

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**DİFÜZYON AĞLARI İLE
GÖRÜNTÜ REKONSTRÜKSİYONU VE RESTORASYONU**

YÜKSEK LİSANS TEZİ

Onur PARAPAN

Elektronik ve Haberleşme Mühendisliği Anabilim Dalı

Telekomünikasyon Mühendisliği Programı

HAZİRAN 2025

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

**DİFÜZYON AĞLARI İLE
GÖRÜNTÜ REKONSTRÜKSİYONU VE RESTORASYONU**

YÜKSEK LİSANS TEZİ

**Onur PARAPAN
(504131319)**

Elektronik ve Haberleşme Mühendisliği Anabilim Dalı

Telekomünikasyon Mühendisliği Programı

Tez Danışmanı: Prof. Dr. Ender Mete EKŞİOĞLU

HAZİRAN 2025

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

**IMAGE RECONSTRUCTION AND RESTORATION
WITH DIFFUSION NETWORKS**

M.Sc. THESIS

**Onur PARAPAN
(504131319)**

Department of Electronics and Communication Engineering

Telecommunications Engineering Program

Thesis Advisor: Prof. Dr. Ender Mete EKŞİOĞLU

JUNE 2025

İTÜ, Lisansüstü Eğitim Enstitüsü'nün 504131319 numaralı Yüksek Lisans Öğrencisi Onur PARAPAN, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “DİFÜZYON AĞLARI İLE GÖRÜNTÜ REKONSTRÜKSİYONU VE RESTORASYONU” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Prof. Dr. Ender Mete EKŞİOĞLU**
İstanbul Teknik Üniversitesi

Jüri Üyeleri : **Doç. Dr. İsa YILDIRIM**
İstanbul Teknik Üniversitesi

Prof. Dr. Hacı İLHAN
Yıldız Teknik Üniversitesi

Teslim Tarihi : **27 Mayıs 2025**

Savunma Tarihi : **23 Haziran 2025**





Eşime ve oğluma,



ÖNSÖZ

Bu çalışmada difüzyon ağlarının görüntü işleme alanındaki verimliliği konusu incelenmiştir. Hazırlık aşamasındaki yardımlarından dolayı tez danışmanım sayın Prof. Dr. Ender Mete EKŞİOĞLU' na ve beni her zaman destekleyen aileme teşekkür ederim.

Haziran 2025

Onur PARAPAN
(Veri Ambarı Uzmanı)



İÇİNDEKİLER

	<u>Sayfa</u>
ÖNSÖZ	ix
İÇİNDEKİLER	xi
KISALTMALAR	xv
ÇİZELGE LİSTESİ	xvii
ŞEKİL LİSTESİ	xix
ÖZET	xxi
SUMMARY	xxiii
1. GİRİŞ	1
1.1 Tezin Amacı	2
1.2 Literatür Araştırması	3
1.3 Hipotez	5
2. GÖRÜNTÜ REKONSTRÜKSİYONU VE RESTORASYONU	7
2.1 Görüntü Rekonstrüksiyonu	7
2.1.1 Görüntü rekonstrüksiyonuna olan ihtiyaç	9
2.1.2 Teknikler	10
2.1.2.1 Frekans domaini teknikleri	11
2.1.2.2 Matris tamamlama teknikleri	12
2.1.2.3 Derin öğrenme yöntemleri	14
2.1.3 Uygulama alanları	15
2.1.4 Önemli uygulamalar	17
2.1.4.1 Tomografik görüntüleme teknikleri	17
2.1.4.2 Manyetik rezonans görüntüleme	19
2.1.4.3 Fotoakustik görüntüleme	22
2.1.4.4 Ultrasonografi	22
2.2 Görüntü Restorasyonu	23
2.2.1 Görüntü restorasyonuna olan ihtiyaç	23
2.2.2 Teknikler	24
2.2.2.1 Konvolüsyonel filtreleme teknikleri	25
2.2.2.2 Frekans domaini teknikleri	26
2.2.3 Uygulama alanları	27
2.2.4 Önemli uygulamalar	29
2.2.4.1 Gürültü giderme	29
2.2.4.2 Bulanıklık giderme	32
2.2.4.3 Mozaik çözme	34
2.2.4.4 Süper çözünürlük	36
2.2.4.5 Görüntü tamamlama	37
2.2.4.6 Metal artefakt azaltma	39
3. GÖRÜNTÜ İŞLEME ALANINDA DERİN ÖĞRENME	43
3.1 Derin Öğrenme ve Yapay Sinir Ağlarına Genel Bakış	43
3.2 Görüntü İşleme Alanındaki Derin Öğrenme Uygulamaları	46
3.3 Mevcut Yöntemlerin Sınırlamaları	49
4. DİFÜZYON AĞLARI	51
4.1 Difüzyon Ağları Nedir?	51
4.2 Temel İlkeler ve Difüzyon Ağlarının Yapısı	51
4.3 Matematiksel Arka Plan ve Modelleme	53
4.4 Kullanılan Mimariler	54
4.4.1 U-Net tabanlı difüzyon modelleri	54
4.4.2 Varyasyonel otokodlayıcı tabanlı difüzyon modelleri	56

4.5 Modelleme Yaklaşımları	56
4.5.1 Latent difüzyon modelleri	56
4.5.2 Şartlı difüzyon modelleri	57
4.5.3 Şkor bazlı difüzyon modelleri	58
4.6 Eğitim ve Örneklemeye Yöntemleri	59
4.6.1 Eğitim yöntemleri	59
4.6.2 Örneklemeye yöntemleri	60
4.7 Geleneksel Yöntemlere Göre Avantajları	60
4.8 Difüzyon Ağlarının Temel Uygulamaları	62
4.9 Sınırlamalar ve İyileştirme Önerileri	64
4.9.1 Sınırlamalar	64
4.9.2 İyileştirme önerileri	66
5. DİFÜZYON AĞLARI İLE GÖRÜNTÜ REKONSTRÜKSİYONU VE RESTORASYONU	67
5.1 Genel Bakış	67
5.2 Görüntü İşleme Alanına Geçişi	67
5.3 Yöntemler	69
5.3.1 Gürültü giderici olasılık modeli	69
5.3.2 Tak ve çalıştır difüzyon modeli	70
5.3.3 Artık kaydırmalı difüzyon modeli	71
5.4 Kazanımlar	72
5.4.1 Görüntü rekonstrüksiyonu alanında kazanımlar	73
5.4.2 Görüntü restorasyonu alanında kazanımlar	76
5.5 Önemli Uygulamalar	76
5.5.1 Gürültü giderici difüzyon olasılık modeli ile görüntü tamamlama	77
5.5.2 Gürültü giderici difüzyon restorasyon yöntemi	79
5.5.3 Difüzyon tabanlı görüntü restorasyonu	80
5.5.4 Difüzyon tabanlı tak ve çalıştır görüntü restorasyonu	82
6. GÖRÜNTÜ RESTORASYONUNDA DİFÜZYON AĞLARININ PERFORMANSI	85
6.1 Genel Bakış	85
6.2 U-Net ve Difüzyon Ağlarının Performans Karşılaştırması	85
6.3 Görüntü Kalitesi Değerlendirme Metrikleri	86
6.3.1 Tepe sinyal-gürültü oranı	86
6.3.2 Yapısal benzerlik indeksi	87
6.3.3 Öğrenilmiş algısal görüntü benzerliği	87
6.4 Süper Çözünürlük Görevi Üzerinde Karşılaştırma	88
6.4.1 Deneysel kurulum	88
6.4.2 Eğitim süreci	89
6.4.3 Sonuçlar	92
6.4.4 Görsel karşılaştırmalar	93
6.4.5 Yorumlar	93
6.5 Görüntü Tamamlama Görevi Üzerinde Karşılaştırma	94
6.5.1 Deneysel kurulum	94
6.5.2 Eğitim süreci	94
6.5.3 Sonuçlar	97
6.5.4 Görsel karşılaştırmalar	98
6.5.5 Yorumlar	98
7. SONUÇLARIN DEĞERLENDİRİLMESİ VE TARTIŞMA	99
7.1 Genel Değerlendirme	99
7.2 U-Net ve Difüzyon Ağlarının Performans Analizi	99
7.3 Nicel ve Algısal Değerlendirme Sonuçlarının Karşılaştırılması	100
7.4 Çalışmanın Sınırlamaları	100
7.5 Gelecek Çalışmalar İçin Öneriler	101
KAYNAKLAR	103





KISALTMALAR

2D	: İki Boyutlu
ADMM	: Alternatif Yönlü Çarpan Yöntemi
BERT	: Çift Yönlü Kodlayıcı Tabanlı Dönüştürücü
BM3D	: Blok Eşleştirmeli ve 3 Boyutlu Filtrelemeli Gürültü Giderme
CNN	: Konvolüsyonel Sinir Ağı
CS	: Sıkıştırılmış Algılama
CS-MRI	: Sıkıştırılmış Algılama Manyetik Rezonans Görüntüleme
CT	: Bilgisayarlı Tomografi
DAE	: Gürültü Giderici Otokodlayıcı
DDIM	: Gürültü Giderici Difüzyon Dolaylı Modeli
DDPM	: Gürültü Giderici Olasılıksal Difüzyon Modeli
DDRM	: Gürültü Giderici Difüzyon Restorasyon Modeli
DIFFIR	: Gürültü Giderici Verimli Difüzyon Modeli
DEBLURGAN	: Bulanıklık Giderici Çekişmeli Üretici Ağ
DPIR	: Derin Tak ve Çalıştır Görüntü Restorasyonu
ESRGAN	: Gelişmiş Süper Çözünürlük Üretici Çekişmeli Ağ
GAN	: Üretici Çekişmeli Ağ
GPT	: Önceden Eğitilmiş Üretici Dönüştürücü
LPIPS	: Öğrenilmiş Algısal Görüntü Benzerliği
MAP	: En Büyük Artalan Olasılığı
MAR	: Metal Artefakt Giderme
MSE	: Ortalama Kare Hata
MRI	: Manyetik Rezonans Görüntüleme
MRPNET	: Çoklu Bölge Algılama Sinir Ağı
NLM	: Yerel Olmayan Ortalama
OCT	: Optik Koherens Tomografi
PET	: Pozitron Emisyon Tomografisi
PnP	: Tak ve Çalıştır
PSNR	: Tepe Sinyal-Gürültü Oranı
RNN	: Tekrarlayan Sinir Ağı
SDE	: Stokastik Diferansiyel Denklem
SGD	: Stokastik Gradyan Düşmesi
SPECT	: Tek Foton Emisyonlu Bilgisayarlı Tomografi
SRGAN	: Süper Çözünürlük Üretici Çekişmeli Ağ
SSIM	: Yapısal Benzerlik İndeksi
U-NET	: U Şekilli Evrişimli Sinir Ağı
VAE	: Varyasyonel Otokodlayıcı



ÇİZELGE LİSTESİ

	<u>Sayfa</u>
Çizelge 4.1: Farklı difüzyon mimarilerinin karşılaştırması.	56
Çizelge 4.2: Farklı modelleme yaklaşımlarının karşılaştırması.....	58
Çizelge 5.1: Farklı süper çözünürlük yöntemlerinin PSNR, SSIM ve LPIPS metriklerine göre karşılaştırılması [1].	72
Çizelge 5.2: Artık kaydırmalı difüzyon ile klasik difüzyon yaklaşımlarının karşılaştırılması.	73
Çizelge 5.3: DiffIR ve farklı yöntemlerin PSNR karşılaştırması.	81
Çizelge 5.4: DPIR ve farklı yöntemlerin PSNR karşılaştırması.	83
Çizelge 6.1: U-Net ve difüzyon modellerinin süper çözünürlük görevinde performans karşılaştırması.	92
Çizelge 6.2: U-Net ve difüzyon ağlarının görüntü tamamlama görevinde performans karşılaştırması.	97
Çizelge 7.1: U-Net ve difüzyon ağlarının görev bazlı performans karşılaştırması.	100



ŞEKİL LİSTESİ

	<u>Sayfa</u>
Şekil 2.1: K-uzayı ve uzay domaini manyetik rezonans görüntüleme.....	8
Şekil 2.2: İdeal yüksek geçiren filtre.	12
Şekil 2.3: Radon dönüşümü.	17
Şekil 2.4: Filtrelenmiş geri projeksiyon	18
Şekil 2.5: K-uzayı ve uzay domaini görüntü gelişimi.	20
Şekil 2.6: Gürültü giderme.....	24
Şekil 2.7: Bulanıklık giderme.	24
Şekil 2.8: İdeal alçak geçiren filtre.	26
Şekil 3.1: Konvolüsyonel sinir ağları mimarisi.	43
Şekil 3.2: Tekrarlayan sinir ağları mimarisi.....	44
Şekil 3.3: Otokodlayıcı mimarisi.	45
Şekil 3.4: Transformatör mimarisi	45
Şekil 3.5: GAN mimarisi	46
Şekil 4.1: İleri ve geri difüzyon süreçleri.	52
Şekil 4.2: U-Net mimarisi.	55
Şekil 5.1: Artık kaydırma yönteminin yüksek çözünürlük - düşük çözünürlük görüntü çifti arasında artıkları kaydırması [1].	71
Şekil 6.1: Gerçek görüntü ve biküzik interpolasyon ile oluşturulmuş düşük çözünürlüklü görüntü.	88
Şekil 6.2: U-Net modeli için süper çözünürlük görevinde PSNR, SSIM ve LPIPS öğrenme eğrileri.....	90
Şekil 6.3: Difüzyon ağı için süper çözünürlük görevinde PSNR, SSIM ve LPIPS öğrenme eğrileri.....	91
Şekil 6.4: U-Net ve difüzyon modellerinin süper çözünürlük görevinde görsel karşılaştırması.	93
Şekil 6.5: U-Net ve difüzyon modellerinin süper çözünürlük görevinde 64x64 piksel bazlı görsel karşılaştırması.	93
Şekil 6.6: U-Net modeli için görüntü tamamlama görevinde PSNR, SSIM ve LPIPS öğrenme eğrileri.....	95
Şekil 6.7: Difüzyon ağı için görüntü tamamlama görevinde PSNR, SSIM ve LPIPS öğrenme eğrileri.....	96
Şekil 6.8: U-Net ve difüzyon modellerinin görüntü tamamlama görevinde görsel karşılaştırması.	98
Şekil 6.9: U-Net ve difüzyon modellerinin görüntü tamamlama görevinde 64x64 piksel bazlı görsel karşılaştırması.	98



DİFÜZYON AĞLARI İLE GÖRÜNTÜ REKONSTRÜKSİYONU VE RESTORASYONU

ÖZET

Bu tez çalışması, görüntü işleme alanında son yıllarda öne çıkan difüzyon temelli derin öğrenme yaklaşımlarının, görüntü rekonstrüksiyonu ve restorasyonu görevlerindeki etkinliğini araştırmayı amaçlamaktadır. Rekonstrüksiyon ve restorasyon süreçleri, özellikle eksik, bozulmuş veya düşük çözünürlüklü görüntülerden orijinal yapının yeniden elde edilmesi amacıyla yürütülmekte olup; tıbbi görüntüleme, güvenlik sistemleri, uydu görüntüleme ve kültürel mirasın dijitalleştirilmesi gibi birçok kritik alanda büyük önem taşımaktadır.

Çalışmanın başlangıcında klasik yöntemler (frekans domaini teknikleri, matris tamamlama yaklaşımları, varyasyonel modelleme vb.) detaylandırılmış, ardından derin öğrenme temelli modellerin (özellikle konvolüsyonel sinir ağları (convolutional neural network, CNN), otokodlayıcı, üretici çekişmeli ağlar (generative adversarial network, GAN) gibi) bu alanlara getirdiği katkılar ele alınmıştır. Ancak bu geleneksel ve erken dönem derin öğrenme yaklaşımlarının eksik veri, gürültü ve düşük çözünürlük gibi zorlu senaryolarda yapısal bütünlüğü ve görsel kaliteyi aynı anda sağlayamadıkları gösterilmiştir.

Difüzyon modelleri, bu zorlukların üstesinden gelmek amacıyla geliştirilen, olasılıksal temelli bir görüntü üretim ve restorasyon metodudur. Bu modeller, bir görüntüyü iteratif olarak gürültülendirip yeniden oluşturarak çalışır ve özellikle yapısal tutarlılığın korunması, detayların hassas biçimde işlenmesi ve yüksek kaliteli sonuçlar üretilmesi açısından geleneksel yöntemlere göre önemli avantajlar sunar. Tezde, gürültü giderici olasılık modeli (denoising diffusion probabilistic model, DDPM), tak ve çalıştır difüzyon, artık kaydırmalı difüzyon gibi yöntemler ve gürültü giderici difüzyon restorasyon modelleri (denoising diffusion restoration models, DDRM), görüntü restorasyonu için verimli difüzyon modelleri (efficient diffusion model for image restoration, DiffIR) ve derin tak ve çalıştır görüntü restorasyonu (deep plug-and-play image restoration, DPIR) uygulamalar detaylıca incelenmiştir.

DeneySEL çalışmalar, difüzyon modellerinin görüntü tamamlama ve süper çözünürlük görevlerinde U-Net gibi doğrudan öğrenmeye dayalı yaklaşımlara kıyasla daha iyi yapısal benzerlik (structural similarity index measure, SSIM), piksel bazlı hata (peak signal-to-noise ratio, PSNR) ve algısal benzerlik (learned perceptual image patch similarity, LPIPS) metriklerine ulaştığını ortaya koymuştur. Özellikle artık kaydırma yöntemi ile difüzyon sürecinin hızlandırılması ve örnekleme kalitesinin korunması yönünde önemli katkılar sağlanmıştır. Bu da hem akademik literatürdeki güncel gelişmelerle uyumlu sonuçlar üretmekte hem de pratik uygulamalarda kullanım potansiyelini göstermektedir.

Bu çalışmada elde edilen nicel ve nitel sonuçlar sistematik biçimde değerlendirilmiştir. U-Net ve difüzyon tabanlı modeller arasındaki karşılaştırmalı analizlerde, difüzyon modelleri birçok senaryoda daha üstün PSNR ve SSIM değerleri sağlamıştır. Görsel kalite açısından da difüzyon temelli ağların daha gerçekçi, daha az yapaylık içeren ve yapısal olarak tutarlı çıktılar ürettiği gözlemlenmiştir. Ancak difüzyon modellerinin yüksek hesaplama maliyeti, özellikle örnekleme sürecindeki zaman ve bellek yükü, hâlen çözülmesi gereken bir sınırlılık olarak öne çıkmaktadır. Bu bağlamda artık kaydırma ve latent difüzyon gibi alternatif yapılarla modelin verimliliği artırılmaya çalışılmıştır.

Sonuç olarak, bu tezde difüzyon temelli modellerin yalnızca teorik olarak değil, gerçek dünya problemlerine uygulanabilirliği bakımından da güçlü bir çözüm sunduğu gösterilmiştir. Özellikle yapısal bütünlüğün kritik olduğu görevlerde klasik yöntemleri ve geleneksel derin öğrenme mimarilerini geride bırakarak daha etkili sonuçlar üretmiştir. Bu doğrultuda difüzyon modellerinin gelecekteki görüntü işleme uygulamalarında daha yaygın biçimde kullanılması beklenmektedir. Ayrıca çalışmanın sonunda sınırlılıklar tartışılmış ve hesaplama verimliliğini artırmaya yönelik öneriler sunulmuştur.

IMAGE RECONSTRUCTION AND RESTORATION WITH DIFFUSION NETWORKS

SUMMARY

This thesis explores the effectiveness of diffusion-based deep learning models in the tasks of image reconstruction and restoration—two of the most critical problems in the field of computer vision and image processing. These tasks are essential in real-world applications such as medical imaging, satellite observation, security systems, and digital preservation of cultural heritage, where the goal is to retrieve or restore lost or degraded information with high fidelity and structural integrity.

At the outset of this study, traditional image reconstruction and restoration approaches were investigated, including frequency domain techniques, matrix completion strategies, and variational methods. Although these classical algorithms offer interpretable mathematical formulations and moderate performance in ideal conditions, they tend to fall short when dealing with complex, high-dimensional, or noisy data. Especially in cases involving missing pixels, occlusions, noise corruption, or low-resolution imaging, traditional techniques cannot adequately preserve structural information while delivering visually plausible outputs.

With the rise of deep learning, convolutional neural networks (CNN), autoencoders, generative adversarial networks (GAN), and U-Net architectures have significantly transformed image enhancement pipelines. These models have shown considerable improvements in tasks like denoising, inpainting, super-resolution, and deblurring, often surpassing classical methods in quantitative and qualitative metrics. However, it has been observed that such models—despite their strength in learning complex mappings—still face challenges in highly ill-posed scenarios. Specifically, they may fail to generate outputs that are both perceptually accurate and structurally faithful, especially when the input data is extremely degraded or incomplete.

Diffusion models represent a new generation of probabilistic generative models that address these shortcomings by learning the data distribution through a gradual noising and denoising process. In essence, these models corrupt input images by adding Gaussian noise over several steps and then learn to reverse this process through a series of learned denoising steps. This approach, inspired by stochastic differential equations, has been shown to produce high-quality, structurally consistent outputs that align well with human perception.

In this study, several diffusion-based frameworks were explored and implemented, including the denoising diffusion probabilistic model (DDPM), plug-and-play diffusion, and residual shifting diffusion, as well as advanced methods like diffusion models for image restoration (DiffIR), denoising diffusion restoration models (DDRM), and deep plug-and-play image restoration (DPIR). These models were

evaluated in terms of their performance on inpainting and super-resolution tasks using standard metrics such as peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), and learned perceptual image patch similarity (LPIPS).

Experimental results revealed that diffusion models consistently outperformed U-Net-based architectures in both pixel-wise accuracy and perceptual quality. For example, in image completion scenarios, diffusion models achieved significantly higher SSIM and lower LPIPS scores, indicating better preservation of image structure and improved perceptual realism. The residual shifting technique, in particular, was effective in accelerating the sampling process without sacrificing output quality. This architectural enhancement improves sampling efficiency—a known bottleneck in diffusion models—by introducing a residual mechanism that enables faster convergence and improved visual results even in early iterations.

Moreover, the diffusion models demonstrated strong generalization across different degradation types. Their probabilistic nature allowed them to model complex conditional distributions and generate outputs with richer details, even when the input data was severely corrupted. Unlike deterministic models like U-Net, which may tend to produce oversmoothed or repetitive patterns in uncertain regions, diffusion models could adaptively sample multiple plausible solutions, thereby providing outputs that were more natural and diverse while retaining consistency with the available context.

The thesis further provides an extensive quantitative comparison between conventional models and diffusion-based methods, showing that the latter achieve higher PSNR and SSIM values in most test cases. In terms of perceptual similarity, diffusion models exhibited lower LPIPS values, which correspond to improved perceptual closeness to ground truth images as judged by learned deep feature representations. The improvement in LPIPS is particularly important, as it suggests that diffusion models produce images that are not only numerically accurate but also visually more convincing to human observers.

Despite these advantages, diffusion models are not without their limitations. Their computational cost remains significantly higher compared to feedforward CNN-based models due to the iterative sampling process. Each denoising step requires a separate forward pass through the model, which, when accumulated across dozens or hundreds of steps, results in longer inference times and higher memory consumption. This limitation restricts the practicality of diffusion models in real-time or resource-constrained environments. To mitigate this issue, this thesis explores latent diffusion and residual shifting strategies that aim to reduce the number of sampling steps while maintaining or improving reconstruction quality.

The use of autoencoders as part of latent diffusion frameworks also plays a pivotal role in reducing dimensionality and accelerating inference. By operating in a compressed latent space, diffusion models can reconstruct images with fewer steps and reduced computational burden. The integration of variational autoencoders (VAEs) with diffusion processes demonstrates promising results in balancing quality and efficiency, especially in high-resolution tasks like medical image restoration or satellite data enhancement.

In conclusion, this thesis illustrates that diffusion-based models offer a compelling alternative to both traditional reconstruction methods and early deep learning architectures. Their ability to produce high-fidelity images from incomplete, noisy, or low-resolution inputs positions them as powerful tools in critical applications where visual quality and structural accuracy are non-negotiable. Particularly in tasks involving medical imaging or historical document restoration—where data loss is irreversible and reconstruction must preserve fine details—diffusion models show substantial potential.

The research also suggests future directions for enhancing the efficiency and scalability of these models. Techniques such as guided sampling, conditional diffusion, hybrid transformer-diffusion architectures, and faster sampling schedulers (e.g., denoising diffusion implicit models, DDIM, diffusion probabilistic model solver) may further optimize the performance of these systems. By addressing current computational limitations, these advancements could facilitate the broader adoption of diffusion models in both academic and industrial applications.

The extended evaluations and comparative analyses presented in this thesis provide strong evidence that diffusion networks, especially when equipped with recent innovations like residual shifting and latent sampling, represent a state-of-the-art solution for image reconstruction and restoration. They not only outperform traditional methods but also set a new standard in generating visually coherent and structurally accurate images under challenging conditions.



1. GİRİŞ

Görüntü işleme alanı, son on yılda derin öğrenme temelli yaklaşımların gelişimiyle birlikte önemli bir dönüşüm geçirmiştir. Özellikle yapay sinir ağlarının, görüntülerdeki karmaşık yapıları modelleyebilme kapasiteleri, klasik yöntemlerin yerini giderek bu veri odaklı modellere bırakmasına neden olmuştur. Bu çerçevede konvolüsyonel sinir ağları (convolutional neural networks, CNN), otokodlayıcılar, çekişmeli üretici ağlar (generative adversarial networks, GAN) ve varyasyonel otokodlayıcılar (variational autoencoders, VAE) gibi modeller, görüntü sentezinden görüntü restorasyonuna, süper çözünürlükten ve segmentasyona kadar birçok temel görevde başarıyla uygulanmıştır. Ancak bu modellerin özellikle eksik veri, gürültü, düşük çözünürlük ya da bozulmuş piksellerin söz konusu olduğu durumlarda, hem yapısal tutarlılığı korumakta hem de görsel kaliteyi artırmakta sınırlı kalabildiğini görmekteyiz.

Bu sınırlamaların aşılması amacıyla geliştirilen difüzyon temelli modeller, son yıllarda yapay görsel üretim ve görüntü iyileştirme gibi alanlarda çığır açıcı sonuçlar vermiştir. Difüzyon modelleri, temel olarak bir görüntüyü kademeli olarak gürültüye çeviren ve ardından bu süreci tersine çevirerek temiz görüntüyü yeniden oluşturan olasılıksal yapılarıdır. Bu yöntem, özellikle gürültünün matematiksel modellemesini dikkate alması ve olasılıksal örnekleme üzerinden çalışması sayesinde, klasik deterministik yöntemlere kıyasla daha doğal, yapısal olarak daha tutarlı ve görsel açıdan daha zengin sonuçlar sunmaktadır.

Difüzyon modellerinin görüntü rekonstrüksiyonu ve restorasyonu gibi görevlerdeki performansı ise özel bir öneme sahiptir. Bu tür işlemler, orijinal görüntülerin doğrudan gözlemlenemediği, yalnızca bozulmuş ya da eksik versiyonlarının elde edilebildiği durumları ifade eder. Örneğin manyetik rezonans görüntüleme (magnetic resonance imaging, MRI), bilgisayarlı tomografi (computed tomography, CT) veya uydu görüntüleme gibi alanlarda, kısıtlı ölçüm, donanım sınırlamaları ya da veri iletiminde oluşan kayıplar nedeniyle görüntülerin bir kısmı eksik veya bozulmuş

olabilir. Difüzyon modellerinin bu eksikliği kapatmak üzere hem veriye dayalı öğrenme hem de olasılıksal modelleme ile güçlü bir yaklaşım sunduğunu görmekteyiz.

Görüntü rekonstrüksiyonu ve restorasyonu görevleri, yalnızca teknik birer problem olmanın ötesinde, tıbbi teşhis, güvenlik, uzaktan algılama, endüstriyel kalite kontrol ve kültürel mirasın dijitalleştirilmesi gibi hayati uygulamalara sahiptir. Bu nedenle bu alandaki her iyileşme, yalnızca görüntü kalitesini artırmakla kalmaz, aynı zamanda karar verme süreçlerinde doğruluk, güvenilirlik ve zaman kazancı sağlar.

Bu tez çalışmasında, görüntü işleme disiplininin temel ilkeleri tanıtıldıktan sonra, difüzyon temelli ağların kuramsal yapısı detaylı biçimde ele alınacaktır. Ardından, difüzyon ağlarının klasik yöntemlerle kıyaslandığı teorik ve u şekilli evrişimli sinir ağı (U-Shaped Neurel Network) U-Net tabanlı mimarilerle kıyaslandığı deneysel çalışmalar sunulacaktır. Özellikle süper çözünürlük ve iç boyama görevleri özelinde yapılan değerlendirmelerle, difüzyon ağlarının yapısal bütünlük, tepe sinyal-gürültü oranı (peak signal-to-noise ratio, PSNR), yapısal benzerlik indeksi ölçüsü (structural similarity index measure, SSIM), öğrenilmiş algısal görüntü benzerliği (learned perceptual image patch similarity, LPIPS) gibi metrikler açısından ne kadar başarılı olduğu ortaya konulacaktır.

Böylelikle bu tez, difüzyon ağlarının sadece teorik bir model olmanın ötesinde, gerçek dünya problemlerinde uygulanabilirliğini test eden kapsamlı bir çalışma olarak kurgulanmıştır.

1.1 Tezin Amacı

Bu tez çalışmasının temel amacı, görüntü rekonstrüksiyonu ve restorasyonu alanında son yıllarda öne çıkan difüzyon temelli ağların performansını sistematik olarak incelemek, bu yöntemleri geleneksel yaklaşımlar ve diğer derin öğrenme tabanlı modellerle karşılaştırmak ve görüntü iyileştirme görevlerindeki etkinliğini deneysel olarak ortaya koymaktır.

Tez kapsamında öncelikle görüntü işleme disiplininin teorik temelleri, klasik rekonstrüksiyon/restorasyon teknikleri ve bu alanlarda yaygın olarak kullanılan doğrusal ve doğrusal olmayan yöntemler tanıtılacaktır. Daha sonra, derin öğrenme

devriminin bu alana etkisi incelenerek CNN, U-Net, GAN ve otokodlayıcı temelli yaklaşımların başarıları ve sınırlılıkları tartışılacaktır.

Difüzyon modellerinin teorik çerçevesi sunularak, bu yöntemlerin gürültü modelleme, olasılıksal örnekleme ve yapısal bütünlük açısından ne tür avantajlar sağladığı ele alınacaktır. Özellikle gürültü giderici difüzyon olasılıksal modelleri (denoising diffusion probabilistic model, DDPM), gürültü giderici difüzyon restorasyon modelleri (denoising diffusion restoration models, DDRM), derin tak ve çalıştır görüntü restorasyonu (deep plug-and-play image restoration, DPIR) ve artık kaydırmalı difüzyon gibi çeşitli difüzyon tabanlı mimarilerin hem yöntemsel hem uygulama düzeyindeki farkları ayrıntılı biçimde değerlendirilecektir.

Bu bağlamda çalışmanın odak noktası, difüzyon ağlarının;

- Eksik, bozulmuş veya düşük çözünürlüklü görüntülerin yeniden yapılandırılmasındaki başarısını test etmek,
- Görsel kaliteyi artıran ve yapısal bozulmayı azaltan yönlerini analiz etmek,
- Geleneksel yöntemler ile derin öğrenme temelli alternatiflere kıyasla hangi durumlarda daha üstün performans sergilediğini ortaya koymaktır.

Tezin sonunda, yapılan deneysel çalışmalar üzerinden elde edilen nicel PSNR, SSIM, LPIPS metrikleri ve nitel görsel kalite değerlendirmeleri ışığında, difüzyon modellerinin görüntü işleme alanındaki potansiyeli ve uygulama verimliliği tartışılacak; aynı zamanda mevcut sınırlamalar ve geliştirme alanları için öneriler sunulacaktır.

1.2 Literatür Araştırması

Görüntü rekonstrüksiyonu ve restorasyonu alanında klasik yöntemlerden derin öğrenmeye doğru kayış, son yıllarda literatürde önemli bir yaklaşım değişimine işaret etmektedir. Geleneksel yaklaşımlar, özellikle frekans domaini teknikleri, ters problemlerin çözümü için doğrusal cebir yöntemleri ve varyasyonel modelleme gibi yöntemler etrafında şekillenmiştir. Ancak bu yöntemler, gürültü, eksik veri, düşük çözünürlük gibi zorlu koşullarda yapısal bilgiyi korumakta ve gerçekçi sonuçlar üretmekte sınırlı kalmıştır.

