Website: https://www.Sepehr.org Online ISSN: 2588-3879; Print ISSN: 2588-3860

Scientific - Research Quarterly

Geographical Data (SEPEHR) Vo.35, No.138, Summer 2026

P.P 45-64

Research paper

https://doi.org/10.22131/sepehr.2026.2086394.3173

A novel method for extracting two dimensional floor plans of residential buildings from point cloud data

Zahra Dalvand[1], Morteza Heidarimozaffar*[2], Hasan Khotanlou[3]

1- PhD Student, Candidate in computer engineering, Faculty of engineering, Bu-Ali Sina University, Hamedan, Iran. Email: zz.dalvand@eng.basu.ac.ir

2- (*Corresponding author) Assistant professor, Department of civil engineering, University of Bu Ali Sina, Hamedan, Iran. Email: m.heidarimozaffr@basu.ac.ir

3- Professor, Department of computer science, Faculty of engineering, Bu-Ali Sina University, Hamedan, Iran. Email: khotanlou@basu.ac.ir

Article Info

Date of receive:

2026/02/24

Date of last review:

2026/06/21

Date of accept:

2026/07/25

Date of online publication:

2026/07/27

Keywords:

Point cloud,
Normal vector,
Deep learning,
Vector reconstruction

Extended Abstract

Introduction

Extraction of two-dimensional floor plans of residential buildings is a fundamental requirement in architectural, civil engineering, and surveying applications. A 2D floor plan provides a vector-based representation of a building's interior layout, accurately depicting the geometric configuration and spatial arrangement of walls and rooms. The process of extracting residential floor plans is typically time-consuming and requires extensive human involvement. The implementation of semi-automated approaches can significantly facilitate this process, improving both the accuracy and efficiency of floor plan extraction. Point cloud data consist of a discrete set of three-dimensional points that accurately describe the geometric and spatial characteristics of an environment. With the advancement of 3D scanning technologies, rapid, precise, and comprehensive acquisition of indoor building spaces in the form of three-dimensional point cloud data has become increasingly feasible. Accordingly, the use of point cloud data as input for the problem of 2D floor plan extraction in indoor surveying has been steadily increasing. Despite recent advances in deep learning methods, most existing approaches have been developed for structured data or outdoor environments, and their performance remains challenging in enclosed indoor spaces with complex structural layouts. Therefore, the extraction of two-dimensional floor plans of residential buildings from three-dimensional point cloud data—particularly with an emphasis on corner detection, wall boundary delineation, and room topology reconstruction—remains a challenging and open problem. The objective of this study is to propose a novel approach for the vector-based representation of indoor residential spaces from point cloud data, grounded in principal component analysis, deep learning techniques, and morphological operators. ... **Page 46**

How to Cite:

Dalvand, Z. Heidarimozaffar, M. Khotanlou, H. (2026). A novel method for extracting two dimensional floor plans of residential buildings from point cloud data. Scientific-Research Quarterly Geographical Data (SEPEHR). 35(138), 45-64.

Materials & Methods

The main contribution of this research lies in the development of an effective framework for extracting two-dimensional floor plans of indoor residential spaces from three-dimensional laser-scanned point clouds. The proposed floor plan extraction process consists of three principal stages: data preprocessing, segmentation, and vector-based reconstruction. In the first stage, the normal vector of each point is estimated using the Principal Component Analysis (PCA) algorithm. The normal vector at a point is defined as the vector perpendicular to the surface at that location. The main idea is to utilize the orientation of the normal vectors as a criterion for surface type identification. Points are classified based on the nz component of their normal vectors in order to accurately distinguish between horizontal and vertical surfaces. By removing points corresponding to horizontal surfaces—such as floors, ceilings, and horizontal furniture elements—only vertical surfaces are preserved for subsequent processing. to reduce computational complexity in terms of execution time and memory consumption, the processed point cloud is transformed into a density image, and the floor plan extraction procedure is subsequently carried out in the two-dimensional image domain. Focusing on points associated with vertical wall surfaces, which define the structural boundaries of the space, enables more accurate and efficient reconstruction of the floor plan. The density image represents the spatial distribution of points projected onto the horizontal X–Y plane and serves as the basis for subsequent data analysis stages.

In the second stage, two deep learning models—U-Net and SegFormer—are employed to extract a pixel-based two-dimensional floor plan from the density image. In the third stage, a set of morphological operations is applied to convert the pixel-wise model output into vector-based representations.

To evaluate the performance of the proposed method and to gain deeper insight into its capabilities and limitations, the Structured3D point cloud dataset was employed.

Results & Discussion

The U-Net and SegFormer models were trained to segment the density images. In this stage, model performance was evaluated on the test dataset using the Dice Score and Mean Intersection over Union (Mean IoU) metrics. The U-Net and SegFormer models achieved Dice Scores of 97% and 94%, respectively, and Mean IoU values of 94% and 90%, respectively. The results of the study indicated that, although SegFormer performs well in capturing overall structural regions, it exhibits lower accuracy in reconstructing thin lines and precise wall boundaries compared to the U-Net model. Ultimately, U-Net was selected as the final model due to its superior performance across the evaluation metrics.

Morphological operators were used for the vector representation of the maps. To evaluate the accuracy of the vector reconstruction of the building maps, the standard metrics Precision and Recall were measured at the corner and room levels and were compared with the results of the RoomFormer model. The research findings showed that the proposed method achieved accuracies of 84% and 85% in the Precision and Recall metrics at the room level, and 75% and 82% at the corner level. The RoomFormer model achieved accuracies of 97.9% and 96.7% in the Precision and Recall metrics at the room level, and 89.1% and 85.3% at the corner level. Unlike previous methods that rely on the two-stage training of Mask R-CNN and transformer models to generate vector maps, the proposed method has lower computational complexity and simultaneously provides stable results with desirable quality.

Conclusion

In this study, a semi-automated method for extracting two-dimensional floor plans from laser-scanned point cloud data is presented. In the first stage, the model output is generated as a pixel-based image, which is then converted into a vector representation suitable for CAD software using morphological operations. The subsequent floor plan generation process still requires human-guided drafting and revision. This approach, however, provides significant time savings and simplifies the subsequent stages of map refinement. This approach enables easy editing and refinement of terrestrial scan results by surveying engineers. The main advantage of the proposed method lies in its semi-automated workflow, user interactivity, and substantial reduction of the manual processing time required to generate floor plans from raw surveying scanner data.

The experimental results demonstrated that the proposed method, with its simpler architecture, lower computational resource requirements, and satisfactory quality in vector-based floor plan representation, can serve as a practical and cost-effective alternative to transformer-based approaches.

شاپای الکترونیکی: 2588-3879 / شاپای چاپی: 2588-3860
تارنمای نشریه https://www.Sepehr.org

فصلنامه علمی - پژوهشی
اطلاعات جغرافیایی (سپهر) دوره ۳۵، شماره ۱۳۸، تابستان ۱۴۰۵

مقاله پژوهشی
صفحات ۶۴ - ۴۵
https://doi.org/10.22131/sepehr.2026.2086394.3173

روشی نوین برای استخراج نقشه دوبعدی ساختمان‌های مسکونی از داده‌های ابرنقاط

زهرا دالوند[۱]، مرتضی حیدری‌مظفر*[۲]، حسن ختن‌لو[۳]

۱- دانشجوی دکتری، گروه کامپیوتر، دانشکده مهندسی، دانشگاه بوعلی سینا، همدان، ایران. zz.dalvand@eng.basu.ac.ir
۲- (*نویسنده مسئول) استادیار گروه عمران، دانشکده مهندسی، دانشگاه بوعلی سینا. m.heidarimozaffr@basu.ac.ir
۳- استاد گروه کامپیوتر، دانشکده مهندسی، دانشگاه بوعلی سینا، همدان، ایران. khotanlou@basu.ac.ir

اطلاعات مقاله

تاریخ دریافت:
۱۴۰۴/۱۲/۰۵
تاریخ آخرین بازنگری:
۱۴۰۵/۰۳/۳۱
تاریخ پذیرش:
۱۴۰۵/۰۵/۰۳
تاریخ انتشار:
۱۴۰۵/۰۵/۰۵

واژه‌های کلیدی:

ابرنقاط؛
بردار نرمال؛
یادگیری عمیق؛
بازسازی برداری

چکیده

استخراج نقشه دوبعدی ساختمان‌های مسکونی، یک نیاز اساسی در مطالعات و کاربردهای حوزه‌های معماری، مهندسی عمران و نقشه‌برداری به‌شمار می‌رود. نقشه دوبعدی، نمایش برداری از چیدمان داخلی ساختمان است که موقعیت و شکل هندسی دقیق دیوارها و اتاق‌ها را بازنمایی می‌کند. استخراج نقشه ساختمان‌های مسکونی، فرآیندی زمان‌بر است و به مشارکت انسانی گسترده نیاز دارد. پیاده‌سازی روش‌های نیمه خودکار می‌تواند این روند را تسهیل کرده و دقت و سرعت استخراج نقشه‌ها را به‌طور قابل توجهی افزایش دهد. داده‌های ابرنقاط، بازنمایی دقیق هندسه سه‌بعدی محیط هستند و جزئیات ساختار داخلی ساختمان را به‌خوبی ارائه می‌دهند. هدف این تحقیق، ارائه روشی نوین برای بازنمایی برداری فضای داخلی ساختمان‌های مسکونی از داده‌های ابرنقاط بر پایه تحلیل مؤلفه‌های اصلی، یادگیری عمیق و عملگرهای مورفولوژیکی است. در این روش با تحلیل مؤلفه‌های اصلی، بردار نرمال نقاط که ویژگی اصلی برای تفکیک دیوارها است، محاسبه می‌شود. ابتدا تصاویر تراکم حاصل از پیش‌پردازش ابرنقاط، با مدل یادگیری عمیق U-Net قطعه‌بندی می‌شوند. سپس خروجی پیکسل-مبنای قطعه‌بندی با استفاده از عملگرهای مورفولوژیکی به صورت برداری بازسازی می‌شود. عملکرد روش پیشنهادی، بر روی مجموعه‌داده Structure3D مورد ارزیابی قرار گرفت. برای سنجش دقت بازسازی برداری نقشه ساختمان، معیارهای استاندارد Precision و Recall در سطح گوشه‌ها و اتاق‌ها اندازه‌گیری شد. نتایج آزمایش‌ها نشان داد که نرخ دقت و نرخ بازخوانی روش پیشنهادی در سطح گوشه‌ها به ترتیب ۷۵٪ و ۸۲٪ است. برخلاف روش‌های پیشین که وابسته به آموزش دو مرحله‌ای مدل‌های Mask R-CNN و ترنسفورمر برای تهیه نقشه‌های برداری هستند، روش پیشنهادی پیچیدگی محاسباتی کمتری دارد و همزمان نتایج پایدار و با کیفیت مطلوب ارائه می‌دهد.

