



Düşük Çözünürlüklü Termal Yüz Görüntü Çözünürlüğünün Derin Öğrenme İle Artırılması

Fatih Mehmet Şenalp^{1*}, Murat Ceylan²

¹ Konya Teknik Üniversitesi, Mühendislik Fakültesi, Elektrik Elektronik Mühendisliği Bölümü, Konya, Türkiye (ORCID: 0000-0001-7831-6724)

² Konya Teknik Üniversitesi, Mühendislik Fakültesi, Elektrik Elektronik Mühendisliği Bölümü, Konya, Türkiye (ORCID: 0000-0001-6503-9668)

(1st International Conference on Computer, Electrical and Electronic Sciences ICCEES 2020 – 8-10 Ekim 2020)

(DOI: 10.31590/ejosat.802174)

ATIF/REFERENCE: Şenalp, F. M. & Ceylan, M. (2020). Düşük Çözünürlüklü Termal Yüz Görüntü Çözünürlüğünün Derin Öğrenme İle Artırılması. *Avrupa Bilim ve Teknoloji Dergisi*, (Özel Sayı), 131-135.

Öz

Termal görüntüleme, gözle görülmeyen kızılötesi enerjisi (ısıyı) esas alan ve görüntünün genel yapısını kızılötesi enerjiye göre oluşmuş renkler ve şekillerin belirlediği görüntüleme sistemidir. Genelde güvenlik amaçlı kullanılmakla birlikte çok çeşitli sektörlerin de kullanımına açıktır. Özellikle son yıllarda termal görüntüleme sistemleri medikal alanda da oldukça geniş bir kullanım alanı bulmaktadır. Termal görüntüleme, tasarımı zor ve maliyeti yüksek bir görüntüleme sistemidir. Bundan dolayı termal görüntülerin çözünürlüğünün artırılması için termal görüntüleme alanındaki süper çözünürlük uygulamalarına ilgi son yıllarda oldukça artmıştır. Burada derin öğrenme alanındaki gelişmeler bu çalışmaları hızlandırmış ve başarıyı artırmıştır. Bu çalışmada da insan yüzlerine ait RGB termal görüntüler üzerinde çekişmeli üretici ağlar kullanılarak süper çözünürlük uygulaması gerçekleştirilmiştir. Bu çalışmada Variocam HD termal kameradan elde edilen görüntüler yüksek çözünürlüklü, Flir One Pro termal kameradan elde edilen görüntüler ise düşük çözünürlüklü görüntüler olarak kullanılmıştır. Bu proje için 12 kişiye ait 5' er adet termal görüntü çifti (yüksek çözünürlüklü-düşük çözünürlüklü) kullanılmıştır. Bu görüntülerden 45' er çift görüntü eğitim veri seti olarak, 15' er çift görüntü ise test veri seti olarak ayrılmıştır. Çekişmeli üretici ağın eğitilmesi sırasında gradyanın yok olması probleminin önüne geçmek ve ağın daha hızlı eğitilmesini sağlamak amacıyla hem üretici ağ hem de ayırt edici ağ kısmında batch normalizasyon katmanları kullanılmıştır. Artık (residual) bloklar çok derin ağların eğitim zorluklarını kolaylaştırarak başarı performansını artırdığı için üretici ağında, ResNet' e benzer şekilde atlamalı bağlantı (skip connection) uygulanmıştır. Ağın eğitilmesi sonucunda elde edilen sonuçların başarı performansı PSNR (tepe sinyal gürültü oranı) ve SSIM (yapısal benzerlik indeksi ölçümü) görüntü kalite metrikleri kullanılarak değerlendirilmiştir. Sonuç olarak biküçük interpolasyona kıyasla PSNR değerlerinde yaklaşık 1,5dB'lik bir artış ve SSIM değerlerinde yaklaşık % 6'lık bir artış gözlemlenmiştir. İlerde yapılacak çalışmalarda, derin ağın daha fazla veri kullanılarak ve iki farklı kameradan elde edilen görüntülerin renk tonları birbirine benzetilerek eğitilmesi sonucunda yüksek maliyetli termal kamera sistemlerine bir alternatif çözüm sunulabilecektir.

Anahtar Kelimeler: Termal görüntüleme, Derin öğrenme, Süper çözünürlük.