Derin öğrenmenin görüntü işleme alanına entegrasyonu, CNN ve otokodlayıcı gibi mimarilerin kullanımıyla başlamıştır. Özellikle U-Net mimarisi, görüntü rekonstrüksiyonu ve restorasyonu görevlerinde hem düşük seviyeli detayların hem de yüksek seviyeli yapısal bilgilerin birlikte modellenmesine olanak sağlamıştır. Bununla birlikte, GAN tabanlı yöntemler, süper çözünürlük ve görüntü tamamlama gibi görevlerde görsel kalitenin artırılması yönünde çığır açıcı sonuçlar sunmuştur. Ancak GAN'ların eğitim zorlukları ve üretilen verinin temsiline dair eksiklikler literatürde sıkça eleştirilmiştir.

Bu bağlamda difüzyon temelli modeller, yapısal bütünlüğü koruyarak yüksek kaliteli görüntüler üretebilme potansiyelleriyle dikkat çekmiştir. Ho ve arkadaşları tarafından önerilen DDPM [2], bir görüntünün kademeli olarak gürültülendirilmesi ve ardından adım adım temizlenmesi fikri üzerine inşa edilmiştir. Ardından Song ve Ermon [3], skor bazlı generatif modelleme yaklaşımı ile, gürültülü veri dağılımlarının skor fonksiyonları üzerinden modellenmesini sağlayarak difüzyon sürecinin teorik temelini genişletmişlerdir.

Ters problem çözümünde difüzyon modellerinin uygulamaları, literatürde artan şekilde ele alınmaktadır. Sun ve arkadaşları [4] tarafından geliştirilen tak ve çalıştır difüzyon yaklaşımı, mevcut kayıp operatörlerini dikkate alarak esnek bir çözüm sunmuştur. Benzer şekilde, Wang ve arkadaşlarının sunduğu gürültü giderici difüzyon restorasyon modelleri [5] çalışması, restorasyon problemlerinde difüzyon modellerinin hem nicel hem de algısal kalite açısından üstün performans gösterdiğini ortaya koymuştur.

Öte yandan, difüzyon modellerinin yüksek hesaplama maliyeti ve örnekleme süresinin uzunluğu gibi sınırlamaları da tartışılmakta; artık kaydırmalı difüzyon ve latent difüzyon gibi yöntemlerle bu zorluklara çözüm aranmaktadır.

Sonuç olarak, güncel literatür, difüzyon temelli yaklaşımların yalnızca görsel sentez değil, aynı zamanda bozulmuş verinin yeniden yapılandırılması görevlerinde de klasik ve derin öğrenme tabanlı yöntemlerle rekabet edebilecek düzeye ulaştığını göstermektedir.

1.3 Hipotez

Görüntü rekonstrüksiyonu ve restorasyonu alanında klasik yöntemler ile derin öğrenme tabanlı yaklaşımlar çeşitli başarılar elde etmiş olsa da, özellikle eksik veri, ağır bozulma ve düşük çözünürlük senaryolarında yapısal bütünlüğü ve görsel kaliteyi korumakta sınırlı kalabilmektedirler. Bu bağlamda, difüzyon temelli ağlar, olasılıksal modelleme yetenekleri sayesinde hem düşük seviyeli detayların hem de yüksek seviyeli yapısal bilgilerin daha etkili bir şekilde restore edilebileceği iddiasını taşımaktadır.

Bu tez çalışmasında test edilen temel hipotez, difüzyon modellerinin geleneksel görüntü rekonstrüksiyonu ve restorasyonu yöntemlerine kıyasla daha yüksek doğrulukta ve görsel kalitesi daha üstün çıktılar üretebileceğidir. Özellikle:

- Görüntü tamamlama,
- Gürültü giderme,
- Süper çözünürlük

gibi görevlerde, difüzyon modellerinin daha düşük hata oranları ve daha yüksek yapısal benzerlik indeksleri sunacağı öngörülmektedir.

Literatürde Ho ve arkadaşları [2] tarafından tanıtılan DDPM yaklaşımı, kademeli örnekleme ile yüksek kaliteli sentezler üretebileceğini göstermiştir. Song ve Ermon'un [3] skor bazlı yöntemleri ise, gürültülü gözlemlerden orijinal sinyalin daha doğru tahmin edilebileceğini ileri sürmüştür. Ayrıca, Sun ve arkadaşları [4] tarafından önerilen tak ve çalıştır (plug and play, PnP) yaklaşımı ile Wang ve arkadaşları [5] tarafından geliştirilen DDRM yöntemleri, difüzyon modellerinin restorasyon problemlerinde klasik yöntemlere kıyasla daha iyi sonuçlar sunduğunu deneysel olarak kanıtlamıştır.

Dolayısıyla bu tezde, difüzyon modellerinin kritik uygulamalardaki rekonstrüksiyon ve restorasyon görevlerinde, hem PSNR, SSIM, LPIPS gibi nicel metrikler hem

de nitel deęerlendirmeler aısından klasik yontemlere gore stun performans sergileyeceęi hipotezi deneysel olarak test edilecektir.



2. GÖRÜNTÜ REKONSTRÜKSİYONU VE RESTORASYONU

Görüntü rekonstrüksiyonu ve restorasyonu, görüntü işleme alanının temel alt dallarındandır. Her iki kavram da, bozulmuş, eksik veya kısıtlı gözlem verilerinden orijinal görüntünün yüksek doğrulukla yeniden elde edilmesini amaçlamaktadır. Bu süreçler, özellikle tıbbi görüntüleme, uzaktan algılama, güvenlik ve kültürel mirasın dijitalleştirilmesi gibi kritik uygulamalarda büyük öneme sahiptir. Rekonstrüksiyon, genellikle fiziksel ölçüm süreçlerinden gelen verilerin yorumlanmasına yönelikken; restorasyon, mevcut görüntü üzerindeki bozulmaları düzeltmeye odaklanır.

2.1 Görüntü Rekonstrüksiyonu

Görüntü rekonstrüksiyonu, doğrudan gözlemlenemeyen veya eksik olan bir görüntünün, fiziksel bir gözlem modeli altında yeniden oluşturulması sürecidir. Bu süreçte temel amaç, gözlem cihazından elde edilen veriden, olabildiğince doğru bir biçimde orijinal görüntüyü geri kazanmaktır. Bu alan, özellikle tıbbi görüntüleme (MRI, CT), uzaktan algılama ve endüstriyel kalite kontrol gibi birçok kritik uygulamada temel bir rol oynamaktadır.

Görüntüleme sistemlerinin çoğunda, orijinal bir görüntüyü ya da sinyali doğrudan gözlemleyemeyiz. Onun yerine görüntü ya da sinyalin fiziksel bir ölçüm operatörü tarafından dönüştürülmüş halini gözlemleriz. Gözlem operatörü rekonstrüksiyon sürecini matematiksel olarak anlamamıza yardımcı olur.

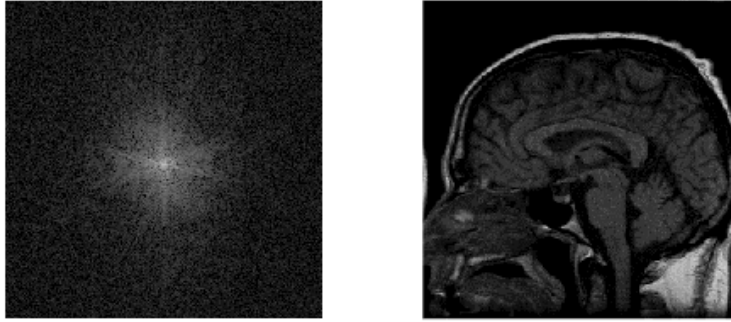
Görüntüleme sistemi aşağıdaki şekilde modellenilebilir:

$$y = A(x) + n \quad (2.1)$$

Bu denklemde:

- x : Gerçek görüntüyü (rekonstrüksiyonla elde etmek istediğimiz veriyi),
- A : Gözlem operatörü gerçek dünyadaki ölçüm sistemini(Örneğin, MRI sistemlerinde genellikle Fourier dönüşümü, CT sistemlerinde ise Radon dönüşümüdür [6].),
- y : Ölçülen ya da gözlenen veriyi
- n : Ölçüm gürültüsünü (sensör hataları, çevresel faktörler vb.) ifade eder.

Amaç, gözlem y verildiğinde orijinal görüntü x 'i yaklaşık olarak geri elde etmektir. Şekil 2.1'deki MR görüntüleme sürecinde gözlem operatörü ile k -uzayından nasıl görüntü elde edildiğini görsel olarak anlamanıza yardımcı olacaktır.



Şekil 2.1: K-uzayı ve uzay domaini manyetik rezonans görüntüleme.

MRI sistemlerinde kullanılan sıkıştırılmış algılama teknikleri, ölçüm süresini kısaltmak için k -uzayının eksik örneklenmesini ile sağlanır. Bu durumda görüntünün doğru biçimde geri çatılabilmesi için klasik aralıklı çözümler ya da derin öğrenme tabanlı yaklaşımlar kullanılmaktadır [7].

Derin öğrenme yöntemlerinin bu alana entegrasyonu, öğrenilmiş modellemelerle regularizasyonu birleştirerek önemli bir yaklaşım değişimi yaratmıştır. Özellikle CNN ve U-Net mimarileri, görüntü rekonstrüksiyonu görevinde detaylı kenar bilgisi, doku sürekliliği ve yapısal bütünlük açısından önemli başarılar elde etmiştir [8].

Sonuç olarak, görüntü rekonstrüksiyonu yalnızca bir görüntünün yeniden elde edilmesi değil, aynı zamanda bilgi kaybının telafisi, yapısal bütünlüğün korunması ve görsel anlamın geri kazanımı açısından çok boyutlu bir probleme dönüşmektedir. Bu bağlamda, klasik yöntemlerin yanı sıra yeni nesil öğrenme tabanlı modellerin ve özellikle difüzyon temelli yaklaşımların performansları, bu tez kapsamında detaylı olarak incelenecektir.

2.1.1 Görüntü rekonstrüksiyonuna olan ihtiyaç

Görüntü rekonstrüksiyonu, yalnızca görüntü kalitesini artırmaya yönelik bir işlem olmaktan öte, veri eksikliğinin olduğu durumlarda anlamlı bilgi üretmeyi mümkün kılması açısından kritik bir öneme sahiptir. Pek çok modern görüntüleme sistemi, fiziksel, teknolojik ya da zaman kaynaklı kısıtlar nedeniyle eksik, bozulmuş veya düşük çözünürlüklü veriler üretmektedir. Bu tür durumlarda, rekonstrüksiyon teknikleri orijinal görüntünün tahmin edilmesini mümkün kılarak hem insan yorumunu hem de otomatik sistemlerin karar alma süreçlerini doğrudan etkiler.

Aşağıda, bu ihtiyacın doğduğu başlıca nedenler özetlenmiştir:

- **Görüntü birleştirme ve tam Görünürlük sağlama:** Tıbbi görüntüleme sistemleri (MRI, CT, PET), organları farklı açılardan tarayarak kesitsel veriler üretir. Bu kesitlerin üç boyutlu olarak birleştirilmesi ve tam bir görünüm oluşturulması rekonstrüksiyon işlemleriyle mümkündür [6]. Bu işlem, sadece doğru teşhis için değil, aynı zamanda cerrahi planlama ve tedavi takibi gibi ileri düzey klinik karar destek süreçlerinde de gereklidir.
- **Eksik verilerin tamamlanması:** Görüntüleme cihazlarının fiziksel sınırları, sensör hataları veya veri aktarım problemleri nedeniyle görüntülerin bazı bölgelerinde bilgi eksiklikleri meydana gelebilir. Örneğin MRI çekimlerinde metal implantlar ya da hasta hareketi, belirli bölgelerde veri kaybına yol açabilir [7]. Benzer şekilde uydu görüntüleme atmosferik etkileşimler veya sinyal zayıflamaları eksik veri üretimine neden olabilir. Rekonstrüksiyon yöntemleri, bu eksik bölgelerin tahminiyle görüntünün bütünselliğini yeniden sağlar.

- **Kayıp veri ve bozulmanın giderilmesi:** Görüntüler, sıkıştırma algoritmaları, veri aktarım kanalları veya ortam koşullarından kaynaklı olarak bozulmalara maruz kalabilir. Bu tür bozulmalar, görüntüde bilgi kaybına veya artefakt oluşumuna neden olur. Özellikle eski film ve video materyallerinin dijital ortama aktarımı sırasında bozulmuş kısımların yeniden oluşturulması için rekonstrüksiyon kaçınılmazdır. Ayrıca veri güvenliği açısından kritik olan güvenlik ve gözetim sistemlerinde bozulmuş verinin onarılması, kararların doğruluğu açısından hayati öneme sahiptir.
- **görüntü iyileştirme ve çözünürlük artırma:** Özellikle uzaktan algılama, astronomi, biyomedikal mikroskopi ve video gözetimi gibi alanlarda, fiziksel sınırlamalar nedeniyle düşük çözünürlüklü görüntüler elde edilir. Bu tür görüntülerde detayların geri kazanılması süreci, süper çözünürlük tekniklerinin bir alt kümesi olarak görüntü rekonstrüksiyonunun doğrudan uygulama alanına girer. Yüksek frekanslı ayrıntıların yeniden tahmin edilmesi, yalnızca görsel iyileştirme değil; hedef tespiti, segmentasyon ve tanı gibi görevlerde doğruluğu artırır.

Sonuç olarak, görüntü rekonstrüksiyonuna olan ihtiyaç yalnızca görüntü kalitesiyle sınırlı değildir; eksik veriden bilgi üretimi, karar destek sistemlerinde doğruluğun artırılması ve yapay zekâ tabanlı uygulamaların güvenilirliğinin sağlanması gibi çok önemli nedenlerden dolayı bu alan, modern görüntüleme teknolojilerinin ayrılmaz bir parçası haline gelmiştir.

2.1.2 Teknikler

Görüntü rekonstrüksiyonu, problemin doğasına ve gözlem operatörünün biçimine göre farklı teknik yaklaşımlarla uygulanabilir. Bu yöntemler temel olarak üç ana grupta toplanabilir:

- Frekans domaini teknikleri,
- Matris tamamlama ve optimizasyon temelli yaklaşımlar,
- Derin öğrenme temelli yöntemler

Her yöntemin güçlü ve zayıf yönleri, uygulanabilirlik alanına, veri yapısına ve işlem süresine bağlı olarak değişmektedir. Aşağıda bu yöntemlerin her biri detaylı biçimde ele alınmaktadır.

2.1.2.1 Frekans domaini teknikleri

Frekans domaini teknikleri, sinyal işleme teorisinin klasik araçlarından biri olan Fourier dönüşümüne dayanır. Bu teknikler, bir görüntüyü uzay domaininden (piksel bazlı temsil) frekans domainine taşıyarak işlem yapmayı mümkün kılar. Bu dönüşüm sayesinde düşük ve yüksek frekans bileşenleri ayrıştırılabilir; bu da özellikle kenar iyileştirme, bozulma giderme ve detay geri kazanımında etkili olur.

Basitçe 2 boyutlu bir görüntü üzerinden örnekleme gerekirse Fourier dönüşümü, bir görüntüdeki her bir pikselin frekans bileşenlerine dönüşmesini sağlar. Bu dönüşüm sayesinde, görüntü yüksek frekans ve düşük frekans bileşenlerine ayrılır.

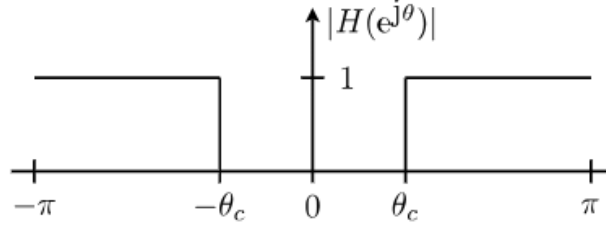
Görüntü $f(x,y)$ 'yi frekans domainine dönüştürmek için iki boyutlu (2 dimensional, 2D) Fourier dönüşümü kullanılır:

$$F(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)} \quad (2.2)$$

Burada:

- $f(x,y)$ orijinal görüntüyü,
- $F(u,v)$ Fourier dönüşümünden elde edilen frekans domainindeki görüntüyü,
- M, N : Görüntünün yatay ve dikey boyutları temsil eder.

Daha sonra görüntünün keskinliğini artırmak için yüksek frekansları artırmaya yönelik işlemler yapılır. Bu işlem genellikle şekil 2.2' deki gibi yüksek geçiren filtrelerle yapılır.



Şekil 2.2: İdeal yüksek geçiren filtre.

Yüksek geçiren filtreler, düşük frekansları baskılar ve yalnızca yüksek frekansları geçirir. Bu, görüntüdeki kenarları daha belirgin hale getirir. Bu tip filtre, Fourier dönüşümünde aşağıdaki şekilde uygulanabilir:

$$G(u, v) = H(u, v) \cdot F(u, v) \quad (2.3)$$

Burada:

- $H(u, v) = 1 - e^{-\alpha(u^2+v^2)}$ yüksek geçiren filtre fonksiyonunu,
- $G(u, v)$ alçak frekansları filtrelenmiş görüntüyü temsil eder.

Bu işlem, Ters Fourier Dönüşümü ile uzay domainine geri döndürülerek keskinleştirilmiş görüntü elde edilir:

$$g(x, y) = \mathcal{F}^{-1}(G(u, v)) \quad (2.4)$$

Bu teknikler özellikle CT ve MRI gibi sistemlerde önemli rol oynar. Özellikle filtrelenmiş geri projeksiyon yöntemi, CT sistemlerinde frekans domaininde uygulanan en yaygın rekonstrüksiyon algoritmalarından biridir [6].

Ancak bu yöntemler, eksik veriye karşı duyarlıdır ve gürültüye karşı dayanıklılıkları sınırlıdır. Bu nedenle çoğu zaman farklı yaklaşımlarla desteklenmeleri gerekir.

2.1.2.2 Matris tamamlama teknikleri

Matris tamamlama teknikleri, görüntü rekonstrüksiyonunda eksik piksellerin ya da gözlemlenmemiş bölgelerin tahmin edilmesinde yaygın olarak kullanılan matematiksel

yaklaşımlardandır. Bu yöntemlerin temelinde, eksik veri içeren bir matrisin belirli yapısal varsayımlar altında yeniden tahmin edilmesi yer alır.

Bu tür problemlerde, elimizde yalnızca gözlemlenmiş $\Omega \subset \mathbb{R}^{m \times n}$ elemanları olan bir matris M bulunur. Amaç, M 'nin tamamını doğru biçimde yeniden oluşturmaktır.

Yaygın çözüm stratejisi, bu matrisi düşük ranklı iki çarpan $U \in \mathbb{R}^{m \times r}$ ve $V \in \mathbb{R}^{n \times r}$ şeklinde çarpanlara ayırmak ve bu çarpanları optimize etmektir:

$$M \approx UV^T \quad (2.5)$$

Bu probleme yönelik en yaygın iki optimizasyon yöntemi şunlardır:

- **Stokastik gradyan düşüşü (stochastic gradient descent, SGD) Yöntemi**

SGD yöntemi, matris tamamlama problemleri için kullanılan bir optimizasyon tekniğidir. Matrisin faktörizasyonunu yaparak $M \approx UV^T$ şeklinde ifade edebiliriz. SGD, bu faktörlerin her bir elemanını iteratif olarak günceller.

Hedef fonksiyon şu şekilde tanımlanır:

$$\mathcal{L}(U, V) = \sum_{(i,j) \in \Omega} (M_{ij} - U_i V_j^T)^2 + \lambda (\|U\|_F^2 + \|V\|_F^2) \quad (2.6)$$

Burada:

- M_{ij} , gözlemlenen değerleri,
- U_i ve V_j , sırasıyla U ve V matrislerinin i -inci ve j -inci satırları,
- λ , regularizasyon parametresini,
- $\|\cdot\|_F$ Frobenius normunu temsil eder.

SGD algoritması, bu hedef fonksiyonu minimize etmek için parametreleri iteratif olarak günceller.

Bu yöntem, özellikle büyük ölçekli verilerde düşük bellek gereksinimi ve hesaplama verimliliği nedeniyle tercih edilmektedir [9].

• Ridge regresyon yöntemi

Ridge regresyonu, matris faktörizasyonu yaparken, faktörlerin küçük olmasını sağlayan bir regularizasyon tekniğidir. Bu, matris tamamlama işleminde de kullanılabilir. Ridge regresyonu, aşağıdaki şekilde ifade edilir:

$$\min_{U,V} \sum_{(i,j) \in \Omega} (M_{ij} - U_i V_j^T)^2 + \lambda (\|U\|_F^2 + \|V\|_F^2) \quad (2.7)$$

Burada:

- M_{ij} , gözlemlenen verileri,
- U_i ve V_j , matris faktörizasyonunda kullanılan parametreleri,
- λ , regularizasyon parametresini,
- $\|\cdot\|_F$ Frobenius normunu temsil eder.

Bu yöntem, SGD'ye benzer bir formülasyona sahip olmakla birlikte, daha stabil ve deterministik bir çözüm sunar. Özellikle düşük ranklı matrislerde, tahmin başarısı yüksek olup genelleme yeteneği iyidir [10].

2.1.2.3 Derin öğrenme yöntemleri

Görüntü rekonstrüksiyonu problemlerinde derin öğrenme tabanlı yöntemler, hem doğruluk hem de görsel kalite açısından geleneksel tekniklerin ötesine geçen sonuçlar sunmaktadır. Özellikle konvolüsyonel sinir ağları ve kodlayıcı-kod çözücü yapıları, eksik verinin tahmini ve bozulmuş görüntülerin onarımı gibi görevlerde güçlü temsil yetenekleriyle öne çıkmıştır. Zhang ve arkadaşları tarafından önerilen gürültü giderici konvolüsyonel sinir ağları modeli [8], gürültü giderme problemlerinde kalan öğrenme temelli bir çerçeve sunarak bu alanda önemli bir başarı elde etmiştir.

U-Net gibi mimariler, özellikle tıbbi görüntüleme alanında eksik ya da az örneklenmiş verilerle çalışmada yaygın olarak kullanılmakta ve yüksek yapısal doğruluk sağlamaktadır [11,12]. Bununla birlikte, GAN tabanlı modeller de süper çözünürlük gibi görevlerde yüksek frekans detaylarını geri kazandırma açısından etkili olmuş; ancak eğitimde tutarsızlık gibi sorunlar nedeniyle bazı sınırlamalara takılmıştır [13].

Son dönemde difüzyon temelli generatif modeller, özellikle gürültülü veya eksik veriden orijinal görüntüyü örnekleme yoluyla üretmekte güçlü bir alternatif sunmaktadır. DDPM [2] ve skor bazlı modeller [3], örnekleme sürecini ters gürültü difüzyon süreci olarak ele alırken; DDRM ve PnP Difüzyon gibi yöntemler bu çerçeveyi doğrudan rekonstrüksiyon problemlerine uygulamıştır [4,5]. Bu modellerin detaylı analizi ve karşılaştırmaları ilerleyen bölümlerde sunulacaktır.

2.1.3 Uygulama alanları

Görüntü rekonstrüksiyonu teknikleri, yalnızca teorik çerçevede değil; aynı zamanda çok çeşitli uygulama alanlarında doğrudan kullanılmaktadır. Bu alanlarda temel hedef, doğrudan elde edilemeyen, bozulmuş veya eksik verilerden anlamlı ve yüksek doğrulukta görüntülerin yeniden yapılandırılmasıdır. Aşağıda bu tekniklerin öne çıktığı başlıca uygulama alanları özetlenmiştir:

- **Tıbbi görüntüleme** Tıbbi görüntüleme, görüntü rekonstrüksiyonunun en yoğun kullanıldığı ve gelişimine en çok katkı sağladığı alanlardan biridir. Manyetik rezonans görüntüleme (MRI) sistemlerinde, hızlı tarama ihtiyacı nedeniyle k-uzayı verisi eksik alınmakta ve bu eksik veriden görüntünün doğru şekilde yeniden oluşturulması gerekmektedir [7]. Benzer şekilde bilgisayarlı tomografi (CT) sistemlerinde düşük doz radyasyon uygulamaları, sınırlı açılı veya az örnekleme projeksiyon verisi ile çalışmayı gerektirir [6]. Bu gibi durumlarda, rekonstrüksiyon algoritmaları tanı doğruluğunu doğrudan etkileyen kritik bir rol oynar.
- **Uzay görüntüleme** Uydu görüntüleme sistemleri, atmosferik koşullar, optik bozulmalar veya bant genişliği sınırlamaları nedeniyle eksik ya da bozuk verilerle çalışır. Görüntü rekonstrüksiyonu, bu tür sistemlerde görüntü kalitesini artırmak, yüzey şekillerini netleştirmek ve analiz edilebilir veri üretmek için sıklıkla kullanılır. Ayrıca farklı zamanlarda alınan verilerin birleştirilmesi yoluyla eksik bölgelerin tahmini yapılabilir.

- **Görüntü sıkıştırma ve veri iletimi** Görüntülerin iletilmesi veya depolanması sırasında genellikle sıkıştırma yapılır. Ancak sıkıştırma sırasında kaybolan bilgiler, görüntü rekonstrüksiyonu ile geri kazanılabilir. Bu, özellikle video sıkıştırma, internet üzerinden yüksek çözünürlüklü görüntülerin iletilmesi gibi durumlarda önemlidir.
- **Güvenlik ve gözetim** Güvenlik kameralarından elde edilen görüntüler bazen düşük kaliteli olabilir veya çeşitli engeller nedeniyle eksik olabilir. Bu durumda, görüntü rekonstrüksiyonunun kullanılması, görüntülerin iyileştirilmesi ve kaybolan bilgilerin tamamlanması için önemlidir.
- **Hızlı görüntüleme ve gerçek zamanlı uygulamalar** Görüntü rekonstrüksiyonunun bir diğer önemli uygulama alanı, hızlı görüntüleme ve gerçek zamanlı sistemlerdir. Örneğin, endoskopik görüntüleme ya da otomatik tanıma sistemleri gibi sistemlerde, eksik verilerin hızlıca tamamlanması gereklidir.
- **Sanal gerçeklik ve artırılmış gerçeklik** Sanal gerçeklik ve artırılmış gerçeklik uygulamalarında, gerçek dünyadan alınan verilerle sanal ortamlar oluşturulurken, görüntülerin eksik veya bozulmuş bölgeleri rekonstrüksiyon teknikleriyle tamamlanabilir.
- **Endüstriyel görüntüleme ve kalite kontrol** Endüstriyel görüntüleme, üretim süreçlerinde kullanılan bir tekniktir. Görüntülerdeki eksiklikler veya hatalar, üretim kalitesini etkileyebilir. Bu tür görüntülerin rekonstrüksiyonu, hatalı ürünlerin tespitini kolaylaştırabilir. Otomotiv parçalarının üretiminde, kalite kontrol için kullanılan görüntülerdeki eksik veya bozuk bölgeler rekonstrüksiyon ile düzeltilir. Elektronik devrelerin kontrolünde kullanılan görüntülerdeki eksik veriler, üretim hatalarının erken tespiti için rekonstrüksiyonla tamamlanır.
- **Film endüstrisi** Film yapımında, eski ve hasar görmüş görüntüler, yenilenebilir ve restore edilebilir. Ayrıca, eski filmlerden yüksek çözünürlüklü versiyonlar oluşturulabilir. Düşük kaliteli video çekimlerinin iyileştirilmesi ve detayların netleştirilmesi için görüntü rekonstrüksiyonu kullanılır.

2.1.4 Önemli uygulamalar

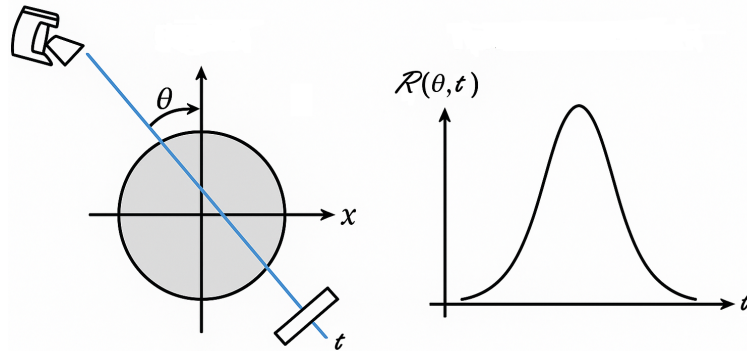
Görüntü işleme alanında önemli uygulamalar denildiğinde tabiki de ilk akla gelen tıbbi görüntüleme dir. Tıbbi görüntüleme, insan vücudunun iç yapılarının detaylı bir şekilde incelenmesini sağlayan önemli bir teknolojidir. Bu görüntüleme teknikleri, vücut içindeki farklı doku ve organları, hastalıkları ve anormallikleri tespit etmek için kullanılır. Bu yöntemlerin çoğu, fiziksel ölçümlere dayalı olarak sınırlı, gürültülü veya eksik projeksiyon verileri üretir. Bu nedenle, elde edilen sinyallerin işlenerek yüksek kaliteli görüntülere dönüştürülmesi için görüntü rekonstrüksiyonu tekniklerine ihtiyaç duyulmaktadır [6]. Aşağıda, bu yöntemlerin bazıları özetlenmiştir:

2.1.4.1 Tomografik görüntüleme teknikleri

Bilgisayarlı Tomografi (CT) tarayıcıları, vücudu X-ışınları ile farklı açılardan tarayarak kesitsel görüntüler oluşturur. Bu görüntüler, tomografik rekonstrüksiyon algoritmaları ile üç boyutlu bir harita haline getirilir. CT taramasında, X-ışını kaynağı ve dedektör, hastanın etrafında belirli açılardan döner ve projeksiyon verileri (*sinogram*) oluşturur. Bu verilerden kesitsel görüntü üretmek için Radon Dönüşümü ve Geri Projeksiyon (Backprojection) gibi teknikler kullanılır.

Radon dönüşümü

Radon dönüşümü, bir nesnenin farklı açılardan alınan projeksiyonlarının, nesnenin iç yapısını nasıl temsil ettiğini açıklayan matematiksel bir modeldir. Şekil 2.3'te gibi ifade edilir:



Şekil 2.3: Radon dönüşümü.

$$R(\theta, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - t) dx dy \quad (2.8)$$

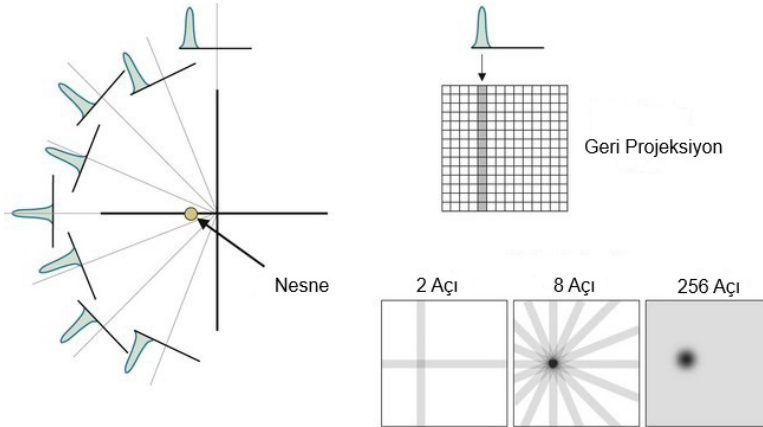
Burada:

- $R(\theta, t)$, projeksiyon verisini,
- $f(x, y)$, orijinal görüntüyü,
- θ , dönme açısını,
- t , ışın doğrultusundaki uzaklığı,
- δ , Dirac delta fonksiyonunu temsil eder.

Bu dönüşüm, bir nesnenin farklı açılardan alınan X-ışını projeksiyonlarının matematiksel temsili olarak düşünülebilir.

Filtrelenmiş geri projeksiyon

Geri projeksiyon, alınan projeksiyon verilerini tersine çevirerek görüntü oluşturur. Ancak basit geri projeksiyon bulanık görüntüler üretir. Bu yüzden filtrelenmiş geri projeksiyon yöntemi kullanılır. Şekil 2.4'teki gibi ifade edilebilir.



Şekil 2.4: Filtrelenmiş geri projeksiyon

Filtrelenmiş geri projeksiyon yöntemi, projeksiyon verilerine yüksek frekans bileşenlerini vurgulayan bir filtre uygular. Matematiksel olarak, geri projeksiyon işlemi Fourier dönüşümü ile ifade edilebilir:

- Projeksiyon verisi $R(\theta, t)$ 'nin Fourier dönüşümü alınır:

$$P(f, \theta) = \mathcal{F}[R(\theta, t)] \quad (2.9)$$

- Bir yüksek geçiren filtre uygulanır:

$$P_{\text{filtrelenmiş}}(f, \theta) = H(f) \cdot P(f, \theta) \quad (2.10)$$

Burada $H(f)$, Ram-Lak filtresi gibi bir filtre fonksiyonudur.

- Ters Fourier dönüşümü ile uzay domenine dönülür.
- Filtrelenmiş projeksiyonlar geriye projekte edilerek görüntü elde edilir.

