

بهره‌گیری از یک شبکه عصبی پیچشی باقیمانده عمیق پالایش شده به منظور ارزیابی و مقایسه دقت آشکارسازی عارضه راه حاصل از تصاویر راداری سنتینل-۱ (مطالعه موردی: کلان‌شهرهای تهران و شیراز)

پرویز ضیائی‌ان فیروزآبادی^۱، سید هدایت شیخ‌قادری^{۲*}، منوچهر کلارستانی^۳

^۱ دانشیار گروه سنجش‌ازدور و سیستم اطلاعات جغرافیایی - دانشکده علوم جغرافیایی - دانشگاه خوارزمی تهران
p.zeaiean@gmail.com

^۲ کارشناس ارشد سنجش‌ازدور و سیستم اطلاعات جغرافیایی - دانشکده علوم جغرافیایی - دانشگاه خوارزمی تهران
hedayatp90@gmail.com

^۳ استادیار گروه آموزشی مهندسی برق و کامپیوتر - دانشکده فنی و مهندسی - دانشگاه خوارزمی تهران
kelarestaghi@khu.ac.ir

(تاریخ دریافت آبان ۱۴۰۰، تاریخ تصویب فروردین ۱۴۰۱)

چکیده

در سال‌های اخیر آشکارسازی و استخراج عارضه راه از تصاویر ماهواره‌ای با پیشرفت و توسعه‌ی الگوریتم‌های یادگیری عمیق در بخش تقسیم‌بندی معنایی بیش‌ازپیش مورد توجه پژوهشگران قرار گرفته است. در این راستا بیشتر مطالعات انجام‌شده در زمینه آشکارسازی و شناسایی عارضه راه با استفاده از تصاویر اپتیک بوده و در این میان مطالعات معدودی با استفاده از تصاویر راداری در سطح جهان انجام شده است. لذا هدف این مقاله بهره‌گیری از یک شبکه عصبی پیچشی باقیمانده عمیق پالایش‌شده (RDRCNN) به منظور ارزیابی و مقایسه دقت آشکارسازی عارضه راه حاصل از تصاویر راداری سنتینل-۱ در کلان‌شهرهای تهران و شیراز در شرایط برابر از نظر تعداد نمونه‌های آموزشی، اعتبارسنجی و معماری یکسان می‌باشد. در این تحقیق جهت آشکارسازی عارضه راه با استفاده از شبکه‌های عصبی عمیق (DNN)، از ترکیب رنگی ۷۷-VH تصاویر راداری سنتینل-۱ از ۸ شهر مختلف (تهران، مشهد، اصفهان، شیراز، تبریز، ارومیه، بغداد و پکن) بهره گرفته شد. در نهایت برای آموزش و آشکارسازی عارضه راه از مدل RDRCNN با داشتن یک واحد متصل باقی‌مانده (RCU) و یک واحد ادراک متسع شده (DPU) استفاده شد. یافته‌های پژوهش حاکی از آن است که مدل RDRCNN در فرایند شناسایی و آشکارسازی راه در دو شهر تهران و شیراز تقریباً یکسان عمل کرده و به‌طور کلی مدل فوق در شهر شیراز اندکی بهتر عمل نموده است؛ به صورتی که از نظر متریک‌های ارزیابی صحت برای تصاویر شهر تهران معیارهای کامل بودن ۵۷٫۶۶٪، صحت ۵۱٫۲۹٪، امتیاز F1 ۵۴٫۴۳٪ و دقت کلی ۹۲٫۷۸٪ و برای تصاویر شهر شیراز معیارهای کامل بودن ۶۰٫۷۷٪، صحت ۵۴٫۷۱٪، امتیاز F1 ۵۷٫۴۰٪ و دقت کلی ۹۵٫۶۳٪ به دست آمد. یافته‌های این پژوهش دقت پایین آموزش و آشکارسازی عارضه راه از تصاویر راداری سنتینل-۱ برای دو مورد از کلان‌شهرهای ایران را نشان می‌دهد. به‌طور کلی با مقایسه نتایج حاصل از این پژوهش و مطالعات قبلی می‌توان به این مورد پی برد که یکی از مهم‌ترین علل دقت پایین نتایج، کم‌عرض بودن راه‌ها در شهرهای ایران می‌باشد؛ اما به علت فقدان مطالعات لازم در زمینه آشکارسازی عارضه راه با تصاویر راداری سنتینل-۱، نمی‌توان با قاطعیت در مورد نتایج آن اظهار نظر کرد و پیشنهاد می‌شود مطالعات بیشتری در این زمینه انجام گیرد.

