نشریه علمی مهندسی فناوری اطلاعات مکانی

سال سیزدهم، شماره دوم، تابستان ۱۴۰۴

Vol.13, No.2, Summer 2025

۱۰۳ -۱۲۳

مقاله پژوهشی

ارزیابی مدل *RandLA-Net* برای بخش‌بندی معنایی ابرنقاط سه‌بعدی *LiDARHD* در محیط‌های شهری پیچیده از طریق بهینه‌سازی پیش‌پردازش

جواد سدیدی[1*]، حمزه رفیع زاده [2]، هانی رضائیان [3]

۱- دانشیار گروه سنجش از دور و سیستم‌های اطلاعات جغرافیایی، دانشکده علوم جغرافیا، دانشگاه خوارزمی

۲- دانشجوی کارشناسی ارشد سنجش از دور و سیستم‌های اطلاعات جغرافیایی، دانشکده علوم جغرافیا، دانشگاه خوارزمی

۳- استادیار گروه سنجش از دور و سیستم‌های اطلاعات جغرافیایی، دانشکده علوم جغرافیا، دانشگاه خوارزمی

تاریخ دریافت مقاله: ۱۴۰۴/۰۲/۱۲ تاریخ پذیرش مقاله: ۱۴۰۴/۰۶/۱۲

چکیده

بخش‌بندی معنایی ابر نقاط سه‌بعدی یکی از چالش‌های کلیدی در پردازش داده‌های لایدار(*LiDAR*) و کاربردهای مرتبط با بینایی کامپیوتری، رباتیک و نقشه‌برداری شهری محسوب می‌شود. در این پژوهش، عملکرد مدل یادگیری عمیق *RandLA-Net* برای بخش‌بندی معنایی ابر نقاط سه‌بعدی مورد بررسی قرار گرفته است. مجموعه داده‌ی مورد استفاده در این پژوهش، داده‌های *LiDARHD* هلند موسوم به *AHN4* بوده که شامل اطلاعات ارتفاعی دقیق از مناطق مختلف، به‌ویژه محیط‌های شهری و طبیعی، می‌باشد. برای بهبود دقت مدل، فرآیند پیش‌پردازش داده‌ها شامل حذف نویز، نرمال‌سازی و افزایش داده اعمال شد. به‌منظور سنجش عملکرد مدل، از شاخص نسبت اشتراک به اجتماع (*IoU*) استفاده شد که درصد همپوشانی بین بخش پیش‌بینی‌شده و بخش واقعی هر کلاس را نشان می‌دهد و معیاری دقیق برای ارزیابی کیفیت بخش‌بندی به شمار می‌رود. نتایج نشان داد که مدل *RandLA-Net* در بخش‌بندی کلاس‌هایی مانند زمین (*IoU*:۹۲٫۳٪) ، پوشش گیاهی(*IoU*:۹۳٫۶٪) و ساختمان(*IoU*:۹۳٫۷٪) عملکرد بالایی داشته است. با این حال، کلاس‌های پیچیده و کم‌نمونه مانند پل (*IoU*:۶۵٫۴٪) با دقت کمتری شناسایی شدند. مقایسه‌ی مدل پیاده‌سازی‌شده با روش‌های پایه مانند *PointNet++* و *KPConv* نشان می‌دهد که *RandLA-Net* توانسته است دقت بالاتر و هزینه‌ی محاسباتی کمتری را ارائه دهد. نتایج کلی این پژوهش، با کسب میانگین نسبت اشتراک به اجتماع (*mIoU*)کلی برابر با ۷۸٫۱٪ بر روی مجموعه داده*LiDARHD* ، اثربخشی یادگیری عمیق را در افزایش دقت و کارایی بخش‌بندی معنایی ابر نقاط سه‌بعدی برای کلاس‌های پرتکرار تأیید می‌کند. در عین حال، این یافته‌ها بر چالش‌های مداوم در زمینه تفکیک صحیح کلاس‌های کم‌نمونه (نادر) در محیط‌های پیچیده شهری تأکید دارد.

کلیدواژه‌ها : ابر نقاط سه‌بعدی، بخش‌بندی معنایی، یادگیری عمیق، *RandLA-Net* ، لایدار(*LiDAR*)، پیش‌پردازش.

* نویسنده مکاتبه‌کننده: تهران، خیابان شهید مفتح، نرسیده به خیابان انقلاب، پلاک ۴۳، دانشگاه خوارزمی.

تلفن: ۰۹۹۲۵۹۵۴۳۰۵

Email: jsadidikhu@khu.ac.ir

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]

[DOI: 10.66224/jgit.13.2.103]

۱- مقدمه

ابر نقاط سه‌بعدی یکی از مهم‌ترین روش‌های نمایش داده در حوزه‌های مختلفی مانند رباتیک، رانندگی خودران، واقعیت افزوده و سیستم‌های اطلاعات مکانی محسوب می‌شود. این نوع داده‌ها شامل مجموعه‌ای از نقاط در فضای سه‌بعدی هستند که مختصات هندسی و ویژگی‌های اضافی مانند شدت بازتاب، رنگ و دیگر ویژگی‌ها را شامل می‌شوند. یکی از چالش‌های اساسی در پردازش این داده‌ها، بخش‌بندی معنایی آن‌هاست که شامل تخصیص برچسب‌های معنایی به هر نقطه از ابر نقاط بر اساس ویژگی‌های هندسی و بافتی آن است[۱].

در سال‌های اخیر، یادگیری عمیق به عنوان یک ابزار قدرتمند برای بخش‌بندی معنایی ابر نقاط سه‌بعدی مطرح شده است. روش‌های مبتنی بر شبکه‌های عصبی عمیق به‌ویژه شبکه‌های عصبی پیچشی[۱](*CNN*) و شبکه‌های عصبی بازگشتی[۲](*RNN*) توانسته‌اند دقت بالایی را در این حوزه ارائه دهند. در مقایسه با روش‌های پایه مانند *PointNet++* و *KPConv* ، مدل‌های یادگیری عمیق قادر به استخراج ویژگی‌های پیچیده و ارتباطات فضایی در داده‌های سه‌بعدی هستند و در نتیجه عملکرد بهتری در دسته‌بندی اشیاء با اشکال پیچیده دارند. به‌ویژه مدل‌هایی مانند *RandLA-Net* و *PointNet* موفق به بهبود دقت و کاهش هزینه‌های محاسباتی شده‌اند. با وجود پیشرفت‌های قابل‌توجه در این زمینه، چالش‌هایی همچنان باقی مانده است. مدیریت ابر نقاط نامنظم، تأثیر تراکم داده‌ها، سطح نویز و عدم تعادل کلاس‌ها از جمله عواملی هستند که می‌توانند بر عملکرد مدل‌های یادگیری عمیق تأثیر بگذارند [۲].

بخش‌بندی معنایی ابر نقاط سه‌بعدی در سال‌های اخیر مورد توجه بسیاری از پژوهشگران قرار گرفته است. پژوهش لاندریو و سیمونوفسکی (۲۰۱۸) نشان داد که سازماندهی داده‌های سه‌بعدی می‌تواند دقت مدل‌های یادگیری ماشین را بهبود بخشد[۳]. آن‌ها یک چارچوب مبتنی بر نمودار[۳]را برای پردازش داده‌های ابر نقاط معرفی کردند که امکان استخراج ویژگی‌های ساختاری را فراهم می‌کند. در راستای بهبود دقت مدل‌های یادگیری عمیق، خان و همکاران (۲۰۲۰) شبکه چگال-اتصالی سلسله‌مراتبی کاهش‌دهنده نویز[۴](*DHDN*) را معرفی کرد که عملکرد بهتری در کاهش نویز و بهبود کیفیت داده‌های تصویری داشت [۴].

تچاپمی و همکاران (۲۰۱۷) نشان دادند که شبکه‌های عصبی مصنوعی[۵](*ANN*) می‌توانند برای برچسب‌گذاری معنایی داده‌های سه‌بعدی از حسگرها عملکرد بهتری داشته باشند[۵]. با این حال، روش‌های اولیه آن‌ها در پردازش داده‌های دارای جزئیات بالا محدودیت داشت. پژوهش وو و همکاران (۲۰۲۳) یک مدل چندمقیاسی ترکیبی مبتنی بر *HRNet*[۶] را معرفی کرد که دقت بالاتری در بخش‌بندی معنایی داشت[۶]. همچنین، لی و همکاران (۲۰۱۸) روش *PointCNN*[۷] را پیشنهاد دادند که همبستگی مکانی-محلی را در داده‌های سه‌بعدی بهتر مدیریت می‌کرد[۷].

پژوهش وو و همکاران (۲۰۱۸) بر بهبود دقت در بخش‌بندی اشیاء جاده‌ای با استفاده از مدل‌های مبتنی بر *CNN* تمرکز داشت و نشان داد که این روش‌ها می‌توانند دقت و سرعت پردازش داده‌های لایدار[۸] را افزایش دهند[۸]. وانگ و همکاران (۲۰۱۹) ماژول *EdgeConv* را برای حفظ اطلاعات مکانی نقاط معرفی کردند که می‌توانست ویژگی‌های هندسی شکل‌ها را بهتر ثبت کند[۹]. لی و همکاران (۲۰۲۲) نیز روش[۹] *WSPointNet* را ارائه کردند که امکان یادگیری ضعیف

[1] *Convolutional Neural Network*
[2] *Recurrent Neural Network*
[3] *Graph-based Framework*
[4] *Densely Connected Hierarchical Denoising Network*
[5] *Artificial Neural Network*
[6] *High-Resolution Network*
[7] *Point Convolutional Neural Network*
[8] *LiDAR*
[9] *Weighted Sampling PointNet*

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]
[DOI: 10.66224/jgit.13.2.103]

نظارت شده را برای درک بهتر صحنه‌های شهری فراهم می‌ساخت[۱۰].

پژوهش های مرتبط با بخش‌بندی معنایی ابر نقاط سه‌بعدی نشان می‌دهند که یادگیری عمیق نقش مهمی در بهبود دقت مدل‌ها ایفا کرده است. مقاله گیدون و روشه (۲۰۲۴) به ارائه‌ی یک مجموعه داده‌ی بزرگ‌مقیاس از تصاویر هوایی و داده‌های لایدار برای طبقه‌بندی گونه‌های درختی در جنگل‌های تک‌گونه‌ای پرداخته است [۱۱]. این پژوهش نشان می‌دهد که استفاده از ترکیب داده‌های لایدار و تصاویر هوایی می‌تواند دقت مدل‌های یادگیری عمیق را افزایش داده و به بهبود تفکیک گونه‌های مختلف در محیط‌های جنگلی کمک کند. در مقاله ماترون و همکاران (۲۰۲۴)، محققان بر توضیح‌پذیری مدل‌های یادگیری عمیق در تحلیل ابر نقاط تمرکز کرده‌اند[۱۲]. آن‌ها نشان داده‌اند که مدل‌های پیچیده مانند *RandLA-Net* و *KPConv* اگرچه دقت بالایی دارند، اما تفسیر خروجی آن‌ها چالش‌برانگیز است. این مطالعه به بررسی روش‌هایی برای افزایش شفافیت مدل‌ها پرداخته و اهمیت درک تصمیمات شبکه‌های عصبی در کاربردهای عملی را نشان داده است. پژوهش یان، لاو و فن (۲۰۲۵) نیز عملکرد مدل‌های یادگیری عمیق را در بخش‌بندی معنایی داده‌های ابر نقاط شهری ارزیابی کرده و نشان داده است که مدل‌هایی مانند *RandLA-Net* و *MinkowskiNet* در محیط‌های شهری، به دلیل ساختارهای پیچیده و تنوع زیاد اشیا، عملکرد متفاوتی دارند[۱۳]. این پژوهش‌ها نشان می‌دهند که برای افزایش دقت مدل‌ها، نیاز به استفاده از ترکیب داده‌های چندمنبعی وجود دارد.

