, sinyal-gürültü oranını

düşürerek gerçek nabza bağlı ince renk değişimlerin maskelenmesine yol açmaktadır. Bu durum fizyolojik akış tarafından üretilen sinyalin zamansal kararlılığını azaltmaktadır.

Bununla birlikte bu aşamada elde edilen ham RGB sinyalleri solunum hareketleri, istemsiz kas hareketleri ve kamera sensör gürültüsü gibi fizyolojik olmayan bileşenleri de barındırmaktadır. Bu gürültü unsurları ve ortam aydınlatmasındaki değişimler nedeniyle ham veriler doğrudan nabız analizi için elverişli bir yapı sunmamaktadır. Bu nedenle sinyalin analiz edilebilir hale getirilmesi ve saf BVP bilgisinin ayrıştırılması gerekmektedir. Bu amaçla elde edilen ham veriler ilerleyen bölümlerde detaylandırılan rPPG çıkarım algoritmalarına ve çok aşamalı sinyal işleme filtrelerine girdi vektörü olarak aktarılmaktadır.

6.3.2.2 Anatomik Segmentasyon Tabanlı Seçici (Maskeli) ROI Yaklaşımı

Fizyolojik sinyal analizinde sinyal-gürültü oranını artırmak ve mimik kaynaklı hareket gürültülerini elimine etmek amacıyla anatomik segmentasyon tabanlı seçici ROI stratejisi geliştirilmiştir. Fizyolojik sinyaldeki mimik kaynaklı gürültüyü elimine etmek amacıyla Google MediaPipe Face Mesh kütüphanesi kullanılarak geliştirilmiş bir maskeleme yaklaşımı uygulanmıştır. Geliştirilen yaklaşımda, kas hareketlerinin nispeten sınırlı kaldığı ve aynı zamanda yüz anatomisi üzerinde kılcal damar yoğunluğunun yüksek olduğu bölgeler seçilmiştir. Bu doğrultuda rPPG sinyal çıkarımı için alın, sol yanak ve sağ yanak olmak üzere üç temel anatomik bölge seçilmiştir. Alın bölgesi, kılcal damar yoğunluğunun yüksek olması ve ağız/çene kaynaklı mimik hareketlerinden görece daha az etkilenmesi nedeniyle tercih edilmiştir. Bu bölgenin sınırlarının belirlenmesinde 10, 338, 297 ve 332 numaralı nirengi indeksleri referans alınmıştır. Sol ve sağ yanak bölgeleri ise elmacık kemikleri üzerindeki deri dokusunun kanlanma açısından zengin olması ve ışık yansımasına karşı daha kararlı bir yapı sunması nedeniyle seçilmiştir. Sol yanak için 116 ve 117, sağ yanak için ise 345 ve 346 numaralı indeksler kullanılmıştır. Belirlenen bu nirengi setleri yardımıyla her bir anatomik bölge için konveks çokgenler oluşturulmuş ve bu çokgenler üzerinden ikili maskeleme işlemi uygulanmıştır. Maskeleme aşamasında seçilen ROI alanları 1 değerini alacak şekilde beyaz olarak işaretlenmiştir. rPPG sinyalini olumsuz etkileyebilecek gözler, ağız, burun delikleri ve saç çizgisi gibi gürültü üreten

bölgeler ise 0 değerine karşılık gelecek biçimde siyah olarak maske dışına bırakılmıştır. Bu yaklaşım çıkarılan rPPG sinyalinin fizyolojik içeriğini güçlendirirken hareket ve aydınlatma kaynaklı bozulmaların önemli ölçüde azaltılmasını sağlamıştır.

rPPG sinyal çıkarımı yalnızca bu maskelenmiş bölgelerdeki piksellerin ağırlıklı ortalaması alınarak gerçekleştirilmiştir. Bu yaklaşım sayesinde kişinin konuşması veya mimik yapması durumunda dahi kararlı ve gürültüden arındırılmış bir nabız sinyali elde edilmesi amaçlanmıştır. Buna karşılık görsel özellik akışında bütünsel yüz ROI yaklaşımı korunmuştur. Bütünsel yüz ROI'nin tercih edilmesinin nedeni PAI tespitinde kritik öneme sahip göz kırpma tutarsızlıkları, göz çerçevesi hareketleri, dudak deformasyonları, maske kenar geçişleri ve cilt dokusuna ilişkin anomaliler gibi pek çok görsel ipucunun yüzün tamamına yayılarak ortaya çıkmasıdır. Görsel akışta maskelenmiş bir ROI kullanılması durumunda sahteciliğin belirlenmesini sağlayan bu önemli ipuçlarının önemli bir kısmı modelden soyutlanacak ve tespit performansı olumsuz yönde etkilenecektir.

Maskeleme ve Sinyal Çıkarım Algoritması

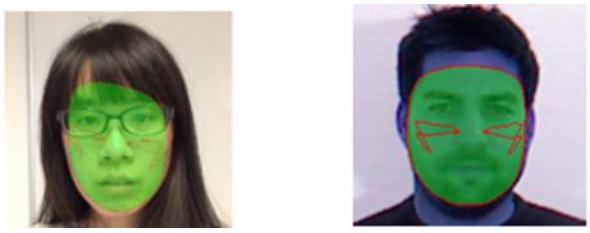
Önerilen yöntemin algoritmik işleyişi, öncelikle I_t anlık görüntü karesi üzerinde 468 adet üç boyutlu nirengi noktasının (x_i, y_i) tespit edilmesiyle başlamaktadır. Bu işlemin ardından seçilen nirengi indeks kümeleri (S_{ROI}) için görüntü düzleminde dışbükey çokgenler oluşturulmakta ve görüntü boyutlarında $(H \times W)$ bir maske matrisi (M_{maske}) üretilmektedir. Elde edilen maske matrisi üzerinde, ilgi alanı (ROI) içerisinde kalan pikseller aktif (1), bu alanların dışında kalan pikseller ise pasif (0) olacak şekilde ikili maskeleme işlemi uygulanmaktadır. Bu işlem matematiksel olarak Denklem 6.1'de belirtilen şekilde tanımlanmaktadır.

$$M_{maske}(x, y) = \begin{cases} 1, & (x,y) \in ROI \\ 0, & diğer \end{cases} \tag{6.1}$$

Maske üretiminin ardından yalnızca maske altında kalan pikseller dikkate alınarak her bir renk kanalı için uzamsal ortalama RGB sinyalleri hesaplanmaktadır. Bu işlem Denklem 6.2'de gösterildiği şekilde ifade edilmektedir.

$$C_{avg}(t) = \frac{\sum_{x,y} I_t(x,y) \cdot M_{maske}(x,y)}{\sum_{x,y} M_{maske}(x,y)},\ C \in \{R, G, B\} \quad (6.2)$$

Bu işlem sayesinde göz kırpma, dudak hareketleri ve arka plan gürültüleri sinyal kaynağından fiziksel olarak ayrıştırılmaktadır. Uygulanan maskeleme ve ilgi alanı çıkarma algoritmasının etkinliğini görselleştirmek amacıyla maskeli olarak işlenmiş örnek görüntüler incelenmiştir. Bu görüntülerde, modelin fizyolojik sinyali çıkarma sırasında odaklandığı anatomik bölgeler ve maskeleme sonrası gürültüden arındırılmış alanlar net bir şekilde gözlemlenmektedir. Gelişmiş ROI stratejisi kapsamında elde edilen örnek maskeli görüntüler Şekil 6.7'de sunulmuştur.



Şekil 6.7 Gelişmiş ROI stratejisi kapsamında elde edilen örnek maskeli görüntüler

6.3.2.3 rPPG Sinyal Çıkarımı ve Sinyal İşleme

Fizyolojik analiz sürecinde seçilen ROI bölgelerinden elde edilen ham RGB sinyalleri doğrudan saf nabız bilgisini temsil etmemektedir. Bu sinyaller ortam aydınlatması değişimleri, istemsiz mimik hareketleri, solunum etkileri ve sensör kaynaklı gürültüler gibi fizyolojik olmayan bileşenler içermektedir. Bu nedenle elde edilen verilerin modele uygun gürültüden arındırılmış ve zamansal olarak kararlı bir yapıya dönüştürülmesi gerekmektedir. Bu amaçla öncelikle sinyal çıkarım algoritmaları uygulanarak ham rPPG sinyali elde edilmiş ardından çok aşamalı bir sinyal işleme hattı ile bu sinyal filtrelenmiştir.

rPPG Sinyal Çıkarım Yöntemleri ve Algoritmik Karşılaştırma

Elde edilen ham RGB sinyallerinden saf nabız bilgisini ayrıştırmak için kullanılan sinyal işleme yöntemi, modelin sunum saldırısı tespiti başarımı üzerinde kritik bir etkiye

sahiptir. Bu çalışmada fizyolojik sinyal kalitesinin ve seçilen algoritmanın model performansına etkisini sistematik biçimde analiz edebilmek amacıyla her iki ROI stratejisi için de Yeşil Kanal, CHROM ve POS algoritmaları ayrı ayrı uygulanmış ve karşılaştırmalı bir performans değerlendirmesi gerçekleştirilmiştir.

Kullanılan algoritmaların temel çalışma prensipleri incelendiğinde en temel yaklaşım olan Yeşil Kanal yöntemi, oksihemoglobinin yeşil ışığı diğer dalga boylarına kıyasla daha yüksek oranda soğurması prensibine dayanmaktadır. Bu yöntemde yalnızca yeşil kanalın uzamsal ortalaması rPPG sinyali olarak kabul edilmektedir. Hesaplama maliyeti oldukça düşük olmakla birlikte ortam aydınlatması değişimlerine karşı hassas olmaktadır.

Buna karşılık CHROM yöntemi, cilt yüzeyindeki speküler yansımaları elimine etmek amacıyla renk farkı sinyallerini kullanan istatistiksel bir yaklaşımdır. Yöntem standart bir cilt tonu varsayımı üzerinden RGB kanallarının doğrusal kombinasyonunu kullanmaktadır. Bu sayede hareket gürültülerine karşı tek kanallı yöntemlere kıyasla daha dirençli bir yapı sunmaktadır.

Daha gelişmiş bir algoritma olan POS yöntemi ise RGB renk uzayında tanımlanan cilde dik olan bir düzlem üzerine projeksiyon yaparak nabız sinyalini gürültüden matematiksel olarak ayrıştırmaktadır. Bu yöntem farklı cilt tonu varyasyonlarına ve aydınlatma değişimlerine karşı daha kararlı bir başarım sergilemektedir. Söz konusu algoritmalar aracılığıyla elde edilen ham rPPG sinyalleri solunum, istemsiz kas hareketleri ve sensör gürültüleri içerdiğinden bir sonraki aşamada detaylandırılan sinyal işleme modülüne aktarılmaktadır.

Sinyal Filtreleme ve Ön İşleme

Sinyal çıkarım algoritmaları aracılığıyla elde edilen ham rPPG sinyalleri (S_{raw}), solunum hareketleri, istemsiz kas aktiviteleri ve sensör kaynaklı gürültüler gibi fizyolojik olmayan bileşenler içermektedir. Bu nedenle, sinyalin analiz edilebilir ve kararlı bir yapıya kavuşturulması amacıyla çok aşamalı bir filtreleme hattı uygulanmıştır. Uygulanan temel adımlar aşağıda ayrıntılı olarak açıklanmaktadır.

Eğilim Giderme

rPPG sinyalleri, çevresel aydınlatmadaki yavaş değişimler nedeniyle sıklıkla düşük frekanslı eksen kaymalarına maruz kalmaktadır. Fizyolojik nabız bileşenlerini bu aydınlatma artefaktlarından izole etmek için sinyale eğilim giderme işlemi gerçekleştirilmiştir. İlk adımda, ham sinyalin genel eğilim bileşeni matematiksel olarak hesaplanmaktadır. Daha sonra bu eğilim değeri orijinal ham sinyalden çıkarılarak nabız dalgası gürültüden arındırılmaktadır. İşlem matematiksel olarak Denklem 6.3 ile ifade edilmektedir.

$$S_{detrend}(t) = S_{raw}(t) - \lambda(t) \tag{6.3}$$

Denklem 6.3'te $S_{detrend}(t)$ eğilimden arındırılmış sinyali, $S_{raw}(t)$ ham rPPG sinyalini $\lambda(t)$ ise sinyale uygulanan yumuşatma operatörü ile elde edilen düşük frekanslı trend bileşenini temsil etmektedir.

Butterworth Bant Geçiren Filtreleme

Eğilimden arındırılmış sinyal, insan kalp atış hızının fizyolojik sınırlarına odaklanmak amacıyla bant geçiren filtreleme işlemine tabi tutulmuştur. Bu çalışmada, yetişkin bireyler için kabul edilen 40-180 BPM aralığına karşılık gelen 0.67-3.0 Hz frekans bandı referans alınmıştır. Bu amaçla düz geçiş bandı karakteristiği sayesinde sinyal formunu bozmadan filtreleme yapabilen 2. dereceden Butterworth bant geçiren filtre kullanılmıştır. Filtrenin frekans uzayındaki genlik transfer fonksiyonu Denklem 6.4'te verilmiştir.

$$|H(j\omega)|^2 = \frac{1}{1+\left(\frac{\omega}{\omega_C}\right)^{2n}} \tag{6.4}$$

Denklem 6.4'te n filtrenin derecesini, ω sinyalin açısal frekansını ve ω_c ise fizyolojik nabız frekans aralığını hedefleyecek şekilde belirlenen kesim frekansını temsil etmektedir. Bu filtreleme adımı sayesinde kalp atımına karşılık gelen frekans bileşenleri

korunurken hareket ve sensör kaynaklı istenmeyen bileşenler etkin biçimde bastırılmaktadır.

Sıfır Fazlı Filtreleme

Biyometrik sinyallerin analizinde filtreleme işlemi sırasında oluşabilecek faz kaymaları sinyalin zamansal yapısını bozarak kalp atım tepe noktalarının hatalı tespit edilmesine neden olabilmektedir. Bu sorunun önüne geçmek amacıyla önerilen yöntemde sıfır fazlı filtreleme sağlamak için filtfilt (Forward-Backward Filtering) tekniği uygulanmıştır.

Bu teknikte sinyal öncelikle ileri yönde filtrelenmektedir. Bu işlem belirli bir faz gecikmesi yaratmaktadır. Ardından elde edilen filtrelenmiş sinyal, zaman ekseninde ters çevrilip tekrar aynı filtreden geçirilmektedir. İkinci filtreleme adımı, ilk aşamada oluşan faz gecikmesini tam tersi yönde telafi etmektedir. Sonuç olarak, genlik spektrumu filtrenin karesi kadar değiştirilirken, toplam faz kayması sıfırlanmaktadır. Bu işlem sayesinde rPPG sinyalindeki tepe noktaları orijinal zaman konumlarıyla tam olarak örtüşmekte ve kalp atımına ilişkin zamansal analizlerin doğruluğu artırılmaktadır.

Z-Skor Normalizasyonu

Filtreleme işlemi sonrasında elde edilen rPPG sinyalleri, farklı deneklerin cilt tonu, yüzey yansıtıcılığı ve ortam aydınlatması gibi değişkenlere bağlı olarak değişen farklı genlik aralıklarına sahip olabilmektedir. Bu genlik farklılıkları, derin öğrenme modelinin eğitim sürecinde gradyan kararlılığını bozarak yakınsama problemlerine yol açmaktadır. Bu nedenle filtrelenmiş rPPG sinyallerinin istatistiksel özelliklerini standartlaştırmak amacıyla Z-skor normalizasyonu uygulanmıştır.
Z-skor normalizasyonu, sinyalin ortalamasını sıfıra ve standart sapmasını bire eşitleyerek veriyi standart normal dağılıma dönüştürmektedir. Her bir zaman adımındaki sinyal değeri için normalizasyon işlemi Denklem 6.5 kullanılarak gerçekleştirilmektedir.

$$S_{norm}(t) = \frac{S(t) - \mu}{\sigma} \tag{6.5}$$

Denklem 6.5'te $S(t)$ t anındaki ham sinyal değerini, μ sinyalin tüm zaman serisi boyunca hesaplanan aritmetik ortalamasını, σ ise sinyalin standart sapmasını temsil etmektedir. Elde edilen $S_{norm}(t)$ değeri ise normalize edilmiş rPPG sinyalini ifade etmektedir.

Sinyalin ortalama değeri olan μ, tüm sinyal örneklerinin toplamının toplam örnek sayısına bölünmesiyle hesaplanmaktadır. Bu işlemin matematiksel ifadesi Denklem 6.6'da verilmiştir.

$$\mu = \frac{1}{N}\sum_{i=0}^{N} S(i) \tag{6.6}$$

Standart sapma (σ) değeri ise her bir örneğin ortalamadan farkının karelerinin ortalamasının karekökü alınarak Denklem 6.7 ile hesaplanmaktadır.

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(S(i) - \mu)^2} \tag{6.7}$$

Bu denklemde N sinyaldeki toplam örnek sayısını temsil etmektedir. $S(i)$ ise sinyalin i. örnek değerini temsil etmektedir. Z-skor normalizasyonu sayesinde sinyallerin genlik varyasyonları elimine edilerek modelin sadece nabız ritmindeki morfolojik ve frekans tabanlı değişimlere odaklanması sağlanmıştır.

