T.C.
HARRAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

DİKKAT TABANLI OTOKODLAYICI KULLANARAK TEK GÖRÜNTÜDEN YÜKSEK DİNAMİK ARALIKLI GÖRÜNTÜ OLUŞTURMA

CEVHER RENK

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

Şanlıurfa
2025

T.C.
HARRAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

DİKKAT TABANLI OTOKODLAYICI KULLANARAK TEK GÖRÜNTÜDEN YÜKSEK DİNAMİK ARALIKLI GÖRÜNTÜ OLUŞTURMA

CEVHER RENK

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
Tez Danışmanı: Dr. Öğr. Üyesi SERDAR ÇİFTÇİ

Şanlıurfa
2025

TEŞEKKÜR

Yüksek lisans sürecim boyunca bilgi birikimi, rehberliği ve akademik vizyonuyla her zaman yolumu aydınlatan; karşılaştığım her sorunda sabırla destek olan, çalışma disiplininden ilham aldığım değerli tez danışmanım Dr. Öğr. Üyesi Serdar Çiftçi'ye en içten teşekkürlerimi sunarım. Kendisinin akademik rehberliği bu tezin en önemli yapı taşlarından biri olmuştur.

Ayrıca, yüksek lisans eğitimim süresince bana bilgi ve deneyimleriyle katkı sağlayan, akademik gelişimime yön veren Bilgisayar Mühendisliği Anabilim Dalı'ndaki tüm değerli hocalarıma teşekkür ederim.

İÇİNDEKİLER

ÖZET

YÜKSEK LİSANS TEZİ

DİKKAT TABANLI OTOKODLAYICI KULLANARAK TEK GÖRÜNTÜDEN YÜKSEK DİNAMİK ARALIKLI GÖRÜNTÜ OLUŞTURMA

CEVHER RENK

HARRAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

Tez Danışman: Dr. Öğr. Üyesi SERDAR ÇİFTÇİ
Yıl: 2025, Sayfa : 68

Tek görüntüden Yüksek Dinamik Aralık (High Dynamic Range – HDR, YDA) görüntü yeniden yapılandırma, modern görüntü işleme alanında hâlen çözüm bekleyen önemli zorluklardan biridir. Düşük Dinamik Aralık (Low Dynamic Range – LDR, DDA) görüntülerde karşılaşılan pozlama yetersizliği, bilgi kaybının önlenmesi açısından ciddi bir engel oluşturmaktadır. Bu tez çalışmasında, çoklu pozlama sentezi esasına dayalı U-Net benzeri Kodlayıcı–Çözücü (Encoder–Decoder) mimarili bir YDA üretim sistemi temel alınarak, farklı Dikkat (Attention) mekanizmalarının bu yapıya entegrasyonu araştırılmıştır. Önerilen mimari, Otokodlayıcı (Autoencoder) tabanlı bir çerçevede yapılandırılmış; Uzamsal Dikkat (Spatial Attention), Kanalsal Dikkat (Channel Attention), Darboğaz Dikkat (Bottleneck Attention), Sıkıştırma–Uyarma Dikkat (Squeeze-and-Excitation – SE Attention) ve Öz Dikkat (Self-Attention) olmak üzere beş farklı dikkat mekanizması, ayrı ayrı model varyantları olarak sisteme entegre edilmiştir. Ayrıca, modele eklenen Pozlama Tutarlılığı Kaybı (Exposure Consistency Loss) başta olmak üzere farklı kayıp fonksiyonları ile performans iyileştirmeleri gerçekleştirilmiştir. Geliştirilen modeller, DrTMO veri kümesi kullanılarak eğitilmiş ve Tepe Sinyal-Gürültü Oranı (Peak Signal-to-Noise Ratio – PSNR), Yapısal Benzerlik İndeksi (Structural Similarity Index – SSIM) ve Öğrenilmiş Algısal Görüntü Benzerliği (Learned Perceptual Image Patch Similarity – LPIPS) metrikleri üzerinden nicel olarak değerlendirilmiştir. Elde edilen sonuçlar, dikkat mekanizmalarının ve yeni kayıp fonksiyonlarının YDA üretim kalitesine önemli katkılar sağladığını; özellikle Uzamsal Dikkat tabanlı modelin hem yapısal doğruluk hem de algısal kalite açısından en yüksek performansı gösterdiğini ortaya koymaktadır. Bulgular, ablasyon analizi ve görsel karşılaştırmalar ile detaylı biçimde desteklenmiştir.

ANAHTAR KELİMELER: Otokodlayıcı, YDA yeniden yapılandırma , Dikkat mekanizmaları

ABSTRACT

MASTER THESIS

SINGLE-IMAGE HIGH DYNAMIC RANGE RECONSTRUCTION USING ATTENTION-BASED AUTOENCODER

CEVHER RENK

HARRAN UNIVERSITY
INSTITUTE OF GRADUATE EDUCATION
DEPARTMENT OF COMPUTER ENGINEERING

Thesis Supervisor: Assist. Prof. Dr. SERDAR ÇİFTÇİ
Year: 2025, Page : 68

Single-image High Dynamic Range (HDR) reconstruction remains a significant challenge in modern image processing. In Low Dynamic Range (LDR) images, exposure limitations constitute a major obstacle to preventing information loss. In this thesis, a U-Net-like Encoder–Decoder based HDR generation framework was adopted, and the integration of different attention mechanisms into this structure was investigated. The proposed architecture was implemented within an autoencoder framework, where five attention mechanisms—Spatial Attention, Channel Attention, Bottleneck Attention, Squeeze-and-Excitation (SE), and Self-Attention—were incorporated as separate model variants. In addition, performance improvements were achieved by introducing different loss functions, most notably the Exposure Consistency Loss, into the training process. The developed models were trained on the DrTMO dataset and quantitatively evaluated using Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS). The results demonstrate that the proposed attention mechanisms and additional loss functions significantly improve HDR reconstruction quality, with the Spatial Attention–based model achieving the highest performance in terms of both structural fidelity and perceptual quality. The findings are further supported by ablation studies and detailed visual comparisons.

KEYWORDS: Autoencoder, HDR reconstruction , Attention mechanisms

ŞEKİLLER DİZİNİ

ÇİZELGELER DİZİNİ

KISALTMALAR

ANN	Artificial Neural Network (Yapay Sinir Ağı)
BAM	Bottleneck Attention Module (Dar Boğaz Dikkat Modülü)
CBAM	Convolutional Block Attention Module (Evrişimsel Blok Dikkat Modülü)
CNN	Convolutional Neural Network (Evrişimsel Sinir Ağı)
GAN	Generative Adversarial Network (Çekişmeli Üretici Ağ)
HDR	High Dynamic Range (Yüksek Dinamik Aralık)
HDR-AttNet	HDR Attention Network (YDA Dikkat Ağı)
KNN	K-Nearest Neighbor (En Yakın Komşu)
LDR	Low Dynamic Range (Düşük Dinamik Aralık)
LPIPS	Learned Perceptual Image Patch Similarity (Öğrenilmiş Algısal Görüntü Benzerliği)
MRI	Magnetic Resonance Imaging (Manyetik Rezonans Görüntüleme)
PSNR	Peak Signal-to-Noise Ratio (Tepe Sinyal-Gürültü Oranı)
SI-HDR	Single Image HDR (Tek Görüntüden Yüksek Dinamik Aralık)
SSIM	Structural Similarity Index (Yapısal Benzerlik İndeksi)
SVM	Support Vector Machine (Destek Vektör Makinesi)
VAE	Variational Autoencoder (Varyasyonel Otokodlayıcı)
Vit	Vision Transformer (Görü Dönüştürücü)

1. GİRİŞ

Yüksek Dinamik Aralıklı (YDA) görüntüleme, sahnedeki çok parlak ve karanlık bölgelerin aynı anda ve ayrıntılı biçimde temsil edilmesini sağlayarak klasik görüntüleme sistemlerinin temel sınırlamalarını aşmayı hedefleyen gelişmiş bir tekniktir (Debevec ve Malik, 1997, Mantiuk vd., 2015; Reinhard, 2020). YDA görüntüleme teknolojisi, insan görsel sisteminin geniş parlaklık algısını taklit ederek sahnelerin daha doğal, kontrastı dengeli ve detay açısından zengin biçimde yeniden üretilmesine olanak tanır. Bu özellik, YDA'yı fotoğrafçılık, medikal görüntüleme, otonom sistemler ve bilgisayarlı görü gibi birçok alanda vazgeçilmez bir bileşen hâline getirmiştir.

Geleneksel YDA üretim yöntemleri, aynı sahnenin farklı pozlama değerlerinde çekilmiş birden fazla görüntüsünün birleştirilmesine dayanır. Ancak bu yöntem, hareketli sahnelerde nesne kaymaları ve parlaklık farklılıkları nedeniyle hayaletlenme (ghosting) gibi bozulmalara yol açabilir (Kalantari ve Ramamoorthi, 2017; Yan vd., 2023). Bu sınırlılıklar, tek görüntüden YDA üretimi (Single-Image HDR – SI-HDR) kavramını ön plana çıkarmıştır. SI-HDR, yalnızca tek bir Düşük Dinamik Aralıklı (DDA) görüntüden yüksek dinamik aralıklı sahne bilgisi üretmeyi amaçlayarak hem çekim süresini kısaltmakta hem de hareketli sahnelerde bozulma oluşumunu azaltmaktadır (Eilertsen vd., 2017; Liu vd., 2020). Son yıllarda derin öğrenmeye dayalı yöntemler, bu alandaki en başarılı yaklaşımlardan biri hâline gelmiştir. Özellikle kodlayıcı–çözücü (encoder–decoder) yapıları, DDA görüntülerden YDA sahne bilgisi elde etmede etkin bir çözüm sunmuş; bu yapılara entegre edilen dikkat (attention) mekanizmaları ise modelin sahnedeki kritik bölgelere odaklanmasını sağlayarak hem yapısal hem de algısal kaliteyi artırmıştır.

Bu tezde, Le vd. (2023) tarafından önerilen tek görüntüden çoklu sahte pozlama üretimine dayalı YDA rekonstrüksiyonu çalışması temel alınmıştır. Söz konusu model, tek bir DDA görüntüden sahte pozlama (synthetic exposure) sentezi gerçekleştirerek yüksek dinamik aralıklı sahne rekonstrüksiyonu sağlamaktadır. Bu çalışmada, söz konusu mimari genişletilerek U-Net tabanlı kodlayıcı–çözücü yapıya beş farklı dikkat mekanizması entegre edilmiştir (Ronneberger vd., 2015; Luong, Pham ve Manning, 2015; Choromanski vd., 2020). Böylece modelin bilgi odaklı öğrenme kabiliyeti artırılmış, dikkat modüllerinin etkileri sistematik olarak karşılaştırılmıştır. Uygulanan dikkat mekanizmaları sırasıyla; Uzamsal Dikkat (Spatial Attention) (Woo vd., 2018), Kanalsal Dikkat (Mekruksavanich ve Jitpattanakul, 2023), Sıkıştırma–Uyarma Dikkat (Squeeze-and-Excitation-SE Attention) (Hu vd., 2018), Darboğaz Dikkat (Bottleneck Attention) (Park vd.,

2018) ve Öz Dikkat (Self-Attention) (Vaswani vd., 2017; Zhang vd., 2019; Lin vd., 2017) modülleridir. Bu modüller, sahnedeki kritik bölgeleri ön plana çıkararak hem yapısal hem de algısal kalitenin artırılmasına katkı sağlamaktadır. Modelin performansı, DrTMO veri kümesi (Endo vd., 2017) üzerinde Tepe Sinyal–Gürültü Oranı (Peak Signal-to-Noise Ratio – PSNR) , Yapısal Benzerlik İndeksi (Structural Similarity Index – SSIM) (Wang vd., 2004) ve Öğrenilmiş Algısal Görüntü Benzerliği (Learned Perceptual Image Patch Similarity – LPIPS) (Zhang vd., 2018) metrikleriyle değerlendirilmiştir. Ek olarak, klasik kayıp fonksiyonlarına Pozlama Tutarlılığı Kaybı (Exposure Consistency Loss) eklenerek parlaklık tutarlılığı ve renk doğruluğu geliştirilmiştir (Kalantari ve Ramamoorthi, 2017; Yan vd., 2021).

Çalışmanın erken aşamalarında Varyasyonel Otokodlayıcı (Variational Autoencoder – VAE) tabanlı bir mimari (Kingma ve Welling, 2013) test edilmiş; ancak yüksek dinamik aralıklı sahnelerde beklenen performans elde edilememiştir. Bu nedenle araştırma, kodlayıcı–çözücü temelli yapı üzerine odaklanmış ve dikkat mekanizmalarının entegrasyonu ile ağın temsil gücünün artırılması hedeflenmiştir.

1.1. Görüntü İşleme ve Görüntü İyileştirme Tekniklerinin Gelişimi

Görüntü işleme, dijital görsellerin analiz edilmesi, dönüştürülmesi ve iyileştirilmesine odaklanan disiplinler arası bir araştırma alanıdır. Alanın temelleri, dijital görüntü işlemenin erken dönem kuramsal ve uygulamalı çalışmalarıyla atılmıştır (Rosenfeld, 1976; Ballard ve Brown, 1982). Bu dönemde geliştirilen ilk sistemler, analog verilerin sayısallaştırılması ve temel morfolojik işlemlerle kenar belirleme, segmentasyon ve gürültü azaltma gibi görevlerin gerçekleştirilmesine odaklanmıştır. 1960'lı yıllarda NASA tarafından geliştirilen uydu görüntü analiz sistemleri, sayısal görüntülemenin ilk örneklerini sunarak modern dijital görüntü işlemenin gelişimine öncülük etmiştir (Hall, 1979). Donanım kapasitesindeki artış, yüksek çözünürlüklü sensörlerin geliştirilmesi ve dijital depolama olanaklarının genişlemesiyle birlikte sistemler giderek daha karmaşık sahne analizleri yapabilir hâle gelmiştir (Burger ve Burge, 2010; Dongur vd., 2022; Date vd., 2022). Bu ilerlemeler, görüntü işlemenin yalnızca görsel kalite artırımı değil; aynı zamanda medikal görüntüleme (Litjens vd., 2017; Sharma ve Mishra, 2022), endüstriyel kalite kontrol (Miskdjian vd., 2024), uzaktan algılama (Jensen, 1986; Liu, 2024) ve güvenlik sistemleri (Sinha vd., 2018) gibi farklı alanlarda da yaygın biçimde kullanılmasını sağlamıştır. Görüntü işlemeye dayalı yapısal hasar tespiti gibi uygulamalar da alanın mühendislikteki kapsamını genişletmiş ve afet sonrası izleme süreçlerinde etkin biçimde kullanılmaya başlanmıştır (Crognale vd., 2023).

Tarihsel olarak, görüntü iyileştirme (image enhancement), görüntü işlemenin en temel alt alanlarından biri olarak kabul edilmektedir. Bu yöntemler, görüntülerin görsel kalitesini artırmak ve bilgi çıkarımını kolaylaştırmak amacıyla geliştirilmiştir. Klasik yaklaşımlar; histogram eşitleme, kontrast artırma ve gürültü azaltma gibi işlemlere dayanmakta olup, özellikle düşük kontrastlı sahnelerde ayrıntı kaybını önlemede yetersiz kalmıştır (Burger ve Burge, 2010). Bilgisayar teknolojilerindeki gelişmelerle birlikte, görüntü iyileştirme teknikleri istatistiksel ve öğrenme tabanlı modellere evrilmiş; bu sayede medikal, endüstriyel ve uzaktan algılama uygulamalarında daha yüksek doğruluk elde edilmiştir (Lore vd., 2017; Zhang, 2022; Wang vd., 2022; Chen vd., 2023; Zangana vd., 2024). Derin öğrenme tabanlı görüntü iyileştirmeye ilişkin kapsamlı derlemeler, bu alandaki yöntemsel eğilimleri detaylı biçimde incelemiştir (Manaa vd., 2023). Benzer şekilde, tıbbi görüntülerde kontrast artırımı ve tanısal doğruluğun geliştirilmesine yönelik çalışmalar da disiplinler arası etkinliği göstermektedir (Sharma ve Mishra, 2022). Ayrıca, görüntü restorasyonu ve iyileştirmede difüzyon modellerine dayalı yaklaşımlar da son dönemde dikkat çekmekte ve yeni nesil öğrenme paradigması olarak değerlendirilmektedir (Li ve vd., 2023). Bu gelişmeler, YDA görüntüleme teknolojilerinin ortaya çıkışına da zemin hazırlamıştır. YDA kavramı, sahnelerdeki çok parlak ve karanlık bölgelerin aynı anda temsil edilmesini sağlayan bir teknik olarak ilk kez Debevec ve Malik (1997) tarafından sistematik biçimde tanımlanmıştır. Klasik çoklu pozlama tabanlı yöntemler, bu yaklaşımı temel alarak farklı pozlama seviyelerinde çekilmiş görüntüleri birleştirmeyi önermiş; ancak hareketli sahnelerde hayaletlenme gibi bozulmalarla karşılaşılmıştır. Bu sınırlılıklar, daha sonra tek görüntüden YDA üretimi yaklaşımlarının geliştirilmesine zemin hazırlamıştır. Bununla birlikte, tek görüntüden YDA dönüştürme görevine yönelik koşullu difüzyon modellerinin de başarıyla uygulandığı görülmektedir (Dalal vd., 2023).

Son yıllarda geliştirilen YDA yöntemleri, gürültü azaltma, parlaklık dengesi ve kontrast iyileştirme gibi hedefleri bir araya getirerek daha tutarlı sonuçlar sunmaktadır. Derin ağlarda uzamsal çözünürlüğü korumaya yönelik genişletilmiş konvolüsyon gibi tasarımlar, bağlamsal farkındalığın artırılmasında önemli bir rol oynamıştır (Chen vd., 2017). HDRUNet modeli (Chen vd., 2021) gürültü giderme ve bit derinliği iyileştirme işlemlerini tek bir ağ yapısında bütünleştirmiştir. UPHDR-GAN (Li vd., 2022) yapısal ve algısal kaliteyi optimize etmiş; FHDR (Khan vd., 2019) kenar koruma ve pozlama dengesini güçlendirmiştir. Çoklu çözünürlük temelli Derin ağ serileri (Deep Network Series) yaklaşımı (Aghabiglou vd., 2023) ise parlaklık tutarlılığında belirgin gelişim ortaya koymuştur. Böylece görüntü işleme sistemleri, analog tabanlı filtreleme yöntemlerinden bağlamsal farkındalığa sahip

derin öğrenme modellerine evrilmiştir.

1.2. Makine Öğrenmesi ve Derin Öğrenme Yaklaşımları

Makine öğrenmesi (Machine Learning), dışarıdan açık kurallar tanımlamaksızın sistemlerin deneyimden öğrenmesini ve bu öğrenimi gelecekteki benzer durumlara genelleyebilmesini sağlayan bir yapay zekâ yaklaşımıdır. Klasik programlama yöntemlerinden farklı olarak, veri içerisindeki istatistiksel örüntüleri tanımlayarak karar verme süreçlerini otomatikleştirir (Mitchell, 1997). Tipik bir makine öğrenmesi modeli, giriş ve çıkış verileri arasındaki ilişkiyi tanımlayan bir fonksiyon öğrenir ve belirli bir hata ölçütünü minimize etmeye çalışır. Bu yöntemler genel olarak üç temel kategoriye ayrılır: denetimli öğrenme (supervised learning), denetimsiz öğrenme (unsupervised learning) ve pekiştirmeli öğrenme (reinforcement learning). Denetimli öğrenme, etiketli veri kümeleri üzerinden girdilerle çıktılar arasındaki bağıntıyı öğrenmeyi hedeflerken; denetimsiz öğrenme, veri içerisindeki gizli yapıları ve kümeleri keşfetmeye odaklanır. Pekiştirmeli öğrenme ise bir ajanın çevresiyle etkileşimi sonucunda ödül sinyalleri üzerinden optimal karar politikalarını öğrenmesini sağlar (Sutton ve Barto, 1998). Klasik makine öğrenmesi algoritmaları Karar Ağaçları (Decision Trees), Destek Vektör Makineleri (Support Vector Machines - SVM) ve En Yakın Komşu (K-Nearest Neighbors - KNN)—yapılandırılmış veri problemlerinde etkili sonuçlar vermekle birlikte, yüksek boyutlu ve çok kanallı görsel verilerde sınırlı başarı göstermektedir. Görsel verilerdeki karmaşık örüntülerin öğrenilmesi gereksinimi, derin öğrenme yaklaşımlarının (Deep Learning) ortaya çıkışını hızlandırmıştır.

Derin öğrenme, çok katmanlı yapay sinir ağları (Artificial Neural Networks – ANN) aracılığıyla verilerin çok düzeyli temsillerinin öğrenilmesini sağlayan gelişmiş bir makine öğrenmesi yöntemidir. Bu yaklaşım, modelin öznitelikleri doğrudan ham veriden çıkarabilmesini mümkün kılarak öznitelik mühendisliği ihtiyacını büyük ölçüde azaltmıştır. Özellikle Evrişimsel Sinir Ağları (Convolutional Neural Networks – CNN), yerel uzamsal özellikleri (kenar, doku, desen) yakalama kabiliyetleri sayesinde görüntü sınıflandırma, nesne tespiti ve bölütleme gibi görevlerde önemli başarılar elde etmiştir (Krizhevsky vd., 2012). Daha karmaşık görevlerde, kodlayıcı–çözücü yapıları, giriş verilerinden çıkarılan özellikleri düşük boyutlu bir temsile dönüştürüp yeniden yapılandırarak süper çözünürlük, gürültü azaltma, görüntü iyileştirme ve YDA görüntü sentezi gibi görevlerde etkin biçimde kullanılmaktadır. Bu tür ağ yapıları, görsel bilgi kaybını en aza indirerek hem yapısal hem de algısal kaliteyi artırmaktadır.

Son yıllarda, derin öğrenme mimarilerine dikkat mekanizmalarının entegre edilmesi, model performansında önemli gelişmelere yol açmıştır. İlk olarak doğal dil işleme alanında önerilen dikkat mekanizması (Bahdanau vd., 2014), modellerin uzun bağımlılıkları öğrenme kapasitesini artırmış ve kısa sürede görsel görevlere de uyarlanmıştır. Dikkat tabanlı modeller, sahnedeki kritik bölgelere seçici biçimde odaklanarak bilgi temsilinin tutarlılığını geliştirmekte ve görsel ayrıntıların korunmasını sağlamaktadır. Bu yaklaşım, özellikle YDA görüntüleme gibi yüksek dinamik aralık gerektiren zorlu görevlerde yapısal doğruluk ve algısal kalite açısından etkili sonuçlar sunmaktadır. Dikkat mekanizmasının en kapsamlı biçimde uygulandığı yapı, dönüştürücü (Transformer) mimarisidir. Vaswani vd. (2017) tarafından önerilen bu model, küresel bağlamı modelleme kabiliyeti sayesinde yalnızca dil işleme değil, görüntü iyileştirme ve YDA üretimi gibi alanlarda da yüksek başarı göstermektedir. Ayrıca, son dönem çalışmalarda Otokodlayıcı tabanlı yapılar, sahne bilgisini düşük boyutlu bir temsilde kodlayıp yeniden yapılandırarak pozlama dengesi ve renk doğruluğunu iyileştirmek amacıyla kullanılmaktadır.

