



POLİTEKNİK DERGİSİ

JOURNAL of POLYTECHNIC

ISSN: 1302-0900 (PRINT), ISSN: 2147-9429 (ONLINE)

URL: <http://dergipark.org.tr/politeknik>



Termal yüz görüntülerinden oluşan yeni bir veri seti için derin öğrenme tabanlı süper çözünürlük uygulaması

Deep learning based super resolution application for a new data set consisting of thermal facial images

Yazar(lar) (Author(s)): Fatih Mehmet ŞENALP¹, Murat CEYLAN²

ORCID¹: 0000-0001-7831-6724

ORCID²: 0000-0001-6503-9668

Bu makaleye şu şekilde atıfta bulunabilirsiniz (To cite to this article): Şenalp F. M., Ceylan M., “Termal yüz görüntülerinden oluşan yeni bir veri seti için derin öğrenme tabanlı süper çözünürlük uygulaması”, *Politeknik Dergisi*, *(*) : *, (*).

Erişim linki (To link to this article): <http://dergipark.org.tr/politeknik/archive>

DOI: 10.2339/politeknik.904675

Termal Yüz Görüntülerinden Oluşan Yeni Bir Veri Seti için Derin Öğrenme Tabanlı Süper Çözünürlük Uygulaması

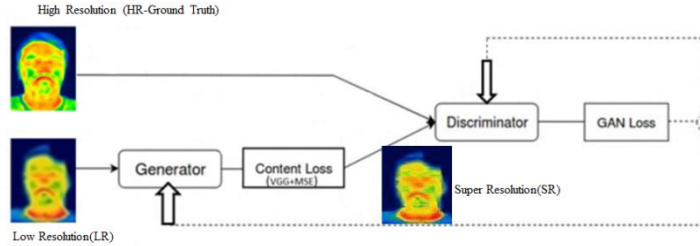
Deep Learning Based Super Resolution Application for a New Data Set Consisting of Thermal Facial Images

Önemli noktalar (Highlights)

- ❖ Derin öğrenme tabanlı süper çözünürlük uygulaması / Deep learning based super resolution implementation.
- ❖ Geleneksel yöntemlerden farklı bir şekilde elde edilen veri seti ile süper çözünürlük uygulaması gerçekleştirilmiştir / Super resolution application was carried out with the data set obtained differently from traditional methods.
- ❖ Elde edilen sonuçlar görüntü kalite metrikleri olan PSNR ve SSIM değerleri kullanılarak değerlendirilmiştir / The obtained results obtained were evaluated using PSNR and SSIM values which are image quality metrics.

Grafik Özet (Graphical Abstract)

Bu çalışmada, termal görüntüler kullanılarak derin öğrenme tabanlı süper çözünürlük uygulamaları gerçekleştirilmiştir./ In this study, deep learning based super resolution applications were carried out using thermal images.



Şekil 1. Geliştirilen modelin mimarisi / Figure 1. Developed model architecture

Amaç (Aim)

Bu çalışmanın esas amacı, termal yüz görüntüleri üzerinde derin öğrenme tabanlı bir süper çözünürlük uygulaması gerçekleştirmektir./ The main purpose of this study is to realize a deep learning based super resolution application on thermal face images.

Tasarım ve Yöntem (Design & Methodology)

Geliştirilen GAN tabanlı model, iki ayrı veri seti üzerinde görüntü kalite metrikleri kullanılarak SRGAN ile karşılaştırılmıştır./ Developed model was compared with SRGAN using image quality metrics on two datasets.

Özgünlük (Originality)

Bu çalışma, birbirinden bağımsız iki farklı termal kamera kullanılarak elde edilen ve termal yüz görüntülerinden oluşan veri seti üzerinde gerçekleştirilen ilk çalışmadır./ This study is the first study conducted on a dataset consisting of thermal facial images obtained using two independent thermal cameras.

Bulgular (Findings)

Elde edilen sonuçlar, geliştirilen modelin SRGAN modelinden daha başarılı olduğu ve veri setindeki düzenlemelerin başarıyı artırdığı görülmüştür. / The results show that the developed model is more successful than the SRGAN model and the arrangements in the data set increase the success.

Sonuç (Conclusion)

Düşük maliyete sahip termal kameradan elde edilen görüntülerin çözünürlüğünü, yüksek maliyete sahip termal kameradan elde edilen kaliteli görüntülerin çözünürlüğüne yaklaştırma açısından başarılı sonuçlar elde edilmiştir./

Etik Standartların Beyanı (Declaration of Ethical Standards)

Bu makalenin yazar(lar)ı çalışmalarında kullandıkları materyal ve yöntemlerin etik kurul izni ve/veya yasal-özel bir izin gerektirmediğini beyan ederler. / The author(s) of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.

Termal Yüz Görüntülerinden Oluşan Yeni Bir Veri Seti için Derin Öğrenme Tabanlı Süper Çözünürlük Uygulaması

Araştırma Makalesi / Research Article

Fatih Mehmet ŞENALP*, Murat CEYLAN

Mühendislik ve Doğa Bilimleri Fakültesi, Elektrik Elektronik Mühendisliği Bölümü, Konya Teknik Üniversitesi, Türkiye
(Geliş/Received : 28.03.2021 ; Kabul/Accepted : 17.01.2022 ; Erken Görünüm/Early View : 26.01.2022)

ÖZ

Termal kamera sistemleri, ısı değişiminin tespitini gerektiren her türlü uygulamada faydalanılabilmemesine rağmen termal görüntüleme sistemleri oldukça yüksek maliyete sahip sistemlerdir ve bu durum termal sistemlerin yaygın bir şekilde kullanımını zorlaştırmaktadır. Ayrıca termal görüntüler elde edilirken düşük kalitede bulanık görüntüler meydana gelebilmektedir. Bu makalede, iki farklı termal kameradan elde edilen termal yüz görüntülerinden oluşan bir veri seti üzerinde süper çözünürlük uygulaması gerçekleştirilmiştir. Belirtilen veri seti geleneksel yöntemlerden farklı bir şekilde oluşturulmuş olup, düşük çözünürlüklü (LR) termal görüntüler 160x120 termal çözünürlüğe sahip kameradan elde edilirken yüksek çözünürlüklü (referans) görüntüler ise 640x480 termal çözünürlüğe sahip kameradan elde edilmiştir. Daha sonra bu görüntülerdeki gereksiz kısımlar kırılarak sadece yüz bölgesine odaklanılarak başka bir çalışma daha gerçekleştirilmiştir. Bu uygulamalar için çekişmeli üretici ağlar (GAN) tabanlı bir derin öğrenme modeli geliştirilmiştir. Sonuçların başarı performansını görüntü kalite metrikleri PSNR (tepe sinyal gürültü oranı) ve SSIM (yapısal benzerlik endeksi) ile değerlendirmeye alınmıştır. Sadece yüz bölgelerine odaklanılarak gerçekleştirilen uygulamanın sonuçları orijinal görüntülerle yapılan uygulama sonuçlarına kıyasla daha iyi olduğu görülmüştür. Bunun yanı sıra bu çalışma, daha az maliyetli termal kamera tarafından elde edilen termal görüntülerin çözünürlüğünü, yüksek maliyete sahip olan ve yüksek kalitede görüntüler elde edilebilen termal kameranın çözünürlüğüne bilhassa görsel olarak yaklaşırma yönünden olumlu sonuçlar vermiştir.

Anahtar Kelimeler: Termal görüntüleme, süper çözünürlük, derin öğrenme, veri setleri.

