MARMARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

UZAKTAN ALGILAMA GÖRÜNTÜLERİNDE DEPREM SONRASI HASARIN TRANSFORMER TABANLI SEGMENTASYONU

TUBA KÖYLÜ

YÜKSEK LİSANS TEZİ

Bilgisayar Mühendisliği

Anabilim Dalı

Bilgisayar Mühendisliği Programı

DANIŞMAN

Dr. Öğr. Üyesi Gözde KARATAŞ BAYDOĞMUŞ

EŞ-DANIŞMAN

Prof. Dr. Kazım YILDIZ

İSTANBUL, 2026

MARMARA UNIVERSITY

INSTITUTE FOR GRADUATE STUDIES

TRANSFORMER-BASED SEGMENTATION OF POST-EARTHQUAKE DAMAGE IN REMOTE SENSING IMAGES

TUBA KÖYLÜ

523622024

MASTER THESIS

Department of Computer Engineering

Thesis Supervisor

Dr. Öğr. Üyesi Gözde KARATAŞ BAYDOĞMUŞ

Thesis CO- Supervisor

Prof. Dr. Kazım YILDIZ

ISTANBUL, 2026

TEŞEKKÜR

Bu tezin planlanması, yürütülmesi ve tamamlanması sürecinde akademik rehberliği, kıymetli yönlendirmeleri ve sabrı ile çalışmama yön veren danışman hocalarım Dr. Öğr. Üyesi Gözde KARATAŞ BAYDOĞMUŞ ve Prof. Dr. Kazım YILDIZ'a en içten teşekkürlerimi sunarım. Uzmanlıkları ve değerli geri bildirimleri, akademik gelişimim için her zaman ilham kaynağı olmuştur.

Araştırma süreci boyunca alandaki değerli bilgi, birikimi ve tecrübesiyle yolumu aydınlatan, özverili desteğini hiç esirgemeyen kıymetli arkadaşım Sabri BOLKAR'a sağladığı önemli katkılar ve sabrı için çok teşekkür ederim.

Çalışmamın dil ve akademik yazım açısından yetkinleşmesi için değerli vaktini ayırarak titiz bir inceleme yapan saygıdeğer hocam Prof. Dr. Ahmet Talat SAYGAÇ'a katkıları ve nezaketi için şükranlarımı sunarım.

Akademik yolculuğumun zorlu anlarında beni her zaman dinleyen, motivasyonumu yüksek tutan ve tavsiyeleriyle yanımda olan arkadaşım Ayşe AZAKLI'ya; hayatım boyunca benden koşulsuz sevgisini ve desteğini esirgemeyen annem Fatma KÖYLÜ'ye ve kardeşlerime yürekten teşekkür ederim.

Ayrıca TÜBİTAK Bilim İnsanı Destek Programları Başkanlığı (BİDEB) 2210 Ulusal Yüksek Lisans Burs Programı tarafından sağlanan destek için TÜBİTAK'a da teşekkür etmek istiyorum.

Bu çalışma, akademik yolculuğumda bana bir şekilde destek olan herkese adanmıştır.

Tuba KÖYLÜ

MARMARA ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

Marmara Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Öğrencisi Tuba KÖYLÜ'nün "**Uzaktan Algılama Görüntülerinde Deprem Sonrası Hasarın Transformer Tabanlı Segmentasyonu**" başlıklı tez çalışması, ……..tarihinde savunulmuş ve jüri üyeleri tarafından başarılı bulunmuştur.

Jüri Üyeleri

Dr. Öğr. Üyesi Gözde KARATAŞ BAYDOĞMUŞ (Danışman)

Marmara Üniversitesi ..... (İMZA).....

Prof. Dr. Serhat ÖZEKES (Üye)

Marmara Üniversitesi ..... (İMZA).....

Prof. Dr. Güray YILMAZ (Üye)

Biruni Üniversitesi ..... (İMZA).....

ONAY

Marmara Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun tarih ve sayılı kararı ile Tuba KÖYLÜ'nün Bilgisayar Mühendisliği Anabilim Dalı Bilgisayar Mühendisliği Programında Yüksek Lisans derecesi alması onanmıştır.

Fen Bilimleri Enstitüsü Müdürü

Prof. Dr. Bülent Mertoğlu

İÇİNDEKİLER

TEŞEKKÜR iii
ÖZET viii
ABSTRACT ix
SEMBOLLER x
KISALTMALAR xi
ŞEKİL xii
TABLO LİSTESİ xiii
1. GİRİŞ 1
1.1. Problemin Tanımı ve Motivasyon 1
1.2. Çalışmanın Amacı ve Kapsamı 4
1.3. Çalışmanın Özgün Değeri ve Katkıları 5
1.4. Tezin Organizasyonu 6
2. KURAMSAL TEMELLER VE LİTERATÜR ÖZETİ 8
2.1. Semantik Segmentasyon ve Görüntü İşleme Prensipleri 8
2.1.1. Evrişimli Sinir Ağları ve Geleneksel Mimari Yaklaşımlar 9
2.1.2. Kodlayıcı–Kod Çözücü (Encoder–Decoder) Yapısı ve Öznitelik Füzyonu . 11
2.2. Transformer Tabanlı Yaklaşımlar ve Vision Transformer (ViT) Ailesi 13
2.2.1. Self-Attention Mantığı ve Görüntüye Uyarlanması 14
2.2.2. Hiyerarşik Transformer'lar: Swin ve PVT Çizgisi 17
2.2.3. Segmentasyon İçin Transformer Yaklaşımları 20
2.2.4. Transformer Tercihinin Gerekçesi 21
2.3. SegFormer Mimarisi: Verimli Bir Segmentasyon Tasarımı 22
2.3.1. Hiyerarşik Transformer Kodlayıcı ve Örtüşmeli Yama Birleştirme 23
2.3.2. Tümüyle MLP Tabanlı Kod Çözücü (All-MLP Decoder) ve Hesaplama Verimliliği 25
2.3.3. SegFormer'ın Hasar Tespiti İçin Avantajları 26
2.4. Literatürde Deprem Sonrası Hasar Tespiti Çalışmaları 27
2.4.1. Tek Zamanlı (Post-Event) Hasar Tespiti Yaklaşımları 29
2.4.2. Çift Zamanlı (Pre-Post) Değişim Tabanlı Hasar Tespiti Yaklaşımları 30
3. MATERYAL VE METOD 32
3.1. Veri Kümesi ve Problem Tanımı 32

3.2. Veri Hazırlığı ve Protokol 34
3.3. Model Mimarileri ve Deney Versiyonları 36
3.3.1. SegFormer: MiT omurga ve all-MLP decoder 37
3.3.2. Swin Transformer ve UPerNet 38
3.3.3. Kayıp Fonksiyonları ve IoU Odaklı Öğrenme Motivasyonu 38
3.3.4. Deney Versiyonları 39
3.3.5. Olasılık haritası, Eşikleme ve (Opsiyonel) Test-Time Augmentation 42
3.4. Performans Metrikleri 42
3.4.1. Karışıklık Matrisi 43
3.4.2. Hassasiyet (Precision), Duyarlılık (Recall) ve F1 Skoru 43
3.4.3. Birleşme Üzerinden Kesişim (IoU) ve Ortalama IoU (mIoU) 44
3.4.4. Doğruluk (Accuracy) 45
3.4.5. Eşikleme (Threshold) ve Olasılık Haritası 45
4. BULGULAR 46
4.1. Deneylerin Genel Özeti ve Raporlama Protokolü 46
4.2. Nicel Sonuçların Karşılaştırması (VAL ve TEST) 48
4.2.1. Karşılaştırma Senaryoları ve Adil Protokol 48
4.2.2. VAL ve TEST Sonuçları 49
4.2.3. Confusion Matrix (Karmaşıklık Matrisi) Analizi 50
4.2.4. Nitel Sonuçların Görsel Analizi 55
4.2.5. Baseline ve Kayıp Fonksiyonu Değişikliğinin Etkisi (Lovász) 56
4.2.6. Veri Artırma (Augmentation) Etkisi: En Büyük Kazanç 57
4.2.7. TTA'nin Katkısı: Çıkarım Aşamasında Ek Kazanç 57
4.2.8. Swin-UPerNet Sonuçları: Transformer Tabanlı Alternatif Kıyas 58
4.2.9. Bölüm Özeti 58
4.3. Eşik (Threshold) Seçimi ve Testte Sabitleme 59
4.4. KATE-PD Veri Seti Tanıtım Çalışması ile Karşılaştırma ve Raporlama Tutarlılığı 61
4.4.1. KATE-PD Referans Değerleri ve Baseline Analizi 62
4.4.2. Bu Tezde Elde Edilen Sonuçlarla Kıyas 62
4.5. Test Zamanı Artırma ve Son İşleme Adımları 63
4.6. İstatistiksel Analiz ve Model Karşılaştırması 65
4.6.1. Friedman Testi Sonuçları 66
4.6.2. Nemenyi Post-hoc Testi Sonuçları 67

5. SONUÇ VE ÖNERİLER 70
5.1. Sonuçlar 70
5.2. Gelecek Çalışmalar İçin Öneriler..... 71
6. KAYNAKLAR..... 73

ÖZET

UZAKTAN ALGILAMA GÖRÜNTÜLERİNDE DEPREM SONRASI HASARIN TRANSFORMER TABANLI SEGMENTASYONU

Bu tezde, 6 Şubat 2023 Kahramanmaraş depremleri sonrasında oluşan yapı hasarlarının yüksek çözünürlüklü optik uydu görüntülerinden, yalnızca afet sonrası görüntüye dayalı tek zamanlı senaryoda otomatik çıkarımı ele alınmıştır. Afet öncesi görüntüye erişimin her zaman mümkün olmaması ve sahada hızlı durum farkındalığı gereksinimi nedeniyle problem, arka plan ve hasar olmak üzere iki sınıflı semantik segmentasyon olarak formüle edilmiştir. Çalışmanın amacı, aynı veri bölünmesi ve raporlama protokolü altında transformer tabanlı iki yaklaşımı karşılaştırmak ve SegFormer tabanlı hattı yöntemsel olarak iyileştirerek performansı artırmaktır.

Deneyler KATE-PD veri seti üzerinde yürütülmüş, görüntüler veri setinin parça tabanlı yapısına uygun olarak 512×512 yamalar üzerinden işlenmiştir. Model girişinde yeniden ölçekleme, yeniden boyutlandırma ve normalize etme adımları uygulanmıştır. Ana mimari olarak SegFormer seçilmiş; hiyerarşik kodlayıcı yapısıyla çok ölçekli temsilleri verimli üretmesi ve hafif çözücü tasarımıyla uygulanabilir bir hesaplama profili sunması bu tercihin temel gerekçesini oluşturmuştur. Karşılaştırma amacıyla Swin Transformer + UPerNet hattı aynı değerlendirme koşulları altında test edilmiştir. SegFormer hattında kayıp fonksiyonu, veri artırma ve çıkarım stratejileri kontrollü karşılaştırmalarla incelenmiş; eşik değeri doğrulama kümesi üzerinde belirlenerek test aşamasında sabitlenmiştir. Performans mIoU ve F1 metrikleriyle raporlanmıştır.

En iyi SegFormer konfigürasyonu, KATE-PD test kümesinde 0.7463 mIoU ve 0.6732 F1 değerlerine ulaşmıştır. Bulgular, başarının yalnızca model büyüklüğünden değil, eğitim ve çıkarım kararlarının tutarlı biçimde tasarlanmasından da etkilendiğini göstermiştir. Bu çalışma, tek zamanlı optik uydu görüntüleriyle deprem sonrası hasar segmentasyonu için tekrarlanabilir bir deney çerçevesi sunmaktadır.

Anahtar Kelimeler: Deprem sonrası hasar tespiti, uzaktan algılama, semantik segmentasyon, Transformer, SegFormer, Swin Transformer, KATE-PD, test-zamanı artırımı.

ABSTRACT

TRANSFORMER-BASED SEGMENTATION OF POST-EARTHQUAKE DAMAGE IN REMOTE SENSING IMAGES

This thesis investigates the automatic extraction of structural damage from high-resolution optical satellite imagery acquired after the 6 February 2023 Kahramanmaraş earthquakes, under a single post-event setting where only post-disaster imagery is available. Since pre-event imagery cannot always be obtained in practice and rapid situational awareness is essential in emergency response, the problem is formulated as a binary semantic segmentation task with two classes: background and damage. The main objective is to compare two transformer-based approaches under the same data split and reporting protocol, and to improve a SegFormer-based pipeline through methodological refinements.

Experiments are conducted on the KATE-PD dataset, and images are processed using the dataset's patch-based structure with 512×512 tiles. Standard input processing steps, including rescaling, resizing, and normalization, are applied prior to model inference. SegFormer is selected as the primary architecture due to its efficient multi-scale representations enabled by a hierarchical encoder and its lightweight decoder design, which offers a practical computational profile. For comparison, a Swin Transformer + UPerNet pipeline is evaluated under identical conditions. Within the SegFormer pipeline, the effects of the loss function, data augmentation, and inference strategies are examined through controlled comparisons; the probability threshold for binary mask generation is selected on the validation set and fixed during testing. Performance is reported using pixel-wise metrics, primarily mIoU and F1 score.

The best SegFormer configuration achieves 0.7463 mIoU and 0.6732 F1 on the KATE-PD test set. The findings indicate that performance is influenced not only by model capacity but also by coherent design choices across training and inference. Overall, this work provides a reproducible experimental framework for post-event damage segmentation using single-date optical satellite imagery.

Keywords: Post-earthquake damage assessment, remote sensing, semantic segmentation, Transformer, SegFormer, Swin Transformer, KATE-PD, test-time augmentation.

SEMBOLLER

TP	**:** Doğru pozitif (true positive)
FP	**:** Yanlış pozitif (false positive)
FN	**:** Yanlış negatif (false negative)
TN	**:** Doğru negatif (true negative)
P	**:** Hassasiyet (precision)
R	**:** Duyarlılık (recall)
F1	**:** F1 skoru (F1-score)
IoU_c	**:** c sınıfı IoU değeri (class-wise IoU)
mIoU	**:** Ortalama IoU (mean IoU)
Acc	**:** Doğruluk (accuracy)
C	**:** Sınıf kümesi (class set)
|C|	**:** Sınıf sayısı (number of classes)
Σ	**:** Toplam sembolü (summation)
H, W	**:** Görüntü boyutları (height, width)
N	**:** Token sayısı (number of tokens)
D	**:** Gömme boyutu (embedding dimension)
h	**:** Başlık sayısı (number of heads)
Q, K, V	**:** Sorgu–anahtar–değer (query–key–value)

KISALTMALAR

AI : Yapay zekâ (Artificial Intelligence)

CD : Değişim tespiti (Change Detection)

CNN : Evrişimli sinir ağı (Convolutional Neural Network)

DL : Derin öğrenme (Deep Learning)

F1 : F1 skoru

FP : Yanlış pozitif

FN : Yanlış negatif

GT : Zemin gerçekliği (Ground Truth)

IoU : Kesişim/Birleşim (Intersection over Union)

KATE-PD : Kahramanmaraş depremi sonrası veri seti

LN : Katman normalizasyonu (Layer Normalization)

ML : Makine öğrenmesi (Machine Learning)

MLP : Çok katmanlı algılayıcı (Multi-Layer Perceptron)

mIoU : Ortalama IoU

MiT : Mix Transformer (SegFormer omurgası)

PVT : Pyramid Vision Transformer (piramidal ViT)

Reg-aware CD : Kayıt hatasına dayanıklı değişim tespiti

RS : Uzaktan algılama (Remote Sensing)

SAR : Sentetik açıklıklı radar (Synthetic Aperture Radar)

SRA : Uzamsal indirgemeli dikkat (Spatial-Reduction Attention)

SW-MSA : Kaydırmalı pencere dikkat (Shifted Window MSA)

TR : Türkiye

TTA : Test zamanı artırımı (Test-Time Augmentation)

UPerNet : UPerNet (segmentasyon çözücüsü)

ViT : Görüntü Transformer'ı (Vision Transformer)

VHR : Çok yüksek çözünürlük (Very High Resolution)

W-MSA : Pencere tabanlı dikkat (Window MSA)

xBD : xBD (xView2) hasar veri seti

ŞEKİL

Şekil 1.1. Şubat 2023 Kahramanmaraş depremi (7,7) etki alanı 1

Şekil 2.1. Self-Attention mantığı ve görüntüye uyarlanması 15

Şekil 2.2. Swin Transformer blok yapısı (W-MSA → SW-MSA) ve artık bağlantılar 18

Şekil 2.3. Hiyerarşik Transformer'lar: Swin ve PVT çizgisi (Çok-Ölçekli Temsil) 19

Şekil 2.4. SegFormer: MiT hiyerarşik kodlayıcı + All-MLP decoder (Verimli Segmentasyon) 22

Şekil 3.1. KATE-PD veri setinden alınmış örnek RGB ve ikili maske görüntüleri. 33

Şekil 3.2. ADE20K'da performans ve model verimliliği karşılaştırması 37

Şekil 3.3. Deney hattının genel akışı ve karşılaştırmalı modelleme çerçevesi. 40

Şekil 4.1. Değerlendirme akışı. 46

Şekil 4.2. Baseline'a göre TEST mIoU değişim karşılaştırması. 50

Şekil 4.3. Örnek tahminlerin görsel olarak karşılaştırılması 55

Şekil 4.4. SegFormer Augmentation TTA deneyinin threshold değerlerine göre f1 değişimi. 60

Şekil 4.5. SegFormer Augmentation TTA deneyinin threshold değerlerine göre mIoU değişimi. 60

Şekil 4.6. Test zamanı artırma çıkarım akışı. 64

Şekil 4.7. Patch-bazlı IoU_1 dağılımları (Test seti, N = 26)..... 67

Şekil 4.8. Nemenyi post-hoc testi p-değerleri ısı haritası..... 69

TABLO LİSTESİ

Tablo 2.1. Deprem sonrası hasar tespitinde temel çalışmaların kronolojik özeti. 28

Tablo 3.1. KATE-PD veri setinin bölümleme sayı ve oranları..... 35

Tablo 4.1. Çalışmaya ait deneylerin VAL/TEST sonuçlarının karşılaştırılması. 49

Tablo 4.2. Piksel bazlı Confusion Matrix sonuçları (VAL+TEST)..... 51

Tablo 4.3. Yapı bazlı Confusion Matrix sonuçları (TEST). 53

Tablo 4.4. Piksel ve yapı bazlı recall değerlerinin karşılaştırılması (TEST). 54

Tablo 4.5. Modellerin hesaplama karmaşıklığı karşılaştırması. 58

Tablo 4.6. KATE-PD veri seti tanıtım çalışması Baseline'ları ile bu tezde elde edilen sonuçların karşılaştırılması (TEST)..... 63

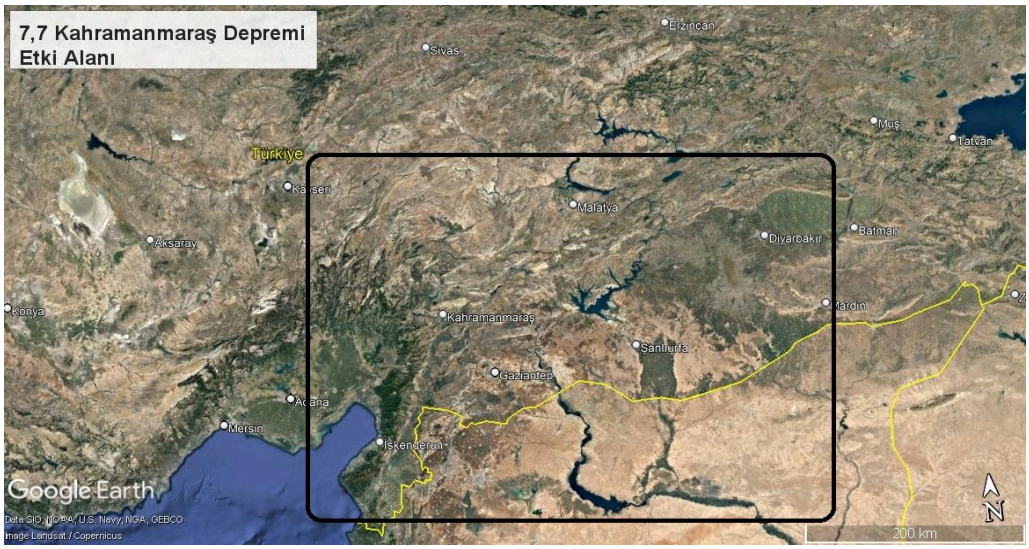
Tablo 4.7. Friedman testi sonuçları..... 66

Tablo 4.8. Nemenyi post-hoc testi p-değerleri ($p < 0.05$ anlamlı farkı göstermektedir). 67

1. GİRİŞ

1.1. Problemin Tanımı ve Motivasyon

Doğal afetler içerisinde depremler, özellikle hızlı kentleşme süreçlerinin yaşandığı ve nüfus yoğunluğunun yüksek olduğu bölgelerde, çok kısa süreler içerisinde büyük ölçekli yapısal yıkım ve yüksek ekonomik kayıp üretebilen en kritik risk türlerinden biridir [1]. Bu durumun en çarpıcı ve yakın dönemli örneği olan 6 Şubat 2023'te Türkiye'nin güneyinde meydana gelen Kahramanmaraş merkezli deprem dizisi, deprem tehlikesinin yalnızca sarsıntı büyüklüğüyle değil; kırılan fayın uzunluğu, artçı etkinliğin yoğunluğu, yerleşim alanlarının maruziyeti ve afet sonrası ilk saatlerdeki erişim/koordinasyon kısıtlarıyla birlikte ele alınması gerektiğini çarpıcı biçimde göstermiştir. Bu olayda aynı gün içinde iki büyük ana şok gerçekleşmiş; büyüklük kestirimleri kurumlara göre küçük farklılıklar göstermekle birlikte, AFAD raporlarında depremler Mw 7.7 (Pazarcık) ve Mw 7.6 (Elbistan) olarak, USGS kayıtlarında ise sırasıyla M 7.8 ve M 7.5 olarak raporlanmıştır [2]. Deprem dizisi Doğu Anadolu Fay Sistemi ile ilişkili doğrultu atımlı kırılma mekanizmalarıyla açıklanmakta; USGS değerlendirmeleri, bu olayların Doğu Anadolu fay sistemi içerisinde geliştiğini vurgulamaktadır [3]. Nitekim Şekil 1.1'de verilen etki alanı görselleştirmesi, deprem dizisinin yalnızca merkez üssü çevresiyle sınırlı kalmadığını; çok sayıda yerleşim alanını kapsayan geniş bir coğrafi çerçevede değerlendirilmesi gerektiğini göstermektedir.



Şekil 1.1. Şubat 2023 Kahramanmaraş depremi (7,7) etki alanı [4].

Kentleşmenin hızlanmasıyla birlikte yapı stoğunun yoğunlaşması, yapı kalitesi ve denetim farklılıklarıyla birleştiğinde depremlerin etkisi yalnızca fiziksel yıkım düzeyinde kalmamakta; aynı zamanda sosyal, lojistik ve ekonomik sistemlerde zincirleme kırılmalara yol açmaktadır [1, 5]. Bu nedenle deprem riskinin azaltılması ve afet sonrası yönetim kapasitesinin güçlendirilmesi, yalnızca mühendislik açısından değil; afet lojistiği, karar destek ve bilgi sistemleri açısından da çok boyutlu bir problem alanı olarak ele alınmaktadır [6]. Afet sonrası ilk saatler ve günler, arama-kurtarma ve acil müdahale faaliyetlerinin yoğunlaştığı, zaman baskısının en yüksek olduğu kritik dönemdir. Bu evrede müdahale ekiplerinin en temel ihtiyacı; hasarın nerede, hangi yoğunlukta ve hangi yayılım biçiminde gerçekleştiğini mümkün olduğunca hızlı biçimde belirleyerek ekip, araç ve tıbbi-insani yardım kaynaklarını öncelikli bölgelere yönlendirmektir [6, 7]. Ancak geleneksel saha incelemeleri; etkilenen alanın büyüklüğü, erişim kısıtları, güvenlik riskleri, artçı sarsıntılar ve saha koşullarının belirsizliği gibi nedenlerle her zaman yeterince hızlı ve kapsamlı yürütülememektedir [7, 8]. Bu durum, erken dönemde karar vericilerin ihtiyaç duyduğu durum farkındalığının (situational awareness) gecikmesine neden olabilmekte ve müdahalenin etkililiğini azaltabilmektedir [7].

Bu kısıtlar karşısında uzaktan algılama verileri; geniş alanları kısa sürede gözlemleyebilmesi, sahaya fiziksel erişim gerektirmemesi ve mekânsal olarak tutarlı bir "kuş bakışı" perspektif sunması nedeniyle afet yönetiminde stratejik bir araç haline gelmiştir [9]. Özellikle afet sonrası hasar değerlendirmesinde yüksek çözünürlüklü uydu/hava görüntülerinin kullanımı, yalnızca akademik bir araştırma konusu değil; aynı zamanda acil durum haritalama ve hızlı haritalama (rapid mapping) gibi operasyonel süreçlerde doğrudan karşılığı olan bir uygulamadır [10, 11]. Uydu tabanlı hızlı haritalama, afet sonrasında görüntü edinimi–işleme–analiz döngüsünü hızlandırarak; hasar yoğunluğu, erişilebilirlik, kritik altyapı etkilenmesi gibi başlıklarda karar vericilere kısa sürede coğrafi ürün sunmayı hedefler [10, 11]. Bu bağlam, uzaktan algılama temelli otomatik/yarı otomatik hasar tespit yöntemlerinin önemini artırmakta; saha doğrulamasıyla desteklenen uydu analizlerinin kaynak tahsisi ve planlama süreçlerinde değer ürettiğini göstermektedir [10, 11].

Uzaktan algılama görüntülerinden bina hasarı tespiti yapmak, bilgisayarlı görü açısından oldukça zorlayıcı bir problemdir. Deprem kaynaklı hasarlar; binaların tamamen yıkılması (enkaz), kısmi çökmeler, çatı deformasyonları, yayılmış moloz yığınları ve çevresel

değişimler gibi farklı görsel karakterlerle ortaya çıkabilmektedir [12, 13]. Bu görsel ipuçları; görüntünün çözünürlüğüne, güneşin geliş açısına ve gölge desenlerine, yapı malzemesine ve kentsel dokunun karmaşıklığına bağlı olarak ciddi değişkenlik gösterir [12-14]. Ayrıca hasarlı bölgelerin çıplak toprak, gölge, geçici yıkım görüntüsü, moloz kümeleri gibi çevresel unsurlarla karışabilmesi, sınıf dengesizliği ve yanlış alarm (false positive) üretimini artıran temel teknik zorluklar arasında raporlanmıştır [13, 14]. Bu nedenle uydu görüntüsünden hasar tespiti çoğu senaryoda yalnızca hasarlı-hasarsız sınıflandırma probleminden öte, mekânsal bağlamın kritik olduğu yoğun tahmin ile semantik segmentasyon karakteri taşıyan bir problemdir [15].

Afet hasarının uzaktan algılama ile tespitinde yöntemsel açıdan önemli bir ayrım, afet öncesi ve sonrası görüntüleri birlikte kullanan çift-zamanlı (bi-temporal) yaklaşımlar ile yalnızca afet sonrası veriyi kullanan tek-zamanlı (post-event only) yaklaşımlar arasındadır [8]. Çift-zamanlı yöntemler değişim sinyalini daha doğrudan yakalayabilse de, operasyonel gerçeklikte afet öncesi görüntülerin her zaman uygun çözünürlükte, aynı sensör/aynı bakış geometrisinde veya yeterince güncel biçimde bulunması garanti değildir [8, 16]. Ayrıca afet öncesi görüntünün çok eski olması durumunda depremle ilgisiz kentsel değişimler (yeni inşaatlar, yıkımlar, arazi düzenlemeleri vb.) "hasar" olarak yanlış algılanabilmektedir [16]. Bu kısıtlar, modelin herhangi bir referans görüntüye ihtiyaç duymadan doğrudan hasar imzasını afet sonrası tek görüntüde öğrenmesini hedefleyen tek-zamanlı yaklaşımların operasyonel değerini artırmaktadır [13, 14].

Son yıllarda derin öğrenme tabanlı yaklaşımlar, uzaktan algılama görüntülerinde hasar tespiti ve semantik segmentasyon görevlerinde belirgin ilerlemeler sağlamıştır [13, 14]. Bununla birlikte afet bağlamında veri çeşitliliği, sınıf dengesizliği ve genelleme sorunları devam etmekte; bu nedenle model mimarisi kadar eğitim stratejileri, veri artırma ve uygun kayıp fonksiyonları da performansı belirleyen önemli unsurlar olarak öne çıkmaktadır [13, 17, 18]. Özellikle sınıf dengesizliği ve segmentasyon kalitesini doğrudan optimize etme ihtiyacı, çapraz entropi gibi standart kayıpların yanında Lovász gibi IoU tabanlı ölçütlerle ilişkili yaklaşımların kullanımını gündeme getirmiştir [17]. Ek olarak, test zamanında dönüşüm tabanlı çıkarım (Test-Time Augmentation, TTA) gibi stratejiler, modelin çıktı kararlılığını artırmak için pratikte sıkça kullanılan yöntemlerdendir [18].

Bu tez çalışması, yukarıda özetlenen operasyonel ve teknik gerekçeler doğrultusunda, "yalnızca afet sonrası" görüntüden bina hasarı tespiti senaryosuna odaklanmaktadır. Bu

kapsamda, semantik segmentasyon alanında son yıllarda yaygınlaşan Transformer tabanlı mimarilerin, özellikle de SegFormer'ın, hasar tespiti problemindeki potansiyeli incelenmektedir [19]. SegFormer, çok ölçekli özellik temsili ve verimli tasarımı ile farklı çözünürlüklerde semantik segmentasyon görevlerinde güçlü sonuçlar üretmeyi hedefleyen bir mimaridir [19]. Çalışmada SegFormer tabanlı bir hat üzerinden, temel eğitim (baseline), uygun kayıp fonksiyonları ile iyileştirme, veri artırma, test zamanı stratejileri ve model karşılaştırmaları gibi deney tasarımlarıyla performans analizleri gerçekleştirilmektedir. Böylece hem afet yönetimi bağlamında pratik değer taşıyan bir problem senaryosu ele alınmakta, hem de modern segmentasyon mimarilerinin bu senaryodaki davranışı sistematik biçimde değerlendirilmektedir [13, 14, 19].