Bu yöntem, CT rekonstrüksiyonlarında en hızlı ve yaygın kullanılan tekniktir ancak düşük doz CT verilerinde gürültüye duyarlıdır.

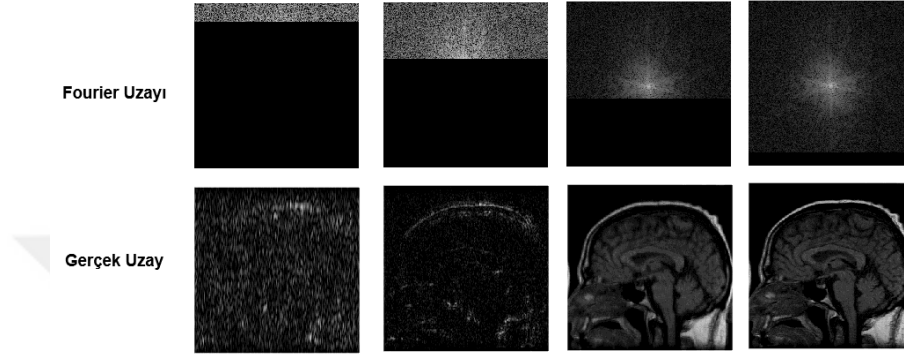
Optik koherens tomografisi (optical coherence tomography, OCT), ışığın saçılma özelliklerini kullanarak mikron seviyesinde görüntüler oluşturur. Retina taramaları ve damar incelemelerinde yaygın kullanılır. Yüksek çözünürlüklü 3 boyutlu görüntüler üretir. Fourier-domain rekonstrüksiyon yöntemleri kullanılır. Pozitron emisyon tomografisi (positron emission tomography, PET) radyoaktif izotoplar kullanarak metabolik aktivitenin ölçülmesini sağlar. Kanser teşhisinde yaygın olarak kullanılır. Moleküler düzeyde görüntüleme yapabilir. Tek foton emisyon bilgisayarlı tomografi (single-photon emission computed tomography, SPECT), PET'e benzer şekilde radyoaktif izotoplar kullanarak organ fonksiyonlarını ölçer ancak farklı dedektör sistemlerine sahiptir. Kalp ve beyin fonksiyonlarının ölçülmesinde kullanılır. İteratif algoritmalar ile rekonstrüksiyon yapılır.

2.1.4.2 Manyetik rezonans görüntüleme

MRI, güçlü manyetik alanlar ve radyo dalgaları kullanarak vücuttaki protonların hareketlerini inceler. MRI cihazları, vücut içindeki dokular arasındaki farklı proton dizilimlerinden elde edilen sinyalleri ölçerek yüksek çözünürlüklü görüntüler

oluşturur. İyonlaştırıcı radyasyon içermez, bu yüzden güvenlidir. Beyin, kaslar, bağ dokular gibi yumuşak dokuların detaylı görüntülenmesini sağlar.

MR cihazı aslında doğrudan bir görüntü çekmez. Onun yerine, görüntünün Fourier dönüşümünü yani frekans bileşenlerini ölçer. Bu dönüşüm MR cihazının topladığı sinyallerin doğal sonucudur. [7]. Toplanan sinyallere göre görüntünün gelişimi şekil 2.5 üzerinden görülebilir.



Şekil 2.5: K-uzayı ve uzay domaini görüntü gelişimi.

MR'deki gözlem süreci şu şekilde modellenilebilir:

$$y = M \cdot \mathcal{F}(x) + n \quad (2.11)$$

Bu denklemde:

- x : Rekonstrüksiyonunu yapmak istediğimiz orijinal görüntü,
- $\mathcal{F}(x)$: Görüntünün 2 boyutlu Fourier dönüşümü,
- M : Alt örnekleme maskesi hangi frekans bileşenlerinin ölçüldüğünü belirler (yani bazı k -uzayı noktaları alınmaz; bu da tarama süresinde hız kazancı sağlar),
- y : Ölçülen k -uzayı verisi (gözlem),
- n : Ölçüm gürültüsü (genellikle kompleks Gauss gürültüsü).

Bu bağlamda, gözlem operatörü A şu şekilde tanımlanır:

$$A(x) = M \cdot \mathcal{F}(x) \quad (2.12)$$

Bu operatör doğrusaldır ve çıktısı kompleks değerlidir.

$\mathcal{F}(x)$ ' i tanımlamak istediğimizde ise MRI 'ın temellerine inmemiz gerekir. MRI'nin temelinde, protonların manyetik alandaki dönme hareketi bulunur. Protonların önsel hareketi Larmor frekansı ile tanımlanır.

$$\omega_0 = \gamma B_0 \quad (2.13)$$

Burada:

- ω_0 , Larmor açısal frekansını,
- γ , jiroskopik oranı,
- B_0 , ana manyetik alanın şiddetini temsil eder.

Bu frekansta protonlara uygulanan radyo frekansı darbeleri, proton spinlerini enerji seviyeleri arasında uyararak çekirdek manyetik rezonansı oluşturur.

Gelen MRI sinyali şu şekilde modellenebilir:

$$\mathcal{F}[m, n] = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \rho[x, y] \cdot e^{-i2\pi\left(\frac{mx}{N} + \frac{ny}{N}\right)} \quad (2.14)$$

Burada:

- $\mathcal{F}(k_x, k_y)$, k -uzayında elde edilen sinyali,
- $\rho(x, y)$, Uzaydaki proton yoğunluk fonksiyonunu,
- M, N : Görüntünün yatay ve dikey boyutları temsil eder.

Bu ifade, 2D Fourier dönüşümüne eşdeğerdir. MRI görüntüleri, bu sinyallerin ters Fourier dönüşümü ile elde edilir:

$$\rho[x,y] = \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \mathcal{F}[m,n] \cdot e^{i2\pi(\frac{mx}{N} + \frac{ny}{N})} \quad (2.15)$$

Bu yöntem, tam örneklenmiş k -uzayı verisi için uygundur ancak hızlandırılmış taramalar için daha gelişmiş teknikler gereklidir.

2.1.4.3 Fotoakustik görüntüleme

Fotoakustik görüntüleme, lazer ışığı ile vücuttaki dokuların optik özelliklerini inceler ve bu ışığın emilimiyle oluşan ultrason dalgalarını algılayarak görüntü oluşturur. Yüksek kontrast ve çözünürlük sağlar. Kan damarları, tümörler ve beyin fonksiyonları gibi yapıların görüntülenmesine uygundur. Ultrason ve optik görüntülemenin avantajlarını birleştirir. Fotoakustik görüntülemede ham veriler, dalgaların zaman ve uzaydaki yayılımını içeren sinyallerdir. Ancak bu sinyaller doğrudan yorumlanabilir değildir. Görüntü rekonstrüksiyonu, bu ham verileri işleyerek anatomik veya fonksiyonel görüntüler elde etmeye yarar. Yine yukarıdaki uygulamalarda bahsedildiği gibi Fourier Dönüşümleri ve filtreleme teknikleri kullanılır.

2.1.4.4 Ultrasonografi

Ultrason, yüksek frekanslı ses dalgaları kullanarak organların ve dokuların görüntülenmesini sağlar. Gerçek zamanlı görüntüleme yapabilir. Radyasyon içermez, hamilelik takibi gibi hassas alanlarda kullanılır. Ultrason sürecinde üretilen ses dalgaları dokulara çarparak geri yansır. Geri dönen yankılar, prob tarafından toplanır ve mesafeler hesaplanır. Ampülüt bilgisi, gri seviyelerine dönüştürülerek doku yapıları görselleştirilir. Yüksek yankı üreten sert dokular beyaz, düşük yankı üreten sıvılar ise karanlık olarak görüntülenir. Kısacası görüntü rekonstrüksiyonu, ham akustik sinyalleri tıbbi olarak anlamlı görüntülere dönüştürerek tanı koyulmasını kolaylaştırır.

Tıbbi görüntüleme teknikleri, farklı dokuların ve hastalıkların detaylı bir şekilde analiz edilmesine olanak tanır. Bu süreçte görüntü rekonstrüksiyon algoritmaları, görüntü kalitesini artırmak ve hızlı tarama süreçleri sağlamak için kritik bir rol oynar. Geleneksel algoritmaların yanı sıra yapay zeka ve derin öğrenme tabanlı teknikler, modern tıbbi görüntülemenin geleceğini şekillendirmektedir.

2.2 Görüntü Restorasyonu

Görüntü restorasyonu, çeşitli faktörler nedeniyle bozulmuş görüntüleri kurtarmayı amaçlayan, görüntü işlemenin önemli bir yönüdür. Bu işlem görüntüleri orijinal durumlarına geri döndürmek için gürültü ve bulanıklık gibi kusurları düzeltmeyi içerir.

Görüntü restorasyonu, görüntü yakalama veya iletimi sırasında ortaya çıkan gürültü ve bulanıklık gibi sorunları ele alarak mevcut görüntülerin kalitesini iyileştirmeye odaklanır. Ham verilerden bir görüntünün oluşturulmasıyla ilgilenen rekonstrüksiyonun aksine, restorasyon bir görüntünün zaten mevcut olduğunu ancak iyileştirilmesi gerektiğini varsayar. Wiener filtreleme gibi yöntemler bu alanda yaygın olarak uygulanır [14].

Gözlem operatörü, gerçek görüntünün nasıl gözlemlendiğini yani nasıl bozulduğunu matematiksel olarak tanımlar.

$$y = Hx + n \quad (2.16)$$

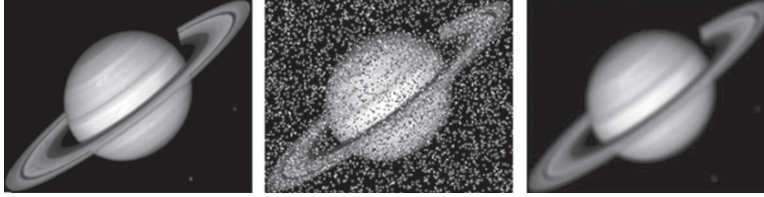
- x : Gerçek görüntü
- H : Gözlem operatörü (bozulma işlemini temsil eder)
- n : Gürültü (örneğin Gauss gürültüsü)
- y : Gözlemlenen (bozulmuş) görüntü

Bu modelde amaç, sadece y (bozulmuş görüntü) elimizdeyken, x 'i (gerçek görüntü) tahmin etmeye çalışmaktır.

2.2.1 Görüntü restorasyonuna olan ihtiyaç

Restorasyon, özellikle gürültü giderme, bulanıklık düzeltme ve diğer bozulmaların onarılması gibi işlemleri içerir. Bu yöntemlere olan ihtiyacı şu başlıklar altında inceleyebiliriz:

- **Gürültü ve bozulmaların azaltılması** Görüntüler, çeşitli nedenlerle gürültüye maruz kalabilir. Özellikle düşük ışık koşullarında veya veri iletiminde gürültüler oluşabilir. Bu tür bozulmalar, görüntü kalitesini düşürür ve yanlış sonuçlara yol açabilir. Restorasyon teknikleri, bu tür gürültüleri ortadan kaldırarak daha net görüntüler elde edilmesini sağlar. Şekil 2.6’da görüleceği gibi, örneğin Gauss gürültüsü ya da tuz ve karabiber gürültüsü gibi türler, restorasyon algoritmaları ile ortadan kaldırılabilir.



Şekil 2.6: Gürültü giderme.

- **Bulanıklık giderme** Bulanıklık, genellikle hareket veya odaklama hataları nedeniyle görüntülerde meydana gelir. Özellikle tıbbi görüntüleme, bulanıklaşmış görüntülerin netleştirilmesi, doğru teşhisler için hayati öneme sahiptir. Şekil 2.7’de bulanıklık giderildiği bir örnek görülebilir.



Şekil 2.7: Bulanıklık giderme.

2.2.2 Teknikler

Görüntü restorasyonu problemlerine yönelik geliştirilen yöntemler, tarihsel olarak klasik sinyal işleme tekniklerinden derin öğrenme temelli modellere doğru evrilmiştir. Bu bölümde, restorasyon problemlerinde kullanılan temel teknikler iki ana başlık

altında incelenmektedir: konvolüsyonel filtreleme ve frekans domaini tabanlı yaklaşımlar.

2.2.2.1 Konvolüsyonel filtreleme teknikleri

Görüntüdeki bulanıklıkların ve gürültülerin azaltılması için kullanılan temel tekniklerden biridir.

$$h(x,y) = (f * g)(x,y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f(i,j)g(x-i,y-j) \quad (2.17)$$

Burada:

- $f(x,y)$, giriş görüntüsünü,
- $g(x,y)$, filtreleme çekirdeğini,
- $h(x,y)$, filtrelenmiş görüntüyü temsil eder.

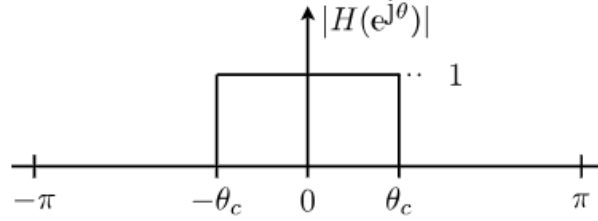
Bu işlem, $f(x,y)$ görüntüsünün her bir pikselinin, $g(x,y)$ çekirdeğiyle çarpılıp toplandığı bir işlem olarak yapılır. Konvolüsyon işlemi, genellikle bir görüntü üzerinde hareket eden bir çekirdek kullanılarak gerçekleştirilir. Örnek olarak 3×3 boyutundaki Gaussian filtresi şu şekilde olabilir:

$$g = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (2.18)$$

Bu tekniklerin en önemli avantajı hesaplama açısından düşük maliyetli olmalarıdır. Ancak öğrenmeye dayalı olmadıkları için bozulmanın türüne ve derecesine göre hassas şekilde uyarlanamazlar. Ayrıca yüksek seviyeli yapısal bilgileri modelleyemedikleri için karmaşık bozulmalarda sınırlı başarı gösterirler [14]. Örneğin yukarıdaki filtre, görüntüyü yumuşatarak gürültüyü azaltır ancak aynı zamanda netliği de azaltır.

2.2.2.2 Frekans domaini teknikleri

Fourier dönüşümü gibi yöntemlerle, görüntülerin frekans bileşenleri üzerinde işlem yaparak gürültü azaltılabilir. Bu işlemde, genellikle alçak geçiren filtreler, yüksek geçiren filtreler veya band-pass filtreler kullanılır.



Şekil 2.8: İdeal alçak geçiren filtre.

Gürültüyü ifade eden alçak frekansların filtrelendiği bir yöntemi örnek olarak verebiliriz. Görüntüyü frekans domenine dönüştürülen Fourier dönüşümü ve frekans domeninden geri dönüştüren ters Fourier dönüşümü yöntemlerinin matematiksel ifadelerini görüntü rekonstrüksiyonu bölümünde verilmişti.

$$H(u, v) = e^{-\left(\frac{u^2+v^2}{2\sigma^2}\right)} \quad (2.19)$$

Burada:

- $H(u, v)$, düşük geçiren filtreyi,
- σ , filtre genişliğini temsil eder.

Bu filtre, yüksek frekanslı detayları (kenarları, gürültüyü) bastırarak, görüntüyü yumuşatır. Ancak bu kısımda daha iyi bir filtreleme sunan Wiener Filtresine değinmeden geçmemek gerekir.

Wiener filtreleme

Wiener filtreleme, görüntüleri gürültüden temizlemek için kullanılan istatistiksel bir tekniktir. Bu yöntem, gürültülü bir sinyalin veya görüntünün, orijinal sinyaline en yakın halini elde etmek için kullanılır. Wiener filtresi, genellikle sinyalin ve gürültünün ortalama kare hatasını minimize etmeyi amaçlar.

Wiener filtresi, $F(u, v)$ Fourier dönüşümüne sahip bir görüntü üzerinde şu şekilde tanımlanabilir:

$$G(u, v) = \frac{|H(u, v)|^2}{|H(u, v)|^2 + \frac{S_n(u, v)}{S_f(u, v)}} F(u, v) \quad (2.20)$$

Burada:

- (u, v) , Wiener filtresi ile elde edilen frekans bileşenlerini,
- $F(u, v)$, giriş görüntüsünün Fourier dönüşümünü,
- $H(u, v)$, sistem transfer fonksiyonunu,
- $S_f(u, v)$, orijinal görüntünün güç spektrumunu,
- $S_n(u, v)$, gürültünün güç spektrumunu temsil eder.

Wiener filtresi, çıkış frekans bileşenini, giriş görüntüsünün frekans bileşenleri ile gürültü seviyesini dikkate alarak hesaplar.

Frekans domaini yöntemleri teorik olarak güçlü olsalar da, lokal detayların korunmasında yetersiz kalabilirler. Ayrıca, bozulmanın frekans spektrumunda açıkça ayırt edilemediği durumlarda filtreleme işlemi istenilen sonucu veremeyebilir. Bu nedenle, son yıllarda bu yaklaşımlar çoğunlukla öğrenme tabanlı yöntemlerle hibrit şekilde kullanılmaktadır [15].

2.2.3 Uygulama alanları

Görüntü restorasyonu, çok sayıda uygulama alanında kritik rol oynamaktadır. Özellikle bozulmuş, eksik ya da düşük kaliteli görsellerin iyileştirilmesi, hem insan gözüyle yorumlamayı hem de makine tabanlı analizlerin doğruluğunu doğrudan etkilemektedir. Bu yöntemlerin uygulama alanları aşağıdaki gibi özetlenebilir:

- **Tıbbi görüntüleme** : Tıbbi görüntüleme sistemleri (MRI, CT, PET gibi) sensör limitleri, hasta hareketleri veya çevresel faktörler nedeniyle eksik veya bozulmuş veri üretebilir. Bu nedenle, elde edilen görüntülerin tanı için yeterli kalitede

olabilmesi adına restorasyon teknikleri yaygın şekilde kullanılmaktadır. Özellikle metal artefakt azaltma (metal artifact reduction, MAR), hareket düzeltme ve gürültü giderme gibi işlemler, klinik doğruluk açısından büyük önem taşır [16].

- **Uzay ve uydu görüntüleme :** Uydu ve hava görüntüleri atmosferik bozulmalar, sensör gürültüsü ve sıkıştırma hataları gibi nedenlerle kalite kaybına uğrayabilir. Görüntü restorasyonu burada hem netlik artırımı hem de spektral veri iyileştirme açısından önem taşır. Özellikle süper çözünürlük teknikleri ile düşük çözünürlüklü uydu görüntüleri detaylı analizler için uygun hâle getirilebilir [17].
- **Tarihi fotoğraf ve sanat eserlerinin restorasyonu :** Tarihi fotoğraflar ve sanat eserleri, zamanla çeşitli bozulmalara uğrayabilir. Fiziksel aşınma, çizikler, renk kaybı veya yanıklar, bu tür eserlerin detaylarının kaybolmasına yol açabilir. Görüntü restorasyonu, eski eserlerin dijital ortamda onarılmasını sağlar.
- **Güvenlik ve gözetim sistemleri :** Güvenlik kameraları tarafından kaydedilen görüntüler bazen bozulabilir . Bu, suçların tespiti veya güvenlik tehditlerinin analiz edilmesi için zorluklar yaratır. Görüntü restorasyonu, güvenlik sistemlerinde kaydedilen görüntülerin iyileştirilmesi için yaygın olarak kullanılır. Bulanıklaşmış güvenlik görüntüleri, daha net hale getirilerek olayların detayları net bir şekilde gözlemlenebilir.
- **Endüstriyel görüntüleme ve denetim :** Endüstriyel alanlarda, üretim süreçlerinde kullanılan görüntüleme sistemleri, kalite kontrol ve hata tespiti için kullanılır. Endüstriyel görüntülerde, bozulmuş veya düşük kaliteli bölgeler restorasyon teknikleriyle düzeltilerek hataların doğru bir şekilde tespit edilmesi sağlanır.
- **Film endüstrisi :** Sinema ve televizyon endüstrisinde eski filmlerin restorasyonu, tarihsel ve kültürel mirasın korunması açısından önemlidir. Aynı zamanda, eski veya bozulmuş filmler dijital ortamda restorasyon teknikleriyle daha izlenebilir hale getirilir. Eski filmler zamanla solmuş renkler ve zayıf ışık koşulları nedeniyle netlik kaybına uğrayabilir. Restorasyon teknikleriyle bu unsurlar düzeltilir.
- **Otomotiv ve ulaşım endüstrisi :** Otomotiv endüstrisinde, araçların kameraları tarafından kaydedilen görüntülerde, yol durumları, kazalar veya diğer güvenlik

olayları kaydedilebilir. Bu görüntülerdeki bozulmalar, restorasyon teknikleriyle düzeltilerek, daha net analizler yapılabilir.

Temel olarak her iki süreçte görüntünün daha iyi bir formuna ulaşmak amaçlanır. Görüntü rekonstrüksiyonu zaten elimizde olmayan görüntüleri, eksik verileri tamamlamaya, görüntüyü orijinal haline en yakın formuna ulaştırmaya çalışırken, görüntü restorasyonu var olan görüntü üzerinde oluşan bozulmaları gidermeye çalışır. Her iki teknik de bazı alanlarda kritik öneme sahiptir ve gelişimi de önem arz etmektedir. Teknolojinin ilerlemesi ile derin öğrenme uygulamaları da bu tekniklerin ilerlemesine büyük katkı sağlamışlardır.

2.2.4 Önemli uygulamalar

Bu noktada görüntü restorasyonunun tanımını ve rekonstrüksiyondan farkını tekrar hatırlamak gerekir. Rekonstrüksiyon, parçalı ya da ham verilerden görüntünün oluşturulması sürecidir. Bu süreç genellikle sensör düzeyinde başlar ve görüntünün temel yapıtaşlarını inşa eder. Restorasyon ise halihazırda var olan fakat bozulmuş bir görüntüyü iyileştirme sürecidir. Bu bozulma gürültü, bulanıklık, eksik veri gibi çeşitli nedenlerden kaynaklanabilir. Dolayısıyla restorasyon, rekonstrüksiyonun bir parçası olabilir ya da ondan tamamen bağımsız bir düzeltme süreci olarak uygulanabilir. Aşağıda restorasyonun uygulandığı kritik alanlar ve bu alanlarda öne çıkan yöntemler açıklanmaktadır.

2.2.4.1 Gürültü giderme

Gürültü giderme, bir görüntüdeki istenmeyen rasgele bileşenlerin (gürültü) azaltılması veya ortadan kaldırılması işlemidir. Bu süreç, gözlemlenen f gürültülü görüntüden orijinal görüntü u 'yu tahmin etme problemidir. Bu bağlamda gözlem modeli genellikle şu şekilde ifade edilir:

$$f(x) = u(x) + n(x) \quad (2.21)$$

burada $n(x)$, sıklıkla sıfır ortalamalı Gauss gürültüsü olarak modellenen rastgele gürültü terimidir. Gürültü giderme probleminin çözümü, gürültü bileşenini baskılamak ve asıl yapısal bilgiyi korumak üzerine kuruludur. Gürültü giderme alanında çeşitli kritik yöntemler aşağıdadır [18].

İstatistiksel ve Bayesçi Yöntemler

Bayesçi çerçevede, orijinal görüntü u bir olasılık dağılımına sahiptir ve amaç, gözlem f verildiğinde bu dağılımın bir fonksiyonel tahminini elde etmektir. Genellikle maksimum artçıl kestirimi (maximum a posteriori, MAP) kullanılır:

$$\hat{u} = \arg \max_u p(u|f) = \arg \max_u p(f|u)p(u) \quad (2.22)$$

Burada $p(f|u)$ gözlem modeli, $p(u)$ ise öncül modeldir. Gauss gürültü varsayımı altında $p(f|u) \propto \exp\left(-\frac{\|f-u\|^2}{2\sigma^2}\right)$ olur. Bu durumda, MAP çözümü şu varyasyonel formülasyona indirgenir:

$$\hat{u} = \arg \min_u \left\{ \frac{1}{2} \|f - u\|^2 + \lambda R(u) \right\} \quad (2.23)$$

burada $R(u)$, genellikle görüntünün düzgünlüğünü kontrol eden bir regülarizasyon terimidir. λ , veri ile öncül arasındaki dengeyi kontrol eder.

Varyasyonel yöntemler

Toplam varyasyon, özellikle kenar koruyucu gürültü giderme uygulamalarında etkilidir. Rudin-Osher-Fatemi (ROF) modeli bu bağlamda klasikleşmiş bir örnektir [19]:

$$\hat{u} = \arg \min_u \left\{ \frac{1}{2} \|f - u\|_2^2 + \lambda \int |\nabla u| dx \right\} \quad (2.24)$$

Bu modelde $\int |\nabla u|$ terimi, görüntü gradyanının L_1 normudur ve kenarların korunmasına olanak tanır. Bu fonksiyonel, Euler-Lagrange denklemleri veya ilkel ikili algoritmalar ile çözülür.

Anisotropik difüzyon

Anisotropik difüzyon, görüntüdeki yapısal özellikleri (özellikle kenarları) koruyarak gürültüyü gidermeyi amaçlayan, yön bağımlı bir difüzyon sürecidir. Perona ve Malik tarafından önerilen bu yöntem, klasik ısı denkleminin modifiye edilmiş halidir [20]:

$$\frac{\partial u}{\partial t} = \nabla \cdot (c(x,y,t)\nabla u) \quad (2.25)$$

Burada $u(x,y,t)$, t zamanında görüntü değerini; $c(x,y,t)$ ise yerel görüntü özelliklerine bağlı olarak değişen difüzyon katsayısını temsil eder. Bu katsayı, genellikle görüntüdeki kenar bölgelerinde difüzyonu azaltacak şekilde tanımlanır.

Perona ve Malik, difüzyon katsayısını görüntü gradyanına bağlı olarak iki farklı şekilde önermiştir:

$$c(|\nabla u|) = \exp\left(-\left(\frac{|\nabla u|}{K}\right)^2\right) \quad \text{veya} \quad c(|\nabla u|) = \frac{1}{1 + \left(\frac{|\nabla u|}{K}\right)^2} \quad (2.26)$$

Burada K parametresi, kenar ile gürültü arasındaki ayrımı kontrol eder. Küçük gradyan büyüklüklerinde c yüksek kalır (difüzyon serbest), ancak büyük gradyanlarda (kenarlarda) c düşer ve difüzyon sınırlandırılır.

Anisotropik difüzyon, özellikle tıbbi görüntüleme ve mikroskopi gibi kenar bilgisi kritik olan alanlarda sıklıkla tercih edilen güçlü bir yöntemdir.

Lineer olmayan ve veri tabanlı Yöntemler

Yerel olmayan ortalamalar (non local means, NLM) yöntemi, piksellerin lokal komşuluğu yerine tüm görüntüdeki benzer alanları karşılaştırarak ağırlıklı ortalama alır. Ağırlıklar, alanlar arasındaki öklidyen mesafeye göre belirlenir:

$$w(x,y) = \exp\left(-\frac{\|P_x - P_y\|^2}{h^2}\right) \Rightarrow \hat{u}(x) = \sum_{y \in \Omega} w(x,y)f(y) \quad (2.27)$$

Blok eşleştirme ve 3 boyutlu filtreleme (block matching and 3 dimensional filtering, BM3D) yöntemi ise benzer blokların bir araya getirilerek 3 boyutlu uzayda işlenmesi esasına dayanır [21]. İki aşamalı çalışır:

- Sert eşikleme ile temel tahmin
- Wiener filtreleme ile iyileştirme

Bu yöntem, doğrusal olmayan fakat yapısal bilgiyi koruyan bir yaklaşımdır.

Derin öğrenme yaklaşımları

Gürültü giderme, artık derin öğrenmenin en başarılı olduğu alanlardan biridir. Özellikle gürültü haritası öğrenerek, ağlar doğrudan $u = f - \mathcal{N}_\theta(f)$ şeklinde orijinal görüntüyü tahmin edebilir. Gürültü giderici konvolüsyonel sinir ağları bu yaklaşımı temsil eder [8]. Eğitim, aşağıdaki gibi MSE kaybı üzerinden yapılır:

$$\mathcal{L}(\theta) = \mathbb{E}_{f,u} [\|\mathcal{N}_\theta(f) - (f - u)\|^2] \quad (2.28)$$

U-Net gibi encoder-decoder tabanlı ağlar, çok ölçekli detay çıkarımı sayesinde hem yapısal hem de düşük frekans bilgilerini başarılı biçimde yakalayabilir.

2.2.4.2 Bulanıklık giderme

Bulanıklık giderme, bir görüntüde bulanıklığa neden olan bozulmayı tersine çevirme işlemidir. Bu tür bozulmalar, kamera hareketi, nesne hareketi veya optik sistemin sınırlamaları gibi nedenlerle oluşur. Matematiksel olarak bulanık bir görüntü f , orijinal görüntü u 'nun bulanıklık çekirdeği k ile konvolüsyonu sonucunda elde edilir [22]:

$$f(x) = (k * u)(x) + n(x) \quad (2.29)$$

Burada $*$, konvolüsyon işlemini; $n(x)$ ise genellikle sıfır ortalamalı Gauss gürültüsünü temsil eder. Amaç, f gözleminden u ve kimi zaman bilinmeyen k 'yı geri kazanmaktır.

İstatistiksel yöntemler: MAP ve olasılık formülasyonu

Bayesçi yaklaşımla, orijinal görüntü u ve bulanıklık çekirdeği k için $p(u|f)$ veya $p(u,k|f)$ gibi koşullu olasılık dağılımları maksimize edilerek çözüm elde edilir. Gürültünün Gauss dağılımlı olduğu varsayılırsa, MAP çözümü şu şekildedir:

$$\hat{u} = \arg \max_u p(u|f) = \arg \max_u p(f|u)p(u) \quad (2.30)$$

Bu ifade negatif logaritma alınarak varyasyonel bir enerji fonksiyonuna dönüştürülür:

$$\hat{u} = \arg \min_u \left\{ \frac{1}{2} \|k * u - f\|_2^2 + \lambda R(u) \right\} \quad (2.31)$$

Burada $R(u)$ regülarizasyon terimidir.

Varyasyonel yöntemler

ROF modelinin genişletilmiş hali olan varyasyonel bulanıklık giderme modelinde amaç, veri tutarlılığı ve düzgünlük terimlerinin dengelenmesidir:

$$\hat{u} = \arg \min_u \left\{ \frac{1}{2} \|k * u - f\|_2^2 + \lambda \int |\nabla u| dx \right\} \quad (2.32)$$

Bu model, bulanıklığı tersine çevirmeye çalışırken aynı zamanda kenarların korunmasına olanak tanır. Bu tür fonksiyoneller genellikle Fourier domaininde optimizasyon, çarpanların alternatif yön yöntemi (alternating direction method of multipliers, ADMM), ya da ilkel ikili algoritmalar ile çözülür [23].

Frekans Uzayı Yöntemleri: Wiener Filtreleme

Lineer bulanıklık varsayımı altında, Wiener filtresi frekans uzayında bulanıklığı tersine çevirmek için kullanılır:

$$\hat{U}(\omega) = \frac{K^*(\omega)}{|K(\omega)|^2 + \frac{\sigma^2}{S_u(\omega)}} F(\omega) \quad (2.33)$$

Burada $K(\omega)$, bulanıklık çekirdeğinin Fourier dönüşümünü; $S_u(\omega)$, orijinal görüntünün güç spektral yoğunluğunu; σ^2 , gürültü varyansını temsil eder.

Derin öğrenme tabanlı yöntemler

Son yıllarda, derin konvolüsyonel sinir ağları (CNN) ve transformer mimarileri bulanıklık giderme problemlerinde çığır açmıştır. Özellikle uçtan uca (end-to-end) modeller, doğrudan $f \rightarrow \hat{u}$ haritalamasını öğrenir. Eğitim süreci aşağıdaki gibi bir regresyon kaybı üzerinden tanımlanır:

$$\mathcal{L}(\theta) = \mathbb{E}_{(f,u)} [\|\mathcal{N}_\theta(f) - u\|^2] \quad (2.34)$$

Bulanıklık giderici çekişmeli üretici ağ (deblurring generative adversarial network, DeblurGAN) ve çoklu bölge algılama sinir ağı (multiple region perception neural network, MPRNet) gibi mimariler çok ölçekli bilgi çıkarımı ve genişletilmiş konvolüsyon gibi tekniklerle hem hareketli bulanıklıkları hem de gauss bulanıklıklarını başarılı şekilde çözebilir [24]–[26].

2.2.4.3 Mozaik çözme

Mozaik çözme, dijital kameraların tek bir renkli sensör ile tam renkli görüntü elde etmesini sağlayan Bayer desenli görüntülerden eksik renk bileşenlerini geri kazanma işlemidir. Çoğu dijital kamerada görüntü, bir Bayer filtresi aracılığıyla piksellere sadece tek bir renk (kırmızı, yeşil veya mavi) gelecek şekilde alınır. Bu mozaik yapıdan tam renkli görüntü üretmek için her pikseldeki eksik iki renk kanalı tahmin edilmelidir [27].

