KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YAZILIM MÜHENDİSLİĞİ ANABİLİM DALI

DERIN ÖĞRENME TABANLI SAHTE YÜZ TESPİTİ: MODEL GELİŞTİRME, EĞİTİM OPTİMİZASYONU VE PERFORMANS DEĞERLENDİRMESİ İÇİN KAPSAMLI BİR YAKLAŞIM

YÜKSEK LİSANS TEZİ

Rehman ALI

EKİM 2025
TRABZON

KARADENİZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

YAZILIM MÜHENDİSLİĞİ ANABİLİM DALI

DERIN ÖĞRENME TABANLI SAHTE YÜZ TESPİTİ: MODEL GELİŞTİRME, EĞİTİM OPTİMİZASYONU VE PERFORMANS DEĞERLENDİRMESİ İÇİN KAPSAMLI BİR YAKLAŞIM

Rehman ALI

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünce
"YAZILIM YÜKSEK MÜHENDİSİ"
Unvanı Verilmesi İçin Kabul Edilen Tezdir.

Tezin Enstitüye Verildiği Tarih : 17 / 09 / 2025

Tezin Savunma Tarihi : 17 / 10 / 2025

Tez Danışmanı : Dr. Öğr. Üyesi Asuman GÜNAY YILMAZ

Trabzon 2025

KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Yazılım Mühendisliği Anabilim Dalında
Rehman ALI Tarafından Hazırlanan

DERIN ÖĞRENME TABANLI SAHTE YÜZ TESPİTİ: MODEL GELİŞTİRME, EĞİTİM OPTİMİZASYONU VE PERFORMANS DEĞERLENDİRMESİ İÇİN KAPSAMLI BİR YAKLAŞIM

başlıklı bu çalışma, Enstitü Yönetim Kurulunun 25 / 09 / 2025 gün ve 2121 sayılı kararıyla oluşturulan jüri tarafından yapılan sınavda
YÜKSEK LİSANS TEZİ
olarak kabul edilmiştir.

Jüri Üyeleri

Başkan : Doç Dr. Yunus SANTUR

Üye : Dr. Öğr. Üyesi Asuman GÜNAY YILMAZ

Üye : Dr. Öğr. Üyesi Sefa ARAS

Prof. Dr. Asim KADIOĞLU

Enstitü Müdürü

ÖNSÖZ

“Derin Öğrenme Tabanlı Sahte Yüz Tespiti: Model Geliştirme, Eğitim Optimizasyonu ve Performans Değerlendirmesi İçin Kapsamlı Bir Yaklaşım” isimli bu tez Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsü Yazılım Mühendisliği Anabilim Dalı, Yüksek Lisans Programı’nda hazırlanmıştır.

Tez çalışmam boyunca bilgi ve tecrübeleriyle yolumu aydınlatan, değerli katkıları ve yönlendirmeleriyle çalışmamın her aşamasında bana destek olan danışman hocam Dr. Öğr. Üyesi Asuman Günay Yılmaz’a en içten teşekkürlerimi sunarım.

Ayrıca tez çalışmam sürecinde fikirleri, eleştirileri ve desteği ile katkı sağlayan tüm hocalarıma ve arkadaşlarıma teşekkür ederim.

Bu tez çalışmasının ileride yapılacak araştırmalara katkı sağlamasını temenni ederim.

REHMAN ALI

Trabzon 2025

TEZ ETİK BEYANNAMESİ

Yüksek Lisans Tezi olarak sunduğum “Derin Öğrenme Tabanlı Sahte Yüz Tespiti: Model Geliştirme, Eğitim Optimizasyonu ve Performans Değerlendirmesi İçin Kapsamlı Bir Yaklaşım” başlıklı bu çalışmayı baştan sona kadar danışmanım Dr. Öğr. Üyesi Asuman Günay Yılmaz’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 kurallara uygun olarak davrandığımı ve aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ederim. 17 / 10 / 2025

Rehman ALI

İÇİNDEKİLER

Sayfa No

Yüksek Lisans Tezi

ÖZET

DERIN ÖĞRENME TABANLI SAHTE YÜZ TESPİTİ: MODEL GELİŞTİRME, EĞİTİM OPTİMİZASYONU VE PERFORMANS DEĞERLENDİRMESİ İÇİN KAPSAMLI BİR YAKLAŞIM

REHMAN ALI

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
Yazılım Mühendisliği Anabilim Dalı
Danışman: Dr. Öğr. Üyesi Asuman Günay Yılmaz
2025, 74 Sayfa

Bu tez çalışması, giderek artan derin sahtecilik tehdidine karşı etkili bir tespit sistemi geliştirmek amacıyla yürütülmüştür. Üretken çekişmeli ağların ve evrişimli sinir ağlarının gerçekçi sahte yüz içeriği üretme yeteneği, siber güvenlik ve sosyal medya platformları için ciddi zorluklar oluşturmaktadır. Bu araştırmada, ConvNeXt mimarisi derin sahtecilik tespiti için optimize edilmiş, %61 daha az parametreye (29.1M'den 11.3M'e) sahip yenilikçi aşama konfigürasyonu ile modelin tespit kabiliyeti korunmuştur. Ağın ayırt edici bölgelere odaklanmasını sağlamak üzere Sıkıştır-ve-Uyar dikkat mekanizması kullanılmıştır. Farklı modellerin topluluk öğrenmesi ve Test Zamanı Veri Artırma tekniklerinin birleşimi, tespit doğruluğunda önemli iyileşmeler sağlamıştır. Gerçek görüntü veri seti FaceForensics++ ve CelebA veri setlerinden, manipüle edilmiş görüntüler ise FaceForensics++ sahte videolarından (Deepfakes, Face2Face, FaceSwap, NeuralTextures) elde edilmiştir. Performans değerlendirmesinde kesinlik, duyarlılık, F1-skoru ve AUC-ROC metrikleri kullanılmıştır. Deneysel sonuçlar, önerilen topluluk tabanlı yaklaşımların üstün performans sergilediğini göstermiştir. Bu çalışma, parametre-verimli model tasarımı ve topluluk optimizasyon stratejilerinde katkılar sunarak derin sahtecilik tehditlerine karşı etkili savunma mekanizmaları geliştirmede önemli bir adım teşkil etmektedir.

Anahtar Kelimeler: sahte yüz tespiti, Derin Öğrenme, model geliştirme, eğitim optimizasyonu.

Master's Thesis

SUMMARY

DEEP LEARNING-BASED FAKE FACE DETECTION: A COMPREHENSIVE APPROACH FOR MODEL DEVELOPMENT, TRAINING OPTIMIZATION, AND PERFORMANCE EVALUATION

REHMAN ALI

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Department of Software Engineering
Supervisor: Dr. Öğr. Üyesi Asuman Günay Yılmaz
2025, 74 Pages

This thesis study has been conducted with the aim of developing an effective detection system against the escalating threat of deepfakes. The capability of generative adversarial networks and convolutional neural networks to produce realistic fake facial content poses serious challenges for cybersecurity and social media platforms. In this research, the ConvNeXt architecture has been optimized for deepfake detection, and the model's detection capability has been preserved through an innovative stage configuration with 61% fewer parameters (from 29.1M to 11.3M). The Squeeze-and-Excitation attention mechanism has been employed to enable the network to focus on discriminative regions. The combination of ensemble learning methods and Test-Time Augmentation techniques has provided significant improvements in detection accuracy. The genuine image dataset consists of frames from FaceForensics++ and CelebA datasets, while manipulated images have been obtained from fake videos (Deepfakes, Face2Face, FaceSwap, NeuralTextures) in the FaceForensics++ dataset. Precision, recall, F1-score, and AUC-ROC metrics have been utilized for performance evaluation. Experimental results have demonstrated that the proposed ensemble-based approaches exhibit superior performance. This study constitutes a significant step in developing effective defense mechanisms against deepfake threats by offering contributions in parameter-efficient model design and ensemble optimization strategies.

Key Words: fake face detection, Deep Learning, model development, training optimization.

ŞEKİLLER DİZİNİ

Sayfa No

TABLOLAR DİZİNİ

Sayfa No

SEMBOLLER VE KISALTMALAR DİZİNİ

Matematiksel Semboller:

θ : Ağ parametreleri

α : Öğrenme oranı

ε : Sayısal kararlılık sabiti

σ : Sigmoid aktivasyon fonksiyonu

δ : ReLU aktivasyon fonksiyonu

γ : Yığın normalizasyonu ölçekleme parametresi

β : Yığın normalizasyonu kaydırma parametresi

μ : Ortalama değer

σ^2 : Varyans

$\odot$: Eleman-eleman çarpım

$\sum$: Toplam operatörü

$\mathbb{E}$: Beklenen değer operatörü

$\sim$: Dağılım gösterimi

$\lfloor \rfloor$: Taban fonksiyonu

Kısaltmalar:

SU : Sıkıştır-ve-Uyar

EBDM : Evrişimli Blok Dikkat Modülü

TZA : Test Zamanı Artırımı

VO : Varyasyonel Otokodlayıcı

YN : Yanlış Negatif

YP : Yanlış Pozitif

AUC - ROC : Area Under the Curve - Receiver Operating Characteristics Curve (Eğri Altında Kalan Alan - Alıcı İşletim Karakteristiği)

ÇÜA : Çekişmeli Üretici Ağlar

ESA : Evrişimli Sinir Ağı

1. GENEL BİLGİLER

1.1. Giriş

Bu Dijital çağ, görsel medyanın üretilme, tüketilme ve kullanılma biçimlerine dair köklü değişiklikler getirmektedir. Günümüzde gerçek ve yapay içerik arasındaki sınır giderek bulanıklaşmaktadır. Bilgisayarlı görü alanında yaşanan gelişmeler, son on yılda yüz analizi ve manipülasyon teknolojilerinde önemli atılımlar sağlamaktadır. Bu ilerlemeler yaratıcılık ve teknolojik yenilik açısından önemli fırsatlar sunarken, siber güvenlik, gazetecilik ve sosyal medya alanlarında ciddi zorluklar yaratmaktadır.

Derin öğrenme algoritmaları, gerçeğinden ayırt edilmesi güç yüz görüntüleri üretebilmektedir. Bu durum, uzman kişiler için bile gerçek ve yapay içerik ayrımını zorlaştırmaktadır. Çekişmeli Üretici Ağlar (ÇÜA), Varyasyonel Otokodlayıcılar (VO) ve gelişmiş Evrişimli Sinir Ağları (ESA) gibi teknolojiler, neredeyse gerçek fotoğraflardan ayırt edilemeyen kaliteli insan yüzleri üretebilmektedir. Bu teknolojik başarılar, film yapımı ve dijital sanat gibi meşru alanlarda faydalı olurken, dijital güvenlik ve gizlilik konularında önemli endişeler yaratmaktadır. İkna edici sahte yüz içeriği yaratma yeteneği, günümüzde sadece akademik bir konu olmaktan çıkmaktadır. Toplumun giderek dijitalleştiği bu dönemde, güven ve bilgi doğruluğunun temel taşları etkilenmektedir. Bu güçlü araçların kullanıcı dostu uygulamalar aracılığıyla herkesin erişimine açılması, durumu daha da karmaşık hale getirmektedir. Teknik bilgisi olmayan kişiler bile ileri yüz manipülasyonları yapabilmektedir. Derin sahte teknolojisinin yaygınlaşması, üretim ve tespit yöntemleri arasında sürekli bir yarış yaratmaktadır. Üretici modeller daha gelişmiş hale geldikçe, tespit sistemleri de buna ayak uydurmaya çalışmaktadır. Bu dinamik süreç, modern tespit sistemlerinin geliştirilmesinde dikkat mekanizmaları, topluluk öğrenimi ve çok ölçekli özellik çıkarma gibi gelişmiş tekniklerin kullanılmasına yol açmaktadır (Rossler et al., 2019; Li et al., 2020; Wang et al., 2020).

Yüz manipülasyonu ve tespitinin matematiksel temelleri, optimizasyon teorisi, olasılık teorisi ve sinyal işleme alanlarının kesişiminde bulunmaktadır. Bu teorik temellerin anlaşılması, güvenilir tespit sistemleri geliştirmek için kritik önem taşımaktadır. Modern üretici modeller milyonlarca hatta milyarlarca parametre içermektedir. Bu karmaşıklık, eşit derecede gelişmiş tespit yaklaşımları gerektirmektedir.

1.2. Problem Tanımı ve Araştırma Motivasyonu

Bu araştırmada ele alınan temel problem, gerçek yüzler ile sahte yüzlerin birbirinden ayırt edilmesinin zorluğudur. Bu problem, derin sahte üretim araçlarının yaygınlaşması ve kalitelerinin sürekli artması nedeniyle kritik önem kazanmıştır. Çalışma, çeşitli veri setleri ve manipülasyon teknikleri için, gerçek ve sahte yüz görüntülerini yüksek güvenilirlikle ayırt eden sistemler geliştirmeyi amaçlamaktadır.

Araştırma motivasyonu, hızla gelişen üretim tekniklerine ayak uydurabilen tespit mekanizmaları geliştirme ihtiyacından kaynaklanmaktadır. Bu sistemler yüksek doğruluk, düşük yanlış pozitif oranı ve düşük hesaplama gereksinimi özelliklerini sağlamalıdır. Mevcut tespit sistemleri birkaç temel zorlukla karşılaşmaktadır. İlk olarak, üretici modellerin sürekli gelişmesi tespit algoritmaları için hareketli bir hedef yaratmaktadır. Bu durum, tespit stratejilerinin sürekli adaptasyon gerektirmesine neden olmaktadır. Üretici modeller genellikle çekişmeli olarak eğitilmekte ve mevcut tespit sistemlerini kandırmayı öğrenmektedir. İkinci olarak, çok çeşitli manipülasyon teknikleri mevcuttur. Tam yüz değişiminden ince ifade transferine kadar uzanan bu teknikler, tekniğe özgü eğitim gerektirmeden genelleşebilen tespit sistemleri gerektirmektedir. Üçüncü olarak, eğitim veri setlerinin kalitesi tespit modellerinin genelleme yeteneğini önemli ölçüde etkilemektedir. Daha önce görülmemiş manipülasyon teknikleri, farklı veri dağılımları veya eğitim verilerinde temsil edilmeyen üretim modelleriyle karşılaşıldığında bu durum daha da kritik hale gelmektedir. Son olarak, yüksek doğrulukla tespit için gerekli hesaplama kaynakları, mobil ve gömülü ortamların kısıtlamaları ile çelişmektedir. Bu ortamlarda gecikme, güç tüketimi ve hesaplama kaynakları ciddi şekilde sınırlanabilmektedir.

Teknik değerlendirmelerin ötesinde, derin sahte teknolojisinin rızasız görüntü yaratımında kullanımı, mağdurlarda ciddi psikolojik zararlar yaratmaktadır. Bu tür faaliyetler birçok ülkede suç sayılmaktadır. Derin sahte teknolojisinin politik dezenformasyon, sosyal mühendislik saldırıları ve kimlik hırsızlığı alanlarındaki kullanımı, güvenilir tespit sistemlerine olan ihtiyacı daha da önemli hale getirmiştir (Chesney & Citron, 2019; Vaccari & Chadwick, 2020; Kietzmann et al., 2020).

1.3. Literatür İncelemesi ve Güncel Durum

Yüz manipülasyonu teknolojisi çok hızlı bir gelişim göstermektedir. Başlangıçta önemli manuel müdahale gerektiren basit görüntü düzenleme tekniklerinden, minimal kullanıcı girdisiyle gerçekçi içerik üreten sgelişmiş sinir ağı tabanlı yaklaşımlara kadar uzanan bu süreç, teknolojik ilerlemenin hızını göstermektedir. Erken yöntemler geleneksel bilgisayar grafikleri, şekil değiştirme (morfing) algoritmaları ve manuel düzenleme prosedürlerine dayanmaktaydı. Bu yaklaşımlar önemli teknik uzmanlık gerektirmekteydi ve sonuçları genellikle dikkatli görsel inceleme ile tespit edilebilmekteydi. Destek Vektör Makineleri (DVM) ve basit sinir ağları gibi Makine Öğrenmesi (MÖ) metodolojilerinin tanıtımı, otomatik yüz manipülasyonuna doğru ilk adımı işaret etmektedir. Ancak bu yaklaşımlar ikna edici sonuçlar üretmede sınırlı kalmıştır. Goodfellow ve arkadaşları (2014) tarafından geliştirilen ÇÜA, bu alanda devrimsel bir değişim yaratmıştır. Bu atılım, video içeriğinde benzeri görülmemiş gerçekçilikle yüz değişimi yapabilen ilk derin sahte sistemlerinin yaratılmasını mümkün kılmıştır. Daha sonra geliştirilen ilerici ÇÜA'lar, StyleGAN ve çeşitli oto kodlayıcı yaklaşımları, üretim kalitesinin sınırlarını zorlamaya devam etmiştir. Bu ilerlemeler yüksek kaliteli manipüle edilmiş içerik yaratmanın teknik engellerini azaltmıştır.

Günümüz üretici modelleri son derece karmaşık mimariler kullanmaktadır. Bu sistemler zamansal tutarlılık, yüksek çözünürlük ve minimal yapı bozukluklarıyla gerçekçi yüz sentezi elde etmek için birden fazla sinir ağı bileşenini birleştirmektedir. Tipik olarak dikkat mekanizmalarıyla güçlendirilmiş kodlayıcı-kod çözücü mimarileri, ilerici üretim teknikleri, çekişmeli eğitim prosedürleri ve gelişmiş kayıp fonksiyonları kullanmaktadırlar. Bu yaklaşımların gelişmişliği, üretilmiş içeriğin sadece sıradan gözlemcileri değil aynı zamanda eğitimli uzmanları ve mevcut bazı tespit sistemlerini bile kandırabildiği bir seviyeye ulaşmıştır (Karras et al., 2021; Brock et al., 2019; Choi et al., 2020).

Tespit alanında da araştırmacılar manipüle edilmiş içeriği tanımlamak için giderek daha gelişmiş yaklaşımlar geliştirmişlerdir. Başlangıçtaki tespit yöntemleri, görüntü yapı bozuklukları, çözünürlük tutarsızlıkları ve zamansal titreme gibi belirgin sorunları tanımlamaya odaklanmıştır. Üretim kalitesi iyileştikçe, tespit sistemleri daha ince örüntüleri, istatistiksel tutarsızlıkları ve mikro-ifadeleri analiz etme yönünde kaymaya zorlanmıştır. Bu evrim, modern tespit sistemlerinin temelini oluşturan topluluk tabanlı yaklaşımlar, dikkat

güçlendirilmiş mimariler, alan adaptasyonu teknikleri ve ileri özellik öğrenimi yöntemlerinin geliştirilmesine yol açmıştır.

Derin sahte tespit alanı son yıllarda hızlı bir büyüme göstermiştir. Dünya çapında araştırmacılar manipüle edilmiş yüz içeriğinin tanımlanmasının zorluklarını ele almak için çeşitli yaklaşımlar önermişlerdir. Literatür belirgin ancak genellikle örtüşen birkaç kategoride incelenebilmektedir. Bunlar, el yapımı özelliklere dayanan geleneksel makine öğrenmesi yaklaşımları, ayırt edici temsilleri otomatik öğrenen derin öğrenme tabanlı yöntemler, video kareleri boyunca tutarsızlıklardan faydalanan zamansal tutarlılık analizi, spektral karakteristikleri inceleyen frekans alanı analizi ve birden fazla yaklaşımı birleştiren hibrit topluluk teknikleri olarak sıralanabilir.

1.3.1. Geleneksel Makine Öğrenmesi Yaklaşımları

Geleneksel makine öğrenmesi yaklaşımları, birleştirilmiş el yapımı özellik çıkarma tekniklerinin, DVM, rastgele ormanlar ve lojistik regresyon gibi klasik algoritmalarla birlikte kullanımına odaklanmıştır. Bu yaklaşımlar görüntülerin istatistiksel özelliklerini, doku örüntülerini, frekans alanı karakteristiklerini ve geometrik tutarsızlıkları analiz etmiştir. Yöntemler erken üretim tekniklerine karşı etkili olmakla birlikte, önceden tanımlanmış özelliklere bağımlılıkları ve evrimleşen üretim tekniklerine adapte olamamaları nedeniyle, modern derin öğrenme tabanlı manipülasyon yöntemlerine karşı sınırlı kalmıştır (Matern et al., 2019; Yang et al., 2019; Li & Lyu, 2018).

1.3.2. Derin Öğrenme Tabanlı Tespit Yöntemleri

Derin öğrenme tabanlı tespit yöntemleri günümüzde en etkili yöntemler olarak öne çıkmaktadır. Bu yaklaşımlar gerçek ve manipüle edilmiş yüz içeriği arasındaki ayırt edici özellikleri otomatik öğrenmek için, derin ESA'ların temsil gücünden yararlanmaktadır. Bu kategoride tipik olarak ResNet, EfficientNet, DenseNet ve Görüntü Dönüştürücüleri (GD) gibi önceden eğitilmiş mimariler kullanılmaktadır. Bu mimariler üzerinde transfer öğrenimi prosedürleri aracılığıyla, derin sahte tespit veri setleri için ince ayar yapılmaktadır. Yöntemlerin etkinliği, mikro-ifadeler, ince renk tutarsızlıkları, sıkıştırma yapı bozuklukları ve yüksek frekanslı gürültü örüntüleri gibi görünmez ince manipülasyon yapı bozukluklarını

yakalama yeteneklerinden kaynaklanmaktadır (Dang et al., 2020; Nguyen et al., 2021; Bonettini et al., 2021).

1.3.3. Zamansal Analiz Yaklaşımları

Zamansal analiz yaklaşımları özellikle video tabanlı derin sahte içeriği hedeflemektedir. Yöntemler zamansal sekanslar boyunca tekrar eden tutarsızlıkları ve yapı bozukluklarını belirlemeye çalışmaktadır. Kare-kare varyasyonları, yüz ifadelerinin zamansal tutarlılığını, aydınlatma koşullarını ve video çerçeveleri boyunca yüz işaretlerinin kararlılığını analiz etmektedir. Temel varsayım, derin sahte videolarında bireysel kareler gerçekçi görünebilirken, tüm yönler boyunca zamansal tutarlılığı korumanın üretim algoritmaları için zor olmasıdır. Bu yöntemler video derin sahte tespiti için umut verici olsa da, statik görüntü manipülasyonuna uygulanamamaktadır (Güera & Delp, 2018; Sabir et al., 2019; Amerini et al., 2019).

1.3.4. Evrişimli Sinir Ağı Mimarileri

ESA'lar, neredeyse tüm etkili derin sahte tespit sistemlerinin omurgasını oluşturmaktadır. Bu modeller görsel bilgiyi işleme, birden fazla ölçekte hiyerarşik özellikler çıkarma ve manipülasyon yapı bozukluklarının göstergesi olan karmaşık uzamsal örüntüleri öğrenmedeki etkinlikleri nedeniyle tercih edilmektedir. ESA mimarilerinin evrimi, tespit sistemlerinin gelişimini doğrudan etkilemiştir. Her büyük mimari yenilik, derin sahte tespit alanına yeni yetenekler getirmiştir. Çok derin ağların eğitimini sağlayan kalıntı bağlantılar, özellik yeniden kullanımını teşvik eden yoğun bağlantı yapıları, hesaplama kaynaklarını odaklayan dikkat mekanizmaları ve farklı uzamsal çözünürlüklerde yapı bozukluklarını yakalayan çok ölçekli işleme yaklaşımları sistematik olarak keşfedilmiştir. ResNet mimari ailesi, derin sahte tespit araştırmalarında özellikle etkili olmuştur. Bu mimari, çok derin ağları eğitme yetenekleri aracılığıyla birçok tespit sistemi için güçlü bir temel sağlamıştır. Kalıntı bağlantıların temel yeniliği, bilginin doğrudan atlama bağlantıları yoluyla akmasına olanak tanıyarak kaybolan gradyanlar problemlerinden kaçınmasıdır. Bu yetenek, manipülasyon yapı bozukluklarının ince doğası göz önüne alındığında kritik önem taşımaktadır. Güvenilir tespit genellikle çok derin özellik hiyerarşileri gerektirmektedir.

Birçok etkili tespit sistemi ResNet mimarileri üzerine inşa edilmiştir (He et al., 2016; Huang et al., 2017; Zagoruyko & Komodakis, 2016).

Son dönemde, evrişimli işlemlerin hesaplama verimliliğini GD'den ilham alan tasarım ilkeleriyle birleştiren ConvNeXt mimarisi önemli bir ilerleme göstermektedir. Bu hibrit yaklaşım, geleneksel evrişimli ağların çağdaş tasarım ilkeleriyle modernize edilmesinin bir örneğidir. ConvNeXt temsil kapasitesini korurken hesaplama yükünü azaltan derinlik bazlı evrişimler, katman normalizasyonu, GELU etkinleştirme fonksiyonları ve ters darboğaz yapıları kullanmaktadır. Bu yaklaşım derin sahte tespitinde dikkat çekici performans sergilemektedir (Liu et al., 2022; Dosovitskiy et al., 2021; Touvron et al., 2021).

Topluluk öğrenimi, tespit doğruluğu, güvenilirliği ve genelleme yeteneğini geliştirmek için özellikle etkili bir yaklaşım olarak ortaya çıkmıştır. Farklı mimari karakteristiklere sahip birden fazla modeli birleştirerek, önemli ölçüde daha iyi genelleme performansı elde edilmesini sağlamaktadır. Topluluk öğrenmenin teorik temeli önyargı-varyans ayrışımına dayanmaktadır. Bu teoriye göre, farklı tahmin edicileri birleştirmek hem önyargıyı hem de varyansı azaltabilmektedir. Sert oylama, yumuşak oylama, ağırlıklı birleştirme ve çok ölçekli füzyon teknikleri gibi çeşitli stratejiler geliştirilmiştir (Zhou et al., 2021; Feng et al., 2022; Verdoliva, 2020).

1.4. Matematiksel Temeller

Yüz manipülasyonu ve tespitinin matematiksel ilkelerinin anlaşılması, etkili tespit sistemleri geliştirmek için önemlidir. Teorik çerçeve sinir ağlarını eğitmek için optimizasyon teorisi, belirsizliği modelleme için olasılık teorisi, temel sınırları anlamak için bilgi teorisi, frekans alanı karakteristikleri için sinyal işleme ve çekişmeli dinamikleri anlamak için oyun teorisi gibi birbirine bağlı alanları kapsamaktadır. Modern derin sahte sistemlerinin çoğunun temelini oluşturan ÇÜA'lar, oyun teorisi ilkelerinde dayalı olup yüksek boyutlu uzaylarda minimax optimizasyonunun gelişmiş uygulamasını temsil etmektedir. Çekişmeli eğitim süreci, üretici ve ayırt edici ağların mükemmel bilgiye sahip sıfır toplamlı oyunda rekabet ettiği karmaşık bir minimax optimizasyon problemini içermektedir.

$$\min_G \max_D V(D,G) = E_{x\sim p_{data(x)}}[\log D(x)] + E_{z\sim p_{z(z)}}\left[\log\left(1-D\left(G(z)\right)\right)\right] \quad (1)$$

Denklem (1), üretici çekişmeli ağların eğitim dinamiklerini yöneten temel amaç fonksiyonunu temsil eder; burada G gerçekçi yüz içeriği yaratmaktan sorumlu üretici ağı, D gerçek ve üretilmiş örnekler arasında ayrım yapmaya çalışan ayırt edici ağı temsil eder. $p_{data}(x)$ eğitim veri setindeki gerçek yüz görüntülerinin dağılımını, $p_z(z)$ üreticiye girdi olarak hizmet eden gürültü vektörleri üzerindeki öncül dağılımı temsil eder ve beklenti operatörleri ilgili dağılımlarından tüm olası örnekler üzerindeki ortalama performansı yakalar. Üretici bu amaç fonksiyonunu minimize etmeye çalışırken aynı zamanda ayırt ediciyi kandırma yeteneğini maksimize eder, ayırt edici ise üretilmiş içeriği tanımlamada giderek daha etkili hale gelerek amacı maksimize etmeye çalışır, bu da rekabetçi öğrenme yoluyla üretim kalitesini sürekli olarak iyileştiren karakteristik çekişmeli eğitim dinamiğine yol açar.