1.2. Çalışmanın Amacı ve Kapsamı

Bu çalışmanın temel amacı, uzaktan algılama literatüründe operasyonel açıdan en zorlu senaryolardan biri olarak kabul edilen tek-zamanlı (yalnızca post-disaster) uydu görüntüleri üzerinden bina hasarlarının piksel düzeyinde segmentasyonunu, modern derin öğrenme yaklaşımlarıyla mümkün olduğunca güvenilir ve tekrarlanabilir şekilde gerçekleştirmektir. Bu amaçla, 6 Şubat 2023 depremleri sonrasına ait deprem sonrası görüntüler ve hasar anotasyonları içeren KATE-PD (Kahramanmaraş Türkiye Earthquake – Post Disaster) veri seti kullanılmıştır [20, 21]. KATE-PD, özellikle Kahramanmaraş depremlerinden ağır etkilenen bölgeleri kapsayan Maxar, Open Data ve Airbus Pleiades yüksek çözünürlüklü uydu görüntülerini araya getirerek, öncesi görüntüye ihtiyaç duymadan yalnızca afet sonrası görüntüler üzerinden hasar alanlarının etiketlenmesini mümkün kılar [20, 21]. Veri setinde, görüntüler grid tabanlı bir yaklaşımla 512×512 yamalara bölünerek, manuel poligon anotasyonları aracılığıyla hasarlı bina alanları işaretlenmiş; bu poligonlar üzerinden pikseller ikili (0/1) biçimde etiketlenerek model eğitimine uygun hale getirilmiştir [20, 21]. Modelleme tarafında, uzun menzilli bağlamı daha etkin kullanabilen Transformer tabanlı yaklaşımlar öne çıktığından, ana mimari olarak semantik segmentasyon için tasarlanmış SegFormer tercih edilmiştir [19]. SegFormer'ın hiyerarşik Transformer encoder + hafif MLP decoder tasarımı, farklı ölçekteki temsil düzeylerini birleştirerek yüksek çözünürlükte maske üretimine uygun, yalın ama güçlü bir çerçeve sunar [19].

Bu tezin kapsamı, statik bir model eğitim sürecinin ötesine geçerek, bina hasarı tespiti problemi üzerinde bütüncül ve kademeli bir iyileştirme hattı kurmayı hedeflemektedir. Bu doğrultuda; veri hazırlığı süreçlerinden eğitim stratejilerine, veri artırma kayıp fonksiyonu ve optimizasyon tercihlerine kadar her aşama sistematik bir biçimde ele alınmıştır. Ayrıca, elde edilen sonuçların güvenilirliğini sağlamak amacıyla değerlendirme protokolleri adım adım optimize edilerek, modelin gerçek dünya senaryolarındaki başarısını en üst düzeye çıkaracak metodolojik bir yol haritası izlenmiştir.

1.3. Çalışmanın Özgün Değeri ve Katkıları

Bu tezin özgün değeri, Türkiye'de yakın dönemde meydana gelen yıkıcı bir afet bağlamında oluşturulmuş tek zamanlı (post-event) bir veri seti üzerinde, Transformer tabanlı semantik segmentasyon yaklaşımlarının bina hasarı tespitindeki performansını sistematik biçimde sınaması ve bu deneysel süreci şeffaf bir şekilde raporlamasıdır. Literatürde afet sonrası hasar tespiti için xBD gibi küresel ölçekli veri setleri bulunsa da, olaya ve bölgeye özgü kentsel doku, bina tipolojileri, enkazın görsel karakteristiği, görüntüleme zamanı ve sensör koşulları gibi faktörler model çıktısını doğrudan ve ciddi biçimde etkileyebilmektedir [8]. Bu nedenle, yalnızca genel amaçlı veri setleri üzerinde elde edilen bulguların, belirli bir afete ve ülkeye özgü sahaya doğrudan taşınması her zaman mümkün olmayabilir.

Bu bağlamda KATE-PD gibi belirli bir afete ve coğrafyaya odaklanan veri setleri üzerinde aynı mimarinin farklı eğitim ve iyileştirme stratejileri altında nasıl davrandığını incelemek, iki açıdan kritik bir katkı üretir. Birincisi, modelin hangi koşullarda başarımının arttığını veya düştüğünü ortaya koyarak literatüre deneysel birikim kazandırır. İkincisi, afet sonrası hızlı durum tespiti gibi zaman kritik operasyonlarda uygulanabilirliği yüksek, tekrarlanabilir bir yaklaşımın oluşturulmasına yardımcı olur [20, 21]. Bu tez, tam da bu boşluğa odaklanarak, tek zamanlı uydu görüntülerinden hasar segmentasyonu problemini yalnızca bir model eğitme adımı olarak değil, uçtan uca bir metodolojik süreç olarak ele almaktadır. Çalışmanın katkıları bu çerçevede üç ana eksende toplanmaktadır. İlk olarak, KATE-PD veri seti üzerinde SegFormer tabanlı bir temel yaklaşım kurulmakta; veri hazırlama, eğitim ve değerlendirme adımları açık ve izlenebilir biçimde tanımlanmaktadır [20, 21].

İkinci olarak, performansı artırmak amacıyla eğitim hattı kademeli olarak geliştirilmekte; veri artırma stratejileri, kayıp fonksiyonu ve optimizasyon tercihleri, eğitim planı ve değerlendirme protokolü gibi bileşenler kontrollü bir şekilde değiştirerek her adımın etkisi ölçülmektedir. Bu yaklaşım, "hangi iyileştirme neden işe yaradı?" sorusuna deneysel olarak cevap üretmeyi hedefler ve sonuçların yorumlanabilirliğini güçlendirir. Üçüncü olarak, elde edilen çıktılar sayısal metriklerle düzenli biçimde raporlanarak karşılaştırılabilir bir deney seti oluşturulmaktadır; böylece hem gelecekteki çalışmalar için referans niteliğinde bir temel sağlanmakta hem de gerçek dünya uygulamalarına daha yakın bir performans okuması sunulmaktadır.

Sonuç olarak bu tez, deprem sonrası hasar segmentasyonu için bilimsel olarak temellendirilmiş, tekrarlanabilir ve kademeli iyileştirmeye açık bir boru hattı (pipeline) ortaya koymayı; aynı zamanda Türkiye'deki büyük ölçekli bir afet örneği üzerinden, Transformer tabanlı segmentasyon yaklaşımlarının güçlü ve zayıf yönlerini görünür kılmayı amaçlamaktadır. Bu yönüyle çalışma, yalnızca bir model denemesi değil, saha koşullarına duyarlı bir değerlendirme ve geliştirme çerçevesi sunan bütüncül bir deneysel katkı niteliği taşımaktadır.

1.4. Tezin Organizasyonu

Beş ana bölümden oluşan bu tez, okuyucuyu problem tanımından başlayarak veri ve yöntem detaylarına, ardından deneysel bulgulara ve sonuçların sentezine taşıyan sistematik bir akışla kurgulanmıştır. Birinci bölümde; problemin tanımı, çalışmanın temel motivasyonu, amaç ve kapsam çerçevesi ile literatüre sağlanan özgün değer ve katkılar sunulmuştur. Bu bölüm, aynı zamanda 6 Şubat 2023 depremleri bağlamında geliştirilen yaklaşımın hangi ihtiyaca cevap verdiğini ve çalışmanın hangi araştırma sorularına odaklandığını netleştirmektedir.

İkinci bölümde, uzaktan algılama teknolojileri ile afet sonrası hasar tespiti alanında yürütülen çalışmalar ele alınarak, bu tezde kullanılan yaklaşımın literatürdeki konumu tartışılmaktadır. Bu kapsamda, özellikle semantik segmentasyonun kuramsal temelleri, problem tanımı ve temel kavramlar derli toplu biçimde özetlenmektedir [22]. Bölüm, aynı zamanda uzaktan algılama görüntülerinde hasar tespitinin zorluklarını ve tek zamanlı görüntü senaryosunun neden daha zorlayıcı bir problem alanı oluşturduğunu açıklayarak sonraki bölümler için kavramsal bir zemin sunmaktadır.

Üçüncü bölümde; çalışma kapsamında kullanılan KATE-PD veri setinin özellikleri, veri hazırlığı ve ön işleme adımları, deneylerde kullanılan mimari tercihleri ve eğitim-değerlendirme protokolleri ayrıntılandırılmaktadır [19-21]. Bu bölümün temel amacı, deneylerin tekrarlanabilirliğini sağlayacak şekilde tüm metodolojik adımları açık biçimde tanımlamak ve modelleme sürecinde yapılan tercihlerin gerekçelerini sunmaktır. Ayrıca, veri artırma stratejileri, kayıp fonksiyonu ve optimizasyon tercihleri gibi eğitim hattını oluşturan bileşenlerin nasıl yapılandırıldığı bu bölümde sistematik olarak açıklanmaktadır.

Dördüncü bölümde, yürütülen deneylerden elde edilen nicel sonuçlar belirlenen metrikler üzerinden raporlanmakta ve farklı deneme kurulumları karşılaştırmalı biçimde değerlendirilmektedir. Bunun yanında, yalnızca sayısal performansın değil, modelin çıktılarının sahaya yakın senaryolarda nasıl davrandığının anlaşılabilmesi için nitel örnekler üzerinden yorumlar sunulmaktadır. Böylece, hangi iyileştirme adımlarının ne ölçüde katkı sağladığı, hangi durumlarda hata örüntülerinin ortaya çıktığı ve modelin sınırlılıklarının nerede yoğunlaştığı görünür hale getirilmektedir.

Beşinci bölümde ise çalışma kapsamında ulaşılan bulgular bir bütün olarak sentezlenmekte; elde edilen sonuçlar problem bağlamında yorumlanmakta ve çalışmanın sınırlılıkları açıkça tartışılmaktadır. Son olarak, gelecekteki araştırmalar için geliştirme önerileri sunularak, özellikle veri çeşitliliği, genelleme gücü, farklı mimari seçenekler ve operasyonel kullanım senaryoları gibi başlıklarda olası ilerleme yolları değerlendirilmektedir.

2. KURAMSAL TEMELLER VE LİTERATÜR ÖZETİ

2.1. Semantik Segmentasyon ve Görüntü İşleme Prensipleri

Semantik segmentasyon, bir görüntüdeki her piksele bir sınıf etiketi atayarak (örneğin hasarlı bina, sağlam bina, zemin, yol gibi) sahnenin sınıf-temelli haritasını üretmeyi amaçlayan piksel düzeyinde sınıflandırma problemidir [23]. Bu problem tanımı, uzaktan algılama uygulamalarında özellikle değerlidir; çünkü tek bir görüntü içinde hem küçük ölçekli detayların (bina kenarları, çatı çizgileri, enkaz dokusu) hem de geniş ölçekli bağlamın (mahalle dokusu, yol ağı, açık alanlar) birlikte değerlendirilmesi gerekir. Uzaktan algılama verilerinde bina sınırlarının keskinliği, karmaşık arka plan, gölge etkisi ve ölçek değişimi gibi zorluklar sık görüldüğünden, semantik segmentasyon klasik görüntü sınıflandırma veya nesne tespiti yaklaşımlarına kıyasla daha ayrıntılı ve mekânsal olarak anlamlı bir çıktı sunar [24]. Semantik segmentasyonda hedef yalnızca "nesne var mı?" sorusuna cevap vermek değildir; nesnenin geometrisini mümkün olduğunca doğru biçimde çıkarmak, yani bina ayak izini ve hasar izini piksel düzeyinde belirginleştirmektir. Bu nedenle mimari tasarımlar, bir yandan sahnenin bağlam bilgisini yakalayan daha derin ve soyut özellikleri öğrenmeye çalışırken, diğer yandan sınır ve kenar gibi ince ayrıntıları koruyan yüksek çözünürlüklü bilgiyi kaybetmemeyi hedefler [23]. Bu iki ihtiyacın aynı anda karşılanması kolay değildir; çünkü derin ağlar genellikle havuzlama ve örnekleme azaltma adımlarıyla daha geniş görüş alanı ve daha güçlü bağlamsal temsil kazanırken, bu işlemler uzamsal çözünürlüğü düşürerek sınırların belirsizleşmesine neden olabilir [25]. Bu motivasyon, literatürde tam evrişimli ağların ortaya çıkışının temel nedenlerinden biridir. Long ve arkadaşları, sınıflandırma için tasarlanmış evrişimsel omurgaları pikselden piksele tahmin üretecek şekilde dönüştürerek, görüntü boyutundan bağımsız yoğun tahmin yapabilen "fully convolutional network" yaklaşımını önermiştir [23]. Bu yaklaşımda, ağın derin katmanları semantik olarak güçlü fakat kaba çözünürlüklü temsil üretirken, daha sığ katmanlar görsel ayrıntıyı daha iyi taşıyabildiği için farklı seviyelerden gelen bilginin birleştirilmesiyle hem doğru sınıf kararı hem de daha düzgün sınır üretimi hedeflenir [23].

Benzer biçimde encoder–decoder ailesi, bir taraftan bağlamı yakalayan daraltma yolu, diğer taraftan konumsal doğruluğu geri kazandıran genişletme yolu ile yoğun tahmin problemlerinde yaygın bir çerçeve sunar; U-Net mimarisi bu çizginin en bilinen örneklerinden biridir [26]. U-Net'teki skip bağlantıları, erken katmanlardaki ayrıntılı

uzamsal bilginin tekrar üst katmanlara taşınmasını sağlayarak sınır ve ince yapıların korunmasına katkı verir [26]. Uzaktan algılama görüntülerinde segmentasyonun bir diğer temel gereksinimi, çok ölçekli yapıları birlikte ele almaktır. Aynı sahnede küçük yapıların yanı sıra büyük yapı blokları ve geniş açık alanlar bulunabildiği için, modelin farklı ölçeklerdeki örüntülere duyarlı olması beklenir. DeepLab ailesi, görüş alanını büyütürken çözünürlüğü daha kontrollü tutabilen atrous convolution yaklaşımını ve çok ölçekli bağlam yakalamayı hedefleyen ASPP gibi bileşenleri bu amaçla ön plana çıkarmıştır [25]. Ayrıca, yalnızca bağlamı güçlendirmek yeterli olmadığından, sınır doğruluğunu artırmak için segmentasyon çıktısının uzamsal olarak rafine edilmesi fikri de literatürde önemli bir yer tutar; DeepLab çalışmaları bu amaçla olasılıksal rafinman yaklaşımlarını da tartışır [25].

Son olarak, uzaktan algılama bağlamında bina odaklı semantik segmentasyonun güçlükleri yalnızca mimari tasarımla sınırlı değildir. Arka planın karmaşıklığı, bina çeşitliliği ve sahne içi görünüm değişkenliği, özellikle bina çıkarımı ve bina sınıflandırması gibi görevlerde belirgin hata örüntüleri üretir [24]. Bu nedenle semantik segmentasyon problemini ele alırken, görüntü işleme prensipleri kadar veri temsili, ölçek ve bağlam ilişkisi, sınırların korunması ve değerlendirme yaklaşımı birlikte düşünülmelidir. Bu tezde izlenen yöntemsel yaklaşım da bu kuramsal çerçeveye dayanarak, tek zamanlı afet sonrası uydu görüntülerinde hasar izlerini ayrıştırmayı hedefleyen bir segmentasyon boru hattını adım adım yapılandırmaktadır.

2.1.1. Evrişimli Sinir Ağları ve Geleneksel Mimari Yaklaşımlar

Evrişimli sinir ağları (convolutional neural networks, CNN) tabanlı segmentasyon yaklaşımlarında temel fikir, görüntüden çok katmanlı bir öznitelik hiyerarşisi çıkarmaktır. İlk katmanlar kenar, doku ve basit geometrik ipuçlarına duyarlı temsil üretirken, daha derin katmanlar nesne parçaları ve sınıf ayrımı için daha ayırt edici örüntüleri taşır. Ancak ağ derinleştikçe yaygın olarak kullanılan havuzlama ve stride işlemleri uzamsal çözünürlüğü düşürür; bu durum sınıflandırma görevlerinde faydalı bir değişmezlik sağlasa da, semantik segmentasyonda konum hassasiyetini zayıflatır. Bu nedenle küçük nesnelerin kaybı, ince yapıların silinmesi ve sınırların bulanıklaşması gibi etkiler literatürde çözünürlük kaybı ve detay kaybı problemi olarak ele alınır. DeepLab

çizgisindeki çalışmalar, özellikle bu aşağı örneklemenin yerelleştirme doğruluğu üzerinde olumsuz bir etkisi olduğuna dikkat çeker [25].

Bu kırılmayı sistematik bir çerçeveye taşıyan temel yaklaşımlardan biri, sınıflandırma ağlarını pikselden piksele çıktı üretecek biçimde dönüştüren Fully Convolutional Networks (FCN) mimarisidir. FCN, tam bağlı katmanları kaldırarak tüm ağı evrişimsel hale getirir ve böylece giriş boyutundan bağımsız olarak yoğun tahmin üretir. Ayrıca farklı katman seviyelerinden gelen temsilleri birleştirme fikrini vurgular; derin katmanların güçlü semantik bilgisini, sığ katmanların daha ince uzamsal detaylarıyla birleştirerek daha keskin maske üretimi hedefler. Bu yaklaşım, modern segmentasyon mimarilerinde yaygınlaşan çok seviyeli füzyon ve skip bağlantısı fikrinin önemli bir başlangıç noktasıdır [23].

Encoder–decoder ailesini popülerleştiren güçlü bir diğer dönüm noktası U-Net mimarisidir. U-Net, aşağı örnekleme yapan bir encoder yoluyla bağlamı toplarken, simetrik bir decoder yoluyla uzamsal çözünürlüğü kademeli biçimde geri kazanır. Aynı çözünürlük seviyesindeki encoder ve decoder katmanlarının birleştirilmesi, hem bağlam bilgisinin hem de lokal detayların birlikte kullanılmasını sağlar; bu sayede özellikle sınırların korunması ve küçük yapıların kaçırılmaması açısından daha kararlı çıktılar üretilebilir. U-Net'in güçlü veri artırma ile birlikte sınırlı etiketli senaryolarda bile etkili olabildiği, orijinal çalışmada özellikle vurgulanmıştır [26].

CNN tabanlı segmentasyonun pratikte sık kullanılan başka bir hattı, verimli geri-örnekleme ve bellek maliyeti açısından optimize edilmiş encoder–decoder tasarımlarıdır. SegNet, encoder tarafındaki max-pooling adımlarında hesaplanan indeksleri decoder tarafında kullanarak daha verimli bir upsampling stratejisi önerir ve bu sayede yoğun tahmin üretimini daha hesaplı hale getirmeyi amaçlar. Bu tip tasarımlar, özellikle büyük çözünürlüklü uzaktan algılama görüntülerinde eğitim ve çıkarım maliyetlerinin yönetimi açısından önem taşır [27].

Çok ölçekli bağlamı yakalamaya odaklanan yaklaşımları da CNN tabanlı segmentasyon literatürünün önemli bir bileşenidir. DeepLab ailesi, atrous evrişim ile görüş alanını büyütürken çözünürlük kaybını kontrol etmeyi ve ASPP modülü ile farklı ölçeklerde bağlamı aynı anda toplamayı hedefler; DeepLabv3+ ise buna ek olarak basit bir decoder modülü ile sınır rafinmanını güçlendirmeyi amaçlar. Bu tasarım, uzaktan algılamada aynı sahnede farklı bina boyutları, farklı yoğunluklar ve değişken arka plan koşulları

bulunması nedeniyle kavramsal olarak özellikle uygundur [25]. Benzer şekilde PSPNet, sahnenin küresel bağlamını daha iyi taşımak için pyramid pooling fikrini öne çıkarır ve bu tür modüllerin geniş ölçekli sahnelerde faydasını gösterir [28].

Özetle CNN tabanlı segmentasyon, yerel örüntüleri yakalama ve çok-seviyeli temsil üretme konusunda oldukça güçlüdür; ancak uzaktan algılamada sık görülen geniş alan bağlamı, büyük ölçekli mekânsal ilişkiler ve uzak bölgeler arası bağımlılıkları modelleme konusunda yapısal sınırlılıklar gösterebilir. Bu zemin, sonraki alt başlıkta ele alınacak Transformer tabanlı yaklaşımların neden yükseldiğini anlamak için önemli bir başlangıç çerçevesi sunar [23].

2.1.2. Kodlayıcı–Kod Çözücü (Encoder–Decoder) Yapısı ve Öznitelik Füzyonu

Semantik segmentasyonda encoder–decoder yaklaşımı, bir görüntüden elde edilen temsilleri iki tamamlayıcı hedef üzerinden düzenleyen temel bir tasarım şeması sunar. Encoder tarafı, "ne var?" sorusunu cevaplayacak şekilde sınıflar arasında ayrım yapmayı kolaylaştıran daha soyut ve daha güçlü semantik öznitelikler üretir. Bu amaçla genellikle ardışık evrişim ve aşağı örnekleme adımları kullanılır; böylece alıcı alan büyütülerek daha geniş bir bağlam yakalanır. Decoder tarafı ise "nerede?" sorusunu hedefleyerek, encoder tarafından üretilen düşük çözünürlüklü fakat güçlü temsilleri kademeli biçimde yükseltir ve çıktı çözünürlüğünü geri kazandırarak piksel düzeyinde konumlandırma yapar. U-Net, bu yaklaşımı açık ve etkili biçimde tanımlayan temel örneklerden biridir; daralan yolun bağlamı, genişleyen yolun ise hassas yerelleştirmeyi hedeflediğini vurgular [26].

Encoder–decoder tasarımını segmentasyonda kritik yapan unsur, tek başına çözünürlük geri kazanımının yeterli olmamasıdır. Encoder boyunca uygulanan aşağı örnekleme, sınıflandırma açısından faydalı olsa da sınırların netliği ve küçük yapıların korunması açısından dezavantaj üretir. Bu noktada decoder yalnızca büyütme işlemi yapan bir bileşen gibi çalışırsa, çıktı maske "genel olarak doğru bölgede" olsa bile sınırlar yumuşayabilir veya küçük ayrıntılar kaybolabilir. Bu nedenle modern segmentasyon mimarileri, decoder'ın encoder'dan gelen bilginin yalnızca ölçeğini değil içeriğini de rafine edebilmesini sağlayacak mekanizmalara ihtiyaç duyar. DeepLabv3+ çalışması bu iki yaklaşımı açık biçimde özetler: çok ölçekli bağlam toplama modülleri bağlamı güçlendirirken, encoder–decoder yapısı uzamsal bilgiyi kademeli olarak geri kazandırarak daha keskin sınırlar üretmeye yardımcı olur [29].

Bu şemanın merkezindeki kavram öznitelik füzyonudur. Tek seviyeden öznitelik kullanmak segmentasyonun iki temel gereksinimini aynı anda karşılamakta zorlanır. Yalnızca derin katmanlardan gelen öznitelikleri kullanmak sınıf ayrımını güçlendirebilir; ancak bu temsiller düşük çözünürlüklü olduğu için sınır doğruluğu ve ince detaylar zayıflayabilir. Buna karşılık yalnızca sığ katmanlara ağırlık vermek sınırları görece daha iyi taşısa da bağlam zayıf kalabilir ve özellikle karmaşık arka planlarda sınıf karışmaları artabilir. FCN yaklaşımı bu dengeyi kurmak için farklı derinliklerden gelen temsilleri birleştiren bir "skip" mimarisini ön plana çıkarır; derin katmandaki kaba ama semantik güçlü temsili, sığ katmandaki ince ama daha zayıf semantik temsille birleştirerek daha ayrıntılı segmentasyon hedefler [23]. U-Net ise aynı fikri encoder ve decoder arasında eş çözünürlük seviyelerinde doğrudan bağlantılar kurarak uygular ve bu sayede decoder'ın yerel detayları yeniden öğrenmek yerine encoder'dan hazır olarak almasını sağlar [26].

Öznitelik füzyonu, yalnızca erken katmanla geç katmanı birleştirme olarak düşünülmemelidir. Pratikte bu füzyon, çok ölçekli temsillerin aynı tahmin başlığına taşınması, farklı çözünürlüklerdeki öznitelik haritalarının bir araya getirilmesi ve sınıf kararının hem bağlam hem de sınır bilgisiyle desteklenmesi anlamına gelir. Bu sayede model, örneğin bir binanın çatı dokusunu veya enkaz görünümünü yerel paternlerle yakalarken, çevresindeki kentsel doku ve komşu yapılarla ilişkisini de bağlamsal olarak değerlendirebilir. Uzaktan algılama verilerinde bu birleşim özellikle önemlidir; çünkü aynı hasar sınıfı farklı bölgelerde farklı dokusal ipuçları gösterebilir ve sadece yerel dokuya bakmak, zemin, gölge veya araç yolları gibi yapıların yanlış pozitif üretilmesine yol açabilir. Bu nedenle encoder–decoder ve öznitelik füzyonu, uzaktan algılama temelli hasar segmentasyonu problemlerinde yaygın ve varsayılan bir omurga yaklaşımı olarak öne çıkmıştır.

Decoder tarafında çözünürlük geri kazanımı farklı biçimlerde uygulanabilir. En temel yöntemler interpolasyon tabanlı büyütme, transpoze evrişim veya öğrenilebilir upsampling bloklarıdır. Hangi yöntem seçilirse seçilsin, amaç yalnızca piksel sayısını artırmak değil; büyütülen uzayda sınıf sınırlarını doğru konumlandıracak ayırt edici bilgiyi taşımaktır. DeepLabv3+ örneğinde, decoder modülünün özellikle nesne sınırları boyunca sonucu rafine etmek için eklendiği açıkça belirtilir; bu, encoder–decoder tasarımının sınır kalitesiyle doğrudan ilişkilendirildiğini gösterir [29].

Sınır doğruluğu konusunun segmentasyon performansına etkisi, kullanılan değerlendirme metrikleriyle daha görünür hale gelir. Özellikle IoU gibi metrikler, tahmin maskesi ile yer gerçeği maskesi arasındaki örtüşmeye dayanır ve sınır boyunca oluşan sistematik kaymalar, örtüşme alanını azaltarak metriği belirgin biçimde düşürebilir. Bu durum, bina ve hasar gibi geometrik sınırları görece keskin sınıflarda daha kritiktir; çünkü küçük bir sınır hatası bile nesnenin alanına oranla anlamlı bir fark yaratabilir. Nitekim sınır hatalarına duyarlılığı daha doğrudan ölçmeyi hedefleyen Boundary IoU çalışması, standart IoU ölçümünün sınır hatalarına karşı davranışını analiz eder ve sınır kalitesinin segmentasyon değerlendirmesinde ayrı bir boyut olduğunu vurgular [30]. Bu nedenle encoder–decoder tasarımı ve öznitelik füzyonu, yalnızca görsel kaliteyi artıran tasarım tercihleri değil, aynı zamanda metrikler üzerinden ölçülen başarımı doğrudan etkileyen metodolojik bileşenler olarak görülmelidir.

Bu tez bağlamında encoder–decoder yaklaşımı ve öznitelik füzyonu, tek zamanlı afet sonrası uydu görüntülerinden hasar izlerini ayrıştırma hedefi için kritik bir temel sunar. Tek zamanlı senaryoda modelin referans olarak afet öncesi görüntüden faydalanamaması, sınırların ve dokusal ipuçlarının doğru yorumlanmasını daha da önemli hale getirir. Bu nedenle takip eden alt başlıklarda CNN tabanlı mimarilerde bu yapının nasıl kurulduğu, Transformer tabanlı yaklaşımlarda çok ölçekli temsillerin nasıl ele alındığı ve seçilen mimarinin bu problemdeki rolü daha ayrıntılı şekilde tartışılacaktır.

2.2. Transformer Tabanlı Yaklaşımlar ve Vision Transformer (ViT) Ailesi

Derin öğrenmede CNN tabanlı mimariler uzun süre görsel görevlerde standart yaklaşım olmuş olsa da, özellikle geniş alan bağlamının kritik olduğu problemlerde uzun menzilli bağımlılıkları yakalamada yapısal sınırlılıklar gösterebilir. CNN'ler yerel alıcı alanlar üzerinden ilerlediği için, uzaktaki bölgeler arasındaki ilişkiyi kurmak çoğu zaman çok sayıda katman üzerinden dolaylı biçimde gerçekleşir. Bu durum, uzaktan algılama görüntülerinde sık görülen büyük ölçekli sahnelerde ve geniş alan örüntülerinde belirginleşir; örneğin hasarın mahalle ölçeğinde kümelenmesi, aynı sokak hattındaki yapıların benzer etkilenme paterni veya büyük bir blok içindeki tekrar eden bina dokuları gibi durumlarda yalnızca yerel filtrelerle temsil kurmak zorlaşabilir.