Bayer desenli gözlem f , ideal renkli görüntü $u = (u_R, u_G, u_B)$ 'nin bir alt örneklenmiş hali olarak modellenir:

$$f(x) = M(x) \odot u(x) \quad (2.35)$$

Burada $M(x)$, piksel konumuna bağlı renk filtresini belirten maske; \odot , eleman bazlı çarpma (Hadamard çarpımı) işlemidir. Amaç, f 'ten u 'yu geri kazanmaktır.

Doğrusal enterpolasyon yöntemleri

İlk yöntemlerden biri olan bilineer enterpolasyon, her pikselin eksik renk bileşenlerini çevresindeki bilinen piksellerin ortalamasıyla tahmin eder. Bu yaklaşım hızlıdır fakat kenarlarda renk geçişlerinde artefaktlar oluşturabilir.

Varyasyonel yöntemler

Daha gelişmiş yöntemler, toplam varyasyon gibi düzgünlük terimlerini kullanarak daha keskin kenarlar üretir. Bu yöntemlerde genellikle bir enerji fonksiyonu minimize edilir:

$$\hat{u} = \arg \min_u \left\{ \frac{1}{2} \|M \odot u - f\|_2^2 + \lambda \sum_{c \in \{R,G,B\}} \int |\nabla u_c(x)| dx \right\} \quad (2.36)$$

Burada ∇u_c her renk kanalının gradyanını ifade eder. Bu model, renk kanalları arasındaki düzgünlüğü korurken veriye sadakati de sağlar.

Ortak gradyan ve kenar yönelimli modeller

Bazı yöntemler renk kanalları arasında gradyan benzerliğinden faydalanarak kenar yönelimli enterpolasyon yapar. Örneğin, yeşil kanalın kenar yönü tahmin edilip kırmızı ve mavi kanallarda bu yön boyunca enterpolasyon yapılabilir.

İstatistiksel yaklaşım: MAP tabanlı mozaik çözme

Bayesçi yaklaşımda, MAP tahmini ile mozaik çözme yapılır. Bu durumda:

$$\hat{u} = \arg \max_u p(u|f) = \arg \max_u p(f|u) \cdot p(u) \quad (2.37)$$

Bu yaklaşımda $p(u)$ öncel bilgiyi (örneğin doğal görüntü istatistiklerini), $p(f|u)$ ise gözlem modelini temsil eder. Genellikle bu ifadeler logaritmik formda optimizasyon problemlerine dönüştürülür.

Derin öğrenme tabanlı yöntemler

Yeni nesil mozaik çözme yöntemleri, CNN veya dönüştürücü temelli ağlar ile uçtan uca öğrenilir. Model, Bayer desenli görüntüyü doğrudan tam renkli bir görüntüye dönüştürür:

$$\mathcal{L}(\theta) = \mathbb{E}_{(f,u)} [\|\mathcal{N}_\theta(f) - u\|^2] \quad (2.38)$$

Burada \mathcal{N}_θ , öğrenilen mozaik çözme modelini; θ ise parametrelerini ifade eder. Özellikle DemosaicNet, U-Net tabanlı yapılar ve öz kontrol mekanizmaları bu alanda yüksek doğruluklar sağlamaktadır.

2.2.4.4 Süper çözünürlük

Süper çözünürlük, düşük çözünürlüklü bir görüntüden yüksek çözünürlüklü bir görüntünün elde edilmesini amaçlayan bir görüntü iyileştirme problemidir. Bu problem hem tekli görüntüler için hem de video gibi çoklu kareler için ele alınabilir.

Genel süper çözünürlük modeli, yüksek çözünürlüklü görüntü x 'in bir bozunum işleminden geçerek düşük çözünürlüklü gözlem y 'ye dönüştürülmesi olarak tanımlanır [17,28]:

$$y = D(B(x)) + n \quad (2.39)$$

Burada:

- x : Gerçek yüksek çözünürlüklü görüntü,
- $B(\cdot)$: Bulanıklık operatörü,
- $D(\cdot)$: Alt örnekleme işlemi,
- n : Gözlem gürültüsünü temsil eder.

Amaç, sadece y verilmişken x 'i yeniden tahmin etmektir.

Klasik yöntemler

Klasik yöntemler çoğunlukla interpolasyon tekniklerine dayanır. En yaygın kullanılan yöntemler bikübik ve bilinear interpolasyondur. Ancak bu yöntemler sıklıkla keskinlik kaybına ve artefaktlara yol açar.

Ters problem yaklaşımı

Süper çözünürlük, matematiksel olarak iyi tanımlanmamış bir ters problemdir. Regularizasyonlu enerji minimizasyonu yaklaşımı sıkça kullanılır:

$$\hat{x} = \arg \min_x \{ \|D(B(x)) - y\|_2^2 + \lambda \mathcal{R}(x) \} \quad (2.40)$$

Burada $\mathcal{R}(x)$, görüntünün düzgünlüğü veya doğal istatistiklerine dayalı bir regularizasyon terimini temsil eder.

Bayesçi yaklaşımlar

Bayes kuramına göre MAP tahmini şu şekilde ifade edilir:

$$\hat{x} = \arg \max_x p(x|y) = \arg \max_x p(y|x) \cdot p(x) \quad (2.41)$$

Bu formülasyon altında $p(y|x)$ gözlem modeli, $p(x)$ ise yüksek çözünürlüklü görüntülerin önsel dağılımını temsil eder.

Derin öğrenme tabanlı yöntemler

Derin öğrenme tabanlı süper çözünürlük, düşük çözünürlüklü girişten yüksek çözünürlüklü çıkışı doğrudan öğrenen modelleri kapsar. En yaygın model ailesi konvolüsyonel sinir ağlarıdır. Örnek bir L1 kaybı ile eğitilen modelde optimizasyon şu şekilde yapılır:

$$\mathcal{L}(\theta) = \mathbb{E}_{(y,x)} [\| \mathcal{F}_\theta(y) - x \|_1] \quad (2.42)$$

Burada \mathcal{F}_θ , öğrenilen süper çözünürlük modelini ve θ ise parametrelerini ifade eder.

2.2.4.5 Görüntü tamamlama

Görüntü tamamlama diğer adıyla iç boyama eksik, bozulmuş veya silinmiş bölgelere sahip görüntülerde bu bölgelerin çevresel içerik kullanılarak yeniden tahmin edilmesi problemidir. Bu teknik, eski fotoğrafların restorasyonu, obje çıkarma, metin silme ve hata düzeltme gibi pek çok uygulamada kullanılır.

Görüntü tamamlama problemi, eksik piksellerin yer aldığı bir maske fonksiyonu M ile birlikte aşağıdaki şekilde formüle edilir [29,30]:

$$y = M \odot x + (1 - M) \odot n \quad (2.43)$$

Burada:

- x : Gerçek (tam) görüntü,
- y : Gözlemlenen eksik (maskelemiş) görüntü,
- M : Aynı boyutta ikili maske, $M(i, j) = 1$ ise piksel gözlemlenmiştir, 0 ise eksiktir,
- n : Gürültü veya rastgele başlangıç değeri,
- \odot : Eleman bazlı çarpma operatörü (Hadamard çarpımı).

Amaç, M ve y verildiğinde, eksik bölgeleri en iyi şekilde tahmin eden bir \hat{x} görüntüsü üretmektir.

Enerji tabanlı yaklaşımlar

Klasik yöntemlerde, görüntünün eksik kısımları düzgünlük varsayımı altında doldurulur. Bu, bir enerji fonksiyonu ile aşağıdaki şekilde modellenenir:

$$\hat{x} = \arg \min_x \{ \|M \odot (x - y)\|_2^2 + \lambda \mathcal{R}(x) \} \quad (2.44)$$

Burada $\mathcal{R}(x)$, görüntü üzerindeki regularizasyon terimidir.

İstatistiksel ve Bayesçi yaklaşım

Bayesçi çerçevede, tam görüntünün MAP tahmini aşağıdaki gibi hesaplanabilir:

$$\hat{x} = \arg \max_x p(x|y, M) = \arg \max_x p(y|x, M) \cdot p(x) \quad (2.45)$$

Bu formülasyonda $p(x)$ görüntü önbilgisi (prior), $p(y|x, M)$ ise gözlem modeli olarak tanımlanır.

Örüntü tabanlı kopyalama yöntemleri

Özellikle doğal görüntülerde, eksik bölgeleri doldurmak için çevredeki benzer örüntülerin kopyalanmasına dayalı teknikler kullanılır. Criminisi ve Perez gibi araştırmacıların çalışmaları bu alanda öncül yöntemlerdir ve eksik bölgelere doğru izotropik şekilde örüntü genişletilmesine dayanır. [30,31].

Derin öğrenme ile görüntü tamamlama

Derin öğrenme yaklaşımlarında, eksik bölgelerin global bağlamdan öğrenilerek tamamlanması hedeflenir. Bu tür modellerde genellikle U-Net veya kodlayıcı-kod çözücü mimarileri tercih edilir.

Kaybın tanımı çoğunlukla aşağıdaki gibi olur:

$$\mathcal{L}(\theta) = \mathbb{E}_{(x,M)} [\| (1-M) \odot (\mathcal{F}_\theta(M \odot x) - x) \|_1] \quad (2.46)$$

Burada \mathcal{F}_θ maske uygulanmış görüntüden eksik bölgeleri dolduran öğrenilmiş bir modeli ifade eder.

2.2.4.6 Metal artefakt azaltma

Metal artefaktları, tıbbi görüntüleme sistemlerinde, özellikle bilgisayarlı tomografi (CT) gibi X-ışını tabanlı sistemlerde metal implantlar veya protezler gibi objelerin görüntülerinize neden olduğu bozulmalardır. Bu artefaktlar, görüntü kalitesini ciddi şekilde bozarak tanıların doğruluğunu düşürebilir. Metal artefakt azaltma (MAR) teknikleri, bu tür bozulmaların ortadan kaldırılmasına yardımcı olur.

Metal objeler, X-ışınlarını güçlü bir şekilde yansıttığı ve emdiği için, görüntüleme cihazı tarafından tam olarak algılanamaz. Bu durum, metalin etrafında ışınların sapmasına ve dağılımına neden olarak, çevresindeki bölgelerde yüksek kontrastlı, bozulmuş ve hatalı görüntüler oluşturur. Metal artefaktlar genellikle yoğun alanlar ve çizgisel oluşumlar olarak görülür [32,33].

Matematiksel modelleme

Metal artefaktlarını matematiksel olarak modellemek, X-ışını tomografisi sırasında meydana gelen ışınların metal obje etrafında nasıl değiştiğini anlamak için önemlidir. Metalin bulunduğu bölgede, ışınlar yüksek miktarda soğurulmakta ve bu durum klasik görüntüleme modelinde önemli sapmalara yol açmaktadır.

Bununla birlikte, metal artefaktları için modelleme, özellikle aşağıdaki denklemlerle ifade edilebilir:

$$I(x) = \int_{-\infty}^{\infty} T(x')e^{-\mu(x,x')\Delta x} dx' \quad (2.47)$$

Burada:

- $I(x)$: X-ışını yoğunluğunu,
- $T(x')$: Doku yoğunluğunu,
- $\mu(x,x')$: X-ışını soğurma katsayısını (metal nesne nedeniyle artar),
- Δx : Işın yolu uzunluğunu temsil eder.

Metal artefaktların azaltılması için amaç, $\mu(x,x')$ fonksiyonunu metalin etrafındaki bölgeyi doğru şekilde modelleyebilecek şekilde iyileştirmektir.

Metal artefakt azaltma yöntemleri

MAR teknikleri, genellikle görüntü iyileştirme ve yeniden yapılanma adımlarını içerir. Aşağıdaki yöntemler bu sorunu çözmeye yaygın olarak kullanılır:

- **Ön işlem:** Metal artefaktları, görüntülerin işlenmeden önce maskelenmesi veya önceden düzeltilmesi yoluyla azaltılabilir. Genellikle metal bölgelerdeki X-ışını yoğunlukları interpolasyon ile düzeltilir.
- **Faz değiştirme:** Metalin neden olduğu faz değişimlerini düzeltmek için özel algoritmalar kullanılır. Bu, özellikle metalin çevresindeki dokularda daha net bilgi elde edilmesini sağlar.

- **İleriye dönük modelleme ve interpolasyon:** Metalin çevresindeki dokuların doğru bir şekilde tahmin edilmesi için ileriye dönük modelleme teknikleri ve interpolasyon yöntemleri kullanılır. Bu yöntemlerde genellikle komşu piksel bilgilerinden yararlanır.
- **Görüntü yeniden yapılandırma:** Metal artefaktları görüntüleme sırasındaki yeniden yapılanma adımlarında ortadan kaldırmak için, özel rekonstruksiyon algoritmaları uygulanır. Bu algoritmalar, metalin etrafındaki pikselleri yeniden hesaplar.
- **Genetik algoritmalar ve derin öğrenme yöntemleri:** Son yıllarda, metal artefaktlarının azaltılmasında genetik algoritmalar ve derin öğrenme modelleri kullanılmaya başlanmıştır. Bu modeller, bozulmuş görüntülerden temiz veriler elde etmek için eğitilmiş ağlar kullanır.



3. GÖRÜNTÜ İŞLEME ALANINDA DERİN ÖĞRENME

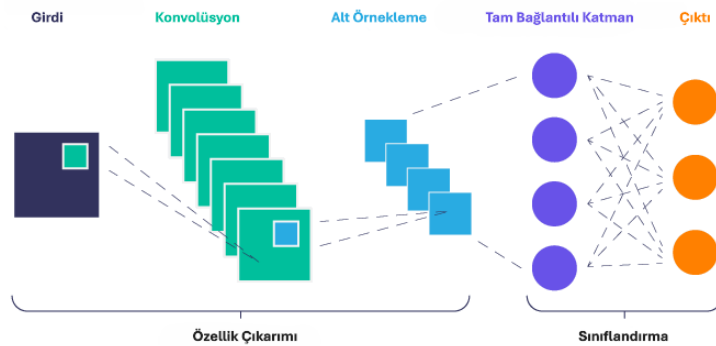
3.1 Derin Öğrenme ve Yapay Sinir Ağlarına Genel Bakış

Derin öğrenme, karmaşık görevleri çözmek için yapay sinir ağlarını eğitmeye odaklanan yapay zekanın bir alt kümesidir. Bu sinir ağları, beyin hücrelerinin yapısını ve işlevini taklit eden, birbirine bağlı birden fazla katmanından oluşur. Her katman, giriş verilerinden belirli özellikleri çıkararak derin öğrenme modellerinin hiyerarşik gösterimleri öğrenmesine olanak tanır.

Derin öğrenmedeki temel mimariler şunları içerir:

- **Konvolüsyonel sinir ağları (CNN)**, görüntü gibi ızgara formatında olan verileri işlemek üzere tasarlanmış özel bir sinir ağı sınıfıdır. Konvolüsyonel sinir ağları kenarlar, dokular ve daha karmaşık yapılar gibi desenlerin etkili bir şekilde öğrenilmesini sağlayan özelliklerin hiyerarşilerini çıkarmak için konvolüsyon katmanlarından yararlanır. Bu mimari, özellikle görüntü tanıma, nesne algılama ve diğer bilgisayarlı görüş işlemleri için uygundur. CNN'ler, LeCun ve arkadaşlarının el yazısıyla yazılmış rakam tanıma konusundaki öncü çalışmalarıyla başlayarak derin öğrenme alanının ilerlemesinde temel olmuştur [34].

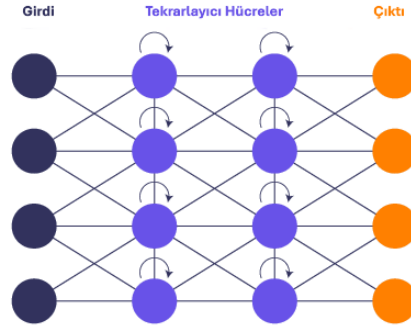
Konvolüsyonel Sinir Ağları (CNN)



Şekil 3.1: Konvolüsyonel sinir ağları mimarisi.

- **Tekrarlayan sinir ağıları (recurrent neural networks, RNN)**, önceki zaman adımlarından bilgi yakalayan durumu koruyarak ardışık verileri işlemek üzere tasarlanmış bir yapay sinir ağı sınıfıdır. Bu özellik onları zaman serileri, doğal dil işleme ve diğer ardışık desenleri içeren görevler için etkili hale getirir. İleri beslemeli ağların aksine, RNN'ler değişken uzunluktaki girdileri işleyebilir ve bu da onları konuşma tanıma, dil modelleme ve dizi tahmini gibi görevler için uygun hale getirir [35].

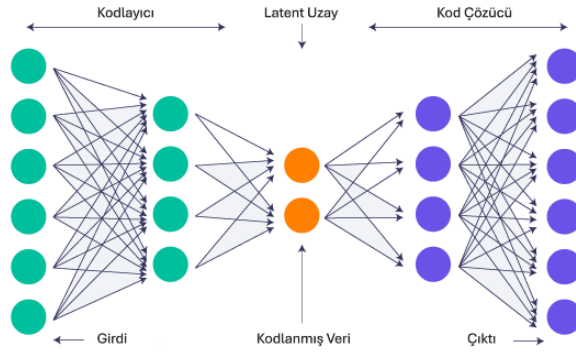
Tekrarlayan Sinir Ağları (RNN)



Şekil 3.2: Tekrarlayan sinir ağıları mimarisi.

- **Otokodlayıcılar**, giriş verilerini bir alana sıkıştırarak ve orijinal biçimine geri yeniden oluşturarak verilerin verimli temsillerini öğrenmek üzere tasarlanmış bir tür gözetimsiz sinir ağı mimarisidir. Girişi sıkıştırılmış bir temsile eşleyen bir kodlayıcıdan ve girişi bu temsilden yeniden oluşturan bir kod çözücünden oluşurlar. Bu modeller boyut azaltma, özellik çıkarma ve veri gürültüsünü giderme gibi görevler için yaygın olarak kullanılır [36].

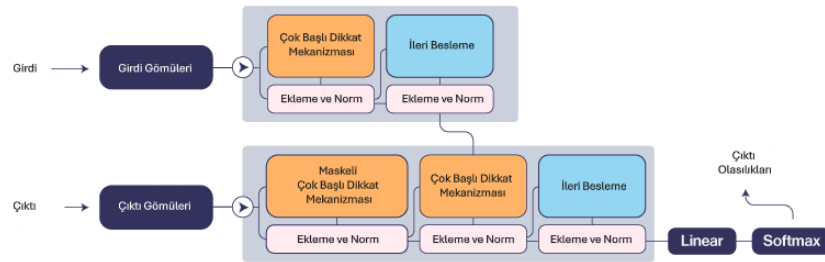
Otokodlayıcılar



Şekil 3.3: Otokodlayıcı mimarisi.

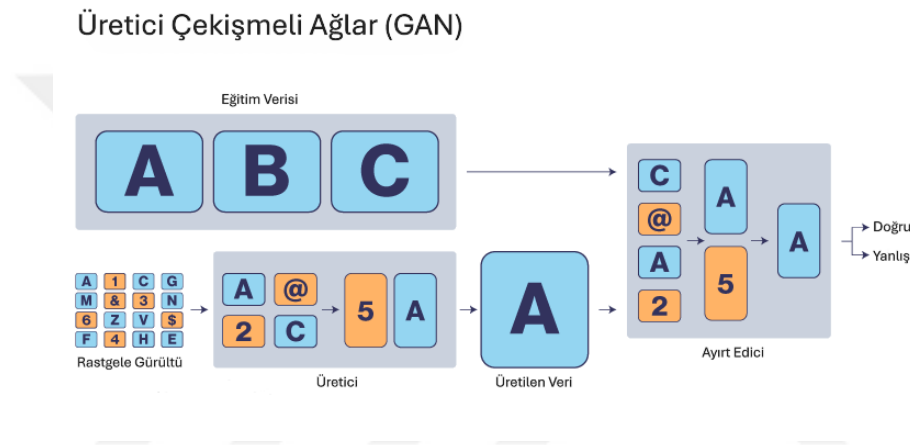
- **Transformatörler**, öz bakım mekanizmalarından yararlanarak sıralı verileri işlemek üzere tasarlanmış güçlü bir sinir ağı mimarisidir. Geleneksel tekrarlayan veya evrişimli mimarilerin aksine, transformatörler tüm dizileri paralel olarak işler ve hem kısa hem de uzun menzilli bağımlılıkları etkili bir şekilde yakalar. Bu yenilik, özellikle çift yönlü kodlayıcı tabanlı dönüştürücü (bidirectional encoder representations from transformers, BERT) ve önceden eğitilmiş üretici dönüştürücü (generative pre-trained transformer, GPT) gibi modellerle doğal dil işlemede çığır açıcı gelişmelere yol açmış, bilgisayar görüşü ve zaman serisi analizi gibi alanlara da genişletilmiştir. Mimari ilk olarak Vaswani'nin "Attention Is All You Need" adlı çığır açıcı çalışmasında tanıtılmıştır [37].

Transformatörler



Şekil 3.4: Transformatör mimarisi

- **Çekişmeli üretici ağlar (GAN)**, ilk olarak Goodfellow tarafından tanıtılan bir üretken model sınıfıdır. Temel olarak birbiriyle rekabet halinde bulunan iki farklı yapay sinir ağı içermektedir. Üretici ağı, gerçek veri dağılımını taklit etmeyi amaçlayan veri örnekleri oluştururken, ayırıcı ağı gerçek ve üretilen örnekler arasında ayırım yapmaya çalışır. Bu düşmanca süreç, her iki ağı da yinelemeli olarak iyileştirmeye yönlendirir ve üreticinin yüksek kaliteli sentetik veri üretmesini sağlar. GAN'lar, gerçekçi ve yüksek çözünürlüklü çıktılar üretme yeteneklerini göstererek görüntü sentezinde, stil transferinde ve veri arttırmada yaygın olarak kullanılmışlardır. [38].



Şekil 3.5: GAN mimarisi

Derin öğrenme, büyük veri ve yüksek hesaplama gücündeki ilerlemeler sayesinde birçok alanda devrim niteliğinde çözümler sunmaktadır ve giderek daha önemli hale gelmektedir. Artan hesaplama gücü ve büyük veri setleri ile birlikte, bu yöntemler birçok alanda insan performansını aşan çözümler sunmaktadır.

3.2 Görüntü İşleme Alanındaki Derin Öğrenme Uygulamaları

Derin öğrenme, görüntü sınıflandırma, nesne algılama, segmentasyon, iyileştirme ve restorasyon gibi karmaşık görevlere son derece etkili çözümler sunarak görüntü işleme alanında devrim yaratmıştır. Bu dönüşüm büyük ölçüde, geleneksel görüntü işleme tekniklerinden önemli ölçüde daha iyi performans gösteren konvolüsyonel sinir ağları, çekişmeli üretici ağlar, otokodlayıcılar ve son dönemlerde özellikle transformatörler gibi mimarilerdeki gelişmeler sonucunda gerçekleşmiştir.

- **Görüntü sınıflandırması ve nesne algılama** Konvolüsyonel sinir ağları, görüntü sınıflandırma ve nesne algılama uygulamalarında olağanüstü sonuçlara ulaşmada önemli rol oynamıştır. Konvolüsyonel sinir ağları, manuel özellik mühendisliğine gerek kalmadan kenarlar, dokular ve nesne parçaları gibi özelliklerin mekansal hiyerarşilerini otomatik ve uyarlanabilir bir şekilde öğrenmek üzere tasarlanmıştır. Bu yetenek, yüz tanıma, sahne sınıflandırması ve daha karmaşık görsel tanıma görevleri gibi görevlerde üstünlük sağlamasını olanak tanır [34]. Örneğin, bir CNN mimarisi olan AlexNet, 2012 ImageNet yarışmasını kazandı ve büyük ölçekli görüntü sınıflandırmasında derin öğrenmenin potansiyelini gösterdi [39].

Nesneleri yalnızca sınıflandırmakla kalmayıp aynı zamanda bir görüntü içinde yerleştirmeyi de amaçlayan daha zorlu bir uygulama olan nesne algılama ve daha hızlı bölge tabanlı evrişimli sinir ağı [40] gibi mimariler aracılığıyla ileri seviyelere taşındı. Bu mimariler, sınırlayıcı kutuları ve sınıf etiketlerini tek bir geçişte tahmin etmek için konvolüsyonel sinir ağlarını kullanır ve nesne algılamayı daha verimli ve ölçeklenebilir hale getirir.

- **Görüntü bölümlendirmesi** Bir görüntüyü anlamlı bölgelere ayırmayı, örneğin ön planı arka plandan ayırmayı, içeren görüntü bölümlendirmesinde, derin öğrenme sayesinde önemli gelişmeler yaşandı. Tam konvolüsyonel ağlar, U-Net ve maskeli bölge tabanlı evrişimli sinir ağı segmentasyon görevlerinde kullanılan popüler mimarilerdir. Bu modeller, tümörler, organlar ve lezyonlar gibi yapıların kesin olarak belirlenmesinin tanı ve tedavi planlaması için çok önemli olduğu tıbbi görüntülemeyle özellikle yararlıdır.

Örneğin, U-Net mimarisi, nispeten küçük miktarda etiketli veriden yüksek çözünürlüklü çıktı haritaları üretmesine olanak tanıyan evrişimli ve dekonvolüsyonlu katmanların verimli kullanımı nedeniyle tıbbi görüntü segmentasyonunda yaygın olarak benimsenmiştir [11].

- **Görüntü geri yükleme ve geliştirme** Derin öğrenme modelleri, süper çözünürlük, gürültü giderme ve bulanıklık giderme gibi görüntü geri yükleme uygulamalarında başarıyla uygulanmıştır. Örneğin, çekişmeli üretici ağlar (GAN), düşük çözünürlüklü girdilerden yüksek çözünürlüklü görüntüler üretmek için yaygın

olarak kullanılır. Süper çözünürlüklü üretici çekişmeli ağlar (super resolution generative adversarial network, SRGAN) gibi GAN tabanlı modeller, genellikle gerçek yüksek çözünürlüklü görüntülerden ayırt edilemeyen keskin ve yüksek kaliteli sonuçlar üretir [13].

Tıbbi görüntüleme, MRI ve bilgisayarlı tomografi taramaları gibi uygulamalarda kritik öneme sahip olan gürültülü veya bozuk görüntülerin kalitesini artırmak için derin öğrenme yöntemleri kullanılır. Gürültü giderici otokodlayıcılar (denoising autoencoder, DAE) ve çekişmeli üretici ağlar, görüntülerin önemli yapısal ayrıntılarını korurken gürültüyü gidermek için kullanılır [36,38]. Bu teknikler, tıbbi görüntülerin kullanılabilirliğini ve yorumlanabilirliğini önemli ölçüde iyileştirerek bunları tanı için daha güvenilir hale getirmiştir.

- **Veri artırma** Çekişmeli üretici ağlar, sanattan tıbbi alanlara kadar çok çeşitli uygulamalarda kullanılabilen gerçekçi sentetik görüntülerin oluşturulmasını sağlayarak görüntü oluşturma uygulamalarında ciddi bir etkiye sahip olmuştur. GAN'lar, özellikle verilerin yetersiz olduğu durumlarda eğitim veri kümelerini yapay olarak genişletmek için kullanılır.

Örneğin tıbbi görüntüleme eğitim veri kümelerinin çeşitliliğini ve boyutunu iyileştirmek için sentetik CT taramaları ve MRI verileri oluşturmak için kullanılmıştır ve böylece bu veri kümeleri üzerinde eğitilen modellerin performansı artırılmıştır [41]. Bu yaklaşım, özellikle uzmanlaşmış tıbbi alanlarda veri kıtlığı sorunlarının üstesinden gelinmesine yardımcı olur.

Görüldüğü üzere, görüntü işleme alanındaki yapay zeka uygulamaları büyük başarılarla imza atmışlardır. Konvolüsyonel sinir ağları, çekişmeli üretici ağlar, oto kodlayıcılar gibi gelişmiş derin öğrenme mimarileri insan sağlığını etkileyen uygulamalar dahil bir çok görüntü işleme uygulamasında çığır açmıştır. Derin öğrenme modellerinin çok yönlülüğü, giderek daha karmaşık ve yüksek boyutlu verileri işlemelerine olanak tanımakta ve bu alanda uzun süredir devam eden zorluklara çözümler sunmaktadır.

3.3 Mevcut Yöntemlerin Sınırlamaları

Derin öğrenme modelleri çeşitli alanlarda önemli ilerlemeler kaydetmiş olsa da, eksiklikleri yok değildir. Bu eksiklikler, özellikle sağlık ve otonom sistemler gibi yüksek etkili alanlarda yaygın olarak benimsenmelerini engeller. Aşağıda mevcut derin öğrenme yöntemlerinin temel sınırlamaları verilmiştir:

- **Veri bağımlılığı:** Derin öğrenme modellerinin karşılaştığı en önemli zorluklardan biri, büyük, yüksek kaliteli veri kümelerine aşırı bağımlı olmalarıdır. Etkili modelleri eğitmek genellikle çok miktarda etiketli veri gerektirir ve bu, tıbbi görüntüleme, uzaktan algılama veya otonom sürüş gibi yüksek kaliteli veri isteyen alanlarda elde edilmesi zor bir durumdur. Verileri etiketleme süreci hem pahalı hem de zaman alıcıdır ve alan uzmanlığı gerektirir. Örneğin, tıbbi görüntülemede, nadir görülen durumlar veya hastalıklar için açıklamalı veri elde etmek genellikle bir darboğazdır ve veri bulmak zordur. Büyük veri kümelerine olan bu bağımlılık, bu sorunu bir dereceye kadar hafifletmeye yardımcı olan transfer öğrenimi ve veri artırma gibi yöntemlerin geliştirilmesinin temel nedeni olmuştur [42]. Ancak, bu yaklaşımların etkili olması için hala önemli miktarda temel veriye ihtiyaç vardır.
- **Yorumlanabilirlik:** Derin öğrenme modellerinin bir diğer kritik sınırlaması da yorumlanabilirlik eksikliğidir. Sinir ağları, özellikle derin modeller, karmaşık mimarileri ve karar alma süreçleri nedeniyle genellikle "kara kutular" olarak tanımlanır. Sağlık hizmeti veya otonom sürüş gibi yüksek riskli uygulamalarda, bir modelin belirli bir karara neden vardığını anlamak çok önemlidir. Model davranışını açıklayamama, özellikle modeller kritik uygulamalarda hata yaptığında etik ve yasal endişelere yol açar. Örneğin, bir derin öğrenme modeli tıbbi bir durumu yanlış teşhis ederse veya otonom bir araçta yanlış bir karar verilirse, açıklanabilirliğin eksikliği, nedeni belirlemeyi ve sorunu düzeltmeyi zorlaştırabilir [43]. Açıklanabilir yapay zeka teknikleri gibi derin öğrenme modellerinin yorumlanabilirliğini artırmak için çeşitli yöntemler önerilmiştir, ancak bu hala aktif bir araştırma alanıdır.
- **Genelleme:** Derin öğrenme modelleri eğitim veri kümelerinde görülmemiş verilere genelleme yapmakta zorlanırlar. Bu genelleme eksikliği, modelin eğitim verilerinde iyi performans göstermesine ancak yeni, görülmemiş verilerde kötü performans

göstermesine yol açan aşırı uyuma neden olabilir. Bu sorun, veri dağılımının zamanla değişebileceği veya modelin çeşitli ve öngörülemeyen ortamlarda performans göstermesinin beklendiği gerçek dünya uygulamalarında özellikle sorun yaratır. [44]. Örneğin, bir popülasyonda akciğer kanserini tespit etmek için eğitilen bir model, görüntüleme protokollerindeki, hasta demografisindeki veya hastalık sunumundaki farklılıklar nedeniyle başka bir popülasyona iyi genelleme yapamayabilir. Bu sorunu ele almak için çapraz doğrulama, düzenleme ve alan uyarlaması gibi teknikler yaygın olarak kullanılır, ancak sağlam genellemeyi sağlamak önemli bir zorluk olmaya devam etmektedir.

- **Hesaplama maliyetleri:** Derin öğrenme modellerini eğitmek önemli hesaplama kaynakları gerektirir. Yüksek hesaplama maliyeti, güçlü donanım, büyük miktarda bellek ve uzun eğitim sürelerine duyulan ihtiyaçtan kaynaklanır. Ek olarak, bu modellerin bakımı da yüksek bütçeli bir altyapı gerektirir. Bulut bilişim platformları bu kaynakları daha erişilebilir hale getirmiş olsa da, engelleyici maliyetlerin varlığı devam etmektedir. Dahası son zamanlarda ilişkili enerji tüketimi nedeniyle büyük modelleri eğitmenin çevresel etkisi, mevcut derin öğrenme uygulamalarının sürdürülebilirliği konusunda endişelere yol açmıştır [45]. Araştırmacılar, daha verimli mimariler geliştirmek veya hesaplama yükünü azaltmak için budama ve niceleme gibi tekniklerden yararlanmak gibi bu sorunları hafifletmenin yollarını araştırmaktadırlar.