استناد به این مقاله:

دالوند، ز؛ حیدری مظفر، م؛ ختن لو، ح (۱۴۰۵) روشی نوین برای استخراج نقشه دوبعدی ساختمان‌های مسکونی از داده‌های ابرنقاط؛ فصلنامه علمی پژوهشی اطلاعات جغرافیایی(سپهر)، ۳۵(۱۳۸)، ۶۴ – ۴۵.

۱- مقدمه

نقشه‌های دوبعدی در طراحی، درک و بازسازی فضای داخلی ساختمان‌های مسکونی نقش حیاتی دارند. داده ابرنقاط مجموعه‌ای گسسته از نقاط سه‌بعدی است که ویژگی‌های هندسی و فضایی محیط را با دقت بالا توصیف می‌کند. این نوع داده، منبع غنی اطلاعات و پایه‌گذار برنامه‌های تصمیم‌گیری است (Poux, 2019). در ابرنقاط، هر نقطه دارای مختصات X، Y و Z است. نحوه اندازه‌گیری، پردازش و استخراج اطلاعات قابل اطمینان از مجموعه داده‌های جمع‌آوری شده، حائز اهمیت است. داده‌های ابرنقاط با روش‌های فتوگرامتری و از طریق تناظریابی تصاویر و یا با بهره‌گیری از فناوری اسکن سه‌بعدی تهیه می‌شوند. این دو رویکرد با دقت و سازوکار متفاوت، امکان بازسازی دقیق هندسه محیط را فراهم می‌کنند (Westoby et al., 2012). فناوری اسکن لیزری به دلیل برخورداری از دقت میلی‌متری پرتو لیزر و تکیه بر جهت‌یابی دقیق حسگر، که توسط سامانه‌های یکپارچه موقعیت و جهت‌گیری[1] پشتیبانی می‌شود، نسبت به روش‌های فتوگرامتری بسیار دقیق‌تر است (McManamon, 2019). استخراج نقشه دوبعدی ساختمان‌های مسکونی از این داده‌ها، کاربرد گسترده‌ای در مستندسازی و مدل‌سازی اطلاعات ساختمان[2] (Lu & Lee, 2017)، مدیریت بحران و امداد و نجات (Kumar et al., 2021)، ناوبری رباتیک و تولید واقعیت مجازی (Gerstweiler et al., 2018)، شبیه‌سازی نورپردازی (He et al., 2021)، تهیه سامانه‌های کمک به افراد نابینا و کم‌بینا (Li et al., 2018) دارد.

با گسترش فناوری اسکن سه‌بعدی، مانند اسکنرهای لیزری زمینی[3]، سامانه‌های برداشت متحرک[4]، حسگرهای عمق‌سنج[5] و سیستم‌های مکان‌یابی و نقشه‌سازی هم‌زمان[6]، امکان برداشت سریع، دقیق و جامع از فضای داخلی ساختمان به صورت داده‌های ابرنقاط سه‌بعدی فراهم شده است (Riveiro & Lindenbergh, 2019). ازاین‌رو، روند استفاده از ابرنقاط به عنوان ورودی مسأله استخراج نقشه دوبعدی در نقشه‌برداری داخلی افزایش یافته است (Chen et al., 2019) و (Yue et al., 2023). با این حال، در مواجهه با نویز و نقاط خارج از محدوده، داده‌های غیرساخت‌یافته و بدون ترتیب و چالش‌های ذاتی دیگر در ابرنقاط خام، بازسازی نقشه دوبعدی همچنان یک مسأله باز و دشوار است (Guo et al., 2020).

با وجود پیشرفت‌های اخیر در روش‌های یادگیری عمیق، اغلب این روش‌ها برای داده‌های ساخت‌یافته یا محیط‌های بیرونی طراحی شده‌اند و عملکرد آن‌ها در محیط‌های بسته با ساختار پیچیده داخلی با چالش مواجه است. ازاین‌رو، استخراج نقشه دوبعدی ساختمان‌های مسکونی از داده ابرنقاط سه‌بعدی، با تمرکز بر شناسایی گوشه‌ها، مرز دیوارها و بازسازی توپولوژی اتاق‌ها، یک مسئله باز و دشوار است.

برای کاهش پیچیدگی مسئله، اغلب ابرنقاط سه‌بعدی در امتداد قائم تصویر شده[7] و از تصویر تراکم حاصل برای پردازش استفاده می‌شود (Chen et al., 2022) و (Yue et al., 2023). بااین‌حال چگونگی تشکیل تصویر تراکم به دلیل تأثیر مستقیم بر کیفیت خروجی، بسیار اهمیت دارد. در این رویکرد، فرآیند استخراج نقشه به دو زیرمسئله قطعه‌بندی تصاویر تراکم و سپس بازنمایی برداری نقشه تقسیم می‌شود. بسیاری از مدل‌های یادگیری عمیق به صورت پیکسل–مبنا به پردازش و قطعه‌بندی تصاویر می‌پردازند. قطعه‌بندی تصویر به معنای تخصیص یک برچسب معنایی به هر پیکسل است، به‌گونه‌ای که نواحی همگن در تصویر بر اساس ساختار یا معنای مشترک تفکیک شوند. در کاربرد نقشه‌برداری، تصاویر تراکم حاصل از ابرنقاط به عنوان ورودی در نظر گرفته شده و قطعه‌بندی آن‌ها موجب شناسایی بخش‌های مختلف مانند دیوارها، کف یا سایر اجزای محیط می‌شود. این رویکرد علاوه بر کاهش پیچیدگی داده‌های سه‌بعدی، زمینه را برای بازنمایی برداری و استخراج نقشه‌های دقیق و ساخت‌یافته فراهم می‌نماید.

1. Position and orientation system
2. Building information model
3. Terrestrial laser scanning
4. Mobile mapping system
5. RGB-D sensors
6. Simultaneous localization and mapping
7. Vertical projection

به‌طور کلی روش‌های ارائه شده برای این مسئله بر اساس نوع عملیات اصلی یا هدف میانی هر روش در فرآیند استخراج نقشه را می‌توان به چهار دسته زیر طبقه‌بندی کرد: (۱) روش‌های مبتنی بر قواعد گرامری، (۲) روش‌های مبتنی بر استخراج گوشه‌ها، (۳) روش‌های مبتنی بر تقسیم‌بندی فضا و (٤) روش‌های مبتنی بر اتصال اجزای ساختاری.

• دسته نخست، شامل رویکردهای سنتی است که پیش از فراگیری الگوریتم‌های یادگیری ماشین به‌کار گرفته می‌شدند. این روش‌ها بر پایه مجموعه‌ای از قواعد صریح و الگوهای از پیش تعریف شده عمل کرده و هدف آن‌ها بازسازی نقشه دوبعدی ساختمان است. بااین‌حال، به علت ضرورت تنظیم دستی پارامترها، انجام مراحل متعدد پیش‌پردازش متناسب با نوع داده و طراحی قواعد سازگار با سبک‌های مختلف معماری، سطح اتوماسیون در این روش‌ها پایین بوده و معمولاً به دخالت مستقیم کاربر وابسته‌اند ***(Pizarro et al., 2022)***. در ***(Cai et al., 2022)*** یک روش برای بازسازی نقشه ساختمان به صورت از پایین به بالا ارائه شده است که از پیش‌فرض‌های هندسی استخراج شده از ابرنقاط خام صحنه‌های داخلی استفاده می‌کند. اگرچه این دسته از روش‌ها، در محیط‌هایی با ساختار هندسی ساده و منظم عملکرد مطلوب نشان می‌دهند، اما در مواجهه با فضاهای پیچیده، نامتعارف یا طراحی‌ها و معماری‌های مدرن با محدودیت‌های جدی روبه‌رو هستند.

• دسته دوم، با تشخیص گوشه‌ها و اتصال آن‌ها به صورت دستی یا نیمه خودکار، نقشه‌ها را بازسازی می‌کنند. مدل HEAT[1] مبتنی بر معماری ترنسفورمر، برای شناسایی گوشه‌ها و استنتاج روابط اتصال میان آن‌ها به‌منظور استخراج گراف مسطح نقشه ارائه شده است. این روش‌ها به شدت به دقت در تشخیص گوشه‌ها وابسته‌اند، به‌گونه‌ای که عدم شناسایی یک گوشه می‌تواند منجر به بازسازی ناقص و حذف بخش قابل‌توجهی از نقشه شود.

• دسته سوم، با تقسیم فضای دوبعدی به چندضلعی‌های کوچک‌تر و اختصاص برچسب‌های اتاق به هرچندضلعی، نقشه دوبعدی را استخراج می‌کنند. در پژوهشی مدل SLIBO–[2] Net ارائه شده است که با استفاده از رگرسیون جعبه‌ای، اطلاعات محلی شکل هر اتاق را به طور مستقیم استخراج می‌کند ***(Su et al., 2023)***. اگرچه این مدل قابلیت نمایش صریح هندسه اتاق‌ها را فراهم می‌کند، اما در نمایش ساختارهای غیرمنهتن[3] با محدودیت مواجه است. در برخی روش‌های پیشین، با استفاده از مدل Mask R–CNN ***(Kaiming He, 2017)*** ابتدا محدوده چندضلعی اتاق‌ها را استخراج کرده و سپس با تعریف یک تابع بهینه‌ساز به استخراج دیوارها و تشخیص مرز اتاق‌ها پرداختند ***(Chen et al., 2019)***، ***(Fang et al., 2021)*** و ***(Stekovic et al., 2021)***. با یک رویکرد مشابه، یو و همکاران. ***(Yue et al., 2023)*** مدل RoomFormer را ارائه دادند. آن‌ها ابتدا با مدل Mask R–CNN محدوده چندضلعی اتاق‌ها را تشخیص داده و سپس با آموزش یک مدل ترنسفورمر به اصلاح چندضلعی‌ها و بازسازی برداری نقشه دوبعدی پرداختند. آن‌ها با تعریف دو سطح پرسش و مدل‌سازی نقشه دوبعدی به عنوان مجموعه‌ای از چندضلعی‌ها با گوشه‌های مرتب شده و طول متغیر، گامی مؤثر در بهبود کیفیت بازسازی برداشتند. بااین‌حال، 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 است.

