

KOCAELİ ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ

ELEKTRONİK VE HABERLEŞME MÜHENDİSLİĞİ  
ANABİLİM DALI

DOKTORA TEZİ

HDR GÖRÜNTÜ ÜRETEBİLEN KAMERA VERİ YOLUNUN  
DERİN AĞ MİMARİSİ KULLANILARAK MODELLENMESİ

KORAY UĞUR ERBAŞ

KOCAELİ 2025

**KOCAELİ ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**

**ELEKTRONİK VE HABERLEŞME MÜHENDİSLİĞİ**  
**ANABİLİM DALI**

**DOKTORA TEZİ**

**HDR GÖRÜNTÜ ÜRETEBİLEN KAMERA VERİ YOLUNUN**  
**DERİN AĞ MİMARİSİ KULLANILARAK MODELLENMESİ**

**KORAY UĞUR ERBAŞ**

**Doç.Dr.Aysun TAŞYAPI ÇELEBİ**  
**Danışman, Kocaeli Üniversitesi**

.....

**Prof.Dr.Oğuzhan URHAN**  
**Jüri Üyesi, Kocaeli Üniversitesi**

.....

**Doç.Dr.Orhan AKBULUT**  
**Jüri Üyesi, Kocaeli Üniversitesi**

.....

**Prof.Dr.Mehmet Kemal GÜLLÜ**  
**Jüri Üyesi, Bakırçay Üniversitesi**

.....

**Doç.Dr.Ali Can KARACA**  
**Jüri Üyesi, Yıldız Teknik Üniversitesi**

.....

**Tezin Savunulduğu Tarih: 14.01.2025**

## ETİK BEYAN VE ARAŞTIRMA FONU DESTEĞİ

Kocaeli Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- Bu tezin bana ait, özgün bir çalışma olduğunu,
- Çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı,
- Bu çalışma kapsamında elde edilen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi,
- Bu çalışmanın Kocaeli Üniversitesi'nin abone olduğu intihal yazılım programı kullanılarak Fen Bilimleri Enstitüsü'nün belirlemiş olduğu ölçütlere uygun olduğunu,
- Kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- Tezin/Projenin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı,

beyan ederim.

- Bu tez/proje çalışmasının herhangi bir aşaması hiçbir kurum/kuruluş tarafından maddi/alt yapı desteği ile desteklenmemiştir.
- Bu tez/proje çalışması kapsamında üretilen veri ve bilgiler ..... Tarafından ..... No'lu proje kapsamında maddi/alt yapı desteği alınarak gerçekleştirilmiştir.

Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçları kabul ettiğimi bildiririm.

.....

Koray Uğur ERBAŞ

## YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI

Fen Bilimleri Enstitüsü tarafından onaylanan lisansüstü tezimin/projemin tamamını veya herhangi bir kısmını, basılı ve elektronik formatta arşivleme ve aşağıda belirtilen koşullarla kullanıma açma izninin Kocaeli Üniversitesi'ne verdiğimi beyan ederim. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarımın bende kalacak, tezimin/projemin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanımını bana ait olacaktır.

Tezin/projenin kendi özgün çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin/projenin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yüksek öğretim kurulu tarafından yayınlanan **“Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge”** kapsamında tezim aşağıda belirtilen koşullar haricinde YÖK Ulusal Tez Merkezi/ Kocaeli Üniversitesi Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü yönetim kurulu kararı ile tezimin/projemin erişime açılması mezuniyet tarihinden itibaren 2 yıl ertelenmiştir.
- Enstitü yönetim kurulu gerekçeli kararı ile tezimin/projemin erişime açılması mezuniyet tarihinden itibaren 6 ay ertelenmiştir.
- Tezim/projem ile ilgili gizlilik kararı verilmemiştir.

.....  
Koray Uğur ERBAŞ

## ÖNSÖZ VE TEŞEKKÜR

Günümüzde kameraların kullanımının yaygınlaşması ile birlikte sahneye ait kaliteli bir görüntünün elde edilebilmesine yönelik talep artarak devam etmektedir. Söz konusu kalite ise insanın görme sistemi yetisi ile doğrudan ilişkilidir. Bu nedenle kamera dizaynında; oluşturulan görüntünün, mümkün olduğunca sahnenin doğrudan insanın kendisi tecrübe ediliyormuşçasına kullanıcıya sunulması hedeflenmektedir. Bu nedenle sahneye ait yüksek ve düşük pozlanmış bölgelerdeki detayları da ortaya çıkaracak, sahneye ait renk bilgisini doğru şekilde yansıtacak ve mümkün oldukça geniş dinamik aralığa sahip görüntünün oluşturulması nihai amaçtır.

Bu tez çalışmasında, ton dönüşümü yapılmış Yüksek Dinamik Aralıklı (High Dynamic Range, HDR) görüntüyü oluşturacak şekilde kamera ISP veri yolunun geliştirilmesi amaçlanmaktadır. Kameradan alınan Bayer paterne sahip 10 bitlik ham veriden, herhangi bir kullanıcı müdahalesine ya da parametre girişine ihtiyaç duyulmadan 8 bitlik çıktı görüntüyü oluşturabilecek derin ağ mimarisi tasarlanacak, eğitimi gerçekleştirilecek ve test aşamasına ait deneysel sonuçlar paylaşılacaktır.

Tez çalışmamda desteğini esirgemeyen, çalışmalarına yön veren, ve bana güvenen danışmanım saygı değer hocam Doç.Dr.Aysun TAŞYAPI ÇELEBİ'ye, süreç içerisinde değerli yorumlarını benden esirgemeyen Prof. Dr.M.Kemal GÜLLÜ ve Prof. Dr.Oğuzhan URHAN'a sonsuz teşekkürlerimi sunarım.

Bununla birlikte asla hakkını ödeyemeyeceğim, bu zorlu süreçte de bana güvenen ve daima arkamda duran sevgili aileme teşekkür ederim. Bununla birlikte Gölcük Tersanesi Komutanlığı Savaş Sistemleri Baş Mühendisliği'nde birlikte görev yaptığımız değerli iş arkadaşlarıma, ihtiyaç duyduğumda her zaman yanımda bulduğum arkadaşım Sayın Ali ÜYRÜŞ'e, uzun koşullarda fotoğrafçılık bilgisiyle bana ilham veren Sayın Burak UZUN'a ve diğer tüm koşu arkadaşlarıma teşekkürler ederim.

Son olarak "Hayatta en hakiki mürşit ilimdir, fendir." sözüyle doğruyu daima akıl ve bilimde aramama rehber olan Ulu Önder Gazi Mustafa Kemal ATATÜRK'e şükranlarımı bir borç bilirim.

Şubat – 2025

Koray Uğur ERBAŞ

## İÇİNDEKİLER

ETİK BEYAN VE ARAŞTIRMA FONU DESTEĞİ.....	i
YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI .....	ii
ÖNSÖZ VE TEŞEKKÜR.....	iii
İÇİNDEKİLER.....	iv
ŞEKİLLER DİZİNİ .....	vi
TABLolar DİZİNİ.....	viii
SİMGELER VE KISALTMALAR DİZİNİ .....	ix
ÖZET .....	xii
ABSTRACT .....	xiii
1. GİRİŞ .....	1
1.1. Çalışmanın Amacı .....	4
2. ÖNCEKİ ÇALIŞMALAR .....	6
2.1. Kamerada Görüntünün Oluşturulması ve Geleneksel Kamera ISP Veri Yolunun Yapısı .....	6
2.2. Kamera ISP Veri Yolunun Derin Öğrenme Mimarisi ile Oluşturulması .....	10
2.2.1 PyNET Mimarisi .....	12
2.2.2 Lightweight PyNET Mimarisi.....	14
2.3. HDR Görüntünün Oluşturulması.....	15
2.3.1. Çoklu pozlamalarla elde edilen LDR görüntülerinin birleştirilmesi.....	15
2.3.2. Tek Bir LDR Görüntüsünden HDR Görüntü Oluşturma .....	16
2.4. Ton Dönüştürme Operatörü (Tone Mapping Operator (TMO)).....	18
3. GÖRÜNTÜ KALİTESİ DEĞERLENDİRME METRİKLERİ.....	20
3.1. Referansa İhtiyaç Duyan Görüntü Kalitesi Değerlendirme Metrikleri (FR-IQA) .....	20
3.1.1. Ortalama Karesel Hata (MSE) .....	21
3.1.2. PSNR (Peak Signal to Noise Ratio) .....	21
3.1.3. SSIM (Structural Similarity Index) .....	22
3.1.4. VIF (Visual Information Fidelity).....	23
3.1.5. DIQam-FR (Deep Image Quality Measure-Full Reference).....	23
3.2. Referansa İhtiyaç Duymayan Görüntü Kalitesi Değerlendirme Metrikleri (NR-IQA).....	24
3.2.1. Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) .....	24
3.2.2. The Natural Image Quality Evaluator (NIQE) .....	25
3.2.3. Perception based Image Quality Evaluator (PIQE).....	26
3.2.4. Neural Image Assessment (NIMA).....	27
3.2.5. Multi-scale Image Quality Transformer (MUSIQ).....	28
3.2.6. Multi-dimension Attention Network for No-Reference Image Quality Assessment (MANIQA) .....	29
3.2.7. Language-Image Quality Evaluator (LIQE) .....	30
3.3. Kullanılan Kalite Metrikleri .....	31
4. VERİ SETİNİN HAZIRLANMASI .....	32
5. ÖNERİLEN YÖNTEM .....	37
5.1. Model Mimarisi .....	37
5.2. Eğitimde Kullanılan Kayıp Fonksiyonları ve Eğitim Parametreleri.....	41
6. DENEYSEL SONUÇLAR .....	44
6.1. Nicel Sonuçlar .....	45

6.2. Nitel Sonuçlar .....	50
6.3. Etkinlik Analizi (Ablation Study).....	60
6.3.1. Model Mimarisi Alt Modül Etkinlik Analizi.....	60
6.3.2. Histogram Kayıp Fonksiyonu Teriminin Ağırlık Katsayısının Belirlenmesi .....	63
6.3.3. Veriseti Çözünürlüğü Etkinlik Analizi.....	74
6.3.4. Kayıp Fonksiyonu Terimlerinin Öğrenilebilir Parametreler ile Ağırlıklandırılması.....	79
6.3.5. Veri Çoğaltma İşlemi Etkinlik Analizi.....	85
7. SONUÇLAR VE ÖNERİLER.....	89
7.1. Sonuçların Değerlendirilmesi .....	89
7.2. Gelecek Dönem Çalışmaları .....	90
KAYNAKLAR.....	92
KİŞİSEL YAYIN VE ESERLER.....	102
ÖZGEÇMİŞ.....	103



## ŞEKİLLER DİZİNİ

Şekil 1.1.	LDR ve HDR görüntü .....	2
Şekil 1.2.	Model genel akış diyagramı .....	4
Şekil 2.1.	Görüntü oluşturma/anlamlandırma diyagramı .....	7
Şekil 2.2.	Sensör (Photodetector) ve CFA Yapısı .....	7
Şekil 2.3.	Kamera ISP Veri Yolu Genel Yapısı .....	8
Şekil 2.4.	Geri Mozaikleme yapılmamış ve yapılmış görüntü .....	8
Şekil 2.5.	Beyaz Dengeleme uygulanmamış ve uygulanmış görüntü .....	9
Şekil 2.6.	PyNET model mimarisi .....	13
Şekil 2.7.	Lightweight PyNET model mimarisi .....	14
Şekil 3.1.	SSIM indeksi ölçüm diyagramı .....	22
Şekil 3.2.	BRISQUE metriği diyagramı .....	25
Şekil 3.3.	NIQE metriği diyagramı .....	26
Şekil 3.4.	NIMA metriği diyagramı .....	28
Şekil 3.5.	MUSIQ metriği diyagramı .....	29
Şekil 3.6.	MANIQA metriği diyagramı .....	29
Şekil 3.7.	LIQE metriği diyagramı .....	31
Şekil 4.1.	Veriseti hazırlama aşamalarına ilişkin görseller .....	33
Şekil 4.2.	Topaz Gigapixel AI fotoğraf düzenleme yazılımı kullanıcı arayüzü .....	34
Şekil 4.3.	Photomatix Pro 5.0 fotoğraf düzenleme yazılımı kullanıcı arayüzü .....	34
Şekil 4.4.	PYNET çıktısı ve Topaz Gigapixel AI ve Photomatix Pro 5.0 kullanılarak elde edilen görüntü .....	35
Şekil 4.5.	Test amacıyla ayrılan görüntüler ve imge kimlikleri .....	36
Şekil 5.1.	Önerilen çalışmanın genel mimarisi .....	38
Şekil 5.2.	Effective Channel Attention (ECA) blok diyagramı .....	40
Şekil 5.3.	Convolutional Block Attention Modules (CBAM) blok diyagramı .....	40
Şekil 5.4.	Information Extraction Module (IEM) blok diyagramı .....	41
Şekil 6.1.	Önerilen ve referans model çıktılarının görsel karşılaştırması-1 .....	51
Şekil 6.2.	Önerilen ve referans model çıktılarının görsel karşılaştırması-2 .....	52
Şekil 6.3.	Önerilen ve referans model çıktılarının görsel karşılaştırması-3 .....	53
Şekil 6.4.	Önerilen ve referans model çıktılarının görsel karşılaştırması-4 .....	54
Şekil 6.5.	Önerilen model ile farklı kameraların ISP çıktılarının görsel karşılaştırması-1 (imge 5) .....	56
Şekil 6.6.	Önerilen model ile farklı kameraların ISP çıktılarının görsel karşılaştırması -2 (imge 8) .....	57
Şekil 6.7.	Önerilen model ile farklı kameraların ISP çıktılarının görsel karşılaştırması-3 (imge 3) .....	58
Şekil 6.8.	Önerilen model ile farklı kameraların ISP çıktılarının görsel karşılaştırması-4 (imge 0) .....	59
Şekil 6.9.	Alt-modül etkinlik analizi görsel karşılaştırma-1 (imge 7) .....	64
Şekil 6.10.	Alt-modül etkinlik analizi görsel karşılaştırma-1 (Şekil 6.9'a ait yakınlaştırılmış sonuçlar) .....	65
Şekil 6.11.	Alt-modül etkinlik analizi görsel karşılaştırma-2 (imge 8) .....	66

Şekil 6.12.	Alt-modül etkinlik analizi görsel karşılaştırma-2 (Şekil 6.11'e ait yakınlaştırılmış sonuçlar) .....	67
Şekil 6.13.	Alt-modül etkinlik analizi görsel karşılaştırma-3 (imge 6) .....	68
Şekil 6.14.	Alt-modül etkinlik analizi görsel karşılaştırma-3 (Şekil 6.13'e ait yakınlaştırılmış sonuçlar) .....	69
Şekil 6.15.	Histogram Kayıp Fonksiyonu ağırlık parametresi analizine ait görsel sonuçlar-1 .....	71
Şekil 6.16.	Histogram Kayıp Fonksiyonu ağırlık parametresi analizine ait görsel sonuçlar-2 .....	72
Şekil 6.17.	Histogram Kayıp Fonksiyonu ağırlık parametresi analizine ait görsel sonuçlar-3 .....	73
Şekil 6.18.	Eğitimde kullanılan veriseti giriş çözünürlüğüne ait etkinlik analizinin nitel sonuçları-1 (imge 6) .....	77
Şekil 6.19.	Eğitimde kullanılan veriseti giriş çözünürlüğüne ait etkinlik analizinin nitel sonuçları-2 (imge 1) .....	78
Şekil 6.20.	Öğrenilebilir kayıp fonksiyonlarına ait ağırlık parametrelerine ait deneysel sonuçlar-1 (imge 2) .....	83
Şekil 6.21.	Öğrenilebilir kayıp fonksiyonlarına ait ağırlık parametrelerine ait deneysel sonuçlar -2 (imge 7) .....	84
Şekil 6.22.	Veri çoğaltma işlemi etkinlik analizi görsel sonuçlar .....	87

## TABLolar DİZİNİ

Tablo 1.1.	IGS için parlaklık skalası .....	1
Tablo 1.2.	Model genel akış diyagramı .....	2
Tablo 5.1.	Kullanılan eğitim parametreleri ve toplam kayıp fonksiyonları .....	42
Tablo 6.1.	Genel model performans karşılaştırma .....	45
Tablo 6.2.	Önerilen model ile referans modellerin ile karşılaştırılması .....	46
Tablo 6.3.	Dinamik Aralık (dB) performans karşılaştırma .....	47
Tablo 6.4.	Önerilen model ile farklı kamera ISP çıktılarının karşılaştırılması.....	49
Tablo 6.5.	Alt modül etkinlik analizi sonuçlarının karşılaştırılması.....	61
Tablo 6.6.	Histogram kayıp fonksiyonu etkinlik analizi (NIMA/MUSIQ).....	70
Tablo 6.7.	Histogram kayıp fonksiyonu etkinlik analizi (Dinamik Aralık, dB).....	74
Tablo 6.8.	Eğitim amacıyla kullanılan veriseti çözünürlüğü etkinlik analizi (Dinamik Aralık, dB) .....	75
Tablo 6.9.	Eğitimde kullanılan veriseti giriş çözünürlüğüne ait etkinlik analizinin nicel sonuçları.....	76
Tablo 6.10.	Öğrenilebilir kayıp fonksiyonlarına ait ağırlık parametreleri (Deney-1).....	80
Tablo 6.11.	Öğrenilebilir kayıp fonksiyonlarına ait ağırlık parametreleri (Deney-2) .....	81
Tablo 6.12.	Öğrenilebilir kayıp fonksiyonlarına ait ağırlık parametrelerine ait deneysel sonuçlar .....	82
Tablo 6.13.	Veri seti çoğaltma işlemi etkinlik analizine ait nicel sonuçlar.....	86

## SİMGELER VE KISALTMALAR DİZİNİ

$C_1$	: Sabit bir sayı
$C_2$	: Sabit bir sayı
$\sigma_x$	: $x$ görüntüsünün standart sapması
$\sigma_x^2$	: $x$ görüntüsünün varyansı
$\sigma_{xy}$	: $x$ ve $y$ görüntülerin çagraz kovaryansı
$\mu_x$	: $x$ görüntüsüne ait ortalama piksel deęeri
$\log_{10}(x)$	: Onluk logaritma fonksiyonu
$\max_x$	: $x$ sinyalindeki maksimum deęer
$MSE(x, y)$	: $x$ ve $y$ görüntüleri arasındaki ortalama karesel hata fonksiyonu
$P_{\text{signal}}$	: Sinyalin gücü, (Watt)
$P_{\text{noise}}$	: Gürültünün gücü, (Watt)

### Kısaltmalar

AI	: Artificial Intelligence (Yapay Zeka)
AWNet	: Attentive Wavelet Network (Dikkatli Dalgacık Aęı)
BM3D	: Block-Matching and 3D (Blok Eşitleme ve 3D)
BRISQUE	: Blind/Referenceless Image Spatial Quality Evaluator (Kör/Referanssız Uzamsal Görüntü Kalite Deęerlendiricisi)
CBAM	: Convolutional Block Attention Modules (Konvolüsyonel Blok Dikkat Modülleri)
CCD	: Charge-Coupled Device (Yük Baęlaşımli Cihaz)
cdf	: Cumulative Distribution Function (Kümülatif Daęılım Fonksiyonu)
CFA	: Color Filter Array (Renk Filtre Dizini)
cGAN	: Conditional Generative Adversarial Networks (Koşullu Çekişmeli Üretici Aęlar)
CLS	: Classification Token (Sınıflandırma Token)
CMOS	: Complementary Metal Oxide Semiconductor (Bütünleyici Metal Oksit Yarı İletken)
CNN	: Convolutional Neural Network (Evrişimsel Sinir Aęları)
CRF	: Camera Response Function (Kamera Cevap Fonksiyonu)
CSANet	: Channel Spatial Attention Network (Kanal Uzamsal Dikkat Aęı)
DDPM	: Denoising Diffusion Probabilistic Model (Gürültü Giderme Yayılım Olasılıksal Modeli)
DFC-DIT	: The Deep Feature Consistent-Deep Image Transformation (Derin Öznitelik Tutarlı-Derin Görüntü Dönüşümü)
DIQam	: Deep Image Quality Measure (Derin Görüntü Kalitesi Ölçümü)
DSLR	: Digital Single-Lens Reflex (Sayısal Tek Mercek Yansımali)
EAM	: Enhancement Attention Module (İyileştirme Dikkat Modülü)
ECA	: Effective Channel Attention (Etkin Kanal Dikkat)
EDLLIE-Net	: Enhanced Deep convolutional Low-Light Image Enhancement Network (Geliştirilmiş Derin Evrişimli Düşük Işık Görüntü Geliştirme Aęı)
EMD	: Earth Moving Distance (Yer Hareket Mesafesi)

FR-IQA	: Full Reference - Image Quality Assessment (Tam Referanslı - Görüntü Kalitesi Değerlendirme)
FSIM	: Feature Similarity (Öznitelik Benzerliği)
GAN	: Generative Adversarial Networks (Çekişmeli Üretici Ağlar)
GFFM	: Global Feature Fusion Module (Küresel Özellik Füzyon Modülü)
HDR	: High Dynamic Range (Yüksek Dinamik Aralık)
HSE	: Hash Based 2D Spatial Embedding (Hash Tabanlı 2D Uzamsal Gömme)
HSV	: Hue-Saturation-Value (Ton-Doyma-Parlaklık)
HVS	: Human Vision System (İnsan Görme Sistemi)
IEM	: Information Extraction Modules (Bilgi Çıkarma Modülleri)
IQA	: Image Quality Assessment (Görüntü Kalitesi Değerlendirme)
ISP	: Image Signal Processor (Görüntü Sinyal İşlemcisi)
İGS	: İnsan Görme Sistemi
LDR	: Low Dynamic Range (Düşük Dinamik Aralık)
LIQE	: Language-Image Quality Evaluator (Dil-Görüntü Kalitesi Değerlendiricisi)
LW-ISP	: Lightweight-Image Signal Processor (Hafif Görüntü Sinyal İşlemcisi)
MANIQA	: Multi-dimension Attention Network for No-Reference Image Quality Assessment
MATLAB	: Matrix Laboratory (Matris Laboratuvarı)
MEF	: Multi-Exposure Image Fusion (Çoklu Pozlama Görüntü Füzyonu)
MP	: Megapixel (Mega Pixel)
MSCN	: Mean Subtracted Contrast Normalized (Ortalama Çıkarılmış Kontrast Normalleştirilme)
MUSIQ	: Multi-scale Image Quality Transformer (Çok Ölçekli Görüntü Kalitesi Dönüştürücüsü)
MOS	: Mean Opinion Score (Ortalama Görüş Puanı)
MSE	: Mean Square Error (Ortalama Karese Hata)
MS-SSIM	: Multi Scale Structural Similarity Index (Çok Ölçekli Yapısal Benzerlik İndeksi)
NIMA	: Neural Image Assessment (Nöral Görüntü Değerlendirmesi)
NIQE	: The Natural Image Quality Evaluator (Doğal Görüntü Kalitesi Değerlendiricisi)
NLM	: Non Local Means (Lokal Olmayan Ortalama)
NR-IQA	: Non-Referenced - Image Quality Assessment (Referanssız - Görüntü Kalitesi Değerlendirme)
NSS	: Natural Scene Statistics (Doğal Sahne İstatistikleri)
NTIRE	: New Trends in Image Restoration and Enhancement
PIQE	: Perception based Image Quality Evaluator (Algıya Dayalı Görüntü Kalitesi Değerlendiricisi)
PSNR	: Peak Signal-to-Noise Ratio (Tepe Sinyal-Gürültü Oranı)
pmf	: Probability Mass Function (Olasılık Yoğunluk Fonksiyonu)
RGB	: Red-Green-Blue (Kırmızı-Yeşil-Mavi)
RGBG	: Red-Green-Blue-Green (Kırmızı-Yeşil-Mavi-Green)
RR-IQA	: Reduced Referenced - Image Quality Assessment (Azaltılmış Referanslı - Görüntü Kalitesi Değerlendirme)

SCA	: Spatial and Channel Attention (Uzamsal ve Kanal Dikkat)
SCE	: Scale Embedding (Ölçek Gömme)
SSIM	: Structural Similarity Index (Yapısal Benzerlik İndeksi)
SSTB	: Scale Swin Transformer Block
SVR	: Support Vector Regression (Destek Vektör Regresyonu)
TAB	: Transposed Attention Block
TCRDN	: Task-specific Channel Adaptive Residual Dense Network (Görev-belirli Adaptif Artık Sık Ağ)
TMO	: Tone Mapping Operator (Ton Dönüştürme Operatörü)
VGG	: Visual Geometry Group
VIF	: Visual Information Fidelity (Görsel Bilgi Doğruluğu)
ViT	: Vision Transformer



# HDR GÖRÜNTÜ ÜRETEBİLEN KAMERA VERİ YOLUNUN DERİN AĞ MİMARİSİ KULLANILARAK MODELLENMESİ

## ÖZET

Bugün, geleneksel bir dijital kamera ISP veri yolu, Geri Mozaikleme, Gürültü Giderme, Beyaz Dengeleme, Renk İyileştirme, Ton Dönüştürme vb. gibi belirli bir görev için tasarlanmış, sıralı olarak birbirine bağlanan görüntü işleme modülleri tarafından oluşturulmaktadır. Bu modüller bağımsız olarak tasarlandığından ve veri yolunun baştan sona optimizasyonu yapılamadığından, oluşturulan nihai görüntüde bozulmaların gözlemlenmesi muhtemeldir. Bu etkileri ortadan kaldırmak için literatürde çeşitli kamera ISP veri yolu modelleri önerilmiş ve evrişimli sinir ağları kullanılarak daha yüksek görüntü kalitesi sonuçları rapor edilmiştir. Ancak hiç bir çalışmada HDR Ton Dönüştürme işleminin aynı zamanda uygulanması amaçlanmamıştır.

Bu tezde, belirli bir Bayer paterne sahip ham görüntülerden Ton Dönüştürme işlemi yapılmış yüksek kaliteli HDR görüntü eşleniğini oluşturmak üzere yeni bir ağ mimarisini önerilmektedir. Bu tez kapsamında ayrıca Zurich veri setinde yer alan görüntülerin alt kümesi Topaz Gigapixel AI ve Photomatrix Pro 5.0 yazılımı ile düzenlenerek uygun veri seti oluşturulmuştur. Oluşturulan görüntülerde daha kararlı renk bilgisi elde etmek amacıyla, literatürde kamera ISP veriyolu çalışmalarından farklı olarak ilk defa kayıp fonksiyonunda histogramdan faydalanılmıştır.

Tez kapsamında gerçekleştirilen deneyler sonucunda, belirli bir Bayer paterne sahip ham görüntülerden referanssız görüntü kalite metrik performansına göre yüksek kaliteli çıktılarının oluşturulabileceği görülmektedir. Referans model çıktıları ve ticari kameralardaki mevcut ISP çıktıları ile yapılan karşılaştırma sonuçları göz önüne alındığında, çözümümüzün ileride gerçek bir kameraya entegre edilebilir olduğunu göstermektedir.

**Anahtar Kelimeler:** Derin Öğrenme, Evrişimli Sinir Ağları, Görüntü İşleme, Kamera ISP Veri Yolu.

# MODELING OF CAMERA PIPELINE THAT CAN PRODUCE HDR IMAGES USING DEEP NETWORK ARCHITECTURE

## ABSTRACT

Today, a conventional digital camera ISP pipeline is constructed by sequentially connected image processing modules which are each designed for a specific task, such as De-mosaicking, De-noising, White Balance, Color Enhancement, Tone-mapping etc. Since these modules are designed independently and the pipeline cannot be optimized end-to-end, artifacts are likely to be observed in the final generated image. To eliminate these effects, various camera ISP pipeline models have been proposed in the literature and higher image quality results have been reported using convolutional neural networks. However, it is not intended to apply HDR Tone-mapping at the same time in any study.

In this thesis, a new network architecture is proposed to create high-quality HDR Tone-mapped images from raw images with a specific Bayer pattern. Within the scope of this thesis, a subset of the images in the Zurich dataset was edited with Topaz Gigapixel AI and Photomatix Pro 5.0 software to create suitable dataset. In order to obtain more stable color information, histogram was used in the loss function for the first time, unlike camera ISP pipeline studies in the literature.

As a result of the experiments carried out within the scope of the thesis, it is seen that high quality outputs can be created from raw images with a certain Bayer pattern according to no-reference image quality metric performance. Considering the comparison results with reference model outputs and existing ISP outputs in commercial cameras, it shows that our solution can be integrated into a real camera in the future.

**Keywords:** Deep Learning, Convolutional Neural Networks, Image Processing, Camera ISP Pipeline.

## 1. GİRİŞ

İnsan Görme Sistemi (İGS),  $10^{-6}$  'dan  $10^8$   $cd/m^2$  değeri skalasında ışıklılığı algılayabilmektedir; fakat standart monitörler günümüzde  $0.5-500$   $cd/m^2$  aralığında ışıklılığı yaklaşık olarak sergileyebilmektedir. Referans teşkil etmesi açısından doğal ışık kaynaklarının ışıklılık değerleri ölçeği Tablo 1.1.'de sunulmuştur.

Tablo 1.1. İGS için parlaklık skalası

Yıldız Işığı		Ay Işığı	İç Mekan Aydınlatma		Dış Mekan Aydınlatma	Gün Işığında Dış Mekan	Güneş
Işıklılık $cd/m^2$	$10^{-6}$	0,001	1	100	10000	$10^6$	$10^8$

Dinamik aralık, bir fotoğraftaki piksel değerlerinin yansıtmakta olduğu en yüksek parlaklık değerinin en düşük parlaklık değerine oranı olarak ifade edilmekte olup, sahneye ait ışıklılık ölçeğinin görüntüleme sistemindeki karşılılığı ile ilişkilidir. Bir ışık ölçme cihazının, sahnedeki minimum ve maksimum ışıklılık değerlerini ihtiva edecek şekilde geniş dinamik aralıkta görüntü elde etmesi, günümüzde mümkün değildir. Kullanılan kameranın sensör, boyut kısıtları ile standart monitörlerin kısıtları, limitli bir parlaklık ve renk uzayında çalışılmasına olanak sağlayabilmektedir.

Fotoğrafçılar kameranın pozlama değerini yani film üzerinde metre kareye düşen ışık miktarını ayarlayarak fotoğrafın dinamik aralığını sınırlar çerçevesinde isteklerine göre değiştirebilirler. Modern kameralar otomatik pozlama özelliğini de sunmaktadır; bu mod optimum performansı sağlamak üzere sahnedeki objenin orta tonu ile elde edilecek resmin orta tonunu mümkün olan en iyi şekilde uyumlandıracak pozlama değerinin kullanılmasıyla uygulanır. Fakat orijinal sahnede genelde geniş bir dinamik aralık bulunması nedeniyle, elde edilen görüntüde yüksek ve düşük pozlama bölgelerinin oluşması kaçınılmazdır.

Günümüzde verimli görüntü işleme teknikleri kullanılarak LDR (Low Dynamic Range) görüntüleri daha geniş bir parlaklık ölçeğinde oluşturmak mümkündür. HDR (High Dynamic Range) görüntü oluşturma teknikleri, sunulan resimdeki renk ve parlaklıkları neredeyse direkt olarak sahnenin kendisinden İGS ile tecrübe ediyormuşçasına bir kalitede algılamamıza olanak sağlar. Günümüzde kullanılan standart monitörlerde HDR

görüntünün görüntülenebilmesi için ton dönüştürme işlemi yapılmasına ihtiyaç duyulmaktadır. Şekil 1.1'de LDR ve ton dönüşümü yapılmış HDR görüntü örnek olarak sunulmuştur. Sağdaki görüntü soldaki görüntü ile karşılaştırıldığında; görüntünün içerisindeki aşırı parlak ve karanlık bölgelerdeki pikseller daha geniş kontrast skalasında ifade edilebildiği gözlemlenmektedir.



Şekil 1.1. LDR (solda) ve HDR (sağda) görüntü (URL-1)

Günümüzde bilimsel bir sınıflandırma tanımı bulunmamakla birlikte standart (düşük) ve yüksek dinamik aralığa sahip görüntülerin genel özellikleri (maksimum parlaklık, dinamik aralık, kontrast, bağlılık ve kuantalama) Tablo 1.2'de sunulmuştur.

Tablo 1.2. HDR ve LDR görüntü karşılaştırma tablosu

	Standart (Düşük) Dinamik Aralığa Sahip Görüntü (LDR)	Yüksek Dinamik Aralığa Sahip Görüntü (HDR)
Maksimum Parlaklık	500 cd/m <sup>2</sup> 'ye kadar	2000-10000 cd/m <sup>2</sup>
Dinamik Aralık	50 dB	120 dB
Görüntü Kontrastı	1:1000	1:1000000
Bağlılık	Görüntüleme Cihazı Sınırı	İnsan Görüş Sistemi sınırı
Kuantalama	6 – 16 bit	Kayan Sayı (Floating point)

HDR görüntüler sanal gerçeklik uygulamalarında, reklam sektöründe, makina görmesi uygulamalarında (medikal resim oluşturma, otonom araç sürüsü vb.), bilimsel resim görüntüleme işlemlerinde (astronomik gözlemler, arkeoloji ve müze çalışmalarında) kullanılmaktadır. Günümüzde yüksek dinamik aralığa sahip görüntüleri oluşturma

tekniklerinin gelişmesiyle birlikte HDR görüntü kullanımının gündelik hayatta da yaygınlaşması beklenmektedir.

Şimdiye kadar birçok görüntü işleme tekniği ve çözüm önerisi daha geniş bir dinamik aralığı ihtiva eden fotoğrafların elde edilebilmesi için önerilmiştir. Bu yöntemlerden bazıları Kamera ISP veri yolu içerisinde uygulanırken, bazıları ise lensten geçerek sensör üzerine düşen ışık miktarını sahnenin ışıklılığına göre açıklık (aperture), perde hızı (shutter speed) parametreleri ile optimum ayarlanması şeklinde uygulanır. Daha güncel çalışmalarda ise farklı pozlama değerleri ile elde edilen görüntünün sinyal işleme teknikleri ile birleştirilmesi ya da tek bir görüntüden HDR görüntü elde edilmesine yönelik çalışmaların yapıldığı görülmektedir. Yapay Sinir Ağlarının ve Derin Öğrenme mimarilerinin Bilgisayarlı Görü (Computer Vision) problemlerinde kullanımının yaygınlaşması ile birlikte daha geniş dinamik aralığa sahip görüntü elde etmeyi amaçlayan çözümlerin, veriden öğrenme metotları ile uygulanmasına başlanmıştır.

Bu çalışmaların paralelinde, özellikle görüntü iyileştirme (image enhancement), süper çözünürlük (super resolution), görüntü gürültü giderici (image denoising) derin öğrenme çalışmalarının yaygınlaştığı bu dönemde, kamera ISP veri yolunun derin ağ mimarisi ile modellenmesine yönelik çalışmalara da başlandığı literatürde görülmektedir.

