

Çiçek Görüntülerinde Derin Öğrenmeyle Süper Çözünürlük: SRCNN ve SRGAN Modellerinin Karşılaştırılması

Super Resolution with Deep Learning on Flower Images: Comparison of SRCNN and SRGAN Models

Ahmet ÇINAR ^{*,1}, Elif Işılay ÜNLÜ²

*: acinar@firat.edu.tr, ORCID: 0000-0001-5528-2226

¹: Bilgisayar Mühendisliği, Fırat Üniversitesi, TÜRKİYE

²: Bilgisayar Mühendisliği, Fırat Üniversitesi, TÜRKİYE

Özet: Günümüzde, görüntü işleme ve derin öğrenme teknikleri, düşük çözünürlüklü görüntülerin yüksek çözünürlüklü hale getirilmesi amacıyla yaygın olarak kullanılmaktadır. Süper çözünürlük, birçok görüntü işleme algoritması için daha fazla veri zenginliği sağlar. Bu da öğrenme algoritmalarının ve yapay zeka modellerinin daha iyi performans göstermesine olanak tanır. Bu nedenlerle, süper çözünürlük, birçok uygulama alanında görüntü kalitesini artırmak ve analiz yeteneklerini geliştirmek için önemli bir araç olarak kabul edilir. Bu bağlamda, bu çalışmada çiçek görüntüleri üzerinde süper çözünürlük yöntemleri uygulanarak, estetik değeri yüksek ve detayları zengin görüntüler elde etme potansiyelini artırmak amaçlanmıştır. Çiçek görüntülerine uygulanan SRCNN ve SRGAN gibi farklı süper çözünürlük yöntemleri arasındaki karşılaştırmalar ve avantajlar da ele alınmıştır. Bu yöntemler arasında SRGAN (Super-Resolution Generative Adversarial Network) gibi GAN tabanlı modellerin, doğal görüntü üretimindeki yetenekleri sayesinde öne çıktığı görülmüştür.

Anahtar Kelimeler: Süper Çözünürlük, SRCNN, SRGAN

Abstract: Nowadays, image processing and deep learning techniques are widely used to convert low-resolution images to high-resolution. Super-resolution provides more data richness for many image processing algorithms. This allows learning algorithms and artificial intelligence models to perform better. For these reasons, super resolution is considered an important tool to improve image quality and enhance analysis capabilities in many application areas. In this context, this study aims to increase the potential of obtaining images with high aesthetic value and rich details by applying super-resolution methods on flower images. Comparisons and advantages between different super-resolution methods such as SRCNN and SRGAN applied to floral images are also discussed. Among these methods, GAN-based models such as SRGAN (Super-Resolution Generative Adversarial Network) were found to be prominent due to their ability to produce natural images.

Keywords: Super Resolution, SRCNN, SRGAN

I. GİRİŞ

Günümüzde, görsel içeriklerin yaygın olarak paylaşıldığı ve kullanıldığı bir çağda, görüntü kalitesi önemli bir etken haline gelmiştir. Ancak, sınırlı donanım çözünürlüğü veya düşük kaliteli kaynaklardan elde edilen görüntüler, görsel deneyimi sınırlandırabilir. İşte bu noktada,

süper çözünürlük yöntemleri devreye girerek, düşük çözünürlüklü görüntüleri daha yüksek kaliteli ve ayrıntılı hale getirmenin önemli bir aracı olmuştur. Görüntü süper-çözünürlüğü (SR), bilgisayarla görme alanında bir görüntünün çözünürlüğünü iyileştiren hayati görüntü işleme yöntemlerinden biridir[1]. Süper çözünürlük, bir görüntünün veya video görüntüsünün çözünürlüğünün artırılması için kullanılan bir teknik olarak tanımlanabilir. Bu teknik, bir dizi düşük çözünürlüklü görüntüyü birleştirerek yüksek çözünürlüklü bir görüntü oluşturur. Süper çözünürlük algoritmaları, yüksek çözünürlüklü bir görüntüyü elde etmek için düşük çözünürlüklü görüntülerin piksel yoğunluğunu ve benzerliklerini analiz eder. Bu analiz sonucunda, yüksek çözünürlüklü bir görüntü elde etmek için kayıp bilgi ve ayrıntılar yeniden oluşturulabilir.

Çok sayıda elektronik görüntü uygulamasında, yüksek çözünürlüklü görüntüler daha yüksek piksel yoğunluğuna sahiptir, bu nedenle birçok pratik uygulamada çok yardımcı olan daha fazla görüntü ayrıntısı sağlayabilirler. Örneğin, yüksek çözünürlüklü tıbbi görüntüler, doktorların hastalara daha doğru teşhis koymasına yardımcı olabilir; Dünya veya Ay gibi uydu görüntülerinde, yüksek çözünürlüklü görüntüler daha karmaşık coğrafi ayrıntılar elde edebilir; yol planlama ve otomatik pilota süper çözünürlüklü görüntüler vazgeçilmezdir. Süper çözünürlüklü görüntü, avantajlarını diğer birçok yerde de tam olarak gösterebilir [2].

