

Tek RGB Görüntüden 3B Rekonstrüksiyon ve BIM-Uyumlu Nesnelere: Dijital İkiz Projeleri için Bir Uygulama

From a Single RGB Image to 3D Reconstruction and BIM-Compatible Objects: An Application for Digital Twinning Projects

Serdar ERİŞEN ¹

¹Hacettepe Üniversitesi, Mimarlık Fakültesi, Mimarlık Bölümü, Ankara, Türkiye

Öz

Akıllı kampüsler için dijital ikiz projeleri genellikle yoğun sensör ağları ve pahalı 3B tarama altyapıları ile yürütülmektedir. Bu çalışma, yalnızca tek bir RGB görüntüden hareketle 2B sahne ayrıştırma ve monoküler derinlik kestirimini birleştirip 3B rekonstrüksiyon, nesne düzeyi ayrıştırma ve Bina Bilgisi Modelleme (BIM)'e otomatik nesne aktarımını mümkün kılan ve ek sensörlere ihtiyaç duymayan pratik bir yaklaşım sunar. Temelde SERNet-Former_v2 ile üretilen sınır-duyarlı 2B sınıf haritaları ve tek-görüntü derinlik çıktıları, geri izdüşüm ile nokta bulutuna dönüştürülür. Düzlem çıkarımı ile duvar, zemin, tavan gibi nesneler ve kapı, pencere gibi açıklıklar ayrıştırılır. Kolon, duvar gibi yapısal elemanlar nesne düzeyinde tanımlanır, BIM uyumlu geometri ve parametreler elde edilir. Bu yaklaşım, akıllı kampüs hedefiyle çok kapsamlı dijital ikiz çalışması olarak geliştirilen TUM2TWIN projesine öncülük eden ScanNet ve TUM CMS Indoor Point Cloud gibi veri kümeleri üzerinde sınanmıştır: Tek-görüntü koşulunda 3B doğruluk 78.8%, 2B→3B projeksiyon tutarlılığı ise 81.9% oranında ölçülmüştür. Derinlik için RMSE, 3B için ICP-RMSE ve 3D RMSE kullanılmış, 0.02 RMSE değeri elde edilmiştir. Tek A100-40GB GPU'da tekli RGB girdilerde 512×512 piksel kırpma boyutu ile örnek başına altı saniyenin altında çalışma süresi sağlanmıştır. Elde edilen BIM-uyumlu mekansal parametreler semantik modellere ve Model Öngörülü Denetim tasarımlarına programatik girdi sağlayarak, erken faz değerlendirmelerini düşük maliyet ve hızlı geri bildirimle ölçeklenebilir kılacaktır. Ek alan uyarlaması gerektirmeyen bu hat, mimarlık ve mühendislik uygulamalarında sahaya hızlı geçiş ve dijital belgeleme için uygulanabilir bir alternatif sunar.

Anahtar Kelimeler: Semantik ayrıştırma, Monoküler derinlik, 3B rekonstrüksiyon, Bina Bilgisi Modelleme

Abstract

Digital twinning projects for smart campuses are often carried out using dense sensor networks (IoT), advanced building and energy management systems (BMS/EMS), and expensive 3D scanning infrastructures such as LiDAR or multi-camera rigs. This study presents a practical, sensor-free approach that starts from a single RGB image and integrates 2D scene parsing and monocular depth estimation into 3D reconstruction, object-level segmentation, and automatic BIM transfer. Boundary-sensitive semantic maps generated by SERNet-Former_v2 and single-image depth outputs are reprojected into point clouds. Planar extraction then identifies major architectural elements (walls, floors, ceilings), while openings (doors, windows) and structural components (columns) are segmented at the object level. The pipeline produces BIM-compatible geometry and parameters. The method is evaluated on benchmark datasets such as ScanNet and TUM CMS Indoor Point Cloud, pioneering comprehensive digital twinning projects such as TUM2TWIN, with an aim of smart campus for making the TUM campuses as the best-surveyed places. Under the single-image setting, 3D accuracy reaches 78.8%, while 2D→3D projection consistency achieves 81.9%. Depth estimation is measured by RMSE, and 3D reconstruction by ICP-RMSE and 3D RMSE, with a minimal reconstruction error of 0.02 RMSE. The entire workflow runs on a A100-40GB GPU with single inputs having 512×512 pixel crop sizes under 6s per sample. The resulting BIM-compliant spatial parameters will serve as programmatic inputs to semantic models and Model Predictive Control (MPC) schemes, enabling cost-efficient and scalable early-stage assessments. Without requiring domain adaptation, this workflow offers a rapid and effective alternative for field deployment and digital documentation in architecture and engineering applications.

Keywords: Semantic segmentation, Monocular depth, 3D reconstruction, Building Information Modeling

I. GİRİŞ

Akıllı kampüs projelerinde yapay zeka tabanlı dijital ikiz ve algılama-anlama çözümleri, enerji verimliliği, kullanıcı konforu, iş sağlığı-güvenliği ve üretkenlik hedeflerini veri-merkezli karar süreçleriyle desteklemeyi amaçlar [1]. Uygulamada bu hedefe çoğu zaman yoğun sensör ağları, gelişmiş bina, enerji yönetim sistemleri ile

LiDAR, çoklu kamera düzenekleri gibi pahalı 3-boyutlu (3B) tarama altyapıları üzerinden yaklaşırlar [2–4]. Bununla birlikte söz konusu altyapılar, toplam sahip olma maliyeti (donanım + kurulum + kalibrasyon + bakım), devreye alma süresi ve işletim karmaşıklığı nedeniyle özellikle ilk faz değerlendirmelerinde ve hızlı geri bildirim gerektiren senaryolarda ciddi engeller oluşturur. Scan-to-BIM literatüründeki kapsamlı derlemeler de bu maliyet/zorluk profilini ayrıntılı biçimde raporlamaktadır [5, 6]. Öte yandan mevcut güvenlik kameraları ile mobil cihazların varlığı ve görsel verinin yaygın bulunurluğu, kapalı mekan veri kümelerinin ölçeği ve çeşitliliğiyle birlikte yalnızca görüntüye dayalı hatların sahada uygulanabilirliğini artırmaktadır [7, 8]. Son olarak 2-boyutlu (2B) semantik ayrıştırma (*semantic segmentation*) ve monoküler derinlik (*monocular depth*) kestirimindeki güncel ilerlemeler, tek görüntüden üretilen 3B geometrinin doğruluk ve kararlılığını pratik olarak kullanılabilir seviyelere taşımış, bu sayede ek sensör olmadan ve yalnızca görüntüye dayalı bir veri işleme hattının gerçekçi bir alternatif haline gelmesini sağlamıştır [9–12].

Son yıllarda kampüs ve hatta şehir ölçeğinde dijital ikizler, farklı kaynaklardan gelen görüntü, nokta bulutu, semantik model gibi verilerin ortak bir referans sisteminde bütünleştirilmesini ve akış analizleri destekleyen veri setlerinin oluşmasını hızlandırmıştır. Bu bağlamda TUM2TWIN, Münih Teknik Üniversitesi (TUM) kampüslerinin yeryüzünde en iyi belgelenen mekanlar olması hedefiyle, çoklu veri temsillerini görüntüler, nokta bulutları ve semantik modelleri bir araya getiren kapsamlı bir dijital ikiz projesi olarak öne çıkmaktadır [13]. Bu makalenin amacı, TUM2TWIN benzeri dijital ikiz projeleri ve akıllı mekan uygulamaları kapsamında, mekan ve nesne parametrelerini, tek RGB görüntüden doğrudan türeten 2B→3B→BIM veri işleme hattı tasarlamak ve bunu mühendislik ölçütleri ile değerlendirmektir. Önerilen hat dört bileşenden oluşur: (i) 2-boyutlu (2B) sahne ayrıştırma (*scene parsing*) ile zemin, duvar, tavan, kapı, pencere, kolon, mobilya, teçhizat gibi iç mekanın temel sınıflarının piksel düzeyinde haritalanması, (ii) tek görüntüden derinlik kestirimi ile kaba geometri çıkarımı, (iii) bu iki çıktının nokta bulutuna 3B rekonstrüksiyon vasıtasıyla dönüştürülmesi ve geometrik tutarlılığın artırılması, (iv) elde edilen yüzey, nesne bilgisinin bina bilgisi modelleme (*Building Information Modeling*) (BIM)-uyumlu (*Industry Foundation Classes*) (IFC/ifcJSON) nesne ve parametrelere çevrilmesidir [14, 15]. Önerilen bu yapı sayesinde ek donanım gerektirmeden, duvar, döşeme, tavan, kapı ve pencere gibi temel yapı elemanlarının, oluşturulan nokta bulutlarından elde edilen x, y, z koordinatları ve renk bilgisine bağlı olarak, şekil, konum ve boyut bilgisi programatik olarak elde edilmektedir. Bu sayede tasarım–uygulama döngüleri kısılır ve saha kararları ölçülebilir hale gelir.

Yöntemin çekirdeğinde, iç mekan sınır duyarlılığı yüksek, kapı, pencere, kolon gibi nadir sınıflarda kararlı davranan *SERNet-Former_v2* adlı yapay sinir ağı bulunmaktadır [16]. Çok-ölçekli bağlamı (*multi-scale context*) yoğun biçimde birleştiren ve sınıf sınırlarında seçici güçlendirme yapan dikkat-temelli (*attention-based*) modüller (*DABM/DABG*, *DAFN*), 2B maskelerin kenar bütünlüğünü büyük ölçüde koruyarak 3B aşamaya aktarır [16]. Derinlik modülü tek görüntüden görece ölçekli bir derinlik haritası üretir ve semantik maskelerle geri izdüşüm yoluyla birleşerek nokta bulutu oluşturur [17]. Eşikleme, morfolojik temizlik, düzlemsellik ve diklik gibi geometrik kısıtlarla yüzey sürekliliği güçlendirilir. Düzlem kümeleri ve açıklık bölgeleri kural setleri ile *IfcWall*, *IfcSlab*, *IfcCeiling*, *IfcDoor*, *IfcWindow* gibi BIM sınıflarına dönüştürülür. Çıktılar *ifcJSON*, *OBJ*, *glTF* gibi formatlarla dışa aktarılabilir [15]. Bu çalışmada IFC opsiyonel bir hedef olarak anılmaktadır [14].

