KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

YÜZ BİYOMETRİK SİSTEMLERİNDE SAHTECİLİKLERİN TESPİTİ

DOKTORA TEZİ

Uğur TURHAL

OCAK 2024
TRABZON

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

YÜZ BİYOMETRİK SİSTEMLERİNDE SAHTECİLİKLERİN TESPİTİ

Uğur TURHAL

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünce
"DOKTOR (BİLGİSAYAR MÜHENDİSLİĞİ)"
Unvanı Verilmesi İçin Kabul Edilen Tezdir.

Tezin Enstitüye Verildiği Tarih : 11 / 12 / 2024
Tezin Savunma Tarihi : 03 / 01 / 2024

Tez Danışmanı : Doç. Dr. Vasif NABİYEV

Trabzon 2024

ÖNSÖZ

Bu tez çalışmasında, yüz biyometrik sistemlerinde sahteciliklerin tespiti üzerine araştırmalar yapılmış ve problemi çözümlemek üzere bir dizi yöntem önerilmiştir.

Tez sürecim boyunca bilgisini ve desteklerini benden esirgemeyen danışmanım Doç. Dr. Vasif NABİYEV'e, sunduğu değerli katkılar için teşekkür ederim.

Doktora sürecimin başından sonuna kadar çalışmalarımla ilgilenen, yol gösteren ve her soruma yılmadan cevap veren, veri setlerinin ediniminde ve kullanılacak yöntemlerin açıklanmasında katkılarını sunan kıymetli hocam Dr. Öğr. Üyesi Asuman GÜNAY YILMAZ'a teşekkürlerimi bir borç bilirim.

Tezime katkıda bulunan değerli hocalarım Prof. Dr. Ayhan İSTANBULLU'ya, Prof. Dr. Bilal ALATAŞ'a ve Doç. Dr. Bekir DİZDAROĞLU'na, hoşgörülü yaklaşımları, değerli yorumları ve anlayışlarından dolayı teşekkürlerimi saygı ve hürmetle sunarım.

Çalışmamın her aşamasında bilgisini ve desteğini benimle paylaşan kıymetli arkadaşım Dr. Öğr. Üyesi Ramazan Özgür DOĞAN'a şükranlarımı sunarım.

2018 yılından bu yana yanımda olan ve her durumda varlıklarını hissettiren değerli arkadaşlarım Dr. Öğr. Üyesi Bilal TAYFUR'a, Dr. Öğr. Üyesi Nurullah ÖKSÜZER'e, Öğr. Gör. Dr. Muhammed Reşit ÇORAPSIZ'a, Öğr. Gör. İbrahim ARSLAN'a, Öğr. Gör. Recep ASLANOĞLU'na ve isimlerini yazmaya imkân bulamadığım tüm arkadaşlarıma teşekkür ederim.

Tez sürecimin en yoğun olduğu anlarda yanımda olan, her ihtiyaç duyduğumda yardıma koşan ve bana zaman ayıran değerli arkadaşım Arş. Gör. Esma GAVGALI'ya içtenlikle teşekkür ederim.

Hayatımın her döneminde desteğini yanımda hissettiğim kıymetli ablam Elif TURHAL UÇURUM'a ve hayatımıza güzellikler, tez çalışmama anlam katan Hira ve Mert Deha'ya gönülden teşekkür ederim.

Ve her şeyden öte, hayatımdaki en büyük değere sahip anneme ve babama, yaptıkları her fedakârlık, sundukları her imkân ve yanımda oldukları her an için sonsuz teşekkürlerimi sunarım.

Uğur TURHAL

Trabzon 2024

TEZ ETİK BEYANNAMESİ

Doktora Tezi olarak sunduğum "Yüz Biyometrik Sistemlerinde Sahteciliklerin Tespiti" başlıklı bu çalışmayı baştan sona kadar danışmanım Doç. Dr. Vasif NABİYEV'in sorumluluğunda tamamladığımı, verileri/örnekleri kendim topladığımı, deneyleri/analizleri ilgili laboratuvarlarda yaptığımı/yaptırdığımı, başka kaynaklardan aldığım bilgileri metinde ve kaynakçada eksiksiz olarak gösterdiğimi, çalışma sürecinde bilimsel araştırma ve etik kurallarına uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim. 03/01/2024

Uğur TURHAL

İÇİNDEKİLER

Sayfa No

ÖZET

Doktora Tezi

YÜZ BİYOMETRİK SİSTEMLERİNDE SAHTECİLİKLERİN TESPİTİ

Uğur TURHAL

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı
Danışman: Doç. Dr. Vasif NABİYEV
2024, 128 Sayfa

Bu tez çalışmasında, standart bir görüntüleme aygıtı kullanan yüz biyometrik doğrulama sisteminde, düşük hata oranı ile yüz sahteciliği tespiti yapan modeller önerilmektedir. Çalışma iki başlık altında değerlendirilmiştir. İlk aşamada, geleneksel doku tanımlama yöntemlerinden elde edilen öznitelikler, klasik öğrenme algoritmaları yardımıyla sınıflandırılarak sahtecilik tespiti gerçekleştirilmiştir. Bu aşamada ilk olarak, bir giriş görüntüsünde sahtecilik tespitine en çok etki eden yüz bölgesinin bulunması hedeflenmiştir. Ardından, renk uzaylarının probleme etkileri araştırılmış ve düşük öznitelik vektörü boyutuna sahip bir örüntü kodlayıcı önerilmiştir. İkinci aşamada ise derin öğrenme algoritmalarının yüz sahteciliği tespitine etkisi araştırılmıştır. Bu kısımda, literatürde sıklıkla kullanılan Xception ağının parametre sayısı azaltılarak, indirgenmiş bir versiyonu üretilmiştir. Oluşturulan bu modele dikkat blokları eklenerek, öznitelik kodlayıcılar üzerinde kanallardaki bilginin ağırlıklandırılması sağlanmıştır. Son olarak, klasik öznitelik çıkarma yöntemi ile derin öğrenme modellerini içeren karma girişli bir derin öğrenme ağı önerilmiş ve sistemlerin performansları değerlendirilmiştir. Elde edilen sonuçlar, renk bilgisinin sahtecilik tespitinde etkili olduğunu, önerilen öznitelik kodlayıcının güçlü tanımlama yeteneğinin yanında düşük hesaplama karmaşıklığına sahip olduğunu ve derin öğrenme modellerinin yüz sahteciliği tespitindeki başarımının, klasik yöntemlerden daha iyi olduğunu ortaya koymuştur.

Anahtar Kelimeler: Yüz Sahteciliği Tespiti, Yerel İkili Örüntüler, SVM, Derin Öğrenme, Xception Ağı

SUMMARY

PhD. Thesis

SPOOFINGS DETECTION IN FACIAL BIOMETRIC SYSTEMS

Uğur TURHAL

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Computer Engineering Graduate Program
Supervisor: Assoc. Prof. Dr. Vasif NABİYEV
2024, 128 Pages

In this thesis, we propose models for face spoofing detection with low error rate in a face biometric verification system using a standard imaging device. The study is evaluated under two headings. In the first stage, spoofing detection is performed by classifying the features obtained from traditional texture identification methods with the help of classical learning algorithms. In this stage, firstly, it is aimed to find the face region in an input image that has the most effect on spoofing detection. Then, the effects of color spaces on the problem were investigated and a pattern encoder with low feature vector size was proposed. In the second stage, the effect of deep learning algorithms on face spoofing detection is investigated. In this section, a reduced version of the Xception network, which is frequently used in the literature, is produced by reducing the number of parameters. Attention blocks were added to this model and the information in the channels was weighted on the feature coders. Finally, a hybrid deep learning model including classical feature extraction method and deep learning models is proposed and the performances of the systems are evaluated. The results show that color information is effective in spoofing detection, the proposed feature encoder has low computational complexity with strong identification capability, and the performance of deep learning models in face spoofing detection is better than classical methods.

Key Words: Face Spoofing Detection, Local Binary Patterns, SVM, Deep Learning, Xception Network

ŞEKİLLER DİZİNİ

Sayfa No

TABLOLAR DİZİNİ

Sayfa No

SEMBOLLER VE KISALTMALAR DİZİNİ

τ	Eşik Değeri
r	Kanal azaltma oranı
2B	2 Boyutlu
3B	3 Boyutlu
AUROC	Alıcı İşletim Karakteristiği Eğrisi Altındaki Alan (Area Under the Receiver Operating Characteristic)
BSIF	İkili İstatistiksel Görüntü Özellikleri (Binarized Statistical Image Features)
CCMF	Renk Kanalı Markov Özniteliği (Color Channel Markov Feature)
CCDMF	Renk Kanalı Farkı Markov Özniteliği (Color Channel Difference Markov Feature)
CCoLBP	Yerel İkili Örüntülerin Kromatik Birleşimi (Chromatic Co-Occurrence of Local Binary Pattern)
CLBP	Renkli Yerel İkili Örüntüler (Color Local Binary Patterns)
CNN	Evrişimli Sinir Ağları (Convolutional Neural Networks)
CoALBP	Komşu Yerel İkili Örüntülerin Birleşimi (Co-occurrence of the Adjacent Local Binary Patterns)
CTMF	Renk Dokusu Markov Özniteliği (Color Texture Markov Feature)
ÇB_YİÖ	Çok Bloklu Yerel İkili Örüntüler
ÇS_YİÖ	Çok Seviyeli Yerel İkili Örüntüler
ÇRÇS_YİÖ	Çok Renkli Çok Seviyeli Yerel İkili Örüntüler
DA	Alan Uyarlaması (Domain Adaptation)
DCT	Ayrık Kosinüs Dönüşümü (Discrete Cosine Transform)
DGP	Alan Kılavuzlu Budama (Domain Guided Pruning)
DHFD	Dinamik Yüksek Frekans Tanımlayıcısı (Dynamic High Frequency Descriptor)
DL	Derin Öğrenme (Deep Learning)
dLBP	Yön Kodlu Yerel İkili Örüntüler (Direction Coded Local Binary Patterns)
DMD	Dinamik Mod Ayrıştırma (Dynamic Mode Decomposition)
DoG	Gaussların Farkı (Difference of Gaussians)
DR	Ayrıştırılmış Temsil (Disentangled Representation)

DWT	Ayrık Dalgacık Dönüşümü (Discrete Wavelet Transform)
ED-LBP	Yerel İkili Örüntü Eşitlik Farkı (Equilibrium Difference Local Binary Pattern)
EER	Eşit Hata Oranı (Equal Error Rate)
FAR	Yanlış Kabul Oranı (False Acceptance Rate)
FC	Tam Bağlantılı (Fully Connected)
FCN	Tam Evrişimli Sinir Ağı (Fully Convolutional Network)
FCN-DA-LSA	Alan Uyarlamalı ve Kayıpsız Boyut Uyarlamalı Tam Evrişimli Ağ (Fully Convolutional Network with Domain Adaptation and Lossless Size Adaptation)
FN	Yanlış Negatifler (False Negatives)
FP	Yanlış Pozitifler (False Positives)
FPR	Yanlış Pozitif Oranın (False Positive Rate)
FRR	Yanlış Reddetme Oranı (False Rejection Rate)
G	Genişlik
GLCM	Gri Seviye Birliktelik Matrisi (Gray-Level Co-Occurrence Matrix)
GMM	Gauss Karışım Modeli (Gaussian Mixture Model
GS	Kılavuzlu Uzay
HoG	Yönlü Gradyanların Histogramı (Histogram of Oriented Gradients)
HOOF	Yönlendirilmiş Optik Akış Histogramı (Histogram of Oriented Optical Flow)
HTER	Yarı Toplam Hata Oranı (Half Total Error Rate)
IDA	Görüntü Bozulma Analizi (Image Distortion Analysis)
IQA	Görüntü Kalitesi Analizi (Image Quality Analysis)
IQM	Görüntü Kalitesi Ölçekleri (Image Quality Metrics)
K	Kanal sayısı
LaF	Laptop – Floresan
LaG	Laptop – Gün Işığı
LBPnet	Yerel İkili Örüntüler Ağı (Local Binary Pattern Network)
LBP-TOP	Üç Dikey Düzlemin Yerel İkili Örüntüleri (Local Binary Patterns of Three Orthogonal Planes)
LDA	Lineer Diskriminant Analizi (Linear Discriminant Analysis)
LDN-TOP	Üç Dikey Düzlemin Yerel Yönlü Sayı Dokusu (Local Directional Number Pattern Three Orthogonal Planes
LDP	Yerel Türev Örüntüsü (Local Derivative Pattern)
LFLBP	Işık Alanı Yerel İkili Örüntüler (Light Field Local Binary Patterns)

LGO	Öğrenilebilir Gradyan Operatörü (Learnable Gradient Operatör)
LPQ	Yerel Faz Niceleme (Local Phase Quantization)
LSTM	Uzun-Kısa Vadeli Bellek (Long Short-Term Memory)
MBSIF	Çok Ölçekli Dinamik İkili İstatistiksel Görüntü Öznitelikleri (Multi-Scale Dynamic Binarized Statistical Image Features)
MC-CNN	Çok Kanallı Evrişimsel Sinir Ağı (Multi-Channel Convolutional Neural Network)
MD	Çok Alanlı (Multi-Domain)
mLBP	Modifiye Yerel İkili Örüntüler (Modified Local Binary Patterns)
MLPQ	Çok Ölçekli Dinamik Yerel Faz Niceleme (Multi-Scale Dynamic Local Phase Quantization)
MSBP	Çok Ölçekli Yerel İkili Örüntü (Multi-Scale Local Binary Pattern)
MSRCR	Renk Restorasyonlu Çok Ölçekli Retineks (Multi-Scale Retinex with Color Restoration)
MTCNN	Çoklu Görev Basamaklı Evrişimsel Sinir Ağı (Multi-Task Cascade Convolutional Neural Network)
P	Görüntüde merkez piksel ile işleme sokulan komşu sayısı
PPV	Pozitif Tahmin Değeri (Positive Predictive Value)
RBF	Radyal Tabanlı Fonksiyon (Radial Basis Function)
ResNet	Artık Ağ (Residual Network)
R	Görüntüdeki merkez piksel komşularının merkez piksele olan uzaklığı
RFE	Özyinelemeli Öznitelik Eleme (Recursive Feature Elimination)
RI-LBP	Rotasyona Dirençli Yerel İkili Örüntüler (Rotation-Invariant Local Binary Pattern)
RNN	Yinelemeli Sinir Ağı (Recurrent Neural Network)
rPPG	Uzaktan Fotopletismografi (Remote Photoplethysmography)
S	Renk uzayı
S_C	Renk uzayı kanal sayısı
SAPLC	Piksel Düzeyinde Yerel Sınıflandırıcıların Uzamsal Birleştirilmesi (Spatial Aggregation of Pixel-level Local Classifiers)
SE	Sıkıştırma ve Uyarma (Squeeze and Excitation)
SGD	Stokastik Gradyan İnişi (Stochastic Gradient Descent)
SSR-FCN	Öz Denetimli Bölgesel Tam Evrişimli Ağ (Self-Supervised Regional Fully Convolutional Network)
SURF	Hızlandırılmış Güçlü Öznitelikler (Speeded-Up Robust Features)
SVM	Destek Vektör Makineleri (Support Vector Machines)
T	Alt bölge sayısı

TaF	Tablet – Floresan
TaG	Tablet – Gün Işığı
TBA	Temel Bileşenler Analizi (Principal Component Analysis)
TCoALBP	Zamansal Birleşimli Komşu Yerel İkili Örüntüler (Temporal Co-occurrence Adjacent Local Binary Patterns)
TeF	Telefon – Floresan
TeG	Telefon – Gün Işığı
TL	Öğrenme Aktarımı (Transfer Learning)
tLBP	Geçiş Kodlu Yerel İkili Örüntüler (Transition Coded Local Binary Patterns)
TP	Gerçek Pozitifler (True Positives)
TPR	Gerçek Pozitif Oranı (True Positive Rate)
TSS	Zamansal Dizi Örneklemesi (Temporal Sequence Sampling)
TSViT	İki Akışlı Vizyon Dönüştürücüler (Two-Stream Vision Transformers)
Vb	Vektör boyutu
X	Giriş katmanı
$\tilde{X}$	Çıkış katmanı
x_c	Görüdeki merkez piksel
x_p	Görüntüdeki merkez pikselin komşuları
Y	Yükseklik
YA_ÇRÇS_YİÖ	Yüz Ağırlıklı Çok Renkli Çok Seviyeli Yerel İkili Örüntüler
$YA\_ÇRÇS\_Y\text{İÖ}_{8,1}^{U2}$	8 komşu 1 yarıçap değerine sahip düzgün yüz ağırlıklı çok renkli çok seviyeli yerel ikili örüntüler
YİÖ	Yerel İkili Örüntüler
$Y\text{İÖ}_{8,1}$	8 komşu 1 yarıçap değerine sahip tüm yerel ikili örüntüler
$Y\text{İÖ}_{8,1}^{U2}$	8 komşu 1 yarıçap değerine sahip düzgün yerel ikili örüntüler
YST	Yüz Sahteciliği Tespiti
VLBP	Hacimsel Yerel İkili Örüntüler (Volume Local Binary Patterns)

1. GENEL BİLGİLER

1.1. Giriş

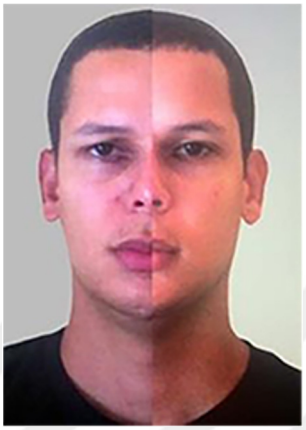
Son yirmi yılda, elektronik ve bilgisayar bilimlerinde teknolojinin ilerlemesi, dünya nüfusunun önemli bir kısmına uygun fiyatlarla üst düzey teknoloji cihazlarına erişim imkânı sağlamıştır. Çevrim içi ödeme ve e-ticaret güvenliği, akıllı telefon tabanlı kimlik doğrulama, güvenli erişim kontrolü ve biyometrik pasaport gibi çeşitli biyometrik sistemler, gerçek hayattaki uygulamalarda yaygın olarak kullanılmaktadır (Ming vd., 2020). Yüz tanıma, temel olarak diğer biyometrik özniteliklere kıyasla avantajları nedeniyle, 90'lardan beri en çok çalışılan biyometrik teknolojiler arasında yer almaktadır (Turk & Pentland, 1991).

Kolay kullanımları, mükemmele yakın performansları ve yüksek güvenlik seviyeleri sayesinde, yüz tanıma sistemleri, iris ve parmak okuyucu gibi diğer biyometrilere kıyasla piyasadaki en yaygın biyometrik sistemler arasındadır (De Luis-García vd., 2003). Temassız yapıları ve kullanıcı iş birliğine olan düşük bağımlılıkları nedeniyle biyometrik doğrulama gerektiren çeşitli alanlarda yaygın olarak kullanılan bu teknoloji üç farklı avantaj sunmaktadır: (1) uzaktan kimlik doğrulamaya olanak sağlar; (2) akıllı telefon veya tablet gibi herhangi bir evrensel ekipmanla çalışır ve herhangi bir ek donanım gerektirmez ve (3) parmak izi kimlik doğrulamasına kıyasla, doğrulama beklenmedik bir şekilde başarısız olursa, doğrulama cihazına ihtiyaç duyulmadan devam ettirilebilir (Imaoka vd., 2021). Ancak bu sistemler, kolay uygulanabilirliklerine rağmen saldırılara karşı savunmasızdır.

Yüz görüntüsü, bir yüz tanıma sisteminin kritik bileşenidir ve bu bilgi özeldir. Günümüzde artan sosyal medya kullanımı, saldırganların bu bilgilere kolay bir şekilde erişmesine ve bunları değiştirmesine olanak sağlamaktadır. Teknolojinin doğal evrimiyle birlikte, saldırı türleri de çeşitlilik göstermiştir. Bu alanda yapılan ilk çalışmalarda saldırılar, kullanıcı görüntülerinin fotoğraf kâğıtlarına basılmasıyla gerçekleştirilmişken daha sonra, kişinin yüz görüntüsünü tanıma sistemine sunmak için dijital cihazlar (dizüstü bilgisayarlar, tabletler, cep telefonları) kullanılmıştır. Yakın zamanlarda ise saldırılar, tüm yüz detaylarının plastik makyajla taklit edilmesine kadar ilerlemiştir.

Yüz tanımaya dayalı kontrol erişim sistemlerinin artan kullanımı, yüz sahteciliği saldırılarını tespit etmek için daha hassas sistemlere olan ihtiyacı arttırmıştır. Yüz sahteciliğinde saldırgan, geçerli bir kullanıcının kimliğini taklit ederek onaylanmaya çalışır.

İnsan yüzlerini kopyalama prosedürleri günümüzde oldukça basit olduğundan (örneğin, fotoğraf ve 3B baskı), sahtecilik tespiti, yüz tanıma sistemlerinde zorunlu hale gelmiştir. Şekil 1, yüz sahteciliği tespiti (YST) probleminin karmaşıklığını gösteren örnek bir görsel içermektedir (L. Souza vd., 2018). Görselde sol taraf gerçek, sağ taraf ise sahte yüz görüntüsüne aittir.



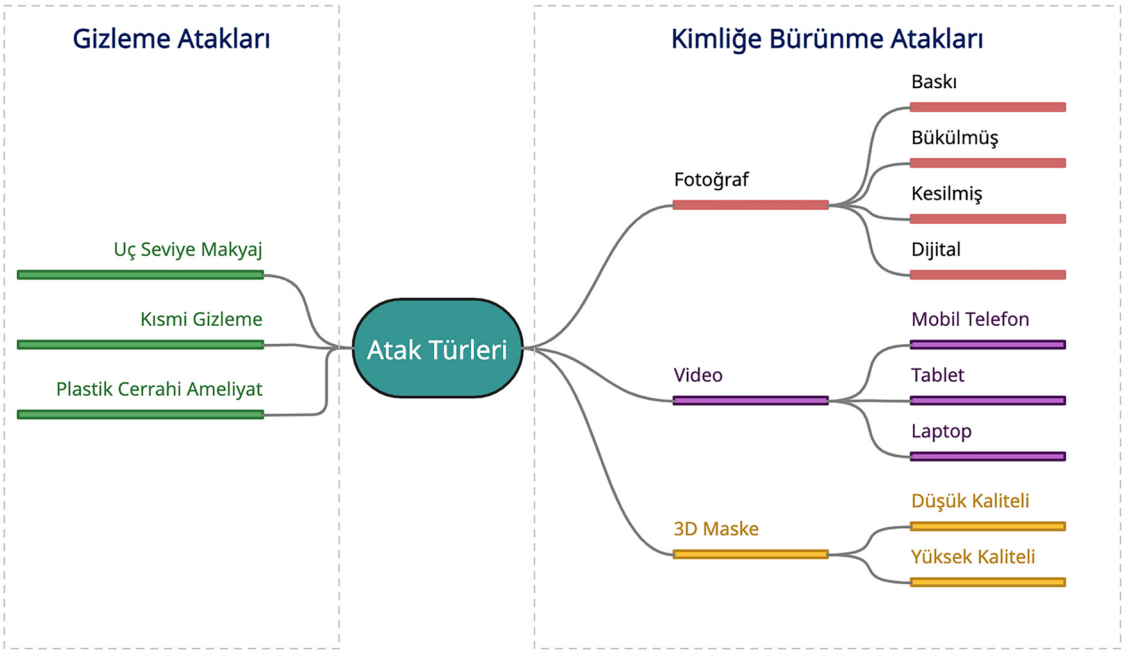
Şekil 1. Gerçek ve sahte yüz görüntülerinin karşılaştırmalı gösterimi

Günümüzde, bu tip sistemleri aldatmak için kullanılabilecek sahte maskelerin nasıl oluşturulacağına dair ayrıntılı bilgi veren öğretici videolar içeren web sitelerine erişmek kolaylaşmıştır (Galbally vd., 2014a). Bu durum, mevcut sistemlerin karşılaşabileceği saldırı türlerinde çeşitliliğin artmasına sebep olmaktadır.

Biyometrik sistemlere yapılan saldırılar yalnızca teorik veya akademik alanlarla sınırlı olmamakta, gerçek uygulamalara karşı da gerçekleştirilmektedir. Bu duruma en iyi örnek, Apple firmasının bütünleşik parmak izi okuyucusu donanımına sahip iPhone 5S model akıllı telefonlarının, satışa sunulduktan yalnızca bir gün sonra sahte parmak izi örnekleriyle aldatılmasıdır (Curtis, 2013). Son kullanıcıların özel bilgilerini içeren kişisel cihazlarına yapılan bu saldırı, hedefin yalnızca küresel kullanımdaki sistemler olmadığını göstermekte ve bu durum, biyometrik doğrulama sistemlerinde sahtecilik tespitinin önemini ortaya koymaktadır.

Yüz sahteciliği problemi temelde kimliğe bürünme atakları ve gizleme atakları olmak özere iki saldırı türü üzerinden değerlendirilebilir. Kimliğe bürünme saldırılarında erişim yetkisine sahip olmayan kişi, yetki sahibi bireyin yüz görüntüsünden ya da videosundan faydalanarak sisteme sızmaya çalışır. Günümüzde artan sosyal medya kullanımı, kişilerin bu platformlara yüz görüntülerini içeren görselleri ve videoları daha sıklıkla yüklemesi, kötü

niyetli kullanıcıların bu verilere erişimini kolaylaştırmakta ve bu durum, yüz kimlik doğrulama sistemlerini kandırmak için ilgili görsellerin ya da videoların kullanılmasına olanak sağlamasına neden olmaktadır. Gizleme ataklarında ise saldırgan, sistem tarafından tanınmaktan kaçınmak için çeşitli hilelere başvurmakta ve giriş görüntüsünün çeşitli bölgelerini manipüle ederek, tespit sistemini hataya zorlamaktadır. Yapılan çalışmalar, genellikle kimliğe bürünme atakları üzerinde yoğunlaşmıştır. Şekil 2, yüz sahteciliği probleminde kullanılan yaygın atak türlerini ve bu türlerin içerdiği yöntemleri göstermektedir.



Şekil 2. Yüz sahteciliğinde kullanılan yaygın atak türleri

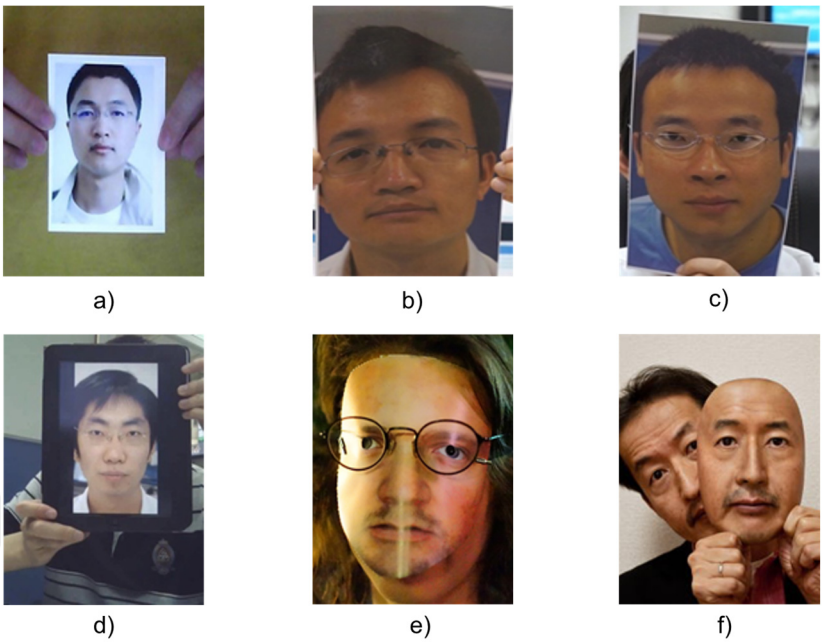
Yüz sahteciliğinde saldırganlar genellikle kimliğe bürünme ataklarını tercih etmektedir. Bu atak türü; fotoğraf saldırıları, video yeniden oynatma saldırıları ve 3B maske saldırıları olmak üzere üç başlıkta kategorize edilebilir. Bu üç ana kategori kendi içlerinde çözünürlük, aydınlatma, giriş cihazı ve konum gibi alt gruplara ayrılır. Bu nedenle, YST sistemi gerçek uygulamalarda kullanıldığında, ortaya çıkacak senaryo türünü kesin olarak tahmin etmek zordur.

Gizleme atakları, genellikle kullanıcının gerçek kimliğini gizlemek için yüz makyajı, plastik cerrahi veya yüzdeki bölgelerin saklanması gibi hilelere dayanır ve diğer yöntemlere kıyasla daha zahmetlidir. Fotoğraf saldırıları (literatürde baskı saldırıları olarak da

adlandırılır) ve video yeniden oynatma saldırıları, internette sürekli artan yüz görüntü akışı ve düşük maliyetli yüksek çözünürlüklü dijital cihazların yaygınlığı nedeniyle en çok tercih edilen saldırı türleridir. Bu yöntemi kullanan saldırganlar, gerçek kullanıcıların yüz örneklerini kolayca toplayabilir ve yeniden kullanabilir.

Fotoğraf saldırıları, yüz kimlik doğrulama sistemine gerçek bir kullanıcının resmi gösterilerek gerçekleştirilir. Saldırganlar tarafından genellikle birkaç strateji kullanılır: Basılı fotoğraf saldırıları, bir kâğıda basılmış bir resmi (örneğin, A3 / A4 kâğıdı, bakır kâğıt veya profesyonel fotoğraf kâğıdı) sunmayı içerir. Şekil 3(a), NUAA veri setinden alınmış basılı fotoğraf saldırısı örneğini içermektedir. Öte yandan fotoğraf görüntüleme saldırılarında görsel; akıllı telefon, tablet veya laptop gibi dijital bir cihazın ekranında görüntülenir ve ardından sisteme sunulur. Bir başka saldırı türünde fotoğrafa derinlik kazandırmak için basılı fotoğraflar yatay ve dikey eksen boyunca bükülebilir. Bu stratejiye bükülmüş fotoğraf saldırısı denir. Şekil 3(b), CASIA-FASD veri setinden alınmış bükülmüş (warped photo) yüksek kaliteli maske fotoğraf atağı saldırısı örneğini içermektedir. Canlılık tespitine yönelik inceleme yapan YST modellerinde izinsiz girişi sağlamak adına, fotoğrafın arkasındaki bireyin yüzünden göz kırpma veya ağız hareketi gibi bazı canlılık ipuçları vermek için ağız, göz ve burun bölgelerinin kesildiği fotoğraf maskelerini içeren atak türleri bulunmaktadır. Şekil 3(c), CASIA-FASD veri setinden alınmış, göz bölgeleri kesik yüksek kaliteli maske fotoğraf saldırısı örneğini içermektedir. Statik fotoğraf saldırıları ile karşılaştırıldığında video oynatma saldırıları, canlılığı taklit etmek için göz kırpma, ağız hareketleri ve yüz ifadelerindeki değişiklikler gibi içsel dinamik bilgileri sunmaları sebebiyle daha karmaşıktır (Sebastien vd., 2014). Şekil 3(d), CASIA-FASD veri setinden alınmış, normal kaliteli video oynatma saldırısı örneğini içermektedir. Bükülmüş fotoğraf saldırıları hariç, fotoğraf veya video oynatma saldırılarının en büyük dezavantajları, iki boyutlu düzlemsel saldırılar olmalarıdır. Bu durumun aksine 3B maske saldırıları, yüz yapaylıklarını yeniden oluşturur. Düşük kaliteli 3B maskeler (Şekil 3(e)) ve yüksek kaliteli 3B maskeler (Şekil 3(f)) arasında ayrım yapılabilir. "Yüz benzeri" 3B yapının yüksek gerçekçiliği ve insan cildini canlı bir şekilde taklit edebilme yeteneği, bu saldırı türlerinin, geleneksel yöntemlerle tespit edilmesini daha zor hale getirir (Tirunagari vd., 2015). Günümüzde, yüksek kaliteli bir 3B maske üretmek hala pahalı (Galbally & Satta, 2016) ve karmaşıktır. Ayrıca bu yöntem, genellikle kullanıcı iş birliğini gerektirir (Erdogmus & Marcel, 2014). Bu nedenle, 3B maske saldırıları, fotoğraf veya video yeniden oynatma saldırılarından çok daha az sıklıkta gerçekleştirilir. Bununla birlikte, 3B edinim sensörlerinin

yaygınlaşmasıyla, önümüzdeki yıllarda 3B maske saldırılarının giderek daha sık hale gelmesi beklenmektedir.



Şekil 3. Yaygın yüz sahteciliği türlerine örnekler

Özel donanım gerektirmeyen ve standart görüntüleme cihazları kullanarak YST gerçekleştiren sistemler, ortam ışığı, görüntü kalitesi, gürültü, yüz görüntülerindeki tutarsızlık (gözleri kısma veya kapama, yüz ifadesini değiştirme) ve yaşlanmaya bağlı yüz değişiklikleri gibi çok sayıda değişkenden etkilenmektedir. Bu nedenle problemin çözümü zorlaşmaktadır.

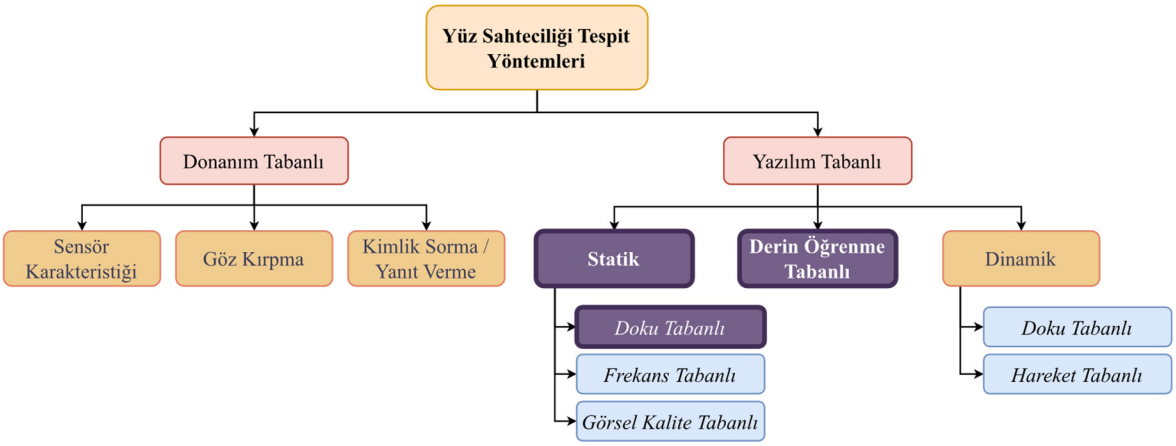
Bu tez çalışmasında, standart görüntüleme cihazlarından elde edilen görüntüler üzerinde YST başarımını arttırmaya yönelik çalışmalar yapılmış ve farklı uygulama alanlarında kullanılabilecek modeller önerilmiştir.

1.2. Tezin Kapsamı ve Katkısı

Literatüre genel bir perspektiften bakıldığında, YST yöntemleri genel olarak donanım ve yazılım tabanlı olmak üzere iki ana başlık altında kategorize edilmektedir. Kullanım amacına göre bu yöntemler alt gruplara ayrılmaktadır.

Donanım tabanlı yaklaşımlarda YST sistemlerine entegre edilen donanım cihazları yardımıyla sahtecilik tespiti yapılırken, yazılım tabanlı yaklaşımlarda, giriş aygıtlarından elde edilen yüz görüntülerinden çıkarılan öznitelikler yardımıyla sahtecilik tespiti

gerçekleştirilir. Bu sebeple, donanım tabanlı yaklaşımlar, yazılım tabanlı yaklaşımlara göre daha maliyetlidir ve daha az tercih edilmektedirler. Şekil 4, iki ana grup altında toplanan YST yöntemlerini ifade etmektedir (Sharma & Selwal, 2023).



Şekil 4. Yüz saldırı tespit yöntemleri

Bu tez kapsamında, yazılım tabanlı yüz sahteciliği tespit yöntemlerinden statik doku tabanlı yöntemler ile derin öğrenme tabanlı yöntemler üzerinde çalışılmıştır.

Tez çalışmasının katkıları aşağıda verilen maddelerle özetlenebilir;

- Bir yüz görüntüsünden elde edilen ağız, burun ve göz bölgelerinin, YST başarımına etkileri incelenmiştir.
- Düzgün Yerel İkili Örüntüler (YİÖ) özniteliklerinin dışında, tüm YİÖ özniteliklerinin YST başarımına etkileri incelenmiştir.
- Veri seti sağlayıcıları tarafından sunulan tüm senaryolarda deneyler gerçekleştirilmiş ve önerilen modellerin, çeşitli senaryolar ve tüm senaryoların bir arada bulunduğu durumlardaki başarımları incelenmiştir.
- Renk kanallarından üretilen statik doku özniteliklerinin YST başarımına etkisi incelenmiş ve probleme etki düzeyleri göz önüne alınarak, bu renk kanallarının bir arada kullanılmasıyla elde edilecek öznitelik kümelerinin problemin çözümüne etkileri araştırılmıştır.
- YST probleminin çözümü için statik doku çıkarma tekniğine dayalı Yüz Ağırlıklı Çok Renkli Çok Seviyeli Yerel İkili Örüntüler (YA_ÇRÇS_YİÖ) yöntemi önerilmiştir.

- YST probleminin çözümü için derin öğrenme tabanlı, düşük parametre sayısı ve dikkat mekanizması içeren derin öğrenme ağı tasarlanmıştır.
- YA_ÇRÇS_YİÖ ile üretilen özniteliklerin, önerilen derin öğrenme ağı ile birleşimi sonucu oluşturulan karma girişli derin öğrenme ağ modeli tasarlanmıştır.

1.3. Literatür Araştırması

1.3.1. Donanım Tabanlı Yüz Sahteciliği Tespiti

Bu yaklaşımlar, yüz tanıma tabanlı sistemlerin sensörüne entegre edilmiş ek bir donanım cihazı kullanarak, insan yüzünün; kan basıncı, iletkenlik, ter, yüz termogramı vb. gibi temel özelliklerini ölçer (Dubey vd., 2016). Donanım cihazı tarafından yakalanan sinyaller, gerçek yüz görüntülerini sahtelerden yeterli doğrulukla ayırt etmek için kullanılır (Galbally vd., 2014b). Bu yöntemler, ek bir donanım cihazına ihtiyaç duyulduğundan dolayı pahalıdır (Bagga & Singh, 2016). Bu teknikleri üç başlık altında incelemek mümkündür: a) Sensör özellikleri, b) Göz kırpma tespiti ve c) Kimlik sorma ve yanıt.

A. Sensör Karakteristikleri

Bu yöntemler, sensör özelliklerine dayanmakta ve özelliklerini araştırmaktadır. Gözlemlenen özellikler, verilerin elde edilmesi için kullanılan görüntü yakalama cihazının tasarımına bağlıdır. Benzer şekilde, 3D taramanın yansıma ölçümleri, sensörün özelliklerini incelemek için iyi bir örnektir.

Tan ve ark. (Tan vd., 2010), gerçek ve sahte yüzler arasında ayrım yapmak için Lambertian yansıma modelini kullanmıştır. Kim ve ark. (S. Kim vd., 2013), yapay ve gerçek yüz örneklerini ayırt etmek için değişken odaklama kavramını (kamera/sensör özelliklerinden biri) kullanmıştır. Belirli bir odak içindeki en yakın ve en uzak nesneler arasındaki aralık, alan derinliği derecesine dayanır. Her görüntüde hem bulanık hem de odaklanmış alanlar bulunurken, kullanıcının seçimi hangi alanın odaklandığına karar verir. Ek olarak, her lensin farklı bir odak uzaklığı vardır. Bu özellik, her örnekten iki ardışık görüntü alındığında, gerçek ve sahte görüntünün odak değerlerinin farklı olduğunu göstermiştir. Gerçek yüz fotoğraflarında odaklanılan bölgeler net ve geri kalanlar derin bilgi nedeniyle bulanıktır. Sahte fotoğraflarda ise bulanıklık oluşmamaktadır. Önerilen yöntem, alan derinliği derecesi daha küçük olduğunda daha iyi sonuçlar alındığını ortaya koymuştur.

Yang (L. Yang, 2014), bir önceki çalışmada önerilen modelin değiştirilmiş bir versiyonunu sunmuştur. Bu çalışmada, burun ve kulağın yerine iki ardışık fotoğrafın farkını artırmak için yüz ve arka planın bulanıklık derecesi dikkate alınmıştır. Bir nesneyi gerçek olarak sınıflandırmak için, burun ve arka plana odaklanan iki ardışık görüntü alınır ve analiz edilir. Bu iki görüntünün bulanıklık derecesi aynı değilse, nesne gerçek olarak sınıflandırılır. Bununla birlikte, iki fotoğrafın benzerliği, nesneyi sahte olarak tanımlar. Kim ve ark. (S. Kim vd., 2015) çalışmalarında odak eksikliğinin etkisini araştırmıştır. Sahtecilik tespiti için üç farklı öznitelik çıkarılmıştır. Bunlar; i) odak, ii) güç histogramı ve gradyan konumu ve iii) yönelimdir. Li ve ark. (X. Li vd., 2016), insanların canlılık özniteliklerini belirlemek için yapay yüz maskelerinde veya basılı fotoğraflarda bulunmayan nabız sinyallerinin tespiti üzerine çalışmalar yürütmüştür. Canlılık tespitine dayalı yöntemler, videolardaki yüz ifadeleri ve göz kırpmaları gibi fiziksel yaşam belirtilerini bulmaya çalışmaktadır. Bununla birlikte, bu yöntemler çok sayıda kullanıcıdan oluşan ekip çalışmasına, ek cihazlara ve video dizilerine ihtiyaç duyar. Ayrıca zaman alıcı ve hesaplama açısından karmaşıktırlar. Raghavendra ve ark. (Raghavendra vd., 2015), ışık alanı kamerasını bir yakalama cihazı olarak kullanarak tanıma sistemine yönelik sahte saldırıları algılayabilen ve bunlara karşı önlem alabilen yeni bir teknik tasarlamıştır. Önerilen yaklaşımda, sunum saldırılarını tespit etmek için keşfedilen bir dizi derinlik görüntüsünden görüntü derinliği (odak) değişimini hesaplamak için üç yöntem tanıtılmıştır: Derinlikteki mutlak değişim, maksimum ve minimum derinliğin oranı, bu değerlerin uç uca birleşimi. Gerçek ve sahte yüz görüntülerini ayırt etmek için bir Destek Vektör Makineleri (Support Vector Machines- SVM) sınıflandırıcısı kullanılmıştır. Önerilen teknik, çeşitli cinsiyet, yaş ve etnik kökene sahip 80 denekten oluşan kendi oluşturdukları bir veri tabanı üzerinde test edilmiştir. Moghaddam ve ark. (Sepas-Moghaddam vd., 2017), ışık alanı görüntülerini analiz ederek, öznitelikleri belirlemek için Işık Alanı Yerel İkili Örüntüler (Light Field Local Binary Patterns- LFLBP) tanımlayıcısını kullanan bir yaklaşım önermiştir. LFLBP, uzamsal verilerin elde edilmesinin dışında, veri seti görüntüleriyle bağlantılı ışık alanı açısal bilgilerin çıkarılması için kullanılmış ve bu verilerle birlikte yüz sahteciliği tespiti yapılmıştır. Tang ve ark. (Tang vd., 2018), rastgele parlayan görüntülerin yansıyan ışık için analiz edildiği bir teknik geliştirmiştir. Bu teknik, dokusal öznitelikleri, üç boyutlu şekilleri ve farklı ışık hızlarında yansımanın işlenmesi gibi çeşitli insan yaşam özelliklerini incelemektedir. Dijital kameraların iş birliği ile önerilen teknik, bir saldırı prosedürünün bıraktığı izleri tespit edebilmektedir.

B. Göz Kırpma Tespiti

Bu yöntemler, kişinin kendiliğinden göz kırpmasını sürekli olarak takip eder. Gözlerin temel işlevi olan göz kapaklarının kapanması veya açılması, göz kırpma olarak bilinen fizyolojik bir süreçtir. Yüz tanıma sistemlerinde göz kırpmanın bu özelliği, sunum saldırılarını önlemek için kullanılır.

Bhaskar ve ark. (Bhaskar vd., 2003) tarafından optik akış hesaplaması ile birleştirilmiş çerçeve farklılaştırma kullanılarak göz kırpma tespiti için bir yöntem kullanılmıştır. Göz kırpmayı diğer yüz hareketlerinden ayırt etmek için, akış vektörlerinin hem büyüklüğünü hem de yönünü kullanan optik akış hesaplanır. Bu yöntemin, algoritmanın karmaşıklığı açısından yüz canlılığı tespiti için maliyetli olduğu kanıtlanmıştır. Pan ve ark. (G. Pan vd., 2007), genel bir web kamerası kullanarak, fotoğrafik yüzü bulmak için bir yöntem sunmuştur. Bu yöntem, ek donanım gerektirmez. Koşullu rastgele alan adını verdikleri çerçeve, göz kırpma tespiti için kullanılmıştır. Bir yüzü gerçek veya sahte olarak ayırt etmek için iki irisin merkezi arasındaki mesafe kullanılmıştır. Önerilen yöntemin, Adaboost ve gizli markov modelinden daha iyi sonuçlar ürettiği belirtilmiştir. Nema (Nema, 2020), yüz canlılığı tespiti için göz kırpma sayısı ve Yönlü Gradyanların Histogramı (Histogram of Oriented Gradients- HoG) öznitelik tanımlayıcısını incelemiş ve ORL ve CASIA-FASD olmak üzere halka açık iki veri kümesi için %96 sınıflandırma doğruluğu elde etmiştir.

C. Kimlik Sorma ve Yanıt

Bu yöntemler, herhangi bir aydınlatma olayından sonra göz bebeğinin kasılması veya önceden belirlenmiş harici uyaranlara doğru bakışın izlenmesi gibi, dış uyaranlara verilen istemsiz veya gönüllü tepkileri tespit etmek için kullanıcının yardımına ihtiyaç duyarlar.

Kollreider ve ark. (Kollreider vd., 2008), video saldırılarını tespit etmek için faydalı olan meydan okuma-yanıt tabanlı bir yüz canlılığı tespit tekniği önermiştir. Ali ve ark. (Ali vd., 2013), görsel uyarıcıya dayalı bir teknik önermiştir. Fotoğraf sahteciliği saldırılarının olup olmadığını belirlemek için bir kişinin bakış açısını değerlendirir. Daha sonra doğrusal öznitelikler kullanılarak gerçek ve sahte girişimler ayırt edilir. Lagorio ve ark. (Lagorio vd., 2013), yüz biyometrik sistemler için 3B yüz yapısı tabanlı bir yaklaşım sunmuştur. Çıkarılan verilerin 3B eğriliği, canlı ve sahte yüzleri ayırt etmek için işlenir. Önerilen teknik, yüzey eğriliğinin birinci dereceden istatistik tahminine dayanmaktadır ve herhangi bir kullanıcı iş birliği gerektirmemektedir. Smith ve ark. (Smith vd., 2015), ekrandaki görüntülerin meydan okuma oluşturmak için kullanıldığı ve bu nedenle dinamik olarak yakalanan yansımaların yanıt verdiği bir yaklaşım tasarlamıştır. Görüntülerin bir dizisini ve bunlara eşlik eden

yansımaları videoyu filigranlar. Yansımanın ekranda gösterilen görüntü setiyle uyumlu olup olmadığını belirlemek için yansıma bölgesi öznitelikleri kullanılır.

1.3.2. Yazılım Tabanlı Yüz Sahteciliği Tespiti

Yazılım tabanlı teknikler, düşük veya yüksek çözünürlüklü kameralar tarafından yakalanan yüz görüntülerinden çıkarılan öznitelikler yardımıyla sahte yüz görüntülerini gerçek yüz görüntülerinden ayırt etmeye çalışır. Bu yöntemler, sensör seviyesindeki yöntemlerden farklıdır. Sensör seviyesindeki teknikler, canlı bir bireyden elde edilen öznitelikleri kullanarak gerçek ve yapay yüzleri ayırt eder. Öznitelik çıkarma yöntemlerinde ise bir yüz görüntüsünden elde edilen öznitelikler işlenerek, gerçek veya yapay yüzler arasındaki farklılıkları bulmak için kullanılır (Galbally vd., 2014b). Bu teknikleri üç başlık altında incelemek mümkündür: a) Statik Yaklaşımlar, b) Dinamik Yaklaşımlar ve c) Derin Öğrenme Tabanlı Yaklaşımlar.

A. Statik Yaklaşımlar

Bu yaklaşımlar, yüz saldırılarını tespit etmek için zamansal veri gerektirmeden yalnızca belirli bir yüz görüntüsünün bir kopyasını kullanır. Bu yöntemler, her bir karenin ayrı ayrı analiz edildiği ve çoğunluk oylama yöntemi kullanılarak nihai kararın alındığı video dizileri için de kullanılabilir. Dinamik yöntemlerle karşılaştırıldığında, statik yaklaşımlar daha hızlıdır ve düşük hesaplama maliyeti ile daha iyi performans gösterirler (Ramachandra & Busch, 2017). Bu teknikler üç alt başlık altında gruplanabilir: a) Doku Tabanlı Yaklaşımlar, b) Frekans Tabanlı Yaklaşımlar ve c) Görüntü Kalitesi Tabanlı Yaklaşımlar.

1) Doku Tabanlı Yaklaşımlar

Doku tabanlı yaklaşımlar, bir yüz görüntüsünün doku öznitelikleri üzerinden gerçek ve sahte ayrımı yapmak için kullanılırlar. Bu yöntemler, mevcut bir yüz görüntüsü örneğinden mikro dokusal desenleri inceler. Baskı hatalarından kaynaklanan pigmentler veya ekran saldırılarından kaynaklanan gölge gibi yapay öznitelikleri kolayca ayırt edebilirler (Ramachandra & Busch, 2017). YST literatüründe, YİÖ güçlü ve hesaplama açısından verimli bir görüntü tanımlayıcısı olduğundan, çalışmaların çoğunda öznitelik çıkarımı için bu operatör kullanılmıştır. Ojala ve ark. (Ojala, Pietikäinen, vd., 2002) tarafından tanıtılan orijinal YİÖ tanımlayıcısı, piksellerin yerel komşuluklarıyla olan ilişkisinden türetilen gri ölçekli bir doku ölçüsüdür. Hem hesaplama hem de gri seviye değişikliklerine karşı daha fazla dayanıklıdır. Her pikselin komşuluğunu merkezi değerle

eşikleyerek görüntüdeki piksel etiketlerini keşfeden operatör, ikili sayıları dikkate alarak piksel etiketlerini belirler.