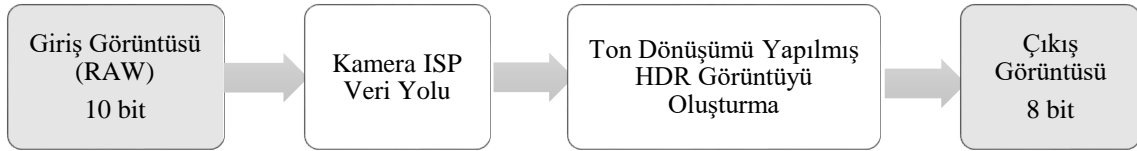
Geleneksel Kamera ISP Kamera Veri Yolu birçok sıralı sinyal işleme modülünden oluşmakta ve her biri farklı bir görevi/fonksiyonu yerine getirmektedir. Kameralarda başlıca kullanılmakta olan modüller; Otomatik Beyaz Dengeleme (Automatic White Balance), Renk Düzeltme, Kenar Pekiştirme, Gürültü Azaltıcı olarak sıralanabilir. Günümüzde bu bloklar çeşitli sinyal işleme yöntemi kullanılarak oluşturulmakta, amacına göre farklı renk uzaylarında görüntüyü işlemekte ve sonrasında ise uç uca eklenmiş (cascade) yapıda birleştirilmektedir. Bu geleneksel yöntemde herhangi bir alt-modülde oluşan arıza bir sonraki bloğa taşınmakta ve kullanıcıya sunulan görüntüde çeşitli bozulmalar ile karşılaşabilmektedir. Bu nedenle tüm modellerin en iyi görüntüyü sunmak üzere optimize edilmesine ihtiyaç duyulmaktadır. Bununla birlikte küçük sensörlerin ve kompakt lenslerin kullanılması nedeniyle, geleneksel kamera ISP veri yolu ile giderilemeyen çeşitli bozulmalar (detayların kaybedilmesi, yüksek gürültü

seviyesi, renk bozulmaları vb.) nihai görüntüde oluşabilmektedir (Ignatov ve diğ., 2020).

Kamera ISP veri yolunun derin öğrenme mimarileri ile modellenmesine yönelik çalışmalarda, veri yolunun baştan sona optimizasyonu sağlanarak, söz konusu bozucu etkilerin giderilmesi ve kullanıcıya sunulan nihai görüntü kalitesinin artırılması hedeflenmektedir.

### 1.1. Çalışmanın Amacı

Bu tez çalışması kapsamında, 10 bitlik Bayer Paterne sahip kamera sensöründen alınan ham veri giriş verisi olarak kullanılarak, HDR ton dönüşümü yapılmış 8 bitlik görüntüyü çıkış görüntüsü olarak elde edebilecek Kamera ISP veri yolu, konvolüsyonel yapay sinir ağlarından oluşan bir derin ağ mimarisi ile modellenmiştir. Bu maksatla geliştirilen modelin genel akış diyagramı Şekil 1.2’de sunulmuştur. Diyagramda belirtilen Kamera ISP Veri Yolu ve Ton Dönüşümü Yapılmış HDR Görüntüyü Oluşturma blokları, geliştirilen ağ mimarisinde bütünleşik olarak uygulanmıştır. Geliştirilen ağın ileride kameraya entegre edilebilecek ve gerçek zamanlı olarak görüntüyü kullanıcıya sunabilmesi için düşük çıkarım zamanına sahip olması hedeflenmiştir. Eğitim ve test esnasında Kamera ISP Veri Yolu’nun modellenmesi amacıyla oluşturulan Zürich Veri Seti’nde yer alan tam çözünürlükteki ham veri ve RGB görüntü çiftleri kullanılmıştır. Veri seti içerisinde yer alan RGB görüntüler, Topaz Gigapixel AI ve Photomatix Pro 5.0 fotoğraf düzenleme yazılımları ile düzenlenmiş; böylelikle görüntülerdeki gürültü azaltılmış, düşük/yüksek pozlama bölgelerindeki detaylar iyileştirilerek eğitim ve test aşamaları için hedef görüntüler hazır hale getirilmiştir.



Şekil 1.2. Model genel akış diyagramı

Bu tez çalışmasının hedefleri ve katkıları aşağıdaki gibi özetlenmiştir:

- Ham veriden ton dönüşümü yapılmış HDR görüntüyü elde edebilecek konvolüsyonel

yapay sinir ağlarından oluşan ağ mimarisi geliştirilmiştir.

- Fotoğraf düzenleme yazılımları ile Zürich Veri Seti'nde yer alan tam çözünürlükteki görüntü çiftlerinden HDR ton dönüşümü yapılmış veri seti türetilerek diğer araştırmacıların kullanımına sunulmuştur.
- Geliştirilen model içerisinde yer alan Dikkat Modülleri ve Kayıp Fonksiyonlarının etkisi eğitim ve test aşamalarında irdelenmiştir.
- Çıktı görüntülerinin başarımı referansa ihtiyaç duymayan görüntü kalitesi metrikleri ile başarımı incelenmiştir.

Tez çalışması, yeni nesil kameraların daha az elektronik donanım ile derin öğrenme temelli bir model ile tasarlanabileceği hedefi doğrultusunda gerçekleştirilmiştir. Bu çalışmanın gelecekte daha kompakt yapıda modellerin geliştirilmesi, daha geniş veri seti ile birlikte daha yüksek kalitede çıkış görüntüsü elde edilmesi ve geliştirilen ağın kameraya entegrasyonun gerçekleştirilmesi amacıyla yapılacak çalışmalara katkı sağlaması amaçlanmıştır.

Tezin bölüm bölüm organizasyonu şu şekildedir:

- Bölüm 2, tez kapsamı dâhilinde ilgili çalışmaları inceler. Kapsam dâhilinde derin öğrenme mimari ile HDR görüntü elde etme yöntemleri ve Kamera ISP Veri Yolu'nun modellenmesi konusunda literatür özeti verir.
- Bölüm 3, görüntü kalitesini değerlendirmek üzere yaygın kullanılan referansa ihtiyaç duyan ve duymayan görüntü kalitesi değerlendirme yöntemlerini açıklar.
- Bölüm 4, tez kapsamında kullanılan veri seti ve veri oluşturma yöntemini açıklar.
- Bölüm 5, geliştirilen ağ mimarisinin tasarımı ve eğitim aşaması açıklar.
- Bölüm 6, test aşaması çıktılarını ve performans sonuçlarını verir.
- Bölüm 7, tez çalışmasının neticesi ve gelecek çalışmalar hakkında bilgi verir.

## 2. ÖNCEKİ ÇALIŞMALAR

Bu bölüm, tezin temelini oluşturan temel kavramları açıklamayı ve literatürde yer alan öncü çalışmaların kapsamını aktarmaya çalışmaktadır. Tezin kapsadığı çalışma alanı, çeşitli araştırma alanlarını içermektedir ve dört ana sınıfta toplanabilir:

- Kamerada Görüntünün Oluşturulması ve Geleneksel ISP Veri Yolu,
- Derin öğrenme ile Kamera ISP Veri Yolu'nun oluşturulması çalışmaları,
- Derin öğrenme teknikleri ile HDR Görüntü Oluşturma çalışmaları,
- Ton Dönüştürme Operatörü çalışmaları,

Bu bölüm içerisinde öncelikle dijital kamerada görüntünün nasıl oluşturulduğu ve geleneksel olarak kullanılan kamera ISP veri yolu içerisinde yer alan sinyal işleme bloklarının tanıtımı yapılacaktır. Sonrasında ise sırasıyla kamera ISP veri yolunun geliştirilmesi, HDR görüntü oluşturma ve Ton Dönüştürme işlevinin derin öğrenme mimarileri ile yerine getirilmesine yönelik literatürde yer alan çalışmalar hakkında bilgi verilecektir.

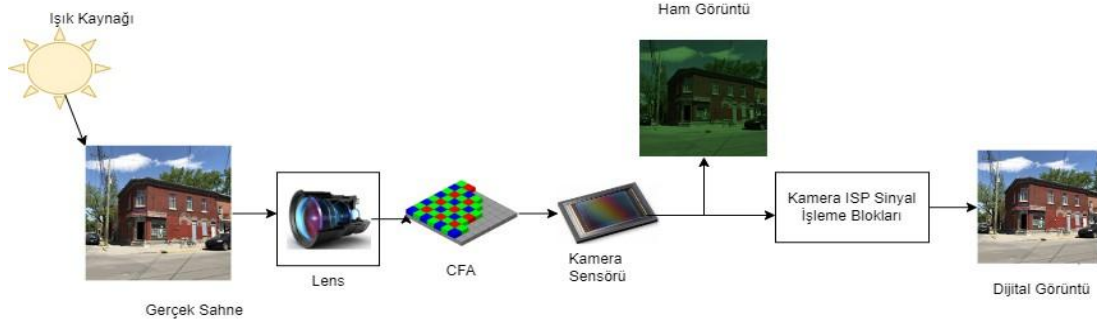
### 2.1. Kamerada Görüntünün Oluşturulması ve Geleneksel Kamera ISP Veri Yolunun Yapısı

En basit tanımıyla kamera ışık ölçüm aletidir. Işık aslında 390-57700 nm dalga boyuna sahip elektromanyetik dalgalar için kullanılan bir tanımlamadır. Günümüzde, ışığın direkt olarak kimyasal filmlerin üzerine düşürülerek görüntünün oluşturulduğu analog kameraların aksine, dijital kameralar üretilmekte ve kullanılmaktadır.

Gerçek sahneden elde edilen görüntünün dijital görüntüye dönüştürülme ve sonrasında amaca göre anlamlandırma süreci Şekil 2.1'de gösterilmiştir. Işık kaynağı sayesinde aydınlatılan ve sahnede yer alan objeler tarafından yansıtılan ışınlar ilk olarak kamera lensine ulaşmaktadır.

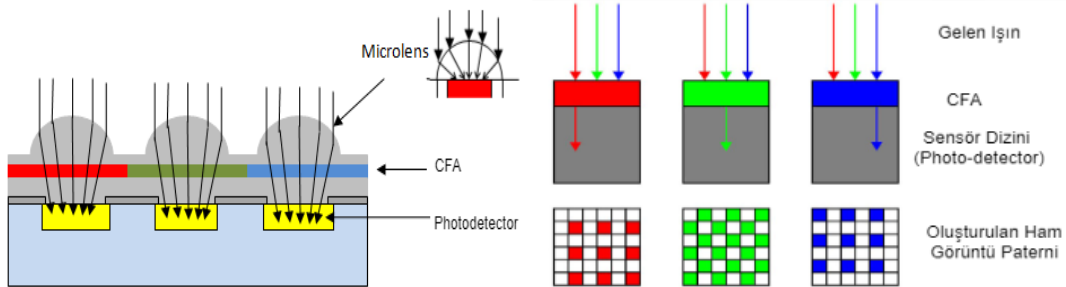
Obtüratörün kapanması ile lensten geçerek toplanan ışınlar, renk bilgisinin elde edilebilmesi amacıyla CFA (Color Filter Array) dizininden geçirilerek Şekil 2.2.'de gösterildiği gibi sensör üzerine düşürülür. CFA dalga boyuna göre bant geçiren filtrelerdir ve genellikle İnsan Görme Sistemi (HVS-Human Vision System)'nin en

duyarlı olduğu (Bayer, 1976) Kırmızı (R), Yeşil (G) ve Mavi (B) renklere göre dizayn edilmiştir. İGS'nin yeşile daha fazla duyarlılık göstermesi nedeniyle diğerlerine oranla iki kat fazla yeşil geçiren filtrenin kullanıldığı Bayer Paterni, en sık kullanılan CFA paternidir.



Şekil 2.1. Görüntü oluşturma/anlamlandırma diyagramı

Kimyasal film yapılarının kullanıldığı analog kameraların aksine, dijital kameralarda CCD (Charge-Coupled Device) ya da CMOS (Complementary Metal Oxide Semiconductor) sensörler vasıtasıyla sahneye ait bilgiyi ihtiva eden fotonlar, elektriksel yük ya da akım olarak elektriksel sinyallere dönüştürülür.



Şekil 2.2. Sensör (Photodetector) ve CFA yapısı (URL-2)

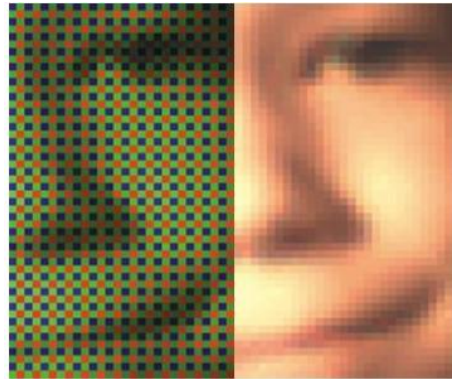
Elde edilen sinyaller analog dijital çeviriciler vasıtası ile dijital sinyallere dönüştürüldükten sonra, nihai görüntünün oluşturulması amacıyla sensör tarafından ISP Veri Yolu içerisinde yer alan çeşitli sinyal işleme bloklarında işlenir. En basit haliyle Kamera ISP Veri Yolu Şekil 2.3.'de (Liang ve diğ., 2019) belirtilen Renk Geri Mozaikleme (Color Demosaicing), Gürültü Giderme, Beyaz Dengeleme ve Renk Uzayı Dönüştürme (White Balance ve Color Space Conversion), Renk İyileştirme (Color Enhancement) ve Ton Dönüştürme (Tone Mapping) bloklarından oluşmaktadır. Piyasada yer alan kameralara ait veri yolunda yer alan blokların akış sırası, dizaynı ve

parametreleri farklılıklar gösterebilmekte ve bilgilerin ticari gizli olması nedeniyle paylaşımına açık olarak sunulmamaktadır.



Şekil 2.3. Kamera ISP Veri Yolu Genel Yapısı

Renk Geri Mozaikleme modülü, CFA kullanımını nedeniyle renk bilgisinin eksik olduğu pikseller için değerlerini tahmin etmek maksadıyla kullanılmaktadır, bu nedenle bir enterpolasyon işlemi olarak kabul edilmektedir. Bu işlem genel olarak komşu piksel değerleri, kenar ve gradyant gibi görüntüden elde edilen bilgiler vasıtasıyla gerçekleştirilir. Geleneksel kamera veri yolunda Laplacian, Bilinear Filter, Bicubic gibi temel enterpolasyon yöntemleri kullanılmakla birlikte, literatürde sonradan önerilen Malvar metodu (Malvar ve diğ., 2004), adaptif homojen-yönlü geri demozikleme metodu (Hirakawa ve Parks, 2005) gibi çeşitli yöntemler de bu amaçla kullanılabilir. Geri Mozaikleme işlemi yapılmış ve yapılmamış görüntünün temsili örneği Şekil 2.4’de sunulmuştur.



Şekil 2.4. Geri Mozaikleme yapılmamış (solda) ve yapılmış (sağda) görüntü

Gürültü Giderme işlemi sensör etkisinden ve Geri Mozaikleme işleminden kaynaklı olan gürültünün, görüntü içerisinden yok edilmesi/azaltılması işlemidir. Bilateral Filters, Median, Non Local Means (NLM), Block-Matching and 3D (BM3D) filtering gibi geleneksel yöntemlerin yanısıra; literatürde sonradan önerilen örneğin Y.Baek’in (Baek ve diğ., 2008) çalışması gibi yöntemler de kullanılabilir.

Beyaz Dengeleme Bloğu, İGS'nin görüntü içerisindeki beyaz rengi farklı ve değişen ışık kaynakları altında doğru anlamlandırma yeteneğini taklit eden işlevi yerine getirir. Yani piksellerin ışık şiddetlerini sahnedeki ışık kaynağına ve renklerin sıcaklığına göre ayarlanır. Literatürde W.Hsu'nun (Hsu ve Cheng, 2021), S.Wang'ın (Wang ve diğ.,2011), Z.Lou'nun (Lou ve diğ., 2015), Y.Hu'nun (Hu ve diğ., 2017), Barron'un (Barron, 2015) gibi önerilen görüntü işleme ve derin öğrenme temelli çeşitli çalışmalar mevcuttur. Beyaz Dengeleme uygulanmamış ve uygulanmış görüntü Şekil 2.5'te sunulmuştur. Beyaz Dengeleme bloğunun ve sonraki Renk İyileştirme bloğunun farklı renk uzayında görüntüyü işlemesi uygulamalarda tercih edilebilmektedir.



Şekil 2.5. Beyaz Dengeleme uygulanmamış (solda) ve uygulanmış (sağda) görüntü (Barron, 2016)

Renk İyileştirme Bloğunda da önceki bloklarda olduğu gibi görüntü işleme ve/veya makine öğrenmesi, derin öğrenme temelli önerilen birçok çalışma literatürde yer almaktadır. Bu bloğun temel işlevi sensör tarafından elde edilen ve veri yolu boyunca işlenen görüntüdeki renk bilgisindeki bozucu etkilerin giderilmesi ve iyileştirilmesidir. Literatürdeki çalışmalara S.Nikam'ın (Nikam ve Yawale, 2015), Y.Ueda'nın (Ueda ve Suetake, 2019) ve Y.Kinoshita'nın (Kinoshita ve diğ., 2020) çalışmaları örnek olarak verilebilir.

Şekil 2.3'te sunulan kamera ISP veri yolu içerisinde yer alan son bloğun işlevi olan Ton Dönüştürme işlemi, görüntünün dinamik aralığının monitör ve projektörlerde gösterilebilmesi amacıyla kaçınılmaz olarak kullanılan daraltılması amacıyla uygulanmaktadır. Bu işlem esnasında görüntüdeki detayların, renk bilgisinin ve yerel

kontrast özelliklerinin mümkün olduğunca korunmasına ihtiyaç duyulmaktadır. Bu alandaki öncü çalışmalar E.Reinhard tarafından (Reinhard ve diğ., 2002), F.Durand tarafından (Durand ve Dorsey, 2002), R.Mantiuk (Mantiuk ve diğ., 2008) olarak gösterilebilir. Günümüzde HDR görüntü/videoların, LDR temelli medyalarda görüntülenme ihtiyacı nedeniyle uygulanan metotlar da Ton Dönüştürme işlemi olarak anılmaktadır. J.Tan'ın çalışması (Tan ve diğ., 2018) ve K.Panetta'nın çalışması (Panetta ve diğ.,2021) bu alandaki çalışmalara örnek olarak verilebilir. Aynı zamanda görüntü içerisindeki detayları ortaya çıkarmak ve görüntü kontrastını artırmak işlemi amacıyla uygulanan çeşitli teknikler de aynı isimle anılmaktadır. Bu nedenle yerel ton dönüştürme işlemine yönelik çalışmalar da literatürde yaygın olarak yer almaktadır. Bilcu'nun (Bilcu ve diğ., 2012) ve Koscevic'in (Koscavic ve diğ., 2021) çalışmaları yerel ton dönüştürme çalışmalarına örnek olarak gösterilebilir.

Yukarıda bahsedilen bloklara ek olarak kamera ISP veri yolu içerisinde; Kenar Bilgisi Pekiştirme işlemi, görüntünün kaydedilme formatına göre sıkıştırma ve Gamma Düzeltmesi işlemleri de yapılabilmektedir.

## **2.2. Kamera ISP Veri Yolunun Derin Öğrenme Mimarisi ile Oluşturulması**

Teknolojinin günlük yaşamdaki kullanımı arttıkça, markette yer alan elektronik ürünlerin rekabet çerçevesinde daha kullanışlı yapıda kompakt olarak kullanıcıya sunulması kaçınılmaz olmuştur. Kameralarda yer alan optik elemanların daha küçük boyutlarda seçilmesinin, elde edilen görüntünün kalitesini doğrudan olumsuz olarak etkilemesi kaçınılmazdır. Bu nedenle, kompakt tasarımların kamera üzerindeki elektronik donanımda gerçekleştirilmesi daha uygulanabilir bir çözümdür. Bu nedenle geleneksel olarak Bölüm 2.1.'de belirtilen sinyal işleme bloklarının baştan-sona derin öğrenme mimarisi ile modellenmesi geleneksel kamera veri yoluna bir alternatif olarak literatürde yer almaya başlamıştır. Sunulan çözümlerin performans sonuçları, geleneksel yöntemler ile karşılaştırıldığında, ham veriden çıkış görüntüsünü elde etmek üzere baştan sona derin öğrenme ile eğitilen ağ mimarilerinin daha başarılı sonuçlar verebildiği görülmektedir.

Literatürde konuya ilişkin önerilen ilk çalışma Chen (Chen ve diğ., 2018) tarafından düşük pozlama ile elde edilen görüntülerde oluşan gürültü ve renk yanlılık (bias)

problemını gidermek amacıyla önerilmiştir. Bu çalışma kapsamında sahneye ait kısa pozlama değeri ile elde edilen ham veriler ile aynı sahneye ait uzun pozlamalı RGB hedef görüntülerden oluşan SID veri seti araştırmacıların kullanımına sunulmuştur. Önerilen CNN tabanlı model mimarisi ile ağır baştan sona eğitimi gerçekleştirilmiş ve geleneksel Kamera ISP veriyoluna alternatif olarak oluşturulan model sunulmuştur.

Literatürde bir diğer çalışma olan DeepISP (Schwartz ve diğ., 2019) çalışmasında, iki katmanlı konvolüsyonel bir ağ mimarisi kullanılarak düşük ve yüksek seviyeli özneliklerin öğrenilmesi hedeflenmiştir. Kayıp fonksiyonu olarak MS-SSIM (Multi Scale Structural Similarity Index) (Wang ve Bovik, 2003) ve  $L_1$  kayıp fonksiyonu birlikte kullanılmış ve eğitilen modelin MOS ve DeepIQA (Bosse ve diğ., 2018) metrikleri ile elde edilen sonuçları göz önüne alındığında Samsung S7'ye ait ISP'ye kıyasla daha başarılı sonuç üretebildiği gösterilmiştir.

Bir diğer öncü çalışma olan CameraNet ise (Liang ve diğ., 2019); RestoreNet ve EnhanceNet olarak adlandırılan iki adet U-Net mimarisine sahip alt modül kullanılarak ham veriden hedef görüntülerin oluşturulması hedeflenmiştir. Bu maksatla HDR+, SID and FiveK verisetleri kullanılarak modelin başarımı PSNR, SSIM ve renk hatası bilgisi metrikleri ile DeepISP çalışmasının sonuçlarıyla kıyaslanmış ve başarımın artırıldığına vurgu yapılmıştır.

U-Net yapısına sahip ağ yapılarının çeşitli bilgisayarlı görü problemlerinde kullanımının yaygınlaşmaya başlamasıyla birlikte, çeşitli görüntü iyileştirme ve düşük ışıklığa sahip görüntülerin tekrardan oluşturulmasına yönelik problemlerde kullanımının arttığı görülmektedir. U-Net mimarisi tabanlı önerilen modellerden biri, dikkat mekanizmasının ve dalgacık dönüşümünün kullanıldığı AWWNet (Dai ve diğ., 2020) olmuştur. Kanal ve Uzamsal Dikkat Modülleri ile oluşturulan Çift Dikkat Modülleri'nin kullanımıyla, CSANet'te (Hsyu ve diğ., 2021) daha az parametre ile daha iyi performans elde edilebildiği görülmüştür.

Daha sonraki çalışmalarda özellikle gerçek zamanlı uygulamalar için model karmaşıklığını azaltmak üzere çeşitli yöntemler önerilmiştir. Bu çalışmalardan bazıları Lightweight PyNet (Cheng ve diğ., 2021), PyNet-V2 Mobile (Ignatov ve diğ., 2022),

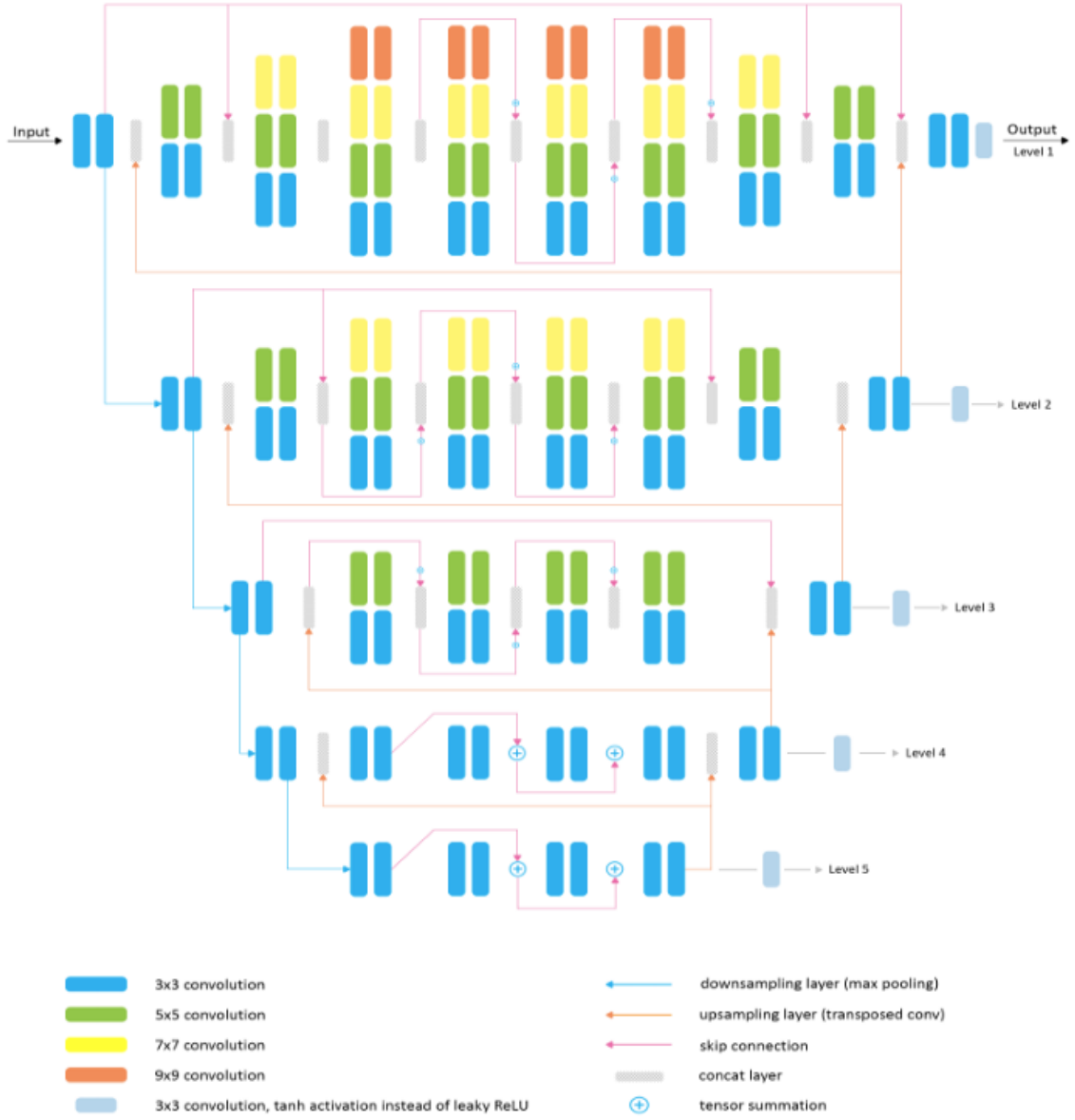
LW-ISP (Chen ve Ma, 2022) ve MicroISP (Ignatov ve diğ., 2022) olarak örneklendirilebilir.

Del-Net (Gupta ve diğ., 2021) ile SCA (Spatial and Channel Attention) Modülü ve EAM (Enhancement Attention Module) bloklarının konuşlandırıldığı çok ölçekli bir mimari önerilmiştir. Bir başka ilham verici çalışma, kullanıcıların bellek kısıtlamaları nedeniyle büyük miktarlarda ham veri depolamaktan kaçınmasına yardımcı olmak için gerektiğinde sRGB görüntüsünden ham veri elde edebilen ve gerektiğinde tam tersini yapabilen InvISP (Xing ve diğ., 2021) ile önerilmiştir. ReconfigISP'de (Yu ve diğ., 2021) ise ISP modüllerinin tam optimizasyonu amaçlanmış, bu çalışma kapsamında farklı ISP modülleri modellenmiş ve proxy ayarlama mekanizması ile optimize edilmiştir. UnISPNet (Mosleh ve diğ., 2023) ile kamera ISP veriyolu, yuvarlanmamış optimizasyon tekniği kullanılarak oluşturulmuştur. UnISPNet modeli, Gürültü Giderici CNN bloklarının cascade yapıda eklenmesiyle oluşturulmuştur. Bu çalışmada ayrıca görüntünün histogram bilgisi çıktı görüntüsünü oluşturmak amacıyla Tone Map Encoder aracılığıyla kullanılmıştır.

### **2.2.1 PyNET Mimarisi**

Tez kapsamında referans çalışma olarak kabul PyNET (Ignatov ve diğ., 2020) çalışmasında, Şekil 2.6'da sunulan ters piramit yapısında konvolüsyonel yapay sinir ağı mimarisi oluşturulmuştur. Modelin eğitimi önerilen çalışma kapsamında araştırmacılara sunulan Zurich veriseti ile gerçekleştirilmiştir. Model, her bir katmanın ayrı ayrı eğitildiği beş adet katmandan oluşmakta olup; modelin beşinci katmandan başlayarak eğitimi gerçekleştirilmektedir. Her bir katmanda farklı çekirdek boyutuna sahip konvolüsyon blokları yer almaktadır. Öğrenmenin kolay gerçekleştirilebilmesi için bloklar arasında atlama bağlantıları kullanılmıştır.

Giriş ve hedef görüntüsünün uzamsal çözünürlüğü ikinci katmandan başlayarak iki çarpanıyla en büyük havuzlama yöntemiyle azaltılmaktadır. Böylelikle ilk katman hariç diğer katmanların eğitimi giriş görüntüsünden daha düşük bir çözünürlükte gerçekleştirilmektedir. Hedef görüntüler de alt katmanların eğitimi için bu doğrultuda yeniden ölçeklendirilmektedir.



Şekil 2.6. PyNET model mimarisi (Ignatov ve diğ., 2020)

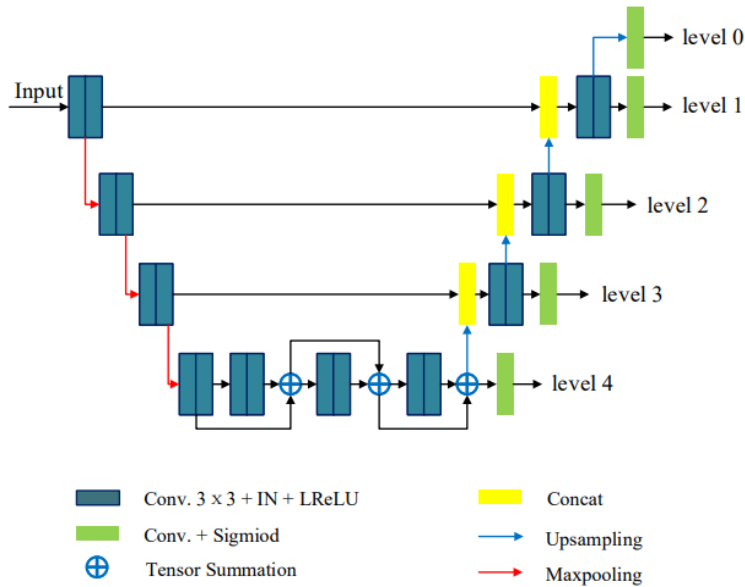
Birinci katmanda ise eğitim işlemi giriş görüntüsünün uzamsal çözünürlüğünde gerçekleştirilmektedir. Eğitim esnasında alt katmanlardan her çıkarılan öznetelik haritası upscale edilir ve bir sonraki katmanın girişi için çıkartılan öznetelik haritası ile birleştirilir. Bu sayede düşük seviyeli özelliklerden başlayarak ağ yapısının eksik detayları öğrenmesi ve oluşturulan nihai görüntüdeki detay bilgilerde iyileştirme sağlanması hedeflenir. Model mimarisinde Relu aktivasyon fonksiyonu, her katmandaki son çıkış blokları hariç, tüm bloklarda kullanılmaktadır. Her bir katmandaki çıkış blokları için ise Sigmoid aktivasyon fonksiyonundan yararlanılmaktadır. Bahse konu

çıkış blokları her bir tur sayısı (epoch) sonunda elde edilen çıktılarının görselleştirilmesi ve eğitim durumunun takip edilebilmesi amacıyla kullanılmaktadır.

PyNET modelinin referans olarak seçilmesindeki en önemli etken bu alanda öncü bir çalışma olmasıdır. Bununla birlikte; (448, 448) çözünürlükteki giriş görüntüleri kullanılarak her bir katmanın ayrı ayrı eğitilebiliyor olması işlemci gücünün daha verimli kullanılabilmesine imkân sağladığı görülmüştür. Bununla birlikte modelde toplamda 47.5 milyon eğitilebilir parametre bulunmaktadır. Önerilen model başarımının, 2020 yılında literatürde yer alan çalışmalara kıyasla daha yüksek olmasına rağmen, model karmaşıklığının kamera donanımına entegre edilebilmesi için uygun olmadığı görülmektedir.

## 2.2.2 Lightweight PyNET Mimarisi

PyNET modelini referans alarak, parametre sayısını artırmayı amaçlayan ilk çalışma Lightweight PyNET (Cheng ve diğ., 2021) olmuştur. Lightweight PyNET mimarisi Şekil 2.7’de sunulmuş olup, model içerisinde toplamda 5.7 milyon parametre yer almaktadır. Model başarımının daha hafif yapıda olmasına rağmen, sezgiye uygun olarak PyNET’den daha iyi olmadığı görülmektedir.



Şekil 2.7. Lightweight PyNET model mimarisi (Cheng ve diğ., 2021)

Sunulan bu çalışma ile kamera entegre edilebilir hafif yapıda bir ağ mimarisinin başarımdan ödün verilerek geliştirilmesi hedeflenmiştir.