واژگان کلیدی: یادگیری عمیق، RDRCNN، سنتینل-۱، آشکارسازی راه، تهران، شیراز

۱- مقدمه

شناسایی و استخراج خودکار راه‌ها یک موضوع مهم و چالش‌برانگیز در سنجنش‌ازدور، فتوگرامتری، پردازش رقومی تصاویر و بینایی کامپیوتر می‌باشد [۱]. با این حال، توسعه قابل توجه فن‌آوری سنجنش‌ازدور در سال‌های اخیر، زمینه را برای آشکارسازی و استخراج خودکار راه از مناطق تحت پوشش گسترده را فراهم نموده است. به‌طور کلی ماکروسکوپی بودن، پویایی و فراهم آوردن داده‌های زمان واقعی از مهم‌ترین مزایای فن‌آوری سنجنش‌ازدور ماهواره‌ای می‌باشد. آشکارسازی و استخراج خودکار راه از تصاویر ماهواره‌ای در کاربردهایی همچون حمل‌ونقل، سیستم‌های هدایت خودکار وسایل نقلیه، مدیریت ترافیک، واکنش‌ها در مواقع وقوع بحران مورد استفاده قرار می‌گیرند [۳-۱]. در این میان، روش‌های تولید و استخراج اطلاعات مکانی در حوزه ژئوماتیک و نقشه‌برداری و در رأس آن‌ها فتوگرامتری و سنجنش‌ازدور بیشترین سهم در تولید نقشه و اطلاعات مکانی را بر عهده دارند. با در نظر گرفتن تصاویر مختلف (اپتیک و رادار)، مناطق مختلف (حومه یا روستا) و انواع مختلف راه‌ها (بزرگراه‌ها، راه‌های روستایی یا خیابان‌ها) الگوریتم‌های شناسایی و استخراج جاده را به چند نمونه طبقه‌بندی (ویژگی) مبتنی بر دانش، مورفولوژی ریاضی، مدل کانتور فعال، برنامه‌نویسی پویا و غیره طبقه‌بندی کرده‌اند [۴-۵]. در همین راستا، ظهور روش‌های یادگیری عمیق، در کنار دقت و سرعت پردازش بالا، نتایج قابل قبولی نیز از خود نشان می‌دهند. یادگیری عمیق، جزء آن دسته از روش‌هایی است که با استفاده از مدل‌های مختلف، اقدام به مشخص کردن موقعیت عوارض موجود در تصاویر می‌کند. این روش، مفاهیم انتزاعی سطح بالا اشیاء را از تصاویر، دریافت و آن‌ها را به لایه‌هایی با مفاهیم سطح پایین‌تر تبدیل می‌کند [۶]. از مزایای یادگیری عمیق، می‌توان به آموزش کاملاً خودکار ویژگی‌ها به صورت نظارت‌نشده و دقت بالا در نتایج خروجی و زمان آموزش پایین اشاره نمود [۷-۸]. اخیراً، با توسعه قدرت محاسبه (GPU)^۱ و مفهوم داده‌های بزرگ، شبکه عصبی پیچشی عمیق^۲ (DCNN) به‌طور گسترده‌ای در مطالعات تصویر مورد استفاده قرار گرفته و در این میان

قطعه‌بندی معنایی به‌منظور آشکارسازی و استخراج راه پیشرفت زیادی داشته است [۹-۱۰]. در واقع قطعه‌بندی معنایی باهدف برچسب‌گذاری هر پیکسل با احتمال مالکیت در طبقات مختلف انجام می‌شود که کار چالش‌برانگیزی است [۱۱]. با پیشرفت‌های چشم‌گیر در هوش مصنوعی، خصوصاً شبکه‌های عصبی عمیق، ایده‌های جدیدی برای آشکارسازی راه‌ها فراهم شده است [۱۲]. در این میان تحقیقات زیادی به‌منظور آشکارسازی و استخراج عارضه راه با استفاده از روش‌های یادگیری عمیق انجام گرفته است که در ادامه به‌مرور مختصری از آن‌ها پرداخته خواهد شد.