Määttä ve ark. (Määttä vd., 2011), YST probleminde yüz dokusunu analiz etmek için YİÖ özniteliklerini kullanmıştır. Gerçek ve sahte yüzler arasındaki farkları yakalamak için çok ölçekli YİÖ operatörleri ile çıkarılan öznitelikler SVM ile sınıflandırılmıştır. Başka bir çalışmada (Määttä vd., 2012), hem doku tabanlı (YİÖ, Gabor) hem de gradyan tabanlı (HoG) yüz tanımlayıcılarının YST performansları incelenmiştir. Waris ve ark. (Waris vd., 2013), yüz görüntülerindeki çeşitli dokusal öznitelikleri göz önünde bulundurarak fotoğraf ve video oynatma saldırılarına karşı yeni bir yaklaşım önermiştir. Görüntülerden elde edilen YİÖ, Gabor ve Gri Seviye Birliktelik Matrisi (Gray-Level Co-Occurrence Matrix- GLCM) öznitelikleri yardımıyla eğitilen SVM modeli ile önerilen yöntemin YST başarımını incelemiştir. Yang ve ark. (J. Yang vd., 2013), bileşen tabanlı yüz kodlaması kullanmıştır. Önerilen yaklaşımda, yüz bölgesi içeren giriş görüntüsü; göz, burun, yüz bölgesi, kanonik yüz bölgesi, ağız vb. gibi birden fazla bileşene ayırmaya odaklanmıştır. Elde edilen bu bölgeler arasındaki farkı keşfetmek için Fisher analizi yapılmıştır. Tüm bu bileşenlerden öznitelikler çıkarılmış ve daha sonra bu alt düzey özniteliklerden üst düzey öznitelikleri temsil etmek için vektör nicelemeye dayalı bir kodlama şeması uygulanmıştır. Raghavendra ve Bush (Raghavendra & Busch, 2014), bir görüntünün global ve yerel özniteliklerini araştırmış ve ağırlıklı puan seviyesi füzyonu ile göz ve yüz bölgelerinden hem İkili İstatistiksel Görüntü Öznitelikleri (Binarized Statistical Image Features- BSIF) (Kannala & Rahtu, 2012) hem de YİÖ özniteliklerini çıkarmak için bir algoritma önermiştir. BSIF tanımlayıcısı YST literatüründe de yaygın olarak kullanılmaktadır. Erdoğmuş ve Marcel (Erdogmus & Marcel, 2014), 3B maske saldırılarıyla mücadele etmek için YİÖ dokusal özniteliğinin Geçiş Kodlu YİÖ (transition coded LBP- tLBP), Modifiye YİÖ (modified LBP- mLBP), Yön Kodlu YİÖ (direction coded LBP- dLBP) gibi çeşitlerini kullanmıştır. Gerçek veya maskeli sınıflandırma için LDA ve SVM gibi sınıflandırıcıları eğitmişlerdir. Blok tabanlı ve çok ölçekli LBP özniteliklerinin, dLBP dışında iki boyutlu görüntüler için daha iyi performans göstermesi, çalışmanın ana sonuçlarındandır. Boulkenafet ve ark. (Boulkenafet vd., 2016b), renk dokusu analizini kullanarak bir yüz sahteciliği önleme yöntemi önermiştir. HSV ve YC_bC_r gibi farklı renk uzaylarının parlaklık ve renklilik kanallarından doku özniteliklerini hesaplamak için farklı doku tanımlayıcıları (YİÖ, Yerel Faz Niceleme (Local Phase Quantization- LPQ), ve BSIF) kullanılmıştır. Patel ve ark. (Patel vd., 2016), mobil cihazlarda çeşitli renk uzaylarının yüz sahteciliği tespiti üzerindeki

etkilerini araştırmışlardır. HSV renk uzayının tüm kanallarından, asimetri, ortalama ve standart sapma gibi bilgiler üretilmiş ve yüz dokusu, YİÖ kullanılarak elde edilmiştir. Ek olarak, RGB kanallarının yüz sahteciliği tespitine etkisi araştırılmıştır. Sonuçlara göre, kırmızı kanal diğer kanallara göre daha etkili olmuştur. Peng ve ark. (F. Peng vd., 2018b), Gauss ölçek uzayının kenar koruma filtreleme yeteneğini geliştirmek için bir Kılavuzlu Uzay (Guided Scale- GS) önermiştir. Bu uzayda, yüzün kenarları korunurken gürültü bileşenleri yumuşatılır. Çalışmada giriş görüntüsü bloklara ayrıştırılmış ve bu blokların kılavuzlu ölçek uzayından GS-YİÖ histogramları elde edilmiştir. En iyi sonuçların HSV + YC_bC_r renk uzayında elde edildiği belirtilmiştir. Zhang ve ark. (L.-B. Zhang vd., 2018), Renk Dokusu Markov Özniteliği (Color Texture Markov Feature- CTMF) öznitelik tanımlayıcısı olarak adlandırılan komşu piksel tutarsızlığını kullanmıştır. Araştırmacılar, komşu pikselleri göz önünde bulundurarak gerçek ve sahte öznitelikler arasında ayrım yapmak için renk dokusunun içsel özniteliklerini araştırmıştır. İlk olarak, yüz görüntüsü alanı tespit edilip normalleştirilmiş, ardından iki önemli öznitelik "Renk Kanalı Markov Özniteliği" (Color Channel Markov Feature- CCMF) ve "Renk Kanalı Farkı Markov Özniteliği" (Color Channel Difference Markov Feature-CCDMF) çıkarılmıştır. Destek Vektör Makineleri tabanlı Özyinelemeli Öznitelik Eleme (Support Vector Machines Recursive Feature Elimination- SVM-RFE) olarak bilinen bir öznitelik seçim tekniği kullanılarak boyutu azaltılan öznitelik kümesi SVM sınıflandırıcısının eğitimi için kullanılmıştır. Peng ve ark. (F. Peng vd., 2018a), kromatik yüz dokusu farklılıklarını ölçmüş ve kanallar arası yüz dokusunu incelemek için yeni bir Yerel İkili Örüntülerin Kromatik Birleşimi (Chromatic Co-Occurrence of Local Binary Pattern - CCoLBP) tanımlayıcısı tasarlamıştır. Hasan ve ark. (Hasan vd., 2019), yüz canlılığı tespiti için bir öznitelik çıkarıcı olarak YİÖ varyansı (YİÖV) ile bir ön işleme adımı olarak Gaussların Farkı (Difference of Gaussians- DoG) filtrelemesini kullanmıştır. DoG filtreleme, yüz görüntüsünden gürültüyü gidermek için uygulanmıştır. YİÖV'yi çıkarmak için kontrast öznitelikleri kullanılmış ve ardından düzenli YİÖ operatörü uygulanmıştır. Sthevanie ve Ramadhani (Sthevanie & Ramadhani, 2018), YST probleminin çözümünde YİÖ ve GLCM özniteliklerini bir arada kullanmıştır. 4 farklı test senaryosunun kullanıldığı çalışmada en iyi sonuçlar, YİÖ ve GLCM matrislerinin göz ve burun bölgelerine uygulanmasıyla elde edilmiştir. Khurshid ve ark. (Khurshid vd., 2019), dokusal öznitelikler üzerinden gerçek zamanlı YST yapan bir sistem geliştirmiştir. İlk olarak kameradan elde edilen RGB görüntüler gri seviyeye ve YC_bC_r renk uzayına dönüştürülmüştür. Ardından bu görüntülerden YİÖ öznitelikleri ve yalnızca gri seviye görüntüden Komşu Yerel İkili

Örüntülerin Birleşimi (Co-occurrence of the Adjacent Local Binary Patterns- CoALBP) öznitelikleri üretilmiştir. Son olarak elde edilen öznitelik kümesi SVM ile sınıflandırılmıştır. Shu ve ark. (Shu vd., 2021), kromatik Yerel İkili Örüntü Eşitlik Farkı (Equilibrium Difference Local Binary Pattern- ED-LBP) doku özniteliğine dayalı YST yöntemi önermiştir. Çalışmada, bir yüz görüntüsündeki komşu piksel uyuşmazlığı dikkate alınmış ve elde edilen bilgiler YİÖ ile kodlanmıştır. Farklı renk kanallarındaki öznitelik histogramları her bir görüntü bandı üzerinde ayrı ayrı hesaplanmıştır. Daha sonra, kromatik ED-LBP histogramları ve iki seviyeli uzamsal piramidin yardımıyla giriş görüntüsünde bulunan yüz bölgesinin yerel yapı bilgisi çıkarılmıştır. Son olarak, farklı renk alanlarından gelen ED-LBP histogramları, SVM ile sınıflandırılmıştır. Wang ve ark. (C. Wang vd., 2023), ham piksellerden detaylı (uzamsal gradyan büyüklüğü gibi) ayırt edici ipuçlarını verimli bir şekilde çıkarmak için halihazırda kullanılan gradyan operatörlerinin bir genellemesi olan Öğrenilebilir Gradyan Operatörünü (Learnable Gradient Operatör- LGO) oluşturmuştur. Optimizasyonu iyileştirmek için eş zamanlı olarak uyarlanabilir bir gradyan kaybı sunmuşlardır. Yaygın olarak kullanılan REPLAY-ATTACK, CASIA-FASD, OULU-NPU ve SiW veri setleri üzerinde elde edilen sonuçlar, önerilen tekniğin YST probleminde oldukça başarılı olduğunu göstermektedir.

2) Frekans Tabanlı Yaklaşımlar

Frekans seviyesi özniteliklerini yüz fotoğraflarından çıkarmak için yüksek ve düşük frekanslı sinyaller ve bunların varyasyonları gibi görüntülerin belirli öznitelikleri kullanılır. Bu kısımda, Frekans tabanlı YST tekniklerini kullanan bazı çalışmalar açıklanmıştır.

Daha önce, Li ve ark. (J. Li vd., 2004), 2D Fourier Spektrumu tabanlı bir YST tekniği tanıtmıştır. Önerilen yöntem iki olguya dayanmaktadır; birincisi, canlı yüzün boyutunun fotoğraf baskısına kıyasla farklı olması, diğeri ise canlı ve sahte yüzler arasındaki poz ve ifade farklılıklarını içermesidir. Teja (Teja, 2011), fotoğraf saldırılarına karşı koymak için göz bebeği ve göz kırpma algılama yardımıyla bir görüntünün Ayrık Kosinüs Dönüşümü (Discrete Cosine Transform- DCT) enerjisinin araştırıldığı bir yöntem tasarlamıştır. Zhang ve ark. (Z. Zhang vd., 2012), yeni bir çoklu DoG filtresi kullanmıştır. DoG, bir yüz görüntüsünden yüksek frekansla ilgili bilgileri çıkarır. Kendi oluşturdukları özel veri setinde gerçekleştirilen deneyler neticesinde %0,17 EER değeri elde edilmiştir. Pithadia ve ark. (Pithadia vd., 2012), yüz görüntülerinden köşe, eğri ve kenar öznitelikleri araştırmış ve düşük çözünürlükte çıkarılan bu özniteliklerin yerel geometrisinin, yüksek çözünürlükteki karşılıklarıyla aynı olduğunu belirtmiştir. Daha sonra YİÖ, süper çözünürlüklü geometrik

bilgiyi modellemek için kullanılmıştır. Peng ve Chan (J. Peng & Chan, 2014), aydınlatmalı ve aydınlatmasız yüz görüntüleri arasındaki yüksek frekanslı enerji bileşenlerinin farkını hesaplayan bir Dinamik Yüksek Frekans Tanımlayıcısı (Dynamic High Frequency Descriptor- DHFD) önermiştir. Ekstra aydınlatma, cilt ve saçın aşırı ayrıntılarını ortaya çıkararak canlı yüzün enerjisini yükseltir. Deneyler, kendi oluşturdukları veri seti üzerinde gerçekleştirilmiş ve sonuçlar, önerilen DHFD'nin orijinal yüksek frekans tanımlayıcısından daha güçlü olduğunu göstermiştir.

3) Görüntü Kalitesi Tabanlı Yaklaşımlar

Bu yöntemlerde, yüz canlılığını tespit etmek için Görüntü Kalitesi Analizi (Image Quality Analysis- IQA) tekniği kullanılır. Gerçek erişimlerin ve izinsiz saldırıların görüntü kalite özellikleri farklı olduğundan, görüntü kalitesi analizine dayalı yöntemler, renk çeşitliliği, bulanıklık, kenar bilgisi, kromatik moment gibi öznitelikleri kullanmaktadır. Bu yöntemlerin uygulanması kolay, hesaplama maliyetleri düşüktür ve kullanıcı iş birliğine ihtiyaç duymazlar. Fakat performansları büyük ölçüde görüntülerin kalitesine bağlıdır.

Köse ve Dugelay (Kose & Dugelay, 2013), yüz maskesi saldırılarıyla mücadele için yansıma özelliklerini analiz etmiştir. Varyasyonel retinex algoritması, gri seviyeli bir görüntüyü aydınlatma ve yansıma tabanlı bileşenlere ayrıştırmak için kullanılır. Galbally ve ark. (Galbally vd., 2014b), gerçek ve sahte yüzleri birbirinden ayırt etmek için 25 farklı Görüntü Kalitesi Ölçeği (Image Quality Metrics- IQM) (ortalama kare hatası, tepe sinyali / gürültü oranı, maksimum fark, ortalama fark, vb.) kullanmıştır. Görüntüler, lineer ve ikinci derece diskriminant analizi ile gerçek ya da sahte olarak sınıflandırılmıştır. Bhogal ve ark. (Bhogal vd., 2017), yapay ve gerçek görüntüleri ayırt etmek için altı farklı IQM kullanmıştır. Wen ve ark. (Wen vd., 2015), Görüntü Bozulma Analizi (Image Distortion Analysis- IDA) tabanlı YST yöntemi önermiştir. Çalışmada, yüz görüntülerinden 4 farklı IDA özniteliği (aynasal yansıma, bulanıklık, renk momenti ve renk çeşitliliği) elde edilmiş ve SVM sınıflandırıcısı kullanılarak görüntü için gerçek veya sahte kararı verilmiştir. Agarwal ve ark. (Agarwal vd., 2016), ayrık dalgacık dönüşümü uygulanmış görüntü dizilerinden blok tabanlı Haralick doku öznitelikleri (korelasyon, kontrast, entropi, fark varyansı, toplam ortalaması, vb.) çıkarmış ve bunları SVM ile sınıflandırarak sahtecilik tespiti gerçekleştirmiştir. Boulkenafet ve ark. (Boulkenafet vd., 2017), farklı renk uzaylarından (HSV, YC_bC_r) Hızlandırılmış Güçlü Öznitelikler (Speeded-Up Robust Features- SURF) vektörlerini doğrusal sınıflandırmaya daha uygun yüksek boyutlu bir alana transfer etmek için Fisher vektör kodlaması kullanmıştır. Wang ve ark. (S.-Y. Wang vd., 2017), basılı

fotoğraf ve video oynatma saldırılarına karşı iki yeni öznitelik önermiştir. İlk öznitelik, görüntünün yeşil ve kırmızı kanalları arasındaki farkı bulmak için kullanılırken, diğeri yerel bölgelerdeki renk dağılımını yaklaşık olarak hesaplamaktadır. Daha sonra, her iki öznitelik Çok Ölçekli Yerel İkili Örüntü (Multi-Scale Local Binary Pattern- MSBP) oluşturmak için birleştirilmiştir. Nikisins ve ark. (Nikisins vd., 2018), farklı IQM'ler kullanarak öznitelik uzayı kullanan bir sistem önermiştir. Gerçek yüz örneklerinin olasılık dağılımını sunmak için bir Gauss Karışım Modeli (Gaussian Mixture Model- GMM) eğitilmiştir. Nguyen ve ark. (H. P. Nguyen vd., 2019), gerçek ve sahte görüntüler arasındaki farkları yüz derisi arasında var olan gürültü istatistikleri temelinde araştırmıştır. Çalışmanın sonucunda, yüksek kaliteli saldırı ve gerçek görüntü örneklerini içeren yeni bir yüz sahteciliği önleme veri kümesi tanıtılmıştır. Chang ve Yeh (Chang & Yeh, 2022), çok ölçekli algısal görüntü kalitesi değerlendirme özelliklerini kullanan bir YST algoritması önermiştir. Sahtecilik tespiti için yüz görüntülerinden çıkarılan Gauss yoğunluğu tabanlı, asimetrik genelleştirilmiş Gauss yoğunluğu tabanlı ve üst gradyan benzerlik sapması tabanlı öznitelikler kullanılmıştır. Elde edilen 21 çok ölçekli öznitelik, SVM kullanılarak sınıflandırılmıştır. CASIA ve REPLAY-ATTACK veri setleri için sırasıyla %12,7 EER ve %5,38 HTER elde edilmiştir.

B. Dinamik Yaklaşımlar

Dinamik yaklaşımlarda, yüz biyometrik özelliğinin birden fazla örneği, verilen görüntü kümesinden öznitelikleri çıkarmak için kullanılır. Bazı dinamik YST teknikleri, video tabanlı sahtecilik saldırılarının tespiti için özel olarak önerilmiştir. Bu teknikler, yüz videolarının hem zamansal hem de statik bilgilerinden yararlanmak için önerildiklerinden, genellikle çok rekabetçi bir performans elde ederler. Ancak bu yaklaşımlar, tek yüz görüntüsünün mevcut olduğu senaryolarda başarılı değildir (Galbally vd., 2014a). Dinamik tabanlı yaklaşımlar, doku ve hareket tabanlı olmak üzere iki alt sınıfa ayrılır.

1) Doku Tabanlı Yaklaşımlar

Bu yaklaşımlar, çoklu kareler veya yakalanan video kareleri boyunca dinamik doku bilgisini keşfeder.

Dinamik doku çalışmaları, farklı ortogonal düzlemlerde YİÖ tanımlayıcısının (Local Binary Patterns of Three Orthogonal Planes- LBP-TOP) kullanılmasıyla başlamıştır (De Freitas Pereira vd., 2013, 2014; Komulainen vd., 2013). Genişletilmiş YİÖ, orijinal YİÖ tanımlayıcısının zaman-uzamsal bir uzantısı olan Hacimsel Yerel İkili Örüntüsüne (Volume Local Binary Patterns- VLBP) verilmiştir (Ojala, Pietikäinen, vd., 2002). VLBP operatörü, görünüm ve hareketi dokunun dinamik tanımında birleştirir. Phan ve ark. (Phan vd., 2016)

tarafından YST için daha yüksek dereceli yerel ikili öznitelik tanımlayıcısı olan Yerel Türev Örüntüsü (Local Derivative Pattern- LDP) önerilmiştir. Merkez ve komşu pikseller arasındaki ilişkinin kodlandığı YİÖ'nün aksine, LDP tanımlayıcısı belirli bir bölgeden farklı uzamsal ilişkileri kodlayarak daha yüksek dereceli yerel bilgileri çıkarır. Bharadwaj ve ark. (Bharadwaj vd., 2013), diğer doku tabanlı tekniklere kıyasla daha iyi performans sağlamak ve hareket tahmini için Yönlendirilmiş Optik Akış Histogramı (Histogram of Oriented Optical Flow- HOOF) algoritmasının önerildiği bir teknik tanıtmıştır. Arashloo ve ark. (Arashloo vd., 2015), yüz canlılığı tespiti için Çok Ölçekli Dinamik İkili İstatistiksel Görüntü Öznitelikleri (Multi-Scale Dynamic Binarized Statistical Image Features- MBSIF) ve Çok Ölçekli Dinamik Yerel Faz Niceleme (Multi-Scale Dynamic Local Phase Quantization- MLPQ) özniteliklerinin birleştirildiği bir yaklaşım ortaya koymuştur. Daha sonra bu tanımlayıcıların üç ortogonal düzlemdeki öznitelikleri birleştirilmiştir. Bu iki tanımlayıcıdan elde edilen birleştirilmiş bilginin, birçok yöntemden daha başarılı sonuçlar ürettiği görülmüştür. Tirunagari ve ark. (Tirunagari vd., 2015), göz kırpma, yüz dinamikleri, dudak hareketi gibi canlılık ipuçlarını yakalayan bir Dinamik Mod Ayrıştırma (Dynamic Mode Decomposition- DMD) algoritması uygulamıştır. DMD, YİÖ ve kesişim çekirdeğine sahip SVM'den oluşan bir ardışık düzen sınıflandırma hattı önerilmiştir. Oldukça verimli olan bu yöntemin kullanımı, herhangi bir ayar gerektirmediği için kolaydır. Arashloo ve Kittler (Arashloo & Kittler, 2018), anomali tespitine dayalı yeni bir YST yöntemi önermiştir. Bu yaklaşımda, eğitim verileri sadece pozitif sınıftan alınırken, test verileri hem pozitif hem de negatif sınıflardan gelmektedir. Farklı doku tanımlayıcıları (LBP-TOP, LPQ-TOP) kullanılarak video dizilerinden dinamik öznitelikler çıkarılmıştır. Pan ve Deravi (S. Pan & Deravi, 2018), görüntüleri iki sınıfa ayırmak için dokusal bilgilerdeki zamansal değişiklikleri temel alan Zamansal Birleşimli Komşu Yerel İkili Örüntüler (Temporal Co-occurrence Adjacent Local Binary Patterns- TCoALBP) adlı yeni bir tanımlayıcı tanıtmıştır. Zhao ve ark. (Zhao vd., 2017), dinamik öznitelikleri temsil etmek için Hacimsel Yerel İkili Örüntü Sayısı (Volume Local Binary Count- VLBC) doku tanımlayıcısını önermiştir. Yöntemde herhangi bir t çerçevesindeki bir merkez piksel için; t-1, t ve t+1 çerçevelerinde, R yarıçap mesafesinde eşit olarak konumlanmış P komşu piksel birlikte kullanılarak, merkez piksele göre eşiklenmekte ve bir kod üretilmektedir. Zhou ve ark. (Zhou vd., 2021), üç ortogonal düzlemden Yerel Yönlü Sayı Dokusu (Local Directional Number Pattern Three Orthogonal Planes- LDN-TOP) örüntülerini kullanarak, hareket bilgisi elde etmek için doku tabanlı bir tanımlayıcı kullanmıştır. Renkli görüntülerin tüm kanallarından gelen

birleştirilmiş LDN-TOP histogramları, olasılıksal işbirlikçi temsil tabanlı bir sınıflandırıcı yardımıyla değerlendirilmiştir. Sonuçlar, HSV ve YC_bC_r renk uzaylarının birleştirilmesinin yüz sahteciliği tespit performansını artırdığını göstermektedir. Zhang ve ark. (Y.-J. Zhang vd., 2022), CoALBP, LPQ ve LDN-TOP öznitelik çıkarma yöntemlerinin birleşimlerinden oluşan YST yöntemi önermiştir. Giriş görüntüsünden elde edilen yüz bölgesi YC_bC_r renk uzayına dönüştürüldükten sonra kanallara ayrılmış ve tüm kanallardan CoALBP, LPQ ve LDN-TOP öznitelikleri elde edilmiştir. Son olarak, veri seti SVM kullanılarak sınıflandırılmıştır. Önerilen yöntem, REPLAY-ATTACK veri setinde %0,07 EER ve %1,30 HTER sonuçlarını üretmiştir.

2) Hareket Tabanlı Yaklaşımlar

Hareket analizi tabanlı yöntemler, video dizilerinden hesaplanan optik akış yardımıyla, ağız, baş, göz vb. gibi noktaların hareketlerinden elde edilen karakteristikleri dikkate alırlar (Chakka vd., 2011). Bu yöntemlerin taklit edilmesi zordur ve düşük kullanıcı iş birliği gerektirir. Ancak, yüksek hareket etkinliğine sahip video dizilerine ihtiyaç duyulması ve yüksek hesaplama karmaşıklığı, bu yaklaşımların ana dezavantajlarıdır.

Kim ve ark. (Y. Kim vd., 2011), sahte yüz tespiti için benzerlik ve hareket tabanlı bir algoritma sunmuştur. Başlangıçta bir giriş videosu iki farklı ön plan ve arka plan bölgesine ayrılır. Benzerlik, yüz karesinin arka planları ve yüz bölgeleri arasındaki farkı belirlemek için hesaplanır. Ardından, arka plan bölgesindeki hareket miktarına kıyasla ön plandaki hareket miktarını ölçmek için bir hareket indeksi hesaplanır. Hem benzerlik hem de hareketin kombinasyonu, sahte veya gerçek yüzü tespit etmek için kullanılır. Benzer bir yaklaşımda, Yan ve ark. (Yan vd., 2012), canlılık tespiti için hareket, görüntüleme bantlama etkisi ve yüz-arka plan tutarlılığını içeren üç doğal ipucu kullanmıştır. Bölge dışı ipuçları, göz kırpma olan gerçek bir yüzde sergilenen hareket miktarını gösterirken, yüz-arka plan tutarlılığı, yüz ve arka plan hareketinin gerçek yüz için düşük tutarlılığa sahip olduğunu ve sahte için yüksek olduğunu göstermektedir. Sahte bir yüzün yeniden üretiminde ortaya çıkan görüntü kalitesi kusuru, görüntü bantlama etkisi ile belirlenmiştir. Bantlama, pürüzsüz, soluk bir gradyan yerine bir renkten diğerine gözle görülür değişimi ifade eder. Tüm bu üç ipucunun birleştirilmesi, hatasıza yakın yüz canlılığı tespit doğruluğu sağlanmıştır. Canlılığı tespit etmek için hareket ve dokusal özniteliklere dayalı tamamlayıcı karşı önlemler Komulainen ve ark. (Komulainen vd., 2013) tarafından araştırılmıştır. Hareket tabanlı karşı önlem, arka plan sahnesi ile kafa hareketleri arasındaki korelasyonlar olarak kullanılır. Sahtecilik tespitinin nihai sonucunu sağlamak için oylama yöntemi kullanılır. Anjos ve ark.

(Anjos vd., 2014), optik akış kullanarak ön plan/arka plan hareket korelasyonuna dayanan bir yöntem önermiştir. Önce yatay ve dikey yönler kullanılarak her piksel için hareket yönü elde edilmiştir. Daha sonra yüz ve arka plan bölgeleri için normalize edilmiş histogramlar üretilmiş ve yüz ile arka plan bölgelerinin açı histogramları arasındaki mesafeler hesaplanmıştır. Son olarak, bu değerlerin *N* çerçeveli pencere boyutu üzerinden ortalaması alınarak, sahtecilikler tespit edilmeye çalışılmıştır. Cai ve ark. (Cai vd., 2015), gerçek yüzün bakış yörüngesinin sahte yüze kıyasla daha yüksek belirsizlik seviyesine sahip olduğu varsayımına dayanan bakış tahmini yaklaşımını kullanmıştır. Çalışmanın arkasındaki fikir, bir bakış histogramı elde etmek için hareketli bir videodan bakış özniteliklerini ve hareketi çıkarmaya dayanmaktadır. Bakış histogramından elde edilen bilgi, kullanıcının bakış hareketinin belirsizlik seviyesini belirlemek için kullanılır ve bu da canlılığın tahmin edilmesine yardımcı olur. Singh ve Arora (Singh & Arora, 2017), yüz canlılığı tespiti için göz kırpma ve ağız hareketlerini kullanmıştır. Morfolojik açma işlemleri, yüz videosundan ardışık olarak yakalanan çerçevelerdeki soyut yanlış bölgeleri çıkarmak için kullanılır. Morfolojik açma, görüntüdeki daha büyük nesnelerin şeklini ve boyutunu korurken bir görüntüden küçük nesneleri ve ince çizgileri kaldırmak için kullanılır. Yüz canlılığı, bir kişinin gözlerinin ve/veya ağzının hareketi kontrol edilerek tespit edilmiş ve yaklaşım hem özel hem de kullanıma açık yüz veri kümeleri üzerinde değerlendirilmiştir. Zhang ve Xiang (W. Zhang & Xiang, 2020), bir videonun gerçek olup olmadığını değerlendirmek için Ayrık Dalgacık Dönüşümü (Discrete Wavelet Transform- DWT), YİÖ ve DCT özniteliklerini bir arada kullanmıştır. DWT-YİÖ öznitelikleri, her çerçevedeki DWT bloklarının YİÖ histogramlarından elde edilmiştir. Daha sonra bu öznitelikler üzerinde dikey olarak DCT işlemi gerçekleştirilerek DWT-YİÖ-DCT öznitelikleri üretilmiştir. Bu öznitelikler, Radyal Tabanlı Fonksiyon (Radial Basis Function- RBF) çekirdeğine sahip SVM sınıflandırıcısı ile eğitilmiş ve YST probleminde kullanılmıştır.

C. Derin Öğrenme Tabanlı Yaklaşımlar

Evrişimli Sinir Ağları (Convolutional Neural Networks- CNN) modellerinin ortaya çıkması ve örüntü tanımadaki yaygın uygulamaları ile derin öğrenmeye dayalı yüz sahteciliği saldırı tespit yöntemleri son yıllarda önemli bir atılım gerçekleştirmiş ve geleneksel el yapımı öznitelik tabanlı dedektörleri geride bırakmıştır. Elle çıkarılmış öznitelik tabanlı YST yaklaşımlarıyla karşılaştırıldığında, derin öğrenme (DL) tabanlı teknikler, otomatik olarak öğrenilen katmanlı öznitelik ifadeleri üretir.

Derin öğrenme tabanlı yaklaşımların YST problemine yönelik son gelişmeleri, karmaşık senaryolarda daha yüksek sınıflandırma doğruluğunu göstermekte ve probleme karşı sağlamlık ve etkinlik kazandırmaktadır.

1) CNN Tabanlı Yaklaşımlar

CNN modelleri, özellikle örüntü sınıflandırması için birçok bilgisayarla görme faaliyetinde yaygın olarak kullanılmaktadır. CNN'in YST görevinde kullanılması, yüz görüntülerinin sahte veya gerçek sınıflara etkili bir şekilde ayrılmasına yardımcı olur. İlk olarak 2014 yılında Yang ve ark. (J. Yang vd., 2014), YST alanında DL'nin kullanımıyla derin seviyeli öznitelikleri çıkarmak için CNN modeli kavramını tanıtmıştır. Xu ve ark. (Xu vd., 2016), Uzun-Kısa Vadeli Bellek (Long Short-Term Memory- LSTM) modelini CNN ile birleştiren bir model önermiştir. Bu model, yüz sahteciliğini önleme görevi için videolardan zamansal yapıları öğrenmektedir. Alotaibi ve Mahmood (Alotaibi & Mahmood, 2017), derinlik bilgisi elde etmek ve sınır konumlarını korumak için doğrusal olmayan difüzyon kullanan bir yüz canlılığı tespit yöntemi önermiştir. Ardından, görüntülerin ayırt edici ve yüksek seviyeli özniteliklerini çıkarmak için evrişimsel sinir ağı kullanılmıştır. Tu ve Fang (Tu & Fang, 2017) öğrenme aktarımı (TL) teorisinin kullanıldığı LSTM birimini uygulamış ve önceden eğitilmiş bir ResNet-50 CNN modeli, görüntü çerçevelerinden uzamsal öznitelikleri çıkarmıştır. Bu öznitelikler, zamansal öznitelikler elde etmek için LSTM'ye girdi olarak sunulmakta ve daha sonra sınıflandırma görevi için kullanılmaktadır. Bir çalışmada, temel CNN yapısı Souza ve ark. (G. B. De Souza vd., 2017) tarafından değiştirilmiş ve Yerel İkili Örüntüler Ağı (Local Binary Pattern Network- LBPnet) olarak bilinen yeni bir CNN modeli önerilmiştir. LBPnet'in ilk katmanı YİÖ öznitelik bilgisi ile oluşturulur ve evrişim işlemi, YİÖ değerleri üzerinde gerçekleştirilir. Wang ve ark. (Y. Wang vd., 2017), iki boyutlu derinlik ve dokusal bilginin ortak temsilini kullanan güçlü bir yöntem geliştirmiştir. Doku öznitelikleri CNN'den öğrenilirken, derinlik bilgisi Kinect kullanılarak elde edilmiştir. Yöntem üç ana bileşenden oluşmaktadır; ilkinde 2D görüntüler kullanılarak genelleştirilmiş doku öznitelikleri keşfedilir; ikincisi derinlik görüntüsü kullanılarak bir öznitelik tasvirini içerir ve son olarak bir birleştirme tekniği uygulanır. Canlı ve sahte yüz görüntülerini ayırt etmek için popüler YİÖ öznitelikleri seçilmiştir ve videolar için oylama stratejisi kullanılmıştır. Li ve ark. (H. Li, He, vd., 2018), öğrenilebilir evrişim katmanlarını, sabit parametreli YİÖ katmanlarıyla birleştirerek, ağ parametrelerinin sayısını büyük ölçüde azaltabilen YİÖ tabanlı uçtan uca öğrenilebilir bir ağ tasarlamıştır. Ağ, seyrek ikili filtreler ve türetilebilir simüle edilmiş çıkış fonksiyonlarından oluşmaktadır. Mevcut

DL tespit yaklaşımlarıyla karşılaştırıldığında, önerilen CNN yapısının parametre sayısını 64 kata kadar azaltabileceği sonucuna varılmıştır. Nguyen ve ark. (D. T. Nguyen vd., 2018), yüz derisinin ayrıntılarını çıkarmak için çok seviyeli bir YİÖ özniteliğini CNN ile birleştirmiştir. Benzer şekilde Zhang ve ark. (L.-B. Zhang vd., 2018), gerçek ve sahte yüz görüntüleri arasındaki renk dağılımını iyileştirmek için normalleştirilmiş yüzlerin renk uzayını dönüştürdükleri bir şema sunmuşlardır. Renk kromatik uzayı tespit edildikten sonra, sınıflandırma için bir topluluğu eğitmek üzere YİÖ öznitelikleri çıkarılır. Dokusal özniteliklerin yanı sıra, derin seviye öznitelikleri de keşfetmek için Chen ve ark. (F.-M. Chen vd., 2019), rotasyona dirençli YİÖ (Rotation-Invariant Local Binary Pattern- RI-LBP) aracılığıyla çıkarılan doku öznitelikleri ile CNN ağı aracılığıyla çıkarılan derin seviye öznitelikleri birleştirmiştir. Elde edilen derin özniteliklerin boyutu, Temel Bileşenler Analizi (TBA) algoritması kullanılarak azaltılmıştır. Ardından bu öznitelik kümeleri, RBF çekirdekli bir SVM'yi eğitmek için kullanılmıştır. Benzer şekilde, Grover ve Mehra (Grover & Mehra, 2019), canlı ve sahte yüz görüntülerini ayırt etmek için YİÖ özniteliklerini ve CNN modelini kullanmıştır.

Li ve ark. (L. Li vd., 2020), geçerli renk uzaylarında üst üste binen örnekler sorununun üstesinden gelmek için bir CompactNet yapısı önermiştir. Model, bir uzay üreteci, öznitelik çıkarıcı ve üçlü kayıp fonksiyonu olmak üzere üç aşamadan oluşmaktadır. Başlangıçta, RGB görüntüsü kompakt uzay üretecine daha az parametre ile verilir ve mevcut renk uzaylarının başka bir yeni uzaya eşlenmesine yardımcı olur. Ardından oluşturulan yüz görüntüsü, derin seviyeli özniteliklerin hesaplanması için öznitelik çıkarıcı modülüne girdi olarak verilir. Son olarak, noktadan merkeze entegrasyon mekanizması, sınıflar arası farklılığı en üst düzeye çıkarmak ve sınıf içi farklılığı en aza indirmek için kullanılan üçlü kayıp fonksiyonu ile eğitim örneklerini seçmek için uygulanır. Ma ve ark. (Ma vd., 2020) çok bölgeli CNN tabanlı bir yaklaşım kullanmış ve saçılma gradyan dağılımı için yeni bir yerel sınıflandırma kaybı kavramı ortaya koymuştur. Yakın zamanda, Pei ve ark. (Pei vd., 2023), yüz tanıma işleminden sonra sahteciliğin tespit edildiği kişiye özgü bir yaklaşım sunmuştur. Bu yaklaşımda, yüz saldırılarını tespit etmek için kullanıcıların çift taraflı fotoğrafları üzerinde çalışan bir Siyam Ağı önerilmiştir. Rusia ve Singh (Rusia & Singh, 2022), yüz sahteciliği saldırılarını tespit etmek için renk-doku tabanlı bir derin sinir ağı önermiştir. Bu yaklaşım, yüz görüntülerinin renk ve doku bilgilerini analiz etmek ve olası sahtecilik girişimlerini belirlemek için CNN'leri kullanmaktadır.

2) Tam Evrişimli Sinir Ağı Tabanlı Yaklaşımlar

Tam Evrişimli Sinir Ağı (Fully Convolutional Network- FCN), ilk olarak bilgisayarla görmede semantik segmentasyon görevleri için kullanılan bir sinir ağı mimarisi türüdür. Ağ boyunca uzamsal bilgileri koruyarak ve piksel bazında tahminlere olanak sağlayarak geleneksel evrişimli sinir ağlarından ayrılır.

FCN'lerin temel özelliği, isteğe bağlı boyutlardaki giriş görüntülerini alabilmeleri ve görüntüdeki her piksel için yoğun tahminler üretebilmeleridir. FCN'ler bunu, tipik CNN'lerdeki tam bağlı katmanları evrişimli katmanlarla değiştirerek, ağın farklı boyutlardaki girdileri kabul etmesine ve girdi görüntüsüyle aynı boyutlarda çıktılar üretmesine olanak tanıyarak başarır.

Atoum ve ark. (Atoum vd., 2017) ve Liu ve ark. (Y. Liu vd., 2018), FCN modellerini eğitmek için yardımcı bir işaret olarak yerel etiketler haritasına benzer derinlik haritası kullanmıştır. Atoum ve ark. çalışmasında iyi sonuçlar elde edebilmek adına iki CNN akışı birleştirilmiştir. İlk model yama tabanlı yüz görüntülerine dayanırken, diğeri derinlik bilgisini çıkarmak için tam yüz görüntülerini kullanır. Benzer şekilde, Liu ve ark. çalışmasında, CNN ve Yinelemeli Sinir Ağı (Recurrent Neural Network- RNN) mimarileri entegre edilmiştir. CNN, sahte ve canlı yüz görüntüleri için ayrık derinliklere yol açan doku tabanlı öznitelikleri keşfetmek için derinlik denetimi uygular. Öznitelik vektörü ve tahmini derinlik bilgisi, hizalanmış öznitelik haritalarının oluşturulması için giriş katmanına bir girdi olarak sağlanır. Elde edilen haritalar ve Uzaktan Fotopletismografi (Remote Photoplethysmography- rPPG) denetimi, verilen video karelerindeki zaman değişkenliğini inceleyen RNN modelini eğitmek için uygulanır. Sahtecilik tespit mekanizması için rPPG sinyali ve derinlik haritası kullanılır. George ve Marcel (George & Marcel, 2019) ve Sun ve ark. (Sun, Song, Chen, vd., 2020), yüz canlılığı tespiti için global etiketlerin yerel etiketlere kıyasla daha az verimli olduğunu ortaya koymuştur. George ve Marcel'in çalışmasında oluşturulan bir Yoğun FCN (DenseFCN) modeli ikili denetim ile eğitilmiştir. Çıktı haritaları üzerindeki denetim, ağ yapısını farklı yamalardan gelen verilerden yararlanarak paylaşılan temsilleri öğrenmeye zorlamıştır. DenseFCN, birden fazla karenin işlenmesinin aciliyeti olmadığından, hızlı karar verme kabiliyeti için uygun hale getirilen kare düzeyinde bilgileri işler. Sun ve ark. çalışmasında ise bir FCN ve toplama parçası içeren "Piksel Düzeyinde Yerel Sınıflandırıcıların Uzamsal Birleştirilmesi" (Spatial Aggregation of Pixel-level Local Classifiers- SAPLC) tasarlanmıştır. FCN, tüm yamalar için yerel etiketleri tahmin eder. Ardından, tahmin edilen etiketler görüntü düzeyinde bir karar vermek için birleştirilir. Deb

ve Jain (Deb & Jain, 2021), ayırt edici yüz görüntüleri ipuçları üzerinde eğitilen Öz Denetimli Bölgesel Tam Evrişimli Ağ (Self-Supervised Regional Fully Convolutional Network- SSR-FCN) adlı denetimli öğrenme tabanlı bir yaklaşım önermiştir. Başlangıçta, FCN modeli, saldırıya eğilimli bölgeleri belirlemek ve ayırt edici ipuçlarını öğrenmek için genel görüntülerle eğitilir. Daha sonra model, yalnızca sunum saldırı alanlarını bulmak için yüz görüntüsünün belirli bölgelerini sunarak yerel olarak mevcut ipuçlarını öğrenmek üzere eğitilir. Son olarak test gerçekleştirilir ve gerçek ve sahte yüz görüntülerini ayırt etmek için nihai sınıflandırma puanı hesaplanır. Arora ve ark. (Arora vd., 2022) görüntülerin boyutlarını azaltmak için evrişimsel otomatik kodlayıcılar kullanmıştır. Kodlayıcı ağırlıkları düzleştirici katmanı ve Tam Bağlantılı (Fully Connected- FC) katmanından oluşan başka bir ağa yüklenmiştir. Boyut azaltma işleminden sonra elde edilen görüntüler, gerçek ve sahte sınıf etiketlerine sınıflandırılması için bu modele verilir. Muhammad ve ark. (Muhammad vd., 2022), iki boyutlu yüz saldırıları için Zamansal Dizi Örneklemesi (Temporal Sequence Sampling- TSS) olarak adlandırılan bir video ön işleme tekniği önermiştir. Çalışmada, çeşitli zamansal uzunluklardaki video klipler üzerinde biriken stabilize edilmiş karelerin denetim görevi gördüğü ve etiketlerin TSS yöntemi tarafından otomatik olarak oluşturulduğu, kendi kendine denetimli bir temsil öğrenme şeması sunulmuş ve CNN modelinin özniteliklerinden yararlanılmıştır.

3) Öğrenme Aktarımı Tabanlı Yaklaşımlar

FCN tabanlı yüz sahteciliği dedektörleri CNN metodolojisinden daha üstün olmasına rağmen, FCN çerçeveleriyle ilişkili bazı sınırlamalar vardır. Bu ağlar sabit boyutta girdi kullanır, bu nedenle yüz görüntülerinin rastgele sabit boyuta kırpılması gerekir. Bu, transfer öğrenme kavramını kullanan YST mekanizmalarında alan adaptasyonu ve alan genellemesinin kullanılmasına işaret etmektedir. Alan Uyarlaması (Domain Adaptation-DA), çok boyutlu verilerle ilgilenen bir TL türüdür. Bilgiyi bir kaynaktan hedef alana aktarmayı amaçlar ve farklı bir uygulama senaryosunda sınırlı eğitim verisi olduğunda faydalı olabilir (Csurka, 2017; Ganin & Lempitsky, 2015).

YST mekanizmaları, eğitim ve test alanlarının yüz pozu, yakalama cihazı, yüz görünümü, aydınlatma vb. açısından farklı veri dağılımlarına sahip olması gibi alanlar arası tutarsızlık sorununa dayanır. Sonuç olarak, YST ile ilgili son çalışmalar, anormallik tespiti (yani, zayıf genelleme yeteneği) sorunlarını karşılamak için etki alanı uyarlamasının yanı sıra etki alanı genellemesi kavramını da kullanmaktadır. Etki alanı genelleştirme tekniği ile öğrenilen genelleştirilmiş öznitelikler ayırt edici olmalı ve birden fazla kaynak etki alanı

arasında paylaşılmalıdır. Böylece, çıkarılan öznitelikler çoklu kaynak örneklerinde YST görevi için ortak farklılaştırılmış ipuçlarını keşfedebilir.

Yang ve ark. (J. Yang vd., 2015) ilk olarak, sahteciliğe karşı sınıflandırıcıları gelişmiş performansla eğitmeyi uyumlu hale getiren sanal öznitelikler üretmek için alt nesne etki alanı uyarlaması kavramını kullanmıştır. Daha sonra, Li ve ark. (H. Li, Li, vd., 2018) tarafından denetimsiz bir etki alanı uyarlama yöntemi başarıyla kullanılmıştır. Çalışmada, etiketli kaynaktan, etiketsiz hedef etki alanına yüz öznitelik uzayı dönüşümüne değinilmiştir. Deneyler CASIA-FASD, REPLAY-ATTACK, MSU ve yeni oluşturulan SiW veri setleri üzerinde gerçekleştirilmiştir. Elde edilen sonuçlar, etki alanı uyarlaması olmadan öğrenme yaklaşımına kıyasla etki alanı uyarlamasının uygulanmasıyla %20 iyileşme olduğunu ortaya koymuştur. Sun ve ark. (Sun, Song, Zhao, vd., 2020), "Alan Uyarlamalı ve Kayıpsız Boyut Uyarlamalı Tam Evrişimli Ağ" (Fully Convolutional Network with Domain Adaptation and Lossless Size Adaptation- FCN-DA-LSA) tasarlamıştır. Model, kayıpsız bir boyut uyarlama ön işlemcisini ve ardından alan uyarlama katmanına gömülü olan piksel düzeyinde sınıflandırma için bir FCN'yi hesaplar. Kayıpsız boyut uyarlama katmanı, yüz yakalama işlemi sırasında ortaya çıkan yüksek sıklıktaki ipuçlarının sürdürülmesine yardımcı olur. DA katmanı, aydınlatma koşulları, yüz saldırıları, yüz veri kümeleri ve kameralardaki farklılıklar nedeniyle farklı yüz alanları arasında genelleme yeteneklerini geliştirir. FCN-DA-LSA, DA ve LSA kullanan iki hibrit protokolün altında %11,22 ve %21,92'lik bir HTER ile sonuçlanmıştır. Mohammadi ve ark. (Mohammadi vd., 2020), önceden eğitilmiş YST modelini hedef veri setine uyarlamak için alan kılavuzlu budamanın (Domain Guided Pruning- DGP) kullanıldığı yeni bir tek sınıf DA yaklaşımı önermiştir. Bu budama, CNN modelinin ilk katmanlarında, öğrenilen bazı filtrelerin hedef veri kümesi için iyi bir genelleme yeteneğine sahip olması nedeniyle uygulanırken, diğerleri kaynak veriye özgüdür ve hedef etki alanındaki ağ performansını iyileştirmek için budanması gerekir. Bir başka çalışmada Wang ve ark. (G. Wang vd., 2020), Ayrıştırılmış Temsil (Disentangled Representation- DR-Net) ve Çok Alanlı öznitelik öğrenmeden (Multi-Domain- MD-Net) oluşan bir yaklaşım önermiştir. Ayrıştırılmış temsil, her bir özelliği ayrı, daha düşük boyutlu değişkenlere ayrıştıran bir denetimsiz öğrenme tekniğidir. Üretken modeller aracılığıyla farklı alanlardan çıkarılan ayrıştırılmış öznitelikler, YST mekanizması için alandan bağımsız öznitelikleri öğrenen MD-Net'e girdi olarak verilmiştir. El-Din ve ark. (El-Din vd., 2021), kaynak alan verilerinin çapraz entropi kaybı ile sınıflandırıcıyı eğitmek için kullanıldığı, denetimsiz hedef alan örneklerinin ise alan uyarlamasında kullanıldığı bir yaklaşım

önermiştir. Kotwal ve ark. (Kotwal vd., 2022), 9 katmanlı bir CNN çerçevesi tasarlayarak, binek araçlar için bir YST tekniği önermiştir. Temel amaç, alana özgü katmanları ve temel ağların göreve özgü ayarlarını uyarlayarak veri kümesi kıtlığı sorununu hafifletmektir. Peng ve ark. (F. Peng vd., 2022), mevcut YST yöntemlerinin aydınlatma değişiminden kaynaklanan performans kaybını gidermek için transfer öğrenmeye temelli İki Akışlı Vizyon Dönüştürücüler (Two-Stream Vision Transformers- TSViT) temelli bir çerçeve tasarlamıştır. TSViT'i eğitmek için RGB renk uzayı ve renk restorasyonlu çok ölçekli retineks (Multi-Scale Retinex with Color Restoration- MSRCR) uzayında yüz görüntüleri girdi olarak verilmiştir. İki kaynaktan gelen öznitelikleri başarılı bir şekilde birleştirmek için iki özelliğin tamamlayıcılığını başarılı bir şekilde yakalayabilen, dikkat mekanizması tabanlı etkili bir öznitelik birleştirme yöntemi geliştirilmiştir. Oulu-NPU, CASIA-MFSD ve REPLAY-ATTACK veri setleri üzerinde yapılan çalışmalar sonucunda, veri seti içi testlerde mevcut yaklaşımların çoğundan daha iyi performans gösterdiği belirtilmiştir. Bunun yanında, genelleme amacıyla uygulanan veri setleri arası testlerde de oldukça iyi performans gösterdiğini ortaya koymaktadır. Mevcut yöntemlerin çoğu, etki alanı değişkenliğini en aza indirmek için etki alanı uyarlamasını kullanmaktadır.

Abdullakutty ve ark. (Abdullakutty vd., 2022), el yapımı "Renkli Yerel İkili Örüntüler" (Colour Local Binary Patterns- CLBP) öznitelik çıkarma yöntemlerini derin öğrenme modelleriyle birlikte kullanmıştır. Önceden eğitilmiş derin öğrenme modelleri kullanılarak elde edilen yüksek seviyeli öznitelikler ve CLBP yöntemi ile çıkarılan düşük seviyeli öznitelikler birleştirilerek bir sınıflandırıcıya verilmiştir. CLBP öznitelik vektörü, giriş görüntüsünün YC_bC_r ve HSV renk uzaylarına dönüştürülmesi, tüm kanallardan YİÖ histogramlarının çıkarılması ve bunların birleştirilmesiyle elde edilmiştir. Sonuçlar, derin ve CLBP özniteliklerinin birlikte kullanılmasıyla YST performansının arttığını göstermiştir. ResNet50-CLBP modeli yardımıyla CASIA ve REPLAY-ATTACK veri setleri için en iyi EER ve HTER sonuçları sırasıyla %8,68 ve %2,64 olmuştur. Buna ek olarak, önerilen modelin temel modele kıyasla hesaplama süresinde önemli azalmalar sağladığı belirtilmiştir.

4) Çoklu Kanal Tabanlı Yaklaşımlar

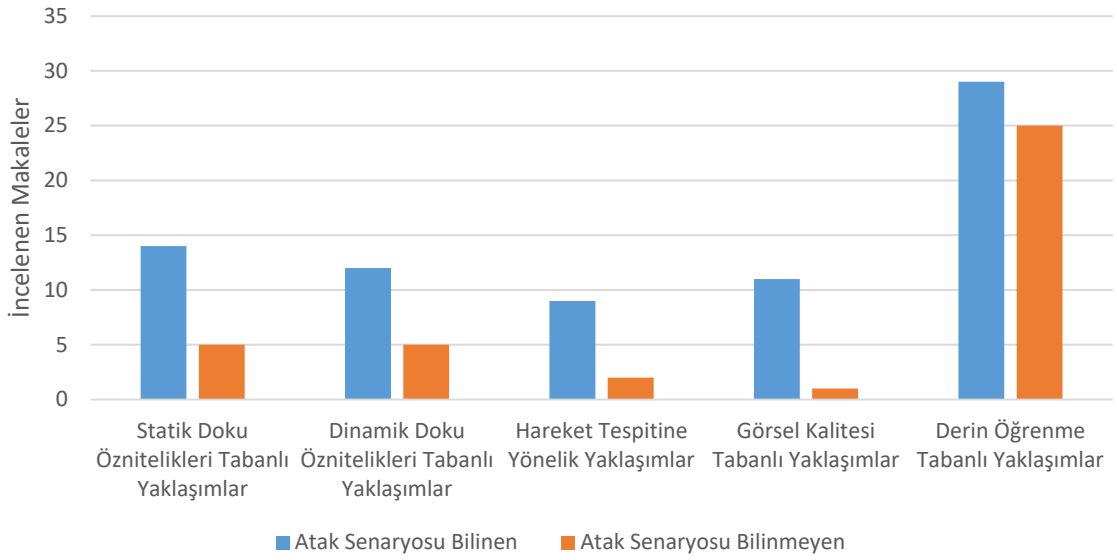
Önceki paragraflarda sunulan görünür spektrum tabanlı gelişmiş YST teknikleriyle ilgili ek bir konu da görüntü yakalama cihazlarının kalitesinin artmasının sırasıyla düşük ve yüksek kaliteli sahte ve gerçek yüz görüntüleri arasındaki ince farkı en aza indirmesidir. 3B maskeler üretmeye yönelik gelişmiş teknolojiler artık kolayca kullanılabilir hale gelmiştir. Bu durum, gerçek ve sahte yüz eserlerini kategorize etmeyi zorlaştırmaktadır. Bu sebeple,

birçok araştırmacı çok kanallı veya genişletilmiş aralıklı görüntüleme sistemi önermiştir. Bhattacharjee ve Marcel (Bhattacharjee & Marcel, 2017), görünür spektrum, termal, kızılötesi ve derinlik kanallarını ele almış ve sırasıyla derinlik ve termal kanallarda 2B ve 3B maske saldırılarını tespit etmenin basit olduğunu göstermişlerdir. Saldırıların çoğu, öğrenme tabanlı bir metodoloji kullanarak farklı kanalların entegrasyonu ile tespit edilmiştir. Benzer şekilde, başka bir çalışmada Bhattacharjee ve ark. (Bhattacharjee vd., 2018), yüz saldırılarını tespit etmek için yüz bölgesinin sıcaklığını araştırmış ve ayrıca özelleştirilmiş silikon tabanlı maskelerle ticari yüz tanıma sistemlerini taklit etme olasılığını göstermiştir. Daha sonra George ve ark. (George vd., 2019), önceden eğitilmiş bir yüz tanıma ağından gelen öğrenme aktarımını dikkate alarak, çoklu kanallardan ortak temsil uygulamak için Çok Kanallı Evrişimsel Sinir Ağı (Multi-Channel CNN- MC-CNN) tasarlamıştır. MC-CNN mimarisinin en büyük avantajı, alana özgü birimlerin (alt katman öznitelikleri) ve üst düzey FC katmanlarının eğitim aşamasında ayarlanmasıdır. Çalışmada LightCNN olarak bilinen bir alt ağ kullanılmıştır. Çalışmanın genel yapısı iki aşamaya ayrılmıştır; ön işleme ve ağ mimarisi. Ön işleme sırasında, yüz tespiti renkli kanalda Çoklu Görev Basamaklı CNN (Multi-Task Cascade CNN- MTCNN) algoritması kullanılarak gerçekleştirilirken, RGB olmayan kanallar için yüz görüntülerinin hem zamansal hem de uzamsal olarak hizalanması gerekmektedir. Yazarlar; birden fazla cihazdan elde ettikleri derinlik, renk, termal ve kızılötesi kanallarını kullanarak, WMCA adlı kendi veri tabanlarını oluşturmuşlardır. Sonuçlardan, önerilen algoritmanın performansının tek renk kanalı kullanıldığında yeterli olmadığı, birden fazla kanalın bir arada kullanılmasının performansı önemli ölçüde artırdığı görülmüştür. Li ve ark. (L. Li vd., 2022), yüz saldırılarını tespit etmek için yüz hareketi ve doku ipuçlarını araştırmıştır. İlk olarak, kesintisiz bir video dizisinden hareketin genliğini ve yönünü karakterize etmek için optik akışlar çıkarılmıştır. Ardından, optik akışlar video kareleriyle birleştirilmiş ve ağın girdisi olarak kullanılmıştır. Bir sonraki adımda, kategorizasyon ağırlıklarını adaptif olarak tahsis etmek için birleşik bölge ve kanal dikkat yöntemleri kullanılmıştır.

Literatürde yapılan çalışmalar incelendiğinde, YST probleminin tüm tekniklerle ele alındığı görülmektedir. Bunun yanında, özellikle güçlü tanımlama özellikleri sebebiyle derin öğrenme tabanlı yaklaşımların son yıllarda popülerleştiği ve çalışmaların bu yöntemler üzerinden gerçekleştirildiği söylenebilir. Derin öğrenme tabanlı yaklaşımları sırasıyla; statik doku öznitelikleri, dinamik doku öznitelikleri, görsel kalitesine dayalı yaklaşımlar ve hareket

tespitine yönelik yaklaşımlar izlemektedir. Şekil 5, YST alanında yapılan çalışmaların dağılımını göstermektedir (Sharma & Selwal, 2023).

Bu tez kapsamında, YST alanında en çok çalışmanın bulunduğu derin öğrenme tabanlı yaklaşımlar ile statik doku öznitelikleri tabanlı yaklaşımlar incelenmiştir.