A Deep Learning-Based Super Resolution Approach for Thermal Face Images Using New Datasets

ABSTRACT

Although thermal camera systems can be used in any application that requires the detection of temperature change, thermal imaging systems are highly costly systems and this situation makes difficult the common use of thermal systems. In addition, blurry images of low quality can occur when thermal images are obtained. In this article, super resolution application has been carried out on a data set consisting of thermal face images obtained from two different thermal cameras. The specified data set was created differently from traditional methods, low resolution (LR) thermal images were obtained from a 160x120 thermal resolution camera, while high resolution (reference) images were obtained from a camera with a thermal resolution of 640x480. Later, unnecessary parts of these images were cropped and another study was carried out by focusing only on the face area. A deep learning model based on adversarial generative networks (GAN) has been developed for these applications. The success performance of the results was evaluated by the image quality metrics PSNR (peak signal to noise ratio) and SSIM (structural similarity index). It has been observed that the results of the application performed by focusing only on the facial areas are better than the results of the application with original images. In addition, this study gave positive results in terms of approximating the resolution of the thermal images obtained by the less costly thermal camera to the resolution of the thermal camera, which has a high cost and can obtain high quality images, especially visually.

Keywords: Thermal imaging, super resolution, deep learning, datasets.

1. GİRİŞ (INTRODUCTION)

Kızılötesi termografi, elektromanyetik spektrum üzerinde görünür alanın dışındaki ısı yayımını renkli bir görüntü veya tonlama biçiminde görünür hale getirmektedir. Bu sistemler temassız bir şekilde sıcaklık ölçüm imkanı sağladığı için insanlar üzerinde gerçekleştirilen uygulamalarda non-invazif ve zararsız

bir yöntem olarak önemli bir yere sahiptir. Termal görüntüleme cihazları başta askeri ve sivil alanlarda olmak üzere ısı değişkenliklerin tespitini gerektiren hemen hemen her alanda önemli bir problem teşhis aracı olarak kullanılabilmektedir. Yaygın kullanım alanına sahip olmasına karşın düşük maliyetli termal görüntüleme sistemleri detay bilgisi yeterli olmayan bulanık görüntüler meydana getirmektedirler. Bundan dolayı termal görüntülerin çözünürlüğünün iyileştirilmesine yönelik çalışmalar, görüntü işleme

*Fatih Mehmet Şenalp (Corresponding Author)
e-posta : fatih.senalp@gmail.com.tr

alanında odak noktası haline gelmiştir. Termal görüntüler üzerinde süper çözünürlük tekniğinin uygulanması sonucunda yüksek maliyete sahip termal görüntüleme sistemlerine alternatif bir çözüm ortaya koyulabilecektir [1].

Süper çözünürlük, elde edilmiş düşük çözünürlüklü görüntülerden daha yüksek çözünürlüğe sahip bir görüntü elde etme tekniğidir. Süper çözünürlüklü bu görüntü, yüksek çözünürlüklü (referans) bir görüntüye karşılık gelen düşük çözünürlüklü (LR) görüntüden tahmini olarak ifade edilmektedir [2]. Süper çözünürlük alanında gerçekleştirilen ilk uygulamalarda interpolasyon tekniklerinden faydalanılmıştır. Başlıca kullanılan teknikler en yakın komşu, lineer ve biküçük interpolasyon yöntemleridir [3]. En yakın komşu interpolasyonu, interpolasyon noktasına en yakın tek bir pikseli dikkate almaktadır. Çift doğrusal interpolasyonda ise bilinmeyen piksele en yakın 2x2 (4 piksel) komşuları dikkate alınmaktadır. Belirtilen 3 yöntem içerisinde en başarılı sonuçları veren biküçük interpolasyon yönteminde bilinmeyen piksele en yakın bilinen 4x4 (16 piksel) komşuları dikkate alınırken, bilinmeyen piksele daha yakın olan komşulara daha yüksek ağırlık verilmektedir. En yakın komşu yönteminde elde edilen görüntülerde mozaikleşmeler meydana gelmektedir. Çift doğrusal interpolasyonda ise bulanık görüntüler oluşmaktadır [4]. Bu gibi durumları minimize edebilmesi açısından biküçük interpolasyon yöntemi daha çok tercih edilmektedir; fakat son yıllardaki donanımsal ilerlemeler sebebiyle derin ağların süper çözünürlük alanında kullanımı yaygın bir hale gelmiştir [5].

Süper çözünürlük metotları, uydu veya hava görüntülerinin [6], yüz imgelerinin [7], iris ve göz imgelerinin [8], yazı çıktılarının [9], işaret ve plaka tanımanın [10], astronomi imgelerinin [11], parmak izi görüntülerinin [12] iyileştirilmesi gibi birçok uygulama alanına sahiptir. Ayrıca, biyomedikal görüntü işleme alanında ise son yıllarda süper çözünürlük ile ilgili çok sayıda uygulama ortaya konmuştur [13].

Derin öğrenme tabanlı gerçekleştirilen ilk süper çözünürlük çalışmalarında evrişimli sinir ağları (CNN) tabanlı modeller geliştirilmiştir. Dong ve ark. SRCNN modelini geliştirerek düşük çözünürlüklü ve yüksek çözünürlüklü görüntüler arasında baştan uca haritalama tekniğinden faydalanmışlardır [5]. Ayrıca, Kim ve ark. SRCNN modeline dayanarak geliştirdikleri VDRS modelinde artık öğrenme (residual learning) yönteminden faydalanmışlardır. Burada, evrişim katmanlarının sayısını 3 yerine 20 olarak düzenlemişler ve daha başarılı sonuçlar gözlemlemişlerdir. Dört farklı veri seti ile gerçekleştirilen uygulamalarda PSNR değeri 0.5-1 dB aralığında artış göstermiştir. Ayrıca SSIM değerleri dikkate alındığında %1-3 aralığında artış gerçekleşmiştir [14]. Lim et al. ise yine artık öğrenme yöntemiyle geliştirilen EDSR modelini önermişler ve hem başarıyı hem de ağırlık eğitim hızını artırmışlardır. Burada EDSR modelinin PSNR değerinin VDRS modeline göre 0.5-1.5 dB aralığında arttığı, SSIM değerinin ise yaklaşık %2 arttığı görülmüştür [15].

Bahsedilen evrişimli sinir ağları tabanlı bu uygulamalar daha çok ortalama kare hata (MSE) değerini minimize etmeye yöneliktir; fakat son yıllarda görsel olarak daha kaliteli görüntüler elde edebilmek amacıyla çekişmeli üretici ağlar (GAN) tabanlı birçok model önerilmiştir. Burada, Ledig et al. çekişmeli üretici ağlar tabanlı geliştirdikleri SRGAN modelini önermişler ve geliştirdikleri bu modelin üretici ağ kısmında atlama bağlantı (skip connection) tekniği kullanmışlardır. SRGAN modeli, görsel olarak evrişimli sinir ağları tabanlı yöntemlere kıyasla daha başarılı sonuçlar elde edilmesini sağlamıştır. Böylece çekişmeli üretici ağlar tabanlı süper çözünürlük uygulamalarına olan ilgiyi artırmıştır [16].

Görünür görüntüler üzerinde gerçekleştirilen süper çözünürlük uygulamalarından sonra termal görüntüler üzerinde de süper çözünürlük uygulamaları gerçekleştirilmiştir. İlk olarak evrişimli sinir ağları tabanlı geliştirilen TEN (Thermal Enhancement Network) modeli önerilmiştir. Fakat bu çalışmada veri setinin yetersizliği sebebiyle derin ağırlık eğitiminde görünür görüntüler kullanılmıştır [17]. Daha sonra Rivadeneira ve ark. eğitim aşamasında da termal görüntülerden faydalanarak başarıyı artırmışlardır [18]. Mandanici ve ark. MISR (Multi-Image Super Resolution) modelini tasarlayarak süper çözünürlüklü termal görüntüleri elde etmişlerdir [19]. Evrişimli sinir ağları tabanlı bu çalışmalar sonucu elde edilen başarılı sonuçların ardından görsel kalitedeki iyileşmeye odaklanabilmek amacıyla çekişmeli üretici ağlar tabanlı süper çözünürlük çalışmaları odak noktası haline gelmiştir. Guei ve ark. hem yakın kızılötesi (NIR) hem de uzun dalga boyu kızılötesi (LWIR) görüntüler üzerinde süper çözünürlük uygulaması geliştirmişlerdir [7]. Sonraki bir çalışmada ise Chudasama ve ark. üç farklı standart termal veri seti üzerinde x2, x3 ve x4 üst ölçekleme tekniklerini kullanarak süper çözünürlüklü termal görüntüleri elde etmişlerdir. Burada, geliştirdikleri TherISuRNet modelini kullanmışlar ve elde edilen sonuçları literatürdeki en yeni modeller ile karşılaştırmışlardır. Sonuçlar incelendiğinde, veri setleri ve üst ölçekleme tekniklerinin genelinde TherISuRNet modeli en iyi PSNR ve SSIM değerlerini elde etmiştir [20].

