



İZMİR BAKIRÇAY ÜNİVERSİTESİ

MÜHENDİSLİK VE MİMARLIK FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

DERİN ÖĞRENME FINAL PROJESİ

AUTOENCODER TABANLI
GÖRÜNTÜ ÇÖZÜNÜRLÜĞÜ ARTIRMA
(SUPER-RESOLUTION)

Hazırlayanlar :

Sudenaz ONARAN 220601060

Ece KARAALP 220601046

İÇİNDEKİLER

1.	GİRİŞ	3
1.1.	Problem Tanımı.....	3
1.2.	Motivasyon.....	3
1.3.	Veri Seti Seçimi	4
2.	METODOLOJİ.....	5
2.1.	Veri Ön İşleme	5
2.2.	Model Mimarileri	5
2.2.1.	Versiyon 1 – Basit Autoencoder	6
2.2.2.	Versiyon 2 – Strided convolution.....	6
2.2.3.	Versiyon 3 – Global Skip Connection.....	7
2.2.4.	Versiyon 4 – Final Model.....	7
2.3.	Hiperparametreler ve Eğitim Ayarları.....	8
3.	BULGULAR VE DEĞERLENDİRME	8
3.1.	Eğitim Sonuçları.....	8
3.2.	Performans Ölçütleri	10
3.3.	Görsel Sonuçlar	11
4.	TARTIŞMA VE SONUÇ	12
4.1.	Bulguların Yorumu	12
4.2.	Gelecek Çalışmalar	13
4.3.	Sonuç.....	13

1. GİRİŞ

1.1. Problem Tanımı

Bu projede ele alınan problem, görüntü çözünürlüğünü artırma (Image Super-Resolution) problemidir. Super-Resolution, düşük çözünürlüklü (Low Resolution – LR) bir görüntüden, daha fazla detay içeren yüksek çözünürlüklü (High Resolution – HR) bir görüntü üretmeyi amaçlayan önemli bir bilgisayarlı görü problemidir.

Görüntü çözünürlüğünü artırma problemi; yüz tanıma sistemleri, medikal görüntüleme, güvenlik kameraları ve eski ya da düşük kaliteli görüntülerin iyileştirilmesi gibi birçok alanda kritik bir rol oynamaktadır. Bu nedenle problem, derin öğrenme yöntemlerinin uygulanması açısından uygun ve güncel bir çalışma alanı sunmaktadır.

Bu çalışmada, düşük çözünürlüklü yüz görüntülerinden yüksek çözünürlüklü yüz görüntülerini üretmek amacıyla CNN tabanlı bir Autoencoder mimarisi kullanılmıştır. Model, giriş olarak verilen düşük çözünürlüklü görüntülerden anlamlı özelliklerini öğrenerek, çıkışta daha net ve detaylı görüntüler üretmeyi hedeflemektedir.

1.2. Motivasyon

Son yıllarda Convolutional Neural Network (CNN) tabanlı modellerin görüntü restorasyonu alanında başarılı sonuçlar vermesi, bu problemin derin öğrenme ile ele alınmasını teşvik etmiştir. Özellikle autoencoder tabanlı mimariler, görüntüsünden anlamlı özellikler çıkarma ve bu özelliklerden yeniden görüntü üretme konusunda etkili bir yapı sunmaktadır.

Bu projenin temel motivasyonları şunlardır:

- Süper-çözünürlük problemini adım adım geliştirilen bir model süreci ile ele almak
- Basit bir mimariden başlayarak, modelin zayıf yönlerini analiz edip iyileştirmek
- Kenar ve detay bilgisini koruyarak görsel kaliteyi artırmak
- Eğitim ve test aşamalarında gerçekçi düşük çözünürlük senaryoları oluşturmak

1.3.Veri Seti Seçimi

Bu çalışmada CelebA (CelebFaces Attributes Dataset) veri seti kullanılmıştır. CelebA veri seti, farklı poz, aydınlatma ve yüz ifadelerine sahip yüksek çözünürlüklü yüz görüntülerinden oluşmaktadır.

- Toplam görüntü sayısı: **202.599**
- Görüntü boyutu: **178 × 218 × 3 (RGB)** (Şekil 2.1)
- Bu projede hesaplama maliyetini düşürmek amacıyla veri setinden **rastgele seçilmiş 500 görüntü** kullanılmıştır.
- Tüm görüntüler eğitim öncesinde **128×128 boyutuna** normalize edilmiştir.