Derin sahte tespit sistemleri tipik olarak ikili sınıflandırma için çapraz-entropi kaybını kullanmaktadır. Bu kayıp fonksiyonu tahmin edilen olasılık dağılımları ile gerçek etiketler arasındaki benzemezliği ölçmektedir.

$$L_{CE} = -\left(\frac{1}{N}\right)\sum_{i=1}^{N}[y_i \log(p_i) + (1-y_i)\log(1-p_i)] \quad (2)$$

Denklem (2)'de, L_{CE} tespit sistemi tarafından yapılan tahminlerin kalitesini nicelleyen çapraz-entropi kayıp fonksiyonunu temsil eder. N eğitim yığınındaki toplam örnek sayısını gösterir. y_i sahte içeriğin 0 ve gerçek içeriğin 1 olduğu i. örnek için gerçek ikili etiketi temsil eder ve p_i i. örneğin tespit ağı tarafından tahmin edildiği şekliyle gerçek olma olasılığını temsil eder. Bu kayıp fonksiyonu, güvenli yanlış tahminlere daha yüksek kayıp değerleri ve doğru tahminlere daha düşük kayıp değerleri atayarak yanlış tahminleri etkili bir şekilde cezalandırır, böylece modeli gerçek dağılıma yakından uyan olasılık tahminleri çıkarmaya teşvik eder.

Derin sahte tespiti için sinir ağlarının optimizasyonu, karmaşık optimizasyon uzayında stokastik gradyan inişi ve türevlerine dayanan ileri yöntemler içermektedir. Modern optimizasyon algoritmaları yakınsamayı hızlandıran momentum terimlerini, adaptif

öğrenme oranı ayarlamalarını ve aşırı öğrenmeyi önleyen düzenlileştirme tekniklerini birleştirmektedir.

$$\theta_{t+1} = \theta_t - \left(\alpha\sqrt{\hat{v}_t + \varepsilon}\right) \times \hat{m}_t \quad (3)$$

Denklem (3), Adam optimizasyon algoritması için parametre güncelleme kuralını temsil eder. Burada θ_t zaman adımı *t*'deki ağ parametrelerini, α parametre güncellemelerinin adım boyutunu kontrol eden öğrenme oranını, $\hat{m}_t$ ve $\hat{v}_t$ sırasıyla gradyanların birinci ve ikinci momentlerinin önyargı-düzeltilmiş tahminlerini, ε ise sıfıra bölmeyi önlemek için eklenen küçük bir sabiti temsil eder. Adam algoritması, momentum tabanlı yöntemlerin faydalarını adaptif öğrenme oranı ayarlamasıyla birleştirerek, farklı parametre ölçeklerine ve gradyan karakteristiklerine adapte olurken karmaşık optimizasyon uzaylarında verimli bir gezinme sağlar. Bu da onu derin sahte tespit araştırmalarında kullanılan çeşitli ve zorlu veri setleri üzerinde derin sinir ağlarını eğitmek için etkili kılar.

Dikkat mekanizmaları, derin sahte tespit sistemlerinde giderek önemli hale gelmiştir. Hesaplama kaynaklarını, girdi görüntülerinin en ayırt edici bölgeleri üzerinde seçici olarak odaklama konusundaki yetenekleri nedeniyle, yaygın olarak benimsenmektedirler. Dikkat mekanizmaları, sinir ağlarının girdinin daha ayırt edici özellikleri vurgulamasını sağlamaktadır. Bu yetenek, ince manipülasyon yapı bozukluklarının belirli yüz bölgelerine lokalize olabileceği derin sahte tespit uygulamalarında özellikle değerlidir. Sıkıştır-ve-Uyar (SU) dikkat mekanizması özellikle etkili bir yaklaşımdır. Belirli bir sınıflandırma görevi için, özellik haritalarını ayırt edici önemlerine dayanarak adaptif olarak yeniden kalibre eden kanal bazlı dikkat stratejisi kullanmaktadır.

$$SE(X) = X \odot \sigma\left(W_2\delta\left(W_1 \cdot GAP(X)\right)\right) \quad (4)$$

Denklem (4), girdi özellik haritalarına uygulanan tam Sıkıştır-ve-Uyar işlemini tanımlamaktadır. Denklemde *X* yığın boyutu, kanallar, yükseklik ve genişlik boyutlarına sahip girdi özellik tensörünü temsil eder. GAP (Global Average Pooling) kanal bilgisini korurken uzamsal boyutları sıkıştıran global ortalama havuzlama işlemini gösterir. W_1 ve W_2 sırasıyla boyut azaltma ve genişletme uygulayan, öğrenilen ağırlık matrisleridir. δ ReLU aktivasyon fonksiyonunu, σ ise dikkat ağırlıklarının 0 ve 1 arasında sınırlanmasını sağlayan sigmoid aktivasyon fonksiyonunu gösterir. ⊙, orijinal özellik haritalarının, öğrenilen dikkat

ağırlıklarıyla öğe bazlı çarpılmasını ifade eder. Bu mekanizma, ağın hesaplama yükünü önemli ölçüde artırmadan ağın temsil gücünü geliştirir.

Evrişim işlemi, neredeyse tüm derin sahte tespit mimarilerinin temel yapı taşını oluşturmaktadır. Bu işlem görsel girdiden yerel örüntüleri ve hiyerarşik özellikleri çıkarmak için matematiksel olarak güçlü ve hesaplama açısından verimli bir yaklaşım sunmaktadır.

$$(I * K)(x, y) = \sum_{i=-a}^{a} \sum_{j=-b}^{b} I(x - i, y - j) \cdot K(i, j) \tag{5}$$

Denklem (5), görüntü işleme ve bilgisayarlı görünün temelini oluşturan ayrık 2B evrişim işlemini temsil eder. Burada *I* girdi görüntüsünü veya özellik haritasını, *K* evrişim çekirdeğini veya filtreyi gösterir. (*x,y*) çıktı özellik haritasındaki uzamsal koordinatlardır. Bu matematiksel işlem, çekirdeğin her uzamsal konumda yerel görüntü bölgeleriyle iç çarpımını hesaplayarak yerel özelliklerin çıkarılmasını sağlar. Böylece manipülasyon yapı bozukluklarının veya gerçek içerik karakteristiklerinin göstergesi olabilecek kenarları, dokuları ve diğer örüntüleri tespit edebilen bir şablon eşleştirme biçimini etkili bir şekilde uygular.

1.5. Veri Seti Karakteristikleri ve Önişleme Metodolojileri

Derin sahte tespiti araştırmaları için en kapsamlı kıyaslama verisini FaceForensics++ veri seti (Rossler et al., 2019) sunmaktadır. Bu veri seti, araştırmacılara birden fazla manipülasyon tekniği, kalite seviyesi ve demografik karakteristiği kapsayan çeşitli manipüle edilmiş ve gerçek yüz videolarına erişim imkanı sağlamaktadır. Bu veri seti, tespit sistemi performansını değerlendirmek ve farklı yaklaşımlar arasında adil karşılaştırmalar sağlamak için standart haline gelmiştir. Veri seti bir kişinin yüzünü başkasınınkiyle değiştiren FaceSwap, yüz ifadelerini aktaran Face2Face, yüksek kaliteli yüz değişimi gerçekleştiren FaceShifter ve yüz ifadelerini manipüle eden NeuralTextures gibi çeşitli manipülasyon tekniklerini içermektedir (Rossler et al., 2019; Dolhansky et al., 2020; Li et al., 2021).

Önişleme metodolojileri, derin sahte tespit sistemlerinin nihai etkinliğini belirlemede kritik rol oynamaktadır. Girdi verilerinin kalitesi, sinir ağlarının ayırt edici özellikleri öğrenme ve yeni örneklere genelleme yeteneğini doğrudan etkilemektedir. Önişleme aşaması birkaç adımı içermektedir. Çok Görevli Evrişimli Sinir Ağı (ÇESA) veya

RetinaFace gibi algoritmaları kullanan sağlam yüz tespiti, önemli bağlamsal bilgiyi korurken yüz bölgelerini izole eden hassas yüz çıkarma, poz varyasyonlarını normalize eden geometrik hizalama prosedürleri ve aydınlatma varyasyonlarının etkisini azaltan fotometrik normalizasyon bunların başlıcalarıdır. Yüz tespiti algoritmaları, yeterli çözünürlüğü koruyarak video karelerindeki yüz bölgelerini belirlemektedir. Hizalama prosedürleri, tespit performansını etkileyebilecek poz varyasyonları, ölçek farklılıkları ve rotasyon tutarsızlıklarını normalize etmektedir. Bu önişleme adımları, tespit modellerinin öğrenme kapasitelerini, gerçek manipülasyon yapı bozukluklarını tanımlamaya odaklamalarını sağlamak için gereklidir. Uygun önişleme, tespit performansını önemli ölçüde iyileştirebilirken, yetersiz önişleme ciddi şekilde bozabilmektedir. Diğer yandan veri artırma teknikleri, model genellemesini iyileştirmek ve eğitim veri setlerinin boyutunu genişletmek için yaygın şekilde kullanılmaktadır. Bu teknikler geometrik dönüşümleri, fotometrik manipülasyonları ve gürültü enjeksiyonunu içermektedir.

$$X_{aug} = T\left(X + \mathcal{N}(0, \sigma^2)\right) \tag{6}$$

Denklem (6), eğitim sırasında uygulanan veri artırma işlemleri için genel bir matematiksel ifadedir. Denklemde X orijinal girdi örneğini (tipik olarak bir yüz görüntüsü), T ise bir dönüşüm fonksiyonunu (geometrik işlemler, fotometrik ayarlamalar veya diğer değişiklikleri içerebilen) gösterir. $\mathcal{N}(0,\sigma^2)$ sıfır ortalama ve σ^2 varyans ile eklemeli Gauss gürültüsünü, X_{aug} sınıf üyeliğini belirleyen temel karakteristikleri koruyan artırılmış örneği temsil eder. Bu matematiksel formülasyon, eğitim sinyalinin bütünlüğünü tehlikeye atmadan modelin sağlamlığını ve genelleme yeteneğini iyileştirir.

1.6. Özellik Çıkarma ve Temsil Öğrenimi

Özellik çıkarma, derin sahte tespit sistemlerinin en kritik bileşenini oluşturmaktadır. Sınıflandırma kararları için kullanılan dahili temsillerin kalitesi, ayırt edici gücü ve genelleme yeteneği, nihai başarıyı belirlemektedir. Tüm tespit sisteminin etkinliği büyük ölçüde çeşitli veri setleri, manipülasyon teknikleri ve dağıtım senaryoları boyunca gerçek ve manipüle edilmiş yüz içeriği arasında güvenilir ayrım yapabilen anlamlı, sağlam ve ayırt edici özellikler çıkarma yeteneğine bağlıdır. Modern tespit sistemleri derin ESA'lar aracılığıyla öğrenilmiş özellik temsillerine dayanmaktadır. Bu ağlar özellik çıkarma ve

sınıflandırmayı aynı anda optimize eden uçtan uca eğitim prosedürleri aracılığıyla, gerçek ve manipüle edilmiş içerik arasında ayrım yapmak için ilgili hiyerarşik örüntüleri otomatik olarak keşfetmektedir. Bu otomatik özellik öğrenimi yaklaşımı hem sınıflandırma doğruluğu hem de genelleme yeteneği açısından, geleneksel el yapımı özellik yöntemlerine göre üstün olduğunu kanıtlamıştır. Öğrenilmiş özelliklerin üstünlüğü, manuel olarak tasarımı zor veya imkansız olan ince örüntüleri keşfetme ve farklı veri setleri ve manipülasyon yöntemlerinin özelliklerine otomatik olarak uyum sağlama yeteneğinden kaynaklanmaktadır.

Derin öğrenme yaklaşımları, sıkıştırma yapı bozukluklarını, karıştırma tutarsızlıklarını, zamansal düzensizlikleri, frekans alanı anomalilerini ve insan gözüyle fark edilemeyecek manipülasyon göstergelerini otomatik olarak öğrenebilmektedir. Hiyerarşik özellik çıkarma, manipülasyon tespiti için, uzamsal ve anlamsal ölçekte birden fazla örüntünün eş zamanlı bulunmasını sağlamaktadır. Alt ağ katmanları tipik olarak kenar örüntüleri, doku varyasyonları, renk tutarsızlıkları gibi temel görsel öğeleri yakalar. Orta katmanlar yüz işaretleri, ifade özellikleri ve bölgesel tutarsızlıklar gibi daha karmaşık örüntüleri öğrenir. Üst katmanlar anlamsal ilişkileri, kimlik özelliklerini ve global tutarlılık örüntülerini kodlayan soyut temsilleri öğrenir. Bu çok ölçekli özellik çıkarma yeteneği, manipülasyon yapı bozukluklarının kullanılan tekniğe bağlı olarak farklı uzamsal ölçeklerde ortaya çıkabileceği derin sahte tespit uygulamaları için özellikle değerlidir. Bazı yapı bozuklukları yüksek frekanslı gürültü kalıpları olarak görünürken, diğerleri yalnızca yüksek seviyeli anlamsal anlayış gerektiren ince ifade tutarsızlıkları yoluyla görünür hale gelebilmektedir. Bu karmaşık özellikleri etkili şekilde öğrenebilmek için, sinir ağı eğitimini stabilize etmek, yakınsamayı hızlandırmak ve genel optimizasyon dinamiklerini iyileştirmek amacıyla yığın normalizasyonu kullanılmaktadır.

$$BN(x) = \gamma \cdot \left((x - \mu_B) / \sqrt{\sigma_B^2 + \varepsilon} \right) + \beta \tag{7}$$

Denklem (7)'de, $BN(x)$ girdi x için yığın-normalize edilmiş çıktı , μ_B ve σ_B^2 sırasıyla mevcut mini-yığın üzerinde hesaplanan ortalama ve varyans değerleridir. γ ve β normalize edilmiş değerlerin ölçeklenmesi ve kaydırılmasını sağlayan öğrenilebilir parametreleri ve ε yığın varyansı çok küçük olduğunda sıfıra bölmeyi önlemek için eklenen küçük bir sabiti temsil eder. Bu teknik, neredeyse tüm modern sinir ağı mimarilerinde standart ve temel bir bileşen haline gelmiştir.

1.7. Performans Değerlendirme Metrikleri ve İstatistiksel Analiz

Derin sahte tespit sistemlerinin kapsamlı bir şekilde değerlendirilmesi için, çalışma koşulları boyunca sistem performansının farklı yönlerini yakalayan birden fazla tamamlayıcı metriğin kullanılması gerekmektedir. Geleneksel doğruluk metrikleri sistem için temel ölçüm sağlarken, dengesiz veri setleri olan senaryolarda veya farklı hata türleriyle ilişkili değişen maliyetlerde sistem davranışını yeterince yansıtmayabilir. Bu nedenle, modern değerlendirme çerçeveleri kesinlik, duyarlılık, F1-skoru, Eğri Altında Kalan Alan - Alıcı İşletim Karakteristiği (Area Under the Curve - Receiver Operating Characteristic, AUC-ROC) gibi birçok metriği birlikte kullanmaktadır.

Kesinlik ve duyarlılık metrikleri, gerçek dünya dağıtım senaryolarında tespit sisteminin davranışını anlamak için önemli bilgiler sağlamaktadır (Powers, 2020; Sokolova & Lapalme, 2009; Davis & Goadrich, 2006).

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

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

Denklem (8) ve (9) sırasıyla kesinlik ve duyarlılık metriklerini, karışıklık matrisi öğeleri cinsinden tanımlar; burada TP gerçek pozitif tahminlerin sayısını (doğru tanımlanmış manipüle edilmiş örnekler), FP yanlış pozitif tahminlerin sayısını (manipüle edilmiş olarak yanlış sınıflandırılan gerçek örnekler) ve FN yanlış negatif tahminlerin sayısını (gerçek olarak yanlış sınıflandırılan manipüle edilmiş örnekler) temsil eder. Bu metrikler sistem davranışının farklı yönleri hakkında bilgiler sağlar ve farklı çalışma koşulları boyunca sistem davranışını anlamak için kritik olan F1-skoru metriği için matematiksel temel oluşturur. F1-skoru (Denklem (10)), kesinlik ve duyarlılık metriklerinin harmonik ortalamasını temsil etmektedir. Her iki metriği eşit değerlendiren ve herhangi birine önyargılı olmayan dengeli bir ölçüt sunmaktadır.

$$F1\_skoru = 2 \cdot Kesinlik \cdot \frac{Duyarlılık}{Kesinlik + Duyarlılık} \tag{10}$$

F1-skoru, özellikle doğruluğun çoğunluk sınıfı üzerindeki doğru tahminler nedeniyle yanıltıcı derecede yüksek olabileceği dengesiz veri setleri olan senaryolarda, sistem

performansının daha dengeli değerlendirilmesi sağlar. Bu metrik, farklı sınıf dağılımlarına sahip veri setleri boyunca tespit sistemlerini değerlendirirken ve karşılaştırırken veya değişen kesinlik-duyarlılık dengeleri sergileyebilen birden fazla yaklaşımı değerlendirirken özellikle değerlidir.

AUC-ROC, tüm olası karar eşikleri boyunca sistem performansının kapsamlı şekilde değerlendirilmesini sağlayan bir metriktir. Bu metrik, sistemin eşik seçimine bakılmaksızın gerçek ve manipüle edilmiş içerik arasında ayrım yapabilme yeteneğini değerlendirdiği için, derin sahte tespit sistemleri açısından özellikle değerlidir.

İstatistiksel anlamlılık testi, gözlemlenen performans farklılıklarının gerçek iyileştirmeler yerine rastgele varyasyon olmamasını sağlamada önemli rol oynamaktadır. Deneysel sonuçların güvenilirliğini ve tekrarlanabilirliğini doğrulamak için gereklidir. Geleneksel olarak, bağımsız örneklemler için kullanılan *t*-testi Denklem (11)'de verilmiştir.

$$t = \frac{\mu^1 - \mu^2}{\sqrt{\frac{s^{12}}{n^1} + \frac{s^{22}}{n^2}}} \tag{11}$$

Denklemde, μ_1 ve μ_2 karşılaştırılan iki sistem için örnek ortalamalarını, s_1^2 ve s_2^2 örnek varyanslarını ve n_1 ve n_2 değerlendirme için kullanılan ilgili örnek boyutlarını temsil eder. Aynı test veri kümesi üzerinde değerlendirilen modeller karşılaştırıldığında ise, bağımsız örneklem *t*-testi yerine eşleştirilmiş istatistiksel testler kullanılır:

- McNemar Testi: İki model arasındaki sınıflandırma doğruluğu farklılıklarının istatistiksel anlamlılığını değerlendirmek için kullanılır. Bu test, aynı test örnekleri üzerinde iki sınıflandırıcının hata oranlarını karşılaştırır.
- DeLong Testi: İki model arasındaki AUC-ROC farklılıklarının istatistiksel anlamlılığını test etmek için kullanılır. Bu test, aynı test veri kümesi üzerinde hesaplanan iki ROC eğrisini karşılaştırır.
- Bootstrap Güven Aralıkları: Tüm performans metrikleri (doğruluk, kesinlik, duyarlılık, F1-skoru) için %95 güven aralıkları, bootstrap yeniden örnekleme yöntemi (10,000 iterasyon) ile hesaplanır. Bu, her metriğin belirsizlik aralığını ve istatistiksel güvenilirliğini gösterir.

İstatistiksel anlamlılık seviyesi $\alpha = 0.05$ olarak belirlenmiş olup, $p < 0.05$ değeri anlamlı farklılıkları göstermektedir.

1.8. Topluluk Öğrenimi Metodolojileri ve Birleştirme Stratejileri

Derin sahte tespit sistemlerinin sağlamlığını, doğruluğunu ve genelleme yeteneğini iyileştirmek için sıklıkla kullanılan yöntemlerden biri topluluk öğrenimidir. Yöntem, tamamlayıcı güce sahip birden fazla bireysel modelin ürettiği tahminlerin, sistematik birleşimi yoluyla bu iyileştirmeyi sağlamaktadır. Topluluk öğrenimi yöntemlerinin teorik temeli istatistiksel öğrenme teorisine dayanmaktadır. Teori, çeşitli tahmin edicileri birleştirmenin, tahmin hatasının hem önyargı hem de varyans bileşenlerini aynı anda azaltabileceği ilkesine dayanmaktadır. Bu yaklaşım, bireysel modellerin çeşitli manipülasyon tekniklerine veya farklı veri setlerine uygulandığında farklı güçlü ve zayıf yönler sergileyebileceği derin sahte tespit uygulamalarında özellikle etkilidir.

Derin sahte tespiti için çeşitli topluluk öğrenme stratejileri uygulanmıştır. Sert oylama yöntemi, bireysel model hatalarına karşı sayıca üstünlüğe göre karar vermektedir. Yumuşak oylama yaklaşımı, birden fazla modelin olasılık tahminlerini birleştirerek daha nüanslı karar vermeyi sağlamaktadır.

$$P_{ensemble} = \left(\frac{1}{M}\right) \Sigma \, P_i \tag{12}$$

Yumuşak oylama topluluk tahmininin matematiksel formülasyonu Denklem (12)'de verilmiştir. Denklemde $P_{ensemble}$ topluluk sistemi tarafından üretilen nihai olasılık tahminini, M topluluktaki model sayısını ve P_i sınıflandırılan örnek için i. model tarafından üretilen olasılık tahminini temsil eder. Bu ortalama alma yaklaşımı, bireysel model hatalarının etkisini azaltırken birden fazla modelin bilgi ve uzmanlığını etkili bir şekilde birleştirir ve böylece daha sağlam ve güvenilir tahminler sağlar.

Ağırlıklı topluluk yöntemleri, bireysel performans özelliklerine dayalı olarak topluluktaki modellere farklı önem ağırlıkları atayarak temel ortalama alma yaklaşımının daha gelişmiş halini temsil etmektedir.

$$P_{weighted} = \sum_{i=1}^{M} w_i P_i \, , where \sum_{i=1}^{M} w_i = 1 \tag{13}$$

Ağırlıklı topluluk tahmini Denklem (13)'te verilmiştir. Denklemde w_i tahmin kalitesine veya uygunluğuna göre i. modele atanan ağırlığı, P_i sınıflandırılan örnek için i. model tarafından üretilen olasılık tahminini, M topluluktaki model sayısını temsil eder.

Model ağırlıklarının optimal seçimi, doğrulama seti performans analizi, çapraz doğrulama yaklaşımları, meta-öğrenme teknikleri veya ağırlık tahminlerindeki belirsizliği hesaba katan Bayesian model ortalaması gibi daha gelişmiş yöntemler dahil olmak üzere çeşitli optimizasyon yöntemleri aracılığıyla belirlenebilir.

Test Zamanı Artırımı (TZA), çıkarım sırasında test örneklerine çeşitli artırımlar uygulayarak, modelin sağlamlığını ve tahmin doğruluğunu geliştiren bir tekniktir. Bu yaklaşım, eğitim sırasında kullanılan veri artırma tekniklerinden yararlanır ancak çıkarım zamanında uygular. Her test örneğinin birden fazla dönüştürülmüş versiyonunu üretir, böylece bireysel örnek seviyesinde topluluk öğrenimi yaklaşımını etkili şekilde uygular. TZA'nın matematiksel temeli, girdi örneklerindeki küçük, gerçekçi bozulmaların, sağlam modeller için sınıflandırma kararını önemli ölçüde etkilememesi gerektiği varsayımına dayanmaktadır (Denklem 14).

$$P_{TTA} = \left(\frac{1}{K}\right)\sum_{k=1}^{K} f\left(T_{k(x)}\right) \tag{14}$$

Denklem (14)'te P_{TTA} test zamanı artırımı yoluyla elde edilen nihai tahmini, K girdi örneğine uygulanan farklı artırım dönüşümlerinin sayısını, f girdileri olasılık tahminlerine eşleyen eğitilmiş model fonksiyonunu, T_k orijinal girdiye uygulanan k. artırım dönüşümünü ve x sınıflandırılan orijinal test örneğini temsil eder. Denklemden görüldüğü gibi, TZA yönteminde, girdinin birden fazla dönüştürülmüş versiyonu için üretilen tahminlerin ortalamasını alınmaktadır. Buna göre, girdinin farklı varyasyonları için modelin tutarlı tahminler üretmesi, daha güvenilir ve sağlam olduğunu göstermektedir.

1.9. Mimari Yenilikler, Tasarım İlkeleri ve Optimizasyon Stratejileri

Derin sahte tespiti için geliştirilen sinir ağı mimarileri, problemin özel gereksinimlerine göre gelişimini sürdürmektedir. Son mimari yenilikler özellik çıkarma yeteneklerini iyileştirmeye, dikkat mekanizmalarını güçlendirmeye ve daha verimli hesaplama yapıları geliştirmeye odaklanmaktadır. ConvNeXt mimarisi, geleneksel evrişimli işlemlerin hesaplama verimliliğini korurken, GD'nin tasarım ilkelerini içeren önemli bir ilerlemeyi temsil etmektedir. Bu mimari, modern mimari ilkeler ve eğitim metodolojileriyle tasarlandığında evrişimli ağların dönüştürücü tabanlı yaklaşımlara kıyasla rekabetçi performans elde edebileceğini göstermektedir. ConvNeXt derinlik bazlı evrişimler, katman

normalizasyonu, GELU etkinleştirme fonksiyonları ve ters darboğaz yapıları dahil birkaç temel yenilik kullanmaktadır. Dikkat mekanizmaları, hesaplama kaynaklarını girdi görüntülerinin en ilgili bölgelerine odaklama konusundaki yetenekleri nedeniyle, modern bilgisayarlı görü uygulamalarında önemli hale gelmiştir.

Derin sahte tespiti için derin sinir ağlarının optimizasyonu, çeşitli veri setleri boyunca genelleme karakteristiklerini korurken, tahmin hatalarını minimize etmeye çalışan gelişmiş matematiksel prosedürler içermektedir. Modern optimizasyon algoritmaları, momentum, adaptif öğrenme oranları ve stokastik gradyan inişinin gelişmiş varyantlarını kullanmaktadır. Bu algoritmalar karmaşık, yüksek boyutlu optimizasyon uzaylarında gezinirken ağ parametrelerini yinelemeli olarak güncellemektedir. Adam optimizasyon algoritması, momentum tabanlı yöntemlerin faydalarını adaptif öğrenme oranı mekanizmalarıyla birleştiren en yaygın yaklaşımlardan biridir.

$$lr(t) = lr^0 \times (0.5)^{\left\lfloor \frac{t}{T} \right\rfloor} \tag{15}$$

Eğitim sırasında düzenli aralıklarla öğrenme oranının azaltılması Denklem (15)'te görülmektedir. Denklemde $lr(t)$ eğitim adımı t'deki öğrenme oranını, lr^o başlangıçtaki öğrenme oranı değerini, T öğrenme oranı azalmaları arasındaki adım boyutunu ve $\lfloor \rfloor$ en yakın tam sayıya yuvarlama fonksiyonunu ifade eder. Bu yaklaşım, parametre güncellemelerinin adım boyutunu azaltarak eğitim ilerledikçe parametrelerin ince ayarını sağlar ve optimizasyon algoritmasının hem eğitim hem de doğrulama veri setlerinde daha iyi yakınsamasına olanak tanır.

Öğrenme oranı, tüm eğitim süreci boyunca parametre güncellemelerinin adım boyutunu kontrol ederek, optimal eğitim sonuçları ve nihai model performansı elde etmede kritik bir rol oynamaktadır. Önceden belirlenmiş aralıklarla öğrenme oranını azaltan adım azalma yaklaşımları, öğrenme oranını üstel fonksiyonlara göre sürekli azaltan üstel azalma yöntemleri, kosinüs şeklindeki azalma kalıplarını izleyen kosinüs tavlama yaklaşımları ve öğrenme oranlarını doğrulama performansı veya diğer eğitim metriklerine dayalı olarak ayarlayan adaptif yöntemler dahil olmak üzere çeşitli stratejiler geliştirilmiştir.