Öğrenme tabanlı geleneksel süper çözünürlük uygulamaları için genellikle hazır bir görüntü veri seti ve bu veri setindeki görüntülerin alt ölçeklenmesiyle oluşturulan görüntü çiftleri (LR –HR) kullanılmaktadır. Bu hazır veri seti yüksek çözünürlüklü (referans) veri seti olarak değerlendirilmektedir. Bu çalışmada ise 2 farklı termal kameradan elde edilen yüksek çözünürlüklü (referans-ground truth) ve düşük çözünürlüklü termal yüz görüntü çiftleri üzerinde çalışmalar gerçekleştirilmiştir. Yüksek maliyetli Variocam HD termal kamera kullanılarak yüksek çözünürlüklü (HR-referans) görüntüler elde edilirken, maliyeti çok daha az olan Flir One Pro termal kamera vasıtasıyla düşük çözünürlüklü görüntüler (LR) elde edilmiştir. Bu çalışmanın sonuçları PSNR ve SSIM görüntü kalite metrikleri kullanılarak

değerlendirilmiştir. Fakat görüntülerin kaynakları farklı olduğundan bu değerlerden ziyade görsel olarak iyileşmenin olması daha önemlidir. Burada yüksek frekans bilgilerinin elde edilmesi (kenar bulma) ve bulanıklığın giderilmesi gibi parametreler önem teşkil etmektedir. Bu çalışmanın esas hedefi, geleneksel çalışmalardaki gibi yüksek kalitede bir termal görüntüyü bozup o görüntünün çözünürlüğünü artırmak yerine, çok daha düşük maliyete sahip bir termal kameradan elde edilen termal görüntülerin çözünürlüğünü yüksek maliyete sahip termal kameradan elde edilen görüntülerin çözünürlüğüne yaklaştırmaktır.

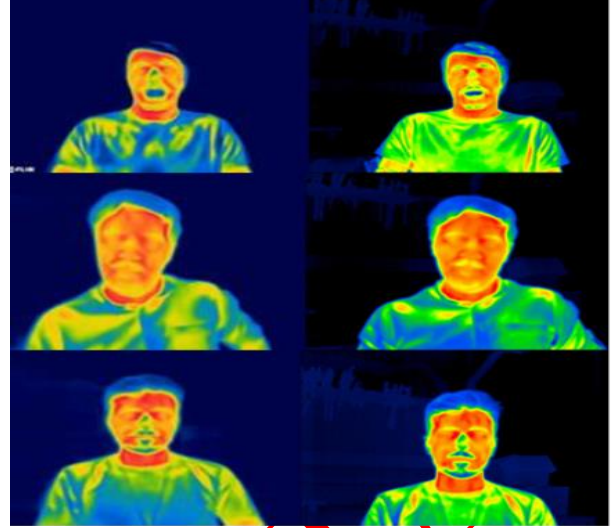
2. ÖNERİLEN YAKLAŞIM (PROPOSED APPROACH)

Bu bölümde termal görüntüler üzerinde gerçekleştirilen iki farklı çalışmaya ait detaylı bilgiler yer almaktadır. Bölüm 2.1 uygulamalarda kullanılan veri setlerinin nasıl elde edildiği ve hazırlandığı ile ilgili bilgileri sunmaktadır. Bölüm 2.2 ise kullanılan derin ağ modeli ve ağı eğitimi ile ilgili bilgiler vermektedir.

2.1. Verilerin Elde Edilmesi (Data Collection)

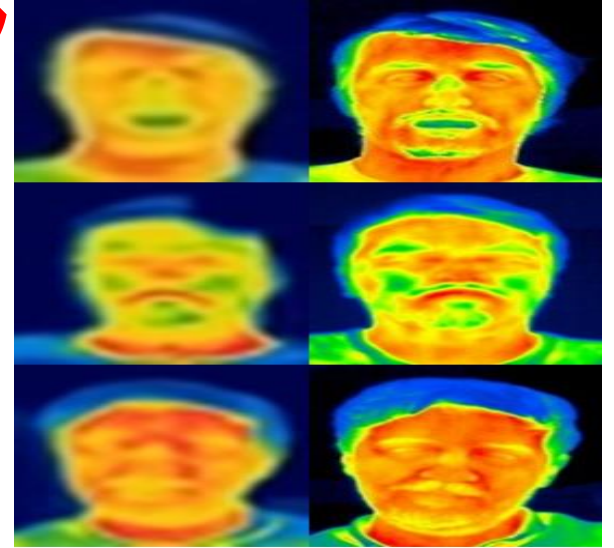
Gerçekleştirilen çalışmada, bazı teknik özellikleri Çizelge 1’ de verilmiş olan iki farklı termal kamera kullanılarak 20 farklı kişiye ait RGB formatında 450 adet yüksek ve 450 adet düşük çözünürlüklü termal yüz görüntüleri elde edilmiştir. Termal görüntülerin elde edilmesi için gerçekleştirilen çekimler yaklaşık 24°C ortam sıcaklığına sahip kapalı bir ortamda gerçekleştirilmiştir. Ayrıca termal yüz görüntüleri çekilecek insanlar, ortam sıcaklığına uyum sağlamaları amacıyla çekim işleminden en az 5 dk önce ortamda hazır bulunmuşlardır. İki farklı termal kameranın kullanıldığı görüntü çekimleri aynı kişi için önce Variocam kamera sonra Flir kamera ile vakit kaybetmeden sırayla gerçekleştirilmiştir. Ayrıca, Flir Tools programı kullanılarak yüksek ve düşük çözünürlüklü görüntüler arasındaki renk tonu farkları minimize edilmiştir.

Yüksek çözünürlüklü (HR-referans) termal görüntüler, 640x480 termal çözünürlüğe sahip Variocam HD termal kamera kullanılarak 1 metre ve 0,5 metre mesafeden çekimler yapılarak elde edilmiştir. Bu görüntüler bilgisayar ortamında 640x480 görüntü boyutuna sahiptir. Burada, farklı mesefalardan görüntü çekimleri gerçekleştirilerek pozlama çeşitliliğinden faydalanılabilmesi sağlanmıştır. Düşük çözünürlüklü görüntüler ise 160x120 termal çözünürlüğe sahip Flir One Pro termal kamera kullanılarak, aynı tekniklerin uygulanmasıyla çekimler gerçekleştirilmiştir. Bu görüntüler bilgisayar ortamına aktarıldığında 1440x1080 görüntü boyutuna sahiptir. Bu görüntüler Matlab ortamında yeniden boyutlandırılarak 160x120 görüntü boyutuna sahip düşük çözünürlüklü görüntüler elde edilmiştir [21]. Farklı termal kameralardan elde edilen bu orjinal görüntülere ait örnek görüntüler Şekil 1’de gösterilmiştir.



Şekil 1. Sol sütunda Flir One Pro termal kamera kullanılarak, sağ sütunda ise Variocam termal kamera kullanılarak elde edilen termal görüntüler (Thermal images obtained using the Flir One Pro thermal camera in the left column, thermal images obtained using the Variocam thermal camera in the right column)

Daha sonra bu görüntüler sadece yüz bölgesine odaklanabilmek için kırılarak 120x160 boyutuna sahip yüksek çözünürlüklü termal görüntülerden ve 30x40 boyutuna sahip düşük çözünürlüklü termal görüntülerden meydana gelen başka bir veri seti elde edilmiştir. Şekil 2’de ham görüntülerin kırılmasıyla oluşturulan veri setine ait örnek görüntüler gösterilmiştir.

