
GAZİ ÜNİVERSİTESİ
Gazi
1926

SÜPER ÇÖZÜNÜRLÜK ALGORİTMALARINDA DERİN ÖĞRENMENİN ETKİSİ

Mehmet Can ÖZ

YÜKSEK LİSANS TEZİ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANA BİLİM DALI

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

TEMMUZ 2024

ETİK BEYAN

Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmasında;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Mehmet Can ÖZ

18/07/2024

SÜPER ÇÖZÜNÜRLÜK ALGORİTMALARINDA DERİN ÖĞRENMENİN ETKİSİ

(Yüksek Lisans Tezi)

Mehmet Can ÖZ

GAZİ ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

Temmuz 2024

ÖZET

Süper çözünürlük uygulamalarında renkli görüntüler için geleneksel yöntemler ve derin öğrenme tabanlı yöntemler uzun zamandır kullanılmaktadır ancak son gelişmeler bunların kullanımını termal(kızılötesi) görüntülere de genişletmiştir. Renkli görüntü sensörlerin aksine, termal sensörler genellikle düşük çözünürlüklüdür. Yüksek çözünürlüklü kızılötesi sensörlerin üretimi daha karmaşıktır. Yüksek çözünürlüklü termal sensörleri için gereken malzemeler görünür ışık sensörleri için kullanılanlara kıyasla daha maliyetlidir ve üretim süreçleri görünür ışık sensörlerine kıyasla çok daha zordur çünkü kızılötesi sensörlerin küçük sıcaklık farklarını algılayacak kadar hassas olması gerekir, bu da genellikle sisteme daha fazla gürültü sokar, dolayısıyla yüksek hassasiyeti korurken gürültüyü azaltmak, özellikle çözünürlük arttıkça çok zordur. Bu gibi kısıtlamalar da derin öğrenme yöntemleri kullanarak termal görüntülerde çözünürlüğünü artırma ihtiyacını doğurmuştur. Bu çalışmada, farklı modelleri termal görüntüler üzerine uygulayarak süper çözünürlük üzerindeki etkisini araştırdık. Özellikle, geleneksel bir yöntem olan bikübik enterpolasyonu, GAN mimarisi kullanan derin öğrenme tabanlı süper çözünürlük yöntemlerinden SRGAN (Super-Resolution Generative Adversarial Network) ve ESRGAN'ı (Enhanced Super-Resolution Generative Adversarial Network) Swin dönüştürücü (transformer) mimarisini kullanan SwinIR'ı (Swin Transformer-based Image Restoration) karşılaştırdık. Swin dönüştürücü ve GAN mimarisi kullanan bu modeller, düşük çözünürlüklü girdilerden yüksek çözünürlüklü çıktılar üreterek görüntü çözünürlüğünü artırmada umut verici sonuçlar göstermiştir. Araştırmamız, bu modellerin termal görüntülerden oluşan bir veri kümesi üzerinde eğitilmesini ve değerlendirilmesini de içermektedir. Modellerin performanslarını görüntü kalitesi, keskinlik ve termal detayların korunması açısından analiz ettik. Çalışmamızdan elde edilen bulgular, termal görüntü çözünürlüğünü iyileştirme alanına katkıda bulunmakta ve süper çözünürlük için derin öğrenme tekniklerinin etkinliği hakkında derinlik sağlamaktadır. Termal görüntüleme, askeri, tıbbi ve endüstriyel gibi çeşitli alanlarda önem kazandıkça, termal görüntülerin çözünürlüğünü iyileştirmek çok önemli hale gelmektedir. Çalışmamız, farklı yaklaşımların potansiyel faydalarına ve sınırlamalarına ışık tutarak uygulayıcılara ve araştırmacılara yol göstermektedir.

Bilim Kodu	:	93438
Anahtar Kelimeler	:	Derin öğrenme, süper çözünürlük, ölçekleme, termal görüntü, CNN, VDSR, GAN, SRGAN, ESRGAN, SwinIR
Sayfa Adedi	:	129
Danışman	:	Prof. Dr. Tuğba Selcen NAVRUZ

THE EFFECT OF DEEP LEARNING IN SUPER RESOLUTION ALGORITHMS

(M. Sc. Thesis)

Mehmet Can ÖZ

GAZİ UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

July 2024

ABSTRACT

Traditional methods and deep learning-based methods have long been used for color images in super-resolution applications, but recent developments have extended their use to thermal (infrared) images. Unlike color image sensors, thermal sensors have generally low resolution. High-resolution infrared sensors are more complex to manufacture. The materials required for high-resolution thermal sensors are more costly than those used for visible light sensors, and the manufacturing processes are much more difficult than for visible light sensors because infrared sensors need to be sensitive enough to detect small temperature differences, which usually introduces more noise into the system, so reducing noise while maintaining high sensitivity is very difficult, especially as the resolution increases. Such limitations have led to the need to improve the resolution of thermal images using deep learning methods. In this study, we investigated the impact of different models on super-resolution by applying them to thermal images. In particular, we compared bicubic interpolation, a traditional method, SRGAN (Super-Resolution Generative Adversarial Network) and ESRGAN (Enhanced Super-Resolution Generative Adversarial Network), deep learning-based super-resolution methods using GAN architecture, and SwinIR (Swin Transformer-based Image Restoration) using Swin Transformer architecture. These models using Swin Transformer and GAN architectures have shown promising results in increasing image resolution by generating high-resolution outputs from low-resolution inputs. Our research involves training and evaluating these models on a dataset of thermal images. We analysed the performance of the models in terms of image quality, sharpness and preservation of thermal details. The results of our study contribute to the field of improving the resolution of thermal images and provide insight into the effectiveness of deep learning techniques for super-resolution. As thermal imaging becomes increasingly important in various fields such as military, medical and industrial, improving the resolution of thermal images becomes crucial. Our work provides guidance to practitioners and researchers by highlighting the potential benefits and limitations of different approaches.

Science Code	:	93438
Key Words	:	Deepl earning, super resolution, scaling, thermal ımage, CNN, VDSR, GAN, SRGAN, ESRGAN, SwinIR
Page Number	:	129
Supervisor	:	Prof. Dr. Tuğba Selcen NAVRUZ

TEŞEKKÜR

Bu tezin hazırlanması sürecinde bana destek olan, yardımını ve bilgisini esirgemeyen birçok kişi ve kuruma teşekkür etmek isterim.

Öncelikle, tez danışmanım Prof. Dr. Tuğba Selcen NAVRUZ'a, her aşamada sabır ve anlayışla destek olduğu, bana rehberlik ettiği, sabırla beni yönlendirdiği ve bilimsel düşünceye olan tutkumu sürekli olarak beslediği için sonsuz teşekkürlerimi sunarım. Onun değerli önerileri ve geri bildirimleri olmadan bu çalışmayı tamamlamak mümkün olmazdı.

Bu tez çalışması boyunca bana sağladıkları sonsuz destek ve anlayış için sevgili eşime ve aileme içtenlikle teşekkür ederim. Her zaman yanımda oldular, beni teşvik ettiler ve bu yolculukta bana güç verdiler. Onların desteği olmadan bu çalışmayı tamamlamak çok daha zor olurdu.

Ayrıca, bu süreçte bana yardımcı olan ve her zaman yanımda olan arkadaşlarıma da teşekkürü bir borç bilirim. Bilimsel tartışmalarımız, fikir alışverişlerimiz ve birlikte geçirdiğimiz zamanlar, bu tez çalışmasının gelişmesine büyük katkı sağladı. Destekleriniz ve dostluğunuz için minnettarım.

Son olarak, araştırmam sırasında bana kaynak sağlayan ve çalışmamı kolaylaştıran çalıştığım şirketime yani ASELSAN ailesine teşekkürlerimi sunarım. Sağladığı imkanlar ve desteği sayesinde bu çalışmayı gerçekleştirebildim.

Herkese en içten teşekkürlerimle,

İÇİNDEKİLER

Sayfa

ÖZET iv

ABSTRACT v

TEŞEKKÜR vi

İÇİNDEKİLER vii

ÇİZELGELERİN LİSTESİ x

ŞEKİLLERİN LİSTESİ xii

SİMGELER VE KISALTMALAR xvi

1. GİRİŞ 1

2. LİTERATÜR TARAMASI 5

3. GÖRÜNTÜLEME SİSTEMLERİ 9

3.1. Gündüz Görüntüleme Sistemleri 11

3.2. Termal Görüntüleme Sistemleri 12

3.2.1. Termal kamera çalışma prensibi 13

3.2.2. Termal kamera kullanım alanları ve avantajları 14

4. SÜPER ÇÖZÜNÜRLÜĞE GENEL BAKIŞ 19

4.1. Süper Çözünürlükte Temel Kavramlar ve Uygulamaları 19

4.2. Geleneksel Süper Çözünürlük Yöntemleri 20

4.2.1. En yakın komşu enterpolasyonu 20

4.2.2. Bilineer enterpolasyon 21

4.2.3. Bikübik enterpolasyon 22

5. DERİN ÖĞRENMEYE GENEL BAKIŞ 25

5.1. Derin Öğrenmede Temel Kavramlar ve Uygulamalar 25

5.2. Derin Öğrenmenin Temelleri 29

Sayfa

5.2.1. Yapay sinir ağları 29
5.2.2. Kayıp fonksiyonları 32
5.2.3. Aktivasyon fonksiyonları 36
5.2.4. Optimize ediciler 43
5.2.5. Geriye yayılım ve gradyan inişini anlama 44
5.2.6. Veri arttırma 44
5.2.7. Yığın normalleştirme 45
6. TERMAL GÖRÜNTÜLERDE SÜPER ÇÖZÜNÜRLÜK 47
6.1. Geleneksel Süper Çözünürlük Metodlarının Termal Görüntüere Uygulanması.. 47
6.2. Derin Öğrenme Tabanlı Süper Çözünürlük Yöntemleri 51
6.2.1. VDSR 51
6.2.2. SRGAN 54
6.2.3. ESRGAN 56
6.2.4. SwinIR 58
6.2.5. Diğer yöntemler 62
7. DENEYSEL SONUÇLAR 63
7.1. Veri Seti 63
7.2. Değerlendirme Metrikleri 64
7.2.1. PSNR 64
7.2.2. SSIM 64
7.2.3. NIQE 65
7.2.4. BRISQUE 66
7.3. Model Karşılaştırmaları 67
7.3.1. Bikübik enterpolasyon 67

Sayfa

7.3.2. VDSR 74

7.3.3. SRGAN 83

7.3.4. ESRGAN 91

7.3.5. SwinIR 99

7.4. Bilineer Alt Örneklemeli Görüntülerin Karşılaştırmaları 107

7.5. Bikübik Alt Örneklemeli Görüntülerin Karşılaştırmaları 113

8. SONUÇLAR VE GELECEKTEKİ ÇALIŞMALAR 119

KAYNAKLAR 121

ÖZGEÇMİŞ 129

ÇİZELGELERİN LİSTESİ

Çizelge **Sayfa**

Çizelge 5.1. Öğrenme çeşitleri 28

Çizelge 5.2. Optimize edicilerin karşılaştırılması 44

Çizelge 7.1. Bilineer enterpolasyon ile alt örnekleme yapılmış görüntünün bikübik enterpolasyon yöntemi sonucunda orijinal görüntü ile karşılaştırılması 69

Çizelge 7.2. Bilineer enterpolasyon ile alt örnekleme yapılmış görüntünün bikübik enterpolasyon yöntemi sonucunda orijinal görüntü ile karşılaştırılması 71

Çizelge 7.3. Bikübik enterpolasyon ile alt örnekleme yapılmış görüntünün bikübik enterpolasyon yöntemi sonucunda orijinal görüntü ile karşılaştırılması 73

Çizelge 7.4. Bikübik enterpolasyon ile alt örnekleme yapılmış görüntünün bikübik enterpolasyon yöntemi sonucunda orijinal görüntü ile karşılaştırılması 74

Çizelge 7.5. Bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün VDSR sonucunda orijinal görüntü ile karşılaştırılması 77

Çizelge 7.6. Bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün VDSR sonucunda orijinal görüntü ile karşılaştırılması 79

Çizelge 7.7. Bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün VDSR sonucunda orijinal görüntü ile karşılaştırılması 81

Çizelge 7.8. Bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün VDSR sonucunda orijinal görüntü ile karşılaştırılması 83

Çizelge 7.9. Bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün SRGAN sonucunda orijinal görüntü ile karşılaştırılması 85

Çizelge 7.10. Bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün SRGAN sonucunda orijinal görüntü ile karşılaştırılması 87

Çizelge 7.11. Bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün SRGAN sonucunda orijinal görüntü ile karşılaştırılması 89

Çizelge 7.12. Bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün SRGAN sonucunda orijinal görüntü ile karşılaştırılması 91

Çizelge 7.13. Bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün ESRGAN sonucunda orijinal görüntü ile karşılaştırılması 93

Çizelge 7.14. Bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün ESRGAN sonucunda orijinal görüntü ile karşılaştırılması 95

Çizelge **Sayfa**

Çizelge 7.15. Bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün ESRGAN sonucunda orijinal görüntü ile karşılaştırılması 97

Çizelge 7.16. Bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün ESRGAN sonucunda orijinal görüntü ile karşılaştırılması 99

Çizelge 7.17. Bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün SwinIR sonucunda orijinal görüntü ile karşılaştırılması 101

Çizelge 7.18. Bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün SwinIR sonucunda orijinal görüntü ile karşılaştırılması 103

Çizelge 7.19. Bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün SwinIR sonucunda orijinal görüntü ile karşılaştırılması 105

Çizelge 7.20. Bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün SwinIR sonucunda orijinal görüntü ile karşılaştırılması 107

Çizelge 7.21. Bilineer alt örneklemeli birinci görüntünün karşılaştırılması 109

Çizelge 7.22. Bilineer alt örneklemeli ikinci görüntünün karşılaştırılması 112

Çizelge 7.23. Bikübik alt örneklemeli birinci görüntünün karşılaştırılması 115

Çizelge 7.24. Bikübik alt örneklemeli ikinci görüntünün karşılaştırılması 117

ŞEKİLLERİN LİSTESİ

Şekil **Sayfa**

Şekil 1.1. Geleneksel yöntemlerle çözünürlüğü arttırılmış örnek görüntü 3

Şekil 1.2. Derin öğrenme yöntemlerinden SRGAN mimarisi kullanılarak çözünürlüğü arttırılmış örnek görüntü 3

Şekil 3.1. Elektromanyetik tayf cetveli 9

Şekil 3.2. Termal kameraların çalıştığı dalga boyu aralıkları 10

Şekil 3.3. Gündüz görüntüleme sistemleri çalışma prensibi 12

Şekil 3.4. Gündüz görüntüleme ve termal görüntüleme karşılaştırması 13

Şekil 3.5. Termal kamera çalışma prensibi 14

Şekil 3.6. Termal kamera kullanım alanları 15

Şekil 3.7. Düşük ışık koşullarında termal kameranın kullanımı 16

Şekil 3.8. Dumanlı havada termal kameranın kullanımı 16

Şekil 4.1. En yakın komşu enterpolasyon örneği 20

Şekil 4.2. Bilineer enterpolasyon örneği 21

Şekil 4.3. Bikübik enterpolasyon örneği 22

Şekil 5.1. Tek nöronlu yapay sinir ağı 25

Şekil 5.2. Denetimli öğrenme işleyişi 26

Şekil 5.3. Denetimsiz öğrenme işleyişi 27

Şekil 5.4. Nöron yapısı 29

Şekil 5.5. Yapay sinir ağı yapısı 30

Şekil 5.6. GAN mimarisi 31

Şekil 5.7. RRDB mimarisi 32

Şekil 5.8. Sigmoid fonksiyonu ve türevi 37

Şekil 5.9. Tanh fonksiyonu ve türevi 38

Şekil **Sayfa**

Şekil 5.10. ReLU fonksiyonu ve türevi 39

Şekil 5.11. PReLU fonksiyonu ve türevi 40

Şekil 5.12. Leaky ReLU fonksiyonu ve türevi 41

Şekil 5.13. Swish fonksiyonu ve türevi 42

Şekil 6.1. Bilineer enterpolasyon tekniği kullanılarak yüksek çözünürlüklü görüntüden 4 kat alt örnekleme sonucu elde edilen düşük çözünürlüklü görüntü 46

Şekil 6.2. Bikübik enterpolasyon tekniği kullanılarak yüksek çözünürlüklü görüntüden 4 kat alt örnekleme sonucu elde edilen düşük çözünürlüklü görüntü 46

Şekil 6.3. Geleneksel süper çözünürlük yöntemlerinin bilineer enterpolasyon ile alt örnekleme yapılmış termal görüntülere uygulanması 48

Şekil 6.4. Geleneksel süper çözünürlük yöntemlerinin bikübik enterpolasyon ile alt örnekleme yapılmış termal görüntülere uygulanması 49

Şekil 6.5. VDSR mimarisi 51

Şekil 6.6. VDSR ile çözünürlüğü iyileştirilmiş termal görüntü 52

Şekil 6.7. SRGAN mimarisi 53

Şekil 6.8. SRGAN ile çözünürlüğü iyileştirilmiş termal görüntü 54

Şekil 6.9. ESRGAN üreteç ağı mimarisi 55

Şekil 6.10. Göreceli ayırt edici model 56

Şekil 6.11. ESRGAN ile çözünürlüğü iyileştirilmiş termal görüntü 57

Şekil 6.12. SwinIR mimarisi 59

Şekil 6.13. SwinIR ile çözünürlüğü iyileştirilmiş termal görüntü 61

Şekil 7.1. Termal görüntü veri seti 63

Şekil 7.2. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, bikübik enterpolasyon ile üst örnekleme uygulanmış hali 68

Şekil 7.3. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, bikübik enterpolasyon ile üst örnekleme uygulanmış hali 70

Şekil 7.4. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, bikübik enterpolasyon ile üst örnekleme uygulanmış hali 73

Şekil **Sayfa**

Şekil 7.5. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, bikübik enterpolasyon ile üst örnekleme uygulanmış hali 74

Şekil 7.6. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, VDSR ile üst örnekleme uygulanmış hali 76

Şekil 7.7. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, VDSR ile üst örnekleme uygulanmış hali 78

Şekil 7.8. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, VDSR ile üst örnekleme uygulanmış hali 80

Şekil 7.9. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, VDSR ile üst örnekleme uygulanmış hali 82

Şekil 7.10. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, SRGAN ile üst örnekleme uygulanmış hali 84

Şekil 7.11. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, SRGAN ile üst örnekleme uygulanmış hali 86

Şekil 7.12. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, SRGAN ile üst örnekleme uygulanmış hali 88

Şekil 7.13. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, SRGAN ile üst örnekleme uygulanmış hali 90

Şekil 7.14. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, ESRGAN ile üst örnekleme uygulanmış hali 92

Şekil 7.15. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, ESRGAN ile üst örnekleme uygulanmış hali 94

Şekil 7.16. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, ESRGAN ile üst örnekleme uygulanmış hali 96

Şekil 7.17. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, ESRGAN ile üst örnekleme uygulanmış hali 98

Şekil 7.18. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, SwinIR ile üst örnekleme uygulanmış hali..... 100

Şekil 7.19. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, SwinIR ile üst örnekleme uygulanmış hali..... 102

Şekil 7.20. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, SwinIR ile üst örnekleme uygulanmış hali..... 104

Şekil **Sayfa**

Şekil 7.21. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, SwinIR ile üst örnekleme uygulanmış hali..... 106

Şekil 7.22. Orijinal görüntü ile bikübik enterpolasyon, VDSR, SRGAN, ESRGAN ve SwinIR modellerine verilen düşük çözünürlüklü görüntülerin çıktıları 108

Şekil 7.23. Şekil 7.22'de her bir model için sunulan görüntülerin aynı bölgelerinden kırpılan kesitler 109

Şekil 7.24. Orijinal görüntü ile bikübik enterpolasyon, VDSR, SRGAN, ESRGAN ve SwinIR modellerine verilen düşük çözünürlüklü görüntülerin çıktıları 111

Şekil 7.25. Şekil 7.24'de her bir model için sunulan görüntülerin aynı bölgelerinden kırpılan kesitler 112

Şekil 7.26. Orijinal görüntü ile bikübik enterpolasyon, VDSR, SRGAN, ESRGAN ve SwinIR modellerine verilen düşük çözünürlüklü görüntülerin çıktıları 114

Şekil 7.27. Şekil 7.26'da her bir model için sunulan görüntülerin aynı bölgelerinden kırpılan kesitler 115

Şekil 7.28. Orijinal görüntü ile bikübik enterpolasyon, VDSR, SRGAN, ESRGAN ve SwinIR modellerine verilen düşük çözünürlüklü görüntülerin çıktıları 116

Şekil 7.29. Şekil 7.28'de her bir model için sunulan görüntülerin aynı bölgelerinden kırpılan kesitler 117

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler Açıklamalar

dB	Desibel
hz	Hertz
m^2	Metrekare
m^3	Metreküp

KısaltmalarAçıklamalar

ESRGAN	Gelişmiş süper çözünürlüklü çekişmeli üretici ağ
CNN	Evrişimsel sinir ağı
RNN	Özyinelemeli sinir ağı
RaGAN	Relativistik GAN
SVR	Destek vektör regrasyonu
SRCNN	Süper çözünürlüklü evrişimsel sinir ağı
SRGAN	Süper çözünürlüklü çekişmeli üretici ağ
SwinIR	Swin dönüştürücü ağ
VDSR	Çok derin süper çözünürlük

1. GİRİŞ

Son yıllarda görüntüleme ve yazılım teknolojilerindeki gelişmelerle birlikte özellikle savunma sanayii alanında hedef takibi, tespiti ve teşhisi uygulamaları sıklıkla kullanılmaya başlanmıştır. Bu uygulamalarda kullanılan ve görünür dalga boylarında çalışan sensörler bir ışık kaynağına ihtiyaç duydukları için olumsuz hava koşullarında veya karanlıkta görüntüyü yakalayamama problemi ile karşılaşmaktadırlar. Bu problemi çözmek için cisimlerin sahip oldukları sıcaklıklara bağlı olarak yaptıkları kızılötesi ışımayı tespit eden termal görüntüleme sistemleri geliştirilmiştir. Termal görüntülemede her ne kadar görünür bölgede ışığa ihtiyaç olmasa da termal kameraların düşük çözünürlük, yüksek gürültü ve yüksek maliyet gibi dezavantajları bulunmaktadır. Bu dezavantajları ortadan kaldırabilmek için düşük çözünürlüklü bir görüntüyü daha yüksek çözünürlüklü bir görüntüye dönüştürme veya iyileştirme süreci olarak tanımlanan süper çözünürlük yöntemleri kullanılmaktadır. Süper çözünürlük, özellikle nesne/hedef tespiti, tıbbi görüntüleme ve diğer uygulama alanlarında faydalı olmaktadır. Geleneksel süper çözünürlük algoritmaların yetersiz kaldığı yerlerde derin öğrenme tabanlı süper çözünürlük yöntemleri kullanılmaktadır.

Termal kameralar algıladıkları dalga boyu aralıklarına göre sınıflandırılır. 1-3 mikrometre arasındaki dalga boylarını algılayabilen termal kameralara Short Wave Infrared (SWIR) yani Yakın Dalga Kızılötesi, 3-5 mikrometre arasındaki dalga boylarını algılayabilen termal kameralara Mid-Wave Infrared (MWIR) yani Orta Dalga Kızılötesi, 8-14 mikrometre arasındaki dalga boylarını algılayabilen termal kameralara Long Wave Infrared (LWIR) yani Uzun Dalga Kızılötesi termal kamera denir. Savunma sanayiindeki SWIR ve MWIR termal kameralarda genellikle soğutmalı termal dedektör kullanılırken LWIR kameralarda soğutmasız termal dedektörler kullanılmaktadır. Çünkü LWIR (Uzun Dalga Kızılötesi) dalga boyu aralığı (8-14 mikrometre) genellikle ortam sıcaklığındaki nesnelerin en yoğun termal radyasyon yaydığı aralıktır. Bu nedenle, LWIR kameraların bu termal radyasyonu algılaması için yüksek hassasiyet gerekmez ve soğutmasız dedektörler yeterli performans sağlar. Ayrıca, termal kameraların çıkış görüntü çözünürlükleri, görüntü formatları ve bit derinliği gibi parametreler kameranın içindeki sensöre ve elektronik karta göre değişir. Maliyet ve sonradan işlemenin (post processing) daha kolay olması nedeniyle, birçok LWIR termal kameranın video çıkışları genellikle analog formatta 8 bit görüntüler sağlar. Buna

karşılık SWIR ve MWIR kameralar daha yüksek bit derinlikli görüntüler sağlayabilir ve sıklıkla HD-SDI gibi dijital arayüzler kullanır.

Termal kameraların en büyük dezavantajlarından birkaçı maliyet ve düşük çözünürlüktür. Bu dezavantajlar, donanım mimarilerini değiştirmeden ve daha düşük bir maliyetle, yalnızca yazılım kullanarak çözünürlüğü artırarak daha net, pürüzsüz ve ayrıntılı görüntüler sağlayan süper çözünürlük ihtiyacını doğurmuştur. Süper çözünürlük çalışmaları, geleneksel yöntemler olan en yakın komşu enterpolasyonu, bilineer enterpolasyon ve bikübik enterpolasyon ile başlamıştır. Geleneksel olan bu çalışmalarda bir görüntü oluşturmak için ara konumlardaki piksel değerleri belli bir matematiksel formüle göre tahmin edilir (Şekil 1.1). En yakın komşu enterpolasyonunda ara pikseller kendilerine en yakın dört pikselin değerini alırken, bilineer enterpolasyonda kendilerine en yakın dört pikselin ağırlıklı ortalamasını almaktadır. Bikübik enterpolasyonda ise ara pikseller kendilerine en yakın 16 pikselin ağırlıklı ortalamasını almaktadır. Geleneksel yöntemler görüntünün farklı bölgelerindeki farklı detaylar için optimum olmayan tek tip bir ölçeklendirme yaklaşımı uyguladıkları, sadece mevcut verilere dayalı matematiksel bir eşitlik ile tahmin yapabildikleri ve gerçek dünya verileri doğrusal olmayan özelliklere sahip olduğu için görüntü çözünürlüğünü ve netliğini arttırmada yetersiz kalmıştır. Bu yetersizliğin sonucunda derin öğrenme tabanlı süper çözünürlük yöntemleri ortaya çıkmıştır. Bu tezde de kullanılan Çok Derin Süper Çözünürlük Yöntemi'nde (Very Deep Super Resolution, VDSR) artık öğrenme stratejisi kullanılarak orijinal görüntü ile bikübik enterpolasyon yöntemi ile çözünürlüğü arttırılmış düşük kaliteli görüntüler arasındaki farklar öğrenilir. Öğrenilen bu farklar sayesinde VDSR modeli düşük kaliteli görüntülere süper çözünürlük uygulayarak görüntünün kalitesini arttırmayı amaçlar. GAN mimarisi kullanan SRGAN ve ESRGAN derin öğrenme tabanlı süper çözünürlük yöntemleri ise karşılıklı rekabet halinde olan üretici ve ayırt edici iki ağın karşılıklı olarak eğitilmesiyle görüntülere süper çözünürlük uygular (Şekil 1.2). Üretici ağ orijinal resme en yakın resmi üretmeye çalışırken, ayırt edici ağ ise üreticinin ürettiği görüntünün gerçek mi sahte mi olduğuna karar vererek üretici ağı gerçeğe en yakın görüntü üretmesi için teşvik eder. Tezde kullanılan diğer bir model olan SwinIR ise Swin Dönüştürücü mimarisini kullanarak sırasıyla görüntülerin yüzeysel özellik çıkarımı, derin özellik çıkarımı ve yüksek kaliteli görüntüyü yeniden oluşturma adımlarıyla görüntülere süper çözünürlük uygulayıp görüntülerin çözünürlüğünü ve kalitesini iyileştirmeyi amaçlar.

Bu tezde sunulan derin öğrenme tabanlı süper çözünürlük yöntemlerinden VDSR, SRGAN, ESRGAN ve SwinIR modelleri literatürde RGB görüntüler üzerinde geliştirilmiş ve gösterilmiştir ancak bu yöntemler RGB görüntülerindeki üç kanal yerine termal görüntülerin sahip olduğu tek kanallı girişi işleyecek şekilde uyarlanmış ve termal görüntüler üzerinde sunulmuştur.



Şekil 1.1. Geleneksel yöntemlerle çözünürlüğü arttırılmış örnek görüntü



Şekil 1.2. Derin öğrenme yöntemlerinden SRGAN mimarisi kullanılarak çözünürlüğü arttırılmış örnek görüntü

Bu tez çalışmasında, termal görüntüleme için derin öğrenme tabanlı süper çözünürlük yöntemlerinden VDSR, SRGAN, ESRGAN ve SwinIR kullanımı araştırılmıştır. Bu yöntemlerin avantajları ve dezavantajları incelenmiş, FLIR termal kamerasından alınan termal görüntü veri seti üzerinde her bir model ayrı ayrı eğitilmiş, eğitilen modeller termal görüntüler üzerine uygulanmış, termal görüntüler üzerine uygulanan modeller kendi arasında ve geleneksel süper çözünürlük yöntemleri ile PSNR, SSIM, NIQE ve BRISQUE gibi değerlendirme metrikleri sonucuna göre karşılaştırılmıştır. Bu karşılaştırma sonucunda termal görüntüler için her bir modelin avantajları ve dezavantajları tespit edilmiştir. Bu

araştırma sayesinde, termal görüntüleme sistemlerinin kalitesini ve etkinliğini artırmaya yönelik devam eden çabalara katkı sağlama amaçlanmıştır.

Bu tez çalışmasının Bölüm 2'de Literatür Taraması sunulup Bölüm 3, 4 ve 5'te Görüntüleme Sistemleri, Derin Öğrenme ve Süper Çözünürlük konularının temeli ve teorisi anlatılmıştır. Bölüm 6 ve 7'de ise derin öğrenme tabanlı süper çözünürlük yöntemlerinden bahsedilip eğitilen modeller termal görüntüler üzerine uygulanmış ve modellerin başarı metriklerine göre değerlendirilmesi ile karşılaştırılması yapılmıştır. Tez çalışmasının son bölümü olan Bölüm 8'de ise başarı metriklerine göre değerlendirilen modeller arasında termal görüntüler için her bir modelin avantajları ve dezavantajları sunulmuş ve gelecekte yapılması düşünülen çalışmalardan bahsedilmiştir.

Bu tez çalışmasında veri setindeki görüntülere bilineer ve bikübik olmak üzere iki farklı alt örnekleme yöntemi uygulanmış ve elde edilen görüntüler bikübik enterpolasyon matematiksel modeline, VDSR, SRGAN, ESRGAN ve SwinIR derin öğrenme tabanlı süper çözünürlük modellerine girdi olarak verilmiştir. Bu 5 modelin çözünürlüğü iyileştirme oranları karşılaştırılmıştır.

2. LİTERATÜR TARAMASI

Süper çözünürlük, düşük çözünürlüklü bir görüntüyü daha yüksek çözünürlüklü bir görüntüye dönüştürme veya iyileştirme sürecidir. Süper çözünürlük çalışmaları özellikle nesne/hedef tespiti, tıbbi görüntüleme vb. uygulama alanlarında kullanılır. Tıbbi görüntülemede, taramaların kalitesini artırarak daha iyi teşhis ve tedavi planlamasına yardımcı olur [1-3]. Uydu ve havadan görüntülemede süper çözünürlük, tarım ve çevresel izleme gibi uygulamalar için daha ayrıntılı ve doğru bilgiler sağlar [4-6]. Savunma sanayiinde gelişmiş çözünürlük, nesnelerin ve bireylerin daha iyi tanımlanmasına ve tanınmasına böylece tespit, teşhis ve tanıma menzillerinin geliştirilmesine yardımcı olur [7]. Makine öğrenimi modelleri için eğitim verilerinin kalitesini artırarak ve düşük çözünürlüklü senaryolarda iyi performans göstererek, süper çözünürlük yapay zekâ ve makine öğrenimindeki teknolojik gelişmeleri destekler [2, 8, 9]. Ekonomik açıdan, pahalı yüksek çözünürlüklü donanımlara uygun maliyetli bir alternatif sağlar ve sosyal açıdan, yüksek kaliteli görüntüleme teknolojisine erişimi artırır [10].

Süper çözünürlük alanı, düşük çözünürlük sorununun üstesinden gelmek için önerilen çeşitli yöntemlerle yıllar içinde önemli bir gelişme göstermiştir. Geleneksel süper-çözünürlük yöntemleri enterpolasyon tabanlı, yeniden yapılandırma tabanlı ve örnek tabanlı yaklaşımlar olarak kategorize edilebilir [11, 12]. Bikübik, bilineer, en yakın komşu, Lanczos ve spline enterpolasyonu gibi enterpolasyon tabanlı yöntemler, dijital görüntü işlemenin ilk dönemlerine kadar uzanan, görüntü yükseltme için kullanılan en eski teknikler arasındadır [13]. Bu yöntemler, düşük çözünürlüklü görüntüdeki komşu piksellerin piksel değerlerini kullanarak yüksek çözünürlüklü ayrıntıları tahmin etmeyi amaçlamaktadır [14]. Geri İteratif (Iterative Back-Projection, IBP) ve Konveks Kümelere İzdüşüm (Projection onto Convex Sets, POCS) gibi tekniklerle örneklendirilen yeniden yapılandırma tabanlı yöntemler 1980'lerin sonu ve 1990'ların başında ortaya çıkmıştır [15, 16]. Bu yöntemler süper çözünürlüğü bir ters problem olarak formüle eder ve yüksek çözünürlüklü görüntüyü gözlemlenen düşük çözünürlüklü görüntülerle ve önceki kısıtlamalarla tutarlılığa dayalı olarak iteratif olarak iyileştirir [16]. Örnek tabanlı süper-çözünürlük ve komşu gömme gibi örnek tabanlı yöntemler 1990'ların sonu ve 2000'lerin başında önem kazanmıştır [17]. Bu yöntemler, yüksek frekanslı ayrıntıların sentezlenmesini sağlayan eşleme işlevlerini veya gömülmeleri öğrenmek için düşük ve yüksek çözünürlüklü görüntü yamalarının veri

tabanlarını kullanır [18]. Geleneksel süper çözünürlük algoritmaların yetersiz kaldığı yerlerde ise derin öğrenme tabanlı süper çözünürlük yöntemleri kullanılmaktadır [15, 19].

Makine öğreniminin bir alt kümesi olan derin öğrenme, büyük miktarda veriden öğrenme ve karmaşık görevleri yerine getirme yeteneği nedeniyle çok önemli bir teknoloji haline gelmiştir. Bu yaklaşım, ham verilerden ilgili özellikleri otomatik olarak öğrenmek ve çıkarmak için çok katmanlı yapay sinir ağlarını kullanır; bu da görüntü tanıma, doğal dil işleme ve konuşma tanıma gibi görevlerde genellikle insan düzeyindeki doğruluğu aşan üstün performansa yol açmıştır [20, 21]. Derin öğrenme algoritmalarının ölçeklenebilirliği, geniş veri kümelerini işlemelerine ve güçlü GPUlar ve özel donanımlarla paralel işlemeden (birden fazla işlemi aynı anda gerçekleştirme) yararlanmalarına olanak tanıyarak onları gerçek dünya uygulamaları için pratik hale getirir [21, 22].

Derin öğrenmenin çok yönlülüğü, bilgisayarla görme, sağlık, finans, otonom sürüş ve savunma sanayii dahil olmak üzere çok sayıda alanda kendini göstermektedir. Derin öğrenme, modellerin belirli görevlere ve veri kümelerine göre ayarlanmasına olanak tanıyarak son derece özelleştirilebilir ve etkili çözümler sağlar [11, 23]. Bu esneklik, gerçek zamanlı dil çevirisi, gelişmiş robotik, gerçekçi görüntüler ve metinler oluşturan üretken modeller gibi yeni yeteneklerin geliştirilmesine yol açmıştır [24]. Evrişimsel sinir ağları (CNN), tekrarlayan sinir ağları (RNN) ve dönüştürücüler gibi mimarilerdeki sürekli ilerlemeler, yapay zekâ araştırma ve geliştirmesindeki yenilikleri daha da artırmaktadır [25, 26].

Derin öğrenme tabanlı süper çözünürlüğün ilk aşamalarında Dong ve arkadaşları 2014 yılında Süper Çözünürlük Evrişimsel Sinir Ağını (SRCNN) tanıtmıştır. Bu öncü çalışma, görüntü süper çözünürlüğü için derin öğrenme tekniklerini kullanma potansiyelini göstererek alanda önemli bir ilerlemeye işaret etmiştir [2, 6]. SRCNN modeli, düşük çözünürlüklü ve yüksek çözünürlüklü görüntüler arasında uçtan uca bir eşlemeyi doğrudan öğrenmek için üç katmanlı bir CNN kullanmış ve hem hız hem de doğruluk açısından geleneksel yöntemlerden önemli ölçüde daha iyi performans göstermiştir [2, 27].

SRCNN'nin başarısını temel alan Kim ve arkadaşları, 2016 yılında, süper çözünürlük performansını daha da ileri taşımayı hedefleyen Çok Derin Süper Çözünürlük (VDSR) ağını önermiştir. [28]. VDSR modelinde, 20 evrişimli katmandan oluşan daha derin bir ağ

kullanılmış ve bu ağların eğitilmesiyle ilgili zorlukları hafifletmek için artık öğrenme kavramı tanıtılmıştır [6, 29, 30]. Bu yaklaşım, ağın yeniden yapılandırma doğruluğunu artırma ve yakınsama hızını optimize etme yeteneği sayesinde, süper çözünürlük görevlerine yönelik yeni bir standart belirlemiştir [6, 28].

Bu gelişmelerin ardından, 2016 yılında Jiwon Kim ve arkadaşları, az sayıda parametre ile yüksek performans elde etmek için özyinelemeli katmanlardan yararlanan ve süper çözünürlük için derin ve özyinelemeli öğrenmenin etkinliğini gösteren Derin Özyinelemeli Evrişimsel Ağını (DRCN) tanıtmıştır [31]. Aynı dönemde Lim ve arkadaşları, önceki mimarilerden gereksiz modülleri kaldırıp modelin kapasitesini artıran ve yığın normalizasyonu (batch normalization) olmadan modele artık (residual) blokları dahil edip performansı iyileştiren Geliştirilmiş Derin Süper Çözünürlük (Enhanced Deep Super Resolution, EDSR) ağını önermiştir [7, 31, 32]. Bu yenilik hem eğitim verimliliğinde hem de süper çözünürlük uygulanmış görüntülerin kalitesinde önemli iyileşmeler sağlamıştır [31, 32].

Çekişmeli Üretici Ağlar (Generative Adversarial Networks, GAN) ile süper çözünürlük çalışmaları ve deneylerinin hız kazandığı 2014 yılında, Ian Goodfellow ve arkadaşları tarafından önerilen GAN modeli derin öğrenme alanında önemli bir atılım yapmıştır [8]. Bu atılım, bir üretici (generator) ve bir ayırt edici (discriminator) olmak üzere iki ağın rekabetçi bir ortamda eşzamanlı olarak eğitildiği derin öğrenme modellerinin eğitimi için yeni bir çerçeve sunarak son derece gerçekçi görüntü üretimi sağlamıştır [33]. Bu gelişmelerin ardından, 2017 yılında Christian Ledig ve arkadaşları Süper Çözünürlüklü Çekişmeli Üretici Ağlar (Super Resolution Generative Adversarial Networks, SRGAN) mimarisini önermiş ve süper çözünürlük araştırmalarına yeni bir bakış açısı getirmiştir [33, 34]. SRGAN, önceden eğitilmiş evrişimsel sinir ağlarından çıkarılan üst düzey özellik temsillerine dayalı bir algısal kayıp fonksiyonu sunmuş ve bu da üretilen yüksek çözünürlüklü görüntülerin görsel kalitesini önemli ölçüde artırmıştır [33, 34]. Ardından, 2018'de X. Wang ve arkadaşları, süper çözünürlükte daha gerçekçi ve ayrıntılı sonuçlar elde etme potansiyelini vurgulayarak GAN tabanlı Geliştirilmiş Süper Çözünürlüklü Çekişmeli Üretici Ağlar (Enhanced Super Resolution Generative Adversarial Networks, ESRGAN) mimarisini önermişlerdir [9, 35]. ESRGAN, ağ mimarisini ve eğitim kaybını daha da geliştirerek, yığın normalizasyon olmadan bir “Artık İçinde Artık Yoğun Blok” (Residual in Residual Dense Block, RRDB) ve göreceli GAN (Relativistic GAN, RaGAN) yaklaşımı dahil ederek daha ayrıntılı dokular

ve daha gerçekçi görüntü çıktıları elde etmiştir [9, 36].

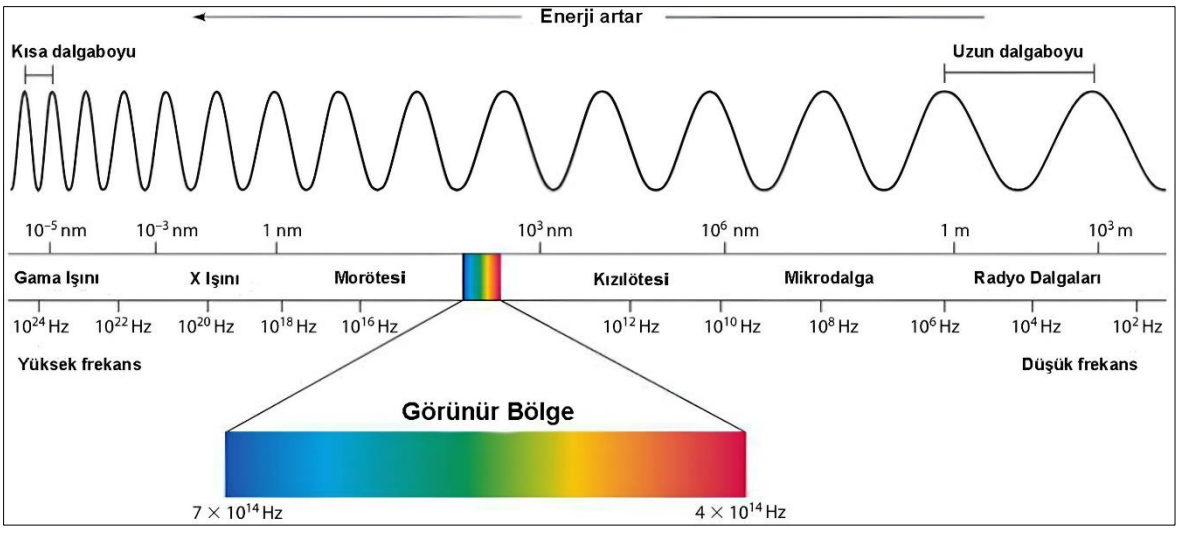
SRGAN ve ESRGAN tarafından getirilen yenilikleri temel alan araştırmacılar, süper çözünürlük yeteneklerini daha da geliştirmeyi amaçlayan dönüştürücü tabanlı mimarileri keşfetmeye devam etti. 2021 yılında SwinIR (Swin Dönüştürücü) modeli, Liang ve arkadaşları tarafından süper çözünürlük de dahil olmak üzere görüntü restorasyon görevleri için transformatörlerin gücünden yararlanan yeni bir yaklaşım olarak tanıtıldı [37]. SwinIR, hem yerel hem de küresel bağlamı etkili bir şekilde yakalamasını sağlayan hiyerarşik bir tasarıma ve kaydırılmış pencereleme (shifted window) şemasına sahip olan Swin Transformer mimarisine dayanmaktadır [10, 38]. Model, görüntüleri daha esnek ve ölçeklenebilir bir şekilde işlemek için öz dikkat (self-attention) mekanizmaları ve evrişimsel katmanların bir kombinasyonunu kullanır [10, 37, 38]. SwinIR, görüntüyü örtüşmeyen pencerelere bölerek ve her bir pencere içinde öz dikkat (self-attention) gerçekleştirerek yüksek çözünürlüklü görüntüleri verimli bir şekilde işleyebilir. Hiyerarşik yapı, modelin farklı ölçeklerdeki bilgileri bir araya getirmesine olanak tanıyarak yeniden yapılandırılan yüksek çözünürlüklü görüntülerin kalitesini önemli ölçüde artırır. SwinIR, geleneksel evrişimsel ağların sınırlamalarını ele alarak ve görüntü süper çözünürlüğünde dönüştürücü tabanlı mimarilerin potansiyelini sergileyerek çeşitli kıyaslamalarda dikkate değer bir performans göstermiştir [10, 37, 38, 39, 40].