Enhancement of Low Resolution Thermal Face Image Resolution Using Deep Learning

Abstract

Thermal imaging is an imaging system based on invisible infrared energy (heat) and the general structure of the image is determined by colors and shapes formed according to infrared energy. Although it is generally used for security purposes, it is open to use in a wide variety of sectors. Especially in recent years, thermal imaging systems have found a wide range of use in the medical field. Thermal imaging is an imaging system that is difficult to design and expensive. Therefore, interest in super resolution applications in the field of thermal imaging to increase the resolution of thermal images has increased considerably in recent years. Here, developments in the field of deep learning have accelerated these studies and increased the success. In this study, super resolution application was carried out on RGB thermal images of human faces by using adversarial generating networks. In this study, the images obtained from the Variocam HD thermal camera were used as high resolution images, while the images obtained from the Flir One Pro thermal camera were used as low resolution images. For this project, 5 pairs of thermal images (high resolution-low

* Sorumlu Yazar: Konya Teknik Üniversitesi, Mühendislik Fakültesi, Elektrik Elektronik Mühendisliği Bölümü, Konya, Türkiye, ORCID: 0000-0001-7831-6724, fatih.senalp@gmail.com

resolution) belonging to 12 people were used. Of these images, 45 are separated as double image training data set, and 15 double images as test data set. Batch normalization layers were used in both the generative network and the discriminator network part in order to avoid the problem of gradient disappearance during the training of the generative adversarial network and to provide faster training of the network. Since residual blocks facilitate the training difficulties of very deep networks and increase the success performance, a skip connection has been applied in the generator network similar to ResNet. The success performance of the results obtained as a result of training the network was evaluated using the image quality metrics PSNR (peak signal to noise ratio) and SSIM (structural similarity index measure). As a result, an increase of approximately 1.5dB in PSNR values and an increase of approximately 6% in SSIM values were observed compared to bicubic interpolation. In future studies, an alternative solution to high cost thermal camera systems can be offered as a result of training the deep network by using more data and by simulating the color tones of images obtained from two different cameras.

Keywords: Thermal imaging, Deep learning, Super resolution.

1. Giriş

Kızılötesi termal görüntüleme, non-invazif, iyonize olmayan, zararsız ve hızlı bir yöntemdir. Temassız sıcaklık ölçüm imkanı sunmaktadır. Bu sebeplerden dolayı termal görüntüleme cihazları çeşitli uygulamalarda önemli bir problem teşhis aracı olarak değerlendirilmektedir. Termal görüntüleme, görüntü işleme alanında kamu güvenliği, sağlık, savunma gibi alanlarda kullanılabildiği için avantajlı bir yapıya sahiptir. Fakat termal görüntülerde, yüksek frekans bölgeleri tam olarak tespit edilememekte ve detay seviyesi düşük olan bulanık görüntüler elde edilmektedir. Bundan dolayı termal kamera görüntülerinde süper çözünürlük tekniklerinin uygulanması, tasarımı oldukça zor ve pahalı olan yüksek çözünürlüklü termal kameralara alternatif çözümler sunabilmektedir (Zhang vd., 2018).

Son yıllarda daha fazla görüntü kalitesi sağlayan kameralara olan talep çok hızlı bir şekilde artış göstermiştir. Ancak donanım seviyesinde elde edilecek çözünürlük miktarının sınırlı kalması bu konuda farklı çalışmaları tetiklemiştir. SR(süper çözünürlük), gözlemlenmiş düşük çözünürlüklü görüntüden yüksek çözünürlüklü görüntü elde etme tekniğidir. Başka bir deyişle yüksek çözünürlüklü(HR) bir görüntünün, düşük çözünürlüklü(LR) karşılığından tahmin edilmesidir (Yue vd., 2018). Süper çözünürlük ile ilgili temel çalışmalarda bazı interpolasyon teknikleri kullanılmaktadır. Interpolasyon, bilinen verileri kullanarak bilinmeyen noktadaki değerleri tahmin etme yöntemidir (Toyran, 2008). Son yıllarda ise derin ağ modelleriyle gerçekleştirilen süper çözünürlük veya görüntü kalitesi artırma uygulamaları popüler bir hale gelmiştir (Dong vd., 2016).

