نشریه علمی مهندسی فناوری اطلاعات مکانی

سال دوازدهم، شماره نخست، بهار ۱۴۰۳

Vol.12, No.1, Spring 2024

۸۳ – ۹٤

مقاله پژوهشی

رویکرد یادگیری عمیق مبتنی بر بلوک‌های تخصصی کانولوشن در استخراج جاده‌های شهری

محمدعرفان امتی[۱]، فاطمه طبیب محمودی[۲*]

۱- دانشجوی کارشناسی ارشد سنجش از دور، گروه مهندسی نقشه‌برداری، دانشکده مهندسی عمران، دانشگاه تربیت دبیر شهید رجایی

۲- استادیار گروه مهندسی نقشه‌برداری، دانشکده مهندسی عمران، دانشگاه تربیت دبیر شهید رجایی

تاریخ دریافت مقاله: ۱۴۰۲/۱۰/۱۶ تاریخ پذیرش مقاله: ۱۴۰۳/۰۳/۲۳

چکیده

استخراج جاده‌ها از تصاویر سنجش از دور با وضوح زیاد در طیف گسترده‌ای از کاربردها مانند تهیه و به‌هنگام‌رسانی نقشه راه‌ها مورد استفاده قرار گرفته است. به دلیل طول زیاد و عرض کم و همچنین سایه‌های ناشی از پوشش گیاهی و ساختمان‌ها، تشخیص جاده‌ها با چالش همراه است. وجود تنوع در جاده‌های موجود در یک منطقه شهری اعم از مجاورت معابر دارای طول کوتاه و بزرگراه‌ها و آزادراه‌های طویل باعث بروز مشکلاتی در تشخیص و طبقه‌بندی خودکار تمامی انواع معابر گردیده است. لذا، با هدف بهبود قابلیت اطمینان و دقت استخراج جاده‌ها، یک مدل شبکه عصبی در این مقاله پیشنهاد شده است که به قطعه‌بندی دقیق پیکسلی دست می‌یابد. شبکه پیشنهادی به‌طور مستقیم تصویر ورودی را پردازش می‌کند و از چهار بلوک کانولوشن تخصصی (*SCB*) در طول نمونه‌برداری کاهشی استفاده می‌کند که با رویکرد نمونه‌برداری با عمق کم برای تولید یک ماسک باینری، کلاس جاده تکمیل می‌شود. با توجه به اینکه شبکه‌های قطعه‌بندی معنایی مرسوم عمیق بوده و دارای پارامترهای قابل آموزش زیادی هستند، شبکه پیشنهادی در این تحقیق از نمونه‌برداری با عمق کم استفاده می‌کند که به کاهش عمق شبکه کمک می‌کند و در نتیجه کاهش تعداد پارامترها را به همراه دارد. قابلیت اجرایی مدل پیشنهادی در این تحقیق با استفاده از مجموعه داده ماساچوست ارزیابی گردید و نتایج ارزیابی به وضوح عملکرد برتر مدل پیشنهادی را در مقایسه با سایر شبکه‌های عصبی با پارامترهای کمتر نشان می‌دهد. مدل پیشنهادی توانست شاخص‌های *IOU* و *F-Score*- را در مجموعه داده ماساچوست در مقایسه با سایر شبکه‌های عصبی نظیر *DeeplabV3+*، *U-Net* و *D-LinkNet* به ترتیب به میزان ۱٫۹۸ و ۳٫۰۳ بهبود بخشد.

کلیدواژه‌ها : قطعه‌بندی معنایی، استخراج جاده، تصاویر سنجش از دور، بلوک‌های تخصصی کانولوشن.

* نویسنده مکاتبه کننده: گروه مهندسی نقشه‌برداری، دانشکده مهندسی عمران، دانشگاه تربیت دبیر شهید رجایی، لویزان، تهران، ایران.

تلفن: ۰۲۱۲۲۹۷۰۰۲۱

Email: fmahmoudi@sru.ac.ir

[Downloaded from jgit.kntu.ac.ir on 2025-07-13]

[DOI: 10.61186/jgit.12.1.83]

۱- مقدمه

به عنوان یکی از اجزای اصلی حمل و نقل، جاده‌ها تأثیر قابل‌توجهی بر جنبه‌های مختلف زندگی معاصر دارند که مشتمل بر مدیریت ترافیک، توسعه شهری و روستایی و ناوبری خودروها می‌باشد [۱]. همزمان با تکامل فناوری سنجش از دور، تصاویر سنجش از دور با وضوح زیاد به عنوان یک دارایی حیاتی برای پردازش رقومی تصاویر در عصر مدرن به شمار می‌آیند [۲]. در این راستا، استخراج جاده با دقت بالا از تصاویر ماهواره‌ای با وضوح زیاد، توجه محققان را در سال‌های اخیر به خود جلب کرده است. با این حال، آثار ناشی از عواملی مانند روشنایی، سایه و نویز، ویژگی‌های متفاوتی را در بین اهداف جاده‌ای یکسان ایجاد می‌کند. به علاوه، ابهام طیفی ناشی از تأثیر اهداف غیرجاده‌ای دارای ویژگی‌های مشابه نیز دشواری مسیر را در استخراج دقیق و جامع جاده‌ها از تصاویر با وضوح زیاد افزایش می‌دهد [۳ و ۴].

در سال‌های اخیر همراه با پیشرفت سریع الگوریتم‌های یادگیری عمیق، تعداد فزاینده‌ای از محققین از روش‌های یادگیری عمیق در حوزه‌های طبقه‌بندی تصویر و قطعه‌بندی معنایی انواع عوارض از جمله جاده-ها استفاده کرده‌اند [۵، ۶، ۷، ۸، ۹ و ۱۰]. شبکه‌های عصبی پیچشی[۱] کارایی خود را در کاربردهای استخراج جاده به دلیل قابلیت‌های قابل‌توجه استخراج ویژگی نشان داده‌اند [۶]. در مقایسه با رویکردهای مرسوم که به استخراج دستی ویژگی‌های سطحی نیاز دارند؛ روش‌های یادگیری عمیق نه تنها دقت استخراج جاده را افزایش می‌دهند؛ بلکه در استخراج عوارض بزرگ مقیاس، عملکرد بهتری دارند [۷].

علی‌رغم قابلیت‌های روش‌های یادگیری عمیق، هنوز هم چالش‌های بزرگی برای قطعه‌بندی جاده‌ها از تصاویر سنجش از دوری وجود دارد که برطرف نمودن آنها نیازمند انجام تحقیقات بیشتر است. از جمله چالش‌های مطرح در تشخیص جاده‌های شهری می‌توان به موارد زیر اشاره نمود:

۱) جاده‌ها شباهت هندسی به رودخانه‌ها و خندق‌ها دارند و این امر مستعد طبقه‌بندی اشتباه است.

۲) جاده‌های کوچک معمولاً مساحت بسیار کوچکی از تصویر را اشغال می‌کنند که منجر به عدم تعادل شدید بین پیش‌زمینه و پس‌زمینه می‌شود.

۳) جاده‌های کوچک در برابر تداخل، آسیب‌پذیرتر هستند که منجر به عدم شناسایی و ناپیوستگی می‌گردد.