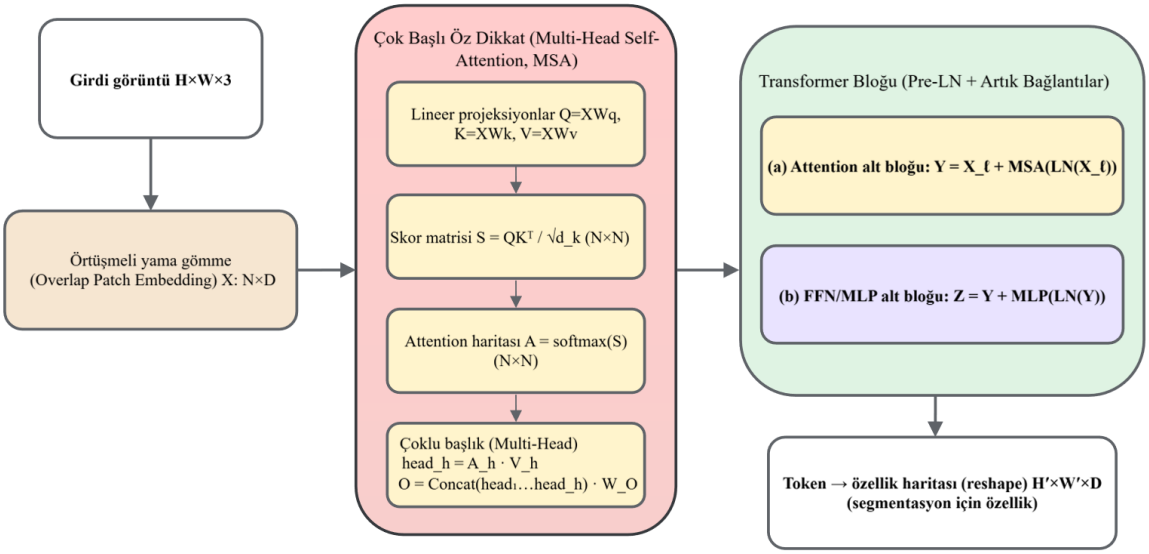
Transformer mimarisi, dizisel verilerde dikkat mekanizmasıyla token'lar arası etkileşimi doğrudan modelleyerek bu tür bağımlılıkları güçlü biçimde temsil eder ve bu yaklaşımın

temelini self-attention mekanizması oluşturur [31]. Self-attention, her bir token'ın diğer tüm token'larla ilişkisini öğrenerek bilgi akışını sadece komşuluk üzerinden değil, küresel bir bağlam üzerinden kurar [31]. Bu nedenle birbirinden uzak bölgelerin birlikte değerlendirilmesi gereken sahnelerde, modelin karar verirken daha geniş bir bağlamı tek adımda dikkate alabilmesi teorik olarak avantaj sağlar. Uzaktan algılama bağlamında bu avantaj, tek tek piksellerin veya küçük yamaların ötesinde sahne düzenini, yapı kümelerini ve çevresel bağlamı birlikte yorumlama ihtiyacında karşılık bulur.

Transformer'ların görsel görevlere uyarlanmasında temel fikir, görüntüyü işlenebilir bir token dizisine dönüştürmektir. Bu dönüşüm sayesinde, görüntüdeki her token diğer token'larla etkileşime girebilir ve model sınıf kararını yalnızca yerel benzerliklere değil, küresel ilişkiler ağına dayanarak şekillendirebilir. Ancak görsel veride token sayısı metne kıyasla çok hızlı arttığı için, bu yaklaşımın pratikte verimli hale getirilmesi kritik bir tasarım hedefi olarak ortaya çıkar. Bu noktada ViT ailesi, görüntüyü sabit boyutlu parçalara ayırma ve bu parçaları token olarak ele alma yaklaşımıyla Transformer encoder'ı doğrudan görsel alana taşıyan en temel çerçevelerden biridir [32].

2.2.1. Self-Attention Mantığı ve Görüntüye Uyarlanması

Transformer'ın temel yapıtaşı olan self-attention, sorgu, anahtar ve değer temsillerini kullanarak ağırlıklı bir toplam üretir. Vaswani ve arkadaşlarının tanımladığı ölçeklenmiş nokta çarpımı dikkat mekanizması şu biçimde ifade edilir: $\text{Attention}(Q, K, V) = \text{softmax}(QK^T / \sqrt{d_k})\, V$ [31]. Burada Q, K ve V matrisleri giriş temsillerinin farklı doğrusal dönüşümleridir ve softmax ile elde edilen dikkat ağırlıkları, her token'ın diğer token'lardan ne ölçüde bilgi alacağını belirler [31]. Çok başlıklı dikkat tasarımı ise bu etkileşimi tek bir ilişki türüne indirgemek yerine, farklı alt-uzaylarda paralel dikkat başlıklarıyla çeşitli ilişki örüntülerini aynı anda yakalamayı hedefler [31]. Görsel bağlamda bu, aynı sahnede hem benzer doku bölgelerini, hem geometriyle ilişkili sınırları, hem de sahne düzenine dair daha küresel ilişkileri farklı başlıklar üzerinden temsil edebilme potansiyeli anlamına gelir.



Şekil 2.1. Self-Attention mantığı ve görüntüye uyarlanması (Kaynak [31]'dan uyarlanmıştır.).

Şekil 2.1'de [31], self-attention mekanizmasının görüntü verisine uyarlanmış temel işlem hattı özetlenmektedir. Girdi görüntüden elde edilen token temsilleri üzerinden sorgu (Q), anahtar (K) ve değer (V) projeksiyonları üretilir; ardından Q–K benzerlikleri üzerinden dikkat ağırlıkları hesaplanır ve bu ağırlıklar kullanılarak token temsilleri yeniden ağırlıklandırılır. Çoklu başlık (multi-head) yapısında paralel olarak yürütülen bu işlem, başlık çıktılarının birleştirilmesiyle tek bir temsil hâline getirilir ve artık bağlantılar ile katman normalizasyonu içeren Transformer bloğundan geçirilerek segmentasyon için kullanılacak özellik temsilleri elde edilir. Burada N token sayısını, D gömme (embedding) boyutunu, d ise her bir başlığın (head) boyutunu ifade eder. Örtüşmeli yama gömme yerel sürekliliği koruyarak küçük detayların yakalanmasını desteklerken, çoklu başlık yapısı farklı uzamsal ilişkilerin aynı anda öğrenilmesine olanak tanır [31].

Görüntüye uyarlamada ilk zorluk, görsel verinin düzenli bir ızgara yapısına sahip olması ve konumsal bilginin doğrudan önemli olmasıdır. CNN'lerde konum bilgisi yerel filtrelerin kaydırmalı uygulanmasıyla örtük biçimde taşınırken, Transformer token dizisi üzerinde çalıştığı için konum bilgisinin ayrıca temsile eklenmesi gerekir. ViT yaklaşımı, görüntüyü eş boyutlu patch'lere ayırıp her patch'i bir token olarak ele alır; ardından bu token'lara konumsal bilgiyi taşıyan eklemeler yaparak Transformer encoder katmanlarına besler [32]. Bu sayede model, hem patch içeriklerini hem de patch'lerin sahnedeki göreli düzenini dikkate alarak temsil öğrenebilir. Patch tabanlı temsilin bir sonucu olarak, model

artık pikseller yerine patch düzeyinde etkileşim kurar; bu, hesaplama açısından daha yönetilebilir bir token sayısı üretirken, çok küçük ayrıntıların temsilini patch boyutuna bağlı hale getirir.

Self-attention'ın görsel alandaki ikinci temel zorluğu hesaplama maliyetidir. Token sayısı N olduğunda, standart self-attention mekanizmasının hesaplama ve bellek maliyeti QK^T etkileşimi nedeniyle yaklaşık N^2 ölçeğinde büyür. Yüksek çözünürlüklü görüntülerde N hızla arttığından, bu maliyet pratik bir sınırlılık haline gelir. Bu nedenle bilgisayarlı görü için geliştirilen birçok Transformer varyantı, maliyeti düşürmek ve çok ölçekli temsilleri daha doğal biçimde üretmek için hiyerarşik veya yerel dikkat temelli tasarımlara yönelmiştir. Swin Transformer, self-attention'ı pencere tabanlı yerel bölgelerle sınırlandırarak maliyeti düşürmeyi ve kaydırılmış pencere şemasıyla pencereler arası bilgi akışını güçlendirmeyi öneren önemli örneklerden biridir [33]. Bu tür hiyerarşik tasarımlar, uzaktan algılama gibi çok ölçekli sahnelerde hem verimlilik hem de farklı çözünürlük seviyelerinde özellik üretimi açısından özellikle anlamlıdır.

ViT (Vision Transformer) yaklaşımı, görüntüyü patch'lere ayırıp bu patch'leri token olarak ele alarak Transformer encoder ile işler ve CNN'in yerel filtrelerine doğrudan ihtiyaç duymadan küresel bağlamı öğrenebileceğini gösterir [32]. Ancak ViT'in orijinal formu, güçlü performansı büyük ölçüde büyük ölçekli ön eğitim ve uygun aktarım stratejileriyle birlikte sergiler. Bu nedenle veri ölçeği ve eğitim stratejileri, ViT ailesinin pratik başarımında belirleyici faktörlerdir [32]. Daha sınırlı veri senaryolarında Transformer'ların nasıl daha verimli eğitilebileceği üzerine çalışmalar da bu motivasyonla ortaya çıkmıştır; DeiT çalışması, ImageNet ölçeğinde veri bazında daha verimli eğitim ve distilasyon stratejileriyle ViT'in rekabetçi biçimde eğitilebileceğini gösteren örneklerden biridir [34].

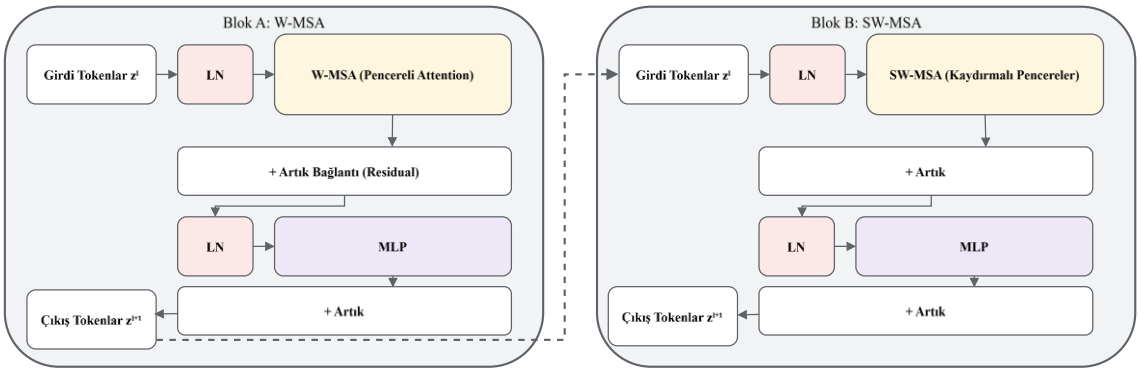
ViT ile birlikte Transformer tabanlı modellerin yalnızca sınıflandırma değil, segmentasyon gibi yoğun tahmin problemlerine uyarlanması da hız kazanmıştır [32]. Bununla birlikte, yoğun tahminde piksellerin uzamsal düzeninin korunması, çok ölçekli özelliklerin üretilmesi ve yüksek çözünürlüklü çıktı ihtiyacı nedeniyle, saf bir patch dizisi temsilinin doğrudan kullanımı çoğu zaman ek tasarım kararları gerektirir. Bu gereksinimler, bir sonraki aşamada ele alınacak hiyerarşik Transformer omurgalarına ve segmentasyon odaklı tasarımlara zemin hazırlar; özellikle çok ölçekli temsil üretimi ve

verimli dikkat mekanizmaları, uzaktan algılama tabanlı hasar segmentasyonu gibi problemler için belirleyici hale gelir.

2.2.2. Hiyerarşik Transformer'lar: Swin ve PVT Çizgisi

Segmentasyon gibi yoğun tahmin görevlerinde modelin iki şeye aynı anda ihtiyacı vardır: yüksek çözünürlüklü yerel detay (özellikle sınırlar) ve farklı ölçeklerde bağlam (bina kümeleri, mahalle dokusu, çevresel örüntüler gibi). Bu nedenle, tek bir çözünürlükte ilerleyen düz token dizisi yaklaşımı pratikte sınırlayıcı kalabilmekte; bunun yerine çok ölçekli özellik haritaları üreten hiyerarşik Transformer omurgaları öne çıkmaktadır. Hiyerarşik yapıların temel avantajı, CNN'lerdeki özellik piramidine benzer biçimde, düşük seviyelerde daha yüksek çözünürlük ve detay; üst seviyelerde daha düşük çözünürlük fakat daha güçlü bağlam üretmeyi doğal şekilde sağlamasıdır.

Swin Transformer bu ihtiyaca doğrudan yanıt veren mimarilerden biridir. Swin'in ana fikri, self-attention hesaplamasını tüm görüntü üzerinde global yapmak yerine, görüntüyü yerel pencerelere bölerek attention'ı pencere içinde hesaplamak ve bir sonraki katmanda pencereleri kaydırarak pencereler arası bilgi alışverişini mümkün kılmaktır. Bu pencere tabanlı dikkat (window-based attention) ve kaydırılmış pencere (shifted window) tasarımı, hesaplama maliyetini kontrol altında tutarken katmanlar boyunca daha geniş bağlama ulaşılmasını sağlar. Swin ayrıca patch merging benzeri aşamalarla çözünürlüğü kademeli olarak düşürerek hiyerarşik temsil üretir ve bu yönüyle nesne tespiti ve semantik segmentasyon gibi görevlerde genel amaçlı bir backbone olarak konumlandırılır. Swin makalesi, bu tasarımın görüntü boyutuna göre daha ölçeklenebilir bir maliyet profili sunduğunu ve dense prediction görevlerinde güçlü sonuçlar verdiğini raporlamaktadır [33].

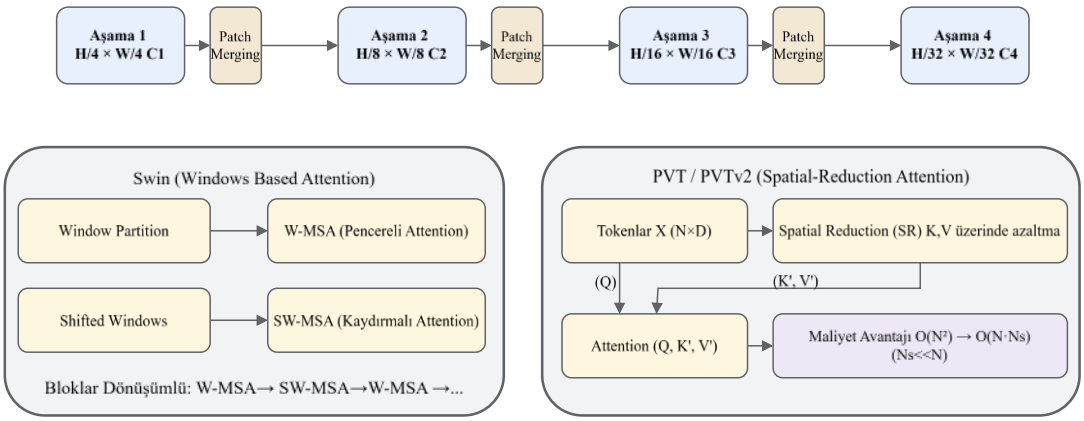


Şekil 2.2. Swin Transformer blok yapısı (W-MSA → SW-MSA) ve artık bağlantılar (Kaynak: [32]'den uyarlanmıştır.).

Şekil 2.2, Swin Transformer'da bir Swin bloğunun (W-MSA → SW-MSA) iç işleyişini ve her alt adımda kullanılan artık bağlantıları göstermektedir. Blok girişindeki belirteçler (tokenlar) önce Katman Normalizasyonu (LN) ile ölçeklenmekte, ardından pencere tabanlı çok başlıklı öz-dikkat mekanizması (W-MSA) uygulanmaktadır. W-MSA çıktısı, bilgi kaybını azaltmak ve gradyan akışını güçlendirmek amacıyla girişle toplanarak artık bağlantı (+) üzerinden birleştirilir. Devamında aynı yapı, pencerelerin konumunun kaydırılmasıyla komşu pencereler arası etkileşimi artıran kaydırılmış pencere dikkat modülü (SW-MSA) ile tekrarlanır. Her iki dikkat adımından sonra, tekrar LN uygulanır ve uzamsal olarak karışan temsilin kanal-boyutunda zenginleştirilmesi için MLP/FFN (çok katmanlı algılayıcı / beslemeli ağ) katmanı kullanılır; MLP çıktısı da yine artık bağlantı ile birleştirilerek blok çıkış belirteçleri elde edilir. Bu şekil, Swin'in temel farkını oluşturan W-MSA ile SW-MSA'nın ardışık kullanımı ve bunun üzerine kurulu Pre-LN + artık bağlantı tasarımını özetlemektedir. Hiyerarşik yapıda çözünürlüğün aşamalı azaltılması ve kanal boyutunun artırılmasını sağlayan Patch Merging (yama birleştirme) ile aşama geçişleri şekil 2.3'de ayrıca gösterilmektedir.

Pyramid Vision Transformer (PVT) çizgisi ise hiyerarşik temsili "piramidal" bir kurgu ile ele alır. PVT'nin motivasyonu, CNN omurgalarında standart hale gelmiş çok ölçekli piramit özellik çıkarımını Transformer alanına taşımak ve dense prediction görevleri için doğrudan kullanılabilir bir özellik seti üretmektir. PVT, token sayısını aşamalı olarak azaltıp çözünürlüğü kademeli şekilde düşürerek çok ölçekli bir yapı oluşturur. Ayrıca self-attention maliyetini azaltmak için spatial-reduction attention (SRA) fikrini önerir; böylece yüksek çözünürlükte token sayısının hızla büyümesinden doğan maliyet baskısını

hafifletmeyi amaçlar. PVT çalışması, bu tasarımla Transformer tabanlı omurgaların segmentasyon ve benzeri yoğun tahmin problemlerine daha uygun hale getirilebileceğini vurgular [35].



Şekil 2.3. Hiyerarşik Transformer'lar: Swin ve PVT çizgisi (Çok-Ölçekli Temsil) (Kaynak: [33, 35]'den uyarlanmıştır.).

Şekil 2.3'de, hiyerarşik Transformer kodlayıcı yaklaşımı bağlamında Swin ve PVT/PVTv2 çizgisinin çok ölçekli temsil mantığı özetlenmektedir. N token sayısı, Dgömme boyutu, N_sSR sonrası azaltılmış token sayısı olmak üzere, üst kısımda patch merging ile çözünürlüğün aşamalı olarak azalırken kanal sayısının artması gösterilmiştir. Swin tarafında dikkat, pencereler içinde (W-MSA) hesaplanıp kaydırmalı pencereler (SW-MSA) ile pencereler arası etkileşim desteklenirken; PVT/PVTv2'de uzamsal azaltma (SR) ile özellikle Kve Vüzerinde token sayısı düşürülerek dikkat hesaplamasının maliyeti $O(N^2)$yerine $O(N \cdot N_s)$seviyesine indirilmektedir.

Uzaktan algılama bağlamında bu iki çizgi de kavramsal olarak önemlidir. Çünkü uydu görüntülerinde aynı sahnede hem küçük yapıların hem büyük yapı bloklarının bulunması, arka planın çok çeşitli olması ve görüntüleme koşullarının değişkenliği, modeli çok ölçekli düşünmeye zorlar. Hiyerarşik Transformer omurgaları bu çeşitliliği tek bir temsil düzeyine sıkıştırmak yerine, farklı çözünürlük seviyelerinde tamamlayıcı özellikler üreterek segmentasyon başlığının işini kolaylaştırır. Bu özellik, özellikle yüksek çözünürlüklü afet sonrası görüntülerde hem hız hem bellek hem de doğruluk dengesini kurmak açısından pratik bir avantaj sağlar [33].

2.2.3. Segmentasyon İçin Transformer Yaklaşımları

Transformer'ların segmentasyona uygulanmasında iki ana yönelim öne çıkar. Birinci yaklaşımda Transformer, özellik çıkarıcı bir encoder olarak kullanılır ve ardından piksel düzeyinde sınıf tahmini üreten bir segmentasyon başlığı ile dense prediction'a inilir. İkinci yaklaşımda ise uçtan uca Transformer tabanlı, maskeyi doğrudan üreten tasarımlar geliştirilir; bu tasarımlar genellikle attention tabanlı mask üretimi veya mask token'ları üzerinden sınıf maskesi çıkarımı gibi stratejiler içerir.

Segmenter, ViT tabanlı temsilin semantik segmentasyona doğrudan uyarlanabileceğini gösteren erken ve etkili örneklerden biridir. Segmenter, ViT benzeri bir encoder'dan gelen patch embedding'leri üzerine basit bir doğrusal decoder ile piksel etiketlemesi yapabilmenin mümkün olduğunu, ayrıca mask transformer decoder ile sınıf maskesi üretimini güçlendirmenin performansı artırabildiğini raporlar. Buradaki önemli nokta, global bağlamın daha ilk katmandan itibaren temsilin içine taşınabilmesi ve segmentasyon gibi belirsizliği yüksek görevlerde bu bağlamın karar tutarlılığına katkı sağlayabilmesidir [36].

Bununla birlikte pratikte segmentasyon için belirleyici konu, doğruluk ile hesaplama maliyeti ve hız arasındaki dengedir. Özellikle yüksek çözünürlüklü görüntülerde, hem encoder'ın maliyeti hem de decoder'ın maliyeti hızla artabilir. Bu nedenle, literatürde ağır decoder'lar yerine daha sade decoder tasarımları, çok ölçekli özellikleri verimli biçimde birleştiren füzyon stratejileri ve attention'ı daha yönetilebilir hale getiren omurga tasarımları öne çıkmıştır. Bu eğilim, doğrudan senin problemin gibi yüksek çözünürlüklü uydu verilerinde daha da kritik hale gelir; çünkü patch sayısı arttıkça maliyet hızla büyür ve gerçekçi eğitim-çıkarım süreleri için tasarımın verimli olması gerekir [33].

SegFormer bu noktada basit ve verimli bir segmentasyon tasarımı sunar. SegFormer'ın encoder tarafı, Mix Transformer (MiT) omurgasıyla hiyerarşik ve çok ölçekli özellikler üretir; decoder tarafında ise ağır bir encoder–decoder yapısı yerine, çok katmanlı algılayıcı tabanlı hafif bir füzyon/çıktı üretim modülü kullanır. Makale, bu sade tasarımın temel avantajlarını iki başlık altında açıklar: hiyerarşik encoder ile çok ölçekli özellik üretimi ve karmaşık decoder'lardan kaçınılarak verimliliğin korunması. Ayrıca SegFormer, belirli bir konumsal kodlama tasarımına bağımlı olmadan farklı çözünürlüklerde daha kararlı davranmayı hedefleyen bir kurgu vurgular [19]. Bu

özellikler, özellikle eğitim ve test çözünürlüğünün tam örtüşmediği veya sahada farklı çözünürlüklerin karşılaşıldığı senaryolarda pratik bir artı olarak değerlendirilebilir.
Özetle, segmentasyon için Transformer yaklaşımında yalnızca Transformer kullanmak değil, bu Transformer temsilini dense prediction'a verimli biçimde indirgemek ve çok ölçekli bilgiyi decoder tarafında sade ama etkili şekilde birleştirmek kritik hale gelmiştir. SegFormer'ın seçimi de bu pratik gereksinimlerle doğrudan örtüşür [19].

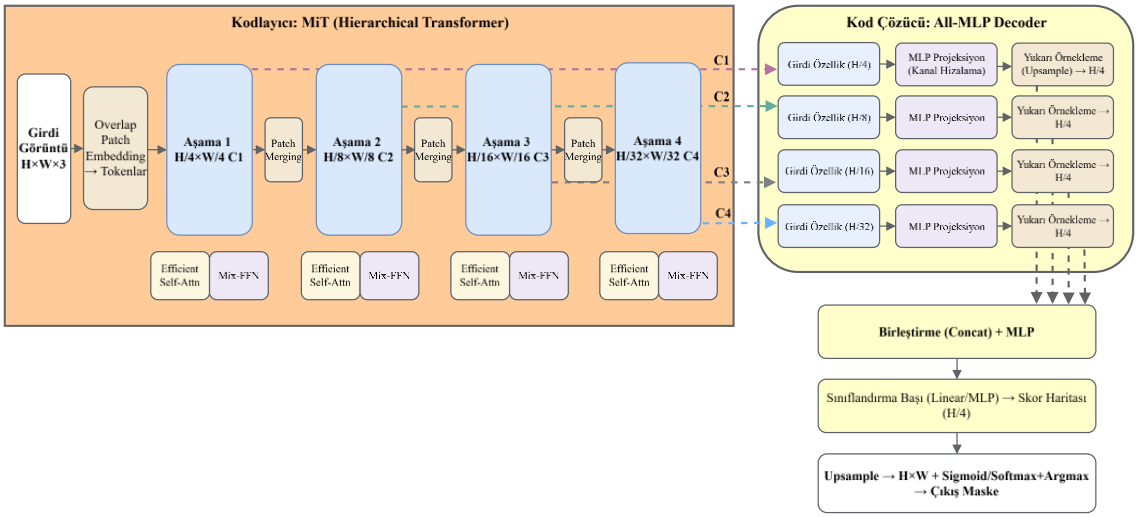
2.2.4. Transformer Tercihinin Gerekçesi

Deprem sonrası hasar tespiti probleminde yıkım izleri çoğu zaman yalnızca tekil bina sınırlarıyla tanımlanamaz; komşu binaların durumu, sokak çizgileri, boşluk ve enkaz alanları, gölge ve ışık koşulları gibi çevresel bağlamla birlikte anlam kazanır. Tek zamanlı, sadece afet sonrası görüntü senaryosunda bu ihtiyaç daha da belirginleşir; çünkü model, "önce–sonra farkı" gibi güçlü bir ipucuna sahip değildir ve kararını doğrudan sahnenin mevcut görünümünden üretmek zorundadır. Bu yüzden, uzak bölgeler arası ilişkileri aynı temsil içinde değerlendirebilme kapasitesi, bu çalışmanın probleminde metodolojik olarak değerli bir motivasyon oluşturur. Transformer mimarisi, self-attention ile token'lar arası küresel etkileşimi doğrudan modelleyebildiği için bu tür bağlamsal ilişkileri temsil etme potansiyeline sahiptir [33].
Bu tez kapsamında SegFormer'ın seçilmesi üç ana gerekçeyle tutarlı bir tercih olarak değerlendirilebilir. Birincisi, MiT omurgasıyla çok ölçekli özellik üretmesidir; bu, uydu görüntülerindeki ölçek çeşitliliği ve kentsel doku değişkenliği için uygun bir zemindir. İkincisi, segmentasyon odaklı tasarlanmış bir çerçeve olmasıdır; yani Transformer'ı yalnızca bir encoder olarak almak yerine, doğrudan segmentasyon gereksinimleriyle uyumlu bir mimari bütünlük sunar. Üçüncüsü, hafif decoder tasarımı sayesinde yüksek çözünürlüklü görüntülerde pratik verimlilik hedeflemesidir; bu da operasyonel senaryolara yakın bir kullanım hedefiyle örtüşür. SegFormer çalışması bu tasarım ilkelerini açık biçimde ortaya koyar ve farklı model ölçeklerinde doğruluk–verimlilik dengesini raporlar [19].
Son olarak, Vision Transformer ailesinin hızla gelişmesi ve segmentasyona yönelik hiyerarşik omurga tasarımlarının yaygınlaşması, bu yaklaşımı güncel ve güçlü bir araştırma hattı haline getirmiştir. Swin ve PVT gibi omurgaların dense prediction görevlerinde backbone olarak konumlandırılması, Transformer tabanlı özellik

çıkarımının pratikte olgunlaşan bir çizgiye dönüştüğünü göstermektedir. Bu çerçeve içinde SegFormer, hem akademik literatürde güçlü bir referans noktası hem de tez kapsamında kurulacak sistematik deney hattı için uygun bir ana mimari seçimi sunmaktadır [33].

2.3. SegFormer Mimarisi: Verimli Bir Segmentasyon Tasarımı

SegFormer, semantik segmentasyon problemini Transformer tabanlı bir omurga ile ele alırken, klasik segmentasyon mimarilerinde sık görülen ağır ve maliyetli çözücü tasarımlarına bağımlı kalmadan güçlü sonuçlar üretmeyi hedefleyen bir yaklaşımdır [19]. Temel fikir, görüntüden farklı çözünürlük seviyelerinde çok ölçekli öznitelikler üreten hiyerarşik bir Transformer kodlayıcı kullanmak ve bu öznitelikleri çok hafif bir çok katmanlı algılayıcı tabanlı çözücü ile birleştirerek piksel düzeyinde sınıf tahmini üretmektir [19].



Şekil 2.4. SegFormer: MiT hiyerarşik kodlayıcı + All-MLP decoder (Verimli Segmentasyon) (Kaynak: [19]'den uyarlanmıştır.).

Şekil 2.4 [19], SegFormer mimarisinin genel akışını özetlemektedir. Girdi görüntü, MiT tabanlı hiyerarşik Transformer kodlayıcı tarafından dört farklı ölçekte özellik haritasına ayrıştırılır ve bu temsiller C1'den C4'e doğru çözünürlük azalırken kanal derinliği artacak şekilde elde edilir. Şekilde kesikli oklarla gösterilen bağlantılar, çok ölçekli özelliklerin kod çözücüye taşınarak birlikte kullanılmasını ifade eder. All-MLP kod çözücü aşamasında her bir ölçekteki özellik haritası önce MLP projeksiyon ile kanal uzayında hizalanır, ardından ortak bir ölçekte yukarı örneklenerek birleştirme adımına aktarılır.

Birleştirilen temsil MLP ile işlenir ve sınıflandırma başlığı üzerinden bir skor haritası üretilir; son aşamada bu harita çıktı maske çözünürlüğüne yükseltilir ve ikili sınıflamada sigmoid, çok sınıflı durumda ise softmax/argmax ile nihai maske elde edilir. Bu tasarım, ağır kod çözücüler yerine hafif All-MLP kod çözücü kullanarak hesaplama verimliliğini hedefler; MiT kodlayıcı ise hiyerarşik yapısı sayesinde çok ölçekli temsilleri üretir ve örtüşmeli yama gömme yerel sürekliliği koruyarak küçük detayların taşınmasına katkı sağlar [19]. Bu mimari tercihlerin pratik karşılığı, doğruluk ile hesaplama maliyeti arasındaki dengeyi daha yönetilebilir hâle getirmesidir.