Bu çalışmanın itki noktasını, ek sensörlere ihtiyaç duymadan görsel veri işleme hattının akıllı mekanlarda erken aşama karar desteği için düşük giriş maliyeti ve hızlı geri bildirim sağlaması oluşturmaktadır. Kavramsal düzeyde veri-merkezlilik, ölçülebilirlik ve ölçeklenebilirlik hedefleri doğrultusunda görüntüden 3B ve BIM-uyumlu bilgiye giden doğrudan bir yol önerilmektedir. Mühendislik düzeyinde ise ayrıştırma, derinlik, 3B ve BIM katmanlarında nicel ve nitel değerlendirme yapılarak katkılarının ölçülebilir biçimde ortaya konması sağlanmaktadır. Bu yönüyle çalışma, mimarlık, tasarım ve yapı disiplini ile bilgisayarlı görü, veri işleme ve mekansal bilişim mühendisliği arasında ortak bir zemin kurmaktadır.

Kısaca bu çalışma şu sorulara odaklanır: Tek bir RGB görüntüsünden, iç mekanın temel elemanları için güvenilir 2B ayrıştırma ile görece ölçekli derinlik üretimi nasıl birlikte kurgulanabilir? 2B ayrıştırma ve derinlik birleşimi, tek-görüntü koşulunda yeterli kalitede nokta bulutu üretilip BIM-uyumlu nesne ve parametreleri çıkarılabilir mi? Sınır duyarlılığı ve nadir sınıfların korunması, 3B rekonstrüksiyon ve BIM nesne doğruluğunu hangi ölçüde iyileştirir? Önerilen hattın çıktıları, semantik veri modelleme ve Model Öngörülü Denetim (*Model Predictive Control*) (MPC) gibi ileri denetim çerçevelerine programatik girdi sağlayacak olgunluğa getirilebilir mi?

Akıllı kampüs projelerine yeni bir boyut katan TUM2TWIN projesine başlangıç noktası oluşturan ScanNet ve TUM CMS Indoor Point Cloud veri setleri bahsedilen araştırma soruları çerçevesinde önerilen yöntemle ele alınmıştır. Buna bağlı olarak TUM CMS veri setinin nesne etiketlerinin oluşumunda referans aldığı S3DIS veri seti de çalışma kapsamına dahil edilmiştir [19]. Bu çerçevede desteklenen 3B rekonstrüksiyon seviyesi iki bağlamda sınanmıştır. ScanNet_v2 ve S3DIS gibi referans veri tabanlarında, üretilen nokta bulutları üzerinden 3B doğruluk,

ScanNet_v2 üzerinde 78.8%, S3DIS üzerinde 75.3% seviyesine ulaşılmıştır [7, 17, 17]. Ayrıca, TUM CMS Indoor Point Cloud veri setine bağlı nokta bulutlarından 2B görüntü ve derinlik çiftleri türetilmiş, bu görüntüler üzerinde 2B sahne ayrıştırma gerçekleştirilmiş ve 2B→3B projeksiyon tutarlılığı piksel-bazlı düzlemde 81.9% oranında gözlemlenmiştir [17, 19]. Bu sonuçlar, 3B öğrenmeye gerek kalmadan 2B→3B etiket üretimi ve değerlendirmenin pratik olduğunu, nadir sınıflar ve açıklıkların korunmasının 3B eşleme aşamasında ölçülebilir fayda sağladığını göstermektedir. Çalışmada alan uyarlaması kullanılmamıştır. Genelleme için veri artırma ve sınır-duyarlı son-işleme tercih edilmiştir.

Elde edilen BIM-uyumlu mekan ve nesne parametreleri, Brick/Project Haystack gibi semantik veri modelleme katmanı ile köprü kurma ve MPC gibi ileri denetim çerçevelerine programatik girdi sağlama olanağı sağlar [3, 4, 20]. Zon topolojisi, alan ve hacim, açıklık oranları, konfor ve kapasite kısıtları otomatik olarak desteklenebilir. Böylece manuel etiketleme ve haritalama yükü azalır. Denetim tasarımları yinelenen ve ölçeklenebilir hale gelir. Önerilen hattın düşük giriş maliyeti ve hızlı geri bildirim özellikleri, kampüs ölçeğinde dijital belgeleme ve varlık yönetimi akışlarını destekler.

Araştırmanın başlıca katkıları şu şeklide özetlenebilir:

- Yalnızca görüntü girdisiyle çalışan, 2B→3B→BIM zincirini uçtan uca kuran ve sensöre ihtiyaç duymayan görsel veri işleme hattı.
- *SERNet-Former_v2* ile sınır-duyarlılığı yüksek 2B ayrıştırma, açıklıkların korunması ve düzlem tutarlılığının artışı sayesinde 3B kalitede gözle görülür kazanımlar.
- Düzlem saptama ve açıklık analiziyle temel yapı elemanlarının BIM sınıf ve şemalarına eşlenmesi ve *ifcJSON* ve *OBJ*'e aktarım.
- TUM2TWIN gibi dijital ikiz ve akıllı mekan projeleri için yeni yöntemleri destekleyen veri setleri ve örneklerin oluşturulması.

Bölüm 2'de ilgili çalışmalar özetlenmiştir. Bölüm 3'te yöntem, 2B ayrıştırma, derinlik, 3B rekonstrüksiyon ve BIM nesne üretimi ile ayrıntılandırılmıştır. Bölüm 4'te deney protokolleri ve uygulama bağlamı sunulmuş, Bölüm 5'te bulgular ve tartışma verilmiştir. Bölüm 6 araştırmayı sonuçlandırmaktadır.

II. İLGİLİ ÇALIŞMALAR

Bu bölüm, alan yazınına beş doğrultuda özetler ve önerilen tek görüntü tabanlı 2B→3B→BIM yaklaşımının konumunu netleştirir: (i) tek görüntüden geometrik çıkarım ve monoküler derinlik çalışmalarında ölçek belirsizliği ve iç mekana özgü zorluklar, (ii) 2B sahne ayrıştırmada (*Convolutional Neural Network*) CNN-Transformer hibritleri, çok-ölçekli füzyon ve sınır-duyarlı işlemler, (iii) 2D semantiğin geri izdüşüm ve füzyon ile 3B

rekonstrüksiyona bağlanması, (iv) Scan-to-BIM hattında düzlem ve boşluk çıkarımı ve BIM şemalarına (*IFC/ifcJSON*) eşleme, (v) Brick/Project Haystack gibi semantik veri modelleri ve MPC bağlamında mekansal parametrelerin kullanımı bu başlıklar arasında sayılabilir. İnceleme kapsamı, kapalı mekan odaklı, RGB veya düşük duyuşal gereksinimli girdileri ele alan, sınır duyarlılığı ve açıklıkların korunmasına vurgu yapan, çıktıları ölçülebilir 3B geometri ve BIM'e taşıyan çalışmaları önceler. Bu çerçeve, yalnız RGB girdisiyle uçtan uca ve hafif bir tasarımla 2B→3B→BIM zincirini kuran bu önerinin literatürdeki boşluğu nasıl adreslediğini ortaya koyar. Ek sensöre ihtiyaç duymayan kurulum BIM'e doğrudan aktarımda, ortalama kesişim-birleşim oranı (*mean Intersection over Union*) (*mIoU*), karekök ortalama kare hata (*Root Mean Square Error*) (*RMSE*), yinelenen en yakın nokta (*Iterative Closest Point*) (*ICP*)-*RMSE* gibi sade ölçüt setleri üzerinden değerlendirilmiştir [12, 21].

2.1. Tek Görüntüden Geometrik Çıkarım

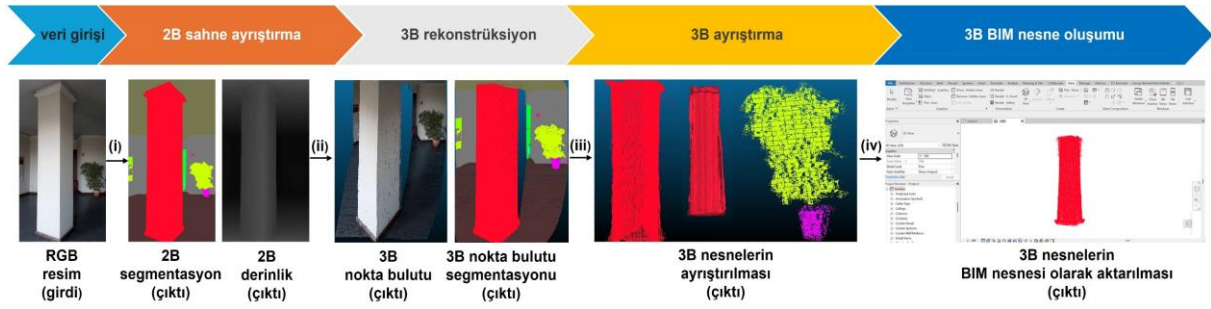
Monoküler derinlik kestirimi, geniş ölçekli ön-eğitim (*pre-training*) ve sahaya özgü ince-ayar (*fine-tuning*) ile iç mekanlarda görece ölçekli derinlik haritaları üretmeyi mümkün kılmıştır [11, 17]. Ölçek belirsizliği, düşük aydınlatma, yansımalar, saydam yüzeyler gibi iç mekana özgü zorluklar ve kapı/pencere çerçeveleri, kolonlar gibi nadir veya ince nesnelerin yapay zeka modelleri ile tam olarak algılanması hala kritik problemlerdir. Ölçüm ve doğrulama tarafında *RMSE* gibi metrikler yaygın olarak kullanılmakta, sınır bölgelerindeki duyarlılık için ek kenar düzenleyicileri tercih edilmektedir [12, 17]. Bu bağlamda bu çalışma, tek görüntü senaryosunu temel alır. Gerektiğinde az-atımlı veya video kiplerini yalnızca rekonstrüksiyon kararlılığını göstermek amacıyla devreye alır.

2.2. 2-Boyutlu (2B) Sahne Ayrıştırma

İç mekan sahnelerinde semantik ayrıştırma, CNN-Transformer hibritleri ve çok-ölçekli füzyon katmanlarıyla önemli ilerleme kaydetmiştir [9, 10]. Başarılı yaklaşımların ortak özelliği, (i) sınıf sınırlarını koruyan kenar-duyarlı işlemler, (ii) nadiren görülen sınıflar için dengesiz veri stratejileri ve (iii) hafif mimari bileşenlerdir. *SERNet-Former_v2*, bu üç gereksinimi birlikte ele alır: DABG/DABM ile seçici güçlendirme, DAFN ile çok-ölçekli bağlam birleştirme ve çıktıda güven haritaları üretimi bunlar arasında sayılabilir [16]. Bu sayede kapı/pencere gibi açıklıkların sınır bütünlüğü korunur ve 3B aşamada düzlem/boşluk ayrımı kolaylaşır.

2.3. 2B'den 3B'ye Rekonstrüksiyon

2B semantik haritalar ve tek görüntü derinlik, geri izdüşüm ile nokta bulutu (*point cloud*) üretimine bağlanır [11]. Literatürde tek-görüntüde mutlak ölçek belirsizliği, düzlemsellik ve diklik gibi geometrik öncüller ile kısmen dengelenir. Çoklu görüntü varlığında TSDF füzyonu ile yüzey sürekliliği artırılır



Şekil 1. Önerilen işlem akışı. Veri girişi → (i) 2B sahne ayrıştırma → (ii) 3B rekonstrüksiyon → (iii) 3B nesne ayrıştırma → (iv) BIM nesne oluşumu.