Sonuç olarak, makine öğrenmesi ve derin öğrenme tabanlı yaklaşımlar, görüntü işleme alanında geleneksel yöntemlerin ötesine geçerek yeni bir yaklaşım oluşturmuştur. Bu yöntemler, veriden doğrudan öğrenebilme ve karmaşık sahneleri yüksek doğrulukla modelleyebilme yetenekleri sayesinde, klasik görüntüleme sistemlerinin sınırlarını aşarak YDA görüntüleme gibi ileri düzey görevlerde yüksek başarı potansiyeli sunmaktadır.

1.3. Literatür Analizi, Araştırma Sınırlılıkları ve Problem Tanımı

Tek görüntüden YDA görüntü üretimi, klasik çoklu pozlama yöntemlerine güçlü bir alternatif olarak öne çıkmakta; hareketli sahnelerde hayaletlenme bozulmalarını azaltma, donanımsal maliyeti düşürme ve gerçek zamanlı sistemlere entegrasyon gibi önemli avantajlar sunmaktadır. Bununla birlikte, mevcut literatür incelendiğinde bu alanda hâlen bazı araştırma sınırlılıklarının ve geliştirmeye açık yönlerin bulunduğu görülmektedir. Mevcut yaklaşımların önemli bir kısmı, sahnedeki kritik bölgelere yeterince odaklanamamakta; özellikle aşırı pozlanmış veya düşük aydınlatmalı alanlarda ayrıntı kayıpları meydana gelmektedir. Bu durumu gidermek amacıyla geliştirilen çeşitli dikkat mekanizmaları, modelin sahnedeki önemli bölgeleri ön plana çıkararak bilgi odaklı öğrenmeyi güçlendirmeyi amaçlamıştır. Ancak bu mekanizmaların YDA üretimi üzerindeki etkileri çoğunlukla dolaylı biçimde incelenmiş, farklı dikkat türlerinin aynı mimari üzerinde sistematik olarak karşılaştırıldığı kapsamlı çalışmalar sınırlı kalmıştır. Bu durum, sahne bağlamına göre hangi dikkat yapısının daha etkili olduğunu belirlemeyi

güçleştirmektedir.

Ayrıca, birçok derin öğrenme tabanlı yöntem model karmaşıklığı, hesaplama yükü ve bellek gereksinimi gibi pratik kısıtları yeterince dikkate almamaktadır. Özellikle mobil cihazlar ve gömülü sistemler için optimize edilmiş, düşük hesaplama maliyetli YDA modellerinin eksikliği, teknolojinin gerçek zamanlı uygulamalara entegrasyonunu sınırlamaktadır. Derin ağlar yüksek doğruluk sağlasa da, parametre sayısındaki artış doğrudan hesaplama verimliliğini olumsuz etkilemektedir (Shen ve Wu, 2021). Buna ek olarak, literatürde kullanılan veri setlerinin çoğunlukla statik sahnelerden oluşması ve pozlama çeşitliliğinin sınırlı olması, modellerin farklı aydınlatma koşullarına genellenebilirliğini zayıflatmakta ve gerçek dünya uygulamalarında performans düşüşlerine neden olmaktadır.

Bu tez çalışmasında, söz konusu sınırlılıklar dikkate alınarak beş farklı dikkat mekanizması, Le vd. (2023) tarafından önerilen U-Net tabanlı kodlayıcı–çözücü mimari üzerine bağımsız biçimde entegre edilmiştir. Ayrıca, parlaklık tutarlılığını ve renk doğruluğunu artırmak amacıyla Pozlama Tutarlılığı Kaybı eklenmiştir. Böylece farklı dikkat türlerinin YDA üretimi üzerindeki etkileri sistematik biçimde değerlendirilmiş; yapısal doğruluk, algısal kalite ve hesaplama verimliliği birlikte analiz edilerek literatürdeki mevcut sınırlamalara yönelik özgün ve bütüncül bir çözüm ortaya konulmuştur.

1.4. Araştırmanın Bilimsel Hedefi, Özgün Katkısı ve Tezin Yapısı

YDA görüntüleme alanında, tek bir DDA görüntüden eksik sahne bilgisinin tamamlanması ve yüksek görsel kaliteye sahip YDA çıktılar elde edilmesi, güncel araştırmalarda önemli bir hedef hâline gelmiştir. Bu tez çalışmasında, yalnızca yapısal doğruluğu değil, aynı zamanda algısal kaliteyi de artıran özgün bir yöntem önermektedir.

Çalışmanın özgün katkıları şu şekilde özetlenebilir:

1. **Dikkat mekanizmalarının mimariye entegrasyonu ve karşılaştırılması:** Mimariye beş farklı dikkat modülü entegre edilerek her biri için ayrı model varyantları oluşturulmuştur. Bu yaklaşım, ağın sahnedeki kritik bölgelere odaklanma, uzun menzilli bağıntıları modelleme ve kanal düzeyinde bilgi akışını düzenleme yeteneğini güçlendirmiştir. Böylece hem

yapısal doğruluk hem de görsel detayların korunması açısından dikkat modüllerinin YDA sentezine katkısı ilk kez sistematik biçimde incelenmiştir.

2. **Kayıp fonksiyonunun iyileştirilmesi:** Pozlama Tutarlılığı Kaybı, parlaklık tutarlılığı ve renk doğruluğunu artırarak YDA üretim kalitesine katkı sağlamıştır.
3. **Çok boyutlu performans değerlendirmesi:** Tüm model varyantları PSNR, SSIM ve LPIPS metrikleriyle DrTMO veri kümesi üzerinde test edilmiş; en yüksek performans uzamsal dikkat tabanlı modelden elde edilmiştir.
4. **Literatürdeki sınırlılığın giderilmesi:** Dikkat mekanizmalarının YDA üretimi üzerindeki etkileri ilk kez karşılaştırmalı biçimde incelenerek, literatürdeki önemli bir eksiklik giderilmiştir.

Sonuç olarak, önerilen yöntem, YDA üretiminde yapısal doğruluk, algısal kalite ve hesaplama verimliliği açısından önemli gelişmeler sağlamış; dikkat mekanizmalarının rolünü sistematik biçimde ortaya koyarak alana bilimsel ve özgün bir katkı sunmuştur.

1.5. YDA Görüntüleme Sistemleri ve Tek Görüntüden YDA Yaklaşımları

Geleneksel görüntüleme sistemleri, sahnelerdeki parlak ve karanlık bölgeleri aynı anda doğru biçimde temsil edemediğinden, hem aşırı pozlanmış hem de gölgede kalan alanlarda bilgi kaybı yaşanır. Bu sınırlılığı aşmak için geliştirilen YDA görüntüleme teknikleri, farklı pozlama seviyelerinde çekilen çoklu görüntülerin birleştirilmesiyle geniş parlaklık aralığını temsil etmeyi hedefler. Böylece hem yüksek ışıklı hem de düşük aydınlatmalı bölgelerdeki detaylar korunarak, insan görsel sistemine daha yakın temsiller elde edilir. Ancak çoklu pozlama yöntemleri, hareketli sahnelerde hayaletlenme olarak bilinen bozulmalara neden olmakta ve uzun çekim süresi nedeniyle pratik uygulamalarda zorluk yaratmaktadır. Ayrıca donanımsal gereksinimlerin artması, bu yöntemlerin taşınabilir cihazlarda kullanılmasını güçleştirmektedir. Bu kısıtlamalar, özellikle dinamik ortamlarda daha esnek ve yazılımsal çözümlere duyulan ihtiyacı ortaya çıkarmıştır.

Bu gereksinim doğrultusunda geliştirilen tek görüntüden YDA yaklaşımları, yalnızca tek bir DDA görüntüden YDA sahne bilgisi üretmeyi amaçlamaktadır. Tek görüntüden YDA yöntemleri, çoklu pozlama gereksinimini ortadan kaldırarak hem çekim süresini kısaltmakta hem de hareketli sahnelerde oluşan hayaletlenme

bozulmalarını azaltmaktadır. Günümüzde bu yöntemler, kaybolan parlaklık ve gölge detaylarını tahmin ederek, daha doğal kontrast dağılımına sahip YDA çıktılar üretebilmektedir (Sharif vd., 2021). Son dönemde yapılan çalışmalar, YDA sentezine yönelik uçtan uca diferansiyellenebilir öğrenme yaklaşımları (Kim vd., 2021), süper çözünürlük ve ters ton eşleme görevlerini eş zamanlı optimize eden GAN tabanlı YDA video modelleri (Kim vd., 2020) ve YDA görüntüleme için derin öğrenme tabanlı iyileştirme stratejilerine ilişkin genel bir değerlendirme sunulan çalışmalar (Park ve vd., 2022) ile alanın gelişiminde önemli katkılar sağlamıştır. Bu gelişmeler, YDA üretimini donanımsal sınırlamalardan bağımsız hâle getirerek, özellikle mobil görüntüleme ve gerçek zamanlı uygulamalarda yeni bir araştırma alanı oluşturmuştur.

2. ÖNCEKİ ÇALIŞMALAR

Görüntüleme teknolojilerinin evrimsel sürecinde, yüksek dinamik aralıklı sahnelerin daha doğru ve detaylı biçimde temsil edilmesi, hem akademik araştırmaların hem de endüstriyel uygulamaların öncelikli hedeflerinden biri hâline gelmiştir. Bu doğrultuda geliştirilen YDA üretim yöntemleri, zaman içerisinde önemli bir çeşitlenme göstermiş; fiziksel temelli pozlama birleştirme algoritmalarından, derin öğrenme tabanlı veri odaklı mimarilere doğru önemli bir dönüşüm yaşanmıştır. İlk nesil yaklaşımlar, çoklu DDA görüntülerin birleştirilmesiyle yüksek parlaklık aralığına ulaşmayı hedeflerken; günümüz yöntemleri, sahne bilgisini tek bir görüntü üzerinden modelleyerek YDA üretimini daha esnek, verimli ve gerçek zamanlı hâle getirmeyi amaçlamaktadır. Bu bölümde, YDA görüntüleme literatürü dört ana başlık altında kapsamlı şekilde ele alınmaktadır. İlk olarak, geleneksel YDA üretim teknikleri tarihsel bağlamda incelenmekte ve bu yöntemlerin sunduğu avantajlar ile karşılaştıkları sınırlamalar ortaya konulmaktadır. Devamında, derin öğrenme temelli YDA çözümleri analiz edilerek, modern mimarilerin YDA sentezine olan katkıları değerlendirilecektir. Üçüncü olarak, otokodlayıcı yapılar ve bu yapıların YDA bağlamında ne şekilde uyarlandıkları tartışılacaktır. Son bölümde ise, dikkat mekanizmalarının YDA üretiminde nasıl konumlandığı; bu mekanizmaların mimari düzeydeki rolü ve performansa etkisi detaylandırılacaktır. Literatürün bu dört eksende sistematik biçimde taranmasıyla, mevcut yöntemlerin kapsamı, sınırlılıkları ve gelişim eğilimleri net biçimde ortaya konulacak; aynı zamanda bu tez kapsamında önerilen modelin konumlandığı araştırma aralığı belirginleştirilecektir.

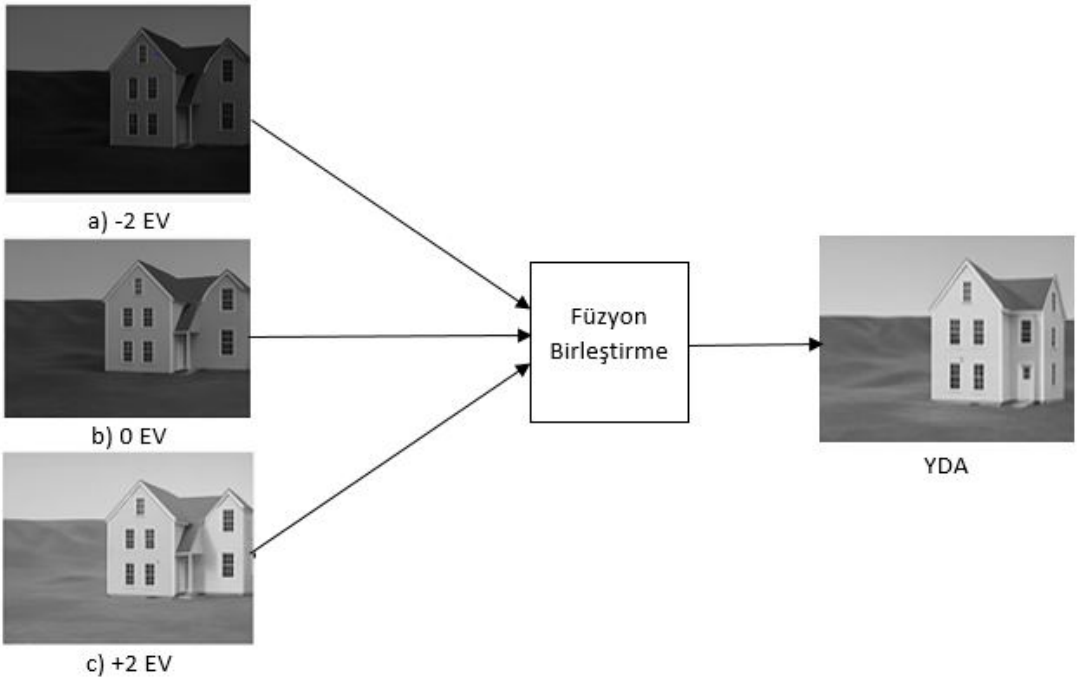
2.1. Klasik YDA Üretim Teknikleri

Görüntüleme teknolojilerinin evriminde, sahnelerdeki geniş parlaklık aralıklarını kayıpsız biçimde temsil edebilmek tarihsel olarak önemli bir teknik problem olmuştur. Görüntüdelerde aşırı parlak ve karanlık bölgelerin birlikte doğru şekilde yakalanması kritik bir ihtiyaç hâline gelmiştir. Bu ihtiyaç doğrultusunda geliştirilen YDA görüntüleme teknikleri, başlangıçta çoklu pozlama yöntemine dayalı olarak şekillenmiştir. Bu yöntem, aynı sahnenin farklı aydınlık seviyelerini yansıtan çoklu görüntüler elde edilmesini ve bu görüntülerin birleştirilerek dinamik aralığın genişletilmesini temel alır (Mantiuk vd., 2015). Klasik YDA sistemlerinin ilk uygulamaları sabit kamera ve üçayak (tripod) destekli ortamlarda gerçekleştirilmiş, daha sonra dijital fotoğraf makinelerinde yer alan otomatik pozlama özellikleriyle yaygınlık kazanmıştır. Bu sistemler özellikle durağan sahnelerde, detaylı ışık dağılımını yakalamakta oldukça başarılıdır. Ancak sahnede hareketli nesnelerin

bulunması veya çekim sürecinde kamera titremesi gibi faktörler, bu yöntemlerin etkinliğini ciddi biçimde sınırlandırmaktadır.

YDA görüntüleme süreci genellikle dört temel adımdan oluşur: pozlama basamaklama (bracketing), görüntü hizalama (alignment), görüntü birleştirme (merging/fusion) ve ton eşleme (tone mapping). İlk adım olan pozlama basamaklama, aynı sahnenin farklı pozlama sürelerinde ardışık görüntülerinin yakalanmasını içerir. Genellikle -2EV, 0EV ve +2EV gibi pozlama değerleriyle (exposure value - EV) 3 ila 7 arasında kare elde edilir. Bu görüntülerin her biri, sahnenin farklı aydınlık bölgelerini temsil eder: düşük pozlamalı görüntüler parlak detayları korurken, yüksek pozlamalı görüntüler karanlık alanlardaki bilgiyi ortaya çıkarır.

Bu adımların genel akışı Şekil 2.1'de gösterilmektedir.



Şekil 2.1. Pozlama basamaklama (bracketing) temelli klasik YDA üretim süreci (Yapay Zekâ destekli görselleştirme ile oluşturulmuştur).

Pozlama basamaklama sürecinin başarılı bir şekilde gerçekleştirilebilmesi, sahnenin sabit kalmasına ve kameranın sabitlenmesine bağlıdır. El titremesi, hareketli nesneler ve ışık değişimleri gibi faktörler, pozlama görüntüleri arasında geometrik farklılıkların oluşmasına neden olur. Modern sistemlerde bu sorun, yüksek hızlı çekim sistemleri ve optik görüntü sabitleme teknolojileriyle kısmen azaltılabilse de, pozlama basamaklama sonrasındaki hizalama ve birleştirme aşamaları sürecin

doğruluğunu belirleyen kritik adımlar olmaya devam etmektedir. İkinci aşama olan görüntü hizalama, farklı pozlama değerlerinde çekilen görüntülerin aynı geometrik referans sistemine dönüştürülmesini amaçlar. YDA üretim zincirinin en kritik bileşenlerinden biri olan bu adım, hem kamera kaynaklı titreşimler hem de sahnedeki nesne hareketleri nedeniyle oluşabilecek konumsal farklılıkları giderir. Görüntü hizalama yöntemleri, küresel (global) ve yerel (local) olarak ikiye ayrılır. Küresel hizalama yöntemleri, tüm görüntüye tek bir dönüşüm (örneğin öteleme, döndürme, ölçekleme) uygular ve sabit sahneler için yeterlidir. Ancak hareketli nesne içeren sahnelerde yerel hizalama gereklidir. Bu yöntemde sahne parçalara ayrılır ve her bölgeye farklı geometrik dönüşümler uygulanır (Eilertsen vd., 2017).

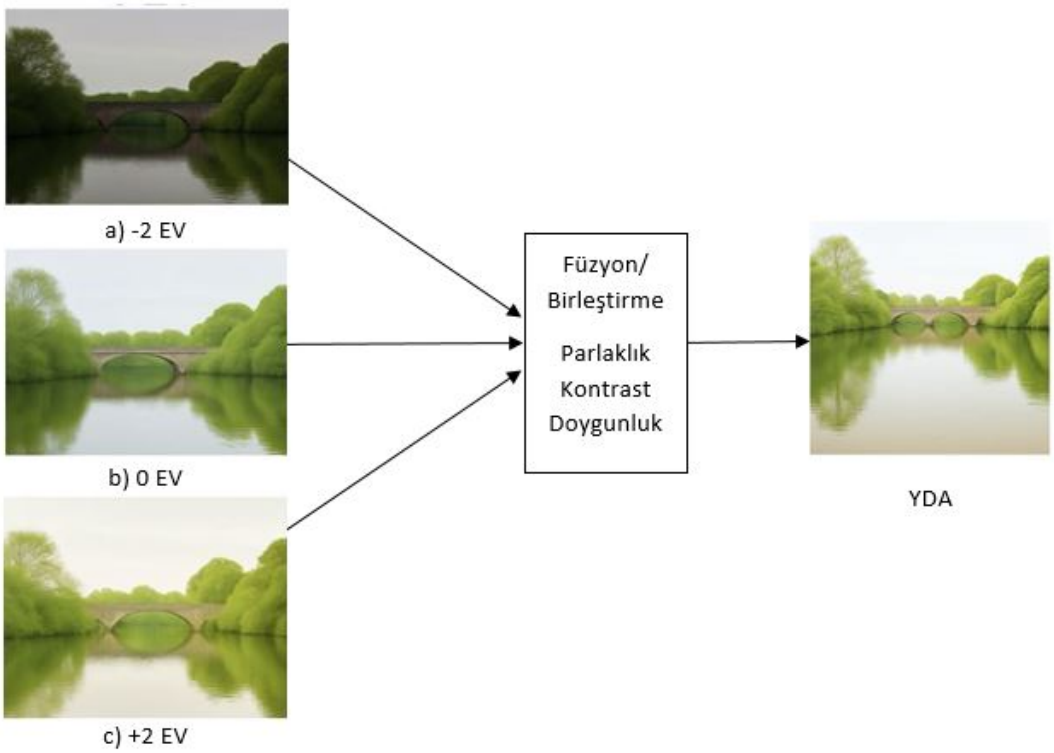
Küresel ve yerel hizalama yöntemlerinin karşılaştırması, Şekil 2.2'de gösterilmektedir.

Şekil 2.2. Küresel hizalama ve yerel hizalama ile sahnenin doğru hizalanması (Yapay Zekâ destekli görselleştirme ile oluşturulmuştur).

Optik akış (optical flow) algoritmaları, yerel hizalama aşamalarında en yaygın kullanılan yöntemler arasında yer alır. Bu algoritmalar, farklı pozlama değerlerinde çekilen görüntüler arasındaki piksel hareketini tahmin ederek bir akış haritası oluşturur. Günümüzde derin öğrenme tabanlı hizalama yaklaşımları, karmaşık hareketleri, örtüşen bölgeleri ve değişken aydınlatma koşullarını daha doğru biçimde modelleyebilme yetenekleri sayesinde geleneksel yöntemlere göre daha yüksek doğruluk sağlamaktadır. Yetersiz hizalama, hayaletlenme ya da çift görüntü gibi bozulmalara yol açarak YDA görüntü kalitesini olumsuz etkileyebilir. Bu nedenle

hizalama aşaması, YDA görüntüleme zincirinin genel başarımında belirleyici bir rol oynamaktadır. Üçüncü adım olan görüntülerin ağırlıklı birleştirme süreci, hizalanmış pozlama görüntülerinden tek bir YDA çıktısı üretmeyi amaçlar. Bu aşamada, her pozlama düzeyinden en yüksek bilgi içeren bölgeler seçilerek nihai görüntü oluşturulur. birleştirme sürecinde genellikle piksel tabanlı ağırlıklandırma stratejileri uygulanır; bu stratejiler, parlaklık, kontrast ve renk doygunluğu gibi ölçütler üzerinden her pozlama seviyesinin katkı oranını belirler. Böylece aşırı pozlanmış veya karanlık bölgeler daha düşük ağırlıkla temsil edilirken, detay açısından zengin alanlar ön plana çıkar.

Bu sürecin genel akışı Şekil 2.3'te gösterilmektedir.



Şekil 2.3. Çoklu pozlama görüntülerinin ağırlıklı birleştirme (füzyon) yoluyla YDA çıktıya dönüştürülmesi (Yapay Zekâ destekli görselleştirme ile oluşturulmuştur).

Gelişmiş ağırlıklı birleştirme teknikleri arasında çok ölçekli analiz, yapısal benzerlik ölçümleri ve frekans bileşen ayrıştırmaları yer almaktadır. Bu tür yöntemler, sahne yapısındaki ince geçişlerin korunmasına katkı sağlar. Ağırlıklı birleştirmenin ardından ortaya çıkan YDA görüntü, doğrudan ekranlarda gösterilemeyecek kadar geniş dinamik aralığa sahip olduğu için, bir sonraki adım olan ton eşleme uygulanır. Ton eşleme işlemi, YDA görüntünün DDA ekranlarda gösterilmesini mümkün kılar. Bu yöntemler küresel ve yerel olmak üzere ikiye

ayrılır. Küresel ton eşleme yöntemleri, tüm görüntüye tek bir fonksiyon uygular ve bu yönüyle hızlıdır. Ancak sahne içeriğindeki yerel kontrastları koruma konusunda sınırlıdır. Buna karşın, yerel ton eşleme yöntemleri görüntüyü parçalara ayırarak her bölgeye özel dönüşüm uygular; bu sayede yerel detaylar daha iyi korunur. Küresel ton eşleme fonksiyonu, logaritmik sıkıştırma ve gama düzeltmesi yoluyla YDA görüntünün doğal bir şekilde DDA ortamına aktarılmasını sağlar (Reinhard, 2020).