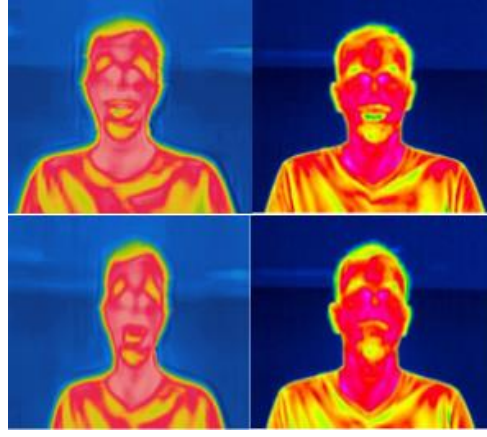
Süper çözünürlüğün oldukça geniş bir uygulama alanı vardır ve bu kavramın başlıca uygulamaları şu şekilde sıralanabilir; uydu veya hava görüntülerinin (Dong vd., 2016), yüz imgelerinin(Guei vd., 2018), iris ve göz imgelerinin (Nguyen vd., 2013), yazı çıktılarının (Glasner vd., 2009), işaret veya plaka bilgilerinin (Glasner vd., 2009), parmak izi imgelerinin geliştirilmesi (Singh vd., 2015). Bunlara ek olarak biyomedikal imge işleme alanında da süper çözünürlük kavramına ait çok sayıda uygulama bulunmaktadır (Gu vd., 2020).

Süper çözünürlük alanında yapılan çalışmalarda genellikle yüksek çözünürlüklü görüntülerin biküzik interpolasyon gibi yöntemlerle boyutları düşürülerek düşük çözünürlük görüntüler elde edilmektedir ve veri seti bu şekilde oluşturulmaktadır. Bu çalışmada ise birbirinden bağımsız 2 farklı kameradan alınan görüntüler veri seti olarak kullanılmıştır. Variocam HD termal kamera yüksek maliyetli ve kaliteli görüntüler elde edilmesini sağlarken, maliteyi çok daha düşük Flir One Pro termal kameradan elde görüntüler daha düşük kalitede görüntüler sağlamaktadır. Buradaki amaç, orijinal bir görüntüyü bozup o görüntünün çözünürlüğünü iyileştirmekten ziyade düşük maliyetli kameradan elde görüntülerin çözünürlüğünü yüksek maliyetli ve kaliteli sonuçlar veren kameranın performansına yaklaştırmaktır.

2. Materyal ve Metot

2.1. Veri Setinin Elde Edilmesi

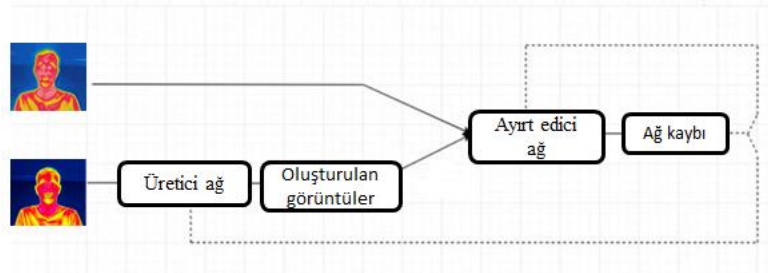
Gerçekleştirilen çalışmada, 12 farklı insan yüzüne ait iki farklı termal kameradan elde edilen görüntülerle eğitim ve test veri seti oluşturulmuştur. Görüntüler elde edilirken, pozlama farklılıklarının sağlayabileceği bilgi çeşitliliğinden yararlanmak amacıyla her kişi için 5 farklı RGB termal görüntü alınmıştır. Görüntü çekimleri yaklaşık olarak 1 metre mesafeden yapılmıştır. Gerçekleştirilen çalışmada, Variocam HD termal kamera kullanılarak 640x480 boyutunda yüksek çözünürlüklü termal görüntüler elde edilmiştir. Ayrıca Flir One Pro termal kamera ile elde edilen düşük çözünürlüklü termal görüntüler yeniden boyutlandırılarak 160x120 boyutuna sahip olacak şekilde düzenlenmiştir. Böylece yüksek çözünürlüklü görüntülerden ve buna karşılık gelen düşük çözünürlüklü görüntülerden oluşan veri seti elde edilmiştir. Şekil 1' de bir kişiye ait iki farklı düşük-yüksek çözünürlüklü görüntü çifti gösterilmektedir. Eğitim ve test verileri birbirinden farklı olacak şekilde, 45' şer adet yüksek ve düşük çözünürlüklü RGB termal görüntü eğitim verisi olarak belirlenmiş olup, 15'şer adet yüksek ve düşük çözünürlüklü RGB termal görüntü ise test verisi olarak belirlenmiştir.