CelebA veri seti, yüz detaylarının belirgin olması nedeniyle süper-çözünürlük problemi için oldukça uygun bir veri setidir.



Şekil 2.1 Veri setinden örnek görüntü , boyut ve mod değerleri

2. METODOLOJİ

2.1. Veri Ön İşleme

Modelin eğitimi için her bir yüksek çözünürlüklü görüntünden yapay olarak düşük çözünürlüklü görüntüler üretilmiştir (Şekil 2.2). Bu yaklaşım, süper-çözünürlük problemlerinde yaygın olarak kullanılan bir tekniktir.

- **HR (High Resolution):**
 - Görüntüler 128×128 boyutuna yeniden boyutlandırılmıştır.
 - Tensor formatına dönüştürülmüştür.
- **LR (Low Resolution):**
 - Görüntüler önce 64×64 boyutuna küçültülmüştür.
 - Gaussian Blur uygulanarak detay kaybı simül edilmişdir.
 - Son versiyonda ColorJitter eklenerek parlaklık ve kontrast değişimleri uygulanmıştır.

Bu sayede model, gerçek hayatı karşılaşabilecek düşük kaliteli görüntülerini daha iyi öğrenmiştir.



Şekil 2.2 LR ve HR örnek görüntü

2.2. Model Mimarileri

Bu çalışmada görüntü süper-çözünürlük problemi, artımlı (iteratif) bir model geliştirme yaklaşımı ile ele alınmıştır. Model, basit bir autoencoder yapısından başlanarak her versiyonda elde edilen sonuçlar analiz edilmiş ve mimari buna göre geliştirilmiştir. Amaç, bulanıklığı azaltmak, kenar ve detay bilgisini daha iyi koruyan bir yapı elde etmektir.

2.2.1. Versiyon 1 – Basit Autoencoder

İlk aşamada, süper-çözünürlük problemi için temel bir Convolutional Autoencoder mimarisi kullanılmıştır. Bu yapı, bir Encoder ve bir Decoder bölümünden oluşmaktadır.

Encoder kısmında:

- Convolution katmanları ile temel görsel özellikler çıkarılmış,
- ReLU aktivasyon fonksiyonu kullanılmış,
- MaxPooling katmanları ile uzamsal boyutlar azaltılarak sıkıştırılmış bir temsil elde edilmiştir.

Decoder kısmında ise:

- Upsampling katmanları ile görüntü boyutu kademeli olarak büyütülmüş,
- Convolution katmanları ile görüntü yeniden oluşturulmuştur.

Bu versiyonda kayıp fonksiyonu olarak Mean Squared Error (MSE Loss) kullanılmıştır. MSE, piksel bazlı farkları minimize etmeye odaklandığı için model genel görüntü yapısını öğrenebilmiştir. Eğitim ve doğrulama kayipları (train / validation loss) düzenli olarak düşmesine rağmen, görsel kalitede belirgin bir iyileşme gözlemlenmemiştir (Şekil 2.2.2) bunun temel nedeni encoder aşamasında kullanılan pooling işlemleri, yüksek frekanslı detayların kalıcı olarak kaybolmasına neden olmaktadır. Decoder, bu kaybolan bilgiyi yalnızca interpolasyon ile geri üretmeye çalıştığı için, keskin kenarlar ve ince dokular yeniden oluşturulamamaktadır. (Şekil 2.2.3)

2.2.2. Versiyon 2 – Strided convolution

Bu versiyonda, klasik pooling katmanları yerine strided convolution kullanılarak downsampling işlemi öğrenilebilir hale getirilmiştir. Böylece ağı, yalnızca uzamsal boyutu küçültmekle kalmayıp, hangi özelliklerin korunacağına da karar verebilmektedir. Bu yaklaşım, encoder aşamasında özellikle yüksek frekanslı detayların tamamen kaybolmasını kısmen engellemiştir.