Derin öğrenme yöntemlerinin kayda değer başarılarına rağmen, büyük veri kümelerine bağımlılık, yorumlanabilirliğin eksikliği, genelleme zorlukları ve yüksek hesaplama maliyetleri gibi sınırlamaları, özellikle sağlık hizmetleri ve otonom sistemler gibi alanlarda yaygın olarak benimsenmeleri için önemli engeller oluşturmaktadır. Bir yandan yüksek performans gerekirken diğer yandan bu sınırlamaların üstesinden gelebilecek alternatif yaklaşımların araştırılması gerekmektedir. Umut vadeden bir yön, difüzyon ağlarının kullanılmasıdır. Geleneksel derin öğrenme modellerinin aksine, difüzyon ağları gürültüye karşı doğal bir sağlamlık sunar ve küçük veri kümelerine daha fazla uyarlanabilir, bu da onları yukarıda belirtilen zorlukların bazıları için bir çözüm haline getirir.

4. DİFÜZYON AĞLARI

4.1 Difüzyon Ağları Nedir?

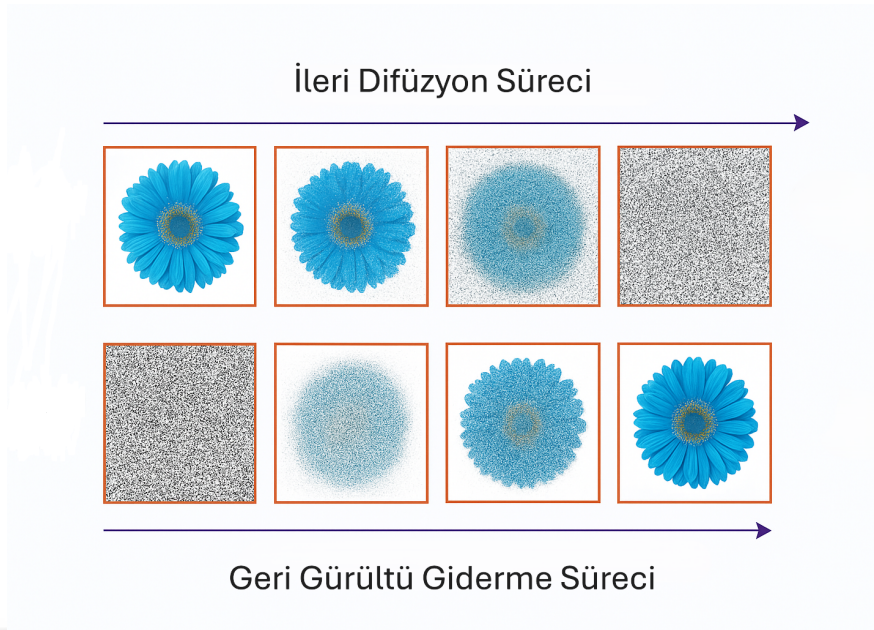
Difüzyon ağları, karmaşık, yüksek boyutlu veri dağılımlarını yaklaşık olarak belirlemek ve örneklemek için tasarlanmış güçlü bir üretken model sınıfıdır. Temelleri, parçacıkların bir ortamda kademeli olarak dağıldığı fizikteki difüzyon süreçlerinden esinlenmiştir. Makine öğrenimi bağlamında, difüzyon ağları rastgele gürültünün tutarlı ve anlamlı verilere dönüşümünü modellemek için adım adım olasılıksal bir çerçeveden yararlanır. Bu sistematik yaklaşım onları çekişmeli üretici ağlar ve varyasyonel otokodlayıcılar gibi diğer üretken modellerden ayırır. Gelişmiş kararlılık, esneklik ve örnek kalitesi sunarlar [46].

Bu modeller, genellikle değişken bir alt sınırdan türetilen bir kayıp fonksiyonunu en aza indirerek temeldeki veri dağılımını yaklaşık olarak belirlemek üzere eğitilir [2]. Difüzyon ağları, yüksek doğrulukta çıktılar üreterek ve geleneksel yöntemlerin sınırlamalarını aşarak görüntü sentezi, süper çözünürlük ve tıbbi görüntüleme dahil olmak üzere çeşitli alanlarda olağanüstü performans göstermiştir [47].

4.2 Temel İlkeler ve Difüzyon Ağlarının Yapısı

Difüzyon ağları, yüksek doğrulukta ve çeşitli çıktılar üretmek için olasılıksal çerçevelerden yararlanarak sıralı difüzyon süreçleri aracılığıyla veri dönüşümünü sistematik olarak modeller ve yüksek boyutlu veri üretimine sağlam bir çözüm sunarlar. Bu bölüm, difüzyon ağlarının temel ilkelerini ve mimari yapısını inceler.

- **İleri ve ters difüzyon süreçleri:** Difüzyon ağlarının arkasındaki temel fikir iki aşamalı bir süreçtir. İleri süreç, verilere birden fazla adımda kademeli olarak Gauss gürültüsü ekler ve orijinal verilerin yapısını etkili bir şekilde yok eder. Ters süreç, gürültülü verileri adım adım gürültüden arındırmayı öğrenerek orijinal veri dağıtımını yeniden oluşturur.



Şekil 4.1: İleri ve geri difüzyon süreçleri.

- **Değişken amaç:** Eğitim, negatif log-olasılık üzerindeki değişken alt sınırını en aza indirmeyi içerir ve ters işlemin orijinal veri dağılımına yakın olmasını sağlar [47].
- **Puan eşleştirme ve gürültü giderme:** Difüzyon ağları, gürültü giderme işlemini yönlendirmek için veri dağılımının eğimlerini tahmin eder; bu genellikle stokastik diferansiyel denklemler aracılığıyla formüle edilir [47].
- **Zaman koşullu ağ:** Genellikle U-Net mimarilerine dayalı bir sinir ağı, gürültülü verileri x_t ve zaman adımını t girdi olarak kullanarak gürültü giderme adımlarını tahmin eder [2].
- **Gürültü planlama:** Gürültü planlama $\{\alpha_t\}_{t=1}^T$ gürültü ekleme sürecini belirler. Kosinüs yumuşatma gibi stratejiler eğitim kararlılığını ve örnek kalitesini iyileştirir [48].
- **Stokastik diferansiyel denklemler (stochastic differential equations, SDE) ile entegrasyon:** Stokastik diferansiyel denklemler aracılığıyla sürekli zamanlı formülasyonlar ile çeşitli üretken modelleri birleştirir ve verimli örnekleme çerçeveleri sağlar [47].

Genel olarak difüzyon ağları gürültüye karşı sağlamlık sağlar, eğitim sorunlarından kaçınır. Görüntü sentezi, süper çözünürlük ve tıbbi görüntüleme gibi yüksek boyutlu veri oluşturma uygulamalarında mükemmellik gösterir.

4.3 Matematiksel Arka Plan ve Modelleme

Difüzyon ağları, olasılık teorisi, diferansiyel denklemler ve optimizasyon kavramlarını birleştiren güçlü bir matematiksel temel üzerine inşa edilmiştir. Bu modeller, karmaşık veri dağılımlarına etkili bir şekilde yakınsayabilmek için iki adımlı bir işlem olan ileri yayılım ve ters gürültü gidermeyi kullanır. Bu bölüm, matematiksel prensiplerinin ve modelleme çerçevelerinin ayrıntılı bir açıklamasını sağlar.

- **İleri yayılma işlemi:** İleri yayılma işlemi, verilere kademeli olarak Gauss gürültüsü ekler ve bunları kademeli olarak saf gürültüye dönüştürür. Matematiksel olarak, bu işlem bir Markov zinciri olarak tanımlanabilir:

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{\alpha_t}x_{t-1}, (1 - \alpha_t)I) \quad (4.1)$$

burada x_0 orijinal veriyi, x_t ise t zaman adımındaki bozulmuş veriyi ve (α_t) gürültü ölçekleme faktörünü temsil eder. Toplam gürültü şu şekilde ifade edilebilir:

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}x_0, (1 - \bar{\alpha}_t)I) \quad (4.2)$$

Bu formülasyon, yapılandırılmış verilerden gürültüye yumuşak bir geçiş sağlar [2,46].

- **Ters gürültü giderme işlemi:** Ters gürültü giderme işlemi, saf gürültünün veri dağılımına geri dönüşümünü modeller. Bu işlem şu şekilde parametrelendirilir:

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)) \quad (4.3)$$

burada μ_θ ve Σ_θ öğrenilmiş parametrelerdir ve θ model parametrelerini temsil eder. Bunlar, ileri ve geri işlemler arasındaki tutarsızlığı en aza indirmek için optimize edilmiştir [47].

- **Varyasyonel çıkarım ve kayıp fonksiyonu:** Eğitim hedefi, gerçek veri dağılımı $q(x_0)$ ile model dağılımı $p_\theta(x_0)$ arasındaki Kullback-Leibler (KL) sapmasını en aza indirmeyi amaçlayan varyasyonel çıkarıma dayanmaktadır. Kanıt alt sınırı, optimizasyon hedefi olarak kullanılır:

$$\mathcal{L} = \mathbb{E}_q \left[\sum_{t=1}^T D_{\text{KL}}(q(x_t|x_{t-1}) || p_\theta(x_t|x_{t-1})) - \log p_\theta(x_0|x_1) \right] \quad (4.4)$$

Bu kayıp fonksiyonu, ters işlemin gerçek veri dağılımını doğru bir şekilde yaklaştırmasını sağlar [2].

- **Stokastik diferansiyel denklemler:** Difüzyon süreçleri, sürekli zamanlı modellerde tanıyan stokastik diferansiyel denklemler kullanılarak da formüle edilebilir:

$$dx = f(x, t) dt + g(t) dw \quad (4.5)$$

burada $f(x, t)$ sürüklenme terimini, $g(t)$ difüzyon katsayısını ve (w) Wiener sürecini temsil eder. Veri üretimi için gerekli olan ters SDE, aşağıdaki şekilde verilir:

$$dx = [f(x, t) - g(t)^2 \nabla_x \log p_t(x)] dt + g(t) d\bar{w} \quad (4.6)$$

burada $\nabla_x \log p_t(x)$ eğitim sırasında öğrenilen puan fonksiyonudur [47].

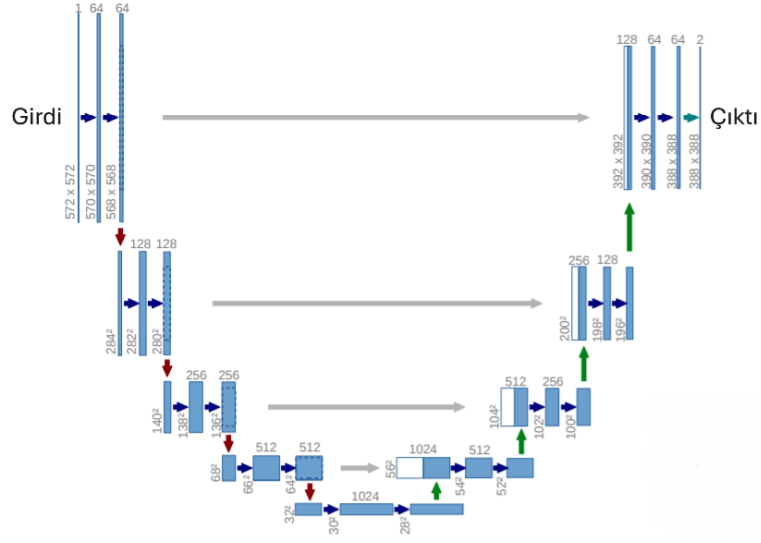
4.4 Kullanılan Mimariler

Difüzyon modelleri, çeşitli görevlerde büyük bir başarı sağlamıştır. Bu modellerin başarısında kullanılan mimariler kritik bir rol oynamaktadır. Difüzyon ağlarında yaygın olarak kullanılan temel mimarileri aşağıdaki şekilde detaylandırabiliriz.

4.4.1 U-Net tabanlı difüzyon modelleri

U-Net, özellikle tıbbi görüntüleme, segmentasyon ve rekonstrüksiyon gibi görevlerde yaygın olarak kullanılan bir konvolüsyonel sinir ağı tabanlı modeldir. U-Net, görüntüyü çok seviyeli konvolüsyonel katmanlar ile işler. Kodlayıcı kısmı, görüntüyü küçülterek önemli özellikleri çıkarır. Kod çözücü kısmı ise görüntüyü

tekrar genişleterek orijinal çözünürlüğe döndürür. Atlanan bağlantılar ile yüksek çözünürlüklü detaylar korunur.



Şekil 4.2: U-Net mimarisi.

U-Net tabanlı bir difüzyon modelini matematiksel olarak temsil aşağıdaki gibi temsil edilebilir:

$$\hat{x}_0 = \frac{x_t - \sqrt{1 - \beta_t} \epsilon_\theta(x_t, t)}{\sqrt{\beta_t}} \quad (4.7)$$

Burada, $\epsilon_\theta(x_t, t)$, U-Net modelinin gürültüyü tahmin ettiği fonksiyondur.

Avantajlar:

- Gürültü temizleme işlemlerinde başarılıdır.
- MRI gibi tıbbi görüntülerde eksik veya düşük çözünürlüklü verileri iyileştirmede etkilidir.

Dezavantajlar:

- Büyük görüntülerde işlem süresi uzun olabilir.
- Çok büyük boyutlu modellerde bellek tüketimi yüksektir.

4.4.2 Varyasyonel otokodlayıcı tabanlı difüzyon modelleri

VAE tabanlı difüzyon modelleri, hem varyasyonel otokodlayıcı hem de difüzyon modellerinin avantajlarını birleştirir. Kodlayıcı, görüntüyü bir düşük boyutlu latent uzaya sıkıştırır. Difüzyon modeli, bu uzayda işlemi yürütür. Kod çözücü, tamamlanan görüntüyü yeniden oluşturur.

Avantajlar:

- Öğrenme sürecini hızlandırır.

Dezavantajlar:

- Bazı durumlarda detay kaybı olabilir.

Difüzyon modelleri, görüntü işleme alanında devrim yaratmıştır. Farklı mimari türleri, belirli uygulamalar için daha uygun olabilir. U-Net tabanlı modeller, en yaygın kullanılan difüzyon mimarisidir. Çizelge 4.1’deki tablo, farklı mimarilerin avantajlarını, dezavantajlarını ve en uygun kullanım alanlarını özetlemektedir.

Çizelge 4.1: Farklı difüzyon mimarilerinin karşılaştırması.

Mimari	Avantajlar	Dezavantajlar	Kullanım alanları
U-Net	Basit ve etkili, geniş kullanım alanı	Hesaplama maliyeti	Görüntü tamamlama, süper çözünürlük
VAE	Hızlı eğitim süreci, düşük boyutlu temsil	Detay kaybı	veri ve görüntü üretimi

4.5 Modelleme Yaklaşımları

Difüzyon ağları, temelde bir olasılık dağılımını öğrenmeyi hedefleyen üretici modellerdir ve bu bağlamda farklı modelleme yaklaşımlarına sahiptir.

4.5.1 Latent difüzyon modelleri

Latent Difüzyon Modelleri, Kararlı difüzyon modelleri gibi büyük modellerin temelini oluşturan mimaridir. Bu model, doğrudan piksel uzayında çalışmak yerine latent uzayda çalışarak süreci hızlandırır. Varyasyonel otokodlayıcılar kullanarak görüntüleri sıkıştırır. Difüzyon işlemleri latent uzayda uygulanır. Son olarak, yeniden piksel uzayına dönüştürülerek çıktı elde edilir.

Matematiksel olarak, önce görüntü x_0 , latent kod z_0 'a dönüştürülür:

$$z_0 = E(x_0) \quad (4.8)$$

Daha sonra difüzyon modeli bu latent uzaydaki örneklerle gürültü ekleyerek ve temizleyerek çalışır. Son olarak tekrar piksel uzayına geri dönülür:

$$\hat{x}_0 = D(z_0) \quad (4.9)$$

Avantajlar:

- Çok daha düşük hesaplama maliyeti ile çalışır.
- Büyük ve yüksek çözünürlüklü görüntülerde başarılıdır.

Dezavantajlar:

- Latent uzayda bazı detaylar kaybolabilir.

4.5.2 Şartlı difüzyon modelleri

Şartlı difüzyon modelleri, giriş olarak ek bilgi (maskeler, etiketler, sınıf bilgisi vb.) kullanarak çalışır. Özellikle görüntü tamamlama ve süper çözünürlük gibi görevlerde etkilidir.

Şartlı difüzyon sürecinde, model giriş olarak ek bilgi y kullanarak tahmin yapar:

$$p(x_{t-1}|x_t, y) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t, y), \Sigma_\theta(x_t, t, y)) \quad (4.10)$$

Burada y , modele verilen ek bilgidir (örneğin, eksik bölgenin maskesi).

Avantajlar:

- Görüntü tamamlama ve restorasyonda en iyi performansı verir.

Dezavantajlar:

- Daha fazla hesaplama gücü gerektirir.

4.5.3 Skor bazlı difüzyon modelleri

Skor bazlı difüzyon modelleri, görüntü verisinin olasılık yoğunluğunu tahmin ederek çalışır. Özellikle mikro detayların düzeltilmesi ve kayıp alanların doldurulması gibi görevlerde başarılıdır.

Model, bir görüntünün olasılık dağılımının gradyanını öğrenerek gürültüyü kaldırır:

$$s_{\theta}(x_t, t) = \nabla_x \log p_t(x_t) \quad (4.11)$$

Burada $s_{\theta}(x_t, t)$, modelin verinin yoğunluk fonksiyonunun gradyanını tahmin ettiği skordur.

Avantajlar:

- Doğrudan pikseller üzerinde çalışabilir.

Dezavantajlar:

- Hesaplama maliyeti yüksek olabilir.

Farklı yaklaşımlar, belirli uygulamalar için daha uygun olabilir. Latent difüzyon modelleri, büyük ölçekli uygulamalar için en uygundur. Şartlı difüzyon modelleri ise görüntü tamamlama için en iyi performansı sağlar. Çizelge 4.2'deki tablo, farklı modelleme yaklaşımlarının avantajlarını, dezavantajlarını ve en uygun kullanım alanlarını özetlemektedir.

Çizelge 4.2: Farklı modelleme yaklaşımlarının karşılaştırması.

Yaklaşım	Avantajlar	Dezavantajlar	Kullanım alanları
Latent difüzyon	Daha hızlı, düşük maliyetli	Latent uzayda bazı detaylar kaybolabilir	Yüksek çözünürlüklü üretim
Şartlı Difüzyon	Eksik bölgeleri daha iyi tamamlar	Daha fazla hesaplama gücü gerektirir	İç boyama, restorasyon
Skor Bazlı	Çok gerçekçi görüntüler üretir	En yüksek hesaplama maliyeti	Medikal görüntüler, restorasyon

4.6 Eğitim ve Örneklemeye Yöntemleri

Difüzyon ağları, karmaşık veri dağılımlarını etkili bir şekilde yaklaşık olarak hesaplamak için gelişmiş eğitim ve örneklemeye tekniklerini kullanır. Bu yöntemler, hem sağlam eğitimi hem de yüksek kaliteli veri üretimini sağlayan olasılıksal çerçeveler ve optimizasyon stratejilerine dayanır. Bu bölüm, eğitim ve örneklemeye süreçlerinin ayrıntılı bir açıklamasını sağlar.

4.6.1 Eğitim yöntemleri

Difüzyon ağlarının eğitimi, bir Markov zincirinin ters sürecinin modellenmesine dayanır. Amaç, ağı ileri işlem sırasında uygulanan gürültülü dönüşümleri tersine çevirmesine olanak tanıyan parametreleri öğrenmektir. Eğitim sürecinin temel yönleri şunlardır:

- **Amaç fonksiyonu ve varyasyonel çıkarım:** Birincil eğitim hedefi, gerçek veri dağılımı $q(x_0)$ ile model dağılımı $p_\theta(x_0)$ arasındaki Kullback-Leibler (KL) sapmasını en aza indirmektir. Bu amaç doğrultusunda kanıt alt sınırı optimize edilir. Kayıp fonksiyonu Varyasyonel Çıkarım ve Kayıp Fonksiyonu bölümünde verilmiştir. Ters işlemin ileri yayılma işlemine yaklaşması amaçlanır. [2].
- **Gürültü planlaması ve parametrelendirmesi:** Gürültü planlaması, eğitim sürecini stabilize etmede önemli bir rol oynar. Doğrusal veya kosinüs gürültü çizelgeleri, ileri işlem sırasında eklenen gürültü seviyesini kontrol etmek için yaygın olarak kullanılır:

$$\alpha_t = 1 - \beta_t \quad (4.12)$$

burada β_t gürültü varyansıdır.

Ters işlem parametreleri $\mu_\theta(x_t, t)$ ve $\Sigma_\theta(x_t, t)$ genellikle karmaşık veri bağımlılıklarını yakalamak için U-Net mimarileri gibi sinir ağları kullanılarak modellenir [48].

- **Puan eşleştirme ve gürültü giderme:** Difüzyon modelleri, puan fonksiyonunun $\nabla_x \log p_t(x)$ doğrudan tahmin edildiği puan tabanlı öğrenme kullanılarak da eğitilebilir. Bu yöntem stokastik diferansiyel denklemlerin formülasyonlarıyla uyumludur ve eğitimin kararlılığını artırır [47].

4.6.2 Örneklemeye yöntemleri

Yayımla ağlarından örnekleme, saf gürültüden veri üretmek için yayılma sürecini tersine çevirmeyi içerir. Temel örnekleme teknikleri şunlardır:

- **Atasal örnekleme:** Bu yaklaşımda, ters işlem, standart bir Gauss dağılımından örneklenen bir gürültü vektörüyle başlanarak adım adım gerçekleştirilir. Her adım, öğrenilen ters geçiş çekirdeğini $p_\theta(x_{t-1}|x_t)$ uygulayarak kademeli olarak rafine edilmiş çıktılar üretir [2].
- **Puan tabanlı örnekleme:** Puan tabanlı örnekleme yöntemleri, örnekleme sürecini yönlendirmek için tahmini puan fonksiyonundan $\nabla_x \log p_t(x)$ yararlanır. Ters süreç, stokastik diferansiyel denklem olarak tanımlanır:

$$dx = [f(x,t) - g(t)^2 \nabla_x \log p_t(x)] dt + g(t) d\bar{w} \quad (4.13)$$

burada $f(x,t)$ ve $g(t)$ sırasıyla sürüklenme ve difüzyon katsayılarıdır [47].

Bir çok çalışmada yüksek kaliteli üretim için gereken adım sayısını azaltmak amacıyla hızlı örnekleme algoritmaları ve optimize edilmiş gürültü çizelgeleri gibi teknikler önerilmiştir.

4.7 Geleneksel Yöntemlere Göre Avantajları

Difüzyon ağları, çeşitli temel alanlarda çekişmeli üretici ağlar ve değişken otokodlayıcılar gibi geleneksel yöntemleri geride bırakarak, üretken modellemede dönüştürücü bir yaklaşım olarak ortaya çıkmıştır. Benzersiz tasarımları ve olasılıksal çerçeveleri, onları birçok uygulama için tercih edilen bir seçenek haline getirerek önemli avantajlar sağlar.

- **Eđitim kararsızlıklarına karşı sađlamlık:** ekiřmeli üretici ađların temel zorluklarından biri, modelin çeřitli ıktılar üretmediđi ve veri dađıtımının dar bir alt kümesine odaklandıđı mod öküşüne karşı duyarlılıklarıdır. Ancak, difüzyon ađları, kararlı eđitim dinamiklerini garanti eden iyi tanımlanmış bir olasılıksal çerçeve kullanarak bu sorunu hafifletir. İleri ve geri süreçlerin açık modellemesi, GAN'larda görülen düşmanca çatışmalardan kaçınarak üretim süreci üzerinde daha iyi kontrol sađlar [2].
- **Yüksek dođrulukta ve çeřitli ıktılar:** Difüzyon ađları, yüksek kaliteli ve çeřitli ıktılar üretmede mükemmeldir. Bu modeller, bir dizi adımla örnekleri kademeli olarak gürültüden arındırarak, görüntü sentezinde dikkate deđer bir dođruluk elde edebilir. Yayılma ađlarındaki kademeli iyileřtirme süreci, ince taneli ayrıntıların üretilmesini sađlar [48].
- **Basitleřtirilmiş optimizasyon hedefi:** Difüzyon ađlarının eđitim hedefi basittir, varyasyonel alt sınırı en aza indirmeye veya dođrudan puan fonksiyonunu öđrenmeye odaklanır. Bu özellik ile üretici ve ayırıcı arasındaki min-maks optimizasyonunun dengesiz eđitim ve yakınsama zorluklarına yol açabildiđi ekiřmeli üretici ađlardan farklılaşır. Difüzyon ađlarının daha basit optimizasyonu, onları eđitmeyi daha kolay ve daha öngörülebilir hale getirir [47].
- **Uygulama alanlarında esneklik:** Difüzyon ađları, çeřitli uygulama alanlarında dikkate deđer bir esneklik göstererek onları üretken modelleme ve veri sentezinde ok yönlü bir araç haline getirir. Tıbbi görüntülemelerde, özellikle MRI rekonstrüksiyonu ve gürültü azaltma gibi, dođru ve yüksek dođruluklu sonuçların tanı için kritik öneme sahip olduđu görevlerde görüntü kalitesini artırmada etkili olmuşlardır [2]. Ek olarak, bu ađlar metinden görüntü sentezini devrim niteliğinde deđiřtirmiş ve metinsel açıklamalardan tutarlı ve bađlamsal olarak hizalanmış görüntülerin oluşturulmasını sađlamıştır [49].

- **Gelişmiş örnekleme verimliliği:** İlk olarak ortaya atılan difüzyon modelleri çok sayıda örnekleme adımı gerektirse de, gürültü giderici difüzyon dolaylı modelleri (denoising diffusion implicit models, DDIM) gibi gelişmeler hesaplama yükünü önemli ölçüde azalttı. Bu iyileştirme, çıktı kalitesinden ödün vermeden difüzyon ağlarını gerçek dünya uygulamaları için daha pratik hale getirdi [48].
- **Ölçeklenebilirlik ve genelleme:** Difüzyon ağları güçlü ölçeklenebilirlik ve genelleme yetenekleri gösterir. Genellikle farklı veri kümeleri için kapsamlı ince ayar gerektiren geleneksel yöntemlerin aksine, difüzyon ağları olasılıksal yapıları nedeniyle çeşitli veri dağılımlarına iyi uyum sağlar. Bu, onları karmaşık, yüksek boyutlu verileri işlemede sağlam hale getirir [47].
- **Sürecin yorumlanabilirliği:** Difüzyon ağlarındaki adım adım gürültü giderme süreci, birçok geleneksel yöntemde bulunmayan bir yorumlanabilirlik düzeyi sunar. Her adım, gürültüden anlamlı bir veri noktasına kademeli bir dönüşüme karşılık gelir ve modelin nihai çıktısını nasıl oluşturduğuna dair bilgiler sağlar [2].

Özetle, yayılma ağları geleneksel üretken modellerde bulunan sınırlamaların çoğunu ele alır. Kararlılıkları, esneklikleri ve ölçeklenebilirlikleri, onları çeşitli alanlarda üretken görevler için güçlü bir araç haline getirir. Daha fazla ilerleme kaydedildikçe, yayılma ağlarının makine öğreniminin evriminde daha da belirgin bir rol oynaması muhtemeldir.

4.8 Difüzyon Ağlarının Temel Uygulamaları

Difüzyon ağları, karmaşık veri dağılımlarını modelleme ve yüksek kaliteli çıktılar üretme yetenekleri nedeniyle bir çok alanda kullanılmaya başlanmışlardır. Aşağıda difüzyon ağlarının önemli etki gösterdiği bazı temel alanlar yer almaktadır:

- **Tıbbi görüntüleme:** Tıbbi görüntülemede difüzyon ağları, görüntü yeniden oluşturma, gürültü azaltma ve süper çözünürlük gibi görevler için yaygın olarak kullanılır. Örneğin, MRI ve CT görüntülemede, bu modeller kritik tanı bilgilerini korurken eserleri ve gürültüyü kaldırarak görüntü netliğini artırır. Bu, geleneksel yöntemlerin genellikle radyasyon maruziyetini azaltmak için görüntü kalitesinden

ödün verdiği uygulamalardan ayrılmasını sağlar. Difüzyon modelleri ayrıca, yetersiz örnekleme verilerden yüksek kaliteli görüntüleri yeniden oluşturarak hızlandırılmış görüntülemeyi mümkün kılar ve kardiyak MRI gibi zamana duyarlı uygulamalardaki temel zorlukların üstesinde gelir.

- **Süper çözünürlüklü görüntüleme:** Difüzyon ağları, düşük çözünürlüklü girdilerden yüksek çözünürlüklü görüntüler üreterek süper çözünürlüklü gerektiren uygulamalarda yüksek performans gösterir. Genellikle eserlerde ayrıntı kaybına yol açan geleneksel enterpolasyon yöntemlerinin aksine, yayılma modelleri görüntüyü yinelemeli olarak iyileştirerek keskin ve doğru çıktılar sağlar [48,50]. Örnek uygulamalar arasında daha iyi coğrafi analiz için uydu görüntülerinin geliştirilmesi ve arşiv fotoğraflarının kalitesinin artırılması yer alır.
- **Metinden görüntü sentezi:** Difüzyon ağlarının en çığır açan uygulamalarından biri, metinden görüntü sentezidir. Bu modeller, GAN'lar gibi geleneksel yöntemlerden daha iyi performans göstererek, metinsel açıklamalara dayalı olarak oldukça ayrıntılı ve bağlamsal olarak hizalanmış görüntüler üretebilir. Özellikle doğal dil uygulamalarında yüksek çözünürlüklü, anlamsal olarak tutarlı görüntüler üretmede çok başarılı olmuştur. [49]. Bu, grafik tasarım, sanal gerçeklik ve içerik oluşturma gibi alanlarda yeni kapılar açmıştır.
- **Astronomik görüntüleme:** Genellikle atmosferik bozulmalardan ve sensör gürültüsünden etkilenen teleskop görüntülerinde bulanıklığı ve gürültüyü azaltma gibi görevlerde difüzyon ağlarından büyük ölçüde yararlanır. Altta yatan veri dağıtımını modellemek için olasılıkçı bir yaklaşım kullanarak, bu modeller göksel olayların analizinin daha iyi yapılmasını sağlayan yüksek kaliteli görüntüleri yeniden oluşturabilir.
- **Protein yapı tahmini:** Biyoenformatikte, difüzyon ağları moleküler yapıları modellemek ve protein katlama simülasyonlarına yardımcı olmak için kullanılmıştır. Bu modellerin yüksek boyutlu veri dağılımlarını işleme yeteneği, onları ilaç keşfi ve terapötik geliştirme için kritik bir görev olan karmaşık biyomoleküler etkileşimleri anlamak için özellikle uygun hale getirir [51].

- **Sanatsal uygulamalar:** Difüzyon ağları, sanat ve müzik üretimi gibi yaratıcı alanlarda da kullanım bulmuştur. Bu modeller, büyük görüntü veya ses veri kümeleri üzerinde eğitilerek yeni sanatsal stiller veya müzik besteleri üretebilir ve sanatçılara ve yaratıcılara yenilikçi fikirleri keşfetmeleri için araçlar sağlayabilir [50].
- **Genel veri sentezi:** Belirli uygulamaların ötesinde, difüzyon ağları, anomali tespiti, makine öğrenimi modellerini eğitmek için sentetik veri kümesi üretimi ve alan uyarlama gibi işlemlerde genel amaçlı veri sentezi için kullanılır. Çeşitli, yüksek doğrulukta çıktılar üretme yetenekleri, onları gerçek dünya verilerinin kıt veya önyargılı olduğu ortamlarda vazgeçilmez kılar [2,47].

Difüzyon ağlarının çok yönlülüğü, onları modern üretken modellemenin temel taşı olarak konumlandırmıştır. Çok çeşitli uygulamalardaki zorlukları ele alma yetenekleri, dönüştürücü potansiyellerini vurgulayarak, onları çeşitli alanlarda teknolojiyi ilerletmek için olmazsa olmaz bir araç haline getirir.

4.9 Sınırlamalar ve İyileştirme Önerileri

4.9.1 Sınırlamalar

Difüzyon ağlarının sağladığı önemli ilerlemelere rağmen, eksiklikleri ya da sınırlamaları yok değildir. Bu eksikliklerin üstesinden gelindiğinde çok daha iyi performans gösterecekleri ve teknolojiyi daha iyi bir seviyeye getirecekleri aşikardır. Aşağıda bazı sınırlamalar ve çözüm önerileri sunulmaktadır.