Küresel ve yerel ton eşleme yaklaşımlarının görsel karşılaştırması Şekil 2.4'te gösterilmektedir.

Şekil 2.4. Küresel ve yerel ton eşleme yöntemlerinin görsel çıktılarının karşılaştırması (Yapay Zekâ destekli görselleştirme ile oluşturulmuştur).

Son yıllarda ton eşleme süreçleri, yapay zekâ ve derin öğrenme yaklaşımlarının entegrasyonuyla önemli bir dönüşüm geçirmiştir. Bu yeni nesil yöntemlerde sahne içeriği, derin ağlar aracılığıyla analiz edilmekte; insan görsel algısıyla daha uyumlu ve içerik farkındalıklı ton dönüşümleri gerçekleştirilmektedir (Wang vd., 2022; Chen vd., 2023). Böylece görsel tutarlılık, renk doğruluğu ve algısal kalite klasik yöntemlere kıyasla belirgin biçimde artırılmıştır. Ayrıca bu yaklaşımlar, mobil cihazlar ve gerçek zamanlı sistemlerde yüksek verimlilik ve düşük hesaplama maliyeti avantajı sunmaktadır.

2.2. Derin Öğrenme Yaklaşımlarıyla YDA Görüntü Oluşturma

YDA görüntüleme alanında son yıllarda yaşanan gelişmeler, sahne bilgisinin yalnızca fiziksel pozlama farklarıyla değil, doğrudan veriden öğrenilen temsiller aracılığıyla yeniden yapılandırmasını mümkün kılmıştır. Bu yaklaşım, geleneksel

çoklu pozlama tabanlı yöntemlerden, sahne aydınlatmasını, kontrast ilişkilerini ve yapısal geçişleri öğrenebilen derin sinir ağı modellerine doğru önemli bir yönelim yaratmıştır. örneğin, görüntü sentezinde kullanılan GAN tabanlı derin üretici modeller (Goodfellow vd., 2014) ve YDA görüntüleme için derin öğrenme eğilimleri (Wang ve Yoon, 2021). Derin öğrenme tabanlı yöntemler, görüntüleme sürecini birleştirme ve ton eşleme gibi ardışık işlemler olarak değil, veriden temsili öğrenme (representation learning) problemi olarak ele alır. Bu sayede model, sahneye ilişkin parlaklık, kontrast, renk doygunluğu ve algısal tutarlılık gibi özellikleri doğrudan öğrenebilir; klasik sistemlerde ayrı yürütülen hizalama, ağırlıklandırma ve ton eşleme adımları tek bir bütünleşik ağ yapısı içinde optimize edilebilir (Liu vd., 2020; Chen vd., 2023).

Bu alandaki önemli çalışmalardan biri, Kalantari ve Ramamoorthi (2017) tarafından geliştirilen derin füzyon modelidir. Bu çalışma, çoklu pozlama görüntülerini sabit kurallar yerine öğrenilebilir ağırlıklarla birleştirerek klasik füzyon süreçlerini otomatikleştirmiş; özellikle hareketli sahnelerde ortaya çıkan hayaletlenme bozulmalarını önemli ölçüde azaltmıştır. Benzer biçimde, Eilertsen vd. (2017) tek bir DDA görüntüden YDA sahne tahmini yapabilen ilk CNN tabanlı mimarilerden birini sunmuş ve tek görüntüden YDA üretimi (SI-HDR) kavramının temelini atmıştır.

Bu doğrultuda geliştirilen çalışmalar, ağın sahne bilgisini fiziksel süreçlerle birlikte öğrenebilmesini hedeflemiştir. Liu vd. (2020), kameranın görüntüleme hattını (camera imaging pipeline) tersine çeviren derin bir model önermiştir. Bu yapı, kameranın pozlama, ton eşleme ve kuantizasyon (quantization) aşamalarını öğrenerek ters yönde modellemiş; doygun bölgelerde kaybolan bilgileri yeniden tahmin edebilmiştir. Metzler vd. (2020) ise Derin Optik YDA (Deep Optics HDR) yaklaşımıyla optik sistemleri derin öğrenmeyle birleştirmiş, fiziksel kamera modelini ve ağ parametrelerini ortak bir optimizasyon sürecinde eğitmiştir. Bu yöntemler, fiziksel modelleme ile yapay öğrenmenin birbirini tamamlayıcı biçimde kullanılabileceğini göstermiştir.

HDRUNet (Chen vd., 2021), kodlayıcı–çözücü yapıya entegre edilen gürültü azaltma (denoising) ve kuantizasyon geri çevirme (dequantization) modülleriyle ayrıntı kaybını en aza indirmiştir. FHDR (Khan vd., 2019) geri beslemeli (feedback) ağ yapısı kullanarak çok aşamalı YDA sentezi gerçekleştirmiş, önceki katman çıktılarının sonraki aşamalarda yeniden değerlendirilmesini sağlamıştır. UpHDR-GAN (Li vd., 2022) ise GAN temelli bir yapı kullanarak eşleşmemiş (unpaired)

DDA–YDA çiftleriyle YDA üretimi gerçekleştirmiş, denetimsiz YDA öğrenimini mümkün kılmıştır.

Bunun yanında, ağların yalnızca pikseller arası ilişkiyi değil, sahnenin anlamsal yapısını da öğrenebilmesini hedefleyen yeni nesil yaklaşımlar geliştirilmiştir. Yan vd. (2021) pozlamalar arası hizalama hatalarını azaltmak için dikkat tabanlı birleşim modülleri önermiş; Liu vd. (2022) kanal ve uzamsal dikkat bileşenlerini birleştirerek hem yerel hem de küresel bilgi akışını dengeli biçimde modellemiştir. Bu gelişmeler, YDA üretiminde dikkat mekanizmalarının etkinliğini açıkça ortaya koymuştur.

Semantik farkındalık ve dikkat temelli modeller, bu eğilimin en güncel aşamasını temsil etmektedir. Yan vd. (2025), sahneye ait semantik özellikleri YDA üretiminde rehber olarak kullanan bilgi aktarımı (semantic knowledge transfer) temelli bir ağ geliştirmiştir. Bu yaklaşım, doygunluk bölgelerinde bilgi kaybını azaltarak daha doğal sahne rekonstrüksiyonu sağlamıştır. Ma ve Zhang (2025), kanal ve mekânsal düzeyde çift yönlü dikkat (dual attention) uygulayarak, özellikle ışık geçiş bölgelerinde ton sürekliliğini iyileştirmiştir. Lopez-Cabrejos vd. (2025) tarafından önerilen dönüştürücü tabanlı düşük karmaşıklıklı YDA sistemi, küresel bağlamı dikkate alan öz-dikkat bloklarıyla geniş sahnelerde daha yüksek doğruluk elde etmiştir.

Sonuç olarak, modern YDA mimarileri yalnızca aydınlatma farklarını değil, sahne bağlamını, nesne ilişkilerini ve algısal nitelikleri öğrenerek insan görsel sistemine daha yakın, doğal YDA sonuçlar üretmektedir. Böylece, gerçek zamanlı, yüksek doğruluklu ve hesaplama açısından verimli YDA üretim sistemleri için sağlam bir temel oluşturulmuştur (Maity vd., 2023; Zangana vd., 2024).

2.3. Otokodlayıcı ve U-Net Tabanlı Yapılar

Derin öğrenme tabanlı YDA görüntü üretiminde, otokodlayıcı ve U-Net temelli mimariler, sahneye ait anlamlı özniteliklerin çıkarılması ve yeniden yapılandırılmasında temel rol oynar. Bu yapılar, özellikle doygun bölgelerdeki bilgi kaybı ve kontrast dengesizliği gibi problemleri veriye dayalı biçimde çözebilme avantajı sunar. Otokodlayıcılar, giriş görüntüsünü kodlayıcı aracılığıyla düşük boyutlu bir temsile dönüştürür ve çözücü yardımıyla yeniden inşa eder. Bu yapı, pozlama kaybı yaşanan sahne bölgelerinde eksik bilgilerin tamamlanmasını sağlar (Lore vd., 2017). Güncel modellerde bu yapı genellikle artık bloklar ve dikkat mekanizmaları ile güçlendirilmiştir. Örneğin Roy vd. (2018), sahnedeki bilgi yoğun

bölgeleri vurgulayan Sıkıştırma-Uyarma yapısını önermiştir.

U-Net mimarisi, otokodlayıcı yapının bir türevi olup, kodlayıcı ve çözücü arasına atlamalı bağlantılar (skip connections) ekleyerek düşük seviyeli detayların korunmasını sağlar (Ronneberger vd., 2015). Bu yapı, YDA sahnelerde kenar, doku ve ton geçişlerinin korunmasına olanak tanır. U-Net'in evrimleşmiş sürümleri olan Residual U-Net, Attention U-Net ve Nested U-Net, dikkat modülleriyle modelin önemli sahne bölgelerine odaklanmasını sağlamıştır. HDRNet (Eilertsen vd., 2017) ve HDR-DANet (Ma ve Zhang, 2025) gibi modern yapılar, bu prensipleri birleştirerek hem yapısal doğruluk hem de algısal kalite açısından yüksek performans elde etmiştir. Kodlayıcı-çözücü yapılar, parametrik verimliliği sayesinde mobil sistemlerde de uygulanabilir düzeye getirilmiş, böylece gerçek zamanlı YDA üretimi mümkün olmuştur.

Sonuç olarak, otokodlayıcı ve U-Net tabanlı mimariler, YDA görüntüleme alanında hesaplama verimliliği, detay koruma ve doğal ton sürekliliği açısından günümüzdeki derin öğrenme modellerinin temelini oluşturmaktadır.

2.4. Dikkat Mekanizmaları: Kavramsal Temeller ve Uygulama Örnekleri

Dikkat mekanizmaları genel olarak, uzamsal, kanal, karma (hybrid) ve öz dikkat olmak üzere dört temel grupta incelenmektedir. Her biri, ağın bilgi işleme sürecinde farklı bir boyutu önceliklendirmekte ve YDA görüntüleme bağlamında özgün katkılar sunmaktadır.

Kanalsal dikkat, ağın farklı özellik haritalarındaki bilgi yoğunluğunu değerlendirerek hangi kanalların daha önemli olduğuna karar verir. Bu mekanizma, özellikle YDA üretiminde renk doygunluğu ve parlaklık ilişkilerinin daha hassas biçimde modellenmesini sağlar (Mekruksavanich ve Jitpattanakul, 2023).

Hu vd. (2018) tarafından geliştirilen Sıkıştırma-Uyarma dikkat yapısı, her kanalın küresel ortalama havuzlama sonucu elde edilen özet bilgisine göre ağırlık ataması yapar ve böylece bilgi açısından zayıf kanalların etkisini azaltır. YDA modellerinde Sıkıştırma-Uyarma katmanlarının eklenmesi, özellikle aşırı pozlanmış veya gölgeli bölgelerde parlaklık tutarlılığını artırmaktadır.

Uzamsal dikkat ise, sahnedeki konumsal ilişkileri ön plana çıkararak modelin hangi bölgelere odaklanması gerektiğini belirler. Woo vd. (2018) tarafından önerilen Evrişimsel Blok Tabanlı Dikkat Modülü (Convolutional Block Attention Module -

CBAM), kanal ve uzamsal dikkat mekanizmalarını ardışık biçimde birleştirerek bu süreci iki aşamalı hale getirmiştir. YDA bağlamında, bu yaklaşım özellikle ışık geçiş bölgelerinde ve ince yapısal detaylarda modelin duyarlılığını artırır.

Karma dikkat yapıları, kanal ve uzamsal dikkat stratejilerini paralel veya etkileşimli biçimde birleştirerek, her iki bilginin avantajlarını aynı anda kullanır. Roy vd. (2018) tarafından önerilen Eşzamanlı Uzamsal ve Kanalsal Sıkıştırma - Uyarma (Concurrent Spatial and Channel Squeeze & Excitation – scSE) yapısı, bu anlayışın tipik bir örneğidir. Bu tür mekanizmalar, YDA üretiminde doygunluk bölgelerinde bilgi kaybını azaltma ve kontrast dengeleme açısından güçlü sonuçlar sunmuştur.

Öz dikkat mekanizması ise, sahne içindeki her pikselin diğer tüm piksellerle olan bağıntısını modelleyerek küresel bağlam farkındalığı kazandırır. Bu yapı, özellikle karmaşık sahnelerde uzun menzilli bağımlılıkların öğrenilmesini sağlar. Zhang vd. (2019) tarafından önerilen öz dikkat GAN (SAGAN) modeli, bu yaklaşımın YDA üretiminde kullanılabileceğini göstermiştir. Daha sonra Dosovitskiy vd. (2020) tarafından geliştirilen görü dönüştürücü mimarisi, öz dikkati tamamen konvolüsyonsuz biçimde uygulayarak görsel görevlerde çığır açmıştır. YDA sistemlerinde dönüştürücü tabanlı yaklaşımlar, küresel sahne modelleme kabiliyeti sayesinde özellikle geniş aydınlatma varyasyonlarında daha dengeli sonuçlar üretmektedir (Lopez-Cabrejos vd., 2025).

Günümüzde geliştirilen hibrit YDA sistemleri, bu dikkat türlerini bir arada kullanarak sahnenin hem yerel dokusal detaylarını hem de genel parlaklık dağılımını optimize edebilmektedir. Örneğin, Ma ve Zhang (2025) tarafından geliştirilen HDR-DANet, hem kanal hem uzamsal dikkat bileşenlerini çift yönlü olarak entegre etmiş ve ışık geçiş bölgelerinde ton sürekliliğini önemli ölçüde iyileştirmiştir.

Sonuç olarak, dikkat mekanizmaları YDA görüntüleme modellerinin bağlamsal farkındalığını, algısal doğruluğunu ve yapısal kararlılığını artıran kilit bileşenler hâline gelmiştir. Farklı dikkat türlerinin uygun biçimde birleştirilmesi, modern YDA sistemlerinde hem kalite hem de verimlilik açısından belirleyici bir rol oynamaktadır.

3. GEREÇ VE YÖNTEM

Bu bölümde, tez kapsamında önerilen YDA görüntü üretim sisteminin geliştirildiği yöntemsel altyapı ayrıntılı biçimde sunulmaktadır. Araştırmanın başlangıcında Varyasyonel Otokodlayıcı tabanlı bir mimari üzerinde çalışmalar yapılmış, ancak bu yaklaşımın YDA üretiminde beklenen performansı sağlayamaması, sonraki aşamalarda U-Net benzeri kodlayıcı-çözücü yapısının temel iskelet olarak tercih edilmesine yol açmıştır. Önerilen sistem, tek bir DDA görüntüden sanal pozlama çeşitliliği üreterek sahnenin farklı aydınlık düzeylerini dijital olarak simüle etmeyi ve bu pozlamaları birleştirerek nihai YDA çıktıyı elde etmeyi amaçlamaktadır. Bu kapsamda ilk olarak kodlayıcı-çözücü yapısının temel bileşenleri açıklanacak; ardından önerilen mimarinin katmanlı tasarımı, dikkat mekanizmalarının entegrasyonu ve öğrenme sürecinde kullanılan kayıp fonksiyonları ayrıntılı biçimde ele alınacaktır. Ayrıca, modelin eğitim sürecinde kullanılan veri kümesi ile test aşamasında tek bir giriş görüntüsünden YDA çıktının nasıl üretildiği özetlenecektir. Böylece hem genel teknik yaklaşım hem de sistemin özgün yönleri sistematik bir biçimde ortaya konulacaktır.

3.1. Derin Öğrenme Tabanlı Görüntü İşleme Sistemlerinin Temelleri

YDA görüntüleme alanında, sahne içeriğini doğru temsil edebilmek için güçlü yapısal öğrenme kapasitesine sahip mimariler kullanılmaktadır. Bu kapsamda kodlayıcı-çözücü tabanlı derin öğrenme yapıları, DDA görüntülerden YDA sahneler üretmek amacıyla yeniden yapılandırma süreçlerinde yaygın olarak tercih edilmektedir. Bu kısımda, özellikle otokodlayıcı yaklaşımları üzerinden gidilerek, önerilen sistemin dayandığı yapısal çerçeve teknik düzeyde ele alınacaktır.

3.1.1. Kodlayıcı-Çözücü Mimarileri ve Otokodlayıcı Tabanlı Görüntü Yeniden Yapılandırma

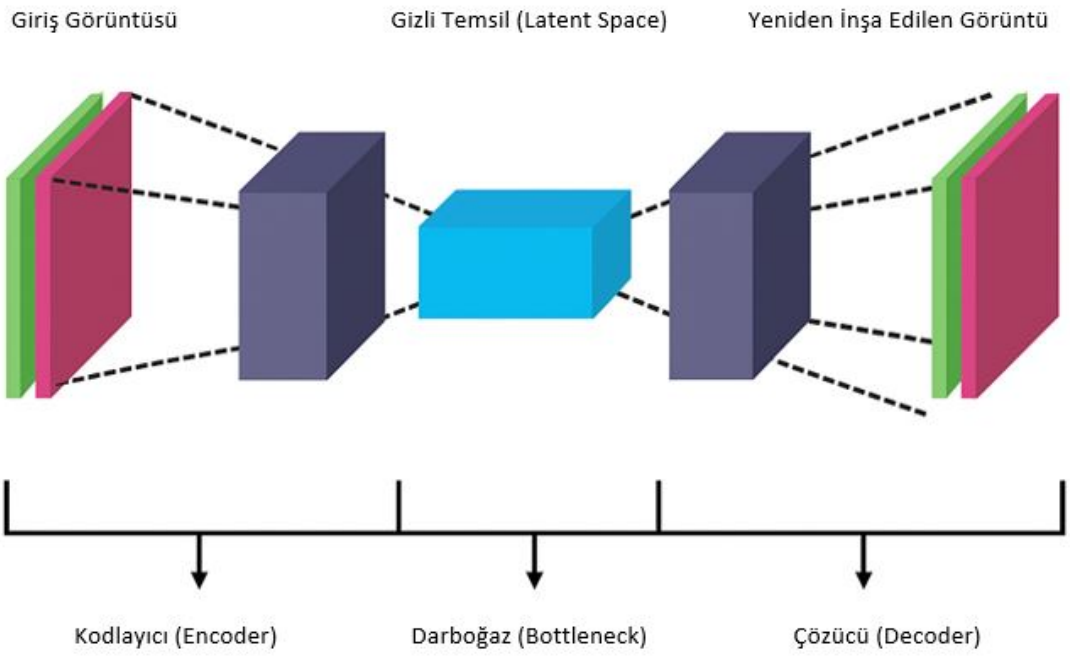
Kodlayıcı-çözücü mimarileri, özellikle görüntü işleme ve yeniden yapılandırma görevlerinde derin öğrenme modellerinin en etkili yapı taşlarından biridir. Bu yapılar, giriş görüntüsünü daha düşük boyutlu temsillere dönüştürerek sıkıştırma işlemi gerçekleştirir ve ardından bu temsilleri kullanarak görüntüyü yeniden üretmeyi amaçlar. Bu süreçte temel hedef, kayıplı sıkıştırma koşullarında bile semantik ve yapısal bütünlüğü koruyarak yüksek kaliteli bir çıktıya ulaşmaktır (Goodfellow vd., 2016).

Bu mimari iki temel bileşenden oluşur:

Kodlayıcı, giriş görüntüsünü daha düşük boyutlu bir gizli (latent) vektöre dönüştürür. Bu süreç genellikle konvolüsyonel katmanlar ve aktivasyon fonksiyonları kullanılarak gerçekleştirilir.

Çözücü, latent temsili kullanarak orijinal görüntüye yakın bir çıktı üretir. Bu süreçte transpoze konvolüsyon (deconvolution), yukarı örnekleme (upsampling) ve aktivasyon katmanları kullanılır.

Otokodlayıcı, kodlayıcı-çözücü mimarisinin denetimsiz öğrenme senaryolarındaki en yaygın uygulamasıdır. Temel olarak, giriş verisini yeniden üretmeye çalışırken arada bir darboğaz katmanı oluştururlar. Bu katman, modelin yalnızca en temel bilgileri yakalamasını zorunlu kılar. Bu yapının genel akışı Şekil 3.1'de gösterilmektedir.



Şekil 3.1. Kodlayıcı-Çözücü tabanlı bir Otokodlayıcı mimarisi (Yapay Zekâ destekli görselleştirme ile oluşturulmuştur).

Şekil 3.1'de gösterilen mimari, yalnızca girişin yeniden üretimiyle sınırlı kalmayıp, gürültü giderme, sıkıştırılmış temsil öğrenimi ve eksik veri tamamlama

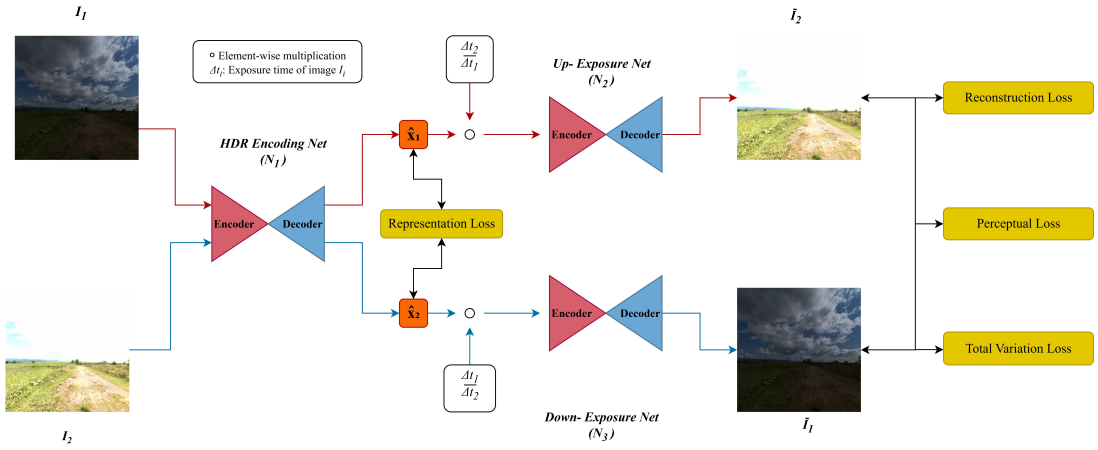
gibi görevlerde de etkili biçimde kullanılmaktadır. YDA görüntüleme bağlamında, otokodlayıcı yapılar DDA girişlerden kayıp bilgiyi semantik düzeyde geri kazanma potansiyeli sunmakta; bu yönüyle hem klasik yöntemlerin sınırlarını aşmakta hem de yüksek kaliteli yeniden yapılandırmalar üretmektedir (Le vd., 2023). Otokodlayıcı'ların öğrenme hedefi, orijinal giriş x ile yeniden üretilen çıktı $\hat{x}$ arasındaki farkı minimize etmektir. Bu fark genellikle ortalama karesel hata (MSE – Mean Squared Error) ile ölçülür:

$$\mathcal{L}(x, \hat{x}) = \|x - \hat{x}\|_2^2 \tag{3.1}$$

Bu kayıp fonksiyonu, modelin giriş verisini mümkün olan en düşük hatayla yeniden üretmesini sağlar. Kodlayıcı ağı giriş verisini düşük boyutlu bir temsile dönüştürürken, çözücü bu temsili kullanarak çıkışı oluşturur. Eğitim süreci boyunca model, bu kaybı minimize etmeye çalışır.

