T.C.
GÜMÜŞHANE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YAPAY ZEKÂ VE AKILLI SİSTEMLER ANA BİLİM DALI

EVRİŞİMLİ SİNİR AĞLARI TABANLI SÜPER ÇÖZÜNÜRLÜK YÖNTEMLERİ İLE AVUÇ İZİ GÖRÜNTÜ KALİTESİNİN İYİLEŞTİRİLMESİ

YÜKSEK LİSANS

Hüseyin Furkan MACAN

HAZİRAN - 2025
GÜMÜŞHANE

T.C.
GÜMÜŞHANE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YAPAY ZEKÂ VE AKILLI SİSTEMLER ANA BİLİM DALI

EVRİŞİMLİ SİNİR AĞLARI TABANLI SÜPER ÇÖZÜNÜRLÜK YÖNTEMLERİ İLE AVUÇ İZİ GÖRÜNTÜ KALİTESİNİN İYİLEŞTİRİLMESİ

ENHANCING PALMPRINT IMAGE QUALITY USING CONVOLUTIONAL NEURAL NETWORK-BASED SUPER-RESOLUTION METHODS

YÜKSEK LİSANS

Hüseyin Furkan MACAN

HAZİRAN - 2025
GÜMÜŞHANE

T.C.
GÜMÜŞHANE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YAPAY ZEKÂ VE AKILLI SİSTEMLER ANA BİLİM DALI

EVRİŞİMLİ SİNİR AĞLARI TABANLI SÜPER ÇÖZÜNÜRLÜK YÖNTEMLERİ İLE AVUÇ İZİ GÖRÜNTÜ KALİTESİNİN İYİLEŞTİRİLMESİ

ENHANCING PALMPRINT IMAGE QUALITY USING CONVOLUTIONAL NEURAL NETWORK-BASED SUPER-RESOLUTION METHODS

YÜKSEK LİSANS

Hüseyin Furkan MACAN

Danışman: Dr. Öğr. Üyesi Özkan BİNGÖL

HAZİRAN - 2025
GÜMÜŞHANE

BİLİMSEL ETİĞE UYGUNLUK BEYANI

Yüksek Lisans Tezi olarak hazırlamış olduğum "**Evrişimli Sinir Ağları Tabanlı Süper Çözünürlük Yöntemleri İle Avuç İzi Görüntü Kalitesinin İyileştirilmesi**" isimli bu tezimin, tamamen kendi çalışmam olduğunu, her alıntıya kaynak gösterdiğimi, alıntı yaptığım tüm çalışmaları kaynakçada belirtiğimi ve Gümüşhane Üniversitesi'nin lisanslı kullanıcısı olduğu intihal yazılım programı ile Lisansüstü Eğitim Enstitüsü'nün belirlediği kıstaslara uygun olarak raporladığımı taahhüt ederim. Tezimin kâğıt ve elektronik kopyalarının Gümüşhane Üniversitesi Lisansüstü Eğitim Enstitüsü arşivinde saklanmasına izin verdiğimi onaylarım.

Lisansüstü Eğitim ve Öğretim Yönetmeliği'nin ilgili maddeleri uyarınca gereğinin yapılmasını arz ederim.

12/06/2025

…….

Hüseyin Furkan MACAN

TEŞEKKÜR

"Evrişimli Sinir Ağları Tabanlı Süper Çözünürlük Yöntemleri İle Avuç İzi Görüntü Kalitesinin İyileştirilmesi" adlı çalışmam süresince yapmış olduğum çalışmalarımı denetleyen, bilgi ve tecrübelerini benimle paylaşan tez danışmanım Dr. Öğr. Üyesi Özkan BİNGÖL'e teşekkür ederim. Bu çalışma, TÜBİTAK 1001 - Bilimsel ve Teknolojik Araştırma Projelerini Destekleme Programı tarafından desteklenen 122E402 nolu "Kısıtlamasız Ortamda Alınan El Görüntülerinden Biyometrik Doğrulama İçin Güvenilir Avuç İzi Bölgesinin Tespiti" isimli proje kapsamında gerçekleştirilmiştir. Bu desteklerinden dolayı TÜBİTAK'a teşekkürlerimi sunarım. Tez çalışması boyunca beni maddi ve manevi olarak destekleyen aileme ve bölüm hocalarıma da bilgilendirme ve desteklerinden dolayı teşekkürlerimi bir borç bilirim.

Hüseyin Furkan MACAN

GÜMÜŞHANE – 2025

ÖZET

Bu tez çalışması, avuç izi tanıma sistemlerinde kullanılan biyometrik görüntülerin kalitesini iyileştirmek amacıyla, derin öğrenme tabanlı süper-çözünürlük yöntemlerinin performansını incelemektedir. Literatürde yaygın olarak kullanılan GAN, DCGAN, BGAN, BEGAN ve özellikle çekişmeli otokodlayıcı (AAE) mimarileri değerlendirilmiş; elde edilen süper-çözünürlüklü çıktılar hem PSNR gibi referanslı hem de BRISQUE gibi referanssız metriklerle analiz edilmiştir.

Yapılan deneylerde, AAE modeli en başarılı sonuçları verdiğinden, geliştirme süreci bu yapı üzerine odaklanmıştır. Modelin kod çözücü kısmı konvolüsyonel katmanlarla yeniden tasarlanarak 2 kat çözünürlük artırımı sağlayacak şekilde optimize edilmiş, IIT Delhi avuç içi veri seti ile eğitilmiştir. Görüntü kalitesi iyileştirme sürecinin yalnızca görsel kaliteyi değil, aynı zamanda biyometrik tanıma doğruluğunu da artırdığı yapılan testlerle ortaya konmuştur.

Anahtar Kelimeler: Avuç İzi Tanıma, Çekişmeli Otokodlayıcı, Derin Öğrenme, Görüntü Kalitesi İyileştirme, Süper-çözünürlük

SUMMARY

This thesis investigates the performance of deep learning-based super-resolution methods to enhance the quality of biometric images used in palmprint recognition systems. Widely used architectures in the literature—such as GAN, DCGAN, BGAN, BEGAN, and particularly the Adversarial Autoencoder (AAE)—were evaluated. The resulting super-resolved images were analyzed using both reference-based metrics like PSNR and no-reference metrics such as BRISQUE. Among these models, the AAE architecture achieved the best results; thus, the development process focused on this structure. The decoder component of the model was redesigned with convolutional layers and optimized to achieve 2× resolution enhancement. The model was trained on the IIT Delhi palmprint dataset. Experimental results demonstrated that the image quality enhancement process not only improved visual quality but also significantly increased biometric recognition accuracy.

Keywords: Palmprint Recognition , Adversarial Autoencoder, Deep Learning, Image Quality Enhancement ,Super Resolution

İÇİNDEKİLER

TABLOLAR DİZİNİ

ŞEKİLLER DİZİNİ

SİMGELER VE KISALTMALAR DİZİNİ

AAE : Adversial Autoencoder- Çekişmeli Otokodlayıcı

AE : Autoencoder-Otokodlayıcı

BRISQUE : Blind/Referenceless Image Spatial Quality Evaluator -Kör/Referanssız Görüntü Mekansal Kalite Değerlendiricisi

CNN : Convolutional Neural Networks- Evrişimsel Sinir Ağları

FSRCNN : Fast Super-Resolution Convolutional Neural Network – Hızlı Süper-Çözünürlük Evrişimsel Sinir Ağı

GAN : Generative Adversial Network- Üretken Çekişmeli Ağ

GPU : Graphics Processing Unit-Grafik İşlem Birimi

HR : High Resolution- Yüksek Çözünürlük

LR : Low Resolution- Düşük Çözünürlük

MISR : Multi-Image Super-Resolution- Çoklu Görüntü Süper-Çözünürlüğü

PSNR : Peak Signal-to-Noise Ratio- Tepe Sinyal-Gürültü Oranı

SR : Super Resolution- Süper Çözünürlük

SRCNN : Super-Resolution Using Deep Convolutional Neural Network

SISR : Single-Image Super-Resolution- Tekil Görüntü Süper Çözünürlüğü

1. GENEL BİLGİLER

1.1. Süper Çözünürlük Kavramı

Süper Çözünürlük (Super-Resolution, SR), görüntü işleme ve bilgisayarla görme alanlarında son yıllarda giderek artan bir öneme sahip olan bir araştırma alanıdır. Bu teknoloji, düşük çözünürlüklü (Low-Resolution, LR) bir görüntüden daha yüksek çözünürlüklü (High-Resolution, HR) ve görsel açıdan daha zengin bir görüntü üretmeyi amaçlamaktadır. Düşük çözünürlüklü görüntüler, genellikle bulanıklık, detay kaybı ve düşük netlik gibi kalite sorunları içerir. Bu durum, özellikle kritik görsel bilgilerin yorumlanmasını zorlaştırmakta ve görüntü tabanlı karar destek sistemlerinde güvenilirliği azaltmaktadır.

SR algoritmaları, bu tür kalite problemlerini telafi ederek, daha keskin, detay bakımından zengin ve yapısal bütünlüğü daha iyi korunmuş görüntüler elde edilmesini sağlar (Şekil 1). Bu süreçte temel amaç, görüntüde kaybolmuş olan yüksek frekans bileşenleri geri kazandırmak ve orijinal görüntünün ayrıntı seviyesini artırmaktır. Özellikle tekil görüntü süper çözünürlüğü (Single Image Super-Resolution, SISR), yalnızca bir adet LR görüntü kullanarak HR görüntü üretimini hedeflemesi açısından birçok alanda uygulama kolaylığı sunar.

Şekil 1. Süper Çözünürlük örneği

SR teknolojisinin temel amacı yalnızca görüntülerin piksel boyutlarını artırmakla sınırlı değildir; daha da ötesinde, görsel içeriğin kalitesini iyileştirerek görüntüleri daha anlamlı, yorumlanabilir ve çeşitli görevler için daha işlevsel hale getirmektir. Yüksek çözünürlüklü görüntüler, özellikle ayrıntı düzeyinin kritik olduğu analizlerde –örneğin biyometrik doğrulama, tıbbi görüntüleme, adli bilişim veya uzaktan algılama gibi– daha güvenilir bilgi sunabilmektedir. Bu nedenle SR teknikleri, düşük kaliteli görüntülerin bilgi potansiyelini artırmak adına stratejik bir dönüştürme aracı olarak değerlendirilmektedir.

Düşük kaliteli verilerin yüksek doğrulukla analiz edilmesinin beklendiği alanlarda, SR algoritmaları görüntü kalitesine ilişkin sınırlılıkları büyük ölçüde aşarak, veri analiz performansını doğrudan etkilemektedir. Düşük çözünürlüklü bir görüntünün, detay yoksunluğu nedeniyle analiz sistemleri tarafından yanlış sınıflandırılması veya yetersiz yorumlanması olasılığı oldukça yüksektir. Bu noktada SR algoritmaları, kaybolmuş veya bastırılmış detayları yeniden yapılandırarak, görüntünün taşıdığı semantik bilgiyi daha belirgin hale getirir. Böylelikle, görüntülerin içerdiği bilgi yoğunluğu artırılmış olur ve sonraki görsel işlem adımları için daha sağlam bir temel sunulur.

Geleneksel yöntemlerle kaydedilmiş düşük kaliteli görüntülerin SR teknolojileri aracılığıyla yüksek çözünürlüklü örnekleriyle karşılaştırılabilir seviyede kaliteye ulaşması, özellikle arşiv verilerinin değerlendirilmesi, düşük bant genişliğinde alınan görüntülerin analizi ve düşük kaliteli kamera sistemlerinden elde edilen görsellerin iyileştirilmesi gibi senaryolarda önemli avantajlar sunmaktadır. Ayrıca, SR tekniklerinin modern derin öğrenme modelleriyle entegre edilmesi, bu süreçteki yeniden yapılandırma doğruluğunu daha da artırarak, çıktıların yalnızca görsel olarak değil, ölçülebilir kalite metrikleri açısından da tatmin edici olmasını mümkün kılmaktadır.

Görüntülerde SR işlemleri, yalnızca klasik görüntü işleme yöntemleriyle sınırlı kalmaksızın, son yıllarda gelişen ileri düzey algoritmalar ve öğrenme tabanlı stratejilerle birlikte çok daha kapsamlı bir şekilde ele alınmaktadır. Geleneksel SR yaklaşımları, LR görüntülerde yer alan sınırlı piksel bilgilerinden yararlanarak, HR görüntülere ait detayları tahmin etmeye odaklanır. Bu süreç genellikle istatistiksel modelleme, yapısal örüntü tanıma ve interpolasyon teknikleri gibi yöntemlerle yürütülür. Ancak bu tür yaklaşımlar, özellikle karmaşık sahneler, düzensiz yapılar veya yoğun detay içeren bölgelerde sınırlı başarı göstermekte ve görsel gerçekçilik açısından tatmin edici sonuçlar üretmekte zorlanmaktadır.

Bu sınırlamaların temel nedeni, geleneksel yöntemlerin görüntüde bulunmayan detayları yalnızca mevcut veriler üzerinden öngörmeye çalışmaları ve herhangi bir dış

veri veya öğrenme süreci içermemeleridir. Ayrıca, bu yöntemler çoğunlukla sabit matematiksel kurallara dayalı olduğu için adaptif davranma kapasiteleri düşüktür ve genelleme yetenekleri zayıftır. Bu durum, özellikle gerçek dünya görüntülerinde rastlanan varyasyonlar karşısında sınırlayıcı bir faktör olarak öne çıkmaktadır.

Son yıllarda yapay zekâ (AI), derin öğrenme (DL) ve yüksek performanslı hesaplama (HPC) alanlarında kaydedilen hızlı gelişmeler, SR alanını kökten dönüştürmüştür. Özellikle konvolüsyonel sinir ağları (Convolutinonal Neural Networks-CNN), üretken çekişmeli ağlar (Generative Adversarial Networks, GAN) ve dönüştürücü tabanlı modeller, düşük çözünürlüklü görüntülerden yüksek kaliteli detayların öğrenilerek üretilmesini mümkün kılan yeni nesil SR çözümlerini beraberinde getirmiştir. Bu yöntemler, yalnızca mevcut görüntü bilgilerine dayanmakla kalmayıp, büyük veri kümeleri üzerinde öğrenilen örüntüleri de modelleyerek kayıp detayların daha doğru bir biçimde yeniden yapılandırılmasına olanak tanır.

SISR, yalnızca tek bir LR görüntüden HR versiyonunu tahmin etmeyi amaçlayan bir yaklaşımdır. Bu yöntem, görüntüdeki detayları tahmin etmek için genellikle uzamsal bağlamı ve derin öğrenme tabanlı yapay sinir ağlarını kullanır. Öte yandan çoklu görüntü süper çözünürlük (Multi-Image Super-Resolution, MISR) ise aynı sahnenin farklı zamanlarda veya açılardan alınmış birden fazla LR görüntüsünü kullanarak daha doğru ve detaylı bir HR görüntü elde etmeyi hedefler. MISR yaklaşımları, görüntüler arasındaki hareketleri hizalayarak (image registration) ortak bilgileri birleştirir ve bu sayede yüksek frekanslı detayların geri kazanımını kolaylaştırır.

MISR, genellikle daha yüksek kaliteli sonuçlar sunmasına rağmen, hizalama ve veri eşleme (alignment) gibi ek karmaşıklıklar içerir. Buna karşılık, SISR daha yaygın ve uygulaması kolaydır, çünkü yalnızca tek bir görüntüye ihtiyaç duyar. Bu bağlamda, her iki yaklaşımın avantajları ve sınırlılıkları uygulama alanına göre değişiklik gösterir (Wang vd., 2020).

1.2. Biyometrik Sistemler ve Görüntü Kalitesinin Rolü

Biyometrik SR, düşük çözünürlüklü biyometrik görüntülerin kalitesini artırmak suretiyle tanıma ve kimlik doğrulama sistemlerinin doğruluğunu ve güvenilirliğini yükseltmeyi hedefleyen özel bir görüntü iyileştirme yaklaşımıdır. Bu yöntem, özellikle güvenlik, adli bilişim, sınır kontrolü ve mobil biyometrik sistemler gibi uygulama alanlarında oldukça kritik bir rol oynamaktadır. Çünkü bu tür sistemlerde kullanılan düşük kaliteli görüntüler —örneğin güvenlik kameralarından alınan avuç içi, yüz, iris ya da parmak izi görüntüleri—, tanıma algoritmalarının performansını ciddi ölçüde

düşürebilmektedir. Bu bağlamda, biyometrik SR, yalnızca görüntü kalitesini iyileştirmeyi değil, aynı zamanda bu iyileştirmenin biyometrik eşleşme sistemleri üzerindeki etkisini de optimize etmeyi amaçlamaktadır.

Biyometrik SR'nin genel süper çözünürlük uygulamalarından temel farkı, yalnızca giriş görüntülerinin istatistiksel ya da yapısal özelliklerinin iyileştirilmesiyle sınırlı olmamasıdır. Genel SR teknikleri, çoğunlukla bir görüntünün daha net, yüksek frekanslı detaylar içeren ve insan gözüyle estetik açıdan daha iyi algılanan bir versiyonunu üretmeyi hedefler. Bu tür yaklaşımlar, piksel bazlı benzerlik, yapısal benzerlik ölçütleri (SSIM), PSNR gibi metrikler doğrultusunda değerlendirilir. Oysa biyometrik SR'de nihai hedef, görsel kalite artışından ziyade, bu iyileştirmenin tanıma performansına doğrudan katkı sağlamasıdır. Başka bir ifadeyle, görüntünün "görsel olarak gerçekçi" olması değil, "biyometrik olarak ayırt edici" özellikler taşıması esastır.

Biyometrik tanıma sistemlerinde elde edilen görüntülerin kalitesi, sistemin genel doğruluk performansı üzerinde belirleyici bir rol oynar. Düşük çözünürlüklü veya bozulmuş görüntüler, özellikle avuç içi, parmak izi, yüz gibi detaylı doku bilgisi gerektiren biyometrik modalitelerde, özellik çıkarımı ve eşleştirme süreçlerini olumsuz etkileyerek tanıma doğruluğunu ciddi ölçüde düşürebilir. Bu durum, hem yanlış pozitif hem de yanlış negatif tanılama oranlarının artmasına yol açar. Görüntü kalitesindeki artış ise, daha tutarlı ve güvenilir biyometrik şablonların oluşturulmasına katkı sağlar. Bu nedenle, süper çözünürlük gibi görüntü iyileştirme tekniklerinin entegrasyonu, biyometrik sistemlerin doğruluk, dayanıklılık ve kullanıcı memnuniyeti açısından daha başarılı sonuçlar üretmesini mümkün kılmaktadır (Zhao vd., 2019).

Avuç içi biyometrisi, geniş yüzey alanı üzerinde yer alan karmaşık damar yapıları, çizgiler ve doku desenleri sayesinde yüksek ayırt edicilik kabiliyetine sahip olmasıyla öne çıkar. Temassız tarama imkânı sunması, hijyen ve kullanıcı konforu açısından önemli bir avantaj sağlarken, sahteciliğe karşı dirençli olması da güvenlik açısından güçlü bir tercih sebebidir (Aykut ve Ekinci, 2013; Bingöl ve Ekinci, 2017). Bununla birlikte, avuç içi biyometrisinin başarısı büyük ölçüde görüntü kalitesine bağlıdır. Düşük çözünürlük, kötü aydınlatma, elin konumlandırma hataları ve hareket bulanıklığı gibi faktörler, damar ve doku desenlerinin yeterince ayrıştırılamamasına neden olabilir. Bu durum, tanıma doğruluğunu olumsuz etkileyerek sistemin güvenilirliğini azaltır. Dolayısıyla, yüksek kaliteli görüntü elde etmeye yönelik donanımsal iyileştirmeler ve süper çözünürlük gibi yazılımsal teknikler, bu zorlukların aşılmasında kritik rol oynamaktadır (Kumar ve Zhang, 2012).

Düşük çözünürlüklü biyometrik veriler, tanıma sistemlerinde önemli performans kayıplarına neden olabilecek çeşitli sorunları beraberinde getirir. Görüntüdeki detay seviyesinin yetersiz olması, özellikle avuç içi, parmak izi ve iris gibi karmaşık desen bilgisi içeren biyometrik modalitelerde özellik çıkarımını zorlaştırır. Bu durum, ayırt edici nitelikteki mikroskobik yapıların bulanıklaşmasına ve dolayısıyla benzer bireyler arasında yanlış eşleşme oranlarının artmasına yol açar. Ayrıca düşük çözünürlük, derin öğrenme tabanlı tanıma sistemlerinin öğrenme süreçlerinde zayıf temsillere neden olarak modelin genelleme kapasitesini sınırlar. Bu tür verilerle eğitilen modeller, özellikle sahada karşılaşılan değişken ortamlarda düşük doğruluk ve yüksek hata oranları gösterebilir. Sonuç olarak, düşük çözünürlüklü biyometrik veriler yalnızca doğruluk değil, aynı zamanda güvenlik ve sistem verimliliği açısından da önemli kısıtlamalar oluşturmaktadır.