Normalize rPPG Sinyalinin LSTM ile Modellenmesi

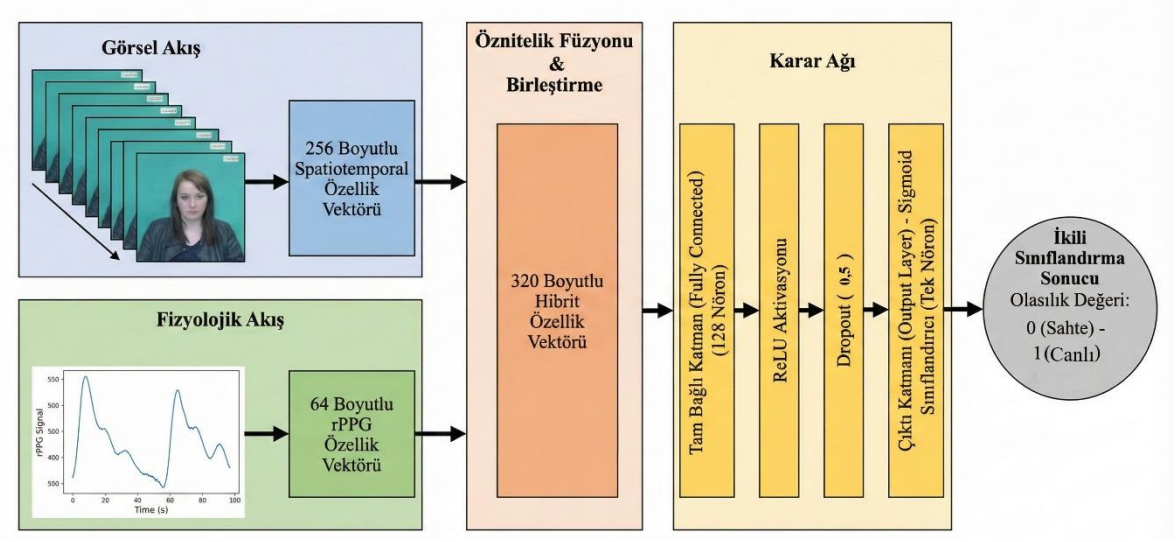
Ön işleme hattından geçirilerek temizlenen ve normalize edilen 1 boyutlu rPPG sinyali nabız ritminin zamansal karakteristiğinin modellenmesi amacıyla tek katmanlı ve 64 gizli birime sahip bir LSTM ağına girdi olarak verilmiştir. LSTM katmanı, canlı yüzlerde gözlenen düzenli, periyodik ve biyofiziksel olarak tutarlı rPPG dalgalanmaları ile sahte yüzlerde karşılaşılan düzensiz, gürültülü veya ritimsiz sinyal davranışlarını ayırt etmeyi öğrenmektedir. Bu ayrımın temelinde, LSTM mimarisinin zamansal bağımlılıkları hücre durumu üzerinden modelleyebilme yeteneği bulunmaktadır. LSTM hücreleri, rPPG sinyali içindeki periyodiklik, faz sürekliliği ve dar bantlı spektral enerji yoğunlaşması gibi kardiyovasküler ritme özgü yapıları koruyarak işlemektedir. Canlı yüzlerde görülen

sistol–diyastol döngüsüne bağlı ritmik değişimler, zaman ekseni boyunca tutarlı bir faz ilişkisi göstermektedir. Buna karşılık sahte yüzlerden elde edilen sinyaller genellikle periyodik değildir, faz sürekliliği bozulmuştur ve spektral enerji belirli bir frekans bandında yoğunlaşmamaktadır. LSTM ağı, bu yapısal farklılıkları hücre durumu içerisinde zamanla biriktirerek öğrenmekte ve fizyolojik olarak tutarlı nabız sinyalinin varlığını güçlü bir canlılık göstergesi olarak kullanarak sahte-canlı yüz ayrımını gerçekleştirebilmektedir.

6.3.3 Özellik Füzyonu ve Karar Mekanizması

Önerilen çift akışlı mimarinin son aşamasında, görsel ve fizyolojik analiz modüllerinden elde edilen ayrık öznitelik vektörleri ortak bir karar düzleminde birleştirilmiştir. Bu aşamanın temel amacı, yüz sahteciliğinin hem uzamsal doku bozulmalarını hem de zamansal fizyolojik tutarsızlıklarını tek bir temsil uzayı içerisinde eş zamanlı olarak değerlendirmektir. Öznitelik füzyonu sürecinde, uzamsal-zamansal görsel akıştan elde edilen 256 boyutlu derin öznitelik vektörü ile fizyolojik akıştan elde edilen 64 boyutlu rPPG tabanlı temsil vektörü, öznitelik -seviyesinde uç uca eklenerek birleştirilmiştir. Bu işlem sonucunda, yüzün hem görsel hem de biyolojik karakteristiklerini temsil eden 320 boyutlu hibrit bir öznitelik uzayı oluşturulmuştur.

Elde edilen bu hibrit temsil vektörü, modelin nihai kararını üretmesi amacıyla 128 nörondan oluşan tam bağlantılı bir katmana aktarılmıştır. Bu katmanda, doğrusal olmayan ilişkilerin etkin biçimde modellenebilmesi için ReLU aktivasyon fonksiyonu kullanılmıştır. Ayrıca modelin aşırı öğrenmesini önlemek ve genelleme yeteneğini artırmak amacıyla bu katmana %50 oranında unutma işlemi uygulanmıştır. Mimarinin çıkış katmanında ise tek nöronlu bir yapı tercih edilmiş ve aktivasyon fonksiyonu olarak sigmoid fonksiyonu kullanılmıştır. Bu son katman, hibrit öznitelik vektörünü işleyerek her bir video örneği için 0 ile 1 aralığında nihai bir olasılık değeri üretmektedir. Üretilen değerin 0'a yakın olması örneğin sahte, 1'e yakın olması ise örneğin canlı sınıfına ait olduğunu göstermektedir. Önerilen öznitelik füzyonu ve karar mekanizmasının genel mimarisi Şekil 6.8’de detaylı olarak sunulmuştur.



Şekil 6.8 Önerilen hibrit mimaride öznitelik füzyonu ve karar mekanizması

7. BULGULAR

Bu bölümde geliştirilen hibrit yüz canlılık tespit modelinin performansı farklı rPPG sinyal çıkarım yöntemleri, uygulanan ROI stratejileri ve zorlu test senaryoları altında detaylı olarak analiz edilmiştir. Modelin başarımı, kişiye özgü biyometrik özelliklerin ezberlenmesini engellemek ve gerçek dünya senaryolarındaki genellenebilirliğini ölçmek amacıyla kişiden bağımsız 5-katlı grup çapraz doğrulama protokolü ile değerlendirilmiştir. Bu doğrulama yöntemi, eğitim ve test kümelerinde aynı kişilere ait görüntülerin bulunmamasını garanti altına alarak veri sızıntısını önlemekte ve modelin genellenebilirliğini en gerçekçi şekilde ölçmektedir. Elde edilen sonuçlar, yüz biyometrisi sunum saldırı tespiti alanında uluslararası referans standart olarak kabul edilen ISO/IEC 30107-3 çerçevesinde tanımlanan APCER, BPCER ve ACER metrikleri ile birlikte doğruluk değerleri üzerinden değerlendirilmiştir. Ayrıca modelin ayırt edicilik performansını görselleştirmek amacıyla ROC eğrileri incelenmiş ve yanlış sınıflandırılan örnekler üzerinden kalitatif bir hata analizi gerçekleştirilmiştir.

Tüm deneysel çalışmalar Python programlama dili ve PyTorch derin öğrenme kütüphanesi kullanılarak NVIDIA GPU (CUDA) altyapısı üzerinde gerçekleştirilmiştir. Eğitim sürecinde modelin ağırlıklarını güncellemek için Adam (Adaptive Moment Estimation) optimizasyon algoritması tercih edilmiştir. Modelin global minimuma kararlı bir şekilde yakınsaması için öğrenme oranı 0.0001 olarak sabitlenmiştir. Donanım bellek kısıtları ve gradyan hesaplamalarının kararlılığı da gözetilerek yığın boyutu 2 olarak belirlenmiştir.

Eğitim süreci çapraz doğrulama protokolündeki her bir fold için toplam 15 epoch boyunca sürdürülmüştür. Her eğitim döngüsünde sabit uzunlukta video klipleri modele sunulmuş ve hata fonksiyonu olarak ikili sınıflandırma problemlerinde standart olan BCEWithLogitsLoss (Binary Cross Entropy with Logits) kullanılmıştır.

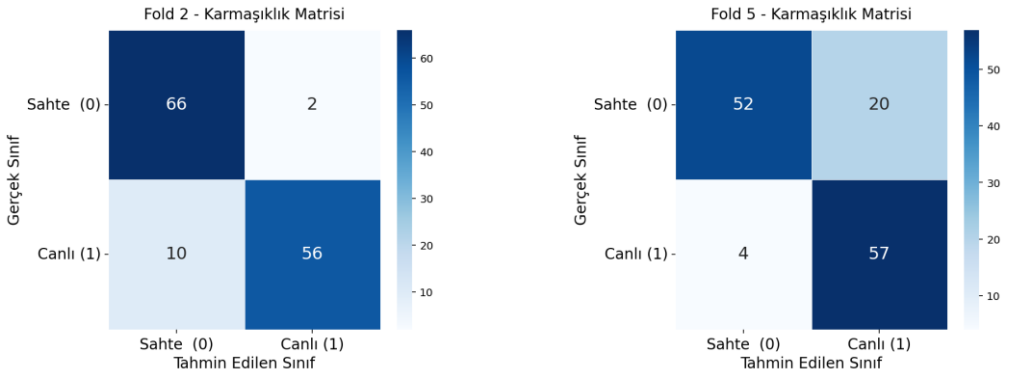
7.1 CHROM Tabanlı rPPG Yöntemi ile Elde Edilen Sonuçlar

Bu bölümde, CHROM tabanlı uzaktan fotopletismografi sinyal çıkarım yönteminin iki farklı ROI stratejisi altında gösterdiği performans ayrıntılı ve karşılaştırmalı olarak değerlendirilmiştir. Analizler, bütünsel yüz ROI ve anatomik segmentasyon tabanlı seçici ROI yaklaşımlarını kapsamaktadır. Her bir ROI stratejisi için ROC eğrileri, karmaşıklık matrisleri ve ISO/IEC 30107-3 standardı ile uyumlu temel performans ölçütleri birlikte ele alınmıştır. Özellikle her yaklaşım için elde edilen en iyi ve en kötü fold sonuçları karşılaştırmalı biçimde incelenerek yöntemin kararlılığı ve genellenebilirliği ortaya konulmuştur.

7.1.1 CHROM Tabanlı Bütünsel Yüz ROI

Bütünsel yüz ROI stratejisinde CHROM yöntemi yüzün tamamındaki renk değişimlerinden yararlanarak rPPG sinyalini çıkarmaktadır. Bu yaklaşım geniş bir doku ve renk çeşitliliğini kapsaması nedeniyle fizyolojik sinyal bileşenlerini yakalama potansiyeline sahiptir. Bununla birlikte mimik hareketleri, baş pozisyonu değişimleri ve konuşma gibi yüz deformasyonları kaynaklı gürültülere karşı daha hassas bir yapı sergilemektedir.

Yapılan beş katmanlı çapraz doğrulama testleri sonucunda bu strateji ile ortalama %95,37 AUC, %14,67 EER ve %14,39 ACER değerlerine ulaşılmıştır. Bu ortalama değerler yöntemin genel başarısını göstermekle birlikte aşağıda detaylandırılan fold bazlı analizler performansın çevresel koşullara göre değişkenlik gösterdiğini ortaya koymaktadır. Bütünsel Yüz ROI stratejisi kullanıldığında elde edilen en iyi ve en kötü sınıflandırma performansına ait karmaşıklık matrisleri Şekil 7.1'de sunulmuştur.

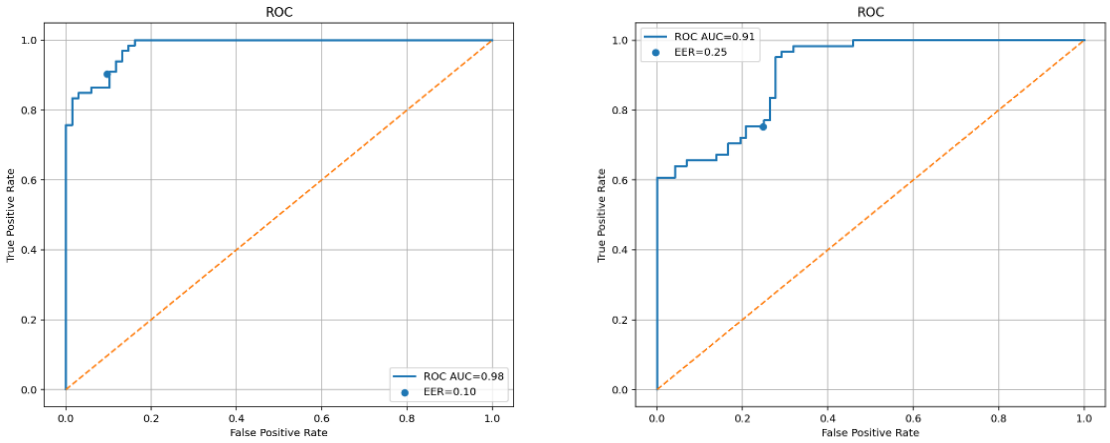


(a) (b)

Şekil 7.1 CHROM yöntemi ve bütünsel yüz ROI stratejisi ile elde edilen karmaşıklık matrisleri a) En iyi fold performansı, b) En kötü fold performansı

Şekil 7.1 (a), CHROM tabanlı bütünsel yüz ROI sonucunda elde edilen sınıflandırma kararlarının en başarılı olduğu durumu göstermektedir. Modelin sahte örnekleri doğru biçimde ayırt etmede yüksek bir performans sergilediği görülmüştür. Ayrıca yanlış pozitif hata oranlarının düşük kaldığı, buna karşın bazı canlı örneklerin sahte olarak sınıflandırıldığı görülmüştür. Şekil 7.1 (b) ise CHROM yönteminin bütünsel yüz ROI altında en zayıf performans sergilediği durumu göstermektedir. Bu fold içerisinde sahte örneklerin bir kısmı yanlışlıkla gerçek sınıfına atandığı görülmüştür. Özellikle 3D maske saldırılarında yanlış pozitif oranı belirgin biçimde arttığı görülmektedir. Bu durum bütünsel yüz yaklaşımının mimik ve yüz hareketlerinden kaynaklı gürültülere duyarlılığının doğrudan bir sonucu olarak değerlendirilmektedir.

Bütünsel yüz ROI stratejisi için elde edilen ROC eğrileri, en iyi ve en kötü fold performansları Şekil 7.2'de gösterilmektedir. Şekil 7.2 a'da sunulan yüksek AUC değeri, yöntemin canlı ve sahte örnekler arasında güçlü bir ayırt edicilik sağlayabildiğini doğrulamaktadır. Buna karşın Şekil 7.2 b'de sunulan en düşük performanslı fold incelendiğinde AUC değerinde anlamlı bir düşüş gözlemlenmektedir.



(a) (b)

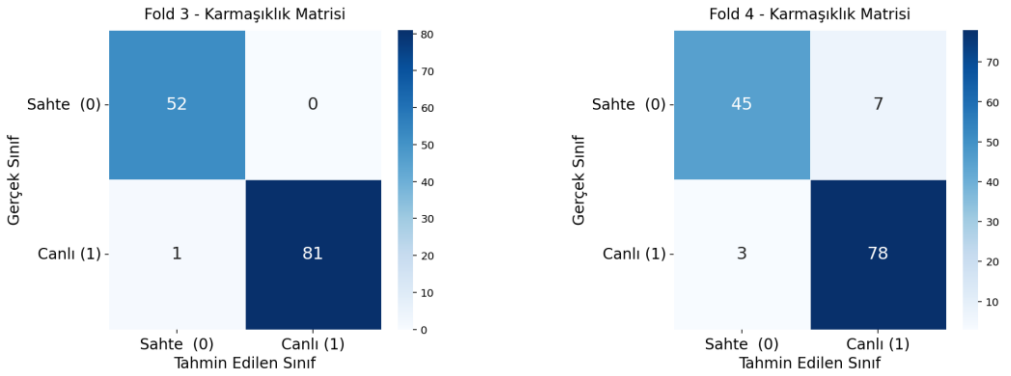
Şekil 7.2 CHROM yöntemi ve bütünsel yüz ROI stratejisi ile elde edilen ROC eğrileri a) En iyi fold performansı b) En kötü fold performansı

Bu performans kaybı, bütünsel yüz ROI yaklaşımının yüz mimiklerine ve istemsiz hareketlere karşı yüksek hassasiyete sahip olduğunu göstermektedir. Bu hareketlerin sinyal periyodikliği üzerinde ani genlik dalgalanmalarına neden olduğu ve sinyalin periyodik yapısını bozduğu değerlendirilmektedir. CHROM yöntemi bütünsel yüz ROI üzerinde uygulandığında, mimik ve yüz ifadelerinden kaynaklanan fizyolojik olmayan renk dalgalanmaları nedeniyle sinyal-gürültü oranının düşmesine yol açmaktadır. Özellikle baş hareketleri ve konuşma esnasında meydana gelen yüz deformasyonları, sinyalin genlik kararlılığını olumsuz etkilemektedir. CHROM yöntemi, RGB kanalları arasındaki korelasyonlara dayalı bir renk projeksiyon tekniği olduğundan, baş hareketleri ve konuşma esnasında oluşan dinamik deformasyonlar sırasında ortaya çıkan kanal uyumsuzluklarını hatalı biçimde nabız sinyali olarak yorumlayabilmektedir.

Tüm bu kısıtlara rağmen, CHROM yönteminin projeksiyon tabanlı matematiksel yapısı ritmik nabız bileşenini belirli bir düzeyde izole edebilme yeteneğine sahiptir. Bu nedenle bütünsel yüz ROI yaklaşımı, özellikle kontrollü ve düşük hareket içeren senaryolarda kabul edilebilir bir performans sergilemektedir.

7.1.2 CHROM Tabanlı Anatomik Segmentasyon Tabanlı Seçici ROI

Anatomik segmentasyon tabanlı seçici ROI stratejisinde, CHROM yöntemi yalnızca alın ve yanak bölgelerinden çıkarılan sinyallerle çalışmaktadır. Bu bölgeler yüksek kılcal damar yoğunluğu ve düşük mimik etkisi nedeniyle rPPG analizi için en uygun fizyolojik kaynakları sunmaktadır. Bu strateji ile gerçekleştirilen testlerde, bütünsel yüz ROI yaklaşımına kıyasla daha yüksek bir başarım elde edilerek ortalama %99,36 AUC, %2,86 EER ve %4,16 ACER değerlerine ulaşılmıştır. Elde edilen bu sonuçlar, anatomik bölge seçiminin CHROM yönteminin gürültü toleransını maksimize ettiğini ve fizyolojik olmayan renk dalgalanmalarını etkin biçimde bastırdığını ve hata oranlarını minimum düzeye indirdiğini göstermektedir. Anatomik segmentasyon tabanlı seçici ROI stratejisi ile elde edilen en iyi ve en kötü fold sonuçlarına ait karmaşıklık matrisleri Şekil 7.3'te karşılaştırmalı olarak verilmiştir.



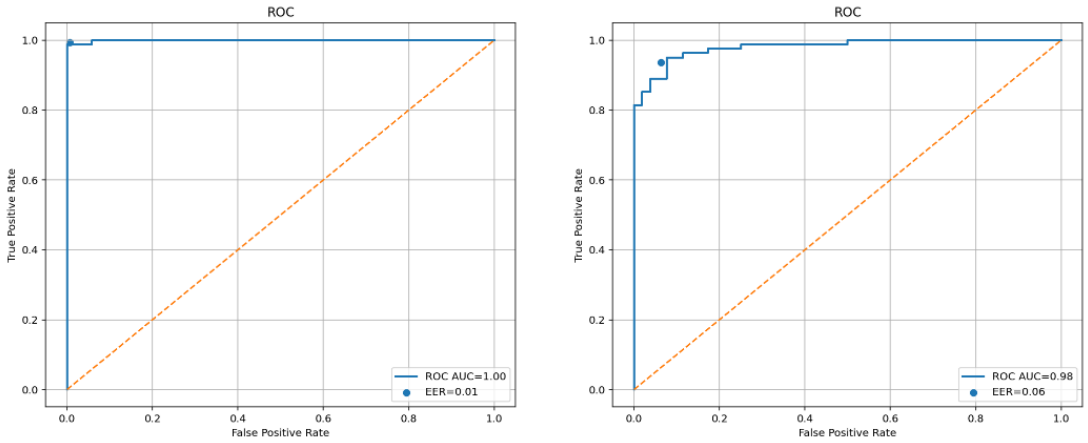
(a) (b)

Şekil 7.3 CHROM yöntemi ve anatomik segmentasyon tabanlı seçici ROI stratejisi ile elde edilen karmaşıklık matrisleri a) En iyi fold performansı b) En kötü fold performansı

Şekil 7.3 (a) iki ROI stratejisi arasındaki en yüksek başarıyı temsil eden en iyi fold sonucunu göstermektedir. Bu foldda modelin neredeyse kusursuz bir ayrım başarısına ulaştığı görülmektedir. Bu fold içerisinde sahte örneklerin tamamı doğru tespit edilerek %0,00'lık bir APCER değerine ulaşılmıştır. Canlı örneklerde ise yaklaşık %1,22'lik bir BPCER oranı ile sadece ihmal edilebilir düzeyde bir hata gözlemlenmiştir.