Geçmişte süper çözünürlük elde etmek için istatistiksel yöntemler, tahmine dayalı yöntemler, blok tabanlı yöntemler, kenar tabanlı ve seyrek gösterim yöntemleri gibi klasik SR yöntemleri kullanıldı. Ancak bilgi işlem gücü ve büyük verilerdeki son gelişmeler, araştırmacıların SR sorunlarını çözmek için derin öğrenmeyi (DL) kullanmalarına olanak tanıdı. Son yıllarda derin öğrenme tekniklerinin hızla gelişmesiyle, derin öğrenmeye dayalı SR modelleri aktif olarak keşfedildi ve genellikle SR'nin çeşitli ölçütlerinde en son teknoloji performansı elde etti [3].

Konvolüsyonel Sinir Ağları (CNN) tabanlı yöntemlerden (örneğin, SRCNN [4] Generative Adversarial Nets (GAN) kullanan son umut verici SR yaklaşımlarına kadar değişen, SR görevlerinin üstesinden gelmek için çeşitli derin öğrenme yöntemleri uygulanmıştır (örn. SRGAN [5]).

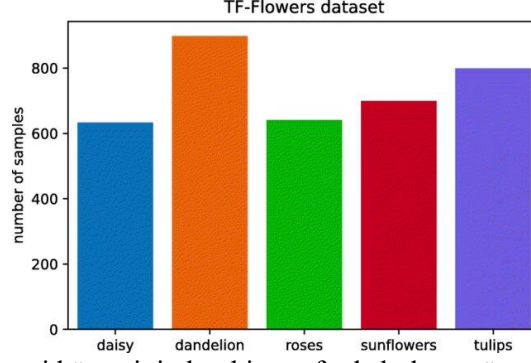
Bu çalışmada çiçek görüntülerine uygulanan süper çözünürlük yöntemleri, doğanın renk paletini ve detaylarını yakalamayı amaçlayarak görüntüleme alanında çeşitli uygulama potansiyelleri sunmaktadır. Çiçekler, doğanın estetik zenginliğini temsil eden eşsiz görseller ve renkler sunar. Ancak, çeşitli faktörlere bağlı olarak elde edilen düşük çözünürlüklü çiçek görüntüleri, bu güzellikleri tam olarak ifade etmekte yetersiz kalabilir. Bu noktada, süper çözünürlük yöntemleri, bilgisayarlı görü algoritmaları ve derin öğrenme tekniklerini kullanarak, düşük çözünürlüklü çiçek görüntülerini daha yüksek çözünürlüklü, canlı ve detaylı hale getirme imkanı sağlanır. Çiçek görüntülerinde süper çözünürlük uygulamaları incelenerek, bu yöntemlerin teknik özellikleri, avantajları ve elde edilen sonuçlar ele alınacaktır.

MATERYAL VE YÖNTEM

A. Veri Seti

Bu çalışmada veri seti olarak 3670 görüntüden oluşan Tensorflow çiçek veri seti kullanılmıştır. Tensorflow çiçek veri kümesi, çiçek görüntülerinden oluşan büyük bir veri kümesidir. Çiçekler veri seti, 5 olası sınıf etiketine sahip çiçek görüntülerinden oluşur. Farklı boyutlarda ve en boy oranlarında çiçeklerin yüksek kaliteli görüntülerini içerir. Tüm görüntüler özellikle derin öğrenme iş yükleri için toplanmıştır. Bu veri seti yalnızca bir eğitim setine bölünmüştür. Tf_flowers veri setini yüklemek için tensorflow_datasets modülü kullanılmış ve ilk 600 görüntü doğrulama verisi olarak alınmıştır. Sınıf 0 (Papatya), Sınıf 1 (Karahindiba), Sınıf 2 (Laleler), Sınıf 3 (Güller) ve Sınıf 4 (Ayciçekleri), sınıf dağılımları Şekil 1'de

gösterilmiştir. Veri setindeki beş sınıfa ait görüntü örnekleri Şekil 2’de verilmiştir.



Şekil 1. TF-Flowers veri kümesinin her bir sınıfında bulunan örnek sayısının dağılımı