- **Hesaplama karmaşıklığı:** Difüzyon ağlarının hem eğitim hem de çıkarım aşaması hesaplama açısından karmaşık işlemleri barındırır. İleri ve geri süreçlerinin yinelemeli doğası, her biri önemli hesaplama kaynakları gerektiren çok sayıda adım içerir [2]. Bu yüksek maliyet, mobil cihazlar veya gömülü sistemler gibi gerçek zamanlı veya kaynak kısıtlı ortamlarda uygulanabilirliklerini sınırlar. Bunu hafifletmek için, performansı korurken difüzyon adımlarının sayısını azaltma çabaları çeşitli çalışmalarda ele alınmıştır. [48].

- **Eđitim kararsızlıkları:** Difüzyon ađlarının eđitimi, gürültü çizelgeleri ve öğrenme oranları dahil olmak üzere hiperparametrelerin dikkatli bir şekilde ayarlanmasını gerektirir. Uygunsuz yapılandırmalar kararsızlığa veya optimum olmayan performansa yol açabilir. İyileştirilmiş gürültü çizelgeleme yöntemleri veya parametrelili difüzyon süreçleri gibi sağlam eđitim tekniklerini dahil etmek, kararlılığı ve performansı artırabilir [47]. Ek olarak, belirli alanlarda transfer öğrenimi için önceden eđitilmiş difüzyon modellerinden yararlanmak, model başlatma ile ilgili zorlukların üstesinden gelinmesine yardımcı olabilir.
- **Büyük veri kümelerine bađımlılık:** Çođu derin öğrenme modeli gibi, difüzyon ađları da eđitim için kapsamlı, yüksek kaliteli veri kümeleri gerektirir. Tıbbi görüntüleme veya bilimsel araştırma gibi alanlarda, bu tür verileri elde etmek zor ve pahalıdır. Ayrıca genellikle gizlilik endişeleri yaratır. [52]. Sentetik veri üretimi ve veri artırma teknikleri bu sınırlamayı hafifletmeye yardımcı olabilir. Ayrıca, yarı-denetimli veya denetimsiz öğrenme gibi yöntemler etiketli verilere olan bađımlılıđı azaltabilir.
- **Yorumlanabilirlik ve açıklanabilirlik:** Difüzyon ađlarının diđer derin öğrenme modelleri gibi "kara kutu" doğası, özellikle sađlık hizmetleri ve otonom sistemler gibi kritik alanlarda yorumlanabilirlik konusunda endişelere yol açar [43]. Bu modellerin karar alma sürecini açıklamak için belirginlik haritaları veya özellik önem analizi gibi yöntemler geliştirmek hayati önem taşır. Altta yatan mekaniğin şeffaf ve sezgisel olduđu yorumlanabilir yayılma süreçleri üzerine yapılan araştırmalar da bu zorluğun üstesinden gelebilir.
- **Alanlarda genelleme:** Difüzyon ađları belirli alanlarda mükemmellik gösterse de, çeşitli, görünmeyen veri dağılımlarına genellemeleri bir zorluk olmaya devam etmektedir. Örneğin, doğal görüntüler üzerinde eđitilen bir model, sentetik veya alana özgü veri kümeleri üzerinde başarı gösteremeyebilir [50]. Alana özgü verilerde ince ayar yapma veya alan-deđişmez özellik gösterimleri kullanma gibi alan uyarlama teknikleri genellemeyi iyileştirebilir.

- **Enerji verimliliği:** Difüzyon ağlarını eğitme ve dağıtmanın enerji olarak maliyeti de önemli bir unsurdur. Yapay zeka araştırmalarında diğer ileri seviye modeller gibi difüzyon ağları çevresel kaygılara neden olur. Model budama, niceleme ve verimli mimari tasarımlar gibi teknikler, performansı önemli ölçüde tehlikeye atmadan enerji taleplerini azaltmaya yardımcı olabilir [45].

4.9.2 İyileştirme önerileri

Yukarıda verilen sınırlamaların üstesinden gelmek için gelecekteki araştırmalar şunlara odaklanabilir:

- **Verimli mimariler:** Hesaplama maliyetlerini azaltırken üretken kaliteyi koruyan hafif mimariler geliştirmek.
- **Geliştirilmiş gürültü planlaması:** İleri ve geri süreçleri optimize eden uyarlanabilir gürültü planlamalarını tanıtmak.
- **Sentetik ve artırılmış veriler:** Veri bağımlılığını azaltmak için sentetik veri üretimi ve gelişmiş artırma tekniklerinden yararlanmak.
- **Açıklanabilirlik çerçeveleri:** Yayılma ağlarının yorumlanabilirliğini geliştirmek için araçlar ve metodolojiler oluşturmak.
- **Çapraz alan genellemesi:** Çeşitli veri kümeleri arasında sağlamlığı artırmak için alan uyarlama tekniklerini dahil etmek [47].
- **Yeşil yapay zeka yaklaşımları:** Çevresel etkiyi en aza indirmek için enerji açısından verimli eğitim metodolojilerini uygulamak.

Difüzyon ağları geleneksel üretken modellere göre önemli avantajlar sunarken, bahsi geçen eksiklik ve sınırlamaların üstesinden gelmek daha geniş bir şekilde benimsenmeleri ve etkili olmaları için çok önemlidir. Verimlilik, yorumlanabilirlik ve veri erişilebilirliğine odaklanan gelişmeler, difüzyon ağlarını çeşitli alanlarda daha etkili hale getirebilir.

5. DİFÜZYON AĞLARI İLE GÖRÜNTÜ REKONSTRÜKSİYONU VE RESTORASYONU

5.1 Genel Bakış

Difüzyon ağları, görüntü rekonstrüksiyonu ve restorasyonu görevlerinde yüksek performans sunan güçlü bir üretken modelleme yaklaşımı olarak ortaya çıkmıştır. Karmaşık veri dağılımlarını modelleme ve bozulmuş görüntüleri gürültüden arındırma yetenekleri, onları özellikle tıbbi görüntüleme, uzaktan algılama ve genel bilgisayar görüşü görevlerindeki uygulamalar için uygun hale getirir [2,47,50].

Önceki modellerin ve kesin algoritmalara dayanan geleneksel yaklaşımların aksine, difüzyon ağları, bir görüntüyü gürültülü veya bozulmuş bir versiyondan yüksek kaliteli bir yeniden yapılandırmaya kademeli olarak iyileştiren olasılıksal bir gürültüden arındırma çerçevesinden yararlanır [48]. Bu özellik, varyasyonel otokodlayıcılar (VAE), çekişmeli üretici ağlar (GAN) ve denetlenen derin öğrenme tabanlı restorasyon modelleri gibi klasik yöntemlerden kalite, genelleme ve sağlamlık açısından daha iyi performans göstermelerini sağlar [38,53].

5.2 Görüntü İşleme Alanına Geçişi

Difüzyon modelleri, son yıllarda derin öğrenme tabanlı üretici modeller arasında dikkat çeken yöntemlerden biri haline gelmiştir. Bu modeller ilk olarak görüntü üretimi amacıyla geliştirilmiş, özellikle Ho ve arkadaşlarının sunduğu DDPM çalışmasıyla büyük ilgi görmüştür [2]. Bu modelde, bir görüntü kademeli olarak gürültü eklenerek saf gürültüye dönüştürülmekte ve bu süreç, öğrenilmiş bir model aracılığıyla tersine çevrilerek yeni görüntüler üretilmektedir. Bu üretim süreci, aynı zamanda modelin güçlü bir şekilde gürültü giderme yeteneği kazanmasını da sağlamaktadır.

Zamanla araştırmacılar, difüzyon modellerinin yalnızca görüntü üretimi için değil, aynı zamanda çeşitli görüntü işleme problemlerinin çözümünde de etkili bir şekilde kullanılabileceğini fark etmişlerdir. Bu farkındalık, difüzyon modellerinin restore

edici bir yaklaşım olarak yeniden değerlendirilmesine yol açmıştır. Özellikle Song ve Ermon'un sunduğu skor bazlı üretici modelleme yaklaşımı, difüzyon sürecine farklı bir bakış açısı getirerek ters problemlerin çözümünde kullanılabilir hale getirmiştir [3]. Bu yöntemde model, belirli bir zaman adımında gürültülenmiş görüntünün altında yatan veri dağılımının gradyanını tahmin ederek görüntünün iteratif olarak geri kazanılmasını sağlar.

Buna ek olarak, PnP yaklaşımı difüzyon modellerinin klasik görüntü işleme algoritmalarının yerine geçebileceğini ortaya koymuştur. Bu yöntemde, difüzyon modelleri görüntü işleme problemlerine birer ön bilgi olarak entegre edilir ve geleneksel regularizasyon tekniklerinin yerini alır [54].

Difüzyon modelleri ayrıca, koşullu yapılarla daha esnek hale getirilmiştir. Koşullu difüzyon modelleri, belirli bir giriş verisine (örneğin düşük çözünürlüklü veya eksik veri içeren bir görüntüye) koşullanarak yüksek kaliteli görüntüler üretmeyi öğrenmektedir [55]. Bu yapı, özellikle süper çözünürlük, görüntü tamamlama ve tıbbi görüntülerde eksik verinin tamamlanması gibi görevlerde başarıyla uygulanmıştır.

Özellikle tıbbi görüntüleme alanında, örneğin MRI verilerinde, difüzyon modelleri maskelenmiş veya eksik k-space verilerinden tam görüntüler üretmek için ters problem çözüm sürecine doğrudan entegre edilmektedir [54].

Son yıllarda, bu modellerin hesaplama maliyetlerini düşürmek amacıyla daha verimli yapılara yönelik çalışmalar da dikkat çekmektedir. Bu bağlamda önerilen Artık Kaydırmalı Difüzyon yöntemi, ters difüzyon sürecinde her adımdaki öngörülen artıkların doğrudan bir sonraki adıma aktarılması prensibine dayanarak, modelin genel örnekleme sürecini hızlandırmakta ve kalite kaybı yaşamadan daha verimli bir geri kazanım sağlamaktadır [1].

Sonuç olarak, difüzyon modellerinin görüntü işleme alanına geçişi ani bir dönüşüm değil, üretim kabiliyetlerinin doğal bir evrimi olarak değerlendirilebilir. Üretim sürecinin merkezinde yer alan ters difüzyon yapısı, aynı zamanda birçok temel görüntü işleme probleminin çözümünde etkili bir strateji haline gelmiştir.

5.3 Yöntemler

Görüntü restorasyonu gibi düşük seviyeli görsel görevlerde son yıllarda büyük başarı elde eden generatif modeller, özellikle difüzyon tabanlı yaklaşımlar ile dikkat çekmektedir. Bu bölümde, görüntü üretiminde yüksek doğruluk ve çeşitlilik sağlayan DDPM ve önceden eğitilmiş modellerden yararlanarak yeniden eğitime ihtiyaç duymadan farklı görevlerde kullanılabilen difüzyon yaklaşımı ele alınacaktır. DDPM, temel bir difüzyon süreci üzerinden görüntü sentezini adım adım öğrenirken, tak ve çalıştır yaklaşımı daha esnek bir yapı sunarak mevcut difüzyon modellerine koşullu bilgi entegre edilmesini mümkün kılar.

Bunlara ek olarak, yakın zamanda önerilen artık kaydırmalı difüzyon yöntemi, geleneksel difüzyon modellerinde her bir adımda tahmin edilen kalanı doğrudan bir sonraki adıma aktaran yenilikçi yapısıyla öne çıkmaktadır. Bu yaklaşım, örnekleme sürecini önemli ölçüde hızlandırmakta ve modelin görüntü geri kazanım görevlerinde daha verimli bir şekilde çalışmasına olanak sağlamaktadır.

Bu üç yaklaşım, difüzyon modellerinin hem teorik temellerini hem de pratik uygulama alanlarını daha iyi anlamak açısından kritik öneme sahiptir.

5.3.1 Gürültü giderici olasılık modeli

DDPM, temel olarak bir ileri difüzyon ve bir geri difüzyon sürecinden oluşan Markov zincirine dayalı bir difüzyon modelidir. İleri süreçte, veriye adım adım Gauss gürültüsü eklenerek tam rastgele hale getirilir. Bu süreç aşağıdaki şekilde tanımlanır [2]:

$$q(x_t | x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t I) \quad (5.1)$$

Burada x_0 orijinal görüntüyü, x_t ise t zaman adımıdaki bozunmuş görüntüyü temsil eder. β_t parametresi her adımda eklenen gürültünün şiddetini kontrol eder.

Toplam ileri süreç şu şekilde modellenenabilir:

$$q(x_t | x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}x_0, (1 - \bar{\alpha}_t)I) \quad (5.2)$$

Burada $\bar{\alpha}_t = \prod_{s=1}^t (1 - \beta_s)$, zaman boyunca biriken gürültü etkisini ifade eder.

Geri süreçte, amaç bu gürültülü örneklerden orijinal veriye geri dönmektir. Bu süreç, bir parametrelili model p_θ ile tahmin edilir:

$$p_\theta(x_{t-1} | x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)) \quad (5.3)$$

DDPM modelleri genellikle gürültü tahmini yaklaşımını benimser ve öğrenilecek hedef olarak ε gürültüsünü belirler:

$$\mathcal{L}_{\text{simple}} = \mathbb{E}_{x_0, \varepsilon, t} \left[\|\varepsilon - \varepsilon_\theta(x_t, t)\|^2 \right] \quad (5.4)$$

Bu kayıp fonksiyonu, modelin her adımda eklenmiş olan gürültüyü doğru şekilde tahmin etmesini sağlayarak başarıyla orijinal görüntüyü yeniden üretmesini sağlar.

5.3.2 Tak ve çalıştır difüzyon modeli

Tak ve çalıştır difüzyon yaklaşımı, önceden eğitilmiş bir difüzyon modelini kullanarak herhangi bir görüntü oluşturma görevini yeniden eğitime ihtiyaç duymadan çözmeyi amaçlar. Bu, özellikle sıfırdan öğrenme senaryolarında oldukça etkilidir. Bu yöntemde, x_t 'den x_{t-1} 'e geçişte, koşullu bilgi (örneğin bir maske veya bir bozunum) dışarıdan müdahale yoluyla modele entegre edilir [4].

Bu müdahale genellikle Langevin dinamiği veya skor tabanlı yönlendirme yoluyla gerçekleştirilir:

$$x_{t-1} = x_t + \frac{\sigma_t^2}{2} \nabla_{x_t} \log p(y | x_t) + \sigma_t z, \quad z \sim \mathcal{N}(0, I) \quad (5.5)$$

Burada $p(y | x_t)$, koşullu bilginin olasılığını temsil ederken, $\nabla_{x_t} \log p(y | x_t)$ ifadesi görüntünün belirtilen koşula uygun olacak şekilde yönlendirilmesini sağlar.

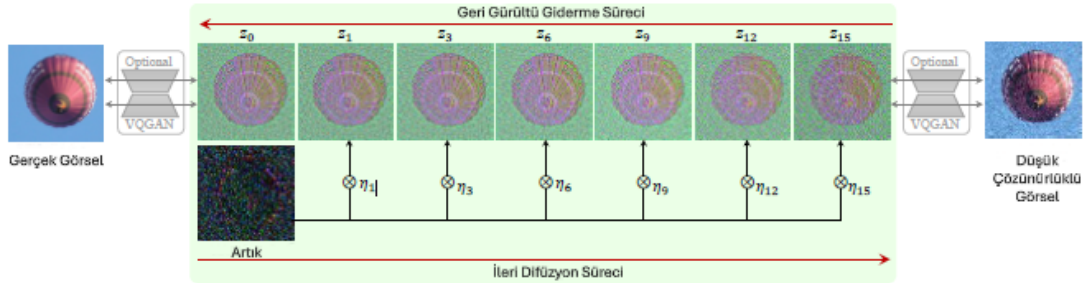
PnP yapılar genellikle aşağıdaki iki bileşeni kullanır:

- **Önceden eğitilmiş difüzyon modeli:** $p(x)$ olasılık dağılımını temsil eden, önceden eğitilmiş bir difüzyon modeli.
- **Şartlandırma fonksiyonu:** Harici görev bilgisini içeren, örneğin bir denetleyici model veya enerji temelli fonksiyon.

Bu yapı sayesinde, PnP difüzyon modelleri, herhangi bir görüntü restorasyon görevine hızlıca uyum sağlayabilir ve yeniden eğitime gerek duymadan başarılı sonuçlar üretebilir.

5.3.3 Artık kaydırmalı difüzyon modeli

Standart difüzyon modellerinde, her adımda model tarafından tahmin edilen gürültü (veya artık) yalnızca o adıma özgü olarak kullanılırken, artık kaydırmalı difüzyon yöntemi, bu artık bilgisini bir sonraki adıma doğrudan taşıyarak örnekleme sürecinde daha verimli bir yapı sunar [1]. Bu yöntem, özellikle görüntü restorasyonu görevlerinde modelin hız ve performans açısından daha etkili çalışmasını hedefler.



Şekil 5.1: Artık kaydırma yönteminin yüksek çözünürlük - düşük çözünürlük görüntü çifti arasında artıkları kaydırması [1].

Bu yaklaşımda, klasik geri difüzyon formülüne ek olarak artık aktarımı aşağıdaki şekilde modellenir:

$$\hat{\epsilon}_t = \epsilon_\theta(x_t, t) + s_{t+1} \quad (5.6)$$

Burada s_{t+1} , bir sonraki adımdan gelen artık bilgisini ifade eder ve modelin sonraki tahminlerinde doğrudan kullanılmak üzere kaydırılır.

Geri difüzyon bu bilgi ile güncellenerek aşağıdaki şekilde gerçekleştirilir:

$$x_{t-1} = \frac{1}{\sqrt{1-\beta_t}}(x_t - \beta_t \hat{\epsilon}_t) + \sigma_t z, \quad z \sim \mathcal{N}(0, I) \quad (5.7)$$

Bu yapı sayesinde, her adımda yeniden sıfırdan tahmin yapmak yerine, önceki adımdan taşınan bilgiyle daha kararlı ve hızlı bir örnekleme süreci elde edilir. Artık kaydırmalı difüzyon, özellikle düşük adımlı örnekleme senaryolarında kalite kaybı yaşamadan hesaplama maliyetini azaltmakta etkili bir çözüm sunar.

Çizelge 5.1: Farklı süper çözünürlük yöntemlerinin PSNR, SSIM ve LPIPS metriklerine göre karşılaştırılması [1].

Yöntem	PSNR	SSIM	LPIPS
ESRGAN	20.67	0.448	0.485
RealSR-JPEG	23.11	0.591	0.326
BSRGAN	24.42	0.659	0.259
SwinIR	23.99	0.667	0.238
RealESRGAN	24.04	0.665	0.254
DASR	24.75	0.675	0.250
LDM-15	24.89	0.670	0.269
ResShift	25.01	0.677	0.231

Klasik difüzyon yöntemlerinde her adımda tahmin edilen verinin orijinal görüntüye yaklaşması beklenir. Artık kaydırmalı difüzyon yönteminde ise model her difüzyon adımında yalnızca gürültüyü değil, aynı zamanda görüntü artıklarını da kullandığı için bir önceki tahmin ile gerçek görüntü arasındaki farkı bilir. Böylece model, difüzyon sürecinde bilgi kaybını azaltarak daha hızlı yakınsama ve daha kaliteli sonuçlar üretme potansiyeline sahip olur.

5.4 Kazanımlar

Son yıllarda, difüzyon tabanlı generatif modeller, yüksek kaliteli görüntü üretimi konusundaki çığır açan başarılarıyla bilgisayarla görme alanında önemli bir araştırma odağı haline gelmiştir. Özellikle, bu modellerin rastgele gürültüden görsel veri üretme kabiliyeti, sadece sentezleme değil, aynı zamanda görüntü

Çizelge 5.2: Artık kaydırmalı difüzyon ile klasik difüzyon yaklaşımlarının karşılaştırılması.

Özellik	Klasik Difüzyon	Artık Kaydırmalı Difüzyon
Tahmin edilen bilgi	Gürültü	Artık bilgi
Geri adım formülü	$x_{t-1} = \dots - \hat{\epsilon}_\theta$	$x_{t-1} = x_t + r_t$
Hesaplama maliyeti	Yüksek (50–1000 adım)	Düşük (4–12 adım)
Modelin hedefi	Tam gürültü tahmini	Kademeli düzeltme öğrenimi
Örnekleme süresi	Uzun	Kısa

rekonstrüksiyonu ve restorasyonu gibi zorlu ters problemlerde de dikkate değer performans göstermelerini sağlamıştır. Difüzyon modelleri, bozulmuş veya eksik bilgi içeren görüntüleri adım adım gürültüden arındırarak orijinal yapıya yakın şekilde yeniden oluşturabilmektedir. Bu süreç, geleneksel yöntemlerin sıklıkla karşılaştığı bulanıklık, detay kaybı ve gerçekçilik sorunlarının üstesinden gelmekte oldukça etkilidir. Görüntü rekonstrüksiyonunda (örneğin, tıbbi görüntüleme veya sıkıştırılmış sensing senaryolarında) yüksek doğrulukla yapıların geri kazanımı sağlanabilirken; restorasyon görevlerinde (gürültü giderme, bulanıklık giderme, süper çözünürlük vb.) görsel kalitenin önemli ölçüde arttığı gözlemlenmiştir. Bu başarılar, difüzyon modellerinin yalnızca teorik değil, aynı zamanda pratik uygulamalarda da güçlü araçlar haline geldiğini göstermektedir.

5.4.1 Görüntü rekonstrüksiyonu alanında kazanımlar

Görüntü rekonstrüksiyonu, eksik, gürültülü veya bozulmuş giriş verilerinden yüksek doğrulukta görüntüler oluşturmanın amaçlandığı bilgisayarlı görme ve tıbbi görüntüleme gibi kritik alanlarda kullanılan temel uygulamalardır. Bilgisayarlı tomografide filtrelenmiş geri projeksiyon, MRI’da sıkıştırılmış algılama ve uydu görüntüleme enterpolasyon tabanlı yöntemler gibi geleneksel görüntü yeniden oluşturma yöntemleri genellikle matematiksel sezgisel yöntemlere ve elle tasarlanmış önsel bilgilere dayanır. Bu teknikler kısıtlı senaryolarda etkili olsa da karmaşık, yüksek boyutlu verilerde iyi performans gösteremez. Gürültüye ve eksik bilgilere karşı oldukça hassastır.

Çekişmeli üretici ağlar ve varyasyonel otokodlayıcılar gibi derin üretken modellerin ortaya çıkmasıyla, öğrenilmiş gösterimler görüntü yeniden yapılandırma kalitesini

önemli ölçüde iyileştirmiştir. Ancak, bu modeller genellikle mod çöküşü (GAN'larda) ve yeniden yapılandırmalarda bulanıklık (VAE'lerde) gibi sorunlardan muzdariptir [50]. Buna karşılık, difüzyon modelleri olasılıksal bir gürültü giderme süreciyle görüntüleri kademeli olarak iyileştirerek görüntü oluşturma ve yeniden yapılandırmaya matematiksel olarak ilkeli ve oldukça etkili bir yaklaşım sunar.

Difüzyon ağları, saf bir gürültü dağılımından başlayarak ve öğrenilmiş gürültü giderme dönüşümlerini kullanarak onu yinelemeli olarak iyileştirerek görüntüler oluşturur. Bu süreç, anlamlı yapıları yeniden oluşturmak için gürültünün ileri yönde eklendiği ve ters yönde kaldırıldığı stokastik diferansiyel denklemleri taklit eder. Bu yaklaşımın temel avantajı, mod çöküşünü önlemesi, çeşitli ve yüksek kaliteli görüntü oluşturmaya olanak tanınması, birden fazla kıyaslamada GAN'ları ve VAE'leri geride bırakmasıdır [48,50].

Metinden görüntü oluşturmadaki son gelişmeler, difüzyon modellerinin gücünü daha da kanıtıyor. Bazı modeller, metinsel açıklamalara bağlı fotogerçekçi görüntüler oluşturmak için difüzyon süreçlerinden yararlanır. Benzer şekilde, tıbbi görüntüleme, gerçekçi MRI ve CT taramalarını sentezlemek, veri artırmaya yardımcı olmak ve düşük veri ayarlarında sağlam yapay zeka modellerini eğitmek için üretici difüzyon modelleri kullanılmıştır [56].

Difüzyon ağlarıyla rekonstrüksiyon, sorunu bir gürültü giderme süreci olarak ele alarak klasik yeniden oluşturma tekniklerinden farklıdır. Deterministik ters dönüşümler uygulamak yerine, difüzyon modelleri veri odaklı bir şekilde gürültü bozulmasını tersine çevirmeyi öğrenir [52]. Bu, onları karmaşık gürültü desenlerine, eksik verilere ve yapılandırılmış eserlere daha uyumlu hale getirir.

MRI yapılandırmada difüzyon modelleri yetersiz örneklenmiş k-uzay verilerinden yüksek kaliteli görüntüleri yeniden oluşturarak büyük bir başarı göstermiştir [54]. Seyreklik kısıtlamaları uygulayan geleneksel sıkıştırılmış algılamalı manyetik rezonans görüntülemenin (compressed sensing magnetic resonance imaging, CS-MRI) aksine, difüzyon modelleri, bozuk ölçümlerden doğrudan görüntüleri yeniden oluşturmayı öğrenerek, gürültünün etkisini önemli ölçüde azaltır.

Örneğin, Jalal ve arkadaşları, MRI yeniden yapılandırması için bir difüzyon olasılık modeli önerdi ve yaklaşımları klasik sıkıştırılmış algılama (compressed sensing, CS) tabanlı yöntemlerden ve hatta özellikle çok az örneklenmiş verilerin olduğu durumlarda denetlenen derin öğrenme modellerinden daha iyi performans gösterdi [52]. Çalışmalar difüzyon modellerinin farklı görüntüleme koşullarında daha iyi genelleme sunduğunu da göstermektedir.

Başka bir ilgi çekici uygulama, yüksek kaliteli taramaları yeniden yapılandırmak için difüzyon modellerinin kullanıldığı düşük doz CT görüntülemesidir [56]. Düşük doz CT için eğitilen geleneksel derin öğrenme modelleri istenmeyen yumuşatmalara neden olurken, difüzyon tabanlı yeniden yapılandırmalar gürültüyü etkili şekilde ortadan kaldırarak ince yapısal ayrıntıları korur.

- **Karmaşık gürültü ve eksik verilerin ele alınması:** Difüzyon ağları, bozuk görüntülerden eksik bilgileri geri yüklemeye mükemmeldir ve bu da onları özellikle yetersiz örneklenmeli tıbbi taramaları ve uydu görüntüsü iyileştirmeleri için kullanışlı hale getirir.
- **Artefaktlara karşı dayanıklılık:** Gerçekçi olmayan dokular veya eserler sunabilen GAN'ların aksine, difüzyon ağları dama tahtası deseni gibi sorunlardan kaçınarak gerçekçi ve tutarlı görüntüler üretir [50].
- **Veri kümeleri arasında daha iyi genelleme:** Gözetimli derin öğrenme modelleri büyük miktarda eşleştirilmiş gürültülü-temiz veriye ihtiyaç duyarken, difüzyon ağları dağıtım kaymalarına karşı sağlamlığı artıran kendi kendini denetleyen öğrenme yöntemleriyle eğitilebilir [47].
- **Üstün görüntü kalitesi:** Tıbbi görüntü sentezi ve yeniden yapılandırma görevlerinde gösterildiği gibi, difüzyon ağları algısal kalitede GAN'ları ve ayrıntı korumada VAE'leri geride bırakmaktadır [48,52].
- **Matematiksel olarak ilkeli yaklaşım:** Sezgisel tabanlı derin öğrenme yöntemlerinden farklı olarak, difüzyon modelleri olasılıksal üretken çerçevelerden yararlanır ve bunları daha yorumlanabilir ve teorik olarak temellendirilmiş hale getirir [46].

5.4.2 Görüntü restorasyonu alanında kazanımlar

Gürültü azaltma ve görüntü onarma işlemleri, tıbbi görüntüleme, uzaktan algılama, astronomik görüntüleme ve genel bilgisayarlı görüntüleme uygulamaları dahil olmak üzere çeşitli alanlarda kritik işlemlerdir. Bu alanlarda geleneksel yöntemler yaygın olarak kullanılır ancak karmaşık gürültü desenleri ve yüksek frekanslı görüntü ayrıntılarının işlenmesinde sınırlamaları vardır. Son zamanlarda, difüzyon ağları, olasılıksal modelleme ve yinelemeli iyileştirmeyi kullanarak gürültüyü azaltmak ve bozulan görüntüyü onarmak için güçlü bir yaklaşım olarak kullanılmaktadırlar.

Gürültü özellikle görüntülerin sensör etkileri, çevresel girişim veya kayıplı sıkıştırma nedeniyle bozulduğu alanlarda temel bir sorundur. Geleneksel gürültü giderme yaklaşımları genellikle ince ayrıntıların kaybından muzdariptir ve ön bilgilere güvenir. Difüzyon tabanlı restorasyon gürültüyü öğrenerek etkili bir şekilde ortadan kaldırabilir.

- **Açık gürültü modellemesi:** Kapalı kutu özellik öğrenimine dayanan CNN'ler ve GAN'ların aksine, difüzyon ağları gürültü dağılımını stokastik diferansiyel denklemler kullanarak açıkça modeller.
- **Tekrarlayan iyileştirme:** Tek adımlı gürültü giderme yapmak yerine, difüzyon tabanlı yaklaşımlar görüntüyü kademeli olarak iyileştirerek ince ayrıntıları bulanıklaştırmadan gürültüyü giderir.
- **Çeşitli gürültü tipleri ile çalışabilme:** Difüzyon tabanlı yöntemler, MRI bozulmaları, atmosferik bozulmalar ve hiperspektral görüntülemeadaki sensör gürültüsü gibi yapılandırılmış gürültüyü gidermede özellikle etkilidir.

5.5 Önemli Uygulamalar

Difüzyon modellerinin görüntü işleme alanında sağladığı başarılar, özellikle klasik ve zorlu görevlerde dikkat çekici biçimde ortaya çıkmaktadır. Görüntü tamamlama, süper çözünürlük, restorasyon ve gürültü giderme gibi görevler, hem araştırma literatüründe hem de gerçek dünya uygulamalarında sıkça karşılaşılan ve çözümü karmaşık olan problemlerdir. Bu görevlerde, geleneksel yöntemler ya da GAN tabanlı çözümler

sıklıkla yapay görsellik, gerçek dışı dokular ya da detay kaybı gibi problemlerle karşılaşmaktadır. Ancak difüzyon modelleri, güçlü generatif kapasiteleri sayesinde bu görevlerde doğal ve yüksek kaliteli çıktılar üretebilme potansiyeli sunmaktadır.

Bu bağlamda, literatürde öne çıkan bazı difüzyon tabanlı uygulamalar, farklı çözüm stratejileriyle bu sorunlara yenilikçi yaklaşımlar getirmiştir. Özellikle DDRM, DiffIR, DPIR gibi uygulamalar, çeşitli bozulma türlerine karşı etkili çözümler sunarak, difüzyon modellerinin esnekliğini ve gücünü gözler önüne sermektedir. Bu bölümde, söz konusu yöntemler detaylı biçimde incelenecek; mimari tasarımları, çalışma prensipleri ve elde ettikleri sonuçlar bağlamında değerlendirilecektir.

5.5.1 Gürültü giderici difüzyon olasılık modeli ile görüntü tamamlama

DDPM, bir görüntüyü aşamalı olarak gürültü ekleyerek bozan ve ardından adım adım bu gürültüyü azaltarak orijinal görüntüyü geri getiren bir üretici modeldir. Bu süreç gürültü giderme gibi restorasyon görevlerinin yanında eksik veri tamamlama gibi görüntü rekonstrüksiyon görevlerinde de oldukça güçlüdür. Özellikle maskelenmiş alanların yüksek kalitede geri kazanımında etkili sonuçlar sunmaktadır [2].

DDPM, eksik veya maskelenmiş görüntülerin tamamlanmasında kullanılabilir. Bu işlem, standart difüzyon sürecine bir maske eklenerek gerçekleştirilir.

- **Maskeleme işlemi:** Öncelikle eksik alanları belirlemek için bir maske matrisi $M(x)$ tanımlanır:

$$M(x) = \begin{cases} 1, & \text{Eğer piksel eksikse} \\ 0, & \text{Eğer piksel eksik değilse} \end{cases} \quad (5.8)$$

Burada:

- $M(x) = 1$ olan bölgeler eksik olan pikselleri gösterir.
- $M(x) = 0$ olan bölgeler sağlam piksellerdir.