Düzenlileştirme yöntemleri, aşırı öğrenmeyi önlemede ve derin sahte tespit modellerinin genelleme yeteneğini iyileştirmede kritik rol oynamaktadır. Bu teknikler modelin daha genelleştirilebilir kalıplar öğrenmesini teşvik etmektedir.

1.10. Zorluklar, Sınırlılıklar ve Gelecek Yönleri

Derin sahte tespit alanı, üretici modeller geliştikçe yeni zorluklarla karşılaşmaya devam etmektedir. Üretim-tespit dinamiğinin çekişmeli doğası, sürekli ve yükselen bir yarış yaratmaktadır. Bu dinamik evrimsel süreç, değişen tehdit manzaraları, yeni manipülasyon teknikleri ve gelişen üretici teknolojileri karşısında tespit yaklaşımlarının uzun vadeli sürdürülebilirliği konusunda endişeler yaratmaktadır. Dolayısı ile değişen koşullarda başarımını koruyabilen adaptif sistemlere ihtiyaç duyulmaktadır. Farklı manipülasyon teknikleri ve değişen karakteristiklere sahip veri setleri üzerinde genelleme yeteneği, en önemli zorluklardan birini temsil etmektedir. Belirli manipülasyon yöntemleri veya veri setleri üzerinde eğitilen modeller, daha önce görülmemiş tekniklere veya farklı veri dağılımlarına uygulandığında zayıf performans sergileyebilmektedir. Diğer yandan hesaplama verimliliği hususları, sınırlı hesaplama kaynakları ve güç tüketimi kısıtlamaları olan gerçek dünya uygulamalarında giderek önemli hale gelmektedir.

Derin sahte tespit araştırmalarının geleceğinin, mevcut sınırlılıkları ele alırken daha sağlam, doğru ve pratik tespit sistemleri için yeni imkanlar sunan gelişmelerle şekillendirilme olasılığı yüksektir. Ses karakteristikleri, görsel içerik, zamansal dinamikler ve meta veriler gibi birden fazla bilgi modalitesinden yararlanan çapraz-modal tespit yaklaşımları, tespit sağlamlığını ve güvenilirliğini iyileştirmek için umut verici bir yönü temsil etmektedir (Li et al., 2021; Khalid et al., 2021; Cozzolino et al., 2021). Federasyon öğrenimi yaklaşımları, şu anda derin sahte tespit araştırmalarını sınırlayan gizlilik ve veri paylaşım zorluklarına potansiyel çözümler sunmaktadır. Merkezi veri paylaşımı gerektirmeyen işbirlikçi model eğitimi sayesinde, gizliliği korurken daha sağlam tespit sistemlerinin geliştirilmesi mümkün olacaktır.

Çekişmeli eğitim teknikleri, eğitim süreci sırasında çekişmeli örnekleri ve ileri saldırı senaryolarını birleştiren başka bir umut verici araştırma yönünü temsil etmektedir. Tespit modellerini çekişmeli örneklere sürekli maruz bırakmak, kasıtlı saldırılara karşı sağlamlığı iyileştirebilmektedir.

Standartlaştırılmış değerlendirme protokolleri, kapsamlı kıyaslama veri setleri ve sağlam metodoloji çerçevelerinin geliştirilmesi, alan için önemli ve devam eden bir zorluk olmaya devam etmektedir. Tutarlı değerlendirme metodolojileri, farklı yaklaşımlar arasında adil karşılaştırmalar sağlamak ve zaman içinde ilerlemeyi takip etmek için gereklidir.

1.11. Araştırma Hedefleri ve Katkıları

Bu araştırma, hesaplama verimliliğini korurken iyi bir performans elde eden kapsamlı, sağlam ve pratik uygulanabilir bir derin sahte tespit sistemi geliştirmeyi amaçlamaktadır. Birincil hedef, kabul edilebilir doğruluk ve güvenilirlikle yüz görüntülerindeki ince manipülasyon yapı bozukluklarını tanımlayan yenilikçi topluluk tabanlı tespit çerçevesi tasarlamak, uygulamak ve sistematik olarak değerlendirmektir.

Bu çalışmanın temel katkısı, yenilikçi tasarım ilkeleri ve kapsamlı deneysel çalışmalar aracılığıyla, model karmaşıklığı ve tespit performansı arasındaki dengeyi optimize eden yeni mimari konfigürasyonların geliştirilmesidir.

Bu araştırmada geliştirilen topluluk metodolojileri, birden fazla veri seti ve değerlendirme protokolü boyunca bireysel model yaklaşımlarına kıyasla tespit doğruluğunda önemli ve tutarlı iyileştirmeler göstermektedir. TZA tekniklerinin entegrasyonu, sistem sağlamlığını ve doğruluğunu daha da artırmaktadır.

Araştırma, derin sahte tespit uygulamalarının özel bağlamında model mimarisi, eğitim metodolojisi, optimizasyon stratejileri ve tespit performansı arasındaki karmaşık ilişkileri anlamaya yönelik önemli teorik ve pratik katkılar sağlamaktadır. Kapsamlı deneysel çalışmaları, sistematik parametre analizi, istatistiksel anlamlılık testi ve detaylı hata analizi aracılığıyla, tespit doğruluğunu ve genelleme yeteneğini etkileyen faktörler hakkında detaylı bilgiler sunmaktadır.

2. YAPILAN ÇALIŞMALAR

2.1. Giriş

Bu tez çalışmasında sahte yüz tespiti için derin öğrenme modellerinin geliştirilmesinde kullanılan yaklaşım; veri kümesi hazırlığı, çeşitli evrişimli sinir ağı (ESA) mimarilerinin uygulanması, dikkat mekanizması entegrasyonu, topluluk öğrenimi stratejileri ve performans optimizasyonu tekniklerini kapsamaktadır. Çalışmadaki derin öğrenme modelleri, gelişmiş üretici modeller tarafından oluşturulan ve yalnızca görsel incelemeyle gerçek insan yüzlerinden ayırt edilmesi zorlaşan sentetik yüzlerin tespit edilmesi probleminin daha az kaynak ile çözülmesi amacıyla tasarlanmıştır.

Araştırma metodolojisi, veri kümesi oluşturma ve ön işleme ile başlayan ve ardından birden fazla güncel derin öğrenme mimarisinin uygulanması ve değerlendirilmesiyle devam eden sistematik bir yaklaşım izlemektedir. Çalışmanın başlangıcında önceden eğitilmiş modellerle transfer öğrenmesi yaklaşımları incelenmiş, ancak sıfırdan eğitim yöntemi ile sahte yüz tespiti için göreve özgü öznitelik öğreniminin etkinliğini gösteren sonuçlar elde edilmiştir. Ayrıca, model performansını sahte yüz tespiti görevi için optimize etmek amacıyla, özel aşama konfigürasyonları ve dikkat mekanizması entegrasyonları dahil olmak üzere çeşitli mimariler geliştirilmiştir. Deneysel çerçeve birden fazla modelin güçlü yönlerini birleştiren topluluk öğrenimi yönteminin uygulanmasının yanı sıra, TZA ve veri kümesi iyileştirme gibi gelişmiş optimizasyon tekniklerini de içermektedir.

2.2. Veri Kümesi Oluşturma

2.2.1. Veri Kümesinin Yapısı

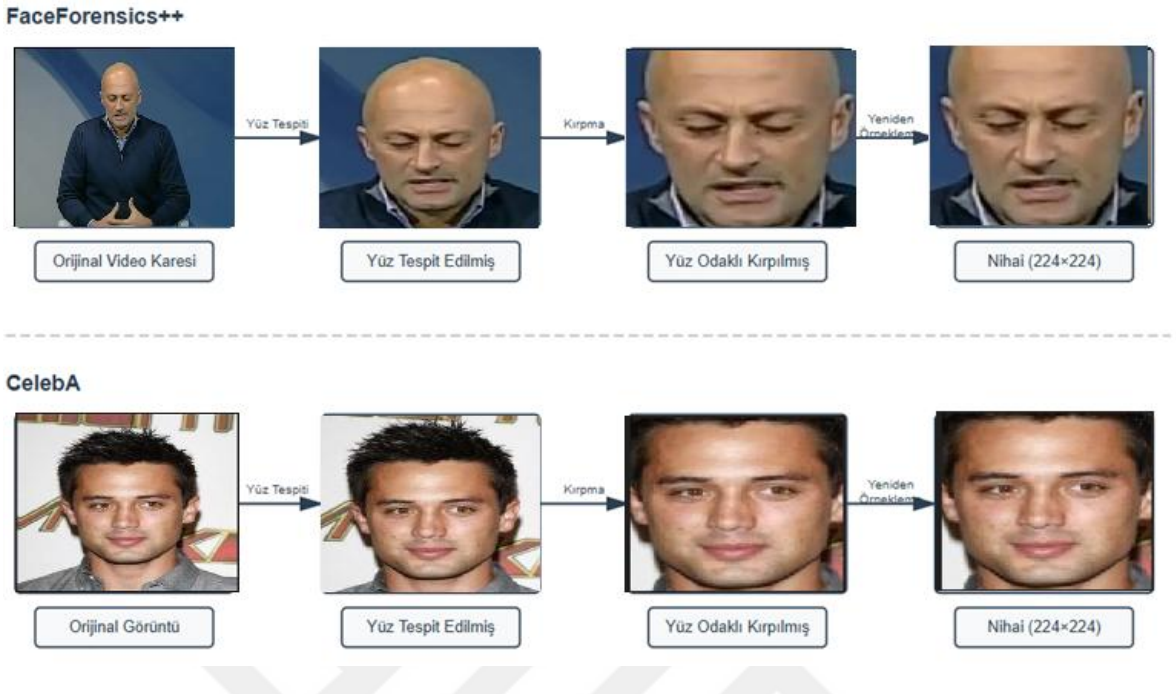
Makine öğrenimi uygulamalarının başarımları, eğitim ve değerlendirme için kullanılan veri kümesinin kalitesi ve yapısıyla doğrudan ilişkilidir. Bu çalışmada hem gerçek hem de yapay yüz görüntülerinden oluşan bir veri kümesi oluşturulmuştur. Veri kümesi, eğitim ile test aşamaları arasında veri sızıntısını önlemek için üç farklı alt küme halinde organize edilmiştir. Eğitim seti, 2.400 gerçek insan yüzü ve 2.468 sentetik olarak üretilmiş yüz olmak üzere 4.868 görüntüden oluşmaktadır. Sahte yüzler gerçek dünya senaryolarında karşılaşılan

üretme tekniklerinin çeşitliliğini yansıtmaktadır. Doğrulama seti, 300 gerçek ve 300 sahte görüntü ile eğitim süreci boyunca tarafsız bir değerlendirme metriği sağlamaktadır. Benzer şekilde, test seti de 300 gerçek ve 300 sentetik yüzün eşit dağılımını içermektedir. Eğitim seti, derin sinir ağlarının anlamlı öznitelik temsilleri öğrenmesi için yeterli veri sağlarken, dengeli doğrulama ve test setleri, model performans metriklerinin her iki sınıf arasında eşit ayrım yapma yeteneğini doğru bir şekilde yansıtmasını sağlamaktadır.

2.2.2. Veri Ön İşleme ve Artırma

Ön işleme süreci, ham video verilerini ve çeşitli görüntü kaynaklarını derin öğrenme modellerinin girişine uygun standartlaştırılmış bir formata dönüştürmek için kullanılmaktadır. FaceForensics++ (Rossler et al., 2019) veri seti, hem gerçek videoları hem de çeşitli derin sahte üretme teknikleri kullanılarak manipüle edilmiş yüz videolarını içermektedir. Hem gerçek hem de sahte yüz görüntülerinden oluşan bir veri seti oluşturmak için bu videolardan kareler çıkarılmıştır. Çıkarılan her kare içindeki yüz bölgelerini belirlemek için yüz tespit algoritmaları kullanılmıştır. Daha sonra, arka plan unsurlarını ortadan kaldırmak ve tüm örnekler boyunca tutarlı bir yapı oluşturmak için yüz odaklı kırpma uygulanmıştır.

Veri kümesi çeşitliliğini artırmak ve modelin genelleme yeteneğini iyileştirmek için veri kümesine CelebA (Liu et al., 2015) veri kümesinden gerçek yüz görüntüleri dahil edilmiştir. Bu ek görüntüler, FaceForensics++'tan türetilen örneklere uygulanan ön işleme sürecinden geçirilmiştir. Bu çift kaynaklı yaklaşım, modelin eğitim sırasında daha çeşitli yüz özellikleri, aydınlatma koşulları ve demografik temsillerle karşılaşmasını sağlamaktadır. FaceForensics++'tan türetilen örnekler ve CelebA görüntülerinin kombinasyonu, tüm deneyler boyunca kullanılan nihai kapsamlı veri kümesini oluşturmuş, sağlam model geliştirme için gerekli olan hem sentetik manipülasyon tekniklerinin hem de gerçek yüz özelliklerinin çeşitliliğini sağlamıştır. Veri setinden örnek görüntüler ve ön işleme süreci Şekil 1’ de görülmektedir.

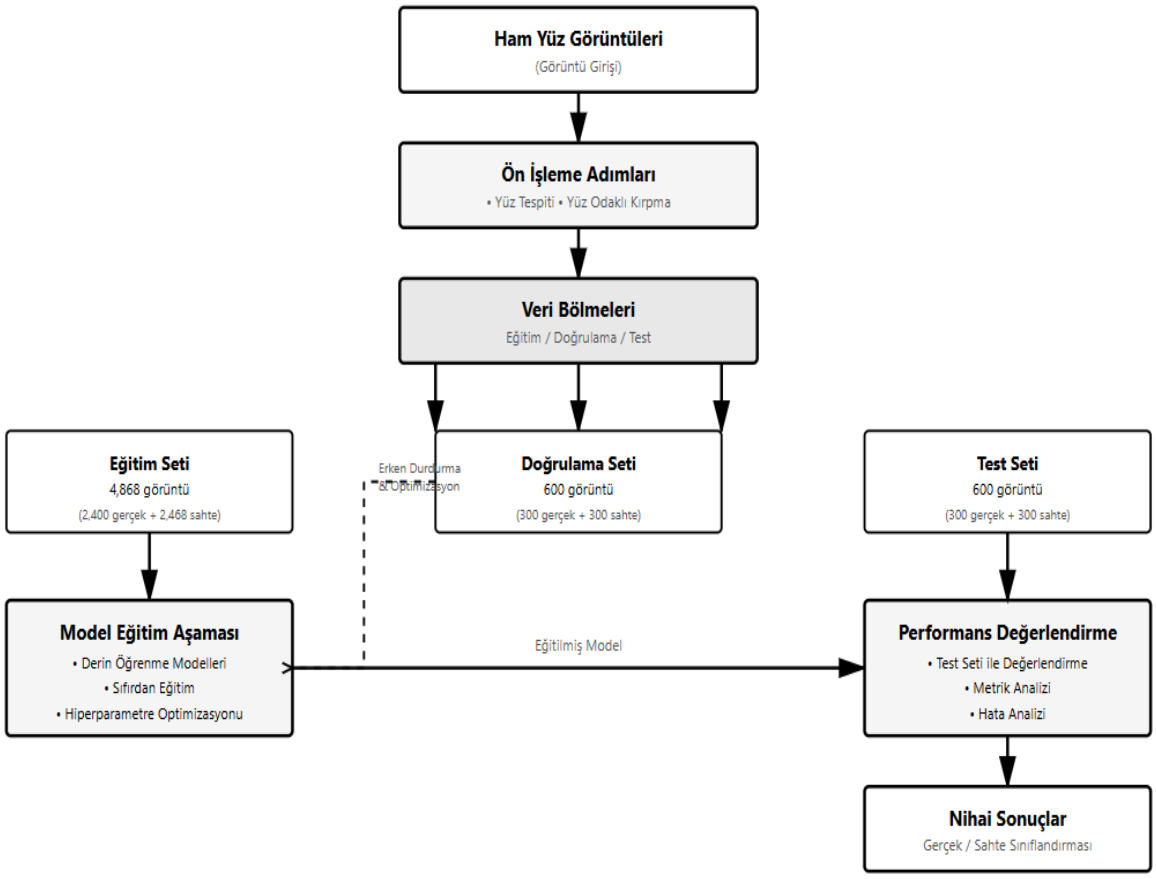


Şekil 1. Ön işleme adımları

Normalizasyon, kırmızı, yeşil ve mavi kanallar için sırasıyla 0.485, 0.456 ve 0.406 ortalama değerler ve 0.229, 0.224 ve 0.225 standart sapmalar ile ImageNet istatistikleri kullanılarak piksel değerlerinin standartlaştırıldığı ön işleme aşamasındaki önemli bir adımı oluşturmaktadır. Bu normalizasyon yaklaşımı, önceden eğitilmiş modellerle uyumluluk sağlamakta ve giriş dağılımlarını beklenen aralıklarda tutarak kararlı bir eğitimi kolaylaştırmaktadır. Veri artırım stratejileri, model genellemesini ve sağlamlığını iyileştirmede çok önemli bir rol oynamaktadır. Artırım işlemi, eğitim sırasında olasılıksal olarak uygulanan, modellerin yönelimden bağımsız özellikleri öğrenmesine yardımcı olan rastgele yatay çevirmeler, hafif poz varyasyonlarına karşı sağlamlık sağlayan döndürme dönüşümleri ve farklı aydınlatma koşullarını simüle eden parlaklık ve kontrast ayarlamaları gibi fotometrik artırımlar dahil olmak üzere çeşitli dönüşümleri içermektedir. Bu artırım teknikleri, sahte yüz tespiti için özellikle önemlidir. Çünkü sentetik görüntüler, ayırt edici özellikler olarak yanlışlıkla öğrenilebilecek ince yapı bozuklukları sergileyebilir.

2.3. Sistem Mimarisine Genel Bakış

Çalışmada tasarlanan sistemin genel mimarisi Şekil 2' de verilmiştir. Mimari ham yüz görüntülerinin sisteme girdiği ve yukarıda açıklanan ön işleme adımlarından geçtiği görüntü girişi ile başlamaktadır. Ön işlenmiş görüntüler daha sonra eğitim, doğrulama veya test kümelerine ayrılmaktadır. Model eğitim aşaması, çeşitli derin öğrenme modellerinin hazırlanan eğitim verilerini kullanarak gerçek ve sahte yüzler arasında ayrım yapmayı öğrendiği sistem mimarisinin çekirdeğini temsil etmektedir. Bu aşamada, başlangıçta önceden eğitilmiş modellerden aktarım öğrenmesi stratejileri kullanılmış, ancak sonraki deneyler, sıfırdan eğitimin sahte yüz tespiti görevi için daha başarılı sonuçlar sağladığını ortaya koymuştur. Eğitim süreci, aşırı öğrenmeyi önlemek ve en uygun genelleme performansını elde etmek için, erken durdurma ve hiper-parametre optimizasyonunu sağlayan doğrulama seti performansıyla yönlendirilmektedir.



Şekil 2. Sistemin genel yapısı

Sistem mimarisinin son bileşeni, eğitilmiş modellerin ayrılmış test seti kullanılarak değerlendirildiği performans değerlendirme modülüdür. Bu değerlendirme aşaması detaylı hata analizi ve istatistiksel anlamlılık testi dahil olmak üzere, birden fazla metrik ve analiz tekniği içermektedir. Çalışmada tüm modellerin aynı koşullar altında değerlendirilmesi sağlanarak farklı yaklaşımlar arasında anlamlı karşılaştırmalar yapılmış ve çeşitli metodolojik seçimlerin göreceli etkinliğine dair güvenilir çıkarımlar sunulmuştur.

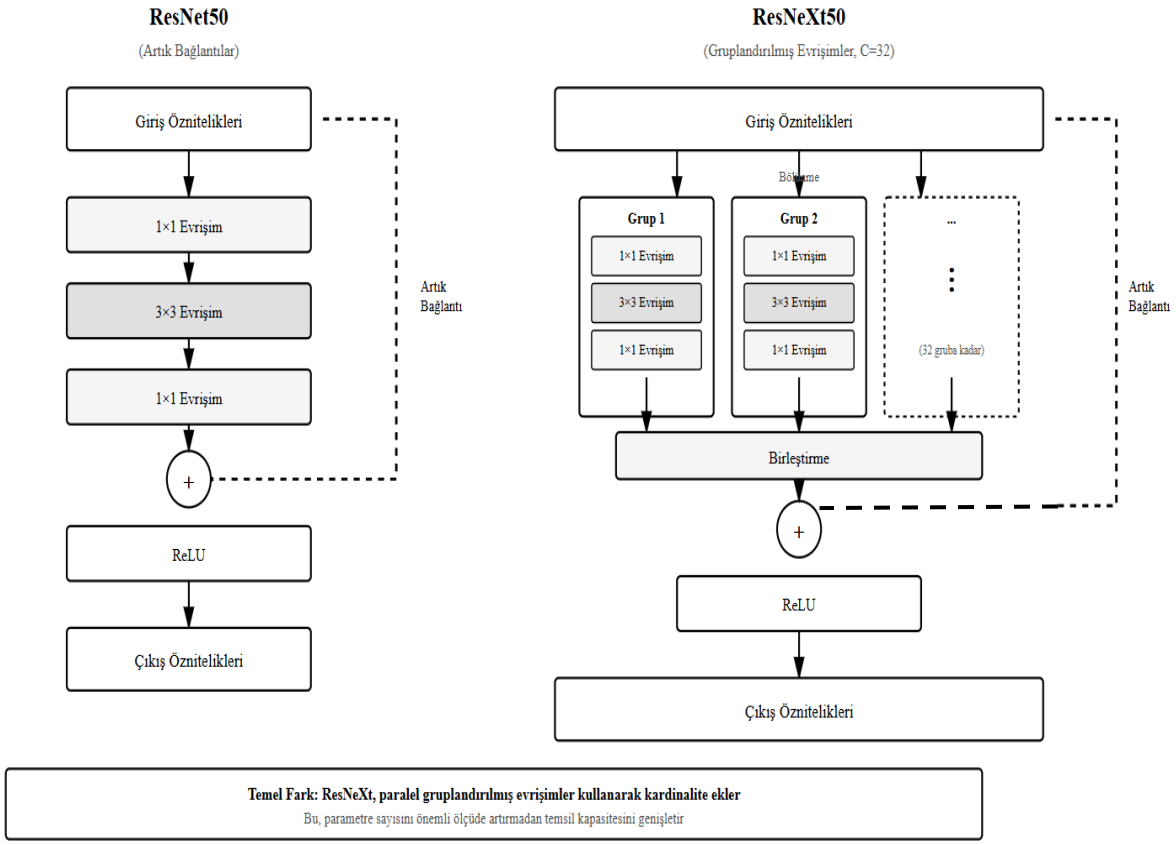
2.4. Sahte Yüz Tespiti Modelinin Geliştirilmesi

Çalışmada nihai sahte yüz tespiti model mimarisinin geliştirilmesi, temel mimarilerle başlayan ve sahte yüz tespiti görevine ilişkin performans analizi ve deneysel çıkarımlara göre aşamalı modifikasyonları içeren bir süreç izlemiştir. ResNet50 ve ResNeXt50 mimarileriyle yapılan ilk temel deneyler, performans kıyaslamalarını belirlemiş ancak parametre verimliliği ve göreve özgü öznitelik öğreniminde sınırlamalar ortaya çıkarmıştır. Bu bulgular, daha modern mimarilerin kullanımını motive etmiş ve ConvNeXt varyantlarının birincil mimari aile olarak araştırılmasına yol açmıştır. Daha sonra ConvNeXt Aşama-3'teki katman sayısının değiştirilmesi yoluyla oluşturulan konfigürasyonların ([3,3,3,3]'ten [3,3,18,3]'e kadar) sahte yüz tespitindeki başarımı incelenmiştir. Buradaki temel hedef modellerin hesaplama karmaşıklığını azaltmaktır. Sonraki aşamada indirgenmiş modellere dikkat mekanizmasının entegrasyonu ile performanstaki değişim incelenmiştir. Çalışmada ayrıca ConvNeXt modeline ek olarak güncel FocalNet mimarisinin de sahte yüz tespitindeki başarımı araştırılmıştır. Model mimarisinin geliştirilmesi sürecindeki bu aşamalar alt bölümlerde detaylı olarak açıklanmıştır.

Ayrıca model geliştirme sürecinde, tespit doğruluğu ile hesaplama verimliliği arasındaki dengenin sağlanmasına özel önem verilmiştir. Bu doğrultuda, modelin genelleme kabiliyetini artırmak amacıyla farklı optimizasyon stratejileri, öğrenme oranı planları ve veri artırma teknikleriyle kapsamlı deneyler yürütülmüştür. Aydınlatma, poz ve çözünürlük değişimlerine karşı dayanıklılığı sağlamak amacıyla yüz hizalama, normalizasyon ve renk uzayı dönüşümleri gibi önişleme adımlarının etkisi de analiz edilmiştir. Bu tamamlayıcı incelemeler, yalnızca eğitim hattının iyileştirilmesine katkı sağlamamış, aynı zamanda sahte yüz tespiti modellerinin veri kümesi özelliklerine ve mimari tasarım tercihleri karşısındaki duyarlılıklarına ilişkin önemli çıkarımlar sunmuştur.

2.4.1. ResNet50 ve ResNeXt50 Modelleri

Çalışmadaki ilk deneylerde, temel performans kıyaslamalarını oluşturmak ve sonraki mimari kararları belirlemek için, yerleşik derin öğrenme mimarileri değerlendirilmiştir. ResNet50 ve ResNeXt50, bilgisayarlı görü görevlerindeki kanıtlanmış etkinliklerinden dolayı temel modeller olarak uygulanmıştır. ResNet50, derin öznitelik öğrenimini sağlamak için artık bağlantılar kullanırken, ResNeXt50, kardinalite yoluyla ek temsilkapasitesi sağlayan gruplandırılmış evrişimler aracılığıyla bu yaklaşımı genişletmektedir (Şekil 3).



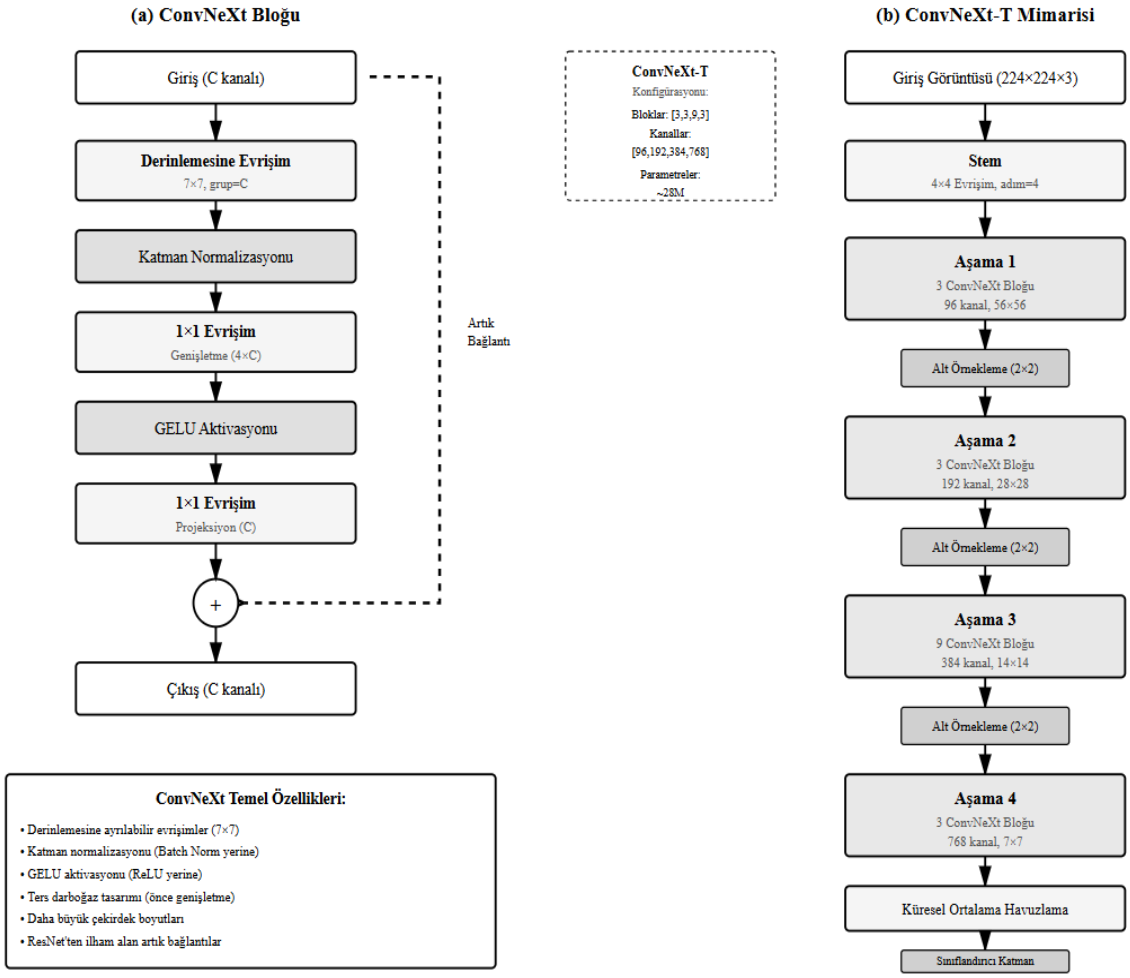
Şekil 3. ResNet ve ResNneXt mimarileri arasındaki farklar

Her iki mimari de son katmanlarının değiştirilmesi yoluyla ikili sınıflandırma için uyarlanmış ve ImageNet önceden eğitilmiş ağırlıklarından aktarım öğrenimi kullanılarak ince ayar yapılmıştır. Bu modeller sahte yüz tespitinde makul bir performans göstermiş olsa da, elde edilen sonuçlar, genel amaçlı mimarilerin, aktarım öğrenimiyle bile, sahte yüz tespiti için parametre verimliliği ve göreve özgü öznitelik ayrımında sınırlamalara sahip olduğunu göstermiştir.