Bu çalışmada da kullanılan FLIR veri kümesinin [41] kullanıldığı bir başka çalışmada [42], termal görüntülerin süper çözünürlük problemini çözmek için TherISuRNet olarak adlandırılan bir model de termal görüntülere uygulanmıştır.

3. GÖRÜNTÜLEME SİSTEMLERİ

Görüntüleme sistemleri, vücutların veya nesnelerin iç veya dış kısımlarının görsel temsillerini oluşturmak için kullanılan teknolojilerdir. Tıp, savunma sanayii ve endüstriyel gibi çeşitli alanlarda kullanılırlar. Görüntüleme sistemlerinde cisimlerden yayınlan ışımayı algılayan sensörler kullanılır. Görüntüleme sistemleri bu sensörlerin algıladığı dalga boylarına göre görünür bölge ve termal görüntüleme sistemleri olmak üzere sınıflandırılabilir.

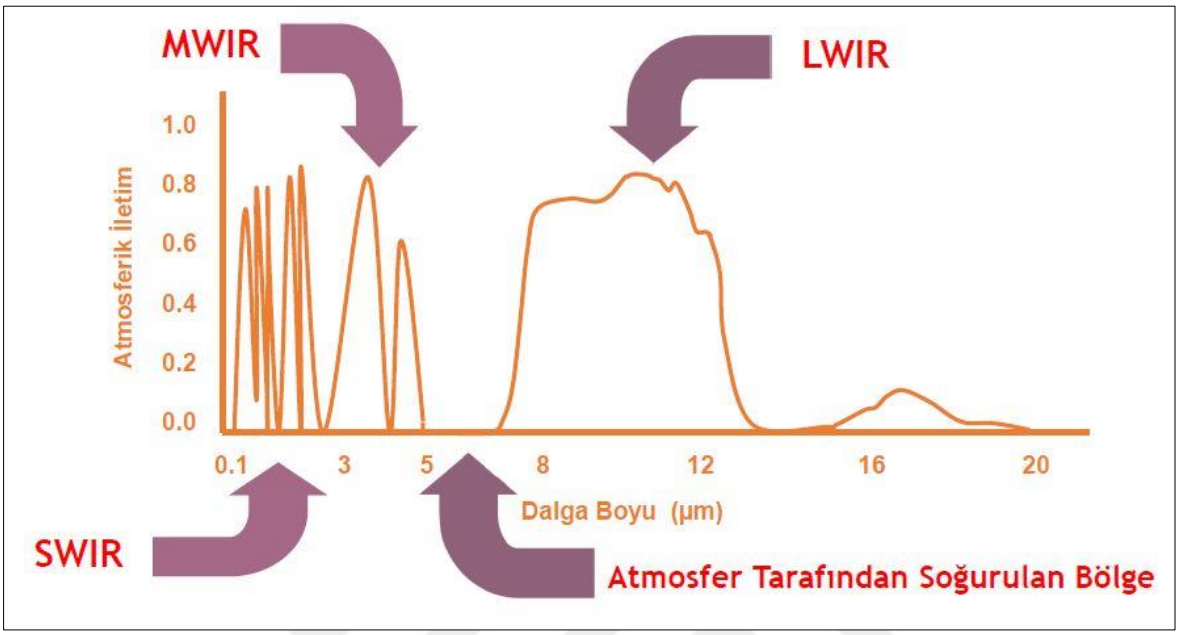
Elektromanyetik dalgalar elektrik ve manyetik alanların bileşimi olarak kabul edilir. Bu elektromanyetik dalgalar dalga formunda hareket ederler. Dalga boyu da bu elektromanyetik alanlarla bağlantılıdır. Elektromanyetik dalgaların tamamının gösterildiği cetvele Şekil 3.1'de sunulduğu üzere Elektromanyetik Tayf denir. İnsan gözünün görebileceği bölgenin hemen ötesinde bulunan ve kırmızı renkten sonra gelen kısma kızılötesi tayf denir. İnsan gözünün görebildiği 0.37 µm ile 0.72 µm arasındaki dalga boyu aralığına görünür tayf denir [43].



Şekil 3.1. Elektromanyetik tayf cetveli [44]

Kızılötesi tayf kendi arasında kısa boylu kızıötesi dalga (Short Wave Infrared, SWIR), orta boylu kızılötesi dalga (Midwave Infrared, MWIR) ve uzun boylu kızılötesi dalga (Longwave Infrared, LWIR) yani yakın, orta ve uzak dediğimiz dalga boylarına göre üçe ayrılır. SWIR kameralar genellikle 1-3 µm dalga boyu aralığında çalışırken, MWIR kameralar 3-5 µm

dalga boyu aralığında çalışır. LWIR kameralar ise genellikle 8-14 µm dalga boyu aralığında çalışır (Şekil 3.2).



Şekil 3.2. Termal kameraların çalıştığı dalga boyu aralıkları

Tüm elektromanyetik ışımalar atmosferden geçerken atmosferle etkileşim halindedir. Bu etkileşim sonucunda enerjinin bir kısmı atmosfer tarafından soğurulur ve elektromanyetik enerjinin zayıflamasına neden olur [45]. Atmosferdeki zayıflama ışıyan sinyalin dalga boyuna bağlıdır [46]. Elektromanyetik tayfın bazı bölgelerinde bulunan bantlar atmosferik pencere olarak tanımlanır [46]. Bu pencerelerde atmosferik soğurma çok düşüktür. Böyle bir pencereye örnek, tayfın 1-3 µm, 3-5 µm ve 7-14 µm dalga boyları arasında kalan bölgere verilebilir. Bu nedenle termal görüntüleme sistemleri 1-3 µm, 3-5 µm ve 8-14 µm dalga boyları arasında çalışacak şekilde tasarlanırlar. Bu bölgede çalışan sistemler, atmosferik soğurmanın az olması sebebi ile daha iyi görüntü elde ederler.

Bu bölümde “Gündüz Görüş Görüntüleme Sistemleri” ve “Termal Görüş Görüntüleme Sistemleri” anlatılmıştır.

3.1. Gündüz Görüntüleme Sistemleri

Gündüz görüntüleme sistemleri öncelikle bir ışık kaynağından, tipik olarak güneşten, nesneler üzerine gelen ışık enerjisini yakalayarak çalışır. Işığın madde ile etkileşimi bir görüntünün oluşumu için temeldir. Bu etkileşim üç ana yolla gerçekleşebilir: soğurma (emilim), iletim ve yansıma [47].

Soğurma

Işık bir nesneyle karşılaştığında, ışığın bazı dalga boyları nesne tarafından emilir. Soğurulan ışık, ısı gibi diğer enerji biçimlerine dönüştürülür. Gün ışığında algıladığımız bir nesnenin rengi büyük ölçüde emilmeyen ve bu nedenle gözlerimize ulaşan ışığın dalga boyları tarafından belirlenir [47].

İletim

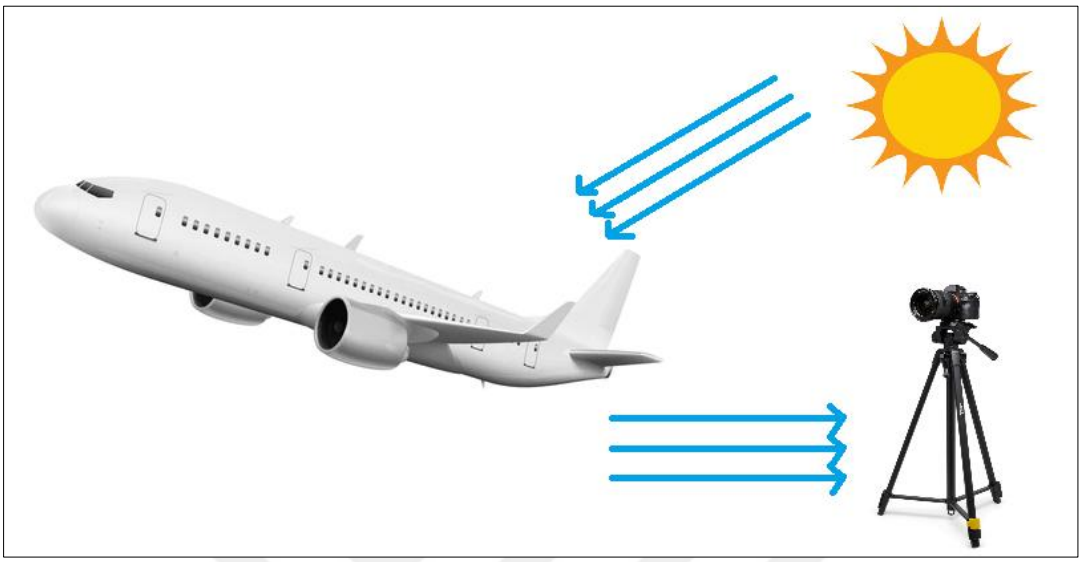
Bazı nesneler ışığın içlerinden geçmesine izin verir, bu süreç iletim olarak bilinir. Işığın geçirilme derecesi, kalınlığı ve malzemenin yapısı da dahil olmak üzere nesnenin özelliklerine bağlıdır. Örneğin, şeffaf cam görünür ışığın çoğunu geçirirken, ahşap gibi daha yoğun bir malzeme geçirmez [47].

Yansıma

Işık bir nesne tarafından da yansıtılabilir. Işık dalgaları pürüzsüz bir yüzeyle karşılaştığında, speküler yansıma olarak bilinen bir süreçle, yüzeye çarptıkları açıyla aynı açıda sekerler. Pürüzlü yüzeyler ışığı her yöne saçar. Bu durum saçılma ya da dağınık yansıma olarak adlandırılır. Yansıyan ışığın açısı ve yoğunluğu, nesnelerin şeklini ve dokusunu nasıl algıladığımız konusunda çok önemli bir rol oynar [47].

Gündüz görüntüleme sistemleri, insan gözü ile aynı temel prensiple çalışır (Şekil 3.3). Bir görüntü oluşturmak için nesnelerden yansıyan ışığı yakalar. Görüntüleme sistemindeki sensör bu yansıyan ışığı algılar ve bir elektrik sinyaline dönüştürür. Bu sinyal daha sonra sahneyi temsil eden bir görüntü üretmek için işlenir. Görüntünün kalitesi ve ayrıntısı, ışığın

yoğunluğu ve açısı, sahnedeki nesnelerin özellikleri ve görüntüleme sisteminin kendi özellikleri gibi çeşitli faktörlere bağlıdır.



Şekil 3.3. Gündüz görüntüleme sistemleri çalışma prensibi

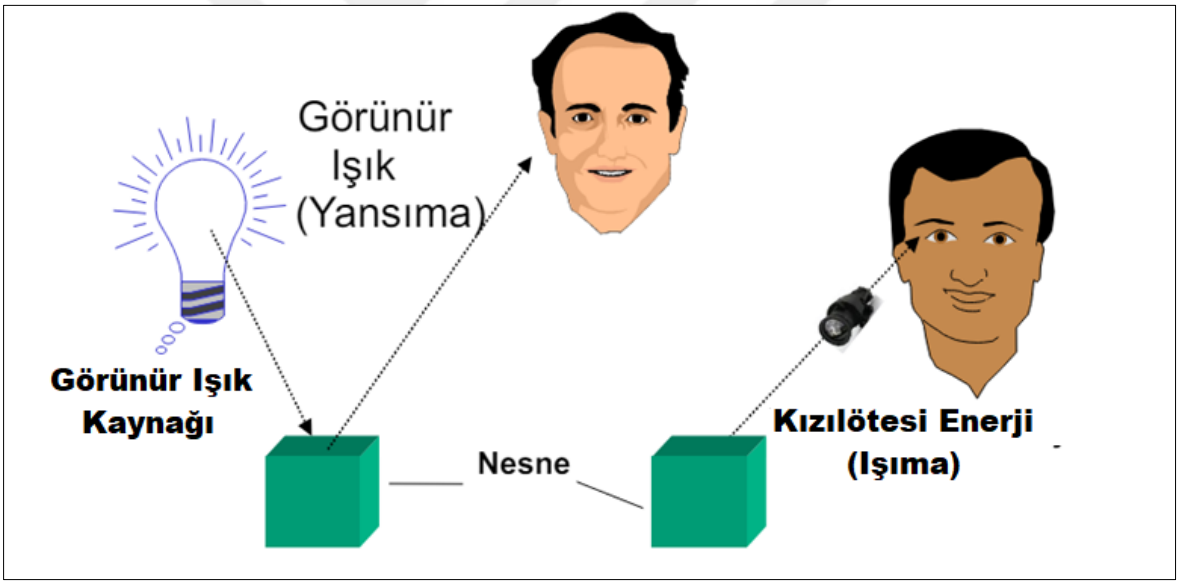
3.2. Termal Görüntüleme Sistemleri

Mutlak sıfır 0 K veya -273 C üzerinde sıcaklığa sahip tüm maddeler elektromanyetik enerji yayarlar. Yayılan bu enerji, elektromanyetik tayftaki mikrodalga, ışık ve X ışınları enerjileri gibi madde içerisindeki molekül etkileşimleri sonucu ortaya çıkan doğal ışıma ile aynı niteliktedir. Işımanın miktarı, sıcaklık arttıkça artmakta ve frekansı yükselmektedir (dalga boyu azalmakta) [48]. Kızılötesi ışıma, elektromanyetik tayfın 0,72-1000 µm arasındaki yakın, orta ve uzak kızılötesi olarak tanımlanan bölgesinde oluşmakta ve bazen ısı ışıması olarak da isimlendirilmektedir. Isı enerjisi denmesinin nedeni ise, bu ışımayı soğuran maddelerin ısınmaları gerçeğidir. Bu gerçeğe karşın kızılötesi enerjiyi, herhangi bir cismi yakabilen enerji ile karıştırmamak gereklidir. Çünkü kızılötesi enerji elektromanyetik ışımanın bir şeklidir [49].

Kızılötesi görüş genellikle termal görüş veya termal görüntüleme ile ilişkilendirilir, çünkü bir vücuttan veya nesneden gelen kızılötesi emisyonlar doğrudan sıcaklıklarıyla ilişkilidir. Daha sıcak nesneler kızılötesi spektrumda daha soğuk olanlardan daha fazla enerji yayar. Bu da günün hangi saatinde olursa olsun kızılötesi dedektörle tespit edilmelerini nispeten kolaylaştırır.

Kızılötesi enerjinin çoğu biçimi ancak kızılötesi kameralar veya gece görüş gözlükleri gibi özel ekipmanlar kullanılarak görülebilir ve ölçülebilir. Bu teknolojiler, bir nesne tarafından salınan veya yayılan ısıyı ölçmek için kızılötesi dalgaları kullanır. Bir nesnenin sıcaklığı arttıkça, atomlarının hareketi artar ve nesne daha fazla kızılötesi enerji yayar. Nesne tarafından yayılan kızılötesi enerji algılanabilir ve daha sonra görsel veya termal bir görüntü üretmek için kullanılan elektronik bir sinyale dönüştürülebilir [50].

Dolayısıyla kısaca özetleyecek olursak; gün ışığı kameralarında (Day TV) görünür ışık enerjisi bir nesneye çarpar, ondan seker, bir dedektör yansıyan ışığı alır ve sonra onu bir görüntüye dönüştürürken, kızılötesi (IR) görüş olarak da bilinen termal görüş ise Şekil 3.4'de sunulduğu üzere nesnelerin yansıttıkları görünür ışık yerine yaydıkları elektromanyetik dalga ile nesnelerin ısısını görmemizi sağlar.

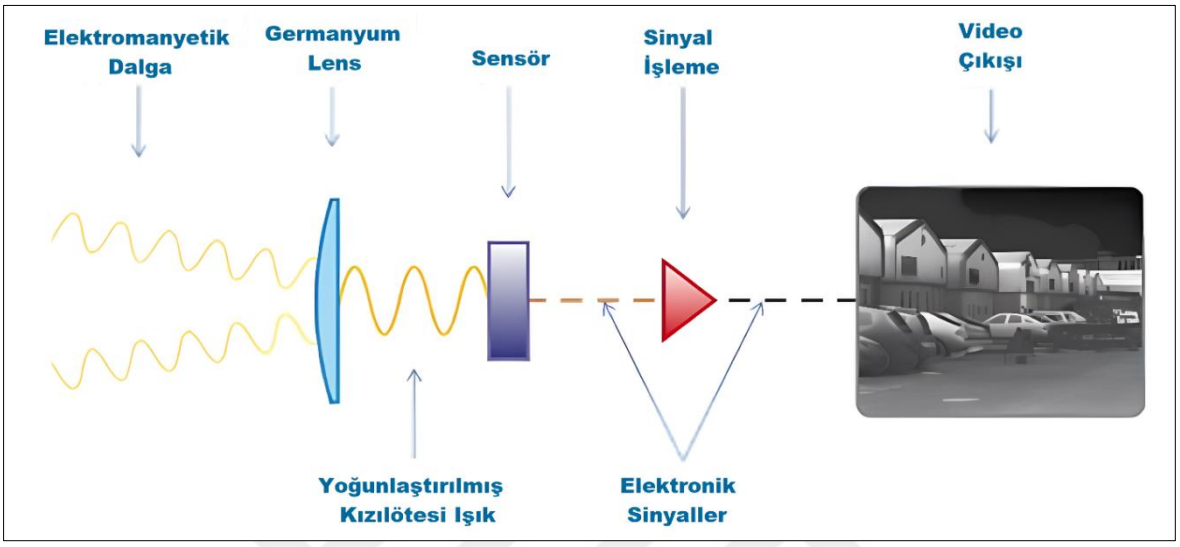


Şekil 3.4. Gündüz görüntüleme ve termal görüntüleme karşılaştırması

3.2.1. Termal kamera çalışma prensibi

Bir termal görüntüleme kamerası, kızılötesi enerjinin içinden geçmesine izin veren bir lense sahiptir. Termal kamerada bulunan özel lens elektromanyetik dalgayı geçirerek odaklı ve yoğunlaştırılmış halde dalganın sensörün üzerine düşmesini sağlar. Bu işlem sayesinde, karmaşık bir sıcaklık modeli oluşturulur. Daha sonra bu karmaşık sıcaklık modeli işlemci yardımıyla elektrik dürtülerine (impuls) ve görüntüye dönüşür (Şekil 3.5). Genellikle termogram olarak adlandırılan sonuç görüntü, farklı sıcaklıkların farklı tonlar veya renklerle

temsil edildiği gri tonlamalı veya renk kodlu bir görüntüdür [51]. Oluşturulan görüntü, yayılan kızılötesi enerji miktarına göre bir kontrast farkını gösterir. Bir nesne ne kadar sıcaksa, o kadar fazla kızılötesi radyasyon üretir.



Şekil 3.5. Termal kamera çalışma prensibi [52]

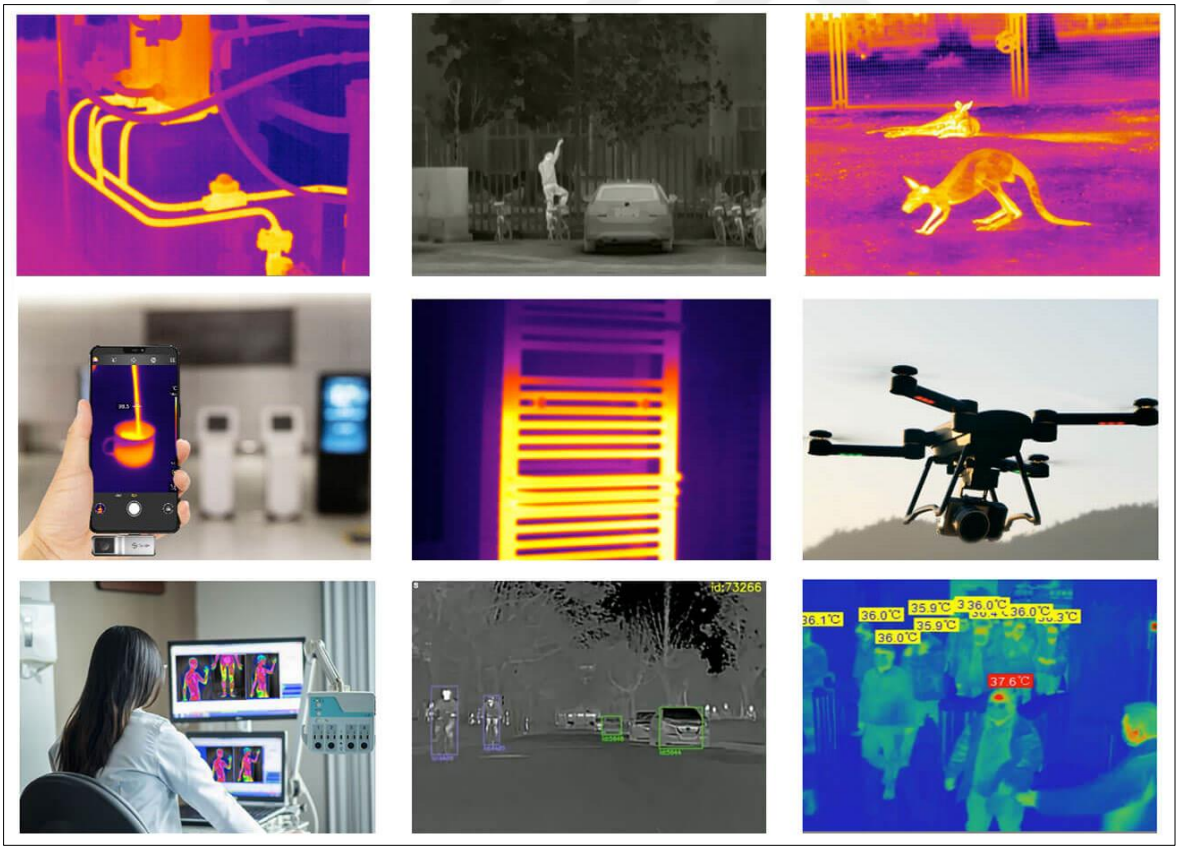
3.2.2. Termal kamera kullanım alanları ve avantajları

Termal kameralar, nesneler tarafından yayılan kızılötesi radyasyonu sıcaklıklarına bağlı olarak algılama ve görselleştirme yetenekleri nedeniyle çeşitli alanlarda geniş bir uygulama yelpazesine sahiptir. Termal kameralar durumsal farkındalığı arttırdığı hedef tespiti izleme ve keşif yeteneklerini arttırıp askeri operasyonlarda önemli bir avantaj sağladığı için savunma sanayii alanında kullanılır. Savunma sanayinde termal kameralar gözetleme ve keşif için hayati önem taşır; tamamen karanlıkta, duman, sis ve yeşilliklerin arasından personel ve araçların tespit edilmesine olanak tanıyarak durumsal farkındalığı ve hedef tespitini artırır. Ayrıca ısı izlerini takip etmek ve hassas hedefleme sağlamak için güdümlü füze sistemlerinde de kullanılırlar [51, 53].

Bina denetimleri alanında, termal kameralar ısı kaçaklarını, eksik yalıtımı ve nem biriken alanları tespit etmek için kullanılır (Şekil 3.6); bu da enerji verimliliğini artırmaya ve yapısal hasarı önlemeye yardımcı olabilir. Elektrik denetimlerinde, bu kameralar aşırı ısınan bileşenleri, hatalı bağlantıları ve elektrik yükü dengesizliklerini tespit etmek için çok önemlidir, böylece potansiyel arızaları ve yangın tehlikelerini önler [53, 54, 55].

Tıp alanında termal görüntüleme, anormal vücut ısısı modellerini tespit etmek için kullanılır (Şekil 3.6) ve enfeksiyonlar, iltihaplanma ve dolaşım sorunları gibi durumların teşhisine yardımcı olur. Güvenlik ve gözetim uygulamaları, davetsiz misafirlerin ve yetkisiz faaliyetlerin tamamen karanlıkta veya duman ve sisin içinden tespit edilmesini sağlayarak termal kameralardan faydalanır. Ayrıca, termal kameralar yangınla mücadelede yanan binalardaki kişilerin yerini tespit etmek ve yangın söndürüldükten sonra yeniden alevlenebilecek sıcak noktaları belirlemek için kullanılır [53, 54, 55].

Endüstriyel ortamlarda, termal kameralar makine ve ekipmanları izleyerek (Şekil 3.6) güvenli sıcaklık aralıklarında çalışmalarını sağlar ve böylece arıza sürelerini ve bakım maliyetlerini azaltır. Genel olarak, termal kameraların çok yönlülüğü ve doğası, verimlilik ve güvenilirliği artırarak savunma sanayii, güvenlik vb. daha birçok kritik uygulama alanlarında onları paha biçilmez araçlar haline getirmektedir [53, 54, 55].



Şekil 3.6. Termal kamera kullanım alanları [53]

Termal kameralar, özellikle diğer görüntüleme sistemlerinin yetersiz kaldığı koşullarda, birçok farklı avantajı nedeniyle sıklıkla kullanılmaktadır.

Termal kameralar, görünür ışıktan bağımsız olarak çalışıp nesnelerden yayılan elektromanyetik dalgayı algıladıkları için düşük ışık koşullarında veya tamamen karanlıkta diğer görüntüleme sistemlerine oranla Şekil 3.7'de sunulduğu üzere çok daha efektif bir şekilde çalışırlar. Gündüz görüntüleme sistemleri ile görüntülenemeyen veya ayırt edilemeyen nesneleri kolaylıkla tespit ve teşhis edebilirler [56]. Bu da onları keşif, gözetleme, arama ve kurtarma operasyonları için çok değerli kılar.



Şekil 3.7. Düşük ışık koşullarında termal kameranın kullanımı [57]

Termal kameralar, standart bir gündüz kamerasının görüşünü engelleyecek veya bloke edecek dumanlı veya sisli gibi olumsuz hava koşullarından etkilenmezler (Şekil 3.8). Bu da termal kamera kullanıcısının görüş alanını arttırır ve kullanıcıya görüş kabiliyeti kazandırır.



Şekil 3.8. Dumanlı havada termal kameranın kullanımı [58]

Termal kameralar, pasif izleme sağlar. Kendi sinyallerini yaymak yerine nesneler tarafından

yayılan kızılötesi radyasyonu tespit ettikleri için pasif olarak çalışırlar [59]. Bu pasif yapı, askeri ve savunma gibi gizliliğin gerekli olduğu uygulamalarda özellikle avantajlıdır.

4. SÜPER ÇÖZÜNÜRLÜĞE GENEL BAKIŞ

Süper çözünürlük, Bölüm 1'de açıklandığı üzere bir görüntünün çözünürlüğünü artırmayı, düşük çözünürlüklü görüntüleri yüksek çözünürlüklü benzerlerine dönüştürmeyi amaçlayan bir görüntü işleme tekniğidir. Bu süreç, yüksek kaliteli görüntülerin çok önemli olduğu tıbbi görüntüleme, uydu görüntüleri, savunma sanayii ve video geliştirme gibi çeşitli uygulamalarda hayati öneme sahiptir. Geleneksel süper çözünürlük yöntemleri enterpolasyon tekniklerine dayanır, bu da çoğu zaman ayrıntı kaybına ve bulanıklığa neden olur bu bağlamda derin öğrenmenin ortaya çıkışıyla birlikte süper çözünürlükte önemli ilerlemeler görülmüştür. Derin öğrenme modelleri, özellikle evrişimli sinir ağları (CNN), büyük veri kümeleri aracılığıyla düşük çözünürlüklü görüntülerden yüksek çözünürlüklü görüntülere kadar karmaşık eşlemeleri öğrenerek süper çözünürlükte literatüre kayda değer çalışmalar kazandırmıştır [2, 20, 28, 32, 34]. Bu modeller, karmaşık desenleri ve ayrıntıları yakalayarak üstün kalite ve netliğe sahip görüntüler üretebilmiştir. Bu bölümde, Süper Çözünürlükte Temel Kavramlar ve Uygulamaları ile Geleneksel Süper Çözünürlük Yöntemleri başlıkları sunulmuştur.

4.1. Süper Çözünürlükte Temel Kavramlar ve Uygulamaları

Süper çözünürlük, düşük çözünürlüklü girdileri yüksek çözünürlüklü çıktılara dönüştürerek görüntülerin çözünürlüğünü artırmaya odaklanan gelişmiş bir görüntü işleme tekniğidir. Bu süreç, her biri daha net, daha ayrıntılı görüntüler üretme hedefine katkıda bulunan birkaç temel kavramı içerir.

Enterpolasyon, bilinen veri noktaları aralığında yeni veri noktalarını veya değerlerini tahmin etmek için kullanılan bir yöntemdir. Geleneksel süper çözünürlük yöntemleri en yakın komşu, bilineer ve bikübik enterpolasyon yöntemleridir. Bu yöntemler, çevredeki piksellerin değerlerine göre yeni piksel değerlerini tahmin eder. Enterpolasyon basit ve hesaplama açısından verimli bir tekniktir.

Derin öğrenmenin ortaya çıkmasıyla birlikte, öğrenmeye dayalı yöntemler süper çözünürlük alanında ilerleme kaydetmiştir. Konvolüsyonel Sinir Ağları (CNN) ve Üretken Çekişmeli

Ağlar (GAN), düşük çözünürlüklü görüntülerden yüksek çözünürlüklü görüntülere karmaşık eşlemeleri öğrenmek için büyük veri kümeleri üzerinde eğitilebilir. Bu modeller karmaşık dokuları ve ince ayrıntıları yakalayabilmekte ve geleneksel yöntemlerden önemli ölçüde daha iyi performans gösterebilmektedir. Öğrenme tabanlı yaklaşımlar farklı görüntü türlerine uyum sağlayabilmekle beraber çok iyi sonuçlar da üretebilmektedir.

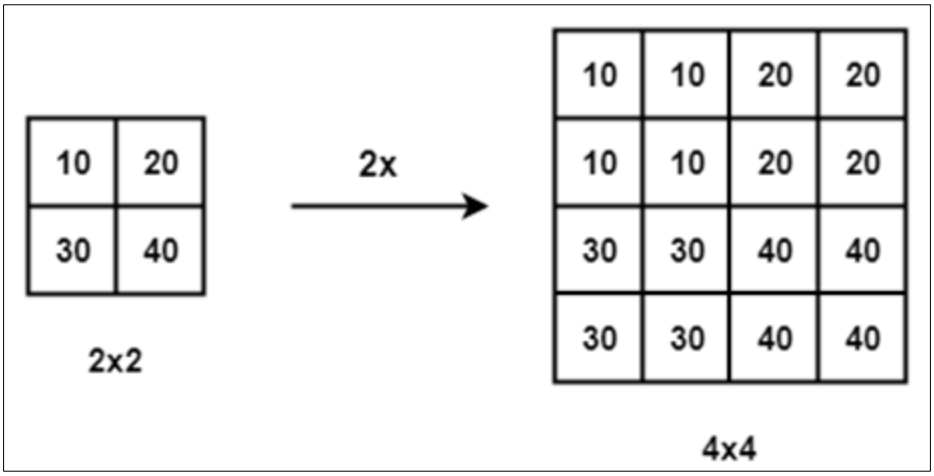
Bu kavramlar süper çözünürlük tekniklerinin temelini oluşturur ve her biri görüntü çözünürlüğünü artırma sürecine katkıda bulunur. Süper çözünürlük, çeşitli uygulama alanlarında kullanılmaktadır. Örneğin, tıbbi görüntüleme alanında, bu teknoloji görüntülerin netliğini artırarak daha detaylı incelemeler yapılabilmesini sağlar. Savunma sanayinde, süper çözünürlük kameraların çözünürlüğünü artırır ve böylece daha hassas tespitler yapabilme yeteneği kazandırır. Ayrıca, coğrafi verilerin ayrıntılarını ve doğruluğunu artırarak daha kesin analizler yapılmasına olanak sağlar.

4.2. Geleneksel Süper Çözünürlük Yöntemleri

Bu bölümde Bilineer Enterpolasyon, Bikübik Enterpolasyon ve En Yakın Komşu Enterpolasyon alt başlıkları sunulmuştur.

4.2.1. En yakın komşu enterpolasyonu

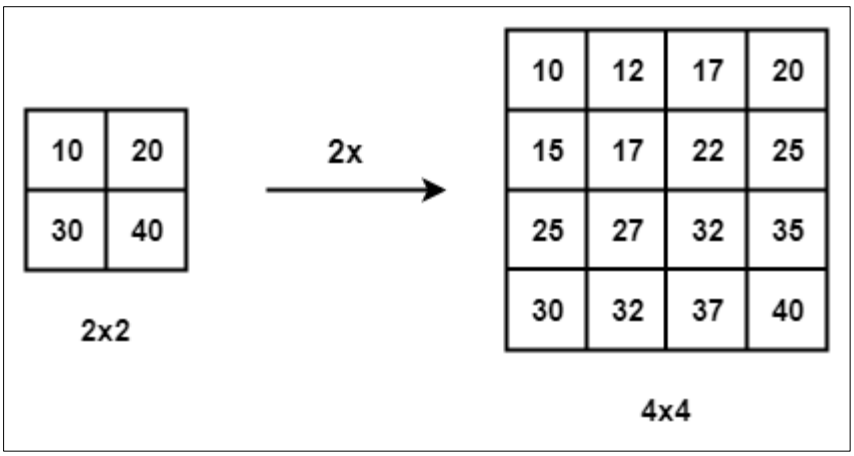
En yakın komşu enterpolasyonu, görüntü işlemede en yakın orijinal pikselin değerini yeni bir piksele atayan ilkel bir yeniden örnekleme tekniğidir (Şekil 4.1).



Şekil 4.1. En yakın komşu enterpolasyon örneği [60]
Basitliği nedeniyle minimum hesaplama kaynağı gerektirir, sadece bir pikseli dikkate alması en yakın komşu enterpolasyonunu zamana duyarlı uygulamalar için ideal hale getirir ancak bu yöntem, çevredeki piksel değerlerini göz ardı etmesi nedeniyle özellikle büyütme sırasında pürüzlü kenarlı görüntülere neden olabilir [47]. Bu yöntemin diğer bir dezavantajı da enterpolasyon sonucunda artefaktların (görüntüdeki bozulmalar) ortaya çıkabilmesidir.

4.2.2. Bilineer enterpolasyon

Bilineer enterpolasyon, bir hedef pikselin değerini belirlerken, bu hedef pikseli çevreleyen en yakın dört pikselin değerlerini kullanır. Bu değerler, ağırlıklı bir ortalama hesaplamasıyla birleştirilir (Şekil 4.2). Bu işlem, iki aşamalı bir doğrusal enterpolasyon sürecini içerir: İlk olarak yatay eksen boyunca, ardından dikey eksen boyunca enterpolasyon yapılır. Bilineer enterpolasyon, bir hedef pikselin en yakın dört komşu pikselin değerlerini kullanarak, düzgün bir şekilde değişen bir yüzey oluşturur. Bu yöntem, en yakın komşu enterpolasyonu gibi daha basit tekniklere göre daha doğru ve pürüzsüz sonuçlar verir. Bilineer enterpolasyon, görüntülerin boyutlandırılması veya döndürülmesi gibi dönüşümler sırasında görüntü kalitesinin korunmasında kritik bir rol oynar. Bir görüntü boyutlandırıldığında veya döndürüldüğünde, yeni piksel konumları genellikle tam sayı olmayan koordinatlara denk gelir. Enterpolasyon kullanılmadığında, görüntü pürüzlü ve düşük kaliteli görünebilir [47].



Şekil 4.2. Bilineer enterpolasyon örneği [61]

Piksellerin en yakın 2x2 komşuluğunu dikkate alarak en yakın komşu enterpolasyonuna kıyasla daha pürüzsüz görüntüler üretmesi, hesaplama karmaşıklığı ve görüntü kalitesi

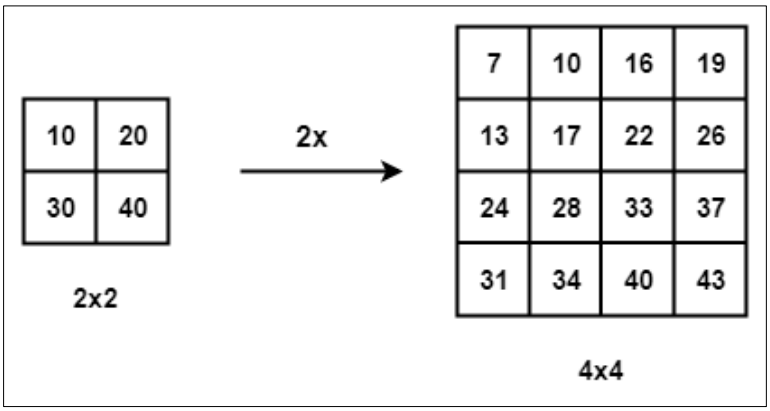
arasında iyi bir denge sunması bilineer enterpolasyonun avantajları arasında yer alırken, özellikle kenarlar boyunca fark edilebilen hafif bulanık bir görüntüye neden olabilmesi ve birden fazla pikselin dikkate alınması nedeniyle en yakın komşu enterpolasyonuna göre hesaplama açısından biraz daha yoğun olması bilineer enterpolasyonun dezavantajları arasında gösterilebilir. Süper çözünürlük yöntemi için kullanılan bilineer enterpolasyonun matematiksel ifadesi Eşitlik (4.1) 'de verilmiştir.

$$f(x,y) = f(Q_{11})(1-x)(1-y) + f(Q_{21})(1-y)x + f(Q_{12})(1-x)y + f(Q_{22})xy \quad (4.1)$$

Eşitlik 4.1'de $f(x,y)$, enterpolasyon yapılacak noktanın tahmini değeridir ve $Q_{11}, Q_{21}, Q_{12}, Q_{22}$ bilinen dört noktanın değerleri iken x ve y ise enterpolasyon yapılacak noktanın koordinatlarıdır.

4.2.3. Bikübik enterpolasyon

Bikübik enterpolasyon, komşu piksellerin değerlerini enterpole etmek için kübik polinomların kullanılmasını içeren, görüntülerin yeniden örneklenmesine yönelik gelişmiş bir yöntemdir (Şekil 4.3).



Şekil 4.3. Bikübik enterpolasyon örneği [62]

Piksellerin yalnızca en yakın 2x2 komşuluğunu dikkate alan bilineer enterpolasyonun aksine, bikübik enterpolasyon daha geniş bir 4x4 komşuluğu dikkate alır ve bu da daha yumuşak gradyanlar ve daha keskin ayrıntılar sağlar [47]. Görsel ayrıntıların kalitesini kaybetmeden görüntüleri yeniden boyutlandırmamız gerektiğinde bikübik enterpolasyon kullanılır. En yakın komşu veya bilineer enterpolasyon gibi daha basit yöntemlerle

bulanıklaşabilecek pürüzsüz kenarları ve ince ayrıntıları korumaya yardımcı olduğu için görüntüleri büyütürken özellikle kullanışlıdır. Bikübik enterpolasyonun matematiksel ifadesi Eşitlik (4.2) 'de verilmiştir.

$$f(x,y) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij} x^i \, y^j \tag{4.2}$$

Eşitlik 4.2'de x ve y enterpolasyon yapılacak noktanın koordinatlarıdır. $f(x,y)$, enterpolasyon yapılacak noktanın tahmini değeridir. a_{ij} ise x ve y'nin çevresindeki piksel değerlerinden hesaplanan katsayılardır.

5. DERİN ÖĞRENMEYE GENEL BAKIŞ

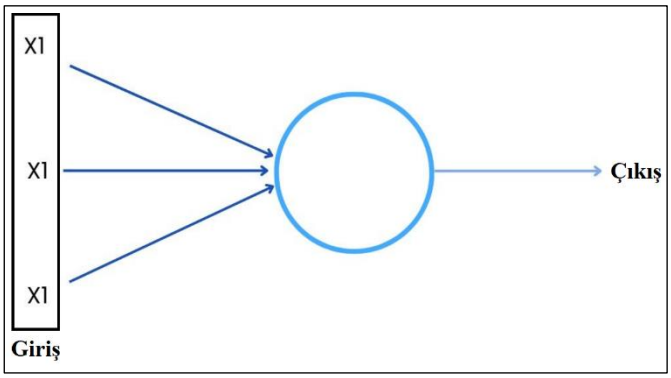
Yapay zekânın bir alt kümesi olan derin öğrenme; bilgisayarla görme, dil işleme, robotik ve görüntü işleme gibi alanlarda iş yapma şeklimizi değiştiren güçlü bir araçtır. Eski yöntemlerin aksine, derin öğrenme algoritmaları, bizim onlara ne aramaları gerektiğini söylememize gerek kalmadan ham verilerden kalıpları otomatik olarak öğrenebilir. Bu da görüntülerin çözünürlüğünü arttırma, konuşmaları anlama ve dilleri tercüme etme gibi görevlerde büyük gelişmeler sağlamıştır.

Derin öğrenme modelleri, beynimizin yapısına benzer şekilde birbirine bağlı nöron katmanlarından oluşur. Bu modeller, verilerdeki karmaşık örüntüleri bulma konusunda geleneksel yöntemlere göre oldukça başarılıdır [21].