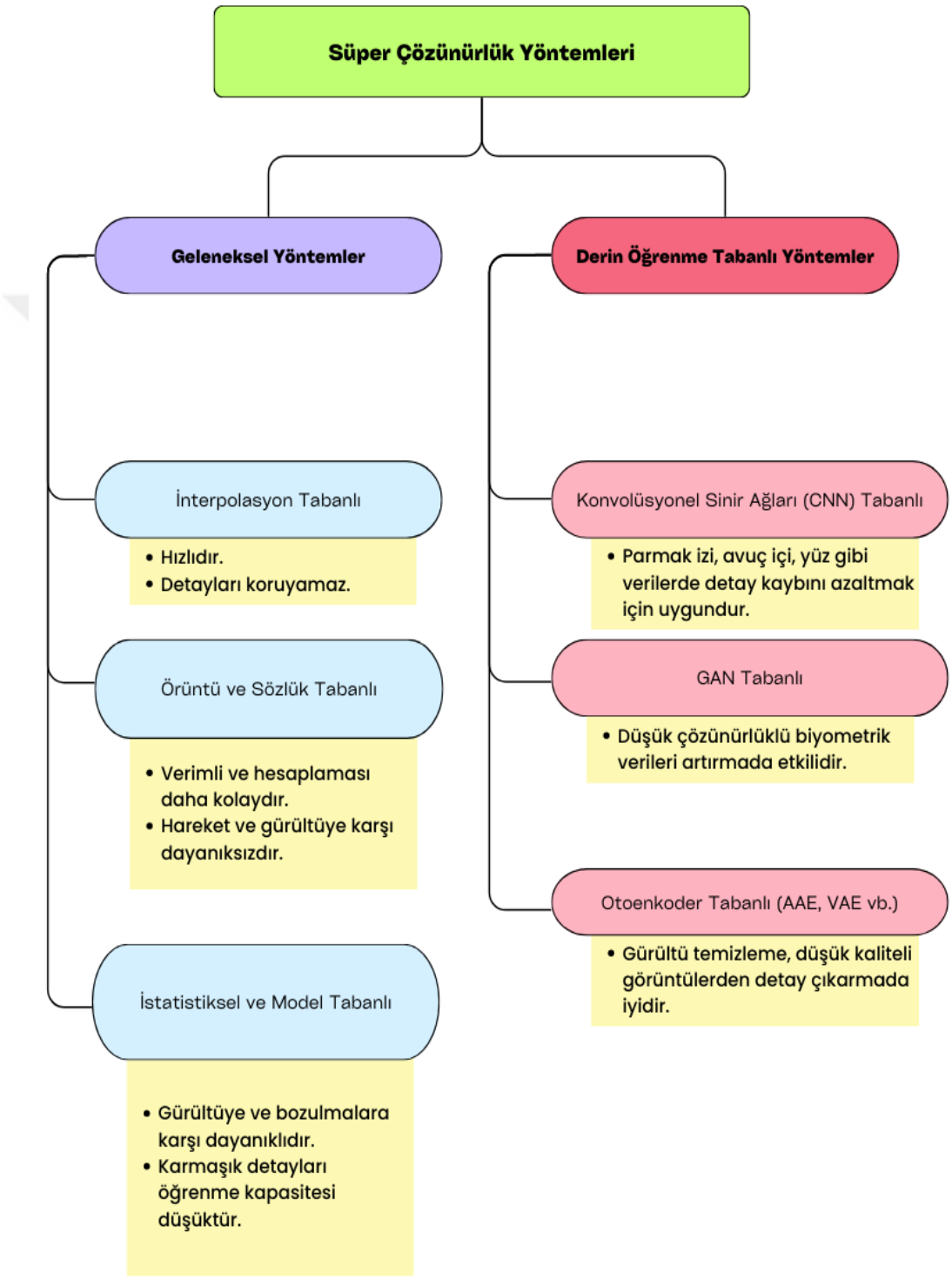
1.3. Derin Öğrenme Tabanlı SR Yöntemleri

Derin öğrenme temelli yöntemlerin SR problemine uygulanmasıyla birlikte, geleneksel tekniklerin sınırlamalarını aşan çok sayıda veri odaklı yaklaşım geliştirilmiştir. Bu gelişim, LR görüntülerin HR karşılıklarına dönüştürülmesinde daha esnek, bağlamsal ve içerik temelli çözümler sunan öğrenme modellerinin önünü açmıştır. Özellikle regresyon tabanlı yapay sinir ağı modelleri, GAN'lar (Goodfellow vd., 2014), akış tabanlı (flow-based) modeller ve olasılık temelli (probabilistic) çerçeveler, SR literatüründe son yıllarda öne çıkan mimariler arasında yer almaktadır (Şekil 2).

CNN'ler, özellikle görüntü işleme ve bilgisayarla görme uygulamalarında yaygın olarak tercih edilen derin öğrenme mimarileridir (Dong vd., 2016). Bu yapılar, görüntülerdeki düşük seviyeli kenar, doku, desen gibi özelliklerden başlayarak daha soyut ve yüksek seviyeli temsillere kadar bilgi çıkarımı yapabilen güçlü katmanlar içerir. CNN'ler, evrişim işlemlerine dayalı olarak yapılandırılmış katmanlar aracılığıyla LR görüntülerdeki mekânsal bilgileri analiz eder, bu analiz sonucu elde edilen öznitelikleri derinleştirerek HR çıktılar üretir. Bu dönüşüm sırasında model, görüntüdeki yapısal bütünlüğü korumaya ve eksik detayları geri kazandırmaya çalışır.

Tipik bir CNN mimarisi, girişten çıkışa doğru sıralanmış birden fazla konvolüsyon katmanı, bu katmanlar arasında yer alan aktivasyon fonksiyonları (çoğunlukla ReLU), havuzlama (pooling) katmanları ve son katmanlarda yer alan tam bağlantılı (fully connected) yapılarından oluşur. Konvolüsyon katmanları, öğrenilebilir filtreler aracılığıyla yerel özellikleri çıkarırken; ReLU gibi aktivasyonlar modele doğrusal olmayanlık kazandırır. Havuzlama katmanları, boyut indirgeme ve öznitelik

yoğunlaştırma işlevi görürken; tam bağlantılı katmanlar ise öğrenilen bilgilerin sınıflandırma veya regresyon gibi görevlerde kullanılmasını sağlar. Bu çok katmanlı yapı, CNN'lerin görüntüleri yalnızca tanımakla kalmayıp, aynı zamanda iyileştirme ve yeniden üretme görevlerinde de yüksek performans göstermesini sağlar.



Şekil 2. Süper Çözünürlük yöntemlerinin sınıflandırılması

Regresyon tabanlı yöntemler, doğrudan LR görüntülerden HR çıktılar üretmek amacıyla eğitilen CNN kullanarak, piksel düzeyinde minimum hata (örn. MSE, L1 loss) hedefler. Bu yaklaşım, ilk SR derin öğrenme modeli olan SRCNN (Dong vd., 2014) ile birlikte yaygınlık kazanmıştır. Ancak bu tür yöntemler sıklıkla görsel bulanıklık sorunuyla karşı karşıya kalmakta, çünkü sadece piksel bazlı benzerlik ölçütlerine odaklandıkları için algısal gerçeklik düzeyini yeterince temsil edememektedir.

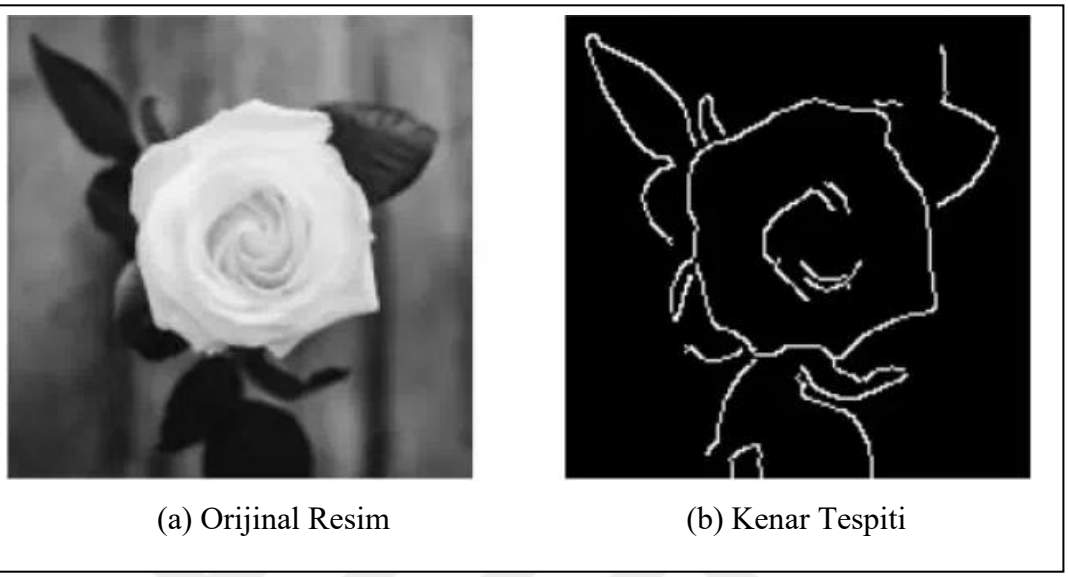
Bu sınırlamaları aşmak üzere önerilen GAN'lar ise, süper çözünürlük alanına devrim niteliğinde katkılarda bulunmuştur. GAN tabanlı modellerde iki farklı ağdan oluşan bir yapı söz konusudur: görüntü üretiminden sorumlu olan üretici (generator) ağ ve bu çıktının gerçekliğini değerlendiren ayrıştırıcı (discriminator) ağ. SRGAN (Ledig vd., 2017) gibi modeller, yalnızca yapay olarak yüksek çözünürlük üretmekle kalmaz; aynı zamanda insan algısına yakın görsel kalite sunmayı hedefleyen algısal kayıpları (perceptual loss) da entegre ederek, daha doğal ve estetik açıdan tatmin edici sonuçlar üretir.

Bunun ötesinde, akış tabanlı (flow-based) SR modelleri ve varyasyonel otokodlayıcılar (variational autoencoders, VAE'ler) gibi olasılık temelli modeller, belirsizliklerin modellenmesi ve çoklu olası çözüm üretme yetenekleri ile öne çıkar. Bu tür yapılar, özellikle görüntüde kaybolan detayların birden çok geçerli yeniden yapılandırması olabileceği durumlarda, deterministik modellerin tekil tahminlerinin ötesine geçerek daha zengin bir çıktı uzayı sunar.

Derin öğrenme tabanlı bu modellerin ortak özelliği, büyük ölçekli eğitim veri kümelerinden örüntü öğrenerek, LR görüntüleri yalnızca piksel düzeyinde benzer değil; aynı zamanda semantik bağlam açısından tutarlı, yüksek kaliteli HR görüntülere dönüştürme kapasiteleridir. Özellikle GAN tabanlı ve algısal kayıplar kullanan modeller, insan görsel algısına yakınlığı artırarak yapay olarak üretilen görüntülerin daha doğal görünmesini sağlamaktadır. Bu yönüyle SR teknolojisi, yalnızca teknik doğruluk değil, aynı zamanda görsel tatmin ve uygulama düzeyinde güvenilirlik açısından da önemli bir eşiği aşmıştır.

Geleneksel SR yöntemleri, görüntüde yüksek frekans bileşenlerini tahmin etmek amacıyla çeşitli istatistiksel ve yapısal stratejiler kullanırlar. Bunlar arasında, görüntü istatistiklerine dayalı modeller (Kim ve Kwon, 2010), kenar tespiti (Şekil 3) ve yapısal detayların analizi (Freedman ve Fattal, 2011), tahmine dayalı yeniden yapılandırma yöntemleri, seyrek gösterim (sparse representation) ve örüntü eşleme gibi teknikler yer almaktadır. Bu yöntemler, mevcut pikseller arasındaki korelasyonları kullanarak eksik

detayları modellemeye çalışırlar. Ancak bu yaklaşımlar, genellikle gerçek dünyadaki karmaşık sahnelerde başarılı olamazlar ve sıkça gürültü, bulanıklık ya da yapay artefaktların ortaya çıkmasına neden olabilirler. Bu durum, özellikle büyütme oranı arttıkça daha belirgin hale gelir ve görüntü kalitesini olumsuz etkiler.



(a) Orijinal Resim (b) Kenar Tespiti

Şekil 3. Kenar Tespiti örneği

Buna karşın, modern SR yaklaşımları derin öğrenme algoritmalarıyla çok daha güçlü sonuçlar sunmaktadır. Derin sinir ağlarının özellikle konvolüsyonel yapıları (CNN), görüntüleri uçtan uca bir şekilde öğrenerek LR'den HR'ye dönüşüm sağlayabilmektedir. Bu alandaki öncü çalışmalardan biri olan SRCNN (Super-Resolution using Convolutional Neural Networks) (Dong vd., 2014), SR problemini doğrudan öğrenmeye dayalı olarak çözmeye yönelik ilk denemelerden biridir. SRCNN, üç katmanlı sade bir CNN yapısıyla LR görüntüleri HR'ye dönüştürebilmekteydi. Ardından geliştirilen FSRCNN (Fast SRCNN), daha derin yapılar ve daha küçük filtreler ile hem hız hem de doğruluk açısından önemli geliştirmeler sunmuştur (Dong vd., 2016). Bu modeller, görüntü SR'sine derin öğrenmenin uygulanabilirliğini göstererek, daha karmaşık ve güçlü yapılar için zemin hazırlamıştır.

İlerleyen çalışmalarla birlikte, artık yalnızca katman sayısının artırılması değil, aynı zamanda daha gelişmiş sinir ağı mimarilerinin kullanımı gündeme gelmiştir. Örneğin, ResNet ve DenseNet gibi bilgisayarla görme alanında başarı gösteren mimariler, SR modellerine entegre edilmiştir (Ledig vd., 2016). Bu modeller, özellik haritaları arasında atlamalı bağlantılar kurarak gradyan akışını iyileştirir ve daha derin modellerin eğitilmesine olanak tanır. Ortaya çıkan modellerin çoğu, düşük çözünürlüklü görüntülerden yüksek çözünürlüklü çıktılar üretmek için regresyon tabanlı stratejiler

izler. Bu modellerin eğitiminde genellikle L1 ya da L2 kayıp fonksiyonları kullanılır. Ancak bu tür kayıplar, piksel bazlı hataları minimize etmeye çalıştığı için, çoğu zaman pürüzsüz ve detaydan yoksun görüntüler üretir. Bu durum özellikle büyütme oranı yüksek ($s > 4$) senaryolarda daha belirgin hale gelmektedir.

Bu zayıflığı gidermek için, gerçekçi ve doğal görünen görüntüler üretebilen üretici modeller tercih edilmeye başlanmıştır. GAN tabanlı SR yaklaşımları, görsel kaliteyi artırmak adına yalnızca piksel doğruluğuna değil, aynı zamanda içerik tutarlılığına ve görsel algı kalitesine de odaklanmaktadır. Bu sayede, görüntülerde daha fazla detay, doku ve gerçekçilik elde edilebilmektedir. Örneğin, SRGAN, VGG tabanlı algısal kayıpları ve GAN mimarisini birleştirerek detaylı ve doğal görünümlü SR çıktıları üretmiştir (Ledig vd., 2016b). Daha güncel çalışmalar ise bu yapıyı geliştirerek ESRGAN, Real-ESRGAN gibi daha yüksek kaliteli modeller ortaya koymuştur (Moser vd., 2022).

1.4. Süper Çözünürlüğün Biyometrik Sistemlerde Kullanımı

SR teknolojileri, düşük çözünürlüklü biyometrik verilerin yapısal bütünlüğünü ve detay zenginliğini artırmak amacıyla güçlü bir araç sunmaktadır. Özellikle avuç içi, yüz, iris gibi biyometrik modalitelerde, düşük çözünürlükten kaynaklanan çizgi silinmeleri, desen bozulmaları ve kenar belirsizlikleri gibi problemler tanıma sürecinde önemli bilgi kayıplarına yol açmaktadır. SR modelleri bu eksiklikleri gidermek üzere, kaybolmuş yüksek frekans bileşenlerini yeniden tahmin ederek görsel bütünlüğü güçlendirir. Böylece orijinal verinin taşıdığı ayırt edici özellikler daha belirgin hale gelir. Geleneksel interpolasyon yöntemlerinin aksine, öğrenme tabanlı SR modelleri örüntü ve doku bilgisini içerik duyarlı biçimde yeniden yapılandırarak, biyometrik sistemlerde kullanılabilecek nitelikte yüksek kaliteli veriler üretme potansiyeline sahiptir. Bu yönüyle SR, yalnızca çözünürlük artırımı değil, aynı zamanda biyometrik veri kalitesinin semantik olarak güçlendirilmesi için de stratejik bir dönüşüm aracı olarak öne çıkmaktadır.

Biyometrik SR modelleri, geleneksel SR yaklaşımlarının ötesine geçerek, tanıma doğruluğunu maksimize edecek şekilde tasarlanır. Bu tasarım süreci, genellikle biyometrik özelliklere özel ön bilgi (domain-specific prior knowledge) ile yönlendirilir. Örneğin, yüz için mimik değişkenliği, avuç içi için çizgi yapısı, iris için desen yoğunluğu gibi özgün yapılar modelin öğrenme sürecine entegre edilir. Bu yapısal öncüller, SR modeline doğrudan ya da dolaylı yollarla aktarılabilir: özel kayıp fonksiyonları, çoklu görevli öğrenme (multi-task learning), dikkat mekanizmaları ya da tanıma ağı ile birlikte eğitilen hibrit modeller (joint optimization) gibi stratejilerle.

Biyometrik SR alanındaki çalışmaların temelleri oldukça eskiye dayanmaktadır. Bu alandaki ilk önemli katkı, 1974 yılında Gerchberg tarafından gerçekleştirilmiştir. Gerchberg'in önerdiği yaklaşım, eksik frekans bileşenlerinin tahmini yoluyla görüntü iyileştirme prensibine dayanmakta olup, günümüzde kullanılan bazı frekans temelli SR yöntemlerinin teorik temelini oluşturmuştur. Bu yöntem, özellikle Fourier alanında düşük çözünürlüklü görüntülerden daha fazla detay elde edilebileceğini öne sürerek, SR kavramının ilk sistematik uygulamalarından biri olarak literatüre geçmiştir (Gerchberg, 1974).

Biyometrik modaliteler için süper çözünürlük uygulamaları ise bu temel yaklaşımlardan yıllar sonra geliştirilmeye başlanmıştır. Bu alandaki ilk örneklerden biri, 1985 yılında Mjolsness tarafından ortaya konmuştur. Mjolsness, doktora tezinde, düşük kaliteli parmak izi görüntülerinin iyileştirilmesine yönelik bir SR yöntemi geliştirmiştir. Önerdiği model, düşük çözünürlükten kaynaklanan yapısal bozulmaların, parmak izi sırt ve vadilerinin düzgün bir şekilde temsil edilebilmesini engellediğini vurgulamış ve bu eksikliğin giderilmesi adına SR'nin potansiyelini ortaya koymuştur. Bu çalışma, biyometrik tanıma sistemlerinde SR'nin doğruluğu artırma potansiyelini ilk gösteren yaklaşımlardan biri olması bakımından öncü niteliği taşımaktadır (Mjolsness, 1986).

Biyometrik SR alanında önemli bir diğer dönüm noktası, 2000 yılında Baker ve Kanade tarafından gerçekleştirilmiştir. Bu çalışmada, yüz görüntüleri için çoklu düşük çözünürlüklü görüntülerden yüksek çözünürlüklü yüz görüntüsü üretme amacı güdülmüştür. Yazarlar, insan yüzlerinin belirli yapısal ve istatistiksel özelliklere sahip olduğunu, bu nedenle öğrenilmiş modellerle bu yapıların eksik detaylarının tahmin edilebileceğini savunmuşlardır. Bu modelin temel amacı, görsel gerçekçiliğin ötesinde, yüz tanıma performansını da artıracak ayrıntılı görsel temsiller elde etmekti (Baker ve Kanade, 2000).

SR'nin iris biyometrisi özelindeki ilk uygulaması ise 2003 yılında Huang ve arkadaşları tarafından gerçekleştirilmiştir. Bu çalışmada, özellikle düşük çözünürlüklü iris görüntülerinden yüksek doğrulukta kimlik doğrulama yapabilmenin zorlukları ele alınmış ve SR algoritmalarının bu eksikliği nasıl telafi edebileceği değerlendirilmiştir. Yazarlar, iris desenlerinin yüksek frekanslı bileşenlerinin kimlik doğrulamada kritik rol oynadığını ve bu detayların SR aracılığıyla geri kazanılmasının tanıma doğruluğuna doğrudan katkı sağladığını göstermiştir. Böylece, iris biyometrisi için SR kullanımının ilk uygulamalı örneklerinden biri literatüre kazandırılmıştır (Huang vd., 2003).

1.5. Tezin Amacı ve Kapsamı

Bu çalışmanın temel amacı, öğrenme tabanlı derin mimarilerin kapasitesinden faydalanarak düşük çözünürlüklü avuç içi biyometrik görüntülerin kalitesini artırmak ve böylece bu görüntülerin kimlik doğrulama sistemlerindeki kullanılabilirliğini ve güvenilirliğini yükseltmektir. Özellikle otomatik tanıma sistemlerinde, düşük çözünürlüklü veriler biyometrik ayırt ediciliği ciddi şekilde azaltmakta ve sistem performansını olumsuz etkilemektedir. Bu bağlamda, çalışmada Otokodlayıcı ve GAN tabanlı süper çözünürlük yöntemleri kullanılarak avuç içi görüntülerinin görsel kalitesinin ve yapısal bütünlüğünün iyileştirilmesi hedeflenmektedir. Böylece hem veri güvenliği hem de tanıma doğruluğu açısından daha üst düzey performans elde edilmesi amaçlanmaktadır.