Şekil 7.3 (b)'de sunulan en düşük performans incelendiğinde ise hata oranının bir miktar arttığı görülmektedir. Bu katmandaki performans düşüşünün temel nedeni, sahte

örneklerin canlı olarak sınıflandırılmasından kaynaklanan ve %13,46 seviyesine yükselen APCER değeridir. Bu sapmaya rağmen en kötü senaryoda dahi genel doğruluk oranının %98'nin üzerinde kalması yöntemin kararlılığını koruduğunu göstermektedir. Modelin özellikle yüksek kaliteli fotoğraf ve video tekrar saldırılarının büyük bir kısmının doğru sınıflandırdığı gözlemlenmiştir. Ayrıca mimik kaynaklı gürültünün maskeleme yoluyla elimine edilmesi, pozitif sınıf doğruluğuna doğrudan katkı sağlamıştır. Yöntemin ROC performansları ise Şekil 7.4'te sunulmuştur.



(a) (b)

Şekil 7.4 CHROM yöntemi ve anatomik segmentasyon tabanlı seçici ROI stratejisi ile elde edilen ROC eğrileri a) En iyi fold performansı, b) En kötü fold performansı

Şekil 7.4 (a), iki farklı ROI stratejisi arasındaki en yüksek başarıyı temsil eden en iyi fold sonucunu göstermektedir. Bu grafikteki ROC eğrisi ve ulaşılan %99,93'lük yüksek AUC değeri seçici anatomik bölgelerin CHROM yönteminin renk-projeksiyon yapısıyla uyumlu çalıştığını ve modelin canlı/sahte sınıflarını birbirinden ayırma gücünü maksimize ettiğini göstermektedir.

Anatomik segmentasyon tabanlı ROI seçimi ile CHROM tabanlı rPPG çıkarım yaklaşımı birlikte kullanıldığında, hata oranlarında bütünsel yüz ROI yaklaşımına kıyasla belirgin bir düşüş sağlanmıştır. Bu strateji sayesinde sinyal kalitesini bozabilecek saç, sakal ve gözlük gibi dış unsurlar analiz dışında bırakılarak, sinyallerin periyodik ve gürültüden arındırılmış bir yapıda elde edilmesi mümkün olmuştur. Elde edilen bulgular, özellikle aydınlatma değişimlerine karşı CHROM algoritmasının yapısal stabilitesini ön plana

çıkarmaktadır. Sonuç olarak anatomik segmentasyon tabanlı ROI seçim yönteminin sistemin genel başarımında belirgin bir performans artışı sağladığı net bir şekilde görülmüştür.

7.1.3 ROI Stratejilerinin Karşılaştırması

CHROM tabanlı rPPG çıkarım yönteminin performansı, 5-katlı kişi tabanlı çapraz doğrulama protokolü esas alınarak iki farklı ROI stratejisi altında karşılaştırmalı olarak analiz edilmiştir. Elde edilen bulgular uluslararası ISO/IEC 30107-3 standardında belirlenen metrikler çerçevesinde değerlendirilmiştir. Her iki stratejiye ait ortalama performans göstergeleri ve standart sapma değerleri Tablo 7.1'de özetlenmiştir.

Tablo 7.1 CHROM Yöntemi kullanılarak farklı ROI stratejilerinin performans karşılaştırması

ROI Stratejisi	ACER (%)	EER (%)	AUC (%)	Doğruluk (%)
Bütünsel Yüz ROI	14,39 ± 3,16	14,67 ± 5,26	95,37 ± 2,40	86,22 ± 3,82
Anatomik Segmentasyon Tabanlı Seçici ROI	4,16 ± 2,84	2,86 ± 1,94	99,36 ± 0,70	96,25 ± 2,66

Analiz sonuçları incelendiğinde her iki ROI yaklaşımının da rPPG tabanlı fizyolojik sinyal çıkarımında etkili olduğu ancak anatomik segmentasyon tabanlı seçici ROI stratejisinin daha kararlı ve genellenebilir bir performans sunduğu görülmektedir. Tablo 7.1'deki verilere göre seçici ROI yaklaşımının bütünsel yüz ROI stratejisine kıyasla ACER değerinde yaklaşık %10,23, EER değerinde ise %11,81 puanlık mutlak bir iyileşme sağladığı görülmektedir. Benzer şekilde AUC değerinde %3,99 puanlık ve doğruluk oranında ise %10,03 puanlık mutlak bir artış elde edilmiştir. Bu sonuçlar modelin benzer dağılımlarda daha yüksek başarımla ayrım yapabildiğini göstermektedir.

Ayrıca performans metriklerindeki standart sapma değerlerindeki düşüş dikkate alındığında anatomik segmentasyon tabanlı seçici ROI yaklaşımının farklı veri setleri ve senaryolar karşısında daha tutarlı sonuçlar ürettiği anlaşılmaktadır. Bu yaklaşımda yalnızca kan akışının en tutarlı şekilde izlendiği yüz segmentlerinin kullanılması, sinyal gürültüsünün azaltılmasına ve nabız kaynaklı mikro varyasyonların daha güçlü biçimde

yakalanmasına imkân sağlamıştır. Bütünsel yüz ROI yaklaşımında gözlenen kararsızlıklar, fizyolojik sinyal açısından zengin olan alın ve yanak bölgelerine odaklanılması ile elimine edilmiştir. Sonuç olarak, anatomik segmentasyon tabanlı seçici ROI stratejisinin CHROM tabanlı rPPG çıkarım yöntemiyle birlikte kullanılması hem sahte hem de canlı örnekler için yüksek güvenilirliğe sahip, kararlı ve etkin bir saldırı tespit sistemi elde edilmesini sağlamıştır.

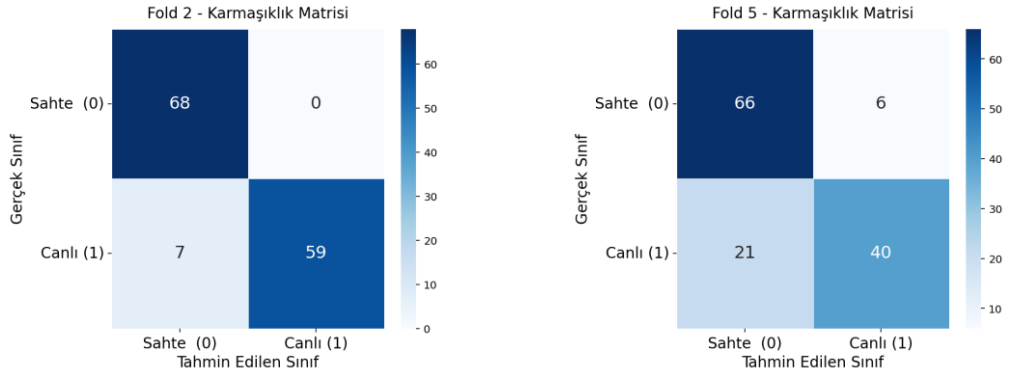
7.2 POS Tabanlı rPPG Yöntemi ile Elde Edilen Sonuçlar

Bu bölümde POS tabanlı rPPG sinyal çıkarım yönteminin, iki farklı ilgi alanı stratejisi olan bütünsel yüz ROI ve anatomik segmentasyon tabanlı seçici ROI altındaki performansı karşılaştırmalı ve ayrıntılı olarak değerlendirilmiştir. Analiz sürecinde her bir ROI yaklaşımı için 5-katlı çapraz doğrulama protokolü içerisindeki en iyi ve en kötü performans gösteren fold durumları ayrı ayrı incelenmiştir. Elde edilen deneysel bulgular ROC eğrileri, karmaşıklık matrisleri ve ISO/IEC 30107-3 standartlarına uygun canlılık metrikleri üzerinden analiz edilmiştir. Özellikle POS yönteminin renk uzayında tanımlı düzlem üzerine projeksiyon yapan matematiksel doğası, yüz bölgesindeki diferansiyel renk değişimlerini ve kan hacmi varyasyonlarını hassas şekilde ayrıştırma yeteneğine sahiptir.

7.2.1 POS – Bütünsel Yüz ROI

Bütünsel yüz ROI stratejisi yaklaşımında POS yöntemi, yüzün tamamındaki piksel yoğunluklarının renk uzayında tanımlanan düzleme izdüşümlerini kullanarak rPPG sinyali üretmektedir. Bu yaklaşım geniş bir doku çeşitliliği ve zengin bir veri havuzu sağlamakla birlikte, özellikle ağız, göz ve burun çevresinde yoğunlaşan mimik hareketlerinden kaynaklanan fizyolojik olmayan gürültüler nedeniyle sinyal kararlılığını olumsuz etkilemektedir. Gerçekleştirilen kişi tabanlı çapraz doğrulama testleri sonucunda bu strateji ile ortalama %95,59 AUC, %12,94 EER ve %15,22 ACER değerlerine ulaşılmıştır. Bu ortalama değerler yöntemin genel başarısını göstermekle birlikte metriklere ait standart sapmanın yüksekliği performansın hareket, mimik ve aydınlatma gibi çevresel faktörlere göre dalgalandığını göstermektedir. Bütünsel yüz ROI stratejisi

kullanıldığında elde edilen en iyi ve en kötü sınıflandırma performansına ait karmaşıklık matrisleri Şekil 7.5'te sunulmuştur.

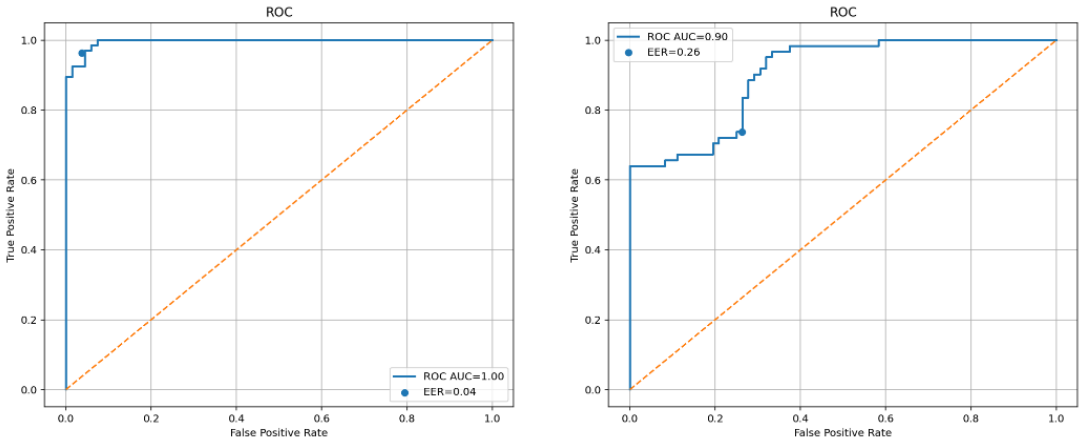


(a) (b)

Şekil 7.5 POS yöntemi ve bütünsel yüz ROI stratejisi ile elde edilen karmaşıklık matrisleri a) En iyi fold performansı, b) En kötü fold performansı

Şekil 7.5 (a), POS yönteminin bütünsel yüz bölgesinden elde edilen zengin renk varyasyonlarını etkin biçimde işleyebildiği durumu temsil etmektedir. Bu fold içerisinde sahte sınıfına ait hata oranları belirgin biçimde azalmıştır. Özellikle tekrar video oynatma saldırılarında yüksek bir ayrım başarısı elde edilerek tüm örnekler doğru şekilde sınıflandırılmıştır.

Buna karşın Şekil 7.5 (b)'de sunulan en kötü performans senaryosu incelendiğinde bütünsel yüz bölgesine uygulanan uzamsal ortalama alma işleminin bazı sahte örneklerde çevresel gürültüleri ve parlama etkilerini de sinyale dâhil ettiği değerlendirilmektedir. Bu durum modelin söz konusu gürültüleri canlılık belirtisi olarak yorumlamasına ve yanlış pozitif oranının artmasına neden olmuştur. Bu hatalar özellikle yüksek yüzey parlaklığı içeren baskı saldırılarında yoğunlaşmaktadır. Ayrıca bu fold özelinde dikkat çeken bir diğer önemli hata türü de gerçek yüz görüntülerinin yanlış sınıflandırılarak sahte olarak etiketlenmesidir. Model bazı gerçek yüz örneklerini yoğun mimik veya ışık kırılmaları nedeniyle 3D maske saldırılarında karşılaşılan sabit ve donuk doku yapısına benzetmiş ve canlı doku olarak doğrulayamamıştır. Elde edilen bu bulgular bütünsel ROI yaklaşımının sahte örneklerin gözden kaçırılmasının yanı sıra gerçek örneklerin hatalı biçimde reddedilmesi riskini de içerdiğini ortaya koymaktadır. Yöntemin sahte ve canlı sınıfları ayırt etme gücünü gösteren ROC eğrileri Şekil 7.6'da sunulmuştur.



(a) (b)

Şekil 7.6 POS yöntemi ve bütünsel yüz ROI stratejisi ile elde edilen ROC eğrileri a) En iyi fold performansı b) En kötü fold performansı

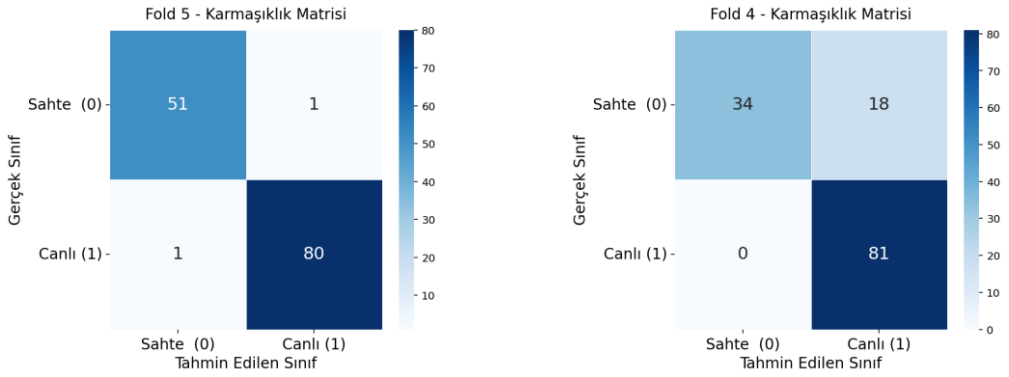
Şekil 7.6 (a)'da sunulan en iyi fold performansına ait ROC eğrisi incelendiğinde AUC değerinin %99 seviyelerine yaklaşması, POS yönteminin ideal koşullarda sahte ve canlı sınıfları ayırma gücünün oldukça yüksek olduğunu göstermektedir.

Buna karşın Şekil 7.6 (b)'de sunulan en kötü performans senaryosunda AUC değerinin yaklaşık %90 seviyesine gerilediği gözlemlenmiştir. Bu düşüşün temel nedeni konuşma, göz kırpma ve baş hareketi içeren dinamik videolarda rPPG sinyalinin faz sürekliliğinin bozulmasıdır. Mimik kaynaklı gürültü, POS algoritmasının cilt rengine dik düzlem üzerindeki projeksiyonunda ayrım kapasitesini zayıflatmış ve bunun sonucunda EER değerinde belirgin bir artış meydana gelmiştir. Sonuç olarak, bütünsel yüz ROI yaklaşımının POS algoritması altında hareketli ve kontrolsüz senaryolara karşı kırılgan bir yapı sergilediği deneysel olarak doğrulanmıştır.

7.2.2 POS – Anatomik Segmentasyon Tabanlı Seçici ROI

Anatomik segmentasyon tabanlı seçici ROI stratejisinde, mimik yoğunluğunun yüksek olduğu ağız ve göz çevreleri sinyal kaynağından çıkarılarak yalnızca alın ve yanak bölgeleri analize dahil edilmiştir. Seçilen bu anatomik segmentlerin hem yüksek damar yoğunluğuna sahip olması hem de düşük kas hareketi içermesi POS yönteminin kararlılığını belirgin şekilde artırmıştır. Yapılan testlerde bu yaklaşım, bütünsel yüz ROI

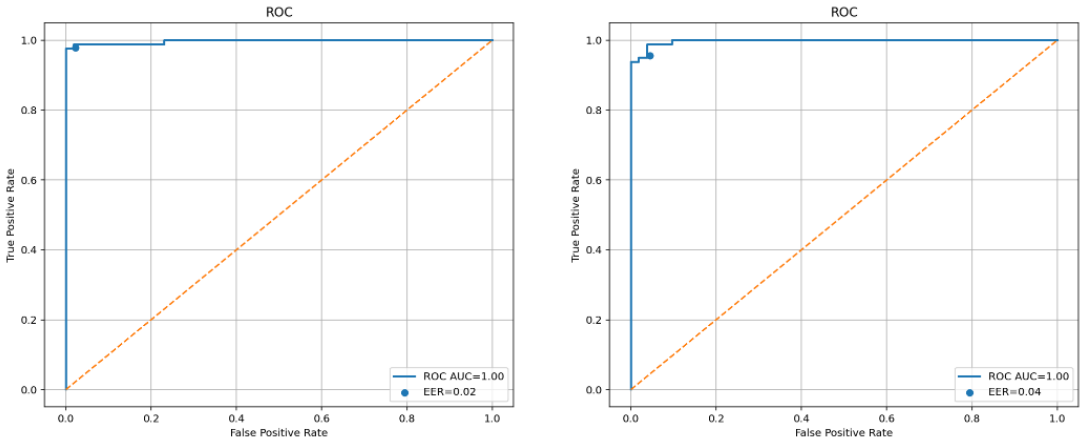
stratejisini geride bırakarak ortalama %99,71 AUC, %2,52 EER ve %7,61 ACER değerlerine ulaşmıştır. Özellikle EER değerinin %2,5 seviyelerine inmesi yöntemin ideal karar eşiğinde sahte ve canlı ayrımını son derece keskin gerçekleştirebildiğini göstermektedir. Anatomik segmentasyon tabanlı seçici ROI stratejisi ile elde edilen en iyi ve en kötü sınıflandırma performanslarına ait karmaşıklık matrisleri Şekil 7.7'de karşılaştırmalı olarak verilmiştir.