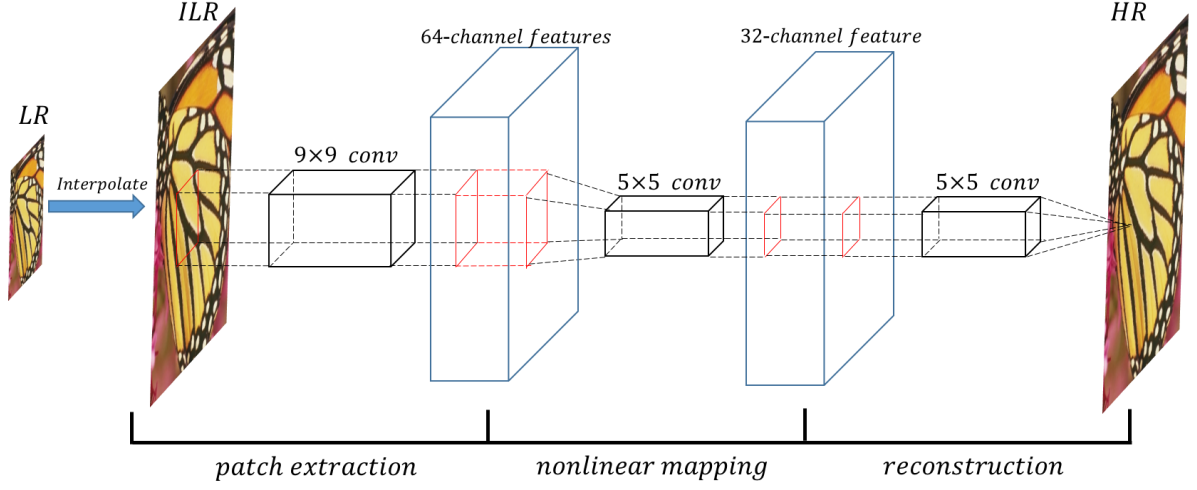


Şekil 2. TF-Flowers Veri Seti

B. SRCNN Mimarisi

SRCNN, bir kilometre taşı olarak adlandırılabilir sinir ağı modelini kullanan ilk süper çözünürlüklü teknolojidir. SRCNN, süper çözünürlük alanında önemli bir dönemeç oluşturmuş ve daha sonraki modellerin geliştirilmesine ilham vermiştir. SRCNN fikri, seyrek kodlama tabanlı süper çözünürlük yöntemlerinden esinlenmiştir SRCNN, her katmanın filtre boyutlarının $64 \times 1 \times 9 \times 9$, $32 \times 64 \times 5 \times 5$ ve $1 \times 32 \times 5 \times 5$ olduğu üç katmanlı bir CNN'dir. Önerilen mimari kavramsal olarak 3 bileşenden oluşmaktadır: Özellik çıkarıcı, doğrusal olmayan haritalama, yeniden yapılandırma. Her biri düşük çözünürlüklü özelliklerin çıkarılmasından, yüksek çözünürlüklü özelliklere eşlemeden ve yeniden yapılanmadan sorumludur. Yama çıkarma katmanı, girişten yoğun yamaları çıkarmak ve bunları evrişimli filtreler kullanarak temsil etmek için kullanılır. Doğrusal olmayan haritalama katmanı, kanal sayısını değiştirmek ve doğrusal olmama eklemek için kullanılan 1×1 evrişimli filtrelerden oluşur. Son yeniden yapılandırma katmanı, yüksek çözünürlüklü görüntüyü yeniden oluşturur. Ağı eğitmek için MSE kaybı işlevi kullanılır ve sonuçları değerlendirmek için PSNR kullanılır [4]. SRCNN, SR gelişiminde bir dönüm noktasıdır, ancak yine de genellikle sığ yapısı nedeniyle suçlanmaktadır. SRCNN'in giriş görüntüsünden hedef yüksek çözünürlüklü görüntüye kadar süper çözünürlük problemini tek bir end-to-end modelde çözmesi ve konvolüsyon katmanları, öğrenilebilir filtreleri kullanarak düşük çözünürlüklü görüntülerdeki özellikleri daha iyi temsil etmeye çalışması avatanjlarının yanı sıra eksiklikleri de vardır. Daha ince detayları ve düzensiz yapıları ele almakta bazı zorlukları olabilir. Özellikle çok karmaşık ve detaylı görüntülerde, modelin

performansı düşebilir. SRCNN'nin performansı, eğitildiği veri setine bağlıdır. Eğer model, geniş bir veri setiyle eğitilmediyse, farklı türdeki görüntülerde düşük başarı elde edebilir. Bazı durumlarda, SRCNN ve benzeri süper çözünürlük modelleri, yüksek çözünürlükteki görüntülerde istenmeyen sanal kenarlar veya artefaktlar üretebilir. Ayrıca, aşırı öğrenme (overfitting) riskiyle başa çıkmak ve dengelemek zor olabilir. SRCNN mimarisi Şekil 3'de verilmiştir.