• دسته چهارم، با تشخیص اجزای ساختاری مانند دیوارها و اتصال آن‌ها به صورت خودکار یا نیمه‌خودکار، نقشه دوبعدی را بازسازی می‌کنند. وظیفه اصلی در این روش‌ها تشخیص لبه است. در یک پژوهش، روشی مبتنی بر ترکیب الگوریتم‌های یادگیری ماشین شامل خوشه‌بندی جابه‌جایی میانگین ***(Comaniciu & Meer, 2002)***، پوشش‌دهی با استفاده از پوسته‌های غیرمحدب و اجماع نمونه‌برداری تصادفی ***(Fischler & Bolles, 1981)*** برای استخراج نقشه دوبعدی ارائه شده است. ***(Kim et al., 2023)***

1. Holistic Edge Attention Transformer

2. Slicing Box Representation with Local Geometry Regularization

3. non-Manhattan structures

به طور کلی، روش‌های پیشین، در استخراج نقشه دوبعدی ساختمان دارای مزایا و محدودیت‌های متعددی بوده‌اند. روش‌های کلاسیک پردازش تصویر مبتنی بر الگو، اگرچه پیاده‌سازی ساده دارند و نیازی به داده آموزشی ندارند، اما دقت آن‌ها پایین است. روش‌های یادگیری ماشین با وجود توانایی در نیمه‌خودکارسازی فرآیند، نیازمند تنظیم دقیق پارامترها هستند و در ساختمان‌های پیچیده عملکرد ضعیفی از خود نشان داده‌اند. در برخی مطالعات، از یادگیری عمیق تنها برای تشخیص چندضلعی‌های اولیه استفاده شده و بازسازی برداری نقشه با کمک توابع بهینه‌ساز انجام گرفته است. با این حال این دسته از روش‌ها به تابع بهینه‌سازی مبتنی بر پیش‌فرض و زمان طولانی بهینه‌سازی محدود هستند. در نهایت، روش‌هایی که در هر دو مرحله استخراج پیکسل-مبنا و بازنمایی برداری از یادگیری عمیق استفاده کرده‌اند، اگرچه دقت بالاتری داشته‌اند، اما به آموزش دو مرحله‌ای مدل، پیچیدگی محاسباتی بالا و داده آموزشی گسترده وابسته هستند.

نوآوری این تحقیق، ارائه‌ی یک چارچوب مؤثر برای استخراج نقشه دوبعدی فضای داخلی ساختمان‌های مسکونی از ابرنقاط اسکن لیزری سه‌بعدی است. مراحل انجام کار به صورت زیر است:

- در گام اول، به کمک الگوریتم تحلیل مؤلفه‌های اصلی ***(Hotelling, 1933)***، بردار نرمال نقاط را محاسبه می‌کنیم. نقاط براساس مؤلفه nz بردار نرمال گروه‌بندی می‌شوند تا سطوح افقی و عمودی به درستی شناسایی شوند. با حذف نقاط مربوط به سطوح افقی مانند کف، سقف و هم‌چنین سطوح افقی مبلمان داخلی، سطوح عمودی حفظ می‌شود. سپس تصویر تراکم مربوط به ابرنقاط پردازش شده تشکیل می‌شود.
- در گام دوم، با استفاده از مدل یادگیری عمیق U-NET ***(Ronneberger et al., 2015)***، نقشه دوبعدی پیکسل-مبنا از تصویر تراکم استخراج می‌شود.
- در گام سوم، با استفاده از عملگرهای مورفولوژیکی، نقشه دوبعدی به صورت برداری استخراج می‌شود.

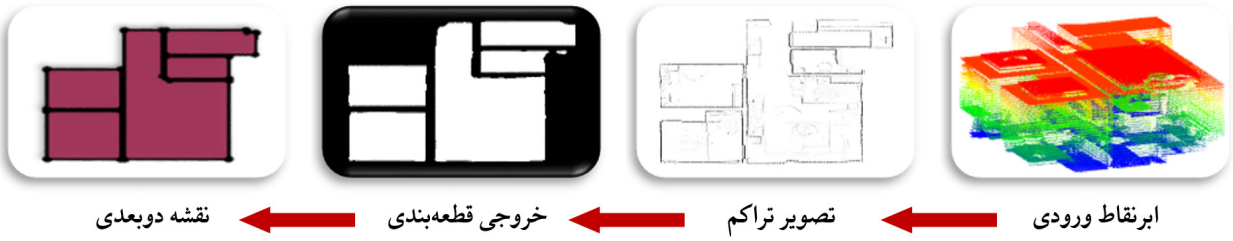
هدف اصلی این تحقیق، ارائه روشی جدید برای استخراج نقشه دوبعدی ساختمان‌های مسکونی با نمایه‌سازی برداری از داده‌های ابرنقاط است. در روش پیشنهادی، برخلاف روش‌های مبتنی بر گرامر، از یادگیری عمیق برای درک زمینه استفاده می‌شود. شبکه عصبی U-NET به تشخیص کامل تک‌گوشه وابسته نیست و در برابر حذف یک گوشه مقاوم‌تر است. علاوه بر این، روش پیشنهادی با ترکیب نقشه تراکم، شبکه عصبی و عملگرهای مورفولوژیکی، فرآیندی یکپارچه ارائه می‌دهد که نیازی به بهینه‌سازی جداگانه یا مراحل تکراری ندارد. مراحل اصلی استخراج نقشه به روش پیشنهادی، در نگاره ۱ نشان داده شده است. در ادامه، ابتدا مجموعه داده‌های مورد استفاده و معیارهای ارزیابی معرفی می‌شوند، سپس به ارائه جزئیات روش پیشنهادی پرداخته و در نهایت، نتایج حاصل از اجرای این روش‌ها گزارش و مورد بحث قرار می‌گیرند.

۲- داده‌ها و روش تحقیق

۲-۱- مجموعه داده‌ها

به‌منظور بررسی عملکرد روش پیشنهادی و شناخت بهتر قابلیت‌ها و محدودیت‌های آن، از دیتاست ابرنقاط Structured3D[1] استفاده شده که در دسترس عموم قرار دارد. این مجموعه داده شامل داده‌های مصنوعی از محیط‌های داخلی است. مجموعه داده Structured3D، یک مجموعه داده مصنوعی از محیط‌های داخلی ساختمان است که شامل ۳۵۰۰ برداشت و نقشه‌های مختلف طبقه می‌شود. در این دیتاست، هر صحنه فضای داخلی یک طبقه را نمایش می‌دهد. داده‌ها دارای برچسب‌گذاری ساختاری سه‌بعدی هستند که برای شبیه‌سازی محیط‌های واقعی طراحی شده‌اند. این مجموعه داده شامل صحنه‌های ساده و پیچیده است***(Zheng et al., 2020)***. به منظور آموزش شبکه عصبی پیشنهادی، تعداد کل نمونه‌ها به صورت ۳۰۰۰ نمونه برای آموزش، ۲۵۰ نمونه برای اعتبارسنجی و ۲۵۰ نمونه برای ارزیابی تقسیم‌بندی شده‌اند.

1. https://github.com/bertjiazheng/Structured3D



نگاره ۱: مراحل اصلی استخراج نقشه دوبعدی در روش پیشنهادی

در این پژوهش، به‌منظور ارزیابی قابلیت تعمیم روش پیشنهادی روی داده‌های واقعی، از مجموعه‌داده SceneCAD استفاده شد. مجموعه‌داده SceneCAD شامل داده‌های واقعی از فضاهای داخلی ساختمان است. برخلاف مجموعه‌داده‌هایی که نقشه کامل یک طبقه را ارائه می‌کنند، هر نمونه در SceneCAD مربوط به یک فضای تک‌اتاقه است. نمونه‌های این مجموعه‌داده از محیط‌های واقعی استخراج شده‌اند و دارای ویژگی‌هایی نظیر نویز اندازه‌گیری، نقاط پرت، تغییرات تراکم نقطه‌ای، انسدادهای موضعی و نامنظم هندسی هستند. داده‌های مصنوعی بازنمایی کاملی از این ویژگی‌ها ارائه نمی‌دهند. هدف اصلی این پژوهش استخراج ساختار فضای داخلی ساختمان در مقیاس نقشه طبقه است. با توجه به اینکه در زمان انجام پژوهش، مجموعه‌داده واقعی عمومی که شامل نقشه کامل طبقه باشد در دسترس نبود، از مجموعه داده SceneCAD برای ارزیابی تکمیلی روش پیشنهادی روی نمونه‌های واقعی تک‌اتاقه استفاده شد.

۲-۲- معیارهای ارزیابی

در این تحقیق، ارزیابی عملکرد در دو مرحله انجام شده است: نخست، بر اساس نتایج پیکسل-مبنای حاصل از مدل قطعه‌بندی و دوم، بر اساس نتایج برداری شده که به صورت نقشه دوبعدی بازنمایی شده‌اند. در ادامه، معیارهای مورد استفاده در هر مرحله معرفی می‌شوند.

