VİDEO KAYITLARINDAN DERİN ÖĞRENME İLE NESNE TESPİTİ VE GERÇEK DÜNYA KOORDİNATLARIYLA 3B DİJİTAL İKİZ GÖRSELLEŞTİRMESİ

YÜKSEK LİSANS TEZİ

Ahmet Sezgin AKTAŞ

Danışman
Prof. Dr. İbrahim YILMAZ

HARİTA MÜHENDİSLİĞİ ANABİLİM DALI

Temmuz 2026

AFYON KOCATEPE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

VİDEO KAYITLARINDAN DERİN ÖĞRENME İLE NESNE TESPİTİ VE GERÇEK DÜNYA KOORDİNATLARIYLA 3B DİJİTAL İKİZ GÖRSELLEŞTİRMESİ

Ahmet Sezgin AKTAŞ

Danışman
Prof. Dr. İbrahim YILMAZ

HARİTA MÜHENDİSLİĞİ ANABİLİM DALI

Temmuz 2026

TEZ ONAY SAYFASI

Ahmet Sezgin AKTAŞ tarafından hazırlanan "Video Kayıtlarından Derin Öğrenme ile Nesne Tespiti ve Gerçek Dünya Koordinatlarıyla 3B Dijital İkiz Görselleştirmesi" adlı tez çalışması lisansüstü eğitim ve öğretim yönetmeliğinin ilgili maddeleri uyarınca 07 / 07 / 2026 tarihinde aşağıdaki jüri tarafından **oy birliği** ile Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsü **Harita Mühendisliği Anabilim Dalı'nda YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Danışman **:** Prof. Dr. İbrahim YILMAZ

Başkan **:** Prof. Dr. İbrahim Öztuğ BİLDİRİCİ
Konya Teknik Üniversitesi, ... İmza .....
Mühendislik ve Doğa Bilimleri Fakültesi

Üye **:** Prof. Dr. İbrahim YILMAZ
Afyon Kocatepe Üniversitesi, ... İmza .....
Mühendislik Fakültesi

Üye **:** Prof. Dr. Murat UYSAL
Afyon Kocatepe Üniversitesi, ... İmza .....
Mühendislik Fakültesi

Afyon Kocatepe Üniversitesi
Fen Bilimleri Enstitüsü Yönetim Kurulu'nun
..... /..... /..... tarih ve
……. sayılı kararıyla onaylanmıştır.

…….
Prof. Dr. Bekir YALÇIN
Enstitü Müdürü

BİLİMSEL ETİK BİLDİRİM SAYFASI

Afyon Kocatepe Üniversitesi

Fen Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmasında;

- Tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- Görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- Başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- Atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- Kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- Ve bu tezin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

07 / 07 / 2026

İmza

Ahmet Sezgin AKTAŞ

ÖZET

Yüksek Lisans Tezi

VİDEO KAYITLARINDAN DERİN ÖĞRENME İLE NESNE TESPİTİ VE GERÇEK DÜNYA KOORDİNATLARIYLA 3B DİJİTAL İKİZ GÖRSELLEŞTİRMESİ

Ahmet Sezgin AKTAŞ
Afyon Kocatepe Üniversitesi
Fen Bilimleri Enstitüsü
Harita Mühendisliği Anabilim Dalı
Danışman: Prof. Dr. İbrahim YILMAZ

Bu tez çalışmasında, sabit kamera görüntülerinden tespit edilen hareketli nesnelerin coğrafi koordinatlarıyla ilişkilendirilerek 3B dijital ikiz ortamında görselleştirilmesine yönelik bütünleşik bir iş akışı geliştirilmiştir. Çalışma kapsamında Ankara ili Gölbaşı ilçesi Dikilitaş bölgesinde İHA fotogrametrisi ve LiDAR verileri kullanılarak 3B dijital ikiz altlığı oluşturulmuş, sabit kamera görüntüleri üzerinde YOLO tabanlı nesne tespiti uygulanmıştır. Tespit edilen nesnelerin piksel koordinatları, farklı koordinat dönüşüm yaklaşımlarıyla coğrafi koordinatlara dönüştürülmüş ve RTK ölçümleriyle doğruluk açısından değerlendirilmiştir. Elde edilen sonuçlar, nesne konumunun zemindeki temsilinde referans nokta seçiminin ve 3B model geometrisinin konumsal doğruluğu doğrudan etkilediğini göstermiştir. Ayrıca video tabanlı konum serilerindeki ani değişimleri azaltmak amacıyla Kalman filtreleme uygulanmış ve daha kararlı yörünge gösterimi elde edilmiştir. Sonuç olarak çalışma, derin öğrenme tabanlı nesne tespiti, İHA-LiDAR destekli 3B modelleme, piksel koordinatlarından coğrafi koordinatlara dönüşüm ve CesiumJS tabanlı web görselleştirmeyi tek bir uygulama akışı içinde birleştiren geliştirilebilir bir dijital ikiz yaklaşımı sunmaktadır.

2026, xi + 166 sayfa

Anahtar Kelimeler: Dijital İkiz, YOLO Tabanlı Nesne Tespiti, İHA Fotogrametrisi, LiDAR, Işın Kesişim Tabanlı Konum Kestirimi, Kalman Filtresi.

ABSTRACT

M.Sc. Thesis

OBJECT DETECTION FROM VIDEO FOOTAGE USING DEEP LEARNING AND 3D DIGITAL TWIN VISUALIZATION WITH REAL WORLD COORDINATES

Ahmet Sezgin AKTAŞ

Afyon Kocatepe University

Graduate School of Natural and Applied Sciences

Department of Geomatics Engineering

Supervisor: Prof. İbrahim YILMAZ

In this thesis, an integrated workflow was developed to associate moving objects detected from fixed camera images with geographic coordinates and visualize them in a 3D digital twin environment. Within the scope of the study, a 3D digital twin base was generated using UAV photogrammetry and LiDAR data in the Dikilitaş region of Gölbaşı district, Ankara, and YOLO-based object detection was applied to fixed camera images. The pixel coordinates of the detected objects were transformed into geographic coordinates using different coordinate transformation approaches, and their positional accuracy was evaluated with reference to RTK measurements. The results showed that the selection of the reference point representing the object position on the ground and the geometry of the 3D model directly affect positional accuracy. In addition, Kalman filtering was applied to reduce sudden changes in video-based position series, resulting in a more stable trajectory representation. As a result, this study presents an extendable digital twin approach that integrates deep learning-based object detection, UAV-LiDAR-supported 3D modeling, pixel-to-geographic coordinate transformation, and CesiumJS-based web visualization within a single application workflow.

2026, xi + 166 pages

Keywords: Digital Twin, YOLO-Based Object Detection, UAV Photogrammetry, LiDAR, Ray-Casting-Based Position Estimation, Kalman Filter.

TEŞEKKÜR

Bu tez çalışmasının konusunun belirlenmesi, deneysel çalışmaların yönlendirilmesi, elde edilen sonuçların değerlendirilmesi ve tez yazım sürecinin her aşamasında sağladığı değerli katkılarından dolayı danışmanım Sayın Prof. Dr. İbrahim YILMAZ'a en içten teşekkürlerimi sunarım. Çalışma kapsamında kullanılan LiDAR ve oblik fotogrametri verilerinin temin edilmesi sürecindeki katkılarından dolayı Harita Genel Müdürlüğü'ne teşekkür ederim. Ayrıca araştırma ve yazım süreci boyunca bilgi, öneri ve eleştirileriyle desteklerini esirgemeyen hocalarıma, meslektaşlarıma ve arkadaşlarıma teşekkür ederim.

Bu süreçte maddi ve manevi desteklerini her zaman yanımda hissettiğim aileme gönülden teşekkür ederim. Tez süreci boyunca uzun çalışma saatlerime, bitmeyen denemelerime, "az kaldı" cümlesinin defalarca uzamasına ve bilgisayar başında geçen sayısız geceye sabır gösteren sevgili eşime ayrıca teşekkür ederim. Onun desteği, anlayışı ve sabrı olmasaydı bu çalışmanın tamamlanması çok daha zor olurdu.

Ahmet Sezgin AKTAŞ
Afyonkarahisar 2026

İÇİNDEKİLER DİZİNİ

Sayfa

ÖZET i
ABSTRACT ii
TEŞEKKÜR iii
İÇİNDEKİLER DİZİNİ iv
SİMGELER ve KISALTMALAR DİZİNİ vii
ŞEKİLLER DİZİNİ ix
ÇİZELGELER DİZİNİ xi
1. GİRİŞ 1
1.1 Çalışmanın Amacı ve Kapsamı 2
1.2 Problemin Tanımı ve Özgün Değeri 3
2. LİTERATÜR BİLGİLERİ 6
2.1 İnsansız Hava Araçları ve Fotogrametrik Veri Üretimi 6
2.1.1 İHA Fotogrametrisinin Temel Esasları 7
2.1.2 Eğik (Oblique) Fotogrametri ve Akıllı Uçuş Planlaması 8
2.2 LiDAR Teknolojisi ve Nokta Bulutu İşleme 9
2.2.1 Havadan Lazer Tarama (ALS) Prensipleri 11
2.2.2 Nokta Bulutlarından 3B Model Üretimi 12
2.3 Derin Öğrenme ve Nesne Tespiti 13
2.3.1 Konvolüsyonel Sinir Ağları (CNN) ve Temel Yapısı 15
2.3.2 YOLO (You Only Look Once) Algoritmaları ve Mimarisi 18
2.3.2.1 YOLO Versiyon Mimarilerinin Karşılaştırılması 20
2.3.2.2 Transfer Öğrenme (Transfer Learning) Yaklaşımı 22
2.3.3 Nesne Tespitinde Performans Değerlendirme Ölçütleri 23
2.4 Görüntü Koordinatlarından Gerçek Dünya Koordinatlarına Dönüşüm 25
2.4.1 Kamera Geometrisi ve Homografi Yaklaşımı 26
2.4.2 Referans Noktalarına Dayalı Konum Kestirimi ve Enterpolasyon 28
2.4.3 Kamera Kalibrasyonu ve Ray-Casting Yaklaşımı 30
2.5 Video Tabanlı Konum Kestirimi ve Filtreleme Yaklaşımları 31
2.5.1 Konum Kestiriminde Gürültü ve Zamansal Kararlılık 32
2.5.2 Kalman Filtreleme Yaklaşımı 33
2.6 Dijital İkiz ve Web Tabanlı 3B Coğrafi Bilgi Sistemleri 35
2.6.1 Dijital İkiz Kavramı ve Kent Modelleri 36

2.6.2 Web Tabanlı CBS Platformları ve Görselleştirme Teknolojileri 38
2.6.3 WebGL Mimarisi ve CesiumJS Kütüphanesi 40
2.6.4 3D Tiles Veri Formatı ve Hiyerarşik Yapısı 41
3. MATERYAL ve METOT 43
3.1 Çalışma Alanı, Veri Setleri ve Donanım Bileşenleri 44
3.1.1 Çalışma Alanının Özellikleri 44
3.1.2 Kullanılan İHA, LiDAR Sensörleri ve Sabit Kameralar 46
3.1.3 Referans Veriler, Koordinat Sistemi ve RTK Ölçümleri 48
3.2 3B Dijital İkiz Altlığının Üretilmesi 49
3.2.1 Smart Oblique (Akıllı Eğik) Uçuş ile Veri Toplama 49
3.2.2 LiDAR Verisi, Veri Entegrasyonu ve 3D Tiles Üretimi 51
3.3 Derin Öğrenme ile Nesne Tespiti Uygulaması 53
3.3.1 Kullanılan YOLO Model Varyantları ve Model Seçim Yaklaşımı 54
3.3.2 Test Görüntüleri, Mesafe/Görünürlük Sınıfları ve Etiketleme Süreci 55
3.3.3 Çıkarım Parametreleri ve Model Çalıştırma Ayarları 56
3.3.4 Video Akışı Üzerinden Tespit ve Piksel Referans Noktalarının Çıkarılması 57
3.4 Piksel Koordinatlarından Gerçek Dünya Koordinatlarına Dönüşüm 59
3.4.1 Referans Noktaları ve Homografi Matrisi Hesabı 60
3.4.2 IDW Tabanlı Lokal Enterpolasyon Yöntemi ile Konum Kestirimi 61
3.4.3 CesiumJS Ortamında Ray-Casting Yaklaşımı 63
3.4.4 Konumsal Doğruluk Değerlendirme Yaklaşımı 65
3.5 Video Tabanlı Konum Yumuşatma ve Kalman Filtreleme 66
3.5.1 Ham Konum Serilerinde Aykırı Değer ve Hız Sıçraması Kontrolü 67
3.5.2 UTM Düzleminde Sabit Hız Modeline Dayalı Kalman Filtresi 67
3.5.3 Ham ve Filtreli Yörüngelerin Karşılaştırma Ölçütleri 68
3.6 CesiumJS Ortamında Entegrasyon ve Görselleştirme 69
3.6.1 Dijital İkizin Web Platformuna Aktarılması 70
3.6.2 Nesne Tespit Çıktılarının CesiumJS Ortamına Aktarılması İş Akışı 71
4. BULGULAR 72
4.1 3B Model ve Dijital İkiz Üretim Sonuçları 72
4.1.1 Fotogrametrik Model Kalitesi ve LiDAR Nokta Bulutu Yoğunluğu Analizi 74
4.2 Nesne Tespit Modeli Performans Analizleri 77

4.2.1 Model Varyantlarının Genel Sayım Tabanlı Tespit Başarımı 78
4.2.2 Mesafe ve Görünürlük Koşullarına Göre Tespit Başarımı 81
4.2.3 Etiketli BBox ve IoU Tabanlı Model Doğrulaması 83
4.2.4 Confidence Eşik Değerinin Tespit Başarımına Etkisi 85
4.2.5 Model Seçimi ve Dijital İkiz Uygulamasına Uygunluk Değerlendirmesi ... 87
4.3 Konum Doğruluğu Analizleri 88
4.3.1 Homografi, IDW Tabanlı Lokal Enterpolasyon ve Ray-Casting Tabanlı Koordinat Dönüşüm Yaklaşımları 90
4.3.2 BBox Referans Noktasının Konum Sonuçlarına Etkisi 93
4.3.3 RTK Referans Ölçümlerine Göre Yöntemlerin Konumsal Doğruluk Karşılaştırması 95
4.3.4 Mesafe, Görünürlük ve Fotoğraf Bazlı Hata Davranışının Değerlendirilmesi 100
4.4 Kalman Filtreleme ile Konum Yumuşatma ve Yörünge Kararlılığı Analizi 105
4.4.1 Ham Konum Kestirimlerinde Gürültü ve Hız Sıçraması Problemi 106
4.4.2 Kalman Filtreleme Uygulaması ve Parametre Seçimi 110
4.4.3 Ham ve Kalman Filtreli Sonuçların Karşılaştırılması 114
4.5 Bütünleşik Konum Kestirimi ve 3B Görselleştirme Sistem Mimarisi 119
5. TARTIŞMA ve SONUÇ 124
5.1 Bulguların Değerlendirilmesi 124
5.2 Çalışmanın Katkıları 126
5.3 Sınırlılıklar 127
5.4 Sonuçlar 129
5.5 Öneriler 130
6. KAYNAKLAR 133
ÖZGEÇMİŞ 139
EKLER 139

SİMGELER ve KISALTMALAR DİZİNİ

Simgeler

(u, v)	Görüntü düzlemindeki piksel koordinatları
(u_b, v_b)	BBox alt orta noktası koordinatları
(u_c, v_c)	BBox merkez noktası koordinatları
(x_{max}, y_{max})	BBox sağ alt köşe koordinatları
(x_{min}, y_{min})	BBox sol üst köşe koordinatları
(E, N)	UTM Easting ve Northing koordinatları
e_i	i. noktaya ait yatay konum hatası
(X, Y, Z)	Üç boyutlu gerçek dünya koordinatları
x_i	Kalman filtresi durum vektörü
z_i	Kalman filtresi ölçüm vektörü
δ_t	Ardışık iki kare arasındaki zaman farkı
ha	Hektar
F	Kalman filtresi durum geçiş matrisi
H	Homografi matrisi
K	Kalman kazancı
P	Kovaryans matrisi
Q	Süreç gürültüsü matrisi
R	Ölçüm gürültüsü matrisi
W	Homojen koordinat ölçek katsayısı
c	Işık hızı
cm	Santimetre
d	Sensör ile hedef yüzey arasındaki mesafe
km^2	Kilometrekare
km	Kilometre
km/h	Kilometre/saat
m	Metre
ms	Milisaniye
px	Piksel
s	Saniye
t	Lazer darbesinin gidiş-dönüş uçuş süresi

Kısaltmalar

2B	İki Boyutlu
3B	Üç Boyutlu
ALS	Airborne Laser Scanning (Havadan Lazer Tarama)
AP	Average Precision
BBox	Bounding Box (Sınırlayıcı Kutu)
BÖHHBÜY	Büyük Ölçekli Harita ve Harita Bilgileri Üretim Yönetmeliği
CBS	Coğrafi Bilgi Sistemleri
CNN	Convolutional Neural Network (Konvolüsyonel Sinir Ağı)
COCO	Common Objects in Context
DFL	Distribution Focal Loss
FN	False Negative (Yanlış Negatif)

Kısaltmalar (Devam)

FOV	Field of View (Görüş Alanı)
FP	False Positive (Yanlış Pozitif)
FPS	Frames Per Second (Saniyedeki Kare Sayısı)
glTF	GL Transmission Format
GNSS	Global Navigation Satellite System (Küresel Navigasyon Uydu Sistemi)
HGM	Harita Genel Müdürlüğü
HLOD	Hierarchical Level of Detail (Hiyerarşik Ayrıntı Düzeyi)
IDW	Inverse Distance Weighting
IMU	Inertial Measurement Unit (Ataletsel Ölçüm Birimi)
IoU	Intersection over Union
İHA	İnsansız Hava Aracı
LiDAR	Light Detection and Ranging (Işık Algılama ve Mesafe Ölçümü)
LOD	Level of Detail (Ayrıntı Düzeyi)
mAP	Mean Average Precision (Ortalama Hassasiyetlerin Ortalaması)
NMS	Non-Maximum Suppression
OGC	Open Geospatial Consortium
OpenCV	Open Source Computer Vision Library
ReLU	Rectified Linear Unit
RMSE	Root Mean Square Error (Karesel Ortalama Hata)
RTK	Real-Time Kinematic
SAM	Sayısal Arazi Modeli
SYM	Sayısal Yüzey Modeli
ToF	Time-of-Flight (Uçuş Süresi)
TP	True Positive (Doğru Pozitif)
UTM	Universal Transverse Mercator
WebGL	Web Graphics Library
WGS 84	World Geodetic System 1984
YKN	Yer Kontrol Noktası
YOLO	You Only Look Once
YÖA	Yer Örnekleme Aralığı (Ground Sampling Distance)
YZ	Yapay Zekâ

ŞEKİLLER DİZİNİ

Sayfa

Şekil 2.1 İHA tabanlı hava fotogrametrisi uygulama süreci. 7

Şekil 2.2 Geleneksel nadir çekim ve eğik fotogrametrik yaklaşımın 3B model üretimine etkisi (Verykokou ve Ioannidis 2024'ten uyarlanmıştır). 8

Şekil 2.3 Havadan Lazer Tarama (ALS) sisteminin temel çalışma prensibi (Ekercin ve Üstün 2004'ten uyarlanmıştır)..... 10

Şekil 2.4 LiDAR çoklu yansıma özelliği (White vd. 2013'ten uyarlanmıştır)..... 12

Şekil 2.5 Yapay zekâ, makine öğrenmesi ve derin öğrenme disiplinleri arasındaki temel sınırlar ile bu kavramların alt kümeler halindeki hiyerarşik ilişkisi (Goodfellow vd. 2016'dan uyarlanmıştır). 14

Şekil 2.6 Makine öğrenmesi ve derin öğrenme çalışma prensipleri (Goodfellow vd. 2016'dan uyarlanmıştır). 14

Şekil 2.7 CNN mimari örneği (Karatzinis vd. 2023'ten uyarlanmıştır). 16

Şekil 2.8 YOLO mimarisi. 18

Şekil 2.9 YOLO mimarisi ölçeklendirme karşılaştırması. 19

Şekil 2.10 COCO veri seti üzerinde YOLO modellerinin mAP^{50-95} karşılaştırması (İnt. Kyn. 1). 22

Şekil 2.11 Sınırlayıcı kutular arasındaki IoU hesaplama mantığı. 24

Şekil 2.12 Görüntü düzlemi ile yer düzlemi arasındaki homografi ilişkisi. 27

Şekil 2.13 Gerçek dünya verilerinin (sensörler, IoT) 3B kent modelleri ile entegrasyonunu gösteren dijital ikiz mimarisi (White vd. 2021'den uyarlanmıştır). 37

Şekil 2.14 CityGML standartlarına göre farklı ayrıntı düzeylerinin (LOD0 - LOD4) gösterimi (Gröger vd. 2012'den uyarlanmıştır). 38

Şekil 3.1 Çalışma alanının konumu ve sınırları (Google My Maps üzerinde oluşturulmuştur). 45

Şekil 3.2 Çalışma alanında gerçekleştirilen Smart Oblique uçuşu sonucunda üretilen SYM. 50

Şekil 3.3 Smart Oblique uçuş verilerinden üretilen çalışma alanının dijital ortomozaik görüntüsü 51

Şekil 3.4 LiDAR nokta bulutu verilerinden elde edilen renk kodlu SAM ve eşyükselti eğrileri. 52

Şekil 3.5 Önceden eğitilmiş YOLO modeli kullanılarak video akışı üzerinde gerçekleştirilen gerçek zamanlı nesne tespiti. 58

Şekil 3.6 Kamera projeksiyon merkezinden piksel noktasına uzanan matematiksel ışının 3B model ile kesişim (ray-casting) prensibi. 65

Şekil 3.7 CesiumJS ortamında tespit edilen nesneler ve ray-casting işleminin 3D Tiles altlığı üzerinde görselleştirilmesi. 72

Şekil 4.1 Üretilen 3B dijital ikiz altlığının genel görünümü. 74

Şekil 4.2 Zenmuse P1 kamerasına ait piksel bazlı geri çatım hata vektörü dağılımı. 75

Şekil 4.3 Zenmuse L2 sistemine entegre RGB kameraya ait piksel bazlı geri çatım hata vektörü dağılımı. 76

Şekil 4.4 YOLO model varyantlarının F1-skor ve görüntü başına ortalama çıkarım süresi ilişkisi. 80

Şekil 4.5 Mesafe durumuna göre model varyantlarının F1-skor değişimi. 82

Şekil 4.6 Görünürlük durumuna göre model varyantlarının F1-skor değişimi. 83

Şekil 4.7 Model varyantlarının F1-skor ve çıkarım süresi dengesi. 88

Şekil 4.8 Kamera aşağı bakış açısının konum hatası ve kullanılabilir kapsama oranı grafiği. 92

Şekil 4.9 BBox referans noktasına göre ortalama konum hatası karşılaştırması. 94

Şekil 4.10 RTK referans noktaları ile tahmini koordinatların UTM düzleminde karşılaştırılması. 96

Şekil 4.11 BBox alt orta noktası için yöntemlere göre ortalama hata ve RMSE karşılaştırması. 98

Şekil 4.12 BBox alt orta noktası için yöntemlere göre konum hatası dağılımı. 99

Şekil 4.13 Sahne grubuna göre ortalama konum hatası karşılaştırması. 101

Şekil 4.14 Görünürlük/örtülme durumuna göre ortalama konum hatası karşılaştırması. 102

Şekil 4.15 Fotoğraf bazlı konum hatalarının yöntemlere göre karşılaştırılması. 103

Şekil 4.16 Aykırı değer süzmesi ve Kalman filtreleme öncesi ham hız değişimi. 109

Şekil 4.17 Ham ve Kalman filtreli yörünge karşılaştırması. 115

Şekil 4.18 Konum zaman serisi karşılaştırması 116

Şekil 4.19 Ham ve Kalman filtreli hız değişimi 117

Şekil 4.20 Referans doğrultuya dik uzaklık değişimi. 118

Şekil 4.21 Gerçek zamanlı görüntü tabanlı konum kestirimi ve görselleştirme için önerilen sistem mimarisi. 120

Şekil 4.22 Araç, yaya, tır ve otobüs modellerinin CesiumJS ortamında 3B dijital ikiz altlığı üzerinde gösterimi. 122

Şekil 4.23 Geliştirilen görüntü tabanlı nesne konumlandırma ve 3B dijital ikiz görselleştirme sisteminin örnek çıktısı. 122

ÇİZELGELER DİZİNİ

Sayfa

Çizelge 3.1 Çalışmada kullanılan temel donanım ve veri toplama bileşenleri. 47

Çizelge 4.1 Fotogrametrik ve LiDAR uçuşlarına ait genel veri üretim parametreleri ve analiz sonuçları..... 73

Çizelge 4.2 YOLO model varyantlarının genel sayım tabanlı tespit başarımı..... 79

Çizelge 4.3 Etiketli BBox verisi üzerinde confidence ≥ 0,30 için model performansları. 84

Çizelge 4.4 Confidence ≥ 0,30 ve confidence ≥ 0,50 eşik değerlerinin genel performansa etkisi. 86

Çizelge 4.5 Video tabanlı ham konum kestirimi için kullanılan temel parametreler... 107

Çizelge 4.6 Kalman filtreleme modelinde kullanılan temel bileşenler ve parametreler 113

1. GİRİŞ

Günümüzde kentleşme oranının hızla artması, şehirlerin yönetimi, güvenliği ve sürdürülebilirliği açısından daha bütünleşik ve dinamik sistemlere duyulan ihtiyacı artırmaktadır. Geleneksel iki boyutlu (2B) haritalar ve statik veri setleri, kent yaşamındaki değişken süreçlerin izlenmesi ve mekânsal olarak analiz edilmesi açısından sınırlı kalmaktadır. Bu bağlamda dijital ikiz teknolojisi, fiziksel çevrenin sanal ortamda temsil edilmesini, izlenmesini ve farklı senaryolar altında değerlendirilmesini sağlayan önemli yaklaşımlardan biri hâline gelmiştir (Alan 2025). Dijital ikizler, fiziksel varlıkların, süreçlerin ve sistemlerin sanal ortamdaki dinamik temsilleri olarak tanımlanmakta ve gerçek dünya ile dijital ortam arasındaki etkileşimi güçlendirmeyi amaçlamaktadır (Tao vd. 2018). Bu yaklaşımın etkin biçimde kullanılabilmesi için yüksek doğruluklu üç boyutlu (3B) mekânsal verilerin üretilmesi, görselleştirilmesi ve anlık ya da zamana bağlı veri akışlarıyla desteklenmesi gerekmektedir.

İnsansız Hava Araçları (İHA) ve Işık Algılama ve Mesafe Ölçümü (Light Detection and Ranging - LiDAR) teknolojilerinde yaşanan gelişmeler, kent ölçeğindeki fiziksel çevrenin yüksek ayrıntı düzeyinde modellenmesine olanak tanımaktadır. İHA tabanlı fotogrametri ve LiDAR verilerinin birlikte kullanılması, karmaşık kentsel sahnelerin yoğun nokta bulutları, sayısal yüzey modelleri ve 3B modeller olarak yeniden oluşturulmasını sağlamaktadır (Vosselman ve Maas 2010, Colomina ve Molina 2014, Zhou vd. 2021). Bununla birlikte, yalnızca statik bir 3B modelin üretilmesi, dinamik şehir süreçlerinin izlenmesi için yeterli değildir. Bu modellerin video kameralar gibi sensörlerden elde edilen verilerle ilişkilendirilmesi, dijital ikiz yapısının güncellenebilir ve izlenebilir bir mekânsal bilgi ortamına dönüşmesini sağlamaktadır (Grieves 2014, White vd. 2021).

Bu noktada yapay zekâ ve derin öğrenme yaklaşımları, video görüntülerinden anlamlı bilgi çıkarımı yapılmasında önemli bir bileşen olarak öne çıkmaktadır (LeCun vd. 2015). You Only Look Once (YOLO) gibi modern nesne tespit algoritmaları, görüntülerdeki nesnelerin sınıflarını ve görüntü düzlemindeki konumlarını gerçek zamanlıya yakın hızlarda belirleyebilmektedir (Redmon vd. 2016, Wang vd. 2023).

Ancak görüntü düzleminde nesne konumlarının tek başına değerlendirilmesi, nesnelerin gerçek dünyadaki mekânsal karşılığını ortaya koymak için yeterli değildir. Bu nedenle tespit edilen nesnelerin piksel koordinatlarından gerçek dünya koordinatlarına dönüştürülmesi ve 3B dijital ikiz ortamında görselleştirilmesi, görüntü tabanlı tespit sonuçlarının coğrafi bağlamla ilişkilendirilmesi açısından önem taşımaktadır.

Bu tez çalışması, İHA ve LiDAR verileriyle üretilen 3B dijital ikiz altlığı üzerinde, sabit kamera görüntülerinden derin öğrenme tabanlı nesne tespiti yapılmasını, tespit edilen nesnelerin gerçek dünya koordinatlarıyla ilişkilendirilmesini ve bu sonuçların web tabanlı 3B ortamda görselleştirilmesini amaçlamaktadır. Böylece statik 3B kent modeli ile dinamik video analiz çıktıları bütünleştirilerek, görüntü tabanlı nesne tespit sonuçlarının mekânsal farkındalık sağlayacak biçimde sunulmasına yönelik bir yaklaşım ortaya konulmaktadır.

1.1 Çalışmanın Amacı ve Kapsamı

Bu tez çalışmasının temel amacı, İHA ve LiDAR verileri kullanılarak oluşturulan 3B dijital ikiz altlığı üzerinde, sabit kamera görüntülerinden derin öğrenme tabanlı nesne tespiti yapılması, tespit edilen nesnelerin görüntü düzlemindeki konumlarının gerçek dünya koordinatlarıyla ilişkilendirilmesi ve elde edilen sonuçların web tabanlı 3B ortamda görselleştirilmesidir. Bu kapsamda çalışma, video görüntülerinden elde edilen nesne tespit çıktılarının yalnızca görüntü düzleminde değerlendirilmesi yerine, coğrafi koordinat sistemiyle ilişkilendirilerek 3B dijital ikiz ortamında mekânsal farkındalık sağlayacak biçimde sunulmasına odaklanmaktadır.

Çalışmada, çalışma alanına ait 3B dijital ikiz altlığının üretilmesi, sabit kamera görüntülerinden nesne tespiti yapılması, tespit edilen nesnelerin piksel koordinatlarından gerçek dünya koordinatlarına dönüştürülmesi, farklı konum kestirim yaklaşımlarının doğruluk açısından değerlendirilmesi ve elde edilen sonuçların CesiumJS tabanlı web ortamında görselleştirilmesi hedeflenmiştir. Ayrıca, nesne tespit modellerinin çalışma sahasına uygunluğu ile video tabanlı konum kestirimlerinin zamansal kararlılığı da çalışma kapsamında incelenmiştir.

Çalışmanın kapsamı aşağıdaki temel bileşenlerden oluşmaktadır:

- 3B Dijital İkiz Altlığının Üretilmesi: Çalışma alanına ait fotogrametrik görüntülerin İHA ile toplanması, LiDAR verileriyle desteklenmesi ve web ortamında kullanılabilir 3B model altlığının oluşturulması.

- Nesne Tespiti: Sabit kamera görüntüleri üzerinde YOLO tabanlı nesne tespit modellerinin uygulanması ve tespit edilen nesnelere ait görüntü düzlemindeki konum bilgilerinin elde edilmesi.

- Koordinat Dönüşümü ve Konum Kestirimi: Görüntü üzerinde tespit edilen nesnelere ait piksel konumlarının, çalışma kapsamında değerlendirilen dönüşüm yaklaşımları kullanılarak gerçek dünya koordinatlarıyla ilişkilendirilmesi.

- Doğruluk ve Performans Değerlendirmesi: Nesne tespit başarımının ve koordinat dönüşüm sonuçlarının saha ölçümleri, referans noktaları ve hata analizleri doğrultusunda değerlendirilmesi.

- Konum Verilerinin İyileştirilmesi: Video tabanlı konum kestirimlerinde oluşabilecek gürültü ve ani değişimlerin azaltılması amacıyla filtreleme yaklaşımının değerlendirilmesi.

- Web Tabanlı 3B Görselleştirme: Üretilen 3B dijital ikiz altlığı ile nesne tespitinden elde edilen konum bilgilerinin CesiumJS ortamında bütünleşik olarak sunulması.

1.2 Problemin Tanımı ve Özgün Değeri

Geleneksel kamera tabanlı gözetleme sistemlerinde elde edilen veriler çoğunlukla iki boyutlu görüntü düzleminde değerlendirilmekte ve olayların gerçek dünya koordinatlarıyla doğrudan ilişkilendirilmesi sınırlı kalmaktadır. Bu durum, görüntüde tespit edilen nesnelerin coğrafi bağlamının belirlenmesini ve mekânsal analiz süreçlerine dâhil edilmesini güçleştirmektedir. Özellikle trafik izleme, güvenlik takibi,

afet yönetimi ve akıllı şehir uygulamaları gibi alanlarda, hareketli nesnelerin yalnızca görüntü üzerindeki konumlarının değil, gerçek dünyadaki mekânsal karşılıklarının da belirlenmesi önem taşımaktadır (Biljecki vd. 2015, White vd. 2021).

Görüntü koordinatlarından gerçek dünya koordinatlarına geçiş süreci; kamera geometrisi, referans noktalarının doğruluğu, görüntü çözünürlüğü, nesne-kamera mesafesi, görüş açısı, örtülme durumu ve kullanılan 3B modelin geometrik kalitesi gibi birçok etkene bağlıdır (Hartley ve Zisserman 2004). Bu nedenle, video görüntülerinden elde edilen nesne tespit sonuçlarının doğrudan coğrafi koordinatlarla ilişkilendirilmesi teknik açıdan önemli bir problemdir. Bunun yanında, yüksek hacimli 3B mekânsal verilerin web ortamında performanslı biçimde sunulması ve dinamik nesne konumlarıyla bütünleştirilmesi de ayrı bir teknik zorluk oluşturmaktadır (White vd. 2021, Zhong vd. 2024).

Bu problem, ulusal akıllı şehir politikaları açısından da güncel bir önem taşımaktadır. 2024–2030 Ulusal Akıllı Şehirler Stratejisi ve Eylem Planı'nda akıllı şehir uygulamalarının toplum, teknoloji, sürdürülebilirlik ve dirençlilik eksenlerinde geliştirilmesi hedeflenmektedir (İnt. Kyn. 10).

Bu kapsamda kamera görüntülerinden elde edilen dinamik nesne bilgilerinin gerçek dünya koordinatlarıyla ilişkilendirilmesi ve 3B dijital ikiz ortamında sunulması; veri tabanlı şehir yönetimi, trafik izleme, güvenlik takibi, afet ve acil durum yönetimi ile karar destek süreçleri açısından uygulanabilir bir yaklaşım sunmaktadır.

Bu tez çalışmasının özgün değeri, İHA ve LiDAR verileriyle oluşturulan 3B dijital ikiz altlığını, derin öğrenme tabanlı nesne tespiti, görüntü-koordinat dönüşümü, konumsal doğruluk analizi ve web tabanlı 3B görselleştirme süreçleriyle bütünleşik olarak ele almasıdır. Çalışma kapsamında, sabit kamera görüntülerinden elde edilen nesne tespit çıktıları yalnızca görüntü düzleminde değerlendirilmemiş; bu çıktıların gerçek dünya koordinatlarıyla ilişkilendirilerek 3B dijital ikiz ortamında sunulmasına yönelik bir iş akışı ortaya konulmuştur.

Çalışmanın literatüre ve uygulamaya sağladığı başlıca katkılar şu şekilde özetlenebilir:

- Hibrit 3B Veri Altlığının Kullanılması: Çalışma alanına ait 3B dijital ikiz altlığının oluşturulmasında fotogrametrik verilerin LiDAR verileriyle desteklenmesi, modelin geometrik temsil gücünü artırmakta ve konum kestirimi süreçleri için daha güvenilir bir mekânsal altlık sağlamaktadır.

- Görüntü Tabanlı Konum Kestirim Yaklaşımlarının Değerlendirilmesi: Sabit kamera görüntülerinde tespit edilen nesnelere ait piksel konumlarının gerçek dünya koordinatlarıyla ilişkilendirilmesi amacıyla farklı dönüşüm ve konum kestirim yaklaşımlarının değerlendirilmesi, görüntü tabanlı nesne tespit sonuçlarının coğrafi bağlamla ilişkilendirilmesine yönelik uygulanabilir bir yaklaşım sunmaktadır.

- Nesne Tespiti ve Konumsal Doğruluğun Birlikte İncelenmesi: Çalışmada yalnızca nesne tespit başarımı değil, tespit edilen nesnelerin gerçek dünya koordinatlarına aktarılması sonucunda oluşan konumsal doğruluk da değerlendirilmiştir. Bu yönüyle çalışma, görüntü tabanlı tespit performansı ile mekânsal doğruluk arasındaki ilişkiyi birlikte ele almaktadır.

- Web Tabanlı 3B Görselleştirme: Derin öğrenme tabanlı nesne tespitinden elde edilen konum bilgilerinin CesiumJS ortamında 3B dijital ikiz altlığı üzerinde sunulması, video analiz çıktılarının mekânsal farkındalık sağlayacak biçimde görselleştirilmesine katkı sağlamaktadır.

2. LİTERATÜR BİLGİLERİ

Bu bölümde, tez çalışmasının kuramsal altyapısını oluşturan temel kavramlar, yöntemler ve teknolojiler ele alınmıştır. İlk olarak İHA tabanlı fotogrametrik veri üretimi ve eğik fotogrametri yaklaşımları açıklanmış, ardından LiDAR teknolojisi ve nokta bulutu işleme süreçleri değerlendirilmiştir. Devamında, video görüntülerinden nesne tespiti amacıyla kullanılan derin öğrenme yaklaşımları, YOLO tabanlı nesne tespit algoritmaları ve nesne tespitinde kullanılan performans değerlendirme ölçütleri incelenmiştir. Daha sonra, görüntü koordinatlarından gerçek dünya koordinatlarına dönüşüm sürecinde kullanılan kamera geometrisi, homografi, enterpolasyon ve ray-casting olarak da bilinen ışın-kesişim yaklaşımına ilişkin kuramsal çerçeve sunulmuştur. Ayrıca, video tabanlı konum kestirimlerinde ortaya çıkabilecek gürültü ve zamansal kararlılık problemleri ile bu problemlerin azaltılmasında kullanılan filtreleme yaklaşımlarına değinilmiştir. Son olarak, dijital ikiz kavramı, web tabanlı 3B Coğrafi Bilgi Sistemleri (CBS), CesiumJS ve 3D Tiles veri yapısı açıklanarak çalışmanın görselleştirme altyapısına ilişkin literatür çerçevesi ortaya konulmuştur.

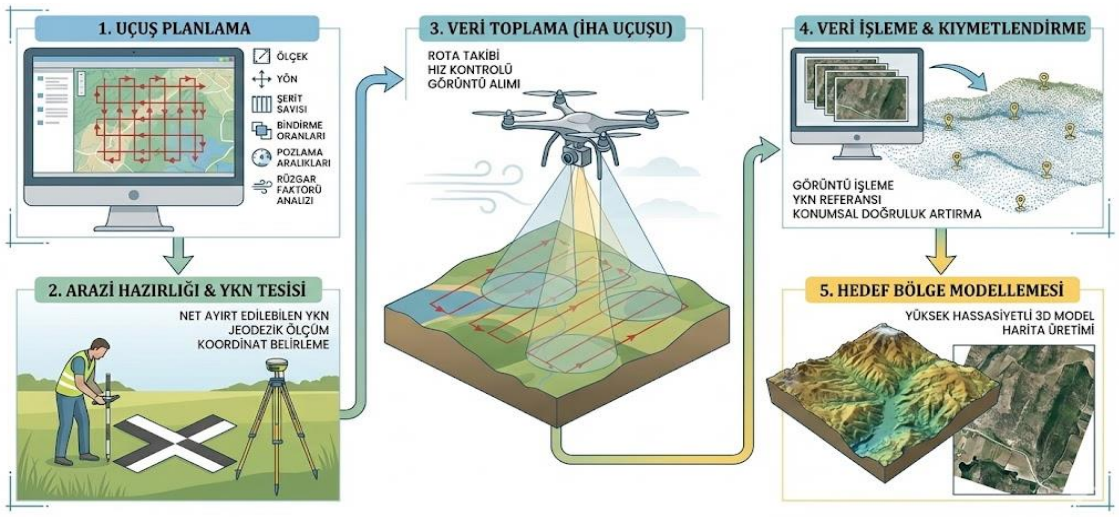
2.1 İnsansız Hava Araçları ve Fotogrametrik Veri Üretimi

İHA'lar, üzerinde insan pilot bulundurmadan, önceden tanımlanan uçuş planları doğrultusunda otonom olarak hareket edebilen veya uzaktan kumanda sistemiyle manuel olarak yönetilebilen hava platformlarıdır. Fotogrametrik veri üretimi açısından İHA'lar; dijital kamera, Küresel Navigasyon Uydu Sistemi (GNSS) ve Ataletsel Ölçüm Birimi (IMU) gibi bileşenlerle birlikte çalışarak yüksek çözünürlüklü görüntülerin hızlı, esnek ve ekonomik biçimde elde edilmesine olanak sağlamaktadır. Bu özellikleri nedeniyle İHA teknolojileri, haritalama, modelleme, izleme ve analiz çalışmalarında önemli bir veri toplama platformu hâline gelmiştir (Özemir ve Uzar 2016). Geleneksel hava fotogrametrisinde balon, uçurtma, helikopter ve insanlı uçak gibi platformlar kullanılırken, günümüzde İHA'lar düşük uçuş yüksekliği, yüksek mekânsal çözünürlük, operasyonel esneklik ve tekrarlı veri toplama imkânı sağlaması nedeniyle birçok uygulamada yaygın olarak tercih edilmektedir. İHA üzerinde taşınabilen kamera ve LiDAR gibi algılayıcılar sayesinde çalışma alanlarına ait ortofoto, Sayısal Yüzey

Modeli (SYM), nokta bulutu ve 3B model gibi fotogrametrik ürünler üretilebilmektedir (Bemis vd. 2014).

2.1.1 İHA Fotogrametrisinin Temel Esasları

İHA fotogrametrisi uygulama süreci genel olarak uçuş planlaması, arazi hazırlığı, veri toplama, görüntü işleme ve hedef bölgenin modellenmesi aşamalarından oluşmaktadır (Şekil 2.1). Uçuş planlaması aşamasında çalışma alanının büyüklüğü, istenen yer örnekleme aralığı, uçuş yüksekliği, uçuş yönü, şerit sayısı, bindirme oranları, pozlama aralıkları ve çevresel koşullar dikkate alınmaktadır. Bu parametreler, elde edilecek görüntülerin kalitesini ve fotogrametrik ürünlerin konumsal doğruluğunu doğrudan etkilemektedir (Colomina ve Molina 2014).



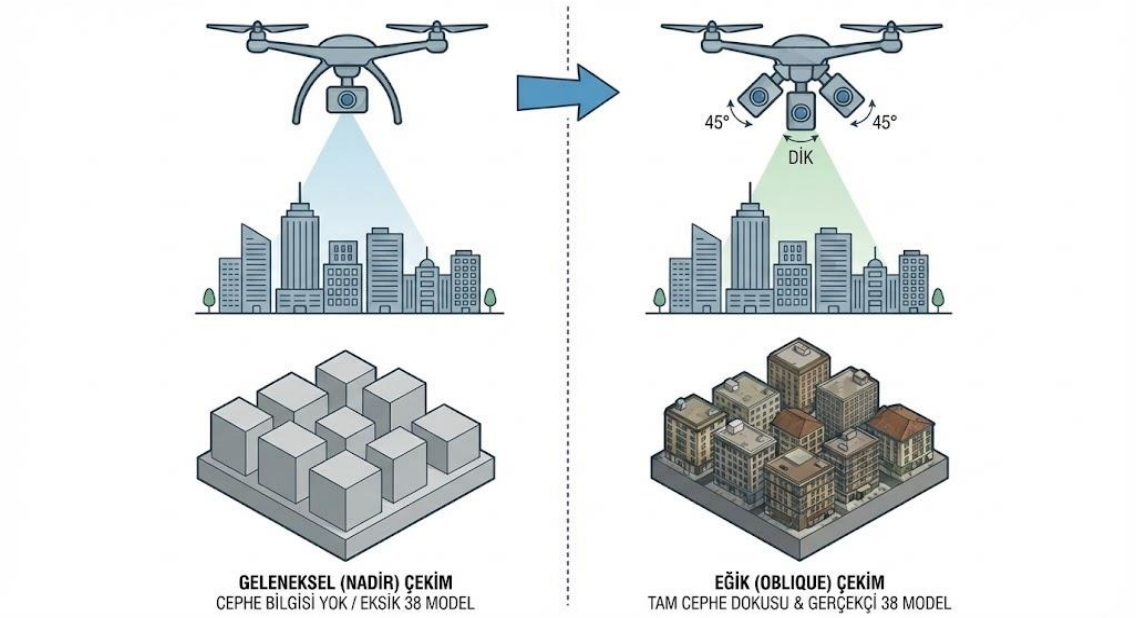
Şekil 2.1 İHA tabanlı hava fotogrametrisi uygulama süreci.

Fotogrametrik üretimde konumsal doğruluğu artırmak amacıyla çalışma alanında görüntüler üzerinde net biçimde seçilebilen Yer Kontrol Noktaları (YKN) tesis edilmektedir. GNSS tabanlı ölçme yöntemleriyle koordinatları belirlenen bu noktalar, görüntülerin dengelenmesi ve üretilen modelin gerçek dünya koordinat sistemiyle ilişkilendirilmesi sürecinde referans olarak kullanılmaktadır. Böylece İHA uçuşu ile elde edilen görüntüler, fotogrametrik işleme yazılımlarında değerlendirilerek ortofoto, SYM, nokta bulutu ve 3B model gibi fotogrametrik ürünlere dönüştürülmektedir (Bemis vd. 2014).

2.1.2 Eğik (Oblique) Fotogrametri ve Akıllı Uçuş Planlaması

Geleneksel hava fotogrametrisinde görüntüler çoğunlukla nadir doğrultuda, diğer bir ifadeyle yüzeye yaklaşık dik açıyla alınmaktadır. Bu yaklaşım çatı ve yatay yüzeylerin modellenmesinde etkili olmakla birlikte, bina cepheleri ve düşey yüzeylere ait ayrıntıların temsil edilmesinde sınırlı kalabilmektedir. Eğik fotogrametri ise nadir görüntülere ek olarak farklı eğim açılarıyla alınan görüntülerin kullanılmasına dayanmakta; böylece yapı cepheleri, düşey detaylar ve dokusal özellikler daha ayrıntılı biçimde modellenebilmektedir (Remondino ve Gerke 2015, Aicardi vd. 2016).

Eğik görüntüler, çoklu kamera sistemleriyle elde edilebildiği gibi, kamera yöneliminin uçuş sırasında değiştirilmesiyle de toplanabilmektedir. Bu yaklaşım, özellikle bina cephelerinin doku kaplamasında, yalnızca nadir çekimlerde ortaya çıkabilecek görüntülenemeyen alanların azaltılmasında ve daha gerçekçi 3B şehir modellerinin üretilmesinde önemli katkı sağlamaktadır (Şekil 2.2).



Şekil 2.2 Geleneksel nadir çekim ve eğik fotogrametrik yaklaşımın 3B model üretimine etkisi (Verykokou ve Ioannidis 2024'ten uyarlanmıştır).

Eğik fotogrametri ve 3B modelleme çalışmalarında sonuç kalitesi, uçuş planlamasıyla doğrudan ilişkilidir. Bu süreçte bindirme oranları, kamera açısı, uçuş yönü, yer örnekleme aralığı, uçuş yüksekliği ve çevresel koşullar birlikte değerlendirilmelidir.

Görüntü eşleştirme başarısının artırılması ve modelde boşlukların azaltılması için yüksek bindirme oranları tercih edilmekte; uygulamanın amacına ve çalışma alanının özelliklerine bağlı olarak yaklaşık %80 boyuna ve %60 enine bindirme değerleri yaygın olarak kullanılmaktadır (Avşar 2006, Maboudi vd. 2023).

Kamera açısı ve uçuş yönü de eğik fotogrametri uygulamalarında önemli parametreler arasındadır. Binaların farklı cephelerinin görüntülenebilmesi için uçuş güzergâhı çoğunlukla ızgara yapısında veya birbirini tamamlayan yönlerde planlanmaktadır. Smart Oblique (Akıllı Eğik) olarak adlandırılan akıllı eğik uçuş modları, kamera yönelimi ve çekim doğrultusunu hedef alanın geometrisine göre düzenleyerek veri toplama sürecinin daha etkin yürütülmesine katkı sağlamaktadır.

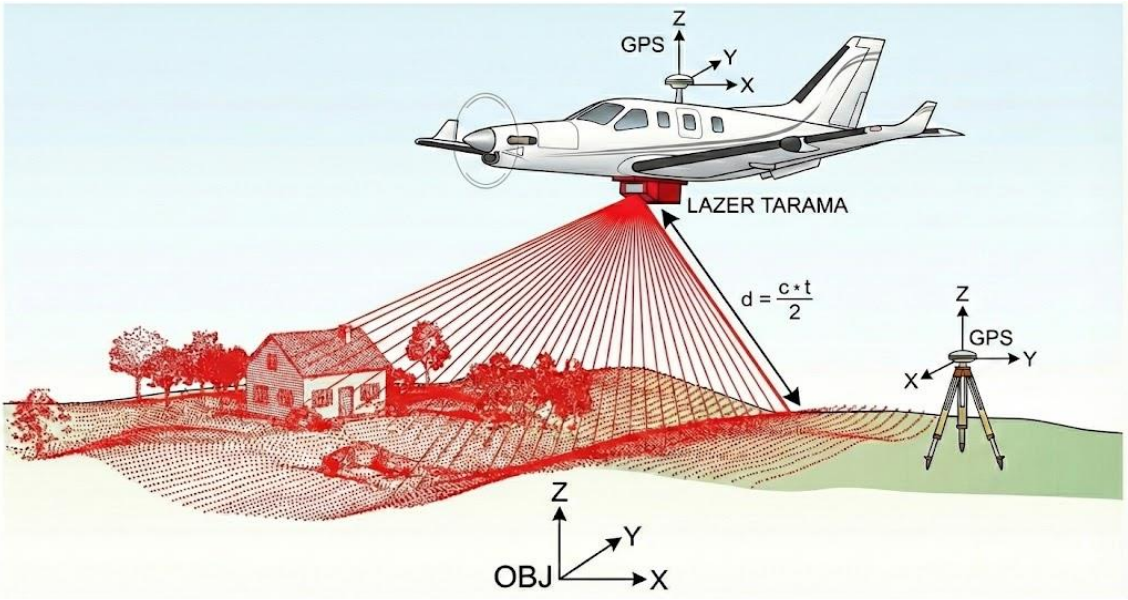
Yer Örnekleme Aralığı (YÖA) ve uçuş yüksekliği, üretilecek modelin ayrıntı düzeyini etkileyen temel faktörlerdendir. Daha düşük uçuş yüksekliği genellikle daha küçük YÖA değeri sağlayarak daha yüksek mekânsal çözünürlük sunarken, daha yüksek uçuşlar geniş alanların daha kısa sürede görüntülenmesine olanak tanımaktadır. Bu nedenle uçuş yüksekliği, çalışmanın amacı, hedeflenen Level of Detail (LOD), çalışma alanının büyüklüğü ve istenen model doğruluğu dikkate alınarak belirlenmelidir (Aicardi vd. 2016, Maboudi vd. 2023).

Eğik fotogrametriyle elde edilen veriler, geometrik model üretiminin yanında yapı cephelerine ait doku bilgisini de içerdiğinden; akıllı şehir uygulamaları, kentsel dönüşüm, afet yönetimi, kültürel miras belgeleme ve 3B kadastro gibi alanlarda görsel ve metrik analizler için güçlü bir altlık sağlamaktadır (Remondino ve Gerke 2015, Aicardi vd. 2016).

2.2 LiDAR Teknolojisi ve Nokta Bulutu İşleme

LiDAR, lazer ışınları kullanarak nesnelerin veya yüzeylerin konumunu, uzaklığını ve yüzey özelliklerini yüksek doğrulukla belirlemeye olanak tanıyan aktif bir uzaktan algılama teknolojisidir. Elektromanyetik spektrumun görünür veya kızılötesi bölgesinde çalışan bu sistemler, hedef yüzeye gönderilen lazer darbelerinin geri dönüş süresini

ölçerek sensör ile hedef yüzey arasındaki mesafeyi hesaplamaktadır (Wehr ve Lohr 1999, Vosselman ve Maas 2010). Bu ölçüm prensibi, Time-of-Flight (ToF) olarak adlandırılan uçuş süresi yaklaşımına dayanmaktadır. Lazer darbesinin hedef yüzeye ulaşıp sensöre geri dönmesi için geçen süre, ışık hızı ile birlikte değerlendirilerek mesafe bilgisi elde edilmektedir (Şekil 2.3).



Şekil 2.3 Havadan Lazer Tarama (ALS) sisteminin temel çalışma prensibi (Ekercin ve Üstün 2004'ten uyarlanmıştır).

Mesafe hesaplama aşağıdaki eşitlik ile ifade edilmektedir:

$$d = \frac{c \times t}{2} \tag{2.1}$$

Eşitlikte d, sensör ile hedef yüzey arasındaki mesafeyi; c, ışık hızını; t ise lazer darbesinin gidiş-dönüş uçuş süresini ifade etmektedir. Eşitlikte ikiye bölme işlemi, ölçülen sürenin lazer darbesinin hedefe ulaşması ve geri dönmesini kapsayan toplam süre olmasından kaynaklanmaktadır. Bu prensip sayesinde elde edilen mesafe ölçümleri, çalışma alanının 3B olarak modellenmesinde kullanılan nokta bulutu verilerinin temelini oluşturmaktadır (Roriz vd. 2022).

LiDAR sistemleri, yüksek nokta yoğunluğu ve konumsal doğruluk sağlaması nedeniyle

arazi modelleme, ormancılık, kentsel planlama ve afet yönetimi gibi birçok alanda önemli bir veri kaynağı olarak kullanılmaktadır (Makineci 2016). Fotogrametrik yöntemlerden farklı olarak aktif algılama prensibine dayanması, LiDAR sistemlerinin dış aydınlatma koşullarından daha az etkilenerek gece veya gündüz veri toplayabilmesine olanak tanımaktadır (Yastıklı 2003). Elde edilen veriler; X, Y, Z koordinatlarının yanı sıra yansıma şiddeti (intensity), dönüş sayısı ve sınıflandırma bilgisi gibi öznitelikleri içeren nokta bulutları biçiminde depolanmaktadır (Vosselman ve Maas 2010).