Bu bölümde, derin öğrenmenin temelleri, derin öğrenmenin yapı taşları olan yapay sinir ağları, kayıp fonksiyonları, aktivasyon fonksiyonları ve optimize ediciler gibi derin öğrenmeyi oluşturan unsurlardan bahsedilmiştir. Bu temel bilgileri anlamak, derin öğrenmenin nasıl çalıştığını ve onu farklı alanlarda nasıl kullanabileceğini görmekte oldukça faydalıdır.

5.1. Derin Öğrenmede Temel Kavramlar ve Uygulamalar

Derin öğrenme, birkaç temel kavram üzerine kurulmuştur. Derin öğrenmenin merkezinde, insan beyninin yapısından esinlenilen hesaplama modelleri olan yapay sinir ağları yer almaktadır (Şekil 5.1).

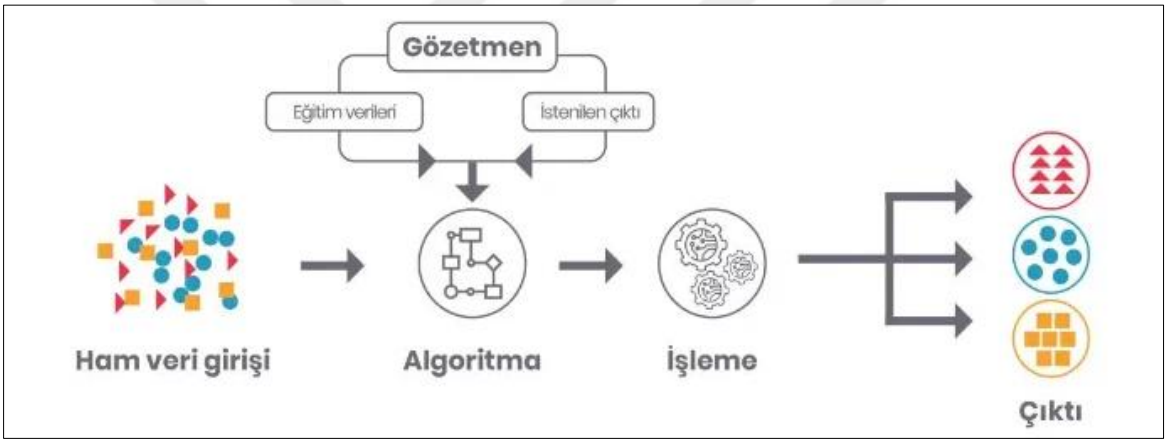


Şekil 5.1. Tek nöronlu yapay sinir ağı [63]

Bu ağlar, bilgiyi işleyen ve ileten nöron adı verilen birbirine bağlı düğüm katmanlarından oluşur. Nöronlar arasındaki bağlantıları ayarlayarak, sinir ağları kalıpları tanımayı ve verilerden tahminler yapmayı öğrenebilir [8].

Denetimli öğrenme

Derin öğrenme ve yapay zekâda yaygın bir yaklaşım, modelin etiketli örneklerden öğrendiği denetimli öğrenmedir (Şekil 5.2). Denetimli öğrenmede algoritma, girdi ve çıktı çiftlerini içeren bir veri kümesi üzerinde eğitilir. Yinelemeli optimizasyon, bir modelin hedef çıktıya ulaşmak için iteratif olarak güncellendiği bir süreçtir. Böylece yinelemeli optimizasyon, girdi ve çıktı arasındaki ilişkiyi öğrenmeyi mümkün kılar [64].



Şekil 5.2. Denetimli öğrenme işleyişi [65]

Regresyon ve sınıflandırma

Denetimli öğrenmeyi regresyon ve sınıflandırma olmak üzere iki ana başlık altında inceleyebiliriz. Regresyon, girdi özelliklerine dayalı olarak sayısal değerin tahmin edildiği bir denetimli öğrenme türüdür. Girdi değişkenleri ile çıktı değişkeni arasındaki ilişkiyi modellemeye odaklanır. Regresyondaki temel amaç, girdi uzayından çıktı uzayına bir eşleme fonksiyonu öğrenmektir, böylece tahmin edilen değerler gerçek değerlere oldukça yakın olur [66].

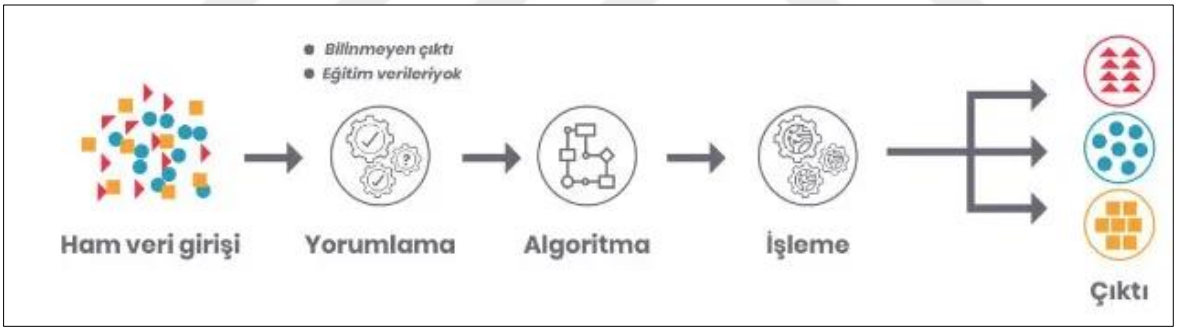
Regresyonda model, girdi-çıktı çiftleri içeren bir veri kümesi kullanarak eğitilir. Her girdi bir dizi özellik, çıktı ise bir değerdir. Model, girdiler ve çıktılar arasındaki ilişkiyi anlamayı

öğrenir, böylece yeni veriler için doğru tahminler yapabilir.

Regresyon için en sık kullanılan yaklaşımlar doğrusal regresyon, polinom regresyon ve sinir ağları olarak sıralanabilir. Regresyon modelimizin ne kadar iyi çalıştığını ölçmek için, tahmin edilen değerlerin gerçek değerlere ne kadar yakın olduğunu söyleyen Ortalama Karesel Hata (Mean Square Error, MSE) ve Ortalama Mutlak Hata (Mean Square Absolute Error, MAE) gibi ölçütleri kullanırız [66].

Denetimsiz öğrenme

Denetimli öğrenmenin aksine, denetimsiz öğrenme, açık bir yönlendirme veya önceden tanımlanmış çıktılar olmadan veri içindeki yapıları ortaya çıkarmayı amaçlayan etiketsiz veriler üzerinde bir modelin eğitilmesini içerir (Şekil 5.3). Dolayısıyla, denetimsiz öğrenme etiketli verilere ihtiyaç duymaz [67]. Denetimsiz öğrenme teknikleri genellikle kümeleme ve boyut indirgeme olacak şekilde ikiye ayrılır.



Şekil 5.3. Denetimsiz öğrenme işleyişi [68]

Kümeleme ve Boyut İndirgeme

Kümeleme algoritmaları benzer veri noktalarını özelliklerine göre bir araya getirir. Amaç, verileri aynı gruptaki noktaların birbirine olabildiğince yakın, farklı gruplardaki noktaların ise birbirinden net bir şekilde ayrıştığı şekilde gruplara ayırmaktır [68].

Boyut indirgeme, verilerin ön işlenmesinde uygulanan istatistiksel bir tekniktir. Öncelikli amacı, veriyi bilgilendirici özelliklerine indirgeyerek veri kümesini basitleştirmek ve aynı zamanda gerçek değerini korumaktır. Bu süreç, hesaplama karmaşıklığını azaltarak ve işlem

sürelerini hızlandırarak model performansını artırır [68]. Bu teknik, sonuçları en iyi tahmin eden belirgin özelliklerin altını çizerek daha odaklı bir analiz yapılmasını sağlar.

Artık öğrenme

Artık öğrenme, özellikle çok derin ağlarda eğitimi daha etkili hale getirmek için derin sinir ağlarında kullanılan bir tekniktir. Artık öğrenme, girdiden çıktıya kadar olan süreçteki tüm eşlemeleri doğrudan öğrenmeye çalışmak yerine, istenen çıktı ile giriş arasındaki farkı öğrenmeye odaklanır [69]. Artık Öğrenmede "Atlamalı Bağlantılar" (Skip Connections), bu öğrenme sürecini destekler. Bir atlamalı bağlantı, bir katmanın girişini doğrudan çıktısına ekler, böylece ağın önceden de belirtildiği üzere giriş ile hedef çıktı arasındaki farkı, yani "artığı" öğrenmesini sağlar. Bu, bilginin gradyanları kaybetmeden daha derin katmanlara aktarabilmesini sağlayarak derin sinir ağlarının daha karmaşık bilgileri öğrenebilmesine ve sinir ağlarının daha hızlı eğitilmesine yardımcı olur [69].

Artık, denetimli ve denetimsiz öğrenmenin bir karşılaştırılması Çizelge 5.1'de sunulmuştur.

Çizelge 5.1. Öğrenme çeşitleri

Öğrenme Çeşitleri	Açıklama
Artık Öğrenme	Bir artık bloğun girişini doğrudan çıkışına ekleyen ve ağın artığı (giriş ile istenen çıkış arasındaki fark) öğrenmesini sağlayan kısayol bağlantıları sunar.
Denetimli Öğrenme	Model, etiketli veriler üzerinde eğitilir ve sağlanan etiketlere dayalı olarak girdileri doğru çıktılarla eşleştirmeyi öğrenir.
Denetimsiz Öğrenme	Model etiketsiz veriler üzerinde eğitilir ve açık etiketler olmadan veri içindeki kalıpları ve yapıları tanımlamayı öğrenir.

Uygulama alanları

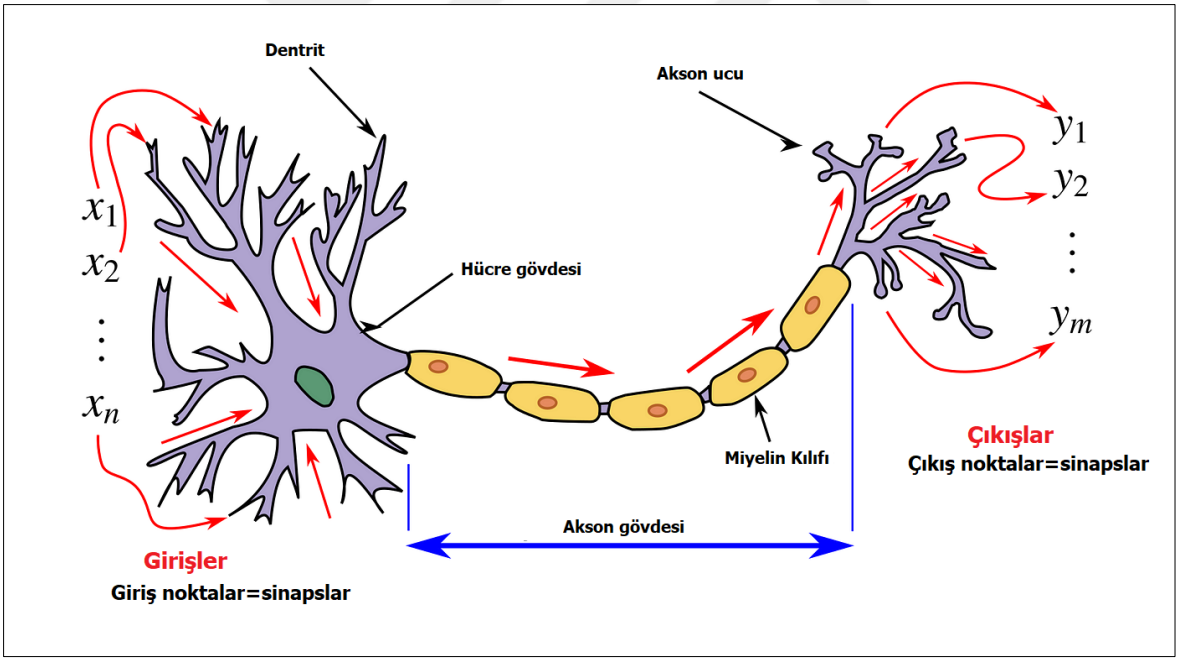
Derin öğrenme uygulamaları, derin öğrenme algoritmalarının büyük veri kümelerinden karmaşık özellikleri öğrenebilme yeteneği sayesinde çok farklı alanlarda yaygın hale gelmiştir. Bilgisayarla görmede, derin öğrenme modelleri görüntü sınıflandırma, nesne algılama ve görüntü segmentasyonu gibi görevlerde başarısını ispat etmiştir [70]. Doğal dil işlemede, derin öğrenme teknikleri dil çevirisi, duygu analizi ve metin oluşturma gibi görevler için kullanılmaktadır. Ayrıca, derin öğrenme sağlık, finans, savunma sanayii ve otonom araçlar gibi alanlarda önemli katkılarda bulunmaktadır.

5.2. Derin Öğrenme Temelleri

Bu bölümde Sinir Ağları, Kayıp Fonksiyonları, Aktivasyon Fonksiyonları ve Optimize Ediciler alt başlıkları sunulmuştur.

5.2.1. Yapay sinir ağları

Yapay sinir ağları, derin öğrenmenin temel yapı taşlarıdır. İnsan beyninin yapısı ve işlevinden esinlenen hesaplama modelleridir. Yapay sinir ağları, Şekil 5.4'te görüldüğü gibi bilgiyi işleyen ve ileten nöronlar olarak bilinen birbirine bağlı düğüm katmanlarından oluşur. Her nöron bir önceki katmandaki birden fazla nörondan girdi alır, bu girdiyi ağırlıklı bir toplam yoluyla işler ve sonucu bir sonraki katmandaki nöronlara iletmeden önce bir aktivasyon fonksiyonundan geçirir [22].



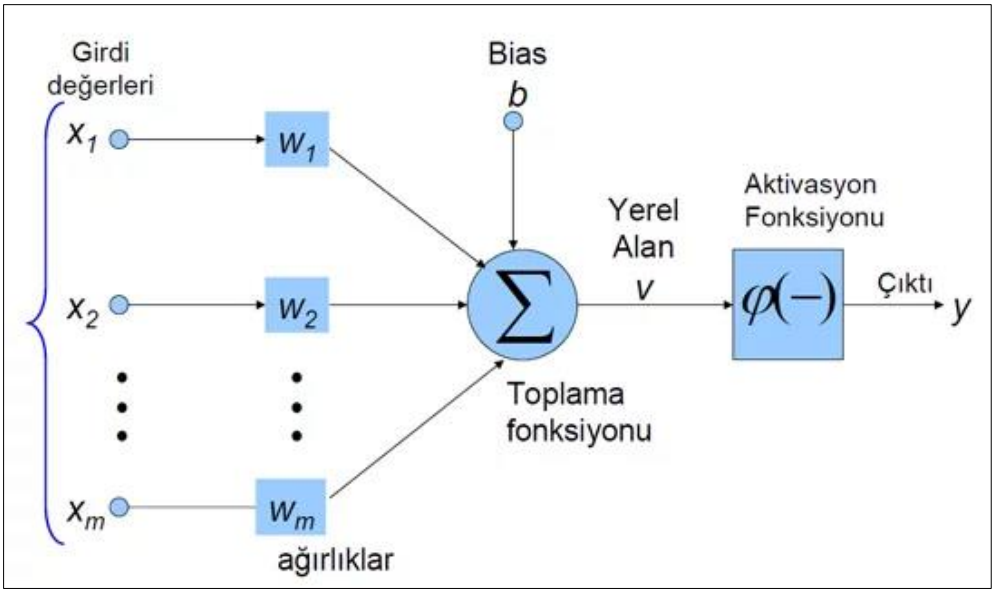
Şekil 5.4. Nöron yapısı [71]

Bir yapay sinir hücresine gelen çıktı, Şekil 5.5'de belirtildiği gibi toplanır ve bir aktivasyon fonksiyonundan geçirilerek çıktı oluşturulur. Aktivasyon fonksiyonu, çıktı değerini hesaplamak için kullanılan önemli bir bileşendir. Bir sinir ağının mimarisi tipik olarak bir giriş katmanı, bir veya daha fazla gizli katman ve bir çıkış katmanı içerir. Giriş katmanı ham verileri alır, özellikleri çıkarmak ve dönüştürmek için karmaşık hesaplamaların yapıldığı

gizli katmanlara aktarır. Çok sayıda nörondan oluşan gizli katmanlar, ağın girdi verilerinin hiyerarşik temsillerini öğrenmesini sağlar. Son çıktı katmanı tahmin veya sınıflandırma sonucunu üretir [21]. Aktivasyon fonksiyonunun girdi ve çıktıları için matematiksel ifadeler aşağıda sunulmuştur.

$$v = \sum_{i=1}^{m} w_i\, x_i + b\,, y = \phi(v) \tag{5.1}$$

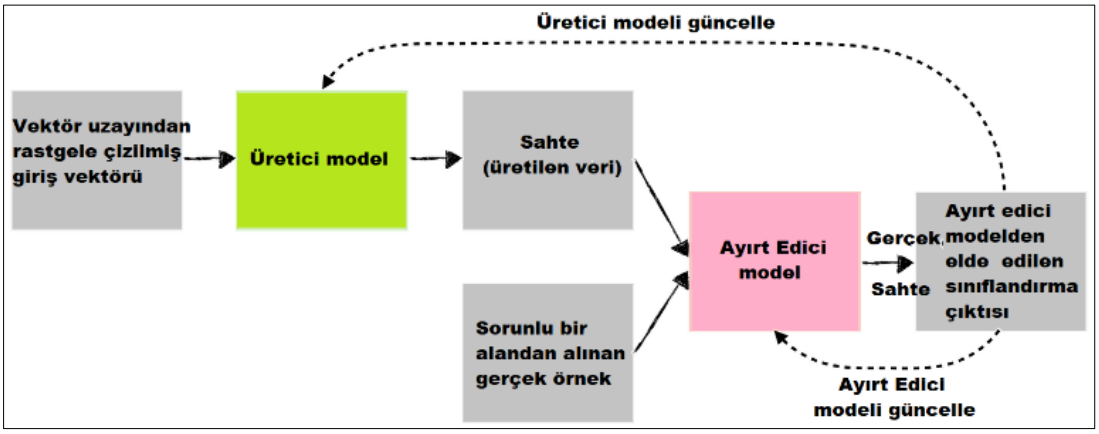
Eşitlik 5.1'deki x_i nöronun girdi özellikleridir. Her bir girdi, işlenmekte olan verinin belirli bir özelliğini veya niteliğini temsil eder. w_i her bir girdi özelliği ile ilişkili ağırlıklardır. Ağırlıklar, nöronun çıktısını tahmin etmede her bir girdinin önemini belirler ve eğitim süreci sırasında öğrenilir. b, bias değerini ifade eder. Bias, yapay sinir ağlarında bir ek parametredir ve nörona girdi olarak verilen ağırlıklı toplamın yanı sıra çıktıyı ayarlamak için kullanılır. Bu nedenle, girişe uygulanan veriye en iyi şekilde uyan bir model oluşturmak için yardımcı olan bir sabittir. $\phi(v)$, aktivasyon fonksiyonudur. Toplam girdi değeri olan v'ye doğrusal olmayan bir dönüşüm uygular. Sinir ağının verilerdeki karmaşık örüntüleri yakalamasını ve modellemesini sağlar. Yani, sinir ağının içindeki her bir nöronun çıktısını hesaplarken bu fonksiyon kullanılır. y ise ağırlıklı toplama aktivasyon fonksiyonu uygulandıktan sonra nöronun çıktısıdır. Çok katmanlı bir ağda başka bir nörona girdi olarak veya tek katmanlı bir ağda nihai çıktı olarak kullanılabilir [22].



Şekil 5.5. Yapay sinir ağı yapısı [72]

Sinir ağları, eğitim adı verilen bir süreç aracılığıyla nöronlar arasındaki bağlantıların ağırlıklarını ayarlayarak öğrenir. Eğitim sırasında ağ büyük bir veri kümesine maruz bırakılır ve tahminleri ile gerçek hedef değerler arasındaki hatayı en aza indirmek için ağırlıkları yinelemeli olarak güncellenir. Bu optimizasyon süreci tipik olarak, ağırlıklara göre kayıp fonksiyonunun gradyanını hesaplayan ve ağırlıkları hatayı azaltan bir yönde güncellemek için kullanan geri yayılım gibi algoritmalar kullanılarak gerçekleştirilir [11].

Her biri farklı görev türleri için uygun olan çeşitli sinir ağı türleri vardır. Düğümler arasındaki bağlantıların döngü oluşturmadığı ileri beslemeli sinir ağları en basit türdür ve genellikle regresyon ve sınıflandırma gibi görevler için kullanılır. Evrişimsel Sinir Ağları (Convolutional Neural Networks, CNN), görüntüler gibi matris benzeri verileri işleyen bir sinir ağı türüdür. Evrişimsel Sinir Ağları görsel örüntüleri tanımlayabildikleri için görüntü sınıflandırma ve nesne algılama gibi görevlerde daha iyidirler. Tekrarlayan Sinir Ağları (Recurrent Neural Networks, RNN) sıralı veriler için tasarlanmıştır ve zaman içinde bilgileri hatırlayabildikleri için dil modelleme ve makine çevirisi gibi görevlerde kullanılırlar. Üretken Çekişmeli Ağlar (Generative Adversarial Networks, GAN), görüntüler gibi gerçekçi veri örnekleri oluşturmak için birbirlerine karşı çalışan iki ağdan (üreteç ve ayırt edici) oluşur (Şekil 5.6). Bu yapı, Üretken Çekişmeli Ağları yeni görüntüler üretme ve verileri artırma konusunda iyileştirme sağlar [9]. Bu farklı sinir ağı mimarileri, derin öğrenmeyi birçok alanda kullanışlı hale getirir.



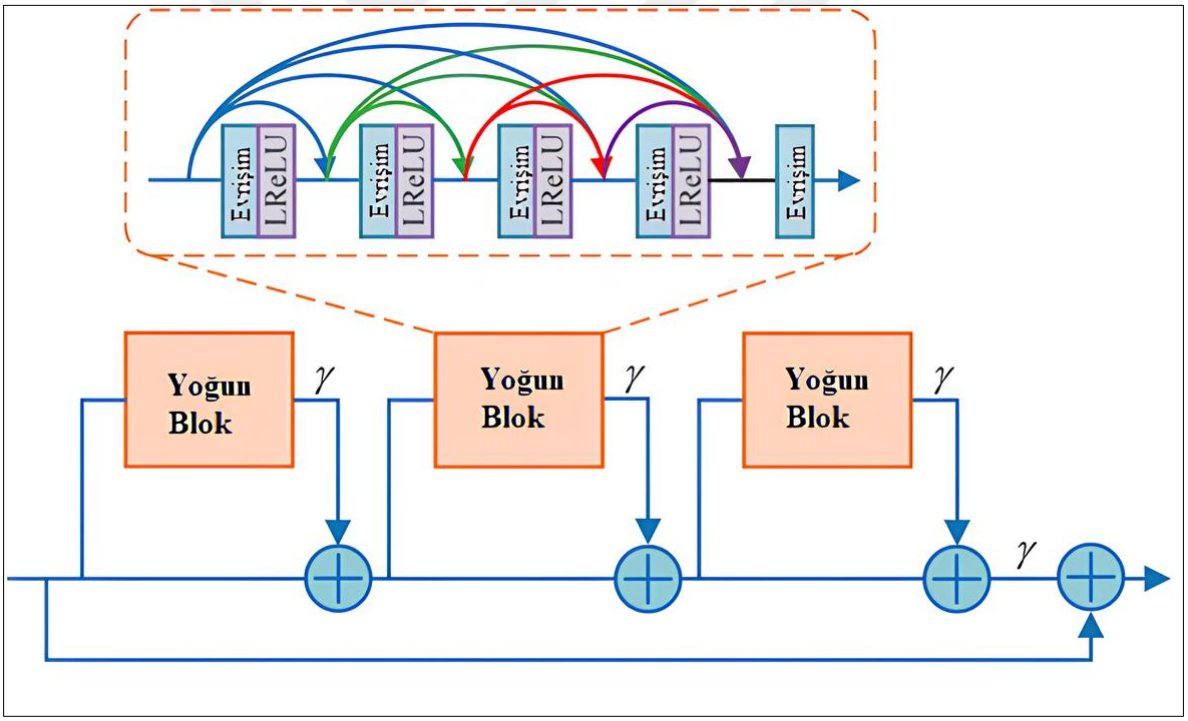
Şekil 5.6. GAN mimarisi

"Artık Blok içinde Artık Yapı" (Residual in Residual Dense Block, RRDB), derin öğrenme mimarilerinde, özellikle de görüntü süper çözünürlüğü ve restorasyonu gibi görevlerde

kullanılan özel bir yapıdır. Her blok içinde artık öğrenme ve yoğun bağlantıları entegre ederek evrişimli sinir ağlarının yeteneklerini geliştirir [73].

Bu sinir ağında, birden fazla katman yoğun bağlantılı bir şekilde düzenlenir ve her katman sırayla eklenen artık özellikleri hesaplayarak çalışır. Bu yoğun bağlantı, derin modelleri eğitmek için önemli olan gradyanların ağ boyunca etkili bir şekilde korunmasına ve yayılmasına yardımcı olur [73]. RRDB Şekil 5.7'de sunulduğu üzere hem blok düzeyinde hem de katman düzeyinde artık bağlantıları içerir.

Bu, her katmanın yalnızca kalıntıları hesaplamakla kalmadığı, aynı zamanda kalıntıların blok içindeki birden fazla katman boyunca toplandığı anlamına gelir. Bu mimari, hesaplama verimliliğini korurken ağın karmaşık eşlemeleri öğrenme yeteneğini geliştirir.



Şekil 5.7. RRDB mimarisi [74]

5.2.2. Kayıp fonksiyonları

Kayıp fonksiyonlarının bir sinir ağının öngörülen çıktısı ile gerçek hedef çıktı arasındaki farkı ölçmek için kullanılan matematiksel fonksiyonlardır. Kayıp fonksiyonları, optimizasyon sürecini gelişmiş model performansına doğru yönlendirir. Dolayısıyla, kayıp

fonksiyonu seçimi modelin performansını ve yakınsamasını önemli ölçüde etkilemektedir.

Burada Kayıp Fonksiyonları iki kategoriye ayrılarak Regresyon Kayıpları ve Sınıflandırma Kayıpları olarak sunulmuştur.

Regresyon Kayıpları

Regresyon, Bölüm 5.1'de anlatıldığı üzere iki veya daha fazla nicel değişken arasındaki ilişkinin belirlenmesini sağlayan yaygın bir analiz yöntemidir.

Regresyon analizi, bağımlı ve bağımsız değişkenler arasındaki ilişkiyi modellemek için kullanılır. Bağımlı değişken, tahmin yapmak istediğimiz değeri temsil ederken, bağımsız değişkenler bu tahmin üzerinde etkisi olan faktörleri ifade eder. Bu modellemeler esnasında bazı regresyon problemleri ortaya çıkmaktadır ve bu problemleri ele almak için spesifik kayıp fonksiyonları geliştirilmiştir [75].

Ortalama Mutlak Hata (Mean Absolute Error, MAE) tahmin edilen değerler ile gerçek değerler arasındaki hataların ortalama büyüklüğünü ölçer. Başka bir deyişle, tahmin edilen ve gerçek değerler arasındaki mutlak farkları hesaplar ve ardından bu farkların ortalamasını alır. MAE fonksiyonu aşağıda verilmiştir [22].

$$Ortalama\ Mutlak\ Hatası = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{5.2}$$

Eşitlik 5.2'de n yüksek çözünürlüklü görüntüdeki piksel sayısını ifade ederken y_i yüksek çözünürlüklü görüntüdeki i konumundaki piksel değerini ifade etmektedir. $\hat{y}_i$ süper çözünürlük ile üretilen görüntünün i konumundaki piksel değerini, $|y_i - \hat{y}_i|$ ise i konumundaki gerçek ve öngörülen piksel değerleri arasındaki mutlak farkı belirtmektedir.

Ortalama Kare Hataların (Mean Squared Error, MSE) karelerinin ortalamasını, yani tahmin edilen değerler ile gerçek değerler arasındaki ortalama karesel farkı ölçer. MSE aşağıdaki şekilde tanımlanır [22].

$$Ortalama\ Kare\ Hatası = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{5.3}$$

Eşitlik 5.3'de n yüksek çözünürlüklü görüntüdeki piksel sayısını ifade ederken y_i yüksek çözünürlüklü görüntünün i konumundaki piksel değerini ifade etmektedir. $\hat{y}_i$ süper çözünürlük ile üretilen görüntünün i konumundaki piksel değerini, $(y_i - \hat{y}_i)^2$ ise i konumundaki gerçek ve öngörülen piksel değerleri arasındaki farkın karesi olan karesel hatayı ifade eder.

Ortalama sapma hatası (Mean Bias Error, MBE), verilerin tahmin edilen değerleri ile gerçek değerler arasındaki ortalama farkı hesaplar ve tahminlerin sistematik olarak gerçek değerlerden daha yüksek veya daha düşük olma eğiliminde olup olmadığını gösterir. Verilerin gerçek ve tahminsel değer sonuçları arasındaki farkların ortalama toplamları alınarak bu hata hesaplanır. MBE fonksiyonu aşağıda verilmiştir [22].

$$\text{Ortalama Sapma Hatası} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \tag{5.4}$$

Eşitlik 5.4'de n yüksek çözünürlüklü görüntüdeki piksel sayısını ifade ederken y_i yüksek çözünürlüklü görüntünün i konumundaki piksel değerini ifade etmektedir. $\hat{y}_i$ süper çözünürlük ile üretilen görüntünün i konumundaki piksel değerini, $(\hat{y}_i - y_i)$ ise i konumundaki gerçek ve öngörülen piksel değerleri arasındaki farkı ifade etmektedir.

Ortalama Kare Logaritmik Hatası (Mean Squared Logarithmic Error, MSLE), tahmin edilen değerlerin logaritmaları ile gerçek değerlerin logaritmaları arasındaki ortalama karesel farkı ölçer ve aşağıdaki gibi ifade edilir [22].

$$\text{Ortalama Kare Logaritmik Hatası } = \frac{1}{n}\sum_{i=1}^{n}(log_e(1 + \hat{y}_i) - log_e(1 + y_i))^2 \tag{5.5}$$

Eşitlik 5.5'de n yüksek çözünürlüklü görüntüdeki piksel sayısını ifade ederken y_i yüksek çözünürlüklü görüntünün i konumundaki piksel değerini ifade etmektedir. $\hat{y}_i$ süper çözünürlük ile üretilen görüntünün i konumundaki piksel değerini, $log_e(1 + \hat{y}_i)$ öngörülen piksel değerinin doğal logaritmasını $log_e(1 + y_i)$ ise orijinal görüntüdeki piksel değerini ifade etmektedir.

Sınıflandırma kayıpları

Sınıflandırma, genel bir veri setindeki verileri yapısal veya karakteristik benzerliklerine göre ortak kategorilerde gruplama işlemidir. Bu süreçte, belirli özelliklere dayalı olarak analiz edilen her bir gözlem için uygun kategoriyi belirlemek üzere sınıflandırma problemleri ortaya çıkar. Bu problemler farklı sınıf çıktılarının tahmin edilmesini sağlar. Sınıflandırma, genel veri setlerinin çeşitli kritik parametreler kullanılarak belirli kategorilere ayrılmasını içerir. Sınıflandırma kaybı fonksiyonları, amacın girdileri ayrı kategorilere sınıflandırmak olduğu görevler için özel olarak tasarlanmıştır. Bu fonksiyonlar, sınıflandırma problemlerinde tahmin doğruluğunun veya yanlışlığının doğrudan hesaplanmasını sağlar [76].

İkili Çapraz Entropi Kaybı, ikili sınıflandırma görevlerinde yaygın olarak kullanılan bir kayıp fonksiyonudur. Çıktısı 0 ile 1 arasında bir olasılık değeri olan bir sınıflandırma modelinin performansını ölçer. Tahmin edilen olasılık gerçek etiketten uzaklaştıkça kayıp artar. İkili Çapraz Entropi Kaybı fonksiyonu aşağıdaki gibidir [76].

$$\text{İkili Çapraz Entropi} = -\frac{1}{n}\sum_{i=1}^{n}[y_i \log(\hat{y}_i) + (1-y_i) log(1-\hat{y}_i)] \quad (5.6)$$

Eşitlik 5.6'da n veri noktalarının sayısı (örneklerin sayısını), y_i i konumundaki veri noktasının gerçek sınıf etiketini (genellikle 0 veya 1), $\hat{y}_i$ ise i konumundaki veri noktası için modelin öngördüğü olasılık veya skor değerini ifade eder.

İkili Çapraz Entropi Kaybı genellikle spam algılama (spam veya spam değil), tıbbi teşhis (hastalık var veya yok) ve ikili duygu analizi (olumlu veya olumsuz duygu) gibi ikili sınıflandırma görevlerinde kullanılır.

Kategorik Çapraz Entropi Kaybı, her girdinin birkaç olası kategoriden birine sınıflandırıldığı çok sınıflı sınıflandırma görevlerinde kullanılan bir kayıp fonksiyonudur. Bu kayıp fonksiyonu, birden fazca çıktısı olan modelin sınıflandırma modelinin performansını ölçer. Temel olarak ikili çapraz entropinin daha kapsamlı halidir. İki sınıf yerine çoklu sınıflandırma süreçlerinde kullanılır. Birden fazla sınıfın olduğu durumlarda, bu entropi türü sınıflar arasındaki belirsizliği değerlendirmek için kullanılır [21]. Kategorik Çapraz Entropi

Kaybı fonksiyonu aşağıdaki gibidir.

$$\text{Kategorik Çapraz Entropi } = -\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{k} y_{ij} \log\left(\hat{y}_{ij}\right) \quad (5.7)$$

Eşitlik 5.7'de n veri kümesindeki örnek sayısını, k ise sınıf sayısını ifade etmektedir. y_{ij}, belirli bir veri noktasının belirli bir sınıfa ait olup olmadığını belirten bir ikili göstergedir. Eğer veri noktası gerçekten o sınıfa aitse, y değeri 1 olur. Aksi takdirde, y değeri 0 olur. $\hat{y}_{ij}$ ise bir örneğin belirli bir sınıfa ait olma olasılığını ifade eder. Bu olasılık değeri, model tarafından tahmin edilir ve 0 ile 1 arasında bir değer alır. Kısaca bu ifade, modelin o örneğin o sınıfa ait olduğunu düşündüğü olasılığı belirtir.

5.2.3. Aktivasyon fonksiyonları

Aktivasyon fonksiyonu, derin öğrenmede sinir ağlarının temel bir bileşenidir ve ağın verilerdeki karmaşık örüntüleri öğrenme ve modelleme becerisinde kritik bir rol oynar. Aktivasyon fonksiyonu, ağa doğrusal olmayan bir özellik katarak karmaşık karar sınırlarının oluşturulmasını ve doğrusal dönüşümlerin tek başına başaramayacağı ilişkilerin yakalanmasını sağlar [22]. Aktivasyon fonksiyonları olmadan, bir sinir ağı sadece bir dizi doğrusal dönüşüm gerçekleştirir ve bunun sonucunda ağın görüntü tanıma, doğal dil işleme ve karmaşık örüntü tanıma gerektiren alanlarda sinir ağı istenilen performansı gösteremez.

Yapay sinir ağı uygulamalarında doğrusal olan ve olmayan aktivasyon fonksiyonları kullanılmaktadır. Doğrusal aktivasyon fonksiyonları, sinir ağlarında kullanılan en basit aktivasyon fonksiyonlarından biridir. Girdi ve çıktı arasında doğrusal bir ilişki ile tanımlanır. Doğrusal aktivasyon fonksiyonlarına örnek $f(x) = x$ fonksiyonu gösterilebilir [22]. Bu fonksiyonlar doğrusal olmayan ilişkileri modelleme yeteneğinden yoksundur, bu da onları esas olarak regresyon görevleri için uygun hale getirir.

Doğrusal olmayan aktivasyon fonksiyonları ise yapay sinir ağlarındaki karmaşık, nonlineer ilişkileri modellemesini sağladıkları için derin öğrenmede kritik öneme sahiptir. Doğrusal olmayan aktivasyon fonksiyonları kullanılmadığında sinir ağı, katman sayısından bağımsız olarak sadece doğrusal dönüşümler gerçekleştirir ve bu da gerçek dünya verilerinin doğrusal olmayan özelliklerini yakalayamayan bir modelle sonuçlanır [21].

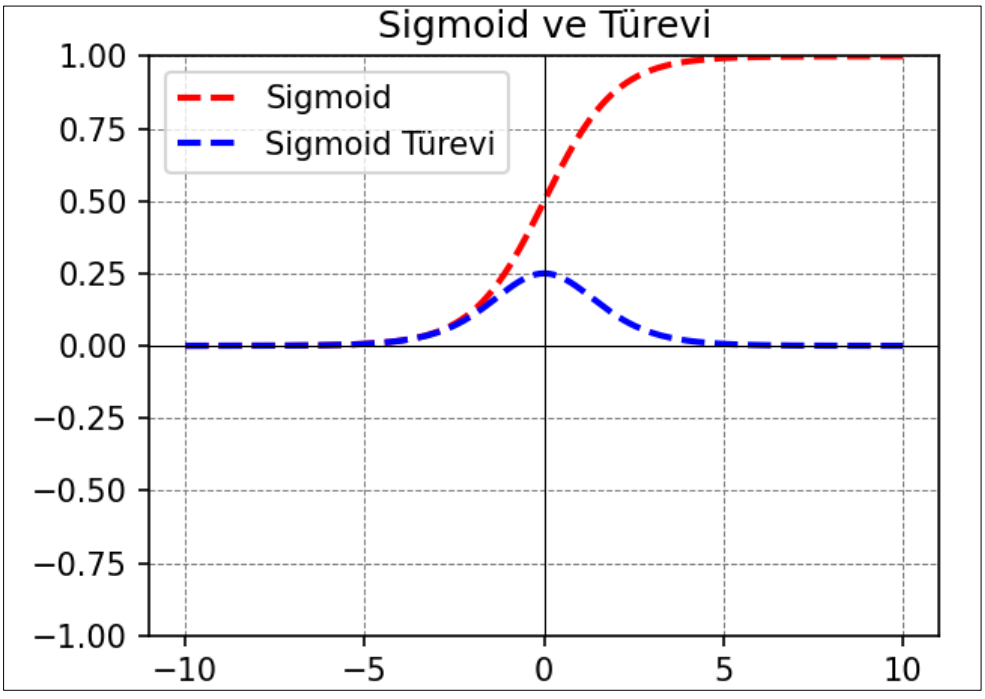
Doğrusal olmayan aktivasyon fonksiyonlarına Sigmoid, Hiperbolik Tanjant, Doğrultulmuş Lineer Birim (ReLU), Sızan Doğrultulmuş Lineer Birim (Leaky ReLU), Parametrik Doğrultulmuş Lineer Birim (PReLU), Swish ve Softmax fonksiyonları örnek verilebilir.

Sigmoid fonksiyonu

Sigmoid fonksiyonu, sinir ağlarında, özellikle ikili sınıflandırma problemlerinde yaygın olarak kullanılan bir aktivasyon fonksiyonudur. Eşitlik 4.8'de verilen Sigmoid Fonksiyonu her giriş değerini 0 ile 1 arasındaki bir çıkış aralığına eşler [22].

$$\sigma(x) = -\frac{1}{1+e^{-x}} \tag{5.8}$$

Eşitlik 5.8'da x terimi, nörona gelen girdilerin ağırlıklı toplamının bias ile toplamını kısaca aktivasyon fonksiyonunun girdisini temsil eder. Sigmoid Fonksiyonunun türev değeri, Şekil 5.8' de görüldüğü üzere, uç noktalarda sıfıra yakınsar (Vanishing Gradient). Bu durum geri yayılım sırasında öğrenmenin durmasına sebebiyet verir. Bu sebeple çok katmanlı ağlarda kullanılması tavsiye edilmez [22].



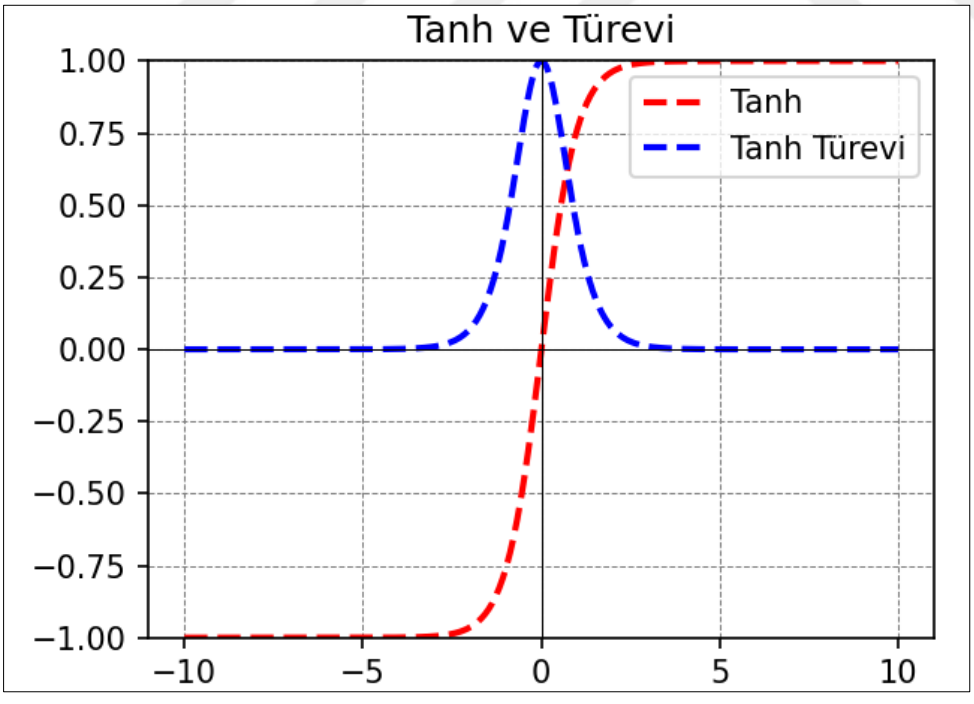
Şekil 5.8. Sigmoid fonksiyonu ve türevi

Hiperbolik Tanjant (Tanh) fonksiyonu

Hiperbolik tanjant (tanh) fonksiyonu, sinir ağlarında yaygın olarak kullanılan bir aktivasyon fonksiyonudur. Sigmoid fonksiyonuna oldukça benzeyen tanh fonksiyonu da aynı sigmoid fonksiyonu gibi ikili sınıflandırmada kullanılır ancak sigmoid fonksiyonundan farklı olarak -1 ile 1 arasında bir değer alır. Tanh fonksiyonu aşağıdaki gibi ifade edilir [22].

$$tanh(x) = -\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{5.9}$$

Eşitlik 5.9'da x terimi, nörona gelen girdilerin ağırlıklı toplamının bias ile toplamını kısaca aktivasyon fonksiyonunun girdisini temsil eder. Tanh fonksiyonunun doğrusal olmama özelliği onu verilerdeki karmaşık örüntüleri öğrenmek için uygun hale getirir ancak Tanh Fonksiyonunun türev değeri, Şekil 5.9'da görüldüğü üzere, uç noktalarda sıfıra yakınsar. Özetle Tanh fonksiyonunda da gradyan kaybı (Vanishing Gradient) sorunu görülür [22]. Bu sebeple çok katmanlı ağlarda bu fonksiyonun da kullanılması tavsiye edilmez.



