




Tek Görüntü Süper Çözünürlük Uygulamaları İçin Görsel Kaliteyi İyileştirmeye Yönelik Yeni Bir Yaklaşım

Fatih Mehmet Şenalp^{1*} 

¹ Konya Teknik Üniversitesi, Elektrik Elektronik Mühendisliği, Konya, Türkiye

fatih.senalp@gmail.com

Öz

Tek görüntü süper çözünürlük problemi, literatürde çeşitli derin öğrenme tabanlı teknikler kullanılarak kapsamlı çalışmalar yapılmıştır. Derin evrişimli ağlar tabanlı süper çözünürlük, çok sayıda pratik uygulama ile beraber hızla büyüyen bir ilgi alanı haline gelmiştir. Bununla birlikte derin öğrenme tabanlı ilk çalışmalar evrişimli sinir ağları tabanlı olup, tepe sinyal gürültü oranı odaklı çalışmalardır. Son yıllardaki çekişmeli üretici ağlar tabanlı geliştirilen modeller sayesinde görsel kaliteyi artırmak esas amaç olarak belirlenmiştir; fakat bu durum görüntü kalite metrikleri incelendiğinde görülmektedir. Bu çalışmada ise eğitim sırasında kullanılan ağ kaybı için hem ortalama kare hata hem de algısal kayıp değerlerinden faydalanılmıştır. Ayrıca, üç farklı eğitim veri setinin birleşimi yeni bir eğitim veri seti olarak kullanılmıştır. Bu etmenlerin sonucunda hem görsel kalite artırılmış hem de görüntü kalite metrik değerlerinde ciddi bir artış yakalanmıştır. Ek olarak, yığın normalleştirme katmanları ağ mimarisine dahil edilmemiş ve bağlantı atlama tekniği kullanılarak derin ağ mimarisinin eğitim hızı artırılmıştır. Önerilen modelin başarı performansı literatürde yer alan önemli modeller ile karşılaştırılmıştır. Burada, tepe sinyal gürültü oranı ve yapısal benzerlik indeksi değerleri literatürde yaygın kullanılan üç farklı test veri seti için ayrı ayrı hesaplanmış ve değerlendirilmiştir. Elde edilen sonuçlar değerlendirildiğinde önerilen modelin diğer modellere göre daha başarılı olduğu ve daha kaliteli görüntüler oluşturduğu görülmektedir. Tüm bulgular değerlendirildiğinde önerilen modelin diğer modellere kıyasla hem başarı hem de eğitim hızı bakımından daha verimli bir model olduğu görülmektedir.

Anahtar kelimeler: Süper çözünürlük, Derin öğrenme, Görüntü kalite metrikleri, Yeni eğitim veri seti

A New Approach to Improve Visual Quality for Single Image Super Resolution Applications

Abstract

The single image super resolution problem has been extensively studied in the literature using various deep learning-based techniques. Super resolution based on deep convolutional networks has become a rapidly growing area of interest with many practical applications. However, the first studies based on deep learning were based on convolutional neural networks and peak signal-to-noise ratio (PSNR) oriented. Thanks to the models developed based on generative adversarial networks (GAN) in recent years, it has been determined as the main objective to increase the visual quality; however, this is not seen when the image quality metrics are examined. In this study, both mean square error and perceptual loss values were used for the network loss used during the training of the network. Also, the combination of three different training datasets was used as a new training dataset. As a result of these factors, both the visual quality has been increased and a significant increase has been achieved in the image quality metric values. In addition, batch normalization layers are not included in the network architecture and the training speed of the deep network architecture is increased by using the skip connection technique. The success performance of the proposed model was compared with the state-of-the-art models in the literature. Here, the peak signal to noise ratio (PSNR) and structural similarity index measure (SSIM) values were calculated and evaluated separately for three different test data sets commonly used in the literature. When the obtained results are evaluated, it is seen that the proposed model is more successful than other models and generates higher quality images. When all the findings are evaluated, it is seen that the proposed model is a more efficient model in terms of both success and training speed compared to state-of-the-art models.