در مرحله اول به منظور ارزیابی عملکرد الگوریتم قطعه‌بندی پیشنهادی، معیارهای استاندارد شباهت‌سنجی شامل دو شاخص Dice score و MIOU[1] استفاده شد. این شاخص‌ها، هم‌پوشانی بین تصاویر واقعیت زمینی[2] و خروجی الگوریتم قطعه‌بندی پیشنهادی را اندازه‌گیری می‌کنند. تعریف این دو شاخص در روابط (۱) و (۲) ارائه شده است:

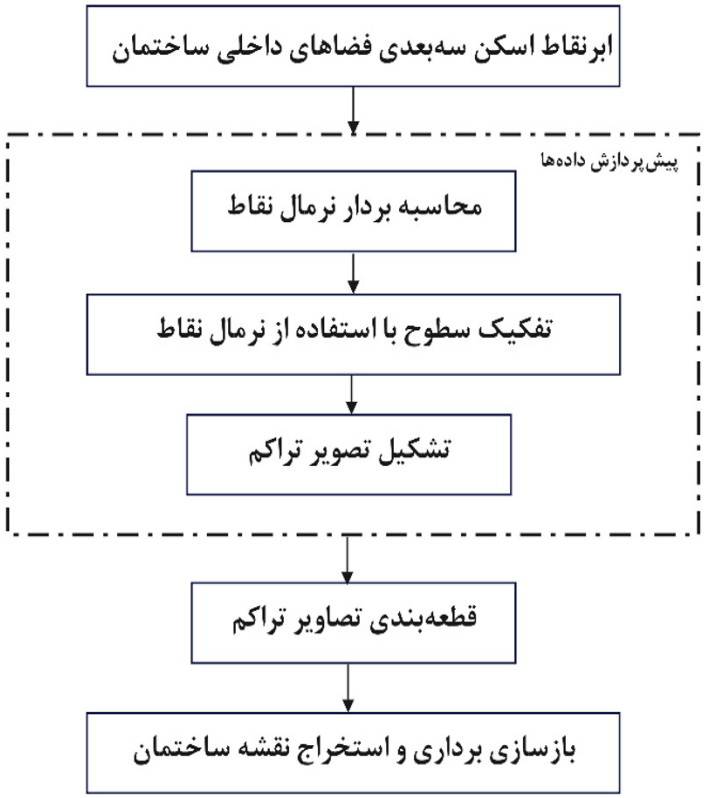
$$\text{Dice score } (A,B) = \frac{2|A \cap B|}{|A|+|B|} \qquad (۱)$$

$$\text{Mean IOU} = \frac{1}{n}\sum_{k=1}^{n}\frac{|A_k \cap B_k|}{|A_k \cup B_k|} \qquad (۲)$$

در اینجا A و B به ترتیب تصویر واقعیت زمینی 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. در مرحله دوم پس از تبدیل خروجی پیکسل-مبنا به فرمت برداری، ارزیابی عملکرد مدل در دو حوزه شناسایی گوشه‌ها[3] و بازسازی اتاق‌ها[4] با استفاده از معیارهای Precision و Recall انجام گرفته است. این معیارها به منظور مقایسه با روش‌های پیشین و سنجش عملکرد مدل به کار می‌روند. یک گوشه در صورتی به درستی بازسازی شده محسوب می‌شود که در فاصله‌ای کمتر از ۱۰ پیکسل از گوشه واقعی قرار گیرد. این حد آستانه مطابق کارهای پیشین *(Stekovic et al., 2021)* و *(Yue et al., 2023)* انتخاب شد. در صورتی که چندین گوشه در نزدیکی یک گوشه واقعی شناسایی شوند، تنها نزدیک‌ترین گوشه به عنوان صحیح در

1. Mean Intersection over Union

2. Ground Truth
3. Corners
4. Rooms



نگاره ۲: روندنمای روش پیشنهادی

نظر گرفته شده و سایر گوشه‌ها به عنوان مثبت کاذب[1] در نظر گرفته می‌شوند. تعریف این دو شاخص در روابط (۳) و (٤) ارائه شده است.

$$\text{Precision}_{\text{corner}} = \frac{\text{TP}_{\text{corner}}}{\text{TP}_{\text{corner}} + \text{FP}_{\text{corner}}} \qquad (۳)$$

$$\text{Recall}_{\text{corner}} = \frac{\text{TP}_{\text{corner}}}{\text{TP}_{\text{corner}} + \text{FN}_{\text{corner}}} \qquad (٤)$$

که در آن:

TP_{corner} تعداد گوشه‌های صحیح شناسایی شده است.

FP_{corner} تعداد گوشه‌های مثبت کاذب است.

FN_{corner} تعداد گوشه‌های واقعی که به درستی شناسایی نشده‌اند.

برای سنجش دقت بازسازی نقشه و تشخیص اتاق‌ها نیز همان معیارهای Precision و Recall مورد استفاده قرار گرفته‌اند. در اینجا، یک اتاق پیش‌بینی شده صحیح تلقی می‌شود اگر:

- شکل آن به صورت چندضلعی بسته معتبر باشد.
- با یکی از اتاق‌های واقعی تطبیق بیش از ٦٠ درصد داشته باشد. این حد آستانه مطابق کارهای پیشین ***(Stekovic et al., 2021)*** و ***(Yue et al., 2023)*** انتخاب شد.

۲-۳- روش پیشنهادی

فرآیند استخراج نقشه در روش پیشنهادی شامل سه مرحله اصلی پیش‌پردازش داده‌ها، قطعه‌بندی و بازسازی برداری است. در مرحله پیش‌پردازش، ابرنقاط سه‌بعدی پردازش شده و تصاویر تراکم تهیه می‌شود. سپس فرآیند قطعه‌بندی روی این تصاویر انجام شده و در نهایت، خروجی پیکسلی حاصل از این مرحله به یک نمایش برداری ساختاریافته تبدیل می‌شود. روندنمای روش پیشنهادی، در نگاره ۲ نشان داده شده است. در ادامه به تشریح روش پیشنهادی خواهیم پرداخت.

1. False Positives

۲-۳-۱- محاسبه بردار نرمال نقاط

برآورد نرمال‌های معتبر، یکی از مراحل اساسی پیش پردازش ابرنقاط است. بردار نرمال یک نقطه، بردار عمود بر سطح در آن نقطه است. برای محاسبه بردار نرمال نقاط، از الگوریتم آنالیز مؤلفه‌های اصلی استفاده شد (*El-Sayed et al., 2018*). الگوریتم PCA[1] یکی از روش‌های آماری است که برای کاهش ابعاد داده‌ها و شناسایی الگوهای نهفته در داده‌ها به کار می‌رود. این الگوریتم با تشکیل ماتریس کوواریانس داده‌ها و محاسبه مقادیر ویژه[2] و بردارهای ویژه[3] آن، اطلاعات مهم و مرتبط با ساختار نقاط را استخراج می‌کند. مراحل اجرای الگوریتم PCA به صورت زیر است:

• جمع‌آوری داده‌ها: ابرنقاط ورودی دارای تعداد N نقطه است که برای هر نقطه مختصات سه‌بعدی X، Y و Z ثبت شده است.

• تعیین همسایگی نقاط: برای به دست آوردن ماتریس کوواریانس و مقادیر ویژه، شناسایی همسایگان نزدیک برای هر نقطه ضروری است. همسایگان نزدیک هر نقطه با استفاده از الگوریتم K نزدیک‌ترین همسایه[4] تعیین شده است. به‌منظور تعیین مقدار بهینه K، مقادیر مختلف این پارامتر به‌صورت تجربی آزمایش شد و درنهایت بر اساس پایداری نرمال‌های حاصل، مقدار k برابر ۳۰ در نظر گرفته شد. نرمال‌های نهایی، با اجرای PCA بر روی همسایگی‌های تولید شده حاصل می‌شوند.

• محاسبه ماتریس کوواریانس: در این مرحله برای هر نقطه، براساس اطلاعات همسایگان آن ماتریس کوواریانس تشکیل می‌شود. ماتریس کوواریانس در این حالت یک ماتریس مربعی با ابعاد ۳*۳ است. این ماتریس ویژگی‌هایی چون همبستگی بین مختصات X، Y و Z را به نمایش می‌گذارد. رابطه (۵) نحوه محاسبه کوواریانس را نشان می‌دهد:

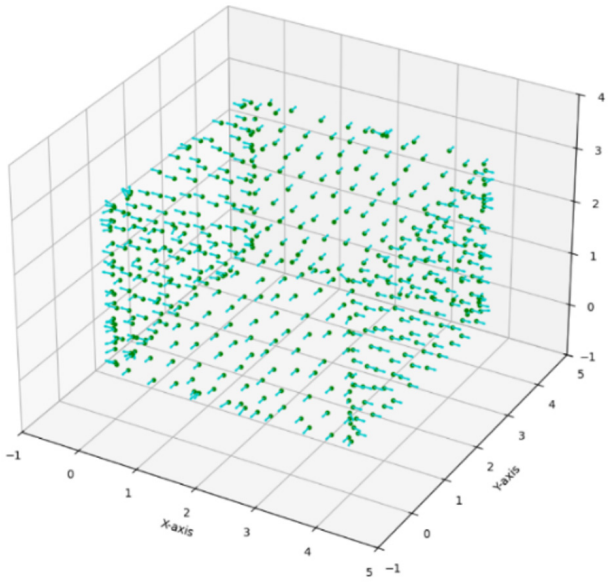
$$C = \frac{1}{N-1} \sum_{i=1}^{n} (P_i - \bar{P})(P_i - \bar{P})^T \qquad (۵)$$

که در آن Pi نماینده نقاط و میانگین نقاط است.

• محاسبه مقادیر ویژه و بردارهای ویژه: با محاسبه مقادیر ویژه λ و یافتن بردارهای ویژه v مربوط به ماتریس کوواریانس، جهت‌های اصلی داده‌ها مشخص می‌شود. بردارهای ویژه به ما کمک می‌کنند تا درک بهتری از توزیع نقاط داشته باشیم.

• انتخاب بردار نرمال: بردار نرمال هر نقطه، برابر با بردار ویژه متناظر با کوچکترین مقدار ویژه ماتریس کوواریانس نقاط همسایه است. این امر به دلیل این است که این بردار در واقع نماینده جهت عمود بر صفحه‌ای است که نقاط در آن قرار گرفته‌اند.

نگاره ۳ تصویر ابرنقاط یک اتاق به همراه بردار نرمال نقاط را نشان می‌دهد. بردار نرمال نقاط، جهت سطوح دیوارها را بازتاب داده و هم‌راستایی مناسبی با صفحات تشکیل دهنده دیوارها دارند.



