

## Yüz Görüntüleri İçin Yeni Bir Süper Çözünürlük Yaklaşımı ve Farklı Görüntü Kalitesindeki Veri Setlerinin Başarı Performansı Üzerindeki Etkileri

### A New Super Resolution Approach for Face Images and the Effects of the Datasets of Different Image Quality on the Success Performance

<sup>1</sup>Fatih Mehmet ŞENALP 

<sup>1</sup>Konya Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Selçuklu/Konya, Türkiye

<sup>1</sup>fatih.senalp@gmail.com.tr

Araştırma Makalesi/Research Article

#### ARTICLE INFO

##### Article history

Received : 13 October 2022

Accepted : 28 January 2023

##### Keywords:

Super Resolution, Deep Learning, Image Quality Metrics, Datasets of Different Image Quality

#### ABSTRACT

The term of resolution is of great importance for various computer vision applications. In recent years, super resolution applications to increase image resolutions have become the focus of researchers, thanks to hardware advances. In this paper, a new deep learning-based super resolution model (SISRGAN) is proposed. In addition, three different datasets with different quality levels were created from the CelebA dataset for super resolution applications. The results obtained as a result of the studies were compared with the state-of-the-art models in the literature using image quality metrics. The proposed deep network model showed a superior success in terms of both visual quality improvement and metric values. The proposed model showed a success performance of approximately 0.5 db in PSNR values, 2% in SSIM values and 1 unit in BRISQUE values compared to the TSRGAN model, which achieved the second best values in terms of metric values. In addition, it has been seen that the low resolution image quality in which the super resolution image will be created directly affects the success.

© 2023 Bandırma Onyedi Eylül University, Faculty of Engineering and Natural Science. Published by Dergi Park. All rights reserved.

#### MAKALE BİLGİSİ

##### Makale Tarihleri

Gönderim : 13 Ekim 2022

Kabul : 28 Ocak 2023

##### Anahtar Kelimeler:

Süper Çözünürlük, Derin Öğrenme, Görüntü Kalite Metrikleri, Farklı Görüntü Kalitesindeki Veri Setleri

#### ÖZET

Çözünürlük kavramı, çeşitli bilgisayarlı görü uygulamaları için büyük önem arz etmektedir. Son yıllarda donanımsal ilerlemeler sayesinde görüntü çözünürlüklerini artırmaya yönelik süper çözünürlük uygulamaları araştırmacıların odak noktası haline gelmiştir. Bu çalışmada ise yeni bir derin öğrenme tabanlı süper çözünürlük modeli (SISRGAN) önerilmiştir. Ayrıca, süper çözünürlük uygulamaları için CelebA veri setinden farklı kalite seviyelerinde üç farklı veri seti oluşturulmuştur. Çalışmalar sonucunda elde edilen sonuçlar görüntü kalite metrikleri kullanılarak literatürde yer alan önemli modeller ile karşılaştırılmıştır. Önerilen derin ağ modelinin hem görsel kalitedeki iyileşme hem de metrik değerleri açısından daha üstün bir başarı ortaya koymuştur. Önerilen model, metrik değerleri bakımından en iyi ikinci değerleri elde eden TSRGAN modeline göre PSNR değerlerinde yaklaşık 0.5 db, SSIM değerlerinde %2 ve BRISQUE değerlerinde 1 birim daha üstün bir başarı performansı ortaya koymuştur. Ek olarak, süper çözünürlüklü görüntünün oluşturulacağı düşük çözünürlüklü görüntü kalitesinin başarıyı doğrudan etkilediği görülmüştür.

© 2023 Bandırma Onyedi Eylül Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi. Dergi Park tarafından yayınlanmaktadır. Tüm Hakları Saklıdır.

## 1. GİRİŞ

Çözünürlük, belirli bir birim içindeki piksel yoğunluğunun bir ölçüsüdür. Daha yüksek çözünürlüklü görüntüler, görüntü hakkında daha fazla ayrıntı sağlamaktadır. Bazı alanlarda, üzerinde çalışılacak görsellerin ayrıntılarını artırmaya önemli derecede ihtiyaç duyulmaktadır [1].

Süper çözünürlük, düşük çözünürlüklü görüntüleri minimum kayıpla geliştirmek için kullanılan bir tekniktir. Bu teknik, gürültü giderme ve bulanıklaştırma gibi çeşitli işlemleri içerebilmektedir. Görüntü iyileştirme, insan anlayışını geliştirmek ve hesaplamalı görevlerden daha yüksek doğruluk değerleri elde etmek için önemlidir. Daha yüksek çözünürlük, sahne hakkında daha fazla ayrıntı sağlamaktadır [2]. Süper çözünürlükteki ilk algoritmada Fourier dönüşümü tanıtıldı ve verilen çözüm araştırmacılar tarafından ilgiyle takip edildi. Bir sinir ağı ile çözülen bu çözümde, parmak izi görüntülerinin çözünürlüğünü iyileştirmek için çalışma gerçekleştirilmiştir. Algoritma tabanlı süper çözünürlük uygulamaları için süper çözünürlük terimi yapılacak olursa, yüksek çözünürlük (referans) görüntünün düşük çözünürlüklü karşılığında tahmini ve referans görüntüye yaklaşımı olarak tanımlanabilmektedir [3].