Mevcut literatürde süper çözünürlük yöntemlerinin genel görüntü işleme alanında yaygın şekilde uygulandığı görülmektedir. Ancak biyometrik veriler, özellikle de avuç içi gibi karmaşık yapılar içeren modaliteler için yapılan SR uygulamaları sınırlıdır. Bu bağlamda, literatürde hem AE hem de GAN tabanlı yaklaşımların doğrudan karşılaştırıldığı ve biyometrik tanıma sistemlerine etkisinin deneysel olarak değerlendirildiği kapsamlı çalışmaların eksik olduğu görülmektedir. Bu tez, bu boşluğu doldurmayı hedeflemektedir.

Çalışmada kullanılan AE mimarisi, düşük çözünürlüklü giriş görüntülerinden yüksek çözünürlüklü çıktılar üretmeyi hedefleyen bir yeniden yapılandırma süreci yürütmektedir. Bu mimari, giriş verisinin sıkıştırılmış bir temsili aracılığıyla temel yapısal bileşenleri öğrenmekte ve ardından bu temsili kullanarak daha ayrıntılı ve net görüntüler oluşturmaktadır. Bu yaklaşım sayesinde yapısal bilgi kaybı minimize edilmekte, özellikle bozulma ve gürültü içeren görüntülerde dahi yüksek frekans detayları geri kazanılabilmektedir. AE tabanlı modellerin bu özelliği, gürültü azaltma (denoising) ve çözünürlük artırma (upscaling) işlemlerini aynı anda gerçekleştirebilmesine olanak tanır.

Öte yandan, GAN tabanlı modeller, bir üretici ve bir ayrıştırıcı ağdan oluşan rekabetçi yapıları sayesinde, yalnızca teknik doğruluğa değil aynı zamanda algısal gerçekçiliğe de odaklanır. Bu modeller, görsel olarak daha inandırıcı ve yüksek frekanslı detayları daha başarılı şekilde içeren çıktılar üretmektedir. GAN mimarilerinin modüler doğası, modelin farklı veri bozulma senaryolarına veya uygulama alanlarına göre özelleştirilmesini mümkün kılmakta; perceptual loss, adversarial loss ve content loss gibi farklı kayıp fonksiyonlarının entegre edilmesine olanak sağlamaktadır.

Çalışmada geliştirilen AE ve GAN tabanlı süper çözünürlük sistemleri, farklı çözünürlük faktörleri, gürültü seviyeleri ve veri bozulma tiplerine karşı uyarlanabilir olacak şekilde tasarlanmıştır. Bu yapıların uçtan uca öğrenme prensibine dayanması, klasik SISR yaklaşımlarına göre daha esnek, otomatik ve optimize edilebilir bir çözüm ortaya koymaktadır. Ayrıca, çok görevli öğrenme, paylaşımlı ağırlıklar ve bilgi aktarımı (transfer learning) gibi ileri stratejilerin entegrasyonu sayesinde modelin parametre verimliliği artırılmakta; eğitim süresi kısaltılmakta ve hesaplama kaynakları daha verimli kullanılabilmektedir.

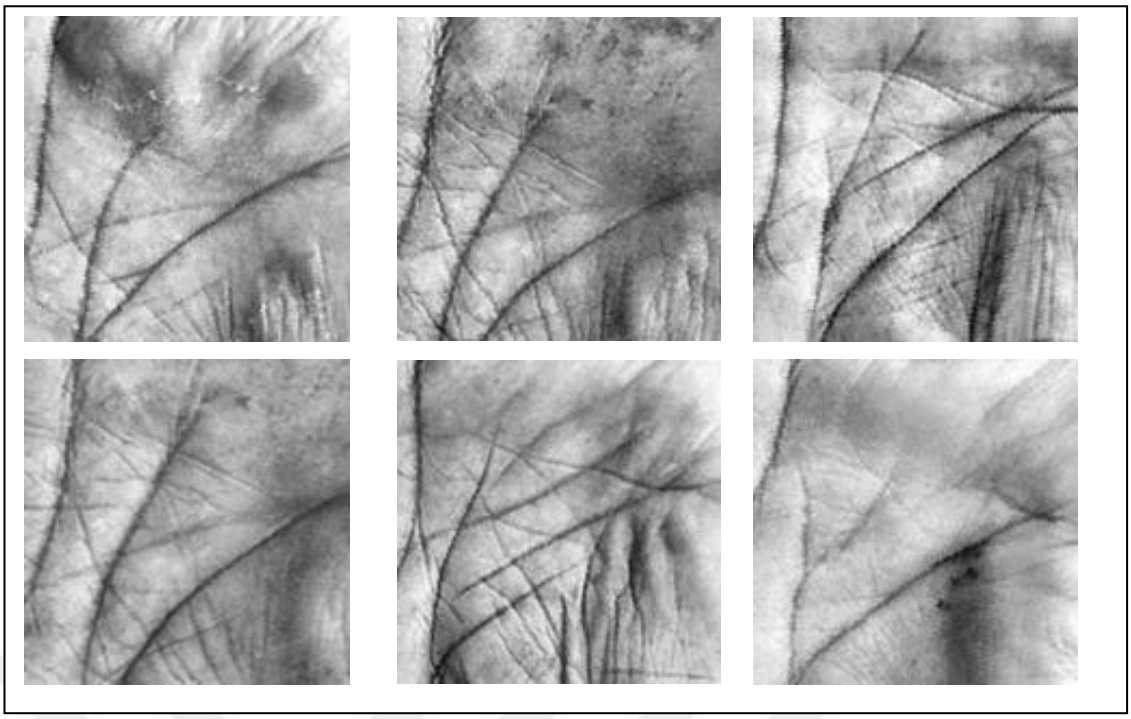
Bu bağlamda, çalışmada önerilen yöntemlerin yalnızca görsel kaliteyi artırmakla kalmayıp; aynı zamanda biyometrik sistemlerin doğruluk, güvenilirlik ve ölçeklenebilirlik gibi temel performans ölçütlerinde de anlamlı iyileştirmeler sağlaması beklenmektedir.

2.YAPILAN ÇALIŞMALAR

2.1. Veri Kümesi Tanımlaması

IIT Delhi avuç içi görüntü veritabanı, Hindistan'ın Yeni Delhi kentindeki IIT Delhi öğrencileri ve personelinden toplanan el görüntülerini içermektedir. Temmuz 2006 ile Haziran 2007 tarihleri arasında, IIT Delhi kampüsünde geliştirilen temassız bir görüntüleme sistemi kullanılarak oluşturulmuştur. Görüntüler, kapalı bir ortamda, kamera lensi etrafında dairesel floresan aydınlatma ile çekilmiştir. Veritabanı içerisinde avuç izi ilgin bölgelerinin (ROI) bulunduğu ek görüntüler de içermektedir. Bu görüntüler Zhang vd. 2003 çalışmasında kullanılan ROI tespiti yaklaşımıyla elde edilmiştir ve örnek görüntüler Şekil 4'te verilmiştir. ROI tespiti için daha etkin yöntemlerin kullanıldığı farklı çalışmalar (Aykut ve Ekinci, 2015; Sünnetçi vd., 2024) bulunmasına rağmen, bu çalışmada orijinal veritabanındaki görüntülerin kullanılması tercih edilmiştir.

Bu veri kümesi, 230 farklı kullanıcıya ait görüntülerden oluşmakta olup, tüm veriler bitmap (.bmp) formatında saklanmaktadır. Katılımcılar 12 ile 57 yaş arasındadır. Her bireyin hem sol hem de sağ elinden yedişer farklı görüntü alınmış ve çeşitli el pozisyonları kaydedilmiştir. Görüntüleme sürecinde, katılımcılara ellerini doğru bir şekilde konumlandırmaları için canlı geri bildirim sağlanmıştır. Temassız görüntüleme yöntemi kullanılması nedeniyle, görüntülerde belirgin ölçekte değişkenlikler gözlemlenebilir.

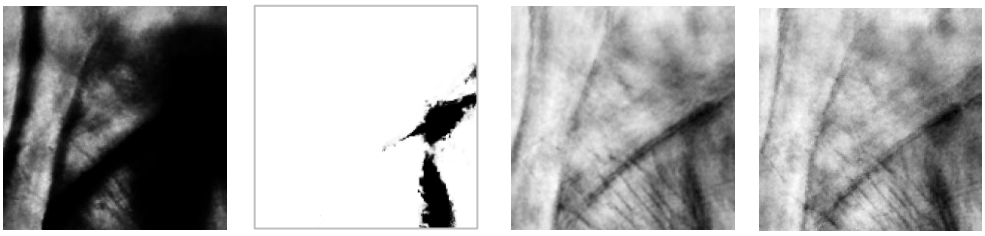


Şekil 4. IITD Avuçizi Veritabanından örnek görüntüler

Toplanan tüm görüntüler, her kullanıcıya atanmış benzersiz bir kimlik numarasıyla sıralı olarak numaralandırılmıştır. Orijinal görüntülerin çözünürlüğü 800 × 600 piksel olup, ek olarak 150 × 150 piksel boyutlarında otomatik kırpılmış ve normalize edilmiş avuç içi görüntüleri de veri kümesine dahil edilmiştir (Kumar, 2008).

2.2. Kayıp Fonksiyonları

Bu çalışmada AE mimarisine dayalı olan AAE modeli kullanılmıştır. Modelin eğitiminde farklı kayıp (loss) fonksiyonları denenmiş ve elde edilen çıktı görüntüler görsel kalite açısından karşılaştırılmıştır. Yapılan bu incelemeler sonucunda, en başarılı sonuçların Binary Cross-Entropy Loss (BCE Loss) ve L1 Loss kullanıldığında elde edildiği gözlemlenmiştir (Şekil 5). Bu iki kayıp fonksiyonu, hem ayrıştırıcının eğitimi hem de kod çözücü çıktısının daha doğal ve gerçekçi görüntüler üretmesi açısından en uygun kombinasyonu sunmuştur.



Şekil 5. AAE Modelinde denenen bazı kayıp fonksiyonları(a, b ve c) ve L1+BCE fonksiyonu(d)

BCE kayıp fonksiyonu, genellikle ikili sınıflandırma problemlerinde kullanılan ve özellikle GAN tabanlı mimarilerde, ayrıştırıcının çıktısını değerlendirmek için tercih edilen bir kayıp fonksiyonudur. Matematiksel olarak formülü Eşitlik 1'de olduğu gibi ifade edilir.

$$\mathcal{L}_{BCE} = -\left[y \cdot \log(\hat{y}) + (1 - y) \cdot \log(1 - \hat{y})\right] \quad \text{(Eşitlik 1)}$$

Burada y gerçek etiketi, ŷ ise modelin tahmin ettiği değeri ifade eder. AAE mimarisinde, ayrıştırıcıdan alınan BCE kayıp fonksiyonunun değeri, gizli uzaydaki örneklerin gerçek dağılıma yakınsamasını sağlamak amacıyla kullanılır (I. Goodfellow vd., 2016).

L1 kayıp fonksiyonu, piksel düzeyinde üretilen görüntü ile gerçek görüntü arasındaki farkın mutlak değeri üzerinden ortalama alınarak hesaplanır (Eşitlik 2).

$$\mathcal{L}_{L1} = \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{x}_i| \quad \text{(Eşitlik 2)}$$

Burada x_i gerçek görüntü piksellerini, $\hat{x}_i$ ise model tarafından üretilen pikselleri temsil eder. L1 Loss, görüntünün yapısal bütünlüğünü korumada ve keskinlik sağlamada önemli rol oynamaktadır (Goodfellow vd., 2016).

Modelin toplam kaybı, ayrıştırıcıdan gelen BCE kayıp fonksiyonu ve kod çözücünün çıktısından gelen L1 kayıp fonksiyonunun bir kombinasyonu olarak hesaplanmaktadır. Bu iki bileşenin birbirine göre ağırlığı, modelin hem gerçeğe yakın görüntüler üretmesi hem de detaylı, yapısal olarak doğru görüntüler oluşturması açısından kritik öneme sahiptir.

Bu çalışmada yapılan deneysel analizler sonucunda, BCE kayıp fonksiyonu için 0.001, L1 kayıp fonksiyonu için ise 0.999 ağırlıkları kullanılmıştır. Bu oranların

belirlenmesinde temel motivasyon, modelin temel hedefinin yüksek kaliteli, yapısal olarak doğru görüntüler üretmek olmasıdır. Ayrıştırıcı tarafından gelen BCE kayıp fonksiyonunun ağırlığı düşürülerek, modelin sadece gerçek/sahte ayrımına odaklanması engellenmiş, bunun yerine kod çözücü çıkışının görsel kalitesine öncelik verilmiştir. L1 kayıp fonksiyonunun yüksek ağırlığı, modelin giriş görüntüsüne benzer ve doğal görünümlü detaylar üreten bir kod çözücü eğitilmesini sağlamıştır.

Bu oranlar ile model, hem örtük uzayda anlamlı örnekler üretmeyi öğrenmiş hem de görsel olarak yüksek kaliteli görüntüler elde etmiştir. Düşük BCE kayıp fonksiyonunun ağırlığı aynı zamanda ayrıştırıcının aşırı baskın hale gelmesini engelleyerek, modelin stabil bir şekilde eğitilmesine olanak tanımıştır.

2.3. AE Tabanlı Süper Çözünürlük Modeli

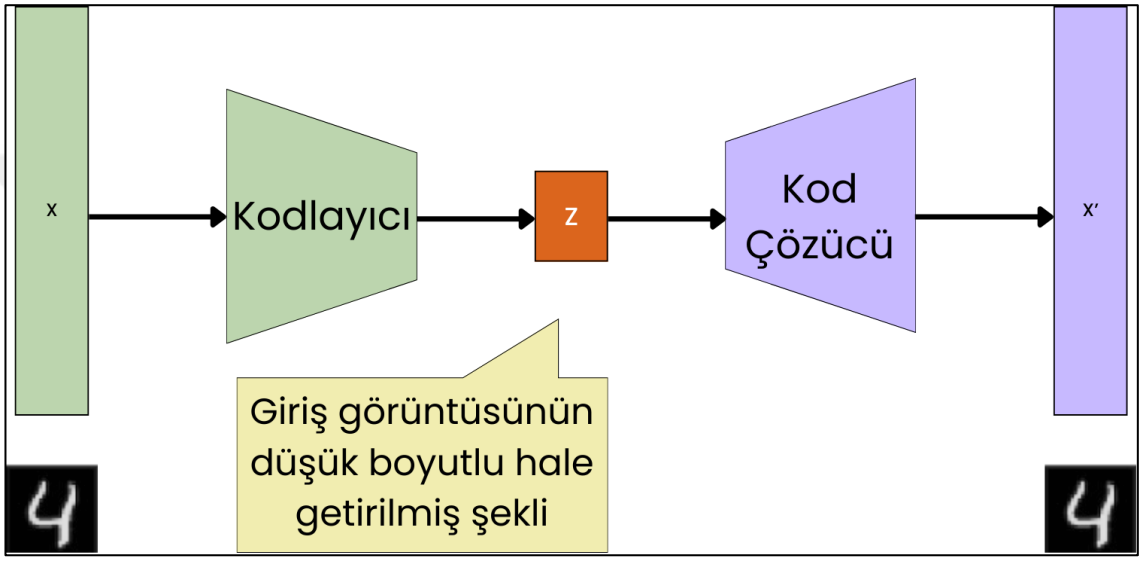
Bir AE, özellikle yüksek boyutlu verilerin daha kompakt ve anlamlı temsillerini öğrenmek amacıyla kullanılan denetimsiz bir yapay sinir ağı mimarisidir. AE'ler temel olarak iki ana bileşenden oluşur: kodlayıcı (encoder) ve kod çözücü (decoder). Kodlayıcı, girdi verisini anlamlı bir şekilde sıkıştırarak daha düşük boyutlu bir temsil olan örtük vektöre (latent vector) dönüştürür. Bu örtük uzay (latent space), girdinin en ayırt edici ve önemli özelliklerini temsil eder. Ardından, kod çözücü bu örtük temsili kullanarak orijinal girdi verisini yeniden oluşturmaya çalışır. Böylece AE, veriyi hem sıkıştırmayı hem de kayıpsız veya minimum kayıpla yeniden üretmeyi hedefler (Kramer, 1991).

AE'ler çoğunlukla gürültü giderme, boyut indirgeme, anlamsal özellik çıkarımı ve veri yeniden yapılandırması gibi görevlerde yaygın biçimde kullanılmaktadır. Özellikle yüksek boyutlu görüntü verileri için AE'ler, veri içerisindeki anlamsal yapıyı koruyarak sadeleştirme yapabilmeleri sayesinde önemli avantajlar sunar. Gürültü giderme AE'leri (denoising AE), bozulmuş ya da eksik verilere karşı oldukça dayanıklıdır ve girişteki rastgele gürültüyü filtreleyerek daha temiz bir temsil elde etmeye olanak tanır. Benzer şekilde, varyasyonel otokodlayıcılar (VAE'ler), AE yapısının olasılıksal bir genellemesi olarak çalışır ve özellikle görüntü sentezi, örnekleme ve istatistiksel temsiller üretimi gibi alanlarda kullanılmaktadır.

GAN'lar ile karşılaştırıldığında, AE'ler veri üretiminden ziyade veri temsilini öğrenmeye odaklanır. GAN mimarisi veri üretiminde yüksek görsel gerçeklik sağlamayı amaçlarken, AE'ler giriş verisinin altında yatan yapısal özellikleri kompakt bir biçimde modellemeye çalışır. Bu yönüyle AE'ler, özellikle SR gibi görevlerde giriş görüntüsünün temel özelliklerini öğrenmek ve daha sonra bu özelliklerden yüksek çözünürlüklü versiyonlar üretmek için bir ön adım ya da temel yapı taşı olarak kullanılabilir. Nitekim

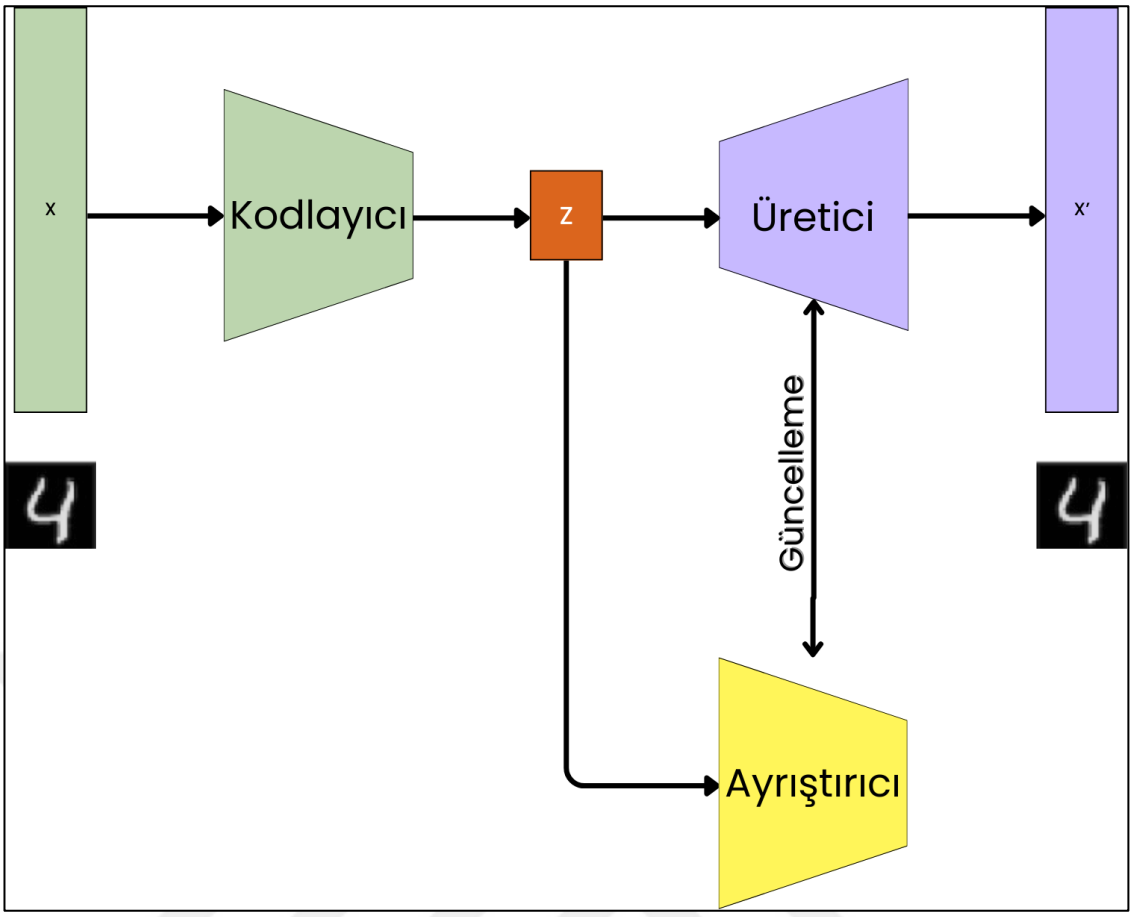
bazı SR yaklaşımları, AE mimarisini GAN ile birleştirerek hem görsel gerçekliği hem de yapısal doğruluğu aynı anda artırmayı hedeflemiştir.