### **2.3. HDR Görüntünün Oluşturulması**

Bugüne kadar bir fotoğraf içerisinde yer alan yüksek ve düşük pozlama bölgelerindeki detay bilgileri ortaya çıkararak aynı zamanda daha geniş dinamik aralıklı görüntüler elde etmek için birçok görüntü işleme tekniği önerilmiştir. HDR görüntü elde etmek için temelde iki farklı yaklaşım vardır;

- Çoklu pozlamalarla elde edilen LDR görüntülerinin birleştirilmesi
- Tek bir LDR görüntüsünden HDR görüntüsü oluşturma

#### **2.3.1. Çoklu pozlamalarla elde edilen LDR görüntülerinin birleştirilmesi**

İlk yöntem, derin öğrenme yaklaşımlarından önce farklı pozlama değerleriyle aynı sahneye ait elde edilen LDR görüntülerden özneliklerin çıkartılması ve optimal ağırlıklandırma fonksiyonlarının kullanılmasıyla LDR görüntülerdeki piksellerin HDR görüntünün oluşturulması sağlanarak literatürde önerilmiştir. Bu çalışmalardan en bilineni Debevec'in önerdiği metottur (Debevec ve Malik, 1997). Bu alana katkı yapan diğer öncü çalışmalara ise Mann'ın (Mann ve Picard, 1995), Mitsunaga'nın (Mitsunaga ve Nayar, 1999) ve Robertson'un (Robertson ve diğ., 2000) sunduğu yöntemler örnek olarak verilebilir. Anılan bu yöntemlerin ilk amacı genel olarak, bilinmeyen kamera tepki fonksiyonunu (CRF) tahmin etmek, ardından mümkün olduğunca sahne ışıklılığına ulaşmak için görüntüleri doğrusallaştırmaktır. Bir sonraki adımda ise, doğrusallaştırılmış farklı pozlama değerleri ile elde edilen LDR görüntüler, HDR görüntüsünü oluşturmak için piksel piksel uygun ağırlıklandırma ile birleştirilir. K.Kirk (Kirk ve Andersen, 2005), Tsin (Tsin ve diğ., 2001) ve M. Granados (Granados ve diğ., 2010) tarafından önerilen çalışmalarda HDR görüntüleri oluşturmak için daha az parametre ve daha verimli birleştirme tekniklerinin kullanılmasına odaklanmıştır. Bu çabalar, LDR görüntülerin mükemmel şekilde hizalandığı durumlarda kabul edilebilir çıktılar üretebilse de; gerçek dünya senaryolarında, özellikle hareketli nesnelere kaynaklanan hizalama bozuklukları olması durumunda, oluşturulan HDR görüntüsünde "ghosting effect" adı verilen bozulmalar meydana gelir. Bu etkiyi gidermek üzere

literatürde çeşitli çözümler sunulmuştur; Wu'nun (Wu ve diğ., 2018), Kalantari'nin (Kalantari ve Ramamoorthi, 2017), Q.Yan'ın çalışmaları (Yan ve diğ., 2019) ile HDR-GAN (Niu ve diğ., 2021) bu yaklaşımlara örnek olarak verilebilir.

Wang tarafından (Wang ve diğ., 2021), düşük ve yüksek pozlanmış görüntü çiftlerinden orta pozlama görüntülerini öğrenmek için iki ağ yapısının eğitilmesi önerilmiştir. Ağın eğitilmesinde kullanılan toplam kayıp fonksiyonu; iki bağımsız ağ tarafından da öğrenilen aynı pozlama değerine sahip LDR görüntüler arasında “LDR-to-LDR loss” olarak tanımlanan  $L_1$  kayıp fonksiyonununun, farklı pozlama değerlerine sahip görüntüler arasında renk bilgisinin doğru şekilde öğrenilmesi için “HSV Loss” ve “RGB Loss” olarak tanımlanan  $L_1$  kayıp fonksiyonlarının ve son olarak yapısal bütünlüğü sağlamak için “MEF-SSIM Loss” olarak tanımlanan kayıp fonksiyonunun toplamı olarak oluşturulmuştur. Daha sonra ise Mertens'in (Mertens ve diğ., 2007) önerdiği pozlama değerlerine ihtiyaç duymayan yöntemi ile öğrenilen orta pozlama değerlerine sahip LDR görüntülerden HDR görüntü oluşturulmuştur. Model herhangi bir hedef HDR görüntüsüne ihtiyaç duymadığından, Wang tarafından önerilen bu yöntem, denetimsiz bir çözüm olarak anılmıştır.

### **2.3.2. Tek Bir LDR Görüntüsünden HDR Görüntü Oluşturma**

Tek bir LDR görüntüsünden HDR görüntü oluşturma işlemi, düşük ve yüksek pozlama bölgelerindeki detayların tek bir görüntüden ortaya çıkarılması ve aynı zamanda yapısal ve renk bütünlüğünün korunması gerekliliği nedeniyle çözülmesi daha zor bir problemdir. DrTMO modelinde (Endo ve diğ., 2017) beş farklı CRF'ye sahip kameralar simüle edilerek, veritabanındaki HDR görüntülerinden farklı pozlamalara sahip LDR görüntüleri öğrenilir ve daha sonra HDR görüntüler, (Mertens ve diğ., 2007) tarafından önerilen pozlama birleştirme kullanılarak birleştirilir. HDRCNN (Eilertsen ve diğ., 2017) çalışmasında ise, kamera sensörünün saturasyona uğraması nedeniyle LDR görüntü içerisinde aşırı parlak bölgelerdeki detayların autocoder ağ mimarisi ile kurtarılması hedeflenmiştir. ExpandNet'te (Marnerides ve diğ., 2018), sırasıyla yüksek frekanslı, piksel bazında komşuluk ilişkilerinde ve bütünsel anlamda ayrıntıları çıkarmak için “Local Branch”, “Expansion Branch” ve “Global Branch” adı verilen üç katmanlı CNN tabanlı bir model önerilmiştir. S.Lee (Lee ve diğ., 2018) ve S.Y.Kim

(Kim ve diğ., 2020) tarafından ise HDR görüntüleri oluşturmak için GAN mimarisi tekniklerinin kullanılması önerilmiştir.

Y.Liu (Liu ve diğ., 2020) probleme başka bir şekilde yaklaşarak, HDR görüntüden LDR görüntüsünü elde etmek için kullanılan ISP veriyolundaki üç bloğun fonksiyonel terslerinin, oluşturulan ayrı ağ mimarileri ile öğrenilmesini önermiştir. Çalışmanın ana fikri, LDR görüntülerinin elde edilmesinde kullanılan modullerin (dinamik aralık kırpma, CRF ile doğrusal olmayan dönüşüm ve kuantizasyon) neden olduğu etkileri “De-quantization”, “Refinement”, “Enhancement” and “Hallucination” olarak adlandırılan alt ağlar ile ortadan kaldırmaktır. Bu çalışmayı, Channel Adaptive Residual Dense Network mimarisinin önerildiği TCRDN çalışması (G.Chen ve diğ., 2021) takip etmiştir.

HDRUNet çalışmasında (Chen ve diğ., 2021), uzamsal olarak dinamik bir encoder-decoder ağ yapısı aracılığıyla aynı anda gürültü giderme ve dequantization işlevleri dikkate alınarak hedef HDR görüntüler oluşturulmuştur. Öte yandan Lee (Lee ve Sunwoo, 2021), giriş LDR görüntüsünün dikkatlice tasarlanmış bir maske ile parlak ve karanlık bölgelere bölünmesini önermiştir. Böylece hem parlak hem de karanlık bölgeler “Dynamic Range Extension Stage” modulünde ayrı ayrı eğitilerek elde edilen sonuçlar, giriş LDR görüntüsünden çıkartılan genel öznitelikler ile “Blending and Generation Stage” modulünde birleştirilerek son HDR görüntü oluşturulmuştur.

Şu ana kadar bahsi geçen çalışmalarda RGB renk uzayında HDR görüntülerin oluşturulması hedeflenirken, renk bilgisinin saturasyon bölgelerinde daha doğru oluşturulabilmesi maksadıyla HSVNet (Lee ve diğ., 2022) çalışmasında, HSV kanalında yer alan “S” ve “V” kanallarından bir maske türetilmiş ve HSV renk uzayında uygun bir kayıp fonksiyonundan yararlanılmıştır.

Huo (Huo ve Yu, 2023), doğru pozlama ile elde edilen detayları koruyan “Weighting” katmanının yanısıra, farklı skalalarda yerel ve global özniteliklerin birleştirildiği “Global Feature Fusion Module (GFFM)” modülünün U-Net tabanlı model mimarisinde kullanılmasını önermiştir. Guo (Guo ve Lin, 2023) tarafından GAN yapısının da kullanıldığı iki aşamalı yöntem sunulmuştur; ilk aşamada farklı pozlama değerlerine sahip LDR görüntülerin düşük ve yüksek pozlanmış bölgelerindeki detay bilgileri dikkat

mekanizmalarının da yer aldığı GAN yapısı ile öğrenilirken, ikinci aşamada HDR görüntüyü oluşturmak farklı pozlama değerlerine sahip görüntülerin çok katmanlı CNN blokları aracılığıyla birleştirildiği bir ağ yapısı geliştirilmiştir. Dalal (Dalal ve diğ., 2023) problemi bir görüntüden görüntüye çeviri problemi (image-to-image translation) olarak ele almıştır ve koşullu DDPM (Denoising Diffusion Probabilistic Model) temelli bir ağ mimarisi kullanmıştır. Bu çalışmada CNN temelli autoencoder yapısı kullanılarak gizli (latent) gösterimin kalitesi artırılmaya çalışılmış ve LDR-HDR görüntüler arasında saturasyon bölgelerindeki detayların ortaya çıkarılmasını amaçlayan yeni bir kayıp fonksiyonunun kullanılması önerilmiştir.

#### **2.4. Ton Dönüştürme Operatörü (Tone Mapping Operator (TMO))**

HDR görüntü elde etmedeki tüm gelişmelere rağmen, HDR monitörlerin yaygın olmaması nedeniyle oluşturulan HDR görüntülerin LDR monitörlerde görüntülenmesini sağlamak üzere ton dönüşümü işlemi uygulanmasına ihtiyaç duyulmaktadır.

Geleneksel kamera ISP veri yollarında da kullanılan ton dönüştürme operatörü görüntü işleme alanındaki popüler araştırma konularından biridir. Şu ana kadar çok sayıda yöntem önerilmiştir, ancak literatürde bu yaklaşımları temel olarak global ve yerel TMO operatörleri olarak kategorize etmek mümkündür. Pattanaik'in (Pattanaik ve diğ., 2000) ve Jung'un (Jung ve Wang, 2015) önermiş olduğu çalışmalardaki gibi genel TMO'lar, dönüşümü yapılacak görüntüdeki renk ve parlaklığı İGS ihtiyacına uyumlu olarak değiştirerek LDR monitörde görüntülenmesini sağlamak amacıyla tüm görüntünün bütününe uygulanacak optimum monotonik dönüştürme eğrişini elde etmeye amaçlamaktadır. Öte yandan Yerel TMO'lar Reinhard (Reinhard ve diğ., 2002) ve Ledda (Ledda ve diğ., 2004) tarafından önerildiği üzere bölgesel ton dönüştürme operatörünü uygulamak için piksel bazında komşuluk bilgisini kullanmaya çalışmaktadır.

Giriş görüntüsünün önce uzamsal, Yee tarafından önerildiği gibi (Yee ve Pattanaik, 2003), veya frekans bazında, Liu (Liu ve diğ., 2018) önerildiği gibi, alt bölgelere bölündüğü ve daha sonra bölgeleri birleştirmeden önce TMO'ların dikkatlice uygulandığı segmentasyona dayalı TMO'lar da mevcuttur. Durand (Durand ve Dorsey, 2002) temel ve detay katmanları olarak adlandırılan iki ölçekli ayrıştırmayı önererek, HDR

görüntüsünün ayrıntılarını korurken kontrastı azaltmayı başarmıştır. Wu (Wu ve diğ., 2019) ise füzyon adımından önce kontrast ve parlaklık katmanı olarak segmente edilmiş veriye iki TMO operatörü uygulamıştır.

CNN bloklarının kullanıldığı ve DFC-DIT (The Deep Feature Consistent Deep Image Transformation) olarak adlandırılan bir başka öncü çalışma ise Hou (Hou ve diğ., 2018) tarafından önerilmiştir. DeepTMO'da (Rana ve diğ., 2019) ise, hızlı TMO uygulanmasını sağlamak maksadıyla için koşullu GAN (cGAN) ağ mimarisinin kullanılması önerilmiştir. Öte yandan Panetta (Panetta ve diğ., 2021) tarafından dikkat mekanizmalarının kullanıldığı derin öğrenme mimarisi önerilmiştir.

Bu tez çerçevesinde önerilen model, doğrudan kamera sensöründen alınan verilerden yüksek kaliteli ton dönüşümü yapılmış HDR görüntülerin üretilmesi problemine ilişkin çözümü ele almaktadır. Yani bu makalede sunulan yöntemle ISP kamera veri yolu, görüntü kalitesi açısından daha iyi çıktılar sağlayacak şekilde modellenmekte, düşük ve yüksek pozlama bölgelerindeki detaylar ortaya çıkarılarak daha geniş dinamik aralığa sahip görüntü çıktıları oluşturulmaktadır.

### 3. GÖRÜNTÜ KALİTESİ DEĞERLENDİRME METRİKLERİ

Bu bölümde tez kapsamında geliştirilen modelin doğrulanması için kullanılan görüntü kalitesi değerlendirme metrikleri tanıtılacak ve söz konusu metriklerin seçilme nedeni açıklanacaktır.

Literatürde, çeşitli amaçlar doğrultusunda görüntüyü oluşturan ya da düzenleyen derin öğrenme mimarilerinin başarımları, çoğunlukla oluşturulan ve hedef görüntüler arasında tanımlanan çeşitli referansa ihtiyaç duyan görüntü değerlendirme (FR-IQA) metrikleri ile ölçülmüştür. Her ne kadar hedef görüntülerden çıkartılan parametre seti ile kalite değerlendirmesi yapan azaltılmış referansa ihtiyaç duyan (RR-IQA) modeller mevcut olsa da, günümüzde çoğu uygulamada hedef görüntünün elde edilmesindeki güçlükler nedeniyle, MOS puanı ile yüksek korelasyona sahip referansa ihtiyaç duymayan görüntü kalitesi değerlendirme metriklerinin (NR-IQA) kullanılması kaçınılmaz hale gelmiştir.

Aşağıda başlıca kullanılmakta olan FR-IQA ve NR-IQA metrikleri açıklanmıştır.

#### 3.1. Referansa İhtiyaç Duyan Görüntü Kalitesi Değerlendirme Metrikleri (FR-IQA)

Bu tür metrikler genelde Varga'nın (D.Varga, 2022) tanımladığı üzere beş kategoride incelenebilir: (i) hata görünürlük, (ii) yapısal benzerlik, (iii) bilgi-teorik, (iv) öğrenmeye dayalı ve (v) füzyon yöntemler.

Hata görünürlük metriklerinde, görünürlük üzerine sunulan yöntemlerde değerlendirmesi yapılan görüntünün ve hedef görüntünün piksel değerleri üzerinden bir mesafe fonksiyonu tanımlanır. Böylelikle algılayıcılar tarafından farklılığın ölçülmesi ile oluşturulan görüntünün kalitesi tespit edilmeye çalışılır. Yapısal benzerlik temelli yaklaşımda ise değerlendirilen görüntünün referans görüntüye olan benzirliğinin ölçülmesi hedeflenmektedir. Referans görüntü ile değerlendirilen görüntünün paylaşmış olduğu bilginin ölçülerek kalite skorunun belirlendiği yöntemler ise bilgi-teorik temelli yaklaşımlar olarak; makine öğrenmesi ya da derin ağ mimarileri yardımıyla görüntü kalitesinin belirlendiği çalışmalar ise öğrenmeye dayalı metotlar olarak anılmaktadır.

Son olarak bu yöntemlerin birlikte kullanıldığı çalışmalar ise füzyon yöntemler olarak sınıflandırılmaktadır.

Referansa ihtiyaç duyan görüntü kalite metriklerinden en bilinen olanları; Ortalama Karesel Hata (Mean Square Error, MSE), PSNR (Peak Signal to Noise Ratio), SSIM (Structural Similarity Index), VIF (Visual Information Fidelity) ve DIQam-FR (Deep Image Quality Measure – Full Reference) 'dir. Bahse konu metrikler aşağıda özetlenmiştir.

### 3.1.1. Ortalama Karesel Hata (MSE)

MSE Denklem (3.1)'de belirtilen yöntem ile hesaplanır; denklemde  $N$  toplam piksel sayısını,  $x_i$ , oluşturulan görüntüye ait  $i$ 'nci piksel değerini ve  $y_i$ , hedef görüntüdeki  $i$ 'nci piksel değerini ifade etmektedir. MSE değerinin düşük olması, hedef görüntü ile karşılaştırıldığında oluşturulan görüntüde daha az hata olması anlamına gelmektedir. Fakat sonuçların insan algıyla her zaman korelasyon içinde olduğu söylenemez. Görüntüdeki piksel farklılıkları üzerinden tanımlanması nedeniyle, referans bir resme göre aynı hata değerine sahip iki görüntü İGS için tamamen farklı tip bozulmalara sahip olabilir.

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2 \quad (3.1)$$

### 3.1.2. PSNR (Peak Signal to Noise Ratio)

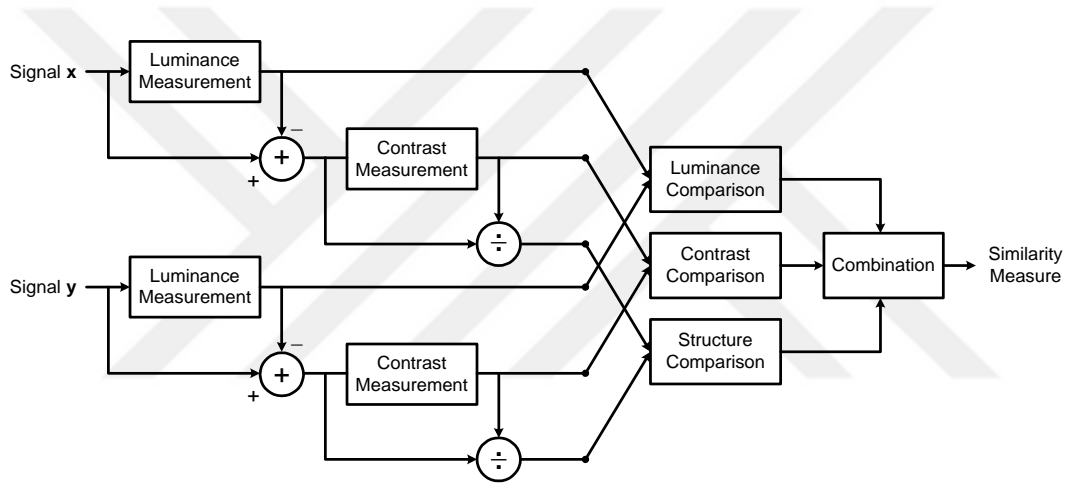
Sinyal Gürültü Oranı (Signal to Noise Ratio, SNR) Denklem (3.2)'de olduğu gibi bir sinyal gücünün,  $P_{\text{signal}}$ , sinyal içerisinde var olan gürültüye ait güç değerine,  $P_{\text{noise}}$ , oranıyla “dB” skalasında ifade edilmektedir. PSNR ise bir görüntünün referans görüntü ile karşılaştırılarak kalite değerlendirmesini yapmak amacıyla kullanılmaktadır. PSNR, mümkün olan en yüksek piksel değerinin karesinin, iki görüntü arasındaki,  $X$  ve  $Y$  ortalama karesel hataya oranıyla “dB” skalasında, Denklem (3.3)'de belirtildiği üzere formüle edilmektedir. PSNR metriğinin de insan algıyla her zaman korelasyon içinde olduğu söylenemez ve piksel değerlerindeki değişikliklerden İGS algısından bağımsız olarak doğrudan etkilenir.

$$\text{SNR} = 10 \log_{10} (P_{\text{signal}}/P_{\text{noise}}) \quad (3.2)$$

$$\text{PSNR} = 10 \log_{10} (\max_x^2 / \text{MSE}(X, Y)) \quad (3.3)$$

### 3.1.3. SSIM (Structural Similarity Index)

SSIM metriği, UIQI metriğinin (Wang ve Bovik, 2002) daha kararlı bir sonuç vermesi üzerine önerilen yapısal benzerlik temelli FR-IQA metriğidir. Şekil 3.1'de görüldüğü üzere hesaplama üç farklı karşılaştırma bloğundan oluşur; parlaklık, kontrast ve yapı karşılaştırması (Wang ve diğ., 2004). SSIM indeksi Denklem (3.4) ile hesaplanır. Denklemdaki  $x$  değerlendirilen görüntüyü,  $y$  ise referans görüntüyü belirtmektedir.



Şekil 3.1. SSIM indeksi ölçüm diyagramı (Z.Wang ve diğ., 2004)

$$\text{SSIM}(x, y) = \frac{(2 \mu_x \mu_y + C_1)(2 \sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3.4)$$

$\mu$  ortalama piksel değerini,  $\sigma$  tahmin edilen sinyal kontrastının standart sapmasını,  $\sigma_{xy}$  korelasyon katsayısını,  $C_1$  ve  $C_2$  ise kararlılık ve verimli bir optimizasyon için tanımlanan sabit parametrelerdir.

Tek bir ölçekte SSIM indeksi hesaplamasının neden olduğu belirsizlikleri ortadan kaldırmak üzere MS-SSIM (Wang ve diğ., 2003) metriği de önerilmiştir. Söz konusu MS-SSIM metriği, referans ve değerlendirilmesi yapılan görüntünün alt örneklenmiş

versiyonlarında yapı ve kontrast karşılaştırmalarının tamamlanarak final skoru için birleştirilmesi olarak özetlenebilir.

Daha sonra ise referans görüntü ile olan benzerliği İGS ile daha uyumlu ölçmek üzere Faz Uyumluluğu (Phase Congruency) ve Gradyan Büyüklüğü (Gradient Magnitude) kriterleri de eklenerek FSIM metriği (Zhang ve diğ., 2011) önerilmiştir.

Yapısal benzerliği dikkate alarak türetilen kalite değerlendirme metrikleri, MSE ve PSNR'a kıyasla daha etkin ve ISG ile kısmen uyumlu değerlendirme yapmaya imkân sağlamaktadır.

#### **3.1.4. VIF (Visual Information Fidelity)**

VIF (Sheikh ve Bovik, 2006) metriği IGS algısı ile uyumlu olarak referans görüntü ile karşılaştırma yapabilecek bilgi-teorik bir çözüm olarak sunulmuştur. Sunulan metrik ile görüntünün kalitesi, referans görüntüler arasında paylaşılan bilginin distorsiyon modelinden ve doğal sahne istatistiklerinden (Natural Scene Statistics (NSS)) faydalanılarak ölçülmesi ile ilişkilendirilir. Distorsiyon modeli parçacık etki alanında sinyal zayıflaması ve ilave gürültü olarak modellenirken, HVS kaynaklı her iki görüntünün algılanmasındaki belirsizlik beyaz gauss gürültüsü olarak modellenmiştir.

#### **3.1.5. DIQam-FR (Deep Image Quality Measure-Full Reference)**

Bosse (Bosse ve diğ., 2018) tarafından önerilen DIQam-FR metriği derin öğrenme temelli referansa ihtiyaç duyan bir değerlendirme metriğidir. Model, Siamese alt-ağ yapısı olarak adlandırılan toplam 10 adet konvolüsyonel bloktan, 5 adet örnekleme (pooling) katmanından ve 2 adet tam bağlantılı (fully connected) katmanından oluşan bir ağın eğitilmesiyle oluşturulmuştur. Referans görüntü ve kalite değeri bilinen bozulmuş görüntü çiftlerinden patch'ler halinde çıkartılan görüntüler oluşturulan ağ yapısının eğitiminde kullanılmıştır. Test aşamasında ise bütün görüntünün kalite değeri, çıkartılan patchlerin ortalama skoru olarak sunulmuştur. LIVE, TID2013 ve CISQ verisetleri ile geliştirilen metriğin etkinliği değerlendirilmiştir. Bosse tarafından sunulan bu çalışmada ayrıca, Siamese alt-ağ mimarisi modifiye edilerek referansa ihtiyaç duymayan görüntü kalite metriği, DIQam-NR, geliştirilmiş ve model başarımı LIVE In the Wild Image Quality Challenge Database (CLIVE) ile değerlendirilmiştir.

### **3.2. Referansa İhtiyaç Duymayan Görüntü Kalitesi Değerlendirme Metrikleri (NR-IQA)**

Görüntü kalitesinin referansa ihtiyaç duymadan belirlenebilmesi çeşitli bilgisayarlı görü uygulamaları için önemli bir amaçtır. Gerçek zamanlı uygulamalarda referans görüntünün kullanılması çok mümkün olmamaktadır. Bu nedenle bir görüntünün kalitesinin referansa ihtiyaç duyulmadan belirlenmesi vazgeçilemeyecek bir ihtiyaçtır. Nihai amacın görüntü kalitesinin İGS algısına uyumlu olarak değerlendirilmesi olduğu göz önüne alındığında, herhangi bir referans görüntü olmadan bu amaca ulaşmanın zorluğu daha kolay anlaşılır.

NR-IQA metriklerinde Doğal Sahne İstatistikleri (NSS - Natural Scene Statistics) ve İGS algısı temelli metotlar kullanılmaktadır. NSS, uzamsal düzlemde normalleştirilmiş parlaklık katsayılarına dayanır ve çok boyutlu Gauss dağılımı ile modellenir. Görüntüdeki bozulmalar, Gauss dağılımında da bozulmalar olarak tespit edilir; kullanılan çeşitli algoritmalar, bir modeli eğitmek ve kalite puanını hesaplamak için NSS özniteliklerini nasıl kullandıklarına göre farklılık gösterir. İGS temelli yöntemlerde ise insan algısına uyumlu özniteliklerin görüntüden çıkartılmasıyla kalite skoru belirlenir.

Derin öğrenme mimarilerinin kullanıldığı yöntemlerde model kadar kullanılan verisetinin yapısı da metrik performansını doğrudan etkilemektedir. Sentetik bozulmalar ile hazırlanmış verisetlerinde sınırlı bozulma çeşitleri bulunmasının yanısıra, bozulmaların eşzamanlı meydana gelmesi durumu da çoğunlukla kapsamamaktadır. Bu nedenle gerçek dünya sahneleri ile oluşturulan verisetleri problemin çözümünde daha faydalı olabilmektedir.

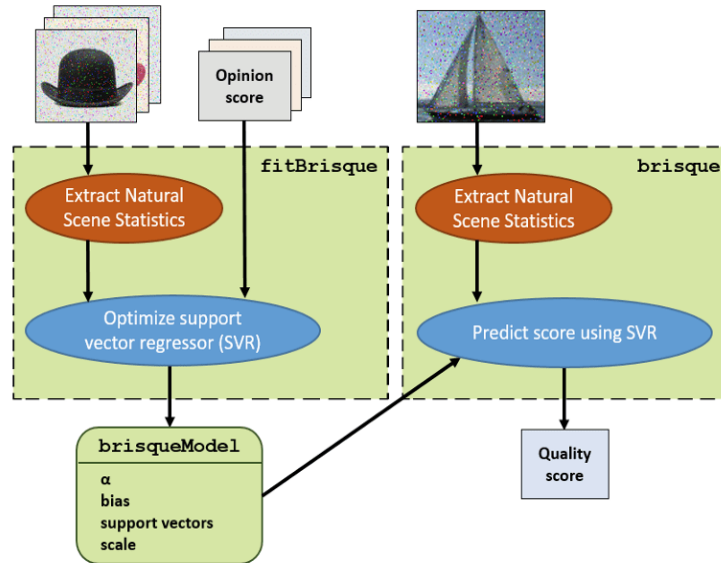
Aşağıda referansa ihtiyaç duymayan ve literatürde sıklıkla kullanılan önemli görüntü kalite metrikleri kısaca tanıtılmıştır.

#### **3.2.1. Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE)**

BRISQUE metriği (Mittal ve diğ., 2012), NSS özniteliklerini kullanarak kalite skorunu tahmin eder. Model, aynı tür bozulmaya sahip görüntülerin kalitesini ölçmekle sınırlıdır. BRISQUE modeli, verisetinde yer alan görüntülerin insan tarafından

değerlendirilmesiyle elde edilen öznel görüş puanları kullanılarak eğitilir; bu nedenle BRISQUE metriği, insanın kalite algısıyla iyi bir şekilde ilişkilendirilmesi avantajına sahiptir.

Bir BRISQUE modelini eğitmek için; bilinen bozulmalara sahip görüntülerini ve bu görüntülerin bozulmamış kopyalarını içeren bir veri setine ihtiyaç duyulmaktadır. BRISQUE modelinin akış diyagramı Şekil 3.2’de sunulmuştur. Veri setinde yer alan görüntülerin, bloklara bölünmeden, NSS öznitelikleri çıkarılır. Sonrasında destek vektör makinesi regresyon (SVR) modelini eğitmek için veri setindeki görüntülerden çıkarılan NSS özellikleri ve her bozuk görüntü için öznel görüş puanı kullanılır. Sonrasında BRISQUE puanını hesaplamak için eğitimi yapılan SVR model parametreleri saklanır. Kalite skorunun tahmini ise, giriş görüntüsünden çıkarılan NSS öznitelikleri ve eğitimi yapılan SVR modeli kullanılarak gerçekleştirilir. BRISQUE skoru ne kadar küçükse görüntü o kadar kalitelidir.

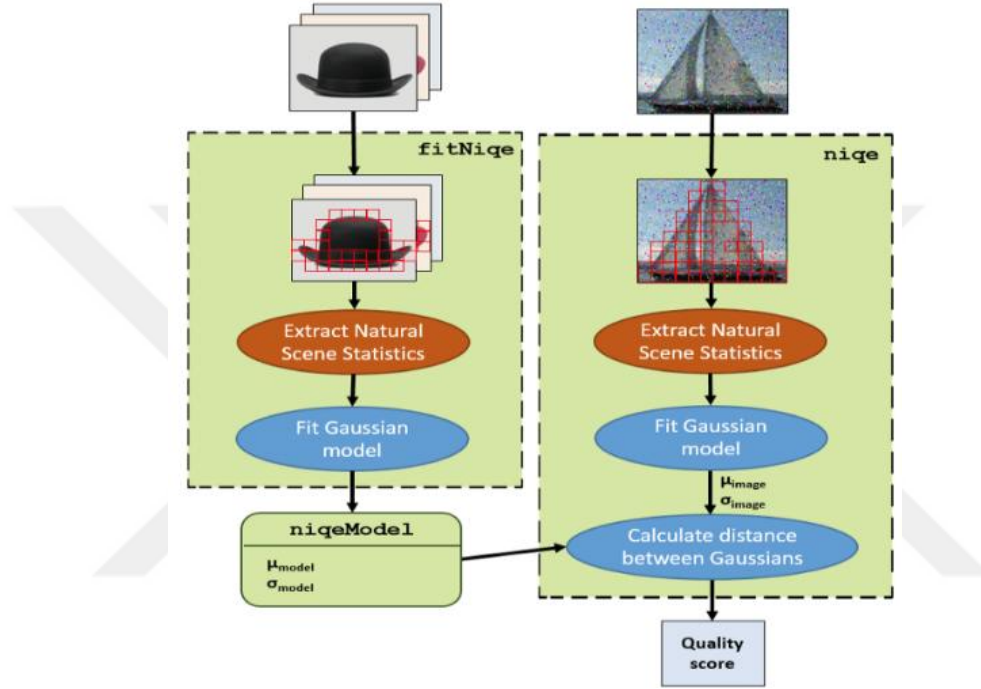


Şekil 3.2. BRISQUE metriği diyagramı (URL-3)

### 3.2.2. The Natural Image Quality Evaluator (NIQE)

NIQE ölçüğü (Mittal ve diğ., 2013), doğal sahnelere ait görüntülerden çıkarılan NSS özniteliklerini kullanır, bu nedenle herhangi bir bozulma tipinden bağımsız olarak skor üretir. Eğitim aşamasında her bir doğal görüntü istatistiksel olarak anlamlı özelliklere sahip bloklara bölünür ve belirlenen bloklar için NSS öznitelikleri hesaplanır. Sonrasında, Şekil 3.3.’de sol tarafta belirtildiği şekilde bu özniteliklerden türetilen çok

değişkenli Gauss dağılımının ortalaması ve standart sapması oluşturulan modelde depolanır. Görüntünün değerlendirilmesi aşamasında, eğitim aşamasında olduğu gibi bloklar belirlenerek bloklara ait NSS öznelikleri çıkartılır ve Şekil 3.3’de sağ tarafta belirtildiği şekilde çok değişkenli Gauss dağılımı belirlenir. NIQE skoru belirlenen gauss dağılımları üzerinden tanımlanan uzaklık fonksiyonu ile belirlenir. NIQE değeri sayısal olarak ne kadar küçükse görüntü o kadar kaliteli.



Şekil 3.3. NIQE metriği diyagramı (URL-3)

### 3.2.3. Perception based Image Quality Evaluator (PIQE)

PIQE metriği (Venkatanath ve diğ., 2015) öznel görüş puanları kullanılmadan ve herhangi bir eğitilmiş modele ihtiyaç duymadan görüntü kalitesini kullanıcıya sunar. İnsan davranışına paralel olarak görüntüdeki bazı bölgelerden yerel öznelikler çıkartılması prensibiyle, aşağıdaki işlem adımları takip edilerek nihai kalite skoru belirlenir. (URL-4)

- Görüntüdeki her piksel için Ortalama Çıkarılmış Kontrast Normalleştirilme (MSCN-Mean Subtracted Contrast Normalized) katsayısı hesaplanır,
- Giriş görüntüsünü 16x16 boyutunda örtüşmeyen bloklara bölünür,

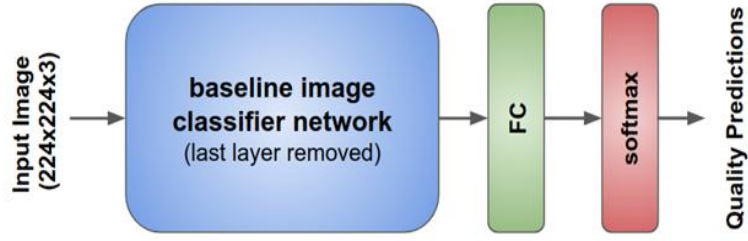
- MSCN katsayılarının varyansına bağılı olarak uzamsal olarak aktif bloklar tespit edilir,
- Tespit edilen bloklarda, MSCN katsayıları kullanılarak yapay menfi etkiler ve gürültüden kaynaklanan bozulmalar değerlendirilir,
- Eşik değeri kullanılarak bloklar, Gauss gürültüsü içeren bozulmuş bloklar ve bozulmamış bloklar olarak sınıflandırılır,
- Giriş görüntüsü için PIQE puanı, bozuk bloklardaki puanların ortalaması olarak hesaplanır.

PIQE değeri sayısal olarak ne küçük olması, görüntünün daha kaliteli olduğu anlamına gelmektedir.

### **3.2.4. Neural Image Assessment (NIMA)**

NIMA metriği (Talebi ve Milanfar, 2018) 0-10 arasında değerler ile görüntünün kalite skorunu referansa ihtiyaç duymadan sunmaktadır. Nümerik olarak yüksek NIMA skoruna sahip olan görüntü daha iyi kaliteye sahip olduğu anlamına gelmektedir. NIMA metriği Şekil 3.4.'te yer alan diyagramda sunulduğu üzere, Öğrenme Aktarımı (Transfer Learning) konsepti ile ImageNet, ResNet, MobileNet gibi bir "base" network sonuna 10 sınıflı Softmax aktivasyon fonksiyonuna sahip tam bağlantılı (fully connected) çoklu sınıflandırıcı eklenerek oluşturulan ağı eğitilmesiyle elde edilmektedir. NIMA metriğinin benzer çalışmalardan en önemli farkı; ortalama skorun yanında insan algısına uyumlu olarak skorun olasılıksal dağılımının tahmin edilmesidir. Eğitim esnasında kayıp fonksiyonu olarak EMD (Earth Moving Distance) kayıp fonksiyonu kullanılmaktadır.