İnterpolasyon gibi süper çözünürlüğe benzer bazı geleneksel görüntü iyileştirme yöntemleri vardır. İnterpolasyon, süper çözünürlüğe benzer bir tekniktir ancak interpolasyon yüksek frekanslı ayrıntıları başarılı bir şekilde elde edememektedir. İnterpolasyon tekniği, en yakın komşu interpolasyonu, çift doğrusal interpolasyonu ve biküçük interpolasyonu gibi birkaç basit ve uygulaması kolay yöntemleri kapsamaktadır; fakat görsel kalite açısından da iyi sonuçlar göstermemektedirler. Bu yöntemler arasında en iyi sonuçları biküçük interpolasyon tekniği elde etmekte ve bu yöntem 4x4 (16 piksel değeri) komşuluk ilişkisinden yararlanmaktadır. Buna karşın çift doğrusal interpolasyon tekniğinde 2x2 komşuluk ilişkisi kullanılırken, en yakın komşu tekniğinde ise en yakın piksel noktasının değeri baz alınmaktadır [4].

Son yıllardaki donanım teknolojilerindeki ve derin öğrenme alanındaki gelişmeler süper çözünürlük uygulamalarına olan ilgiyi artırmıştır. Burada, derin ağ mimarilerinin geleneksel yöntemlere göre daha başarılı sonuçlar elde ettiği görülmüştür. Süper çözünürlük, birçok alanda önemli uygulamalara sahiptir. Örnek olarak; sahnelerde nesne tespiti (özellikle küçük nesnelere), gözetleme videolarında yüz tanıma, tıbbi görüntüleme, uzaktan algılamada görüntülerin yorumlanmasının iyileştirilmesi, astronomik görüntüler ve adli tıp gibi uygulama alanları gösterilebilmektedir [5]. Bunların yanı sıra son yıllarda termal ve hiperspektral görüntüleme gibi alanlarda da süper çözünürlük metodlarından yararlanılmaktadır [6].

Derin öğrenme tabanlı süper çözünürlük çalışmaları için ilk olarak evrişimli sinir ağları (CNN) tabanlı modeller kullanılmıştır. Dong vd. SRCNN modelini sunmuşlar ve düşük çözünürlüklü- yüksek çözünürlüklü görüntüler arasında uçtan uca haritalama (end to end mapping) tekniğini uygulamışlardır [7]. Ancak, bu tür sığ sinir ağları, giriş görüntülerinin çok seviyeli özelliklerini elde etme yeteneğini sınırlı tutmuştur. Önceki kenar görüntü süper çözünürlüğü için elverişli olmasına dikkat ederek, Liang vd. Derin ağ modelini eğitmek için ilk olarak düşük çözünürlüklü (DÇ) görüntüleri ile Sobel kenarlarını kullanmışlardır [8]. Bununla birlikte modelin başarısında bariz bir iyileşme elde edilememiştir. Daha sonra Kim vd. SRCNN modelinden esinlenerek tasarladıkları VDRS modelinde kalıntı öğrenme (residual learning) yönteminden yararlanmışlardır. Ayrıca evrişim katmanlarının sayısını 3 yerine 20 olarak düzenlemişler ve daha başarılı sonuçlara ulaşmışlardır. Burada, PSNR değeri 0,5-1 dB ve SSIM değeri ise %1-3 aralığında artış göstermiştir [9].

CNN tabanlı gerçekleştirilen uygulamalarda, daha çok ortalama kare hatayı (MSE) olabildiğince minimize etmek amaçlanmıştır. Son yıllarda ise görsellik açısından daha kaliteli görüntüler elde edebilmek amacıyla çekişmeli üretici ağlar (GAN) tabanlı modeller önerilmiştir [10]. Ledig vd. SRGAN modelini geliştirmişler ve bu modelin üretici ağına bağlantı atlama tekniği uygulamışlardır [11]. SRGAN modeli, görsellik açısından CNN tabanlı modellere göre daha başarılı sonuçlara ulaşılmasını sağlamıştır. Bu durum GAN tabanlı süper çözünürlük uygulamalarına olan ilgiyi artırmıştır. Park vd. yalnızca oluşturulan görüntüleri değil, aynı zamanda özellik alanındaki hiyerarşik özellikleri de ayırt etmek için iki ayırt ediciye sahip GAN yapısına benzer bir model olan SRFeat'i sunmuşlardır [12]. Bu ek ayırt edici ağ, süper çözünürlüklü görüntüleri oluştururken üretici ağı özellik yaklaşımına dikkat çekmeye zorlamaktadır. Daha sonra, Wang vd. SRGAN modelini geliştirerek yoğun blok (dense block) kullandıkları ESRGAN modelini önermişlerdir. Bu modelin yapısında, her yoğun bloğun girişi aynı zamanda sıradaki bloğun çıkışına da bağlıdır ve artık öğrenme (residual learning) tekniğini güçlendirecek şekilde bağlantılar düzenlenmiştir [13]. Burada, SRGAN modeli ile karşılaştırıldığında 0.1-0.4 dB aralığında artışlar elde ettiği gözlenmiştir. Yakın zamanda gerçekleştirilen diğer bir çalışmada ise Senalp ve Ceylan geliştirdikleri TSRGAN ile süper çözünürlük uygulamaları gerçekleştirilmiş ve oldukça başarılı sonuçlar elde etmişlerdir. Burada, SRGAN modeline göre PSNR değerinde 0.6 dB, SSIM değerinde ise %2 daha başarılı sonuçlar elde edilmiştir [14].

Bu çalışmada, literatürde yaygın olarak kullanılan veri seti üzerinde süper çözünürlük uygulamaları gerçekleştirmek için yeni bir derin öğrenme tabanlı model önerilmiştir. Ayrıca, kullanılan veri setindeki görüntü kalitesinin başarı performansı üzerindeki etkisini incelemek amacıyla, kullanılan veri setinden yararlanarak daha düşük kalite görüntülerden oluşan iki farklı yeni veri seti daha oluşturulmuştur. Sonraki bölümlerde bu veri setleri, geliştirilen derin ağ modeli ve sonuçların değerlendirilmesi hakkında detaylı bilgiler verilmiştir.