[11, 12, 22]. Çoğu yaklaşım, ölçüm sensörlerine (RGB-D/LiDAR) dayanır. Bu çalışma, (i) 2D sınır duyarlılığını yüksek tutarak açıklıkların 3B’de korunmasını, (ii) geometrik kurallarla düzlemler ve açıklıkların ayrıştırılmasını ve (iii) elde edilen nokta bulutundan ölçülebilir BIM-uyumlu çıktılara geçişi hedefler.

2.4. Scan-to-BIM ve Semantik Bütünleşme

Scan-to-BIM araştırmalarında, düzlem saptama (RANSAC), açıklık çıkarımı (morfoloji ve oran kısıtları) ve BIM-uyumlu (IFC/ifcJSON) şemalara aktarım temel adımdır [5, 14, 15, 19, 23]. Nesne kümeleri, açıklık ölçüleri ve konum ilişkileri için kural kümeleri yaygındır. ifcJSON ve OBJ gibi formatlara dışa aktarım, CAD/BIM kanallarına entegrasyonu kolaylaştırır. Mevcut literatürde sensör verisine bağımlılık yaygındır. Buna karşılık bu çalışmada önerilen yaklaşım, yalnızca RGB görüntü ile 2B→3B→BIM zincirini çalıştırarak düşük maliyet ve hızlı devreye alma avantajı sağlar.

2.5. Semantik Veri Modelleri ve MPC Bağı

Akıllı binalarda semantik veri modelleme (Brick, Project Haystack) ekipman-zon-nokta ilişkilerini bilgi grafi olarak düzenler ve denetim tasarımına programatik girdi sağlar [3, 4, 20]. Zon topolojisi, alan,

hacim ve açıklık oranları gibi mekansal büyüklüklerin otomatik elde edilmesi, ölçeklenebilir denetim kurulumlarının ön koşuludur. Bu çalışma, 2B→3B hattı üzerinden BIM-uyumlu mekansal öğeleri ve temel geometrik parametreleri üretmek, semantik grafiğe ve MPC’de durum ve çıktı tanımları ve kısıt bileşenlerine veri sağlama potansiyeli taşır.

2.6. Alan Yazınındaki Boşluk ve Bu Çalışmanın Konumlanması

Özetle literatür, 2D ayrıştırma, tek görüntü derinlik ve 3B rekonstrüksiyonun ayrı ayrı olgunlaştığını, ancak sensörsüz koşulda sınır-duyarlı 2D → ölçülebilir 3D → BIM-uyumlu nesne ve parametre zincirinin uçtan uca ve hafif bir tasarımla gösterilmesinin sınırlı kaldığını işaret eder [11, 12]. Bu çalışmanın katkısı, (i) yalnızca tek bir RGB görüntüsünden başlayarak kararlı bir 2D→3D→BIM akışı kurması, (ii) açıklıklar ve nadir sınıflarda sınır bütünlüğünü koruyarak 3B ve BIM düzeyinde ölçülebilir kazanımlar üretmesi ve (iii) elde edilen çıktıların semantik veri modelleri üzerinden dijital ikiz projeleri ve MPC gibi ileri denetim katmanlarına veri sağlayabilecek nitelikte olmasıdır.

III. YÖNTEM

Bu çalışmada, 3B mekansal bilgi ve nesne bilgisinin, görsel algılama dışında sensörlere ihtiyaç duymayan iç

Tablo 1. Önerilen yöntemin veri işlem akışı (i–iv): Girdi–çıkı ve ölçütler.

Aşama	Amaç	Girdi	Çıkı	Ölçüt(ler)
(i) 2B Segmentasyon +Derinlik	2B sınıf haritaları ve görece ölçekli derinlik üretmek	RGB tek görüntü I	Olasılık hacmi S , güven Q_s ; derinlik D , güven Q_d	$mIoU$, $pAcc$, $RMSE$
(ii) 3B Rekonstrüksiyon	Nokta bulutu ve düzlem kümeleri elde etmek	(S, Q_s, D, Q_d, K) , varsa (R, t)	Semantik nokta bulutu P , düzlemler $\{\Pi_k\}$	$ICP-RMSE$, $3D-RMSE$
(iii) Nesne Ayrıştırma	3B buluttan nesne örnekleri çıkarmak	P , $\{\Pi_k\}$	$\{O_j\}$ (sınıf, konum, boyut, ilişkiler)	$mIoU$, $3D-RMSE$
(iv) BIM Aktarım	Nesneleri makine-okunur BIM’e taşımak	$\{O_j\}$	$ifcJSON$, OBJ	$ifcClass$

Sembol tanımları: I : tek RGB görüntü, S : sınıf olasılık hacmi, Q_s : segmentasyon güveni, D : monoküler derinlik, Q_d : derinlik güveni, K : kamera içsel matrisi, (R, t) : kamera pozunu (dışsallar), P : sınıf/güven etiketli 3B nokta bulutu, $\{\Pi_k\}$: RANSAC ile çıkarılan düzlem kümeleri, $\{O_j\}$: nesne örnekleri (sınıf, konum, boyut, ilişkiler).

Ölçütler: $mIoU$: ortalama sınıflar arası IoU (2B), $RMSE$: kök-ortalama-kare hata (derinlik), $ICP-RMSE$: ICP sonrası $RMSE$ (3B hizalama), $3D-RMSE$: 3B nokta/parametre hatası.

mekanlarda yapay zeka ve ileri düzey veri işleme yöntemleriyle yüksek isabet oranı ve tekrarlanabilir biçimde görüntüden üretilmesi amaçlanmıştır. Bu amaç doğrultusunda, tek bir RGB görüntüyü girdi alan bir işlem akışı tasarlanmış, sırasıyla (i) 2B sahne ayrıştırma ve monoküler derinlik üretimi, (ii) 3B rekonstrüksiyon, (iii) semantik işaretli 3B buluttan nesne bazında ayrıştırma, (iv) sonuçların BIM-uyumlu (*ifcJSON*) nesne ve parametre verisine dönüştürülmesi adımları planlanmıştır (Şekil 1). Her adımın girdi-çıkışı ve bu çalışmada benimsenen sadeleştirilmiş değerlendirme ölçütleri Tablo 1’de özetlenmektedir.

3.1. Genel İşlem Akışı

Tek girdili çerçevede, $I \in \mathbb{R}^{H \times W \times 3}$ RGB görüntüsü, zemin, duvar, tavan, kapı, pencere, kolon, mobilya, teçhizat gibi iç mekan sınıflarının piksel düzeyi dağılımını veren olasılık hacmi S ve buna eşlik eden güven haritası Q_s ile sonuçlanacak biçimde bu çalışmada 2B görsel veri işleme ağı olarak kullanılan *SERNet-Former_v2* sinir ağına iletilir. Monoküler derinlik kestirim modülü aynı görüntüden görece ölçekli derinlik haritası D ve güven Q_d üretir. Bu çiftler, kamera içsel parametreleri K (varsa dışsal poz R, t) ile birlikte geri izdüşüm adımına aktarılır.

Her piksel (u, v) ve derinlik $z = D(u, v)$ için kamera koordinatlarında 3B nokta

$$x_c = z K^{-1} [u \ v \ 1]^T \quad (1)$$

olarak elde edilir. Poz biliniyorsa dünya eksenlerine dönüşüm

$$x_w = R^T (x_c - t) \quad (2)$$

şeklinde yapılır. Elde edilen ham nokta bulutunda RANSAC ile düzlem kümeleri ayrıştırılır [23]. Diklik, koşutluk kısıtlarıyla duvar–zemin–tavan yüzeyleri kararlı hale getirilir ve gerekli durumlarda *Delaunay/Poisson* temelli yüzeyleme uygulanır [24, 25].

Üretilen sınıf işaretli 3B dağılım P , tohumlama–bölge büyütme–düzlem bağlama dizisiyle nesne örneklerine ayrılır. Son aşamada, her nesnenin sınıfı, konumu ve boyut parametreleri BIM-uyumlu bir şemaya eşlenir. Çıktı, projelerde yaygın kullanılan *ifcJSON* veya *OBJ* biçimlerinde dışa aktarılır. Bu akış Şekil 1’de görselleştirilmiştir ve ayrıntıları aşağıda sunulmuştur.

3.2. 2B Sahne Ayrıştırma

Bu adımın hedefi, iç mekan sınıflarında sınır duyarlılığı yüksek, kapı, pencere çerçeveleri ve kolon gibi nadir sınıflarda kararlı ayrıştırmalar üretmektir. *SERNet-Former_v2*, konvolüsyon tabanlı yerel özellik çıkarımı ile dikkat katmanlarının küresel bağlam avantajını birleştiren modern bir gövde üzerinde çalışır. Model, sahne sınırlarında seçici güçlendirme yapan *DABG/DABM* blokları ve çok-ölçekli semantiği kayıpsız biçimde birleştiren *DAFN* füzyon katmanı ile

donatılmıştır [16]. Böylece, 2B maske tahmini sırasında kenar bütünlüğü korunur. Kapı, pencere gibi açıklıklar net biçimde ayırt edilir ve bu nitelikler doğrudan 3B aşamaya taşınır. Eğitim (*training*) ve veri hazırlığı aşamasında girdiler 512×512 piksel kırpma boyutlu (*crop size*) çözünürlükte normalize edilir. Çok ölçekli kırpma (0.5–1.5), $\pm 10^\circ$ döndürme, yatay çevirme ve hafif renk titreşimi (*jitter*) ile çeşitlilik sağlanır. Sınıf-tabanlı karışım teknikleriyle kenar sürekliliği bozulmadan yeni örnekler üretilir. Kayıp (*loss*) fonksiyonu olarak Çapraz Entropi ve *Dice* bileşimi tercih edilir. Sahne sınırlarında hata duyarlılığını artırmak için sınır piksel ağırlıklı bir varyant uygulanır. Sınıf dengesizliği, sınıf frekanslarına ters-log ağırlıklandırma ile dengelenir [16]. Model çıktısı olasılık hacmi S ile güven haritası Q_s ’dir. Ayrıştırmanın 2B değerlendirmesi *mIoU* ile raporlanmıştır. Başka bir deyişle *mIoU*, 2B sahnelerde elde edilen nesne etiketlerinin konumunun verilen referansla birebir eşleşip eşleşmediğinin doğrudan ölçütü olarak ele alınmaktadır.