(a) (b)

Şekil 7.7 POS yöntemi ve anatomik segmentasyon tabanlı seçici ROI stratejisi ile elde edilen karmaşıklık matrisleri a) En iyi fold performansı b) En kötü fold performansı

Şekil 7.7 (a), gerçekleştirilen POS deneyleri arasında en yüksek performansın elde edildiği durumu göstermektedir. Seçici anatomik bölgeler sayesinde EER oranı minimum düzeyde kalmıştır. Özellikle 3D maske saldırılarında yüksek doğruluk sağlanmıştır. Şekil 7.7 (b)'de sunulan en kötü fold performansı incelendiğinde modelin bazı sahte örnekleri canlı olarak yanlış etiketlediği görülmektedir. Özellikle belirli 3D maske ve video tekrar saldırılarının, sistemin ayırt edici özellik çıkarma kapasitesini zorladığı ve bu saldırı türlerinde sınıflandırma hatalarının belirgin biçimde arttığı değerlendirilmektedir. Yöntemin sahte ve canlı sınıfları ayırt etme başarısını gösteren ROC eğrileri Şekil 7.8'de sunulmuştur.



(a) (b)

Şekil 7.8 POS yöntemi ve anatomik segmentasyon tabanlı seçici ROI stratejisi ile elde edilen ROC eğrileri a) En iyi fold performansı b) En kötü fold performansı

Şekil 7.8 (a), %99,69 AUC değeri ile POS yönteminin fizyolojik döngülerini yüksek doğrulukla yakalayabildiğini göstermektedir. Ölçüm bölgelerinin sabit ve düşük gürültü düzeyine sahip olması sinyal-gürültü oranının artmasına katkı sağlamıştır. Bu durum POS projeksiyonunun doğruluğunu olumlu yönde etkilemiştir. Bu analizde en dikkat çekici bulgulardan biri, en yüksek sınıflandırma hatasına sahip olan fold 4'ün dahi ROC eğrisinde %99,71 gibi oldukça yüksek bir AUC değerine ulaşmasıdır. Bu durum POS yönteminin sahte ve gerçek sinyalleri ayırt etme yeteneğinin son derece güçlü olduğunu ancak fold 4 özelinde karar eşiğinin optimal noktadan sapması nedeniyle sahte kabul oranının arttığını göstermektedir. Başka bir ifadeyle, yöntem ayrıştırma açısından güçlü olmakla birlikte, katı bir eşik uygulanması durumunda hata oranı artabilmektedir. Buna rağmen en kötü fold performansında dahi AUC değerinin %99 bandında kalması anatomik segmentasyon tabanlı seçici ROI yaklaşımının, bütünsel yüz ROI yaklaşımının en olumsuz senaryosuna kıyasla belirgin bir üstünlük sunduğunu ortaya koymaktadır.

7.2.3 ROI Stratejilerinin Karşılaştırması

POS tabanlı rPPG sinyal çıkarım yöntemi için uygulanan bütünsel yüz ROI ve anatomik segmentasyon tabanlı seçici ROI stratejileri, kişi tabanlı 5-katlı çapraz doğrulama protokolü kapsamında elde edilen ortalama performans değerleri üzerinden karşılaştırmalı olarak analiz edilmiştir. Bütünsel yüz ROI ve anatomik segmentasyon

tabanlı seçici ROI yaklaşımlarının ortalama performans sonuçları Tablo 7.2 sunulmaktadır.

Tablo 7.2 POS Yöntemi kullanılarak farklı ROI stratejilerinin performans karşılaştırması

ROI Stratejisi	ACER (%)	EER (%)	AUC (%)	Doğruluk (%)
Bütünsel Yüz ROI	15,22 ± 5,38	12,94 ± 7,85	95,59 ± 3,08	86,06 ± 5,56
Anatomik Segmentasyon Tabanlı Seçici ROI	7,61 ± 6,05	2,52 ± 1,23	99,71 ± 0,13	93,56 ± 5,24

Tablo 7.2'deki veriler incelendiğinde, anatomik segmentasyon tabanlı seçici ROI stratejisinin bütünsel yüz ROI yaklaşımına kıyasla hata oranlarında belirgin bir düşüş sağladığı görülmektedir. Anatomik segmentasyon tabanlı seçici ROI stratejisinin bütünsel ROI yaklaşımına göre ACER değerinde %7,61, EER değerinde ise %10,42 puanlık mutlak bir iyileşme kaydedilmiştir. Benzer şekilde AUC değerindeki yaklaşık %4,12 ve doğruluk değerindeki %7,50 puanlık artış, POS algoritmasının gürültüsüz anatomik bölgelerde fizyolojik sinyali yakalama kapasitesinin maksimize edildiğini doğrulamaktadır.

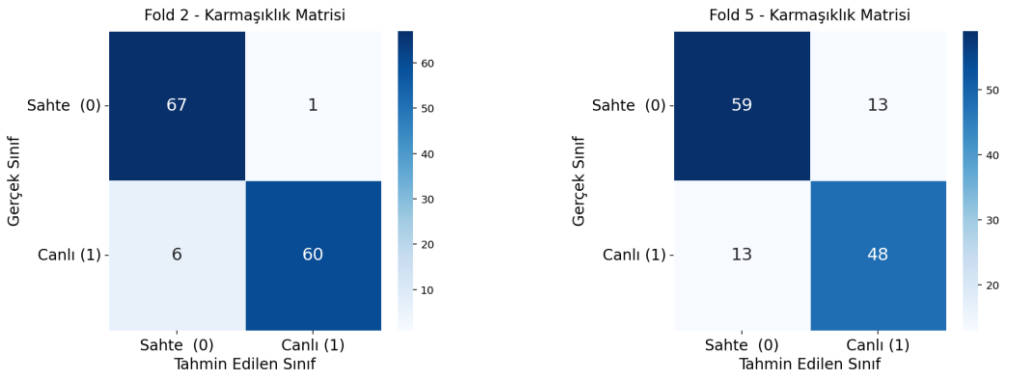
Ayrıca standart sapma değerlerindeki belirgin düşüş, sinyal çıkarım sürecinin fold'lar arasında daha kararlı ve tekrarlanabilir hâle geldiğini göstermektedir. Ancak ACER standart sapmasındaki %6,05'lik artış yöntemin bazı zorlu sahnelerde eşik değeri hassasiyeti yaşadığını, genel ayrım gücü mükemmel olsa bile karar mekanizmasında nadir sapmalar olabildiğini işaret etmektedir. Bu bulgular POS yönteminin doku özelliklerinden ziyade fizyolojik kaynağa dayalı bir yöntem olması nedeniyle ROI seçimine duyarlı olduğunu doğrulamaktadır. Sonuç olarak anatomik olarak optimize edilen segmentler sayesinde mimik ve aydınlatma kaynaklı gürültülerin modele olan olumsuz etkisi en aza indirilmiş ve sahte-canlı sınıflarını birbirinden ayırt etme kapasitesi belirgin şekilde artırılmıştır. Bu bağlamda POS algoritması ve anatomik segmentasyon tabanlı seçici ROI kombinasyonu bu çalışmada incelenen rPPG algoritmaları arasında en stabil sinyal çıkarımını sağlayan yöntemlerden biri olarak öne çıkmaktadır.

7.3 Yeşil Kanal Tabanlı rPPG Yöntemi ile Elde Edilen Sonuçlar

Bu bölümde oksihemoglobinin ışık emilim karakteristiği açısından en duyarlı dalga boyuna karşılık gelen yeşil kanalın kullanıldığı Yeşil Kanal tabanlı rPPG sinyal çıkarım yönteminin, iki farklı ROI stratejisi altında sergilediği performans ayrıntılı ve karşılaştırmalı olarak değerlendirilmiştir. Analiz sürecinde her bir ROI stratejisi için ROC eğrileri, karmaşıklık matrisleri ve hata metrikleri üzerinden kapsamlı analizler yapılmıştır.

7.3.1 Yeşil Kanal – Bütünsel Yüz ROI

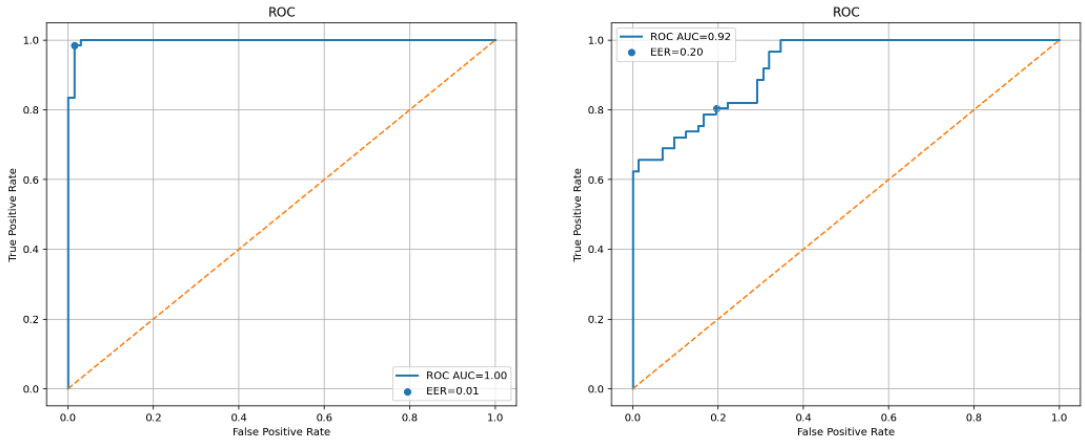
Yeşil Kanal yönteminin yüzün tamamında uygulandığı bütünsel yüz ROI yaklaşımı, geniş bir doku alanından renk bilgisi toplaması sayesinde yüksek genlikli ve sürekli bir sinyal üretme potansiyeline sahiptir. Bununla birlikte yeşil kanal aydınlatma değişimlerine olan doğal hassasiyeti nedeniyle gürültüye en açık yöntemlerden biridir. Özellikle yüz mimikleri, baş hareketleri ve kamera açısındaki değişimlerden kaynaklanan sinyal bozulmaları bu yöntemi doğrudan etkilemektedir. Gerçekleştirilen 5-katlı kişi tabanlı çapraz doğrulama testleri neticesinde bu strateji ile ortalama %95,55 AUC, %12,89 EER ve %15,17 ACER değerlerine ulaşılmıştır. Ortalama ACER değerindeki yüksek standart sapma bütünsel yüz ROI yaklaşımının çevresel koşullara karşı kırılgan olduğunu göstermektedir. Bütünsel yüz ROI stratejisi kullanıldığında elde edilen en iyi ve en kötü sınıflandırma performanslarına ait karmaşıklık matrisleri Şekil 7.9'da sunulmuştur.



(a) (b)

Şekil 7.9 Yeşil Kanal yöntemi ve bütünsel yüz ROI stratejisi ile elde edilen karmaşıklık matrisleri a) En iyi fold performansı b) En kötü fold performansı

Şekil 7.9 (a), Yeşil Kanal yöntemi ve bütünsel yüz ROI stratejisi ile elde edilen en iyi sonuçları göstermektedir. İlgili karmaşıklık matrisinde sahte örneklerin neredeyse tamamının doğru sınıflandırıldığı görülmektedir. Bu durum özellikle düz yüzeyli fotoğraf ve video tekrar saldırılarında Yeşil Kanal'dan elde edilen sinyalin periyodik yapısının sahte yüzey gürültüsünden net biçimde ayrışmasından kaynaklanmaktadır. Buna karşın Şekil 7.9 (b)'de sunulan en düşük performanslı fold sonucu incelendiğinde yüksek varyanslı aydınlatma ve mimik kaynaklı bozulmalar nedeniyle yeşil kanal sinyalinde bütünlük kaybı yaşandığı ve yanlış sınıflandırma oranlarının arttığı gözlemlenmiştir. Literatürde de bilindiği üzere yeşil kanalın ortam aydınlatmasındaki değişimlere karşı yüksek hassasiyet barındırması, mimik hareketleri gibi değişkenlerle birleştiğinde ağın karar mekanizmasını zorlayabilmektedir. Nitekim bu senaryoda, söz konusu çevresel faktörlerin modelin ayırt edici özellik çıkarma kapasitesini kısıtladığı ve sınıflandırma kararlılığını olumsuz etkilediği değerlendirilmektedir. Yeşil Kanal yönteminin sahte ve canlı sınıfları ayırt etme gücünü gösteren ROC eğrileri Şekil 7.10'da sunulmuştur.



(a) (b)

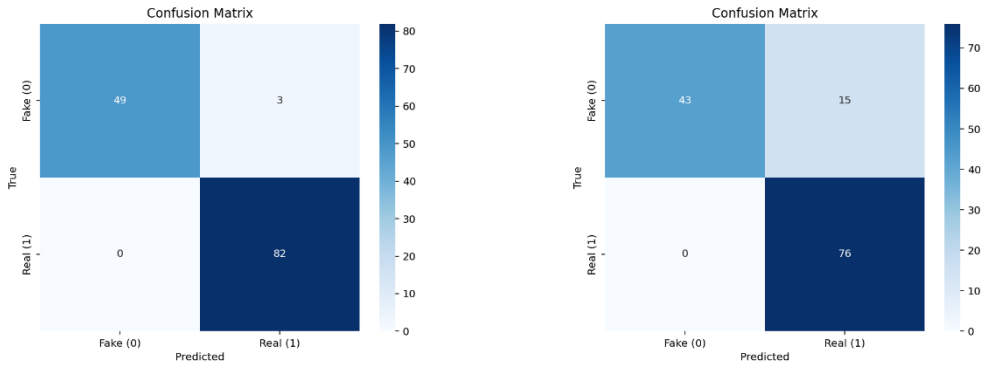
Şekil 7.10 Yeşil Kanal yöntemi ve bütünsel yüz ROI stratejisi ile elde edilen ROC eğrileri a) En iyi fold performansı, b) En kötü fold performansı

Şekil 7.10 (a), en iyi başarımı gösteren foldda Yeşil Kanal yönteminin %99,73'lük AUC değerine ulaştığını, yöntemin sahte ve canlı sınıfları ayırt etme yeteneğinin oldukça yüksek olduğunu ortaya koymaktadır. Buna karşılık Şekil 7.10 (b)'de sunulan en olumsuz senaryoda AUC değeri %92,09 seviyesine gerilerken EER değeri %19,56'ya yükselmiştir. Bu nicel bulgular, bütünsel yüz ROI stratejisinin zorlayıcı gürültü koşulları

altında modelin ayrım gücünde belirgin düşüşlere yol açabildiğini ve sistemin genel sınıflandırma kararlılığını zayıflattığını göstermektedir.

7.3.2 Yeşil Kanal – Anatomik Segmentasyon Tabanlı Seçici ROI

Bu yaklaşımda Yeşil Kanal yöntemi yalnızca rPPG sinyalini en kararlı ve güvenilir biçimde yansıttığı bilinen alın ve yanak bölgelerine uygulanmıştır. Yüksek kılcal damar yoğunluğuna sahip olan bu bölgeler mimiklerden, çene hareketlerinden ve gölge değişimlerinden daha az etkilenerek yeşil kanalın fizyolojik sinyal taşıma kapasitesini maksimize etmektedir. Yapılan testler sonucunda anatomik segmentasyon tabanlı seçici ROI stratejisi ile ortalama %99,33 AUC, %4,23 EER ve %5,90 ACER değerleri elde edilmiştir. Bu sonuçlar anatomik bölge seçiminin yöntemin kararlılığını artırdığını doğrulamaktadır. Anatomik segmentasyon tabanlı seçici ROI stratejisi ile elde edilen en iyi ve en kötü fold performanslarına ait karmaşıklık matrisleri karşılaştırmalı olarak Şekil 7.11'de verilmiştir.



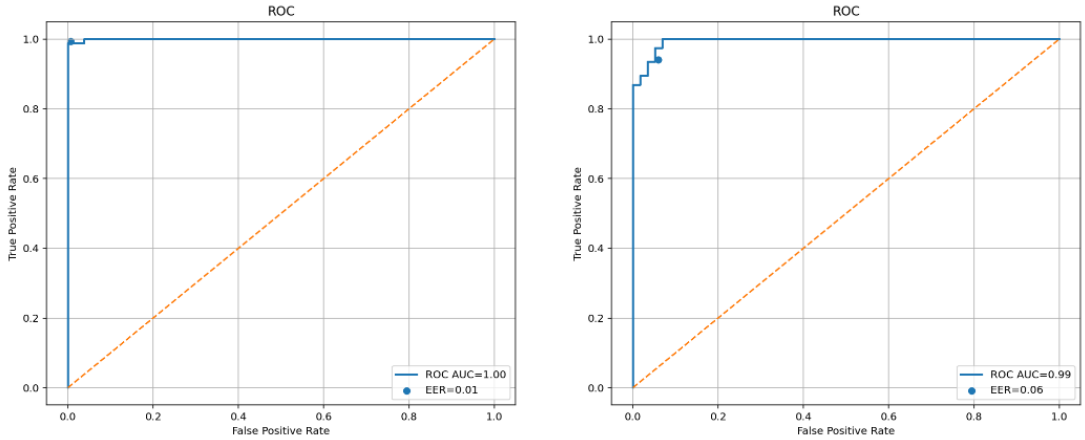
(a) (b)

Şekil 7.11 Yeşil Kanal yöntemi ve anatomik segmentasyon tabanlı seçici ROI stratejisi ile elde edilen karmaşıklık matrisleri a) En iyi fold performansı b) En kötü fold performansı.

Şekil 7.11 (a) en iyi fold için elde edilen sonuçları göstermektedir. Sistemin %99,95 AUC değeri ile neredeyse kusursuz bir performans sergilediği görülmektedir. Bu fold'da sahte örneklerin %94'ü ve canlı örneklerin tamamı doğru şekilde sınıflandırılmıştır. Buna karşılık ACER değerinin %2,88 seviyesine kadar düştüğü gözlemlenmiştir. Bu sonuçlar yüksek kaliteli ekran saldırıları altında dahi sınıflandırma doğruluğunun arttığını göstermektedir. Bunun temel nedeni seçici ROI yaklaşımında yeşil kanalın ışık

yansımalarından ziyade biyolojik ritme duyarlı sinyalleri ön plana çıkarabilmesidir.

Buna karşılık Şekil 7.11 (b)'de sunulan en kötü fold performansı incelendiğinde sahte kabul oranının %25,86 seviyesine yükseldiği görülmektedir. Bu durum söz konusu fold içerisindeki bazı saldırı türlerinin seçici ROI altında dahi yeşil kanalda canlı dokuya benzer renk değişimleri oluşturabildiğini göstermektedir. Bununla birlikte hata oranlarındaki artışa rağmen seçici ROI yaklaşımının bütünsel yüz ROI yönteminin en olumsuz senaryosuna kıyasla performans kaybını belirgin ölçüde sınırladığı görülmektedir. Yöntemin sahte ve canlı sınıfları ayırt etme başarısını gösteren ROC eğrileri Şekil 7.12'de sunulmuştur.