Şekil 1. Düşük ve yüksek çözünürlüklü görüntü çifti

2.2. Derin Öğrenme Modeli

Bu çalışmada, derin ağ modeli olarak ÇÜA (çekişmeli üretici ağlar) temelli geliştirilen bir model olan SRGAN(super resolution generative adversarial networks) kullanılmıştır (Ledig vd., 2017). ÇÜA mimarisi üretici(generator) ağ ve ayırt edici(discriminator) ağ olmak üzere iki bölümden oluşmaktadır. Düşük çözünürlüklü görüntüler üretici ağ vasıtasıyla üst örneklenir ve süper çözünürlüklü görüntüyü çıktı olarak oluşturmaktadır. Ayırt edici ağ ise süper çözünürlüklü görüntü ile yüksek çözünürlüklü görüntü arasında ayırım yapabilmesi ve ağ kaybını kullanarak geri besleme yöntemiyle(back propagation) ayırt edici ağın ile üretici ağın eğitilmesi için kullanılmıştır. Ağın mimarisi Şekil 2' de gösterilmektedir(Ledig vd., 2017). Bu ağ modeli hem veri setinin büyüklüğü hem de veri setindeki görüntülerin boyutlarına göre yeniden düzenlenerek uygulanmıştır. Görüntülerden rastgele 32x32 boyutunda parçalar alınarak üretici ağda işlenecek ve ayırt edici ağa 128x128 boyutunda üst ölçeklenmiş olarak iletilecek şekilde düzenlenmiştir. Bu şekilde ağın eğitimi 7000 iterasyon boyunca devam ettirilmiştir. Ağın eğitimi aşamasında öğrenme oranı 0.0003 tercih edilmiş olup optimasyon metodu olarak Adam optimizasyon algoritması kullanılmıştır.



Şekil 2. ÇÜA Mimarisi

Düşük çözünürlüklü görüntüler üretici ağ vasıtasıyla üst örneklenir ve süper çözünürlüklü görüntüyü çıktı olarak oluşturmaktadır. Ayırt edici ağın süper çözünürlüklü görüntü ile yüksek çözünürlüklü görüntü arasında ayırım yapabilmesi için ÇÜA kaybı kullanılarak geri besleme yöntemi vasıtasıyla ayırt edici ağ ve üretici ağ eğitilmiştir. Ayrıca oluşturulan süper çözünürlüklü görüntü ile yüksek çözünürlüklü (ground truth) görüntüler arasındaki algısal(perceptual) farkları elde etmek ve değerlendirmek için content ve adversarial kayıpları kapsayan algısal kayıp (perceptual loss) kullanılmaktadır (Johnson, Alahi ve Li, 2016).

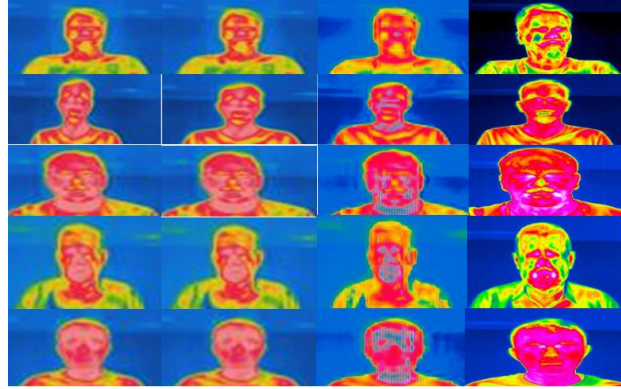
Ağ mimarisi genel olarak konvolüsyon katmanları, batch normalizasyon ve aktivasyon katmanlarından (relu, leakyrelu, swish, tanh) meydana gelmektedir. Üretici ağında swish ve ReLU aktivasyon fonksiyonları kullanılarak hem negatif kısımda öğrenmenin sağlanması hem de eğitim hızında önemli bir düşüşün önüne geçilmesi hedeflenmiştir. Böylece ağın daha az görüntü ile hızlı bir şekilde eğitilmesi sağlanmış olmaktadır. Ayırt edici ağında ise LeakyReLU aktivasyon fonksiyonu kullanılarak negatif öğrenme daha stabil bir şekilde gerçekleşmiştir. Bunun dışında gradyanın yok olması probleminin önüne geçmek ve ağın daha hızlı eğitilmesini sağlamak amacıyla hem üretici (generator) hem de ayırt edici (discriminator) kısmında batch normalizasyon katmanları kullanılmıştır (Ioffe ve Szegedy, 2015). Ayrıca artık(residual) bloklar çok derin ağların eğitim zorluklarını kolaylaştırarak başarı performansını artırdığı için üretici ağında, ResNet' e benzer şekilde atlama bağlantı (skip connection) uygulanmıştır.