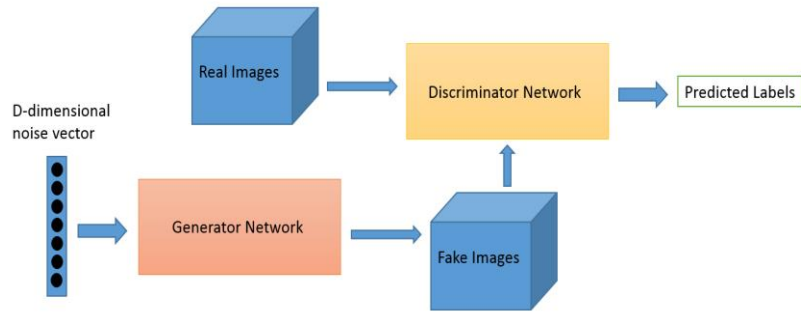
Şekil 3. SRCNN Mimarisi[6]

C. SRGAN Mimarisi

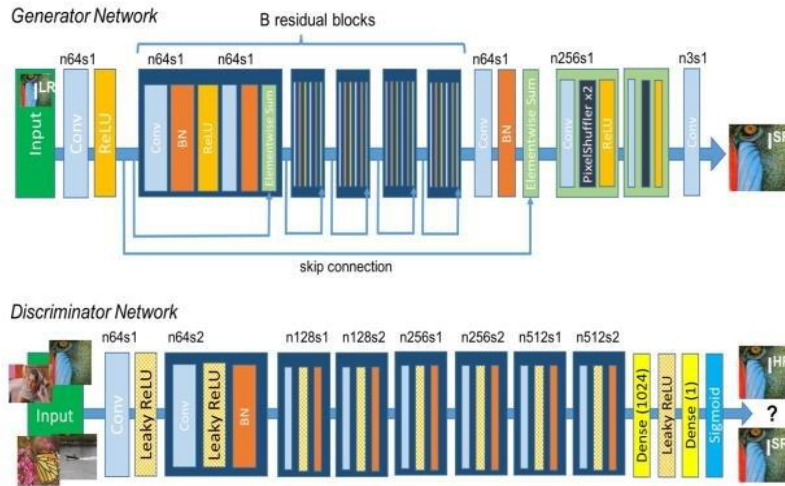
İlk GAN tabanlı SR tekniği SRGAN'dır. SRGAN, geleneksel süper çözünürlük yöntemleriyle karşılaştırıldığında GAN'ları kullanarak daha iyi sonuçlar elde etmektedir. Bu teknik, fotogerçekçi sonuçlar sağlamak için Generative Adversarial Network (GAN) yapısını ve süper çözünürlük elde etmek için gelişmiş bir yapay sinir ağı tasarımını kullanır. Üreteç ve ayırtıcı bu yöntemin iki ana parçasıdır. Üretici, düşük çözünürlüklü görüntüleri yüksek çözünürlüklü görüntülere dönüştürmek için kullanılan bir öğrenme modelidir. Ayırtıcı, üretici tarafından oluşturulan görüntüleri gerçek yüksek çözünürlüklü görüntülerden ayırmaya çalışır[5]. GAN mimarisi şekil 4'teki gibidir.

SRGAN, süper çözünürlük elde etmek için bu iki unsur arasındaki çatışmadan yararlanır. Ayırtıcı, daha doğru ayırmalar yapmak amacıyla üreticinin çıktısını gerçek yüksek çözünürlüklü görüntülerle karşılaştırırken, üretici de ayırtıcı tarafından üretilen yüksek çözünürlüklü görüntülerle karşılaştırılabilir görüntüler üretmeye çalışır.

SRGAN'larda öne çıkan faktör algısal kayıp fonksiyonudur. Üretici ve ayırtıcı GAN mimarisine dayalı olarak eğitilirken, SRGAN'lar hedeflerine ulaşmak için başka bir kayıp fonksiyonunun yardımını kullanır: algısal/içerik kaybı fonksiyonu. SRGAN mimarisi şekil 5'te verilmiştir.



Şekil 4. GAN Mimarisi



Şekil 5. SRGAN Mimarisi[5]

