



ORIGINAL RESEARCH PAPER

2D Floorplan Extraction of Indoor Spaces from Mobile Laser Scanner Point Clouds Using Deep Learning Methods

M. Heidarimozaffar^{*,1}, Z. Dalvand²¹ Department of Civil Engineering, Faculty of Engineering, Bu-Ali Sina University, Hamadan, Iran² Department of Computer Science, Faculty of Engineering, Bu-Ali Sina University, Hamedan, Iran

ABSTRACT

Received: 04 May 2025
 Reviewed: 10 June 2025
 Revised: 02 August 2025
 Accepted: 06 September 2025

KEYWORDS:

Floorplan
 Point Clouds
 Density IMAGE
 Deep Learning

* Corresponding author

✉ m.heidarimozaffar@basu.ac.ir

☎ (+98912) 5185925

Background and Objectives: The extraction of 2D floorplans of building interiors plays a vital role in various domains, including architecture, surveying, building information modeling (BIM), robotics, and virtual reality. Mobile laser scanners capture the geometric structure of indoor environments with millimeter-level accuracy and record the results as point cloud data. Point clouds are a rich source of information for generating 2D floorplans of indoor spaces. However, surface reflection noise, occlusions caused by indoor objects, and the non-uniform density of points pose significant challenges for processing such data. Initially, 2D floorplan extraction relied on classical geometric methods. In recent years, however, deep learning-based approaches have gained increasing attention due to their strong ability to understand complex patterns and their robustness to noise. The main objective of this study is to present an effective framework for extracting 2D floorplans of building interiors from point cloud data using deep learning methods and to compare its performance with that of classical techniques.

Methods: In this study, an effective framework is proposed for extracting 2D floorplans of indoor building spaces from point cloud data, consisting of three sequential steps: data preprocessing, model implementation, and final evaluation. This framework enables a direct comparison between classical methods and deep learning approaches within a unified setting. Point cloud data are inherently discrete and unstructured, making direct processing challenging. In the preprocessing step, point clouds were projected onto a 2D space to generate density images, thereby reducing computational complexity. In the second step, two deep learning models, U-Net and Pix2Pix, as well as the classical Hough Transform algorithm, were implemented, with the density images serving as a common input for all methods. In the third step, the proposed framework was evaluated using publicly available datasets, including FloorNet and Structure3D. The input data were split into training, validation, and test sets, and data augmentation techniques were applied to improve model generalization. The performance of the models was assessed using the Dice Score and Intersection over Union (IoU) metrics.

Findings: Deep learning models demonstrated satisfactory performance on samples without occlusions, achieving accuracy levels above 90%. In particular, the U-Net model achieved a Dice Score of 97% on the Structure3D dataset. However, in samples containing occlusions, the models were unable to fully extract the floorplans. In contrast, the Hough Transform algorithm performed reasonably well in line detection but exhibited limitations in generating coherent and topologically valid outputs suitable for indoor map modeling due to its inability to capture topological structure. Moreover, the trial-and-error process required to tune the algorithm's parameters significantly increased its runtime.

Conclusion: The findings of this study indicate that deep learning methods, when provided with complete data, are capable of accurately and structurally extracting 2D floorplans from point clouds. However, under real-world conditions where occlusion is inevitable, developing models that are robust to incomplete data becomes essential. To address this challenge, future research directions include employing hybrid architectures and incorporating complementary data sources such as RGB images or depth maps. The proposed framework in this study serves as an effective step toward the systematic comparison of 2D floorplan extraction methods and provides a foundation for developing more advanced models suitable for real-world applications.



NUMBER OF REFERENCES

31



NUMBER OF FIGURES

14



NUMBER OF TABLES

2

مقاله پژوهشی

استخراج نقشه دوبعدی فضاهای داخلی ساختمان با روش‌های یادگیری عمیق از ابرنقاط لیزر اسکنرهای همراه

مرتضی حیدری مظفر^{۱*}، زهرا دالوند^۲^۱ گروه مهندسی عمران، دانشکده فنی و مهندسی، دانشگاه بوعلی سینا، همدان، ایران^۲ گروه کامپیوتر، دانشکده مهندسی، دانشگاه بوعلی سینا، همدان، ایران

چکیده

پیشینه و اهداف: استخراج نقشه دوبعدی فضاهای داخلی ساختمان در حوزه‌های مختلف از جمله معماری، نقشه‌برداری، مدل‌سازی اطلاعات ساختمان (BIM)، رباتیک و واقعیت مجازی نقش حیاتی دارد. اسکنرهای لیزری همراه، ساختار هندسی محیط را با دقت میلی‌متری اندازه‌گیری کرده و نتایج حاصل را به صورت داده‌های ابرنقاط ثبت می‌کنند. داده‌های ابرنقاط، منبع غنی اطلاعات برای استخراج نقشه دوبعدی فضاهای داخلی ساختمان هستند. با این حال، عواملی مانند نویز ناشی از بازتاب‌های سطحی، انسداد دید توسط اجسام درون ساختمان و تراکم غیریکنواخت نقاط، پردازش این نوع داده را با چالش مواجه می‌کنند. در ابتدا استخراج نقشه دوبعدی ساختمان، با تکیه بر روش‌های کلاسیک هندسی انجام می‌شد و در سال‌های اخیر، روش‌های مبتنی بر یادگیری عمیق به دلیل توانایی بالا در درک الگوهای پیچیده و مقاومت در برابر نویز به‌طور فزاینده‌ای مورد توجه قرار گرفته‌اند. هدف این پژوهش، ارائه یک چارچوب موثر برای استخراج نقشه دوبعدی فضاهای داخلی ساختمان از داده‌های ابرنقاط با استفاده از روش‌های یادگیری عمیق و مقایسه عملکرد آن با روش‌های کلاسیک است.

روش‌ها: در این پژوهش، یک چارچوب موثر برای استخراج نقشه دوبعدی فضاهای داخلی ساختمان از داده‌های ابرنقاط پیشنهاد شده است که شامل سه گام متوالی پیش‌پردازش داده‌ها، پیاده‌سازی و ارزیابی نهایی مدل‌ها است. این چارچوب امکان مقایسه مستقیم روش‌های کلاسیک و یادگیری عمیق را در یک بستر مشترک فراهم می‌کند. داده‌های ابرنقاط دارای ساختار گسسته و غیرساخت‌یافته هستند و پردازش مستقیم آن‌ها دشوار است. در گام پیش‌پردازش داده‌ها، ابتدا با نگاشت ابرنقاط به فضای دوبعدی، تصاویر تراکم تولید شد تا پیچیدگی محاسباتی کاهش یابد. در گام دوم، دو مدل یادگیری عمیق U-Net و Pix2Pix و الگوریتم کلاسیک تبدیل هاف پیاده‌سازی شد و تصاویر تراکم به‌عنوان ورودی مشترک این روش‌ها استفاده شد. در گام سوم به منظور ارزیابی عملکرد روش پیشنهادی، آزمایش‌ها روی مجموعه داده‌های در دسترس FloorNet و Structure3D انجام شد. داده‌های ورودی به سه مجموعه داده آموزشی، اعتبارسنجی و آزمون تقسیم شده‌اند و به منظور ارتقای تعمیم‌پذیری، تکنیک‌های داده‌افزایی اعمال شده است. ارزیابی عملکرد مدل‌ها با استفاده از معیارهای Dice Score و IoU انجام شد.

یافته‌ها: مدل‌های یادگیری عمیق در نمونه‌های فاقد انسداد، عملکرد قابل قبولی داشته و دقت بیش از ۹۰٪ نشان دادند. به‌ویژه، مدل U-Net در مجموعه داده Structure3D به دقت ۹۷٪ در معیار Dice Score دست یافت. با این حال، مدل‌ها در نمونه‌هایی که شامل عارضه انسداد بودند، نتوانستند نقشه را به طور کامل استخراج کنند. در مقابل، الگوریتم تبدیل هاف در تشخیص خطوط عملکرد قابل قبول داشت اما به دلیل عدم تشخیص ساختار توپولوژیک، در تولید خروجی‌های منسجم و قابل استفاده برای مدل‌سازی نقشه‌های داخلی محدودیت دارد. همچنین آزمون و خطا برای تنظیم مقادیر پارامترها، باعث افزایش قابل توجه زمان اجرای الگوریتم شد.

نتیجه‌گیری: یافته‌های این پژوهش نشان داد که روش‌های یادگیری عمیق، در صورت وجود داده‌های کامل، قادر به استخراج دقیق و ساخت‌یافته نقشه دوبعدی از داده‌های ابرنقاط هستند اما در شرایط واقعی که عارضه انسداد اجتناب‌ناپذیر است، توسعه مدل‌های مقاوم نسبت به داده ناقص ضروری است. برای این منظور بهره‌گیری از معماری‌های ترکیبی، استفاده از منابع اطلاعاتی مکمل نظیر تصاویر RGB یا داده‌های عمق مسیر پژوهش‌های آینده است. چارچوب

تاریخ دریافت: ۱۴ اردیبهشت ۱۴۰۴
تاریخ داوری: ۲۰ خرداد ۱۴۰۴
تاریخ اصلاح: ۱۱ مرداد ۱۴۰۴
تاریخ پذیرش: ۱۵ شهریور ۱۴۰۴

واژگان کلیدی:

نقشه دوبعدی
ابر نقاط
تصویر تراکم
یادگیری عمیق

* نویسنده مسئول

m.heidarimozaffar@basu.ac.ir

۰۹۱۲-۵۱۸۵۹۲۵

پیشنهادی این پژوهش، گامی مؤثر در جهت مقایسه‌ی نظام‌مند روش‌های استخراج نقشه دوبعدی و ایجاد بستر توسعه برای مدل‌های پیشرفته‌تر در کاربردهای واقعی تلقی می‌شود.

مقدمه

استخراج خودکار نقشه دوبعدی محیط‌های داخلی ساختمان‌ها از داده‌های خام اسکنر لیزری، نقطه عطف بزرگی در تحقیقات نقشه‌برداری است. مجموعه داده‌های فضایی گسسته که به عنوان ابرنقاط شناخته می‌شوند، اغلب پایه‌گذار برنامه‌های تصمیم‌گیری هستند [۱]. در ابرنقاط، هر نقطه دارای مختصات X ، Y و Z است. نکته حائز اهمیت نحوه اندازه‌گیری، پردازش و تولید اطلاعات قابل اطمینان از مجموعه داده‌های جمع‌آوری شده است [۲]. داده‌های ابرنقاط به صورت مستقیم به وسیله پوششگرهای لیزری اندازه‌گیری می‌شوند و یا به صورت غیر مستقیم با روش‌های فتوگرامتری تولید می‌شوند [۳]. برخلاف روش‌های فتوگرامتری، پوششگرهای لیزری به دلیل دقت محدوده لیزری در سطح میلی‌متر و جهت‌گیری دقیق پلتفرم حسگر (LIDAR) که با سیستم تعیین موقعیت و جهت‌گیری یکپارچه (position and orientation system) پشتیبانی می‌شود، بسیار دقیق است [۴].