نگاره ۳: نمایش بردار نرمال نقاط

۲-۳-۲- تفکیک سطوح با استفاده از نرمال نقاط

ابرنقاط ورودی دارای مختصات سه‌بعدی (x, y, z) و بردار نرمال متناظر هر نقطه دارای سه مؤلفه (nx, ny, nz) است. سطوح اصلی سازه با راستای یکی از محورهای مختصات هم‌راستا یا بسیار نزدیک به آن است، بنابراین می‌توان با تحلیل مؤلفه‌های اصلی بردار نرمال صفحات افقی و عمودی را به خوبی تفکیک کرد. ایده‌ی اصلی استفاده از جهت نرمال‌ها به عنوان معیاری برای تشخیص نوع سطوح است. در داده‌های ابرنقاط، سطوح افقی مانند کف و سقف و یا سطوح افقی مبلمان داخلی دارای

1. Principal Component Analysis
2. Eigenvalues
3. Eigenvectors
4. k-nearest neighbors

|nz| نزدیک به یک هستند، در حالی که سطوح عمودی |nz| نزدیک به صفر دارند. برای تفکیک این دو دسته از سطوح، یک آستانه تجربی برابر با ۰/۲ برای مؤلفه عمودی بردار نرمال در نظر گرفته شد. به‌منظور بررسی حساسیت روش پیشنهادی نسبت به آستانه‌ی مؤلفه‌ی عمودی بردار نرمال، مقادیر مختلف $\tau_{nz} = 0, 0.1, 0.2, 0.3$ آزمایش شدند. نتایج آزمایش نشان داد که انتخاب $\tau_{nz} = 0$ بسیار سخت‌گیرانه است، زیرا در داده‌های ابرنقاط به دلیل نویز اندازه‌گیری، خطای تخمین نرمال و ناهمواری سطح، مقدار nz برای نقاط دیوار برابر صفر نیست. از طرف دیگر افزایش بیش از حد آستانه مانند $\tau_{nz} = 0.3$، باعث ورود نقاط سطوح غیرعمودی به مجموعه‌ی کاندیدهای دیوار می‌شود. بر اساس نتایج آزمایش، $\tau_{nz} = 0.2$ تعادل مناسبی میان حفظ نقاط دیوار و جلوگیری از ورود نویز و سطوح غیراصلی فراهم کرد، بنابراین به‌عنوان مقدار نهایی در این پژوهش انتخاب شد. با مقایسه مؤلفه عمودی nz با حد آستانه تعریف شده، نقاط به سه دسته زیر تفکیک می‌شوند:

- اگر nz > threshold باشد، بردار نرمال تقریباً در راستای قائم به سمت بالا است. نقاط با این ویژگی، روی صفحات افقی مانند کف سازه قرار دارند.
- اگر nz > threshold – باشد، بردار نرمال تقریباً در راستای قائم به سمت پایین است. نقاط با این ویژگی، روی صفحات افقی مانند سقف سازه قرار دارند.
- اگر |nz| < threshold باشد، بردار نرمال تقریباً به حالت افقی قرار دارد. نقاط با این ویژگی، روی سطوح عمودی مانند دیوارها قرار دارند.

براساس این گروه‌بندی، نقاط مربوط به صفحات عمودی حفظ شده و نقاط متعلق به صفحات افقی مانند کف، سقف و هم‌چنین سطوح افقی مبلمان داخلی مانند سطح یک میز حذف می‌شوند.

۲-۳-۳- تشکیل تصویر تراکم از داده‌های ابرنقاط

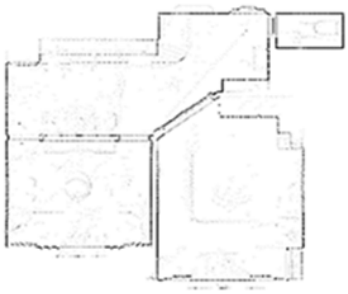
به منظور کاهش پیچیدگی محاسباتی از نظر زمان اجرا و میزان مصرف حافظه، ابرنقاط اولیه به تصویر تراکم[1] تبدیل شده و فرآیند استخراج نقشه در فضای دوبعدی تصویر انجام می‌گیرد. تمرکز بر نقاط مربوط به سطوح عمودی دیوارها که مرزهای ساختاری فضا را تعریف می‌کنند، امکان بازسازی دقیق‌تر و سریع‌تر نقشه را فراهم می‌نماید. تصویر تراکم اطلاعات توزیع نقاط را در صفحه افقی X–Y نمایش می‌دهد و مبنایی برای مراحل بعدی تحلیل داده است. برای این منظور در ابتدا، مختصات نقاط در صفحه دوبعدی 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 روابط (٦) و (٧) تعریف می‌شود:

$$\acute{x} = \frac{x - x_{min}}{x_{max} - x_{min}} * W \qquad (٦)$$

$$\acute{y} = \frac{y - y_{min}}{y_{max} - y_{min}} * H \qquad (٧)$$

که در آن H و W به ترتیب عرض و ارتفاع تصویر، $\acute{x}$ و $\acute{y}$ مختصات نرمال شده نقاط در تصویر هستند.

در این روش برای هر پیکسل تصویر دوبعدی، تعداد نقاطی که در آن پیکسل 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، تجمیع می‌شوند. نتیجه این پردازش یک نقشه تراکم است که در آن مقدار هر پیکسل نشان دهنده تراکم نقاط در آن مکان خاص است. نگاره ٤ یک نمونه تصویر تراکم حاصل از ابرنقاط را نشان می‌دهد.



نگاره ٤: نمایش تصویر تراکم

1. Density Image

۲-۳-٤- قطعه‌بندی تصاویر

در این تحقیق، فرایند قطعه‌بندی بر روی تصاویر تراکم حاصل از ابرنقاط انجام شده است. این تصاویر به دلیل نمایش واضح خطوط ساختاری دیوارها، ورودی مناسبی برای مدل‌های یادگیری عمیق در استخراج مرزهای هندسی محسوب می‌شوند. برای انجام این وظیفه، از معماری U–Net استفاده شد. مدل U–Net با بهره‌گیری از ساختار رمزگذار – رمزگشا و مسیرهای اتصال میان‌بر[1] در بازسازی جزئیات لبه‌ها و مرزهای باریک عملکرد موفقی دارد. افزون بر این، به‌منظور دستیابی به ارزیابی جامع‌تر و مقایسه عملکرد، مدل SegFormer[2] به‌عنوان یکی از شبکه‌های شناخته‌شده در قطعه‌بندی معنایی مبتنی بر ترنسفورمر *(Xie et al., 2021)*نیز مورد آزمایش و ارزیابی قرار گرفت.

شبکه عصبی U–Net یک معماری قدرتمند برای پردازش تصاویر است که به طور خاص برای وظایف قطعه‌بندی طراحی شده است. این مدل دارای دو بخش اصلی رمزگذار[3] و رمزگشا[4] است. بخش رمزگذار شامل چندین لایه کانولوشن و لایه‌های کاهش ابعاد[5] 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[6] و کانولوشن، ابعاد تصویر را به اندازه اولیه بازمی‌گرداند. هم‌چنین، U–Net از اتصالات مستقیم 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 مهم را به خوبی شناسایی کند.

یک پارامتر مهم در شبکه‌های عصبی، تابع فعال‌سازی است. توابع فعال‌سازی به طور ساده توابع انتقالی هستند که برای شناسایی روابط پیچیده بین داده‌ها استفاده می‌شوند و خروجی غیرخطی ارائه می‌دهند. بدون توابع فعال‌سازی، شبکه‌های عصبی تنها مدل‌های رگرسیون خطی خواهند بود که قادر به نمایش چندجمله‌ای‌های درجه یک هستند. درشبکه‌های عصبی عمیق، هدف اصلی لایه Dropout، کاهش بیش برازش[7] است. بیش برازش، 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.

علاوه بر معماری U–Net، در این تحقیق از مدل SegFormer نیز برای انجام قطعه‌بندی استفاده شد. ویژگی اصلی SegFormer، بهره‌گیری از backbone مبتنی بر (Mix Transformer MiT) برای استخراج ویژگی‌های چندمقیاس و استفاده از decoder مبتنی بر MLP است که با طراحی بهینه، دقت بالای قطعه‌بندی را با پیچیدگی محاسباتی پایین ترکیب می‌کند. این سازوکار موجب می‌شود مدل بتواند همزمان الگوهای محلی و سراسری را در سطح عمیق‌تری درک کند. در تصاویر تراکم حاصل از ابرنقاط، که اغلب شامل نواحی با تراکم متفاوت و الگوهای خطی باریک هستند، وجود چنین درکی از زمینه مکانی نقش مهمی در تشخیص صحیح مرز اتاق‌ها و تفکیک سطوح مجاور ایفا می‌کند. معماری SegFormer با ترکیب چندمقیاسی اطلاعات، توانایی بالایی در استخراج ویژگی‌های پایدار نسبت به تغییرات ابعادی و هندسی دارد.

1. Skip Connections
2. Simple and Efficient Design for Semantic Segmentation with Transformers
3. Encoder
4. Decoder
5. Pooling Layer
6. Upsampling
7. Overfitting

۲-۳-۵- بازنمایی برداری خروجی

هدف اصلی بازنمایی برداری، استخراج نواحی معنادار از ماسک‌های پیکسلی و بازسازی آن‌ها به صورت چندضلعی‌های برداری با گوشه‌های دقیق است. این بردارها قابلیت ذخیره‌سازی در قالب‌های استاندارد نظیر json و بهره‌برداری در کاربردهایی مانند تحلیل نقشه‌های ساختمانی را فراهم می‌کنند.

برای تبدیل خروجی پیکسلی مدل به اشکال برداری، از مجموعه‌ای از عملیات مورفولوژیکی بهره گرفته شد. در نخستین گام، عملگر فرسایش[1] به‌کار گرفته شد تا نواحی به هم چسبیده‌ی حاصل از پیش‌بینی نادرست مرزها توسط مدل به درستی از یکدیگر جدا شوند. این فرآیند با ایجاد نوارهای باریک بین نواحی مجزا، زمینه‌ی تفکیک دقیق‌تر بخش‌های مستقل محیط داخلی را فراهم می‌کند. در ادامه، از الگوریتم اجزای متصل[2] برای شناسایی و برچسب‌گذاری نواحی مستقل استفاده شد تا امکان پردازش مجزای هر ناحیه فراهم شود. سپس، به‌منظور جبران اثر انقباضی ناشی از فرسایش و بازسازی ساختار اصلی هر ناحیه، عملگر گسترش[3] بر روی هر بخش اعمال شد. در آخر، با استفاده از روش استخراج مرزها[4] مرز خارجی هر ناحیه شناسایی شد. مرزهای به دست آمده معمولاً شامل تعداد زیادی نقطه هستند که ناشی از ماهیت پله‌پله‌ی لبه‌ها در تصاویر رستری است. برای ساده‌سازی این مرزها و حذف اعوجاج‌های زائد، از الگوریتم داگلاس-پیکر[5] استفاده شد. این الگوریتم با کاهش تعداد نقاط مرزی به مجموعه‌ای از گوشه‌های کلیدی، شکل کلی چندضلعی را حفظ کرده و نمایی ساده‌تر و دقیق‌تر از ساختار هندسی ارائه می‌دهد.