Şekil 5. YST alanında yapılan çalışmaların dağılımı

1.4. Sinir Ağları

Sinir ağları, insan beyninin yapısından ve işlevselliğinden esinlenen hesaplama modelleridir. Katmanlar halinde organize edilmiş nöronlar olarak bilinen birbirine bağlı düğümlerden oluşurlar. Verilerden öğrenme, karmaşık ilişkileri ele alma ve değişen ortamlara uyum sağlama yetenekleri, onları çeşitli alanlardaki zorlu problemleri çözmek için değerli araçlar haline getirmektedir.

Bu bölümde, Yapay Sinir Ağları (YSA) ve derin öğrenme ağları açıklanmaktadır.

1.4.1. Yapay Sinir Ağları

Yapay sinir ağları, yapay zekanın temel bir bileşenidir ve çeşitli alanlarda yaygın bir uygulama alanı bulmuştur. Birbirine bağlı nöronlardan oluşurlar ve girdi ve çıktı veri kümeleri arasındaki karmaşık ilişkileri öğrenme ve modelleme yeteneğine sahiptirler (Dong vd., 2019). Bu ağlar üç temel katman türünden oluşur: girdi, gizli ve çıktı. Girdi katmanı

ham verileri alırken, gizli katmanlar bilgileri işler, girdilere ağırlıklı toplamlar uygular ve sonuçları doğrusal olmayan bir yapıya sokmak için aktivasyon fonksiyonlarından geçirir. Son olarak, çıktı katmanı, sınıflandırmalar veya tahminler gibi ağın istenen çıktısını üretir.

YSA'ların temel özelliklerinden biri, ağın çıktısındaki hataya dayalı olarak nöronlar arasındaki bağlantıların ağırlıklarını ayarlamayı içeren geri yayılım adı verilen bir süreç aracılığıyla öğrenme ve performansı iyileştirme yetenekleridir (Rumelhart vd., 1986). Eğitim sırasında, bu ağırlıkları gradyan inişi ve geriye yayılma gibi optimizasyon yöntemlerini kullanarak ayarlarlar. Geriye yayılma, hataları hesaplar ve ağırlıkları ağ boyunca geriye doğru güncelleyerek modelin tahminlerini yinelemeli olarak iyileştirir. Bu öğrenme süreci, YSA'ların verilerdeki karmaşık ilişkileri yakalamasına ve gürültü ve belirsizliğin varlığında bile doğru tahminler yapmasına olanak tanır (Ahmed, 2020; Mo, 2021). Bu ağlar tahmin, kontrol sistemleri, örüntü tanıma ve optimizasyon gibi çeşitli uygulamalarda yaygın olarak kullanılmıştır (Boubaker vd., 2022; Mo, 2021; Ranganayaki & Deepa, 2016).

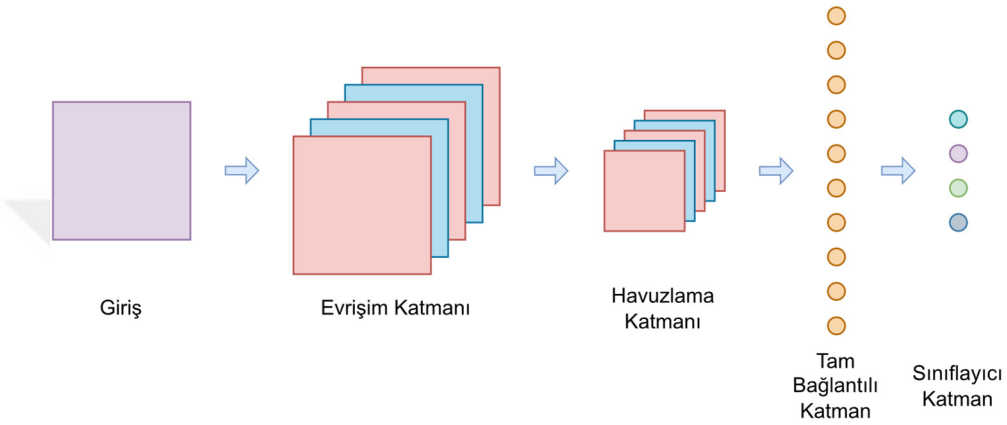
Verilerden öğrenme, karmaşık örüntüleri tanıma ve örneklerden genelleme yapma yetenekleri, sinir ağlarını birçok alanda karmaşık problemleri çözmede güçlü bir araç haline getirmektedir.

1.4.2. Derin Öğrenme Ağları

Derin öğrenme, verilerden temsilleri otomatik olarak öğrenme yeteneği nedeniyle büyük ilgi gören yapay zekâ ve makine öğreniminin bir alt kümesidir. Geleneksel makine öğrenimi algoritmalarının aksine, derin öğrenme, eldeki göreve dayalı olarak ham verilerden doğrudan öznitelik çıkarmayı öğrenebildiğinden, önceden tanımlanmış elle hazırlanmış kurallar aracılığıyla açık öznitelik çıkarımı gerektirmez (Mantach vd., 2022). Geleneksel yöntemlerle karşılaştırıldığında bu durum, özellikle karmaşık ve yapılandırılmamış verileri içeren görevlerde derin öğrenmeye önemli bir avantaj sağlar.

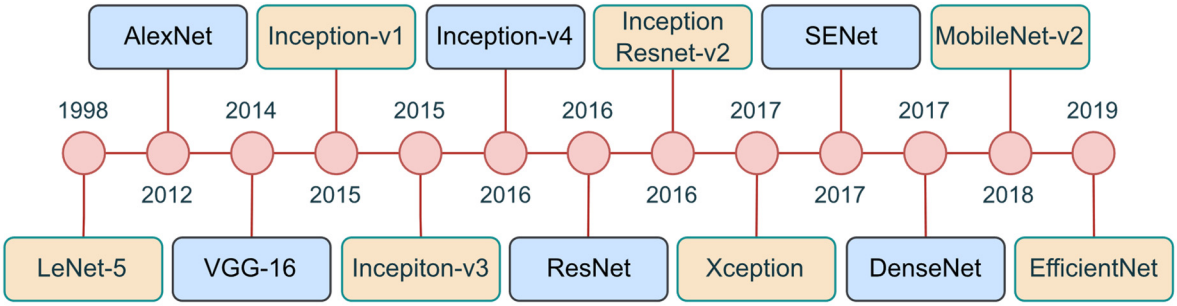
Derin öğrenmeyi diğerlerinden ayıran en önemli unsur, verilerin hiyerarşik temsillerini öğrenme yeteneğidir. Çoklu katmanları sayesinde ağ, ham girdiden daha üst düzey öznitelikleri soyutlayabilir. Eğitim sırasında ağ, tahmin edilen ve gerçek çıktılar arasındaki farkı en aza indirmek için geri yayılım gibi algoritmalar kullanarak parametrelerini yinelemeli olarak ayarlar.

Derin öğrenme teknikleri bilgisayarla görme, doğal dil işleme ve tıbbi görüntü analizi gibi çeşitli alanlar üzerinde önemli bir etkiye sahiptir (Saufi vd., 2018; Y. Wang vd., 2021; S. Zhang vd., 2021). Evrişimli sinir ağları ve uzun-kısa dönemli bellek modelleri gibi modeller, birçok örüntü sınıflandırma uygulamasında büyük başarılar elde ederek, elle oluşturulmuş özniteliklere sahip makine öğrenimi modellerini gölgede bırakmıştır (R. Liu vd., 2020). Şekil 6, evrişimli sinir ağlarının genel yapısını göstermektedir.



Şekil 6. Evrişimli Sinir ağlarının genel yapısı

Evrişimli sinir ağları, güçlü tanımlayıcı özellikleri sebebiyle, özellikle görüntü işleme uygulamalarında sıklıkla tercih edilmektedir. Bu alandaki ilk yaklaşım, LeCun ve Bengio (LeCun vd., 1998) tarafından yapılan çalışmada LeNet olarak bilinen ve elle yazılmış rakamları tanımak için önerilen ağdır. CNN kavramı ilk olarak bu çalışmada ortaya atılmıştır. Krizhevsky ve ark. (Krizhevsky vd., 2012), AlexNet adıyla tanınan ve görüntü sınıflandırması için önerilen CNN modelini 2012 yılında önermiş ve bunu Simonyan ve Zisserman (Simonyan & Zisserman, 2015) tarafından 2014 yılında önerilen VGGNet, 2015 yılı içerisinde Szegedy ve ark. (Szegedy vd., 2015) tarafından önerilen GoogLENet, He ve ark. (He vd., 2016) tarafından 2016 yılında önerilen ResNet gibi sayısız diğerleri izlemiştir. Şekil 7, literatürde sıklıkla kullanılan derin öğrenme ağlarının kronolojik çizelgesini sunmaktadır (Alzubaidi vd., 2021).



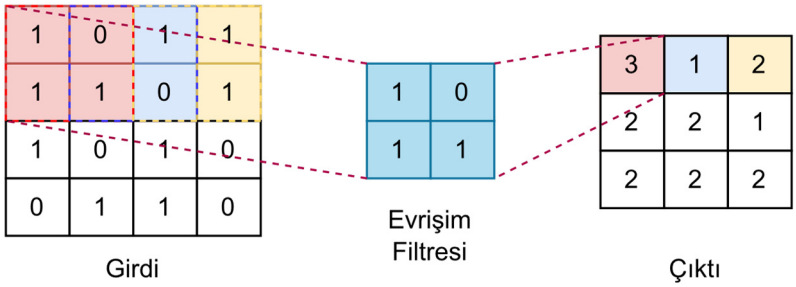
Şekil 7. Derin öğrenme ağlarının kronolojik çizelgesi

1.4.3. Derin Öğrenme Ağlarında Genel Bilgiler

Derin öğrenme ağlarının tasarımında kullanılan katmanlar, mimarinin tanımlama yeteneğinin, öğrenme karakteristiğinin ve özellikle parametre sayısının belirlenmesinde oldukça önemlidir. Bunun yanında, kullanılacak hiper parametrelerin de dikkatli seçilmesi gerekir. Bu kısımda, derin öğrenme ağlarında sıklıkla kullanılan katmanlar ve hiper parametreler hakkında bilgiler verilmiştir.

A. Evrişim Katmanı

Evrişim katmanı, girdi verilerini işlemek ve bunlardan öznitelikler çıkarmak için kullanılan CNN'ler içindeki temel bir yapı taşıdır. Evrişim katmanı, konvolüsyon adı verilen matematiksel bir işlem aracılığıyla giriş verilerine bir dizi öğrenilebilir filtre uygular. Her filtre, tüm girdi boyunca kayarak girdi verilerinin farklı kısımlarını tarar, öznitelik haritaları oluşturmak için eleman bazında çarpma ve toplama işlemlerini hesaplar. Bu öznitelik haritaları, girdi içindeki kenarlar, dokular veya daha karmaşık yapılar gibi öğrenilmiş kalıpları veya öznitelikleri temsil eder. Evrişimsel katmanın temel özellikleri arasında, özniteliklerin uzamsal hiyerarşilerinin öğrenilmesine yardımcı olan parametre paylaşımı ve girdi uzayındaki konumlarına bakılmaksızın yerel örüntüleri yakalama yeteneği yer alır. Şekil 8, evrişim katmanı uygulamasının bir örneğini içermektedir.



Şekil 8. Evrişim uygulaması

Yukarıda örneği verilen evrişim uygulamasında amaç, giriş nöron verilerinin yerel özniteliklerini çıkarmaktır. Giriş verisinin boyutu $n \times n$, evrişim filtresinin boyutu $k \times k$ ve çıkış verisinin boyutu $m \times m$ olarak verildiğinde, girişe ve evrişim filtresine bağlı olarak elde edilecek çıkış verisinin boyutu Denklem 1 kullanılarak hesaplanır.

$$m = n - k + 1 \tag{1}$$

Buna göre, giriş verisinin X ve çıkış verisinin Y ile ifade edildiği bir sistemde, girişe ve evrişim filtresine bağlı elde edilen çıkış verisi Denklem 2 kullanılarak hesaplanır.

$$Y_{ij} = \sum_{i=1}^{k} \sum_{j=1}^{k} (X_{ij} \times K_{ij}) \tag{2}$$

Denklemde X_{ij} ve Y_{ij}, sırasıyla giriş ve çıkış öznitelik haritalarına, K_{ij} ise evrişim filtresine karşılık gelen elemanlardır.

Evrişim katmanı çoğunlukla bir aktivasyon fonksiyonu yardımıyla kullanılır. Bu durum, ağa doğrusal olmayan bir yapı kazandırır ve ağın, girdi verilerinden karmaşık örüntüler ve temsiller öğrenmesini sağlar. Aktivasyon fonksiyonları, bir sinir ağının çıktısının belirlenmesine yardımcı olur. Yaygın olarak kullanılan aktivasyon fonksiyonlarından bazıları şunlardır:

ReLU (Rectified Linear Unit): Basitliği ve etkinliği nedeniyle CNN'lerde en yaygın kullanılan aktivasyon fonksiyonlarından biri haline gelmiştir. Negatif girişler için sıfır çıkış vererek ve pozitif girişleri değişmeden geçirerek doğrusal olmayan bir özellik sunar. ReLU aktivasyon fonksiyonu çıktısı, Denklem 3 kullanılarak hesaplanır.

$$f(x) = \max(0, x) \tag{3}$$

Leaky ReLU, eğitim sırasında bazı nöronların devre dışı kalabileceği "ölen ReLU" sorununu ele alan bir ReLU çeşididir. Negatif girdiler için küçük bir gradyana izin vererek nöronların ölmesini önler. Leaky ReLU aktivasyon fonksiyonu çıktısı, Denklem 4 kullanılarak hesaplanır. Denklemde α küçük bir sabittir (örneğin, 0,01).

$$f(x) = \max(\alpha x, x) \tag{4}$$

Sigmoid: Giriş değerlerini (0, 1) aralığına sıkıştırır ve ikili sınıflandırma problemlerinde kullanışlıdır. Ancak, özellikle derinliğin arttığı ağlarda kaybolan gradyanlardan fazlaca etkilenir. Sigmoid aktivasyon fonksiyonu çıktısı, Denklem 5 kullanılarak hesaplanır.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{5}$$

Tanh: Girdi değerlerini (-1, 1) aralığına sıkıştırarak, sigmoid ile karşılaştırıldığında ortalanmış çıktılar sağlar. Sigmoid fonksiyonunda olduğu gibi, derinliğin arttığı ağlarda kaybolan gradyanlardan fazlaca etkilenir. Tanh aktivasyon fonksiyonu çıktısı, Denklem 6 kullanılarak hesaplanır.

$$f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{6}$$

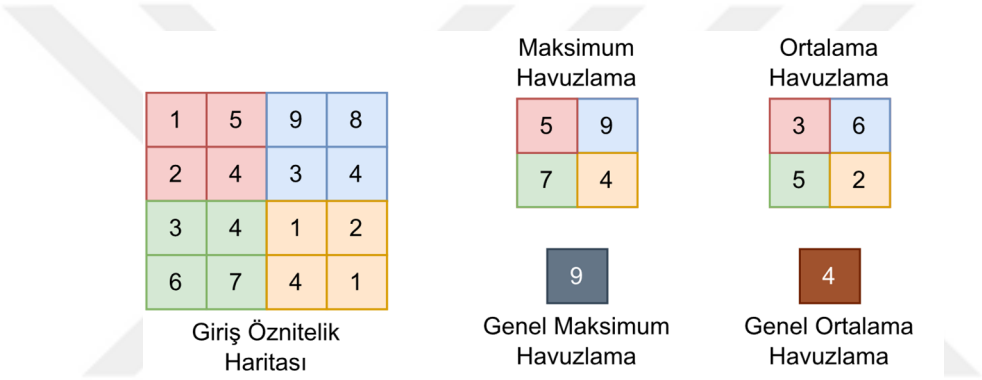
Aktivasyon fonksiyonunun seçimi probleme, ağ mimarisine ve eğitim sırasındaki performansa bağlıdır. ReLU ve Leaky ReLU gibi varyantları, hesaplama verimliliği ve birçok senaryoda iyi performans göstermeleri nedeniyle yaygın olarak tercih edilmektedir.

B. Havuzlama Katmanı

Havuzlama katmanı, evrişim katmanlarından elde edilen öznitelik haritalarının uzamsal boyutlarını küçültmek ve azaltmak için kullanılan bir tekniktir. Maksimum ve ortalama olmak üzere iki yaygın havuzlama türü mevcuttur. Maksimum havuzlama, giriş öznitelik haritasını dikdörtgen bölgelere ayırır ve her bölgedeki maksimum değeri koruyarak en önemli öznitelikleri korurken uzamsal boyutu etkili bir şekilde azaltır. Öte yandan

ortalama havuzlama, her bir bölge içindeki ortalama değeri hesaplar. Bu teknikler, öznitelik haritasının ayrıştırılmış bölgelerine uygulanabildiği gibi, tüm veriye de uygulanabilir. Bu durumda isimleri genel maksimum havuzlama ve genel ortalama havuzlama olarak adlandırılır.

Havuzlama işlemi, parametre ve işlem sayısını azaltarak hesaplama karmaşıklığını azaltır ve bu da bir tür düzenleme sağlayarak aşırı uyumu önlemeye yardımcı olur. Ayrıca, bazı uzamsal bilgileri atarken en önemli öznitelikleri soyutlayarak ve bunlara odaklanarak, öznitelik hiyerarşileri oluşturmaya yardımcı olur. Şekil 9, giriş öznitelik haritasına uygulanan maksimum havuzlama, genel maksimum havuzlama, ortalama havuzlama ve genel ortalama havuzlama işlemlerini göstermektedir.

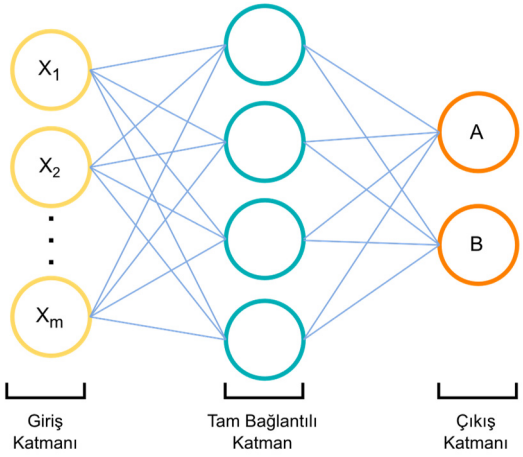


Şekil 9. Havuzlama operatörü uygulaması

C. Tam Bağlantılı Katman

Yoğun katman olarak da bilinen tam bağlı katman, katmandaki her nöronun bir önceki katmandaki her nörona bağlı olduğu bir yapay sinir ağı katmanı türüdür. Bu, bir katmandaki tüm nöronların bir sonraki katmandaki her nörona bağlı olduğu ve her bağlantının kendi ağırlığına sahip olduğu anlamına gelir. Tam bağlantılı katmanlar tipik olarak sinir ağı mimarisinin sonunda bulunur ve önceki katmanlardan gelen öznitelikleri çıktı katmanına bağlamak için kullanılır, bu da onları sınıflandırma ve regresyon gibi görevler için uygun hale getirir.

Sinir ağları bağlamında, birden fazla evrişimsel katmandan sonra, tam bağlantılı katmanlar genellikle evrişimsel katmanlar tarafından çıkarılan öznitelikleri entegre etmek ve nihai çıktıyı üretmek için kullanılır. Son katmanda kullanılacak aktivasyon fonksiyonu, problemde bulunan giriş sayısına göre belirlenir. Çok sınıflı problemlerde genellikle softmax fonksiyonu kullanılmaktadır. Şekil 10, tam bağlantılı katman tasarımına bir örnektir.



Şekil 10. Tam bağlantılı katmanın gösterimi

D. Sınıflayıcı Katman

Sınıflandırma görevleri için kullanılan bir CNN'de, sınıflandırma katmanları tipik olarak ağ mimarisinin sonuna doğru ortaya çıkar. Bu katmanlar, önceki evrişim ve havuzlama katmanları tarafından öğrenilen öznitelikleri alır ve bunları, giriş verilerini farklı sınıflara veya kategorilere sınıflandırmak için kullanır. Bir CNN'in sınıflandırma katmanlarında aşağıdaki bileşenler bulunmaktadır:

Düzleştirme Katmanı: Evrişim ve havuzlama katmanlarından elde edilen çıktı tipik olarak 3D veya daha yüksek boyutlu bir formattadır (yükseklik, genişlik, kanallar/öznitelikler). Tam bağlantılı katmanlara geçmeden önce, bu uzamsal bilgiyi tek boyutlu bir vektöre dönüştürmek için bir düzleştirme katmanı kullanılır. Bu katman, çok boyutlu öznitelik haritalarını, yoğun (tam bağlantılı) katmanlara beslenebilecek tek bir sürekli vektöre dönüştürür.

Yoğun Katmanlar: Öznitelikler düzleştirildikten sonra bir veya daha fazla tam bağlantılı katmandan geçirilir. Bu katmanlar tam bağlantılı katman olarak çalışır ve bir katmandaki her nöronu bir sonraki katmandaki her nörona bağlar. Yoğun katmanlar, ağın daha önceki katmanlarda öğrenilen öznitelikleri birleştirerek, karmaşık örüntüleri öğrenmesini sağlar. Son yoğun katman genellikle sınıflandırma problemindeki sınıf sayısına eşit sayıda nörona sahiptir.

Çıktı Katmanı: Ağın son katmanı, tipik olarak problemin niteliğine bağlı olarak belirli bir aktivasyon fonksiyonu kullanan çıktı katmanıdır. Bu alanda, iki sınıflı problemler için Sigmoid, çok sınıflı problemler için Softmax aktivasyon fonksiyonları kullanılır.

- Sigmoid Aktivasyon: İkili sınıflandırma problemleri için tek bir nöron çıkışı kullanılır. Çıktıyı 0 ile 1 arasında sıkıştırır ve girdinin pozitif sınıfa ait olma olasılığını temsil eder.
- Softmax Aktivasyonu: Çok sınıflı sınıflandırma için çıktı katmanında kullanılır. Çıktıyı birden fazla sınıf üzerinde bir olasılık dağılımına normalleştirir ve her çıktı, girdinin belirli bir sınıfa ait olma olasılığını temsil eder. En yüksek olasılığa sahip sınıf, tahmin edilen sınıf olarak kabul edilir.

E. İnce Ayar

Önceden eğitilmiş bir sinir ağını alıp yeni bir görev veya veri kümesi üzerinde daha fazla eğitme sürecini ifade eder. İnce ayar sırasında, öğrenilen öznitelikleri yeni verilerin belirli özelliklerine uyarlamak için ağdaki katmanların bir kısmı veya tamamı güncellenebilir. İnce ayar genellikle önceden eğitilmiş model büyük bir veri kümesi üzerinde faydalı öznitelikler öğrendiğinde ve kullanıcı bu öznitelikleri daha küçük bir veri kümesiyle ilgili bir göreve uygulamak istediğinde kullanılır.

F. Öğrenme Aktarımı

Bir görev üzerinde eğitilen bir modelin ikinci, ilgili bir görevi yerine getirmek üzere uyarlandığı bir makine öğrenimi tekniğidir. Derin öğrenmede bu genellikle önceden eğitilmiş bir sinir ağının yeni bir görev için öznitelik çıkarıcı olarak kullanılmasını içerir.

Tipik iş akışı, bir modelin büyük bir veri kümesi (kaynak görev) üzerinde önceden eğitilmesini, son sınıflandırma katmanlarının kaldırılmasını ve öğrenilen özniteliklerin daha sonra hedef görev üzerinde eğitilen yeni bir modele aktarılmasını (gerekirse ince ayar) içerir.

G. Veri Büyütme

Eğitim verilerine döndürme, çevirme ve yakınlaştırma gibi rastgele dönüşümlerin uygulanmasını içerir. Bu teknik, eğitim setinin çeşitliliğini artırmaya yardımcı olur ve modelin genellemesini geliştirir. Özellikle eğitim veri setinin boyutu sınırlı olduğunda faydalıdır. Aşırı uyumu önlemeye yardımcı olur ve modelin giriş verilerindeki varyasyonları ele alma yeteneğini geliştirir.

1.4.4. Derin Öğrenme Kavramında Hiper Parametreler

Derin öğrenmede hiper parametreler, eğitim verilerinden öğrenilmeyen ancak eğitim sürecinden önce ayarlanan ayarlar veya yapılandırmalardır. Bunlar modelin genel

davranışını kontrol eder ve öğrenme sürecini etkiler. Uygun hiper parametrelerin seçilmesi, bir derin öğrenme modelinin başarısı için çok önemlidir.

Derin öğrenme modelleri çeşitli uygulamalar için güçlü araçlardır, ancak performansları büyük ölçüde hiper parametrelerin seçimine bağlıdır. Hiper parametreler, verilerden öğrenilmeyen ancak eğitim süreci başlamadan önce ayarlanan parametrelerdir Modelin belirli bir veri kümesinden öğrenme yeteneğini belirlerler. Arama uzayı çok geniş olabileceğinden ve her bir hiper parametrenin modelin performansı üzerindeki etkisi genellikle doğrusal olmayan ve birbirine bağlı olduğundan, hiper parametreleri optimize etmek zorlu bir görevdir (Agrawal vd., 2021).

Derin öğrenme uygulamalarında sıklıkla kullanılan bazı parametreler aşağıda açıklanmıştır:

A. Öğrenme Oranı

Eğitim sırasında modelin parametrelerinin güncellendiği adım boyutunu belirleyen önemli bir hiper parametredir. Optimizasyon algoritmasının (Örn. gradyan inişi) her iterasyonunda, ağırlıklara yapılan güncellemelerin büyüklüğünü kontrol eden skaler bir değerdir.

İyi ayarlanmış bir öğrenme oranı etkili eğitim için çok önemlidir. Çok yüksek bir öğrenme oranı optimizasyon sürecinin minimumun üzerine çıkmasına neden olabilirken, çok düşük bir öğrenme oranı yavaş yakınsamaya veya yerel minimumlarda takılıp kalmaya neden olabilir. Öğrenme oranları tipik olarak 0,1 ila 0,0001 aralığında ayarlanır. Eğitim sürecini izlemek ve model etkili bir şekilde öğrenmiyorsa öğrenme oranını ayarlamak gerekir. Çalışmada kullanılan tüm ağlarda öğrenme oranı 0,001 olarak belirlenmiştir.

B. Yığın Boyutu

Optimizasyon algoritmasının bir iterasyonunda kullanılan eğitim örneklerinin sayısını ifade eder. Modelin parametreleri güncellenmeden önce kaç örneğin işleneceğini belirler.

Yığın boyutu hem bellek gereksinimleri hem de eğitim sürecinin kararlılığı üzerinde etkisi vardır. Daha küçük parti boyutları daha hızlı yakınsamaya yol açabilir ancak güncellemelerde gürültüye neden olabilir. Daha büyük parti boyutları daha kararlı güncellemeler sağlayabilir ancak daha fazla bellek gerektirebilir. Yığın boyutları genellikle 2'nin kuvvetleri olarak seçilir (Örn. 32, 64, 128). Hesaplama kaynaklarını, bellek kısıtlamalarını ve güncellemelerde istenen kararlılık düzeyini dengelemek önemlidir.

C. Devir Sayısı

Devir sayısı, Devir sayısı, eğitim aşamasında model eğitiminin kaç iterasyon süreceğini gösteren değerdir. Bu değer, modelin eğitim sırasında tüm veri kümesini kaç kez gördüğünü belirler. Çok az sayıda devir sayısı için eğitim yapmak, modelin verilerden yeterince öğrenmediği yetersiz uyuma yol açabilir. Çok fazla devir sayısı için eğitim, modelin eğitim verilerini ezberlemeye başladığı aşırı uyuma yol açabilir.

Devir sayısı, problemin karmaşıklığı, veri kümesi boyutu ve model mimarisi gibi faktörlere bağlıdır. Eğitimin ne zaman durdurulacağını belirlemek için eğitim ve doğrulama kaybını izlemek yaygınca tercih edilen yöntemlerdendir. Çalışmada kullanılan tüm ağlarda devir sayısı değeri 100 olarak belirlenmiştir.

D. Aktivasyon Fonksiyonları

Aktivasyon fonksiyonları ağa doğrusal olmayan bir özellik katar. Her bir nörondaki girdilerin ağırlıklı toplamını o nöronun çıktısına dönüştürürler. Doğrusal olmayan özellikler, derin öğrenme modellerinin verilerdeki karmaşık örüntüleri ve ilişkileri öğrenmesi için gereklidir. Yaygın aktivasyon fonksiyonları arasında ReLU, sigmoid ve tanh bulunur.

ReLU, basitliği ve etkinliği nedeniyle gizli katmanlar için popüler bir seçimdir. Bununla birlikte, çıktı katmanı için aktivasyon fonksiyonunun seçimi probleme bağlıdır. Yapılan çalışma iki sınıflı bir problem olduğundan, çıktı katmanlarında sigmoid, ağ içerisindeki katmanlarda ise ReLU ve sigmoid aktivasyon fonksiyonları kullanılmıştır.

E. Katman ve Birim Sayısı

Katman sayısı ve her katmandaki birim (nöron) sayısı dahil olmak üzere sinir ağının mimarisini ifade eder. Daha fazla birime sahip daha derin ağlar potansiyel olarak daha karmaşık kalıpları öğrenebilir, ancak aynı zamanda aşırı uyuma daha yatkın olabilirler. Daha basit problemler çok fazla katman gerektirmeyebilir.

Mimari, sorunun karmaşıklığına ve mevcut veri miktarına bağlıdır. Daha karmaşık problemler daha derin ağlardan faydalanabilirken, daha basit problemler çok fazla katman gerektirmeyebilir.

F. Düzenlileştirme Teknikleri

Düzenli hale getirme teknikleri, modelin parametrelerine kısıtlamalar ekleyerek aşırı uyumu önlemeye yardımcı olur. Yaygın teknikler arasında L1 ve L2 düzenlileştirme, bırakma ve toplu normalleştirme yer alır.

Aşırı uyum, bir model eğitim verilerine çok yakından uymayı öğrendiğinde ortaya çıkar ve görünmeyen verilere zayıf genelleme ile sonuçlanır. Düzenlileştirme tekniğinin ve hiper parametrelerinin seçimi, spesifik probleme ve model karmaşıklığına bağlıdır.

G. İyileştiriciler

İyileştiriciler (optimizer), kayıp fonksiyonunu en aza indirmek için eğitim sırasında modelin parametrelerini (ağırlıklarını) güncellemekten sorumludur. Bunu, ağırlıklara göre kaybın gradyanlarını hesaplayarak ve bunları buna göre ayarlayarak yapar.

Her iyileştirici kendi güncelleme kurallarına sahiptir. Bu durum, eğitimin yakınsama hızını ve kararlılığını etkileyebilir. Yaygın iyileştiriciler arasında Stokastik Gradyan İnişi (Stochastic Gradient Descent- SGD), Adam, RMSprop bulunur. Bu parametrenin seçimi; problem, veri seti ve model mimarisi gibi faktörlere bağlıdır. Belirli bir görev için hangi iyileştiricinin en iyi şekilde çalıştığını bilmek mümkün değildir. Bu sebeple farklı teknikler kullanılarak çalışmanın tekrar edilmesi ve karşılaştırmalar sonucu kullanılacak tekniğin belirlenmesi gerekebilir. Çalışmadaki tüm ağlarda Adam iyileştiricisi kullanılmıştır.

H. Başlangıç Ağırlıkları

Ağın ağırlıklarına atanan başlangıç değerleri öğrenme sürecini etkileyebilir. Doğru başlatma, kaybolan veya patlayan gradyanlar gibi sorunların önlenmesine yardımcı olabilir. Ağırlıklar başlangıçta çok büyük veya çok küçükse, yakınsamada zorluklara ve eğitimde istikrarsızlığa yol açabilir. Başlangıç ağırlıklarının rastgele belirlenmesinin dışında kullanılan farklı teknikler de mevcuttur. Seçim, aktivasyon fonksiyonuna ve spesifik probleme bağlıdır.

İ. Düğüm Seyreltme Oranı

Eğitim sırasında rastgele seçilen nöronların göz ardı edildiği bir düzenleme tekniğidir. Bu değer, belirli nöronlara olan bağımlılığı azaltarak aşırı uyumu önlemeye yardımcı olur. Düğüm seyreltme oranı, ağı daha sağlam olmaya teşvik eder ve herhangi bir özelliğe veya nörona çok fazla güvenmesini önler. Düğüm seyreltme oranı için yaygın değerler 0,2 ila 0,5 arasında değişir, ancak optimum oran belirli probleme ve mimariye bağlıdır. Çalışmadaki ağlarda farklı düğüm seyreltme oranları (0,2 ~ 0,5) kullanılmıştır.

J. Öğrenme Oranını Zamanlama

Eğitim sırasında öğrenme oranının zaman içinde nasıl değiştiğini tanımlar. Sabit olabilir, zamanla azalabilir veya belirli koşullara göre ayarlanabilir. Öğrenme oranının uyarlanması, eğitimin stabilize edilmesine ve optimizasyon sürecinin ince ayarının yapılmasına yardımcı olabilir. Örneğin, eğitim ilerledikçe öğrenme oranını azaltmak

modelin daha doğru bir şekilde yakınsamasına yardımcı olabilir. Farklı öğrenme hızı programları, farklı problem türleri için daha etkili olabilir.

K. Kayıp Fonksiyonu

Modelin tahmin edilen çıktıları ile eğitim verilerindeki gerçek etiketler arasındaki farkı ölçer. Modelin ne kadar iyi performans gösterdiğinin bir ölçüsünü sağlar. Uygun bir kayıp fonksiyonunun seçilmesi, problemin doğasına ve modelin istenen davranışına bağlıdır. Kayıp fonksiyonunun seçimi genellikle çözülmekte olan spesifik problem tarafından belirlenir. Çalışmadaki tüm ağlarda olasılık tabanlı "BinaryCrossentropy" kayıp fonksiyonu kullanılmıştır. Bu fonksiyon, ikili sınıflandırma uygulamalarında gerçek etiketler ile tahmin edilen etiketler arasındaki çapraz entropi kaybını hesaplar.

L. Erken Durdurma

Belirlenen devir sayısı boyunca, doğrulama kaybında iyileşme olmaması gibi belirli bir koşul karşılandığında eğitim sürecini durdurarak aşırı uyumu önlemek için kullanılan bir tekniktir. Erken durdurma, modelin çok uzun süre eğitilmesini önlemeye yardımcı olur, bu da aşırı uyuma yol açabilir. Model doğrulama kümesinde iyi bir performans elde ettiğinde eğitimin durdurulmasını sağlar.

Eğitimin ne zaman durdurulacağının seçimi probleme ve veri setine bağlıdır. Doğrulama kaybını izlemek ve sabit kaldığı noktaya ulaşıldığında veya artmaya başladığında eğitimi durdurmak yaygın kullanılan tekniklerdendir. Çalışmadaki tüm ağlarda erken durdurma kriteri olarak geliştirme setinin kayıp değeri takip edilmiş ve sabır parametresi 10 olarak belirlenmiştir. Bu durum, geliştirme setinde elde edilen her kayıp değerinin, üzerinden 10 devir sayısı boyunca iyileştirilmesini takip eder ve en düşük değerin elde edilmesinin akabinde bu sayaç yeniden başlar. Durdurma kriterine ulaşılmadığı noktada devir sayısı sınırlayıcısı devreye girer.

M. Ağırlık Azalması

Ağırlıkların büyüklüğüne bağlı olarak kayıp fonksiyonuna bir ceza terimi ekleyen bir L2 düzenlemesi biçimidir. Büyük ağırlıkları bastırarak aşırı uyumu önlemeye yardımcı olur. Ağırlık azaltma, modelin karmaşıklığını kontrol etmek ve verilerdeki gürültüye uymasını önlemek için yararlı bir tekniktir. Ağırlık azalmasının gücü belirli probleme ve veri kümesine göre ayarlanmalıdır.

1.5. Kullanılan Veri Setleri

Bu bölümde, çalışmada kullanılan veri setleri hakkında genel bilgiler paylaşılmıştır.

1.5.1. NUAA

NUAA basılı fotoğraf saldırıları için halka açık ilk YST veri kümesidir. Orijinal yüz görüntülerinin kaydında bir web kamerası kullanılmıştır. Veri seti, 15 denekten elde edilen görüntülerden oluşmaktadır. Bu görüntülerin ediniminde deneklerden, göz kırpmalarından kaçınmaları ve nötr yüz ifadeleriyle önden poz vermeleri istenmiştir. Saldırılar, fotoğraf ya da A4 kâğıda basılı fotoğraflar kullanılarak gerçekleştirilmiştir. Veri kümesi, eğitim ve test için iki ayrı alt gruba ayrılmıştır. Eğitim seti; 9 gerçek kullanıcıdan elde edilen 1.743 yüz görüntüsünü ve bu kullanıcıları taklit eden 1.748 sahte görüntüyü içermektedir. Test seti; 3.362 gerçek görüntüyü ve 5.761 sahte görüntüyü içermektedir. Veri seti üç farklı şekilde sunulmaktadır: i) Ham giriş görüntülerini içeren veri seti, ii) Yüz bölgelerinin Viola-Jones dedektörü (Viola & Jones, 2004) ile belirlenmesi ve kırpılmasıyla oluşturulan veri seti, iii) Tespit edilen yüz bölgelerinin göz noktalarına göre hizalanması (Tan vd., 2009) ile normalize edilmiş ve 64×64 piksel boyutlarına ölçeklenmiş veri seti. Şekil 11, NUAA veri setine ait örnek görüntüleri içermektedir (Tan vd., 2010).

Şekil 11. NUAA veri setine ait örnek görüntüler

1.5.2. CASIA-FASD

CASIA-FASD veri seti hem basılı fotoğraf hem de video yeniden oynatma saldırıları sağlayan halka açık ilk yüz SST veri kümesidir. CASIA-FASD veri seti, üç tür saldırıdan oluşmaktadır: i) Bükülmüş fotoğraflar (kâğıt maske saldırılarını taklit eder), ii) Göz bölgeleri

kesilmiş basılı fotoğraflar, iii) Video saldırıları (göz kırpma, ağız ve kafa hareketleri gibi canlılık belirtileri içerir).

Her gerçek yüz ve saldırısı videosu; düşük, normal ve yüksek olmak üzere 3 farklı kalitede toplanır. Yüksek kaliteli videoların çözünürlüğü 1280×720 piksel, düşük / normal kalitedeki videoların çözünürlüğü 640×480 pikseldir. Veri kümesi, eğitim ve test için iki ayrı alt gruba ayrılmıştır. Eğitim seti 20 kullanıcıdan elde edilen gerçek ve sahte görüntüleri, test seti ise 30 kullanıcıdan elde edilen gerçek ve sahte görüntüleri içermektedir. Üç farklı görüntü kalitesi için üç farklı saldırı türü (bükülmüş / kesilmiş fotoğraf saldırısı ve video yeniden oynatma saldırısı) ve tüm verilerin bir arada olduğu genel test ile birlikte toplamda yedi test senaryosu bulunmaktadır. Şekil 12, CASIA-FASD veri setine ait örnek görüntüleri içermektedir (Z. Zhang vd., 2012).

Şekil 12. CASIA-FASD veri setine ait örnek görüntüler

1.5.3. REPLAY-ATTACK

REPLAY-ATTACK veri seti, PRINT-ATTACK (Anjos & Marcel, 2011) veri setini geliştiren ekip tarafından geliştirilmiş, kapsamlı bir yüz sahteciliği tespiti veri setidir. REPLAY-ATTACK, PRINT-ATTACK veri seti ile karşılaştırıldığında, Phone-Attack ve Tablet-Attack olmak üzere fazladan iki saldırıya sahiptir. Phone-Attack, video veya fotoğraf saldırısı için bir iPhone ekranının kullanıldığı atak türüdür. Tablet-Attack ise; bir iPad yardımıyla yüksek çözünürlüklü (1024×768) dijital fotoğrafların veya videoların kullanıldığı atak türüdür. Böylece REPLAY-ATTACK veri seti, basılı fotoğraf veya ekranlar kullanılarak yapılan fotoğraf saldırılarını ve video tekrarlama saldırılarını değerlendirmek için kullanılabilmektedir. Veri kümesi; fotoğraf ve video saldırılarını tespit etmek amacıyla farklı ışıklandırma koşulları altında 50 denekten elde edilen 1.300 farklı

görseli içeren, eğitim, geliştirme ve test gruplarından oluşmaktadır. Şekil 13, REPLAY-ATTACK veri setine ait örnek görüntüleri içermektedir (Chingovska vd., 2012).

Şekil 13. REPLAY-ATTACK veri setine ait örnek görüntüler

Tablo 1, çalışmada kullanılan veri setlerine ait örnek sayıları, aydınlatma koşulları, saldırı türleri gibi detaylı bilgileri içermektedir.

Tablo 1. Çalışmada kullanılan veri setlerine ait bilgiler

Veri Seti	Kişi Sayısı	Aydınlatma Koşulları	Cihaz Konumu	Saldırı Türleri	Gerçek/Sahte Video* Sayısı
NUAA	15	Kontrolsüz	Elde	Fotoğraf	5.105/7.509
CASIA-FASD	50	Kontrolsüz	Elde	Düşük kalite Normal Kalite Yüksek Kalite Bükülmüş Fotoğraf Kesilmiş Fotoğraf Video Oynatma Tümü	150/450
REPLAY-ATTACK	50	Kontrollü Kontrolsüz	Elde Sabit Tümü	Yüksek Çözünürlük Mobil Basılı Fotoğraf Dijital Fotoğraf Video Oynatma Tümü	200/1.000

* NUAA veri setinde Fotoğraf

1.6. Boyut İndirgeme

Bir sınıflandırma işleminde hesaplama karmaşıklığı, hesaplama süresi ve depolama gibi maliyetleri en aza indirmek esastır. Öznitelik vektörünün boyutunun azaltılması bu sürecin temel aşamalarından biridir. Boyut indirgeme, veri kümesini en iyi şekilde temsil edecek ve en iyi sınıflandırma doğruluğunu elde edecek alt kümeleri belirlemeyi amaçlar.

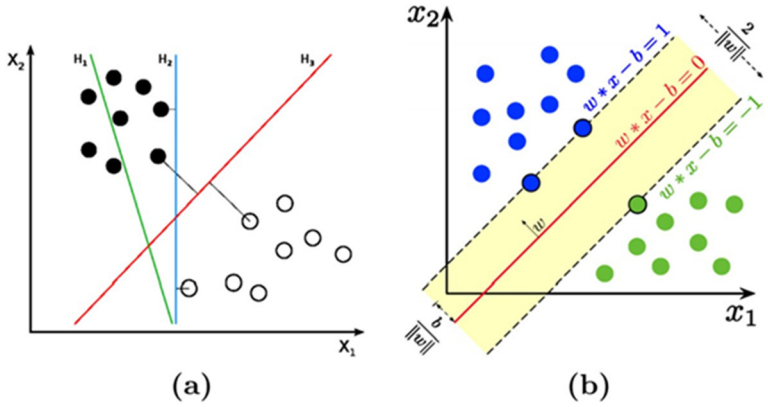
Literatürde birçok boyut indirgeme tekniği mevcuttur. Çalışmada Temel Bileşen Analizi (TBA) yöntemi kullanılmıştır.

TBA, çok boyutlu bir uzaydaki verilerin varyansı en üst düzeye çıkaracak şekilde daha düşük boyutlu bir uzaya yansıtılması yöntemidir (Alpaydin, 2010). Temel amaç, veri kümesini en iyi şekilde temsil eden ve temel bileşenler olarak adlandırılan değişkenlerin doğrusal kombinasyonlarını bulmaktır. Bu temel bileşenler, üzerlerine yansıtılan verilerin varyansını maksimize eden öz vektörlere karşılık gelir (Croux vd., 2013). Veri kovaryans matrisinin (S) öz vektörleri $W_{opt} = argmax_{\|W\|=1} W^T SW$ denklemi ile elde edilir. Denklem çözüldüğünde, S'nin en büyük d $(d \geq D)$ öz değerine karşılık gelen öz vektörleri (W) elde edilir. Daha sonra $y_i = W^T x_i (y_i \in R^d)$ kullanılarak boyut azaltma işlemi gerçekleştirilir (Günay & Nabiyev, 2017). Çalışmada %95 öz değere sahip temel bileşenler kullanılmıştır.

1.7. Sınıflandırma

Makine öğrenimi algoritmaları genel olarak sınıflandırmayı içeren denetimli öğrenme ve kümelemeyi içeren denetimsiz öğrenme olmak üzere iki ana kategoriye ayrılır. Sınıflandırmanın önemi, verileri belirli özelliklere dayalı olarak farklı sınıflara veya gruplara ayırma, kalıpların tanımlanmasını ve tahminlerin yapılmasını sağlama yeteneğinde yatmaktadır. Sınıflandırma, bilgisayarların verilerden öğrenmesini, tahminlerde veya kararlarda bulunmasını ve öğrenilen örüntülere dayalı olarak yeni örnekleri sınıflandırmasını sağlayan temel bir makine öğrenimi problemidir. Bu çalışmada, elle çıkarılan özniteliklerin sınıflandırılmasında SVM algoritması kullanılmıştır.

SVM, istatistiksel öğrenme teorisine dayanan denetimli bir sınıflandırma algoritmasıdır. Sınıflandırıcının matematiksel algoritmaları başlangıçta iki sınıflı verilerin doğrusal sınıflandırılması için tasarlanmış ve daha sonra çok sınıflı ve doğrusal olmayan verilerin sınıflandırılması için genelleştirilmiştir. Sınıflandırıcıda temel amaç, iki sınıfı birbirinden en iyi şekilde ayırabilen hiper düzlemi bulmaya dayanır (Kavzoğlu & Çölkesen, 2010). Şekil 14, iki sınıfı ayıran SVM modelini temsil etmektedir.



Şekil 14. SVM sınıflandırıcısında hiper düzlem belirleme

Şekil 14(a)'da H1 düzlemi sınıfları doğru bir şekilde ayıramamaktadır. H2 düzlemi sınıfları başarılı bir şekilde ayırmıştır, ancak örnekler ile hiper düzlem arasındaki mesafe minimumdur. H3 düzlemi ise sınıf örneklerini maksimum mesafe ile ayırmıştır. Şekil 14(b)' de iki sınıfın en yakın örnekleri arasından geçen düzlem bu sınıfları ayırmak için optimum çözümdür. Düzlemden eşit uzaklıktaki hayali noktaları kesen örnekler destek vektörleri olarak adlandırılır. Çalışmada, giriş görüntüsünü gerçek/sahte olarak sınıflandırmak için RBF çekirdeğine sahip SVM kullanılmıştır.

1.8. Değerlendirme Ölçütleri

Bir YST modelinin başarımı; Doğruluk, Kesinlik, Duyarlılık, F1-Skoru, Özgüllük, Yanlış Kabul Oranı (False Acceptance Rate- FAR), Yanlış Reddetme Oranı (False Rejection Rate- FRR), Eşit Hata Oranı (Equal Error Rate- EER), Yarı Toplam Hata Oranı (Half Total Error Rate- HTER) ve Alıcı İşletim Karakteristiği Eğrisi Altındaki Alan (Area Under the Receiver Operating Characteristic- AUROC) gibi performans ölçütleri yardımıyla değerlendirilir. Metriklerin seçimi, uygulamanın özel gereksinimlerine ve değerlendirme için kullanılan veri kümesine bağlıdır. Bu metrikler, bir yüz sahteciliği tespit sisteminin performansı hakkında farklı bakış açıları sağlar. Bu çalışmada; FAR, FRR, EER, HTER, Kesinlik, Duyarlılık, F1-Skoru ve AUROC metrikleri kullanılmıştır.

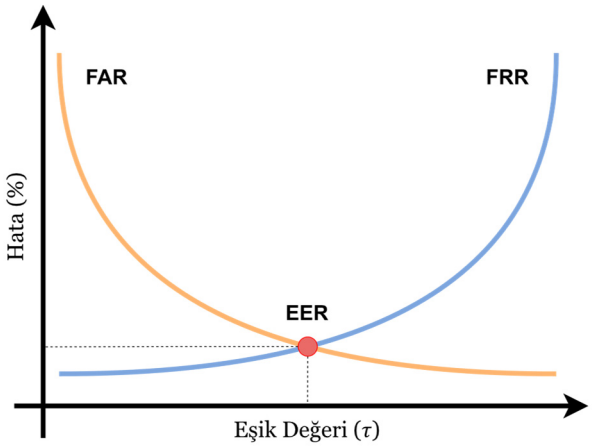
Yanlış Kabul Oranı (FAR): Sistemin bir sahtekarı (gerçek olmayan bir kullanıcı) yanlışlıkla gerçek bir kullanıcı olarak kabul etme oranıdır. Biyometrik sistemler bağlamında bu, yetkisiz bir kişinin yetkili bir kişi olarak yanlış tanımlanması anlamına gelir. Daha düşük bir FAR, yetkisiz kullanıcıların erişim kazanma olasılığının daha düşük olduğu anlamına geldiğinden, sistemde daha yüksek bir güvenlik seviyesine işaret eder.

Yanlış Reddetme Oranı (FRR): Sistemin gerçek bir kullanıcıyı yanlışlıkla reddetme oranıdır. Biyometrik sistemler bağlamında bu, yetkili bir kullanıcıyı tanımada başarısız olmak anlamına gelir. FRR, özellikle güvenliğin çok önemli olduğu uygulamalarda kritik bir metriktir. FRR ile FAR'ın dengelenmesi, biyometrik sistemin her iki hata türünü de özel kullanım durumu için gerekli olduğu ölçüde en aza indirmesini sağlar.

Eşit Hata Oranı (EER): Farklı biyometrik kimlik doğrulama yöntemlerinin doğruluk seviyelerini ölçmek ve karşılaştırmak için kullanılan bir performans ölçütüdür. EER, çaprazlama hata oranı olarak da adlandırılabilir. EER, FRR ve FAR değerlerinin grafiklerinin kesiştiği noktadan elde edilir (Teh vd., 2016).

EER oldukça önemli bir metriktir çünkü sistem performansının dengeli bir görünümünü sağlar. EER noktasında, yanlış kabul ve yanlış reddetmenin maliyeti eşit kabul edilir. Pratik anlamda bu, EER noktasında sistemin sahtekarları kabul ederken de gerçek kullanıcıları reddederken olduğu kadar çok hata yaptığı anlamına gelir.

Daha düşük bir EER, sistemin gerçek ve sahtekâr kullanıcıları ayırt etmede daha doğru olduğu anlamına geldiğinden, daha iyi bir sistem performansına işaret eder. Yetkisiz erişimi en aza indirmek ve meşru kullanıcılar için aşırı rahatsızlıktan kaçınmak arasında bir denge kurmaya yardımcı olduğu için hem güvenliğin hem de kullanıcı rahatlığının kritik olduğu senaryolarda önemli bir metriktir. Şekil 15, FAR ve FRR eğrilerinden EER değerinin tespitini göstermektedir.



Şekil 15. EER değerinin elde edilmesi

Yarı Toplam Hata Oranı (HTER): Bir YST sistemi temelde iki tür hatayla karşı karşıyadır. Bunlar; gerçek erişim reddi (yanlış ret) ve bir saldırının kabulüdür (yanlış kabul).

Bu sistemlerin performansı genellikle HTER metriği ile ölçülür. HTER, FAR ve FRR toplamının yarısıdır ve Denklem 7 kullanılarak hesaplanır.

$$HTER(\tau) = \frac{FAR(\tau) + FRR(\tau)}{2} \tag{7}$$

FAR ve FRR, τ eşik değerine bağlı olduğundan, FAR'ın azalması FRR'nin artmasına neden olur. Bu nedenle, sonuçlar genellikle farklı τ eşik değerleri için FAR'ın, FRR'ye göre değişimini gösteren grafiklerle temsil edilir. Geliştirme setinde EER'ye karşılık gelen noktada elde edilen τ eşik değeri kullanılarak test setinden HTER hesaplanır.

Kesinlik: Pozitif Tahmin Değeri (Positive Predictive Value- PPV) olarak da bilinen Kesinlik, gerçek pozitiflerin (True Positives- TP), gerçek pozitifler ve yanlış pozitiflerin (False Positives- FP) toplamına oranıdır. Kesinlik, Denklem 8 kullanılarak elde edilir:

$$Kesinlik = \frac{TP}{TP + FP} \tag{8}$$

Yüksek kesinlik, pozitif bir tahminin doğru olma ihtimalinin yüksek olduğunu gösterir.

Duyarlılık: İsabet Oranı veya Gerçek Pozitif Oranı (True Positive Rate- TPR) olarak da bilinen Duyarlılık, gerçek pozitiflerin gerçek pozitifler ve yanlış negatiflerin (False Negatives- FN) toplamına oranıdır. Duyarlılık, Denklem 9 kullanılarak elde edilir:

$$Duyarlılık = \frac{TP}{TP + FN} \tag{9}$$

Yüksek duyarlılık, gerçek pozitiflerin büyük bir kısmının doğru tahmin edildiğini gösterir. Sınıflandırma modellerinin performansını değerlendirmek için kullanılan bu iki metrik, özellikle dengesiz sınıfların veya yanlış pozitif/negatiflerin farklı etkileri olduğu durumlarda önemlidir.