Kingma ve Welling (2013) tarafından tanıtılan VAE modeli, AE mimarisinin önemli bir evrimsel adımı olarak kabul edilir. Bu yaklaşım, AE'nin deterministik yapısını olasılıksal bir çerçeveye taşıyarak, daha güçlü ve kontrollü bir veri üretim yeteneği sağlamıştır. AE mimarilerinin bu esnekliği ve öğrenme kapasitesi, SR başta olmak üzere pek çok bilgisayarla görü ve görüntü işleme probleminde güçlü bir temel sunmaktadır (Şekil 6).



Şekil 6. AE mimari şeması

Çekişmeli Otokodlayıcı (Adversarial Autoencoder-AAE), geleneksel AE' lerin bir uzantısı olup, gizli uzaydaki veri dağılımını şekillendirmek için adversaryal öğrenmeyi kullanır. AAE, VAE gibi olasılıksal bir modelleme yaklaşımı benimser ancak Kullback-Leibler (KL) uzaklığı yerine bir ayrıştırıcı ile gizli değişkenin belirlenen öncel dağılıma uymasını zorlar (Şekil 7). Böylece, elde edilen örtük uzay dağılımı daha esnek ve kontrol edilebilir hale gelir. AAE'nin en büyük avantajlarından biri, VAE gibi karmaşık türev hesaplamaları gerektirmemesi ve GAN'lar gibi eğitim dengesizliği sorunlarını daha az yaşamasıdır. Aynı zamanda geleneksel AE'lere kıyasla daha anlamlı ve yapılandırılmış bir örtük uzay sunarak, veri üretimi ve özellik öğrenimi açısından daha iyi sonuçlar verebilmektedir (Makhzani vd., 2015).



Şekil 7. Çekişmeli AE mimari şeması

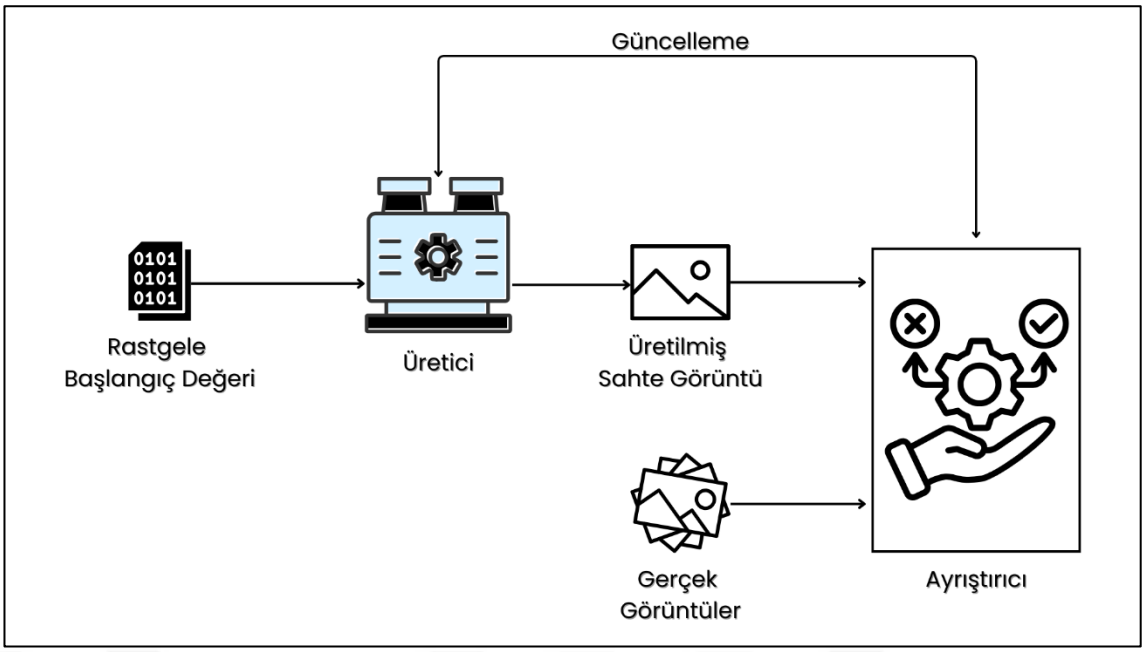
Eğitim süreci 2000 döngü boyunca yürütülmüş ve her bir güncellemede 64 boyutunda veri yığınları kullanılmıştır. Öğrenme oranı 0.0001, momentum katsayıları ise sırasıyla $\beta1=0.5$ ve $\beta2=0.999$ olarak belirlenmiştir. Latent vektörün boyutu 10 olarak seçilmiş ve gri tonlamalı (1 kanal) 150×150 çözünürlüklü görüntülerle çalışılmıştır. Modelin çıktılarındaki kaliteyi artırmak ve yapısal tutarlılığı korumak amacıyla önceki bölümde de bahsedildiği gibi L1 + BCE kayıp fonksiyonlarının kombinasyonu tercih edilmiştir. L1 Loss, çıktı görüntü ile hedef görüntü arasındaki piksel farklarının mutlak değerini minimize ederek daha keskin ve detaylı sonuçlar üretmeyi hedefler. Ayrıca deneylerde karşılaştırmalı olarak L2 (Mean Squared Error) ve SSIM (Structural Similarity Index) gibi farklı kayıplar da değerlendirilmiş; ancak L2'nin yumuşak geçişlere neden olduğu, SSIM'in ise yapısal bütünlükte avantaj sağlamakla birlikte algısal kaliteyi yeterince temsil edemediği gözlemlenmiştir. Sonuç olarak, L1 + BCE kombinasyonunun hem yapısal detayların korunmasında hem de ayrıştırıcı üzerinden gerçekçilik sağlamada en dengeli çözümü sunduğu belirlenmiştir.

2.4. GAN Tabanlı Süper Çözünürlük Modeli

GAN'lar, görüntü üretimi ve iyileştirme alanlarında çığır açan modeller arasında yer alır. Özellikle SR problemlerinde, geleneksel yöntemlerin aksine daha doğal, detaylı ve görsel olarak tatmin edici sonuçlar üretme yetenekleriyle öne çıkarlar. GAN mimarisi iki temel bileşenden oluşur: biri veri üretmeye odaklanan üretici, diğeri ise bu verilerin gerçek mi yoksa sahte mi olduğunu ayırt etmeye çalışan ayırt edici modeldir. Bu iki model, birbirlerine karşıt görevlerle eğitilir ve bu çekişmeli yapı, modelin zamanla daha gerçekçi çıktılar üretebilmesini sağlar.

Bu model, birbiriyle rekabet eden iki CNN kullanır: Üretici (Generator) ve Ayrıştırıcı (Discriminator). Eğitim sürecinin temelinde, üreticinin ayrıştırıcıyı aldatma çabası yatmaktadır(Şekil 8). Başlangıçta üretici, düşük çözünürlüklü girişten yola çıkarak yüksek çözünürlüklü ama sahte görüntüler üretir. Bu görüntüler başlangıçta oldukça yapay ve düşük kalitede olabilir; ancak üretici, zamanla ayırt edicinin geri bildirimlerine göre kendini geliştirir ve daha gerçekçi görüntüler üretmeye başlar. Aynı zamanda ayırt edici de gelişerek üretici tarafından oluşturulan bu sahte örnekleri daha iyi ayırt etmeye başlar. Bu iki ağ arasındaki dinamik denge, modelin hem görsel hem de yapısal doğruluğu yüksek sonuçlar üretmesini mümkün kılar.

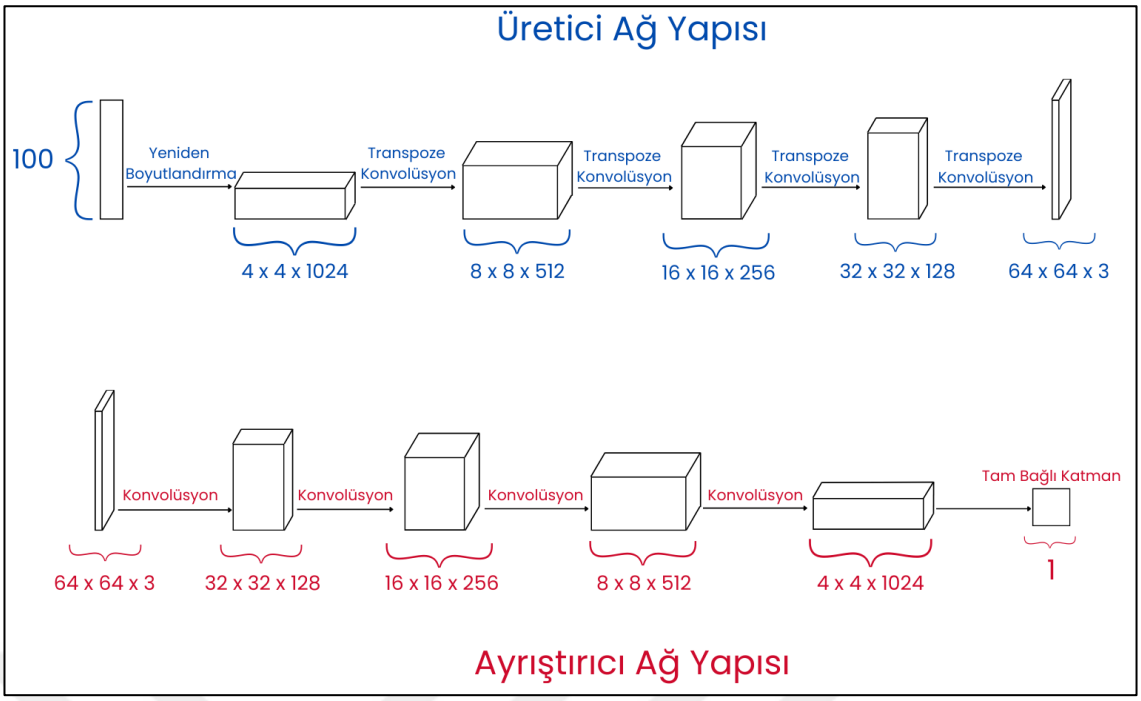
GAN tabanlı SR yöntemleri, geleneksel L1 veya L2 kayıplarının sebep olduğu pürüzsüz ve detaydan yoksun görüntüler yerine, daha keskin, ayrıntılı ve algısal olarak tatmin edici çıktılar üretebilir. Bu yöntemlerde kullanılan algısal kayıplar (perceptual losses) ve içerik kayıpları (content losses), özellikle insan gözüyle anlamlı farklar yaratabilecek detayların korunmasını sağlar. Böylece elde edilen görüntüler sadece piksel benzerliği açısından değil, aynı zamanda görsel gerçeklik bakımından da üstündür.



Şekil 8. Çekişmeli Üretken Ağlar (GAN)

Bu yapı, SRGAN(Super-Resolution Using a Generative Adversarial Network) (Ledig vd., 2016b) ve ESRGAN (Enhanced Super-resolution Generative Adversarial Networks) (Wang vd., 2018) gibi modellerde kullanılır ve daha az yumuşatılmış, daha keskin görüntüler üretmek için çekişmeli kayıp (adversarial loss) ve içerik kaybı (content loss) kombinasyonu ile optimize edilir. Son teknoloji GAN modelleri, yüksek kalitede ve ayrıntılı görüntüler oluşturma yetenekleri sayesinde büyük ilgi görmektedir. Ancak, yüksek hesaplama maliyeti gerektirmeleri, bazen kararlı bir şekilde öğrenememeleri ve eğitim sürecinde istikrarsızlık yaşamaları gibi dezavantajları bulunmaktadır.

Derin Örtüşük Üretici Ağlar (Deep Convolutional Generative Adversarial Networks - DCGAN), GAN'ların bir türevidir ve özellikle görüntü üretimi ve sentezi için optimize edilmiştir(Şekil 9). DCGAN, geleneksel GAN mimarisine derin evrişimsel katmanlar ekleyerek, üretilen görüntülerin daha keskin ve gerçekçi olmasını sağlar (Radford vd., 2016). DCGAN, özellikle yüz görüntüsü üretimi, veri artırma, stil transferi ve tıbbi görüntü sentezi gibi alanlarda yaygın olarak kullanılır. Avantajları arasında derin öğrenme tabanlı güçlü görüntü üretme yeteneği, özniteliklerin hiyerarşik öğrenimi ve istikrarlı eğitim süreci yer alırken, dezavantajları arasında mod çökmesi (mode collapse), eğitim sürecinin hassasiyeti ve büyük veri gereksinimi bulunur (Wang vd., 2018) .



Şekil 9. DCGAN modelinde Üretici (üstteki) ve Ayrıştırıcının(alttaki) mimarisi

Eğitim süreci parametreleri AE modelinin eğitim parametreleriyle aynı olacak şekilde ayarlanmıştır. Eğitim 2000 döngü boyunca yürütülmüş ve her bir güncellemede 64 boyutunda veri yığınları kullanılmıştır. Öğrenme oranı 0.0001, momentum katsayıları ise sırasıyla $\beta1=0.5$ ve $\beta2=0.999$ olarak belirlenmiştir. Latent vektörün boyutu 10 olarak seçilmiş ve gri tonlamalı (1 kanal) 150×150 çözünürlüklü görüntülerle çalışılmıştır. L1 + BCE kayıp fonksiyonu kombinasyonu kullanılmıştır.

2.5. Değerlendirme Metrikleri ve Karşılaştırma

AAE, Sınır Arayan Çekişmeli Üretici Ağ (Boundary-Seeking Generatıve Adversarıal Networks – BGAN), Sınır Dengesi Çekişmeli Üretici Ağ (Boundary Equilibrium Generative Adversarial Networks – BEGAN) ve Derin Evrişimsel Çekişmeli Üretici Ağ (Deep Convolutional Generative Adversarial Networks - DCGAN) modelleri bu çalışmada kullanılmıştır.

PSNR (Peak Signal-to-Noise Ratio), bir görüntünün kalite düzeyini ölçmek amacıyla yaygın olarak kullanılan objektif bir metriktir. Bu metrik, elde edilen (yeniden oluşturulan ya da işlenmiş) görüntüyü orijinal referans görüntüyle piksel düzeyinde karşılaştırarak, görüntüdeki bozulma oranını sayısal olarak ifade eder. PSNR değeri ne kadar yüksekse, elde edilen görüntünün referans görüntüye o kadar yakın olduğu ve dolayısıyla bozulmanın düşük seviyede gerçekleştiği kabul edilir. Ancak PSNR metriği,

yalnızca referans görüntünün mevcut olduğu durumlarda kullanılabilir; bu nedenle uygulama alanı, referans görüntünün bulunabildiği senaryolarla sınırlıdır.

Öte yandan BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator – Kör/Referanssız Görüntü Mekansal Kalite Değerlendiricisi) metriği, adından da anlaşılacağı üzere, herhangi bir referans görüntüye ihtiyaç duymadan görüntü kalitesini değerlendirebilen bir ölçüm yöntemidir. BRISQUE, görüntünün doğal sahne istatistiklerini analiz ederek, uzamsal alanda meydana gelen bölgesel bozulmaları istatistiksel yöntemlerle değerlendirir ve bu doğrultuda insan görsel algısına yakın bir kalite tahmini sunar. Bu özellikleri sayesinde BRISQUE, özellikle gerçek hayattaki referanssız görüntü işleme senaryolarında tercih edilen güçlü bir değerlendirme metriği olarak öne çıkmaktadır.

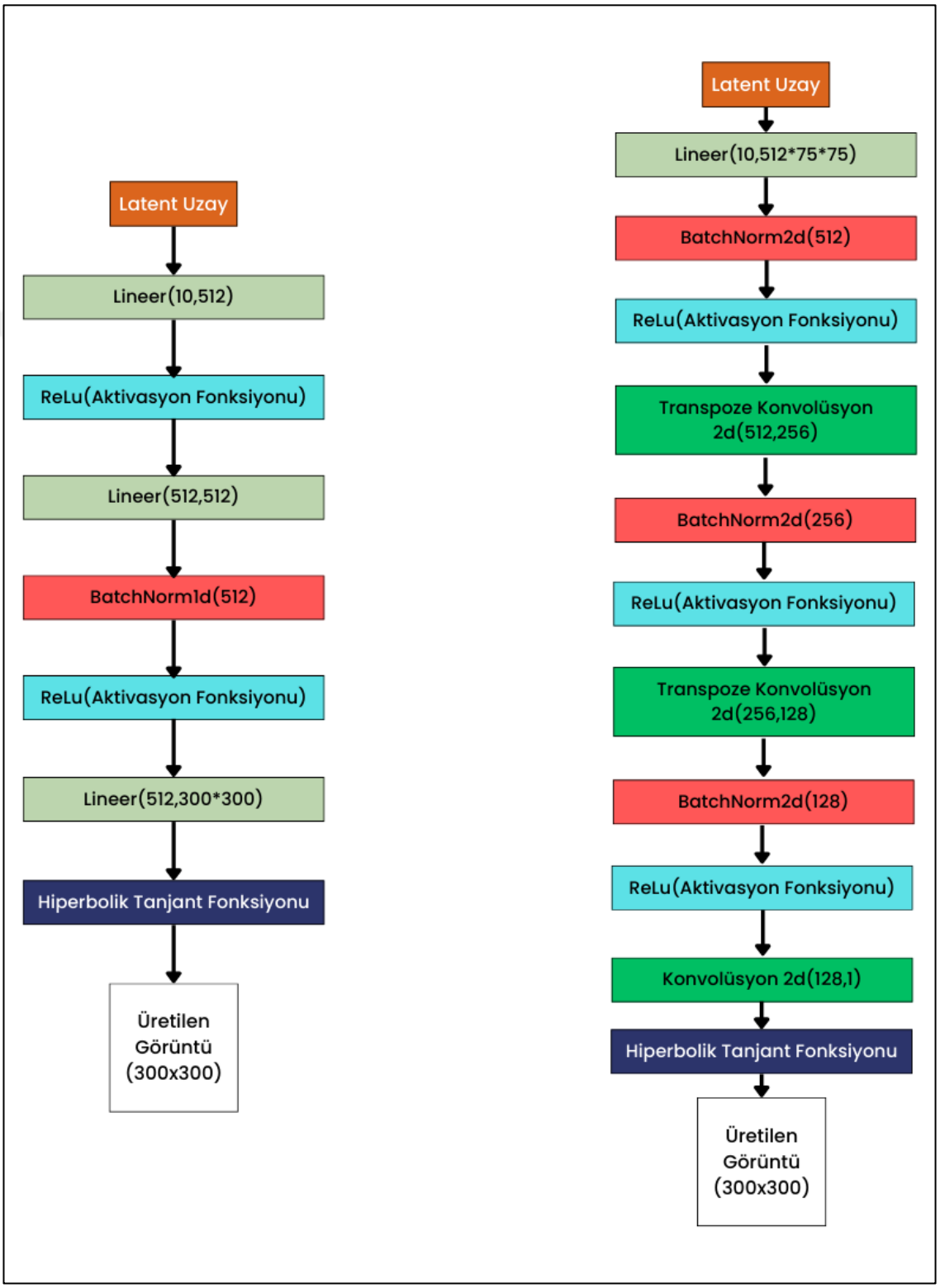
Her bir modelin eğitimi için girdi olarak 150x150 çözünürlüklü 235 farklı kişiye ait toplamda 1391 adet IITD sağ el avuç içi görüntüsü bmp formatında gri tonlamalı olarak verilmiştir. Modeller 2000 epoch boyunca 0.0001 öğrenme oranıyla NVIDIA Geforce RTX 2070 SUPER model Grafik İşlem Birimi (Graphics Processing Unit- GPU) kullanılarak eğitilmiştir. Bu eğitimin sonuçları, "BULGULAR VE SONUÇLAR" bölümünde incelenecektir. AAE Modelinin PSNR ve BRISQUE ölçeği sonuçları ve grafik değerlerinin diğer modellere göre üstün olması nedeniyle çalışmanın devamında AAE modeli üzerinde yoğunlaşılmıştır.

2.6. Geliştirilmiş AAE Kod Çözücü

Çalışmada, AAE yapısında kullanılan tam bağlantılı katmanlardan oluşan standart kod çözücü mimarisi yerine iki boyutlu konvolüsyonel katmanlar içeren geliştirilmiş bir kod çözücü mimarisi önerilmektedir(Şekil 10).

Standart AAE kod çözücü yapısında, gizli uzaydan elde edilen z vektörü ardışık lineer (nn.Linear) katmanlardan geçirilerek doğrudan bir vektör haline getirilmekte ve sonrasında hedef görüntü boyutuna ölçeklenmektedir. Bu yöntem düşük çözünürlüklü veri üretimi için basit ve hızlı bir çözüm sunarken, bazı sınırlamalara sahiptir: Tam bağlantılı katmanlar, girdiye ait yerel uzamsal ilişkileri koruyamadığından, özellikle karmaşık yapılı ve detaylı görüntülerin yeniden inşasında kalite kaybına neden olur. Ayrıca, yüksek çözünürlüklü görüntüler üretmek istendiğinde her piksel için ayrı ağırlık öğrenilmesi gerektiğinden, modelin parametre sayısı hızla artar ve bu durum öğrenmeyi zorlaştırır. Bununla birlikte, bu katmanlar kenar ve doku gibi yüksek frekans bileşenlerini temsil etmede yetersiz kalarak detay eksikliğine yol açar.

Önerilen yeni kod çözücü yapısı, örtük vektörün önce bir lineer katman ile düşük çözünürlüklü bir özellik haritasına (feature map) dönüştürülmesini sağlar. Daha sonra bu harita, ardışık transpoze konvolüsyon (dekonvolüsyon) katmanları ve normal konvolüsyon katmanları ile aşamalı olarak büyütülür.