Elde edilen sonuçlarda, önceki modele kıyasla kenarların daha belirgin olduğu ve genel görüntü keskinliğinin arttiği gözlemlenmiştir. Bununla birlikte, residual bağlantılarının bulunmaması

nedeniyle derin katmanlarda detay üretimi hâlâ sınırlı kalmıştır ve ince yapılar tam olarak geri kazanılamamıştır. (Şekil 2.2.5)

Kayıp grafiği, modelin eğitim sürecinde istikrarlı bir şekilde yakınsadığını göstermektedir. Kayıp değerinin düzenli olarak azalması, strided convolution kullanımının öğrenme sürecine olumlu katkı sağladığını göstermektedir. Ancak loss değerinin belirli bir seviyeden sonra yavaş azalması, mimarinin ifade gücünün sınırlı olduğuna ve daha gelişmiş bağlantı yapılarının (örneğin residual bağlantılar) gerekli olabileceğine işaret etmektedir. (Şekil 2.2.4)

2.2.3. Versiyon 3 – Global Skip Connection

Üçüncü versiyonda, modelin tüm görüntüyü yeniden üretmeye çalışması yerine yalnızca eksik detayları öğrenmesini sağlamak amacıyla Global Skip Connection yapısı eklenmiştir. (Şekil 2.2.7)

Bu yaklaşımda:

- Girişteki düşük çözünürlüklü görüntü, bilinear interpolation ile büyütülmüş,
- Bu büyütülmüş görüntü model çıktısına doğrudan eklenmiştir.

Bu sayede:

- Model, zaten mevcut olan düşük frekanslı bilgiyi tekrar öğrenmek zorunda kalmamış,
- Öğrenme süreci yüksek frekanslı detaylara (kenarlar, dokular) odaklanmıştır.

Loss grafiği, Global Skip Connection kullanımıyla modelin daha hızlı yakınsadığını ve daha düşük hata seviyelerine ulaştığını göstermektedir. (Şekil 2.2.6)

2.2.4. Versiyon 4 – Final Model

Son versiyonda, önceki tüm versiyonlardan elde edilen kazanımlar bir araya getirilmiş ve model daha da güçlendirilmiştir.(Şekil 2.2.9)

Bu versiyonda yapılan başlıca geliştirmeler:

- 8 adet Residual Block kullanılarak model kapasitesi artırılmıştır.
- Düşük çözünürlüklü görüntüler daha gerçekçi olacak şekilde:

- Downsampling,
- Gaussian Blur,
- ColorJitter (parlaklık ve kontrast değişimi)
uygulanarak üretilmiştir.
- Kayıp fonksiyonuna Edge Loss eklenmiştir. (2.2.8)

Edge Loss sayesinde model:

- Kenar ve kontur bölgelerine daha fazla önem vermiş,
- Yüz hatları ve geçiş bölgelerinde daha keskin sonuçlar üretmiştir.

2.3.Hiperparametreler ve Eğitim Ayarları

Tüm deneylerde kullanılan temel hiperparametreler aşağıda verilmiştir:

- **Optimizer:** Adam
- **Learning Rate:** 1e-4
- **Batch Size:** 8

Final modelde kullanılan kayıp fonksiyonu:

$$\text{Loss} = \text{L1} + 0.1 \times \text{EdgeLoss}$$

Bu yapı sayesinde:

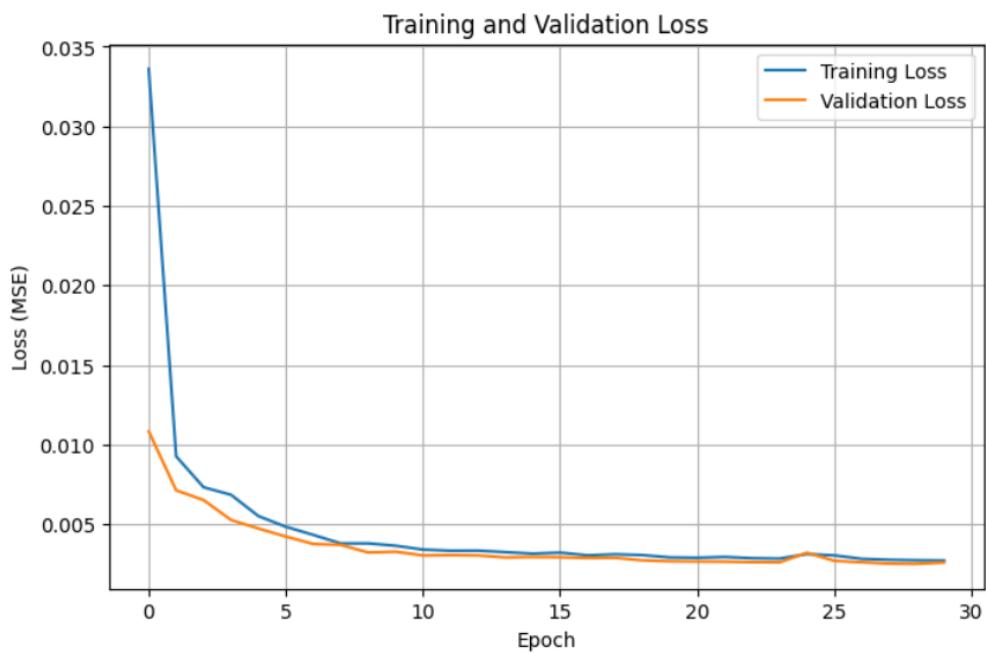
- **L1 Loss**, genel piksel benzerliğini sağlamış,
- **Edge Loss**, kenar ve kontur bilgisini güçlendirmiştir.