Keywords: Super resolution, Deep learning, Image quality metrics, New training dataset

* Sorumlu yazar.
E-posta adresi: fatih.senalp@gmail.com

Alındı : 3 Ekim 2022
Revizyon : 13 Ocak 2023
Kabul : 16 Ocak 2023

1. Giriş (Introduction)

Süper Çözünürlüğün (SC) temel amacı, daha düşük çözünürlüklü görüntülerden daha yüksek çözünürlüklü bir görüntü oluşturmaktır. Yüksek çözünürlüklü görüntü, yüksek piksel yoğunluğu ve dolayısıyla orijinal sahne hakkında daha fazla ayrıntılar sunmaktadır (Toyran, 2008). Görüntülerin örüntü tanıma ve analizinde daha iyi performans için bilgisayarlı görü uygulamalarında yüksek çözünürlük ihtiyacı yaygındır. Medikal görüntüleme için yüksek çözünürlük önemlidir (Gu vd., 2020). Ayrıca, gözetim, adli ve uydular görüntüleme uygulamaları gibi birçok uygulama, yüksek çözünürlüğün gerekli olduğu görüntüde belirli bir ilgi alanının yakınlaştırılmasını gerektirmektedir. Ancak, yüksek çözünürlüklü görüntüleri elde etmek her zaman mümkün olmamaktadır. Bunun nedeni, yüksek çözünürlüklü görüntüleme sistemlerinin kurulumunun pahalı olması ve ayrıca sensörün, optik üretim teknolojisinin doğal sınırlamalarıdır (Zhang vd., 2018). Bu problemler, nispeten ucuz olan ve süper çözünürlük kavramına yol açan görüntü işleme algoritmalarının kullanılmasıyla aşılabilir. Daha az maliyetli olabileceği ve mevcut düşük çözünürlüklü görüntüleme sistemlerinden yararlanılabildiği için avantaj sağlamaktadır (Anwar vd., 2020).

Süper çözünürlük, bir sahnenin düşük çözünürlüklü (gürültülü) görüntü dizisinin bir kombinasyonunun, yüksek çözünürlüklü bir görüntü veya görüntü dizisi oluşturmak için kullanılabileceği fikrine dayanmaktadır. Yani, daha düşük çözünürlükte gözlenen bir dizi görüntü verildiğinde, orijinal sahne görüntüsü yüksek çözünürlükle yeniden oluşturulmaya çalışılmaktadır. Süper çözünürlük uygulamalarındaki genel yaklaşım, düşük çözünürlüklü görüntülerin yüksek çözünürlüklü bir görüntünün yeniden örneklenmesinden elde edildiği şeklindedir (Chudasama vd, 2020). Daha sonra, girdi görüntülerine ve görüntüleme modeline dayalı olarak daha yüksek çözünürlüğe sahip süper çözünürlüklü görüntüler yeniden oluşturulmaktadır. Bu nedenle, görüntüleme modelinin doğruluğu, süper çözünürlük için hayati önem taşımakta ve yanlış bir modelleme görüntüyü daha da bozabilmektedir. Süper çözünürlük uygulamalarında, gözlenen görüntüler bir veya birden fazla kameradan alınabildiği gibi bir video dizisinin kareleri şeklinde de olabilmektedir. Bu görüntülerin ortak bir referans çerçevesine eşlenmesi gerekmektedir. Bu işlem kayıt olarak ifade edilmektedir. Ayrıca süper çözünürlük prosedürü istenildiği takdirde hizalanmış bileşik görüntüdeki ilgili bir bölgeye de uygulanabilmektedir (Glasner vd., 2009).

Görüntü işleme alanında önemli bir yere sahip olan süper çözünürlük tekniği; tıbbi görüntüleme, uydular görüntüleme, gözetleme ve güvenlik, astronomik görüntüleme, askeri görüntüleme vb. uygulama alanlarında kullanılabilmektedir (Yue vd., 2018). Ayrıca, son yıllarda termal görüntüleme üzerinde de

süper çözünürlük uygulamaları yaygınlaşmıştır (Yan vd., 2021).