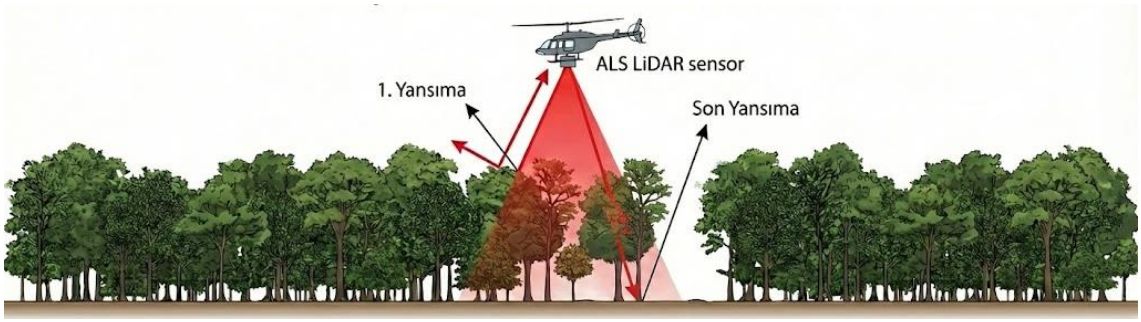
2.2.1 Havadan Lazer Tarama (ALS) Prensipleri

Havadan lazer tarama, uçak, helikopter ya da İHA gibi hava platformlarına entegre edilen LiDAR sensörleri aracılığıyla geniş alanların hızlı ve yüksek doğrulukla haritalanmasını sağlayan bir ölçme yöntemidir. ALS sistemleri, lazer tarayıcı, GNSS ve IMU bileşenlerinin birlikte çalışmasına dayanır. Lazer tarayıcı yüzeye lazer darbeleri göndererek mesafe bilgisini üretirken, GNSS platformun anlık konumunu, IMU ise platformun uçuş sırasındaki yönelim ve hareket bilgilerini belirlemektedir. Bu bileşenlerden elde edilen verilerin bütünleştirilmesiyle lazer darbelerinin yer yüzeyindeki konumu hesaplanmakta ve çalışma alanına ait 3B nokta bulutu oluşturulmaktadır (Wehr ve Lohr 1999).

ALS sistemlerinin önemli özelliklerinden biri çoklu yansıma kabiliyetidir. Gönderilen bir lazer darbesi, geçirgen veya karmaşık yüzeylerden birden fazla yansıma üretebilir. Bu durumda sistem, bitki örtüsünün üst kısımlarından gelen ilk yansımaları ve zemine daha yakın yüzeylerden gelen son yansımaları ayrı ayrı kaydedebilmektedir (Şekil 2.4).

Bu özellik, özellikle yoğun bitki örtüsünün bulunduğu alanlarda çıplak zemin yüzeyinin belirlenmesine ve Sayısal Arazi Modeli (SAM) üretimine katkı sağlamaktadır (Kraus 2007). Çoklu yansıma özelliği aynı zamanda arazi, bitki örtüsü ve yapı yüzeylerinin ayrıştırılmasına yardımcı olmakta; LiDAR verilerinin fotogrametrik verilerle birlikte kullanılması, 3B model üretiminde geometrik temsilin güçlendirilmesine katkı sağlamaktadır (Vosselman ve Maas 2010, Zhou vd. 2021).

Modern ALS sistemleri, farklı tarama desenleri, çoklu darbe teknolojisi ve tam dalga formu analizi gibi yaklaşımlar sayesinde yüksek nokta yoğunluğuna sahip veri üretebilmektedir. Bu özellikleriyle ALS verileri; arazi modelleme, ormancılık, kentsel modelleme, afet yönetimi ve 3B dijital ikiz üretimi gibi uygulamalarda önemli bir mekânsal veri kaynağı olarak kullanılmaktadır (Vosselman ve Maas 2010).



Şekil 2.4 LiDAR çoklu yansıma özelliği (White vd. 2013'ten uyarlanmıştır).

2.2.2 Nokta Bulutlarından 3B Model Üretimi

LiDAR sensörlerinden elde edilen ham nokta bulutları, yoğun, düzensiz dağılıma sahip ve yer yer gürültü içerebilen 3B veri kümeleridir. Bu verilerin görselleştirme ve analiz amacıyla kullanılabilir 3B modellere dönüştürülebilmesi için belirli işleme adımlarından geçirilmesi gerekmektedir. Bu süreç genel olarak ön işleme, zemin ve zemin dışı noktaların ayrıştırılması, sınıflandırma, yüzey modelleme ve veri optimizasyonu aşamalarından oluşmaktadır.

Ön işleme aşamasında, ham nokta bulutunda bulunan hatalı ölçümlerin, aykırı değerlerin ve sistem kaynaklı gürültülerin azaltılması amaçlanır. Ardından zemin ve zemin dışı noktaların ayrıştırılmasıyla arazi yüzeyi, yapı yüzeyleri, bitki örtüsü ve diğer nesneler daha anlamlı biçimde analiz edilebilir hâle getirilmektedir. Bu işlem, SAM, SYM, nokta bulutu sınıflandırması ve 3B model üretimi için temel bir aşamadır (Zhang vd. 2016).

Sınıflandırma ve segmentasyon aşamasında nokta bulutları; arazi, bina, yol, bitki örtüsü ve benzeri sınıflara ayrılabilmektedir. Bu süreçte noktaların yükseklik, komşuluk

ilişkisi, yüzey eğimi, düzlemsellik ve yansıma şiddeti gibi geometrik ve öznitelik bilgileri kullanılmaktadır. Sınıflandırılmış nokta bulutları, çalışma alanındaki fiziksel çevrenin daha düzenli biçimde temsil edilmesini sağlamakta ve sonraki modelleme aşamalarına altlık oluşturmaktadır (Vosselman ve Maas 2010).

3B model üretimi aşamasında, nokta bulutlarından yüzeyler oluşturularak üçgen ağ model, SYM veya yoğun nokta bulutu tabanlı görselleştirme ürünleri elde edilebilmektedir. Bu ürünler, çalışma amacına bağlı olarak fotogrametrik verilerle birlikte değerlendirilerek daha bütüncül bir 3B temsil sağlayabilmektedir. LiDAR verilerinin sunduğu konumsal doğruluk ve yüzey bilgisi, fotogrametrik verilerin doku ve görsel zenginliğiyle birlikte kullanıldığında dijital ikiz üretimi için güçlü bir mekânsal altlık oluşturmaktadır (Zhou vd. 2021, White vd. 2021).

Elde edilen sonuç ürünler, CBS yazılımlarında analiz edilebilmekte veya CesiumJS gibi web tabanlı 3B platformlarda kullanılmak üzere optimize edilebilmektedir. Bu kapsamda nokta bulutlarının ve 3B model verilerinin uygun formatlara dönüştürülmesi, veri boyutunun yönetilmesi ve web ortamında performanslı biçimde sunulması, dijital ikiz uygulamalarının kullanılabilirliği açısından önemli bir aşamadır.

2.3 Derin Öğrenme ve Nesne Tespiti

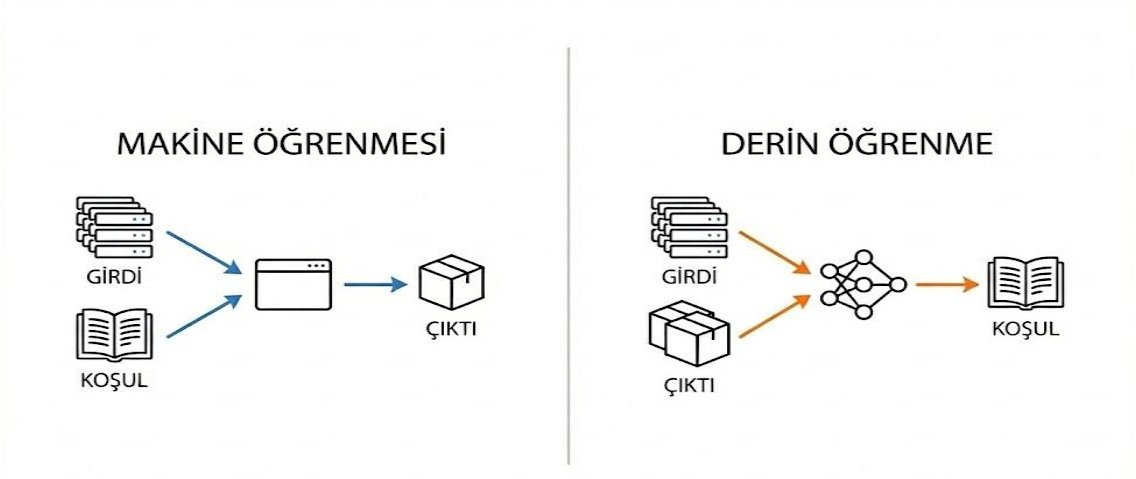
Yapay zekâ (YZ), insan benzeri karar verme, öğrenme ve problem çözme süreçlerinin bilgisayar sistemleri aracılığıyla modellenmesini amaçlayan geniş bir araştırma alanıdır. Makine öğrenmesi, YZ'nin bir alt alanı olarak verilerden örüntü öğrenen ve bu örüntüler üzerinden tahmin veya sınıflandırma yapabilen algoritmaları kapsamaktadır.

Derin öğrenme ise makine öğrenmesi içinde yer alan, çok katmanlı yapay sinir ağları aracılığıyla karmaşık veri temsillerini öğrenebilen bir yaklaşımdır. Özellikle görüntü, ses ve metin gibi yüksek boyutlu verilerde özniteliklerin otomatik olarak öğrenilebilmesi, derin öğrenmenin bilgisayarlı görü ve nesne tespiti uygulamalarında yaygın biçimde kullanılmasını sağlamaktadır (LeCun vd. 2015). YZ, makine öğrenmesi ve derin öğrenme arasındaki kavramsal hiyerarşi Şekil 2.5'te gösterilmiştir.



Şekil 2.5 Yapay zekâ, makine öğrenmesi ve derin öğrenme disiplinleri arasındaki temel sınırlar ile bu kavramların alt kümeler halindeki hiyerarşik ilişkisi (Goodfellow vd. 2016'dan uyarlanmıştır).

Geleneksel makine öğrenmesi yöntemlerinde sınıflandırma başarımı çoğunlukla veriden uygun özniteliklerin çıkarılmasına bağlıdır. Bu süreçte öznitelik çıkarımı çoğu zaman kullanıcı tarafından tanımlanan veya uzman bilgisine dayanan yöntemlerle gerçekleştirilmektedir. Buna karşılık derin öğrenme yöntemleri, özellikle görüntü, ses ve metin gibi yüksek boyutlu verilerden öznitelikleri otomatik olarak öğrenebilme yeteneğine sahiptir. Bu yönüyle derin öğrenme, bilgisayarlı görü ve nesne tespiti gibi alanlarda yaygın biçimde kullanılmaktadır (LeCun vd. 2015, Gündüz 2023). Makine öğrenmesi ile derin öğrenme yaklaşımlarının genel çalışma prensipleri Şekil 2.6'da verilmiştir.



Şekil 2.6 Makine öğrenmesi ve derin öğrenme çalışma prensipleri (Goodfellow vd. 2016'dan uyarlanmıştır).

Derin öğrenme modelleri, çok katmanlı mimarileri sayesinde veriyi hiyerarşik bir biçimde temsil etmektedir. Görüntü işleme uygulamalarında ilk katmanlar kenar, köşe ve doku gibi düşük düzeyli özellikleri öğrenirken, daha derin katmanlar nesne parçaları ve nesne bütününe ilişkin daha soyut temsiller oluşturabilmektedir. Bu özellik, derin öğrenme yaklaşımlarını video görüntülerinden araç, yaya ve benzeri nesnelerin tespit edilmesi gibi karmaşık görsel tanıma problemleri için uygun hâle getirmektedir (Gündüz 2023).

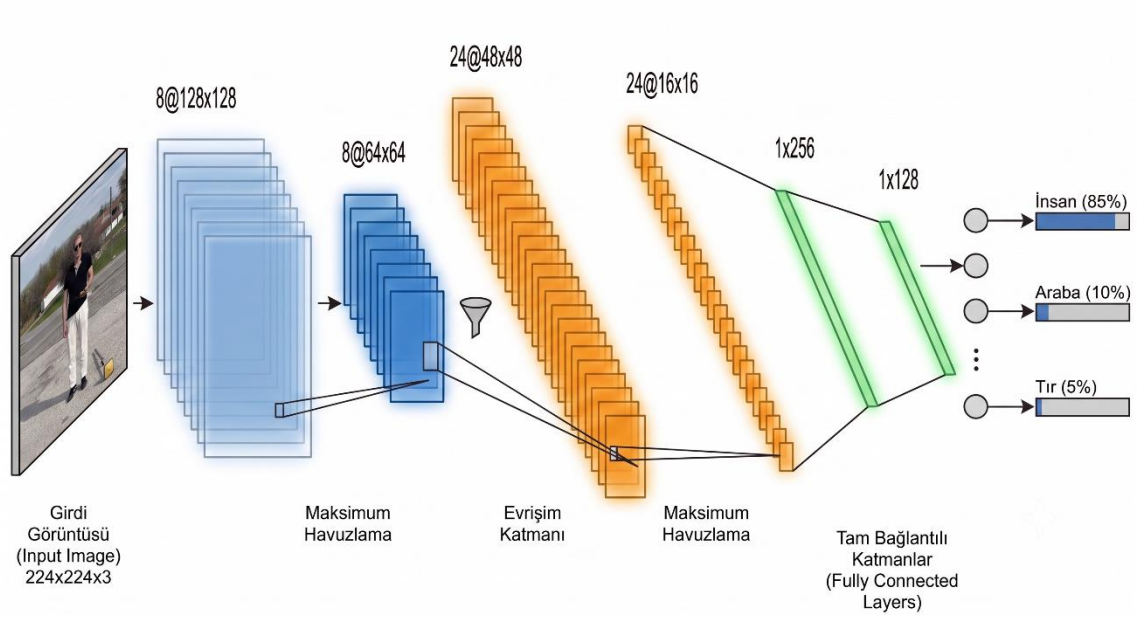
2.3.1 Konvolüsyonel Sinir Ağları (CNN) ve Temel Yapısı

Bilgisayarlı görü ve nesne tespiti problemlerinde yaygın olarak kullanılan derin öğrenme mimarilerinden biri Konvolüsyonel Sinir Ağlarıdır. Literatürde Convolutional Neural Networks (CNN) olarak adlandırılan bu yapılar, özellikle görüntü gibi ızgara düzenine sahip verilerin işlenmesi için geliştirilmiştir. CNN mimarileri, görsel verilerdeki kenar, köşe, doku, renk geçişi ve nesne parçaları gibi örüntüleri katmanlar boyunca hiyerarşik biçimde öğrenebilmesi nedeniyle görüntü sınıflandırma, nesne tanıma ve nesne tespiti uygulamalarında etkili sonuçlar vermektedir (LeCun vd. 1998, LeCun vd. 2015).

CNN'ler, temel çalışma prensibi bakımından yapay sinir ağlarına benzemekle birlikte, görüntü verilerindeki mekânsal ilişkileri koruyarak öğrenme yapabilmeleri açısından klasik tam bağlantılı ağlardan ayrılmaktadır. Geleneksel yöntemlerde görüntüye ait özniteliklerin çoğu kullanıcı tarafından tanımlanırken, CNN mimarilerinde bu öznitelikler eğitim sürecinde model tarafından otomatik olarak öğrenilmektedir. Bu durum, özellikle büyük veri kümelerinde ve karmaşık görsel örüntülerin bulunduğu problemlerde CNN tabanlı yaklaşımların güçlü sonuçlar üretmesine katkı sağlamaktadır (O'Shea 2015).

Şekil 2.7'de gösterildiği üzere, tipik bir CNN mimarisi evrişim katmanları, aktivasyon fonksiyonları, havuzlama katmanları ve tam bağlantılı katmanlardan oluşmaktadır. Evrişim katmanları, giriş görüntüsü üzerinde filtreler kullanarak öznitelik haritaları üretir. Bu öznitelik haritaları, görüntü içerisindeki yerel örüntülerin farklı soyutlama

düzeylerinde temsil edilmesini sağlar. Aktivasyon fonksiyonları, ağın doğrusal olmayan ilişkileri öğrenebilmesine olanak tanırken; havuzlama katmanları, öznitelik haritalarının boyutunu azaltarak hesaplama yükünü düşürür ve modelin daha kararlı temsiller üretmesine katkı sağlar. Tam bağlantılı katmanlar ise önceki katmanlardan elde edilen öznitelikleri bütünleştirerek sınıflandırma veya tahmin sonucunu üretir.



Şekil 2.7 CNN mimari örneği (Karatzinis vd. 2023'ten uyarlanmıştır).

CNN mimarisinde en temel işlem evrişim işlemidir. Evrişim, bir çekirdek ya da filtre matrisinin giriş görüntüsü üzerinde kaydırılması ve her konumda ilgili piksel değerleri ile filtre ağırlıklarının çarpılıp toplanması esasına dayanmaktadır. Bu işlem sonucunda görüntüdeki sınır, doku, ton değişimi ve şekil gibi yerel özellikleri temsil eden öznitelik haritaları elde edilmektedir. Giriş matrisi (X) ve filtre matrisi (W) olmak üzere iki boyutlu evrişim işlemi Eşitlik (2.2) ile ifade edilebilir.

$$(X * W)[i,j] = \sum_{m} \sum_{n} X[i+m, j+n] \cdot W[m,n] \tag{2.2}$$

Eşitlikte $X[i + m, j + n]$, giriş matrisindeki ilgili değeri; $W[m,n]$, filtre matrisindeki ağırlık katsayısını; i ve j, çıktı matrisindeki satır ve sütun konumlarını; m ve n ise filtre matrisinin satır ve sütun indislerini ifade etmektedir. Filtrenin giriş matrisi üzerinde

kaydırılmasıyla elde edilen çıktı, ilgili öznitelik haritasını oluşturmaktadır.

CNN mimarilerinde evrişim işleminin ardından doğrusal olmayan ilişkilerin öğrenilebilmesi için aktivasyon fonksiyonları kullanılmaktadır. Bu amaçla yaygın biçimde tercih edilen Rectified Linear Unit (ReLU), $f(x) = max(0, x)$ eşitliği ile tanımlanmaktadır. ReLU fonksiyonu, negatif değerleri sıfıra eşitleyerek pozitif değerleri korumakta ve bu sayede ağın karmaşık örüntüleri öğrenebilmesine katkı sağlamaktadır.

Evrişim katmanlarından sonra kullanılan havuzlama katmanları, öznitelik haritalarının boyutunu küçültmek ve hesaplama yükünü azaltmak amacıyla uygulanmaktadır. Örneğin maksimum havuzlama işleminde, belirli bir pencere içerisindeki en yüksek değer seçilerek daha küçük boyutlu bir çıktı elde edilir. Bu işlem, görüntü içerisindeki temel özelliklerin korunmasına yardımcı olurken, modelin küçük konum değişimlerine karşı daha dayanıklı temsiller üretmesini sağlar. Havuzlama işlemi sonrasında çıktı boyutu Eşitlik (2.3) ile hesaplanabilir.

$$n_{çıktı} = \frac{n_{girdi} + 2p - f}{s} + 1 \tag{2.3}$$

Eşitlikte n_{girdi}, giriş öznitelik haritasının boyutunu; p, dolgu miktarını; f, filtre boyutunu; s, kayma adımını; $n_{çıktı}$ ise işlem sonucunda oluşan çıktı boyutunu ifade etmektedir. Dolgu miktarı, görüntünün kenar bölgelerindeki bilginin korunmasına yardımcı olurken; kayma adımı, filtrenin giriş üzerinde hangi aralıklarla hareket edeceğini belirlemektedir.

CNN mimarisinin son aşamasında, evrişim ve havuzlama katmanlarından elde edilen çok boyutlu öznitelik haritaları tek boyutlu bir vektöre dönüştürülerek tam bağlantılı katmanlara aktarılır. Bu katmanlarda yerel öznitelikler bütünleştirilir ve modelin sınıflandırma ya da tahmin çıktısı üretilir. Bu yapı sayesinde CNN modelleri, görüntülerden otomatik öznitelik çıkarımı yapabilmekte ve nesne tespiti gibi karmaşık bilgisayarlı görü problemlerinde etkin biçimde kullanılabilmektedir (Gu vd. 2018).

2.3.2 YOLO (You Only Look Once) Algoritmaları ve Mimarisi

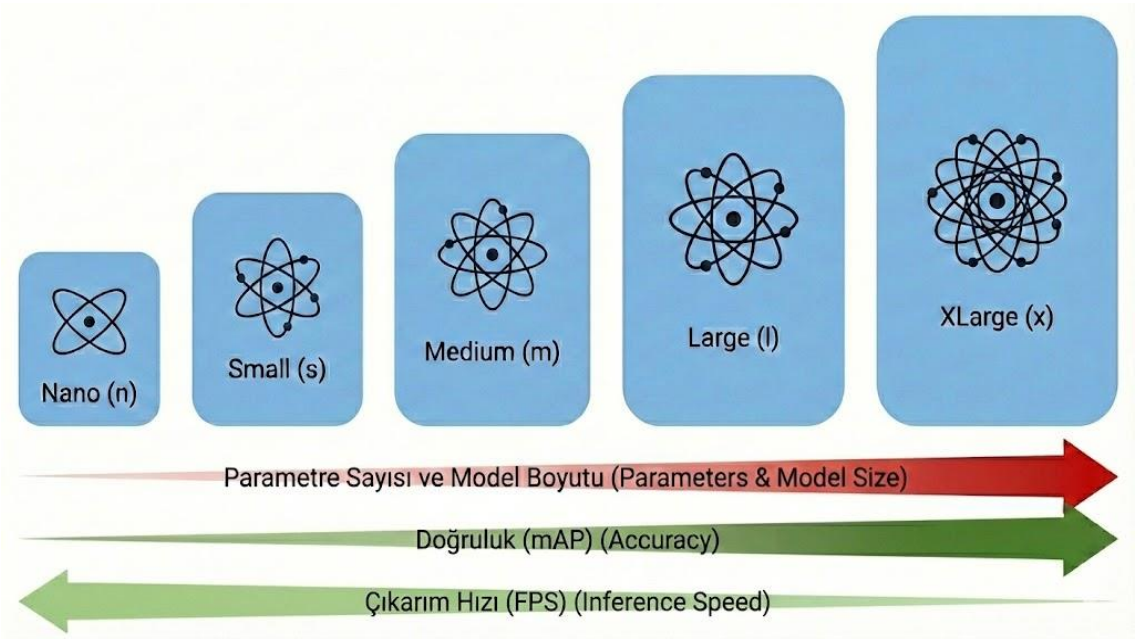
YOLO, nesne tespit problemini tek aşamalı bir regresyon problemi olarak ele alan derin öğrenme tabanlı bir nesne tespit yaklaşımıdır. Bölge önerisi tabanlı iki aşamalı yöntemlerde görüntü öncelikle aday bölgelere ayrılmakta, ardından bu bölgeler sınıflandırılmaktadır. YOLO mimarisinde ise görüntünün tamamı tek bir ağ üzerinden değerlendirilmekte; nesnelerin sınıf olasılıkları, güven skorları ve sınırlayıcı kutu koordinatları aynı işlem sürecinde tahmin edilmektedir. Bu yapı, nesne tespit sürecinin daha hızlı yürütülmesine olanak tanımakta ve YOLO algoritmalarını gerçek zamanlıya yakın uygulamalar için uygun hâle getirmektedir (Redmon vd. 2016).

YOLO'nun temel çalışma mantığında girdi görüntüsü ızgara yapısına ayrılmakta ve her bir ızgara hücresi, merkez noktası kendi bölgesine düşen nesnelerin tespitinden sorumlu olmaktadır. Model, her bir nesne için sınırlayıcı kutu konumlarını, ilgili kutuya ait güven skorunu ve sınıf olasılıklarını üretmektedir. Böylece görüntüde yer alan nesnelerin hem hangi sınıfa ait olduğu hem de görüntü düzlemindeki konumu aynı anda belirlenebilmektedir. Bu yaklaşım, özellikle video görüntüleri üzerinde ardışık karelerde nesne tespiti yapılması gereken uygulamalarda hız açısından önemli bir avantaj sağlamaktadır. Modern YOLO mimarileri genel olarak omurga, boyun ve baş olmak üzere üç temel bileşenden oluşmaktadır (Şekil 2.8).

Şekil 2.8 YOLO mimarisi.
Omurga bölümü, girdi görüntüsünden farklı düzeylerde öznitelikler çıkaran temel ağ yapısıdır. Bu bölümde düşük seviyeli kenar ve doku özelliklerinden başlayarak daha karmaşık nesne temsillerine kadar farklı özellikler öğrenilmektedir. Boyun bölümü, omurga tarafından elde edilen farklı ölçeklerdeki öznitelik haritalarını birleştirerek küçük, orta ve büyük nesnelerin daha etkili biçimde algılanmasına katkı sağlamaktadır. Baş bölümü ise sınıf olasılıkları, güven skoru ve sınırlayıcı kutu konumlarını üreten çıktı katmanı olarak görev yapmaktadır (Wang vd. 2023).

YOLO ailesindeki modeller, farklı donanım kapasiteleri ve uygulama gereksinimleri dikkate alınarak çeşitli ölçeklerde sunulmaktadır. Nano, small, medium, large ve xlarge

gibi model varyantları; parametre sayısı, hesaplama maliyeti, çıkarım hızı ve doğruluk bakımından farklı özellikler göstermektedir. Küçük ölçekli modeller daha az parametre içerdiğinden daha hızlı çalışmakta ve düşük donanım kapasitesine sahip sistemlerde kullanılabilmektedir. Buna karşılık büyük ölçekli modeller, daha fazla parametre ve işlem gücü gerektirmesine rağmen özellikle küçük nesnelerin tespiti, karmaşık arka planlar ve yoğun sahnelerde daha yüksek doğruluk sağlayabilmektedir (İnt. Kyn. 1). Şekil 2.9'da YOLO model ölçekleri arasındaki parametre sayısı, doğruluk ve çıkarım hızı ilişkisi şematik olarak gösterilmektedir.



Şekil 2.9 YOLO mimarisi ölçeklendirme karşılaştırması.

Model ölçeğinin artırılması genellikle mean Mean Average Precision (Ortalama Hassasiyetlerin Ortalaması - mAP) değerinde iyileşme sağlarken, Frames Per Second (FPS) cinsinden ifade edilen çıkarım hızında azalmaya ve hesaplama maliyetinde artışa neden olmaktadır. Bu nedenle YOLO tabanlı uygulamalarda model seçimi yalnızca doğruluk başarımına göre değil, aynı zamanda çalışma ortamının donanım kapasitesi, görüntü çözünürlüğü, hedef nesne boyutu ve gerçek zamanlıya yakın çalışma gereksinimi dikkate alınarak yapılmalıdır.

Bu tez çalışmasında YOLO tabanlı nesne tespit modellerinin tercih edilmesinin temel nedeni, video görüntülerinden nesne konumlarının hızlı biçimde elde edilebilmesi ve bu

sonuçların 3B dijital ikiz ortamına aktarılabilecek konumsal bilgiler üretmeye uygun olmasıdır. Bu nedenle YOLO mimarisi, çalışmada sabit kamera görüntülerinden elde edilen hareketli nesne tespitlerinin temelini oluşturmaktadır.

2.3.2.1 YOLO Versiyon Mimarilerinin Karşılaştırılması

YOLO algoritması, ilk sürümünden itibaren nesne tespitinde hız ve doğruluk dengesini iyileştirmek amacıyla farklı mimari düzenlemelerle geliştirilmiştir. İlk YOLO sürümü, nesne tespitini tek aşamalı bir problem olarak ele alarak gerçek zamanlı nesne tespiti yaklaşımının temelini oluşturmuştur. Daha sonraki sürümlerde ise omurga yapısı, öznitelik birleştirme stratejileri, çıktı katmanı, kayıp fonksiyonları ve eğitim yaklaşımları üzerinde çeşitli iyileştirmeler yapılmıştır (Redmon vd. 2016, Wang vd. 2023).

YOLOv1–YOLOv7 aralığındaki sürümler, YOLO mimarisinin temel gelişim çizgisini oluşturmaktadır. İlk sürümler daha çok hız avantajını öne çıkarırken, sonraki sürümlerde çok ölçekli tespit, daha güçlü omurga yapıları, gelişmiş öznitelik birleştirme katmanları ve veri artırma stratejileri kullanılmaya başlanmıştır. YOLOv4 ve YOLOv5 ile birlikte mozaik veri artırımı, CSP tabanlı yapıların kullanımı ve eğitim sürecine yönelik iyileştirmeler öne çıkmıştır. YOLOv7 ise E-ELAN yapısı ve eğitim sırasında kullanılan ek iyileştirmelerle gerçek zamanlı nesne tespitinde hız ve doğruluk dengesini geliştirmeyi amaçlamıştır (Wang vd. 2023).
YOLOv8, Ultralytics tarafından geliştirilen ve önceki YOLO sürümlerine göre daha esnek bir mimari sunan modellerden biridir. YOLOv8'de anchor-free tespit yaklaşımı benimsenmiş, sınıflandırma ve kutu regresyonu işlemlerinin daha ayrık biçimde yürütülmesine olanak tanıyan bir çıktı yapısı kullanılmıştır. Ayrıca C3 modülleri yerine daha verimli C2f modüllerinin kullanılması, modelin öznitelik çıkarımı ve bilgi akışı bakımından daha etkin çalışmasına katkı sağlamaktadır (Çelik 2024, İnt. Kyn. 1).

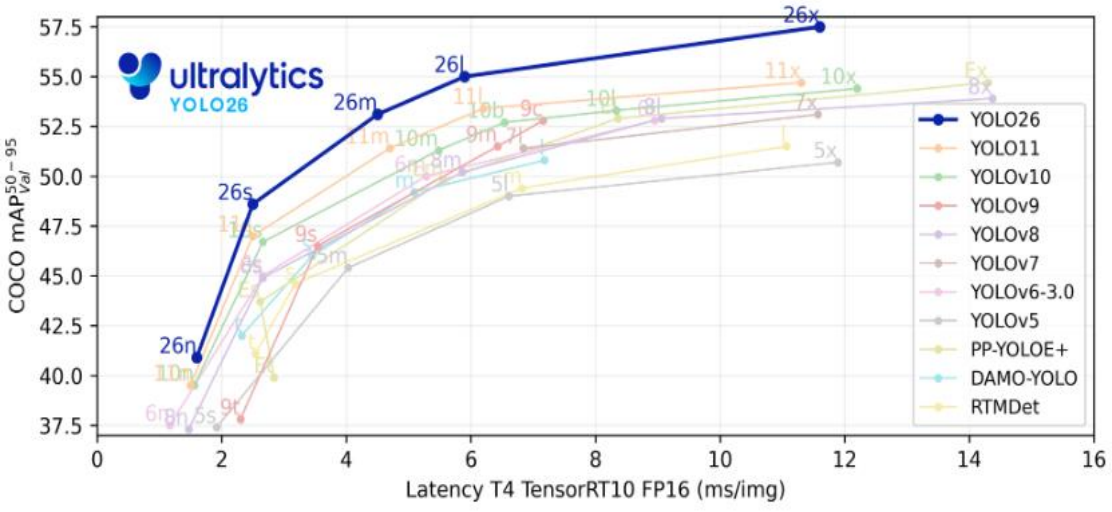
YOLO11, YOLOv8 sonrasında geliştirilen Ultralytics tabanlı YOLO ailesi içinde daha verimli parametre kullanımı ve gelişmiş öznitelik çıkarımı hedefleyen bir sürüm olarak sunulmuştur. Bu sürümde C2f blokları yerine C3k2 blokları kullanılmış, ayrıca uzamsal

ve kanal düzeyindeki bilgileri daha etkin değerlendirmek amacıyla C2PSA dikkat mekanizmasına yer verilmiştir. Bu yapı, özellikle farklı ölçeklerdeki nesnelerin temsil edilmesi ve modelin doğruluk-verimlilik dengesinin iyileştirilmesi açısından önem taşımaktadır (İnt. Kyn. 1).

YOLO26 ise YOLO ailesinde uçtan uca gerçek zamanlı nesne tespiti ve kaynak verimliliği yönüyle öne çıkan güncel bir model ailesidir. YOLO26'da klasik YOLO yaklaşımlarında yaygın biçimde kullanılan NMS tabanlı işlem sonrası adımlar yerine NMS-free çıkarım yapısı benimsenmiş ve DFL kaldırılarak daha sade bir çıktı başlığı elde edilmiştir. Ayrıca MuSGD optimizasyon yaklaşımı, Progressive Loss ve Small-Target-Aware Label Assignment stratejileriyle eğitim sürecinde kararlılık ve küçük nesne tespiti başarımının artırılması hedeflenmiştir (Jocher vd. 2026).

YOLO26 ailesi, farklı model ölçekleriyle hem doğruluk hem de çıkarım hızı bakımından çeşitli donanım ve uygulama senaryolarına yönelik seçenekler sunmaktadır (Jocher vd. 2026, İnt. Kyn. 1).

Şekil 2.10'da YOLO model ailelerinin COCO veri seti üzerindeki $mAP^{50\text{-}95}$ değerleri ve model ölçeklerine göre performans değişimleri karşılaştırmalı olarak gösterilmektedir. Bu tür karşılaştırmalar, yalnızca doğruluk değerlerinin değil, model boyutu, çıkarım hızı ve donanım gereksinimi gibi ölçütlerin de model seçiminde dikkate alınması gerektiğini göstermektedir.



Şekil 2.10 COCO veri seti üzerinde YOLO modellerinin mAP^{50-95} karşılaştırması (İnt. Kyn. 1).

2.3.2.2 Transfer Öğrenme (Transfer Learning) Yaklaşımı

Derin öğrenme modellerinin rastgele başlatılmış ağırlıklarla sıfırdan eğitilmesi, çoğunlukla büyük miktarda etiketlenmiş veri, yüksek işlem gücü ve uzun eğitim süresi gerektirmektedir. Bu gereksinimler, özellikle sınırlı veri setine sahip uygulamalarda model eğitimi açısından önemli bir kısıt oluşturmaktadır. Bu nedenle transfer öğrenme yaklaşımı, derin öğrenme tabanlı çalışmalarda yaygın biçimde kullanılan yöntemlerden biri hâline gelmiştir. Transfer öğrenme, büyük ve genel amaçlı veri setleri üzerinde önceden eğitilmiş bir modelin öğrendiği ağırlıkların, daha özel bir probleme başlangıç bilgisi olarak aktarılması esasına dayanmaktadır (Pan ve Yang 2010, Weiss vd. 2016).

Konvolüsyonel sinir ağı tabanlı mimarilerde ilk katmanlar genellikle kenar, çizgi, renk geçişi ve doku gibi genel görsel öznitelikleri öğrenirken, daha derin katmanlar nesne parçaları ve sınıfa özgü daha karmaşık temsilleri öğrenmektedir.

Transfer öğrenme yaklaşımında, önceden eğitilmiş modelin genel öznitelik çıkarım yeteneğinden yararlanılarak yeni bir veri seti üzerinde daha hızlı ve kararlı bir öğrenme süreci elde edilebilmektedir. Uygulamanın amacına göre modelin bazı katmanları sabit tutulabilir, bazı katmanları ise hedef veri setine uyum sağlayacak şekilde yeniden eğitilebilir. Bu işlem literatürde ince ayar olarak adlandırılmaktadır.
Bu yaklaşımın başlıca avantajları; eğitim süresinin kısalması, sınırlı veri setleriyle çalışılabilmesi ve modelin genelleme yeteneğinin artırılmasıdır. Ayrıca, önceden eğitilmiş modellerin kullanılması, küçük veri setlerinde aşırı öğrenme riskinin azaltılmasına katkı sağlayabilmektedir. Nesne tespiti uygulamalarında transfer öğrenme, ImageNet veya COCO gibi geniş veri setleri üzerinde eğitilmiş modellerin belirli bir hedef probleme uyarlanmasına olanak tanımaktadır. Bu nedenle YOLO tabanlı modellerde önceden eğitilmiş ağırlıkların kullanılması, özellikle sınırlı sayıda görüntüyle çalışılan uygulamalarda yaygın bir başlangıç yaklaşımı olarak değerlendirilmektedir.

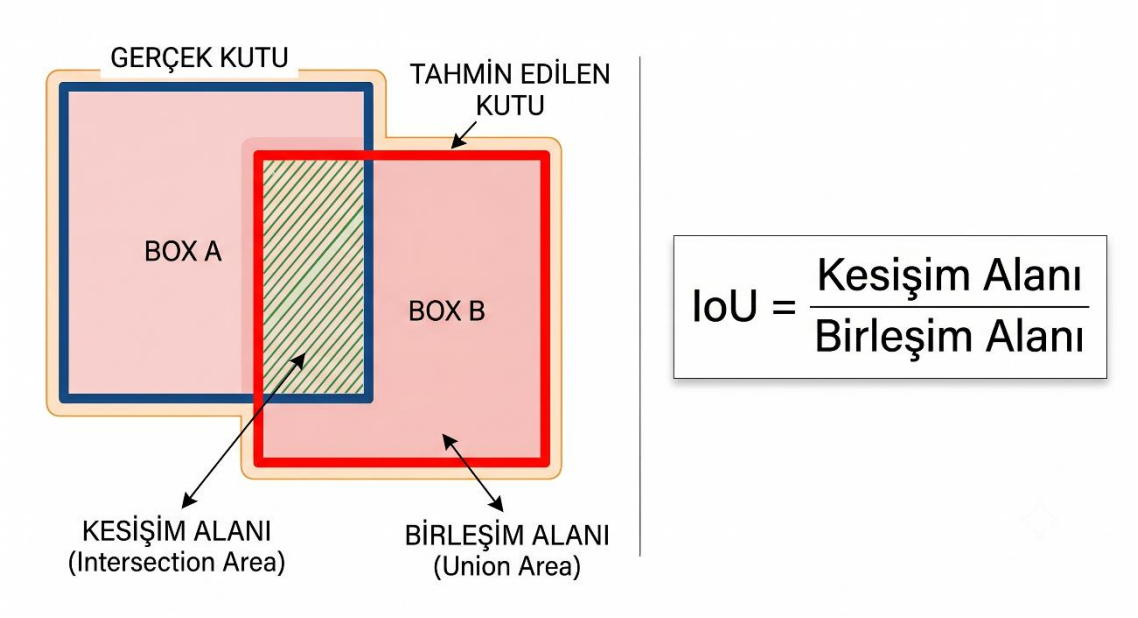
2.3.3 Nesne Tespitinde Performans Değerlendirme Ölçütleri

Nesne tespit modellerinin başarımı, yalnızca görüntüdeki nesnelerin doğru sınıflandırılmasıyla değil, aynı zamanda bu nesnelerin görüntü düzlemindeki konumlarının ne kadar doğru belirlendiğiyle de ilişkilidir. Bu nedenle nesne tespiti problemlerinde sınıflandırma başarımı ile konumlandırma başarımını birlikte değerlendiren ölçütler kullanılmaktadır. Literatürde yaygın olarak kullanılan temel ölçütler arasında kesinlik, duyarlılık, F1 skoru, Intersection over Union (IoU), Average Precision (AP) ve mAP yer almaktadır (Rezatofighi vd. 2019, Padilla vd. 2021).

Nesne tespitinde model çıktıları genellikle tahmin edilen sınıf etiketi, güven skoru ve sınırlayıcı kutu koordinatlarından oluşmaktadır. Bir tahminin doğru kabul edilebilmesi için tahmin edilen sınıfın doğru olması ve tahmin edilen sınırlayıcı kutunun gerçek kutuyla yeterli düzeyde örtüşmesi gerekmektedir. Bu örtüşme IoU değeri ile değerlendirilmektedir. IoU, tahmin edilen sınırlayıcı kutu ile gerçek sınırlayıcı kutu arasındaki kesişim alanının birleşim alanına oranı olarak tanımlanmaktadır. IoU değerinin 1'e yaklaşması, tahmin edilen kutunun gerçek kutuyla daha yüksek oranda örtüştüğünü göstermektedir.

$$IoU = \frac{A\left(B_{tahmin} \cap B_{gerçek}\right)}{A\left(B_{tahmin} \cup B_{gerçek}\right)} \tag{2.4}$$

Eşitlikte B_{tahmin}, model tarafından tahmin edilen sınırlayıcı kutuyu; $B_{gerçek}$, etiketlenmiş gerçek sınırlayıcı kutuyu; $\cap$ kesişim alanını ve $\cup$ birleşim alanını ifade etmektedir. Belirlenen IoU eşik değerinin üzerinde kalan ve doğru sınıfa ait olan tahminler doğru pozitif; yanlış sınıfa ait olan veya yeterli örtüşme sağlamayan tahminler ise yanlış pozitif olarak değerlendirilmektedir. Tahmin edilen sınırlayıcı kutu ile gerçek sınırlayıcı kutu arasındaki örtüşme ilişkisi Şekil 2.11'de gösterilmiştir.



Şekil 2.11 Sınırlayıcı kutular arasındaki IoU hesaplama mantığı.

Kesinlik, model tarafından pozitif olarak tahmin edilen sonuçların ne kadarının gerçekten doğru olduğunu göstermektedir. Duyarlılık ise veri setindeki gerçek nesnelerin ne kadarının model tarafından tespit edilebildiğini ifade etmektedir. Bu iki ölçüt Eşitlik (2.5) ve Eşitlik (2.6) ile hesaplanmaktadır.

$$Kesinlik = \frac{TP}{TP + FP} \tag{2.5}$$

$$Duyarlılık = \frac{TP}{TP + FN} \tag{2.6}$$

Burada True Positive (TP) doğru pozitif tahminleri, False Positive (FP) yanlış pozitif tahminleri, False Negative (FN) ise gerçekte var olduğu hâlde model tarafından tespit edilemeyen nesneleri ifade etmektedir. Kesinlik değerinin yüksek olması, modelin yanlış alarm üretme eğiliminin düşük olduğunu; duyarlılık değerinin yüksek olması ise gerçek nesnelerin büyük bölümünün tespit edilebildiğini göstermektedir. Bu iki ölçütün birlikte değerlendirilmesi amacıyla kullanılan F1 skoru, kesinlik ve duyarlılık değerlerinin harmonik ortalaması olarak tanımlanmaktadır.

$$F1 = 2 \times \left(\frac{Kesinlik * Duyarlılık}{Kesinlik + Duyarlılık}\right) \quad (2.7)$$

F1 skoru, yanlış pozitif ve yanlış negatif tahminlerin birlikte değerlendirilmesi gereken durumlarda model başarımını daha dengeli biçimde yorumlamaya olanak tanımaktadır. Nesne tespiti uygulamalarında güven skoru eşiği de model çıktılarının değerlendirilmesinde önemli bir parametredir. Güven skoru eşiğinin düşük seçilmesi daha fazla nesnenin tespit edilmesini sağlayabilirken yanlış pozitif sayısını artırabilmektedir. Eşiğin yüksek seçilmesi ise daha güvenilir tahminler üretse de bazı nesnelerin kaçırılmasına neden olabilmektedir. Bu nedenle güven skoru eşiği, uygulamanın kesinlik ve duyarlılık beklentilerine göre belirlenmelidir.

AP, farklı güven skoru eşiklerinde elde edilen kesinlik-duyarlılık ilişkisinin özetlenmesiyle hesaplanan bir performans ölçütüdür. Birden fazla sınıfın bulunduğu nesne tespiti problemlerinde her sınıf için hesaplanan AP değerlerinin ortalaması alınarak mAP değeri elde edilmektedir. mAP, farklı nesne tespit modellerinin aynı veri seti üzerinde karşılaştırılmasında yaygın biçimde kullanılan ölçütlerden biridir (Everingham vd. 2010, Lin vd. 2014).

$$mAP = \left(\frac{1}{N}\right) \times \sum_{k=1}^{N} AP_k \quad (2.8)$$

Eşitlikte N, veri setindeki sınıf sayısını; AP_k, k. sınıfa ait AP değerini ifade etmektedir. Bu tez çalışmasında nesne tespit modellerinin değerlendirilmesinde kesinlik, duyarlılık, F1 skoru, IoU, güven skoru ve çıkarım süresi gibi ölçütler birlikte ele alınmıştır. Böylece yalnızca tespit doğruluğu değil, modelin video tabanlı uygulamalarda kullanılabilirliği ve gerçek zamanlıya yakın çalışma potansiyeli de değerlendirilmiştir.

2.4 Görüntü Koordinatlarından Gerçek Dünya Koordinatlarına Dönüşüm

Kamera görüntülerinde nesne tespiti sonucunda elde edilen konum bilgisi, öncelikle görüntü düzlemindeki piksel koordinatlarıyla ifade edilmektedir. Ancak piksel

koordinatları, nesnenin görüntü üzerindeki konumunu belirtmekle birlikte, bu nesnenin gerçek dünyadaki mekânsal karşılığını doğrudan vermez. Bu nedenle görüntü tabanlı nesne tespit sonuçlarının harita, 3B model veya dijital ikiz ortamında değerlendirilebilmesi için görüntü koordinatlarının gerçek dünya koordinatlarıyla ilişkilendirilmesi gerekmektedir.

Görüntü koordinatlarından gerçek dünya koordinatlarına geçiş, kamera geometrisi, kamera konumu ve yönelimi, görüntü çözünürlüğü, lens bozulmaları, YKN doğruluğu, nesne-kamera mesafesi ve kullanılan 3B modelin geometrik niteliği gibi birçok etkene bağlıdır. Bilgisayarlı görü literatüründe bu ilişki, genellikle perspektif projeksiyon, kamera kalibrasyonu, homografi ve 3B yüzeyle ışın-kesişim gibi geometrik yaklaşımlar üzerinden ele alınmaktadır (Hartley ve Zisserman 2004, Szeliski 2022).

Kamera görüntüsü, 3B dünyadaki noktaların 2B görüntü düzlemine izdüşümü olarak düşünülebilir. Bu izdüşüm sürecinde derinlik bilgisi kaybolduğu için, yalnızca tek bir görüntüden doğrudan tam 3B konum elde etmek çoğu durumda mümkün değildir. Bu nedenle görüntüdeki bir pikselin gerçek dünyadaki konumunun belirlenebilmesi için ek bilgilere ihtiyaç duyulmaktadır. Bu bilgiler; kamera kalibrasyon parametreleri, görüntü üzerinde seçilen referans noktaları, yer düzlemi varsayımı, 3B model yüzeyi veya araziye ait kontrol noktaları olabilir.

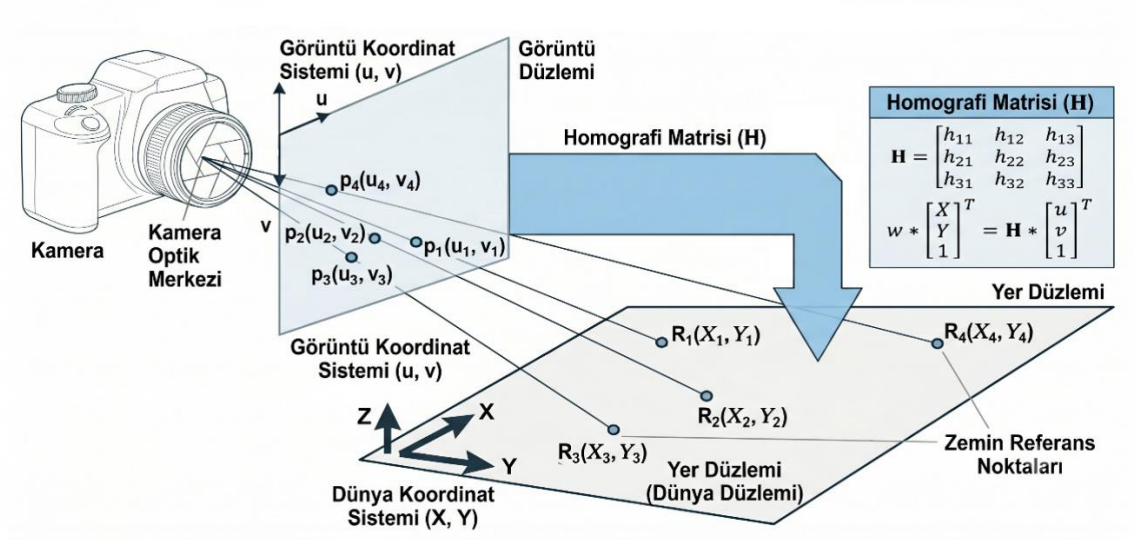
Bu tez çalışmasının konusu açısından görüntü koordinatlarının gerçek dünya koordinatlarıyla ilişkilendirilmesi, video görüntülerinden tespit edilen hareketli nesnelerin dijital ikiz ortamında konumlandırılabilmesi için kritik bir aşamadır. Bu kapsamda literatürde yaygın olarak kullanılan homografi, referans noktalarına dayalı enterpolasyon ve kamera ışını ile 3B yüzey kesişimine dayalı ray-casting yaklaşımı, görüntü tabanlı konum kestiriminde değerlendirilebilecek temel yöntemler arasında yer almaktadır.

2.4.1 Kamera Geometrisi ve Homografi Yaklaşımı

Kamera geometrisi, 3B uzaydaki noktaların kamera görüntüsündeki 2B piksel

koordinatlarına izdüşümünü açıklayan temel yaklaşımlardan biridir. Perspektif projeksiyon sürecinde derinlik bilgisi görüntü düzleminde doğrudan korunmadığından, bir pikselin gerçek dünyadaki konumunun belirlenebilmesi için kamera parametreleri, referans noktaları veya sahne geometrisine ilişkin ek bilgilere ihtiyaç duyulmaktadır (Hartley ve Zisserman 2004, Szeliski 2022).

Homografi, aynı düzlem üzerinde bulunan noktalar arasında projektif dönüşüm kurmak için kullanılan geometrik bir yaklaşımdır. Görüntü düzlemi ile yer düzlemi arasında yeterli sayıda karşılıklı nokta bilindiğinde, bu ilişki homografi matrisiyle ifade edilebilmektedir. Bu nedenle homografi, özellikle düzlemsel kabul edilen zemin veya yol yüzeylerinde piksel koordinatlarının gerçek dünya koordinatlarıyla ilişkilendirilmesinde kullanılmaktadır. Görüntü düzlemi ile yer düzlemi arasındaki homografi ilişkisi Şekil 2.12'de gösterilmiştir.



Şekil 2.12 Görüntü düzlemi ile yer düzlemi arasındaki homografi ilişkisi.
Homografi dönüşümü genel olarak Eşitlik (2.9) ile gösterilebilir:

$$s\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{2.9}$$

Eşitlikte u ve v, görüntü düzlemindeki piksel koordinatlarını; X ve Y, yer düzlemindeki gerçek dünya koordinatlarını; H, görüntü düzleminden yer düzlemine dönüşümü

sağlayan homografi matrisini; s ise ölçek katsayısını ifade etmektedir. Eşitlik (2.9), görüntü düzlemindeki bir piksel noktasının yer düzlemindeki karşılığının homografi matrisi yardımıyla ifade edilmesini göstermektedir. Homografi matrisinin hesaplanabilmesi için görüntü düzlemindeki noktalar ile gerçek dünya düzlemindeki karşılıkları arasında en az dört adet aynı doğru üzerinde bulunmayan nokta eşleşmesine ihtiyaç duyulmaktadır. Nokta sayısının artırılması ve noktaların çalışma alanına dengeli biçimde dağılması, dönüşümün doğruluğunu olumlu yönde etkileyebilmektedir.

Nesne tespiti uygulamalarında homografi yaklaşımı kullanılırken, tespit edilen sınırlayıcı kutunun hangi noktasının gerçek dünya konumunu temsil edeceği önemlidir. Özellikle yaya veya araç gibi nesnelerde, sınırlayıcı kutunun merkezi her zaman nesnenin zeminle temas noktasını temsil etmeyebilir. Bu nedenle görüntüden gerçek dünya koordinatına dönüşümde, çoğu uygulamada sınırlayıcı kutunun alt orta noktası veya nesnenin zemin temasına en yakın noktası referans alınmaktadır. Bu seçim, görüntü düzlemindeki tespit sonucunun yer düzlemindeki konumla daha tutarlı ilişkilendirilmesine katkı sağlamaktadır.

Homografi yaklaşımının temel sınırlılığı, dönüşümün düzlemsel yüzey varsayımına dayanmasıdır. Çalışma alanında eğim değişimleri, kot farkları, yüksek yapılar, kamera görüş açısındaki bozulmalar veya referans noktalarının hatalı belirlenmesi durumunda dönüşüm doğruluğu azalabilmektedir. Bu nedenle homografi, özellikle yol, meydan veya düz zemin gibi yaklaşık düzlemsel kabul edilebilen alanlarda etkili bir yöntem olarak değerlendirilmektedir. Daha karmaşık 3B sahnelerde ise kamera kalibrasyonu, 3B model yüzeyi veya ray-casting gibi yaklaşımlarla desteklenmesi gerekebilmektedir.

2.4.2 Referans Noktalarına Dayalı Konum Kestirimi ve Enterpolasyon

Görüntü koordinatlarından gerçek dünya koordinatlarına geçişte kullanılan yaklaşımlardan biri de görüntü üzerinde belirlenen referans noktalarına dayalı konum kestirimidir. Bu yöntemde, görüntü düzleminde koordinatları bilinen piksel noktaları ile bu noktaların sahadaki gerçek dünya koordinatları arasında ilişki kurulmaktadır. Referans noktaları, genellikle arazi ölçmeleri, YKN'ler, fotogrametrik dengeleme

sonuçları veya harita tabanlı koordinat verileri kullanılarak belirlenmektedir. Böylece görüntüde tespit edilen bir nesneye ait piksel konumu, çevresindeki bilinen referans noktaları yardımıyla gerçek dünya koordinatlarıyla ilişkilendirilebilmektedir.

Referans noktalarına dayalı konum kestirimi, özellikle kamera kalibrasyon parametrelerinin tam olarak bilinmediği veya çalışma alanının sınırlı bir bölgesinde pratik bir dönüşüm yaklaşımına ihtiyaç duyulduğu durumlarda kullanılabilmektedir. Bu yaklaşımda dönüşüm doğruluğu; referans noktalarının konumsal doğruluğuna, görüntü üzerinde doğru seçilmesine, çalışma alanına dengeli biçimde dağılmasına ve tespit edilen nesnenin bu referans noktalarına olan uzaklığına bağlıdır. Referans noktalarının yalnızca belirli bir bölgede yoğunlaşması, görüntünün diğer kısımlarında konum kestirim hatalarının artmasına neden olabilmektedir.

Enterpolasyon yaklaşımı, bilinen referans noktalarından yararlanarak bilinmeyen bir noktanın konum değerinin tahmin edilmesine dayanır. Görüntü üzerinde tespit edilen nesne konumu, çevresindeki referans noktalarının gerçek dünya koordinatları kullanılarak hesaplanabilir. Bu işlemde yakın referans noktalarının, uzak referans noktalarına göre daha fazla etkiye sahip olması beklenir. Bu nedenle ters mesafe ağırlıklı enterpolasyon gibi yöntemler, konum kestiriminde kullanılabilecek temel yaklaşımlar arasında yer almaktadır (Shepard 1968).