با توسعه فناوری‌های برداشت سه‌بعدی نظیر اسکنرهای لیزری زمینی (Terrestrial Laser Scanning)، سامانه‌های برداشت متحرک (Mobile Mapping system)، حسگرهای عمق‌سنج RGB-D و سیستم‌های مکان‌یابی و نقشه‌سازی هم‌زمان (Simultaneous Localization and Mapping)، امکان برداشت سریع، دقیق و جامع از محیط‌های داخلی ساختمان به صورت داده‌های ابرنقاط سه‌بعدی فراهم شده است [۵]. استخراج نقشه دوبعدی از این داده‌ها نقش کلیدی در کاربردهایی نظیر نقشه‌برداری، مدل‌سازی شهری و صنعتی [۶]، مدل‌سازی اطلاعات ساختمان (Building Information Model) [۷]، مدیریت بحران و امداد و نجات [۸]، ناوبری رباتیک و تولید واقعیت مجازی [۹]، مستندسازی و حفظ میراث فرهنگی [۱۰]، فیلم‌ها و بازی‌های رایانه‌ای دارد. چالش اصلی، تبدیل داده‌های ابرنقاط به ساختارهای منسجم و قابل تفسیر مانند دیوارها، درها و اتاق‌های بسته است. تفسیر و تبدیل این داده‌ها به نقشه، به دلیل وجود عواملی مانند نویز و نقاط خارج از محدوده، انسداد (occlusion)، تراکم ناکافی و تغییرپذیری وضوح، عدم پیوستگی داده و استانداردسازی همچنان یک مسئله باز و دشوار است [۱۱]. تحقق این هدف مستلزم بهره‌گیری از روش‌هایی دقیق برای تشخیص مرزهای ساختاری، جبران نواحی دچار انسداد و حفظ یکپارچگی و مقیاس هندسی محیط است.

قطعه‌بندی معنایی (Semantic Segmentation) یکی از مهم‌ترین مراحل پردازش برای تبدیل داده‌های ابرنقاط به نقشه‌های قابل استفاده محسوب می‌شود [۱۲]. قطعه‌بندی در ابرنقاط، فرآیندی است که طی آن مجموعه‌ای از نقاط سه‌بعدی به زیرمجموعه‌هایی با ویژگی‌های مشترک تقسیم می‌شوند، به گونه‌ای که هر قطعه نشان‌دهنده یک ناحیه یا ساختار همگن در صحنه باشد. ویژگی‌های مشترک می‌توانند شامل

خصوصیات هندسی (مانند موقعیت، نرمال سطح، انحنا)، ویژگی‌های رادیومتریکی (مانند شدت بازتاب یا رنگ) یا خصوصیات معنایی (مانند تعلق به یک شیء خاص) باشند. هدف از قطعه‌بندی، ساده‌سازی، ساختاردهی و تسهیل درک و تفسیر داده‌های حجیم و غیرساخت‌یافته ابرنقاط است، به طوری که بتوان از آن‌ها برای تحلیل‌های بعدی مانند استخراج نقشه‌های دوبعدی و سه‌بعدی استفاده کرد.

روش‌های ارائه‌شده برای این مسئله را می‌توان به طور کلی در دو رویکرد اصلی طبقه‌بندی کرد: (۱) رویکردهای مبتنی بر قواعد گرامری و (۲) رویکردهای مبتنی بر یادگیری. دسته نخست عمدتاً شامل روش‌های کلاسیک است که پیش از توسعه گسترده الگوریتم‌های یادگیری ماشین مورد استفاده قرار می‌گرفتند. این روش‌ها با تکیه بر مجموعه‌ای از قواعد صریح و الگوهای از پیش تعیین‌شده سعی در استخراج نقشه دوبعدی دارند. با این حال، این دسته از روش‌ها معمولاً به دلیل پیچیدگی در تنظیم دقیق پارامترها، پیش‌پردازش سطح پایین متناسب با نوع داده، و طراحی قواعد منطبق با سبک‌های معماری مختلف، دارای سطح اتوماسیون پایین و نیازمند مداخلات انسانی قابل توجهی هستند [۱۳]. افزون بر آن، وابستگی شدید این الگوریتم‌ها به فرمت نقشه‌های ورودی، انعطاف‌پذیری آن‌ها را در مواجهه با داده‌های متنوع کاهش می‌دهد. در مطالعه‌ای با استفاده از پیش‌فرض‌های هندسی استخراج شده از ابرنقاط محیط‌های داخلی، الگوریتمی مبتنی بر بازسازی از پایین به بالا ارائه داد [۱۴]. همچنین، در پژوهشی دیگر با بهره‌گیری از الگوها و قواعد از پیش تعریف شده، اقدام به تقسیم‌بندی فضایی محیط‌های بسته نمود [۱۵]. این دسته از روش‌ها، هرچند در محیط‌هایی با ساختار هندسی منظم عملکرد مناسبی از خود نشان می‌دهند، اما در فضاهایی با هندسه‌های پیچیده، نامنظم یا طراحی‌های غیرسنتی با محدودیت‌های جدی مواجه‌اند.

رویکردهای مبتنی بر یادگیری به دو دسته روش‌های یادگیری ماشین و یادگیری عمیق تقسیم می‌شود. در روش‌های یادگیری ماشین، الگوریتم‌هایی نظیر خوشه‌بندی آماری، بهینه‌سازی عددی و تکنیک‌های استخراج ویژگی کلاسیک به منظور شناسایی الگوهای ساختاری موجود در داده به کار گرفته می‌شوند [۱۶]. ساختارهای هندسی اتاق‌ها با ترکیبی از الگوریتم‌های یادگیری ماشین نظیر خوشه‌بندی مبتنی بر میانگین، پوشش‌دهی با استفاده از پوسته‌های غیرمقعر و نیز الگوریتم اجماع نمونه‌برداری تصادفی (RANSAC) استخراج شده‌اند [۱۷].

روش‌های یادگیری عمیق با بهره‌گیری از معماری‌های پیشرفته شبکه‌های عصبی، توانایی مدل‌سازی ارتباطات غیرخطی و پیچیده بین داده‌های ورودی و ساختارهای نقشه دوبعدی را دارند. از معماری ترنسفورمر جهت شناسایی گوشه‌ها و استنتاج روابط اتصال میان آن‌ها استفاده شده تا گراف مسطح نقشه به دست آید [۱۸]. با این حال، عدم

ضروری در فرآیندهای پردازش ابرنقاط در حوزه معماری و نقشه‌برداری دیجیتال به‌شمار می‌رود.

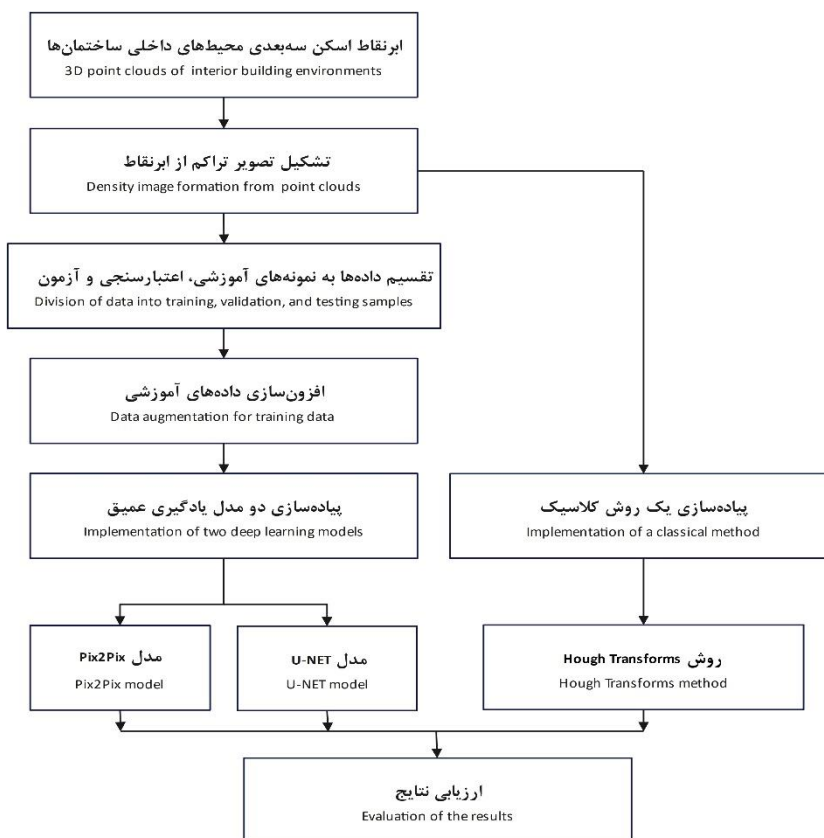
نوآوری این پژوهش، ارائه‌ی یک چارچوب مؤثر برای استخراج نقشه‌ی دوبعدی محیط‌های داخلی ساختمان‌ها از ابرنقاط اسکن لیزری سه‌بعدی است. در شکل (۱) روندنمای فرآیند استخراج نقشه در روش پیشنهادی نشان داده شده است. مراحل انجام کار به صورت زیر است:

- در گام اول، تصویر تراکم مربوط به ابرنقاط ورودی تشکیل می‌شود.
- در گام دوم، استخراج نقشه دوبعدی از تصویر تراکم با یک روش کلاسیک و سپس با استفاده از رویکردهای یادگیری عمیق پیاده‌سازی می‌شود.
- در گام سوم، به ارزیابی نتایج مدل‌های پیاده‌سازی شده خواهیم پرداخت.

اغلب پژوهش‌های پیشین تنها بر یکی از دو رویکرد کلاسیک یا یادگیری عمیق تمرکز داشته‌اند، در این پژوهش یک چارچوب جامع برای مقایسه، تحلیل و ارزیابی متقابل این دو دسته روش ارائه شده است. هدف اصلی این پژوهش، استخراج نقشه دوبعدی فضاهای داخلی ساختمان از داده‌های ابرنقاط اسکن لیزری، از طریق پیاده‌سازی و مقایسه دو رویکرد مستقل شامل روش‌های کلاسیک و یادگیری عمیق است. در ادامه، ابتدا مجموعه داده‌های مورد استفاده معرفی می‌شوند؛ سپس به ارائه جزئیات روش پیشنهادی پرداخته و در نهایت، نتایج حاصل از اجرای این روش‌ها گزارش و مورد بحث قرار می‌گیرند.