## 2. MATERYAL VE METOT

Bu bölümde, gerçekleştirilen süper çözünürlük uygulamaları için kullanılan ve farklı kalitede oluşturulan veri setleri, geliştirilen derin ağ mimarisi (SISRGAN) ve başarı performansı değerlendirme ölçütleri ile ilgili detaylı bilgiler verilmektedir.

### 2.1. Veri Setleri

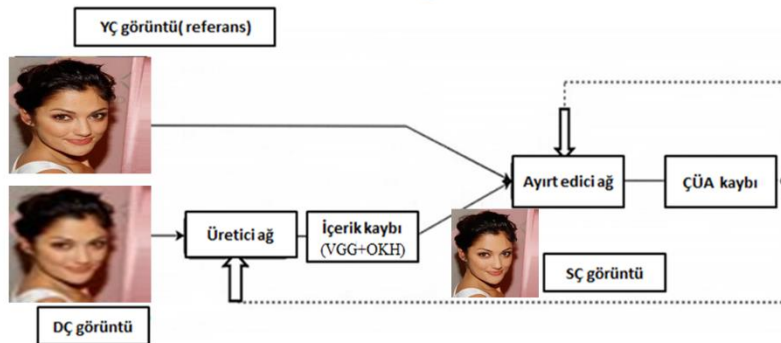
Bu çalışmada, önerilen derin ağ mimarisinin eğitim ve test işlemleri literatürde yaygın bir şekilde kullanılan CelebA veri seti [15] üzerinde yerine getirilmiştir. Bu veri setinden rastgele seçilen 2000 görüntüden 1900 tanesi eğitim, 100 tanesi ise test süreçleri için ayrılmıştır. Bu görüntüler yüksek çözünürlüklü (YÇ-referans) görüntüler olarak değerlendirilmiş ve bazı örnekler Şekil 1'de paylaşılmıştır. CelebA veri setine ait bu yüksek çözünürlüklü görüntüler 178x218 boyutlarındadırlar. Daha sonra bu yüksek çözünürlüklü (referans) görüntüler MATLAB ortamında 1/4 oranında alt ölçeklenerek (down-scaling) düşük çözünürlüklü görüntü veri seti oluşturulmuştur. Böylece 1900 yüksek-düşük çözünürlüklü görüntü çifti eğitim süreci için hazır hale getirilmiştir. Oluşturulan bu veri setine ek olarak, yüksek çözünürlüklü görüntülerin alt ölçeklenmesi sırasında %90 kalite faktörü ile kaydetme ve 4x4 maksimum havuzlama (max. pooling) yöntemleri kullanılarak daha az kalitede düşük çözünürlüklü görüntüler elde edilmiştir. Böylece üç farklı yüksek-düşük çözünürlüklü eğitim ve test veri setleri meydana getirilmiştir. Ayrıca, gerçekleştirilen süper çözünürlük uygulamaları sırasında eğitim ve test veri setleri birbirinden tamamen farklı olacak şekilde düzenlenmiştir.



Şekil 1. CelebA veri setine ait örnek görüntüler.

### 2.1. Derin Ağ Mimarisi

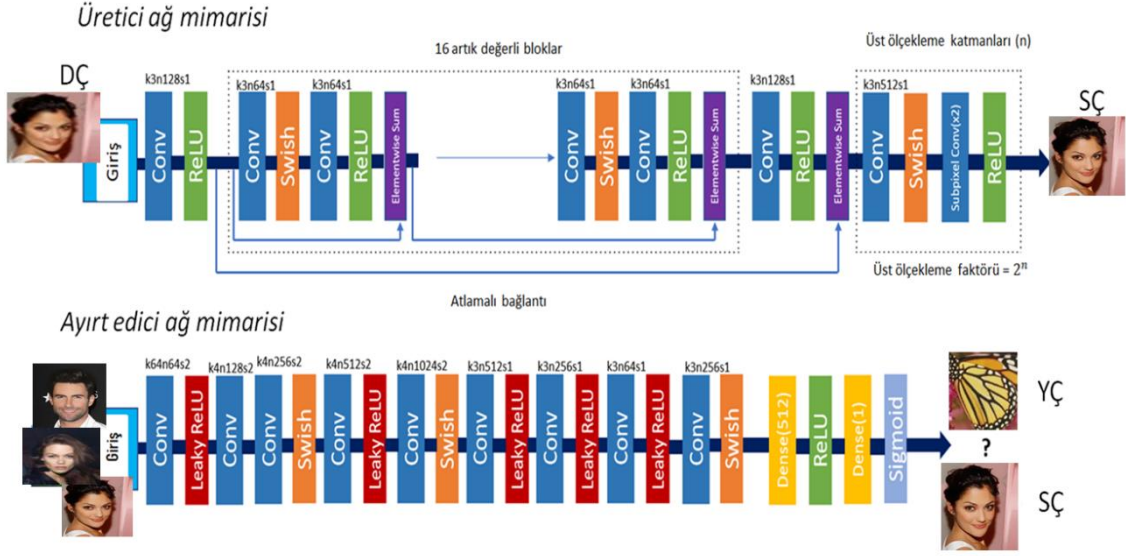
Bu bölüm, süper çözünürlük çalışmalarının başarısını artırmak ve daha hızlı bir eğitim süreci yerine getirebilmek amacıyla geliştirilen SISRGAN (Single Image Super Resolution Generative Adversarial Networks) modeli ile ilgili detaylı bilgiler vermektedir. SISRGAN modeli literatürde önemli yere sahip TSRGAN [4] modelinden esinlenerek daha yüksek kalitede süper çözünürlüklü görüntüler elde etmek amacıyla geliştirilmiştir. Önerilen bu model çekişmeli üretici ağlar (GAN) tabanlı bir model olup, üretici (generator) ve ayırt edici (discriminator) ağ kısımlarından meydana gelmektedir [16]. Üretici ağ, x4 üst ölçekleme yaparak süper çözünürlüklü görüntüyü (SÇ) meydana getirmektedir. Modelin çalışma prensibine göre ağ kaybından yararlanarak geri yayılım (backpropagation) tekniği uygulanmaktadır. SISRGAN mimarisinin genel gösterimi Şekil 2'de verilmiştir.