Alıcı Operatör Karakteristik Eğrisi Altında Kalan Alan (AUROC), makine öğrenimi ve sınıflandırma görevlerinde yaygın olarak kullanılan bir metriktir. Yüz sahteciliği tespiti bağlamında AUROC, modelin farklı eşik değerlerinde gerçek ve sahte yüzler arasında ayrım yapma yeteneğini ölçer. ROC eğrisi, yanlış pozitif oranının (False Positive Rate- FPR)

gerçek pozitif oranına göre değişimini göstermektedir. Bu değerler, Denklem 10 kullanılarak elde edilir.

$$TPR = \frac{TP}{TP + FN}$$
$$FPR = \frac{FP}{TN + FP} \quad (10)$$

Denklemde TP; doğru tahmin edilen pozitif örnekleri, FP; yanlış tahmin edilen pozitif örnekleri, TN; doğru tahmin edilen negatif örnekleri ve FN; yanlış tahmin edilen negatif örnekleri temsil etmektedir. AUROC skorunun 1 olması, mükemmel bir sınıflandırıcıya işaret ederken, 0,5 olması rastgele bir sınıflandırıcıya işaret eder. Bu durum, özellikle dengesiz dağılıma sahip veri setlerinde, yüksek popülasyona sahip sınıfın tam olarak doğru tahmin edilmesine karşılık, düşük popülasyona sahip sınıfın tümüyle hatalı tahmini sebebiyle oluşacak yüksek sınıflandırma doğruluğunun, sistem başarısı hakkında hatalı karar verilmesinin önüne geçilmesinde önemlidir. Böyle bir durumda 0,5 olarak elde edilecek AUROC değeri, doğruluk değerinde belirtildiğinin aksine, sınıflandırma başarımının rastgele oluşunu ifade etmektedir.

Yüz sahteciliği tespitinde, daha yüksek bir AUROC, modelin gerçek ve sahte yüzleri ayırt etmede daha iyi olduğunu gösterir. Bu durum, önerilen modelin hem gerçek hem de sahte örnekleri doğru bir şekilde sınıflandırmada daha iyi bir yeteneğe sahip olduğu anlamına gelir.

1.9. Tezin Genel Yapısı

Bu tez, beş bölümden oluşmaktadır ve bölüm içerikleri aşağıda verilmiştir.

Bölüm 1: Araştırma alanının kapsamlı bir özetini içerir. Literatürdeki ilgili çalışmalar gözden geçirilmiş ve bu çalışmalar hakkında genel bilgi verilmiştir. Bu bölümde, önerilen modelin etkinliğini test etmek için kullanılan veri setleri açıklanmış ve performansı değerlendirmek için kullanılan değerlendirme metrikleri ayrıntılı olarak ele alınmıştır.

Bölüm 2: Giriş görüntülerinde ön işleme, yüz bölgesinin tespiti ve normalizasyon, doku özniteliklerinin çıkarımı, boyut indirgeme, sınıflandırma, derin öğrenme modelleri ve karma girişli model tasarımı ayrıntılı olarak sunulmaktadır.

Bölüm 3: Tüm deneylerin detayları, önerilen modeller ve YST başarımları, önerilen modellerin güncel yöntemlerle karşılaştırılması, zaman değerlendirmesine ait sonuçlar verilmiştir.

Bölüm 4: Önerilen tüm yöntemlerin sonuçları yer almaktadır.

Bölüm 5: Gelecekte yapılacak çalışmalar için önerileri içermektedir.

2. YAPILAN ÇALIŞMALAR

Bu bölümde, giriş görüntülerinin işlenmesi, renk kanalları, öznitelik çıkarma ve sınıflandırma ile ilgili bilgiler verilmektedir.

2.1. Görüntü Ön İşleme

Gerçek ortamda yüz sahteciliğinin nasıl gerçekleştirileceği tam olarak bilinmediğinden, giriş görüntüsünden yüz bölgesinin tespit edilmesi, hizalanması ve yeniden boyutlandırılması, tam otomatik bir YST sisteminin performansını önemli ölçüde etkiler. Bu sebeple, kullanılacak modelin eğitiminden önce, öznitelikleri çıkarılacak giriş görüntülerinin bir ön işlemden geçirilmesi gerekmektedir.

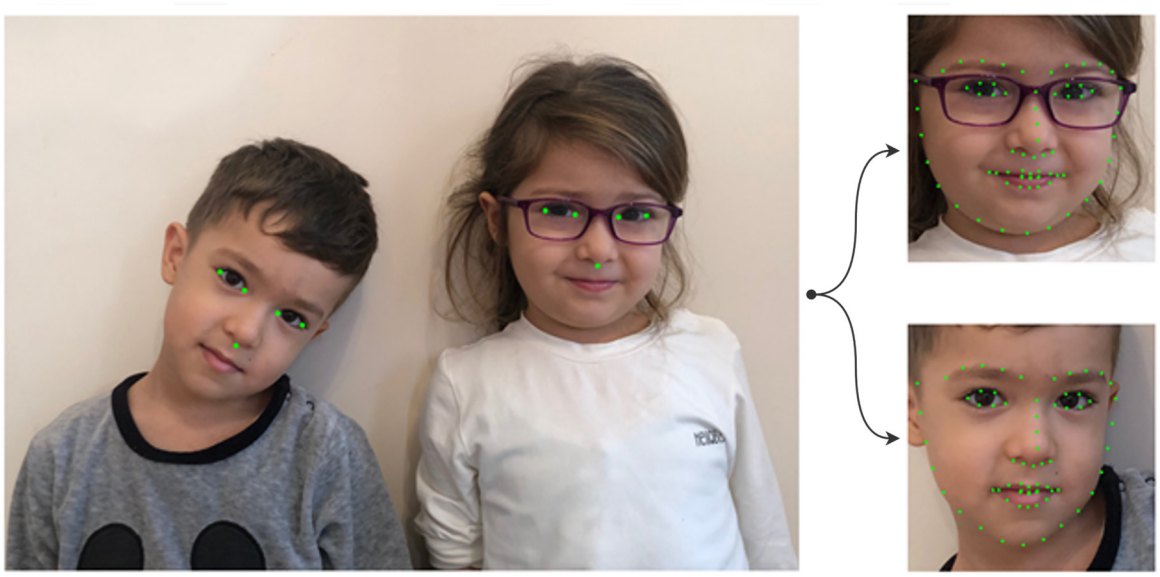
Yüz sahteciliklerinin tespitine yönelik geliştirilen bir modelde, yüz bölgesinin tespiti oldukça önemlidir. Literatürde yaygınca kullanılan veri setlerinden bazıları, giriş görüntüsü üzerinde bulunan yüz bölgelerinin koordinatlarını içermektedir. Ancak tüm veri setleri için bu bilgilerin sağlandığı söylenemez. Örneğin, NUAA seti; ham giriş görüntüsü, yüz bölgesi kırpılmış giriş görüntüsü ve yüz bölgesinin 64×64 piksel boyutlarına indirgenerek normalize edilmiş giriş görüntüsü olmak üzere üç ayrı formatta sunulmuştur. REPLAY-ATTACK seti ise tek formatta sunulmuş olmasına rağmen, görüntüler içerisindeki yüz konumları geliştiriciler tarafından sağlanmıştır. Ancak CASIA veri seti için bu verilerin hiçbiri sunulmamıştır. Bu bağlamda, tam otomatik bir yüz sahteciliği tespit sisteminin geliştirilmesinde, yüz bölgesinin tespit edilmesi bir gereklilik haline gelmektedir.

Bu çalışmada, giriş görüntülerinden yüz bölgelerinin tespitinde ve normalizasyonunda, Dlib kütüphanesinin (King, 2009) önceden eğitilmiş 5 noktalı yüz işaretçisi kullanılmıştır.

Dlib; makine öğrenimi, bilgisayarla görme ve görüntü işlemede çeşitli görevler için araçlar ve algoritmalar sağlayan, C++ ile yazılmış açık kaynaklı bir yazılım kütüphanesidir. Algoritma, giriş görüntülerinde yüz bölgesinin tespiti ve hizalama amacıyla önceden eğitilmiş 5 noktalı yüz işaretleyicisinden faydalanır. Bu yüz işaret algılama modeli, bir insan yüzündeki sol göz köşesi, sağ göz köşesi, burun ucu, sol ağız köşesi ve sağ ağız köşesinden oluşan beş önemli noktayı bulmayı hedefler. Bu noktalar, yüzün temel özelliklerini yakalamak için stratejik olarak konumlandırılmıştır. Özellikle yüz hizalama, duygu tanıma ve çeşitli yüz analizi uygulamaları gibi görevlerde bu bilgiler sıklıkla kullanılmaktadır.

2.2. Ağız, Burun ve Göz Bölgelerinin Tespiti

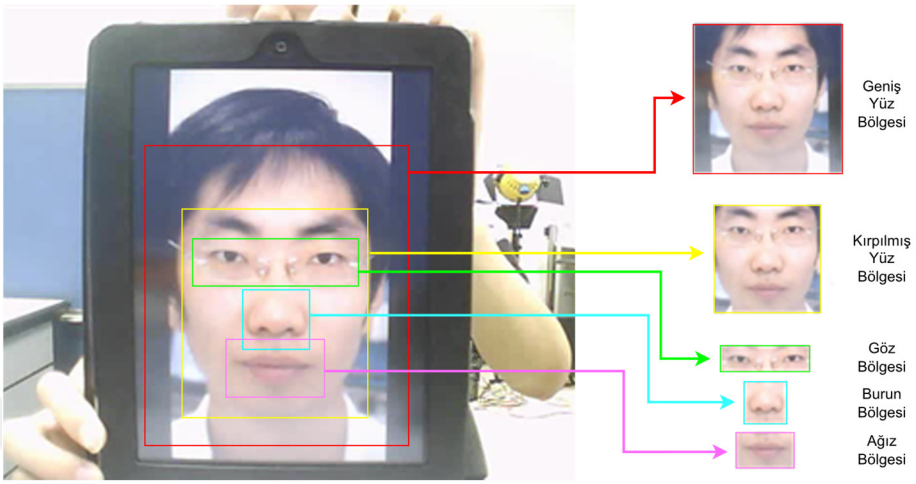
Bir YST probleminin çözümünde yüz bölgesi anahtar bileşendir ancak çalışmanın amaçlarından biri olan ağız, burun ve göz bölgelerinin YST problemine etkisinin belirlenmesi açısından bu bölgelerin elde edilmesi gerekmektedir. Bu aşamada yine Dlib kütüphanesinin sunduğu 68 noktalı yüz işaretleyicinden faydalanılmıştır. Bu yüz işaret algılama modeli, bir insan yüzündeki çene çizgisi noktaları, kaşlar, gözler, burun ve ağız bölgeleri de dahil olmak üzere 68 kilit noktayı bulma yeteneğine sahiptir. Şekil 16, bir giriş görüntüsünde bulunan yüz bölgelerinin tespitini, hizalanmasını ve 68 noktalı yüz işaretleyicisinin uygulanmasını ifade etmektedir.



Şekil 16. Yüz bölgelerinin tespiti ve 68 noktalı yüz işaretleyicisinin uygulanması

Şekil 16(a), farklı duruş pozisyonları içeren bir giriş görüntüsünde, 5 noktalı yüz işaretçisi yardımıyla yüz bölgelerinin tespitini temsil etmektedir. Burada amaç, kullanılan işaretçinin rotasyondan bağımsız olduğunu göstermektir. Şekil 16(b) ise, bir önceki aşamada tespit edilen yüz bölgelerinin normalizasyonu ve bu bölgelere uygulanan 68 noktalı yüz işaretçisini temsil etmektedir. İlk durumda birbirinden farklı konumlarda olan yüz bölgeleri, ağız, burun ve göz noktaları benzer konumlara gelecek şekilde hizalanmıştır. Bu durum, özellikle elle çıkarılan doku özniteliklerine dayalı bir modelin eğitimi ve başarımı için oldukça önemlidir. İkinci durumda ise hizalanmış görüntülere uygulanan yüz maskesi yardımıyla bölgelerin elde edilmesi sağlanmıştır. Şekil 17, CASIA-FASD veri setine ait

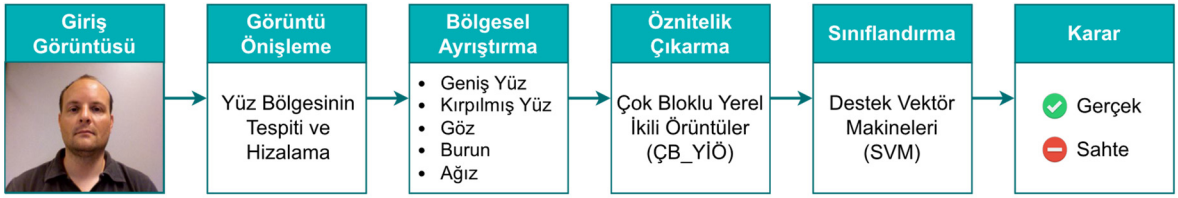
örnek giriş görüntüsünden yüz maskesi yardımıyla üretilen ve YST başarımları değerlendirilen "Geniş Yüz, Kırpılmış Yüz, Ağız, Burun ve Göz" bölgelerini temsil etmektedir.



Şekil 17. Giriş görüntüsünden elde edilen yüz bölgeleri

2.3. Yüz Bölgeleri ile Yüz Sahteciliği Tespiti

Covid-19 nedeniyle önceki birkaç yılda hayatımızın ayrılmaz parçaları haline gelen cerrahi maskeler, yüz tanıma sistemlerinin performansını doğrudan etkilemektedir. Bu nedenle, maske kullanımının zorunlu olduğu durumlarda, maske dışında kalan yüz bölgesi üzerinden yüz tanıma sistemlerinin geliştirilmesi kaçınılmazdır. Yüz görüntüsünün belirli bir bölgesinden tanıma gerçekleştiren sistemler, yine bu bölgeler üzerinden sahteciliğe uğrayacaktır. Bu nedenle bu çalışmada çeşitli yüz bölgelerinin (geniş yüz, kırpılmış yüz, ağız, burun ve göz) YST performansları araştırılmıştır. Şekil 18, Önerilen sistem modelinin genel yapısını içermektedir.



Şekil 18. Yüz bölgeleri ile YST sistemi

Önerilen yöntemde, giriş görüntülerinde yüz bölgeleri tespit edilmiş ve bu bölgeler hizalanmıştır. Ardından hizalanmış görüntüdeki geniş yüz, kırpılmış yüz, ağız, burun ve göz bölgeleri tespit edilmiş ve bölgesel ayrıştırma yapılmıştır. Tüm bu bölgeler, $1 \times 1, 2 \times 2$ ve 4×4 alt bölgelere ayrıştırılmış ve Bölüm 2.4.3'te detaylı açıklaması verilen Çok Bloklu Yerel İkili Örüntüler (ÇB_YİÖ) algoritması ile tüm bu alt bölgelerden öznitelikler çıkarılmıştır. Elde edilen veri kümesi boyutu TBA kullanılarak küçültülmüştür. Son aşamada, öznitelik kümesi SVM ile sınıflandırılmış ve yüz bölgelerinin YST problemine bireysel etkileri araştırılmıştır.

2.4. Doku Öznitelikleri ile Yüz Sahteciliği Tespiti

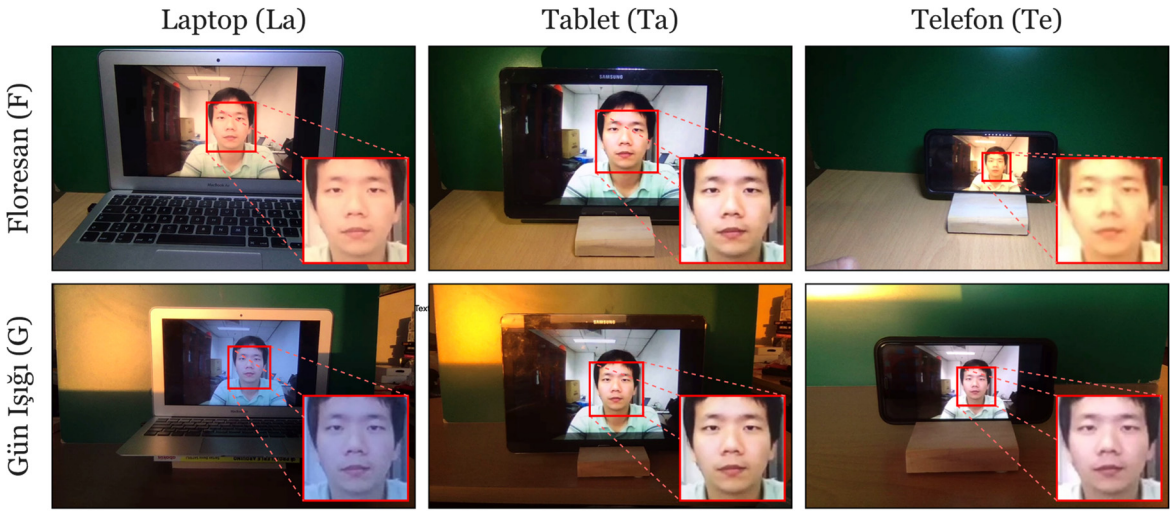
2018 yılına kadar gerçekleştirilen çoğu sahtecilik karşıtı çözüm doku analizi tabanlıdır (Boulkenafet vd., 2018) ancak günümüzde derin öğrenme tabanlı yaklaşımların yaygınlaşması sebebiyle bu alanda yapılan çalışmalar daha çok yerini derin öğrenme algoritmalarına bırakmıştır. Bunun yanında, kolay uygulanabilirlikleri ve düşük hesaplama maliyetleri, araştırmacıların bu alanda çalışmalar yürütmesine en önemli sebeplerden gösterilebilir.

Doku özniteliklerine dayalı YST uygulamalarında temel amaç, giriş görüntülerinden elle çıkarılan öznitelikler yardımıyla YST gerçekleştirmektir. Bu alanda yapılan ilk çalışmalar, özellikle kullanılan doku çıkarma algoritmalarının kısıtlamaları sebebiyle gri renk uzayında gerçekleştirilmiştir. Ancak bu görüntüler, çoğunlukla RGB renk uzayındaki renkleri kodlayan aygıtlar tarafından toplanmaktadır. Bununla birlikte, farklı spesifik özelliklere sahip birçok renk uzayı vardır ve sınıflandırma performanslarının, sınıflandırıcının içinde çalıştığı renk uzaylarının seçimine bağlı olduğu bilinmektedir. Her bir renk alanı belirli fiziksel, fizyolojik ve psiko-görsel özellikleri hesaba katar ancak hiçbir renk uzayı için "tüm veri setlerinin ayırt edilmesinde etkilidir" demek mümkün değildir (Porebski vd., 2013).

Birçok araştırmacı, renk dokusu sınıflandırmasında performans iyileştirmek için “en iyi” renk uzayını belirlemeye çalışmaktadır. Bunun için farklı renk uzaylarında tüm renk kanalları için kodlanan görüntülere sınıflandırma işlemi uygulanır ve her biri ile ulaşılan performanslar karşılaştırılır. Bu yaklaşım sonucunda en iyi sınıflandırma doğruluğunu sağlayan renk uzayı, problemin çözümü için seçilir. Ancak yukarıda da belirtildiği gibi dikkate alınan doku öznitelikleri ne olursa olsun, tüm veri setleri için ayırt edici tek bir renk uzayının belirlenmesi mümkün değildir. Örneğin, 2016 yılında yapılan bir çalışma (Sandid & Douik, 2016), en iyi renk uzayı seçiminin, kullanılan veri setine ve dolayısıyla dikkate alınan uygulamaya bağlı olduğunu ortaya koymaktadır. Uygulamada, BarkTex (Lakmann, 1998) ve OuTex-13 (Ojala, Maenpaa, vd., 2002) veri setlerinin renk dokularını sınıflandırmak için aynı yöntem ve öznitelikler kullanılmış ve en iyi sonuçların, bu iki veri setinin her biri için farklı renk uzayları kullanılarak elde edildiği belirtilmiştir. Bu durum, en iyi renk uzayının, dikkate alınan uygulama ve yaklaşıma bağlı olduğu hipotezini doğrulamaktadır (Porebski vd., 2013).

YST yöntemlerinin performansları genellikle standart veri setleri üzerinde değerlendirilir. Bu veri setleri, benzer koşullar altında sınırlı sayıda cihazla çekilen görüntü veya videolardan oluşur. Gerçek dünyadaki yüz saldırılarında ise çeşitli cihazlar kullanılmakta ve saldırılar farklı aydınlatma koşullarında gerçekleşmektedir. Literatürde renk doku analizi içeren çalışmalarda RGB, HSV ve YC_bC_r gibi cihaza bağlı renk uzaylarının sıklıkla kullanıldığı görülmektedir. Bu uzayların içerdiği kanallar, kullanılan donanım, parametreler ve görüntüleme cihazlarına göre değişiklik göstermektedir. Girdi görüntüsünün cihaza bağlı uzaylara dönüştürülmesi, cihaz ve görüntüleme koşullarına bağlı olarak oldukça farklı görüntüler elde edilmesine ve buna bağlı olarak YST sisteminin düşük performans göstermesine neden olabilmektedir. Bununla birlikte, cihazdan bağımsız görüntüleme sağlayan renk uzayları, elde edildikleri cihazın özelliklerinden etkilenmemektedir. Dolayısıyla YST sisteminin performansına katkı sunacakları düşünülmüş ve çalışmada, farklı giriş cihazları kullanılarak oluşturulan veri tabanları olması sebebiyle, giriş cihazından etkilenmeyen ek bir renk uzayının kullanılmasının avantajlı olacağı düşünülmüştür.

Bu durumu analiz etmek için CASIA-FASD veri setinden alınan gerçek bir görüntü, iki farklı aydınlatma koşulunda (gün ışığı ve floresan lamba) üç farklı cihaz (telefon, tablet ve dizüstü bilgisayar) yardımıyla kameraya sunularak sahte görüntüler oluşturulmuştur. Şekil 19, oluşturulan saldırı görüntülerini göstermektedir.



Şekil 19. Renk uzayı kanallarının benzerlikleri

Üretilen altı sahte görüntü RGB, HSV ve L*a*b* renk uzaylarına dönüştürüldükten sonra, her renk uzayındaki bileşenlerin yapısal benzerlik indeksi (Structural Similarity Index Measure- SSIM) (Z. Wang vd., 2003) Denklem 11 kullanılarak hesaplanmıştır.

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (11)$$

Burada μ_x: x' in ortalaması, μ_y: y' nin ortalaması, σ_x^2: x' in varyansı, σ_y^2: y' nin varyansı, σ_{xy} x ve y' nin kovaryansıdır. c_1 ve c_2: zayıf paydalı bölmeyi dengeleyen katsayılardır. Telefon (Te) ekranından gün ışığında (G) elde edilen görüntüye (TeG) ait renk kanallarının, diğer görüntülerin renk kanalları ile yapısal benzerlik indeks metrikleri hesaplanmıştır. Tablo 2, elde edilen yüzdelik benzerlikleri ifade etmektedir.

Tablo 2. Renk kanalları arası yapısal benzerlikler

	Telefon - Gün Işığı (TeG)								
Atak Türü	**HSV**			**L*a*b***			**RGB**		
	H	**S**	**V**	**L***	**a***	**b***	**R**	**G**	**B**
TeF	20,89	64,79	77,26	87,42	92,61	95,00	76,94	87,73	86,55
TaG	22,12	68,85	84,48	85,36	96,93	95,91	84,38	85,08	82,48
TaF	24,72	67,55	81,52	83,09	96,24	94,20	81,14	82,78	81,88
LaG	42,06	63,73	81,56	82,81	97,75	94,92	75,27	81,80	81,81
LaF	20,32	67,87	82,35	87,85	96,63	95,92	82,28	88,34	84,38
Standart Sapma	9,12	2,18	2,63	2,35	1,99	0,73	3,79	2,90	2,04

Tabloda görüldüğü üzere, L*a*b* renk uzayında yer alan kanallardan elde edilen görüntüler, farklı cihazlarda görüntülenseler veya farklı aydınlatma koşullarında elde edilseler dahi diğer renk uzaylarından elde edilen görüntülere göre benzerlik göstermektedir. Bu durum, özellikle değişken ışıklandırma koşulları altında farklı giriş cihazları ile elde edilen görüntülerde sahteciliklerin belirlenmesinde, saldırılarda kullanılan cihaza bağımlı olmayan ve gerçek dünya senaryolarına daha uygun bir YST sistemi geliştirilmesi açısından L*a*b* renk uzayının kullanılmasının avantajlı olacağı düşüncesinin oluşmasında etkili olmuştur.

2.4.1. Renk Uzayları

Renk uzayları genel olarak cihaza bağımlı ve cihazdan bağımsız olmak üzere iki gruba ayrılabilir (Bora vd., 2015). Cihaza bağlı bir renk uzayında üretilen renkler hem kullanılan parametrelere hem de görüntüleme için kullanılan cihaza bağlıdır. RGB, CMYK ve HSV bu renk uzayının en yaygın örnekleridir. Cihazdan bağımsız bir renk uzayında ise, bir dizi parametre hangi donanım kullanılırsa kullanılsın aynı rengi üretmektedir. Bu renk uzaylarından L*a*b*, yaygın olarak kullanılmaktadır. Bu çalışmada RGB, HSV, YC_bC_r ve L*a*b* renk uzayları kullanılmıştır.

A. RGB

RGB (Kırmızı, Yeşil, Mavi) renk uzayı, dijital cihazlarda geniş bir renk yelpazesi oluşturmak için kullanılan temel bir yöntemdir. Işığın üç ana rengini bir araya getirerek yoğunluklarını ayarlamak suretiyle çeşitli tonlar oluşturma prensibiyle çalışır. Bu renk modelinde her bir kanal tipik olarak 8 bitlik bir aralıkta çalışır ve her renk için 256 farklı yoğunluk seviyesine izin verir. RGB renk uzayı dijital ekranlarda, bilgisayar grafiklerinde, dijital kameralarda ve renklerin ışık kullanılarak oluşturulduğu diğer çeşitli cihazlarda yaygın olarak kullanılır.

Bu renk uzayı; renk algılama, temsil ve görüntüleme için en sık kullanılan renk uzayıdır. Bununla birlikte, üç renk bileşeni (kırmızı, yeşil ve mavi) arasındaki yüksek korelasyon, parlaklık ve kromatik bilginin tam ayrıştırılamaması nedeniyle görüntü analizindeki uygulaması oldukça sınırlıdır (Boulkenafet vd., 2015).

B. HSV

HSV renk uzayının üç boyutlu temsili, merkezi dikey eksenin yoğunluğu temsil ettiği bir hekzatondur. Ton, kırmızı eksene göre 0 açısında kırmızı, $2\pi/3$'te yeşil, $4\pi/3$'te mavi ve

yine 2π'de kırmızı olan [0, 2π] aralığında bir açı olarak tanımlanır. Doygunluk, rengin derinliği veya saflığıdır ve merkezde 0 ile dış yüzeyde 1 arasındaki bir değerle merkezi eksenden radyal mesafe olarak ölçülür. S=0 için, yoğunluk ekseni boyunca daha yükseğe çıkıldıkça, çeşitli gri tonlarında siyahtan beyaza gidilmektedir (Z. K. Huang & Liu, 2007). RGB renk uzayından HSV renk uzayına dönüşüm için Denklem 12 kullanılmaktadır.

$$f(x) = \begin{cases} 2 - cos^{-1}\left\{\dfrac{[(R-G)+(R-B)]}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right., & B < G \\ cos^{-1}\left\{\dfrac{[(R-G)+(R-B)]}{2\sqrt{(R-G)^2+(R-G)(G-B)}}\right., & aksi\ halde \end{cases} \tag{12}$$

$$S = \frac{V - \min(R,G,B)}{V}$$

$$V = \max(R,G,B)$$

C. YC_bC_r

YC_bC_r renk uzayı, video bilgilerinin işlenmesinde dijital algoritmalara yönelik artan taleplere yanıt olarak tanımlanmış ve dijital videolarda yaygın olarak kullanılan bir model haline gelmiştir. Parlaklığı temsil eden Y bileşeni RGB değerlerinin ağırlıklı toplamı olarak elde edilir. C_b ve C_r ise sırasıyla mavi ve kırmızı renk bileşenlerinin parlaklık bileşeni ile farkını ifade etmektedir (Vezhnevets vd., 2003). RGB renk uzayından YC_bC_r renk uzayına dönüşüm için Denklem 13 kullanılmaktadır.

$$\begin{aligned} \mathrm{Y} &= 0.299\mathrm{R} + 0.587\mathrm{G} + 0.114\mathrm{B} \\ \mathrm{C_b} &= 128 - 0.168736\mathrm{R} - 0.331264\mathrm{G} + 0.5\mathrm{B} \\ \mathrm{C_r} &= 128 + 0.5\mathrm{R} - 0.418688\mathrm{G} - 0.081312\mathrm{B} \end{aligned} \tag{13}$$

D. L*a*b*

L*a*b* renk uzayındaki L* kanalı parlaklığı temsil etmekte ve gri tonlarına karşılık gelen ve 0-100 arasında değişen değerler içermektedir. a* ve b* kanalları ise -128 ile +127 arasında değerlere sahip olup sırasıyla kırmızı-yeşil oranını ve sarı-mavi oranını temsil etmektedir. Bu nedenle, a* veya b* kanalındaki yüksek bir değer, daha yüksek oranda kırmızı veya sarı olan bir rengi gösterirken, düşük bir değer, daha yüksek oranda yeşil veya mavi olan bir tonu belirtir (Murali & Govindan, 2013).

L*a*b* renk uzayı, algılanabilir tüm renkleri içerir. Modelin en önemli özelliklerinden biri cihaz bağımsızlığıdır. Bu, renklerin yaratılışlarından veya üzerinde görüntülendikleri cihazdan bağımsız olarak tanımlandığı anlamına gelir. Ayrıca, cihaz bağımsızlığı açısından farklı cihazlar arasında bir değişim formatı olarak kullanılır.

L*a*b* renk uzayı, insan görüş sistemine göre tasarlanmıştır. a* ve b* bileşenlerindeki çıktı eğrileri değiştirilerek renk dengelemek, L bileşenini kullanarak kontrastı ayarlamak mümkündür. İnsan görsel algısından ziyade fiziksel cihazların çıktısını modelleyen RGB veya CMYK uzaylarında bu dönüşümler ancak bir görüntü düzenleme uygulamasındaki uygun karışım seçenekleri yardımıyla yapılabilir (Baldevbhai & Anand, 2012).

2.4.2. Yerel İkili Örüntüler

Ojala ve ark. (Ojala, Pietikäinen, vd., 2002) tarafından önerilen YİÖ, bir görüntünün her pikselini, komşu pikselleriyle eşikleyerek etiketleyen ve sonucu ikili bir sayı olarak üreten basit ve verimli bir doku operatörüdür. YİÖ kodları Denklem 14 kullanılarak üretilir.

$$Y\dot{I}\ddot{O}_{P,R}(x_c) = \sum_{p=0}^{P-1} u(x_p - x_c) \times 2^p$$
$$u(y) = \begin{cases} 1, & y \geq 0 \\ 0, & y < 0 \end{cases} \tag{14}$$

Denklemde x_c, x_p, R ve P sırasıyla merkez pikseli, merkez pikselin komşularını, komşuların merkez piksele olan uzaklığını ve işleme sokulan komşu sayısını ifade etmektedir. Şekil 20, bir piksele karşılık gelen YİÖ kodunun üretilme sürecini göstermektedir.

Şekil 20. YİÖ kodlarının üretilmesi

Üretilen YİÖ kodları, düzgün ve düzgün olmayan olmak üzere iki başlık altında değerlendirilir. Düzgün örüntüler ($Y\dot{I}\ddot{O}_{8,1}^{U2}$), ikili YİÖ kodunda 0-1 geçişinin iki veya daha az olduğu durumları ifade etmektedir. Geçiş işlemi, her 0→1 ve 1→0 değişimlerini ifade etmektedir. Örnekle, 00000000 ve 11111111 dizileri 0 geçişe, 00110000, 11000011 ve 11110000 örüntüleri 2 geçişe sahip olduğundan, bu örüntüler düzgündür. Ancak, sırasıyla 3, 4, 5 ve 6 geçişe sahip olan 00110011, 00110100, 01001101 ve 01010010 örüntüleri düzgün değildir.

Görüntü doku bilgisini temsil etmek için, ikili kodların görüntüde meydana gelme sayıları bir histogramla gösterilir. Ojala ve ark. (Ojala, Pietikäinen, vd., 2002), düzgün yerel ikili örüntülerin, yerel görüntü dokusunun temel özelliklerini taşıdığını ve bunlardan üretilen histogramın, görüntüyü tanımlamada oldukça güçlü olduğunu belirtmiştir. Düzgün örüntüler görüntüdeki, kenarlar, köşeler ve noktalar gibi mikro özniteliklere karşılık geldiği için doku bilgisini daha çok taşımaktadır.

YİÖ histogramı çıkarılırken ($Y\dot{I}\ddot{O}_{8,1}^{U2}$) kullanıldığında, histogramda her bir düzgün örüntü için bir bölme bulunurken, düzgün olmayan örüntülerin hepsi tek bir bölmede toplanmaktadır. Tüm örüntüler incelendiğinde $P = 8$ komşuluk için $2^P = 256$ faklı örüntü üretilebileceği ve bunlardan sadece 58 tanesinin düzgün olduğu görülmektedir.

$I(x, y)$ görüntüsü için R piksel uzaklıktaki P komşu için YİÖ histogramı Denklem 15 kullanılarak üretilir.

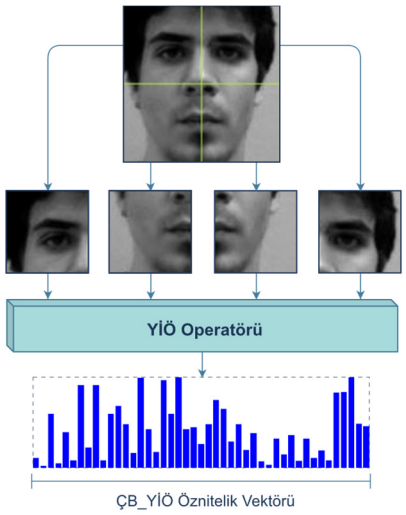
$$H_i = \sum_{i = 0,1,\dots,n-1} f\{Y\dot{I}\ddot{O}_{P,R}(x_c) = U(i)\} \qquad (15)$$

$$f(y) = \begin{cases} 1, & \text{eğer } y \text{ doğru ise} \\ 0, & \text{eğer y yanlış ise} \end{cases}$$

Histogram üretiminde sadece düzgün örüntüler dikkate alındığında $n = 58$ olacaktır. Üretilen özniteliklerden düzgün olanlar ise $U(i)$ dizisinde saklanmaktadır. Düzgün olmayan örüntüler bir arada tutulduğundan, (8,1) komşuluk için üretilen YİÖ histogramının boyutu 59 olur. Tüm örüntülerin ($Y\dot{I}\ddot{O}_{8,1}$) kullanıldığı durumda ise $n = 256$ olacağından, YİÖ histogramının boyutu da 256 olacak ve böylece her örüntü, kendisine karşılık gelen histogram bölmesinde saklanacaktır. Bu çalışmada hem düzgün örüntülerin hem de tüm örüntülerin YST problemine etkisi araştırılmıştır.

2.4.3. Çok Seviyeli Yerel İkili Örüntüler

YİÖ operatörü, güçlü doku tanımlama yeteneğinin yanında basit ve hızlı hesaplanması sebebiyle tercih edilmektedir ancak bazı problemlerde daha detaylı özniteliklere ihtiyaç duyulmaktadır. Bu durumda, Çok Bloklu YİÖ (ÇB_YİÖ) operatörü kullanılır. ÇB_YİÖ, giriş görüntüsünün $n \times m = T$ adet alt bölgeye ayrılması ve her bölgeden elde edilen düzgün YİÖ histogramlarının uç uca eklenmesi esasına dayanır. Sonuç olarak tüm bloklardan üretilen YİÖ öznitelik histogramları yan yana eklenerek $T \times Vb$ elemanlı ÇB_YİÖ öznitelik vektörü elde edilir. İfade içerisindeki Vb değeri, kullanılan örüntü türüne ait vektör boyutunu temsil etmektedir ve bu değer düzgün örüntüler için 59, tüm örüntüler için ise 256'dır. Şekil 21, giriş görüntüsünün 2×2 bölgeye ayrılması sonrası ÇB_YİÖ histogramının elde edilişini ifade etmektedir.



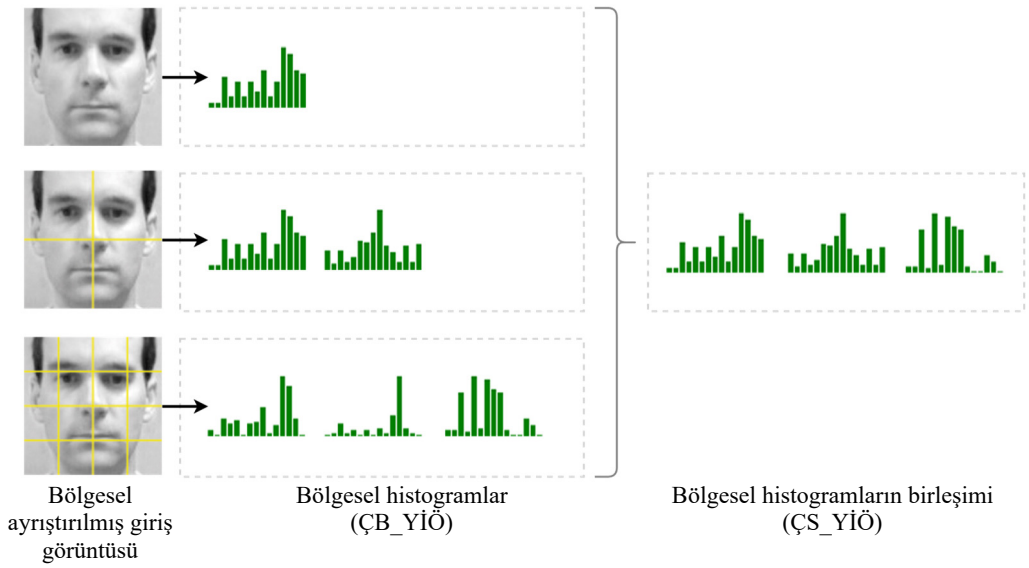
Şekil 21. Çok bloklu YİÖ özniteliklerinin elde edilmesi

Çok Seviyeli YİÖ (ÇS_YİÖ) öznitelikleri ise, bir giriş görüntüsünden farklı seviyelerde ÇB_YİÖ özniteliklerinin üretilmesi ve birleştirilmesi ile elde edilen daha detaylı bir doku tanımlama yöntemidir. Burada, giriş görüntüsü $1 \times 1, 2 \times 2$ ve 4×4 gibi N adet alt bölgeye ayrılır ve bu bölgelerin her birinden ÇB_YİÖ öznitelikleri çıkarılarak uç uca eklenir. Sonuç olarak tüm seviyelerdeki tüm bloklara karşılık gelen YİÖ histogramları uç uca birleştirilerek, ÇS_YİÖ öznitelik vektörü elde edilir. Bu vektörün uzunluğu Denklem 16 kullanılarak hesaplanır.

$$ÇS\_YİÖ_{uzunluk} = \sum_{i=1}^{N} T_i \times Vb \tag{16}$$

$$T = (1 \times 1, \quad 2 \times 2, \quad 4 \times 4, \quad \ldots\ldots, \quad n \times m)$$

Denklemde n ve m, giriş görüntüsünün satırda ve sütunda ayrıştırıldığı bölge sayısını, Vb ise vektör boyutunu temsil etmektedir. Şekil 22, ÇB_YİÖ histogramlarının birleşiminden elde edilen ÇS_YİÖ histogramını ifade etmektedir.



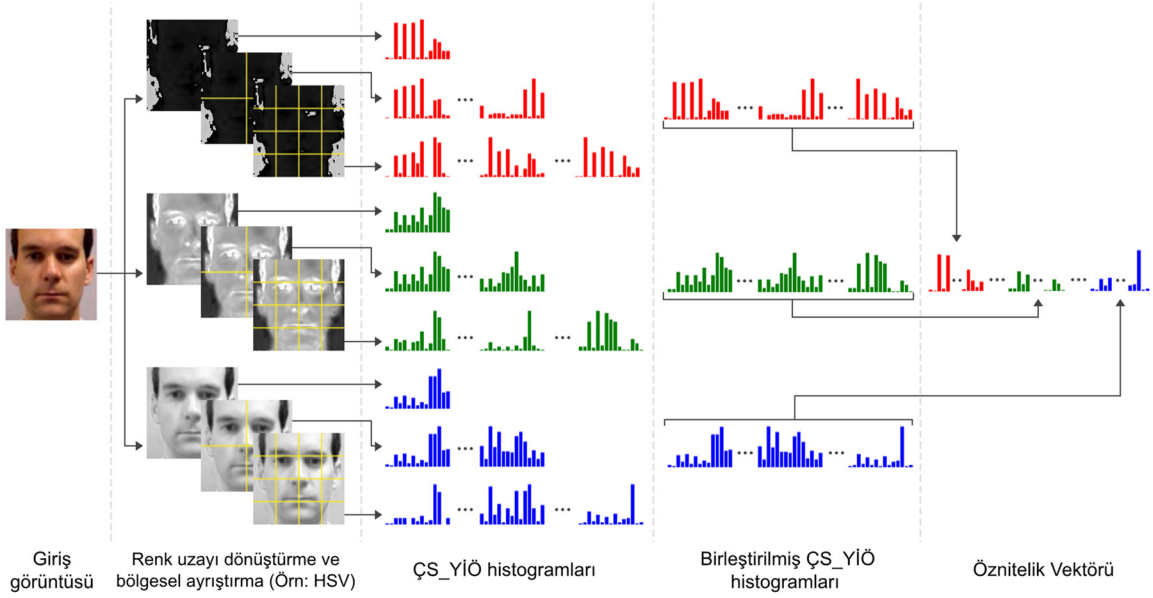
Şekil 22. ÇS_YİÖ öznitelik vektörünün üretimi

2.4.4. Çok Renkli Çok Seviyeli Yerel İkili Örüntüler

Yapılan araştırmalar, farklı renk kanallarının YST probleminin çözümünde daha fazla bilgi taşıdığını ortaya koymuştur. Bu amaçla, çalışmada giriş görüntüleri renk kanallarına

ayrılmış, her kanaldan elde edilen ÇS_YİÖ öznitelikleri uç uca eklenerek, giriş görüntüsünü temsil edecek öznitelik vektörü elde edilmiştir. Bu işlem birden fazla renk uzayındaki kanallara uygulanıp uç uca eklenerek Çok Renkli ÇS_YİÖ (ÇRÇS_YİÖ) öznitelik vektörü üretilmiştir.

ÇRÇS_YİÖ histogramları, giriş görüntüsünün aynı bölgeleri için farklı renk kanallarından üretilen bilgileri barındırmaktadır. Bu durum, özellikle bir renk kanalından edinilen bilginin yetersiz olduğu durumlarda farklı kanallardan üretilen ek bilgiler yardımıyla YST probleminde daha detaylı bir analiz yapmaya olanak sağlar. Şekil 23, renk kanallarından üretilen ÇS_YİÖ histogramlarını ifade etmektedir.



Şekil 23. Renk kanallarından üretilen ÇS_YİÖ

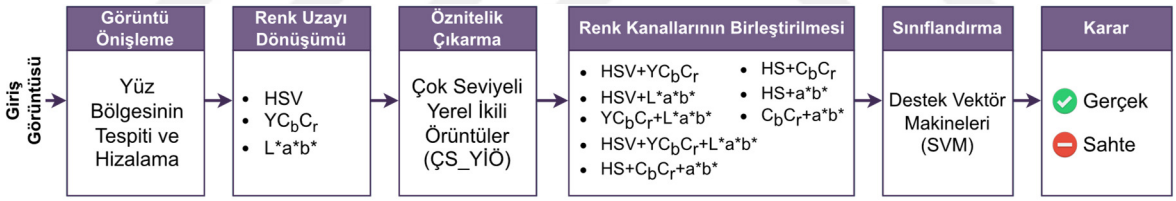
ÇRÇS_YİÖ histogramlarının üretiminde giriş görüntüsü öncelikle bir renk uzayına dönüştürülmektedir (Örn: HSV). Ardından, renk uzayının her kanalı ayrı bir görüntü olarak ele alınır ve bu görüntüler üzerinde bölgesel ayrıştırma işlemi uygulanır. Bu işlem sonucunda üretilen ÇB_YİÖ histogramları uç uca eklenerek üretilen ÇS_YİÖ histogramları, kendi aralarında uç uca eklenerek birleştirilmiş ÇS_YİÖ histogramları elde edilir. Bu histogramların her biri, renk kanallarını ($ÇS\_YİÖ_H$, $ÇS\_YİÖ_S$, $ÇS\_YİÖ_V$) temsil etmektedir. Son olarak üretilen tüm histogramlar uç uca eklenir ve ilgili renk uzayı için giriş vektörü elde edilir. Bu işlemin birden çok renk uzayı için tekrarlanması sonucunda ise ÇRÇS_YİÖ histogramları oluşturulur. ÇRÇS_YİÖ histogramlarının boyutu, kullanılan renk uzayı sayısı, bu renk uzaylarının içerdiği kanal sayısı ve bölgesel ayrıştırma sayısına bağlı olarak

değişkenlik gösterir. Bir giriş görüntüsü için ÇRÇS_YİÖ histogramının uzunluğu Denklem 17 kullanılarak hesaplanır.

$$ÇRÇS\_YİÖ_{uzunluk} = \sum_{k=1}^{S}\sum_{j=1}^{S_C}\sum_{i=1}^{T} T_i \times Vb \qquad (17)$$

$$T = (1 \times 1, \quad 2 \times 2, \quad 4 \times 4, \quad \ldots\ldots, \quad n \times m)$$

Denklemde S ve S_C sırasıyla kullanılan renk uzaylarını ve her bir renk uzayının içerdiği kanal sayısını ifade etmektedir. Yukarıdaki denkleme göre, giriş görüntüsünün iki farklı renk uzayına dönüştürülmesi (Örn: HSV, L*a*b*) ve bu renk uzaylarının her kanalının (iki renk uzayı da üç kanal içermektedir) sırasıyla $1 \times 1, 2 \times 2$ ve 4×4 bölgelere ayrılması sonucunda elde edilecek öznitelik vektörü boyutu $S \times S_C \times \sum T \times Vb$ formülünden; $2 \times 3 \times (1 + 4 + 16) \times Vb = 126 \times Vb$ olarak hesaplanır. Şekil 24, HSV, YC_bC_r ve L*a*b* renk uzaylarının kullanıldığı renk kanalı tabanlı modele ait sistemin genel yapısını içermektedir.



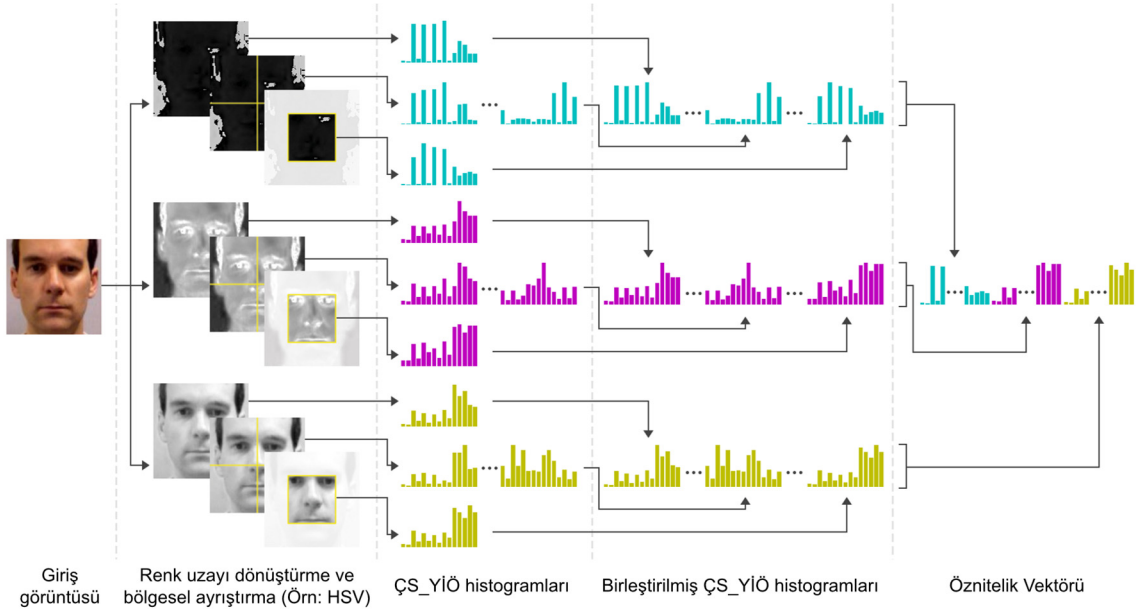
Şekil 24. Önerilen renk kanalı tabanlı modele ait sistemin genel yapısı

Önerilen yöntemde, giriş görüntülerindeki geniş yüz bölgeleri tespit edildikten ve bu bölgeler hizalandıktan sonra bu görüntüler HSV, YC_bC_r ve L*a*b* renk uzaylarına dönüştürülmüştür. Ardından, renk uzaylarındaki her renk kanalı için ÇS_YİÖ algoritması yardımıyla öznitelikler çıkarılmıştır. Bir sonraki aşamada, farklı renk kanallarından çıkarılan bu öznitelikler çeşitli kombinasyonlarla birleştirilmiş ve boyutları TBA kullanılarak küçültülmüştür. Son aşamada, elde edilen veriler SVM ile sınıflandırılmış ve farklı renk kanallarının birleşiminin, YST problemine etkisi incelenmiştir.

2.4.5. Yüz Ağırlıklı Çok Renkli Çok Seviyeli Yerel İkili Örüntüler

ÇRÇS_YİÖ öznitelikleri ayrıntılı bilgiler içerebilir, ancak eğitimde kullanılacak vektörlerin boyutunu da artırır. Bu da yüksek düzeyde hesaplama karmaşıklığı ve depolama maliyeti ile sonuçlanır.

Yüz görüntüsü bir YST sisteminin kritik bileşenidir ve bu bilginin verimli kullanımı performansı önemli ölçüde artırabilir. Bu tez çalışması kapsamında, giriş görüntüsünün YST problemine etkisini arttırmak için Yüz Ağırlıklı Çok Renkli Çok Seviyeli YİÖ (YA_ÇRÇS_YİÖ) öznitelik çıkarma yöntemi önerilmiştir. Önerilen yöntemde giriş görüntüsü 128×128 piksel boyutlarına ölçeklendirildikten sonra hedef renk uzayına dönüştürülür. Renk uzayının tüm kanalları sırasıyla 1×1 ve 2×2 alt bölgelere ayrıştırılır. Ardından bu bölgelerden YİÖ öznitelikleri çıkarılır ve uç uca eklenir. Son olarak, ağız, burun ve göz bölgelerinin etkisini arttırmak ve bölgesel ayrıştırmalarda oluşabilecek kaybı en aza indirmek amacıyla giriş görüntüsü, merkez noktasından 64×64 piksel boyutlarında kırpılır ve elde edilen yeni görselden YİÖ öznitelikleri çıkarılarak bir önceki gruba eklenir. Bu işlem, kullanılan diğer renk uzayları için tekrarlanır ve sonucunda, giriş görüntüsü için nihai YA_ÇRÇS_YİÖ öznitelik vektörü oluşturulur. Şekil 25, tek bir renk uzayı üzerinden önerilen öznitelik çıkarma yöntemine ait üretim sürecini temsil etmektedir.



Şekil 25. YA_ÇRÇS_YİÖ özniteliklerinin üretim süreci

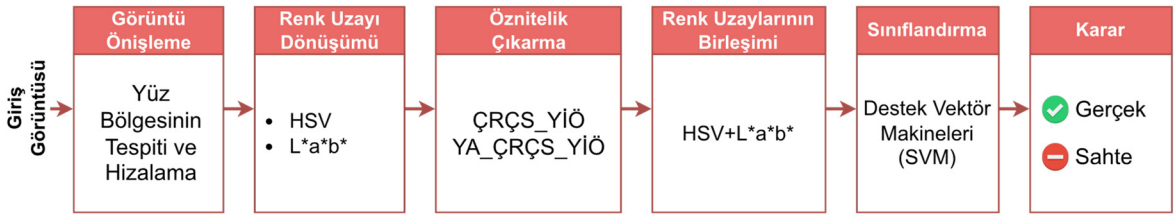
Önerilen bu yöntem, bir YST sistemi için ihtiyaç duyulan anahtar yüz bölgelerini içeren alanların etkisini artırmayı ve bununla birlikte, hesaplama karmaşıklığı ile depolama maliyetlerini en aza indirgemek için giriş vektör uzayının boyutunu azaltmayı amaçlamaktadır. YA_ÇRÇS_YİÖ histogramlarının boyutu, kullanılan renk uzayı sayısı ve bu renk uzaylarının içerdiği kanal sayısına bağlı olarak değişkenlik gösterir. Bir giriş görüntüsü için YA_ÇRÇS_YİÖ histogramının uzunluğu Denklem 18 kullanılarak hesaplanır.

$$YA\_ÇRÇS\_YİÖ_{uzunluk} = \sum_{k=1}^{S}\sum_{j=1}^{S_C}\left(\left(\sum_{i=1}^{2} T_i \times Vb\right) + Vb\right) \tag{18}$$

$$T = (1 \times 1, \quad 2 \times 2)$$

Yukarıdaki denkleme göre, giriş görüntüsünün iki farklı renk uzayına dönüştürülmesi (Örn: HSV, L*a*b*) ve bu renk uzaylarının her kanalının sırasıyla $1 \times 1, 2 \times 2$ ve merkez etrafından kırpılan 64×64 piksellik görüntünün de 1×1 bölgelere ayrılması sonucunda elde edilecek öznitelik vektörünün boyutu $S \times S_C \times (\sum T \times Vb + Vb)$ formülünden; $2 \times 3 \times ((1 + 4) \times Vb + Vb) = 36 \times Vb$ olarak hesaplanır. Bu vektör; YA_ÇRÇS_YİÖ$_H$, YA_ÇRÇS_YİÖ$_S$, YA_ÇRÇS_YİÖ$_V$, YA_ÇRÇS_YİÖ$_{L*}$, YA_ÇRÇS_YİÖ$_{a*}$ ve YA_ÇRÇS_YİÖ$_{b*}$ özniteliklerinin uç uca eklenmesiyle oluşturulmuştur.