Şekil 5.9. Tanh fonksiyonu ve türevi

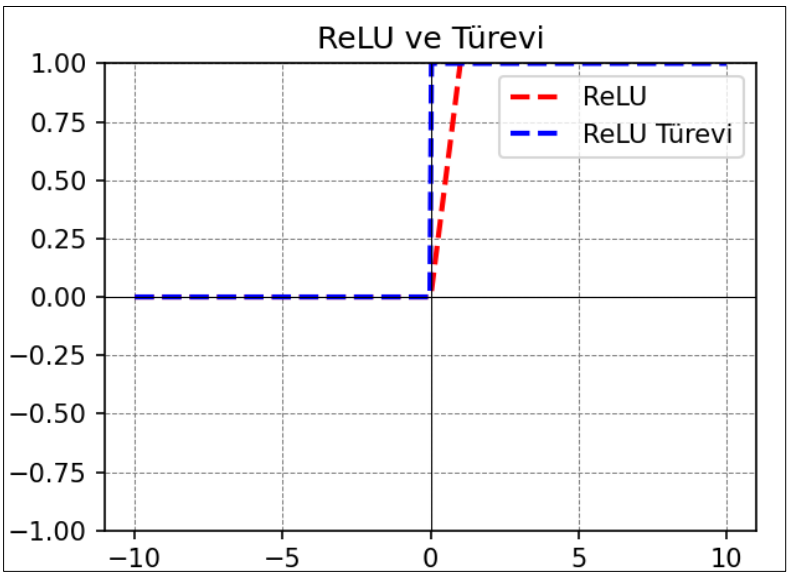
Doğrultulmuş Lineer Birim (ReLU) fonksiyonu

Doğrultulmuş Lineer Birim (ReLU), derin öğrenme modelleri için yapay sinir ağlarında yaygın olarak kullanılan bir aktivasyon fonksiyonudur. Birincil rolü, modele doğrusal olmayan bir özellik katarak öğrenme kapasitesini artırmaktır.

ReLU, girdi pozitifse girdiyi doğrudan çıkarır. Giriş negatifse, çıkışı sıfır yapar. Bu basitlik ReLU'yu hesaplama açısından verimli kılar. Sigmoid ve hiperbolik tanjant gibi diğer aktivasyon fonksiyonlarını etkileyen kaybolan gradyan sorununun üstesinden gelmeye yardımcı olur. Bu etkinliği ReLU'yu, evrişimli sinir ağları gibi birçok sinir ağı mimarisi için varsayılan aktivasyon fonksiyonu haline getirmiştir. Ancak Şekil 5.10'da görüldüğü üzere ReLU'nun negatif değerler için türevinin sıfır olması sebebi ile bazı nöronlar eğitim sırasında pasif hale gelebilir. Özellikle, bir ReLU nöronuna yapılan girişlerin ağırlıklı toplamı tüm eğitim örnekleri için negatifse, geriye yayılma sırasındaki gradyan sıfır olur. Sonuç olarak, nöron öğrenmeyi durdurur ve ölür. Yani gradyan yok olması gerçekleşir. ReLU fonksiyonu matematiksel olarak aşağıdaki gibidir [22].

$$ReLU(x) = x < 0 \; için \; 0, \;\; x \geq 0 \; için \; x \tag{5.10}$$

Eşitlik 5.10'da x terimi, nörona gelen girdilerin ağırlıklı toplamının bias ile toplamını kısaca aktivasyon fonksiyonunun girdisini temsil eder.



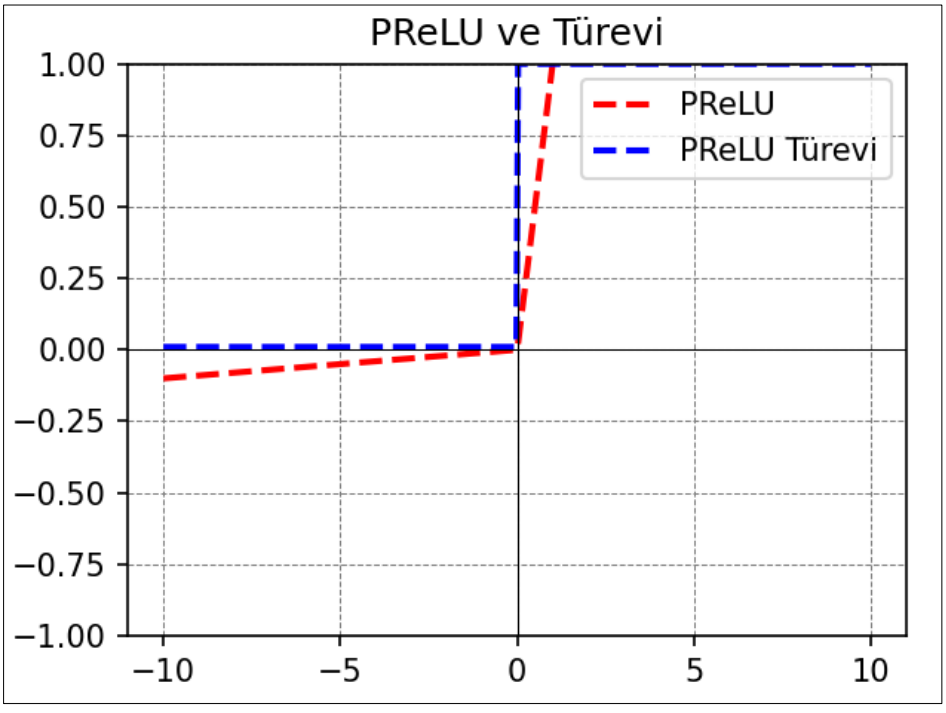
Şekil 5.10. ReLU fonksiyonu ve türevi

Parametrik Doğrultulmuş Lineer Birim (PReLU) fonksiyonu

Parametrik Doğrultulmuş Lineer Birim (PReLU), ReLU'nun bazı sınırlamalarını, özellikle de "ölen ReLU" sorununu ele almak için tasarlanmış, geleneksel Doğrultulmuş Lineer Birim (ReLU) aktivasyon fonksiyonunun gelişmiş bir çeşididir (Şekil 5.11). Ölen ReLU sorunu, nöronlar tüm negatif giriş değerleri için sıfır çıkış verdiğinde ortaya çıkar ve bu nöronların eğitim sırasında öğrenmeyi etkin bir şekilde durdurmasına neden olur. PReLU, negatif girdi değerleri için eğimi belirleyen öğrenilebilir bir parametre olan α'yı ekleyerek bu sorunu hafifletir. PReLU fonksiyonu aşağıdaki gibidir [22].

$$PReLU(x) = x > 0 \; için \; x, \;\; x \leq 0 \; için \; \alpha x \tag{5.11}$$

Eşitlik 5.11'de x terimi, nörona gelen girdilerin ağırlıklı toplamının bias ile toplamını kısaca aktivasyon fonksiyonunun girdisini temsil eder. α ise fonksiyonun negatif kısmının eğimini belirleyen öğrenilebilir bir parametredir. Parametrik ReLU (PReLU) kullanılan bir modelde, α gibi ek parametrelerin öğrenilmesi gerektiği için hesaplama maliyetinde bir artış yaşanır. Eğer model uygun şekilde düzenlenmemişse, bu ek karmaşıklık aşırı uyum riskine yol açabilir.



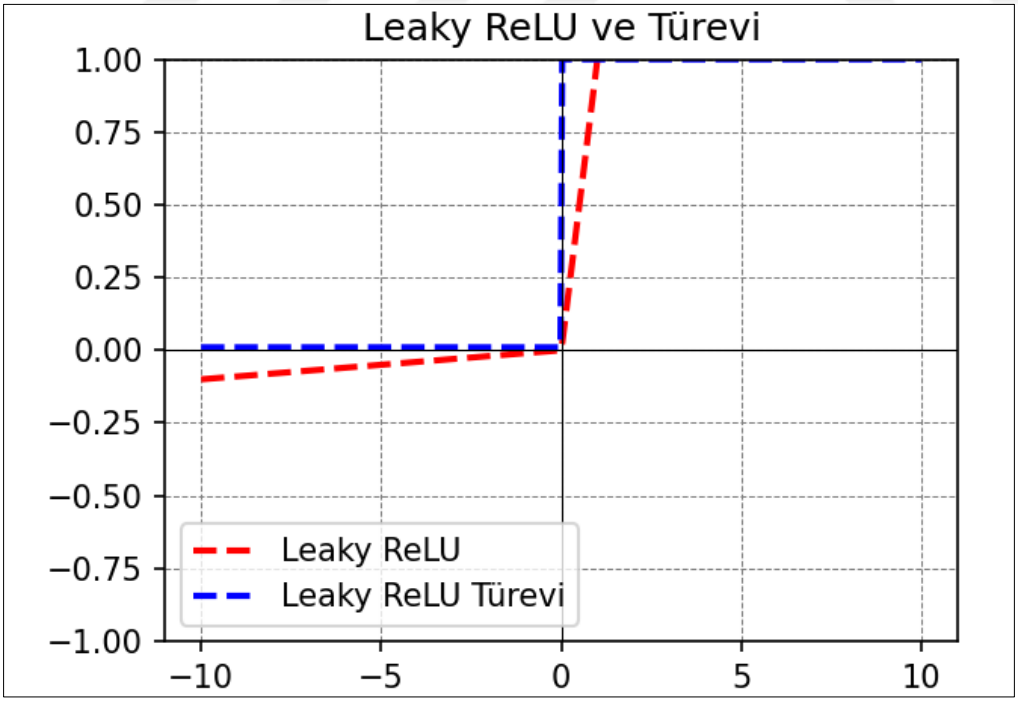
Şekil 5.11. PReLU fonksiyonu ve türevi

Sızan Doğrultulmuş Lineer Birim (Leaky ReLU) fonksiyonu

sızan doğrultulmuş lineer birim (sızan relu), sinir ağlarında "ölen ReLU" sorunu olarak bilinen geleneksel Doğrultulmuş Lineer Birim (ReLU) ile ilgili yaygın bir sorunu ele almak için kullanılan bir aktivasyon fonksiyonudur (Şekil 5.12). Sızan ReLU, negatif girdi değerleri için küçük, sıfır olmayan bir gradyana izin vererek ReLU işlevini değiştirir, böylece nöronları negatif girdiler aldıklarında bile aktif tutar. Negatif değerleri tamamen sıfırlamak yerine, Sızan ReLU bunları küçük bir sabit değerle, tipik olarak 0,01 ile çarpar. Bu küçük gradyan, ağ boyunca bilgi ve gradyan akışının korunmasına yardımcı olarak nöronların eğitim sırasında öğrenmeye devam etmesini sağlar [22]. Leaky ReLU fonksiyonu aşağıdaki gibi tanımlanır.

$$Leaky\ ReLU(x) = x < 0\ için\ 0.01x,\ \ x \geq 0\ için\ x \tag{5.12}$$

Eşitlik 5.12'de x terimi, nörona gelen girdilerin ağırlıklı toplamının bias ile toplamını temsil eder.



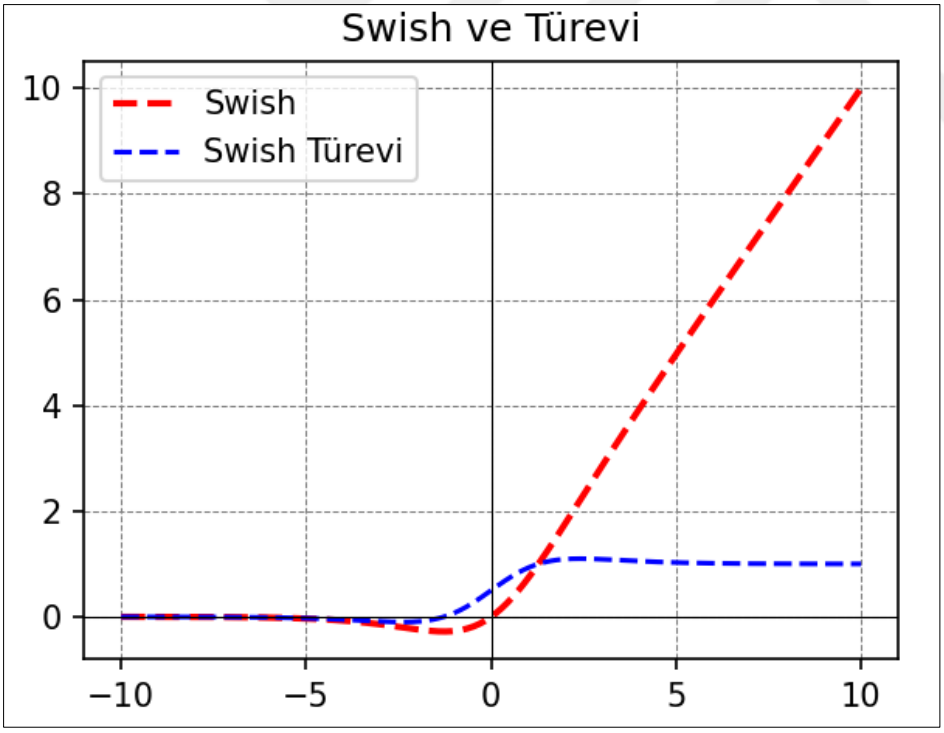
Şekil 5.12. Leaky ReLU fonksiyonu ve türevi

Özetle, hem Sızan ReLU hem de PReLU, negatif girdiler için sıfır olmayan bir gradyan

sağlayarak ölmekte olan ReLU sorununu hafifletmek için tasarlanmış olsa da, yaklaşımlarında farklılık gösterirler. Sızan ReLU negatif eğim için küçük bir sabit kullanır, bu da onu basit ve hesaplama açısından verimli kılar, ancak daha az uyarlanabilir hale getirir. Öte yandan PReLU, eğitim sırasında negatif eğimi öğrenir, daha fazla esneklik ve potansiyel olarak daha iyi performans sunar, ancak ek karmaşıklığa sahiptir [22]. Bu aktivasyon fonksiyonları arasındaki seçim, görevin ve veri kümesinin özel gereksinimlerinin yanı sıra hesaplama verimliliği ve uyarlanabilirlik arasındaki dengeye bağlıdır.

Swish fonksiyonu

Google'daki araştırmacılar tarafından tanıtılan Swish aktivasyon fonksiyonu, sürekli türevlenebilir bir aktivasyon fonksiyonudur (Şekil 5.13). Bu da geri yayılma sırasında kararlı gradyan yayılımını kolaylaştırır ve daha verimli eğitimle sonuçlanır.



Şekil 5.13. Swish fonksiyonu ve türevi

ReLU ve türevlerinin aksine, Swish, bir tür kendi kendini düzenleyen mekanizma içerir. Bu, girdiyi, girdiye bağlı bir faktörle ölçeklendirerek, ağın girdinin farklı bölümlerini esnek bir şekilde bastırmasına veya güçlendirmesine olanak sağlar. Bu özellik, sigmoid bileşeninin

hiçbir zaman sıfıra ulaşmaması nedeniyle, nöronların negatif girdiler için bile sıfır olmayan gradyanlar almasını sağlar. Bu, "ölen ReLU" sorununu hafifletir. Sonuç olarak Swish, derin sinir ağlarının performansını ve öğrenme verimliliğini artırmak için güçlü bir araçtır. Swish fonksiyonu aşağıdaki gibi tanımlanır [77].

$$Swish(x) = x * sigmoid(x) \tag{5.13}$$

x terimi, nörona gelen girdilerin ağırlıklı toplamının bias ile toplamını temsil eder.

Softmax fonksiyonu

Normalleştirilmiş üstel fonksiyon olarak da bilinen Softmax fonksiyonu, genellikle sınıflandırma için kullanılan bir sinir ağının son katmanında, farklı sınıfların olasılıklarını temsil etmek için kullanılır. Softmax fonksiyonu aşağıdaki gibidir [22].

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \tag{5.14}$$

Burada, z softmax fonksiyonunun giriş vektörü ve K sınıf sayısıdır. Softmax işlevi, potansiyel sonuçlar listesinin olasılık dağılımlarını temsil eden bir vektör çıktısı verir. Bu, bir ağın çıktısını, tahmin edilen çıktı sınıflarının olasılık dağılımına normalize eden bir yöntemdir. Softmax fonksiyonunun çıktısı her zaman (0, 1) aralığındadır ve çıktı olasılıklarının toplamı her zaman 1'e eşittir [22].

5.2.4. Optimize ediciler

Derin öğrenmede optimize ediciler, modelleri eğitmek için temel araçlardır. Bunlar, modelin hatalarını ölçen kayıp fonksiyonunu en aza indirmek için eğitim sırasında modelin parametrelerini ayarlayan algoritmalardır. Bir optimize edicinin ana görevi, verilere dayalı olarak ağırlıkları ayarlayarak bu hataları azaltmak ve modelin performansını iyileştirmektir.

Her birinin kendi artıları ve eksileri olan çeşitli optimize edici türleri vardır (Çizelge 5.2). Bazı yaygın kullanılanlar arasında Stokastik Gradyan İnişi (Stochastic Gradient Descent, SGD), Adam ve RMSprop bulunur. Bu optimize ediciler en iyi model parametrelerini

bulmak için farklı kurallar, öğrenme oranları ve momentum stratejileri kullanır [22].

Çizelge 5.2. Optimize edicilerin karşılaştırılması

Özellik	SGD	RMSProp	ADAM
Öğrenme Hızı	Sabit	Dinamik	Dinamik
Ağırlıkların Güncellenmesi	Basit	Daha karmaşık	Daha karmaşık
Bellek Kullanımı	Düşük	Orta	Yüksek
Hiperparametreler	Öğrenme hızı	Öğrenme hızı, decay	Öğrenme hızı, β1, β2, ε
İyi Performans Gösterdiği Durumlar	Düzgün kayıp yüzeyleri	Gürültülü/dağınık kayıp yüzeyleri	Çeşitli kayıp yüzeyleri

Derin öğrenmede süper çözünürlük gibi görevlerde, optimize edici seçimi model performansını büyük ölçüde etkileyebilir. Bu görevlerde yaygın olarak kullanılan optimize ediciler SGD ve Adam'dır. SGD, basitliği ve verimliliği ile öne çıkarken Adam, diğer SGD varyasyonlarının güçlü yönlerini birleştirerek karmaşık ve gürültülü verileri işleme yeteneği nedeniyle tercih edilmektedir [22].

5.2.5. Geriye yayılım ve gradyan inişi

Geriye Yayılma ve Gradyan İnişi, derin öğrenmede temel algoritmalardır ve süper çözünürlük görevlerinde kullanılanlar da dahil olmak üzere sinir ağlarını eğitmek için çok önemlidir. Geriye yayılma, sinir ağının ağırlıklarını ayarlayarak hatayı en aza indirmek için kullanılan bir optimizasyon algoritmasıdır. Zincir kuralını uygulayarak her bir ağırlığa göre kayıp fonksiyonunun gradyanını hesaplar ve ağın hatayı yinelemeli olarak azaltarak öğrenmesini sağlar. Gradyan İnişi ise bu ağırlıkları gradyanın tersi yönünde hareket ettirerek güncelleyen, kayıp fonksiyonunun yerel veya küresel minimumlarını bulmayı amaçlayan bir yöntemdir [21].

5.2.6. Veri arttırma

Veri artırımı, mevcut verilerin değiştirilmiş kopyalarını oluşturarak eğitim veri kümesinin boyutunu ve çeşitliliğini yapay olarak artıran bir tekniktir. Derin öğrenmede sınırlı eğitim verisi sorununu çözmek ve aşırı uyumu (overfitting) önlemek için kullanılır. Veri artırımı, çeşitli dönüşümler yoluyla ek eğitim örnekleri oluşturarak modelin daha çeşitli ve değişmez

özellikler öğrenmesine yardımcı olur.

5.2.7. Yığın normalleştirme

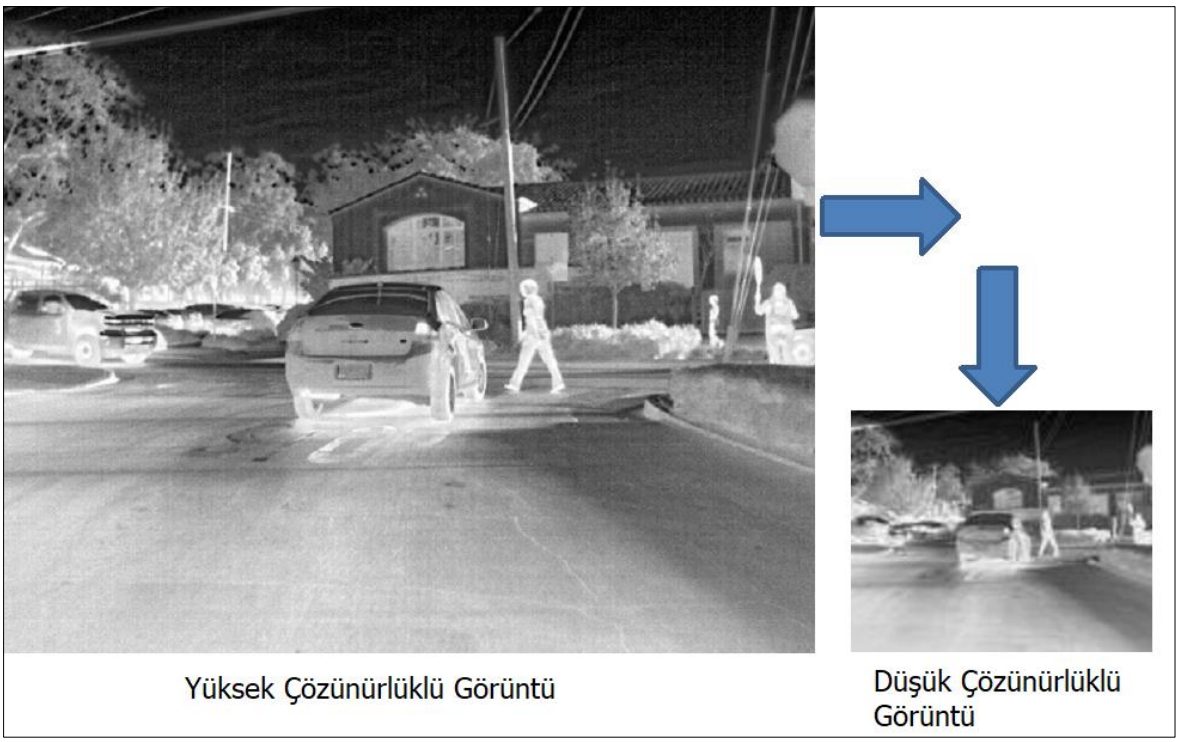
2015 yılında Sergey Ioffe ve Christian Szegedy tarafından tanıtılan Yığın Normalleştirme, derin öğrenme sürecinde her bir katmanın girdilerinin dağılımının eğitim sırasında değişmesi sorununu, yani dahili ortak değişken kaymasını ele almak için kullanılan bir tekniktir [78]. Yığın Normalleştirme, bir önceki aktivasyon katmanının çıktısını normalize eder. Bu, yığın ortalamasını çıkararak ve yığın standart sapmasına bölerek gerçekleştirilir. Bu işlem, eğitim sürecini daha stabil hale getirir ve hızlandırır. Daha yüksek öğrenme oranlarına olanak sağlar ve başlangıç hassasiyetini azaltır. Aynı zamanda, modelin ağırlıklarını eğitim sırasında güncellerken kullanılan gradyanların çok küçük olması (vanishing gradients) veya çok büyük olması (exploding gradients) durumlarını önlemeye yardımcı olur [21]. Eğer gradyanlar çok küçükse, ağırlıklar çok yavaş bir şekilde güncellenir ve bu da modelin eğitiminin çok yavaş ilerlemesine veya tamamen durmasına neden olabilir. Öte yandan, eğer gradyanlar çok büyükse, ağırlıklar çok hızlı bir şekilde güncellenir ve bu da modelin ağırlıklarının istikrarsız bir şekilde dalgalanmasına ve modelin optimum çözümü bulamamasına neden olabilir. Bu durum genellikle "exploding gradients" problemi olarak adlandırılır [21]. Yığın Normalleştirme (Batch Normalization) tam da bu tür problemleri çözmek için kullanılır. Her bir küçük yığındaki (mini-batch) aktivasyonların ortalamasını 0 ve standart sapmasını 1 olacak şekilde normalleştirir. Bu, hem "vanishing gradients" hem de "exploding gradients" problemlerini önlemeye yardımcı olur, böylece modelin eğitim süreci daha kararlı ve hızlı hale gelir. Bu nedenle, Yığın Normalleştirme'nin kullanılması genellikle modelin eğitim sürecini hızlandırır ve modeli daha kararlı hale getirir [21].

6. TERMAL GÖRÜNTÜLERDE SÜPER ÇÖZÜNÜRLÜK

Bu bölümde Geleneksel Süper Çözünürlük Yöntemlerinden Bikübik Enterpolasyon, En Yakın Komşu Enterpolasyonu, Bilineer Enterpolasyon ve Derin Öğrenme Tabanlı Süper Çözünürlük Yöntemlerinden VDSR, SRGAN, ESRGAN ve SwinIR modellerinden bahsedilip termal görüntüler üzerine uygulanmıştır.

6.1. Geleneksel Süper Çözünürlük Yöntemlerinin Termal Görüntülere Uygulanması

640x512 çözünürlüğe sahip olan bir termal görüntüye, bikübik ve bilineer enterpolasyon teknikleri kullanılarak 4 kat alt örnekleme uygulanmış ve bu işlem sonucunda 160x128 çözünürlüğe sahip düşük çözünürlüklü bir görüntü elde edilmiştir. Bu düşük çözünürlüklü görüntü, Şekil 6.1 ve 6.2'de gösterilmiştir. Daha sonra, bu düşük çözünürlüklü görüntüye Bilineer ve Bikübik Enterpolasyon teknikleri uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlem sonucunda elde edilen görüntüler, Şekil 6.3 ve 6.4'te sunulmuştur.



Şekil 6.1. Bilineer enterpolasyon tekniği kullanılarak yüksek çözünürlüklü görüntüden 4 kat alt örnekleme sonucu elde edilen düşük çözünürlüklü görüntü



Şekil 6.2. Bikübik enterpolasyon tekniği kullanılarak yüksek çözünürlüklü görüntüden 4 kat alt örnekleme sonucu elde edilen düşük çözünürlüklü görüntü

Şekil 6.1'deki düşük çözünürlüklü görüntüye En Yakın Komşu Enterpolasyonu, Bilineer Enterpolasyon ve Bikübik Enterpolasyon teknikleri uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlem sonucunda elde edilen görüntüler, Şekil 6.3'te sunulmuştur.



Şekil 6.3. Geleneksel süper çözünürlük yöntemlerinin bilineer enterpolasyon ile alt örnekleme yapılmış termal görüntülere uygulanması

Şekillerden görüldüğü gibi orijinal görüntüye en yakın görüntüyü bikübik enterpolasyon sağlarken, en yakın komşu enterpolasyonu en bulanık görüntüye sebep olmuştur. Düşük çözünürlüklü görüntü elde edilirken bilineer enterpolasyon kullanıldığı için bikübik

enterpolasyon ile elde edilmiş yüksek çözünürlüklü görüntü orijinal görüntüye göre hala yüksek kayıp içerir.

Şekil 6.2'deki düşük çözünürlüklü görüntüye En Yakın Komşu Enterpolasyonu, Bilineer Enterpolasyon ve Bikübik Enterpolasyon teknikleri uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlem sonucunda elde edilen görüntüler, Şekil 6.4'te sunulmuştur.



Şekil 6.4. Geleneksel süper çözünürlük yöntemlerinin bikübik enterpolasyon ile alt örnekleme yapılmış termal görüntülere uygulanması

Şekillerden görüldüğü üzere orijinal görüntüye en yakın görüntüyü bikübik enterpolasyon sağlarken, en yakın komşu enterpolasyonu en bulanık görüntüye sebep olmuştur. Ayrıca, düşük çözünürlüklü görüntü elde edilirken bikübik enterpolasyon kullanıldığı için bikübik enterpolasyon ile elde edilmiş yüksek çözünürlüklü görüntülerin Şekil 6.3'teki çözünürlüğü yükseltilmiş şekillere kıyasla daha pürüzsüz ve keskin olduğu görülmektedir.

6.2. Derin Öğrenme Tabanlı Süper Çözünürlük Yöntemleri

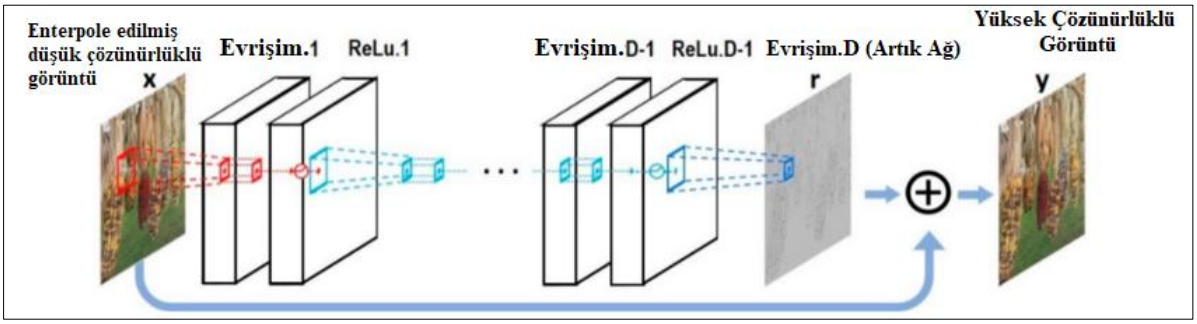
Bu bölüm altında Derin Öğrenme Tabanlı Süper Çözünürlük Yöntemlerinden VDSR, SRGAN, ESRGAN ve SwinIR modelleri tanıtılıp ardından termal görüntüler üzerindeki örnekleri sunulmuştur.

6.2.1. VDSR

Çok Derin Süper Çözünürlük (Very Deep Super Resolution, VDSR) modeli, derin öğrenmenin gücünden yararlanarak görüntü süper çözünürlüğü alanında önemli bir ilerlemeyi temsil etmektedir [6, 22, 79]. Kim ve arkadaşları tarafından 2016 yılında tanıtılan VDSR, tipik olarak 20 katmandan oluşan derin bir konvolüsyonel sinir ağı mimarisiyle tasarlanmıştır [6]. Bu derinlik, modelin yüksek kaliteli süper çözünürlük için gerekli olan görüntülerdeki karmaşık desenleri ve ince ayrıntıları yakalamasını sağlar [22]. VDSR'deki önemli bir yenilik Şekil 6.5'te sunulduğu üzere artık öğrenmenin kullanılmasıdır. Ağ, düşük çözünürlüklü (LR) bir görüntüyü doğrudan yüksek çözünürlüklü (HR) muadiliyle eşlemek yerine, HR görüntüsü ile LR görüntüsü arasındaki farklar olan kalıntıları tahmin etmeyi öğrenir [22, 79]. Bu yaklaşım öğrenme sürecini basitleştirir ve yakınsamayı hızlandırarak çok derin ağların etkili bir şekilde eğitilmesini mümkün kılar [6, 22, 79].

VDSR modelindeki her bir evrişimsel katman, hesaplama karmaşıklığı ile ayrıntılı görüntü özelliklerini öğrenme yeteneği arasındaki dengeyi sağlayan bir yapılandırma olan 3x3 boyutunda 64 filtre kullanır [6, 22, 79]. Ağın girdisi, düşük çözünürlüklü görüntünün bikübik enterpolasyon kullanılarak istenen boyuta yükseltilmiş enterpolasyonlu bir versiyonudur. Bu yaklaşım, ağın yükseltme işlemini kendisi yapmak yerine yükseltilmiş görüntüyü iyileştirmeye odaklanmasını sağlar [22, 79]. Model, tahmin edilen piksel değerleri ile gerçek değerleri arasındaki farkların karelerinin ortalamasını ölçen kayıp fonksiyonu olarak ortalama karesel hata (MSE) kullanılarak eğitilir. Eğitim, stokastik gradyan inişi (SGD)

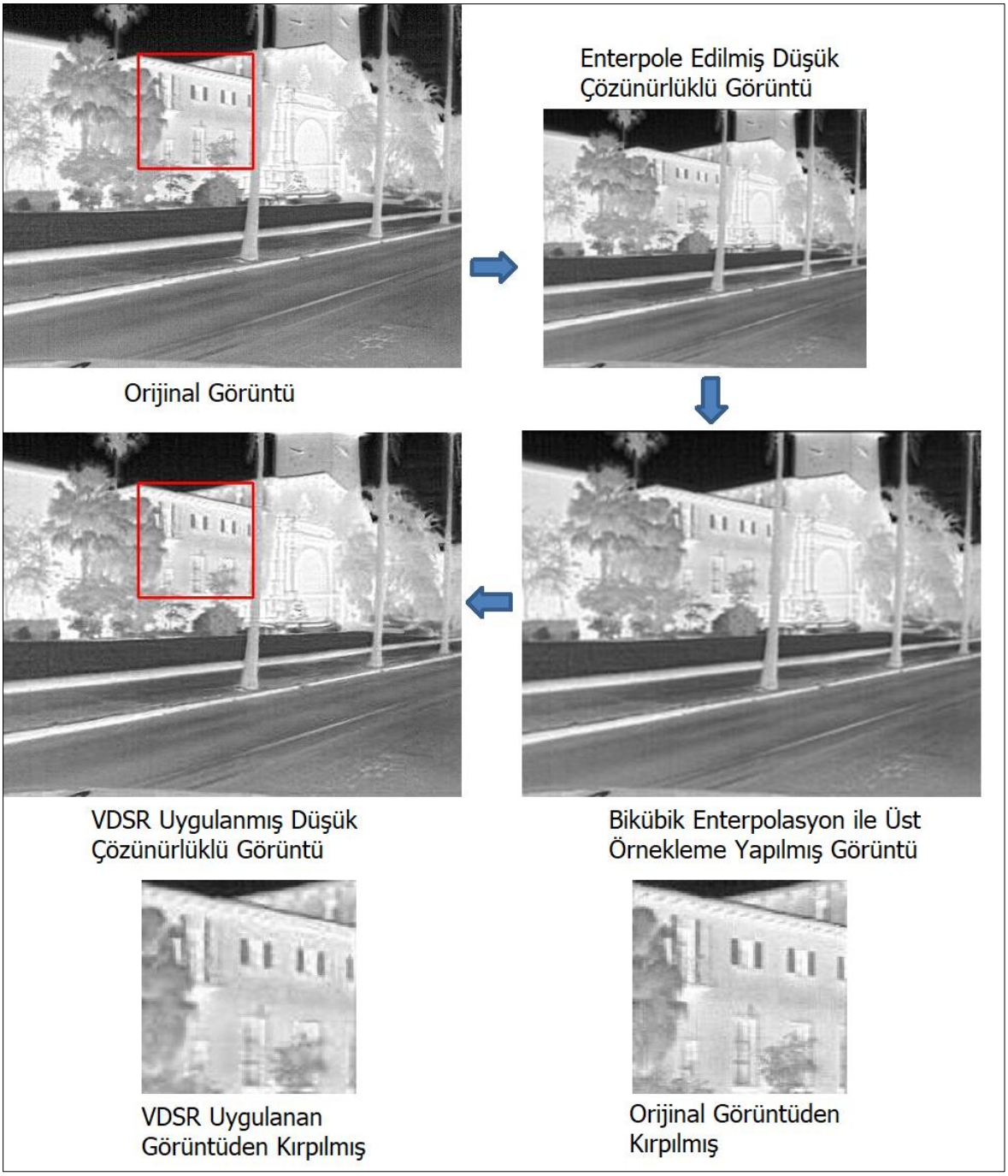
kullanılarak gerçekleştirilir ve eğitim esnasında yakınsamayı artırmak için ayarlanabilir öğrenme oranı kullanılır [6, 22].



Şekil 6.5. VDSR mimarisi [80]

VDSR modelinin görüntü çiftlerinden oluşan büyük veri kümelerinden öğrenme ve genelleme yeteneği, onu çeşitli görüntü türleri ve çözünürlükler arasında oldukça etkili kılmaktadır [6, 22]. Döndürme, çevirme ve ölçekleme gibi veri artırma teknikleri, eğitim setinin çeşitliliğini artırmak ve modelin sağlamlığını geliştirmek için eğitim sırasında sıklıkla kullanılır. VDSR'nin performansı, görsel kalite açısından VDSR'dan önce literatüre kazandırılan süper çözünürlük yöntemlerine kıyasla daha iyi sonuçlar verdiği standart kıyaslama veri kümeleri üzerinde kapsamlı bir şekilde doğrulanmıştır [6, 22, 79]. Derin ağ mimarisi, artık öğrenme ve verimli eğitim tekniklerinin birleşimi, VDSR'yi görüntü süper çözünürlüğü alanında önemli bir model olarak konumlandırmakta ve bu alandaki sonraki ilerlemelerin önünü açmaktadır [6, 22, 79].

Orijinal görüntüden bikübik enterpolasyon ile alt örnekleme yapılmış düşük çözünürlüklü görüntü bikübik enterpolasyon ile tekrardan orijinal görüntü boyutlarına getirilmiş ardından VDSR modeli orijinal görüntü ile aynı boyutlarda olan düşük çözünürlüklü görüntüye uygulanmıştır. Bu işlem sonucunda elde edilen görüntüler, Şekil 6.6'da sunulmuştur.

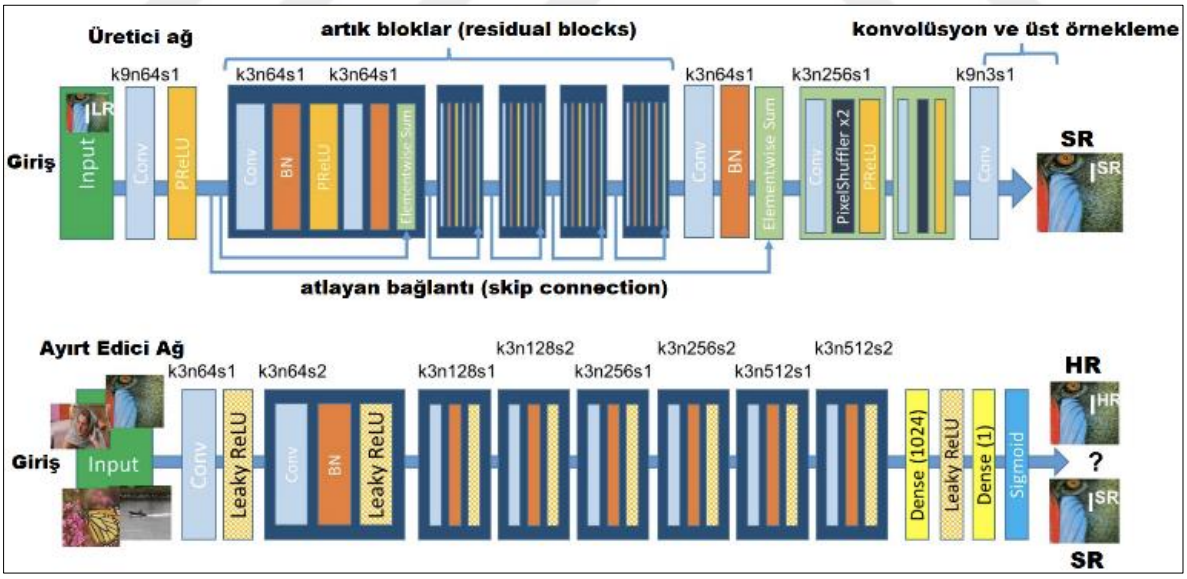


Şekil 6.6. VDSR ile çözünürlüğü iyileştirilmiş termal görüntü

Şekil 6.6'da görüldüğü üzere düşük çözünürlüklü resme VDSR modeli uygulanarak elde edilen yüksek çözünürlüklü görüntü orijinal görüntüye yakın olsa da çözünürlük yükseltme sırasında görüntüde bazı bölgelerde bozulmaların ve bulanıkların oluştuğu görülmektedir.

6.2.2. SRGAN

Süper Çözünürlüklü Üretken Çekişmeli Ağ (Super Resolution Generative Adversarial Network, SRGAN), Ledig ve arkadaşları tarafından 2017 yılında tanıtılan görüntü süper çözünürlüğü alanında önemli bir sıçramayı temsil etmektedir. SRGAN, düşük çözünürlüklü girdilerden yüksek çözünürlüklü görüntüler oluşturan Bölüm 4.2.1 altında da bahsedildiği ve Şekil 6.6'da gösterildiği gibi bir üretici ağ ile gerçek ve üretilen yüksek çözünürlüklü görüntüler arasında ayrım yapan bir ayırt edici ağ içeren üretken çekişmeli bir modeldir. Üretici ağ, her biri 3x3 boyutunda 64 filtreli konvolüsyonel katmanlar, yığın normalizasyonu (Batch Normalizasyonu) ve PReLU aktivasyon fonksiyonları içeren artık blokları kullanır [81]. Bu artık bloklar, yüksek frekanslı ayrıntıları geliştirirken düşük frekanslı içeriği koruyarak öğrenme sürecini kolaylaştırır [81]. Ek olarak, SRGAN'ın üretici ağı, alt piksel konvolüsyonu yoluyla uygulanan ve görüntü çözünürlüğünü verimli bir şekilde artıran yukarı örnekleme katmanları içerir [34]. Öte yandan, ayırt edici ağ, artan filtre boyutlarına ve Sızan ReLU aktivasyonlarına sahip konvolüsyonel katmanlardan oluşur ve görüntü orijinalliği için bir olasılık puanı veren yoğun katmanlarla tamamlanır.