Ters mesafe ağırlıklı enterpolasyonun genel ifadesi Eşitlik (2.10) ile gösterilebilir:

$$P = \frac{\sum_{i=1}^{n} w_i P_i}{\sum_{i=1}^{n} w_i} \tag{2.10}$$

$$w_i = \frac{1}{d_i^p} \tag{2.11}$$

Eşitliklerde P, konumu tahmin edilen noktayı; P_i, i. referans noktasına ait gerçek dünya koordinatını; w_i, i. referans noktasına ait ağırlık değerini; d_i, tahmin edilecek piksel konumu ile i. referans noktasının görüntü düzlemindeki piksel konumu arasındaki mesafeyi; p ise mesafenin etkisini belirleyen güç parametresini ifade etmektedir. Bu

yaklaşımda tahmin edilecek noktaya daha yakın olan referans noktaları daha yüksek ağırlık almakta, uzak noktaların etkisi ise azalmaktadır.

Bu yöntemin önemli avantajlarından biri, karmaşık kamera parametrelerine ihtiyaç duymadan görüntü ile gerçek dünya koordinatları arasında pratik bir ilişki kurulabilmesidir. Bununla birlikte enterpolasyon tabanlı konum kestirimi, referans noktalarının doğruluğuna ve dağılımına duyarlıdır. Ayrıca çalışma alanında eğim değişimleri, perspektif bozulmaları, kamera görüş açısı farklılıkları ve nesnenin zemin temas noktasının doğru belirlenememesi gibi durumlar konumsal hataları artırabilmektedir. Bu nedenle referans noktalarına dayalı enterpolasyon yaklaşımı, referans noktalarının dağılımına ve doğruluğuna bağlı olarak değerlendirilmelidir. Daha karmaşık sahnelerde ise homografi veya ray-casting gibi geometrik yaklaşımlarla karşılaştırılması, konum kestirim sonuçlarının daha sağlıklı yorumlanmasına katkı sağlayabilir.

2.4.3 Kamera Kalibrasyonu ve Ray-Casting Yaklaşımı

Görüntü koordinatlarından gerçek dünya koordinatlarına geçişte kullanılan geometrik yaklaşımlardan biri kamera kalibrasyonuna dayalı ray-casting yöntemidir. Kamera kalibrasyonu, kameranın iç parametreleri ile dış yöneltme parametrelerinin belirlenmesi sürecidir. İç parametreler odak uzaklığı, asal nokta ve piksel ölçeği gibi kameraya ait bilgileri içerirken, dış yöneltme parametreleri kameranın gerçek dünya koordinat sistemindeki konumunu ve yönelimini ifade etmektedir.

Bu parametrelerin bilinmesi, görüntü düzlemindeki bir pikselin kamera merkezinden çıkan bir ışın olarak 3B uzayda modellenmesine olanak sağlamaktadır (Zhang 2000, Hartley ve Zisserman 2004).

Ray-casting yaklaşımında, görüntü üzerinde seçilen veya nesne tespiti sonucunda elde edilen bir piksel koordinatından kamera merkezine bağlı bir ışın oluşturulur. Bu ışın, kamera parametreleri kullanılarak 3B uzaya yönlendirilir ve zemin, mesh yüzeyi veya 3B model ile kesiştirilir.

Elde edilen kesişim noktası, ilgili pikselin gerçek dünya koordinat sistemindeki karşılığı olarak değerlendirilebilir. Bu yaklaşım, özellikle sahnenin 3B geometrisinin bilindiği durumlarda görüntü tabanlı tespit sonuçlarının mekânsal olarak konumlandırılmasında kullanılabilmektedir.

Kamera merkezinden çıkan ışın genel olarak Eşitlik (2.12) ile ifade edilebilir:

$$P(\lambda) = C + \lambda d \tag{2.12}$$

Eşitlikte $P(\lambda)$, ışın üzerindeki herhangi bir 3B noktayı; C, kamera merkezini; d, ışının yön vektörünü; λ ise ışın boyunca ilerleme parametresini ifade etmektedir. Ray-casting işleminde λ değerinin belirlenmesi, ışının 3B model yüzeyiyle veya zemin düzlemiyle kesiştiği noktanın bulunmasına karşılık gelmektedir. Eğer yüzey üçgen ağ yapısında temsil ediliyorsa, kamera ışını ile mesh üzerindeki üçgen yüzeyler arasında kesişim hesaplanabilmektedir (Möller ve Trumbore 1997).

Ray-casting yaklaşımının temel avantajı, homografide olduğu gibi yalnızca düzlemsel yüzey varsayımına bağlı kalmadan 3B sahne geometrisinden yararlanabilmesidir. Bu nedenle eğim değişimi, kot farkı veya karmaşık yüzey yapısının bulunduğu alanlarda daha esnek bir konum kestirim yaklaşımı sunabilmektedir.

Bununla birlikte yöntemin doğruluğu; kamera kalibrasyonunun hassasiyetine, kamera konum ve yönelim bilgilerinin doğruluğuna, kullanılan 3B modelin geometrik kalitesine ve nesnenin görüntü üzerindeki temsil noktasının doğru belirlenmesine bağlıdır.

2.5 Video Tabanlı Konum Kestirimi ve Filtreleme Yaklaşımları

Video tabanlı nesne tespiti uygulamalarında her bir görüntü karesinden elde edilen sınırlayıcı kutu bilgileri, nesnenin görüntü düzlemindeki anlık konumunu temsil etmektedir. Bu konum bilgileri; homografi, referans noktalarına dayalı enterpolasyon veya ray-casting gibi yöntemlerle gerçek dünya koordinatlarıyla ilişkilendirildiğinde, hareketli nesnelere ait zamana bağlı konum serileri elde edilebilmektedir.

Ancak ardışık video karelerinden üretilen konum değerleri, tespit modelinin kare bazlı değişken çıktıları, sınırlayıcı kutu konumundaki küçük farklılıklar, örtülme durumu, kamera görüş açısı ve koordinat dönüşüm sürecindeki belirsizlikler nedeniyle gürültü içerebilmektedir. Bu durum, dijital ikiz ortamında hareketli nesnelerin kararsız veya titreşimli biçimde görünmesine neden olabilir.

Bu nedenle video tabanlı konum kestirimlerinde yalnızca anlık konum doğruluğu değil, ardışık kareler arasındaki zamansal tutarlılık da dikkate alınmalıdır. Filtreleme yaklaşımları, konum serilerindeki ani dalgalanmaların azaltılması ve nesnenin gerçek hareket eğilimine daha uygun bir yörüngeyle temsil edilmesi amacıyla kullanılmaktadır. Bu kapsamda hareketli ortalama, üstel yumuşatma ve Kalman filtresi gibi yöntemler zamana bağlı konum verilerinin düzenlenmesinde kullanılabilecek temel yaklaşımlar arasında yer almaktadır (Kalman 1960, Welch ve Bishop 2006).

2.5.1 Konum Kestiriminde Gürültü ve Zamansal Kararlılık

Video tabanlı konum kestiriminde, hareketli nesnelere ait konum bilgileri ardışık görüntü karelerinden elde edilmektedir. Her bir karede nesne tespiti sonucu üretilen sınırlayıcı kutu, nesnenin görüntü düzlemindeki anlık temsilini sağlamaktadır. Ancak bu temsil her zaman sabit ve hatasız değildir. Nesnenin görüntüdeki görünümü, kamera açısı, ışık koşulları, hareket bulanıklığı, örtülme durumu ve tespit modelinin kare bazlı değişken çıktıları nedeniyle sınırlayıcı kutu konumunda küçük farklılıklar oluşabilmektedir (Yilmaz vd. 2006).

Sınırlayıcı kutu konumundaki bu küçük değişimler, görüntü koordinatlarından gerçek dünya koordinatlarına dönüştürme aşamasında daha belirgin konumsal dalgalanmalara neden olabilmektedir. Özellikle nesnenin alt orta noktası, merkez noktası veya zemin temas noktası gibi belirli bir temsil noktası üzerinden konum kestirimi yapıldığında, tespit kutusundaki birkaç piksellik değişim gerçek dünya koordinatlarında ani sıçramalar şeklinde görülebilmektedir. Bu durum, nesnenin gerçekte düzgün bir hareket izlediği durumlarda dahi konum serisinin düzensiz görünmesine yol açabilir.

Konum kestirimindeki gürültü yalnızca nesne tespitinden kaynaklanmamaktadır. Görüntüden gerçek dünya koordinatına geçişte kullanılan homografi, referans noktalarına dayalı enterpolasyon veya ray-casting gibi yöntemlerdeki belirsizlikler de konumsal hatalara neden olabilmektedir. Referans noktalarının görüntü üzerinde hatalı seçilmesi, 3B model yüzeyinin geometrik doğruluğu, kamera kalibrasyonundaki küçük sapmalar ve nesnenin zemin temas noktasının doğru belirlenememesi bu hataların başlıca nedenleri arasında yer almaktadır.

Zamansal kararlılık, ardışık karelerde elde edilen konum kestirimlerinin fiziksel hareket sürekliliğiyle uyumlu olması anlamına gelmektedir. Gerçek dünyada hareket eden bir yaya veya aracın konumu, kısa zaman aralıklarında genellikle süreklilik gösterir. Buna karşılık ham konum kestirimleri, tespit veya dönüşüm hataları nedeniyle kareler arasında ani ve fiziksel olarak gerçekçi olmayan değişimler içerebilir. Bu tür sıçramalar, özellikle dijital ikiz ortamında hareketli nesnelerin görselleştirilmesi sırasında nesnenin titreşimli, kararsız veya gerçek hareketinden kopuk görünmesine neden olabilir.

Bu nedenle video tabanlı konum kestirimlerinde yalnızca tekil karelerde elde edilen konum doğruluğu değil, ardışık kareler arasındaki zamansal tutarlılık da dikkate alınmalıdır. Konum serilerinin filtrelenmesi veya yumuşatılması, anlık ölçüm hatalarının etkisini azaltarak hareketli nesnenin daha kararlı bir yörüngeyle temsil edilmesine katkı sağlayabilmektedir. Bununla birlikte aşırı yumuşatma, nesnenin gerçek yön değişimlerini veya ani hareketlerini bastırabileceğinden, filtreleme işleminin konumsal gürültüyü azaltırken gerçek hareket dinamiklerini koruyacak biçimde uygulanması gerekmektedir.

2.5.2 Kalman Filtreleme Yaklaşımı

Kalman filtresi, gürültü içeren ölçümlerden yararlanarak bir sistemin zamanla değişen durumunu tahmin etmek amacıyla kullanılan doğrusal kestirim yöntemlerinden biridir. Bu yöntem, sistemin önceki durumundan hareketle mevcut durumu tahmin etmekte ve daha sonra yeni ölçüm bilgisi geldiğinde bu tahmini güncellemektedir. Bu yönüyle Kalman filtresi, özellikle konum, hız, hareket takibi ve zamana bağlı ölçüm verilerinin

işlenmesi gibi problemlerde yaygın biçimde kullanılmaktadır (Kalman 1960, Welch ve Bishop 2006).

Kalman filtresinin temel çalışma mantığı tahmin ve güncelleme olmak üzere iki aşamadan oluşmaktadır. Tahmin aşamasında, sistemin bir önceki durumundan yararlanılarak mevcut duruma ilişkin ön kestirim yapılır. Güncelleme aşamasında ise yeni ölçüm bilgisi dikkate alınarak bu tahmin düzeltilir. Böylece filtre, yalnızca anlık ölçüme bağlı kalmadan, hem önceki hareket eğilimini hem de mevcut ölçüm bilgisini birlikte değerlendirir. Bu yaklaşım, ölçüm gürültüsünün bulunduğu durumlarda daha kararlı bir durum kestirimi elde edilmesine katkı sağlamaktadır.

Kalman filtresi genel olarak durum ve ölçüm modelleri üzerinden tanımlanır. Durum modeli, sistemin zaman içindeki değişimini; ölçüm modeli ise gözlenen ölçümlerin sistem durumuyla ilişkisini ifade etmektedir. Bu ilişki genel biçimde Eşitlik (2.13) ve Eşitlik (2.14) ile gösterilebilir:

$$x_i = F_i x_{\{i-1\}} + w_i \tag{2.13}$$

$$z_i = H_i x_i + v_i \tag{2.14}$$

Eşitliklerde x_i, $i.$ zamandaki sistem durumunu; F_i, durum geçiş matrisini; w_i, süreç gürültüsünü; z_i, ölçüm vektörünü; H_i, ölçüm matrisini; v_i ise ölçüm gürültüsünü ifade etmektedir. Süreç gürültüsü, modelin gerçek hareketi tam olarak temsil edemediği durumları; ölçüm gürültüsü ise sensör veya ölçüm kaynaklı belirsizlikleri temsil etmektedir.

Video tabanlı konum kestirimi açısından Kalman filtresi, ardışık karelerden elde edilen ham konum ölçümlerinin zamansal olarak daha kararlı hâle getirilmesinde kullanılabilmektedir. Nesne tespitinden elde edilen konumlar, her karede küçük dalgalanmalar gösterebildiğinden, bu ölçümlerin doğrudan görselleştirilmesi dijital ikiz ortamında titreşimli veya düzensiz hareket izlenimine neden olabilir. Kalman filtresi bu noktada, ölçülen konum değerlerini önceki hareket eğilimiyle birlikte değerlendirerek daha tutarlı bir konum serisi elde edilmesine yardımcı olur.

Bununla birlikte Kalman filtresinin başarımı, seçilen hareket modeline ve gürültü parametrelerine bağlıdır. Hareket modeli gerçek hareket davranışını yeterince temsil etmiyorsa veya ölçüm ve süreç gürültüsü uygun biçimde belirlenmemişse, filtrelenmiş konum serisi gerçek hareketten sapabilir. Bu nedenle video tabanlı konum kestiriminde Kalman filtresi uygulanırken, ölçüm güvenilirliği, beklenen hareket dinamiği ve konum serisinin zamansal davranışı birlikte değerlendirilmelidir.

2.6 Dijital İkiz ve Web Tabanlı 3B Coğrafi Bilgi Sistemleri

İnternet ve web teknolojilerindeki gelişmeler, mekânsal verilerin farklı kullanıcılar tarafından daha erişilebilir, paylaşılabilir ve etkileşimli biçimde kullanılmasına olanak sağlamıştır. Bu gelişmelerle birlikte CBS uygulamaları yalnızca masaüstü yazılımlarla sınırlı kalmamış; web tarayıcıları üzerinden çalışan, farklı kullanıcıların aynı mekânsal veriye uzaktan erişebildiği web tabanlı yapılara da dönüşmüştür. Web tabanlı CBS uygulamaları, kullanıcıların ek yazılım kurulumuna ihtiyaç duymadan mekânsal verileri görüntüleyebilmesine, sorgulayabilmesine ve belirli analiz işlemlerini gerçekleştirebilmesine katkı sağlamaktadır (Boulos vd. 2011).

Geleneksel 2B haritalar, kent ölçeğindeki karmaşık yapıların, düşey ilişkilerin ve hacimsel mekânsal bilgilerin temsilinde sınırlı kalabilmektedir. Bu nedenle kentsel modelleme, çevre planlama, afet yönetimi, ulaşım ve akıllı şehir uygulamaları gibi alanlarda 3B mekânsal veri kullanımına duyulan ihtiyaç artmıştır. 3B kent modelleri, yapıların yalnızca yatay konumlarını değil; yükseklik, cephe, hacim ve çevresel ilişkilerini de temsil edebildiğinden, mekânsal analiz ve görselleştirme süreçlerinde daha zengin bir veri altyapısı sunmaktadır (Biljecki vd. 2015).

Web tabanlı 3B CBS yaklaşımları, yüksek hacimli 3B mekânsal verilerin internet ortamında görselleştirilmesi ve farklı veri kaynaklarıyla bütünleştirilmesi açısından önemli bir altyapı sağlamaktadır. Bu altyapı, statik 3B kent modellerinin yanı sıra sensör verileri, video analiz çıktıları ve zamana bağlı konum bilgileri gibi dinamik verilerin de mekânsal bağlam içinde sunulmasına olanak tanımaktadır. Böylece dijital

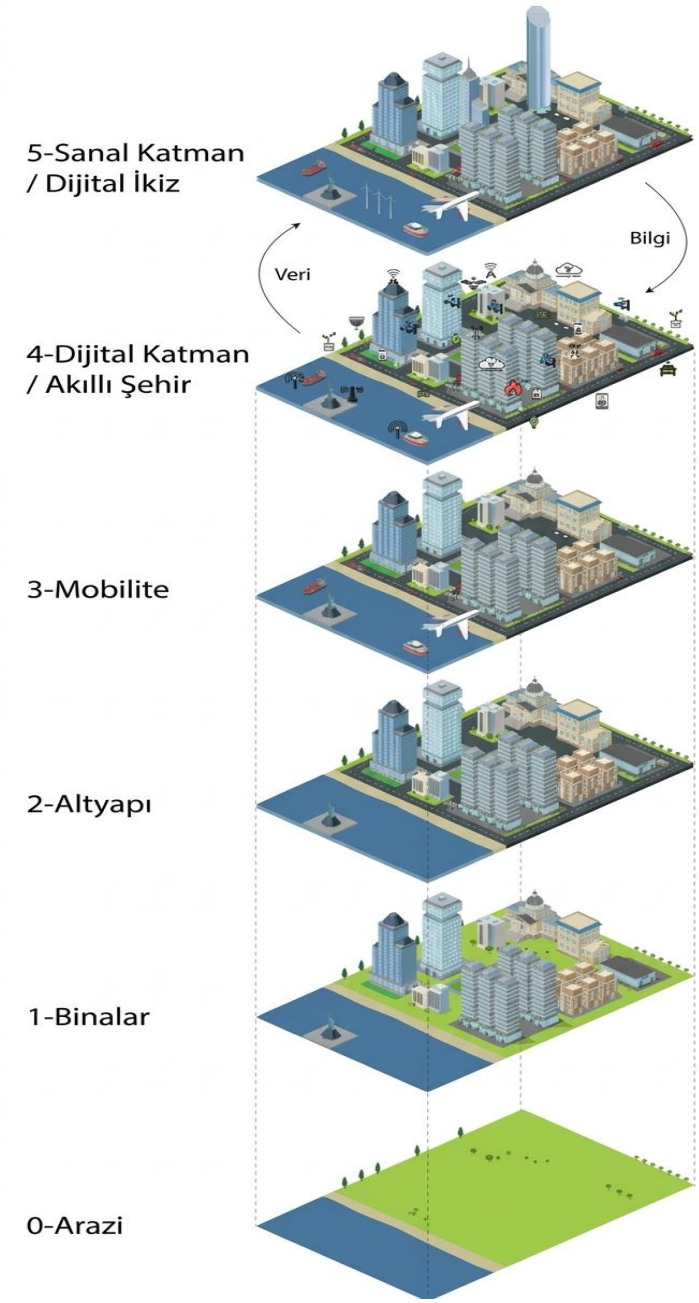
ikiz uygulamalarında fiziksel çevrenin 3B temsili ile gerçek veya zamana bağlı veri akışlarının birlikte değerlendirilmesi mümkün hâle gelmektedir (White vd. 2021, Zhong vd. 2024).

Bu tez çalışması açısından web tabanlı 3B CBS altyapısı, İHA ve LiDAR verileriyle oluşturulan 3B dijital ikiz altlığının sunulması ve sabit kamera görüntülerinden elde edilen nesne tespit sonuçlarının mekânsal olarak görselleştirilmesi için temel bir bileşen olarak değerlendirilmektedir. Bu kapsamda 3B model, nesne konumları ve video tabanlı analiz çıktılarının aynı ortamda bütünleştirilmesi, görüntü tabanlı tespit sonuçlarının coğrafi bağlamla ilişkilendirilmesine katkı sağlamaktadır.

2.6.1 Dijital İkiz Kavramı ve Kent Modelleri

Dijital ikiz, fiziksel bir varlığın, sürecin veya sistemin dijital ortamda temsil edilmesini ve bu temsilin gerçek dünyadan elde edilen verilerle desteklenmesini ifade eden bir yaklaşımdır. İlk olarak ürün yaşam döngüsü yönetimi kapsamında ele alınan bu kavram, zamanla akıllı şehirler, altyapı yönetimi, afet yönetimi, ulaşım ve kentsel planlama gibi alanlarda da kullanılmaya başlanmıştır (Grieves 2014, White vd. 2021). Kent ölçeğinde dijital ikizler, fiziksel çevreye ait 3B mekânsal verilerin sensör verileri, kamera görüntüleri, trafik bilgileri ve diğer zamana bağlı veri kaynaklarıyla ilişkilendirilmesiyle daha dinamik bir karar destek ortamı sunabilmektedir.

Dijital ikiz mimarisinde topografya, yapılar, altyapı ağları, mobilite verileri ve dijital veri akışları birlikte değerlendirilebilmektedir. Topografya katmanı arazi yüzeyi, eğim ve doğal çevre özelliklerini temsil ederken; yapı katmanı binalar ve kentsel dokunun 3B temsiline karşılık gelmektedir. Altyapı katmanı ulaşım, enerji ve iletişim ağlarını; mobilite katmanı ise yaya, araç ve lojistik hareketlerini kapsamaktadır. Sensörler, kameralar ve diğer veri kaynaklarından elde edilen dinamik bilgiler ise sanal ortamın güncellenmesine ve fiziksel çevredeki değişimlerin dijital model üzerinde izlenmesine katkı sağlamaktadır (White vd. 2021). Gerçek dünya verileri ile 3B kent modelleri arasındaki dijital ikiz ilişkisi Şekil 2.13'te gösterilmiştir.

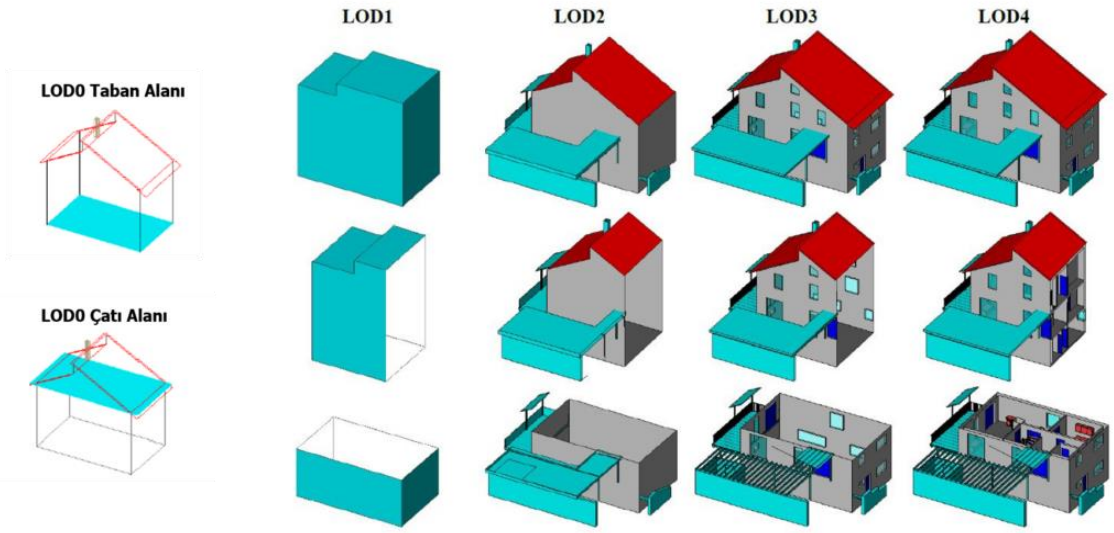


Şekil 2.13 Gerçek dünya verilerinin (sensörler, IoT) 3B kent modelleri ile entegrasyonunu gösteren dijital ikiz mimarisi (White vd. 2021'den uyarlanmıştır).

Kent ölçeğindeki dijital ikizlerin temel mekânsal altlığını yüksek doğruluklu 3B kent modelleri oluşturmaktadır. 3B kent modelleri, yalnızca yapıların görsel temsilini sunmakla kalmamakta; aynı zamanda yükseklik, cephe, hacim, topoloji ve semantik bilgi gibi bileşenler aracılığıyla görünürlük analizi, enerji modellemesi, taşkın analizi ve kentsel planlama gibi uygulamalara altlık sağlayabilmektedir (Biljecki vd. 2015). Bu modellerin farklı sistemler arasında paylaşılabilmesi ve standart bir yapıda temsil edilebilmesi amacıyla Open Geospatial Consortium (OGC) tarafından geliştirilen

CityGML gibi modelleme standartları kullanılmaktadır (Gröger vd. 2012).

3B kent modelleri, kullanım amacına ve ihtiyaç duyulan geometrik ayrıntı düzeyine göre farklı Level of Detail (LOD) seviyelerinde üretilebilmektedir. CityGML yaklaşımında LOD0'dan LOD4'e kadar farklı ayrıntı düzeyleri tanımlanmaktadır. LOD0 genellikle arazi modeli ve bina taban izlerini içerirken, LOD1 basit blok yapı modellerini temsil etmektedir. LOD2 çatı geometrileri ve daha ayrıntılı dış yapı temsilini, LOD3 cephe elemanları gibi mimari ayrıntıları, LOD4 ise iç mekân bileşenlerini kapsayan en ayrıntılı model düzeyini ifade etmektedir (Gröger vd. 2012). CityGML standartlarına göre farklı LOD seviyeleri Şekil 2.14'te gösterilmiştir.



Şekil 2.14 CityGML standartlarına göre farklı ayrıntı düzeylerinin (LOD0 - LOD4) gösterimi (Gröger vd. 2012'den uyarlanmıştır).

Bu tez çalışması açısından dijital ikiz kavramı, çalışma alanına ait 3B model altlığının yalnızca statik bir görselleştirme ortamı olarak değil, video tabanlı nesne tespit sonuçlarıyla ilişkilendirilebilen mekânsal bir temsil ortamı olarak ele alınmasını sağlamaktadır. Bu bağlamda 3B kent modeli, kamera görüntülerinden elde edilen nesne konumlarının gerçek dünya koordinatlarıyla birlikte sunulabilmesi için temel mekânsal referans altyapısını oluşturmaktadır.

2.6.2 Web Tabanlı CBS Platformları ve Görselleştirme Teknolojileri

Mekânsal verilerin web ortamında sunulması amacıyla farklı veri türlerini, görselleştirme yaklaşımlarını ve kullanıcı etkileşimlerini destekleyen çeşitli platformlar geliştirilmiştir. Bu platformların seçiminde projenin 2B veya 3B görselleştirme ihtiyacı, kullanılacak veri türü, veri hacmi, performans gereksinimi, lisanslama koşulları ve mevcut sistemlerle entegrasyon olanakları belirleyici olmaktadır.

OpenLayers, web tabanlı harita uygulamalarında yaygın olarak kullanılan açık kaynaklı JavaScript kütüphanelerinden biridir. Vektör ve raster harita katmanlarının sunulması, farklı veri kaynaklarının harita üzerinde gösterilmesi ve kullanıcı etkileşimlerinin yönetilmesi açısından güçlü bir altyapı sağlamaktadır. Bu nedenle OpenLayers, özellikle 2B web haritacılığı ve vektörel veri yönetimi odaklı uygulamalarda tercih edilmektedir (İnt. Kyn. 2).

ArcGIS tabanlı web çözümleri, Esri ekosistemiyle bütünleşik çalışabilen ve 2B harita ile 3B sahne görselleştirme olanakları sunan ticari platformlar arasında yer almaktadır. Bu yapı, kurumsal CBS altyapılarıyla entegrasyon, mekânsal analiz ve veri yönetimi açısından kapsamlı olanaklar sağlamakla birlikte, lisanslama ve ekosistem bağımlılığı açısından açık kaynaklı çözümlerden ayrılmaktadır (İnt. Kyn. 3).

Mapbox GL JS, vektör döşeme tabanlı harita sunumu, donanım hızlandırmalı görselleştirme, terrain gösterimi ve 3B bina katmanları gibi özellikleriyle web tabanlı harita uygulamalarında kullanılan bir diğer platformdur. Özellikle stil yönetimi, hızlı harita sunumu ve etkileşimli görselleştirme açısından güçlü bir yapı sunmaktadır. Bununla birlikte, büyük hacimli fotogrametrik 3B modeller, LiDAR nokta bulutları ve dijital ikiz altlıkları gibi heterojen 3B mekânsal verilerin küresel ölçekte yönetimi için her durumda tek başına yeterli bir çözüm olmayabilir (İnt. Kyn. 4).

Three.js ise doğrudan bir CBS kütüphanesi olmamakla birlikte, web tarayıcılarında WebGL tabanlı 3B grafiklerin oluşturulması için yaygın olarak kullanılan bir JavaScript kütüphanesidir. Mimari modelleme, ürün görselleştirme ve yerel 3B sahne oluşturma gibi uygulamalarda esnek bir grafik altyapısı sunmaktadır. Ancak küresel koordinat sistemleri, jeodezik dönüşümler, arazi verileri ve büyük ölçekli mekânsal veri yönetimi

gibi CBS odaklı gereksinimler için ek geliştirmelere ihtiyaç duyulabilmektedir (İnt. Kyn. 5).

Bu platformlar incelendiğinde, web tabanlı mekânsal görselleştirme çözümlerinin farklı kullanım senaryolarına göre güçlü yönler sunduğu görülmektedir. Bu tez çalışmasında ise yüksek hacimli 3B dijital ikiz altlığının gerçek dünya koordinatlarıyla birlikte sunulması, video tabanlı nesne tespit sonuçlarının 3B ortamda görselleştirilmesi ve zamanla değişen konum bilgilerinin sahne üzerinde izlenebilmesi gerekmektedir. Bu nedenle küresel koordinat sistemleriyle uyumlu, 3D Tiles veri yapısını destekleyen ve web ortamında 3B mekânsal verilerin etkileşimli olarak sunulmasına olanak sağlayan CesiumJS altyapısı, çalışmanın görselleştirme bileşeni açısından uygun bir seçenek olarak değerlendirilmektedir.

2.6.3 WebGL Mimarisi ve CesiumJS Kütüphanesi

Büyük boyutlu 3B kent modellerinin, nokta bulutlarının ve dijital ikiz altlıklarının web ortamında etkileşimli biçimde sunulabilmesi için tarayıcı tabanlı grafik işleme teknolojilerine ihtiyaç duyulmaktadır (Cozzi ve Ring 2011). Bu kapsamda Web Graphics Library (WebGL), web tarayıcılarında eklenti gerektirmeden 2B ve 3B grafiklerin oluşturulmasına olanak sağlayan temel teknolojilerden biridir. WebGL, grafik işlemlerinin donanım hızlandırmalı olarak yürütülmesini desteklediğinden, büyük hacimli mekânsal verilerin web ortamında görselleştirilmesi açısından önemli bir altyapı sunmaktadır (İnt. Kyn. 6).

CesiumJS, WebGL tabanlı çalışan ve 3B mekânsal verilerin küresel koordinat sistemi üzerinde görselleştirilmesine olanak sağlayan açık kaynaklı bir JavaScript kütüphanesidir (İnt. Kyn. 7). Arazi verileri, görüntü katmanları, 3B modeller, nokta bulutları ve zamana bağlı konum verileri gibi farklı veri türlerinin aynı sahne içinde sunulabilmesi, CesiumJS'i web tabanlı 3B CBS ve dijital ikiz uygulamaları açısından önemli bir platform hâline getirmektedir.

CesiumJS ortamında temel olarak sahne, kamera, katmanlar ve veri kaynakları birlikte

çalışmaktadır. Sahne yapısı; arazi, binalar, 3B modeller ve diğer mekânsal unsurların görselleştirildiği etkileşimli ortamı ifade etmektedir. Kamera bileşeni, kullanıcının 3B sahne içinde farklı bakış açılarıyla gezinmesini sağlamaktadır. Veri kaynakları ise görüntü servisleri, arazi modelleri, 3D Tiles verileri ve zamana bağlı nesne konumları gibi farklı mekânsal içeriklerin sahneye eklenmesine olanak tanımaktadır (İnt. Kyn. 7).

CesiumJS'in önemli özelliklerinden biri, zamana bağlı değişen verilerin 3B sahne üzerinde görselleştirilebilmesidir. Bu özellik, hareketli nesnelerin konumlarının zaman ekseniyle birlikte izlenmesi, sensör verilerinin mekânsal bağlam içinde sunulması ve dijital ikiz ortamında dinamik süreçlerin temsil edilmesi açısından önem taşımaktadır. Bu nedenle CesiumJS, yalnızca statik 3B kent modellerinin gösterimi için değil, aynı zamanda video tabanlı nesne tespit sonuçlarından elde edilen konum bilgilerinin 3B dijital ikiz ortamında sunulması için de uygun bir görselleştirme altyapısı sağlamaktadır (Zhong vd. 2024).

Bu tez çalışmasında CesiumJS, İHA ve LiDAR verileriyle oluşturulan 3B dijital ikiz altlığının web ortamında sunulması ve sabit kamera görüntülerinden elde edilen nesne konumlarının gerçek dünya koordinatlarıyla birlikte görselleştirilmesi amacıyla değerlendirilmiştir. Böylece statik 3B model altlığı ile video tabanlı dinamik nesne bilgileri aynı web tabanlı ortamda bütünleştirilebilmektedir.

2.6.4 3D Tiles Veri Formatı ve Hiyerarşik Yapısı

Yüksek çözünürlüklü fotogrametrik modeller, LiDAR nokta bulutları ve 3B kent modelleri, veri hacmi bakımından oldukça büyük dosya boyutlarına ulaşabilmektedir. Bu tür verilerin web tarayıcısı üzerinden tek seferde yüklenmesi, bellek kullanımı, ağ bant genişliği ve grafik işleme performansı açısından önemli sınırlılıklar oluşturmaktadır. Bu nedenle büyük hacimli 3B mekânsal verilerin web ortamında parça parça yüklenmesini ve kullanıcı bakış açısına göre dinamik biçimde görüntülenmesini sağlayan hiyerarşik veri yapıları kullanılmaktadır.

3D Tiles, büyük ölçekli ve heterojen 3B mekânsal verilerin web, masaüstü ve mobil

ortamlarda akışkan biçimde sunulabilmesi amacıyla geliştirilen açık bir veri standardıdır (İnt. Kyn. 8). Fotogrametrik modeller, 3B binalar, nokta bulutları ve yapı modeli verileri gibi farklı veri türlerinin hiyerarşik bir yapı içinde sunulmasına olanak sağlamaktadır. 3D Tiles yapısında veri, tek parça hâlinde yüklenmek yerine küçük karo parçalarına ayrılmakta; kullanıcının görüş alanına, kamera konumuna ve ayrıntı ihtiyacına göre yalnızca gerekli karolar yüklenmektedir. Bu yaklaşım, web ortamında büyük hacimli 3B verilerin daha verimli biçimde görselleştirilmesine katkı sağlamaktadır.

3D Tiles standardının temelinde Hiyerarşik Ayrıntı Düzeyi (Hierarchical Level of Detail - HLOD) yaklaşımı yer almaktadır. Bu yaklaşımda, kameraya uzak bölgelerde daha düşük ayrıntı düzeyindeki karolar kullanılırken, kullanıcı modele yaklaştıkça daha ayrıntılı karolar yüklenmektedir. Böylece hem ağ üzerinden aktarılan veri miktarı hem de tarayıcıda işlenmesi gereken geometri yükü azaltılabilmektedir. Bu yapı, özellikle şehir ölçeğindeki 3B modellerin, yoğun nokta bulutlarının ve fotogrametrik ağ modellerin web tabanlı ortamlarda sunulmasında önemli bir performans avantajı sağlamaktadır.

3D Tiles veri yapısı genel olarak bir tileset dosyası ve bu dosyada tanımlanan karo hiyerarşisinden oluşmaktadır. Tileset dosyası, karoların konumunu, sınır hacimlerini, geometrik hata değerlerini ve alt karolarla olan ilişkilerini tanımlamaktadır. Karo hiyerarşisinin oluşturulmasında dörtlü ağaç, sekizli ağaç veya veri dağılımına bağlı farklı uzaysal bölme yaklaşımları kullanılabilmektedir. Dörtlü ağaç yapıları daha çok yüzey ve harita tabanlı bölme işlemlerinde, sekizli ağaç yapıları ise 3B hacimsel verilerin bölümlenmesinde kullanılabilmektedir.

3D Tiles standardı farklı 3B veri türlerinin temsil edilebilmesine olanak tanımaktadır. 3D Tiles 1.0 kapsamında b3dm, i3dm, pnts ve cmpt gibi içerik türleri kullanılmıştır. b3dm dokulu 3B modellerin, i3dm tekrar eden model örneklerinin, pnts nokta bulutlarının, cmpt ise birden fazla içerik türünün tek bir karo içinde birleştirilmesinin temsilinde kullanılmıştır. Güncel 3D Tiles yaklaşımlarında ise GL Transmission Format (glTF) tabanlı içerik yapısı öne çıkmakta ve 3B verilerin daha standart, verimli ve birlikte çalışabilir biçimde aktarılmasına olanak sağlamaktadır (İnt. Kyn. 8, İnt. Kyn. 9).

Bu tez çalışması kapsamında oluşturulan 3B dijital ikiz altlığının web ortamında verimli biçimde sunulabilmesi için 3D Tiles veri yapısından yararlanılması uygun görülmektedir. Böylece yüksek çözünürlüklü fotogrametrik model ve diğer 3B mekânsal veriler, CesiumJS ortamında kullanıcı bakış açısına bağlı olarak kademeli biçimde yüklenebilmekte ve video tabanlı nesne tespit sonuçlarıyla aynı 3B sahne içerisinde görselleştirilebilmektedir.

3. MATERYAL ve METOT

Bu bölümde, tez çalışması kapsamında kullanılan veri toplama araçları, çalışma alanı, donanım bileşenleri, veri işleme süreçleri ve uygulanan yöntemler açıklanmaktadır. Çalışmanın genel iş akışı; çalışma alanından İHA tabanlı fotogrametrik ve LiDAR verilerinin toplanması, bu verilerden 3B dijital ikiz altlığının oluşturulması, sabit kamera görüntülerinden YOLO tabanlı nesne tespiti yapılması, tespit edilen nesnelerin

piksel koordinatlarından gerçek dünya koordinatlarına dönüştürülmesi ve sonuçların CesiumJS ortamında görselleştirilmesi aşamalarından oluşmaktadır. Ayrıca, konum kestirimlerinde ortaya çıkan zamansal dalgalanmaların azaltılması amacıyla filtreleme yaklaşımları da değerlendirilmiştir.

3.1 Çalışma Alanı, Veri Setleri ve Donanım Bileşenleri

Bu tez çalışması kapsamında, kentsel ve kırsal özelliklerin bir arada bulunduğu bir bölgede 3B dijital ikiz altlığının oluşturulması ve video görüntülerinden elde edilen nesne tespit sonuçlarının bu dijital ikiz ortamı ile ilişkilendirilmesi amaçlanmıştır. Bu doğrultuda veri toplama sürecinde İHA platformları, fotogrametrik kamera sistemi, LiDAR sensörü ve sabit konumdan görüntü kaydı alabilen video kayıt cihazları kullanılmıştır.

Çalışma alanına ait mekânsal verilerin üretilmesinde İHA tabanlı eğik fotogrametri ve LiDAR ölçümlerinden yararlanılmıştır. Fotogrametrik görüntüler, çalışma alanının dokulu 3B modelinin ve ortomozaik ürünlerinin oluşturulmasında; LiDAR verileri ise nokta bulutu, arazi yüzeyi ve geometrik doğruluk değerlendirmelerinde kullanılmıştır. Sabit kamera görüntüleri ise hareketli nesnelerin tespiti, piksel koordinatlarının çıkarılması ve bu koordinatların gerçek dünya koordinatlarıyla ilişkilendirilmesi amacıyla değerlendirilmiştir.

Veri işleme ve analiz süreçleri; fotogrametrik değerlendirme, nokta bulutu işleme, 3B model üretimi, YOLO tabanlı nesne tespiti, koordinat dönüşümü, konumsal doğruluk analizi ve web tabanlı görselleştirme aşamalarını kapsamaktadır. Bu işlemler, yüksek hacimli görüntü ve nokta bulutu verilerinin işlenmesine olanak sağlayan bir iş istasyonu üzerinde gerçekleştirilmiştir. Derin öğrenme aşamasında ise önceden eğitilmiş YOLO model varyantları kullanılmış; modellerin çıkarım performansları ve tespit başarımları karşılaştırmalı olarak değerlendirilmiştir.

3.1.1 Çalışma Alanının Özellikleri

Araştırma sahası olarak Ankara ili Gölbaşı ilçesi sınırları içerisinde yer alan Dikilitaş bölgesi belirlenmiştir. Çalışma alanının seçiminde, kentsel ve kırsal özelliklerin bir arada bulunması, değişken arazi morfolojisine sahip olması ve farklı yapı yoğunluklarını içermesi etkili olmuştur. Bu özellikler, İHA tabanlı fotogrametrik veri üretimi, LiDAR nokta bulutu işleme, 3B model üretimi ve video tabanlı nesne konumlandırma süreçlerinin aynı çalışma alanı üzerinde birlikte değerlendirilmesine olanak sağlamaktadır. Çalışma alanı konumu ve sınırları Şekil 3.1'de sunulmuştur.



Şekil 3.1 Çalışma alanının konumu ve sınırları (Google My Maps üzerinde oluşturulmuştur).

Gerçekleştirilen İHA uçuşları sonucunda çalışma alanının yaklaşık 1,503 km² büyüklüğünde olduğu belirlenmiştir. Topoğrafik olarak değişken eğim değerlerine sahip olan bölgede, çoğunlukla az katlı ve düzensiz yerleşim düzenine sahip yapılar bulunmaktadır. Arazi eğimi, yapı dağılımı, yol ağı ve açık alanların birlikte yer alması; fotogrametrik model üretimi, LiDAR tabanlı arazi yüzeyi çıkarımı ve 3B dijital ikiz altlığının oluşturulması açısından uygun bir veri ortamı sağlamaktadır.

Çalışma alanındaki topoğrafik değişkenlik, görüntü koordinatlarından gerçek dünya koordinatlarına dönüşüm sürecinde de önem taşımaktadır. Düzlemsel yüzey varsayımına dayalı yaklaşımların sınırlılıklarının görülebilmesi, referans noktalarına dayalı konum kestirimi ve ray-casting gibi 3B yüzeyden yararlanan yöntemlerin değerlendirilmesi açısından bölgenin bu yapısı çalışmaya katkı sağlamaktadır. Bu nedenle Dikilitaş bölgesi, hem 3B dijital ikiz altlığı üretimi hem de video tabanlı nesne tespit sonuçlarının gerçek dünya koordinatlarıyla ilişkilendirilmesi için uygun bir çalışma alanı olarak değerlendirilmiştir.

3.1.2 Kullanılan İHA, LiDAR Sensörleri ve Sabit Kameralar

Çalışma kapsamında 3B dijital ikiz altlığının üretilmesi, nokta bulutu ve fotogrametrik model verilerinin işlenmesi, sabit kamera görüntülerinden nesne tespiti yapılması ve bu tespitlerin gerçek dünya koordinatlarıyla ilişkilendirilmesi amacıyla farklı veri toplama ve işlem donanımlarından yararlanılmıştır. Bu kapsamda İHA platformu, fotogrametrik kamera, LiDAR sensörü, sabit video kayıt cihazları ve veri işleme süreçlerinde kullanılan iş istasyonu temel donanım bileşenlerini oluşturmaktadır.

- **İHA platformu:** Fotogrametrik ve LiDAR veri toplama uçuşlarında DJI Matrice 300 RTK platformu kullanılmıştır. Uçuşlar sırasında platformun Gerçek Zamanlı Kinematik (Real-Time Kinematic - RTK) özelliğinden yararlanılmış ve görüntü/poz verilerinin büyük bölümü FIX (sabit) çözümle kaydedilmiştir. Fotogrametrik uçuşta 2311 görüntünün tamamı, LiDAR uçuşunda ise 1223 poz verisinin 1219'u FIX çözümle elde edilmiştir. Uçuşlar yaklaşık 137–153 m aralığındaki ortalama irtifalarda ve yaklaşık 7,0–7,3 m/s uçuş hızlarında gerçekleştirilmiştir.
- **Fotogrametrik kamera sistemi:** Çalışma alanına ait yüksek çözünürlüklü ortomozaik, SYM ve dokulu 3B model üretimi amacıyla DJI Zenmuse P1 tam kare kamera kullanılmıştır. Smart Oblique uçuş stratejisiyle gerçekleştirilen veri toplama sürecinde toplam 2311 adet görüntü elde edilmiştir. Fotogrametrik değerlendirme sonucunda ortalama YÖA değeri 4,44 cm olarak hesaplanmış, blok dengelemesi sonrasında fotogrametrik modelin konumsal Karesel Ortalama Hata (Root Mean Square Error - RMSE) değeri 0,031 m olarak elde edilmiştir.

- **LiDAR sensörü:** Çalışma alanındaki arazi yüzeyinin, yapı geometrilerinin ve nokta bulutu tabanlı 3B temsilin elde edilmesi amacıyla DJI Zenmuse L2 LiDAR sensörü kullanılmıştır. LiDAR uçuşunda sensör 240 kHz darbe hızı ve 1200 Hz tarama hızıyla çalıştırılmıştır. Elde edilen nokta bulutu verilerinde çalışma alanı genelindeki ortalama nokta yoğunluğu 205,3 nokta/m² olarak hesaplanmıştır. LiDAR verilerinin georeferanslama süreci sonucunda RMSE değeri 0,056 m olarak raporlanmıştır.

- **Sabit video kayıt cihazları:** Video tabanlı nesne tespiti ve konum kestirimi aşamalarında, çalışma alanındaki yol ve açık alanları görecek şekilde sabitlenen mobil cihaz kameralarından yararlanılmıştır. Bu cihazlar tripod yardımıyla sabitlenerek sahadaki sabit kamera kurulumunu temsil edecek biçimde kullanılmıştır. Analiz gereksinimlerine bağlı olarak 4K ve 1080p çözünürlüklerde video kayıtları alınmış; bu görüntüler yaya ve araç gibi hareketli nesnelerin tespiti, piksel koordinatlarının çıkarılması ve gerçek dünya koordinatlarına dönüştürülmesi süreçlerinde kullanılmıştır.

- **İş istasyonu ve işlem altyapısı:** Fotogrametrik değerlendirme, LiDAR veri işleme, 3B model üretimi, YOLO tabanlı nesne tespiti, koordinat dönüşümü ve görselleştirme işlemleri bir iş istasyonu üzerinde gerçekleştirilmiştir. Kullanılan sistemde Intel Core i7-8700K işlemci, NVIDIA GeForce GTX 1080 grafik kartı ve 64 GB RAM bulunmaktadır. Bu donanım altyapısı; DJI Terra yazılımında fotogrametrik ve LiDAR veri işleme süreçlerinin yürütülmesi, Python ortamında nesne tespiti ve koordinat dönüşümü işlemlerinin gerçekleştirilmesi ve web tabanlı görselleştirme uygulamalarının geliştirilmesi için kullanılmıştır.

Çalışmada kullanılan temel donanım ve veri toplama bileşenleri Çizelge 3.1'de özetlenmiştir.

Çizelge 3.1 Çalışmada kullanılan temel donanım ve veri toplama bileşenleri.

Bileşen	Kullanılan sistem / model	Çalışmadaki kullanım amacı

İHA platformu	DJI Matrice 300 RTK	Fotogrametrik ve LiDAR veri toplama uçuşlarının gerçekleştirilmesi
Fotogrametrik kamera	DJI Zenmuse P1	Smart Oblique görüntü alımı, ortomozaik, SYM ve dokulu 3B model üretimi
LiDAR sensörü	DJI Zenmuse L2	Nokta bulutu üretimi, arazi yüzeyi çıkarımı ve geometrik 3B temsil
Sabit video kayıt cihazları	Tripodla sabitlenen mobil cihaz kameraları	Video tabanlı nesne tespiti, piksel koordinatı çıkarımı ve konum kestirimi
İş istasyonu	Intel Core i7-8700K, NVIDIA GTX 1080, 64 GB RAM	Fotogrametrik değerlendirme, LiDAR işleme YOLO çıkarımı, koordinat dönüşümü ve görselleştirme işlemleri

3.1.3 Referans Veriler, Koordinat Sistemi ve RTK Ölçümleri

Çalışmada üretilen fotogrametrik model, LiDAR nokta bulutu, kamera tabanlı konum kestirimleri ve dijital ikiz görselleştirmeleri ortak bir koordinat referans sistemi içerisinde değerlendirilmiştir. Bu kapsamda çalışma alanına ait mekânsal veriler World Geodetic System 1984 (WGS 84) / Universal Transverse Mercator (UTM) Zone 36N koordinat sisteminde işlenmiş ve analiz edilmiştir. Böylece farklı veri kaynaklarından elde edilen konum bilgilerinin aynı metrik düzlemde karşılaştırılabilmesi sağlanmıştır.
Konumsal doğruluk değerlendirmelerinde referans veri olarak RTK ölçümleri kullanılmıştır. Video görüntülerinde tespit edilen yaya konumlarının gerçek dünya koordinatlarıyla karşılaştırılabilmesi amacıyla, sahada belirlenen örnek noktalarda RTK yöntemiyle koordinat ölçümleri gerçekleştirilmiştir. Bu ölçümler, görüntü tabanlı konum kestirimlerinden elde edilen sonuçların doğruluğunu değerlendirmek için referans değer olarak kabul edilmiştir.

Kamera görüntülerinden elde edilen piksel koordinatları; homografi, lokal enterpolasyon ve ray-casting yaklaşımlarıyla gerçek dünya koordinatlarına

dönüştürülmüştür. Elde edilen tahmini koordinatlar, RTK ölçümleriyle aynı koordinat sisteminde değerlendirilerek karşılaştırılmıştır. Hata analizlerinin metre cinsinden yapılabilmesi için coğrafi koordinatlar gerektiğinde UTM düzlemine dönüştürülmüş; yatay konum hataları, tahmini konum ile RTK referans konumu arasındaki iki boyutlu mesafe üzerinden hesaplanmıştır.

Bu yaklaşım sayesinde, farklı koordinat dönüşüm yöntemlerinin konumsal başarımı aynı referans veri üzerinden karşılaştırmalı olarak değerlendirilebilmiştir. Ayrıca çalışma alanındaki topoğrafik değişkenliğin görüntü tabanlı konum kestirimi üzerindeki etkisinin incelenebilmesi için RTK referans verileri, homografi, lokal enterpolasyon ve ray-casting yöntemlerinin doğruluk analizlerinde temel doğrulama verisi olarak kullanılmıştır.

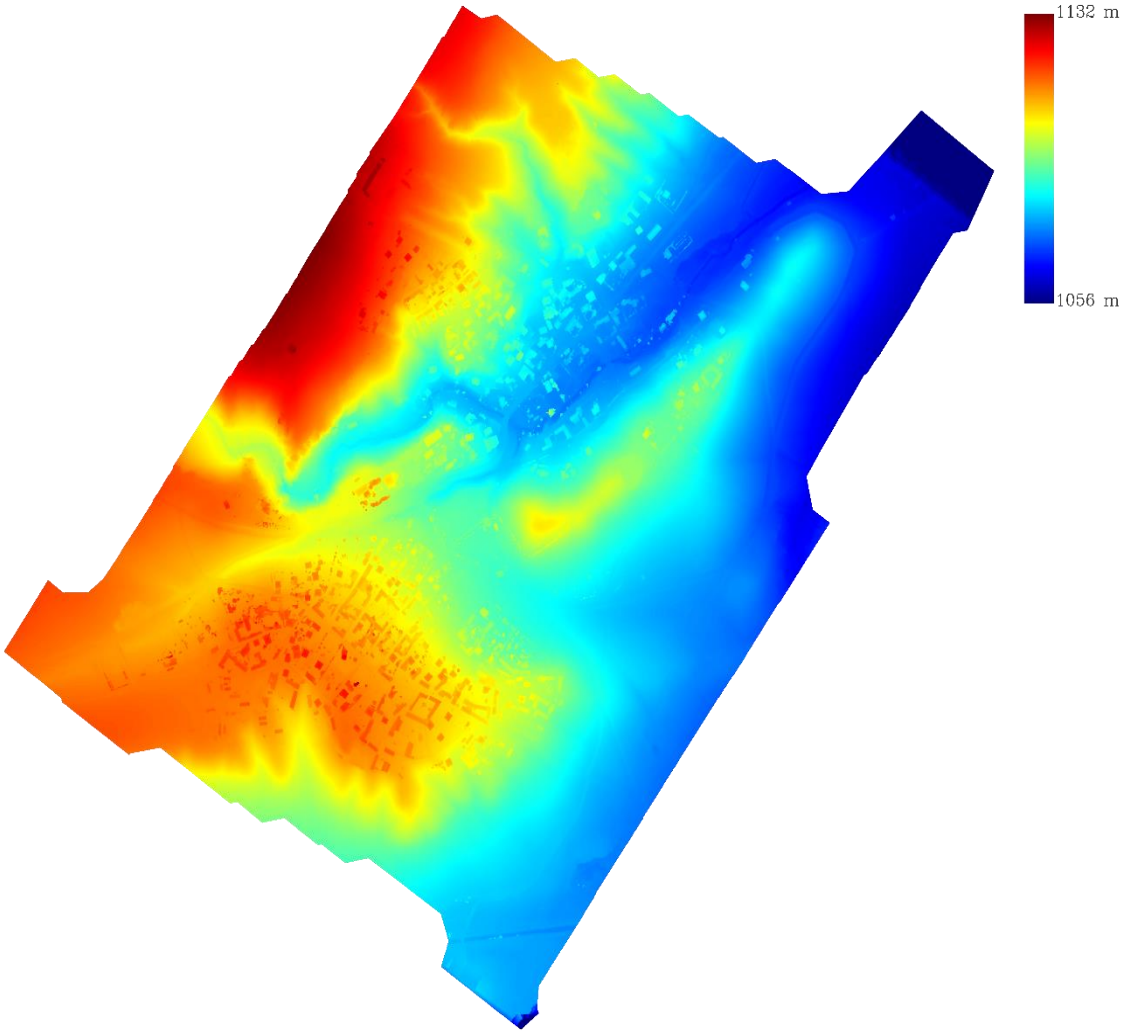
3.2 3B Dijital İkiz Altlığının Üretilmesi

Bu bölümde, çalışma alanından elde edilen fotogrametrik görüntüler ve LiDAR verileri kullanılarak 3B dijital ikiz altlığının oluşturulmasına yönelik veri işleme süreci açıklanmaktadır. Bu kapsamda Smart Oblique uçuş verilerinin değerlendirilmesi, LiDAR nokta bulutunun işlenmesi, fotogrametrik ve LiDAR verilerinin aynı koordinat sistemi içerisinde değerlendirilmesi, 3B mesh modelin üretilmesi ve web tabanlı görselleştirmeye uygun 3D Tiles veri yapısına dönüştürülmesi aşamaları ele alınmıştır.
Üretilen 3B altlık model, çalışma alanının topoğrafik yapısını, yapı geometrilerini ve görsel dokusunu temsil eden temel mekânsal veri katmanı olarak değerlendirilmiştir. Bu altlık, sonraki aşamalarda sabit kamera görüntülerinden elde edilen nesne tespit sonuçlarının gerçek dünya koordinatlarıyla ilişkilendirilmesi ve CesiumJS ortamında görselleştirilmesi için referans mekânsal zemin olarak kullanılmıştır.