3. BULGULAR VE DEĞERLENDİRME

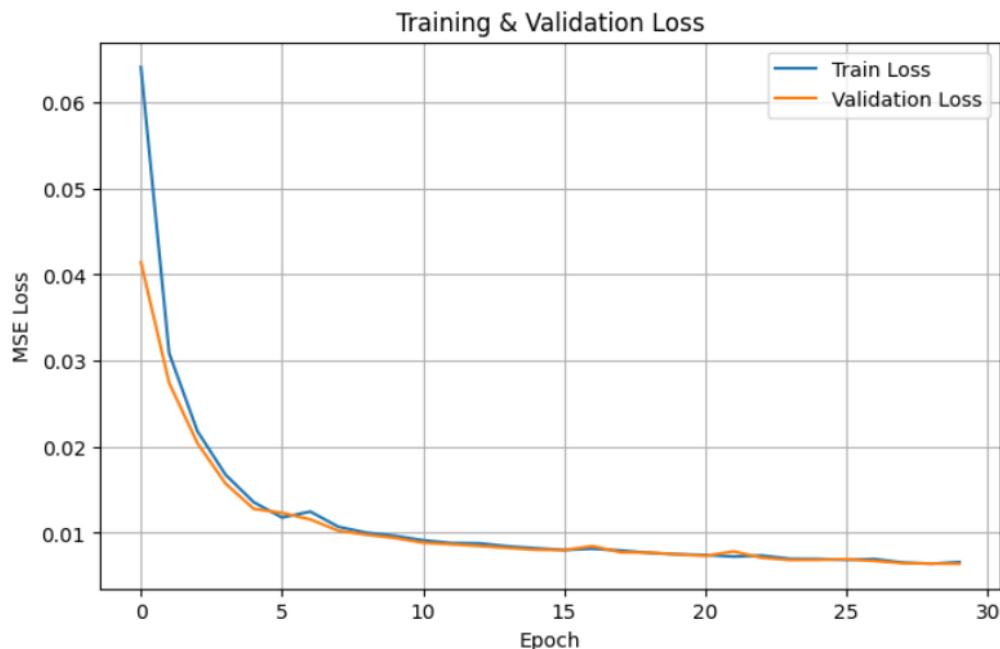
3.1.Eğitim Sonuçları

Eğitim sürecinde train ve validation loss değerleri birlikte azalmıştır. Bu durum, modelin overfitting yapmadığını ve genelleme yeteneğinin iyi olduğunu göstermektedir.

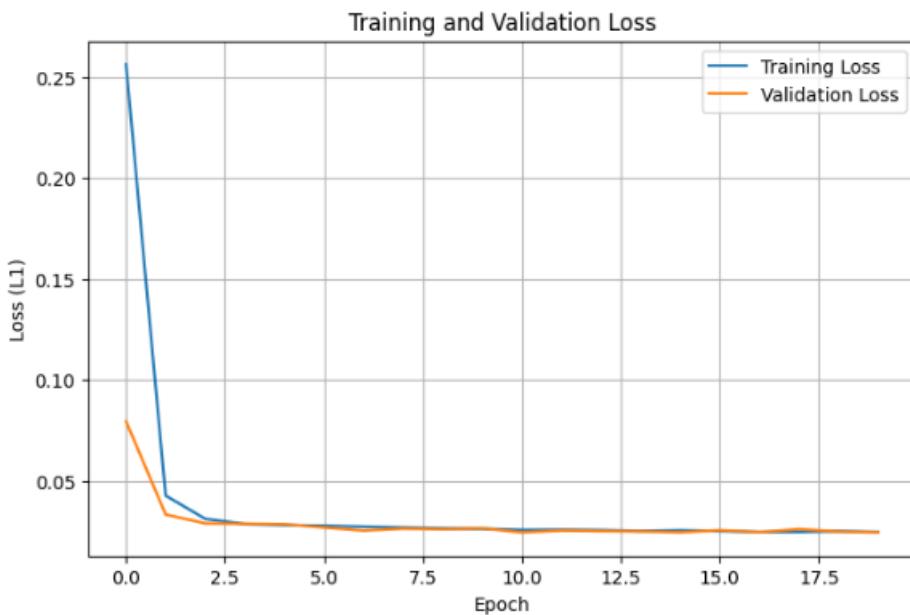
- Train loss ve validation loss eğrileri paralel ilerlemiştir
- Eğitim stabil bir şekilde tamamlanmıştır.