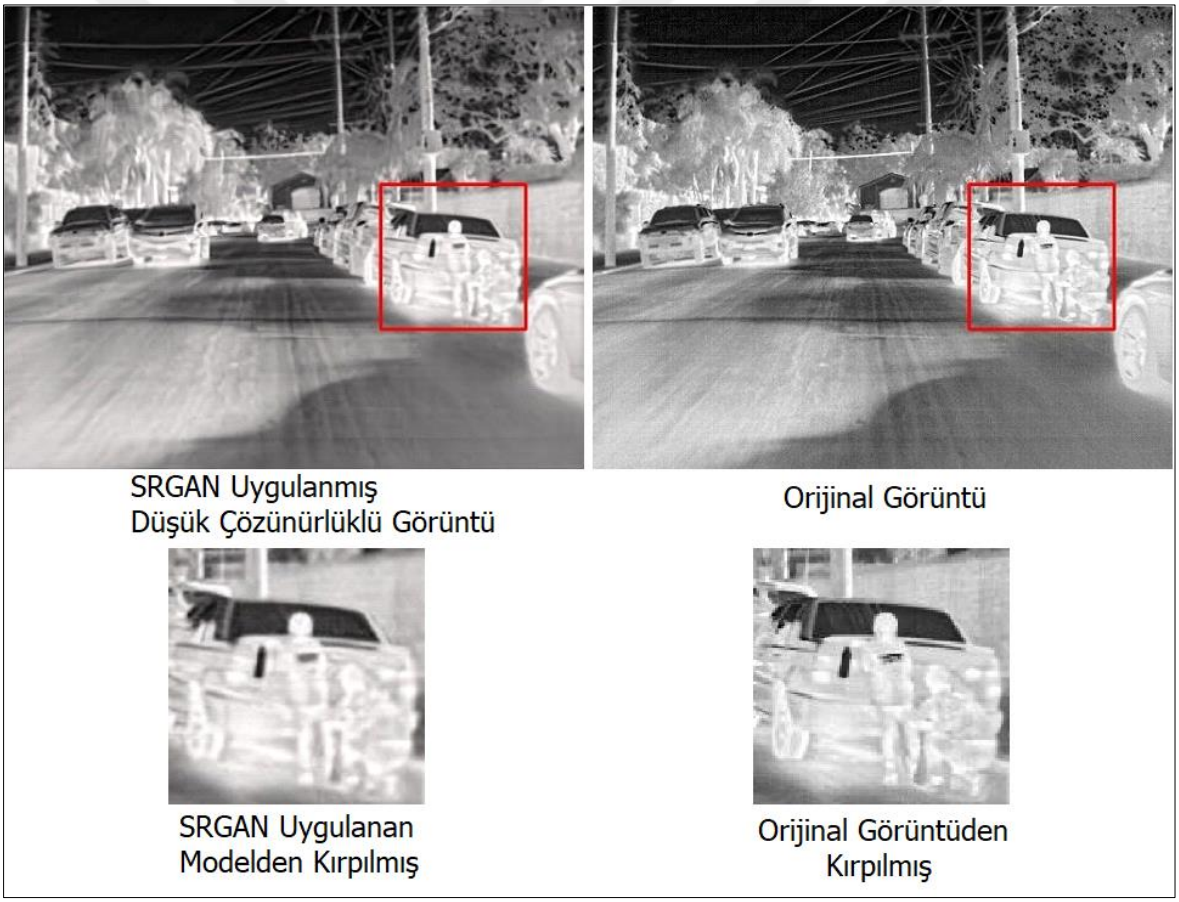


Şekil 6.7. SRGAN mimarisi [82]

SRGAN, içerik kaybını ölçen ortalama karesel hata (MSE), ayırt edici ağdan gelen çekişmeli kayıp ve önceden eğitilmiş bir VGG ağından elde edilen üst düzey özellik haritalarıyla hesaplanan algısal kayıp fonksiyonlarının bir kombinasyonunu kullanır [2, 8]. Model, ADAM optimizatörü [8] kullanılarak tipik öğrenme oranı olan 1e-4 öğrenme oranıyla

optimize edilmiş ve modele stabil bir şekilde yakınsama sağlamak için dinamik olarak öğrenme oranı azaltımı yöntemi uygulanmıştır [8]. Artık öğrenme ile basit bir derin evrişimli ağ kullanan ve piksel bazında kaybı en aza indirmeye odaklanan VDSR modeliyle karşılaştırıldığında, SRGAN'ın GAN tabanlı mimarisi ve çok yönlü kayıp fonksiyonları daha gerçekçi görüntüler üretmesini sağlamıştır. Artan eğitim karmaşıklığına rağmen, SRGAN'ın görsel olarak daha net ve yüksek kaliteli sonuçlar üretme yeteneği, onu görüntü süper çözünürlüğü alanında önemli bir model haline getirmiştir [2, 81].

Düşük çözünürlüklü görüntüye SRGAN modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlem sonucunda elde edilen görüntü, Şekil 6.8'de sunulmuştur.

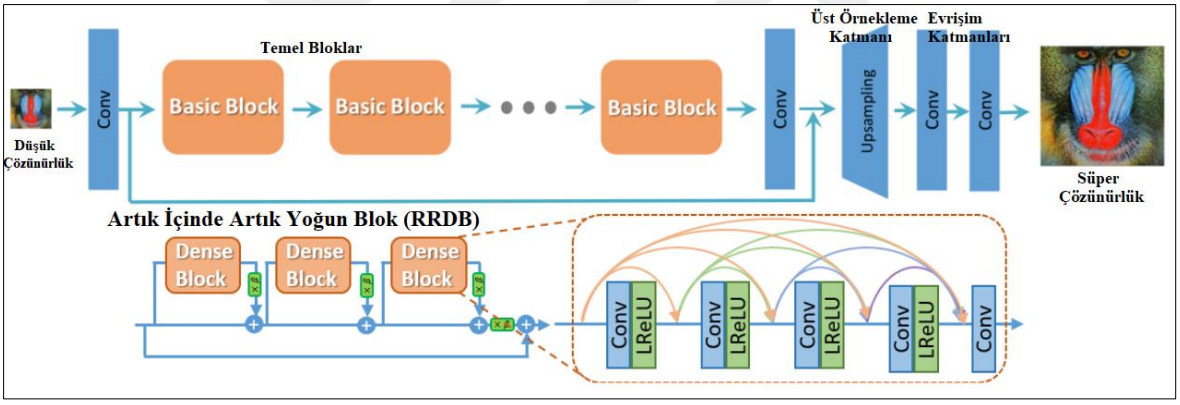


Şekil 6.8. SRGAN ile çözünürlüğü iyileştirilmiş termal görüntü

Şekil 6.8'de görüldüğü üzere düşük çözünürlüklü resme SRGAN modeli uygulanarak elde edilen yüksek çözünürlüklü görüntü orijinal görüntüye oldukça benzemektedir. Ancak her ne kadar benzer bir görüntü olarak gözükse de çözünürlük yükseltme sırasında görüntüde bazı bölgelerde pürüzlerin oluştuğu görülmektedir.

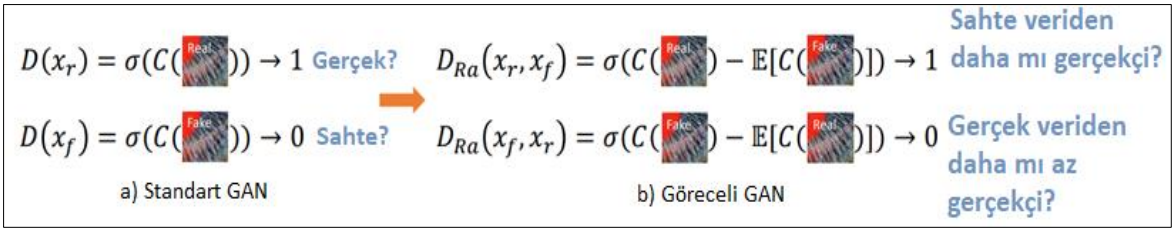
6.2.3. ESRGAN

Geliştirilmiş Süper Çözünürlüklü Üretken Çekişmeli Ağ (Enhanced Super-Resolution Generative Adversarial Network, ESRGAN), Wang ve arkadaşları tarafından 2018 yılında tanıtılan görüntü süper çözünürlüğü alanında önemli bir ilerlemeyi temsil etmektedir [9]. SRGAN tarafından sunulan bilgiler üzerine inşa edilen ESRGAN, görüntü kalitesini ve gerçekçiliğini artırmak için çeşitli mimari geliştirmeler içermektedir. ESRGAN'da üretici ağın temel bileşeni, özellik çıkarımını geliştirmek ve bilgi akışını korumak için artık öğrenmeyi yoğun bağlantılarla birleştiren Artık İçinde Artık Yoğun Bloktur (RRDB) (Şekil 6.9) [9]. Her bir RRDB, yoğun bağlantılara sahip birden fazla katmandan oluşur ve eğitimi stabilize etmek için artık ölçekleme uygulanır [26]. Üretici ağ ayrıca görüntü çözünürlüğünü verimli bir şekilde artırmak için alt piksel konvolüsyonu kullanan yukarı örnekleme katmanları içerir [83].



Şekil 6.9. ESRGAN üreteç ağı mimarisi [84]

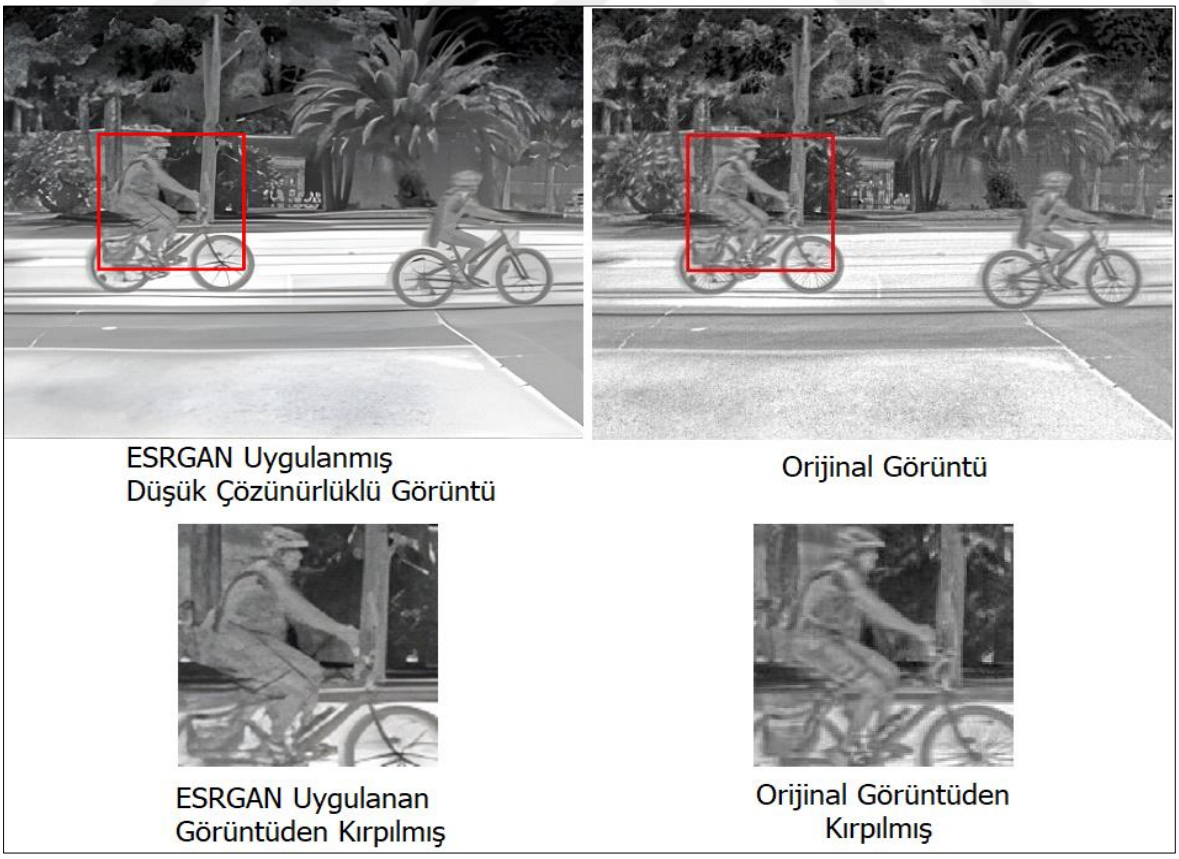
ESRGAN'daki ayırt edici ağ, SRGAN'a benzer şekilde, artan filtre boyutları ile Sızan ReLU aktivasyonlarına sahip konvolüsyonel katmanlardan ve ardından bir olasılık puanı çıkarmak için tam bağlı katmanlardan oluşur [85]. ESRGAN, içerik kaybı, çekişmeli kayıp ve algısal kaybın bir kombinasyonunu kullanır [9]. Ayrıca, çekişmeli eğitim için Relativistik GAN (RaGAN) kullanır [85]. Bu yaklaşım, ayırt edicinin Şekil 6.10'da sunulduğu gibi gerçek ve üretilen görüntüleri daha etkili bir şekilde karşılaştırmasını sağlar [85]. Model, ADAM optimizer kullanılarak hem üretici hem de ayırt edici ağ için tipik olarak 1e-4'e ayarlanmış bir öğrenme oranı ile optimize edilir ve kararlı yakınsama için öğrenme oranı düşüşü uygulanır [9].



Şekil 6.10. Göreceli ayırt edici model [9]

SRGAN ile karşılaştırıldığında, ESRGAN'ın artık içinde artık yoğun blokları, yoğun bağlantıları ve artık ölçeklendirmeyi kullanması, özellik çıkarma ve öğrenme verimliliğini önemli ölçüde artırarak daha yüksek kaliteli ve daha gerçekçi görüntüler elde edilmesini sağlar. Bu geliştirmeler ESRGAN'ı görüntü süper çözünürlüğü alanında, öncekilere göre daha az yapaylıkla daha gerçekçi sonuçlar üretebilen, çok önemli bir model haline getirmektedir [9].

Düşük çözünürlüklü görüntüye ESRGAN modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlem sonucunda elde edilen görüntü, Şekil 6.11'de sunulmuştur.



Şekil 6.11. ESRGAN ile çözünürlüğü iyileştirilmiş termal görüntü

Şekil 6.11’de görüldüğü üzere düşük çözünürlüklü resme ESRGAN modeli uygulanarak elde edilen yüksek çözünürlüklü görüntünün orijinal görüntüdeki bulanıklığı giderdiği görülmektedir. Buna rağmen resmin bazı bölgelerindeki küçük detayların da süper çözünürlük esnasında kaybolduğu görülmektedir.

6.2.4. SwinIR

SwinIR (Swin Transformer for Image Restoration), 2021 yılında Liu ve arkadaşları tarafından tanıtılan, süper çözünürlük, denoising ve artefakt giderme gibi görevler için tasarlanmış gelişmiş bir modeldir [38, 40]. Kaydırılmış pencere tabanlı bir dikkat mekanizması kullanarak hem bölgesel (yerel) hem de genel (küresel) özellikleri yakalayarak görüntüleri işleyen Şekil 6.12’de sunulan Swin Transformer mimarisinden yararlanır [39, 40].

Anahtar kavramlar

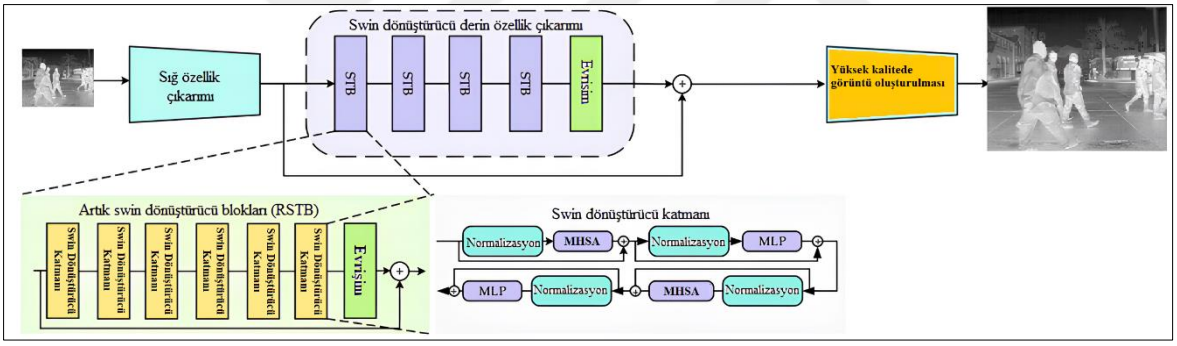
Dönüştürücü, sinir ağları bağlamında bir dönüştürücü, önemli özelliklerini anlamak için bir görüntünün farklı kısımlarına odaklanan akıllı bir filtre gibi davranır. Görüntüleri sabit bir şekilde işleyen geleneksel yöntemlerin aksine, Transformatörler görüntünün kapsamlı bir temsilini yakalamak için odaklarını dinamik olarak ayarlayabilirler [86].

Sinir ağlarında, bir Transformatör akıllı bir filtre gibi davranır ve bir görüntünün kapsamlı özelliklerini yakalamak için odağını dinamik olarak ayarlar. Swin Dönüştürücüsü, bu özel Transformatör türü, bir görüntüyü küçük yamalara (bir yapbozun parçaları gibi) böler ve her yamayı bağımsız olarak işler yani öz dikkat işlemi (self attention) greçekleştirir. Daha sonra görüntüyü daha iyi anlamak için bu yamaları hafifçe kaydırarak hem küçük ayrıntıları hem de genel yapıyı yakalar. Bu yapıya da Kaydırmalı Pencere Dikkat Mekanizması (Shifted Window Attention) denir [87, 88].

Swin Transformer katmanlarındaki Çok Katmanlı Algılayıcı (Multi Layer Perceptron, MLP) bileşenleri, doğrusal olmayan aktivasyon fonksiyonları tarafından takip edilen çoklu doğrusal dönüşümlerden oluşan bir sinir ağı katmanı türüdür [37]. SwinIR modelinde MLPler, öz dikkat (self mekanizması tarafından çıkarılan özellikleri daha fazla işlemek için kullanılır [37]. Böylece MLP sayesinde daha karmaşık ve zengin özellikler öğrenilir

Çok Başlı Öz Dikkat (Multi-Head Self-Attention, MHSA) mekanizması, modelin aynı anda birden fazla odaklama işlemi gerçekleştirmesini sağlar; burada her kafa bağımsız olarak görüntünün farklı kısımlarına odaklanmayı öğrenir [89]. Swin Transformer katmanlarındaki Çok Katmanlı Algılayıcılar (MLP), öz dikkat mekanizması tarafından çıkarılan özellikleri işleyerek karmaşık ve zengin özellikleri öğrenir [89]. Katman normalizasyonu (LN), her veri örneği için özellikler arasında girdileri normalleştirerek eğitimi stabilize eder ve hızlandırır, her katmana gelen girdilerin tutarlı bir dağılıma sahip olmasını sağlayarak eğitim sırasında istikrarlı gradyanların korunmasına yardımcı olur [90].

Artık öğrenme, modelin düşük kaliteli bir görüntü ile yüksek kaliteli bir görüntü arasındaki farkı öğrenmesini sağlayarak öğrenme sürecini basitleştirir ve performansı artırır [10]. Bir gömme katmanı, giriş görüntüsünün yamalarını daha yüksek boyutlu vektörlere dönüştürerek öz dikkat gibi daha karmaşık işlemleri kolaylaştırır [89].



Şekil 6.12. SwinIR mimarisi [91]

SwinIR birkaç temel modülden oluşmaktadır.

Sığ özellik çıkarma

İlk aşama, bir konvolüsyonel katmanın temel özellikleri tanımlamak için görüntünün ön taramasını gerçekleştirdiği yüzeysel özellik çıkarımını içerir. Bu süreç, bir yapbozun kenar parçalarını düzenlemeye benzetilebilir ve daha sonraki, daha ayrıntılı analiz için temel bir yapı sağlar.

Bunu takiben, giriş görüntüsü örtüşmeyen yamalara bölünür. Her bir yama daha sonra düzleştirilir ve doğrusal bir dönüşüme tabi tutularak daha yüksek boyutlu bir vektöre

dönüştürülür [37]. Matematiksel olarak, x_p görüntünün bir yamasını temsil ediyorsa, gömme katmanı Eşitlik 6.1'deki gibi tanımlanır.

$$F_0 = Lineer(x_p) \tag{6.1}$$

Burada F_0, orijinal görüntü yamasının daha ayrıntılı ve daha yüksek boyutlu bir temsilini sağlayan gömme vektörünü ifade eder. Bu dönüşüm, modelin yamaları sabit boyutlu vektörler olarak ele almasını ve böylece öz dikkat gibi daha karmaşık işlemleri kolaylaştırmasını sağlar [86].

Derin özellik çıkarımı

Bu aşamada görüntü yamalarını ayrıntılı olarak işlemek için Artık Swin Dönüştürücü Blokları (Residual Swin Transformer Blocks, RSTB) kullanılır [10]. Swin Dönüştürücüsü görüntüyü küçük yamalara böler, her birini bağımsız olarak işler ve Kaydırılmış Pencere Dikkat mekanizması aracılığıyla görüntüyü daha iyi anlamak için bu yamaları kaydırır [25]. Bu kısımda Çok Başlı Öz Dikkat (Multi Head Self Attention, MHSA), Çok Katmanlı Algılayıcı (Multi Layer Perceptron, MLP), Katman Normalizasyonu (Layer Normalization, LN), Artık Bağlantı (Residual Connection) ve Üst Örnekleme (Upsampling) işlemleri gerçekleşir [89].

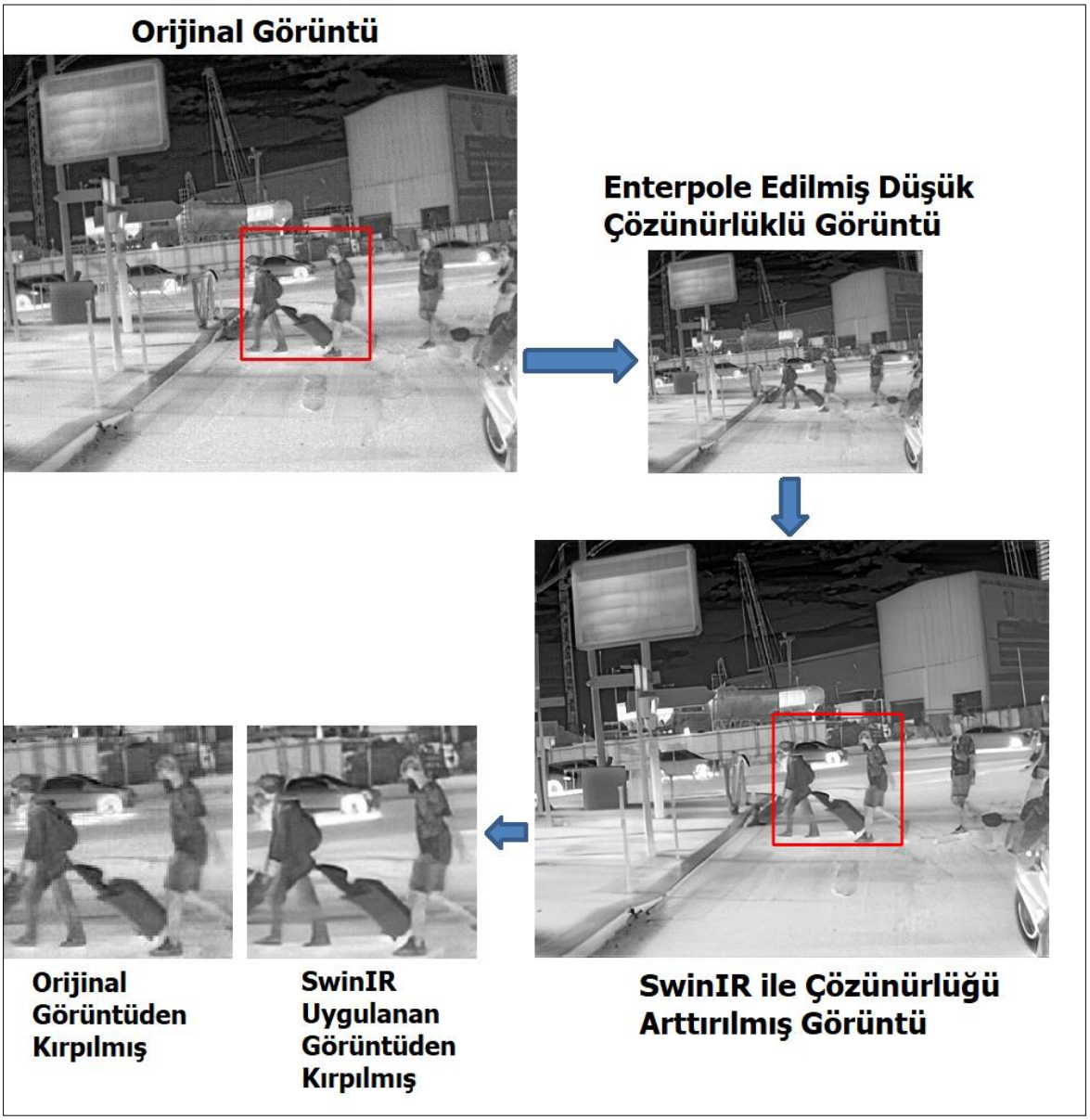
Yüksek kaliteli görüntü yapılandırması

Son modül, iyileştirilmiş özellikleri kullanarak yüksek kaliteli görüntüyü yeniden yapılandırır. Bu adım, iyileştirilmiş görüntüyü üretmek için önceki aşamalardan elde edilen çıktıları birleştirir. Yüksek kaliteli görüntü yeniden oluşturulduktan sonra SwinIR tipik olarak, L1 kayıp fonksiyonunu (tahmin edilen değer ile gerçek değer arasındaki farkın mutlak değeri bir nevi MAE) kullanır.

SwinIR, görüntüleri küçük parçalara bölerek ve bu parçaları kaydırarak hem ince detayları hem de genel yapıları doğru bir şekilde temsil eder [10, 37]. Bu yöntem, farklı ölçeklerdeki ayrıntıları yakalamayı mümkün kılar. Artık öğrenme süreci eğitimi hızlandırarak modeli

daha verimli hale getirirken, sığ ve derin özellik çıkarma kombinasyonu hem temel hem de karmaşık ayrıntıları yakalayarak üstün görüntü kalitesi sunmayı amaçlar [10].

Orijinal görüntüden bikübik enterpolasyon ile oluşturulan düşük çözünürlüklü görüntüye SwinIR modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlem sonucunda elde edilen görüntü, Şekil 6.13'te sunulmuştur.



Şekil 6.13. SwinIR ile çözünürlüğü iyileştirilmiş termal görüntü

6.2.5. Diğer yöntemler

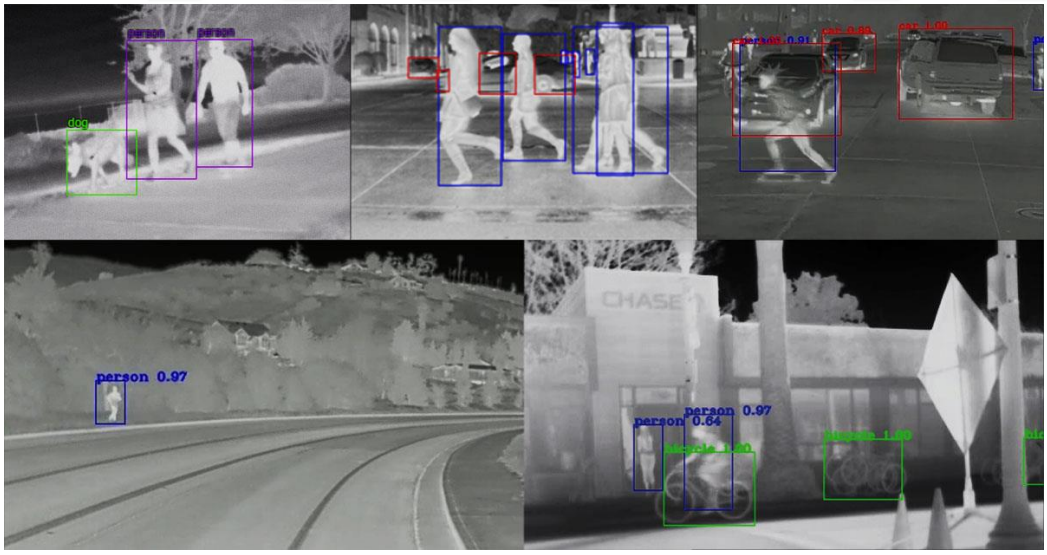
VDSR, SRGAN, ESRGAN ve SwinIR'e ek olarak, literatürde tartışılan başka derin öğrenme tabanlı süper çözünürlük yöntemleri de vardır. Bunlar arasında Süper Çözünürlüklü Evrişimsel Sinir Ağı (Super Resolution Convolutional Neural Network, SRCNN), Tek Görüntü Süper Çözünürlüğü için Geliştirilmiş Derin Artık Ağları (Enhanced Deep Residual Network, EDSR) da bulunmaktadır [27, 32]. Bu yöntemler süper çözünürlük tekniklerinin gelişimine önemli katkılar sağlamış olsa da bu tez kapsamında detaylı olarak incelenmemiştir. Bunun temel nedeni, yenilikçi yaklaşımlarına rağmen, VDSR, SRGAN, ESRGAN ve SwinIR modelleri kadar başarılı veya etkili sonuçlar elde edememeleridir. Sonuç olarak, tez kapsamında daha etkili modellere odaklanılarak, görüntü süper çözünürlüğündeki en son teknolojinin daha net ve özlü bir şekilde değerlendirilmesini amaçlanmıştır.

7. DENEYSEL SONUÇLAR

Bu bölüm altında, veri kümesi ve görüntü değerlendirme metrikleri ayrıntılı olarak açıklanmıştır. Ayrıca, Bikübik Enterpolasyon gibi geleneksel süper çözünürlük yöntemlerinin yanı sıra VDSR, SRGAN, ESRGAN ve SwinIR gibi derin öğrenme tabanlı süper çözünürlük modelleri de veri kümemizdeki termal görüntülere uygulanmıştır. Bu yöntemlerin karşılaştırma sonuçları kapsamlı bir şekilde sunulmuştur.

7.1. Veri Seti

Renkli görüntüler için literatürde geniş veri seti bulunurken, termal görüntüler için uygun veri seti sayısı kısıtlıdır. Bu tezde, eğitim ve testlerde kullanmak üzere FLIR'ın veri seti olan Teledyne FLIR Tau 2 640 x 512 çözünürlüğündeki soğutmasız bir dedektör kullanan bir termal kameradan alınan ve Şekil 7.1'de sunulan FLIR Termal Veri Seti modellere girdi olarak verilmiştir [41]. Eğitim için 1000 adet, doğrulama (validasyon) için ise 150 adet 640x512 çözünürlüğe sahip 8 bitlik termal görüntü bulunmaktadır.



Şekil 7.1. Termal görüntü veri seti [41]

Modele verilmek üzere eğitim ve validasyon setleri bikübik ve bilineer enterpolasyonla 160x128 çözünürlüklü görüntüler haline getirilmiştir. 256x128 çözünürlüğe sahip görüntüler düşük çözünürlüklü, 640x512 çözünürlüklü Şekiller ise yüksek çözünürlüklü görüntüler olacak şekilde eğitim ve validasyon veri seti olarak kullanılmıştır.

7.2. Değerlendirme Metrikleri

Süper çözünürlük teknikleri, çözünürlüğü arttırılmış görüntülerin kalitesini kapsamlı bir şekilde değerlendirmeyi amaçlar. Bu bölüm altında da bu amaçla referans gerektiren (full-reference) değerlendirme metriklerinden PNSR ve SSIM incelenirken referans gerektirmeyen (no-reference) metriklerden ise NIQE ve BRISQUE incelenmiştir.

7.2.1. PSNR

Tepe Sinyal Gürültü Oranı (Peak Signal to Noise Ratio, PSNR) görüntüler arasındaki aslına uygunluk ve yapısal benzerlikler hakkında bilgi sağlar [93]. PSNR, bir sinyalin maksimum potansiyel gücü ile sinyalin gücünü etkileyen bozucu gürültü arasındaki oranı ölçer. Sıkıştırılmış ve yeniden yapılandırılmış görüntülerin kalitesini değerlendirmek için sıklıkla kullanılır ve birimi desibeldir (dB) [93]. Daha yüksek PSNR değerleri daha iyi kaliteyi (daha az gürültü) gösterir. PSNR formülü Eşitlik 7.1'de verilmiştir.

$$PSNR = 10\, log_{10} \frac{MAX^2}{MSE} \tag{7.1}$$

Burada MAX orijinal görüntünün piksel değerlerinin maksimum değeridir. MSE (Mean Squared Error), Bölüm 5.2.2'nin de altında bahsedildiği gibi piksel düzeyinde orijinal ve iyileştirilmiş görüntüler arasındaki ortalama karesel hata miktarını temsil eder ve önceki bölümlerde yer alan Eşitlik 5.3'teki gibi ifade edilir.

7.2.2. SSIM

Yapısal Benzerlik İndeksi Ölçüsü (Structural Similarity Index Measure, SSIM), iki görüntü arasındaki benzerliği değerlendirmek için kullanılan bir metriktir [93]. Ortalama Karesel Hata (MSE) veya Tepe Sinyal-Gürültü Oranı (PSNR) gibi yalnızca piksel-piksel farklılıkları ölçen geleneksel yöntemlerin aksine, SSIM yapısal bilgi, parlaklık ve kontrasttaki değişiklikleri dikkate alarak algılanan görsel kaliteyi hesaba katar [93]. SSIM değerleri -1 ila 1 arasındadır. 1 değeri, iki görüntü arasındaki mükemmel benzerliği gösterir. 0'a yakın veya negatif değerler benzerliğin düşük olduğunu gösterir.

SSIM üç bileşene göre hesaplanır: Parlaklık (l), iki görüntü arasındaki parlaklık benzerliğini ölçer. Kontrast (c), iki görüntü arasındaki kontrastın benzerliğini ölçer. Yapılar, görüntülerdeki yapıların benzerliğini ölçer. Bu bileşenlerin her biri genel SSIM endeksine katkıda bulunur. SSIM aşağıdaki eşitlikte verilmiştir.

$$SSIM(x,y) = [l(x,y)]^{\alpha} . [c(x,y)]^{\beta} . [s(x,y)]^{\gamma} \quad (7.2)$$

Burada l parlaklığı, c kontrastı ifade ederken s ise yapısal benzerliği ifade etmektedir.

7.2.3. NIQE

Herhangi bir referans görüntüye ihtiyaç duymayan Natural Image Quality Evaluator (NIQE) yani Doğal Görüntü Kalite Değerlendiricisi, doğal sahne istatistiklerinden sapmalara dayalı olarak bir görüntünün kalitesini ölçen referanssız görüntü kalitesi değerlendirme metriğidir [94]. Doğal görüntülerin, görüntü bozulduğunda değişen belirli istatistiksel özelliklere sahip olduğu varsayımıyla çalışır.

Doğal Sahne İstatistikleri (NSS), doğal görüntülerden çıkarılan istatistiksel özelliklerdir. NIQE bu özellikleri doğal görüntülerin dağılımını modellemek için kullanır [94, 95]. Ayrıca NIQE, yüksek kaliteli görüntülerden NSS dağılımını temsil etmek için bir Çok Değişkenli Gauss (MVG) Modeli kullanır [94]. NIQE değerlendirme metriğinin formülü aşağıda Eşitlik 7.3'de verilmiştir.

$$NIQE(I) = \sqrt{(\mu_I - \mu_N)^T (\frac{\sum_I + \sum_N}{2})^{-1} (\mu_I - \mu_N)} \quad (7.3)$$

Bir I görüntüsü verildiğinde, $\mu_N$ve $\sum_N$ yüksek kaliteli doğal görüntülerden oluşan bir modelden elde edilen ortalama ve kovaryans matrisini, $\mu_I$ve μ_N ise test görüntüsünün özelliklerinden hesaplanan ortalama ve kovaryans matrisini temsil etsmektedir. NIQE puanı, test görüntüsünün Doğal Sahne İstatistikleri ile model arasındaki istatistiksel mesafe olarak hesaplanır ve test görüntüsünün doğal görüntü modelinden ne kadar saptığını gösterir [96]. Daha düşük bir NIQE puanı, test görüntüsünün istatistiksel özelliklerinin doğal görüntülere daha yakın olduğunu gösterdiğinden daha yüksek görüntü kalitesine işaret eder.

7.2.4. BRISQUE

Orijinal görüntü bilgilerine başvurmadan görüntünün özelliklerinin doğrusal olmayan istatistiklerini kullanarak kalite değerlendirmesi yapan BRISQUE Blind/Referenceless Image Spatial Quality Evaluator) yani Kör/Referanssız Görüntü Mekansal Kalite Değerlendiricisi, görüntülerin kalitesini mekansal alan doğal sahne istatistiklerine dayalı olarak değerlendiren başka bir referanssız görüntü kalitesi değerlendirme ölçümüdür [94, 97, 98].

NIQE'den farklı olarak BRISQUE, bilinen kalite puanlarına sahip bozuk görüntüler üzerinde eğitilmiş bir Support Vector Regression (SVR) yani destek vektör regresyon modeli kullanır [97, 98]. BRISQUE değerlendirme metriğinin formülü aşağıda Eşitlik 7.4'de verilmiştir.

$$BRISQUE(I) = g(f) \tag{7.4}$$

Eşitlik 7.4'teki I, değerlendirilen görüntüyü temsil eder. Kalite puanını belirlemek istediğimiz giriş görüntüsüdür. I görüntüsü, algısal kalitesini yansıtan ilgili özellikleri çıkarmak için ön işleme tabi tutulur.

f, I görüntüsünden çıkarılan özellik vektörüdür. f özellik vektörü, görüntünün uzamsal alanından türetilen istatistiksel özelliklerden oluşur. Bu özellikler Mean Subtracted Contrast Normalized (MSCN) yani Ortalama Çıkarılmış Kontrast Normalleştirilmiş katsayılarını ve bu katsayıların çeşitli yönlerde (yatay, dikey ve çapraz) ikili çarpımlarını içerir [99]. Çıkarılan özellikler, görüntünün doğal sahne istatistiklerini ve yapısal bilgilerini yakalamak için tasarlanmıştır.

g, BRISQUE puanını tahmin etmek için kullanılan önceden eğitilmiş Destek Vektör Regresyonu (SVR) modelini gösterir. SVR modeli g, bilinen kalite puanlarına sahip görüntülerden oluşan büyük bir veri kümesi üzerinde eğitilir ve çıkarılan özellikler f ile görüntülerin algısal kalitesi arasındaki ilişkiyi öğrenmesini sağlar [99]. f özellik vektörü SVR modeline g girildiğinde, I görüntüsünün algısal kalitesini temsil eden bir kalite puanı çıkarır.

Özetlemek gerekirse, I giriş görüntüsüdür. f, I'dan çıkarılan özellik vektörüdür. g, f

özelliklerine dayalı olarak BRISQUE kalite puanını tahmin eden SVR modelidir. Yüksek BRISQUE skoru görüntünün kalitesiz ve daha fazla bozulduğunu gösterirken, düşük BRISQUE skoru, görüntünün kaliteli ve bozulmaların az olduğunu yani görüntünün insan gözüyle algılanan doğal görüntülere daha yakın olduğunu belirtir.

7.3. Model Karşılaştırmaları

Bu bölümde, termal görüntülerden oluşan veri setine geleneksel süper çözünürlük yöntemlerinden bikübik enterpolasyon, derin öğrenme tabanlı süper çözünürlük yöntemlerinden CNN, GAN ve Swin mimarileri kullanan ve FLIR termal veri setiyle eğitilmiş VDSR, SRGAN, ESRGAN ve SwinIR modelleri uygulanarak elde edilen sonuçlar paylaşılmıştır. Modeller için oluşturulan düşük çözünürlüklü görüntülerin detayları Bölüm 7.1 altında verilmiştir. Ayrıca tüm büyütme değerleri dört kat olarak uygulanmıştır.

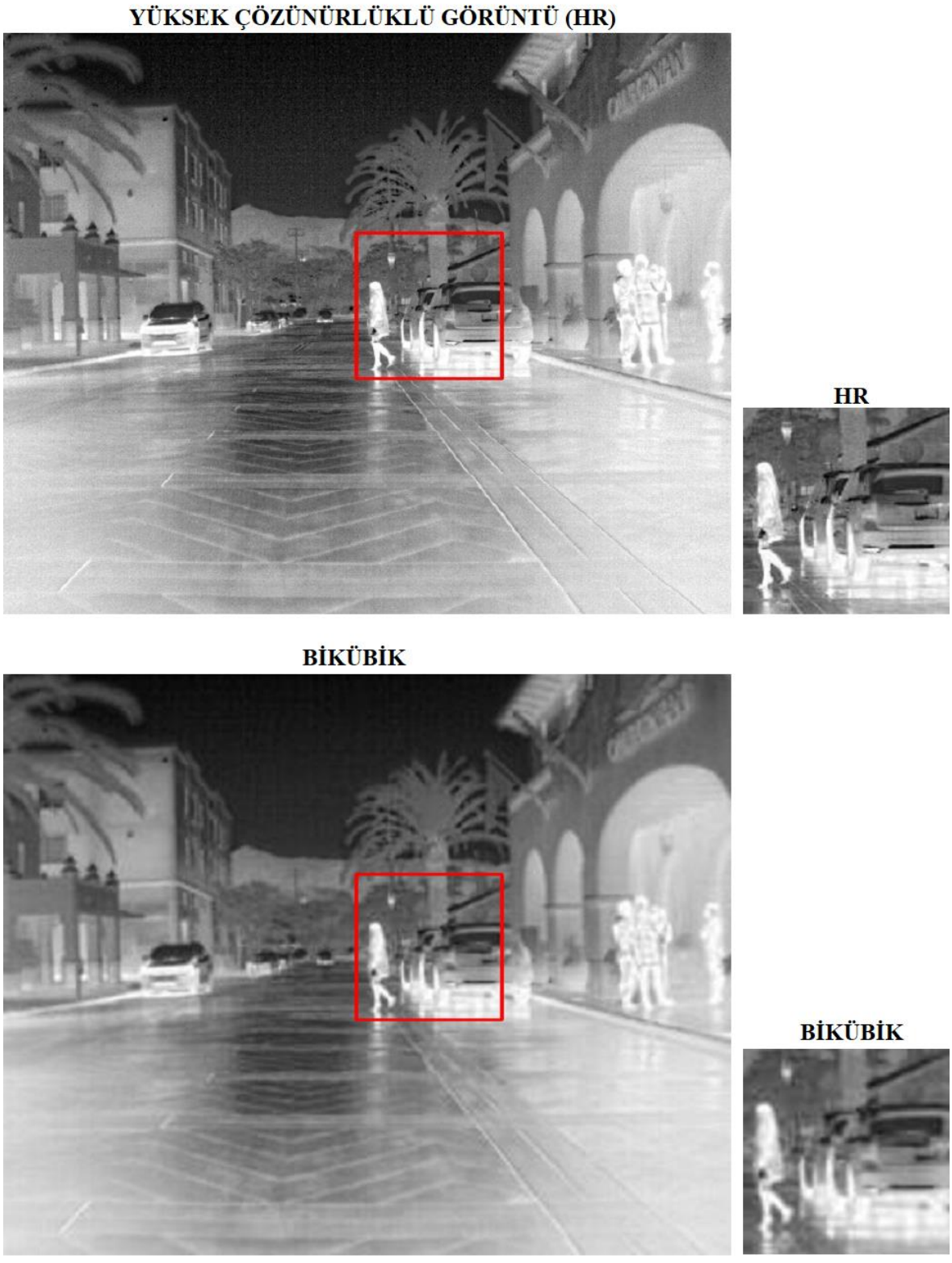
7.3.1. Bikübik Enterpolasyon

Bu bölümde bilineer ve bikübik enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntülere Bikübik Enterpolasyon yöntemi ile 4 kat üst örnekleme yapılmış ve görüntülerin başarı metriklerinin karşılaştırması sunulmuştur.