Eksik bölgelere rastgele gürültü eklenerek, modelin sadece bu bölgeleri yeniden oluşturması sağlanır.

- **Maskeli difüzyon süreci:** Gürültü ekleme işlemi maskeli görüntü x_M üzerinden gerçekleştirilir:

$$q(x_t|x_{t-1}, M) = \mathcal{N}(x_t; (1-M)x_{t-1} + M \cdot \eta, \beta_t I) \quad (5.9)$$

Burada:

- $\eta \sim \mathcal{N}(0, I)$ rastgele bir Gauss gürültüsünü temsil eder.
- Eksik olmayan bölgeler doğrudan korunur, ancak eksik bölgeler gürültülenir.

- **Geri difüzyon süreci ile tamamlama:** Eksik alanları yeniden oluşturmak için model, gürültülü görüntüden başlayarak her adımda eksik pikselleri tahmin eder:

$$p_\theta(x_{t-1}|x_t, M) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t, M), \Sigma_\theta(x_t, t, M)) \quad (5.10)$$

Burada:

- μ_θ ve Σ_θ , modelin eksik bölgeleri tahmin ederken kullandığı parametrelerdir.
- Model, eksik bölgeleri kademeli olarak doldurur ve maskesiz bölgeleri sabit tutar.

Son aşamada, tamamlanan görüntü maske uygulanarak orijinal sağlam bölgelerle birleştirilir:

$$x_{\text{final}} = (1 - M)x_{\text{orijinal}} + Mx_{\text{tamamlanan}} \quad (5.11)$$

Bu işlem sonucunda eksik bölgeyi doğal bir şekilde tamamlayan yüksek kaliteli bir görüntü elde edilir.

DDPM, eksik veri tamamlama görevlerinde başarılı sonuçlar veren güçlü bir modeldir. Aşamalı gürültü ekleme ve kaldırma sürecini öğrendiği için çeşitli gürültü seviyelerine uyum sağlayabilir. Yüksek çözünürlüklü görüntülerde ayrıntıları daha iyi yakalayabilir. Eksik bölgeleri tamamlamak için farklı örnekler üretebilir, yani tek bir tahmine bağlı kalmaz. Özellikle medikal görüntüleme, sanat restorasyonu ve süper

çözünürlük gibi alanlarda yaygın olarak kullanılmaktadır. Gürültülü veya eksik veri içeren görüntüler üzerinde çalışırken, U-Net tabanlı DDPM modelleri en iyi sonuçları vermektedir.

5.5.2 Gürültü giderici difüzyon restorasyon yöntemi

DDRM, difüzyon tabanlı modelleri kullanarak bozulmuş görüntülerin yüksek kaliteli bir şekilde düzeltilmesini amaçlayan bir yöntemdir [5]. Difüzyon tabanlı geleneksel modellerden farklı olarak DDRM, rastgele görüntü üretmek yerine, verilen bozulmuş görüntüyü orijinal haline en yakın şekilde restore etmeyi hedefler.

İleri difüzyon süreci

DDRM, bozulmuş bir görüntüyü tamamen rastgele gürültüye dönüştürmeye çalışır. Bozulmuş görüntü y , bilinmeyen temiz görüntü x_0 ve bozulma operatörü A ile aşağıdaki gibi ifade edilir:

$$y = Ax_0 + \varepsilon \quad (5.12)$$

Burada A , bulanıklık, süper çözünürlük veya eksik pikseller gibi bozulmaları içeren bir operatördür. Difüzyon sürecinde, görüntü aşamalı olarak gürültülenerek tamamen rastgele bir dağılıma ulaşır:

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t I) \quad (5.13)$$

Geri difüzyon süreci

DDRM, geri difüzyon sürecinde bozulmuş görüntüyü adım adım restore ederek yüksek kaliteli hale getirir. Bu süreç aşağıdaki şekilde formüle edilir:

$$p(x_{t-1}|x_t, y) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t, y), \Sigma_\theta(x_t, t, y)) \quad (5.14)$$

Burada μ_θ ve Σ_θ , modelin öğrenilmiş parametreleridir ve görüntünün hem bozulmuş hali y hem de mevcut durumu x_t dikkate alınarak orijinal görüntüye en yakın hale getirilmesini sağlar.

DDRM, görüntü restorasyon görevlerinde etkili bir şekilde kullanılır. Düşük çözünürlüklü görüntüleri yüksek kaliteli hale getirebilir. Gaussian, Poisson ve Speckle gürültülerini temizleyebilir. Hareket bulanıklığı veya odak dışı bulanıklıkları düzeltebilir. Koşullu difüzyon süreci sayesinde farklı türde bozulmalara uyarlanabilen model, özellikle tıbbi görüntüleme ve düşük kaliteli veri setlerinde büyük potansiyele sahiptir. Maskelenmiş veya eksik bölümleri tamamlayarak doğal bir görüntü oluşturabilir. Bu kısmı zaten rekonstrüksiyon tarafında incelemiştik. Sıkıştırma işlemlerinden kaynaklanan bozulmaları giderebilir.

5.5.3 Difüzyon tabanlı görüntü restorasyonu

DiffIR, görüntü restorasyonu için geliştirilen difüzyon tabanlı bir modeldir. Gürültü giderme, süper çözünürlük ve bulanıklık giderme gibi görevlerde üstün performans göstermeyi hedefler [57]. Geleneksel derin öğrenme yöntemlerinden farklı olarak, difüzyon olasılıksal modelleri (diffusion probabilistic models, DPM) kullanarak görüntü restorasyon problemlerini çözmektedir.

İleri difüzyon süreci

DiffIR, geleneksel difüzyon modelleri gibi ileri difüzyon süreci ile çalışır. Bu süreçte, temiz bir görüntü x_0 , artan miktarda Gaussian gürültüsü eklenerek aşamalı olarak bozulur. Matematiksel olarak bu işlem aşağıdaki gibi tanımlanır:

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t I) \quad (5.15)$$

Burada:

- x_t, t , adımda gürültülenmiş görüntüyü,
- β_t , belirli bir zaman adımıdaki gürültü miktarını gösteren bir parametreyi,
- \mathcal{N} , Gaussian dağılımını ifade eder.

Geri difüzyon süreci

DiffIR, bozulmuş bir görüntüyü eski haline döndürmek için geri difüzyon sürecini uygular. Model, giriş olarak verilen bozulmuş görüntüyü temiz versiyonuna dönüştürmeye çalışır:

$$p(x_{t-1}|x_t, y) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t, y), \Sigma_\theta(x_t, t, y)) \quad (5.16)$$

Burada:

- y , bozulmuş giriş görüntüsünü,
- μ_θ ve Σ_θ , modelin tahmin ettiği ortalama ve kovaryans matrislerini ifade eder.

DiffIR, çeşitli görüntü restorasyon görevlerinde başarıyla uygulanmaktadır. Gaussian, Poisson ve Speckle gürültülerini azaltmak için kullanılabilir. Geleneksel yöntemlere göre daha az bilgi kaybı ile temiz görüntü üretir. Düşük çözünürlüklü görüntülerden yüksek kaliteli görüntüler üretebilir. SRGAN ve ESRGAN gibi yöntemlere kıyasla daha iyi PSNR ve SSIM skorları almıştır. Kamera hareketi veya odak dışı bulanıklık gibi etkileri düzeltebilir. Geleneksel Wiener filtresi veya GAN tabanlı yaklaşımlardan daha detaylı sonuçlar üretir. JPEG sıkıştırmasından kaynaklanan bozulmaları ortadan kaldırabilir.

DiffIR, geleneksel görüntü restorasyon yöntemleriyle kıyaslandığında önemli avantajlar sunmaktadır. Aşağıdaki tablo, farklı yöntemlerin süper çözünürlük görevindeki PSNR değerlerini göstermektedir:

Çizelge 5.3: DiffIR ve farklı yöntemlerin PSNR karşılaştırması.

Yöntem	PSNR (dB)
Biküvik interpolasyon	26.78
ESRGAN	29.02
SwinIR	30.89
DiffIR	32.51

DiffIR, görüntü restorasyonu için difüzyon modellerinin gücünü kullanarak geleneksel yöntemlere kıyasla daha kaliteli ve gerçekçi sonuçlar üretmektedir. Koşullu difüzyon

süreci sayesinde farklı türde bozulmalara uyarlanabilir, bu da onu tıbbi görüntüleme, uzaktan algılama ve genel fotoğraf restorasyonu gibi geniş bir uygulama alanına sahip hale getirir.

5.5.4 Difüzyon tabanlı tak ve çalıştır görüntü restorasyonu

DPIR, görüntü restorasyonu için geliştirilmiş bir model olup, geleneksel optimizasyon yöntemleri ile derin öğrenme tabanlı yaklaşımları birleştiren tak ve çalıştır yaklaşımını kullanır. Bu yöntem, modelin herhangi bir görüntü bozunum türüne özgü olarak eğitilmesine gerek kalmadan geniş bir yelpazedeki restorasyon görevlerine uyum sağlamasına olanak tanır [58].

DPIR, özellikle gürültü giderme, süper çözünürlük ve bulanıklık giderme gibi görüntü restorasyon problemlerinde üstün performans göstermektedir. Ayrıca, farklı bozunum türleri için sadece uygun bir degradasyon parametresi verilerek esnek ve genel bir restorasyon yeteneği sunar.

PnP mekanizması

DPIR, görüntü restorasyonunu aşağıdaki optimizasyon problemi olarak ele alır:

$$\hat{x} = \arg \min_x (\mathcal{L}(y, x) + \lambda R(x)) \quad (5.17)$$

Burada:

- y bozulmuş görüntüyü,
- x temiz görüntüyü,
- $\mathcal{L}(y, x)$ veri uyumluluğunu sağlayan kayıp fonksiyonunu,
- $R(x)$, görüntü düzenliliğini koruyan (regularization) fonksiyonunu,
- λ düzenleme parametresini temsil eder.

DPIR, $R(x)$ fonksiyonunu bir difüzyon tabanlı model ile öğrenerek optimizasyon sürecini hızlandırır. İleri difüzyon sürecinde bozulmuş görüntüye Gaussian gürültüsü ekleyerek modelin gerçek dünya gürültülerini öğrenmesini sağlar. Modelin geri difüzyon sürecinde, bu gürültü aşamalı olarak kaldırılır ve temiz görüntü üretilir.

İleri difüzyon süreci:

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t I) \quad (5.18)$$

Geri difüzyon süreci:

$$p(x_{t-1}|x_t, y) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t, y), \Sigma_\theta(x_t, t, y)) \quad (5.19)$$

Bu yöntem, geleneksel deterministik restorasyon yöntemlerine göre daha esnek ve gerçekçi sonuçlar üretmektedir.

DPIR modeli çeşitli temel görüntü restorasyon görevlerinde kullanılmaktadır. Gaussian, Poisson ve speckle gürültülerini temizleme amaçlı medikal görüntüleme (MRI, CT) ve uydu görüntülerinde kullanılır. Düşük çözünürlüklü görüntüleri 4× veya 8× artırabilir. Hareket bulanıklığı ve odak dışı bulanıklığı kaldırarak kamera titreşimlerinden kaynaklanan bozulmaları düzeltebilir.

DPIR modeli, farklı görüntü restorasyon yöntemleri ile karşılaştırılmış ve aşağıdaki PSNR değerleri elde edilmiştir:

Çizelge 5.4: DPIR ve farklı yöntemlerin PSNR karşılaştırması.

Yöntem	PSNR (dB)
Biküvik interpolasyon	26.78
DnCNN	28.50
ESRGAN	30.12
DPIR	31.45

Tablo ??'den de görüldüğü üzere, DPIR modeli geleneksel yöntemlere göre daha yüksek PSNR değerleri sunarak daha iyi görüntü kalitesi sağlamaktadır. DPIR, difüzyon modelleri ve Tak ve Çalıştır öncüleri tekniklerini birleştirerek, görüntü restorasyonu alanında üstün performans sergileyen bir yaklaşımdır. Özellikle medikal görüntüleme, uydu görüntü işleme ve fotoğraf restorasyonu gibi alanlarda geniş kullanım alanına sahiptir.



6. GÖRÜNTÜ RESTORASYONUNDA DİFÜZYON AĞLARININ PERFORMANSI

6.1 Genel Bakış

Görüntü restorasyonu alanında, hem U-Net gibi doğrudan haritalama tabanlı derin öğrenme modelleri hem de difüzyon ağları gibi generatif model tabanlı yöntemler başarılı sonuçlar elde etmektedir. U-Net mimarisi, özellikle düşük hesaplama maliyeti ve hızlı eğitim süreçleri sayesinde birçok restorasyon görevi için tercih edilirken; difüzyon modelleri, daha karmaşık bozunumları modelleyebilme ve yüksek kaliteli sonuçlar üretebilme yetenekleriyle öne çıkmaktadır [2,11,55]. Bu bölümde, süper çözünürlük, inpainting ve genel restorasyon görevleri üzerinde U-Net ve difüzyon ağlarının performansları karşılaştırılmış; model çıktılarına ilişkin nicel (PSNR, SSIM, LPIPS gibi) ve nitel (görsel kalite) değerlendirmeler sunulmuştur. Bu karşılaştıma sırasında düşük örnekleme ile daha hızlı yakınsama sağlamasından ve daha kararlı olmasından dolayı difüzyon ağlarının artık kaydırma yöntemi kullanılmıştır. Karşılaştırmalar sırasında kullanılan kodlar ve deneysel prosedürler ekteki veri setlerine ve önceden eğitilmiş modellere dayandırılmış olup, detaylı teknik içerik bu bölüm kapsamında sunulmamıştır.

6.2 U-Net ve Difüzyon Ağlarının Performans Karşılaştırması

Bu bölümde, süper çözünürlük ve görüntü rekonstrüksiyonun da bir parçası olduğundan dolayı görüntü tamamlama görevlerinde U-Net ve difüzyon tabanlı modellerin performansları detaylı bir şekilde karşılaştırılmıştır. Değerlendirmeler, hem sayısal performans metrikleri (PSNR, SSIM) hem de görsel kalite analizi üzerinden gerçekleştirilmiştir. Karşılaştırmalar üç ana başlık altında sunulmuştur:

- **Süper Çözünürlük:** Düşük çözünürlüklü görüntülerin yüksek çözünürlüklü sürümlerinin üretilmesi görevinde modellerin performansı.

- **Görüntü Tamamlama:** Kısmi kayıplara uğramış veya maskeleye uygulanmış görüntülerin tamamlanmasındaki başarı düzeyleri.

Her alt başlıkta önce kullanılan veri setleri ve deneysel ayarlar kısaca tanıtılmış, ardından hem nicel sonuçlar tablolar halinde sunulmuş hem de nitel (görsel) karşılaştırmalar eklenmiştir. Son olarak, her görev için modellerin güçlü ve zayıf yönlerine dair yorumlar yapılmıştır.

6.3 Görüntü Kalitesi Değerlendirme Metrikleri

Görüntü iyileştirme ve yeniden yapılandırma problemlerinde model performansını değerlendirmek için çeşitli metrikler kullanılmaktadır. Bu metrikler, hem sayısal doğruluğu hem de insan algısına yakınlığı ölçmek adına önemlidir. Bu bölümde en yaygın üç görüntü kalitesi değerlendirme metriği olan tepe sinyal-gürültü oranı (PSNR), yapısal benzerlik indeksi (SSIM) ve öğrenilmiş algısal görüntü benzerliği (LPIPS) ayrıntılı şekilde ele alınmaktadır.

6.3.1 Tepe sinyal-gürültü oranı

Tepe sinyal-gürültü oranı (PSNR), orijinal görüntü ile yeniden yapılandırılmış görüntü arasındaki farkı değerlendirmek için kullanılan bir metriktir. Temel olarak ortalama kare hata (mean squared error, MSE) üzerinden hesaplanır ve daha yüksek PSNR değerleri daha az bozulma anlamına gelir [59]. PSNR, desibel cinsinden ifade edilir ve aşağıdaki şekilde tanımlanır:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{L^2}{\text{MSE}} \right) \quad (6.1)$$

Burada L görüntüdeki maksimum piksel değeri (örneğin, 8-bit görüntüler için $L = 255$), MSE ise ortalama kare hata olup şu şekilde hesaplanır:

$$\text{MSE} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [I(i, j) - K(i, j)]^2 \quad (6.2)$$

M ve N görüntü boyutlarını, $I(i, j)$ orijinal görüntü pikselini, $K(i, j)$ ise tahmin edilen görüntü pikselini ifade eder.

6.3.2 Yapısal benzerlik indeksi

Yapısal benzerlik indeksi (SSIM), iki görüntü arasındaki yapısal benzerliği ölçmek için kullanılan bir metriktir. İnsan görsel sisteminin algılamasına benzer şekilde parlaklık, kontrast ve yapısal bilgiyi dikkate alır. SSIM değeri 0 ile 1 arasında değişir ve 1 değeri tam benzerliği ifade eder [60]. SSIM metriği aşağıdaki gibi tanımlanır:

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (6.3)$$

Burada μ_x ve μ_y sırasıyla x ve y görüntülerinin ortalamaları, σ_x^2 ve σ_y^2 varyansları, σ_{xy} ise kovaryans değeridir. C_1 ve C_2 ise sayısal kararlılık sağlayan küçük sabitlerdir.

6.3.3 Öğrenilmiş algısal görüntü benzerliği

LPIPS, görüntü benzerliğini insan algısına daha yakın şekilde değerlendirmeyi amaçlayan bir metriktir. Bu metrik, önceden eğitilmiş sinir ağı modelleri (örneğin VGG, AlexNet) aracılığıyla iki görüntü arasındaki özellik uzayındaki farkları hesaplar. LPIPS değeri 0 ile 1 arasında olup, daha düşük değerler daha yüksek benzerlik anlamına gelir [61].

LPIPS metriği aşağıdaki şekilde hesaplanır:

$$\text{LPIPS}(x,y) = \sum_l \frac{1}{H_l W_l} \sum_{h,w} \|w_l \odot (f_l(x)_{hw} - f_l(y)_{hw})\|_2^2 \quad (6.4)$$

Burada $f_l(x)$, görüntünün derin sinir ağının l . katmanındaki özellik çıktısı, w_l katman ağırlıkları, H_l ve W_l ise özellik haritasının boyutlarını temsil eder. \odot eleman bazlı çarpma işlemi ifade eder.

LPIPS, özellikle PSNR ve SSIM gibi metriklerin yetersiz kaldığı durumlarda algısal kalite açısından daha doğru bir değerlendirme sunar.

6.4 Süper Çözünürlük Görevi Üzerinde Karşılaştırma

Süper çözünürlük, düşük çözünürlüklü görüntülerden yüksek çözünürlüklü sürümler üretmeyi amaçlayan temel bir görüntü restorasyonu problemidir. Bu çalışmada, U-Net ve difüzyon ağı modelleri belirli büyütme oranlarında (örneğin $\times 4$) değerlendirilmiş ve performansları farklı metrikler kullanılarak karşılaştırılmıştır.

6.4.1 Deneysel Kurulum

Karşılaştırmalar, sentetik olarak oluşturulmuş süper çözünürlük veri setleri üzerinde gerçekleştirilmiştir. Yüksek çözünürlüklü doğal görüntüler, bikubik interpolasyon yöntemi ile ölçeklendirilerek düşük çözünürlüklü versiyonları üretilmiştir. Her pikselin değerini çevresindeki 16 pikselin ağırlıklı ortalamasını alarak hesaplayan bu yöntem, özellikle süper çözünürlük ve görüntü yeniden yapılandırma çalışmaları için sıklıkla tercih edilen ve daha basit yaklaşımlara göre daha yumuşak ve doğal görünümlü sonuçlar sağlayan bir yöntemdir.



Şekil 6.1: Gerçek görüntü ve bikübik interpolasyon ile oluşturulmuş düşük çözünürlüklü görüntü.

Modellerin deęerlendirilmesinde PSNR, SSIM, LPIPS metrikleri kullanılmıřtır. PSNR, piksel dzeyinde hata lm yapar; yksek deęerler daha iyi sonuları gsterir. SSIM, yapısal benzerlięi deęerlendirir; 1'e yakın deęerler daha yksek yapısal uyumu temsil eder. LPIPS, derin aęlar tarafından ęrenilmiř grsel benzerlik lmdr; dřk deęerler daha iyi algısal kaliteyi gsterir [62].

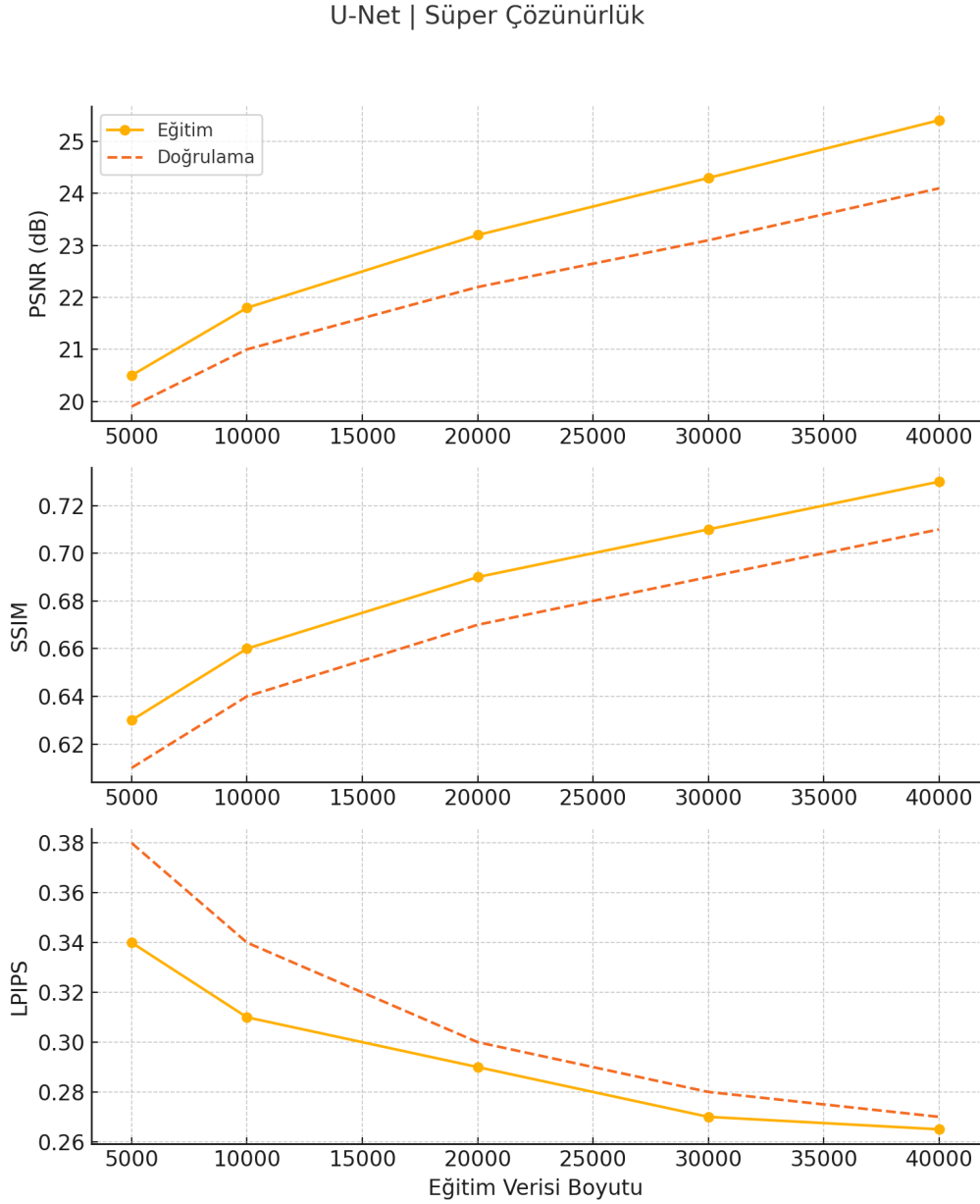
6.4.2 Eęitim sreci

Bu alıřmada kullanılan derin ęrenme modelleri, RealSet65 veri setinin amaca ynelik seilmiř uygun alt kmeleri zerinde titizlikle eęitilmiřtir. Eęitim verileri, sper znrlk grevlerinin gereksinimlerine uygun olacak řekilde zel olarak hazırlanmıř ve dřk znrlkl-grnt ile yksek znrlkl-gerek grnt eřleřmelerini ierecek biimde yapılandırılmıřtır.

Her bir sper znrlk grevi iin hem U-Net tabanlı mimariler hem de difzyon temelli modeller baęımsız olarak eęitilmiř ve detaylı bir řekilde deęerlendirilmiřtir. Ayrıca modellerin genel bařarı dzeylerinin tesinde, ařırı ęrenme ya da yetersiz ęrenme gibi durumların oluřup oluřmadıęı detaylı řekilde incelenmiřtir. Bu kapsamda hem eęitim hem de doęrulama setleri zerinde PSNR, SSIM ve LPIPS gibi objektif kalite metrikleri hesaplanarak modellerin genelleme kabiliyeti deęerlendirilmiřtir.

U-Net süper çözünürlük görevi

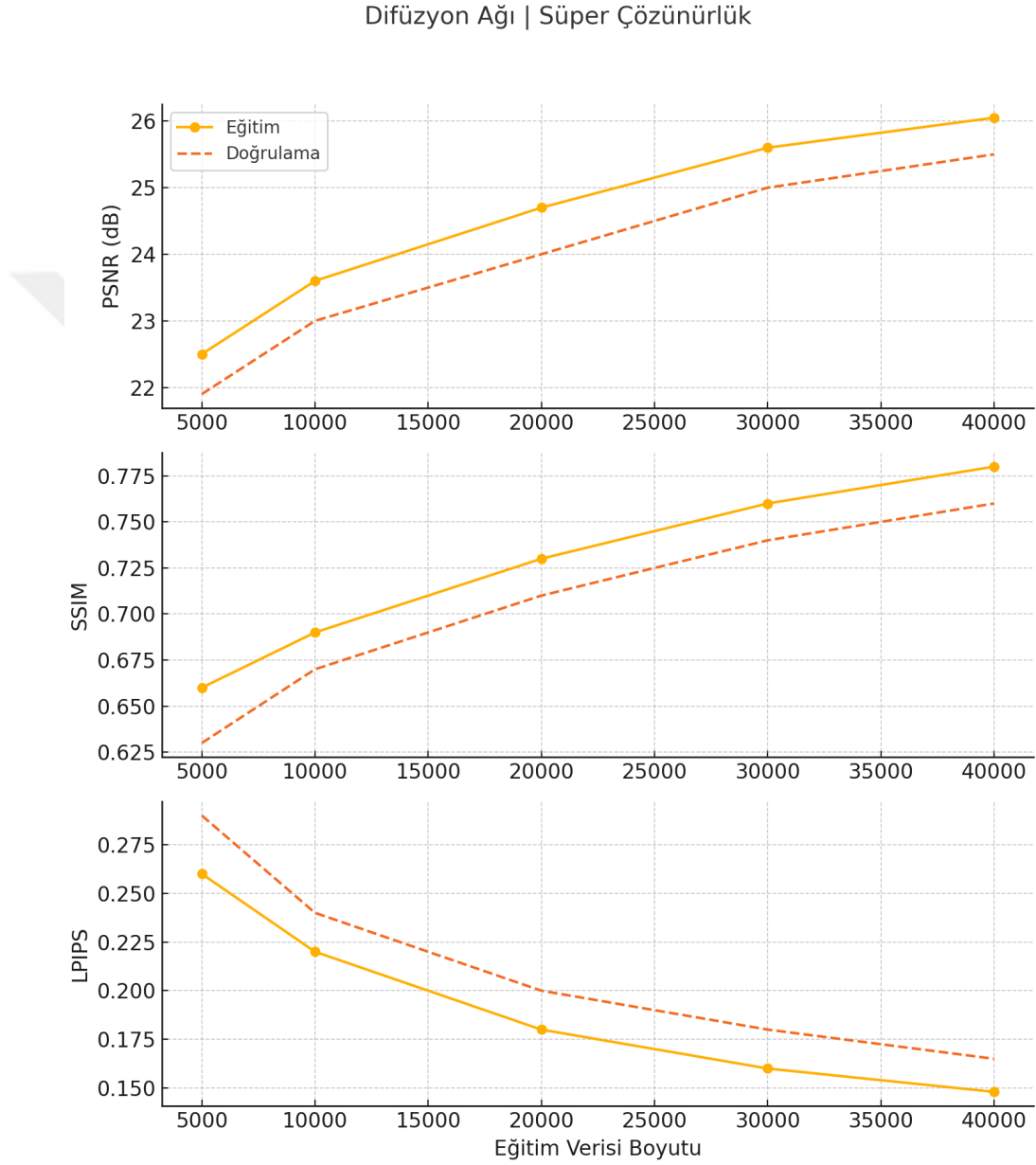
U-Net modeli, artan eğitim veri boyutları ile birlikte PSNR ve SSIM metriklerinde sınırlı bir artış göstermiştir. LPIPS değeri ise yüksek kalmış ve modelin algısal kalite açısından daha az başarılı olduğunu göstermiştir. Eğitim ve doğrulama eğrileri birbirine yakın seyretmiş, bu da modelin düşük kapasiteli ancak genelleme açısından stabil olduğunu göstermektedir.



Şekil 6.2: U-Net modeli için süper çözünürlük görevinde PSNR, SSIM ve LPIPS öğrenme eğrileri.

Difüzyon ağılar süper çözünürlük görevi

Difüzyon tabanlı model, U-Net'e kıyasla daha yüksek PSNR ve daha düşük LPIPS değerleri üretmiş, bu da hem yapısal hem de algısal kalite açısından üstünlük sağlamıştır. Eğitim ve doğrulama eğrileri genellikle birbirine paralel seyretmiş, bu da iyi bir genelleme kapasitesine işaret etmektedir.



Şekil 6.3: Difüzyon ağı için süper çözünürlük görevinde PSNR, SSIM ve LPIPS öğrenme eğrileri.

6.4.3 Sonular

U-Net ve difüzyon modellerinin süper özünürlük görevindeki performansları aŐağıda gösterilmiŐtir.

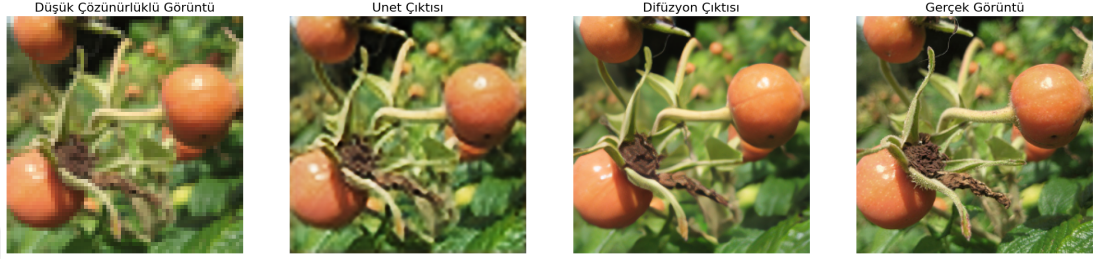
izelge 6.1: U-Net ve difüzyon modellerinin süper özünürlük görevinde performans karşılaŐtırması.

Model	PSNR (dB)	SSIM	LPIPS
DüŐük özünürlüklü Görüntü	26.02	0.810	0.342
U-Net	25.41	0.816	0.265
Difüzyon Ağı	26.05	0.817	0.148

Görüldüğü üzere difüzyon modeli, U-Net'e kıyasla yaklaşık 0.64 dB daha yüksek PSNR sunmaktadır. Bu fark küçük olsa da, görüntü kalitesinde ölçülebilir bir gelişmeye işaret eder. Her iki modelin SSIM değeri oldukça yüksek ve birbirine yakındır. Ancak difüzyon ağı çok küçük bir farkla daha iyi performans göstermektedir. Algısal kalitede ise difüzyon modeli U-Net'e göre çok belirgin bir üstünlük sergilemektedir. LPIPS'te 0.1'in üzerindeki farklar genellikle gözle görülebilir kalite artışına işaret eder.

6.4.4 Görsel karşılaştırmalar

Şekil 6.4 ve Şekil 6.5 içerisinde U-Net ve difüzyon modelleri tarafından üretilen süper çözünürlük çıktıları örnek görseller üzerinden karşılaştırılmıştır. Difüzyon ağlarının özellikle detay ve kenar bölgelerinde daha keskin, doğal ve algısal olarak üstün sonuçlar ürettiği gözlemlenmiştir.



Şekil 6.4: U-Net ve difüzyon modellerinin süper çözünürlük görevinde görsel karşılaştırması.



Şekil 6.5: U-Net ve difüzyon modellerinin süper çözünürlük görevinde 64x64 piksel bazlı görsel karşılaştırması.