(a) (b)

Şekil 7.12 Yeşil Kanal yöntemi ve anatomik segmentasyon tabanlı seçici ROI stratejisi ile elde edilen ROC eğrisi a) En iyi fold performansı b) En kötü fold performansı

Şekil 7.12 (a) anatomik bölge seçiminin Yeşil Kanal tabanlı rPPG sinyal kalitesini optimize etmedeki başarısını somut bir biçimde ortaya koymaktadır. ROC eğrisinin ideal referans noktasına neredeyse tam yakınsadığı bu senaryoda %99,95 gibi yüksek bir AUC değerine ulaşılmıştır. Bu durum, modelin canlı ve sahte dokuları birbirinden ayırmada yüksek bir kararlılık sergilediğini göstermektedir. Buna karşılık Şekil 7.12 (b)'de sunulan en kötü fold başarımında AUC değeri %99,43 seviyesine gerilemiştir. Hatalı sınıflandırılan video örnekleri manuel olarak incelendiğinde lateral baş hareketleri, dengesiz yan aydınlatmalar ve göz/çene bölgesindeki mikro-gölge değişimleri gibi zorlayıcı çevresel faktörlerin yoğunlaştığı tespit edilmiştir. Söz konusu çevresel ve

fiziksel bozunmaların, modelin ayırt edici öznitelik öğrenme kapasitesini baskılayarak sınıflandırma başarımındaki bu kısmi düşüşe yol açtığı değerlendirilmektedir.

7.3.3 ROI Stratejilerinin Karşılaştırması

Yeşil Kanal tabanlı rPPG yöntemi için uygulanan iki farklı ROI stratejisi, beş fold üzerinden hesaplanan ortalama değerler ve standart sapmalar dikkate alınarak ISO/IEC 30107-3 standardında tanımlanan canlılık metrikleri çerçevesinde karşılaştırılmıştır. Bütünsel yüz ROI ve anatomik segmentasyon tabanlı seçici ROI yaklaşımlarının ortalama performans özeti Tablo 7.3'te sunulmaktadır.

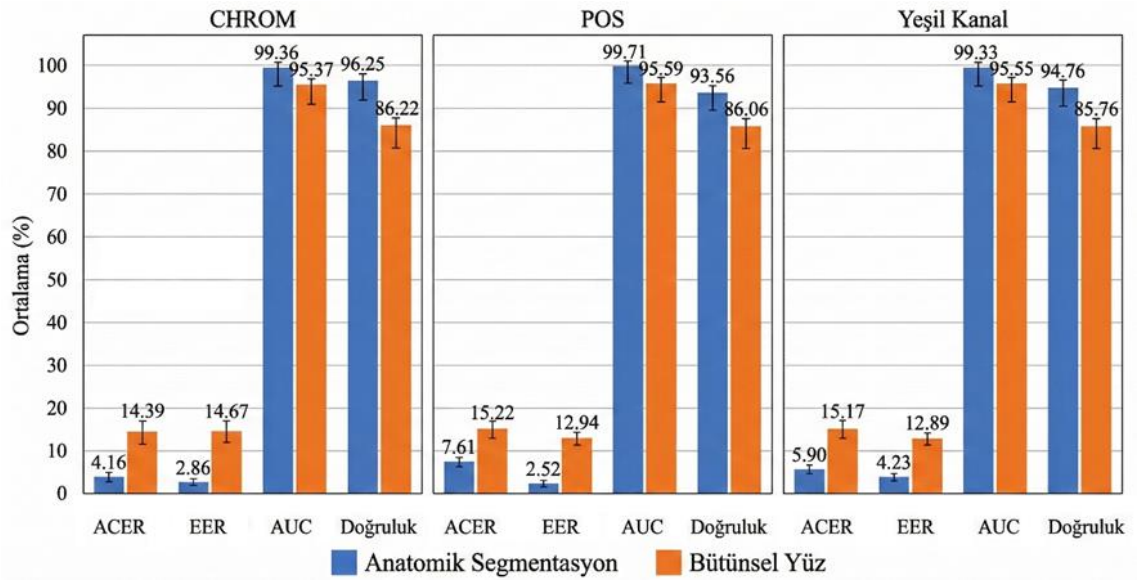
Tablo 7.3 Yeşil Kanal yöntemi kullanılarak farklı ROI stratejilerinin performans karşılaştırması

ROI Stratejisi	ACER (%)	EER (%)	AUC (%)	Doğruluk (%)
Bütünsel Yüz ROI	15,17 ± 5,30	12,89 ± 6,07	95,55 ± 2,44	85,76 ± 5,59
Anatomik Segmentasyon Tabanlı Seçici ROI	5,90 ± 3,78	4,23 ± 2,87	99,33 ± 0,43	94,76 ± 3,54

Tablo 7.3'teki veriler incelendiğinde anatomik segmentasyon tabanlı seçici ROI stratejisinin Yeşil Kanal yönteminin belirgin biçimde iyileştirdiği görülmektedir. Seçici ROI kullanımıyla ACER değerinde %9,27, EER değerinde ise %8,66 puanlık mutlak bir iyileşme sağlanmıştır. Benzer şekilde modelin genel ayrım gücünü ifade eden AUC değerinde %3,78 puanlık ve doğruluk oranında ise %9,00 puanlık mutlak artış elde edilmiştir. AUC değerindeki artış seçici ROI yaklaşımının yöntemin genel ayrım gücünü anlamlı düzeyde artırdığını göstermektedir. Bütünsel yüz ROI yaklaşımında yüzün tamamından kaynaklanan gürültü Yeşil Kanal yönteminin performansını sınırlarken anatomik segmentasyon sayesinde bu gürültüler büyük ölçüde elimine edilmiş ve yöntem daha kararlı bir yapıya kavuşmuştur. Sonuç olarak Yeşil Kanal yöntemi için anatomik bölgelerin seçilerek işlenmesi, yöntemin güvenilirliğini ve canlılık tespit başarısını önemli ölçüde artırmıştır.

7.4 ROI Stratejilerinin Performansının Genel Değerlendirilmesi

Bu bölümde rPPG tabanlı fizyolojik sinyal çıkarımında kullanılan iki farklı ROI stratejisinin model performansı üzerindeki etkisi sistematik biçimde değerlendirilmiştir. Analiz süreci kapsamında CHROM, POS ve Yeşil Kanal olmak üzere üç farklı rPPG yöntemi için ACER, EER, AUC ve doğruluk performans metrikleri hesaplanmıştır. Doğrulama protokolü olarak, sonuçların güvenilirliğini garanti altına alan kişi tabanlı 5-katlı çapraz doğrulama yöntemi benimsenmiştir. Gerçekleştirilen tüm deneylerden elde edilen ortalama performans değerleri ve standart sapma değerleri Şekil 7.13'te özetlenmiştir.



Şekil 7.13 Farklı ROI stratejileri ve rPPG algoritmalarının performans karşılaştırması

Bütünsel yüz ROI stratejisi altında elde edilen sonuçlar incelendiğinde, kullanılan yöntemden bağımsız olarak ortalama sınıflandırma hatası değerlerinin %14,39 ile %15,22 aralığında seyrettiği ve standart sapmaların yüksek olduğu görülmektedir. Yüzün tamamının analize dahil edildiği bu yaklaşımda göz kırpma, dudak hareketleri ve mimiklerden kaynaklanan asimetrik kas deformasyonları algoritmaların ayırt edici fizyolojik öznitelikleri yakalama kapasitesini baskılamaktadır. Bütünsel yüz yaklaşımının neden olduğu bu gürültü artefaktları, modelin öğrenme sürecini sınırlandırmakta ve AUC değerlerinin %95 bandında kalmasına neden olmaktadır. Buna karşılık anatomik segmentasyon tabanlı seçici ROI stratejisine geçildiğinde performans metriklerinde

istatistiksel olarak anlamlı ve belirgin bir iyileşme gözlemlenmiştir. Kılcal damar yoğunluğunun yüksek ve kas hareketliliğinin düşük olduğu alın ve yanak bölgelerine odaklanılması sonucunda, ACER değerlerinin %4,16 ile %7,61 aralığına düştüğü ve AUC değerlerinin %99 seviyesinin üzerine çıktığı gözlemlenmiştir. Elde edilen sonuçlar, rPPG tabanlı sistemlerde ROI bölge seçiminin, kullanılan algoritma tercihine kıyasla daha belirleyici bir optimizasyon faktörü olabileceğini göstermektedir.

Anatomik segmentasyon tabanlı seçici ROI stratejisi altında yöntemlerin kendi içindeki performansları değerlendirildiğinde ise farklı algoritmaların farklı performans metriklerinde öne çıktığı gözlemlenmiştir. CHROM yöntemi özelinde incelendiğinde, bütünsel yüz ROI kullanımı ile %14,39 olarak ölçülen ortalama ACER değerinin anatomik segmentasyon tabanlı seçici ROI yaklaşımıyla %4,16'ya gerilediği tespit edilmiştir. Hata oranında sağlanan 10,23 puanlık bu mutlak iyileşme CHROM yönteminin bütünsel yüz ROI analizde mimik kaynaklı gürültülere karşı hassasiyetini doğrulamakla birlikte doğru bölge seçimiyle en düşük standart sapma değerine ulaşarak en kararlı yöntem haline geldiğini göstermektedir. Benzer şekilde, bütünsel yaklaşımda %95,37 olarak ölçülen AUC değeri gelişmiş ROI stratejisinin kullanılmasıyla %99,36'ya yükselmiştir. Bu 3,99 puanlık artış, göz ve ağız çevresindeki mimik, konuşma veya doğal olmayan hareketlerden kaynaklanan fizyolojik olmayan gürültülerin filtrelenmesi sayesinde modelin canlı ve sahte sınıfları ayırt etme kapasitesinin önemli ölçüde arttığını doğrulamaktadır.

ROI stratejisinin etkisi, POS yönteminin AUC skorunda belirgin biçimde gözlemlenmiştir. Bütünsel yaklaşımda %95,59 olarak ölçülen AUC değerinin, seçici ROI stratejisinin uygulanmasıyla %99,71 seviyesine yükseldiği görülmüştür. ACER metriğinde sağlanan 7,61 puanlık iyileşme ise POS algoritmasının matematiksel projeksiyon yapısının, homojen cilt bölgelerinden elde edilen daha temiz fizyolojik sinyallerle beslendiğinde potansiyel performansını önemli ölçüde artırdığını göstermektedir.

Şekil incelendiğinde dikkat çeken bir diğer husus CHROM ve Yeşil Kanal yöntemlerinde standart sapma değerlerinin belirgin düşüş göstermesidir. Bu durum, anatomik

segmentasyon tabanlı ROI seçiminin modelin veri setindeki varyasyonlara karşı direncini artırdığını doğrulamaktadır. Ancak POS yönteminin standart sapma analizi detaylandırıldığında dikkat çekici bir eğilim gözlemlenmektedir. Yöntemin EER standart sapması 7,85'ten 1,23'e ve AUC standart sapması 3,08'den 0,13'e gerileyerek belirgin ölçüde azalırken, ACER standart sapması 5,38'den 6,05'e çıkarak bir miktar artış göstermiştir. Elde edilen bu bulgu, POS algoritmasının sahte ve canlı sınıfları birbirinden ayırt etme yeteneğini temsil eden AUC değerinin oldukça kararlı hâle geldiğini göstermektedir. Bununla birlikte bazı zorlu test senaryolarında karar eşiğine karşı hassasiyetin devam ettiği ve bu nedenle eşik kaynaklı kısmi hata dalgalanmalarının oluşabildiği görülmektedir.

Genel bir değerlendirme yapıldığında, önerilen ROI stratejisinin yalnızca ortalama sınıflandırma başarısını artırmakla kalmadığı görülmektedir. Ayrıca EER değerlerinde yaklaşık 10 ila 12 puanlık mutlak iyileşmeler sağlayarak sistemin kararlılığını literatürdeki güncel çalışmalarla rekabet edebilir üst bir seviyeye taşıdığı anlaşılmaktadır. Elde edilen düşük standart sapma değerleri modelin farklı aydınlatma koşulları ve özneler arası varyasyonlar altında tutarlı performans sergilediğini ortaya koymaktadır. Sonuç olarak, önerilen ROI seçiminin fizyolojik öznitelik çıkarımı açısından kritik bir öneme sahip olduğu ortaya konulmuştur. Ayrıca bu şekilde optimize edilen rPPG akışının, görsel akıştan elde edilen tam yüz bilgisiyle yüksek düzeyde tamamlayıcılık sağladığı belirlenmiştir.

Önerilen ROI stratejilerinin ve rPPG algoritmalarının doğruluk oranlarındaki iyileşmelerin altında yatan temel dinamikleri ve potansiyel zafiyetleri daha şeffaf bir biçimde değerlendirmek amacıyla karmaşıklık matrislerinde yer alan hatalı sınıflandırılan örnekler detaylı olarak incelenmiştir. Üç farklı rPPG yöntemi ve iki farklı ROI stratejisi altındaki başarısız tespitler analiz edildiğinde hataların rastgele dağılmadığı aksine veri seti ve saldırı türü bazında belirgin kümelenmeler oluşturduğu tespit edilmiştir.

Replay-Mobile veri setine ait ortak hatalar incelendiğinde, modelin yanlış ret hatasına düştüğü örneklerin büyük çoğunluğunun yetersiz veya ters ışık koşullarında kaydedildiği belirlenmiştir. Yetersiz ortam ışığının neden olduğu donanımsal gürültü ve düşük

kontrast, modelin canlılık tespiti için ihtiyaç duyduğu ayırt edici öznitelikleri çıkarmasını zorlaştırmaktadır. Bu durum, modelin karar sınırlarını doğrudan etkileyerek hatalı sınıflandırmalara yol açmaktadır. Benzer şekilde, ışık kaynağının yetersiz olduğu bazı video tekrar saldırılarında sistemin yanlış kabul hatası ürettiği görülmüştür. Özellikle tablet ve telefon üzerinden gerçekleştirilen video tekrar oynatma saldırılarında, ekran yüzeyinin mat olduğu senaryoların sistemi önemli oranda zorladığı tespit edilmiştir. Mat ekranlar, dijital sahteciliği ele veren en belirgin bağlamsal ipuçlarından biri olan ekran parlamalarını büyük ölçüde engellemektedir. Bu yansımaların yokluğu, modelin ekran üzerindeki zamansal değişimleri hatalı bir şekilde canlı doku özniteliği olarak sınıflandırmasına neden olabilmektedir. Sınıflandırma sürecinde modelin en çok zorlandığı Replay-Mobile veri setine ait zorlayıcı aydınlatma ve ekran koşullarını yansıtan görsel örnekler Şekil 7.14'te sunulmuştur.

Şekil 7.14 Replay-Mobile veri setine ait ortak hatalar

Bütünsel ROI stratejisi altındaki başarısız dosyalar incelendiğinde, her üç rPPG yönteminde de hataların büyük çoğunluğunun 3DMAD veri setine ait Session 1 ve Session 2 kayıtlarından kaynaklandığı görülmüştür. Veri setinin yapısı gereği Session 1 ve Session 2, canlı yani maskesiz insan görüntülerini içermektedir. Bu durum, bütünsel yaklaşımın 3DMAD veri setindeki gerçek yüzleri yüksek oranda sahte olarak

sınıflandırdığını göstermektedir. Bu başarısızlığın temel nedeninin, saç, arka plan sınırları, mimik ve konuşma kaynaklı kas hareketleri gibi yüzün tamamına ait değişkenlerin analize dâhil edilmesi olduğu değerlendirilmektedir. Ancak anatomik segmentasyon yaklaşımı uygulandığında, bu iki alt kümedeki sınıflandırma hatalarının tamamen ortadan kalktığı ve doğru etiketleme oranının maksimum seviyeye ulaştığı tespit edilmiştir. Buna karşılık insan derisinin optik yansıma özelliklerini yüksek oranda taklit eden tamamen sahte maske saldırılarından oluşan Session 03 örneklerinde modelin bazı maskeleri canlı olarak sınıflandırdığı gözlemlenmiştir. Bu durum, önerilen stratejinin yapısal deformasyonları ve hareket gürültülerini filtrelemede son derece başarılı olduğunu ancak materyal benzerliğinin çok yüksek olduğu senaryolarda sınıflandırma sınırlarının aşılabildiğini göstermektedir. Sınıflandırma sürecinde modelin zorlandığı 3DMAD sahte maske senaryolarına ait görsel örnekler Şekil 7.15'te sunulmuştur.

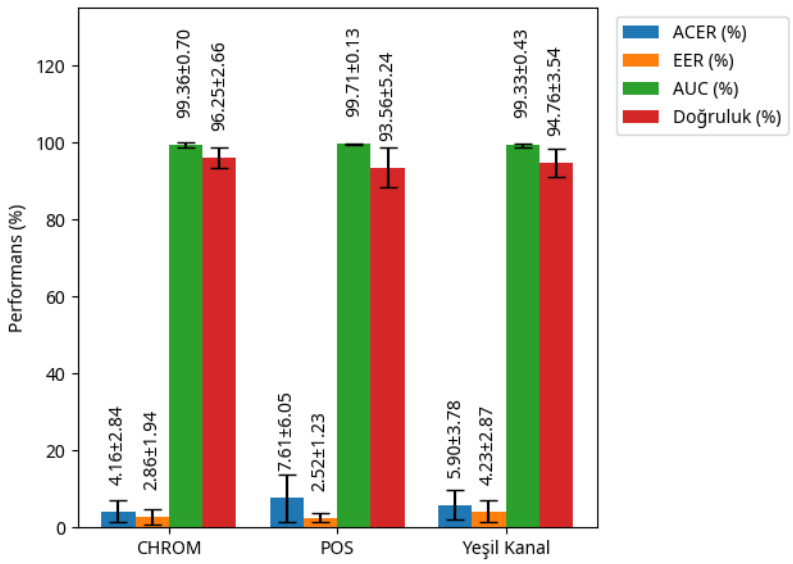
Şekil 7.15 3DMAD veri setine ait ortak hatalar

Hatalı sınıflandırma örnekleri üzerinde gerçekleştirilen bu analiz, rPPG temelli anatomik bölge seçiminin gerçek yüzlerin doğru tespit edilmesinde önemli bir katkı sağladığına işaret etmektedir. Bununla birlikte, yalnızca gürültüden arındırılmış spesifik deri bölgelerine odaklanmanın, önceden kaydedilmiş fizyolojik ipuçları barındıran yüksek kaliteli 2B mat ekran saldırılarına ve çok gerçekçi 3D maskelere karşı modelin karar mekanizmasını zaman zaman yanıltabildiği gözlemlenmiştir. Elde edilen bu bulgular, tez kapsamında önerilen ve fizyolojik rPPG akışını uzamsal görsel doku analiziyle entegre eden çift akışlı hibrit mimarinin teorik altyapısını destekler niteliktedir. Önerilen mimaride rPPG akışının bölgesel sinyal odaklı yapısından kaynaklanabilecek potansiyel zafiyetlerin görsel akıştaki evrişimsel sinir ağlarının moiré desenleri, ekran yansımaları

ve çerçeve bozulmaları ve dokusal anomaliler gibi uzamsal artefaktları tespit etme yeteneği ile dengelenmesi hedeflenmiştir. Bu tamamlayıcı füzyon yaklaşımı sayesinde, heterojen sunum saldırılarına karşı donanımdan bağımsız yüksek ayırt ediciliğe sahip ve daha bütüncül bir savunma mekanizmasının oluşturulması amaçlanmıştır.