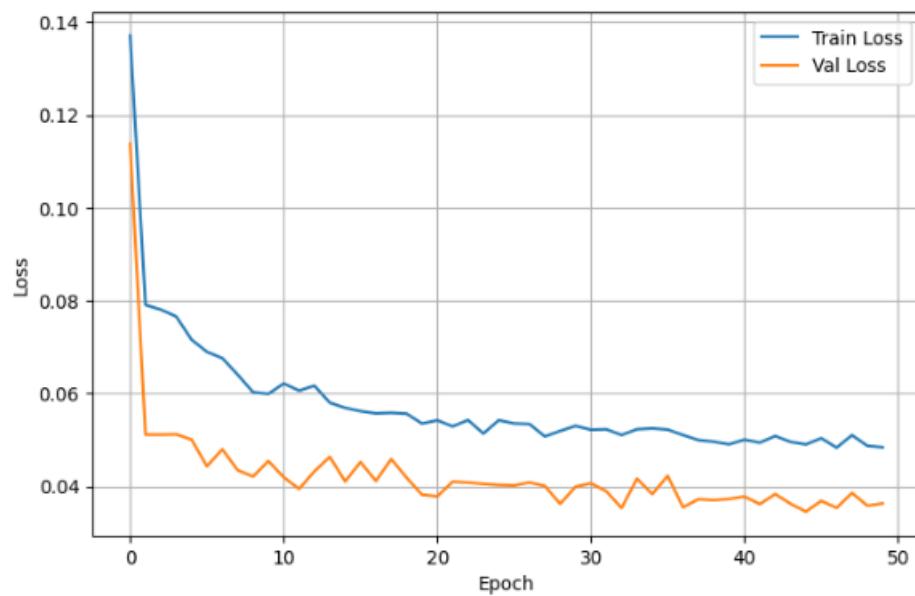
Şekil 2.2.2 Basit Autoencoder Eğitim ve Doğrulama Kayıp Grafiği



Şekil 2.2.4 Strided convolution eğitim ve doğrulama kayıp grafiği



Şekil 2.2.6 Global Skip Connection eğitim ve doğrulama kayıp grafiği



Şekil 2.2.8 Final Model eğitim ve doğrulama kayıp grafiği

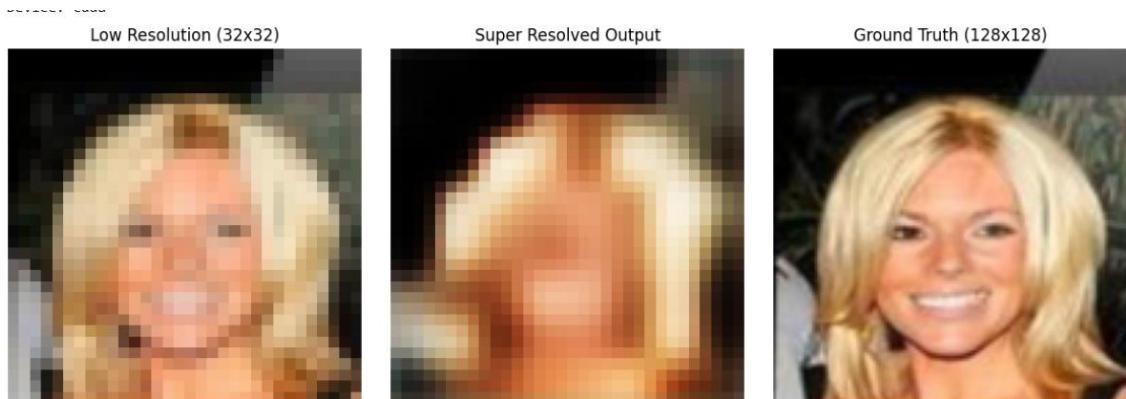
3.2.Performans Ölçütleri

- **Loss Grafikleri:** Eğitim süresince düzenli düşüş
- Görsel değerlendirme, nicel metriklerle uyumlu sonuçlar vermiştir