Süper çözünürlük kavramı üzerine gerçekleştirilen ilk çalışmalarda faydalanılan basit yöntemler en yakın komşu (nearest neighbor), çift doğrusal (bilinear), biküçük interpolasyon metotlarıdır. Bu teknikler görüntüdeki piksellerin komşuluk ilişkisine dayanmaktadır. Biküçük interpolasyon yöntemi diğer iki yöntemle kıyasla daha yüksek kalitede görüntüler elde edilmesini sağlamaktadır. Buna rağmen kenar detay bilgilerinin elde edilmesi gibi etmenler için yetersiz kalmaktadır. 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 (Dong vd., 2016).

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 (Dong vd., 2015). 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 değerleri ise %1-3 aralığında artış gerçekleşmiştir (Kim vd., 2016).

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 (Goodfellow vd., 2014). Ledig vd. SRGAN modelini geliştirmişler ve bu modelin üretici ağında bağlantı atlama tekniği uygulamışlardır. SRGAN modeli, görsellik açısından CNN tabanlı modellere göre daha başarılı sonuçlara ulaşılmasını sağlamıştır (Ledig vd., 2017). Bu durum GAN tabanlı süper çözünürlük uygulamalarına olan ilgiyi artırmıştır. Wang vd. SRGAN modelini geliştirerek yoğun blok (dense block) kullandıkları ESRGAN modelini önermişlerdir. Burada, SRGAN modeli ile karşılaştırıldığında 0.1-0.4 dB aralığında artışlar elde ettiği gözlenmiştir (Wang vd., 2020). 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 (Senalp ve Ceylan, 2021).

2. Materyal ve Metot (Material and Method)

Bu bölümde, gerçekleştirilen süper çözünürlük uygulamalarında kullanılan veri setleri, geliştirilen derin ağ mimarisi ve başarı performansı değerlendirme ölçütleri ile ilgili detaylı bilgiler paylaşılmaktadır.

2.1. Veri setleri (Datasets)

Bu çalışmada, derin ağ mimarisinin eğitim sürecini gerçekleştirmek amacıyla BSD300, T91 ve General100 veri setlerinin (Dong vd., 2016) birleştirilmesiyle yeni bir veri seti oluşturulmuştur. Böylece farklı görüntü boyutu ve pozlama özelliklerinden faydalanılmıştır. Eğitim veri seti 391 adet görüntüden meydana gelmektedir. Bu veri setine ait örnek görüntüler Şekil 1'de paylaşılmıştır. Derin ağ test sürecinde ise Set5, Set14 ve BSD100 veri setleri (Ledig vd., 2017) ayrı ayrı olacak şekilde kullanılmıştır. Böylece eğitilen derin ağ 119 görüntü üzerinde test edilmiştir.

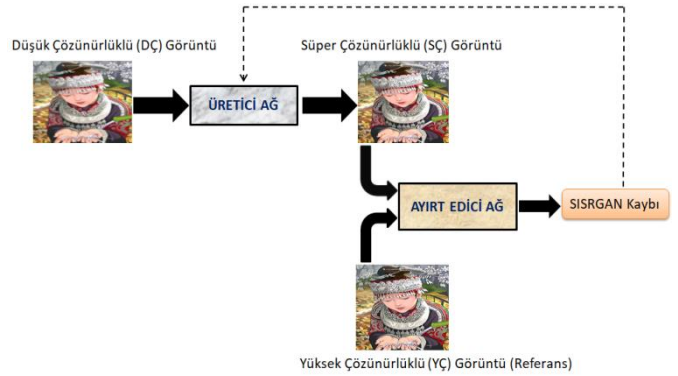


Şekil 1. Eğitim veri setinden örnek görüntüler (Sample images from the training dataset)