7.5 rPPG Yöntemlerinin Karşılaştırılması

Anatomik segmentasyon tabanlı seçici ROI stratejisi sabit bir değişken olarak kabul edilerek, literatürdeki üç temel rPPG sinyal çıkarım yönteminin performansı karşılaştırmalı olarak analiz edilmiştir. Elde edilen ortalama sonuçlar Şekil 7.16'da özetlenmiştir.



Şekil 7.16 Anatomik segmentasyon tabanlı seçici ROI stratejisi altında farklı rPPG yöntemlerinin performans karşılaştırması

Bu çalışmada gerçekleştirilen deneysel analizler anatomik segmentasyon tabanlı seçici ROI stratejisinin rPPG yöntemlerinin sinyal kalitesi ve canlılık tespiti performansı üzerindeki belirleyici rolünü somut verilerle ortaya koymaktadır.

Deney setinde POS yöntemi, %99,71 AUC ve %2,52 EER değerlerine ulaşarak sınıfları birbirinden ayırma kapasitesi bakımından en yüksek performansı sergilemiştir POS yönteminin bu başarısı diğer yöntemlerin aksine sabit katsayılı renk kanalı

kombinasyonlarına dayanmak yerine fizyolojik yüz haritasına uyarlanan düzlem-ortogonal bir projeksiyon tekniği kullanmasından kaynaklanmaktadır. Bu teknik zamanla güncellenen dinamik bir projeksiyon düzlemi tanımlayarak yüzeyden yansıyan ışığı yansıma kaynaklı gürültüyü oluşturan speküler bileşenler ile asıl nabız bilgisini taşıyan difüze bileşenlere etkin bir biçimde ayrıştırmaktadır. Bu sayede veri setindeki heterojen aydınlatma değişimlerine ve farklı deneklerin cilt tonu varyasyonlarına karşı yüksek düzeyde adaptasyon sağlanmıştır. POS yöntemi, sinyal enerjisini maksimize ederek sahte ve canlı doku arasındaki dağılım farkını daha belirgin hâle getirmiş, bu durum da teorik ayrım gücünü ifade eden AUC performansının en yüksek seviyeye ulaşmasına katkı sağlamıştır.

Sınıfları ayırma gücünde POS yöntemi öne çıkarken, nihai karar başarısını temsil eden ACER metriğinde ise CHROM yöntemi %4,16 hata oranı ve ± 2,84 standart sapma ile en başarılı ve en kararlı yöntem olarak belirlenmiştir. CHROM yönteminin renk farkı tabanlı yaklaşımı, özellikle yüzdeki mimik ve bölgesel kas hareketlerinin oluşturduğu hareket artefaktlarını azaltma konusunda yüksek bir dayanıklılık göstermektedir. POS yönteminin ACER metriğinde ±6,05 gibi görece yüksek bir standart sapma göstermesi, bu yöntemin genel ayırt ediciliği yüksek olsa da karar eşiğine karşı daha hassas bir yapı sergilediğini göstermektedir. Ayrıca bu durum, veri setinin bazı zorlu alt kümelerinde yöntemin optimal çalışma noktasından sapma gösterebildiğini ortaya koymaktadır. Buna karşılık CHROM yöntemi, farklı aydınlatma ve senaryo koşulları altında canlı ve sahte dokuları daha yüksek kararlılıkla ayırt edebilmekte ve daha stabil bir kabul/ret performansı sunmaktadır. Bu durum, CHROM yaklaşımının gerçek dünya senaryoları ve pratik uygulamalar açısından güvenilir bir mimari sunduğunu göstermektedir.

Yeşil kanalın tek başına kullanıldığı Yeşil Kanal yöntemi ise %5,90 ACER değeri ile POS ve CHROM arasında dengeli bir performans sergilemiştir. Hemoglobinin absorpsiyon spektrumunun yeşil ışıkta tepe noktası oluşturması nedeniyle güçlü bir fizyolojik bilgi taşısa da tek kanallı yapısı ortam ışığındaki beyaz dengesi dalgalanmaları gibi ani değişimlere karşı çok kanallı yöntemlere kıyasla ağın daha hassas kalmasına neden olmaktadır. Ancak anatomik segmentasyon ile gürültülü bölgelerin dışarıda bırakılması, Yeşil Kanal yönteminin de bütünsel yaklaşıma göre performansını belirgin

ölçüde artırmasını sağlamıştır.

Sonuç olarak anatomik segmentasyon yaklaşımı, yüzün tamamından elde edilen sinyaldeki varyansı minimize ederek rPPG algoritmalarının sadece nabız kaynaklı değişimlere odaklanmasını sağlamıştır. Bu strateji, modelin farklı çapraz doğrulama katmanları boyunca performans dalgalanmasını azaltmış ve sistemin bütünsel kararlılığını artırmıştır. Önerilen hibrit yapıda, yüksek ayırt ediciliğin kritik olduğu durumlarda POS yönteminin, sistem kararlılığı ve düşük hata oranının öncelikli olduğu senaryolarda ise CHROM tabanlı yöntemin tercih edilmesi gerektiği bu çalışmanın en önemli deneysel çıktılarından biridir. Bu durum, fizyolojik akışın uygun ROI stratejisi ile optimize edilmesinin sunum saldırısı tespit performansı üzerinde doğrudan ve kritik bir etkiye sahip olduğunu doğrulamaktadır.

7.6 Çapraz Doğrulama Bazlı Performans Analizi

Bir önceki bölümde gerçekleştirilen karşılaştırmalı analizlerde, sistemin nihai karar başarısını temsil eden ACER metriği baz alındığında en üstün performansı CHROM yönteminin sergilediği tespit edilmiştir. Bu doğrultuda, en kararlı mimari olarak öne çıkan CHROM yönteminin farklı veri alt kümelerindeki davranışını detaylandıran çapraz doğrulama analizleri bu bölümde sunulmaktadır. Önerilen modelin genelleme yeteneğini ve veri dağılımındaki varyasyonlara karşı kararlılığını ölçmek amacıyla uygulanan 5-katlı kişi tabanlı çapraz doğrulama sonuçları Tablo 7.4'te sunulmuştur. Bu sonuçlar CHROM yönteminin anatomik segmentasyon desteği ile öğrenme sürecinde farklı veri alt kümelerine nasıl tepki verdiğini ve güvenlik-konfor dengesini nasıl optimize ettiğini ortaya koymaktadır.

Tablo 7.4 CHROM yöntemi için fold bazlı sonuçlar

Fold	APCER (%)	BPCER (%)	ACER (%)	AUC (%)	Doğruluk (%)
Fold 1	**%0,00**	%4,23	%2,11	%99,93	%97,76
Fold 2	%6,90	**%0,00**	%3,45	%99,61	%97,01
Fold 3	**%0,00**	%1,22	**%0,61**	**%99,93**	**%99,25**
Fold 4	%13,46	%3,70	%8,58	%98,03	%92,48
Fold 5	%9,62	%2,47	%6,04	%99,31	%94,73

Tablodaki veriler incelendiğinde modelin tüm foldlarda %98'in üzerinde AUC değeri sergileyerek veri setinin dağılımından bağımsız yüksek bir ayırt edicilik gücüne sahip olduğu görülmektedir. Bununla birlikte farklı test kümeleri üzerinde gerçekleştirilen analizlerde sistemin güvenlik düzeyini temsil eden APCER ile kullanıcı erişilebilirliğini ifade eden BPCER metrikleri arasında karakteristik değişimler gözlemlenmiştir.

Birinci ve üçüncü foldlar istemin güvenlik mimarisinin en kararlı çalıştığı ve en başarılı olduğu senaryoları temsil etmektedir. Her iki fold'da da APCER değeri %0,00 olarak ölçülmüştür. Bu durum yüksek çözünürlüklü maskeler ve video tekrar oynatma girişimleri de dahil olmak üzere test setindeki hiçbir sahte saldırının sistemi atlatamadığını göstermektedir. Bir başka deyişle sistem sıfır yanlış kabul hatasıyla mutlak bir koruma performansı sergilemiştir. Özellikle üçüncü fold %1,22 gibi çok düşük bir BPCER ve %0,61 seviyesindeki minimal ACER değeri ile güvenliğin maksimize edildiği aynı zamanda gerçek kullanıcı erişim konforundan neredeyse hiç ödün verilmediği en ideal çalışma noktasını oluşturmuştur.

İkinci fold sonuçları %0,00 seviyesindeki BPCER değeri ile sistemin hiçbir canlı kullanıcıyı hatalı olarak reddetmediğini yani sıfır yanlış ret hatası ile çalıştığını göstermektedir. Bu senaryoda model kullanıcı deneyimi ve erişim kolaylığı açısından kusursuz bir performans sergilemiştir. Ancak erişilebilirlikteki bu yüksek başarı güvenlik performansında bir denge değişimi yaratarak APCER değerinin %6,90 seviyesine yükselmesine neden olmuştur. Diğer bir ifadeyle sistemin gerçek kullanıcılara gösterdiği yüksek tolerans canlı doku karakteristiğini taklit eden bazı sofistike sahte örneklerin de sistem tarafından kabul edilmesine yol açmıştır.

Dördüncü fold %13,46 APCER ve %8,58 ACER değeri ile model performansının en çok zorlandığı veri kümesi olarak öne çıkmaktadır. Bu fold içerisindeki sahte örneklerin önemli bir kısmının model tarafından canlı olarak sınıflandırılması ilgili veri setinde yoğunlaşan sofistike saldırı türlerinin CHROM yönteminin ayırt edici öznitelik çıkarma kapasitesini kısıtladığına işaret etmektedir. Buna karşın AUC değerinin %98,03 gibi yüksek bir seviyede korunması önemli bir bulgudur. Bu durum modelin sahte ve canlı örnekleri birbirinden ayırma başarısının halen güçlü olduğunu ancak bu spesifik katman

özelinde karar eşiğinin optimum noktadan saptığını işaret etmektedir. Benzer şekilde beşinci fold sonucunda ölçülen %9,62 APCER değeriyle sistemin güvenlik hassasiyetinin zorlandığı diğer bir senaryo olarak kaydedilmiştir.

Bütünsel yüz ROI yaklaşımlarında %10-15 bandında seyreden hata oranlarının aksine anatomik segmentasyon tabanlı bu yaklaşımda en kötü senaryoda dahi hata oranı tek haneli seviyelerde olduğu görülmektedir. Ortalama ACER'in %4,16 seviyesinde gerçekleşmesi ve AUC değerlerindeki kararlılık önerilen yöntemin farklı aydınlatma ve özneler arası varyasyonlara karşı dirençli olduğunu göstermektedir. Elde edilen bulgular CHROM yönteminin anatomik bölgelere odaklandığında sunum saldırısı tespitinde sıfır hata hedefine yaklaşabildiğini somut verilerle doğrulamaktadır.

7.7 Hata Analizi ve Zorlu Senaryolar

Her bir çapraz doğrulama katmanı için kaydedilen hata log dosyaları ile yanlış sınıflandırılan video örneklerinin niteliksel analizi, önerilen hibrit modelin sınırlarının ve zayıf noktalarının belirlenmesinde önemli rol oynaktadır. Elde edilen bulgular doğrultusunda hata kaynakları ve zorlu senaryoların temel olarak üç ana faktör etrafında yoğunlaştığı belirlenmiştir.

İlk ve en belirgin hata kaynağı zorlu ve dengesiz aydınlatma koşullarıdır. Özellikle veri setlerinde “Adverse” ve “Lightoff” olarak etiketlenen senaryolardaki videolarda hata oranlarının anlamlı biçimde arttığı tespit edilmiştir. Bu tür koşullar hibrit mimarinin hem görsel hem de fizyolojik akışını eş zamanlı olarak olumsuz etkilemektedir. Düşük ışık seviyelerinde görsel akış, yüzey dokusu, mikro-yansımalar ve kenar geçişleri gibi sahtecilik tespitinde kritik rol oynayan ince uzamsal ipuçlarını kaybetmektedir. Buna paralel olarak fotometrik açıdan kamera sensöründe artan kuantum gürültüsü ve okuma hataları, rPPG sinyalinin genliğini ve faz kararlılığını bozarak sinyal-gürültü oranını kritik eşiklerin altına düşürmektedir. Bu durum özellikle fizyolojik akışın nabız kaynaklı periyodik bileşenleri güvenilir biçimde ayrıştırmasını engellemektedir. Sonuç olarak bu senaryolarda yanlış negatif oranının belirgin biçimde arttığı özellikle düşük ışık altında kaydedilmiş video tekrar oynatma saldırılarının tespitinin zorlaştığı gözlemlenmiştir. Bu

bulgu düşük aydınlatmanın sistem için birincil fiziksel kısıt oluşturduğunu ve gerçek dünya uygulamalarında ek önlemler gerektirdiğini göstermektedir.

İkinci önemli hata kaynağı, yüksek gerçekliğe sahip 3D maske saldırılarıdır. Özellikle 3DMAD veri setinde yer alan kişiye özel üretilmiş maskeler, doğru negatif oranını düşüren en kritik saldırı türü olarak öne çıkmaktadır. Bu maske saldırıları cilt dokusu, yüzey geometrisi ve ışık yansıma özelliklerini yüksek doğrulukla taklit edebilmekte ve bu sayede görsel akışın sahteciliğe özgü ipuçlarını bastırabilmektedir. Buna ek olarak 3DMAD veri setinde kullanılan Kinect sensörünün RGB renk tepkisinin, standart web kameraları ve mobil cihazlara kıyasla farklı bir spektral karakteristik sergilediği belirlenmiştir. Bu farklılık rPPG sinyal akışında renk kanalları arasındaki korelasyonu değiştirmekte ve fizyolojik sinyal çıkarımını zorlaştırmaktadır. Özellikle POS ve CHROM gibi projeksiyon tabanlı yöntemlerde bu sensör farkının karar eşiği hassasiyetini artırdığı gözlemlenmiştir. Bu bulgu önerilen modelin saldırı türlerinin yanı sıra kullanılan sensör ve donanım çeşitliliğine karşı da duyarlı olduğunu göstermektedir. Dolayısıyla farklı kamera sistemleriyle genellenebilirliğin artırılması gelecekteki çalışmalar için kritik bir araştırma alanı olarak öne çıkmaktadır.

Üçüncü ve son hata kaynağı konuşma, yoğun mimik kullanımı ve ani baş hareketleri içeren dinamik senaryolardır. Anatomik segmentasyon tabanlı seçici ROI stratejisi uygulanmasına rağmen büyük ölçekli ve ani hareketlerin etkisi tamamen izole edilememiştir. Bu tür hareketler rPPG sinyalinde genlik sıçramalarına ve spektral yayılmaya neden olarak kalp atımına özgü dar bantlı yapıyı bozmaktadır. Oluşan bu güçlü hareket artefaktları, modelin fizyolojik akışta elde edilen sinyali periyodik olmayan veya düzensiz bir desen olarak yorumlamasına yol açabilmektedir. Bu durum özellikle gerçek kullanıcıların sahte olarak etiketlenmesine neden olmakta ve yanlış ret oranlarını artırmaktadır. Yapılan incelemelerde bu hataların çoğunlukla konuşma sırasında ağız ve çene bölgesinde oluşan gölge ve deformasyonlardan kaynaklandığı belirlenmiştir. Bu kısıt ticari biyometrik sistemler açısından kullanıcı deneyimi, erişim konforu ve sistem kabul edilebilirliği bakımından en kritik iyileştirme alanı olarak öne çıkmaktadır. Hareket telafisi, adaptif ROI güncelleme veya ivme tabanlı hareket bastırma mekanizmalarının entegrasyonu, bu sınırlamaların aşılmasına yönelik gelecek çalışmalarda

değerlendirilmesi gereken öncelikli çözüm yolları arasında yer almaktadır.

7.8 Sistem Yanıt Süresi ve Hesaplama Maliyeti

Önerilen hibrit yüz canlılık tespit sisteminin gerçek zamanlı uygulanabilirliğini değerlendirmek amacıyla kapsamlı yanıt süresi ve hesaplama maliyeti analizleri gerçekleştirilmiştir. Bu kapsamda anatomik segmentasyon tabanlı seçici ROI stratejisi ile çift akışlı hibrit mimarinin bütünleşik biçimde çalıştığı senaryolar esas alınmıştır. Yapılan ölçümler ortalama 90 kareden oluşan tek bir video kesitini işleme süresinin yaklaşık 1,30 saniye olduğunu göstermektedir. Elde edilen bu yanıt süresi havaalanı pasaport kontrol noktaları veya bankacılık giriş sistemleri gibi yüksek güvenlik ve hızlı akış gerektiren geçiş kontrol senaryoları için literatürde kabul gören standartlar dahilindedir. Sistemin hesaplama maliyeti bileşenlerine ayrıştırıldığında MediaPipe tabanlı ROI çıkarım sürecinin toplam işlem süresi içinde ihmal edilebilir düzeyde olduğu görülmüştür. Buna karşılık asıl işlem yükünün rPPG sinyal dönüşümleri ve zamansal bağımlılıkları analiz eden LSTM katmanlarında yoğunlaştığı tespit edilmiştir. Bu durum sistemin uç cihazlarda dahi kabul edilebilir bir gecikme ile çalışabileceğini doğrulamaktadır.

8. TARTIŞMA ve SONUÇLAR

Bu tez çalışmasında yüz tanıma sistemlerini hedef alan sunum saldırılarına karşı dayanıklı ve yüksek doğruluklu bir doğrulama mekanizması geliştirmek amacıyla görsel ve fizyolojik öznitelikleri birleştiren hibrit bir derin öğrenme mimarisi tasarlanmıştır. Önerilen hibrit derin öğrenme mimarisi rPPG sinyalinden türetilen fizyolojik canlılık göstergeleri ile görsel akıştan çıkarılan uzamsal-zamansal doku temelli yüz temsillerini iki paralel kanalda işleyerek bütünleşik bir karar mekanizması oluşturmaktadır.

Çalışmanın literatüre sunduğu en temel ve özgün katkı rPPG yöntemlerinin literatürde bilinen en kritik zayıflıkları arasında yer alan baş hareketi, mimik deformasyonu ve aydınlatma değişimleri kaynaklı sinyal bozulmalarının anatomik segmentasyon tabanlı seçici ROI yaklaşımı ile sistematik biçimde minimize edilmesidir. Deneysel bulgular bütünsel yüz ROI yaklaşımında %14-15 bandında seyreden hata oranlarının önerilen seçici ROI stratejisi ile CHROM yöntemi için %4,16 seviyesine Yeşil Kanal yöntemi için ise %5,90 seviyesine kadar düşürülebildiğini göstermektedir. Bu yaklaşım sayesinde yalnızca nabız kaynaklı mikro varyasyonların en tutarlı biçimde izlendiği yüz segmentleri işleme alınmıştır. Böylece sinyal-gürültü oranı arttırılarak ritmik bileşenin faz sürekliliği korunmuş ve modelin fizyolojik canlılık göstergesini ayrıştırma kapasitesi güçlendirilmiştir.