Şekil 10. Standart AAE Kod Çözücü mimarisi (solda) ve Geliştirilmiş AAE Kod Çözücü mimarisi (sağda)

Konvolüsyonel işlemler sayesinde verinin lokal yapısı korunarak, özellikle avuç içi gibi karmaşık desenlerin detaylı rekonstrüksiyonu mümkün hale gelir. Bu işlemler, ağırlıkları uzamsal boyutlarda paylaşarak parametrelerin daha verimli kullanılmasını sağlar ve modelin yüksek çözünürlüklü görüntüler üretirken kompakt kalmasına katkıda bulunur. Lokal özelliklerin öğrenilmesi, modelin farklı avuç içi yapılarında detay kaybetmeden üretim yapabilmesini sağlayarak genel geçerliği artırır. Ayrıca, çözünürlük seviyeleri her adımda istikrarlı biçimde yükseltildiği için, çıkış görüntülerinde bulanıklık, mozaiklenme veya yapaylık gibi istenmeyen bozulmalar minimize edilir.

Bu deneyde, düşük çözünürlüklü avuç içi görüntülerinin süper çözünürlük yoluyla kalitesinin artırılması hedeflenmiştir. Bu amaç doğrultusunda, temel AAE mimarisi, çıkış aşamasında konvolüsyonel katmanlar içerecek şekilde geliştirilmiştir. Önerilen kod çözücü mimarisi, örtük vektörlerden 75×75 çözünürlükte bir öznitelik haritası üretmekte ve bu haritayı kademeli olarak ConvTranspose2D katmanlarıyla 300×300 boyutuna yükseltmektedir. Böylece, ağın daha yapısal ve ayrıntılı görüntüler üretmesi sağlanmaktadır. Ayrıca, modelin ifade kapasitesini artırmak ve daha zengin temsiller öğrenmesini sağlamak amacıyla gizli uzay boyutu 128'e yükseltilmiştir. Bu sayede modelin, daha ayrıntılı yapıları kodlayarak süper çözünürlük sürecinde daha gerçekçi çıktılar üretebilmesi hedeflenmiştir.

Modelin eğitimi 200 eğitim döngüsü boyunca, 8'lik küme boyutu ile gerçekleştirilmiştir. Öğrenme oranı olarak 0.0001 seçilmiştir. Tüm giriş görüntüleri gri tonlamalı (1 kanal) olarak 150×150 çözünürlüğe normalize edilmiştir. Model, bu görüntülerden yüksek çözünürlüklü çıktılar üretmeyi öğrenmiştir.

Eğitim sürecinde, modelin üretim kalitesini izlemek adına her bir eğitim döngüsü sonunda elde edilen g_loss değerleri kaydedilmiştir. Bu değerler, "Bulgular ve Sonuçlar" bölümünde incelenip yorumlanacaktır.

2.7. Düşük Çözünürlüklü Görüntülerin Üretilmesi ve Ölçekleme Modelleri

Bu çalışmada, modelin eğitiminde ve değerlendirilmesinde düşük çözünürlüklü görüntülerin yapay olarak üretilmesi önemli bir adımı oluşturmaktadır. Özellikle modelin çıktıları ile gerçek görüntüler arasındaki farkları doğru biçimde ölçebilmek için, çözünürlük düşürme işlemi uygulanmıştır. Bu işlem sırasında yüksek çözünürlüklü giriş görüntüleri, bilinear interpolasyon yöntemiyle 75×75 piksele indirgenmiş ve kayıp fonksiyonlarının hesaplanmasında bu düşük çözünürlüklü versiyonlar referans olarak kullanılmıştır. Böylece, modelin çıkışı ile hedef görüntü arasında doğrudan karşılaştırma

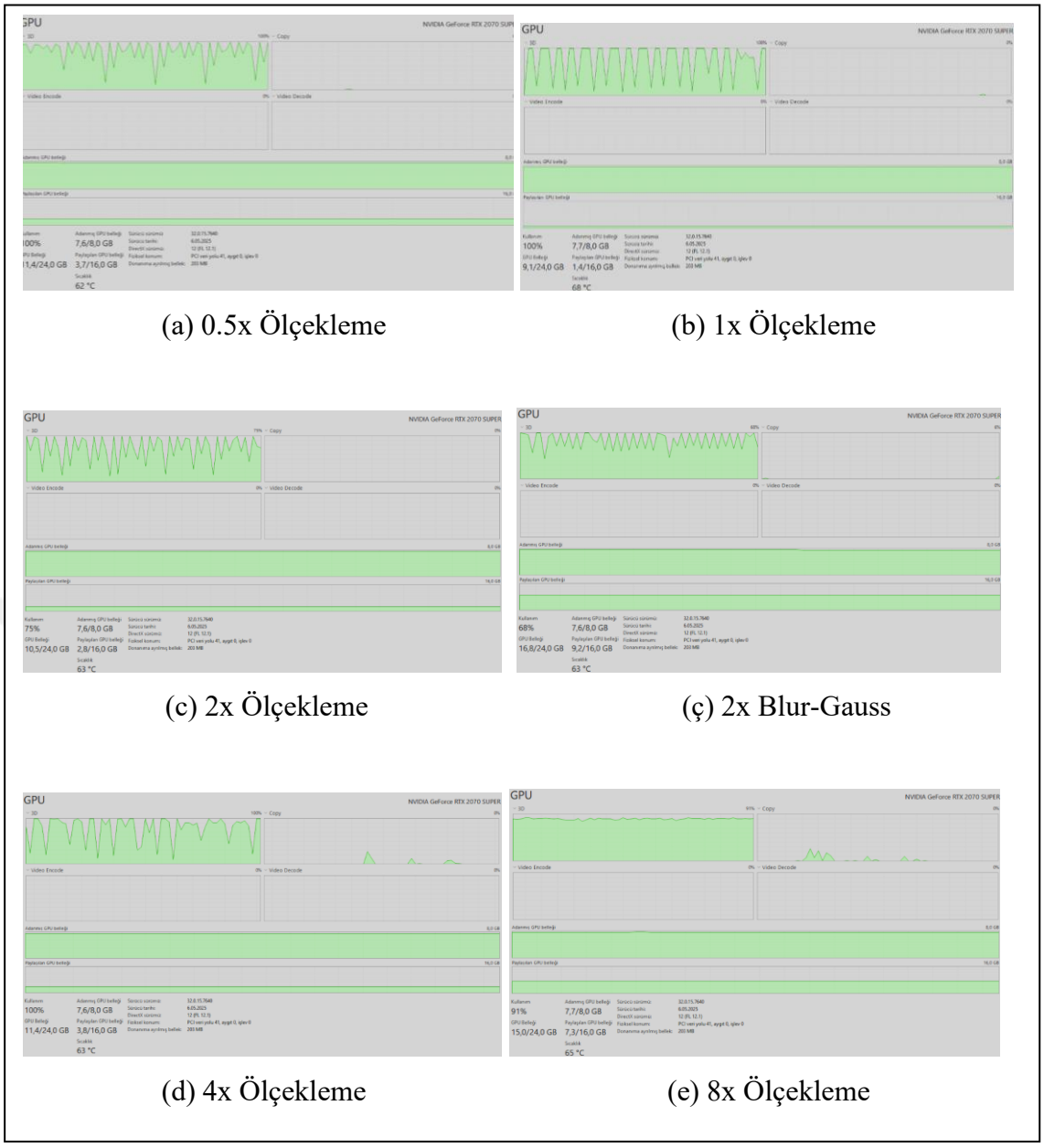
yapılabilmiş hem piksel temelli L1 kaybı hem de PSNR metriği daha anlamlı hale getirilmiştir.

Görüntülerin yalnızca çözünürlük açısından değil, içerik kalitesi açısından da bozulduğu senaryolar oluşturulmuştur. Bu kapsamda, gürültü ve bulanıklık gibi yaygın bozulma türleri kullanılarak eğitim verisi çeşitlendirilmiştir. Rastgele seçilen görüntülere Gauss gürültüsü, tuz-biber gürültüsü ve bulanıklaştırma filtreleri uygulanmış, her bireye ait bir görüntü temiz bırakılırken kalanlar bu yöntemlerle bozulmuştur. Bu yaklaşım, modelin bozulmuş veri karşısındaki dayanıklılığını test etmek ve daha gerçekçi bir eğitim ortamı sağlamak amacıyla tercih edilmiştir.

Modelin çözünürlük artırma kapasitesi 2×, 4× ve 8× olmak üzere üç farklı seviyede test edilmiştir. Her bir çözünürlük oranı için kod çözücü yapısı yeniden tasarlanmış ve çıkış çözünürlüğüne ulaşmak amacıyla farklı sayıda transpoze konvolüsyon katmanları kullanılmıştır. 2× çözünürlük deneyinde 75×75'lik temsil, iki aşamalı katmanlarla 300×300 boyutuna çıkarılmıştır. 4× modelde bu yapı üç aşamaya çıkarılmış ve 600×600 çözünürlüğe ulaşılmıştır. 8× modelde ise dört katmanlı daha derin bir yapı ile 1200×1200 çözünürlüğe kadar büyütme gerçekleştirilmiştir. Bu farklı mimari tasarımlar, çözünürlük düzeyi arttıkça hem hesaplama yükünü hem de model karmaşıklığını önemli ölçüde artırmıştır. Özellikle 8× deneyinde eğitim süresinin ve GPU kullanımının ciddi şekilde yükseldiği gözlemlenmiştir.

2.8. Hesaplama Performansı ve Kaynak Kullanımı

Bu çalışmada kullanılan AAE tabanlı süper çözünürlük modelleri, NVIDIA GeForce RTX 2070 SUPER GPU donanımı üzerinde eğitilmiştir. Eğitim sürecinde, farklı çözünürlük ölçekleri (0.5×, 2×, 4× ve 8×) için GPU kullanım durumu, bellek tüketimi ve işlem süresi dikkatle izlenmiş ve kayıt altına alınmıştır. Elde edilen GPU performans ekran görüntüleri Şekil arasında sunulmaktadır.



(a) 0.5x Ölçekleme (b) 1x Ölçekleme

(c) 2x Ölçekleme (ç) 2x Blur-Gauss

(d) 4x Ölçekleme (e) 8x Ölçekleme

Şekil 11. AAE modeli eğitimi sırasında GPU kullanım durumu

Ölçekleme katsayısı arttıkça modelin parametre sayısındaki artış ve eğitim verisinin detay seviyesi, GPU kaynak kullanımını doğrudan etkilemiştir. Özellikle 8× ölçekleme denemesinde GPU bellek kullanımının 11.4 GB'a kadar ulaştığı ve GPU kullanım oranının %100'e yakın seyrettiği gözlemlenmiştir. Bu durumda GPU sıcaklığı 65 °C seviyelerinde stabilize olmuştur. Diğer ölçeklemelerde de GPU kullanımı genel olarak %75–100 aralığında gerçekleşmiştir. Bu gözlemler Tablo 1'de özetlenmiştir.

Tablo 1. Farklı ölçekleme düzeylerinde eğitim süresi ve GPU kaynak kullanımı

Ölçekleme	Eğitim Süresi	GPU Kullanımı	Bellek Tüketimi	Sıcaklık (ort.)
0.5×	~6 saat	%100	7.6–11.4 GB	62 °C
2×	~13 saat	%100	7.7–9.1 GB	62 °C
4×	~25 saat	%75	10.5 GB	63 °C
8×	~52 saat	%68–91	15.0 GB	65 °C

Ayrıca, 2× ölçekli modelin bir varyantında, eğitim verisine tuz-biber, Gauss gürültüsü ve bulanıklaştırma gibi bozulmalar eklenerek modelin dayanıklılığı test edilmiştir. Bu eğitim sırasında GPU kullanımı %100 seviyesine ulaşmış; bellek tüketimi 11.4 GB'a ve sıcaklık değeri 63 °C'ye kadar çıkmıştır. Bu durum, bozulmuş veri ile yapılan eğitimin daha yoğun işlem gücü gerektirdiğini göstermektedir.

Sonuç olarak, kullanılan modellerin GPU kaynaklarını yüksek düzeyde verimli kullandığı; bellek kapasitesi, sıcaklık ve kullanım oranı açısından sistemin kararlı çalıştığı gözlemlenmiştir. Bu durum, önerilen yöntemlerin orta seviyeli GPU sistemlerinde uygulanabilirliğini ve ölçeklenebilirliğini desteklemektedir.

3.BULGULAR VE SONUÇLAR

3.1. Nicel Performans Sonuçları

Tablo 2 (Macan vd., 2024)'deki ölçümler ve Şekil 12'deki grafiklerde görüldüğü gibi, elde edilen sonuçlar, dört farklı modelin (AAE, BEGAN, DCGAN ve BGAN) süper çözünürlük görevindeki başarımlarını hem nicel metrikler hem de eğitim süreci boyunca elde edilen kayıplar ve kalite göstergeleri bakımından karşılaştırmalı olarak değerlendirme imkânı sunmaktadır.

AAE modeline ait grafiklerde BRISQUE değerinin eğitim süreci boyunca büyük oranda azaldığı, PSNR değerinin ise istikrarlı şekilde arttığı gözlemlenmektedir. Bu durum, AAE'nin hem yapay hem de gerçekçi görsel kaliteyi optimize etmede başarılı olduğunu göstermektedir. BRISQUE dalgalanmaları düşük düzeyde seyretmiş ve özellikle 2000 epoch civarında oldukça dengeli sonuçlar elde edilmiştir. Ayrıca D ve G kayıplarının düşük ve kararlı olması, modelin discriminator ve generator arasındaki dengeyi başarıyla koruduğunu ortaya koymaktadır.

BEGAN modelinde (Şekil 12.b) PSNR değerleri düşük kalmasına rağmen BRISQUE ölçütünde göreceli olarak daha pozitif bir seyir izlenmiştir. Ancak grafikte dikkat çeken en önemli unsur BRISQUE değerinin çok keskin dalgalanmalar göstermesi, modelin bazı dönemlerde görüntü kalitesini kaybettiğine işaret etmektedir. Bu da modelin kararlılık sorunları yaşadığını düşündürmektedir. G kaybının düşük olması, çıktının discriminator açısından yeterince gerçekçi bulunduğunu gösterse de, PSNR'nin ve BRISQUE'in birlikte değerlendirildiğinde bu gerçekçiliğin yapısal doğruluk açısından zayıf olabileceğini göstermektedir.

BGAN modeli (Şekil 13.c), genel olarak yüksek G loss ve çok değişken BRISQUE değerleri ile dikkat çekmektedir. Özellikle reconstructed BRISQUE değerlerinin yüksek olması ve negatif değerlere inmemesi, bazı durumlarda görüntülerin aşırı düzeyde yapaylaştığını göstermektedir. Bununla birlikte PSNR değeri DCGAN'a kıyasla biraz daha yüksek seviyede sabitlenmiştir. Bu durum, BGAN'ın çıkışlarında gürültü azaltımı yerine kontrastı artırmaya yönelik eğilimleri olabileceğini düşündürmektedir.

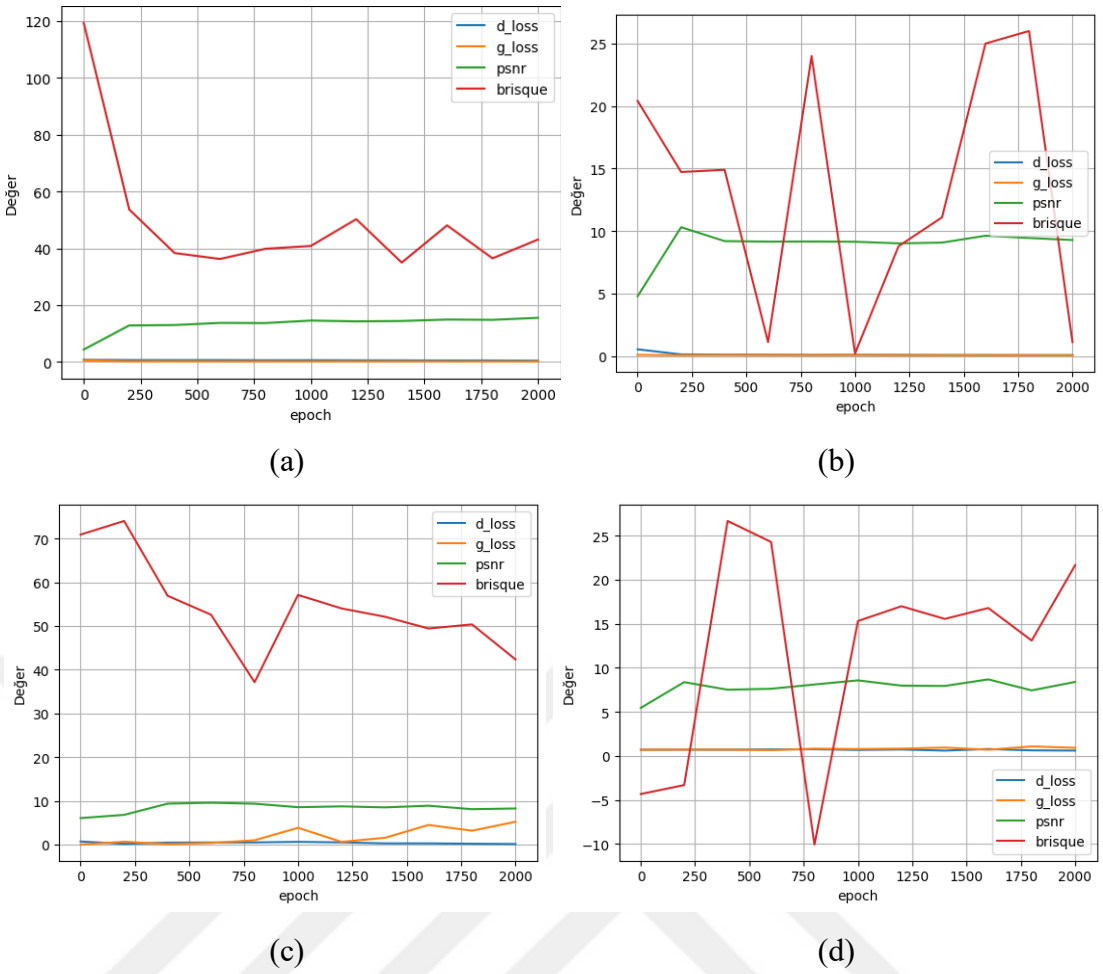
DCGAN ise görsel kalite açısından en dengesiz sonuçları vermiştir. BRISQUE değerleri birçok noktada negatif değerlere inmiş ve bu durum yapay görüntülerin referanssız kalite ölçümünde zayıf bulunduğunu ortaya koymuştur. Ayrıca PSNR değerleri de diğer modellere kıyasla en düşük seviyelerde seyretmektedir. Grafiklerde de görüldüğü üzere BRISQUE değişimleri oldukça yüksek genlikli dalgalanmalar

göstermektedir. Bu sonuç, DCGAN'ın SR görevine doğrudan adapte edilmediği durumlarda öğrenme sürecinin kararsızlaşabileceğine işaret etmektedir.

Sonuç olarak, genel kalite değerlendirmesi açısından AAE modeli tüm metriklerde en istikrarlı ve dengeli sonuçları sunmuştur. BRISQUE ve PSNR gibi tamamlayıcı kalite göstergelerinde sağlanan uyum, AAE'nin hem yapay görüntü üretiminde hem de yapısal benzerlik bakımından daha başarılı olduğunu göstermektedir. Diğer modellere kıyasla daha düşük kayıp değerleriyle ve daha kararlı bir öğrenme eğrisiyle öne çıkmaktadır. Bu bağlamda, biyometrik süper çözünürlük uygulamaları gibi hassas alanlarda AAE tabanlı çözümlerin daha tutarlı ve güvenilir sonuçlar üretebileceği değerlendirilmektedir.