شناسایی کامل برخی گوشه‌ها منجر به حذف ناخواسته برخی یال‌ها و در نتیجه، ناقص‌ماندن اتاق‌های حاصل گردید. در پژوهشی دیگر، نمایه‌ای مبتنی بر تقسیم‌بندی پلان به مجموعه‌ای از جعبه‌ها معرفی شد که با استفاده از رگرسیون جعبه‌ای، اطلاعات محلی شکل هر اتاق به‌طور مستقیم استخراج می‌شد [۱۹]. اگرچه این رویکرد قابلیت نمایش صریح هندسه اتاق‌ها را فراهم می‌کرد، اما در نمایش ساختارهای نامنظم با محدودیت مواجه بود. با تعریف دو سطح پرسش و مدل‌سازی نقشه دوبعدی به‌عنوان مجموعه‌ای از چندضلعی‌ها با گوشه‌های مرتب‌شده و طول متغیر، گامی مؤثر در بهبود کیفیت بازسازی برداشت [۲۰]. با این حال، این روش همچنان با چالش‌هایی مانند تولید توالی‌های نامنظم ناشی از چندضلعی‌های اولیه تصادفی و همچنین خطا در بازسازی دقیق چندضلعی‌ها به دلیل حذف یا سوگیری در شناسایی گوشه‌های منفرد روبه‌رو بود.

با وجود پیشرفت‌های اخیر در روش‌های یادگیری عمیق و الگوریتم‌های مبتنی بر شبکه‌های عصبی، اغلب این روش‌ها برای داده‌های ساخت‌یافته یا محیط‌های بیرونی طراحی شده‌اند و عملکرد آن‌ها در محیط‌های بسته با ساختار پیچیده داخلی، به‌ویژه در شرایط برداشت واقعی، با چالش مواجه است. از این رو، توسعه‌ی روشی دقیق و قابل تعمیم برای استخراج خودکار نقشه‌های دوبعدی محیط‌های داخلی از ابرنقاط سه‌بعدی، با تمرکز بر شناسایی مرز دیوارها و بازسازی توپولوژی اتاق‌ها، نیازی



شکل ۱: روش پیشنهادی در این تحقیق برای استخراج نقشه دوبعدی محیط‌های داخلی ساختمان‌ها از ابرنقاط سه‌بعدی

Fig. 1: The proposed method in this study for extracting 2D floorplans of indoor building environments from 3D point clouds

روش تحقیق

مجموعه داده‌ها

جهت بررسی عملکرد روش پیشنهادی و شناخت بهتر قابلیت‌ها و محدودیت‌های آن، از دو دیتاست ابرنقاط استفاده شده که در دسترس عموم قرار دارند. این مجموعه‌ها شامل داده‌های مصنوعی و واقعی از محیط‌های داخلی هستند و هر کدام ویژگی‌ها و چالش‌های خاص خود را دارند:

- مجموعه داده Structured3D، یک مجموعه داده مصنوعی از محیط‌های داخلی است که شامل ۳۵۰۰ برداشت و نقشه‌های مختلف طبقه می‌باشد. داده‌ها دارای برچسب‌گذاری ساختاری سه‌بعدی هستند که برای شبیه‌سازی محیط‌های واقعی طراحی شده‌اند. این مجموعه داده شامل صحنه‌های ساده و پیچیده است [۲۱]. به منظور آموزش شبکه عصبی پیشنهادی، تعداد کل نمونه‌ها به صورت ۳۰۰۰ اسکن برای آموزش، ۲۵۰ اسکن برای اعتبارسنجی و ۲۵۰ اسکن برای ارزیابی تقسیم‌بندی شده است.

- مجموعه داده FloorNet، یک مجموعه داده واقعی از محیط‌های داخلی است که شامل ۱۵۵ برداشت و نقشه‌های مختلف طبقه می‌باشد. داده‌ها دارای برچسب‌گذاری ساختاری سه‌بعدی هستند که به صورت دستی تهیه شده است [۲۲]. به منظور آموزش شبکه عصبی، تعداد کل نمونه‌ها به صورت ۱۳۵ اسکن برای آموزش، ۲۰ اسکن برای اعتبارسنجی و ۲۰ اسکن برای ارزیابی تقسیم‌بندی شده است.

پیش پردازش داده‌ها

تشکیل تصویر تراکم از داده‌های ابرنقاط

به منظور کاهش پیچیدگی محاسباتی از نظر زمان اجرا و میزان مصرف حافظه، ابرنقاط اولیه به تصویر تراکم (Density Image) تبدیل شده و فرآیند استخراج نقشه در فضای دوبعدی تصویر انجام می‌گیرد. تصویر تراکم اطلاعات توزیع نقاط را در صفحه افقی X-Y نمایش دهد و مبنایی

برای مراحل بعدی تحلیل داده است. برای این منظور در ابتدا، مختصات نقاط در صفحه X-Y نرمال‌سازی می‌شوند تا در محدوده تصویر دوبعدی قرار گیرند. این کار با محاسبه کمینه و بیشینه مختصات نقاط و مقیاس‌بندی آن‌ها به اندازه تصویر انجام می‌شود. این فرآیند به صورت رابطه (۱) و (۲) تعریف می‌شود:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} * W \quad (1)$$

$$y' = \frac{y - y_{min}}{y_{max} - y_{min}} * H \quad (2)$$

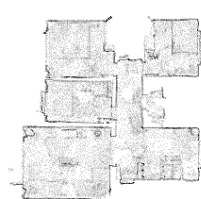
که در آن H و W به ترتیب ارتفاع و عرض تصویر، x' و y' مختصات نرمال شده نقاط در تصویر هستند.

در این روش برای هر پیکسل تصویر دوبعدی، تعداد نقاطی که در آن محدوده قرار می‌گیرند شمارش می‌شود. این شمارش براساس مختصات x و y نقاط انجام شده و تمام نقاط دارای مقدار Z متفاوت که در یک موقعیت (x, y) قرار دارند، جمع می‌شوند. نتیجه این پردازش یک نقشه چگالی است که در آن مقدار هر پیکسل نشان دهنده تراکم نقاط در آن مکان خاص است. نمونه تصاویر تراکم دیتاست‌های ورودی به همراه ماسک واقعیت زمینی آن‌ها در شکل (۲) نشان داده شده است.

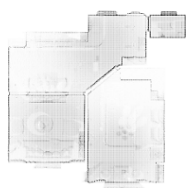
الگوریتم‌های پیشنهادی برای استخراج نقشه دوبعدی

روش کلاسیک Hough Transform

یکی از تکنیک‌های پرکاربرد در پردازش تصویر و بینایی ماشین تبدیل هاف است که برای شناسایی و تحلیل اشکال هندسی مانند خطوط، دایره‌ها و سایر منحنی‌ها در تصاویر مورد استفاده قرار می‌گیرد. در این روش، به‌ویژه برای تشخیص خطوط در نقشه‌های طبقات ساختمان، برای هر نقطه‌ای در فضای تصویر، یک مجموعه از پارامترها در فضای ویژگی ترسیم می‌شود که معرف خطوط قابل ترسیم از آن نقطه هستند [۲۳]. به طور خاص، تبدیل هاف به صورت زیر عمل می‌کند:



الف



ب

شکل ۲: نمونه تصویر تراکم دیتاست ورودی به همراه نقشه دوبعدی مربوطه، الف - دیتاست FloorNet، ب - دیتاست Structured3D

Fig. 2: Sample density images from the input dataset along with their corresponding 2D maps — (a) FloorNet dataset, (b) Structured3D dataset

مدل دارای دو بخش اصلی رمزگذار (Encoder) و رمزگشا (Decoder) است. بخش رمزگذار شامل چندین لایه کانولوشن و لایه‌های کاهش ابعاد (Pooling Layer) است که وظیفه استخراج ویژگی‌ها و کاهش ابعاد تصویر ورودی را به عهده دارد. در این بخش، ویژگی‌های مختلف تصویر به تدریج استخراج می‌شوند و ابعاد تصویر کاهش می‌یابد، بنابراین مدل می‌تواند الگوهای پیچیده در داده‌ها را یاد بگیرد. سپس، بخش رمزگشا با استفاده از لایه‌های افزایش ابعاد (Upsampling) و کانولوشن، ابعاد تصویر را به اندازه اولیه بازمی‌گرداند. همچنین، U-Net از اتصالات مستقیم (Skip Connections) بین بخش‌های رمزگذار و رمزگشا استفاده می‌کند که به مدل اجازه می‌دهد جزئیات دقیق‌تر و اطلاعات محلی را حفظ کند. این ویژگی باعث می‌شود که U-Net در قطعه‌بندی تصاویر، عملکرد بهتری داشته باشد و بتواند لبه‌ها و بافت‌های اصلی را به خوبی شناسایی کند. شکل (۳) معماری شبکه عصبی U-Net را نمایش می‌دهد. توابع فعال‌ساز در شبکه‌های عصبی، پارامتری مهم است. توابع فعال‌ساز، برای شناسایی روابط پیچیده بین داده‌ها استفاده می‌شوند و خروجی غیرخطی ارائه می‌دهند. بدون توابع فعال‌ساز، شبکه‌های عصبی تنها مدل‌های رگرسیون خطی خواهند بود که قادر به نمایش چندجمله‌ای‌های درجه یک هستند [۲۶]. در شبکه‌های عصبی عمیق، هدف اصلی لایه Dropout، کاهش بیش‌برازش (Overfitting) است. بیش‌برازش، زمانی اتفاق می‌افتد که مدل به داده‌های آموزشی به شدت وابسته می‌شود و پارامترها را طوری تنظیم می‌کند که به خوبی روی داده‌های آموزشی عملکرد داشته باشد، اما در داده‌های جدید عملکرد ضعیفی از خود نشان می‌دهد. لایه Dropout در طول آموزش، با خاموش کردن نورون‌های مختلف، هر بار یک زیرمدل متفاوت از شبکه ایجاد می‌کند. این وضعیت شبیه به مجموعه‌ای از مدل‌ها با وزن‌های مختلف است که خود باعث بهبود عمومی الگوی یادگیری می‌شود [۲۷].

نمایش فضای پارامتر: هر خط در فضای تصویر می‌تواند با استفاده از معادله‌ی خط به شکل $y = mx + b$ یا فرم معادله عمومی‌تر $Ax + By + C = 0$ توصیف شود. در این روش، معمولاً از فرم قطبی $r = x \cos(\theta) + y \sin(\theta)$ استفاده می‌شود. که در آن r فاصله از مبدأ θ زاویه‌ای است که خط نسبت به محور x تشکیل می‌دهد [۲۴].

تبدیل نقاط تصویر به فضای پارامتر: برای هر نقطه‌ی موجود در تصویر، مقدار r و θ محاسبه شده و در فضای پارامتر ثبت می‌شود. بدین ترتیب، هر نقطه‌ی موجود در تصویر به یک منحنی در فضای پارامتر تبدیل می‌شود.