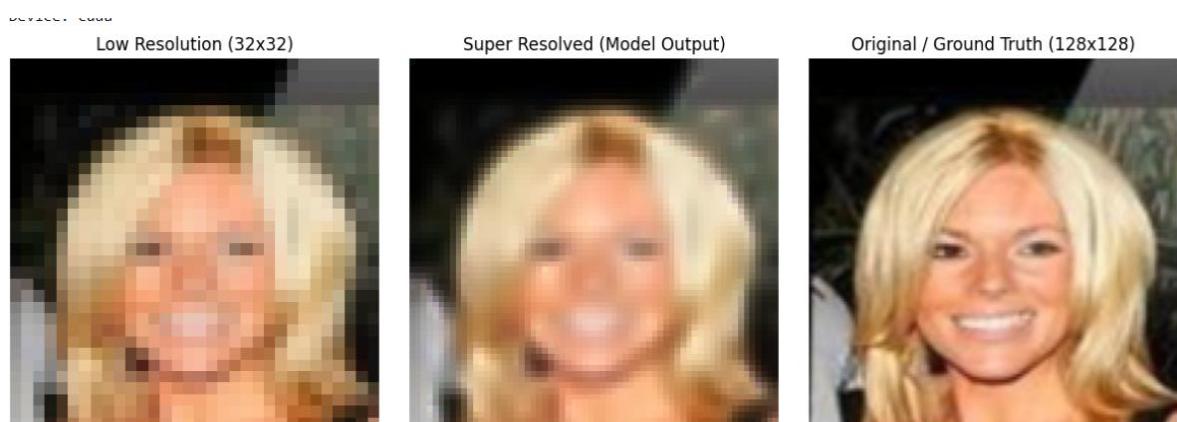
3.3.Görsel Sonuçlar

- İlk versiyon: Bulanık, detay kaybı yüksek
- Orta versiyonlar: Kenar ve yüz detaylarında iyileşme
- Final versiyon: Daha keskin kenarlar, daha doğal doku yapısı

Üretilen örnek görüntüler, modelin süper-çözünürlük görevini başarıyla yerine getirdiğini göstermektedir.



Şekil 2.2.3 Basit autoencoder görüntü sonuçları



Şekil 2.2.5 Strided convolution görüntü sonuçları



Şekil 2.2.7 Global Skip Connection görüntü sonuçları



Şekil 2.2.9 Final Model görüntü Sonuçları

4. TARTIŞMA VE SONUÇ

4.1. Bulguların Yorumu

Bu çalışmada, basit bir autoencoder yapısından başlanarak model adım adım geliştirilmiştir. Her versiyonda önceki modelin eksikleri analiz edilerek mimari, veri ön işleme ve kayıp fonksiyonları iyileştirilmiştir.

Özellikle:

- Residual learning
- Global skip connection
- Edge loss

bileşenlerinin süper-çözünürlük performansına önemli katkı sağladığı gözlemlenmiştir.

4.2. Gelecek Çalışmalar

Gelecek çalışmalarında aşağıdaki geliştirmeler yapılabilir:

- GAN tabanlı (SRGAN / ESRGAN) mimarilerin kullanılması
- Perceptual Loss (VGG Loss) eklenmesi
- Daha büyük veri setiyle eğitim
- PSNR ve LPIPS gibi ek metriklerin kullanılması

4.3. Sonuç

Bu proje kapsamında geliştirilen autoencoder tabanlı süper-çözünürlük modeli, düşük çözünürlüklü yüz görüntülerinden görsel olarak tatmin edici yüksek çözünürlüklü görüntüler üretmeyi başarmıştır. Elde edilen sonuçlar, derin öğrenmenin görüntü restorasyonu problemlerinde etkili bir çözüm sunduğunu göstermektedir.