Bu aşamanın önemli bir tasarım kararı, 2B ayrıştırma ile 3B kullanım arasında net bir arayüz tanımlamaktır. Haritalar tek bir renk paleti ve tekdüze sınıf kümesi ile üretilir. İlerideki adımlarda (geri izdüşüm, nesne ayrıştırma) kayan eşiklerin neden olacağı sapmaları azaltmak için güven haritası referans alınarak düşük güvenli pikseller temizlenir. Bu sayede kapı, pencere gibi açıklık sınıfları için boşluk–kütle ayrımı daha az belirsizlik içerir.

3.3. Monoküler Derinlik ve 3B Rekonstrüksiyon

Monoküler derinlik üretimi, tek bir görüntüden görece ölçekli bir derinlik alanı D elde etmeyi amaçlar. Tekrarlayan dokular, parlama, cam yüzeyler gibi iç mekana özgü örüntüler dikkate alınarak eğitilen model, kenar koruma ve ölçek kararlılığı odaklı düzenleyicilerle desteklenir. Bu bileşenin değerlendirmesi *RMSE* ile verilir. Çalışmanın odağı gereği *RMSE*, 2B sahnelerde metrik bazlı hata ölçütü olarak benimsenmiştir.

Geri izdüşüm birleşimi (2B’den 3B’ye). Geri izdüşüm adımı, 2B ayrıştırma ve derinliğin geometrik birleşimidir. Kamera içsel matrisi K ve piksel (u, v) için 3B nokta

$$x_c = z K^{-1} [u \ v \ 1]^T, \rightarrow z = D(u, v) \quad (3)$$

şeklinde hesaplanır. Poz bilgisi varsa dünya koordinatlarına dönüşüm

$$x_w = R^T (x_c - t) \quad (4)$$

ile yapılır. Böylelikle her piksel, semantik etiketi ve güven skoru ile ilişkilendirilmiş 3B bir örneğe dönüşür.

Yüzey ve açıklık modelleme. Komşuluk tabanlı normal kestirimi, yüzey eğilimlerini belirler. RANSAC

ile duvar–zemin–tavan gibi büyük ölçekli düzlemler ayrıştırılır ve diklik, koşutluk ilişkileriyle tutarlılık artırılır. Kapı, pencere sınıfları için 2B maske bölgeleri 3B’ye taşınır. Yüzeyleme gerekiyorsa *Delaunay* veya *Poisson/SPSR* ile nokta bulutu üretilir. Kenar-koruyan düzeltmeler ve küçük boşluk doldurmalarla rekonstrüksiyon yüzeyi iyileştirilir.

Değerlendirme ölçütleri. 3B rekonstrüksiyonun performansı, kayıt sonrası *ICP-RMSE* ve nokta–nokta/parametre bazlı *3D RMSE* ile raporlanır [25]. Sonuç olarak *3D RMSE* nokta bazında ölçülen geometrik sapmaları ifade eder. Bu kapsamda elde edilen değerler, hem küresel hizalama kalitesini hem de yerel doğruluğu anlaşılır biçimde özetler. Başka bir anlatımla *3D RMSE*, hem konum, hem de derinlik ve yönelim ile ilgili hata ölçütü olup, oluşturulan 3B hakkında kesinlik bilgisi vermektedir.

3.4. Semantik 3B Nokta Bulutundan Nesne Bazında Ayrıştırma

Bu aşama, sınıf işaretli 3B nokta bulutunu nesne örneklerine ayrıştırarak üst katmandaki parametrik gereksinimler için hazır hale getirir. Girdi

$$\mathcal{P} = \{(x_i, c_i, q_i)\}_{i=1}^N \quad (5)$$

şeklinde. Burada $x_i \in \mathbb{R}^3$ nokta konumu, c_i sınıf etiketi ve $q_i \in [0,1]$ güven skorudur. Çıktı kümesi $\{O_j\}_{j=1}^M$ her nesne için sınıf c_j , konum (kütle merkezi μ), boyut (*Oriented Bounding Box*, yani yönelimli sınır kutusu, (*OBB*) ile (w, h, t)), yönelim ve belirsizlik (σ ; güven–ağırlıklı kovaryans temelli) alanlarını; ayrıca ilişkileri, örneğin, “kapı $\subset$ duvar”, içerir.

Tohumlama. 2B ayrıştırmadan gelen kapı, pencere, kolon maske bölgeleri 3B’de tohum altkümeleri başlatır. Duvar–zemin–tavan sınıflarındaki noktalar, Bölüm 3.3’te bulunan düzlemlere $|d(x_i, \Pi)| < \tau_\Pi$ koşuluyla bağlanır. τ_Π tipik olarak 2–5 cm aralığındadır. Düşük güvenli noktalar ($q_i < \tau_q$) ve çok küçük bileşenler, ilerideki adımlarda yanlış birleştirmeleri önlemek üzere elenir.

Bölge büyütme ve bağlama. k -NN komşuluk grafinde geometrik yakınlık, sınıf uyumu ve normal benzerliği birlikte değerlendirilir. Sığ öncelikli arama (*Breadth-first-search*) ile bölgeler genişletilir. Periyodik olarak düzlem yakalama uygulanarak yüzey kümeleri kararlı tutulur. İnce ayırım evresinde küçük bileşenler kaldırılır, delik, boşluk durumları semantik olarak doğrulanır.

Parametre çıkarımı. Nokta kümelerinin kovaryansı ve *Principle Component Analysis*, yani temel bileşenler analizi (*PCA*) tabanlı yönelim hesaplanır. En küçük hacimli *OBB* ile (w, h, t) boyutları elde edilir, kütle merkezi μ ve yerel eksenlere göre yönelim raporlanır. Belirsizlik matrisi

$$\Sigma = \frac{1}{\sum_{i \in I_j} q_i} \sum_{i \in I_j} q_i (x_i - \mu)(x_i - \mu)^T \quad (6)$$

şeklinde modellenir. Burada I_j , O_j nesnesine ait indeksler kümesini ifade eder. Ana eksenlerdeki karekök özdeğerler σ olarak raporlanır.

Tutarlılık ve değerlendirme. Tutarlılık denetimleri, duvar normaline diklik, paralel yüzeyler arasında koşutluk ve açıklıkların ilgili duvarla ilişkilendirilmesi kısıtlarını içerir. Nesne düzeyi başarımlar *3D RMSE* ile verilir. Gerektiğinde *ICP* sonrası hizalama sonucuna göre sonuç güncellenir [17].

3.5. BIM-Uyumlu Nesne/Parametre Üretimi

Nesne örnekleri O_j , BIM-uyumlu bir şemaya eşlenerek makine-okunur hale getirilir. Düzlem kümeleri sırasıyla duvar, zemin, tavan sınıflarına; açıklık kümeleri kapı, pencere sınıflarına; kolon kümeleri yapısal eleman kategorisine karşılık gelecek şekilde haritalanır. Her nesne için kalınlık, yükseklik, açıklık ebatları, konum ve yönelim parametreleri elde edilen 3D rekonstrüksiyonlar vasıtasıyla kaydedilir. Diklik, koşutluk ve kesit oranı denetimleriyle geometrik tutarlılık korunur. Çıktı veri modeli, ID, sınıf, geometri, öznitelikler ve ilişkiler alanlarını içerir. Gereksinime göre *ifcJSON* veya *OBJ* biçimlerine dışa aktarım yapılır. Ayrıca bu nesnelerin semantik veri katmanları ile programatik bağ kurabilmesi için ilişki alanları kaydedilir.

Bu süreç, tasarım, elle etiketleme ve manuel eşleme yükünü azaltırken, dijital ikizleme ve denetim senaryolarında hızlı geri bildirim sağlar. Özellikle HVAC (heating, ventilating and air conditioning) için Model Öngörülü Denetim gibi üst seviye uygulamalarda, zon topolojisi, alan, hacim ve açıklık oranları gibi büyüklükler doğrudan bu veriden türetilir [2, 20].

3.6. Hesaplama Profili ve Yeniden Üretilirlik

Önerilen işlem akışının, tek A100–40GB GPU üzerinde her RGB veri için 512×512 piksel ebatlı kırılmış girdilerle çalışması amaçlanmıştır. Bu bölümde ortaya konan yapı, iç mekanlarda görüntü→2B→3B→BIM doğrultusunda düşük maliyetli, hızlı ve ölçeklenebilir bir yol önermektedir. Deneyler ve Sonuçlar bölümünde, veri bağlamı, eğitim, ince-ayar protokolleri ve değerlendirme senaryoları ayrıntılandırılacaktır.

IV. DENEYLER VE SONUÇLAR

Bu bölüm, önerilen görüntü→2B→3B→BIM işlem akışının deneysel değerlendirmesini, uygulama ayrıntılarını ve sayısal çıktıları birlikte sunar. Önce veri kümeleri kısaca tanıtılır. Ardından seçilen model ve yöntemlerin ön-eğitim ve uygulama ayrıntıları verilir. Devamında, 2B ayrıştırma ve 3B rekonstrüksiyon deneylerinin protokolleri ve sonuçları ilgili tablo ve şekillerle paylaşılır. Bölüm, BIM’e aktarım deney düzeni ve çıktılarıyla

sonuçlandırılmıştır. Ayrıntılı yorumlar, karşılaştırmalı tartışma ve sınırlılıklar ise 5. Bölüm olan Bulgular ve Tartışma'da ele alınacaktır.

4.1. Veri Kümeleri

ADE20K (2B, sahne ayrıştırma). İç ve dış mekanlardan oluşan geniş ölçekli bir 2B semantik ayrıştırma kümesidir. Sahnelerdeki mimari elemanlar, mobilya ve teçhizat gibi sınıflar piksel düzeyinde etiketlenmiştir [27]. Modelin genel sahne semantiği ve sınıf sınırları açısından olgunlaşması için ön-eğitim kaynağı olarak kullanılmıştır.

SUN RGB-D (2B, derinlik, iç mekan). İç mekan odaklı ve RGB'ye ek olarak derinlik bilgisinin de bulunduğu bir referans veri tabanıdır [28]. Nesne çeşitliliği yüksektir. İç mekan derinlik ipuçlarının öğrenilmesi ve 2B semantik ile birlikte geometrik süreklilik kazanımı için ön-eğitim ve ince-ayar aşamasında kullanılmıştır.

NYU Depth v2 (2B, derinlik, iç mekan). İç mekan sahnelerinden oluşan derinlik ve RGB veri kümesidir [29]. Nitelikli etiketlemeleri ve standartlaştırılmış sahneleri sayesinde monoküler derinlik modülünün kalibrasyonu için tercih edilmiştir.

ScanNet_v2 (2B RGB-D, 3B, iç mekan taramaları). İç mekanlarda gezici tarayıcılarla elde edilen RGB-D video sekansları ve bunlardan üretilmiş 3B bulutları barındırır [7]. Tek-atımlı, az-atımlı, video tabanlı rekonstrüksiyon kiplerini kıyaslamak, ICP tabanlı hizalama ve 3B doğruluk değerlendirmeleri yapmak için kullanılmıştır.