تجمع و شناسایی خطوط: در فضای پارامتر، نقاطی که به یکدیگر نزدیک هستند، نشان‌دهنده وجود یک خط در تصویر اصلی هستند. با استفاده از یک ناحیه‌ی تجمع (accumulator)، تعداد دفعاتی که هر مجموعه از مقادیر r و θ برمی‌گردد، شمارش می‌شود. مقادیر با بیشترین تعداد تجمع، معرف خطوطی هستند که در تصویر اصلی وجود دارند.

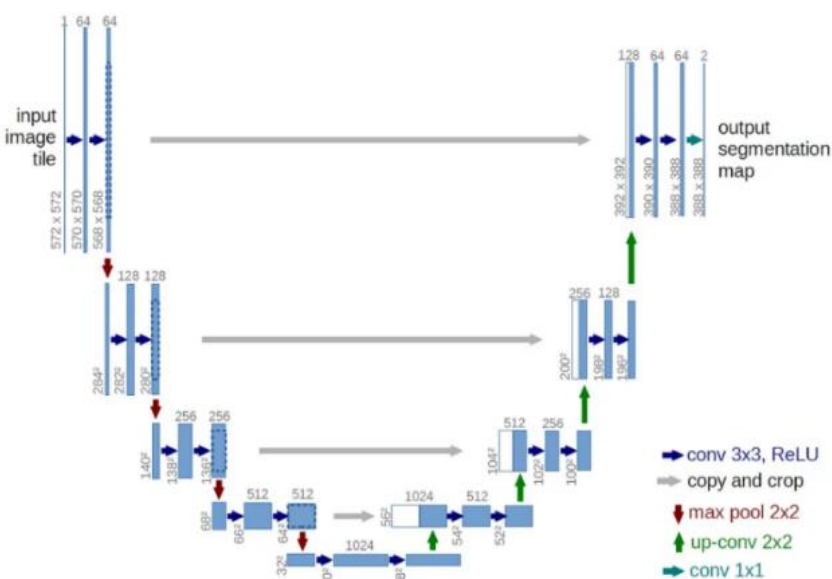
شناسایی و بازسازی خطوط: پس از شناسایی نقاط حداکثر در فضای پارامتر، می‌توان با استفاده از این مقادیر، خطوط شناسایی‌شده را به وضوح در تصویر ورودی بازسازی کرد.

این روش به دلیل قابلیت زیاد آن در شناسایی خطوط به‌ویژه در شناسایی ساختارهای مستطیلی و دیوارها در نقشه‌های طبقات، بسیار مورد توجه قرار گرفته است. با این حال، ضروری است که پارامترهای الگوریتم به درستی تنظیم شوند تا دقت و کارایی آن افزایش یابد.

مدل‌های یادگیری عمیق

مدل U-NET

شبکه عصبی U-Net یک معماری قدرتمند برای پردازش تصاویر است که به‌طور خاص برای وظایف قطعه‌بندی طراحی شده است [۲۵]. این



شکل ۳: معماری مدل U-Net

Fig. 3: Architecture of the U-Net Model

مدل Pix2Pix

- زیان بازسازی: برای اطمینان از اینکه خروجی تولیدشده شباهت نزدیکی به تصویر هدف دارد، از زیان L1 بین تصویر تولیدشده و تصویر واقعی استفاده می‌شود. رابطه (۴) نحوه محاسبه زیان بازسازی را نشان می‌دهد.

$$\begin{aligned} Loss_{L1G} &= E_{x,y} [|y \\ &- G(x)|_1] \end{aligned} \quad (4)$$

ترکیب این دو تابع زیان، تابع کلی آموزش مدل را تشکیل می‌دهد. رابطه (۵) نحوه محاسبه زیان کل را نشان می‌دهد.

$$G^* = \arg \min \max L_{CGAN}(G, D) + \lambda Loss_{L1G} \quad (5)$$

که در آن λ یک ضریب وزنی برای کنترل تأثیر زیان بازسازی است و به طور معمول برابر ۱۰۰ در نظر گرفته می‌شود.

معیارهای ارزیابی

به منظور ارزیابی عملکرد الگوریتم‌های قطعه‌بندی پیشنهادی، از معیارهای استاندارد شباهت‌سنجی شامل دو شاخص Dice score و IOU استفاده شد. این شاخص‌ها، هم‌پوشانی بین تصاویر Ground Truth و خروجی الگوریتم قطعه‌بندی پیشنهادی را اندازه‌گیری می‌کنند. تعریف این دو شاخص در رابطه (۶) و (۷) آمده است:

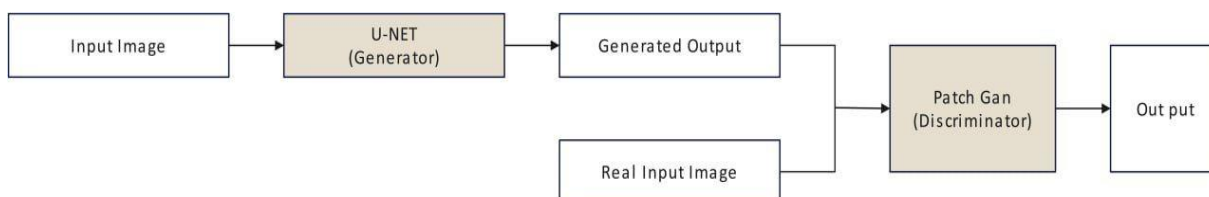
$$Dice\ Score(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (6)$$

$$Mean\ IOU = \frac{1}{n} \sum_{k=1}^n \frac{|A_k \cap B_k|}{|A_k \cup B_k|} \quad (7)$$

که در اینجا A و B به ترتیب تصویر Ground Truth و تصویر پیش‌بینی شده توسط مدل است. n تعداد کلاس‌ها و برابر ۲ است. هرچه عدد به دست آمده به مقدار یک نزدیکتر باشد، نشان دهنده هم‌پوشانی بیشتر و عملکرد مطلوب قطعه‌بندی است.

نتایج و بحث

در مرحله پیش‌پردازش، مختصات نقاط در صفحه X-Y نرمال‌سازی شده و یک تصویر تراکم دوبعدی (Density Image) با ابعاد 256×256 تشکیل گردید. این تصویر میزان تراکم نقاط در هر موقعیت را نشان داده و مبنایی برای مراحل بعدی استخراج نقشه محسوب می‌شود. شکل (۵) فرآیند کلی پیش‌پردازش ابرنقاط را به صورت شماتیک نمایش می‌دهد.



شکل ۴: مدل Pix2Pix

Fig. 4: The Pix2Pix Model

مدل Pix2Pix روش ترجمه تصویر به تصویر بر پایه شبکه‌های GAN شرطی است که قابلیت یادگیری نگاشت‌های پیچیده بین تصویر ورودی و تصویر هدف را دارد. برخلاف مدل‌های GAN کلاسیک که به صورت عمومی از داده‌های بدون نظارت استفاده می‌کنند، در Pix2Pix از جفت تصاویر ورودی و هدف برای یادگیری نگاشت‌ها استفاده می‌شود [۲۸]. شکل (۴) مدل Pix2Pix را به صورت شماتیک نمایش می‌دهد.

مدل Pix2Pix از دو شبکه اصلی تشکیل شده است:

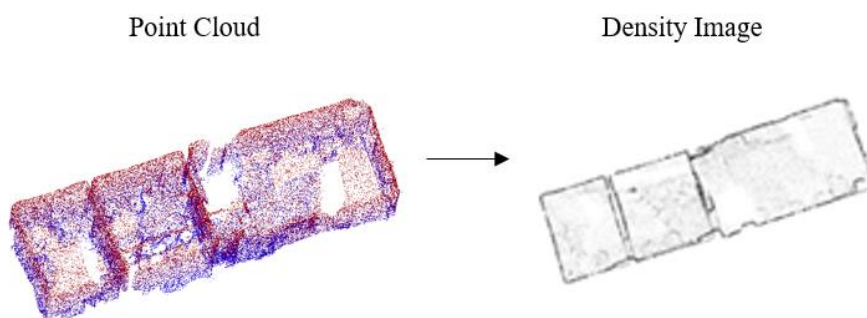
- Generator: شبکه مولد با معماری U-Net طراحی شده است که از یک مسیر نزولی برای استخراج ویژگی‌ها و یک مسیر صعودی برای بازسازی تصویر استفاده می‌کند. در این معماری، اتصالات میان‌بر بین لایه‌های متناظر در دو مسیر، باعث حفظ جزئیات محلی و مرزهای دقیق در خروجی تصویر می‌شود. این ویژگی در ترجمه تصویر به تصویر، به ویژه برای بازسازی دقیق جزئیات و ساختارهای پیچیده بسیار مهم است. مدل مولد تلاش می‌کند که با یادگیری روابط پیچیده بین ورودی و خروجی، تصویری تولید کند که از نظر محتوا و ساختار مشابه تصویر هدف باشد. در اینجا از شبکه‌های کانولوشنی (CNN) و دریافت ویژگی‌های چندمقیاسی برای بهینه‌سازی و افزایش دقت استفاده می‌شود.

- Discriminator: شبکه تمایزدهنده با ساختار PatchGAN طراحی شده است که تصویر تولیدی را به صورت قطعه‌های $N \times N$ بررسی می‌کند. این مدل به جای تشخیص کل تصویر، تمرکز خود را روی صحت محلی تصاویر می‌گذارد، یعنی تمایز را در مقیاس کوچک و محلی بین تصاویر واقعی و تولیدی یاد می‌گیرد. این ویژگی کمک می‌کند تا جزئیات دقیق‌تر و ویژگی‌های ظریف‌تر در تصویر تولیدی با دقت بیشتر حفظ شود.

در مدل Pix2Pix، دو تابع زیان اصلی برای شبکه مولد و تمایزدهنده استفاده می‌شود:

- Conditional GAN Loss: این تابع زیان سعی می‌کند شبکه مولد را به تولید تصاویری ترغیب کند که شبکه تمایزدهنده نتواند آن‌ها را از تصاویر واقعی تشخیص دهد. رابطه (۳) نحوه محاسبه تابع زیان Conditional GAN را نشان می‌دهد.

$$\begin{aligned} L_{CGAN}(G, D) &= E_{x,y} [\log D(x, y)] \\ &+ E_x [\log(1 \\ &- D(x, G(x)))] \end{aligned} \quad (3)$$



شکل ۵: تصویر شماتیک فرآیند کلی پیش‌پردازش ابرنقاط
Fig. 5: Schematic Illustration of the Overall Point Cloud Processing Workflow

نتایج روش Hough Transform

در این پژوهش ابتدا از روش Hough Transform به منظور استخراج نقشه دوبعدی ساختمان استفاده کردیم. پیاده‌سازی این روش با استفاده از تابع cv2.HoughLinesP در کتابخانه OpenCV انجام گرفته است. این تابع قابلیت شناسایی خطوط پیوسته در تصاویر را دارد و با استفاده از پارامترهای مختلف، تنظیمات لازم برای دقت و کارایی شناسایی خطوط را فراهم می‌آورد. پارامترهای کلیدی این تابع به شرح زیر تنظیم شده است:

edges: تصویر لبه‌ها است که معمولاً به واسطه مراحل پیش‌پردازش تصویر، مانند فیلتر Canny، حاصل می‌شود [۲۹]. این تصویر لبه‌ها، شامل اطلاعاتی در مورد موقعیت لبه‌های موجود در تصویر است که شناسایی خطوط بر اساس آن انجام می‌پذیرد.