3.2.1 Smart Oblique (Akıllı Eğik) Uçuş ile Veri Toplama

Geleneksel hava fotogrametrisinde yaygın olarak kullanılan nadir görüntüleme yaklaşımı, özellikle çatı ve yatay yüzeylerin modellenmesinde etkili olmakla birlikte, bina cepheleri ve düşey yüzeylere ait ayrıntıların temsilinde sınırlı kalabilmektedir. Bu çalışmada, yapıların düşey yüzeylerinin daha ayrıntılı temsil edilebilmesi ve çalışma

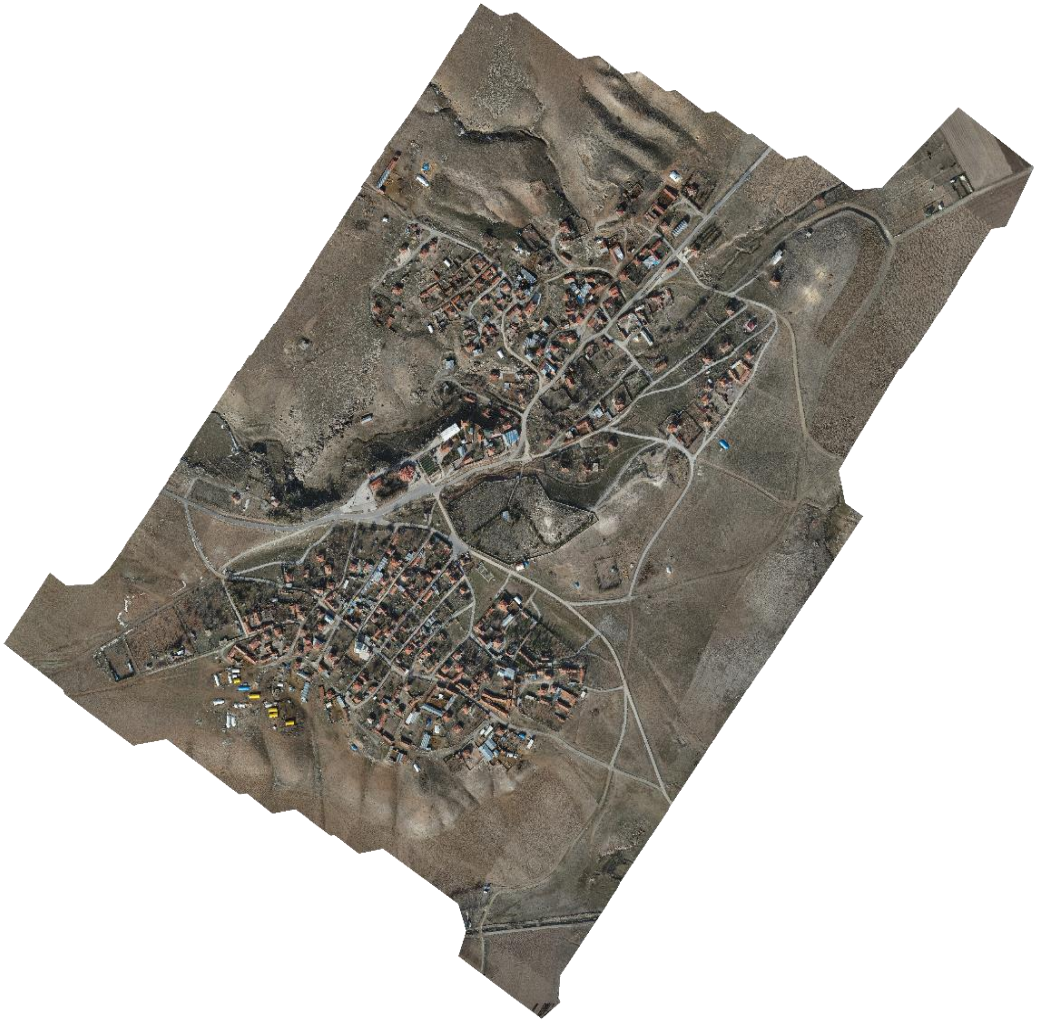
alanına ait dokulu 3B modelin üretilebilmesi amacıyla DJI Matrice 300 RTK platformuna entegre DJI Zenmuse P1 kamera ile Smart Oblique uçuş stratejisi uygulanmıştır. Bu uçuş sonucunda çalışma alanına ait fotogrametrik görüntüler elde edilmiş ve bu verilerden SYM, ortomozaik ve dokulu 3B model üretiminde yararlanılmıştır. Smart Oblique uçuş verilerinden üretilen SYM Şekil 3.2'de gösterilmiştir.



Şekil 3.2 Çalışma alanında gerçekleştirilen Smart Oblique uçuşu sonucunda üretilen SYM.

Ortalama 152 m uçuş irtifasında gerçekleştirilen fotogrametrik veri toplama süreci sonucunda, yaklaşık 1,503 km² büyüklüğündeki çalışma alanından toplam 2311 adet görüntü elde edilmiştir. Uçuş süresince RTK destekli konumlandırma kullanılmış ve tüm görüntüler FIX çözümle kaydedilmiştir. Böylece görüntülerin dış yöneltme parametrelerinin belirlenmesinde kullanılacak konum bilgileri yüksek doğruluklu olarak elde edilmiştir. Fotogrametrik değerlendirme süreci DJI Terra yazılımı kullanılarak

gerçekleştirilmiştir. Görüntü hizalama, hava üçgenlemesi, yoğun nokta bulutu üretimi ve yüzey modelleme işlemleri sonucunda çalışma alanına ait ortalama YÖA 4,44 cm olarak hesaplanmıştır. Bu çözünürlük değeri, çalışma alanındaki yapı cepheleri, yol yüzeyleri ve arazi dokularının 3B model üzerinde temsil edilmesi açısından uygun bir veri altyapısı sağlamıştır. Smart Oblique uçuş verilerinden üretilen ortomozaik görüntü Şekil 3.3'te verilmiştir.

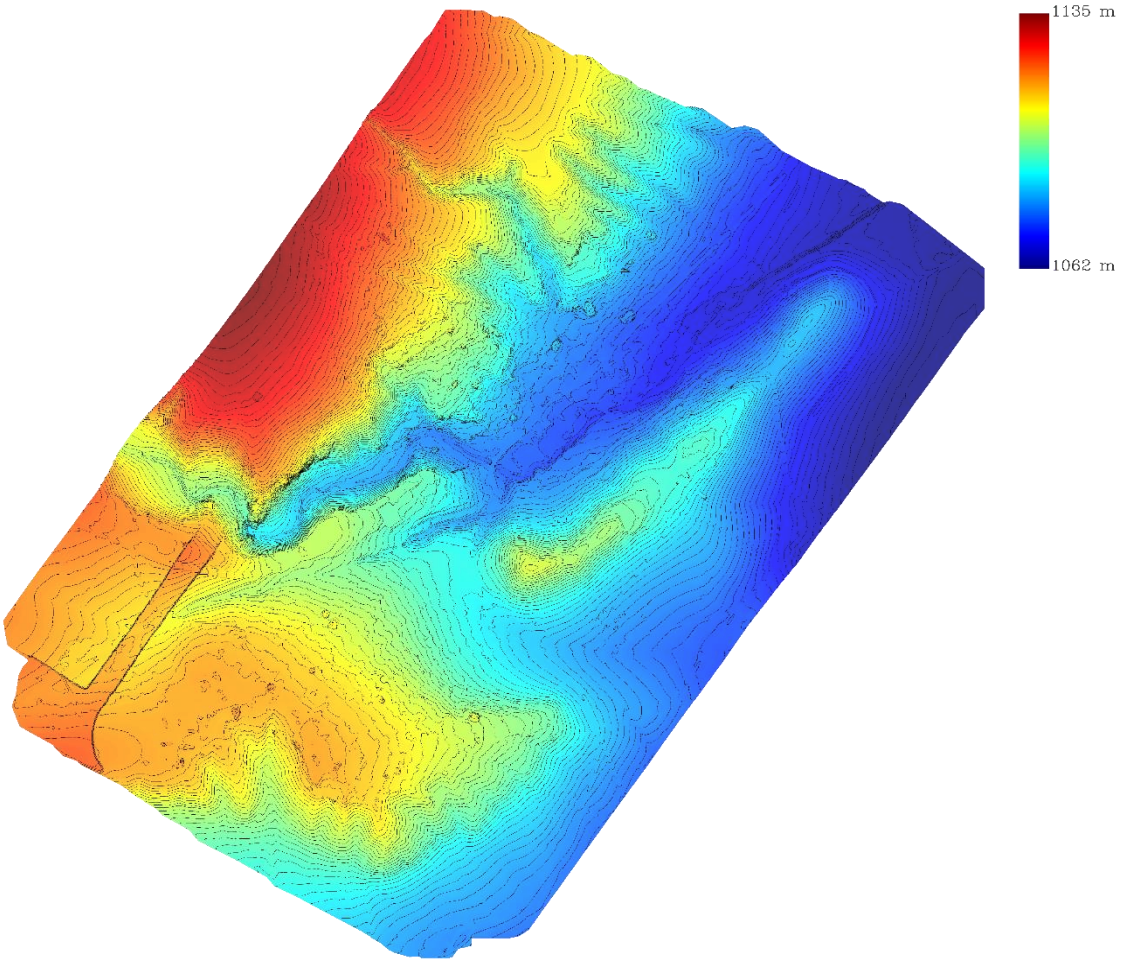


Şekil 3.3 Smart Oblique uçuş verilerinden üretilen çalışma alanının dijital ortomozaik görüntüsü

3.2.2 LiDAR Verisi, Veri Entegrasyonu ve 3D Tiles Üretimi

Fotogrametrik yöntemler, homojen dokuya sahip yüzeylerde veya yoğun bitki örtüsünün bulunduğu alanlarda görüntü eşleme sürecine bağlı olarak eksik ya da gürültülü veri üretebilmektedir. Bu nedenle çalışma kapsamında fotogrametrik veriler, DJI Zenmuse L2 sensöründen elde edilen LiDAR verileriyle birlikte değerlendirilmiştir.

LiDAR verileri, arazi yüzeyinin temsil edilmesi, nokta bulutu tabanlı geometrik kontrolün sağlanması ve fotogrametrik modelin desteklenmesi amacıyla kullanılmıştır. LiDAR veri toplama sürecinde toplam 1223 poz verisi elde edilmiştir. Sensör, 240 kHz darbe hızı ve 1200 Hz tarama hızıyla çalıştırılmıştır. Üretilen nokta bulutu verileri üzerinde yapılan değerlendirmede, çalışma alanı genelindeki ortalama nokta yoğunluğu 205,3 nokta/m^2 olarak hesaplanmıştır. LiDAR nokta bulutu verilerinden elde edilen renk kodlu Sayısal Arazi Modeli (SAM) ve eşyükselti eğrileri Şekil 3.4'te gösterilmiştir.



Şekil 3.4 LiDAR nokta bulutu verilerinden elde edilen renk kodlu SAM ve eşyükselti eğrileri.

Fotogrametrik görüntülerden üretilen model ve LiDAR nokta bulutu, WGS 84 / UTM Zone 36N koordinat sisteminde değerlendirilmiştir. İHA platformundan elde edilen RTK/GNSS ve IMU verileri kullanılarak her iki veri setinin aynı koordinat referans sistemi içerisinde konumlandırılması sağlanmıştır. Fotogrametrik model için konumsal RMSE değeri 0,031 m, LiDAR verisi için georeferanslama RMSE değeri ise 0,056 m olarak elde edilmiştir. Bu sayede fotogrametrik verilerin sağladığı görsel doku bilgisi ile

LiDAR verilerinin sağladığı nokta bulutu tabanlı geometrik bilgi aynı çalışma alanı üzerinde birlikte değerlendirilebilmiştir.

Araştırma kapsamında fotogrametrik ve LiDAR verileri tek bir modele indirgenmemiş; aynı koordinat sisteminde konumlandırılmış iki tamamlayıcı veri katmanı olarak kullanılmıştır. Fotogrametrik verilerden çalışma alanına ait dokulu 3B mesh model üretilmiş, LiDAR verileri ise yoğun nokta bulutu ve arazi yüzeyi bilgisi açısından değerlendirilmiştir. Bu yaklaşım, hem görsel temsil hem de geometrik kontrol açısından dijital ikiz altlığının daha bütüncül biçimde oluşturulmasına katkı sağlamıştır.

Üretilen 3B mesh model ve nokta bulutu verilerinin web ortamında verimli biçimde sunulabilmesi için veriler 3D Tiles formatına dönüştürülmüştür. Büyük hacimli 3B verilerin web tarayıcısında tek seferde yüklenmesi bellek, ağ bant genişliği ve grafik işleme performansı açısından sınırlılıklar oluşturabileceğinden, 3D Tiles veri yapısından yararlanılmıştır. Bu yapı, 3B modeli mekânsal olarak daha küçük karo parçalarına ayırmakta ve kullanıcı bakış açısına göre yalnızca gerekli ayrıntı düzeyindeki karoların yüklenmesine olanak sağlamaktadır.

CesiumJS ortamında 3D Tiles verileri, kullanıcının kamera konumu, görüş alanı ve modele olan uzaklığına bağlı olarak kademeli biçimde yüklenebilmektedir. Bu sayede yüksek çözünürlüklü fotogrametrik model ve LiDAR nokta bulutu web tabanlı 3B sahnede daha verimli biçimde görselleştirilebilmiştir. Oluşturulan bu 3B dijital ikiz altlığı, sonraki aşamalarda YOLO tabanlı nesne tespit sonuçlarının gerçek dünya koordinatlarıyla birlikte CesiumJS ortamında gösterilmesi için temel mekânsal referans katmanı olarak kullanılmıştır.

3.3 Derin Öğrenme ile Nesne Tespiti Uygulaması

Bu tez çalışmasında, çalışma alanını gören sabit kamera görüntülerinden hareketli nesnelerin tespit edilmesi ve bu tespitlerden elde edilen piksel koordinatlarının 3B dijital ikiz ortamında kullanılabilecek konum bilgilerine dönüştürülmesi amaçlanmıştır. Bu kapsamda bilgisayarlı görü yöntemlerinden yararlanılarak video görüntüleri

üzerinde YOLO tabanlı nesne tespiti uygulanmıştır.

Nesne tespiti aşamasında, önceden eğitilmiş YOLO model varyantları kullanılmış ve modellerin çalışma alanındaki görüntüler üzerindeki tespit başarımı karşılaştırmalı olarak değerlendirilmiştir. Model değerlendirme sürecinde hem sayım tabanlı ön değerlendirme hem de etiketli sınırlayıcı kutular üzerinden IoU tabanlı doğrulama yaklaşımı uygulanmıştır. Böylece modeller yalnızca nesne sayısını doğru tahmin etme başarımı açısından değil, aynı zamanda sınırlayıcı kutu konum doğruluğu ve çıkarım süresi bakımından da incelenmiştir. Nesne tespiti sonucunda üretilen sınıf bilgisi, güven skoru, Sınırlayıcı Kutu (Bounding Box - BBox) koordinatları, merkez noktası ve alt orta nokta bilgilerini içeren çıktı dosyası veri yapısı Ek 1'de verilmiştir.

Geliştirme ve analiz süreçleri Python programlama dili kullanılarak yürütülmüş; video karelerinin okunması, görüntü işleme adımları ve tespit sonuçlarının görselleştirilmesi gibi işlemlerde Open Source Computer Vision Library (OpenCV) kütüphanesinden yararlanılmıştır. Video karelerinde tespit edilen yaya ve araç gibi nesnelere ait sınırlayıcı kutuların merkez ve alt orta noktaları hesaplanarak, koordinat dönüşüm aşamasında kullanılacak piksel referans noktaları oluşturulmuştur.

3.3.1 Kullanılan YOLO Model Varyantları ve Model Seçim Yaklaşımı

Çalışma kapsamında, video görüntülerinden yaya ve araç gibi hareketli nesnelerin tespit edilmesi amacıyla YOLO tabanlı nesne tespit modellerinden yararlanılmıştır. YOLO mimarileri, görüntünün tamamını tek bir ileri geçişte değerlendirerek sınıf olasılıkları ve sınırlayıcı kutu koordinatlarını aynı anda üretebildiğinden, video tabanlı uygulamalarda hız ve doğruluk dengesi açısından uygun bir yaklaşım sunmaktadır.
Bu tez çalışmasının temel amacı yeni bir derin öğrenme modeli geliştirmek değil, önceden eğitilmiş nesne tespit modellerinden elde edilen sonuçların gerçek dünya koordinatlarıyla ilişkilendirilerek 3B dijital ikiz ortamında görselleştirilmesidir. Bu nedenle çalışma kapsamında sıfırdan model eğitimi yapılmamış, COCO veri seti üzerinde önceden eğitilmiş YOLO model ağırlıklarından yararlanılmıştır. COCO veri setinde yer alan person ve car sınıfları, çalışma kapsamında izlenmesi hedeflenen yaya

ve araç nesneleriyle doğrudan ilişkili olduğundan, hazır model ağırlıkları uygulama için uygun bir başlangıç altyapısı sağlamaktadır.

Model seçimi sürecinde yalnızca tek bir YOLO modeli kullanılmamış; YOLOv8 ve YOLO26 ailelerine ait farklı ölçeklerdeki model varyantları karşılaştırmalı olarak değerlendirilmiştir. Bu kapsamda nano, small, medium, large ve xlarge ölçeklerindeki modeller; tespit başarımı, sınırlayıcı kutu konum doğruluğu, güven skoru davranışı ve çıkarım süresi bakımından incelenmiştir. Böylece model seçimi yalnızca doğruluk değerine göre değil, aynı zamanda gerçek zamanlıya yakın dijital ikiz entegrasyonu için gerekli hız-doğruluk dengesi dikkate alınarak yapılmıştır.

Önceden eğitilmiş model ağırlıkları kullanıldığından, bu çalışma kapsamında projeye özgü yeni bir eğitim veri seti oluşturularak model eğitimi gerçekleştirilmemiştir. Bununla birlikte, modellerin çalışma alanındaki başarımını değerlendirmek amacıyla seçilen test görüntüleri üzerinde manuel sayım ve etiketli sınırlayıcı kutu doğrulaması yapılmıştır. Bu doğrulama süreci, model eğitimi için değil; kullanılan hazır modellerin çalışma alanındaki tespit ve konumlandırma başarımını değerlendirmek amacıyla uygulanmıştır.

3.3.2 Test Görüntüleri, Mesafe/Görünürlük Sınıfları ve Etiketleme Süreci

YOLO tabanlı nesne tespit modellerinin çalışma alanındaki başarımını değerlendirmek amacıyla sabit kamera görüntülerinden temsilî test görüntüleri seçilmiştir. Test görüntülerinin seçiminde, yaya ve araç gibi hedef nesnelerin farklı uzaklık, görünürlük ve sahne koşullarında yer almasına dikkat edilmiştir. Böylece model performansının yalnızca genel tespit başarımı açısından değil, aynı zamanda sahne koşullarına bağlı değişimi bakımından da incelenmesi amaçlanmıştır.

İlk değerlendirme aşamasında, çalışma alanından elde edilen 12 adet temsilî görüntü kullanılmıştır. Bu görüntülerde yer alan yaya nesneleri manuel olarak sayılmış ve her görüntü için gerçek yaya sayısı belirlenmiştir. Test görüntüleri mesafe bakımından yakın, orta ve uzak; görünürlük bakımından ise açık ve örtülü sahneler olarak

sınıflandırılmıştır. Bu sınıflandırma, model varyantlarının küçük, uzak veya kısmen örtülü hedefler üzerindeki davranışının karşılaştırılabilmesi amacıyla yapılmıştır.

Sayım tabanlı ön değerlendirmede, her modelin görüntü üzerinde ürettiği yaya tespit sayısı, manuel olarak belirlenen gerçek yaya sayısı ile karşılaştırılmıştır. Bu aşamada amaç, model varyantlarının çalışma alanına ilk uygunluğunu ve farklı sahne koşullarındaki genel tespit eğilimlerini belirlemektir. Sayım tabanlı değerlendirme, sınırlayıcı kutuların gerçek nesneyle örtüşme düzeyini doğrudan göstermediğinden, nihai model seçimi için tek başına yeterli kabul edilmemiştir.

Bu nedenle ikinci aşamada, model tahminlerinin sınırlayıcı kutu konum doğruluğunu değerlendirmek amacıyla manuel etiketleme işlemi yapılmıştır. Seçilen test görüntülerindeki nesneler etiketleme aracı kullanılarak sınırlayıcı kutularla işaretlenmiş ve etiketler YOLO formatında kaydedilmiştir. Etiketleme sürecinde person, car, bus ve truck sınıfları tanımlanmıştır. Ancak test görüntülerinde yer alan nesne dağılımına bağlı olarak değerlendirmeler ağırlıklı olarak person ve car sınıfları üzerinden gerçekleştirilmiştir.

Etiketli veri seti kullanılarak yapılan doğrulama aşamasında, model tahminleri ile manuel etiketler Kesişimin Birleşime Oranı (Intersection over Union - IoU) ölçütü üzerinden eşleştirilmiştir. Bir tahminin doğru kabul edilebilmesi için hem sınıf bilgisinin doğru olması hem de tahmin edilen sınırlayıcı kutunun (bounding box - BBox) gerçek kutu ile belirlenen IoU eşik değerinin üzerinde örtüşme sağlaması beklenmiştir. Bu yaklaşım sayesinde model varyantları, yalnızca nesne sayısını doğru tahmin etme başarımı bakımından değil, aynı zamanda sınırlayıcı kutu konum doğruluğu bakımından da karşılaştırılmıştır.

3.3.3 Çıkarım Parametreleri ve Model Çalıştırma Ayarları

YOLO model varyantlarının video görüntüleri ve test görüntüleri üzerindeki tespit başarımını değerlendirmek amacıyla çıkarım işlemleri aynı donanım ve yazılım ortamında gerçekleştirilmiştir. Modellerin karşılaştırılabilir sonuçlar üretmesi için test görüntüleri, giriş çözünürlüğü, confidence (güven) eşik değerleri ve değerlendirme

koşulları sabit tutulmuş; her modelin çıktıları sınıf etiketi, güven skoru ve sınırlayıcı kutu koordinatları üzerinden kaydedilmiştir.

Çıkarım işlemlerinde, model tarafından üretilen düşük güven skoruna sahip tahminlerin elenebilmesi amacıyla confidence eşik değeri kullanılmıştır. Bu çalışmada model davranışını değerlendirmek için farklı confidence eşikleri incelenmiş; özellikle 0,30 ve 0,50 eşik değerlerinin tespit başarımı üzerindeki etkisi karşılaştırılmıştır. Daha düşük confidence eşiği, küçük, uzak veya kısmen örtülü nesnelerin tespit edilme olasılığını artırabilirken, yanlış pozitif tahminlerin artmasına da neden olabilmektedir. Daha yüksek eşik değeri ise daha güvenilir tahminleri korumakta, ancak bazı gerçek nesnelerin tespit dışı kalmasına yol açabilmektedir.

Aynı nesne üzerinde birden fazla örtüşen tahmin üretilmesini azaltmak amacıyla YOLO çıkarım sürecinde NMS işlemi uygulanmıştır. Bu işlem sayesinde, aynı nesneye ait yüksek oranda çakışan sınırlayıcı kutular arasından güven skoru daha yüksek olan tahmin korunmakta ve tekrar eden tespitlerin etkisi azaltılmaktadır.

Model varyantlarının hız bakımından karşılaştırılabilmesi için çıkarım süreleri görüntü başına milisaniye cinsinden kaydedilmiştir. Elde edilen çıkarım süreleri kullanılarak modellerin saniyedeki kare işleme kapasitesi hesaplanmış ve hız-doğruluk dengesi değerlendirilmiştir. Böylece model seçimi yalnızca tespit doğruluğuna göre değil, aynı zamanda video tabanlı dijital ikiz uygulamasında kullanılabilirlik açısından da ele alınmıştır.

Çıkarım işlemleri Python ortamında yürütülmüş; video karelerinin okunması, görüntü işleme adımları ve tespit sonuçlarının görselleştirilmesi gibi işlemlerde OpenCV kütüphanesinden yararlanılmıştır. Model çalıştırma ve değerlendirme süreçleri, 3.1.2'de teknik özellikleri verilen iş istasyonu üzerinde gerçekleştirilmiştir.

3.3.4 Video Akışı Üzerinden Tespit ve Piksel Referans Noktalarının Çıkarılması

Önceden eğitilmiş YOLO model varyantları, sabit kamera görüntülerinden elde edilen

4K ve 1080p çözünürlüklü video kayıtları üzerinde çalıştırılmıştır. Python ortamında yürütülen işlem akışında, video kayıtlarından ardışık kareler okunmuş ve her bir kare model girişine aktarılmıştır. Model çıktısı olarak tespit edilen nesnelere ait sınıf etiketi, güven skoru ve sınırlayıcı kutu koordinatları elde edilmiştir. Video görüntüsü üzerinde gerçekleştirilen nesne tespit sürecine ait örnek çıktı Şekil 3.5'te gösterilmiştir.

Şekil 3.5 Önceden eğitilmiş YOLO modeli kullanılarak video akışı üzerinde gerçekleştirilen gerçek zamanlı nesne tespiti.

Tespit edilen nesnelerin 3B dijital ikiz ortamında konumlandırılabilmesi için, sınırlayıcı kutulardan piksel referans noktaları çıkarılmıştır. Bu çalışmada iki farklı piksel referans noktası değerlendirilmiştir. Birincisi sınırlayıcı kutunun geometrik merkezi, ikincisi ise nesnenin zemin temas noktasına daha yakın olduğu kabul edilen alt orta noktadır. Bu iki noktanın birlikte değerlendirilmesi, piksel referans noktası seçiminin gerçek dünya koordinatlarına dönüşüm üzerindeki etkisini incelemek amacıyla yapılmıştır.
YOLO algoritması tarafından üretilen sınırlayıcı kutunun sol üst (x_{min} , y_{min}) ve sağ alt (x_{max} , y_{max}) köşe koordinatları kullanılarak aşağıdaki bağıntılarla hesaplanmıştır:

$$u_c = \frac{(x_{min} + x_{max})}{2} \tag{3.1}$$

$$v_c = \frac{y_{min} + y_{max}}{2} \tag{3.2}$$

Sınırlayıcı kutunun alt orta noktası ise Eşitlik (3.3) ve Eşitlik (3.4) ile hesaplanmıştır:

$$u_b = \frac{x_{min} + x_{max}}{2} \tag{3.3}$$

$$v_b = y_{max} \tag{3.4}$$

Burada u_c ve v_c, sınırlayıcı kutu merkezinin piksel koordinatlarını; u_b ve v_b ise sınırlayıcı kutunun alt orta noktasına ait piksel koordinatlarını ifade etmektedir. Alt orta

nokta, özellikle yaya ve araç gibi nesnelerde zeminle temas noktasına daha yakın bir temsil noktası olarak kabul edilmiştir. Bu nedenle koordinat dönüşüm analizlerinde hem merkez noktası hem de alt orta nokta kullanılarak, gerçek dünya konum doğruluğu üzerindeki etkileri karşılaştırılmıştır.

Elde edilen piksel koordinatları, sonraki aşamada homografi, lokal enterpolasyon ve ray-casting yöntemleriyle gerçek dünya koordinatlarına dönüştürülmek üzere kaydedilmiştir. YOLO tespit çıktılarından BBox merkez ve alt orta referans noktalarının çıkarılmasına ilişkin örnek kod kesiti Ek 10'da verilmiştir.

3.4 Piksel Koordinatlarından Gerçek Dünya Koordinatlarına Dönüşüm

YOLO tabanlı nesne tespiti sonucunda elde edilen hareketli nesne konumları, öncelikle video karesi üzerindeki iki boyutlu piksel koordinatları ile temsil edilmektedir. Ancak bu piksel koordinatları, nesnenin görüntü üzerindeki konumunu göstermekte; nesnenin gerçek dünyadaki coğrafi karşılığını doğrudan ifade etmemektedir. Bu nedenle tespit edilen nesnelerin 3B dijital ikiz ortamında doğru konumlandırılabilmesi için piksel koordinatlarının gerçek dünya koordinatlarıyla ilişkilendirilmesi gerekmektedir.

Bu çalışmada piksel koordinatlarından gerçek dünya koordinatlarına dönüşüm süreci üç farklı yaklaşım üzerinden ele alınmıştır. İlk olarak, görüntü düzlemi ile yer düzlemi arasında ilişki kurmak amacıyla referans noktalarına dayalı homografi matrisi hesaplanmıştır. İkinci olarak, görüntü üzerinde bilinen referans noktalarından yararlanılarak lokal enterpolasyon yaklaşımı uygulanmıştır. Üçüncü olarak ise kamera parametreleri ve 3B dijital ikiz yüzeyi kullanılarak kamera ışını ile 3B yüzey kesişimine dayalı ray-casting yaklaşımı değerlendirilmiştir.

Elde edilen konum kestirimleri, WGS 84 / UTM Zone 36N koordinat sistemi içerisinde değerlendirilmiş ve RTK ölçümleriyle karşılaştırılarak konumsal doğruluk analizleri gerçekleştirilmiştir. Bu sayede farklı koordinat dönüşüm yaklaşımlarının, çalışma alanındaki topoğrafik yapı ve kamera görüş koşulları altında ürettiği konum sonuçları karşılaştırmalı olarak incelenmiştir.

3.4.1 Referans Noktaları ve Homografi Matrisi Hesabı

Görüntü düzlemi ile yer düzlemi arasındaki ilişki, çalışma alanının yaklaşık düzlemsel kabul edilebildiği yerel bölgelerde homografi yaklaşımı ile ifade edilebilmektedir. Homografi, iki düzlem üzerindeki karşılıklı nokta çiftleri arasındaki projektif dönüşümü tanımlayan 3 × 3 boyutlu bir dönüşüm matrisidir. Bu çalışmada homografi yaklaşımı, sabit kamera görüntülerinde belirlenen piksel koordinatlarının UTM düzlemindeki gerçek dünya koordinatlarıyla ilişkilendirilmesi amacıyla kullanılmıştır.

Homografi matrisinin hesaplanabilmesi için, kamera görüş alanı içerisinde yer alan ve çalışma alanında konumu belirlenebilen sabit detaylar referans noktası olarak seçilmiştir. Bu kapsamda yol çizgisi kesişimleri, kaldırım köşeleri, rögar kapakları ve benzeri görüntü üzerinde ayırt edilebilen sabit detaylardan yararlanılmıştır. Referans noktalarının görüntü düzlemindeki piksel koordinatları (u, v) ile UTM düzlemindeki Easting ve Northing koordinatları (E, N) eşleştirilmiştir. Homografi matrisinin sekiz serbestlik derecesine sahip olması nedeniyle çözüm için en az dört adet aynı doğru üzerinde bulunmayan nokta eşleşmesine ihtiyaç duyulmaktadır. Noktaların çalışma alanına dengeli biçimde dağılması, dönüşüm doğruluğu açısından önem taşımaktadır.

Piksel koordinatlarından UTM düzlemi koordinatlarına dönüşüm modeli Eşitlik (3.5) ile ifade edilmiştir:

$$\begin{bmatrix} E' \\ N' \\ W \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{3.5}$$

Elde edilen homojen koordinatlar ölçek katsayısı olan (W) değerine bölünerek gerçek dünya koordinatları Eşitlik (3.6) ile hesaplanmıştır:

$$E = E'/W, N = \frac{N'}{W} \tag{3.6}$$

Eşitliklerde u ve v, görüntü düzlemindeki piksel koordinatlarını; E ve N, UTM düzlemindeki Easting ve Northing koordinatlarını; h_{ij}, homografi matrisinin elemanlarını; W ise homojen koordinatlardaki ölçek katsayısını ifade etmektedir.

Homografi matrisinin hesaplanmasında görüntü düzlemindeki piksel koordinatları ile UTM düzlemindeki koordinatlar arasındaki nokta eşleşmeleri kullanılmıştır. Bu çalışmada homografi matrisi, OpenCV kütüphanesindeki findHomography fonksiyonu ile standart çözüm yöntemi kullanılarak hesaplanmıştır. Kontrol noktası dosyasında görüntü adı bilgisi bulunduğu durumlarda her görüntü için ayrı homografi matrisi hesaplanmış; aksi durumda tüm görüntüler için ortak bir homografi matrisi kullanılmıştır.

Bu nedenle dönüşüm doğruluğunun artırılabilmesi için referans noktalarının görüntü üzerinde doğru işaretlenmesine, gerçek dünya koordinatlarının güvenilir biçimde belirlenmesine ve noktaların çalışma alanına dengeli biçimde dağılmasına dikkat edilmiştir. Homografi sonucunda elde edilen UTM koordinatları, karşılaştırma ve raporlama süreçlerinde gerektiğinde WGS 84 koordinat sistemine dönüştürülmüştür. Homografi matrisi hesabı ve piksel koordinatlarının UTM düzlemine aktarılmasına ilişkin örnek kod kesiti Ek 12'de verilmiştir.

3.4.2 IDW Tabanlı Lokal Enterpolasyon Yöntemi ile Konum Kestirimi

Homografi yaklaşımı, görüntü düzlemi ile yer düzlemi arasında tek bir projektif dönüşüm matrisi kullanarak koordinat dönüşümü gerçekleştirmektedir. Bu yaklaşım, çalışma alanının yaklaşık düzlemsel kabul edilebildiği bölgelerde kullanılabilmekle birlikte, arazi eğiminin değiştiği veya referans noktalarının görüntü içerisinde homojen dağılmadığı alanlarda konumsal hataların artmasına neden olabilmektedir. Dikilitaş çalışma alanının değişken topoğrafik yapısı dikkate alındığında, tek bir homografi dönüşümünün tüm görüntü alanı için aynı doğrulukta sonuç üretmesi beklenmemektedir.

Bu nedenle çalışmada, homografi yaklaşımına ek olarak referans noktalarına dayalı lokal enterpolasyon yöntemi de uygulanmıştır. Bu yaklaşımda, görüntü üzerinde piksel koordinatları bilinen referans noktaları ile bu noktalara karşılık gelen gerçek dünya koordinatları kullanılmıştır. Öncelikle referans noktalarının WGS 84 formatındaki enlem ve boylam değerleri, metrik hesaplamaların yapılabilmesi için UTM koordinat düzlemine dönüştürülmüştür. Böylece enterpolasyon işlemi, metre biriminde ifade edilen Easting ve Northing koordinatları üzerinden gerçekleştirilmiştir.

Lokal enterpolasyon sürecinde Ters Mesafe Ağırlıklı Enterpolasyon (Inverse Distance Weighting - IDW) yöntemi kullanılmıştır. Bu yöntemde, tespit edilen nesneye ait piksel referans noktası ile kontrol noktaları arasındaki piksel düzlemi mesafeleri hesaplanmış ve en yakın kontrol noktaları seçilmiştir. Çalışmada IDW için en yakın kontrol noktası sayısı $(k = 6)$, güç parametresi ise $(p = 2{,}0)$ olarak belirlenmiştir. Tahmin edilecek piksel noktasına daha yakın olan kontrol noktaları daha yüksek ağırlık alırken, uzak noktaların etkisi daha düşük olmuştur.
IDW yöntemiyle tahmin edilen koordinatlar, kontrol noktalarının UTM Easting ve Northing değerlerinin mesafe ağırlıklı ortalaması alınarak hesaplanmıştır. Nesne tespitlerinden elde edilen hem sınırlayıcı kutu merkezi hem de alt orta nokta referansları için ayrı ayrı enterpolasyon yapılmıştır. Böylece piksel referans noktası seçiminin konum kestirim sonucu üzerindeki etkisi de değerlendirilebilmiştir.

Enterpolasyon sonucunda her bir tespit için UTM koordinatları (E, N) elde edilmiş, daha sonra bu değerler WGS 84 koordinat sistemine dönüştürülerek enlem ve boylam bilgileri üretilmiştir. Elde edilen sonuçlar, homografi ve ray-casting yöntemlerinden elde edilen konum kestirimleriyle karşılaştırılmak üzere kaydedilmiştir. Homografi ve lokal enterpolasyon işlemlerinde kullanılan kontrol noktalarına ait veri yapısı Ek 3'te verilmiştir. Bu yöntem, çalışma alanındaki referans noktalarına dayalı olarak görüntü tabanlı konum kestirimi için homografi yaklaşımına alternatif bir değerlendirme yöntemi olarak kullanılmıştır. IDW tabanlı lokal enterpolasyon işlemine ait örnek kod kesiti Ek 13'te sunulmuştur.

3.4.3 CesiumJS Ortamında Ray-Casting Yaklaşımı

Homografi ve IDW tabanlı lokal enterpolasyon yaklaşımları, görüntü düzlemindeki piksel koordinatlarını yer düzlemi veya referans noktaları üzerinden gerçek dünya koordinatlarıyla ilişkilendirmektedir. Ancak çalışma alanında eğim değişimleri, kot farkları ve düzensiz 3B yüzeylerin bulunması, yalnızca düzlemsel dönüşüm veya referans noktası tabanlı kestirimlerle elde edilen konumların doğruluğunu etkileyebilmektedir. Bu nedenle çalışmada, tespit edilen nesne piksel noktalarının 3B dijital ikiz yüzeyiyle doğrudan ilişkilendirilmesi amacıyla CesiumJS ortamında ray-casting yaklaşımı uygulanmıştır.

Ray-casting işlemi için öncelikle YOLO tabanlı nesne tespitinden elde edilen sınırlayıcı kutu merkezi ve alt orta nokta koordinatları kullanılarak ray girdi verisi oluşturulmuştur. Bu veri içerisinde görüntü boyutu, kamera konumu, kamera yönelimi, görüş açısı ve görüntü oranı bilgileri ile her bir tespit nesnesine ait piksel koordinatları yer almaktadır. Böylece her bir piksel referans noktası, CesiumJS ortamında kamera konumu ve yönelimiyle birlikte değerlendirilebilecek bir ray-casting girdisine dönüştürülmüştür.

CesiumJS ortamında kamera; enlem, boylam, yükseklik, heading, pitch, roll, Field of View (FOV) ve aspect ratio parametreleri kullanılarak konumlandırılmıştır. Daha sonra kameranın yön, sağ ve yukarı vektörleri alınarak görüntü düzlemindeki piksel koordinatı normalize edilmiş ve ilgili pikselden geçen ışın yönü hesaplanmıştır.

Piksel koordinatlarının normalize edilmesi Eşitlik (3.7) ve Eşitlik (3.8) ile ifade edilebilir:

$$u_n = 2\left(\frac{u}{w}\right) - 1 \tag{3.7}$$

$$v_n = 1 - 2\left(\frac{v}{h}\right) \tag{3.8}$$

Burada u ve v, görüntü düzlemindeki piksel koordinatlarını; w ve h, görüntünün

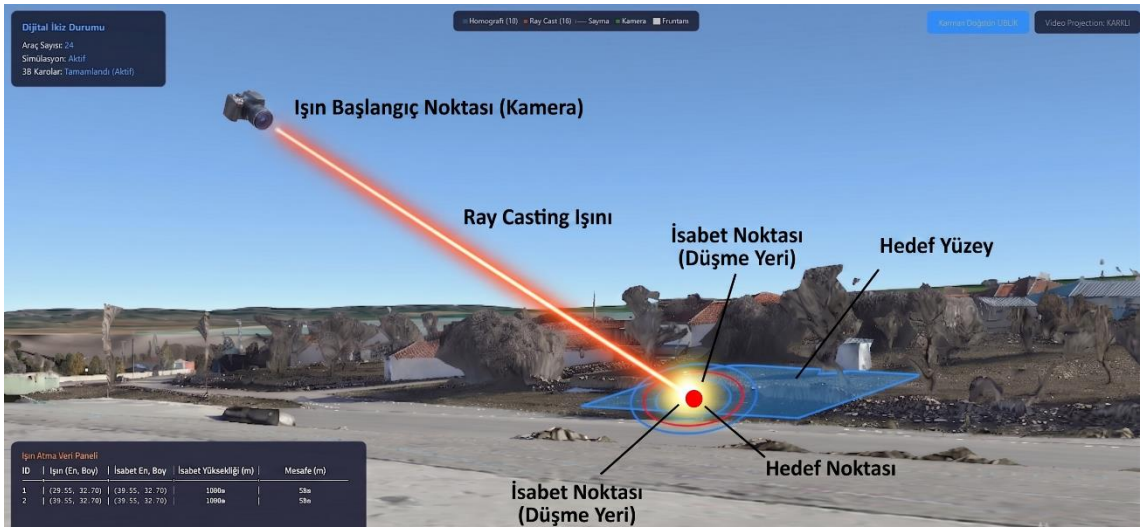
genişlik ve yükseklik değerlerini; u_n ve v_n ise normalize edilmiş görüntü koordinatlarını ifade etmektedir. Görüntü koordinat sisteminde başlangıç noktası sol üst köşe kabul edildiğinden, düşey eksendeki yön v_n ifadesinde ters çevrilmiştir.

Normalize edilmiş piksel koordinatları kullanılarak kamera merkezinden geçen ışın yönü hesaplanmış ve genel ışın modeli Eşitlik (3.9) ile ifade edilmiştir:

$$P(\lambda) = C + \lambda d \tag{3.9}$$

Eşitlikte C, kamera konumunu; d, ilgili piksel noktasından geçen normalize edilmiş ışın yön vektörünü; λ, ışın boyunca ilerleme parametresini; $P(\lambda)$ ise ışın üzerindeki herhangi bir 3B noktayı göstermektedir. Ray-casting işleminde amaç, bu ışının 3B dijital ikiz yüzeyi ile kesiştiği noktayı belirlemektir.

Bu çalışmada ray-casting işlemi, CesiumJS ortamında yüklenen 3D Tiles veri katmanları üzerinde uygulanmıştır. Işın öncelikle aktif 3D Tiles sahne geometrisi ile kesiştirilmiş, kesişim bulunamadığı durumlarda arazi/globe yüzeyi ile kesişim kontrol edilmiştir. Kesişim sonucunda elde edilen 3B nokta WGS 84 koordinat sisteminde enlem, boylam ve yükseklik bilgisine dönüştürülerek kaydedilmiştir. Ray-casting yaklaşımının uygulama arayüzü ve ışın-yüzey kesişim mantığı Şekil 3.6'da gösterilmiştir.



Şekil 3.6 Kamera projeksiyon merkezinden piksel noktasına uzanan matematiksel ışının 3B model ile kesişim (ray-casting) prensibi.

Bu yaklaşımda LiDAR ve oblik fotogrametri tabanlı 3D Tiles katmanları ayrı ayrı değerlendirilebilmiştir. Böylece aynı piksel referans noktası için farklı 3B yüzey altlıkları üzerinde kesişim noktaları hesaplanmış ve elde edilen sonuçlar konumsal doğruluk analizlerinde karşılaştırılmıştır. Bununla birlikte ray-casting sonuçlarının doğruluğu; kamera konumu ve yöneliminin sahadaki gerçek kamera duruşuyla uyumuna, FOV ve görüntü oranı değerlerinin doğruluğuna, 3B modelin geometrik kalitesine ve tespit edilen nesne için seçilen piksel referans noktasının uygunluğuna bağlıdır. Bu nedenle ray-casting sonuçları homografi ve IDW tabanlı lokal enterpolasyon sonuçlarıyla birlikte karşılaştırmalı olarak değerlendirilmiştir.

3.4.4 Konumsal Doğruluk Değerlendirme Yaklaşımı

Piksel koordinatlarından gerçek dünya koordinatlarına dönüşümde kullanılan homografi, IDW tabanlı lokal enterpolasyon ve ray-casting yöntemlerinin konumsal başarımı, RTK ölçümleri referans alınarak değerlendirilmiştir. Bu kapsamda her yöntemle elde edilen tahmini konumlar, aynı nesne veya kontrol konumu için ölçülen RTK koordinatlarıyla karşılaştırılmıştır.

Hata analizlerinin metre cinsinden yapılabilmesi için tahmini konumlar ve RTK referans koordinatları UTM düzleminde değerlendirilmiştir. Her bir tespit için yatay konum hatası, tahmini Easting ve Northing koordinatları ile RTK referans koordinatları arasındaki iki boyutlu Öklid mesafesi olarak hesaplanmıştır.

$$e_i = \sqrt{\left(E_i - E_{RTK,i}\right)^2 + \left(N_i - N_{RTK,i}\right)^2} \tag{3.10}$$

Eşitlikte e_i, i. tespit için yatay konum hatasını; E_i ve N_i, yöntemle elde edilen tahmini UTM koordinatlarını; $E_{RTK,i}$ ve $N_{RTK,i}$ ise aynı noktaya ait RTK referans koordinatlarını ifade etmektedir.

Bu çalışmada homografi ve IDW tabanlı lokal enterpolasyon yöntemleri, temel olarak piksel koordinatlarından yatay konum bileşenlerinin (E, N) kestirilmesi amacıyla kullanılmıştır. Bu yöntemler doğrudan 3B yüzey kesişimine dayanmadığından, yükseklik bileşeni yöntem çıktısının asli bir parçası olarak değerlendirilmemiştir. Buna karşılık ray-casting yaklaşımında, kamera merkezinden ilgili piksel noktasına uzanan ışının 3B dijital ikiz yüzeyiyle kesiştiği nokta hesaplandığından, yatay konumla birlikte yükseklik bilgisi de elde edilebilmiştir. Bu nedenle yöntemlerin konumsal doğruluk karşılaştırmalarında temel değerlendirme yatay konum hatası üzerinden yapılmış; ray-casting sonuçları ise 3B yüzeyle ilişkili konum üretme kapasitesi açısından ayrıca değerlendirilmiştir.

3.5 Video Tabanlı Konum Yumuşatma ve Kalman Filtreleme

Video tabanlı konum kestiriminde her bir görüntü karesinden elde edilen nesne konumu, ray-casting sonucunda WGS 84 koordinatları olarak elde edilmiştir. Ancak ardışık karelerde üretilen ham konum değerleri; nesne tespit kutularındaki küçük değişimler, piksel referans noktası seçimi, kamera görüş geometrisi ve 3B model yüzeyiyle kesişim sürecindeki belirsizlikler nedeniyle kısa süreli konum sıçramaları içerebilmektedir. Bu tür sıçramalar, hareketli nesnenin 3B dijital ikiz ortamında izlenmesi sırasında titreşimli veya fiziksel hareket sürekliliğiyle uyumsuz yörüngelerin oluşmasına neden olabilmektedir.

Bu nedenle çalışmada, ray-casting ile elde edilen ham konum serilerinin zamansal olarak daha kararlı hâle getirilmesi amacıyla Kalman filtreleme yaklaşımı uygulanmıştır. Filtreleme işlemi doğrudan enlem ve boylam değerleri üzerinde değil, metre biriminde hesaplama yapılabilmesi için UTM düzlemindeki Easting ve Northing bileşenleri üzerinde gerçekleştirilmiştir. Ray-casting sonucunda elde edilen WGS 84 koordinatları önce UTM düzlemine dönüştürülmüş, filtreleme ve hız hesapları bu metrik koordinatlar üzerinden yapılmıştır.

Kalman filtresi, ham konum ölçümlerini doğrudan nihai konum olarak kabul etmek yerine, önceki karelerden gelen hareket eğilimiyle birlikte değerlendirerek daha kararlı

bir konum kestirimi üretmek amacıyla kullanılmıştır. Bu kapsamda sabit hız hareket modeli esas alınmış; konum bileşenleri ölçüm olarak kullanılırken, hız bileşenleri model tarafından kestirilmiştir. Filtreleme sonucunda elde edilen konum serileri, ham konum sonuçlarıyla karşılaştırılarak video tabanlı dijital ikiz görselleştirmesinde zamansal kararlılık açısından değerlendirilmiştir.

3.5.1 Ham Konum Serilerinde Aykırı Değer ve Hız Sıçraması Kontrolü

Kalman filtreleme öncesinde, ray-casting sonucunda elde edilen ham konum serileri fiziksel hareket sürekliliği açısından kontrol edilmiştir. Bu amaçla her video karesi için elde edilen WGS 84 koordinatları UTM düzlemine dönüştürülmüş ve ardışık geçerli konumlar arasındaki mesafe ile zaman farkı kullanılarak anlık hız değerleri hesaplanmıştır. Hız hesabı, Easting ve Northing bileşenleri arasındaki iki boyutlu mesafe değişimine dayalı olarak yapılmıştır.

Fiziksel olarak gerçekçi olmayan ani konum sıçramalarının filtre sonucunu olumsuz etkilemesini önlemek amacıyla hız eşiği tabanlı aykırı değer kontrolü uygulanmıştır. Bu kontrolde her nokta yalnızca bir önceki satırla değil, son kabul edilen geçerli konumla karşılaştırılmıştır. Böylece arka arkaya gelen hatalı konum ölçümlerinin filtreye geçerli ölçüm olarak aktarılması engellenmiştir.

Belirlenen hız eşiğini aşan konumlar aykırı değer olarak işaretlenmiş ve Kalman filtresine doğrudan ölçüm olarak verilmemiştir. Ancak zaman dizisi korunmuş, böylece geçersiz ölçüm bulunan karelerde filtre yalnızca hareket modeline dayalı tahmin üretmeye devam etmiştir. Geçerli yeni bir ölçüm geldiğinde ise filtre tahmini yeniden ölçümle güncellenmiştir. Bu yaklaşım, ham ray-casting sonuçlarındaki ani sıçramaların yörüngeye doğrudan yansımasını azaltmak amacıyla uygulanmıştır.

3.5.2 UTM Düzleminde Sabit Hız Modeline Dayalı Kalman Filtresi

Bu çalışmada kullanılan Kalman filtresi, iki boyutlu UTM koordinat düzleminde sabit hız hareket modeline dayalı olarak uygulanmıştır. Filtre durum vektörü, Easting ve

Northing konum bileşenleri ile bu doğrultulardaki hız bileşenlerinden oluşturulmuştur. Bu yapı sayesinde filtre, yalnızca konum ölçümlerini kullanmakla kalmamış, aynı zamanda ardışık karelerdeki hareket eğilimini de dikkate alarak hız bileşenlerini kestirmiştir.

Durum vektörü Eşitlik (3.11) ile tanımlanmıştır:

$$x_i = \begin{bmatrix} E_i \\ N_i \\ v_{E,i} \\ v_{N,i} \end{bmatrix} \tag{3.11}$$

Ölçüm vektörü ise yalnızca konum bileşenlerinden oluşmaktadır:

$$z_i = \begin{bmatrix} E_i \\ N_i \end{bmatrix} \tag{3.12}$$

Burada E_i ve N_i, i. kareye ait UTM Easting ve Northing koordinatlarını; $v_{E,i}$ ve $v_{N,i}$, ise Easting ve Northing doğrultularındaki hız bileşenlerini ifade etmektedir. Ölçüm vektöründe yalnızca konum bileşenleri yer aldığından, hız değerleri doğrudan ölçülmemiş; sabit hız modeli ve ardışık konum ölçümleri üzerinden kestirilmiştir.

Kalman filtreleme sürecinde tahmin ve güncelleme adımları her video karesi için sırasıyla uygulanmıştır. Geçerli ölçüm bulunan karelerde filtre tahmini ölçümle güncellenmiş, aykırı değer olarak belirlenen veya ölçüm bulunmayan karelerde ise yalnızca tahmin adımı kullanılmıştır. Bu sayede filtre, ham konum ölçümlerindeki kısa süreli dalgalanmaları azaltırken hareket sürekliliğini koruyacak şekilde uygulanmıştır.

3.5.3 Ham ve Filtreli Yörüngelerin Karşılaştırma Ölçütleri

Kalman filtreleme sonucunda her video karesi için filtrelenmiş Easting ve Northing konum bileşenleri ile bu doğrultulardaki hız bileşenleri elde edilmiştir. Elde edilen filtreli UTM koordinatları, görselleştirme ve raporlama amacıyla tekrar WGS 84

koordinat sistemine dönüştürülmüştür.

Ham ve filtreli konum serilerinin karşılaştırılmasında; ardışık kareler arasındaki adım mesafesi, hız değişimi ve referans doğrultuya olan dik uzaklık gibi ölçütler kullanılmıştır. Referans doğrultu, video sahnesinde hareketin beklenen başlangıç ve bitiş konumları arasındaki çizgi olarak tanımlanmış; ham ve filtreli konumların bu doğrultuya göre sapmaları karşılaştırılmıştır. Böylece filtreleme işleminin konum serisindeki ani sıçramaları ne ölçüde azalttığı ve yörüngeyi fiziksel hareket sürekliliğine ne kadar yaklaştırdığı değerlendirilmiştir.

Bu kapsamda ham konum, Kalman filtreli konum, ham hız, filtreli hız ve referans doğrultuya uzaklık değerleri ayrı ayrı hesaplanmış ve sonuçlar çizelge ve grafiklerle değerlendirilmiştir. Bu değerlendirme, Kalman filtreleme yaklaşımının video tabanlı dijital ikiz uygulamasında zamansal kararlılık üzerindeki etkisini ortaya koymak amacıyla yapılmıştır.

3.6 CesiumJS Ortamında Entegrasyon ve Görselleştirme

Çalışmanın son aşamasında, üretilen 3B dijital ikiz altlığı ile derin öğrenme tabanlı nesne tespiti ve konum kestirimi çıktıları CesiumJS ortamında bir araya getirilmiştir. Bu aşamadaki temel amaç, çalışma alanına ait fotogrametrik model ve LiDAR nokta bulutu verilerinin web tabanlı 3B sahnede görselleştirilmesi ve video görüntülerinden elde edilen nesne konumlarının aynı mekânsal referans sistemi içerisinde değerlendirilebilmesidir.

Bu kapsamda fotogrametrik mesh model ve LiDAR nokta bulutu verileri 3D Tiles veri yapısı üzerinden CesiumJS ortamına aktarılmıştır. Nesne tespiti sonucunda elde edilen sınırlayıcı kutu referans noktaları, kamera konumu, kamera yönelimi, görüntü çözünürlüğü, FOV ve aspect ratio parametreleriyle birlikte ray-casting sürecinde kullanılmak üzere JSON formatında hazırlanmıştır. CesiumJS ortamında ray-casting işlemi için kullanılan örnek girdi JSON veri yapısı Ek 8'de sunulmuştur. CesiumJS tarafında bu veriler okunarak ilgili piksel noktaları için kamera merkezinden 3B dijital

ikiz yüzeyine doğru ışınlar oluşturulmuş ve ışın-yüzey kesişim noktaları hesaplanmıştır.

Bu entegrasyon sayesinde, nesne tespiti sonucunda elde edilen piksel tabanlı çıktılar yalnızca görüntü düzleminde kalmamış; homografi, IDW tabanlı lokal enterpolasyon ve ray-casting yöntemlerinden elde edilen konumsal sonuçlarla birlikte 3B dijital ikiz ortamında değerlendirilebilir hâle getirilmiştir. Böylece çalışma kapsamında geliştirilen iş akışı, video tabanlı nesne tespit sonuçlarının 3B coğrafi ortamda görselleştirilmesi ve konumsal doğruluk analizleriyle birlikte yorumlanması için kullanılabilir bir uygulama altyapısı sunmuştur.