ÇRÇS_YİÖ algoritmasında ise bu değer $126 \times Vb$ olarak hesaplanmıştır. Bu sonuç, önerilen YA_ÇRÇS_YİÖ yönteminin, ÇRÇS_YİÖ yöntemine oranla %71,43 daha düşük boyutlu olduğunu ortaya koymaktadır. Şekil 26, HSV ve L*a*b* renk uzaylarının kullanıldığı önerilen Yüz Ağırlıklı Çok Renkli Çok Seviyeli Yerel İkili Örüntüler modeline ait sistemin genel yapısını, Şekil 27, YA_ÇRÇS_YİÖ öznitelik çıkarma yöntemine ait sözde kodu içermektedir.



Şekil 26. Önerilen yüz ağırlıklı modele ait sistemin genel yapısı

```
Giriş: Görüntü img,
Çıkış: Öznitelik Vektörü f

I ← oku img
I_ch ← I görüntüsünü H,S,V,L*,a*,b* renk kanallarına ayrıştır
Döngü I_ch değişkenindeki her veriyi img_ch içerisinde tut:
    Her img_ch için Öznitelik vektörlerini tutacak fv değişkenini tanımla
    fv.ekle(img_ch için hesaplanan YİÖ öznitelikleri)
    patch_img_ch ← img_ch görselini 2 satır 2 sütuna böl
    Döngü patch_img_ch içerisindeki her veriyi patch içerisinde tut:
        fv.ekle(patch için hesaplanan YİÖ öznitelikleri)
    img_ch ← img_ch görselini, ortadan 64 x 64 piksel boyutlarında kırp
    fv.ekle(img_ch için hesaplanan YİÖ öznitelikleri)
f ← Tüm fv verilerini birleştir

Oluşan f vektörünü ana programa gönder.
```

Şekil 27. Yüz Ağırlıklı Çok Renkli Çok Seviyeli Yerel İkili Örüntüler algoritmasına ait sözde kod

2.5. Derin Öznitelikler ile Yüz Sahteciliği Tespiti

Yüz sahteciliği tespiti bağlamında derin öznitelikler, derin sinir ağları tarafından otomatik olarak öğrenilen yüz görüntülerinin üst düzey temsillerini ifade eder. Bu öznitelikler, yüzün gerçek mi yoksa sahte mi (örneğin basılı bir fotoğraf veya dijital bir ekran) olduğunu gösteren yüz görüntülerinin karmaşık desenlerini yakalar.

Derin öznitelikler, sahtekarlık saldırılarının tespit edilmesini sağladıkları ve yüz tanıma sistemlerinin sağlamlığına katkıda bulundukları için yüz sahtekarlığı sorununu ele almada çok önemlidir. Son araştırmalar, yüz sahteciliği saldırılarını tespit etmek için ayırt edici öznitelikler çıkarmak üzere derin öğrenme tekniklerinden yararlanmaya odaklanmıştır. Örneğin, farklı renk uzaylarından ve dokulardan öznitelikleri analiz ederek yüz sahteciliği saldırılarını tespit etmek için renk-doku tabanlı bir derin sinir ağı yaklaşımı önerilmiştir

(Rusia & Singh, 2022). Çalışma, NUAA, REPLAY-ATTACK, CASIA-FASD ve MSU-MFSD gibi çeşitli halka açık karşılaştırmalı yüz sahteciliği veri kümelerinden öznitelikleri analiz etmek için makine öğrenimi ve derin öğrenme tabanlı yaklaşımların kullanımını vurgulamaktadır. Ayrıca yapılan bir başka çalışmada, farklı veri kümelerinde yüz tanıma için güçlü yüz özelliği temsillerini öğrenmede derin öğrenme algoritmalarının avantajları gösterilmiştir (Zhu vd., 2020). Çalışma, derin öğrenmenin çok sayıda veri örneğinden ayırt edici öznitelikler çıkarma yeteneğinin altını çizerek, yüz tanıma sistemlerinin etkinliğine katkıda bulunmaktadır. Bu çalışmalar, yüz tanıma sistemlerinin sahtecilik saldırılarına karşı dayanıklılığını artırmada derin özniteliklerin önemini vurgulamaktadır.

Tez çalışmasında derin öğrenme tabanlı yüz sahteciliği tespiti kapsamında, literatürde yaygınca kullanılan Xception ağının indirgenmiş bir versiyonu üretilmiştir. Bunun yanında, üretilen modele dikkat mekanizmalarını entegre etmek amacıyla SE blokları eklenmiştir. Ek olarak, basit bir SE bloklu artık ağ oluşturulmuş ve tüm bu ağların YST problemine etkileri araştırılmıştır.

2.5.1. Xception Ağı

Xception, sınırlı eğitim verileriyle segmentasyon görevlerini yerine getirme kabiliyetiyle bilinen bir evrişimsel sinir ağıdır (Rusin vd., 2021). Keras kütüphanesinin yaratıcısı François Chollet tarafından önerilen bu mimari, derinlemesine ayrılabilir evrişim katmanları kullanmasıyla karakterize edilen Inception modelinin gelişmiş bir versiyonudur (Fadli & Herlistiono, 2020). Xception mimarisinin arkasındaki ana fikir, uzamsal ve kanal bazlı filtreleme işlemlerini birbirinden ayırarak evrişimsel sinir ağlarının verimliliğini ve performansını artırmaktır.

Geleneksel CNN'lerde, evrişimsel katmanlar hem uzamsal filtreleme (kanallar arası korelasyonlar) hem de kanal bazında filtreleme (uzamsal korelasyonlar) gerçekleştirir. Bu, çok sayıda parametreye ve hesaplama karmaşıklığına yol açabilir. Xception bu sorunu, uzamsal ve kanal bazlı filtrelemeyi ayrı işlemlere bölen derinlik bazlı ayrılabilir evrişimler kullanarak çözmektedir.

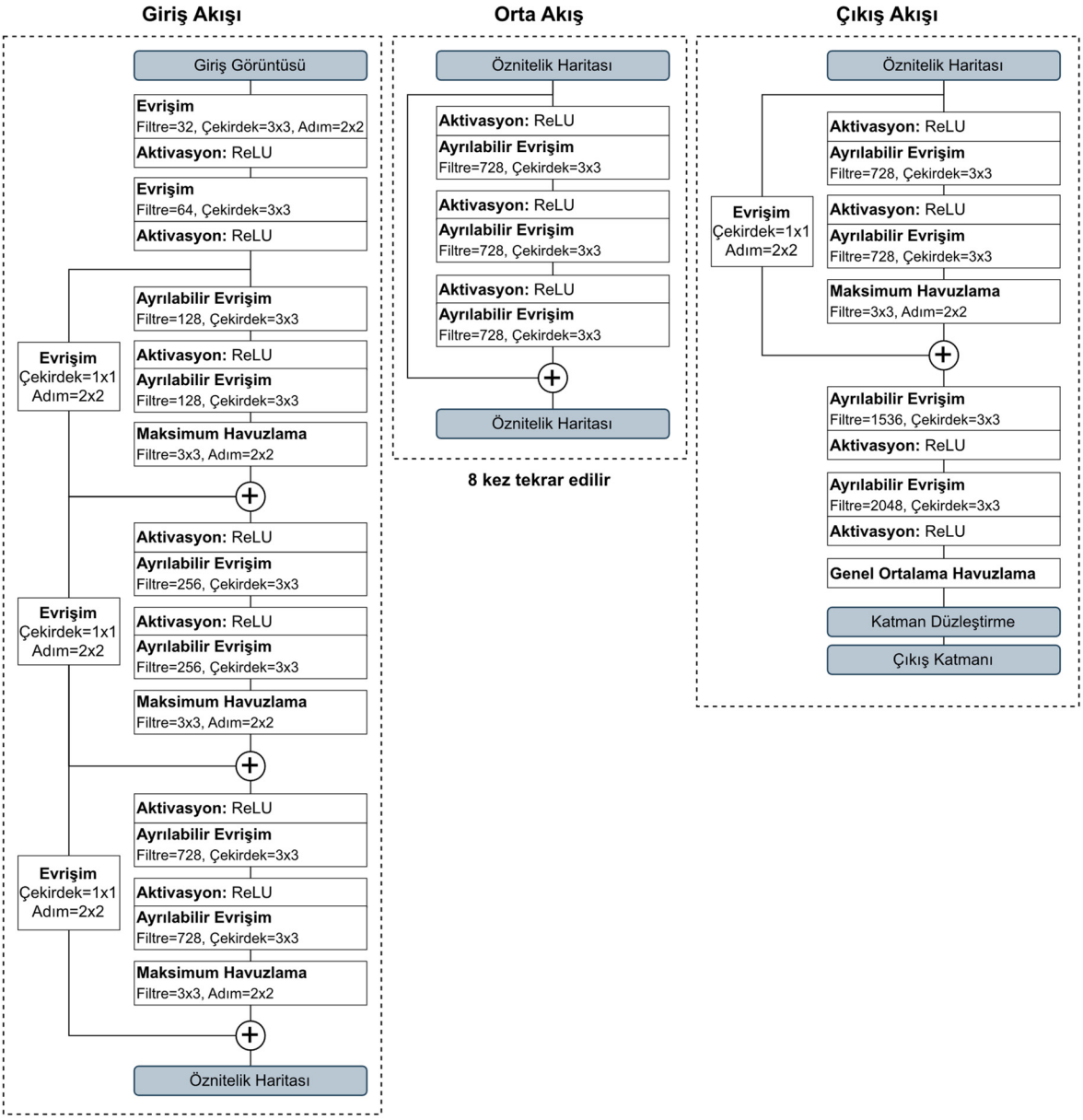
Derinlemesine ayrılabilir evrişimler iki adımdan oluşur: derinlemesine evrişimler ve noktasal evrişimler. Derinlemesine evrişim adımında, her bir giriş kanalı ayrı bir filtre ile evrişir ve bir dizi ara öznitelik haritası elde edilir. Ardından, noktasal evrişim adımında, ara

öznitelik haritalarını nihai çıkış kanallarında birleştirmek için 1×1 evrişim uygulanır. Temel Xception mimarisi giriş, orta ve çıkış olmak üzere üç akıştan meydana gelmektedir.

Giriş Akışı, giriş verilerinden düşük seviyeli özniteliklerin elde edildiği kısmıdır. Tipik olarak birkaç evrişimsel katmandan ve ardından giriş görüntülerinden kenarlar, dokular ve basit desenler gibi temel görsel öznitelikleri çıkarmayı amaçlayan maksimum havuzlama işlemlerinden oluşur. Bu akış, giriş verilerini kademeli olarak aşağı örneklemeyi amaçlamaktadır.

Orta Akış, Xception ağının çekirdeğini oluşturur. Derinlemesine ayrılabilir evrişimlerin çoklu bloklarından oluşur. Bu bloklar farklı soyutlama seviyelerinde öznitelik çıkarımı gerçekleştirir. Giriş akışından elde edilen öznitelik haritalarını işleyerek daha karmaşık ve daha üst düzey görsel öznitelikleri yakalamak için tasarlanmıştır.

Çıkış Akışı, ağın küresel öznitelik toplama ve sınıflandırmadan sorumlu son kısmıdır. Uzamsal bilgileri bir vektörde yoğunlaştırmak için küresel ortalama havuzlama gibi işlemleri ve ardından sınıflandırma veya regresyon görevleri için tam bağlı katmanları içerir. Çıkış akışı, orta akış tarafından çıkarılan üst düzey öznitelikleri alır ve bunları çıktı sınıflarına veya tahminlere eşler. Şekil 28, Xception ağının genel mimarisini göstermektedir.

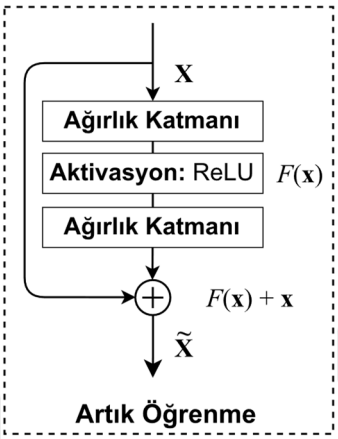


Şekil 28. Xception ağının genel mimarisi

2.5.2. Artık Ağlar

Artık ağlar (ResNets), makine öğrenimi ve bilgisayarla görme alanında büyük ilgi gören bir tür derin sinir ağı mimarisidir. Bu mimari, çok derin sinir ağlarında eğitimi engelleyen kaybolan gradyanlar sorununu ele almak için önerilmiştir. ResNet'lerin temel yeniliği, eğitim sırasında gradyanın daha kolay akmasını sağlayan atlama bağlantıları içeren ve böylece çok derin ağların eğitimini mümkün kılan artık blokların kullanılmasıdır. Bu yaklaşımın, sinir ağlarının daha kolay ve daha hızlı yakınsamasına yardımcı olmada oldukça etkili olduğu gösterilmiştir. Bu durum, kaybolan gradyanlar gibi sorunlar nedeniyle çok

derin ağların eğitilmesinin zor olabildiği derin öğrenme bağlamında özellikle önemlidir. ResNets'te artık blokların kullanılmasının bu sorunları hafiflettiği ve son derece derin ağların başarılı bir şekilde eğitilmesine olanak sağladığı görülmüştür (L. Huang vd., 2017). Şekil 29, artık ağlarda öğrenme işleminin gerçekleştiği temel yapı taşını göstermektedir.



Şekil 29. Artık ağların temel yapı taşı

Şekil 29 incelendiğinde, artık bloklar kullanılmadan önce oluşan girdi (X), katmanın ağırlıkları ile çarpılır ve bir bias terimi bu çarpıma eklenir. Ardından bu ifade F aktivasyon fonksiyonundan geçer ve $F(X)$ çıktısı elde edilir. Son aşamada bu çıktı artık blok ile toplanır ve nihai $F(X) + X$ çıkışı elde edilir.

Atlama bağlantıları ve artık öğrenmeyi bir araya getiren artık ağlar, gelişmiş optimizasyonla çok derin ağların eğitilmesini sağlayarak bozulma sorunuyla karşılaşmadan derinliğin artmasına olanak tanır. Bu mimari, daha derin ve daha güçlü sinir ağlarının eğitilmesi için etkili bir çözüm sağlayarak derin öğrenme alanında önemli bir etkiye sahip olmuştur.

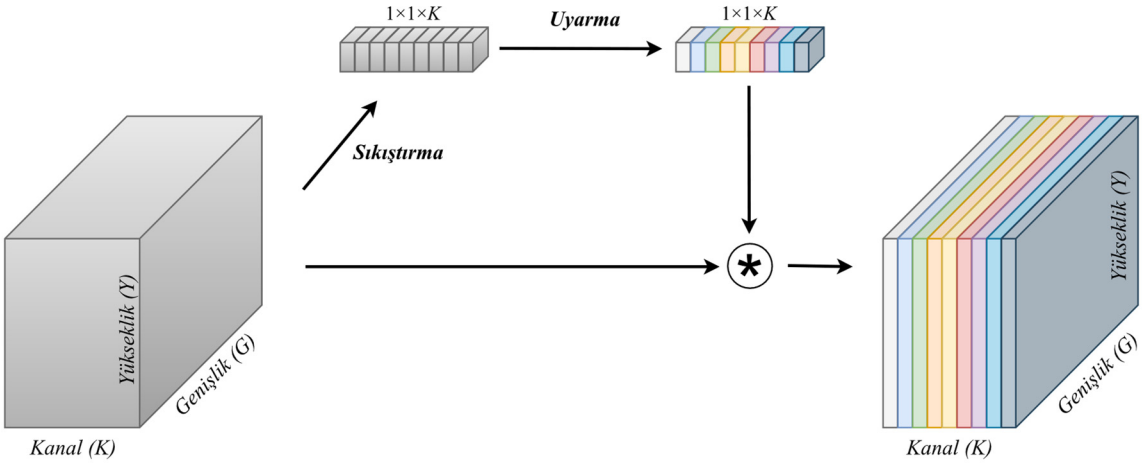
2.5.3. Sıkıştırma ve Uyarma Blokları

Sıkıştırma ve Uyarma (Squeeze and Excitation- SE) blokları, görüntü sınıflandırma, süper çözünürlük ve stok tahmini gibi çeşitli derin öğrenme görevlerinde yaygın olarak kullanılmaktadır (Jeon & Rhee, 2021; Sarigul, 2023). Bu bloklar, kanal bazlı bağımlılıkları açıkça modelleyerek ve öznitelik haritalarını uyarlamalı olarak yeniden kalibre ederek evrişimli sinir ağlarının temsil gücünü artırmayı amaçlamaktadır.

SE bloğu iki ana işlemden oluşur: sıkıştırma ve uyarma. Sıkıştırma işleminde, öznitelik haritalarının uzamsal boyutlarını kanal başına tek bir değere indirgeyen küresel ortalama havuzlama uygulanarak her kanalın küresel bilgisi yakalanır. Bu işlem kanal bazında istatistiklerin yakalanmasına ve önemli bilgilerin özetlenmesine yardımcı olur. Uyarma işleminde, kanal bazlı bağımlılıklar tam bağlı bir katman ve ardından bir aktivasyon fonksiyonu kullanılarak modellenir. Bu adım, ağın her bir kanalın önemini öğrenmesini ve kanal dikkat ağırlıklarını oluşturmasını sağlar.

SE bloğu, artık ağlar gibi mevcut CNN mimarilerine kolayca entegre edilebilir (Jeon & Rhee, 2021). SE blokları artık öğrenme çerçevesine dahil ederek, ağ artık bilgileri etkili bir şekilde öğrenilebilir ve öznitelik haritaları uyarlamalı olarak yeniden kalibre edebilir. Bu da görüntü sınıflandırma ve süper çözünürlük gibi görevlerde daha iyi performans elde edilmesini sağlar.

Yapılan çalışmalar, SE bloğunu da içeren dikkat mekanizmalarının, çeşitli derin öğrenme görevlerinde etkili olduğunu ortaya koymuştur. Dikkat mekanizmaları, ağın ilgili bilgilere odaklanmasını ve alakasız veya gürültülü öznitelikleri göz ardı etmesini sağlar (F. Wang vd., 2017; H. Wang vd., 2022; Watkins vd., 2019; Y. Zhang vd., 2020). Bu mekanizmalar, özellikle görüntü sınıflandırma ve yaya algılama gibi belirli özniteliklerin veya kanalların diğerlerinden daha bilgilendirici olduğu görevlerde kullanışlıdır (Sarigul, 2023; Y. Zhang vd., 2020). Şekil 30, sıkıştırma ve uyarma bloklarının uygulama örneğini göstermektedir.



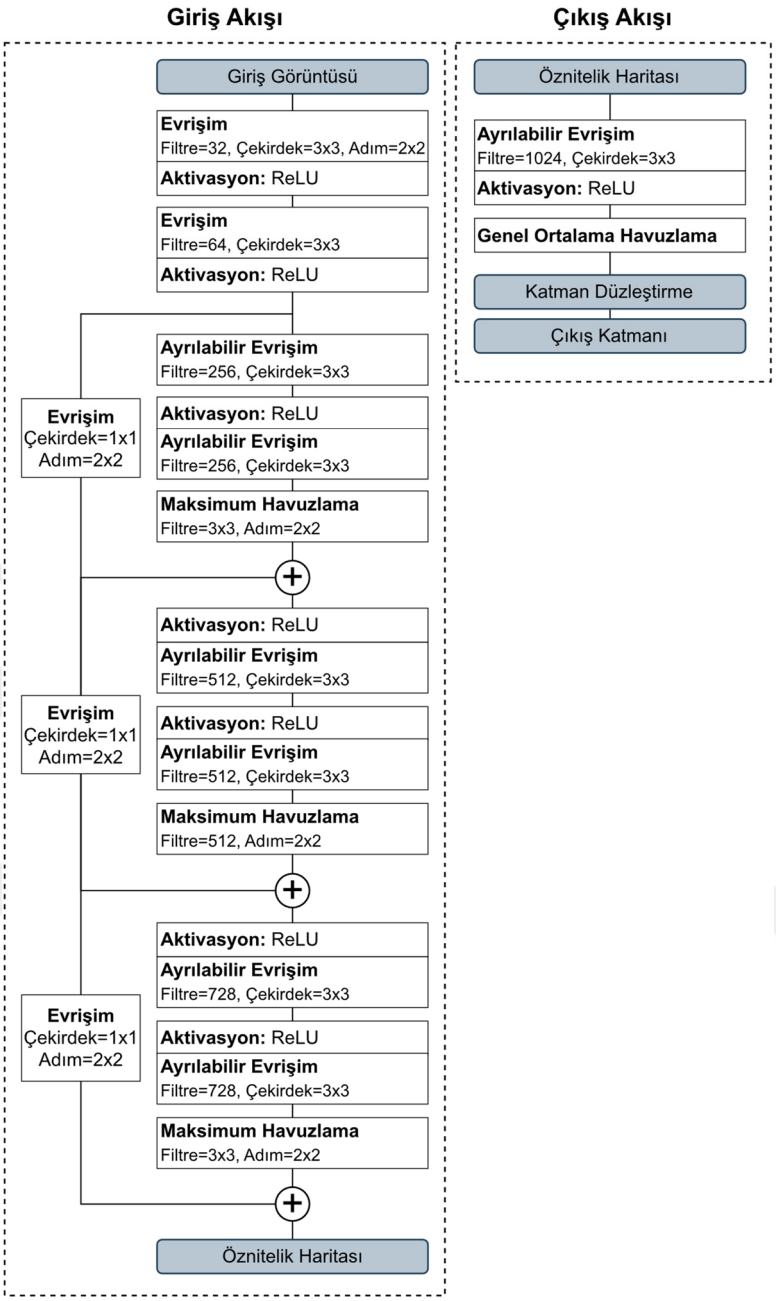
Şekil 30. Sıkıştırma ve uyarma bloğu

2.5.4. İndirgenmiş Xception Ağı

Xception ağı, birçok problemde başarısını ispatlamış güçlü bir mimariye sahiptir. Bu gücü, sınırlı eğitim verilerini işlemedeki verimliliğinden, ayırt edici öznitelik haritalarını öğrenme kabiliyetinden ve çeşitli uygulamalarda diğer CNN yapılarına kıyasla üstün performansından kaynaklanmaktadır.

Xception, derinlemesine ayrılabilir evrişimler kullanarak, temsil kapasitesini korurken hesaplama karmaşıklığını önemli ölçüde azaltır. Bu mimari, verimli öznitelik çıkarımına olanak tanıyarak benzer performansa sahip diğer ağlara kıyasla, hesaplama açısından daha az karmaşık olmasını sağlar. Ancak tüm bu avantajlara rağmen, Xception ağının toplam parametre sayısı 20.863.529'dur. Bir derin öğrenme uygulamasında bu parametrelerin tümünün eğitilmesi zorunlu değildir. Öğrenme aktarımı ve ince ayar gibi birtakım işlemler yardımıyla eğitilecek parametre sayısında azalmaya gidilebilir ancak özellikle ince ayarlamada eğitimin hangi katmandan itibaren yapılacağını belirlemek başlı başına bir optimizasyon problemidir. Bu sebeple, çalışmada kullanılmak üzere Xception ağının indirgenmiş bir versiyonu oluşturulmuştur.

Xception mimarisi üç akıştan oluşmaktadır. Bunlardan ilki olan giriş akışı, düşük seviyeli ayrıntıları yakalar, Orta akış bu ayrıntıları daha soyut ve karmaşık özniteliklere dönüştürür ve son olarak çıkış akışı tahminler veya sınıflandırmalar yapmak için bu yüksek seviyeli öznitelikleri işlemek için kullanılır. Özellikle yüksek seviyeli özniteliklerin çıkarılmasında sekiz kez tekrarlanan orta akış, Xception ağının parametre sayısında ciddi miktarda artışa neden olmaktadır. Bu sebeple, çalışmada kullanılan indirgenmiş Xception ağında orta akış kullanılmamıştır. Bunun yerine, giriş akışındaki filtre boyutları arttırılarak, düşük parametre sayısı yardımıyla tanımlama kabiliyetinin arttırılması hedeflenmiştir. Parametre sayısının azalması sebebiyle, çıkış akışında bulunan evrişim, ayrılabilir evrişim ve maksimum havuzlama katmanları da mimariden çıkarılmış ve giriş akışından üretilen düşük seviyeli öznitelikler çıkış akışında bir kez ayrılabilir evrişimden geçirilerek, genel ortalama havuzlama operatörüne uygulandıktan sonra sınıflandırma için kullanılmıştır. Bu işlemler sonucunda tasarlanan ağın parametre sayısı 2.731.065 olarak hesaplanmıştır. Şekil 31, çalışmada kullanılan indirgenmiş Xception ağı mimarisini temsil etmektedir.

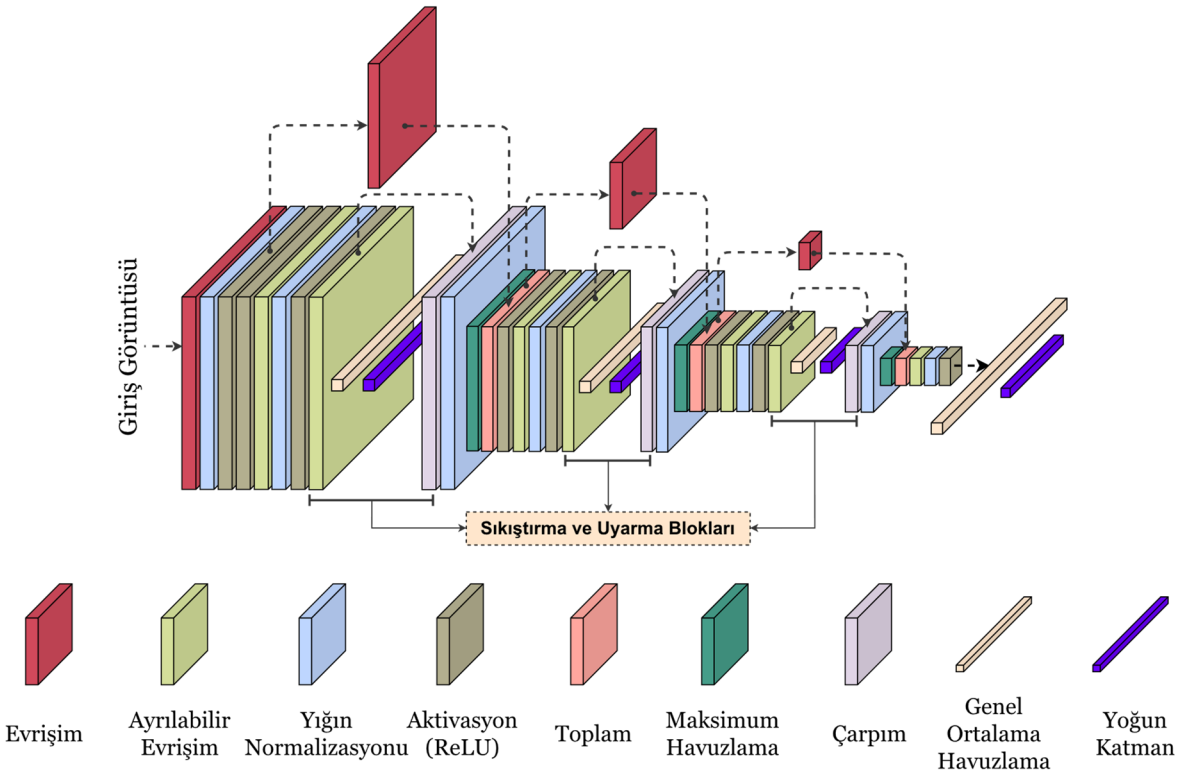


Şekil 31. İndirgenmiş Xception ağı

2.5.5. SE Bloklu İndirgenmiş Xception Ağı

SE bloğunun genel amacı, her kanaldaki özniteliklerin temsilini uyarlanabilir şekilde geliştirmektir. Bu blok, kanal bazlı ilişkileri öğrenerek ve öznitelik yanıtlarını önemlerine göre yeniden kalibre ederek, ağın bilgilendirici özniteliklere daha fazla odaklanmasına ve daha az ilgili olanları bastırmasına yardımcı olur. Bu sayede, evrişimli ağlarda gelişmiş öznitelik öğrenimine imkân sağlar.

Çalışmada kullanılan indirgenmiş Xception ağında kanal etkilerini araştırmak ve sahtecilik tespitinde başarımı arttırmak amacıyla, bu ağın giriş akışındaki evrişim işlemlerinden hemen sonra SE blokları kullanılmıştır. Bu sayede, giriş görüntülerinden elde edilen yüksek kanal sayısına sahip düşük seviyeli özniteliklerde, kanal bazında önemli bilgiler özetlenmiş ve her kanalın önemi belirlenerek, kanal dikkat ağırlıklarının oluşturulması sağlanmıştır. Böylece, problemin çözümüne etkisi daha düşük olan kanallarda bu ağırlıklar bastırılmış ve etkili kanallara olan dikkat arttırılmıştır. Şekil 32, sıkıştırma ve uyarma bloklarını içeren indirgenmiş Xception ağını göstermektedir.



Şekil 32. SE bloklu indirgenmiş Xception ağı

Bu blokların kullanımıyla, önerilen modelde toplam parametre sayısı 3.590.225 olarak hesaplanmıştır.

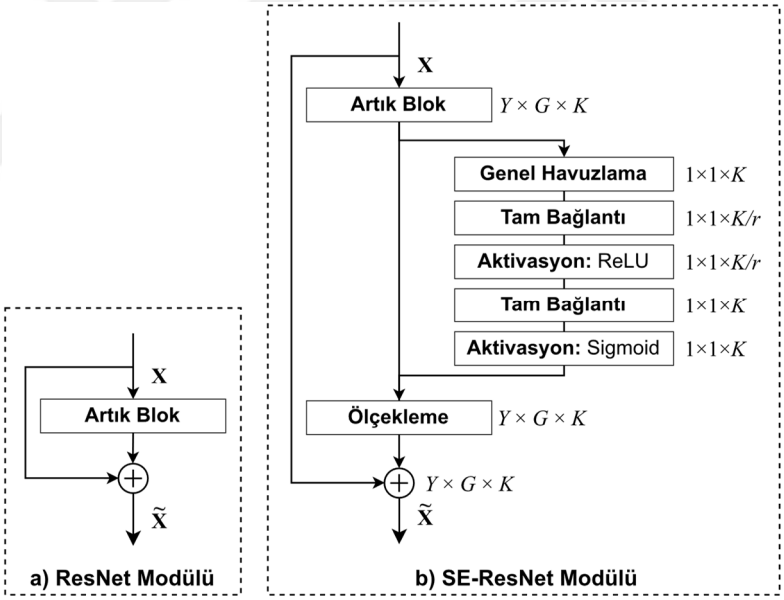
2.5.6. SE Bloklu Artık Ağ

SE Bloklu Artık Ağ (SE-ResNet), sıkıştırma ve uyarma bloklarını orijinal ResNet yapısına dahil eden artık ağ mimarisinin bir varyantını ifade eder. Bu mimaride, ağın temsil gücünü artırmak için ResNet'ten kalan bağlantılar SE bloklarıyla birleştirilir. SE bloklarının

eklenmesi SE-ResNet'in önemli özniteliklere odaklanmasını ve her kanalda daha az ilgili olanları bastırmasını sağlar.

SE-ResNet mimarisi, ağın artık işlevleri öğrenmesini sağlayan ve genellikle kimlik kısa yolları olarak bilinen atlama bağlantılarına sahip artık blokları içeren ResNet'in temel yapısını takip eder. Bu kısa yollar, çok derin ağların eğitimi sırasında gradyanların akışını kolaylaştırarak daha kolay optimizasyona yardımcı olur.

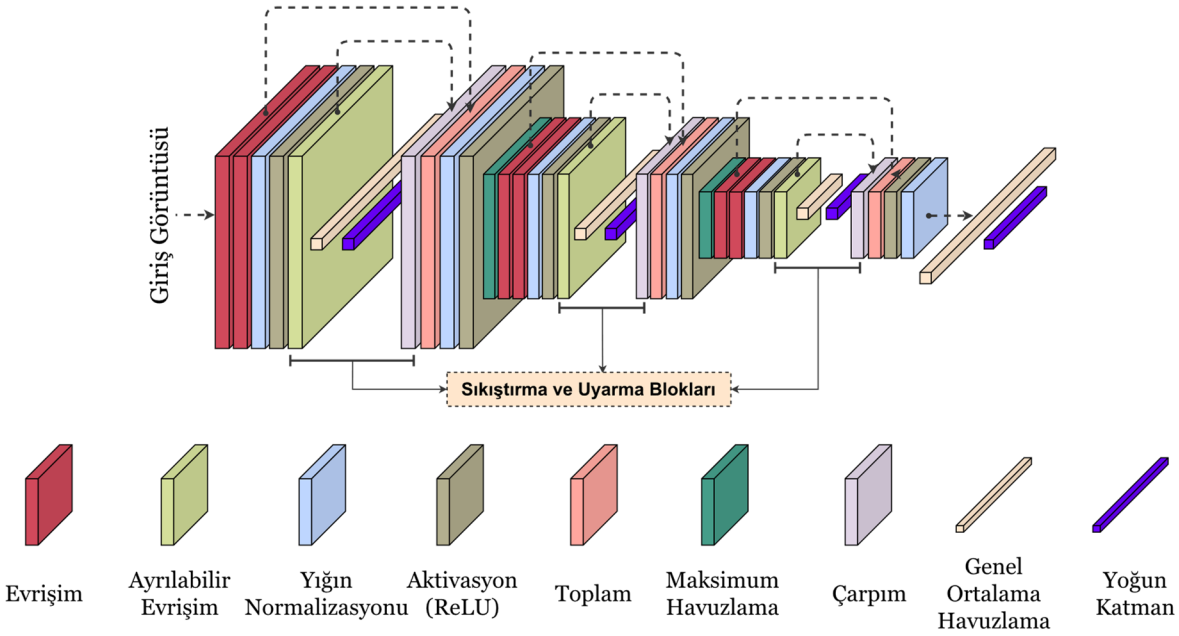
SE-ResNet, özellikle ImageNet gibi görüntü sınıflandırma kriterlerinde olmak üzere çeşitli bilgisayarla görme görevlerinde standart ResNet mimarilerine kıyasla gelişmiş performans göstermiştir (Hu vd., 2018). Bu ağ yapısı, ResNet'in artık bağlantılarının avantajlarını SE bloklarının uyarlanabilir kanal bazlı yeniden kalibrasyonuyla birleştirerek, verilerden bilgilendirici özniteliklerin öğrenilmesinde daha iyi doğruluk ve verimlilik elde eder. Şekil 33, standart bir ResNet modülünü ve SE blokları içeren ResNet modülünü temsil etmektedir.



Şekil 33. Standart ResNet modülü (a), SE bloklu ResNet modülü (b)

Şekilde bulunan X ve $\tilde{X}$ sırasıyla giriş ve çıkış katmanlarını, Y, G, K ve r sırasıyla yükseklik, genişlik, kanal sayısı ve kanal azaltma oranını temsil etmektedir. Kanal azaltma oranı için belirlenen kesin bir değer olmamakla birlikte, araştırmacılar $r = 16$ değerinin doğruluk ve hesaplama karmaşıklığı bağlamında iyi sonuçlar ürettiğini belirtmiştir (Hu vd., 2018). Çalışmada düşük parametreye sahip ağlar tasarlandığından, kullanılan SE bloklarında

$r = 1$ olarak alınmış ve kanal küçültme işlemi uygulanmamıştır. Şekil 34, çalışmada kullanılan SE blokları içeren basit ResNet mimarisini ifade etmektedir.



Şekil 34. SE blokları içeren ResNet mimarisi

Bu mimaride kullanılan filtre sayısının oldukça az olması (16, 32 ve 64) nedeniyle, önerilen modelde toplam parametre sayısı 64.161 olarak hesaplanmıştır.

2.6. Karma Girişli Derin Öğrenme Modeli ile Yüz Sahteciliği Tespiti

Çok girdili derin öğrenme modelleri, belirli bir görevi yerine getirmek için birden fazla girdi verisi alan sinir ağı mimarileridir. Bu modeller, ağın farklı kaynaklardan aynı anda veya paralel olarak öğrenmesini sağlayarak çeşitli veri türlerini işlemek üzere tasarlanmıştır.

Görüntüler ve sayısal veriler gibi karışık girdi türlerini derin öğrenme modellerine dahil ederken, mimari tasarım bu farklı veri türlerini etkili bir şekilde işleme etrafında döner. Buradaki zorluk, modelin her iki kaynaktan aynı anda anlamlı temsiller öğrenmesini sağlarken, her bir girdi türünün farklı yapılarını ve özelliklerini ele almakta yatmaktadır.

Yaygın bir strateji, sinir ağı mimarisi içinde her biri belirli bir girdi türünü işlemeye adanmış paralel yollar oluşturmayı içerir. Örneğin, evrişimsel katmanlar görüntü verilerini işlemek, uzamsal bilgileri yakalamak ve ilgili öznitelikleri çıkarmak için kullanılabilirken,

tam bağlantılı katmanlar gibi ayrı dallar sayısal verileri işleyebilir, sayısal özniteliklerdeki kalıpları ve ilişkileri yakalayabilir.

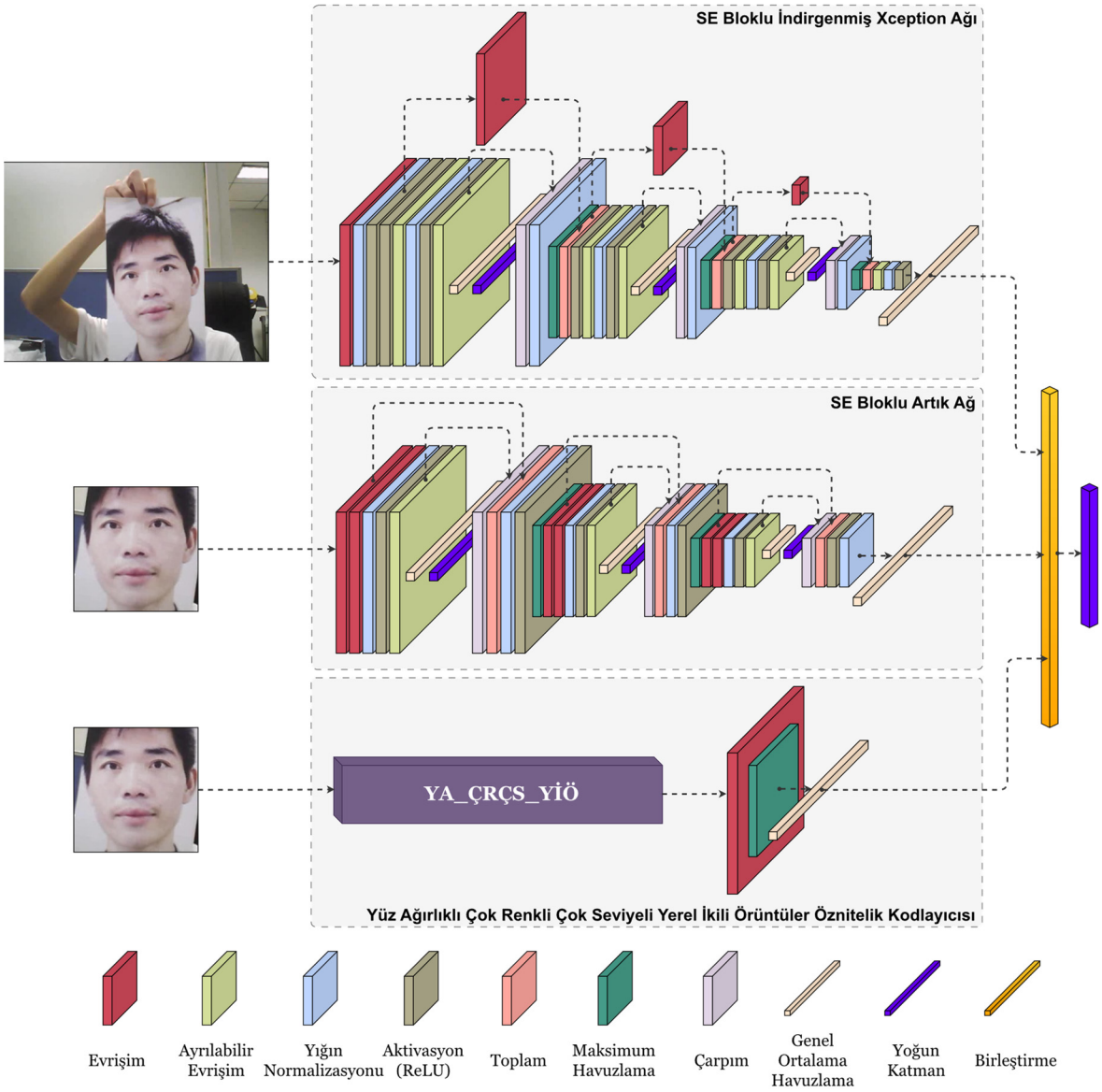
Bir başka yaklaşım da farklı girdi kaynaklarından elde edilen öznitelikleri birleştirmektir. Bu birleştirilmiş temsiller, ortak öğrenme için sonraki katmanlara beslenmeden önce birleştirilebilir veya bir araya getirilebilir. Görüntülerden ve sayısal verilerden elde edilen bilgileri bu şekilde birleştirerek, model potansiyel olarak tamamlayıcı bilgiler çıkarabilir ve daha güçlü temsiller oluşturabilir. Ayrıca, öğrenme süreci sırasında farklı girdi kaynaklarının veya özniteliklerin önemini dinamik olarak vurgulamak veya tartmak için dikkat mekanizmaları entegre edilebilir. Bu, modelin hem görüntü hem de sayısal verilerin ilgili yönlerine odaklanmasını sağlayarak bilinçli tahminler veya sınıflandırmalar yapma kabiliyetini artırır.

Mimari seçimi, karışık girdileri işlemek için en etkili yaklaşımı belirlemek üzere verilerin özel yapısına, görev gereksinimlerine bağlıdır. Derin öğrenme modelleri, farklı bilgi türlerini etkili bir şekilde entegre ederek, genel performansı ve tahmin doğruluğunu artırmak için hem görüntü hem de sayısal verilerin güçlü yönlerinden yararlanabilir. Şekil 35, çalışmada kullanılan çok girişli derin öğrenme modeline ait mimariyi ifade etmektedir.

Önerilen mimari, SE Bloklu İndirgenmiş Xception Ağı, SE Bloklu Artık Ağ ve Yüz Ağırlıklı Çok Renkli Çok Seviyeli Yerel İkili Örüntüler Kodlayıcısı olmak üzere üç kanaldan meydana gelmektedir.

Xception ağından esinlenilerek üretilen ve diğerlerine oranla çok daha fazla parametre içeren ilk kanalda, giriş görüntülerinde yüz bölgesi dışında kalan alanların probleme etkisini arttırmak amacıyla, veri setlerinde bulunan işlenmemiş ham görüntüler giriş verisi olarak kullanılmaktadır. Kullanılan SE blokları, evrişimler sonucu üretilen kanallar üzerinde ağırlıklandırmalar yaparak, güçlü özniteliklerin bir sonraki katmanlarda etkisini arttırmayı hedeflemektedir. Bu kanalda, artık bloklarda kullanılan evrişim blokları, giriş verisinin çıkışa eşitlenmesi ve ağın, daha karmaşık öznitelikler öğrenmesini ve çıkarmasını sağlamak amacıyla kullanılmıştır.

İkinci kanalda kullanılan SE Bloklu Artık Ağ, karşılaştırılan tüm ağlar arasında en düşük parametreye sahip olanıdır. Oldukça düşük sayıda filtreye sahip olması, giriş görüntülerinde öğrenmeyi zorlaştıracak verilerden kaçınılmasını mecbur kılmaktadır. Bu sebeple, önerilen mimarinin bu kanalında kullanılan derin öğrenme ağı için giriş görüntüsü olarak yalnızca “yüz görüntüleri” kullanılmıştır.



Şekil 35. Önerilen karma girişli modelin yapısı

Üçüncü kanalda ise numerik verileri işleyen basit bir ağ tasarımı gerçekleştirilmiştir. Derin öğrenme ağları varsayılan olarak RGB giriş uzayı üzerinde işlemler gerçekleştirmektedir. Çalışmada önerilen YA_ÇRÇS_YİÖ öznitelik kodlayıcısı, düşük vektör boyutunun yanında, farklı renk uzaylarını kullanarak, YST probleminde etkili bir çözüm sunmaktadır. Bu aşamada, HSV ve L*a*b* renk uzaylarında YA_ÇRÇS_YİÖ kodlayıcısının güçlü tanımlama yeteneği kullanılarak, derin öğrenme ağına farklı renk uzaylarının etkileri eklenmiş ve sonuçta YST performansında başarım artışı hedeflenmiştir.

Önerilen bu mimaride, birden fazla modelin bir arada bulunması nedeniyle, toplam parametre sayısı 3.817.777 olarak hesaplanmıştır.

3. BULGULAR VE TARTIŞMA

Bu bölümde, tez çalışması boyunca elde edilen bulgular açıklanmış ve literatürdeki sonuçlarla karşılaştırmalar yapılmıştır.

3.1. Yüz Bölgelerinin Sahtecilik Tespiti Başarımları

Çalışmanın ilk aşamasında yüz bölgelerinin sahtecilik tespitine etkileri araştırılmıştır. Giriş görüntüsünden elde edilen geniş yüz, kırpılmış yüz, ağız, burun ve göz bölgeleri yardımıyla YST gerçekleştirilmiş, bu aşamada en iyi sonuçları üreten yüz bölgesi, diğer çalışmalar için giriş olarak kullanılmıştır. Bunun yanında, geniş yüz ve kırpılmış yüz görüntülerine kıyasla çok daha az bilgi içeren ağız, burun ve göz bölgelerinin kendi aralarında bir değerlendirmesi de yapılmıştır.

Çalışmada kullanılan NUAA ve CASIA-FASD veri setlerinde, videolardaki tüm çerçeveler dikkate alınmış ve bunlardan yüz bölgeleri tespit edilen kareler deneylerde kullanılmıştır. REPLAY-ATTACK veri setinde ise örneklerin ve test senaryolarının çokluğu nedeniyle, işlem süresi ve karmaşıklığını azaltmak için videolardan 125ms aralıklarla çerçeveler alınarak görüntü sayısı azaltılmıştır. Veri setlerinde kullanılan kare sayısı sırasıyla; 12.614, 110.379 ve 101.201 olarak belirlenmiştir.

NUAA ve CASIA-FASD veri setleri eğitim ve test kümelerinden oluşmaktadır. REPLAY-ATTACK veri setinde ise eğitim ve test kümeleri dışında bir de geliştirme kümesi bulunmaktadır. Bu nedenle NUAA ve CASIA-FASD veri setleri için geliştirme kümesi, 5 kat çapraz doğrulama metodu yardımıyla eğitim setinden elde edilmiştir. Çapraz doğrulamanın her aşamasında 5 eşit parçaya bölünen eğitim kümesinden, 4 parça eğitim seti, kalan parça ise geliştirme seti olarak kullanılmıştır. Bu işlem 5 kez tekrar edilip elde edilen sonuçların ortalaması hesaplanmıştır.

NUAA veri setinde tek saldırı türü (basılı fotoğraf) için 3.459 eğitim, 9.067 test örneği bulunmaktadır. CASIA-FASD veri setinde üç farklı görüntü kalitesi, üç farklı saldırı türü ve tüm verilerin bir arada olduğu genel test ile toplamda yedi test senaryosu bulunmaktadır. Bunlar, düşük kaliteli örnek saldırısı, normal kaliteli örnek saldırısı, yüksek kaliteli örnek saldırısı, bükülmüş fotoğraf saldırısı, kesilmiş fotoğraf saldırısı, video oynatma saldırısı ve bu saldırı türlerinin tümünün bir arada olduğu tümü saldırısı şeklindedir. Çalışmada bu yedi senaryo için ayrı ayrı sonuçlar elde edilmiştir. REPLAY-ATTACK veri

seti ise temelde 6 farklı saldırı türü içermektedir. Bunlar sırasıyla; yüksek çözünürlüklü, mobil, basılı fotoğraf, dijital+basılı fotoğraf, video oynatma ve bu saldırı türlerinin tümünün bir arada olduğu karma saldırılarıdır. Diğer yandan, veri setinde saldırı görüntülerinin/videolarının kameraya sunulması sırasında iki farklı yöntem kullanılmıştır: cihazın elde tutulması ve cihazın sabit bir yere konumlandırılması. Konumlandırma türüne göre saldırı senaryoları ele alındığında (elde, sabit ve iki konumlandırma türünün bir arada kullanılması), REPLAY-ATTACK veri seti için $6 \times 3 = 18$ farklı test senaryosu değerlendirilmiştir.

Tablo 3, üç veri seti için deneylerde kullanılan gerçek ve sahte görüntü sayılarını içermektedir.

Tablo 3. Deneylerde kullanılan örnek sayıları

Veri Seti	Saldırı Türü	Gerçek Örnek Sayıları			Sahte Örnek Sayıları		
		Eğitim	Geliştirme	Test	Eğitim	Geliştirme	Test
NUAA	Basılı Fotoğraf	1.743	-	3.356	1.746	-	5.711
CASIA-FASD	Düşük Kalite	3.140		5.297	11.015		16.088
	Normal Kalite	3.197		4.830	11.231		16.085
	Yüksek Kalite	4.577		5.782	11.846		17.291
	Bükülmüş Fotoğraf		-		12.859	-	19.165
	Kesilmiş Fotoğraf	10.914		15.909	9.499		14.731
	Video Oynatma				11.734		15.568
	Tümü				34.092		49.464
REPLAY-ATTACK	H_Yüksek Çözünürlük	7.185	7.197	9.596	4.761	4.633	6.106
	F_Yüksek Çözünürlük				4.757	4.651	6.173
	A_Yüksek Çözünürlük				9.518	9.284	12.279
	H_Mobil				4.703	4.706	6.084
	F_Mobil				4.717	4.462	5.731
	A_Mobil				9.420	9.168	11.815
	H_Basılı Fotoğraf				7.137	7.104	9.218
	F_Basılı Fotoğraf				7.170	6.927	9.142
	A_Basılı Fotoğraf				14.307	14.031	18.360
	H_Dijital+Basılı Fotoğraf				2.395	2.396	3.114
	F_Dijital+Basılı Fotoğraf				2.370	2.342	3.122
	A_Dijital+Basılı Fotoğraf				4.765	4.738	6.236
	H_Video				4.722	4.631	6.086
	F_Video				4.674	4.528	5.884
	A_Video				9.396	9.159	11.970
	H_Tümü				11.859	11.735	15.304
	F_Tümü				11.844	11.455	15.026
	A_Tümü				23.703	23.190	30.330

Tablo 3 incelendiğinde, H_: Giriş cihazı elde tutularak gerçekleştirilen saldırıları, F_: Giriş cihazı sabit bir destek üzerinde konumlandırılarak gerçekleştirilen saldırıları, A_: bu iki kümenin birleşimini ifade etmektedir.

Çalışmada, yüz bölgelerinden yapılacak yüz sahteciliği tespitinde bölgesel YİÖ öznitelikleri kullanılmıştır. Giriş görüntüleri (Geniş Yüz, Kırpılmış Yüz, Ağız, Burun ve

Göz) çeşitli sayıda alt bölgelere ayrıştırılarak bölgesel öznitelikler çıkarılmıştır. Önceki bir çalışmada (Yılmaz vd., 2020), 8×8 bölgesel ayrıştırma işleminin benzer giriş görüntülerinde problemin çözümüne etkisi olmayan özniteliklerin üretilmesine ve vektör boyutunun oldukça artmasına neden olduğu görülmüştür. Bu sebeple çalışmada, giriş görüntülerinin $1 \times 1, 2 \times 2$ ve 4×4 alt bölgelerinden elde edilen öznitelik vektörleri kullanılmıştır. Görüntüler alt bölgelere ayrıldıktan sonra bu bölgelere $Y\dot{I}\ddot{O}_{8,1}^{U2}$ operatörü uygulanmıştır. Bu bölgelerden elde edilen $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1}^{U2}$ öznitelikleri SVM ile sınıflandırılmıştır. Daha sonra aynı işlemler, TBA yardımıyla boyutu indirgenen öznitelik kümesi üzerinde tekrar gerçekleştirilmiştir.

Tablo 4, NUAA, CASIA-FASD ve REPLAY-ATTACK veri setleri üzerindeki YST deneylerinde elde edilen sınıflandırma sonuçlarından en iyilerini içermektedir. En iyi kavramı, bir giriş görüntüsünün farklı bölgesel ayrıştırma işlemlerine tabi tutulması ve sınıflandırıcının, her saldırı senaryosu için hangi bölgesel ayrıştırmada en iyi sonucu ürettiğini temsil etmektedir. Farklı test senaryoları ve kullanılan yüz bölgeleri nedeniyle, elde edilen sonuçların sayısı oldukça fazladır. Örneğin; yalnızca REPLAY-ATTACK veri seti, 5 farklı yüz bölgesi (Geniş Yüz, Kırpılmış Yüz, Göz, Ağız, Burun), 3 farklı bölgesel ayrıştırma ($1 \times 1, 2 \times 2, 4 \times 4$), 18 farklı test senaryosu ve 2 farklı öznitelik kümesi ($Y\dot{I}\ddot{O}_{8,1}^{U2}$, $Y\dot{I}\ddot{O}_{8,1}^{U2} + TBA$) için, $5 \times 3 \times 18 \times 2 = 540$ sınıflandırma sonucu içermektedir. Bu nedenle tabloda yalnızca en iyi sonuçlar paylaşılmıştır.