ρ : وضوح دقت در فضای قطبی را تعیین می‌کند و نشان‌دهنده فاصله بین خطوط در واحد پیکسل است. یک مقدار کوچک‌تر برای ρ دقت بیشتری را در شناسایی خطوط باریک‌تر فراهم می‌آورد. در اینجا از ۱ پیکسل استفاده شده است که تعادل مناسبی بین دقت و زمان پردازش برقرار می‌کند.

θ : تعیین‌کننده وضوح زاویه در فضای قطبی است و به درجه‌های مربوط به هر خط اختصاص داده می‌شود. با اختصاص این مقدار به $\text{np.pi} / 180$ رادیان، به ازای هر درجه کامل، یک مقدار متناوب محاسبه می‌شود که امکان شناسایی خطوط با زوایای دقیق‌تری را فراهم می‌آورد.

Threshold: مقدار آستانه تعیین‌کننده تعداد حداقل رأی‌های لازم برای شناسایی یک خط است. مقدار این پارامتر برابر ۳۰ تنظیم شد. به بیان ساده‌تر، اگر حداقل ۳۰ رأی برای یک خط جمع‌آوری شده باشد، به عنوان یک خط معتبر شناسایی خواهد شد. افزایش این مقدار می‌تواند منجر به کاهش تعداد خطوط شناسایی‌شده و در عوض افزایش دقت آن‌ها شود.

minLineLength: نشان‌دهنده حداقل طول یک خط است که باید برای شناسایی معتبر باشد. مقدار این پارامتر برابر ۶ تنظیم شد. هر خطی که کمتر از ۶ پیکسل طول داشته باشد، نادیده گرفته می‌شود. این پارامتر

تنظیمات پیاده‌سازی

در این پژوهش، به منظور استخراج نقشه دوبعدی فضاهای داخلی ساختمان از داده‌های ابرنقاط، یک روش کلاسیک و دو رویکرد مبتنی بر یادگیری عمیق شامل مدل U-NET و Pix2Pix پیاده‌سازی و ارزیابی شدند. تمام مدل‌ها در چارچوب pytorch پیاده‌سازی شد. تمام آزمایش‌ها بر روی یک GPU Nvidia 4060 با ۸ گیگابایت حافظه انجام شد. در ادامه، تنظیمات پیاده‌سازی و آموزش هر یک از مدل‌ها به تفصیل ارائه می‌شود:

تنظیمات مدل U-NET: مدل U-NET با ساختار کلاسیک شامل مسیرهای رمزگذار و رمزگشا متقارن پیاده‌سازی شد. تابع Adam عنوان بهینه‌ساز و نرخ یادگیری برابر $1e-4$ تنظیم شد. از تابع هزینه BCEWithLogitsLoss استفاده شد تا تعادل بین دقت و فراخوانی در بخش‌های مختلف نقشه حفظ شود. تعداد دوره‌های آموزشی برابر ۲۰۰ در نظر گرفته شد. برای جلوگیری از بیش‌برازش مدل، نرخ Dropout برابر ۰/۲ در لایه‌های میانی اعمال شد. تابع فعال‌ساز RELU در تمام لایه‌های کانولوشن به کار گرفته شد، در حالی که از تابع sigmoid در خروجی جهت تولید نقشه استفاده شد. ورودی‌ها به اندازه 256×256 پیکسل تغییر اندازه داده شدند و مقادیر پیکسل‌ها به بازه $[0,1]$ نرمال‌سازی شد.

تنظیمات مدل Pix2Pix: مدل Pix2Pix به عنوان یک شبکه GAN شرطی برای ترجمه تصویر به تصویر پیاده‌سازی شد. ساختار generator بر پایه U-NET با عمق ۸ لایه و discriminator از نوع PatchGAN با پنجره‌های ۷۰ در ۷۰ طراحی گردید. برای بهینه‌سازی دو شبکه از تابع بهینه‌ساز Adamw با نرخ یادگیری $5e-5$ و پارامتر $\beta = 0.5$ استفاده شد. تابع هزینه generator ترکیبی از Binary Cross-Entropy برای تابع زیان خصمانه و L1 Loss با ضریب $\lambda = 10$ برای حفظ شباهت ساختاری بین تصویر ورودی و خروجی بود. تعداد دوره‌های آموزشی برابر ۲۰۰ تعیین شد. به منظور بهبود هم‌گرایی مدل، نرخ یادگیری در هر ۱۰ دوره به میزان ۱۰٪ کاهش داده شد. اندازه تصاویر ورودی 256×256 پیکسل است و مقادیر پیکسل‌ها به بازه $[0,1]$ نرمال‌سازی شد.

نتایج عددی و کیفی ارائه شده نشان‌دهنده دقت مناسب مدل در تفکیک نواحی ساختاری مختلف نقشه می‌باشد. جدول (۱) و (۲) نتایج ارزیابی مدل پیشنهادی را بر روی داده‌های test نشان می‌دهد.

جدول ۱: نتایج ارزیابی مدل پیشنهادی براساس شاخص Dice Score
Table 1: Evaluation Results of the Proposed Model Based on the Dice Score Metric

Dataset	Dice Score	
	U-NET	Pix2Pix
FloorNet	0.925	0.90
Structured3D	0.971	0.957

جدول ۲: نتایج ارزیابی مدل پیشنهادی براساس شاخص Mean IOU
Table 2: Evaluation Results of the Proposed Model Based on the IOU Metric

Dataset	Mean IOU	
	U-NET	Pix2Pix
FloorNet	0.90	0.87
Structured3D	0.94	0.91

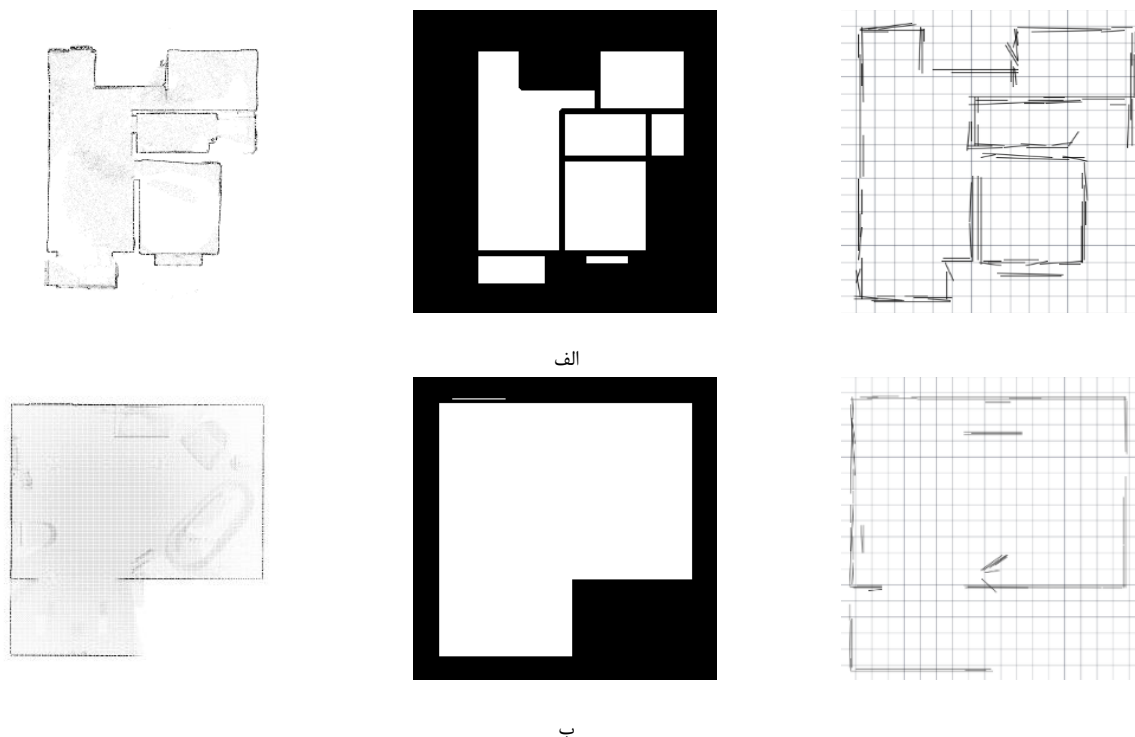
در خصوص ارزیابی مدل‌های یادگیر، ماتریس ابهام اطلاعات جزئی‌تر از نتایج را به ما نشان می‌دهد [۳۰]. شکل ۷ و ۸ confusion matrix مدل پیشنهادی بر روی مجموعه داده‌های ورودی را نشان می‌دهد.

کمک می‌کند تا خطوط کوتاه که ممکن است ناشی از نویز باشند از بین بروند.

maxLineGap: حداکثر فاصله‌ای را که دو نقطه درست برای تشکیل یک خط می‌توانند داشته باشند را، تعیین می‌کند. مقدار این پارامتر برابر ۵ تنظیم شد. اگر فاصله بین دو نقطه به قدری باشد که بیشتر از ۵ پیکسل باشد، این نقاط به عنوان بخشی از یک خط واحد در نظر گرفته نخواهند شد. این ویژگی به تسهیل شناسایی خطوط پیوسته کمک می‌کند. استفاده بهینه از این پارامترها به شناسایی دقیق و مؤثر خطوط در تصویر کمک کرده و موجب بهبود کیفیت استخراج اطلاعات هندسی از نقشه‌های طبقه شده است. شکل (۶) نمونه نتایج این روش بر روی دیتاست‌های مورد مطالعه را نمایش می‌دهد. از اصلی‌ترین معایب تبدیل عمومی هاف، عدم محدودیت فضای پارامتر است که منجر به هزینه پردازشی و زمانی بالا در این روش می‌گردد.