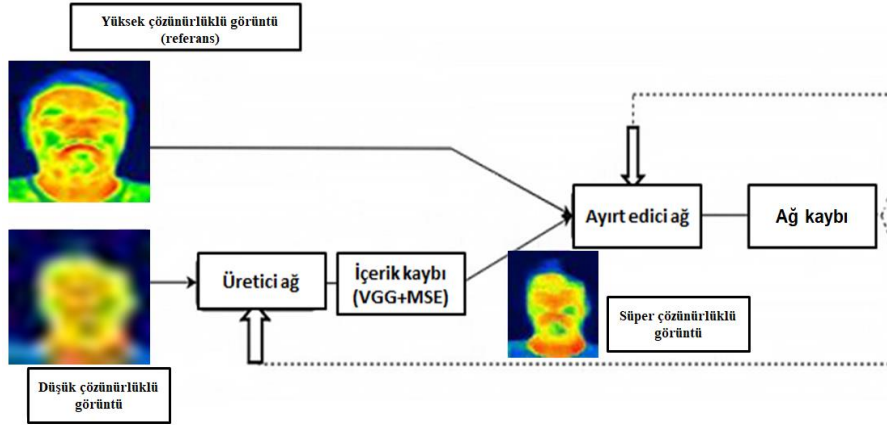


Şekil 2. Sol sütunda Flir One Pro termal kamera kullanılarak, sağ sütunda ise Variocam termal kamera kullanılarak elde edilen kırılmış termal görüntüler (Cropped thermal images obtained using the Flir One Pro thermal camera in the left column, cropped thermal images obtained using the Variocam thermal camera in the right column)

Çizelge 1. Termal kamera özellikleri (Thermal camera specifications)

Teknik Özellikler	Variocam HD	Flir One Pro
Termal Çözünürlük	640x480	160x120
Görsel Çözünürlük	640x480	1440 x 1080
Hassasiyet	1 °C	0,07 °C
Ölçüm Aralığı	-20°C - 400°C	-40°C - 2000°C

2.2. Derin Öğrenme Mimarisi ve Eğitim Süreci (Deep Learning Architecture and Training Process)



Şekil 3. TSRGAN mimarisi (TSRGAN architecture)

Elde edilen görüntülerdeki görsel kaliteyi artırmak amacıyla içerik (content) ve çekişmeli (adversarial) kayıpların birleşiminden oluşan algısal kayıp (perceptual loss) kullanılmıştır [23]. Böylece süper çözünürlüklü termal görüntüler ile yüksek çözünürlüklü (referans) termal görüntüler arasındaki görsellik farkı minimize edilmeye çalışılmıştır. Derin ağ kaybı (GAN kaybı) MSE, içerik ve çekişmeli kayıplarının toplamından oluşmaktadır. GAN kaybının formülü Denklem 2' de verilmiştir.

MSE (ortalama kare hata) değeri her zaman pozitif değere sahip olup, derin ağ modelinin performansını değerlendirmek için kullanılmaktadır. Burada kayıp değeri sıfıra ne kadar çok yaklaşırsa derin ağ modelinin o derece başarılı olduğu anlaşılmaktadır. Denklem 1'de MSE değerinin formülü verilmiş olup, n veri setinin boyutunu, e_j ise esas değer ile tahmin edilen değer arasındaki hata değerini göstermektedir [24].

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

İçerik kaybı, önceden eğitilmiş 19 katmanlı ağ (VGG19) ile hesaplanan VGG kayıp ve MSE kayıp değerinin birleşiminden oluşmaktadır. Burada VGG kayıp görüntülerin görsel olarak algılanmasını artırmak için

Bu bölüm, gerçekleştirilen uygulamalarda kullanılan derin ağ modeli olan TSRGAN (Thermal Super

Resolution GAN) modelini tanıtmaktadır. Çekişmeli üretici ağlar (GAN - generative adversarial networks) tabanlı geliştirilen bu derin ağ modeli, üretici (generator) ve ayırt edici (discriminator) ağları içeren iki ayrı bölümden meydana gelmektedir [22]. Bu derin ağ modelinde üretici ağ, düşük çözünürlüklü termal görüntüleri x4 üst ölçekleyerek süper çözünürlüklü termal görüntüleri meydana getirmektedir. Ayırt edici ağ ise elde edilen süper çözünürlüklü görüntü ile yüksek çözünürlüklü referans görüntü arasındaki farkı belirleyerek derin ağ modelinin kaybını (GAN kaybı) hesaplamaktadır. Burada, geri besleme yöntemi (back propagation) ile derin ağın eğitim süreci gerçekleştirilmektedir. Ağın mimarisi Şekil 3' te gösterilmektedir.

kullanılırken, MSE kayıp ise görsel olarak iyileşmenin yanında PSNR/SSIM değerlerini de artırabilmek için kullanılmıştır. Böylece hem piksel hem de algısal farklılıkları minimize etmek amaçlanmıştır [22].

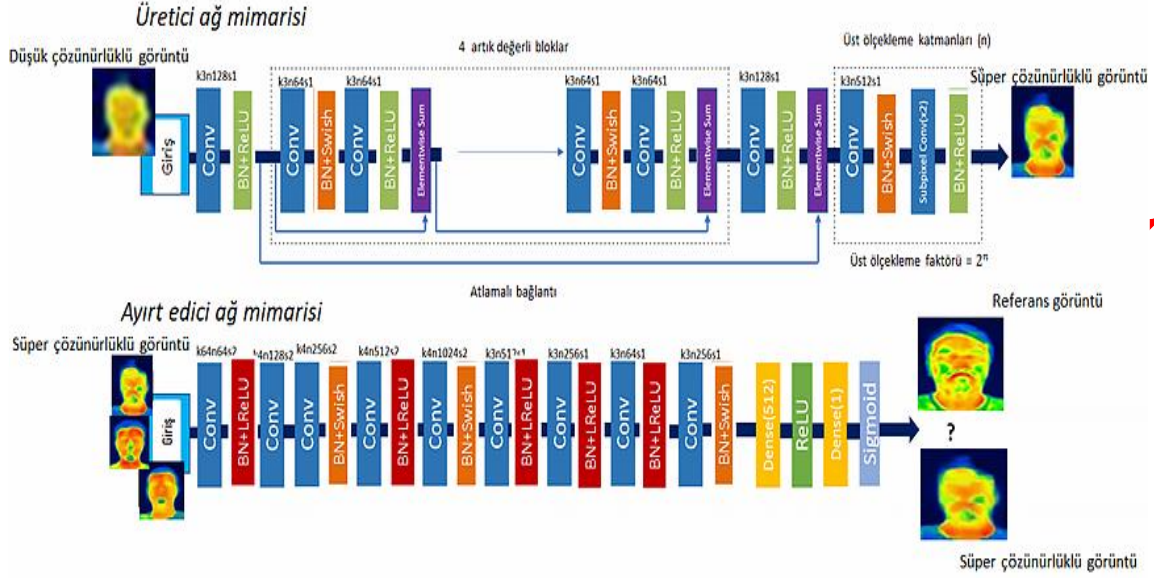
Çekişmeli kayıp, ayırt edicinin tüm eğitim örnekleri üzerindeki olasılıklarına dayalı olarak ifade edilmektedir. Bu kaybın kullanılma amacı yapısal farklılığı azaltarak görüntü kalitesini geliştirmektir. Yani termal görüntülerde sadece PSNR/SSIM değerlerine odaklanmak yerine görselliği artırmak hedeflenmiştir [25].

$$l_{\text{toplam(gan)}} = l_{\text{içerik}}(l_{\text{mse}} + l_{\text{vgg}}) + l_{\text{çekişmeli}} \quad (2)$$