3.2. Le vd. (2022) Tarafından Önerilen Tek Görüntüden Yüksek Dinamik Aralıklı Görüntü Yeniden Yapılandırma Yaklaşımı

Le vd. (2022) tarafından geliştirilen "Single-Image HDR Reconstruction by Multi-Exposure Generation" yöntemi, tek bir DDA görüntüden YDA sahnelerin yeniden yapılandırılmasını amaçlayan yenilikçi bir derin öğrenme yaklaşımıdır. Yöntem, kamera görüntüleme sürecinin fiziksel modelini tersine çevirerek sahne ışınımını (irradiance) tahmin etmekte ve bu ışıma bilgisinden farklı pozlama seviyelerinde sahte görüntüler üretmektedir. Bu mimarinin genel yapısı Şekil 3.2'de gösterilmektedir.



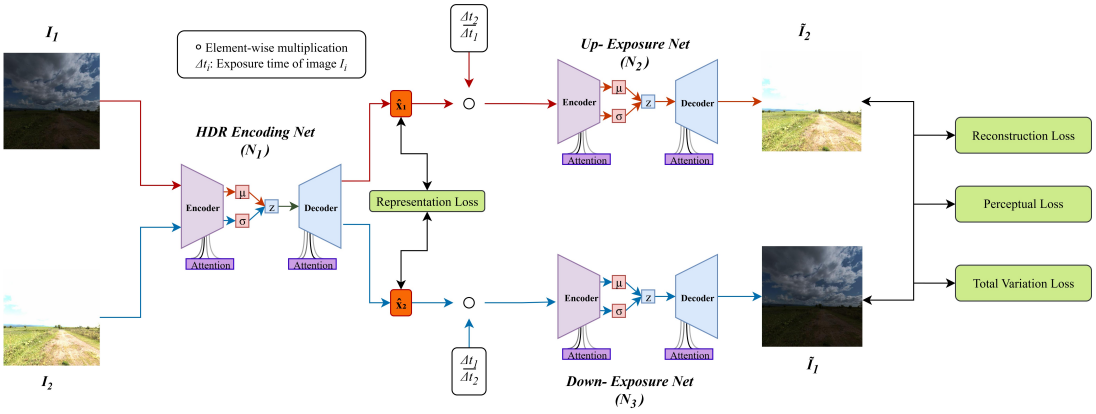
Şekil 3.2. Le vd. (2022) tarafından önerilen mimarinin genel yapısı.

Model, U-Net benzeri bir kodlayıcı–çözücü mimari üzerine inşa edilmiş ve üç ana alt ağdan oluşmaktadır:

- **HDR Encoding Net**: Girdi görüntüsünü sahne ışınımını temsil eden gizli (latent) bir temsile dönüştürür.
- **Up-Exposure Net**: Bu temsilden yüksek pozlamalı görüntüler üretir.
- **Down-Exposure Net**: Aynı temsilden düşük pozlamalı görüntüler oluşturur.

3.3. Varyasyonel Otokodlayıcı (Variational Autoencoder - VAE) Tabanlı Mimari.

Çalışmanın başlangıç aşamasında, tek görüntüden YDA üretimi için Varyasyonel Otokodlayıcı (Variational Autoencoder - VAE) tabanlı bir yaklaşım denenmiştir. Bu yapıda kodlayıcı, giriş DDA görüntüsünü ortalama (μ) ve varyans (σ) parametreleriyle tanımlanan bir gizil dağılıma dönüştürmüş; yeniden parametrizasyon (reparametrizasyon) adımı ile bu dağılımdan örneklenen gizli temsil vektörü, çözücü aracılığıyla yeniden YDA tahmini olarak üretilmiştir. Ancak yapılan deneyler, VAE'nin yüksek dinamik aralığı temsil etmede yetersiz kaldığını göstermiştir. Özellikle parlaklık geçişlerinde detay kaybı, renk doygunluğu hataları ve yapısal bozulmalar gözlenmiştir. Bu sınırlılıklar nedeniyle VAE, sonraki aşamalarda temel yöntem olarak kullanılmamış; yerine daha kararlı özellik aktarımı sunan U-Net tabanlı kodlayıcı–çözücü mimarisi tercih edilmiştir. Bu mimarinin genel yapısı Şekil 3.3'te gösterilmektedir.



Şekil 3.3. Varyasyonel Otokodlayıcı Tabanlı Mimari (Le vd., 2023'ten esinlenilmiştir).

Şekil 3.3'te gösterilen VAE tabanlı mimarinin performansı, baz alınan temel otokodlayıcı tabanlı modelle karşılaştırmalı olarak değerlendirilmiştir. Deneysel sonuçlar, VAE'nin yüksek dinamik aralıklı sahnelerde parlaklık geçişlerini ve renk doygunluğunu korumada yetersiz kaldığını göstermiştir. Özellikle PSNR ve SSIM değerlerindeki düşüş, bu mimarinin detay kayıplarına ve gürültü artışına neden olduğunu ortaya koymuştur. Çizelge 3.1'de görüldüğü üzere, VAE tabanlı mimarinin tüm metriklerde Temel (Base) modele göre daha düşük performans sergilemesi, bu yapının YDA sahne rekonstrüksiyonu için uygun olmadığını doğrulamaktadır.

Çizelge 3.1. VAE tabanlı mimarinin performans değerlendirmesi

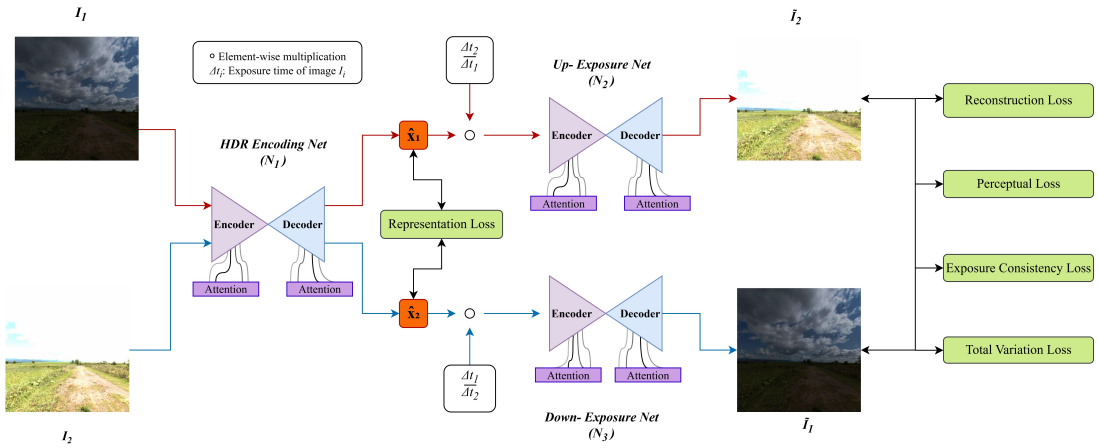
Model	PSNR (↑)	SSIM (↑)	LPIPS (↓)
VAE	20.30	0.890	0.131
Base	21.50	0.913	0.109

3.4. Önerilen YDA Üretim Sisteminin Katmanlı Mimarisi

YDA görüntüleme bağlamında tek bir DDA görüntüden yüksek kaliteli YDA sahnelerin üretimi, hem semantik bilgi hem de pozlama çeşitliliği açısından derin temsil öğrenme gerektiren zorlu bir problemdir. Bu çalışmada önerilen sistem, bu problemi çözmek amacıyla çok modüllü bir mimari yapılandırma sunmaktadır. Temel olarak U-Net benzeri bir kodlayıcı-çözücü iskeleti üzerine inşa edilen sistemde, sanal pozlama çeşitliliği üretimi hedeflenmiş ve sahnenin farklı pozlama senaryolarındaki karşılıklarının modellenmesi amaçlanmıştır. Bu bağlamda, yapı üç

ana bileşene ayrılmıştır: YDA Kodlama (HDR Encoding Net) Ağı, Artırılmış Pozlama (Up-Exposure Net) Ağı ve Azaltılmış Pozlama (Down-Exposure Net) Ağı. Her bir ağ, sahnenin belirli pozlama koşullarındaki detaylarını yeniden oluşturarak nihai YDA sentezine katkıda bulunur.Böylelikle, model tek bir DDA görüntüden farklı pozlama seviyelerine sahip görüntüler üretir.

Elde edilen bu çoklu pozlama görüntüleri, Photomatix veya benzeri bir YDA birleştirme yazılımı kullanılarak bir araya getirilir ve YDA görünüm elde edilir. Bu yaklaşım, makalede önerilen yöntemle paralel olarak, fiziksel .hdr dosyası oluşturmak yerine görsel olarak YDA karakteristiği taşıyan bir sahne rekonstrüksiyonu sağlamaktadır. Söz konusu yaklaşım, özellikle düşük pozlamadan kaynaklanan detay kayıpları veya aşırı pozlamanın neden olduğu doygunluk problemleriyle başa çıkmak üzere tasarlanmıştır. Bu çok modüllü sistemin genel yapısı Şekil 3.4'te gösterilmektedir.



Şekil 3.4. YDA Dikkat Ağı (HDR-AttNet) mimarisi (Le vd., 2023'ten esinlenilmiştir).

Şekil 3.4'te görüldüğü üzere, yapıda YDA Kodlama (HDR Encoding Net), Artırılmış Pozlama (Up-Exposure Net) ve Azaltılmış Pozlama (Down-Exposure Net) ağları ile bu alt ağlar arasında gerçekleşen pozlama oranına dayalı sanal dönüşüm işlemleri yer almaktadır. Kodlayıcı-çözücü yapısı üzerine kurulu olan her alt modül, dikkat mekanizmaları ile zenginleştirilmiş ve modelin çıktıları birden fazla kayıp fonksiyonunun ortak etkisiyle optimize edilmiştir. Bu kapsamda, rekonstrüksiyon kaybı yapısal doğruluğu, algısal kayıp görsel benzerliği, toplam varyasyon kaybı gürültü azaltımı ve pürüzsüzlüğü, temsil kaybı ara katman temsillerinin tutarlılığını, pozlama tutarlılığı kaybı ise farklı sanal pozlamalar arasında parlaklık tutarlılığını sağlamaktadır. Bu bütünsel yapı, sahne detaylarının pozlamaya duyarlı biçimde yeniden inşasını gerçekleştirirken aynı zamanda görsel

tutarlılığı da korumayı hedeflemektedir.

3.4.1. YDA Kodlama Ağı (HDR Encoding Net)

Önerilen sistemin ilk ve en temel bileşeni olan YDA Kodlama Ağı (HDR Encoding Net), giriş olarak verilen YDA görüntüyü pozlamadan bağımsız, sahneye ait yüksek seviyeli bir temsil düzlemine dönüştürmek üzere tasarlanmıştır. Bu yapı, kodlayıcı-çözücü tabanlı bir mimariye sahip olup, derin konvolüsyon katmanları aracılığıyla sahne bilgisini çok boyutlu öznitelik alanına sıkıştırır. Kodlayıcı bölümünde her seviyede sırasıyla konvolüsyon, aktivasyon (ReLU) ve Batch Normalization işlemleri gerçekleştirilirken, çözüzü tarafında yeniden örnekleme işlemleriyle birlikte konvolüsyon ve aktivasyon katmanları yer alır. Böylece, sahneye ait global yapılar ve yerel detaylar kodlayıcı katmanlarında yakalanırken, çözücü kısmı bu bilgileri kullanarak pozlama duyarlılığı yüksek ara temsiller üretir.

Mimari, toplam yedi seviyeden oluşan bir U-Net benzeri yapıdır ve her seviyede kanal sayısı arttırılarak öznitelik çözünürlüğü derinleştirilir. Kodlayıcı tarafında giriş kanal sayısı 16 ile başlarken, en derin katmanda bu sayı 256'ya ulaşır. Çözücü kısmında ise, yeniden örnekleme işlemleri alt piksel konvolüsyonu (sub-pixel convolution) yöntemiyle gerçekleştirilir. Bu yöntem, klasik dekonvolüsyon tekniklerine kıyasla hem daha düşük işlem maliyeti sunar hem de daha keskin görüntü çıktıları sağlar. Ayrıca, kodlayıcı ve çözücü blokları arasında kurulan simetrik geçiş bağlantıları (skip connections), düşük seviyeli detayların kaybolmasını engelleyerek daha kaliteli YDA çıktılar elde edilmesine katkıda bulunur.

YDA kodlama ağı, modelin sanal pozlama çeşitliliğini oluşturmasını sağlayan temel mekanizmalardan birini barındırır. Bu amaçla, aynı sahneye ait farklı pozlama seviyelerine sahip iki DDA görüntü I_1 ve I_2 kullanılarak aşağıdaki şekilde kodlama işlemi gerçekleştirilir:

$$\hat{X}_1 = \mathcal{N}_1(I_1) \tag{3.2}$$

$$\hat{X}_2 = \mathcal{N}_1(I_2) \tag{3.3}$$

Burada $\mathcal{N}_1$, YDA kodlama ağı; $\hat{X}_1$ ve $\hat{X}_2$ ise giriş görüntülerinin gizli (latent) öznitelik düzeyindeki temsillerini ifade eder. Bu temsiller, sahnenin farklı pozlama

koşullarındaki varyasyonunu simüle etmek üzere pozlama oranlarına göre normalize edilir:

$$Z_1 = \hat{X}_1 \left(\frac{\Delta t_2}{\Delta t_1} \right) \tag{3.4}$$

$$Z_2 = \hat{X}_2 \left(\frac{\Delta t_1}{\Delta t_2} \right) \tag{3.5}$$

Buradaki Δt_i, her bir giriş görüntüsüne ait pozlama süresini ifade etmektedir. Bu dönüşüm adımı sayesinde, model sahneyi sanki farklı pozlama koşullarında gözlemlemiş gibi davranabilir ve YDA sentezi için zengin ara temsiller oluşturabilir. Böylece model, eksik bilgi içeren karanlık veya doygun bölgelerde detay kurtarımı için gerekli içerik çeşitliliğini elde eder.

3.4.2. Artırılmış Pozlama Ağı (Up-Exposure Net)

YDA kodlama ağından elde edilen pozlama-normalize edilmiş öznitelik temsilleri, sahnedeki düşük aydınlatmalı (under-exposed) alanlardaki bilgi kaybını telafi etmek üzere Artırılmış pozlama ağına yönlendirilir. Bu ağın temel amacı, karanlık bölgelerde eksik kalan detayları geri kazanmak ve sahneyi daha aydınlık bir pozlama seviyesinde temsil eden ara görüntüler oluşturmaktır.

Modelde I_1 düşük pozlamalı bir giriş görüntüsü olarak ele alındığında, bu görüntüden elde edilen $\hat{X}_1 = N_1(I_1)$ öznitelik temsili pozlama oranı ile normalize edilerek aşağıdaki şekilde sanal bir dönüşüm gerçekleştirilir:

$$Z_1 = \hat{X}_1 \left(\frac{\Delta t_2}{\Delta t_1} \right) \tag{3.6}$$

Elde edilen Z_1, ardından Artırılmış Pozlama Ağı (N_2)'e giriş olarak verilir ve bu ağ daha yüksek pozlama seviyesine karşılık gelen sahne görüntüsünü üretir:

$$\hat{I}_2 = \mathcal{N}_2(Z_1) = \mathcal{N}_2 \left(\hat{X}_1 \left(\frac{\Delta t_2}{\Delta t_1} \right) \right) \tag{3.7}$$

Bu yapı da tıpkı YDA kodlama ağı gibi kodlayıcı-çözücü mimarisi üzerine kuruludur. Kodlayıcı kısmı, Z_1 içindeki düşük seviyeli karanlık detayları güçlendirerek sahne hakkında daha zengin bir temsil elde etmeye çalışırken; çözücü kısmı, bu temsil üzerinden daha parlak ve detaylı bir sahne rekonstrüksiyonu gerçekleştirmektedir.

Ağın başarısını artırmak amacıyla, kodlayıcı ve çözücü blokları arasına U-Net benzeri geçiş bağlantıları yerleştirilmiştir. Bu bağlantılar sayesinde, karanlık bölgelerdeki ince detaylar doğrudan çözücü katmanlarına aktarılır ve rekonstrüksiyon sırasında bilgi kaybı minimize edilir. Böylece, ağ sadece sahne aydınlığını artırmakla kalmaz, aynı zamanda semantik ve yapısal bütünlüğü de koruyarak YDA sentezine katkı sağlar.

Artırılmış Pozlama Ağı, YDA üretiminin özellikle düşük pozlamadan kaynaklı bilgi eksikliği yaşayan alanlarında kritik rol oynar. Yetersiz ışık alan bölgelerde sahne bilgisinin eksiksiz şekilde yeniden yapılandırılmasını mümkün kılan bu modül, YDA Dikkat Ağı (HDR-AttNet) mimarisinin aydınlık dengeleme kapasitesinin temel taşıdır.

3.4.3. Azaltılmış Pozlama Ağı (Down-Exposure Net)

Azaltılmış Pozlama Ağı (Down-Exposure Net), önerilen sistemin yüksek pozlamadan kaynaklı doygunluk problemlerini hedef alan ikinci alt modülüdür. Bu ağın temel amacı, aşırı aydınlatılmış (over-exposed) bölgelerdeki bilgi kaybını gidermek ve sahnenin daha düşük pozlamaya sahip versiyonunu yeniden inşa etmektir. Böylece, doymuş parlak alanlardaki kayıp detayların geri kazanılması sağlanır.

Modelde I_2 yüksek pozlamalı bir görüntü olarak kabul edildiğinde, YDA kodlama ağı üzerinden elde edilen öznitelik temsili aşağıdaki şekilde hesaplanır:

$$\hat{X}_2 = \mathcal{N}_1(I_2) \tag{3.8}$$

Ardından bu temsil, pozlama oranına bağlı olarak normalize edilir:

$$Z_2 = \hat{X}_2 \left(\frac{\Delta t_1}{\Delta t_2} \right) \tag{3.9}$$

Elde edilen Z_2 temsili, Azaltılmış Pozlama Ağı $(\mathcal{N}_3)$'e giriş olarak verilerek daha düşük pozlamaya sahip bir sahne görüntüsü sentezlenir:

$$\hat{I}_1 = \mathcal{N}_3(Z_2) = \mathcal{N}_3\left(\hat{X}_2\left(\frac{\Delta t_1}{\Delta t_2}\right)\right) \tag{3.10}$$

Tıpkı Artırılmış Pozlama Ağında olduğu gibi, burada da kodlayıcı-çözücü tabanlı bir yapı kullanılmıştır. Kodlayıcı tarafında, sahnedeki doygunluk seviyesini aşan bölgelerde hâlâ mevcut olan yapısal bilgiler analiz edilir. Çözücü kısmı ise bu bilgilerden daha düşük pozlamaya sahip, detaylı bir ara sahne görüntüsü oluşturmaya çalışır. Ağın mimarisinde yer alan geçiş bağlantıları, kodlayıcı katmanlarında çıkarılan düşük seviyeli yapısal bilgilerin çözücü katmanlarına doğrudan aktarılmasını sağlayarak bilgi kaybını önler. Özellikle parlak bölgelerde ayrıntı kaybının sıkça yaşandığı durumlarda bu mekanizma, rekonstrüksiyonun kalitesini artırır.

Azaltılmış Pozlama Ağı, YDA sahne üretiminde aydınlık dengeleme açısından tamamlayıcı bir rol üstlenir. Bu ağ sayesinde, YDA kodlama ağı ile modellenen sahne temsilleri daha düşük pozlama düzeyinde yeniden üretilerek doygun alanlardaki ayrıntılar geri kazanılır ve genel sahne dinamik aralığı zenginleştirilmiş olur

3.4.4. YDA Dikkat Ağı (HDR-AttNet) Katman Yapısı ve Akış Diyagramı

YDA Dikkat Ağı (HDR-AttNet) mimarisi, kodlayıcı-çözücü temelli üç ana bileşenden oluşmaktadır: YDA kodlama ağı, Artırılmış Pozlama Ağı ve Azaltılmış Pozlama Ağı. Bu alt ağların her biri, konvolüsyonel katmanlar, aktivasyon fonksiyonları, normalizasyon adımları ve belirli noktalara yerleştirilmiş dikkat modülleri ile yapılandırılmıştır. Kodlayıcı birimleri, 3×3 konvolüsyon katmanı, ReLU aktivasyonu ve Batch Normalization işlemlerinden oluşan bloklar halinde düzenlenmiştir. Yedi seviyeye kadar derinleştirilen kodlayıcı yapısında kanal sayısı girişten çıkışa doğru kademeli olarak artırılmakta (örneğin: 16 → 32 → 64 → 128 → 256) ve bu sayede düşük seviyeli kenar ve doku bilgisinden yüksek seviyeli semantik temsillere geçiş sağlanmaktadır. Çözücü katmanları ise bu temsilleri ters yönde işleyerek yüksek çözünürlüklü görüntü sentezini gerçekleştirir. yukarı örnekleme (Upsampling) adımları, alt piksel konvolüsyonu veya en yakın komşu (nearest-neighbor) interpolasyonu ile yapılmakta; bloklar konvolüsyon, aktivasyon ve normalizasyon katmanlarıyla desteklenmektedir.

YDA Kodlama Ağı, sahneyi pozlamadan bağımsız öznitelik haritaları ile temsil eder. Bu temsil, pozlama oranı (Δt) kullanılarak ölçeklenir ve iki ayrı kola yönlendirilir: $\Delta t > 0$ durumunda temsil artırılmış pozlama ağına aktarılır ve ağ, sahneyi daha uzun pozlama süresiyle çekilmiş gibi yeniden yapılandırır; $\Delta t < 0$ durumunda ise temsil azaltılmış pozlama ağına gönderilir ve sahne daha kısa pozlama süresiyle modellenir. Böylece artırılmış pozlama ağı karanlık bölgelerde eksik ayrıntıları geri kazanırken, azaltılmış pozlama ağı aşırı parlak alanlardaki doygunluğu azaltarak detay kurtarımı sağlar.

Eğitim sürecinde, model aynı sahnenin iki farklı pozlama değerine sahip DDA görüntüsünü ve bu görüntülere ait EV değerlerini giriş olarak almaktadır. YDA kodlama ağı bu görüntüleri latent temsillere dönüştürmekte, ardından iki pozlama arasındaki fark Δt hesaplanmaktadır. Bu fark, latent temsiller üzerinde ölçekleme (multiplication) işlemiyle uygulanarak farklı pozlama seviyeleri arasındaki dönüşüm öğrenilmektedir. $\Delta t > 0$ olduğunda temsil artırılmış pozlama ağına, $\Delta t < 0$ olduğunda ise azaltılmış pozlama ağına yönlendirilir. Böylece ağ, hem karanlık hem de aşırı parlak bölgelerin yeniden yapılandırılmasını öğrenmektedir.