Bu tezde kullanılan değerlendirme metriklerinden biri NIMA skorudur. Sonuçların elde edilmesinde kullanılan NIMA metriğine ait model, LIVEInTheWild veri seti kullanılarak ve MobileNet-v2 (Sandler ve diğ., 2018) temel ağ yapısı kullanılarak eğitilmiştir. LIVEInTheWild verisetinin içerisinde, 1162 adet mobil telefon ile elde edilmiş görüntüler ile bu görüntülerin ortalama 175 kişi tarafından değerlendirilmesi ile elde edilmiş öznel kalite skorları bulunmaktadır. Değerlendirmede kullanılan model, "mat" dosyası olarak MATLAB tarafından kullanıma sunulmuştur. (URL-5)

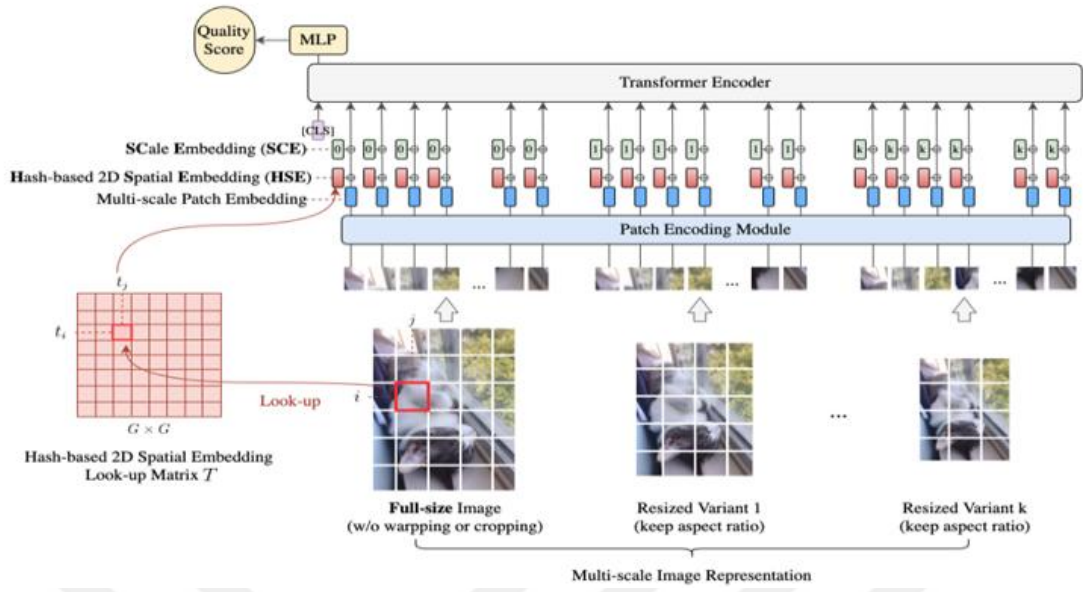


Şekil 3.4. NIMA metriği diyagramı (Talebi ve Milanfar, 2018)

### 3.2.5. Multi-scale Image Quality Transformer (MUSIQ)

MUSIQ (Ke ve diğ., 2021) skoru, literatürde yer alan CNN tabanlı referansa ihtiyaç duymayan kalite metriklerinin aynı ölçeklerde girişe ihtiyaç duyması ve bu nedenle skor tahmininden önce tekrar ölçeklendirilmesindeki (resize) bozucu etkilerin skora etki etmemesini sağlamak amacıyla önerilmiştir. Önerilen yöntemde kalite değerlendirmesi yapılan herhangi bir çözünürlüğe sahip giriş görüntüsü, önceden belirlenmiş en/boy oranları kullanarak farklı çözünürlükteki ölçeklendirilmiş versiyonları ile birlikte giriş görüntüsü olarak kullanılır. Daha sonra giriş görüntüleri Yama Kodlama Modülü (Patch Encoding Module) aracılığıyla sabit boyutlu yamalara (patches) bölünür. Görüntünün yapısal bütünlüğünü algılamak ve farklı en/boy oranlarına sahip görüntüleri işleyebilmek amacıyla; yamaların uzamsal konumları HSE (Hash Based 2D Spatial Embedding) tarafından, ölçek bilgisinin elde edilmesi için ise yamalar SCE (Scale Embedding) tarafından kodlanır. Bu kodlanan verilere öğrenilebilir CLS (Classification Token) durumu da eklenerek, orijinal giriş görüntüsünün Transformer Encoder çıkışında son temsili oluşturulur. Son olarak tam bağlantılı (fully connected) MLP (multi-layer perceptron) ağ yapısı ile görüntünün MUSIQ skoru tahmin edilir. Modele ait ağ mimarisi Şekil 3.5'te sunulmuştur. Giriş görüntüsünün MUSIQ skoru 0-100 aralığında tahmin edilmekte olup, sayısal olarak büyük olan skor daha iyi görüntü kalitesine sahip olduğu anlamına gelmektedir.

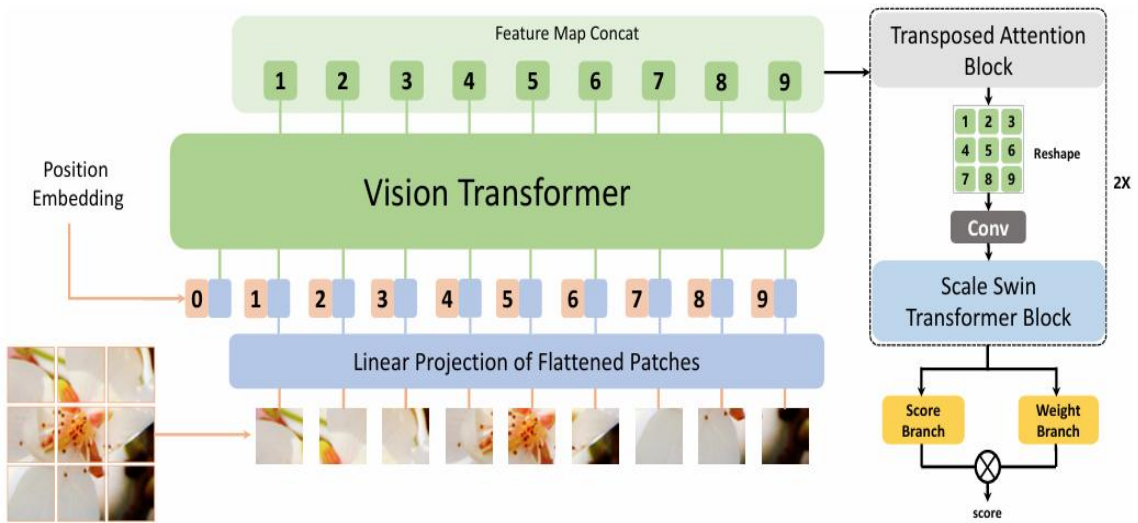
Metriğin deneysel sonuçları PaQ-2-PiQ (Ying ve diğ., 2020), SPAQ (Fang ve diğ.,2020), and KonIQ10k (Hosu ve diğ.,2020) verisetleri ile doğrulanmıştır. Bu tezde kullanılan kalite değerlendirme metriklerinden biri MUSIQ skoru olarak belirlenmiş olup, KonIQ10k veri seti ile eğitilmiş model parametreleri kullanılmıştır (URL-6).



Şekil 3.5. MUSIQ metriği diyagramı (Ke ve diğ., 2021)

### 3.2.6. Multi-dimension Attention Network for No-Reference Image Quality Assessment (MANIQA)

MANIQA metriği (Yang ve diğ., 2022), referans görüntüye ihtiyaç duymayan görüntü kalitesi değerlendirme metriklerine yönelik gerçekleştirilen NTIRE 2022 (Gu ve diğ., 2022) yarışmasında en iyi performansı sergileyen metrik olarak öne çıkmıştır. İnsan görüş algısıyla uyumlu kalite skorunu tahmin etmek amacıyla önerilen MANIQA metriğine ait model diyagramını Şekil 3.6’da sunulmuştur.



Şekil 3.6. MANIQA metriği diyagramı (Yang ve diğ., 2022)

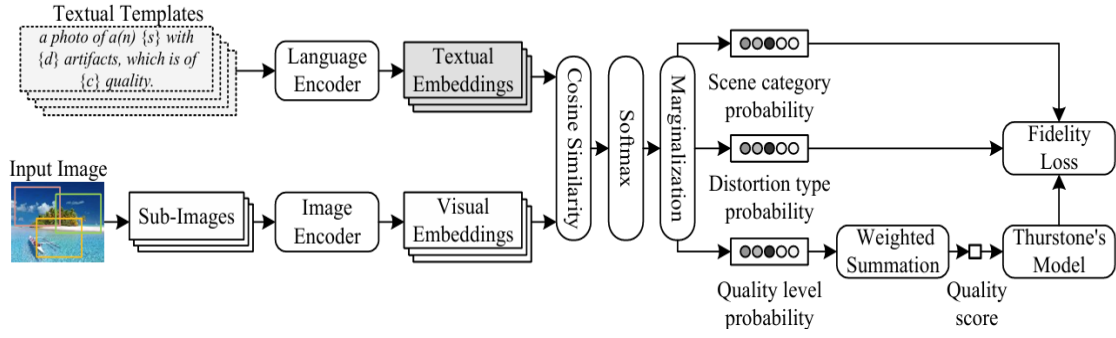
Kullanılan metriğin amacı kanal ve uzamsal boyutlarda yer alan bilginin kullanılmasıyla kalite skorunun üretilmesidir. Bu amaç doğrultusunda (8, 8)'lik görüntü yamaları çıkartılarak ViT aracılığıyla öznelik haritaları çıkartılır ve sonrasında TAB (Transposed Attention Block) ve SSTB (Scale Swin Transformer Block) aracılığı ile dikkat mekanizmalarının kullanılması sağlanır. Bu sayede genel ve yerel olarak farklı bölgelerdeki çıkartılan öznelikler arasındaki etkileşim sağlanmış olur. Nihai kalite skoru ise her görüntü yamasının puanı ve ağırlık katsayısı kullanılarak tahmin edilir.

MANIQA metriği, giriş görüntüsünün 0-1 aralığında kalite skorunu tahmin etmektedir. Bu tez kapsamında nicel değerlendirmeler yapmak amacıyla, URL-7 ile sunulan ve KADID-10k (Lin ve diğ., 2019) veriseti ile eğitilen model parametreleri kullanılmıştır.

### **3.2.7. Language-Image Quality Evaluator (LIQE)**

LIQE metriği (Zhang ve diğ., 2023), çok görevli öğrenme metodunu kullanarak, metinsel ve görsel olarak tasarlanmış kodlayıcıların (encoder) eş zamanlı eğitilmesiyle literatüre kazandırılmıştır. Sunulan çalışma ile görüntünün kalite skorunun tahminin yanısıra, 11 adet bozulma ve 9 adet sahne kategorisi içerisinde sınıflandırılmasını da gerçekleştirilmektedir. Belirlenen görevleri gerçekleştirebilmek amacıyla oluşturulan modelde metinsel ve giriş görüntüsüne ait parçalardan elde edilen indirgenmiş öznelik vektörleri, CLIP (Contrastive Language Image Pre-training) modeliyle elde edilir. Daha sonra kosinüs benzerliğiyle ortak olasılık dağılımları hesaplanır ve sonrasında her bir görev için marjinalleştirilmiş olasılık dağılımları hesaplanır. Şekil 3.7'de sunulan ağ mimarisi her bir görev için tanımlanmış kayıp fonksiyonu ve genel tanımlanmış bir kayıp fonksiyonu ile eğitilmiştir. Görüntünün kalite skoru hesaplanan olasılık dağılımının ağırlıksal toplama işlemi sonucunda elde edilir. LIQE skorunun sayısal olarak daha büyük olması, görüntü kalitesinin daha iyi olduğu anlamına gelmektedir. LIQE skorunun dinamik aralığı ise 0 - 5 arasındadır.

Bu tezde LIQE metriği de oluşturulan model mimarisinin başarımını değerlendirmek amacıyla kullanılmıştır. Kullanılan LIQE metriğine ait model, KonIQ10k veri seti ile eğitilmiştir. (URL-8).



Şekil 3.7. LIQE metriği diyagramı (Zhang ve diğ., 2023)

### 3.3. Kullanılan Kalite Metrikleri

Bu tez kapsamında oluşturulan modele ait çıktıların performansı, referansa ihtiyaç duymayan metriklerden MUSIQ, NIMA, LIQE ve MANIQA kullanılarak değerlendirilmiştir. Referansa ihtiyaç duymayan metriklerin kullanılmasının ilk nedeni, veri setinin oluşturulması aşamasında fotoğraf düzenleme yazılımları kullanılarak hedef görüntülerin oluşturulmasıdır. Bu aşamada standart parametreler yerine, görüntüye özgü düzenleme işlemi uygulanmıştır. Kullanılan veriseti, doğrudan herhangi bir kamera vasıtasıyla elde edilmemiş; öznel fotoğraf düzenlenmesi işlemleri gerektiren karmaşık bir süreç sonunda oluşturulabilmiştir.

Referansa ihtiyaç duyulmayan metriklerin kullanılmasının bir diğer nedeni ise, söz konusu metriklerin model başarımını, doğrudan İGS algısını ifade eden MOS ile uyumlu olarak ölçebilmesidir.

Referansa ihtiyaç duyan metriklerden MSE ve SSIM fonksiyonları eğitim esnasında kullanılmıştır. Özellikle eğitim aşamasında her bir katmandaki iterasyonun durdurulması esnasında MSE ve SSIM sonuçlarından faydalanılmıştır.

#### 4. VERİ SETİNİN HAZIRLANMASI

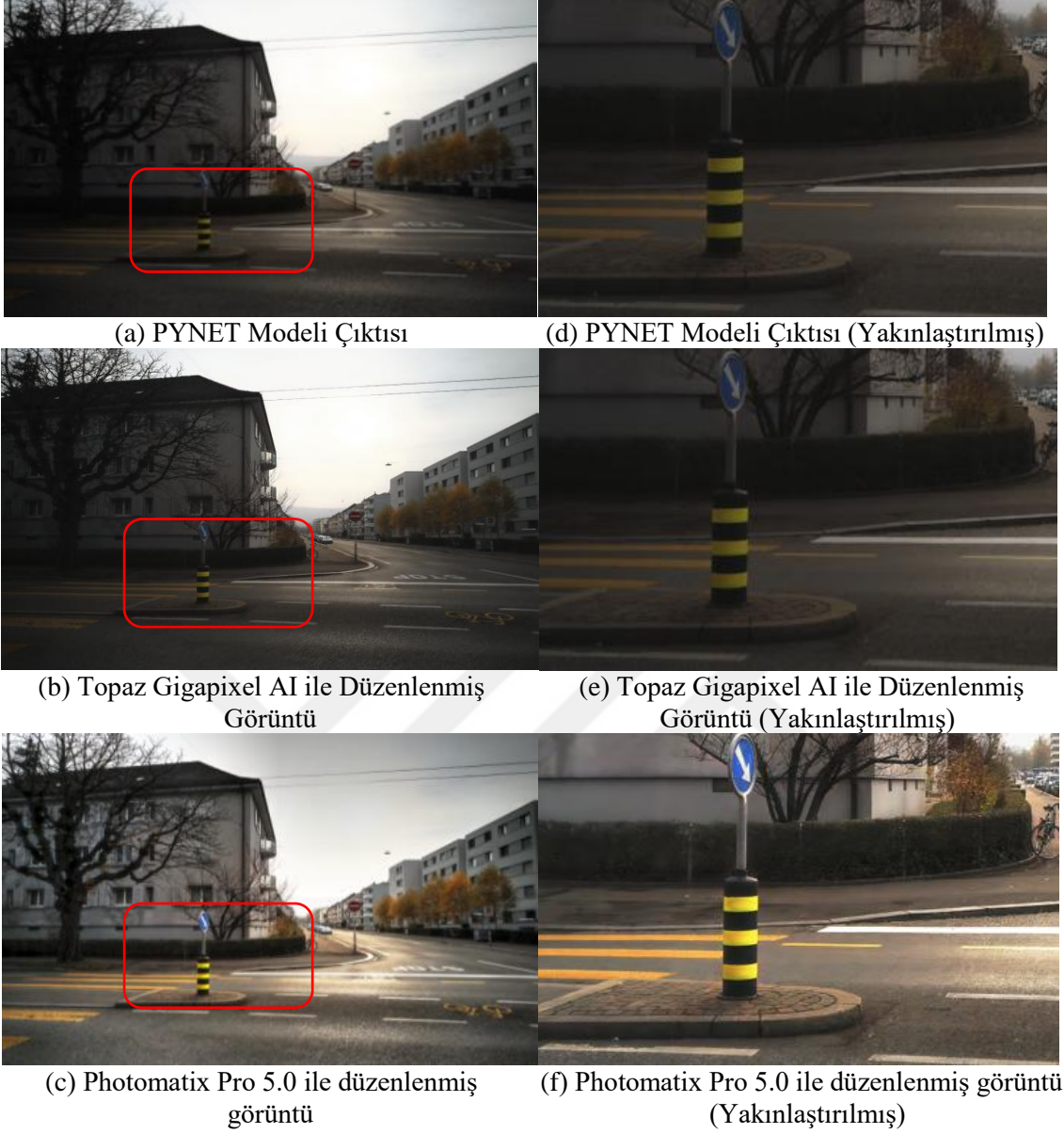
Bu tez kapsamında, kamera ISP veri yolunun derin ağ mimarisi ile modellenmesi amacıyla oluşturulmuş Zürich Veri Seti (Ignatov ve diğ., 2020)'nden faydalanılmıştır. Zürich Veriseti'nde yaklaşık 20 bin görüntü yer almakta olup, içerisinde;

- Sony 12 MP Exmor IMX380 Bayer paterne sahip sensörün kullanıldığı Huawei P20'den elde edilen RAW görüntüler,
- Alınan ham görüntülerin Huawei ISP veri yolu ile işlenmesi ile elde edilen RGB görüntüler,
- 24mm f/1.4L lense sahip Canon 5D Mark IV kamera ile elde edilen hedef RGB görüntüler, bulunmaktadır.

Bu tez kapsamında, Zürich veri setindeki 168 adet tam çözünürlüğe sahip (2976, 3968) RGB görüntüye karşılık gelen 10 bitlik RAW görüntülerin 8 bitlik PyNET model çıktıları ilave fotoğraf düzenleme işlemlerine tabi tutulmuş ve daha sonra eğitim aşaması için 12600 adet (448, 448) çözünürlükte görüntü parçaları çıkarılmıştır. Aynı kırpma işlemi RAW görüntülere de uygulanarak eğitim için gerekli görüntü çiftlerinden oluşan veriseti hazırlanmıştır. Eğitim amacıyla kullanılan görüntü yamalarının çözünürlüğüne yönelik etkinlik analizi Bölüm 6.3.3'te sunulmuştur.

Eğitim esnasında her bir eğitim tur sayısı bitiminde doğrulama amacıyla 1204 adet (448, 448) çözünürlüğe sahip RAW-RGB görüntü çifti kullanılmıştır. Modelin eğitilmesi hitamında, test amacıyla ise 10 adet tam çözünürlüğe sahip RAW görüntü kullanılmıştır.

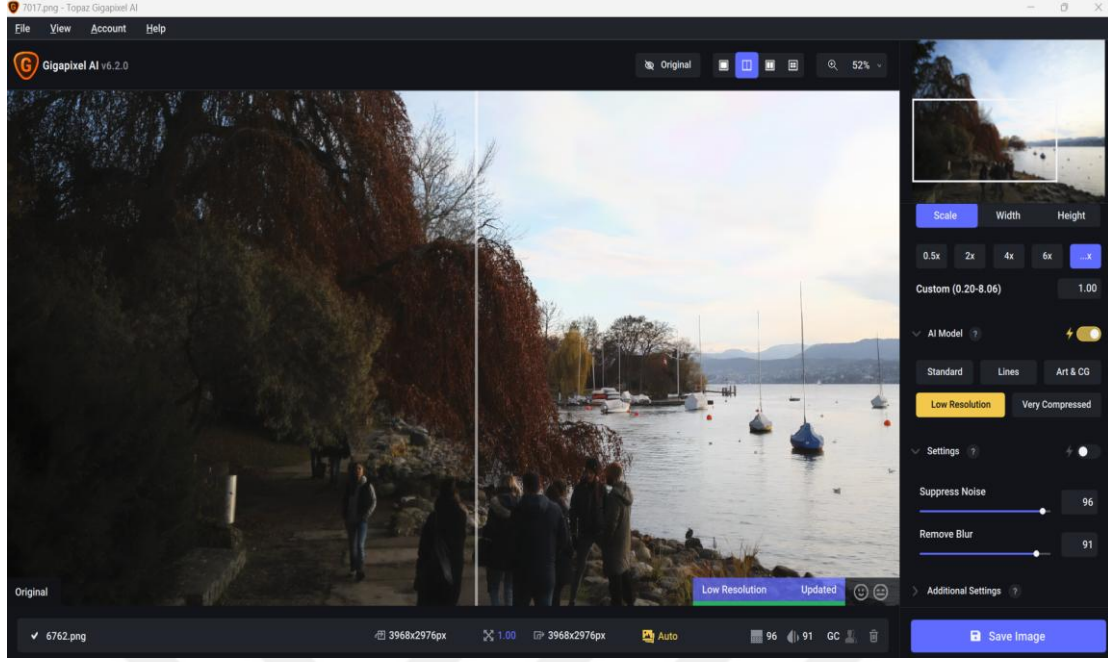
Verisetinin hazırlanması aşamasında çeşitli kullanıcı yorumları ve yaygınlığı göz önünde bulundurularak Topaz Gigapixel AI ve Photomatix Pro 5.0 yazılım paketleri sırasıyla hedef RGB görüntüleri oluşturmak amacıyla kullanılmıştır. Hedef görüntünün oluşturulması esnasında her bir aşamanın örnek çıktı görüntüleri Şekil 4.1'de sunulmuştur. Özellikle Topaz Gigapixel AI gürültü azaltımında, Photomatix Pro 5.0 ise düşük pozlama ve yüksek pozlama bölgelerindeki detayları ortaya çıkarmakta başarılı sonuçlar verdiği görülmektedir.



Şekil 4.1. Veri seti hazırlama aşamalarına ilişkin görseller

Düzenleme işlemi sırasında bazı örneklerde olumsuz bozucu etkilerin görülmesi nedeniyle Topaz Gigapixel AI ve Photomatix Pro 5.0 yazılım paketlerinin varsayılan (default) parametreleri kullanılmamıştır. Söz konusu fotoğraf düzenleme yazılımları ile düzenleme işlemi, düzenlenen görüntüye özel olarak seçilen parametrelerle gerçekleştirilmiştir. Topaz Gigapixel AI yazılımının kullanıcı arayüzü Şekil 4.2’de; Photomatix Pro 5.0 yazılımının kullanıcı arayüzü Şekil 4.3’de sunulmuştur.

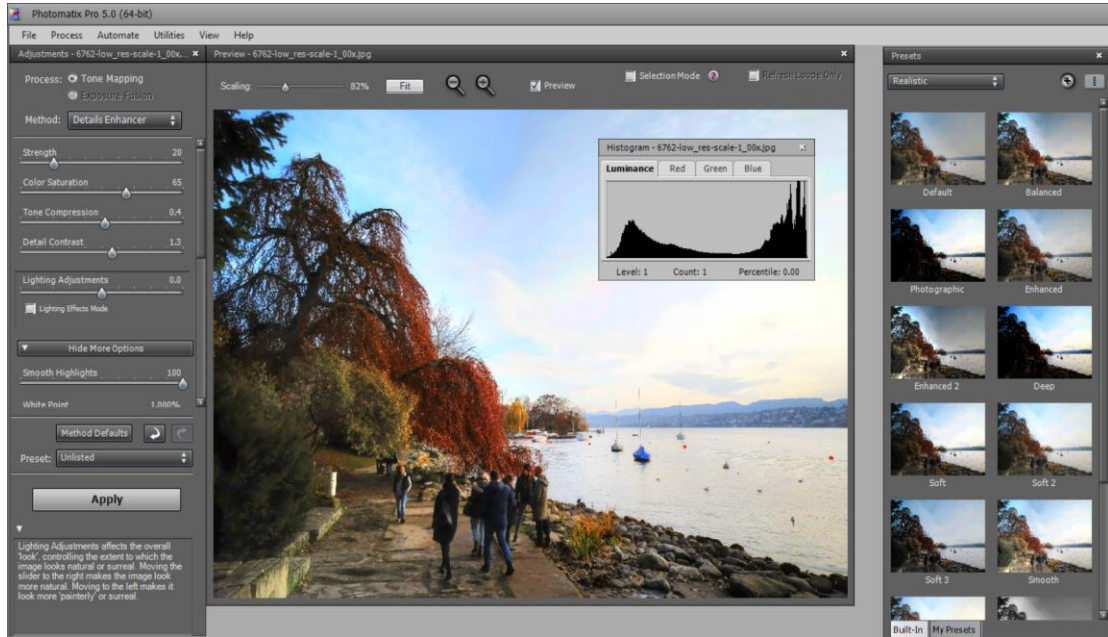
Topaz Gigapixel AI kullanıcı arayüzünden anlaşılacağı üzere görüntü önceden belirlenmiş birkaç varsayılan seçenek arasından seçim yapabilmektedir. Ayrıca görüntü



Şekil 4.2. Topaz Gigapixel AI fotoğraf düzenleme yazılımı kullanıcı arayüzü

giderimi, bulanık giderici dereceleri 0-100 skalasında seçilebilmektedir. İlaveten yüz detaylarını iyileştirici seviye ayarlamalı bir seçenek bulunmaktadır.

Photomatrix Pro 5.0 farklı pozlama değerlerinden HDR görüntü oluşturmak üzere yaygın olarak kullanılmaktadır. Bununla birlikte tek görüntüden ise tonemapping yapma imkanı vermektedir.



Şekil 4.3. Photomatrix Pro 5.0 fotoğraf düzenleme yazılımı kullanıcı arayüzü

Şekil 4.3.'te sunulan kullanım arayüzü sunulmuştur. “Presets” menüsünden varsayılan ayarlar kullanılabilen, aynı zamanda sol tarafta yer alan menüden manuel kullanıcı girdileri yapılabilmektedir. Photomatix Pro 5.0 düşük pozlanmış bölgelerdeki detayları ortaya çıkarmakta başarılı sonuçlar verdiği görülmektedir.

Tez kapsamında hedef görüntülerin İGS algısına uygun olarak düzenlenmesi aşamasında en optimum çözümün; görüntünün ilk önce Topaz Gigapixel AI ile, sonrasında ise Photomatix Pro 5.0 ile düzenlenmesi olduğu görülmüştür. Tez kapsamında hazırlanan verisetinde yer alan düzenleme maksadıyla kullanılan örnek bir görüntü ile bu görüntünün her iki yazılım ile düzenlendikten sonraki son hali Şekil 4.4.'te sunulmuştur.



Şekil 4.4. PYNET çıktısı (solda), Topaz Gigapixel AI ve Photomatix Pro 5.0 kullanılarak elde edilen görüntü (sağda)

Test amacıyla kullanılan 10 adet yüksek çözünürlüklü (2976, 3968) görüntü imge numaraları ile birlikte Şekil 4.5'te sunulmuştur. Bölüm 6'da yer alan deneysel sonuçlarda nicel ve nitel olarak yapılan karşılaştırmalar söz konusu imge numaraları ile birlikte paylaşılmıştır. Tez kapsamında kullanılan test görüntülerinin belirlenmesi esnasında çeşitli kriterler gözönünde bulundurulmuştur. Şekil 4.5'te yer alan test görüntüleri incelendiğinde; seçilen sahnelerin farklı ışıklılık seviyelerine sahip olduğu, sahnede yer alan renk dağılımının değişkenlik gösterdiği ve ışık kaynağının sahnedeki konumlarının da farklı olduğu görülebilmektedir.



(a) imge 0



(b) imge 1



(c) imge 2



(d) imge 3



(a) imge 4



(a) imge 5



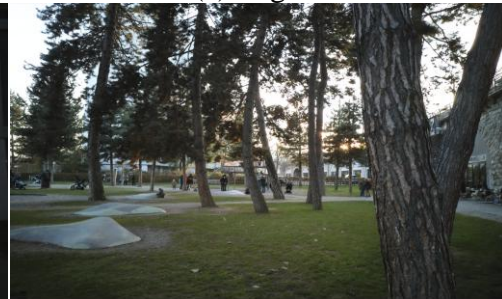
(a) imge 6



(a) imge 7



(a) imge 8



(a) imge 9

Şekil 4.5. Test amacıyla ayrılan görüntüler ve imge kimlikleri

## 5. ÖNERİLEN YÖNTEM

Bu bölümde önerilen modelin ağ mimarisinin yapısı, ağ mimarisi içerisinde yer alan dikkat modüllerinin yapısı, modelin eğitimi esnasında kullanılan eğitim parametreleri ve kayıp fonksiyonları tanıtılacaktır.

### 5.1. Model Mimarisi

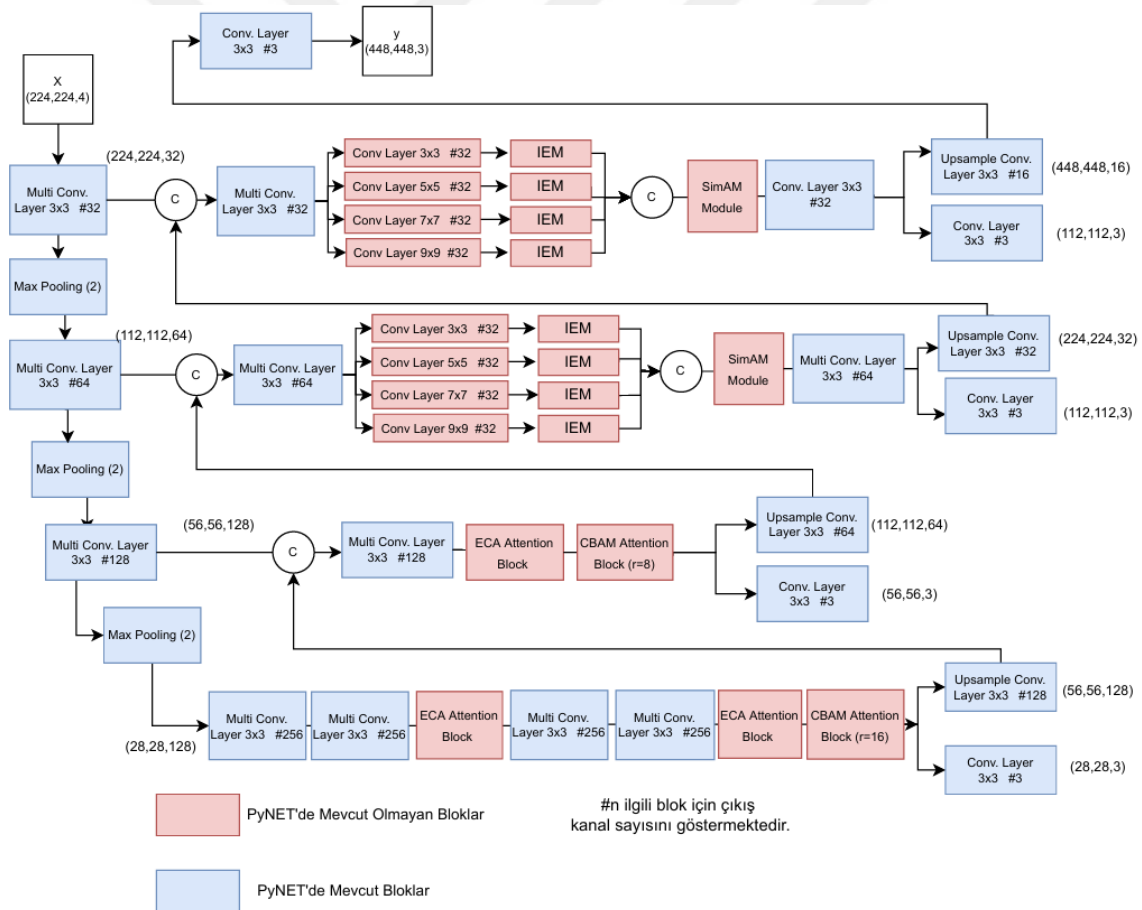
Önerilen model mimarisi Şekil 5.1'de gösterilmektedir. Ağın geliştirilmesi sırasında referans temel model olarak PyNET (Ignatov ve diğ., 2020) kullanılmıştır; dolayısıyla önerilen modelde de piramit yapısından faydalanılmıştır. Sunulan model beş katmandan oluşmaktadır ve her bir katman, Layer-4 olarak adlandırılan en alt katmandan başlayarak ayrı ayrı eğitilmektedir. Model mimarisi ve eğitimin gerçekleştirildiği kodlar URL-9'de paylaşılmıştır. Eğitim sırasında; eğitimin gerçekleştirildiği katman ve eğitim için gerekli diğer hiperparametreler (tur sayısı, adım boyu, paket büyüklüğü) “utils.py” ile tanımlanmaktadır. Layer-4 hariç diğer katmanlarda, kendinden önce eğitilen model parametreleri geri yüklenmektedir. Böylelikle tüm model parametrelerinin aynı anda eğitilmemesi sağlanmakta ve eğitim süreci daha verimli şekilde gerçekleştirilebilmektedir.

Şekil 5.1.'de sunulan model mimarisinde görülebileceği üzere, Max Pooling bloğu ile uzamsal öznitelik çözünürlüğü Layer-2, Layer-3 ve Layer-4 girişinde ilaveten iki çarpanıyla azaltılmaktadır. Layer-1 (upsample bloğu çıktısı) ve Layer-0'da eğitim, hedef görüntünün giriş uzamsal çözünürlüğünde (448, 448) gerçekleştirilmektedir.

Eğitim esnasında her katmanda çıkarılan öznitelik haritası büyütülür ve bir sonraki katmanın girişi için çıkartılan öznitelik haritası ile birleştirilir. Bu sayede düşük seviyeli özelliklerden başlayarak ağ yapısının eksik detayları öğrenmesi sağlanmış olur ve oluşturulan görüntünün iyileştirilmesi sağlanır. ReLU aktivasyon fonksiyonu, her seviyedeki son çıkış blokları hariç tüm evrişimli ve üst örneklemeli evrişimli bloklarda kullanılır. Her bir katmanda çıkış blokları için sigmoid aktivasyon fonksiyonundan yararlanılmaktadır. Bahse konu çıkış blokları her bir eğitim tur sayısı sonunda elde edilen çıktıların görselleştirilmesi ve eğitim durumunun takip edilebilmesi amacıyla kullanılmaktadır. PyNET'te olduğu önerilen modelde de evrişimli katmanlarda “instance

normalization” kullanılmıştır. Eğitim süreci sırasında giriş görüntüsü her zaman (224, 224, 4) çözünürlüğündedir, dört adet kanal sayısının olması nedeni giriş görüntüsünün Bayer paterne sahip olmasıdır, bu tezde kullanılan veriseti “RGBG” desenindedir.