S3DIS (3B, iç mekan). Büyük ofis, koridor, oda tiplerini kapsayan iç mekan taramalarından oluşur [18]. Düzlemsel yüzeylerin ve açıklıkların yoğun olduğu tipik mimari sahnelerde 3B değerlendirme yapmak üzere ScanNet_v2 ile birlikte kullanılmıştır.

TUM CMS Indoor Point Cloud (3B, pozlar, türetilmiş 2B set). Bu çalışmada, mevcut 3B nokta bulutları ve kayıtlı kamera pozları aracılığıyla 2B görüntü ve derinlik çiftleri türetilmiş, 2B sahne ayrıştırma haritaları üretilmiştir [19]. Bu haritalar geri izdüşüm ile 3B'ye taşınarak kalite denetimi yapılmış, böylece 3B öğrenmeye gerek duymadan 3B değerlendirmeye elverişli bir düzenek elde edilmiştir. Söz konusu veri seti S3DIS nesne etiketlerini benimsemiştir.

Bu çalışma kapsamında alan uyarlaması kullanılmamıştır. Veri çeşitliliği artırma ve sınır-duyarlı son-işleme ile sağlanmıştır.

4.2. Seçilen Model ve Yöntem: Ön-Eğitim ve Uygulama Ayrıntıları

Yöntem bölümünde bahsedilen akış, detayları verilen aşamalar, algoritmalar ve işlemler ardı sıra çalışan

kodlar şeklinde Python3 ortamında, tek A100-40GB GPU kullanılarak gerçekleştirilmiş ve elde edilen nokta bulutları, BIM arayüzüne görselleştirilmek üzere aktarılmıştır.

2B Sahne Ayrıştırma. Bu çalışmada 2B sahne ayrıştırma *SERNet-Former_v2* tercih edilmiştir. Hedef, iç mekan sınıflarında sınır duyarlılığını yükseltmek ve kapı, pencere, kolon gibi nadir sınıflarda kararlılığı artırmaktır. ADE20K, SUN RGB-D ve NYU-Depth v2 üzerinde *SERNet-Former_v2* modeli uygulanmıştır (Tablo 2). Bu kapsamda iç mekan sınıflarında sınır duyarlılığı ve nadir sınıf kararlılığı hedeflenmiştir. ADE20K üzerinde temel semantik temsiller, SUN RGB-D, NYU-Depth v2 üzerinde iç mekan bağlamı ve derinlik ipuçlarıyla hizalama sağlanmıştır. İnce-ayar aşamasında iç mekan sınıf kümesi (zemin, duvar, tavan, kapı, pencere, kolon, mobilya, teçhizat) ile tek görüntü koşulu hedeflenmiştir. Veri setleri üzerindeki ön-eğitim aşamalarında girdi çözünürlüğü 512×512 pikseldir. AdamW ile $\eta = 3 \times 10^{-4}$, cosine öğrenme hızı zamanlaması, 30–35 bölüm aralığında uygulanmıştır. Veri artırmada çok ölçekli kırpma (0.5–1.5), $\pm 10^\circ$ dönme, yatay çevirme ve hafif renk titreşimi (*jitter*) kullanılmış, seyrek sınıflar zenginleştirilmiştir. Çıktılar olasılık hacmi S ve güven Q_s olup, özet tablo, $mIoU$ ve piksel doğruluğu $pAcc$ değerlerini gösterir (Tablo 2).

Tablo 2. SERNet-Former_v2 2B sonuç karşılaştırması (ADE20K, SUN RGB-D, NYU-D_v2).

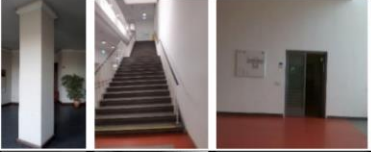
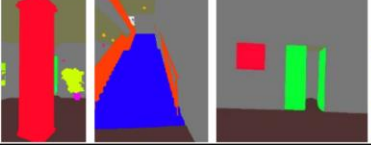
2B Segmentasyon		
Veri Kümesi	mIoU (%)	Parametre (Milyon)
ADE20K	60.9	245
2B Derinlik		
Veri Kümesi	pAcc (%)	Parametre (Milyon)
SUN RGB-D	~84	245
NYU-D_v2	~78	245

Monoküler Derinlik Kestirimi. İç mekan odaklı tek görüntü derinlik modeli, SUN RGB-D ve NYU-Depth v2 bağlamlarıyla ön-eğitim ve kalibrasyon görmüştür (Tablo 2). Çıktılar derinlik haritası D ve güven Q_d 'dir. 3B rekonstrüksiyon adımının derinlik girdisini oluşturur. Değerlendirme ölçütü $pACC$ ve $RMSE$ 'dir (Tablo 2 ve 3).

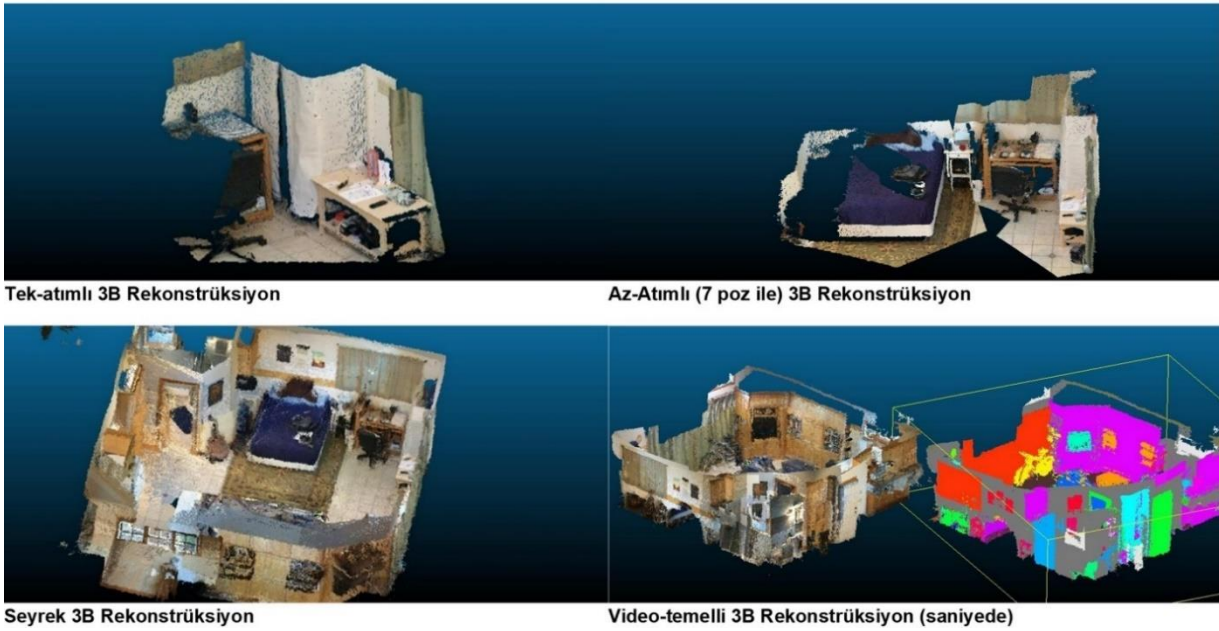
Tablo 3. 2B ve 3B hata payı.

2B Veri Seti	RMSE
SUN RGB-D	0.06–0.08
NYU-D_v2	0.07–0.09
3B Veri Seti	RMSE
ScanNet_v2	0.03
TUM CMS Indoor Point Cloud	0.02

3B Rekonstrüksiyon Hattı uygulaması. Geri izdüşümle üretilen nokta bulutu üzerinde komşuluk tabanlı normal kestirimi, RANSAC ile düzlem çıkarımı ve diklik, koşutluk denetimleri uygulanır. Gerekli sahnelerde *Delaunay* veya *Poisson/SPSR* ile nokta bulutu üretilir. Kenar koruyan düzeltme, delik

veri seti	TUM CMS Indoor Point Clouds	ScanNet_v2
RGB resim (girdi)		
derinlik (çıktı)		
segmentasyon (çıktı)		

Şekil 2. TUM CMS Indoor Point Cloud ve ScanNet_v2 veri setleri üzerinde 2B ayrıştırma örnekleri: [7, 16, 20, 30].



Şekil 3. ScanNet_v2 veri seti ile 3B Rekonstrüksiyonlar: Tek-atımlı, az-atımlı, seyrek, video-temelli.

doldurma ile yüzey sürekliliği artırılır. Kullanılan veri setleri, elde edilen sonuçların doğruluğunu değerlendirmede kullanılmıştır. Bu aşamada referans buluta ICP ile kayıt yapılır [26]. 3B değerlendirme iki metrikle raporlanmıştır: *mIoU* ve *RMSE* (Tablo 3 ve 4).

Tablo 4. 3B doğruluk (ScanNet, S3DIS, TUM CMS Indoor Point Cloud)

3D Segmentasyon / Rekonstrüksiyon	mIoU (%)
S3DIS	75.4
ScanNet_v2	78.8
TUM CMS Indoor Point Cloud	81.9

4.3. Deney Protokolleri, Uygulama Ayrıntıları ve Sonuçların Sunumu

2B Sahne Ayrıştırma: Protokol ve Sonuçlar. Bu çalışma boyunca tüm 2B deneyler tek görüntü koşulunda yürütülmüştür. Her veri kümesi ayrımları, literatürde yaygın standart bölünmelere göre oluşturulmuş, zemin, duvar, tavan, kapı, pencere, kolon, mobilya, teçhizat gibi sınıflamalar ayrıştırılmıştır. Çalışma kapsamında değerlendirme yalnızca sınıf başına ortalama olan *mIoU* üzerinden Tablo 2’de raporlanmıştır. Temsili ayrıştırma görselleri ve sınır duyarlılığı örnekleri Şekil 2’de verilmiştir.



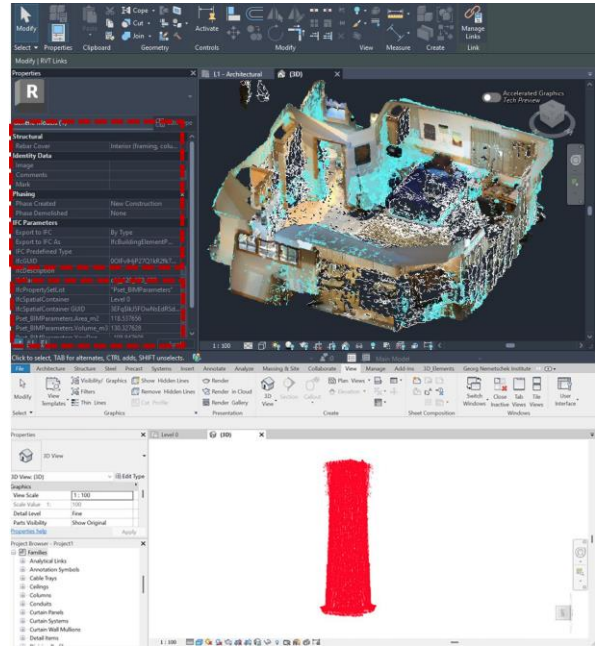
Şekil 4. TUM CMS Indoor Point Cloud için 3B rekonstrüksiyon örnekleri [20, 30].