2.2. Derin ağ mimarisi (Deep network architecture)

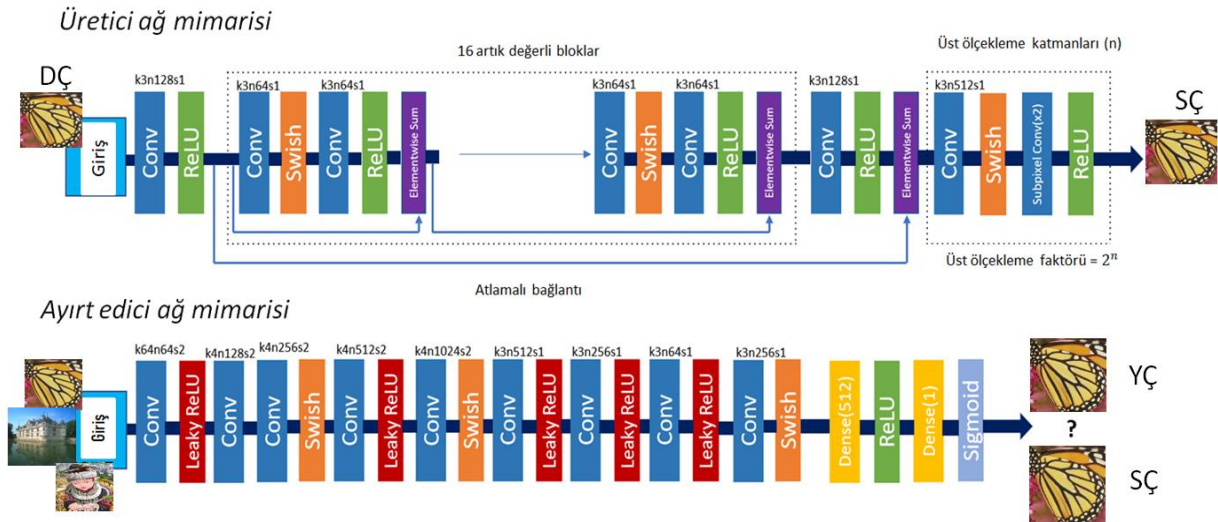
Bu bölüm, süper çözünürlük çalışmalarının başarısını artırmak ve daha hızlı bir eğitim süreci

yerine getirebilmek amacıyla geliştirilen SISRGAN (Single Image Super Resolution Generative Adversarial Networks) modeli ile ilgili detaylı bilgiler vermektedir. Önerilen bu model çekişmeli üretici ağlar (GAN) tabanlı bir model olup, üretici (generator) ve ayırt edici (discriminator) ağ kısımlarından meydana gelmektedir (Radford vd., 2015). Üretici ağ, x4 üst ölçekleme yaparak süper çözünürlüklü görüntüyü (SÇ) meydana getirmektedir. Modelin çalışma prensibine göre ağ kaybindan yararlanarak geri yayılım (backpropagation) tekniği uygulanmaktadır (Senalp ve Ceylan, 2020). SISRGAN mimarisinin genel gösterimi Şekil 2' de verilmiştir.



Şekil 2. SISRGAN mimarisi (SISRGAN architecture)

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.



Şekil 3. Üretici ve ayırt edici ağ mimarileri (Generator and discriminator networks)

Matematiksel İfade:

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

Denklem 1’de verilen formülde δ_G , derin ağ modelinin ağırlık ve sapma değerlerini ifade ederken G ise SÇ görüntüyü temsil etmektedir. Ayrıca, I_n^{DC} ve I_n^{YC} 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 (Johnson vd., 2016).

$$L_{\text{toplam}(SISRGAN)} = L_{mse} + \underbrace{L_{vgg}}_{L_{icerik}} + L_{cekişmeli} \quad (2)$$

Ortalama Kare Hata (MSE), derin ağ modellerinin başarısını gösteren pozitif değerli bir ölçüttür. MSE değerinin azalması modelin başarısının arttığını göstermektedir. 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. $G_{\phi}(I^{DC})$, SÇ görüntüyü, I^{YC} 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 (Gu vd., 2020).