Bu yaklaşım, özellikle uydu ve drone gibi yüksek çözünürlüklü görüntülerde segmentasyon yapılırken önem kazanan doğruluk–hız dengesine doğrudan karşılık gelir; çünkü çözünürlük arttıkça hem bellek hem hesaplama maliyeti hızla yükselir ve modelin pratikte kullanılabilirliği zorlaşır.

SegFormer'ın tasarım felsefesi, Transformer'ların küresel bağlamı güçlü biçimde modelleyebilmesine karşın, dikkat hesaplamasının belirteç sayısıyla hızlı büyümesi sorununa verimli bir yanıt üretmektir [19, 31]. Bu nedenle mimari, tek bir çözünürlükte ilerleyen düz bir temsil yerine, çözünürlüğü katmanlar boyunca kademeli azaltan hiyerarşik bir kurgu benimser; böylece hem daha geniş bağlam yakalanır hem de çok ölçekli öznitelikler doğal biçimde elde edilir [19]. Bu genel mantık, segmentasyon gibi yoğun tahmin görevlerinde öne çıkan hiyerarşik Transformer omurgalarının (Swin ve PVT çizgisi) yaklaşımıyla da uyumludur; her iki çizgi de farklı çözünürlük seviyelerinde temsil üreterek hem detay hem bağlam bilgisini aynı anda taşımayı amaçlar [28, 33].

SegFormer'ın pratik açıdan önemli bir diğer noktası da, farklı çözünürlüklerde test yapıldığında performansın daha kararlı olmasına dönük tasarım tercihlerini vurgulamasıdır. Makalede konumsal kodlama gereksinimini azaltan bir kurgu tercih edilerek, eğitim ve test çözünürlükleri birebir aynı olmadığında ortaya çıkabilen uyum sorunlarının önüne geçilmesinin hedeflendiği belirtilir [19]. Bu tür bir esneklik, uzaktan algılamada sahaya yakın senaryolar için özellikle değerlidir; çünkü pratikte görüntü çözünürlüğü, çekim koşulları ve sahne ölçeği değişkenlik gösterebilir.

2.3.1. Hiyerarşik Transformer Kodlayıcı ve Örtüşmeli Yama Birleştirme

Görüntünün Transformer tarafından işlenebilmesi için, görüntünün önce belirteçlere dönüştürülmesi gerekir. Klasik ViT yaklaşımında bu dönüşüm çoğu zaman örtüşmeyen

yamalar üzerinden yapılır; ancak semantik segmentasyon gibi yoğun tahmin görevlerinde komşu bölgeler arasındaki süreklilik ve sınır bilgisi kritik olduğundan, SegFormer'ın MiT omurgası örtüşmeli yama gömme yaklaşımını kullanır [19]. Bu sayede yan yana bölgeler arasındaki ilişki tamamen kopmadan, daha "görüntüye yakın" bir başlangıç temsili elde edilir. Özellikle bina sınırları gibi keskin geçişlerin bulunduğu problemlerde, bu tür bir giriş temsili sınırların ve ince yapıların daha tutarlı taşınmasına yardımcı olabilecek bir zemin sağlar.

MiT omurgası, birden fazla aşamada çalışarak CNN'lerdeki özellik piramidi fikrine benzer şekilde farklı çözünürlük seviyelerinde öznitelik haritaları üretir [19]. Alt aşamalarda daha yüksek çözünürlükte ve daha detaylı öznitelikler korunurken, üst aşamalara doğru çözünürlük kademeli olarak düşürülür ve daha güçlü bağlam bilgisi taşıyan temsiller elde edilir. Bu çok ölçekli üretim, uzaktan algılama için doğrudan anlamlıdır; çünkü aynı sahnede hem küçük detaylar (bina kenarı, küçük enkaz izleri) hem de daha geniş örüntüler (bina kümeleri, blok dokusu, yıkımın çevresel paterni) birlikte değerlendirilmek zorundadır. Hiyerarşik kodlayıcı yaklaşımı, bu ihtiyacı tek bir temsil düzeyine sıkıştırmadan çözerek çözücüye farklı ölçeklerden tamamlayıcı bilgi aktarır [19].

Transformer tabanlı omurgalarda en önemli pratik sorunlardan biri, yüksek çözünürlükte belirteç sayısının çok büyümesi ve bunun dikkat hesaplamasını pahalı hale getirmesidir. SegFormer bu maliyeti kontrol altında tutmak için, özellikle erken aşamalarda dikkat hesabını daha verimli hale getiren bir azaltma stratejisi kullanır. Makalede bu yaklaşım, uzamsal azaltmalı dikkat (SRA) olarak ele alınır ve ana fikir, dikkat hesabında kullanılan bazı temsillerin uzamsal çözünürlüğünü düşürerek hesaplama yükünü sınırlamaktır [19]. Böylece hem yüksek çözünürlükten gelen detay tamamen kaybedilmez, hem de eğitim ve çıkarım maliyeti pratik sınırlar içinde tutulabilir. Bu tür verimlilik odaklı tasarım tercihleri, özellikle uydu veya drone görüntülerinde yama sayısı arttıkça maliyetin hızla büyümesi nedeniyle, mimarinin sahaya yakın kullanımını daha gerçekçi hale getirir.

Özetle SegFormer'ın MiT kodlayıcısı, örtüşmeli yama temsiliyle yerel sürekliliği daha iyi koruyan bir giriş sağlar; hiyerarşik yapısıyla çok ölçekli öznitelikler üretir; verimlilik odaklı dikkat tasarımıyla da yüksek çözünürlükte maliyetin kontrol altında tutulmasına katkı verir [19]. Bu birleşim, tezin problem tanımında yer alan tek zamanlı afet sonrası

görüntülerde hasar izlerinin çıkarımı gibi bağlam ve sınır duyarlılığı yüksek bir görev için metodolojik olarak güçlü bir gerekçe oluşturur.

2.3.2. Tümüyle MLP Tabanlı Kod Çözücü (All-MLP Decoder) ve Hesaplama Verimliliği

SegFormer'ın ayırt edici yönlerinden biri, kod çözücü tarafında klasik segmentasyon mimarilerinde görülen karmaşık ve maliyetli tasarımlar yerine, "All-MLP" olarak adlandırılan oldukça hafif bir birleştirme yaklaşımı kullanmasıdır [19]. Bu yaklaşımın temel varsayımı, sınıfları ayıran asıl temsil gücünün büyük ölçüde kodlayıcı (MiT) içinde öğrenildiği; kod çözücünün ise yeni bir temsil "inşa etmekten" çok, farklı ölçeklerden gelen bilgiyi aynı düzleme getirip birleştirerek piksel düzeyinde sınıf tahmini üretmesi gerektiğidir [19].

All-MLP kod çözücüde süreç genel olarak şu mantıkla ilerler: Kodlayıcıdan gelen her çözünürlük seviyesindeki öznitelik haritası önce ortak bir kanal boyutuna indirgenir; ardından bu haritalar aynı uzamsal çözünürlüğe getirilerek birleştirmeye hazır hale gelir. Son adımda birleştirilen bu temsil üzerinden her piksel için sınıf skorları üretilir [19]. Burada dikkat edilmesi gereken nokta şudur: SegFormer, doğruluk artışını çoğunlukla kodlayıcı taraftaki çok ölçekli ve bağlamı güçlü temsilden bekler; kod çözücü ise bu temsili "taşıyan ve birleştiren" sade bir bileşen olarak tutulur [19].

Bu tercih, pratikte hesaplama verimliliği açısından anlamlıdır. Segmentasyon mimarilerinde ağır çözücü yapıları; yüksek çözünürlükte ara katmanların bellek tüketimini artırabilir ve yukarı örnekleme adımlarıyla işlem süresini uzatabilir. SegFormer'ın çözücüyü hafif tutması, özellikle yüksek çözünürlüklü uydu veya drone verisinde eğitim ve çıkarımı daha yönetilebilir hale getirmeyi hedefler [19]. Tez bağlamında bu verimlilik, yalnızca "hız" demek değildir; aynı zamanda deney döngüsünün hızlanması, daha fazla iyileştirme denemesinin yapılabilmesi ve her denemenin etkisinin daha net izlenebilmesi anlamına gelir.

İkinci bir kazanım da genelleme tarafında görülebilir. Veri miktarının sınırlı olduğu veya etiket belirsizliğinin yüksek olduğu senaryolarda, aşırı karmaşık çözücü tasarımlar modele gereksiz serbestlik kazandırarak aşırı uyum riskini büyütebilir. SegFormer'ın yaklaşımı ise temsil öğrenmenin ağırlığını kodlayıcıya verip çözücüde sade bir birleştirme hattı kullanarak bu riski azaltmayı amaçlar [19]. Bu özellik, hasar gibi sınıf

sınırlarının net olmadığı ve görüntü içinde benzer doku tuzaklarının çok bulunduğu problemlerde daha kontrollü bir davranışa zemin hazırlayabilir.

2.3.3. SegFormer'ın Hasar Tespiti İçin Avantajları

Deprem sonrası hasar segmentasyonu, görsel açıdan düzenli ve sınırları net nesnelerden çok; düzensiz, parçalı ve bağlama bağlı izler içerir. Enkaz yığınları, gölge bölgeleri, boşluklar, üst üste binmiş yapı dokuları ve farklı aydınlatma koşulları sınıflar arası ayrımı zorlaştırır. Ayrıca bu tezde hedeflenen tek zamanlı (yalnızca afet sonrası) senaryoda model, "önce-sonra farkı" gibi güçlü bir ipucuna sahip olmadan karar üretmek zorundadır. Bu koşullar altında SegFormer'ın avantajları üç başlıkta özetlenebilir:

Çok ölçekli temsil üretimi: MiT kodlayıcı, farklı çözünürlük seviyelerinde öznitelikler üreterek hem küçük detayların hem de daha geniş örüntülerin birlikte değerlendirilmesine imkan tanır [19]. Uydu görüntülerinde bina boyutlarının ve kentsel dokunun değişken olması, tek ölçekli bir temsilin yetersiz kalmasına yol açabilir. Bu nedenle çok ölçekli temsil, özellikle bina kenarı gibi ince ayrıntıları kaçırmadan, daha geniş sahne bağlamını da taşımak açısından önemlidir [19].

Geniş bağlamı kullanabilme potansiyeli: Transformer tabanlı yapılarda dikkat mekanizması, görüntünün farklı bölgeleri arasında ilişki kurmayı kolaylaştıran bir çerçeve sunar [31]. Bu, hasar tespitinde sık karşılaşılan benzer doku problemlerinde yardımcı olabilir. Örneğin gölge ile enkaz dokusu yerel olarak benzeşebilir; ancak çevredeki bina düzeni, sokak dokusu ve komşu yapı örüntüsü birlikte değerlendirildiğinde ayrım daha tutarlı hale gelebilir [31], [19].

Verimlilik ve deneysel iterasyon kolaylığı: SegFormer, ağır bir çözücü yerine hafif bir birleştirme tasarımı kullandığı için yüksek çözünürlüklü veride eğitim/çıkarım maliyetini daha yönetilebilir tutmayı hedefler [19]. Bu verimlilik, tez kapsamında kademeli iyileştirme yaklaşımına doğrudan uyumludur; çünkü veri artırma stratejileri, kayıp fonksiyonu ayarları, öğrenme oranı düzeni ve değerlendirme protokollerinin sistematik şekilde denenmesi daha pratik hale gelir. Böylece yalnızca en iyi sonucu aramak değil, hangi müdahalenin hangi etkiyi ürettiğini adım adım izlemek de kolaylaşır.

Sonuç olarak SegFormer, tek zamanlı afet sonrası görüntülerden hasar maskesi üretme hedefi için hem güncel literatürle uyumlu hem de uygulama açısından yönetilebilir bir mimari çerçeve sunar [19]. Çok ölçekli temsil üretimi, bağlamı daha geniş ölçekte

değerlendirebilme potansiyeli ve verimlilik odaklı çözücü tasarımı, deprem sonrası hasar segmentasyonu gibi zor bir görev için SegFormer'ı metodolojik olarak tutarlı bir tercih haline getirir [19], [31].

2.4. Literatürde Deprem Sonrası Hasar Tespiti Çalışmaları

Deprem sonrası hasar tespiti, uzaktan algılama literatüründe uzun süredir ele alınan bir problem olmasına rağmen, pratik koşullar nedeniyle hâlâ zorlu ve çok boyutlu bir araştırma alanı olarak görülmektedir [12]. Bu zorluğun temelinde yalnızca model mimarisi değil, veri edinim zincirinin kendisi yer alır: afet öncesi görüntünün bulunabilirliği, görüntülerin farklı sensörlerden gelmesi, çekim geometrisi ve ortorektifikasyon farkları, eş-kayıt (co-registration) hataları, optik görüntülerde aydınlatma–gölge etkileri, atmosferik koşullar ve mekânsal çözünürlüğe bağlı detay kayıpları hasar sinyalini doğrudan etkileyebilir [12]. Ayrıca bina ölçeğinde hasarın otomatik çıkarımı çoğu zaman ciddi sınıf dengesizliği barındırır; geniş bir sahnede "hasarlı piksel" oranı düşük olabilir ve modelin yanlış alarm üretme eğilimi artabilir. Bu durum, eğitim stratejisi ve değerlendirme protokolü kurulumunu da kritik hale getirir [12].

Buna karşın deprem sonrası hasar tespitine yönelik araştırma ilgisi giderek artmaktadır; çünkü uzaktan algılama tabanlı hızlı haritalama, sahaya erişimin sınırlı olduğu ilk saatler ve ilk günlerde durum farkındalığını artırma potansiyeli taşır [12]. Özellikle bina ölçeğinde "hasar nerede yoğunlaşıyor?" sorusuna yanıt verebilen yaklaşımlar, insani yardım ve kaynak planlamasını hızlandırma açısından değerli görülmektedir [12]. Literatürde bu hedefe iki ana senaryo üzerinden yaklaşıldığı görülür: yalnızca afet sonrası görüntüye dayalı tek zamanlı çıkarım ve afet öncesi–sonrası görüntü çiftine dayalı değişim tabanlı çıkarım. Bu iki senaryo, pratik veri kısıtlarına ve operasyonel ihtiyaçlara göre farklı avantajlar ve riskler taşır [12]. Literatürde deprem sonrası hasar tespiti çalışmalarının; kullanılan sensör türü (optik/SAR), veri senaryosu (tek-zamanlı/çift-zamanlı) ve model ailesi (klasik yöntemler, CNN tabanlı yaklaşımlar, Transformer tabanlı mimariler) açısından farklılaştığı görülmektedir. Bu çerçevede, alandaki temel çalışmalar Tablo 2.1'de veri/sensör, mimari-yöntem ve pratik güçlü yön–sınırlılık boyutlarıyla kronolojik olarak özetlenmiştir.

Tablo 2.1. Deprem sonrası hasar tespitinde temel çalışmaların kronolojik özeti.

Kaynak	Veri/Sensör	Mimari/Yöntem	Güçlü Yön	Sınırlılık
Townshend 1992 [37]	Teorik: misregistration (Hizalama Hatası)	CD analizi	Kayıt hatası etkisi net	Deprem özel değil
Yamazaki 2007 [38]	Çoklu çalışma (derleme)	Derleme (review)	Afet RS panoraması	Güncel DL az
Dell'Acqua 2012 [12]	Çoklu deprem (derleme)	Derleme	Deprem odaklı çerçeve	Yeni model yok
Dong & Shan 2013 [22]	Çoklu bina hasarı (derleme)	Derleme	Hatları toplar	Deney kıyası az
Gong 2016 [39]	TerraSAR-X VHR (sonrası)	SAR + ML/özellik	Bulut/ışık bağımsız	SAR yorumu zor
Ji 2018 [13]	Haiti 2010 optik VHR (sonrası)	CNN tabanlı	Afet öncesi veri gerektirmez	Gölge/benzer doku
Gupta 2019 [8]	xBD (öncesi/sonrası optik)	Benchmark veri seti	Standart kıyas	Domain shift, öncesi/sonrası
Weber 2020 [40]	Multi-temporal uydu	DL füzyon	Değişim sinyali güçlü	Hizalama + afet öncesi şart
Ge 2020 [40]	Çoklu SAR çalışma (derleme)	SAR derleme	Operasyonel SAR özeti	Benchmark sınırlı
Zheng 2021 [41]	Çoklu afet RS (öncesi/sonrası)	DL semantik CD	Hızlı yanıt çerçevesi	Eş-kayıt kritik
SegFormer 2021 [19]	Standart benchmarklar	Transformer (MiT+MLP)	Verimli + çok ölçek	Ayara duyarlı
Bai 2023 [42]	Çoklu veri seti (CD derleme)	CD derleme	Güncel toplama	Deprem özel değil
Macchiarulo 2024 [43]	VHR SAR (sonrası)	SAR + ML	Afet sonrası veri uygulanır	Etiket/önbilgi ihtiyacı

Kaynak	Veri/Sensör	Mimari/Yöntem	Güçlü Yön	Sınırlılık
Jing 2025 [44]	ChangeRD (pre/post)	Reg-aware CD	Misregistration toleransı	Karmaşıklık artar
Musaoglu 2025 [20]	KATE-PD (sonrası optik, TR)	Benchmark (afet sonrası)	Yerel doku, zor örnek	Dengesiz sınıf, gölge

Tablo 2.1'de görüldüğü üzere, erken dönem çalışmalar daha çok sensör/ön işleme ve değişim tespiti kısıtlarına odaklanırken, son yıllarda büyük ölçekli veri setleri ve derin öğrenme tabanlı mimarilerle birlikte model karşılaştırmaları daha sistematik biçimde ele alınmaktadır.

2.4.1. Tek Zamanlı (Post-Event) Hasar Tespiti Yaklaşımları

Tek zamanlı hasar tespiti, afet öncesi eşleşmiş görüntünün bulunmadığı, uygun eş-kayıt sağlanamadığı veya zaman baskısı nedeniyle önce–sonra çifti üretmenin mümkün olmadığı durumlarda devreye giren en erişilebilir senaryolardan biridir [12]. Bu senaryoda model, yalnızca afet sonrası görüntüdeki görsel ipuçlarından hasarı tahmin etmeye çalışır. Optik görüntülerde bu ipuçları çoğu zaman çatı düzenindeki bozulmalar, parçalı enkaz dokusu, bina formundaki deformasyon, çevresel yayılım ve yerleşim dokusundaki düzensizlikler gibi, doğrudan hasar sınıfına işaret eden örüntüler üzerinden temsil edilir [12]. Ancak tek zamanlı yaklaşım, değişim sinyalini doğrudan görmediği için karar daha fazla doku, şekil ve bağlam temsiline dayanır; bu da özellikle gölge–enkaz benzerliği gibi görsel tuzaklarda hataya açık bir ortam yaratır [12].

Bu bağlamda, tek zamanlı çalışmalarda SAR verisinin optiğe alternatif veya tamamlayıcı olarak kullanıldığı güçlü bir literatür hattı oluşmuştur. SAR'ın gün ışığına bağımlı olmaması ve bulutluluk gibi koşullardan daha az etkilenmesi, afet sonrası ilk edinimlerde veri sürekliliği açısından avantaj sağlayabilir [45]. Özellikle yalnızca post-event çok yüksek çözünürlüklü SAR görüntüsüyle bina ölçeğinde ayakta / çökmüş ayrımı hedefleyen yaklaşımlar; bina ayak izi (footprint) gibi yardımcı katmanlarla birlikte SAR özniteliklerini kullanarak tek zamanlı senaryoda anlamlı sonuçlar üretilebileceğini göstermiştir [39, 45]. Benzer biçimde daha yeni çalışmalar, yalnızca post-event SAR ve envanter-ayak izi bilgisiyle bölgesel ölçekte hasar tahmini yapmayı ve afet öncesi görüntü

bağımlılığını azaltmayı hedefleyen yöntemlerin giderek güçlendiğini vurgulamaktadır [43].

Optik görüntülerde tek zamanlı hasar tespiti ise çoğunlukla hasar dokusunu yakalamaya dayansa da, aydınlatma ve gölge etkileri, çekim zamanı farkları, enkazın zemin ve çevresel nesnelerle karışması gibi faktörler karar verme sürecini zorlaştırır [22]. Bu nedenle post-event only senaryosunda başarı, yalnızca mimariye değil; modelin bağlamı ne kadar iyi temsil edebildiğine, veri artırma ve ön işleme ile genellemenin ne kadar güçlendirildiğine ve sınıf dengesizliğine karşı eğitim stratejisinin ne kadar sağlam kurulduğuna bağlıdır [22]. Bu tezde tek zamanlı post-event senaryoya odaklanılması da, operasyonel açıdan öncesi yoksa da çalışabilen bir sistem hedefiyle uyumludur.

2.4.2. Çift Zamanlı (Pre-Post) Değişim Tabanlı Hasar Tespiti Yaklaşımları

Çift zamanlı hasar tespiti, afet öncesi ve sonrası görüntü arasındaki farktan yararlanarak değişim sinyalini doğrudan öğrenmeyi amaçlar. Bu yaklaşım teorik olarak daha güçlü bir ipucu sunar; çünkü model yalnızca hasarın sonraki görüntüde nasıl göründüğünü değil, aynı zamanda öncesine göre neyin değiştiğini de yakalayabilir [42, 46]. Bununla birlikte pratikte en kritik kırılganlık, iki tarihli görüntülerin geometrik olarak doğru hizalanmasıdır. Eş-kayıt hataları, piksel düzeyinde değişim maskesinde sahte sinyaller üretebilir; bu sahte sinyaller özellikle bina sınırlarında ve ince yapısal detaylarda belirginleşerek performansı düşürebilir [46]. Literatürde bu nedenle, eş-kayıt hatalarına dayanıklı değişim tespiti tasarımlarının ve kayıt hatasının değişim ölçümüne etkisini azaltan stratejilerin önemli bir araştırma başlığı olarak öne çıktığı görülür [42, 46].

Değişim tespiti daha genel anlamda uzaktan algılama literatüründe uzun süredir çalışılmaktadır. Klasik çizgide görüntü farkı/oranı ve istatistiksel karar kuralları gibi yöntemler yer alırken, modern yaklaşımlar derin öğrenme tabanlı temsil öğrenme ve mekânsal bağlam modelleme üzerine yoğunlaşmıştır [42]. Özellikle çok yüksek çözünürlüklü görüntülerde, tasarımın öznitelik uzayı seçimi, mekânsal bağlamın kullanımı ve karar kuralının birlikte düşünülmesi gerektiği vurgulanır; çünkü bina ölçeğinde küçük bir hizalama hatası bile değişim varmış gibi yorumlanabilir [46]. Bu bağlamda, bina hasarı için semantik değişim tespiti yaklaşımını nesne tabanlı bir çerçeveyle birleştiren ve hızlı afet yanıtı hedefleyen yöntemler de literatürde öne çıkmıştır [41].

Deprem hasarı özelinde pre-post yaklaşımı, “yıkıldı mı?” sorusu için güçlü bir sinyal sunabilse de her afet için kaliteli pre-event görüntüye erişim mümkün olmayabilir. Ayrıca pre-post görüntü çiftlerinde veri temini, kayıt doğruluğu ve çekim geometrisi kısıtları operasyonel süreçleri zorlaştırır; bu nedenle tek zamanlı yaklaşımlar önemini korumaktadır [12, 46]. Buna rağmen, çoklu sensör kullanımıyla optik ve SAR’ın tamamlayıcı doğasından yararlanan çalışmalar, deprem hasarı çıkarımında daha sağlam bir temsil kurmayı hedeflemiş ve bu hattın uzun süredir literatürde yer aldığı görülmüştür [38].

Son yıllarda büyük ölçekli açık veri setleri, pre/post hasar değerlendirmesini hızlandırmış ve yöntemlerin karşılaştırılabilir biçimde raporlanmasını kolaylaştırmıştır. Özellikle xBD, farklı afet türlerinden pre- ve post-event görüntüler ile bina anotasyonları ve hasar seviyeleri etiketleri sunarak hem değişim tespiti hem de hasar değerlendirme çalışmaları için önemli bir referans oluşturmuştur [8]. xBD’nin xView2 gibi benchmark ve yarışma ekosistemlerini beslemesi, yöntemlerin standart protokoller üzerinden kıyaslanmasına ve pratikte uygulanabilir boru hatlarının gelişmesine katkı sağlamıştır [8]. Bu tez post-event odaklı olsa da, bu benchmark çizgisi; sınıf dengesizliği, genelleme, veri çeşitliliği ve değerlendirme protokolü gibi başlıklarda metodolojik bir referans noktası sunması açısından önemlidir [8].

Bu bölümde, deprem sonrası bina hasarı ve yıkımının uzaktan algılama görüntülerinden çıkarımını mümkün kılan literatür çerçevesi geniş bir perspektifle özetlenmiştir. Öncelikle tek zamanlı post-event senaryoda hasarın doku, şekil ve bağlam ipuçları üzerinden çıkarıldığı; bu yaklaşımın operasyonel olarak erişilebilir olmakla birlikte gölge, aydınlatma ve sınıf dengesizliği gibi etkenlere duyarlı olduğu vurgulanmıştır [12]. Ardından çift zamanlı pre-post senaryonun değişim sinyali açısından daha güçlü bir zemin sunduğu, ancak eş-kayıt doğruluğu ve veri bulunabilirliği gereksinimleri nedeniyle pratikte önemli kısıtlar barındırdığı ele alınmıştır [37, 44]. Son olarak, xBD gibi büyük ölçekli veri setlerinin ve benchmark ekosistemlerinin yöntemlerin karşılaştırılabilir biçimde raporlanmasına katkı sunduğu belirtilmiştir [8, 40]. Bu çerçevenin, bir sonraki bölümde sunulacak veri seti, veri hazırlama adımları, deney tasarımı ve değerlendirme protokolünün neden bu şekilde kurgulandığını temellendirmek üzere bir referans haritası işlevi gördüğü değerlendirilmektedir [8, 12, 19].

3. MATERYAL VE METOD

Bu bölümde, tezin deneysel aşamasında kullanılan veri seti ve deney hattı genel çerçevesiyle sunulmaktadır. Öncelikle KATE-PD veri setinin yapısı ve çalışmanın hedefi olan tek zamanlı uydu görüntülerinden bina hasarının piksel düzeyinde çıkarımı tanımlanır. Ardından görüntü ve maskelerin modele uygun hale getirilmesi için uygulanan temel ön işleme adımları ile ikili maske üretim mantığı özetlenir. Veri artırma kısmında, genellemeyi güçlendirmek amacıyla kullanılan başlıca geometrik dönüşümler ve maske–görüntü hizasının korunması prensibi belirtilir. Son olarak, SegFormer tabanlı yaklaşımın kademeli iyileştirme mantığıyla eğitilip değerlendirildiği, karşılaştırma amacıyla alternatif bir mimarinin de aynı protokol altında test edildiği ve değerlendirme-çıkarım kararlarının temel hatlarıyla ele alındığı ifade edilir.

3.1. Veri Kümesi ve Problem Tanımı

Bu çalışmanın deneysel aşamasında kullanılan temel veri kaynağı, 6 Şubat 2023 tarihinde meydana gelen ve "Yüzyılın Felaketi" olarak adlandırılan Kahramanmaraş merkezli depremlerin ardından hazırlanan KATE-PD (KAhramanmaras Turkiye Earthquake Post Disaster) veri setidir [20, 21].

KATE-PD veri seti, 6 Şubat 2023 depremlerinden sonra tekil afet sonrası uydu görüntülerinden yıkık/enkazlı bina alanlarını piksel düzeyinde çıkarmayı hedefleyen bir semantik segmentasyon benchmark'ıdır. Literatürdeki mevcut hasar tespit veri setlerinin büyük bir kısmı genellikle afet öncesi ve sonrası görüntü çiftlerine ihtiyaç duyan değişim tespiti (change detection) odaklıyken, KATE-PD veri seti özellikle afet sonrası çekilen tekil görüntülerden hızlı ve hassas analiz yapabilmek amacıyla tasarlanmıştır [20, 21]. Veri seti, Adıyaman, Gaziantep, Hatay, Kahramanmaraş, Kilis, Osmaniye ve Malatya olmak üzere yedi ildeki ağır etkilenen bölgeleri kapsar ve toplamda 3.293 km²'lik geniş bir coğrafi alan üzerinde 1.952 adet yıkık bina anotasyonunu içerir [20, 21].

KATE-PD'nin ham görüntü kaynağı 14 adet afet sonrası yüksek çözünürlüklü optik uydu görüntüsünden oluşur; farklı sensör sağlayıcıları (Airbus Pleiades 1A/1B ve Maxar WorldView serisi gibi), farklı aydınlatma ve bakış açıları ile 0.3–0.5 m aralığında yer örnekleme mesafesi (GSD) çeşitliliği sunar [20, 21]. Görüntüler RGB bantlarda, 8-bit radyometrik çözünürlükte olup EPSG:4326 (WGS84) koordinat sisteminde

standartlaştırılmış ve GeoTIFF formatında sağlanmıştır; kaynak görüntülerin il/ilçe, tarih, sağlayıcı ve çözünürlük bilgileri tablo halinde raporlanmıştır [20, 21].

Etiketleme sürecinde çalışma alanı çakışmasız 512×512 piksel karolara bölünür; bu karolar Label Studio ortamına yüklenerek işbirlikçi bir anotasyon kampanyası yürütülür. Anotasyonlar yalnızca görsel yorumla yapılır; etiketleyicinin emin olmadığı düşük kaliteli örnekler veri setinden elenir ve "emin olunan" hasar bölgeleri poligonlarla işaretlenir [20, 21]. Yıkılmış bir bina çatısı kısmen sağlam görünse bile çevresindeki molozla birlikte bütün yapı alanı etiketlenir; poligonlar moloz yayılımını mümkün olduğunca kapsayacak şekilde çizilir, ayırt edilemeyen bitişik yıkımlarda tek bir poligonla birleşik alan işaretlenebilir ve enkazla iç içe geçmiş bitki örtüsü gibi unsurlar da ilgili poligona dahil edilebilir [20, 21].