3. Araştırma Sonuçları ve Tartışma

3.1. Araştırma Sonuçları

Bu çalışmada, görsel kalitenin iyileştirilmesinin yanı sıra başarı performansının değerlendirilmesi amacıyla PSNR ve SSIM görüntü kalite metrikleri göz önüne alınmıştır. Gerçekleştirilen çalışmada giriş görüntüsünün boyutu x4 üst örneklenerek süper çözünürlüklü görüntü elde edilmiştir. Giriş görüntülerinin boyutları 160x120 iken referans (ground truth) ve çıktı olarak elde edilen

süper çözünürlüklü görüntülerin boyutları 640x480'dir. Elde edilen sonuçlardan bazıları Şekil 3' de gösterilmektedir. Burada ilk sütun düşük çözünürlüklü giriş görüntülerini, ikinci sütun bikübik interpolasyon sonucu oluşan görüntüleri, üçüncü sütun derin ağı çıkıtsı olan süper çözünürlüklü görüntüleri ve son sütun temel referans olarak kullanılan yüksek çözünürlüklü görüntüleri temsil etmektedir.



Şekil 3. Örnek görüntü sonuçları

Görüntü kalitesinde en önemli etkilerden biri gürültüdür. Burada PSNR değerinin yüksekliği, görüntünün o derece kaliteli olduğunu ve gürültünün de o derece az olduğunu ifade etmektedir. PSNR değeri hesaplanırken referans bir giriş görüntüsü alınır ve bu görüntüdeki gürültüyü ölçmek için bir çıkış görüntüsüyle karşılaştırılır. PSNR değeri Denklem 1'de verilen formülle hesaplanmaktadır (Javaid vd., 2013).

$$PSNR = 20 \times \log_{10} (255 \sqrt{MSE(I,I_0)}) \quad (1)$$

Yapısal benzerlik indeksi(SSIM), iki çerçeve arasındaki yapısal benzerliği ölçen bir kalite ölçütüdür. Bu görüntü kalite metriği, iki çerçeve arasındaki yapısal bilgiye ait varyasyondaki değişimleri dikkate aldığı için çerçevelerdeki kalite bozulmasını ifade etmektedir. SSIM değeri Denklem 2' de verilen formülle hesaplanmaktadır (Javaid vd., 2013).

$$SSIM(n) = [2\mu_I \mu_D (n) + C_1] / [\sigma_I^2 (n) + \sigma_D^2 (n) + C_2] \quad (2)$$

Derin ağı modelinin eğitilmesi sonucu elde edilen görüntülere ait PSNR/SSIM değerleri, bikübik interpolasyon sonucu elde edilen görüntülere ait PSNR/SSIM değerleri ve düşük çözünürlüklü görüntüye ait PSNR/SSIM değerleri ile karşılaştırılmıştır. Tablo 1' de sonuçlar gösterilmektedir. Düşük çözünürlüklü giriş görüntüleriyle mukayese edildiğinde, bikübik interpolasyon sonucunda PSNR değerlerinde yaklaşık 0,5 dB'lik bir artış sağlanırken süper çözünürlüklü görüntülerde yaklaşık 2 dB'lik bir artış gözlemlenmiştir. SSIM değerlerinde ise sırasıyla %1,5 ve %7 civarında bir artış gözlemlenmiştir.

Tablo 1. PSNR/SSIM Sonuçları

Düşük çözünürlüklü görüntü		Bikübik interpolasyon		SRGAN	
PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
23,357	0,738	23,992	0,753	25,525	0,816

3.2. Tartışma

Bu çalışma sonucundan elde edilen sonuçlar veri setinin küçüklüğü göz önüne alındığında tatmin edici boyutlardadır. İlerleyen çalışmalarda daha fazla termal görüntünün kullanılması ve farklı mesafelerde yapılacak çekimlerden elde edilecek görüntü zenginliği ile birlikte termal kameradan elde edilen termal görüntülerin renk tonlarının birbirine daha fazla benzetilmesi sağlanarak maliyeti yüksek olan termal görüntüleme sistemlerinin görüntü kalitesine yaklaşılması öngörülmektedir.