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

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

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

2.3. Görüntü kalite metrikleri (Image quality metrics)

Tepe sinyal gürültü oranı (PSNR), iki sinyal arasındaki 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 (Javaid vd., 2013).

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

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, $\mu_{I_R}(n)$ ve $\mu_{I_D}(n)$ sembolleri, referans (I_R) ve bozuk (I_D) görüntü dizisinin n. çerçevesinin ortalamasını temsil etmektedir. Ayrıca, $\sigma_{I_R}(n)$ ve $\sigma_{I_D}(n)$, referans (I_R) ve bozuk (I_D) görüntü dizisinin n. çerçevesinin standart sapmasını temsil etmektedir. C1, C2 ise kararsızlıktan kaçınmak için kullanılan sabitlerdir (Dosovitskiy ve Brox, 2016).

$$SSIM = \frac{[2\mu_{I_R}(n)\mu_{I_D}(n)+C1][2\sigma_{I_R}I_D(n)+C2]}{[\mu_{I_R}^2(n)+\mu_{I_D}^2(n)+C1][\sigma_{I_R}^2(n)+\sigma_{I_D}^2(n)+C2]} \quad (7)$$

3. Süper Çözünürlük Uygulamalarının Sonuçları (The Results of Super Resolution Implementations)

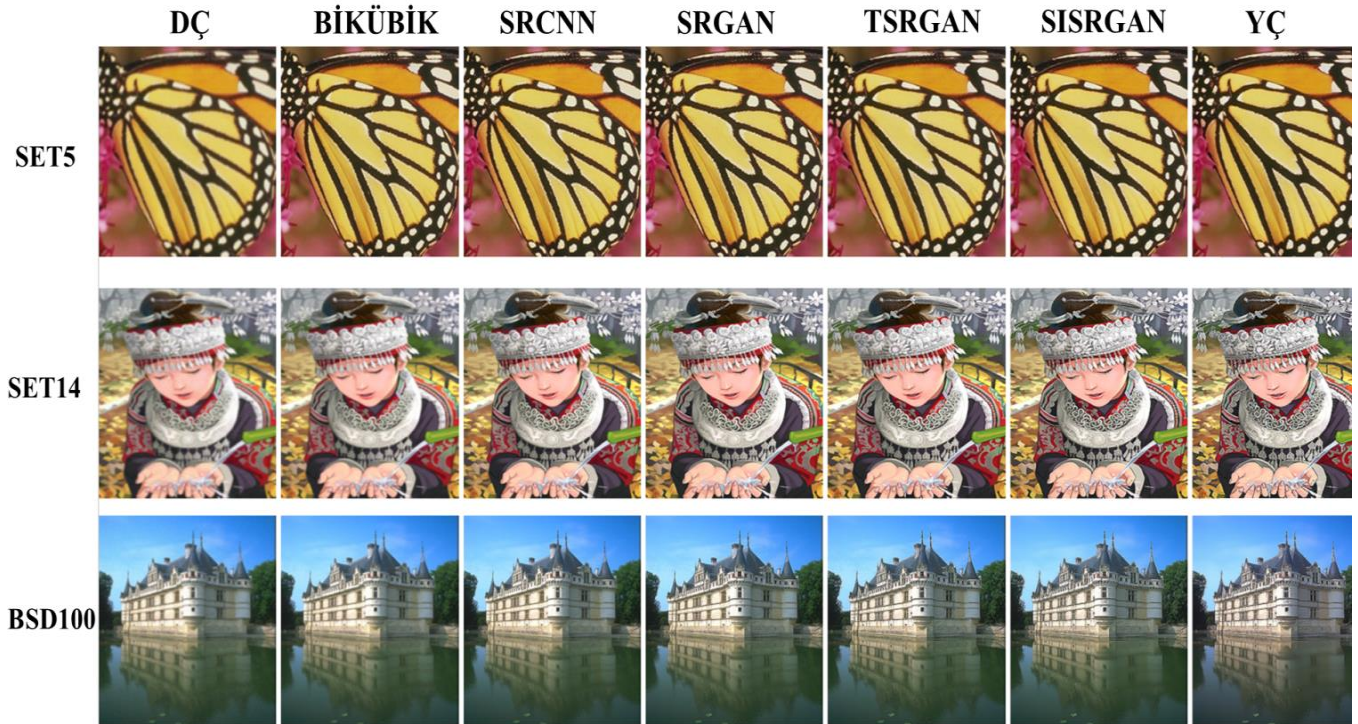
Bu bölümde, gerçekleştirilen uygulamaların eğitim süreçleri ve uygulamalar sonucunda elde edilen sonuçlar hakkında detaylı bilgiler yer almaktadır.