نیه و هینتون (۲۰۱۰) در پیاده‌سازی تکنیک‌های یادگیری عمیق در استخراج جاده‌ها پیشگام شدند و روشی را پیشنهاد کردند که از ماشین‌های محدود بولتزمن برای شناسایی جاده‌ها در تصاویر ماهواره‌ای با وضوح زیاد استفاده می‌کرد [۸]. در تحقیق انجام شده توسط لی و همکاران (۲۰۱۶) ابتدا شبکه عصبی پیچشی برای تعیین پیکسل‌های احتمالی متعلق به جاده‌ها استفاده می‌شود و به دنبال آن الگوریتم هموارسازی مبتنی بر انتگرال خطی برای اتصال شکاف‌های کوچک طراحی شده است[۹]. در تحقیق باستانی و همکاران (۲۰۱۸) یک روش جستجوی تکراری در تابع تصمیم‌گیری شبکه پیچشی برای کاهش پیچیدگی‌های انجام عملیات پس‌پردازش و به منظور بهبود دقت نتایج پیشنهاد شده است [۱۰].

شبکه‌های عصبی پیچشی عمیق[۲] *(DCNN)* تکامل یافته شبکه‌های عصبی مصنوعی هستند که در سال‌های اخیر، رایج‌ترین روش‌های قطعه‌بندی معنایی تصاویر با وضوح زیاد براساس این نوع از شبکه‌ها طراحی شده‌اند [۱۱، ۱۲، ۱۳، ۱۴، ۱۵، ۱۶، ۱۷، ۱۸، ۱۹، ۲۰، ۲۱، ۲۲، ۲۳، ۲۴، ۲۵ و۲۶]. مزیت اصلی این الگوریتم‌ها این است که می‌توانند ویژگی‌های سطح بالا

[۱] *Convolutional Neural Networks*

[۲] *Deep Convolutional Neural Network*

[Downloaded from jgit.kntu.ac.ir on 2025-07-13]

[DOI: 10.61186/jgit.12.1.83]

را بیاموزند و استخراج کنند که این امر منجر به بهبود عملکرد آنها شده و دقت قطعه‌بندی معنایی را بهبود می‌بخشد [۱۲].

در ساختار شبکه‌های کاملاً پیچشی[۱] [۱۷]، لایه‌های کاملاً متصل با لایه‌های کانولوشن جایگزین می‌شوند تا نقشه‌های ویژگی، حاوی اطلاعات موقعیتی باشند. در قطعه‌بندی معنایی مبتنی بر *DCNN*، تکرار عملیات پیچش و فرآیند نمونه‌برداری کاهشی، منجر به از دست‌دادن ویژگی‌های مکانی مهم و چالش‌های قطعه‌بندی مانند لبه‌های نامشخص و مرزهای تار می‌شود. برای مقابله با این محدودیت‌ها، شبکه‌های *U-Net* معرفی شدند[۱۳ و ۲۶]. این رویکرد به طرز ماهرانه‌ای جزییات مکانی فراوانی را بازیابی می‌کند و از اطلاعات بافت جاده به طور بهینه بهره‌برداری می‌نماید. با این حال، با افزایش عمق *U-Net*، مسائلی مانند ناپدیدشدن گرادیان ممکن است رخ دهد. در نتیجه، تحقیقات موجود بر بهبود عملکرد و پایداری شبکه با استفاده از معماری منحصر به فرد *U-Net* تمرکز می‌کنند؛ در حالیکه به طور همزمان نیز مدل شبکه را اصلاح می‌نمایند [۲۷]. برای مثال، شاو و همکاران (۲۰۲۲) در مکانیزم توجه به کانال[۲] و توجه به مکان[۳] را برای بهبود استفاده از اطلاعات طیفی و مکانی در شبکه *U-Net* و با هدف استخراج جاده‌ها معرفی نمودند [۲۶]. به علاوه، بلوک‌های متصل متراکم باقی‌مانده[۴] برای افزایش استفاده مجدد از ویژگی و انتقال جریان اطلاعات معرفی شدند، و یک ماژول باقیمانده کانولوشن گسترده[۵] نیز برای استخراج اطلاعات شبکه جاده‌ای در مقیاس‌های مختلف معرفی شد.

ژانگ و همکاران با تشکیل بلوک‌های *ResNet* و ترکیب با معماری *U-Net*، شبکه‌های *ResUNet* را ایجاد کردند[۲۸]. علاوه براین ژو و همکاران مدل *D-LinkNet* را توسعه دادند که در آن ماژول‌های کانولوشن گسترده در مدل *U-Net* ادغام می‌شود تا پذیرندگی میدان گیرنده را افزایش داده و با مقیاس‌گذاری و برش تصویر ورودی به ابعاد مختلف، قابلیت تشخیص عوارض در مقیاس بزرگ را بهبود بخشد [۲۹ و ۳۰].

هنگام استخراج عوارضی نظیر جاده‌ها که دارای ابعاد مختلفی هستند، استراتژی دیگر شامل بزرگ‌کردن میدان‌های گیرنده از طریق عملیات کانولوشن با نرخ‌های مختلف یا همان عملیات کانولوشن گسترده است. به عنوان نمونه، می‌توان به شبکه‌های *DeepLabv3+* در این زمینه اشاره کرد [۳۱]. استفاده از کانولوشن گسترده، میدان دریافت بزرگ‌تری را بدون افزایش تعداد پارامترها فراهم می‌کند و به دنبال آن اطلاعات غنی برای پیش‌بینی جزئیات تصاویر بدست می‌آید. با این حال، روش استخراج ویژگی مبتنی بر عملیات کانولوشن گسترده می‌تواند به راحتی اطلاعات مربوط به اهداف کوچک مقیاس را از دست بدهد؛ زیرا هسته‌های پیچشی[۶] فقط ویژگی‌ها را از مناطق محدود استخراج می‌کنند.

تحقیق انجام‌شده در این مقاله برای بررسی و بهبود چالش‌های ذکرشده و ارتقاء عملکرد استخراج جاده، یک مدل شبکه یادگیری عمیق را ارائه می‌کند که برای ویژگی‌های جاده در تصاویر سنجش از دور طراحی شده است. شبکه پیشنهادی در بخش رمزگذار، از چهار بلوک کانولوشن متشکل از لایه‌های کانولوشن، توابع فعال‌ساز[۷] و نرمال‌سازی دسته‌ای[۸] و عملیات

[۱] *Fully Convolutional Network (FCN)*

[۲] *Channel Attention*

[۳] *Spatial Attention*

[۴] *Residual densely connected blocks*

[۵] *Residual Dilated Convolution Modules*

[۶] *Convolution Kernels*

[۷] *Activation Function*

[۸] *Batch Normalization(BN)*

[Downloaded from jgit.kntu.ac.ir on 2025-07-13]

[DOI: 10.61186/jgit.12.1.83]

نمونه‌برداری کاهشی[۱] برای کاهش ابعاد نقشه‌های ویژگی استفاده می‌کند و برای بازسازی اطلاعات مکانی و تولید خروجی نهایی، از عملیات نمونه‌برداری افزایشی[۲] نسبتا ساده‌تر با هدف تولید یک ماسک باینری استفاده می‌نماید. برای پایین نگه‌داشتن هزینه محاسباتی، از تعداد محدودی کانولوشن‌های ترانهاده[۳] در روند نمونه‌برداری افزایشی استفاده می‌شود. با توجه به اینکه شبکه‌های قطعه‌بندی معنایی مرسوم عمیق هستند و دارای پارامترهای قابل آموزش زیادی هستند، شبکه پیشنهادی در این تحقیق از نمونه‌برداری افزایشی با عمق کم[۴] استفاده می‌کند که به کاهش عمق شبکه کمک می‌کند و در نتیجه کاهش تعداد پارامترها را به همراه دارد. به منظور مشاهده اثربخشی چارچوب پیشنهادی، کارایی مدل پیشنهادی در مقایسه با معماری‌های *U-Net*، *D-LinkNet* و *DeepLabV3+* بررسی می‌شود.