نتایج روش یادگیری عمیق

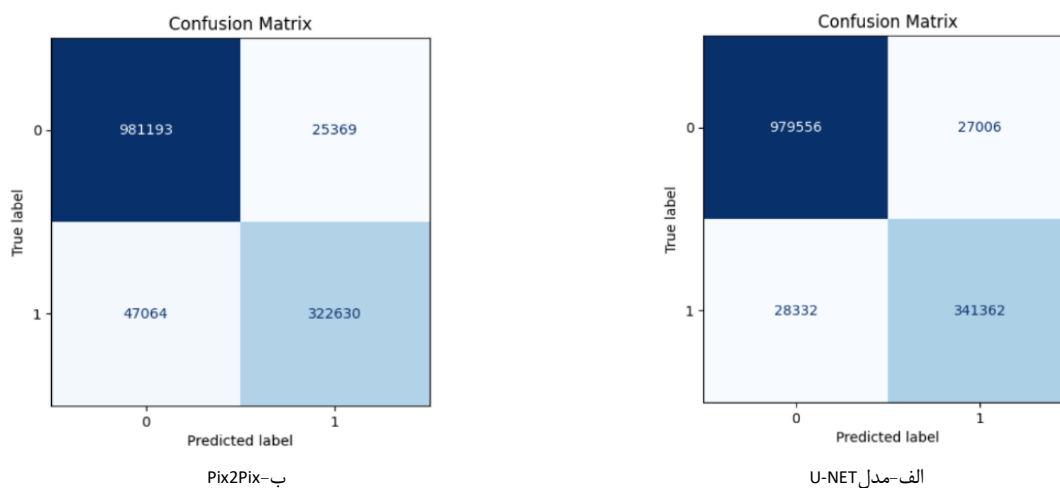
روش پیشنهادی در این پژوهش مبتنی بر بازسازی پیکسلی و بر پایه یادگیری عمیق است. از این‌رو، ارزیابی ما بر مبنای شاخص‌های پیکسلی مانند Dice Score، IOU و همچنین تحلیل کیفی بصری انجام شده است.



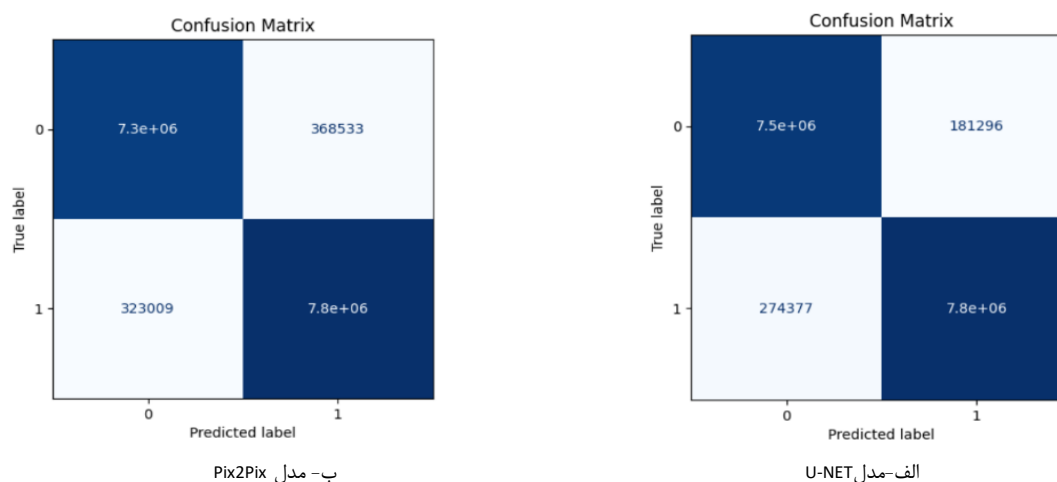
شکل ۶: نمونه نتایج تبدیل هاف (سمت چپ تصویر تراکم، وسط واقعیت زمینی نمونه‌ها و سمت راست خروجی پیش‌بینی شده است)،

الف-نتایج دیتاست FloorNet ب- نتایج دیتاست Structured3D

Fig. 6: Sample results of the Hough transform (left: density image, middle: ground truth samples, right: predicted output);
(a) Results on the FloorNet dataset (b) Results on the Structured3D dataset



شکل ۷: ماتریس ابهام مدل‌های پیشنهادی بر روی دیتاست FloorNet
Fig. 7: Confusion Matrix of the Proposed Models on the FloorNet Dataset



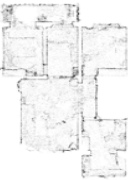
شکل ۸: ماتریس ابهام مدل‌های پیشنهادی بر روی دیتاست Structure3D
Fig. 8: Confusion Matrix of the Proposed Models on the Structured3D Dataset

محدودیت‌ها

اگرچه روش پیشنهادی مبتنی بر یادگیری عمیق با استفاده از مدل‌های U-Net و Pix2Pix عملکرد مطلوبی در استخراج نقشه دوبعدی داشته و به میانگین دقت بیش از ۹۰ درصد بر اساس معیار Dice Score دست یافته است، اما همچنان با برخی محدودیت‌ها مواجه است. نخست، نقشه بازسازی شده به پروجکشن تصاویر دو بعدی از ابرنقاط سه بعدی وابسته است. کیفیت ابرنقاط ورودی و روش پروجکشن تصویر می‌تواند دقت نتایج را به طرز معناداری تحت تأثیر قرار دهد. دوم، مدل‌های یادگیری عمیق پیشنهادی، تصویر ورودی را به صورت پیکسل به پیکسل قطعه‌بندی می‌کند و ممکن است این فرآیند منجر به ترسیم خطوط کج و نامطلوب شود. این نواقص در دقت مرزها، قابلیت کاربردی نتایج را برای مصارف نقشه‌برداری کاهش می‌دهد. سوم، به‌طور خاص هر دو مدل در نمونه‌هایی که داده ابرنقطه اولیه دارای انسداد (Occlusion) شدید بوده است، عملکرد ضعیف‌تری نشان داده‌اند. عارضه انسداد در نتیجه محدودیت دید حسگر یا نویزهای محیطی رخ می‌دهد.

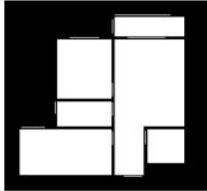
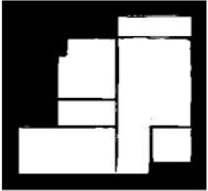
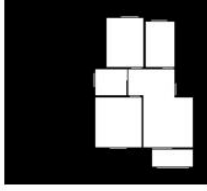
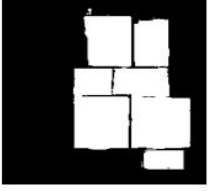
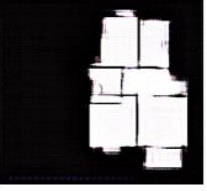
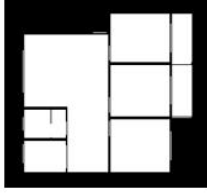
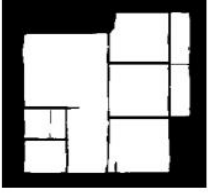
نتایج کیفی حاصل از اجرای روش پیشنهادی، در ادامه ارائه شده است. به منظور ارزیابی بصری عملکرد این مدل‌ها، نمونه‌هایی از نتایج پیش‌بینی شده به همراه تصاویر ورودی و نقشه‌های مرجع (Ground Truth) نمایش داده می‌شوند. این تصاویر، نمایانگر توانایی مدل‌ها در بازسازی مرزهای دقیق فضاهای داخلی ساختمان و تطابق با ساختارهای واقعی هستند. شکل (۹) و (۱۰) نمونه نتایج روش پیشنهادی روی دیتاست FloorNet و Structure3d را نشان می‌دهد.

علاوه بر تحلیل‌های جداگانه فوق، شکل‌های (۱۱) و (۱۲) خروجی سه روش تبدیل هاف، U-Net و Pix2Pix را به صورت مقایسه‌ای روی یک نمونه از هر دیتاست نمایش می‌دهند. مقایسه نتایج نشان می‌دهد که تبدیل هاف صرفاً خطوط منفرد و گسسته تولید می‌کند و قادر به بازسازی ساختارهای توپولوژیک نیست، در حالی که مدل‌های یادگیری عمیق مرزهای پیوسته و چندضلعی‌های بسته را با دقت بازسازی کرده‌اند.

INPUT	Ground Truth	Output (U-Net)	Output (Pix2Pix)
Sample 1			
			
Sample 2			
			
Sample 3			
			

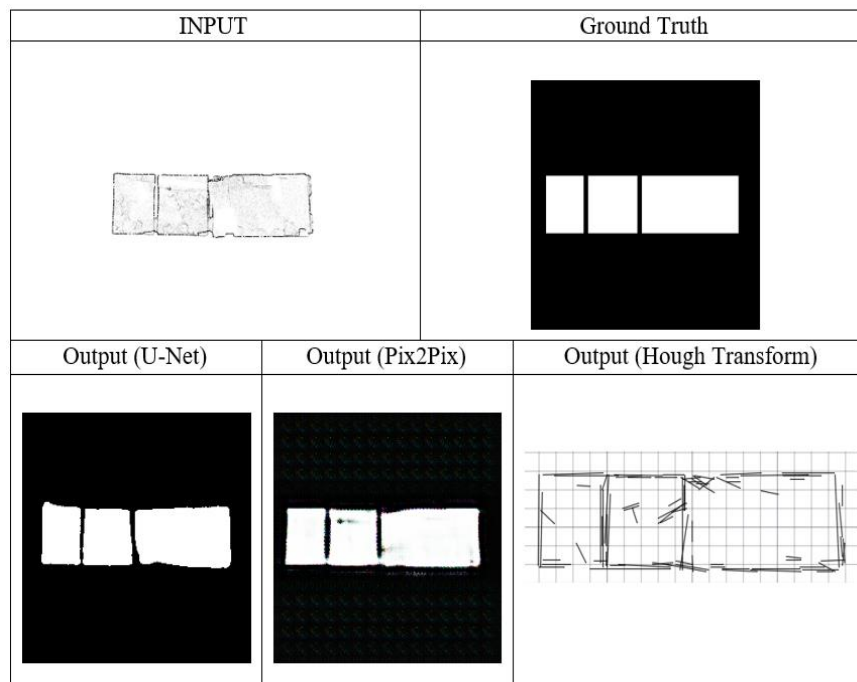
شکل ۹: نتایج روش پیشنهادی بر روی سه نمونه تست از دیتاست FloorNet

Fig. 9: Results of the proposed method on three test samples from the FloorNet dataset

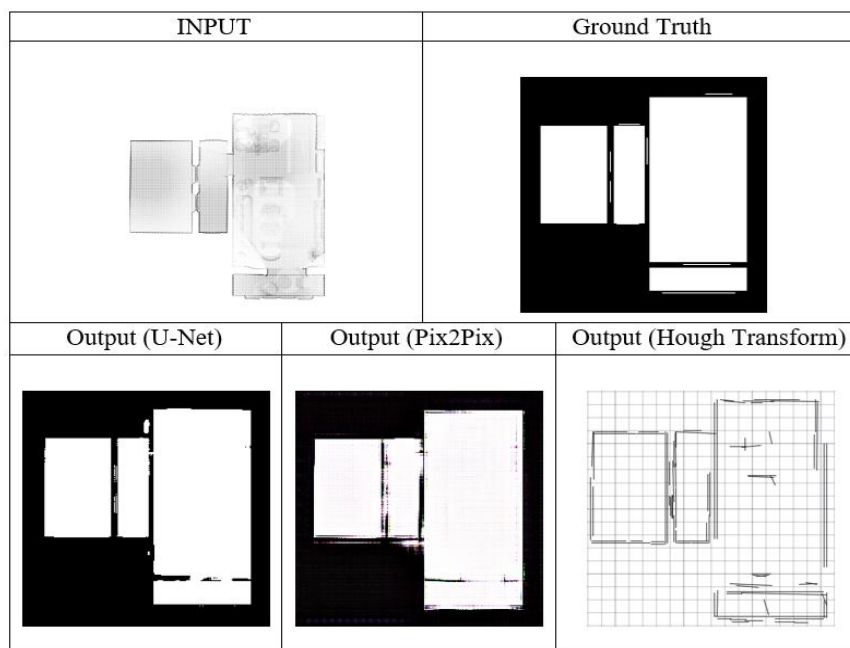
INPUT	Ground Truth	Output (U-Net)	Output (Pix2Pix)
Sample 1			
			