3.1. Eğitim sürecinin detayları (Training details)

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, ağın eğitimi 14 saat sürmüştür. Ek olarak, TSRGAN ağının eğitimi 18 saat sürerken, SRGAN ağının eğitimi ise 19 saatte tamamlanmıştır. Ayrıca, önerilen modelde optimizasyon algoritması olarak Adam algoritması kullanılmış ve öğrenme oranı 0.001 tercih edilmiştir.

3.2. Elde edilen sonuçlar (The results obtained)

Bu bölümde eğitilmiş derin ağ modellerinin Set5, Set14 ve BSD100 test veri setleri üzerinde elde edilen sonuçlar hem görsel olarak hem de metrik değerleri olarak sunulmaktadır. Üç test verisine ait sonuçlar ayrı ayrı paylaşılmıştır. Burada, önerilen SISRGAN derin ağ modelinin sonuçları literatürde yer alan yeni ve gelişmiş modellerle (SRCNN, SRGAN, TSRGAN) karşılaştırılmıştır. Modellerin test işlemleri sonucu elde edilen süper çözünürlüklü sonuç görüntülerinden örnekler Şekil 4'te görülmektedir. Bu görüntülerin yanı sıra düşük çözünürlüklü ve yüksek çözünürlüklü (referans) görüntülere ait örnekler de paylaşılmıştır.



Şekil 4. Elde edilen görüntülerden örnekler (Examples of the obtained images)

Daha sonra, farklı test veri setlerindeki görüntülerden oluşturulmuş süper çözünürlüklü görüntülerin başarı performansları PSNR ve SSIM görüntü kalite metrikleri dikkate alınarak değerlendirilmiştir. Elde edilen görüntülerden hesaplanan metrik değerlerinin sonuçları Tablo 1’de verilmiştir. Elde edilen en iyi sonuçlar kalın biçimde, ikinci en iyi değerler ise altı çizili olarak gösterilmiştir. PSNR ve SSIM sonuçları incelendiğinde önerilen modelin SRCNN, SRGAN ve TSRGAN modellerine göre daha üstün bir başarı elde ettiği görülmektedir. PSNR değerleri dikkate alındığında, önerilen SISRGAN modeli TSRGAN modeline göre 0.5 dB, SRGAN modeline göre 0.8-1.3 dB aralığında ve SRCNN modeline göre ise 0.2-0.6 dB aralığında daha

başarılı sonuçlara ulaştığı gözlemlenmiştir. Ayrıca SSIM değerleri incelendiğinde, SISRGAN modelinin TSRGAN ve SRGAN modelleri ile karşılaştırıldığında 2-6% aralığında daha başarılı olduğu görülmüştür. Bunun yanı sıra evrişimli sinir ağıları tabanlı SRCNN modeli ile karşılaştırıldığında ise yine daha başarılı olmasına rağmen daha yakın değerler elde edilmiştir. BSD100 veri seti üzerinde yürütülen test çalışmaları sonucunda SRCNN modeli en iyi SSIM değerini elde ederken ikinci en iyi değeri önerilen SISRGAN modeli elde etmiştir. Bunun dışında tüm test çalışmalarının neticesinde önerilen SISRGAN modeli en iyi sonuçlara ulaşmayı başarmıştır. TSRGAN modeli ise genel sonuçlarda en iyi ikinci değerlere ulaştığı gözlenmiştir.