تا کنون مدل *RandLA-Net* عمدتاً بر روی مجموعه داده‌های معیار مانند *Semantic3D* و *SemanticKITTI* ارزیابی شده و نشان داده که در این داده‌ها عملکرد برتری دارد[۱۴] . با این حال، داده‌های *LiDARHD* به عنوان یک مجموعه داده هوایی با تراکم بسیار بالا (حدود ۱۰ پالس در هر متر مربع) دارای کلاس‌های متفاوتی مانند پل و سازه‌های بالای سطح زمین (سازه‌های دائمی) هستند . ارزیابی *RandLA-Net* روی این داده‌های جدید تاکنون گزارش نشده است. از این‌رو، یک شکاف مهم تحقیقاتی وجود دارد که ضرورت مطالعه و ارزیابی مدل پیاده‌سازی‌شده بر روی داده‌های *LiDARHD* را توجیه می‌کند[۱۴].

این پژوهش با هدف ارزیابی و پیاده سازی مدل‌های یادگیری عمیق در بخش‌بندی معنایی ابر نقاط سه‌بعدی انجام شده است. به‌طور خاص، این مطالعه مدل *RandLA-Net* را بررسی کرده و عملکرد آن را در مقایسه با روش‌های پایه مانند *PointNet++* و *KPConv* و سایر مدل‌های یادگیری عمیق تحلیل می‌کند. علاوه بر این، تأثیر عوامل مختلف مانند تراکم داده، نویز و عدم تعادل کلاس‌ها بر دقت بخش‌بندی معنایی مورد بررسی قرار می‌گیرد .نوآوری‌های این پژوهش شامل بررسی جامع چالش‌های موجود در بخش‌بندی معنایی، ارزیابی روش‌های یادگیری عمیق بر روی مجموعه داده‌های مختلف در پردازش ابر نقاط است. همچنین، برای بهبود عملکرد مدل در کلاس‌های کم‌نمونه مانند پل و ساختار دائمی، از تکنیک‌های افزایش داده مانند چرخش، جابجایی، نویزگذاری و نمونه‌برداری مجدد استفاده شد که موجب بهبود شناسایی این کلاس‌ها و افزایش شاخص‌های اعتبارسنجی آن‌ها شد[۱۵] .

با وجود عملکرد موفق مدل *RandLA-Net* بر روی مجموعه داده‌های استاندارد، ارزیابی کارایی آن بر روی داده‌های جدید و چالش‌برانگیز *LiDARHD* کمتر مورد توجه قرار گرفته است. داده‌های *LiDARHD*، مانند مجموعه داده *AHN4* هلند، به دلیل تراکم بسیار بالا، مقیاس جغرافیایی وسیع و تنوع ساختاری پیچیده در محیط‌های شهری و طبیعی، چالش‌های منحصربه‌فردی را ایجاد می‌کنند. این چالش‌ها شامل مدیریت حجم عظیم داده و عدم توازن شدید بین کلاس‌های مختلف (مانند فراوانی کلاس‌های زمین و ساختمان در مقابل کمبود نمونه برای کلاس پل) است. این موضوع به‌ویژه

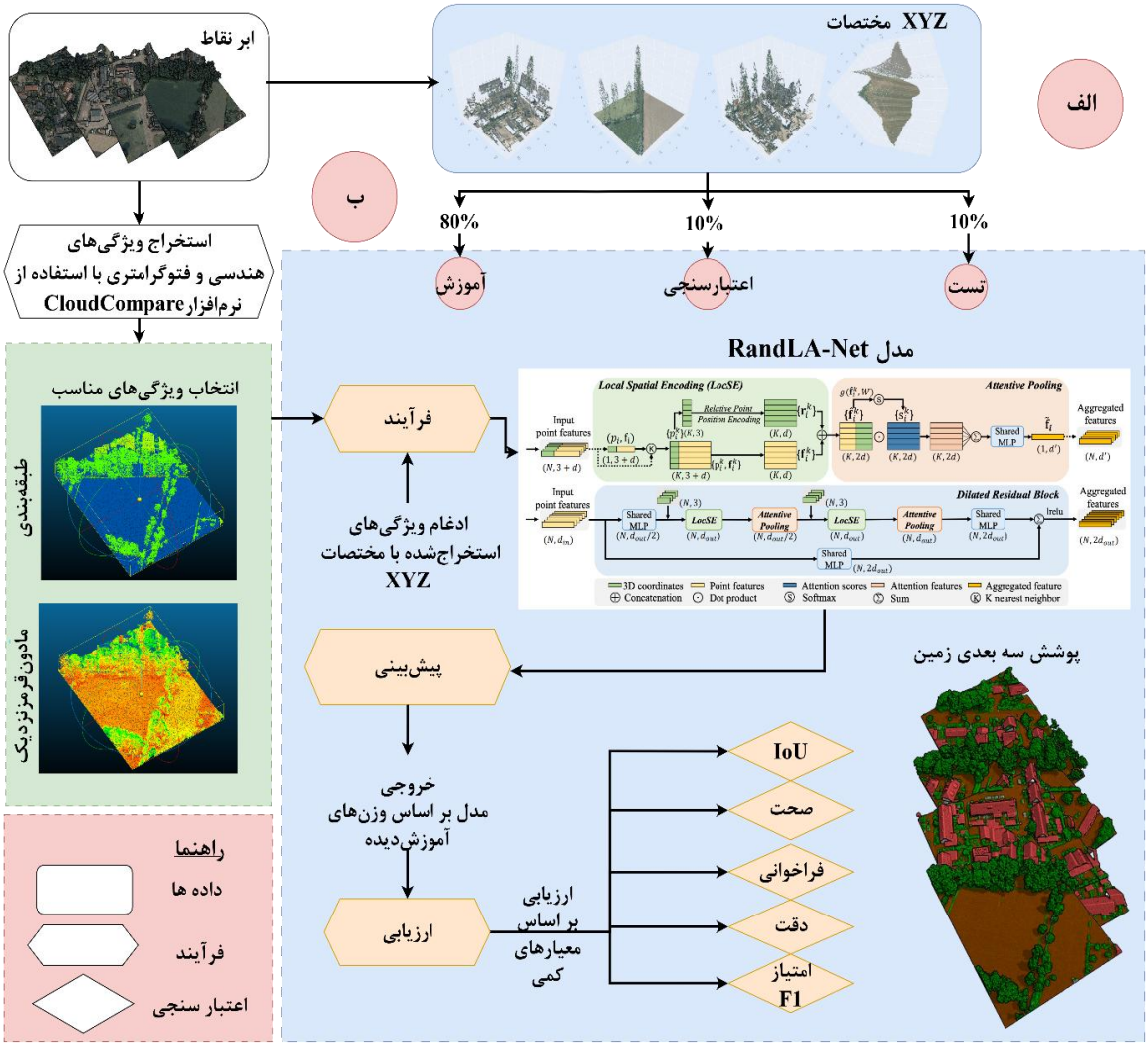
[Downloaded from jgit.kntu.ac.ir on 2026-09-16]

[DOI: 10.66224/jgit.13.2.103]

در محیط‌های پیچیده شهری و برای کلاس‌های نادر مانند پل و سازه‌های خاص نمود بیشتری دارد. بنابراین، شکاف تحقیقاتی اصلی، عدم وجود یک ارزیابی جامع از توانایی‌ها و محدودیت‌های مدل *RandLA-Net* در مواجهه با این نوع داده‌های مقیاس بزرگ و دنیای واقعی است. این پژوهش با پیاده‌سازی و ارزیابی دقیق این مدل بر روی داده‌های *LiDARHD*، تلاش می‌کند تا این شکاف را پر کرده و درک عمیق‌تری از عملکرد آن در سناریوهای کاربردی ارائه دهد.

۲- مواد و روش‌ها

در تحقیق حاضر، از داده‌های دقت بالای لایدار از محیط‌های شهری و طبیعی استفاده شده است تا عملکرد مدل *RandLA-Net* در بخش‌بندی معنایی ابرنقاط سه‌بعدی ارزیابی شود. ابتدا مجموعه داده‌های مورد استفاده و ویژگی‌های آن‌ها معرفی می‌شوند. سپس فرآیندهای پیش‌پردازش داده‌ها توضیح داده شده و مدل‌سازی با استفاده از شبکه‌های عصبی عمیق بررسی خواهد شد. در ادامه، معیارهای ارزیابی مدل ارائه خواهد شد. شکل (۱)، فرآیند استفاده از مدل *RandLA-Net* برای بخش‌بندی معنایی ابر نقطه سه‌بعدی را نمایش می‌دهد.



شکل ۱: فرآیند کامل استفاده از مدل *RandLA-Net* برای بخش‌بندی معنایی ابرهای نقطه سه‌بعدی در این پژوهش

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]
[DOI: 10.66224/jgit.13.2.103]

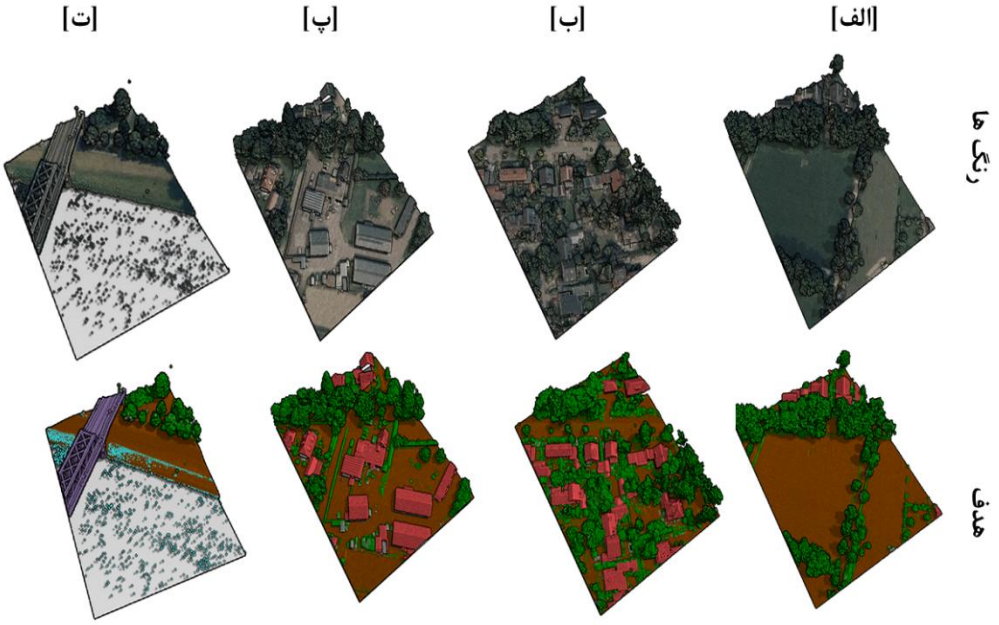
در ابتدا، تصاویر هوایی منطقه مورد نظر به عنوان ابر نقاط ورودی در نظر گرفته شده و اطلاعات مختصات سه‌بعدی (*XYZ*) از این ابر نقاط استخراج می‌شوندسپس با استفاده از نرم‌افزار *CloudCompare*، مجموعه‌ای از ویژگی‌های هندسی (نظیر شیب، انحنای سطح، تراکم نقاط) و ویژگی‌های فتوگرامتریک (مانند شدت بازتابش لیزر[1] و باند مادون‌قرمز نزدیک(*NIR*)[2] از ابرنقاط استخراج شدند. پس از تحلیل اولیه، ویژگی‌هایی که بیشترین تأثیر را در تمایز بین کلاس‌ها داشتند، از جمله ویژگی طبقه‌بندی[3] و مقادیر بازتابش در باند مادون‌قرمز نزدیک ، به‌عنوان ورودی‌های کلیدی برای مدل یادگیری انتخاب شدند. داده‌ها به سه بخش شامل داده‌های آموزش[4] ٪۸۰، اعتبارسنجی[5] ٪۱۰ و تست[6] ٪۱۰ تقسیم شده است. در مرحله بعد، ویژگی‌های استخراج‌شده ترکیب و مدل *RandLA-Net* با استفاده از لایه‌های مختلف برای استخراج ویژگی‌ها و پیش‌بینی بخش‌بندی معنایی ابر نقاط اجرا شده است. پس از آموزش مدل، پیش‌بینی‌های مربوط به داده‌های تست انجام شده و عملکرد مدل با استفاده از معیارهای ارزیابی مانند نسبت اشتراک به اجتماع (*IoU*) [7]، صحت[8]، دقت[9]، فراخوانی[10] و امتیاز *F1*-[11] مورد ارزیابی قرار می‌گیرد. در نهایت، خروجی به صورت پوشش زمین سه‌بعدی[12] ارائه می‌شود که نتیجه نهایی بخش‌بندی معنایی ابر نقاط را نشان می‌دهد. این دیاگرام به طور کامل فرآیند را از ورودی اولیه تا خروجی نهایی شرح می‌دهد و نحوه استفاده از مدل *RandLA-Net* را به وضوح توضیح می‌دهد.