Şekil 2. SISRGAN mimarisi.

Önerilen SISRGAN modelinde yığın normalleştirme (batch normalization) katmanları kullanılmayarak eğitim hızı artırılmıştır. Ayrıca kullanılan swish aktivasyon fonksiyonu sayesinde negatif kısımda da öğrenme sağlanmıştır. Ek olarak, üretici ağ kısmında bağlantı atlama (skip connection) tekniği uygulanarak hem modelin başarısı artırılmış olup hem de eğitim süreci kolaylaştırılmıştır. Üretici ve ayırt edici ağların mimarileri Şekil 3'te gösterilmektedir. SISRGAN modeline ait derin ağ mimarisi genel olarak konvolüsyon, aktivasyon ve yoğun

(dense) katmanlarından meydana gelmektedir. Ayrıca, üretici ağ mimarisinde üst ölçekleme (up-scaling) işlemini gerçekleştirmeye yarayan alt piksel konvolüsyon (sub-pixel convolution) katmanları kullanılmaktadır. Geliştirilen modelde x4 üst ölçekleme faktörü için iki adet alt piksel konvolüsyon katmanı yer almaktadır.



Şekil 3. Üretici ve ayırt edici ağ mimarileri.

Matematiksel İfade:

$$\delta_G = \min \frac{1}{N} \sum_{n=1}^N L_{SISRGAN}(G_{\delta_G}(I_n^{DÇ}), I_n^{YÇ}) \quad (1)$$

Denklem 1’de verilen formülde  $\delta_G$ , derin ağ modelinin ağırlık ve sapma değerlerini ifade ederken G ise SÇ görüntüyü temsil etmektedir. Ayrıca,  $I_n^{DÇ}$  ve  $I_n^{YÇ}$  ise sırasıyla eğitim veri setine ait DÇ ve YÇ örnek görüntüleri temsil etmektedir. Algısal kayıp (perceptual loss) değerini temsil eden  $L_{SISRGAN}$ , aynı zamanda geri yayılım yönteminde kullanılan derin ağın kaybıdır ve metrik değerlerindeki iyileşmenin yanı sıra görsel kaliteyi de artırabilmek amacıyla kullanılmaktadır. Bu algısal kayıp değeri Denklem 2’de verilen formülle hesaplanmaktadır [17].

$$L_{\text{toplam}(SISRGAN)} = L_{\text{mse}} + L_{\text{vgg}} + L_{\text{çekişmeli}} \quad (2)$$

Ortalama Kare Hata (MSE), derin ağ modellerinin başarısını gösteren pozitif değerli bir ölçüttür. MSE değerinin azalması modelin başarısının arttığını göstermektedir [18]. MSE, Denklem 3’te verilen formülle hesaplanmaktadır ve n veri setinin boyutunu,  $e_j$  ise tahmin ile esas değer arasındaki hata değerini temsil etmektedir.

$$MSE = \frac{1}{n} \sum_{j=1}^n e_j^2 \quad (3)$$

VGG kayıp değeri hesaplanırken önceden eğitilmiş VGG19 modelinden yararlanılmış olup, hesaplama formülü Denklem 4’te verilmiştir.  $G_{\phi}(I^{DÇ})$ , SÇ görüntüyü,  $I^{YÇ}$  ise YÇ görüntüyü temsil etmektedir. Ayrıca  $\psi_{i,j}$ , i. maksimum havuzlama (maxpooling) katmanından önce j. konvolüsyon katmanındaki özellik haritalarını gösterirken,  $W_{i,j}$  ve  $H_{i,j}$  özellik haritalarının boyutlarını göstermektedir [19].

$$L_{VGG_{i,j}}^{SÇ} = \frac{1}{W_{i,j} H_{i,j}} \sum_{m=1}^{W_{i,j}} \sum_{n=1}^{H_{i,j}} (\psi_{i,j}(I^{YÇ})_{m,n} - \psi_{i,j}(G_{\phi}(I^{DÇ}))_{m,n})^2 \quad (4)$$

Çekişmeli kayıp (adversarial loss), daha kaliteli görüntüler oluşturabilmek amacıyla yapısal olarak benzerliği artırma esasına dayanarak hesaplanan bir kayıp değeridir. Bu kayıp değerinin formülü Denklem 5’te verilmiştir.