TUM CMS Indoor Point Cloud'dan 2B Ayrıştırma Seti Türetilmesi ve Denetimi. TUM CMS nokta bulutları ile 2B görüntü ve derinlik çiftleri türetilmiştir [30] (Şekil 2). Bu görüntülerde *SERNet-Former_v2* ile 2B maske haritaları üretilmiş, geri izdüşüm ile 3B'ye taşınarak projeksiyon tutarlılığı referans veri seti üzerinden denetlenmiştir. 2B→3B projeksiyon karşılaştırmasında piksel-bazlı doğruluk 81.9% olarak elde edilmiştir. Bu bulgu, 2B sınıf haritalarının 3B rekonstrüksiyonla yeterli uyum içinde olduğunu ve 3B öğrenmeye gerek kalmadan 3B etiket üretimini mümkün kıldığını göstermektedir. Nicel özet Tablo 4'te, örnek görseller Şekil 2'de sunulmuştur.

3B Rekonstrüksiyon: Protokol, Parametreler ve Sonuçlar. Bu bölümde 3B Rekonstrüksiyon dört ayrı kipte daha değerlendirilmiştir: Tek-atımlı, az-atımlı, seyrek ve video-bazlı. Tüm kiplerde geri izdüşüm formülasyonu aynıdır. Çoklu görüntü kiplerinde yalnızca yüzey sürekliliğini göstermek için öğrenme fonksiyonu kullanılmadan TSDF füzyonu ile birleştirme kullanılabilir [22]. Şekil 3, kipler arası görsel karşılaştırmayı sunar.

Bu çalışmada rekonstrüksiyon için kullanılan sabit parametreler şöyledir: Düzlem eşik mesafesi τ_{Π} 0.02-0.05 m aralığında, güven eşiği τ_q 0.4-0.6 aralığında, k -NN komşuluk sayısı k 8-16 arasında ve normal benzerliği eşiği 10° - 15° aralığındadır. Bu eşik değerleri sahne tipinden bağımsız olarak tek biçimde seçilmiş ve tüm veri setleri boyunca sabit tutulmuştur.

TUM CMS Indoor Point Cloud: Tek-görüntüden 3B rekonstrüksiyon ve semantik zenginleştirme. Bu düzenekte amaç, mevcut TUM CMS nokta bulutlarını



Şekil 5. RGB görüntüden üretilen 3B nesneler, bütün sahne içinde veya ayrı olarak BIM ortamına aktarılabilir.

yalnızca RGB girdilerden üretilen tek-atımlı rekonstrüksiyonlarla besleyip sahne kapsamını artırmaktır (Şekil 4). Akış şu adımlarla yürütülür: Kayıtlı kamera pozlarından temsili RGB kareler seçilir (koridor, ofis, lab tipleri). Her kare I için monoküler modelden görece ölçekli derinlik D ve güven Q_d elde edilir. Eşzamanlı olarak 2B ayrıştırma S , Q_s alınır. D ve K ile geri izdüşüm yapılarak sınıf ve güvenle işaretli nokta bulutu $\mathcal{P}_j$ üretilir. Düşük güvenli pikseller ($Q_s < \tau_q$) ve sınıf-dışı artıklar temizlenir. Komşuluk tabanlı normal kestirimi ve RANSAC ile duvar-zemin-tavan düzlemleri ayrıştırılır, açıklık maskelemesi 3B'ye taşınır. Az-atımlı, seyrek, video kiplerinde, aynı sahneye ait çoklu pozlar için öğrenmesiz birleştirme uygulanır. Önce *ICP* ile çerçeve bulutları hizalanır, ölçek belirsizliği, benzerlik dönüşümü içinde optimize edilir, ardından yalnız yüzey sürekliliğini göstermek üzere TSDF füzyonu ile deliklerin kapanması sağlanır. Gerekğinde *Delaunay/Poisson* rekonstrüksiyonu ile nokta bulutu oluşturulur. Kenar koruyan düzeltmeler ve küçük boşluk doldurmalarla yüzey kalitesi iyileştirilir. Şekil 4, farklı sahne tiplerinden örnek çıktıları göstermektedir. Kapı, pencere açıklıklarının korunması ve duvar düzlemlerinin sürekliliği korunmuştur.

Sonuçlar. Değerlendirme elde edilen temsili nokta bulutlarının referans veri setinde bulunan hedef noktalara göre kayıt sonrası hata, *ICP-RMSE*, ve son olarak nihai geometrik sapma, *3D RMSE*, ölçülerek yapılmış, sonuçlar (3B) *RMSE* ve *mIoU* olarak raporlanmıştır. ScanNet_v2 üzerinde 3B doğruluk 78.8% mertebesindedir. Bu oran, tek-görüntü koşulunda tutarlı bir geometrik kalite göstermektedir. TUM CMS dataseti için 81.9% oranı, düzlemsel nokta bazında bire bir örtüşme durumu optimize edilerek elde

edilmiştir. Kıyas sonuçları Tablo 3 ve 4'te, kip görselleri Şekil 3 ve 4'te verilmiştir.

Nesne Bazında Ayırıştırma. Sınıf işaretli 3B bulut, tohumlama → bölge büyütme → düzlem bağlama sırası ile nesne örneklerine ayırıştırılmıştır. Her nesne için konum (kütle merkezi μ), yönelim ve topolojik ilişkiler (örneğin “kapı $\subset$ duvar”) ayrıca kaydedilmiştir [17]. Nesne düzeyi değerlendirme genel 3D RMSE ile raporlanmış, nesne örnekleri ve BIM görünümü Şekil 1 ve Şekil 5'te sunulmuştur.

4.4. BIM'e Aktarım

Nesne bazında ayırıştırma sonrasında üretilen örnekler $\{O_j\}$ BIM'e aktarımın girdisini oluşturur. Bu aşamada deneyler, hem duvar kalınlığı, yükseklik, açıklık boyutları, konum ve yönelim gibi geometrik parametrelerin güvenilir biçimde üretilebilmesi hem de “kapı $\subset$ duvar”, “pencere $\subset$ dış duvar”, “zemin $\leftrightarrow$ duvar” gibi topolojik ilişkilerin hatasız kurulabilmesi üzerine kurgulanmıştır [17]. Tüm sahnelerde dünya koordinat çerçevesi (x, y, z) sağ el kuralına göre tanımlanmış, birim metre, açı birimi derece olarak kullanılmıştır. İşlem boyunca 3B veriler ve BIM çıktı alanları bu birimlerde tutarlı kalacak şekilde denetlenmiştir.

Eşleme kuralları deneysel olarak sabitlenmiştir: RANSAC'tan gelen büyük düzlemler normal yönlerine göre ve segmentasyon ile atanan renk paletine göre “duvar/zemin/tavan” sınıflarına, silindirik/dikdörtgen kesitli taşıyıcı kümeler “kolon” sınıfına, 2B'den taşınmış ve 3B'de doğrulanmış boşluk kümeleri “kapı/pencere” sınıflarına eşlenir. Duvar kalınlığı, karşılıklı ve eş-normal iki duvar yüzey kümesinin en yakın nokta istatistiğiyle çıkarılır. Eşiklendirme ile gürültü bastırılır ve 2–5 cm tolerans içinde düzlemler “çift” kabul edilir. Zemin ve tavan yüzeyleri için diklik, koşutluk denetimi, yerçekimi yönü varsayımı altında uygulanır. Böylece eğik planlı sahnelerde bile “yatay” kabulü kontrollü biçimde yapılır. Açıklıkların konum ve boyutları, 2B maskeden gelen sınırların 3B'ye taşınıp ilgili duvar düzlemine dik izdüşümü ile belirlenir. Tutarlılık saptandığında açıklık, “aday” olarak işaretlenir.

Parametreleştirme sonrasında her nesneye benzersiz bir kimlik (ID), sınıf etiketi, geometri özeti (w, h, t), kütle merkezi μ , yönelim (yerel eksenler) ve “Door-123 $\subset$ Wall-A” gibi ilişkiler atanır. Bu alanlar bir BIM sözlüğü üstünden şemalaştırılır ve ifcJSON biçiminde dışa aktarılır. Aynı nesneler görsel doğrulama için OBJ olarak da üretilir. ifcJSON tarafında sınıf karşılıkları IfcWall, IfcSlab, IfcCeiling, IfcDoor, IfcWindow olarak düzenlenmiş, parametrik alanlar “Profile, Thickness, Height, Area, OpeningWidth, OpeningHeight, LocalPlacement” gibi öğelerle beslenmiştir. Dışa aktarım sırasında yerel-küresel dönüşümler, eksen uzlaşısı (Z ↑), ve birim dönüşümleri otomatik testlerle doğrulanır.

Koşullar bozulduğunda izlenen istisnai işlemler de raporlanır. Cam ve ayna yüzeylerinde derinlik tahmininin zayıfladığı durumlarda açıklık sınırları 3B'de salınım gösterebilir. Tek-atımdan kaynaklanan eksik veri varsa, düzlemler RANSAC artıklarına göre yeniden kestirilir ve düzlemlerin birleşimi ilişkilendirilir. Böylece BIM'de düzlem tutarlılığı korunur. Elde edilen BIM-uyumlu çıktılar, sınıf karşılıklarına göre tasnif edilen nesneleri, gerek sahne içinde gerekse ayrı ayrı olarak uyumlu yazılım ortamlarında alan, hacim, ve ölçüleri başta olmak üzere tanımlı sayısal veriler eşliğinde sunmaya yardımcı olur (Şekil 5).

Çalışma zamanı açısından aktarım adımı hat boyunca düşük maliyetlidir. 512×512 piksel çözünürlük ve tek A100-40GB GPU düzeninde semantik segmentasyon ve derinlik tahmini, girdiye bağlı olarak 1 ile 3 saniye arasında, geriye kalan 3B rekonstrüksiyon, nesne ayırıştırma ve ifcJSON yazımı yaklaşık 1 ile 2 saniye arasında sürmektedir. Özetle, nesne sayısına bağlı olarak aktarımın sahne başına yaklaşık 1–6 s aralığında kaldığı gözlenmiştir. Bellek tüketimi nesne sayısına lineer bağlıdır ve 6–8 GB VRAM bütçesi içinde kalır. Tüm dışa aktarımlar için deterministik kimlik üretimi, yani sahne–zaman damgası + sınıf + artan sayaç, kullanılmış, böylece farklı çalıştırmalarda BIM grafinin izlenebilirliği sağlanmıştır.