۲-۱- مجموعه داده

مجموعه داده مورد استفاده در این پژوهش شامل ابر نقاط سه‌بعدی به‌دست‌آمده از مجموعه داده‌های هوایی لایدار کشور هلند است. این مجموعه داده با اطلاعات دقیق ارتفاع برای کل کشور هلند است. اطلاعات ارتفاع از هلیکوپترها و هواپیماها با استفاده از فناوری لیزر با دقت عمودی ۵ سانتی متر جمع آوری شده و یک شبکه بندی داده‌های هوایی لایدار کشور هلند[13] (*AHN*) ارائه شده است که هر گرید را به ۲۵ گرید ۱٫۲۵×۱ کیلومتر تقسیم می کند تا کار راحت تر انجام شود، که هدف آن ترسیم نقشه کل کشور با استفاده از داده های هوابرد با چگالی بالا لایدار است. این داده‌ها دارای ویژگی‌های متعددی هستند که شامل مختصات هندسی هر نقطه، شدت بازتاب، رنگ و سایر اطلاعات مرتبط با محیط می‌شود. کلاس‌های کم‌نمونه و چالش برانگیز، مانند پل‌ها، مناطق کوهستانی و محیط‌های شهری، به طور خاص بیش از حد نمونه برداری می‌شوند تا معیاری قوی برای الگوریتم های بخش‌بندی ارائه کنند [۱۶]. برای مقابله با مشکل عدم توازن داده‌ها، از تکنیک نمونه‌برداری بیش از حد[14] برای کلاس‌های کم‌تعداد[15] استفاده شد. در این روش، نمونه‌های مربوط به کلاس‌هایی که تعداد نقاط کمتری داشتند (مانند پل و سازه‌های دائمی)، در مرحله آموزش تکرار شدند تا تأثیر آن‌ها در فرآیند یادگیری مدل افزایش یابد و از سوگیری مدل به سمت کلاس‌های پرتعداد جلوگیری شود. شکل (۲) نمونه‌ای از این داده‌ها را به صورت تقسیم بندی شده برای کاربری‌های مختلف نشان می‌دهد.

[1] *Laser reflection intensity*
[2] *Near Infrared*
[3] *Classification*
[4] *Training*
[5] *Validation*
[6] *Testing*
[7] *Intersection over Union*
[8] *Accuracy*
[9] *Precision*
[10] *Recall*
[11] *F1-score*
[12] *3D Land Cover*
[13] *Actueel Hoogtebestand Nederland*
[14] *Oversampling*
[15] *Minority Classes*

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]
[DOI: 10.66224/jgit.13.2.103]



شکل۲: نمونه مجموعه داده *AHN* که در آن رنگ‌ها، قطعه ابر نقاط هوایی رنگی بوده و هدف، برچسب های منطبق با ابر نقاط هوایی می‌باشد که در کاربری‌های (الف) طبیعی و روستایی، (ب) و (پ) شهری و (ت)ساحلی نشان داده شده‌اند.

انتخاب این مجموعه داده به دلیل داشتن تنوع بالا در اشکال، سطوح و انواع مختلف اشیاء انجام شده است که امکان ارزیابی دقیق‌تر مدل پیاده‌سازی‌شده را فراهم می‌آورد. کلاس‌های این مجموعه داده شامل ۷ برچسب طبقه‌بندی نشده[1]، زمین[2]، پوشش گیاهی[3]، ساختمان[4]، آب[5]، پل[6] و سازه غیر موقت[7] است. در نهایت، مجموعه داده به سه بخش آموزش (۸۰٪) ، اعتبارسنجی (۱۰٪) و آزمایش (۱۰٪) تقسیم گردید تا عملکرد مدل در شرایط جدید مورد ارزیابی قرار گیرد.

آموزش با استفاده از *Myria3D*، یک کتابخانه یادگیری عمیق سه بعدی که به روشنی برای پردازش داده های لایدار توسعه یافته است، انجام می شود. گردش کار آموزشی بر روی *PyTorch-Lightning* ساخته شده است که آموزش توزیع شده[8] و ردیابی آزمایش‌ها[9] را ساده می کند.

۲-۲- پیش‌پردازش داده

برای بهبود کیفیت داده و افزایش کارایی مدل، مجموعه‌ای از فرآیندهای پیش‌پردازش بر روی داده‌ها انجام شده است. در مرحله نخست، داده‌ها پاک‌سازی شده و نقاط پرت، داده‌های نویزی و نقاط نامعتبر حذف شدند تا تأثیر منفی این موارد بر عملکرد مدل کاهش یابد. سپس، برای مدیریت حجم بالای داده‌ها و کاهش میزان محاسبات، فرآیند نمونه‌برداری انجام شد که به کاهش داده‌های اضافی بدون از دست دادن اطلاعات کلیدی منجر گردید .

در مرحله بعد، نرمال‌سازی داده‌ها انجام شد که طی آن مقادیر ویژگی‌ها به یک بازه استاندارد تبدیل شدند تا از

[1] *Unclassified*
[2] *Ground*
[3] *Vegetation*
[4] *Building*
[5] *Water*
[6] *Bridge*
[7] *Permanent structure*

[8] *Distributed learning*
[9] *Experiment Tracking*

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]
[DOI: 10.66224/jgit.13.2.103]

تأثیر نامتوازن برخی ویژگی‌ها بر فرآیند یادگیری جلوگیری شود. همچنین، به منظور افزایش تنوع داده‌ها و بهبود توانایی تعمیم مدل، از تکنیک‌های افزایش داده مانند چرخش، انتقال و نویزگذاری استفاده گردید. به منظور متعادل‌سازی داده‌ها که در آن کلاس‌های معنایی خاصی مانند آب، پل‌ها و سازه‌های دائمی در مقایسه با موارد رایج‌تر مانند پوشش گیاهی یا زمین نادر هستند، نمونه‌گیری هدفمند انجام شد که از کلاس‌های کم‌نمونه یا پیچیده بیش از حد نمونه‌برداری می‌کند تا از نمایش کافی در مجموعه داده نهایی اطمینان حاصل شود.

به‌طور مشخص، به منظور افزایش تنوع داده‌های آموزشی و جلوگیری از بیش‌برازش[1]، تکنیک‌های افزایش داده[2]بر روی نمونه‌های آموزشی اعمال شد. این تکنیک‌ها شامل موارد زیر بودند:

- چرخش تصادفی[3]: چرخش هر نمونه ابر نقطه حول محور عمودی[4]با زاویه‌ای تصادفی در بازه [۱۵- و ۱۵+] درجه.
- تغییر مقیاس تصادفی[5]: تغییر مقیاس هر نمونه به صورت یکنواخت با یک ضریب تصادفی در بازه[۰/۹۵ و ۱/۰۵].
- لرزش تصادفی[6]: افزودن نویز گوسی با انحراف معیار ۰/۰۱ به مختصات هر نقطه. این تکنیک‌ها به مدل کمک می‌کنند تا ویژگی‌های هندسی پایدار را مستقل از موقعیت و مقیاس بیاموزد.

این تغییرات منجر به افزایش تنوع داده‌های آموزشی و در نتیجه افزایش جزیی *IoU* مدل گردید، امری که با نتایج مشابه در ادبیات پیشین مطابقت دارد[۱۵].

[1] *Overfitting*
[2] *Data Augmentation*
[3] *Random Rotation*
[4] *Z-axis*
[5] *Random Scaling*
[6] *Random Jittering*

۲-۳- مدل‌سازی با شبکه‌های عصبی عمیق

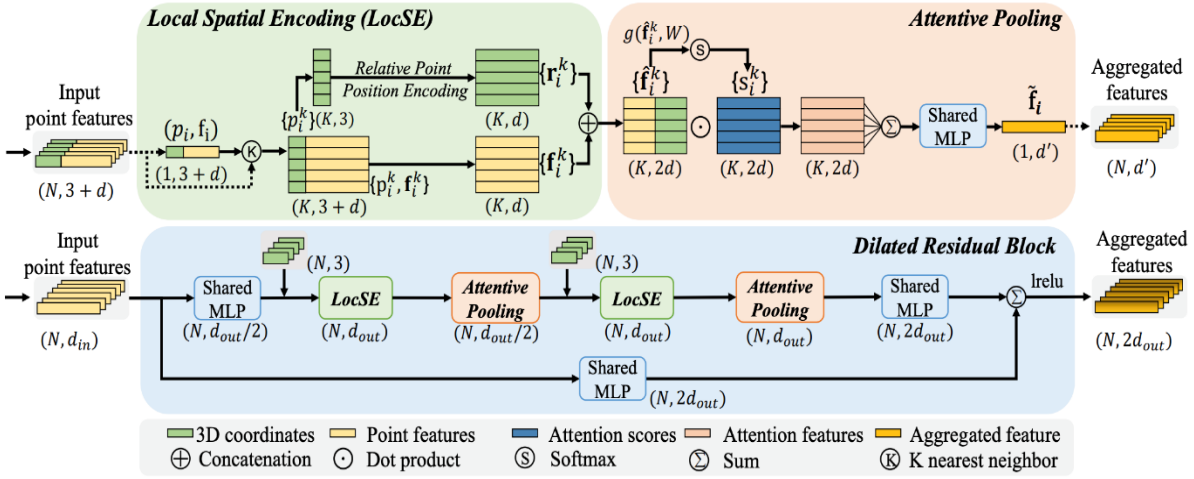
در این پژوهش، برای بخش‌بندی معنایی ابر نقاط سه‌بعدی از یک مدل مبتنی بر شبکه‌های عصبی عمیق استفاده شده است. هدف این مدل، استخراج ویژگی‌های هندسی و بافتی داده‌های سه‌بعدی و تخصیص برچسب‌های معنایی به نقاط است. مدل انتخاب‌شده *RandLA-Net* است که به دلیل عملکرد مناسب در پردازش ابر نقاط و کاهش هزینه‌های محاسباتی مورد استفاده قرار گرفته است. *RandLA-Net* یک شبکه برای بخش‌بندی معنایی ابرهای نقطه سه بعدی است که از ساختار رمزگذار-رمزگشا به شکل *U* پیروی می‌کند و شامل اجزای کلیدی زیر است: نمونه‌گیری تصادفی در لایه‌های رمزگذار برای کاهش تعداد نقاط و افزایش کارایی در ابرهای بزرگ، ماژول تجمع ویژگی‌های محلی برای ضبط اطلاعات هندسی و متنی با رویکردی سبک‌وزن،

پرسپترون‌های چندلایه[7](*MLP*) برای یادگیری ویژگی‌های نقطه‌ای با ابعاد بالاتر، اتصالات پرش برای حفظ اطلاعات مکانی در فرآیند نمونه‌گیری، و درون‌یابی نزدیکترین همسایه در فاز رمزگشا برای بازیابی وضوح ابر نقطه ورودی اصلی [۱۴].

شکل(۳) فلوچارت کلی روش آموزش مدل شبکه *RandLA-Net* برای بخش‌بندی معنایی ابرهای نقطه سه‌بعدی را نشان می‌دهد. *RandLA-Net* یک معماری کارآمد مبتنی بر شبکه‌های عصبی پیچشی است که به طور خاص برای بخش‌بندی معنایی ابر نقاط در مقیاس بزرگ طراحی شده است. برخلاف روش‌هایی که به پیش‌پردازش‌های پیچیده یا نمونه‌برداری‌های غیراستاندارد نیاز دارند، این مدل از یک معماری تک‌مسیره[8] و کارآمد بهره می‌برد.