Önerilen derin ağ mimarisi konvolüsyon katmanlarından, batch normalizasyon katmanlarından ve aktivasyon katmanlarından oluşmaktadır. Çekişmeli üretici ağlar tabanlı modele (TSRGAN) ait üretici ve ayırt edici ağ mimarileri Şekil 4' te gösterilmiştir. Üretici ağ kısmında negatif öğrenmenin de gerçekleşmesi için swish aktivasyon fonksiyonu kullanılırken, ReLU aktivasyon fonksiyonunun da kullanılmasıyla eğitim hızında önemli bir düşüşün önüne geçilmiştir. Bu sayede kısıtlı veri seti ile hızlı bir şekilde uygulama gerçekleştirilebilmektedir. Bunun yanı sıra artık değerli ağlara (ResNet) benzer şekilde atlamalı bağlantı uygulanarak derin ağın eğitim

hızı ve başarısı artırılmaya çalışılmıştır [26]. Daha sonra, üretici ağı son kısmında istenilen üst ölçekleme oranına göre alt piksel evrişim (sub-pixel convolution) katmanları kullanılmıştır. Bu çalışmada, x4 üst ölçekleme oranında süper çözünürlük uygulamaları gerçekleştirildiği için iki tane alt piksel evrişim katmanı

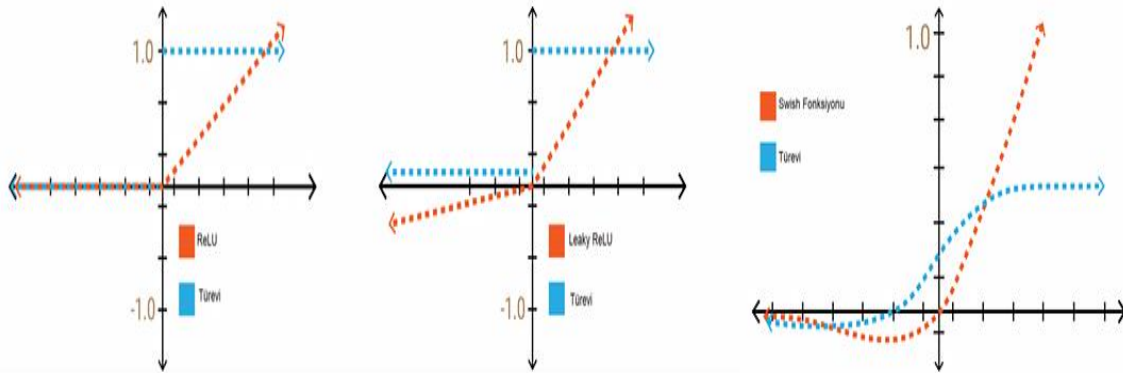
kullanılmıştır. Ayırt edici ağ kısmında ise negatif öğrenmenin daha stabil gerçekleşmesi için LeakyReLU aktivasyon fonksiyonu tercih edilmiştir. Ayrıca, batch normalizasyon katmanlarının hem üretici hem de ayırt edici ağında kullanılmasıyla gradyanın yok olması probleminin önüne geçmek amaçlanmıştır [27].



Şekil 4. Üretici ve ayırt edici ağlar (Generator and discriminator networks)

TSRGAN modelinin mimarisinde kullanılan ReLU, LeakyReLU ve swish aktivasyon fonksiyonları ve bu fonksiyonların türevleri Şekil 5'te gösterilmektedir. ReLU fonksiyonu $[0, +\infty)$ aralığında değer almaktadır. Derin ağ yapılarında hesaplama yükünün verimli olması ve ağı eğitiminin hızlı olması istenmektedir. Bu sebeple, ReLU fonksiyonunun ve türevinin negatif ekseninde 0 değerini almasından dolayı derin ağ yapılarında yaygın bir şekilde kullanılmaktadır. Bu avantajlı durumuna karşın negatif kısımda öğrenme gerçekleşmemesi en büyük dezavantajdır. LeakyReLU fonksiyonu ise $(-\infty,$

$+\infty)$ aralığında değer almaktadır. Böylece ReLU fonksiyonundaki yok olan gradyanlar bu fonksiyonda mevcut olup, negatif kısımda da öğrenme sağlanmaktadır. Son olarak swish fonksiyonu da ReLU'dan farklı olarak negatif ekseninde değer almaktadır; fakat negatif ekseninde doğrusal olmayan değer aldığı için LeakyReLU fonksiyonundan ayrılmaktadır. Yani swish fonksiyonunun girdi değeri artsa bile çıktı değeri düşebilmektedir. Bu karakteristiği sayesinde gradyanın yok olması probleminin önüne geçme konusunda büyük öneme sahiptir [28].



Şekil 5. Aktivasyon fonksiyonları ve türevleri (Activation functions and derivatives)

Önerilen derin ağ modeli, Bölüm 3'te detaylıca anlatılan uygulamalarda kullanılmıştır. Ağ modelinin hiper parametreleri olarak öğrenme oranı (learning rate) 0,0003 belirlenmiştir ve optimizasyon yöntemi olarak Adam algoritması tercih edilmiştir. Bunun dışında ağı eğitim hızını yüksek tutmak ve başarı performansını geliştirebilmek için küçük parçalar halinde eğitim süreci gerçekleştirilmiştir. Burada üretici ağına girişinde parça (patch) boyutları (16, 16, 3) iken ayırt edici ağına girişine (64, 64, 3) boyutlarında parçalar verilmektedir. Gerçekleştirilen uygulamalarda derin ağ 30.000 döngü (epoch) boyunca eğitilmiştir. Derin ağı eğitim süreci CPU E5-2680 işlemcisine ve 32 gb GeForce GTX 1080 Ti ekran kartına sahip iş istasyonu vasıtasıyla tamamlanmıştır. Bölüm 3.1' de detayları verilen uygulamalar için ağı eğitim süresi yaklaşık 112 saat sürmüştür. Gerçekleştirilen her iki çalışmanın başarı performanslarının değerlendirilmesi, görüntü kalite metrikleri olan PSNR ve SSIM değerleri kullanılarak gerçekleştirilmiştir. Değerlendirme aşamasında yüksek çözünürlüklü görüntüler referans görüntü olarak belirlenmiştir ve düşük çözünürlüklü, biküçük ve süper çözünürlüklü görüntüler birbiriyle karşılaştırılmıştır. PSNR değeri, bir görüntünün ne kadar az gürültüye sahip olduğunun bir ölçütüdür. Bu değer ne kadar yüksek olursa görüntü o oranda kalitelidir ve düşük gürültü oranına sahiptir. PSNR değeri, referans bir giriş görüntüsü ile gürültü oranı hesaplanacak çıkış görüntüsü karşılaştırılarak hesaplanmaktadır. Denklem 3'te PSNR değerinin formülü gösterilmektedir. Burada I referans giriş görüntüsünü, I_0 ise çıkış görüntüsünü temsil etmektedir [29].

$$PSNR = 20 \times \log_{10} \left(255 \sqrt{MSE(I, I_0)} \right) \quad (3)$$

SSIM değeri, iki farklı görüntüden alınan çerçeveler arasındaki yapısal benzerliği tespit eden bir görüntü kalite metriğidir. SSIM metriği farklı görüntülere ait çerçeveler arasındaki yapısal bilgideki değişimleri göz önüne almaktadır. Dolayısıyla iki farklı görüntüye ait çerçevelerde meydana gelen bozulmayı göstermektedir. Denklem 4' te SSIM değerinin formülü gösterilmiştir. Burada $\mu_R(n)$, $\mu_D(n)$ işaretleri, referans (I_R) ve bozuk (I_D) görüntü dizisinin n . çerçevesinin ortalamasını, $\sigma_R(n)$ ve $\sigma_D(n)$ ise referans (I_R) ve bozuk (I_D) görüntü dizisinin n . çerçevesinin standart sapmasını ifade etmektedir. C_1 , C_2 ise kararsızlıktan kaçınmak için kullanılan sabitlerdir [29].