Tablo 1. PSNR ve SSIM sonuçları (PSNR and SSIM results)

	DÇ	Biküçük İnterpolasyon	SRCNN	SRGAN	TSRGAN	SISRGAN	YÇ (Referans)
Set 5							
PSNR	25.256	27.372	28.549	28.328	<u>28.785</u>	29.216	∞
SSIM	0,7742	0,8211	0,8522	0,8492	<u>0,8754</u>	0,9065	1
Set 14							
PSNR	22.981	25.140	27.289	26.654	27.213	27.487	∞
SSIM	0,6685	0,7366	<u>0,7914</u>	0,7841	0,7879	0,8092	1
BSD100							
PSNR	24.162	25.583	26.251	25.578	<u>26.492</u>	26.817	∞
SSIM	0,6514	0,7052	0,7486	0,6993	0,737	<u>0,7461</u>	1

4. Sonuçlar (Conclusions)

Bu makalede görünür görüntüler için derin öğrenme tabanlı yeni bir süper çözünürlük modeli önerilmiştir. Bu model literatürde önemli yere sahip modellerle karşılaştırılmış ve hem görsel algı kalitesi hem de görüntü kalite metrikleri açısından daha başarılı sonuçlar elde ettiği görülmüştür. Bu çalışmalar sırasında üç farklı eğitim veri seti birleştirilerek yeni bir eğitim veri seti oluşturulmuştur. Önerilen SISRGAN modeli ve literatürdeki güncel başarıya sahip modeller için gerçekleştirilen test aşamaları sırasında literatürde yaygın bir şekilde kullanılan üç farklı test veri setinden faydalanılmıştır. Böylece farklı görüntüler üzerinde modelin başarısı ispatlanmıştır. Önerilen modelin başarısındaki en önemli etmenlerden birisi de ağı eğitimini sırasında farklı görüntülerden yararlanılmasıdır.

Ayrıca, önerilen modelde yığın normalleştirme katmanları çıkarılıp, konvolüsyon katmanlarının

artırılması sonucu modelin başarısı daha net bir şekilde ortaya konulmuştur. Burada, bağlantı atlama tekniğinin kullanılmasıyla beraber ağı eğitim hızı bakımından da daha verimli bir model elde edilmiştir. Ek olarak, önerilen model çekişmeli üretici ağılar tabanlı bir model olmasına rağmen tepe sinyal gürültü oranı (PSNR) ve yapısal benzerlik indeksi (SSIM) metrikleri açısından da evrişimli sinir ağıları tabanlı SRCNN modeline göre üstün bir başarı göstermiştir. Bunun önemli olmasının sebebi ise evrişimli sinir ağıları tabanlı modellerin literatürde daha çok PSNR-odaklı modeller olarak ifade edilmesidir. Buradaki amaç görsel kaliteden ziyade metrik değerlerinde başarı elde etmektir. Buna karşın bu makalede gerçekleştirilen çalışmalarda amaç görsel olarak daha kaliteli süper çözünürlüklü görüntüler oluşturmak ve bunu metrik değerleri ile desteklemektir. Gerçekleştirilen süper çözünürlük uygulamaları neticesinde, kenar detay bilgilerinin

(yüksek frekans bilgileri) başarılı şekilde elde edildiği ve bulanıklığın giderilerek daha net süper çözünürlüklü görüntülerin üretildiği görülmüştür. Bu durum hesaplanan görüntü kalite metriklerinin değerleri de incelendiğinde açık bir şekilde ortaya konmuştur.