2.4.2. ConvNeXt Mimari Ailesi

ConvNeXt modeli, evrişimli işlemlerin hesaplama verimliliğini korurken, dönüştürücü mimarisinin avantajlarını birleştiren bir yaklaşımı temsil etmektedir. Mimari, evrişimli ağlar ve dönüştürücülerin tasarım ilkelerinin birleştirerek standart ResNet tasarımını sistematik olarak geliştirmekte, bu da gelişmiş performans ve eğitim kararlılığı ile sonuçlanmaktadır.

ConvNeXt mimarisi, birkaç temel tasarım ilkesini kullanmaktadır. Derinlemesine ayrılabilir evrişimlerin kullanılması, temsil kapasitesini korurken hesaplama karmaşıklığını azaltmakta ve toplu normalizasyon yerine katman normalizasyonunun dahil edilmesi, daha kararlı eğitim süreci sağlamaktadır. GELU aktivasyon fonksiyonunun ReLU yerine seçilmesi, daha büyük çekirdek boyutlarının kullanılmasıyla birlikte, ağın giriş görüntülerindeki daha karmaşık uzamsal ilişkileri yakalamasını sağlamaktadır. ConvNeXt mimarisinin Mini (Tiny-ConvNeXt-T), Küçük (Small-ConvNeXt-S), Temel (Base-ConvNeXt-B) ve Geniş (Large- ConvNeXt-L) olmak üzere 4 varyantı bulunmaktadır.



Şekil 4. a) Convnext bloğunun yapısı b) ConvNeXt-T mimarisi

Bir ConvNeXt bloğunun yapısı ve ConvNeXt-T mimarisi Şekil 4'te, tüm varyantlar için katmanlardaki blok/kanal sayıları ve parametre sayılarına ait bilgiler ise Tablo 1'de verilmiştir.

Tablo 1. ConvNeXt varyantları

ConvNeXt Modeli	Parametre Sayısı	Kanal Sayısı	Blok Sayısı
T	27.871.224	(96, 192, 384, 768)	(3, 3, 9, 3)
S	49.505.784	(96, 192, 384, 768)	(3, 3, 27, 3)
B	87.634.456	(128, 256, 512, 1024)	(3, 3, 27, 3)
L	196.332.120	(192, 384, 768, 1536)	(3, 3, 27, 3)

Tablodan görüldüğü gibi yaklaşık 28 milyon parametreye sahip ConvNeXt-T, güçlü temsil yeteneklerini korurken kaynak kısıtlı ortamlar için en uygun varyantı temsil etmektedir. Dört aşamalı hiyerarşik tasarım, aşamalı olarak daha yüksek seviyeli öznitelik soyutlamaları aracılığıyla giriş görüntülerinin verimli bir şekilde işlenmesini sağlamaktadır. ConvNeXt-S, yaklaşık 50 milyon parametreyle ConvNeXt-T varyantını genişletmekte, makul hesaplama gereksinimlerini korurken geliştirilmiş performans sunan dengeli bir mimari sağlamaktadır. Yaklaşık 88 milyon parametreye sahip ConvNeXt-B, ConvNeXt ailesi içindeki standart temel mimari olarak hizmet vermektedir. Yaklaşık 197 milyon parametreye sahip ConvNeXt-L ise maksimum performansın hesaplama verimliliğine göre önceliklendirildiği senaryolar için tasarlanmıştır. Bu varyantların değerlendirilmesi, ConvNeXt-T'nin, sahte yüz tespiti veri kümesinde daha büyük varyantlarda gözlemlenen aşırı öğrenme eğilimlerinden kaçınırken, göreve özgü modifikasyonlar için en uygun temeli sağladığını ortaya koymuştur.

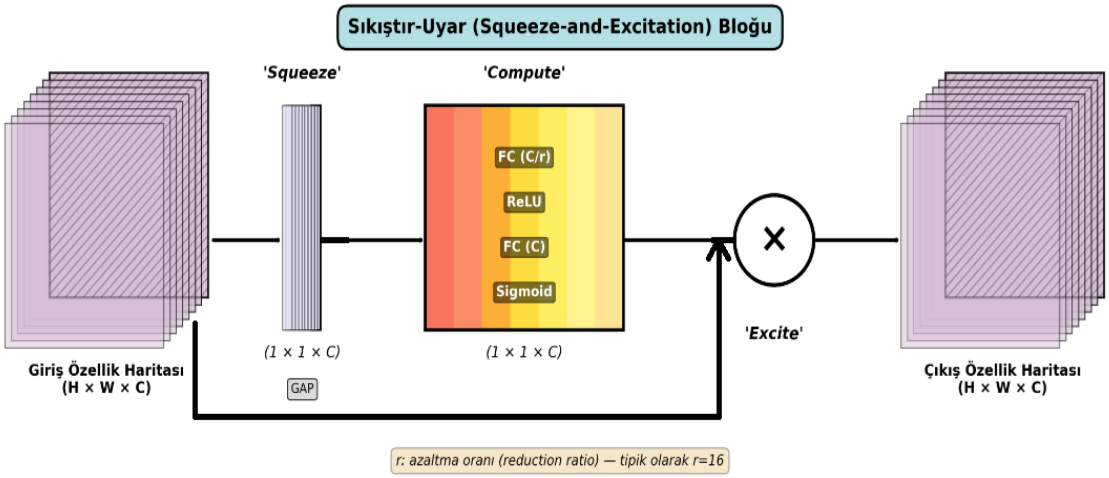
2.4.3. ConvNeXt Varyantı Geliştirme

Çalışmanın bu aşamasında, ConvNeXt mimarisinin en derin aşaması olan Aşama-3'teki blok sayısının değiştirilmesi yoluyla özel ConvNeXt mimarilerinin geliştirilmesi hedeflenmiştir. Böylece sahte yüz tespiti için hesaplama karmaşıklığı ve performans açısından en uygun ağ derinliğinin belirlenmesi sağlanmıştır. Mimarinin 3. aşamasındaki blok sayısı sırasıyla 3, 6, 9, 12, 15, 18 olarak değiştirilerek ConvNeXt [3 3 3 3], ConvNeXt

[3 3 6 3], ConvNeXt [3 3 9 3], ConvNeXt[3 3 12 3], ConvNeXt [3 3 15 3], ConvNeXt [3 3 18 3] mimarileri üretilmiştir. Bu mimarilerde hesaplama yükünü artırmamak için kanal sayıları ConvNeXt-T modelindeki gibi (96, 192, 384, 768) kullanılmıştır. Bu deneyler önemli bir sonuç ortaya çıkarmıştır: belirli bir derinlik eşiğinin ötesinde, ek Aşama-3 blokları azalan getiriler sağlamış ve aşırı öğrenme riskini artırmıştır. Bu bulgu, parametre verimliliği hususlarıyla birleştiğinde, dördüncü aşamanın tamamen ortadan kaldırılması ve dikkat mekanizmalarının dahil edilmesinin daha iyi bir performans-verimlilik dengesi sağlayacağı hipotezine yol açmıştır. Bu hipotez nihai modellerin temeli haline gelen [3,3,3,0] mimarisinin geliştirilmesine yol açmıştır.

2.4.4. Dikkat Mekanizması Entegrasyonu

Geleneksel evrişimli katmanlar, farklı kanalların eldeki özel görev için farklı öneme sahip öznitelikleri yakalamasına rağmen, tüm kanalları eşit şekilde ele alma sınırlılığına sahiptir. Sıkıştır-Uyar (SU) bloklarının entegrasyonu, kanal bazlı dikkat mekanizmaları aracılığıyla evrişimli sinir ağlarının temsil kapasitesini artırmaya yönelik bir yaklaşımı temsil etmektedir. Sahte yüz tespiti bağlamında, bu kanal bazlı dikkat mekanizması, modelin gerçek ve sahte yüzler arasında ayrım yapmak için en ayrıştırıcı olan öznitelik kanallarını vurgulamasını sağladığı için özellikle değerli hale gelmektedir. SU bloğu mimarisi, kanal önem ağırlıklarını verimli bir şekilde hesaplayan işlemler dizisinden oluşmaktadır (Şekil 5).



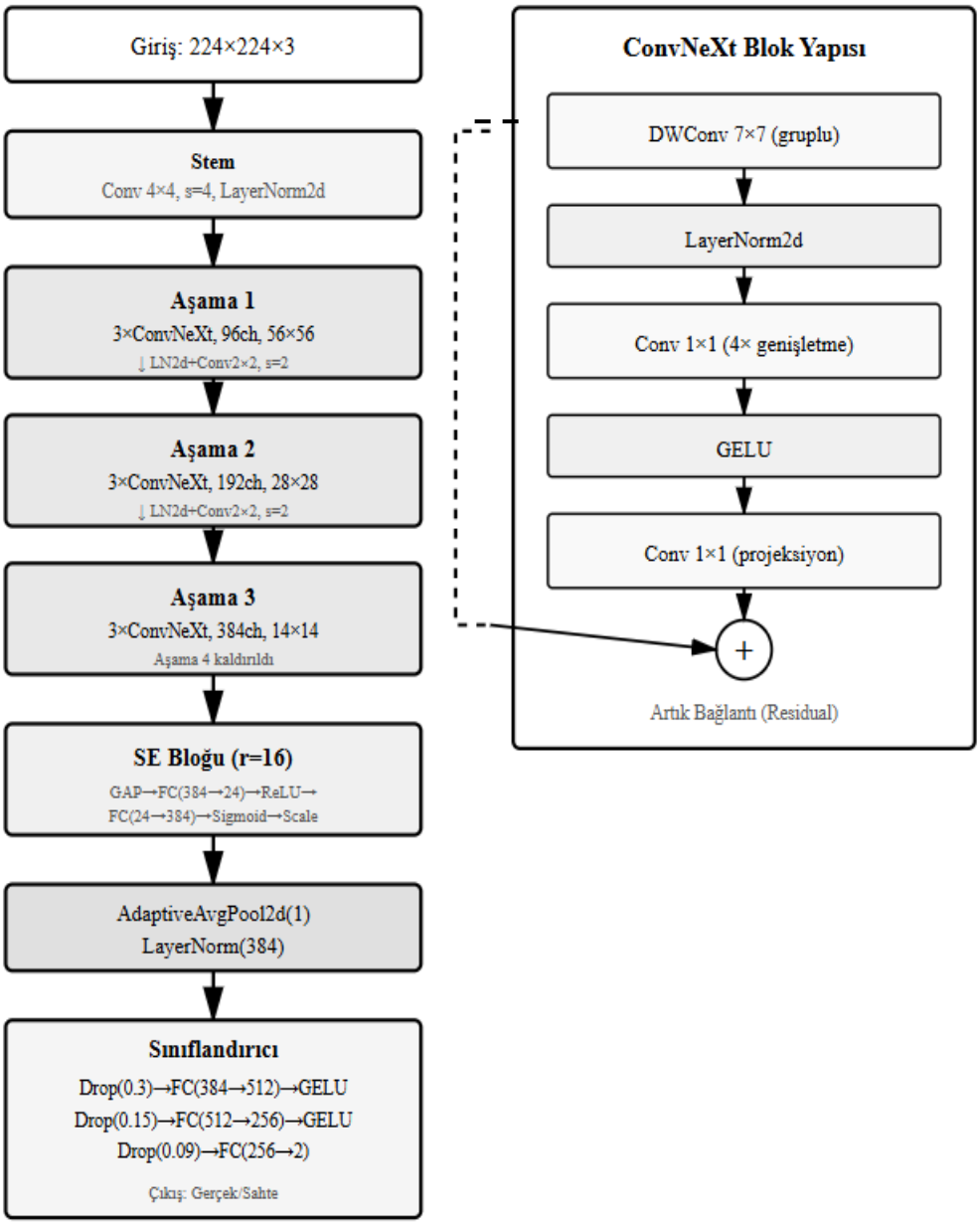
Şekil 5. SU bloğunun yapısı

Global ortalama havuzlama işlemi, uzamsal bilgiyi kanal başına tek bir değere sıkıştırarak, her kanalın genel aktivasyon gücünün bir temsilini oluşturur. Bu sıkıştırma adımını, farklı öznitelik kanalları arasındaki karmaşık ilişkileri öğrenen iki aşamalı çok katmanlı algılayıcı takip eder. İlk katman, ağın verimli kanal ilişki temsillerini öğrenmeye zorlayan bir darboğaz oluşturmak için boyutu azaltırken, ikinci katman, kanal başına dikkat ağırlıklarını üretmek için orijinal kanal sayısına genişletmektedir. SU bloklarının model mimarisine uygulanmasında, hesaplama verimliliğini temsil kapasitesiyle dengeleyen 16'lık bir azaltma oranı kullanılmaktadır.

SU bloklarının [3,3,3,0] mimarisine yerleştirilme işlemi hem hesaplama verimliliğini hem de model performansını önemli ölçüde etkilemektedir. SU bloklarının farklı aşamalar arasına ya da mimarinin sonuna yerleştirilmesiyle yapılan kapsamlı deneyler, sahte yüz tespiti görevi için uç yerleştirmenin en uygun performans-verimlilik dengesini sağladığını göstermiştir. Dikkat mekanizmalarının uç yerleştirilmesi, modelin evrişimli aşamalar tarafından çıkarılan öznitelik haritalarının önemini belirlemesini sağlamaktadır. Bu yaklaşım, dikkat hesaplamalarının, düşük seviyeli kenar dedektörlerinden yüksek seviyeli anlamsal temsillere kadar öğrenilen özniteliklerin tümüne erişmesini sağlamaktadır. Dikkat mekanizması, ayırt edici özelliklerin, farklı soyutlama seviyelerindeki öznitelikler arasındaki ince ilişkileri içerebildiği sahte yüz tespiti için özellikle değerlidir.

2.4.5. Önerilen Sahte Yüz Tespiti Modeli

Çalışmada önerilen sahte yüz tespiti modeli Şekil 6'da görülmektedir. Önerilen mimari, son aşamanın tamamen dikkat mekanizmasıyla değiştirildiği, [3,3,3,0] konfigürasyonuna sahip değiştirilmiş bir ConvNeXt tasarımını benimsemiştir. Mimari, ilk üç aşama için [96, 192, 384] kanal boyutlarını kullanmakta ve dikkat tabanlı öznitelik iyileştirmesi amacıyla dördüncü aşamayı ortadan kaldırmaktadır. Bu modifikasyon, rekabetçi performans seviyelerini korurken, yaklaşık 29 milyon parametreden 5.5 milyon parametreye düşüş sağlamıştır. Ağın sonuna SU bloğu yerleştirilerek, evrişimli öznitelik hiyerarşisine dayalı kanal dikkat hesaplamasına olanak tanınmıştır.



Şekil 6. Sahte yüz tespiti için geliştirilen [3 3 3 0] + SU modeli

Önerilen sahte yüz tespiti modelinin özellikleri şunlardır:

- Değiştirilmiş Aşama Konfigürasyonu: [3,3,3,0] tasarımı, geleneksel dördüncü aşamayı ortadan kaldırır.
- Dikkat Entegrasyonu: SU blokları, uzamsal dikkat yükü olmaksızın kanal bazlı öznitelik yeniden kalibrasyonu sağlar.

- Parametre Verimliliği: Temel ConvNeXt-T'ye kıyasla performans korunurken parametre sayısının %81.1 azalmasını sağlar.
- Dışlama Düzenlileştirmesi: Sınıflandırıcıdaki aşamalı dışlama (dropout) oranları (0.3, 0.15, 0.09), aşırı öğrenmeyi önler.
- GELU Aktivasyonu: Geleneksel ReLU'ya kıyasla gelişmiş gradyan akışı sağlar.
- LayerNorm2d: Evrişimli öznitelik haritalarına uyarlanmış kararlı normalizasyon stratejisidir.

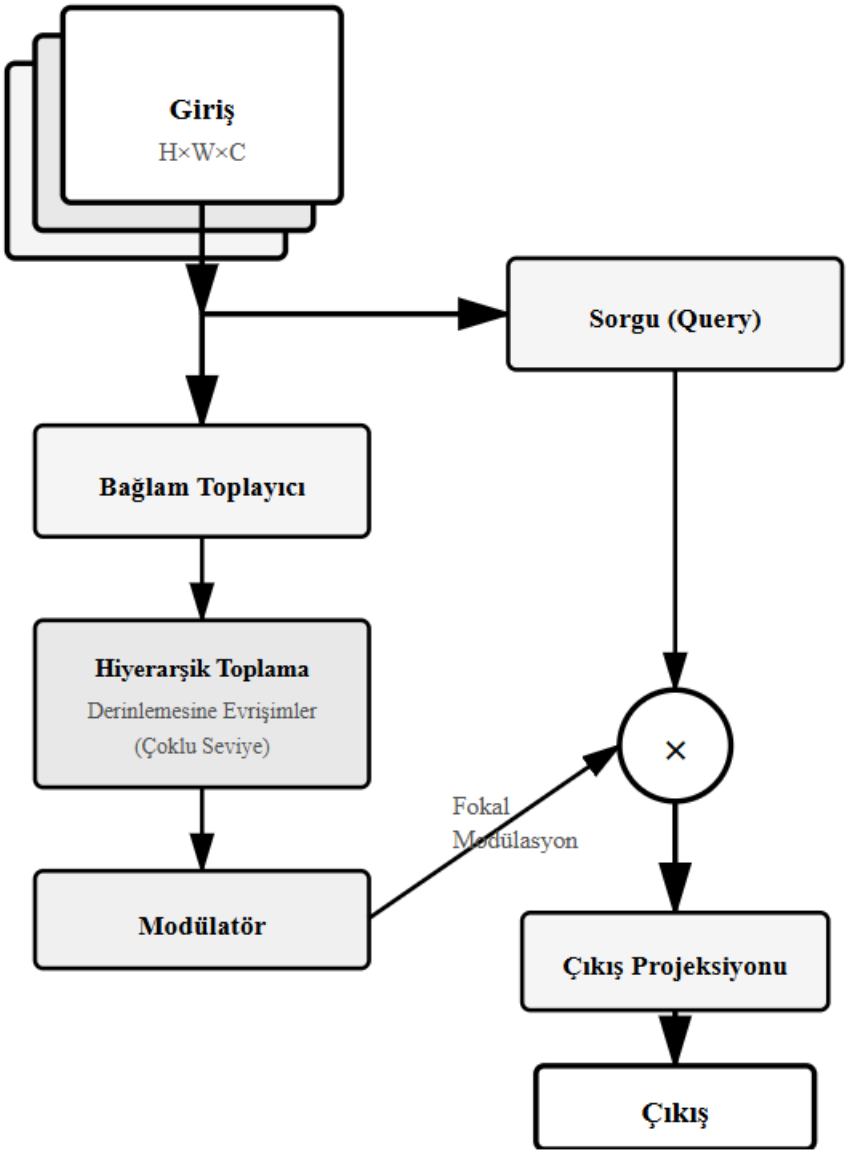
Modelin sınıflandırıcı bölümü, evrişimli temsillerden sınıflandırma kararlarına etkili öznitelik dönüşümü sağlayan üç tam bağlantılı katman (384→512→256→2) içermektedir. Aşamalı dışlama oranlarının kullanılması, aşırı öğrenme riskinin en yüksek olduğu erken sınıflandırıcı katmanlarında, uyarlanabilir düzenlileştirme sağlamaktadır.

Çalışmada son aşamada önerilen [3 3 3 0] + SU modelinin de birkaç varyantı geliştirilmiştir. Bunlar aşağıda maddeler halinde verilmiştir. Daha sonraki aşamada bu varyantlar topluluk öğrenme için girdi oluşturacaktır.

- Model-1 ([3,3,3,0]+SU): En uygun [3,3,3,0] konfigürasyonu ve sonda konumlandırılmış SU bloklarına sahip temel mimariyi temsil etmektedir.
- Model-2 ([4,3,3,0]+SU): Artan erken aşama işleme kapasitesinin öznitelik çıkarımını iyileştirip iyileştirmediğini araştırmak için Aşama-1‘e ek bir blok eklenmiştir.
- Model-3 ([3,4,3,0]+SU): Gelişmiş ara seviye öznitelik çıkarma işlemenin etkisini araştırmak için Aşama-2‘ye ek bir blok eklenmiştir.

2.4.6. FocalNet Mimarisi

FocalNet, geleneksel öz-dikkat yerine fokal modülasyon (Şekil 7) mekanizmalarını kullanan, hesaplama verimliliği ve hiyerarşik dikkat hesaplamasında teorik avantajlar sunan bir modeldir. Tez çalışmasında bu özelliklerinden dolayı FocalNet mimaris, ConvNeXt tabanlı yaklaşımlara potansiyel alternatif olarak değerlendirilmiştir.



Şekil 7. Focal modülasyon mekanizması

FocalNet-T-SRF (Küçük Alıcı Alan), FocalNet-T-LRF (Büyük Alıcı Alan) ve daha büyük ölçekli konfigürasyonlar dahil olmak üzere birden fazla FocalNet varyantı uygulanmış, her biri hesaplama verimliliği ve temsil kapasitesi arasında farklı bir dengeyi temsil etmiştir. Bu mimariler, fokal modülasyon mekanizmaları hem büyük ölçekli veri kümeleri ile önceden eğitilmiş ağırlıklarla hem de sıfırdan eğitim yoluyla değerlendirilmiştir. FocalNet ilginç teorik özellikler sergilemiş olsa da, deneysel sonuçlar bu mimarilerin sahte yüz tespiti görevi için ConvNeXt tabanlı yaklaşımlara göre önemli avantajlar sağlamadığını göstermiştir. Performans metrikleri, ConvNeXt varyantlarından sürekli olarak daha düşük

kalmış ve eğitim kararlılığı daha düşük kalmıştır. Bu bulgulara dayanarak, araştırmada daha başarılı sonuçlar veren ve göreve özgü optimizasyon için daha büyük potansiyel gösteren dikkat mekanizması entegrasyonu ile yalnızca ConvNeXt varyantlarına odaklanma kararı alınmıştır.

2.5. Topluluk Öğrenimi Yöntemleri

Topluluk öğrenimi, makine öğrenmesinde birden çok modelin birleştirilerek, herhangi bir bireysel modele kıyasla daha iyi tahmin performansı elde etmesini sağlayan bir yaklaşımdır. Bu yaklaşımın temelindeki prensip, kalabalıkların bilgeliği (wisdom of crowds) kavramına dayanır: Uygun şekilde tasarlandığında, farklı modellerden oluşan bir topluluk, daha doğru ve sağlam tahminler yapabilir.

Topluluk öğreniminin teorik temeli, model çeşitliliği aracılığıyla tahmin hatalarını azaltmaya dayanır. Aynı görev üzerinde eğitilen bireysel modeller, mimarilerindeki, eğitim prosedürlerindeki veya öğrendikleri temsillerdeki farklılıklar nedeniyle farklı türde hatalar yapabilirler. Topluluk öğrenimi ise birden çok modelden gelen tahminleri bir araya getirerek, varyansı, yanlılığı (bias) veya her ikisini de etkin bir şekilde azaltabilir. Topluluk öğrenimi yöntemleri, her biri birden çok modeli birleştirmek için belirli mekanizmalara sahip birkaç farklı yaklaşıma ayrılabilir. Bu yaklaşımlar aşağıdaki alt bölümlerde açıklanmıştır.

2.5.1. Torbalama Topluluk Öğrenimi

Torbalama, aynı model mimarisinin birden çok örneğini, eğitim verisinin farklı alt kümeleri üzerinde eğiterek varyansı azaltır. Her alt küme, veri noktalarının yerine konularak rastgele seçildiği önyükleme örneklemesi yoluyla oluşturulur. Bireysel modeller paralel olarak bağımsız bir şekilde eğitilir ve tahminleri ortalama (regresyon için) veya oylama (sınıflandırma için) yoluyla birleştirilir. Rastgele Ormanlar, torbalamanın en bilinen örneğidir; burada birden çok karar ağacı, ön yüklenmiş örnekler ve rastgele öznitelik alt kümeleri üzerinde eğitilir. Torbalamanın temel avantajı, özellikle derin sinir ağları gibi yüksek varyansa sahip modeller için aşırı öğrenmeyi azaltma yeteneğidir. Farklı veri alt kümeleri üzerinde eğitim yaparak, her model farklı karar sınırları geliştirir ve bunların ortalaması, herhangi bir tek modeldeki aykırı değerlerin veya gürültünün etkisini azaltır.

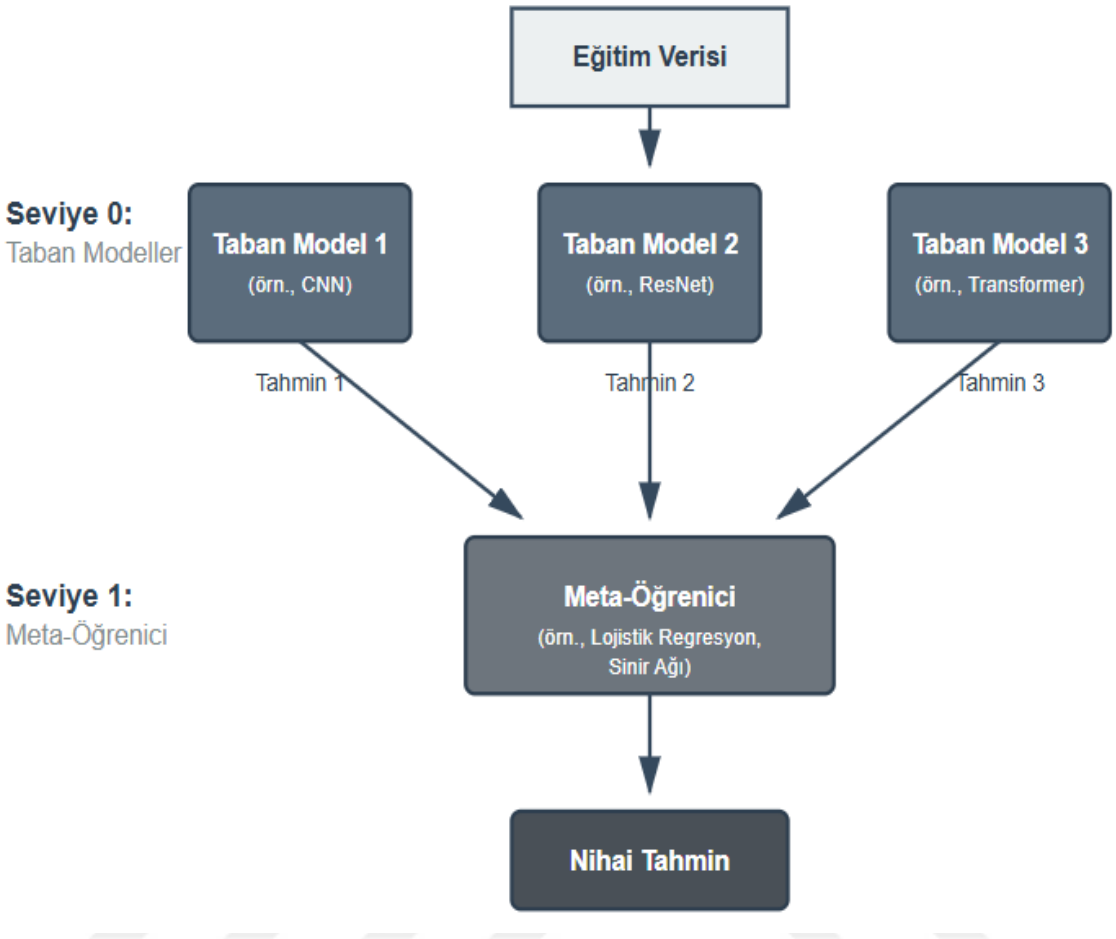
Torbalama, bireysel modellerin kararsız veya aşırı öğrenmeye eğilimli olduğu durumlarda özellikle etkilidir.