Test aşamasında ise yalnızca tek bir DDA görüntü kullanılmaktadır. Eğitim sürecinde öğrenilen Δt-tabanlı dönüşümler bu giriş görüntüsüne uygulanarak sanal artırılmış pozlama ve azaltılmış pozlama temsilleri üretilmekte; bu varyantlar birleştirilerek nihai YDA sentezi gerçekleştirilmektedir.

YDA dikkat ağının genel akış diyagramı Şekil 3.3'te sunulmuştur. Şekilde modüller arası veri akışı, pozlama oranına göre yapılan dönüşümler ve elde edilen sahne çıktılarının kayıp fonksiyonlarıyla nasıl optimize edildiği görselleştirilmiştir. Bu bütünsel mimari, tek bir DDA görüntüden gerçek çoklu pozlama çekilmişçesine davranarak daha geniş dinamik aralığa sahip sahneler üretilmesini mümkün kılmaktadır.

3.5. Dikkat Mekanizmalarının Model Mimarisine Entegrasyonu

Derin ağların özellikle görüntü sentezi görevlerinde daha anlamlı ve detaylı temsil öğrenebilmesi için dikkat mekanizmaları önemli katkılar sağlamaktadır. Bu çalışmada, YDA üretim sisteminin hem kodlayıcı hem de çözücü bloklarına beş farklı dikkat mekanizması entegre edilmiştir: Uzamsal Dikkat, Kanalsal Dikkat, Darboğaz Dikkat, Sıkıştırma–Uyarma Dikkat ve Öz Dikkat. Her bir dikkat türü, ağın belirli özellikleri daha iyi öğrenmesini sağlayarak hem yapısal bütünlük hem de görsel kalite açısından çıktıyı iyileştirmeye yönelik olarak kullanılmıştır.

Dikkat modülleri, kodlayıcı katmanlarında daha anlamlı öznitelik çıkarımı yapılmasına, çözücü tarafında ise detayların daha etkili biçimde yeniden inşa edilmesine katkı sunar. Her bir dikkat mekanizması, önerilen YDA dikkat ağı mimarisi içine bağımsız olarak entegre edilerek ayrı model varyantları oluşturulmuştur. Bu sayede farklı dikkat türlerinin model performansı üzerindeki etkileri deneysel olarak karşılaştırılabilmiştir. Aşağıdaki alt başlıklarda, söz konusu dikkat mekanizmalarının mimari düzeyde nasıl uygulandığı detaylandırılmaktadır.

3.5.1. Uzamsal Dikkat (Spatial Attention)

Uzamsal dikkat mekanizması, sahnede konumsal olarak öne çıkan bölgelerin belirlenmesini ve bu bölgelere odaklanılmasını sağlamak amacıyla önerilen mimariye entegre edilmiştir. Bu modül, YDA dikkat ağı sisteminin hem kodlayıcı hem de çözücü katmanlarının çıkışlarında uygulanmakta olup, özellikle detay yoğunluğu yüksek alanların vurgulanmasında etkin rol oynamaktadır. Uzamsal dikkat modülünün temel işleyişi, bir öznitelik haritası üzerinde hem ortalama hem de maksimum değerlerin kanallar boyunca havuzlanmasıyla başlar. Elde edilen iki farklı 2D harita kanal ekseninde birleştirilir ve ardından 3×3 konvolüsyon uygulanarak uzamsal önem haritası çıkarılır. Bu harita, sigmoid aktivasyon fonksiyonu yardımıyla normalize edilir ve giriş öznitelik haritası ile çarpılarak sadece önemli bölgelerin öne çıkarılması sağlanır.

Uzamsal dikkat işlemi şu şekilde formüle edilir:

$$M_{spatial} = \sigma\left(f_{3\times3}\left([\mathrm{AvgPool}(F); \mathrm{MaxPool}(F)]\right)\right) \tag{3.11}$$

Burada:

- F giriş öznitelik haritasıdır.
- $AvgPool(F)$ ve $MaxPool(F)$ sırasıyla ortalama ve maksimum kanal havuzlamadır.
- $f_{3\times3}$, 3×3 konvolüsyon işlemini temsil eder.
- σ, sigmoid aktivasyon fonksiyonudur.
- $[AvgPool,(F),;MaxPool,(F)]$, kanal boyunca birleştirmeyi (concatenation) ifade eder.

Bu dikkat maskesi, giriş öznitelikleri ile çarpılarak şu şekilde ağırlıklandırılmış çıktı elde edilir:

$$F' = M_{spatial} \odot F \tag{3.12}$$

Burada $\odot$, eleman bazında çarpımı temsil eder.

Yapının kodlayıcı tarafında uygulanması, giriş görüntüsündeki dikkat gerektiren bölgelerin daha doğru şekilde temsil edilmesini sağlarken; çözücü tarafında uygulanması, sahne detaylarının daha iyi geri kazanılmasına katkıda bulunmuştur. Yapılan deneysel karşılaştırmalarda, uzamsal dikkat içeren model varyantlarının, özellikle ince yapısal detayların korunmasında daha başarılı çıktılar ürettiği gözlemlenmiştir.

3.5.2. Kanalsal Dikkat (Channel Attention)

Kanalsal dikkat mekanizması, her bir öznitelik kanalının göreceli önemini modelleyerek, YDA sahnelerde renk doğruluğu ve parlaklık tutarlılığını artırmayı hedeflemektedir. Bu modül, önerilen mimaride hem kodlayıcı hem de çözücü bloklarının sonuna entegre edilerek öznitelik haritalarının kanal düzeyindeki ağırlıklarının dinamik olarak yeniden ölçeklendirilmesini sağlamıştır. Bu dikkat türü, giriş öznitelik haritası F'yi kanallar boyunca küresel ortalama havuzlamaya (Global Average Pooling) tabi tutarak sıkıştırılmış bir kanal tanımlayıcı vektör s üretir. Ardından bu vektör iki tam bağlantılı (fully connected – FC) katmandan geçirilerek kanal dikkat vektörü elde edilir. Bu işlem aşağıdaki gibi formüle edilir:

$$s = GAP(F) \in \mathbb{R}^C \tag{3.13}$$

$$Z = \sigma\left(W_2 \cdot \mathrm{ReLU}(W_1 \cdot s)\right) \in \mathbb{R}^C \tag{3.14}$$

Burada:

- $GAP(\cdot)$, küresel ortalama havuzlamaya işlemidir.
- $W_1 \in \mathbb{R}^{\frac{C}{r} \times C}$, ilk tam bağlantılı katman ağırlıklarıdır.
- $W_2 \in \mathbb{R}^{C \times \frac{C}{r}}$, ikinci tam bağlantılı katman ağırlıklarıdır.
- σ, sigmoid aktivasyon fonksiyonudur.
- r, sıkıştırma oranını temsil eder (genellikle 16'dır).

Elde edilen kanal dikkat vektörü Z, orijinal öznitelik haritasıyla kanal bazında çarpılarak girişin yeniden ölçeklendirilmiş hâli F' elde edilir:

$$F'_c = Z_c \cdot F_c \quad \forall c \in [1, C]$$

(3.15)

Bu çarpım kanal düzeyinde gerçekleştiği için yalnızca her kanalın genel katkısı yeniden ayarlanır; böylece model, daha anlamlı olan kanalları vurgulayarak, daha az önemli olanları bastırabilir. YDA dikkat ağı yapısına entegre edilen kanalsal dikkat mekanizması, özellikle sahnelerde renk dengesinin korunması ve ışık yoğunluğunun doğru yansıtılması açısından önemli katkılar sağlamıştır. Model varyantları karşılaştırıldığında, bu dikkat türü özellikle renk geçişlerinde yumuşaklık ve parlak alanlarda stabilite sağlamasıyla öne çıkmıştır.

3.5.3. Darboğaz Dikkat (Bottleneck Attention)

Darboğaz Dikkat mekanizması, yüksek boyutlu öznitelik temsillerinde yer alan gereksiz veya tekrar eden bilgileri filtreleyerek, yalnızca kritik yapısal bilgilerin öne çıkarılmasını sağlayan bir mekanizmadır. YDA dikkat ağı mimarisine bu modül, özellikle kodlayıcı ve çözücü arasındaki geçiş noktalarına yerleştirilmiş ve bilgi yoğunluğunun en fazla olduğu derin katmanlarda kullanılmıştır. Bu Mimari Şekil 3.5'te gösterilmiştir. Bu dikkat türü, giriş öznitelik haritasını $F \in \mathbb{R}^{C \times W \times H}$ olarak kabul eder ve önce kanallar boyunca küresel ortalama havuzlama işlemi uygular. Ardından, elde edilen özet vektör iki adet tam bağlantılı katmandan geçirilerek bir dikkat skoru vektörü hesaplanır:

$$s = GAP(F) \in \mathbb{R}^C \tag{3.16}$$

$$M_{bottleneck} = \sigma\left(W_2 \cdot \mathrm{ReLU}(W_1 \cdot s)\right) \in \mathbb{R}^C \tag{3.17}$$

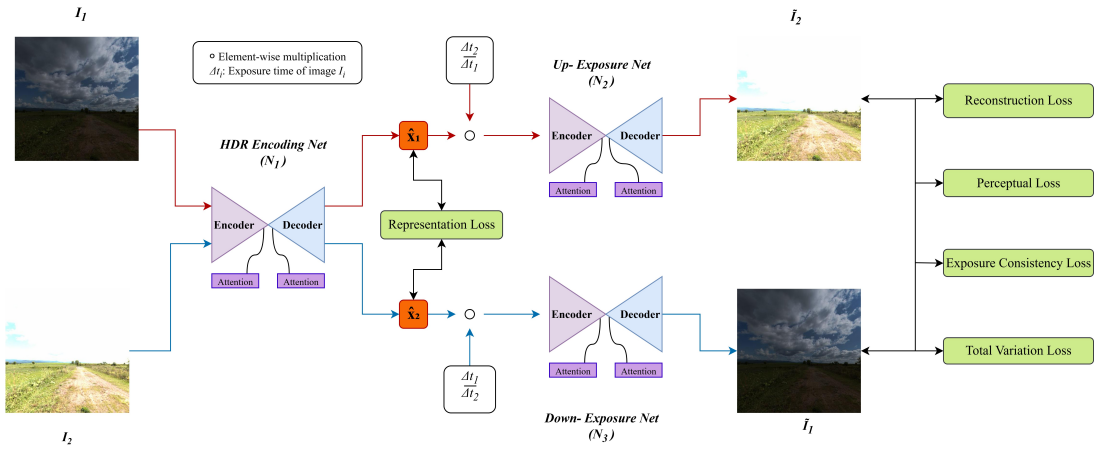
Burada:

- $W_1 \in \mathbb{R}^{\frac{C}{r} \times C}$, ilk tam bağlantılı katmanın ağırlıklarıdır.
- $W_2 \in \mathbb{R}^{C \times \frac{C}{r}}$, ikinci tam bağlantılı katmanın ağırlıklarıdır.
- σ, sigmoid aktivasyon fonksiyonudur.
- r, kanal sıkıştırma oranıdır (örneğin 16).

Daha sonra, bu dikkat maskesi giriş öznitelikleriyle kanal bazında çarpılarak yeniden ölçeklendirilmiş çıktı üretilir:

$$F'_c = M_{bottleneck,c} \cdot F_c \quad \forall c \in [1, C] \tag{3.18}$$

Bu yapı sayesinde ağ, yalnızca önemli yapısal bilgileri geçirmekte; önemsiz, gürültülü veya tekrarlı kanalları baskılamaktadır. Darboğaz dikkat, özellikle YDA kodlama ağının derin katmanlarında öznitelik yoğunluğunun yüksek olduğu bölgelerde modelin bilgi akışını optimize etmek amacıyla kullanılmıştır.



Şekil 3.5. YDA Darboğaz Dikkat Ağı mimarisi (Le vd., 2023'ten esinlenilmiştir).

3.5.4. Sıkıştırma–Uyarma Dikkat (Squeeze-and-Excitation Attention)

Sıkıştırma–Uyarma Dikkat mekanizması, kanal boyutundaki öznitelik ilişkilerini modelleyerek, her kanalın göreli önemini öğrenmeye dayalı bir yaklaşımdır. YDA dikkat ağı mimarisinde bu modül, kodlayıcı ve çözücü katmanlarının sonlarında uygulanarak her kanalın katkısının adaptif olarak yeniden ağırlıklandırılmasını sağlamıştır. Böylece, ağın renk tutarlılığı, parlaklık dengesi ve yapısal detaylar açısından daha güçlü temsil öğrenmesi mümkün olmuştur. Sıkıştırma–Uyarma dikkat modülü iki temel aşamadan oluşur: sıkıştırma (squeeze) ve uyarma (excitation). İlk olarak, giriş öznitelik haritası üzerinde kanal boyutunda global ortalama havuzlama uygulanarak kanal tanımlayıcı vektör elde edilir:

$$s = GAP(F),\ s \in \mathbb{R}^C \tag{3.19}$$

Ardından bu vektör, iki adet tam bağlantılı katmandan geçirilerek kanal ağırlık vektörü oluşturulur:

$$w = \sigma\left(W_2 \cdot \mathrm{ReLU}(W_1 \cdot s)\right) \in \mathbb{R}^C \tag{3.20}$$

Burada:

- $W_1 \in \mathbb{R}^{\frac{C}{r} \times C}$ ve $W_2 \in \mathbb{R}^{C \times \frac{C}{r}}$ tam bağlantılı katmanların ağırlıklarıdır.
- r, sıkıştırma oranıdır (genellikle 16),
- σ, sigmoid aktivasyon fonksiyonudur.

Son olarak, elde edilen kanal ağırlık vektörü giriş öznitelik haritası ile çarpılarak yeniden ölçeklendirilmiş öznitelikler üretilir:

$$F'_c = w_c \cdot F_c \quad \forall c \in [1, C]$$

(3.21)

Bu işlem sayesinde, daha anlamlı ve katkı sağlayan kanallar öne çıkarılırken; önemsiz ya da gürültülü kanalların etkisi azaltılır.

Sıkıştırma–Uyarma Dikkat mekanizması, önerilen mimaride özellikle renk geçişlerinin yumuşatılması, ışık dağılımının dengelenmesi ve YDA görüntülerde görsel kalite kayıplarının önlenmesi amacıyla kullanılmıştır.

3.5.5. Öz Dikkat (Self Attention)

Öz Dikkat mekanizması, sahnedeki uzak konumlar arasında uzun menzilli bağıntıların modellenmesini sağlayarak, ağın daha bütüncül ve bağlamsal bilgiye dayalı temsil üretmesine olanak tanır. Bu özellik, YDA görüntü sentezinde sahnenin genel yapısının korunması ve anlamlı bölgeler arasındaki ilişkilerin doğru kurulması açısından son derece önemlidir. YDA dikkat ağı mimarisinde öz dikkat modülü, kodlayıcı katmanlarının orta ve derin seviyelerine entegre edilerek global bağlamsal farkındalık kazandırılmıştır.

Öz Dikkat mekanizması, giriş öznitelik haritasından üç ayrı temsil çıkararak çalışır: Sorgu (Query), Anahtar (Key), Değer (Value) matrisleri. Bu matrisler, W_Q, W_K, W_V ağırlıkları kullanılarak aşağıdaki gibi elde edilir:

$$Q = W_Q \cdot F$$

(3.22)

$$K = W_K \cdot F \tag{3.23}$$

$$V = W_V \cdot F \tag{3.24}$$

Daha sonra bu temsil birimleri kullanılarak dikkat matrisi hesaplanır:

$$Attention(Q, K, V) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right) V \tag{3.25}$$

Burada:

- $Q, W, V \in \mathbb{R}^{d_k \times N}$, öznitelik haritalarının yeniden düzenlenmiş temsilleridir (burada $N = H \times W$),
- d_k, sorgu ve anahtar matrislerinin boyutudur (ölçekleme faktörü),
- $softmax$, her konumdaki dikkat ağırlıklarını normalleştirir.

Bu yapı sayesinde, sahnedeki her bir konum, tüm diğer konumlarla ilişkili olarak yeniden ağırlıklandırılır. Yani bir pikselin temsilini üretirken yalnızca kendi çevresini değil, tüm sahneyi dikkate alır. Öz Dikkat, bu yönüyle konvansiyonel konvolüsyon işlemlerine göre daha kapsamlı bağlamsal temsil üretir. YDA dikkat ağı içerisine entegre edilen öz dikkat modülü, özellikle büyük nesneler, yumuşak geçiş bölgeleri ve sahneler arası tutarlılığın sağlanmasında etkili olmuştur.

3.6. Kayıp Fonksiyonları ve Öğrenme Kriterleri

Derin öğrenme tabanlı YDA görüntü üretiminde, sadece piksel düzeyinde doğruluk değil, aynı zamanda algısal tutarlılık, yapısal süreklilik ve sahneye dair anlamlı temsillerin korunması da son derece önemlidir. Bu hedefleri dengelemek amacıyla önerilen YDA dikkat ağ modelinin eğitimi sırasında birden fazla kayıp

fonksiyonu birlikte optimize edilmiştir. Her bir kayıp fonksiyonu, YDA sentezinin farklı bir yönüne odaklanarak nihai çıktının hem teknik hem de görsel olarak yüksek kalitede olmasını sağlamaktadır. Modelin eğitim süreci beş temel kayıp fonksiyonu ile yönlendirilmiştir: Yeniden Yapılandırma Kaybı (Reconstruction Loss), Algısal Kayıp (Perceptual Loss), Toplam Varyasyon Kaybı (Total Variation Loss), Temsil Kaybı (Representation Loss) ve Pozlama Tutarlılığı Kaybı (Exposure Consistency Loss). Bu kayıplar, uygun ağırlık katsayıları ile toplam kayıp fonksiyonuna dahil edilerek modelin çok yönlü öğrenme başarısı desteklenmiştir.

3.6.1. Yeniden Yapılandırma Kaybı (Reconstruction Loss)

Yeniden Yapılandırma Kaybı (Reconstruction Loss), önerilen YDA dikkat ağ modelinin ürettiği sahte pozlamalı DDA görüntüler ile karşılık gelen gerçek DDA görüntüler arasındaki piksel düzeyindeki farkı minimize etmeyi amaçlar. Bu kayıp fonksiyonu, ağın çıktılarının doğrudan doğruluk düzeyini ölçen temel metriklerden biridir ve modelin düşük seviyeli yapısal doğruluğunu korumasına olanak tanır.

Bu çalışmada yeniden yapılandırma kaybı olarak ortalama mutlak hata (Mean Absolute Error – MAE), yani ℓ_1 normu kullanılmıştır. ℓ_1 kaybı, ℓ_2 normuna kıyasla kenar yapılarının korunmasında daha etkilidir ve yüksek genlikli hatalara karşı daha az duyarlıdır. Bu özelliği nedeniyle YDA görüntülerdeki keskin geçişlerin ve parlaklık değişimlerinin daha doğru temsil edilmesine katkı sağlar.

Modelin eğitiminde kullanılan yeniden yapılandırma kaybı şu şekilde tanımlanır:

$$\mathcal{L}_{rec} = \left\| \hat{I} - I \right\|_1 \tag{3.26}$$

Burada:

- $\hat{I}$, model tarafından üretilen (sahte) DDA görüntüyü,
- I, karşılık gelen gerçek DDA hedef görüntüyü temsil eder,
- $\|\cdot\|_1$, piksel başına ortalama mutlak farkı ifade eder.

Bu kayıp fonksiyonu hem artırılmış pozlama $(\hat{I}_2)$ hem de azaltılmış pozlama $(\hat{I}_1)$ tahminleri için ayrı ayrı hesaplanabilir ve toplam kayba ağırlıklı olarak dahil edilir. yeniden yapılandırma kaybı sayesinde model, ürettiği görüntülerin piksel düzeyinde gerçeğe yakın olmasını öğrenir ve YDA üretim sürecinin doğrusal bileşenlerinde yüksek hassasiyet sağlar.

3.6.2. Algısal Kayıp (Perceptual Loss)

Algısal Kayıp (Perceptual Loss), klasik piksel bazlı kayıp fonksiyonlarının ötesine geçerek, insan görsel sistemiyle daha uyumlu bir değerlendirme yapmayı amaçlar. Piksel düzeyinde küçük farklılıklar teknik olarak yüksek benzerlik gösterebilirken, bu farklar insan gözüyle algılandığında belirgin hale gelebilir. Algısal Kayıp bu boşluğu kapatarak, modelin daha anlamlı, detaylı ve gerçekçi görüntüler üretmesini sağlar. Bu kayıp fonksiyonu, modelin ürettiği görüntü ile gerçek görüntü arasındaki farkı derin bir ağdan (VGG-19 gibi) çıkarılan ara öznitelik haritaları düzeyinde ölçer. Böylece sadece piksel benzerliği değil, aynı zamanda semantik ve yapısal benzerlik de gözetilir. Bu yöntem özellikle doku kalitesi, kenar netliği ve stil uyumu gibi yüksek seviyeli özelliklerin korunmasında etkilidir.

Önerilen modelde algısal kayıp şu şekilde tanımlanır:

$$\mathcal{L}_{perc} = \sum_{l} \left\| \phi_l(\hat{I}) - \phi_l(I) \right\|_2^2 \tag{3.27}$$

Burada:

- $\phi_l(\cdot)$, önceden eğitilmiş VGG-19 ağının l. katmanından çıkarılan öznitelik haritasıdır,
- $\hat{I}$, model tarafından üretilen görüntüdür,
- I, karşılık gelen hedef görüntüdür,
- $\|\cdot\|_2^2$, kareli ℓ_2 normunu ifade eder.

Bu kayıp, genellikle VGG-19 ağının "conv1_2", "conv2_2", "conv3_4" gibi erken ve orta seviye katmanlarında hesaplanır. Bu katmanlar görsel detayları hem

yerel hem de global ölçekte yansıttığı için YDA yeniden yapılandırma sürecine görsel kalite açısından anlamlı katkı sağlar. algısal kayıp, YDA dikkat ağı mimarisi içerisinde yalnızca yeniden yapılandırma doğruluğunu değil, aynı zamanda görsel estetiği ve gerçekçiliği hedefleyen bir optimizasyon bileşeni olarak önemli rol oynamaktadır. Özellikle yapay sınırlar, keskin geçişler ve dokusal kayıpların önlenmesinde bu kaybın etkisi deneysel olarak gözlemlenmiştir.

3.6.3. Toplam Varyasyon Kaybı (Total Variation Loss)

Toplam Varyasyon Kaybı (Total Variation Loss), görüntü çıktılarındaki yüksek frekanslı parazitleri, yapay geçişleri ve keskin kenar dışı bozulmaları azaltmak amacıyla kullanılan düzenleyici bir kayıp fonksiyonudur. YDA görüntüleme sürecinde, özellikle pozlama dönüşümleri sonrası oluşabilecek yapay desenleri (checkerboard artefact gibi) baskılayarak görsel pürüzsüzlük sağlar. Toplam varyasyon kaybı, bir görüntüdeki yatay ve düşey komşu pikseller arasındaki parlaklık farklarını minimize etmeye çalışır. Böylece, modelin daha düzgün geçişler üretmesi ve YDA görüntünün doğal bir görünüm kazanması sağlanır.