Sample 2			
			
Sample 3			
			

شکل ۱۰: نتایج روش پیشنهادی بر روی سه نمونه تست از دیتاست Structure3D

Fig. 10: Results of the proposed method on three test samples from the Structure3D dataset



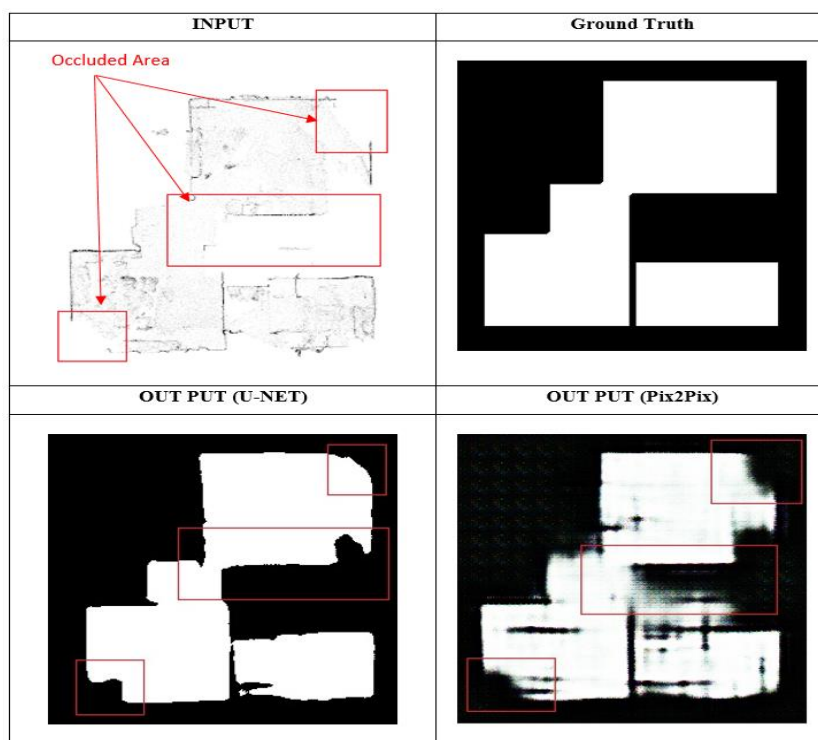
شکل ۱۱: مقایسه نتایج روش کلاسیک و مدل‌های یادگیری عمیق روی دیتاست FloorNet
 Fig. 11: Comparison of the results of the classical method and deep learning models on the FloorNet dataset



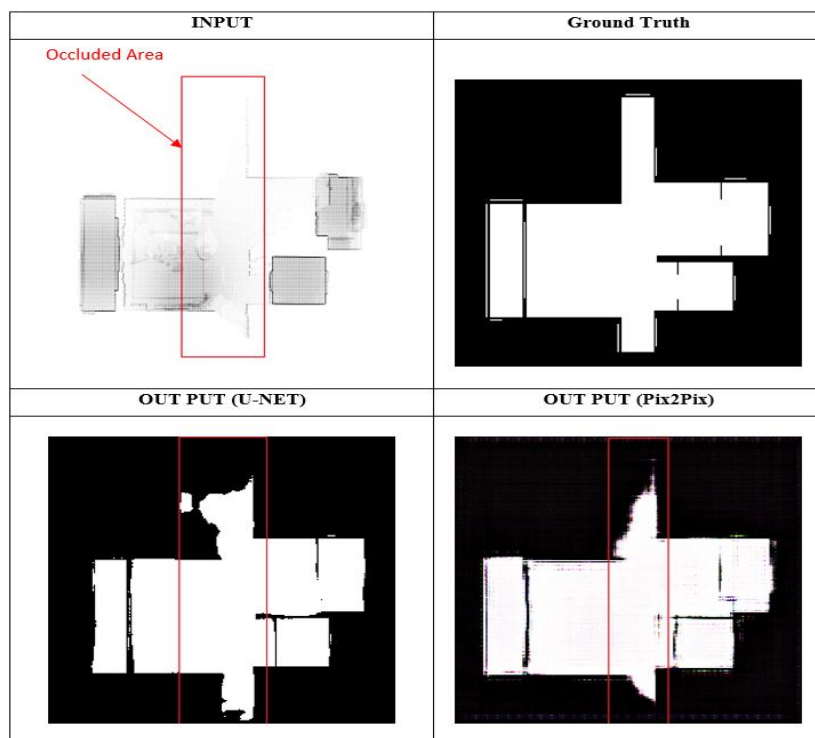
شکل ۱۲: مقایسه نتایج روش کلاسیک و مدل‌های یادگیری عمیق روی دیتاست Structure3D
 Fig. 12: Comparison of the results of the classical method and deep learning models on the Structure3D dataset

(۱۳) و (۱۴) ارائه شده‌اند تا این محدودیت به صورت عینی نمایش داده شود. در این تصاویر، نواحی دارای انسداد با کادر قرمز رنگ نمایش داده شده است. این مسئله نشان می‌دهد که علی‌رغم دقت زیاد مدل‌ها در شرایط داده‌های کامل و بدون نقص، روش پیشنهادی نسبت به داده‌های ناقص یا دارای انسداد، که در کاربردهای واقعی مانند اسکن محیط‌های داخلی بسیار رایج هستند، از پایداری کافی برخوردار نیست.

نواحی دارای انسداد به طور معمول فاقد اطلاعات پیکسلی کافی هستند [۳۱]. در این موارد، تصاویر تراکمی حاصل از ابرنقاط ناقص، فاقد اطلاعات ساختاری کافی برای استنتاج صحیح بخش‌های گم‌شده بودند. این کمبود اطلاعات باعث شد مدل‌ها نتوانند به درستی بخش‌های حذف شده از تصویر تراکم دوبعدی مانند دیوارها یا گوشه‌ها را بازسازی کنند، و در نتیجه دچار اختلال در پیوستگی خطوط و ناتوانی در پیش‌بینی شکل واقعی مرزها می‌شود. نمونه‌هایی از این موارد شکست در شکل



شکل ۱۳: نتایج بصری مدل‌های پیاده‌سازی شده روی نمونه دارای انسداد در مجموعه داده FloorNet
 Fig. 13: Visual results of the implemented models on an occluded sample from the FloorNet dataset



شکل ۱۴: نتایج بصری مدل‌های پیاده‌سازی شده روی نمونه دارای انسداد در مجموعه داده Structure3d
 Fig. 14: Visual results of the implemented models on an occluded sample from the Structure3D dataset

نتیجه‌گیری

چارچوب شامل پیاده‌سازی و ارزیابی دو رویکرد کلاسیک و مبتنی بر یادگیری عمیق است. روش کلاسیک تبدیل هاف، با وجود قابلیت اولیه در شناسایی خطوط مستقیم در برابر نویز، انسداد و ناپیوستگی‌های

در این پژوهش، یک چارچوب موثر جهت استخراج نقشه دوبعدی فضاهای داخلی ساختمان‌ها از داده‌های ابرنقاط ورودی ارائه شد. این

مشارکت نویسندگان

مرتضی حیدری مظفر ایده اولیه کار را مطرح و با ایجاد شرایط لازم برای انجام کار تحقیقاتی، سرکار خانم زهرا دالوند را در جمع‌آوری داده‌ها، تجزیه و تحلیل، برنامه‌نویسی و پردازش داده‌ها و جمع‌بندی نتایج و نگارش مقاله راهنمایی کردند.

تشکر و قدردانی

مقاله‌ی ارسالی، بخشی از کار پایان‌نامه دکتری است که در دانشگاه بوعلی سینا انجام شده است. از همین رو، از گروه عمران و کامپیوتر دانشکده مهندسی که شرایط را فراهم نمودند، تشکر و قدردانی به عمل می‌آید.

تعارض منافع

«هیچ‌گونه تعارض منافع توسط نویسندگان بیان نشده است.»

منابع و مأخذ

- [1] Poux F. The smart point cloud: Structuring 3D intelligent point data [dissertation]: Universite de Liege (Belgium); 2019. doi:10.13140/RG.2.2.20457.75367.
- [2] Heidari Mozaffar M, Varshosaz M, Saadatseresht M. Combining 2D and 3D Edges for Detecting Occluded Regions in Terrestrial Laser Scanner Point Clouds. *Journal of Soft Computing and Information Technology*. 2015;4(4):64-76.
- [3] Westoby M, Brasington J, Glasser N, Hambrey M, Reynolds M. 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*. 2012;179:300-314. doi:10.1016/j.geomorph.2012.08.021.
- [4] McManamon P. Lidar technologies and Systems. USA: Society of Photo-Optical Instrumentation Engineers (SPIE); 2019. doi:10.1117/3.2518254
- [5] Riveiro B, Lindenbergh R. Laser scanning: an emerging technology in structural engineering. China: CRC Press/Balkema; 2019. doi:10.1201/9781351018869
- [6] Heidari Mozaffar M, Hossein SA. Extracting Façade Points of Urban Buildings from Mobile Laser Scanner Point Clouds. *Remote Sensing and Geospatial Information Research*. 2024;2(2):179-186. doi:10.22061/jrsgr.2023.1990.
- [7] Lu Q, Lee S. A semi-automatic approach to detect structural components from CAD drawings for constructing as-is BIM objects. *Computing in Civil Engineering* 2017. p. 84-91. doi:10.1061/9780784480823.011.
- [8] Kumar A, Masane Y, Dhakhwa S. Disaster management plan for residential-cum-commercial property in Mumbai: a case study. *Environment, Development and Sustainability*. 2021;23(8):12598-12614. doi:10.1007/s10668-020-01129-7.

ساختاری داده‌های واقعی عملکرد قابل اعتمادی ندارد. همچنین، خروجی این روش به صورت مجموعه‌ای از خطوط منفرد و بدون انسجام توپولوژیک بود که بازسازی یک نقشه دقیق و یکپارچه را با دشواری مواجه می‌کند. همچنین، نتایج این روش به شدت به مقادیر پارامترها وابسته است و تنظیم دقیق پارامترها منجر به افزایش قابل توجه هزینه محاسباتی و زمانی می‌گردد.