Bilineer alt örneklemeli birinci görüntünün çıktıları

Bilineer Enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntüye Bikübik Enterpolasyon tekniği uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.2'de gösterilmiştir.

Çizelge 7.1'de, Bikübik Enterpolasyon yöntemi ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da enterpole edilmiş görüntü için çizelgede yer verilmiştir.



Şekil 7.2. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, bikübik enterpolasyon ile üst örnekleme uygulanmış hali

Çizelge 7.1. Bilineer enterpolasyon ile alt örnekleme yapılmış görüntünün bikübik enterpolasyon yöntemi sonucunda orijinal görüntü ile karşılaştırılması

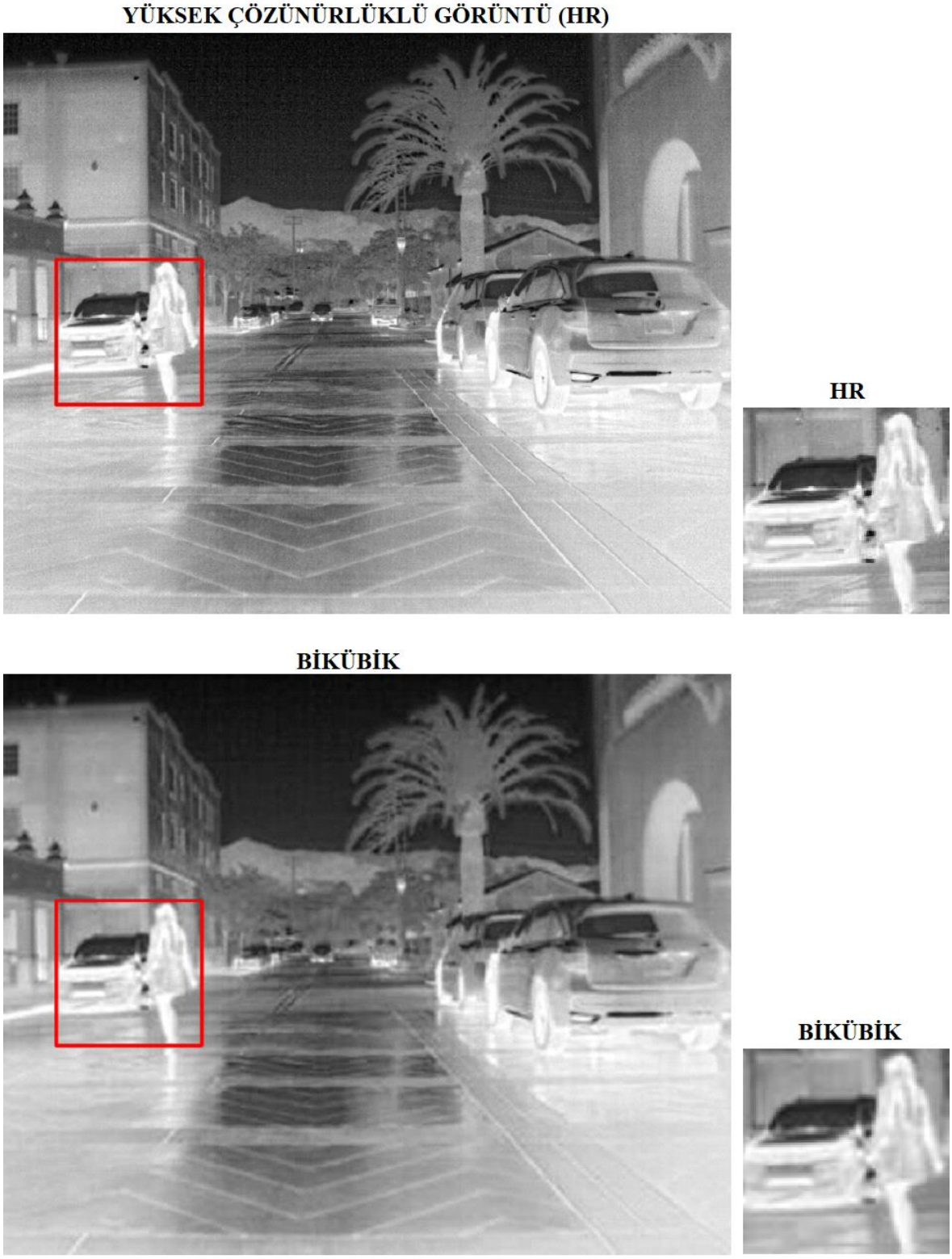
Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
Bikübik	4x	30.12 dB	0.740	4.51	51.40

Çizelge 7.1 ve Şekil 7.2 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.12 dB, SSIM değeri ise 0.74 olarak hesaplanmıştır. Enterpole edilmiş görüntünün NIQE skoru 4.51 iken BRISQUE skoru ise 51.40 olarak ölçülmüştür.

Bilineer alt örneklemeli ikinci görüntünün çıktıları

Bilineer Enterpolasyon ile 4 kat alt örnekleme yapılmış birinci görüntüden farklı bir düşük çözünürlüklü görüntüye de Bikübik Enterpolasyon tekniği uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.3'de gösterilmiştir.

Çizelge 7.2'de, Bikübik Enterpolasyon yöntemi ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da enterpole edilmiş görüntü için çizelgede yer verilmiştir.



Şekil 7.3. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, bikübik enterpolasyon ile üst örnekleme uygulanmış hali

Çizelge 7.2. Bilineer enterpolasyon ile alt örnekleme yapılmış görüntünün bikübik enterpolasyon yöntemi sonucunda orijinal görüntü ile karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
Bikübik	4x	30.32 dB	0.737	4.40	47.50

Çizelge 7.2 ve Şekil 7.3 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.32 dB, SSIM değeri ise 0.737 olarak hesaplanmıştır. Enterpole edilmiş görüntünün NIQE skoru 4.40 iken BRISQUE skoru ise 47.50 olarak ölçülmüştür.

Bikübik alt örneklemeli birinci görüntünün çıktıları

Bikübik Enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntüye Bikübik Enterpolasyon tekniği uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.4'de gösterilmiştir.

Çizelge 7.3'de, Bikübik Enterpolasyon yöntemi ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da enterpole edilmiş görüntü için çizelgede yer verilmiştir.



Şekil 7.4. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, bikübik enterpolasyon ile üst örnekleme uygulanmış hali

Çizelge 7.3. Bikübik enterpolasyon ile alt örnekleme yapılmış görüntünün bikübik enterpolasyon yöntemi sonucunda orijinal görüntü ile karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
Bikübik	4x	30.72	0.750	4.34	47.09

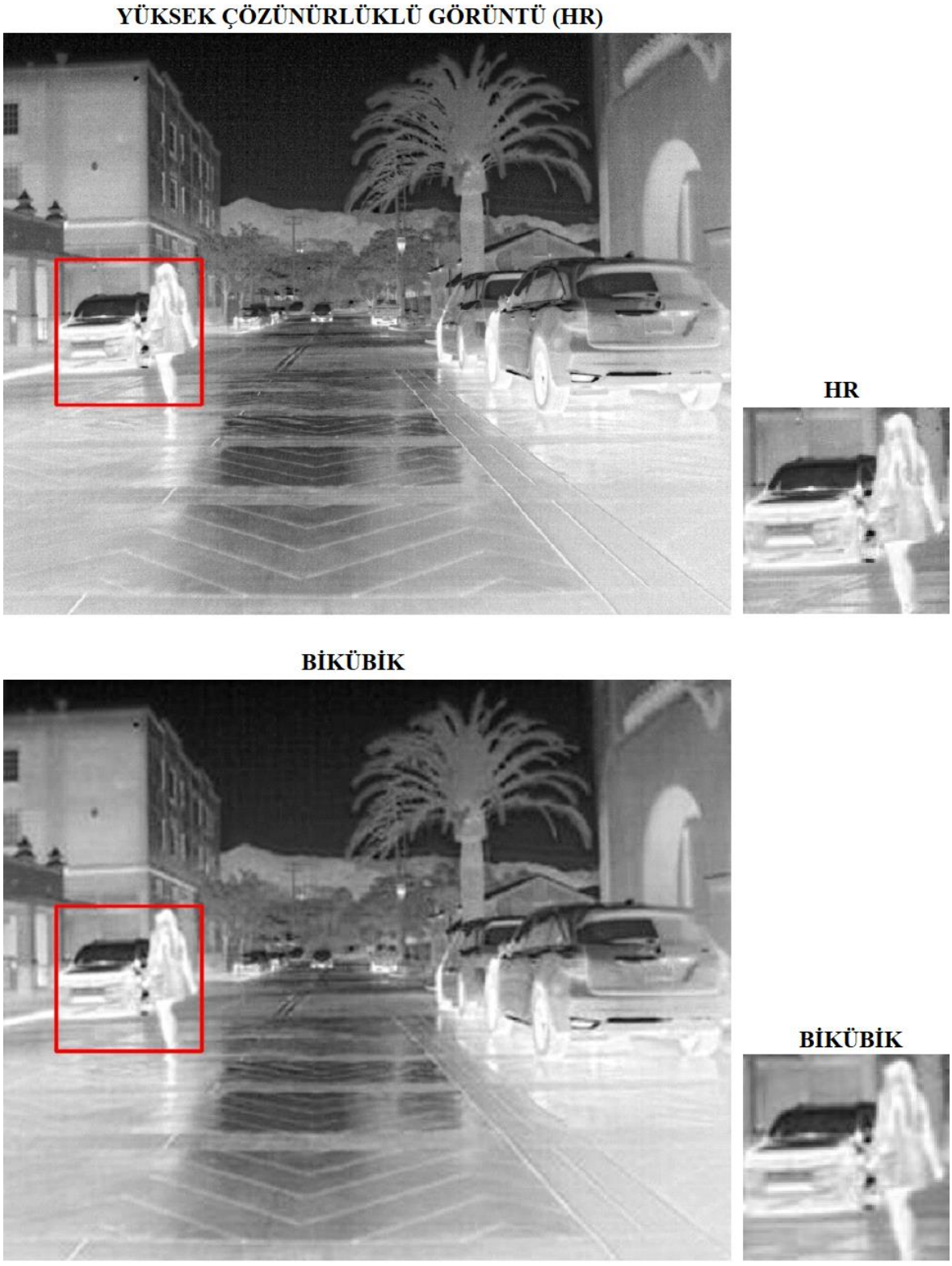
Çizelge 7.3 ve Şekil 7.4 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.72 dB, SSIM değeri ise 0.750 olarak hesaplanmıştır. Enterpole edilmiş görüntünün NIQE skoru 4.34 iken BRISQUE skoru ise 47.09 olarak ölçülmüştür.

Girdi olarak bilineer enterpolasyon yerine bikübik ile alt örnekleme yapılmış görüntü verildiğinde, bikübik enterpolasyon yönteminin PSNR, NIQE ve BRISQUE metriklerinde daha yüksek başarıma sahip olduğu ve SSIM metriğinde ise sonucun hemen hemen aynı kaldığı görülmüştür.

Bikübik alt örneklemeli ikinci görüntünün çıktıları

Bikübik Enterpolasyon ile 4 kat alt örnekleme yapılmış birinci görüntüden farklı bir düşük çözünürlüklü görüntüye de Bikübik Enterpolasyon tekniği uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.5'de gösterilmiştir.

Çizelge 7.4'de, Bikübik Enterpolasyon yöntemi ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da enterpole edilmiş görüntü için çizelgede yer verilmiştir.



Şekil 7.5. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, bikübik enterpolasyon ile üst örnekleme uygulanmış hali

Çizelge 7.4. Bikübik enterpolasyon ile alt örnekleme yapılmış görüntünün bikübik enterpolasyon yöntemi sonucunda orijinal görüntü ile karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
Bikübik	4x	30.45	0.743	4.03	46.54

Çizelge 7.4 ve Şekil 7.5 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.45 dB, SSIM değeri ise 0.743 olarak hesaplanmıştır. Enterpole edilmiş görüntünün NIQE skoru 4.03 iken BRISQUE skoru ise 46.54 olarak ölçülmüştür.

Girdi olarak bilineer enterpolasyon yerine bikübik ile alt örnekleme yapılmış görüntü verildiğinde, bikübik enterpolasyon yönteminin PSNR, NIQE ve BRISQUE metriklerinde daha yüksek başarıma sahip olduğu ve SSIM metriğinde ise sonucun hemen hemen aynı kaldığı görülmüştür.

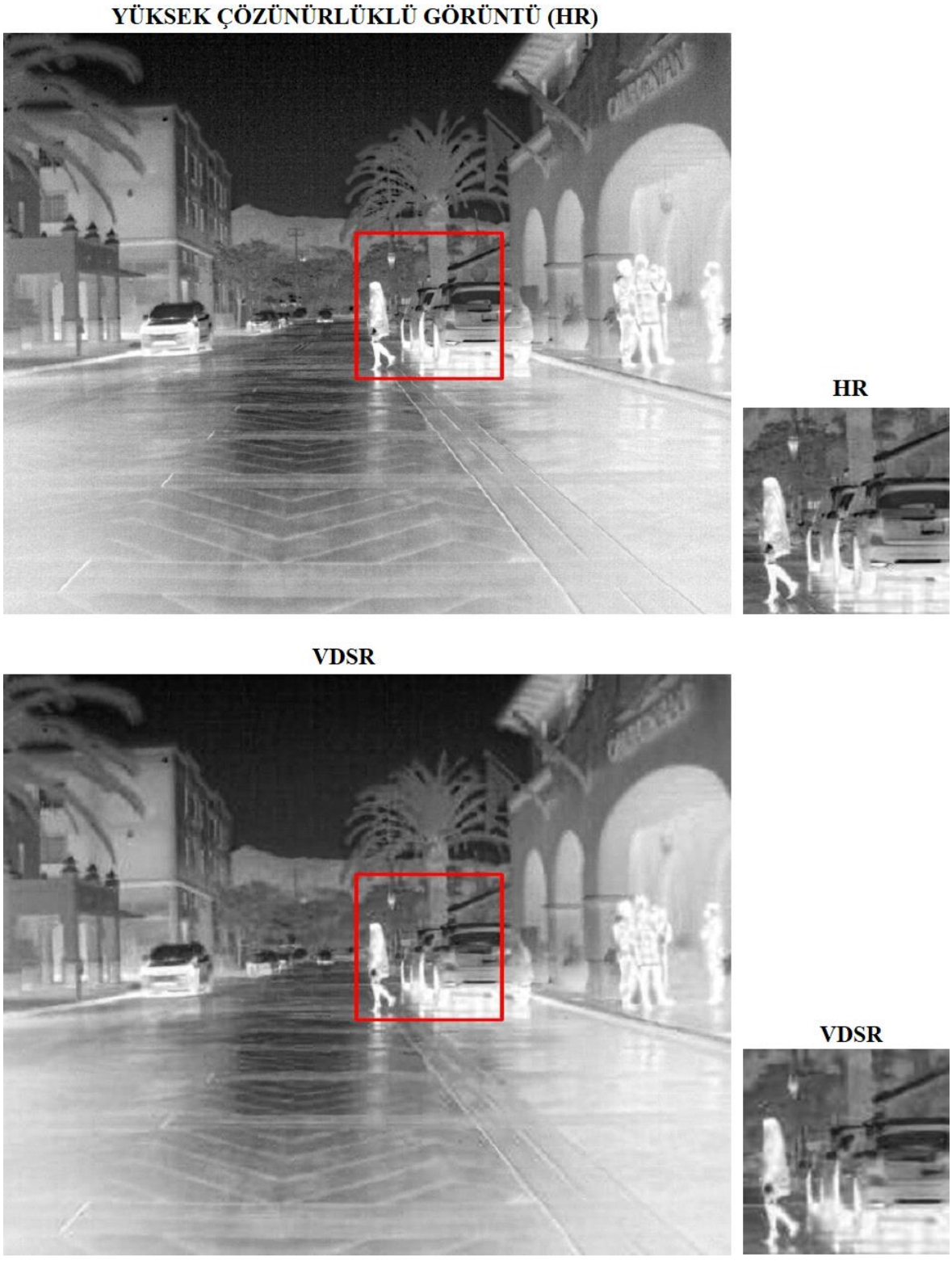
7.3.2. VDSR

Bu bölümde bilineer ve bikübik enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntülere VDSR modeli uygulanarak görüntülerin başarı metriklerinin karşılaştırması sunulmuştur.

Bilineer alt örneklemeli birinci görüntünün çıktıları

Bilineer Enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntüye VDSR modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.6'da gösterilmiştir.

Çizelge 7.5'de, VDSR modeli ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da VDSR uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.6. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, VDSR ile üst örnekleme uygulanmış hali

Çizelge 7.5. Bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün VDSR sonucunda orijinal görüntü ile karşılaştırılması

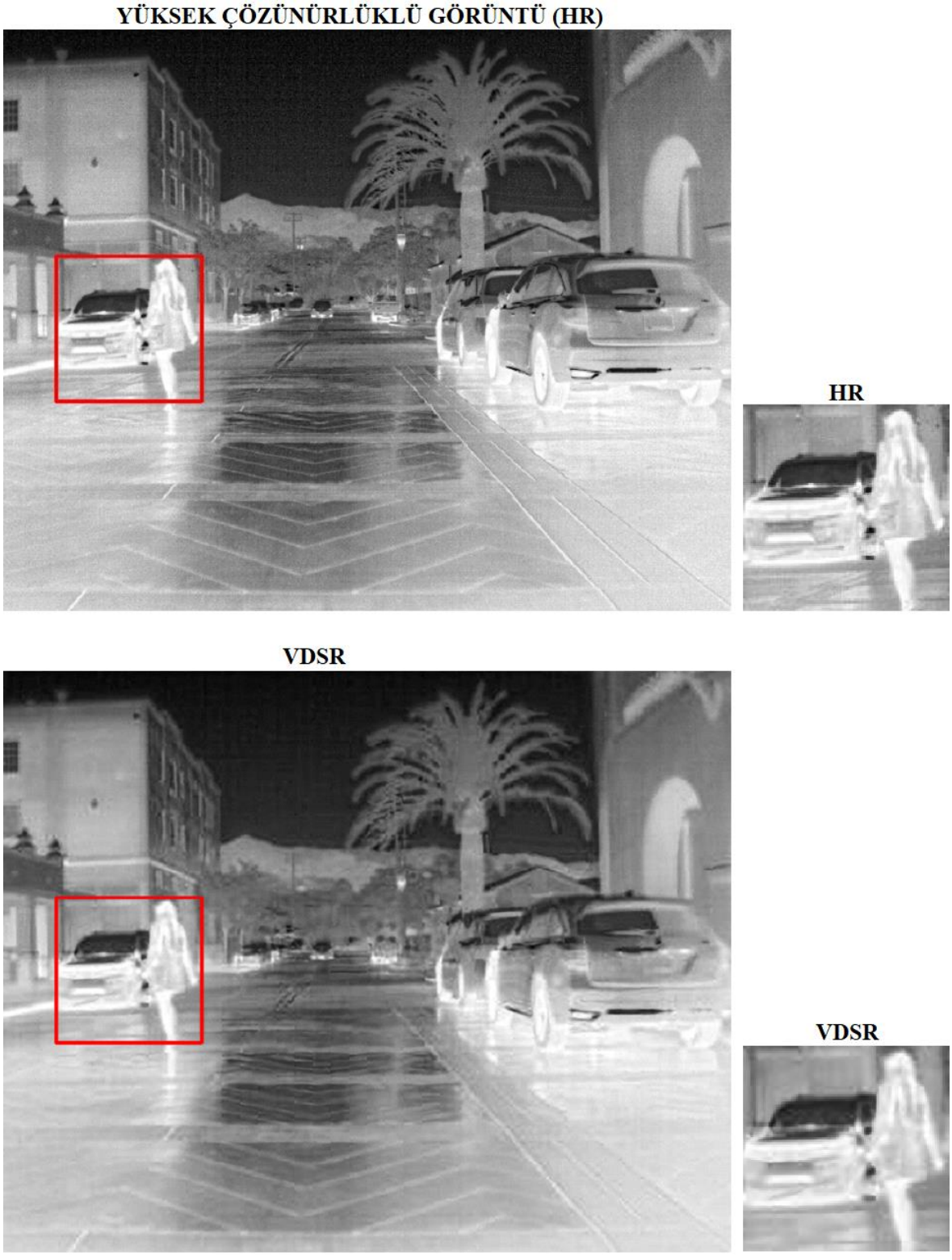
Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
VDSR	4x	30.819 dB	0.741	4.22	45.69

Çizelge 7.5 ve Şekil 7.6 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.819 dB, SSIM değeri ise 0.741 olarak hesaplanmıştır. Enterpole edilmiş görüntünün NIQE skoru 4.22 iken BRISQUE skoru ise 45.69 olarak ölçülmüştür.

Bilineer alt örneklemeli ikinci görüntünün çıktıları

VDSR ile 4 kat alt örnekleme yapılmış birinci görüntüden farklı bir düşük çözünürlüklü görüntüye de VDSR modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.7'de gösterilmiştir.

Çizelge 7.6'de, Bikübik Enterpolasyon yöntemi ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da VDSR uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.7. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, VDSR ile üst örnekleme uygulanmış hali

Çizelge 7.6. Bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün VDSR sonucunda orijinal görüntü ile karşılaştırılması

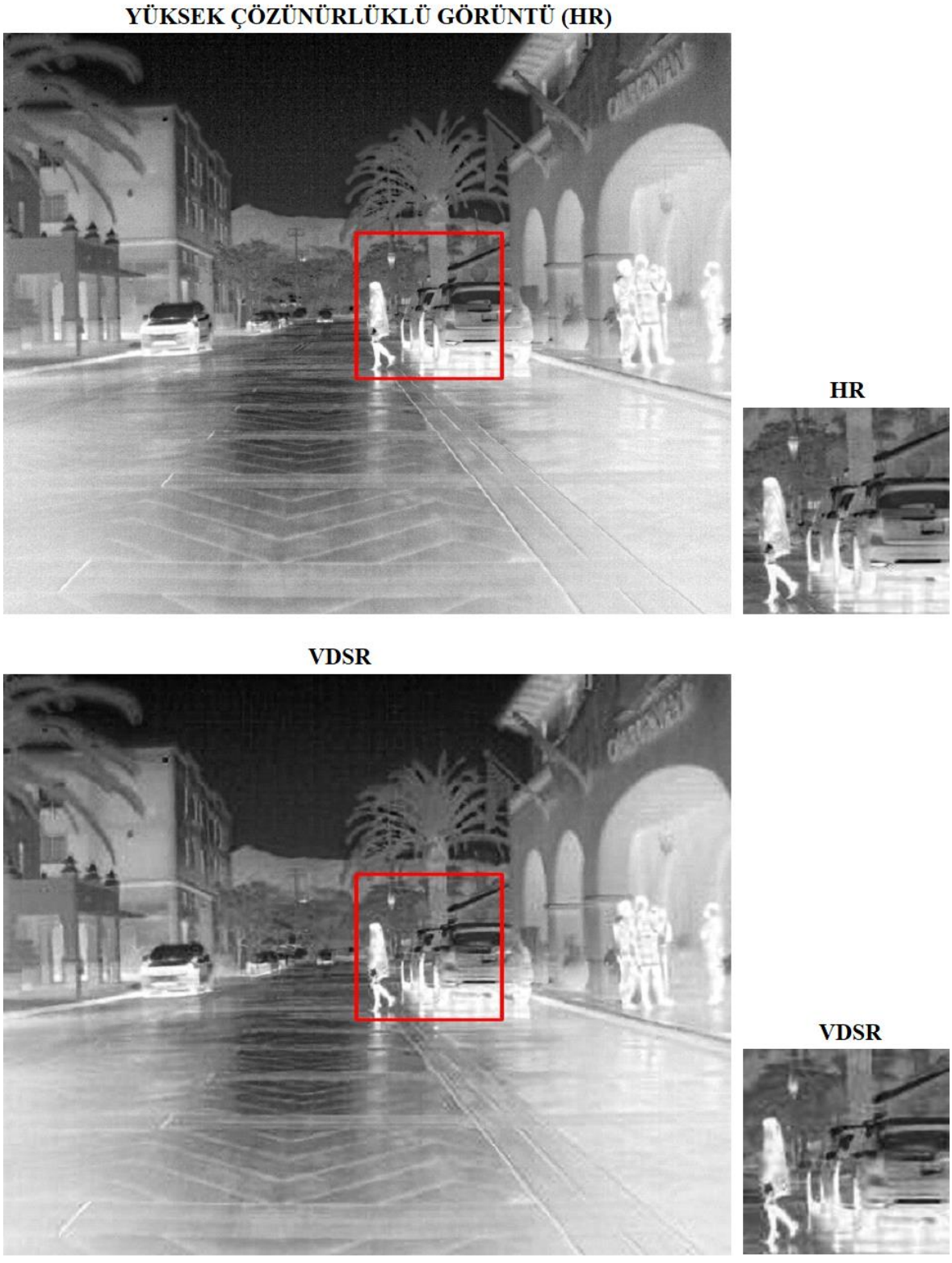
Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
VDSR	4x	30.481	0.739	4.07	38.56

Çizelge 7.6 ve Şekil 7.7 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.481 dB, SSIM değeri ise 0.739 olarak hesaplanmıştır. Enterpole edilmiş görüntünün NIQE skoru 4.07 iken BRISQUE skoru ise 38.56 olarak ölçülmüştür.

Bikübik alt örneklemeli birinci görüntünün çıktıları

Bikübik Enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntüye VDSR modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.8’de gösterilmiştir.

Çizelge 7.7'de, VDSR modeli ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da VDSR uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.8. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, VDSR ile üst örnekleme uygulanmış hali

Çizelge 7.7. Bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün VDSR sonucunda orijinal görüntü ile karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
VDSR	4x	30.98	0.752	4.06	38.92

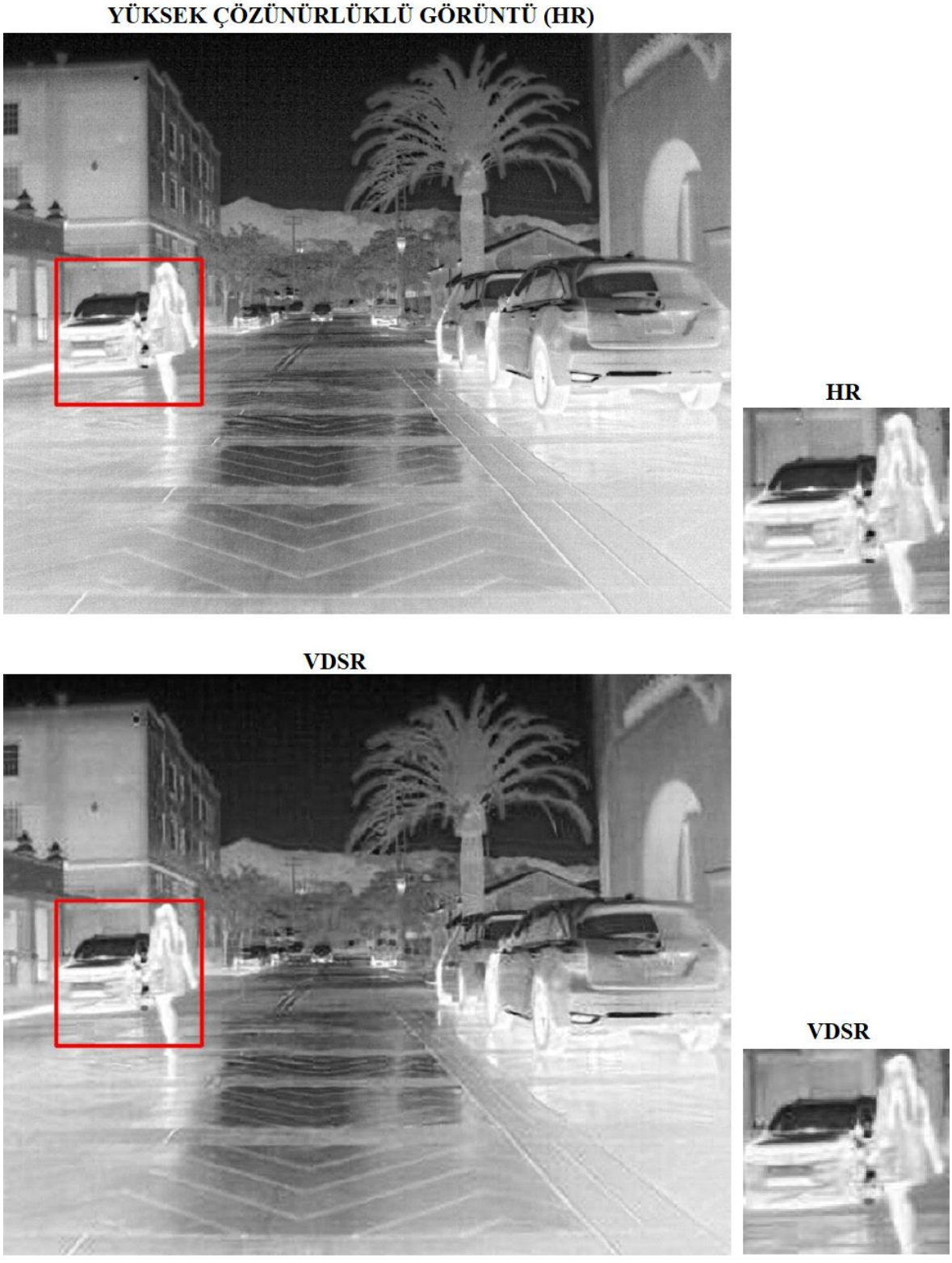
Çizelge 7.7 ve Şekil 7.8 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.98 dB, SSIM değeri ise 0.752 olarak hesaplanmıştır. Enterpole edilmiş görüntünün NIQE skoru 4.06 iken BRISQUE skoru ise 38.92 olarak ölçülmüştür.

Girdi olarak bilineer enterpolasyon yerine bikübik ile alt örnekleme yapılmış görüntü verildiğinde, bikübik enterpolasyon yönteminin tüm metriklerde daha yüksek başarıma sahip olduğu görülmüştür.

Bikübik alt örneklemeli ikinci görüntünün çıktıları

VDSR ile 4 kat alt örnekleme yapılmış birinci görüntüden farklı bir düşük çözünürlüklü görüntüye de VDSR modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.9'de gösterilmiştir.

Çizelge 7.8'de, Bikübik Enterpolasyon yöntemi ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da VDSR uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.9. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, VDSR ile üst örnekleme uygulanmış hali

Çizelge 7.8. Bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün VDSR sonucunda orijinal görüntü ile karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
VDSR	4x	30.51	0.745	3.78	37.36

Çizelge 7.8 ve Şekil 7.9 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.51 dB, SSIM değeri ise 0.745 olarak hesaplanmıştır. Enterpole edilmiş görüntünün NIQE skoru 3.78 iken BRISQUE skoru ise 37.36 olarak ölçülmüştür.

Girdi olarak bilineer enterpolasyon yerine bikübik ile alt örnekleme yapılmış görüntü verildiğinde, bikübik enterpolasyon yönteminin tüm metriklerde daha yüksek başarıma sahip olduğu görülmüştür.

7.3.3. SRGAN

Bu bölümde bilineer ve bikübik enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntülere SRGAN modeli uygulanarak görüntülerin başarı metriklerinin karşılaştırması sunulmuştur.

Bilineer alt örneklemeli birinci görüntünün çıktıları

Bilineer Enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntüye SRGAN modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.10'da gösterilmiştir.

Çizelge 7.9'da, SRGAN modeli ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da SRGAN uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.10. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, SRGAN ile üst örnekleme uygulanmış hali

Çizelge 7.9. Bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün SRGAN sonucunda orijinal görüntü ile karşılaştırılması

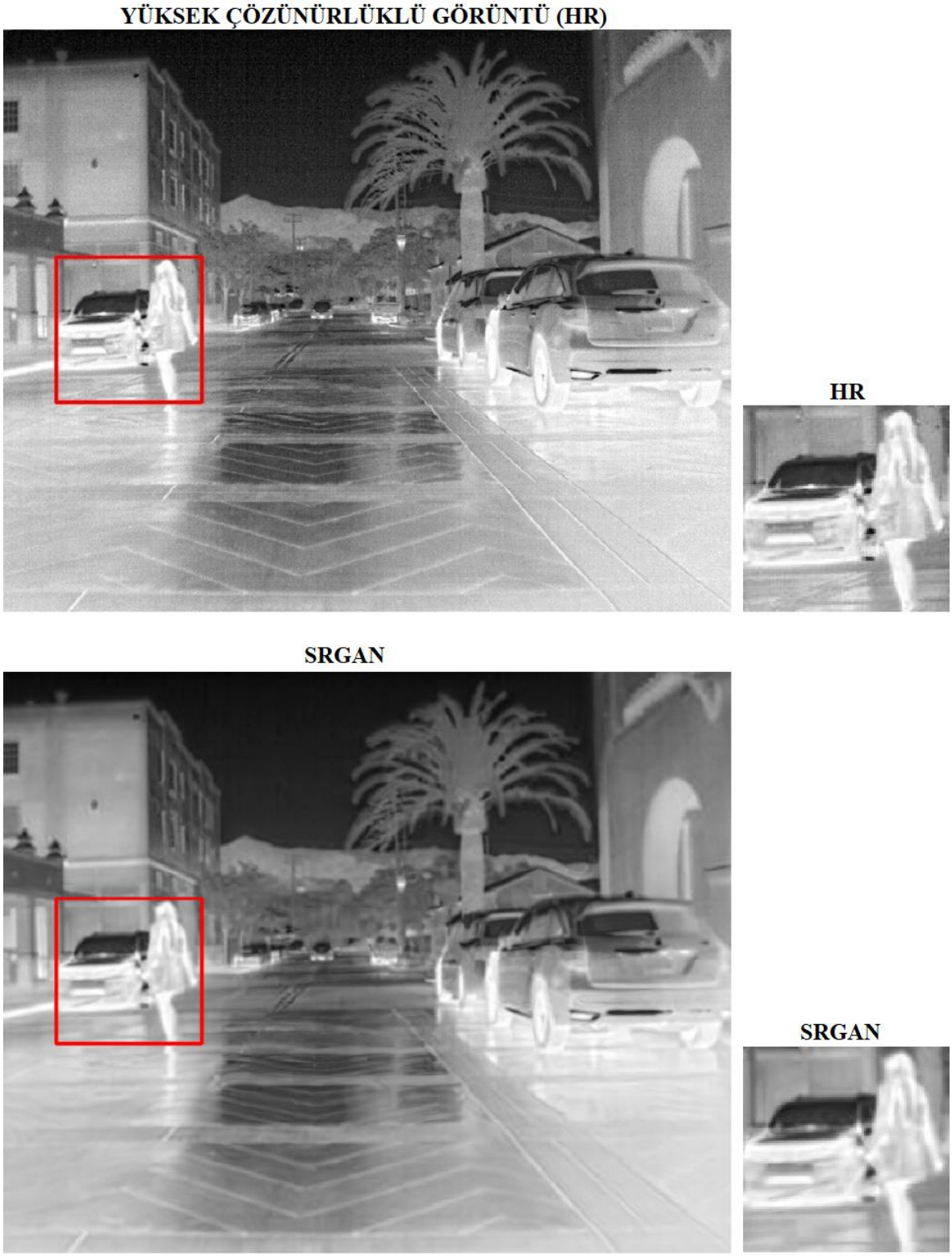
Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
SRGAN	4x	30.798 dB	0.742	4.16	33.63

Çizelge 7.9 ve Şekil 7.10 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.798 dB, SSIM değeri ise 0.742 olarak hesaplanmıştır. SRGAN uygulanmış görüntünün NIQE skoru 4.16 iken BRISQUE skoru ise 33.63 olarak ölçülmüştür.

Bilineer alt örneklemeli ikinci görüntünün çıktıları

Bilineer ile 4 kat alt örnekleme yapılmış birinci görüntüden farklı bir düşük çözünürlüklü görüntüye de SRGAN modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.11'de gösterilmiştir.

Çizelge 7.10'da, SRGAN modeli ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da SRGAN uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.11. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, SRGAN ile üst örnekleme uygulanmış hali

Çizelge 7.10. Bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün VDSR sonucunda orijinal görüntü ile karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
SRGAN	4x	30.70	0.750	3.40	32.57

Çizelge 7.10 ve Şekil 7.11 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.70 dB, SSIM değeri ise 0.750 olarak hesaplanmıştır. SRGAN uygulanmış görüntünün NIQE skoru 3.40 iken BRISQUE skoru ise 32.57 olarak ölçülmüştür.

Bikübik alt örneklemeli birinci görüntünün çıktıları

Bikübik Enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntüye VDSR modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.12'de gösterilmiştir.

Çizelge 7.11'de, SRGAN modeli ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da SRGAN uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.12. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, SRGAN ile üst örnekleme uygulanmış hali

Çizelge 7.11. Bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün SRGAN sonucunda orijinal görüntü ile karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
Bikübik	4x	31.18	0.755	3.68	36.39

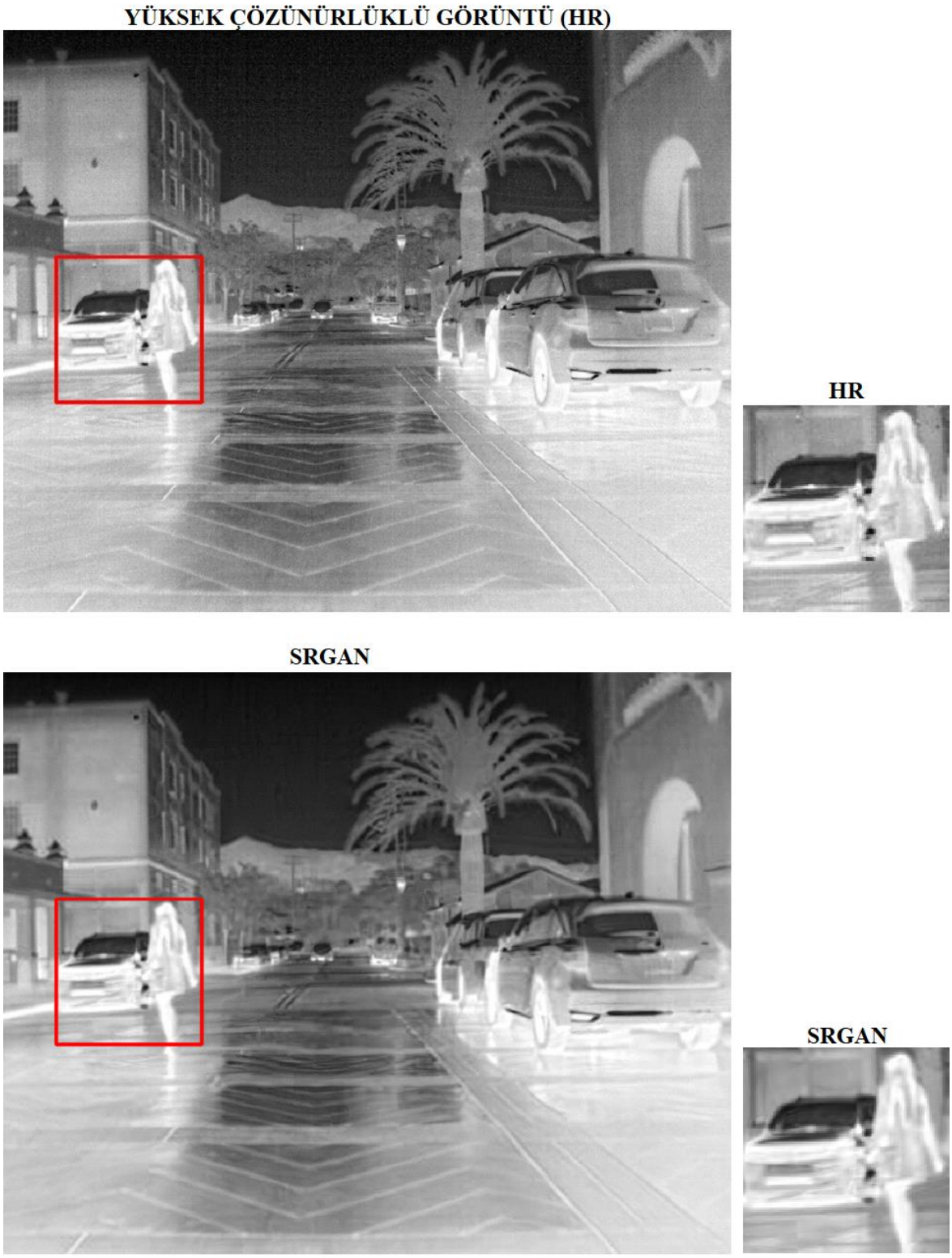
Çizelge 7.11 ve Şekil 7.12 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 31.18 dB, SSIM değeri ise 0.755 olarak hesaplanmıştır. SRGAN uygulanmış görüntünün NIQE skoru 3.68 iken BRISQUE skoru ise 36.39 olarak ölçülmüştür.

Girdi olarak bilineer enterpolasyon yerine bikübik ile alt örnekleme yapılmış görüntü verildiğinde, bikübik enterpolasyon yönteminin tüm metriklerde daha yüksek başarıma sahip olduğu görülmüştür.

Bikübik alt örneklemeli ikinci görüntünün çıktıları

Bikübik ile 4 kat alt örnekleme yapılmış birinci görüntüden farklı bir düşük çözünürlüklü görüntüye de SRGAN modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.13'de gösterilmiştir.

Çizelge 7.12'de, Bikübik Enterpolasyon yöntemi ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da SRGAN uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.13. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, SRGAN ile üst örnekleme uygulanmış hali

Çizelge 7.12. Bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün srgan sonucunda orijinal görüntü ile karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
SRGAN	4x	30.55	0.750	3.40	34.83

Çizelge 7.12 ve Şekil 7.13 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.55 dB, SSIM değeri ise 0.750 olarak hesaplanmıştır. SRGAN uygulanmış görüntünün NIQE skoru 3.40 iken BRISQUE skoru ise 34.83 olarak ölçülmüştür.