[7] *Multi Layer Perceptrons*
[8] *Single-path*

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]
[DOI: 10.66224/jgit.13.2.103]



شکل ۳: فلوچارت روش آموزش مدل شبکه *RandLA-Net* برای بخش‌بندی معنایی ابرهای نقطه سه‌بعدی[۱۴]

دو جزء کلیدی در معماری *RandLA-Net* عبارتند از:

۱) نمونه‌برداری تصادفی[1]: برای مدیریت ابر نقاط با مقیاس بزرگ (میلیون‌ها نقطه)، *RandLA-Net* به جای استفاده از روش‌های نمونه‌برداری پرهزینه مانند[2] (*KNN*) یا[3] (*FPS*)، از نمونه‌برداری کاملاً تصادفی استفاده می‌کند. این رویکرد به طرز چشمگیری هزینه‌های محاسباتی را کاهش می‌دهد و امکان پردازش سریع داده‌ها را فراهم می‌سازد.

۲) ماژول تجمیع ویژگی محلی[4] (*LFA*): برای جبران از دست رفتن اطلاعات ساختاری ناشی از نمونه‌برداری تصادفی، از ماژول *LFA* استفاده می‌شود. این ماژول یک واحد قدرتمند برای استخراج و تجمیع هوشمندانه ویژگی‌های هندسی از همسایگی هر نقطه است. فرآیند کار *LFA* شامل چندین مرحله است:

– یافتن همسایگان: برای هر نقطه، نزدیک‌ترین همسایگان آن با استفاده از الگوریتم(*KNN*) شناسایی می‌شوند.

– کدگذاری موقعیت نسبی: موقعیت نسبی هر نقطه همسایه نسبت به نقطه مرکزی محاسبه و کدگذاری می‌شود. این کار به مدل کمک می‌کند تا روابط فضایی و ساختار هندسی محلی را درک کند.

– تجمیع ویژگی‌ها: ویژگی‌های هر نقطه با اطلاعات موقعیت نسبی آن ترکیب شده و سپس از طریق یک پرسپترون چندلایه (*MLP*) مشترک عبور داده می‌شود تا ویژگی‌های سطح بالا استخراج شوند.

– ادغام مبتنی بر توجه[5]: در نهایت، یک مکانیزم توجه[6] به صورت خودکار به ویژگی‌های مهم‌تر وزن بیشتری اختصاص می‌دهد و آن‌ها را برای تولید یک بردار ویژگی نهایی برای نقطه مرکزی، تجمیع می‌کند. این رویکرد تضمین می‌کند که اطلاعات کلیدی همسایگی حفظ شود.

معماری کلی مدل از یک ساختار رمزگذار-رمزگشا[7] پیروی می‌کند. در بخش رمزگذار، چندین لایه *LFA* به همراه نمونه‌برداری تصادفی به صورت متوالی اعمال می‌شوند تا به تدریج وضوح مکانی ابر نقاط کاهش یافته و یک نمایش غنی از ویژگی‌ها در ابعاد پایین‌تر استخراج شود. در بخش رمزگشا، ویژگی‌های

[1] *Random Sampling*
[2] *K-Nearest Neighbors*
[2] *Farthest Point Sampling*
[4] *Local Feature Aggregation*
[5] *Attention-based Pooling*
[6] *Attention*
[7] *Encoder-Decoder*

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]
[DOI: 10.66224/jgit.13.2.103]

استخراج‌شده با استفاده از درون‌یابی[1] به ابعاد بالاتر بازگردانده شده و با ویژگی‌های استخراج‌شده از مراحل متناظر در رمزگذار ترکیب می‌شوند تا بخش‌بندی دقیق در وضوح اصلی ابر نقاط حاصل شود.

۲-۴- ارزیابی مدل

پس از آموزش مدل، عملکرد آن با استفاده از معیارهای مختلف ارزیابی شده است. در ارزیابی عملکرد یک مدل طبقه‌بندی، پنج معیار مهم مورد استفاده قرار می‌گیرد که هر یک معنا و کاربرد خاص خود را دارند. در ادامه، توضیح هر یک از این معیارها به همراه رابطه ریاضی مربوطه آورده شده است. به‌منظور ارزیابی دقت و کارایی الگوریتم‌ها در پیش‌بینی، از جمله شاخص‌هایی که مورد بررسی قرار گرفته‌اند شامل: صحت (رابطه(۱))، دقت (رابطه(۲))، فراخوانی(رابطه(۳))، امتیاز-$F1$ (رابطه(۴)) و میانگین نسبت اشتراک به اجتماع[2] ($mIoU$) و (IoU) (رابطه(۵) و (۶)) هستند که روابط ریاضی هر کدام به شرح زیر است:

رابطه(۱)

$$Accuracy = \frac{TPc + TNc}{TPc + TNc + FPc + FNc}$$

رابطه(۲)

$$Precision = \frac{TPc}{TPc+FPc}$$

رابطه(۳)

$$Recall = \frac{TPc}{TPc + FNc}$$

رابطه(۴)

$$F1 - score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

رابطه(۵)

$$IoU = \frac{TPc}{TPc + FPc + FNc}$$

رابطه(۶)

$$Mean\ IoU = IoU_i \sum_{i=1}^{c} \frac{1}{C}$$

C = کلاس

TP = الگوریتم نمونه را در دسته مثبت طبقه‌بندی کرده و نمونه هم مثبت است.

TN = الگوریتم نمونه را در دسته منفی طبقه‌بندی کرده و نمونه هم منفی است.

FP = الگوریتم نمونه را در دسته مثبت طبقه‌بندی کرده اما نمونه منفی است.

FN = الگوریتم نمونه را در دسته منفی طبقه‌بندی کرده اما نمونه مثبت است.

۲-۵- تنظیم هایپرپارامترها:

جدول (۱) تنظیمات به‌دست‌آمده برای *RandLA-Net* را نمایش می‌دهد. معیار انتخاب این پارامترها بر اساس تاثیر آنها بر زمان اجرای الگوریتم و بهبود احتمالی معیارهای ارزیابی انتخاب شده‌اند. جدول ارائه‌شده شامل لیستی از هایپرپارامترهای کلیدی است که به طور خاص برای عملکرد *RandLA-Net* انتخاب شده‌اند. هر هایپرپارامتر با توصیفی خلاصه و محدوده تست‌شده مشخص شده است. این جدول نه تنها به عنوان یک راهنمای عملی برای تنظیم هایپرپارامترها عمل می‌کند، بلکه به تصمیم‌گیری‌های آگاهانه‌تر در مورد تأثیر هر پارامتر بر عملکرد مدل کمک می‌کند. به عنوان مثال، تنظیم پارامتر نرخ یادگیری[3] به صورت پیوسته به مقدار دقیق‌تری منجر شده است که در روش‌های سنتی مانند جستجوی شبکه‌ای[4] ممکن است نادیده گرفته شود.

[1] *Interpolation*
[2] *Mean Intersection over Union*
[3] *Learning Rate*
[4] *Grid Search*

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]
[DOI: 10.66224/jgit.13.2.103]

جدول۱: مربوط به هاپرپارامترهای تنظیم شده در روش *RandLA-Net*

هایپرپارامتر	توصیف خلاصه	محدوده تست شده	مقدار
اندازه دسته[1]	تعداد نمونه‌هایی که در هر بار آموزش به شبکه داده می‌شود.	[۸ و ۶۴]	۱۶
نرخ یادگیری[2]	نرخ یادگیری برای به‌روزرسانی وزن‌های شبکه.	[۰٫۰۱ و ۰٫۰۰۰۱]	۰.۰۰۲
تعداد همسایه‌ها[3] *(K)*	تعداد نقاط همسایه‌ای که در ماژول تجمع ویژگی‌های محلی استفاده می‌شود.	[۸ و ۶۴]	۳۲
نسبت نمونه‌گیری[4]	نسبت نمونه‌گیری تصادفی در هر لایه رمزگذار برای کاهش تعداد نقاط.	[۰٫۱ و ۱٫۰]	۰.۳
لایه‌های *MLP*[5]	تعداد لایه‌های پرسپترون چندلایه *(MLP)*برای استخراج ویژگی‌های نقطه‌ای.	[۵ و ۱]	۳
واحدهای مخفی *MLP*[6]	تعداد یکان‌های پنهان در هر لایه *MLP*.	[۲۵۶ و ۳۲]	۲۵۶
تکرار[7]	تعداد دوره‌های آموزش .	[۲۰۰ و ۱۰۰]	۱۶۱
بهینه‌ساز[8]	الگوریتم بهینه‌سازی برای آموزش شبکه.	*[SGD,[9] AdamW, Adam, RMSprop]*	*Adam*
تابع فعال‌سازی[10]	نوع تابع فعال‌سازی استفاده‌شده در لایه‌های *MLP*.	*[ReLU, LeakyReLU, ELU]*	*ReLU*
تابع زیان[11]	تابع زیان برای محاسبه خطای بین پیش‌بینی و برچسب واقعی.	*[Cross-Entropy, Focal Loss, Dice Loss]*	*Cross-Entropy*

۳- نتایج

در این فصل، نتایج به‌دست‌آمده از اجرای مدل بر روی داده‌های ابر نقاط سه‌بعدی ارائه شده و عملکرد آن در بخش‌بندی معنایی کلاس‌های مختلف مورد تحلیل قرار می‌گیرد.

۳-۱- عملکرد مدل

برای اثبات کارایی مدل پیاده‌سازی‌شده، عملکرد آن با دو روش پایه *PointNet++* و *KPConv* مقایسه شد. نتایج نشان می‌دهد که *RandLA-Net* به‌طور قابل ملاحظه‌ای بهتر عمل می‌کند. به‌عنوان نمونه، در مجموعه *Semantic3D* مدل *RandLA-Net* به *mIoU*

[1] *Batch Size*
[2] *Learning Rate*
[3] *Number of Neighbors (k)*
[4] *Sampling ratio*
[5] *MLP Layers*
[6] *MLP Hidden Units*
[7] *Epoch*
[8] *Optimizer*
[9] *Stochastic Gradient Descent*
[10] *Activation Function*
[11] *Loss Function*

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]
[DOI: 10.66224/jgit.13.2.103]

برابر ٪۷۸٫۱ دست یافته در حالی که *KPConv* حدود ٪۷۴٫۶ و *PointNet++* تنها ٪۲۰٫۱ کسب کرده‌اند. نتایج مقایسه مذکور در جدول(۲) آورده شده است:

جدول ۲: مربوط مقایسه روش *RandLA-Net* دو روش پایه *PointNet++* و *KPConv*

روش	*mIoU (%)*
RandLA-Net	۷۸٫۱
KPConv [۱۷]	۷۴٫۶
PointNet++ [۱۸]	۲۰٫۱

همان‌طور که مشاهده می‌شود، مدل *RandLA-Net* در دو مورد نخست برتری معناداری دارد. مقادیر بالا از منابع معتبر بازیابی شده است و بیانگر بهبود نسبی مدل پیاده‌سازی‌شده است.

شکل (۴) مجموعه نمودارهای عملکرد مدل را در شناسایی و طبقه‌بندی کلاس‌های مختلف بر اساس پنج معیار ارزیابی شامل امتیاز-*F1*، فراخوانی، دقت ، صحت و *IoU* نمایش می‌دهد. نتایج نشان می‌دهد که عملکرد مدل بسته به کلاس‌های مختلف، تغییرات محسوسی دارد. در نمودار اول، مربوط به داده‌های طبقه‌بندی نشده ، عملکرد مدل ضعیف‌تر از سایر کلاس‌ها است. مقدار امتیاز -*F1* برابر با ۶۴٫۴ است که نسبت به سایر کلاس‌ها بسیار پایین‌تر است. همچنین، مقدار فراخوانی برابر با ۵۴٫۹ بوده که نشان‌دهنده تعداد زیادی از نمونه‌های این کلاس است که مدل قادر به شناسایی آن‌ها نبوده است. مقدار دقت این کلاس، یعنی ۷۷٫۸، بالاتر از سایر معیارها است، اما همچنان در مقایسه با سایر کلاس‌ها در سطح پایینی قرار دارد. مقدار *IoU* نیز در این کلاس تنها ۴۷٫۸ است که نشان می‌دهد مدل در تعیین دقیق محدوده این کلاس عملکرد ضعیفی داشته است.