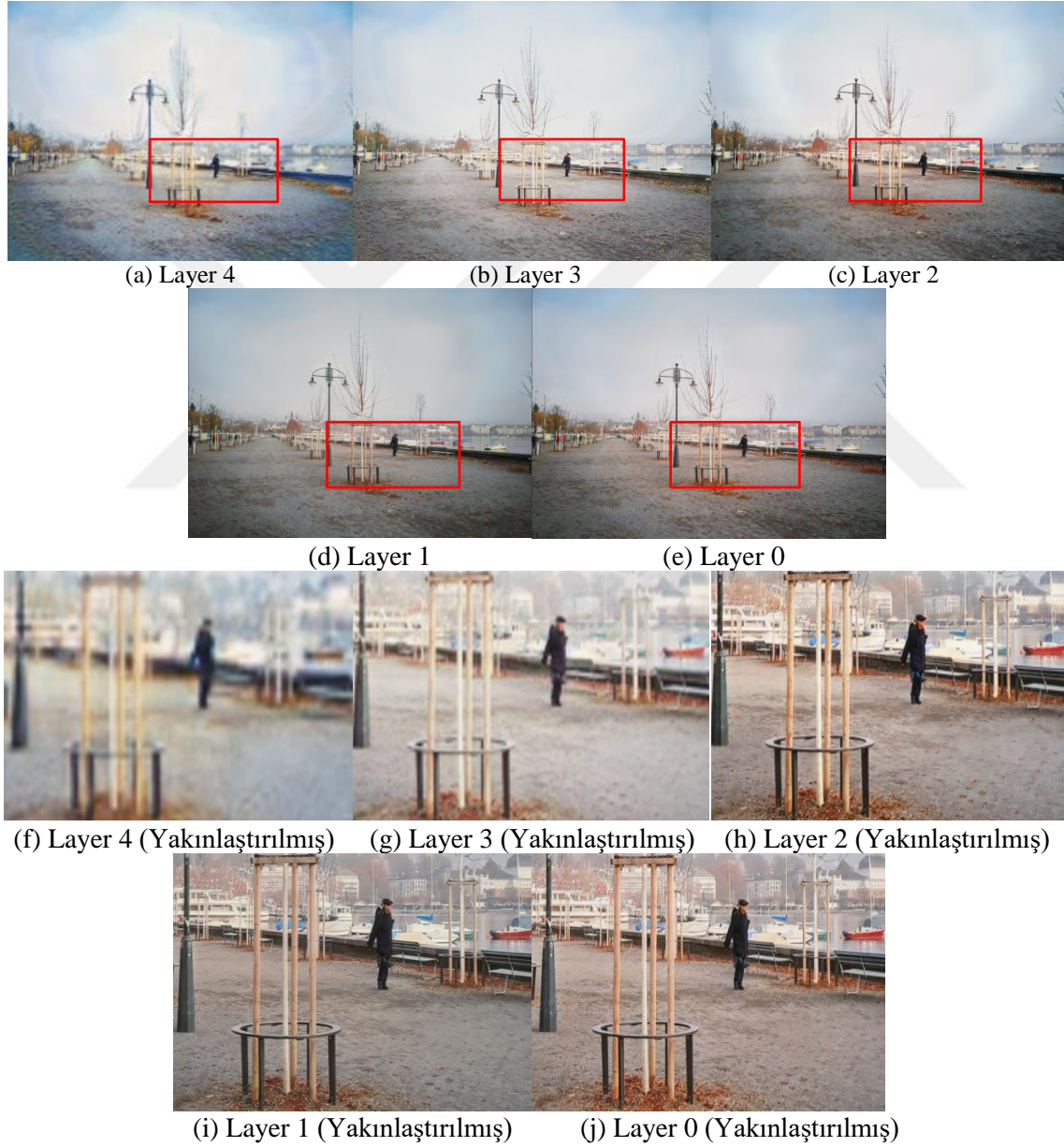
Bu tezin özgün katkılarından biri önerilen yeni model mimarisidir. PyNET ve Lightweight PyNET'in (Cheng ve diğ.,2021) aksine; önerilen modelde Effective Channel Attention (ECA) Blocks, Convolutional Block Attention Modules (CBAM), Information Extraction Modules (IEM) ve SimAM Attention Block olarak adlandırılan çeşitli dikkat modülleri kullanılmıştır. Bu modüller, Şekil 5.1'de gösterildiği gibi spesifik çoklu evrişimli bloklardan sonra konuşlandırılmıştır. Ayrıca önerilen model mimarisinde Layer-1 ve Layer-2 olarak adlandırılan katmanlarda SimAM dikkat blokları kullanılmıştır.



Şekil 5.1. Önerilen çalışmanın genel mimarisi

Her bir katmanda kullanılan dikkat modüllerini tanıtmadan önce, hedef görüntü ile aynı çözünürlükteki nihai çıktıyı oluşturma aşamasında her bir katmanın etkisi Şekil 5.2'de

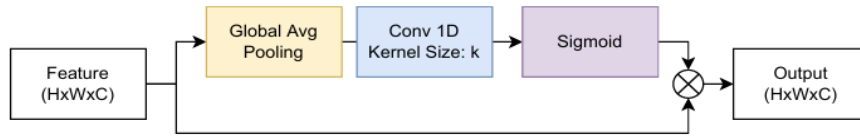
sunulmuştur. Sunulan örnek görüntü test amaçlı model başarımını belirlemede kullanılan görüntülerden biri olup, önerilen nihai modelin parametreleri ile oluşturulan katmanlara ait çıktılardır. Şekil 5.2. de sunulan (a), (b), (c) ve (d) sırasıyla giriş görüntüsünün 0.0625x, 0.125x, 0.25x ve 0.5x çözünürlükteki model çıktılarıdır. Şekil 5.2. (e) ile sunulan görüntü ise hedef görüntünün çözünürlüğündedir, “x” değeri bu çözünürlüğü ifade etmektedir. Şekil 5.2.’de sunulan (f), (g), (h), (i) ve (j) kırmızı kutuda işrate edilen yakınlaştırılmış görüntülerdir.



Şekil 5.2. Katmanlara ait çıktıların görsel karşılaştırılması. (imge 2)

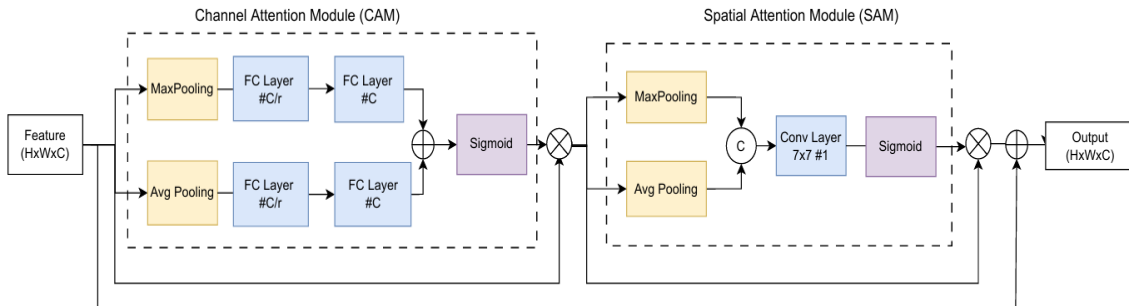
Renk, parlaklık gibi global özelliklerin çıkarıldığı Layer-4 ve Layer-3 katmanında ECA-Net (Wang ve diğ., 2020) ve CBAM (Woo ve diğ., 2018) dikkat modülleri kullanılmıştır. ECA-Net bloğu ile kanallar arası öznitelikler Şekil 5.3'de gösterildiği gibi çıkarılır. Layer-4 ve Layer-3'te giriş özelliklerindeki kanal sayısı diğer katmanlara göre nispeten daha fazladır. Boyutsallığın azalmasını önlemek amacıyla, kanal bazında ilişkilerin 1D evrişim yoluyla çıkarıldığı ECA-Net kullanılmıştır. Önerilen modelde, uyarlanabilir çekirdek boyutu  $k$ 'nin belirlenmesi için Denklem (5.1)'de yer alan  $b=1$  ve  $\gamma=2$  hiperparametreleri kullanılmıştır. Denklem (5.1)'de yer alan  $C$  kanal sayısını,  $f(x) = |x|_{\text{odd}}$  ise en yakın pozitif tek tamsayıyı elde etmek için kullanılan fonksiyonu ifade etmektedir.

$$k = f(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (5.1)$$



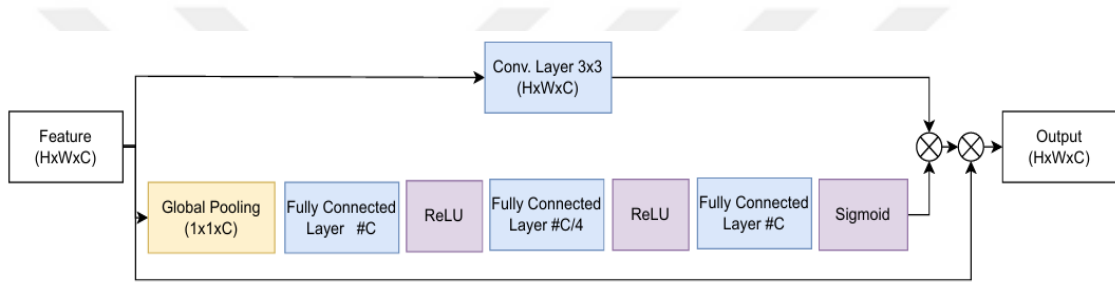
Şekil 5.2. Effective Channel Attention (ECA) blok diyagramı

CBAM modülü Şekil 5.4.'te blok diyagramı sunulduğu üzere seri olarak bağlı CAM kanal dikkat modülü ve SAM olarak adlandırılan uzamsal dikkat modülünden oluşmaktadır. CAM modülü kanallar arası öznitelikleri çıkartmayı amaçlarken, SAM modülü ise giriş olarak verilen öznitelik haritasından uzamsal ilişkileri çıkartmaktadır. Önerilen modelde, SAM için  $7 \times 7$  çekirdek boyutu kullanılmıştır. CAM bloğu için ise “multi-layer perceptron ratio” olarak anılan “ $r$ ” hiperparametresi, Layer-4’te  $r=16$ , Layer-3’te ise  $r=8$  olarak kullanılmıştır.



Şekil 5.3. Convolutional Block Attention Modules (CBAM) blok diyagramı

Layer-2 ve Layer-1'de sahnedeki görüntüde yer alan şekillere ait öznelikler ve nesnelere renk detayları öğrenilir. Bu iki katmanda, ilk çoklu evrişimli katmandan sonra, EDLLIE-Net (Ke ve diğ.,2020) çalışmasından esinlenilerek çok ölçekli öznelik çıkarma işlemi uygulanmıştır, daha sonra ise Şekil 5.5'te blok diyagramı sunulan IEM (Information Extraction Module) alt modülü kullanılmıştır. IEM, aynı anda hem uzamsal hem de kanallar arası yararlı bilgilerin çıkarılmasını sağlar. Söz konusu modülde kanallar arası bilgiden çıkartılan parametreleri ağırlık olarak kullanarak, çekirdek boyutu 3 olan evrişim katmanının çıktısıyla çarpılarak öznelik haritasının değerini atar. Son olarak giriş öznelik haritasıyla eleman bazında çarpma işlemi uygulanarak çıkış katmanı elde edilir.



Şekil 5.4. Information Extraction Module (IEM) blok diyagramı

Layer-1 ve Layer-2'de çoklu ölçekte çıkarılan özelliklerin birleştirilmesinden sonra parametresiz SimAM (Yang ve diğ.,2021) dikkat modülü uygulanır. SimAM, sinirbilim teorisine dayalı olarak tanımlanmış enerji fonksiyonu ile çıkartılan üç boyutlu tahmini dikkat ağırlıklarını kullanır. Layer-1 ve Layer-2 katmanlarında işlenen öznelik haritasına ait yükseklik ve genişlik boyutlarının daha büyük olması nedeniyle, ilave eğitilebilir parametre gerektirmeyen SimAM dikkat bloğu kullanılmıştır. Böylelikle ilave model karmaşıklığına sebep olmadan, ağırlık öğrenme esnasında önemli uzamsal bölgelere ve kanallara odaklanması sağlanmıştır.

## 5.2. Eğitimde Kullanılan Kayıp Fonksiyonları ve Eğitim Parametreleri

Önerilen modelin her katmanı farklı kayıp fonksiyonu terimlerinin toplamıyla ifade edilen kayıp fonksiyonu kullanılarak eğitilmiştir. Eğitim sırasında kullanılan toplam kayıp fonksiyonları ve eğitim parametreleri Tablo 5.1'de gösterilmektedir.

Layer-4'ün eğitimi esnasında, 8x ile örneklemlenmiş hedef görüntünün kullanıldığı göz

önünde bulundurulurken, algısal kayıp fonksiyonlarının hiçbiri kullanılmamaktadır. Layer-3'ten başlayarak VGG tabanlı (Johnson ve diğ., 2016) algısal kayıp kullanılmıştır. Öğrenme işleminin daha yüksek çözünürlükte gerçekleştirildiği Layer-1 ve Layer-0'da, giriş ve çıkış arasındaki yapısal benzerliği sağlamak için SSIM (Wang ve diğ., 2003) kayıp fonksiyonundan yararlanılmıştır. Her katmanın eğitiminde, toplam kayıp fonksiyonunda ortalama karesel hata (Mean Square Error - MSE) kayıp fonksiyonu ve histogram temelli kayıp fonksiyonu kullanılmıştır. Kayıp fonksiyonuna ait ağırlıkların öğrenilebilir parametre olarak tanımlanarak modelin eğitilmesi ile elde edilen sonuçlar Bölüm 6.3.4'te sunulmuştur.

Tablo 5.1. Kullanılan eğitim parametreleri ve toplam kayıp fonksiyonu

Katman	Batch Size	Epoch	Toplam Kayıp Fonksiyonu
4	50	15	$L_{Level\ 4} = L_{MSE} + 0,25 \cdot L_{Histogram}$
3	35	20	$L_{Level\ 3} = 10 \cdot L_{MSE} + 0,25 \cdot L_{Histogram} + L_{VGG}$
2	25	25	$L_{Level\ 2} = 10 \cdot L_{MSE} + 0,25 \cdot L_{Histogram} + L_{VGG}$
1	15	33	$L_{Level\ 1} = 10 \cdot L_{MSE} + 0,4 \cdot (1 - L_{SSIM}) + 0,25 \cdot L_{Histogram} + L_{VGG}$
0	10	60	$L_{Level\ 0} = 10 \cdot L_{MSE} + 0,25 \cdot L_{Histogram} + L_{VGG} + 0,4 \cdot (1 - L_{SSIM})$

Bu tezin literatüre yaptığı katkılardan bir diğeri ise histogram tabanlı kayıp fonksiyonunun eğitim esnasında Kamera ISP veriyolunun eğitiminde kullanılmasıdır. Hedef görüntünün histogram bilgisinden yararlanabilmek amacıyla, RGB-uv özniteliklerinin çıkartılması, HistoGAN'da (Afifi ve diğ., 2021) tanımlanan metodoloji ile gerçekleştirilmiştir. Belirli bir, I, giriş görüntüsü için, yeşil kanal birincil kanal seçilerek log-kroma uzayındaki terimler Denklem (5.2) ve (5.3) 'de sunulduğu gibi hesaplanır,

$$I_{uG}(x) = \log \left( \frac{I_G(x) + \epsilon}{I_R(x) + \epsilon} \right) \quad (5.2)$$

$$I_{vG}(x) = \log \left( \frac{I_G(x) + \epsilon}{I_B(x) + \epsilon} \right) \quad (5.3)$$

$R$ ,  $G$ ,  $B$  indisleri giriş görüntüsünün ilgili renk kanalını belirtmektedir,  $\epsilon$  parametresi ise formülün kararlılığını sağlamak için kullanılmakta olup  $10^{-6}$  olarak seçilmiştir.  $x$  uzamsal düzlemde piksel konumunu;  $(uG, vG)$  ise yeşil kanal birincil kanal olduğunda log-kroma uzayındaki koordinatları ifade etmektedir. Türevi alınabilir histogram özneliği Denklem (5.4)'te sunulduğu gibi hesaplanmaktadır,

$$H(u, v, c) = \alpha \sum_x k(I_{uG}(x), I_{vG}(x), u, v) I_y(x) \quad (5.4)$$

Denklem (5.4)'te yer alan,  $I_y(x) = \sqrt[2]{I_R^2(x) + I_G^2(x) + I_B^2(x)}$  piksel yoğunluğunu,  $\alpha$  parametresi ise normalizasyon maksadıyla kullanılan ağırlıklığı ifade etmektedir. Bu ağırlıklandırma ile  $\sum_{u,v,G} H(u, v, G) = 1$  olması sağlanır. Hesaplamada kullanılan  $k(\cdot)$  fonksiyonu “inverse-quadratic kernel function” olarak anılan ve Denklem (5.5)'de sunulan çekirdek fonksiyonudur.

$$k(I_{uG}, I_{vG}(x), u, v) = \left(1 + \left(\frac{|I_{uG} - u|}{\tau}\right)^2\right)^{-1} \times \left(1 + \left(\frac{|I_{vG} - v|}{\tau}\right)^2\right)^{-1} \quad (5.5)$$

Denklem (5.5) yer alan çekirdek fonksiyonunda,  $\tau$ , histogramın pürüzsüzlüğünü kontrol etmek amacıyla kullanılan bir parametredir ve 0.02 olarak seçilmiştir. Ayrıca, Histogram boyut büyüklüğü 64 olarak belirlenmiştir. Histogram kayıp fonksiyonu, hedef görüntü ve tahmin edilen görüntüye ait türevi alınabilir histogramlar kullanılarak Denklem (5.6)'da sunulan Hellinger Uzaklık fonksiyonu kullanılarak tanımlanmıştır,

$$L_{\text{hist}} = \frac{1}{\sqrt{2}} \|H_{\text{gt}}^{1/2} - H_{\text{pred}}^{1/2}\|_2 \quad (5.6)$$

Eğitim esnasında her bir katman için kullanılan veri kümesi büyüklüğü (Batch Size) ve veriseti üzerindeki tur sayısı Tablo 5.1'de sunulmuştur. Model ağ yapısında yer alan toplamda 6.5M parametere ADAM (Knigam ve Ba, 2015) algoritması ile optimize edilmiştir. Adım boyu olarak  $5 \times 10^{-5}$  kullanılmıştır. Müteakip bölümlerde sonuçları sunulmuş olan önerilen modelin eğitimi NVIDIA RTX A5000 GPU ile Pytorch kütüphanesi kullanılarak gerçekleştirilmiştir.

## 6. DENEYSEL SONUÇLAR

Bu bölümde tez kapsamında gerçekleştirilen denelere ait nicel ve nitel sonuçlar paylaşılmaktadır. Bu kapsamda ilk olarak önerilen modelin niceliksel ve niteliksel performansı referans olarak kabul edilen model performanları ile karşılaştırılmıştır. Zurich veri setinde bulunan 10 adet tam çözünürlüklü (2976, 3968) ham görüntü test amacıyla kullanılmıştır. Karşılaştırma maksadıyla bu bölümde kullanılan model mimarileri aşağıda sunulmuştur;

- PyNET (Ignatov ve diğ., 2020) mimarisinin tüm Zurich veri seti ile eğitilmesiyle elde edilen modelin performansı
- PyNET (Ignatov ve diğ., 2020) mimarisinin, bu tezde belirtilen metot ile hazırlanan veriseti kullanılarak eğitilmesiyle elde edilen modelin performansı (Erbaş ve Çelebi, 2023)
- Lightweight PyNET (Cheng ve diğ.,2021) mimarisinin hazırlanan veri seti ile eğitilmesi ile elde edilen modelin performansı

Daha sonra aynı 10 adet test görüntüsü kullanılarak, geliştirilen model performansı markette kullanıcı tarafından kullanılan Huawei P20 ve Canon 5D Mark IV DSLR fotoğraf makinesine ait kamera ISP veri yolu performansları ile karşılaştırılmıştır. Bu karşılaştırma işlemine yönelik elde edilen sonuçlar, önerilen model ile referans modeller arasındaki karşılaştırmada olduğu gibi niceliksel ve niteliksel olarak sunulmuştur.

Sonrasında ise çeşitli etkinlik analizi çalışmalarından elde edilen sonuçlar okuyucu ile paylaşılmıştır. Böylelikle, model mimarisinin, verisetinin ve kayıp fonksiyonu ağırlık parametrelerinin model başarımına etkisi irdelenmiştir. Bu kapsamda gerçekleştirilen etkinlik analiz çalışmaları aşağıda listelenmiştir;

- Model mimarisi alt modül etkinlik analizi
- Histogram kayıp fonksiyonu terimine ait ağırlık katsayısının belirlenmesi
- Veriseti Çözünürlüğü etkinlik analizi
- Kayıp fonksiyonu terimlerinin öğrenilebilir parametre olarak ağırlıklandırılması
- Veri çoğaltma işlemi etkinlik analizi

## 6.1. Nicel Sonular

Önerilen modelin, referans modeller ile karşılaştırmalı olarak başarımının ortaya konulabilmesi maksadıyla test için ayrılan 10 adet tam çözünürlüklü (2976 x 3968) ham görüntü, giriş verisi olarak kullanılmıştır. Bahse konu ham giriş verilerinin kullanılması ile önerilen modelden ve referans modellerden elde edilen çıktı görüntüleri, Tez'in 3'ncü bölümünde tanıtılan referans görüntüye ihtiyaç duymayan görüntü kalite metriklerinden MUSIQ (Ke ve diğ., 2021), NIMA (Talebi ve Milanfar, 2018), LIQE (Zhang ve diğ., 2023) ve MANIQA (Yang ve diğ., 2022) kullanılarak niceliksel olarak değerlendirilmiştir.

Her dört metrik de insan algısı ile uyumlu olarak, MOS puanları ile yüksek korelasyona sahip sonuçlar üretebilmesi nedeniyle tez kapsamında kullanımı tercih edilmiştir. Ayrıca kullanılan bahse konu metriklere ait yayınların güncelliği, atıf sayıları ve benzer çalışmalarda kullanım yaygınlığı da dikkate alınmıştır.

Modellerin referansa ihtiyaç duymayan görüntü kalite metrikleri ile karşılaştırılmasından önce, Tablo 6.1'de karşılaştırılan modellere ait; toplam eğitilen parametre sayısı, model boyutu ve tam çözünürlükteki (2976 x 3968) bir görüntü için çıkarım zamanı (inference time) karşılaştırmaları sunulmuştur. Sonuçlardan görüleceği üzere Önerilen Model PyNET'e göre model karmaşıklığı daha az ve çıkarım zamanı gerçek zamanlı uygulamalar için daha uygundur. Sonraki bölümlerde bahsedilecek model başarımları dikkate alındığında ise, Lightweight PyNET'e kıyasla kabul edilebilir sayıda parametre artışı mevcuttur.

Tablo 6.1. Genel model performans karşılaştırma

Model	PyNET	Lightweight PyNET	Önerilen Model
Eğitilen Parametre Sayısı	47 554 738	5 705 167	6 473 474
Model Dosyası Büyüklüğü	185,8 MB	22,3 MB	24,7 MB
Çıkarım Zmanı (Inference Time)	1,9 s	0,71 s	0,83 s

Tablo 6.2. Önerilen model ile referans modellerin ile karşılaştırılması

Model Veriseti	PyNET Zurich Veriseti				PyNET Hazırlanmış Veriseti				Lightweight PyNET Hazırlanmış Veriseti				Önerilen Model Hazırlanmış Veriseti			
	NIMA	MUSIQ	LIQE	MANIQA	NIMA	MUSIQ	LIQE	MANIQA	NIMA	MUSIQ	LIQE	MANIQA	NIMA	MUSIQ	LIQE	MANIQA
imge 0	6,38	60,01	2,34	0,3530	<u>6,93</u>	<b>67,58</b>	<b>4,26</b>	<u>0,4167</u>	6,72	62,93	3,18	0,3890	<b>7,17</b>	<u>65,97</u>	<u>3,32</u>	<b>0,4199</b>
imge 1	6,81	60,91	3,66	0,3717	<u>6,85</u>	<b>66,75</b>	<b>4,89</b>	<u>0,3906</u>	6,70	62,68	3,57	0,3721	<b>6,99</b>	<u>64,91</u>	<u>4,25</u>	<b>0,4090</b>
imge 2	4,18	<u>61,36</u>	3,82	0,3519	<u>4,54</u>	60,18	<b>4,57</b>	<u>0,3606</u>	4,57	60,00	4,03	0,3273	<b>5,21</b>	<b>62,02</b>	<u>4,19</u>	<b>0,3673</b>
imge 3	6,47	58,54	3,08	0,4248	<u>6,65</u>	<u>63,38</u>	<b>4,81</b>	<u>0,4980</u>	6,56	60,61	3,30	0,4240	<b>6,74</b>	<b>63,57</b>	<u>3,58</u>	<b>0,5047</b>
imge 4	5,49	61,13	3,17	0,3899	<u>5,99</u>	<b>66,65</b>	<b>4,88</b>	<u>0,4343</u>	5,80	63,24	3,61	0,3945	<b>6,16</b>	<u>64,81</u>	<u>3,99</u>	<b>0,4378</b>
imge 5	<u>6,73</u>	61,84	2,42	0,3219	6,44	<b>64,47</b>	<b>4,47</b>	<u>0,3606</u>	6,56	62,84	2,97	0,3219	<b>6,83</b>	<u>63,74</u>	<u>3,59</u>	<b>0,3642</b>
imge 6	6,85	<b>63,01</b>	3,16	0,3375	<b>7,46</b>	62,38	<b>4,48</b>	<b>0,3952</b>	<u>7,16</u>	62,07	3,25	0,3531	<b>7,46</b>	<u>62,63</u>	<u>3,39</u>	<u>0,3918</u>
imge 7	5,15	61,93	2,84	0,3881	<u>5,30</u>	<b>65,76</b>	<b>4,05</b>	<u>0,3885</u>	4,97	63,23	3,05	0,3721	<b>5,56</b>	<b>65,76</b>	<u>3,54</u>	<b>0,4050</b>
imge 8	6,20	60,92	3,32	0,4057	<u>6,68</u>	<b>63,79</b>	<b>4,76</b>	<u>0,4167</u>	6,32	61,62	3,30	0,3918	<b>6,86</b>	<u>62,95</u>	<u>3,73</u>	<b>0,4179</b>
imge 9	6,01	53,55	2,35	0,3219	<u>6,22</u>	<b>60,63</b>	<b>3,47</b>	<u>0,3272</u>	6,20	56,39	2,77	0,2971	<b>6,43</b>	<u>58,12</u>	<u>3,03</u>	<b>0,3563</b>
Ort. Skor	6,03	60,32	3,02	0,3669	<u>6,30</u>	<b>64,15</b>	<b>4,46</b>	<u>0,3989</u>	6,16	61,56	3,30	0,3643	<b>6,54</b>	<u>63,45</u>	<u>3,67</u>	<b>0,4073</b>

Önerilen modele ve referans modellere ait çıktı görüntülerin NR-IQA metriklerinden olan, MUSIQ, NIMA, LIQE ve MANIQA skorları karşılaştırılmalı olarak Tablo 6.2’de sunulmuştur. NIMA skoru için, MATLAB tarafından sağlanan “Live in the Wild” veri seti ile eğitilmiş model parametreleri kullanılmıştır (URL-5). MUSIQ (URL-6) ve LIQE (URL-8) skorları için ise “konIQ-10K” veriseti ile eğitilen model parametreleri kullanılmıştır. MANIQA (URL-7) skoru için ise KADID-10k veriseti ile eğitilen model parametreleri kullanılmıştır.

NIMA, MUSIQ, LIQE ve MANIQA metriklerinde, yüksek nümerik değere sahip skorlar, daha yüksek kaliteli sonucu ifade etmektedir. Tablo 6.2’de ve sonrasında sunulacak deney sonuçlarında; en iyi sonuç kalın, ikinci en iyi sonuç ise altı çizili ve italik olarak gösterilmiştir. Tablo 6.2’de ye alan sonuçlardan görüleceği üzere, önerilen model ortalamada en iyi NIMA ve MANIQA skoruna sahiptir. MUSIQ ve LIQE skorları dikkate alındığında ise önerilen metodun en iyi ikinci performansa sahip olduğu görülmektedir. Birinci ve ikinci sütunlar karşılaştırıldığında bu tezin dördüncü bölümünde bahsedilen yöntem ile hazırlanan verisetinin sonuca etkisi görülebilmektedir. Diğer taraftan Lighthweight PyNET (Cheng ve diğ.,2021) ve önerilen modelin sonuçları karşılaştırıldığında ise, önerilen model mimarisinin etkisi görülebilmektedir.

Tablo 6.3. Dinamik Aralık (dB) performans karşılaştırma

Model	PyNET	Lightweight PyNET	Önerilen Model
imge 0	3,57	3,45	<b>3,65</b>
imge 1	3,87	3,68	<b>4,04</b>
imge 2	2,70	2,69	<b>2,91</b>
imge 3	3,81	3,92	<b>4,37</b>
imge 4	3,88	<b>3,89</b>	3,78
imge 5	<b>4,62</b>	4,59	4,55
imge 6	3,47	<b>3,53</b>	3,52
imge 7	<b>3,73</b>	3,19	3,54
imge 8	4,01	4,01	<b>4,28</b>
imge 9	3,69	<b>4,26</b>	3,99
Ortalama Skor	3,74	3,72	<b>3,85</b>

Önerilen ve referans model çıktılarının sahip olduğu dinamik aralıklar ise Tablo 6.3’de sunulmuştur. Bir görüntünün dinamik aralığı, sahnenin maksimum ve minimum ışık yoğunlukları arasındaki oranı tanımlayan bir diğer önemli faktördür.

Dinamik aralığı hesaplamak için öncelikle RGB renk kanal değerlerinden bağıl ışıklılık (luminance) değeri Denklem (6.1) ile hesaplanır ve ardından Denklem (6.2)'deki gibi olasılık yoğunluk fonksiyonu (probability mass function, pmf) elde edilir,

$$Y = 0.2126 \cdot R + 0.7152 \cdot G + 0.0722 \cdot B \quad (6.1)$$

$$p(Y_k) = \text{pmf}(Y) = \frac{\text{Total pixels with intensity } Y_k}{\text{Total pixels in image}} \quad (6.2)$$

Sonrasında bağıl ışıklılık değerinden histograma ait kümülatif dağılım fonksiyonu (cumulative distribution function, cdf) Denklem (6.3) ile hesaplanır,

$$\text{cdf}_Y = \text{cdf}(Y) = \sum_{k=0}^{L-1} p(Y_k) \quad (6.3)$$

Sonrasında ise minimum ve maksimum değerler %1'lik kısımlar göz ardı edilerek Denklem (6.4) ve (6.5) ile belirlenir. %1' lik kısımların göz ardı edilmesinin sebebi olası gürültü etkilerinden sakınmaktır.

$$\min_{\text{cdf}_Y} = \min(\text{cdf}_Y > 0.01) \quad (6.4)$$

$$\max_{\text{cdf}_Y} = \min(\text{cdf}_Y > 0.99) \quad (6.5)$$

Belirlenen minimum ve maksimum yoğunluklardan logaritmik ölçekte Denklem (6.6) ile dinamik aralık hesaplanır.

$$\text{DR} = \log_2(\max_{\text{cdf}_Y}) - \log_2(\min_{\text{cdf}_Y}) \quad (6.6)$$

Tablo 6.3'de sunulan sonuçlarda ortalama sonuçlar dikkate alındığında, referans modellere kıyasla önerilen model mimarisi ile daha yüksek dinamik aralığa sahip görüntülerin oluşturulabildiği görülmüştür.

Önerilen model çıktıları ile aynı sahnenin Huawei P20 VE Canon 5D Mark IV DSLR kamera ISP çıktıları, NIMA, MUSIQ, LIQE ve MANIQA skorları ile Tablo 6.4'de karşılaştırılmıştır. Karşılaştırma maksadıyla aynı 10 adet test görüntüsü kullanılmıştır.

Tablo 6.4. Önerilen model ile farklı kamera ISP çıktılarının karşılaştırılması

ISP Görüntü ID	Huawei P20 Kamera ISP				Canon 5D Mark IV DSLR				Önerilen Model			
	NIMA	MUSIQ	LIQE	MANIQA	NIMA	MUSIQ	LIQE	MANIQA	NIMA	MUSIQ	LIQE	MANIQA
imge 0	<u>7,06</u>	<u>59,36</u>	<u>2,26</u>	<u>0,3622</u>	7,05	55,22	1,15	0,3036	<b>7,17</b>	<b>65,97</b>	<b>3,32</b>	<b>0,4199</b>
imge 1	<b>7,53</b>	<u>60,10</u>	2,75	<u>0,3513</u>	<u>7,01</u>	57,59	<u>3,85</u>	0,3210	6,99	<b>64,91</b>	<b>4,25</b>	<u>0,4090</u>
imge 2	4,66	<b>64,22</b>	2,77	0,3298	<u>4,92</u>	<u>63,51</u>	<u>4,04</u>	<b>0,3747</b>	<b>5,21</b>	62,02	<b>4,19</b>	<u>0,3673</u>
imge 3	6,71	<u>60,51</u>	2,79	0,3868	<b>6,82</b>	54,65	<u>3,43</u>	<u>0,4563</u>	<u>6,74</u>	<b>63,57</b>	<b>3,58</b>	<b>0,5047</b>
imge 4	<u>5,74</u>	<u>61,86</u>	2,56	<b>0,4984</b>	5,73	58,86	<u>3,59</u>	0,4105	<b>6,16</b>	<b>64,81</b>	<b>3,99</b>	<u>0,4378</u>
imge 5	6,62	<u>61,18</u>	2,43	<u>0,3578</u>	<u>6,63</u>	59,37	<u>2,55</u>	0,3518	<b>6,83</b>	<b>63,74</b>	<b>3,59</b>	<b>0,3642</b>
imge 6	<b>7,54</b>	61,45	2,27	<b>0,4241</b>	<u>7,51</u>	<b>63,70</b>	<u>3,20</u>	0,4172	7,46	<u>62,63</u>	<b>3,39</b>	<u>0,3918</u>
imge 7	<b>6,3</b>	<u>60,08</u>	2,02	0,3345	<u>6,04</u>	56,83	<u>2,87</u>	<u>0,3636</u>	5,56	<b>65,76</b>	<b>3,54</b>	<b>0,4050</b>
imge 8	6,84	<u>59,67</u>	2,52	0,3592	<b>6,93</b>	56,13	<u>2,86</u>	<u>0,3919</u>	<u>6,86</u>	<b>62,95</b>	<b>3,73</b>	<b>0,4179</b>
imge 9	<b>6,49</b>	<u>54,72</u>	2,28	<u>0,3312</u>	6,25	52,74	<u>2,70</u>	0,3200	<u>6,43</u>	<b>58,12</b>	<b>3,03</b>	<b>0,3563</b>
Ortalama Skor	<b>6,55</b>	<u>60,31</u>	2,46	<u>0,3735</u>	6,49	57,8	<u>3,02</u>	0,3711	<u>6,54</u>	<b>63,45</b>	<b>3,67</b>	<b>0,4073</b>

Ancak karşılaştırılan görüntülerin çözünürlükleri, sensör boyutu ve lens kısıtlamaları nedeniyle birbirinden farklıdır. Adil bir karşılaştırma yapılabilmesi maksadıyla doğallığı etkileyebilecek yeniden boyutlandırma ve kırpma işlemleri uygulanmamıştır.