Matematiksel olarak toplam varyasyon kaybı şu şekilde tanımlanır:

$$\mathcal{L}_{tv} = \sum_{i,j} \left(\left| \hat{I}_{i+1,j} - \hat{I}_{i,j} \right|^2 + \left| \hat{I}_{i,j+1} - \hat{I}_{i,j} \right|^2 \right) \tag{3.28}$$

Burada:

- $\hat{I}$, model tarafından üretilen görüntüyü temsil eder,
- i, j, piksel koordinatlarıdır,
- Toplam varyasyon, hem yatay (sağ) hem de düşey (aşağı) yönlü parlaklık değişimlerini içerir.

Bu kayıp, modelin detayları tamamen düzleştirmeden, yalnızca ani ve yapay geçişleri yumuşatacak şekilde öğrenmesine yardımcı olur. Görsel kalitenin artırılması, YDA sahnelerde özellikle düşük pozlamalı alanlarda yumuşak geçişlerin sağlanması ve gereksiz parazitlerin engellenmesi adına oldukça etkilidir. YDA dikkat

ağı mimarisinde toplam varyasyon kaybı, özellikle çözücü tarafında üretilen çıktılar üzerinde uygulanmış ve diğer kayıplarla birlikte toplam optimizasyon fonksiyonuna entegre edilmiştir. Böylece model, yalnızca yapısal doğruluğa değil, aynı zamanda görsel akıcılığa da odaklanmıştır.

3.6.4. Temsil Kaybı (Representation Loss)

Temsil Kaybı (Representation Loss), YDA kodlama ağı tarafından üretilen öznitelik temsillerinin pozlamadan bağımsız olarak tutarlı kalmasını sağlamak amacıyla kullanılan bir yapısal kayıptır. Bu kayıp, farklı pozlama seviyelerine sahip giriş görüntülerinden elde edilen latent temsillerin birbirine yakın olmasını teşvik ederek, sahne içeriğinin pozlama değişiminden etkilenmeden öğrenilmesini sağlar. Bu kaybın temel varsayımı, aynı sahneden farklı pozlama değerleriyle elde edilen görüntülerin, aynı semantik içeriği taşımasıdır. Dolayısıyla, modelin bu farklı pozlamalardan çıkardığı latent öznitelik temsilleri de birbirine yakın olmalıdır. Bu tutarlılığı sağlamak üzere ℓ_2 normuna dayalı bir eşleme kaybı tanımlanmıştır:

$$\mathcal{L}_{rep} = \left\| \hat{X}_1 - \hat{X}_2 \right\|_2^2 \tag{3.29}$$

Burada:

- $\hat{X}_1 = N_1(I_1)$, düşük pozlamalı görüntüden elde edilen öznitelik temsili,
- $\hat{X}_2 = N_1(I_2)$, yüksek pozlamalı görüntüden elde edilen öznitelik temsili,
- N_1, YDA kodlama ağını ifade eder.

Bu kayıp fonksiyonu sayesinde model, aynı sahneye ait farklı pozlama seviyelerinde öğrenilen temsilleri yakınsatarak, pozlama bağımlılığını minimize eder. Özellikle pozlama oranlarına göre normalleştirilmiş temsillerin oluşturulmasında bu kaybın etkisi belirgin hale gelir. Temsil kaybı, YDA sahnelerin içeriksel bütünlüğünün korunması açısından son derece kritiktir. Modelin, yalnızca görüntü çıktılarında değil, aynı zamanda öz nitelik düzeyinde semantik tutarlılık sağlamasını

mümkün kılar. Bu yapı sayesinde YDA dikkat ağı, sahnenin pozlama değişimlerine karşı daha kararlı ve genellenebilir bir öğrenme performansı sergiler.

3.6.5. Pozlama Tutarlılığı Kaybı (Exposure Consistency Loss)

Pozlama Tutarlılığı Kaybı (Exposure Consistency Loss), YDA Dikkat Ağı modelinde farklı pozlama seviyelerinde üretilen sahte görüntülerin ışık ölçeği açısından tutarlı olmasını sağlamak amacıyla tanımlanmıştır.

Bu kaybın temel varsayımı, aynı sahnenin farklı pozlama sürelerinde (t_1,t_2) elde edilen temsillerinin aydınlatma bakımından orantılı olması gerektiğidir.

Yani model, düşük pozlamalı ($\hat{I}_1$) ve yüksek pozlamalı ($\hat{I}_2$) sahte görüntüler ürettiğinde, bu iki görüntü pozlama süresiyle normalize edildiğinde benzer sahne parlaklık dağılımını korumalıdır.

$$L_{\text{exp}} = \left\| \frac{\hat{I}_1}{t_1} - \frac{\hat{I}_2}{t_2} \right\|_1 \tag{3.30}$$

Burada:

$\hat{I}_1$: modelin azaltılmış pozlama ağı (Down-Exposure Net) üzerinden ürettiği sahte görüntü,

$\hat{I}_2$: modelin artırılmış pozlama ağı (Up-Exposure Net) üzerinden ürettiği sahte görüntü,

t_1 ve t_2 : bu görüntülere karşılık gelen pozlama süreleri (EV) değerleridir,

$\|\cdot\|_1$: piksel başına ortalama mutlak fark (MAE) işlemini göstermektedir.

4. BULGULAR

Bu bölümde, önerilen YDA Dikkat Ağı modeline ait deneysel bulgular ayrıntılı olarak sunulmaktadır. Modelin farklı dikkat mekanizmalarıyla zenginleştirilmiş varyantları ayrı ayrı eğitilmiş ve hem nicel (PSNR, SSIM, LPIPS) hem de nitel (görsel karşılaştırmalar) ölçütler üzerinden değerlendirilmiştir. İlk olarak, eğitim süreci, kullanılan donanım altyapısı ve yapılandırma parametreleri açıklanmaktadır. Ardından, literatürdeki yöntemlerle karşılaştırmalı performans analizleri verilmekte; ablasyon çalışmaları aracılığıyla her bir dikkat modülünün ve eklenen kayıp fonksiyonlarının modele katkısı sistematik olarak incelenmektedir. Ayrıca, parametre sayısı ve işlem süreleri gibi karmaşıklık analizleri sunularak modelin verimliliği ortaya konulmuştur. Son kısımda ise çeşitli sahneler üzerinde yapılan görsel karşılaştırmalar, modelin sayısal metriklerin ötesinde detay koruma, renk tutarlılığı ve algısal kalite bakımından sağladığı üstünlükler gözlemlenmiştir. Elde edilen sonuçlar, YDA Dikkat Ağının hem yapısal doğruluk hem de algısal kalite açısından güçlü bir performans sergilediğini ve mevcut literatürdeki yöntemlere kıyasla anlamlı iyileştirmeler sunduğunu göstermektedir.

4.1. Eğitim Yapılandırması ve Deneysel Ortam

Bu çalışmada, önerilen YDA Dikkat Ağı modeli YDA sahne yeniden yapılandırması amacıyla, YDA yeniden yapılandırma araştırmalarında yaygın olarak kullanılan standart kıyaslama (benchmark) setlerinden biri olan DrTMO veri kümesi üzerinde eğitilmiştir. Veri kümesi, toplam 1.043 YDA sahne içermekte olup, her sahne dokuz farklı pozlama değeri (EV) ile DDA formatına dönüştürülmüştür. Böylece sahnelerin hem aydınlık hem de karanlık bölgeleri için zengin çeşitlilik elde edilmiştir. Toplamda 46.935 DDA görüntü eğitim, 6.210 DDA görüntü ise test için ayrılmıştır. Tüm görüntüler eğitim sürecinde 512×512 piksel çözünürlüğe yeniden boyutlandırılmıştır. DrTMO veri kümesinin önemli bir avantajı, sahne çeşitliliğinin yüksek olması ve her YDA kaynağının çoklu pozlama varyantları ile desteklenmesidir. Bu özellik, hem modelin genellenebilirliğini artırmakta hem de literatürde yer alan farklı yöntemlerle adil karşılaştırma yapılmasına olanak tanımaktadır.

Model eğitimi, her biri ortalama 31.500 adım içeren toplam 6 epoch boyunca, yaklaşık 200.000 step olacak şekilde yapılandırılmıştır. Eğitim süreci, 16 GB bellek kapasiteli NVIDIA Tesla V100-SXM2 GPU üzerinde gerçekleştirilmiş ve tüm deneyler PyTorch çerçevesi kullanılarak kodlanmıştır. Beş farklı dikkat mekanizmasıyla oluşturulan tüm model varyantları tekil olarak eğitilmiş olup, eğitim

süreleri modelin karmaşıklığına göre değişiklik göstermiştir (Çizelge 4.1). Modelin eğitimi ayrıca beş farklı kayıp fonksiyonu ile yönlendirilmiş ve her bir kayıp için uygun ağırlık katsayıları atanmıştır. Eğitim parametrelerinin ayrıntıları Çizelge 4.2'de sunulmaktadır.

Eğitim sürecinde model, aynı sahneye ait iki farklı pozlama seviyesinde DDA görüntüyü girdi olarak almıştır. Bu giriş çiftleri, sahnedeki pozlama farkını öğrenebilmesi amacıyla rastgele seçilmiştir. Model, bu iki pozlama arasındaki ilişkiyi kullanarak ışık yoğunluğu, kontrast ve renk doğruluğu açısından tutarlı bir temsili öğrenmiştir. YDA Kodlama Ağı sahneye ait temel aydınlatma bilgisini kodlarken, Artırılmış Pozlama ve Azaltılmış Pozlama alt ağları farklı pozlama seviyelerine ait detayları yeniden yapılandıracak biçimde optimize edilmiştir. Bu yapı sayesinde model, tek bir sahneden farklı pozlama düzeylerini tahmin edebilecek şekilde genellenmiştir.

Test aşaması, önerilen YDA Dikkat Ağı modelinin tek bir DDA giriş görüntüden sahnenin farklı pozlama seviyelerindeki varyantlarını tahmin etme yeteneğini değerlendirmek amacıyla gerçekleştirilmiştir. Bu aşamada model, yalnızca bir DDA görüntü ve hedef pozlama değeri bilgisi ile çalışmaktadır. Model, verilen hedef pozlama oranına göre sahnenin ya daha aydınlık (over-exposed) ya da daha karanlık (under-exposed) varyantını tahmin eder. Böylece aynı sahne için birden fazla pozlama seviyesinde sentetik DDA görüntü oluşturulur. Test işlemleri, 16 GB bellek kapasiteli NVIDIA Tesla V100-SXM2 GPU üzerinde gerçekleştirilmiş olup, model tek bir DDA girdi için sanal pozlama görüntülerini ortalama 49 saniyede üretmiştir.

YDA kodlama modülü, giriş görüntüden sahnenin temel aydınlatma bilgisini ve yapısal özelliklerini çıkarır. Ardından Artırılmış Pozlama ve Azaltılmış Pozlama alt ağları, bu temsilleri hedef pozlama değerine göre dönüştürerek farklı aydınlatma koşullarına ait görüntü varyantları üretir. Test sürecinde, model tarafından tahmin edilen çoklu pozlama görüntüleri, Photomatix YDA birleştirme (HDR merge) algoritması kullanılarak tek bir YDA sahneye dönüştürülür. Bu işlem, farklı pozlama düzeylerinden gelen parlaklık ve gölge bilgilerinin dengeli biçimde birleştirilmesini sağlamaktadır.

Elde edilen YDA görüntü, görünür aralığa dönüştürülmek üzere Reinhard küresel ton eşleme operatörüyle normalize edilir. Test aşamasında herhangi bir referans (ground-truth) YDA görüntü kullanılmamıştır. Model, yalnızca tek bir DDA girdi ve hedef pozlama değeri üzerinden sahnenin yüksek dinamik aralıklı karşılığını tahmin etmiştir. Bu yapı, önerilen YDA Dikkat Ağı mimarisinin tek görüntüden YDA sahne yeniden yapılandırması görevinde hem aydınlatma tutarlılığını hem de görsel ayrıntı bütünlüğünü başarılı biçimde koruyabildiğini göstermektedir. Model, 16 GB bellek kapasiteli NVIDIA Tesla V100-SXM2 GPU üzerinde test edilmiştir. Tek bir DDA girdi için dokuz farklı sanal pozlama varyantı ortalama 49 saniyede üretilmiş olup, bu süre modelin tek görüntüden çoklu pozlama sentezleme verimliliğini göstermektedir.

Çizelge 4.1. YDA Dikkat Ağı (HDR-AttNet) Model Karmaşıklığı ve Eğitim Süreleri

Model	Eğitilebilir Parametre (Trainable Params)	Toplam Parametre (Total Params)	Model Boyutu	Temel (Base) + Dikkat (dk / saat)	Temel (Base) + Dikkat+ Pozlama Tutarlılığı Kaybı (dk / saat)
Temel (Base)	62.2 M	96.9 M	387 MB	960 dk (16 saat)	–
Uzamsal Dikkat	67.7 M	102 M	410 MB	2231 dk (37.2 saat, 1g 13s)	2255 dk (37.6 saat, 1g 13.5s)
Kanalsal Dikkat	62.9 M	97.6 M	390 MB	1342 dk (22.4 saat)	1366 dk (22.8 saat)
Darboğaz Dikkat	87.7 M	122 M	489 MB	1367 dk (22.8 saat)	1391 dk (23.2 saat)
Sıkıştırma-Uyarma Dikkat	62.9 M	97.6 M	390 MB	1296 dk (21.6 saat)	1320 dk (22.0 saat)
Öz Dikkat	65.6 M	100 M	401 MB	1392 dk (23.2 saat)	1416 dk (23.6 saat)

Çizelge 4.2. Eğitim Yapılandırması ve Hiperparametreler

Parametre	Değer
Eğitim Veri Kümesi (Training Dataset)	DrTMO
Toplam Eğitim Adımı (Total Training Steps)	200,000
Epoch Sayısı (Total Epochs)	6
Epoch Başına Adım (Steps per Epoch)	~31,500
Batch Boyutu (Batch Size)	8
Optimizasyon Yöntemi (Optimizer)	Adam
Öğrenme Oranı (Learning Rate)	0.0001
Kayıp Ağırlıkları (Loss Weights)	λ_1=100.0, λ_2=1.0, λ_3=0.1, λ_4=0.00001
Donanım (Hardware)	NVIDIA Tesla V100-SXM2 (16 GB)
Geliştirme Ortamı (Framework)	PyTorch

Modelin üç alt ağı (YDA Kodlama, Artırılmış Pozlama ve Azaltılmış Pozlama Ağı) 7 seviyeli U-Net benzeri kodlayıcı-çözücü mimarilerinden oluşmaktadır. Her seviyede iki adet 3×3 konvolüsyon katmanı, ardından batch normalization ve ReLU (bazı bloklarda Leaky ReLU) aktivasyon fonksiyonları uygulanmıştır. YDA kodlama modülünde kanal sayısı 16'dan başlayarak 256'ya kadar artırılırken; Artırılmış Pozlama ve Azaltılmış Pozlama modüllerinde bu sayı 32'den 512'ye kadar çıkmaktadır. Yukarı örnekleme işlemleri, klasik dekonvolüsyon yerine alt ağlar konvolüsyon yöntemi ile gerçekleştirilmiş; bu sayede hem ayrıntı korunmuş hem de hesaplama verimliliği artırılmıştır. Atlamalı bağlantılar ve normalizasyon katmanları hariç tutulduğunda, her bir alt ağ toplamda 28 konvolüsyon katmanından oluşmaktadır. Bu alt mimarilerin genel yapısal özeti Çizelge 4.3'te sunulmaktadır.

Çizelge 4.3. YDA Dikkat Ağı (HDR-AttNet) Alt Ağlarının Yapısal Özeti

AltAğ(Sub-Network)	Seviye Sayısı (Levels)	Başlangıç Kanalı (Initial Channels)	Maksimum Kanal (Max Channels)	Upsampling Türü (Upsampling Type)	ÇıkışKanalı (Output Channels)
YDA Kodlama Ağı	7	16	256	Sub-pixel	3
Artırılmış Pozlama Ağı	7	32	512	Sub-pixel	3
Azaltılmış Pozlama Ağı	7	32	512	Sub-pixel	3

4.2. Performans Değerlendirmeleri ve Metrik Karşılaştırmaları

Üretilen YDA sahnelerin kalitesini nesnel biçimde değerlendirmek amacıyla, hem yapısal hem de algısal tutarlılığı ölçen çeşitli metrikler kullanılmıştır. Bu çalışmada, model çıktılarının doğruluğunu analiz etmek için üç temel metrik tercih edilmiştir: PSNR, SSIM, LPIPS. Bu metriklerin her biri, tek bir DDA giriş görüntüsünden elde edilen YDA sahnelerin farklı yönlerini değerlendirmeyi amaçlamaktadır. Karşılaştırmalı analizlerde herhangi bir dikkat mekanizması içermeyen referans yapı olan Temel Model (Base Model) ile diğer dikkat varyantları karşılaştırılmıştır. PSNR, tahmin edilen YDA görüntü ile referans YDA görüntü arasındaki piksel düzeyindeki farkı ölçerek yapısal doğruluğu nicel olarak ortaya koyar. Özellikle parlaklık doğruluğu ve genel yapısal tutarlılık açısından önemli bir göstergedir (Sara vd., 2019). Buna karşılık, SSIM metriği, parlaklık, kontrast ve yapısal bileşenleri aynı anda dikkate alarak insan görsel algısına daha yakın bir benzerlik ölçümü sunar (Wang vd., 2004). Öte yandan, LPIPS metriği derin sinir ağları tarafından öğrenilmiş öznitelik temsillerini karşılaştırarak hesaplanmakta; bu yönüyle piksel düzeyinin ötesine geçerek insan görsel sistemine daha yakın bir algısal benzerlik ölçümüne olanak tanımaktadır (Zhang vd., 2018).

4.2.1. PSNR ile Yapısal Doğruluk Analizi

PSNR, yeniden yapılandırılan bir görüntü ile referans görüntü arasındaki piksel düzeyindeki benzerliği ölçen en yaygın nesnel değerlendirme metriklerinden biridir. Bu metrik, ortalama karesel hata (MSE) temel alınarak hesaplanır ve sonuçlar desibel (dB) cinsinden ifade edilir. PSNR değeri ne kadar yüksekse, yeniden yapılandırılan görüntünün kalite seviyesi o kadar iyidir ve referans görüntüye olan yapısal benzerliği o derece yüksektir.

PSNR değeri aşağıdaki şekilde hesaplanmaktadır:

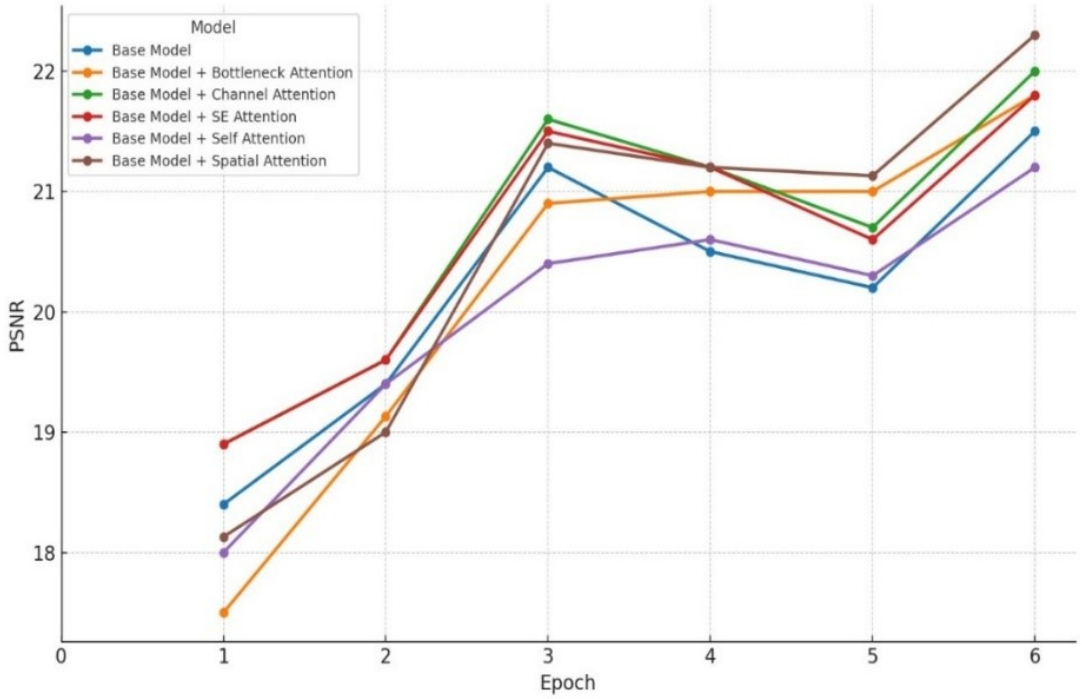
$$PSNR(I, \hat{I}) = 10 \cdot \log_{10} \left(\frac{(MAX)^2}{\frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \left(I(i,j) - \hat{I}(i,j) \right)^2} \right) \quad (4.1)$$

Burada:

- $I(i,j)$, referans YDA görüntüsünün (i,j) piksel konumundaki değerini,
- $\hat{I}(i,j)$, yeniden yapılandırılan görüntünün aynı konumdaki piksel değerini,
- m, n, görüntünün boyutlarını,
- MAX_I mümkün olan maksimum piksel değerini (örneğin normalize edilmiş görüntüler için 1.0 veya 8-bit görüntüler için 255) ifade etmektedir.

Denklemin paydasındaki ifade, referans ve yeniden üretilmiş görüntü arasındaki ortalama karesel hatayı (MSE) göstermektedir. Bu hata azaldıkça, PSNR değeri yükselmekte ve görüntü kalitesi iyileşmektedir.

Her bir dikkat mekanizması varyantı için elde edilen ortalama PSNR performansı Şekil 4.1'de gösterilmektedir.



Şekil 4.1. Farklı dikkat mekanizması varyantları için ortalama PSNR performansının karşılaştırılması.

Daha yüksek PSNR değerleri, yeniden yapılandırılan YDA sahnelerin referans görüntülere yapısal olarak daha yakın olduğunu ve daha az bozulma içerdiğini göstermektedir. Bu durum, modelin parlaklık bilgisini koruma ve gürültü azaltma açısından etkinliğini yansıtmaktadır.

4.2.2. SSIM ile Görsel Tutarlılık Analizi

SSIM, referans görüntü ile yeniden yapılandırılmış görüntü arasındaki algısal benzerliği, yapısal bilgiye dayalı olarak değerlendiren gelişmiş bir görüntü kalite metriğidir. PSNR yalnızca piksel düzeyindeki farklara odaklanırken, SSIM bunun ötesine geçerek parlaklık, kontrast ve yapısal bileşenleri birlikte analiz eder. Bu yönüyle insan görsel algısına daha yakın bir değerlendirme sağlar.