Ayırt edici ağıın tüm eğitim veri seti üzerindeki olasılıklarına bağlı olarak hesaplanmaktadır.  $D_\alpha(G_\beta(I^{DC}))$ , SÇ görüntünün YÇ (referans) görüntüye ne ölçüde benzediğini göstermektedir.

$$L_{\text{Çekişmeli}}^{S\text{Ç}} = \sum_{n=1}^N -\log D_\alpha(G_\beta(I^{DC})) \quad (5)$$

## 2.2. Değerlendirme Kriterleri

Tepe sinyal gürültü oranı (PSNR), iki sinyal arasındaki sinyal kalitesini etkileyen bozulma gürültüsünün etkisine bağlı bir oran hesaplama amacıyla kullanılmaktadır. İki görüntü için hesaplanan bu oran desibel formunda hesaplanmaktadır. PSNR değeri, sinyaller geniş dinamik aralıklara sahip olduğu için desibel ölçeğinin logaritma terimi olarak hesaplanmaktadır. Tepe sinyal gürültü oranı, düşük kalitedeki görüntülerin yeniden oluşturulması sonucu süper çözünürlüklü görüntülerin kalitelerini değerlendirmek için kullanılan en yaygın ve önemli kalite değerlendirme metriklerindedir. Orjinal sinyal referans veri olarak kabul edilirken, gürültü ise sıkıştırma veya bozulmadan kaynaklanan hata olarak tanımlanmaktadır. PSNR, yeniden yapılandırma kalitesine ilişkin insan algısının yaklaşık tahmini olarak ifade edilebilmektedir. PSNR değerinin hesaplama formülü Denklem 6'da verilmiştir. Bu denklemde,  $I_0$  referans görüntüyü,  $I$  ise oluşturulan süper çözünürlüklü görüntüyü temsil etmektedir [20].

$$PSNR = 20 \times \log_{10} \left( 255 \sqrt{MSE(I, I_0)} \right) \quad (6)$$

Yapısal benzerlik indeksi (SSIM), algıya dayalı bir metriktir. Bu yöntemde görüntü bozulması, yapısal bilgilerdeki algısal değişiklik olarak kabul edilmektedir. Yapısal bilgi terimi, birbirine güçlü şekilde bağlı pikseller veya uzamsal olarak kapalı pikseller hakkında vurgu yapmaktadır. Bu güçlü bir şekilde birbirine bağımlı pikseller, görüntü alanındaki görsel nesnelere hakkında daha önemli bazı bilgileri belirtmektedir. SSIM, görüntü ve videoların algılanan görsel kalitesini tahmin etmektedir. Yani, orijinal ve yeniden oluşturulmuş iki görüntü arasındaki benzerliği ölçmektedir. SSIM değerinin hesaplama formülü Denklem 7'de verilmiş olup,  $\mu_R(n)$  ve  $\mu_D(n)$  sembolleri, referans ( $I_R$ ) ve bozuk ( $I_D$ ) görüntü dizisinin  $n$ . çerçevesinin ortalamasını temsil etmektedir. Ayrıca,  $\sigma_R(n)$  ve  $\sigma_D(n)$ , referans ( $I_R$ ) ve bozuk ( $I_D$ ) görüntü dizisinin  $n$ . çerçevesinin standart sapmasını temsil etmektedir.  $C1$ ,  $C2$  ise kararsızlıktan kaçınmak için kullanılan sabitlerdir [21].

$$SSIM = \frac{[2\mu_R(n)\mu_D(n) + C1][2\sigma_R(n)\sigma_D(n) + C2]}{[\mu_R^2(n) + \mu_D^2(n) + C1][\sigma_R^2(n) + \sigma_D^2(n) + C2]} \quad (7)$$

Kör/referanssız görüntü uzamsal kalite değerlendiricisi (BRISQUE), referans istemeyen bir ölçüm metriğidir. Bu algoritma istatistik tabanlı ve referanssız bir görüntü kalite değerlendirme metodudur. BRISQUE algoritmasının içindeki detay kayıpları, istatistiksel gözlem ve yöntemler yardımıyla herhangi bir referansa ihtiyaç duyulmadan hesaplanmaktadır. Hesaplanan bu değer 0'a yaklaştıkça detay kaybı azalacağından dolayı görüntünün iyileştirilmesi çalışmalarının daha başarılı olduğu kabul edilmektedir.