D. Eğitim

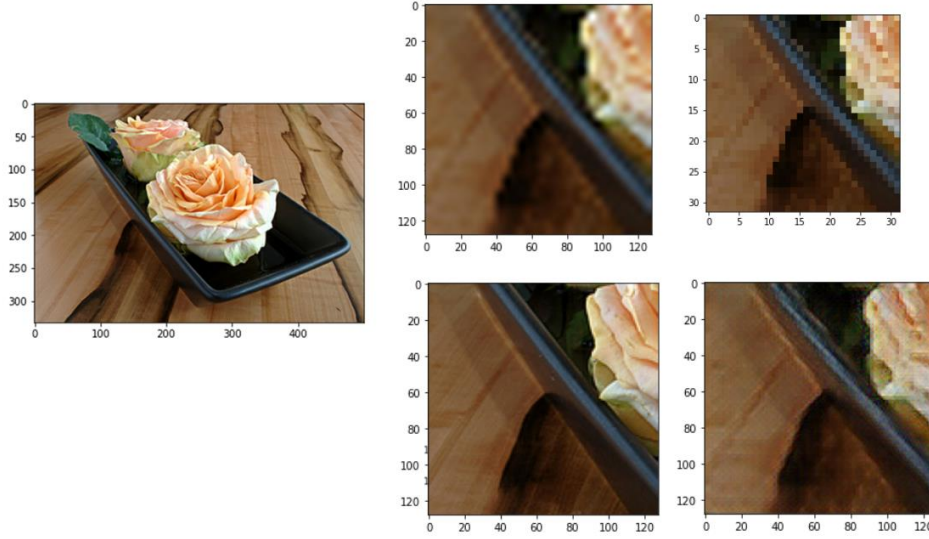
Test verileri veri kümesindeki ilk 600 görüntü olarak bölünür ve verilen görüntüyü Şekil 6'daki gibi 128x128 boyutunda rastgele kırpma ve görüntünün düşük çözünürlüklü ve yüksek çözünürlüklü kopyasını döndürmek için build_data işlevi tanımlanır. Düşük çözünürlüklü kopya biküzik enterpolasyon yoluyla oluşturulur. SRCNN modeli eğitim kaybı pixel_mse_loss olarak tanımlanır ve ayrıca modelin PSNR kaybını değerlendirmek için bir PSNR fonksiyonu tanımlanır. Model mimarisi, çekirdek boyutları 9-1-5 olan 3 katmanlı bir CNN'dir. Her epochta görüntüler yeniden kırılarak yeni görüntü örnekleri oluşturulur. Model, performansı artırmak için daha fazla iterasyon için eğitilebilir, ancak kayıp genellikle 0,0030'un altına düşmemiştir.



Şekil 6. Yeniden boyutlandırılan ve bikübik enterpolasyon uygulanan görüntü örnekleri

Ağdaki ağırlıklar yalnızca görüntü boyutuna bağlı olmayan evrişim filtreleri olduğundan, ağ, toplu eğitim için (128, 128)'e kırılan eğitim verileriyle aynı boyutta olmayan görüntüleri girebilir. Eğitilmiş SRCNN, bikübik enterpolasyonlu görüntüyle karşılaştırıldığında algısal olarak görüntü çözünürlüğünde radikal bir artış göstermez.

SRGAN model eğitimi için veri kümesindeki her görüntüyü (128, 128) ve (32, 32) düşük çözünürlüklü bir kopyasıyla eşlemek için bir fonksiyon tanımlanır (Şekil 7). Bu, her eğitim döneminden önce çalıştırılır. Artık üretici mimarisi Tensorflow kullanarak tanımlanır. Bir artık bloğun tamamını oluşturmak için fonksiyonlar tanımlanmış ve eleman bazında toplam atlama bağlantıları da uygulanmıştır. 5 artık blok bağlanır ve son görüntü Upsample_block işlevinde uygulanan piksel karıştırıcı yöntemiyle yukarı örneklenir. Ayırt edici, HR/SR görüntüleri arasında ayırım yapmak için eğitilir; GAN kaybı ayırt ediciye geri yayılır ve üreticiye de aktarılır. Ağ, görüntüyü giren ve görüntünün gerçekliğine karar veren geleneksel bir CNN'dir.



Şekil 7. Yeniden boyutlandırılan ve bikübik enterpolasyon uygulanan görüntü örnekleri

E. Değerlendirme Metrikleri

SRCNN ve SRGAN modellerinin performansı, yüksek çözünürlüklü (HR) ve SR (Super-Resolution) görüntülerinin kalitesini değerlendirmek için farklı metrikler kullanılarak değerlendirilmiştir. Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM) ve gözlem değerlendirmeleri gibi metrikler kullanılarak model başarıları ölçülmüştür.

Peak Signal-to-Noise Ratio (PSNR), görüntü kalitesini değerlendirmek için sıklıkla kullanılan bir metriktir. PSNR, mümkün olan maksimum sinyal gücü ile bozucu gürültünün gücü arasındaki orandır. İki görüntü arasındaki farkın MSE (Mean Squared Error - Ortalama Kare Hatası) ile ölçüldüğü bir yöntemdir. MSE, her pikselin orijinal ve yeniden yapılandırılmış görüntü arasındaki farkın karesinin ortalamasıdır. PSNR, MSE değerinin logaritması olarak hesaplanır ve genellikle desibel (dB) biriminde ifade edilir[3].