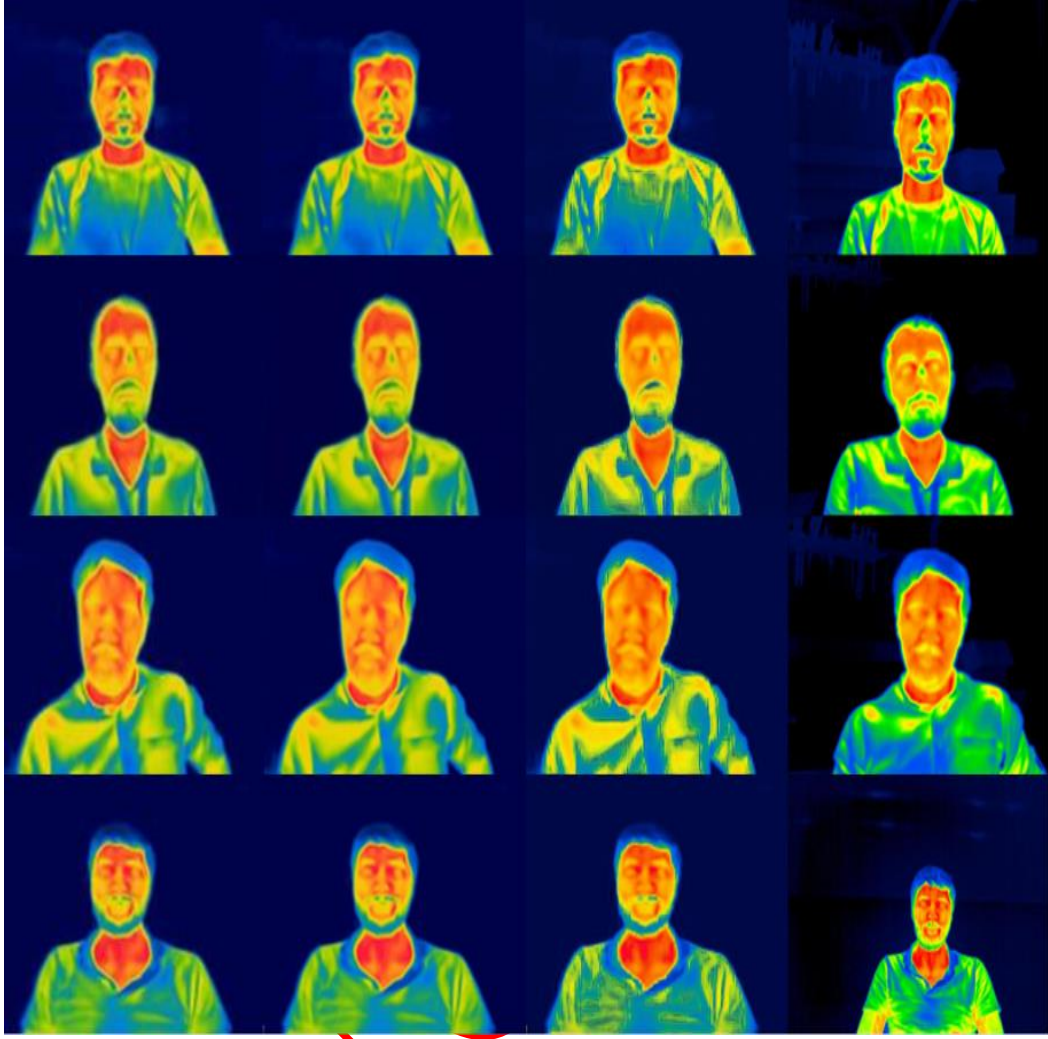
$$SSIM = \frac{[2\mu_R(n)\mu_D(n) + C_1][2\sigma_R(n)\sigma_D(n) + C_2]}{[\mu_R^2(n) + \mu_D^2(n) + C_1][\sigma_R^2(n) + \sigma_D^2(n) + C_2]} \quad (4)$$

3. GERÇEKLEŞTİRİLEN ÇALIŞMALAR (EXPERIMENTAL RESULTS)

Bu bölüm, yapılan çalışmalarla ilgili veri setlerinin kullanımı ve derin ağı eğitimi hakkında detaylı bilgiler sunmaktadır. Bölüm 3.1' de termal yüz görüntüleri üzerinde gerçekleştirilen süper çözünürlük uygulamaları ile ilgili detaylı bilgiler ve elde edilen sonuçlar yer almaktadır.

3.1. Termal Yüz Görüntüleri için Süper Çözünürlük Uygulamaları (Super Resolution Implementations for Thermal Face Images)

Bu çalışmada, daha önceden Bölüm 2.1' de bahsedilmiş olan veri setleri kullanılmıştır. Bu veri setleri 450 yüksek çözünürlüklü ve 450 düşük çözünürlüklü termal görüntü içermekte olup, toplam 900 termal görüntüden oluşmaktadır. Bu görüntülerden 400 yüksek-düşük çözünürlüklü görüntü çifti eğitim veri seti olarak, 50 yüksek-düşük çözünürlüklü görüntü çifti ise test veri seti olarak kullanılmıştır. Ayrıca eğitim ve test görüntüleri birbirinden farklı olacak şekilde düzenlenmiştir. Orijinal termal görüntülerle gerçekleştirilen uygulamada, düşük çözünürlüklü test görüntüleri 160x120 boyutuna sahiptir. Bu görüntüler önceki bölümde bahsedilen derin ağ modeli ile x4 üst ölçeklenerek 640x480 boyutlarında süper çözünürlüklü termal görüntüler elde edilmektedir. Referans görüntüler de aynı şekilde 640x480 görüntü boyutuna sahiptir. Elde edilen sonuçlardan örnek görüntüler sırasıyla düşük çözünürlüklü, biküçük interpolasyon, süper çözünürlüklü, yüksek çözünürlüklü (referans) termal görüntüler olacak şekilde Şekil 6'da gösterilmektedir. Bu çalışmaya ait PSNR/SSIM sonuçları Çizelge 2'de gösterilmiştir. Burada, düşük çözünürlüklü giriş görüntüleriyle karşılaştırıldığında biküçük interpolasyon ve SRGAN yöntemlerinin uygulanması sonucu elde edilen termal görüntülere ait PSNR değerlerinde dikkate değer bir artış gözlemlenmiştir. TSRGAN metodunun uygulanması sonucu elde edilen süper çözünürlüklü termal görüntülerin PSNR değerlerinde ise yaklaşık 0,2 dB'lik bir iyileşme görülmektedir. Ayrıca, SSIM değerlerinde sırasıyla %1,5 ve %4 oranında iyileşmeler olduğu gözlemlenmiştir.