۲- روش پیشنهادی

ساختار کلی رویکرد پیشنهادی در این تحقیق در شکل(۱) نشان داده شده‌است. این شبکه براساس چهار بلوک تخصصی کانولوشن(*SCB*)[۵] شکل می‌گیرد. استراتژی مورداستفاده در این تحقیق با هدف کاهش هزینه‌های محاسباتی شبکه، استفاده از لایه‌های پیچشی عمومی[۶] و لایه‌های پیچشی قابل تفکیک[۷] می‌باشد. دلیل استفاده از لایه‌های پیچشی عمومی، توانمندسازی شبکه برای ثبت ویژگی‌های مکانی در ورودی است. لایه‌های پیچشی قابل تفکیک نیز با تقسیم به دو نوع عملیات مجزا که متشکل از پیچش عمقی[۸] و پیچش نقطه‌ای[۹] هستند، شبکه را از نظر محاسباتی کارآمدتر می‌کنند.

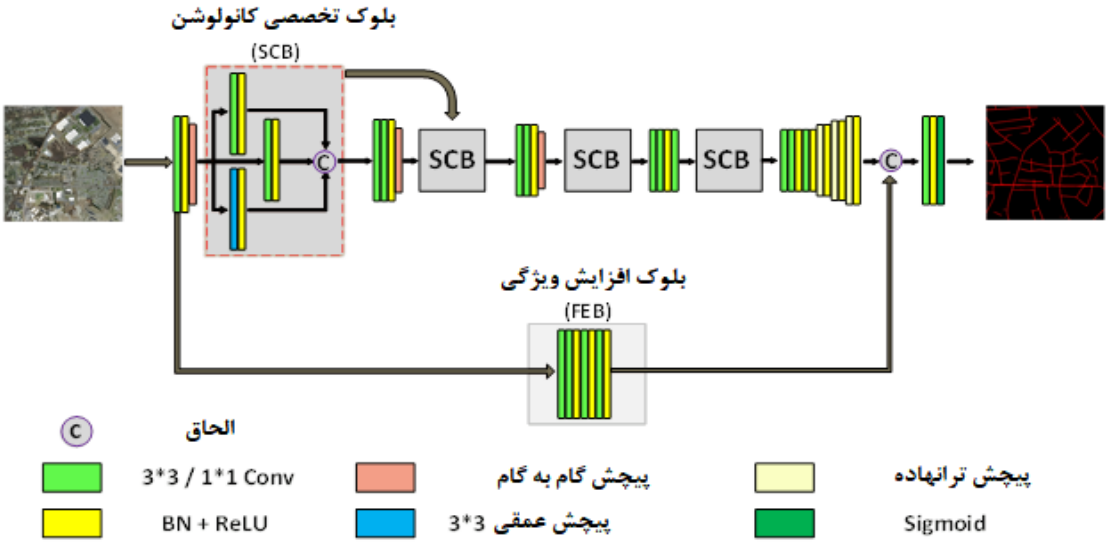
ساختار هر *SCB* در شکل(۲) نشان داده شده‌است. هر *SCB* در شبکه پیشنهادی، شامل یک لایه پیچشی عمومی ۳×۳، یک لایه پیچشی قابل تفکیک ۳×۳ و یک لایه پیچشی عمومی ۱×۱ است. این لایه‌های پیچشی به‌طور مستقیم به یک لایه پیچشی گسترده (لایه پیچشی گام به گام)[۱۰] متصل می‌شوند و از عملیات *BN* و تابع فعال‌ساز *ReLU* برای ترکیب خروجی‌های خود استفاده می‌کنند و در نهایت، عملیات الحاق نیز برای ادغام خروجی‌های این سه لایه استفاده می‌شود.

در شبکه پیشنهادی در این مقاله به موازات طراحی مسیر بلوک‌های *SCB*، بلوک افزایش ویژگی(*FEB*)[۱۱] قرار داده شده است تا ویژگی‌های مکانی را بهبود دهد. دلیل این امر آن است که *FEB* اطلاعات مکانی را در لایه‌های اولیه شبکه افزایش می‌دهد و ویژگی‌های غنی را در مرحله نهایی شبکه فراهم می‌کند. *FEB* از یک بلوک پیچشی با عمق کم تشکیل شده‌است که به‌طور قابل توجهی ویژگی‌های سطح پایین را حفظ می‌کند و آن‌ها را در انتهای شبکه ارائه می‌دهد. این بلوک از چهار کانولوشن با عمق حداکثر ۱۶ کانال استفاده می‌کند و پارامترهای شبکه را به حداقل می‌رساند[۳۲].

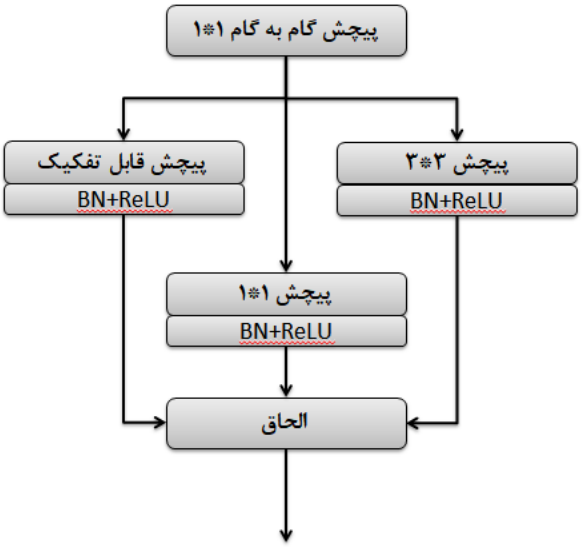
---

[۱] *Down-sampling*
[۲] *Up-sampling*
[۳] *Transposed Convolutions*
[۴] *Shallow Upsampling*
[۵] *Specialized Convolution Blocks (SCB)*
[۶] *General Convolutions*
[۷] *Separable Convolution*
[۸] *Depthwise Convolution*
[۹] *Pointwise Convolution*
[۱۰] *Strided Convolution*
[۱۱] *Feature Enhancement Block (FEB)*

[Downloaded from jgit.kntu.ac.ir on 2025-07-13]
[DOI: 10.61186/jgit.12.1.83]



شکل۱: چارچوب مدل پیشنهادی برای قطعه‌بندی معنایی تصاویر با وضوح زیاد



شکل ۲: ساختار هر بلوک *SCB*

در معماری پیشنهادی، بلوک نمونه‌برداری افزایشی، ویژگی‌ها را از سمت نمونه‌برداری کاهشی دریافت می‌کند و از قسمت نمونه‌برداری با عمق کم، ویژگی *FUS*[1] را که در شبکه‌های قطعه‌بندی معنایی رایج است، خروجی می‌دهد. نمونه‌برداری کاهشی پی در پی و لایه‌های پیچشی متعدد منجر به از دست‌رفتن ویژگی‌ها می‌شود و *FUS*، نمی‌تواند نرخ مثبت واقعی[2] خوبی ارائه دهد. در نتیجه، بلوک *FEB*، مشخصه *Fi* را

[1] *Shallow Up-Sampling Feature (FUS)*

[2] *True Positive*

[Downloaded from jgit.kntu.ac.ir on 2025-07-13]

[DOI: 10.61186/jgit.12.1.83]

از لایه‌های اولیه شبکه می‌گیرد و مشخصه FE را پس از اعمال لایه‌های پیچشی کم‌عمق که شامل ویژگی‌های سطح پایین است، تولید می‌کند.