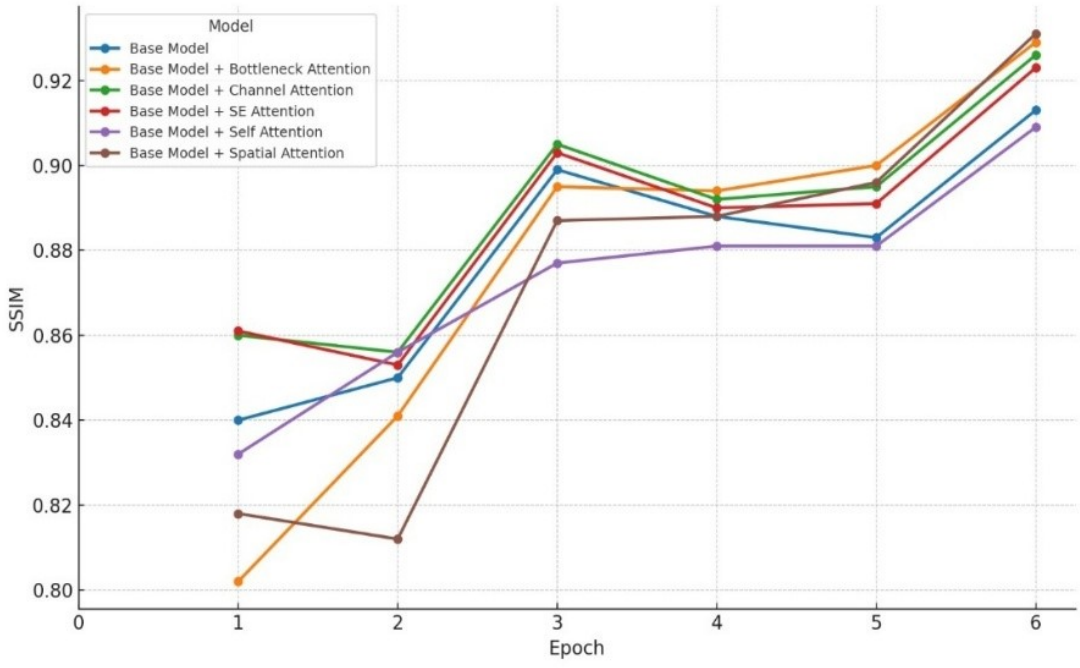
SSIM metriği aşağıdaki şekilde hesaplanmaktadır:

$$SSIM(I, \hat{I}) = \frac{(2\mu_I\mu_{\hat{I}} + C_1)(2\sigma_{I\hat{I}} + C_2)}{(\mu_I^2 + \mu_{\hat{I}}^2 + C_1)(\sigma_I^2 + \sigma_{\hat{I}}^2 + C_2)} \quad (4.2)$$

Burada:

- μ_I, $\mu_{\hat{I}}$: referans ve yeniden yapılandırılmış görüntülerin ortalama parlaklık değerlerini,
- σ_I^2, $\sigma_{\hat{I}}^2$: ilgili varyansları,
- $\sigma_{I\hat{I}}$: iki görüntü arasındaki kovaryansı,
- C_1 ve C_2: zayıf payda koşullarında bölmeyi stabilize eden sabitleri temsil eder.

SSIM değerleri 0 ile 1 arasında değişir ve 1'e daha yakın değerler, yeniden yapılandırılan görüntünün referans görüntüye daha yüksek yapısal benzerliğe ve görsel tutarlılığa sahip olduğunu gösterir.Her bir model varyantının sahne içindeki yapısal tutarlılığı, ortalama SSIM skorları üzerinden değerlendirilmiştir. Elde edilen sonuçlar Şekil 4.2'de gösterilmektedir.



Şekil 4.2. Farklı dikkat mekanizması varyantları için ortalama SSIM skorlarının karşılaştırmalı dağılımı.

Özellikle uzamsal dikkat modülü ile yapılandırılan model, yerel detayları koruma açısından belirgin bir avantaj sağlamış ve bu metrikte en yüksek başarıyı

elde etmiştir.

4.2.3. LPIPS ile Algısal Kalite Analizi

LPIPS, önceden eğitilmiş derin sinir ağlarından çıkarılan yüksek seviyeli öznitelik temsillerini karşılaştırarak görüntüler arasındaki algısal benzerliği ölçen bir metriktir. PSNR ve SSIM gibi klasik metrikler yalnızca piksel düzeyindeki farklara odaklanırken, LPIPS doğrudan öz nitelik (feature) uzayında çalışır ve görüntülerin insan görsel sistemi tarafından nasıl algılandığına daha yakın bir değerlendirme sunar. Bu metrik özellikle derin öğrenme tabanlı görüntü üretim ve iyileştirme çalışmalarında algısal kaliteyi değerlendirmek için daha dayanıklı ve güvenilir bir ölçüt olarak kabul edilmektedir.

LPIPS metriği aşağıdaki şekilde hesaplanmaktadır:

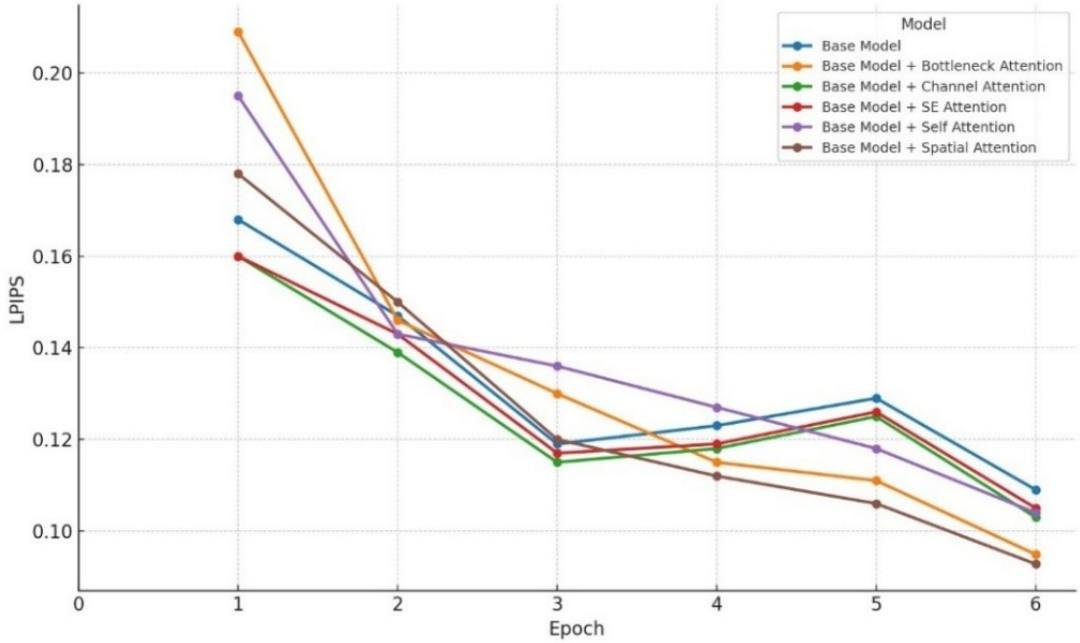
$$LPIPS(I,\hat{I}) = \sum_{l} \frac{1}{H_l W_l} \sum_{h=1}^{H_l} \sum_{w=1}^{W_l} \left\| w_l \odot \left(\phi_l(I)_{hw} - \phi_l(\hat{I})_{hw} \right) \right\|_2^2 \quad (4.3)$$

Burada:

- $\varphi_l(\cdot)$: önceden eğitilmiş bir ağın (örneğin VGG) l. katmanından çıkarılan öznitelik haritalarını,
- H_l, W_l: bu katmandaki öznitelik haritasının boyutlarını,
- w_l: kanal önemini ayarlamak için öğrenilmiş ağırlıkları,
- $\odot$: eleman bazlı çarpımı (element-wise multiplication) ifade etmektedir.

LPIPS skoru ne kadar düşükse, yeniden yapılandırılan görüntü, referans görüntüyle algısal olarak o kadar benzerdir. Bu, insan gözünün sahneyi nasıl algıladığına daha yakın bir eşleşme anlamına gelir. Model çıktılarının insan görsel sistemiyle ne derece uyumlu olduğunu değerlendiren bu metrikte elde edilen sonuçlar Şekil 4.3'te gösterilmektedir.

Elde edilen sonuçlar, dikkat mekanizmalarının entegrasyonunun model başarımını tüm değerlendirme metrikleri açısından anlamlı düzeyde artırdığını ortaya koymaktadır. Özellikle uzamsal dikkat modülü, PSNR, SSIM ve LPIPS metriklerinde en yüksek başarıyı göstererek, modelin hem yapısal hem de algısal doğruluğa sahip YDA sahneler üretebildiğini göstermiştir. PSNR skorlarının yüksek yeniden yapılandırma doğruluğunu ortaya koymasının yanı sıra, bu skorların LPIPS değerleriyle tutarlı olması, yapısal başarımın aynı zamanda görsel kalite artışı ile de desteklendiğini göstermektedir.



Şekil 4.3. Farklı dikkat mekanizması varyantları için ortalama LPIPS skorlarının karşılaştırmalı dağılımı.

Düşük LPIPS skorları, sahnelerin daha doğal, yumuşak geçişli ve estetik açıdan tatmin edici şekilde yeniden üretildiğini ortaya koymaktadır.

4.3. Ablasyon Analizi

Bu bölümde, önerilen YDA Dikkat Ağı mimarisine entegre edilen beş farklı dikkat mekanizmasının model üzerindeki etkisi sistematik olarak incelenmiştir. Her bir dikkat mekanizması, temel model üzerine tekil olarak entegre edilmiş ve bağımsız model varyantları oluşturulmuştur. Elde edilen varyantların performansı, nesnel değerlendirme metrikleri olan PSNR, SSIM ve LPIPS değerleri üzerinden karşılaştırılmıştır. Karşılaştırmalı analiz sonuçları Çizelge 4.4'te sunulmaktadır.

Çizelge 4.4. Dikkat mekanizmalarının performans sonuçları

Model	PSNR (↑)	SSIM (↑)	LPIPS (↓)
Temel (Base) Model	21.50	0.9130	0.1090
Temel (Base) Model + Uzamsal Dikkat	**22.30**	**0.9310**	**0.0928**
Temel (Base) Model + Kanalsal Dikkat	<u>22.00</u>	0.9260	0.1030
Temel (Base) Model + Darboğaz Dikkat	21.80	<u>0.9290</u>	<u>0.0949</u>
Temel (Base) Model+Sıkıştırma-Uyarma Dikkat	21.80	0.9230	0.1050
Temel (Base) Model + Öz Dikkat	21.20	0.9090	0.1040

En yüksek başarı gösteren skorlar kalın, ikinci en iyi skorlar altı çizili olarak gösterilmiştir.

Ayrıca, modelin farklı bileşenlerinin katkısını daha ayrıntılı olarak incelemek amacıyla, aynı varyantlara Pozlama Tutarlılığı Kaybı eklenmiş ve bu ek kayıp fonksiyonunun etkisi bağımsız olarak değerlendirilmiştir. Böylece yalnızca dikkat mekanizmalarının getirdiği iyileştirmeler ile, dikkat mekanizmaları ve ek kayıp fonksiyonu birlikte kullanıldığında elde edilen sonuçlar karşılaştırmalı olarak analiz edilmiştir. İlgili bulgular Çizelge 4.5'te sunulmaktadır.

Çizelge 4.5. Dikkat mekanizmaları ve Pozlama Tutarlılığı Kaybı sonrası performans sonuçları

Model	PSNR (↑)	SSIM (↑)	LPIPS (↓)
Temel (Base) Model	21.50	0.9130	0.1090
Temel (Base) Model + Uzamsal Dikkat + Pozlama Tutarlılığı Kaybı	**22.40**	**0.9310**	**0.0906**
Temel (Base) Model + Kanalsal Dikkat + Pozlama Tutarlılığı Kaybı	<u>22.00</u>	0.9300	0.1020
Temel (Base) Model + Darboğaz Dikkat + Pozlama Tutarlılığı Kaybı	21.80	<u>0.9300</u>	<u>0.0946</u>
Temel (Base) Model+Sıkıştırma-Uyarma Dikkat + Pozlama Tutarlılığı Kaybı	21.80	0.9280	0.1030
Temel (Base) Model + Öz Dikkat + Pozlama Tutarlılığı Kaybı	21.20	0.9170	0.1040

En yüksek başarı gösteren skorlar kalın, ikinci en iyi skorlar altı çizili olarak gösterilmiştir.

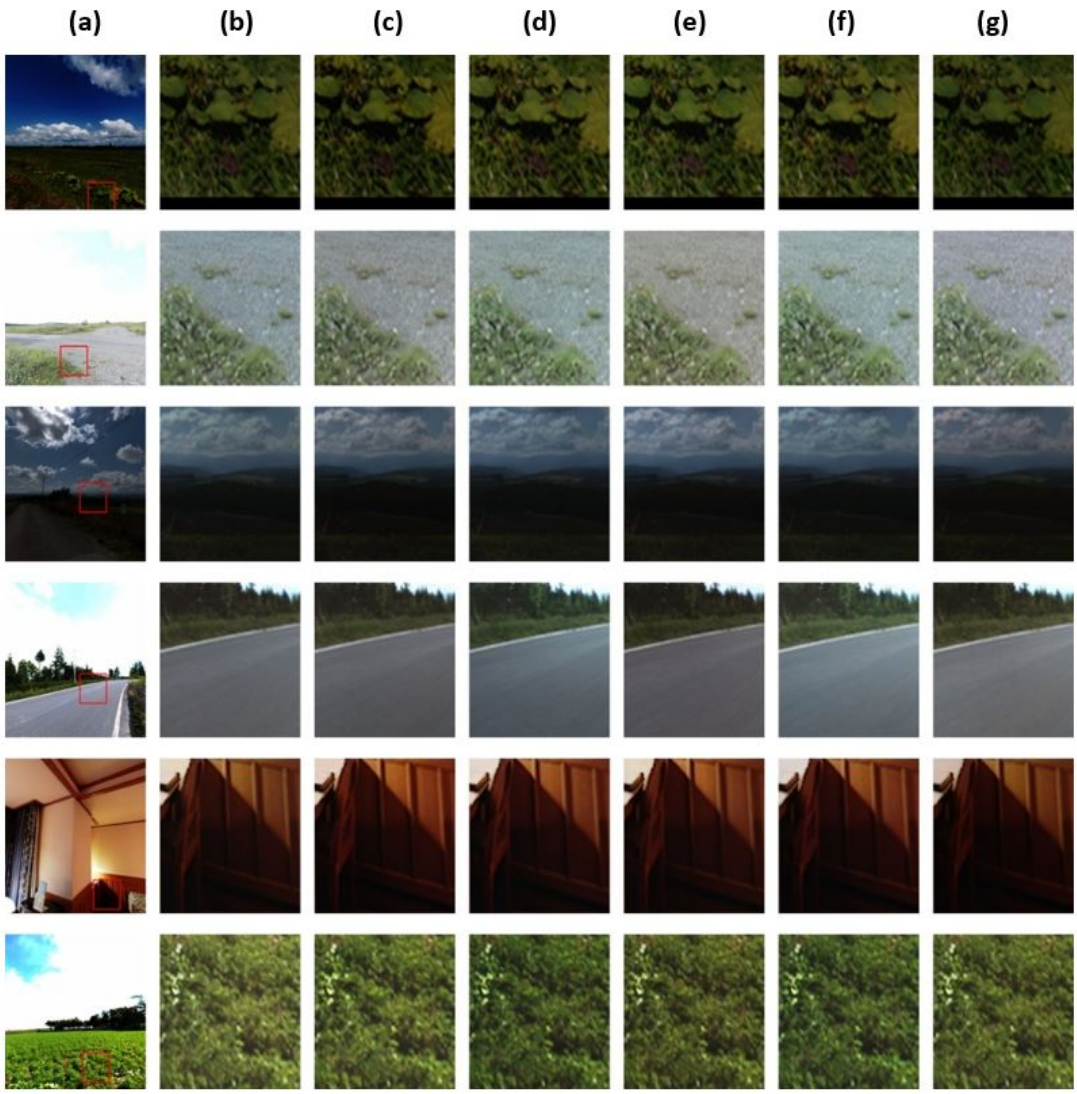
Analizler sonucunda, özellikle Uzamsal Dikkat mekanizmasının tüm metrikler açısından en belirgin iyileştirmeyi sağladığı görülmüştür. Bununla birlikte, Pozlama Tutarlılığı Kaybı eklenmesi, özellikle Uzamsal Dikkat ile birlikte uygulandığında ek performans artışı sağlamış ve modelin pozlamalar arası tutarlılığı daha iyi öğrenmesine katkıda bulunmuştur.

4.4. Görsel Karşılaştırmalar

Sayısal metriklere dayalı değerlendirmenin yanı sıra, farklı dikkat mekanizması varyantları tarafından üretilen YDA görüntülerin görsel kalitesi de analiz edilmiştir. Şekil 4.4'te, çeşitli sahnelere ait giriş DDA görüntüleri ile birlikte,

her bir dikkat mekanizması ile üretilmiş YDA çıktılar karşılaştırmalı olarak sunulmuştur. Her sahne için iyileşmenin belirgin olduğu bölgeler kırmızı kutular ile işaretlenmiş ve bu alanlar ayrıca yakınlaştırılmış görüntüler ile detaylı incelenmiştir. Yapılan karşılaştırmalar, özellikle uzamsal dikkat mekanizmasının keskin kenarların korunması, doku detaylarının artırılması ve renk doygunluğunun iyileştirilmesi açısından üstün performans gösterdiğini açıkça ortaya koymaktadır. Öte yandan, kanalsal dikkat mekanizması renk dengesi açısından güçlü bir başarı sergilemiş; öz dikkat modülü ise sahne bütünlüğünü korumada, özellikle küresel bağlam bilgisi açısından en istikrarlı sonuçları üretmiştir.

Görsel detayların korunması yalnızca sayısal metrikler açısından değil, insan algısı temelli değerlendirme açısından da son derece değerlidir. Bu durum, önerilen modelin gerçek dünya uygulamalarında kullanılabilirliğini ve görsel kaliteye olan katkısını açıkça göstermektedir.



Şekil 4.4. Farklı dikkat mekanizması varyantları ile üretilen YDA çıktılara ait görsel karşılaştırmalar. Her sütun sırasıyla şu çıktılara karşılık gelmektedir: (a) Giriş DDA görüntüsü, (b) Temel (Base) Model, (c) Temel (Base) Model + Uzamsal Dikkat, (d) Temel (Base) Model + Kanalsal Dikkat, (e) Temel (Base) Model + Darboğaz Dikkat, (f) Temel (Base) Model + Sıkıştırma-Uyarma Dikkat, (g) Temel (Base) Model + Öz Dikkat. İlk satırda tam boyutlu görüntüler, ikinci satırda ise kırmızı kutularla işaretlenen bölgelerin yakınlaştırılmış halleri sunulmuştur.

4.5. Literatür Karşılaştırmaları

YDA yeniden yapılandırma literatüründe, özellikle tek görüntüden YDA üretimi üzerine son yıllarda önemli ilerlemeler kaydedilmiştir. Erken dönem çalışmalarda, çoklu pozlama tabanlı yaklaşımlar baskın olsa da (Endo vd., 2017; Lee vd., 2018), bu yöntemler hareketli sahnelerde hayaletlenme bozulmaları ve hizalama sorunları nedeniyle sınırlı başarı göstermiştir. Daha sonraki CNN tabanlı yöntemler

(Liu vd., 2020) tek görüntüden YDA üretiminde dikkate değer gelişmeler sağlamış olsa da, kontrast doğruluğu ve ince detayların korunması açısından yetersiz kalmıştır. Son yıllarda geliştirilen otokodlayıcı tabanlı mimariler (Le vd., 2023), hem görsel kaliteyi hem de PSNR, SSIM ve LPIPS gibi nesnel metrikleri iyileştirerek literatürde yeni bir seviye ortaya koymuştur. Bu çalışmalar, özellikle düşük pozlamalı bölgelerdeki bilgi kaybını azaltma ve aydınlık alanlardaki doygunluğu kontrol etme konusunda önemli katkılar sağlamıştır.

Çizelge 4.6'da, literatürde rapor edilen yöntemlerin DrTMO veri kümesi üzerindeki performans karşılaştırmaları sunulmuştur; tablodaki YDA Dikkat Ağı (HDR-AttNet) sonuçları ise, beş farklı dikkat mekanizmasıyla eğitilen model varyantları arasından en yüksek başarıyı gösteren Uzamsal Dikkat mekanizmasına aittir. Burada PSNR, SSIM ve LPIPS gibi yaygın metrikler dikkate alınmıştır. En yüksek başarı gösteren skorlar kalın olarak belirtilmiştir. Bu sonuçlar, literatürdeki genel eğilimi özetlemenin yanı sıra, önerilen YDA Dikkat Ağı modelinin katkısını değerlendirmek için bir referans noktası sağlamaktadır. Bu karşılaştırmada, hem literatürde temel alınan Le vd. (2023) çalışması hem de önerilen YDA Dikkat Ağı modeli 8 batch size değeri altında çalıştırılmış ve elde edilen sonuçlar bu koşullar altında rapor edilmiştir. PSNR, SSIM ve LPIPS değerleri, eğitim sürecinde her epoch sonunda doğrulama (validation) kümesi üzerinde hesaplanan performans sonuçlarından elde edilmiştir.

Çizelge 4.6. DrTMO veri kümesinde literatür yöntemleri ile önerilen YDA Dikkat Ağı (HDR-AttNet) modelinin performans karşılaştırması.

Yöntem	Veri Kümesi	PSNR (↑)	SSIM (↑)	LPIPS (↓)
Deep Recursive HDRI (Lee vd., 2018)	DrTMO	19.56	0.7920	0.2096
Deep Reverse Tone Mapping Network (Endo vd., 2017)	DrTMO	21.60	0.8493	0.1592
Deep Reverse Imaging Network (Liu vd., 2020)	DrTMO	19.77	0.7832	0.2001
HDR-SYNTH-trained DRIN (Liu vd., 2020)	HDR-SYNTH	21.58	0.8333	0.1427
Exposure Dual Network (Le vd., 2023)	DrTMO	21.50	0.9130	0.1090
YDA Dikkat Ağı (HDR-AttNet)	DrTMO	**22.40**	**0.9310**	**0.0906**

En yüksek başarı gösteren skorlar kalın olarak gösterilmiştir.

Çizelge 4.6'da görüldüğü üzere, önerilen YDA Dikkat Ağı modeli, literatürde yer alan diğer yöntemlere kıyasla genel olarak daha yüksek yapısal benzerlik (SSIM) ve daha düşük algısal hata (LPIPS) değerleri elde etmiştir. Model, sahne içerisindeki hem aydınlık hem gölge bölgelerdeki detayları daha etkin biçimde koruyarak renk doğruluğu ve kontrast sürekliliği açısından güçlü bir performans sergilemiştir. Bu iyileştirmelerde, Uzamsal Dikkat mekanizmasının uzamsal detay farkındalığını artırması ve Pozlama Tutarlılığı Kaybı bileşeninin farklı pozlama seviyeleri arasındaki parlaklık tutarlılığını sağlaması önemli rol oynamıştır. Sonuç olarak, önerilen YDA Dikkat Ağı mimarisi, tek görüntüden YDA yeniden yapılandırma görevinde hem nesnel metrikler hem de algısal kalite bakımından literatürdeki mevcut yöntemlere göre daha dengeli ve başarılı sonuçlar sunmuştur.