## 3. GERÇEKLEŞTİRİLEN SÜPER ÇÖZÜNÜRLÜK UYGULAMALARI

### 3.1. Eğitim Süreçlerinin Detayları

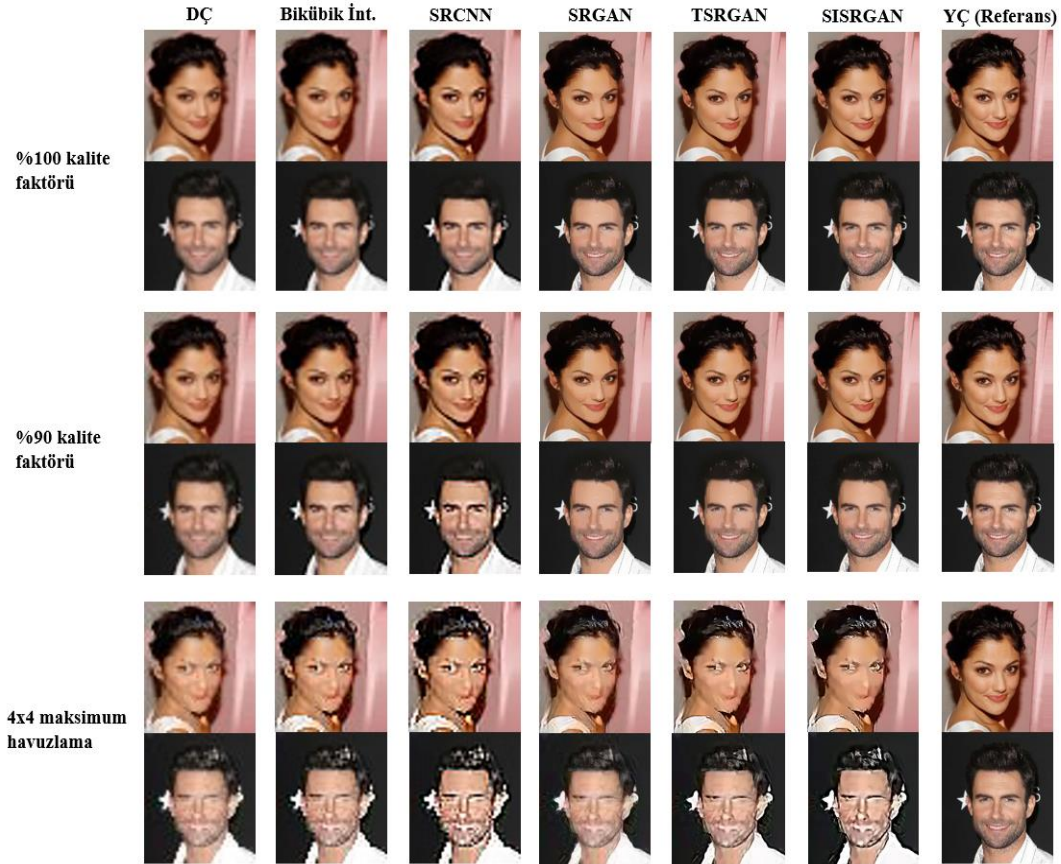
Bu çalışmada, derin öğrenme tabanlı süper çözünürlük modellerinin eğitim süreçleri için CPU E5-2680 işlemci ve 32 gb GeForce GTX 1080 Ti ekran kartına sahip bir iş istasyonu kullanılmıştır. Derin ağların eğitimi 10.000 iterasyon (epoch) sonunda tamamlanmıştır. Burada, önerilen SISRGAN modelinin diğer modellere göre eğitim hızının daha yüksek olduğu gözlemlenmiş olup, ağı eğitimi 19 saat sürmüştür. Ek olarak, TSRGAN ağı eğitimi 23 saat sürerken, SRGAN ağı eğitimi ise 25 saatte tamamlanmıştır. Ayrıca, önerilen modelde optimizasyon algoritması olarak Adam algoritması kullanılmış ve öğrenme oranı 0.001 tercih edilmiştir.

### 3.2. Süper Çözünürlük Uygulamalarının Sonuçları

Bu bölümde, oluşturulan üç farklı veri seti üzerinde gerçekleştirilen süper çözünürlük uygulamalarının sonuçları ayrı ayrı olacak şekilde paylaşılmıştır. CelebA veri seti ve bu veri setinden faydalanarak oluşturulmuş diğer iki veri seti için yürütülen çalışmalarda x4 üst ölçekleme tekniği kullanılmıştır. Çalışmalar sonucu elde edilen süper çözünürlüklü görüntülerin kalitesi görüntü kalite metrikleri olan PSNR ve SSIM ölçütleri baz alınarak değerlendirilmiştir. Ayrıca, görsel bir değerlendirme yapabilmek amacıyla düşük çözünürlüklü, yüksek çözünürlüklü ve üretilen süper çözünürlüklü görüntülere ait bazı örnekler Şekil 4'te paylaşılmıştır. Burada, Bölüm 2.1'de detaylıca açıklanan üç farklı veri seti ile gerçekleştirilen çalışmaların sonuçları ayrı ayrı sunulmuştur. Elde edilen görüntüler incelendiğinde %100 ve %90 kalite faktörü ile oluşturulan veri setleri için yerine getirilen çalışmalarda sonuçların görsel olarak yakın olduğu ve %100 kalite faktörü ile oluşturulan

görüntülerin daha kaliteli olduğu gözlemlenmiştir; fakat 4x4 maksimum havuzlama yöntemi ile elde edilen görüntüler üzerinde yürütülen çalışmada ise başarının daha düşük seviyede kaldığı görülmektedir. Bu durum, elde edilen düşük çözünürlüklü görüntülerde piksel değerlerine dair veri kaybının çok daha yüksek olmasından kaynaklanmaktadır. Bununla birlikte maksimum havuzlama yöntemiyle bir veri seti oluşturularak ilk defa gerçekleştirilen süper çözünürlük uygulamasında, önerilen SISRGAN modelinin literatürde yer alan önemli modellere (SRCNN, SRGAN, TSRGAN) karşı daha üstün bir başarıya ulaştığı gözlenmiştir çünkü kenar detay bilgileri (yüksek frekans bilgileri) daha iyi şekilde oluşturularak daha net süper çözünürlüklü görüntüler meydana getirilmiştir. Şekil 4 incelendiğinde, özellikle SISRGAN modeli olmak üzere süper çözünürlük tekniklerinin görüntüleri iyileştirme konusunda hem bulanıklık giderme hem de yüksek frekans bilgilerinin elde edilmesi açısından başarılı sonuçlar elde ettiği görülmüştür. Elde edilen bu görsel sonuçlar, süper çözünürlüklü görüntü formunun oluşturulacağı düşük çözünürlüklü ham görüntünün çözünürlüğünün başarı performansı üzerinde oldukça etkili olduğunu göstermiştir.