Önerilen seçici ROI yaklaşımı Yu vd. (2021) tarafından sunulan global ilişki madenciliği yaklaşımına kıyasla daha deterministik ve hesaplama açısından daha düşük maliyetli bir çözüm sunmaktadır. Ayrıca MediaPipe landmarklarının her karede yüzü dinamik olarak takip edebilmesi sayesinde baş hareketleri sırasında dahi ROI'nin deri yüzeyi üzerinde sabit kalmasını sağlanmıştır. Bu yaklaşım Sun vd. (2025) tarafından önerilen landmark tabanlı yüz birleştirme yöntemiyle kavramsal olarak örtüşmekte olup gerçek zamanlı uygulamalar açısından çok daha düşük bir hesaplama maliyeti sunan bir yapı ortaya koymaktadır.

Hibrit mimarinin optimizasyonu sırasında görsel ve fizyolojik akışlar arasındaki zamansal senkronizasyonun korunması sağlanmıştır. Füzyon katmanında ise iki akıştan

elde edilen temsil vektörlerinin birleştirilmesiyle sahte–canlı yüz ayrımı tek akışlı modellere kıyasla daha kararlı ve daha genellenebilir bir biçimde gerçekleştirilmiştir. Bu bütünleşik yapı özellikle gerçek zamanlı senaryolarda görülen sinyal dalgalanmalarını minimize ederek düşük kaliteli sensörler ve değişken çevresel koşullar altında dahi tutarlı bir performans sergilemiştir. Elde edilen deneysel bulgular geliştirilen yöntemin fotoğraf, video ve 3D maske gibi heterojen saldırı vektörlerine karşı yüksek genellenebilirlik sunduğunu ortaya koymaktadır.

Nihai karar başarısını temsil eden ACER metriğinde CHROM yöntemi %4,16 hata oranı ve ± 2,84 standart sapma değeri ile en başarılı ve en kararlı yöntem olarak tespit edilmiştir. CHROM yönteminin renk farkı tabanlı yaklaşımı özellikle yüzdeki rijit olmayan hareketleri ve aydınlatma değişimlerinin oluşturduğu artefaktları baskılamada üstün bir dayanıklılık sergilemiştir.

Fizyolojik sinyali gürültüden ayrıştırma kapasitesini gösteren AUC ve EER metrikleri incelendiğinde POS yöntemi %99,71 AUC ve %2,52 EER değerleri ile teorik olarak en başarılı algoritma olarak öne çıkmaktadır. POS yönteminin sahip olduğu düzlem-ortogonal projeksiyon prensibi ideal koşullar altında en temiz sinyali üretme potansiyeline sahiptir. Ancak bu yüksek performansa karşın ± 6,05 olarak ölçülen yüksek standart sapma değeri yöntemin karar eşiğine karşı hassasiyet gösterdiğini ve veri setindeki varyasyonlardan daha fazla etkilendiğini göstermektedir. Bu durum, uygulama senaryosuna bağlı olarak yöntem seçiminin stratejik bir karar olduğunu göstermektedir. Düşük hata toleransı ve yüksek kararlılığın öncelikli olduğu güvenlik odaklı biyometrik sistemlerde CHROM tabanlı yaklaşımın buna karşılık fizyolojik analiz hassasiyetinin ön planda olduğu medikal uygulamalarda ise POS tabanlı yaklaşımın tercih edilmesinin daha uygun olacağı değerlendirilmektedir.

Literatürde öne çıkan güncel rPPG tabanlı ve hibrit yüz canlılık tespiti çalışmalarında kullanılan mimari yapılar, ön işleme teknikleri ve raporlanan performans metrikleri sistematik olarak incelenmiştir. Mevcut çalışmaların sunduğu yöntemler ile bu tez kapsamında geliştirilen özgün mimarinin karşılaştırmalı performans analizi Tablo 8.1'de detaylandırılmıştır.

Tablo 8.1 rPPG ve doku/hareket özelliklerini birleştiren yöntemler

Çalışma / Kaynak	Model Mimarisi	Ön İşleme / Hizalama	rPPG Yöntemi	Füzyon Mekanizması	Kullanılan Veri Setleri (Test)	Doğrulama Tekniği	EER (%)	AUC (%)
(Sun vd. 2025)	Uzamsal - Zamansal CNN ve Çok Ölçekli CDC	OpenFace ve CHROM Uzay Dönüşümü	Faz Odaklı CNN	Mekânsal Dikkat Füzyonu	3DMAD, CSMAD, HKBU-Mars	Veri Kümeleri Arası Test	2,60 (3DMAD)	99,50 (3DMAD)
(Yu vd. 2021)	rPPG Transformer (Görsel Transformer)	MSTmaps	Öğrenilmiş rPPG	İki Akışlı MSTmap (Yüz ve Arka Plan)	3DMAD, HKBU-MARsV2	Veri Kümeleri Arası Test	5,93 (3DMAD)	97,65 (3DMAD)
(Yu vd. 2024)	İki Kollu ViT (Görsel ve Fizyolojik)	MTCNN Yüz Algılayıcı	MSTmap ve WaveletMap	Ağırlıklandırılırmış Parti Normalizasyonu	SiW, 3DMAD, FF++, CelebDFv2	Etki Alanı ve Kümeler Arası Test	12,16 (Joint)	91,54 (Joint)
(Lin vd. 2023)	İki Akışlı Yöntem (Mekânsal ve Zamansal Akış)	Yüz Hizalama, ROI Filtresi	STMap Üretici	Skor Füzyonu	PURE, UBFC-RPPG, RErPPG-Net	Tek Veri Kümesi İçi Test	12,55	94,84
(Gomez vd. 2023)	Evrişimsel Dikkat Ağı (CAN)	rPPG Çıkarımı, ROI Takibi	Öğrenilmiş rPPG (Domain Transfer)	Alan Transferi	Deepfakes Etki Alanı	Etki Alanı İçi ve Arası Test	19,32 (ACER)	—
(Gündoğar ve Erdem 2021)	Geleneksel Özellik Tabanlı (SVM)	Viola -Jones Yüz Tespiti	2SR, CHROM, Li's CVPR14	Tek Akış	3DMAD, Replay-Attack, MSU-MFSD	LOOCV (3DMAD), Grandtest (Replay)	11,25 (Replay)	—
Önerilen Model	MobileNetV2 ve Dual LSTM	Yüz tespiti, Yüz hizalama, Yeniden boyutlandırma	CHROM	Öznitelik Birleştirme	Replay-Mobile, 3DMAD, PURE, UBFC	Kişiden Bağımsız 5-Katlı Çapraz Doğrulama	2,86	99,36

Tablo 8.1'deki sonuçlar incelendiğinde önerilen modelin, özellikle heterojen veri setleri ve kişi bağımsız doğrulama protokolleri altında rekabetçi ve çoğu durumda daha yüksek performans sergilediği görülmektedir. Sun vd. (2025) ve Yu vd. (2021) çalışmalarının özellikle 3DMAD gibi tekil ve kontrollü veri setleri üzerinde sırasıyla %2,60 ve %5,93 gibi düşük hata oranlarına ulaştığı görülmektedir. Ancak bu yöntemlerin başarısı büyük ölçüde homojen veri dağılımına sahip intra-dataset senaryolarla sınırlı kaldığı görülmektedir. Buna karşılık önerilen hibrit model Replay-Mobile, 3DMAD, PURE ve

UBFC-RPPG gibi farklı aydınlatma koşullarını, sensör tiplerini ve hareket örüntülerini içeren heterojen bir veri havuzunda kişiden bağımsız 5-katlı çapraz doğrulama protokolüyle test edilmiştir. Bu zorlu doğrulama şemasına rağmen elde edilen %2,86 EER, %4,16 ACER ve %99,36 AUC değerleri, modelin yüksek genellenebilirlik kapasitesine sahip olduğunu göstermektedir.

Sonuç olarak bu tez kapsamında önerilen mimari, hesaplama maliyeti yüksek transformer tabanlı veya çok ölçekli karmaşık CNN mimarileri yerine MobileNetV2, seçici ROI ve CHROM tabanlı rPPG çıkarımını özgün bir öznitelik birleştirme stratejisiyle birleştirerek daha hafif, daha hızlı ve yüksek doğruluklu bir çözüm sunmaktadır. Elde edilen bulgular, önerilen yaklaşımın gerçek zamanlı biyometrik güvenlik sistemleri için pratik, ölçeklenebilir ve güvenilir bir çözüm olduğunu göstermektedir.

8.1 Çalışmanın Kısıtlılıkları

Bu çalışmada elde edilen yüksek doğruluk ve genellenebilirliğe rağmen, önerilen yaklaşımın literatürdeki diğer yöntemlerle karşılaştırıldığında bazı sınırlılıkları bulunmaktadır. Öncelikle modelin 3DMAD veri setindeki Kinect sensörü ile kaydedilmiş gerçek yüz örneklerini sınıflandırmada standart RGB web kameralarına kıyasla görece daha düşük performans sergilediği gözlemlenmiştir. Bu durum modelin farklı spektral özelliklere sahip kameralar için yeniden kalibre edilmesi gerektiğini ve alan kayması problemine karşı daha güçlü bir uyum mekanizması tasarlanması gerektiğini göstermektedir.

Bir diğer önemli kısıtlılık Replay-Mobile veri setindeki düşük aydınlatma koşullarında rPPG sinyal kalitesinin belirgin biçimde düşmesi ve buna bağlı olarak yanlış reddetme oranlarını artmasıdır. Bu durum aydınlatma değişimlerinin fizyolojik sinyal kalitesine doğrudan etki ettiğini ve düşük ışık koşullarının canlılık tespit sistemleri için kritik bir zorluk olduğunu göstermektedir.

Son olarak önerilen modelin ortalama 1,30 saniye süren işlem süresi, havaalanı veya banka girişlerindeki yüksek güvenlikli geçiş sistemleri için uygun olmakla birlikte anlık

mobil kilit açma veya kullanıcı etkileşimli senaryolar için halen optimizasyona ihtiyaç duymaktadır. Bu durum modelin parametre sayısının ve saniyedeki kayan nokta işlem miktarının daha da azaltılması gerektiğini göstermektedir.

8.2 Gelecek Çalışmalar İçin Öneriler

Elde edilen deneysel bulgular, mevcut çalışmanın sınırları ve literatürdeki güncel eğilimler göz önüne alındığında gelecekte yürütülecek araştırmalar için çeşitli stratejik geliştirme alanları ön plana çıkmaktadır. Görüntüleme sürecinde RGB sensörlerin yanı sıra yakın kızılötesi ve termal kameraların sisteme entegre edilmesi canlılık tespitinin güvenilirliğini önemli ölçüde artırma potansiyeline sahiptir. Özellikle termal damar haritalarının kullanımı, 3D maske ve silikon protezlerin gerçek insan dokusu sıcaklığını taklit edememesi nedeniyle sahtecilik tespitinde belirgin bir ayırt edicilik sağlayacaktır.

Modelin mobil ve gömülü sistemlerde gerçek zamanlı çalışabilirliğini artırmak adına MobileNetV3-Small veya EfficientNet-Lite gibi daha hafif omurga mimarileri tercih edilerek modelin mobil cihazlarda çalışabilir hâle gelmesi sağlanabilir. Ayrıca eğitim sonrası kuantizasyon ve ağ budama teknikleri uygulanarak model başarımından ödün vermeden gecikme süresinin %30-40 oranında azaltılması hedeflenebilir.

Mevcut çalışma sabit koordinatlı anatomik segmentlere dayalı bir ROI stratejisi kullanmaktadır. Gelecek çalışmalarda görüntüdeki baş hareketi, lokal ışık değişimi ve cilt görünürlüğüne göre ilgi alanlarını dinamik olarak güncelleyen dikkat tabanlı bir modül geliştirilebilir. Böyle bir mekanizma sinyal gürültüsünü minimize ederken değişken senaryolarda daha tutarlı rPPG verisi elde edilmesine olanak tanıyabilecektir.

Son olarak bu tez kapsamında oluşturulan heterojen veri havuzu üzerinde gerçekleştirilen kapsamlı deneyler önerilen hibrit mimarinin mevcut saldırı türlerine karşı başarısını istatistiksel olarak doğrulamıştır. Ancak biyometrik tehdit vektörlerinin çeşitliliği ve görüntüleme teknolojilerindeki hızlı ilerlemeler göz önüne alındığında gelecekteki çalışmaların yeni nesil yüksek gerçeklikli silikon maskeler, mikro-doku taklidi yapabilen üretim materyalleri ve ultra-yüksek çözünürlüklü ekran tabanlı saldırılar ile

zenginleştirilmiş veri setleri üzerinde test edilmesi büyük önem taşımaktadır. Bu tür ek doğrulama çalışmaları, sistemin en zorlu sınır koşullarındaki dayanıklılığını ortaya koyarak evrensel ölçekte güvenilirliğini pekiştirecektir.

9. KAYNAKLAR

Adjabi I, Ouahabi A, Benzaoui A, Taleb-Ahmed A, 2020, Past, Present, and Future of Face Recognition: A Review, Electronics, 9, 1188.

Akgül A, 2015, Kişisel Verilerin Korunması Bağlamında Biyometrik Yöntemlerin Kullanımı ve Danıştay Yaklaşımı, Türkiye Barolar Birliği Dergisi, 199-222.

Akın E, Şahin ME, 2024, Derin Öğrenme ve Yapay Sinir Ağı Modelleri Üzerine Bir İnceleme, EMO Bilimsel Dergi, 14, 27-38.

Allen J, 2007, Photoplethysmography and its application in clinical physiological measurement, Physiological Measurement, 28, R1-R39.

Anthony P, Ay B, Aydin G, 2021, A Review of Face Anti-spoofing Methods for Face Recognition Systems, 2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA), 1-9.

Antil A, Dhiman C, 2025, Unmasking Deception: A Comprehensive Survey on the Evolution of Face Anti-spoofing Methods, Neurocomputing, 617, 128992.

Arslan B, Sağıroğlu Ş, 2016, Mobil Cihazlarda Biyometrik Sistemler Üzerine Bir İnceleme, Politeknik Dergisi, 19, 101-114.

Ben Fredj H, Bouguezzi S, Souani C, 2021, Face Recognition in Unconstrained Environment with CNN, The Visual Computer, 37, 217-226.

Bobbia S, Macwan R, Benezeth Y, Mansouri A, Dubois J, 2019, Unsupervised skin tissue segmentation for remote photoplethysmography, Pattern Recognition Letters, 124, 82-90.

Botina-Monsalve D, Benezeth Y, Miteran J, 2022, RTrPPG: An Ultra Light 3DCNN for Real-Time Remote Photoplethysmography, 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2145-2153.

Boyer AO, Boyer RS, 1991, A Biographical Sketch of W. W. Bledsoe, Automated Reasoning, Springer, Dordrecht, 1-29.

Busch C, 2023, Standards for Biometric Presentation Attack Detection, Handbook of Biometric Anti-Spoofing: Presentation Attack Detection and Vulnerability

Assessment, Springer Nature, Singapore, 571-583.

Castellano Ontiveros R, Elgendi M, Menon C, 2024, A Machine Learning-Based Approach for Constructing Remote Photoplethysmogram Signals from Video Cameras, Communications Medicine, 4, 1-8.

Chen W, Yi Z, Lim LJR, Lim RQR, Zhang A, Qian Z, vd., 2024, Deep Learning and Remote Photoplethysmography Powered Advancements in Contactless Physiological Measurement, Frontiers in Bioengineering and Biotechnology, 12, 1420100.

Chingovska I, Anjos A, Marcel S, 2012, On the Effectiveness of Local Binary Patterns in Face Anti-Spoofing, BIOSIG - Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG, 1-7.

Chu S, Xia M, Yuan M, Liu X, Seppanen T, Zhao G, vd., 2025, CodePhys: Robust Video-based Remote Physiological Measurement through Latent Codebook Querying.

Comas J, Alomar A, Ruiz A, Sukno F, 2024, PhysFlow: Skin tone transfer for remote heart rate estimation through conditional normalizing flows, arXiv.

Costa-Pazo A, Bhattacharjee S, Vazquez-Fernandez E, Marcel S, 2016, The Replay-Mobile Face Presentation-Attack Database, 2016 International Conference of the Biometrics Special Interest Group (BIOSIG), Darmstadt, 1-7.

Çiçek M, 2024, Türkiye'de Biyometrik Veri Güvenliği: Kişisel Verilerin Korunması Kanunu Çerçevesinde Etik Bir Değerlendirme, Kişisel Verileri Koruma Dergisi, 6, 54-76.

Davis J, Goadrich M, 2006, The relationship between Precision-Recall and ROC curves, Proceedings of the 23rd international conference on Machine learning, New York, NY, USA, 233-240.

Debnath U, Kim S, 2025, A comprehensive review of heart rate measurement using remote photoplethysmography and deep learning, BioMedical Engineering OnLine, 24, 73.

Demirezen H, 2022, Remote Heart Rate Estımatıon Using Non-Contact Photoplethysmography, Marmara Üniversitesi, ISTANBUL, Yüksek Lisans Tezi,

97 s., ISTANBUL.

Deng J, Guo J, Xue N, Zafeiriou S, 2019, ArcFace: Additive Angular Margin Loss for Deep Face Recognition, 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 4685-4694.

Erdogmus N, Marcel S, 2013, Spoofing in 2D face recognition with 3D masks and anti-spoofing with Kinect, 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), 1-6.

Forman G, Scholz M, 2010, Apples-to-Apples in Cross-Validation Studies: Pitfalls in Classifier Performance Measurement, Association for Computing Machinery, 49-57.

Geng M, Hu G, Wang G, Wang Y, 2024, Efficient rPPG-Based Biometric Authentication Using Lightweight Vision Transformer, 2024 4th International Conference on Electronic Information Engineering and Computer Communication (EIECC), 488-491.

Evlioğlu Gezer, E, 2025, Sosyal Medyada Yapay Zekâ Kullanımı: Uygulama Alanları ve Sosyal Etkileri, İletişim ve Toplum Araştırmaları Dergisi, 5, 516-528.

Gomez LF, Fierrez J, Morales A, Ghafourian M, Tolosana R, Solano I, vd., 2023, PAD-Phys: Exploiting Physiology for Presentation Attack Detection in Face Biometrics, 2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC), Torino, Italy, 1669-1674.

Gomez-Barrero M, Drozdowski P, Rathgeb C, Patino J, Todisco M, Nautsch A, vd., 2022, Biometrics in the Era of COVID-19: Challenges and Opportunities, IEEE Transactions on Technology and Society, 3, 307-322.

Guo G, Zhang N, 2019, A Survey on Deep Learning Based Face Recognition, Computer Vision and Image Understanding, 189, 102805.

Günay Yılmaz A, Gedikli E, Alhori O, 2023, Yüz Görüntülerine Morflemeye Dayalı Maske Giydirme ve Maskeli Yüz Tanıma, Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 27, 12-21.