Sonuç olarak, deneyler geometri (ICP-RMSE, 3D RMSE) ve topoloji (ilişki tutarlılığı) açısından önerilen aktarımın sahne çeşitliliğinde kararlı davrandığını göstermektedir. Düzlemsel öğelerde parametre sapmaları düşük, açıklıklarda ise 2B sınır kalitesi ve derinlik gürültüsüne bağlı orta düzey sapmalar gözlenmiştir. Bu durum, sınır-duyarlı 2B bileşen açıkken belirgin biçimde azalır. Elde edilen BIM-uyumlu veriler, yönetim sistemlerine bağlı Brick/Project Haystack gibi semantik modeller ve MPC gibi kontrol tasarımı için zon, alan, hacim gibi gereksinimlerine girdi sağlayacak niteliktedir (Şekil 5).

V. BULGULAR VE TARTIŞMA

Bu bölüm, önerilen görüntü→2B→3B→BIM hattının deneylerden elde edilen sayısal çıktılarının ve uygulamaya dönük yorumlarının bütüncül bir değerlendirmesini sunar. Değerlendirme kurgusu, Bölüm 4'te tanımlanan protokollerle eşleniktir ve ölçüt seti sade tutulmuştur: 2B için *mIoU*, derinlik için *RMSE*, 3B için *ICP-RMSE* ve *3D-RMSE*. Bulgular, 2B sahne ayırıştırma, tek görüntü derinlik ve 3B rekonstrüksiyon, nesne bazında ayırıştırma ve BIM'e aktarım alt temaları altında ele alınmıştır. Sayısal özetler Tablo 2–4 ve ilgili Şekil 2–5 ile desteklenmiştir.

Deneyler, yalnızca tek bir RGB görüntüsü ile başlayan akışın iç mekan ölçeğinde hem semantik hem de geometrik katmanda tekrar edilebilir bir doğruluk düzeyi yakaladığını göstermektedir. 2B düzlemde *SERNet-Former_v2* mimarisi, iç mekana özgü kapı,

pencere, kolon gibi nadir sınıflar üzerinde sınır sürekliliğini güçlendirirken duvar, zemin, tavan gibi geniş sınıflarda alan bütünlüğünü korur. Bu davranış 3B'ye olumlu taşınmıştır. TUM CMS Indoor Point Cloud üzerinden türetilmiş sette 2B→3B projeksiyon tutarlılığı 81.9% *mIoU* ve hata oranı 0.02 *RMSE* mertebesindedir (Şekil 2, Tablo 3 ve Tablo 4). Bu oran, 3B öğrenmeye gerek duymaksızın etiket üretimini pratikte mümkün kılar. ScanNet v2 üzerinde tek-görüntü kipinde 3B doğruluk 78.8% düzeyindedir (Tablo 4). Seyrek-atım ve video kiplerine geçildiğinde yüzey sürekliliği artmaktadır (Şekil 3).

2B sonuçlar, iç mekan semantiğini taşıyan sınıflarda net sınıf ayrımı ve sınır tutarlılığı sunar. Literatürde sık görülen, kapı, pencere gibi açıklıkların duvarlarla birleşiminde maske taşması sorunu, modelin güven haritası ile düşük güvenli piksellerin filtrelenmesi sayesinde belirgin biçimde azalmaktadır (Şekil 2). Böylece duvar–zemin–tavan gibi büyük sınıflarda sınıf içi süreklilik korunurken, açıklıklarda kenar bütünlüğü güçlenir (Şekil 2). Bu çalışmada sunum ekonomisi gözetilerek sadece *mIoU* raporlanmıştır (Tablo 2 ve Tablo 4).

Derinlik modülü, tek atım koşulunda görece ölçekli bir derinlik haritası (*D*) sağlar. Standart iç mekan sahnelerinde *RMSE* seviyeleri, 3B aşamada uygulanan düzlem/normal kısıtları sayesinde sönümlenir. Deneyler, cam yüzeyleri ve parlama kaynaklı sapmaların 3B'ye aktarılırken düzlem çıkarımı ve açıklık doğrulamaları ile kısmen telafi edildiğini göstermektedir (Şekil 2–4). Rekonstrüksiyon kiplerinin kıyaslamasında tek-atım, yüzey sürekliliği en kısıtlı senaryo olmakla birlikte, *ICP* kaydı sonrası global hizalama kabul edilebilir düzeydedir. Seyrek-atım ve video-temelli rekonstrüksiyon kiplerinde çoklu görüş çeşitliliği, boşluk doldurma, kenar sürekliliği ve gürültü bastırma açısından belirgin iyileşme sağlamıştır (Şekil 3). Bu tablo, ek sensörler kullanılmadığında dahi görüntü zenginliğinin 3B kalitesini doğrudan etkilediğini teyit eder.

3B nesne ayrıştırma adımı, duvar/zemin/tavan gibi düzlemsel küme sınıflarının renk bileşenleri ile ayrılmasında kullanılan 2B sahne ayrımı yöntemi ile desteklenmiştir. Her nesne için raporlanan ve min-OBb tabanlı (*w*, *h*, *t*), kütle merkezi μ , ve topolojik ilişkileri kapsayan parametrik özet BIM aktarımı için gerekli alanları üretmeyi mümkün kılar. Bu sayede duvar kalınlığı, açıklık boyutları ve konum gibi mimari parametreler otomatik elde edilir. İlişkisel düzeyde “kapı⊂duvar” gibi kuralların otomatik kontrolü topoloji bütünlüğünü güçlendirir (Şekil 5).

Deneyler, önerilen eşleme kurallarıyla (düzlemler → duvar/zemin/tavan, boşluklar → kapı/pencere, kesitli kümeler → kolon) parametrik alanların (kalınlık, yükseklik, alan, hacim, açıklık ebatları, konum, yönelim) tutarlı biçimde üretilebileceğini

göstermektedir (Şekil 5). *ifcJSON* ve *OBJ* biçimlerine aktarım, görsel ve sayısal doğrulamayı kolaylaştırır. Yerel–küresel dönüşüm ve birim/eksen uzlaşısı (*Z* ↑) gibi ayrıntılar denetlenir. Elde edilen veri, üst katmanda semantik bilgi grafi, örneğin Brick/Haystack, ve MPC odaklı kontrol tasarımlarına programatik bir köprü kurmayı ve zon bölümleri, alan, hacim ve açıklık oranlarını aynı veri hattından türeterek güncellemeyi olanaklı kılmaktadır (Şekil 5). Bu sayede kurumsal ölçekte, farklı binalar ve zonlar arasında düşük marjinal maliyetle kıyas ve izleme olanağı doğacaktır.

Çalışma zamanı ve ölçeklenebilirlik. Tekli RGB girdilerinin 512×512 piksel ebatlarındaki çözünürlüğe ölçeklendiği durumlarda tek A100-40GB GPU üzerinde azami 6 s/örnek sonucu elde edilmiştir. Bu sonuç, saha çalışmalarında anlık geri bildirim döngülerini destekler niteliktedir. Yapay sınır ağı olarak 245 milyon öğrenilebilir parametre içeren SERNet-Former_v2 kullanılmış olup, farklı girdi ebatları ve projeler için ölçeklenebilir mimariler oluşturulabilir. Tek bir fotoğraftan kısa süre içinde 3B rekonstrüksiyon ve BIM-uyumlu parametrik çıktı üretmek, hem tasarım iterasyonlarını hızlandırır hem de enerji ve konfor odaklı kontrol modellerine hızlı girdi sağlar [2]. Bellek tüketiminin 6–8 GB aralığında kalması, orta sınıf GPU'larda dahi saha kullanılabilirliğini artırır.

Sınırlılıklar ve gelecek çalışmalar. Uygulanan ayrıştırma yöntemleri, cam, ayna yüzeyleri ve parlak malzemeler üzerinde derinlik sapmalarına duyarlıdır. Hareket bulanıklığı ve düşük aydınlatma koşullarında hem 2B maskeler hem de derinlik kalitesi düşmektedir. Çok dar koridorlarda duvar çiftleme hataları karşılaşılan sorunlardandır. Sağlamlaştırma için eş-normal çift yüzey eşlemesi ve histogram temelli kalınlık kestirimi gelecek çalışması olarak önerilmektedir.

Çalışmada elde edilen parametrik IFC/ifcJSON temsilleri nesne düzeyinde geometri ve temel özneliliklerin kaydıyla sınırlandırılmıştır. Önerilen çerçeve, Brick, Project Haystack veya Model Predictive Control (MPC) gibi üst seviye semantik/enerji yönetim sistemlerine doğrudan katkı verebilecek nitelikte olmakla beraber, bu entegrasyon gelecek çalışmaların öncelikli adımı olarak planlanmıştır. Benzer biçimde, önerilen hattın deneyimlenen veri setleri dışında farklı aydınlatma veya ortam koşulları içeren gerçek dünya senaryolarında canlı (*real-time*) veri akışı sağlayabilmesi de gelecek çalışmalar arasında planlanmaktadır.

Son olarak, bu çalışma tek görüntü kipini temel alır. Çoklu ve video kiplerinde kalite yükselse de, veri lojistiği ve koşul kontrolü pratik uygulamada ek maliyet doğurmaktadır. Kapı–duvar, pencere–dış duvar, zemin–duvar birleşimi gibi topolojik ilişkiler

seti bu çalışmada temel kurallarla sınırlandırılmış olup, gelecekte bu ilişkilerin zenginleştirilmesi ve öğrenen otomatik doğrulama mekanizmalarıyla güçlendirilmesi planlanmaktadır.

Akıllı kampüs mekanları açısından yorum. Daha önce bahsedilen TUM2TWIN ve TUM CMS gibi kampüs temelli dijital ikiz projeleri kapsamında, elde edilen örnekler ve çıktılar var olan veri setlerine, 2D Scene Parsing Dataset gibi ek veri setleri oluşturularak, katkı sağlamıştır [30]. Ek olarak, geliştirilen yöntemler bahsi geçen yapıları çevrenin belgelenmesinde ve mekan içerisinde yer alan nesneler hakkında daha fazla sayısal verinin oluşturulmasında katkıda bulunmuştur. Öte yandan, herhangi bir bağlama bağlı kalmaksızın, geliştirilen yöntemler ve elde edilen bulgular, semantik veri okuma gereksinimi ile ölçeklenebilir uygulama hedefi arasında ciddi bir gereksinime hitap etmektedir. Görüntü tabanlı veri işleme ve iş akışının BIM'e doğrudan bağlı hale getirilmesi, akıllı kampüs ve hatta akıllı şehir altyapılarındaki karar verme süreçlerini parametrik verilere dayalı kılar [31]. Zon topolojisi, alan, hacim ve açıklık oranları gibi göstergeler otomatik üretilebilir ve düzenli olarak yenilenebilir (Şekil 5). Bu, manuel etiketleme ve ölçüm yükünü azaltarak denetim tasarımlarında yinelenbilirlik ve standartlaştırma sağlar.