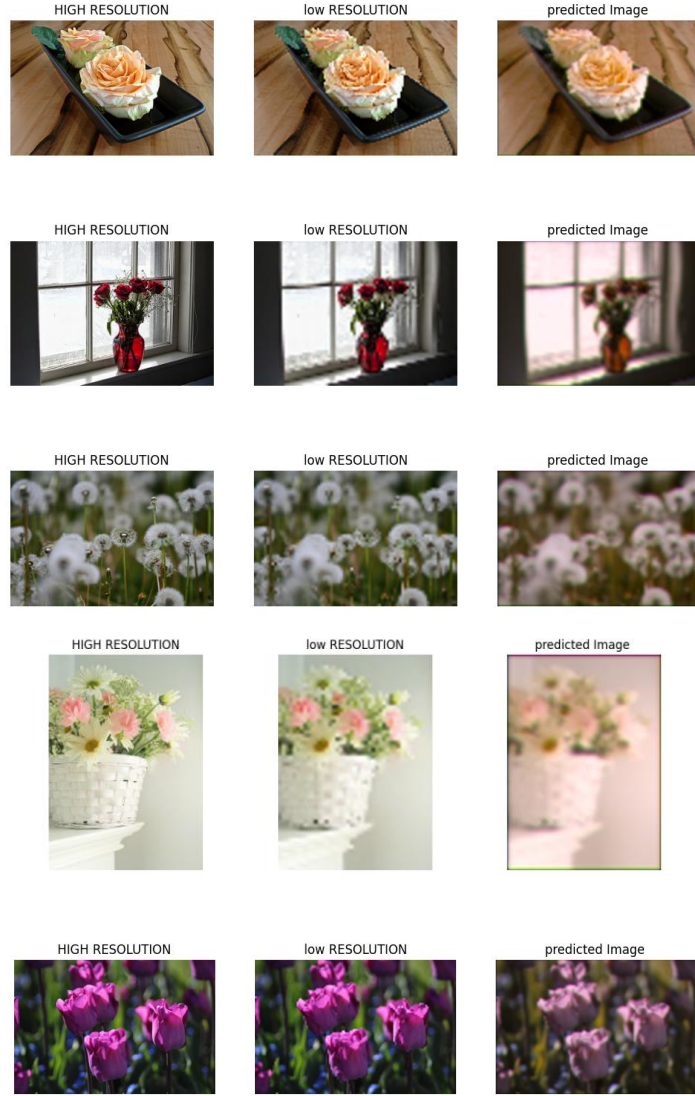
$$PSNR = 10 \log_{10} \frac{M^2}{MSE} \quad (1)$$