2.5.2. Yükseltme Topluluk Öğrenimi

Yükseltme, torbalamadan farklı bir prensiple çalışır: Her yeni model önceki modellerin yaptığı hataları düzeltmeye odaklanarak ardışık olarak eğitilir. Eğitim aşamalar halinde ilerler ve sonraki her modelde, önceki modeller tarafından yanlış sınıflandırılan örneklere daha yüksek ağırlık verilir. Popüler yükseltme algoritmaları arasında AdaBoost, Gradyan Yükseltme ve XGBoost bulunur. Yükseltmenin temel farkı, ardışık doğası ve sınıflandırılması zor örneklere odaklanmasıdır. Bu örneklere yinelemeli olarak odaklanarak, yükseltme hem yanlılığı hem de varyansı önemli ölçüde azaltabilir. Ancak, bu ardışık bağımlılık, aşırı öğrenmeyi önlemek için yükseltmenin daha dikkatli ayar gerektirdiği ve eğitimin torbalama kadar kolay paralelleştirilemediği anlamına gelir. Yükseltme, tipik olarak torbalamadan daha yüksek doğruluk elde eder ancak daha fazla hesaplama kaynağı ve dikkatli hiperparametre seçimi gerektirir.

2.5.3. Yığma Topluluk Öğrenimi

Yığma, birden çok farklı model tahminlerinin, daha yüksek seviyeli bir meta-öğrenicinin girdi öznitelikleri olarak hizmet ettiği bir meta-öğrenme yaklaşımını ifade eder. Süreç iki aşamadan meydana gelir. Birinci aşamada, çeşitli modeller orijinal veri seti üzerinde eğitilir; ikinci aşamada ise, bir meta-model bu tahminleri en iyi şekilde birleştirmeyi öğrenerek nihai çıktıyı üretir. Meta-öğrenici, farklı modellerinin nasıl performans gösterdiğine dair kalıpları keşfeder ve en uygun ağırlıklandırma ve birleştirme stratejisini öğrenir (Şekil 8).



Şekil 8. Yığma topluluk öğrenimi mimarisi

Yığmanın başarısı, farklı model mimarilerinin güçlü yönlerinden yararlanma yeteneğiyle ilişkilidir. Taban modellerin farklı hata kalıplarına sahip olduğu durumlarda, basit ortalama veya oylama yerine, birleştirme stratejisini öğrenerek daha yüksek performans elde edebilir. Ancak yığma, meta-öğrenicinin taban model tahminlerini aşırı öğrenmesini önlemek için dikkatli çapraz doğrulama gerektirir ve ek eğitim aşaması nedeniyle hesaplama karmaşıklığını artırır.

2.5.4. Oylama Topluluk Öğrenimi

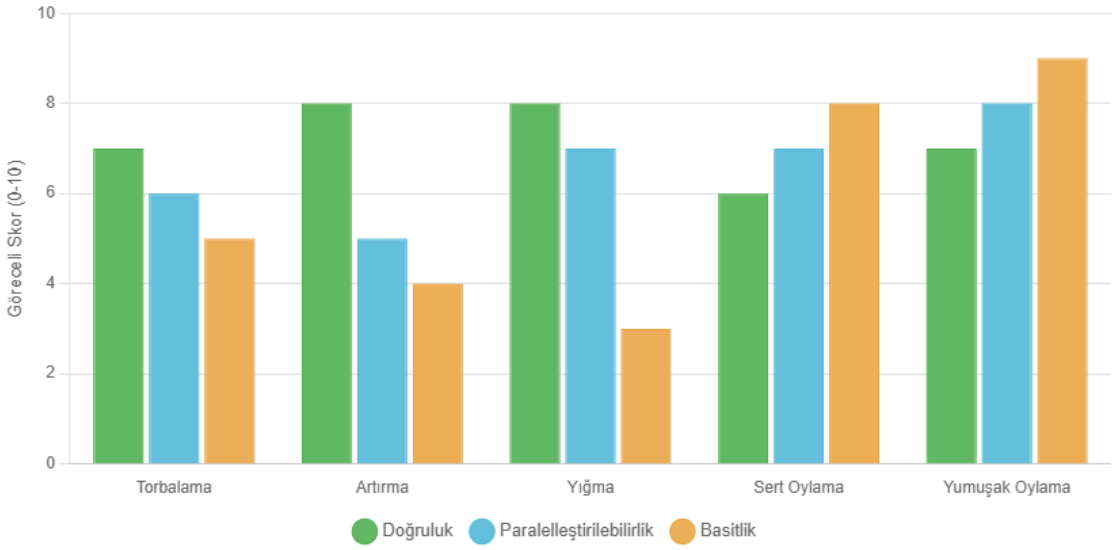
Oylama topluluk öğrenme yöntemi, birden çok eğitilmiş modelin tahminlerini doğrudan birleştiren en basit birleştirme stratejisini temsil eder. İki ana varyantı mevcuttur:

- Sert Oylama: Her modelden gelen sınıf tahminlerini sayar ve çoğunluk sınıfını seçer. Her modelin nihai karara eşit katkıda bulunması, bu yaklaşımı basit ve yorumlanabilir kılar. Sert oylama, modellerin benzer doğruluk seviyelerine sahip olduğu ve bağımsız hatalar yaptığı durumlarda iyi çalışır.
- Yumuşak Oylama: Nihai sınıflandırma kararını vermeden önce her modelden gelen tahmini sınıf olasılıklarının ortalamasını alır. Bu yaklaşım, her modelin tahminlerinin güven seviyelerinden yararlanarak, daha yüksek kesinliğe sahip modellerin nihai karar üzerinde daha fazla etkiye sahip olmasına olanak tanır. Yumuşak oylama, modeller iyi kalibre edilmiş olasılık tahminleri ürettiğinde, genellikle sert oylamadan daha iyi performans gösterir.

2.5.5. Topluluk Öğreniminin Avantajları ve Sınırlamaları

Topluluk öğrenimi, karmaşık sınıflandırma görevleri için birkaç önemli avantaj sunar. Birincil fayda, bireysel modellere kıyasla gelişmiş tahmin doğruluğu ve sağlamlığıdır. Çeşitli modellerden gelen tahminleri bir araya getirerek, bireysel model hatalarının ve yanlılıkların etkisini azaltır. Topluluk yöntemleri ayrıca, birleşik modelin bireysel modelleri etkileyen eğitim verisi özelliklerini aşırı öğrenme olasılığının daha düşük olması nedeniyle, görülmeyen verilere karşı daha iyi genelleme sağlar. Ek olarak, topluluklar, nihai tahminin eğitim verisindeki küçük varyasyonlara daha az duyarlı olması nedeniyle artan kararlılık sunar.

Ancak, topluluk öğrenimi bazı sınırlamaları da beraberinde getirir. En önemli dezavantaj, hem birden çok modeli eğitmek hem de çıkarım sırasında tüm modellerden tahminler üretmek için gereken hesaplama maliyetidir. Bellek gereksinimleri de topluluktaki model sayısıyla doğru orantılı olarak artar. Ayrıca, birleştirilmiş karar verme sürecini anlamak tek bir modeli analiz etmekten daha karmaşık hale geldiği için, topluluk yöntemlerinin yorumlanabilirliği bir miktar düşer. Topluluk hiperparametrelerinin (model sayısı, birleştirme stratejisi, çeşitlilik mekanizmaları) optimal konfigürasyonu, dikkatli deneyler gerektirir ve göreve bağımlı olabilir.



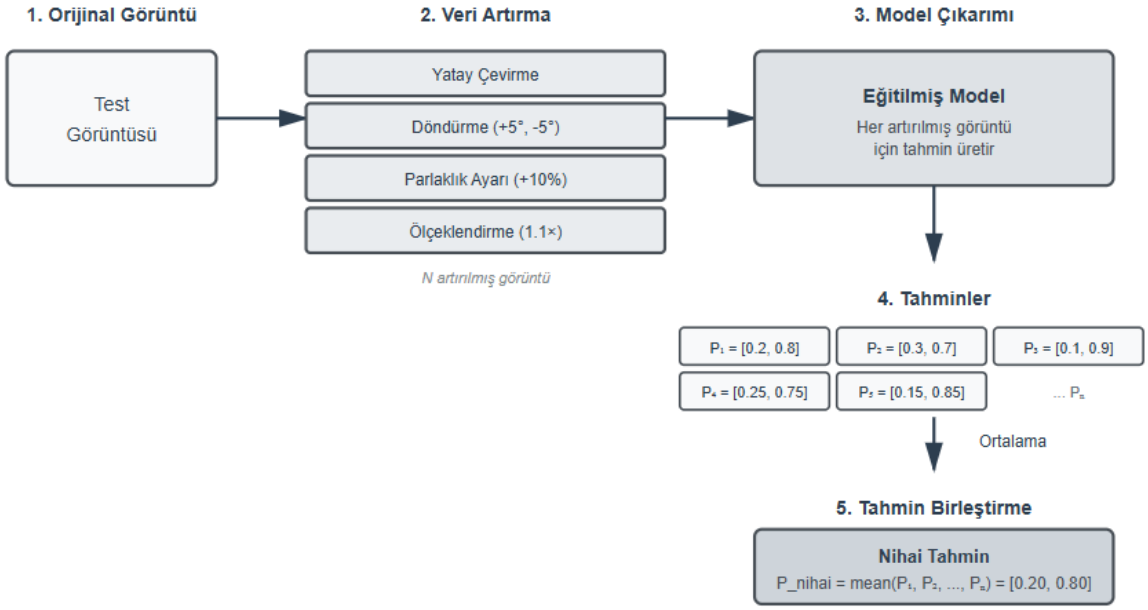
Şekil 9. Topluluk öğrenimi yöntemlerinin karşılaştırması

Şekil 9, farklı topluluk öğrenimi yöntemlerinin doğruluk, paralelleştirilebilirlik ve basitlik açısından karşılaştırmalı bir analizini sunmaktadır. Bu karşılaştırma, her yöntemin güçlü ve zayıf yönlerini göz önünde bulundurarak en uygun topluluk stratejisinin seçilmesine yardımcı olmaktadır. Şekilden görüldüğü üzere, yığma ve yumuşak oylama yöntemleri doğruluk açısından en yüksek skorlara sahiptir. Yığma, meta-öğrenici aracılığıyla taban modellerin tahminlerini optimal şekilde birleştirme yeteneği sayesinde üstün performans sunmaktadır. Yumuşak oylama ise basitliği ve paralelleştirilebilirliği ile öne çıkmakta, aynı zamanda yüksek doğruluk sağlamaktadır. Bu nedenlerle, bu çalışmada sahte yüz tespiti için yığma ve oylama tabanlı topluluk yöntemleri tercih edilmiştir

2.6. Test Zamanı Artırımı

Test Zamanı Artırımı (TZA), her bir test girdisinin birden çok artırılmış versiyonunu üreterek ve bunların tahminlerini birleştirerek çıkarım sırasında model performansını artıran bir tahmin iyileştirme tekniğidir. Veri çeşitliliğini artırmak için yalnızca eğitim sırasında uygulanan geleneksel veri artırımının aksine, TZA daha sağlam ve doğru tahminler üretmek için test veya dağıtım aşamasında artırımları uygular. TZA'nın temel prensibi, iyi eğitilmiş bir modelin, aynı girdinin anlamsal olarak eşdeğer varyasyonları için tutarlı tahminler üretmesi gerektiği varsayımına dayanır. Ancak pratikte, modeller altta yatan içeriği veya

sınıfı değiştirmeyen küçük girdi varyasyonlarına duyarlı olabilir. TZA, dikkatlice tasarlanmış artırımlar aracılığıyla her test girdisinin birden çok versiyonunu oluşturur, her versiyon için tahminler elde eder ve bu tahminleri birleştirerek nihai ve daha güvenilir bir çıktı üretmeye çalışır. TZA yöntemi Şekil 10'da görülmektedir.



Şekil 10. Test Zamanı Artırma (TZA) işlem akışı

Şekilden görüldüğü gibi TZA süreci, her test örneği için sistematik bir süreç izler:

Adım 1: Artırım Üretimi: Her bir test girdisi için, çeşitli dönüşümler uygulanarak birden çok artırılmış versiyon oluşturulur. Bu dönüşümler, görünümde kontrollü varyasyonlar sunarken anlamsal içeriği korur. Yaygın artırım teknikleri arasında yatay çevirme, küçük açılarla döndürme, hafif ölçekleme, rastgele kırpma, parlaklık ve kontrast ayarlamaları ile renk titreşimi (color jittering) bulunur. Artırımların seçimi ve parametreleri, özgül göreve ve etiket değişmezliğini koruyan dönüşümlere bağlıdır.

Adım 2: Model Çıkarımı: Girdinin her bir artırılmış versiyonu, eğitilmiş model aracılığıyla bağımsız olarak işlenerek bir dizi tahmin üretilir. Model, her artırılmış girdiyi, model parametrelerinde veya mimarisinde herhangi bir değişiklik olmaksızın, normal test girdilerini işlediğiyle aynı şekilde işler.

Adım 3: Tahmin Toplama: Tüm artırılmış versiyonlardan gelen tahminler birleştirilerek nihai tahmin üretilir. Sınıflandırma görevleri için standart yaklaşım, tahmini olasılık dağılımlarının ortalamasını almak ve en yüksek ortalama olasılığa sahip sınıfı

seçmektir. Alternatif toplama stratejileri, tahmini sınıfların modu (sert oylama) veya tahmin güvenine dayalı ağırlıklı ortalama almayı içerir. Toplanan tahmin, tipik olarak tek girdilere yapılan tahminlere kıyasla daha düşük varyansa ve daha yüksek sağlamlığa sahip olur.

Adım 4: Nihai Karar: Toplanan tahmin, orijinal test girdisi için nihai çıktı olarak hizmet eder. Bu süreç, artırımların potansiyel olarak rastgeleleştirildiği veya önceden tanımlanmış bir sete göre sistematik olarak uygulandığı her test örneği için bağımsız olarak tekrarlanır.

2.6.1. Test Zamanı Artırımı Türleri

Farklı artırım türleri, TZA'da belirli amaçlara hizmet eder:

- Geometrik Dönüşümler: Yatay çevirme, küçük açılar (tipik olarak ±10−15 derece) içinde döndürme ve ölçekleme gibi işlemleri içerir. Bu dönüşümler, modelin gerçek dünya dağıtım koşullarında ortaya çıkabilecek küçük geometrik varyasyonlara karşı değişmezlik elde etmesine yardımcı olur. Yatay çevirme, birçok görsel görev için anlamsal içeriği koruduğu için özellikle yaygındır.
- Fotometrik Dönüşümler: Parlaklık ayarlamaları, kontrast değişiklikleri, doygunluk modifikasyonları ve ton kaymaları dahil olmak üzere, içeriği değiştirmeden görüntü görünümünü değiştirir. Bu artırımlar, pratikte sıklıkla ortaya çıkan aydınlatma koşullarındaki, kamera ayarlarındaki veya görüntü kalitesindeki varyasyonlara karşı sağlamlığı artırır.
- Uzamsal Dönüşümler: Rastgele kırpma veya en boy oranı değişiklikleri gibi işlemleri içerir. Bunlar, görüntüler içindeki nesnelerin konumlandırmasındaki veya çerçevelemedeki varyasyonları hesaba katmaya yardımcı olur.
- Hibrit Yaklaşımlar: Hem yatay çevirme hem de hafif döndürme uygulama veya geometrik ve fotometrik değişiklikleri birleştirme gibi birden çok dönüşüm türünü birleştirir. Bu artırım türlerinin kombinasyonu fayda sağlayabilir, ancak artırılmış versiyonların sayısı, eklenen her dönüşümle çarpımsal olarak büyür.

2.6.2. TZA Faydaları ve Dikkat Edilmesi Gerekenler

TZA yöntemi, çıkarım sırasında model performansını iyileştirmek için birkaç temel avantaj sunar. Birincil fayda, tahmin varyansının azaltılması yoluyla artan tahmin

doğruluğudur. Birden çok artırılmış versiyon üzerinden tahminleri ortalamak, TZA'nın küçük girdi varyasyonlarına daha az duyarlı, daha kararlı ve güvenilir çıktılar üretmesini sağlar. Bu gelişmiş sağlamlık, girdi koşullarının değişebileceği gerçek dünya dağıtım senaryolarında, modelleri daha güvenilir hale getirir. TZA ayrıca, birden çok farklı model gerektirmeden topluluk benzeri bir fayda sağlar. Model çeşitliliği yerine girdi varyasyonu yoluyla çeşitli tahminler üretmek için tek bir eğitilmiş model kullanılır, bu da geleneksel model topluluklarına kıyasla hesaplama verimliliği sunar. Ek olarak, TZA, yeniden eğitim veya orijinal modelin modifikasyonunu gerektirmeyen, eğitim sonrası bir teknik olarak uygulanabilir, bu da onu dağıtılmış sistemleri iyileştirmek için pratik bir seçenek haline getirir.

Ancak TZA çıkarım sırasında hesaplama yükü getirir, çünkü model her girdinin birden çok versiyonunu işlemek zorundadır. Çıkarım süresi, uygulanan artırım sayısıyla doğrusal olarak artar, bu da gecikmeye duyarlı uygulamalar için sorun oluşturabilir. Artırım türleri ve en uygun sayısı, doğruluk iyileştirmeleri ile hesaplama maliyetleri arasındaki dengeyi sağlamak için dikkatlice seçilmelidir. Ayrıca, belirli artırımların etkinliği göreve bağımlıdır ve belirli bir uygulama için en faydalı dönüşümleri belirlemek deneysel doğrulama gerektirebilir.

3. BULGULAR

3.1. Giriş

Bu bölümde, derin sahte yüz tespit sistemi geliştirme sürecinde gerçekleştirilen deneysel çalışmaların analizi sunulmaktadır. Araştırma metodolojisinin temelini, özellikle ConvNeXt tabanlı modeller ve topluluk öğrenimi tekniklerine odaklanılarak, çeşitli derin öğrenme mimarilerinin sistematik incelenmesi oluşturmaktadır. Gerçekleştirilen bu çalışmalar, hesaplama verimliliği korunarak tespit doğruluğunun aşamalı olarak iyileştirilmesi hedefiyle tasarlanmış olup, nihai olarak test veri kümesinde %99.66 doğruluk oranına ulaşılmıştır.

Deneysel çerçeve, temel model oluşturma ile başlayan ve mimari modifikasyonlar, topluluk entegrasyonu ve gelişmiş optimizasyon teknikleri aracılığıyla ilerleyen çok aşamalı bir yapıya sahiptir. Araştırma kapsamında, farklı model mimarileri, eğitim stratejileri ve değerlendirme metodolojileri ile kapsamlı deneyler yürütülmüştür. Bu deneylerden elde edilen bulgular alt bölümlerde verilmiştir.

3.2. Deneysel Ayarlar

Bu araştırmada, güçlü ve dayanıklı bir derin sahte tespit modeli oluşturulması amacıyla çok sayıda kaynaktan yararlanılarak bir veri kümesi üretilmiştir. Bu süreçte FaceForensics++ veri setindeki video dizilerinden kareler çıkarılmıştır. Anahtar kareler, yüz görünürlüğü, görüntü kalitesi ve zamansal dağılım temelinde seçilerek, yüksek görsel kalite standartları korunmakla birlikte yüz pozları, ifadeleri ve ışık koşullarının çeşitliliği maksimize edilmiştir. Elde edilen görüntülerden yüz bölgelerinin doğru şekilde tespit edilmesi için güncel yüz tespit algoritmalarından yararlanmıştır. Tüm görüntülerin model eğitimi için tutarlı girdi karakteristiklerinin sağlanması amacıyla standardizasyon uygulanmıştır. Şekil 11 ve Şekil 12, sırasıyla FaceForensics++ ve CelebA veri setlerinden gerçek yüzlerin ve FaceForensics++ veri setinden manipüle edilmiş yüzlerin örneklerini göstermektedir.

Şekil 11. FaceForensics++ ve CelebA veri setlerinden gerçek görüntüler

Şekil 12. FaceForensics++ verisetinden sahte görüntüler

Sonraki aşamada, veri sızıntısının önlenmesi için, veri seti eğitim, doğrulama ve test setlerine ayrıştırılmıştır. Ayrıştırma işleminde video bazlı bölümlendirme stratejisi kullanılmıştır. Manipüle edilmiş görüntüler için, aynı kaynak videodan çıkarılan tüm karelerin (video başına 3-6+ kare) aynı alt kümeye (eğitim, doğrulama veya test) atanması

sağlanarak, video düzeyinde veri izolasyonu gerçekleştirilmiştir. Bu yaklaşım, modelin eğitim sırasında bir videodan öğrendiği yapay bozuklukları, test aşamasında aynı videonun farklı karelerinde tanımasından kaynaklanan performans artışını önlemektedir. Gerçek görüntüler için, FaceForensics++ veri setinden video başına bir kare çıkarılmış ve CelebA veri setinden alınan benzersiz kimlik görüntüleriyle desteklenmiştir; bu sayede tüm gerçek görüntülerin farklı bireyleri temsil etmesi garanti altına alınmıştır. Bölümlendirme aşamasında, tüm alt kümelerde farklı manipülasyon tekniklerinin (Deepfakes, Face2Face, FaceSwap, NeuralTextures) ve demografik karakteristiklerin oransal temsilinin korunması sağlanmıştır. Deneylerde kullanılan eğitim, doğrulama ve test kümelerindeki örnek dağılımı Tablo 2'de görülmektedir.

Tablo 2. Veri kümesi dağılımı

Veri Seti Bölümü	**Gerçek Görüntüler**	**Sahte Görüntüler**	**Toplam Görüntüler**	**Gerçek/Sahte Oranı**
Eğitim	2,400	2,468	4,868	%49.3 / %50.7
Doğrulama	300	300	600	%50.0 / %50.0
Test	300	300	600	%50.0 / %50.0
Toplam	3,000	3,068	6,068	%49.4 / %50.6

Tablo 2'de özetlendiği üzere, nihai veri kümesi, her bir alt kümede gerçek ve sahte (manipüle edilmiş) örneklerin dengeli bir temsilini içerecek şekilde 4.868 eğitim görüntüsü, 600 doğrulama görüntüsü ve 600 test görüntüsünden oluşturulmuştur. Bu dengeli dağılım, hem eğitim hem de değerlendirme süreçlerinde herhangi bir sınıfa karşı oluşabilecek önyargıyı önlemek için önemlidir.

Deneysel kurulum, farklı yaklaşımlar arasında adil bir karşılaştırma sağlamak amacıyla tüm deneyler boyunca standardize edilmiştir. Bu bağlamda, tüm modeller aynı donanım konfigürasyonları, optimizasyon parametreleri ve değerlendirme metrikleri kullanılarak eğitilmiştir. Deneysel parametreler Tablo 3'te görülmektedir. Elde edilen sonuçların tekrarlanabilirliği ise, bir başlangıç değeri (seed) yönetimi ve tüm deneysel parametrelerin detaylı dokümantasyonu ile güvence altına alınmıştır.