در نهایت دو مشخصه FUS و FE مطابق رابطه (۱) برای تولید مشخصه CS ترکیب می‌شوند:

$$C_S = F_{US} \oplus F_E \qquad \text{رابطه(۱)}$$

در رابطه (۱)، $\oplus$ نشان‌دهنده الحاق عمقی[1] دو دسته ویژگی می‌باشد. ویژگی‌های به‌دست آمده از CS ، حاوی اطلاعات لبه غنی است که در نتیجه حساسیت بالاتری ایجاد می‌کند. در انتهای شبکه، از تابع *Sigmoid* برای تعیین برچسب جاده به هر پیکسل تصویر استفاده می‌شود.

۳- نتایج اجرایی

در ادامه مباحث، در این بخش ابتدا داده‌های مورداستفاده در این پژوهش معرفی شده و سپس، نتایج اجرایی مراحل مخلف روش پیشنهادی ارائه می‌گردد.

۳-۱- داده مورد استفاده

در این تحقیق، کارایی شبکه قطعه‌بندی را با استفاده از مجموعه داده‌های جاده ماساچوست ارزیابی کردیم. مجموعه داده‌های جاده ماساچوست که در سطح جهانی به عنوان گسترده‌ترین مجموعه داده جاده‌ای قابل‌دسترسی برای عموم شناخته می‌شود توسط نیه و هینتون[2] ایجاد شد[۳۳]. این مجموعه داده تقریباً ۵۰۰ کیلومتر مربع را پوشش می‌دهد و مجموعه‌ای متنوع از مناطق جغرافیایی از شهری تا روستایی را شامل می‌شود و مجموعه‌ای از انواع جاده‌ها مانند بزرگراه‌ها، جاده‌های خاکی روستایی و جاده‌های آسفالت را در برمی‌گیرد. این مجموعه داده شامل ۱۱۷۱ تصویر است که هر کدام با قدرت تفکیک هر پیکسل ۱٫۲ متر و ابعاد ۱۵۰۰ × ۱۵۰۰ پیکسل هستند. مجموعه داده به ۱۱۰۸ تصویر آموزشی، ۴۹ تصویر آزمایشی و ۱۴ تصویر اعتبارسنجی تقسیم‌بندی شده است. مجموعه‌های آموزشی و اعتبارسنجی دارای برچسب‌های تصویر باینری هستند که در آن پیکسل‌های جاده با ۱ و پیکسل‌های پس‌زمینه با ۰ مشخص می‌شوند. با توجه به محدودیت‌های سخت‌افزاری و ضرورت بهینه‌سازی عملکرد مدل، استفاده از تصاویر با اندازه بزرگ برای آموزش ناکارآمد تلقی می‌شود. بنابراین، در این تحقیق، ۱۱۰۸ تصویر آموزشی و برچسب‌های مربوط به آنها، هر کدام به ابعاد ۲۵۶ × ۲۵۶ پیکسل تقسیم شدند. مجموعه آموزشی نهایی ۲۰۷۷۶ تصویر، هر کدام با ابعاد ۲۵۶ × ۲۵۶ پیکسل و مجموعه اعتبارسنجی شامل ۶۹۲۴ تصویر با ابعاد یکسان آماده شدند.

۳-۲- معیارهای ارزیابی

جنبه‌های مختلف اثربخشی رویکرد پیشنهادی با استفاده از چهار معیار ارزیابی زیر مورد بررسی قرار می‌گیرد:

- معیار *IOU*، همپوشانی بین مناطق ساختمانی پیش‌بینی‌شده و زمینی را با تقسیم تقاطع دو منطقه بر اتحاد آنها محاسبه می‌کند (رابطه (۲)).

$$\mathrm{IOU} = \frac{\mathrm{TP}}{\mathrm{FP} + \mathrm{TP} + \mathrm{FN}} \qquad \text{رابطه(۲)}$$

- معیار دقت، به نسبت نمونه‌های مثبت به درستی شناسایی شده در مقایسه با تعداد کل پیش‌بینی‌های مثبت اشاره دارد (رابطه (۳)). دقت بالا نشان‌دهنده نرخ مثبت کاذب پایین است؛ که نشان می‌دهد مدل در شناسایی دقیق نمونه‌های مثبت برتری دارد (رابطه(۳)).

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{FP} + \mathrm{TP}} \qquad \text{رابطه(۳)}$$

در رابطه (۳)، مثبت واقعی *(TP)* به مواردی اشاره دارد که مدل به درستی یک جاده را شناسایی می‌کند. منفی واقعی *(TN)* مربوط به مواردی است که مدل به درستی یک عارضه غیرجاده ای را شناسایی می‌کند. پیش‌بینی مدل با حقیقت واقعی در هر دوی این موارد مطابقت دارد. مثبت کاذب *(FP)* زمانی رخ می‌دهد که مدل به اشتباه یک عارضه غیرجاده‌ای را به عنوان یک جاده شناسایی کند. منفی کاذب *(FN)* زمانی ایجاد می‌شود

[1] *Depth-wise Concatenation*

[2] *Minh and Hinton*

[Downloaded from jgit.kntu.ac.ir on 2025-07-13]

[DOI: 10.61186/jgit.12.1.83]

که مدل به اشتباه یک عارضه جاده‌ای را به عنوان غیرجاده طبقه‌بندی کند. در هر دو موقعیت، پیش‌بینی مدل با حقیقت واقعی در تضاد است.