Tablo 2. DCGAN, BGAN, BEGAN ve AAE ölçüm sonuçları

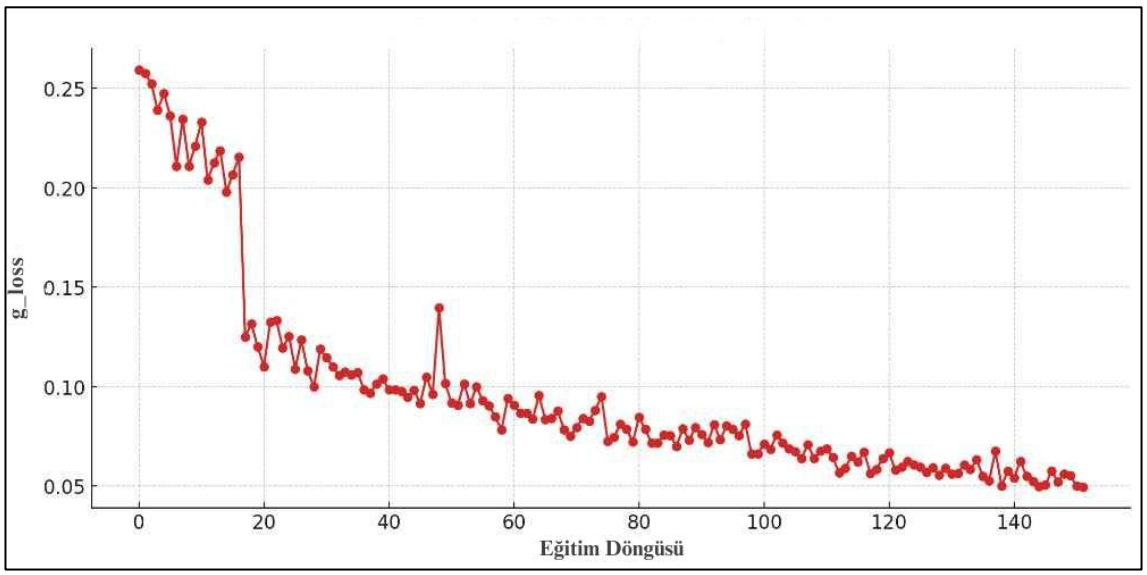
Modeller ve Ölçümler	**Ölçümler**					
	Epoch	***D Loss***	***G Loss***	***PSNR***	***Brisque Original***	***Brisque Reconstructed***
AAE	1945	0.48	0.13	14.7	35.66	38.32
AAE	1963	0.44	0.127	13.8	57.33	42.65
AAE	2000	0.36	0.123	15.49	54.46	36.24
BEGAN	1986	0.82	0.065	9.27	37.68	28.27
BEGAN	1995	0.86	0.064	9.276	30.43	31.8
BEGAN	2000	0.84	0.064	9.28	46.5	1.76
DCGAN	1986	0.54	0.9	8.29	48.15	-24.04
DCGAN	1995	0.51	0.97	8.81	39.28	-23.60
DCGAN	2000	0.59	0.91	8.12	38.80	-21.67
BGAN	1986	0.15	4.34	8.79	24.90	45.37
BGAN	1995	0.23	3.87	8.49	43	49
BGAN	2000	0.14	8.9	8.4	37.01	44.3



Şekil 12. Yöntemlere göre 2000 döngü eğitimindeki metriklerin ilerleyişi a) AAE, b) BEGAN, c) BGAN ve d) DCGAN

3.2. Geliştirilmiş Modelin Eğitimi

Şekil 13'teki grafik, bu g_loss değerlerinin eğitim döngüsü bazında değişimini göstermektedir.



Şekil 13. Geliştirilmiş Kod Çözücülü AAE eğitimi sonucu Kod Çözücü Kayıp Değerleri

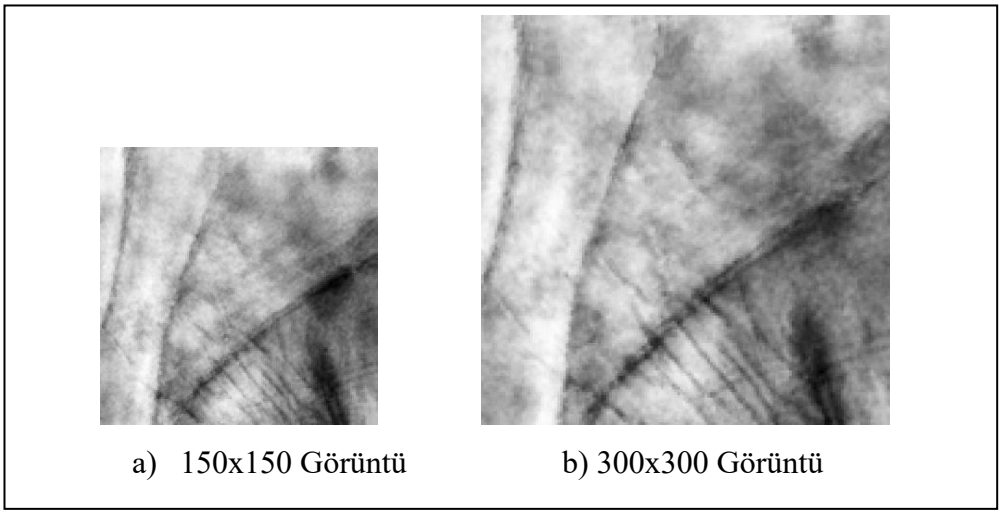
Elde edilen grafik incelendiğinde, g_loss değerinin eğitim boyunca belirgin bir azalma eğilimi sergilediği görülmektedir. Başlangıçta 0.259 seviyelerinde olan g_loss, eğitim sonunda 0.049 seviyelerine kadar gerilemiştir. Bu düşüş, modelin eğitim süreci boyunca hedeflediği dağılıma daha iyi yakınsadığını ve kod çözücü tarafından daha az ayırt edilebilecek çıktılar ürettiğini göstermektedir. Yani model, düşük çözünürlüklü girdilerden yüksek çözünürlükte ve gerçekçi görünümlü çıktılar üretme konusunda önemli bir iyileşme göstermiştir.

g_loss eğrisinin düzenli bir biçimde azalma göstermesi, modelin istikrarlı bir şekilde öğrenme gerçekleştirdiğine ve aşırı dalgalanmaların yaşanmadığına işaret etmektedir. Bu durum aynı zamanda öğrenme oranı, küme boyutu gibi hiperparametrelerin iyi seçilmiş olduğunu ve modelin kararlı bir optimizasyon süreci geçirdiğini desteklemektedir.

3.3. Farklı Parametreler Altında Model Performansı

AAE mimarisinin kod çözücü bölümü, modelin 150×150 boyutundaki giriş görüntülerinden 300×300 çözünürlüğünde çıktılar üretebilmesini sağlayacak şekilde tasarlanmıştır. Bu amaçla, kod çözücü katmanında kullanılan ConvTranspose2d (transpoze konvolüsyon) katmanları ile kademeli olarak çözünürlük artırımı gerçekleştirilmiştir (Şekil 14).

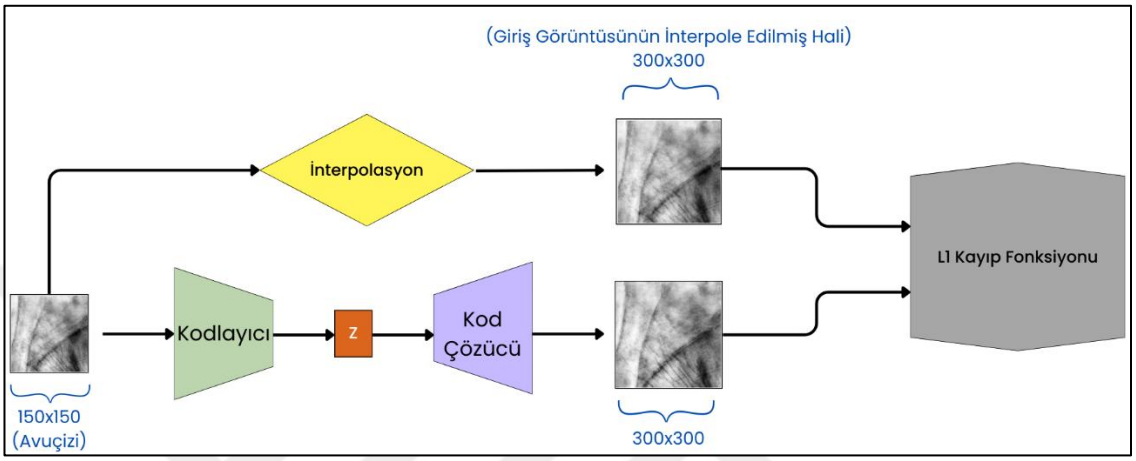
Modelin çıktı boyutunun girişten daha büyük olması, temel AE yapısına göre bir farklılık teşkil etmektedir. Ancak, L1 kayıp fonksiyonu piksel düzeyinde iki görüntü arasındaki farkları hesapladığı için, bu fonksiyonun uygulanabilmesi adına karşılaştırılan görüntülerin aynı boyutta olması gerekmektedir.



a) 150x150 Görüntü b) 300x300 Görüntü

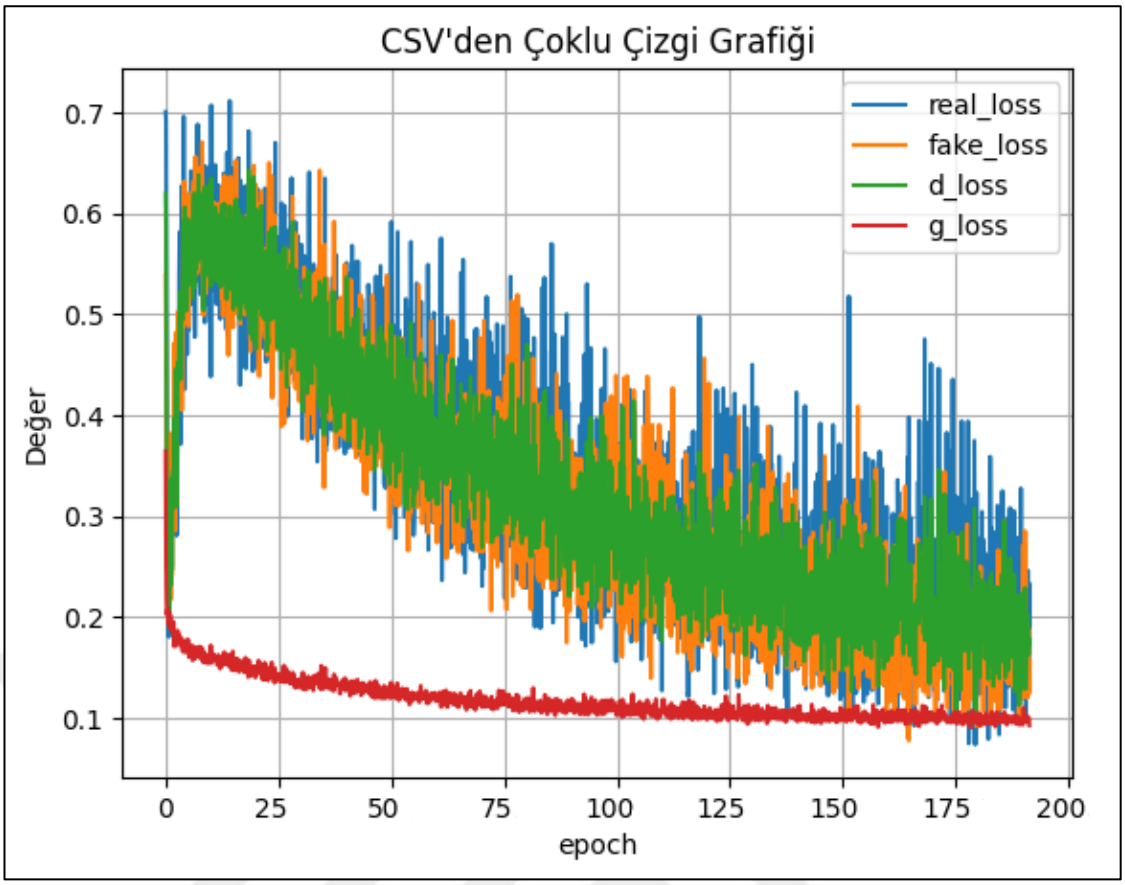
Şekil 14. Orijinal(a) ve boyutu artırılmış(b) AAE çıktı görüntüleri

Bu nedenle, L1 kaybını hesaplayabilmek amacıyla, giriş olarak kullanılan 150×150 boyutundaki görüntüler, bikübik interpolasyon yöntemiyle 300×300 boyutuna ölçeklendirilmiştir. Böylece hem modelin ürettiği 300×300 boyutundaki çıktı, hem de interpolasyonla büyütülen 300×300 boyutundaki giriş görüntüsü karşılaştırılarak L1 kayıp değeri elde edilmiştir. Bu yaklaşım sayesinde hem süper çözünürlük hedeflenmiş hem de ağın eğitiminde daha istikrarlı bir öğrenme süreci sağlanmıştır (Şekil 15).



Şekil 15. İnterpolasyon ve Kod Çözücü çıkışlarının L1 Kayıp Fonksiyonu ile karşılaştırılması

Bu çalışmada, avuç içi görüntülerinin süper çözünürlükle iyileştirilmesi amacıyla geliştirilen AAE modeli, 150×150 çözünürlükteki giriş görüntülerinden 300×300 çözünürlükte çıktılar üretmek üzere tasarlanmıştır. Modelin eğitim sürecinde, piksel düzeyinde doğruluğu sağlayan L1 kayıp fonksiyonu ile sahte ve gerçek görüntüler arasındaki ayrımı teşvik eden BCE kayıp fonksiyonu birlikte kullanılmıştır. Bu iki bileşenli kayıp fonksiyonu, hem görsel kaliteyi artırmayı hem de yapay görüntülerin gerçekliğe yakınlığını maksimize etmeyi hedeflemektedir. Model, toplamda 200 döngü boyunca eğitilerek hem üretici hem de ayrıştırıcı bileşenlerinin performansı izlenmiş ve bu süreçte elde edilen kayıp değerleri detaylı şekilde analiz edilmiştir (Şekil 16). Aşağıda bu analiz sonuçlarına dayalı olarak, eğitim sürecinin farklı aşamalardaki davranışları değerlendirilmiştir.

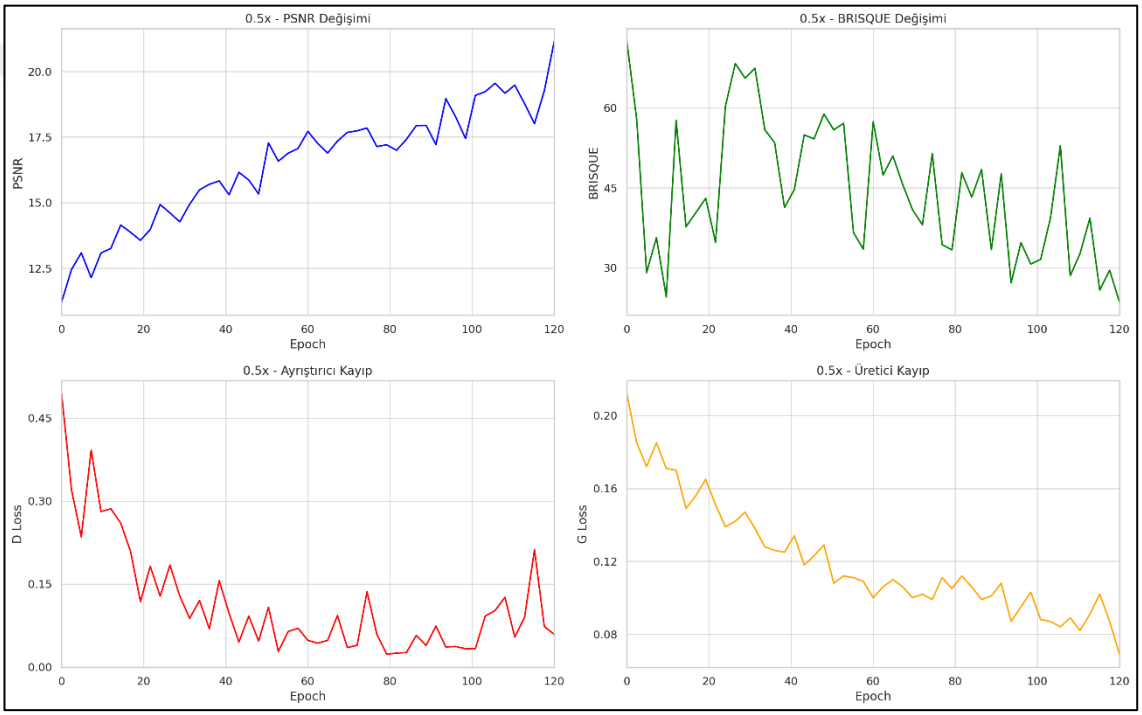


Şekil 16. AAE 2x Ölçekleme deneyi Kayıp Değerleri grafiği

Eğitim süreci genel olarak üç ana aşamada değerlendirilebilir: başlangıç, orta ve geç dönemler. Başlangıç aşamasında (yaklaşık 0–20 döngü aralığı), hem real_loss hem de fake_loss değerlerinde gözlemlenen yüksek dalgalanma, ayrıştırıcının gerçek ve sahte görüntüleri ayırt etmede zorlandığını ve modelin henüz kararlı bir öğrenme sürecine girmediğini göstermektedir. Bu dönemde g_loss'ta gözlenen hızlı düşüş ise, üreticinin erken evrede anlamlı görüntüler üretmeye başladığını işaret etmektedir. Orta aşamada (yaklaşık 20–100 döngü aralığı), ayrıştırıcı kayıplarının (özellikle d_loss) giderek daha istikrarlı bir şekilde azaldığı görülmekte olup, modelin ayrım yapma kabiliyetinin arttığını göstermektedir. Aynı zamanda g_loss değerinin daha yavaş fakat düzenli bir şekilde azalmaya devam etmesi, üreticinin daha kaliteli ve ayrıştırıcıyı yanıltabilecek düzeyde çıktılar üretmeye başladığını ortaya koymaktadır. Son olarak, geç aşamada (yaklaşık 100–200 döngü aralığı), d_loss değerinin oldukça düşük ve stabil bir seviyeye ulaşması, ayrıştırıcının güvenilir bir şekilde çalıştığını göstermektedir. Bu dönemde g_loss da yaklaşık 0.1 seviyelerinde sabitlenmiş olup, modelin başarılı bir şekilde öğrenme sürecini tamamladığına ve gerçekçi çıktılar üretebildiğine işaret etmektedir.

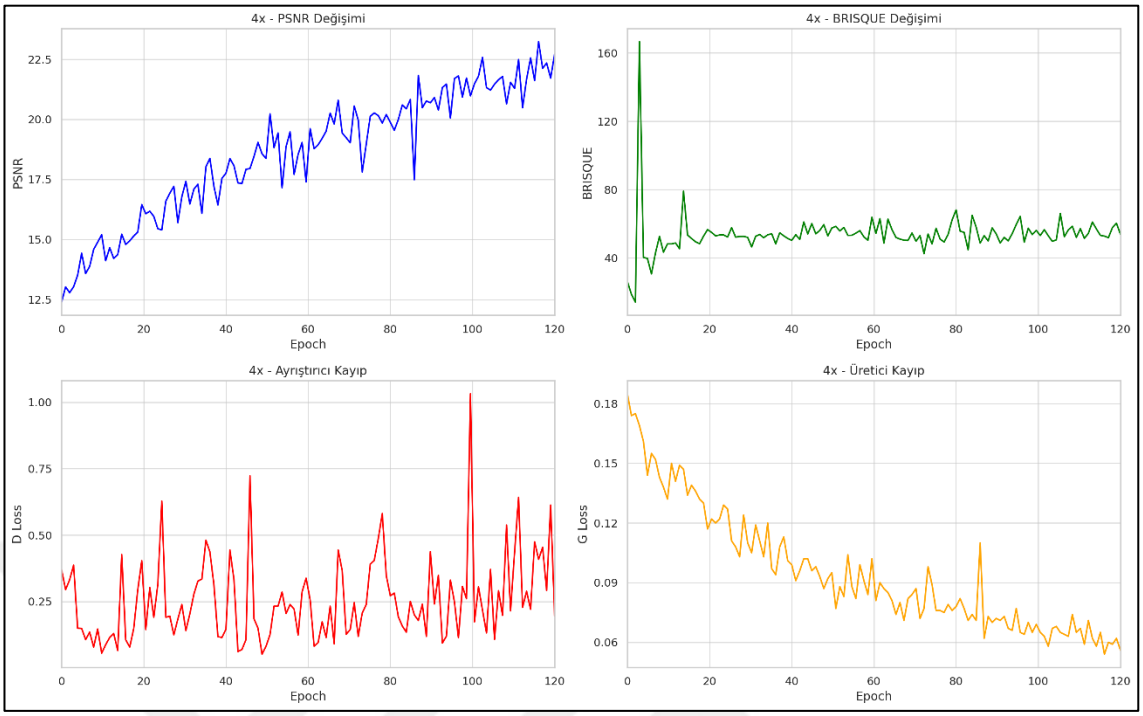
Şekil 17'da da görülen 0.5x çözünürlük seviyesinde eğitilen modelin metrik grafikleri, hem üretici hem de ayrıştırıcı ağların öğrenme süreçlerinde istikrarlı bir gelişim

gösterdiğini ortaya koymaktadır. PSNR eğrisi, düşük başlangıç değerlerinden itibaren kararlı bir şekilde artarak modelin zamanla daha kaliteli ve detaylı görüntüler üretmeyi öğrendiğini göstermektedir. BRISQUE skorları ise ilk döngülerde yüksek seviyelerde dalgalanma gösterse de, eğitim ilerledikçe düşüş eğilimi sergileyerek görüntü kalitesindeki iyileşmenin model tarafından başarılabildiğini yansıtmaktadır. Üretici kayıp eğrisi, başta dalgalı seyretmekle birlikte zamanla azalarak modelin gerçekçi çıktılar üretme yeteneğini pekiştirdiğini, ayrıştırıcı kaybın ise daha erken stabilize olarak üretici üzerindeki baskıyı sürdürdüğünü göstermektedir. Genel olarak, bu çözünürlük oranında modelin hem görüntü kalitesi metrikleri hem de kayıp fonksiyonları açısından dengeli ve öğrenmeye açık bir yapı sergilediği söylenebilir.