M: bir görüntü sinyalinin mümkün olan maksimum gücü

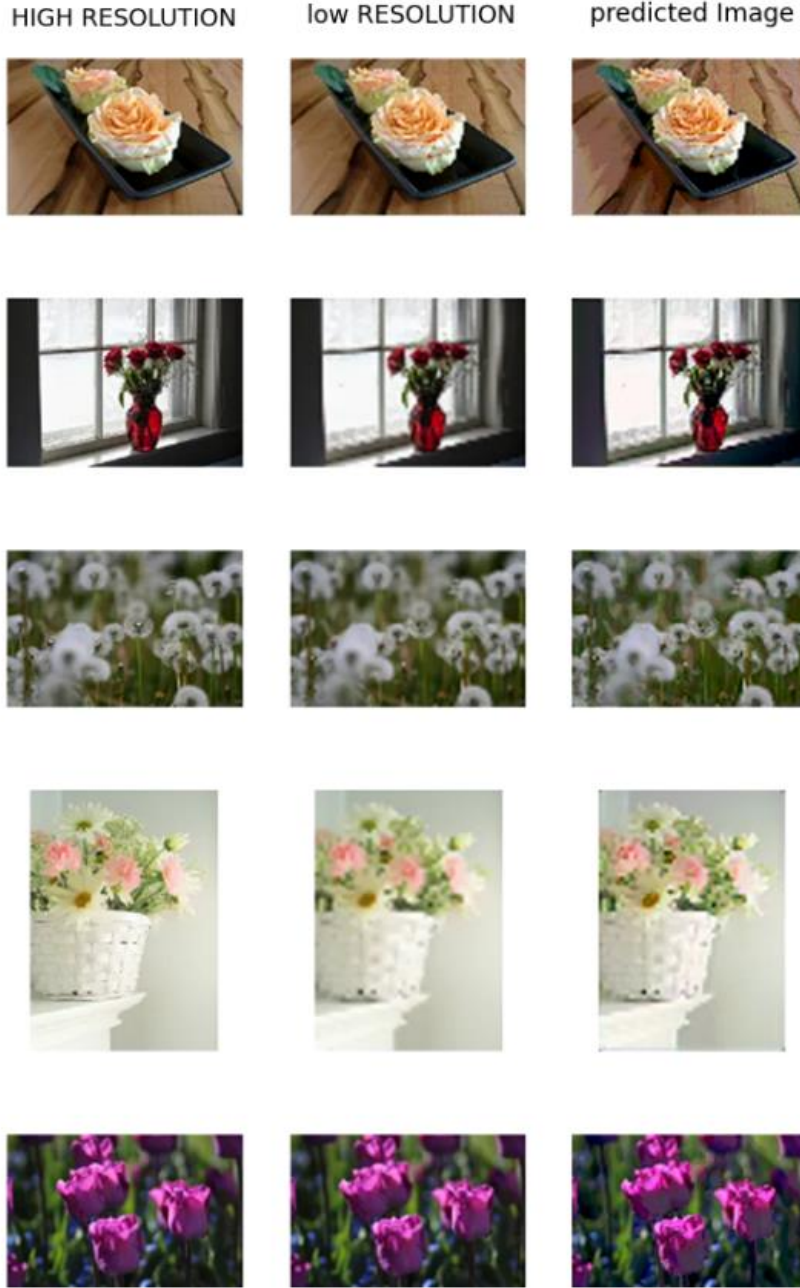
MSE: piksel piksel ortalama kare hatası

BULGULAR

SRCNN ve SRGAN modelleri, eğitim veri kümesi üzerinde eğitilmiş ve modelin performansı, yüksek çözünürlüklü (HR) ve SR (Super-Resolution) görüntülerinin kalitesini değerlendirmek için farklı metrikler kullanılarak değerlendirilmiştir. Her iki modelde de iterasyon sayısı arttıkça tahmin edilen çiçek görüntülerinin kalitesi artmıştır. SRCNN modelinin 1 iterasyon eğitim ve 100 iterasyon eğitim sonuçları şekil 8 ve şekil 9'da verilmiştir. İlk sütundaki görüntüler orijinal yüksek çözünürlüklü HR görüntüleri, ikinci sütundaki görüntüler düşük çözünürlüklü görüntüler, üçüncü sütundaki görüntüler ise tahmin edilen süper çözünürlüklü SR görüntüleridir.

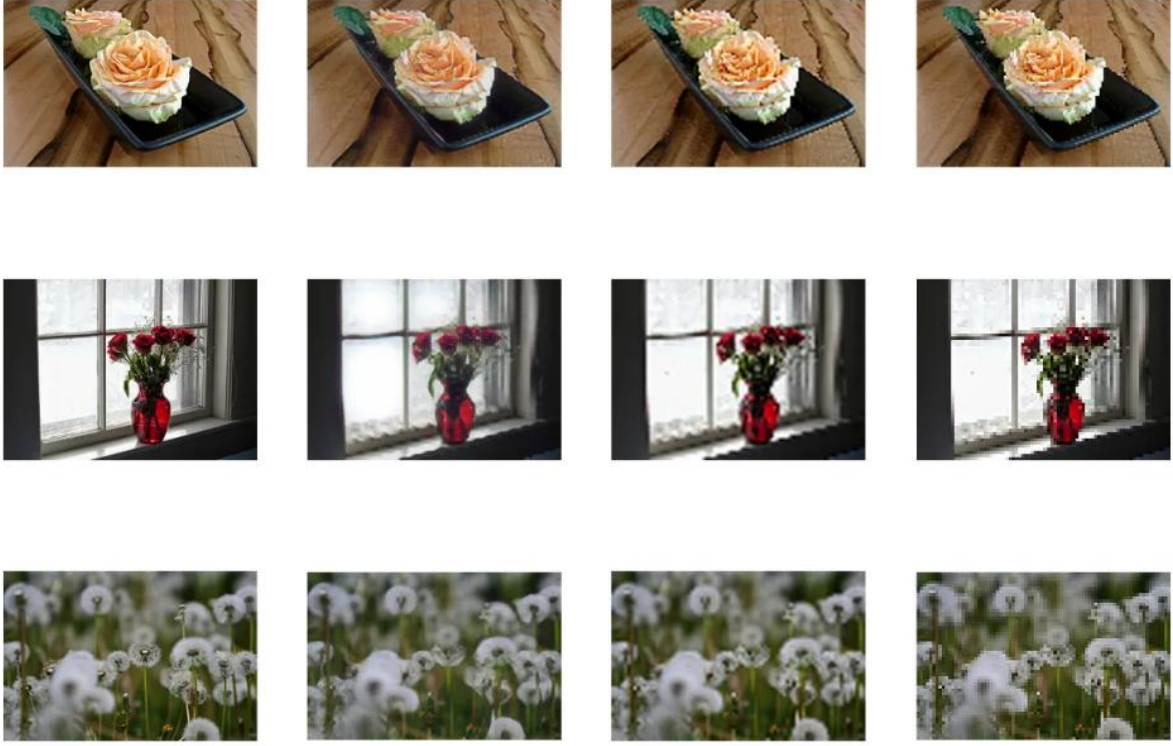


Şekil 8. SRCNN Sonuçları - 1 Epoch



Şekil 9. SRCNN Sonuçları- 100 Epoch

SRGAN modeli, düşük çözünürlüklü (LR) çiçek görüntülerini süper çözünürlüklü (SR) görüntülere dönüştürme yeteneğini başarılı bir şekilde öğrenmiştir. Şekil 8’de ilk sütundaki görüntüler orijinal HR görüntüsü, ikinci sütun görüntüler SR görüntüsü, üçüncü ve dördüncü sütun görüntüler ise düşük çözünürlüklü ve biküçük enterpolasyonlu görüntülerdir. Her model yeterli süre boyunca eğitilmemiş olsa da her modelin performansını karşılaştırdık. VGG ile eğitilmiş modellerle ve rakip kayıplarla oluşturulan görüntülerin kalitesi daha iyi görünüyor. Orijinal çiçek görüntülerine yakın örnekler elde edilmiştir.