3.6.1 Dijital İkizin Web Platformuna Aktarılması

Çalışma alanına ait fotogrametrik mesh model ve LiDAR nokta bulutu verileri, büyük hacimli 3B mekânsal veri niteliği taşıdığından web ortamında doğrudan tek parça hâlinde kullanılmamıştır. Bu nedenle veriler, CesiumJS ortamında daha verimli biçimde görüntülenebilmesi için 3D Tiles veri yapısı kullanılarak aktarılmıştır. 3D Tiles yapısı, büyük 3B verileri mekânsal olarak parçalara ayırmakta ve kullanıcı bakış açısına göre gerekli ayrıntı düzeyindeki karoların yüklenmesine olanak sağlamaktadır.

CesiumJS ortamında LiDAR ve fotogrametrik model katmanları ayrı veri altlıkları olarak tanımlanmıştır. Böylece ray-casting uygulamalarında aktif 3D Tiles katmanı değiştirilebilmiş ve aynı piksel referans noktası için farklı 3B yüzey altlıkları üzerinde kesişim sonuçları elde edilebilmiştir. Bu yapı, LiDAR nokta bulutu ve oblik fotogrametri tabanlı mesh modelin konum kestirim sürecindeki etkilerinin karşılaştırılabilmesine olanak sağlamıştır.

Web tabanlı görselleştirme sürecinde CesiumJS'in kamera kontrolü, 3D Tiles yükleme, sahne geometrisiyle kesişim sorgulama ve koordinat dönüşüm özelliklerinden yararlanılmıştır. Kamera konumu ve yönelimi çalışma alanındaki sabit kamera konfigürasyonuna göre tanımlanmış; FOV ve aspect ratio değerleri kullanılarak video görüntüsündeki piksel koordinatları ile CesiumJS sahnesindeki 3B ışınlar ilişkilendirilmiştir.

3.6.2 Nesne Tespit Çıktılarının CesiumJS Ortamına Aktarılması İş Akışı

YOLO tabanlı nesne tespiti sonucunda elde edilen yaya ve araç gibi hareketli nesnelere ait sınırlayıcı kutu bilgileri, piksel referans noktalarına dönüştürülerek ray-casting sürecine aktarılmıştır. Bu kapsamda her tespit için sınıf adı, güven skoru, görüntü adı, nesne kimliği, piksel koordinatları ve kullanılan sınırlayıcı kutu referans noktası bilgileri kaydedilmiştir. Hazırlanan bu veriler JSON formatında CesiumJS arayüzüne aktarılmıştır.

CesiumJS arayüzünde her piksel referans noktası için kamera merkezinden başlayan ve ilgili piksel doğrultusundan geçen bir ışın oluşturulmuştur. Oluşturulan ışın önce aktif 3D Tiles sahne geometrisiyle, kesişim bulunamadığı durumlarda ise arazi/globe yüzeyiyle ilişkilendirilmiştir. Işının 3B yüzeyle kesiştiği nokta, ilgili nesneye ait konumsal kestirim sonucu olarak kaydedilmiştir. Elde edilen sonuçlar enlem, boylam ve yükseklik bilgileriyle birlikte dışa aktarılmıştır. CesiumJS ortamında gerçekleştirilen ray-casting işlemi sonucunda elde edilen çıktı dosyasına ait veri yapısı Ek 4'te verilmiştir.

Dinamik nesnelerin 3B sahneye model olarak eklenmesi durumunda, konumlandırma işleminin yalnızca enlem, boylam ve yükseklik bileşenleriyle sınırlı kalmaması gerekmektedir. Özellikle araç, tır ve otobüs gibi zeminle sürekli temas hâlinde bulunan nesnelerde, modelin yönelimi sahnenin yerel yüzey eğimiyle uyumlu olmalıdır. Bu nedenle gerçekçi bir görselleştirme için heading açısına ek olarak, yüzey eğimine bağlı pitch ve roll bileşenlerinin de dikkate alınması önerilmektedir.

Geliştirilen arayüzde LiDAR, oblik fotogrametri veya her iki 3B altlığın birlikte görüntülenebilmesi sağlanmıştır. CesiumJS ortamında ray-casting ve GLB/glTF tabanlı 3B model yerleştirme işlemlerine ait örnek kod kesiti Ek 16'da verilmiştir. Ayrıca tespit edilen nesneler, ray-casting ışınları, hedef yüzey ve kesişim noktaları görsel olarak izlenebilmiştir. Bu yapı, nesne tespiti çıktılarının 3B dijital ikiz üzerinde konumsal olarak değerlendirilmesine ve farklı yüzey altlıkları üzerinden elde edilen ray-casting

sonuçlarının karşılaştırılmasına olanak sağlamıştır. Geliştirilen CesiumJS arayüzünde tespit edilen nesneler ve ray-casting işleminin görselleştirilmesine ait örnek Şekil 3.7'de verilmiştir.

Şekil 3.7 CesiumJS ortamında tespit edilen nesneler ve ray-casting işleminin 3D Tiles altlığı üzerinde görselleştirilmesi.

4. BULGULAR

Bu bölümde, araştırma kapsamında Bölüm 3'te açıklanan veri toplama, veri işleme, nesne tespiti, koordinat dönüşümü ve görselleştirme süreçlerinden elde edilen bulgular sunulmaktadır. İlk olarak, Gölbaşı Dikilitaş çalışma alanında gerçekleştirilen İHA tabanlı fotogrametri ve LiDAR ölçümleri sonucunda üretilen 3B dijital ikiz altlığı; veri üretim parametreleri, model çıktıları, nokta bulutu yoğunluğu ve konumsal doğruluk değerleri açısından değerlendirilmiştir.

Ardından, sabit kamera görüntüleri üzerinde kullanılan YOLO tabanlı model varyantlarının nesne tespit başarımı incelenmiştir. Bu kapsamda modeller; sayım tabanlı ön değerlendirme, etiketli sınırlayıcı kutular üzerinden IoU tabanlı doğrulama, güven skoru davranışı ve çıkarım süresi bakımından karşılaştırılmıştır.

Son aşamada ise tespit edilen nesnelere ait piksel koordinatlarından gerçek dünya koordinatlarına geçişte kullanılan homografi, IDW tabanlı lokal enterpolasyon ve ray-casting yöntemlerinin konumsal doğruluk sonuçları değerlendirilmiştir. Elde edilen konum kestirimleri RTK ölçümleriyle karşılaştırılmış; ayrıca ray-casting tabanlı ham konum serilerinin Kalman filtreleme sonrasındaki zamansal kararlılık değişimi incelenmiştir. Böylece 3B dijital ikiz altlığı, nesne tespiti, koordinat dönüşümü ve web tabanlı görselleştirme süreçlerinin çalışma bütünlüğü içerisinde elde edilen sonuçları ortaya konulmuştur.

4.1 3B Model ve Dijital İkiz Üretim Sonuçları

Çalışma alanını oluşturan yaklaşık 1,503 km² büyüklüğündeki Dikilitaş bölgesinde, DJI Matrice 300 RTK platformu kullanılarak fotogrametrik ve LiDAR veri toplama uçuşları gerçekleştirilmiştir. Bu büyüklük yaklaşık 150,3 hektarlık bir alanı temsil etmektedir. DJI Terra (v5.0.2) yazılımı ortamında yürütülen veri işleme aşamaları sonucunda, çalışma alanına ait fotogrametrik mesh model ile LiDAR nokta bulutu aynı koordinat sistemi içerisinde değerlendirilmiş ve 3B dijital ikiz altlığı oluşturulmuştur. Fotogrametrik ve LiDAR veri üretimine ilişkin temel parametreler ile elde edilen sonuçlar Çizelge 4.1'de sunulmuştur.

Çizelge 4.1 Fotogrametrik ve LiDAR uçuşlarına ait genel veri üretim parametreleri ve analiz sonuçları.

Parametre	Fotogrametrik üretim (Zenmuse P1)	LiDAR üretimi (Zenmuse L2)
Uçuş irtifası (Ortalama)	152,6 m	137,7 m – 142,2 m
Girdi / Poz sayısı	2311 Adet	1223 Adet
Yer örnekleme aralığı (YÖA)	4,44 cm (0,044 m)	Aktif sensör
RTK çözüm durumu	2311 FIX	1219 FIX, 4 SINGLE
Coğrafi referanslama hatası (RMSE)	0,031 m	0,056 m
Uçuş/Veri toplama süresi	43 Dakika	43 Dakika
Çıktı formatları (Web uyumlu)	3D Tiles (B3DM)	3D Tiles (PNTS)

Fotogrametrik üretimde DJI Zenmuse P1 kamerası ile elde edilen 2311 görüntü kullanılmış, ortalama YÖA değeri 4,44 cm olarak hesaplanmıştır. Blok dengelemesi sonucunda fotogrametrik model için coğrafi referanslama hatası 0,031 m olarak elde edilmiştir. LiDAR üretiminde ise DJI Zenmuse L2 sensöründen 1223 poz verisi toplanmış, bu pozların 1219'u FIX çözümle, 4'ü ise SINGLE çözümle kaydedilmiştir. LiDAR verisi için coğrafi referanslama hatası 0,056 m olarak belirlenmiştir.

Elde edilen fotogrametrik mesh model ve LiDAR nokta bulutu, web tabanlı görselleştirme gereksinimleri doğrultusunda 3D Tiles formatına dönüştürülmüştür. Bu kapsamda fotogrametrik model b3dm, LiDAR verisi ise pnts formatında dışa aktarılmıştır. Veri hacminin web ortamında daha verimli yönetilebilmesi amacıyla model sadeleştirme oranı %50 olarak uygulanmıştır. Yaklaşık 60 GB veri hacmine

sahip olan bu veri setinin web tarayıcısında tek parça hâlinde yüklenmesi bellek, ağ bant genişliği ve grafik işleme performansı açısından uygun olmadığından, 3D Tiles veri yapısından yararlanılmıştır.

3D Tiles yaklaşımı, büyük hacimli 3B verilerin mekânsal olarak daha küçük karo parçalarına ayrılmasına ve kullanıcının görüş alanına giren gerekli ayrıntı düzeyindeki karoların yüklenmesine olanak sağlamaktadır. Bu kapsamda DJI Terra tarafından üretilen 3D Tiles veri dizinleri Cesium ion ortamına aktarılmış ve CesiumJS tabanlı web arayüzünde görselleştirilmiştir. Böylece fotogrametrik mesh model ile LiDAR nokta bulutu aynı 3B sahne içerisinde birlikte değerlendirilebilmiş ve dijital ikiz altlığı, sonraki aşamalarda nesne tespiti sonuçlarının konumlandırılması için kullanılabilir hâle getirilmiştir. Şekil 4.1'de çalışma alanına ait kamera projeksiyon merkezleri ile üretilen 3B dijital ikiz altlığının genel görünümü sunulmaktadır.

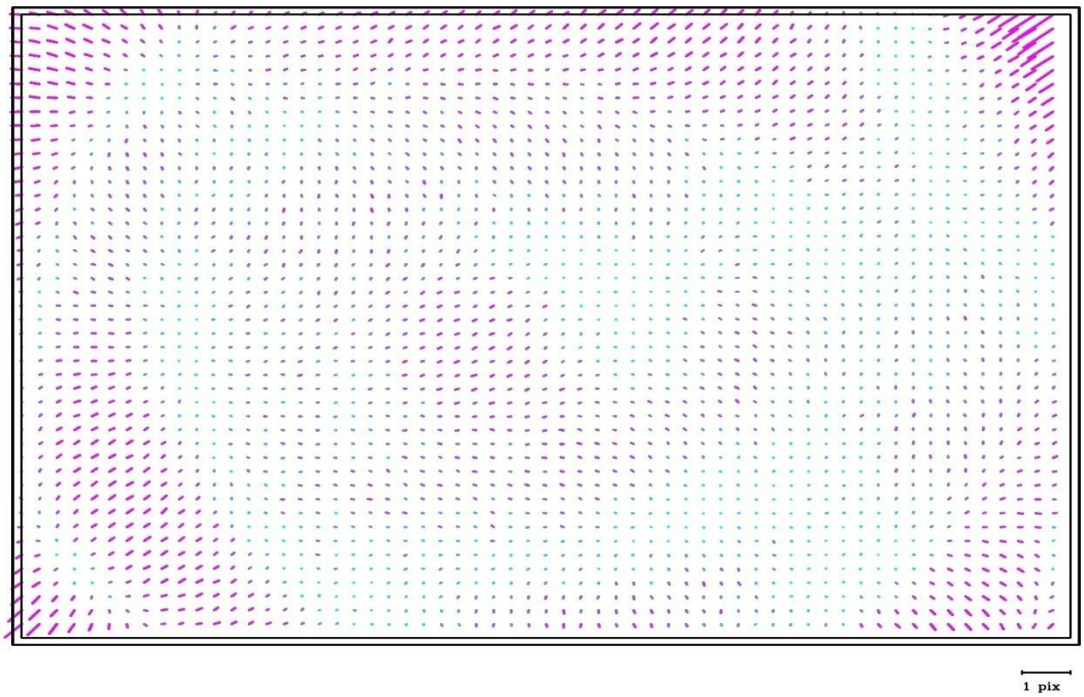


Şekil 4.1 Üretilen 3B dijital ikiz altlığının genel görünümü.

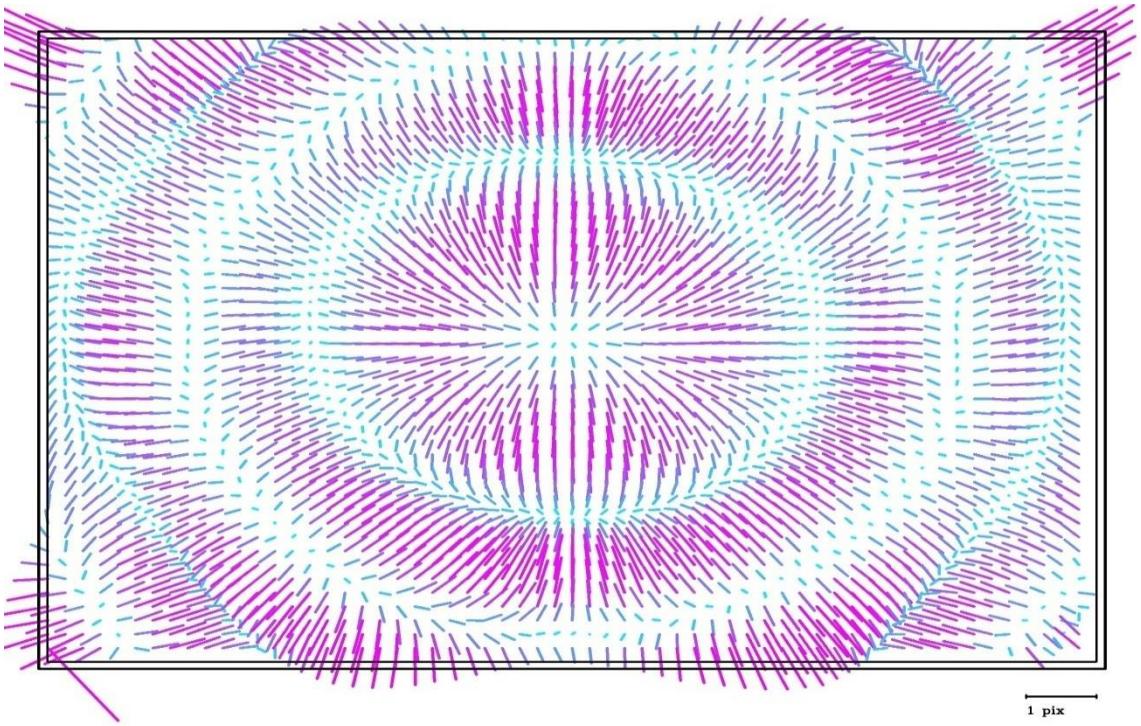
4.1.1 Fotogrametrik Model Kalitesi ve LiDAR Nokta Bulutu Yoğunluğu Analizi

Üretilen 3B dijital ikiz altlığının hareketli nesne konumlarının değerlendirilmesinde kullanılabilmesi için metrik doğruluk açısından güvenilir olması gerekmektedir. Bu nedenle fotogrametrik model üretimine ait hava üçgenlemesi ve blok dengelemesi

sonuçları ile LiDAR nokta bulutunun yoğunluk ve konum doğruluğu değerleri incelenmiştir. Fotogrametrik veri üretiminde DJI Zenmuse P1 kamerası ile toplanan 2311 görüntünün tamamı RTK FIX çözümle kaydedilmiştir. Blok dengelemesi sürecinde kamera yöneltme parametreleri ve mercek distorsiyon katsayıları otokalibrasyon yaklaşımıyla optimize edilmiştir. Bu işlem sonucunda fotogrametrik model için coğrafi referanslama RMSE değeri 0,031 m olarak elde edilmiştir. Zenmuse P1 kamerası ve LiDAR sistemine entegre RGB kamera için elde edilen piksel bazlı geri çatım hata vektörü dağılımları sırasıyla Şekil 4.2 ve Şekil 4.3'te sunulmuştur.



Şekil 4.2 Zenmuse P1 kamerasına ait piksel bazlı geri çatım hata vektörü dağılımı.



Şekil 4.3 Zenmuse L2 sistemine entegre RGB kameraya ait piksel bazlı geri çatım hata vektörü dağılımı.

Geri çatım hata dağılımları incelendiğinde, fotogrametrik blok dengelemesinin genel olarak kararlı bir sonuç verdiği görülmektedir. Zenmuse P1 kamerasına ait hata vektörlerinin sensör yüzeyinde daha dengeli bir dağılım göstermesi, fotogrametrik model üretiminde kullanılan görüntülerin blok dengelemesi açısından uygun bir yapı sunduğunu göstermektedir. Buna karşılık Zenmuse L2 sistemine entegre RGB kameraya ait hata dağılımında daha belirgin bir radyal karakter gözlenmektedir. Bu durum, LiDAR sistemine entegre yardımcı RGB kameranın mercek yapısı ve kullanım amacıyla ilişkili olarak değerlendirilebilir.

Fotogrametrik model üretimi sırasında görüntü hizalama, yoğun nokta bulutu üretimi ve 3B mesh model oluşturma işlemleri DJI Terra yazılımı kullanılarak gerçekleştirilmiştir. Donanım kaynakları ve veri hacmi dikkate alınarak çalışma alanı 16 alt blok hâlinde işlenmiş ve 3B model üretim süreci 1327 dakika sürmüştür. Eğik görüntülerin sağladığı farklı bakış açıları, özellikle bina cepheleri ve düşey yüzeylerin dokulu 3B model üzerinde daha ayrıntılı temsil edilmesine katkı sağlamıştır.

LiDAR veri üretiminde DJI Zenmuse L2 sensöründen elde edilen nokta bulutu kullanılmıştır. Lazer tarayıcının 240 kHz darbe hızı ve çoklu yansıma kabiliyeti

sayesinde çalışma alanı genelinde ortalama nokta yoğunluğu 205,3 nokta/m² olarak hesaplanmıştır. Elde edilen nokta yoğunluğu, çalışma alanının arazi yüzeyi, yapı geometrileri ve nokta bulutu tabanlı 3B temsil açısından ayrıntılı biçimde değerlendirilebilmesine olanak sağlamıştır.

Nokta bulutu yoğunluğu ayrıca Harita Genel Müdürlüğü (HGM) tarafından yayımlanan Büyük Ölçekli Harita ve Harita Bilgileri Üretim Yönetmeliği (BÖHHBÜY) kapsamında tanımlanan büyük ölçekli harita üretim standartları açısından değerlendirilmiştir. Bu değerlendirme sonucunda, LiDAR nokta yoğunluğunun 1:1000 ölçek üretim standardını %97,4 oranında, 1:500 ölçek üretim standardını ise %94,7 oranında karşıladığı belirlenmiştir. Bu sonuç, üretilen LiDAR verisinin çalışma alanının yüksek ayrıntılı arazi ve yüzey temsili için uygun bir yoğunluk düzeyi sunduğunu göstermektedir.

LiDAR verisinin konum doğruluğunu etkileyen bileşenlerden biri de GNSS/IMU yörünge çözümüdür. LiDAR üretiminde 1223 poz verisinin 1219'u RTK FIX çözümle, 4'ü ise SINGLE çözümle kaydedilmiştir. Yörünge çözümüne ait kuzey, doğu ve düşey bileşenlerdeki hata değerlerinin düşük seviyelerde kalması, LiDAR nokta bulutunun coğrafi referanslama doğruluğunu desteklemiştir. Genel değerlendirme sonucunda LiDAR verisi için coğrafi referanslama RMSE değeri 0,056 m olarak elde edilmiştir.

LiDAR verilerinin işlenmesi aşamasında zemin ve zemin dışı nokta ayrımı yapılmıştır. Bu kapsamda ham LiDAR nokta bulutu; zemin, yapı, araç ve bitki örtüsü gibi farklı yüzey özelliklerini temsil eden noktalardan oluşmaktadır.

Zemin sınıflandırması sonucunda elde edilen arazi yüzeyi bilgisi, çalışma alanının topografik yapısının temsil edilmesine katkı sağlamıştır. Bu veri, özellikle ray-casting yaklaşımında 3B yüzeyle kesişim noktasının belirlenmesi ve nesne konumlarının dijital ikiz altlığı üzerinde değerlendirilmesi açısından önemli bir referans katmanı olarak kullanılmıştır.

4.2 Nesne Tespit Modeli Performans Analizleri

Bu bölümde, sabit kamera görüntülerinden yaya nesnelerinin tespit edilmesi amacıyla kullanılan YOLO tabanlı model varyantlarının performansı karşılaştırmalı olarak değerlendirilmiştir. Model karşılaştırması, çalışma alanından seçilen 12 adet temsilî görüntü üzerinde gerçekleştirilmiştir. Seçilen görüntüler, yayanın kameraya olan uzaklığı ve görünürlük durumu dikkate alınarak sınıflandırılmıştır. Bu kapsamda görüntüler mesafe bakımından yakın, orta ve uzak; görünürlük bakımından ise açık ve örtülü olarak gruplandırılmıştır.

Değerlendirme sürecinde yalnızca yaya sınıfı dikkate alınmıştır. Her görüntü için gerçek yaya sayısı manuel olarak belirlenmiş ve farklı YOLO model varyantlarının aynı görüntüler üzerinde ürettiği tespit sayıları ile karşılaştırılmıştır. Bu aşamada yapılan değerlendirme, manuel BBox eşleşmesine dayalı IoU tabanlı nesne tespit doğrulaması değil, sayım tabanlı bir ön değerlendirme niteliğindedir. Bu nedenle doğru pozitif, yanlış pozitif ve yanlış negatif değerleri; gerçek yaya sayısı ile model tarafından tahmin edilen yaya sayısı arasındaki farklar dikkate alınarak hesaplanmıştır.

Bu analizin temel amacı, dijital ikiz sisteminde kullanılacak nesne tespit modelinin yalnızca genel tespit başarımını değil, aynı zamanda uzak, küçük veya kısmen örtülü hedeflerde güvenilir tespit üretip üretemediğini değerlendirmektir. Çünkü sonraki aşamada homografi, IDW tabanlı lokal enterpolasyon ve ray-casting yöntemleriyle gerçek dünya koordinatlarına dönüştürülecek piksel referans noktaları, doğrudan YOLO tarafından üretilen BBox çıktısına bağlıdır. Bu nedenle nesne tespit başarımı, koordinat dönüşüm sürecinin doğruluğunu etkileyen temel ön koşullardan biri olarak değerlendirilmiştir.

4.2.1 Model Varyantlarının Genel Sayım Tabanlı Tespit Başarımı

Nesne tespit modelinin çalışma sahasına ilk uygunluğunu değerlendirmek amacıyla YOLOv8 ve YOLO26 mimarilerine ait farklı ölçeklerdeki model varyantları karşılaştırılmıştır. Değerlendirme, çalışma alanından seçilen 12 temsilî görüntü üzerinde gerçekleştirilmiştir. Bu görüntülerde toplam 13 yaya bulunmakta olup, her görüntüdeki gerçek yaya sayısı manuel olarak belirlenmiştir. Test görüntüleri mesafe bakımından

yakın, orta ve uzak; görünürlük bakımından ise açık ve örtülü olarak sınıflandırılmıştır. Böylece model başarımı yalnızca toplam tespit doğruluğu açısından değil, sahne koşullarına bağlı değişim bakımından da incelenmiştir.

Bu aşamada yapılan değerlendirme, sayım tabanlı bir ön değerlendirme niteliğindedir. Her görüntü için modelin tespit ettiği yaya sayısı, manuel olarak belirlenen gerçek yaya sayısı ile karşılaştırılmıştır. Doğru pozitif değeri, gerçek yaya sayısı ile tahmin edilen yaya sayısının minimumu olarak hesaplanmıştır.

Modelin gerçek yaya sayısından fazla tespit üretmesi yanlış pozitif, eksik tespit üretmesi ise yanlış negatif olarak değerlendirilmiştir. Bu yaklaşımla her model için precision, recall ve F1-skoru değerleri hesaplanmış ve sonuçlar Çizelge 4.2'de verilmiştir.

Çizelge 4.2 YOLO model varyantlarının genel sayım tabanlı tespit başarımı.

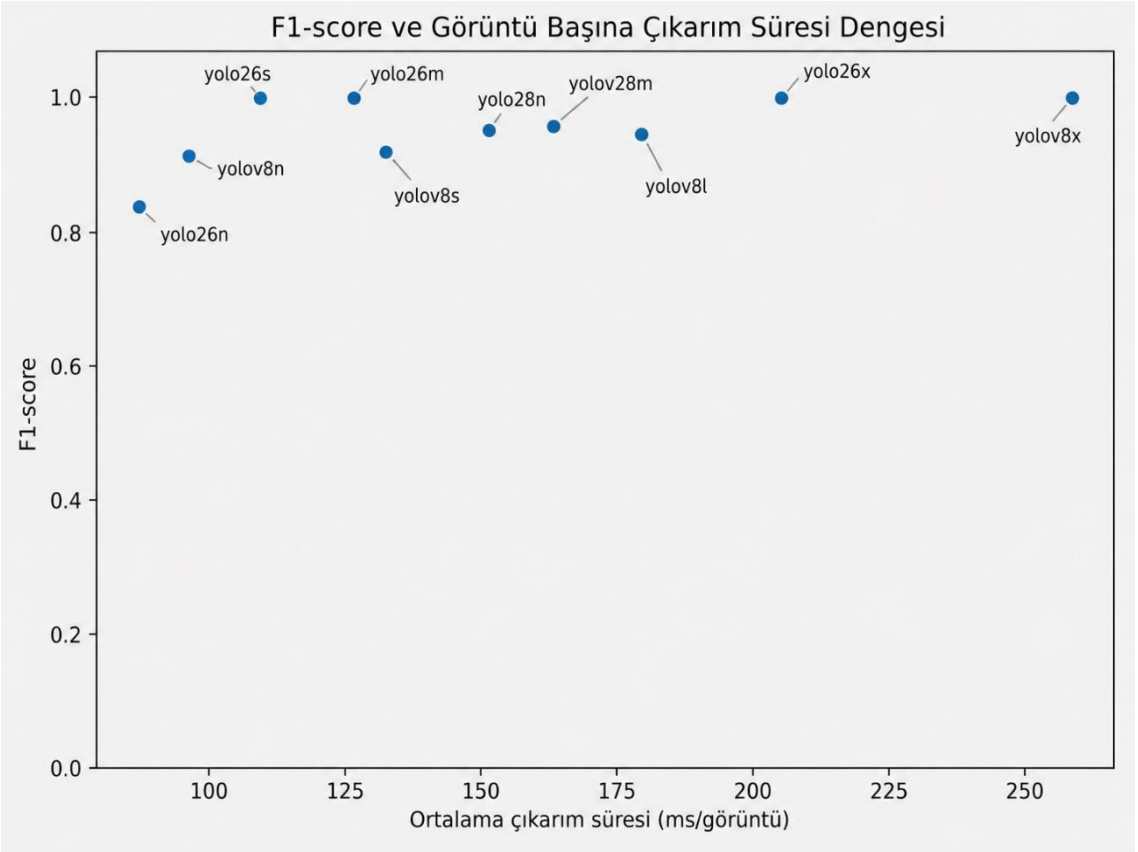
Model	TP	FP	FN	Precision	Recall	F1-skor	Sayım tabanlı yorum
YOLO26-l	12	0	1	1,000	0,923	0,960	Bir uzak yayayı kaçırmıştır
YOLO26-m	13	0	0	1,000	1,000	1,000	Genel olarak hatasız sayım
YOLO26-n	9	0	4	1,000	0,692	0,818	Uzak/örtülü sahnelerde zayıf
YOLO26-s	13	0	0	1,000	1,000	1,000	Dengeli ve başarılı
YOLO26-x	13	0	0	1,000	1,000	1,000	Başarılı, confidence yüksek
YOLOv8-l	12	0	1	1,000	0,923	0,960	Bir uzak yayayı kaçırmıştır
YOLOv8-m	13	1	0	0,929	1,000	0,963	Bir fazla tespit üretmiştir
YOLOv8-n	11	0	2	1,000	0,846	0,917	Bazı yayaları kaçırmıştır
YOLOv8-s	13	2	0	0,867	1,000	0,929	Fazla tespit üretmiştir
YOLOv8-x	13	0	0	1,000	1,000	1,000	Genel olarak hatasız sayım

Çizelge 4.2 incelendiğinde, YOLO26-m, YOLO26-s, YOLO26-x ve YOLOv8-x modellerinin 12 görüntüden oluşan ön değerlendirme setinde sayım hatası üretmediği görülmektedir. Bu modeller sayım tabanlı değerlendirmede precision, recall ve F1-skoru açısından en yüksek değerlere ulaşmıştır. Buna karşılık YOLO26-n ve YOLOv8-n modellerinde özellikle uzak ve örtülü sahnelerde eksik tespitler meydana gelmiştir.

YOLOv8-m ve YOLOv8-s modelleri ise tüm gerçek yayaları tespit etmekle birlikte, bazı görüntülerde fazla tespit üretmiştir.

Bu sonuçlar, model varyantları arasında özellikle küçük, uzak veya kısmen örtülü hedeflerde belirgin performans farklılıkları oluştuğunu göstermektedir. Ancak sayım tabanlı değerlendirme, tespit edilen BBox'ın gerçek nesneyle ne ölçüde örtüştüğünü göstermediğinden, model seçimi için tek başına yeterli değildir. Dijital ikiz sisteminde BBox'ın merkez veya alt orta noktası homografi, IDW tabanlı lokal enterpolasyon ve ray-casting yöntemleriyle gerçek dünya koordinatlarına dönüştürüleceğinden, BBox konum doğruluğunun ayrıca değerlendirilmesi gerekmektedir.

Sayım tabanlı ön değerlendirmede modellerin F1-skorları ile görüntü başına ortalama çıkarım süreleri birlikte incelenmiştir. Model varyantlarının F1-skoru ve çıkarım süresi ilişkisi Şekil 4.4'te verilmiştir.



Şekil 4.4 YOLO model varyantlarının F1-skor ve görüntü başına ortalama çıkarım süresi ilişkisi.

Şekil 4.4, sayım tabanlı değerlendirme kapsamında bazı modellerin yüksek F1-skoru değerine ulaşmasına karşın çıkarım süresi bakımından farklılaştığını göstermektedir. Bu ön analizde YOLO26-s modeli, yüksek F1-skoru değerini düşük çıkarım süresiyle birlikte sunması nedeniyle gerçek zamanlıya yakın dijital ikiz uygulamaları açısından güçlü bir aday olarak öne çıkmıştır. Bununla birlikte bu sonuç yalnızca sayım tabanlı ön değerlendirmeye dayandığından, nihai model seçimi etiketli BBox ve IoU tabanlı doğrulama sonuçlarıyla birlikte yapılmıştır.

4.2.2 Mesafe ve Görünürlük Koşullarına Göre Tespit Başarımı

Model varyantlarının genel tespit başarımına ek olarak, test görüntüleri mesafe ve görünürlük koşullarına göre de değerlendirilmiştir. Bu kapsamda görüntüler mesafe bakımından yakın, orta ve uzak; görünürlük bakımından ise açık ve örtülü olarak sınıflandırılmıştır. Böylece modellerin yalnızca toplam tespit başarımı değil, farklı sahne koşullarındaki davranışları da incelenmiştir.

Mesafe bazlı sonuçlar değerlendirildiğinde, yakın mesafedeki görüntülerde model varyantlarının genel olarak yüksek F1-skor değerlerine ulaştığı görülmektedir. Yakın mesafede yayanın görüntü düzleminde daha büyük piksel alanı kaplaması, tespit işlemini kolaylaştırmakta ve BBox üretimini daha kararlı hâle getirmektedir. Bu nedenle yakın mesafe grubunda modeller arasında belirgin bir başarım farkı oluşmamıştır.

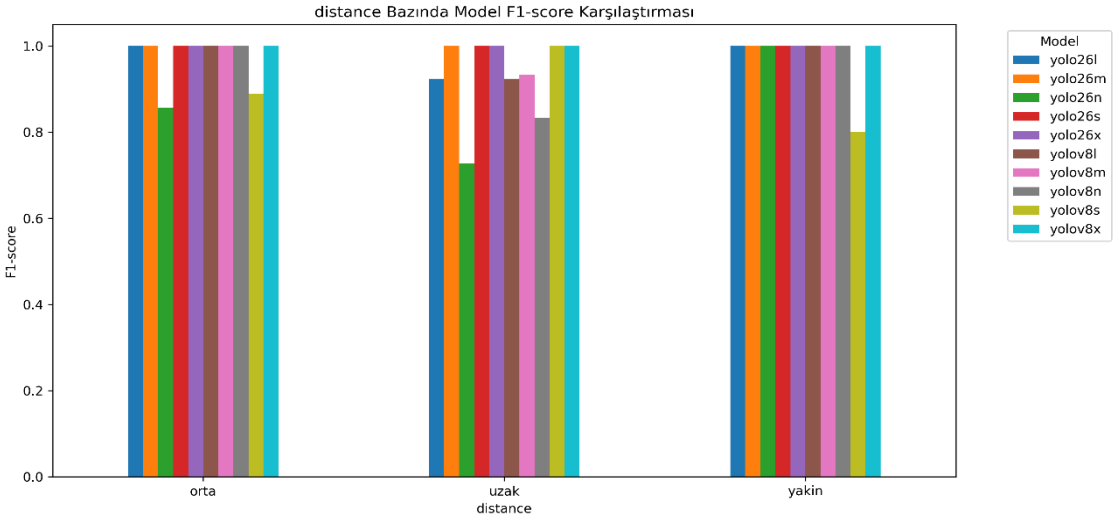
Orta mesafe grubunda da genel tespit başarımı yüksek olmakla birlikte, bazı model varyantlarında eksik veya fazla tespitler gözlenmiştir. Özellikle küçük ölçekli modellerin bazı sahnelerde yayayı kaçırabildiği, bazı modellerin ise gerçek yaya sayısından fazla tespit üretebildiği görülmüştür. Bu durum, orta mesafede nesne boyutunun hâlâ yeterli düzeyde olmasına rağmen arka plan karmaşıklığı, ışık koşulları ve BBox sınırlarının belirsizleşmesi gibi etkenlerin model performansını etkileyebildiğini göstermektedir.

Uzak mesafe grubunda ise model varyantları arasındaki fark daha belirgin hâle gelmiştir. Küçük ölçekli modellerin, özellikle YOLO26-n ve YOLOv8-n varyantlarının,

uzak mesafedeki bazı yayaları kaçırdığı görülmüştür. Buna karşılık YOLO26-m, YOLO26-s, YOLO26-x ve YOLOv8-x modelleri uzak mesafedeki yayaları daha yüksek F1-skor değerleriyle tespit etmiştir.

Bu sonuç, uzak mesafedeki yayaların görüntü üzerinde daha az piksel alanı kaplaması nedeniyle model kapasitesinin tespit başarımı üzerinde daha belirleyici hâle geldiğini göstermektedir.

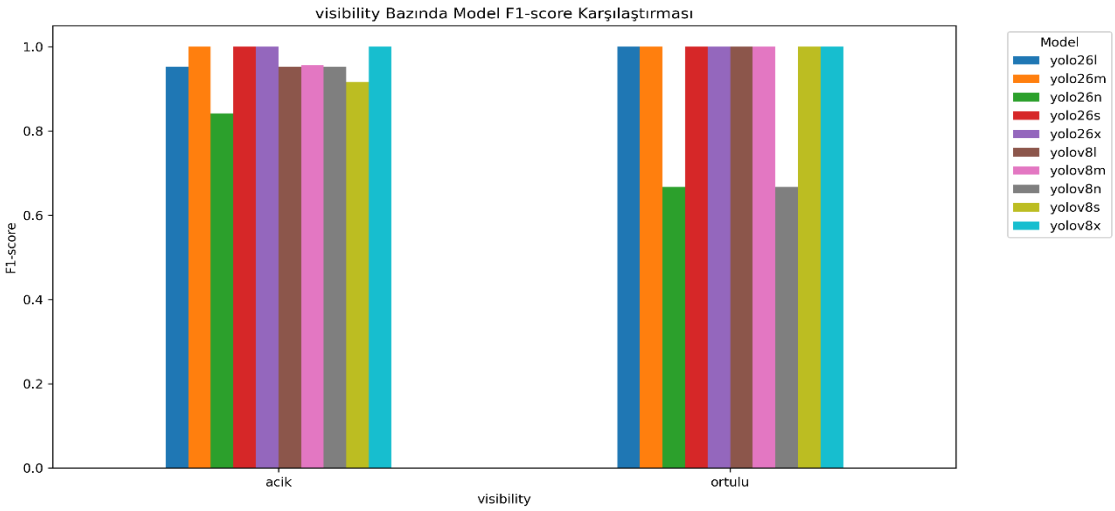
Dijital ikiz uygulaması açısından bu bulgu önemlidir. Çünkü uzak mesafede üretilen BBox'ların küçük ve hassas olması, sonraki aşamada yapılacak homografi, IDW tabanlı lokal enterpolasyon veya ray-casting tabanlı koordinat dönüşümlerinde konumsal belirsizliği artırabilmektedir. Mesafe koşullarına bağlı F1-skor değişimi Şekil 4.5'te sunulmuştur.



Şekil 4.5 Mesafe durumuna göre model varyantlarının F1-skor değişimi.

Görünürlük koşullarına göre yapılan değerlendirmede ise açık sahnelerde modellerin genel olarak daha başarılı sonuçlar ürettiği görülmektedir. Açık görüntülerde yayanın gövde ve ayak bölgesinin daha net seçilebilmesi, modelin hem nesneyi tespit etmesini hem de BBox sınırlarını daha kararlı üretmesini sağlamaktadır. Buna karşılık örtülü sahnelerde, özellikle küçük model varyantlarında tespit başarımının düştüğü görülmüştür. YOLO26-n ve YOLOv8-n modelleri örtülü sahnelerde bazı yayaları kaçırırken, daha büyük kapasiteli modeller örtülü sahnelerde daha yüksek F1-skor

değerleri üretmiştir. Açık ve örtülü sahnelere göre model varyantlarının F1-skor karşılaştırması Şekil 4.6'da verilmiştir.



Şekil 4.6 Görünürlük durumuna göre model varyantlarının F1-skor değişimi.

Bu sonuçlar, kısmi kapanma ve uzak mesafe gibi zorlayıcı koşulların model performansını doğrudan etkilediğini göstermektedir. Özellikle dijital ikiz üzerinde yaya konumlarının gerçek dünya koordinatlarına aktarılması hedeflendiğinden, modelin yalnızca yayayı tespit etmesi değil, aynı zamanda BBox referans noktasını kararlı biçimde üretmesi de önem taşımaktadır. Bu nedenle model seçimi yapılırken genel F1-skor değerinin yanı sıra uzak ve örtülü sahnelerdeki davranışı da dikkate alınmıştır.

4.2.3 Etiketli BBox ve IoU Tabanlı Model Doğrulaması

Sayım tabanlı ön değerlendirmenin ardından, model varyantlarının BBox konum doğruluğunu ve sınıf tahmin başarımını daha ayrıntılı biçimde analiz etmek amacıyla etiketli BBox verisi üzerinde doğrulama yapılmıştır.

Bu aşamada görüntüler LabelImg aracı kullanılarak manuel olarak etiketlenmiş ve gerçek BBox değerleri YOLO formatında kaydedilmiştir. Etiket sınıfları person, car, bus ve truck olarak tanımlanmıştır. Ancak etiketli test setinde bus ve truck sınıflarına ait örnek bulunmadığından, sınıf bazlı değerlendirmeler person ve car sınıfları üzerinden yapılmıştır. Etiketli BBox doğrulamasında kullanılan IoU tabanlı eşleştirme ve temel metrik hesaplama sürecine ait örnek kod kesiti Ek 11'de sunulmuştur.

Bu analizde model tahminleri ile gerçek BBox etiketleri IoU ≥ 0,50 koşulunda eşleştirilmiştir. IoU değeri, tahmin edilen BBox ile gerçek BBox arasındaki örtüşme oranını ifade etmektedir. Bir tahminin doğru kabul edilebilmesi için yalnızca doğru sınıfı üretmesi değil, aynı zamanda gerçek BBox ile belirlenen eşik değerinin üzerinde örtüşme sağlaması gerekmektedir. Bu nedenle bu aşamadaki değerlendirme, önceki sayım tabanlı analize göre BBox konum doğruluğunu da dikkate alan daha ayrıntılı bir doğrulama niteliği taşımaktadır.

Ana değerlendirme için confidence eşiği 0,30 olarak belirlenmiştir. Bu eşik, özellikle uzak veya küçük nesnelerin tamamen elenmesini önlemek amacıyla tercih edilmiştir. Confidence ≥ 0,30 koşulu altında elde edilen model karşılaştırma sonuçları Çizelge 4.3'te sunulmuştur.

Çizelge 4.3 Etiketli BBox verisi üzerinde confidence ≥ 0,30 için model performansları.

Model	Sınıf F1	BBox F1	Ort. IoU	IoU≥ 0,75	FP	FN	ms/ Görüntü	FPS
YOLO26-l	0,849	0,856	0,798	0,664	23	57	51,77	19,32
YOLO26-x	0,847	0,855	0,803	0,685	16	63	86,14	11,61
YOLOv8-x	0,836	0,844	0,805	0,707	23	63	100,45	9,96
YOLOv8-l	0,832	0,840	0,809	0,711	20	67	69,00	14,49
YOLOv8-m	0,824	0,832	0,815	0,734	11	77	43,81	22,83
YOLOv8-s	0,820	0,828	0,811	0,724	18	74	30,91	32,35
YOLO26-s	0,816	0,824	0,812	0,710	15	78	26,27	38,07
YOLO26-m	0,812	0,820	0,817	0,729	13	81	38,32	26,10
YOLOv8-n	0,769	0,776	0,814	0,749	25	92	22,14	45,18
YOLO26-n	0,764	0,772	0,819	0,767	12	102	20,93	47,77

Çizelge 4.3'e göre, confidence ≥ 0,30 koşulunda en yüksek sınıf F1 ve BBox F1 değerleri YOLO26-l modelinde elde edilmiştir. YOLO26-l modeli, 0,849 sınıf F1 ve 0,856 BBox F1 değeri ile en yüksek genel doğruluk performansını üretmiştir. Ayrıca yanlış negatif sayısının diğer modellere göre daha düşük olması, bu modelin etiketli veri setindeki nesneleri kaçırma eğiliminin daha sınırlı olduğunu göstermektedir.

Buna karşılık YOLO26-s modeli, doğruluk bakımından YOLO26-l modelinin gerisinde kalmasına rağmen, görüntü başına 26,27 ms çıkarım süresiyle önemli bir hız avantajı sunmuştur. YOLO26-l modeli yaklaşık 19,32 FPS üretirken, YOLO26-s modeli yaklaşık 38,07 FPS değerine ulaşmıştır. Bu durum, YOLO26-s modelinin gerçek zamanlıya yakın dijital ikiz entegrasyonu açısından daha dengeli bir aday olduğunu göstermektedir.

YOLO26-x ve YOLOv8-x modelleri de yüksek doğruluk değerleri üretmiş olmakla birlikte, çıkarım sürelerinin daha yüksek olması nedeniyle gerçek zamanlıya yakın uygulamalar açısından daha maliyetli seçenekler olarak değerlendirilmiştir. Nano modeller ise daha hızlı çalışmasına rağmen, yanlış negatif sayılarının artması nedeniyle uzak veya küçük nesneleri kaçırma eğilimi göstermiştir. Bu nedenle yalnızca hız dikkate alındığında nano modeller avantajlı görünse de, dijital ikiz ortamında nesne sürekliliği ve konumlandırma doğruluğu açısından yeterli bulunmamıştır. YOLO model varyantlarının etiketli BBox ve IoU tabanlı doğrulamasında kullanılan ayrıntılı sonuç veri yapısı Ek 6'da sunulmuştur.

4.2.4 Confidence Eşik Değerinin Tespit Başarımına Etkisi

Confidence eşik değeri, YOLO modelinin ürettiği tahminlerin değerlendirmeye alınıp alınmayacağını belirleyen önemli bir parametredir. Model, her tespit için nesneye ait sınıf bilgisinin yanında bir güven skoru üretmektedir. Bu skor, ilgili tahminin belirlenen sınıfa ait olma olasılığını temsil etmektedir.

Düşük confidence eşiği kullanıldığında daha fazla tespit değerlendirmeye dâhil edilmekte, bu durum özellikle uzak, küçük veya kısmen örtülü nesnelerin yakalanma olasılığını artırmaktadır. Ancak eşik değerinin düşük tutulması, düşük güven skoruna sahip hatalı tahminlerin de sisteme girmesine neden olabileceğinden yanlış pozitif sayısını artırabilmektedir.

Buna karşılık yüksek confidence eşiği, düşük güven skoruna sahip tahminleri eleyerek daha seçici bir çıktı üretmektedir. Bu durumda değerlendirmeye alınan BBox

tahminlerinin ortalama örtüşme kalitesi artabilmekte, ancak bazı gerçek nesneler yeterli güven skoruna ulaşamadığı için tespit dışı kalabilmektedir. Bu durum özellikle dijital ikiz uygulamaları açısından önemlidir; çünkü tespit dışı kalan bir nesne sonraki aşamada homografi, IDW tabanlı lokal enterpolasyon veya ray-casting süreçlerine aktarılamamaktadır.

Bu çalışmada, seçilen model olan YOLO26-s için confidence ≥ 0,30 ve confidence ≥ 0,50 eşik değerleri karşılaştırılmıştır. Her iki analizde de aynı görüntüler, aynı manuel BBox etiketleri ve IoU ≥ 0,50 eşleşme koşulu kullanılmıştır. Böylece model mimarisi, test veri seti ve IoU eşiği sabit tutulmuş; yalnızca confidence eşik değerinin tespit başarımı, yanlış pozitif/yanlış negatif dengesi ve BBox örtüşme kalitesi üzerindeki etkisi değerlendirilmiştir.

Çizelge 4.4 Confidence ≥ 0,30 ve confidence ≥ 0,50 eşik değerlerinin genel performansa etkisi.

Conf.	FP	FN	Sınıf F1	BBox F1	Ortalama IoU	IoU ≥ 0,75	ms/ Görüntü	FPS
0,30	15	78	0,816	0,824	0,812	0,710	26,27	38,07
0,50	7	100	0,777	0,785	0,825	0,759	33,39	29,95

Çizelge 4.4 incelendiğinde, confidence eşiğinin 0,30'dan 0,50'ye çıkarılmasıyla yanlış pozitif sayısının 15'ten 7'ye düştüğü görülmektedir. Bu durum, daha yüksek confidence eşiğinin düşük güven skoruna sahip hatalı tespitleri azalttığını göstermektedir. Buna karşılık yanlış negatif sayısı 78'den 100'e yükselmiştir. Bu sonuç, confidence eşiği artırıldığında modelin daha seçici hâle geldiğini, ancak bazı gerçek nesneleri de tespit dışı bıraktığını göstermektedir.

F1-skor değerleri de bu durumu desteklemektedir. Confidence ≥ 0,30 koşulunda sınıf F1 değeri 0,816 ve BBox F1 değeri 0,824 iken, confidence ≥ 0,50 koşulunda bu değerler sırasıyla 0,777 ve 0,785'e düşmüştür. Buna karşılık ortalama IoU değeri 0,812'den 0,825'e, IoU ≥ 0,75 oranı ise 0,710'dan 0,759'a yükselmiştir. Bu sonuç, 0,50 eşiğinde değerlendirmeye alınan BBox tahminlerinin gerçek etiketlerle daha yüksek örtüşme sağladığını; ancak daha fazla gerçek nesnenin kaçırıldığını göstermektedir. Dijital ikiz uygulaması açısından yanlış negatif hatalar kritik öneme sahiptir. Çünkü yanlış negatif

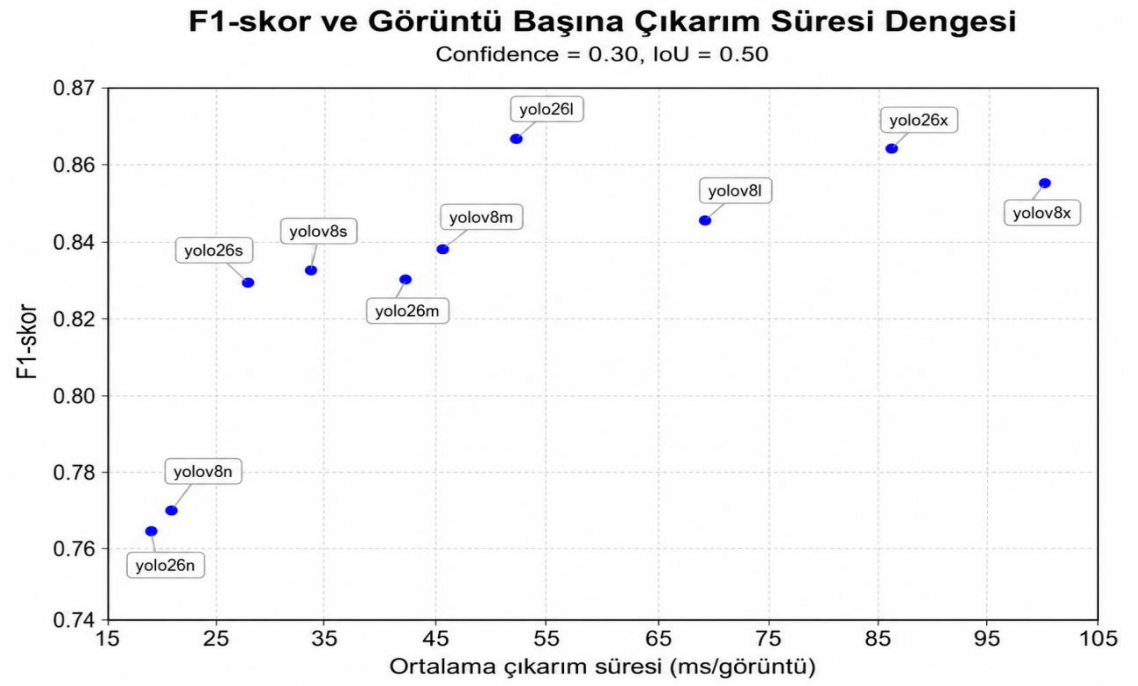
durumunda gerçek bir yaya veya araç CesiumJS ortamında hiç temsil edilmemektedir. Bu nedenle nesne sürekliliğinin korunması ve uzak/küçük hedeflerin tespit dışı kalmaması amacıyla YOLO26-s modeli için confidence ≥ 0,30 eşiği daha uygun bulunmuştur. Ayrıca bu eşikte modelin daha yüksek FPS değerine ulaşması, gerçek zamanlıya yakın dijital ikiz uygulaması açısından bu seçimi desteklemektedir.

4.2.5 Model Seçimi ve Dijital İkiz Uygulamasına Uygunluk Değerlendirmesi

Gerçekleştirilen analizler birlikte değerlendirildiğinde, model seçiminin yalnızca en yüksek F1-skor değerine göre yapılmasının yeterli olmadığı görülmektedir. Bu tez kapsamında nesne tespit modeli, yalnızca görüntü üzerinde nesne tespiti yapmak için değil, tespit edilen nesnelerin homografi, IDW tabanlı lokal enterpolasyon ve ray-casting yöntemleriyle gerçek dünya koordinatlarına dönüştürülerek 3B dijital ikiz ortamında görselleştirilmesi amacıyla kullanılmaktadır. Bu nedenle modelin tespit başarımının yanı sıra çıkarım süresi, BBox konum doğruluğu ve küçük/uzak nesnelerdeki tespit sürekliliği de dikkate alınmıştır.

Etiketli BBox doğrulama sonuçlarına göre en yüksek sınıf F1 ve BBox F1 değerleri YOLO26-l modelinde elde edilmiştir. Bu nedenle YOLO26-l, doğruluk öncelikli kullanım senaryolarında en başarılı model olarak değerlendirilebilir. Ancak gerçek zamanlıya yakın dijital ikiz entegrasyonunda çıkarım süresi de belirleyici olduğundan, YOLO26-s yaklaşık iki kat daha yüksek FPS üretmesiyle hız-doğruluk dengesi bakımından daha uygun bir aday olarak öne çıkmıştır.

Bu kapsamda nihai model değerlendirmesi iki ölçüt üzerinden yapılmıştır: doğruluk açısından YOLO26-l en yüksek performansı üretirken, gerçek zamanlıya yakın uygulama gereksinimi açısından YOLO26-s daha dengeli bir çözüm sunmuştur. Bu nedenle sonraki koordinat dönüşüm ve CesiumJS entegrasyonu analizlerinde temel model olarak YOLO26-s kullanılmış; YOLO26-l ise doğruluk odaklı karşılaştırmalarda referans başarı düzeyi olarak dikkate alınmıştır. Model varyantlarının F1-skor ve çıkarım süresi dengesi Şekil 4.7'de verilmiştir.



Şekil 4.7 Model varyantlarının F1-skor ve çıkarım süresi dengesi.

Şekil 4.7 incelendiğinde, model seçiminde ideal bölgenin yüksek F1-skor ve düşük çıkarım süresi değerlerini birlikte sağlayan bölge olduğu görülmektedir. YOLO26-l modeli doğruluk açısından en yüksek sonucu üretirken, YOLO26-s modeli daha düşük çıkarım süresiyle gerçek zamanlıya yakın uygulamalar açısından daha avantajlı bir konumda yer almaktadır. Bu nedenle sistemin nihai kullanım amacı dikkate alınarak YOLO26-s modeli hız-doğruluk dengesi bakımından tercih edilmiştir.