Tablo 6.4’de yer alan sonuçlar dikkate alındığında, önerilen model ile elde edilen sonuçların her iki ISP kamera veriyolu performansından da daha iyi olduğu görülmektedir. Elde edilen sonuçlara göre, önerdiğimiz yöntem Huawei P20 ISP ile hemen hemen benzer NIMA puanına sahiptir. MUSUQ, LIQE ve MANIQA skorlarına bakıldığında ise önerilen yöntemin her iki kamera ISP performansına kıyasla daha iyi sonuçlar üretebildiği gözlemlenmiştir.

## 6.2. Nitel Sonuçlar

Modelimizin başarısını görsel olarak değerlendirebilmek maksadıyla, önerilen yöntemle oluşturulan görüntüler, Tablo 6.2’de sırasıyla birinci ve üçüncü sütunlara karşılık gelen PyNET (Ignatov ve diğ., 2020) ve Lightweight PyNET (Cheng ve diğ., 2021) model çıktıları ile karşılaştırılmıştır. Örnek olarak sunulan görüntüler Şekil 6.1, Şekil 6.2, Şekil 6.3 ve Şekil 6.4’te sunulmuştur.

Şekil 6.1’de sunulan ilk örnek görüntüler incelendiğinde, önerilen model ile sahne ışıklılığının daha doğru şekilde oluşturulabildiği görülmektedir. Bu etki, Şekil 6.1 (f)’de yer alan sahne ortasındaki mavi arabaya odaklanıldığında fark edilebilmektedir. Ayrıca sağ taraftaki ağaca odaklanıldığında önerilen modelin detay iyileştirme açısından daha başarılı sonuçlar üretebildiği anlaşılmaktadır.

Şekil 6.2.’de yer alan ikinci örnek görüntüye ait görseller incelendiğinde, sahnenin zeminine ait yerel ton haritalama etkisi görülebilmektedir. Aynı zamanda bina dokusuna ait detay bilgilerin önerilen model ile daha doğal şekilde oluşturulabildiği görülmektedir.

Şekil 6.3.’de sunulan görüntüler incelendiğinde ise bina altındaki alanda yer alan bitki örtüsünün önerilen yöntem ile PyNET çıktısına kıyasla daha az gürültülü ve bulanık bir şekilde oluşturulabildiği görülmektedir. Önerilen model mimarisinin renk bilgisini oluşturmadaki pozitif etkisi ve detay bilgisinin daha iyi çıkartılmasındaki etkisi ikinci ve üçüncü sütunlara odaklanıldığında daha belirgin olarak görülebilmektedir.



Şekil 6.1. Önerilen ve referans model çıktılarının görsel karşılaştırması-1

Zurich Veriseti

Hazırlanan Veriseti

Hazırlanan Veriseti



(a) imge 8 (PyNET Çıktısı)

(b) imge 8 (Lightweight PyNET Çıktısı)

(c) imge 8 (Önerilen Model Çıktısı)



(d) imge 8 (PyNET Çıktısı-Yakınlaştırılmış)

(e) imge 8 (Lightweight PyNET Çıktısı-Yakınlaştırılmış)

(f) imge 8 (Önerilen Model Çıktısı-Yakınlaştırılmış)

Şekil 6.2. Önerilen ve referans model çıktılarının görsel karşılaştırması-2

Zurich Veriseti

Hazırlanan Veriseti

Hazırlanan Veriseti



(a) imge 3 (PyNET Çıktısı)

(b) imge 3 (Lightweight PyNET Çıktısı)

(c) imge 3 (Önerilen Model Çıktısı)



(d) imge 3 (PyNET Çıktısı-  
Yakınlaştırılmış)

(e) imge 3 (Lightweight PyNET Çıktısı-  
Yakınlaştırılmış)

(f) imge 3 (Önerilen Model Çıktısı-  
Yakınlaştırılmış)

Şekil 6.3. Önerilen ve referans model çıktılarının görsel karşılaştırması-3



Şekil 6.4. Önerilen ve referans model çıktılarının görsel karşılaştırması-4

Şekil 6.4'te ise (e) ve (f) ile sunulan görüntüler incelendiğinde, önerilen modelin bitki örtüsü bölgesindeki renk bilgisinin insan algısına uyumlu olarak daha doğru ve doğal bir şekilde oluşturulabildiği gözlemlenmektedir.

Önerilen modele oluşturulan görüntüler ile Huawei P20 ve Canon 5D Mark IV DSLR marka/modele sahip kameraların yerleşik ISP'lerinin oluşturduğu görüntülerin görsel olarak karşılaştırması Şekil 6.5, Şekil 6.6, Şekil 6.7 ve Şekil 6.8' de sunulmuştur. Bahse konu görüntülerin nicel sonuçları Bölüm 6.1'de yer alan Tablo 6.4'te sunulmuştur.

Sunulan örnek görüntüler incelendiğinde; önerilen yöntemin, diğer IPS kamera veriyolu çıktılarına kıyasla parlak ve yansıma olan bölgelerde daha canlı ve doğal sonuçlar üretebildiği görülmektedir. Örneğin, Şekil 6.5 (d) ve eşlenik diğer çıktılar incelendiğinde, arabalara ve ışık yansımalarına odaklanarak, hem Canon ISP'de hem de önerilen yöntemde parlaklığın Huawei P20 kamera ISP performansına kıyasla daha iyi oluşturulabildiği görülmektedir. Aynı etki Şekil 6.6'daki sahnede yer alan tabela üzerindeki yansımalarda da görülebilmektedir. Renk bilgisinin insan görme algısı ile uyumlu olarak daha doğal olarak oluşturulabildiği ise Şekil 6.7'de ve Şekil 6.8'de yer alan bitki örtüsü bölgelerine odaklanılarak gözlemlenebilmektedir. Bitki örtüsünü oluşturan yeşil ve kırmızı ağırlıklı renklerin önerilen model ile daha canlı olarak oluşturulabildiği görülmektedir.

Bununla birlikte sahnede yer alan yüksek pozlamalı bölgelere odaklanıldığında, örneğin Şekil 6.6 (b)'de gökyüzünde bulunan bulutlara bakılabilir, Huawei P20 kamerasına ait ISP veriyolunun ve önerilen modelin, Canon 5D Mark'a göre daha fazla detay çıkartabildiği görülebilmektedir. Önerilen model ile Canon 5D Mark kamera ISP çıktıları karşılaştırıldığında ise kırmızı renk bilgisinin oluşturulmasında önerilen modelin daha başarılı olduğu gözlemlenmektedir. Söz konusu etkiyi görebilmek için Şekil 6.7 (f)'de bina altında yer alan yeşil alandaki kırmızı tonlara ve ağaçlardaki yapraklara odaklanılabilir. Önerilen model Canon 5D Mark kamera çıktısına kıyasla daha canlı kırmızı tonların oluşturulabildiği görülmektedir. Şekil 6.8 (f)'de ise bitki örtüsündeki kırmızı tonlarda aynı olumlu etki Şekil 6.8 (e) ile kıyaslandığında görülebilmektedir.



(a) Huawei P20 Kamera ISP Çıktısı,  
Çözünürlük (3840, 5120)

(b) Canon 5D Mark IV DSLR Kamera ISP  
Çıktısı, Çözünürlük (4480, 6720)

(c) Önerilen Model Çıktısı,  
Çözünürlük (2880, 3968)

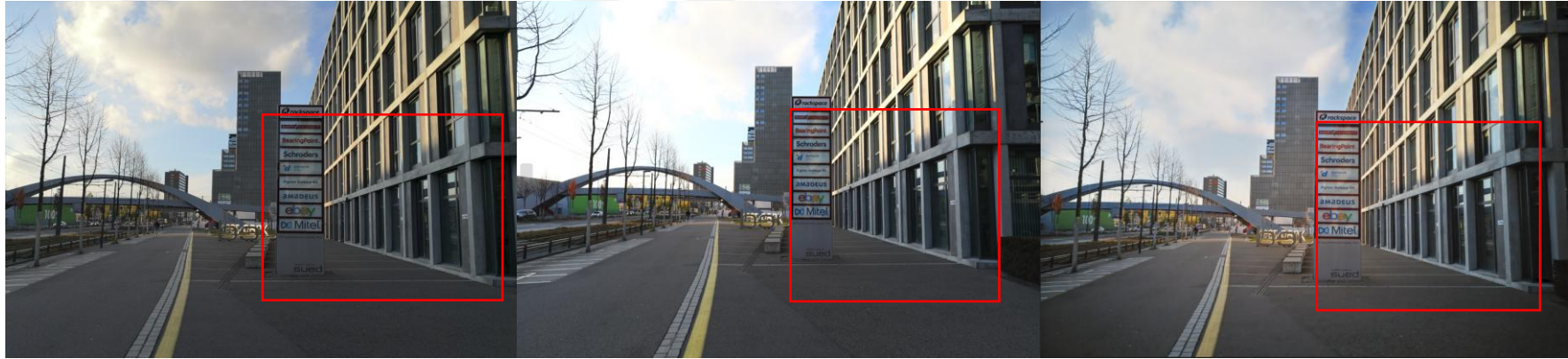


(d) Huawei P20 Kamera ISP Çıktısı,  
(Yakınlaştırılmış)

(e) Canon 5D Mark IV DSLR Kamera ISP  
Çıktısı, (Yakınlaştırılmış)

(f) Önerilen Model Çıktısı,  
(Yakınlaştırılmış)

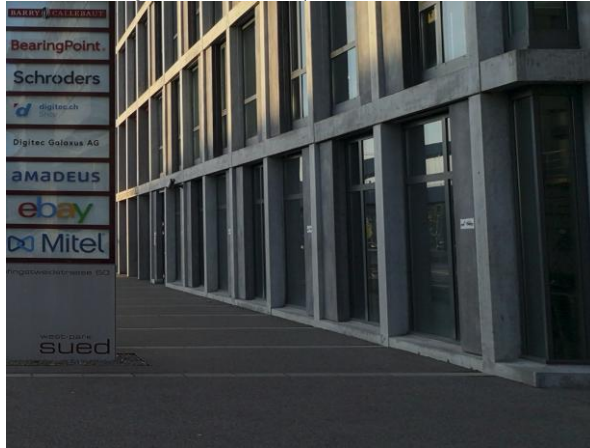
Şekil 6.5. Önerilen model ile farklı kameraların ISP çıktılarının görsel karşılaştırması-1 (imge 5)



(a) Huawei P20 Kamera ISP Çıktısı,  
Çözünürlük (3840, 5120)

(b) Canon 5D Mark IV DSLR Kamera ISP  
Çıktısı, Çözünürlük (4480, 6720)

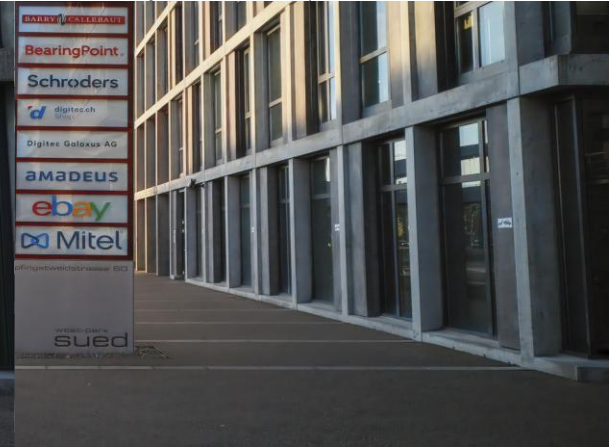
(c) Önerilen Model Çıktısı,  
Çözünürlük (2880, 3968)



(d) Huawei P20 Kamera ISP Çıktısı,  
(Yakınlaştırılmış)



(e) Canon 5D Mark IV DSLR Kamera ISP  
Çıktısı, (Yakınlaştırılmış)



(f) Önerilen Model Çıktısı,  
(Yakınlaştırılmış)

Şekil 6.6. Önerilen model ile farklı kameraların ISP çıktılarının görsel karşılaştırması-2 (imge 8)



(a) Huawei P20 Kamera ISP Çıktısı,  
Çözünürlük (3840, 5120)

(b) Canon 5D Mark IV DSLR Kamera ISP  
Çıktısı, Çözünürlük (4480, 6720)

(c) Önerilen Model Çıktısı,  
Çözünürlük (2880, 3968)



(d) Huawei P20 Kamera ISP Çıktısı,  
(Yakınlaştırılmış)

(e) Canon 5D Mark IV DSLR Kamera ISP  
Çıktısı, (Yakınlaştırılmış)

(f) Önerilen Model Çıktısı,  
(Yakınlaştırılmış)

Şekil 6.7. Önerilen model ile farklı kameraların ISP çıktılarının görsel karşılaştırması-3 (imge 3)



(a) Huawei P20 Kamera ISP Çıktısı,  
Çözünürlük (3840, 5120)

(b) Canon 5D Mark IV DSLR Kamera ISP  
Çıktısı, Çözünürlük (4480, 6720)

(c) Önerilen Model Çıktısı,  
Çözünürlük (2880, 3968)



(d) Huawei P20 Kamera ISP Çıktısı,  
(Yakınlaştırılmış)

(e) Canon 5D Mark IV DSLR Kamera ISP  
Çıktısı, (Yakınlaştırılmış)

(f) Önerilen Model Çıktısı,  
(Yakınlaştırılmış)

Şekil 6.8. Önerilen model ile farklı kameraların ISP çıktılarının görsel karşılaştırması-4 (imge 0)

### 6.3. Etkinlik Analizi (Ablation Study)

Model mimarisinde yer alan alt modüllerinin (dikkat mekanizmaları) ve Histogram Kayıp fonksiyonun ağırlığının etkisini anlayabilmek amacıyla bir dizi etkinlik analizi gerçekleştirilmiştir. Bu kısımda etkinlik analizi kapsamında gerçekleştirilen deneylere ait sonuçlar sunulmuştur.

#### 6.3.1. Model Mimarisi Alt Modül Etkinlik Analizi

Bu bölümde önerilen model mimarimizde yer alan farklı katmanlardaki blokların etkisini gözlemleyebilmek amacıyla gerçekleştirilen bir dizi alt-modül etkinlik deneyine ait sonuçlar sunulmuştur. Bu bölümde sunulan alternatif model mimarileri de önerilen modelde olduğu gibi her katman için Tablo 5.1’de belirtilen parametreler ve kayıp fonksiyonları kullanılarak eğitilmiştir. Referansa ihtiyaç duymayan NIMA ve MUSIQ kalite metriklerine ait skorlar ile nicel olarak sonuçlar Tablo 6.5’te sunulmuştur. İlave olarak görsel örnekler de sunularak alt-modül etkileri açıklanmıştır. Alt-modül etkinlik analizinde de, deneysel sonuçlar için kullanılan tam çözünürlüğe sahip 10 adet test görüntüleri kullanılmıştır.

İlk olarak herhangi bir dikkat modülünün kullanılmadığı ve IEM modülünün yer almadığı temel referans model eğitilmiştir. Daha sonra ECA bloğu Layer-3 ve Layer-4 katmanlarına eklenerek ECA dikkat mekanizmasının etkisi gözlemlenmiştir. Tablo 6.5’te en solda temel referans modele ait sonuçlar, soldan ikinci sütunda ise ECA dikkat mekanizmasının eklenmesiyle elde edilen sonuçlar sunulmaktadır. Müteakiben CBAM dikkat modülü aynı katmanlara eklenmiş ve üçüncü sütunda sunulduğu üzere nicel etkisi kaydedilmiştir.

Tablo 6.5’te ikinci ve üçüncü sütunda yer alan sonuçlar incelendiğinde, NIMA skorlarında iyileşme olmasına rağmen MUSIQ skorlarında anlamlı bir olumlu etki tespit edilmediği görülmektedir. Bu farklılığın nedeninin Layer-4 ve Layer-3’te gerçekleşen düşük çözünürlüklü öğrenme sürecinden kaynaklanabileceği değerlendirilmektedir. MUSIQ metriği ile değerlendirilen görseller öncelikle farklı ölçeklerde yamalar halinde örneklenip puanlanır ve nihai skor farklı ölçeklerdeki yamalara ait puanların ortalaması ile

Tablo 6.5. Alt modül etkinlik analizi sonuçlarının karşılaştırılması

Model	Temel Referans Model		Temel Model + ECA		Temel Model + ECA + CBAM		Temel Model + ECA + CBAM + IEM		Önerilen Model	
	NIMA	MUSIQ	NIMA	MUSIQ	NIMA	MUSIQ	NIMA	MUSIQ	NIMA	MUSIQ
imge 0	6,81	64,28	6,83	63,42	<u>7,04</u>	64,21	6,93	<u>65,75</u>	<b>7,17</b>	<b>65,97</b>
imge 1	6,71	62,97	6,90	62,42	<b>7,03</b>	63,85	6,89	<u>64,12</u>	<u>6,99</u>	<b>64,91</b>
imge 2	4,20	60,10	4,47	60,75	4,51	<u>60,89</u>	<u>4,64</u>	<u>60,89</u>	<b>5,21</b>	<b>62,02</b>
imge 3	6,47	61,63	6,68	61,65	<b>6,83</b>	61,92	6,63	62,52	<u>6,74</u>	<b>63,57</b>
imge 4	5,93	62,76	6,01	62,81	5,96	63,32	<b>6,23</b>	<u>63,99</u>	<u>6,16</u>	<b>64,81</b>
imge 5	6,62	63,25	<b>6,97</b>	63,03	6,68	62,71	6,74	<b>64,43</b>	<u>6,83</u>	<u>63,74</u>
imge 6	7,06	<b>63,03</b>	7,20	61,72	<b>7,61</b>	61,03	7,25	61,98	<u>7,46</u>	<u>62,63</u>
imge 7	5,14	63,63	5,31	64,41	4,99	63,89	<u>5,32</u>	<u>64,83</u>	<b>5,56</b>	<b>65,76</b>
imge 8	6,41	62,69	6,57	60,75	6,66	61,84	<u>6,83</u>	<b>63,76</b>	<b>6,86</b>	<u>62,95</u>
imge 9	6,16	56,14	<u>6,46</u>	55,49	<b>6,50</b>	57,01	6,45	<b>58,18</b>	6,43	<u>58,12</u>
Ortalama Skor	6,15	62,05	6,33	61,64	6,37	62,07	<u>6,38</u>	<u>62,95</u>	<b>6,55</b>	<b>63,45</b>

elde edilmektedir, diğer taraftan NIMA metriğinde ise nihai skor doğrudan yeniden ölçeklenen görüntü üzerinden üretilir. Dolayısıyla Layer-4 ve Layer-3'te genel (global) özneliklerin öğrenilmesi nedeniyle eklenen blokların etkisi, MUSIQ skoruna ait sonuçlar ile ayırt edilememektedir.

Tablo 6.5'te yer alan dördüncü sütunda ise, Layer-1 ve Layer-2'de çok ölçekli öznelik çıkarma işlemi sonrasında kullanılan IEM modülünün etkisi görülmektedir. Sayısal sonuçlardan anlaşılacağı üzere yüksek frekanslı detayların oluşturulmasının etkisi MUSIQ skorundaki artıştan anlaşılabilir. Son olarak en sağda yer alan sütunda SimAM Dikkat Bloku'nun eklenmesi ile elde edilen modelin etkisi görülmektedir.

Öte yandan tezin altıncı bölümünde tanıtılan Dinamik Aralık sonuçlarına bakıldığında, önerilen nihai model, temel referans modele kıyasla 0,60 dB'lik bir iyileşme olduğu gözlemlenmiştir.

Alt Modül Etkinlik Analizi'ne ait sonuçların örnek görsel karşılaştırmaları Şekil 6.9, Şekil 6.10, Şekil 6.11, Şekil 6.12, Şekil 6.13 ve Şekil 6.14'de sunulmaktadır. Blokların etkisini gözlemlenmek için detayların ve tüm görüntü bütünlüğünün aynı anda değerlendirilmesi önemlidir. Sonuçlara dikkatle bakıldığında, Layer-3 ve Layer-4'te kullanılan ECA ve CBAM dikkat modüllerinin detayların iyileştirilmesindeki olumlu etkileri sırasıyla Şekil 6.10 (a), (b) ve (c)'de görülmektedir. Özellikle sahnenin ön planında sağda yer alan ağaca ait dokulara ve sahnenin solunda yürüyen insana ait detaylarda iyileştirme olduğu ve renk bilgisinin daha canlı olarak oluşturulabildiği görülmektedir. Benzer olumlu etki sırasıyla Şekil 6.12 (a), (b) ve (c)'de yer alan tabelada yer alan yazılara odaklanıldığında da görülebilmektedir. Her iki örnekte de mavi ve kırmızı renklerinin daha canlı renklerde oluşturulabildiği görülmektedir. Alt-modül etkinlik analizi deneyleri esnasında ECA ve CBAM dikkat modüllerinin Layer-2 ve Layer-1 katmanlarında da kullanımı test edilmiş; ancak özellikle az ışık alan bölgelerde bölgesel olumsuz etkileri olduğu gözlemlenmiştir.

IEM ve SimAM bloklarının etkileri yüksek çözünürlü görüntüde daha kolay görülebilmektedir. Çok katmanlı öznelik çıkarımından sonra Layer-1 ve Layer-2'de kullanılan IEM modülü, uzamsal ve kanallar arası faydalı bilgilerin öğrenilmesi ve sonraki konvolüsyonel ağlara aktarılması amacıyla kullanılmaktadır. Ancak Şekil 6.9 ve

Şekil 6.11’de yer alan yüksek çözünürlüklü görüntülerde düşük pozlamalı bölgelerde detayların ortaya çıkarılabildiği ve sahnede yer alan objelere ait yüksek frekansa sahip dokularda iyileştirmeler olduğu görülebilmektedir. Hem NIMA, hemde MUSIQ metriklerine ait sonuçlara bakıldığında eklenen IEM bloğunun olumlu etkisi görülebilmektedir. Bununla birlikte Şekil 6.10 (d) ve Şekil 6.12 (d)’de yer alan örnekler incelendiğinde, önceki deney çıktılarına kıyasla renklerin solgunlaştığı fark edilmektedir. Bu durum bir önceki ara sonucun daha kaliteli olduğu izlenimi vermektedir. Ancak Şekil 6.9 (d) ve Şekil 6.11 (d)’de yer alan yüksek çözünürlüklü görüntülere bakıldığında, az pozlanmış bölgelerdeki detayların çıkarılabildiği görülmektedir. Bununla birlikte Şekil 6.14 (c)’de CBAM bloğunun eklenmesiyle gökyüzünde oluşan halo etkisinin, IEM bloğunun eklenmesi ile giderilebildiği görülmektedir.

Son olarak SimAM Dikkat Bloku’nun eklenmesiyle nihai önerilen model mimarisine ulaşılmaktadır. Bir önceki Alt Modül Etkinlik Analizi’nde bahsedilen, çok katmanlı öznelik çıkarımı ve sonrasında IEM bloğu çıktılarının birleştirilmesinden sonra SimAM modülü ağ mimarisine eklenmiştir. Şekil 6.9(d) ve Şekil 6.9(e) çıktıları karşılaştırıldığında, renklerin solgunlaşmasına neden olan önceki mimarideki olumsuz etkilerin iyileştirildiği görülmektedir. Söz konusu etki Şekil 6.10 (e) ‘deki yakınlaştırılmış görüntüde net olarak görülebilmektedir, örneğin yakınlaştırılmış görüntünün sol tarafında yer alan ağaç dallarındaki yaprakların sarı renklerin ve yürüyen insana ait kıyafetlerdeki renklerin model tarafından doğru şekilde oluşturulabildiği görülmektedir. Alt modül etkinlik analizi gerçekleştirilen Şekil 6.11’de sunulan diğer örnek görüntü incelendiğinde de benzer etkiler görülebilmektedir. Şekil 6.12 (e)’de yer alan yakınlaştırılmış görüntüde binaların arkasında kalan gökyüzünün mavi rengi dikkate alındığında, renk dağılımının daha doğru şekilde oluşturulduğu gözlemlenebilmektedir.

### **6.3.2. Histogram Kayıp Fonksiyonu Teriminin Ağırlık Katsayısının Belirlenmesi**

Toplam kayıp fonksiyonu içerisinde yer alan histogram kaybına ait ağırlık katsayısı olan  $\alpha$  parametresinin ağırlığını belirlemek için her hangi bir dikkat modülünün kullanılmadığı, temel ağ mimarisi modeli eğitilmiştir. Öncelikle model histogram kaybı olmadan model eğitilmiş, sonrasında ise sırasıyla 0,25, 0,5 ve 0,75 gibi ağırlık katsayıları kullanılarak eğitim aşaması Tablo 5.1’de yer alan parametreler ile tekrarlanmıştır.



(a) Temel Referans Model

(b) Temel Model + ECA

(c) Temel Model + ECA + CBAM



(d) Temel Model + ECA + CBAM + IEM

(e) Temel Model + ECA + IEM  
CBAM + SimAM (Önerilen Model)

Şekil 6.9. Alt-modül etkinlik analizi görsel karşılaştırma-1 (imge 7)



(a) Temel Referans Model

(b) Temel Model + ECA

(c) Temel Model + ECA + CBAM



(d) Temel Model + ECA + CBAM + IEM

(e) Temel Model + ECA + IEM  
CBAM + SimAM (Önerilen Model)

Şekil 6.10. Alt-modül etkinlik analizi görsel karşılaştırma-1 (Şekil 6.9'a ait yakınlaştırılmış sonuçlar)



(a) Temel Referans Model

(b) Temel Model + ECA

(c) Temel Model + ECA + CBAM



(d) Temel Model + ECA + CBAM + IEM

(e) Temel Model + ECA + IEM  
CBAM + SimAM (Önerilen Model)

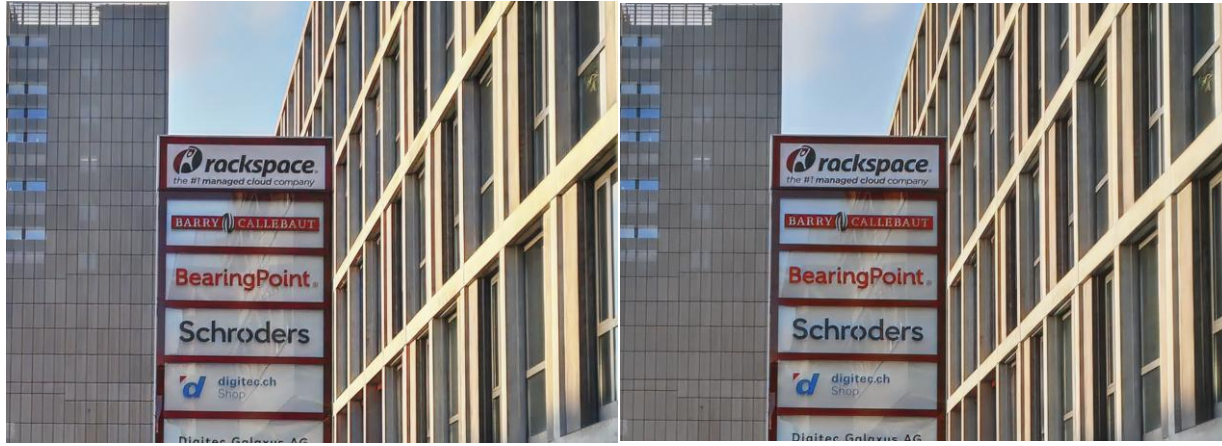
Şekil 6.11. Alt-modül etkinlik analizi görsel karşılaştırma-2 (imge 8)



(a) Temel Referans Model

(b) Temel Model + ECA

(c) Temel Model + ECA + CBAM



(d) Temel Model + ECA + CBAM + IEM

(e) Temel Model + ECA + IEM  
CBAM + SimAM (Önerilen Model)

Şekil 6.12. Alt-modül etkinlik analizi görsel karşılaştırma-2 (Şekil 6.11'e ait yakınlaştırılmış sonuçlar)



(a) Temel Referans Model

(b) Temel Model + ECA

(c) Temel Model + ECA + CBAM



(d) Temel Model + ECA + CBAM + IEM

(e) Temel Model + ECA + IEM  
CBAM + SimAM (Önerilen Model)

Şekil 6.13. Alt-modül etkinlik analizi görsel karşılaştırma-3 (imge 6)



(a) Temel Referans Model

(b) Temel Model + ECA

(c) Temel Model + ECA + CBAM



(d) Temel Model + ECA + CBAM + IEM

(e) Temel Model + ECA + IEM  
CBAM + SimAM (Önerilen Model)

Şekil 6.14. Alt-modül etkinlik analizi görsel karşılaştırma-3 (Şekil 6.13'e ait yakınlaştırılmış sonuçlar)

Tablo 6.6’da, histogram kayıp fonksiyonu ağırlık parametresini belirlemek üzere dört farklı katsayı ile eğitilen modellere ait test görüntü çıktılarının NIMA ve MUSIQ görüntü kalitesi skorları sunulmaktadır.

Tablo 6.6. Histogram kayıp fonksiyonu etkinlik analizi (NIMA/MUSIQ)

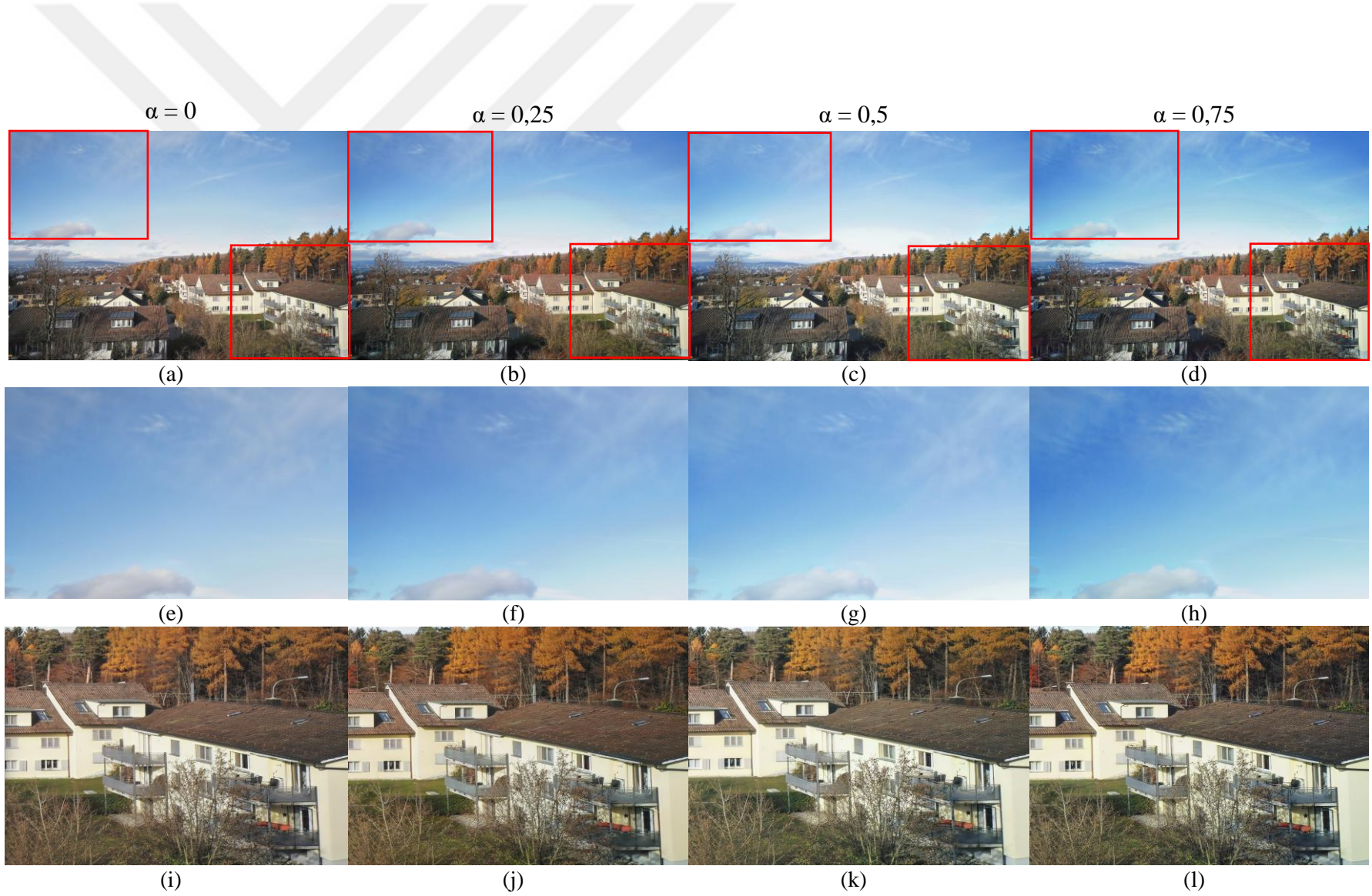
Ağırlık Görüntü No	$(\alpha = 0)$		$(\alpha = 0,25)$		$(\alpha = 0,5)$		$(\alpha = 0,75)$	
	NIMA	MUSIQ	NIMA	MUSIQ	NIMA	MUSIQ	NIMA	MUSIQ
imge 0	6,72	<b>65,56</b>	<b>6,96</b>	<u>64,47</u>	<u>6,77</u>	62,94	6,73	62,34
imge 1	<b>6,86</b>	62,63	<u>6,80</u>	<b>63,71</b>	6,75	<u>63,04</u>	6,61	62,16
imge 2	4,33	<b>61,21</b>	4,25	59,91	<u>4,38</u>	<u>60,57</u>	<b>4,61</b>	59,42
imge 3	6,48	<b>62,52</b>	<b>6,69</b>	<u>61,84</u>	6,49	60,22	<u>6,56</u>	59,78
imge 4	5,94	<b>63,50</b>	<b>6,19</b>	<u>63,49</u>	6,03	61,77	<u>6,08</u>	61,47
imge 5	6,57	<b>63,58</b>	<u>6,74</u>	62,58	6,46	<u>62,76</u>	<b>6,84</b>	61,86
imge 6	7,20	<u>62,27</u>	<u>7,31</u>	<b>62,74</b>	7,19	61,56	<b>7,41</b>	60,58
imge 7	<b>5,35</b>	<u>64,22</u>	<u>5,23</u>	<b>64,57</b>	<u>5,23</u>	63,40	5,11	62,19
imge 8	6,45	<u>61,13</u>	<u>6,65</u>	<b>61,91</b>	6,39	60,99	<b>6,66</b>	60,32
imge 9	6,72	56,43	<b>6,96</b>	56,75	<u>6,77</u>	55,11	6,73	55,34
Ortalama Skor	6,21	<b>62,31</b>	<b>6,31</b>	<u>62,20</u>	6,19	61,23	<u>6,29</u>	60,55

Tablo 6.6’da yer alan sonuçlar incelendiğinde  $\alpha = 0,25$  ile ağırlıklandırılmış Histogram Kayıp Fonksiyonu ile eğitilmiş modelin; NIMA metriği için en yüksek sonucu, MUSIQ metriği için ise en iyi ikinci skoru verdiği görülmektedir. Bununla birlikte, MUSIQ metriği için en iyi sonucun Histogram Kayıp Fonksiyonu uygulanmamış model ile alındığı görülmektedir. Elde edilen nicel sonuçlar, Şekil 6.13, Şekil 6.14 ve Şekil 6.15’te sunulan görsel sonuçlar ile birlikte değerlendirildiğinde, Histogram Kayıp Fonksiyonu ağırlık parametresi,  $\alpha$ , 0,25 olarak belirlenmiştir.