Gündoğar M, Erdem Ç, 2021, Presentation attack detection for face recognition using

remotephotoplethysmography and cascaded fusion, Turkish Journal of Electrical Engineering and Computer Sciences, 29, 3240-3258.

Güneş A, 2025, Deepfake'in Elektronik Ortamda Yapay Zekâ Tabanlı Kimlik Doğrulama Sürecine Etkisi, Bilgi Teknolojileri ve İletişim Dergisi, 3, 1-34.

Haan G de, Jeanne V, 2013, Robust Pulse Rate From Chrominance-Based rPPG, IEEE Transactions on Biomedical Engineering, 60, 2878-2886.

Hanley JA, McNeil BJ, 1982, The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve, Radiology, 143, 29-36.

Hernandez-Ortega J, Daza R, Morales A, Fierrez J, Tolosana R, 2020, Heart Rate Estimation from Face Videos for Student Assessment: Experiments on edBB, 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), 172-177.

Hernandez-Ortega J, Fierrez J, Morales A, Galbally J, 2019, Introduction to Face Presentation Attack Detection, Handbook of Biometric Anti-Spoofing: Presentation Attack Detection, Springer International Publishing, Cham, 187-206.

Hernandez-Ortega J, Fierrez J, Morales A, Tome P, 2018, Time Analysis of Pulse-Based Face Anti-Spoofing in Visible and NIR, 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, 657-6578.

Hertzman AB, 1937, Photoelectric Plethysmography of the Fingers and Toes in Man, Proceedings of the Society for Experimental Biology and Medicine, 37, 529-534.

Hochreiter S, Schmidhuber J, 1997, Long Short-Term Memory, Neural Computation, 9, 1735-1780.

Huang P-K, Chen T-H, Chan Y-T, Chen K-W, Hsu C-T, 2025, DD-rPPGNet: De-Interfering and Descriptive Feature Learning for Unsupervised rPPG Estimation, IEEE Transactions on Information Forensics and Security, 20, 4956-4970.

Jiang F, Li Q, Liu B, Wang W, Shan C, Sun Z, vd., 2025, Learning Knowledge-based Prompts for Robust 3D Mask Presentation Attack Detection, IEEE Transactions on Pattern Analysis and Machine Intelligence.

Jin X, Tan X, 2017, Face alignment in-the-wild: A Survey, Computer Vision and Image Understanding, 162, 1-22.

Kamat C, 2024, Face Anti-Spoofing Methods: A Comparative Analysis through the Lens of a Comprehensive Review, International Journal for Research in Applied Science and Engineering Technology, 12, 514-526.

Karhunen J, Oja E, Wang L, Vigario R, Joutsensalo J, 1997, A Class of Neural Networks for Independent Component Analysis, IEEE Transactions on Neural Networks, 8, 486-504.

Khan M, Saeed M, El Saddik A, Gueaieb W, 2023, ARTriViT: Automatic Face Recognition System Using ViT-Based Siamese Neural Networks with a Triplet Loss, 2023 IEEE 32nd International Symposium on Industrial Electronics (ISIE), 1-6.

Khan S, Siddique THM, Ibrahim MS, Siddiqui AJ, Huang K, 2025, Spatio-temporal deep learning for improved face presentation attack detection, Knowledge-Based Systems, 311, 113059.

Kim D-Y, Lee K, Sohn C-B, 2021, Assessment of ROI Selection for Facial Video-Based rPPG, Sensors, 21, 7923.

Kim S-H, Jeon S-M, Lee EC, 2022, Face Biometric Spoof Detection Method Using a Remote Photoplethysmography Signal, Sensors, 22, 3070.

Kohavi R, 1995, A Study of Cross Validation and Bootstrap for Accuracy Estimation and Model Selection, International Joint Conference on Artificial Intelligence.

Kooij KM van der, Naber M, 2019, An Open-Source Remote Heart Rate Imaging Method with Practical Apparatus and Algorithms, Behavior Research Methods, 51, 2106-2119.

Kortli Y, Jridi M, Al Falou A, Atri M, 2020, Face Recognition Systems: A Survey, Sensors, 20, 342.

Kossack B, Wisotzky E, Hilsmann A, Eisert P, 2021, Automatic Region-Based Heart Rate Measurement Using Remote Photoplethysmography, 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal,

BC, Canada, 2755-2759.

Kuang H, Ao C, Ma X, Liu X, 2023a, Remote photoplethysmography signals enhancement based on generative adversarial networks, 2023 IEEE 3rd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA), 792-796.

Kuang H, Ao C, Ma X, Liu X, 2023b, Shuffle-rPPGNet: Efficient Network With Global Context for Remote Heart Rate Variability Measurement, IEEE Sensors Journal, 23, 15199-15209.

Kurshan E, Mehta D, Balch T, 2024, AI versus AI in Financial Crimes & Detection: GenAI Crime Waves to Co-Evolutionary AI, Proceedings of the 5th ACM International Conference on AI in Finance, New York, NY, USA, 745-751.

Lam A, Kuno Y, 2015, Robust Heart Rate Measurement from Video Using Select Random Patches, 2015 IEEE International Conference on Computer Vision (ICCV), 3640-3648.

Lecun Y, Bottou L, Bengio Y, Haffner P, 1998, Gradient-based learning applied to document recognition, Proceedings of the IEEE, 86, 2278-2324.

Lee K, Oh J, You H, Lee EC, 2023, Improving Remote Photoplethysmography Performance through Deep-Learning-Based Real-Time Skin Segmentation Network, Electronics, 12, 3729.

Lee RJ, Sivakumar S, Lim KH, 2023b, Review on Remote Heart Rate Measurements Using Photoplethysmography, Multimedia Tools and Applications, 83, 44699-44728.

Lee S-H, Yun G, Park SH, Lim MY, Lee YK, 2025, Towards Robust Deepfake Detection Based on Heart Rate Analysis., KSII Transactions on Internet & Information Systems, 19, 191-212.

Li L, Chen C, Pan L, Zhang LY, Wang Z, Zhang J, vd., 2023, A Survey of PPG's Application in Authentication, Computers & Security, 135, 103488.

Li L, Yao Z, Gao S, Han H, Xia Z, 2024, Face Anti-Spoofing via Jointly Modeling Local Texture and Constructed Depth, Engineering Applications of Artificial

Intelligence, 133, 108345.

Li X, Chen J, Zhao G, Pietikainen M, 2014, Remote Heart Rate Measurement from Face Videos under Realistic Situations, 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 4264-4271.

Li X, Komulainen J, Zhao G, Yuen P-C, Pietikäinen M, 2016, Generalized face anti-spoofing by detecting pulse from face videos, 2016 23rd International Conference on Pattern Recognition (ICPR), 4244-4249.

Lin Y, Maiorana E, Li B, Campisi P, 2023, Detection of Photoplethysmography Manipulation in Video Forgery, 2023 31st European Signal Processing Conference (EUSIPCO), Helsinki, Finland, 615-619.

Liu A, Tan Z, Wan J, Liang Y, Lei Z, Guo G, vd., 2021, Face Anti-Spoofing via Adversarial Cross-Modality Translation, IEEE Transactions on Information Forensics and Security, 16, 2759-2772.

Liu S, Yang B, Yuen PC, Zhao G, 2016, A 3D Mask Face Anti-Spoofing Database with Real World Variations, 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Las Vegas, NV, USA, 1551-1557.

Liu S-Q, Lan X, Yuen PC, 2022, Learning Temporal Similarity of Remote Photoplethysmography for Fast 3D Mask Face Presentation Attack Detection, IEEE Transactions on Information Forensics and Security, 17, 3195-3210.

Liu X, Zuo J, Ma X, Kuang H, 2024, Uni-rPPGNet: Efficient and Lightweight Remote Heart Rate Variability Measurement, 2024 IEEE 7th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), 19-23.

Ma R, Chen C, 2025, RF-BayesPhysNet: A Bayesian rPPG Uncertainty Estimation Method for Complex Scenarios, arXiv.

Ma X, Wang Z, Liu X, Kuang H, 2024, CMRPPGFormer: 3-D Spatio-Temporal Convolutional Modulation Transformer Network for Remote Heart Rate Estimation, IEEE Sensors Journal, 24, 30275-30286.

Maity AK, Wang J, Sabharwal A, Nayar SK, 2022, RobustPPG: camera-based robust heart rate estimation using motion cancellation, Biomedical Optics Express, 13,

5447-5467.

Maniatopoulos A, Mitianoudis N, 2021, Learnable Leaky ReLU (LeLeLU): An Alternative Accuracy-Optimized Activation Function, Information, 12, 513.

Marcel S, Nixon MS, Fierrez J, Evans N (ed.), 2019, Handbook of Biometric Anti-Spoofing: Presentation Attack Detection, Springer International Publishing, Cham.

Masood M, Nawaz M, Malik KM, Javed A, Irtaza A, Malik H, 2023, Deepfakes Generation and Detection: State-of-the-Art, Open Challenges, Countermeasures, and Way Forward, Applied Intelligence, 53, 3974-4026.

Menakadevi B, Kumar DS, Nagasaratha P, Parimalam K, 2025, Biometric System Attacks-A Case Study, 2025 International Conference on Computer, Electrical & Communication Engineering (ICCECE), 1-7.

Ming Z, Visani M, Luqman MM, Burie J-C, 2020, A Survey on Anti-Spoofing Methods for Facial Recognition with RGB Cameras of Generic Consumer Devices, Journal of Imaging, 6, 139.

Nair V, Hinton GE, 2010, Rectified linear units improve restricted boltzmann machines, Proceedings of the 27th International Conference on International Conference on Machine Learning, Madison, WI, USA, 807-814.

Nowara EM, Sabharwal A, Veeraraghavan A, 2017, PPGSecure: Biometric Presentation Attack Detection Using Photopletysmograms, 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017), 56-62.

Ontiveros RC, Elgendi M, Missale G, Menon C, 2023, Evaluating RGB Channels in Remote Photoplethysmography: A Comparative Study with Contact-Based PPG, Frontiers in Physiology, 14.

Patel Y, Tanwar S, Gupta R, Bhattacharya P, Davidson IE, Nyameko R, vd., 2023, Deepfake Generation and Detection: Case Study and Challenges, IEEE Access, 11, 143296-143323.

Phillips PJ, Moon H, Rizvi SA, Rauss PJ, 2000, The FERET evaluation methodology for face-recognition algorithms, IEEE Transactions on Pattern Analysis and Machine

Intelligence, 22, 1090-1104.

Phillips PJ, Scruggs WT, O'Toole AJ, Flynn PJ, Boyer KW, Schott CL, vd., 2007, FRVT 2006 and ICE 2006 Large-Scale Results, National Institute of Standards and Technology, Gaithersburg, MD, NIST IR 7408.

Poh M-Z, McDuff DJ, Picard RW, 2010, Non-Contact, Automated Cardiac Pulse Measurements Using Video Imaging and Blind Source Separation., Optics Express, 18, 10762-10774.

Poh M-Z, McDuff DJ, Picard RW, 2011, Advancements in Noncontact, Multiparameter Physiological Measurements Using a Webcam, IEEE Transactions on Biomedical Engineering, 58, 7-11.

Pooshideh M, Beheshti A, Qi Y, Farhood H, Simpson M, Gatland N, vd., 2024, Presentation Attack Detection: A Systematic Literature Review, ACM Comput. Surv., 57, 25:1-25:32.

Savic M, Zhao G, 2025, RS+rPPG: Robust Strongly Self-Supervised Learning for rPPG, IEEE Transactions on Circuits and Systems for Video Technology, 35, 7911-7924.

Schroff F, Kalenichenko D, Philbin J, 2015, FaceNet: A Unified Embedding for Face Recognition and Clustering, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 815-823.

Schuckers S, 2016, Presentations and attacks, and spoofs, oh my, Image Vision Comput., 55, 26-30.

Seibold C, Wisotzky EL, Beckmann A, Kossack B, Hilsmann A, Eisert P, 2025, High-Quality Deepfakes Have a Heart!, Frontiers in Imaging, 4, 1504551.

Shao H, Luo L, Qian J, Yan M, Chen S, Yang J, 2025, Remote Photoplethysmography in Real-World and Extreme Lighting Scenarios, arXiv.

Sharma D, Selwal A, 2023, A Survey on Face Presentation Attack Detection Mechanisms: Hitherto and Future Perspectives, Multimedia Systems, 29, 1527-1577.

Shinde SR, Bongale AM, Dharrao D, Thepade SD, 2025, An enhanced light weight face

liveness detection method using deep convolutional neural network, MethodsX, 14, 103229.

Siddiqi MH, Khan K, Khan RU, Alsirhani A, 2022, Face Image Analysis Using Machine Learning: A Survey on Recent Trends and Applications, Electronics, 11, 1210.

Stricker R, Müller S, Gross H-M, 2014, Non-contact video-based pulse rate measurement on a mobile service robot, The 23rd IEEE International Symposium on Robot and Human Interactive Communication, 1056-1062.

Sun R, Yu X, Feng H, Wang F, Zhang X, 2025, Motion-Robust Mask Face Presentation Attack Detection via Dual-Stream Texture-rPPG Network, The Visual Computer, 41, 4517-4532.

Sun Y, Thakor N, 2016, Photoplethysmography revisited: from contact to noncontact, from point to imaging, IEEE transactions on bio-medical engineering, 63, 463-477.

Sun Z, Li X, Komulainen J, Zhao G, 2024, Biometric Authentication Based on Enhanced Remote Photoplethysmography Signal Morphology, 2024 IEEE International Joint Conference on Biometrics (IJCB), Buffalo, NY, USA, 1-10.

Taylor J, 2019, Major Breach Found in Biometrics System Used by Banks, UK Police and Defence Firms, The Guardian.

Tian Y, Kuang H, Ma X, Liu X, 2024, GSW-UNet: An Improved UNet Model Based on PPG Signal for Arterial Blood Pressure Waveform Estimation, 2024 IEEE 7th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), 24-28.

Tiraki Y, Bakır Ç, Serttaş S, Temurtaş H, 2022, Evrişimsel Sinir Ağları ile Otomatik Yüz Tanıma Sistemi, International Journal of Engineering Research and Development, 14, 219-224.

Trigueros DS, Meng L, Hartnett M, 2018, Face Recognition: From Traditional to Deep Learning Methods, arXiv.Org, https://arxiv.org/abs/1811.00116v1, Erişim tarihi 22.05.2025.

Tsitiridis A, Conde C, Gomez Ayllon B, Cabello E, 2019, Bio-Inspired Presentation

Attack Detection for Face Biometrics, Frontiers in Computational Neuroscience, 13.

Turk M, Pentland A, 1991, Eigenfaces for Recognition, Journal of Cognitive Neuroscience, 3, 71-86.

Verkruysse W, Svaasand LO, Nelson JS, 2008, Remote plethysmographic imaging using ambient light, Optics express, 16, 21434-21445.

Viola P, Jones MJ, 2004, Robust Real-Time Face Detection, International Journal of Computer Vision, 57, 137-154.

Wang K, Tang J, Fan Y, Ji J, Shi Y, Wang Y, 2025, Memory-efficient Low-latency Remote Photoplethysmography through Temporal-Spatial State Space Duality, arXiv.

Wang W, Brinker AC den, Stuijk S, Haan G de, 2017, Algorithmic Principles of Remote PPG, IEEE Transactions on Biomedical Engineering, 64, 1479-1491.

Wang W, Stuijk S, Haan G de, 2016, A Novel Algorithm for Remote Photoplethysmography: Spatial Subspace Rotation, IEEE Transactions on Biomedical Engineering, 63, 1974-1984.

Wu H-Y, Rubinstein M, Shih E, Guttag J, Durand F, Freeman W, 2012, Eulerian Video Magnification for Revealing Subtle Changes in the World, MIT Web Domain, 31, 65:1-65:8.

Wu X, Feng X, Casado CÁ, Liu L, López MB, 2024, Facial Kinship Verification from remote photoplethysmography, arXiv.

Xiang Z, Tan H, Ye W, 2018, The Excellent Properties of a Dense Grid-Based HOG Feature on Face Recognition Compared to Gabor and LBP, IEEE Access, 6, 29306-29319.

Xiao H, Liu T, Sun Y, Li Y, Zhao S, Avolio A, 2024, Remote photoplethysmography for heart rate measurement: A review, Biomedical Signal Processing and Control, 88, 105608.

Yalçın N, Gürbüz F, 2015, Biyometrik Güvenlik Sistemlerinin İncelenmesi, Duzce University Journal of Science and Technology, 3, 398-413.

Yang W, Wang S, Sahri NM, Karie NM, Ahmed M, Valli C, 2021, Biometrics for Internet-of-Things Security: A Review, Sensors, 21, 6163.

Yang Z, Ge W, Zhang Z, 2020, Face Recognition Based on MTCNN and Integrated Application of FaceNet and LBP Method, 2020 2nd International Conference on Artificial Intelligence and Advanced Manufacture (AIAM), 95-98.

Yıldırım S, 2024, Bankacılıkta Uzaktan Kimlik Tespitinde Karşılaşılan Riskler ve Çözüm Önerileri, Mülkiye Dergisi, 48, 243-276.

Yu Z, Cai R, Li Z, Yang W, Shi J, Kot AC, 2024, Benchmarking Joint Face Spoofing and Forgery Detection With Visual and Physiological Cues, IEEE Transactions on Dependable and Secure Computing, 21, 4327-4342.

Yu Z, Li X, Wang P, Zhao G, 2021, TransRPPG: Remote Photoplethysmography Transformer for 3D Mask Face Presentation Attack Detection, IEEE Signal Processing Letters, 28, 1290-1294.

Yün M, 2023, Suç Soruşturmalarında Veri Madenciliği, Yapay Zeka, Uygulamaları ve Geleceği Python ve Opencv ile Gerçek Zamanlı Yüz Tanıma Uygulaması, Hitit Üniversitesi, Çorum, Yüksek Lisans Tezi, 118 s., Çorum.

Zhang Q, Lin X, Zhang Y, Liu Q, Cai F, 2023, Non-Contact High Precision Pulse-Rate Monitoring System for Moving Subjects in Different Motion States, Medical & Biological Engineering & Computing, 61, 2769-2783.

ÖZGEÇMİŞ

Adı Soyadı : Sena Özkara

Doğum Yeri ve Tarihi : Afyonkarahisar

Yabancı Dili : İngilizce

İletişim (Telefon / e-posta) : senaaozkaraa@outlook.com

Eğitim Durumu (Kurum ve Yıl)

Lise : Osmangazi Anadolu Lisesi (2014 – 2018)

Lisans : Burdur Mehmet Akif Ersoy Üniversitesi, Bilgisayar Mühendisliği Bölümü, (2019 – 2023)

Yayınları (SCI ve diğer) : Özkara S, Baysan E, 2024, İlk Okumayı Öğretmeye Yönelik Oyunlaştırılmış Bir Mobil Uygulama Tasarımı, *Dijital Teknolojiler ve Eğitim Dergisi*, 3, 140-153.