در مرحله پردازش مورفولوژیک، از المان‌های ساختاری خطی برای سازگار شدن عملیات با ماهیت هندسی داده‌ها استفاده شد. از آن‌جا که دیوارهای ساختمان ماهیتی عمدتاً افقی و عمودی دارند، کرنل‌های مورفولوژی نیز به‌صورت خطی و در همین راستاها انتخاب شدند. اندازه کرنل‌ها به وضوح[6] تصویر وابسته است. در این پژوهش، خروجی U-Net دارای اندازه ثابت ۲۵۶×۲۵۶ است. ابعاد کرنل‌ها در فضای پیکسلی با این ابعاد تنظیم شد. به‌طور مشخص، در مرحله فرسایش از دو کرنل خطی افقی با ابعاد ۹*۲ و کرنل خطی عمودی با ابعاد ۲*۹ استفاده شد تا ساختارهای باریک و خطی متناظر با مرز دیوارها حفظ شده و نویزهای نقطه‌ای حذف شوند. در مرحله بازیابی شکل نیز این الگو با همان ابعاد به‌کار گرفته شد تا شکل اصلی اتاق‌ها بدون انبساط اضافی بازسازی شود. در صورت تغییر وضوح تصویر، لازم است ابعاد کرنل متناسب با ضخامت متوسط دیوار در فضای پیکسلی تنظیم شود. استفاده از کرنل‌های خطی، در مقایسه با کرنل‌های مربعی یا دایروی، موجب می‌شود که لبه‌های دیوارها صاف‌تر شده، مرز اتاق‌ها از یکدیگر جدا شوند و استخراج چندضلعی‌ها در مرحله کانتورگیری با دقت بیشتری انجام گیرد.

۳- پیاده‌سازی و نتایج

۳-۱- جزئیات پیاده‌سازی

مدل پیشنهادی در محیط Python و چارچوب pytorch پیاده‌سازی شد. تمام آزمایش‌ها بر روی یک GPU Nvidia 4060 با ۸ گیگابایت حافظه انجام شد. مدل U-NET با ساختار کلاسیک شامل مسیرهای رمزگذار و رمزگشا متقارن پیاده‌سازی شد. مدل SegFormer با استفاده از مدل پایه[7] Mix Transformer پیاده‌سازی شد. تابع Adam به 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 بیش برازش مدل، نرخ Dropout برابر ۰/۲۰ در لایه‌های میانی اعمال شد. تابع فعال‌ساز RELU

1. Erosion
2. Connected components
3. Dilation
4. Find contours
5. Douglas-peucker
6. Resolution
7. Backbone

در تمام لایه‌های کانولوشن به کار گرفته شد، در حالی که از تابع sigmoid در خروجی برای تولید نقشه استفاده شد. تصاویر ورودی به اندازه ۲۵۶*۲۵۶ پیکسل تنظیم و مقادیر پیکسل‌ها به بازه [۰,۱] نرمال‌سازی شد.

۳-۲- نتایج قطعه‌بندی

به منظور قطعه‌بندی تصاویر تراکم، مدل‌های U–Net و SegFormer آموزش داده شد. برای این منظور، تصاویر تراکم حاصل از داده‌های ابرنقاط به نسبت ۸۰٪ برای آموزش، ۱۰٪ برای اعتبارسنجی و ۱۰٪ برای آزمون تقسیم شدند. ارزیابی کمّی عملکرد مدل بر روی مجموعه آزمون با استفاده از معیارهای Dice Score و Mean IOU انجام گرفت. نتایج این ارزیابی در جدول ۱ ارائه شده است.

جدول ۱: نتایج کمّی قطعه‌بندی مدل‌های پیشنهادی

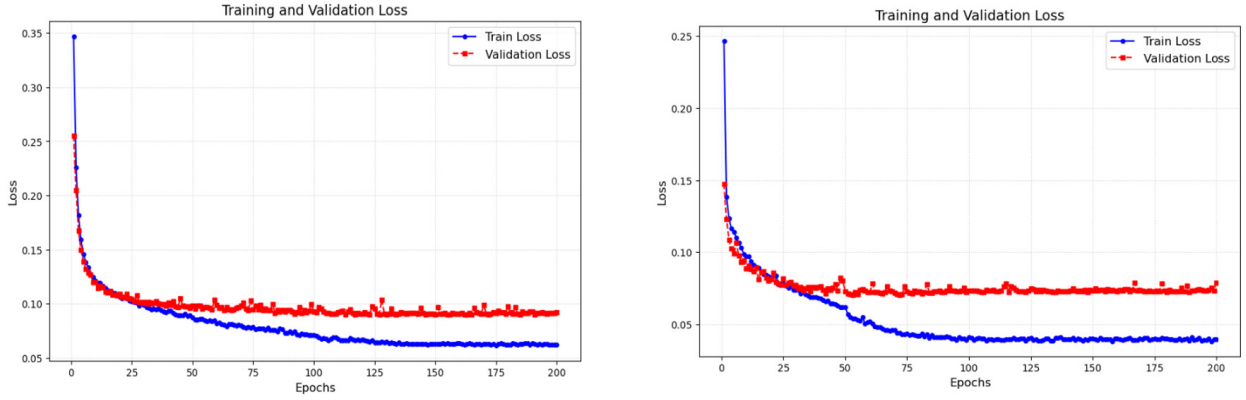
Model	Dice Score	Mean IOU
U-NET	۰/۹۷	۰/۹٤
SegFormer	۰/۹٤	۰/۹۰

جدول ۲، نمونه نتایج ارزیابی کیفی عملکرد مدل‌های قطعه‌بندی پیشنهادی را روی دیتاست Structure3d نشان می‌دهد. این نمونه‌ها شامل تصویر ورودی، برچسب واقعی و خروجی پیش‌بینی شده هستند. نتایج نشان می‌دهند که مدل U–Net در مجموعه داده ورودی عملکرد بهتری داشته و مرز فضاهای داخلی را به درستی بازسازی کرده است.

جدول ۲: نتایج کیفی قطعه‌بندی مدل‌های پیشنهادی

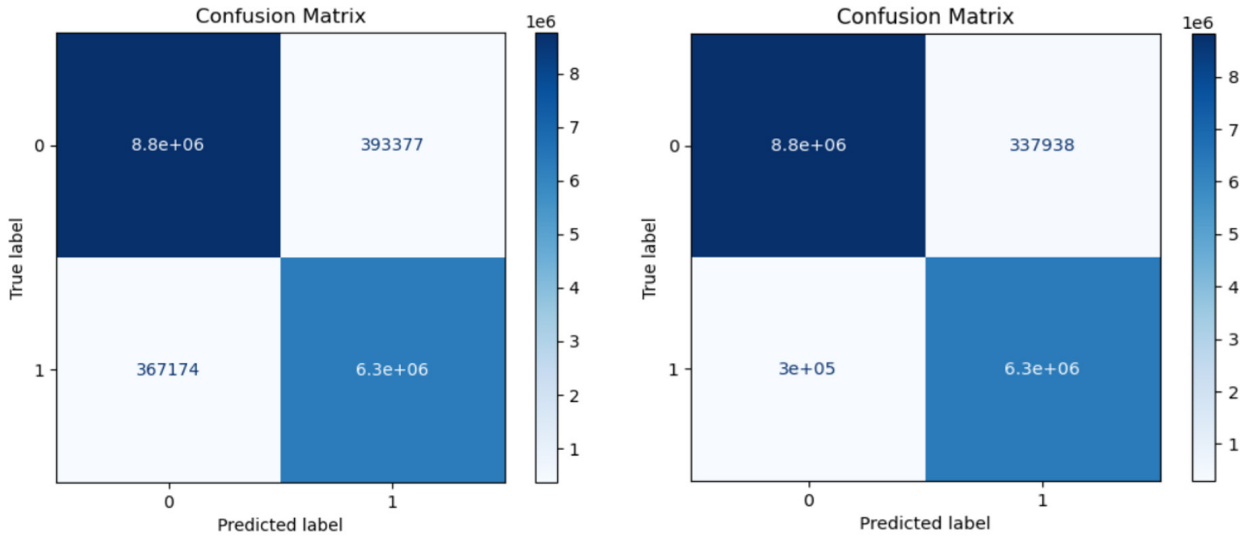
Input	Ground Truth	Output(U-NET)	Output (SegFormer)

در خصوص ارزیابی مدل‌های یادگیر، نمودار Epoch vs Loss و ماتریس درهم ریختگی اطلاعات جزئی‌تر از نتایج را به ما نشان می‌دهد. نگاره ۵ نمودار Epoch vs Loss فرآیند آموزش و نگاره ٦ تصویر confusion matrix مدل‌های U–Net و SegFormer روی نمونه‌های تست مجموعه‌داده ورودی را نشان می‌دهند. برچسب «یک» کلاس هدف شامل فضای داخل اتاق‌ها و برچسب "صفر" فضای خارج اتاق‌ها را نشان می‌دهد. بررسی‌های کمی و کیفی نشان داد که اگرچه SegFormer در استخراج زمینه‌های کلی عملکرد مناسبی دارد، اما در بازسازی خطوط باریک و مرزهای دقیق دیوارها دقت کمتری نسبت به مدل U–Net ارائه می‌کند. در نهایت، مدل U–Net به دلیل عملکرد بهتر در معیارهای شاخص به عنوان مدل نهایی انتخاب شد.