- معیار بازخوانی[1] این معیار به شناسایی صحیح موارد مثبت واقعی نسبت به تعداد کل موارد مثبت واقعی در مجموعه آزمایشی اشاره دارد. مقادیر زیاد این معیار نشان‌دهنده نرخ منفی کاذب پایین است (رابطه(۴)).

$$\text{Recall} = \frac{\text{TP}}{\text{FN} + \text{TP}}$$ رابطه(۴)

- معیار امتیاز-اف[2] نشان‌دهنده میانگین متعادل بین دقت و معیار بازخوانی است که به عنوان میانگین هارمونیک محاسبه می‌شود (رابطه(۵)).

$$\text{F}_{\text{score}} = \frac{2*(\text{Precision}*\text{Recall})}{\text{Precision} + \text{Recall}}$$ رابطه(۵)

۳-۳- تنظیمات اجرایی

بسته برنامه نویسی *PyTorch* بر روی یک *GPU GeForce RTX 3060* در تمام آزمایشات برای اجرای استراتژی پیشنهادی استفاده شد. مقادیر درجات خاکستری پیکسل‌های تصاویر قبل از استفاده برای آموزش، از محدوده اصلی [۰، ۲۵۵] تا محدوده [۰، ۱] نرمال می‌شوند. ابعاد ورودی را ۲۵۶×۲۵۶ در نظر می‌گیریم. معماری‌های تاثیرگذار و کارآمد *U-Net*، *D-LinkNet* و *DeepLabv3+* برای مقایسه با مدل پیشنهادی و ارزیابی اثربخشی آن انتخاب شده‌اند. تحت شرایط مشابه، هر چهار مدل بر روی مجموعه داده جاده ماساچوست آزمایش می‌شوند. از بهینه ساز *Adam* با نرخ یادگیری اولیه ۰٫۰۰۱ برای آموزش شبکه استفاده شده است [۳۴]. رویکرد مورد استفاده برای زمان‌بندی نرخ یادگیری، استراتژی چند جمله‌ای است که مطابق رابطه (۶) محاسبه می‌شود [۳۵]:

$$\text{Lr} = 0.001 * (1 - \frac{\text{iter}}{\text{max_iter}})^{0.75}$$ رابطه(۶)

فرآیند آموزش مشتمل بر ۶۰ اپک است و تابع زیان آنتروپی متقابل[3] برای استخراج جاده و در کنار آن، تابع زیان تاس[4] برای کاهش مشکل نمونه‌های نامتعادل مورد استفاده قرار گرفت. فرمول محاسبه تابع زیان آنتروپی متقابل در رابطه (۷) نشان داده شده است:

$$\text{L}_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}[\, y_i \log\,(\hat{y}_i)] + (1 - y_i)\log(1 - \hat{y}_i)\,]$$ رابطه(۷)

در رابطه (۷)، y مقدار برچسب پیکسل واقعی را نشان می‌دهد و $\hat{y}_i$ نشان‌دهنده مقدار پیکسل برچسب پیش‌بینی‌شده توسط شبکه و N نشان‌دهنده تعداد پیکسل‌ها است.

فرمول ضریب تاس به عنوان یک تابع زیان در رابطه(۸) نشان داده شده است:

$$\text{L}_{dice} = 1 - \frac{2\,|\text{X} \cap \text{Y}|}{|\text{X}| + |\text{Y}|}$$ رابطه(۸)

در رابطه (۸)، X تصویر پیش‌بینی‌شده توسط شبکه پیشنهادی، Y برچسب واقعی تصویر ورودی، $|X|$ تعداد پیکسل‌ها در تصویر پیش‌بینی و $|Y|$ تعداد پیکسل‌ها را در برچسب واقعی نشان می‌دهد.

تابع زیان تاس برای الگوهای نامتعادل دارای عملکردی قوی است و می‌تواند مشکل عدم تعادل نمونه را حل کند. این به‌ویژه برای کارهای استخراج جاده که در آن عدم‌تعادل قابل‌توجهی بین نمونه‌های مثبت و منفی وجود دارد، کاربرد دارد. در چنین مواردی، اگر فقط تابع زیان *CE* در سطح پیکسل استفاده شود، این امکان وجود دارد که شبکه به سمت پیش‌بینی کلاس با نمونه‌های بیشتر حرکت کند. در این تحقیق، تابع زیان آنتروپی متقابل و تابع زیان تاس به عنوان تابع زیان نهایی با یکدیگر ترکیب شده که در رابطه (۹) نشان داده شده‌است:

$$\text{L}_{\text{final}} = \text{L}_{\text{BCE}} + \text{L}_{\text{dice}}$$ رابطه(۹)

[1] *Recall*

[2] *F Score*

[3] *Cross-entropy Loss Function*

[4] *Dice Loss Function*

[Downloaded from jgit.kntu.ac.ir on 2025-07-13]

[DOI: 10.61186/jgit.12.1.83]

۴- مقایسه و تجزیه و تحلیل نتایج

در جدول (۱) نتایج کمی چهار مدل ارزیابی‌شده در مجموعه داده‌های جاده ماساچوست ارائه شده است. از آنجایی که عدم تعادل کلاسی شدید در کار استخراج جاده وجود دارد، نتایج ارزیابی با معیارهای *IoU* و *F Score* مناسب‌تر است. روش پیشنهادی در این مقاله نسبت به سه روش دیگر در معیارهای *IoU* و *F Score* عملکرد بهتری دارد و به ترتیب مقادیر ۸۰٫۸۸ و ۷۸٫۰۰ را به‌دست می‌آورد. شکل(۳) نتایج بصری به‌دست‌آمده توسط مدل‌های مختلف بر روی مجموعه داده جاده ماساچوست را نشان می‌دهد.

جدول۱: ارزیابی کمی عملکرد مدل پیشنهادی در مجموعه داده ماساچوست

مدل شبکه	دقت	بازخوانی	امتیاز-اف	*IOU*
DeepLabV3+	۷۸٫۰۶	۶۹٫۰۶	۷۳٫۲۸	۷۷٫۶۹
U-Net	۷۸٫۲۷	۷۰٫۴۱	۷۴٫۱۳	۷۴٫۵۲
D-LinkNet	۸۷٫۶۱	۶۵٫۵۲	۷۴٫۹۷	۷۸٫۹۰
مدل پیشنهادی	۸۲٫۵۹	۷۳٫۸۹	۷۸٫۰۰	۸۰٫۸۸

جهت بررسی عملکرد مدل پیشنهادی در این مقاله، نتیجه حاصل با خروجی مدل‌های انتخابی مقایسه شد. همان‌طور که از جدول(۱) مشاهده می‌شود؛ مدل پیشنهادی در این مقاله، به ترتیب ۲٪، ۶٫۴٪ و ۳٫۲٪ در شاخص *IOU* و ۳٪، ۳٫۹٪ و ۴٫۷٪ در شاخص *F-SCORE* از *D-LinkNet*، *U-Net* و *DeepLabV3+* بهتر عمل کرده است. اگرچه متریک *Precision* شبکه پیشنهادی، عملکرد ضعیف‌تری نسبت به *D-LinkNet* دارد، اما می‌توان نتیجه گرفت که عملکرد کلی این مدل بهتر از *D-LinkNet* و دو مدل اصلی دیگر است.