Geliştirilen derin ağ modelinin başarısını sayısal değerlerle de destekleyebilmek amacıyla üç farklı veri seti için PSNR ve SSIM değerleri bağımsız olarak hesaplanmıştır. Hesaplanan görüntü kalite metriklerinin sonuçları Tablo 1’de verilmiş ve en iyi değerler kalın, en iyi ikinci değerler ise altı çizili olarak gösterilmiştir. Burada, önerilen SISRGAN modelinin genel çerçevede en başarılı sonuçları elde ettiği gözlenirken, TSRGAN modelinin ise en iyi ikinci başarıya sahip olduğu görülmektedir. SISRGAN modeli TSRGAN modeline göre PSNR değerlerinde 0,3-0,5 dB aralığında üstün bir başarı performansı ortaya koymuştur. Ayrıca, SSIM değerleri dikkate alındığında yaklaşık %2 daha başarılı sonuçlar elde etmiştir; ancak maksimum havuzlama ile elde edilen veri setinde ise TSRGAN modeline yakın bir başarı ile en iyi ikinci değere sahiptir. Ek olarak, PSNR odaklı modeller olarak değerlendirilen evrişimli sinir ağları tabanlı SRCNN modeline karşı da oldukça yüksek değerlere ulaşarak görsel kalitenin yanında metrik değerleri açısından da SISRGAN modeli başarısını kanıtlamıştır. Ayrıca, PSNR ve SSIM metriklerinin yanı sıra BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator) metriği ile sonuçlar değerlendirilmiştir. Referans tabanlı olmayan bu metrik algısal kaliteyi ölçmek amacıyla kullanılmıştır. BRISQUE metrik değeri 0-100 aralığında değerler almakta ve düşük değer daha iyi bir algısal kaliteyi ifade etmektedir. Tablo 1’de paylaşılan sonuçlar önerilen modelin BRISQUE metriği açısından da literatürdeki modellere göre başarılı olduğu ve referans görüntünün değerlerine oldukça yaklaştığı gözlemlenmiştir.



Şekil 4. Farklı veri setleri için elde edilen süper çözünürlüklü yüz görüntüleri.

**Tablo 1.** PSNR ve SSIM sonuçları.

	DÇ	Bikübbik İnterpolasyon	SRCNN [7]	SRGAN [11]	TSRGAN [14]	SISRGAN (Önerilen)	YÇ (Referans)
<b>%100 kalite faktörü</b>							
PSNR	21,437	23,238	24,956	24,663	<u>25,294</u>	<b>25,714</b>	$\infty$
SSIM	0,723	0,749	0,785	0,771	<u>0,794</u>	<b>0,816</b>	<b>1</b>
BRISQUE	36,685	31,603	30,311	27,093	<u>24,967</u>	<b>23,623</b>	<b>17,449</b>
<b>%90 kalite faktörü</b>							
PSNR	21,118	22,826	24,552	24,419	<u>25,046</u>	<b>25,315</b>	$\infty$
SSIM	0,671	0,698	0,758	0,734	<u>0,761</u>	<b>0,783</b>	<b>1</b>
BRISQUE	38,683	35,804	32,543	29,965	<u>27,604</u>	<b>26,970</b>	<b>17,449</b>
<b>4x4 maksimum havuzlama</b>							
PSNR	15,762	16,912	<u>20,364</u>	19,814	20,148	<b>20,554</b>	$\infty$
SSIM	0,604	0,626	0,685	0,668	<b>0,692</b>	<u>0,690</u>	<b>1</b>
BRISQUE	52,578	49,635	48,001	<u>41,877</u>	42,640	<b>41,822</b>	<b>17,449</b>

#### 4. SONUÇ

Bu makalede, görünür yüz görüntüleri için yeni bir derin öğrenme tabanlı süper çözünürlük modeli önerilmiştir. Önerilen SISRGAN modeli çekişmeli üretici ağlar tabanlı bir model olup, hem algısal olarak görsel kaliteyi artırmak hem de görüntü kalite metrikleriyle bunu gösterebilmek amacıyla geliştirilmiştir. Ayrıca, derin ağı eğitimi ve test aşamalarında kullanılan düşük çözünürlüklü görüntülerin kalite seviyesinin, yeniden oluşturulan süper çözünürlüklü görüntülerin kalitesine olan etkisini gözlemleyebilmek için CelebA veri setinden oluşturulmuş üç farklı kalite seviyesine sahip veri setleri üzerinde çalışmalar yerine getirilmiştir. Çalışmalar sonucu elde edilen bütün sonuçlar PSNR ve SSIM değerleri dikkate alınarak literatürde yer alan yeni ve üstün başarılı modeller ile karşılaştırılmıştır. Sonuçlar değerlendirildiğinde önerilen derin ağı modelinin diğer modellere karşı oldukça üstün bir başarıya sahip olduğu gözlemlenmiştir. Bununla birlikte, kalite faktörleri ile oluşturulan veri setleri üzerindeki sonuçlar birbirine yakın değerlerde olmasına rağmen maksimum havuzlama tekniği ile elde edilen veri seti için yürütülen çalışmaların sonuçları daha düşük değerlerdedir. Bu durumun en büyük sebebi, süper çözünürlüklü görüntünün üretileceği düşük çözünürlüklü görüntüdeki veri kaybının çok daha yüksek seviyede olmasıdır. Önerilen model bu veri seti üzerindeki çalışmalarda da en iyi PSNR ve en iyi ikinci SSIM değerlerine ulaşarak, düşük kalitedeki görüntüler üzerinde de uygulanabilir olduğunu kanıtlamıştır. Ek olarak, bulanıklık giderme ve kenar detay bilgilerinin elde edilmesi açısından da başarılı sonuçlar gösteren SISRGAN modeli görsel kalite açısından da önemli oranda iyileşmeler göstermiştir. Burada, maksimum havuzlama ile elde edilmiş düşük çözünürlüklü görüntülerdeki iyileşmeler biraz daha düşük seviyede kalmıştır, çünkü bu görüntüler diğer veri setlerindeki görüntülere göre çok daha düşük görsel kaliteye sahiptir. Sonuç olarak, yüksek çözünürlüklü referans görüntülere yakın kalitede süper çözünürlüklü görüntüler oluşturma bakımından %100 ve %90 kalite faktörleri ile türetilen veri setleri üzerindeki çalışmalar hem görsel hem de metrik değerleri dikkate alındığında daha başarılı sonuçlar ortaya koymuştur. Gerçekleştirilen tüm süper çözünürlük uygulamalarının gösterdiği yüksek başarı performansı, görüntüleme cihazlarının kullanıldığı projelerde daha düşük maliyetli kameralardan faydalanarak proje maliyetini düşürme konusunda araştırmacılara bir alternatif sunmaktadır. Ayrıca, görüntü detaylar bilgilerinin önemli olduğu hastalık teşhisi gibi medikal uygulamalar için süper çözünürlük yöntemlerinin faydalı olabileceği öngörülmektedir.