Tablo 3. Deneysel parametreler

Kategori	Parametre	Değer
Donanım Konfigürasyonu	GPU	NVIDIA L4
	Hesaplama Platformu	Google Colab Pro
Eğitim Parametreleri	Yığın Boyutu	32
	Eğitim Tur Sayısı	150
	Öğrenme Katsayısı	0.001
	Ağırlık Azalması	0.05
	Optimizasyon Yöntemi	AdamW
	Kayıp Fonksiyonu	CrossEntropyLoss (etiket düzeltme: 0.1)
	Öğrenme Oranı Zamanlayıcısı	CosineAnnealingWarmRestarts (T_0=30, T_mult=2, η_min=1e-6)
	Eğim Kırpma	max_norm = 1.0
	Erken Durdurma Sabrı	35 epoch
Model Parametreleri	Dışlama Oranı	0.3
	SU Bloğu İndirgeme Oranı	16
	Aktivasyon Fonksiyonu	GELU
Veri Artırma	Girdi Görüntü Boyutu	224 × 224 piksel
	Eğitim Artırma Teknikleri	RandomCrop, RandomHorizontalFlip (p=0.5), RandomVerticalFlip (p=0.2), RandomRotation (±20°), ColorJitter, RandomGrayscale (p=0.1), RandomAdjustSharpness (p=0.3), RandomErasing (p=0.1)
	Normalizasyon	ImageNet standardı (ortalama=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225])
Tekrarlanabilirlik	Rastgele Tohum	Sabitlenmemiş (çoklu çalıştırmaların ortalaması)
	Çerçeve	Python 3.10.x, PyTorch 2.x
	CUDA Sürümü	11.x
Veri Seti Konfigürasyonu	FaceForensics++ Sıkıştırma Seviyesi	c23/c40
	Gerçek Görüntüler	1 (kare/video)
	Sahte Görüntüler	3-6+ (kare/video)
	Eğitim Görüntüsü	4,868
	Doğrulama Görüntüsü	600
	Test Görüntüsü	600
	Yüz Tespit Yöntemi	MTCNN
Yazılım Ortamı	Python Sürümü	3.10.x

3.3. Ön Eğitimli Modellerin Derin Sahte Tespiti Performansının Değerlendirilmesi

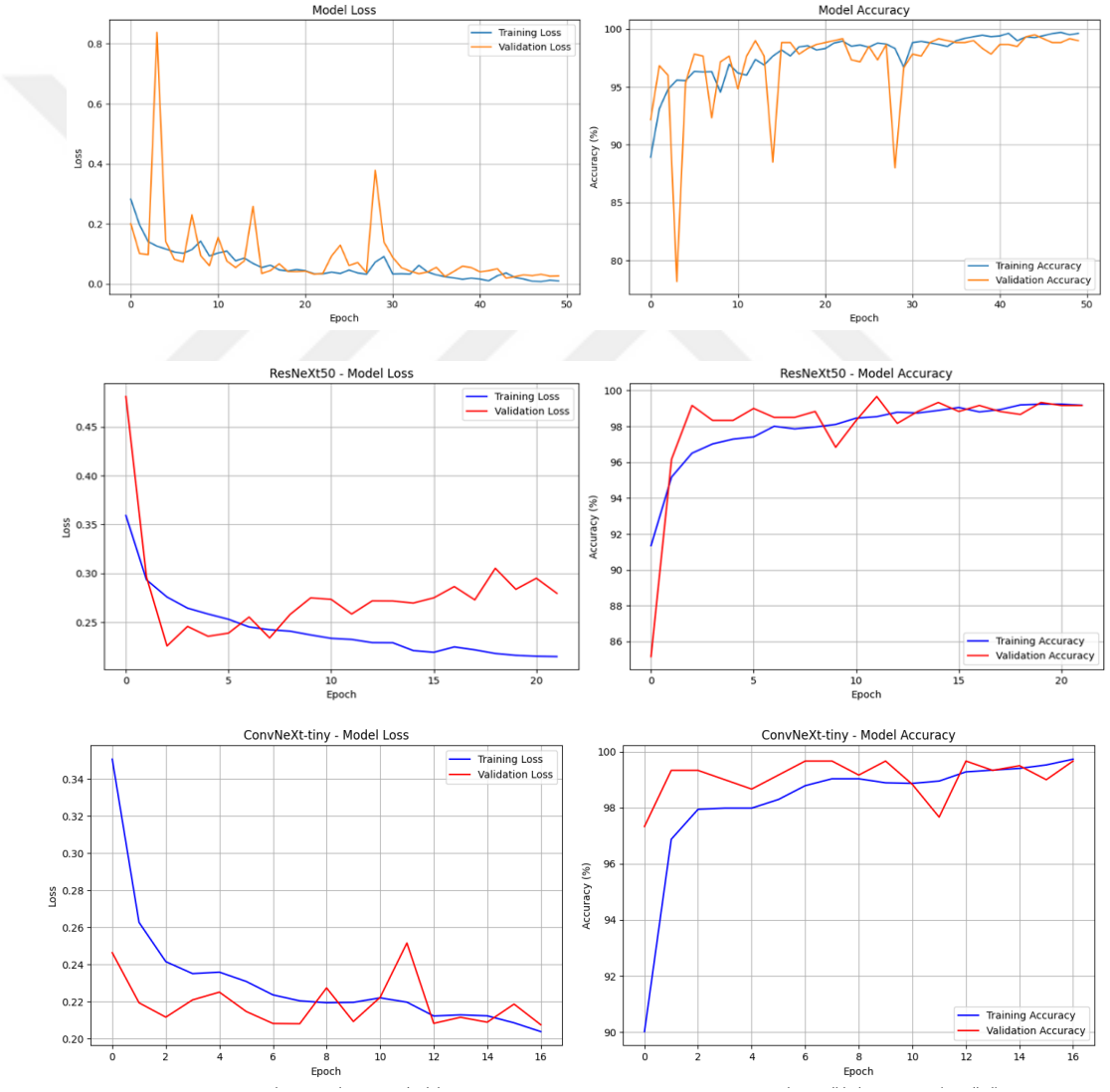
3.3.1. Temel Modellerle Derin Sahte Tespiti

Araştırmanın ilk aşamasında, ön eğitimli derin öğrenme mimarileri kullanılarak temel performans değerlendirilmiştir. Bu modellerin, derin sahte tespiti alanındaki etkinliklerinin anlaşılması ve ilerleyen aşamalardaki iyileştirmeler için güvenilir performans kıyaslamalarının belirlenmesi amacıyla, çeşitli önceden eğitilmiş modeller değerlendirilmiştir. Temel deneyler, ResNet50, ResNeXt50 ve çeşitli ConvNeXt varyantları dâhil olmak üzere önceden eğitilmiş güncel modeller ile gerçekleştirilmiştir. Bu modellerin seçiminde, bilgisayarlı görü görevlerindeki kanıtlanmış başarıları ve standart derin öğrenme çerçevelerindeki yaygın kullanımları temel alınmıştır. Her bir model için, önceden öğrenilmiş özellik temsillerinden faydalanan aktarım öğrenimi teknikleri kullanılarak, derin sahte tespiti veri kümesine uyum sağlaması için ince ayar yapılmıştır. Bu ön eğitimli modellerle doğrulama ve test kümesi üzerinde elde edilen derin sahte tespiti başarımları Tablo 4'te, başarım grafikleri ise Şekil 13'te verilmiştir.

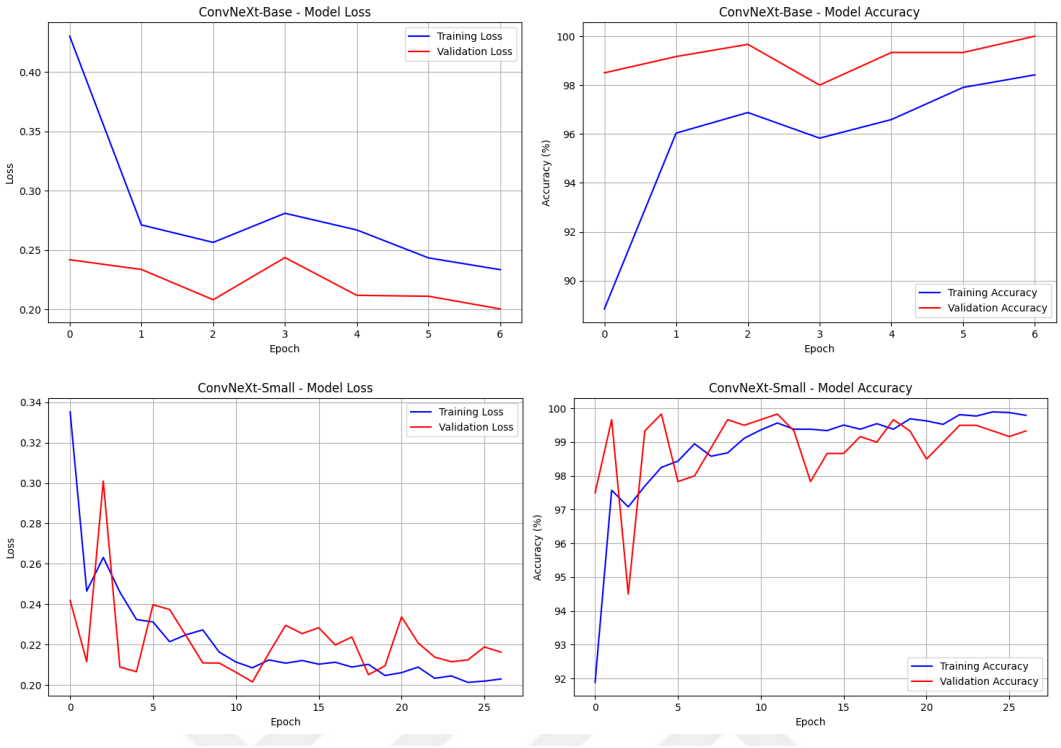
Tablo 4. Önceden eğitilmiş modellerin performans karşılaştırması

Metrik	**ResNet50**	**ResNeXt50**	**ConvNeXt-T**	**ConvNeXt-S**	**ConvNeXt-B**	**ConvNeXt-L**
Eğitim Yaklaşımı	Ön Eğitimli	Ön Eğitimli	Ön Eğitimli	Ön Eğitimli	Ön Eğitimli	Ön Eğitimli
Parametreler	24,558,146	24,160,834	29,114,698	50,749,258	90,297,706	203,669,930
Test Doğruluğu (%)	99.00	99.17	99.67	99.50	100.00	99.83
AUC Skoru	1.000	1.000	1.000	1.000	1.000	1.000
Kesinlik (Gerçek/Sahte)	1.00 / 0.98	1.00 / 0.98	1.00 / 1.00	1.00 / 0.99	1.00 / 1.00	1.00 / 1.00
Duyarlılık (Gerçek/Sahte)	0.98 / 1.00	0.98 / 1.00	1.00 / 1.00	0.99 / 1.00	1.00 / 1.00	1.00 / 1.00
F1-Skoru (Gerçek/Sahte)	0.99 / 0.99	0.99 / 0.99	1.00 / 1.00	0.99 / 1.00	1.00 / 1.00	1.00 / 1.00
En İyi Doğ. Doğruluğu (%)	99.50	99.67	99.67	99.83	100.00	100.00

Tablo 4'te sunulan temel değerlendirme sonuçları, ConvNeXt mimarilerinin geleneksel ResNet tabanlı modellere göre daha iyi performans sergilediğini açıkça ortaya koymuştur. ConvNeXt-Base modelinin, %100 test doğruluğu ile en yüksek performansı elde ederek, gerçek ve derin sahte yüz görüntülerini ayırt etmede çok başarılı olduğu görülmüştür. Bu performans, ConvNeXt'in sonraki deneysel aşamalar için birincil mimari olarak belirlenmesini sağlamıştır. Fakat bu başarım yaklaşık 87 milyon parametreli bir model ile elde edilmiştir. Çalışmada daha hafif modellerin üretilmesi ve başarım ile model ağırlığı arasında denge oluşturulması amaçlandığı için bir sonraki aşamada daha hafif ConvNeXt mimarileri tasarlanmıştır.

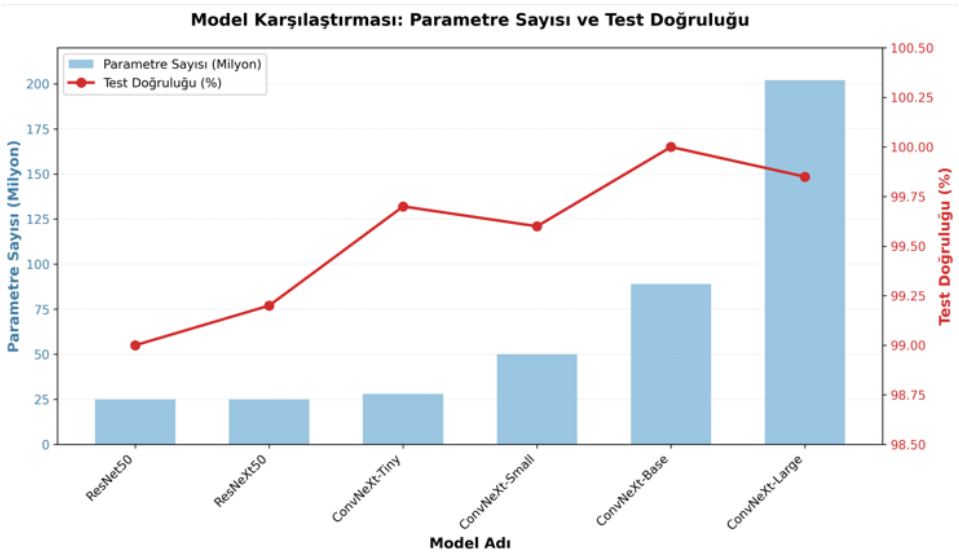


Şekil 13. ResNet, ResNeXt, ConvNeXt Tiny modellerinin başarım grafikleri



Şekil 14. ConvNeXt Base, Small modellerinin başarım grafikleri

İlk temel değerlenmeden sonra tespit performansına en önemli katkıyı sağlayan mimari bileşenleri anlamak amacıyla detaylı bir analiz gerçekleştirilmiştir. Bu analiz, parametre sayısı, model derinliği ve ConvNeXt mimarilerine özgü tasarım elemanları dâhil olmak üzere farklı mimari seçimlerin etkilerini incelemeye odaklanmıştır. Temel modellerin parametre sayısı ve başarım olarak karşılaştırılması Şekil 15'te görülmektedir.



Şekil 15. Temel modellerin parametre sayısı ve başarım olarak kıyaslanması

Şekilden, daha büyük modellerin genellikle daha yüksek doğruluk elde ettiği, fakat bu ilişkinin tam olarak doğrusal olmadığı görülmektedir. ConvNeXt-T'nin, önemli ölçüde daha az parametre ile rekabetçi performans sergilemesi, modelin verimli bir parametre kullanımına sahip olduğunu düşündürmektedir. Bu bulgu, pratik dağıtım senaryolarına uygun hafif ancak başarımı yüksek modellerin geliştirilmesine odaklanan sonraki deneyler için kritik bir öneme sahiptir.

3.3.2. Topluluk Öğrenme ile Model Birleştirme

Bireysel ConvNeXt modellerinin sergilediği güçlü temel performanstan hareketle, çalışmada, tespit doğruluğunun daha da iyileştirilmesi amacıyla topluluk öğrenimi teknikleri uygulanmıştır. Topluluk yöntemlerinin, birden fazla modelin tahminlerinin birleştirilmesi yoluyla aşırı uyumu azaltma ve genelleme kabiliyetini iyileştirme konularında çeşitli makine öğrenimi alanlarında etkinliği kanıtlanmıştır. Topluluk metodolojisinde, nihai tahmine çeşitli bakış açılarının katılımını sağlayabilecek, tamamlayıcı güce sahip modellerin birleştirilmesi ise başarımın yükseltilmesi hedeflenmektedir. Bu amaçla çalışmada, bireysel performans karakteristikleri ve mimari farklılıkları temel alınarak birden fazla ConvNeXt varyantı seçilmiştir. Bu stratejiyle, farklı modellerin güçlü yönlerinden faydalanan ve bireysel zayıflıklarını en aza indiren bir topluluk yapısı oluşturulması amaçlanmıştır.

Çalışmada üç ana oylama stratejisi uygulanmış ve değerlendirilmiştir: sert oylama, yumuşak oylama ve ağırlıklı oylama. Sert oylama, topluluk üyeleri arasında çoğunluk kararı alınması yoluyla tahminleri birleştirirken, yumuşak oylama tahmin edilen olasılıkların ortalaması üzerinden gerçekleştirilmiştir. Ağırlıklı oylama ise, bireysel performans metriklerine dayalı olarak topluluk üyelerine farklı önem ağırlıklarının atanmasıyla yumuşak oylamanın daha gelişmiş bir uzantısını temsil etmektedir. Çalışmada ConvNeXt varyantlarının tahminlerinin birleştirilmesi ile elde edilen derin sahte tespiti başarımları Tablo 5'te verilmiştir.

Tablo 5. Topluluk yöntemleri ile derin sahte tespiti performansı

Topluluk Kombinasyonu	Oylama Yöntemi	Test Doğruluğu (%)	AUC Skoru
ConvNeXt-T+ ConvNeXt-B	Sert Oylama	99.83	1.000
ConvNeXt-T+ ConvNeXt-B	Yumuşak Oylama	100.00	1.000
ConvNeXt-Y + ConvNeXt-B	Ağırlıklı Oylama	100.00	1.000
ConvNeXt-T + ConvNeXt-S	Yumuşak Oylama	100.00	1.000
ConvNeXt-T + ConvNeXt-L	Yumuşak Oylama	100.00	1.000
ConvNeXt-B + ConvNeXt-L	Yumuşak Oylama	100.00	1.000
ConvNeXt-T + B+ S	Sert Oylama	100.00	1.000

Tablo 5'te gösterilen deneysel sonuçlar, topluluk yöntemlerinin bireysel modellere kıyasla tutarlı bir şekilde daha iyi performans sergilediğini, özellikle yumuşak ve ağırlıklı oylama yaklaşımlarının en iyi sonuçları elde ettiğini göstermiştir. En iyi başarım ve parametre sayısı dengesi, ConvNeXt-T + ConvNeXt-S kombinasyonunun yumuşak oylama ile birleştirilmesiyle, %100 test doğruluğu olarak elde edilmiştir. Çünkü bu birleşimin parametre sayısı (79M), aynı başarımı gösteren diğer birleşimlerin ya da yüksek başarımlı ConvNeXt-B modelinin parametre sayısından (90M) daha azdır.

3.4. Önerilen Derin Sahte Tespiti Modelinin Başarımının Değerlendirilmesi

3.4.1. ConvNeXt Aşama Konfigürasyonu Optimizasyonu

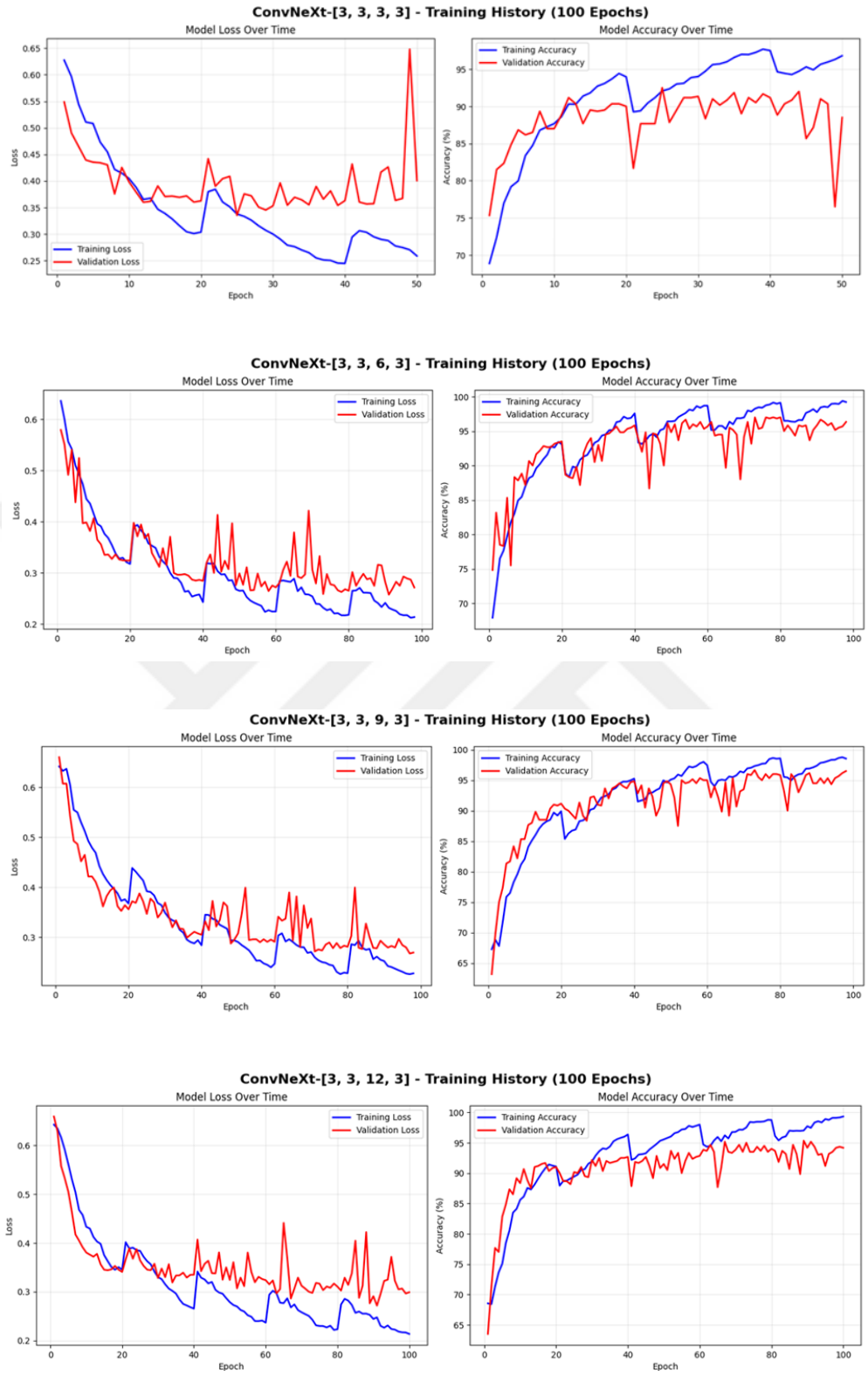
Temel mimari değerlendirmesinin ardından, derin sahte tespiti görevleri için optimal mimari derinliğini belirlemek amacıyla ConvNeXt aşamalarındaki katman sayıları sistematik olarak değiştirilerek deneyler gerçekleştirilmiştir. Deneylerde, gerçek ve manipüle edilmiş yüz içeriğinin ayırt edilmesi için kritik önem taşıyan orta seviye özelliklerin yakalanması nedeniyle, özellikle ConvNeXt mimarisinin Aşama 3'ündeki blok sayısının değiştirilmesine odaklanılmıştır. Çalışmada minimal karmaşıklıktan maksimum karmaşıklığa kadar altı farklı Aşama-3 konfigürasyonu değerlendirilmiştir. Test edilen konfigürasyonlar [3,3,3,3], [3,3,6,3], [3,3,9,3], [3,3,12,3], [3,3,15,3] ve [3,3,18,3] olarak belirlenmiştir. Her bir konfigürasyonda, orta seviye özellik karmaşıklığının tespit performansı üzerindeki etkisini belirlemek amacıyla Aşama-1, Aşama-2 ve Aşama-4 parametreleri aynen korunmuştur.

ConvNeXt'in T, S, B ve L modellerinde, Tablo 1'de görüldüğü gibi hem Aşama-3'teki katman sayısı, hem de aşamalarda kullanılan filtre sayıları artmaktadır. Bu ise yine tablodan görüldüğü gibi modelin ağırlaşmasına ve kaynak kısıtlı ortamlarda verimli çalışamamasına neden olmaktadır. Bu nedenle üretilen ConvNeXt varyantlarının katmanlardaki filtre sayıları ConvNeXt-T modelindeki gibi kullanılmış, Aşama-3'teki katman sayısı değiştirilerek daha hafif modellerin üretilmesi hedeflenmiştir. Bu deneysel çalışmadaki tüm modeller, aynı eğitim prosedürleri, optimizasyon parametreleri ve değerlendirme protokolleri ile 100 tur boyunca eğitilmiştir. Aşırı uyumun önlenmesi ve eğitim kaynaklarının verimli kullanılması için 25 tur sabır parametresi ile erken durdurma mekanizmaları uygulanmıştır. Öğrenme katsayısı, eğitim sürecinde CosineAnnealingWarmRestarts zamanlayıcısı kullanılarak dinamik olarak ayarlanmıştır. Bu zamanlayıcı, her 20 epoch'ta öğrenme katsayısını başlangıç değerine (0.001) geri döndürmekte ve kosinüs benzetim yaklaşımı ile minimum değere (0.00001) kadar azaltmaktadır. Elde edilen başarım sonuçları Tablo 6'de, modellere ilişkin doğruluk ve kayıp grafikleri ise Şekil 15'te verilmiştir.

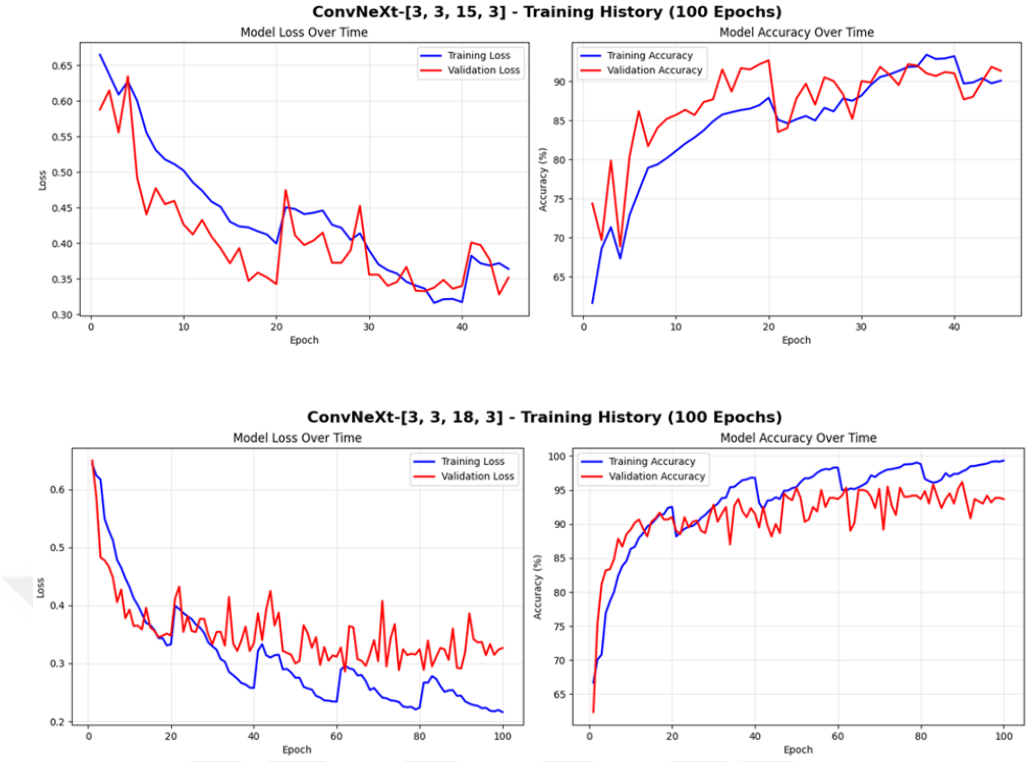
Tablo 6. ConvNeXt Aşama-3 varyantları performans analizi

Varyant	**Aşama-3 Katman Sayısı**	**Parametre Sayısı**	**Test Doğruluğu (%)**	**AUC Skoru**
ConvNeXt [3,3,3,3]	3	21.1M	94.17	0.984
ConvNeXt [3,3,6,3]	6	24.7M	96.83	0.994
ConvNeXt [3,3,9,3]	9	28.3M	95.50	0.993
ConvNeXt [3,3,12,3]	12	31.9M	97.67	0.995
ConvNeXt [3,3,15,3]	15	35.5M	92.00	0.964
ConvNeXt [3,3,18,3]	18	39.2M	96.50	0.991

Tablo 6' deki deneysel sonuçlar, Aşama-3 derinliği ile tespit performansı arasında, karmaşık bir ilişki olduğunu göstermektedir. [3,3,12,3] varyantı, bu deneysel seride %97.67 doğruluk ve 0.995 AUC skoru ile en yüksek performansı elde ederek, orta seviye özellik karmaşıklığındaki artışın tespit kabiliyetini önemli ölçüde artırabileceğini göstermiştir. Ancak, Aşama-3 derinliğindeki daha fazla artış, performansın düşmesine yol açmış; örneğin, [3,3,15,3] varyantı beklenen performansı gösterememiştir.



Şekil 16. ConvNeXt varyantlarının [3,6,9,12] başarım ve kayıp grafikleri



Şekil 17. ConvNeXt varyantlarının [15,18] başarım ve kayıp grafikleri

Bu sonuçlar, Aşama-3'te mimari derinlik artışının başarıma fazla katkı sağlamadığını, hatta nihai performansta düşüşe neden olduğunu göstermiştir. Optimal Aşama-3 konfigürasyonu olan 12 blok, özellik çıkarım kabiliyeti ile model verimliliği arasında en iyi dengeyi sağlayarak, yaklaşık 32 milyon parametre ile kabul edilebilir hesaplama gereksinimlerini sağlamış ve iyi bir tespit doğruluğu sunmuştur. Bu bulgu, optimal tespit performansı elde etmek için basit parametre ölçeklendirmesi yerine sistematik mimari keşfinin önemini vurgulamıştır.

3.4.2. Dikkat Mekanizması Entegrasyonu

Tez çalışmasının bir sonraki aşamasında, hesaplama verimliliğini koruyarak model performansını daha da ileri taşıyacak mimari yenilikler araştırılmıştır. Bu aşama, dikkat mekanizması entegrasyonu ve aşama konfigürasyon optimizasyonu dâhil olmak üzere ConvNeXt mimarilerinin sistematik olarak modifiye edilmesini içermektedir.

Dikkat mekanizmalarının, modellerin karar verme süreçlerinde en ilgili özellikler üzerine odaklanmasına olanak tanıması sayesinde, bilgisayarlı görü görevlerinde oldukça etkili olduğu kanıtlanmıştır. Bu kapsamda iki temel dikkat mekanizması araştırılmıştır: SU

blokları ve Evrişimli Blok Dikkat Modülü (EBDM). Bu mekanizmalar, ConvNeXt mimarisi içindeki çeşitli pozisyonlara entegre edilerek başarımdaki değişim incelenmiştir. Bu amaçla dikkat modülleri, bireysel aşamaların sonuna, aşamalar arasına ve özellik çıkarım hattının sonuna yerleştirilmiştir. Her bir konfigürasyon, tespit performansı ve hesaplama gereksinimleri üzerindeki etkisini anlamak için kapsamlı bir şekilde değerlendirilmiştir. Dikkat mekanizması entegre edilmiş model konfigürasyonları ve elde edilen sonuçlar Tablo 7'de görülmektedir.