در مقابل، داده‌های زمین دارای بالاترین عملکرد در بین تمامی دسته‌ها هستند. مقدار امتیاز -*F1* برابر ۹۵٫۸ بوده و فراخوانی و صحت هر دو مقدار ۹۷٫۷ را نشان می‌دهند که بیانگر دقت بسیار بالای مدل در این بخش است. همچنین، مقدار دقت برابر با ۹۳٫۸ است که نشان می‌دهد مدل توانسته درصد بسیار زیادی از نمونه‌های این کلاس را به درستی شناسایی کند. مقدار *IoU* در این دسته نیز ۹۲٫۳ است که یکی از بالاترین مقادیر در بین تمام کلاس‌ها محسوب می‌شود.

برای کلاس‌های خاص، عملکرد مدل تفاوت‌هایی را نشان می‌دهد. در کلاس پوشش گیاهی ، مقدار امتیاز -*F1* برابر با ۹۶٫۸ است، درحالی‌که فراخوانی مقدار ۹۵٫۶ را دارد، که نسبتاً کمتر از دقت (۹۸) است. این موضوع نشان می‌دهد که مدل در شناسایی برخی از نمونه‌های این کلاس دچار مشکل شده است اما پیش‌بینی‌های انجام‌شده به میزان بالایی صحیح بوده‌اند. مقدار صحت در این کلاس ۹۵٫۶ است که نشان‌دهنده عملکرد کلی قابل قبول است، درحالی‌که مقدار *IoU* برابر با ۹۳٫۶ نشان می‌دهد که همپوشانی مدل در این کلاس نسبتاً مناسب است.

کلاس ساختمان عملکردی مشابه پوشش گیاهی دارد، اما مقادیر فراخوانی و *IoU* اندکی بالاتر هستند (۹۵٫۰ و ۹۳٫۷). مقدار دقت نیز برابر ۹۶٫۲ بوده که نشان می‌دهد مدل در این دسته نیز تعداد بسیار کمی از نمونه‌ها را به اشتباه برچسب‌گذاری کرده است. مقدار صحت در این دسته برابر ۹۳٫۷ است که نشان‌دهنده عملکرد پایدار مدل در این کلاس است.

کلاس آب یکی از بهترین عملکردها را در میان کلاس‌های خاص دارد. مقدار امتیاز -*F1* در این کلاس ۹۴٫۸ است که نزدیک به مقدار کلاس زمین است. مقدار دقت برابر با ۹۷٫۱ بوده که بالاترین مقدار در بین تمامی کلاس‌ها است، به این معنی که مدل تقریباً تمامی نمونه‌های تشخیص داده‌شده در این کلاس را درست برچسب‌گذاری کرده است. مقدار فراخوانی و صحت هر دو برابر ۹۲٫۶ هستند که نشان می‌دهد مدل توانسته اکثر نمونه‌های این کلاس را به درستی شناسایی کند. مقدار *IoU*نیز ۹۰٫۳ است که همپوشانی بالای مدل را در این دسته نشان می‌هد.

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]

[DOI: 10.66224/jgit.13.2.103]

شکل ۴: نمودار ستونی عملکرد شاخص‌های اعتبارسنجی در هر کلاس و میانگین کلی

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]

[DOI: 10.66224/jgit.13.2.103]

مقدار امتیاز-*F1* برای کلاس پل برابر با ۷۹ است، درحالی‌که مقدار دقت در این کلاس ۷۹٫۳ بوده که نشان می‌دهد مدل در پیش‌بینی‌های انجام‌شده در این دسته دقت بالایی نداشته است. مقدار فراخوانی برابر با ۷۸٫۶ است که اندکی پایین‌تر از دقت است. مقدار صحت نیز ۹۶٫۱ بوده که نشان‌دهنده عملکرد قوی مدل است. مقدار *IoU* در این کلاس ۶۵٫۴ است که نسبت به سایر معیارهای این دسته کمی پایین‌تر است.

کلاس ساختار دائمی در مقایسه با سایر کلاس‌ها عملکرد ضعیف‌تری دارد. مقدار امتیاز -*F1* برابر ۷۷٫۷ است، اما مقدار فراخوانی برابر ۷۶٫۶ است که نشان می‌دهد مدل تعدادی از نمونه‌های این کلاس را نتوانسته شناسایی کند. همچنین مقدار دقت برابر ۷۸٫۹ است که نسبتاً بالا ارزیابی نمی‌شود. درنهایت، مقدار *IoU* در این دسته برابر با ۶۳٫۷ است که پایین‌ترین مقدار در بین کلاس‌های خاص محسوب می‌شود.

در نهایت، مقادیر میانگین کلان[1] نشان می‌دهد که عملکرد کلی مدل مطلوب است اما برخی چالش‌ها در شناسایی برخی کلاس‌ها همچنان وجود دارد. مقدار امتیاز -*F1* برابر با ۸۶٫۲ بوده که نشان‌دهنده میانگین کلی عملکرد مدل است. مقدار فراخوانی برابر ۸۴٫۲ بوده که نسبت به سایر معیارها کمی پایین‌تر است، درحالی‌که مقدار دقت برابر ۸۸٫۷ نشان‌دهنده دقت بالا در پیش‌بینی نمونه‌های درست است. مقدار صحت نیز ۸۶٫۷ بوده که نشان می‌دهد مدل به‌طور کلی عملکرد نسبتاً دقیقی در پیش‌بینی کلاس‌ها داشته است، اما همچنان در برخی موارد چالش‌هایی وجود دارد. مقدار *IoU* برابر با ۷۸٫۱ است که نشان می‌دهد همپوشانی کلی مدل در تمام کلاس‌ها در سطح قابل قبولی قرار دارد.

به‌طور کلی نتایج کیفی عملکرد مدل در بخش‌بندی کلاس‌های مختلف، در شکل‌های (۵) تا (۱۱) نمایش داده شده است. به منظور ارزیابی قابلیت تعمیم‌پذیری[2] مدل، مجموعه‌ای از نتایج در قالب چهار زیرشکل (الف) تا (ت) ارائه گردیده است که نمونه‌هایی از محیط‌های مختلف با کاربری‌های متنوع (شهری، روستایی، ساحلی و طبیعی) را در بر می‌گیرد. در رویکرد بصری‌سازی اتخاذشده، نقاط داده‌ای متعلق به کلاس هدف، با یک گرادیان رنگی از زرد (نشان‌دهنده حداقل تراکم یا اطمینان) به قرمز (نشان‌دهنده حداکثر تراکم یا اطمینان) نمایش داده شده‌اند، در حالی که سایر کلاس‌ها برای ایجاد تمایز بصری، به صورت یکنواخت با رنگ آبی مشخص گردیده‌اند. این استراتژی نمایش، امکان تحلیل کیفی عملکرد مدل و شناسایی نقاط قوت و ضعف آن در تشخیص ساختارهای هندسی مختلف را فراهم می‌آورد.

همان‌طور که در تصاویر شکل (۵) قابل مشاهده است، مدل با موفقیت توانسته است پوشش گیاهی با ساختارهای پیچیده و نامنظم، از جمله تاج درختان و بوته‌ها را در محیط‌های مختلف شناسایی کند. دقت بالای مدل در این کلاس (*IoU*: ۹۳٫۶٪) نشان‌دهنده توانایی آن در استخراج ویژگی‌های بافتی و هندسی غیریکنواخت گیاهان است.

تصاویر شکل (۶) عملکرد عالی مدل در تفکیک کلاس "زمین" را نمایش می‌دهند که شامل سطوح هموار و پست‌ترین نقاط در ابر نقاط است. همانطور که در نتایج کمی نیز تأیید شد (*IoU*: ۹۲٫۳٪)، مدل به خوبی سطوح پیوسته زمین را از سایر عوارض مانند ساختمان‌ها و گیاهان متمایز می‌کند.

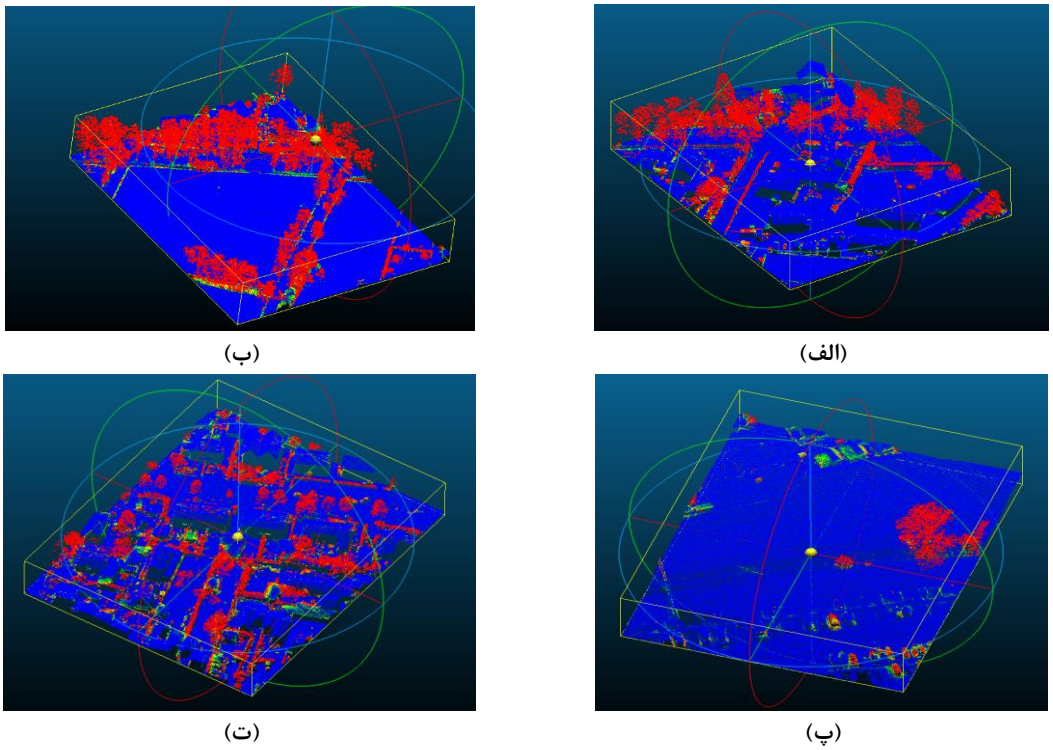
تصاویر شکل (۷)، دقت بالای مدل در شناسایی ساختمان‌ها با اشکال و ابعاد گوناگون را نمایش می‌دهد. مدل با تکیه بر ویژگی‌های هندسی منظم مانند سطوح مسطح (سقف‌ها) و لبه‌های صاف، توانسته است این کلاس را با دقت (*IoU*: ۹۳٫۷٪) بخش‌بندی کند.