Tablo 4, üç veri seti üzerinde 5 farklı yüz bölgesi ve toplamda 26 farklı test senaryosu için YST başarımının sonuçlarını içermektedir. Tabloya göre, 26 test senaryosunun 20'sinde geniş yüz bölgesinin YST başarımı diğer bölgelere göre (kırpılmış yüz, ağız, burun ve göz) daha yüksektir. Kalan 6 test senaryosunda ise kırpılmış yüz bölgesi daha başarılıdır. Diğer yandan, $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1}^{U2}$ özniteliklerinin boyutunun TBA ile küçültülmesi, 26 test senaryosunun 16'sında başarımı artırmaktadır. NUAA veri setinde bulunan tek saldırı tipi için kırpılmış yüz bölgesinden elde edilen $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1}^{U2}$ öznitelikleri, YST probleminde %3,53 HTER değerine sahiptir.

CASIA-FASD veri setinde en yüksek başarım, video oynatma saldırısı için geniş yüz bölgesinden üretilen $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1}^{U2}$ öznitelikleri ile %4,66 HTER olarak elde edilmiştir. Veri setinde tanımlanan 7 test senaryosundan 4'ünde (düşük kalite, normal kalite, basılı fotoğraf, video oynatma saldırıları) geniş yüz bölgesinden elde edilen $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1}^{U2}$ özniteliklerinin yüz sahteciliği tespitinde daha iyi olduğu söylenebilir. Diğer yandan, yüksek kalite ve kesilmiş

fotoğraf saldırılarında en iyi sonuçlar, kırpılmış yüz bölgelerinden sırasıyla $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ ve $ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ öznitelikleri kullanılarak elde edilmiştir.

Tablo 4. NUAA, CASIA ve REPLAY-ATTACK veri setlerinde, yüz bölgeleri ve test senaryoları için elde dilen en iyi HTER değerleri

Veri Seti	Saldırı Test Senaryosu	Bölgeler									
		Geniş Yüz		Kırpılmış Yüz		Ağız		Burun		Göz	
		$ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ (Bölge)
NUAA	Basılı Fotoğraf	7,93 (2x2)	12,65 (1x1)	**3,53 (2x2)**	17,54 (1x1)	11,35 (2x2)	12,54 (2x2)	9,52 (4x4)	23,77 (4x4)	19,78 (4x4)	13,99 (1x1)
CASIA	Düşük Kalite	**7,00 (2x2)**	16,63 (2x2)	10,84 (2x2)	12,81 (1x1)	11,95 (4x4)	12,81 (4x4)	10,69 (2x2)	10,20 (4x4)	21,88 (4x4)	23,26 (4x4)
	Normal Kalite	**9,04 (2x2)**	19,91 (4x4)	11,49 (2x2)	14,41 (4x4)	14,81 (4x4)	15,72 (2x2)	16,01 (2x2)	15,74 (2x2)	24,20 (2x2)	22,33 (2x2)
	Yüksek Kalite	22,55 (4x4)	33,66 (1x1)	**18,72 (4x4)**	29,24 (2x2)	28,42 (4x4)	27,83 (2x2)	31,81 (1x1)	30,30 (2x2)	38,54 (2x2)	37,67 (2x2)
	Bükülmüş Fotoğraf	**11,16 (4x4)**	14,96 (2x2)	16,59 (4x4)	13,41 (4x4)	23,35 (2x2)	24,40 (2x2)	25,97 (2x2)	23,65 (2x2)	32,40 (4x4)	31,28 (4x4)
	Kesilmiş Fotoğraf	13,63 (4x4)	12,38 (2x2)	15,24 (4x4)	**10,17 (4x4)**	18,73 (2x2)	22,68 (2x2)	22,29 (1x1)	21,38 (2x2)	15,44 (4x4)	15,92 (4x4)
	Video Oynatma	**4,66 (2x2)**	7,84 (2x2)	10,37 (4x4)	8,24 (4x4)	20,66 (4x4)	20,84 (2x2)	21,45 (2x2)	21,35 (2x2)	28,60 (1x1)	28,75 (2x2)
	Tümü	13,60 (4x4)	14,85 (2x2)	15,89 (4x4)	**12,75 (4x4)**	24,52 (2x2)	25,11 (2x2)	25,65 (2x2)	23,51 (2x2)	31,64 (4x4)	32,58 (4x4)
REPLAY-ATTACK	H_ Yüksek Çözünürlük	8,40 (2x2)	**7,32 (2x2)**	12,60 (2x2)	9,72 (4x4)	22,82 (2x2)	23,46 (2x2)	27,16 (1x1)	27,61 (2x2)	19,95 (2x2)	17,06 (2x2)
	F_ Yüksek Çözünürlük	10,38 (2x2)	**6,67 (2x2)**	13,33 (1x1)	8,54 (4x4)	20,80 (2x2)	20,10 (2x2)	23,33 (2x2)	24,33 (2x2)	17,83 (2x2)	16,51 (2x2)
	A_Yüksek Çözünürlük	9,09 (2x2)	**6,35 (2x2)**	11,64 (2x2)	9,63 (4x4)	21,44 (2x2)	23,02 (2x2)	24,13 (1x1)	26,32 (2x2)	18,53 (2x2)	17,18 (2x2)
	H_Mobil	4,63 (2x2)	**2,37 (4x4)**	2,82 (2x2)	3,51 (4x4)	10,63 (2x2)	9,91 (1x1)	11,16 (2x2)	11,64 (2x2)	9,34 (2x2)	13,29 (2x2)
	F_Mobil	**0,76 (4x4)**	0,94 (4x4)	1,32 (2x2)	1,68 (4x4)	13,24 (2x2)	12,09 (2x2)	6,76 (2x2)	6,81 (2x2)	9,26 (2x2)	13,31 (1x1)
	A_Mobil	4,36 (2x2)	**2,44 (4x4)**	4,26 (2x2)	3,84 (4x4)	11,90 (2x2)	11,98 (2x2)	10,02 (2x2)	10,41 (2x2)	10,89 (2x2)	14,73 (2x2)
	H_Basılı Fotoğraf	**1,11 (4x4)**	3,06 (4x4)	4,51 (4x4)	5,07 (1x1)	17,49 (1x1)	17,42 (2x2)	7,21 (4x4)	7,20 (2x2)	14,01 (4x4)	12,67 (2x2)
	F_Basılı Fotoğraf	**2,36 (4x4)**	2,57 (2x2)	4,00 (4x4)	5,08 (1x1)	19,38 (1x1)	17,66 (4x4)	7,49 (1x1)	6,61 (2x2)	16,77 (1x1)	16,04 (2x2)
	A_Basılı Fotoğraf	**2,39 (4x4)**	2,87 (4x4)	4,11 (4x4)	4,45 (4x4)	18,78 (1x1)	18,24 (4x4)	7,56 (2x2)	7,05 (2x2)	15,18 (4x4)	13,54 (2x2)
	H_Dijital + Basılı Fotoğraf	9,52 (1x1)	**8,38 (2x2)**	12,47 (2x2)	10,73 (2x2)	23,71 (2x2)	23,90 (2x2)	22,76 (4x4)	23,26 (2x2)	22,25 (2x2)	21,99 (2x2)
	F_Dijital + Basılı Fotoğraf	9,83 (2x2)	**5,14 (2x2)**	8,52 (4x4)	6,45 (4x4)	23,06 (2x2)	21,84 (2x2)	20,53 (4x4)	19,41 (2x2)	24,06 (2x2)	23,32 (2x2)

Tablo 4'ün devamı

Veri Seti	Saldırı Test Senaryosu	Bölgeler									
		Geniş Yüz		Kırpılmış Yüz		Ağız		Burun		Göz	
		$ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ (Bölge)	$ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ (Bölge)
REPLAY-ATTACK	A_Dijital + Basılı Fotoğraf	11,11 (2x2)	**7,89 (2x2)**	11,67 (4x4)	8,90 (4x4)	24,23 (2x2)	23,82 (2x2)	21,67 (4x4)	22,06 (1x1)	22,95 (2x2)	23,66 (2x2)
	H_Video	6,81 (2x2)	5,93 (4x4)	7,96 (2x2)	**5,21 (4x4)**	16,89 (1x1)	15,46 (1x1)	18,35 (2x2)	19,28 (2x2)	12,84 (2x2)	14,57 (2x2)
	F_Video	6,73 (4x4)	**3,33 (2x2)**	8,37 (4x4)	4,01 (4x4)	15,77 (2x2)	16,41 (2x2)	16,43 (2x2)	17,21 (2x2)	12,37 (2x2)	14,53 (1x1)
	A_Video	7,38 (2x2)	5,21 (4x4)	7,71 (2x2)	**4,90 (4x4)**	16,27 (1x1)	15,45 (1x1)	17,21 (2x2)	17,98 (2x2)	12,62 (2x2)	15,38 (1x1)
	H_Tümü	10,06 (1x1)	**7,61 (2x2)**	10,49 (2x2)	8,96 (2x2)	23,73 (2x2)	22,60 (1x1)	23,36 (4x4)	23,61 (2x2)	20,81 (2x2)	19,48 (2x2)
	F_Tümü	9,37 (2x2)	**4,14 (2x2)**	10,01 (4x4)	6,04 (4x4)	21,29 (2x2)	20,00 (2x2)	20,13 (2x2)	19,06 (2x2)	21,75 (2x2)	21,49 (2x2)
	A_Tümü	10,12 (2x2)	**7,09 (2x2)**	10,29 (2x2)	8,47 (2x2)	22,48 (2x2)	23,32 (2x2)	22,44 (2x2)	21,49 (2x2)	21,19 (2x2)	20,88 (2x2)

CASIA-FASD veri setindeki tüm saldırı türlerinin bir arada olduğu senaryo değerlendirildiğinde, kırpılmış yüz bölgesinden üretilen $ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ özniteliklerinin %12,75 ile en düşük HTER değerine sahip olduğu görülmektedir.

Daha fazla örnek, saldırı türü ve test senaryosuna sahip REPLAY-ATTACK veri setinde ise en yüksek başarım, sabit bir kaynak üzerinden gerçekleştirilen mobil saldırıda (F_Mobil), geniş yüz bölgesinden üretilen $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelikleri ile %0,76 HTER olarak elde edilmiştir. Mobil saldırıların tümünde (F + H) ise $ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ öznitelikleri ile %2,44 HTER değerine ulaşılmıştır. Bu veri setindeki 18 test senaryosundan sadece ikisinde (H_Video ve A_Video) kırpılmış yüz bölgesinden üretilen $ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ öznitelikleri YST başarımını artırmıştır. Kalan 16 test senaryosunun dördünde geniş yüz bölgesinden üretilen $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelikleri daha iyi başarıma sahipken, 12 test senaryosunda yine geniş yüz bölgesinden üretilen $ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ özniteliklerinin daha iyi YST performansına sahip olduğu görülmektedir. REPLAY-ATTACK veri setindeki tüm saldırı türlerinin bir arada olduğu senaryoda %7,09 HTER değeri elde edilmiştir. Bu değer, sadece sabit kaynaktan yapılan saldırılarda %4,14'e kadar düşmektedir. Genel olarak, aynı öznitelik çıkarma yönteminin daha çeşitli saldırı türü ve daha kaliteli görüntüler içeren veri seti üzerinde daha başarılı olduğu ifade edilebilir. Tüm bu değerlendirme sonuçları, geniş yüz bölgesinin daha çok bilgi içerdiğini ortaya koymaktadır.

Ağız, burun ve göz bölgelerinin HTER ölçeği çerçevesinde YST başarımları incelendiğinde, elde edilen en iyi sonuçlar; NUAA veri seti için %9,52, CASIA-FASD veri setindeki düşük kalite saldırısı için %10,20 ve REPLAY-ATTACK veri setindeki F_Mobil saldırısı için %6,76 olarak, tüm veri setleri için burun bölgelerinden üretilmiştir. Bu durum, bireysel senaryolar bazında burun bölgesinin YST başarımındaki etkisini ortaya koymaktadır.

CASIA-FASD veri setinde 7 test senaryosunun dördünde ağız bölgesi, ikisinde burun bölgesi ve birinde ise göz bölgesinin yüz sahteciliği tespitinde daha başarılı olduğu görülmektedir. Tüm saldırılar birlikte değerlendirildiğinde, burun bölgesinden çıkarılan $ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ özniteliklerinin %23,51 HTER değerine sahip olduğu görülmüştür.

REPLAY-ATTACK veri setindeki saldırı türleri göz önüne alındığında, yüksek çözünürlük saldırısında göz bölgesinin, burun ve ağız bölgelerine kıyasla YST başarımı daha yüksektir. A_Yüksek Çözünürlük saldırısında göz bölgesinden üretilen $ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ öznitelikleri ile %17,18 HTER değeri elde edilmiştir. Bu saldırı türünde, burun ve ağız bölgeleri için başarımlar sırasıyla %24,13 ve %21,44 olarak hesaplanmıştır. Bu veri setindeki 18 test senaryosundan 10'unda göz bölgesinden üretilen özniteliklere ait YST performansının daha yüksek olduğu görülmektedir. Bu durum, günümüz ve gelecekteki pandemi koşullarında maske kullanımı nedeniyle sadece göz bölgesinden yapılacak tanıma sistemlerine yapılacak saldırıların tespiti başarımları ile ilgili bir öngörü oluşturmaktadır. Kalan 8 test senaryosunda ise burun bölgesinin başarımının yüksek olduğu anlaşılmaktadır. Bu nedenle göz ve burun bölgesinin birlikte değerlendirilmesinin başarımı artırabileceği düşünülebilir. Ağız bölgesi bu veri setindeki test senaryolarında göz ve burun bölgesine göre bir üstünlük göstermemiştir. Tüm saldırı türleri dahil edildiğinde, göz bölgesinden elde edilen $ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ özniteliklerinin sahtecilik tespiti başarımı %20,88 HTER olarak belirlenmiştir. Sonuçlar, göz bölgesinin, burun ve ağız bölgelerine göre YST probleminin çözümünde daha etkili olduğunu göstermektedir. Diğer bir sonuç ise $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelik boyutunun küçültülmesinin başarımı artırdığı şeklindedir. Genel bir değerlendirme olarak, geniş yüz bölgesinden elde edilen başarılı sonuçlar, yüz bölgesi dışında kalan alanların, YST probleminin tespitine olumlu katkılar yaptığını ortaya koymaktadır.

3.2. Tüm Yerel İkili Örüntülerin Sahtecilik Tespiti Başarımları

Bir önceki bölümde, 5 yüz bölgesinden YST problemine etkisi en çok olanın geniş yüz bölgesi olduğu belirlendiğinden, bu aşamadaki deneyler sadece geniş yüz bölgesi üzerinde gerçekleştirilmiştir. Diğer yandan doku tanıma çalışmalarında genellikle $Y\dot{I}\ddot{O}_{8,1}^{U2}$ özniteliklerinin kullanıldığı görülmektedir. Bu çalışmada tüm YİÖ özniteliklerinin ($Y\dot{I}\ddot{O}_{8,1}$) YST başarımına etkisini incelemek amacıyla, geniş yüz bölgesinden çıkarılan $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1}$ öznitelikleri kullanılmıştır.

Şekil 36, geniş yüz bölgesinden çıkarılan $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1}^{U2}$, $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1}^{U2} + TBA$ ve $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1} + TBA$ özniteliklerinin, NUAA, CASIA-FASD ve REPLAY-ATTACK veri setlerindeki test senaryolarına göre YST başarımlarını ifade etmektedir.

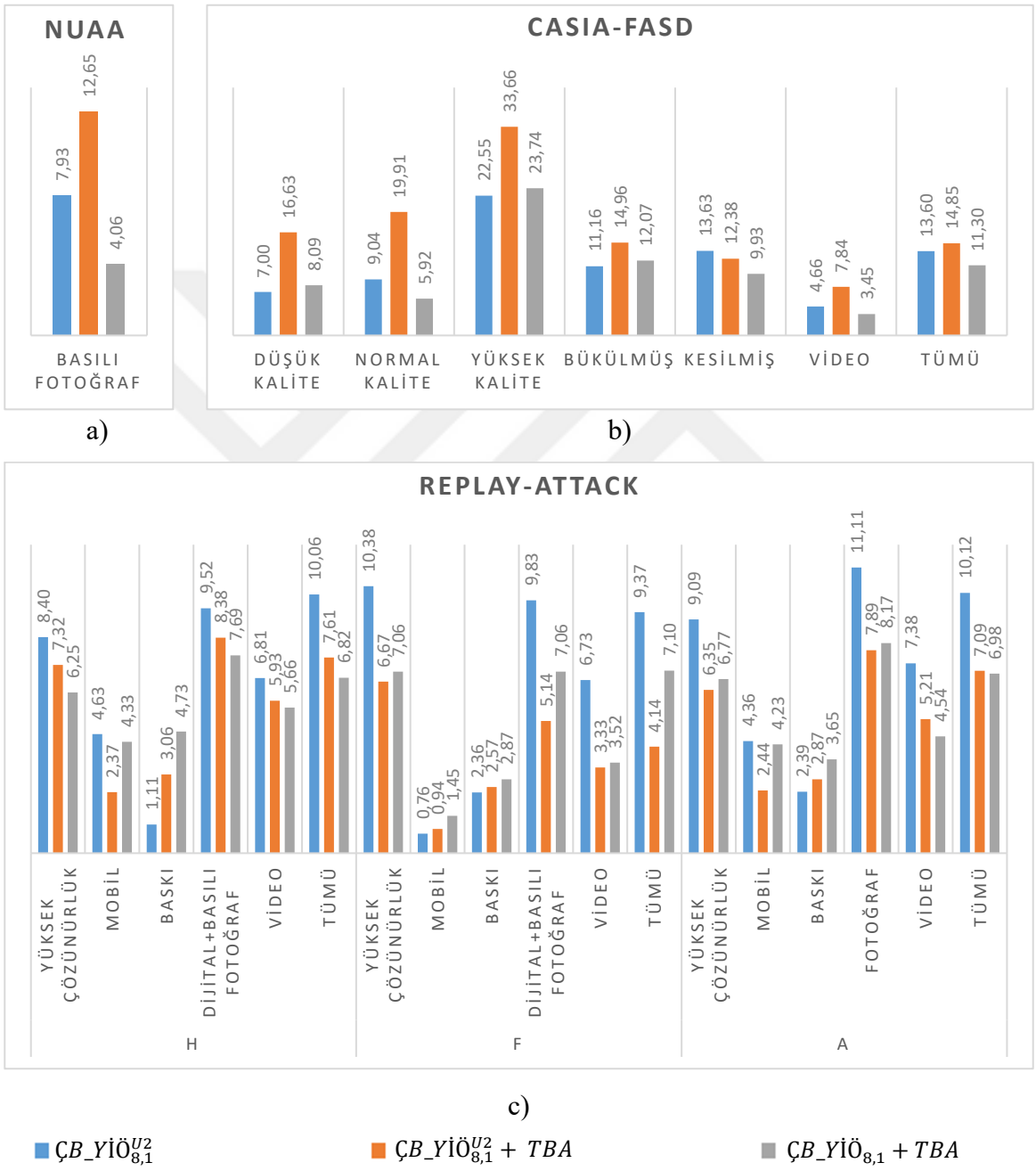
Şekil 36(a) incelendiğinde, NUAA veri setindeki tek senaryo olan basılı fotoğraf saldırısında $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1}^{U2}$ öznitelikler ile %7,93 HTER elde edilmişken, $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1} + TBA$ öznitelikleri kullanılarak bu değer %4,06'ya kadar indirilmiştir. İki yöntem arasında %48,8'lik bir fark bulunmaktadır. Bu durum, bir giriş görüntüsünden üretilen tüm YİÖ örüntülerinin, YST problemine olan etkisini ortaya koymaktadır. Bu test senaryosu için AUROC ve sınıflandırma doğruluğu değerleri sırasıyla %95,94 ve %95,88 olarak hesaplanmıştır.

Şekil 36(b), CASIA-FASD veri setindeki test senaryoları için elde edilen sonuçları içermektedir. Grafikten görüldüğü üzere, 7 test senaryosunun dördünde (normal kalite, kesilmiş fotoğraf, video oynatma ve tümü) $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1} + TBA$ öznitelikleri YST başarımını çeşitli oranlarda artırmaktadır.

Gerçek hayat uygulamalarında saldırganın hangi tipte saldırı gerçekleştireceğini tahmin etmek mümkün olmadığından, önerilen modellerin değerlendirmesinde tüm saldırı türlerinin bir arada bulunduğu test senaryoları oldukça önem kazanmaktadır. Bu sebeple, "tümü" senaryosuna odaklanıldığında, $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1}^{U2}$ öznitelikleriniin kullanımıyla %13,6 olarak elde edilen HTER değerinin, $\text{Ç}B\_Y\dot{I}\ddot{O}_{8,1} + TBA$ öznitelikleri ile %11,3 seviyesine gerilediği görülmektedir. Bu iki yöntem arasında %16,9'luk iyileşme elde edilmesi oldukça önemlidir. Bu test senaryosu için AUROC ve sınıflandırma doğruluğu değerleri sırasıyla %88,70 ve %89,91 olarak hesaplanmıştır.

Şekil 36(c), REPLAY-ATTACK veri setindeki test senaryoları için elde edilen sonuçları içermektedir. Sonuçlar incelendiğinde, özellikle kaynağın elde tutulduğu

H_Yüksek Çözünürlük, H_Dijital+Basılı Fotoğraf, H_Video ve H_Tümü saldırı türlerinde, $ÇB\_Y\dot{I}Ö_{8,1} + TBA$ özniteliklerinin kullanımının, YST başarımını artırdığı görülmektedir. Ek olarak A_Video (hem sabit konumlandırılmış kaynaktan hem de elde tutulan kaynaktan yapılan video oynatma saldırıları) senaryosunda da sahtecilik tespiti başarımı artmıştır.



Şekil 36. Geniş yüz bölgesinden üretilen özniteliklerin sınıflandırma sonuçları

Özellikle kaynağın elde tutulduğu saldırı türlerinin hepsini içeren H_Tümü test senaryosunda, $ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ öznitelikleri kullanılarak %7,61 HTER ile sahtecilik

tespiti yapılmışken, $ÇB\_Y\dot{I}Ö_{8,1} + TBA$ öznitelikleri kullanıldığında bu değer %6,82 seviyesine gerilemiştir. Bu durum, ilgili senaryo için %10,38'lik bir iyileşmeyi ifade etmektedir. Genelde saldırıların elle tutulan cihazlardan yapıldığı düşünüldüğünde, bu iyileşme oldukça önemlidir. Bu test senaryosu için AUROC ve sınıflandırma doğruluğu değerleri sırasıyla %93,18 ve %92,73 olarak hesaplanmıştır. Ayrıca, REPLAY-ATTACK veri setindeki tüm saldırı türlerini bir arada bulunduran A_Tümü test senaryosunda, $ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ öznitelikleri %7,09 HTER değeri üretmişken, $ÇB\_Y\dot{I}Ö_{8,1} + TBA$ öznitelikleri kullanılarak yapılan sınıflandırmada bu değer %6,98 olarak hesaplanmıştır. Bu durum, ilgili senaryo için %1,5'lik bir iyileşmeyi ifade etmektedir. Bu test senaryosu için AUROC ve sınıflandırma doğruluğu değerleri sırasıyla %92,85 ve %92,19 olarak hesaplanmıştır. Elde edilen bu sonuçlardan, $ÇB\_Y\dot{I}Ö_{8,1}$ özniteliklerinin YST probleminde bilgi taşıdığını ortaya koymaktadır. Tablo 5, önerilen yöntemin literatürdeki diğer yöntemler ile karşılaştırmasını göstermektedir.

Tablo 5. Önerilen ÇB_YİÖ$_{8,1}$+TBA yöntemine ait literatür karşılaştırması

YÖNTEM	NUAA		CASIA		REPLAY-ATTACK	
	EER (%)	HTER (%)	EER (%)	HTER (%)	EER (%)	HTER (%)
(Määttä vd., 2011)	2,9	-	-	-	-	-
(Chingovska vd., 2012)	-	19,03	-	18,17	-	15,16
(Määttä vd., 2012)	1,1	-	-	-	-	-
(J. Yang vd., 2013)	1,9	-	11,8	-	-	-
(Komulainen vd., 2013)	-	-	-	-	5,11	-
(Tirunagari vd., 2015)	-	-	-	21,75	-	3,75
(Boulkenafet vd., 2015)	-	-	6,2	-	0,4	2,9
(Tian & Xiang, 2016)	-	-	-	18,06	-	0,0
(I. Kim vd., 2017)	-	-	4,89	-	-	5,5
(G. B. De Souza vd., 2017)	1,8	1,7	-	-	-	-
(W. Zhang & Xiang, 2020)	-	-	5,56	-	0,0	0,0
(Shu vd., 2021)	-	-	1,11	-	0,0	0,0
Önerilen Yöntem	**0,17**	4,06	**0,22**	**11,30**	9,28	6,98

Tablo 5 incelendiğinde, NUAA ve CASIA-FASD veri setleri üzerinde önerilen yöntem ile elde edilen YST sonuçlarının, literatürdeki diğer yöntemlerden başarılı olduğu görülmektedir. Bu veri setleri geliştirme kümeleri içermediğinden, EER değerlerinin üretilmesi adına eğitim kümesinden 5 kat çapraz doğrulama yardımıyla geliştirme kümeleri elde edilmiş ve ardından test seti üzerinde sınıflandırma işlemleri gerçekleştirilmiştir. Elde edilen sonuçlara göre, NUAA veri seti için %0,17 EER ve %4,06 HTER, CASIA-FASD veri seti için %0,22 EER ve %11,3 HTER, REPLAY-ATTACK veri seti için ise %9,28 EER ve

%6,98 HTER değerleri için NUAA ve CASIA-FASD veri setleri için başarılı, REPLAY-ATTACK veri seti için ise karşılaştırılabilir olduğu söylenebilir.

Literatürdeki çalışmalar yüz görüntüsünün tamamı üzerinden YST gerçekleştirdiği için tabloda geniş yüz bölgesi ile elde edilen sonuçlara göre karşılaştırma yapılmıştır. Dolayısıyla, ağız, burun ve göz bölgelerinin YST başarımlarına ilişkin sonuçların kıyaslanabilmesi mümkün olmamıştır. Diğer yandan önceki çalışmalar incelendiğinde, genellikle bütün veri seti üzerinden üretilen sonuçların verildiği, çalışmadaki gibi çeşitli test senaryolarının çalışılmadığı görülmektedir. Bu bakımdan, çalışmada 26 farklı test senaryosuna ve 5 farklı yüz bölgesine göre YST başarımının incelenmesinin, bu konuda yapılacak diğer çalışmalar için faydalı olacağı düşünülmektedir.

3.3. Renk Kanallarının Yüz Sahteciliği Tespiti Başarımına Etkileri

Doku analizi tabanlı yüz sahteciliği tespiti üzerine yapılan çalışmalar, genellikle kullanılan doku tanımlayıcıların sınırlayıcı özellikleri sebebiyle gri renk uzayında gerçekleştirilmiştir. 2015 yılında yapılan bir çalışmada (Boulkenafet vd., 2015), farklı renk uzaylarından çıkarılan doku özniteliklerinin yüz sahteciliği tespitinde oldukça etkili olduğu görülmüştür. Bu sebeple, problemin çözümünde renk uzaylarının kullanımı bu tarihten itibaren yaygınlaşmıştır.

Çalışmanın bu aşamasında, bir önceki çalışmada en iyi sonuçların üretildiği geniş yüz bölgesinin giriş olarak kullanıldığı CASIA-FASD ve REPLAY-ATTACK veri setleri üzerinde, renk uzaylarına ait kanalların YST başarımına etkisi araştırılmıştır.

Giriş görüntüsü HSV, YC_bC_r ve L*a*b* renk uzaylarına dönüştürülmüş, ardından bu renk uzaylarının tüm kanallarından $ÇS\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelikleri ($ÇS\_Y\dot{I}Ö_{8,1_H}^{U2}$, $ÇS\_Y\dot{I}Ö_{8,1_S}^{U2}$, $ÇS\_Y\dot{I}Ö_{8,1_V}^{U2}$, $ÇS\_Y\dot{I}Ö_{8,1_Y}^{U2}$, $ÇS\_Y\dot{I}Ö_{8,1_{C_b}}^{U2}$, $ÇS\_Y\dot{I}Ö_{8,1_{C_r}}^{U2}$, $ÇS\_Y\dot{I}Ö_{8,1_{L^*}}^{U2}$, $ÇS\_Y\dot{I}Ö_{8,1_{a^*}}^{U2}$, $ÇS\_Y\dot{I}Ö_{8,1_{b^*}}^{U2}$) üretilmiştir. Daha sonra bu öznitelikler, ikili ve üçlü kombinasyonlar şeklinde birleştirilmiş ve üretilen bütünleşik öznitelik vektörleri, hesaplama ve depolama maliyetlerinin yanı sıra sınıflandırma sürelerini de azaltmak için TBA kullanılarak küçültülmüştür. Elde edilen veri kümesi SVM yardımıyla sınıflandırılmış ve sahtecilik tespiti gerçekleştirilmiştir. $ÇS\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelik vektörlerinin üretiminde, $1 \times 1, 2 \times 2$ ve 4×4 bölgelerden elde edilen $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelikleri kullanılmıştır.

Bir önceki çalışmada REPLAY-ATTACK veri seti için uygulanan deneylerin yoğunluğu sebebiyle bu çalışmada, ilgili veri setindeki saldırı türleri gruplanmıştır. Böylece,

18 test senaryosu yerini 6 test senaryosuna bırakmıştır. Ancak bu durum, karma saldırı türlerini içermesi sebebiyle bir önceki duruma göre daha zorludur. Tablo 6, çalışmada gerçekleştirilen deneylerde kullanılan örnek sayılarını içermektedir.

Tablo 6. Deneylerde kullanılan örnek sayıları

Veri Seti	Saldırı Türü	Gerçek Örnek Sayıları			Sahte Örnek Sayıları		
		Eğitim	Geliştirme	Test	Eğitim	Geliştirme	Test
CASIA-FASD	Düşük kalite	3.140	-	5.297	11.015	-	16.088
	Normal Kalite	3.197		4.830	11.231		16.085
	Yüksek Kalite	4.577		5.782	11.846		17.291
	Bükülmüş Fotoğraf	10.914		15.909	12.859		19.165
	Kesilmiş Fotoğraf				9.499		14.731
	Video Oynatma				11.734		15.568
	Tümü				34.092		49.464
REPLAY-ATTACK	Yüksek Çözünürlük	7.185	7.197	9.596	9.518	9.284	12.279
	Mobil				9.420	9.168	11.815
	Basılı Fotoğraf				4.765	4.738	6.236
	Dijital Fotoğraf				14.307	14.031	18.360
	Video Oynatma				9.396	9.159	11.970
	Tümü				23.703	23.190	30.330

Tablo 7, CASIA-FASD veri kümesindeki 7 farklı saldırı senaryosu ve REPLAY-ATTACK veri kümesindeki 6 farklı saldırı senaryosu için elde edilen EER, HTER ve AUROC değerlerini içermektedir.

Tablo 7. ÇS_YİÖ$_{HSV}$, ÇS_YİÖ$_{YCbCr}$, ÇS_YİÖ$_{L*a*b*}$ özniteliklerinin CASIA-FASD ve REPLAY-ATTACK veri setlerindeki YST başarımları

Veri Seti	Saldırı Türü	Öznitelik Türü					
		ÇS_YİÖ$_{HSV}$		ÇS_YİÖ$_{YCbCr}$		ÇS_YİÖ$_{L*a*b*}$	
		EER	AUROC	EER	AUROC	EER	AUROC
CASIA-FASD	Düşük Kalite	5,10	96,22	4,40	94,81	6,53	94,71
	Normal Kalite	0,11	99,58	2,00	98,22	2,35	97,73
	Yüksek Kalite	5,15	90,10	6,89	90,60	5,92	92,49
	Bükülmüş Fotoğraf	3,75	96,25	6,06	94,17	4,41	94,84
	Kesilmiş Fotoğraf	2,96	96,01	4,87	94,33	3,01	95,82
	Video Oynatma	1,38	99,02	2,78	97,43	2,68	97,31
	Tümü	2,46	96,38	6,83	92,93	4,49	95,43
		HTER	AUROC	HTER	AUROC	HTER	AUROC
REPLAY-ATTACK	Yüksek Çözünürlük	0,84	99,16	1,90	98,10	1,17	98,84
	Mobil	0,00	100,00	1,12	98,88	3,11	96,90
	Basılı Fotoğraf	1,81	98,19	0,56	99,44	1,11	98,89
	Dijital Fotoğraf	0,80	99,20	1,49	98,51	0,85	99,15
	Video Oynatma	0,62	99,38	2,37	97,63	3,31	96,69
	Tümü	0,71	99,29	1,28	98,72	0,70	99,30

Tablodan görüldüğü gibi genel olarak saldırı senaryolarında, HSV renk uzayının kanallarından çıkarılan ÇS_YİÖ histogramlarının birleşiminden meydana gelen öznitelik vektörleri daha başarılıdır. Bunun yanında, bazı senaryolarda YC_bC_r ve L*a*b* renk

uzaylarından çıkarılan özniteliklerin de başarılı olduğu görülmektedir. Bu nedenle, bu renk uzaylarının kanallarının çeşitli birleşimlerinde üretilen öznitelik vektörleri ile YST başarımının artırılması hedeflenmiştir. Tablo 8 ve Tablo 9, sırasıyla CASIA-FASD ve REPLAY-ATTACK veri kümeleri için renk kanallarının kombinasyonlu birleşimlerinden üretilen öznitelik vektörleri ile elde edilen YST başarımlarını içermektedir.

Tablo 8. CASIA-FASD veri setinde renk kanalı kombinasyonlarının YST başarımı (EER)

Kullanılan Renk Kanalları	**Saldırı Türü**						
	Düşük Kalite	**Normal Kalite**	**Yüksek Kalite**	**Bükülmüş Fotoğraf**	**Kesilmiş Fotoğraf**	**Video Oynatma**	**Tümü**
HSV + L*a*b*	5,70	**0,12**	**4,97**	2,92	2,09	1,50	2,43
HSV + YC_bC_r	5,09	**0,12**	5,60	**2,45**	**2,04**	**1,31**	**2,22**
L*a*b* + YC_bC_r	5,25	2,29	6,30	4,44	2,82	2,60	4,89
HS + a*b*	4,43	0,90	5,40	3,76	2,53	2,55	2,97
HS + C_bC_r	**4,27**	0,89	6,77	3,84	2,93	2,37	3,24
a*b* + C_bC_r	6,21	3,91	10,81	5,84	5,17	4,97	6,05
HS + a*b* + C_bC_r	4,71	1,07	7,14	3,75	2,76	2,73	3,03
HSV + L*a*b* + YC_bC_r	5,93	0,29	5,55	2,75	2,07	1,53	2,55

Tablo 8 incelendiğinde, CASIA-FASD veri kümesinde genel olarak HSV ve YC_bC_r renk uzaylarının kanallarından üretilen öznitelik birleşimlerinin daha iyi bir YST başarımı sağladığını göstermektedir. Yüksek kalite saldırı senaryosunda HSV ve L*a*b* uzaylarından üretilen öznitelik vektörleri daha yüksek bir başarımla YST gerçekleştirmektedir. Düşük kalite saldırı senaryosunda ise renk uzaylarının H, S, C_b ve C_r kanallarından üretilen öznitelik vektörlerinin birleşiminin YST başarımı daha yüksektir. Bu veri kümesinde en iyi YST başarımı %0,11 EER ile Normal kalite saldırı senaryosunda elde edilmiştir. CASIA-FASD veri kümesi için tüm saldırı türlerinin birlikte değerlendirildiği test senaryosunda ise ÇS_YİÖ$_{HSV}$+ ÇS_YİÖ$_{YCbCr}$ öznitelikleri ile %2,22 EER değerine sahip YST başarımı elde edilmiştir. Bu senaryo için elde edilen AUROC değeri ise %97,10'dur.

Tablo 9. REPLAY-ATTACK veri setinde renk kanalı kombinasyonlarının YST başarımı (HTER)

Kullanılan Renk Kanalları	**Saldırı Türü**					
	Yüksek Çözünürlük	**Mobil**	**Basılı Fotoğraf**	**Dijital Fotoğraf**	**Video Oynatma**	**Tümü**
HSV + L*a*b*	**0,36**	0,93	0,23	0,28	0,50	0,20
HSV + YC_bC_r	0,75	0,00	0,91	0,63	0,26	0,54
L*a*b* + YC_bC_r	1,38	1,16	0,35	0,96	2,85	0,60
HS + a*b*	1,02	0,10	0,10	0,03	0,33	0,05
HS + C_bC_r	1,20	**0,00**	0,14	0,11	**0,04**	0,12
a*b* + C_bC_r	1,29	0,55	4,38	1,70	1,44	1,23
HS + a*b* + C_bC_r	1,21	0,00	0,65	**0,00**	0,22	**0,03**
HSV + L*a*b* + YC_bC_r	0,65	0,13	**0,04**	0,26	1,09	0,24

Tablo 9, REPLAY-ATTACK veri kümesi için elde edilen YST başarımlarını içermektedir. Tablo incelendiğinde, çeşitli kanal birleşimleriyle daha iyi sonuçlar elde edildiği görülmektedir. Yüksek çözünürlük saldırısında H, S, V, L*, a* ve b* kanal özniteliklerinin birleşimi, mobil ve video oynatma saldırılarında H, S, C_b ve C_r kanal özniteliklerinin birleşimi, basılı fotoğraf saldırısında H, S, a* ve b* kanal özniteliklerinin birleşimi, dijital fotoğraf saldırısında ise H, S, a*, b*, C_b ve C_r kanal özniteliklerinin birleşimi en iyi YST başarımı sağlamıştır. Tüm saldırı türlerinin birlikte değerlendirildiği test senaryosunda ÇS_YİÖ$_{HS}$ + ÇS_YİÖ$_{a*b*}$ + ÇS_YİÖ$_{CbCr}$ öznitelikleri ile %0,03 HTER değerine sahip YST başarımı ve %99,97 AUROC değeri elde edilmiştir.

Çalışmada tüm deneyler i7 7700k, 3.60 GHz işlemci ve 8GB DDR4 belleğe sahip bilgisayarda gerçekleştirilmiştir. Tablo 10, önerilen YST yönteminin veri setleri için eğitim süreleri ve bir test görüntüsü için karar verme sürelerini içermektedir.

Tablo 10. Önerilen yöntemin veri setleri için eğitim süreleri ve bir test görüntüsü için karar verme süresi

Kullanılan Renk Kanalları	**Öznitelik Türü**	**Eğitim Seti**		**Geliştirme Seti**		**Test (1 Görsel)**
		Görüntü Sayısı	**İşlem Süresi (sn.)**	**Görüntü Sayısı**	**İşlem Süresi (sn.)**	**İşlem Süresi (sn.)**
CASIA-FASD	ÇS_YİÖ$_{HSV}$+ ÇS_YİÖ$_{YCbCr}$	45.006	6.073,42	-	-	0,01
REPLAY-ATTACK	ÇS_YİÖ$_{HS}$ + ÇS_YİÖ$_{a*b*}$ + ÇS_YİÖ$_{CbCr}$	30.888	1.973,96	30.387	220,38	0,01

Çizelgede görüldüğü üzere, sistem eğitimi tamamlandıktan sonra bir görüntünün gerçek/sahte olarak değerlendirilmesi 0,01 saniye içerisinde gerçekleştirilmektedir. Bu sonuç, özellikle gerçek zamanlı sistem tasarımları için oldukça önemlidir. Tablo 11, önerilen YST yönteminin literatürde yapılan çalışmalarla karşılaştırılmasını içermektedir.

Tablo 11. Önerilen renk kanalı tabanlı yönteme ait literatür karşılaştırması

Yöntem	**CASIA-FASD**	**REPLAY-ATTACK**	
	EER	**EER**	**HTER**
Color LBP (YC_bC_r + HSV) (Boulkenafet vd., 2015)	6,20	0,40	2,90
CoALBP + LPQ (YC_bC_r + HSV) (Boulkenafet vd., 2016b)	2,10	0,40	2,80
LBP + Color Moment (Patel vd., 2016)	5,88	-	14,60
Color SURF (YC_bC_r + HSV) (Boulkenafet vd., 2016a)	2,80	0,10	2,20
LGBP (F. Peng vd., 2018b)	7,06	7,08	8,29
LBP + GS-LBP (F. Peng vd., 2018b)	3,05	0,69	3,31
Color Texture (Boulkenafet vd., 2018)	4,60	1,20	4,20
CTMF + SVM-RFE (L.-B. Zhang vd., 2018)	8,00	4,00	4,40
LDN-TOP (Zhou vd., 2021)	**0,37**	-	3,13
Önerilen Yöntem	2,22	**0,01**	**0,03**

Tablodan görüldüğü üzere, önerilen yöntem kullanılarak REPLAY-ATTACK veri seti için gerçekleştirilen YST uygulamasında, renk uzaylarından doku bilgisi üretme konusunda HSV ve YC_bC_r renk uzaylarının birleşiminden üretilen YİÖ özniteliklerini kullanan çalışmalardan daha iyi başarım elde edilmiştir. Literatürdeki çalışmalarda elde edilen sonuçlara paralel olarak, bu çalışmada da renk uzaylarının kromatik bileşenlerinden çıkarılan özniteliklerin yüz sahteciliği tespitinde daha başarılı olduğu görülmüştür. Bu durum, kromatik bileşenlerin, parlaklık bileşenlerine göre daha fazla doku bilgi taşımasının bir sonucudur.

Önerilen yöntem CASIA-FASD veri kümesi için LDN-TOP yönteminden daha düşük bir başarım göstermiştir. LDN-TOP yöntemi, peş peşe gelen çok sayıdaki (15, 25, …, 55, 65, 75) çerçeveyi işleme sokarak öznitelik çıkarmaktadır. Bu yöntemde çerçeve sayısının artışına bağlı olarak iyileşme sağlanmaktadır fakat CASIA-FASD veri kümesinde bazı videolar LDN-TOP özniteliklerini çıkarmak için yeterli sayıda çerçeveye sahip değildir (Zhou vd., 2021). Bu çalışmada önerilen yöntemde sadece bir çerçeve üzerinden gerçek/sahte kararı verilmektedir. Ayrıca LDN-TOP yöntemine göre hesaplama karmaşıklığı çok daha düşük, yorumlanabilme özelliği ise daha yüksektir. Diğer yandan CASIA veri kümesinde yüz bölgesinin belirlenebildiği tüm çerçeveler deneylerde kullanılmıştır. Elde edilen başarım literatürle kıyaslanabilir niteliktedir. Ek olarak, çoğu çalışmada veri kümesindeki hangi çerçevelerin kullanıldığı bilgisi verilmemiştir. Bu durum sonuçların adil bir şekilde kıyaslanmasını doğrudan etkilemektedir. Ayrıca, önerilen yöntem, daha iyi başarım raporlayan yöntemlere göre çok daha basittir ve düşük hesaplama karmaşıklığına sahiptir.

3.4. Yüz Ağırlıklı Çok Renkli Çok Seviyeli Yerel İkili Örüntüler Yönteminin Yüz Sahteciliği Tespiti Başarımı

Önceki bölümlerde, veri setleri üzerinde gerçekleştirilen deneyler sonucunda, farklı test senaryolarında farklı bölgesel YİÖ özniteliklerinin ($ÇB\_YİÖ_{8,1}^{U2}$) daha iyi performans gösterdiği görülmüştür. Bu durumun dışında genel bir değerlendirme yapıldığında, çok renk uzayının bir arada kullanılmasıyla gerçekleştirilen deneylerden elde edilen sonuçların, tek renk uzayından elde edilen sonuçlardan daha başarılı olduğu anlaşılmaktadır. Bu sebeple, her iki veri setinde YST başarımının arttırılması amacıyla ÇRÇS_YİÖ yönteminden üretilen özniteliklerin kullanılması önerilmiştir. Bu yöntem, problemin çözümüne dair oldukça fazla bilgiyi içeriyor olmasına rağmen, üretim sürecindeki hesaplama maliyetleri ve öznitelik

vektörünün boyutu, önerilen yöntemin uygulama alanlarının kısıtlanmasına neden olmaktadır. Bu nedenle çalışmada, daha düşük öznitelik vektör boyutuna ve daha az hesaplama karmaşıklığına sahip ancak bazı senaryolarda çok daha iyi sonuçlar üreten YA_ÇRÇS_YİÖ özniteliklerinin kullanımı önerilmiştir.

Çalışmada giriş görüntüsü HSV ve L*a*b* renk uzaylarına dönüştürülmüş, ardından bu renk uzaylarının tüm kanallarından $ÇS\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelikleri ($ÇS\_Y\dot{I}Ö_{8,1_H}^{U2}$, $ÇS\_Y\dot{I}Ö_{8,1_S}^{U2}$, $ÇS\_Y\dot{I}Ö_{8,1_V}^{U2}$, $ÇS\_Y\dot{I}Ö_{8,1_{L*}}^{U2}$, $ÇS\_Y\dot{I}Ö_{8,1_{a*}}^{U2}$, $ÇS\_Y\dot{I}Ö_{8,1_{b*}}^{U2}$) üretilmiştir. Daha sonra bu öznitelikler uç uca eklenmiş ve üretilen bütünleşik öznitelik vektörleri, hesaplama ve depolama maliyetlerinin yanı sıra sınıflandırma sürelerini de azaltmak için TBA kullanılarak küçültülmüştür. Elde edilen $YA\_ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ öznitelik vektörleri SVM yardımıyla sınıflandırılmış ve sahtecilik tespiti gerçekleştirilmiştir. $YA\_ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelik vektörlerinin üretiminde, $1 \times 1, 2 \times 2$ ve 1×1 bölgelerden elde edilen $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ özniteliklerinin birleşimi ile oluşturulan $ÇS\_Y\dot{I}Ö_{8,1}^{U2}$ kullanılmıştır. Bu işlem tüm renk uzayları için tekrarlandığında $ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelikleri üretilir.

Uygulamanın ilk aşamasında, RGB uzayındaki giriş görüntüleri gri seviyeye dönüştürülmüş, tek kanaldan ÇB_YİÖ öznitelikleri çıkarılmış ve her bir bölgesel öznitelik vektörünün YST başarımı ayrı ayrı hesaplanmıştır. Bu süreçte $1 \times 1, 2 \times 2$ ve 4×4 $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelikleri kullanılmıştır. İkinci aşamada renkli görüntüler RGB kanallarına ayrılmış her kanaldan çıkarılan $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ özniteliklerin birleşimi yardımıyla YST başarımı incelenmiştir. Burada vektör kümesi, tüm renk kanallarından aynı bölgesel ayrıştırma ile elde edilen özniteliklerin uç uca eklenmesi ile oluşturulmuştur. Bir örnekle açıklamak gerekirse, giriş görüntüsünün R, G ve B kanallarına 2×2 bölgesel ayrıştırma işlemi uygulandıktan sonra, bu bölgelerden tüm kanallar için üretilen $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ özniteliklerin birleştirilmesi ile öznitelik vektörü elde edilmektedir. Aynı şekilde renk uzayının diğer kanallarından üretilen 1×1 ve 4×4 $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelikleri kullanılarak, renk uzayının YST başarımı hesaplanmıştır. Daha sonra RGB uzayındaki görüntüler sırasıyla HSV ve L*a*b* renk uzayına dönüştürülmüş ve bu renk uzaylarındaki kanallardan aynı şekilde $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelikleri çıkarılarak birleştirilmiş ve YST başarımları incelenmiştir. Tablo 12 ve

Tablo 13 sırasıyla, CASIA-FASD veri setindeki 7 farklı test senaryosu ve REPLAY-ATTACK veri setindeki 18 farklı test senaryosu için elde edilen en iyi ($1 \times 1, 2 \times 2$ ve 4×4

bölgesel ayrıştırmalardan üretilen $\text{ÇB\_Y İÖ}_{8,1}^{U2}$ öznitelikleri ile yapılan sınıflandırma sonuçları) EER ve HTER sonuçlarını içermektedir.

Tablo 12. CASIA-FASD veri setinde tek renk uzayından üretilen ÇB_YİÖ ve birden çok renk uzayından üretilen ÇB_YİÖ özniteliklerinin YST başarımları (EER)

Renk Uzayı	Saldırı Türü						
	Düşük Kalite	**Normal Kalite**	**Yüksek Kalite**	**Bükülmüş Fotoğraf**	**Kesilmiş Fotoğraf**	**Video Oynatma**	**Tümü**
Gri	6,27 (4x4)	4,50 (4x4)	13,60 (4x4)	11,11 (4x4)	10,39 (4x4)	2,21 (4x4)	11,51 (4x4)
RGB	6,29 (4x4)	3,94 (4x4)	10,87 (4x4)	9,26 (4x4)	9,80 (4x4)	2,02 (4x4)	9,16 (4x4)
HSV	**5,11** (4x4)	0,17 (4x4)	5,28 (4x4)	3,16 (2x2)	3,48 (2x2)	**1,43** (4x4)	2,66 (4x4)
L*a*b*	6,48 (4x4)	2,53 (2x2)	6,26 (4x4)	4,55 (4x4)	2,70 (2x2)	2,76 (2x2)	4,91 (2x2)
HSV+L*a*b*	5,59 (2x2)	**0,11** (2x2)	**3,90** (1x1)	**2,55** (1x1)	**2,09** (2x2)	1,51 (4x4)	**2,31** (2x2)

Tablo 13. REPLAY-ATTACK veri setinde tek renk uzayından üretilen ÇB_YİÖ ve birden çok renk uzayından üretilen ÇB_YİÖ özniteliklerinin YST başarımları (HTER)

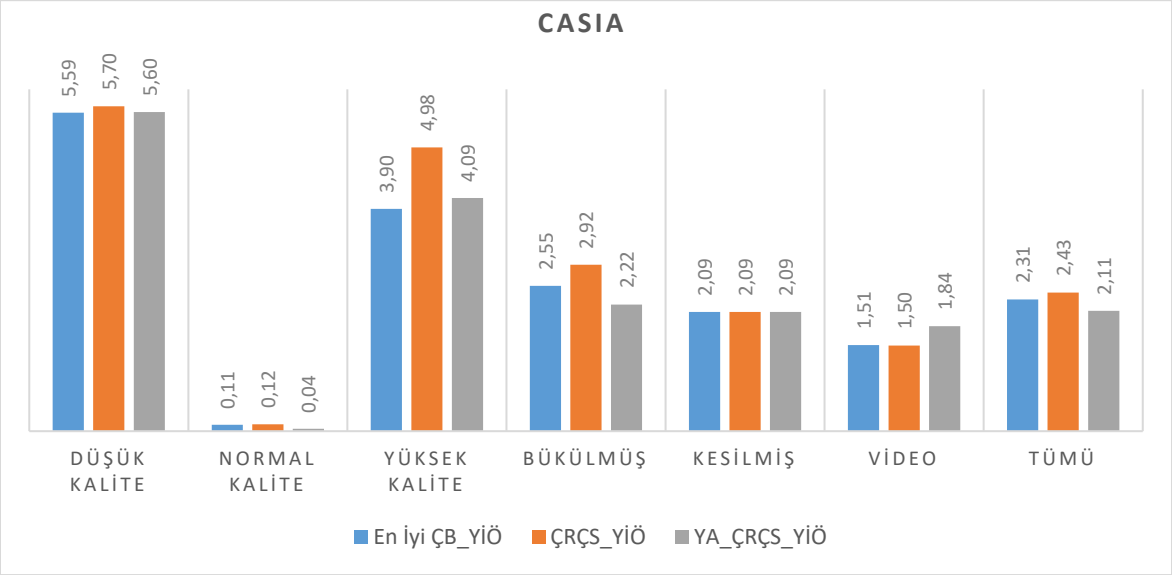
Cihaz Konumu	Renk Uzayı	Saldırı Türü					
		Yüksek Çözünürlük	**Mobil**	**Basılı Fotoğraf**	**Dijital Fotoğraf**	**Video Oynatma**	**Tümü**
H	Gri	5,35 (2x2)	2,17 (4x4)	2,83 (4x4)	7,56 (2x2)	4,49 (2x2)	5,90 (2x2)
	RGB	3,40 (4x4)	0,67 (4x4)	2,61 (4x4)	4,39 (2x2)	2,53 (4x4)	2,42 (2x2)
	HSV	0,68 (4x4)	0,09 (4x4)	1,29 (4x4)	1,38 (4x4)	1,22 (1x1)	1,16 (2x2)
	L*a*b*	1,26 (4x4)	1,83 (1x1)	1,35 (4x4)	1,22 (4x4)	1,57 (2x2)	1,24 (4x4)
	HSV+L*a*b*	**0,43** (4x4)	**0,02** (1x1)	**0,00** (4x4)	**0,51** (4x4)	**0,02** (1x1)	**0,64** (4x4)
F	Gri	5,55 (2x2)	0,25 (4x4)	1,81 (4x4)	5,25 (2x2)	2,93 (2x2)	4,30 (2x2)
	RGB	2,29 (2x2)	0,38 (4x4)	1,42 (4x4)	3,56 (2x2)	1,75 (2x2)	2,55 (2x2)
	HSV	0,66 (2x2)	0,01 (4x4)	1,44 (4x4)	0,82 (4x4)	0,09 (4x4)	0,64 (4x4)
	L*a*b*	0,58 (2x2)	0,25 (4x4)	**0,00** (4x4)	1,09 (2x2)	1,30 (2x2)	0,93 (2x2)
	HSV+L*a*b*	**0,06** (2x2)	**0,00** (1x1)	0,18 (4x4)	**0,15** (2x2)	**0,00** (2x2)	**0,07** (2x2)
A	Gri	4,94 (2x2)	2,14 (4x4)	2,44 (4x4)	7,15 (2x2)	4,00 (2x2)	5,93 (2x2)
	RGB	3,21 (4x4)	0,57 (2x2)	1,97 (4x4)	3,87 (2x2)	2,10 (4x4)	2,62 (2x2)
	HSV	0,67 (4x4)	**0,00** (2x2)	1,75 (4x4)	0,96 (4x4)	0,67 (1x1)	0,89 (2x2)
	L*a*b*	1,05 (2x2)	2,24 (4x4)	0,70 (4x4)	0,78 (4x4)	1,58 (2x2)	1,07 (2x2)
	HSV+L*a*b*	**0,37** (2x2)	**0,00** (1x1)	**0,06** (4x4)	**0,27** (4x4)	**0,03** (2x2)	**0,36** (4x4)

Tablolardan görüldüğü üzere, HSV ve L*a*b* renk uzaylarının her birinden üretilen $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelikleri, gri seviye ve RGB renk uzayına göre daha iyi YST başarımına sahiptir. Genel olarak test senaryolarında HSV renk uzayından çıkarılan öznitelikler daha başarılı iken, L*a*b* renk uzayında çıkarılan özniteliklerin de başarılı olduğu senaryoların olduğu görülmektedir. Fakat bu sonuçların farklı $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelikleriyle elde edildiği görülmektedir. Bu durum, hangi bölgesel ayrıştırmanın daha iyi sonuç vereceğinin tam olarak kestirilememesine ve bu sebeple, gerçek zamanlı bir sistemde hedeflenen modelin oluşturulmasına engel olmaktadır.