Sonuç olarak, nesne tespit modeli seçiminde üç temel bulguya ulaşılmıştır. İlk olarak, sayım tabanlı ön değerlendirme YOLO26-s modelinin yüksek tespit başarımı ve düşük işlem süresi nedeniyle güçlü bir aday olduğunu göstermiştir. İkinci olarak, etiketli BBox ve IoU tabanlı doğrulama doğruluk açısından YOLO26-l modelinin en yüksek performansı sunduğunu ortaya koymuştur. Üçüncü olarak, confidence eşik değeri analizinde 0,30 eşiğinin 0,50 eşiğine göre daha düşük yanlış negatif üretmesi nedeniyle dijital ikiz uygulaması için daha uygun olduğu görülmüştür. Bu bulgular doğrultusunda, gerçek zamanlıya yakın dijital ikiz entegrasyonu ve koordinat dönüşüm analizlerinde YOLO26-s modeli temel model olarak kullanılmıştır.

4.3 Konum Doğruluğu Analizleri

Bu bölümde, sabit kamera görüntülerinden YOLO26-s modeli ile tespit edilen hareketli nesnelerin görüntü düzlemindeki piksel koordinatlarından gerçek dünya koordinatlarına dönüştürülmesi süreci, RTK ölçümlü yaya örnekleri üzerinden değerlendirilmiştir. Nesne tespiti aşamasında elde edilen BBox çıktıları, yalnızca görüntü üzerinde nesnenin konumunu göstermekte; dijital ikiz ortamında kullanılabilmeleri için bu noktaların coğrafi koordinat sistemine aktarılması gerekmektedir. Bu nedenle çalışmanın bu aşamasında, görüntü tabanlı tespit sonuçlarından elde edilen tahmini konumlar ile RTK ölçümleri arasındaki konumsal farklar analiz edilmiştir.

Konumsal doğruluk analizlerinde referans veri olarak arazide RTK yöntemiyle ölçülen yaya konumları kullanılmıştır. Konumsal doğruluk analizlerinde kullanılan RTK referans noktalarına ait veri yapısı Ek 2'de sunulmuştur. Her bir görüntü için ilgili yayanın WGS 84 koordinatları ve yükseklik bilgisi kaydedilmiş, hata analizlerinin metrik düzlemde yapılabilmesi amacıyla koordinatlar UTM Zone 36N sistemine dönüştürülmüştür. Böylece farklı yöntemlerden elde edilen tahmini koordinatlar ile RTK referans koordinatları arasındaki yatay konum hatası metre cinsinden hesaplanmıştır.

Bu kapsamda dört farklı koordinat dönüşüm yaklaşımı karşılaştırılmıştır. Bunlar; homografi dönüşümü, IDW tabanlı lokal enterpolasyon, LiDAR tabanlı ray-casting ve oblik fotogrametri modeli üzerinde gerçekleştirilen ray-casting yöntemleridir. Homografi ve IDW tabanlı lokal enterpolasyon yöntemleri, görüntü düzlemi ile arazi koordinatları arasında iki boyutlu ilişki kurarken; ray-casting yöntemleri CesiumJS ortamındaki 3B dijital ikiz modeli üzerinde kamera ışınının yüzeyle kesiştiği noktayı esas almaktadır. Bu yönüyle ray-casting yöntemi, sahnenin üç boyutlu geometrisini dikkate alması bakımından diğer iki yaklaşımdan ayrılmaktadır.

Konum dönüşümünde kullanılan piksel referans noktasının sonuçlara etkisini incelemek amacıyla her bir BBox için iki farklı nokta değerlendirilmiştir. Birincisi BBox merkezi, ikincisi ise BBox alt orta noktasıdır. BBox merkezi, nesnenin görüntüdeki geometrik merkezini temsil ederken; BBox alt orta noktası, yayanın zeminle temas ettiği ayak bölgesine daha yakın bir referans noktası olarak kabul edilmiştir.

Dijital ikiz ortamında amaç yayanın görüntüdeki gövde merkezini değil, zemindeki konumunu temsil etmek olduğundan, BBox referans noktası seçimi konumsal doğruluk üzerinde doğrudan etkili olmaktadır.

Bu analizlerde temel değerlendirme ölçütü, tahmini konum ile RTK referans noktası arasındaki yatay uzaklıktır. Bunun yanında yöntemlerin genel başarımını ortaya koymak için ortalama hata, RMSE, maksimum hata ve hata dağılımları da incelenmiştir. Ayrıca sahne koşullarının konum doğruluğuna etkisini değerlendirmek amacıyla görüntüler yakın, orta ve uzak mesafe gruplarına; görünürlük açısından ise açık ve örtülü sahnelere ayrılmıştır.

Böylece koordinat dönüşüm yöntemleri yalnızca genel ortalama hata değerleri üzerinden değil, aynı zamanda mesafe, görünürlük ve BBox referans noktası değişkenleri üzerinden de karşılaştırılmıştır.

4.3.1 Homografi, IDW Tabanlı Lokal Enterpolasyon ve Ray-Casting Tabanlı Koordinat Dönüşüm Yaklaşımları

Görüntü düzleminde elde edilen nesne tespit sonuçlarının dijital ikiz ortamında kullanılabilmesi için, BBox referans noktasının arazi üzerindeki karşılığının belirlenmesi gerekmektedir. Bu çalışmada bu amaçla kullanılan yöntemler, temel olarak iki farklı yaklaşıma dayanmaktadır. Homografi ve IDW tabanlı lokal enterpolasyon yöntemleri, görüntü koordinatları ile arazi koordinatları arasında iki boyutlu bir ilişki kurarken; ray-casting yöntemi kamera geometrisi ve 3B model yüzeyi üzerinden doğrudan ışın-yüzey kesişimini esas almaktadır.

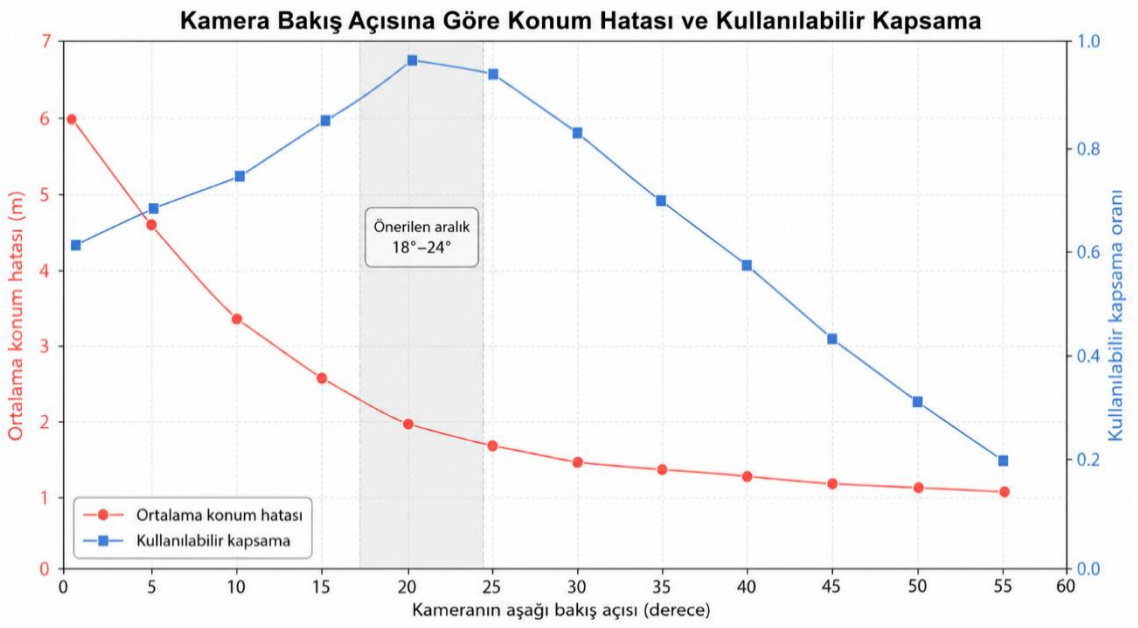
Homografi yaklaşımında, görüntü üzerinde koordinatı bilinen referans noktaları ile bu noktaların arazi üzerindeki gerçek dünya koordinatları kullanılarak projektif bir dönüşüm matrisi hesaplanmaktadır. Bu dönüşüm, görüntü düzlemi ile yer düzlemi arasında matematiksel bir ilişki kurduğu için hızlı ve uygulanabilir bir yöntemdir. Ancak yöntemin temel sınırlılığı, çalışma bölgesinin yaklaşık düzlemsel kabul

edilmesine dayanmasıdır. Bu nedenle arazi eğimi, kot farkı, perspektif bozulması veya referans noktalarının hedef bölgeyi yeterince temsil etmemesi durumunda homografi sonuçlarında konumsal sapmalar oluşabilmektedir.

IDW tabanlı lokal enterpolasyon yönteminde ise tek bir genel dönüşüm matrisi yerine, hedef piksel noktasına yakın referans noktalarının konumsal ilişkisi dikkate alınmaktadır. Bu yaklaşımda, tahmin edilecek piksel noktasına daha yakın olan referans noktaları daha yüksek ağırlık almakta, uzak noktaların etkisi ise azalmaktadır. Böylece hedef noktanın yakın çevresindeki referans noktalarından yararlanılarak yerel bir konum kestirimi yapılmaktadır. Bununla birlikte, referans noktalarının az olduğu, belirli bölgelerde kümelendiği veya hedef noktanın referans noktalarından uzaklaştığı durumlarda enterpolasyon hatası artabilmektedir. Bu nedenle lokal enterpolasyonun başarımı, doğrudan referans noktalarının sayısına, görüntü içerisindeki dağılımına ve hedef nesneye olan uzaklığına bağlıdır.

Ray-casting yaklaşımı, diğer iki yöntemden farklı olarak 3B sahne geometrisini doğrudan kullanmaktadır. Bu yöntemde kamera konumu, yönelimi, görüntü çözünürlüğü, FOV ve görüntü oranı dikkate alınarak BBox referans noktasından geçen bir kamera ışını oluşturulmaktadır. Bu ışının CesiumJS ortamındaki 3B yüzey modeliyle kesiştiği nokta, nesnenin tahmini arazi konumu olarak kabul edilmektedir. Çalışmada bu işlem hem LiDAR tabanlı nokta bulutu hem de oblik fotogrametriyle üretilen 3B mesh model üzerinde uygulanmış ve farklı 3B altlıkların ray-casting tabanlı konum kestirimine etkisi karşılaştırılmıştır.

Ray-casting yönteminin başarımı, 3B model doğruluğunun yanında kamera geometrisine de bağlıdır. Kamera yataya yakın baktığında, BBox üzerindeki küçük bir piksel sapması zeminde daha büyük konumsal hataya dönüşebilmektedir. Bunun nedeni, görüş ışınının zemine düşük açıyla yaklaşması ve kesişim noktasının kamera doğrultusunda daha duyarlı hâle gelmesidir. Kamera daha aşağı yönlendirildiğinde ışın-zemin kesişimi daha kararlı olmakta ve piksel hatasının metre cinsinden etkisi azalmaktadır. Ancak çok dik aşağı bakışta kapsama alanı daraldığından sistemin izleme kapasitesi sınırlanabilmektedir. Bu ilişki Şekil 4.8'de kavramsal olarak gösterilmiştir.



Şekil 4.8 Kamera aşağı bakış açısının konum hatası ve kullanılabilir kapsama oranı grafiği.

Şekil 4.8'de görüldüğü üzere, kameranın aşağı bakış açısı arttıkça ortalama konum hatasının azalma eğiliminde olduğu, buna karşılık kullanılabilir kapsama oranının belirli bir açıdan sonra düşmeye başladığı görülmektedir. Bu nedenle kamera yerleşiminde yalnızca en düşük konum hatasını veren açı değil, konum doğruluğu ile sahne kapsama oranı arasındaki denge dikkate alınmalıdır.

Bu çalışma kapsamında kullanılan kavramsal değerlendirme, yaklaşık 3 m kamera yüksekliği varsayımı altında yorumlanmıştır. Kamera yüksekliği arttıkça aynı pitch açısında daha geniş bir alan görülebilmekte ve ray-zemin kesişim geometrisi değişebilmektedir. Dolayısıyla önerilen açı aralığı sabit bir değer olarak değil, kamera yüksekliği, lens görüş açısı, hedef uzaklığı ve sahne geometrisine bağlı olarak optimize edilmesi gereken bir tasarım parametresi olarak değerlendirilmelidir.

Bu bağlamda yaklaşık 18°–24° aralığı, çalışmadaki kamera yüksekliği ve sahne yapısı dikkate alındığında konum hatası ile kullanılabilir kapsama oranı arasında dengeli bir çalışma aralığı olarak kabul edilmiştir. Daha yatay açılarda piksel hatasının zemindeki etkisi artarken, daha dik açılarda kapsama alanının daralması sistemin izleme performansını sınırlayabilmektedir. Bu nedenle sonraki analizlerde yalnızca yöntemlerin ortalama hata değerleri değil, BBox referans noktası, sahne uzaklığı ve

görünürlük durumu da birlikte değerlendirilmiştir.

4.3.2 BBox Referans Noktasının Konum Sonuçlarına Etkisi

Nesne tespit algoritmalarından elde edilen BBox çıktıları, görüntü düzleminde nesnenin kapladığı dikdörtgensel alanı temsil etmektedir. Ancak bu dikdörtgen alanın hangi noktasının gerçek dünya koordinatına dönüştürüleceği, konumsal doğruluk üzerinde doğrudan etkili olmaktadır.

Özellikle yaya nesnesi için dijital ikiz ortamında temsil edilmesi gereken konum, nesnenin görüntüdeki geometrik merkezi değil, zemindeki ayak veya temas noktasıdır. Bu nedenle bu çalışmada BBox merkezi ve BBox alt orta noktası olmak üzere iki farklı referans noktası karşılaştırılmıştır.
Bu analizlerde kullanılan kamera geometrisi, yaklaşık pitch = 0° kabul edilen yataya yakın bakış koşulunu temsil etmektedir. Bu durum, piksel düzlemindeki küçük sapmaların zemindeki ışın-yüzey kesişim noktasında daha büyük konumsal farklılıklara dönüşebilmesi nedeniyle, yöntemler açısından zorlayıcı bir test durumu oluşturmaktadır. Bu nedenle elde edilen sonuçlar, özellikle yataya yakın kamera bakışında BBox referans noktası seçiminin konum doğruluğu üzerindeki etkisini değerlendirmek amacıyla yorumlanmıştır.

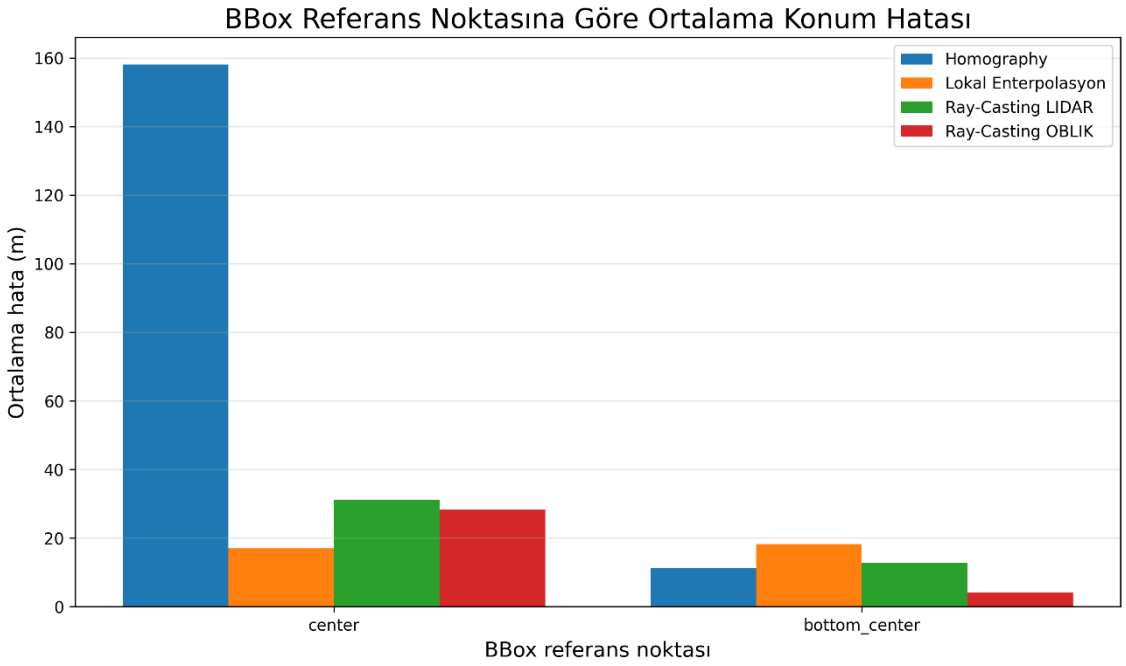
BBox merkezi, nesnenin görüntü üzerindeki geometrik orta noktasını ifade etmektedir. Bu nokta, yayanın yaklaşık gövde merkezine karşılık geldiğinden, doğrudan zemin üzerindeki konumu temsil etmemektedir. Kamera görüntüsünde yaya dikey bir nesne olarak algılandığı için, BBox merkezinden üretilen ışın çoğu zaman zemindeki ayak noktasına değil, nesnenin gövde yüksekliğine karşılık gelen doğrultuya yönelmektedir.

Bu durum özellikle ray-casting yaklaşımında, ışının beklenen ayak noktasından farklı bir yüzeyle kesişmesine veya zeminde daha uzak bir noktaya düşmesine neden olabilmektedir. Homografi ve IDW tabanlı lokal enterpolasyon yöntemlerinde de BBox merkezinin kullanılması, görüntü düzlemindeki zemin temsili ile nesne gövdesi arasında geometrik uyumsuzluk oluşturabilmektedir.

BBox alt orta noktası ise nesnenin görüntüdeki alt sınırının orta noktasını temsil etmektedir. Yaya tespiti özelinde bu nokta, kişinin zeminle temas ettiği ayak bölgesine daha yakın kabul edilmektedir. Bu nedenle gerçek dünya koordinatına dönüştürülecek referans nokta olarak BBox alt orta noktasının kullanılması, özellikle zemin konumu tahmini açısından daha anlamlıdır.

Bu yaklaşım, yayanın dijital ikiz ortamında zemine oturtulması ve konumunun RTK ölçümüyle karşılaştırılması için daha uygun bir temsil sağlamaktadır.

Şekil 4.9'da BBox merkezi ve BBox alt orta noktası kullanılarak elde edilen ortalama konum hataları yöntem bazında karşılaştırılmıştır.



Şekil 4.9 BBox referans noktasına göre ortalama konum hatası karşılaştırması.

Şekil 4.9 incelendiğinde, BBox merkezi kullanıldığında bazı yöntemlerde ortalama konum hatasının belirgin biçimde arttığı görülmektedir. Bu artış özellikle homografi yönteminde daha belirgin hâle gelmiştir. Bunun temel nedeni, BBox merkezinin zemindeki gerçek yaya konumunu değil, yayanın görüntü üzerindeki gövde merkezini temsil etmesidir. Böyle bir noktanın zemin koordinatına dönüştürülmesi, özellikle perspektif etkisinin yüksek olduğu ve kameranın yataya yakın baktığı sahnelerde önemli

sapmalara yol açabilmektedir.

BBox alt orta noktası kullanıldığında ise genel hata değerlerinin önemli ölçüde azaldığı görülmektedir. Bu sonuç, yaya konumunun dijital ikiz ortamında temsil edilmesinde BBox alt orta noktasının daha uygun bir referans olduğunu göstermektedir. Özellikle ray-casting Oblik yönteminde BBox alt orta noktasının kullanılmasıyla hata değerlerinin diğer yöntemlere göre daha düşük seviyede kaldığı gözlenmiştir. Bu durum, oblik fotogrametri modeli üzerinde yapılan ışın-yüzey kesişiminin, yayanın zemindeki konumunu temsil etmede daha kararlı sonuçlar üretebildiğini göstermektedir.

Bununla birlikte BBox alt orta noktasının her durumda gerçek ayak noktasını kusursuz temsil ettiği söylenemez. Yaya kısmen örtülü olduğunda, ayak bölgesi görüntüde net seçilemediğinde veya YOLO tarafından üretilen BBox nesnenin gerçek alt sınırını tam olarak kapsamadığında, alt orta nokta da hatalı konum üretebilmektedir. Ayrıca kameranın bakış açısı, hedefe olan mesafe ve perspektif etkisi BBox geometrisini değiştirebildiğinden, bazı özel durumlarda ayak noktası kestirimi veya insan poz tahmini tabanlı taban noktası çıkarımı gibi yöntemlerin ayrıca değerlendirilmesi gerekebilir.

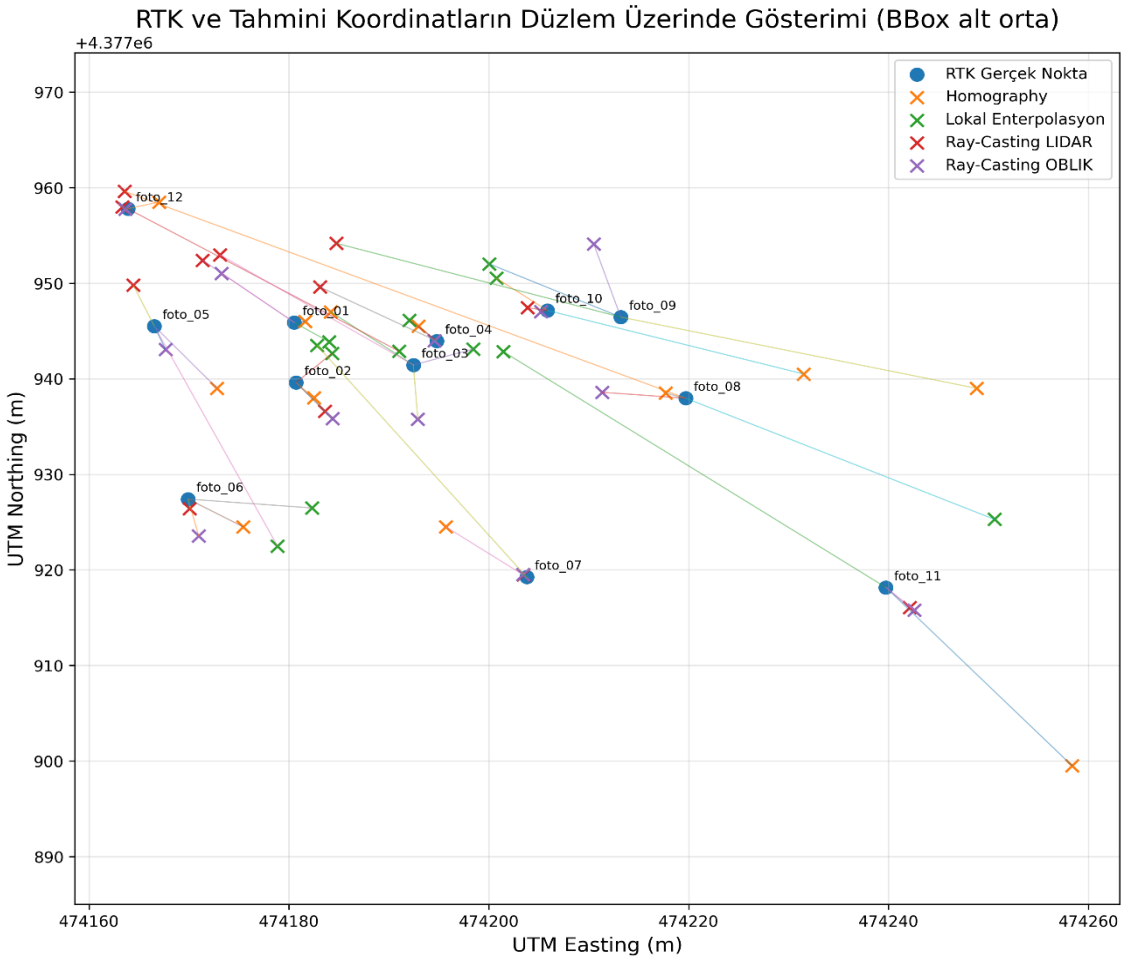
Bu bulgular doğrultusunda, sonraki konumsal doğruluk analizlerinde ana değerlendirme BBox alt orta noktası üzerinden yapılmıştır. BBox merkezi sonuçları ise referans nokta seçiminin konum doğruluğu üzerindeki etkisini göstermek amacıyla karşılaştırmalı olarak değerlendirilmiştir. Bu tercih, dijital ikiz ortamında yayanın gerçek zemin konumunun temsil edilmesi açısından daha tutarlı bir yaklaşım sunmaktadır.

4.3.3 RTK Referans Ölçümlerine Göre Yöntemlerin Konumsal Doğruluk Karşılaştırması

Konum doğruluğu analizinde, her bir görüntü için RTK yöntemiyle ölçülen yaya konumu referans değer olarak kabul edilmiştir. YOLO26-s modeli tarafından üretilen BBox çıktılarından elde edilen alt orta nokta koordinatları; homografi, IDW tabanlı lokal enterpolasyon, ray-casting LiDAR ve ray-casting Oblik yöntemleriyle gerçek

dünya koordinatlarına dönüştürülmüştür. Elde edilen tahmini koordinatlar, RTK referans koordinatları ile UTM düzleminde karşılaştırılmış ve her yöntem için yatay konum hatası metre cinsinden hesaplanmıştır.

Bu aşamada ana değerlendirme BBox alt orta noktası üzerinden yapılmıştır. Bunun nedeni, önceki bölümde de belirtildiği üzere, yaya konumunun dijital ikiz ortamında zemindeki ayak veya temas noktasıyla temsil edilmesinin daha uygun olmasıdır. BBox merkezinden elde edilen sonuçlar, özellikle perspektif etkisinin belirgin olduğu görüntülerde yüksek sapmalar üretebildiğinden, yöntemlerin genel konumsal başarımını yorumlamak için BBox alt orta noktası daha uygun bir referans noktası olarak değerlendirilmiştir. RTK referans noktaları ile yöntemlerden elde edilen tahmini koordinatların düzlem üzerindeki dağılımı Şekil 4.10'da verilmiştir.



Şekil 4.10 RTK referans noktaları ile tahmini koordinatların UTM düzleminde karşılaştırılması.

Şekil 4.10 incelendiğinde, bazı görüntülerde tahmini koordinatların RTK referans noktasına yakın konumlandığı, bazı görüntülerde ise yöntemler arasında belirgin sapmalar oluştuğu görülmektedir. Bu durum, konum dönüşüm doğruluğunun yalnızca kullanılan matematiksel yönteme değil; görüntüdeki nesne uzaklığına, BBox üretim kararlılığına, sahne geometrisine, kamera bakış açısına ve 3B model yüzeyinin temsil kalitesine de bağlı olduğunu göstermektedir.

Homografi yöntemi bazı noktalarda RTK referansına yakın sonuçlar üretmekle birlikte, özellikle perspektif etkisinin ve uzaklık değişiminin arttığı sahnelerde daha belirgin sapmalar göstermiştir. Bu durum beklenen bir sonuçtur; çünkü homografi yöntemi görüntü düzlemi ile arazi düzlemi arasında iki boyutlu projektif bir ilişki kurmaktadır. Arazi yüzeyinin tam düzlemsel olmaması, referans noktalarının hedef bölgeyi yeterince temsil etmemesi veya yayanın referans noktaları ağı dışına yaklaşması durumunda homografi dönüşümünde hata artabilmektedir.

IDW tabanlı lokal enterpolasyon yöntemi, referans noktalarına yakın bölgelerde daha dengeli sonuçlar üretebilmekle birlikte, bazı görüntülerde yüksek sapmalar göstermiştir. Bu durum, yöntemin referans noktalarının dağılımına doğrudan bağımlı olduğunu ortaya koymaktadır.

Özellikle hedef noktanın referans noktalarından uzaklaşması veya referans noktalarının belirli bölgelerde kümelenmesi, lokal enterpolasyon sonucunda tahmin edilen koordinatın RTK referansından uzaklaşmasına neden olabilmektedir.

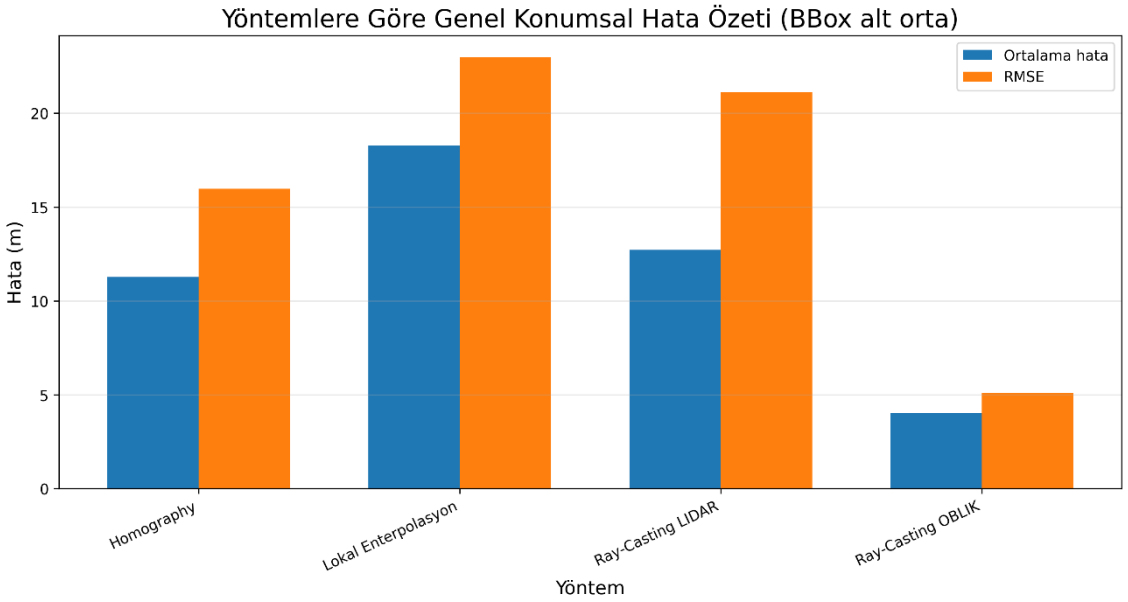
Ray-casting LiDAR yöntemi, 3B sahne geometrisini dikkate alması bakımından güçlü bir yaklaşım olmakla birlikte, bazı görüntülerde yüksek hata değerleri üretmiştir. Bu durumun temel nedenleri arasında LiDAR nokta bulutundaki yerel yüzey süreksizlikleri, ışının zemin yerine farklı bir yüzeyle ilk kesişimi, nokta yoğunluğundaki değişimler ve örtülme koşulları gösterilebilir.

Ray-casting yönteminde hesaplanan konum, kameradan çıkan ışının sahnede ilk temas ettiği yüzeye bağlı olduğundan, küçük bir BBox veya kamera yönelim hatası konum

sonucunu önemli ölçüde değiştirebilmektedir.

Bu veri setinde ray-casting Oblik yöntemi ise genel olarak daha kararlı ve düşük hata değerleri üretmiştir. Oblik fotogrametriyle elde edilen mesh modelin yüzey sürekliliği, ışının zemin veya yakın yüzeyle daha düzenli kesişmesine olanak sağlamış olabilir. Bu nedenle, özellikle BBox alt orta noktası kullanıldığında, ray-casting Oblik yaklaşımı yaya konumunun dijital ikiz ortamında temsil edilmesi açısından daha uygun sonuçlar vermiştir.

Yöntemlerin genel hata düzeylerini daha açık biçimde değerlendirmek amacıyla ortalama hata ve RMSE değerleri Şekil 4.11'de karşılaştırılmıştır.



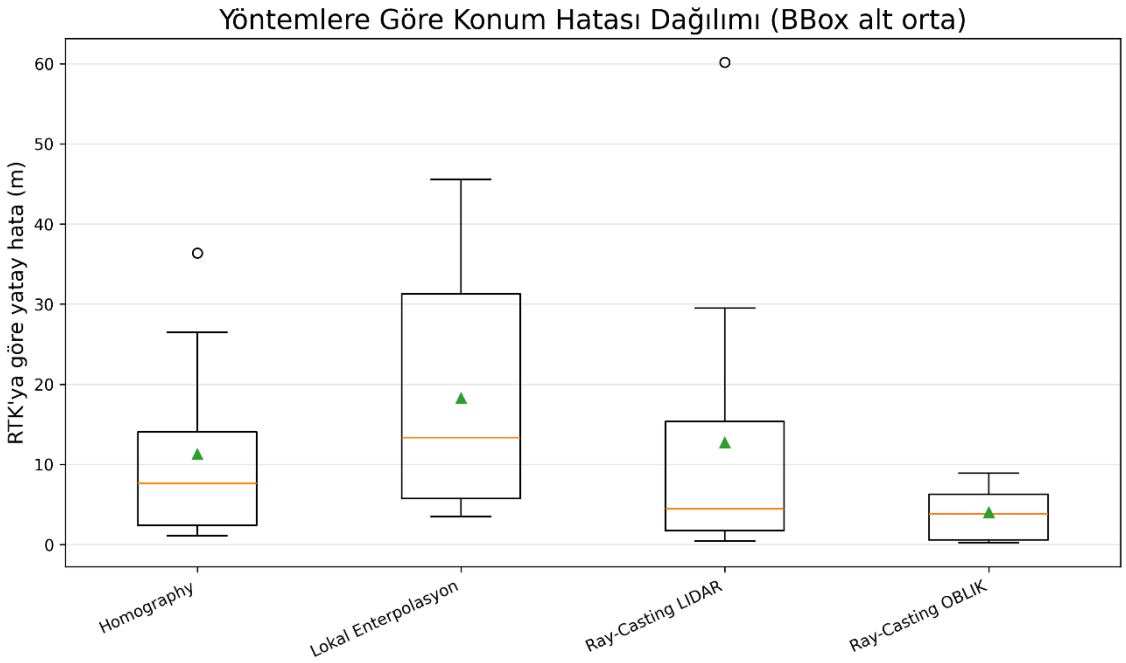
Şekil 4.11 BBox alt orta noktası için yöntemlere göre ortalama hata ve RMSE karşılaştırması.

Şekil 4.11'de görüldüğü üzere, yöntemlerin yalnızca ortalama hata değerleri değil, RMSE değerleri de dikkate alınmalıdır. Ortalama hata, yöntemlerin genel başarım düzeyini gösterirken; RMSE, yüksek sapma üreten uç değerlerin etkisini daha belirgin biçimde ortaya koymaktadır. Bu nedenle RMSE değerinin ortalama hatadan belirgin biçimde yüksek olduğu yöntemlerde, az sayıda görüntüde oluşan büyük konumsal sapmaların genel performansı etkilediği söylenebilir.

BBox alt orta noktası özelinde değerlendirildiğinde, ray-casting Oblik yönteminin

ortalama hata ve RMSE bakımından en dengeli sonuçları sunduğu görülmektedir. Bu yöntem, hem 3B sahne geometrisini dikkate almakta hem de fotogrametrik mesh yüzeyin sürekliliğinden yararlanmaktadır. Buna karşılık ray-casting LiDAR yöntemi bazı sahnelerde başarılı sonuçlar üretmesine rağmen, özellikle örtülme veya yüzey süreksizliği içeren durumlarda daha yüksek hata değerleri oluşturmuştur. Koordinat dönüşüm yöntemlerine ait fotoğraf bazlı hata analizi veri yapısı Ek 7'de verilmiştir. Homografi ve IDW tabanlı lokal enterpolasyon yöntemleri ise uygulama kolaylığı bakımından avantajlı olmakla birlikte, 3B yüzey bilgisini doğrudan kullanmadıkları için sahne geometrisindeki değişimlere daha duyarlıdır. Bu durum, özellikle farklı uzaklık gruplarının ve değişken arazi yüzeyinin bulunduğu görüntülerde hata değerlerinin artmasına neden olmuştur.

Ortalama hata ve RMSE değerleri yöntemlerin genel başarımını göstermesine karşın, hata değerlerinin dağılımını ve uç sapmaları açıklamak için tek başına yeterli değildir. Bu nedenle BBox alt orta noktası için yöntemlere göre hata dağılımları ayrıca incelenmiş ve Şekil 4.12'de verilmiştir.



Şekil 4.12 BBox alt orta noktası için yöntemlere göre konum hatası dağılımı.

Şekil 4.12 incelendiğinde, yöntemlerin yalnızca ortalama hata değerleri bakımından değil, hata dağılımlarının genişliği ve uç değer üretme eğilimleri bakımından da

farklılaştığı görülmektedir.

Daha dar hata dağılımı, yöntemin farklı görüntülerde daha kararlı sonuç ürettiğini göstermektedir. Buna karşılık geniş kutu aralığı veya yüksek uç değerler, yöntemin bazı sahne koşullarında ani sapmalar oluşturabildiğine işaret etmektedir. Bu nedenle yöntem seçimi yapılırken ortalama hata, RMSE ve hata dağılımı birlikte değerlendirilmiştir.

Homografi ve IDW tabanlı lokal enterpolasyon yöntemleri bu çalışmada yatay koordinat kestirimi açısından değerlendirilmiştir. Buna karşılık ray-casting yaklaşımı, 3B model yüzeyiyle doğrudan kesişim kurduğu için yatay konumla birlikte yükseklik bileşenini de üretmiştir. Bu durum, özellikle eğimli ve kot farkı içeren alanlarda ray-casting yönteminin 3B dijital ikiz yüzeyiyle daha doğrudan ilişki kurmasını sağlamaktadır. Konumsal hata, RMSE ve yöntem bazlı özet değerlerin hesaplanmasına ilişkin örnek kod kesiti Ek 14'te verilmiştir.

Bu bulgular, dijital ikiz ortamında gerçek dünya koordinatı üretilecekse, yöntemin yalnızca hızlı çalışmasının yeterli olmadığını göstermektedir. Kullanılan dönüşüm yaklaşımının, BBox referans noktasının ve 3B altlık kalitesinin birlikte değerlendirilmesi gerekmektedir. Bu kapsamda, BBox alt orta noktası ile birlikte kullanılan ray-casting Oblik yaklaşımı, bu veri setinde konumsal doğruluk bakımından en dengeli yöntem olarak öne çıkmıştır.

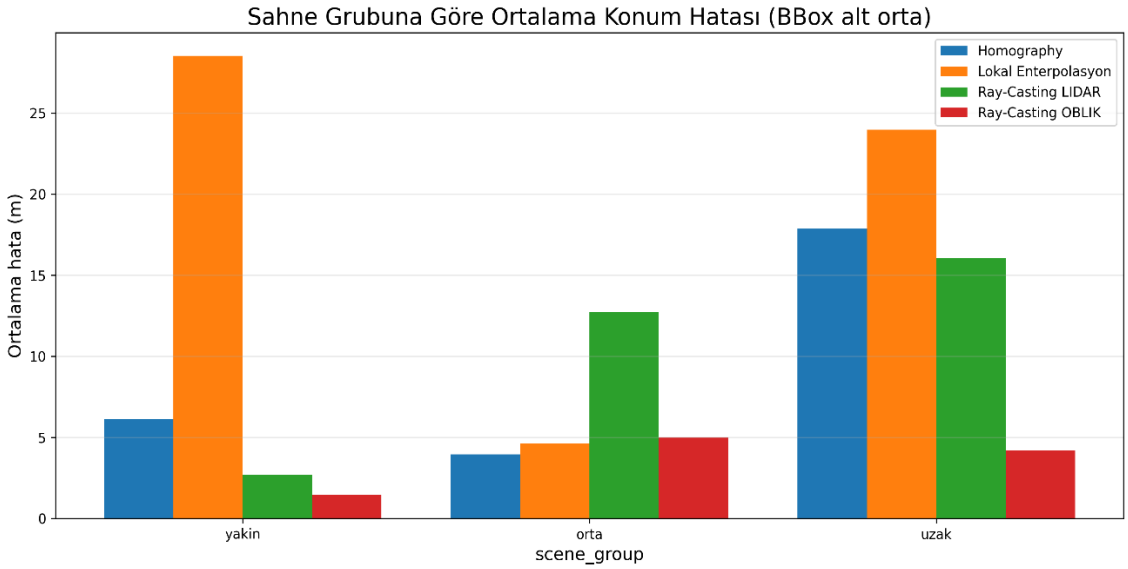
4.3.4 Mesafe, Görünürlük ve Fotoğraf Bazlı Hata Davranışının Değerlendirilmesi

Önceki bölümde yöntemlerin genel konumsal doğruluk performansları RTK referans ölçümleri temel alınarak değerlendirilmiştir. Ancak konum hatası yalnızca kullanılan dönüşüm yöntemine bağlı değildir. Hedefin kameraya olan uzaklığı, görünürlük durumu, BBox kararlılığı, kamera bakış geometrisi ve 3B model yüzeyinin ilgili bölgedeki temsil kalitesi de hata değerlerini doğrudan etkileyebilmektedir.

Özellikle uzak mesafelerde nesnenin görüntü düzleminde kapladığı piksel alanı azalmakta, bu nedenle BBox alt sınırındaki küçük değişimler arazi üzerinde daha büyük

konumsal sapmalara dönüşebilmektedir. Benzer şekilde, yayanın ayak bölgesinin örtülmesi veya BBox alt orta noktasının gerçek zemin temas noktasını temsil edememesi, tüm koordinat dönüşüm yaklaşımlarında hata değerlerini artırabilmektedir.

Bu nedenle bu bölümde, konum hatalarının mesafe gruplarına, örtülme durumuna ve fotoğraf bazlı değişimine göre nasıl farklılaştığı incelenmiştir. İlk olarak görüntüler yakın, orta ve uzak sahne gruplarına ayrılarak yöntemlerin ortalama konum hataları karşılaştırılmıştır. Bu gruplandırma, yöntemlerin yalnızca genel ortalama hata değerleriyle değil, farklı sahne koşullarındaki davranışlarıyla birlikte değerlendirilmesini sağlamaktadır. Sahne grubuna göre elde edilen hata değerleri Şekil 4.13'te verilmiştir.



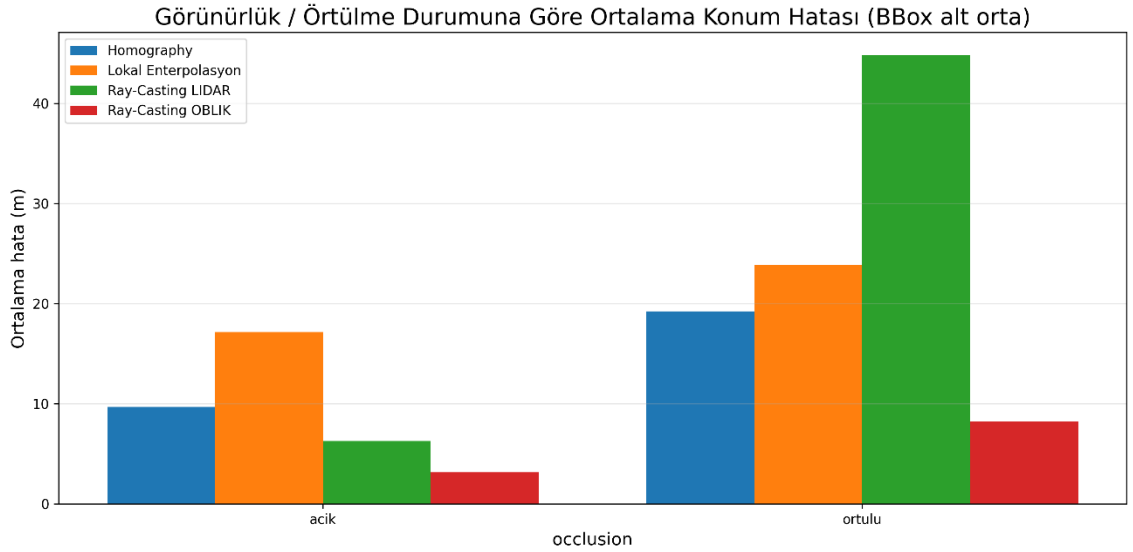
Şekil 4.13 Sahne grubuna göre ortalama konum hatası karşılaştırması.

Şekil 4.13 incelendiğinde, sahne uzaklığının konumsal hata üzerinde belirgin bir etkisi olduğu görülmektedir. Yakın ve orta mesafeli görüntülerde BBox alt orta noktasının zemindeki yaya konumunu temsil etme başarısı daha yüksek olmakta, buna bağlı olarak yöntemler genel olarak daha düşük hata değerleri üretmektedir. Buna karşılık uzak sahnelerde yayanın görüntü düzleminde kapladığı piksel alanı azaldığından, BBox alt sınırında oluşabilecek küçük bir piksel sapması arazi üzerinde daha büyük bir metrik hataya dönüşebilmektedir.

Bu durum görüntü tabanlı koordinat dönüşümü açısından önemlidir. Çünkü uzak

mesafede aynı piksel hatası, yakın mesafeye göre daha büyük bir yatay konum farkı oluşturabilmektedir. Homografi ve IDW tabanlı lokal enterpolasyon yöntemleri bu durumda referans noktalarının sahne üzerindeki temsil gücüne daha fazla bağımlı hâle gelirken, ray-casting yöntemlerinde ise ışının zemin veya 3B yüzey ile kesiştiği nokta daha hassas hâle gelmektedir. Bu nedenle uzak sahnelerde hem BBox üretim kararlılığı hem de kamera geometrisi konumsal doğruluğu doğrudan etkilemektedir. Mesafe etkisine ek olarak, görüntüler açık ve örtülü sahne koşullarına göre de değerlendirilmiştir.

Görünürlük ve örtülme durumuna göre yöntemlerin ortalama hata değerleri Şekil 4.14'te sunulmuştur.



Şekil 4.14 Görünürlük/örtülme durumuna göre ortalama konum hatası karşılaştırması.

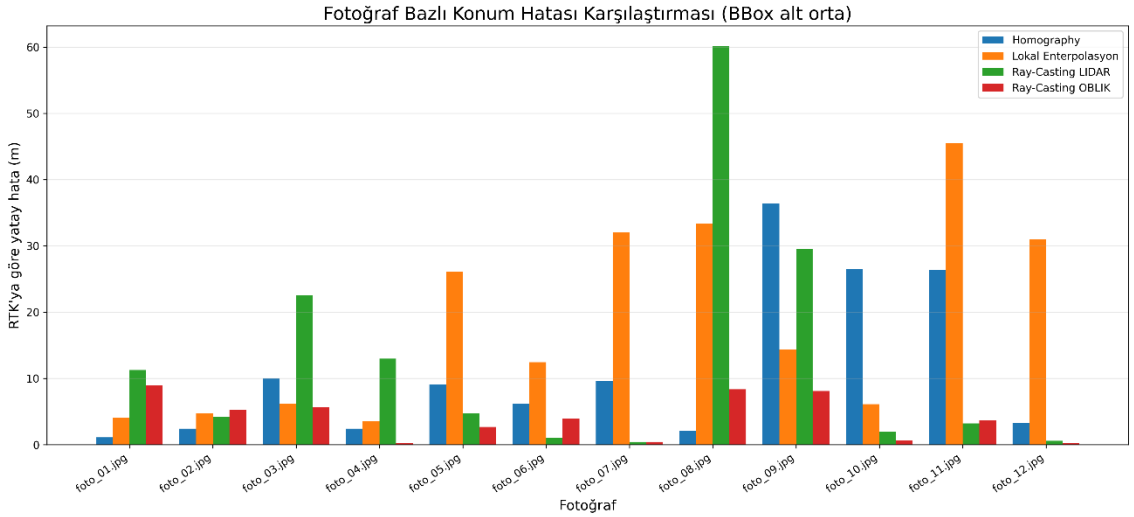
Şekil 4.14'e göre, örtülü sahnelerde hata değerlerinin genel olarak arttığı görülmektedir. Bunun temel nedeni, yayanın ayak veya alt gövde bölgesinin görüntüde tam olarak seçilememesi ve YOLO tarafından üretilen BBox alt sınırının gerçek zemin temas noktasını her zaman doğru temsil edememesidir. BBox alt orta noktasının hatalı belirlenmesi, sonraki koordinat dönüşüm aşamasında doğrudan konum hatasına dönüşmektedir.

Örtülme durumu ray-casting yöntemleri için ayrıca önem taşımaktadır. Ray-casting yaklaşımında kamera merkezinden gönderilen ışının 3B sahnede ilk temas ettiği yüzey,

tahmini koordinat olarak kabul edilmektedir. Eğer BBox alt orta noktası gerçek ayak noktasına değil de örtücü bir nesneye, zeminden farklı bir yüzeye veya 3B modeldeki süreksiz bir bölgeye yönelirse, ray beklenen zemin noktası yerine farklı bir yüzeyle kesişebilir. Bu durum özellikle nokta bulutu tabanlı LiDAR modelinde yerel yüzey boşlukları veya yoğunluk değişimleri nedeniyle daha belirgin hâle gelebilmektedir.

Oblik fotogrametri tabanlı mesh modelde ise yüzey sürekliliğinin daha yüksek olması, bazı sahnelerde ray-yüzey kesişiminin daha kararlı gerçekleşmesine katkı sağlamış olabilir. Bununla birlikte örtülme durumlarında yalnızca 3B model kalitesi değil, BBox referans noktasının doğruluğu da belirleyici olduğundan, bu tür sahnelerde tüm yöntemlerin hata üretme olasılığı artmaktadır.

Mesafe ve görünürlük gruplarına ek olarak, her bir fotoğrafta oluşan hata davranışı ayrıca incelenmiştir. Fotoğraf bazlı konum hatalarının yöntemlere göre karşılaştırılması Şekil 4.15'te verilmiştir.



Şekil 4.15 Fotoğraf bazlı konum hatalarının yöntemlere göre karşılaştırılması.

Şekil 4.15, konum hatasının görüntüler arasında homojen dağılmadığını göstermektedir. Bazı fotoğraflarda yöntemlerin hata değerleri birbirine yakın seyrederken, bazı fotoğraflarda belirli yöntemlerin daha yüksek sapma ürettiği görülmektedir. Bu durum, genel ortalama hata değerlerinin tek başına yeterli olmadığını; sahne bazlı hata davranışının da ayrıca değerlendirilmesi gerektiğini ortaya koymaktadır.

Fotoğraf bazlı değerlendirme, özellikle uç hata üreten görüntülerin belirlenmesi açısından önemlidir. Belirli bir yöntemin genel ortalama hatası düşük olsa bile, bazı sahnelerde yüksek sapmalar üretmesi sistemin kararlılığını olumsuz etkileyebilir. Dijital ikiz ortamında hareketli nesne konumlarının sürekliliği hedeflendiğinden, yalnızca ortalama hata değil, görüntüler arasında hata değerlerinin ne kadar değiştiği de dikkate alınmalıdır.

Bu bulgular birlikte değerlendirildiğinde, konum doğruluğunun yalnızca seçilen koordinat dönüşüm yöntemiyle açıklanamayacağı görülmektedir. Hedefin kameraya uzaklığı, görünürlük durumu, BBox alt orta noktasının ayak bölgesini temsil etme başarısı ve 3B modelin yerel yüzey kalitesi konum hatasını birlikte belirlemektedir. Özellikle ray-casting tabanlı yöntemlerde, kamera merkezinden BBox referans noktasına gönderilen ışının 3B model üzerinde ilk kesiştiği yüzey tahmini koordinat olarak kabul edildiğinden, örtülme veya yanlış yüzey kesişimi durumlarında hata değerleri artabilmektedir.

Bu nedenle ray-casting yöntemi, 3B dijital ikiz üzerinde güçlü bir geometrik konumlandırma yaklaşımı sunmakla birlikte, sahne koşullarına bağlı hata kaynaklarına duyarlı bir yöntem olarak değerlendirilmelidir.

Yayanın ayak bölgesinin görünmemesi, BBox alt sınırının hatalı oluşması, LiDAR nokta bulutunda yerel yüzey boşlukları bulunması veya oblik mesh modelde yüzey bozulmaları oluşması, ray'in beklenen zemin noktası yerine farklı bir yüzeyle kesişmesine neden olabilmektedir. Bu durum, özellikle örtülü sahnelerde ve fotoğraf bazlı hata grafiğinde görülen ani sapmaları açıklayan temel etkenlerden biridir.

CesiumJS ortamında zaman etiketli konumların ardışık biçimde tanımlanmasıyla, hareketli nesnelerin konumları arasında görsel süreklilik sağlanabilmektedir. Ancak bu yapı, sisteme gönderilen konum ölçümlerinin doğruluğunu kendiliğinden artırmamaktadır. Hatalı veya sıçramalı koordinatlar kullanıldığında nesne yalnızca bu hatalı konumlar arasında akıcı biçimde hareket etmektedir. Bu nedenle gerçek zamanlıya yakın dijital ikiz uygulamalarında CesiumJS ortamına aktarılacak

koordinatların öncelikle zamansal olarak yumuşatılması önem taşımaktadır.

Bu sonuçlar, konum dönüşümünde en düşük hata değerinin tek başına yeterli ölçüt olmadığını; yöntemin sahne koşullarına karşı kararlılığının da değerlendirilmesi gerektiğini göstermektedir. Özellikle ray-casting tabanlı yaklaşımlarda bazı sahnelerde oluşan ani konum sapmaları, hareketli nesne takibi için zamansal yumuşatma ihtiyacını ortaya koymaktadır. Çünkü ardışık karelerdeki konum dalgalanmaları, dijital ikiz ortamında nesnenin titreşimli veya süreksiz hareket ediyormuş gibi algılanmasına neden olabilmektedir. Bu nedenle bir sonraki bölümde, konumların Kalman filtreleme ile yumuşatılması ve yörünge kararlılığı üzerindeki etkisi incelenmiştir.

4.4 Kalman Filtreleme ile Konum Yumuşatma ve Yörünge Kararlılığı Analizi

Önceki bölümlerde görüntü tabanlı konum kestiriminin, nesne tespiti sonucunda elde edilen BBox referans noktası ve 3B yüzey modeli üzerinde gerçekleştirilen ray-casting işlemi üzerinden üretildiği açıklanmıştır. Video tabanlı nesne tespiti sonucunda elde edilen frame bazlı örnek JSON çıktı yapısı Ek 5'te verilmiştir. Bu yaklaşım, tekil görüntüler üzerinden konum üretimi açısından uygulanabilir sonuçlar vermekle birlikte, video akışı içerisinde ardışık kareler değerlendirildiğinde ham konum kestirimlerinde kısa süreli dalgalanmalar, ani konum sıçramaları ve fiziksel olarak gerçekçi olmayan hız değerleri oluşabilmektedir. Bu durum özellikle BBox alt orta noktasının her karede tam olarak aynı fiziksel noktayı temsil edememesi, görüntü çözünürlüğü ve perspektif etkileri, model yüzeyindeki küçük geometrik değişimler ve ray-casting işleminin farklı yüzey parçalarıyla kesişebilmesi nedeniyle ortaya çıkabilmektedir.

Bu bölümde, söz konusu ham konum çıktılarının zamansal süreklilik açısından daha kararlı hâle getirilmesi amacıyla sabit hız modeline dayalı doğrusal Kalman filtresi uygulanmıştır. Analizde hedef nesne sınıfı person olarak seçilmiş ve nesnenin konum temsilinde BBox alt orta noktası kullanılmıştır. Video görüntüsü üzerinden her karede nesne tespiti yapılmış, seçilen BBox noktası kamera parametreleri ile birlikte CesiumJS ortamına aktarılmış ve Oblik 3B modeli üzerinde ray-casting işlemiyle WGS 84 koordinatları elde edilmiştir. Daha sonra bu koordinatlar UTM düzlemine

dönüştürülerek ham konum serisi oluşturulmuştur.