ب الف

نگاره ۵: نمودار Epoch vs Loss الف- مدل U-Net ب- مدل SegFormer



ب الف

نگاره ٦: ماتریس درهم ریختگی الف- مدل U-Net ب- مدل SegFormer

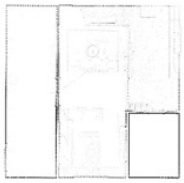
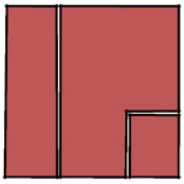
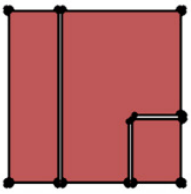
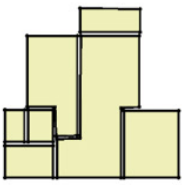
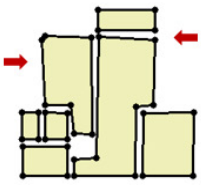
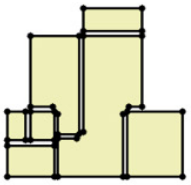
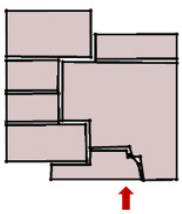
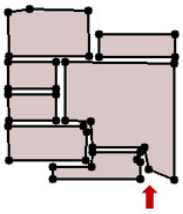
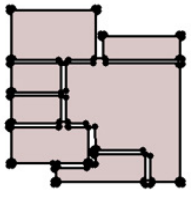
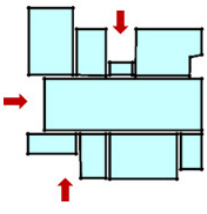
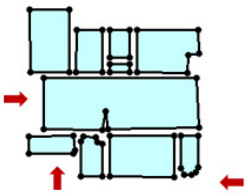
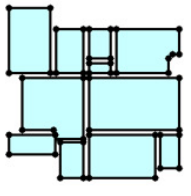
۳-۳- نتایج بازنمایی برداری نقشه دوبعدی

به منظور بازنمایی برداری نقشه‌ها از عملگرهای مورفولوژی استفاده شد. پس از تبدیل خروجی‌های پیکسل-مبنا به فرمت برداری و استخراج گوشه‌ها و مرز اتاق‌ها، عملکرد مدل با استفاده از معیارهای استاندارد Precision و Recall اندازه‌گیری شده است. این معیارها در دو سطح گوشه‌ها و اتاق‌ها محاسبه و با نتایج دو روش پیشین دیگر مقایسه شدند. نتایج ارزیابی کمی بازنمایی برداری در جدول ۳، ارائه شده است.

جدول ۳: مقایسه کمّی عملکرد روش پیشنهادی با روش‌های پیشین در بازسازی برداری نقشه‌ها

Models	Rooms Prec%	Rooms Rec%	Corner Prec%	Corners Rec%
MonteFloor	۹۶	۹۴	۸۹	۷۷
RoomFormer	۹۷/۹	۹۶/۷	۸۹/۱	۸۵/۳
Our model	۸۴/۰	۸۵/۰	۷۵/۰	۸۲/۰

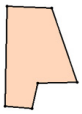
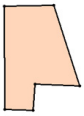
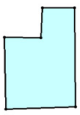
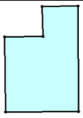
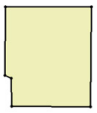
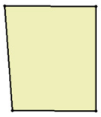
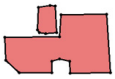
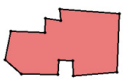
جدول ٤: مقایسه کیفی عملکرد روش پیشنهادی با مدل RoomFormer در بازسازی برداری نقشه‌ها

Density Image	RoomFormer	Ours	Ground Truth
			
			
			
			

جدول ۵: نتایج ارزیابی کمّی روش پیشنهادی روی مجموعه داده SceneCAD

Models	Rooms %Prec	Rooms %Rec	Corner %Prec	Corners %Rec
Our model	۹۹/۰	۶۴/۰	۶۰/۰	۵۷/۰

جدول ٦: نمونه نتایج ارزیابی کیفی روش پیشنهادی روی مجموعه‌داده SceneCAD

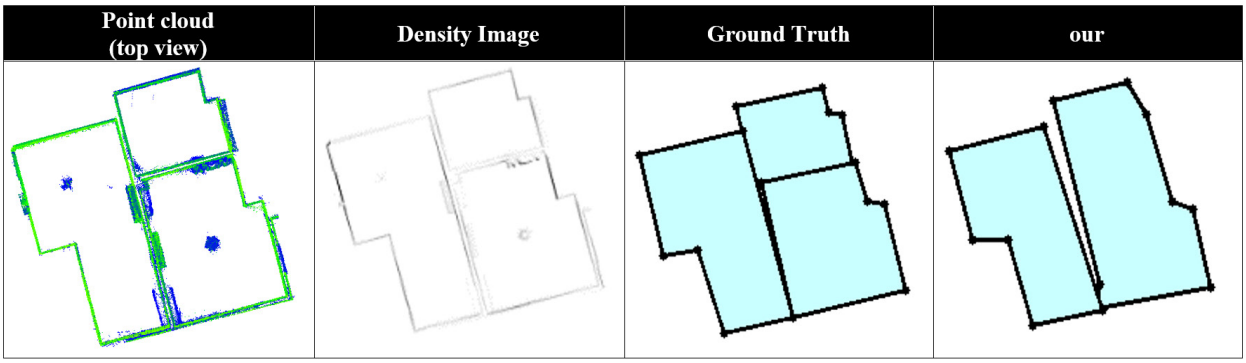
Density Image	Ours	Ground Truth
		
		
		
		

جدول ٤، نمونه نتایج ارزیابی کیفی روش پیشنهادی در بازسازی برداری نقشه را در مقایسه با مدل RoomFormer نشان می‌دهد. به دلیل عدم دسترسی به پیاده‌سازی مدل MonteFloor امکان مقایسه کیفی با این مدل وجود ندارد.

اگرچه مدل RoomFormer درارزیابی‌های کمّی برتری نسبی نشان می‌دهد، تحلیل کیفی خروجی‌ها بیانگر آن است که روش پیشنهادی قادر به تولید نقشه‌هایی با انسجام توپولوژیکی مطلوب و حفظ مرزهای ساختاری اتاق‌ها است. افزون بر این، روش پیشنهادی چارچوب یادگیری ساده‌تر دارد. در روش پیشنهادی تنها یک شبکه عصبی U–Net برای قطعه‌بندی تصاویر تراکم آموزش داده می‌شود، در حالی که مدل RoomFormer مستلزم یک فرآیند دومرحله‌ای شامل آموزش مدل‌های MaskR–CNN و ترنسفورمر است. از نظر پیچیدگی محاسباتی، روش پیشنهادی دارای ۷, ٦٩٩, ۷۱۲ پارامتر یادگیر[1] و زمان آموزش ۲ ساعت و زمان استنتاج[2] ١٤ میلی‌ثانیه است. در مقایسه، مدل RoomFormer دارای ١٧,٣٧١,٤٧٤ پارامتر یادگیر است و علاوه بر آن از مدل Mask R–CNN و backbone مبتنی بر ResNet-50 بهره می‌برد. زمان آموزش آن ۳۳ ساعت و زمان استنتاج ٦٨٩ میلی‌ثانیه است. تمامی آزمایش‌ها برای هر دو روش بر روی سخت‌افزار یکسان NVIDIA RTX 4060 انجام شده است. روش MonteFloor نیز مبتنی بر Mask R–CNN به همراه درخت مونت کارلو است. بنابراین روش پیشنهادی ضمن داشتن فرآیند ساده‌تر، به‌طور قابل‌توجهی سریع‌تر از دو روش پیشین است.

به‌منظور ارزیابی قابلیت تعمیم‌پذیری روش پیشنهادی به داده‌های واقعی، آزمایش تکمیلی روی مجموعه‌داده SceneCAD انجام شد. در این آزمایش مدل آموزش‌دیده با مجموعه‌داده Structured٣D، روی نمونه‌های واقعی مجموعه‌داده SceneCAD ارزیابی شد. نتایج ارزیابی کمّی روش پیشنهادی روی مجموعه داده SceneCAD در جدول ۵، ارائه شده است. جدول ٦، نمونه نتایج ارزیابی کیفی روش پیشنهادی روی مجموعه‌داده SceneCAD را نشان می‌دهد.

1. Learnable parameter
2. Inference time



نگاره ۷: محدودیت روش پیشنهادی

عملکرد پایین‌تر مدل در مجموعه داده SceneCAD در مقایسه با داده‌های مصنوعی Structured3D، اغلب ناشی از پیچیدگی ذاتی و ماهیت غیرساختاریافته داده‌های واقعی است. افزون بر این، فرآیند برچسب‌گذاری در مجموعه‌های داده، توسط نیروی انسانی انجام می‌شود و در برخی نمونه‌ها خطا دارد. عدم تطابق تصویر و برچسب، بر کاهش دقت مدل در مرحله‌ی ارزیابی مؤثر بوده است.

در این تحقیق، روشی نیمه‌خودکار برای استخراج نقشه دوبعدی از داده‌های ابرنقاط لیزراسکنرهای نقشه‌برداری ارائه شده است. در این روش، در مرحله نخست، خروجی مدل به صورت تصویر پیکسلی تهیه و سپس با استفاده از عملگرهای مورفولوژی به یک نمایش برداری قابل استفاده در نرم‌افزارهای طراحی به کمک کامپیوتر (نرم‌افزارهای CAD[1]) تبدیل می‌شود. در ادامه فرایند تهیه نقشه، عملیات ترسیم و بازنگری انسانی را لازم دارد. به این شیوه، از لحاظ زمانی صرفه‌جویی بسیار خوبی صورت گرفته و مراحل بعدی ترسیم نقشه به سادگی صورت می‌پذیرد. این رویکرد امکان ویرایش و تکمیل آسان نتایج عملیات اسکن زمینی، توسط مهندسان نقشه‌برداری را فراهم می‌کند. مزیت اصلی روش پیشنهادی، بر کاربرد نیمه‌اتوماتیک و قابلیت تعامل با کاربر و صرفه‌جویی در زمان پردازش دستی مورد نیاز برای تهیه نقشه از داده‌های اولیه اسکنرهای نقشه‌برداری است.