Xu و همکاران، (۲۰۱۸) مقاله‌ای با عنوان استخراج راه از تصاویر ماهواره‌ای با قدرت تفکیک بالا با استفاده یادگیری عمیق به انجام رساندند. مدل پیشنهادی باهدف استخراج اطلاعات محلی و جهانی راه‌ها در تصاویر سنجنش‌ازدور و بهبود دقت استخراج شبکه راه‌ها انجام گرفت. به‌طور کلی نتایج تحقیق نشان‌دهنده‌ی آن بود که مدل پیشنهادی آن‌ها یعنی Net - U - U - GL راه‌هایی که با مقیاس مختلف در تصاویر سنجنش‌ازدور برچسب‌زنی شده بودند، خوب عمل کرده و راه‌ها با موفقیت از طریق شبکه عصبی عمیق پیچشی شناسایی و استخراج شده است [۱۳]. در تحقیقی Henry و همکاران، (۲۰۱۸) به‌منظور قطعه‌بندی راه تصاویر راداری TerraSAR-X، از شبکه‌های عصبی کاملاً پیچشی استفاده کردند. آن‌ها عملکرد نسبی شبکه‌های اولیه و پیشرفته را پس از افزایش دقیق حساسیت آن‌ها نسبت به اشیاء نازک با اضافه کردن قواعد تلورانس فضایی، مطالعه نمودند. نتایج حاصله حاکی از آن بود که مدل‌های مورد استفاده آن‌ها نتایج امیدوارکننده‌ای را نشان داده که با موفقیت اکثر راه‌های مجموعه داده آزمایشی را استخراج کرده است. آن‌ها دریافتند که با تنظیم دقیق شبکه‌های عصبی کاملاً کانولوشن می‌توان به نتایج مناسبی در زمینه‌ی قطعه‌بندی راه‌ها در تصاویر راداری TerraSAR-X دست یافت [۱۴]. Gao و همکاران، (۲۰۱۹) در تحقیقی یک روش جدید برای شناسایی و استخراج راه‌ها از تصاویر ماهواره‌ای اپتیک با استفاده از یک شبکه عصبی پیچشی باقیمانده عمیق پالایش‌شده^۳ (RDRCNN) با یک مرحله پس پردازش پیشنهاد کردند. RDRCNN پیشنهادی آن‌ها

^۱ Graphics processing unit^۲ deep convolutional neural network^۳ Refined Deep Residual Convolutional Neural Network

مطالعه مقایسه‌ای با هشت روش یادگیری عمیق که اخیراً ایجاد شده بودند در استخراج راه برتر و کاربردی‌تر است [۱۶]. Lian و همکاران، (۲۰۲۱) در پژوهشی با بهره‌گیری از روش‌های یادگیری عمیق به آشکارسازی و استخراج عارضه‌ی راه در مناطق مختلف شهری، بایر، کوهستانی و زراعی در کشور چین از تصاویر راداری و نوری سنتینل-۱ و سنتینل-۲ پرداختند. آن‌ها در پژوهش خود از یک مدل متحد U-Net (UU-Net) برای ترکیب تصاویر سنتینل-۱ و ۲ برای استخراج راه استفاده نمودند. یافته‌های پژوهش حاکی از آن بود که UU-Net دقت بهتری نسبت به شبکه‌های کانولوشن عمیق سنتی با داده‌های نوری یا SAR داشته است. به‌طور کلی نتایج شبکه‌ها نشان داد که SAR در استخراج راه در زمین‌های بایر موفق‌تر و تصاویر نوری در مناطق کوهستانی دقت بالاتری از خود نشان داده است. در این راستا دقت راه‌های استخراجی به‌طور کلی ارتباط مثبتی با مناطق کوهستانی (پرشیب) و مناطق شهری دارد [۱۷].

در این میان علاوه بر تحقیقات فوق که به آن اشاره شد مطالعات دیگری هم به‌منظور آشکارسازی و استخراج عارضه راه در مناطق شهری با استفاده از شبکه‌های عصبی عمیق^۳ (DNN) انجام گرفته است [۱۸-۲۷].

همان‌طور که از ادبیات تحقیق هم قابل استناد است اکثر مطالعات انجام شده در زمینه‌ی آشکارسازی و استخراج عارضه راه با بهره‌گیری از شبکه‌های عصبی عمیق با استفاده از تصاویر اپتیک بوده و در این راستا مطالعات معدودی با استفاده از تصاویر راداری در سطح جهان انجام شده است. در همین راستا در کنار مطالعات محدود انجام شده، تاکنون موارد زیر مورد توجه محققین قرار نگرفته است:

- ۱) استفاده از تصاویر شهرهای مختلف به‌عنوان داده‌های آموزشی
- ۲) عدم استفاده از دیگر مدل‌های مبتنی بر شبکه‌های عصبی عمیق به‌منظور آشکارسازی عارضه راه از تصاویر راداری سنتینل-۱
- ۳) عدم آشکارسازی عارضه راه در شهرهای مختلف با شرایط یکسان از نظر تعداد نمونه‌های آموزشی، اعتبارسنجی و معماری شبکه