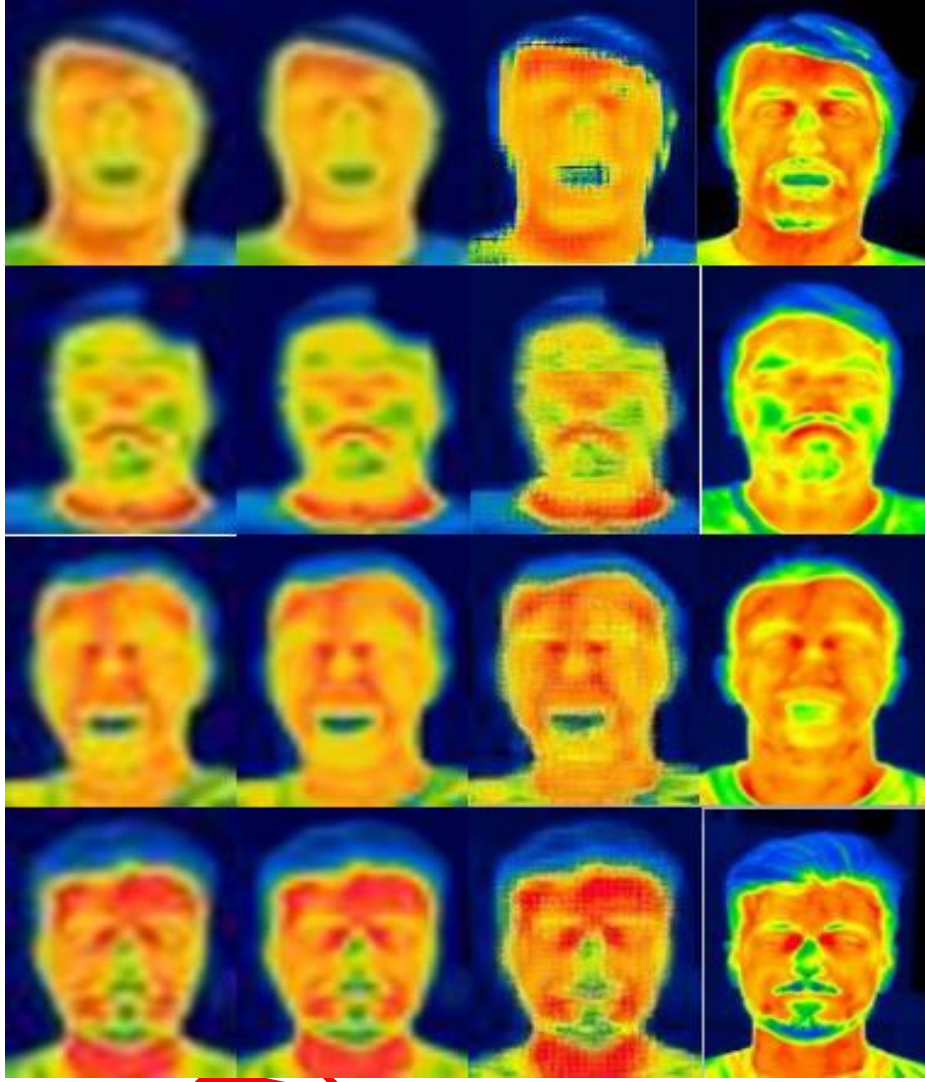
Şekil 6. Soldan sağa sırasıyla örnek yüz görüntüleri (düşük çözünürlüklü, bikübik, süper çözünürlüklü, yüksek çözünürlüklü-referans)/ (Sample images in order from left to right (low resolution, bicubic, super resolution, ground truth))

Çizelge 2. Orjinal termal yüz görüntüler için PSNR/SSIM sonuçları (PSNR/SSIM results for original thermal images)

LR		Bikübik		SRGAN		TSRGAN (önerilen)	
PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
11,041	0,315	11,086	0,328	11,104	0,334	11,217	0,352

Kırpılmış termal görüntülerle gerçekleştirilen uygulamada, düşük çözünürlüklü test görüntüleri 30x40 görüntü boyutuna sahiptir. Bu görüntüler önceki bölümde bahsedilen derin ağ modeli ile x4 üst ölçeklenerek 120x160 boyutlarında süper çözünürlüklü termal görüntüler elde edilmektedir. Yüksek çözünürlüklü (referans) görüntüler de aynı şekilde

120x160 görüntü boyutuna sahiptir. Elde edilen sonuçlardan örnek görüntüler sırasıyla düşük çözünürlüklü, bikübik interpolasyon, süper çözünürlüklü, referans termal görüntüler olacak şekilde Şekil 7'de gösterilmektedir. Elde edilen görüntüleri ait bütün sonuçlar PSNR/SSIM görüntü kalite metrikleri ile karşılaştırılmıştır ve Çizelge 3'te gösterilmektedir.



Şekil 7. Soldan sağa sırasıyla örnek yüz görüntüleri (düşük çözünürlüklü, bikübik, süper çözünürlüklü, yüksek çözünürlüklü-referans)/ (Sample images in order from left to right (low resolution, bicubic, super resolution, ground truth))

Düşük çözünürlüklü giriş görüntüleriyle karşılaştırıldığında, bikübik interpolasyon ve SRGAN yöntemlerinin uygulanması sonucu elde edilen termal görüntüleri ait PSNR değerlerinde önemli bir artış gözlemlenmemiştir. Bunun yanı sıra TSRGAN

metodunun uygulanması sonucu elde edilen süper çözünürlüklü termal görüntülerin PSNR değerlerinde yaklaşık 0,3 dB'lik iyileşme gerçekleşmiştir. Ayrıca SSIM değerlerinde sırasıyla %2 ve %5 oranlarında iyileşmeler olduğu görülmektedir.

Çizelge 3. Kırpılmış termal yüz görüntüleri için PSNR/SSIM sonuçları (PSNR/SSIM results for cropped thermal images)

LR		Bikübik		SRGAN		TSRGAN (önerilen)	
PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
13,159	0,3956	13,181	0,407	13,203	0,418	13,406	0,4455

Bu çalışmalara ek olarak, TSRGAN modelinin sonuçlarının literatürdeki benzer çalışmalarla karşılaştırılması amacıyla GT veri seti üzerinde süper çözünürlük uygulaması gerçekleştirilmiştir [20, 30]. Burada, x4 üst ölçekleme tekniği kullanılmış olup, elde

edilen görüntü kalite metriklerinin sonuçları Çizelge 4'te sunulmuştur. Önerilen TSRGAN modeli en iyi PSNR değerini elde ederken, en iyi SSIM değerini TherISuNet modeli elde etmiştir. Sonuç olarak, önerilen modelin hem standart termal veri setleri üzerinde uygulanabilirliği hem

Çizelge 4. Elde edilen PSNR/SSIM sonuçları (The obtained PSNR/SSIM results)

Veri seti	Metrik	Metot					
		Bikübik	SRResNet[16]	EDSR[15]	RCAN[31]	TherISuRNet[20]	TSRGAN (önerilen)
GT	PSNR	32,665	33,124	34,485	34,420	34,495	34,662
	SSIM	0,862	0,9018	0,906	0,907	0,910	0,903

de literatürdeki modellere karşı başarılı olduğu gösterilmiştir.

4. TARTIŞMA VE SONUÇ (DISCUSSION AND CONCLUSION)

Bu makalede, süper çözünürlük uygulamaları için çekişmeli üretici ağlar tabanlı yeni bir yaklaşım önerilerek termal yüz görüntüleri üzerinde uygulanmıştır. Bu derin ağ modeli birbiriyle ilişkili iki farklı uygulamada kullanılmıştır. İki uygulamada da geleneksel yöntemlerin dışına çıkılarak yüksek çözünürlüklü ve düşük çözünürlüklü termal yüz görüntülerinden oluşan veri seti iki farklı termal kamera kullanılarak elde edilmiştir. Veri seti oluşturulurken yüksek maliyetli olan Variocam HD termal kamera kullanılarak elde edilen termal görüntüler yüksek çözünürlüklü (referans) görüntüler, düşük maliyetli ve bir mobil cihaza kolayca entegre edilerek kullanılabilen Flir One Pro termal kamera kullanılarak elde edilen termal görüntüler düşük çözünürlüklü görüntüler olarak değerlendirilmiştir. Gerçekleştirilen uygulamaların başarı performansını görsel iyileştirmenin yanı sıra PSNR/SSIM değerleri de dikkate alınarak yorumlanmıştır. Sonuçları görsel olarak da değerlendirebilmek için her iki uygulamaya ait bazı örnek görüntüler uygulama bölümünde sunulmuştur. Elde edilen sonuç görüntüleri incelendiğinde görüntü kenar detaylarının netleşmesi ve bulanıklığın azalması gibi özellikler bakımından termal yüz görüntülerin çözünürlüğünün arttığı gözlemlenmiştir. Burada, sadece yüz bölgelerine odaklanılarak gerçekleştirilen çalışma sonuçlarının, orijinal görüntülerle gerçekleştirilen çalışma sonuçlarından daha başarılı olduğu hem PSNR/SSIM değerleri açısından hem de görsel olarak açıkça görülmektedir. Bu durum, termal görüntülerdeki gereksiz kısımların derin ağın başarısını düşürdüğünü göstermektedir. Bu sebeple, gelecekte yapılacak çalışmalarda hedef kısımlara odaklanılmasının süper çözünürlük uygulamalarının başarısını daha da artırabileceği öngörülmektedir. Ayrıca, düşük çözünürlüklü termal görüntüler, renk tonları ve yapı olarak yüksek çözünürlüklü (referans) termal görüntülere ne kadar benzer bir şekilde elde edilebilirse görüntülerdeki iyileşmenin PSNR/SSIM gibi sayısal sonuçlara yansımaları o derecede mümkün olacaktır.