4. Sonuç

Termal görüntülemenin yaygınlaşan kullanım alanına karşılık sistem maliyetinin yüksek olması, araştırmacıların termal görüntülerde süper çözünürlük uygulamalarına odaklanmasına sebep olmuştur. Bu çalışmada da derin öğrenme modeli ile termal görüntüler üzerinde süper çözünürlük uygulaması yapılarak, temel bir görüntü iyileştirme tekniği olan bikübik interpolasyon yöntemiyle karşılaştırılmıştır. Sonuçlar görüntü kalite metrikleri olan PSNR ve SSIM değerleriyle değerlendirilmiş ve başarılı

sonuçlar elde edilmiştir. Elde edilen sonuçlar ileride yapılacak uygulamalarda, yapay zeka ve büyük veri seti kullanımının yüksek maliyetli termal görüntüleme sistemlerine yakın bir performans gösterebileceğine dair umut vermektedir.

Bu çalışmada kullanılan Variocam HD termal görüntüleme sisteminin maliyeti yaklaşık 15.000 Euro civarındadır. Buna karşılık düşük çözünürlüklü görüntüleri elde etmek için kullanılan Flir One Pro termal kameranın maliyeti 600 Euro civarındadır. Görüldüğü üzere iki termal kamera arasında ciddi bir miktarda maliyet farkı mevcuttur. Ayrıca klasik termal görüntüleme sisteminin kurulumu ve görüntülerin elde edilmesi Flir One Pro termal kamera ile karşılaştırıldığında daha fazla zaman almaktadır. Sonuç olarak hem kullanımı daha kolay olan hem de maliyeti çok daha düşük olan bir sistem meydana getirilerek termal görüntüleme sistemlerindeki kullanımın yaygınlaşması amaçlanmaktadır.

Kaynakça

- Dong, C., Loy, C. C., He, K. ve Tang, X.(2016). Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38, 295-307, doi: 10.1109/TPAMI.2015.2439281.
- Dong W., Fu F., Shi G., Cao X., Wu J., Li G. ve Li X.(2016). Hyperspectral image super-resolution via non-negative structured sparse representation. IEEE Transactions on Image Processing, 25 (5), 2337-2352.
- Glasner D., Bagon S. ve Irani M.(2009). Super-resolution from a single image. In Computer Vision, IEEE 12th International Conference on , 349-356.
- Gu, Y. vd.(2020). MedSRGAN: medical images super-resolution using generative adversarial networks. Multimed Tools App.l
- Guei, A. ve Akhlooufi, M.(2018). Deep learning enhancement of infrared face images using generative adversarial networks. Applied Optics, 57 (18), 98.
- Ioffe, S. ve Szegedy, C.(2015). Batch normalization: accelerating deep network training by reducing internal covariate shift. Proceedings of The 32nd International Conference on Machine Learning (ICML),448–456.
- Javaid, H., Babar, T.K., Rasool, A. ve Saghir, R.U.(2013). Video colour variation detection and motion magnification to observe subtle changes, M.Sc.Thesis, Blekinge Institute of Technology, Faisalabad, Pakistan, 57.
- Johnson, J., Alahi, A. ve Li, F.(2016). Perceptual losses for real-time style transfer and super resolution. In European Conference on Computer Vision (ECCV), 694–711. Springer.
- 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, 105-114, doi: 10.1109/CVPR.2017.19.
- Nguyen K., Fookes C., Sridharan S. ve Denman S.(2013). Feature-domain super-resolution for iris recognition, Computer Vision and Image Understanding, 117 (10), 1526-1535.
- Singh K., Gupta A. ve Kapoor R.(2015). Fingerprint image super-resolution via ridge orientation-based clustered coupled sparse dictionaries. Journal of Electronic Imaging, 24 (4), 043015.
- Toyran, M.(2008). Düşük çözünürlüklü görüntülerden süper çözünürlüklü görüntü oluşturma. (Yüksek lisans tezi), Fen Bilimleri Enstitüsü, İstanbul.
- Yue, L., Shen, H., Li, J., Yuan, Q., Zhang, H. ve Zhang, L.(2018), Image super-resolution:the techniques, applications, and future. *Signal Processing*, 128,389-408.
- Zhang, X., Li, C., Meng, Q., Liu,S., Zhang, Y. ve Wang, J.(2018). Infrared image super resolution by combining compressive sensing and deep learning. *Sensors (Basel)*, 18 (8), 2587.