Kalman filtreleme öncesinde, ham ray-casting sonuçlarında insan hareketiyle bağdaşmayan ani sıçramaların filtreleme sürecini olumsuz etkilemesini önlemek amacıyla fiziksel hız sınırı kullanılmıştır. Bu kapsamda yaya hareketi için 35 km/h üzerindeki ardışık konum değişimleri aykırı ölçüm olarak değerlendirilmiş ve Kalman filtresine doğrudan ölçüm olarak verilmemiştir. Bu eşik, normal yaya hareketi, hızlı yürüme ve kısa süreli koşu davranışlarını dışlamayacak; ancak ray-casting veya BBox kaynaklı gerçek dışı sıçramaların etkisini azaltacak şekilde belirlenmiştir.

Bu çalışmada kullanılan filtre, sabit hız hareket modeline dayalı doğrusal Kalman filtresidir. Filtre durum vektörü, yatay düzlemdeki Easting ve Northing koordinatları ile bu doğrultulardaki hız bileşenlerinden oluşmaktadır. Bu nedenle filtre, konumun yalnızca tekil karelerdeki değerini değil, ardışık karelerdeki hareket sürekliliğini de dikkate almaktadır. Filtreleme sürecinde süreç gürültüsü 0,5, ölçüm gürültüsü ise 25 olarak tanımlanmıştır. Süreç gürültüsünün düşük tutulması, yörüngenin ani yön ve hız değişimlerine aşırı duyarlı olmadan daha kararlı ilerlemesini sağlarken; ölçüm gürültüsünün görece yüksek seçilmesi, ham ray-casting ölçümlerindeki kısa süreli dalgalanmaların filtre sonucuna sınırlı yansımasına olanak tanımaktadır.

Bu bölümde yapılan değerlendirme üç aşamada ele alınmıştır. İlk olarak ham konum kestirimlerinde gözlenen gürültü, hız sıçraması ve kısa süreli konumsal kararsızlıklar incelenmiştir. İkinci aşamada kullanılan sabit hız modeline dayalı Kalman filtreleme yaklaşımı, aykırı değer süzme yöntemi ve uygulama süreci açıklanmıştır. Son aşamada ise ham ve Kalman filtreli konum çıktıları; yörünge sürekliliği, hız değişimi, zaman serisi davranışı ve referans doğrultuya uzaklık bakımından karşılaştırılmıştır. Böylece Kalman filtrelemenin, görüntü ve 3B model tabanlı konum kestirimlerinde zamansal kararlılık ve hareket sürekliliği üzerindeki etkisi değerlendirilmiştir.

4.4.1 Ham Konum Kestirimlerinde Gürültü ve Hız Sıçraması Problemi

Video tabanlı konum kestirimi sürecinde her kare için hedef nesne YOLO26-s modeli

ile tespit edilmiş ve nesnenin zeminle temas noktasını temsil etmesi amacıyla BBox alt orta noktası kullanılmıştır. Bu referans noktasının seçilmesindeki temel amaç, yayanın görüntü üzerindeki gövde merkezinden ziyade arazi üzerindeki ayak/temas konumuna daha yakın bir temsil elde etmektir. Her karede elde edilen BBox alt orta noktası, kamera konumu, kamera yönelimi, görüntü çözünürlüğü ve görüş açısı bilgileriyle birlikte 3B model üzerinde ray-casting işlemine girdi olarak verilmiştir.

Ray-casting sonucunda, ilgili piksel noktasından kamera merkezine bağlı olarak oluşturulan ışının Oblik 3B model yüzeyiyle kesiştiği nokta hesaplanmış ve bu noktanın WGS 84 koordinatları elde edilmiştir. Elde edilen coğrafi koordinatlar, ham konum serisinin metrik olarak değerlendirilebilmesi ve ardışık kareler arasındaki mesafe/hız değişimlerinin hesaplanabilmesi amacıyla UTM koordinat sistemine dönüştürülmüştür. Böylece video akışı boyunca her kare için Easting ve Northing bileşenlerinden oluşan ham konum serisi elde edilmiştir. Bu aşamada kullanılan temel görüntü, kamera ve analiz parametreleri Çizelge 4.5'te verilmiştir.

Çizelge 4.5 Video tabanlı ham konum kestirimi için kullanılan temel parametreler.

Parametre	Kullanılan değer
Video çözünürlüğü	1920x1080
Görüntü oranı	16:9
Tespit güven eşiği	0,30
YOLO giriş boyutu	1280
BBox referans noktası	Alt kenar orta noktası
Kamera enlemi	39,550836331186
Kamera boylamı	32,699221043914
Kamera yüksekliği	1105,2500 m
Heading (Yön)	115,3500°
Pitch (Yunuslama)	-3,7000°
Roll (Yatış)	1,3500°

Ham konum kestirimlerinde gözlenen gürültünün temel nedenlerinden biri, BBox alt orta noktasının her karede aynı fiziksel noktayı birebir temsil edememesidir. Yaya

hareketi sırasında ayak konumu, bacak açıklığı, gölge, kısmi örtülme, perspektif etkisi ve BBox sınırının kareden kareye küçük miktarlarda değişmesi, alt orta noktanın görüntü düzlemindeki yerini etkileyebilmektedir. Görüntü düzleminde birkaç piksellik küçük bir değişim, özellikle eğik kamera geometrisinde arazi üzerindeki karşılığını daha belirgin konumsal sapma olarak gösterebilmektedir.

Ham konum serisindeki bir diğer hata kaynağı, ray-casting işleminin 3B model yüzeyi ile olan kesişimidir. İdeal durumda BBox alt orta noktasından gönderilen ışının nesnenin zemin temas noktasına karşılık gelen yüzeyle kesişmesi beklenir. Ancak 3B model yüzeyindeki yerel eğim değişimleri, yüzey kırıklıkları, kaldırım ve yol kenarı gibi geometrik süreksizlikler veya model yüzeyindeki küçük farklılıklar nedeniyle bazı karelerde ışın, hedeflenen zemin yüzeyi yerine yakın çevredeki farklı bir yüzey parçası ile kesişebilmektedir. Bu durum ham koordinatlarda kısa süreli sapmalar ve ani sıçramalar oluşturabilmektedir.

Ham konum kestirimindeki bu dalgalanmalar, ardışık karelerden hız değeri üretildiğinde daha belirgin hâle gelmektedir. Bu çalışmada her kareye ait yatay yer değiştirme ve hız değeri, ardışık UTM koordinatları arasındaki mesafe ve zaman farkı kullanılarak hesaplanmıştır.

$$d_i = \sqrt{\left\{\left(E_i - E_{\{i-1\}}\right)^2 + \left(N_i - N_{\{i-1\}}\right)^2\right\}} \tag{4.1}$$

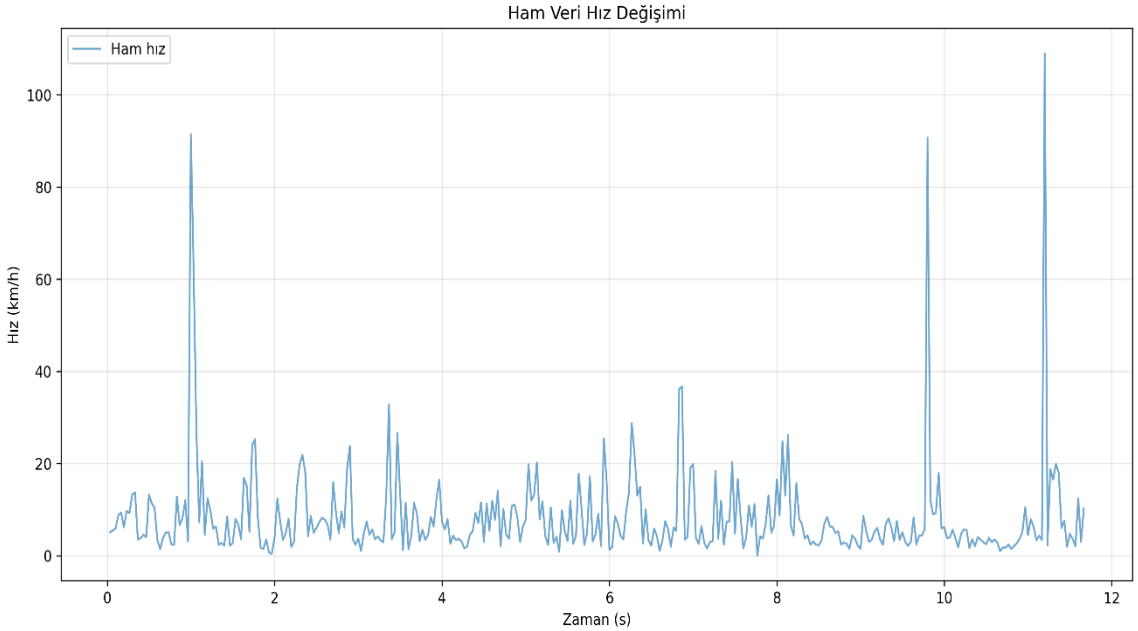
$$v_i = \frac{d_i}{t_i - t_{i-1}} * 3{,}6 \tag{4.2}$$

Burada E_i ve N_i, i. kareye ait UTM Easting ve Northing koordinatlarını; ilgili karenin zaman değerini; d_i, ardışık iki konum arasındaki yatay mesafeyi; v_i ise km/h cinsinden hız değerini ifade etmektedir. Eşitlik (4.2)'de yer alan 3,6 katsayısı, hızın m/s biriminden km/h birimine dönüştürülmesi için kullanılmıştır.

Şekil 4.16'da aykırı değer süzmesi ve Kalman filtreleme uygulanmadan önce ham konum kestirimlerinden türetilen hız değişimi verilmiştir. Grafikte görülen ani hız

pikleri, yayanın gerçek hareketinden ziyade BBox taban noktasındaki küçük değişimler ve ray-casting işlemindeki yüzey kesişimi belirsizlikleriyle ilişkilidir. Ham hız değerlerinin bazı anlarda yaya hareketiyle açıklanamayacak düzeylere ulaşması, ham konum serisinin doğrudan zamansal analiz ve yörünge değerlendirmesi için yeterince kararlı olmadığını göstermektedir.

Bu nedenle Kalman filtreleme aşamasına geçilmeden ham konum serisi fiziksel hareket sınırları bakımından kontrol edilmiştir. Hedef nesne yaya olduğundan, 35 km/h üzerindeki ardışık konum değişimleri kabul edilmemiştir. Bu eşik, normal yürüme, hızlı yürüme ve kısa süreli koşu davranışlarını dışlamayacak; buna karşılık BBox veya ray-casting kaynaklı ani konum sıçramalarını belirleyebilecek şekilde kullanılmıştır.



Şekil 4.16 Aykırı değer süzmesi ve Kalman filtreleme öncesi ham hız değişimi.

Bu nedenle Kalman filtreleme aşamasına geçilmeden önce ham konum serisi fiziksel hareket sınırları bakımından kontrol edilmiştir. Hedef nesne yaya olduğundan, 35 km/h üzerindeki ardışık konum değişimleri gerçekçi kabul edilmemiştir. Bu eşik, normal yürüme, hızlı yürüme ve kısa süreli koşu davranışlarını dışlamayacak; buna karşılık BBox veya ray-casting kaynaklı ani konum sıçramalarını belirleyebilecek şekilde kullanılmıştır.

Aykırı değer kontrolünde her nokta, yalnızca kendisinden önceki kareye göre değil, son

kabul edilen geçerli konuma göre değerlendirilmiştir. Böylece art arda oluşabilecek hatalı konum kümelerindeki sonraki noktaların da aykırı ölçüm olarak belirlenmesi amaçlanmıştır. Hız eşiğini aşan ölçümler Kalman filtresine doğrudan verilmemiş, ancak zaman serisinin sürekliliği korunarak bu adımlarda yalnızca hareket modeline dayalı tahmin üretilmiştir.

Bu değerlendirme sonucunda, ham konum kestirimlerinin genel hareket doğrultusunu temsil etmekle birlikte, kısa süreli gürültü ve hız sıçramaları nedeniyle doğrudan nihai yörünge olarak kullanılmasının uygun olmadığı görülmüştür.

Özellikle ardışık kareler arasında oluşan ani konum değişimleri, yayanın gerçek hareketinden çok BBox referans noktasındaki küçük dalgalanmalar ve ray-casting yüzey kesişimindeki belirsizliklerle ilişkilidir. Bu nedenle sonraki aşamada, konum serisinin daha süreklilik gösteren ve zamansal olarak daha kararlı bir yapıya dönüştürülmesi amacıyla sabit hız modeline dayalı doğrusal Kalman filtresi uygulanmıştır.

4.4.2 Kalman Filtreleme Uygulaması ve Parametre Seçimi

Ham konum kestirimlerinde gözlenen kısa süreli dalgalanmalar ve fiziksel olarak gerçekçi olmayan hız sıçramaları nedeniyle, video tabanlı konum serisinin zamansal olarak yumuşatılması amacıyla doğrusal Kalman filtresi uygulanmıştır. Bu çalışmada kullanılan filtre, sabit hız hareket modeline dayanmaktadır. Bu nedenle yöntem; doğrusal durum geçişi ve doğrusal ölçüm modeli kullanan standart Kalman filtresi olarak değerlendirilmiştir.

Kalman filtresinin temel amacı, her karede ray-casting ile elde edilen ham konum ölçümünü doğrudan nihai konum olarak kabul etmek yerine, önceki karelerdeki hareket sürekliliğini de dikkate alarak daha kararlı bir konum kestirimi üretmektir. Bu yaklaşımda filtre, bir önceki durumdan hareketle mevcut konumu tahmin etmekte, ardından ilgili karede geçerli bir ham ölçüm varsa bu ölçümle tahminini güncellemektedir. Böylece tekil karelerde oluşan kısa süreli konum dalgalanmalarının yörüngeye doğrudan ve keskin biçimde yansıması azaltılmaktadır.

Bu çalışmada Kalman filtresi UTM koordinat düzleminde uygulanmıştır. WGS 84 koordinatları doğrudan filtreye verilmemiş, öncelikle UTM Easting ve Northing bileşenlerine dönüştürülmüştür. Bunun nedeni, Kalman filtreleme ve hız hesabı işlemlerinin metre biriminde daha tutarlı yapılabilmesidir. Filtre durum vektörü Eşitlik (4.3) ile tanımlanmıştır:

$$x_i = \begin{bmatrix} E_i \\ N_i \\ v_{E,i} \\ v_{N,i} \end{bmatrix} \tag{4.3}$$

Burada E_i ve N_i, kareye ait UTM Easting ve Northing konum bileşenlerini; $v_{E,i}$ ve $v_{N,i}$ ise sırasıyla Easting ve Northing doğrultularındaki hız bileşenlerini ifade etmektedir. Ölçüm vektörü ise yalnızca konum bileşenlerinden oluşturulmuştur:

$$z_i = [E_i, N_i]^T \tag{4.4}$$

Bu nedenle filtre, hız bileşenlerini doğrudan ölçmemekte; ardışık konum ölçümleri ve sabit hız modeli üzerinden kestirmektedir. Sabit hız modelinde durum geçişi Eşitlik (4.5) ile ifade edilmiştir:

$$F_i = \begin{bmatrix} 1 & 0 & \Delta t_i & 0 \\ 0 & 1 & 0 & \Delta t_i \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.5}$$

Burada (Δt_i), ardışık iki kare arasındaki zaman farkını göstermektedir. Buna göre tahmin aşaması Eşitlik (4.6) ve Eşitlik (4.7) ile uygulanmıştır:

$$\hat{x}_i^- = F_i \hat{x}_{i-1} \tag{4.6}$$

$$P_i^- = F_i P_{i-1} F_i^T + Q_i \tag{4.7}$$

Burada $\hat{x}_i^-$, ölçüm güncellemesi yapılmadan önceki tahmini durum vektörünü; P_i^-,

tahmin kovaryans matrisini; Q_i ise süreç gürültüsü matrisini ifade etmektedir. Süreç gürültüsü, sabit hız modelinin gerçek hareketi tam olarak temsil edemediği durumlarda modele esneklik kazandırmak için kullanılmıştır. Çalışmada süreç gürültüsü katsayısı 0,5 olarak belirlenmiştir.

Ölçüm modeli yalnızca konum bileşenlerini kullandığı için ölçüm matrisi Eşitlik (4.8) ile tanımlanmıştır:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{4.8}$$

Bu matris, durum vektöründeki E ve N bileşenlerinin ölçüm vektörü ile karşılaştırılmasını sağlamaktadır. Ölçüm güncellemesi aşamasında önce ölçüm artığı hesaplanmıştır:

$$y_i = z_i - H\,\hat{x}_i^- \tag{4.9}$$

Daha sonra artık kovaryansı ve Kalman kazancı Eşitlik (4.10) ve Eşitlik (4.11) ile hesaplanmıştır:

$$S_i = H\,P_i^- H^T + R \tag{4.10}$$

$$K_i = P_i^- H^T (H\,P_i^- H^T + R)^{-1} \tag{4.11}$$

Burada R, ölçüm gürültüsü matrisini; K_i ise Kalman kazancını ifade etmektedir. Ölçüm gürültüsü katsayısı 25 olarak kullanılmıştır. Bu değerin görece yüksek seçilmesi, ham ray-casting ölçümlerinin kısa süreli dalgalanmalar içerebileceği kabulüne dayanmaktadır. Böylece filtre, her ham ölçüme tamamen bağlı kalmak yerine, ölçümü önceki hareket eğilimiyle birlikte değerlendirmiştir.

Son olarak durum vektörü ve kovaryans matrisi Eşitlik (4.12) ve Eşitlik (4.13) ile güncellenmiştir:

$$\hat{x}_i = \hat{x}_i^- + K_i\, y_i \tag{4.12}$$

$$P_i = (I - K_i H) P_i^- \tag{4.13}$$

Bu işlem her video karesi için tekrarlanmış ve sonuçta Kalman filtreli Easting ve Northing konum serisi elde edilmiştir. Elde edilen filtreli UTM koordinatları, görselleştirme ve raporlama amacıyla tekrar WGS 84 koordinat sistemine dönüştürülmüştür.

Kalman filtreleme öncesinde uygulanan aykırı değer kontrolü, filtreye verilen ölçümlerin fiziksel olarak anlamlı olmasını sağlamak amacıyla kullanılmıştır. Bu kapsamda 35 km/h üzerindeki ardışık konum değişimleri gerçekçi yaya hareketi olarak değerlendirilmemiştir. Hız sınırını aşan ölçümler filtreye doğrudan konum ölçümü olarak verilmemiş; ancak zaman dizisi korunarak filtre tahmin aşamasının devam etmesi sağlanmıştır. Böylece filtre, ölçümün geçersiz olduğu karelerde yalnızca hareket modeline dayalı tahmin üretmiş, geçerli ölçüm geldiğinde ise tahminini yeniden güncellemiştir. Kalman filtreleme sürecinde kullanılan temel model bileşenleri ve parametreler Çizelge 4.6'da özetlenmiştir.

Çizelge 4.6 Kalman filtreleme modelinde kullanılan temel bileşenler ve parametreler.

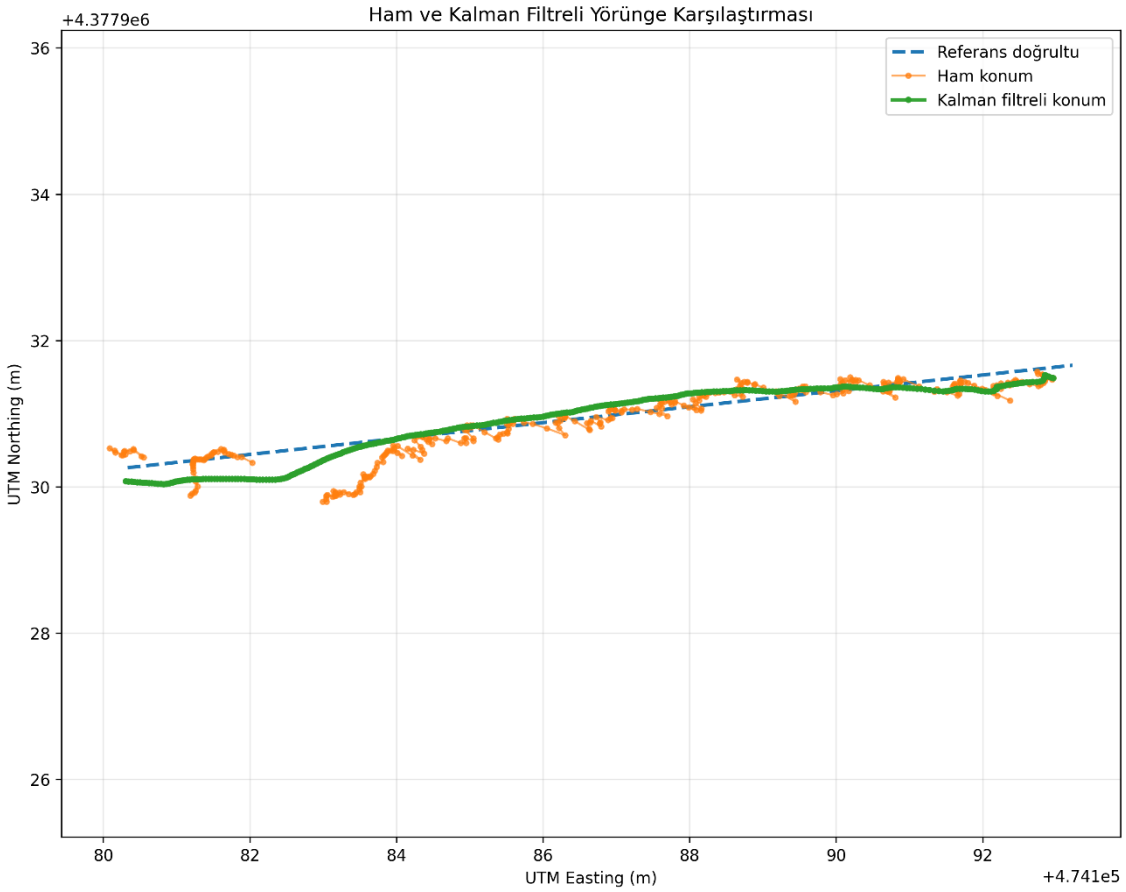
Parametre	Kullanılan değer
Filtre tipi	Doğrusal kalman filtresi
Hareket modeli	Sabit hız modeli
Durum vektörü	$\left[E_i, N_i, v_{E,i}, v_{N,i}\right]^T$
Ölçüm vektörü	$[E_i, N_i]^T$
Ölçüm matrisi	H
Süreç gürültüsü katsayısı	0,5
Ölçüm gürültüsü katsayısı	25
Aykırı hız eşiği	35 km/h
Koordinat düzlemi	UTM
Çıktı	Kalman filtreli (E), (N), (v_E), (v_N)

Bu yapı sayesinde ham konum ölçümleri tamamen yok sayılmamış, ancak ölçüm gürültüsü yüksek kabul edilerek hareket sürekliliğiyle birlikte değerlendirilmiştir. Bu tercih, ani konum sıçramalarının filtre sonucuna doğrudan yansımasını azaltırken, gerçek hareket değişimlerinin de tamamen dışlanmamasını sağlamaktadır. Bu nedenle çalışmada sert bir konum reddetme yaklaşımı yerine, fiziksel hız sınırı ve yüksek ölçüm gürültüsü birlikte kullanılmıştır. Aykırı hız kontrolü ve sabit hız modeline dayalı Kalman filtreleme sürecine ait örnek kod kesiti Ek 15'te sunulmuştur.

4.4.3 Ham ve Kalman Filtreli Sonuçların Karşılaştırılması

Bu bölümde, 35 km/h aykırı hız eşiği uygulanarak fiziksel olarak gerçekçi olmayan ani konum sıçramaları temizlenen ham konum serisi ile Kalman filtreli konum çıktıları karşılaştırılmıştır.

Bu nedenle burada kullanılan "ham konum" ifadesi, aykırı değer süzmesi sonrasında geçerli kabul edilen ray-casting tabanlı ölçümleri ifade etmektedir. Karşılaştırma; yörünge sürekliliği, konum zaman serisi, hız değişimi ve referans doğrultuya dik uzaklık ölçütleri üzerinden yapılmıştır. Şekil 4.17'de ham konum ölçümleri ile Kalman filtreli konumların UTM düzlemindeki yörünge karşılaştırması verilmiştir. Gösterilen referans doğrultu, video sahnesinde başlangıç ve bitiş konumlarını birleştiren doğru olarak tanımlanmıştır, yaya da bu doğrultu üzerinde hareket etmiştir. Ham ve Kalman filtreli konumların bu doğruya göre davranışı karşılaştırılmıştır.

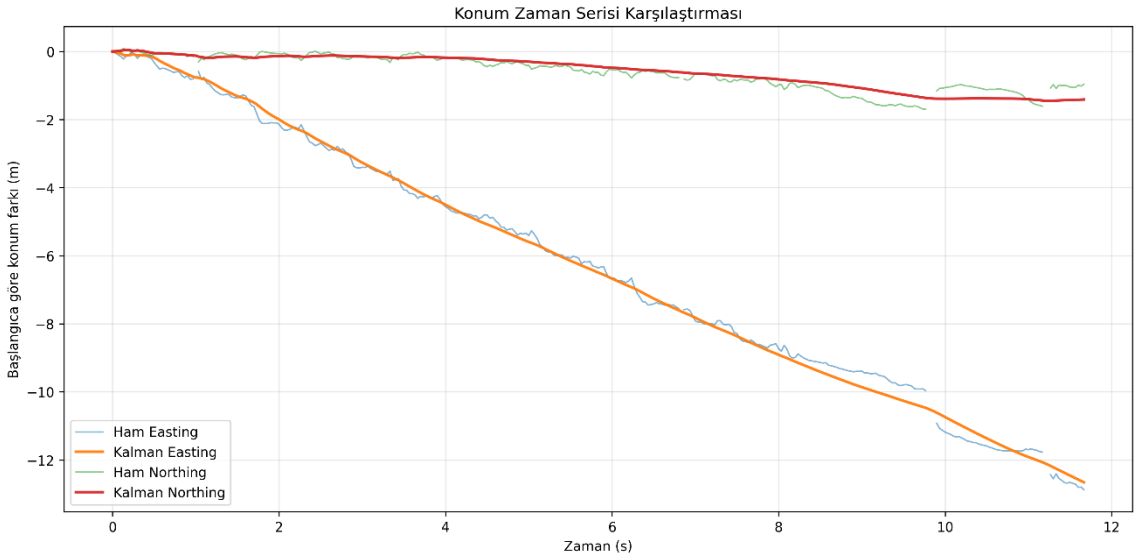


Şekil 4.17 Ham ve Kalman filtreli yörünge karşılaştırması.

Şekil 4.17 incelendiğinde, ham konum ölçümlerinin genel olarak referans hareket doğrultusunu takip ettiği, ancak noktasal düzeyde küçük dalgalanmalar içerdiği görülmektedir. Bu dalgalanmalar özellikle BBox alt orta noktasının kareler arasında küçük değişimler göstermesi ve ray-casting işleminin 3B yüzey üzerindeki yerel kesişim farklılıklarından etkilenmesiyle açıklanabilir. Kalman filtreli yörünge ise ham konum noktalarındaki kısa süreli salınımları bastırarak daha sürekli ve düzgün bir hat oluşturmuştur.

Yörünge grafiğinde Kalman filtreli konumun referans doğrultuya daha düzenli bir biçimde yaklaştığı, ham ölçümlerde görülen küçük ölçekli saçılmaların ise belirgin biçimde azaldığı görülmektedir. Bununla birlikte Kalman filtresi, ham ölçümleri tamamen yok saymamış; genel hareket eğilimini koruyarak konum serisini daha yumuşak hâle getirmiştir. Bu durum, kullanılan ölçüm gürültüsü değerinin ham ray-casting ölçümlerine sınırlı güven verecek şekilde seçilmesiyle ilişkilidir. Şekil 4.18'de

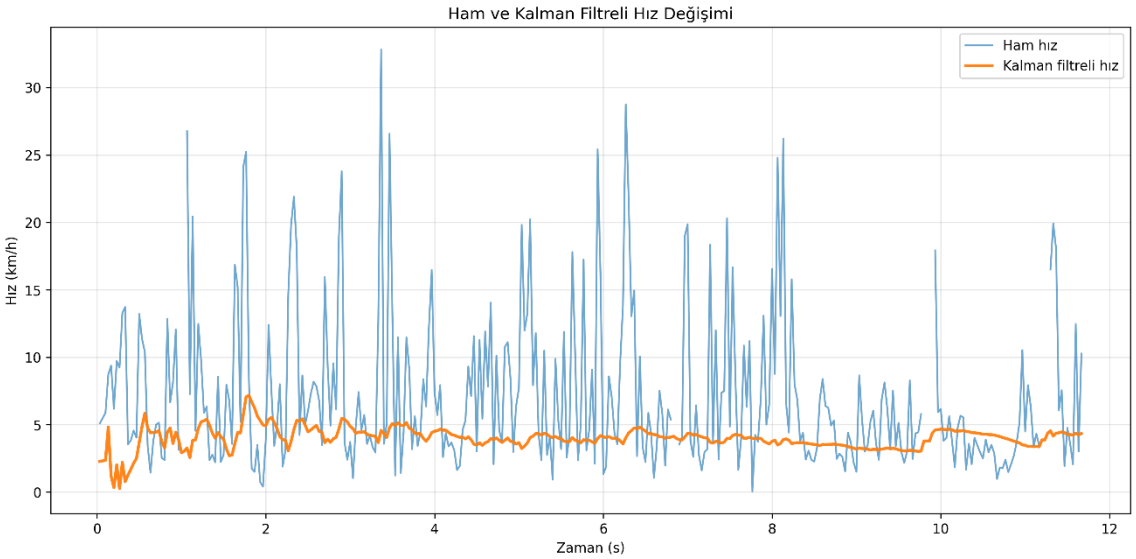
başlangıç konumuna göre normalize edilmiş Easting ve Northing bileşenlerinin zamana bağlı değişimi verilmiştir.



Şekil 4.18 Konum zaman serisi karşılaştırması

Konum zaman serisi incelendiğinde, ham Easting bileşeninin genel hareket eğilimini takip etmekle birlikte küçük dalgalanmalar içerdiği görülmektedir. Kalman filtreli Easting bileşeni ise bu dalgalanmaları azaltarak daha düzgün bir değişim göstermiştir. Northing bileşeninde, benzer şekilde ham ölçümlerde de oynamalar bulunmasına karşın, Kalman filtreli sonuç daha kararlı ve süreklilik gösteren bir yapı sergilemektedir.

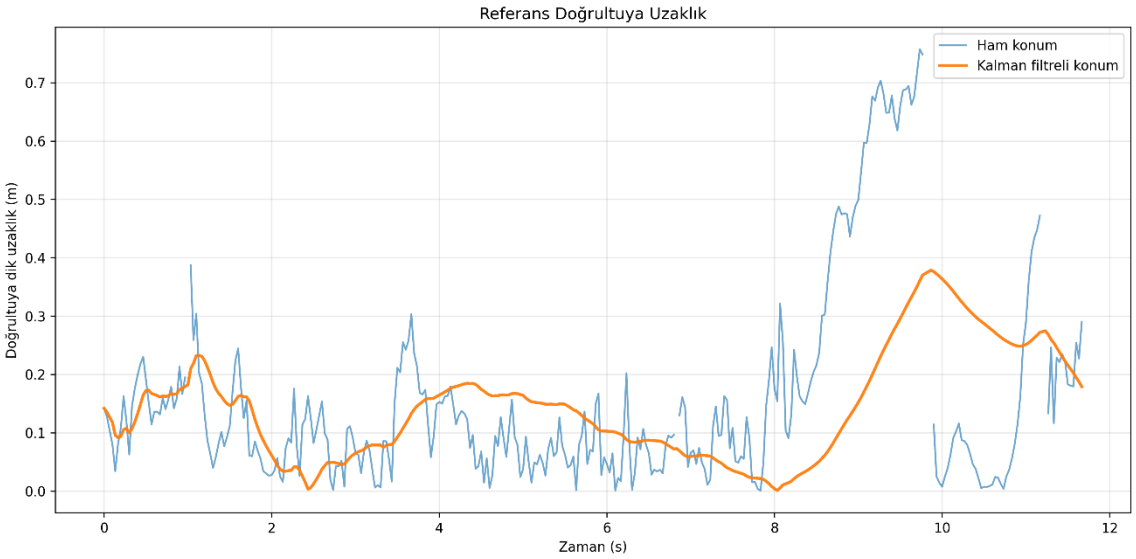
Bu grafikte dikkat edilmesi gereken nokta, Kalman filtresinin yalnızca konum değerlerini yumuşatmadığı, aynı zamanda ardışık karelerdeki hareket sürekliliğini de koruduğudur. Ham konum serisinde kare bazlı küçük kırılmalar bulunurken, Kalman filtreli seride bu kırılmalar daha düşük genlikli hâle gelmiştir. Bu sonuç, sabit hız modelinin özellikle kısa zaman aralıklarında yaya hareketini temsil etmek için uygun bir yaklaşım sunduğunu göstermektedir. Şekil 4.19'da ham hız değerleri ile Kalman filtreli hız değerlerinin zamana bağlı değişimi verilmiştir.



Şekil 4.19 Ham ve Kalman filtreli hız değişimi

Hız grafiği, Kalman filtrelemenin etkisini en belirgin gösteren sonuçlardan biridir. Aykırı hız süzmesi sonrasında dahi ham hız eğrisinde kısa süreli pikler ve ani değişimler gözlenmektedir. Bu pikler, yayanın gerçek hareketindeki ani hızlanmalardan çok, ardışık karelerdeki konum ölçüm gürültüsünün hız hesabına yansımasıyla ilişkilidir. Çünkü hız değeri, ardışık iki konum arasındaki farktan üretildiği için küçük konumsal oynamalar bile hız eğrisinde daha belirgin dalgalanmalar oluşturabilmektedir.

Kalman filtreli hız eğrisi ise ham hız eğrisine göre daha düşük genlikli ve daha sürekli bir yapı göstermektedir. Ham hızda görülen kısa süreli ani yükselmeler filtreli hızda bastırılmış, hareket daha kararlı bir hız bandında temsil edilmiştir. Bu sonuç, Kalman filtresinin yalnızca konum yumuşatma için değil, aynı zamanda konumdan türetilen hız bilgisinin kararlılığını artırmak için de etkili olduğunu göstermektedir.
Bununla birlikte Kalman filtreli hızın tamamen sabit hâle gelmediği görülmektedir. Bu durum beklenen bir sonuçtur; çünkü filtre, gerçek hareket değişimlerini tamamen yok sayacak şekilde tasarlanmamıştır. Çalışmada sert bir konum reddetme eşiği kullanılmamış, bunun yerine fiziksel hız sınırı ve yüksek ölçüm gürültüsü yaklaşımı tercih edilmiştir. Böylece ani ve gerçek dışı sıçramaların etkisi azaltılırken, yayanın gerçek hareketindeki kademeli değişimlerin filtre sonucuna yansımasına izin verilmiştir. Şekil 4.20'de ham ve Kalman filtreli konumların referans doğrultuya olan dik uzaklık değişimi verilmiştir.



Şekil 4.20 Referans doğrultuya dik uzaklık değişimi.

Referans doğrultuya dik uzaklık grafiği, yörüngenin tanımlanan hareket ekseninden ne kadar saptığını göstermektedir. Ham konum serisinde bazı zaman aralıklarında dik uzaklığın ani biçimde arttığı, buna karşılık Kalman filtreli konumda bu değişimlerin daha yumuşak gerçekleştiği görülmektedir. Özellikle ham konumda kısa süreli yükselmeler daha keskin iken, Kalman filtreli eğri bu yükselmeleri daha düşük frekanslı ve daha sürekli bir değişim olarak temsil etmektedir.

Ancak Kalman filtreli konumun her anda ham konumdan daha düşük dik uzaklık üretmesi beklenmemelidir. Çünkü Kalman filtresi noktasal olarak en yakın konumu seçmek yerine, yörüngenin zamansal sürekliliğini ve hareket modelini dikkate almaktadır.
Bu nedenle bazı anlarda ham ölçüm referans doğrultuya daha yakın görünebilir; ancak genel eğilim açısından Kalman filtreli çıktı daha kararlı ve daha az salınımlı bir yörünge sunmaktadır. Bu durum, Kalman filtrelemenin temel amacının tekil karelerdeki en düşük hata değerini üretmekten çok, zaman boyunca tutarlı bir hareket kestirimi sağlamak olduğunu göstermektedir.

Elde edilen sonuçlar birlikte değerlendirildiğinde, Kalman filtrelemenin ham ray-casting tabanlı konum kestirimlerinde görülen kısa süreli gürültüyü azalttığı, hız eğrisindeki ani

pikleri bastırdığı ve yörüngeyi daha sürekli hâle getirdiği görülmektedir. Özellikle video tabanlı uygulamalarda konum çıktılarının yalnızca tekil kare doğruluğu açısından değil, zamansal süreklilik açısından da değerlendirilmesi gerekmektedir. Bu bağlamda sabit hız modeline dayalı doğrusal Kalman filtresi, görüntü ve 3B model tabanlı konum kestirimlerinin daha kararlı bir yörüngeye dönüştürülmesinde etkili bir ara işlem olarak değerlendirilebilir.

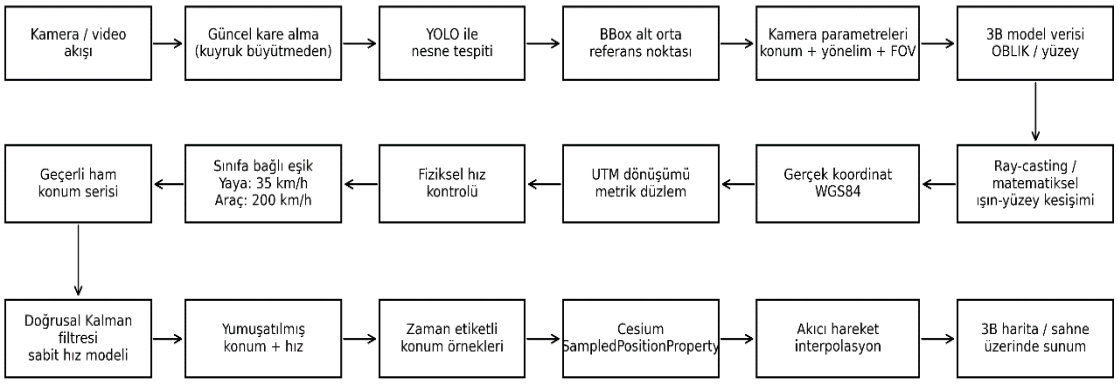
Bununla birlikte Kalman filtreleme, ray-casting veya BBox kaynaklı tüm hataları tamamen ortadan kaldıran bağımsız bir doğruluk artırma yöntemi olarak görülmemelidir. Filtre, geçerli ölçümlerin zamansal tutarlılığını artırmakta ve kısa süreli gürültüyü azaltmaktadır; ancak kamera parametreleri, 3B model yüzey kalitesi, BBox referans noktasının doğruluğu ve hedefin görünürlük durumu konum kestiriminin temel doğruluk sınırlarını belirlemeye devam etmektedir. Bu nedenle Kalman filtreleme, nihai sistemde tek başına mutlak doğruluk sağlayan bir yöntem değil, ham konum serisinin kararlılığını artıran tamamlayıcı bir yumuşatma adımı olarak değerlendirilmiştir.

4.5 Bütünleşik Konum Kestirimi ve 3B Görselleştirme Sistem Mimarisi

Bu çalışmada elde edilen bulgular doğrultusunda, görüntü tabanlı konum kestiriminin yalnızca tekil kareler üzerinden koordinat üretme problemi olarak ele alınmasının yeterli olmadığı değerlendirilmiştir. Video tabanlı uygulamalarda üretilen konumların zamansal olarak kararlı olması, fiziksel hareket sınırlarıyla uyum göstermesi ve görselleştirme ortamında gecikmeyi artırmadan sunulabilmesi önem taşımaktadır. Çünkü tekil karelerde kabul edilebilir görünen konum kestirimleri, ardışık kareler boyunca değerlendirildiğinde titreşimli veya süreksiz hareket davranışına yol açabilmektedir.

Bu durum, konum üretim sürecinin yalnızca geometrik doğrulukla değil, aynı zamanda zamansal tutarlılıkla birlikte ele alınmasını gerekli kılmaktadır. Bu nedenle önerilen sistem mimarisi; kamera görüntüsünden nesne tespiti, BBox referans noktasının gerçek dünya koordinatına dönüştürülmesi, fiziksel olarak gerçekçi olmayan ölçümlerin ayıklanması, Kalman filtreleme ile yörünge yumuşatma ve zaman tabanlı 3B

görselleştirme aşamalarından oluşmaktadır. Bu yapı, hem konum kestirimindeki ani sıçramaların azaltılmasını hem de dijital ikiz ortamında daha tutarlı bir hareket sunumu elde edilmesini hedeflemektedir. Önerilen sistemin genel işlem akışı Şekil 4.21'de verilmiştir.



Şekil 4.21 Gerçek zamanlı görüntü tabanlı konum kestirimi ve görselleştirme için önerilen sistem mimarisi.

Şekil 4.21'de görüldüğü üzere sistemin ilk aşamasında kamera veya video akışından güncel görüntü karesi alınmaktadır. Alınan görüntü üzerinde YOLO tabanlı nesne tespiti uygulanmakta ve hedef nesneye ait BBox bilgisi elde edilmektedir. Bu çalışmada yaya hedefi için BBox alt orta noktası, nesnenin zeminle temas noktasını temsil eden referans nokta olarak kullanılmıştır. Bu referans noktasının seçilmesi, görüntü düzlemindeki nesne tespit sonucunun 3B model üzerinde konum üretimi için kullanılabilmesini sağlamaktadır.

İkinci aşamada, BBox alt orta noktasının görüntü düzlemindeki piksel koordinatı kamera parametreleriyle birlikte gerçek dünya koordinatına dönüştürülmektedir. Bu işlem için kamera konumu, heading, pitch, roll, görüş açısı ve görüntü oranı gibi parametreler kullanılmaktadır. Görüntü düzlemindeki seçilen pikselden 3B modele doğru bir ışın oluşturulmakta ve bu ışının 3B model yüzeyi ile kesiştiği nokta hesaplanmaktadır. Bu işlem CesiumJS ortamında ray-casting araçlarıyla uygulanabileceği gibi, 3B model geometrisi ve kamera projeksiyon parametreleri kullanılarak matematiksel olarak da gerçekleştirilebilir. Bu nedenle temel işlem, kullanılan yazılımdan bağımsız olarak, görüntü noktasının 3B yüzey üzerinde karşılık

geldiği kesişim noktasının belirlenmesidir.

Ray-casting veya eşdeğer matematiksel kesişim hesabı sonucunda elde edilen konum WGS 84 koordinat sisteminde ifade edilmektedir. Ancak hız, mesafe ve filtreleme işlemlerinin metrik düzlemde yapılabilmesi için bu koordinatlar UTM sistemine dönüştürülmektedir. Böylece ardışık konumlar arasındaki mesafe, hız ve yörünge sapması gibi büyüklükler metre cinsinden değerlendirilebilmektedir.

Üçüncü aşamada ham konum serisi fiziksel hareket sınırları açısından kontrol edilmektedir. Bu kontrolün amacı, BBox taban noktası veya 3B yüzey kesişimi kaynaklı ani konum sıçramalarının doğrudan filtreleme ve görselleştirme aşamasına aktarılmasını engellemektir. Bu çalışmada yaya hareketi için 35 km/h üzerindeki ardışık konum değişimleri gerçekçi kabul edilmemiştir. Aynı yaklaşım farklı hedef sınıfları için sınıfa özgü hız sınırlarıyla uygulanabilir. Örneğin araçlar için daha yüksek bir eşik değer tanımlanması gerekmektedir. Bu tür eşiklerin hedef sınıfı, sahne tipi ve beklenen hareket davranışına göre belirlenmesi uygun olacaktır.

Dördüncü aşamada, fiziksel hız kontrolünden geçen konum ölçümleri sabit hız modeline dayalı doğrusal Kalman filtresine aktarılmaktadır. Kalman filtresi, her karede üretilen ham konum ölçümünü doğrudan nihai değer olarak kabul etmek yerine, önceki konum ve hız eğilimini de dikkate alarak daha kararlı bir yörünge üretmektedir. Bu çalışmada kullanılan yaklaşımda sert bir konum reddetme eşiği tercih edilmemiştir. Bunun yerine, fiziksel hız sınırı ile gerçek dışı sıçramalar ayıklanmış; kalan ölçümler ise ölçüm gürültüsü yüksek kabul edilerek Kalman filtresi içinde yumuşatılmıştır. Bu yapı, ani hataların etkisini azaltırken gerçek hareket değişimlerinin tamamen dışlanmamasına olanak sağlamaktadır.

Beşinci aşamada, Kalman filtresinden elde edilen konum çıktılarının 3B sahnede zamana bağlı olarak sunulması gerekmektedir. Bu aşamada yalnızca son konumu sahneye anlık olarak aktarmak yerine, konumların zaman bilgisiyle birlikte değerlendirilmesi daha uygun bir sunum sağlamaktadır.

CesiumJS ortamında bu amaçla SampledPositionProperty yapısı kullanılabilir. Bu yapı, farklı zamanlara ait konum örneklerinin tanımlanmasına ve nesnenin bu örnekler arasında zamana bağlı olarak daha akıcı biçimde gösterilmesine imkân vermektedir. CesiumJS ortamında araç, yaya, tır ve otobüs gibi nesne sınıflarının GLB/glTF tabanlı 3B modellerle temsil edilmesine yönelik örnek model yerleştirme veri yapısı Ek 9'da verilmiştir. Önerilen mimarinin görselleştirme katmanında, elde edilen konum kestirimlerinin yalnızca nokta veya etiket olarak değil, nesne sınıfına uygun 3B modellerle de temsil edilmesi mümkündür. Önerilen sistem mimarisinin 3B görselleştirme çıktısına ait örnek ekran görüntüsü Şekil 4.22'de verilmiştir.



Şekil 4.22 Araç, yaya, tır ve otobüs modellerinin CesiumJS ortamında 3B dijital ikiz altlığı üzerinde gösterimi.

Çalışma kapsamında geliştirilen uygulamada, kamera görüntüsü üzerinde gerçekleştirilen nesne tespitleri ile bu nesnelere ait 3B görselleştirme sonuçları aynı arayüz içerisinde birlikte sunulmuştur. Kamera görüntüsündeki tespit sonuçları ile CesiumJS tabanlı 3B dijital ikiz ortamındaki konumsal karşılıkların birlikte gösterildiği çalışan sistem örneği Şekil 4.23'te verilmiştir.

Şekil 4.23 Geliştirilen görüntü tabanlı nesne konumlandırma ve 3B dijital ikiz görselleştirme

sisteminin örnek çıktısı.

Bu nedenle önerilen yapıda her karenin sırayla işlenmesi yerine, işlem tamamlandığında kamera akışından en güncel karenin alınması daha uygun görülmektedir. Başka bir ifadeyle sistem, bir kare üzerindeki nesne tespiti ve konum kestirim işlemlerini tamamladıktan sonra sıradaki eski kareyi değil, mevcut güncel kareyi işlemelidir. Bu yaklaşım bazı karelerin atlanmasına neden olabilir; ancak konumun güncel sahneyle uyumlu kalmasını sağlar. Böylece işlem süresi arttığında sistem gecikmesi katlanarak büyümez ve görselleştirilen hareketin gerçek zamanla ilişkisi korunur.

Bu mimaride donanım kullanımı da sistem başarısı açısından önemlidir. Nesne tespiti gibi derin öğrenme tabanlı işlemler GPU üzerinde çalıştırıldığında işlem süresi azalabilmektedir. Ray-casting veya 3B model kesişim hesabı, koordinat dönüşümü ve Kalman filtreleme adımları ise sistem mimarisine göre CPU veya GPU üzerinde yürütülebilir. Burada amaç, her kareyi mutlaka işlemek değil, kabul edilebilir gecikme ile güncel ve kararlı konum üretmektir. Bu nedenle gerçek zamanlıya yakın sistem tasarımında model çalışma süresi, kamera kare hızı, donanım kapasitesi ve görselleştirme güncelleme aralığı birlikte değerlendirilmelidir.

Sonuç olarak önerilen mimari; kamera görüntüsünden elde edilen BBox tabanlı tespit bilgisini, 3B model üzerinden gerçek dünya koordinatına dönüştürmekte; fiziksel hız sınırı ve Kalman filtreleme ile bu koordinatları zamansal olarak yumuşatmakta; ardından zaman etiketli konum örnekleriyle 3B ortamda daha akıcı biçimde sunmaktadır. Bu yapı, ham görüntü tabanlı konum ölçümlerindeki ani sıçramaların etkisini azaltmak ve gerçek zamanlıya yakın sunumda gecikmeyi kontrol altında tutmak için kullanılabilecek bir çerçeve sunmaktadır.

5. TARTIŞMA ve SONUÇ

Bu tez çalışmasında, sabit kamera görüntülerinden derin öğrenme tabanlı nesne tespiti yapılarak hareketli nesnelerin gerçek dünya koordinatlarıyla ilişkilendirilmesi ve bu bilgilerin 3B dijital ikiz ortamında değerlendirilebilmesi amaçlanmıştır. Bu amaç doğrultusunda çalışma alanına ait fotogrametrik mesh model ve LiDAR nokta bulutu kullanılarak 3B dijital ikiz altlığı oluşturulmuş; YOLO tabanlı model varyantları nesne tespit başarımı ve işlem süresi açısından karşılaştırılmış; piksel koordinatlarından gerçek dünya koordinatlarına dönüşüm için homografi, IDW tabanlı lokal enterpolasyon ve ray-casting yaklaşımları değerlendirilmiştir. Ayrıca video tabanlı konum serilerinde ortaya çıkan kısa süreli dalgalanmaların azaltılması amacıyla Kalman filtreleme yaklaşımı uygulanmıştır.

5.1 Bulguların Değerlendirilmesi

Elde edilen bulgular, video tabanlı nesne konumlandırma probleminin yalnızca nesne tespiti başarımıyla açıklanamayacağını göstermektedir. Nesne tespit modeli doğru sınıfı üretebilse bile, BBox referans noktasının konumu, kamera bakış geometrisi, hedefin kameraya uzaklığı, örtülme durumu ve kullanılan 3B yüzey altlığının kalitesi konumsal doğruluğu doğrudan etkilemektedir. Bu nedenle görüntü tabanlı dijital ikiz

uygulamalarında nesne tespiti, koordinat dönüşümü, 3B yüzey temsili ve zamansal filtreleme aşamalarının birlikte değerlendirilmesi gerekmektedir.

Model karşılaştırma sonuçları, en yüksek doğruluk değerinin her zaman en uygun sistem seçimi anlamına gelmediğini göstermiştir. Etiketli BBox ve IoU tabanlı doğrulama sonuçlarında YOLO26-l modeli doğruluk açısından öne çıkarken, YOLO26-s modeli daha düşük çıkarım süresi ve yeterli tespit başarımı nedeniyle hız-doğruluk dengesi bakımından daha uygun bulunmuştur. Dijital ikiz uygulamalarında tespit çıktılarının koordinat dönüşümü ve görselleştirme süreçlerine gerçek zamanlıya yakın biçimde aktarılması gerektiğinden, model seçiminde yalnızca F1-skor değil, işlem süresi, yanlış negatif davranışı ve sistem bütünlüğü de dikkate alınmalıdır.

Konum doğruluğu analizleri, BBox alt orta noktasının yaya nesnesinin zemindeki konumunu temsil etmek için BBox merkezine göre daha uygun olduğunu ortaya koymuştur. BBox merkezi, yayanın gövde bölgesine karşılık geldiğinden özellikle ray-casting ve homografi tabanlı dönüşümlerde zemindeki gerçek konumdan sapmalara neden olabilmektedir. Buna karşılık BBox alt orta noktası, yayanın ayak veya temas bölgesine daha yakın bir temsil sunduğundan konumsal doğruluk açısından daha tutarlı sonuçlar vermiştir. Bununla birlikte BBox alt orta noktasının da her durumda gerçek ayak noktasını kusursuz temsil etmediği, özellikle örtülme, uzak mesafe ve BBox sınırındaki dalgalanmalar nedeniyle hataların artabildiği görülmüştür.

Koordinat dönüşüm yöntemleri birlikte değerlendirildiğinde, homografi ve IDW tabanlı lokal enterpolasyon yöntemlerinin uygulanabilir olmakla birlikte sahne geometrisi ve referans noktalarının dağılımına duyarlı olduğu belirlenmiştir. Homografi yöntemi hızlı ve pratik bir çözüm sunmasına rağmen düzlemsel yüzey varsayımına bağlıdır. IDW tabanlı lokal enterpolasyon ise daha bölgesel bir kestirim yapabilmekte, ancak kontrol noktalarının dağılımına ve hedef noktanın bu noktalara olan uzaklığına doğrudan bağlı kalmaktadır.

Ray-casting yaklaşımı ise 3B dijital ikiz yüzeyiyle doğrudan kesişim kurduğu için özellikle eğimli ve kot farkı içeren alanlarda daha anlamlı bir konumlandırma yapısı

sunmuştur. Bu çalışmada ray-casting Oblik yaklaşımı, BBox alt orta noktası ile birlikte kullanıldığında en dengeli konum doğruluğu sonuçlarını üretmiştir.

Kalman filtreleme sonuçları, ham ray-casting tabanlı konum serilerinde görülen kısa süreli dalgalanmaların azaltılabildiğini göstermiştir. Filtreleme sonrasında yörünge daha sürekli hâle gelmiş, hız eğrisindeki ani pikler azalmış ve hareket daha kararlı bir yapı kazanmıştır. Bununla birlikte Kalman filtresi, BBox hatası, kamera parametresi hatası veya 3B model yüzeyindeki bozulmaları ortadan kaldıran bağımsız bir doğruluk artırma yöntemi değildir. Bu nedenle Kalman filtreleme, mutlak konum doğruluğunu garanti eden bir işlem olarak değil, geçerli konum ölçümlerinin zamansal tutarlılığını artıran tamamlayıcı bir yumuşatma adımı olarak değerlendirilmelidir.

5.2 Çalışmanın Katkıları

Bu çalışmanın temel katkısı, video görüntülerinden elde edilen derin öğrenme tabanlı nesne tespit sonuçlarının 3B dijital ikiz ortamında gerçek dünya koordinatlarıyla ilişkilendirilmesine yönelik bütünleşik bir iş akışı sunmasıdır. Çalışmada İHA tabanlı fotogrametri, LiDAR nokta bulutu, YOLO tabanlı nesne tespiti, piksel koordinatlarından gerçek dünya koordinatlarına dönüşüm, CesiumJS ortamında 3B görselleştirme ve Kalman filtreleme yaklaşımları aynı uygulama çerçevesinde ele alınmıştır.