Şekil 10. SRGAN Sonuçları

SONUÇ

Bu çalışmada, düşük çözünürlüklü görüntülerin yüksek çözünürlüklü hale getirilmesi amacıyla kullanılan iki önemli süper çözünürlük modeli olan SRCNN ve SRGAN'ı inceledik. Her iki model de farklı yaklaşımlar kullanarak süper çözünürlük elde etmeye odaklanmıştır. Bu çalışma, çiçek görüntülerine süper çözünürlük uygulamalarını inceleyerek, bu yöntemlerin performansını ve elde edilen sonuçları değerlendirmektedir. Derin öğrenme algoritmaları arasında öne çıkan Convolutional Neural Networks (CNN), çiçeklerin ince detaylarını yakalayarak yüksek çözünürlük elde etme sürecine katkı sağlamaktadır. Ayrıca, bu algoritmaların kullanıldığı örneklerle desteklenerek, renk doğruluğu ve görüntü netliği gibi kriterler üzerinden süper çözünürlük yöntemlerinin etkinliğini ele almaktadır. SRGAN modeli algısal olarak MSE tabanlı SRCNN modelinden daha iyi performans gösterse de, bu durum daha fazla eğitim tekrarı yoluyla çözülebilir. SRCNN'nin avantajları arasında end-to-end öğrenme, öğrenilebilir filtreler ve genel süper çözünürlük performansı bulunmaktadır. Ancak, modelin daha ince detayları ele almakta bazen zorlandığı ve eğitim süresi ile hesaplama gücü ihtiyacının yüksek olabileceği gözlemlenmiştir. SRGAN'ın avantajları arasında daha doğal görüntülerin üretilmesi ve insan gözünün algısal özelliklerini dikkate alan içerik kaybı bulunmaktadır. Ancak, modelin eğitim süresi ve hesaplama gücü ihtiyacı, bazen gerçek zamanlı uygulamalarda kullanımını sınırlayabilir.

Her iki model de süper çözünürlük alanında önemli gelişmelere işaret etmektedir. SRCNN, temel bir süper çözünürlük modeli olarak hizmet ederken, SRGAN daha doğal ve estetik açıdan tatmin edici görüntüler üretme konusunda öne çıkar. Ancak, her iki modelin de belirli durum ve gereksinimlere daha iyi uyduğunu söylemek mümkündür. Çiçek görüntüleri üzerinde yapılan bu incelemeler, süper çözünürlük modellerinin doğal güzellikleri daha iyi ifade etme ve estetik değeri artırma potansiyelini vurgulamaktadır. Farklı çiçek türleri, renk tonları ve yaprak dokuları, model performansını etkileyebilir. Gelecekteki çalışmalarda, çeşitli çiçek türlerine uyarlanmış modeller geliştirmek, genel performansı artırabilir. Modelin daha iyi

genelleme yapabilmesi için daha geniş bir çiçek veri seti kullanılabilir. Farklı ışık koşulları, çekim açıları ve çözünürlük düzeylerini içeren çeşitli veri setleri, modelin gerçek dünya senaryolarına daha iyi uyum sağlamasına yardımcı olabilir. Modellerin daha ince detayları ele alabilmesi için yeni öğrenme yöntemleri ve özellik çıkarma teknikleri geliştirilebilir.

KAYNAKLAR

- [1] Z. Wang, J. Chen, and S. C. H. Hoi, “Deep Learning for Image Super-Resolution: A Survey,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 10, pp. 3365–3387, 2021, doi: 10.1109/TPAMI.2020.2982166.
- [2] L. Zhou and S. Feng, “A Review of Deep Learning for Single Image Super-Resolution,” *ICIIBMS 2019 - 4th Int. Conf. Intell. Informatics Biomed. Sci.*, no. 41671402, pp. 139–142, 2019, doi: 10.1109/ICIIBMS46890.2019.8991477.
- [3] S. M. A. Bashir, Y. Wang, M. Khan, and Y. Niu, “A comprehensive review of deep learningbased single image super-resolution,” *PeerJ Comput. Sci.*, vol. 7, pp. 1–56, 2021, doi: 10.7717/PEERJ-CS.621.
- [4] C. Dong, C.C. Loy, K. He, X. Tang, “Learning a deep convolutional network for image super-resolution,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8692, no. September, pp. 184–199, 2014.
- [5] C. Ledig *et al.*, “Photo-realistic single image super-resolution using a generative adversarial network,” *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 105–114, 2017, doi: 10.1109/CVPR.2017.19.
- [6] W. Yang, X. Zhang, Y. Tian, W. Wang, J. H. Xue, and Q. Liao, “Deep Learning for Single Image Super-Resolution: A Brief Review,” *IEEE Trans. Multimed.*, vol. 21, no. 12, pp. 3106–3121, 2019, doi: 10.1109/TMM.2019.2919431.