٤- محدودیت‌ها

با وجود نتایج مطلوب روش پیشنهادی، این روش همچنان با برخی محدودیت‌ها همراه است. اگرچه مدل U-Net توانایی قطعه‌بندی دیوارها با جهت‌گیری‌های مختلف از جمله دیوارهای مورب را دارد، اما در مرحله بازسازی برداری نقشه، عملگرهای مورفولوژیکی در برخی ساختارهای غیرمتعامد ممکن است عملکرد مطلوب نداشته باشند. در روش پیشنهادی پس‌پردازش مورفولوژیکی بر پایه المان‌های ساختاری خطی افقی و عمودی انجام می‌شود. این انتخاب با توجه به غالب بودن راستاهای متعامد در ساختمان‌های مسکونی انجام شده است. با این حال، در مواردی که یک دیوار مورب به‌عنوان مرز مشترک میان دو اتاق قرار گیرد، کرنل‌های افقی و عمودی ممکن است نتوانند پیوستگی هندسی این دیوار را به‌درستی حفظ یا بازسازی کنند. در نتیجه، احتمال شکستگی یا حذف بخشی از دیوار مورب در مرحله بازسازی برداری وجود دارد که می‌تواند به ادغام نادرست فضاهای مجاور یا کاهش دقت استخراج اتاق‌ها منجر شود. بنابراین، عملکرد روش پیشنهادی برای نقشه‌هایی با دیوارهای مورب یا هندسه غیرمتعامد محدود است. وجود برخی مبلمان عمودی نظیر کمدهای دیواری، کتابخانه‌ها یا پارتیشن‌های بلند می‌تواند در مرحله تولید تصویر تراکم باعث ایجاد پاسخ‌هایی مشابه دیوار شود. از آنجا که این عناصر در ابرنقاط به صورت سطوح عمودی ظاهر می‌شوند، در برخی موارد ممکن است توسط مدل به عنوان دیوار شناسایی شده و منجر به ایجاد مرزهای کاذب در نقشه دوبعدی شوند. نگاره ۷ یک نمونه محدودیت روش پیشنهادی را نشان می‌دهد.

1. Computer Aided Design

۵- نتیجه‌گیری و پیشنهادات

در این تحقیق، روش جدیدی برای استخراج نقشه دوبعدی ساختمان‌های مسکونی از داده‌های ابرنقاط ارائه شد. کارهای پیشین در این حوزه، به استخراج نقشه دوبعدی براساس تصاویر تراکم حاصل از ابرنقاط پرداخته‌اند. هدف ما از این تحقیق، ارائه روشی با پیچیدگی محاسباتی کمتر و دقت مطلوب است. برای تحقق این هدف، در مرحله اول با استفاده از الگوریتم تحلیل مؤلفه‌های اصلی، تصاویر تراکم بهینه تهیه کردیم که به خوبی لبه‌های دیوارها را نشان می‌دهند. سپس این تصاویر با استفاده از مدل U–Net قطعه‌بندی شده و درنهایت با به‌کارگیری عملگرهای مورفولوژیکی، مرزها و ساختارهای استخراج شده در قالب برداری بازنمایی می‌شوند.

در سطح پیکسلی، مدل پیشنهادی به دقت ۹۷% در معیار Dice Score و ۹٤% در معیار Mean IoU دست یافت که نشان می‌دهد مدل پیشنهادی به خوبی لبه‌ها و محدوده چندضلعی اتاق‌ها را تشخیص می‌دهد. همچنین در بازنمایی برداری نقشه، بهره‌گیری از عملگرهای مورفولوژیکی شامل فرسایش، اتصال مؤلفه‌ها، گسترش و استخراج مرزها، موجب انسجام خطوط مرزی و بازسازی دقیق هندسه اتاق‌ها می‌شود، که ثبات و کیفیت نقشه خروجی را تضمین می‌کند. نتایج کیفی جدول ٤ این موضوع را تأیید می‌کند. نتایج آزمایش‌ها نشان داد که روش پیشنهادی با برخورداری از ساختار ساده‌تر، نیاز کمتر به منابع محاسباتی و کیفیت قابل قبول در بازنمایی برداری نقشه‌ها، به عنوان جایگزینی عملی و مقرون‌به‌صرفه برای روش‌های مبتنی بر ترنسفورمر قابل طرح است.

در کارهای آتی، طراحی روش‌های سرتاسری[1] مبتنی بر یادگیری عمیق می‌تواند گام بعدی ارزشمندی در این حوزه باشد. در این رویکرد، مرحله مجزای بازنمایی برداری که در بسیاری از روش‌های موجود ضروری است، به‌صورت ضمنی در فرآیند یادگیری شبکه عصبی ادغام خواهد شد. همچنین ارزیابی عملکرد مدل روی داده‌های واقعی در مقیاس نقشه طبقه، به‌عنوان یکی از مسیرهای پژوهشی آینده مطرح است. از جمله اهداف آتی نویسندگان، تهیه و توسعه یک مجموعه‌داده واقعی از ابرنقاط ساختمان‌های مسکونی است.

1. End-to-End Approach

تعارض منافع

در این پژوهش، حامی مالی و تعارض منافع وجود ندارد.

References

1. Cai, R., Li, H., Xie, J., & Jin, X. (2022). Accurate floorplan reconstruction using geometric priors. Computers & Graphics, Vol.120, pp. 360-369. https://doi.org/10.1016/j.cag.2021.10.011
2. Chen, J., Liu, C., Wu, J., & Furukawa, Y. (2019). Floor-SP: Inverse CAD for Floorplans by Sequential Room-wise Shortest Path. International Conference on Computer Vision (ICCV), pp. 2661-2670. https://doi.org/10.48550/arXiv.1908.06702
3. Chen, J., Qian, Y., & Furukawa, Y. (2022). Heat: Holistic edge attention transformer for structured reconstruction. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,
4. Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. IEEE transactions on pattern analysis and machine intelligence, Vol. 24(5), pp. 603-619. https://doi.org/10.1109/34.1000236
5. El□Sayed, E., Abdel□Kader, R. F., Nashaat, H., & Marei, M. (2018). Plane detection in 3D point cloud using octree□balanced density down□sampling and iterative adaptive plane extraction. IET Image Processing, Vol. 12(9), pp. 1595-1605. https://doi.org/10.1049/iet-ipr.2017.1076
6. Fang, H., Lafarge, F., Pan, C., & Huang, H. (2021). Floorplan generation from 3D point clouds: A space partitioning approach. ISPRS

Journal of Photogrammetry and Remote Sensing, Vol. 175, pp. 44-55. https://doi.org/10.1016/j.isprsjprs.2021.02.012

7. Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, Vol. 24(6), pp. 381-395. https://doi.org/10.1145/358669.358692
8. Gerstweiler, G., Furlan, L., Timofeev, M., & Kaufmann, H. (2018). Extraction of structural and semantic data from 2D floor plans for interactive and immersive VR real estate exploration. Technologies, Vol. 6(4), pp. 101-128. https://doi.org/10.3390/technologies6040101
9. Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., & Bennamoun, M. (2020). Deep learning for 3d point clouds: A survey. IEEE transactions on pattern analysis and machine intelligence, Vol. 43(12), pp. 4338-4364. https://doi.org/10.1109/TPAMI.2020.3005434
10. He, Q., Li, Z., Gao, W., Chen, H., Wu, X., Cheng, X., & Lin, B. (2021). Predictive models for daylight performance of general floorplans based on CNN and GAN: A proof-of-concept study. Building and Environment, Vol. 206, pp. 108-346. https://doi.org/10.1016/j.buildenv.2021.108346
11. Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. Journal of educational psychology, Vol. 24(6), pp. 417-440. https://doi.org/10.1037/h0071325
12. Kaiming He, G. G., Piotr Dollar, Ross Girshick, . (2017). Mask R-CNN International Conference on Computer Vision (ICCV),
13. Kim, M., Lee, D., Kim, T., Oh, S., & Cho, H. (2023). Automated extraction of geometric primitives with solid lines from unstructured point clouds for creating digital buildings models. Automation in Construction, Vol. 145, pp. 1042-1046. https://doi.org/10.1016/j.autcon.2022.104642
14. Kumar, A., Masane, Y., & Dhakhwa, S. (2021). Disaster management plan for residential-cum-commercial property in Mumbai: a case study. Environment, Development and Sustainability, Vol. 23(8), pp. 12598-12614. https://doi.org/10.1007/s10668-020-01129-7
15. Li, B., Munoz, J. P., Rong, X., Chen, Q., Xiao, J., Tian, Y., Arditi, A., & Yousuf, M. (2018). Vision-based mobile indoor assistive navigation aid for blind people. IEEE transactions on mobile computing, Vol. 18(3), pp. 702-714. https://doi.org/10.1145/2736277.2741660
16. Lu, Q., & Lee, S. (2017). A semi-automatic approach to detect structural components from CAD drawings for constructing as-is BIM objects. In Computing in Civil Engineering 2017 (pp. pp. 84-91). https://doi.org/10.1061/9780784480823.011
17. McManamon, P. (2019). Lidar technologies and Systems (1st Edition ed.). Society of Photo-Optical Instrumentation Engineers (SPIE). https://doi.org/10.1117/3.2518254
18. Pizarro, P. N., Hitschfeld, N., Sipiran, I., & Saavedra, J. M. (2022). Automatic floor plan analysis and recognition. Automation in Construction, Vol. 140, pp. 104-348. https://doi.org/10.1016/j.autcon.2022.104348
19. Poux, F. (2019). The smart point cloud: Structuring 3D intelligent point data [Doctoral dissertation, Université de Liège]. https://doi.org/10.2312/diss.20202633000
20. Riveiro, B., & Lindenbergh, R. (2019). Laser scanning: an emerging technology in structural engineering (1st Edition ed.). CRC Press. https://doi.org/10.1201/9781351018869
21. Ronneberger, O., Fischer, P., & Brox, T. (2015). 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,
22. Stekovic, S., Rad, M., Fraundorfer, F., & Lepetit, V. (2021). MonteFloor: Extending MCTS for Reconstructing Accurate Large-Scale Floor Plans International Conference on Computer

Vision (ICCV), https://www.tugraz.at/index.php?id=52770

23. Su, J.-W., Tung, K.-Y., Peng, C.-H., Wonka, P., & Chu, H.-K. J. (2023). Slibo-net: Floorplan reconstruction via slicing box representation with local geometry regularization. Advances in Neural Information Processing Systems, Vol. 36, pp. 48781-48792.
24. Westoby, M., Brasington, J., Glasser, N., Hambrey, M., & Reynolds, M. (2012). 'Structure-from-Motion'photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, Vol. 179, pp. 300-314. https://doi.org/10.1016/j.geomorph.2012.08.021
25. Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P. (2021). SegFormer: Simple and efficient design for semantic segmentation with transformers. Advances in Neural Information Processing Systems, Vol. 34, pp. 12077-12090. https://doi.org/10.48550/arXiv.2105.15203
26. Yue, Y., Kontogianni, T., Schindler, K., & Engelmann, F. (2023). Connecting the dots: Floorplan reconstruction using two-level queries. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,
27. Zheng, J., Zhang, J., Li, J., Tang, R., Gao, S., & Zhou, Z. (2020). 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,

COPYRIGHTS

©2026 by the authors. Published by National Geographical Organization. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-NoDerivs 4.0 International (CC BY-ND 4.0)