Çalışmanın bir diğer katkısı, pikselden gerçek dünya koordinatına dönüşümde farklı yöntemlerin karşılaştırmalı olarak değerlendirilmesidir. Homografi, IDW tabanlı lokal enterpolasyon ve ray-casting yaklaşımlarının aynı veri seti üzerinde test edilmesi, bu yöntemlerin güçlü ve sınırlı yönlerinin ortaya konulmasını sağlamıştır. Ayrıca LiDAR ve oblik fotogrametri tabanlı 3B yüzeylerin ray-casting sonuçlarına etkisinin ayrı ayrı değerlendirilmesi, dijital ikiz altlığı seçiminde yüzey sürekliliğinin ve model geometrisinin önemini göstermiştir.

Bu çalışma ayrıca, BBox referans noktası seçiminin konumsal doğruluk üzerindeki etkisini açık biçimde ortaya koymuştur. Nesne tespiti çalışmalarında BBox genellikle

sınıflandırma ve tespit başarımı açısından değerlendirilirken, bu çalışmada BBox'ın hangi noktasının gerçek dünya koordinatına dönüştürüleceği ayrıca incelenmiştir. Bu yönüyle çalışma, görüntü tabanlı nesne tespit sonuçlarının coğrafi konumlandırma amacıyla kullanılmasında BBox referans noktası seçiminin kritik bir değişken olduğunu göstermektedir.

Önerilen iş akışı, belediyeler ve yerel yönetimler açısından da önemli bir karar destek potansiyeli taşımaktadır. Kamera görüntülerinden elde edilen nesne tespit sonuçlarının 3B coğrafi bağlam içerisinde konumlandırılması; trafik yönetimi, yaya yoğunluğu analizi, tehlikeli bölge ihlali, kaza veya düşme benzeri olağan dışı durumların tespiti, afet ve tahliye yönetimi, güvenlik izleme ve acil durum uyarı sistemleri gibi farklı akıllı şehir uygulamalarına uyarlanabilir. Böylece kamera tabanlı algılama sonuçları yalnızca görüntü üzerinde kalmamakta, gerçek dünya koordinatlarıyla ilişkilendirilerek mekânsal analiz ve görselleştirme süreçlerine dâhil edilebilmektedir. Bu yönüyle önerilen iş akışı, 2024–2030 Ulusal Akıllı Şehirler Stratejisi ve Eylem Planı'nda vurgulanan teknoloji odaklı, sürdürülebilir ve dirençli akıllı şehir yaklaşımıyla ilişkilendirilebilecek bir uygulama altyapısı sunmaktadır (İnt. Kyn. 10).

Bu yaklaşımın en önemli faydalarından biri, tespit edilen nesnenin yalnızca görüntüde var olup olmadığını değil, gerçek dünyada yaklaşık olarak nerede bulunduğunu da belirleyebilmesidir. Bu durum, özellikle belediyelerin saha ekiplerini yönlendirmesi, yoğunluk veya risk bölgelerini izlemesi, trafik ve yaya güvenliği kararlarını desteklemesi, kamusal alanlarda olay farkındalığını artırması ve acil müdahale süreçlerinde konumsal bilgi sağlaması açısından değerlidir. Bu nedenle çalışma, dijital ikiz tabanlı yerel yönetim uygulamaları için görüntü verisi ile 3B mekânsal veri arasında kurulabilecek bir entegrasyon örneği sunmaktadır.

5.3 Sınırlılıklar

Bu çalışmanın bazı sınırlılıkları bulunmaktadır. İlk olarak, görüntü tabanlı konum kestirimi BBox referans noktasına bağlı olduğundan, yayanın kısmen örtülmesi, ayak bölgesinin görüntüde net seçilememesi veya BBox sınırlarının kareler arasında

değişmesi konum hatasına neden olabilmektedir. Bu nedenle BBox alt orta noktası genel olarak uygun bir temsil sunsa da, her durumda gerçek zemin temas noktasını kusursuz biçimde temsil etmemektedir.

İkinci olarak, ray-casting sonuçları kamera konumu, kamera yönelimi, FOV değeri, görüntü oranı ve 3B model yüzeyinin yerel geometrik kalitesine duyarlıdır. Kamera yönelimindeki küçük hatalar veya 3B yüzeydeki yerel bozulmalar, ışın-yüzey kesişim noktasını değiştirebilmektedir. Özellikle yataya yakın kamera bakışlarında küçük piksel sapmaları zeminde daha büyük metrik hatalara dönüşebilmektedir.

Üçüncü olarak, homografi ve IDW tabanlı lokal enterpolasyon yöntemleri bu çalışmada temel olarak yatay konum bileşenlerinin kestirilmesi amacıyla değerlendirilmiştir. Bu yöntemler doğrudan 3B yüzey kesişimine dayanmadığından, yükseklik bileşeni ray-casting yaklaşımında olduğu gibi doğal biçimde elde edilmemektedir. Bu nedenle eğimli ve kot farkı içeren alanlarda bu yöntemlerin kullanımı dikkatli değerlendirilmelidir.

Dördüncü olarak, bu çalışmada değerlendirmeler sınırlı sayıda temsilî görüntü ve video kesiti üzerinden gerçekleştirilmiştir. Farklı kamera yükseklikleri, farklı pitch açıları, farklı FOV değerleri, farklı ışık koşulları, kalabalık sahneler, gece görüntüleri ve yoğun hareket senaryoları üzerinde daha kapsamlı testlerin yapılması gerekmektedir. Ayrıca bu çalışmada temel değerlendirme yaya sınıfı üzerinden yürütülmüş, farklı hareketli nesne sınıfları için ayrı hız modelleri ve eşik değerleri ayrıntılı biçimde test edilmemiştir.

Son olarak, önerilen gerçek zamanlıya yakın sistem mimarisi çalışma bulgularına dayalı olarak ortaya konulmuştur. Tüm bileşenlerin farklı donanım ortamlarında, uzun süreli canlı kamera akışları üzerinde ve operasyonel saha koşullarında ayrıca doğrulanması gerekmektedir. Bu durum, önerilen yaklaşımın uygulanabilirliğini azaltmamakta; ancak sistemin gerçek kullanım senaryolarına aktarılabilmesi için ek test ve optimizasyon ihtiyacını göstermektedir.

5.4 Sonuçlar

Bu tez çalışmasında, sabit kamera görüntülerinden elde edilen nesne tespit sonuçlarının gerçek dünya koordinatlarıyla ilişkilendirilmesi ve 3B dijital ikiz ortamında değerlendirilmesi için bütünleşik bir yöntem önerilmiştir.

Çalışma alanına ait fotogrametrik mesh model ve LiDAR nokta bulutu kullanılarak 3B dijital ikiz altlığı oluşturulmuş; bu altlık, konum dönüşümü ve görselleştirme süreçlerinde temel mekânsal referans olarak kullanılmıştır.

YOLO tabanlı model varyantları üzerinde yapılan karşılaştırmalarda, YOLO26-l modelinin doğruluk açısından en yüksek performansı ürettiği, YOLO26-s modelinin ise hız-doğruluk dengesi açısından dijital ikiz entegrasyonu için daha uygun olduğu belirlenmiştir. Confidence eşik değeri analizinde 0,30 eşiğinin, yanlış negatif sayısını daha düşük tutması nedeniyle çalışma kapsamında daha uygun olduğu görülmüştür.

Konumsal doğruluk analizlerinde BBox alt orta noktasının, BBox merkezine göre yaya konumunun zemindeki temsilinde daha başarılı olduğu belirlenmiştir. Koordinat dönüşüm yöntemleri arasında ray-casting Oblik yaklaşımı, bu veri setinde en dengeli sonuçları üretmiştir.

Homografi ve IDW tabanlı lokal enterpolasyon yöntemleri ise uygulama kolaylığı sağlamasına rağmen, sahne geometrisi ve referans noktalarının dağılımına daha duyarlı sonuçlar vermiştir.

Kalman filtreleme, ham ray-casting konum serisindeki ani dalgalanmaları azaltarak yörünge sürekliliğini artırmış ve video tabanlı dijital ikiz uygulamalarında zamansal kararlılığın önemini ortaya koymuştur. Bununla birlikte, filtreleme işleminin temel işlevinin mutlak konum doğruluğunu tek başına artırmak değil, geçerli konum ölçümlerini zamansal olarak daha kararlı hâle getirmek olduğu sonucuna ulaşılmıştır.

Bunun yanında, 3B dijital ikiz altlığının üretildiği anda çalışma alanında bulunan araç,

yaya veya geçici nesneler, fotogrametrik mesh model ya da nokta bulutu içerisinde statik yüzey parçaları olarak temsil edilebilmektedir. Bu durum, daha sonra aynı sahneye dinamik olarak eklenen araç, yaya veya benzeri 3B modellerle görsel çakışma oluşturabilir. Bu nedenle hareketli nesnelerin doğru ve temiz biçimde temsil edilebilmesi için, 3B model üretiminden sonra geçici nesne kalıntılarının mümkün olduğunca temizlenmesi veya dinamik nesne katmanından ayrıştırılması gerekmektedir.

Genel olarak bu çalışma, derin öğrenme tabanlı nesne tespiti ile 3B dijital ikiz görselleştirmesinin birlikte kullanılabileceğini; ancak güvenilir sonuçlar için nesne tespiti, BBox referans noktası, kamera geometrisi, 3B yüzey kalitesi, koordinat dönüşüm yöntemi ve zamansal filtreleme bileşenlerinin bütüncül olarak ele alınması gerektiğini göstermiştir.

5.5 Öneriler

Gelecek çalışmalarda kamera konumu ve yöneliminin daha hassas biçimde belirlenmesi için sistematik kamera kalibrasyonu ve saha ölçümleri kullanılabilir. Farklı kamera yükseklikleri, pitch açıları ve FOV değerlerinin konum doğruluğuna etkisi daha geniş veri setleri üzerinde test edilebilir. Ayrıca tek kamera yerine çoklu kamera kullanımı, örtülme problemini azaltarak BBox referans noktasının daha kararlı belirlenmesine katkı sağlayabilir.

Gelecek çalışmalarda, 3B dijital ikiz altlığına sonradan eklenen araç, yaya, tır ve otobüs gibi dinamik nesne modellerinin arazi yüzeyiyle daha gerçekçi biçimde ilişkilendirilmesi için yüzey eğimine duyarlı konumlandırma yaklaşımları geliştirilebilir. Özellikle araç ve tır gibi tekerlekli nesnelerde yalnızca yatay doğrultuyu temsil eden heading açısının kullanılması yeterli olmayabilir. Bu tür nesnelerin yol veya arazi yüzeyine gerçekçi biçimde oturabilmesi için, yerel yüzey normalinden veya yakın çevredeki yükseklik örneklerinden yararlanılarak pitch ve roll açılarının da hesaplanması gerekmektedir. Böylece 3B model sahnesine eklenen dinamik nesnelerin eğimli yol yüzeylerinde havada kalması, zemine gömülmesi veya araziyle uyumsuz görünmesi azaltılabilir.

Yaya konumunun zemindeki temas noktasını daha doğru belirlemek için insan poz tahmini veya ayak noktası çıkarımı tabanlı yöntemler değerlendirilebilir. Böylece yalnızca BBox alt orta noktasına bağlı kalmadan, kişinin gerçek ayak veya temas noktası daha güvenilir biçimde tahmin edilebilir. Araç, bisiklet, motosiklet veya farklı hareketli nesne sınıfları için ise sınıfa özgü hız sınırları, hareket modelleri ve Kalman filtre parametreleri ayrıca belirlenmelidir.

Gelecek çalışmalarda sistem farklı akıllı şehir uygulamalarına genişletilebilir. Belediyeler açısından bu yaklaşım; yaya geçidi güvenliği, kavşak ve yol kenarı izleme, kalabalık yoğunluğu analizi, yasaklı alan ihlali, otopark ve durak çevresi hareketliliği, afet sonrası toplanma alanlarının izlenmesi, tahliye güzergâhlarının değerlendirilmesi ve acil müdahale ekiplerinin yönlendirilmesi gibi uygulamalarda kullanılabilecek bir karar destek altyapısına dönüştürülebilir. Böylece kamera görüntülerinden elde edilen tespitler, 3B dijital ikiz ortamında konumsal farkındalık sağlayan bir bilgi katmanına dönüştürülebilir.

Güvenlik odaklı gelecek çalışmalarda, etik ve hukuki gereklilikler gözetilerek, yetkili kurumlar tarafından sağlanan uygun veri setleriyle tehlikeli nesne veya olağan dışı davranış sınıflarının tespiti araştırılabilir. Örneğin silah benzeri tehlikeli nesnelerin, kavga, düşme, ani kaçış veya kalabalıkta olağan dışı hareket örüntülerinin belirlenmesi, erken uyarı ve acil durum yönetimi açısından değerlendirilebilir. Ancak bu tür uygulamalarda yanlış pozitif ve yanlış negatif sonuçların ciddi etkiler oluşturabileceği dikkate alınmalı; sistem çıktıları doğrudan otomatik karar mekanizması olarak değil, insan denetimiyle çalışan karar destek unsuru olarak ele alınmalıdır.

Bunun yanında, gelecek çalışmalarda önerilen yaklaşımın kritik tesis güvenliği, stratejik altyapıların izlenmesi, kampüs güvenliği, sınırlandırılmış alan kontrolü ve yetkisiz giriş tespiti gibi güvenlik odaklı senaryolara uyarlanması değerlendirilebilir. Bu kapsamda sistem, yalnızca nesne konumlarının gösterildiği bir 3B görselleştirme ortamı olarak değil, aynı zamanda riskli bölge ihlali, olağan dışı hareket örüntüsü, yoğunluk değişimi ve acil durum uyarıları üretebilen insan denetimli bir karar destek altyapısı olarak geliştirilebilir.

Yaya ve insan hareketlerinin daha ayrıntılı temsil edilebilmesi için insan poz tahmini (Human Pose Estimation) ve anahtar nokta tespiti (keypoint detection) tabanlı yaklaşımlar da gelecek çalışmalara dâhil edilebilir. Bu sayede kişinin yalnızca BBox veya ayak temas noktası değil; baş, omuz, dirsek, el, kalça, diz ve ayak bileği gibi eklem noktaları da belirlenebilir. Elde edilen eklem noktaları 3B dijital ikiz ortamında iskelet modeli, avatar veya animasyonlu insan modeli ile ilişkilendirilerek yürüme, koşma, düşme, eğilme veya olağan dışı davranışların daha anlaşılır biçimde görselleştirilmesi sağlanabilir.

Görselleştirme tarafında ise CesiumJS tabanlı web gösteriminin yanında, daha gerçekçi ışıklandırma, fizik tabanlı hareket, animasyon, çarpışma kontrolü ve simülasyon kabiliyeti gerektiren senaryolarda Cesium for Unreal ve Unreal Engine tabanlı çözümler değerlendirilebilir. Böylece 3D Tiles veri yapısı, gerçek dünya koordinatları, dinamik nesne konumları ve insan/araç animasyonları daha etkileşimli ve yüksek görsel gerçeklikli bir simülasyon ortamında sunulabilir.

Dijital dönüşüm açısından bakıldığında, bu tür sistemlerin belediyelerde kademeli bir uygulama planıyla geliştirilmesi mümkündür. İlk aşamada mevcut kamera altyapısı, 3B kent modeli ve temel nesne tespiti bileşenleri birlikte test edilebilir. İkinci aşamada belirli pilot bölgelerde canlı kamera akışı, dijital ikiz görselleştirmesi ve konum doğruluğu analizleri entegre edilebilir. Üçüncü aşamada trafik, güvenlik, afet yönetimi ve saha ekipleri gibi farklı birimlerin ihtiyaçlarına göre özel alarm ve raporlama modülleri geliştirilebilir. Beş yıllık bir dijital dönüşüm perspektifinde ise bu altyapı; sürekli güncellenen 3B kent modeli, çoklu kamera entegrasyonu, yapay zekâ tabanlı olay algılama, sensör verileri, IoT sistemleri ve belediye karar destek panelleriyle bütünleşik çalışan bir akıllı şehir platformuna dönüştürülebilir.

Gerçek zamanlıya yakın kullanım açısından model çıkarım süresi, ray-casting süresi, Kalman filtreleme süresi ve CesiumJS görselleştirme güncelleme aralığı birlikte optimize edilmelidir. Bu kapsamda edge computing, GPU hızlandırma ve kare atlama stratejileri kullanılarak gecikme azaltılabilir. Böylece sistemin yalnızca kayıtlı video

veya fotoğraf verileri üzerinde değil, canlı kamera akışları üzerinde de uçtan uca test edilmesi mümkün hâle gelebilir.

Son olarak, önerilen sistemin daha geniş ölçekte uygulanabilmesi için veri güvenliği, kişisel verilerin korunması, mahremiyet, etik kullanım ilkeleri ve insan denetimli karar süreçleri dikkate alınmalıdır. Özellikle kamusal alanlarda kamera tabanlı analiz sistemlerinin kullanımı, teknik başarımın yanında hukuki ve toplumsal kabul açısından da değerlendirilmelidir. Bu nedenle gelecek çalışmaların yalnızca algoritmik doğruluk ve sistem performansına değil, aynı zamanda güvenilir, şeffaf ve sorumlu dijital ikiz uygulamalarının geliştirilmesine odaklanması önerilmektedir.

6. KAYNAKLAR

Aicardi I, Chiabrando F, Grasso N, Lingua A M, Noardo F, Spanò A, 2016, UAV Photogrammetry with Oblique Images: First Analysis on Data Acquisition and Processing, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B1, 835-842.

Alan Ö, 2025, Türkiye'nin Akıllı Şehir Stratejisi: Ulusal Yaklaşım, Küresel Perspektif ve Gelecek Vizyonu, Çevre, Şehir ve İklim Dergisi, Akıllı Şehirler Özel Sayısı, 16-37.

Avşar E Ö, 2006, Tarihi Köprülerin Digital Fotogrametri Tekniği Yardımıyla Modellenmesi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 105s, İstanbul.

Bemis S P, Micklethwaite S, Turner D, James M R, Akciz S, Thiele S T, vd., 2014, Ground-Based and UAV-Based Photogrammetry: A Multi-Scale, High-Resolution Mapping Tool For Structural Geology and Paleoseismology, Journal of Structural Geology, 69, 163-178.

Biljecki F, Stoter J, Ledoux H, Zlatanova S, Çöltekin A, 2015, Applications of 3D City Models: State of the Art Review, ISPRS International Journal of Geo-Information, 4(4), 2842-2889.

Boulos M N K, Blanchard B J, Walker C, Montero J, Tripathy A, Gutierrez-Osuna R, 2011, Web GIS in Practice X: A Microsoft Kinect Natural User Interface for Google Earth Navigation, International Journal of Health Geographics, 10(1), 45.

Colomina I, Molina P, 2014, Unmanned Aerial Systems for Photogrammetry and Remote Sensing: A Review, ISPRS Journal of Photogrammetry and Remote Sensing, 92, 79-97.

Cozzi P, Ring K, 2011, 3D Engine Design for Virtual Globes, A K Peters/CRC Press, 1. Baskı, Boca Raton.

Çelik M T, 2024, YOLOv8 Nesne Tespit Modeli ile Plastik Parça Yüzey Kusurlarının Gerçek Zamanlı Tespit Edilmesi, Bursa Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü, Yüksek Lisans Tezi, 69s, Bursa.

Ekercin S, Üstün B, 2004, Uzaktan Algılamada Yeni Bir Teknoloji: LiDAR, Jeodezi, Jeoinformasyon ve Arazi Yönetimi Dergisi, 91, 34-38.

Everingham M, Van Gool L, Williams C K I, Winn J, Zisserman A, 2010, The Pascal Visual Object Classes (VOC) Challenge, International Journal of Computer Vision, 88, 303-338.

Goodfellow I, Bengio Y, Courville A, 2016, Deep Learning, MIT Press, Cambridge, MA.

Grieves M, 2014, Digital Twin: Manufacturing Excellence Through Virtual Factory Replication, White paper, LLC, 1(2014), 1-7.

Gröger G, Kolbe T H, Nagel C, Hafele K H, 2012, OGC City Geography Markup Language (CityGML) Encoding Standard, Germany.

Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, vd., 2018, Recent Advances in Convolutional Neural Networks, Pattern Recognition, 77, 354-377.

Gündüz M Ş, 2023, Gerçek Zamanlı Videolardan Kalabalık ve Sosyal Mesafe Tespiti İçin YOLO Algoritmalarının Performans Analizi: COVID-19 Uygulaması, Iğdır Üniversitesi, Lisansüstü Eğitim Enstitüsü, Yüksek Lisans Tezi, 75s, Iğdır.

Hartley R, Zisserman A, 2004, Multiple View Geometry in Computer Vision, Cambridge University Press, 2. Baskı, Cambridge.

Jocher G, Qiu J, Liu M, Lyu S, Akyon F C, Kalfaoglu M E, 2026, Ultralytics YOLO26: Unified Real-Time End-to-End Vision Models, arXiv preprint, arXiv:2606.03748.

Kalman R E, 1960, A New Approach to Linear Filtering and Prediction Problems, Journal of Basic Engineering, 82(1), 35-45.

Karatzinis G D, Apostolikas N A, Boutalis Y S, Papakostas G A, 2023, Fuzzy Cognitive Networks in Diverse Applications Using Hybrid Representative Structures, International Journal of Fuzzy Systems, 25, 2534-2554.

Kraus K, 2007, Fotogrametri Cilt 1, Nobel Yayın Dağıtım, 7. Baskı, 462s, Ankara.

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

LeCun Y, Bottou L, Bengio Y, Haffner P, 1998, Gradient-Based Learning Applied to Document Recognition, Proceedings of the IEEE, 86(11), 2278-2324.

Lin T Y, Maire M, Belongie S, Hays J, Perona P, Ramanan D, vd., 2014, Microsoft COCO: Common Objects in Context, European Conference on Computer Vision, 740-755.

Maboudi M, Homaei M R, Song S, Malihi S, Saadatseresht M, Gerke M, 2023, A Review on Viewpoints and Path Planning for UAV-Based 3-D Reconstruction, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16, 5026-5048.

Makineci H B, 2016, İnsansız Hava Araçları Lidar Etkileşimi, Geomatik, 1(1), 19-23.

Möller T, Trumbore B, 1997, Fast, Minimum Storage Ray-Triangle Intersection, Journal of Graphics Tools, 2(1), 21-28.

O'Shea K, 2015, An Introduction to Convolutional Neural Networks, arXiv preprint, arXiv:1511.08458.

Özemir I, Uzar M, 2016, İHA İle Fotogrametrik Veri Üretimi, VI. UZAL-CBS Sempozyumu, 5-7 Ekim 2016, Adana.

Padilla R, Passos W L, Dias T L B, Netto S L, da Silva E A B, 2021, A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit, Electronics, 10(3), 279.

Pan S J, Yang Q, 2010, A Survey on Transfer Learning, IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345-1359.

Redmon J, Divvala S, Girshick R, Farhadi A, 2016, You Only Look Once: Unified, Real-Time Object Detection, IEEE Conference on Computer Vision and Pattern Recognition, 27-30 Haziran 2016, Las Vegas, 779-788.

Remondino F, Gerke M, 2015, Oblique Aerial Imagery – A Review, Photogrammetric Week 2015, Stuttgart, 75-83.

Rezatofighi H, Tsoi N, Gwak J, Sadeghian A, Reid I, Savarese S, 2019, Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression, IEEE/CVF Conference on Computer Vision and Pattern Recognition, 658-666.

Roriz R, Cabral J, Gomes T, 2022, Automotive LiDAR Technology: A Survey, IEEE Transactions on Intelligent Transportation Systems, 23(7), 6282–6297.

Shepard D, 1968, A Two-Dimensional Interpolation Function for Irregularly-Spaced Data, Proceedings of the 1968 23rd ACM National Conference, 517-524.

Szeliski R, 2022, Computer Vision: Algorithms and Applications, Springer, 2. Baskı, Cham.

Tao F, Liu W R, Liu J H, vd., 2018, Digital Twin and Its Potential Application Exploration, Computer Integrated Manufacturing Systems, 24, 1-18.

Verykokou S, Ioannidis C, 2024, Oblique Aerial Images: Geometric Principles, Relationships and Definitions, Encyclopedia, 4, 234-255.

Vosselman G, Maas H G, 2010, Airborne and Terrestrial Laser Scanning, Whittles Publishing, Dunbeath, Scotland.

Wang C Y, Bochkovskiy A, Liao H Y M, 2023, YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors, IEEE/CVF

Conference on Computer Vision and Pattern Recognition, 18-22 Haziran 2023, Vancouver, 7464-7475.

Wehr A, Lohr U, 1999, Airborne Laser Scanning: An Introduction and Overview, ISPRS Journal of Photogrammetry and Remote Sensing, 54(2-3), 68-82.

Weiss K, Khoshgoftaar T M, Wang D, 2016, A Survey of Transfer Learning, Journal of Big Data, 3, 1-40.

Welch G, Bishop G, 2006, An Introduction to the Kalman Filter, University of North Carolina at Chapel Hill, Department of Computer Science, Technical Report TR 95-041.

White G, Zink A, Codecá L, Clarke S, 2021, A Digital Twin Smart City for Citizen Feedback, Cities, 110, 103064.

White J C, Wulder M A, Varhola A, Vastaranta M, Coops N C, Cook B D, vd., 2013, A Best Practices Guide for Generating Forest Inventory Attributes from Airborne Laser Scanning Data Using the Area-Based Approach, Information Report FI-X-10, Canadian Forest Service, Canadian Wood Fibre Centre, Pacific Forestry Centre, 50s, Victoria.

Yastıklı N, 2003, GPS/IMU Verilerini Kullanarak Hava Fotoğraflarının Doğrudan Yöneltilmesi ve Birleştirilmiş Blok Dengeleme Olanakları, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi, 121s, İstanbul.

Yilmaz A, Javed O, Shah M, 2006, Object Tracking: A Survey, ACM Computing Surveys, 38(4), 1-45.

Zhang W, Qi J, Wan P, Wang H, Xie D, Wang X, vd., 2016, An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation, Remote Sensing, 8(6), 501.

Zhang Z, 2000, A Flexible New Technique for Camera Calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11), 1330-1334.

Zhong T, Lin J, Wang H, He J, Zhong L, Chen L, 2024, An Initial Exploration of Digital Twin Technology for Hydraulic Structures Based on Cesium, DEAI '24: Proceedings of the 2024 Guangdong-Hong Kong-Macao Greater Bay Area

International Conference on Digital Economy and Artificial Intelligence, 19-21 Ocak 2024, Dongguan, 762-771.

Zhou T, Hasheminasab S M, Habib A, 2021, Tightly-Coupled Camera/LiDAR Integration for Point Cloud Generation from GNSS/INS-Assisted UAV Mapping Systems, ISPRS Journal of Photogrammetry and Remote Sensing, 180, 336-356.

İnternet Kaynakları

1- https://docs.ultralytics.com/tr/models/yolo26, 12.01.2026

2- https://openlayers.org/en/latest/apidoc/, 09.02.2026

3- https://developers.arcgis.com/javascript/latest/, 09.02.2026

4- https://docs.mapbox.com/mapbox-gl-js/, 10.02.2026

5- https://threejs.org/docs/, 10.02.2026

6- https://www.khronos.org/webgl, 14.02.2026

7- https://cesium.com/platform/cesiumjs/, 14.02.2026

8- https://www.ogc.org/standards/3dtiles/, 16.02.2026

9- https://cesium.com/why-cesium/3d-tiles/, 17.02.2026

10- https://akillisehirler.csb.gov.tr/wp-content/uploads/EylemPlani.pdf, 20.04.2026

ÖZGEÇMİŞ

Diğer konular

EKLER

EK 1. Nesne Tespiti Çıktı Dosyası Veri Yapısı

Alan adı	Örnek değer
image	frame_001.jpg
image_width	1920
image_height	1080
detection_id	1
class_id	0

class_name	person
confidence	0.7983
x1	926.5
y1	639.8
x2	970.6
y2	756.4
bbox_width	44.1
bbox_height	116.6
bbox_area	5142.06
center_x	948.55
center_y	698.10
bottom_x	948.55
bottom_y	756.40
selection_strategy	highest_confidence
person_count_detected	1
needs_review	false

EK 2. RTK Referans Noktaları Veri Yapısı

Alan adı	Örnek değer
image	frame_001.jpg
point_id	RTK_001
scene_group	orta
occlusion	açık
rtk_lat	39.55073550
rtk_lon	32.69948580
rtk_height	1104.85

EK 3. Homografi ve Lokal Enterpolasyon Kontrol Noktaları Veri Yapısı

Alan adı	Örnek değer
image	frame_001.jpg
point_id	CP_001
pixel_x	948.50
pixel_y	756.40
world_lat	39.55073550
world_lon	32.69948580

EK 4. Ray-Casting Çıktı Dosyası Veri Yapısı

Alan adı	Örnek değer
image	frame_001.jpg
object_id	1
class_name	person
confidence	0.7983
bbox_ref	bottom_center
pixel_x	948.50
pixel_y	756.40

active_tileset	OBLIK
ray_lat	39.55073491
ray_lon	32.69948612
ray_height	1104.92
hit	true

EK 5. Video Tabanlı YOLO Tespit Çıktısı Örnek JSON Veri Yapısı

Alan adı	Örnek değer
video_info.fps	30.0
video_info.total_frames	3148
video_info.resolution	[1920, 1080]
video_info.avg_processing_fps	23.63
video_info.avg_gpu_usage_pct	0.0
video_info.total_processing_time_sec	133.23
camera_position.lat	39.55083633118592

camera_position.lon	32.69922104391364
camera_position.height	1105.1012892753522
frames[0].frame_id	0
frames[0].timestamp_ms	0
frames[0].detections[0].track_id	1
frames[0].detections[0].label	person
frames[0].detections[0].confidence	0.7983
frames[0].detections[0].bbox	[926.5, 639.8, 970.6, 756.4]
frames[0].detections[0].bottom_center_px	[948.5, 756.4]
frames[0].detections[0].homography_lat	39.5507355
frames[0].detections[0].homography_lon	32.6994858
frames[1].frame_id	1
frames[1].timestamp_ms	33
frames[1].detections[0].track_id	1
frames[1].detections[0].label	person
frames[1].detections[0].confidence	0.7972
frames[1].detections[0].bbox	[928.0, 638.4, 971.1, 752.4]
frames[1].detections[0].bottom_center_px	[949.5, 752.4]
frames[1].detections[0].homography_lat	39.5507317
frames[1].detections[0].homography_lon	32.6994858

EK 6. YOLO Model Doğrulama Sonuçları Veri Yapısı

Alan adı	Örnek değer
model	YOLO26-s
conf_threshold	0.30
iou_threshold	0.50
frame_count	3148
matched_total	186
diagonal_correct	171
wrong_class_matches	3

unmatched_false_positive	15
unmatched_false_negative	78
class_precision	0.816
class_recall	0.816
class_f1	0.816
bbox_precision	0.824
bbox_recall	0.824
bbox_f1	0.824
mean_iou	0.812
median_iou	0.825
iou_50_rate	0.824
iou_75_rate	0.710
iou_90_rate	0.125
avg_inference_time_ms	26.27
derived_fps	38.07
total_unmatched_errors	93

EK 7. Koordinat Dönüşüm Yöntemlerine Ait Hata Analizi Veri Yapısı

Alan adı	Örnek değer	Alan adı	Örnek değer
image	frame_001.jpg	interp_error_m	0.26
object_id	1	ray_lidar_error_m	0.15
bbox_ref	bottom_center	ray_oblik_error_m	0.08
pixel_x	948.50	hom_interp_diff_m	0.17
pixel_y	756.40	hom_ray_lidar_diff_m	0.55
point_id	RTK_001	hom_ray_oblik_diff_m	0.36

scene_group	orta	best_method	ray_oblik
occlusion	açık		
rtk_lat	39.55073550		
rtk_lon	32.69948580		
rtk_height	1104.85		
rtk_e	474135.42		
rtk_n	4377895.31		
hom_lat	39.55073170		
hom_lon	32.69948580		
hom_e	474135.42		
hom_n	4377894.89		
interp_lat	39.55073324		
interp_lon	32.69948604		
interp_e	474135.44		
interp_n	4377895.06		
ray_lidar_lat	39.55073610		
ray_lidar_lon	32.69948730		
ray_lidar_height	1104.96		
ray_lidar_e	474135.55		
ray_lidar_n	4377895.38		
ray_lidar_hit	true		
ray_oblik_lat	39.55073491		
ray_oblik_lon	32.69948612		
ray_oblik_height	1104.92		
ray_oblik_e	474135.45		
ray_oblik_n	4377895.24		
ray_oblik_hit	true		
hom_error_m	0.42		

EK 8. CesiumJS Ray-Casting Girdi JSON Veri Yapısı

Alan adı	Örnek değer
image_width	1920
image_height	1080
camera_position.lat	39.55083633118592
camera_position.lon	32.69922104391364
camera_position.height	1105.1012892753522
camera_orientation.heading_deg	115.35

camera_orientation.pitch_deg	-3.70
camera_orientation.roll_deg	1.35
cesium.fov	66.35
cesium.aspect	1.7777
items[0].image	frame_001.jpg
items[0].object_id	1
items[0].class_name	person
items[0].confidence	0.7983
items[0].bbox_ref	bottom_center
items[0].pixel_x	948.5
items[0].pixel_y	756.4
items[0].image_width	1920
items[0].image_height	1080

EK 9. 3B Model Yerleştirme JSON Veri Yapısı

Alan adı	Örnek değer	Alan adı	Örnek değer
models[0].id	car_1	models[3].height	1105.00
models[0].class_name	car	models[3].heading_deg	110.0
models[0].lat	39.55083633	models[3].pitch_deg	0.0
models[0].lon	32.69922104	models[3].roll_deg	0.0
models[0].height	1105.25	models[3].scale	1.0
models[0].heading_deg	115.35	models[3].uri	models/bus.glb

models[0].pitch_deg	0.0
models[0].roll_deg	0.0
models[0].scale	1.0
models[0].uri	models/car.glb
models[1].id	person_1
models[1].class_name	person
models[1].lat	39.55081221
models[1].lon	32.69918455
models[1].height	1104.80
models[1].heading_deg	0.0
models[1].pitch_deg	0.0
models[1].roll_deg	0.0
models[1].scale	1.0
models[1].uri	models/person.glb
models[2].id	truck_1
models[2].class_name	truck
models[2].lat	39.55079012
models[2].lon	32.69930124
models[2].height	1105.10
models[2].heading_deg	90.0
models[2].pitch_deg	0.0
models[2].roll_deg	0.0
models[2].scale	1.0
models[2].uri	models/truck.glb
models[3].id	bus_1
models[3].class_name	bus
models[3].lat	39.55076544
models[3].lon	32.69926518

EK 10. YOLO Tespiti Sonucunda BBox Referans Noktalarının Çıkarılması

```
# Gerekli kütüphaneler
import cv2
import pandas as pd
from ultralytics import YOLO

# BBox referans noktalarının hesaplanması
def get_bbox_points(x1, y1, x2, y2):
```

```
    bbox_width = x2 - x1
    bbox_height = y2 - y1
    bbox_area = bbox_width * bbox_height
    center_x = (x1 + x2) / 2.0
    center_y = (y1 + y2) / 2.0
    bottom_x = (x1 + x2) / 2.0
    bottom_y = y2
    return {
        "bbox_width": bbox_width,
        "bbox_height": bbox_height,
        "bbox_area": bbox_area,
        "center_x": center_x,
        "center_y": center_y,
        "bottom_x": bottom_x,
        "bottom_y": bottom_y
    }

# Uygun tespitin seçilmesi
def select_detection(detections, strategy="largest_area"):
    if not detections:
        return None
    if strategy == "highest_conf":
        return max(detections, key=lambda d: d["confidence"])
    if strategy == "largest_area":
        return max(detections, key=lambda d: d["bbox_area"])
    raise ValueError(f"Bilinmeyen seçim stratejisi: {strategy}")

# YOLO çıktılarının veri tablosuna dönüştürülmesi
def detections_to_rows(results, image_name, image_width, image_height):
```

EK 10. (Devam) YOLO Tespiti Sonucunda BBox Referans Noktalarının Çıkarılması

```
    rows = []
    detection_id = 1
    for result in results:
        for box in result.boxes:
            x1, y1, x2, y2 = box.xyxy[0].cpu().numpy()
            confidence = float(box.conf[0])
            class_id = int(box.cls[0])
```

```
            class_name = result.names[class_id]
            bbox_points = get_bbox_points(x1, y1, x2, y2)
            row = {
                "image": image_name,
                "image_width": image_width,
                "image_height": image_height,
                "detection_id": detection_id,
                "class_id": class_id,
                "class_name": class_name,
                "confidence": confidence,
                "x1": float(x1),
                "y1": float(y1),
                "x2": float(x2),
                "y2": float(y2),
                **bbox_points
            }
            rows.append(row)
            detection_id += 1
    return pd.DataFrame(rows)
```

EK 11. Etiketli BBox ve IoU Tabanlı Model Doğrulama

```
# Gerekli kütüphaneler
import numpy as np

# YOLO formatındaki etiketlerin piksel tabanlı BBox formatına dönüştürülmesi
def xywhn_to_xyxy(box, img_w, img_h):
    x_c, y_c, w, h = box
    x1 = (x_c - w / 2) * img_w
```

```
    y1 = (y_c - h / 2) * img_h
    x2 = (x_c + w / 2) * img_w
    y2 = (y_c + h / 2) * img_h
    return [x1, y1, x2, y2]

# İki BBox arasındaki IoU değerinin hesaplanması
def bbox_iou(box_a, box_b):
    ax1, ay1, ax2, ay2 = box_a
    bx1, by1, bx2, by2 = box_b
    ix1 = max(ax1, bx1)
    iy1 = max(ay1, by1)
    ix2 = min(ax2, bx2)
    iy2 = min(ay2, by2)
    iw = max(0, ix2 - ix1)
    ih = max(0, iy2 - iy1)
    intersection_area = iw * ih
    area_a = max(0, ax2 - ax1) * max(0, ay2 - ay1)
    area_b = max(0, bx2 - bx1) * max(0, by2 - by1)
    union_area = area_a + area_b - intersection_area
    if union_area <= 0:
        return 0.0
    return intersection_area / union_area

# Tahmin BBox'ları ile gerçek BBox'ların eşleştirilmesi
def match_predictions_to_gt(gt_items, pred_items, iou_threshold):
    candidates = []
    for g_idx, gt in enumerate(gt_items):
        for p_idx, pred in enumerate(pred_items):
            iou = bbox_iou(gt["box"], pred["box"])
```

EK 11. (Devam) Etiketli BBox ve IoU Tabanlı Model Doğrulama

```
            if iou >= iou_threshold:
                candidates.append((iou, g_idx, p_idx))
    candidates.sort(reverse=True, key=lambda x: x[0])
    used_gts = set()
    used_preds = set()
    matches = []
    for iou, g_idx, p_idx in candidates:
```

```
        if g_idx in used_gts or p_idx in used_preds:
            continue
        used_gts.add(g_idx)
        used_preds.add(p_idx)
        matches.append((g_idx, p_idx, iou))
    unmatched_gts = [i for i in range(len(gt_items)) if i not in used_gts]
    unmatched_preds = [i for i in range(len(pred_items)) if i not in
used_preds]
    return matches, unmatched_gts, unmatched_preds

# Precision, recall ve F1 değerlerinin hesaplanması
def safe_divide(a, b):
    return a / b if b != 0 else 0.0
def compute_f1(tp, fp, fn):
    precision = safe_divide(tp, tp + fp)
    recall = safe_divide(tp, tp + fn)
    f1 = safe_divide(2 * precision * recall, precision + recall)
    return precision, recall, f1
```

EK 12. Homografi Tabanlı Pikselden UTM Koordinatına Dönüşüm

```
# Gerekli kütüphaneler
import cv2
import numpy as np
import pandas as pd
from pyproj import CRS, Transformer

# WGS 84 ve UTM dönüşüm nesnelerinin oluşturulması
```

```
def get_utm_crs_from_lat_lon(lat, lon):
    zone = int((lon + 180) // 6) + 1
    if lat >= 0:
        epsg = 32600 + zone
    else:
        epsg = 32700 + zone
    return CRS.from_epsg(epsg), epsg
def build_transformers(mean_lat, mean_lon):
    utm_crs, epsg = get_utm_crs_from_lat_lon(mean_lat, mean_lon)
    wgs84 = CRS.from_epsg(4326)
    to_utm = Transformer.from_crs(wgs84, utm_crs, always_xy=True)
    to_wgs84 = Transformer.from_crs(utm_crs, wgs84, always_xy=True)
    return to_utm, to_wgs84, epsg

# Kontrol noktalarının UTM koordinatlarına dönüştürülmesi
def prepare_control_points(control_df, to_utm):
    e_values = []
    n_values = []
    for _, row in control_df.iterrows():
        e,        n        =        to_utm.transform(float(row["world_lon"]),
float(row["world_lat"]))
        e_values.append(e)
        n_values.append(n)
    control_df["world_e"] = e_values
    control_df["world_n"] = n_values
    return control_df

# Homografi matrisinin hesaplanması
```

EK 12. (Devam) Homografi Tabanlı Pikselden UTM Koordinatına Dönüşüm

```
def compute_homography(control_df):
    required = ["pixel_x", "pixel_y", "world_e", "world_n"]
    for col in required:
        if col not in control_df.columns:
            raise ValueError(f"Kontrol noktalarında eksik kolon: {col}")
    if len(control_df) < 4:
        raise ValueError("Homografi için en az 4 kontrol noktası gerekir.")
```

```
    src_pixels                    =                    control_df[["pixel_x",
"pixel_y"]].to_numpy(dtype=np.float32)
    dst_world = control_df[["world_e", "world_n"]].to_numpy(dtype=np.float32)
    H, mask = cv2.findHomography(src_pixels, dst_world, method=0)
    if H is None:
        raise RuntimeError("Homografi matrisi hesaplanamadı.")
    return H

# Piksel koordinatının homografi ile UTM düzlemine aktarılması
def apply_homography(H, pixel_x, pixel_y):
    point = np.array([[[float(pixel_x), float(pixel_y)]]], dtype=np.float32)
    transformed = cv2.perspectiveTransform(point, H)
    e = float(transformed[0, 0, 0])
    n = float(transformed[0, 0, 1])
    return e, n
```

EK 13. IDW Tabanlı Lokal Enterpolasyon ile Konum Kestirimi

```
# Gerekli kütüphaneler
import numpy as np

# IDW tabanlı lokal enterpolasyon
def   idw_interpolate(pixel_x,   pixel_y,   control_df,   k=6,   power=2.0,
min_distance=1e-6):
    dx = control_df["pixel_x"].to_numpy(dtype=float) - float(pixel_x)
```

```
    dy = control_df["pixel_y"].to_numpy(dtype=float) - float(pixel_y)
    distances = np.sqrt(dx ** 2 + dy ** 2)
    if np.min(distances) < min_distance:
        idx = int(np.argmin(distances))
        row = control_df.iloc[idx]
        return      float(row["world_e"]),      float(row["world_n"]),      1,
float(np.min(distances))
    k_eff = min(k, len(control_df))
    nearest_idx = np.argsort(distances)[:k_eff]
    nearest_dist = distances[nearest_idx]
    nearest = control_df.iloc[nearest_idx]
    weights = 1.0 / (nearest_dist ** power)
    weights = weights / np.sum(weights)
    interp_e               =               float(np.sum(weights               *
nearest["world_e"].to_numpy(dtype=float)))
    interp_n               =               float(np.sum(weights               *
nearest["world_n"].to_numpy(dtype=float)))
    mean_dist = float(np.mean(nearest_dist))
    return interp_e, interp_n, k_eff, mean_dist
```

EK 14. Konumsal Doğruluk ve Hata Metriklerinin Hesaplanması

```
# Gerekli kütüphaneler
import numpy as np
import pandas as pd

# RTK referans noktasına göre yatay hata hesabı
def distance_m(e1, n1, e2, n2):
    if pd.isna(e1) or pd.isna(n1) or pd.isna(e2) or pd.isna(n2):
```

```
        return np.nan
    return float(np.sqrt((e1 - e2) ** 2 + (n1 - n2) ** 2))

# RMSE hesabı
def rmse(values):
    values = pd.Series(values).dropna()
    if len(values) == 0:
        return np.nan
    return float(np.sqrt(np.mean(values ** 2)))

# Yöntem bazlı hata özetlerinin üretilmesi
def make_method_summary(df, method_error_cols):
    rows = []
    for method_name, error_col in method_error_cols.items():
        if error_col not in df.columns:
            continue
        errors = df[error_col].dropna()
        if len(errors) == 0:
            continue
        rows.append({
            "method": method_name,
            "count": int(len(errors)),
            "mean_error_m": float(errors.mean()),
            "rmse_m": rmse(errors),
            "median_error_m": float(errors.median()),
            "min_error_m": float(errors.min()),
            "max_error_m": float(errors.max()),
            "std_error_m": float(errors.std(ddof=1)) if len(errors) > 1 else
0.0
```

EK 14. (Devam) Konumsal Doğruluk ve Hata Metriklerinin Hesaplanması

```
        })
    return pd.DataFrame(rows)

# En düşük hata veren yöntemin belirlenmesi
def choose_best_method(row, method_error_cols):
    best_method = None
    best_error = None
```

```
    for method_name, error_col in method_error_cols.items():
        if error_col not in row.index:
            continue
        error = row[error_col]
        if pd.isna(error):
            continue
        if best_error is None or error < best_error:
            best_error = error
            best_method = method_name
    return best_method
```

EK 15. Sabit Hız Modeline Dayalı Kalman Filtreleme

```
# Gerekli kütüphaneler
import numpy as np
import pandas as pd

# Hız hesabı
def compute_speed_kmh(e, n, times):
    e = np.asarray(e, dtype=float)
```

```
    n = np.asarray(n, dtype=float)
    times = np.asarray(times, dtype=float)
    speed = np.full(len(e), np.nan)
    step_dist = np.full(len(e), np.nan)
    for i in range(1, len(e)):
        if not all(np.isfinite([e[i], n[i], e[i - 1], n[i - 1], times[i],
times[i - 1]])):
            continue
        dt = times[i] - times[i - 1]
        if dt <= 0:
            continue
        d = np.sqrt((e[i] - e[i - 1]) ** 2 + (n[i] - n[i - 1]) ** 2)
        step_dist[i] = d
        speed[i] = (d / dt) * 3.6
  return speed, step_dist

# Fiziksel hız sınırına göre aykırı değer kontrolü
def filter_outliers_by_speed(df, max_speed_kmh=35.0):
    df = df.copy()
    df["raw_speed_before_filter_kmh"] = np.nan
    df["raw_step_before_filter_m"] = np.nan
    df["is_speed_outlier"] = 0
    last_valid_e = None
    last_valid_n = None
    last_valid_t = None
    for i in range(len(df)):
        e = df.loc[i, "raw_e"]
        n = df.loc[i, "raw_n"]
        t = df.loc[i, "timestamp_sec"]
```

EK 15. (Devam) Sabit Hız Modeline Dayalı Kalman Filtreleme

```
        if not np.isfinite(e) or not np.isfinite(n) or not np.isfinite(t):
            df.loc[i, "is_speed_outlier"] = 1
            continue
        if last_valid_e is None:
            last_valid_e = e
            last_valid_n = n
            last_valid_t = t
```

```
            continue
        dt = t - last_valid_t
        if not np.isfinite(dt) or dt <= 0:
            df.loc[i, "is_speed_outlier"] = 1
            df.loc[i, "raw_e"] = np.nan
            df.loc[i, "raw_n"] = np.nan
            continue
        step_m = np.sqrt((e - last_valid_e) ** 2 + (n - last_valid_n) ** 2)
        speed_kmh = (step_m / dt) * 3.6
        df.loc[i, "raw_speed_before_filter_kmh"] = speed_kmh
        df.loc[i, "raw_step_before_filter_m"] = step_m
        if speed_kmh > float(max_speed_kmh):
            df.loc[i, "is_speed_outlier"] = 1
            df.loc[i, "raw_e"] = np.nan
            df.loc[i, "raw_n"] = np.nan
        else:
            last_valid_e = e
            last_valid_n = n
            last_valid_t = t
    return df

# Sabit hız modeline dayalı Kalman filtresi
def       kalman_filter_2d(times,       measurements_e,       measurements_n,
process_noise=0.5, measurement_noise=25.0):
    times = np.asarray(times, dtype=float)
    z_e = np.asarray(measurements_e, dtype=float)
    z_n = np.asarray(measurements_n, dtype=float)
    valid = np.isfinite(z_e) & np.isfinite(z_n)
    if valid.sum() == 0:
```

EK 15. (Devam) Sabit Hız Modeline Dayalı Kalman Filtreleme

```
        raise ValueError("Kalman filtresi için geçerli ölçüm yok.")
    first_idx = int(np.where(valid)[0][0])
    x = np.array([z_e[first_idx], z_n[first_idx], 0.0, 0.0], dtype=float)
    P = np.eye(4) * 100.0
    H = np.array([
        [1, 0, 0, 0],
        [0, 1, 0, 0]
```

```
], dtype=float)
R = np.eye(2) * float(measurement_noise)
out_e = np.full(len(times), np.nan)
out_n = np.full(len(times), np.nan)
out_ve = np.full(len(times), np.nan)
out_vn = np.full(len(times), np.nan)
last_time = times[first_idx]
for i in range(len(times)):
    if i < first_idx:
        continue
    dt = float(times[i] - last_time)
    if not np.isfinite(dt) or dt <= 0:
        dt = 1.0 / 30.0
    F = np.array([
        [1, 0, dt, 0],
        [0, 1, 0, dt],
        [0, 0, 1, 0],
        [0, 0, 0, 1]
    ], dtype=float)
    q = float(process_noise)
    Q = q * np.array([
        [dt ** 4 / 4, 0, dt ** 3 / 2, 0],
        [0, dt ** 4 / 4, 0, dt ** 3 / 2],
        [dt ** 3 / 2, 0, dt ** 2, 0],
        [0, dt ** 3 / 2, 0, dt ** 2]
    ], dtype=float)
    x = F @ x
    P = F @ P @ F.T + Q
    if valid[i]:
```

EK 15. (Devam) Sabit Hız Modeline Dayalı Kalman Filtreleme

```
        z = np.array([z_e[i], z_n[i]], dtype=float)
        y = z - H @ x
        S = H @ P @ H.T + R
        K = P @ H.T @ np.linalg.inv(S)
        x = x + K @ y
        P = (np.eye(4) - K @ H) @ P
    out_e[i] = x[0]
```

```
        out_n[i] = x[1]
        out_ve[i] = x[2]
        out_vn[i] = x[3]
        last_time = times[i]
    return out_e, out_n, out_ve, out_vn
```

EK 16. CesiumJS Ortamında Ray-Casting ve 3B Model Yerleştirme

```
// Piksel koordinatından kamera ışını oluşturulması
function pixelToRayDirection(px, py, imageWidth, imageHeight) {
  if (!camDir || !camRight || !camUp) {
    captureConfiguredCameraVectors();
  }
  const u = (Number(px) / Number(imageWidth)) * 2.0 - 1.0;
  const v = 1.0 - (Number(py) / Number(imageHeight)) * 2.0;
```

```
  const rightComponent = Cesium.Cartesian3.multiplyByScalar(
    camRight,
    u * tanH,
    new Cesium.Cartesian3()
  );
  const upComponent = Cesium.Cartesian3.multiplyByScalar(
    camUp,
    v * tanV,
    new Cesium.Cartesian3()
  );
  const dir = new Cesium.Cartesian3();
  Cesium.Cartesian3.add(camDir, rightComponent, dir);
  Cesium.Cartesian3.add(dir, upComponent, dir);
  Cesium.Cartesian3.normalize(dir, dir);
  return dir;
}

// Ray-casting ile 3B yüzey kesişiminin hesaplanması
function castRayForItem(item) {
  const imageWidth = Number(item.image_width || rayInput.image_width ||
IMAGE_WIDTH);
  const imageHeight = Number(item.image_height || rayInput.image_height ||
IMAGE_HEIGHT);
  const direction = pixelToRayDirection(
    Number(item.pixel_x),
    Number(item.pixel_y),
    imageWidth,
    imageHeight
  );
```

EK 16. (Devam) CesiumJS Ortamında Ray-Casting ve 3B Model Yerleştirme

```
  const ray = new Cesium.Ray(camPos, direction);
  let hitPosition = null;
  const hit3D = viewer.scene.pickFromRay(ray);
  if (hit3D && hit3D.position) {
    hitPosition = hit3D.position;
  }
  if (!hitPosition) {
```

```
    const hitTerrain = viewer.scene.globe.pick(ray, viewer.scene);
    if (hitTerrain) {
      hitPosition = hitTerrain;
    }
  }
  if (!hitPosition) {
    return {
      image: item.image,
      object_id: item.object_id,
      class_name: item.class_name,
      confidence: item.confidence,
      bbox_ref: item.bbox_ref,
      pixel_x: item.pixel_x,
      pixel_y: item.pixel_y,
      active_tileset: activeTilesetName,
      ray_lat: null,
      ray_lon: null,
      ray_height: null,
      hit: false
    };
  }
  const geo = cartesianToWgs84(hitPosition);
  return {
    image: item.image,
    object_id: item.object_id,
    class_name: item.class_name,
    confidence: item.confidence,
    bbox_ref: item.bbox_ref,
    pixel_x: item.pixel_x,
```

EK 16. (Devam) CesiumJS Ortamında Ray-Casting ve 3B Model Yerleştirme

```
    pixel_y: item.pixel_y,
    active_tileset: activeTilesetName,
    ray_lat: geo.lat,
    ray_lon: geo.lon,
    ray_height: geo.height,
    hit: true
  };
```

```
}

// GLB / glTF tabanlı 3B modelin sahneye eklenmesi
function createModelAtWgs84(options) {
  const lat = Number(options.lat);
  const lon = Number(options.lon);
  const height = Number(options.height || 0.0);
  const headingDeg = Number(options.heading_deg || 0.0);
  const pitchDeg = Number(options.pitch_deg || 0.0);
  const rollDeg = Number(options.roll_deg || 0.0);
  const scale = Number(options.scale || 1.0);
  const uri = options.uri;
  const position = Cesium.Cartesian3.fromDegrees(lon, lat, height);
  const hpr = new Cesium.HeadingPitchRoll(
    Cesium.Math.toRadians(headingDeg),
    Cesium.Math.toRadians(pitchDeg),
    Cesium.Math.toRadians(rollDeg)
  );
  const orientation = Cesium.Transforms.headingPitchRollQuaternion(position,
hpr);
  const entity = viewer.entities.add({
    id: options.id,
    name: options.id,
    position: position,
    orientation: orientation,
    model: {
      uri: uri,
      scale: scale,
      minimumPixelSize: 48,
```

EK 16. (Devam) CesiumJS Ortamında Ray-Casting ve 3B Model Yerleştirme

```
      maximumScale: 20000
    },
    properties: {
      class_name: options.class_name,
      lat: lat,
      lon: lon,
      height: height,
```

```
      uri: uri
    }
  });
  viewer.scene.requestRender();
  return entity;
}
```