[1] *Macro Average*

[2] *generalization*

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]

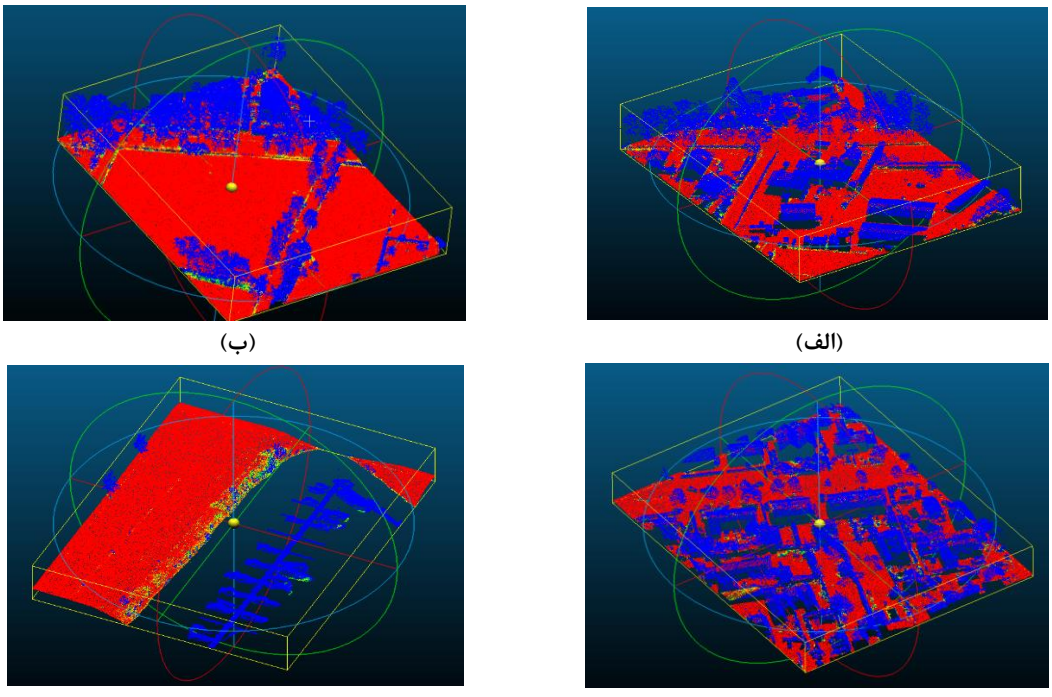
[DOI: 10.66224/jgit.13.2.103]



(الف) (ب)

(پ) (ت)

شکل ۵: نمونه‌ای از نتایج بخش‌بندی مدل برای کلاس "پوشش گیاهی"در کاربری‌های: الف) روستایی، ب) روستایی و طبیعی، پ) ساحلی، ت) شهری



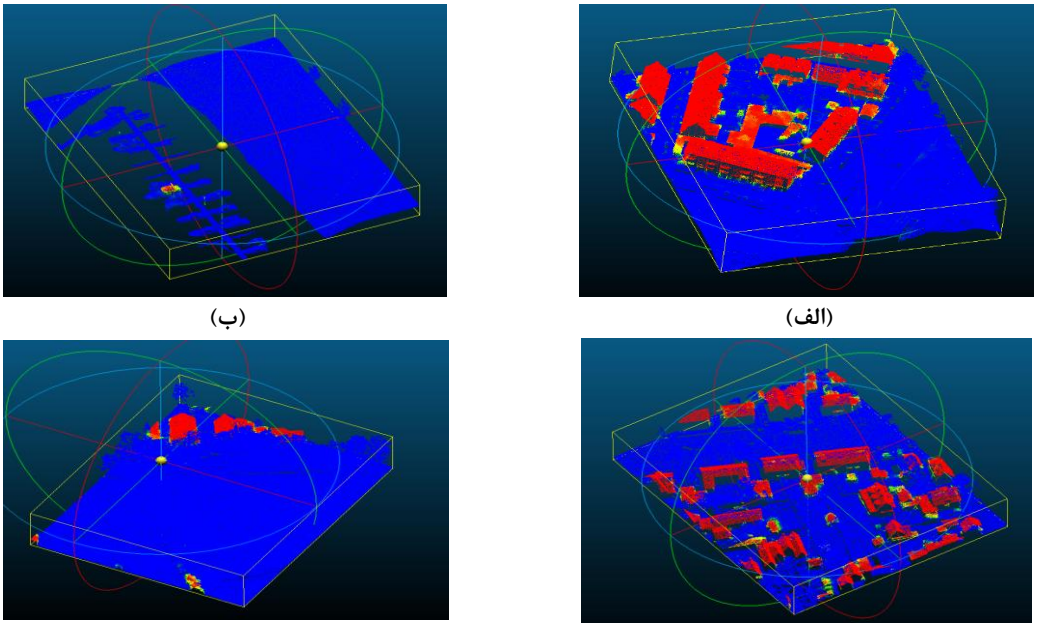
(الف) (ب)

(پ) (ت)

شکل ۶: نمونه‌ای از بخش‌بندی مدل برای کلاس "زمین" در کاربری‌های: الف) روستایی، ب) روستایی و طبیعی، پ) شهری، ت) ساحلی

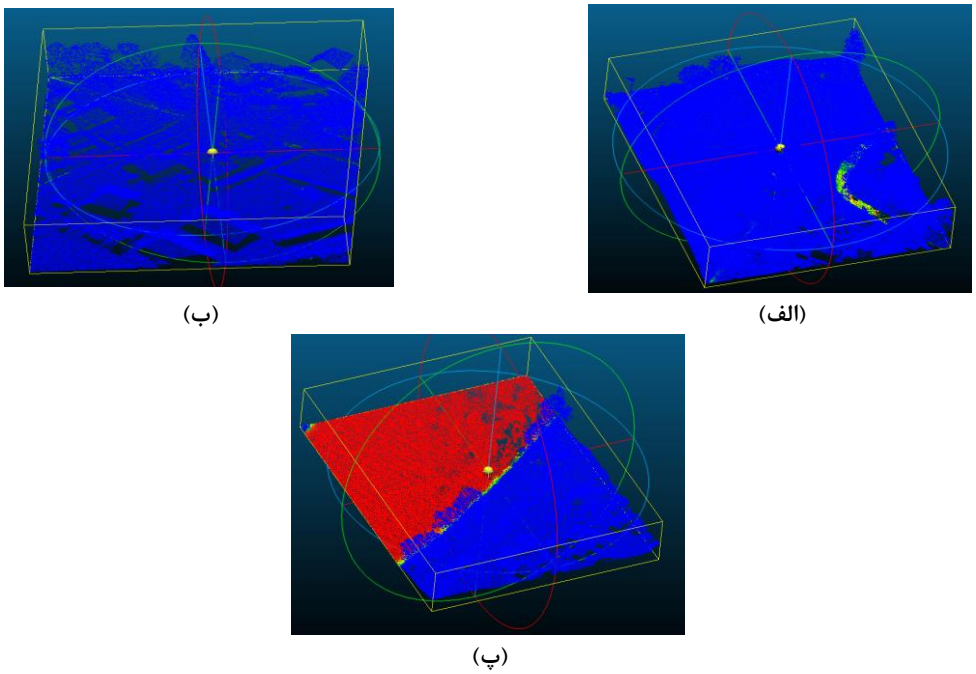
[Downloaded from jgit.kntu.ac.ir on 2026-09-16]

[DOI: 10.66224/jgit.13.2.103]



شکل ۷: نمونه‌ای از بخش‌بندی مدل برای کلاس "ساختمان" در کاربری‌های: الف) شهری، ب) ساحلی، پ) شهری، ت) روستایی

همانطور که در تصاویر شکل (۸) قابل مشاهده است، مدل سطوح آبی را که معمولاً به دلیل جذب بالای امواج لیزر، تراکم نقاط بسیار پایینی دارند، با دقت بالایی شناسایی کرده است. تشخیص صحیح این کلاس (*IoU*: ۹۰٫۳٪) نشان می‌دهد که مدل قادر به یادگیری ویژگی‌های خاص بازتابی و ارتفاعی سطوح آبی است.

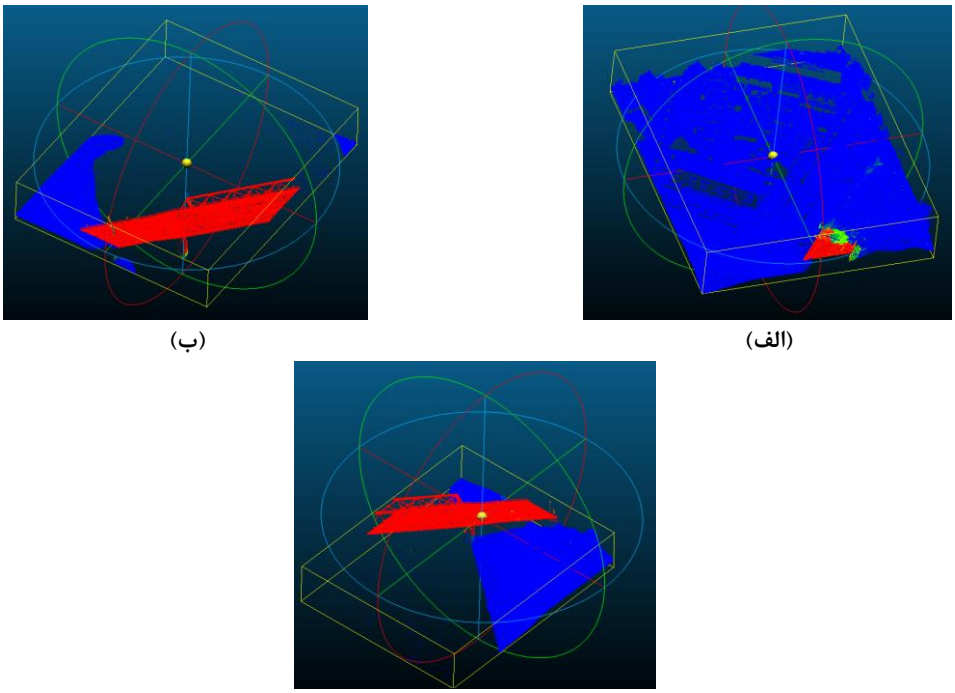


شکل ۸: نمونه‌ای از بخش‌بندی مدل برای کلاس "آب" در کاربری‌های: الف) روستایی و طبیعی ، ب) روستایی ، پ) ساحلی و طبیعی

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]

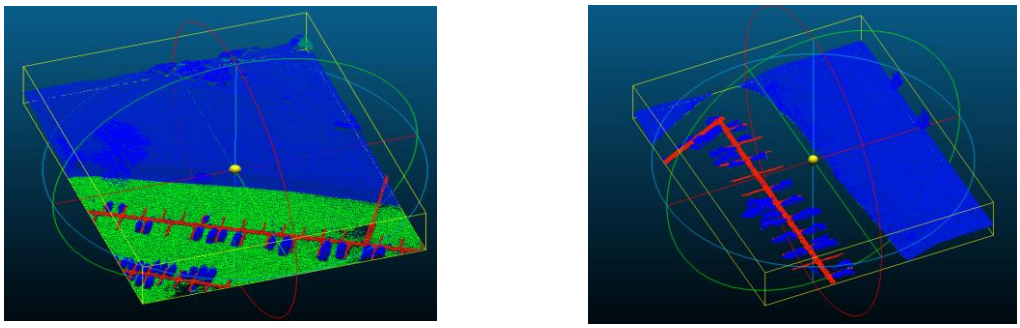
[DOI: 10.66224/jgit.13.2.103]

تصاویر شکل (۹) نشان می‌دهد که چگونه کم‌نمونه بودن کلاس "پل"، منجر به چالش در شناسایی آن توسط مدل شده است. پل‌ها سازه‌هایی مرتفع و اغلب باریک هستند که از نظر هندسی ممکن است با بخش‌هایی از اسکله‌ها یا ساختمان‌ها اشتباه گرفته شوند. عملکرد متوسط مدل (IoU : ۶۵٫۴٪) در این نمونه‌ها مشهود است. همان‌طور که در تصاویر شکل(۱۰) مشاهده می‌شود، این کلاس شامل سازه‌های متفرقه انسان‌ساخت (مانند دکل‌ها، حصارها یا اسکله‌ها) است که تنوع هندسی بالایی دارند. کمبود نمونه‌های آموزشی و نبودِ یک الگوی هندسی ثابت در این نوع سازه‌ها، منجر به عملکرد ضعیف‌تر مدل در این کلاس (IoU : ۶۳٫۷٪) شده است.