Şekil 3.1. KATE-PD veri setinden alınmış örnek RGB ve ikili maske görüntüleri.

Çalışmada kullanılan KATE-PD veri setinden seçilen örnek RGB-maske çiftleri Şekil 3.1'de sunulmaktadır. Şekildeki RGB yamalar, deprem sonrası kentsel dokuda enkaz kümeleri, çatılardaki düzensiz dokular, gölge alanları ve yüksek parlaklık gibi değişken görsel koşulları bir arada barındırarak, tek zamanlı hasar segmentasyonunun neden zorlu bir problem olduğunu somut biçimde göstermektedir. Karşılık gelen ikili maske görüntülerinde ise beyaz pikseller hasar sınıfını, siyah pikseller arka plan sınıfını temsil etmektedir; böylece model, yalnızca post-event görüntüden, hasarın düzensiz ve parçalı geometrisini piksel düzeyinde öğrenmeye zorlanmaktadır. Bu örneklerin birlikte verilmesi, veri setindeki sahne çeşitliliğini ve etiketlerin poligon tabanlı yapısını görsel olarak açıklarken, ilerleyen bölümlerde tanımlanan ön işleme, veri artırma ve değerlendirme adımlarının neden gerekli olduğunu da temellendirmektedir.

Bu çalışmada, deprem sonrası uzaktan algılama görüntülerinden tek görüntü senaryosunda hasar bölgelerinin ayrıştırılması problemi ele alınmıştır. Amaç, her bir girdi görüntüsü için piksel düzeyinde ikili bir maske üretmek ve "hasar" sınıfına ait alanları arka plandan ayırmaktır. Çalışmada kullanılan veri kümesi KATE-PD olup, hasar bölgeleri etiketlenmiş maske görüntüleriyle birlikte sunulmaktadır. Problem, iki sınıflı ayrıştırma olarak tanımlanmıştır: arka plan (0) ve hasar (1).

Bu bölümde tanımlanan tüm deneylerde veri kökü sabit tutulmuş, model çıktıları iki sınıf üzerinden üretilmiştir. Böylece farklı mimari ve eğitim stratejileri arasında karşılaştırılabilir bir deney zemini oluşturulmuştur.

Model geliştirme ve deneyler, Python tabanlı derin öğrenme ekosistemi ile yürütülmüş; temel eğitim ve değerlendirme süreçleri için PyTorch kullanılmıştır. PyTorch'un dinamik hesap grafiğine dayalı yapısı, derin öğrenme araştırmalarında hızlı iterasyon ve hata ayıklama açısından yaygın biçimde tercih edilmektedir [47].

3.2. Veri Hazırlığı ve Protokol

Bu çalışmada deprem sonrası hasar tespiti problemi, tek zamanlı uzaktan algılama görüntülerinde ikili semantik segmentasyon problemi olarak ele alınmıştır. Kullanılan KATE-PD veri seti, 6 Şubat 2023 Kahramanmaraş depremleri sonrasında farklı uydu kaynaklarından elde edilen görüntüler ile bunlara karşılık gelen hasar etiketlerini içermektedir. Veri seti, afet sonrası hızlı durum değerlendirme senaryolarına uygun

olacak şekilde “hasarlı / hasarsız” ikili ayrımını hedefleyen piksel tabanlı maskelerle sunulmaktadır [20, 21].

Veri seti eğitim, doğrulama ve test olmak üzere üç alt kümeye ayrılmıştır. Eğitim kümesi model parametrelerinin öğrenilmesi için kullanılırken, doğrulama kümesi model seçimi ve hiperparametre ayarları için kullanılmıştır. Test kümesi ise model ve ayarlar sabitlendikten sonra yalnızca nihai performansın raporlanması amacıyla kullanılmıştır. Bu ayrım, model değerlendirmesinin tarafsız ve bilimsel olarak savunulabilir şekilde yapılmasını sağlamaktadır [48]. Veri setinin bölümleme bilgileri Tablo 3.1’de sunulmaktadır.

Tablo 3.1. KATE-PD veri setinin bölümleme sayı ve oranları.

Bölüm	Görüntü Sayısı	Oran (Yaklaşık%)
Train	675	80
Validation	78	10
Test	79	10

Veri setinde sağlanan etiketler piksel tabanlı maske formatında kullanılmıştır. Maske görüntüleri tek kanallı (grayscale) biçimde okunmuş ve ikili sınıf yapısına dönüştürülmüştür. Bu dönüşümde maske üzerinde pozitif sınıfı temsil eden değerler (örneğin 255) 1’e dönüştürülmüş, diğer tüm değerler ise 0 olarak kabul edilmiştir. Böylece modelin tahmin edeceği hedef çıktı her piksel için [0,1] kümesinden bir sınıf etiketine indirgenmiştir. Bu yaklaşım, hasar alanını tek bir sınıf altında topladığı için problem tanımını sadeleştirmekte ve ikili segmentasyon metriklerinin (IoU, F1) daha doğrudan yorumlanmasına olanak sağlamaktadır.

Model girişlerinin standartlaştırılması amacıyla Hugging Face Transformers ekosistemi içerisinde bulunan AutoImageProcessor bileşeni kullanılmıştır. Bu bileşen görüntüler üzerinde yeniden boyutlandırma (resize), piksel değerlerinin yeniden ölçeklenmesi (rescale) ve normalizasyon işlemlerini otomatik olarak gerçekleştirmektedir. Normalizasyon işlemi çoğunlukla ImageNet veri setine ait ortalama ve standart sapma değerleri kullanılarak yapılmaktadır. Bu standartlaştırma adımları farklı görüntü kaynaklarından gelen veriler arasındaki dağılım farklılıklarını azaltarak model eğitimini daha kararlı hale getirmektedir [49].

Ayrıca ağ giriş boyutunun sabitlenmesi ve mini-batch eğitim sürecinin mümkün kılınabilmesi için görüntüler model girişine verilmeden önce yeniden boyutlandırılmıştır. KATE-PD veri setinin parça (patch) temelli yapısı (örneğin 512×512 yamalar) bu standardizasyonu kolaylaştırmakta ve eğitim sırasında hesaplama ile bellek maliyetinin yönetilebilir kalmasına yardımcı olmaktadır [20].

Deney tasarımında veri artırma bileşeni yalnızca "augmentation" içeren deney hatlarında etkinleştirilmiş; baseline ve Lovász kaybı odaklı deneylerde veri artırma kapalı tutulmuştur. Böylece veri artırmanın performans üzerindeki etkisi diğer bileşenlerden bağımsız olarak gözlemlenebilmiştir. Segmentasyon görevlerinde veri artırma uygulanırken temel kural, görüntüye uygulanan geometrik dönüşümlerin maskeye de aynı şekilde uygulanmasıdır. Aksi durumda piksel–etiket hizası bozulur ve öğrenme süreci doğrudan olumsuz etkilenir.

Bu çalışmada kullanılan veri artırma stratejisi, maske ile geometrik tutarlılığı koruyan basit geometrik dönüşümlerden oluşmaktadır. Eğitim verisi üzerinde yatay çevirme (horizontal flip) dönüşümü p_hflip = 0.5 olasılıkla uygulanmıştır. Dikey çevirme dönüşümü ise deney tasarımında varsayılan olarak kapalı tutulmuş (p_vflip = 0.0) ve yalnızca gerekli görülen durumlarda etkinleştirilebilecek şekilde tanımlanmıştır. Buna ek olarak, 90° tabanlı döndürme dönüşümü için p_rot90 = 0.5 kullanılmış ve dönüşüm aktif olduğunda açı değeri (0°, 90°, 180°, 270°) kümesinden rastgele seçilmiştir. Bu dönüşümlerde görüntüler için bilinear, maske görüntüleri için ise nearest-neighbor enterpolasyonu kullanılarak sınıf etiketlerinin ara değerlere dönüşmesi engellenmiştir. Bu veri artırma stratejisi, modelin farklı yönelimlere sahip hasar örüntülerine karşı daha dayanıklı bir temsil öğrenmesine katkı sağlamayı amaçlamaktadır.

Model eğitimi sırasında kullanılan veri yükleme ve artırma işlemleri rastgele seçimlere dayalı olduğundan, deneylerin tekrarlanabilirliğini sağlamak amacıyla rastgelelik üreten başlangıç parametresi (seed) sabitlenmiştir. Bu uygulama, konfigürasyon değişmediği sürece deneylerin aynı sonuçları üretmesini sağlayarak deneysel güvenilirliği artırmaktadır.

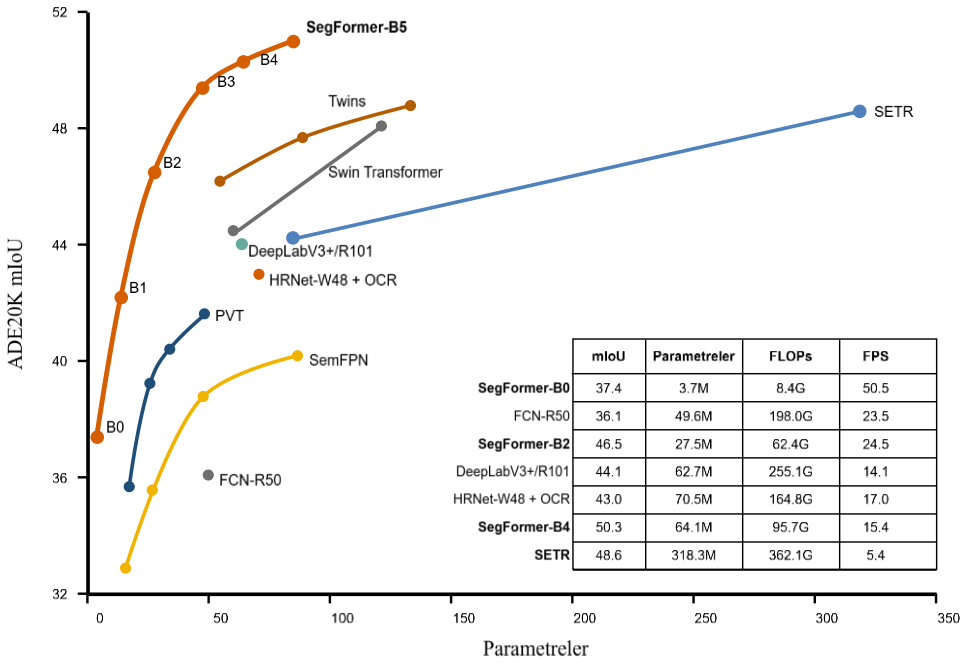
3.3.Model Mimarileri ve Deney Versiyonları

Bu tezde deprem sonrası hasar alanlarının ikili semantik segmentasyon problemine (arka plan / hasar) indirgenmiş bir formülasyonu ele alınmış ve transformer tabanlı iki farklı

mimari ailesi üzerinden karşılaştırmalı bir deney tasarımı yürütülmüştür. Temel amaç; aynı veri bölünmesi ve aynı raporlama protokolü altında, SegFormer tabanlı bir hattın kademeli olarak iyileştirilmesi ve alternatif bir transformer segmentasyon mimarisi olan Swin + UPerNet ile kıyaslanmasıdır. KATE-PD veri seti bağlamında hasar dokularının düzensiz yapısı, sınır belirsizliği ve arka planla benzer görsel örüntülere sahip olması, modelin hem bağlamı hem de ayrıntıyı birlikte öğrenmesini gerektirmektedir [20].

3.3.1. SegFormer: MiT omurga ve all-MLP decoder

SegFormer, semantik segmentasyon için transformer tabanlı bir omurga (MiT: Mix Transformer) ve hafif bir decoder tasarımıyla, çok ölçekli özellikleri verimli biçimde birleştirmeyi hedefleyen bir mimaridir. Kuramsal temeller ve literatür özeti bölümünde ve şekil 2.4 üzerinde detaylı şekilde tanıtıldığı gibi MiT omurgası, farklı çözünürlüklerde hiyerarşik temsil öğrenerek hem geniş bağlam bilgisini hem de yerel ayrıntıları yakalayabilen bir yapı sunar; decoder tarafında ise ağırlıklı olarak MLP tabanlı bir birleştirme yaklaşımı ile segmentasyon çıktısı üretilir. Bu tasarım, klasik ağır decoder'lara kıyasla daha düşük hesaplama maliyetiyle rekabetçi segmentasyon performansı elde etmeyi amaçlar [19].



	mIoU	Parametreler	FLOPs	FPS
SegFormer-B0	37.4	3.7M	8.4G	50.5
FCN-R50	36.1	49.6M	198.0G	23.5
SegFormer-B2	46.5	27.5M	62.4G	24.5
DeepLabV3+/R101	44.1	62.7M	255.1G	14.1
HRNet-W48 + OCR	43.0	70.5M	164.8G	17.0
SegFormer-B4	50.3	64.1M	95.7G	15.4
SETR	48.6	318.3M	362.1G	5.4

Şekil 3.2. ADE20K'da performans ve model verimliliği karşılaştırması [19].

Şekil 3.2'de [19] SegFormer mimarisinin versiyonları ve diğer mimarilerle performans karşılaştırması görülmektedir. Bu tezde SegFormer deneyleri için "mit-b2" omurgası seçilmiştir. B2 omurgasının tercih edilme nedeni, hem güçlü bir yapı hem de küçük-orta ölçekli GPU bellek kısıtları altında uygulanabilir bir yol sunmasıdır.

3.3.2. Swin Transformer ve UPerNet

Bir önceki bölümde şekil 2.1 ve şekil 2.2'de detayları açıklandığı üzere Swin Transformer, görüntüyü aşamalı bir yapı içinde işleyerek her aşamada pencere tabanlı dikkat bloklarıyla öznitelik çıkarır ve "patch merging" adımlarıyla çözünürlüğü kademeli olarak düşürüp kanal sayısını artırarak hiyerarşik bir temsil üretir. Ayrıca aynı şekilde gösterilen ardışık blok yapısında, bir katmanda pencere içi dikkat (W-MSA) uygulanırken takip eden katmanda kaydırılmış pencere dikkatine (SW-MSA) geçilerek pencereler arası bilgi alışverişi sağlanır ve küresel bağlamın daha verimli biçimde taşınması hedeflenir [33].

UPerNet ise çok ölçekli özelliklerin birleştirilmesi için "feature pyramid" ve bağlam modüllerini birlikte kullanarak farklı seviyelerdeki temsilleri tek bir segmentasyon çıktısında toplar. Bu nedenle Swin + UPerNet birleşimi, SegFormer'e göre farklı bir tasarım felsefesini temsil eden güçlü bir karşılaştırma referansı olarak seçilmiştir.

Bu tezde Swin + UPerNet hattı, bir kıyas mimarisi olarak konumlandırılmış; amaç SegFormer hattında yapılan iyileştirmelerin, alternatif bir transformer segmentasyon sistemine göre doğruluk–maliyet dengesi açısından nerede durduğunu göstermektir.

3.3.3. Kayıp Fonksiyonları ve IoU Odaklı Öğrenme Motivasyonu

Segmentasyon problemlerinde piksel düzeyinde sınıflandırma kaybı, sınıfların doğru sınıflandırılmasına odaklanırken; IoU gibi bölge örtüşme tabanlı metrikler doğrudan optimize edilmez. Bu tezde IoU ile daha uyumlu bir optimizasyon hedefi için Lovász-Softmax kaybı incelenmiştir. Lovász-Softmax, IoU/Jaccard benzeri küme örtüşme ölçülerine daha yakın bir surrogate (yaklaştırım) kayıp sağlayarak, özellikle sınıf sınırlarının belirsiz olduğu durumlarda segmentasyon kalitesini artırmayı hedefler. [17].

Bu bağlamda tezde iki ana eğitim stratejisi yer almıştır:

1. Standart eğitim (baseline): çapraz-entropi temelli öğrenme,
2. IoU hizalı eğitim: çapraz-entropi + Lovász-Softmax birleşimi (SegFormer + Lovász hattı).

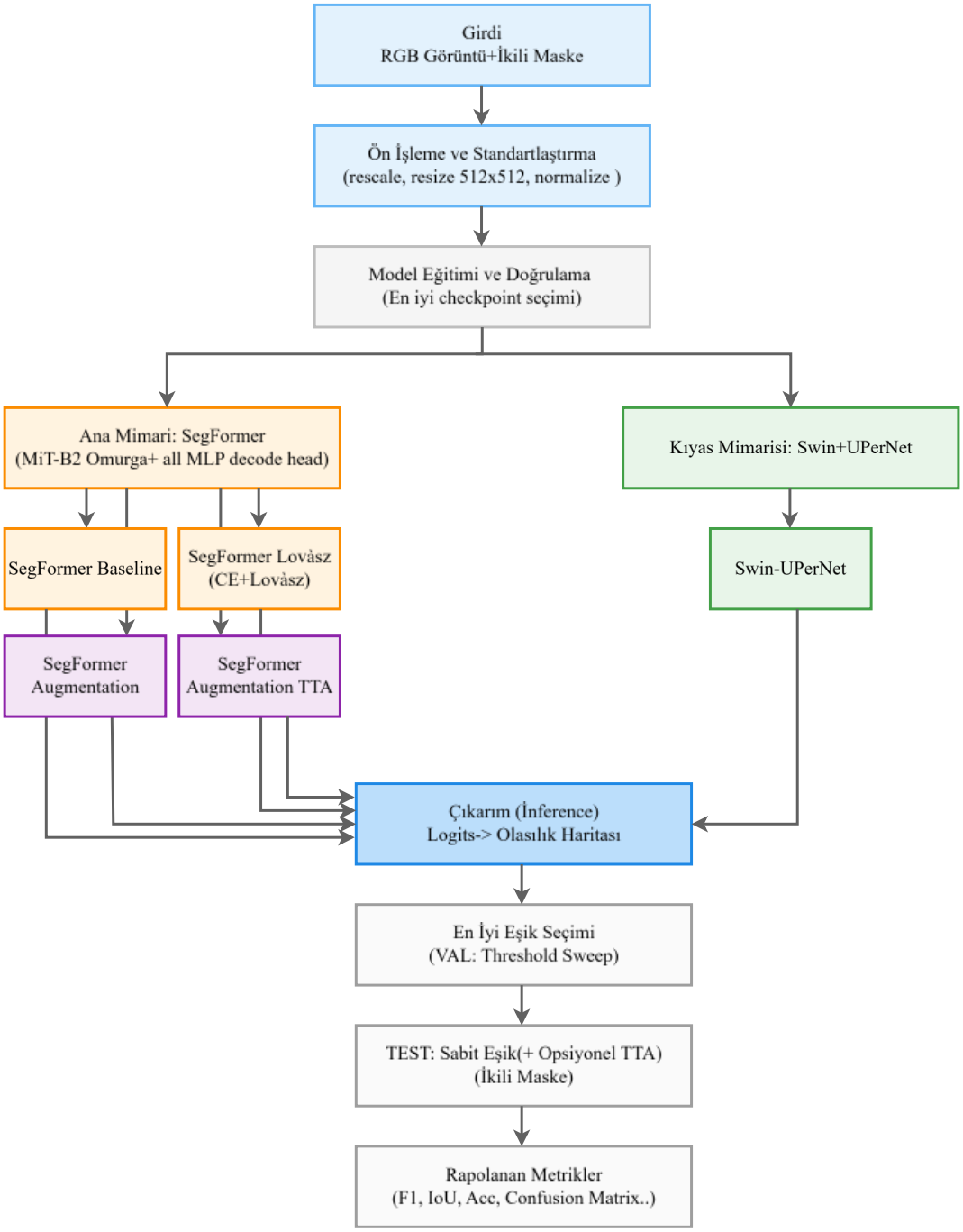
3.3.4. Deney Versiyonları

Bu tezde deneyler, her bir bileşenin etkisini ayrı ayrı görebilmek amacıyla kontrollü ve bağımsız karşılaştırmalar şeklinde kurgulanmıştır. Bu tasarımın temel gerekçesi, hasar segmentasyonu gibi sınır belirsizliği ve arka plan benzerliği nedeniyle farklı hata türlerine açık bir problemde, performans değişiminin hangi müdahaleden kaynaklandığını net biçimde ayırt edebilmektir.

Bu doğrultuda her deney versiyonu, aynı veri bölünmesi ve aynı raporlama protokolü korunarak, ortak bir referans hattına göre yalnızca belirli bir ana bileşeni test edecek şekilde yapılandırılmıştır. Böylece mimari tercihin etkisi, veri artırma stratejisinin etkisi, kayıp fonksiyonu/optimizasyon ayarlarının etkisi ve çıkarım aşamasındaki karar stratejilerinin etkisi birbirinden bağımsız olarak ölçümlenebilir hale getirilmiştir.

Bu yaklaşım, tez raporlamasında katkı analizini de kolaylaştırır. Her bir müdahalenin, referans hatta göre nicel metriklerde ne ölçüde fark yarattığı ayrı ayrı raporlanabildiğinden, elde edilen iyileşmelerin hangi teknik tercihlerle ilişkili olduğu daha güçlü ve izlenebilir biçimde tartışılabilmektedir.

Bu doğrultuda, veri hazırlama adımlarından model eğitimi ve çıkarım stratejilerine kadar izlenen tüm deney süreci, SegFormer ana hattı ile Swin+UPerNet kıyas hattını aynı değerlendirme protokolü altında birleştiren genel akış şeması, Şekil 3.3'te özetlenmiştir.



Şekil 3.3. Deney hattının genel akışı ve karşılaştırmalı modelleme çerçevesi.

Şekil 3.3'deki deney hatlarında belirtilen adımlar sırasıyla uygulanmıştır.

(1) Baseline SegFormer (MiT-B2): Bu hat, çalışmanın başlangıç (referans) noktasıdır. SegFormer mimarisi, veri setinin ikili hasar maske yapısı altında standart eğitim protokolü ile eğitilmiştir. Burada amaç, mimariyi "olduğu gibi" kullanarak başlangıç performansını belirlemek ve sonraki tüm iyileştirmeleri bu referans çizgisine göre

değerlendirmektir. Baseline, aynı veri bölünmesi ve aynı değerlendirme protokolü altında raporlandığı için, eklenen her bileşenin getirdiği artış/azalış doğrudan kıyaslanabilir.

(2) SegFormer + Lovász: İkinci hat, baseline ile aynı mimariyi korur; ancak optimizasyon hedefini güçlendirmek için kayıp fonksiyonuna Lovász-Softmax bileşeni eklenir. Bu hattın motivasyonu, segmentasyonda nihai performans göstergelerinden biri olan IoU'nun, salt piksel düzeyli sınıflandırma kayıplarıyla (ör. CE) her zaman doğrudan hizalı olmamasıdır. Dolayısıyla Lovász bileşeni ile modelin örtüşme kalitesini artıracak yönde daha tutarlı güncellemeler alması hedeflenir. Bu deneyin çıktısı, "aynı mimari + farklı kayıp" değişkeninin izole edilmesini sağlar.

(3) SegFormer + Augmentation: Üçüncü hat, model mimarisini ve eğitim hedefini mümkün olduğunca sabit tutarak, veri tarafında kontrollü bir çeşitlilik artışı sağlar. Afet sahnelerinde hasar dokusu, gölge, çatı malzemesi, ışık koşulları ve görüntüleme açısı gibi faktörler aynı sınıf için farklı görsel örüntüler doğurabilir. Bu nedenle eğitim sırasında uygulanan geometrik dönüşümler (ör. 90° döndürme, yatay çevirme) modelin belirli poz/oryantasyon kalıplarına aşırı uyum göstermesini azaltmayı ve daha iyi genelleme sağlamayı amaçlar. Bu hattın değeri, "aynı mimari + veri artırma" etkisini tek başına görünür kılmasıdır.

(4) UPerNet + Swin : Dördüncü hat, SegFormer'i iyileştirme çizgisinden bağımsız olarak, problem için farklı bir transformer tasarımının kıyas amaçlı değerlendirilmesini içerir. Swin omurgası pencere-tabanlı dikkat mekanizmasıyla hesaplama maliyetini kontrol ederken, UPerNet çok ölçekli özellikleri güçlü bir şekilde birleştiren bir decoder yaklaşımı sunar. Bu hattın tez içindeki rolü, "aynı veri/protokol altında alternatif mimari ailesi ne kadar rekabetçi?" sorusuna cevap vererek, SegFormer tabanlı önerilen hattın bağlamsal değerini güçlendirmektir.

(5) SegFormer + Augmentation + TTA: Son hat, eğitim aşamasında kullanılan SegFormer + Augmentation modelini temel alır ve bu modeli değiştirmeden çıkarım (inference) aşamasında daha kararlı maske üretmeyi hedefler. Burada iki temel mekanizma birlikte ele alınır:

- Eşik (threshold) seçimi: Modelin ürettiği olasılık haritasından ikili maske üretmek için kullanılan eşik değeri, doğrulama (val) seti üzerinde taranır ve seçilir. Bu işlem, model ağırlıklarını değiştirmeden yalnızca karar mekanizmasını ayarladığı için "post-processing / karar katmanı" iyileştirmesi olarak değerlendirilir.

- Test-Time Augmentation (TTA): Görüntünün yatay/dikey çevrilmiş varyantları üzerinden tahmin alınıp olasılık haritaları birleştirilerek, rastlantısal hataların azaltılması ve sınır bölgelerinde daha stabil çıktılar elde edilmesi amaçlanır.

Bu hat, özellikle eğitim tamamlandıktan sonra düşük maliyetli bir performans kazanımı sunabildiği için pratik değere sahiptir; çünkü yeniden eğitim gerektirmeden yalnızca değerlendirme/çıkarım adımlarında iyileştirme sağlar. Bu deney hattı, modelin yalnızca eğitimle değil, karar ve çıkarım stratejileriyle de güçlendirilebildiğini göstermesi açısından önemlidir.

3.3.5. Olasılık haritası, Eşikleme ve (Opsiyonel) Test-Time Augmentation

Transformer segmentasyon modelleri piksel başına logit çıktısı üretir ve bu değerler olasılığa dönüştürülerek ikili hasar maskesi elde edilir. İkili senaryoda pratik yaklaşım; hasar sınıfına ait olasılık haritasından bir karar eşiği ile maske üretmektir. Bu tezde raporlama protokolünde karar eşiği, doğrulama seti üzerinde taranarak seçilmiş; test setinde ise seçilen eşik sabit tutularak tek seferlik nihai performans raporlanmıştır. Bu yaklaşım, test setine ayar yapılmasını engelleyerek değerlendirmede tarafsızlığı korur.

Ayrıca deneylerden birinde test-time augmentation (TTA) uygulanmıştır. TTA'da aynı görüntünün yatay/dikey çevrilmiş varyantları üzerinden tahmin alınır, olasılık haritaları birleştirilir ve daha kararlı bir çıktı hedeflenir [50]. Bu tezde TTA, eğitim değil çıkarım (inference) aşaması iyileştirmesi olarak ele alınmış; dolayısıyla model ağırlıkları değişmeden yalnızca karar katmanı güçlendirilmiştir.

3.4. Performans Metrikleri

Bu tezde modellerin hasar tespiti başarımı, piksel düzeyinde sınıflandırma çıktıları üzerinden nicel olarak değerlendirilmiştir. Değerlendirme, ikili semantik segmentasyon problemine uygun biçimde piksel bazında karışıklık matrisi (confusion matrix) terimlerine indirgenmiş; buradan türetilen metrikler ile hem sınıf-odaklı hem de genel performans raporlanmıştır. Kullanılan temel metrikler: Precision, Recall, F1, IoU, mIoU ve Accuracy'dir. Ayrıca raporlamada, test seti yalnızca nihai değerlendirme için kullanılmış; eşik seçimi doğrulama seti üzerinden yapılmıştır.

3.4.1. Karışıklık Matrisi

İkili segmentasyonda her piksel için model çıktısı "hasar" veya "arka plan" olarak karar verilir. Bu karar, olasılık haritasının bir eşik değerine göre iki sınıfa ayrıştırmasıyla elde edilir. Karar sonucunda piksel bazında dört temel sayı tanımlanır:

- TP (True Positive): Gerçekte hasar olan pikselin hasar tahmin edilmesi
- FP (False Positive): Gerçekte arka plan olan pikselin hasar tahmin edilmesi
- TN (True Negative): Gerçekte arka plan olan pikselin arka plan tahmin edilmesi
- FN (False Negative): Gerçekte hasar olan pikselin arka plan tahmin edilmesi

Tezde kullanılan raporlama yaklaşımı global confusion mantığıyla, bu terimleri tüm görüntülerdeki tüm pikseller üzerinde biriktirerek (dataset-level) hesaplar. Bu tercih, metriklerin tek tek görüntülerdeki varyanstan etkilenmesini azaltır ve benchmark tarzı dataset üzerinden toplu raporlama yaklaşımıyla uyumludur; literatürde "dataset–mIoU" vurgusu da bu noktaya işaret eder [17]. Bununla birlikte, bu tezde yalnızca piksel düzeyinde değerlendirme yapılmamış, aynı zamanda yapı bazlı (object-level) değerlendirme de gerçekleştirilmiştir. Yapı bazlı analizde, hasarlı yapıların bir bütün olarak doğru şekilde tespit edilip edilmediği dikkate alınmakta; bu kapsamda confusion matrix terimleri yapı seviyesinde doğru ve yanlış tespitler üzerinden yeniden tanımlanmaktadır. Bu yaklaşım, modelin yalnızca piksel doğruluğunu değil, aynı zamanda hasarlı yapıların bütüncül olarak yakalanabilme başarısını değerlendirmeyi mümkün kılmaktadır.