Girdi olarak bilineer enterpolasyon yerine bikübik ile alt örnekleme yapılmış görüntü verildiğinde, bikübik enterpolasyon yönteminin tüm metriklerde daha yüksek başarıma sahip olduğu görülmüştür.

7.3.4. ESRGAN

Bu bölümde bilineer ve bikübik enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntülere ESRGAN modeli uygulanarak görüntülerin başarı metriklerinin karşılaştırması sunulmuştur.

Bilineer alt örneklemeli birinci görüntünün çıktıları

Bilineer Enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntüye ESRGAN modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.14'da gösterilmiştir.

Çizelge 7.13'da, ESRGAN modeli ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da ESRGAN uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.14. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, ESRGAN ile üst örnekleme uygulanmış hali

Çizelge 7.13. Bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün ESRGAN sonucunda orijinal görüntü ile karşılaştırılması

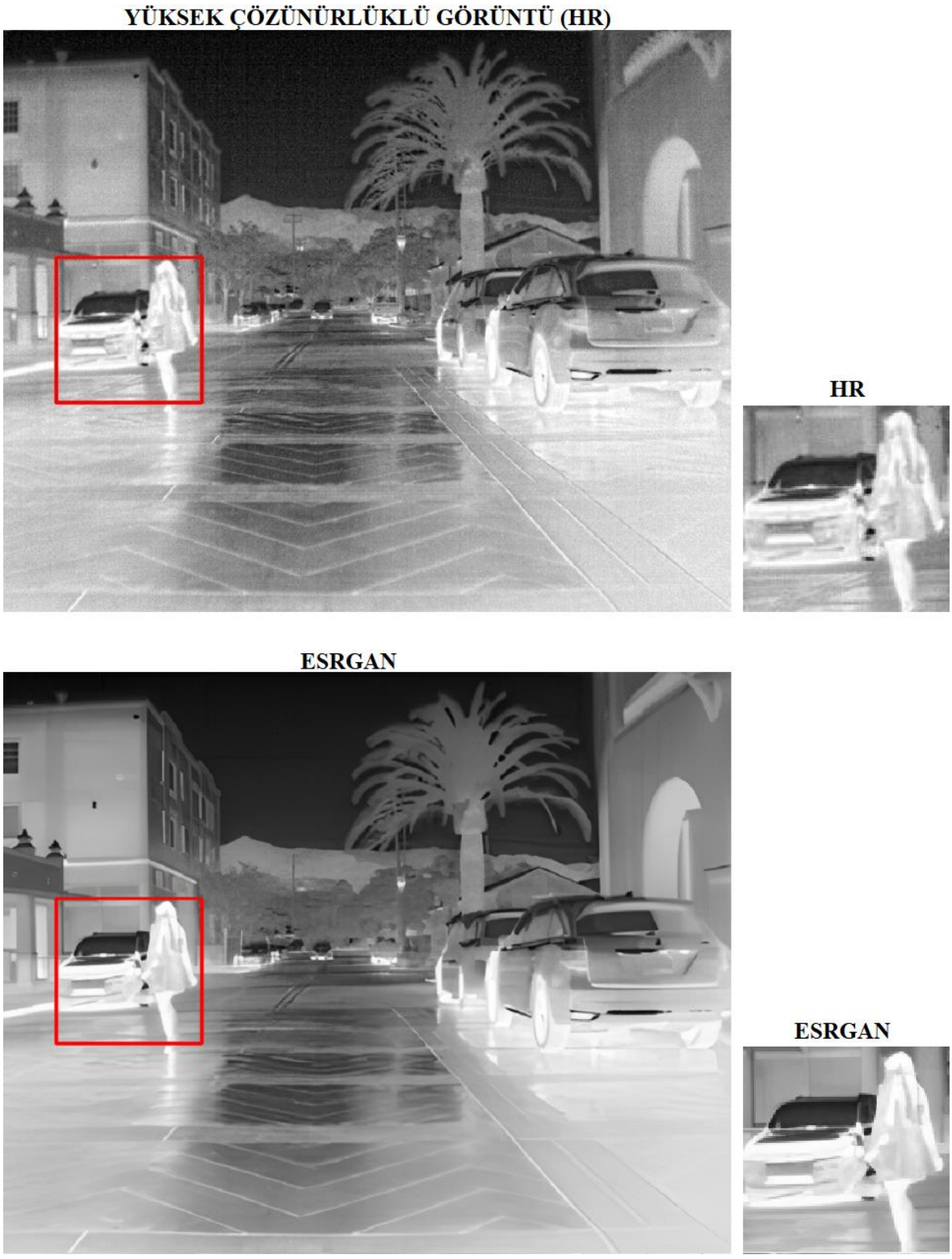
Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
ESRGAN	4x	30.27 dB	0.741	3.03	20.21

Çizelge 7.13 ve Şekil 7.14 değerlendirildiğinde ESRGAN ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.27dB, SSIM değeri ise 0.741 olarak hesaplanmıştır. ESRGAN uygulanmış görüntünün NIQE skoru 3.03 iken BRISQUE skoru ise 20.21 olarak ölçülmüştür.

Bilineer alt örneklemeli ikinci görüntünün çıktıları

Bilineer Enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntüye ESRGAN modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.15'de gösterilmiştir.

Çizelge 7.14'de, SRGAN modeli ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da ESRGAN uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.15. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, ESRGAN ile üst örnekleme uygulanmış hali

Çizelge 7.14. Bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün ESRGAN sonucunda orijinal görüntü ile karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
ESRGAN	4x	30.77	0.741	2.96	18.91

Çizelge 7.14 ve Şekil 7.15 değerlendirildiğinde ESRGAN ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.77 dB, SSIM değeri ise 0.741 olarak hesaplanmıştır. ESRGAN uygulanmış görüntünün NIQE skoru 2.96 iken BRISQUE skoru ise 18.91 olarak ölçülmüştür.

Bikübik alt örneklemeli birinci görüntünün çıktıları

Bikübik Enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntüye ESRGAN modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.16'da gösterilmiştir.

Çizelge 7.15'de, ESRGAN modeli ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da ESRGAN uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.16. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, ESRGAN ile üst örnekleme uygulanmış hali

Çizelge 7.15. Bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün ESRGAN sonucunda orijinal görüntü ile karşılaştırılması

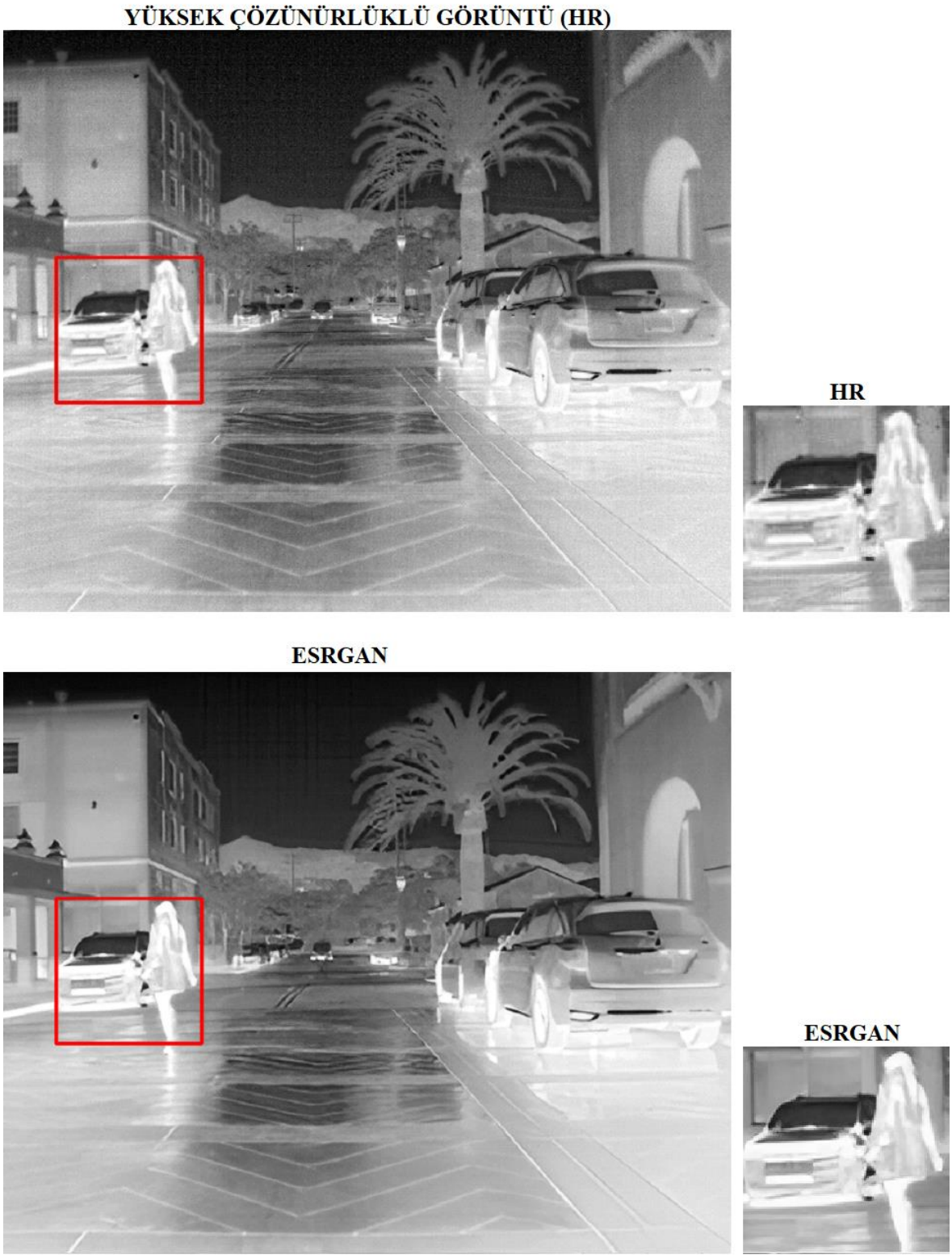
Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
ESRGAN	4x	30.77	0.751	3.72	29.83

Çizelge 7.13 ve Şekil 7.14 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.77dB, SSIM değeri ise 0.751 olarak hesaplanmıştır. ESRGAN uygulanmış görüntünün NIQE skoru 3.72 iken BRISQUE skoru ise 29.83 olarak ölçülmüştür.

Bikübik alt örneklemeli ikinci görüntünün çıktıları

Bikübik Enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntüye ESRGAN modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.17'de gösterilmiştir.

Çizelge 7.16'da, ESRGAN modeli ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da ESRGAN uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.17. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, ESRGAN ile üst örnekleme uygulanmış hali

Çizelge 7.16. Bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün ESRGAN sonucunda orijinal görüntü ile karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
ESRGAN	4x	30.57	0.749	3.47	26.81

Çizelge 7.16 ve Şekil 7.16 değerlendirildiğinde ESRGAN ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.57dB, SSIM değeri ise 0.749 olarak hesaplanmıştır. ESRGAN uygulanmış görüntünün NIQE skoru 3.47 iken BRISQUE skoru ise 26.81 olarak ölçülmüştür.

7.3.5. SwinIR

Bu bölümde bilineer ve bikübik enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntülere SwinIR modeli uygulanarak görüntülerin başarı metriklerinin karşılaştırması sunulmuştur.

Bilineer alt örneklemeli birinci görüntünün çıktıları

Bilineer Enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntüye SwinIR modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.18'de gösterilmiştir.

Çizelge 7.16'de, SwinIR modeli ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da VDSR uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.18. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, SwinIR ile üst örnekleme uygulanmış hali

Çizelge 7.17. Bilineer enterpolasyon ile alt örnekleme yapılmış birinci görüntünün SwinIR sonucunda orijinal görüntü ile karşılaştırılması

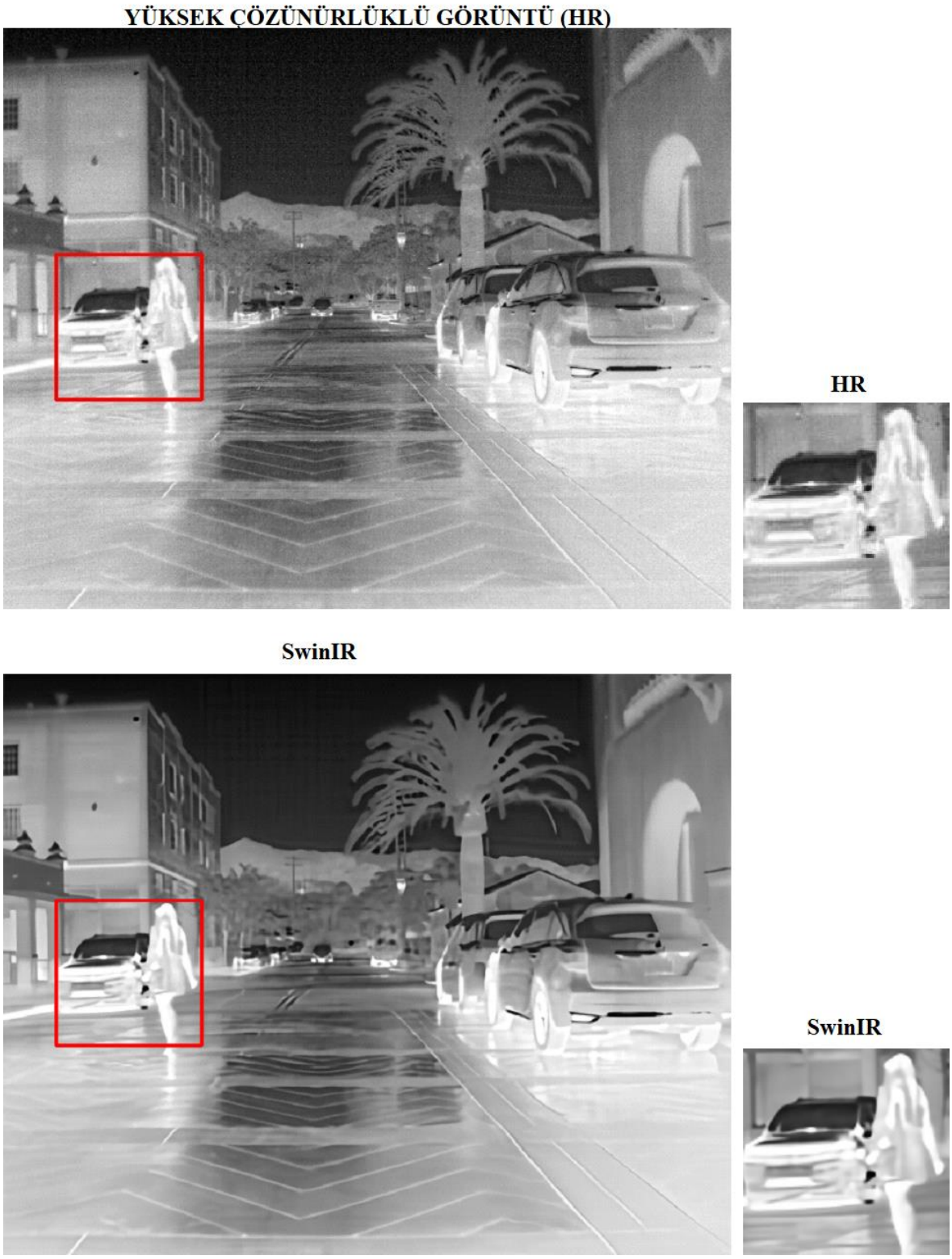
Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
SwinIR	4x	31.32 dB	0.763	3.85	35.40

Çizelge 7.17 ve Şekil 7.17 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 31.32 dB, SSIM değeri ise 0.763 olarak hesaplanmıştır. SwinIR uygulanmış görüntünün NIQE skoru 3.85 iken BRISQUE skoru ise 35.40 olarak ölçülmüştür.

Bilineer alt örneklemeli ikinci görüntünün çıktıları

Bilineer ile 4 kat alt örnekleme yapılmış birinci görüntüden farklı bir düşük çözünürlüklü görüntüye de SwinIR modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.19'de gösterilmiştir.

Çizelge 7.17'de, SwinIR ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da SwinIR uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.19. Orijinal görüntü ve bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, SwinIR ile üst örnekleme uygulanmış hali

Çizelge 7.18. Bilineer enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün VDSR sonucunda orijinal görüntü ile karşılaştırılması

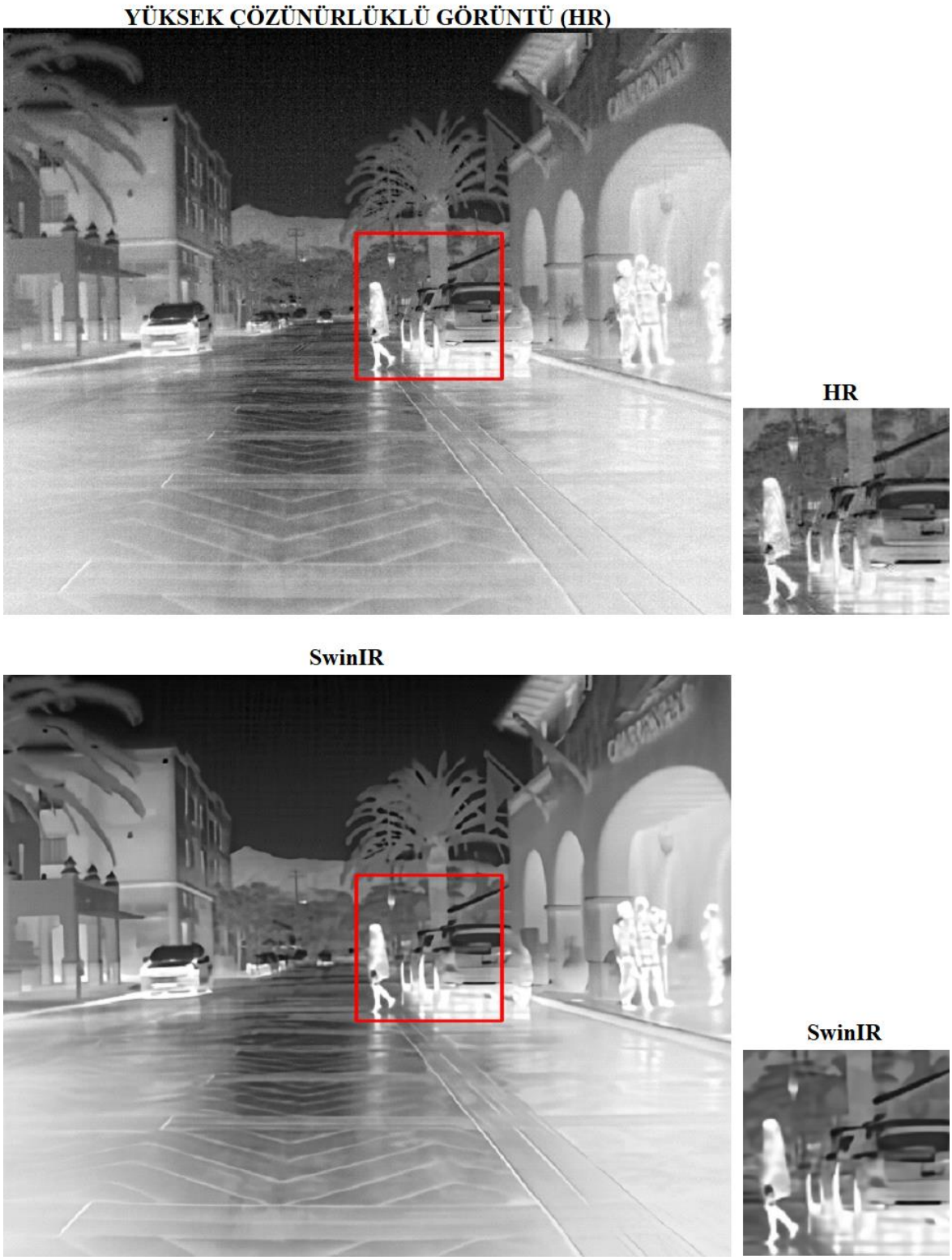
Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
SwinIR	4x	30.71	0.767	4.31	32.51

Çizelge 7.18 ve Şekil 7.18 değerlendirildiğinde bikübik enterpolasyon ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.71 dB, SSIM değeri ise 0.767 olarak hesaplanmıştır. SwinIR uyglanmış görüntünün NIQE skoru 4.31 iken BRISQUE skoru ise 32.51 olarak ölçülmüştür.

Bikübik alt örneklemeli birinci görüntünün çıktıları

Bikübik Enterpolasyon ile 4 kat alt örnekleme yapılmış düşük çözünürlüklü görüntüye ESRGAN modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.20’de gösterilmiştir.

Çizelge 7.18'de, SwiinIR modeli ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da SwinIR uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.20. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün, SwinIR ile üst örnekleme uygulanmış hali

Çizelge 7.19. Bikübik enterpolasyon ile alt örnekleme yapılmış birinci görüntünün SwinIR sonucunda orijinal görüntü ile karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
SwinIR	4x	31.43	0.771	4.15	33.99

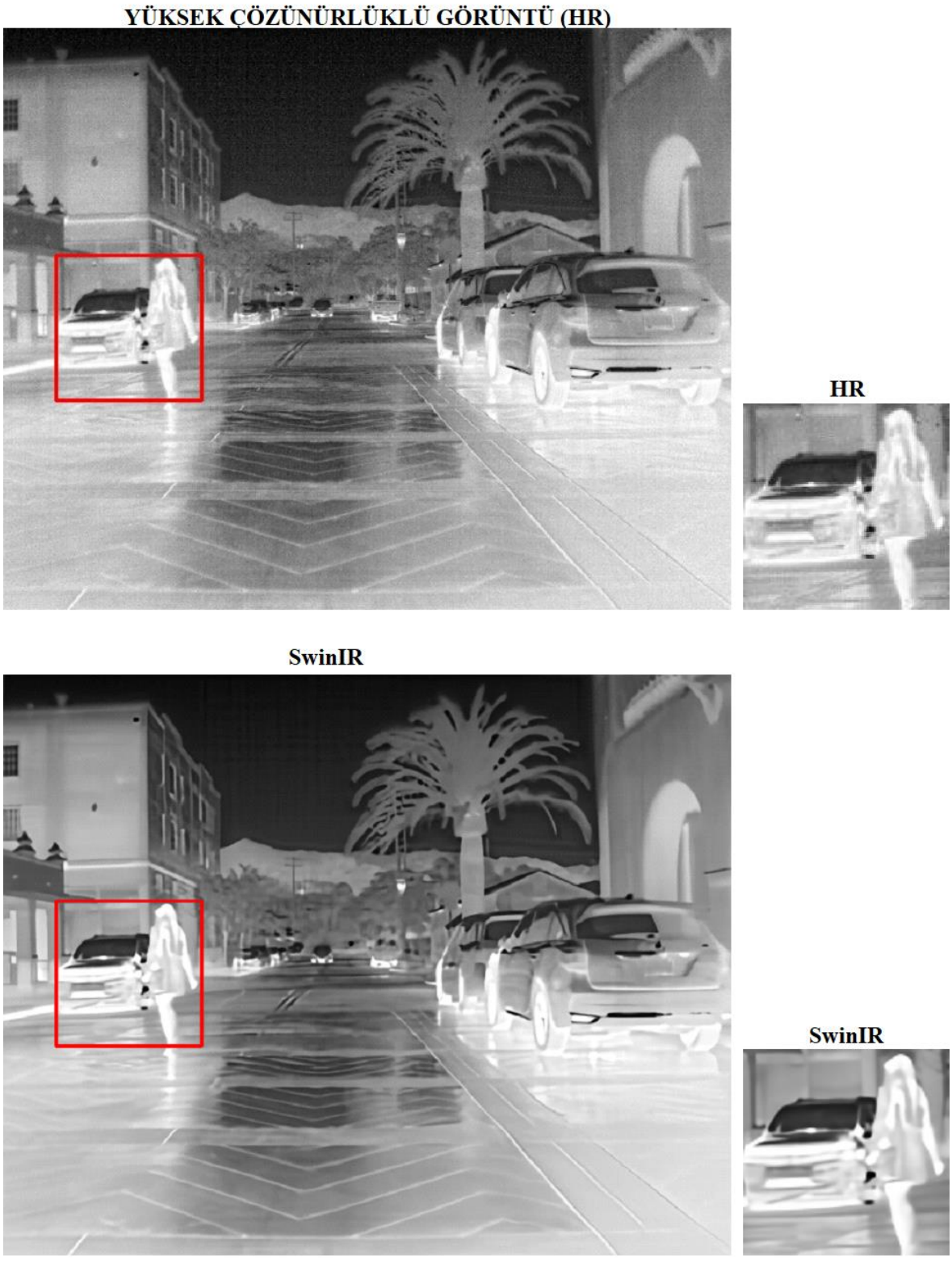
Çizelge 7.19 ve Şekil 7.19 değerlendirildiğinde SwinIR ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 31.43 dB, SSIM değeri ise 0.771 olarak hesaplanmıştır. SwinIR uygulanmış görüntünün NIQE skoru 4.15 iken BRISQUE skoru ise 33.99 olarak ölçülmüştür.

Girdi olarak bilineer enterpolasyon yerine bikübik ile alt örnekleme yapılmış görüntü verildiğinde, bikübik enterpolasyon yönteminin PSNR, NIQE ve BRISQUE metriklerinde daha yüksek başarıma sahip olduğu, SSIM metriğinde ise girdi olarak bilineer enterpolasyon yönteminin daha yüksek başarıma sahip olduğu görülmüştür.

Bikübik alt örneklemeli ikinci görüntünün çıktıları

Bikübik ile 4 kat alt örnekleme yapılmış birinci görüntüden farklı bir düşük çözünürlüklü görüntüye de SwinIR modeli uygulanarak 4 kat üst örnekleme gerçekleştirilmiştir. Bu işlemin sonucunda elde edilen görüntü ile orijinal görüntü ve her iki görüntünün aynı konumlarından kırpılan 128x128 boyutundaki kesitler Şekil 7.21'de gösterilmiştir.

Çizelge 7.18'de, SwinIR ile elde edilen görüntünün orijinal görüntüye göre PSNR ve SSIM değerleri sunulmuş, ayrıca referanssız değerlendirme metriklerinden NIQE ve BRISQUE skorlarına da SwinIR uygulanmış görüntü için çizelgede yer verilmiştir.



Şekil 7.21. Orijinal görüntü ve bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün, SwinIR ile üst örnekleme uygulanmış hali

Çizelge 7.20. Bikübik enterpolasyon ile alt örnekleme yapılmış ikinci görüntünün SwinIR sonucunda orijinal görüntü ile karşılaştırılması

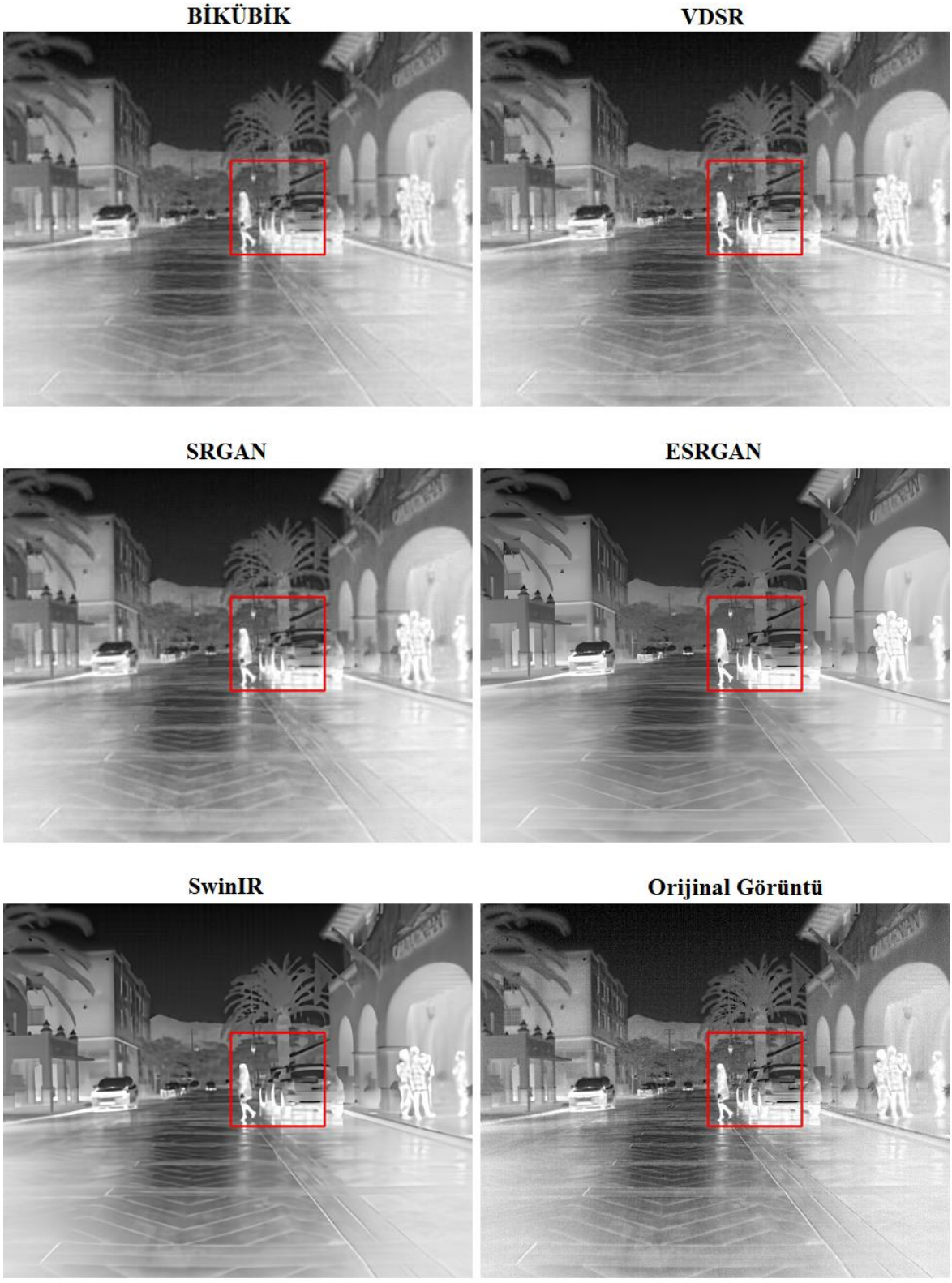
Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
SwinIR	4x	30.71	0.767	4.30	32.48

Çizelge 7.20 ve Şekil 7.19 değerlendirildiğinde SwinIR ile çözünürlüğü yükseltilmiş görüntünün orijinal görüntü ile kıyaslandığında PSNR değer 30.71 dB, SSIM değeri ise 0.767 olarak hesaplanmıştır. Enterpole edilmiş görüntünün NIQE skoru 4.30 iken BRISQUE skoru ise 32.48 olarak ölçülmüştür.

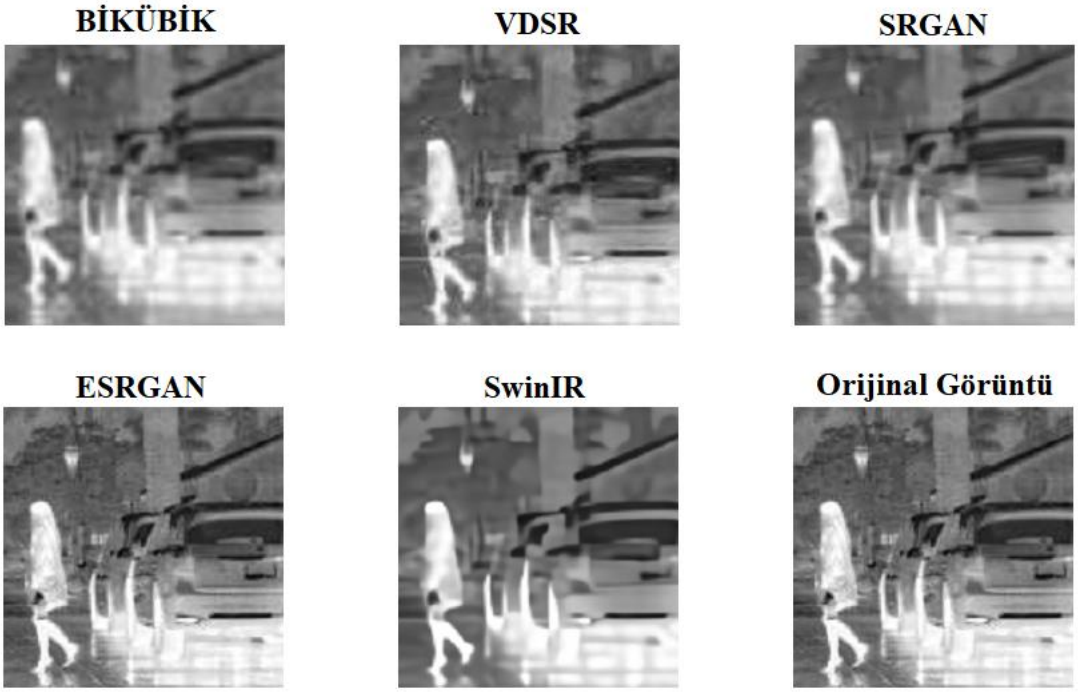
Girdi olarak bilineer enterpolasyon yerine bikübik ile alt örnekleme yapılmış görüntü verildiğinde, bikübik enterpolasyon yönteminin SSIM, NIQE ve BRISQUE metriklerinde daha yüksek başarıma sahip olduğu, PSNR metriğinde ise girdi olarak bilineer enterpolasyon yönteminin daha yüksek başarıma sahip olduğu görülmüştür.

7.4. Bilineer Alt Örneklemeli Görüntülerin Karşılaştırmaları

Bölüm 7.3 altında Bilineer Enterpolasyon ile 4 kat alt örnekleme yapılmış iki farklı görüntü Bikübik, VDSR, SRGAN, ESRGAN ve SwinIR yöntemlerine girdi olarak verilmiş ve bu yöntemler ile 4 kat üst örnekleme yapılmıştır. Şekil 7.22, 7.23, 7.24 ve 7.25'de iki farklı görüntü için her bir yöntemin çıktı görüntüleri yer almaktadır. Çizelge 7.19 ve 7.20'de ise Bilineer Enterpolasyon ile alt örnekleme yapılmış görüntülerin VDSR, SRGAN, ESRGAN ve SwinIR çıktıları başarı metrikleri açısından karşılaştırılmıştır.



Şekil 7.22. Orijinal görüntü ile bikübik enterpolasyon, VDSR, SRGAN, ESRGAN ve SwinIR modellerine verilen düşük çözünürlüklü görüntülerin çıktıları



Şekil 7.23. Şekil 7.22'de her bir model için sunulan görüntülerin aynı bölgelerinden kırpılan kesitler

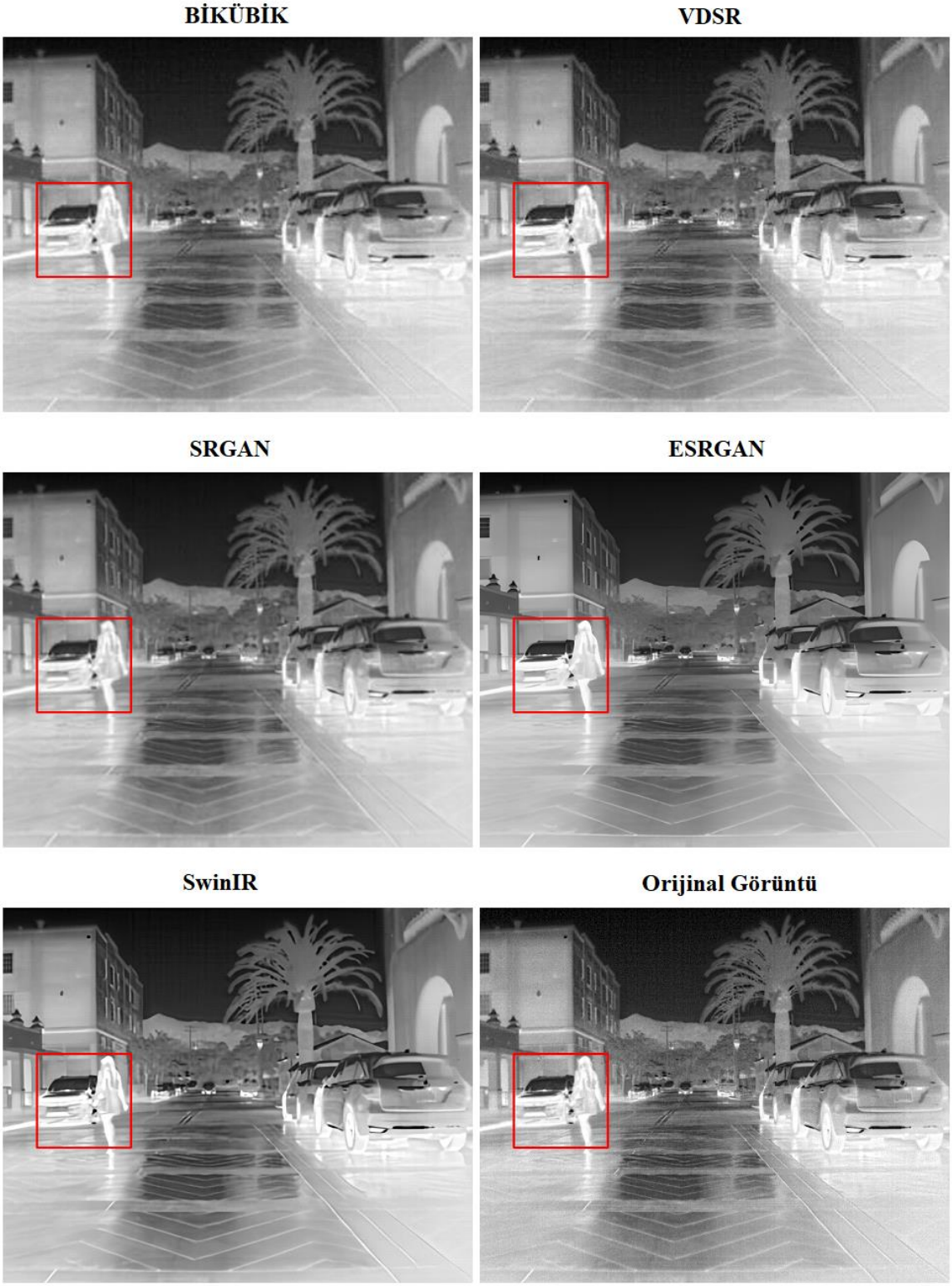
Çizelge 7.21. Bilineer alt örneklemeli birinci görüntünün karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
Bikübik	4x	30.12 dB	0.740	4.51	51.40
VDSR		30.819 dB	0.741	4.22	45.69
SRGAN		30.798 dB	0.742	4.16	33.63
ESRGAN		30.27 dB	0.741	3.03	20.21
SwinIR		31.32 dB	0.763	3.85	35.40

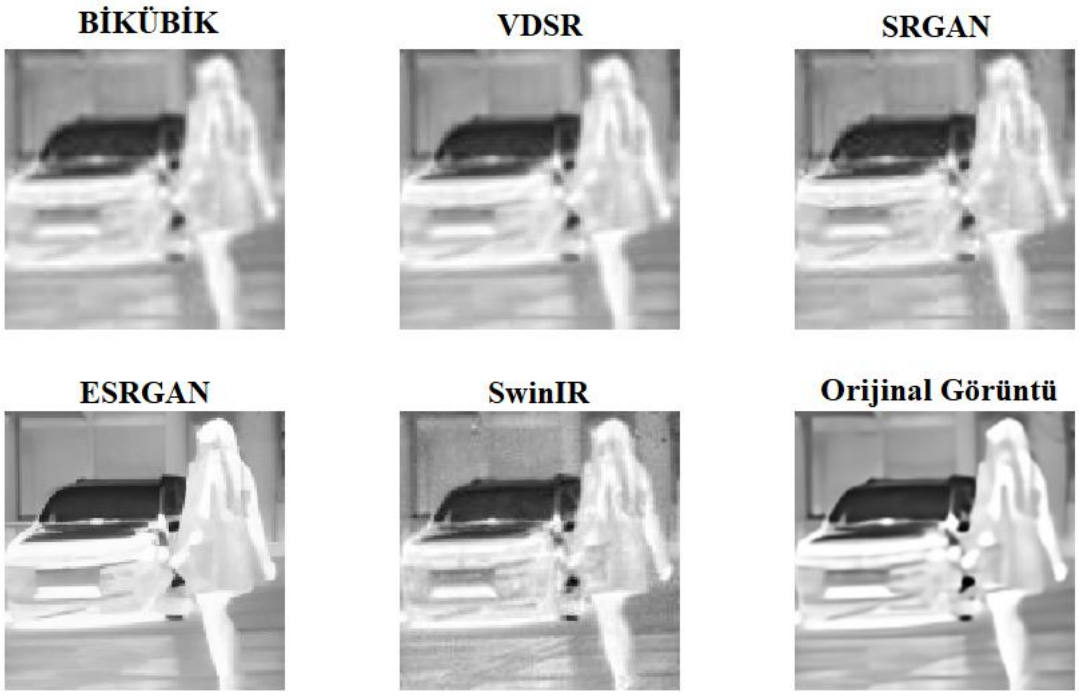
Çizelge 7.21, Şekil 22 ve Şekil 23 değerlendirildiğinde tüm başarı metrikleri arasında en kötü sonucu Bikübik Enterpolasyon yöntemi verirken referans gerektiren başarı metriklerinden olan PSNR ve SSIM için en iyi sonucu SwinIR modelinin verdiği, referans gerektirmeyen başarı metriklerinden olan NIQE ve BRISQUE için en iyi skoru ESRGAN modelinin verdiği görülmüştür.

SwinIR, görüntülerin yapısal bütünlüğünü ve detayları koruma konusunda optimize edilen bir modelken, ESRGAN gürültü ve artefaktları etkili bir şekilde azaltarak, daha doğal görünen görüntüler üretir. Bu yüzden NIQE ve BRISQUE skorları için ESRGAN daha iyi

sonuç verirken SwinIR, PSNR ve SSIM değerleri için daha sonuç üretmiştir. Referans gerektiren başarı metriklerinde SwinIR'ın referans gerektirmeyen metriklerinde ise GAN tabanlı ESRGAN modelinin daha iyi sonuç vermesi literatürdeki çalışmalarla da uyumludur [100]. Literatürde termal görüntüler ile ilgili kısıtlı çalışma olması sebebi ile Nguyen ve arkadaşlarının renkli görüntüler için yaptığı çalışmada [100], SwinIR'ın PSNR ve SSIM için daha iyi sonuçlar verdiği, NIQE ve BRISQUE'de ise ESRGAN'ın daha iyi sonuçlar verdiği görülmüştür. Bu sonuçlar, Çizelge 7.19 ile karşılaştırıldığında benzer sonuçlar elde edildiği görülmüştür. [100, 101].