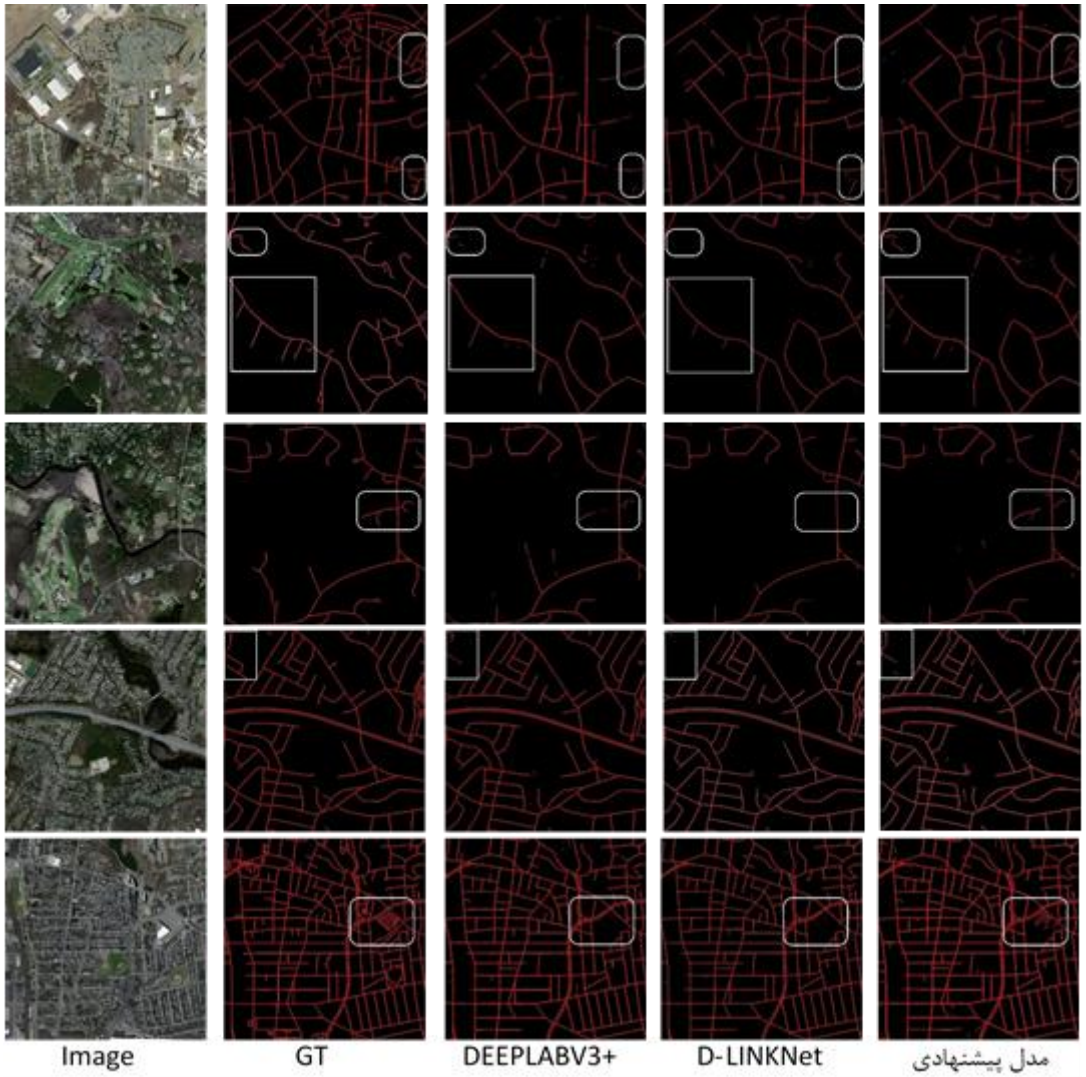
تمایز قابل توجهی بین شبکه پیشنهادی این مقاله و مدل‌های جایگزین فوق الذکر وجود دارد که شایسته توجه است. شبکه پیشنهادی برای دستیابی به تعادل بین عملکرد و مقرون به صرفه بودن طراحی شده است. به طور خاص، این شبکه از ۳٫۵ میلیون پارامتر قابل آموزش تشکیل شده است؛ در حالی که برای مثال، *DeepLabV3+* حدود ۶ میلیون و یا شبکه *D-LinkNet* حدود ۱۱ میلیون پارامتر دارد. این پتانسیل رویکرد مدل پیشنهادی برای دستیابی به نتایج قابل مقایسه با پارامترهای کمتر را نشان می‌دهد که از نظر محاسباتی امیدوارکننده است.

با توجه به شکل (۳)، قاب‌های سفید رنگ به دلیل وجود پوشش گیاهی و اندازه کوچک، قسمت‌هایی که به راحتی از دست می‌روند را برجسته می‌کنند. در این نتایج می‌توان مشاهده کرد که مدل پیشنهادی تشخیص بهتری در جاده‌های با اندازه کوچک نسبت به نمونه‌های مرجع دارد. به عنوان مثال جاده‌های پارکینگ از نظر رنگی مشابه پس‌زمینه هستند (سطر پنجم) و برخی از جاده‌های کوچک پوشیده از پوشش گیاهی نازک هستند (سطر دوم و چهارم). در این دو مورد، مدل‌های مرجع نمی‌توانند به طور موثر این جاده‌های کوچک را شناسایی کنند در حالیکه مدل پیشنهادی نرخ تشخیص بسیار بالاتری را ارائه می‌دهد.

به طور مشابه در مقایسه با مدل‌های مرجع، مدل ارائه شده در این مقاله تداوم بسیار بهتر (خط چهارم)، نرخ کمتر در تشخیص نادرست و نرخ *Recall* بالاتری (سطر

[Downloaded from jgit.kntu.ac.ir on 2025-07-13]
[DOI: 10.61186/jgit.12.1.83]

یکم و سوم) را برای جاده‌های کوچک نشـان مـی‌دهـد. هم معیارهای آزمون و هم نتایج بصری نشـان مـی‌دهـد که شبکه پیشنهادی می‌تواند به انـدازه کـافی اطلاعـات معنایی بافت محلی و جهانی را به تصویر بکشـد و بهتـر از سایر مدل‌های مقایسـه‌ای بـرای اسـتخراج جـاده در محیط‌های پیچیده عمل کند که در آن‌ها جاده‌هایی بـا اندازه‌های گوناگون و جاده‌هایی با اندازه کوچک کـه بـه دلیل پوشش گیاهی اطراف و یا سایه‌های سـاختمان‌هـا مســتعد از دســت رفــتن هســتند؛ وجــود دارنــد.



شکل۳: مقایسه بصری نتایج مدل پیشنهادی و سایر مدل‌ها بر مجموعه داده جاده ماساچوست

۵- بحث و نتیجه‌گیری

در این مقاله، مدلی بـرای افـزایش کارآمـدی اسـتخراج جـاده‌هـا از تصـاویر سـنجش از دوری ارائـه گردیـد. در شبکه پیشنهادی، بلوک‌هـای *SCB* بـه‌طـور ویـژه بـرای استخراج جاده‌های کوچک در محیط‌های شهری دارای جاده‌هایی با ابعاد مختلف، طراحی شده‌انـد. بـا طراحـی ساختار *FEB*، مدل مـی‌توانـد اطلاعـات زمینـه را بهتـر

[Downloaded from jgit.kntu.ac.ir on 2025-07-13]
[DOI: 10.61186/jgit.12.1.83]

دریافت کرده و ویژگی‌های سطح پایین‌تر را حفظ نماید. در نتیجه مدل قادر است جزئیات بیشتری را در ماسک قطعه‌بندی بازیابی کند و عملکرد استخراج جاده را در محیط‌های پیچیده بهبود بخشد.