3.4.2. Hassasiyet (Precision), Duyarlılık (Recall) ve F1 Skoru

Piksel tabanlı segmentasyon performansının değerlendirilmesinde, sadece genel doğruluk değerine bakmak, özellikle hasarlı alanların tüm görüntüye oranla çok küçük olduğu durumlarda yanıltıcı sonuçlar verebilir. Bu nedenle, modelin hasar sınıfı (pozitif sınıf) üzerindeki etkinliğini ölçmek amacıyla hassasiyet (precision) ve duyarlılık (recall) metrikleri kullanılır:

Hassasiyet (precision), modelin hasar olarak tahmin ettiği piksellerin ne kadarının gerçekten hasarlı olduğunu gösterir; dolayısıyla modelin yanlış alarm (false positive) üretme eğilimi bu metrik üzerinden izlenir. Hassasiyet, Eşitlik 3.1 ile hesaplanmaktadır:

$$\text{Precision} = \frac{TP}{TP + FP} \qquad (3.1)$$

Eşitlik 3.1'de yer alan TP doğru pozitifleri (hasar olup hasar tahmin edilen pikseller), FP ise yanlış pozitifleri (hasar değilken hasar tahmin edilen pikseller) temsil etmektedir. Duyarlılık (recall), gerçekte hasarlı olan piksellerin ne kadarının model tarafından yakalandığını ifade eder; bu nedenle hasarı kaçırma (false negative) davranışıyla doğrudan ilişkilidir. Duyarlılık, Eşitlik 3.2 ile tanımlanır:

$$\text{Recall} = \frac{TP}{TP + FN} \qquad (3.2)$$

Eşitlik 3.2'de FN, yanlış negatifleri (hasar olup arka plan tahmin edilen pikseller) göstermektedir.

Hassasiyet ve duyarlılık arasında denge kurmak amacıyla, bu iki ölçütün harmonik ortalaması olan F1 skoru kullanılır. F1 skoru Eşitlik 3.3 ile hesaplanmaktadır:

$$F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (3.3)$$

Harmonik ortalama yapısı sayesinde hassasiyet veya duyarlılıktan biri düşük kaldığında F1 skoru da belirgin biçimde düşmekte; bu durum modelin zayıf yönünü daha görünür kılmaktadır. Bu nedenle F1, özellikle sınıf dengesizliğinin belirgin olduğu segmentasyon problemlerinde performansı tek bir özet değerle raporlamak için pratik bir ölçüt sunar [51].

3.4.3. Birleşme Üzerinden Kesişim (IoU) ve Ortalama IoU (mIoU)

Semantik segmentasyonda yaygın kullanılan ölçütlerden biri, Jaccard indeksine dayanan Birleşme Üzerinden Kesişim (Intersection over Union, IoU) metriğidir. IoU, tahmin edilen sınıf bölgesi ile gerçek sınıf bölgesinin kesişiminin, birleşime oranını ölçer. Sınıf c için IoU, Eşitlik 3.4 ile tanımlanır:

$$IoU_c = \frac{TP_c}{TP_c + FP_c + FN_c} \qquad (3.4)$$

Eşitlik 3.4'te TP_c, FP_c ve FN_c terimleri sırasıyla sınıf c için doğru pozitif, yanlış pozitif ve yanlış negatif piksel sayılarını ifade etmektedir.

Bu tezde problem ikili olarak ele alındığından iki sınıf bulunmaktadır: arka plan ($c = 0$) ve hasar ($c = 1$). Ortalama IoU (mIoU) ise tüm sınıflar için hesaplanan IoU değerlerinin ortalamasıdır ve Eşitlik 3.5 ile verilir:

$$mIoU = \frac{1}{|\,C\,|} \sum_{c \in C} Io\,U_c \qquad (3.5)$$

Burada C, sınıf kümesini; $|C|$ise sınıf sayısını temsil etmektedir. Literatürde "image–mIoU" ve "dataset–mIoU" ayrımı vurgulanmaktadır. Dataset–mIoU yaklaşımında tüm pikseller bir araya getirilerek tek bir küresel hesap yapılır; bazı sınıfların bazı görüntülerde hiç bulunmadığı durumlarda bu yaklaşım daha kararlı raporlama sağlayabilmektedir [17].

3.4.4. Doğruluk (Accuracy)

Doğruluk (accuracy), tüm pikseller üzerinde doğru sınıflanan piksellerin oranını ifade eder ve Eşitlik 3.6 ile tanımlanır:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \qquad (3.6)$$

Eşitlik 3.6'da TN, doğru negatifleri yani, arka plan olup arka plan tahmin edilen pikselleri temsil etmektedir. Ancak segmentasyon problemlerinde arka plan sınıfı çoğu zaman baskın olduğundan accuracy değeri yüksek görünebilir ve bu metrik tek başına yanıltıcı olabilir. Bu tezde accuracy yalnızca tamamlayıcı bir metrik olarak raporlanmış; ana performans değerlendirmesi IoU/mIoU ve F1 üzerinden yapılmıştır.

3.4.5. Eşikleme (Threshold) ve Olasılık Haritası

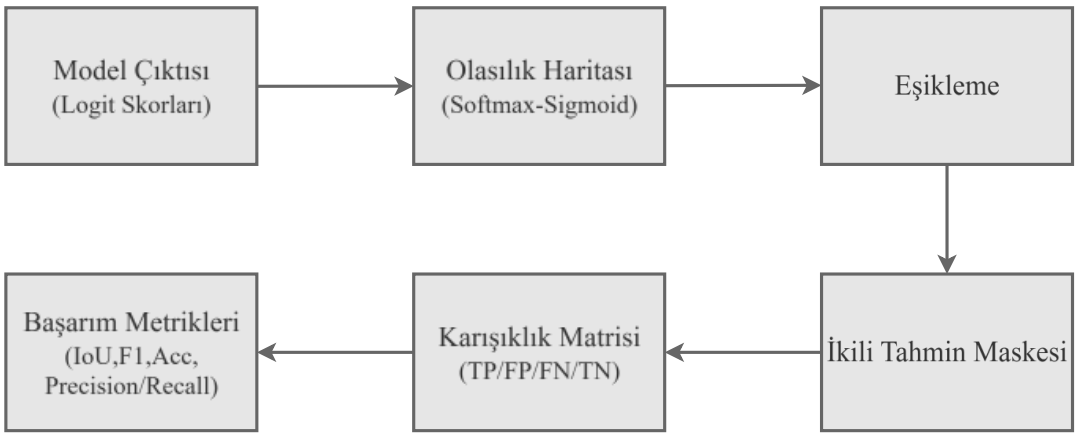
Model çıktısı, her piksel için hasar olasılığına çevrilip bir eşik değeri ile ikili maskeye dönüştürülür. Bu tezde olasılık haritası üretimi için ortak değerlendirme hattında "sigmoid_diff" yaklaşımı kullanılmıştır: iki sınıflı logit üretildiğinde hasar ve arka plan logitleri arasındaki fark sigmoid ile olasılığa çevrilir; bu, ikili karar üretimini tek bir olasılık haritası üzerinden standartlaştırır. Eşik değeri, doğrulama seti üzerinde taranarak (threshold sweep) seçilir ve testte sabit tutulur; böylece test seti ayar yapmak için kullanılmadığından raporlama adil kalır.

4. BULGULAR

4.1. Deneylerin Genel Özeti ve Raporlama Protokolü

Bu tez kapsamında, Türkiye 2023 depremi sonrası uzaktan algılama görüntülerinden bina hasarının piksel seviyesinde tespit edilmesi problemi ele alınmıştır. Problem, ikili semantik segmentasyon biçiminde kurgulanmış; her piksel hasar ve arka plan sınıflarından birine atanmıştır. Deneyler, operasyonel kullanım açısından kritik olan tek görüntü (single-image) senaryosunu hedeflemekte; yani afet öncesi görüntü zorunlu kılınmadan, yalnızca afet sonrası görüntü üzerinden hasar bölgelerinin ayrıştırılması amaçlanmaktadır. Bu bağlamda model çıktısı, görüntü üzerinde hasarlı bölgeleri işaretleyen ikili bir maske üretimi olarak değerlendirilmiştir.

Deneysel çalışma, aynı veri seti bölünmesi ve aynı değerlendirme mantığı altında, farklı modelleme ve eğitim stratejilerini karşılaştıracak şekilde tasarlanmıştır. Bu amaçla beş ana deney hattı oluşturulmuştur: SegFormer tabanlı temel yaklaşım (Baseline), IoU odaklı bir kayıp terimi ile güçlendirilen SegFormer (Lovász-Softmax), veri artırma ile genellemesi güçlendirilen SegFormer (Augmentation), transformer tabanlı alternatif bir segmentasyon mimarisi olan UPerNet + Swin (kıyas modeli) ve SegFormer (Augmentation) versiyonuna çıkarım aşamasında performans artırımı hedefleyen test-time augmentation (TTA) stratejisinin eklendiği SegFormer (Augmentation+TTA) deneyi. Bu deney hatları, hem mimari düzeyde hem de eğitim-çıkarım stratejileri düzeyinde farklılıklar taşımaktadır. Ancak karşılaştırmanın adil olabilmesi için veri erişimi ve raporlama protokolü sabit tutulmaktadır.



Şekil 4.1. Değerlendirme akışı.

Şekil 4.1'de deney süreçlerine ait değerlendirme akışı verilmiştir. Değerlendirme protokolünün temel amacı, farklı denemeler arasında eşit şartlarda karşılaştırmayı sağlayacak tek bir ölçüm standardı oluşturmaktır. Bu nedenle, tüm modeller için raporlanan metrikler aynı mantıkla hesaplanmıştır. Modelin ürettiği sürekli skor haritası, belirlenen bir eşik üzerinden ikili maske haline getirilmiş, ardından piksel seviyesinde global confusion matrix (TP, FP, TN, FN) oluşturulmuştur. Bu yaklaşımda metrikler tek tek görüntü bazında ortalanmak yerine, tüm veri bölünmesi üzerinde piksel sayımları biriktirilerek hesaplanmıştır. Böylece sonuçlar, görüntüler arası örnek büyüklüğü farklarından daha az etkilenmekte ve özellikle sınıf dengesizliği içeren problemlerde daha istikrarlı bir ölçüm çerçevesi sunmaktadır. Başka bir ifadeyle, bu tezde kullanılan değerlendirme yöntemi "global pixel-level" raporlamayı esas alır.

Modelin hasar sınıfı için ürettiği olasılık/puan haritası, doğrulama seti üzerinde eşik taraması yapılarak en uygun eşik değeri seçilecek şekilde optimize edilmiştir. Bu seçim, doğrudan test seti üzerinde yapılmamıştır. Bunun temel gerekçesi, test setinin nihai ve tarafsız raporlama amacı taşımasıdır. Eğer eşik değeri test seti üzerinde optimize edilirse, test performansı gerçek genelleme kabiliyetini yansıtmayabilir ve sonuçlar iyimserleşebilir. Bu nedenle protokol şu şekilde uygulanmıştır: önce validation üzerinde eşik taraması yapılmış, belirlenen eşik değeri sabitlenmiş ve aynı eşik ile test setinde tek seferlik nihai değerlendirme raporlanmıştır. Bu düzen, model seçimi ve küçük ayarların doğrulama seti üzerinde yapılmasını, test setinin ise yalnızca sonuç raporu için kullanılmasını sağlar.

Deneylerin raporlanmasında iki ayrı amaç aynı anda gözetilmiştir. Birinci amaç, modellerin karşılaştırmalı performansını net biçimde göstermektir. Bu kapsamda, tüm denemeler için hem VAL hem TEST üzerinde mIoU ve F1 gibi ana metrikler raporlanmış; ayrıca doğruluk (accuracy) ve piksel seviyesinde confusion matrix tabloları sunulmuştur. İkinci amaç ise, gerçek uygulama senaryolarında karar mekanizmasının nasıl kurulacağını şeffaflaştırmaktır. Bu nedenle raporlamada yalnızca tek bir metrik verilmemiş; eşik seçiminin etkisi, hasar sınıfında yanlış alarm (FP) ve kaçırma (FN) davranışıyla ilişkilendirilerek yorumlanabilir hale getirilmiştir. Özellikle hasar tespiti gibi afet müdahalesi bağlamında, yüksek doğruluk tek başına yeterli bir gösterge olmayabileceğinden, mIoU ve F1 skorlarının birlikte verilmesi tercih edilmiştir.

Bu tezde, en iyi model seçimi yalnızca doğrulama performansına dayandırılmıştır. Eğitim sırasında model ağırlıkları belirli aralıklarla kontrol edilmiş; doğrulama setinde hedef metriğe göre en iyi değer üreten ağırlıklar "best checkpoint" olarak saklanmıştır. Böylece test değerlendirmesinde kullanılan ağırlıkların, test verisini görmeden seçildiği garanti altına alınmıştır. Bu prosedür, tekrarlanabilirlik ve tarafsız raporlama açısından kritik olup, tez boyunca sunulan tüm nihai performans değerleri bu saklanan checkpoint'ler üzerinden üretilmiştir.

Sonuç olarak bu bölümde tanımlanan protokol, tez boyunca yürütülen tüm deneylerin aynı ölçütler altında karşılaştırılmasını sağlayan temel çerçeveyi oluşturur. Bir sonraki bölümde, bu protokole göre elde edilen doğrulama ve test sonuçları nicel olarak sunulacak olup; modeller arası farklar, performans artışının hangi stratejilerle geldiğini gösterecek şekilde sistematik biçimde karşılaştırılacaktır.

4.2. Nicel Sonuçların Karşılaştırması (VAL ve TEST)

Bu bölümde, Bölüm 4.1'de tanımlanan değerlendirme protokolü izlenerek elde edilen nicel sonuçlar sunulmakta ve modeller arası performans farklılıkları doğrulama (VAL) ve test (TEST) kümeleri üzerinde karşılaştırılmaktadır. Karşılaştırma, aynı veri seti bölünmesi, aynı metrik hesaplama mantığı (global pixel-level confusion matrix) ve aynı raporlama düzeni altında yapılmıştır. Böylece, elde edilen değerler doğrudan karşılaştırılabilir nitelik taşımaktadır.

4.2.1. Karşılaştırma Senaryoları ve Adil Protokol

Karşılaştırma kapsamında beş ana deney hattı raporlanmıştır: SegFormer Baseline, SegFormer + Lovász, SegFormer + Augmentation, SegFormer + Augmentation + TTA ve UPerNet + Swin. Bu hatların her biri, aynı eğitim/veri erişimi koşullarında yürütülmüş; farklılıklar yalnızca ilgili yöntemin getirdiği mimari veya stratejik değişikliklerle sınırlı tutulmuştur. Eşik değeri her bir model için doğrulama seti üzerinde seçilmiş; test setinde ise bu değer sabitlenerek tek seferlik nihai performans raporlanmıştır. Bu yaklaşım, test setinin model seçimi veya hiperparametre ayarı amacıyla kullanılmasını engelleyerek, test performansını daha güvenilir bir genelleme göstergesi haline getirir.

Bu tezde raporlanan temel metrikler mIoU ve F1'dir. mIoU, iki sınıflı segmentasyon probleminde hem arka plan hem hasar sınıfının kesişim-birleşim oranlarının ortalaması olarak hesaplanmış ve genel segmentasyon başarısını özetleyen bir ölçüt olarak

kullanılmıştır. F1 skoru ise özellikle hasar sınıfında yanlış alarm (FP) ve kaçırma (FN) dengesini birlikte yansıtması nedeniyle kritik bir tamamlayıcı metriktir. Ek olarak doğruluk (accuracy) ve confusion matrix değerleri, sonuçların daha somut yorumlanabilmesi için raporlamaya dahil edilmiştir.

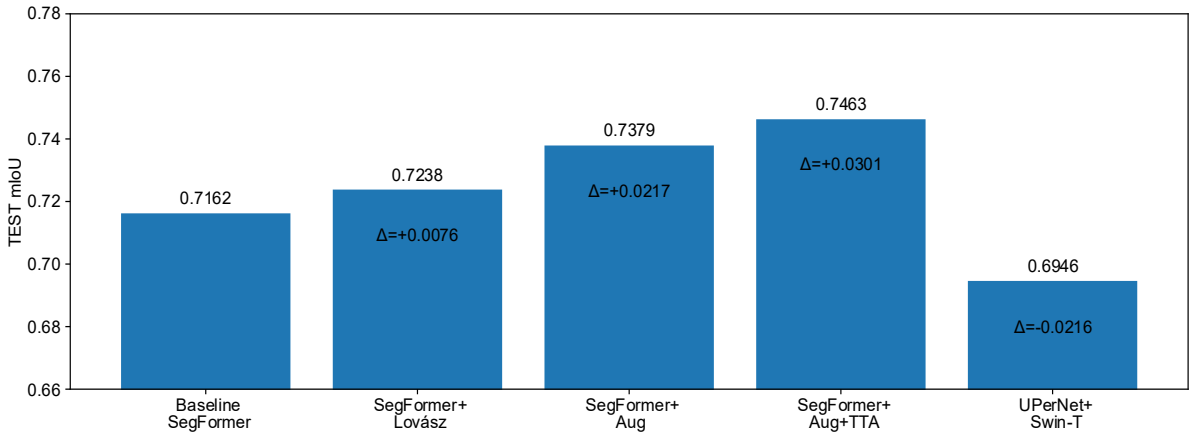
4.2.2. VAL ve TEST Sonuçları

Tablo 4.1'de tüm deney hatlarının VAL ve TEST kümeleri üzerindeki mIoU ve F1 değerleri, ayrıca her bir model için VAL üzerinde seçilen threshold değeri verilmiştir. Bu tablo, modellerin göreli performansını tek bakışta görmeyi ve hangi stratejilerin daha etkili iyileştirme sağladığını değerlendirmeyi mümkün kılmaktadır.

Tablo 4.1. Çalışmaya ait deneylerin VAL/TEST sonuçlarının karşılaştırılması.

Model / Deneme	Threshold	VAL mIoU	VAL F1	TEST mIoU	TEST F1
SegFormer Baseline	0.48	0.7114	0.6131	0.7162	0.6207
SegFormer +Lovász	0.56	0.7138	0.6165	0.7238	0.6338
SegFormer+ Augmentation	0.38	0.7266	0.6400	0.7379	0.6591
SegFormer+ Augmentation + TTA	**0.42**	**0.7294**	**0.6447**	**0.7463**	**0.6732**
Swin + UPerNet	0.30	0.6919	0.5761	0.6946	0.5789

Tablo 4.1'de raporlanan sonuçlar, test kümesinde en yüksek başarımın SegFormer + Augmentation + TTA hattı ile elde edildiğini göstermektedir (mIoU=0.7463, F1=0.6732). Baseline'a kıyasla en belirgin iyileşme, veri artırma stratejilerinin devreye alınmasıyla gözlenmiş hem IoU hem de F1 değerlerinde tutarlı bir artış sağlanarak hasar sınıfındaki kaçırmaların (FN) azaltıldığı ve genel örtüşme başarımının yükseldiği görülmüştür. Buna karşılık Swin+ UPerNet modeli, daha güçlü bir omurga kullanılmasına rağmen bu veri seti ve değerlendirme protokolü altında SegFormer Baseline'ın dahi gerisinde kalmıştır. Bu durum, modelin hasar sınıfında daha fazla kaçırma (FN) üretmesi ve sınıf dengesizliği koşullarında SegFormer tabanlı hattın daha uygun bir hata profili sergilediğini göstermektedir.



Şekil 4.2. Baseline'a göre TEST mIoU değişim karşılaştırması.

Şekil 4.2'de, test kümesindeki mIoU değerlerinin baseline SegFormer hattına göre değişimi karşılaştırmalı olarak sunulmakta ve SegFormer + Augmentation + TTA hattının 0.7463 mIoU ve 0.6732 F1 ile en iyi sonucu verdiği görülmektedir. Bu bulgu, performans artışının tek bir müdahaleden ziyade, eğitim aşamasında genellemeyi güçlendiren veri artırma ile çıkarım aşamasında tahmin kararlılığını artıran TTA'nın birlikte kullanılmasından kaynaklandığını göstermektedir.

4.2.3. Confusion Matrix (Karmaşıklık Matrisi) Analizi

Confusion matrix, sınıflandırma ve segmentasyon modellerinin performansını değerlendirmek amacıyla yaygın olarak kullanılan bir analiz aracıdır. Bu matris, model tarafından yapılan doğru ve yanlış sınıflandırmaları True Positive (TP), False Positive (FP), True Negative (TN) ve False Negative (FN) kategorileri üzerinden göstermektedir. Bu çalışmada, model performansını daha kapsamlı değerlendirebilmek amacıyla hem piksel bazlı hem de yapı bazlı confusion matrix analizleri gerçekleştirilmiştir.

Tablo 4.2. Piksel bazlı Confusion Matrix sonuçları (VAL+TEST).

Segformer-Baseline	Model Tahmini 0	Model Tahmini 1
Gerçek Etiket 0	20054519	171085
Gerçek Etiket 1	189054	294718

Segformer-Lovasz	Model Tahmini 0	Model Tahmini 1
Gerçek Etiket 0	20087826	137778
Gerçek Etiket 1	195419	288353

Segformer-Aug	Model Tahmini 0	Model Tahmini 1
Gerçek Etiket 0	20078711	146893
Gerçek Etiket 1	173783	309989

Segformer-Aug-TTA	Model Tahmini 0	Model Tahmini 1
Gerçek Etiket 0	20098387	127217
Gerçek Etiket 1	173744	310028

Swin-UperNet	Model Tahmini 0	Model Tahmini 1
Gerçek Etiket 0	20082441	143163
Gerçek Etiket 1	228378	255394

Bu çalışmada raporlanan piksel bazlı confusion matrix değerleri Tablo 4.2'te verilmiştir. Bu değerler, KATE-PD probleminde belirgin bir sınıf dengesizliği bulunduğunu açık biçimde göstermektedir: arka plan (sınıf 0) piksel sayısı, hasar (sınıf 1) piksel sayısından çok daha fazladır. Örneğin Baseline modelde TN değerinin 20 milyon seviyesinde olması, yalnızca doğruluk (accuracy) metriğinin yüksek çıkmasına yol açabilmekte; ancak bu durum modelin hasar sınıfını ne kadar iyi yakaladığına dair tek başına güvenilir bir ölçüm sunmamaktadır. Bu nedenle tez boyunca hasar sınıfındaki yanlış alarm (FP) ve kaçırma (FN) davranışını birlikte yansıtan F1, ayrıca örtüşme başarımını doğrudan ölçen IoU metriği öncelikli değerlendirme ölçütleri olarak ele alınmıştır.

Confusion matrix'ler yöntemlerin hata profilini net biçimde ayırt etmektedir. Baseline senaryoda hasar sınıfı için FP=171,085 ve FN=189,054 iken, Lovász eklenmesi FP'yi azaltarak (137,778) modelin "yanlış alarm" üretme eğilimini düşürmüş; buna karşılık FN değerinde sınırlı bir artış (195,419) görülmüştür. Bu durum, Lovász-Softmax teriminin modelin kararlarını daha "temkinli" hale getirerek kesinliği (precision) artırırken (0.6327 → 0.6767), duyarlılığı (recall) bir miktar baskılayabildiğini göstermektedir. Veri artırma (Augmentation) ile birlikte FN'nin belirgin biçimde azalması (173,783) hasar alanlarının daha az kaçırıldığını; hem FP hem FN'nin dengeli biçimde iyileşmesiyle F1'in yükseldiğini göstermektedir. En iyi sonuç veren Augmentation+TTA hattı ise FP'yi en düşük seviyeye indirirken (127,217) TP'yi yüksek tutmuş (310,028) ve bu sayede kesinlik (0.7090) ve F1 (0.6732) açısından en dengeli hata profiline ulaşmıştır. Buna karşılık Swin-UPerNet modelinde FN'nin belirgin biçimde artması (228,378), modelin hasarlı pikselleri kaçırma eğiliminin yükseldiğine işaret etmektedir.
Piksel bazlı confusion matrix sonuçları modelin piksel düzeyindeki sınıflandırma başarımını gösterirken, yapı bazlı analiz modelin nesne bütünlüğünü ne ölçüde koruyabildiğini ortaya koymaktadır. Özellikle afet sonrası hasar tespiti gibi uygulamalarda, yalnızca piksel düzeyinde doğruluğun yüksek olması modelin pratik kullanım açısından yeterli olduğunu garanti etmemektedir.

Yapı bazlı değerlendirme yalnızca test veri kümesi üzerinde gerçekleştirilmiştir. Bunun nedeni, nesne bazlı analizlerin modelin nihai genelleme performansını değerlendirmek amacıyla genellikle bağımsız test verisi üzerinde raporlanmasıdır. Doğrulama kümesi (VAL) bu çalışmada hiperparametre ayarı ve model seçimi için kullanıldığından, yapı bazlı analizlerin aynı veri üzerinde gerçekleştirilmesi model seçim sürecine dolaylı olarak etki edebilecek yanlılıklar oluşturabilir. Bu nedenle yapı bazlı confusion matrix analizi yalnızca bağımsız test kümesi üzerinde gerçekleştirilmiş ve modelin gerçek performansını yansıtacak şekilde raporlanmıştır.

Modelin hasarlı yapıları bütüncül biçimde tespit edebilmesi, operasyonel kullanım açısından önemli bir kriterdir. Bu nedenle piksel bazlı analizlere ek olarak, model çıktılarının yapı bazlı performansını değerlendirmek amacıyla nesne bazlı bir eşleşme analizi gerçekleştirilmiştir. Bu analiz kapsamında ground truth maskeleri içerisindeki her

bir yapı ayrı bir nesne olarak ele alınmış ve model tarafından üretilen tahmin maskeleri ile bu yapılar arasındaki örtüşme oranı hesaplanmıştır.

Nesne tabanlı görüntü analizi çalışmalarında kullanılan örtüşme eşikleri uygulama bağlamına göre değişkenlik gösterebilmektedir ve literatürde tek bir evrensel eşik değeri bulunmamaktadır [52, 53]. Özellikle uzaktan algılama uygulamalarında nesne geometrilerinin düzensiz olması ve segmentasyon çıktılarının parçalı yapıda olabilmesi nedeniyle katı IoU eşikleri çoğu zaman aşırı kısıtlayıcı olabilmektedir. Bu nedenle bina çıkarımı ve afet hasarı tespiti gibi çalışmalarda kısmi örtüşmeye dayalı nesne eşleşme kriterleri kullanılabilmektedir [54].

Bu çalışmada da hasar bölgelerinin düzensiz ve parçalı geometrik yapısı dikkate alınarak %30 örtüşme eşiği kullanılmıştır. Buna göre referans bir yapının en az %30'luk kısmı tahmin maskesi tarafından kapsandığında ilgili yapı doğru tespit edilmiş (True Positive) kabul edilmiştir. Hiçbir tahmin maskesi ile %30 örtüşme sağlayamayan referans yapılar False Negative, referans yapılarla yeterli örtüşmeye sahip olmayan tahmin maskeleri ise False Positive olarak değerlendirilmiştir.

Yapı bazlı değerlendirme sonucunda elde edilen değerler Tablo 4.3'te sunulmuştur.

Tablo 4.3. Yapı bazlı Confusion Matrix sonuçları (TEST).

Model	TP	FP	FN
SegFormer Baseline	107	58	35
SegFormer + Aug	103	60	39
SegFormer + Aug + TTA	112	51	30
SegFormer + Lovasz	105	50	37
Swin-UPerNet	90	59	52

Tablo incelendiğinde farklı model konfigürasyonlarının yapı bazlı tespit performanslarında belirgin farklılıklar gösterdiği görülmektedir. Özellikle SegFormer + Aug + TTA modelinin en yüksek TP ve en düşük FN değerlerine sahip olması, test-time augmentation yaklaşımının farklı bakış açılarından elde edilen tahminleri birleştirerek yapı bütünlüğünü güçlendirdiğini göstermektedir. Buna karşılık bazı deney hatlarında

sınır doğruluğunu artıran yöntemlerin hasar bölgelerini daha parçalı tahmin etmesine bağlı olarak nesne bazlı performansın görece düşebildiği gözlemlenmiştir.

Bu sonuçlar, piksel bazlı metriklerin model performansını değerlendirmede önemli olmakla birlikte, özellikle afet hasarı tespiti gibi uygulamalarda yapı bazlı değerlendirme ile birlikte ele alınmasının daha kapsamlı ve gerçekçi bir analiz sunduğunu ortaya koymaktadır.