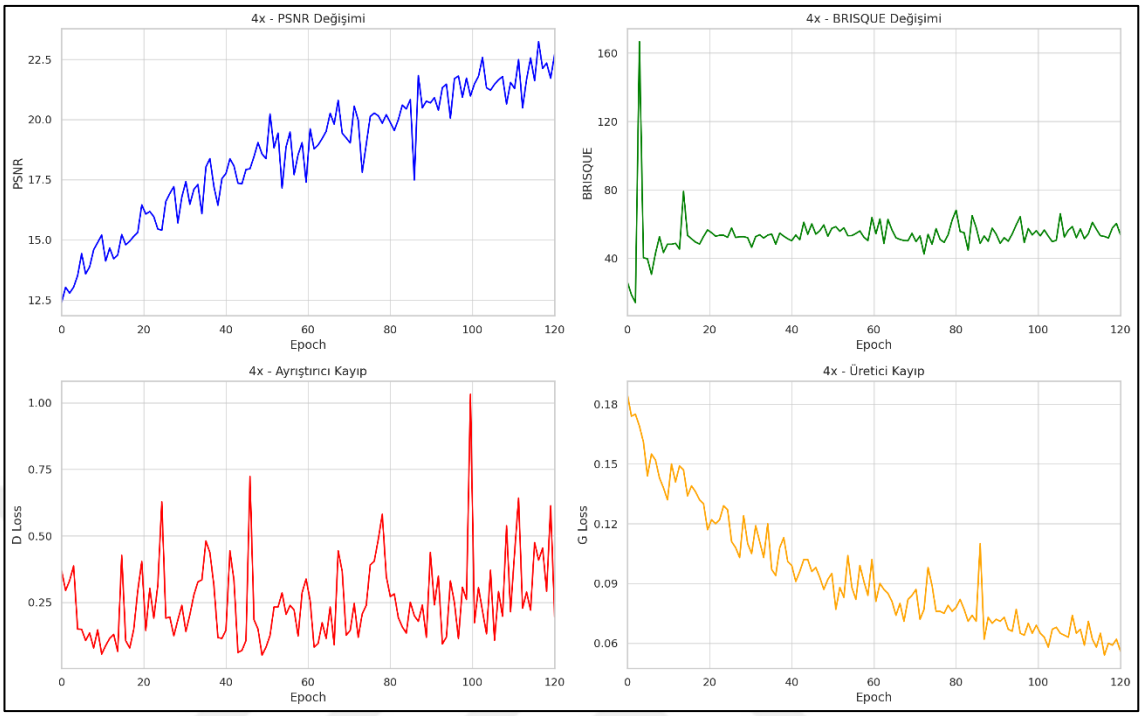
Şekil 17. 0.5x Ölçekleme eğitimi çıktı grafikleri

4x çözünürlük oranı ile eğitilen modelin metrik eğrileri(



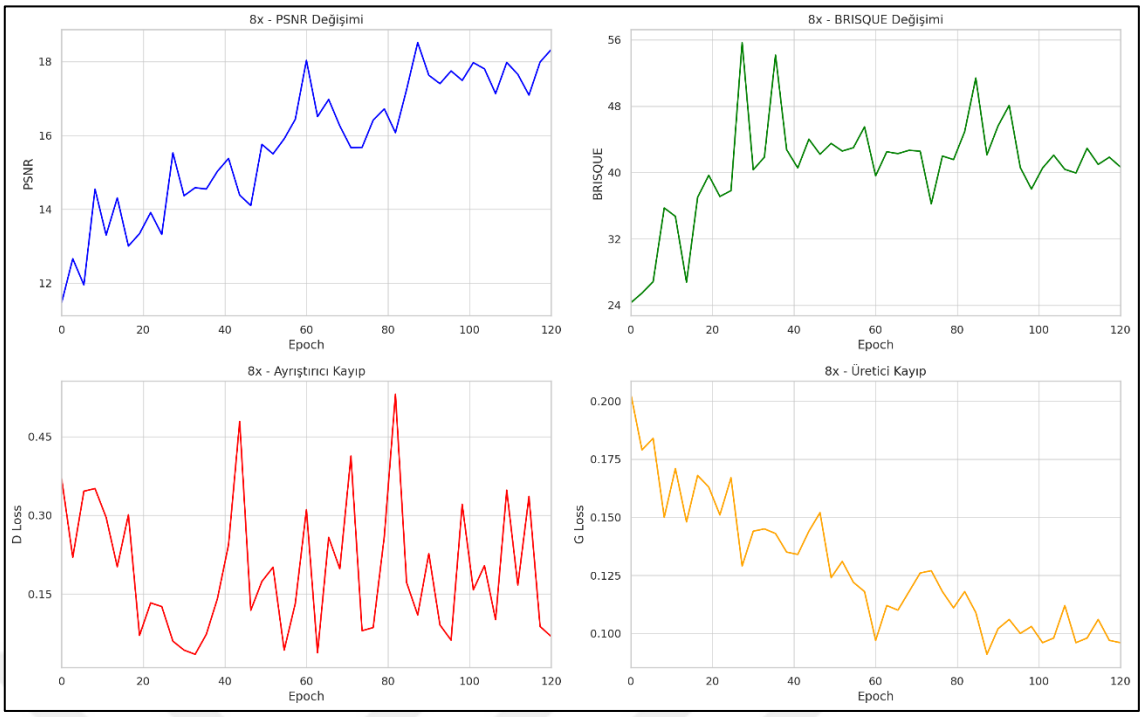
Şekil 18), yüksek çözünürlük üretimi için modelin daha yoğun öğrenme süreci geçirdiğini ve bu süreçte üretici ile ayrıştırıcı ağlar arasında denge kurma çabasının belirgin olduğunu göstermektedir. PSNR değerleri zamanla istikrarlı biçimde artarken, BRISQUE skorlarında özellikle ilk epochlarda dalgalı bir seyir gözlenmiş, ancak eğitim ilerledikçe görüntü kalitesinin iyileştiğini gösteren düşüşler gerçekleşmiştir. Ayrıştırıcı ağın kaybı başlangıçta dalgalı olsa da kısa sürede kararlılığa ulaşırken, üretici ağın kaybı daha uzun süre yüksek seviyelerde seyretmiş ve sonlara doğru daha düşük seviyelere inerek modelin kaliteli çıktı üretme potansiyelini yansıtmıştır. Bu metrik eğilimleri, 4x oranında yüksek çözünürlüklü ve detaylı görüntüler üretmenin daha karmaşık ve zaman alıcı bir süreç

olduğunu, ancak modelin öğrenme kapasitesi doğrultusunda bu hedefe kademeli şekilde ulaşabildiğini göstermektedir.



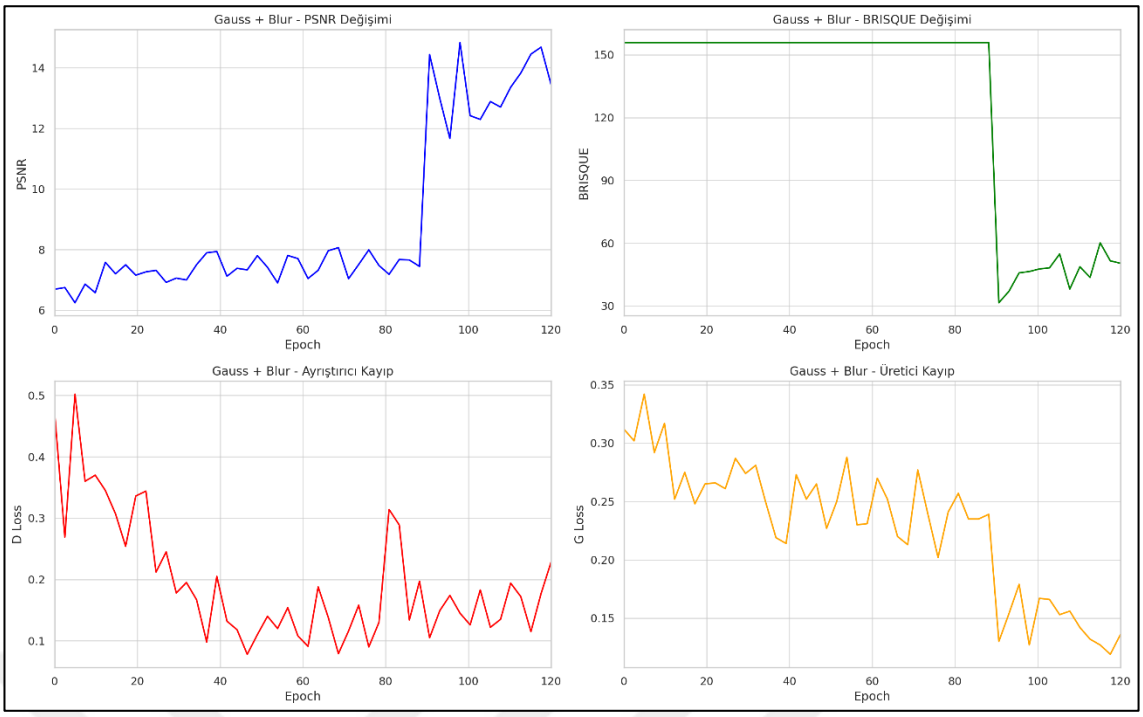
Şekil 18. 4x Ölçekleme eğitimi çıktı grafikleri

8x çözünürlük düzeyinde eğitilen modelin metrik eğrileri (Şekil 19), yüksek çözünürlüklü görüntü üretiminin model üzerinde ciddi bir öğrenme yükü oluşturduğunu açıkça göstermektedir. PSNR eğrisi genel olarak düşük seviyelerde seyredip yalnızca sınırlı artış göstererek modelin detay üretiminde zorlandığını ortaya koyarken, BRISQUE skorları ise belirli bir azalma eğilimi sergilese de daha düşük çözünürlüklü deneylerdeki kadar kararlı bir iyileşme göstermemiştir. Üretici kaybı başlangıçta yüksek seviyelerde dalgalanarak ağın gerçekçi görüntüler üretme sürecine adaptasyonda zorlandığını göstermiş, ancak son döngülerde göreli bir düşüş sergileyerek kısmen istikrar kazanmıştır. Ayrıştırıcı kayıplar ise üreticiye kıyasla daha erken stabilize olmuş ve modelin gerçek ile sahte görüntüleri ayırt etme kapasitesini daha hızlı geliştirdiğini göstermiştir. Bu sonuçlar, 8x seviyesinde süper çözünürlük üretiminin hem hesaplama kaynakları hem de mimari karmaşıklık açısından daha gelişmiş çözümler gerektirdiğini ve geleneksel yaklaşımların bu ölçekte sınırlı başarı gösterebildiğini ortaya koymaktadır.



Şekil 19. 8x Ölçekleme eğitimi çıktı grafikleri

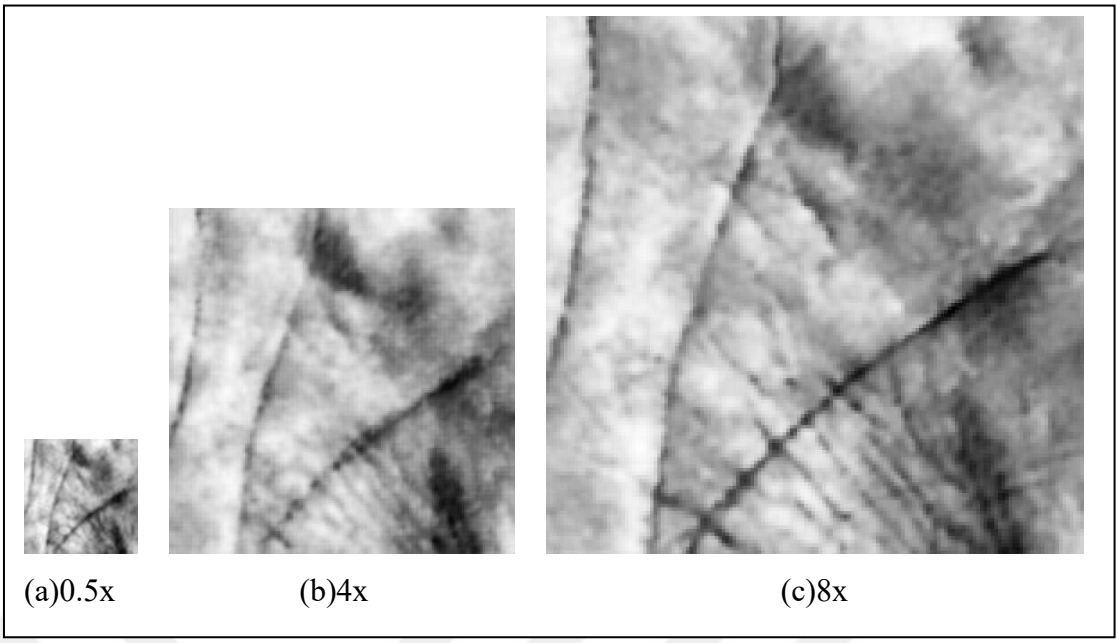
Gauss gürültüsü ve bulanıklık (blur) etkisi altında eğitilen modelin metrik grafikleri (Şekil 20), bozulmuş verilerle başa çıkma kabiliyetinin sınırlı ancak iyileştirilebilir olduğunu ortaya koymaktadır. PSNR eğrisi, klasik veri setlerine kıyasla daha düşük seviyelerde seyretmiş ve yalnızca sınırlı bir artış göstermiştir; bu durum, bozulmuş girişlerin modelin detaylı ve yüksek kaliteli çıktılar üretme kapasitesini zorladığını göstermektedir. BRISQUE skorları da benzer şekilde yüksek başlamış ve eğitim ilerledikçe dalgalanarak kademeli olarak azalmıştır; bu da görüntü kalitesinin zamanla iyileştiğini ancak bozulmalar nedeniyle istenen seviyeye tam olarak ulaşılamadığını göstermektedir. Üretici ağın kaybı başta oldukça yüksektir ve modelin bozuk verilerle gerçekçi çıktılar üretmekte zorlandığını göstermektedir, ancak eğitim ilerledikçe belirgin bir düşüş yaşanarak öğrenme gerçekleşmiştir. Ayrıştırıcı ağ ise üreticiye kıyasla daha erken stabilize olmuş ve bozuk veri ile üretilen sahte görüntüler arasındaki farkları daha hızlı şekilde öğrenebilmiştir. Genel olarak bu sonuçlar, modelin gürültü ve bulanıklık gibi bozulmalara karşı dayanıklılık kazandığını, ancak bu tür verilerin modelin öğrenme sürecini belirgin biçimde zorlaştırdığını göstermektedir.



Şekil 20. Gauss+Blur Veriseti ile eğitimin çıktı grafikleri

3.4. AE ve GAN Yaklaşımlarının Güçlü ve Zayıf Yönleri

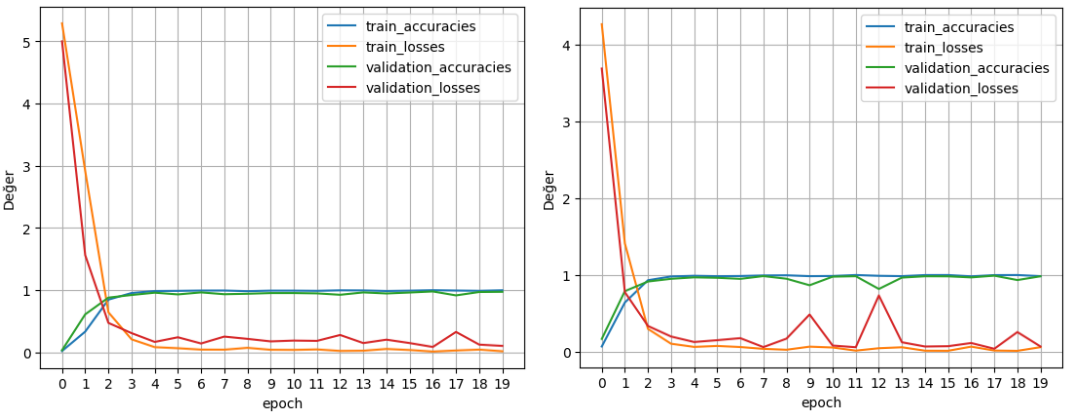
AE temelli mimariler, özellikle düşük çözünürlüklü veya bozulmuş görüntülerin temel yapısal bilgilerini yeniden inşa etme açısından son derece başarılı sonuçlar vermektedir(Şekil 21). Kodlayıcı kısmı giriş görüntüsünü sıkıştırılmış bir latent uzaya aktarırken, kod çözücü bu temsili yeniden yapılandırarak orijinal forma yakın bir çıkış üretir. Bu mekanizma sayesinde AE mimarileri, özellikle düşük çözünürlüklü (ör. 0.5x) ya da bulanıklaştırılmış veri senaryolarında (ör. Gauss+Blur) temel çizgi yapıları ve genel konturları başarıyla koruyarak anlamlı ve yapısal açıdan tutarlı görüntüler üretmiştir. Nitekim BRISQUE ve PSNR gibi metriklerde gözlemlenen kademeli iyileşmeler, AE mimarisinin özellikle “yeniden yapılandırma doğruluğu” bakımından güçlü olduğunu göstermektedir.



Şekil 21. AAE Çözünürlük Ölçekleme çıktı örnekleri

3.5. Uygulama Odaklı Değerlendirme

Modelin tanıma performansının ölçülmesinin temel amacı, geliştirilen AAE tabanlı görüntü iyileştirme sürecinin yalnızca görsel kaliteyi artırmakla kalmayıp, aynı zamanda biyometrik tanıma sistemleri açısından da işlevsel olup olmadığını değerlendirmektir. Görüntü süper çözünürlük veya kalite iyileştirme çalışmaları çoğu zaman sadece görsel benzerlik veya yapay kalite metrikleriyle değerlendirilmekteyken, biyometrik sistemlerde bu yeterli değildir. Bu kısımda, modelin çıktısının tanıma performansına etkisini analiz ederek, elde edilen görüntülerin tanıma açısından ne derece bilgi taşıdığı ve sistemin doğruluk oranlarını ne ölçüde koruyabildiği incelenmiştir. Böylece, iyileştirilen avuç içi görüntülerinin gerçek sistemlerde kullanılabilirliğine dair daha güvenilir ve pratik bir değerlendirme yapılması amaçlanmıştır. Tanıma başarımı, modelin hem üretimsel doğruluğunu hem de biyometrik anlamda geçerliliğini ortaya koyması bakımından kritik öneme sahiptir.

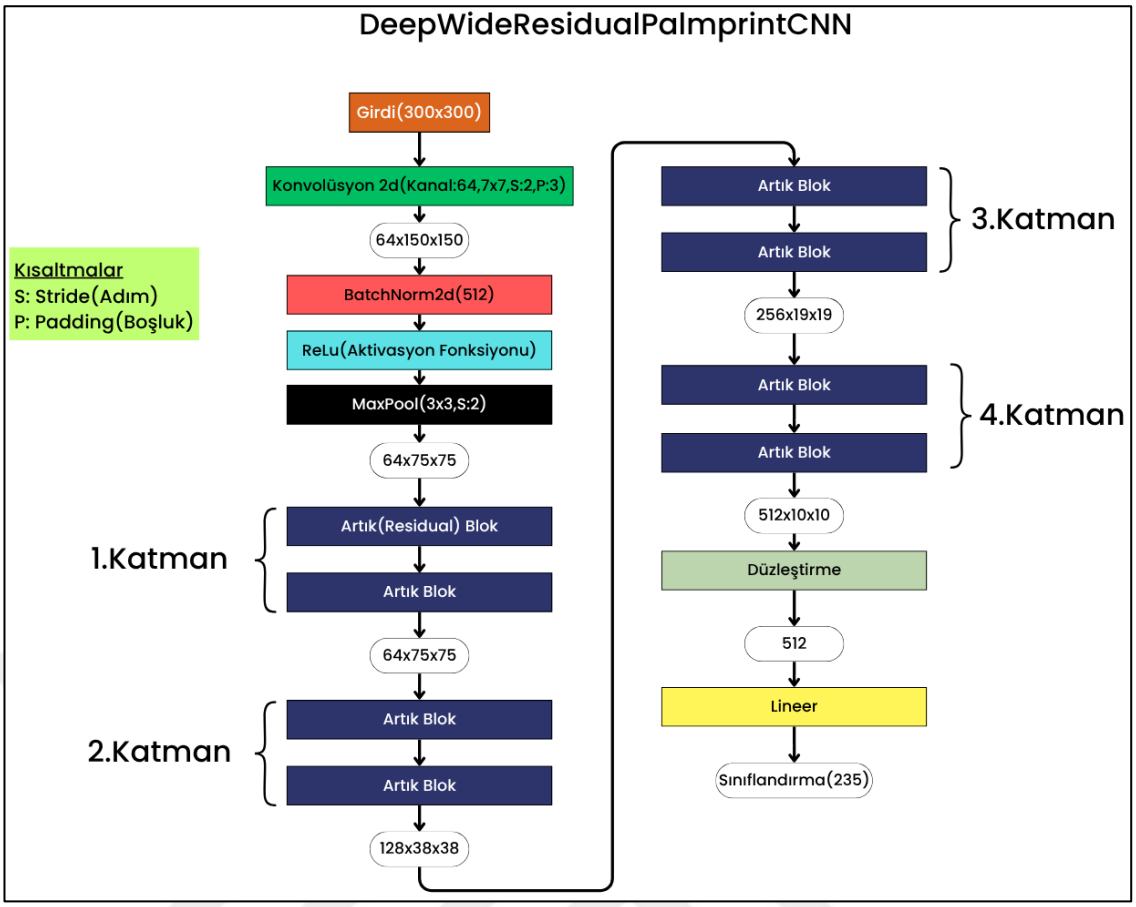


a)Orijinal Görüntülerin Tanıma Performansı Grafiği

b)Üretilmiş Görüntülerin Tanıma Performansı Grafiği

Şekil 22. Orijinal(a) ve üretilmiş(b) görüntülerin Tanıma Performansı grafikleri

Bu değerlendirme sürecinde kullanılan "DeepWideResidualPalmprintCNN" modeli, derin ve geniş yapılı bir evrişimsel sinir ağı olup, özellikle avuç içi biyometrik tanıma görevleri için özelleştirilmiştir (Şekil 23). Model, ilk olarak geniş bir konvolüsyon katmanı ve maksimum havuzlama işlemi ile temel öznitelikleri çıkarır. Ardından art arda yerleştirilmiş dört adet Artık (Residual) Blok katmanı aracılığıyla hem düşük hem yüksek seviyeli özellikleri etkili biçimde öğrenir. Her bir blok, girdi ile çıktı arasındaki farkı öğrenmeye odaklanan artık bağlantılar (residual connections) içerir ve bu sayede derin katmanlarda meydana gelebilecek gradyan sönmesi problemi azaltılır. Modelin sonunda yer alan global ortalama havuzlama ve tam bağlantılı katman, her bir örneği 235 sınıftan birine sınıflandırmak üzere tasarlanmıştır. Böylece hem geniş hem de derin özellik temsili sunularak, avuç içi görüntüleri üzerinden yüksek doğrulukla tanıma yapılması hedeflenmiştir.