5. TARTIŞMA

Bu bölümde, önerilen YDA Dikkat Ağı mimarisine entegre edilen beş farklı attention mekanizmasının ve Pozlama Tutarlılığı Kaybı bileşeninin, YDA sahne yeniden yapılandırma üzerindeki etkileri kapsamlı biçimde tartışılmaktadır. Analizler, sayısal metrik sonuçları (PSNR, SSIM, LPIPS) ve görsel incelemeler birlikte değerlendirilmiştir. Elde edilen bulgular, her bileşenin modelin farklı yönlerini geliştirdiğini, ancak bu katkıların algısal kalite, yapısal doğruluk ve aydınlatma tutarlılığı açısından değişken etkiler oluşturduğunu ortaya koymuştur.

Uzamsal Dikkat mekanizması, YDA sahne yeniden yapılandırma görevinde en yüksek performansı sağlamıştır. Bu yapı, modelin sahnedeki yerel bölgeleri konumsal öneme göre ağırlıklandırmasını sağlar; yani model, nerede daha fazla dikkat göstermesi gerektiğini öğrenir. Bu sayede kenar, doku ve gölge geçişleri gibi kritik görsel detaylar daha doğru şekilde yeniden yapılandırılmıştır. Sonuçlarda gözlenen yüksek SSIM ve PSNR değerleri, modelin detayları bozmadan sahnenin genel yapısını koruyabildiğini göstermektedir. Ayrıca düşük LPIPS skorları,Uzamsal Dikkatin yalnızca yapısal doğruluğu değil, aynı zamanda algısal bütünlüğü de geliştirdiğini kanıtlamaktadır. Görsel incelemelerde, bu varyantın özellikle kontrast geçişlerinde daha pürüzsüz tonlamalar ve renk doygunluğu açısından daha doğal sonuçlar verdiği görülmüştür. Bu nedenle Uzamsal Dikkat, modelin hem detay hassasiyetini hem de görsel tutarlılığını optimize eden en etkili bileşen olarak öne çıkmıştır.

Kanalsal Dikkat mekanizması, sahnedeki her bir kanalın (R, G, B) önemini adaptif biçimde öğrenerek renk doğruluğunu güçlendirmeyi hedeflemiştir. Bu yapı sayesinde model, özellikle yüksek kontrastlı sahnelerde renk kaymalarını azaltmış ve doygunluk seviyesini dengelemiştir. Bu mekanizma, parlak ve gölge alanlar arasında fotometrik tutarlılığı koruyarak renk tonlarının doğal görünmesini sağlamıştır. Nicel sonuçlar, SSIM değerlerinde belirgin bir iyileşme olduğunu göstermekte; görsel değerlendirmeler ise renk bozulmalarının azaldığını ortaya koymaktadır. Bununla birlikte, Kanalsal Dikkat yapısının uzamsal bilgiye doğrudan odaklanmaması, modelin doku ve kenar keskinliğini artırma kapasitesini sınırlamıştır. Dolayısıyla, bu mekanizma renksel doğruluk açısından yüksek katkı sunarken, sahne detaylarının vurgulanması açısından Uzamsal Dikkat mekanizmasına kıyasla daha sınırlı bir etki göstermiştir.

Darboğaz Dikkat modülü, ağın derin katmanlarında özellik yoğunlaştırması yaparak bilgi akışını daha seçici hale getirmeyi hedeflemiştir. Bu mekanizma, önemli

özellikleri ön plana çıkarıp gereksiz aktivasyonları bastırarak temsil gücünü artırmak ve gürültüyü azaltmak amacıyla tasarlanmıştır. Özellikle yüksek seviyeli soyutlamalarda, sahnenin genel aydınlatma dengesini koruma ve global kontrast farklarını düzenleme açısından olumlu katkılar sağlamıştır. Bununla birlikte, mimariye eklenen bu modül, çok katmanlı yapısı ve ek konvolüsyon blokları nedeniyle modelin parametre sayısını ve hesaplama yükünü artırmıştır. Yani bilgi aktarımında seçicilik kazandırmış olsa da, ağın genel karmaşıklığı yükselmiş ve eğitim süresi uzamıştır. Dolayısıyla bu mekanizma, gürültü kontrolü ve genel istikrar açısından faydalı olmakla birlikte, hesaplama verimliliği ve ayrıntı hassasiyeti arasında bir denge yaratmıştır. Bu nedenle, Darboğaz Dikkat modelin soyutlama gücünü artıran, ancak performans kazanımı açısından maliyetli bir bileşen olarak değerlendirilebilir.

Sıkıştırma-Uyarma Dikkat, her kanalın küresel ortalamasına göre önem derecesini yeniden ölçeklendirerek kanal bazlı dikkat sağlar. Bu sayede model, farklı aydınlatma koşullarında renk kontrastı ve ışık dengelemesi bakımından daha kararlı bir performans göstermiştir. Özellikle düşük ışıklı sahnelerde, Sıkıştırma-Uyarma dikkat modülü parlaklık tutarlılığını koruyarak renk dengesini stabilize etmiştir. SSIM metriklerinde görülen istikrarlı artış, bu mekanizmanın sahnenin genel aydınlatma dengesini başarılı şekilde modellediğini göstermektedir. Buna karşın, Sıkıştırma-Uyarma dikkat global ilişkiler üzerine odaklandığından, yerel doku detaylarını Uzamsal Dikkat kadar güçlü biçimde vurgulayamamıştır. Bu nedenle Sıkıştırma-Uyarma modülü, renk ve aydınlatma dengesinde olumlu katkı sunarken, mikro düzeyde keskinlik bakımından sınırlı bir performans sergilemiştir.

Öz Dikkat mekanizması, sahne genelindeki uzun menzilli bağıntıları yakalamayı hedefleyen bir yapıdır. Teorik olarak bu modül, sahne boyunca dağılmış aydınlatma değişimlerini ve renk ilişkilerini daha tutarlı biçimde modelleyebilmelidir. Ancak pratik sonuçlar, bu yapının YDA yeniden yapılandırma görevi için beklenen katkıyı sağlayamadığını göstermiştir. Elde edilen metriklerde, Öz Dikkat varyantı PSNR,SSIM ve LPIPS açısından base modelin gerisinde kalmıştır. Yüksek hesaplama maliyeti ve karmaşık bağıntı modellemesi, özellikle detay yoğun bölgelerde bilgi seyrelmesine neden olmuş; bu da ince kenar yapılarının ve yerel dokuların kaybolmasıyla sonuçlanmıştır. Görsel analizlerde, bu varyantın genel tonlama açısından dengeli görünse de, sahne içi kontrast ve detay keskinliği bakımından zayıf performans sergilediği gözlemlenmiştir. Dolayısıyla, Öz Dikkat mekanizması global tutarlılığı artırma potansiyeline sahip olsa da, YDA yeniden yapılandırma doğası gereği kritik olan yerel detay farkındalığını

sağlayamadığından, genel performansa anlamlı bir katkı sunamamıştır.

Bu sonuç, tek görüntüden YDA üretiminde uzun menzilli dikkat yerine yerel odaklı (spatial) mekanizmaların daha etkili olduğunu göstermektedir.

YDA sahne yeniden yapılandırma sürecinde, yalnızca yapısal doğruluk değil, aynı zamanda farklı pozlama seviyeleri arasındaki aydınlatma tutarlılığı da büyük önem taşır. Klasik kayıp fonksiyonları genellikle tek bir görüntü üzerinde çalıştığı için, model farklı pozlamalardan üretilen tahminler arasında ışık yoğunluğu farkı yaratabilmektedir. Bu durum, YDA birleştirme aşamasında parlak alanlarda doygunluk, gölgeli bölgelerde ise kontrast kaybı olarak gözlemlenmiştir. Bu eksikliği gidermek amacıyla önerilen modele Pozlama Tutarlılığı Kaybı eklenmiştir. Bu bileşen, sahnenin farklı pozlama seviyelerinde üretilen varyantlarının ışık dağılımlarını birbiriyle uyumlu hale getirmeyi amaçlar.

Böylece model, tüm pozlama aralıklarında fiziksel olarak tutarlı bir aydınlatma temsili öğrenir. Bu kayıp fonksiyonu eklendiğinde, PSNR ve SSIM değerlerinde artış, LPIPS değerinde ise anlamlı bir düşüş elde edilmiştir. Özellikle Uzamsam Dikkat mekanizmasıyla birlikte kullanıldığında, model parlak ve karanlık alanlar arasında daha dengeli geçişler üretmiş ve YDA füzyon sonrasında görsel tutarlılığı belirgin biçimde iyileştirmiştir.

Sonuç olarak, Pozlama Tutarlılığı Kaybı, YDA Dikkat Ağının yalnızca detay doğruluğunu değil, aydınlatma sürekliliğini de optimize ederek genel sahne stabilitesini artırmıştır.

6. SONUÇLAR

Bu tez kapsamında geliştirilen çok modüllü kodlayıcı-çözücü tabanlı YDA Dikkat Ağı mimarisi, tek bir DDA görüntüden yüksek doğrulukta YDA çıktılar üretebilme yeteneğini başarıyla ortaya koymuştur. Baz alınan üç kollu yapı: YDA Kodlama Ağı, Artırılmış Pozlama Ağı ve Azaltılmış Pozlama Ağı sahnedeki aşırı pozlanmış ve düşük pozlanmış bölgelerdeki bilgi kayıplarını telafi ederek fiziksel pozlama sürecinin tersini modellemiştir. Beş farklı dikkat mekanizmasının mimariye ayrı ayrı entegre edilmesiyle oluşturulan model varyantları, hem yapısal tutarlılık hem de algısal kalite açısından performans artışı sağlamıştır. Yapılan deneysel analizler, uzamsal dikkat mekanizmasının PSNR, SSIM ve LPIPS metriklerinde en başarılı sonuçları elde ettiğini göstermiş; bu durum modelin hem yerel yapıları koruma hem de görsel bütünlük sağlama açısından üstün performans sergilediğini ortaya koymuştur.

Elde edilen bulgular, dikkat mekanizmalarının yalnızca sınıflandırma ya da segmentasyon gibi görevlerde değil, aynı zamanda YDA yeniden yapılandırma gibi karmaşık görsel sentez problemlerinde de etkili biçimde kullanılabileceğini göstermektedir. Bu yönüyle YDA Dikkat Ağı modeli, hem teorik katkılar sunmakta hem de gerçek dünya uygulamaları için pratik potansiyel taşımaktadır. YDA görüntüleme alanındaki gelecekteki çalışmalar için sağlam bir temel oluşturmaktadır.

7. ÖNERİLER

Bu tez kapsamında geliştirilen YDA Dikkat Ağı mimarisi, dikkat mekanizmalarının YDA görüntü sentezine katkılarını başarılı şekilde ortaya koymuştur. Ancak modelin daha da geliştirilmesi ve farklı uygulama alanlarında test edilebilmesi için çeşitli öneriler sunulabilir:

- Farklı dikkat türlerini birleştiren hibrit mekanizmalar, alternatif yapılandırmalarla değerlendirilerek dikkat modüllerinin birlikte çalışma potansiyeli daha derinlemesine incelenebilir.
- Dönüştürücü (Transformer) tabanlı mimariler ile karşılaştırmalı çalışmalar gerçekleştirilerek, küresel bağlam modellemesi açısından derin evrişimli yapılara alternatif potansiyeller araştırılabilir.
- Mobil ve gömülü sistemlere uygun hafif (lightweight) modeller geliştirilerek gerçek zamanlı YDA uygulamaları desteklenebilir.
- Modelin genellenebilirliğini artırmak amacıyla, farklı aydınlatma koşulları, sahne içerikleri ve pozlama dağılımları içeren daha çeşitli ve gerçek dünya koşullarına yakın YDA veri kümeleri üzerinde çapraz doğrulama yapılabilir.

Tüm bu öneriler, hem YDA Dikkat Ağı mimarisinin daha gelişmiş versiyonlarının oluşturulmasına hem de YDA görüntüleme alanındaki mevcut sınırlılıkların aşılmasına katkı sağlayacaktır.

KAYNAKLAR

A Sharif, S. M., Naqvi, R. A., Biswas, M., & Kim, S. (2021). A two-stage deep network for high dynamic range image reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 550-559).

Aghabiglou, A., Terris, M., Jackson, A., & Wiaux, Y. (2023, June). Deep network series for large-scale high-dynamic range imaging. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1-5). IEEE.

Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.

Ballard, D. H., & Brown, C. M. (1982). *Computer vision*. Prentice Hall Professional Technical Reference.

Burger, W., & Burge, M. J. (2010). *Principles of digital image processing: fundamental techniques*. Springer Science & Business Media.

Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, *40*(4), 834-848.

Chen, X., Liu, Y., Zhang, Z., Qiao, Y., & Dong, C. (2021). Hdrunet: Single image hdr reconstruction with denoising and dequantization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 354-363).

Chen, Z., Pawar, K., Ekanayake, M., Pain, C., Zhong, S., & Egan, G. F. (2023). Deep learning for image enhancement and correction in magnetic resonance imaging—state-of-the-art and challenges. *Journal of Digital Imaging*, *36*(1), 204-230.

Choromanski, K., Likhosherstov, V., Dohan, D., Song, X., Gane, A., Sarlos, T., ... & Weller, A. (2020). Rethinking attention with performers. *arXiv preprint arXiv:2009.14794*.

Crognale, M., De Iuliis, M., Rinaldi, C., & Gattulli, V. (2023). Damage detection with image processing: a comparative study. *Earthquake Engineering and Engineering Vibration*, *22*(2), 333-345.

Dalal, D., Vashishtha, G., Singh, P., & Raman, S. (2023, October). Single image ldr to hdr conversion using conditional diffusion. In *2023 IEEE International Conference on Image Processing (ICIP)* (pp. 3533-3537). IEEE.

Date, S., Rakhde, V. M., & Deharkar, A. B. (2022). A research on digital image processing technology and application. *International Journal of Advanced Research in Computer and Communication Engineering, 11*(11), 263–268.

Debevec, P., Reinhard, E., Ward, G., & Pattanaik, S. (2004). High dynamic range imaging. In *ACM SIGGRAPH 2004 Course Notes* (pp. 14-es).

Dongur, K. R., Tandekar, P., & Purve, S. K. (2022). Digital image processing: Its history and application. *International Journal of Advanced Research in Computer and Communication Engineering, 11*(6), 263–268.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*.

Eilertsen, G., Kronander, J., Denes, G., Mantiuk, R. K., & Unger, J. (2017). HDR image reconstruction from a single exposure using deep CNNs. *ACM transactions on graphics (TOG)*, *36*(6), 1-15.

Endo, Y., Kanamori, Y., & Mitani, J. (2017). Deep reverse tone mapping. *ACM Transactions on Graphics (TOG)*, 36(6), 1–10.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, *27*.

Hall, E. (1979). *Computer image processing and recognition*. Elsevier.

Huang, Y., Wang, W., & Wang, L. (2017). Video super-resolution via bidirectional recurrent convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, *40*(4), 1015-1028.

Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7132-7141).

Jensen, J. R. (1986). Introductory Digital. *Image Processing, New Jersey: Prentice-Hall*.

Kalantari, N. K., & Ramamoorthi, R. (2017). Deep high dynamic range imaging of dynamic scenes. *ACM Trans. Graph.*, *36*(4), 144-1.

Khan, Z., Khanna, M., & Raman, S. (2019, November). Fhdr: Hdr image reconstruction from a single ldr image using feedback network. In *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)* (pp. 1-5). IEEE.

Kim, J., Lee, S., & Kang, S. J. (2021, May). End-to-end differentiable learning to HDR image synthesis for multi-exposure images. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 35, No. 2, pp. 1780-1788).

Kim, S. Y., Oh, J., & Kim, M. (2020, April). Jsi-gan: Gan-based joint super-resolution and inverse tone-mapping with pixel-wise task-specific filters for uhd hdr video. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 34, No. 07, pp. 11287-11295).

Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information*

processing systems, *25*.

Le, P. H., Le, Q., Nguyen, R., & Hua, B. S. (2023). Single-image hdr reconstruction by multi-exposure generation. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision* (pp. 4063-4072).

Lee, S., An, G. H., & Kang, S. J. (2018). Deep recursive HDRI: Inverse tone mapping using generative adversarial networks. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 596–611). Springer.

Li, R., Wang, C., Wang, J., Liu, G., Zhang, H. Y., Zeng, B., & Liu, S. (2022). Uphdr-gan: Generative adversarial network for high dynamic range imaging with unpaired data. *IEEE Transactions on Circuits and Systems for Video Technology*, *32*(11), 7532-7546.

Li, X., Ren, Y., Jin, X., Lan, C., Wang, X., Zeng, W., ... & Chen, Z. (2023). Diffusion Models for Image Restoration and Enhancement--A Comprehensive Survey. *arXiv preprint arXiv:2308.09388*.

Lin, Z., Feng, M., Santos, C. N. D., Yu, M., Xiang, B., Zhou, B., & Bengio, Y. (2017). A structured self-attentive sentence embedding. *arXiv preprint arXiv:1703.03130*.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, *42*, 60-88.

Liu, C. (2024). A review of digital image processing techniques and future prospects. *International Journal of Computer Science and Information Technology*, 4, 223–233. https://doi.org/10.62051/ijcsit.v4n3.22

Liu, Y. L., Lai, W. S., Chen, Y. S., Kao, Y. L., Yang, M. H., Chuang, Y. Y., & Huang, J. B. (2020). Single-image HDR reconstruction by learning to reverse the camera pipeline. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 1651-1660).

Lopez-Cabrejos, J., Paixão, T., Alvarez, A. B., & Luque, D. B. (2025). An Efficient and Low-Complexity Transformer-Based Deep Learning Framework for High-Dynamic-Range Image Reconstruction. *Sensors*, *25*(5), 1497.

Lore, K. G., Akintayo, A., & Sarkar, S. (2017). LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition*, *61*, 650-662.

Luong, M. T., Pham, H., & Manning, C. D. (2015). Effective approaches to attention-based neural machine translation. *arXiv preprint arXiv:1508.04025*.

Ma, J., & Zhang, H. (2025). HDR-DANet: single HDR image reconstruction via dual attention. *Multimedia Systems*, *31*(1), 1-12.

Maity, A., Pious, R., Lenka, S. K., Choudhary, V., & Lokhande, S. (2023). A survey on super resolution for video enhancement using gan. *arXiv preprint arXiv:2312.16471*.

Manaa, M. J., Abbas, A. R., & Shakur, W. A. (2023, December). A systematic

review for image enhancement using deep learning techniques. In *AIP Conference Proceedings* (Vol. 2977, No. 1). AIP Publishing.

Mantiuk, R., Krawczyk, G., Zdrojewska, D., Mantiuk, R., Myszkowski, K., & Seidel, H. P. (2015). *High dynamic range imaging* (pp. 1-42). na.

Mekruksavanich, S., & Jitpattanakul, A. (2023). Hybrid convolution neural network with channel attention mechanism for sensor-based human activity recognition. *Scientific Reports*, *13*(1), 12067.

Metzler, C. A., Ikoma, H., Peng, Y., & Wetzstein, G. (2020). Deep optics for single-shot high-dynamic-range imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1375-1385).

Miskdjian, I., Hodhod, H., Abdeen, M., & Elshabrawy, M. (2024). Investigation and optimization of factors affecting the accuracy of strain measurement via digital image processing. *Journal of Engineering and Applied Science*, *71*(1), 63.

Mitchell, T. M. (1997). Does machine learning really work?. *AI magazine*, *18*(3), 11-11.

Park, J., Woo, S., Lee, J. Y., & Kweon, I. S. (2018). Bam: Bottleneck attention module. *arXiv preprint arXiv:1807.06514*.

Park, Y. I., Song, J. W., & Kang, S. J. (2022, June). 65-3: Invited Paper: Deep Learning-Based Image Enhancement for HDR Imaging. In *SID Symposium Digest of Technical Papers* (Vol. 53, No. 1, pp. 865-868).

Reinhard, E. (2020). High dynamic range imaging. In *Computer Vision: A Reference Guide* (pp. 1-6). Cham: Springer International Publishing.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18* (pp. 234-241). Springer international publishing.

Rosenfeld, A. (1976). *Digital picture processing*. Academic press.

Roy, A. G., Navab, N., & Wachinger, C. (2018). Concurrent spatial and channel 'squeeze & excitation'in fully convolutional networks. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I* (pp. 421-429). Springer International Publishing.

Sara, U., Akter, M., & Uddin, M. S. (2019). Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study. *Journal of Computer and Communications*, *7*(3), 8-18.

Sharma, A., & Mishra, P. K. (2022). Image enhancement techniques on deep learning approaches for automated diagnosis of COVID-19 features using CXR images. *Multimedia Tools and Applications*, *81*(29), 42649-42690.

Shen, J., & Wu, T. (2021). Learning Spatially-Adaptive Squeeze-Excitation Networks for Image Synthesis and Image Recognition. *arXiv preprint arXiv:2112.14804*.

Sinha, R. K., Pandey, R., & Pattnaik, R. (2018). Deep learning for computer vision tasks: a review. *arXiv preprint arXiv:1804.03928*.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction* (Vol. 1, No. 1, pp. 9-11). Cambridge: MIT press.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

Wang, L., & Yoon, K. J. (2021). Deep learning for hdr imaging: State-of-the-art and future trends. *IEEE transactions on pattern analysis and machine intelligence*, *44*(12), 8874-8895.

Wang, Y., Xie, W., & Liu, H. (2022). Low-light image enhancement based on deep learning: a survey. *Optical Engineering*, *61*(4), 040901-040901.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, *13*(4), 600-612.

Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 3-19).

Yan, Q., Hu, T., Chen, G., Dong, W., & Zhang, Y. (2025). Boosting HDR Image Reconstruction via Semantic Knowledge Transfer. *arXiv preprint arXiv:2503.15361*.

Yan, Q., Wang, B., Zhang, L., Zhang, J., You, Z., Shi, Q., & Zhang, Y. (2021). Towards accurate HDR imaging with learning generator constraints. *Neurocomputing*, *428*, 79-91.

Yan, Q., Zhang, S., Chen, W., Tang, H., Zhu, Y., Sun, J., ... & Zhang, Y. (2023). Smae: Few-shot learning for hdr deghosting with saturation-aware masked autoencoders. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5775-5784).

Zangana, H. M., Mustafa, F. M., Mohammed, A. K., & Omar, N. (2024). Enhancing Image Quality With Deep Learning: Techniques And Applications. *Jurnal ELTIKOM: Jurnal Teknik Elektro, Teknologi Informasi dan Komputer*, *8*(2), 119-131.

Zhang, H., Goodfellow, I., Metaxas, D., & Odena, A. (2019, May). Self-attention generative adversarial networks. In *International conference on machine learning* (pp. 7354-7363). PMLR.

Zhang, P. (2022). Image enhancement method based on deep learning. *Mathematical Problems in Engineering*, *2022*(1), 6797367.

Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 586-595).

Zhang, S. F., Zhai, J. H., Xie, B. J., Zhan, Y., & Wang, X. (2019, July). Multimodal representation learning: Advances, trends and challenges. In *2019 International Conference on Machine Learning and Cybernetics (ICMLC)* (pp. 1-6). IEEE.