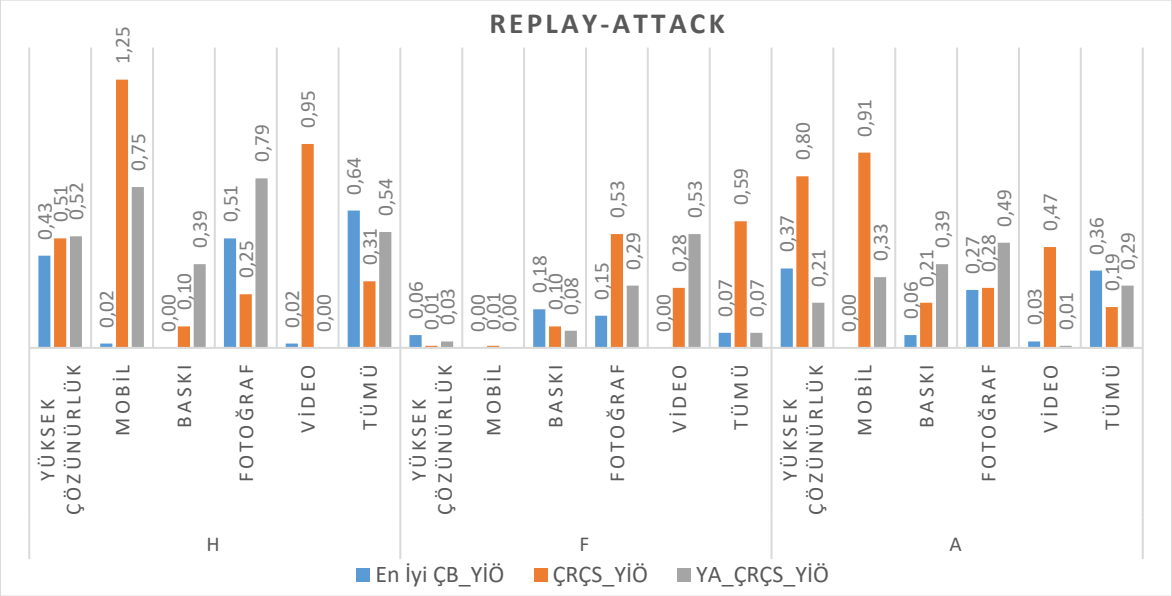
Tüm renk uzayları kendi içlerinde değerlendirildiğinde, CASIA-FASD veri setindeki 7 test senaryosunun 6'sında HSV renk uzayında, birinde ise L*a*b* renk uzayından çıkarılan özniteliklerin en iyi sonuçları ürettiği tablodan görülmektedir. REPLAY-ATTACK veri setinde ise 12 test senaryosunda HSV renk uzayının, 6 senaryoda ise L*a*b* renk uzayının en iyi sonuçları ürettiği anlaşılmaktadır. Bu nedenle, önerilen modelde bu iki renk uzayının birleştirilmesi ile YST başarımının artırılması hedeflenmiştir.

HSV ve L*a*b* renk uzaylarının birleşiminden yani 6 renk kanalından üretilen $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ öznitelikleri, CASIA-FASD veri setindeki 7 test senaryosunun 5'inde (tümü senaryosu dahil), REPLAY-ATTACK veri setindeki 18 test senaryosunun 17'sinde (tümü dahil) YST başarımının iyileşmesini sağlamıştır. Fakat yine sonuçlardan anlaşılacağı üzere, farklı test senaryolarında farklı $ÇB\_Y\dot{I}Ö_{8,1}^{U2}$ özniteliklerinin başarımı söz konusudur. Şekil 37, CASIA-FASD veri setinde $En\ iyi\ ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$, $ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ ve $YA\_ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ öznitelikleri ile gerçekleştirilen YST başarımlarının karşılaştırmasını göstermektedir.

Grafikte görüldüğü üzere, CASIA-FASD veri setinde $ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ öznitelikleri yalnızca bir test senaryosunda (Video) diğer yöntemlerden daha iyi EER sonuçları üretmiş, $YA\_ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ öznitelikleri ise üç test senaryosunda (Normal Kalite, Bükülmüş Fotoğraf ve Tümü) daha iyi başarım göstermiştir. Bir test senaryosunda (Kesilmiş) tüm öznitelik çıkarma yöntemleri en iyi başarımı yakalarken, bir test senaryosunda (Düşük Kalite) ise $YA\_ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ öznitelikleriyle en iyi değerle neredeyse aynı bir başarım sonucu elde edilmiştir. Şekil 38, REPLAY-ATTACK veri setinde $En\ iyi\ ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$, $ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ ve $YA\_ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ öznitelikleri ile gerçekleştirilen YST başarımlarının karşılaştırmasını göstermektedir.



Şekil 37. CASIA-FASD veri setinde En iyi ÇB_YİÖ, ÇRÇS_YİÖ ve YA_ÇRÇS_YİÖ özniteliklerinin YST başarımları (EER)



Şekil 38. REPLAY-ATTACK veri setinde En iyi ÇB_YİÖ, ÇRÇS_YİÖ ve YA_ÇRÇS_YİÖ özniteliklerinin YST başarımları (HTER)

Şekil 38 incelendiğinde, REPLAY-ATTACK veri setinde gerçekleştirilen 4 test senaryosunda (H_Fotoğraf, H_Tümü, F_Yüksek Çözünürlük, A_Tümü) $ÇRÇS\_YİÖ_{8,1}^{U2} + TBA$ özniteliklerinin, 6 test senaryosunda (H_Video, F_Mobil, F_Baskı, F_Tümü,

A_Yüksek Çözünürlük, A_Video) ise $YA\_ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ özniteliklerinin daha iyi YST başarımı sergilediği görülmektedir. Kalan 8 test senaryosunun ikisinde (H_Yüksek Çözünürlük, A_Fotoğraf), önerilen yöntem ile en yüksek başarım değerine çok yakın sonuçlar elde edilmiş, 6'sında (H_Mobil, H_Baskı, F_Fotoğraf, F_Video, A_Baskı, A_Mobil) ise en iyi değere ulaşılamamıştır. Ancak bu senaryolarda sonuçlar farklı bölgesel ayrıştırmalar ile elde edilmiştir. Dolayısıyla, $En\ iyi\ ÇB\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ yöntemi kullanılarak elde edilen sonuçların, $ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ ve $YA\_ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ yöntemleri ile karşılaştırılması adil olmayacaktır. Bu sebeple, tüm senaryolar önerilen yöntemler üzerinden karşılaştırıldığında, 9 test senaryosunda $ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ özniteliklerinin, kalan 9 test senaryosunda ise $YA\_ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ özniteliklerinin başarılı olduğu sonucuna varılmaktadır.

Gerçek dünya uygulamalarında atakların hangi cihazdan, hangi koşullarda ve hangi türde yapılacağının bilinmesi mümkün değildir. Bu sebeple, veri setlerinin tüm saldırı türlerini barındıran atak türleri üzerinden değerlendirmelerin yapılması önemlidir.

Önerilen $YA\_ÇRÇS\_Y\dot{I}Ö_{8,1}^{U2} + TBA$ özniteliklerinin CASIA-FASD ve REPLAY-ATTACK veri setlerinde bulunan tüm saldırı senaryoları altındaki performansı incelendiğinde, %2,11 EER ve %0,19 HTER değerleri ürettiği görülmektedir.

Tablo 14, önerilen yöntemin CASIA-FASD ve REPLAY-ATTACK veri setleri üzerinde tüm saldırı senaryolarının bir arada olduğu atak türü için elde edilen AUROC, Kesinlik, Duyarlılık ve F1-Skoru yüzdelik sonuçlarını içermektedir.

Tablo 14. CASIA-FASD ve REPLAY-ATTACK veri setlerinde önerilen yöntemin AUROC, Kesinlik, Duyarlılık, F1-Skoru sonuçları

Veri Seti	Yöntem	AUROC	Kesinlik	Duyarlılık	F1-Skoru
CASIA-FASD	ÇRÇS_YİÖ	97,23	96,25	95,66	95,96
	YA_ÇRÇS_YİÖ	98,03	97,06	97,00	9703
REPLAY-ATTACK	ÇRÇS_YİÖ	99,81	98,83	100,00	99,41
	YA_ÇRÇS_YİÖ	99,72	98,23	100,00	99,11

Tüm deneyler, 32G RAM ve Intel (R) Core (TM) i7-10700K CPU @ 3,80GHz işlemciye sahip standart bir masaüstü bilgisayarda, Windows 10 üzerinde çalışan Python 3,7 yazılım platformunda gerçekleştirilmiştir. Tablo 15, CASIA-FASD ve REPLAY-ATTACK veri setlerinde görüntü ön işleme, öznitelik çıkarma ve tek bir görüntü için saniye türünden hesaplama sürelerini içermektedir.

Tablo 15. CASIA-FASD ve REPLAY-ATTACK veri setlerinde görüntü başına hesaplama süreleri

Veri Seti	Yöntem	Ön İşleme	Öznitelik Çıkarma	Sınıflandırma	Toplam
CASIA-FASD	ÇRÇS_YİÖ	0,1316	0,1820	0,0080	0,3216
	YA_ÇRÇS_YİÖ	0,1316	0,1690	0,0020	0,3026
REPLAY-ATTACK	ÇRÇS_YİÖ	0,0379	0,1745	0,0040	0,2164
	YA_ÇRÇS_YİÖ	0,0379	0,1676	0,0010	0,2065

Tablo 16, önerilen YA_ÇRÇS_YİÖ YST yönteminin literatürde yapılan benzer çalışmalarla karşılaştırılmasını göstermektedir.

Tablo 16. Önerilen yüz ağırlıklı yönteme ait literatür karşılaştırması

Yöntem	CASIA-FASD	REPLAY-ATTACK	
	EER	EER	HTER
Color LBP (YC_bC_r + HSV) (Boulkenafet vd., 2015)	6,20	0,40	2,90
CoALBP + LPQ (YC_bC_r + HSV) (Boulkenafet vd., 2016b)	2,10	0,40	2,80
LBP + Color Moment (Patel vd., 2016)	5,88	-	14,60
Color SURF (YC_bC_r + HSV) (Boulkenafet vd., 2016a)	2,80	0,10	2,20
LGBP (F. Peng vd., 2018b)	7,06	7,08	8,29
LBP + GS-LBP (F. Peng vd., 2018b)	3,05	0,69	3,31
Color Texture (Boulkenafet vd., 2018)	4,60	1,20	4,20
CTMF + SVM-RFE (L.-B. Zhang vd., 2018)	8,00	4,00	4,40
LDN-TOP (Zhou vd., 2021)	**0,37**	-	3,13
ED-LBP (Shu vd., 2021)	1,11	**0,00**	**0,00**
Önerilen Yöntem (ÇRÇS_YİÖ)	2,43	0,06	0,19
Önerilen Yöntem (YA_ÇRÇS_YİÖ)	2,11	0,22	0,29

Tablodan görüldüğü gibi önerilen yöntem iki veri setinde de önceki çalışmalarla kıyaslanabilir bir performans göstermiştir. Çalışmada önerilen yöntem ile elde edilen sonuçların, renk dokusu konusunda YST alanında ilk kez yapılan ve HSV ve YC_bC_r renk uzaylarının birleşiminden üretilen YİÖ özniteliklerine dayanan çalışmalardan daha iyi başarım sergilediği görülmektedir. Bunun yanında, çoğu çalışmada veri setinde kaç adet çerçevenin kullanıldığı bilgisi verilmemiştir. Bu durum sonuçlar üzerinde adil bir karşılaştırma yapılmasını doğrudan etkilemektedir.

3.5. Derin Öğrenme Ağlarının Yüz Sahteciliği Tespiti Başarımları

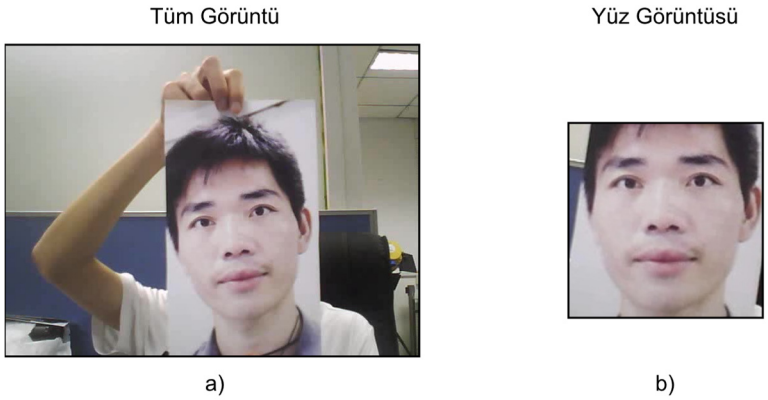
2014 yılında YST alanında ilk kez uygulanan derin öğrenme kavramı, 2018 yılına kadar yavaş bir ilerleyiş göstermiş, bu yıldan itibaren ise derin öğrenme uygulamalarının ve bu modelleri çalıştıracak donanımların yaygınlaşması sebebiyle hız kazanmıştır. Güçlü doku

tanımlama kabiliyetleri, bu uygulamaların, son yıllarda klasik öznitelik çıkarma algoritmalarından çok daha fazla tercih edilmesine neden olmuştur.

Önceki çalışmalarda giriş görüntüsünün çeşitli versiyonları yardımıyla YST problemi incelenmiş ve sonucunda, probleme en çok etki eden giriş formunun, geniş yüz bölgesi olduğu görülmüştür. Bu durum akla: "Giriş görüntüsünde yüz bölgesi dışında kalan alanlar yüz sahteciliği tespitinde etkili midir?" sorusunu getirmiştir.

Bir YST modeli için yüz görüntüsü anahtar veridir ancak bu modeller yalnızca yüzün belirli özelliklerini analiz etmekle kalmaz, aynı zamanda çevresel bilgilere ve bağlamsal ipuçlarına da odaklanabilir. Örneğin, sahte bir yüzün arka planı veya çevresel unsurları, arka plandaki nesnelerin konumu, ışıklandırma özellikleri, perspektif, gölgeler veya yüzey dokusu gibi detaylar, sahteciliğin tespitinde yardımcı olabilir. Geleneksel öznitelik çıkarma yöntemlerinde bellek tüketimi, giriş görüntülerinde yüz bölgelerine odaklanılmasına ve çevresel bilgilerden faydalanılamamasına neden olmaktadır. Ancak derin öğrenme modellerinde bu durum biraz daha farklıdır. Geleneksel modellere kıyasla derin öğrenme modelleri ekran kartı üzerinde çalıştığından, hesaplama ve bellek tüketimi için ekran kartını kullanırlar. Özellikle günümüzde yaygınlaşan yüksek bellek miktarına sahip ekran kartları için bu durumun üstesinden gelmek zor değildir ancak kullanılacak veri miktarı, ağın derinliği, filtre sayısı gibi parametreler, modelin boyunda ciddi rol oynar. Bu sebeple, kullanılacak modelin belirlenmesinde donanım kapasitesinin bilinmesi önemlidir.

Çalışmanın bu aşamasında, veri seti sağlayıcıları tarafından sunulan görüntüler (Şekil 39(a)) ve sadece yüz bölgelerini içeren görüntüler (Şekil 39(b)) kullanılarak, derin öğrenme modellerinin YST başarımları incelenmiştir.



Şekil 39. Derin öğrenme ağlarında kullanılan giriş görüntüleri

Derin öğrenme ağlarında başlangıç şartları önemli olduğundan, kullanılan tüm ağlar veri setlerinin tüm atak türlerini içeren saldırı senaryolarında (tümü) 5 kez eğitime tabi tutulmuş ve elde edilen sonuçlar, ortalama ve standart sapma değerleri ile tablolarda verilmiştir. Tablo 17, tüm ağlarda kullanılan parametrelere ait bilgileri içermektedir.

Tablo 17. Deneylerde kullanılan derin öğrenme parametreleri

Parametre	**Değer**
Giriş Görüntüsü Boyutu	Yüz Görüntüsü: 128 × 128 Tüm Görüntü: 224 × 224
Yığın Boyutu	32
Veri Artırma (Eğitim setinde)	Dikey ve Yatay çevirme: 0,2 Rastgele Parlaklık: 0,2 Rastgele Kontrast: 0,2 Rastgele Döndürme: 0,2
İyileştirici	Adam
Öğrenme Oranı	0,001
Devir Sayısı	100
Kayıp Fonksiyonu	İkili Çapraz-Entropi
Erken Durdurma	Sabır (Patience): 10 En iyi ağırlıkları geri yükle: True

Deneylerin karşılaştırılabilir olması açısından, tüm hesaplamalar 64G RAM, AMD Ryzen™ 5 5600 CPU @ 3,5GHZ işlemci ve NVIDIA GeForce RTX™ 3060 12G GDDR6 @ 1,78 GHz ekran kartına sahip standart bir masaüstü bilgisayarda, Windows 11 üzerinde çalışan Python 3,10 yazılım platformunda gerçekleştirilmiştir.

3.5.1. Xception Ağının Yüz Sahteciliği Tespiti Başarımı

Derin öğrenme ağlarında YST başarımlarının incelendiği bu bölümde, ilk olarak Xception ağı incelenmiştir. Bu aşamada kullanılan derin öğrenme ağının tüm katmanları, Tablo 17'de verilen parametreler kullanılarak eğitilmiştir. Sonuçların geçerliliği açısından bu işlem 5 kez tekrar edilmiştir. Tablo 18, elde edilen sınıflandırma sonuçlarını, ortalamalarını ve standart sapmalarını içermektedir.

Tablo 18. Xception ağının YST başarımı

Veri Seti	Giriş Görüntüsü	Ölçüt	Model Eğitim Aşaması					Ortalama	Standart Sapma
			1	2	3	4	5		
CASIA-FASD	Yüz görüntüsü	EER	2,32	2,45	2,68	2,66	1,93	2,41	0,31
	Tüm görüntü		1,77	2,09	1,55	4,11	2,25	2,35	1,02
REPLAY-ATTACK	Yüz görüntüsü	EER	15,39	3,27	7,85	3,38	8,73	7,72	4,96
		HTER	9,61	3,89	7,69	3,24	3,97	5,68	2,81
	Tüm görüntü	EER	0,00	0,00	0,21	0,00	0,00	0,04	0,09
		HTER	0,00	0,00	0,25	0,00	0,00	0,05	0,11

Tablo 18 incelendiğinde, giriş verisinin yüz bölgesinden kırpılmış görüntülerden oluşması yerine görüntünün tüm haliyle kullanılması, özellikle REPLAY-ATTACK veri setinin sınıflandırma sonuçlarında ciddi bir iyileştirmeye sebep olmuştur. Ancak bu senaryoda, eğitim aşamasında yüz görüntüleri kullanıldığında her bir devir ortalama 9 saniyede, tüm görüntüler kullanıldığında ise ortalama 26 saniyede tamamlanmıştır. Bu durum, eğitim sürelerinde ciddi bir artışa neden olmaktadır.

Xception ağı, oldukça güçlü bir tanımlama yeteneğine sahip olmasının yanında, eğitim için kullanılan yüksek parametre sayısı ve model ağırlık katsayılarının büyüklüğü sebebiyle hem hesaplama hem de depolama maliyetini beraberinde getirir. Bu tür durumlarda, ağın eğiteceği parametre sayısının azaltılması amacıyla öğrenme aktarımı uygulanır. Tablo 19, Xception ağına ImageNet katsayıları yardımıyla öğrenme aktarımı uygulandıktan sonra sadece çıkış katmanının eğitilmesi ile elde edilen sınıflandırma sonuçlarını içermektedir.

Tablo 19. Xception ağında öğrenme aktarımı ile YST başarımı

Veri Seti	Giriş Görüntüsü	Ölçüt	Model Eğitim Aşaması					Ortalama	Standart Sapma
			1	2	3	4	5		
CASIA -FASD	Yüz görüntüsü	EER	20,13	20,91	20,89	20,59	21,59	20,82	0,53
	Tüm görüntü		21,19	20,76	20,75	20,96	20,50	20,83	0,26
REPLAY-ATTACK	Yüz görüntüsü	EER	25,47	21,57	16,74	16,68	17,00	19,49	3,93
		HTER	22,28	17,89	19,47	19,48	19,76	19,78	1,58
	Tüm görüntü	EER	14,76	14,61	9,88	9,62	4,31	10,63	4,31
		HTER	8,95	9,06	8,28	7,55	6,44	8,06	1,09

Tablodan görüleceği üzere, öğrenme aktarımı kullanımı sonucunda YST başarımları ciddi oranda düşmektedir. Bu durumun birden çok sebebi olabilir ancak en büyük sebep, eğitime katılan parametre sayısındaki düşüştür. Xception ağı tümüyle eğitildiğinde, eğitime katılan parametre sayısı 20,809,001 iken, öğrenme aktarımı sonrasında bu sayı 2,049 olarak belirlenmiştir. Bu durumda model yalnızca çıkış katmanında bir güncelleme yapmakta ve bu da sınıflandırma için yeterli olamamaktadır. Böyle durumlarda ince ayar yardımıyla, ağın belirli bir katmandan itibaren eğitilmesi faydalı olabilmektedir ancak bu durum çalışma kapsamı dışında olduğundan gerçekleştirilmemiştir.

3.5.2. İndirgenmiş Xception Ağının Yüz Sahteciliği Tespiti Başarımı

İndirgenmiş Xception ağı, orijinal Xception ağının orta akışının ve çıkış akışındaki bir katmanın çıkarılmasıyla üretilmiş daha hafif bir versiyonudur. Bu ağda, orta akıştan çıkarılan yüksek seviyeli öznitelikler yerine, giriş akışında filtre sayısı arttırılarak, düşük seviyeli özniteliklerin etkisinin arttırılması hedeflenmiştir. Çalışmanın bu aşamasında, Xception ağının indirgenmiş versiyonu üzerinde YST başarımı incelenmiştir. Tablo 20, elde edilen sınıflandırma sonuçlarını, ortalamalarını ve standart sapmalarını içermektedir.

Tablo 20. İndirgenmiş Xception ağının YST başarımı

Veri Seti	Giriş Görüntüsü	Ölçüt	Model Eğitim Aşaması					Ortalama	Standart Sapma
			1	2	3	4	5		
CASIA -FASD	Yüz görüntüsü	EER	6,04	4,82	4,74	4,99	4,99	5,12	0,53
	Tüm görüntü		2,81	3,89	3,31	4,36	2,50	3,37	0,76
REPLAY-ATTACK	Yüz görüntüsü	EER	20,22	4,99	4,52	5,56	4,31	7,92	6,89
		HTER	8,21	4,77	5,46	6,72	5,25	6,08	1,39
	Tüm görüntü	EER	0,00	0,00	0,00	0,00	0,00	0,00	0,00
		HTER	2,34	4,90	2,20	0,00	0,00	1,89	2,03

Tablo 20 incelendiğinde, bir önceki sonuçlara benzer şekilde, yüz görüntüleri yerine görüntünün tamamının giriş olarak kullanıldığı senaryolarda YST başarımı daha yüksektir. Bu modelde elde edilen sonuçlar, öğrenme aktarımı kullanılan Xception ağından daha başarılıdır ancak Xception ağının tümüyle eğitiminden elde edilen sonuçlardan yalnızca REPLAY-ATTACK veri setinde tüm görüntülerin kullanıldığı senaryoda üretilen EER sonuçlarında başarılı olmuştur.

İndirgenmiş Xception ağı, elde ettiği sınıflandırma sonuçları açısından esinlenildiği ağ kadar başarılı gözükmemektedir ancak toplam parametre sayısı, orijinal ağdan yaklaşık %87 daha azdır. Bu sebeple, düşük hesaplama kabiliyetine ve depolama alanına sahip cihazlarda kullanılabilirliği çok daha yüksektir.

3.5.3. SE Bloklu İndirgenmiş Xception Ağının Yüz Sahteciliği Tespiti Başarımı

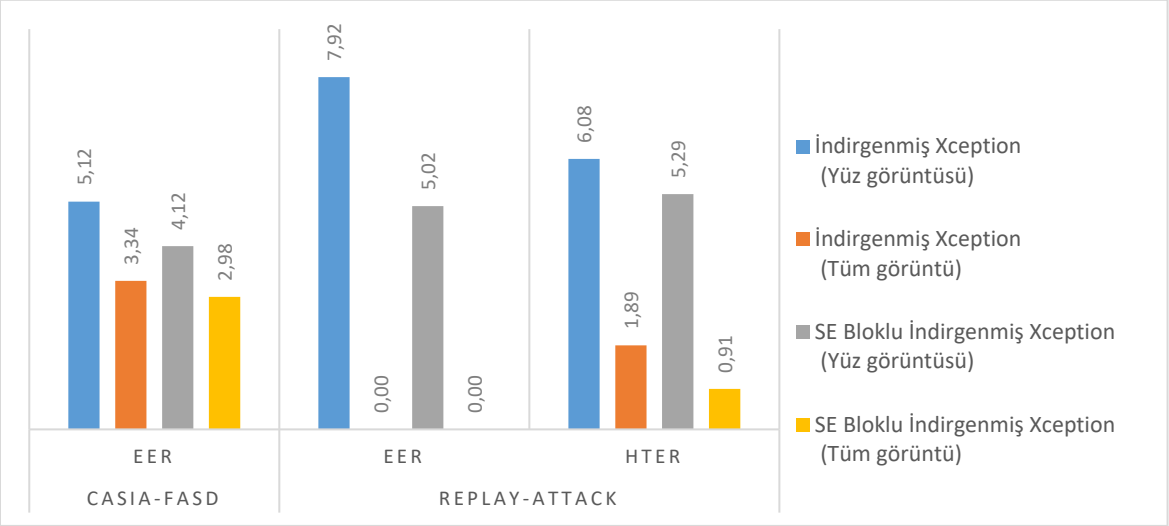
SE Bloklu İndirgenmiş Xception ağı, indirgenmiş Xception ağının artık katmanlarına sıkıştırma ve uyarma bloklarının eklenmesiyle, kanallarda bulunan özniteliklerden, problemin çözümünde etkisi düşük olanları bastırarak, YST başarımını arttırmayı hedeflemektedir. Çalışmanın bu aşamasında, dikkat mekanizmasına sahip indirgenmiş Xception ağı üzerinde YST başarımı incelenmiştir. Tablo 21, elde edilen sınıflandırma sonuçlarını, ortalamalarını ve standart sapmalarını içermektedir.

Tablo 21. SE bloklu indirgenmiş Xception ağının YST başarımı

Veri Seti	Giriş Görüntüsü	Ölçüt	Model Eğitim Aşaması					Ortalama	Standart Sapma
			1	2	3	4	5		
CASIA-FASD	Yüz görüntüsü	EER	4,40	4,59	3,84	3,65	4,12	4,12	0,39
	Tüm görüntü		2,96	2,93	2,55	3,52	2,95	2,98	0,35
REPLAY-ATTACK	Yüz görüntüsü	EER	2,96	4,26	4,00	5,82	8,06	5,02	1,98
		HTER	4,87	5,37	5,94	6,81	3,45	5,29	1,25
	Tüm görüntü	EER	0,00	0,00	0,00	0,00	0,00	0,00	0,00
		HTER	2,27	0,00	2,27	0,00	0,00	0,91	1,24

Tablo 21 incelendiğinde, önceki sonuçlarda olduğu gibi tüm görüntülerin kullanıldığı senaryolarda YST başarımı daha yüksektir. Bu modelde elde edilen sonuçlar, indirgenmiş Xception daha başarılıdır ancak Xception ağı ile kıyaslandığında, bir önceki durumda olduğu gibi yalnızca REPLAY-ATTACK veri setinde tüm görüntülerden üretilen EER sonuçlarında daha başarılıdır.

Önerilen bu model, Xception ağının toplam parametre sayısından yaklaşık %83 daha az, indirgenmiş Xception ağının toplam parametre sayısından ise yaklaşık %31 daha fazladır. Şekil 40, bu iki ağın tüm giriş görüntüsü türlerinde YST başarımlarının karşılaştırmasını içermektedir.



Şekil 40. İndirgenmiş Xception ağı ile SE bloklu indirgenmiş Xception ağının YST başarım karşılaştırması

Şekil 40 incelendiğinde, dikkat mekanizması içeren SE Bloklu İndirgenmiş Xception derin öğrenme modelinin, tüm veri setlerinde ve performans ölçütlerinde, indirgenmiş Xception Ağı modelinden daha başarılı olduğu görülmektedir. Bu durum, dikkat bloklarının YST problemindeki etkisini ortaya koymaktadır.

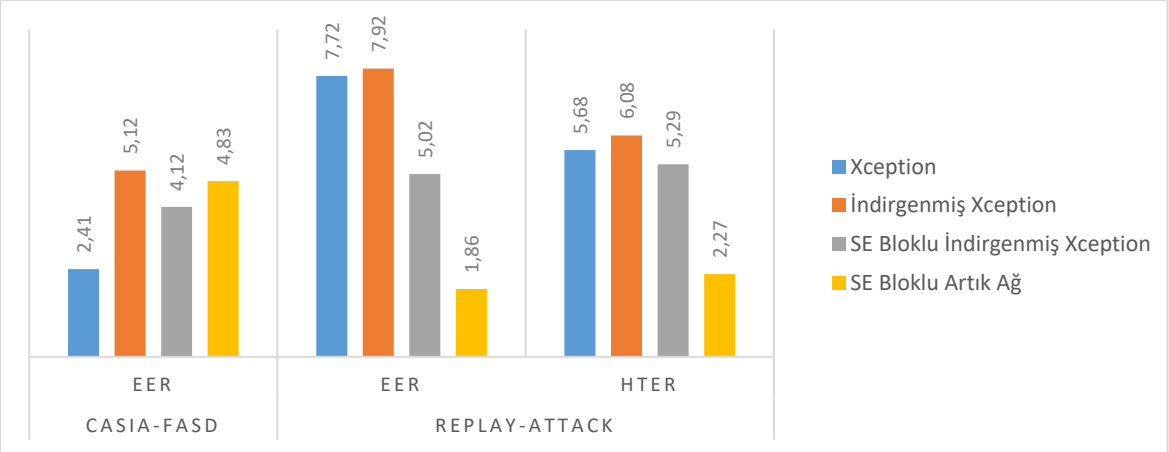
3.5.4. SE Bloklu Artık Ağın Yüz Sahteciliği Tespiti Başarımı

SE Bloklu Artık Ağ, sıkıştırma ve uyarma blokları içeren basit bir artık ağ tasarımına sahiptir. Bu ağın kullanımındaki amaç, veri setlerinin oldukça düşük parametrelere sahip ağlarla eğitilmesi durumunda, önerilen modelin YST başarımının incelenmesidir. Çalışmanın bu aşamasında, dikkat mekanizmasına sahip basit bir artık ağın YST başarımı incelenmiştir. Tablo 22, elde edilen sınıflandırma sonuçlarını, ortalamalarını ve standart sapmalarını içermektedir.

Tablo 22. SE bloklu artık ağın YST başarımı

Veri Seti	Giriş Görüntüsü	Ölçüt	Model Eğitim Aşaması					Ortalama	Standart Sapma
			1	2	3	4	5		
CASIA -FASD	Yüz görüntüsü	EER	4,22	5,12	5,31	4,74	4,77	4,83	0,42
	Tüm görüntü		4,01	3,15	2,38	4,11	3,15	3,36	0,72
REPLAY-ATTACK	Yüz görüntüsü	EER	0,88	0,42	2,60	2,70	2,70	1,86	1,12
		HTER	1,91	2,34	1,29	2,82	2,99	2,27	0,69
	Tüm görüntü	EER	0,00	0,00	0,16	0,00	0,00	0,03	0,07
		HTER	0,00	2,27	0,02	0,71	0,50	0,70	0,93

Tablo 22 incelendiğinde, önceki sonuçlarda olduğu gibi tüm görüntülerin kullanıldığı senaryolarda YST başarımı daha yüksektir. Ancak bu modelde diğerlerinden farklı bir durum meydana gelmiştir. Özellikle düşük parametreye ve dikkat mekanizmasına sahip olmasının yanında, çok düşük filtre sayısı içermesi sebebiyle, diğer senaryolardan farklı olarak, yalnızca yüz görüntülerinin kullanıldığı durumda oldukça ciddi bir iyileşme meydana gelmiştir. Şekil 41, önerilen tüm ağların yüz görüntüleri üzerindeki YST başarımlarının karşılaştırmasını içermektedir.



Şekil 41. Kullanılan derin ağların yüz görüntüleri üzerinde YST performansları

Şekil 41 incelendiğinde, yalnızca yüz görüntülerinin kullanıldığı senaryoda, özellikle REPLAY-ATTACK veri setinde SE bloklu artık ağın ciddi bir performans artışı sağladığı görülmektedir. Üstelik bu model yalnızca 64.161 parametre içermektedir. İçerdiği katman ve filtre sayısı sebebiyle bu kadar düşük parametreye sahip bu model için REPLAY-ATTACK veri seti üzerinde güçlü bir tanımlama yeteneğine sahip olduğu söylenebilir ancak aynı durum CASIA-FASD veri seti için söylenememektedir. Çalışmanın bir sonraki

aşamasında önerilen karma girişli derin öğrenme modelinde bu ağın hafifliğinden ve yüz bölgeleri üzerindeki güçlü tanımlama yeteneklerinden faydalanılmıştır.

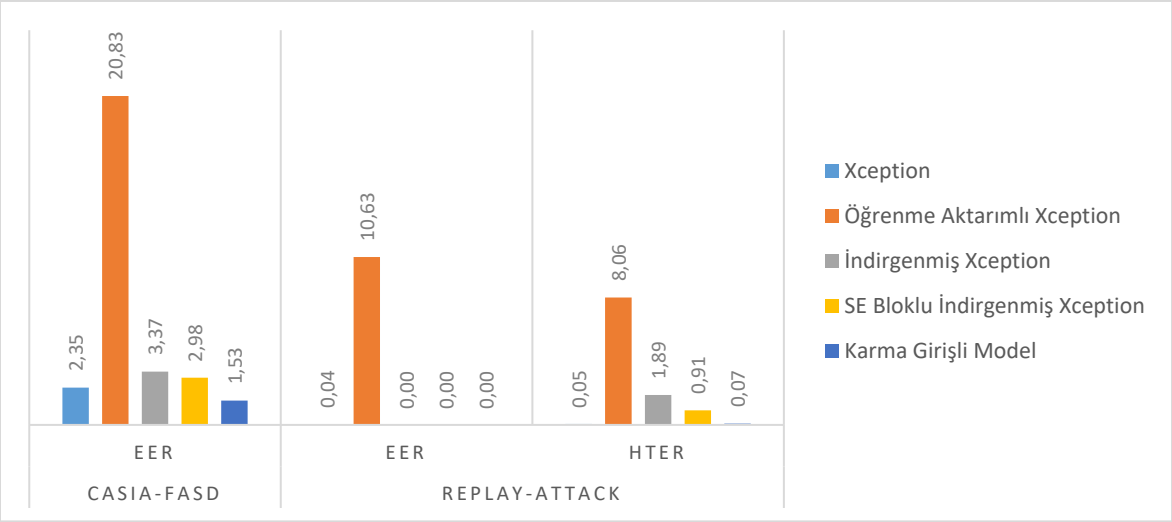
3.6. Karma Girişli Derin Öğrenme Modelinin Yüz Sahteciliği Tespiti Başarımı

Karma girişli derin öğrenme modeli, SE bloklu indirgenmiş Xception ağı, SE bloklu artık ağ ve sayısal verileri işleyen birkaç katmandan oluşan üç girişli bir ağdır. Bu model sırasıyla, tüm görüntü, yüz görüntüsü ve yüz görüntülerinden üretilen YA_ÇRÇS_YİÖ özniteliklerini giriş olarak almaktadır. Çalışmanın bu aşamasında, önerilen karma girişli derin öğrenme modeli üzerinde YST başarımı incelenmiştir. Tablo 23, elde edilen sınıflandırma sonuçlarını, ortalamalarını ve standart sapmalarını içermektedir.

Tablo 23. Karma girişli derin öğrenme modelinin YST başarımı

Veri Seti	Ölçüt	Model Eğitim Aşaması					Ortalama	Standart Sapma
		1	2	3	4	5		
CASIA-FASD	EER	1,61	2,35	1,47	1,03	1,19	1,53	0,51
REPLAY-ATTACK	EER	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	HTER	0,00	0,00	0,36	0,00	0,00	0,07	0,16

Tablo 23 incelendiğinde, önceki sonuçlarda olduğu gibi yüz görüntüleri ve tüm görüntüler için ayrı sonuçlar sunulmamıştır. Bunun sebebi, önerilen modelin iki veri türünü de giriş olarak kullanmasıdır. Bu modelde elde edilen sonuçlar, orijinal Xception ağı ile REPLAY-ATTACK veri setinden üretilen HTER sonuçları haricinde tüm ağlarda ve ölçütlerde daha başarılı çıkmıştır. Şekil 42, çalışmada kullanılan tüm derin öğrenme tabanlı ağların CASIA-FASD ve REPLAY-ATTACK veri setlerinde tüm görüntüler üzerindeki YST başarımlarını ifade etmektedir.



Şekil 42. CASIA-FASD ve REPLAY-ATTACK veri setleri üzerinde önerilen modellerin karşılaştırması

Önerilen karma girişli model REPLAY-ATTACK veri setinde Xception ağından daha düşük HTER performansı göstermiş olmasına rağmen EER değeri bakımından daha iyi sonuç üretmiştir. Bunun yanında, tanımlanması çok daha zor olan CASIA-FASD veri setinde önerilen karma girişli model, en yakın takipçisi olan Xception ağından yaklaşık %35 daha başarılı olmuştur. Üstelik tüm bu işlemleri, eğitime katılan 3.809.777 parametre ile gerçekleştirmiştir. Bu değer, Xception ağındaki 20.809.001 parametrenin yaklaşık %18,31'i kadardır.

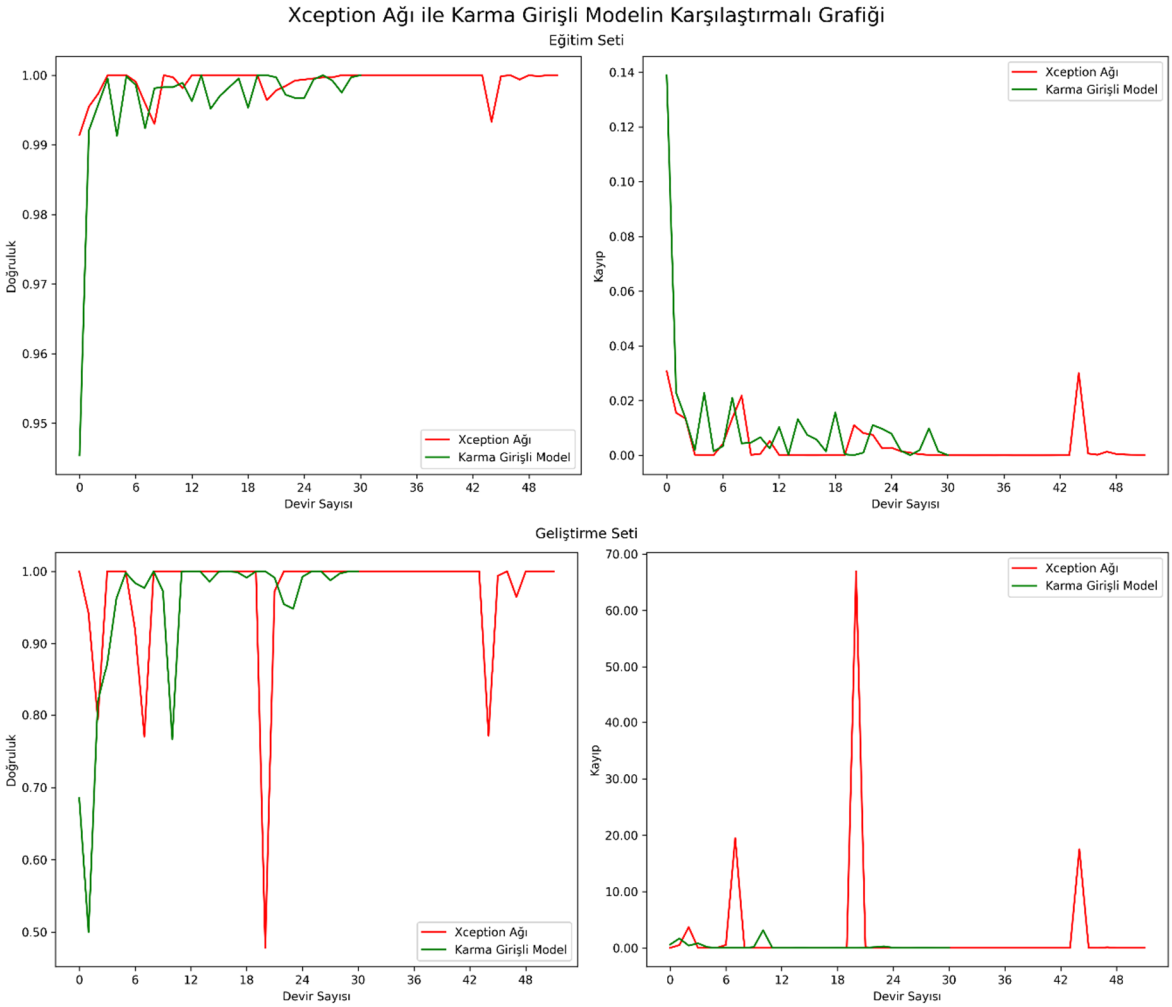
Düşük parametre sayısı ve hesaplama karmaşıklığının yanında kolay uygulanabilirliği ve anlaşılabilir olması sebebiyle, önerilen karma girişli derin öğrenme modelinin YST problemin çözümünde oldukça başarılı olduğu görülmektedir. Çalışmada kullanılan tüm ağların eğitime katılan ve toplam parametre sayıları Tablo 24'de verilmiştir.

Tablo 24. Kullanılan ağların parametre sayıları

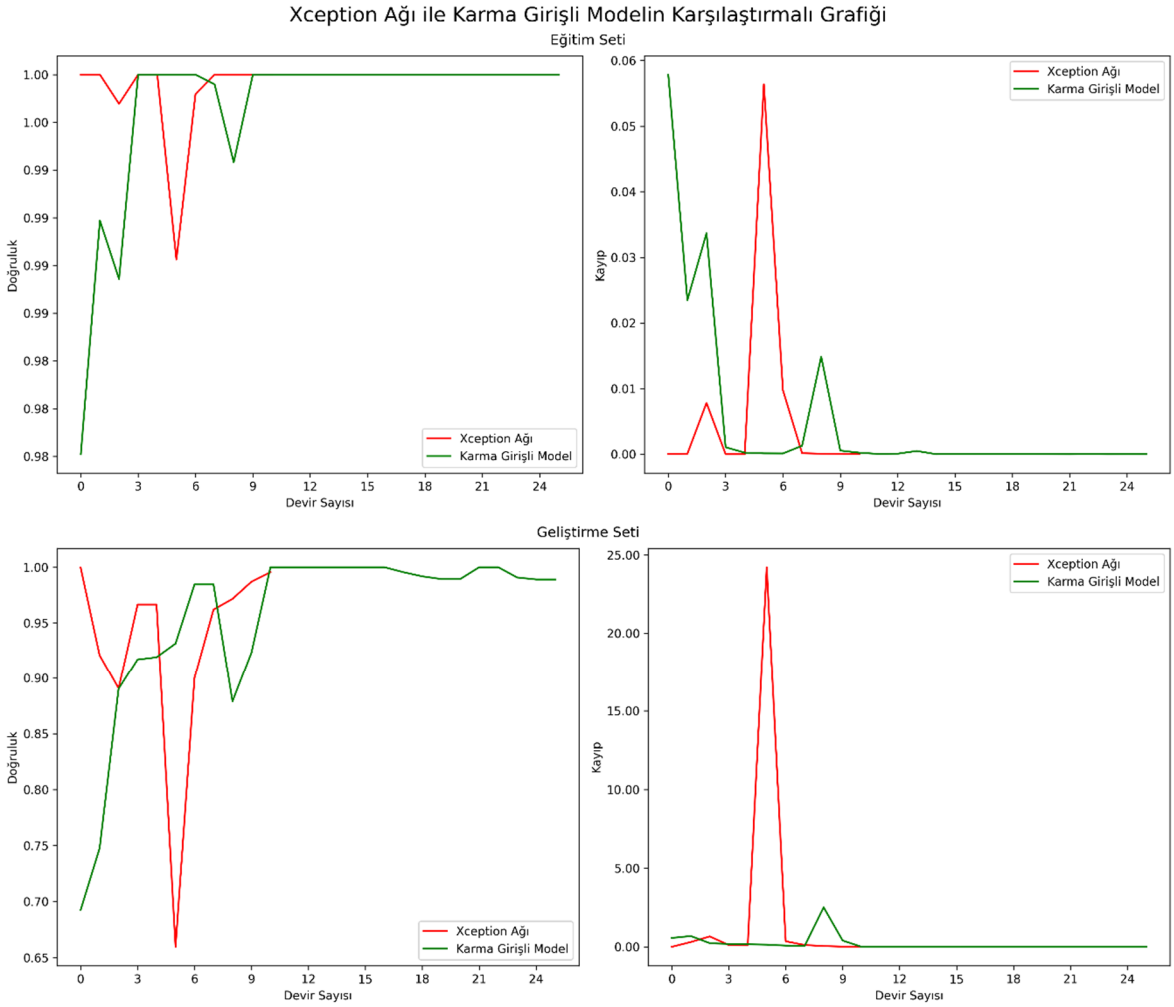
Parametre	**Eğitime Katılan Parametre Sayısı**	**Toplam Parametre Sayısı**
Xception	20.809.001	20.863.529
Xception (Öğrenme Aktarımlı)	2.049	20.863.529
İndirgenmiş Xception	2.722.777	2.731.065
SE Bloklu İndirgenmiş Xception	3.581.937	3.590.225
SE Bloklu Artık Ağ	63.713	64.161
Karma Girişli Model	3.809.041	3.817.777

Tablodan görüldüğü üzere, önerilen tüm ağlardaki parametre sayıları, orijinal Xception ağının toplam parametre sayısından %82 ile %87 arasında daha azdır. Şekil 43 ve

Şekil 44 sırasıyla, CASIA-FASD ve REPLAY-ATTACK veri setleri için önerilen karma girişli model ile Xception ağının YST başarımlarına ait eğitim seti doğruluğu, eğitim seti kaybı, geliştirme seti doğruluğu ve geliştirme seti kaybı grafiklerini içermektedir.



Şekil 43. Karma girişli model ile Xception ağının CASIA-FASD veri setinde YST performansları



Şekil 44. Karma girişli model ile Xception ağının REPLAY-ATTACK veri setinde YST performansları

Şekil 43 ve Şekil 44 incelendiğinde, Xception ağına oranla çok daha az parametreye sahip olan karma girişli modelin daha fazla keskin geçişe sahip değerler ürettiği görülmektedir. Bu durum, önerilen modelin düşük seviyeli özniteliklerle öğrenme işlemini gerçekleştirmesinden kaynaklanmaktadır. Bununla birlikte, üretilen sonuçlarda Xception ağına kıyasla çok daha az dalgalanma oluşmuştur. Bu veriler, önerilen modelin tanımlama yeteneğinin güçlü olduğunu göstermektedir. Grafiklerden, kullanılan ağlarda eğitim sırasında eksik ya da aşırı öğrenme olmadığı çıkarımı yapılabilir.

Tablo 25, tez çalışmasında önerilen ağlar, Xception ağı ve YA_ÇRÇS_YİÖ kodlayıcısının literatürdeki derin öğrenme tabanlı yaklaşımlarla karşılaştırmasını içermektedir.

Tablo 25. Önerilen derin öğrenme tabanlı yöntemlerin literatür karşılaştırması

Yöntem	CASIA-FASD	REPLAY-ATTACK	
	EER	EER	HTER
(Atoum vd., 2017)	2,67	0,79	0,72
(Tu & Fang, 2017)	1,00	1,03	1,18
(Rehman vd., 2018)	4,59	-	5,74
(L. Li vd., 2018)	2,50	0,60	1,30
(H. Li, He, vd., 2018)	1,40	0,30	1,20
(F.-M. Chen vd., 2019)	4,40	2,30	2,60
(H. Chen vd., 2019)	2,36	0,06	0,18
(H. Chen vd., 2020)	3,14	0,21	0,39
(L. Li vd., 2020)	-	0,80	0,70
(Shi vd., 2021)	**0,37**	**0,00**	0,50
(G. Wang vd., 2021)	3,30	-	1,30
(Shu vd., 2023)	2,90	4,70	0,39
(Antil & Dhiman, 2023)	2,37	**0,00**	**0,00**
(Pei vd., 2023)	1,13	**0,00**	**0,00**
YA_ÇRÇS_YİÖ	2,11	0,22	0,29
Xception	2,35	0,04	0,05
İndirgenmiş Xception	3,37	**0,00**	1,89
SE Bloklu İndirgenmiş Xception	2,98	**0,00**	0,91
SE Bloklu Artık Ağ	3,36	0,03	0,70
Karma Girişli Model	1,53	**0,00**	0,07

Tablo 25, önerilen ağların literatürdeki diğer çalışmaların birçoğundan daha başarılı olduğunu ve bazı sonuçlar ile kıyaslanabilir olduğunu ortaya koymaktadır. Bu durum, düşük parametre sayısına sahip ağlar yardımıyla YST başarımını arttırma konusunda hedeflenen noktaya varıldığını göstermektedir. Bunun yanında, karma girişli modelin özellikle CASIA-FASD veri seti üzerinde elde ettiği başarım, literatürdeki birçok çalışmadan açık ara öndedir. Esinlenildiği Xception ağından daha iyi sonuçlar üretmesi, YA_ÇRÇS_YİÖ kodlayıcısının güçlü tanımlama yeteneğini ve dikkat mekanizmasının etkisini ortaya koymaktadır.

4. SONUÇLAR

Bu tez çalışmasında, standart görüntüleme aygıtları yardımıyla yüz sahteciliği tespiti gerçekleştirebilen bir dizi model önerilmiştir.

Çalışmanın ilk aşamasında, giriş görüntüsünden çıkarılan çeşitli yüz bölgelerinin (geniş yüz, kırpılmış yüz, ağız, burun, göz) yüz sahteciliği problemine etkileri araştırılmıştır. Deneyler, 3 veri seti ve toplam 26 farklı test senaryosu için 5 farklı yüz bölgesinden üretilen $ÇB\_Y İÖ_{8,1}^{U2}$ histogramları ve bu histogramlara TBA uygulanarak elde edilen düşük boyutlu öznitelik kümesi üzerinde gerçekleştirilmiştir. Elde edilen sonuçlar, diğer bölgelere oranla çok daha fazla veriye sahip geniş yüz bölgesinin, 20 test senaryosunda en iyi sonuçları ürettiğini göstermiştir. Kalan 6 test senaryosunda ise kırpılmış yüz bölgesi başarılı olmuştur. Bu sonuçlar, YST probleminin çözümünde geniş yüz bölgesinin, kıyaslanan diğer bölgelere göre daha fazla bilgi içerdiğini göstermektedir. Öte yandan, giriş görüntülerinin içerdiği arka plan bilgisinin de performansı olumlu yönde etkilediği anlaşılmaktadır.

2020 yılının başlarında küresel çapta meydana gelen pandemi, cerrahi maske kullanımı zorunluluğunu beraberinde getirmiştir. Yüzün, burun ve ağız bölgelerini tümüyle kapatan bu maskeler, yüz sahteciliği tespitinde standart görüntüleme aygıtlarını kullanan sistemler için büyük bir problem haline gelmiştir. Bu sebeple, giriş görüntüsüne uygulanacak olası bölgesel gizlemeler sonucu YST sistemlerinin başarımlarını ölçmek amacıyla; ağız, burun ve göz bölgeleri kendi aralarında değerlendirmeye tabi tutulmuştur. Bu değerlendirme neticesinde, veri setlerine ve test senaryolarına göre farklı bölgelerin daha iyi performans gösterdiği görülmektedir. REPLAY-ATTACK veri setindeki 18 test senaryosunun 10'unda göz bölgesi diğer bölgelere göre daha başarılı olmuştur. Bu durum, mevcut ve gelecekteki pandemi koşullarında maske kullanımına bağlı olarak sadece göz bölgesinden yapılacak saldırıların tespit performansı hakkında bir öngörü oluşturmaktadır.