دو تابع زیان *LBCE* و *LDICE* تا حد معینی مکمل یکدیگر هستند. در زمینه کاربرد استخراج جاده، اجرای این تابع زیان جدید، بهینه‌سازی مقدار قابل‌توجهی از پیکسل‌های پس‌زمینه را از طریق *LBCE* در نظر می‌گیرد. به طور همزمان، رابطه مکانی بین نمونه‌های مثبت و منفی را با مدیریت تابع *LDICE* تأیید می‌کند. بنابراین، مشکل نمونه‌های نامتعادل کنترل می‌گردد. در مراحل اولیه پیش‌بینی، زمانی که اختلاف بین نتیجه پیش‌بینی‌شده و برچسب واقعی قابل‌توجه است؛ *LBCE* می‌تواند یک گرادیان برای کمک به همگرایی سریع مدل فراهم کند. همانطور که با ادامه آموزش شبکه، پیش‌بینی به برچسب واقعی نزدیک می‌شود؛ گرادیان *LBCE* کاهش می‌یابد. در مقابل *LDICE* می‌تواند یک گرادیان ثابت در این مرحله ارائه دهد که امکان بهینه‌سازی بهتر مدل را فراهم می‌کند.

ارزیابی جامع قابلیت روش پیشنهادی در این مقاله بر روی مجموعه داده ماساچوست جاده‌ها انجام شد و بدین منظور، مقایسه‌ها بین نتایج مدل پیشنهادی و مدل‌های پیشرفته *D-LinkNet*، *U-Net* و *DeepLabV3+* صورت گرفت. در این مقایسه، روش پیشنهادی در این مقاله توانست شاخص های *IOU*، امتیاز-اف و بازخوانی را به ترتیب به میزان ۱٫۹۸، ۳٫۰۳ و ۳٫۴۸ بهبود بخشد. نتایج تجربی نشان داد که مدل پیشنهادی در این مقاله از همه مدل‌های مرجع به‌ویژه برای جاده‌های کوچک مجاور با پوشش گیاهی و پس‌زمینه پیچیده، بهتر عمل می‌کند. بعلاوه، شبکه پیشنهادی با موفقیت عملکرد محاسباتی را متعادل می‌کند و نتایج مطلوبی را با پارامترهای قابل آموزش کمتر در مقایسه با روش‌های موجود ارائه می‌دهد. استفاده از قابلیت‌های سایر توابع زیان در شبکه پیشنهادی و امکان سنجی بهبود دقت نتایج به عنوان پیشنهادی برای تحقیقات آتی ارائه می‌گردد.

مراجع

[1] *Wang, W., Yang, N., Zhang, Y., Wang, F., Cao, T., Eklund, P. "A review of road extraction from remote sensing images," Journal of Traffic and Transportation Engineering (English Edition), Vol.3, No.3, PP. 271-282, 2016. https://doi.org/10.1016/j.jtte.2016.05.005.*

[2] *Huang, X., Zhang, L. "Road centreline extraction from high - resolution imagery based on multiscale structural features and support vector machines," International Journal of Remote Sensing, Vol. 30, No.8, pp. 1977-1987, DOI: 10.1080/01431160802546837.*

[3] *Bicego, M., Dalfini, S., Vernazza G., Murino, V. "Automatic road extraction from aerial images by probabilistic contour tracking," Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429), Barcelona, Spain, 2003, pp. III-585, doi: 10.1109/ICIP.2003.1247312.*

[4] *Baumgartner, A., Steger, C., Mayer, H., Eckstein, W., Ebner, H. "AUTOMATIC ROAD EXTRACTION BASED ON MULTI-SCALE, GROUPING, AND CONTEXT." Photogrammetric Engineering and Remote Sensing, Vol. 65, No.1999, pp. 777-785. 1999.*

[5] *Xu, Y., Xie, Z., Feng, Y., Chen, Z. "Road Extraction from High-Resolution Remote Sensing Imagery Using Deep Learning." Remote Sensing, Vol. 10, No.9, pp.1461, 2018. DOI:10.3390/rs10091461.*

[6] *Abdollahi, A., Pradhan, B., Shukla, N., Chakraborty, S., Alamri, A. "Deep Learning Approaches Applied to Remote Sensing Datasets for Road Extraction: A State-Of-The-Art Review," Remote Sens. Vol.12, No.1444, 2020. doi:10.3390/rs12091444.*

[7] *Choi, S., Do, M. "Development of the Road*

[Downloaded from jgit.kntu.ac.ir on 2025-07-13]

[DOI: 10.61186/jgit.12.1.83]

Pavement Deterioration Model Based on the Deep Learning Method," Electronics 2020, 9, 3; doi:10.3390/electronics9010003.

[8] Mnih, V., Hinton, G.E. "Learning to Detect Roads in High-Resolution Aerial Images," K. Daniilidis, P. Maragos, N. Paragios (Eds.): ECCV 2010, Part VI, LNCS 6316, pp. 210–223, 2010.

[9] Li, P., Zang, Y.,Wang, C., Li, J., Cheng, M., Luo, L., Yu, Y. "ROAD NETWORK EXTRACTION VIA DEEP LEARNING AND LINE INTEGRAL CONVOLUTION," 978-1-5090-3332-4/16/$31.00 ©2016 IEEE.

[10] Bastani, F., He, S., Abbar, S., Alizadeh, M., Balakrishnan, H., Chawla, S., Madden, S., DeWitt, D. "RoadTracer: Automatic Extraction of Road Networks from Aerial Images," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition.

[11] Yoo., H.J. "Deep Convolution Neural Networks in Computer Vision: a Review," IEIE Transactions on Smart Processing and Computing, 4(1), 2015 http://dx.doi.org/10.5573/IEIESPC.2015.4.1.035.

[12] Garcia-Garcia, A., Orts-Escolano, S., Oprea, S.O., Villena-Martinez, V., Garcia-Rodriguez, J. "A Review on Deep Learning Techniques Applied to Semantic Segmentation," Computer Vision and Pattern Recognition (cs.CV); Artificial Intelligence (cs.AI), https://doi.org/10.48550/arXiv.1704.06857.

[13] Ronneberger, O., Fischer, Ph., Brox, Th. "U-Net: Convolutional Networks for Biomedical Image Segmentation," N. Navab et al. (Eds.): MICCAI 2015, Part III, LNCS 9351, pp. 234–241, 2015. DOI: 10.1007/978-3-319-24574-4_28.

[14] Pal, K., Yadav, P., Katal, N. "RoadSegNet: a deep learning framework for autonomous urban road detection," Journal of Engineering and Applied Science, 2022. https://doi.org/10.1186/s44147-022-00162-9.

[15] Safavi, S.H., Sadeghi, M., Ebadpour, M. "DPRSMR: Deep learning-based Persian Road Surface Marking Recognition," J. Electr. Comput. Eng. Innovations, Vol. 11, No.2, PP. 409-418, 2023. Doi: 10.22061/jecei.2023.9496.627.

[16] Shao, Z., Tang, P., Wang, Zh., Saleem, N., Yam, S., Sommai, Ch. "BRRNet: A Fully Convolutional Neural Network for Automatic Building Extraction From High-Resolution Remote Sensing Images," Remote Sens. 2020, 12, 1050; doi:10.3390/rs12061050.

[17] Long, J., Shelhamer, E., Darrell, T. "Fully convolutional networks for semantic segmentation," in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015 pp. 3431-3440. doi: 10.1109/CVPR.2015.7298965.