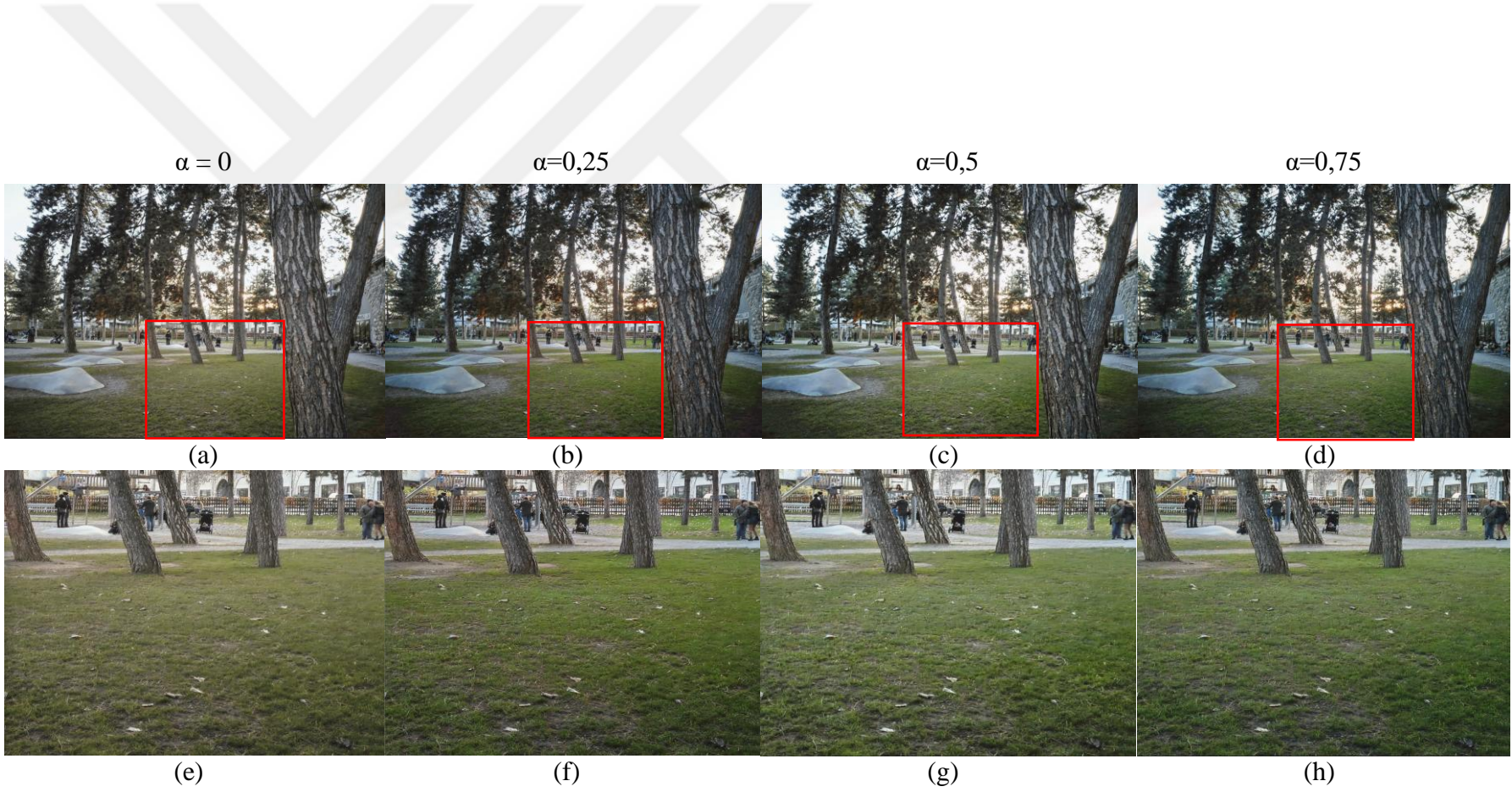
Görsel sonuçlar dikkatli bir şekilde incelendiğinde, histogram kayıp fonksiyonun kullanılmamasının olumsuz etkisi Şekil 6.14 (e)’de gökyüzne ait mavi renk bilgisinin ve Şekil 6.15 (e)’de ise yeşil renk bilgisinin doğru bir şekilde oluşturulmadığından görülebilmektedir. Öte yandan,  $\alpha = 0,75$  ağırlığının artırılmasıyla çıktı görüntülerinde bazı yapaylıklar görülebilir; örneğin Şekil 6.13 (h)’da gökyüzünde bozucu etkiler görülebilmektedir.  $\alpha = 0,5$  ile  $\alpha = 0,25$  arasındaki görsel karşılaştırma diğerleri kadar belirgin değildir; ancak Şekil 6.14 (j)’de çatıdaki kırmızı rengin daha canlı şekilde oluşturulduğu görülmektedir. Benzer şekilde Şekil 6.15 (f)’te yer alan görseldeki çimenlik bölgede de yeşil renk bilgisinin daha doğal bir şekilde oluşturulduğu görülmektedir.



Şekil 6.15. Histogram Kayıp Fonksiyonu ağırlık parametresi analizine ait görsel sonuçlar-1



Şekil 6.16. Histogram Kayıp Fonksiyonu ağırlık parametresi analizine ait görsel sonuçlar-2



Şekil 6.17. Histogram Kayıp Fonksiyonu ağırlık parametresi analizine ait görsel sonuçlar-3

Son olarak, Histogram Kayıp Fonksiyonun ağırlığının belirlenmesine yönelik analizler sonucunda oluşturulan görüntülerin dinamik aralık değerleri Tablo 6.7’de sunulmaktadır. Sonuçlar incelendiğinde  $\alpha = 0.5$  durumunda en geniş,  $\alpha = 0.25$  durumunda ise ikinci en geniş dinamik aralık değerinin elde edildiği görülmektedir.

Tablo 6.7. Histogram kayıp fonksiyonu etkinlik analizi (Dinamik Aralık, dB)

Görüntü No	( $\alpha = 0$ )	( $\alpha = 0,25$ )	( $\alpha = 0,5$ )	( $\alpha = 0,75$ )
imge 0	<u>3,42</u>	3,29	<b>3,44</b>	3,40
imge 1	3,33	<u>3,51</u>	<b>3,74</b>	3,36
imge 2	<u>2,68</u>	2,44	<b>2,77</b>	2,60
imge 3	3,71	<b>4,00</b>	<u>3,78</u>	3,90
imge 4	3,45	3,59	<u>3,85</u>	<b>3,92</b>
imge 5	4,01	<u>4,22</u>	<b>4,25</b>	4,11
imge 6	3,48	<b>3,71</b>	3,61	<u>3,65</u>
imge 7	3,21	<u>3,22</u>	<b>3,53</b>	3,15
imge 8	3,70	<b>4,13</b>	<u>4,04</u>	3,77
imge 9	3,43	<u>3,70</u>	3,58	<b>3,87</b>
Ortalama (dB)	3,44	<u>3,58</u>	<b>3,66</b>	3,54

Niceliksel ve niteliksel karşılaştırmalar dikkate alındığında histogram kayıp fonksiyonu teriminin ağırlık parametresi 0,25 olarak belirlenmiştir.

### 6.3.3. Veriseti Çözünürlüğü Etkinlik Analizi

Bölüm 4’te veriseti hazırlanması aşamasında uygulanan yöntem detaylı olarak anlatılmıştı. Bu tezde önerilen modelin eğitimi, Zürih verisetinde yer alan 168 adet yüksek çözünürlüklü görüntüden çıkartılan (448, 448)’lik görüntü parçaları ile gerçekleştirilmiştir. Bu bölümde, modelin eğitimi esnasında farklı çözünürlüklerde görüntü parçalarının kullanılmasının model başarımına etkisi incelenmiştir. Bu maksatla (448,448) çözünürlüğünde hazırlanan verisetinin dışında; (224, 224) ve (896, 896)’lık çözünürlüğe sahip RAW ve RGB hedef görüntü çiftlerinden oluşan iki ayrı veriseti daha oluşturulmuştur. Hazırlanan verisetlerinde eğitim ve test için ayrılan görüntü çifti adetleri aşağıda sunulmuştur;

- (448, 448) için eğitim 12600 adet, test için 1204 adet
- (224, 224) için eğitim 50400 adet, test için 3360 adet

- (896, 896) için eğitim 3696 adet, test için 336 adet

Çıkartılan görüntü parçalarının adetleri, yüksek çözünürlükteki bir görüntüdeki piksel sayısı ile çıkartılan görüntü parçalarındaki piksel sayılarının oranı göz önünde bulundurularak belirlenmiştir.

Hazırlanan verisetlerinin eğitimi tez kapsamında sunulan model ile eğitilmiştir. (896, 896)'lık patchler ile eğitim gerçekleştirilirken, modelde ilave bir revizyon ihtiyacı ortaya çıkmamıştır. Ancak (224, 224)'lük patchler ile eğitim gerçekleştirilirken, kayıp fonksiyonunda kullanılan VGG19 kayıp fonksiyonu teriminden kaynaklı hata kodunun alınmaması maksadıyla modelde revizyon ihtiyacı ortaya çıkmıştır. Modelin (224, 224) çözünürlükteki görüntü parçaları ile eğitilmesi esnasında Layer-4'teki parametre sayısının yetersizliğini önlemek maksadıyla, Layer-3 ve Layer-4 arasındaki Max Pooling bloğu devre dışı bırakılmıştır. Sonrasında ise modelde yer alan öznelik haritalarının boyut uyumluluğunu sağlayabilmek amacıyla Layer-4 çıkışından Layer-3 girişinde birleştirilmeden önce kullanılan Upscale bloğu da baypas edilmiştir.

Tablo 6.8. Eğitim amacıyla kullanılan veriseti çözünürlüğü etkinlik analizi (Dinamik Aralık, dB)

Model	Veriseti Çözünürlüğü (224,224)	Veriseti Çözünürlüğü (896,896)	Önerilen Çözünürlük (448, 448)
imge 0	<b>3,83</b>	3,80	3,65
imge 1	<b>4,23</b>	3,89	4,04
imge 2	2,96	<b>2,99</b>	2,91
imge 3	3,91	4,03	<b>4,37</b>
imge 4	3,75	3,66	<b>3,78</b>
imge 5	4,47	4,30	<b>4,55</b>
imge 6	<b>3,69</b>	3,51	3,52
imge 7	3,78	<b>3,86</b>	3,54
imge 8	3,84	3,86	<b>4,28</b>
imge 9	3,81	3,69	<b>3,99</b>
Ort.Skor	3,83	3,76	<b>3,85</b>

Modelin eğitilmesi hazırlanan verisetleri ile ayrı ayrı eğitilmesi sonucunda, aynı 10 adet test verisi kullanılarak elde edilen dinamik aralık değerleri Tablo 6.8'de sunulmuştur. Ortalama olarak en yüksek dinamik aralığın (448, 448)'lik yamalarla eğitilen modele ait olduğu görülmektedir.

Tablo 6.9. Eğitimde kullanılan veriseti giriş çözünürlüğüne ait etkinlik analizinin nicel sonuçları

Model	Önerilen Model				Önerilen Model				Önerilen Model			
Veriseti	(896, 896) Veriseti Çözünürlüğü				(224, 224) Veriseti Çözünürlüğü				Önerilen Veriseti Çözünürlüğü (448, 448)			
Görüntü No	NIMA	MUSIQ	LIQE	MANIQA	NIMA	MUSIQ	LIQE	MANIQA	NIMA	MUSIQ	LIQE	MANIQA
imge 0	<u>6,83</u>	64,27	3,00	0,3740	6,82	<u>65,68</u>	<b>4,02</b>	<u>0,4192</u>	<b>7,17</b>	<b>65,97</b>	<u>3,32</u>	<b>0,4199</b>
imge 1	6,79	62,58	3,57	0,3655	<u>6,95</u>	<b>65,05</b>	<b>4,63</b>	<b>0,4091</b>	<b>6,99</b>	<u>64,91</u>	<u>4,25</u>	<u>0,4090</u>
imge 2	4,33	<u>61,97</u>	3,71	0,3231	<u>4,57</u>	61,08	<b>4,50</b>	<b>0,3677</b>	<b>5,21</b>	<b>62,02</b>	<u>4,19</u>	<u>0,3673</u>
imge 3	<b>6,78</b>	60,65	3,40	0,4022	6,60	63,04	<b>3,74</b>	<u>0,4828</u>	<u>6,74</u>	<u>63,57</u>	<u>3,58</u>	<b>0,5047</b>
imge 4	<u>6,15</u>	64,40	3,45	0,4016	6,13	<b>65,99</b>	<b>4,57</b>	<u>0,4314</u>	<b>6,16</b>	<u>64,81</u>	<u>3,99</u>	<b>0,4378</b>
imge 5	<u>6,75</u>	63,25	3,14	0,3250	6,70	<b>64,06</b>	<b>3,79</b>	<u>0,3597</u>	<b>6,83</b>	<u>63,74</u>	<u>3,59</u>	<b>0,3642</b>
imge 6	<b>7,59</b>	<b>62,90</b>	3,27	0,3526	<u>7,57</u>	62,43	<b>3,93</b>	<b>0,4123</b>	7,46	<u>62,63</u>	<u>3,39</u>	<u>0,3918</u>
imge 7	5,37	64,61	3,11	0,3790	<u>5,49</u>	<b>65,82</b>	<b>4,01</b>	<u>0,3983</u>	<b>5,56</b>	<u>65,76</u>	<u>3,54</u>	<b>0,4050</b>
imge 8	<u>6,69</u>	62,12	3,43	0,3611	6,44	<u>62,88</u>	<b>4,33</b>	<u>0,4066</u>	<b>6,86</b>	<b>62,95</b>	<u>3,73</u>	<b>0,4179</b>
imge 9	5,96	<u>58,13</u>	2,92	0,3256	<u>6,42</u>	<b>59,41</b>	<u>3,17</u>	<u>0,3542</u>	<b>6,43</b>	58,12	3,03	<b>0,3563</b>
Ortalama Skor	6,32	62,48	3,3	0,361	<u>6,37</u>	<b>63,54</b>	<b>4,07</b>	<u>0,4041</u>	<b>6,54</b>	<u>63,45</u>	<u>3,67</u>	<b>0,4073</b>



(a) (896, 896) çözünürlüklü veriseti ile eğitilen model çıktısı

(b) (224, 224) çözünürlüklü veriseti ile eğitilen model çıktısı

(c) Önerilen Model Çıktısı (448, 448)



(d) Çıktı (a)'ya ait yakınlaştırılmış görüntü

(e) Çıktı (b)'ye ait yakınlaştırılmış görüntü

(f) Çıktı (c)'ye ait yakınlaştırılmış görüntü

Şekil 6.18. Eğitimde kullanılan veriseti giriş çözünürlüğüne ait etkinlik analizinin nitel sonuçları-1 (imge 6)



(a) (896, 896) çözünürlüklü veriseti ile eğitilen model çıktısı

(b) (224, 224) çözünürlüklü veriseti ile eğitilen model çıktısı

(c) Önerilen Model Çıktısı (448, 448)



(d) Çıktı (a)'ya ait yakınlaştırılmış görüntü

(e) Çıktı (b)'ye ait yakınlaştırılmış görüntü

(f) Çıktı (c)'ye ait yakınlaştırılmış görüntü

Şekil 6.19. Eğitimde kullanılan veriseti giriş çözünürlüğüne ait etkinlik analizinin nitel sonuçları-2 (imge 1)

Ayrıca NIMA, MUSIQ, LIQE ve MANIQA metriklerine ait nicel sonuçlar Tablo 6.9’da paylaşılmıştır. Sonuçlar incelendiğinde (448, 448)’lik görüntü yamaları ile eğitilen modelin NIMA ve MANIQA metrikleri ile daha yüksek kalitede çıktılar üretebildiği görülmektedir. MUSIQ ve LIQE metriklerinde ise en iyi sonucun (224, 224)’lik veriseti ile alındığı; ikinci en iyi sonucun ise (448, 448)’lik verisetine ait olduğu görülmektedir.

Ancak test görüntülerine ait çıktılar görsel olarak incelendiğinde; (224, 224) ve (896, 896) çözünürlüğündeki görüntü yamaları ile eğitilen modellerde bazı bozucu etkilerin olduğu tespit edilmiştir. Örneğin Şekil 6.18 (a) ve (b)’de yer alan çıktılar ve bu görüntülere karşılık gelen yakınlaştırılmış görseller incelendiğinde gökyüzünde halo etkisine benzer bozucu etkilerin olduğu görülmektedir. Benzer bozucu etkiler Şekil 6.19’da da rahatlıkla görülebilmektedir. Genel olarak görsel çıktılar karşılaştırıldığında (224, 224) çözünürlüğe sahip veriseti ile eğitilen modele ait çıktılarda bu etkinin daha yaygın olduğu görülmektedir.

#### **6.3.4. Kayıp Fonksiyonu Terimlerinin Öğrenilebilir Parametreler ile Ağırlıklandırılması**

Eğitim esnasında her bir katmanda kullanılan kayıp fonksiyonlarının ağırlık katsayıları Bölüm 5.2’de sunulmuştur. Sonrasında ise Histogram Kayıp Fonksiyonun ağırlık katsayısının belirlenmesine yönelik deneyler ve deney sonuçları 6.3.2’de sunulmuştur. Bu başlık altında ise her bir katmanın eğitimi esnasında kullanılan ağırlık katsayılarının öğrenilebilir parametreler olarak tanımlanması durumunda elde edilen sonuçlar ve gözlemler paylaşılacaktır. Konunun irdelenmesi maksadıyla öğrenilebilir parametre olarak tanımlanan kayıp fonksiyonu terimlerinin ağırlıklarının ilklendirme değerlerine bağlı olarak iki ayrı deney yapılmıştır;

- Deney-1: Kayıp fonksiyonu ağırlık parametrelerinin Tablo 5.1’de sunulan katsayılar ile ilklendirilmesi
- Deney-2: Kayıp fonksiyonu ağırlık parametrelerinin eşit ağırlık katsayıları ile ilklendirilmesi

İklendirilen ağırlık parametreleri, ADAM optimizasyon algoritmasına dahil edilerek model mimarisi parametreleri ile birlikte eğitilmesi sağlanmıştır. Ağırlık katsayılarını

belirten parametreler ve hangi ilgili kayıp fonksiyonu terimine karşılık geldiği aşağıda listelenmiştir;

- a: Ortalama Karesel Hata kayıp fonksiyonu ağırlık katsayısı
- b: Histogram kayıp fonksiyonu ağırlık katsayısı
- c: VGG tabanlı kayıp fonksiyonu ağırlık katsayısı
- d: SSIM kayıp fonksiyonu ağırlık katsayısı

Deney-1 kapsamında, her bir katmandaki kayıp fonksiyonu ağırlık parametrelerine ait ilklendirilen değerler, eğitim sonunda nihai parametre değerleri ve eğitilen toplam tur sayısı Tablo 6.10'da sunulmuştur. Deney esnasında, Layer-1 ve Layer-0 katmanlarında eğitim kararsız sonuçlar ürettiği ve siyah çıktılar verdiği görülmüştür. Gözlemlenen bu durumun ağırlık parametrelerinin negatif değerlere gitmesiyle ilişkili olduğu değerlendirilmektedir. Bu durumla karşılaşıldığında eğitim kararsız duruma gitmeden önceki model parametreleri geri yüklenerek, ilklendirilen ağırlık parametreleri ile eğitim tekrardan başlatılmıştır. Bu nedenle, özellikle Layer-0 için ağırlık parametrelerinin başlangıç değerlerinden fazla değişmediği görülmektedir.

Tablo 6.10. Öğrenilebilir kayıp fonksiyonlarına ait ağırlık parametreleri (Deney-1)

Katman	İklendirilen Ağırlık Katsayıları	Eğitim Sonu Ağırlık Katsayıları	Epoch
Layer-4	a = 1,0 b = 0,25	a = 0,8369 b = 0,0280	19
Layer-3	a = 10,0 b = 0,25 c = 1,0	a = 9,6807 b = -0,1659 c = 0,6443	24
Layer-2	a = 10,0 b = 0,25 c = 1,0	a = 9,5144 b = -0,3565 c = 0,4619	24
Layer-1	a = 10,0 b = 0,25 c = 1,0 d = 0,4	a = 9,6520 b = -0,1291 c = 0,6365 d = 0,0504	24
Layer-0	a = 10,0 b = 0,25 c = 1,0 d = 0,4	a = 9,8239 b = 0,0676 c = 0,8167 d = 0,2203	10

Deney-1 sürecinde ayrıca eğitim sürecinin kararlı hale getirilmesini sağlamak amacıyla L1 ve L2 düzenleme terimlerinin kullanımı denenmiştir. Bu durumda eğitim

sürecinde gözlemlenen kararsızlığın giderildiği, ancak alınan sonuçların yeterli kalitede olmadığı, sahneye ait renk bilgisinin doğru oluşturulamadığı gözlemlenmiştir.

Deney-2 kapsamında, ağırlık parametreleri eşit değerler ile başlatılmıştır. Bu durumda her bir katmandaki kayıp fonksiyonu ağırlık parametrelerine ait ilklendirilen değerler, eğitim sonunda nihai parametre değerleri ve eğitilen toplam tur sayısı Tablo 6.11’de sunulmuştur. Eğitim esnasında Layer-1’de eğitimin kararsızlaşması sorunu tekrardan görülmüş ve bu durumda son görülen ağırlık parametreleri ile kalan tur sayısından eğitim yeniden başlatılmıştır. İlave olarak optimizasyon adım boyu  $1 \times 10^{-5}$ ’e düşürülmüştür. Layer-0’da ise parametreler  $1 \times 10^{-5}$  adım boyu ile optimize edilmiş, bu nedenle tur sayısı uzatılmıştır.

Tablo 6.11. Öğrenilebilir kayıp fonksiyonlarına ait ağırlık parametreleri (Deney-2)

Katman	İklendirilen Ağırlık Katsayıları	Eğitim Sonu Ağırlık Katsayıları	Epoch
Layer-4	a = 1,0 b = 1,0	a = 0,8380 b = 0,8114	17
Layer-3	a = 1,0 b = 1,0 c = 1,0	a = 0,7308 b = 0,6948 c = 0,7422	18
Layer-2	a = 1,0 b = 1,0 c = 1,0	a = 0,3658 b = 0,3088 c = 0,3501	29
Layer-1	a = 1,0 b = 1,0 c = 1,0 d = 1,0	a = 0,1480 b = 0,0785 c = 0,0278 d = 0,0918	33
Layer-0	a = 1,0 b = 1,0 c = 1,0 d = 1,0	a = 0,1022 b = 0,0455 c = 0,0488 d = 0,795	80

Deney-1 ve Deney-2’ye ait NIMA, MUSIQ, LIQE ve MANIQA metriklerine ait nicel sonuçlar Tablo 6.12’de sunulmuştur. Sonuçlar incelendiğinde tez kapsamında önerilen kayıp fonksiyonu ağırlık parametreleri ile gözlemlenen performansın üstüne çıkılamadığı görülmektedir. Kullanılan referansa ihtiyaç duymayan görüntü kalitesi değerlendirme metriklerine göre sunulan modelin daha yüksek kalitede sonuçlar üretebildiği; Deney-1’de mevcut ağırlık parametreleri ile ilklendirilerek eğitilen model parametrelerin ise Deney-2’ye kıyasla daha iyi sonuçlar üretebildiği anlaşılmaktadır.

Tablo 6.12. Öğrenilebilir kayıp fonksiyonlarına ait ağırlık parametrelerine ait deneysel sonuçlar

Model	Önerilen Model				Önerilen Model				Önerilen Model			
Veriseti	Deney-1 Sonuçları				Deney-2 Sonuçları				Tablo 5.1'de Sunulan Sabit Ağırlık Katsayıları			
Görüntü No	NIMA	MUSIQ	LIQE	MANIQA	NIMA	MUSIQ	LIQE	MANIQA	NIMA	MUSIQ	LIQE	MANIQA
imge 0	<u>6,97</u>	<u>63,0</u>	<u>2,75</u>	<u>0,3433</u>	6,95	62,51	2,53	0,3359	<b>7,17</b>	<b>65,97</b>	<b>3,32</b>	<b>0,4199</b>
imge 1	<u>7,06</u>	60,70	<u>3,27</u>	<u>0,3580</u>	<b>7,14</b>	<u>61,68</u>	3,24	0,3334	6,99	<b>64,91</b>	<b>4,25</b>	<b>0,4090</b>
imge 2	4,84	60,41	<u>3,45</u>	<u>0,3042</u>	<u>4,94</u>	<u>60,69</u>	3,04	0,2990	<b>5,21</b>	<b>62,02</b>	<b>4,19</b>	<b>0,3673</b>
imge 3	6,69	<u>60,05</u>	<u>3,07</u>	<u>0,3914</u>	<u>6,70</u>	59,82	3,01	0,3812	<b>6,74</b>	<b>63,57</b>	<b>3,58</b>	<b>0,5047</b>
imge 4	5,75	<u>62,15</u>	<u>3,03</u>	<u>0,3688</u>	<u>5,86</u>	61,70	2,95	0,3520	<b>6,16</b>	<b>64,81</b>	<b>3,99</b>	<b>0,4378</b>
imge 5	6,68	61,98	<u>2,88</u>	<u>0,3136</u>	<u>6,82</u>	<u>62,49</u>	2,46	0,3001	<b>6,83</b>	<b>63,74</b>	<b>3,59</b>	<b>0,3642</b>
imge 6	<b>7,71</b>	<b>63,75</b>	<u>2,72</u>	<u>0,3168</u>	<u>7,52</u>	62,43	2,67	0,3160	7,46	<u>62,63</u>	<b>3,39</b>	<b>0,3918</b>
imge 7	<u>5,51</u>	<u>64,91</u>	<u>2,78</u>	<u>0,3570</u>	5,35	63,34	2,55	0,3346	<b>5,56</b>	<b>65,76</b>	<b>3,54</b>	<b>0,4050</b>
imge 8	6,78	<u>62</u>	<u>3,15</u>	<u>0,3693</u>	<b>6,96</b>	61,59	2,90	0,3346	<u>6,86</u>	<b>62,95</b>	<b>3,73</b>	<b>0,4179</b>
imge 9	<u>6,35</u>	54,79	<u>2,62</u>	<u>0,3092</u>	6,30	<u>55,98</u>	2,48	0,2913	<b>6,43</b>	<b>58,12</b>	<b>3,03</b>	<b>0,3563</b>
Ortalama Skor	6,43	<u>61,38</u>	<u>2,97</u>	<u>0,3432</u>	<u>6,45</u>	61,22	2,79	0,3278	<b>6,54</b>	<b>63,45</b>	<b>3,67</b>	<b>0,4073</b>



(a) Deney-1 eğitim çıktısı

(b) Deney-2 eğitim çıktısı

(c) Önerilen Model Çıktısı



(d) Çıktı (a)'ya ait  
yakınlaştırılmış görüntü

(e) Çıktı (b)'ye ait  
yakınlaştırılmış görüntü

(f) Çıktı (c)'ye ait  
yakınlaştırılmış görüntü

Şekil 6.20. Öğrenilebilir kayıp fonksiyonlarına ait ağırlık parametrelerine ait deneysel sonuçlar-1 (imge 2)



(a) Deney-1 eğitim çıktısı

(b) Deney-2 eğitim çıktısı

(c) Önerilen Model Çıktısı



(d) Çıktı (a)'ya ait  
yakınlaştırılmış görüntü

(e) Çıktı (b)'ye ait  
yakınlaştırılmış görüntü

(f) Çıktı (c)'ye ait  
yakınlaştırılmış görüntü

Şekil 6.21. Öğrenilebilir kayıp fonksiyonlarına ait ağırlık parametrelerine ait deneysel sonuçlar-2 (imge 7)

Deney-1 ve Deney-2 kapsamında eğitilen modellere ait çıktıların görsel olarak karşılaştırması ise Şekil 6.20 ve Şekil 6.21’de sunulmuştur. Deney-1’e ait çıktılardan Şekil 6.20 (a) incelendiğinde gökyüzünde bozucu etkilerin olduğu görülmektedir. Aynı zamanda Şekil 6.20 (a)’da sahnenin sağ alt tarafında yer alan duvarda ton dönüşümü açısından yeterli iyileştirmenin sağlanamadığı görülmektedir. Aynı sahneye ait Şekil 6.20 (f) çıktısı incelendiğinde ton dönüşümün başarılı şekilde uygulanabildiği görülmektedir. Bununla birlikte Deney-2’ye ait Şekil 6.20 (e) ile sunulan çıktı incelendiğinde aynı bölgede renklerde bozucu etkilerin olduğu görülmektedir. Benzer şekilde, Deney-2 ile elde edilen görsel sonuçlardan Şekil 6.21 (e)’de siyah aracın bulunduğu zeminde mavi renkte oluşan bozulmaların meydana geldiği görülmektedir.

### **6.3.5. Veri Çoğaltma İşlemi Etkinlik Analizi**

Verisetinin hazırlama aşamasında kullanılan yöntemler Bölüm 4’te bahsedilmişti. Bu başlık altında ise çeşitli veriseti çoğaltma tekniklerinin kullanılmasıyla oluşturulan alternatif verisetinin model performansına etkisi incelenecektir.

Modelin eğitimi esnasında kullanılan ham veri, “RGBG” Bayer paternine sahip olması nedeniyle dört kanallı bir veridir. Bu nedenle Pytorch kütüphanesinde yer alan veri çoğaltma tekniklerinin doğrudan kullanımı, kanal diziliminde hatalara ve hedef görüntüler ile uyumsuzluğa neden olmaktadır. Aynı nedenle Rasgele Döndürme veri çoğaltma tekniği kullanılamamaktadır. Veri çoğaltma tekniklerinden Yatay ve Dikey Ters Döndürme (HorizontalFlip ve VerticalFlip) yöntemlerinin ise ham verideki kanal dizilimlerini göz önünde bulundurularak uygulanmasının mümkün olduğu görülmüştür.

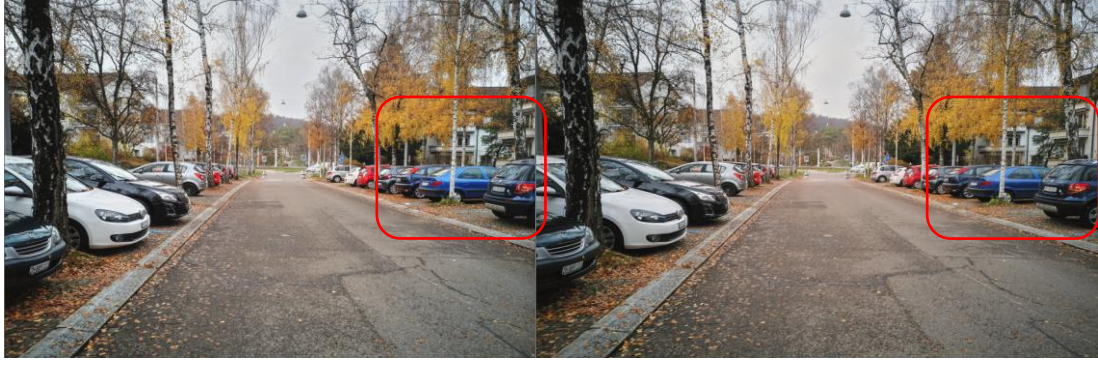
Rastgele kırpma işlemi uygulanan (448, 448) görüntü yamalarına Yatay ve Dikey Ters Döndürme işlemi uygulandığında “RGBG” kanal dizilimi bozulmaktadır. Bu nedenle kanal diziliminin bozulmamasını sağlamak amacıyla, (449, 449) görüntü yamaları çıkartılarak yatay ve dikey ters döndürme işlemi uygulanmıştır. Sonrasında ise görüntü yamalarının (448, 448) çözünürlükte olmasını sağlamak amacıyla yükseklik ve genişlik uzamsal düzleminden birer satır piksel çıkartılmıştır. Bu yöntem ile ilave olarak 12600 adet yatay ters döndürülmüş veri, 12600 de dikey ters döndürülmüş görüntü çiftleri oluşturulmuş; böylelikle toplamda 37800 adet eğitim veriseti oluşturulmuştur.

Veri çoğaltma yöntemleri ile hazırlanan genişletilmiş veriseti kullanılarak önerilen modelin eğitilmesi sonucunda elde edilen nicel sonuçlar Tablo 6.13'te sunulmuştur. Elde edilen sonuçlar Bölüm 6.1'de sunulan önerilen modele ait sonuçlar ile aynı test verisi kullanılarak karşılaştırıldığında; yeni veriseti ile MUSIQ, LIQE ve MANIQA metriklerinde daha yüksek skora sahip çıktılar elde edilebildiği görülmektedir. NIMA metriğinde ise veri çoğaltma işlemi sonunda daha düşük kaliteli çıktılar elde edilebildiği anlaşılmaktadır. Bu durumun NIMA metriğinin giriş görüntüsünü yeniden ölçeklendirmesi ile ilgili olabileceği düşünülmektedir.

Tablo 6.13. Veri seti çoğaltma işlemi etkinlik analizine ait nicel sonuçlar

Model	Önerilen Model				Önerilen Model			
Veriseti	Veri Seti Çoğaltma Teknikleri ile Hazırlanan Veri Seti (37800 adet)				Hazırlanmış Veriseti (12600 adet)			
Görüntü No	NIMA	MUSIQ	LIQE	MANIQA	NIMA	MUSIQ	LIQE	MANIQA
imge 0	7,01	<b>66,86</b>	<b>3,82</b>	<b>0,4397</b>	<b>7,17</b>	65,97	3,32	0,4199
imge 1	6,80	<b>65,91</b>	<b>4,67</b>	<b>0,4194</b>	<b>6,99</b>	64,91	4,25	0,4090
imge 2	4,29	60,92	<b>4,35</b>	0,3615	<b>5,21</b>	<b>62,02</b>	4,19	<b>0,3673</b>
imge 3	<b>6,87</b>	<b>63,86</b>	<b>4,20</b>	<b>0,5366</b>	6,74	63,57	3,58	0,5047
imge 4	<b>6,27</b>	<b>66,22</b>	<b>4,43</b>	0,4219	6,16	64,81	3,99	<b>0,4378</b>
imge 5	6,73	63,47	<b>3,96</b>	<b>0,3666</b>	<b>6,83</b>	<b>63,74</b>	3,59	0,3642
imge 6	7,26	61,35	<b>3,78</b>	<b>0,4248</b>	<b>7,46</b>	<b>62,63</b>	3,39	0,3918
imge 7	5,42	65,76	<b>3,82</b>	0,3885	<b>5,56</b>	65,76	3,54	<b>0,4050</b>
imge 8	6,71	<b>63,16</b>	<b>4,09</b>	<b>0,4296</b>	<b>6,86</b>	62,95	3,73	0,4179
imge 9	6,26	<b>58,95</b>	<b>3,37</b>	<b>0,3993</b>	<b>6,43</b>	58,12	3,03	0,3563
Ortalama Skor	6,36	<b>63,65</b>	<b>4,05</b>	<b>0,4188</b>	<b>6,54</b>	63,45	3,67	0,4073

Şekil 6.22'de ise her iki veri seti ile önerilen modelin eğitilmesi sonucunda elde edilen test görüntülerine ait çıktılar sunulmuştur. Çıktı görüntüler incelendiğinde her iki modelin gürültü ve görüntünün genel ton dönüşümü açısından benzer çıktılar oluşturduğu görülmektedir. Ancak yerel bölgelere odaklanıldığında veri çoğaltma işleminin olumlu etkisi görülebilmektedir.