Kaynaklar (References)

- Anwar, S., Khan, S., Barnes, N., 2020. A deep journey into super-resolution: A Survey. *ACM Computing Surveys* 53:1-34. <http://dx.doi.org/10.1145/3390462>
- Chudasama, V., Patel, H., Prajapati, K., Upla, K., Ramachandra, R., Raja, K., Busch, C., 2020. TherISuRNet- A computationally efficient thermal image super-resolution network. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. pp 388-397. <http://dx.doi.org/10.1109/CVPRW50498.2020.00051>
- Dong, C., Loy, C.C., He, K., Tan, X., 2015. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38:295-307. <http://dx.doi.org/10.1109/TPAMI.2015.2439281>
- Dong, C., Loy, C.C., Tang, X., 2016. Accelerating the Super-Resolution Convolutional Neural Network. *Computer Vision ECCV 2016 Lecture Notes in Computer Science Springer* 9906. https://doi.org/10.1007/978-3-319-46475-6_25
- Dosovitskiy, A., Brox, T., 2016. Generating images with perceptual similarity metrics based on deep networks. In *Advances in Neural Information Processing Systems (NIPS)*. pp 658-666. <https://doi.org/10.48550/arXiv.1602.02644>
- Glasner, D., Bagon, S., Irani, M., 2009. Super-resolution from a single image. *IEEE 12th International Conference on Computer Vision*. pp 349-356. <http://dx.doi.org/10.1109/ICCV.2009.5459271>
- Goodfellow, I., vd., 2014. Generative adversarial networks. In *Advances in Neural Information Processing Systems (NIPS)*. pp 2672-2680. <http://dx.doi.org/10.1145/3422622>
- Gu, Y., vd., 2020. MedSRGAN: medical images super-resolution using generative adversarial networks. *Multimedia Tools and Applications* 79:21815-21840. <http://dx.doi.org/10.1007/s11042-020-08980-w>
- Javaid, H., Babar, T.K., Rasool, A., Saghir, R.U., 2013. Video colour variation detection and motion magnification to observe subtle changes. M.Sc. Thesis, Blekinge Institute of Technology, Faisalabad, Pakistan.
- Johnson, J., Alahi, A., Li, F., 2016. Perceptual losses for real-time style transfer and super resolution. In *European Conference on Computer Vision (ECCV) Springer*. pp 694-711. http://dx.doi.org/10.1007/978-3-319-46475-6_43
- Kim, J., Lee, J.K., Lee, K.M., 2016. Accurate image super-resolution using very deep convolutional networks. *IEEE CVPR*. pp 1646-1654. <http://dx.doi.org/10.1109/CVPR.2016.182>
- Ledig, C., vd., 2017. 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. <http://dx.doi.org/10.1109/CVPR.2017.19>
- Radford, A., Metz, L., Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.
- Senalp, F.M., Ceylan, M., 2020. Enhancement of low resolution thermal face image resolution using deep learning. *European Journal of Science & Tech*. pp 131-135. <https://doi.org/10.31590/ejosat.802174>
- Senalp, F.M., Ceylan, M., 2021. Deep learning based super resolution and classification applications for neonatal thermal images. *Traitement du Signal* 38(5):1361-1368. <https://doi.org/10.18280/ts.380511>
- Toyran, M., 2008. Reconstructing super resolution images from low resolution images. M.Sc. Thesis, Institute of science, Istanbul.
- Wang, M., Chen, Z., Wu, Q.M.J., Jian, M., 2020. Improved face super-resolution generative adversarial networks. *Machine Vision and Applications* 31:22. <https://doi.org/10.1007/s00138-020-01073-6>
- Yan, R., Yang, K., Wang, K., 2021. 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. <http://dx.doi.org/10.1109/ROBIO54168.2021.9739390>
- Yue, L., Shen, H., Li, J., Yuan, Q., Zhang, H., Zhang, L., 2018. Image super-resolution: the techniques, applications, and future. *Signal Processing* 128:389-408. <http://dx.doi.org/10.1016/j.sigpro.2016.05.002>
- Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., Fu, Y., 2018. Image super-resolution using very deep residual channel attention networks. *Proceedings of the European Conference on Computer Vision*. pp 286-301. http://dx.doi.org/10.1007/978-3-030-01234-2_18