Çalışmanın ikinci aşamasında, düzgün örüntülerin haricinde tüm örüntülerin YST performansı incelenmiştir. Bu alandaki deneyler, bir önceki çalışmada en iyi YST performansına sahip olduğu belirlenen geniş yüz bölgesi üzerinde gerçekleştirilmiştir. Bu bölgeden üretilen $ÇB\_Y İÖ_{8,1}$ histogramları üzerinde TBA kullanılarak boyutları indirgenmiş ve SVM ile sınıflandırma yapılmıştır. Elde edilen sonuçlar, $ÇB\_Y İÖ_{8,1}$ histogramlarının 19 test senaryosunda YST performansını artırdığını göstermiştir. Bu veriler, bir giriş

görüntüsünden üretilen tüm YİÖ histogramlarının, YST probleminde önemli bilgiler taşıdığını ortaya koymuştur.

Çalışmanın üçüncü aşamasında, renk uzaylarının YST problemine etkisi incelenmiştir. Literatürde sıklıkla kullanılan cihaza bağımlı renk uzaylarının yanında (HSV, YC_bC_r,), cihazdan bağımsız L*a*b* renk uzayı kullanılarak, tespit performansının arttırılması hedeflenmiştir. Kullanılan renk uzaylarından çıkarılan $ÇB\_Y\text{İÖ}_{8,1}^{U2}$ öznitelikleri üzerinde tüm kombinasyonel birleşimler uygulanmış ve deneyler, CASIA-FASD ve REPLAY-ATTACK veri setleri üzerinde 13 farklı senaryoda gerçekleştirilmiştir. Elde edilen sonuçlar, tüm renk kanallarının farklı problemler üzerinde farklı etkileri olduğunu, tek bir renk kanalının, problemin çözümü için sunulamayacağını ortaya koymuştur. Bunun yanında, renk kanallarını kullanmanın, gri renk uzayı ile gerçekleştirilen yüz sahteciliği tespitine oranla başarımı çok daha fazla arttırdığı görülmüştür.

Çalışmanın dördüncü aşamasında, birden çok renk uzayının kullanılması sebebiyle meydana gelen yüksek boyutlu veri kümelerinde hesaplama ve depolama maliyetlerini indirgemek amacıyla bir örüntü kodlayıcı önerilmiştir. Önerilen algoritma, cihaza bağımlı ve cihazdan bağımsız iki renk uzayını kullanmakta ve renk uzaylarının her kanalından üretilecek $ÇS\_Y\text{İÖ}_{8,1}$ histogramlarında giriş vektör boyutunu azaltmayı hedeflemektedir. Önerilen bu kodlayıcı, giriş görüntüsünün 1×1 ve 2×2 alt bölgelere ayrıştırılmasını ve bu bölgelerden $ÇB\_Y\text{İÖ}_{8,1}$ histogramlarının üretilmesini, sonrasında ise yüz görüntüsündeki temel ayrıştırıcı noktaları içeren (ağız, burun ve göz) bir alt bölgeden $ÇB\_Y\text{İÖ}_{8,1}$ histogramlarının üretilmesini ve son aşamada tüm bu histogramların birleştirilmesini esas almaktadır. Tüm deneyler CASIA-FASD ve REPLAY-ATTACK veri setleri üzerinde 25 farklı test senaryosunda gerçekleştirilmiştir. Elde edilen sonuçlar, önerilen yöntemin son teknoloji yöntemlere kıyasla YST probleminde başarılı olduğunu göstermiştir.

Çalışmanın beşinci aşamasında, Xception derin öğrenme ağının ve bu ağdan esinlenerek üretilen derin öğrenme modellerinin YST başarımına etkileri araştırılmıştır. Derin öğrenme modelleri yüksek hesaplama ve depolama maliyeleri gerektirdiğinden, üretilen yeni ağlarda parametre sayılarının oldukça düşük olmasına dikkat edilmiştir. Bunun yanında, kanal dikkat mekanizmaları yardımıyla katmanlar üzerinde ağırlıklandırmalar gerçekleştirilmiş ve bu durumun YST başarımını arttırdığı görülmüştür. Önerilen modeller Xception ağından daha başarılı olamamıştır ancak parametre sayısı ve çalışma hızları açısından değerlendirildiklerinde, özellikle daha hafif ağların tasarlanma gerekliliğinin oluştuğu günümüzde kullanım alanlarına sahip olabilecekleri düşünülmektedir.

Çalışmanın son aşamasında hem derin öznitelikleri hem de önceki çalışmada önerilen YA_ÇRÇS_YİÖ özniteliklerini kullanan, karma girişli derin öğrenme modeli önerilmiştir. Önerilen model, derin öğrenme ağlarının güçlü tanımlayıcı özelliklerinden ve klasik öznitelik çıkarma yöntemlerinin kolay uygulanabilirliğinden faydalanmaktadır. Elde edilen sonuçlar incelendiğinde, önerilen modelin CASIA-FASD veri setinde Xception ağından daha başarılı olduğu, REPLAY-ATTACK veri setinde ise çok yakın performanslar sergilediği görülmüştür. Üstelik tüm bunları, Xception ağına oranla yaklaşık %82 daha az parametre ile gerçekleştirmiştir. Bu durum, önerilen YA_ÇRÇS_YİÖ kodlayıcı ile üretilen özniteliklerin, farklı modellerle kullanıldığında tanımlama yeteneğini arttırdığını ortaya koymaktadır.

5. ÖNERİLER

- Farklı renk uzaylarının YST problemine etkisi araştırılabilir. Bir renk uzayının her problemde etkili olduğunu söylemek mümkün olmadığından, mümkün olan tüm uzaylarından üretilen doku özniteliklerinin YST etkilerinin araştırılması, sonraki çalışmalarda bu uzaylara odaklanılmasını sağlayabilir.
- Kullanılacak renk uzaylarının birleşimi ve bu birleşimin ağırlıklandırılması, sınıflandırma işleminin birden çok sınıflandırıcı ile gerçekleştirilmesi ve sonuçta oylama mekanizmasının kullanılması, YST problemini çözümlemede etkin rol oynayabilir.
- Farklı öznitelik çıkarma yöntemleri kullanılarak üretilen öznitelikler, derin özniteliklerle birleştirilerek, güçlü tanımlama yeteneğine sahip ağlar oluşturulabilir.
- Daha az parametrelere sahip derin ağlar tasarlanabilir. Özellikle taşınabilir cihazlar seviyesine indirgenen bu sistemlerde hesaplama karmaşıklığının azaltılması oldukça önemlidir.
- Kullanılacak derin öğrenme tabanlı modellerde öğrenme aktarımı ince ayar yardımıyla uygulanabilir. Bu durum, eğitim sürelerinde kısalmaya ve doğrudan öğrenme aktarımı kullanımına kıyasla doğruluk değerlerinde artışa neden olabilir.
- Derin öğrenme ağlarında başlangıç koşullarının belirlenmesi ve hiper parametrelerin incelenmesi adına optimizasyon çalışmaları gerçekleştirilebilir. Bu alanda gerçekleştirilecek tüm çalışmalar, daha sonra yapılacak araştırmalar için bir kılavuz niteliğinde olacaktır.
- Çapraz test senaryoları yardımıyla önerilen modellerin gerçek hayattaki problemlere nasıl tepkiler verdiği ölçülebilir. Modelin istenilen performansı sunmaması durumunda alan transferi yöntemleri ile sahip olunan veri setleri, hedeflenen alandaki görüntülere uyarlanabilir.

6. KAYNAKLAR

Abdullakutty, F., Johnston, P., & Elyan, E. (2022). Fusion Methods for Face Presentation Attack Detection. *Sensors*, *22*(14). https://doi.org/10.3390/s22145196

Agarwal, A., Singh, R., & Vatsa, M. (2016). Face anti-spoofing using Haralick features. *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems, BTAS 2016*. https://doi.org/10.1109/BTAS.2016.7791171

Agrawal, S., Sarkar, S., Alazab, M., Maddikunta, P. K. R., Gadekallu, T. R., & Pham, Q.-V. (2021). Genetic CFL: Hyperparameter Optimization in Clustered Federated Learning. *Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2021/7156420

Ahmed, L. A. (2020). Comparison of Artificial Neural Network and Box- Jenkins Models to Predict the Number of Patients With Hypertension in Kalar. *Ibn Al- Haitham Journal for Pure and Applied Science*. https://doi.org/10.30526/33.4.2516

Ali, A., Deravi, F., & Hoque, S. (2013). Directional sensitivity of gaze-collinearity features in liveness detection. *Proceedings - 2013 4th International Conference on Emerging Security Technologies, EST 2013*, 8–11. https://doi.org/10.1109/EST.2013.7

Alotaibi, A., & Mahmood, A. (2017). Deep face liveness detection based on nonlinear diffusion using convolution neural network. *Signal, Image and Video Processing*, *11*(4), 713–720. https://doi.org/10.1007/s11760-016-1014-2

Alpaydin, E. (2010). Introduction to Machine Learning Third Edition. *Introduction to Machine Learning*. https://doi.org/10.1007/978-1-62703-748-8_7

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data 2021 8:1*, *8*(1), 1–74. https://doi.org/10.1186/S40537-021-00444-8

Anjos, A., Chakka, M. M., & Marcel, S. (2014). Motion-based counter-measures to photo attacks in face recognition. *IET Biometrics*, *3*(3), 147–158. https://doi.org/10.1049/iet-bmt.2012.0071

Anjos, A., & Marcel, S. (2011). Counter-measures to photo attacks in face recognition: A public database and a baseline. *2011 International Joint Conference on Biometrics, IJCB 2011*. https://doi.org/10.1109/IJCB.2011.6117503

Antil, A., & Dhiman, C. (2023). A two stream face anti-spoofing framework using multi-level deep features and ELBP features. *Multimedia Systems*, *29*(3), 1361–1376. https://doi.org/10.1007/s00530-023-01060-7

Arashloo, S. R., & Kittler, J. (2018). An anomaly detection approach to face spoofing detection: A new formulation and evaluation protocol. *IEEE International Joint Conference on Biometrics, IJCB 2017*, *2018-Janua*, 80–89.

https://doi.org/10.1109/BTAS.2017.8272685

Arashloo, S. R., Kittler, J., & Christmas, W. (2015). Face Spoofing Detection Based on Multiple Descriptor Fusion Using Multiscale Dynamic Binarized Statistical Image Features. *IEEE Transactions on Information Forensics and Security*, *10*(11), 2396–2407. https://doi.org/10.1109/TIFS.2015.2458700

Arora, S., Bhatia, M. P. S., & Mittal, V. (2022). A robust framework for spoofing detection in faces using deep learning. *Visual Computer*, *38*(7), 2461–2472. https://doi.org/10.1007/s00371-021-02123-4

Atoum, Y., Liu, Y., Jourabloo, A., & Liu, X. (2017). Face anti-spoofing using patch and depth-based CNNs. *2017 IEEE International Joint Conference on Biometrics (IJCB)*, *2018-Janua*, 319–328. https://doi.org/10.1109/BTAS.2017.8272713

Bagga, M., & Singh, B. (2016). Spoofing detection in face recognition: A review. *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, 2037–2042.

Baldevbhai, P. J., & Anand, R. S. (2012). Color Image Segmentation for Medical Images using L*a*b* Color Space. *IOSR Journal of Electronics and Communication Engineering*, *1*(2), 24–45. https://doi.org/10.9790/2834-0122445

Bengio, Y. (2009). Learning Deep Architectures for AI. *Foundations and Trends® in Machine Learning*. https://doi.org/10.1561/2200000006

Bharadwaj, S., Dhamecha, T. I., Vatsa, M., & Singh, R. (2013). Computationally efficient face spoofing detection with motion magnification. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 105–110. https://doi.org/10.1109/CVPRW.2013.23

Bhaskar, T. N., Keat, F. T., Ranganath, S., & Venkatesh, Y. V. (2003). Blink detection and eye tracking for eye localization. *TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region*, *2*, 821–824.

Bhattacharjee, S., & Marcel, S. (2017). What You Can't See Can Help You - Extended-Range Imaging for 3D-Mask Presentation Attack Detection. *Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft fur Informatik (GI)*. https://doi.org/10.23919/BIOSIG.2017.8053524

Bhattacharjee, S., Mohammadi, A., & Marcel, S. (2018). Spoofing deep face recognition with custom silicone masks. *Tech. Rep.*

Bhogal, A. P. S., Sollinger, D., Trung, P., & Uhl, A. (2017). Non-reference image quality assessment for biometric presentation attack detection. *Proceedings - 2017 5th International Workshop on Biometrics and Forensics, IWBF 2017*. https://doi.org/10.1109/IWBF.2017.7935080

Bora, D. J., Kumar Gupta, A., & Khan, F. A. (2015). Comparing the Performance of L*A*B* and HSV Color Spaces with Respect to Color Image Segmentation. Içinde *International Journal of Emerging Technology and Advanced Engineering* (C. 5, Sayı 2).

www.ijetae.com

Boubaker, H., Canarella, G., Gupta, R., & Miller, S. M. (2022). A Hybrid ARFIMA Wavelet Artificial Neural Network Model for DJIA Index Forecasting. *Computational Economics*. https://doi.org/10.1007/s10614-022-10320-z

Boulkenafet, Z., Komulainen, J., & Hadid, A. (2015). Face anti-spoofing based on color texture analysis. *2015 IEEE International Conference on Image Processing (ICIP)*, 2636–2640. https://doi.org/10.1109/ICIP.2015.7351280

Boulkenafet, Z., Komulainen, J., & Hadid, A. (2016a). Face Anti-Spoofing using Speeded-Up Robust Features and Fisher Vector Encoding. *IEEE Signal Processing Letters*, *24*(2), 1–1. https://doi.org/10.1109/LSP.2016.2630740

Boulkenafet, Z., Komulainen, J., & Hadid, A. (2016b). Face Spoofing Detection Using Colour Texture Analysis. *IEEE Transactions on Information Forensics and Security*, *11*(8), 1818–1830. https://doi.org/10.1109/TIFS.2016.2555286

Boulkenafet, Z., Komulainen, J., & Hadid, A. (2018). On the generalization of color texture-based face anti-spoofing. *Image and Vision Computing*, *77*, 1–9. https://doi.org/10.1016/j.imavis.2018.04.007

Boulkenafet, Z., Komulainen, J., Li, L., Feng, X., & Hadid, A. (2017). OULU-NPU: A Mobile Face Presentation Attack Database with Real-World Variations. *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 612–618. https://doi.org/10.1109/FG.2017.77

Cai, L., Xiong, C., Huang, L., & Liu, C. (2015). A novel face spoofing detection method based on gaze estimation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9005*, 547–561. https://doi.org/10.1007/978-3-319-16811-1_36/COVER

Chakka, M. M., Anjos, A., Marcel, S., Tronci, R., Muntoni, D., Fadda, G., Pili, M., Sirena, N., Murgia, G., Ristori, M., Roli, F., Yan, J., Yi, D., Lei, Z., Zhang, Z., Li, S. Z., Schwartz, W. R., Rocha, A., Pedrini, H., … Pietikäinen, M. (2011). Competition on counter measures to 2-D facial spoofing attacks. *2011 International Joint Conference on Biometrics, IJCB 2011*. https://doi.org/10.1109/IJCB.2011.6117509

Chang, H. H., & Yeh, C. H. (2022). Face anti-spoofing detection based on multi-scale image quality assessment. *Image and Vision Computing*, *121*. https://doi.org/10.1016/j.imavis.2022.104428

Chen, F.-M., Wen, C., Xie, K., Wen, F.-Q., Sheng, G.-Q., & Tang, X.-G. (2019). Face liveness detection: Fusing colour texture feature and deep feature. *IET Biometrics*, *8*(6), 369–377. https://doi.org/10.1049/iet-bmt.2018.5235

Chen, H., Chen, Y., Tian, X., & Jiang, R. (2019). A Cascade Face Spoofing Detector Based on Face Anti-Spoofing R-CNN and Improved Retinex LBP. *IEEE Access*, *7*, 170116–170133. https://doi.org/10.1109/ACCESS.2019.2955383

Chen, H., Hu, G., Lei, Z., Chen, Y., Robertson, N. M., & Li, S. Z. (2020). Attention-Based

Two-Stream Convolutional Networks for Face Spoofing Detection. *IEEE Transactions on Information Forensics and Security*, *15*, 578–593. https://doi.org/10.1109/TIFS.2019.2922241

Chingovska, I., Anjos, A., & Marcel, S. (2012). On the effectiveness of local binary patterns in face anti-spoofing. *2012 BIOSIG - Proceedings of the International Conference of Biometrics Special Interest Group (BIOSIG)*, 1–7.

Croux, C., Filzmoser, P., & Fritz, H. (2013). Robust sparse principal component analysis. *Technometrics*. https://doi.org/10.1080/00401706.2012.727746

Csurka, G. (2017). Domain adaptation for visual applications: A comprehensive survey. *Domain Adaptation for Visual Applications: A Comprehensive Survey*.

Curtis, S. (2013). *iPhone 5s fingerprint sensor "hacked" within days of launch*. The Telegraph. http://www.telegraph.co.uk/technology/apple/iphone/10327635/iPhone-5s-fingerprint-sensor-hacked-within-days-of-launch.html

De Freitas Pereira, T., Anjos, A., De Martino, J. M., & Marcel, S. (2013). LBP-TOP based countermeasure against face spoofing attacks. *Computer Vision - ACCV 2012 Workshops*, *7728 LNCS*(PART 1), 121–132. https://doi.org/10.1007/978-3-642-37410-4_11

De Freitas Pereira, T., Komulainen, J., Anjos, A., De Martino, J. M., Hadid, A., Pietikäinen, M., & Marcel, S. (2014). Face liveness detection using dynamic texture. Içinde *EURASIP Journal on Image and Video Processing* (Sayı 2). http://jivp.eurasipjournals.com/content/2014/1/2

De Luis-García, R., Alberola-López, C., Aghzout, O., & Ruiz-Alzola, J. (2003). Biometric identification systems. *Signal Processing*, *83*(12), 2539–2557. https://doi.org/10.1016/j.sigpro.2003.08.001

Deb, D., & Jain, A. K. (2021). Look Locally Infer Globally: A Generalizable Face Anti-Spoofing Approach. *IEEE Transactions on Information Forensics and Security*, *16*, 1143–1157. https://doi.org/10.1109/TIFS.2020.3029879

Dong, W., Liu, J., Xie, Z., & Dong, L. (2019). *Adaptive Neural Network-Based Approximation to Accelerate Eulerian Fluid Simulation*. https://doi.org/10.1145/3295500.3356147

Dubey, R. K., Goh, J., & Thing, V. L. L. (2016). Fingerprint Liveness Detection From Single Image Using Low-Level Features and Shape Analysis. *IEEE Transactions on Information Forensics and Security*, *11*(7), 1461–1475. https://doi.org/10.1109/TIFS.2016.2535899

El-Din, Y. S., Moustafa, M. N., & Mahdi, H. (2021). Adversarial unsupervised domain adaptation guided with deep clustering for face presentation attack detection. *Adversarial Unsupervised Domain Adaptation Guided with Deep Clustering for Face Presentation Attack Detection*.

Erdogmus, N., & Marcel, S. (2014). Spoofing face recognition with 3D masks. *IEEE*

Transactions on Information Forensics and Security, *9*(7), 1084–1097. https://doi.org/10.1109/TIFS.2014.2322255

Fadli, V. F., & Herlistiono, I. O. (2020). Steel Surface Defect Detection Using Deep Learning. *International Journal of Innovative Science and Research Technology*. https://doi.org/10.38124/ijisrt20jul240

Galbally, J., Marcel, S., & Fierrez, J. (2014a). Biometric antispoofing methods: A survey in face recognition. *IEEE Access*, *2*, 1530–1552. https://doi.org/10.1109/ACCESS.2014.2381273

Galbally, J., Marcel, S., & Fierrez, J. (2014b). Image quality assessment for fake biometric detection: Application to Iris, fingerprint, and face recognition. *IEEE Transactions on Image Processing*, *23*(2), 710–724. https://doi.org/10.1109/TIP.2013.2292332

Galbally, J., & Satta, R. (2016). Three-dimensional and two-and-a-half-dimensional face recognition spoofing using three-dimensional printed models. *IET Biometrics*, *5*(2), 83–91. https://doi.org/10.1049/iet-bmt.2014.0075

Ganin, Y., & Lempitsky, V. (2015). Unsupervised domain adaptation by backpropagation. *32nd International Conference on Machine Learning, ICML 2015*, *2*, 1180–1189. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84969802531&partnerID=40&md5=6e7bda57e1a6acedc8878dba299b2eb2

George, A., & Marcel, S. (2019). Deep Pixel-wise Binary Supervision for Face Presentation Attack Detection. *2019 International Conference on Biometrics, ICB 2019*. https://doi.org/10.1109/ICB45273.2019.8987370

George, A., Mostaani, Z., Geissenbuhler, D., Nikisins, O., Anjos, A., & Marcel, S. (2019). Biometric Face Presentation Attack Detection With Multi-Channel Convolutional Neural Network. *IEEE Transactions on Information Forensics and Security*, *15*(1), 42–55. https://doi.org/10.1109/TIFS.2019.2916652

Grover, K., & Mehra, R. (2019). Face Spoofing Detection using Enhanced Local Binary Pattern. *Int. J. Eng. Adv. Technol.*, *9*(2).

Günay, A., & Nabiyev, V. (2017). Yüz Bölgelerinin Yaş Tahmini Başarımlarının Yaş Gruplarına Göre Değerlendirilmesi. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, *9*(2), 1–10. https://dergipark.org.tr/tr/pub/tbbmd/issue/30339/267320

Hasan, M. R., Hasan Mahmud, S. M., & Li, X. Y. (2019). Face Anti-Spoofing Using Texture-Based Techniques and Filtering Methods. *Journal of Physics: Conference Series*, *1229*(1). https://doi.org/10.1088/1742-6596/1229/1/012044

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-December*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-Excitation Networks. *Proceedings of the*

IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 7132–7141. https://doi.org/10.1109/CVPR.2018.00745

Huang, L., Xu, J., Sun, J., & Yang, Y. (2017). An improved residual LSTM architecture for acoustic modeling. *2nd International Conference on Computer and Communication Systems (ICCCS)*, 101–105.

Huang, Z. K., & Liu, D. H. (2007). Segmentation of color image using EM algorithm in HSV color space. *Proceedings of the 2007 International Conference on Information Acquisition, ICIA*, 316–319. https://doi.org/10.1109/ICIA.2007.4295749

Imaoka, H., Hashimoto, H., Takahashi, K., Ebihara, A. F., Liu, J., Hayasaka, A., Morishita, Y., & Sakurai, K. (2021). The future of biometrics technology: From face recognition to related applications. *APSIPA Transactions on Signal and Information Processing*. https://doi.org/10.1017/ATSIP.2021.8

Jeon, W.-S., & Rhee, S.-Y. (2021). A Restoration Method of Single Image Super Resolution Using Improved Residual Learning With Squeeze and Excitation Blocks. *International Journal of Fuzzy Logic and Intelligent Systems*. https://doi.org/10.5391/ijfis.2021.21.3.222

Jourabloo, A., Liu, Y., & Liu, X. (2018). Face de-spoofing: Anti-spoofing via noise modeling. *ECCV*, 290–306.

Kannala, J., & Rahtu, E. (2012). BSIF: Binarized statistical image features. *Proceedings - International Conference on Pattern Recognition*, 1363–1366. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84874563913&partnerID=40&md5=12e4641bd4d9e597136beb41bbe1b274

Kavzoğlu, T., & Çölkesen, İ. (2010). Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi. *Harita Dergisi*, *144*(7), 73–82.

Khurshid, A., Tamayo, S. C., Fernandes, E., Gadelha, M. R., & Teofilo, M. (2019). A robust and real-time face anti-spoofing method based on texture feature analysis. *International Conference on Human-Computer Interaction*, 484–496.

Kim, I., Ahn, J., & Kim, D. (2017). Face spoofing detection with highlight removal effect and distortions. *2016 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016 - Conference Proceedings*, 4299–4304. https://doi.org/10.1109/SMC.2016.7844907

Kim, S., Ban, Y., & Lee, S. (2015). Face Liveness Detection Using Defocus. *Sensors 2015, Vol. 15, Pages 1537-1563*, *15*(1), 1537–1563. https://doi.org/10.3390/S150101537

Kim, S., Yu, S., Kim, K., Ban, Y., & Lee, S. (2013). Face Liveness Detection using Variable Focusing. *IEEE*, 1–6.

Kim, Y., Yoo, J.-H., & Choi, K. (2011). A motion and similarity-based fake detection method for biometric face recognition systems. *IEEE Transactions on Consumer Electronics*, *57*(2), 756–762. https://doi.org/10.1109/TCE.2011.5955219

King, D. E. (2009). Dlib-ml: A machine learning toolkit. *The Journal of Machine Learning Research*, *10*, 1755–1758.

Kollreider, K., Fronthaler, H., & Bigun, J. (2008). Verifying liveness by multiple experts in face biometrics. *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 1–6.

Komulainen, J., Hadid, A., & Pietikäinen, M. (2013). Face Spoofing Detection Using Dynamic Texture. Içinde J.-I. Park & J. Kim (Ed.), *Computer Vision - ACCV 2012 Workshops* (ss. 146–157). Springer Berlin Heidelberg.

Komulainen, J., Hadid, A., Pietikainen, M., Anjos, A., & Marcel, S. (2013). Complementary countermeasures for detecting scenic face spoofing attacks. *Proceedings - 2013 International Conference on Biometrics, ICB 2013*. https://doi.org/10.1109/ICB.2013.6612968

Kose, N., & Dugelay, J.-L. (2013). Reflectance analysis based countermeasure technique to detect face mask attacks. *2013 18th International Conference on Digital Signal Processing, DSP 2013*. https://doi.org/10.1109/ICDSP.2013.6622704

Kotwal, K., Bhattacharjee, S., Abbet, P., Mostaani, Z., Wei, H., Wenkang, X., Yaxi, Z., & Marcel, S. (2022). Domain-Specific Adaptation of CNN for Detecting Face Presentation Attacks in NIR. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, *4*(1), 135–147. https://doi.org/10.1109/TBIOM.2022.3143569

Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). Imagenet classification with deep convolutional neural networks. *NIPS*, *25*, 1106–1114.

Lagorio, A., Tistarelli, M., Cadoni, M., Fookes, C., & Sridharan, S. (2013). Liveness detection based on 3D face shape analysis. *2013 International Workshop on Biometrics and Forensics (IWBF)*, 1–4.

Lakmann, R. (1998). Barktex benchmark database of color textured images. *Koblenz-Landau University*.

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2323. https://doi.org/10.1109/5.726791

Li, H., He, P., Wang, S., Rocha, A., Jiang, X., & Kot, A. C. (2018). Learning Generalized Deep Feature Representation for Face Anti-Spoofing. *IEEE Transactions on Information Forensics and Security*, *13*(10), 2639–2652. https://doi.org/10.1109/TIFS.2018.2825949

Li, H., Li, W., Cao, H., Wang, S., Huang, F., & Kot, A. C. (2018). Unsupervised Domain Adaptation for Face Anti-Spoofing. *IEEE Transactions on Information Forensics and Security*, *13*(7), 1794–1809. https://doi.org/10.1109/TIFS.2018.2801312

Li, J., Wang, Y., Tan, T., Jain, A. K., Li, J., Wang, Y., Tan, T., & Jain, A. K. (2004). Live face detection based on the analysis of Fourier spectra. *SPIE*, *5404*, 296–303. https://doi.org/10.1117/12.541955

Li, L., Feng, X., Xia, Z., Jiang, X., & Hadid, A. (2018). Face spoofing detection with local binary pattern network. *Journal of Visual Communication and Image Representation*, *54*, 182–192. https://doi.org/10.1016/j.jvcir.2018.05.009

Li, L., Xia, Z., Jiang, X., Roli, F., & Feng, X. (2020). CompactNet: learning a compact space for face presentation attack detection. *Neurocomputing*, *409*, 191–207. https://doi.org/10.1016/j.neucom.2020.05.017

Li, L., Xia, Z., Wu, J., Yang, L., & Han, H. (2022). Face presentation attack detection based on optical flow and texture analysis. *Journal of King Saud University - Computer and Information Sciences*, *34*(4), 1455–1467. https://doi.org/10.1016/j.jksuci.2022.02.019

Li, X., Komulainen, J., Zhao, G., Yuen, P.-C., & Pietikainen, M. (2016). Generalized face anti-spoofing by detecting pulse from face videos. *Proceedings - International Conference on Pattern Recognition*, *0*, 4244–4249. https://doi.org/10.1109/ICPR.2016.7900300

Liu, R., Liu, Y., Yan, Y., & Wang, J. (2020). Iterative Deep Neighborhood: A Deep Learning Model Which Involves Both Input Data Points and Their Neighbors. *Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2020/9868017

Liu, Y., Jourabloo, A., & Liu, X. (2018). Learning Deep Models for Face Anti-Spoofing: Binary or Auxiliary Supervision. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 389–398. https://doi.org/10.1109/CVPR.2018.00048

Ma, Y., Wu, L., Li, Z., & liu, F. (2020). A novel face presentation attack detection scheme based on multi-regional convolutional neural networks. *Pattern Recognition Letters*, *131*, 261–267. https://doi.org/10.1016/j.patrec.2020.01.002

Määttä, J., Hadid, A., & Pietikäinen, M. (2011). Face spoofing detection from single images using micro-texture analysis. *2011 International Joint Conference on Biometrics, IJCB 2011*. https://doi.org/10.1109/IJCB.2011.6117510

Määttä, J., Hadid, A., & Pietikäinen, M. (2012). Face spoofing detection from single images using texture and local shape analysis. *IET Biometrics*, *1*(1), 3–10. https://doi.org/10.1049/iet-bmt.2011.0009

Mantach, S., Lutfi, A., Tavasani, H. M., Ashraf, A. B., El-Hag, A. H., & Kordi, B. (2022). Deep Learning in High Voltage Engineering: A Literature Review. *Energies*. https://doi.org/10.3390/en15145005

Ming, Z., Visani, M., Luqman, M. M., & Burie, J.-C. (2020). *A Survey On Anti-Spoofing Methods For Face Recognition with RGB Cameras of Generic Consumer Devices*. http://arxiv.org/abs/2010.04145

Mo, R. (2021). Review of Neural Network Algorithm and Its Application in Reactive Distillation. *Asian Journal of Chemical Sciences*. https://doi.org/10.9734/ajocs/2021/v9i319073

Mohammadi, A., Bhattacharjee, S., & Marcel, S. (2020). Domain Adaptation for

Generalization of Face Presentation Attack Detection in Mobile Settengs with Minimal Information. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, *2020-May*, 1001–1005. https://doi.org/10.1109/ICASSP40776.2020.9053685

Muhammad, U., Yu, Z., & Komulainen, J. (2022). Self-supervised 2D face presentation attack detection via temporal sequence sampling. *Pattern Recognition Letters*, *156*, 15–22. https://doi.org/10.1016/j.patrec.2022.03.001

Murali, S., & Govindan, V. K. (2013). Shadow detection and removal from a single image: Using LAB color space. *Cybernetics and Information Technologies*, *13*(1), 95–103. https://doi.org/10.2478/cait-2013-0009

Nema, A. (2020). Ameliorated Anti-Spoofing Application for PCs with Users' Liveness Detection Using Blink Count. *2020 International Conference on Computational Performance Evaluation, ComPE 2020*, 311–315. https://doi.org/10.1109/ComPE49325.2020.9200166

Neu, D. A., Lahann, J., & Fettke, P. (2021). A Systematic Literature Review on State-of-the-Art Deep Learning Methods for Process Prediction. *Artificial Intelligence Review*. https://doi.org/10.1007/s10462-021-09960-8

Nguyen, D. T., Pham, T. D., Baek, N. R., & Park, K. R. (2018). Combining deep and handcrafted image features for presentation attack detection in face recognition systems using visible-light camera sensors. *Sensors*, *18*(3). https://doi.org/10.3390/s18030699

Nguyen, H. P., Delahaies, A., Retraint, F., & Morain-Nicolier, F. (2019). Face Presentation Attack Detection Based on a Statistical Model of Image Noise. *IEEE Access*, *7*, 175429–175442. https://doi.org/10.1109/ACCESS.2019.2957273

Nikisins, O., Mohammadi, A., Anjos, A., & Marcel, S. (2018). On effectiveness of anomaly detection approaches against unseen presentation attacks in face anti-spoofing. *Proceedings - 2018 International Conference on Biometrics, ICB 2018*, 75–81. https://doi.org/10.1109/ICB2018.2018.00022

Ojala, T., Maenpaa, T., Pietikainen, M., Viertola, J., Kyllonen, J., & Huovinen, S. (2002). Outex - new framework for empirical evaluation of texture analysis algorithms. *2002 International Conference on Pattern Recognition*, *1*, 701–706 c.1. https://doi.org/10.1109/ICPR.2002.1044854

Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *24*(7), 971–987. https://doi.org/10.1109/TPAMI.2002.1017623

Pan, G., Sun, L., Wu, Z., & Lao, S. (2007). Eyeblink-based anti-spoofing in face recognition from a generic webcamera. *ICCV*, 1–8.

Pan, S., & Deravi, F. (2018). Facial biometrie presentation attack detection using temporal texture co-occurrence. *2018 IEEE 4th International Conference on Identity, Security, and Behavior Analysis, ISBA 2018*, *2018-January*, 1–7.

https://doi.org/10.1109/ISBA.2018.8311464

Patel, K., Han, H., & Jain, A. K. (2016). Secure Face Unlock: Spoof Detection on Smartphones. *IEEE Transactions on Information Forensics and Security*, *11*(10), 2268–2283. https://doi.org/10.1109/TIFS.2016.2578288

Pei, M., Yan, B., Hao, H., & Zhao, M. (2023). Person-Specific Face Spoofing Detection Based on a Siamese Network. *Pattern Recognition*, *135*. https://doi.org/10.1016/j.patcog.2022.109148

Peng, F., Meng, S.-H., & Long, M. (2022). Presentation attack detection based on two-stream vision transformers with self-attention fusion. *Journal of Visual Communication and Image Representation*, *85*. https://doi.org/10.1016/j.jvcir.2022.103518

Peng, F., Qin, L., & Long, M. (2018a). CCoLBP: Chromatic co-occurrence of local binary pattern for face presentation attack detection. *Proceedings - International Conference on Computer Communications and Networks, ICCCN*, *2018-July*. https://doi.org/10.1109/ICCCN.2018.8487325

Peng, F., Qin, L., & Long, M. (2018b). Face presentation attack detection using guided scale texture. *Multimedia Tools and Applications*, *77*(7), 8883–8909. https://doi.org/10.1007/s11042-017-4780-0

Peng, J., & Chan, P. P. K. (2014). Face liveness detection for combating the spoofing attack in face recognition. *International Conference on Wavelet Analysis and Pattern Recognition*, *2014-January*, 176–181. https://doi.org/10.1109/ICWAPR.2014.6961311

Phan, Q. T., Dang-Nguyen, D. T., Boato, G., & De Natale, F. G. B. (2016). FACE spoofing detection using LDP-TOP. *Proceedings - International Conference on Image Processing, ICIP*. https://doi.org/10.1109/ICIP.2016.7532388

Pithadia, P. V., Gajjar, P. P., & Dave, J. V. (2012). Feature preserving super-resolution use of LBP and DWT. *2012 International Conference on Devices, Circuits and Systems, ICDCS 2012*, 456–461. https://doi.org/10.1109/ICDCSYST.2012.6188798

Porebski, A., Vandenbroucke, N., & Macaire, L. (2013). Supervised texture classification: color space or texture feature selection? *Pattern Analysis and Applications*, *16*(1), 1–18. https://doi.org/10.1007/s10044-012-0291-9

Raghavendra, R., & Busch, C. (2014). Robust 2D/3D face mask presentation attack detection scheme by exploring multiple features and comparison score level fusion. *FUSION 2014 - 17th International Conference on Information Fusion*. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84910616626&partnerID=40&md5=a8271d02d9450a89978f94c79cde8ba1

Raghavendra, R., Raja, K. B., & Busch, C. (2015). Presentation attack detection for face recognition using light field camera. *IEEE Transactions on Image Processing*, *24*(3), 1060–1075. https://doi.org/10.1109/TIP.2015.2395951

Ramachandra, R., & Busch, C. (2017). Presentation attack detection methods for face recognition systems: A comprehensive survey. *ACM Computing Surveys*, *50*(1).

https://doi.org/10.1145/3038924

Ranganayaki, V., & Deepa, S. N. (2016). An Intelligent Ensemble Neural Network Model for Wind Speed Prediction in Renewable Energy Systems. *The Scientific World Journal*. https://doi.org/10.1155/2016/9293529

Rehman, Y. A. U., Po, L. M., & Liu, M. (2018). LiveNet: Improving features generalization for face liveness detection using convolution neural networks. *Expert Systems with Applications*, *108*, 159–169. https://doi.org/10.1016/j.eswa.2018.05.004

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning Representations by Back-Propagating Errors. *Nature*. https://doi.org/10.1038/323533a0

Rusia, M. K., & Singh, D. K. (2022). A Color-Texture-Based Deep Neural Network Technique to Detect Face Spoofing Attacks. *Cybernetics and Information Technologies*. https://doi.org/10.2478/cait-2022-0032

Rusin, D. S., Alehina, A., Safonova, A., & Dmitriev, E. V. (2021). Using Deep Learning Algorithms for Texture Segmentation of Ultra-High Resolution Satellite Images. *E3s Web of Conferences*. https://doi.org/10.1051/e3sconf/202133301010

Sandid, F., & Douik, A. (2016). Robust color texture descriptor for material recognition. *Pattern Recognition Letters*, *80*, 15–23. https://doi.org/https://doi.org/10.1016/j.patrec.2016.05.010

Sarigul, M. (2023). Görüntü Sınıflandırma Görevi Için CNN Kanal Dikkat Modüllerinin Performans Analizi. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*. https://doi.org/10.21605/cukurovaumfd.1273688

Saufi, M. M., Zamanhuri, M. A. B., Mohammad, N., & Ibrahim, Z. B. (2018). Deep Learning for Roman Handwritten Character Recognition. *Indonesian Journal of Electrical Engineering and Computer Science*. https://doi.org/10.11591/ijeecs.v12.i2.pp455-460

Sebastien, M., Nixon, M., & Li, S. (2014). *Handbook of biometric anti-spoofing: trusted biometrics under spoofing attacks*.

Sepas-Moghaddam, A., Correia, P. L., & Pereira, F. (2017). Light field local binary patterns description for face recognition. *2017 IEEE International Conference on Image Processing (ICIP)*, 3815–3819. https://doi.org/10.1109/ICIP.2017.8296996

Sharma, D., & Selwal, A. (2023). A survey on face presentation attack detection mechanisms: hitherto and future perspectives. *Multimedia Systems*, *29*(3), 1527–1577. https://doi.org/10.1007/s00530-023-01070-5

Shi, L., Zhou, Z., & Guo, Z. (2021). Face Anti-Spoofing Using Spatial Pyramid Pooling. *2020 25th International Conference on Pattern Recognition (ICPR)*, 2126–2133. https://doi.org/10.1109/ICPR48806.2021.9412407

Shu, X., Li, X., Zuo, X., Xu, D., & Shi, J. (2023). Face spoofing detection based on multi-scale color inversion dual-stream convolutional neural network. *Expert Systems with Applications*, *224*, 119988. https://doi.org/10.1016/J.ESWA.2023.119988

Shu, X., Tang, H., & Huang, S. (2021). Face spoofing detection based on chromatic ED-LBP texture feature. *Multimedia Systems*, *27*(2), 161–176. https://doi.org/10.1007/s00530-020-00719-9

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85083953063&partnerID=40&md5=a5b2d6b3fc9f0a6f92864467079a1280

Singh, M., & Arora, A. S. (2017). A robust anti-spoofing technique for face liveness detection with morphological operations. *Optik*, *139*, 347–354. https://doi.org/10.1016/j.ijleo.2017.04.004

Smith, D. F., Wiliem, A., & Lovell, B. C. (2015). Face recognition on consumer devices: Reflections on replay attacks. *IEEE Transactions on Information Forensics and Security*, *10*(4), 736–745.

Souza, G. B. De, Member, S., Felipe, D., Pires, R. G., Marana, A. N., & Papa, J. P. (2017). Deep Texture for RoboutFace Spoofing Detection. *IEEE Transactions on Circuits and Systems II: Express Briefs*, *64*(12), 1397–1401.

Souza, L., Oliveira, L., Pamplona, M., & Papa, J. (2018). How far did we get in face spoofing detection? *Engineering Applications of Artificial Intelligence*, *72*, 368–381. https://doi.org/10.1016/j.engappai.2018.04.013

Sthevanie, F., & Ramadhani, K. N. (2018). Spoofing detection on facial images recognition using LBP and GLCM combination. *Journal of Physics: Conference Series*, *971*(1). https://doi.org/10.1088/1742-6596/971/1/012014

Sun, W., Song, Y., Chen, C., Huang, J., & Kot, A. C. (2020). Face Spoofing Detection Based on Local Ternary Label Supervision in Fully Convolutional Networks. *IEEE Transactions on Information Forensics and Security*, *15*, 3181–3196. https://doi.org/10.1109/TIFS.2020.2985530

Sun, W., Song, Y., Zhao, H., & Jin, Z. (2020). A Face Spoofing Detection Method Based on Domain Adaptation and Lossless Size Adaptation. *IEEE Access*, *8*, 66553–66563. https://doi.org/10.1109/ACCESS.2020.2985453

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *07-12-June-2015*, 1–9. https://doi.org/10.1109/CVPR.2015.7298594

Tan, X., Li, Y., Liu, J., & Jiang, L. (2010). Face Liveness Detection from a Single Image with Sparse Low Rank Bilinear Discriminative Model. Içinde K. Daniilidis, P. Maragos, & N. Paragios (Ed.), *Computer Vision – ECCV 2010* (ss. 504–517). Springer Berlin Heidelberg.

Tan, X., Song, F., Zhou, Z.-H., & Chen, S. (2009). Enhanced Pictorial Structures for precise eye localization under incontrolled conditions. *2009 IEEE Conference on Computer*

Vision and Pattern Recognition, 1621–1628. https://doi.org/10.1109/CVPR.2009.5206818

Tang, D., Zhou, Z., Zhang, Y., & Zhang, K. (2018). Face flashing: a secure liveness detection protocol based on light reflections. *arXiv preprint arXiv:1801.01949*.

Teh, P. S., Zhang, N., Teoh, A. B. J., & Chen, K. (2016). A survey on touch dynamics authentication in mobile devices. *Computers & Security*, *59*, 210–235. https://doi.org/https://doi.org/10.1016/j.cose.2016.03.003

Teja, M. H. (2011). Real-time live face detection using face template matching and DCT energy analysis. *Proceedings of the 2011 International Conference of Soft Computing and Pattern Recognition, SoCPaR 2011*, 342–346. https://doi.org/10.1109/SoCPaR.2011.6089267

Tian, Y., & Xiang, S. (2016). Detection of video-based face spoofing using LBP and multiscale DCT. *International Workshop on Digital Watermarking*, 16–28.

Tirunagari, S., Poh, N., Windridge, D., Iorliam, A., Suki, N., & Ho, A. T. S. (2015). Detection of face spoofing using visual dynamics. *IEEE Transactions on Information Forensics and Security*, *10*(4), 762–777. https://doi.org/10.1109/TIFS.2015.2406533

Tu, X., & Fang, Y. (2017). Ultra-deep Neural Network for Face Anti-spoofing. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10635 LNCS*, 686–695. https://doi.org/10.1007/978-3-319-70096-0_70/COVER

Turk, M., & Pentland, A. (1991). Eigenfaces for Recognition. *Journal of Cognitive Neuroscience*, *3*(1), 71–86. https://doi.org/10.1162/jocn.1991.3.1.71

Vezhnevets, V., Sazonov, V., & Andreeva, A. (2003). A survey on pixel-based skin color detection techniques. *Proc. Graphicon*, *3*, 85–92.

Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International journal of computer vision*, *57*(2), 137–154.

Wang, C., Yu, B., & Zhou, J. (2023). A Learnable Gradient operator for face presentation attack detection. *Pattern Recognition*, *135*. https://doi.org/10.1016/j.patcog.2022.109146

Wang, F., Jiang, M., Qian, C., Yang, S., Cheng, L., Zhang, H., Wang, X., & Tang, X. (2017). *Residual Attention Network for Image Classification*. https://doi.org/10.1109/cvpr.2017.683

Wang, G., Han, H., Shan, S., & Chen, X. (2019). Improving Cross-database Face Presentation Attack Detection via Adversarial Domain Adaptation. *2019 International Conference on Biometrics, ICB 2019*. https://doi.org/10.1109/ICB45273.2019.8987254

Wang, G., Han, H., Shan, S., & Chen, X. (2020). Cross-domain Face Presentation Attack Detection via Multi-domain Disentangled Representation Learning. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*,

6677–6686. https://doi.org/10.1109/CVPR42600.2020.00671

Wang, G., Han, H., Shan, S., & Chen, X. (2021). Unsupervised Adversarial Domain Adaptation for Cross-Domain Face Presentation Attack Detection. *IEEE Transactions on Information Forensics and Security*, *16*, 56–69. https://doi.org/10.1109/TIFS.2020.3002390

Wang, H., Tan, Y., & Li, P. (2022). *A NLOS Identification Algorithm of UWB Signal via Temporal Convolutional Neural Network and Attention Mechanism*. https://doi.org/10.21203/rs.3.rs-1839518/v1

Wang, S.-Y., Yang, S.-H., Chen, Y.-P., & Huang, J.-W. (2017). Face liveness detection based on skin blood flow analysis. *Symmetry*, *9*(12). https://doi.org/10.3390/sym9120305

Wang, Y., Ge, X., Ma, H., Qi, S., Zhang, G., & Yao, Y.-D. (2021). Deep Learning in Medical Ultrasound Image Analysis: A Review. *Ieee Access*. https://doi.org/10.1109/access.2021.3071301

Wang, Y., Nian, F., Li, T., Meng, Z., & Wang, K. (2017). Robust face anti-spoofing with depth information. *Journal of Visual Communication and Image Representation*, *49*, 332–337. https://doi.org/10.1016/j.jvcir.2017.09.002

Wang, Z., Simoncelli, E. P., & Bovik, A. C. (2003). Multi-scale structural similarity for image quality assessment. *Conference Record of the Asilomar Conference on Signals, Systems and Computers*, *2*(Ki L), 1398–1402. https://doi.org/10.1109/acssc.2003.1292216

Waris, M.-A., Zhang, H., Ahmad, I., Kiranyaz, S., & Gabbouj, M. (2013). Analysis of textural features for face biometric anti-spoofing. *21st European Signal Processing Conference (EUSIPCO 2013)*, 1–5.

Watkins, C. P., Dehaene, O., & Dehaene, S. (2019). *Automatic Construction of a Phonics Curriculum for Reading Education Using the Transformer Neural Network*. https://doi.org/10.1007/978-3-030-23207-8_42

Wen, D., Han, H., & Jain, A. K. (2015). Face spoof detection with image distortion analysis. *IEEE Transactions on Information Forensics and Security*, *10*(4), 746–761. https://doi.org/10.1109/TIFS.2015.2400395

Xu, Z., Li, S., & Deng, W. (2016). Learning temporal features using LSTM-CNN architecture for face anti-spoofing. *Proceedings - 3rd IAPR Asian Conference on Pattern Recognition, ACPR 2015*, 141–145. https://doi.org/10.1109/ACPR.2015.7486482

Yan, J., Zhang, Z., Lei, Z., Yi, D., & Li, S. Z. (2012). Face liveness detection by exploring multiple scenic clues. *2012 12th International Conference on Control, Automation, Robotics and Vision, ICARCV 2012*, 188–193. https://doi.org/10.1109/ICARCV.2012.6485156

Yang, J., Lei, Z., & Li, S. Z. (2014). Learn convolutional neural network for face anti-

spoofing. *CoRR*.

Yang, J., Lei, Z., Liao, S., & Li, S. Z. (2013). Face liveness detection with component dependent descriptor. *2013 International Conference on Biometrics (ICB)*, 1–6. https://doi.org/10.1109/ICB.2013.6612955

Yang, J., Lei, Z., Yi, D., & Li, S. Z. (2015). Person-specific face antispoofing with subject domain adaptation. *IEEE Transactions on Information Forensics and Security*, *10*(4), 797–809.

Yang, L. (2014). Face liveness detection by focusing on frontal faces and image backgrounds. *International Conference on Wavelet Analysis and Pattern Recognition*, *2014-January*, 93–97. https://doi.org/10.1109/ICWAPR.2014.6961297

Yılmaz, A. G., Turhal, U., & Nabiyev, V. V. (2020). EFFECT OF FEATURE SELECTION WITH META-HEURISTIC OPTIMIZATION METHODS ON FACE SPOOFING DETECTION. *Journal of Modern Technology & Engineering*, *5*(1).

Zhang, L.-B., Peng, F., Qin, L., & Long, M. (2018). Face spoofing detection based on color texture Markov feature and support vector machine recursive feature elimination. *Journal of Visual Communication and Image Representation*, *51*, 56–69. https://doi.org/10.1016/j.jvcir.2018.01.001

Zhang, S., Zhang, M., Ma, S., Wang, Q., Qu, Y., Sun, Z., & Yang, T. (2021). Research Progress of Deep Learning in the Diagnosis and Prevention of Stroke. *Biomed Research International*. https://doi.org/10.1155/2021/5213550

Zhang, W., & Xiang, S. (2020). Face anti-spoofing detection based on DWT-LBP-DCT features. *Signal Processing: Image Communication*, *89*. https://doi.org/10.1016/j.image.2020.115990

Zhang, Y.-J., Chen, J.-Y., & Lu, Z.-M. (2022). Face anti-spoofing detection based on color texture structure analysis. *Taiwan Ubiquitous Information*, *7*(2).

Zhang, Y., Yin, Z., & Huang, S. (2020). Attention Based Multi-Layer Fusion of Multispectral Images for Pedestrian Detection. *Ieee Access*. https://doi.org/10.1109/access.2020.3022623

Zhang, Z., Yan, J., Liu, S., Lei, Z., Yi, D., & Li, S. Z. (2012). A face antispoofing database with diverse attacks. *2012 5th IAPR International Conference on Biometrics (ICB)*, 26–31. https://doi.org/10.1109/ICB.2012.6199754

Zhao, X., Lin, Y., & Heikkila, J. (2017). Dynamic Texture Recognition Using Volume Local Binary Count Patterns with an Application to 2D Face Spoofing Detection. *IEEE Transactions on Multimedia*, *20*(3), 552–566. https://doi.org/10.1109/TMM.2017.2750415

Zhou, J., Shu, K., Liu, P., Xiang, J., & Xiong, S. (2021). Face Anti-Spoofing Based on Dynamic Color Texture Analysis Using Local Directional Number Pattern. *2020 25th International Conference on Pattern Recognition (ICPR)*, 4221–4228. https://doi.org/10.1109/ICPR48806.2021.9412323

Zhu, N., Yu, Z., & Kou, C. (2020). A New Deep Neural Architecture Search Pipeline for Face Recognition. *Ieee Access*. https://doi.org/10.1109/access.2020.2994207

ÖZGEÇMİŞ

Uğur TURHAL, Lise eğitimini 2005 yılında Trabzon Anadolu Meslek Lisesinde tamamladıktan sonra 2011 yılında Marmara Üniversitesi Bilgisayar ve Kontrol Öğretmenliği Bölümünden, 2017 yılında İstanbul Üniversitesi Bilgisayar Mühendisliği Bölümünden mezun oldu. Yalova Üniversitesinde 2012 yılında başladığı yüksek lisans eğitimini 2016 yılında tamamladı. 2018 yılında Karadeniz Teknik Üniversitesinde Bilgisayar Mühendisliği Anabilim Dalında doktora programına başladı. 2013-2017 yılları arasında Balıkesir Üniversitesi Bilgisayar Mühendisliği Bölümünde Uzman olarak görev yapan Turhal, 2017 yılından bu yana Bayburt Üniversitesi Bilgisayar Teknolojileri Bölümünde Öğretim Görevlisi olarak görev yapmaktadır.

Tez Kapsamındaki Yayınlar

Uluslararası hakemli dergilerde yayınlanan makaleler (SCI/SCI-E)

1. Günay Yılmaz, A., Turhal, U., & Nabiyev, V. (2023). Face Presentation Attack Detection Performances of Facial Regions with Multi-Block LBP Features. Multimedia Tools and Applications, 82(26), 40039–40063. https://doi.org/10.1007/s11042-023-14453-7
2. Turhal, U., Günay Yılmaz, A., & Nabiyev, V. (2023). A New Face Presentation Attack Detection Method Based on Face-Weighted Multi-Color Multi-Level Texture Features. The Visual Computer, 1–16. https://doi.org/10.1007/s00371-023-02866-2

Diğer dergilerde yayınlanan makaleler

1. Yılmaz, A. G., Turhal, U., & Nabiyev, V. (2020). Effect Of Feature Selection with Meta-Heuristic Optimization Methods on Face Spoofing Detection. Journal of Modern Technology & Engineering, 5(1).
2. Günay Yılmaz, A., Turhal, U., & Nabiyev, V. (2022). Farklı Renk Kanallarında Üretilen Doku Özniteliklerinin Yüz Sahteciliği Tespiti Başarımına Etkisinin İncelenmesi. Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi, 15(1), 56–65. https://doi.org/10.54525/tbbmd.1075383