در مقابل، نتایج حاصل از مدل‌های یادگیری عمیق U-Net و Pix2Pix نشان داد که این روش‌ها قادر به استخراج دقیق نقشه دوبعدی فضاهای داخلی ساختمان‌ها هستند. به ویژه، مدل‌های پیشنهادی توانستند در معیار Dice Score به دقت بیش از ۹۰ درصد و در معیار IoU به دقت بیش از ۸۰ درصد دست یابند، که حاکی از توانایی بالای آن‌ها در بازسازی ساختارهای اصلی محیط مانند دیوارها و گوشه‌ها می‌باشد. افزون بر آن، ارزیابی کیفی از طریق مقایسه بصری خروجی مدل‌ها با نقشه‌های مرجع نیز انجام گرفت. نتایج کیفی به دست آمده که با تصاویر نمونه در متن مقاله آورده شده است، نشان می‌دهد که مدل‌های یادگیری عمیق توانستند خطوط مرزی را با وضوح و انسجام بالاتری بازسازی کنند. به‌طور کلی نتایج کمی و کیفی برتری روش‌های مبتنی بر یادگیری عمیق را نسبت به رویکردهای کلاسیک به خوبی نشان می‌دهد.

با این حال، تحلیل نتایج نشان داد که مدل‌های پیشنهادی در شرایطی که داده‌های ابرنقاط دارای انسداد شدید باشد، با کاهش دقت در پیش‌بینی مواجه می‌شوند. در چنین شرایطی، بخش‌هایی از دیوارها یا ساختارهای مرزی به دلیل فقدان اطلاعات کافی به‌درستی بازسازی نمی‌شوند یا دچار گسستگی و اعوجاج در خروجی نهایی می‌شوند. این موضوع به‌ویژه در فضاهای پیچیده، تأثیر بیشتری بر عملکرد مدل دارد و ضرورت اعمال راهکارهایی برای افزایش مقاومت مدل در برابر نقص داده را آشکار می‌سازد. علاوه بر این، یکی دیگر از محدودیت‌های روش پیشنهادی، ماهیت پیکسلی خروجی مدل‌های یادگیری عمیق است که در آن مرزها و دیوارها به‌صورت ناحیه‌های پیکسلی دودویی (binary masks) بازسازی می‌شوند. خروجی پیکسلی، در کاربردهای عملی مانند تولید نقشه‌های CAD یا سیستم‌های اطلاعات مکانی (GIS)، باید به صورت برداری دقیق (vectorization) بازسازی شود. تبدیل خروجی پیکسلی به مرزهای دقیق برداری بدون از دست دادن جزئیات هندسی یا ایجاد خطاهای ساختاری، همچنان یک چالش باز محسوب می‌شود. در ادامه این پژوهش، تمرکز بر توسعه مدل‌هایی خواهد بود که نه تنها در برابر داده‌های ناقص و دارای انسداد مقاوم باشند، بلکه بتوانند به‌صورت مستقیم یا غیرمستقیم، خروجی‌هایی در قالب بردارهای هندسی تولید کنند. از جمله مسیرهای پیشنهادی می‌توان به طراحی شبکه‌هایی با قابلیت یادگیری ساختارهای برداری، بهره‌گیری از توابع زیان (Loss Functions) مبتنی بر هندسه و استفاده از پس‌پردازش‌های ساختاری مانند گراف‌های توپولوژیک اشاره کرد.

- [22] Liu C, Wu J, Furukawa Y. Floornet: A unified framework for floorplan reconstruction from 3d scans. Proceedings of the European conference on computer vision (ECCV); 2018. Cham: Springer. doi:10.13140/RG.2.2.20457.75367.
- [23] Aggarwal N, Karl W. Line detection in images through regularized Hough transform. IEEE transactions on image processing. 2006;15(3):582-591. doi:10.1109/TIP.2005.863021.
- [24] Yang R, Deng C, Yu K, Li Z, Pan L. A new way for cartesian coordinate transformation and its precision evaluation. Remote Sensing. 2022;14(4):864. doi:10.3390/rs14040864.
- [25] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18; 2015. Cham: Springer. doi:10.1007/978-3-319-24574-4_28.
- [26] Goodfellow I, Bengio Y, Courville A, Bengio Y. Deep learning. USA: MIT press Cambridge; 2016. doi:10.4258/hir.2016.22.4.351
- [27] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research. 2014;15(1):1929-1958. doi:10.5555/2627435.2670313.
- [28] Isola P, Zhu J, Zhou T, Efros A. Image-to-image translation with conditional adversarial networks. Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. doi:10.1109/CVPR.2017.632.
- [29] Nambiyar A, Kannan R. Defective Edge Detection Using Cascaded Ensemble Canny Operator. arXiv preprint arXiv:241114868. 2024. doi:10.48550/arXiv.2411.14868.
- [30] Sathyanarayanan S, Tantri B. Confusion matrix-based performance evaluation metrics. African Journal of Biomedical Research. 2024;4023-4031. doi:10.53555/AJBR.v27i4S.4345.
- [31] Heidari Mozaffar M, Varshosaz M, Saadatseresht M. Occlusion Area as Suitable Guidance for Terrestrial Laser Scanner Localization. ISSGE. 2020;10(2):199-207. doi:10.1001.1.2322102.1399.10.2.15.8.
- [9] Gerstweiler G, Furlan L, Timofeev M, Kaufmann H. Extraction of structural and semantic data from 2D floor plans for interactive and immersive VR real estate exploration. Technologies. 2018;6(4):101. doi:10.3390/technologies6040101.
- [10] Pexman K. Automated Floor Plan and Building Model Creation for Cultural Heritage Buildings from Laser Scanner Data. 2021.
- [11] Guo Y, Wang H, Hu Q, Liu H, Liu L, Bennamoun M. Deep learning for 3d point clouds: A survey. IEEE transactions on pattern analysis and machine intelligence. 2020;43(12):4338-4364. doi:10.1109/TPAMI.2020.3005434.
- [12] Golparvar F, Mohammadzadeh A. Semantic Segmentation of point cloud data acquired from Industrial Sites using Terrestrial Laser Scanner based on Graph Methods. Iran: K. N. Toosi University; 2021.
- [13] Pizarro P, Hitschfeld N, Sipiran I, Saavedra J. Automatic floor plan analysis and recognition. Automation in Construction. 2022;140:104-348. doi:10.1016/j.autcon.2022.104348.
- [14] Cai R, Li H, Xie J, Jin X. Accurate floorplan reconstruction using geometric priors. Computers & Graphics. 2022;102:360-369. doi:10.1016/j.cag.2021.10.011.
- [15] Fang H, Lafarge F, Pan C, Huang H. Floorplan generation from 3D point clouds: A space partitioning approach. ISPRS Journal of Photogrammetry and Remote Sensing. 2021;175:44-55. doi:10.1016/j.isprsjprs.2021.02.012.
- [16] Alpaydin E. Introduction to machine learning. Cambridge, MA: MIT press; 2020. doi:10.7551/mitpress/13811.001.0001
- [17] Kim M, Lee D, Kim T, Oh S, Cho H. Automated extraction of geometric primitives with solid lines from unstructured point clouds for creating digital buildings models. Automation in Construction. 2023;145:104642. doi:10.1016/j.autcon.2022.104642.
- [18] Chen J, Qian Y, Furukawa Y. Heat: Holistic edge attention transformer for structured reconstruction. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2022. doi:10.1109/CVPR52688.2022.00377.
- [19] Su J, Tung K, Peng C, Wonka P, Chu H. Slibo-net: Floorplan reconstruction via slicing box representation with local geometry regularization. Advances in Neural Information Processing Systems. 2023;36:48781-48792.
- [20] Yue Y, Kontogianni T, Schindler K, Engelmann F. Connecting the dots: Floorplan reconstruction using two-level queries. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2023.
- [21] Zheng J, Zhang J, Li J, Tang R, Gao S, Zhou Z. Structured3d: A large photo-realistic dataset for structured 3d modeling. Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IX 16; 2020: Springer. doi:10.1007/978-3-030-58545-7_30.

معرفی نویسندگان

AUTHOR(S) BIOSKETCHES



مرتضی حیدری مظفر در سال ۱۳۹۵، دوره دکتری مهندسی عمران - نقشه‌برداری خود را در دانشگاه صنعتی خواجه نصیرالدین طوسی با موفقیت به اتمام رساند. ایشان از سال ۱۳۹۶ با عنوان استادیار گروه عمران دانشکده مهندسی دانشگاه بوعلی سینا مشغول تدریس و پژوهش هستند. نتایج تحقیقات ایشان در قالب ده‌ها مقاله علمی در نشریات و کنفرانسهای مختلف ارائه شده است.

و پژوهش هستند. نتایج تحقیقات ایشان در قالب ده‌ها مقاله علمی در نشریات و کنفرانسهای مختلف ارائه شده است.



خرم‌آباد دریافت کرد. او هم‌اکنون در حال تکمیل دوره دکتری خود در رشته هوش مصنوعی در دانشگاه بوعلی سینا همدان است. علایق او درباره یادگیری عمیق، بینایی ماشین، پردازش تصاویر، پردازش داده‌های لیزر اسکنر، هوش مصنوعی در اطلاعات مکانی و توسعه کاربردهای سنجش از دور است.

Dalvand, Z. Ph.D. Candidate in Computer Engineering, Faculty of Engineering, Bu-Ali Sina University, Hamedan, Iran

✉ z.dalvand@eng.basu.ac.ir

زمینه‌های تخصصی ایشان عبارتند از : فتوگرامتری، سنجش از دور، هوش مصنوعی، کاربردهای سنجش از دور در مدیریت بلایای طبیعی، پردازش داده‌های لیزر اسکنرهای زمینی و همراه، GIS در صنعت آب و برق.

Heidarimozaffar, M. Assistant Professor, Department of Civil Engineering, Faculty of Engineering, Bu-Ali Sina University, Hamadan, Iran

✉ m.heidarimozaffar@basu.ac.ir

زهرا دالوند مدرک کارشناسی مهندسی کامپیوتر خود را از دانشگاه تربیت معلم تهران و کارشناسی ارشد را از دانشگاه آزاد اسلامی واحد

Citation (Vancouver): Heidarimozaffar M, Dalvand Z. [2D Floorplan Extraction of Indoor Spaces from Mobile Laser Scanner Point Clouds Using Deep Learning Methods]. *J. RS. GEOINF. RES.* 2024; 3(2): 219-234

 <https://doi.org/10.22061/jrsg.2025.12127.1100>



COPYRIGHTS

© 2025 The Author(s). This is an open-access article distributed under the terms and conditions of the Creative Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)
(<https://creativecommons.org/licenses/by-nc/4.0/>)