Şekil 23. DeepWideResidualPalmPrintCNN model şeması

Geliştirilmiş kod çözücü yapısına sahip AAE modelinin çıktıları üzerinde yapılan analiz, modelin avuç içi görüntülerini yüksek doğrulukla ve artırılmış çözünürlükle yeniden üretebildiğini ve biyometrik tanıma açısından başarılı performans sergilediğini göstermektedir. Eğitim süreci boyunca hem eğitim hem de doğrulama doğruluklarının %95'in üzerinde seyretmesi, modelin hem öğrenme kapasitesinin yüksek olduğunu hem de genelleme yeteneğini koruduğunu ortaya koymaktadır. Giriş görüntülerine ait grafiklerde gözlemlenen düşük kayıplar ve stabil doğruluklar, modelin orijinal veriyi kolaylıkla tanıyabildiğini gösterirken, çıkış görüntülerine ait grafiklerde doğruluk değerlerinin benzer şekilde yüksek kalması, geliştirilen kod çözücünün görüntüdeki biyometrik ayırt edici özellikleri başarılı şekilde koruyarak ve yüksek çözünürlükle yeniden üretebildiğini kanıtlamaktadır. Her ne kadar çıkış görüntülerine ait doğrulama kayıplarında bazı dalgalanmalar gözlense de genel olarak kayıpların düşük seviyelerde kalması, modelin kararlı çıktılar üretme eğiliminde olduğunu göstermektedir. Bu sonuçlar, geliştirilen modelin sadece görsel kaliteyi artırmakla kalmayıp aynı zamanda tanıma performansını da koruduğunu, dolayısıyla biyometrik sistemlerde kullanılabilecek güvenilir bir süper çözünürlük yaklaşımı sunduğunu ortaya koymaktadır.

3.6. Genel Yorumlar ve Bilimsel Katkı

Bu çalışmada geliştirilen AE tabanlı süper çözünürlük mimarileri, literatürde yer alan benzer biyometrik görüntü iyileştirme çalışmalarına kıyasla farklı çözünürlük oranlarında (0.5x, 2x, 4x, 8x) ve çeşitli bozulma senaryolarında (Gauss, blur) test edilmiştir. Elde edilen sonuçlar, AE mimarisinin özellikle düşük çözünürlüklerde yapısal koruma ve yeniden yapılandırma başarımı açısından güçlü olduğunu göstermektedir. Literatürde sıkça karşılaşılan GAN tabanlı yöntemlerle yapılan karşılaştırmalar ise GAN modellerinin görsel detay üretiminde daha başarılı olabildiğini, ancak AE mimarisinin yapısal tutarlılığı daha istikrarlı koruduğunu ortaya koymuştur. Bu çalışmanın avuç içi biyometrisi özelinde kapsamlı bir deneysel analiz sunması ve BRISQUE gibi referanssız metriklerle nicel ve nitel değerlendirmeyi birleştirmesi, benzer çalışmalardan ayrılan önemli yönler arasında yer almaktadır. Ayrıca, çeşitli gürültü tipleri altında model dayanıklılığını inceleyen az sayıdaki çalışmadan biri olması bakımından da literatüre katkı sunmaktadır.

Çalışmanın özgünlüğü, AE ve GAN mimarilerini yalnızca yüksek çözünürlük üretme açısından değil, aynı zamanda düşük kaliteli biyometrik verilerde tanıma için kullanılabilirliği artırma perspektifiyle değerlendirmesinde yatmaktadır. Özellikle 0.5x gibi literatürde nadiren ele alınan ultra düşük çözünürlük seviyelerinde model davranışının analiz edilmesi, bilimsel anlamda önemli bir boşluğu doldurmaktadır. Ayrıca Gauss gürültüsü ve bulanıklık gibi bozulmaların dahil edilmesiyle, modellerin gerçek hayattaki zorlu koşullara karşı dirençleri test edilmiş; bu da geliştirilen sistemlerin pratik uygulamalara yönelik geçerliliğini artırmaktadır. Bu tez kapsamında geliştirilen mimariler, düşük çözünürlükteki biyometrik görüntülerin kalite iyileştirme sürecine yalnızca klasik süper çözünürlük bakış açısından değil, aynı zamanda yeniden yapılandırma, görsel gerçekçilik ve dayanıklılık açılarından çok yönlü bir bakış getirmiştir. Böylece hem model mimarisi hem de deneysel tasarım yönüyle alana somut bir katkı sunulmuştur.

4.TARTIŞMA VE ÖNERİ

4.1. Bulguların Yorumlanması

Bu çalışmada kullanılan AE tabanlı süper çözünürlük modeli, farklı çözünürlük oranlarında (0.5x, 2x, 4x, 8x) ve çeşitli bozulma senaryolarında (Gauss, blur) test edilmiştir. Elde edilen sonuçlar, AE mimarisinin özellikle düşük ve orta düzey çözünürlüklerde yapısal bütünlük ve yeniden yapılandırma doğruluğu açısından tutarlı ve stabil çıktılar üretebildiğini göstermektedir. Özellikle gürültülü ve bulanık veri koşullarında dahi AE modeli, temel çizgi ve doku yapılarını koruyarak anlamlı çıktılar üretmiş, BRISQUE ve PSNR gibi metriklerde kademeli bir iyileşme sergilemiştir.

Farklı koşullarda modelin davranışı değerlendirildiğinde, AE mimarisinin düşük çözünürlüklerde (örneğin 0.5x) hızlı bir şekilde kararlı hale geldiği ve yapısal öğeleri koruyabildiği, ancak çözünürlük oranı arttıkça (örneğin 8x) detay üretiminde zorlandığı görülmektedir. Bu durum, mimarinin tasarımı gereği daha çok genel yapıların korunmasına odaklandığını, ince detayların üretimi içinse daha gelişmiş modellerin gerekebileceğini ortaya koymaktadır.

Beklentilere paralel olarak AE modelinin düşük çözünürlüklü ve bozulmuş verilerde daha başarılı çıktılar ürettiği, ancak yüksek ölçekleme oranlarında görsel detay kaybı yaşadığı gözlenmiştir. Bu sonuçlar, modelin genel tasarımıyla uyumlu bir şekilde davranış gösterdiğini doğrulamaktadır. Görsel kalite ile biyometrik doğruluk arasındaki ilişki açısından değerlendirildiğinde, AE modeliyle üretilen görüntülerde çizgi yapılarının netliği ve simetrisi arttıkça kalite metriklerinde iyileşme olduğu kadar, tanınabilirlik potansiyelinin de yükseldiği gözlemlenmiştir.

4.2. Yöntemsel Sınırlılıklar

Bu çalışmada kullanılan veri seti, yaklaşık 230 bireye ait sağ elin beşer görüntüsünden oluşmakta olup, sınırlı sayıda örnekle temsil edilmektedir. Veri kümesi, her ne kadar belirli varyasyonları (poz farkı, ışık değişimi vb.) içerse de, yaş, cinsiyet, etnik köken gibi biyometrik çeşitliliği tam olarak yansıtmamaktadır. Bu durum, geliştirilen modelin genellenebilirliğini kısıtlamakta ve daha geniş, dengeli veri setlerine ihtiyaç duyulduğunu ortaya koymaktadır. Ayrıca sadece sağ el verileriyle çalışılması, iki el arasında oluşabilecek yapısal farklılıkların modellenmesini engellemiştir.

Teknik sınırlamalar açısından değerlendirildiğinde, özellikle yüksek çözünürlük oranlarında (ör. 8×) AE modelinin eğitim süresi oldukça uzamış ve eğitim zamanları sırasıyla 0.5× için yaklaşık 6 saat, 2× için 13 saat, 4× için 25 saat, 8× için ise yaklaşık 52 saati bulmuştur. Bu süreler, donanım olarak kullanılan NVIDIA RTX 2070 Super GPU'nun belleği (8 GB) dikkate alındığında mevcut sistemin sınırlarında çalışıldığını göstermektedir. Ayrıca yüksek çözünürlükteki çıkışlar için kullanılan daha derin kod çözücü mimarileri, hem GPU belleğini zorlamış hem de batch boyutunun düşürülmesine neden olmuştur. Bu da modelin öğrenme kapasitesini sınırlayan bir başka faktör olarak öne çıkmıştır.

GAN tabanlı modeller çalışmada geliştirilmemiş olmasına rağmen, kıyaslama amacıyla kullanılan mevcut GAN mimarilerinin eğitiminde karşılaşılan genel zorluklara da dikkat çekilmiştir. Bu modellerde üretici ve ayrıştırıcı ağlar arasındaki dengenin sağlanması oldukça güç olup, eğitim süreci çoğu zaman kararsız hale gelmektedir. Aşırı öğrenme riski, özellikle küçük veri setlerinde önemli bir problem olarak ortaya çıkmakta ve hiperparametrelerin küçük değişikliklerle bile sonuçları önemli ölçüde etkilediği görülmektedir. Bu nedenlerle, GAN modelleriyle stabil bir eğitim gerçekleştirmek için daha hassas ayar süreçleri ve daha fazla hesaplama kaynağı gereklidir.

4.3. Uygulama Alanlarına Yönelik Öneriler

Bu çalışmada elde edilen bulgular doğrultusunda, geliştirilen AE tabanlı süper çözünürlük modellerinin biyometrik sistemlerde düşük kaliteli giriş verilerinin iyileştirilmesinde etkin bir şekilde kullanılabileceği öngörülmektedir. Özellikle düşük çözünürlüklü ya da bozulmuş avuç içi görüntülerinin tanıma sistemine aktarılmadan önce süper çözünürlükle netleştirilmesi, tanıma doğruluğunu artırabilecek önemli bir adımdır. Bu yöntem, veri kalitesinin yetersiz olduğu sistemlerde yazılım düzeyinde çözüm sunarak donanımsal yükseltme ihtiyacını azaltabilir.

Uygulama açısından bakıldığında, güvenlik sistemlerinde avuç içi tanıma tabanlı erişim kontrolü ve kimlik doğrulama senaryolarında modelin kullanım potansiyeli yüksektir. Aynı zamanda mobil biyometri alanında düşük çözünürlükle çekilmiş avuç içi görüntülerinin işlenmesi, mobil cihazlar üzerindeki tanıma performansını artırabilir. Sağlık teknolojileri bağlamında ise, avuç içi damar yapılarının analizi gibi non-invaziv biyometrik uygulamalarda görüntü kalitesinin artırılması, daha doğru teşhis ve sınıflandırma sistemlerine katkı sağlayabilir.

Gerçek zamanlı uygulamalar için mevcut AE mimarisinin sadeleştirilmiş ve optimize edilmiş sürümleri önerilmektedir. Özellikle daha az sayıda parametreye sahip

kodlayıcı-kod çözücü yapılarının, hesaplama yükü düşük sistemlerde kullanılabilir hale getirilmesi mümkündür. Ayrıca GAN tabanlı modellerin üretici ve ayrıştırıcı bileşenleri sadeleştirilerek mobil platformlara entegre edilebilir. Bu tür sadeleştirilmiş yapılar düşük gecikmeli, güvenilir süper çözünürlük uygulamalarının yolunu açabilir. Gerçek zamanlı ihtiyaçlar göz önüne alındığında, model karmaşıklığı ile doğruluk arasında optimum dengenin kurulması uygulama başarımı açısından kritik önemdedir.

4.4. Gelecek Çalışmalar İçin Öneriler

Gelecek çalışmalarda, transformer tabanlı SR mimarileri ile geliştirilen AE modellerinin karşılaştırmalı performans analizlerinin gerçekleştirilmesi, modelin görsel detay üretimi ve yapısal koruma açısından sınanması açısından önemlidir. Özellikle SwinIR gibi görsel dikkat mekanizmalarına dayalı SR yaklaşımlarının, avuç içi biyometrisi gibi çizgisel yapıya sahip veriler üzerindeki etkisinin incelenmesi, mevcut AE mimarilerinin güçlü ve zayıf yönlerini daha ayrıntılı ortaya koyabilir.

Geliştirilen SR modellerinin sadece görüntü kalitesi bağlamında değil, doğrudan biyometrik doğrulama sistemleri ile entegre olarak değerlendirilmesi, SR işleminin tanıma başarımına olan katkısını doğrudan ölçme fırsatı sunacaktır. Bu kapsamda, SR sonrası elde edilen iyileştirilmiş görüntülerin tanıma doğruluğuna etkisinin ROC eğrileri, CMC analizleri veya FAR/FRR oranları gibi ölçütlerle değerlendirilmesi, modelin pratik faydasını somutlaştıracaktır.

Ayrıca önerilen AE ve karşılaştırmalı GAN modellerinin, yalnızca avuç içi görüntüleriyle sınırlı kalmayıp, yüz, iris ve parmak izi gibi farklı biyometrik modalitelerde de test edilmesi, mimarilerin genellenebilirliğini değerlendirme açısından önemli bir adımdır. Her bir biyometrik modalite kendine özgü yapısal özellikler barındırdığından, bu tür genişletilmiş çalışmalar modelin farklı senaryolardaki etkinliğini ortaya koyacaktır.

Son olarak, gerçek zamanlı uygulamalar için SR modellerinin sadeleştirilmiş versiyonlarının geliştirilmesi ve bu modellerin mobil cihazlara uygulanabilirliği üzerine çalışmalar yapılması önerilmektedir. Bu bağlamda, model karmaşıklığının azaltılması, hesaplama yükünün düşürülmesi ve donanımsal kısıtlar altında SR performansının korunması, SR çözümlerinin pratik hayata aktarılabilirliğini büyük ölçüde artıracaktır. Özellikle edge cihazlar üzerinde çalışan hafif AE/GAN modelleri ile düşük çözünürlüklü biyometrik verilerin yerinde ve gecikmesiz işlenmesi mümkün hale gelebilir.

KAYNAKÇA

Aykut, M., ve Ekinci, M. (2015). Developing a contactless palmprint authentication system by introducing a novel ROI extraction method. *Image and Vision Computing*, *40*, 65–74

Aykut, M., ve Ekinci, M. (2013). AAM-based palm segmentation in unrestricted backgrounds and various postures for palmprint recognition. *Pattern Recognition Letters, 34*(9), 955–962.

Baker, S., ve Kanade, T. (2000). Hallucinating faces. *Proceedings - 4th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2000*, 83-88. https://doi.org/10.1109/AFGR.2000.840616

Bingöl, Ö., ve Ekinci, M. (2017). Stereo-based palmprint recognition in various 3D postures. *Expert Systems with Applications*, *78*, 74-88.

Dong, C., Loy, C. C., He, K., ve Tang, X. (2014). Image Super-Resolution Using Deep Convolutional Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *38*(2), 295-307. https://doi.org/10.1109/TPAMI.2015.2439281

Dong, C., Loy, C. C., ve Tang, X. (2016). Accelerating the Super-Resolution Convolutional Neural Network. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9906 LNCS*, 391-407. https://doi.org/10.1007/978-3-319-46475-6_25

Freedman, G., ve Fattal, R. (2011). Image and video upscaling from local self-examples. *ACM Transactions on Graphics (TOG)*, *30*(2). https://doi.org/10.1145/1944846.1944852

Gerchberg, R. W. (1974). Super-resolution through Error Energy Reduction. *Optica Acta: International Journal of Optics*, *21*(9), 709-720. https://doi.org/10.1080/713818946

Goodfellow, I., Bengio, Y., ve Courville, A. (2016) *Deep Learning*.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., ve Bengio, Y. (2014). Generative adversarial networks. *arXiv*. https://arxiv.org/abs/1406.2661

Huang, J., Ma, L., Tan, T., ve Wang, Y. (2003). Learning based resolution enhancement of iris images. *Proceedings of the 14th British Machine Vision Conference (BMVC)*, 153–162.

Kim, K. I., ve Kwon, Y. (2010). Single-image super-resolution using sparse regression and natural image prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(6), 1127-1133. https://doi.org/10.1109/TPAMI.2010.25

Kramer, M. A. (1991). Nonlinear principal component analysis using autoassociative neural networks. Içinde *AIChE Journal* (C. 37, Sayı 2). https://doi.org/10.1002/aic.690370209

Kumar, A. (2008). Incorporating cohort information for reliable palmprint authentication. *Proceedings of the Sixth Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP)*, 583–590. https://doi.org/10.1109/ICVGIP.2008.73

Kumar, A., ve Zhang, D. (2012). Personal recognition using hand shape and texture. *IEEE Transactions on Image Processing*, *21*(4), 2453–2463. https://doi.org/10.1109/TIP.2011.2175743

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., ve Shi, W. (2016a). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, *2017-January*, 105-114. https://doi.org/10.1109/CVPR.2017.19

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., ve Shi, W. (2016b). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, *2017-January*, 105-114. https://doi.org/10.1109/CVPR.2017.19

Lim, B., Son, S., Kim, H., Nah, S., ve Mu Lee, K. (2017). Enhanced deep residual networks for single image super-resolution. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 136–144. https://arxiv.org/abs/1707.02921

Macan, H. F., Bingol, O., Turk, S., ve Ekinci, M. (2024). Super-resolution based ROI enhancement approaches for palmprint recognition. *Proceedings of the 8th International Artificial Intelligence and Data Processing Symposium (IDAP)*, 399–406. https://doi.org/10.1109/IDAP64064.2024.10710808

Makhzani, A., Shlens, J., Jaitly, N., Brain, G., Openai, I. G., ve Frey, B. (2015). Adversarial Autoencoders. *Elements of Dimensionality Reduction and Manifold Learning*, 577-596. https://doi.org/10.1007/978-3-031-10602-6_21

Mjolsness, E. D. (1986). *Neural Networks, Pattern Recognition, and Fingerprint Hallucination*. https://doi.org/10.7907/M0VQ-DJ43

Moser, B., Raue, F., Frolov, S., Hees, J., Palacio, S., ve Dengel, A. (2022). Hitchhiker's Guide to Super-Resolution: Introduction and Recent Advances. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(8), 9862-9882. https://doi.org/10.1109/TPAMI.2023.3243794

Radford, A., Metz, L., ve Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv*. https://arxiv.org/abs/1511.06434

Sünnetci, B. K., Bingöl, Ö., Gedikli, E., ve Ekinci, M. (2024, September). High Performance Automatic ROI Extraction from Hand Images with Convolutional Neural Networks. In *2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP)* (pp. 1-6). *IEEE*.

Valsesia, D., ve Magli, E. (2021). Permutation invariance and uncertainty in multitemporal image super-resolution. *IEEE Transactions on Geoscience and Remote Sensing*. https://doi.org/10.1109/TGRS.2021.3130673

Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., ve Loy, C. C. (2018). ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11133 LNCS*, 63-79. https://doi.org/10.1007/978-3-030-11021-5_5

Wang, Z., Chen, J., ve Hoi, S. C. (2020). Deep learning for image super-resolution: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10), 3365-3387. https://doi.org/10.1109/TPAMI.2020.2982166

Zhao, Q., Zhang, Y., ve Wang, Y. (2019). *The impact of image quality on biometric recognition performance: A survey.* IEEE Access, 7, 123245–123265. https://doi.org/10.1109/ACCESS.2019.2937965

ÖZGEÇMİŞ

Hüseyin Furkan MACAN Lise eğitimini Gümüşhane Türk Telekom Fen Lisesi'nde tamamlamış, üniversite lisans eğitimini Karadeniz Teknik Üniversitesi Bilgisayar Mühendisliği'nden mezun olmuştur. Gümüşhane Üniversitesi Yapay Zeka ve Akıllı Sistemler ABD'da yüksek lisans eğitimini sürdürmektedir. Bu tezin de çıktısı olduğu ve TUBİTAK 1001 - Bilimsel ve Teknolojik Araştırma Projelerini Destekleme Programı tarafından desteklenen "122E402" numaralı ve "Kısıtlamasız Ortamda Alınan El Görüntülerinden Biyometrik Doğrulama İçin Güvenilir Avuç İzi Bölgesinin Tespiti" adlı projede bursiyer olarak yer almıştır. Gümüşhane Üniversitesi İktisadi ve İdari Bilimler Fakültesi Yönetim Bilişim Sistemleri Bölümünde Araştırma Görevlisi olarak çalışmaktadır.