(الف) (ب) (پ)

شکل ۹: نمونه‌ای از بخش‌بندی مدل برای کلاس "پل" در کاربری‌های: الف) شهری ، ب) ساحلی و شهری ، پ) ساحلی و شهری



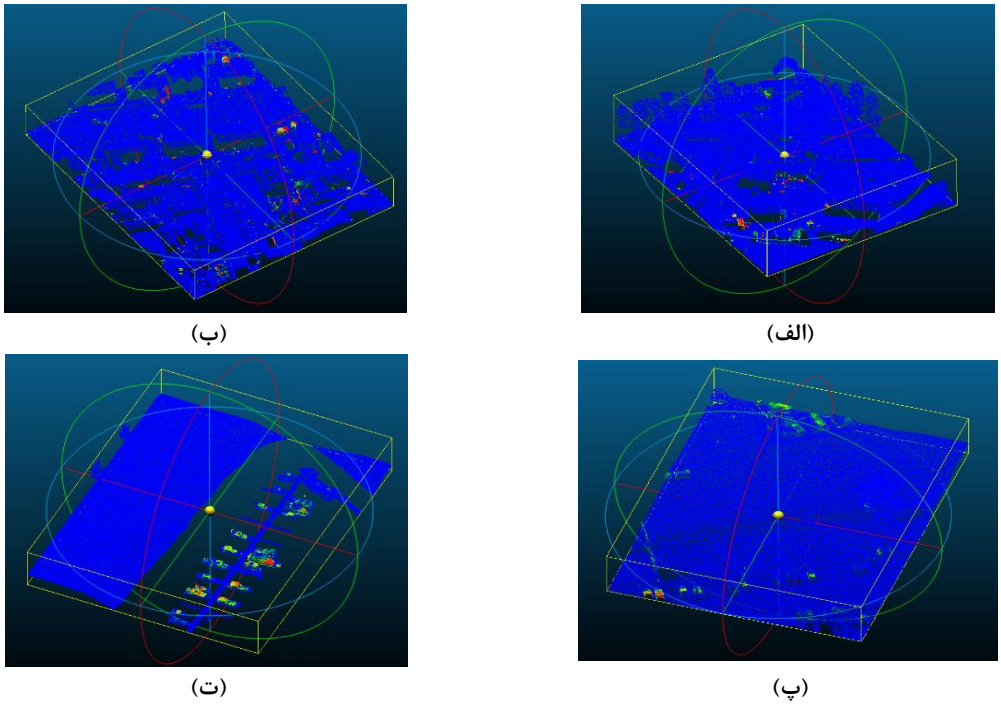
(الف) (ب)

شکل ۱۰: نمونه‌ای از بخش‌بندی مدل برای کلاس "ساختار دائمی" در کاربری‌های: الف) ساحلی ، ب) شهری و ساحلی

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]

[DOI: 10.66224/jgit.13.2.103]

تصاویر شکل (۱۱) عملکرد مدل را برای کلاس "طبقه‌بندی نشده" نشان می‌دهند که پایین‌ترین دقت را دارد (*IoU*: ۴۷٫۸٪). این کلاس شامل اشیاء کوچک، نقاط نویز یا عوارضی است که در هیچ‌یک از دسته‌های دیگر قرار نمی‌گیرند و به همین دلیل، مدل در یافتن یک الگوی مشخص برای آن‌ها با چالش جدی مواجه است.



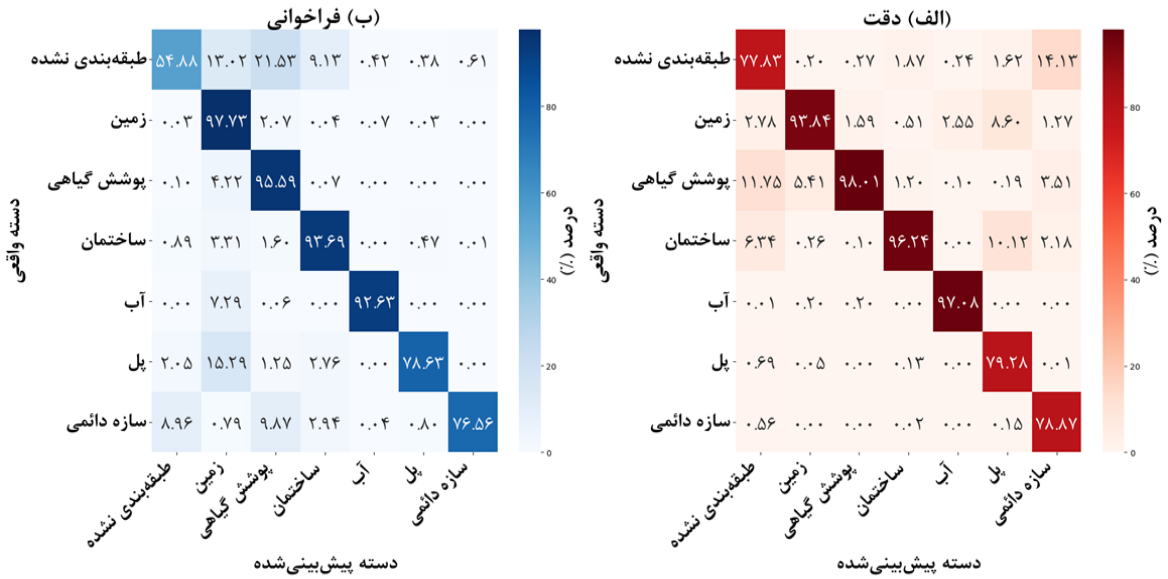
شکل ۱۱: نمونه‌ای از بخش‌بندی مدل برای کلاس "طبقه‌بندی نشده" در کاربری‌های: الف) شهری، ب) شهری، پ) روستایی، ت) ساحلی

ماتریس درهم‌ریختگی (*Confusion Matrix*) ارائه‌شده در شکل (۱۲) نشان می‌دهد که در حالی که مدل در طبقه‌بندی کلاس‌های "زمین" (۹۳٫۸۴٪)، "پوشش گیاهی" (۹۸٫۰۱٪) و "ساختمان" (۹۶٫۲۴٪) به دقت بالایی دست یافته، عملکرد آن برای کلاس "طبقه‌بندی نشده" (۷۷٫۸۳٪) به شکل قابل توجهی پایین‌تر است که بیانگر چالش مدل در شناسایی عوارض نامشخص می‌باشد.

تحلیل متریک فراخوانی از ماتریس درهمریختگی نشان می‌دهد که در حالی که کلاس‌های "زمین" (۹۷٫۷۳٪) و "پوشش گیاهی" (۹۵٫۵۹٪) با نرخ بالایی بازیابی می‌شوند، نرخ فراخوانی پایین برای کلاس "طبقه‌بندی نشده" (۵۴٫۸۸٪) بیانگر ضعف مدل در شناسایی کامل کلاس های کم نمونه این کلاس است.

تحلیل ماتریس درهم‌ریختگی نشان می‌دهد که کلاس‌های کم‌نمونه مانند "پل" و "سازه دائمی" بیشترین خطا را دارند. برای نمونه، بسیاری از نقاط متعلق به کلاس "پل" به‌اشتباه به کلاس‌های مشابه نظیر "ساختمان" یا "زمین" تخصیص یافته‌اند. از آنجا که *LiDARHD* شامل کلاس‌هایی مانند پل و سازه‌های بالای سطح زمین است، خطاهای مشاهده‌شده را می‌توان به کمبود نمونه‌های آموزشی برای این کلاس‌های کم‌نمونه نسبت داد. به‌طور کلی، دقت تفکیک نمونه کلاس‌های کم‌تعداد پایین‌تر بوده و ماتریس درهم‌ریختگی نشانگر تداخل آن‌ها با نمونه کلاس‌های پرتعدادتر است.

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]
[DOI: 10.66224/jgit.13.2.103]



شکل ۱۲: مربوط به ماتریس درهمریختگی هر کلاس: ماتریس (الف) مربوط به دقت و ماتریس (ب) مربوط به فراخوانی است.

۴- بحث و نتایج

نتایج نشان می‌دهد که مدل *RandLA-Net* در طبقه‌بندی کلاس‌های پرتکرار و با ساختار مشخص،مانند زمین، پوشش گیاهی و ساختمان عملکرد بسیار خوبی (*IoU* > ۹۰٪) ارائه داده است. این موفقیت را می‌توان به توانایی ماژول *LFA* در استخراج ویژگی‌های هندسی متمایز و کارایی نمونه‌برداری تصادفی در پردازش حجم بالای نقاط این کلاس‌ها نسبت داد.

در مقابل، عملکرد مدل برای کلاس‌های کم‌نمونه و با هندسه پیچیده‌تر مانند پل (*IoU*: ۶۵٫۴٪) و سازه دائمی (*IoU*: ۶۳٫۷٪) به طور قابل توجهی ضعیف‌تر بود. دلایل اصلی این خطاها را می‌توان به موارد زیر نسبت داد :

- کمبود نمونه آموزشی :تعداد ناکافی نمونه‌ها از این کلاس‌ها در مجموعه داده آموزشی باعث می‌شود مدل نتواند الگوهای پیچیده آن‌ها را به خوبی بیاموزد.
- شباهت هندسی با کلاس‌های دیگر :بخش‌هایی از پل‌ها ممکن است از نظر هندسی به جاده‌ها یا ساختمان‌ها شباهت داشته باشند که منجر به طبقه‌بندی اشتباه می‌شود.

نتایج حاصل از این پژوهش نشان می‌دهد که مدل یادگیری عمیق *RandLA-Net* در بخش‌بندی معنایی ابر نقاط سه‌بعدی عملکرد موفقی داشته است. مدل توانست کلاس‌های زمین، پوشش گیاهی، ساختمان و آب را با دقت بالا شناسایی کند که نشان‌دهنده‌ی توانایی آن در استخراج ویژگی‌های هندسی و متنی داده‌های لایدار است. صحت ۹۷٫۷٪ برای زمین و ۹۵٫۶٪ برای پوشش گیاهی نشان می‌دهد که مدل به‌خوبی قادر به شناسایی الگوهای خاص این کلاس‌ها بوده است. این موفقیت را می‌توان به تفاوت‌های بازتابی و ویژگی‌های هندسی این کلاس‌ها نسبت داد که در فناوری لایدار به‌وضوح قابل تفکیک هستند. علاوه بر این، میزان صحت مدل در بخش‌بندی کلاس‌های ساختمان (۹۶٫۲٪) و آب (۹۷٫۰٪) نیز نشان داد که ویژگی‌های منظم هندسی و بازتاب‌های خاص این کلاس‌ها موجب افزایش کارایی مدل در تشخیص آن‌ها شده است.

فرآیند پیش‌پردازش داده‌ها تأثیر قابل‌توجهی بر عملکرد مدل داشته است. حذف نویز، نرمال‌سازی و نمونه‌برداری مجدد باعث کاهش داده‌های اضافی و بهبود دقت مدل شد. استفاده از تکنیک‌های افزایش

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]
[DOI: 10.66224/jgit.13.2.103]

داده مانند چرخش، انتقال و نویزگذاری موجب شد که مدل بتواند الگوهای پیچیده‌تری را شناسایی کند و قابلیت تعمیم آن افزایش یابد. همچنین، متعادل‌سازی مجموعه داده برای کلاس‌های کمتر نمونه‌برداری شده، به‌ویژه پل‌ها و ساختار دائمی، باعث بهبود شناسایی این کلاس‌ها گردید. در مقایسه با مدل‌هایی که فاقد مرحله‌ی افزایش داده بودند، مدل حاضر توانست دقت بالاتری در تشخیص کلاس‌های کم‌تکرار به دست آورد. به‌طور خاص، صحت مدل در تشخیص کلاس ساختار دائمی به میزان ۷۸/۹٪ رسید.

عملکرد مدل در تشخیص کلاس‌های ساختمان و آب نیز بسیار قابل‌توجه است. مقدار *IoU* برای ساختمان‌ها برابر با ۹۳/۷٪ است که بیانگر توانایی بالای مدل در جداسازی این سازه‌ها از سایر کلاس‌هاست. این دقت را می‌توان به ویژگی‌های منظم هندسی و بازتاب‌های خاص سطوح ساختمان‌ها نسبت داد که مدل توانسته است آن‌ها را به‌درستی شناسایی کند. در مورد کلاس آب، مقدار *IoU* برابر ۹۰/۳٪ ثبت شده که نشان می‌دهد مدل توانسته است این کلاس را با دقت نسبتا بالایی تفکیک کند.

بااین‌حال، عملکرد مدل در شناسایی برخی کلاس‌ها چالش‌برانگیز بوده است. برای مثال، مقدار *IoU* برای کلاس پل ۶۵/۴٪ و برای کلاس ساختار دائمی ۶۳/۷٪ گزارش شده است که نشان می‌دهد مدل در شناسایی این کلاس‌ها تا حدی مشکل داشته است. علت این موضوع می‌تواند پیچیدگی ساختاری و تنوع زیاد این نوع سازه‌ها باشد. همچنین، مقدار *IoU* برای کلاس طبقه‌بندی نشده برابر با ۴۷/۸٪ است که پایین‌ترین مقدار در بین تمامی کلاس‌هاست و نشان‌دهنده‌ی دشواری مدل در تفکیک این دسته از نمونه‌هاست. این ضعف را می‌توان به تنوع بالای اشیای این کلاس و عدم تعریف دقیق آن‌ها در مجموعه داده‌های آموزشی نسبت داد. این مسئله نشان می‌دهد که مدل‌های یادگیری عمیق ممکن است در مواجهه با کلاس‌های دارای تنوع زیاد و بدون ویژگی‌های مشخص، دچار افت دقت شوند.