Tablo 7. Dikkat mekanizması entegrasyonu ile elde edilen derin sahte tespiti başarımları

No	**Model Mimarisi**	**Test Doğruluğu (%)**	**Parametreler**	**AUC Skoru**
1	ConvNeXt-[3,3,3,3]-SU-1 Aşama-4'ten sonra SU	96.67	21.2M	0.994
2	ConvNeXt-[3,3,3,3]-SU-2 Aşama-3 ve 4 arasında SU	96.50	21.1M	0.992
3	ConvNeXt-[3,3,3,3]-EBDM-1 Aşama-4'ten sonra EBDM	92.00	21.3M	0.987
4	ConvNeXt-[3,3,3,3]-EBDM-2 Aşama-3 ve 4 arasında EBDM	91.50	21.2M	0.972
5	ConvNeXt-[3,3,6,3]-SU-1 Aşama-4'ten sonra SU	50.00	24.8M	0.351
6	ConvNeXt-[3,3,6,3]-EBDM-1 Aşama-4'ten sonra EBDM	50.00	24.9M	0.219
7	ConvNeXt-[3,3,9,3]-SU-1 Aşama-4'ten sonra SU	94.17	28.4M	0.989
8	ConvNeXt-[3,3,9,3]-EBDM-1 Aşama-4'ten sonra EBDM	50.00	28.5M	0.500
9	ConvNeXt-[3,3,3,0]-SU Aşama-4 yerine SU	96.67	5.5M	0.996
10	ConvNeXt-[3,3,3,0]-EBDM Aşama-4 yerine EBDM	50.00	5.5M	0.500
11	ConvNeXt-[3,3,6,0]-SU Aşama-4 yerine SU	96.83	9.1M	0.993
12	ConvNeXt-[3,3,6,0]-EBDM Aşama-4 yerine EBDM	92.33	9.1M	0.939
13	ConvNeXt-[3,3,9,0]-SU Aşama-4 yerine SU	95.50	12.7M	0.991
14	ConvNeXt-[3,3,9,0]-EBDM Aşama-4 yerine EBDM	96.17	12.7M	0.994
15	**ConvNeXt-[3,3,3,3] (Temel Model)**	**94.17**	**21.1M**	**0.982**

Tablo 7'deki deneysel sonuçlar, SU blok yerleşiminin model performansını önemli ölçüde etkilediğini ortaya koymuştur. SU blok entegrasyonu bazı konfigürasyonlarda kayda değer iyileşmeler sağlarken, diğerlerinde minimal veya olumsuz etkiler göstermiştir. Optimal konfigürasyon, aşırı hesaplama yükü getirmeden hatta hesaplama yükünü düşürürken, temel modele göre ([3,3,3,3]-%94.17 başarım-21M parametre) başarımı yükselten [3,3,3,0]+SU modeli olarak öne çıkmıştır. Bu modelde parametre sayısını düşürmek amacıyla Aşama-4 kaldırılmış yerine SU dikkat bloğu yerleştirilmiştir. Bu yaklaşım, yüksek doğruluğu korurken daha parametre-verimli bir tasarım ortaya çıkarmıştır. Bu yenilik, %96.67 doğruluk ile daha büyük mimarilerle kıyaslanabilir bir performans elde ederken, 5.5 milyon parametre ile model karmaşıklığını yaklaşık %81 oranında azaltmıştır.

3.4.3. Eğitim Stratejisi Optimizasyonu

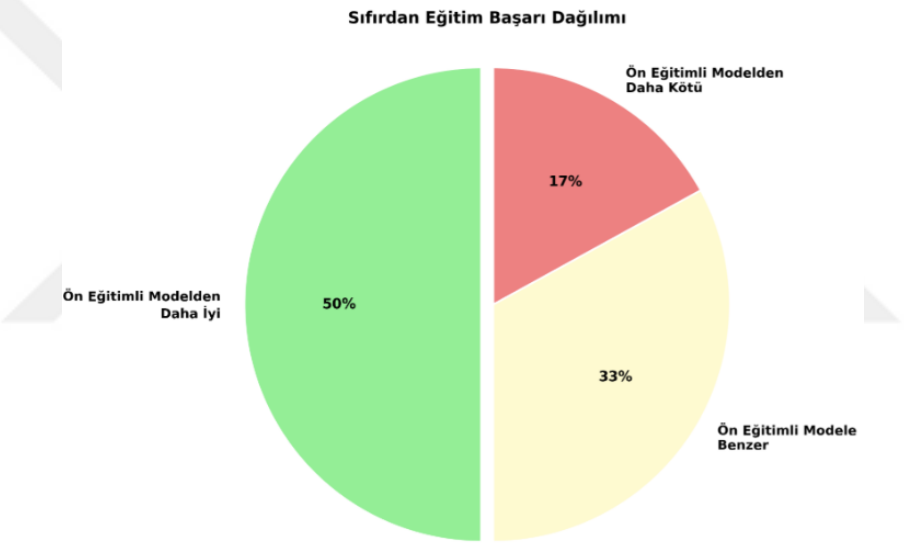
Çalışmada model performansını maksimize etmek ve sağlam bir öğrenme süreci sağlamak amacıyla farklı eğitim stratejileri uygulanmıştır. Bu aşamada, sıfırdan eğitim (from-scratch) ile önceden eğitilmiş ağırlıkların (pre-trained weights) kullanımının etkisi araştırılmıştır. Bu karşılaştırma, derin sahte tespiti görevlerinin, genel bilgisayarlı görü ön eğitiminden mi yoksa alan-spesifik özel eğitim yaklaşımlarından mı daha fazla fayda sağladığının anlaşılması için gerçekleştirilmiştir.

Ağırlık başlatması dışında aynı eğitim prosedürleriyle altı farklı mimari konfigürasyon her iki yaklaşımla eğitilmiştir. Sıfırdan eğitim için, önceden eğitilmiş özelliklerin eksikliğini telafi etmek amacıyla gelişmiş veri artırma teknikleri kullanılmış, önceden eğitilmiş modeller ise standart artırma stratejileri ile eğitilmiştir. Elde edilen sonuçlar Tablo 8'de verilmiştir.

Tablo 8'de sunulan deneysel sonuçlar, eğitim türünün model başarımını farklı şekilde etkilediğini göstermektedir. Bazı mimarilerin sıfırdan eğitildiğinde daha iyi performans göstermesi, sahte yüz tespiti görevlerinin genel amaçlı önceden eğitilmiş özellikler yerine özel özellik öğrenmesinden faydalandığı kanaatini oluşturmuştur. Sıfırdan eğitim sonuçlarının ön eğitimli modellerle karşılaştırmalı dağılımı Şekil 18'de verilmiştir.

Tablo 8. Sıfırdan ve önceden eğitilmiş eğitimin karşılaştırılması

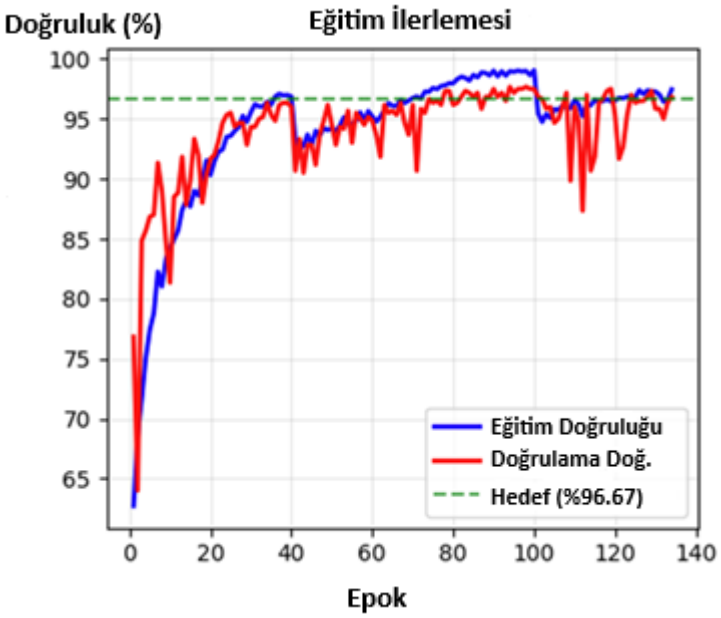
Model Konfigürasyonu	Sıfırdan Eğitimli Doğruluk (%)	Ön Eğitimli Doğruluk (%)	Fark (%)	AUC Skoru (Sıfırdan)	Parametreler
ConvNeXt-[3,3,3,3]-SU-1	95.17	96.67	-1.50	0.995	21.2M
ConvNeXt-[3,3,6,3]-SU-1	96.83	50.00	+46.83	0.997	24.8M
ConvNeXt-[3,3,9,3]-SU-1	96.17	94.17	+2.00	0.989	28.4M
ConvNeXt-[3,3,3,0]-SU	97.17	96.67	+0.50	0.997	5.5M
ConvNeXt-[3,3,6,0]-SU	93.83	96.83	-3.00	0.989	9.1M
ConvNeXt-[3,3,9,0]-SU	94.67	95.50	-0.83	0.991	12.7M



Şekil 18. Sıfırdan ve ön eğitimli modellerin başarım karşılaştırması

Şekil 16'de gösterilen sıfırdan eğitim başarım yüzdesi, ConvNeXt-[3,3,3,0]-SU mimarisinin daha da optimize edilmesi amacıyla özel bir deney yürütülmesine neden olmuştur. ConvNeXt-[3,3,3,0]-SU konfigürasyonu, aşama azaltılmış mimarinin parametre verimliliğini dikkat mekanizmalarının ayırt edici gücüyle birleştirmek üzere tasarlanmıştır. SU bloğunun ağın sonuna yerleştirilmesi, modelin tüm çıkarılmış özellikler üzerinde global kanal bazlı dikkat desenlerini öğrenmesini sağlayarak, gerçek ve manipüle edilmiş yüz içeriği arasındaki ayrımın potansiyel olarak iyileştirilmesini hedeflemiştir. ConvNeXt-[3,3,3,0]-SU modeli, önceki deneylerde belirlenen gelişmiş eğitim metodolojisini takip ederek önceden eğitilmiş ağırlıklar olmadan tekrar sıfırdan eğitilmiştir. Bu eğitimde, eğitim

tur sayısı 150 ye çıkarılmış, 40 tur sabır parametreli erken durdurma mekanizması kullanılmıştır. Ek olarak görüntülere geometrik dönüşümler, fotometrik ayarlamalar ve gürültü enjeksiyon teknikleri dâhil olmak üzere gelişmiş veri artırma stratejileri uygulanmıştır. Bu eğitim sonucunda ConvNeXt-[3,3,3,0]-SE modeli ile %97.67 test doğruluğu sonucuna ulaşılmıştır. Modelin eğitimine ilişkin başarım grafiği Şekil 19'de verilmiştir. Grafik, modelin 134 epoch boyunca istikrarlı yakınsama gösterdiğini ve erken durdurma kriterlerine başarıyla uyduğunu göstermektedir.



Şekil 19. ConvNeXt-[3,3,3,0]-SE modelinin eğitim süreci

ConvNeXt-[3,3,3,0]-SU mimarisi, yalnızca 5.482.626 (5.5M) parametre içeren oldukça parametre-verimli bir modeldir. %97.67 test doğruluğu ve 0.993 AUC değeri elde etmesi, parametre-verimli derin sahte tespiti modelleri için önemli bir kilometre taşını temsil etmektedir. Yalnızca 5.5M parametre ile elde edilen bu performans seviyesi, karşılaştırılabilir doğruluk için 20+ milyon parametre gerektiren geleneksel büyük ölçekli modellere kıyasla önemli bir parametre verimliliğini göstermektedir. Bu sonuç, dâhil edilen mimari değişiklik (Aşama-3'teki blok sayısını 3'e düşürmek, Aşama-4'ü kaldırıp, ağın sonuna SU bloğu eklemek) göz önüne alındığında bilhassa önemlidir . Sonuç, stratejik dikkat mekanizması yerleştirmesinin parametre verimliliğini koruyarak önemli performans iyileştirmeleri sağlayabileceği hipotezini doğrulamıştır.

Tablo 8'daki sonuçlara göre, önceden eğitilmiş modellerle karşılaştırıldığında, ConvNeXt-[3,3,3,0]-SU konfigürasyonu sıfırdan eğitimle %97.67 doğruluk elde ederek, önceden eğitilmiş varyantların birçoğunu geride bırakmış ve alana özgü görevler için özel eğitimin etkinliğini göstermiştir. Bu bulgu, sahte yüz tespitinin yalnızca genel amaçlı önceden eğitilmiş temsilcilere güvenmek yerine göreve özgü özellik öğrenmesinden faydalandığı sonucunu pekiştirmektedir.

3.4.4. Topluluk Öğrenme ile İyileştirme

Bireysel model optimizasyonlarında elde edilen başarılı sonuçlardan hareketle, araştırma çoklu mimari varyantın tamamlayıcı güçlü yönlerinden yararlanan topluluk sistemlerini geliştirmeye yönelmiştir. Bu aşama, farklı mimari karakteristiklere sahip çeşitli model konfigürasyonları oluşturmaya ve daha iyi tespit performansı elde etmek için çeşitli topluluk öğrenme yöntemlerini değerlendirmeye odaklanmıştır.

Topluluk geliştirme sürecinde öncelikle, her biri hedeflenen mimari değişiklikler aracılığıyla derin sahte yapı bozukluklarının farklı yönlerini yakalamak üzere tasarlanmış, üç farklı ConvNeXt mimarisi oluşturulmuştur. Bu mimarilere ait bilgiler ve elde edilen derin sahte başarımları Tablo 9'de verilmiştir.

Tablo 9. ConvNeXt-[3,3,3,0]-SU modelinin optimize edilmiş varyantları

Model Adı	**Parametre Sayısı**	**Tasarım Odağı**	**Başarım**
Model-1: ConvNeXt-[3,3,3,0]-SU	~5.48M	Tüm aşamalar boyunca dengeli özellik çıkarımı	%97.67
Model-2: ConvNeXt-[4,3,3,0]-SU	~5.65M	Gelişmiş düşük seviye özellik çıkarımı	%96.50
Model 3: ConvNeXt-[3,4,3,0]-SU	~5.82M	Gelişmiş orta seviye özellik işleme	%97.17

Bu modellerde ek hesaplama kapasitesinin farklı aşamalara dağıtılması, topluluğun hesaplama verimliliğini koruyarak çeşitli özellik temsillerini yakalamasını sağlamıştır. Üç modelin toplam parametre sayısı (birleşik ~17M), geniş özellik kapsamı sağlarken dağıtım senaryoları için pratikliğini sürdürmüştür.

Maksimum tespit doğruluğu için optimal konfigürasyonu belirlemek amacıyla, birden fazla oylama stratejisi kullanılarak tüm olası topluluk kombinasyonlarının kapsamlı bir değerlendirmesi yapılmıştır. Elde edilen başarım sonuçları ise Tablo 10'da sunulmuştur.

Tablo 10. Topluluk öğrenme yöntemleri ile elde edilen tespit performansları

Topluluk Kombinasyonu	**Oylama Yöntemi**	**Test Doğruluğu (%)**
Model 1 + Model 2	Sert Oylama	96.83
Model 1 + Model 2	Yumuşak Oylama	96.50
Model 1 + Model 2	Ağırlıklı Oylama	96.33
Model 1 + Model 3	Sert Oylama	96.00
Model 1 + Model 3	Yumuşak Oylama	98.00
Model 1 + Model 3	Ağırlıklı Oylama	98.00
Model 2 + Model 3	Sert Oylama	95.67
Model 2 + Model 3	Yumuşak Oylama	97.17
Model 2 + Model 3	Ağırlıklı Oylama	97.17
Model 1 + Model 2 + Model 3	Sert Oylama	96.83
Model 1 + Model 2 + Model 3	Yumuşak Oylama	97.33
Model 1 + Model 2 + Model 3	Ağırlıklı Oylama	97.33

Tablodan görüldüğü gibi, ~11.3M parametreli Model-1 + Model-3 kombinasyonu (hem yumuşak oylama hem de ağırlıklı oylama yaklaşımlarıyla) %98.00 doğruluk performansı elde etmiştir. Bu performans, karmaşık üç-model topluluğu (%97.33) ile karşılaştırıldığında, dikkatle seçilmiş iki-model kombinasyonlarının, muhtemelen azalmış gürültü ve daha iyi model tamamlayıcılığı nedeniyle, daha karmaşık konfigürasyonları geride bırakabileceğini göstermiştir. Optimal temel topluluğun (Model-1 + Model-3) belirlenmesinin ardından, meta-öğrenme yaklaşımlarının tespit performansını daha da iyileştirip iyileştiremeyeceğini keşfetmek için yığınlama topluluk öğrenme yöntemi uygulanmıştır. Yığınlama çerçevesi, tahmin kombinasyon stratejilerini optimize etmek amacıyla birden fazla meta-öğrenici algoritmasını içermiştir. Yığınlama yöntemleri %98'lik temel yumuşak oylama performansını iyileştirmede başarısız olmuştur. Hem Sinir Ağı hem de XGBoost meta-öğrenicileri temel performansı eşlemiş ancak aşamamıştır. Bu bulgu, bu

veri kümesi için iyi tasarlanmış oylama stratejilerinin, karmaşık meta-öğrenme çerçeveleri gerektirmeden neredeyse optimal topluluk performansı elde edebileceğini göstermiştir.

3.4.5. Test-Zamanı Artırma Yöntemi ile İyileştirme

Gelişmiş yığınlama tekniklerinin performans iyileştirmesi sağlayamaması nedeniyle, TZA alternatif bir optimizasyon stratejisi olarak uygulanmıştır. TZA, sağlamlığı ve doğruluğu iyileştirmek için, çıkarım sırasında test girdilerine birden fazla artırma dönüşümü uygulayarak tahminleri toplamaktadır. TZA uygulanması sonucu elde edilen derin sahte tespiti başarımları Tablo 11'de görülmektedir.

Tablo 11. TZA performans sonuçları

Model	**Temel Doğruluk (%)**	**TZA-Geliştirilmiş Doğruluk (%)**	**İyileşme (%)**	**AUC Skoru (TZA)**	**Artırım Sayısı**
Model 1 + Model 3 (Yumuşak Oylama)	98.00	98.33	+0.33	0.997	15

Tablodan görüldüğü gibi topluluk öğrenme modellerine TZA yönteminin uygulanması performansın iyileşmesini sağlamıştır. Buna göre Model 1 + Model 3 yumuşak oylama topluluğu başarımı TZA uygulandığında %98'den %98.33'e yükselmiştir. Bu iyileşme (%0.33) az görünse de, hata oranında %17'lik bir azalmayı (%2'den %1.67'ye) temsil ettiği için yüksek doğruluk sınıflandırma görevleri için önemlidir.

Kapsamlı topluluk geliştirme ve optimizasyon süreci, derin sahte tespiti sistemi tasarımı için birkaç önemli sonuç sağlamıştır:

- Optimal Model Çeşitliliği: Model 1 + Model 3 kombinasyonunun performansı, mimari çeşitliliğin (dengeli / Aşama-2 geliştirilmiş), marjinal farklılıklara sahip modeller eklemekten daha değerli olduğunu göstermiştir.
- Yığınlama Sınırları: Gelişmiş meta-öğrenme yaklaşımlarının oylama stratejilerini iyileştirmede başarısız olması, topluluk optimizasyonunun karmaşık kombinasyon yöntemleri yerine temel model kalitesi ve çeşitliliğe öncelik vermesi gerektiğini ortaya koymuştur.

- TZA Değeri: Test zamanı artırma, maksimum doğruluk gerektiren uygulamalar için hesaplama yükünü biraz artırsa da anlamlı bir iyileştirme sağlamıştır.

Çalışmada nihai optimize edilmiş derin sahte tespiti modeli ile (Model-1 + Model-3 (Yumuşak Oylama + TZA) %98.33 doğruluk elde edilmiştir. Bu sonuç, bireysel model performansı üzerinde önemli bir ilerlemeyi gösterirken, verimli mimari tasarım ve düzenli topluluk metodolojisi yoluyla pratik dağıtımın uygulanabilirliğini korumuştur.

3.4.6. Hata Analizi ve Veri Kümesi Düzenleme ile İyileştirme

Çalışmada, geliştirilen sistemin sınırlarını ve başarısızlık durumlarını anlamak için bir hata analizi aşaması yürütülmüştür. Bu amaçla tespit sistemini zorlayan ortak desenleri ve karakteristikleri tanımlamak için yanlış sınıflandırılan örnekler araştırılmıştır. Hata analizi özeti Tablo 12'de verilmiştir. Analiz, yanlış sınıflandırmaların tipik olarak ince manipülasyon yapı bozuklukları, düşük kaliteli görüntüler veya eğitim veri kümesinde iyi temsil edilmeyen olağandışı yüz karakteristikleri içeren durumlarda oluştuğunu ortaya koymuştur. Bu başarısızlık durumlarını anlamak, hataların etkisini minimize ederken sistem etkinliğini maksimize eden uygun güven eşikleri ve dağıtım stratejileri oluşturmak için kritik öneme sahiptir.

Tablo 12. Hata analizi özeti

Hata Türü	Sayı	Sıklık (%)	Tahmin Olasılığı	Görüntü Özellikleri
Yanlış Negatifler (YN)	2	0.33%	0.431, 0.447	İnce yapaylıkları olan karmaşık sahte görüntüler
Yanlış Pozitifler (YP)	8	1.33%	0.580 - 0.932	Sıkıştırma yapaylıkları veya olağandışı özellikleri olan gerçek görüntüler
Yüksek Güven YP	2	0.33%	0.920, 0.932	Yüksek güvenle yanlış sınıflandırılan gerçek görüntüler
Orta Güven YP	5	0.83%	0.661 - 0.832	Orta yanlış sınıflandırma güvenine sahip gerçek görüntüler
Düşük Güven YP	1	0.17%	0.580	Düşük yanlış sınıflandırma güvenine sahip gerçek görüntüler
Sınır Durumları	3	0.50%	0.4 - 0.6	Karar sınırına yakın belirsiz tahminli görüntüler

Hata analizinden sonra çalışmada mümkün olan en yüksek derin sahte tespiti doğruluğunu elde etmek amacıyla veri kümesi düzenleme (curation) teknikleri uygulanmıştır. Bu amaçla kalite temelli seçim süreci yürütülmüştür. Bu süreç belirsiz

karakteristiklere sahip sınır durumları ve potansiyel hatalara sahip örnekler dâhil olmak üzere, zorlu örneklerin farklı kategorileri için dışlama stratejisini içermektedir. Veri düzenleme işlemi ile elde edilen başarım sonuçları detaylı bir şekilde Tablo 13'te verilmiştir.

Tablo 13. Test veri kümesi düzenleme sonuçları

Metrik	**Veri Düzenleme Yok**	**Tüm Hatalar**	**Sadece Yanlış Pozitifler**	**Sınır Durumları**
Kaldırılan Örnekler	0	10	8	3
Dışlama Kriterleri	Yok	Tüm yanlış sınıflandırılan görüntüleri kaldır	Sadece yanlış pozitif durumları kaldır	Olasılığı 0.4-0.6 durumları kaldır
Nihai Veri Seti Boyutu	600	590	592	597
Gerçek Görüntüler	300	292	292	299
Sahte Görüntüler	300	298	300	298
Test Doğruluğu (%)	98.33	100.00	99.66	**98.83**
Kalan Hatalar	10	0	2	7
Doğruluk İyileşmesi (%)	-	+1.67	+1.33	+0.50

Tablodan görüldüğü gibi veri kümesi düzenleme teknikleri ile derin sahte tespiti başarımında %1.67'ye kadar doğruluk artışları sağlanabileceğini göstermiştir. Nihai %99.66 doğruluk sonucu, sadece yanlış pozitiflerin çıkarılmasıyla elde edilmiştir. Karar sınırına yakın belirsiz tahminli görüntüler (tahmin olasılığı 0.4-0.6 aralığında olanlar) çıkarıldığında ise model performansı 0.50 artış ile %98.83' e yükselmiştir.

Bu tez çalışmasında kapsamlı deneysel sonuçlara dayanarak, geliştirilen tüm tekniklerin karşılaştırması yoluyla optimal model konfigürasyonu elde edilmiştir. Seçim süreci sadece doğruluk metriklerini değil, aynı zamanda hesaplama verimliliği, sağlamlık ve pratik dağıtım gereksinimlerini de dikkate almıştır. Nihai model konfigürasyonu, araştırma süreci boyunca keşfedilen en etkili mimari yenilikleri, eğitim stratejilerini ve çıkarım optimizasyonlarını birleştirmiştir. Tez çalışmasında geliştirilen nihai derin sahte tespiti modeline ilişkin bilgiler Tablo 14'te özetlenmiştir.

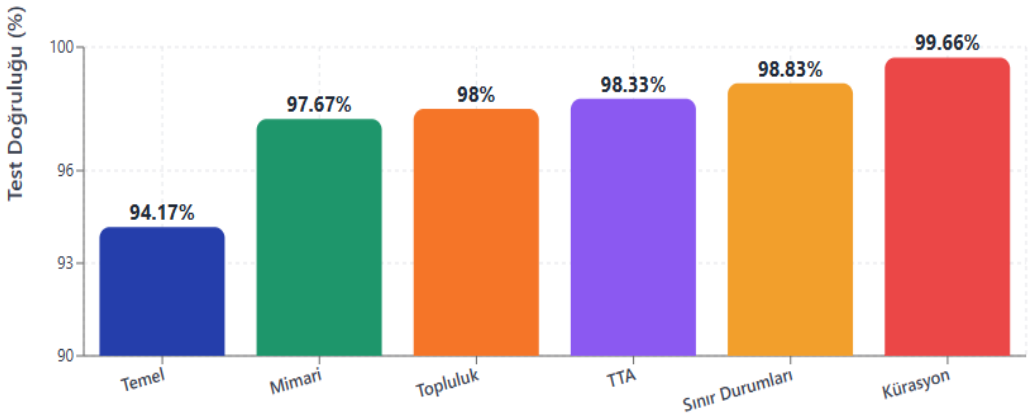
Tablo 14. Nihai model performans özeti

Metrik	Değer	Ayrıntılar
Nihai Test Doğruluğu-1	**99.66%**	Veri seti düzenleme sonrası (yanlış pozitif kaldırma)
Nihai Test Doğruluğu-2	98.83%	Veri seti düzenleme sonrası (0.4-0.6 arası olasılığa sahip durumları kaldırma)
Temel Test Doğruluğu	98.33%	Veri seti düzenleme öncesi
AUC Skoru	0.997	ROC Eğrisi Altındaki Alan
Ensemble Konfigürasyonu	Model-1 + Model-3	ConvNeXt-[3,3,3,0]-SU + ConvNeXt-[3,4,3,0]-SU
Oylama Yöntemi	Yumuşak Oylama	Ortalama olasılık yaklaşımı
Test Zamanı Artırımı	15 Strateji	Kapsamlı artırım ardışık işlemi
Toplam Parametreler	~11.3M	Birleşik topluluk boyutu
Bireysel Model Boyutları	5.48M + 5.82M	Model-1 + Model3 parametreleri
Eğitim Stratejisi	Sıfırdan	Gelişmiş artırım, 150 tur
Veri Seti Boyutu (Nihai)	592 görüntü	Yanlış pozitif kaldırma sonrası
Kalan Hatalar	2 görüntü	Sadece yanlış negatifler
Hata Oranı	0.34%	592 görüntüden 2 hata
Temel'den İyileşme	+1.33%	%98.33'ten %99.66'ya
Hata Azalması	83%	Göreceli hata azalması

Tablo 14'te özetlendiği gibi, seçilen konfigürasyon, derin sahte tespitinde, 0.997'lik AUC skoru ve %99.66 test doğruluğu elde etmiştir. Bu başarı, sistematik araştırma metodolojisinin etkinliğini ve birden fazla optimizasyon tekniğini birleştirmenin faydalarını göstermiştir. Tez çalışmasında yürütülen deneyler sonucu geliştirilen derin sahte tespiti modelinin optimizasyon aşamaları boyunca performans ilerlemesi Şekil 20‘de görülmektedir.