#### Yazar Katkıları

Fatih Mehmet Şenalp: Makalenin kapsamının belirlenmesine, araştırmaların yürütülmesine, makalenin yazımına ve sonuçların değerlendirilmesine katkı sağlamıştır.

#### Çıkar Çatışması

Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan ederler.

#### KAYNAKÇA

- [1] M. Protter, M. Elad, H. Takeda and P. Milanfar "Generalizing the nonlocal-means to super-resolution reconstruction", IEEE Transactions on image processing, vol. 18, no. 1, pp. 36-51, 2008.
- [2] R. W. Gerchberg "Super-resolution through error energy reduction", Optica Acta: International Journal of Optics, vol. 21, no. 9, pp. 709-720, 1974.
- [3] K. Nasrollahi and T. B. Moeslund "Super-resolution: a comprehensive survey. Machine vision

- and applications”, vol. 25, no. 6, pp. 1423-1468, 2014.
- [4] H. Hurkal and Z. Orman “A survey on image super-resolution with generative adversarial networks”, *Acta Infologica*, vol. 4, no. 2, pp. 139-154, 2020.
- [5] L. Yue et al. “Image super-resolution: the techniques, applications, and future”, *Signal Processing* vol. 128, pp. 389-408, 2018.
- [6] R. Yan, K. Yang and K. Wang “NLFNet: Non-Local Fusion Towards Generalized Multimodal Semantic Segmentation across RGB-Depth, Polarization, and Thermal Images”, 2021 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 1129-1135, 2021.
- [7] C. Dong, C.C. Loy, K. He, X. Tan “Image super-resolution using deep convolutional networks”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, pp. 295-307, 2015.
- [8] Y. Liang, J. Wang, S. Zhou, Y. Gong and N. Zheng “Incorporating image priors with deep convolutional neural networks for image super-resolution”, *Neurocomputing*, vol. 194, pp. 340-347, 2016.
- [9] J. Kim, J.K. Lee and K.M. Lee “Accurate image super-resolution using very deep convolutional networks”, *IEEE CVPR*, pp. 1646-1654, 2016.
- [10] I. Goodfellow et al. “Generative adversarial networks”, In *Advances in Neural Information Processing Systems (NIPS)*, pp. 2672-2680, 2014.
- [11] C. Ledig et al. “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network”, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Honolulu, pp. 105-114, 2017.
- [12] S.J. Park, H. Son, S. Cho, K.S. Hong, S. Lee “Srfeat: single image super-resolution with feature discrimination”, In: *Proceedings of the European conference on computer vision (ECCV)*, pp. 439-455, 2018.
- [13] M. Wang et al. “Improved face super-resolution generative adversarial networks”, *Machine Vision and Applications*, vol. 31, p. 22, 2020.
- [14] F.M. Senalp and M. Ceylan “Deep learning based super resolution and classification applications for neonatal thermal images”, *Traitement du Signal* vol. 38, no. 5, pp. 1361-1368, 2021.
- [15] Z. Liu, P. Luo, X. Wang, and X. Tang “Deep Learning Face Attributes in the Wild”, in *Proc. IEEE/CVF International Conference on Computer Vision*, 2015.
- [16] A. Radford, L. Metz and S. Chintala “Unsupervised representation learning with deep convolutional generative adversarial networks”, *arXiv preprint arXiv:1511.06434*, 2015.
- [17] J. Johnson, A. Alahi and F. Li “Perceptual losses for real-time style transfer and super resolution”, In *European Conference on Computer Vision (ECCV)* Springer, pp. 694-711, 2016.
- [18] D. Weixiang, P. Addepalli and Y. Zhao “The Spatial Resolution Enhancement for a Thermogram Enabled by Controlled Sub-pixel Movements”, *IEEE Transactions on Instrumentation and Measurement* vol. 69, no. 6, pp. 3566-3575, 2019.
- [19] Y. Gu et al. “MedSRGAN: medical images super-resolution using generative adversarial networks”, *Multimed Tools Appl.* vol. 79, pp. 21815-21840, 2020.
- [20] H. Javaid et al. “Video colour variation detection and motion magnification to observe subtle changes”, *M.Sc. Thesis Blekinge Institute of Technology Faisalabad Pakistan*, 2013.
- [21] A. Dosovitskiy and T. Brox “Generating images with perceptual similarity metrics based on deep networks”, In *Advances in Neural Information Processing Systems (NIPS)*, pp. 658-666, 2016.