اگرچه در این پژوهش امکان مقایسه مستقیم با سایر مدل‌ها بر روی همین مجموعه داده فراهم نبود، اما نتایج کسب‌شده با یافته‌های موجود در مقالات دیگر همخوانی دارد. به عنوان مثال، روش‌های پایه مانند *PointNet++* و *KPConv* معمولاً در پردازش ابر نقاط مقیاس بزرگ با محدودیت‌های محاسباتی و حافظه مواجه می‌شوند. کارایی *RandLA-Net* در پردازش سریع داده‌های *LiDARHD* ، برتری رویکرد نمونه‌برداری تصادفی آن را نسبت به روش‌های نمونه‌برداری ساختاریافته در این سناریوها تأیید می‌کند. دستیابی به *mIoU* کلی ۷۸/۱٪ بر روی داده چالش‌برانگیز *AHN4* ، نشان‌دهنده پتانسیل بالای این مدل برای کاربردهای عملی نقشه‌برداری شهری است، هرچند که بهبود عملکرد برای کلاس‌های کم‌تعداد نیازمند تحقیقات بیشتر در زمینه تکنیک‌های موازنه داده و طراحی توابع زیان آگاه به کلاس است.

مدل پیاده‌سازی‌شده بر پایه یادگیری عمیق توانسته است دقت بالایی در شناسایی کلاس‌های مختلف به‌دست آورد و فرآیند پیش‌پردازش شامل حذف نویز، افزایش داده و نمونه‌برداری هدفمند نقش مؤثری در بهبود عملکرد آن داشته است. مقایسه غیرمستقیم با روش‌های مرجع نشان می‌دهد که *RandLA-Net* از نظر دقت و کارایی برتری محسوسی نسبت به مدل‌های پایه دارد. با این حال، چالش‌هایی مانند طبقه‌بندی کلاس‌های کم‌نمونه و اشیای نامشخص همچنان پابرجاست؛ مسائلی که می‌توان آن‌ها را با بهره‌گیری از روش‌های پیشرفته‌تر نظیر یادگیری نیمه‌نظارتی، استفاده از شبکه‌های ترانسفورمر و طراحی توابع زیان حساس به کلاس بهبود بخشید. به‌طور کلی، این پژوهش نشان می‌دهد که شبکه‌های عصبی عمیق نقش مؤثری در ارتقاء دقت و کارایی مدل‌های بخش‌بندی معنایی ابر نقاط سه‌بعدی ایفا می‌کنند و مسیرهای نوینی را برای تحقیقات آینده فراهم می‌سازند.

۵- محدودیت‌های پژوهش

علی‌رغم نتایج مطلوب، این پژوهش دارای

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]

[DOI: 10.66224/jgit.13.2.103]

محدودیت‌هایی است که باید در نظر گرفته شوند. نخست، اگرچه از تکنیک‌های افزایش داده و نمونه‌برداری هدفمند استفاده شد، مدل همچنان در شناسایی کلاس‌های بسیار کم‌نمونه مانند پل و سازه دائمی با چالش مواجه است. این ضعف به‌ویژه در محیط‌های پیچیده شهری که کلاس‌های نادر در آن پراکنده‌اند، برجسته‌تر می‌شود.این امر نشان می‌دهد که برای غلبه کامل بر عدم توازن شدید داده‌ها، ممکن است به رویکردهای پیشرفته‌تری مانند توابع زیان حساس به کلاس (*Class-aware Loss Functions*) یا یادگیری نیمه‌نظارتی نیاز باشد.

دوم، عملکرد مدل تنها بر روی مجموعه داده *AHN4* ارزیابی شده است . بنابراین، قابلیت تعمیم‌پذیری نتایج به سایر داده‌های لایدار با تراکم نقاط، سطح نویز یا ویژگی‌های جغرافیایی متفاوت، نیازمند بررسی‌های بیشتر است. در نهایت، مقایسه با روش‌های پایه مانند *PointNet++*و *KPConv* بر اساس نتایج گزارش‌شده در مقالات دیگر بر روی مجموعه داده‌های متفاوت (مانند *Semantic3D*) انجام شده است و یک مقایسه مستقیم و کنترل‌شده بر روی داده‌های *AHN4* صورت نگرفته است. اذعان به این محدودیت‌ها می‌تواند مسیر را برای تحقیقات آتی در این حوزه هموارتر سازد.

.

مراجع

[1] *C. Gaydon, M. Daab, and F. Roche, "FRACTAL: An Ultra-Large-Scale Aerial Lidar Dataset for 3D Semantic Segmentation of Diverse Landscapes," arXiv preprint arXiv:2405.04634, 2024.*

[2] *S. Yang, M. Hou, and S. Li, "Three-Dimensional Point Cloud Semantic Segmentation for Cultural Heritage: A Comprehensive Review," Remote Sensing, vol. 15, no. 3, p. 548, 2023. [Online]. Available: https://www.mdpi.com/2072-4292/15/3/548.*

[3] *L. Landrieu and M. Simonovsky, "Large-Scale Point Cloud Semantic Segmentation with Superpoint Graphs," in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 18-23 June 2018 2018, pp. 4558-4567, doi: 10.1109/CVPR.2018.00479.*

[4] *R. Khan, M. Abbas, R. Anjum, F. Waheed, S. Ahmed, and F. K. Khan, Evaluating Machine Learning Techniques on Human Activity Recognition Using Accelerometer Data. 2020.*

[5] *L. Tchapmi, C. Choy, I. Armeni, J. Gwak, and S. Savarese, "Segcloud: Semantic segmentation of 3d point clouds," in 2017 international conference on 3D vision (3DV), 2017: IEEE, pp. 537-547, doi: https://doi.org/10.48550/arXiv.1710.07563.*

[6] *Y. Wu, Q. Yao, X. Fan, M. Gong, W. Ma, and Q. Miao, "PANet: A Point-Attention Based Multi-Scale Feature Fusion Network for Point Cloud Registration," IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1-13, 2023, doi: 10.1109/TIM.2023.3271757.*

[7] *Y. Li, R. Bu, M. Sun, W. Wu, X. Di, and B. Chen, "Pointcnn: Convolution on x-transformed points," Advances in neural information processing systems, vol. 31, 2018.*

[8] *B. Wu, A. Wan, X. Yue, and K. Keutzer, "Squeezeseg: Convolutional neural nets with recurrent crf for real-time road-object segmentation from 3d lidar point cloud," in 2018 IEEE international conference on robotics and automation (ICRA), 2018: IEEE, pp. 1887-1893.*

[9] *Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, "Dynamic graph cnn for learning on point clouds," ACM Transactions on Graphics (tog), vol. 38, no. 5, pp. 1-12, 2019, doi: https://doi.org/10.48550/arXiv.1801.07829.*

[10] *X. Lei et al., "WSPointNet: A multi-branch weakly supervised learning network for semantic segmentation of large-scale mobile laser scanning point clouds," International journal of applied earth*

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]

[DOI: 10.66224/jgit.13.2.103]

observation and geoinformation, vol. 115, p. 103129, 2022.

[11] C. Gaydon and F. Roche, "PureForest: A Large-Scale Aerial Lidar and Aerial Imagery Dataset for Tree Species Classification in Monospecific Forests," arXiv preprint arXiv:2404.12064, 2024, doi: https://doi.org/10.48550/arXiv.2404.12064.

[12] F. Matrone, M. Paolanti, E. Frontoni, and R. Pierdicca, "Enhancing explainability of deep learning models for point cloud analysis: a focus on semantic segmentation," International Journal of Digital Earth, vol. 17, no. 1, p. 2390457, 2024/12/31 2024, doi: 10.1080/175389.47.2024.2390457.

[13] H. Yan, A. Lau, and H. Fan, "Evaluating Deep Learning Advances for Point Cloud Semantic Segmentation in Urban Environments," KN - Journal of Cartography and Geographic Information, 2025/01/23 2025, doi: 10.1007/s42489-025-00185-1.

[14] Q. Hu et al., "Randla-net: Efficient semantic segmentation of large-scale point clouds," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 11108-11117.

[15] Z. Xie, H. Liu, Y. He, Y. Shi, P. Yu, J. Ai, and L. Zhong, "Cross modal networks for point cloud semantic segmentation of Chinese ancient buildings," npj Heritage Science, vol. 13, no. 1, p. 131, 2025.

[16] Delft University of Technology. "GeoTiles – Geospatial data services." Delft University of Technology. https://geotiles.citg.tudelft.nl (accessed.

[17] H. Thomas, C. R. Qi, J.-E. Deschaud, B. Marcotegui, F. Goulette, and L. J. Guibas, "Kpconv: Flexible and deformable convolution for point clouds," in Proceedings of the IEEE/CVF international conference on computer vision, 2019, pp. 6411-6420.

[18] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," Advances in neural information processing systems, vol. 30, 2017.

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]

[DOI: 10.66224/jgit.13.2.103]

K. N. TOOSI UNIVERSITY OF TECHNOLOGY
FACULTY OF GEODESY AND GEOMATICS ENGINEERING

Journal of Geospatial Information Technology
Vol.13, No.2, Summer 2025

Research Paper

Evaluation of RandLA-Net Model for Semantic Segmentation of 3D LiDARHD Point Clouds in Complex Urban Environments through Preprocessing Optimization

Javad Sadidi [1]*, Hamzeh Rafizadeh [2], Hani Rezaian [3]

1- Associate Professor, Department of Remote Sensing and Geographic Information Systems, Faculty of Geographic Sciences, Kharazmi University.
2- Master's Student, Department of Remote Sensing and Geographic Information Systems, Faculty of Geographic Sciences, Kharazmi University.
3- Assistant Professor, Department of Remote Sensing and Geographic Information Systems, Faculty of Geographic Sciences, Kharazmi University.

Abstract

Semantic segmentation of 3D point clouds is a key challenge in LiDAR data processing and the applications which are related to computer vision, robotics, and urban mapping. In this study, the performance of the RandLA-Net deep learning model for semantic segmentation of 3D point clouds is investigated. The dataset used in this study is the Dutch LiDARHD data called AHN4, which contains accurate elevation information from different areas, especially the urban and natural environments. To improve the model's accuracy, a data preprocessing pipeline including denoising, normalization, and data augmentation was applied. To assess the model's performance, the Intersection over Union (IoU) metric was used, which measures the percentage of overlap between the predicted and ground truth segments for each class and serves as a precise criterion for evaluating segmentation quality. The results indicated that the RandLA-Net model achieved high performance in segmenting classes such as ground (92٫3% IoU), vegetation (93٫6% IoU), and buildings (93٫7% IoU). However, the complex and underrepresented classes like bridges (65٫4% IoU) were identified with lower accuracy. A comparison of the implemented model with baseline methods such as PointNet++ and KPConv shows that RandLA-Net provides higher accuracy with lower computational cost. The overall results of this study, with an overall average community-to-community ratio (mIoU) of 87٫1% on the LiDARHD dataset, confirm the effectiveness of deep learning in increasing the accuracy and efficiency of the semantic segmentation of 3D point clouds for high-frequency classes. At the same time, these findings highlight the ongoing challenges in correct separation of the sparse classes in complex urban environments.

Key words: *3D Point Clouds, Semantic Segmentation, Deep Learning, RandLA-Net, LiDAR, Preprocessing.*

Correspondence Address*: No. 43, Shahid Mofateh Street, Tehran Province, Kharazmi University..*
Tel*: 09925954305*
Email*: jsadidikhu@khu.ac.ir*

[Downloaded from jgit.kntu.ac.ir on 2026-09-16]

[DOI: 10.66224/jgit.13.2.103]