Şekil 7.24. Orijinal görüntü ile bikübik enterpolasyon, VDSR, SRGAN, ESRGAN ve SwinIR modellerine verilen düşük çözünürlüklü görüntülerin çıktıları



Şekil 7.25. Şekil 7.24'de her bir model için sunulan görüntülerin aynı bölgelerinden kırpılan kesitler

Çizelge 7.22. Bilineer alt örneklemeli ikinci görüntünün karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
Bikübik	4x	30.32 dB	0.737	4.40	47.50
VDSR		30.481	0.739	4.07	38.56
SRGAN		30.70	0.750	3.40	32.57
ESRGAN		30.77	0.741	2.96	18.91
SwinIR		30.71	0.767	4.31	32.51

Çizelge 7.22, Şekil 7.24 ve Şekil 7.25 değerlendirildiğinde tüm başarı metrikleri arasında en kötü sonucu Bikübik Enterpolasyon yöntemi verirken referans gerektiren başarı metriklerinden olan PSNR için ESRGAN, SSIM için en iyi sonucu SwinIR modelinin verdiği, referans gerektirmeyen başarı metriklerinden olan NIQE ve BRISQUE için en iyi skoru ESRGAN modelinin verdiği görülmüştür.

PSNR, orijinal yüksek çözünürlüklü görüntü ile süper çözünürlüklü görüntü arasındaki piksel bazlı farkı ölçer. SSIM ise, görüntünün yapısal özelliklerini, yani insan gözünün algıladığı özellikleri korumaya odaklanır.

ESRGAN modeli, giriş görüntüsü detaylı ve karmaşık olduğunda veya görüntüdeki gürültü

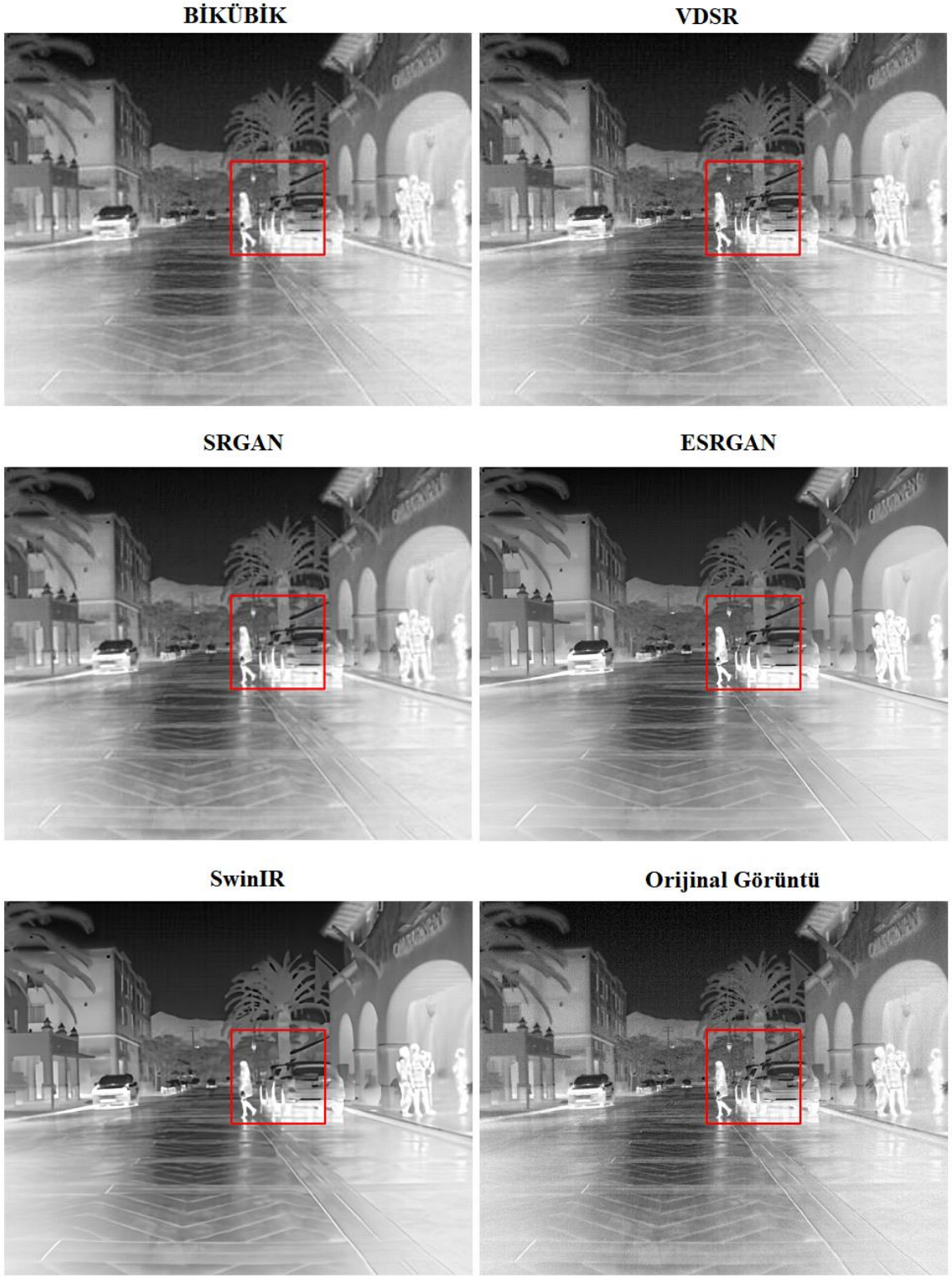
oranı yüksek olduğunda, PSNR değerinde daha iyi sonuçlar verebilir. Bu, ESRGAN'ın karmaşık ve gürültülü verileri daha iyi işleyebildiği anlamına gelir.

Öte yandan, giriş görüntüsü daha düzenli ve yapısal özelliklere sahip olduğunda, yani kenarlar ve köşeler gibi özellikler daha belirgin olduğunda, SwinIR modeli SSIM skorlarında daha iyi performans gösterebilir. Bu, SwinIR'ın bu tür özellikleri daha iyi koruyabildiği anlamına gelir.

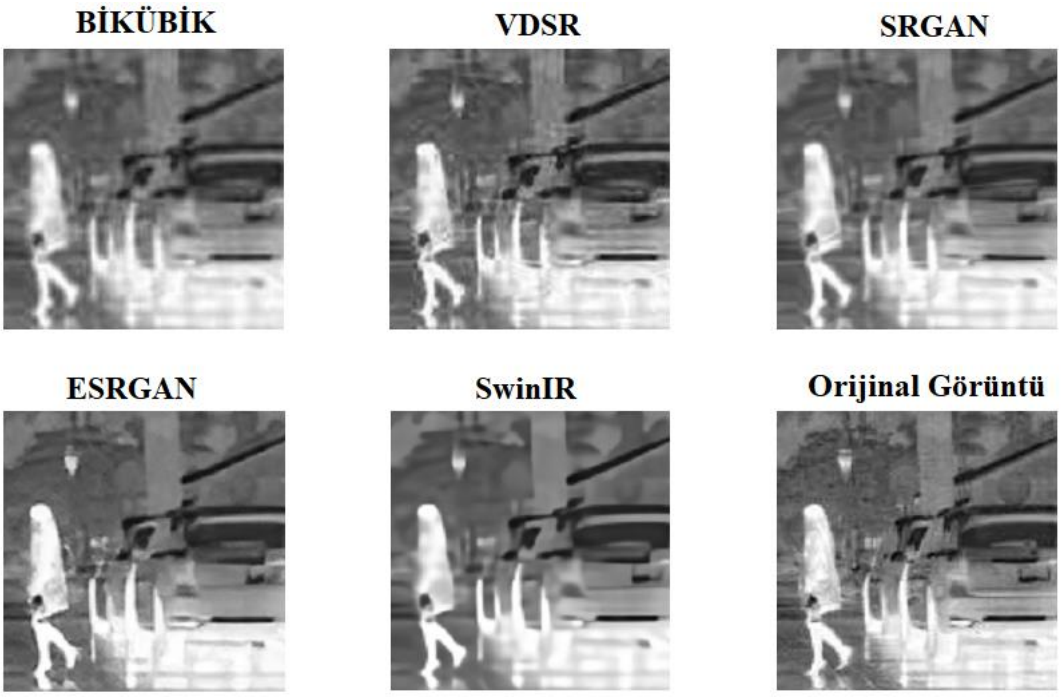
Sonuç olarak, aynı süper çözünürlük modelleri, farklı giriş görüntülerine uygulandığında farklı sonuçlar verebilir. Bu, süper çözünürlük modelinin performansının, giriş görüntüsünün içeriğine ve yapısal özelliklerine bağlı olarak değişebileceği anlamına gelir. Yao ve arkadaşlarının renkli görüntüler için yaptığı çalışmada [101], çoğu görüntüde SwinIR'ın PSNR ve SSIM skorlarında ESRGAN'a göre üstünlük sağladığı, bazı örneklerde ise PSNR özelinde ESRGAN'ın daha iyi sonuçlar verdiği görülmüştür.

7.5. Bikübik Alt Örneklemeli Görüntülerin Karşılaştırmaları

Bölüm 7.3 altında Bikübik Enterpolasyon ile 4 kat alt örnekleme yapılmış iki farklı görüntü Bikübik, VDSR, SRGAN, ESRGAN ve SwinIR yöntemlerine girdi olarak verilmiş ve bu yöntemler ile 4 kat üst örnekleme yapılmıştır. Şekil 7.26, 7.27, 7.28 ve 7.29'da iki görüntü için her bir yöntemin çıktı görüntüleri yer almaktadır. Çizelge 7.21 ve 7.22'de ise Bikübik Enterpolasyon ile alt örnekleme yapılmış görüntüler kendi aralarında olmak üzere görüntülerin başarı metriklerinin karşılaştırması yer almaktadır.



Şekil 7.26. Orijinal görüntü ile bikübik enterpolasyon, VDSR, SRGAN, ESRGAN ve SwinIR modellerine verilen düşük çözünürlüklü görüntülerin çıktıları



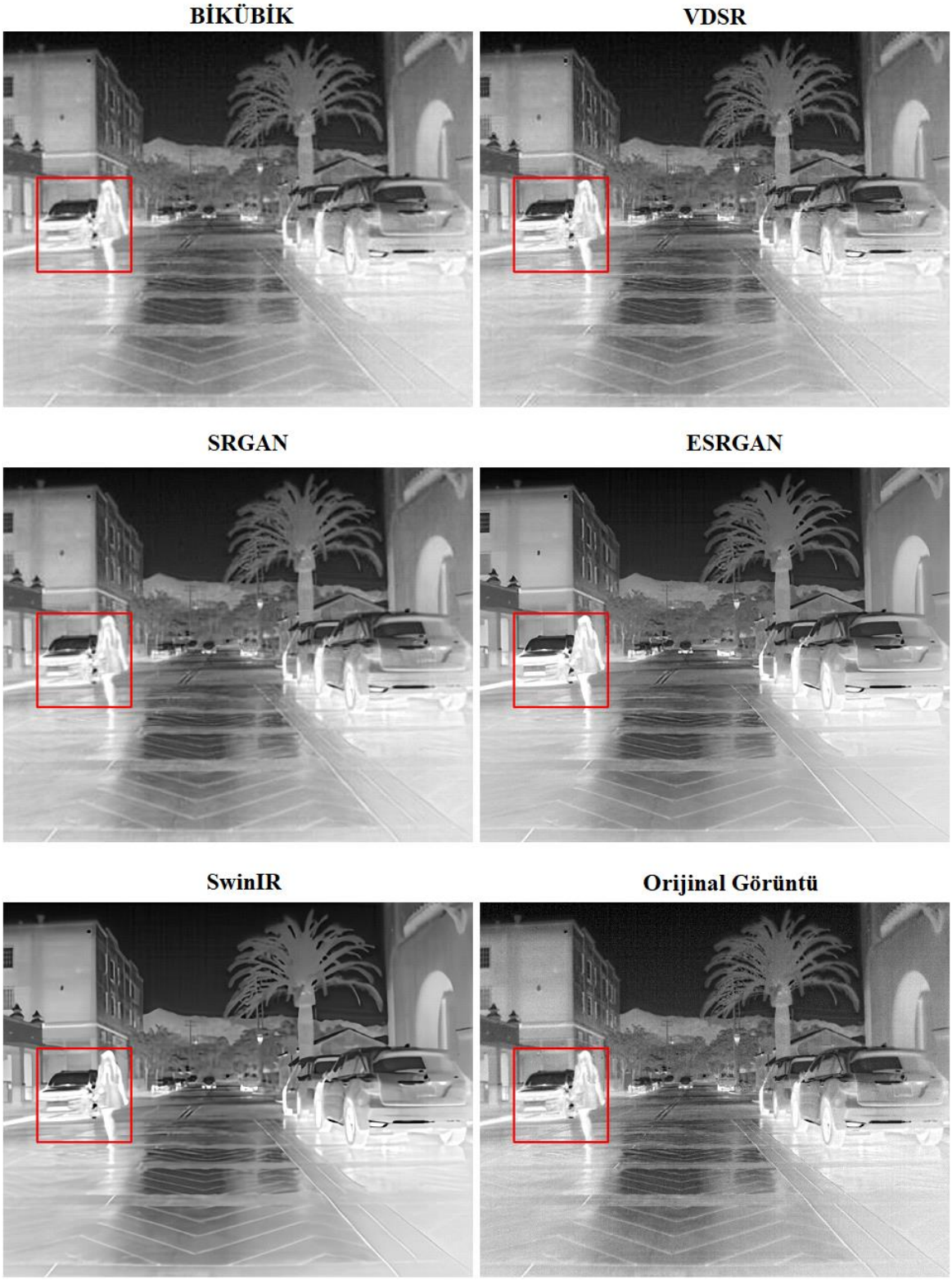
Şekil 7.27. Şekil 7.26'da her bir model için sunulan görüntülerin aynı bölgelerinden kırpılan kesitler

Çizelge 7.23. Bikübik alt örneklemeli birinci görüntünün karşılaştırılması

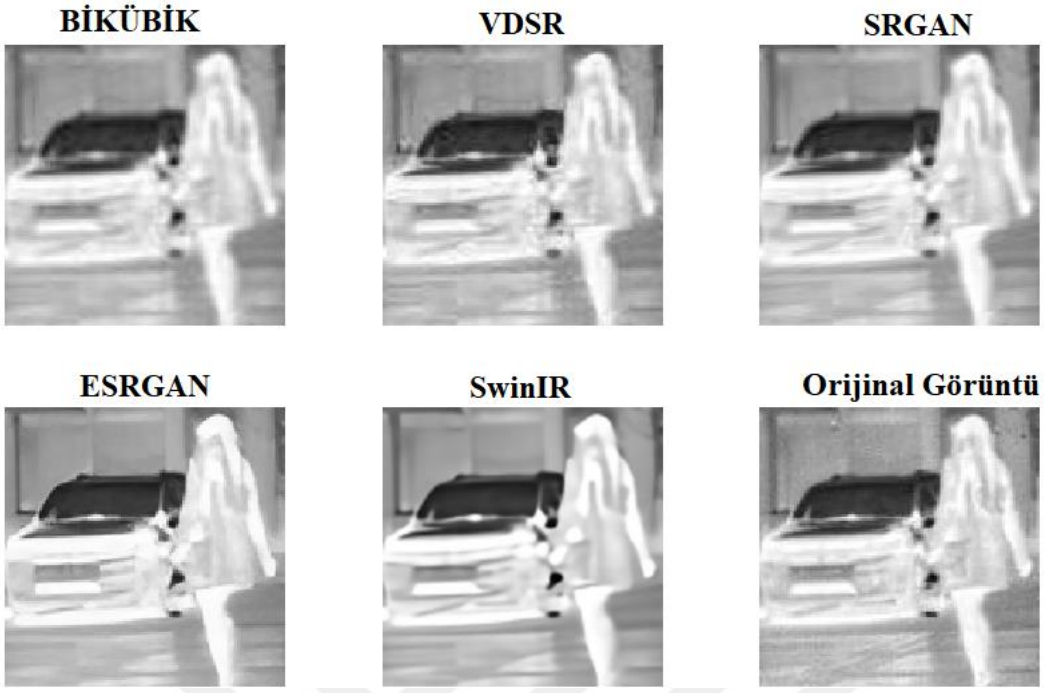
Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
Bikübik	4x	30.72	0.750	4.34	47.09
VDSR		30.98	0.752	4.06	38.92
SRGAN		31.18	0.755	3.68	36.39
ESRGAN		30.77	0.751	3.72	29.83
SwinIR		31.43	0.771	4.15	33.99

Çizelge 7.23, Şekil 7.26 ve Şekil 7.27 değerlendirildiğinde tüm başarı metrikleri arasında en kötü sonucu Bikübik Enterpolasyon yöntemi verirken referans gerektiren başarı metriklerinden olan PSNR ve SSIM için en iyi sonucu SwinIR modelinin verdiği, referans gerektirmeyen başarı metriklerinden olan NIQE için SRGAN ve BRISQUE için en iyi skoru ESRGAN modelinin verdiği görülmüştür.

NIQE bir görüntünün doğallığını ve genel görünümünü değerlendirir. BRISQUE ise görüntüdeki detayların netliği ve gürültü seviyelerine odaklanır. SRGAN, görüntünün doğallığını iyileştirirken, ESRGAN ise görüntünün mekânsal kalitesini iyileştirir. Bu da, NIQE ve BRISQUE skorlarının farklılık göstermesine neden olur.



Şekil 7.28. Orijinal görüntü ile bikübik enterpolasyon, VDSR, SRGAN, ESRGAN ve SwinIR modellerine verilen düşük çözünürlüklü görüntülerin çıktıları



Şekil 7.29. Şekil 7.28'de her bir model için sunulan görüntülerin aynı bölgelerinden kırpılan kesitler

Çizelge 7.24. Bikübik alt örneklemeli ikinci görüntünün karşılaştırılması

Yöntem	Ölçek	PSNR	SSIM	NIQE	BRISQUE
Bikübik	4x	30.45	0.743	4.03	46.54
VDSR		30.51	0.745	3.78	37.36
SRGAN		30.55	0.750	3.40	34.83
ESRGAN		30.57	0.749	3.47	26.81
SwinIR		30.71	0.767	4.30	32.48

Çizelge 7.24, Şekil 7.28 ve Şekil 7.29 değerlendirildiğinde tüm başarı metrikleri arasında en kötü sonucu Bikübik Enterpolasyon yöntemi verirken referans gerektiren başarı metriklerinden olan PSNR için SRGAN, SSIM için en iyi sonucu SwinIR modelinin verdiği, referans gerektirmeyen başarı metriklerinden olan NIQE için SRGAN ve BRISQUE için en iyi skoru ESRGAN modelinin verdiği görülmüştür. Önceki karşılaştırmalarda da sunulduğu üzere, GAN tabanlı mimarinin yüksek gürültü oranına sahip görüntülerde daha iyi sonuçlar vermesi nedeniyle, SwinIR'ın NIQE ve BRISQUE skorlarında GAN tabanlı mimarilere göre daha düşük sonuçlar vermesi normaldir.

PSNR ve SSIM değerlerinde SwinIR modelinin, NIQE ve BRISQUE skorlarında ise

ESRGAN modelinin en iyi sonuçları ürettiği belirlenmiştir. Bu sonuçlar, termal görüntüler üzerinde yapılan bu çalışmanın, literatürde renkli görüntüler üzerinde yapılan çalışmalarla uyumlu olduğunu göstermektedir. [9, 36, 100, 101].

8. SONUÇLAR VE GELECEKTEKİ ÇALIŞMALAR

Bu çalışmada FLIR veri setindeki yüksek çözünürlüklü görüntüler bikübik enterpolasyon yöntemi ile dört kat alt örnekleme yapılarak eğitim ve validasyon seti oluşturulmuştur. Bu veri setleri VDSR, SRGAN, ESRGAN ve SwinIR modelleri ile eğitilmiştir. Bu eğitim esnasında VDSR modelinde SGD optimazyon algoritması, SRGAN, ESRGAN ve SwinIR modellerinde ADAM optimizasyon algoritması kullanılmış olup aşırı uyuma önlem alabilmek için öğrenme hızı dinamik olarak ayarlanmıştır. Eğitim sonrasında, düşük çözünürlüklü termal görüntülere bikübik enterpolasyon, VDSR SRGAN, ESRGAN ve SwinIR olmak üzere beş farklı süper çözünürlük modeli uygulanmıştır. Bunun için her bir modele yüksek çözünürlüklü görüntülerden bilineer enterpolasyon ve bikübik enterpolasyon ile dört kat alt örnekleme yapılmış aynı düşük çözünürlüklü görüntüler verilmiş ve iyileştirilmiş görüntüler elde edilmiştir. VDSR modelinin çıktılarında pürüzlü görüntü kusuru tespit edilmiş olup SRGAN modelinin çıktılarında ise istenmeyen kusurlar (artifact) tespit edilmiştir. ESRGAN ve SwinIR mimarisi ile bu sorun ortadan kaldırılmıştır.

Bu çalışma kapsamında, değerlendirme metrikleri olarak, referans gerektiren değerlendirme metriklerinden olan PSNR ve SSIM ile referans gerektirmeyen metriklerden NIQE ve BRISQUE kullanılmıştır. Öncelikle, derin öğrenme tabanlı ağ modellerinin performansları, bikübik enterpolasyon yöntemi ile analiz edilmiştir. Bu aşamadan sonra, derin ağ modellerinin kendi aralarındaki karşılaştırmalar gerçekleştirilerek performans farklılıkları detaylı bir şekilde incelenmiştir.

Giriş görüntüsünün orijinal görüntüden farklı metodlarla türetilme yöntemi ve bu yöntemlerin farklı kullanımları sonuç görüntülerde belirgin değişiklikler meydana getirebilmektedir. Dolayısıyla, her bir yöntemin bilineer ve bikübik enterpolasyon yöntemi ile alt örnekleme yapılmış görüntülere uygulanmasıyla elde edilen görüntülerin çıktılarında farklılıklar gözlemlenmiştir.

Derin öğrenme tabanlı süper çözünürlük yöntemlerin termal görüntüler üzerinde iyileştirme sağladığı görülmüştür. Değerlendirme metrikleri dikkate alındığında referans gerektiren metriklerden olan PSNR ve SSIM metrikleri için en iyi sonucu SwinIR'ın en kötü sonucu ise bikübik enterpolasyonun verdiği görülmüştür. Referans gerektirmeyen metriklerden olan

NIQE için en iyi sonucu SRGAN ve ESRGAN modeli verirken, BRISQUE için ise en iyi sonucu ESRGAN modelinin verdiği görülmüştür. Bu farklılığın temel nedeninin farklı başarı metriklerinin görüntülerin farklı niteliklerine ve özelliklerine odaklanmasıyla açıklanmıştır.

Çizelge 7.19, Çizelge 7.20, Çizelge 7.21 ve Çizelge 7.22'nin sonuçları ışığında herhangi bir metriğin termal görüntülerde tek başına başarı değerlendirmesi için yeterli olmadığı görülmüştür. Ayrıca 2x ölçek ile de aynı denemeler yapıldığında başarı metrik sonuçları belli bir oranda yükselse bile model başarı sonuçları karşılaştırıldığında bir farklılık gözlemlenememiştir.

Literatürde termal görüntülerde çözünürlük arttırılması çalışmaları kısıtlıdır. Bildiğimiz kadarıyla FLIR veri seti daha önce VDSR, SRGAN, ESRGAN ve/veya SwinIR modellerine uygulanmamıştır. Bu çalışma, VDSR, SRGAN, ESRGAN ve SwinIR mimarilerini "eğiterek" FLIR veri setine uygulayıp çözünürlük arttıran ilk ya da nadir çalışmalardan biri olduğu için literatürdeki renkli görüntüler ile ilgili yapılan çalışmalarla kıyaslama yapılmıştır. Yapılan kıyaslamada sonuçların renkli görüntülerde çıkan model performansları ile tutarlı olduğu gözlemlenmiştir.

Gelecekteki çalışmalarda, daha fazla termal veri seti kullanarak bu çalışmada başarılı sonuçlar veren ESRGAN ve SwinIR, gibi modellerin optimize edilerek eğitilmesi planlanmaktadır. Ayrıca eğitilen her bir modele gerçek hayat verilerinin uygulanması hedeflenmektedir. Ardından bu modellerin karşılaştırılması yapılarak termal görüntüleme için en uygun derin öğrenme tabanlı süper çözünürlük yönteminin bulunması, bulunan bu yöntemin termal kameraların en önemli parametrelerinden biri olan DRI (Detection, Recognition, Idenfication) parametrelerini hangi oranda arttırdığının tespit edilmesi çalışmaları yürütülecektir. Çalışma sonucunda optimize edilmiş model savunma sanayiimizde envanterde yer alan termal kameralarımıza uygulanacaktır.

KAYNAKLAR

1. Huang, G., Liu, Z., Van Der Maaten, L. ve Weinberger, K. Q. (2017). *Densely Connected Convolutional Networks*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 4700-4708.

2. 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(2), 295-307.

3. Greenspan, H. (2009). Super-resolution in medical imaging. *The Computer Journal*, 52(1), 43-63.

4. Bovik, A. (2009). *The essential guide to image processing*. Academic Press.

5. Jain, A. K. (1989). *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ, USA: Prentice-Hall.

6. Kim, J., Lee, J. K. ve Lee, K. M. (2016). *Accurate Image Super-Resolution Using Very Deep Convolutional Networks*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 1646-1654.

7. Lim, B., Son, S., Kim, H., Nah, S. ve Lee, K. M. (2017). *Enhanced Deep Residual Networks for Single Image Super-Resolution*. IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 136-144.

8. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. ve Bengio, Y. (2014). *Generative Adversarial Nets*. Advances in Neural Information Processing Systems (NeurIPS), Montreal, QC, Canada, 2672-2680.

9. Wang, X., Yu, K., Dong, C., Change Loy, C. ve Tang, X. (2018). *ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks*. European Conference on Computer Vision (ECCV), Munich, Germany, 63-79.

10. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S. ve Guo, B. (2021). *Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows*. IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 10012-10022.

11. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85-117.

12. Hardie, R. C., Barnard, K. J. ve Armstrong, E. E. (1997). Joint MAP Registration and High-Resolution Image Estimation Using a Sequence of Undersampled Images. *IEEE Transactions on Image Processing*, 6(12), 1621-1633.

13. Keys, R. (1981). Cubic Convolution Interpolation for Digital Image Processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(6), 1153-1160.

14. Allebach, J. ve Wong, P. W. (1999). *Edge-Directed Interpolation*. IEEE International Conference on Image Processing (ICIP), Kobe, Japan, 707-710.

15. Irani, M., Peleg, S. (1991). Improving Resolution by Image Registration. *CVGIP: Graphical Models and Image Processing*, 53(3), 231-239.

16. Irani, M., Peleg, S. (1990). *Super Resolution from Image Sequences*. 10th International Conference on Pattern Recognition (ICPR), Atlantic City, NJ, USA, 115-120.

17. Freeman, W. T., Pasztor, E. C. Carmichael, O. T. (2000). Learning Low-Level Vision. *International Journal of Computer Vision*, 40(1), 25-47.

18. Hou, H. S. Andrews, H. C. (1978). Cubic Splines for Image Interpolation and Digital Filtering. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(6), 508-517.

19. Simonyan, K. Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. International Conference on Learning Representations (ICLR), San Diego, CA, USA, 1-14.

20. Krizhevsky, A., Sutskever, I. Hinton, G. E. (2012). *ImageNet Classification with Deep Convolutional Neural Networks*. Advances in Neural Information Processing Systems (NeurIPS), Lake Tahoe, NV, USA, 1097-1105.

21. LeCun, Y., Bengio, Y. Hinton, G. (2015). Deep Learning. *Nature*, 521(7553), 436-444.

22. Goodfellow, I., Bengio, Y. Courville, A. (2016). Deep Learning. *MIT Press*.

23. Abadi, M. et al. (2016). *TensorFlow: A System for Large-Scale Machine Learning*. 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI), Savannah, GA, USA, 265-283.

24. Sutskever, I., Vinyals, O. Le, Q. V. (2014). *Sequence to Sequence Learning with Neural Networks*. Advances in Neural Information Processing Systems (NeurIPS), Montreal, QC, Canada, 3104-3112.

25. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. Polosukhin, I. (2017). *Attention is All You Need*. Advances in Neural Information Processing Systems (NeurIPS), Long Beach, CA, USA, 5998-6008.

26. He, K., Zhang, X., Ren, S. Sun, J. (2016). *Deep residual learning for image recognition*. IEEE Conference on Computer Vision and Pattern Recognition, 770-778.

27. Dong, C., Loy, C. C., He, K. Tang, X. (2014*). Learning a Deep Convolutional Network for Image Super-Resolution*. European Conference on Computer Vision (ECCV), Zurich, Switzerland, 184-199.

28. Kim, J., Lee, J. K. Lee, K. M. (2016). *Deeply-Recursive Convolutional Network for Image Super-Resolution*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 1637-1645.

29. Hitawala, S., Li, Y., Wang, X., Yang, D. (2018). *Image Super-Resolution Using VDSR-ResNeXt and SRCGAN*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2814-2823

30. Panagiotopoulou, A., Grammatikopoulos, L., Charou, E., Bratsolis, E., Madamopoulos, N., Petrogonas, J. (2020). *Very Deep Super-Resolution of Remotely Sensed Images with Mean Square Error and Var-norm Estimators as Loss Functions*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2814-2823.

31. Kim, J., Lee, J. K. Lee, K. M. (2018). Deeply-Recursive Convolutional Network for Image Super-Resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), 1138-1151.

32. Lim, B., Son, S., Kim, H., Nah, S. Lee, K. M. (2018). Enhanced Deep Residual Networks for Single Image Super-Resolution. *IEEE Transactions on Image Processing*, 27(3), 1429-1440.

33. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. Bengio, Y. (2020). Generative Adversarial Nets. *Communications of the ACM*, 63(11), 139-144.

34. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., Shi, W. (2017). *Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 105-114.

35. Wang, X., Yu, K., Dong, C., Loy, C. C. (2018). *Recovering Realistic Texture in Image Super-Resolution by Deep Spatial Feature Transform*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 606-615.

36. Song, J., Yi, H., Xu, W., Li, X., Li, B., Liu, Y. (2023). ESRGAN-DP: Enhanced super-resolution generative adversarial network with adaptive dual perceptual loss. *Heliyon*, 9, e15134.

37. Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L., Timofte, R. (2021). *SwinIR: Image Restoration Using Swin Transformer*. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops, Montreal, QC, Canada, 1833-1844.

38. Chen, K., Dai, Y., Li, X., Zhao, Q., Yan, X. (2023). *SwinFSR: Stereo Image Super-Resolution Using SwinIR and Frequency Domain Knowledge*. CVPR Workshops.

39. Xue, F., Wang, T., Li, Y., Li, Q., Zhao, Q. (2023). Rt-swinir: an improved digital wallchart image super-resolution with attention-based learned text loss. *The Visual Computer*.

40. Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B. (2022). Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10), 6982-6995.

41. İnternet: FLIR Systems. (2023). Free FLIR Thermal Dataset for Algorithm Training. URL: http://www.flir.in/oem/adas/adas-dataset-form/. Erişim tarihi: 15 Aralık 2023.

42. Chudasama, V., Shah, M., Bhatt, C., Parikh, S. (2020). *TherISuRNet - A Computationally Efficient Thermal Image Super-Resolution Network*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 86-87.

43. Hecht, E. (2016). *Optics*. Pearson Education.

44. İnternet: Twinkl. (2024). Electromagnetic Spectrum. URL: https://www.twinkl.com.tr/teaching-wiki/electromagnetic-spectrum. Erişim tarihi: 17 Haziran 2024.

45. Preece, D. A. (2015). Atmospheric Attenuation of Electromagnetic Radiation. *IEEE Transactions on Antennas and Propagation*, 63(11), 5070-5075.

46. Shaw, J. A. (2007). *Principles of Thermal Infrared Remote Sensing. Remote Sensing of the Environment: An Earth Resource Perspective*, 2nd ed., Pearson Education.

47. Gonzalez, R. C., ve Woods, R. E. (2017). *Digital Image Processing*, 4th ed. Pearson.

48. Howell, J. R., Siegel, R., Mengüç, M. P. (2010). *Thermal Radiation Heat Transfer*, 5th ed. CRC Press.

49. Rogalski, A. (2011). *Infrared Detectors*, 2nd ed. CRC Press.

50. Goodman, J. W. (2005). *Introduction to Fourier Optics*, 3rd ed. Roberts and Company Publishers.

51. Vollmer, M., Möllmann, K. P. (2017). *Infrared Thermal Imaging: Fundamentals, Research and Applications*. Wiley-VCH.

52. İnternet: Zayıf Akım. (2024). Termal Kamera Nedir?. URL: https://zayifakim.com/termal-kamera-nedir.html. Erişim tarihi: 17 Haziran 2024.

53. İnternet: Global Sensor Technology. (2024). Technical Principles and Applications of Thermal Imaging Modules. URL: https://www.gst-ir.net/news-events/infrared-knowledge/337.html. Erişim tarihi: 25 Mayıs 2024.

54. Gade, R., Moeslund, T. B. (2014). Thermal cameras and applications: a survey. *Machine Vision and Applications*, 25, 245–262.

55. Holst, G. C. (2000). *Common Sense Approach to Thermal Imaging*. JCD Publishing.

56. Vollmer, M. J. (2009). *Advancements in Thermal Imaging Technology*. Proceedings of the SPIE - The International Society for Optical Engineering, 7298.

57. İnternet: Danziger, M. (2020). Autonomous Vehicles Need Thermal Cameras. Foresight, 6 Ağustos 2020. URL: https://www.foresightauto.com/autonomous-vehicles-need-thermal-cameras/. Erişim tarihi: 25 Mayıs 2024.

58. İnternet: iTherml. (2024). Observation, Search and Rescue - Thermal Imaging Cameras. URL: https://www.itherml.com/app/observation-search-rescue-thermal-imaging-cameras-2.html. Erişim tarihi: 25 Mayıs 2024.

59. Xavier, G. D., Parhi, K. S. (2014). Review of passive infrared sensors for urban security. *IEEE Sensors Journal*, cilt 14, no. 12, 4185-4195.

60. İnternet: Khan, M. A. (2023). Single Image Super Resolution Using a Single Camera and Implementation. National University of Sciences and Technology, Temmuz 2023. URL: https://doi.org/10.13140/RG.2.2.17547.05926. Erişim tarihi: 25 Mayıs 2024.

61. İnternet: The AI Learner. (2018). Image Processing: Bilinear Interpolation. 29 Aralık 2018. URL: https://theailearner.com/2018/12/29/image-processing-bilinear-interpolation/ Erişim tarihi: 25 Mayıs 2024.

62. İnternet: The AI Learner. (2018). Image Processing: Bicubic Interpolation. 29 Aralık 2018. URL: https://theailearner.com/2018/12/29/image-processing-bicubic-interpolation/. Erişim tarihi: 25 Mayıs 2024.

63. İnternet: Tüylü, C. (2024). Types of Neural Networks: ANN, RNN, CNN. Medium. URL: https://medium.com/cagataytuylu/types-of-neural-networks-ann-rnn-cnn-b5fd559330c. Erişim tarihi: 29 Haziran 2024.

64. nternet: Ruder, S. (2016). An overview of gradient descent optimization algorithms. Sebastian Ruder Blog. URL: https://www.ruder.io/optimizing-gradient-descent/. Erişim tarihi: 29 Haziranİ 2024.

65. İnternet: Yurtaslan, N. (2024). Denetimli Makine Öğrenmesi Nedir?. Medium. URL: https://nurhayatyurtaslan.medium.com/denetimli-makine-öğrenmesi-nedir-240ec377e6e8. Erişim tarihi: 25 Mayıs 2024.

66. İnternet: Springer Link. (2024). Comparative Study of Supervised Regression Algorithms in Machine Learning. URL: https://link.springer.com/chapter/10.1007/978-3-031-47672-3_35. Erişim tarihi: 25 Mayıs 2024.

67. İnternet: Springer. (2023). On the Philosophy of Unsupervised Learning URL: https://link.springer.com/article/10.1007/s13347-023-00635-6 Erişim tarihi: 25 Mayıs 2024.

68. İnternet: Turhost. (2024). Makine Öğrenmesi (Machine Learning) Nedir?. Turhost Blog. URL: https://www.turhost.com/blog/makine-ogrenmesi-machine-learning-nedir/. Erişim tarihi: 29 Haziran 2024.

69. Fang, W., Yu, Z., Chen, Y., Huang, T., Masquelier, T., Tian, Y. (2021). Deep Residual Learning in Spiking Neural Networks. *Advances in Neural Information Processing Systems (NeurIPS).*

70. Ahmed, S. F., Alam, Md. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., Mofijur, M., Ali, A. B. M. S., Gandomi, A. H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. *Artificial Intelligence Review*, 56, 13521–13617.

71. İnternet: Şehir, G. (2024). Yapay Sinir Ağları (Artificial Neural Network - ANN). Medium. URL:https://medium.com/@gulsahseher/yapay-sinir-aglari-artificial-neural-network-ann-af78c7b8f22e. Erişim tarihi: 29 Haziran 2024.

72. İnternet: Bildirici, F. (2024). Yapay Zeka Eğitim Serisi: Bölüm 14. Medium. URL: https://medium.com/@fatihbildirici.dev/yapay-zeka-eğitim-serisi-bölüm-14 696e146d3d52. Erişim tarihi: 29 Haziran 2024.

73. Xu, C., Yang, X., Li, S., Huang, X., Pan, H., Lei, X. (2023). Improved Deep Plug-and-Play Super-Resolution with Residual-in-Residual Dense Block. *Pattern Analysis and Applications*.

74. Liu, G., Zhou, J., Xu, L., Li, L., Luo, B., Chen, S. (2022). MCTSR: A super-resolution method for defects in gas-insulated switchgear X-ray images based on multi-scale contextual transformers. *High Voltage*, 7(4), 1234-1245.

75. Montgomery, D. C., Peck, E. A., Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. John Wiley & Sons.

76. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

77. İnternet: Ramachandran, P., Zoph, B., Le, Q. V. (2017). Swish: a Self-Gated Activation Function. URL: https://arxiv.org/abs/1710.05941v1. Erişim tarihi: 19 Haziran 2024.

78. Ioffe, S., Szegedy, C. (2015). *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. Proceedings of the 32nd International Conference on Machine Learning (ICML 2015).

79. El-Adawy, O. A., El-Soudani, M. M. E. A., & Amer, H. H. (2019). *Image super-resolution using deep learning: A survey*. 2019 IEEE 31st International Conference on Microelectronics (ICM).

80. Teo, T.-A., Fu, Y.-J. (2021). Spatiotemporal Fusion of Formosat-2 and Landsat-8 Satellite Images: A Comparison of 'Super Resolution-Then-Blend' and 'Blend-Then-Super Resolution' Approaches. *Remote Sensing*, 13(4), Makale no. 606.

81. Johnson, J. Alahi, A., Fei-Fei, L. (2016). *Perceptual losses for real-time style transfer and super-resolution*. Proceedings of the European Conference on Computer Vision (ECCV).

82. Song, L. (2019). Image super-resolution reconstruction algorithm based on EDSR and improved SRGAN. *Applied and Computational Engineering*, 5, 178-185.

83. Shi, W. Caballero, J., Huszár, F., Totz, J., Aitken, A. P., Bishop, R., Wang, Z. (2016). *Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

84. Kishore, L., Kumar, A., Dang, N. (2020). Enhanced Image Restoration by GANs using Game Theory. *Procedia Computer Science*, 173, 225-233.

85. Jolicoeur-Martineau, A. (2018). *The relativistic discriminator: a key element missing from standard GAN*. Proceedings of the 7th International Conference on Learning Representations (ICLR 2019)

86. Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., .Houlsby, N (2020). *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS 2020).

87. Hu, W. Zhang, Z., Xie, Z., Lin, S (2019). *Local Relation Networks for Image Recognition*. Proceedings of the IEEE International Conference on Computer Vision, 3464-3473.

88. Chen, T. Kornblith, S., Norouzi, M.,Hinton, G. (2020). *A Simple Framework for Contrastive Learning of Visual Representations*. Proceedings of the International Conference on Machine Learning, 1597-1607.

89. Carion, N. Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S. (2020). *End-to-End Object Detection with Transformers*. European Conference on Computer Vision, 213-229.

90. Ba, J., Kiros, J., Hinton, G. (2016). *Layer Normalization*. Proceedings of the 30th Conference on Neural Information Processing Systems (NeurIPS 2016).

91. Rong, Z., Jiang, X., Huang, L., Zhou, H. (2023). Swin–MRDB: Pan-Sharpening Model Based on the Swin Transformer and Multi-Scale CNN. *Appl. Sci*., 13(15), 9022.

92. Vaswani, A. (2017). Attention is All You Need. *Advances in Neural Information Processing Systems*, 30.

93. Hore, A., Ziou, D. (2010). *Image quality metrics: PSNR vs. SSIM*. 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 2366-2369.

94. Mittal, A., Soundararajan, R., Bovik, A. C. (2013). Making a 'completely blind' image quality analyzer. *IEEE Signal Processing Letters*, 20(3), 209-212.

95. Moorthy, A. K., Bovik, A. C. (2011). Blind image quality assessment: From natural scene statistics to perceptual quality. *IEEE Transactions on Image Processing*, 20(12), 3350-3364.

96. Rohil, M. K., Gupta, N., Yadav, P. (2019). An improved model for no-reference image quality assessment and a no-reference video quality assessment model based on frame analysis. *Signal, Image and Video Processing*, 13(6), 1147–1154.

97. Mittal, A., Moorthy, A. K., Bovik, A. C. (2011). *Blind/Referenceless Image Spatial Quality Evaluator*. 2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), 06-09 Kasım 2011, 139.

98. Mittal, A., Moorthy, A. K., Bovik, A. C. (2012). No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing*, 21(12), 4695-4708.

99. İnternet: Shrimali, K. R. (2024). Image Quality Assessment - BRISQUE. LearnOpenCV. URL: https://learnopencv.com/image-quality-assessment-brisque/. Erişim tarihi: 29 Haziran 2024.

100. Nguyen, M., Sood, K., Avery, K., Bein, D. (2023). *Deep Learning-based Super-Resolution on the Cloud: Focus on Face and Text Enhancement*. The 5th International Conference on Robotics and Computer Vision.

101. Yao, X., Pan, Y., Wang, J. (2024). An Omnidirectional Image Super-Resolution Method Based on Enhanced SwinIR. *Information*, 15(5), 248.



Gazili olmak ayrıcalıktır