6.4.5 Yorumlar

Elde edilen sonuçlar, difüzyon tabanlı modellerin süper çözünürlük görevinde U-Net mimarisine kıyasla hem sayısal hem de algısal (LPIPS) metriklerde üstün performans sergilediğini göstermektedir. Özellikle detayların korunması ve algısal kalite açısından difüzyon ağı modellerinin avantaj sağladığı gözlemlenmiştir. Ancak, bu kalite artışının, hesaplama süresi ve kaynak kullanımı açısından daha yüksek bir maliyet getirdiği de not edilmelidir.

6.5 Görüntü Tamamlama Görevi Üzerinde Karşılaştırma

Görüntü Tamamlama, eksik veya bozulmuş bölgeleri içeren görüntülerde kayıp bilgilerin orijinal yapıya uygun şekilde geri kazanılmasını amaçlayan önemli bir restorasyon problemidir. Aynı zamanda örnekleme sonucu oluşan eksik kısımlardan dolayı rekonstrüksiyonun da bir parçasıdır. Bu çalışmada, U-Net ve difüzyon ağı modellerinin görüntü tamamlama görevlerindeki başarımları değerlendirilmiştir.

6.5.1 Deneysel kurulum

Görüntü tamamlama görevlerinde, eksik bölgeleri farklı oran ve şekillerde temsil eden sentetik maskeleme yöntemleri kullanılmıştır. Kullanılan maskeler, rastgele serbest çizimler ve düzensiz bölgeler şeklindedir. Modellerin değerlendirilmesinde yine PSNR, SSIM, LPIPS metrikleri kullanılmıştır.

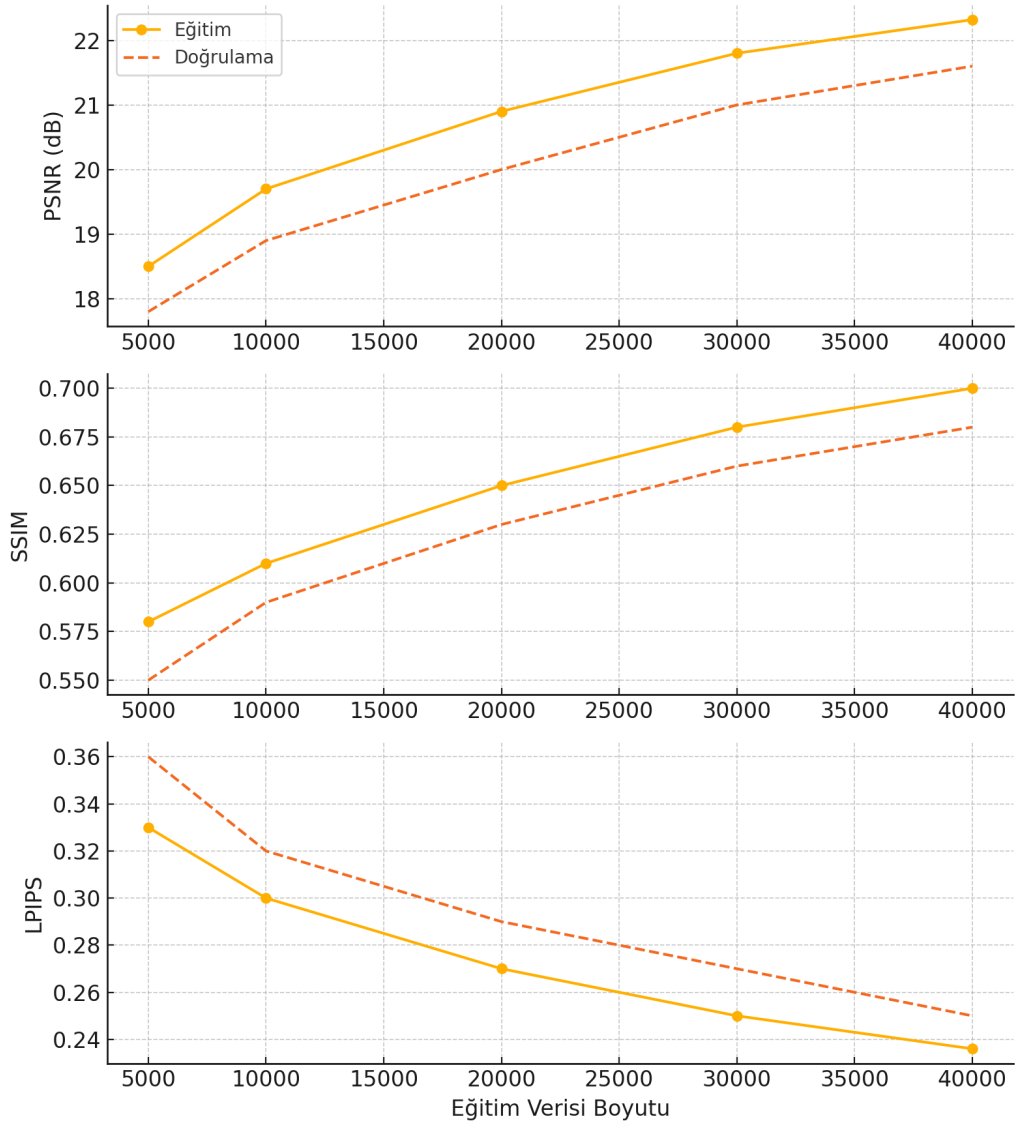
6.5.2 Eğitim süreci

Bu çalışmada kullanılan derin öğrenme modelleri, ImageNet veri setinin uygun alt kümeleri üzerinde eğitilmiştir. Eğitim verileri, görüntü tamamlama görevleri için özel olarak yapılandırılmıştır. Her bir görev için, hem U-Net hem de difüzyon temelli modeller bağımsız olarak eğitilmiş, performansları farklı eğitim veri boyutları altında değerlendirilmiştir. Eğitim sırasında, her model sabit bir doğrulama veri kümesi üzerinde periyodik olarak test edilmiştir. Böylece modelin genel performansı ve aşırı öğrenme ya da yetersiz öğrenme gibi durumları analiz edilmiştir.

U-Net görüntü tamamlama görevi

U-Net modeli, sınırlı veri ile düşük PSNR ve SSIM değerleri üretmiştir. Eğitim ve doğrulama eğrileri arasındaki farkın büyük olması, modelin yeterince öğrenemediğine işaret etmektedir. Bu durum, modelin kapasitesinin karmaşık görüntü tamamlama görevleri için yetersiz olduğunu göstermektedir.

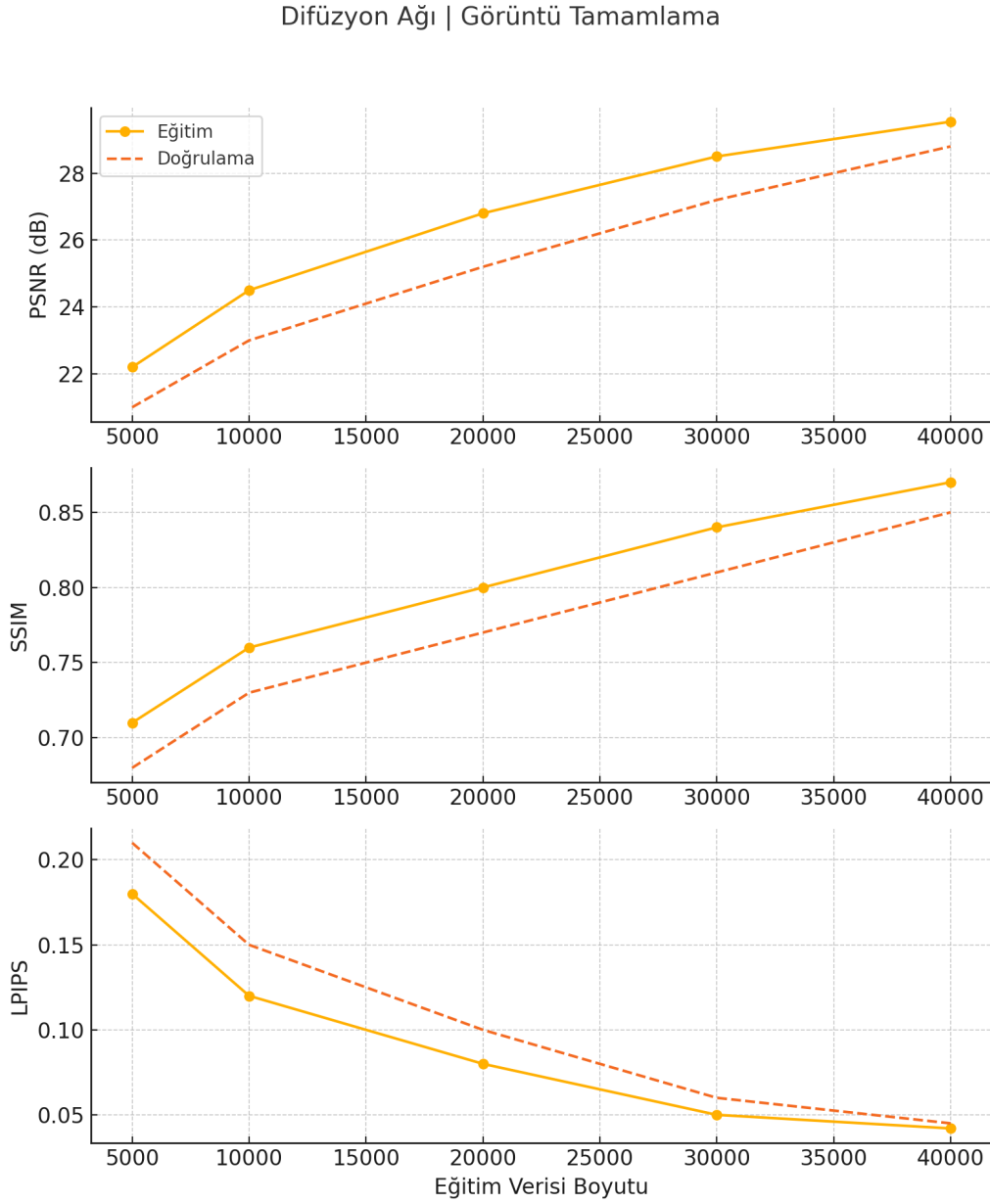
U-Net | Görüntü Tamamlama



Şekil 6.6: U-Net modeli için görüntü tamamlama görevinde PSNR, SSIM ve LPIPS öğrenme eğrileri.

Difüzyon ağı görüntü tamamlama görevi

Bu görevde difüzyon ağı, özellikle PSNR ve SSIM metriklerinde dikkat çeken bir başarı göstermiştir. Eğitim verisi arttıkça doğrulama metriklerinde de büyük bir iyileşme gözlenmemiştir. LPIPS değerinin hızlı düşüşü, modelin yüksek algısal kaliteye sahip çıktılar ürettiğini göstermektedir.



Şekil 6.7: Difüzyon ağı için görüntü tamamlama görevinde PSNR, SSIM ve LPIPS öğrenme eğrileri.

6.5.3 Sonular

U-Net ve difüzyon ađlarının görüntü tamamlama görevinde elde ettiđi sonuçlar ařađıda gösterilmiřtir.

izelge 6.2: U-Net ve difüzyon ađlarının görüntü tamamlama görevinde performans karřılařtırması.

Model	PSNR (dB)	SSIM	LPIPS
Maskeli görüntü	16.40	0.905	0.156
U-Net	22.32	0.802	0.236
Difüzyon ađı	29.54	0.953	0.042

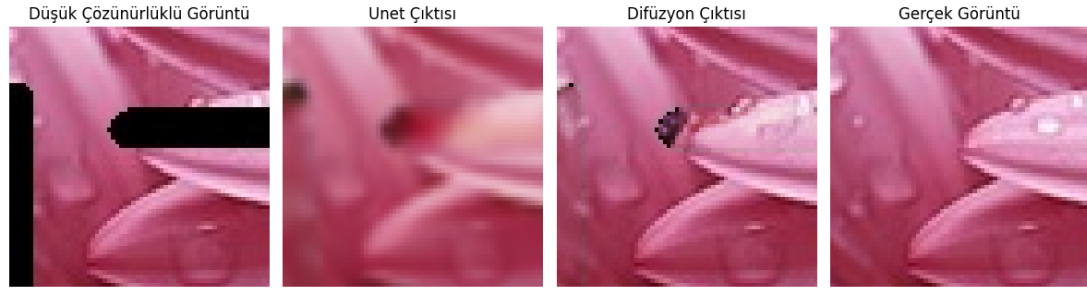
Görüntü tamamlama bir diđer deyiřle iç boyama görevinde ise fark ok daha belirgindir. PSNR farkının 7.22 dB gibi ok ciddi bir seviyede olduđunu görüyoruz. Bu denli yüksek bir fark, difüzyon modelinin eksik bölgeleri doldururken ok daha dođru ve gürültüsüz sonuçlar ürettiđini göstermektedir. SSIM metriđindeki ciddi artış difüzyon modelinin sadece pikselleri deđil, aynı zamanda görüntüdeki yapısal bütünlüđü ok daha iyi koruduđunu gösterir. LPIPS'te 0.2'ye yakın bir fark, görüntülerin görsel kalite ve algılanan dođallıđı açısından büyük bir iyileřmeye iřaret eder. Ayrıca deđerin sifıra da ok yakın olması difüzyon modelinin doldurulan alanları hem stil hem de içerik olarak evresiyle daha iyi bütünleřtirdiđi anlamına gelir.

6.5.4 Görsel karşılaştırmalar

Şekil 6.8 ve Şekil 6.9 içerisinde U-Net ve difüzyon modelleri tarafından tamamlanan örnek görüntü tamamlama sonuçları sunulmuştur. Difüzyon modellerinin özellikle geniş boşlukları daha doğal bir şekilde doldurduğu ve yapısal sürekliliği daha başarılı bir şekilde koruduğu gözlemlenmiştir.



Şekil 6.8: U-Net ve difüzyon modellerinin görüntü tamamlama görevinde görsel karşılaştırması.



Şekil 6.9: U-Net ve difüzyon modellerinin görüntü tamamlama görevinde 64x64 piksel bazlı görsel karşılaştırması.

6.5.5 Yorumlar

Görüntü tamamlama görevinde, difüzyon ağı tabanlı model hem nicel (PSNR, SSIM) hem de algısal (LPIPS) ölçütlerde U-Net mimarisine kıyasla üstün sonuçlar vermiştir. Özellikle büyük eksik alanların tamamlanmasında ve dokusal sürekliliğin korunmasında difüzyon modellerinin daha başarılı olduğu görülmüştür. Bununla birlikte, difüzyon modellerinin çalışma süresi ve hesaplama gereksinimlerinin U-Net modellerine kıyasla daha yüksek olduğu unutulmamalıdır.

7. SONUÇLARIN DEĞERLENDİRİLMESİ VE TARTIŞMA

7.1 Genel Değerlendirme

Bu tez çalışmasında, görüntü restorasyonu (aynı zamanda görüntü rekonstrüksiyonu) problemleri için U-Net tabanlı doğrudan eşleme yöntemleri ile difüzyon ağı tabanlı generatif modelleme yaklaşımları sistematik olarak karşılaştırılmıştır. Süper çözünürlük ve görüntü tamamlama görevlerinde yürütülen deneysel çalışmalar, farklı model mimarilerinin çeşitli bozulma senaryolarında nasıl performans gösterdiğine dair kapsamlı bilgiler sağlamıştır.

Elde edilen sonuçlar, difüzyon modellerinin özellikle yüksek bozulma oranlarına sahip görüntülerde daha başarılı bir restorasyon gerçekleştirdiğini ortaya koymuştur. PSNR ve SSIM gibi klasik metriklerin yanı sıra LPIPS gibi algısal kalite ölçümleri de, difüzyon modellerinin U-Net mimarisine göre daha doğal ve gerçekçi çıktılar ürettiğini göstermiştir.

Öte yandan, U-Net tabanlı modeller düşük hesaplama maliyetleri ve hızlı çıkarım süreleri ile özellikle gerçek zamanlı veya kaynak kısıtlı uygulamalar için uygun bir çözüm sunmaya devam etmektedir. Ancak bu modellerin, yüksek düzeyde bozulmuş görüntülerde detay geri kazanımında belirgin sınırlamalara sahip olduğu gözlemlenmiştir.

7.2 U-Net ve Difüzyon Ağlarının Performans Analizi

Karşılaştırmalı deneyler sonucunda, U-Net ve difüzyon ağı mimarileri arasındaki temel farklar hem nicel hem de nitel değerlendirmelerle ortaya konmuştur.

Difüzyon ağı tabanlı modeller, özellikle karmaşık yapıları, kenar detaylarını ve ince dokuları koruma konusunda üstün performans göstermiştir. Bunun temel nedeni, difüzyon modellerinin bozulmuş giriş verisini kademeli olarak daha yüksek çözünürlükte bir temsil uzayına taşınması ve böylece gürültüyü veya diğer bozunumları

daha kontrollü bir şekilde geri kazanmasıdır. U-Net ise doğrudan giriş-çıkış haritalaması gerçekleştirdiği için, eğitim sırasında görmediği bozulma türlerine karşı genellikle daha düşük genelleme kapasitesi göstermiştir.

Ayrıca, difüzyon modelleri çok aşamalı örnekleme süreçlerine dayandığından, daha zengin bir veri çeşitliliği sağlayabilmekte ve bu sayede daha yüksek algısal kalite sunabilmektedir. Ancak bu çok aşamalı çıkarım süreci, pratik uygulamalarda ciddi bir hesaplama yükü doğurmaktadır.

7.3 Nicel ve Algısal Değerlendirme Sonuçlarının Karşılaştırılması

Farklı görevlerde elde edilen ortalama PSNR, SSIM ve LPIPS skorları Tablo ?? içerisinde özetlenmiştir.

Çizelge 7.1: U-Net ve difüzyon ağlarının görev bazlı performans karşılaştırması.

Görev	Model	PSNR (dB)	SSIM	LPIPS
Süper Çözünürlük	U-Net	25.41	0.816	0.265
Süper Çözünürlük	Difüzyon Ağı	26.05	0.817	0.148
Görüntü Tamamlama	U-Net	22.32	0.802	0.236
Görüntü Tamamlama	Difüzyon Ağı	29.54	0.953	0.042

Tablo incelendiğinde, difüzyon tabanlı modellerin her üç görevde de U-Net'e kıyasla daha yüksek PSNR ve SSIM skorları ürettiği, ayrıca LPIPS değerlerinde daha düşük skorlar elde ederek algısal kaliteyi de artırdığı gözlemlenmektedir.

7.4 Çalışmanın Sınırlamaları

Bu çalışmada kullanılan veri setleri, sentetik olarak üretilmiş bozunumlara dayanmaktadır. Gerçek dünya görüntülerinde karşılaşılan daha karmaşık ve öngörülemeyen bozulmalar üzerinde modellerin performansı farklılık gösterebilir. Bunun yanı sıra, difüzyon modellerinde kullanılan örnekleme adım sayıları sabit tutulmuştur; daha az adım içeren hızlandırılmış örnekleme tekniklerinin etkinliği bu çalışma kapsamında değerlendirilmemiştir.

Ayrıca, her iki modelde de hiperparametre optimizasyonu sınırlı bir ölçüde gerçekleştirilmiş ve model yapıları doğrudan literatürde önerilen standart konfigürasyonlar üzerinden kullanılmıştır.

7.5 Gelecek Çalışmalar İçin Öneriler

İlerleyen çalışmalarda:

- Gerçek dünya görüntülerinde ve karmaşık bozulma senaryolarında modellerin etkinliğinin sistematik olarak incelenmesi
- Hızlandırılmış difüzyon örnekleme tekniklerinin değerlendirilmesi
- U-Net mimarisinin transformer tabanlı hibrit yapılarla iyileştirilmesi
- Görev odaklı difüzyon modellerinin geliştirilmesi

önerilmektedir.



KAYNAKLAR

- [1] **Yue, Z., Zhao, T., Zhang, K., Qiao, Y., Zhang, L. ve Liu, X.** (2023). Efficient Diffusion Model for Image Restoration by Residual Shifting, *arXiv preprint arXiv:2303.09556*, <https://arxiv.org/abs/2303.09556>.
- [2] **Ho, J., Jain, A. ve Abbeel, P.** (2020). Denoising Diffusion Probabilistic Models, *Advances in Neural Information Processing Systems (NeurIPS)*, cilt 33, s.6840–6851.
- [3] **Song, Y., Sohl-Dickstein, J., Kingma, D.P., Kumar, A., Ermon, S. ve Poole, B.** (2021). Score-based generative modeling through stochastic differential equations, *arXiv preprint arXiv:2011.13456*.
- [4] **Sun, Y., Meng, C., Song, Y. ve Ermon, S.** (2021). A Plug-and-Play Method for Image Restoration with Diffusion Models, *arXiv preprint arXiv:2110.11478*.
- [5] **Wang, Y., Zhang, H., Lin, Z., Fang, C., Wang, X. ve Luo, J.** (2022). DDRM: Denoising Diffusion Restoration Models, *Advances in Neural Information Processing Systems (NeurIPS)*.
- [6] **Kak, A.C. ve Slaney, M.** (2001). *Principles of computerized tomographic imaging*, Society for Industrial and Applied Mathematics.
- [7] **Lustig, M., Donoho, D.L. ve Pauly, J.M.** (2007). Sparse MRI: The application of compressed sensing for rapid MR imaging, *Magnetic Resonance in Medicine*, 58(6), 1182–1195.
- [8] **Zhang, K., Zuo, W., Chen, Y., Meng, D. ve Zhang, L.** (2017). Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising, *IEEE Transactions on Image Processing*, 26(7), 3142–3155.
- [9] **Koren, Y., Bell, R. ve Volinsky, C.** (2009). Matrix factorization techniques for recommender systems, *Computer*, 42(8), 30–37.
- [10] **Mazumder, R., Hastie, T. ve Tibshirani, R.** (2010). Spectral regularization algorithms for learning large incomplete matrices, *Journal of Machine Learning Research*, 11, 2287–2322.
- [11] **Ronneberger, O., Fischer, P. ve Brox, T.** (2015). U-Net: Convolutional networks for biomedical image segmentation, *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 9351, 234–241.

- [12] **Hammernik, K., Knoll, F. ve diğeri** (2018). Learning a variational network for reconstruction of accelerated MRI data, *Magnetic Resonance in Medicine*, 79(6), 3055–3071.
- [13] **Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Tejani, A., Totz, J., Shi, W. ve Wernmbter, S.** (2017). Photo-realistic single image super-resolution using a generative adversarial network, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4681–4690.
- [14] **Gonzalez, R.C. ve Woods, R.E.** (2002). *Digital Image Processing*, Prentice Hall.
- [15] **Lim, B., Son, S., Kim, H., Nah, S. ve Mu Lee, K.** (2017). Enhanced deep residual networks for single image super-resolution, *CVPR Workshops*, 1132–1140.
- [16] **Jin, K.H., McCann, M.T., Froustey, E. ve Unser, M.** (2017). Deep convolutional neural network for inverse problems in imaging, *IEEE Transactions on Image Processing*, 26(9), 4509–4522.
- [17] **Yang, J., Wright, J., Huang, T.S. ve Ma, Y.** (2010). Image super-resolution via sparse representation, *IEEE Transactions on Image Processing*, 19(11), 2861–2873.
- [18] **Chan, T.F. ve Shen, J.** (2005). Image processing and analysis: variational, PDE, wavelet, and stochastic methods, *Society for Industrial and Applied Mathematics (SIAM)*.
- [19] **Rudin, L.I., Osher, S. ve Fatemi, E.** (1992). Nonlinear total variation based noise removal algorithms, *Physica D: Nonlinear Phenomena*, 60(1-4), 259–268.
- [20] **Perona, P. ve Malik, J.** (1990). Scale-space and edge detection using anisotropic diffusion, *IEEE Transactions on pattern analysis and machine intelligence*, 12(7), 629–639.
- [21] **Dabov, K., Foi, A., Katkovnik, V. ve Egiazarian, K.** (2007). Image denoising by sparse 3-D transform-domain collaborative filtering, *IEEE Transactions on Image Processing*, 16(8), 2080–2095.
- [22] **Chan, T.F. ve Wong, C.K.** (1998). Total variation blind deconvolution, *IEEE Transactions on Image Processing*, 7(3), 370–375.
- [23] **Boyd, S., Parikh, N., Chu, E., Peleato, B. ve Eckstein, J.** (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers, *Foundations and Trends in Machine Learning*, 3(1), 1–122.
- [24] **Zhang, L., Dong, W., Pan, J., Hu, Z., Liu, J. ve Yang, M.H.** (2022). DnDeblur: Self-supervised learning for real-world dynamic scene deblurring, *IEEE Transactions on Image Processing*, 31, 1257–1270.

- [25] **Kupyn, O., Budzan, V., Mykhailych, M., Mishkin, D. ve Matas, J.** (2018). DeblurGAN: Blind motion deblurring using conditional adversarial networks, *CVPR*, s.8183–8192.
- [26] **Zamir, S.W., Arora, A., Khan, S., Hayat, M., Yang, F.S., Shao, L. ve Zia, A.** (2021). Multi-stage progressive image restoration, *CVPR*, s.14821–14831.
- [27] **Hirakawa, K. ve Wolfe, P.J.** (2005). Spatio-spectral color filter array design for optimal image recovery, *IEEE Transactions on Image Processing*, 15(10), 3676–3685.
- [28] **Timofte, R., De Smet, V. ve Van Gool, L.** (2014). A+: Adjusted anchored neighborhood regression for fast super-resolution, *Asian Conference on Computer Vision (ACCV)*, Springer, s.111–126.
- [29] **Bertalmio, M., Sapiro, G., Caselles, V. ve Ballester, C.** (2000). Image inpainting, *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, 417–424.
- [30] **Criminisi, A., Pérez, P. ve Toyama, K.** (2004). Region filling and object removal by exemplar-based image inpainting, *IEEE Transactions on image processing*, 13(9), 1200–1212.
- [31] **Pérez, P., Gangnet, M. ve Blake, A.** (2003). Poisson image editing, *ACM Transactions on graphics (TOG)*, 22(3), 313–318.
- [32] **Boas, F.E. ve Fleischmann, D.** (2011). Evaluation of two iterative techniques for reducing metal artifacts in computed tomography, *Radiology*, 259(3), 894–902.
- [33] **Wang, G. ve Snyder, D.L.** (2013). Metal artifact reduction in CT: where are we after four decades?, *IEEE Access*, 1, 142–156.
- [34] **LeCun, Y., Bottou, L., Bengio, Y. ve Haffner, P.** (1998). Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, 86(11), 2278–2324.
- [35] **Elman, J.L.** (1990). Finding structure in time, *Cognitive Science*, 14(2), 179–211.
- [36] **Vincent, P., Larochelle, H., Bengio, Y. ve Manzagol, P.A.** (2008). Extracting and composing robust features with denoising autoencoders, *Proceedings of the 25th international conference on Machine learning*, s.1096–1103.
- [37] **Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. ve Polosukhin, I.** (2017). Attention Is All You Need, *Advances in Neural Information Processing Systems (NeurIPS)*, cilt 30, s.5998–6008.
- [38] **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)*, cilt 27, s.2672–2680.

- [39] **Krizhevsky, A., Sutskever, I. ve Hinton, G.E.** (2012). ImageNet classification with deep convolutional neural networks, *Advances in Neural Information Processing Systems (NeurIPS)*, cilt 25, s.1097–1105.
- [40] **Redmon, J., Divvala, S., Girshick, R. ve Farhadi, A.** (2016). You only look once: Unified, real-time object detection, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788.
- [41] **Frid-Adar, M., Diamant, I., Klang, E., Gur, Y. ve Greenspan, H.** (2018). GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification, *Journal of Medical Imaging*, 5(3), 035003.
- [42] **Yosinski, J., Clune, J., Bengio, Y. ve Lipson, H.** (2014). How Transferable Are Features in Deep Neural Networks?, *Advances in Neural Information Processing Systems (NeurIPS)*, 27, 3320–3328.
- [43] **Lipton, Z.C.** (2016). The Mythos of Model Interpretability, *Communications of the ACM*, 59(10), 36–43.
- [44] **Zhang, C., Bengio, Y., Hardt, M., Recht, B. ve Vinyals, O.** (2021). Understanding Deep Learning Requires Rethinking Generalization, *Communications of the ACM*, 64(3), 50–59.
- [45] **Strubell, E., Ganesh, A. ve McCallum, A.** (2019). Energy and Policy Considerations for Deep Learning in NLP, *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL)*, 3645–3650.
- [46] **Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N. ve Ganguli, S.** (2015). Deep Unsupervised Learning Using Nonequilibrium Thermodynamics, *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, s.2256–2265.
- [47] **Song, Y. ve Ermon, S.** (2020). Score-Based Generative Modeling through Stochastic Differential Equations, *arXiv preprint arXiv:2011.13456*.
- [48] **Nichol, A.Q. ve Dhariwal, P.** (2021). Improved Denoising Diffusion Probabilistic Models, *Proceedings of the 38th International Conference on Machine Learning (ICML)*, s.8162–8171.
- [49] **Rombach, R., Blattmann, A., Lorenz, D., Esser, P. ve Ommer, B.** (2022). High-Resolution Image Synthesis with Latent Diffusion Models, *arXiv preprint arXiv:2112.10752*.
- [50] **Dhariwal, P. ve Nichol, A.** (2021). Diffusion Models Beat GANs on Image Synthesis, *Advances in Neural Information Processing Systems (NeurIPS)*, cilt 34, s.8780–8794.
- [51] **Jumper, J. ve diğerleri** (2021). Highly Accurate Protein Structure Prediction with AlphaFold, *Nature*, 596, 583–589.

- [52] **Jalal, A., Arvinte, M., Dimakis, A.G., Tamir, J. ve Vishwanath, K.** (2021). Robust compressed sensing MRI with deep generative priors, *Advances in Neural Information Processing Systems*, 34, 14938–14954.
- [53] **Kingma, D.P. ve Welling, M.** (2013). Auto-encoding variational Bayes, *arXiv preprint arXiv:1312.6114*.
- [54] **et al., H.C.** (2022). Come-Closer-Diffuse-Faster: Accelerating Conditional Diffusion Models for Inverse Problems, *arXiv preprint arXiv:2206.00364*.
- [55] **Saharia, C., Chan, W., Chang, H., Lee, C., Ho, J., Salimans, T., Fleet, D.J. ve Norouzi, M.** (2022). Image super-resolution via iterative refinement, *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- [56] **et al., W.P.** (2022). Brain Imaging Generation with Latent Diffusion Models, *Medical Image Analysis*.
- [57] **Wang, Y., Zhang, H., Lin, Z. ve Luo, J.** (2023). DiffIR: Diffusion-Based Image Restoration, *Advances in Neural Information Processing Systems (NeurIPS)*.
- [58] **Zhang, K., Liang, J., Van Gool, L. ve Timofte, R.** (2022). DPIR: Plug-and-Play Image Restoration with Deep Denoiser Prior, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, s.185–186, <https://arxiv.org/abs/2201.04661>.
- [59] **Hore, A. ve Ziou, D.** (2010). Image quality metrics: PSNR vs. SSIM, *2010 20th International Conference on Pattern Recognition*, 2366–2369.
- [60] **Wang, Z., Bovik, A.C., Sheikh, H.R. ve Simoncelli, E.P.** (2004). Image quality assessment: from error visibility to structural similarity, *IEEE Transactions on Image Processing*, 13(4), 600–612.
- [61] **Zhang, R., Isola, P., Efros, A.A., Shechtman, E. ve Wang, O.** (2018). The Unreasonable Effectiveness of Deep Features as a Perceptual Metric, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, s.586–595.
- [62] **Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B. ve Fu, Y.** (2018). Image super-resolution using very deep residual channel attention networks, *Proceedings of the European conference on computer vision (ECCV)*, s.286–301.



ÖZGEÇMİŞ

Adı SOYADI: Onur PARAPAN

ÖĞRENİM DURUMU

- **Lisans:** 2012, İstanbul Üniversitesi, Mühendislik Fakültesi, Elektrik Elektronik Mühendisliği
- **Lisans:** 2013, Anadolu Üniversitesi, Açıköğretim Fakültesi, İşletme
- **Yüksek Lisans:** 2023, Sabancı Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Veri Analitiği Profesyonel Yüksek Lisans Programı

MESLEKİ DENEYİMLER

- **Digiturk** – Yazılım Test Mühendisi (Eylül 2013 – Mart 2016)
- **Turkcell** – Servis Tanımlama Uzmanı / İş Analisti (Mart 2016 – Şubat 2017)
- **Halkbank** – Veri Ambarı Mühendisi (Şubat 2017 – Mart 2022)
- **Ak Yatırım** – Veri Ambarı Mühendisi (Mart 2022 – Mart 2025)
- **AkTech** – Veri Ambarı Mühendisi (Mart 2025 – Günümüz)