2B'de sınır duyarlılığı ve nadir sınıf kararlılığı, 3B'de düzlem bütünlüğü, açıklıkların korunması ve kabul edilebilir *mIoU*, *ICP-RMSE* profili, BIM'de ise geometrik, parametrik ve ilişkisel bütünlük elde edilmiştir (Şekil 5). Bu tablo, ek donanım gerektirmeden iç mekanların semantik-geometrik karakterizasyonunu yapmak ve bunu makine-okunur BIM katmanına taşımak için yeterli mühendislik uygunluğuna ulaştığını göstermektedir.

VI. SONUÇ

Bu çalışma, tek bir RGB görüntüsünden başlayarak 2B semantik segmentasyonu ile monoküler derinlikten 3B rekonstrüksiyon oluşturarak nesne düzeyi ayrıştırma ve BIM'e aktarım adımlarını pratik bir hat içinde birleştiren bütünlük bir yaklaşım sunmuştur. Hat, 2B'de *mIoU*, derinlikte *RMSE* ve 3B'de *ICP-RMSE* ve *3D-RMSE* ile yalın fakat yeterli bir çerçevede sınanmıştır. Sonuçlar, 2B'de sınır duyarlılığının ve nadir sınıf kararlılığının 3B'ye olumlu taşındığını, açıklıkların duvar yüzeylerinde bozulmadan korunabildiğini göstermektedir. ScanNet_v2 üzerinde tek-görüntü kipinde 3B doğruluk 78.8%, TUM CMS Indoor Point Cloud veri setinde ise piksel-bazlı 2B→3B projeksiyon tutarlılığı 81.9% oranında gözlenmiştir.

Nesne düzeyi ayrıştırmadan üretilen parametrik özet ve ilişkiler *ifcJSON* biçimlerine aktarılmış, böylece BIM-uyumlu çıktılar elde edilmiştir. Uygulama açısından, 512×512 piksel ebatlı veri girişi ile azami 6 s/girdi sonucu ve 6–8 GB VRAM bütçesi, erken aşama tasarım

ve denetim senaryolarında anlık geri bildirimi erişilebilir kılar. Bu sonuçlar, Brick/Project Haystack benzeri semantik katmanlara ve MPC gibi üst düzey denetim tasarımlarına programatik girdi sağlamaya, zon topolojisi, alan, hacim, açıklık oranları gibi göstergelerin elle etiketleme olmadan üretilebilmesine olanak sağlayacaktır.

Bu çalışmada bahsedilen sınırlılıklar, yüzeylerde görülen derinlik sapmaları, düşük ışık koşullarında maskelerin zayıflaması ve dar koridorlarda karşılaşılan duvar çiftleme hatalarıdır. Bunlara karşı pratikte güven bantları, eş-normal yüzey eşleştirmesi ve düzlem tabanlı yeniden kestirim yaklaşımları uygulanmış ve hata etkileri kısmen azaltılmıştır. Gelecekte, daha geniş sahne çeşitliliği içeren çoklu-görüntü ve uzun video kipleri ile genelleme analizlerinin yapılması, canlı veri akışı sağlayan modellerin geliştirilmesi, BIM nesne parametrisasyonu ile otomatik kalite güvencesi ve kurumsal ölçekte ölçeklenmiş saha denemeleri planlanmaktadır. Mevcut çalışmada elde edilen IFC/ifcJSON tabanlı parametrik nesneler temel geometrik ve öznitelik kayıtlarıyla sınırlandırılmış olup, üst seviye semantik veya enerji yönetim sistemleriyle doğrudan entegrasyon gelecek çalışmaların öncelikli konusu olacaktır. Ayrıca, kurulan topolojik ilişkiler seti temel kurallarla elde edilmiş olup, ileride bu ilişkilerin zenginleştirilmesi ve otomatik doğrulama mekanizmalarıyla desteklenmesi öngörülmektedir.

Sonuç olarak, önerilen hat, akıllı kampüs projelerini hedefleyen dijital ikiz projeleri bağlamına katkı sunmanın yanında, iç mekanlarda 2B→3B→BIM dönüşümünü genel mühendislik ölçütleriyle savunulabilir, tekrarlanabilir ve pratiğe dönük bir düzeye taşımakta, aynı zamanda ileri düzey akıllı ve semantik sistemlere geçiş için sağlam bir temel oluşturmaktadır.

KAYNAKLAR

- [1] Erişen, S. (2023). An empirical study of the technoparks in Turkey in investigating the challenges and potential of designing intelligent spaces. *Sustainability*, 15(13), 10150.
- [2] Erişen, S. (2023). A systematic approach to optimizing energy-efficient automated systems with learning models for thermal comfort control in indoor spaces. *Buildings*, 13(7), 1824.
- [3] Balaji, B., Bhattacharya, A., Fierro, G., Gao, J., Gluck, J., Hong, D., Johansen, A., Koh, J., Ploennigs, J., Agarwal, Y., Culler, D., Gupta, R. K., Agarwal, R., vd. (2016). Brick: Towards a unified metadata schema for buildings. *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments (BuildSys)*.
- [4] Project Haystack. (2019). *Project Haystack: Open source initiative for semantic tagging of data.*

- [5] Tang, P., Huber, D., Akinci, B., Lipman, R., ve Lytle, A. (2010). Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review. *Automation in Construction*, 19(7), 829–843.
- [6] Volk, R., Stengel, J., ve Schultmann, F. (2014). Building information modeling (BIM) for existing buildings—Literature review and future needs. *Automation in Construction*, 38, 109–127.
- [7] Dai, A., Chang, A. X., Savva, M., Halber, M., Funkhouser, T., ve Niessner, M. (2017). ScanNet: Richly-annotated 3D reconstructions of indoor scenes. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5828–5839.
- [8] Chang, A. X., Dai, A., Funkhouser, T., Halber, M., Niessner, M., Savva, M., Song, S., Zeng, A., ve Zhang, Y. (2017). Matterport3D: Learning from RGB-D data in indoor environments. *Proceedings of the International Conference on 3D Vision (3DV)*.
- [9] Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alonso, I., Alvarez, J. M., ve Lu, T. (2021). SegFormer: Simple and efficient design for semantic segmentation with transformers. *Advances in Neural Information Processing Systems (NeurIPS)*.
- [10] Cheng, B., Misra, I., Schwing, A., Kirillov, A., ve Girdhar, R. (2022). Masked-attention mask transformer for universal image segmentation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [11] Bhat, S. F., Birkel, R., Wofk, D., Wonka, P., ve Müller, M. (2023). Zoedepth: Zero-shot transfer by combining relative and metric depth. *arXiv preprint arXiv:2302.12288*.
- [12] Yin, W., Zhang, C., Chen, H., Cai, Z., Yu, G., Wang, K., Chen, X., ve Shen, C. (2023). Metric3D: Towards zero-shot metric 3D prediction from a single image. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 9043–9053.
- [13] Wysocki, O., Schwab, B., Biswanath, M. K., Zhang, Q., Zhu, J., Froeh, T., Heeramaaglore, M. ve diğerleri. (2025). TUM2TWIN: Introducing the Large-Scale Multimodal Urban Digital Twin Benchmark Dataset. *arXiv:2505.07396*. <https://tum2t.win/> "TUM2TWIN"
- [14] buildingSMART International. (2023). *Industry foundation classes (IFC) specification (IFC 4.x series)*.
- [15] buildingSMART International. (2021). *ifcJSON: A JSON serialization for IFC*.
- [16] Erişen, S. (2025). Efficient segmentation using attention-fusion modules with dense predictions. *IEEE Access*, 13, 107552–107565.
- [17] Erişen, S., Mehranfar, M., ve Borrmann, A. (2025, Mayıs). A multi-task framework for 2D scene parsing and 3D reconstruction in indoor space documentation. *EG-ICE 2025 Workshop on Intelligent Computing in Engineering*.
- [18] Armeni, I., Sener, O., Zamir, A. R., Jiang, H., Brilakis, I., Fischer, M., ve Savarese, S. (2016). 3D semantic parsing of large-scale indoor spaces. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1534–1543.
- [19] Mehranfar, M., Vega-Torres, M., Braun, A., ve Borrmann, A. (2024). Automated data-driven method for creating digital building models from dense point clouds and images through semantic segmentation and parametric model fitting. *Advanced Engineering Informatics*, 62, 102862.
- [20] Wan, L., Rossa, F., Welfonder, T., Petrova, E., ve Pauwels, P. (2025). Enabling scalable model predictive control design for building HVAC systems using semantic data modelling. *Automation in Construction*, 170, 105929.
- [21] Erişen, S., Mehranfar, M., ve Borrmann, A. (2025). Single image to semantic BIM: Domain-adapted 3D reconstruction and annotations via multi-task deep learning. *Remote Sensing*, 17(16), 2910.
- [22] Curless, B., ve Levoy, M. (1996). A volumetric method for building complex models from range images. *Proceedings of SIGGRAPH*, 303–312.
- [23] Fischler, M. A., ve Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381–395.
- [24] Kazhdan, M., Bolitho, M., ve Hoppe, H. (2006). Poisson surface reconstruction. *Eurographics Symposium on Geometry Processing (SGP)*, 61–70.
- [25] Delaunay, B. (1934). Sur la sphère vide. *Izvestia Akademii Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk*, 7(6), 793–800.
- [26] Besl, P. J., ve McKay, N. D. (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 239–256.
- [27] Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., ve Torralba, A. (2017). Scene parsing through ADE20K dataset. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 633–641.
- [28] Song, S., Lichtenberg, S. P., ve Xiao, J. (2015). SUN RGB-D: A RGB-D scene understanding benchmark suite. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 567–576.

-
- [29] Silberman, N., Hoiem, D., Kohli, P., ve Fergus, R. (2012). Indoor segmentation and support inference from RGBD images. *Proceedings of the European Conference on Computer Vision (ECCV)*, 746–760.
- [30] Erişen, S., Mehranfar, M., ve Borrmann, A. (2025). 2D Scene Parsing Dataset of TUM CMS Indoor Point Cloud. mediaTUM. doi: 10.14459/2025mp1779501
- [31] Erişen, S. (2021). Incremental transformation of spatial intelligence from smart systems to sensorial infrastructures. *Building Research and Information*, 49(1), 113-126.