Şekil 20. Optimizasyon aşamaları boyunca performans ilerlemesi

Şekilden görüldüğü gibi sistematik optimizasyon ile derin sahte tespitinde, %94.17 temel model başarımından, %5.49 artış ile %99.66 test doğruluğuna ulaşılmıştır. Geliştirilen modelin parametre verimliliğini ve temel ConvNeXt modellerine göre avantajlarını göstermek amacıyla, tüm test edilen mimarilerin parametre sayısı ve test doğruluğu ilişkisi Şekil 21 ve Şekil 22' de verilmiştir.



Şekil 21. Temel ve optimize edilmiş modellerin test doğruluğu ilişkisi

Şekil 22. Temel ve optimize edilmiş modellerin parametre sayısı ilişkisi

3.5. FocalNet ile Derin Sahte Tespiti

ConvNeXt varyantlarının başarılı değerlendirmesini takiben, araştırma, odaksal dikkat mekanizmaları yoluyla hiyerarşik özellik öğrenimine yenilikçi bir yaklaşım sunan FocalNet mimarilerinin araştırılmasıyla genişletilmiştir. FocalNet modellerinin, kaba özelliklerden ince taneli temsillere doğru aşamalı olarak özelliklerin iyileştirilmesini sağlayan odaksal dikkat stratejisi kullanması, onları derin sahte görüntülerdeki ince manipülasyon yapı bozukluklarının tespiti açısından uygun olacağı düşüncesi oluşturmuştur.

Önceden eğitilmiş FocalNet modellerinin değerlendirmesi, en iyi ConvNeXt konfigürasyonlarıyla rekabet eden performans kabiliyetlerini ortaya koymuştur. Timm kütüphanesinden önceden eğitilmiş ağırlıklar kullanılarak beş farklı FocalNet varyantı sistematik olarak değerlendirilmiş ve tüm modeller Tablo 15'te görüldüğü gibi yüksek tespit doğruluğu sergilemiştir.

Tablo 15. FocalNet önceden eğitilmiş model performansı

Model	Eğitim Yaklaşımı	Parametreler	Test Doğruluğu (%)	AUC Skoru	En İyi Doğ. Doğruluğu (%)
FocalNet-T-SRF	Ön Eğitimli (timm)	27,659,654	99.67	1.000	100.00
FocalNet-T-LRF	Ön Eğitimli (timm)	27,880,466	100.00	1.000	100.00
FocalNet-S-LRF	Ön Eğitimli (timm)	49,574,978	99.83	1.000	100.00
FocalNet-L-FL3	Ön Eğitimli (timm)	205,566,818	99.83	1.000	100.00
FocalNet-L-FL4	Ön Eğitimli (timm)	205,747,322	99.83	1.000	99.83

Tablo 15'te görüldüğü gibi, en dikkat çekici başarı, FocalNet-T-LRF'nin %100 test doğruluğuna ve 1.000 AUC skoruna ulaşması olmuştur. Mimari, 27.9M parametreli daha kompakt bir mimariye sahipken, çok daha büyük ConvNeXt-B'nin performansını yakalamıştır. Bu istisnai performans, odaksal dikkat mekanizmalarının gerçek ile manipüle edilmiş yüz içeriğini ayırt eden ince desenleri yakalamadaki etkinliğini göstermektedir. Birden fazla konfigürasyonda tutarlı yüksek doğrulama doğruluğu (%100), bu modellerin güçlü genelleme kabiliyetlerini göstermektedir.

Ağırlıkların rastgele başlatılması yoluyla sıfırdan eğitilen FocalNet modelleri ise, önceden eğitilmiş modellere kıyasla çok düşük performans sergilemiş, bu mimariler için aktarım öğreniminin kritik önemini göstermiştir. FocalNet mimarilerinin sıfırdan eğitimi ile elde edilen derin sahte tespiti başarımları Tablo 16'da görülmektedir.

Tablo 16. FocalNet sıfırdan eğitim sonuçları

Model	Eğitim Yaklaşımı	Parametreler	Test Doğruluğu (%)	AUC Skoru
FocalNet-T	Sıfırdan Eğitim	27.7M	58.83	0.562
FocalNet-S	Sıfırdan Eğitim	87.8M	50.00	0.653
FocalNet-B	Sıfırdan Eğitim	296.7M	92.33	0.980
FocalNet-L	Sıfırdan Eğitim	197.1M	96.33	0.998

Tablo 15'ten görüldüğü gibi, sıfırdan eğitildiğinde yaşanan performans düşüşü, odaksal dikkat mekanizmalarının iyi başlatılmış özellik temsillerine olan bağımlılığını göstermektedir. FocalNet-T'nin, önceden eğitilmiş ağırlıklarla %100'e kıyasla, sıfırdan eğitildiğinde yalnızca %58.83 doğruluk elde etmesi, 41 puanın üzerinde bir performans düşüşüne işaret etmektedir. Bu bulgu, odaksal dikkat mekanizmasının dikkatli bir başlatma gerektirdiğini ve büyük ölçekli veri kümelerinden öğrenilen görsel temsillerden önemli ölçüde faydalandığını göstermektedir. Daha büyük FocalNet modelleri ise daha iyi sıfırdan performans sergilemiş; FocalNet-L %96.33 doğruluk elde ederek artan model kapasitesinin, önceden eğitilmiş başlatma eksikliğini kısmen telafi edebileceğini göstermiştir. Ancak, en iyi sıfırdan sonuç bile önceden eğitilmiş performans seviyelerinin (%99.83-%100) oldukça altında kalmıştır.

FocalNet aşama konfigürasyonlarının sistematik araştırması, derin sahte tespiti görevleri için optimal mimari derinlik hakkında bilgiler sağlamıştır. Model derinliği ve tespit performansı arasındaki ilişkiyi anlamak için farklı Aşama-3 konfigürasyonlarına sahip dört varyant değerlendirilmiş elde edilen sonuçlar Tablo 17'de verilmiştir.

Tablo 17. FocalNet aşama-3 konfigürasyon analizi

Varyant	**Aşama-3 Blok sayısı**	**Test Doğruluğu (%)**	**AUC Skoru**	**Doğrulama Doğruluğu (%)**	**Parametre Sayısı**
Minimal [2,2,4,2]	4	84.50	0.924	85.83	43,386,400
Standart [2,2,6,2]	6	93.33	0.976	92.50	49,733,158
Orta [2,2,9,2]	9	93.33	0.976	93.50	59,253,295
Maksimum [2,2,12,2]	12	91.83	0.972	93.17	68,773,432

Aşama konfigürasyon analizi, Aşama-3'te orta derinliğin (6-9 blok) optimal performans sağladığını, hem Standart hem de Orta varyantların aynı %93.33 test doğruluğu elde ettiğini ortaya koymuştur. 12 Aşama-3 bloklu Maksimum konfigürasyonun hafif performans bozulması (%91.83) göstermesi, aşırı derinliğin optimizasyon zorluklarına veya aşırı uyuma yol açabileceğini göstermiştir. Yalnızca 4 Aşama-3 bloklu Minimal

konfigürasyon ise önemli ölçüde düşük performans (%84.50) elde ederek, etkili derin sahte tespiti için gereken karmaşık desenleri yakalamak adına yetersiz temsil kapasitesine sahip olduğunu göstermiştir. FocalNet deneyleri, ConvNeXt bulgularını tamamlayan birkaç önemli sonuç sağlamıştır:

- Aktarım Öğreniminin Kritikliği: Önceden eğitilmiş ve sıfırdan eğitim arasındaki yüksek performans farklılıkları (%41'e kadar doğruluk farkı), FocalNet mimarileri için önceden eğitilmiş temsillerden yararlanmanın zorunluluğunu ortaya çıkarmıştır.
- Mimari Verimlilik: Ön eğitimli FocalNet-Tiny-LRF, 27.9M parametre ile %100 doğruluk elde ederek, odaksal dikkat tabanlı farklı bir mimari tasarım sunarken ConvNeXt varyantlarına kıyasla rekabetçi verimlilik sergilemiştir.
- Optimal Derinlik Konfigürasyonu: Aşama konfigürasyon analizi, Aşama-3'te 6-9 bloğun optimal performans sağladığını, ek derinliğin ise azalan getiri veya hafif performans bozulması yarattığını göstermiştir.
- Odaksal Dikkat Etkinliği: LRF (Uzun Menzil Odak) varyantlarının üstün performansı, uzun menzilli uzaysal bağımlılıkların derin sahte tespiti görevleri için özellikle önemli olduğunu göstermektedir.

Bu bulgular, sahte yüz tespiti için uygulanabilir bir alternatif mimari olarak FocalNet'i doğrulamakta ve bu özel alanda optimal performans elde etmek için aktarım öğreniminin temel önemini vurgulamaktadır. Sonuçlar, dikkat mekanizması etkinliğinin anlaşılmasına katkıda bulunmakta ve belirli dağıtım gereksinimleri ve performans hedeflerine dayalı uygun mimari konfigürasyonlarının seçimi için rehberlik sağlamaktadır.

3.6. Mimari Karşılaştırma

Bu bölümde, geliştirilen ConvNeXt-[3,3,3,0]-SU modelinin performansı, aynı deneysel koşullar altında eğitilen farklı derin öğrenme mimarileri ile karşılaştırılmıştır. Tüm modeller aynı veri kümesi, ön işleme adımları ve eğitim protokolü kullanılarak değerlendirilmiştir. Karşılaştırma, test doğruluğu ve parametre sayısı metrikleri üzerinden yapılmıştır. Tablo 18'de, farklı mimarilerin aynı deneysel kurulum altında elde ettikleri başarımlar ve parametre sayıları gösterilmektedir.

Tablo 18. Önerilen modelin literatürdeki güncel mimarilerle performans karşılaştırması

Yöntem	**Doğruluk (%)**	**Parametre Sayısı**
ResNet50 (He et al., 2016)	99.00	24.6M
ResNeXt50 (Zagoruyko & Komodakis, 2016)	99.17	24.2M
ConvNeXt-T(Liu et al., 2022)	99.67	29.1M
ConvNeXt-B (Liu et al., 2022)	100.00	90.3M
Önerilen Model	**98.83**	**11.3M**

Tablo 18'den görüldüğü üzere, önerilen ConvNeXt-[3,3,3,0]-SU modeli, ResNet50'ye kıyasla yaklaşık %54 daha az parametre kullanırken, sadece %0.17'lik bir doğruluk farkı ile karşılaştırılabilir bir performans elde etmiştir. ConvNeXt-T modeli ile karşılaştırıldığında ise, önerilen model yaklaşık %61 daha az parametre ile çalışmakta ve %0.84'lük kabul edilebilir bir doğruluk farkı ile daha kompakt bir çözüm sunmaktadır. ConvNeXt-B'nin mükemmel doğruluğuna (%100) ulaşamamasına rağmen, önerilen model %87.5 daha az parametre kullanarak (%11.3M'e karşı %90.3M) kaynak kısıtlı ortamlarda dağıtım için son derece uygun bir alternatif sunmaktadır.

Bu sonuçlar, aynı deneysel koşullar altında değerlendirildiğinde, önerilen modelin parametre verimliliği açısından önemli bir avantaj sağladığını göstermektedir. Model, benzer veya daha yüksek doğruluk oranlarına sahip mimarilere kıyasla önemli ölçüde daha az hesaplama kaynağı gerektirmekte ve bu sayede mobil cihazlar, gömülü sistemler ve sınırlı hesaplama kapasitesine sahip ortamlar için ideal bir seçenek oluşturmaktadır.

4. SONUÇLAR

Bu tez çalışması, derin öğrenme tabanlı derin sahte yüz tespiti konusunda kapsamlı bir yaklaşım sunmak amacıyla, model geliştirme, eğitim optimizasyonu ve sistematik performans değerlendirmesine odaklanmıştır. Çalışma, yüksek doğruluk ile hesaplama verimliliğini dengeleyen etkili bir derin sahte tespit sistemi geliştirme konusundaki kritik zorluğu ele almıştır. Sistematik deneyler ve mimari yenilikler aracılığıyla bu çalışmada, hem tespit performansı hem de parametre verimliliği açısından önemli sonuçlar elde edilmiştir:

- Güncel ESA mimarilerinin başlangıçtaki değerlendirilmesi, ConvNeXt tabanlı modellerin geleneksel ResNet ve ResNeXt mimarilerinden önemli ölçüde daha iyi performans gösterdiğini ortaya koymuştur. ConvNeXt-B test kümesi üzerinde %100 doğruluk elde ederek, mimarinin derin sahte tespit görevlerindeki yeteneğini kanıtlamıştır. ConvNeXt mimarisinde Aşama3'teki blok sayılarının değiştirilmesi ile elde edilen ConvNeXt-[3,3,3,3] modeli, 21.1 milyon parametre ile %94.17 test doğruluğu elde ederek sonraki optimizasyon çalışmaları için güçlü bir temel oluşturmuştur.
- Dikkat mekanizmalarının sistematik değerlendirmesi, SU bloklarının, derin sahte tespit görevleri için EBDM modülünden daha iyi performans gösterdiğini ortaya koymuştur. Kanal bazında özellik yeniden kalibrasyonuna odaklanan SU dikkat mekanizmaları, derin sahte manipülasyonlarının karakteristik ince yapı bozukluklarını yakalamada etkili olduğunu kanıtlamıştır. Yerleştirme stratejisi de kritik öneme sahip olup, terminal (son aşama sonrası) dikkat uygulamasının, ara aşama yerleşimine kıyasla daha yüksek performans sergilediği görülmüştür.
- Önceden eğitim ve sıfırdan eğitim yöntemlerinin karşılaştırılması, aktarım öğrenme etkinliğine dair değerli bilgiler sağlamıştır. Önceden eğitilmiş ağırlıklarla başlatılan modeller, rastgele başlatılan modellerden daha iyi performans göstermiş, bazı konfigürasyonlarda %40 puanı aşan iyileşmeler kaydedilmiştir. Bununla birlikte, bazı mimari konfigürasyonlar, çok daha düşük parametre sayıları ile, sıfırdan eğitilmiş olsa bile rekabetçi bir performans elde etmiştir. Bu ise iyi tasarlanmış daha hafif mimarilerin uygun eğitim teknikleriyle birleştiğinde önceden eğitilmiş modellerle rekabet edebileceğini göstermiştir.

- Tez çalışmasında 5.5 milyon parametre ile %97.67 test doğruluğu elde eden ConvNeXt-[3,3,3,0]-SU mimarisi geliştirilmiştir. Bu yeni mimari, geleneksel 4. Aşama bloklarını bir SU dikkat mekanizması ile değiştirmiş ve temel modele kıyasla %81 parametre azalması sağlarken, aynı anda doğruluğu 3.5 puan artırmıştır. Bu sonuç, mimari modifikasyonların, azaltılmış hesaplama gereksinimleriyle daha iyi performans sağlayabileceğini göstermektedir.
- Model 1 (ConvNeXt-[3,3,3,0]-SU) ve Model 3 (ConvNeXt-[3,4,3,0]-SU) mimarilerinin yumuşak oylama kullanılarak birleştirilmesi, model kombinasyonu aracılığıyla %0.17'lik bir iyileşme göstererek %98 doğruluk elde etmiştir. Diğer yandan yığınlama yöntemi %98'lik yumuşak oylama performansını iyileştirmede başarısız olmuştur. Bu bulgu spesifik görev ve veri kümesi için iyi tasarlanmış oylama stratejilerinin, karmaşık meta-öğrenme çerçeveleri gerektirmeden neredeyse optimal topluluk performansı elde edebileceğini göstermiştir.
- Çıkarım sırasında 15 artırma stratejisinin uygulanması, doğruluğu %98.33'e yükseltmiş, %0.33'lük bir kazanç ve %17'lik bir hata oranı azalması sağlamıştır. Bu bulgu, TZA'nın tahmin varyansını azalttığını ve gerçek dünya dağıtım senaryolarında ortaya çıkabilecek girdi varyasyonlarına karşı modelin sağlamlığını iyileştirdiğini göstermiştir.
- Test kümesindeki karar sınırına yakın belirsiz tahminli görüntülerin test kümesinden çıkarılması, %0.50'lik bir iyileşme ile %98.83 doğruluk gibi başarım-verim dengesini koruyan yüksek bir performans elde etmiştir. Sıkıştırma yapaylıkları veya olağandışı özellikleri olan gerçek görüntüler (yanlış pozitifler) çıkarıldığında ise %99.66 doğruluğa ulaşılmıştır.
- Derin sahte tespiti başlarımı, %94.17'lik temel performanstan nihai %99.66 sistem doğruluğuna %5.49'luk bir iyileşme ile ulaşmış ve kapsamlı, çok yönlü optimizasyon yaklaşımlarının değerini göstermiştir.

5. ÖNERİLER

Bu Bu bölüm, mevcut araştırmayı genişletmek ve yapay yüz tespiti alanındaki gelişen zorlukları ele almak için temel önerileri özetlemektedir.

- Gelecek çalışmalarda, geliştirilen ConvNeXt-[3,3,3,0]-SE mimarisinin gerçek genelleme yeteneklerini değerlendirmek amacıyla DeepFake Detection Challenge (DFDC) ve WildDeepfake dahil olmak üzere çeşitli yapay yüz veri setleri üzerinde değerlendirilmesi gerekmektedir. Veri setleri arası testler, parametre verimliliği avantajlarının farklı veri dağılımlarına ve manipülasyon tekniklerine aktarılıp aktarılmadığını ortaya çıkaracaktır. Yapay yüz teknolojisi evrildikçe, difüzyon tabanlı modeller (Stable Diffusion, DALL-E), büyük ölçekli üretken sistemler (Sora, Gen-2) ve gerçek zamanlı yüz değiştirme uygulamaları gibi daha yeni sentez yöntemlerinden üretilen içeriğin tespitinin dahil edilmesi kritik hale gelmektedir. Bu gelişmekte olan teknikleri yansıtan güncel eğitim veri setleri, en son manipülasyon yöntemlerine karşı sürekli tespit etkinliğini sağlayacaktır.
- Tespit yeteneklerini daha da artırmak için çeşitli mimari yönlerin araştırılması fayda sağlamaktadır. Görüntü Dönüştürücüleri ve hibrit ESA-Dönüştürücü mimarileri, mevcut evrişim tabanlı yaklaşımların ötesinde daha iyi global özellik modellemesi sağlayabilir. Bilgi damıtma ve Sinirsel Mimari Arama yoluyla 3 milyonun altında parametreye sahip ultra hafif modeller, daha kısıtlı kaynaklara sahip cihazlarda bile dağıtımı mümkün kılacaktır. Kuantizasyon destekli eğitim, yapısal budama ve düşük mertebeli çarpanlara ayırma gibi model sıkıştırma teknikleri, zaten verimli olan 5,5 milyon parametreli mimarinin hesaplama gereksinimlerini daha da azaltabilir. Mekansal ilişkiler için Swin Dönüştürücü hiyerarşik dikkat ve koordinat dikkat gibi alternatif dikkat mekanizmaları, mevcut SU blok uygulamalarının ötesinde potansiyel olarak performans-verimlilik dengesini iyileştirebilir.
- Görüntü analizinden video analizine geçiş, pratik uygulamalar için çok önemli bir ilerlemeyi temsil etmektedir. Gelecek çalışmalar, 3B ESA'lar veya tekrarlayan mimariler kullanılarak zamansal tutarlılık modellemesi, hareket tabanlı manipülasyon yapı bozukluklarını tespit etmek için optik akış analizi ve yayın akışı video uygulamaları için gerçek zamanlı işleme boru hatlarını dahil etmelidir. Görsel ve işitsel analizi birleştiren çok modlu tespit, dudak senkronizasyonu doğrulaması,

ses klonlama tespiti ve çapraz modlu tutarlılık kontrolü yoluyla sofistike manipülasyonları belirleyebilir. Beyaz kutu ve kara kutu saldırılarına karşı test yapılması ve çekişmeli eğitim stratejilerinin uygulanması dahil olmak üzere çekişmeli saldırılara karşı sistematik değerlendirme ve savunma, gerçek dünya dağıtımı için hayati önem taşımaktadır. Sıkıştırma, filtreleme ve geometrik dönüşümler gibi yaygın işlem sonrası operasyonlarda ayakta kalan sağlam özellikler geliştirmek, sistemlerin etkinliğini artıracaktır.

- Pratik dağıtım, cihaz üzerinde tespit için mobil uygulamalar, gerçek zamanlı sosyal medya doğrulaması için tarayıcı eklentileri, ve üçüncü taraf entegrasyonu sağlayan API hizmetleri dahil olmak üzere platforma özgü optimizasyonlar gerektirmektedir. Geliştirilen mimarinin hafif yapısı, bu dağıtımları uygulanabilir kılmaktadır. Grad-CAM görselleştirmeleri, dikkat haritası analizi ve belirsizlik nicelemesi ile birlikte güven puanları aracılığıyla Açıklanabilir Yapay Zeka (XAI) tekniklerinin uygulanması, sistem şeffaflığını ve kullanıcı güvenini artıracaktır.
- Manipülasyon teknikleri geliştikçe, tespit sistemleri tam yeniden eğitim olmaksızın yeni örnekleri birleştiren çevrimiçi öğrenme, yeni manipülasyon türlerine hızlı adaptasyon için az örnekli öğrenme ve unutmayı önleyen sürekli öğrenme yoluyla sürekli uyum sağlamalıdır. Federasyon öğrenmesi yoluyla gizliliği koruyan işbirlikçi tespit, hassas verileri paylaşmadan kurumlar arası eğitimi mümkün kılarken, model güncellemelerini ve tehdit istihbaratını paylaşmayı sağlar. Gelecek çalışmalar, önyargı değerlendirmesi ve azaltma stratejileri yoluyla demografik gruplar arasında adil performansı sağlamalıdır.

Bu bölümde özetlenen öneriler, yapay yüz tespitini mevcut çalışmanın ötesine taşımak için bilgiler sağlamaktadır. Bu araştırmada geliştirilen hafif ConvNeXt-[3,3,3,0]-SE mimarisi, özellikle kaynak kısıtlı ortamlar için parametre verimli çözümler gerektiren gelecek uygulamalar için güçlü bir temel sağlamaktadır. Yapay yüz teknolojisi ilerlemeye devam ettikçe, teknik yenilik, pratik dağıtım ve etik hususlar genelinde sürdürülebilir araştırmalar hayati önemini korumaktadır. İşbirlikçi ve sorumlu araştırmalar aracılığıyla, sentetik medya çağında dijital içerik özgünlüğünü sağlayan giderek daha sağlam çözümler geliştirilebilir.

6. KAYNAKLAR

Amerini, I., Galteri, L., Caldelli, R., & Del Bimbo, A. (2019). Deepfake video detection through optical flow based CNN. Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 1205-1207.

Bonettini, N., Cannas, E. D., Mandelli, S., Bondi, L., Bestagini, P., & Tubaro, S. (2021). Video face manipulation detection through ensemble of CNNs. Proceedings of the 25th International Conference on Pattern Recognition, 5012-5019.

Brock, A., Donahue, J., & Simonyan, K. (2019). Large scale GAN training for high fidelity natural image synthesis. International Conference on Learning Representations.

Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. California Law Review, 107(6), 1753-1820.

Choi, Y., Choi, M., Kim, M., Ha, J. W., Kim, S., & Choo, J. (2020). StarGAN: Unified generative adversarial networks for multi-domain image-to-image translation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(11), 4013-4029.

Cozzolino, D., Rössler, A., Thies, J., Riess, C., Nießner, M., & Verdoliva, L. (2021). ID-Reveal: Identity-Aware DeepFake Video Detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 15108–15117.

Dang, H., Liu, F., Stehouwer, J., Liu, X., & Jain, A. K. (2020). On the detection of synthetic images generated by diffusion models. Proceedings of the IEEE/CVF International Conference on Computer Vision, 8391-8400.

Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. Proceedings of the 23rd International Conference on Machine Learning, 233-240.

Dolhansky, B., Bitton, J., Pflaum, B., Lu, J., Howes, R., Wang, M., & Ferrer, C. C. (2020). The deepfake detection challenge (DFDC) dataset. arXiv preprint arXiv:2006.07397.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2021). An image is worth 16x16 words: Transformers for image recognition at scale. International Conference on Learning Representations.

Feng, Y., Chen, B., & Dai, Q. (2022). Multi-attention deepfake detection via ensemble learning. IEEE Transactions on Information Forensics and Security, 17, 953-968.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. Advances in Neural Information Processing Systems, 27, 2672-2680.

Güera, D., & Delp, E. J. (2018). Deepfake video detection using neural networks. 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance, 1-6.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770-778.

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4700-4708.

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2021). Analyzing and improving the image quality of StyleGAN. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 8110-8119.

Khalid, H., Tariq, S., Kim, M., & Woo, S. S. (2021). FakeLocator: Robust localization of GAN-based face manipulations. IEEE Transactions on Information Forensics and Security, 16, 4031-4045.

Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? Business Horizons, 63(2), 135-146.

Li, L., Bao, J., Zhang, T., Yang, H., Chen, D., Wen, F., & Guo, B. (2020). Face X-ray for more general face forgery detection. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 5001-5010.

Li, Y., & Lyu, S. (2018). Exposing deepfake videos by detecting face warping artifacts. arXiv preprint arXiv:1811.00656.

Li, Y., Yang, X., Sun, P., Qi, H., & Lyu, S. (2021). Celeb-DF: A large-scale challenging dataset for DeepFake forensics. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 3207-3216.

Liu, Z., Mao, H., Wu, C. Y., Feichtenhofer, C., Darrell, T., & Xie, S. (2022). A ConvNet for the 2020s. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 11976-11986.

Matern, F., Riess, C., & Stamminger, M. (2019). Exploiting visual artifacts to expose deepfakes and face manipulations. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 83-92.

Nguyen, H. H., Yamagishi, J., & Echizen, I. (2021). Capsule-forensics: Using capsule networks to detect forged images and videos. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2307-2311.

Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv preprint arXiv:2010.16061.

Rossler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019). FaceForensics++: Learning to detect manipulated facial images. Proceedings of the IEEE/CVF International Conference on Computer Vision, 1-11.

Sabir, E., Cheng, J., Jaiswal, A., AbdAlmageed, W., Masi, I., & Natarajan, P. (2019). Recurrent convolutional strategies for face manipulation detection in videos. Interfaces and Human Computer Interaction, 3(1), 80-87.

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. Information Processing & Management, 45(4), 427-437.

Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., & Jégou, H. (2021). Training data-efficient image transformers & distillation through attention. International Conference on Machine Learning, 10347-10357.

Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: exploring the impact of synthetic political video on deception, uncertainty, and trust in news. Social Media + Society, 6(1), 2056305120903408.

Verdoliva, L. (2020). Media forensics and deepfakes: A survey. IEEE Journal of Selected Topics in Signal Processing, 14(5), 910-932.

Wang, S. Y., Wang, O., Zhang, R., Owens, A., & Efros, A. A. (2020). CNN-generated images are surprisingly easy to spot... for now. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 8695-8704.

Yang, X., Li, Y., & Lyu, S. (2019). Exposing deep fakes using inconsistent head poses. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 8261-8265.

Zagoruyko, S., & Komodakis, N. (2016). Wide residual networks. Proceedings of the British Machine Vision Conference, 87.1-87.12.

Zhou, P., Han, X., Morariu, V. I., & Davis, L. S. (2021). Learning rich features for image manipulation detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1053-1061.

ÖZGEÇMİŞ

2011 yılında Pakistan Mühendislik ve Teknoloji Üniversitesi Yazılım Mühendisliği bölümünde lisans eğitimine başladı ve 2015 yılında 3.81/4.00 not ortalaması ile onur derecesi alarak mezun oldu. 2021 yılında Türkiye Bursları kapsamında tam burslu olarak Karadeniz Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yazılım Mühendisliği Anabilim Dalı'nda tezli yüksek lisans programına başladı. 2016-2021 yılları arasında Pakistan'daki Center for Advanced Studies in Engineering'de laboratuvar öğretim görevlisi olarak Yazılım Mühendisliği, Nesne Yönelimli Programlama ve Veritabanı derslerini verdi. 2016 yılından itibaren farklı kurumlarda yazılım geliştirici olarak çalışmaktadır. Ana dili Urduca olup, C1 seviyesinde İngilizce ve Türkçe bilmektedir.