(a) Veri çoğaltma işlemi ile hazırlanan veriseti (imge 5)

(b) Veri çoğaltma işlemi olmadan hazırlanan veriseti (imge 5)



(c) Görüntü (a) (Yakınlaştırılmış)

(d) Görüntü (b) (Yakınlaştırılmış)



(e) Veri çoğaltma işlemi ile hazırlanan veriseti ile eğitim çıktısı (imge 3)

(f) Veri çoğaltma işlemi olmadan hazırlanan veriseti ile eğitim çıktısı (imge 3)



(g) Görüntü (e) (Yakınlaştırılmış)

(h) Görüntü (f) (Yakınlaştırılmış)

Şekil 6.22. Veri çoğaltma işlemi etkinlik analizi görsel sonuçlar

Örneğin, Şekil 6.22 (c) ve (d) karşılaştırıldığında açık mavi arabanın daha parlak olarak oluşturulabildiği, sahnenin sağında yer alan koyu mavi araç üzerindeki yansımaların da daha doğru şekilde oluşturulabildiği görülmektedir.

Ancak bir diğer örnekte, imge 3'e ait çıktıda, veri çoğaltma işlemi sonucunda eğitilen modelin gökyüzünde bozucu etkilerin oluşmasına neden olduğu gözlemlenmiştir. Söz konusu etkinin eğitim sürecindeki rastlantısallıktan kaynaklabileceği düşünülmektedir. Ayrıca artan veriseti ile birlikte eğitim sürecinde seçilen hiperparametrelerin tekrardan gözden geçirilmesiyle ileriki çalışmalarda bu bozucu etkinin giderilebileceği değerlendirilmektedir.



## 7. SONUÇLAR VE ÖNERİLER

Bu bölümde tezde anlatılan çalışmaların sonucu değerlendirilecektir. Sonrasında ise elde edilen sonuçların gelecekte nasıl daha da geliştirilebileceği ve literatüre katkı sağlayabilecek olası çalışmalar önerilecektir.

### 7.1. Sonuçların Değerlendirilmesi

Bu tez kapsamında tasarlanan kamera ISP veri yolu ile 10 bitlik Bayer paterne sahip ham veriden, kullanıcıya sunulmak üzere 8 bitlik RGB çıkış görüntüsünün elde edilmiştir. Geliştirilen modelin başarımı, bu tez kapsamında referans olarak belirlenen PyNET (Ignatov ve diğ., 2020) ve Lightweight PyNET (Cheng ve diğ., 2021) modelleri ile karşılaştırılmıştır. NR-IQA kalite metriklerine göre (MUSIQ, NIMA ve LIQE) model başarımının artırıldığı ve daha geniş dinamik aralığa sahip görüntülerin elde edilebildiği görülmüştür. Elde edilen sayısal sonuçlar ayrıca nitel değerlendirmeler ile desteklenmiştir.

Ayrıca geliştirilen modelin başarımı, ticari olarak kullanılmakta olan Huawei P20 ve Canon 5D Mark IV DSLR kamera ISP veri yolu performansları ile karşılaştırılmış ve kullanıcıya sunulmak üzere daha yüksek kaliteli görüntülerin oluşturulabildiği nitel ve nicel sonuçlar ile ortaya konulmuştur.

İcra edilen deneyler kapsamında veri seti hazırlama metodolojisinin etkisi ortaya koyulmuş, ayrıca çeşitli etkinlik analizleri ile model mimarisindeki dikkat modüllerinin model başarımına etkisi araştırılmıştır. Histogram bilgisi üzerinden tanımlanan kayıp fonksiyonu, toplam kayıp fonksiyonu içerisindeki etkisi gözlemlenmiş ve ağırlık parametresi belirlenmiştir. Bununla birlikte eğitimde kullanılan verisetindeki görüntü çiftlerinin çözünürlüğünün model başarımına etkisi irdelenmiş ve kayıp fonksiyonunda yer alan ağırlık parametreleri katsayıları öğrenilebilir parametre olarak tanımlanarak model başarımına etkisi araştırılmıştır.

Hem literatürdeki hem de bu tezdeki elde edilen deneysel sonuçlar değerlendirildiğinde, geleneksel kamera ISP veri yoluna kıyasla derin öğrenme yöntemleri ile geliştirilen modelin teknolojik olgunluk seviyesinin oldukça yüksek olduğu anlaşılmaktadır. Bu

tezde, sahneye ait görüntünün HDR görüntünün ton-dönüşümü yaparak kullanıcıya sunulması hedeflenerek kamera ISP veri yolunun tasarımı yapılmıştır.

## 7.2. Gelecek Dönem Çalışmaları

Geleneksel kamera ISP veri yolu yerine, derin öğrenme mimarisi aracılığıyla geliştirilen bir model kullanılarak nihai görüntüyü oluşturmanın en önemli kısıt, uygun veri kümesinin oluşturularak modelin eğitilmesidir. Kullanılan ham verinin bit derinliği ve Bayer patterni gibi temel karakteristiklerinin yanında veri setinin mümkün olduğunca farklı ortam ışıklılığına sahip sahnelerden toplanmış olması model başarımına doğrudan etki etmektedir. Bu çalışmada, Zurich veri kümesinin yalnızca küçük bir alt kümesi kullanılabilmiştir. PyNET'te eğitim için 46800 adet yama (patch) görüntü kullanılırken, bu tez kapsamında sunulan modelde yalnızca 12600 yama (patch) görüntü kullanılabilmiştir. Bu tez kapsamında hazırlanan verisetinin, model başarımını artırdığı görülmüştür. Ancak verisetinin türetildiği Zurich verisetinde yer alan yüksek çözünürlüklü görüntüler incelendiğinde, sahne ışıklılığı açısından yeterli çeşitliliğin bulunmadığı anlaşılmaktadır. Özellikle yüksek ve düşük pozlama bölgelerdeki detaylarında çıkartılarak ton dönüştürme işlemini de gerçekleştirebilen bir kamera ISP veri yolu modelinin, bu amaçla oluşturulmuş bir veri seti aracılığıyla eğitilmesinin model performansı daha da artıracığı öngörülmektedir.

Bu nedenle gelecekte, ağın yerel ve küresel ton eşleme yeteneğini geliştirmek için özellikle yüksek ve düşük pozlama bölgelerini içeren ham ve ton dönüştürülmüş veri çiftlerinin toplanması hedeflenmektedir. Zurich veriseti, ton eşlemeyle ilgili herhangi bir kriteri dikkate alınarak toplanmadığından, bu tez kapsamında önerilen modelin sınır performansı görülememektedir. İleride bu tezin ve sunulan makalenin (Erbaş ve Çelebi, 2024) çıktıları ile gerekli kaynakların bulunması halinde uygun verisetinin literatüre kazandırılması amacıyla çalışma yapılacaktır.

İleriki dönemde, oluşturulan görüntülerin kalitesini daha da artırabilmek amacıyla dikkat modüllerindeki yeni araştırmalar göz önünde bulundurularak model mimarisinin geliştirilmesi hedeflenmektedir. Bu esnada Bölüm 6.5'te belirtilen veri çoğaltma tekniklerinin model performansına olumlu etkisi de göz önünde bulundurulacak, gözlemlenen kısmi bozucu etkiler de giderilmeye çalışılacaktır.

Bu tezin nihai sonucu kameraya entegre edilebilmeye uygun kamera ISP veri yolunun geliştirilmesidir. Kullanıcının fotoğraf çekme sürecini etkileyen kamera donanım bileşenleri (lens, renk filtresi dizisi, diyafram vb.) dikkate alındığında, tasarlanan kamera ISP veri yolu mimarisinin kameraya özgü veri kümesiyle eğitilmesi veya en azından ince ayarlanması (fine-tuning) ihtiyacı olacaktır. Bu nedenle gelecekte geliştirilen modelin uygun veriseti ile ince ayar yapılarak nihayetlendirilmesi ve kameraya entegre edilmesi hedeflenmektedir. Bu tezde sunulan çalışmanın derin öğrenmeye dayalı kamera ISP arařtırmalarına deęer katabileceęini umuyorum.



## KAYNAKLAR

- Afifi, M., Brubaker, M.A., Brown, M.S. (2021). HistoGAN: Controlling Colors of GAN-Generated and Real Images via Color Histograms. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2021)*, Nashville, Amerika Birleşik Devletleri, 19-25 Haziran 2021.
- Baek, Y., Cho, D., Lee, J., Kim, W. (2008). Noise Reduction for Image Signal Processor in Digital Cameras. *International Conference on Convergence and Hybrid Information Technology (ICHIT)*, Daejeon, Kore Cumhuriyeti, 28-30 Ağustos 2008.
- Barron, J.T. (2015). Convolutional Color Constancy. *IEEE International Conference on Computer Vision (ICCV)*, Massachusetts, Amerika Birleşik Devletleri, 07-13 Aralık 2015.
- Bayer, B.E. (1976). Color Image Array, U.S. Patent No. 3 971 065., *Eastman Kodak Co.*
- Bilcu, R.C., Alenius, S., Vehvilainen, M. (2012). Adaptive Local Tone Mapping of Color Images. *20th European Signal Processing Conference (EUSIPCO)*, Bükreş, Romanya, 27-31 Ağustos 2012.
- Bosse, S., Maniry, D., Muller, K., Wiegand, T., Samek, W. (2018). Deep Neural Networks for No-Reference and Full-Reference Image Quality Assessment. *IEEE Transactions on Image Processing*, 27(1), 206-219.
- Chen, C., Chen, Q., Xu, J., Koltun, V. (2018). Learning to See in the Dark. *Conference on Computer Vision and Pattern Recognition (CVPR 2018)*, Salt Lake City, Amerika Birleşik Devletleri, 19-21 Haziran 2018.
- Chen, G., Zhang, L., Sun, M., Gao, Y., Michelini, P.N., Wu, Y. (2021). Single-Image HDR Reconstruction with Task-specific Network Based on Channel Adaptive RDN. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2021)*, Sanal Konferans, 19-25 Haziran 2021.
- Chen, H., Ma, K. (2022). LW-ISP: A Lightweight Model with ISP and Deep Learning. *Image and Video Processing British Machine Vision Conference (BMVC 2022)*, Londra, Birleşik Krallık, 21-24 Kasım 2022.
- Chen, X., Liu, Y., Zhang, Z., Qiao, Y., Dong, C. (2021). HDRUNet: Single Image HDR Reconstruction with Denoising and Dequantization. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2021)*, Sanal Konferans, 19-25 Haziran 2021.
- Cheng, Y., Yue, H. and Mao, Y. (2021). A Lightweight Convolutional Neural Network for Camera ISP. *IEEE 21st International Conference on Communication Technology (ICCT)*, Tianjin, China, 13-16 Ekim 2021.

- Dai, L., Liu, X., Li, C., Chen, J. (2020). AWWNet: Attentive Wavelet Network for Image ISP. *The European Conference on Computer Vision (ECCV 2020)*, Sanal Konferans, 23-28 Ağustos 2020.
- Dalal, D., Vashishtha, G., Singh, P., Raman, S. (2023). Single Image LDR to HDR Conversion Using Conditional Diffusion. *IEEE International Conference on Image Processing (ICIP)*, Kuala Lumpur, Malezya, 08-11 Ekim 2023.
- Debevec, P.E., Malik, J. (1997). Recovering High Dynamic Range Radiance Maps from Photographs. *Computer Graphics and Interactive Techniques Conference (SIGGRAPH 97)*, Los Angeles, Amerika Birleşik Devletleri, 03-08 Ağustos 1997.
- Durand, F., Dorsey, J. (2002). Fast Bilateral Filtering for the Display of High-Dynamic-Range Images. *Association for Computing Machinery (ACM) Transactions on Graphics (TOG)*, 21(3), 257-266.
- Eilertsen, G., Kronander, J., Denes, G., Mantiuk, R.K., Unger, J. (2017). HDRCNN: HDR Image Reconstruction from a Single Exposure Using Deep CNNs. *Association for Computing Machinery (ACM) Transactions on Graphics*, 36(6), 1-15.
- Endo, Y., Kanamori, Y., Mitani, J. (2017). DrTMO-Deep Reverse Tone Mapping. *Association for Computing Machinery (ACM) Transactions on Graphics (SIGGRAPH ASIA 2017)*, Bangkok, Tayland, 27-30 Kasım 2017.
- Erbaş, K.U., Çelebi, A.T. (2023). Modeling Camera ISP Pipeline with Deep Learning. *Sinyal İşleme ve İletişim Uygulamaları Konferansı (SIU 2023)*, İstanbul, Türkiye, 05-08 Temmuz 2023.
- Erbaş, K.U., Çelebi, A.T. (2024). RAW to Tonemapped HDR Camera ISP. *Signal, Image and Video Processing (SIViP)*, 18(12), 9037-9049.
- Fang, Y., Zhu, H., Zeng, Y., Ma, K., Wang, Z. (2020). Perceptual Quality Assessment of Smartphone Photography. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Sanal Konferans, 14-19 Haziran 2020.
- Granados, M., Ajdin, B., Wand, M., Theobalt, C., Seidel, H.P., Lensch, H.P.A. (2010). Optimal HDR Reconstruction with Linear Digital Cameras. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2010)*, San Francisco, Amerika Birleşik Devletleri, 13-18 Haziran 2010.
- Gu, J., Cai H., Dong C., Ren, J., Timofte, R. (2022). NTIRE 2022 Challenge on Perceptual Image Quality Assessment. *Computer Vision and Pattern Recognition Workshop (CVPR 2022)*, New Orleans, Amerika Birleşik Devletleri, 19-24 Haziran 2022.

- Guo, B.C., Lin, C.H. (2023). Single Image HDR Reconstruction Based on Two-Stage GAN Structure. *IEEE International Conference on Image Processing (ICIP 2023)*, Kuala Lumpur, Malezya, 08-11 Ekim 2023.
- Gupta, S., Srivastava, D., Chaturverdi, U., Jain, A., Khandelwal, G. (2021). *Del-Net: A Single Stage Network for Mobile Camera ISP*. Samsung R&D Institute, India, <https://arxiv.org/abs/2108.01623>, (Ziyaret Tarihi: 10 Haziran 2024).
- Hirakawa, K., Parks, T.W. (2005). Adaptive Homogeneity-Directed Demosaicing Algorithm. *IEEE Transactions on Image Processing*. 14(3), 360-369.
- Hosu, V., Lin, H., Sziranyi, T., Saupe, D. (2020). Koniq-10k: An Ecologically Valid Database for Deep Learning of Blind Image Quality Assessment. *IEEE Transactions on Image Processing*, 29, 4041-4056.
- Hou, X., Duan, J., Qiu, G. (2018). *Deep Feature Consistent Deep Image Transformations: Downscaling, Decolorization and HDR Tone Mapping*. Nottingham, Birleşik Krallık, <https://arxiv.org/abs/1707.09482>, (Ziyaret Tarihi: 14 Haziran 2024).
- Hsu, W., Chang, H. (2021). A Novel Automatic White Balance Method for Color Constancy Under Different Color Temperatures. *IEEE Access*, 9, 111925-111937.
- Hsyu, M., Liu, C., Chen, C., Chen, C. (2021). CSANet: High Speed Channel Spatial Attention Network for Mobile ISP. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2021)*, Nashville, Amerika Birleşik Devletleri, 19-25 Haziran 2021.
- Hu, Y., Wang, B., Lin, S. (2017). FC4: Fully Convolutional Color Constancy with Confidence-Weighted Pooling. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017)*, Honolulu, Amerika Birleşik Devletleri, 21-26 Temmuz 2017.
- Huo, Z., Yu, J. (2023). Reconstructing HDR Image from a Single Exposure LDR Image with Global Feature Fusion. *16th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Taizhou, Çin Halk Cumhuriyeti, 28-30 Ekim 2023.
- Ignatov, A., Gool, L.V., Timofte, R. (2020). Replacing Mobile Camera ISP with a Single Deep Learning Model. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2020)*, Sanal Konferans, 14-19 Haziran 2020.
- Ignatov, A., Malivenko, G., Timofte, R., Tseng, Y., Xu, Y., Yu, P.H., Chiang, C.M., Kuo, H.K., Chen, M.H., Cheng, C.M., Gool, L.V. (2022). PyNet-V2 Mobile: Efficient On-Device Photo Processing with Neural Networks. *International Conference on Pattern Recognition (ICPR 2022)*, Montreal, Kanada, 21-25 Ağustos 2022.

- Ignatov, A., Sycheva, A., Timofte, R., Tseng, Y., Xu, Y.S., Yu, P.H., Chiang, C.M., Kuo, H.K., Chen, M.H., Cheng, C.M., and Van Gool, L., (2022). Microisp: Processing 32mp Photos on Mobile Devices with Deep Learning. *European Conference on Computer Vision (ECCV 2022)*, Tel Aviv, İsrail, 23-27 Ekim 2022.
- Johnson, J., Alahi, A., Fei-Fei, L. (2016). Perceptual Losses for Real-Time Style Transfer and Super-Resolution. *European Conference on Computer Vision (ECCV 2016)*, Amsterdam, Hollanda, 11-14 Ekim 2016.
- Jung, C., Wang, X. (2015). Detail-Preserving Tone Mapping for Low Dynamic Range Displays with Adaptive Gamma Correction. *Conference on Visual Communications and Image Processing (VCIP 2015)*, Singapore, 13-16 Aralık 2015.
- Kalantari, N.K., Ramamoorthi, R. (2017). Deep High Dynamic Range Imaging of Dynamic Scenes. *Association for Computing Machinery (ACM) Transactions on Graphics (TOG)*, 36(4), 1-12.
- Ke, J., Wang, Q., Wang, Y., Milanfar, P., Yang, F. (2021). MUSIQ: Multi-scale Image Quality Transformer. *IEEE/CVF International Conference on Computer Vision (ICCV 2021)*, Sanal Konferans, 11-17 Ekim 2021.
- Ke, X., Lin, W., Chen, G., Chen, Q., Qi, X., Ma, J. (2020). EDLLIE-Net: Enhanced Deep Convolutional Networks for Low-Light Image Enhancement. *IEEE Conference on Image, Vision and Computing, (ICIVC 2020)*, Sanal Konferans, 26 Ağustos 2020.
- Kim, S.Y., Oh J., Kim, M. (2020). JSI-GAN: GAN-Based Joint Super-Resolution and Inverse Tone-Mapping with Pixel-Wise Task Specific Filters for UHD HDR Video, *Conference on Artificial Intelligence (AAAI-20)*, New York, Amerika Birleşik Devletleri, 07-12 Şubat 2020.
- Kinoshita, Y., Seo, K., Kiya, H. (2020). Hue-Correction Scheme Based on Constant-Hue Plane for Deep-Learning-Based Color-Image Enhancement. *IEEE Access*, 8, 9540-9550.
- Kirk, K., Andersen, H.J. (2006). Noise Characterization of Weighting Schemes for Combination of Multiple Exposures, *British Machine Vision Conference (BMVC 2006)*, Edinburgh, İskoçya 04-07 Eylül 2006.
- Knigam, D.P., Ba, J. (2015). Adam: A Method for Stochastic Optimization. *3rd International Conference for Learning Representations (ICLR 2015)*, San Diego, Amerika Birleşik Devletleri, 07-09 Mayıs 2015.
- Koscevic, K., Stipetic, V., Provenzi, E., Banic, N., Subasic, M., Loncaric, S. (2021). HD-RACE: Spray-based Local Tone Mapping Operator. *International Symposium on Image and Signal Processing and Analysis (ISPA 2021)*, Zagreb, Hırvatistan, 13-15 Eylül 2021.

- Ledda, P., Santos, L.P., Chalmers, A. (2004). A Local Model of Eye Adaptation for High Dynamic Range Images. *International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa (AFRIGRAPH 2004)*, Stellenbosch, Güney Afrika Cumhuriyeti, 03-05 Kasım 2004.
- Lee, B.D., Sunwoo, M.H. (2021). HDR Image Reconstruction Using Segmented Image Learning. *IEEE Access*, 9, 142729-142742.
- Lee, M.J., Rhee, C., Lee, C.H. (2022). HSVNet: Reconstructing HDR Image from a Single Exposure LDR Image with CNN. *Applied Sciences*, 12(5), 2370-2387.
- Lee, S., An, G.H., Kang, S.J. (2018). Deep Recursive HDRI: Inverse Tonemapping Using Generative Adversarial Networks. *European Conference on Computer Vision (ECCV 2018)*, Münih, Almanya, 09-14 Eylül 2018.
- Liang, Z., Cai, J., Z., Cao, Zhang, L. (2019). CameraNet: A Two-Stage Framework for Effective Camera ISP Learning. *IEEE Transactions on Image Processing*, 30, 2248-2262.
- Lin, H., Hosu, V., Saupe, D. (2019). KADID-10k: A Large-scale Artificially Distorted IQA Database. *International Conference on Quality of Multimedia Experience (QoMEX 2019)*, Berlin, Almanya, 05-07 Haziran 2019.
- Liu, W., Wang, Q., Liu, Y., Li, N. (2018). High Dynamic Tone Mapping Algorithm Based on Wavelet Domain Image Fusion. *IEEE Conference on Industrial Electronics and Applications (ICIEA 2018)*, Wuhan, Çin Halk Cumhuriyeti, 31 Mayıs-02 Haziran 2018.
- Liu, Y., Lai, W., Chen, Y., Kao, Y., Yang, M., Chuang, Y., Huang, J. (2020). Single-Image HDR Reconstruction by Learning to Reverse the Camera Pipeline. *Computer Vision and Pattern Recognition Conference (CVPR 2020)*, Sanal Konferans, 14-19 Haziran 2020.
- Lou, Z., Gevers, T., Hu, N., Lucassen, M. (2015). Color Constancy by Deep Learning. *British Machine Vision Conference (BMVC)*, Swansea, Birleşik Krallık, 07-10 Eylül 2015.
- Malvar, H.S., He, L., Cutler, R. (2004). Quality Linear Interpolation for Demosaicing of Bayer-Patterned Color Images. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2004)*. Montreal, Kanada, 17-21 Mayıs 2004.
- Mann, S., Picard, R.W. (1995). On Being ‘Undigital’ with Digital Cameras: Extending Dynamic Range by Combining Differently Exposed Pictures. *IST Annual Conference*, Boston, Amerika Birleşik Devletleri, 08-12 Ekim 1995.

- Mantiuk, R., Daly, S., Kerofsky, L. (2008). Display Adaptive Tone Mapping. *Association for Computing Machinery (ACM) Transactions on Graphics (TOG)*, 27(3), 1-10.
- Marnierides, D., Bashford-Rogers, T., Hatchett, J., Debattista, K. (2018). Expandnet: A Deep Convolutional Neural Network for High Dynamic Range Expansion from Low Dynamic Range Content. *Computer Graphics Forum (Eurographics)*, 37(2), 37-49.
- Mertens, T., Kautz, J., Reeth, F.V. (2007). Exposure Fusion. *Pacific Conference on Computer Graphics and Applications (PG'07)*, Maui, Amerika Birleşik Devletleri, 29 Ekim-02 Kasım 2007.
- Mitsunaga, T., Nayar, S.K. (1999). Radiometric Self-Calibration. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Fort Collins, Amerika Birleşik Devletleri, 23-25 Haziran 1999.
- Mittal, A., Moorthy, A.K., Bovik, A.C. (2012). No-Reference Image Quality Assessment in the Spatial Domain. *IEEE Transactions on Image Processing*, 21(12), 4695-4708.
- Mittal, A., Soundararajan, R., Bovik, A.C. (2013). Making a “Completely Blind” Image Quality Analyzer. *IEEE Signal Processing Letters*, 20(3), 209-212.
- Mosleh, A., Tahaei, M.S., Clark, J.J., Nia, V.P. (2023). Towards Low-cost Learning-based Camera ISP via Unrolled Optimization. *Conference on Robots and Vision (CRV 2023)*, Montreal, Canada, 06-08 Haziran 2023.
- Nikam, S., Yawale, R. (2015). Color Image Enhancement Using Daubechies Wavelet Transform and HIS Color Model. *International Conference on Industrial Instrumentation and Control (ICIC 2015)*, Pune, Hindistan, 28-30 Mayıs 2015.
- Niu, Y., Wu, J., Liu, W., Guo, W., Lau, R.W.H. (2021). HDR-GAN: HDR Image Reconstruction from Multi-Exposed LDR Images with Large Motions. *IEEE Transactions on Image Processing*, 30, 3885-3896.
- Panetta, K., Kezebou, L., Oludare, V., Agaian, S. Xia, Z. (2021). TMO-Net: A Parameter-Free Tone Mapping Operator Using Generative Adversarial Network, and Performance Benchmarking on Large Scale HDR Dataset. *IEEE Access*, 9, 39500-39517.
- Pattanaik, S.N., Tumblin, J., Yee, H., Greenberg, D.P. (2000). Time-Dependent Visual Adaptation for fast Realistic Image Display. *Computer Graphics and Interactive Techniques Conference (SIGGRAPH 2000)*, New Orleans, Amerika Birleşik Devletleri, 23-28 Temmuz 2000.

- Rana, A., Singh, P., Valenzise, G., Dufaux, F., Komodakis, N., Smolic, A. (2019). Deep Tone Mapping Operator for High Dynamic Range Images. *IEEE Transactions on Image Processing*, 29, 1285-1298.
- Reinhard, E., Stark, M., Shirley, P., Ferwarda, J. (2002). Photographic Tone Reproduction for Digital Images. *Association for Computing Machinery (ACM) Transactions on Graphics (TOG)*, 21(3), 267-276.
- Robertson, M.A., Borman, S., Stevenson, R.L. (2000). Estimation-Theoretic Approach to Dynamic Range Enhancement Using Multiple Exposures. *Journal of Electronic Imaging*, 12(2), 219-228.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, C.L. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2018)*, Salt Lake City, Amerika Birleşik Devletleri, 19-21 Haziran 2018.
- Schwartz, E., Giryes, R., Bronstein, A.M. (2019). DeepISP: Towards Learning an End-to-End Image Processing Pipeline. *IEEE Transactions on Image Processing*, 28(2), 912-923.
- Sharif, S., Naqvi, R.A., Biswas, M. (2021). Beyond Joint Demosaicking and Denoising: An Image Processing Pipeline for a Pixel-bin Image Sensor. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPR 2021)*, Sanal Konferans, 19-25 Haziran 2021.
- Sheikh, H.R., Bovik, A.C. (2006). Image Information and Visual Quality. *IEEE Transactions on Image Processing*, 15(2), 430-444.
- Talebi, H., Milanfar, P. (2018). NIMA: Neural Image Assessment. *IEEE Transactions on Image Processing*, 27(8), 3998-4011.
- Tan, J., Huang, Y., Wang, K. (2018). Logarithmic Tone Mapping Algorithm Based on Block Mapping Fusion. *International Conference on Audio, Language and Image Processing (ICALIP 2018)*, Shanghai, Çin Halk Cumhuriyeti, 16-17 Temmuz 2018.
- Tsin, Y., Ramesh, V., Kanade, T. (2001). Statistical Calibration of the CCD Imaging Process. *IEEE International Conference on Computer Vision (ICCV 2001)*, Vancouver, Kanada, 07-14 Temmuz 2001.
- Ueda, Y., Suetake, N. (2019). Hue-Preserving Color Image Enhancement on a Vector Space of Convex Combination Coefficients. *IEEE International Conference on Image Processing (ICIP 2019)*, Taipei, Taiwan, 22-25 Eylül 2019.
- URL-1:<https://www.flickr.com/>, (Ziyaret Tarihi: 16 Temmuz 2024).
- URL-2:[https://en.wikipedia.org/wiki/Bayer\\_filter](https://en.wikipedia.org/wiki/Bayer_filter), (Ziyaret Tarihi: 060 Şubat 2025).

- URL-3:<https://www.mathworks.com/help/images/train-and-use-a-no-reference-quality-assessment-model.html>, (Ziyaret Tarihi: 16 Temmuz 2024).
- URL-4:<https://www.mathworks.com/help/images/ref/piqe.html>, (Ziyaret Tarihi: 16 Temmuz 2024).
- URL-5:<https://www.mathworks.com/help/deeplearning/ug/nima-image-quality-assessment.html>, (Ziyaret Tarihi: 16 Temmuz 2024)
- URL-6:<https://www.kaggle.com/code/kerneler/image-aesthetic-scoring-with-musiq-models>, (Ziyaret Tarihi: 14 Mayıs 2024)
- URL-7:<https://github.com/IIGROUP/MANIQA>, (Ziyaret Tarihi: 02 Şubat 2025)
- URL-8:<https://github.com/zwx8981/LIQE>, (Ziyaret Tarihi: 10 Haziran 2024)
- URL-9:[https://github.com/korayerbas/raw\\_to\\_tonemappedHDR.git](https://github.com/korayerbas/raw_to_tonemappedHDR.git), (Ziyaret Tarihi: 27 Ocak 2025)
- Varga, D. (2022). Full-Reference Image Quality Assessment Based on Grünwald–Letnikov Derivative, Image Gradients, and Visual Saliency. *Journal Electronics*, 11(4), 559-575.
- Venkatanath, N. , Praneeth, D., Chandrasekhar, M.Bh., Channappayya, S.S., Medasani, S.S. (2015). Blind Image Quality Evaluation Using Perception Based Features. *National Conference on Communications (NCC 2015)*, Mumbai, Hindistan, 27 Şubat-01 Mart 2015.
- Wang, H., Guoyu Lu, Zhang, T. (2021). Unsupervised HDR Image Reconstruction Based On Over/Under-Exposed LDR Image Pair. *IEEE International Conference on Multimedia and Expo (ICME 2021)*, Shenzen, Çin Halk Cumhuriyeti, 05-09 Temmuz 2021.
- Wang, Q., Wu, B., Zhu, P., Zuo, W., Hu, Q. (2020). ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2020)*, Sanal Konferans, 14-19 Temmuz 2020.
- Wang, S., Zhang, Y., Deng, P., Zhou, F. (2011). Fast Automatic White Balancing Method by Color Histogram Stretching. *International Congress on Image and Signal Processing (CISP 2011)*, Shanghai, Çin Halk Cumhuriyeti, 15-17 Ekim 2011.
- Wang, Z., Bovik, A.C. (2002). A Universal Image Quality Index. *IEEE Signal Processing Letters*, 9(3), 81-84.

- 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.
- Wang, Z., Simoncelli, E.P., Bovik, A.C. (2003). Multi-Scale Structural Similarity for Image Quality Assessment. *IEEE Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, Amerika Birleşik Devletleri, 09-12 Kasım 2003.
- Woo, S., Park, J., Lee, J., Kweon, I.S. (2018). CBAM: Convolutional Block Attention Module. *European Conference on Computer Vision (ECCV 2018)*, Münih, Almanya, 08-14 Eylül 2018.
- Wu, P., Hu, C., Luo, N., Tao, L., Zheng, W. (2019). A New Tone Mapping Operator Based on Two-layer Model. *International Conference on Information Systems and Computer Aided Education (ICISCAE 2019)*, Dalian, Çin Halk Cumhuriyeti, 28-30 Eylül 2019.
- Wu, S., Xu, J., Tai, Y.W., Tang, C.K. (2018). Deep High Dynamic Range Imaging with Large Foreground Motions. *European Conference on Computer Vision (ECCV 2018)*, Münih, Almanya, 08-14 Eylül 2018.
- Xing, Y., Qian, Z., Chen, Q. (2021). Invertible Image Signal Processing. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2021)*, Sanal Konferans, 19-25 Haziran 2021.
- Yan, Q., Gong, D., Shi, Q., Hengel, A., Shen, C., Reid, I., Zhang, Y. (2019). Attention Guided Network for Ghost Free High Dynamic Range Imaging. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2019)*, Long Beach, Amerika Birleşik Devletleri, 16-20 Haziran 2019.
- Yang, L., Zhang, R.Y., Li, L., Xie, X. (2021). SimAM: A Simple Parameter-Free Attention Module for Convolutional Neural Networks. *International Conference on Machine Learning (PMLR 2021)*, Sanal Konferans, 18-24 Haziran 2021.
- Yang, S., Wu, T., Shi, S., Lao, S., Gong, Y., Cao, M., Wang, J., Yang, Y. (2022). MANIQA: Multi-dimension Attention Network for No-Reference Image Quality Assessment. *Conference on Computer Vision and Pattern Recognition Workshops (CVPR 2022)*, New Orleans, Amerika Birleşik Devletleri, 19-24 Haziran 2022
- Yee, Y.H., Pattanaik, S. (2003). Segmentation and Adaptive Assimilation for Detail-Preserving Display of High-Dynamic Range Images. *The Visual Computer*, 19(7-8), 457-466.
- Ying, Z., Niu, H., Gupta, P., Mahajan, D., Ghadiyaram, D., Bovik, A. (2020). From Patches to Pictures (paq-2-piq): Mapping the Perceptual Space of Picture Quality. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2020)*, Sanal Konferans, 14-19 Haziran 2020.

Yu, K., Li, Z., Peng, Y., Loy, C.C., Gu, J. (2021). ReconfigISP: Reconfigurable Camera Image Processing Pipeline. *IEEE International Conference on Computer Vision (ICCV 2021)*, Sanal Konferans, 11-17 Ekim 2021.

Zhang, L., Zhang, L., Mou, X., Zhang, D. (2011). FSIM: A Feature Similarity Index for Image Quality Assessment. *IEEE Transactions on Image Processing*, 20(8), 2378-2386.

Zhang, W., Zhai, G., Wei, Y., Yang, X., Ma, K. (2023). Blind Image Quality Assessment via Vision-Language Correspondence: A Multitask Learning Perspective. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2023)*, Vancouver, Kanada, 18-22 Haziran 2023.



## KİŞİSEL YAYIN VE ESERLER

**Erbaş, K.U., Çelebi, A.T.** (2023). Modeling Camera ISP Pipeline with Deep Learning. *Sinyal İşleme ve İletişim Uygulamaları Konferansı (SIU 2023)*, İstanbul, Türkiye, 05-08 Temmuz 2023.

**Erbaş, K.U., Çelebi, A.T.** (2024). RAW to Tonemapped HDR Camera ISP. *Signal, Image and Video Processing (SIViP)*, 18(12), 9037-9049.



## ÖZGEÇMİŞ

İlk ve orta öğrenimini Kocaeli’de tamamladı. 2007 yılında Deniz Lisesi’nden mezun olarak geçiş yaptığı Deniz Harp Okulu’ndan Elektrik ve Elektronik Mühendisi olarak 2011 yılında mezun oldu. 2015 yılında The University of Texas at Austin üniversitesinde Elektromanyetik ve Akustik alanında Yüksek Lisans eğitimini tamamladı. 2018 yılından beri Kocaeli Üniversitesi Fen Bilimleri Enstitüsü Elektronik ve Haberleşme Mühendisliği Bölümü’nde doktora eğitimine devam etmektedir.