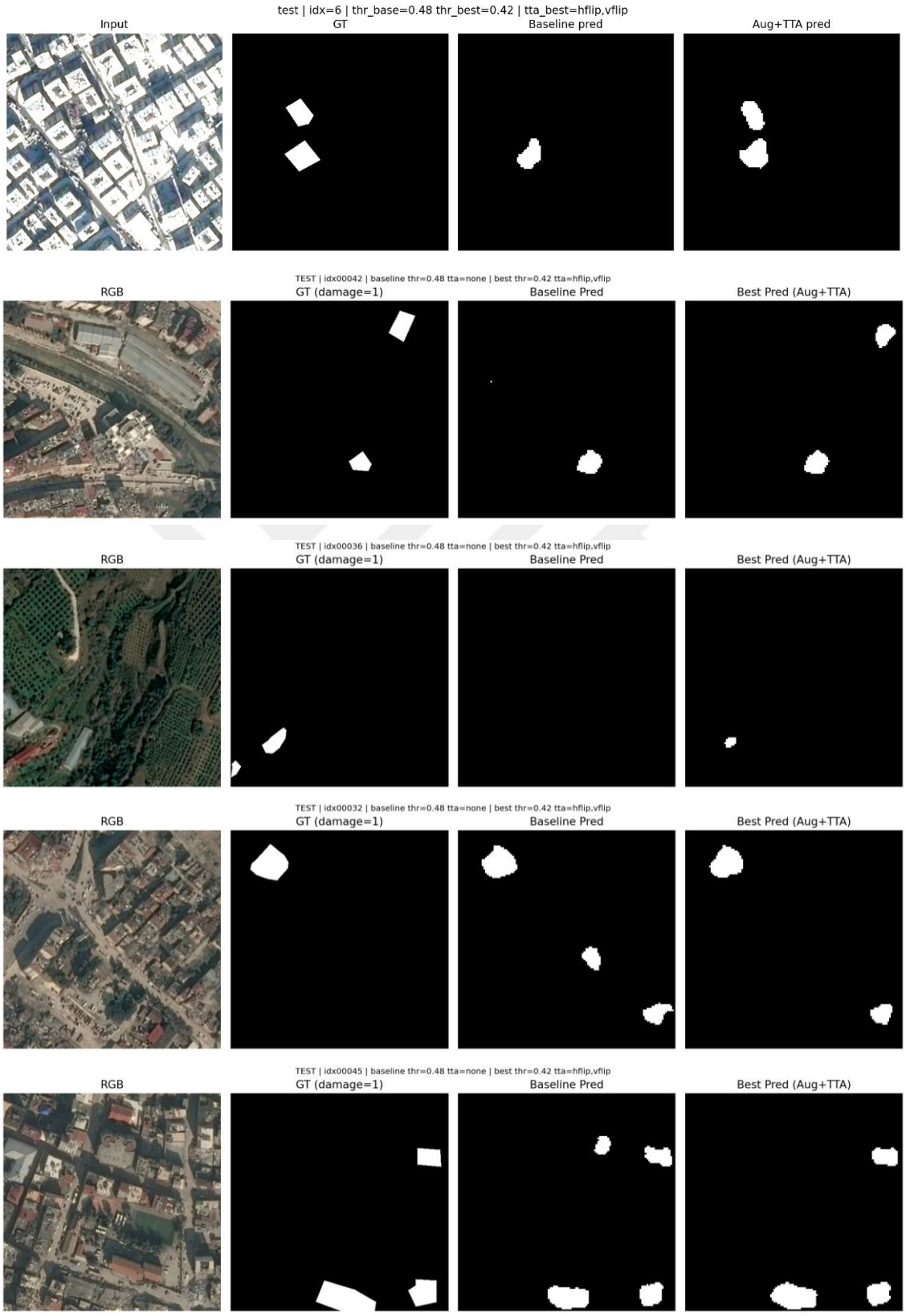
Confusion matrix üzerinden elde edilen bulguların daha yorumlanabilir hale getirilmesi amacıyla, modellerin hasarlı yapı sınıfını yakalama başarısını ifade eden duyarlılık (recall) değerleri hem piksel hem de yapı bazlı olarak hesaplanmıştır. Bu iki farklı değerlendirme düzeyi birlikte ele alınarak, modellerin yalnızca piksel düzeyinde değil, aynı zamanda yapı bütünlüğü açısından da performans farklılıkları ortaya konulmuştur. Elde edilen sonuçlar Tablo 4.4'te sunulmaktadır.

Tablo 4.4. Piksel ve yapı bazlı recall değerlerinin karşılaştırılması (TEST).

Model	Piksel Bazlı Recall	Yapı Bazlı Recall
SegFormer Baseline	0.609	0.754
SegFormer + Aug	0.641	0.725
SegFormer + Aug + TTA	0.641	0.789
SegFormer + Lovasz	0.596	0.739
Swin-UPerNet	0.528	0.634

Piksel ve yapı bazlı recall değerleri birlikte incelendiğinde, bazı modellerin piksel düzeyinde benzer performans göstermesine rağmen yapı bazlı değerlendirmede farklılaştığı görülmektedir. Özellikle veri artırma tek başına piksel bazlı recall'ı artırırken, yapı bazlı recall'da düşüşe yol açmıştır. Buna karşılık TTA, piksel bazlı performansı sabit tutarken yapı bazlı recall'ı artırarak daha tutarlı ve bütüncül tahminler üretmiştir.

4.2.4. Nitel Sonuçların Görsel Analizi



Şekil 4.3. Örnek tahminlerin görsel olarak karşılaştırılması.

Şekil 4.3'te yer alan örnekler, iyileştirilmiş hattın (Aug+TTA) yalnızca metrikleri artırmakla kalmayıp, tahmin maskesinin geometrik bütünlüğünü ve GT ile görsel tutarlılığını da belirgin biçimde iyileştirdiğini göstermektedir. Baseline modelin bazı örneklerde hasarı daha küçük adacıklara bölerek parça parça ürettiği ve sınırları tam oturtamadığı görülürken; Aug+TTA ile tahminlerin daha dolu, daha sürekli ve GT'deki şekle daha yakın hale geldiği izlenmektedir. Bu durum, veri artırmanın eğitim aşamasında modele farklı örüntüler karşısında daha sağlam temsil kazandırması; TTA'nın ise çıkarım sırasında farklı dönüşümlerden gelen tahminleri birleştirerek rastlantısal hataları azaltmasıyla açıklanabilir.

Nitel inceleme, hata türlerinin de daha anlaşılır hale geldiğini göstermektedir: Baseline modelin bazı durumlarda gölge, koyu asfalt/yol parçaları veya çatı üzerindeki düzensiz doku gibi alanları hasar sınıfıyla karıştırarak yanlış alarm (FP) üretebildiği; Aug+TTA hattında ise bu tür yanlış pozitiflerin daha sınırlı kaldığı izlenmektedir. Öte yandan Baseline'ın küçük, dağınık veya düşük kontrastlı hasar izlerini arka plan olarak bırakma eğilimi (FN) de dikkat çekmektedir. Augmentation ve TTA birlikte kullanıldığında, hem hasar bölgelerinin daha bütüncül yakalandığı hem de küçük/dağınık hasar parçalarının kaçırılmasının azaldığı; dolayısıyla GT ile daha tutarlı bir maske üretildiği görülmektedir.

4.2.5. Baseline ve Kayıp Fonksiyonu Değişikliğinin Etkisi (Lovász)

SegFormer Baseline hattı, standart eğitim yaklaşımını temsil etmekte ve diğer tüm yöntemler için referans noktası oluşturmaktadır. Baseline modelin test setindeki mIoU değeri 0.7162 iken, SegFormer + Lovász hattında bu değer 0.7238'e yükselmiştir. Benzer şekilde F1 skoru 0.6207'den 0.6338'e çıkmıştır. Bu artış, Lovász-Softmax teriminin doğrudan IoU hedefiyle ilişkili bir optimizasyon baskısı oluşturarak, hasar sınıfında doğru bölge örtüşmesini (overlap) artırmaya yardımcı olabileceğini düşündürmektedir. Bununla birlikte iyileşme düzeyi, augmentation ve TTA gibi stratejilerin getirdiği kazanıma kıyasla daha sınırlı kalmıştır. Bu durum, bu veri setinde başarımı belirleyen ana faktörlerden birinin modelin görsel çeşitliliğe dayanıklılığı (genelleme) olabileceğini işaret eder.

Bu noktada, Lovász yaklaşımının tezde bir "yapısal katkı" olarak sunulması mümkündür; çünkü kayıp fonksiyonu düzeyinde yapılan bu değişiklik, hedef metrikle daha uyumlu bir optimizasyon yaklaşımı sunmaktadır. Ancak deney sonuçları, tek başına kayıp

fonksiyonu değişikliğinin performans tavanını belirgin biçimde yükseltmediğini, daha çok ince ayar seviyesinde artış sağladığını göstermektedir.

4.2.6. Veri Artırma (Augmentation) Etkisi: En Büyük Kazanç

SegFormer + Augmentation hattı, baseline'a göre en belirgin iyileşmenin elde edildiği stratejilerden biridir. Test setinde mIoU 0.7162'den 0.7379'a; F1 ise 0.6207'den 0.6591'e yükselmiştir. Bu artış, hasar sınıfının görsel olarak yüksek çeşitlilik göstermesi (farklı çatı tipleri, farklı yıkım örüntüleri, gölge/ışık koşulları, çevresel dokular) nedeniyle, modelin eğitim sırasında daha geniş bir örneklem uzayına maruz bırakılmasının özellikle faydalı olabileceğini göstermektedir. Veri artırma, modelin belirli geometrik yönelimlere (ör. bina yönü) veya sınırlı görsel koşullara aşırı uyum göstermesini azaltır; bu da test performansındaki artışla tutarlı bir gözlemdir.

Tez bağlamında augmentation hattı, hem pratik uygulanabilirliği hem de etkisinin büyüklüğü nedeniyle ana katkı unsurlarından biri olarak vurgulanabilir. Çünkü bu strateji, ek veri toplamadan, sadece veri uzayında çeşitlilik üreterek daha iyi genelleme sağlamaktadır.

4.2.7. TTA'nin Katkısı: Çıkarım Aşamasında Ek Kazanç

SegFormer + Augmentation + TTA hattı, augmentation-only hattının bir adım ötesinde performans sunmuştur. Test setinde mIoU 0.7379'dan 0.7463'e, F1 ise 0.6591'den 0.6732'ye yükselmiştir. TTA'nın temel etkisi, tek bir görüntü üzerinden yapılan tahminin rastlantısal gürültüsünü azaltmak ve modelin kararını daha kararlı hale getirmektir. Görüntünün yatay/dikey çevrilmiş versiyonları üzerinde yapılan tahminler birleştirildiğinde, modelin sınır bölgelerinde ve küçük hasar alanlarında daha tutarlı maske üretmesi beklenir. Bu durum özellikle hasarın sınırlarının belirsiz olduğu (moloz-yol/çatı dokusu karışımı) bölgelerde fayda sağlayabilir.

Önemli bir nokta, TTA'nın eğitim sürecini değiştirmemesi, yalnızca çıkarım maliyeti eklemesidir. Bu nedenle tezde TTA, "performansı artıran fakat hesaplama maliyetini yükselten opsiyonel bir inference stratejisi" olarak konumlandırılmalıdır. Bu yönüyle TTA, uygulama bağlamında "doğruluk–maliyet" dengesinde seçilebilir bir iyileştirme olarak sunulabilir.

4.2.8. Swin-UPerNet Sonuçları: Transformer Tabanlı Alternatif Kıyas

UPerNet + Swin hattı, transformer tabanlı ikinci bir segmentasyon yaklaşımı olarak kıyas amacıyla yürütülmüştür. Sonuçlar, bu hattın SegFormer tabanlı en iyi yaklaşımların gerisinde kaldığını göstermektedir. Test setinde Swin-UPerNet 0.6946 mIoU ve 0.5789 F1 değerlerine ulaşırken, SegFormer tabanlı hatlar bu değerlerin üzerinde performans üretmiştir. Bu durum, seçilen veri seti ve eğitim protokolü altında SegFormer'ın çok ölçekli özellik birleştirme yapısı ve hafif decoder tasarımının daha avantajlı bir doğruluk–maliyet dengesi sunabildiğine işaret etmektedir. Modellerin hesaplama karmaşıklığı karşılaştırmasının verildiği tablo 4.5'te de görüldüğü gibi Swin-UPerNet hattının eğitim süresinin daha uzun olması ve kaynak gereksiniminin daha yüksek olması, pratik uygulama açısından SegFormer hattını daha cazip kılan bir faktör olarak tartışılabilir.

Tablo 4.5. Modellerin hesaplama karmaşıklığı karşılaştırması.

Model Parametre Sayısı	Eğitim Süresi (Epoch)	GPU (VRAM)	Operasyonel Avantaj	Model Parametre Sayısı
SegFormer (MiT-B2)	~25M	8–12 dk	RTX 3050 (4GB)	Düşük kaynak tüketimi, yüksek hız
Swin-UPerNet	~50M+	50–70 dk	RTX 3050 (4GB)	Yüksek bellek ve zaman maliyeti

4.2.9.Bölüm Özeti

Bu bölümde sunulan nicel sonuçlar, en etkili iyileştirmenin veri artırma ve çıkarım aşamasında TTA uygulaması ile elde edildiğini ortaya koymuştur. Baseline SegFormer hattı güçlü bir referans performans sunarken, Lovász kayıp terimi sınırlı fakat tutarlı bir artış sağlamıştır. Augmentation ise belirgin bir performans sıçraması üretmiş; TTA ile birlikte en yüksek mIoU ve F1 değerleri elde edilmiştir. Transformer tabanlı alternatif kıyas modeli Swin-UPerNet, bu çalışma koşullarında SegFormer hattının gerisinde kalmıştır. Bir sonraki bölümde, eşik (threshold) seçim sürecinin neden VAL üzerinde yapıldığı, TEST'te neden sabit tutulduğu ve eşik optimizasyonunun sonuçlara etkisi ayrıntılı olarak tartışılacaktır.

4.3. Eşik (Threshold) Seçimi ve Testte Sabitleme

Bu çalışmada modellerin ürettiği çıktı, piksel seviyesinde sürekli bir olasılık/puan haritasıdır. Bu haritanın ikili (hasar/hasarsız) maske haline gelebilmesi için bir eşik değeri tanımlanması gerekir. Eşik seçimi, segmentasyon çıktısının hassasiyet (precision) ve duyarlılık (recall) dengesini doğrudan etkilediğinden, raporlanan mIoU ve F1 değerlerinde belirleyici bir rol oynar. Bu nedenle eşik değeri rastgele sabitlenmemiş; her model için doğrulama (VAL) kümesi üzerinde tarama yapılarak en iyi performansı veren değer seçilmiştir.

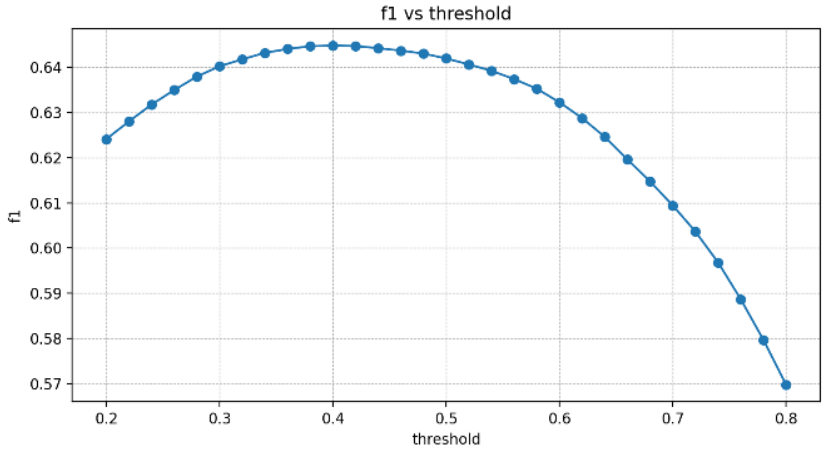
Eşiğin VAL üzerinde seçilmesinin temel gerekçesi, test kümesinin tarafsız bir genelleme değerlendirmesi için korunmasıdır. Test kümesi üzerinde eşik ayarı yapılırsa, test performansı istemeden test verisine uyarlanmış hale gelebilir ve gerçek saha performansını temsil etme gücü azalır. Bu nedenle izlenen protokol şöyledir: (i) VAL üzerinde belirlenen hedef metriğe göre en iyi eşik değeri bulunur, (ii) bu eşik sabitlenir, (iii) TEST kümesinde aynı eşikle tek seferlik nihai metrikler raporlanır. Böylece test sonuçları, model seçimi ve küçük ayarların etkisinden arındırılmış daha adil bir göstergeye dönüşür.

Bu tezde eşik optimizasyonunun pratik etkisi açık biçimde gözlenmiştir. Örneğin SegFormer tabanlı en iyi hat üzerinde yapılan taramada, eşik değerinin 0.38 civarından 0.42 seviyesine taşınması, hasar sınıfında üretilen pozitif alan miktarını ve dolayısıyla FP/FN dengesini değiştirmiş; bunun sonucunda TEST mIoU ve TEST F1 değerlerinde anlamlı bir artış elde edilmiştir. Eşik arttıkça model daha “temkinli” hale gelmekte yani daha az pikseli hasar olarak etiketleme eğiliminde olmakta, bu da yanlış alarmları (FP) düşürürken bazı gerçek hasar piksellerinin kaçırılmasına (FN) yol açabilmektedir. Eşik azaldığında ise model daha agresif hasar tahmini yapar; kaçırmaları azaltabilir fakat yanlış alarmı artırabilir. Bu nedenle tek bir sabit eşik tüm modeller ve tüm veri koşulları için optimal olmayabilir; model bazında VAL üzerinde belirlenen eşik, performansı daha dengeli hale getirmektedir.

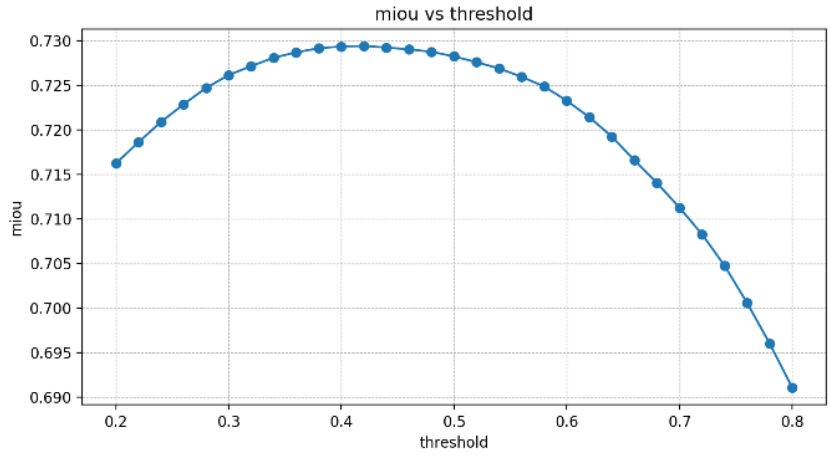
Bu noktada yapılan işlem hiperparametre optimizasyonunun sınırlı bir biçimi olarak değerlendirilebilir; ancak eğitim hiperparametrelerini (öğrenme oranı, batch size, epoch sayısı gibi) kapsamlı şekilde arayan bir süreç yürütülmemiştir. Burada optimize edilen şey model ağırlıkları değil, ağırlıklar sabitken çıkarım aşamasında uygulanan karar eşiğidir. Bu nedenle tez bağlamında bu adım, “model seçim sürecinin bir parçası olan

inference-time kalibrasyon" veya "çıktı eşiği ayarı" olarak ifade edilmelidir. Bu ayrım önemlidir: eğitim hiperparametre optimizasyonu, modelin öğrenme dinamiğini değiştirirken; eşik seçimi, öğrenilmiş skorların karar maskesine dönüşüm kuralını belirler.

Sonuç olarak bu çalışma, eşik seçiminin yalnızca küçük bir teknik ayrıntı olmadığını; özellikle sınıf dengesizliği ve sınır belirsizliği içeren hasar tespiti probleminde raporlanan metrikleri belirgin biçimde etkileyebildiğini göstermektedir. Bu nedenle tezde raporlanan tüm TEST sonuçları, model başına VAL üzerinde seçilen ve TEST'te sabitlenen eşik değerleri ile üretilmiş; böylece yöntemler arası karşılaştırmanın adil kalması hedeflenmiştir.



Şekil 4.4. SegFormer Augmentation TTA deneyinin threshold değerlerine göre f1 değişimi.



Şekil 4.5. SegFormer Augmentation TTA deneyinin threshold değerlerine göre mIoU değişimi.

Şekil 4.4. ve Şekil 4.5.'te bu çalışmada kullanılan deney versiyonlarından SegFormer + Augmentation + TTA'a ait threshold değerlerine göre F1 ve mIoU değişimlerini gösteren grafikler verilmiştir. Tüm deneylerin değerlendirme sürecinde uygulanan threshold sweep işlemi sonrasında en iyi sonucu veren threshold değerleri seçilmiştir.

Doğrulama kümesi üzerinde gerçekleştirilen eşik (threshold) taraması, modelin sürekli skor/olasılık haritasının ikili maske üretimine dönüştürülmesi aşamasında kritik bir adımdır. Bu tezde her model için eşik değeri VAL kümesi üzerinde belirlenmiş; test kümesinde ise aynı eşik sabit tutularak nihai metrikler tek seferde raporlanmıştır. Şekil 5'de görüldüğü üzere farklı deney hatları için optimal eşik değerleri değişebilmektedir (ör. Baseline 0.48, Lovász 0.56, Augmentation 0.38, Aug+TTA 0.42). Bu farklılık, modellerin olasılık kalibrasyonlarının ve "hasar" sınıfına verdikleri güven dağılımının birbirinden farklı olmasından kaynaklanır. Eşik taramasında amaç, FP (yanlış alarm) ve FN (kaçırma) dengesini optimize ederek F1/IoU gibi örtüşme temelli metriklerde en yüksek değeri veren noktayı seçmektir; böylece test değerlendirmesi, hiperparametre ayarı yapılmış bir modele değil, doğrulamada seçilmiş ve testte sabitlenen adil bir karara dayandırılmış olur.

4.4. KATE-PD Veri Seti Tanıtım Çalışması ile Karşılaştırma ve Raporlama Tutarlılığı

Bu tezde elde edilen nicel bulguların bilimsel bir temele oturtulabilmesi amacıyla, KATE-PD veri setini literatüre kazandıran çalışmada raporlanan referans sonuçlarla doğrudan bir karşılaştırma yürütülmüştür. Metodolojik tutarlılığı sağlamak adına, veri setinin eğitim, doğrulama ve test aşamaları için belirlenen %80, %10 ve %10 (675/78/79 karo) oranındaki bölme protokolü bu çalışmada da aynen korunmuştur. İlgili literatürde, ResNet18 omurgalı UNet mimarisi ve YOLO11x-seg segmentasyon modeli olmak üzere iki farklı temel yaklaşım raporlanmaktadır [20]. KATE-PD çalışmasında piksel düzeyindeki hasar olasılığı üretimi sigmoid tabanlı fonksiyonlar ile gerçekleştirilmekte; model başarısı ise mIoU, F1 ve Genel Doğruluk (OA) metrikleri üzerinden değerlendirilmektedir. Bu tez kapsamında yürütülen deneylerde de aynı değerlendirme standartları benimsenmiş; hasar ve arka plan ayrımı için ikili maskeler kullanılarak mIoU ve F1 metrikleri hesaplanmıştır. Ayrıca, karar mekanizmasının kararlılığını artırmak için olasılık haritalarına uygulanan eşik değerleri (threshold), test setinden bağımsız olarak doğrulama seti üzerinde optimize edilmiştir.

4.4.1. KATE-PD Referans Değerleri ve Baseline Analizi

Tez çalışması kapsamında elde edilen sonuçların nesnel bir zeminde değerlendirilebilmesi için, veri setini literatüre tanıtan çalışmada raporlanan referans değerler baz alınmıştır. KATE-PD çalışmasında sunulan bu metrikler; eğitim, doğrulama ve test süreçleri için mIoU, F1 ve Genel Doğruluk (OA) değerlerini kapsamaktadır. İlgili çalışmada raporlanan bulgulara göre, YOLO mimarisi SMP (Segmentation Models PyTorch) tabanlı UNet yaklaşımına kıyasla tüm metriklerde daha yüksek bir performans sergilemiştir [20].

Önemli bir ayrıntı olarak, referans çalışmada bu skorların herhangi bir hiperparametre optimizasyonu yapılmadan elde edildiği vurgulanmıştır. Yazarlar, bu sonuçların temel amacının veri setinin genel karakteristiğini tanıtmak olduğunu ve sunulan değerlerin gelecek araştırmalar için birer "baseline" niteliği taşıdığını belirtmişlerdir [20]. Bu tez çalışmasındaki karşılaştırmalı analizlerde, KATE-PD referans çalışmasındaki nihai test değerleri kriter olarak kabul edilmiştir:

SMP (Test Skoru): mIoU=0.4896, F1=0.0389 ve OA=0.9512 değerleri ile raporlanmıştır.

YOLO (Test Skoru): mIoU=0.6772, F1=0.5699 ve OA=0.9781 değerleri ile SMP modeline göre daha gelişmiş bir performans sunmuştur.

Bu referans değerler, bu tezde önerilen SegFormer tabanlı modellerin ulaştığı doğruluk seviyesinin literatürdeki ilk baseline sonuçlarına göre ne ölçüde bir gelişim gösterdiğini anlamak açısından kritik bir mukayese imkânı sağlamaktadır.

4.4.2. Bu Tezde Elde Edilen Sonuçlarla Kıyas

Bu tezde, aynı veri bölünmesi ve aynı problem tanımı altında SegFormer (MiT-B2) tabanlı bir hat değerlendirilmiş; eğitim stratejileri olarak baseline, Lovász ek kaybı, augmentation ve augmentation + TTA gibi varyantlar ile tranformer ailesinden farklı bir mimari olan Swin yapısının UperNet ile birlikte kullanıldığı model karşılaştırılmıştır. Elde edilen en iyi test sonucu, augmentation tabanlı model üzerinde test-time augmentation uygulanmasıyla elde edilmiştir (test mIoU=0.7463, F1=0.6732).

KATE-PD makalesinde yer alan modeller ile bu çalışmada kullanılan modellerin mIoU ve F1 metrikleri Tablo 4.6'da karşılaştırmalı olarak raporlanmıştır. Bu tezde sunulan değerler, ilgili makaledeki sigmoid tabanlı olasılık üretimi ve eşikleme yaklaşımı ile metodolojik olarak uyumlu şekilde hesaplanmıştır [20].

Tablo 4.6. KATE-PD veri seti tanıtım çalışması Baseline'ları ile bu tezde elde edilen sonuçların karşılaştırılması (TEST).

Yöntem / Model	TEST mIoU	TEST F1
KATE-PD – SMP (ResNet18 + Dice)	0.4896	0.0389
KATE-PD – YOLO11x-seg	0.6772	0.5699
Bu tez – SegFormer Baseline	0.7162	0.6207
Bu tez – SegFormer + Lovász	0.7238	0.6338
Bu tez – SegFormer + Augmentation	0.7379	0.6591
Bu tez – SegFormer + Augmentation + TTA	**0.7463**	**0.6732**
Bu tez – Swin + UPerNet	0.6946	0.5789

Karşılaştırma tablosu, KATE-PD makalesinde raporlanan YOLO baseline'ının (test mIoU=0.6772; F1=0.5699) üzerine, bu tezde geliştirilen SegFormer tabanlı hatların belirgin iyileşme sağladığını göstermektedir [20].

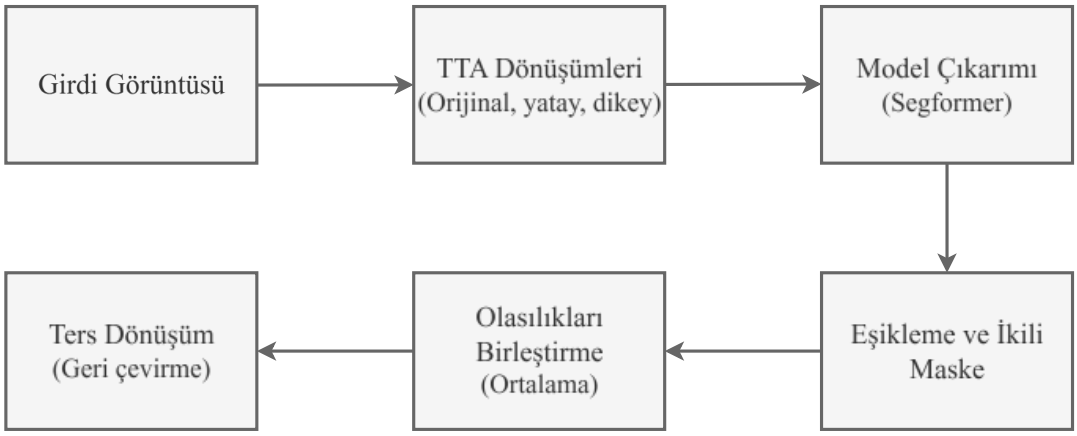
Özellikle veri artırma ile eğitilen SegFormer modelinin test performansını yükselttiği; test-time augmentation eklendiğinde ise mIoU ve F1 değerlerinin daha da arttığı görülmüştür. Bu bulgu, KATE-PD makalesinde de vurgulanan "bu skorların iyileştirilebilir baseline'lar olduğu" notuyla tutarlıdır; zira makalede hiperparametre optimizasyonuna gidilmediği ve amaçlarının veri seti sunumu olduğu belirtilmiştir [20].

Geliştirilen tüm SegFormer varyantları, KATE-PD makalesinde raporlanan en güçlü baseline olan YOLO modelini (%67.72 mIoU) geride bırakarak literatüre yeni bir başarı eşiği kazandırmıştır. Elde edilen bulgular, SegFormer mimarisinin uygun eğitim ve çıkarım stratejileriyle birleştiğinde, KATE-PD veri seti üzerinde literatürdeki mevcut yaklaşımlara göre çok daha güçlü ve kararlı bir alternatif oluşturduğunu kanıtlamaktadır.

4.5. Test Zamanı Artırma ve Son İşleme Adımları

Bu tezde nihai çıktı üretimi, yalnızca eğitim sırasında öğrenilen ağ ağırlıklarına bağlı bırakılmamış; çıkarım aşamasında modele ek bir "karar istikrarı" kazandırmak amacıyla test zamanı artırma (test-time augmentation, TTA) ve hafif son-işleme adımları değerlendirilmiştir. TTA'nın temel fikri, aynı görüntünün basit geometrik dönüşümler (yatay/dikey çevirme) altındaki tahminlerini bir araya getirerek daha kararlı bir olasılık haritası elde etmektir. Bu yaklaşım, pratikte özellikle sınıf sınırlarının belirsiz olduğu ve

küçük veya dağınık hasar bölgelerinin bulunduğu problemlerde gürültüyü azaltarak genelleme performansını artırabilen bir kestirim tekniği olarak raporlanmıştır [50]. Test zamanı arttırma yaklaşımının uygulama akışı şekil 4.5'te gösterilmiştir.



Şekil 4.6. Test zamanı artırma çıkarım akışı.

Bu çalışmada TTA, yatay ve dikey çevirme olmak üzere iki dönüşüm kümesi üzerinden uygulanmıştır. Her bir dönüşüm için modelin ürettiği hasar olasılık haritası, ters dönüşüm ile tekrar özgün koordinat sistemine taşınmış ve piksel-bazlı ortalama alma ile birleştirilmiştir. Bu süreç şu şekilde özetlenebilir:

Girdi görüntüsü için dönüşüm kümesi $\mathcal{T}$ tanımlanır (ör. $\{id, hflip, vflip\}$).