Bu çalışma, düşük maliyetli termal bir kameradan elde edilen görüntülerin çözünürlüğünü, yüksek maliyetli termal bir kameradan elde edilen kaliteli görüntülerin çözünürlüğüne yaklaştırma yönünden bize umut

vermektedir. Böylece termal kameranın farklı ilgi alanlarında ve sektörlerde kullanımının yaygınlaşması için araştırmacıların termal görüntüleme uygulamalarına odaklanması hedeflenmektedir.

TEŞEKKÜR (ACKNOWLEDGEMENT)

Söz konusu çalışma KONYA TEKNİK ÜNİVERSİTESİ Bilimsel Araştırma Projeleri Birimi Koordinatörlüğü tarafından 201102001 proje numaralı “Termal Görüntü Çözünürlüğünün Artırılması İçin Derin Öğrenme Tabanlı Bulut Sisteminin Geliştirilmesi” konusu ile ilgili olup, ilgili birimce desteklenmiştir.

ETİK STANDARTLARIN BEYANI (DECLARATION OF ETHICAL STANDARDS)

Bu makalenin yazar(lar)ı çalışmalarında kullandıkları materyal ve yöntemlerin etik kurul izni ve/veya yasal-özel bir izin gerektirmediğini beyan ederler.

YAZARLARIN KATKILARI (AUTHORS' CONTRIBUTIONS)

Fatih Mehmet ŞENALP: Deneyle ilgili yapılmış, sonuçları analiz etmiş ve makalenin yazım işlemini gerçekleştirmiştir.

Murat CEYLAN: Sonuçları analiz etmiş ve makalenin yazım işlemini gerçekleştirmiştir.

ÇIKAR ÇATIŞMASI (CONFLICT OF INTEREST)

Bu çalışmada herhangi bir çıkar çatışması yoktur.

KAYNAKLAR (REFERENCES)

- [1] Zhang X., Li C., Meng Q., Liu S., Zhang Y. and Wang J., “Infrared image super resolution by combining compressive sensing and deep learning”, *Sensors (Basel)*,18(8):2587(2018).
- [2] Yue L., Shen H., Li J., Yuan Q., Zhang H. and Zhang L., “Image super-resolution:the techniques, applications, and future” *Signal Processing*,128:389-408, (2018).
- [3] Toyran M., “Reconstructing super resolution images from low resolution images”, M.Sc. Thesis, *Institute of Science*, Istanbul, (2008).
- [4] Lobanov A.P., “Resolution limits in astronomical images”, *arXiv*, preprint astro-ph/0503225, (2005).

- [5] Dong C., Loy C.C., He K., Tan X., “Image super-resolution using deep convolutional networks”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38:295-307, (2016).
- [6] Dong W., Fu F., Shi G., Cao X., Wu J., Li G. and Li X., “Hyperspectral image super-resolution via non-negative structured sparse representation”, *IEEE Transactions on Image Processing*, 25(5):2337-2352, (2016).
- [7] Guei A., Akhloufi M., “Deep learning enhancement of infrared face images using generative adversarial networks”, *Applied Optics*, 57(18): 98, (2018).
- [8] Nguyen K., Fookes C., Sridharan S., Denman S., “Feature-domain super-resolution for iris recognition”, *Computer Vision and Image Understanding*, 117(10):1526-1535, (2013).
- [9] Glasner D., Bagon S., Irani M., “Super-resolution from a single image”, *IEEE 12th International Conference on Computer Vision*, 349-356, (2009).
- [10] Lillesand T., Kiefer R. W., and Chipman J., “Remote sensing and image interpretation”, *John Wiley & Sons*, Hoboken, (2014).
- [11] Çiftçi S., Karaman M., “Landsat Uydu Görüntülerinde Derin Öğrenme Tabanlı Tek Görüntülü Süper-Çözünürlük Deneyleri”, *Harran Üniversitesi Mühendislik Dergisi*, 5(3): 194-204, (2020).
- [12] Singh K., Gupta A., Kapoor R., “Fingerprint image super-resolution via ridge orientation-based clustered coupled sparse dictionaries”, *Journal of Electronic Imaging*, 24(4):043015, (2015).
- [13] Gu Y., et al., “MedSRGAN: medical images super-resolution using generative adversarial networks”. *Multimed Tools Appl*, 79:21815–21840, (2020).
- [14] Kim J., Lee J. K., Lee K. M., “Accurate image super-resolution using very deep convolutional networks”, *IEEE CVPR*, 1646–1654, (2016).
- [15] Lim B., Son S., Kim H., Nah S., Lee K.M., “Enhanced deep residual networks for single image super-resolution” *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1132–1140, (2017).
- [16] Ledig C., et al., “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network”, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu 105-114, (2017).
- [17] Choi Y., Kim N., Hwang S., Kweon I.S., “Thermal image enhancement using convolutional neural network”, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 223–230, (2016).
- [18] Rivadeneira R., Sappa A., Vintimilla B., “Thermal Image Super-resolution: A Novel Architecture and Dataset”, *15th International Conference on Computer Vision Theory and Applications*, 111-119, (2020).
- [19] Mandanici E., Tavasci L., Corsini F., Gandolfi S., “A multi-image super-resolution algorithm applied to thermal imagery”, *Applied Geomatics*, 11(3):215–228, (2019).
- [20] Chudasama V., et al., “TherISuRNet- A computationally efficient thermal image super-resolution network”, *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, WA, USA 388-397, (2020).
- [21] Senalp F.M., Ceylan M., “Enhancement of low resolution thermal face image resolution using deep learning”, *European Journal of Science & Technology*, 131-135, (2020).
- [22] Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A. and Bengio Y., “Generative adversarial networks”, *In Advances in Neural Information Processing Systems (NIPS)*, 2672–2680, (2014).
- [23] Johnson J., Alahi A., Li F., “Perceptual losses for real-time style transfer and super resolution”, *European Conference on Computer Vision (ECCV)* Springer:694–711, (2016).
- [24] Dosovitskiy A., Brox T., “Generating images with perceptual similarity metrics based on deep networks”, *In Advances in Neural Information Processing Systems (NIPS)*, 658–666 (2016).
- [25] Anwar S., Khan S., Barnes N. A., “Deep Journey into Super-resolution: A Survey”, *ACM Computing Surveys*, 53:1-34, (2020).
- [26] Senalp F. M., Ceylan M., “Deep learning based super resolution and classification applications for neonatal thermal images”, *Traitement du Signal*, 38:5, pp. 1361-1368, (2021).
- [27] Ioffe S., Szegedy C., “Batch normalization: accelerating deep network training by reducing internal covariate shift”, *Proceedings of The 32nd International Conference on Machine Learning (ICML)*, 448–456, (2015).
- [28] Rasamoelina A. D., Adjailia F. ve Sinčák P., “A Review of Activation Function for Artificial Neural Network”, *IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMII)*, 281-286, (2020).
- [29] Javaid H., Babar T.K., Rasool A., Saghir R.U., “Video colour variation detection and motion magnification to observe subtle changes. M.Sc. Thesis”, *Blekinge Institute of Technology*, Faisalabad, Pakistan, (2013).
- [30] Senalp F.M., Ceylan M., “Effects of the deep learning-based super-resolution method on thermal image classification applications”, *Multimed Tools Appl*, (2022).
- [31] Zhang Y., Li K., Li K., Wang L., Zhong B., Fu Y., “Image super-resolution using very deep residual channel attention networks”, *Proceedings of the European Conference on Computer Vision (ECCV)*, 286–301, (2018).