[18] Alokasi, H., Ahmad, M.B. "Deep Learning-Based Frameworks for Semantic Segmentation of Road Scenes." Electronics 2022, 11, 1884. https://doi.org/10.3390/electronics11121884.

[19] Lin, G., Milan, A., Shen, Ch., Reid, J. "RefineNet: Multi-Path Refinement Networks for High-Resolution Semantic Segmentation," 2017 IEEE Conference on Computer Vision and Pattern Recognition, DOI 10.1109/CVPR.2017.549.

[20] Pan, X., Yang, F., Gao, L., Chen, Zh., Zhang, B., Fan, H., Ren, J. "Building Extraction from High-Resolution Aerial Imagery Using a Generative Adversarial Network with Spatial and Channel Attention Mechanisms," Remote Sens. 2019, 11, 917; doi:10.3390/rs11080917.

[21] Xu, Y., Wu, L., Xie, Z., Chen, Z. 2018. "Building Extraction in Very High Resolution Remote Sensing Imagery Using Deep Learning and Guided Filters," Remote Sens. 2018, 10, 144; doi: 10.3390/rs10010144.

[22] Wang, S., Hou, X., Zhao, X. "Automatic Building Extraction From High-Resolution

[Downloaded from jgit.kntu.ac.ir on 2025-07-13]

[DOI: 10.61186/jgit.12.1.83]

Aerial Imagery via Fully Convolutional Encoder-Decoder Network With Non-Local Block," in IEEE Access, vol. 8, pp. 7313-7322, 2020, doi: 10.1109/ACCESS.2020.2964043.

[23] Wei, S., Ji, S., Lu, M. 2020. "Toward Automatic Building Footprint Delineation From Aerial Images Using CNN and Regularization," in IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 3, pp. 2178-2189, March 2020, doi: 10.1109/TGRS.2019.2954461.

[24] Kumar, R., Amitab, K. 2023. "Semantic Segmentation of Road Scene Using Deep Learning," In: Basu, S., Kole, D.K., Maji, A.K., Plewczynski, D., Bhattacharjee, D. (eds) Proceedings of International Conference on Frontiers in Computing and Systems. Lecture Notes in Networks and Systems, vol 404. Springer, Singapore. https://doi.org/10.1007/978-981-19-0105-8_13.

[25] Yang, H. L., Yuan, J., Lunga, D., Laverdiere, M., Rose, A., Bhaduri, B. "Building Extraction at Scale Using Convolutional Neural Network: Mapping of the United States," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 11, no. 8, pp. 2600-2614, Aug. 2018, doi: 10.1109/JSTARS.2018.2835377.

[26] Shao, S., Xiao, L., Lin, L., Ren, C., Tian, J. "Road Extraction Convolutional Neural Network with Embedded Attention Mechanism for Remote Sensing Imagery," Remote Sens. 2022, 14, 2061. https://doi.org/10.3390/rs14092061.

[27] Hou, Y., Liu, Z., Zhang, T., Li, Y. 2021. " C-UNet: Complement UNet for Remote Sensing Road Extraction," Sensors 2021, 21, 2153. https://doi.org/10.3390/s21062153.

[28] Zhang, Z., Liu, Q., Wang, Y. 2018. " Road Extraction by Deep Residual U-Net," IEEE GEOSCIENCE AND REMOTE SENSING LETTERS.

[29] Zhou, L., Zhang, C., Wu, M. "D-linknet: Linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction," in Proc. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, USA, Jun. 2018, pp. 192–1924.

[30] Yu, F., Koltun, V. 2016. " MULTI-SCALE CONTEXT AGGREGATION BY DILATED CONVOLUTIONS," Published as a conference paper at ICLR 2016

[31] Chen, L.C., Papandreou, G., Schroff, F., Adam, H. 2017. "Rethinking Atrous Convolution for Semantic Image Segmentation," Computer Vision and Pattern Recognition, 2017. https://doi.org/10.48550/arXiv.1706.05587.

[32] Khan, T.M., Arsalan, M., Iqbal, Sh., Razzak, I., Meijering, E. 2023. "Feature Enhancer Segmentation Network (FES-Net) for Vessel Segmentation." Doi: 10.13140/RG.2.2.14160.48642.

[33] Zhong, Z., Li, J., Cui, W., Jiang, H. 2016. "Fully CONVOLUTIONAL NETWORKS FOR BUILDING AND ROAD EXTRACTION: PRELIMINARY RESULTS," https://ieeexplore.ieee.org/document/7729406.

[34] Kingma, D., Ba, J. "Adam: A Method for Stochastic Optimization," International Conference on Learning Representations. 2014.

[35] Mishra, P., Sarawadekar, K. 2019. "Polynomial Learning Rate Policy with Warm Restart for Deep Neural Network," TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), Kochi, India, pp. 2087-2092, doi: 10.1109/TENCON.2019.8929465.

[Downloaded from jgit.kntu.ac.ir on 2025-07-13]

[DOI: 10.61186/jgit.12.1.83]

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

Journal of Geospatial Information Technology
Vol.12, No.1, Spring 2024

Research Paper

A Deep Learning Approach based on the Specialized Convolutional Blocks in Urban Road Extraction

Mohammad Erfan Omati [1], Fatemeh Tabib Mahmoudi [2]*

1- MSc student of remote sensing in Department of Geomatics, Faculty of Civil Engineering, Shahid Rajaee Teacher Training University
2- Assisstant professor in Department of Geomatics, Faculty of Civil Engineering, Shahid Rajaee Teacher Training University

Abstract

Road extraction from high-resolution remote sensing images has been used in a wide range of applications such as traffic management, route planning, and road navigation. Due to their long length and small width, as well as shadows caused by vegetation and buildings, the detection of the roads challenging. As the roads In an area are of different types such as being near short passages, highways and motorways, we face some difficulties in automatic classifying and recognizing different kinds of roads. In order to improve the reliability and accuracy of extraction of the roads with shorter lengths when there are roads of different sizes; a neural network model is proposed in this paper that achieves pixel-accurate segmentation. The proposed network directly processes the input image and uses four specialized convolutional blocks (SCB) during down-sampling which is complemented by a shallow sampling approach to generate a binary mask for the road class. As the common semantic segmentation networks are deep and have various teachable parameters, the proposed network in this research uses shallow sampling which leads to lessen the network depth and as a result the number of the parameters decreases. The performance of the proposed model in this research was evaluated using the Massachusetts dataset, and the evaluation results clearly show the superior performance of the proposed model compared to the other neural networks with fewer parameters. Compared to the other neural networks such as DEEPLAB3+, U_NET and D_LINKNET, the proposed model was able to improve the IOU and F-Score indices in Massachusetts dataset by 1.98 and 3.03, respectively.

Key words: *Semantic Segmentation, road extraction, Deep Learning, specialized convolutional blocks.*

Correspondence Address *: Remote Sensing Group, Department of Geomatics, Faculty of Civil, Shahid Rajaee Teacher Training University, Tehran, Iran.*
Tel*: +98 21 22970021.*
Email*: fmahmoudi@sru.ac.ir*

[Downloaded from jgit.kntu.ac.ir on 2025-07-13]

[DOI: 10.61186/jgit.12.1.83]