Her dönüşüm için tahmin olasılığı üretilir: $P_t = f_\theta(t(x))$.

Ters dönüşümle aynı uzaya getirilir: $\tilde{P}_t = t^{-1}(P_t)$.

Birleştirilmiş olasılık: $\bar{P} = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \tilde{P}_t$.

Son maske: $\hat{Y} = \mathbb{1}[\bar{P} > \tau]$.

TTA sonrasında uygulanan son-işleme adımları, küçük parça temizliği gibi basit gürültü azaltma işlemlerini kapsamaktadır. Özellikle ikili segmentasyon çıktılarında modelin tek tük piksel adacıkları üretmesi, salt-and-pepper benzeri artefaktlar şeklinde yaygındır; literatürde bu küçük artefaktların kaldırılmasının çıktıyı daha tutarlı hale getiren bir son-işleme adımı olarak kullanıldığı bildirilmektedir [55]. Bu tezde de benzer motivasyonla, bağlı bileşen temelli alan eşiğiyle çok küçük adacıkları temizleme opsiyonu "min_area" parametresi üzerinden değerlendirilmiştir. Connected-component tabanlı filtreleme, görüntü işleme literatüründe yaygın bir gürültü bastırma yaklaşımıdır [56].

4.6. İstatistiksel Analiz ve Model Karşılaştırması

Bu çalışmada farklı model konfigürasyonlarının hasar tespiti performansları arasında istatistiksel olarak anlamlı bir fark olup olmadığı incelenmiştir. Analizler, test kümesi üzerinde patch-bazlı hasar sınıfı IoU_1 skorları kullanılarak gerçekleştirilmiştir.

Test kümesi toplam 79 patch içermesine rağmen, yer gerçeğinde hasar bulunan ve IoU_1 değerinin tanımlı olduğu 26 patch istatistiksel analiz kapsamına dahil edilmiştir. Aynı patch'ler üzerinde aşağıdaki beş model konfigürasyonu karşılaştırılmıştır:

- SegFormer Baseline
- SegFormer + Lovász Loss
- SegFormer + Augmentation
- SegFormer + Augmentation + TTA
- Swin-UPerNet

IoU_1 değerlerinin sınırlı aralıkta olması, dağılımın çarpık özellik göstermesi ve örnek sayısının görece düşük olması (N = 26) nedeniyle parametrik testler yerine non-parametrik istatistiksel yöntemler tercih edilmiştir. Non-parametrik yöntemler, veri dağılımının normal olduğu varsayımını gerektirmediği için küçük örneklem büyüklüklerinde ve sıralama tabanlı karşılaştırmalarda daha güvenilir sonuçlar sunmaktadır [57]. İstatistiksel analizlerde hasar sınıfına ait IoU_1 değerleri kullanılmıştır. Bu tercih, arka plan piksel sayısının yüksek olduğu veri setlerinde mIoU değerlerinin arka plan sınıfı tarafından baskılanabilmesi nedeniyle yapılmıştır.

Bu kapsamda çoklu model karşılaştırması için Friedman testi uygulanmıştır. Friedman testi, aynı veri örnekleri üzerinde değerlendirilen birden fazla algoritmanın performanslarını karşılaştırmak amacıyla kullanılan sıralama tabanlı non-parametrik bir istatistiksel testtir [58]. Makine öğrenmesi alanında farklı sınıflandırıcıların veya model konfigürasyonlarının karşılaştırılmasında yaygın olarak kullanılan bir yöntemdir [59].

Friedman testinin anlamlı sonuç vermesi durumunda model çiftleri arasındaki farkların belirlenebilmesi amacıyla Nemenyi post-hoc testi uygulanmıştır. Nemenyi testi, Friedman testi sonrasında gerçekleştirilen çoklu karşılaştırmalar için kullanılan dağılım varsayımı gerektirmeyen bir yöntemdir ve model çiftleri arasındaki istatistiksel anlamlılık ilişkilerini belirlemek amacıyla kullanılır [60].

4.6.1. Friedman Testi Sonuçları

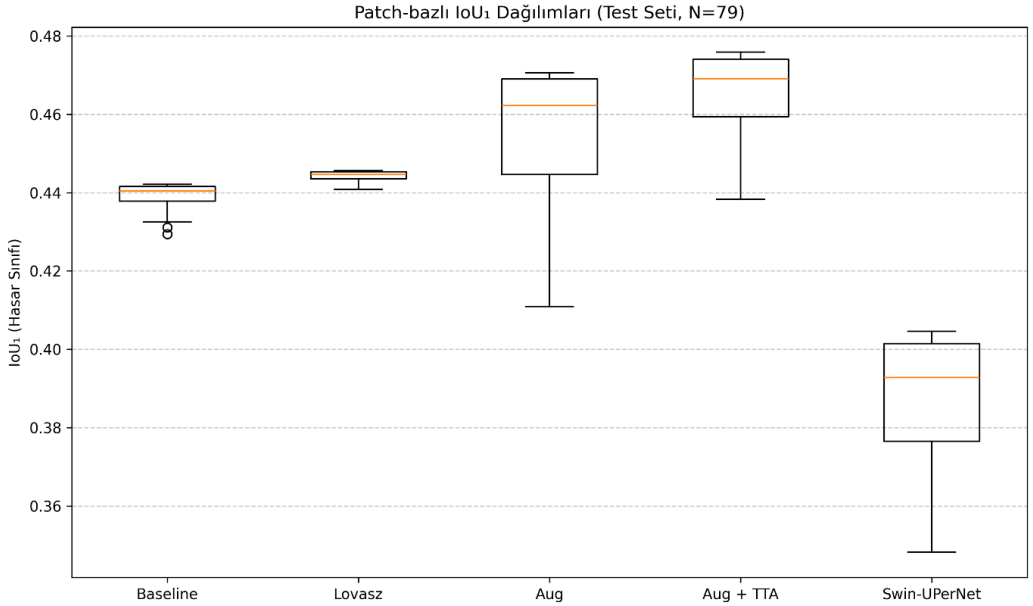
Friedman testi sonuçları Tablo 4.7'de sunulmaktadır.

Tablo 4.7. Friedman testi sonuçları.

Test	İstatistik	p-değeri
Friedman	88.7385	< 0.001

Sonuçlara göre $p < 0.05$ olduğundan karşılaştırılan modeller arasında istatistiksel olarak anlamlı bir fark bulunmaktadır. Bu bulgu, karşılaştırılan model konfigürasyonlarından en az birinin diğerlerinden anlamlı derecede farklı performans sergilediğini göstermektedir. Model performans dağılımlarının daha iyi anlaşılabilmesi amacıyla patch-bazlı IoU_1 değerleri kutu grafikleri (boxplot) ile görselleştirilmiştir. Şekil 4.7'de sunulan boxplot grafiği, farklı model konfigürasyonlarının test kümesi üzerindeki IoU_1 dağılımlarını ve varyanslarını göstermektedir.

Şekil 4.7 incelendiğinde özellikle Augmentation ve Augmentation + TTA konfigürasyonlarının diğer modellere kıyasla daha yüksek medyan IoU_1 değerlerine sahip olduğu görülmektedir. Bununla birlikte bu iki modelin dağılımlarının büyük ölçüde örtüştüğü ve varyanslarının benzer olduğu dikkat çekmektedir. Bu durum iki model arasındaki farkın istatistiksel olarak anlamlı çıkmamasının olası nedenlerinden biri olarak değerlendirilebilir.



Şekil 4.7. Patch-bazlı IoU_1 dağılımları (Test seti, N = 26).

4.6.2. Nemenyi Post-hoc Testi Sonuçları

Friedman testinin anlamlı çıkmasının ardından model çiftleri arasındaki farklar Nemenyi post-hoc testi ile incelenmiştir. Elde edilen p-değerleri Tablo 4.8'de sunulmaktadır.

Tablo 4.8. Nemenyi post-hoc testi p-değerleri ($p < 0.05$ anlamlı farkı göstermektedir).

	Baseline	Lovasz	Aug	Aug+TTA	Swin
Baseline	1	0.065	0.002	2.46e-08	0.0648
Lovasz	0.065	1	0.831	0.0077	1.41e-06
Aug	0.002	0.831	1	0.151	2.7e-09
Aug+TTA	2.46e-08	0.0077	0.151	1	1.11e-16
Swin-UPerNet	0.0648	1.41e-06	2.7e-09	1.11e-16	1

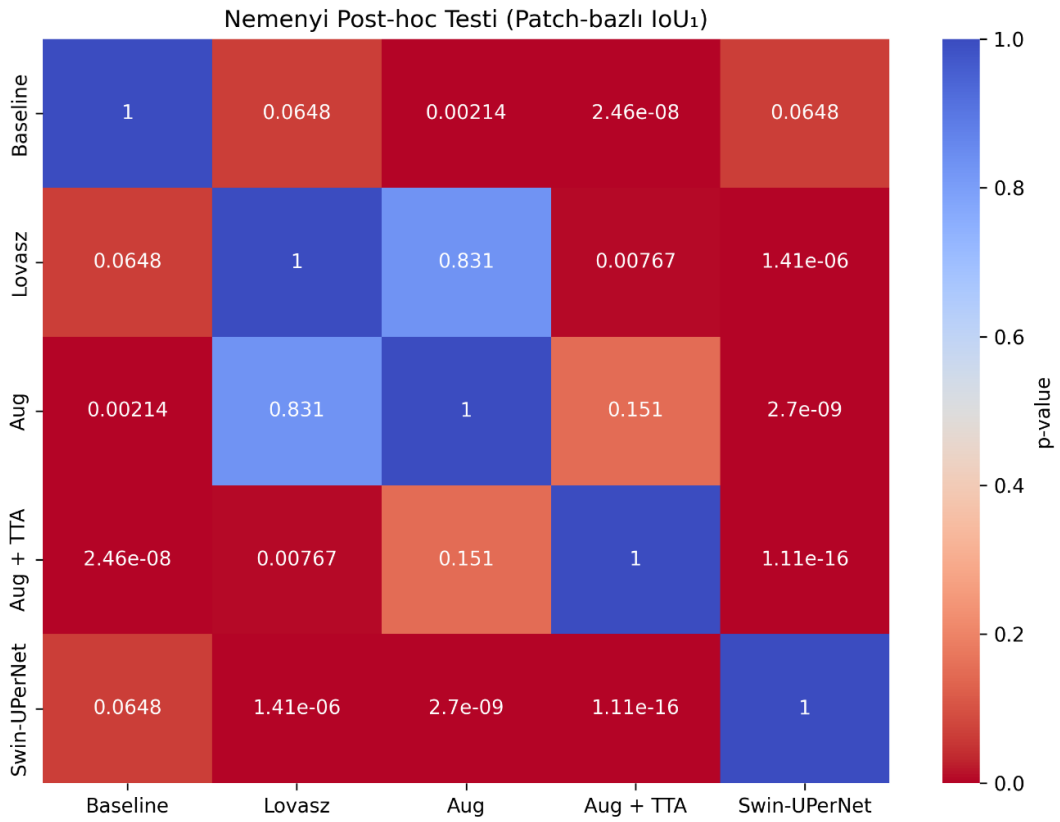
Nemenyi post-hoc testi sonuçlarına göre:

- Aug + TTA modeli, Baseline, Lovász ve Swin-UPerNet modellerine kıyasla istatistiksel olarak anlamlı derecede daha yüksek performans göstermektedir.
- Aug ve Aug + TTA modelleri arasında istatistiksel olarak anlamlı bir fark bulunmamıştır ($p \approx 0.151$). Bu durum test kümesi özelinde TTA'nın patch-bazlı IoU_1 skorlarına katkısının sınırlı kaldığını göstermektedir.
- Swin-UPerNet modeli SegFormer tabanlı modellere kıyasla daha düşük performans sergilemiş ve özellikle Aug ve Aug + TTA modellerinden anlamlı şekilde daha kötü sonuçlar üretmiştir.

Bu bulgular TTA'nın mutlak performans metriklerinde küçük iyileşmeler sağlayabilse de sınırlı sayıda hasar içeren patch ve yüksek varyans nedeniyle bu katkının istatistiksel anlamlılık eşiğini her durumda aşamayabileceğini göstermektedir.

Model çiftleri arasındaki istatistiksel ilişkilerin daha açık biçimde görülebilmesi amacıyla Nemenyi testi sonuçları ayrıca bir ısı haritası (heatmap) ile görselleştirilmiştir. Şekil 4.8'de sunulan bu görselde model çiftleri arasındaki p-değerleri renk ölçeği ile ifade edilmekte ve hangi model çiftleri arasında anlamlı fark bulunduğu görsel olarak özetlenmektedir.

Sonuç olarak gerçekleştirilen bu istatistiksel analiz, model karşılaştırmalarının yalnızca ortalama performans metriklerine dayandırılmadığını, aynı test örnekleri üzerinde gerçekleştirilen non-parametrik istatistiksel testlerle desteklendiğini göstermektedir. Nicel performans metrikleri ve yapı bazlı değerlendirmeler birlikte ele alındığında Aug + TTA modelinin genel olarak en güçlü performansı sunduğu, ancak bazı iyileştirmelerin istatistiksel olarak sınırlı etkiye sahip olduğu görülmektedir.



Şekil 4.8. Nemenyi post-hoc testi p-değerleri ısı haritası.

Sonuç olarak gerçekleştirilen bu istatistiksel analiz, model karşılaştırmalarının yalnızca ortalama performans metriklerine dayandırılmadığını, aynı test örnekleri üzerinde gerçekleştirilen non-parametrik istatistiksel testlerle desteklendiğini göstermektedir. Nicel performans metrikleri ve yapı bazlı değerlendirmeler birlikte ele alındığında, Aug + TTA modelinin genel olarak en güçlü performansı sunduğu, ancak bazı iyileştirmelerin istatistiksel olarak sınırlı etkiye sahip olduğu görülmektedir.

5. SONUÇ VE ÖNERİLER

5.1. Sonuçlar

Bu tez çalışması, Türkiye'nin yakın tarihindeki en yıkıcı doğal afetlerden biri olan 6 Şubat 2023 Kahramanmaraş depremleri sonrasında oluşan yapısal hasarların, yüksek çözünürlüklü uydu görüntüleri üzerinden otomatik olarak saptanması problemini metodolojik bir çerçevede ele almıştır. Araştırma boyunca elde edilen bulgular, afet sonrası acil müdahale süreçlerinde zaman ve veri kısıtı gibi operasyonel gerçeklikler göz önüne alındığında, afet öncesi görüntü zorunluluğunu ortadan kaldıran tek zamanlı analiz yaklaşımının stratejik değerini ortaya koymuştur. Bu kapsamda geliştirilen SegFormer tabanlı segmentasyon hattı, deprem enkazlarının düzensiz dokusu ve çevre unsurlarıyla olan benzerliğinden kaynaklanan zorlukları, transformer mimarisinin sunduğu küresel bağlam bilgisinden yararlanarak daha tutarlı biçimde modelleyebilmiştir. Deneyler sonucunda ulaşılan 0.7463 mIoU ve 0.6732 F1 metrikleri, önerilen yaklaşımın sadece akademik bir başarı değil, aynı zamanda sahada uygulanabilir bir hasar tespit kapasitesi sunduğunu sahaya yakın senaryolarda uygulanabilirliğini göstermektedir.

Çalışmanın metodolojik gelişim süreci incelendiğinde, en belirgin performans artışının tek bir teknik müdahaleden ziyade, eğitim ve çıkarım aşamalarının kademeli olarak optimize edilmesinden kaynaklandığı görülmektedir. Lovász-Softmax kayıp fonksiyonu ile optimizasyon hedefinin doğrudan IoU odaklı hale getirilmesi, hasar sınırlarında daha keskin ve tutarlı maskeler üretilmesini sağlamıştır. Eğitim aşamasında uygulanan zenginleştirilmiş veri artırma stratejileri modelin farklı görsel koşullara karşı genelleme yeteneğini güçlendirirken, çıkarım aşamasında devreye alınan Test-Zamanı Artırımı (TTA) mekanizması tahminlerdeki rastlantısal gürültüyü minimize ederek karar kararlılığını en üst seviyeye taşımıştır. Bu durum, model performansının sadece veri miktarına değil, aynı zamanda çıkarım stratejilerinin kalitesine de doğrudan bağlı olduğunu metodolojik bir şeffaflıkla göstermiştir.

Tezin en çarpıcı çıktılarından biri ise mimari karmaşıklık ile operasyonel verimlilik arasındaki dengeyi sorgulayan karşılaştırmalı analizlerde ortaya çıkmıştır. Parametre sayısı ve hesaplama maliyeti bakımından daha ağır bir yapı olan Swin-UPerNet hattının, optimize edilmiş SegFormer hattının gerisinde kalması; bu görev ve veri koşulları altında "daha büyük" bir modelin her zaman "daha iyi" sonuç vermeyebileceğine işaret etmektedir. SegFormer'ın MiT-B2 omurgası ile sunduğu hiyerarşik temsil ve hafif kod

çözücü tasarımı, hem daha dengeli bir hata profili sergilemiş hem de eğitim ve çıkarım maliyetleri açısından hızlı haritalama ihtiyacına daha uygun bir alternatif oluşturmuştur.
Sonuç olarak bu çalışma, KATE-PD veri seti üzerinde modern transformer tabanlı semantik segmentasyon yaklaşımlarının davranışını sistematik biçimde raporlamış; adil değerlendirme protokolü kapsamında eşik değerinin yalnızca doğrulama kümesi üzerinde belirlenmesiyle test sonuçlarının daha savunulabilir ve tekrarlanabilir bir temele oturtulmasına özen göstermiştir. Elde edilen bulgular, tek zamanlı post-event hasar segmentasyonu için uygulanabilir bir iş akışının mümkün olduğunu ve bu iş akışının gelecekte geliştirilecek otomatik hasar analiz sistemleri için operasyonel bir yol haritası sunduğunu göstermektedir.

5.2. Gelecek Çalışmalar İçin Öneriler

Bu tezde elde edilen bulgular, tek zamanlı optik uydu görüntülerinden hasar segmentasyonu yapılabildiğini gösterse de, yöntemin daha geniş ölçekte ve farklı operasyonel koşullarda güvenilirliğini artırmak için çeşitli araştırma yönleri öne çıkmaktadır. Öncelikle, bu çalışmada yalnızca optik veriler kullanıldığından, bulutluluk ve aydınlatma koşullarından daha az etkilenen Sentetik Açıklıklı Radar (SAR) verisinin Transformer tabanlı mimarilere entegre edilmesi, her hava koşulunda çalışabilen daha dayanıklı hasar haritalama sistemlerinin geliştirilmesine katkı sağlayabilir. Bunun yanında nitel analizlerde belirginleşen gölge, koyu zemin dokuları veya benzer tekstüre sahip yüzeylerle hasarın karışması gibi hata örüntülerinin daha sistematik biçimde karakterize edilmesi ve bu hataları hedefleyen eğitim stratejilerinin değerlendirilmesi, modelin ayırt ediciliğini güçlendirebilir.
Bir diğer önemli yön, geliştirilen hattın gerçek saha koşullarına daha yakın biçimde kullanılabilmesi için çıkarım maliyetinin düşürülmesidir. Özellikle test-zamanı artırımı gibi doğruluğu yükselten ancak hesaplama maliyetini artıran bileşenler yerine, daha hafif çıkarım stratejileri, model sıkıştırma yaklaşımları ya da düşük gecikmeli tasarımlar ile uç cihazlarda çalıştırılabilir çözümler geliştirilmesi, arama-kurtarma ve hızlı durum değerlendirme süreçlerine doğrudan katkı sunabilir. Son olarak, yöntemin farklı coğrafyalara ve farklı kentsel dokulara genellenebilirliğini ortaya koymak amacıyla çapraz veri seti deneyleri yapılması ve alan kayması (domain shift) etkilerini azaltmaya yönelik transfer öğrenme, alan uyarlama veya öz-denetimli ön eğitim gibi yaklaşımların

incelenmesi önerilmektedir. Bu doğrultuda, hem veri çeşitliliğini hem de modelin belirsizlik altında karar verme kapasitesini güçlendiren çalışmalar, deprem sonrası otomatik hasar analiz sistemlerinin daha güvenilir ve operasyonel hale gelmesine yardımcı olacaktır.

6. KAYNAKLAR

1. Bilham, R., *The seismic future of cities.* Bulletin of earthquake engineering, 2009. **7**(4): p. 839-887.

2. Afet, T. and A.D.Y. Başkanlığı, *Şubat 2023 Pazarcık-Elbistan Kahramanmaraş (mw: 7.7–mw: 7.6) Depremleri Raporu.* Deprem ve Risk Azaltma Genel Müdürlüğü Deprem Dairesi Başkanlığı. **2**.

3. U.S. Geological Survey (USGS). *M7.8 and M7.5 Kahramanmaraş Earthquake Sequence struck near Nurdağı, Turkey (Türkiye)*. 2023 28.12.2025]; Available from: https://www.usgs.gov/news/featured-story/m78-and-m75-kahramanmaras-earthquake-sequence-near-nurdagi-turkey-turkiye.

4. GOOGLE. *Google Earth*. 05.01.2026]; Available from: https://earth.google.com.

5. Erdik, M., *Urban Earthquake Risk.* 2006.

6. Altay, N. and W.G. Green III, *OR/MS research in disaster operations management.* European journal of operational research, 2006. **175**(1): p. 475-493.

7. Van Wassenhove, L.N., *Humanitarian aid logistics: supply chain management in high gear.* Journal of the Operational research Society, 2006. **57**(5): p. 475-489.

8. Gupta, R., et al. *Creating xBD: A dataset for assessing building damage from satellite imagery*. in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. 2019.

9. Joyce, K.E., et al., *A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters.* Progress in physical geography, 2009. **33**(2): p. 183-207.

10. Voigt, S., et al., *Global trends in satellite-based emergency mapping.* Science, 2016. **353**(6296): p. 247-252.

11. Wania, A., et al., *Increasing timeliness of satellite-based flood mapping using early warning systems in the Copernicus Emergency Management Service.* Remote Sensing, 2021. **13**(11): p. 2114.

12. Dell'Acqua, F. and P. Gamba, *Remote sensing and earthquake damage assessment: Experiences, limits, and perspectives.* Proceedings of the IEEE, 2012. **100**(10): p. 2876-2890.

13. Ji, M., L. Liu, and M. Buchroithner, *Identifying collapsed buildings using post-earthquake satellite imagery and convolutional neural networks: A case study of the 2010 Haiti earthquake.* Remote Sensing, 2018. **10**(11): p. 1689.

14. Bouchard, I., et al., *On transfer learning for building damage assessment from satellite imagery in emergency contexts.* Remote Sensing, 2022. **14**(11): p. 2532.

15. Yuan, Y., X. Chen, and J. Wang. *Object-contextual representations for semantic segmentation*. in *European conference on computer vision*. 2020. Springer.

16. Zhang, C., et al., *A domain adaptation neural network for change detection with heterogeneous optical and SAR remote sensing images.* International Journal of Applied Earth Observation and Geoinformation, 2022. **109**: p. 102769.

17. Berman, M., A.R. Triki, and M.B. Blaschko. *The lovász-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.

18. Buslaev, A., et al., *Albumentations: fast and flexible image augmentations.* Information, 2020. **11**(2): p. 125.

19. Xie, E., et al., *SegFormer: Simple and efficient design for semantic segmentation with transformers.* Advances in neural information processing systems, 2021. **34**: p. 12077-12090.

20. Musaoglu, N., et al., *A Rapid Damage Assessment Using Remote Sensing: Türkiye 2023 Post-Earthquake Dataset: KATE-PD*. 2025. 1454-1457.

21. Codeocean. *KATE-PD: Kahramanmaraş Türkiye Earthquake-Post Disaster Dataset*. 2025 20.12.2025]; Available from: https://codeocean.com/capsule/8064452/tree/v1.

22. Dong, L. and J. Shan, *A comprehensive review of earthquake-induced building damage detection with remote sensing techniques.* ISPRS Journal of Photogrammetry and Remote Sensing, 2013. **84**: p. 85-99.

23. Long, J., E. Shelhamer, and T. Darrell. *Fully convolutional networks for semantic segmentation*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.

24. Sun, S., et al., *Semantic segmentation for buildings of large intra-class variation in remote sensing images with O-GAN.* Remote Sensing, 2021. **13**(3): p. 475.

25. Chen, L.-C., et al., *Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs.* IEEE transactions on pattern analysis and machine intelligence, 2017. **40**(4): p. 834-848.

26. Ronneberger, O., P. Fischer, and T. Brox. *U-net: Convolutional networks for biomedical image segmentation*. in *International Conference on Medical image computing and computer-assisted intervention*. 2015. Springer.

27. Badrinarayanan, V., A. Kendall, and R. Cipolla, *Segnet: A deep convolutional encoder-decoder architecture for image segmentation.* IEEE transactions on pattern analysis and machine intelligence, 2017. **39**(12): p. 2481-2495.

28. Zhao, H., et al. *Pyramid scene parsing network*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

29. Chen, L.-C., et al. *Encoder-decoder with atrous separable convolution for semantic image segmentation*. in *Proceedings of the European conference on computer vision (ECCV)*. 2018.

30. Cheng, B., et al. *Boundary IoU: Improving object-centric image segmentation evaluation*. in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2021.

31. Vaswani, A., et al., *Attention is all you need.* Advances in neural information processing systems, 2017. **30**.

32. Dosovitskiy, A., *An image is worth 16x16 words: Transformers for image recognition at scale.* arXiv preprint arXiv:2010.11929, 2020.

33. Liu, Z., et al. *Swin transformer: Hierarchical vision transformer using shifted windows*. in *Proceedings of the IEEE/CVF international conference on computer vision*. 2021.

34. Touvron, H., et al. *Training data-efficient image transformers & distillation through attention*. in *International conference on machine learning*. 2021. PMLR.

35. Wang, W., et al. *Pyramid vision transformer: A versatile backbone for dense prediction without convolutions*. in *Proceedings of the IEEE/CVF international conference on computer vision*. 2021.

36. Strudel, R., et al. *Segmenter: Transformer for semantic segmentation*. in *Proceedings of the IEEE/CVF international conference on computer vision*. 2021.

37. Townshend, J., et al., *The impact of misregistration on change detection.* Geoscience and Remote Sensing, IEEE Transactions on, 1992. **30**: p. 1054-1060.

38. Yamazaki, F. and M. Matsuoka, *Remote sensing technologies in post-disaster damage assessment.* Journal of Earthquake and Tsunami, 2007. **1**(03): p. 193-210.

39. Gong, L., et al., *Earthquake-induced building damage detection with post-event sub-meter VHR TerraSAR-X staring spotlight imagery.* Remote Sensing, 2016. **8**(11): p. 887.

40. Weber, E. and H. Kané, *Building Disaster Damage Assessment in Satellite Imagery with Multi-Temporal Fusion*. 2020.

41. Zheng, Z., et al., *Building damage assessment for rapid disaster response with a deep object-based semantic change detection framework: From natural disasters to man-made disasters.* Remote Sensing of Environment, 2021. **265**: p. 112636.

42. Bai, T., et al., *Deep learning for change detection in remote sensing: a review.* Geo-spatial Information Science, 2023. **26**(3): p. 262-288.

43. Macchiarulo, V., et al., *Integrating post-event very high resolution SAR imagery and machine learning for building-level earthquake damage assessment.* Bulletin of Earthquake Engineering, 2024: p. 1-27.

44. Jing, W., et al., *ChangeRD: A registration-integrated change detection framework for unaligned remote sensing images.* ISPRS Journal of Photogrammetry and Remote Sensing, 2025. **220**: p. 64-74.

45. Ge, P., H. Gokon, and K. Meguro, *A review on synthetic aperture radar-based building damage assessment in disasters.* Remote Sensing of Environment, 2020. **240**: p. 111693.

46. Farin, D. and P.H. de With. *Misregistration errors in change detection algorithms and how to avoid them*. in *IEEE International Conference on Image Processing 2005*. 2005. IEEE.

47. Paszke, A., et al., *Pytorch: An imperative style, high-performance deep learning library.* Advances in neural information processing systems, 2019. **32**.

48. Scikit-learn developers. *Cross-validation: evaluating estimator performance*. [cited 02.01.2026; Available from: https://scikit-learn.org/stable/modules/cross_validation.html.

49. Hugging Face. *SegFormer*. 12.12.2025]; Available from: https://huggingface.co/docs/transformers/model_doc/segformer.

50. Kimura, M. *Understanding Test-Time Augmentation*. in *Neural Information Processing*. 2021. Cham: Springer International Publishing.

51. Powers, D., *Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation.* Mach. Learn. Technol., 2008. **2**.

52. Blaschke, T., *Object based image analysis for remote sensing.* ISPRS journal of photogrammetry and remote sensing, 2010. **65**(1): p. 2-16.

53. Radoux, J. and P. Bogaert, *Good practices for object-based accuracy assessment.* Remote Sensing, 2017. **9**(7): p. 646.

54. Gupta, R., et al., *xbd: A dataset for assessing building damage from satellite imagery.* arXiv preprint arXiv:1911.09296, 2019.

55. Asaf, M.Z., et al., *A Modified Deep Semantic Segmentation Model for Analysis of Whole Slide Skin Images.* Scientific Reports, 2024. **14**(1): p. 23489.

56. Jones, R., *Connected Filtering and Segmentation Using Component Trees.* Computer Vision and Image Understanding, 1999. **75**(3): p. 215-228.

57. Conover, W.J., *Practical nonparametric statistics*. 1999: john wiley & sons.

58. Friedman, M., *The use of ranks to avoid the assumption of normality implicit in the analysis of variance.* Journal of the american statistical association, 1937. **32**(200): p. 675-701.

59. Demšar, J., *Statistical comparisons of classifiers over multiple data sets.* Journal of Machine learning research, 2006. **7**(Jan): p. 1-30.

60. Nemenyi, P.B., *Distribution-free multiple comparisons*. 1963: Princeton University.