شامل یک واحد متصل باقی‌مانده^۱ (RCU) و یک واحد ادراک متسع‌شده^۲ (DPU) بود. آن‌ها در تحقیق خویش از مورفولوژی ریاضی و یک الگوریتم رأی‌دهی تانسور برای بهبود عملکرد RDRCNN در طول پس‌پردازش استفاده کردند. آزمایش‌ها بر روی دو مجموعه داده از تصاویر با وضوح بالا انجام گرفت تا عملکرد معماری‌های شبکه پیشنهادی را نشان داده و نتایج معماری‌های پیشنهادی با معماری‌های شبکه دیگر مقایسه شوند. نتایج نشان‌دهنده کارایی مؤثر روش پیشنهادی برای آشکارسازی و استخراج راه‌ها از یک محیط پیچیده بود [۱۵]. Li و همکاران، (۲۰۱۹) در مقاله‌ای، روشی به‌منظور بهبود شبکه‌های عصبی برای استخراج راه‌ها از تصاویر سنجش‌ازدوری پهناد پیشنهاد دادند. آن‌ها با تمرکز بر روی مسئله کارایی محاسباتی پایین D-LinkNet، به پیشرفت‌هایی رسیدند. در نهایت با توجه به برخی تغییرات، شبکه عصبی بهبودیافته دیگری بانام B-D-LinknetPlus ساخته و مقایسه بین شبکه‌های عصبی انجام دادند. نتایج نهایی نشان داد که شبکه‌های عصبی بهبودیافته در کاهش اندازه شبکه و ایجاد دقت موردنیاز برای استخراج راه مثرتر است [۱۰]. Zhang و همکاران، (۲۰۱۹) یک مدل خودکار (U-FCN Net) مبتنی بر DNN را برای استخراج جاده‌ها از تصاویر با پلاریزاسیون Sentinel1-SAR (VV-VH) پیشنهاد دادند. آن‌ها به‌طور دستی داده‌های واقعیت زمینی را با کمک گوگل‌مپ برچسب‌زنی کردند. برای مقایسه CNN اولیه، چهار روش دیگر برای تأیید عملکرد داده‌های SAR مورد استفاده قرار گرفتند. نتایج آزمایش نشان داد که شبکه پیشنهادی (fcn) به‌طور قابل‌توجهی از روش‌های یادگیری ماشین سنتی در دقت و کمتر شدن زمان محاسباتی، نسبت به CNN پیشی گرفته است [۴]. Ren و همکاران، (۲۰۲۰) در پژوهشی یک کیسول جدید (DA - CapsUNet) با توجه به روش U-Net برای استخراج عارضه راه با ترکیب ویژگی‌های مفید نمایش کیسول و ویژگی‌های قدرتمند مکانیسم دقت، پیشنهاد دادند. نتایج حاصله حاکی از آن بود که معیارهای کامل‌بودن، صحت و امتیاز F1 به ترتیب ۹۴٫۸۶٪، ۹۵٫۲۳٪ و ۹۵٫۰۴٪ به‌دست آمده است. ارزیابی‌های کمی در یک مجموعه داده بزرگ نشان داد که DA - CapsUNet در

^۱ Residual Connected Unit

^۲ Dilated Perception Unit

^۳ Deep neural networks

هرکدام از مراحل یادشده در شکل (۱) در ادامه توضیح داده خواهند شد.

۱-۲- پیش‌پردازش و آماده‌سازی تصاویر راداری سنتینل-۱

در این مطالعه از ۸ فریم تصویر GRD از شهرهای مختلف حاصل از سنجنده راداری سنتینل-۱ که در یک‌زمان مشابه تصویربرداری شده‌اند استفاده می‌گردد (جدول ۱).

جدول ۱- مشخصات تصاویر مورد استفاده (راداری سنتینل-۱)

تصاویر	باند	قطبش	رزولوشن	زمان
تهران	C	VV-VH	۱۰ متر	۲۰۲۰/۶
شیراز				
مشهد				
اصفهان				
تبریز				
ارومیه				
بغداد				
پکن				

به‌منظور استفاده از این تصاویر در ادامه کار، یک سری مراحل پیش‌پردازش روی آن‌ها انجام گرفت که به شرح زیر است:

- (۱) تصحیح رادیومتریک
 - (۲) حذف و کاهش نویز اسپیکل
 - (۳) تصحیح هندسی
 - (۴) تبدیل دی سی بل
- به‌منظور حذف نویز اسپیکل از فیلتر لی استفاده می‌شود [۲۸ و ۴]. در این میان جهت حفظ وضوح و عدم کاهش جزئیات تصاویر از اندازه پنجره پیکسل ۳×۳ استفاده می‌گردد. همچنین جهت تصحیح هندسی تصاویر فوق از مدل رقومی ارتفاعی ۹۰ متری SRTM استفاده خواهد شد؛ و نهایتاً جهت تبدیل دی سی بل، ضرایب برگشت رادار (σ^0) به‌عنوان سیگما ۰ یا سیگما ناشناخته) در واحد دی سی بل (db) از مقادیر دیجیتالی، با استفاده از معادله-۱) به دست می‌آید [۴].

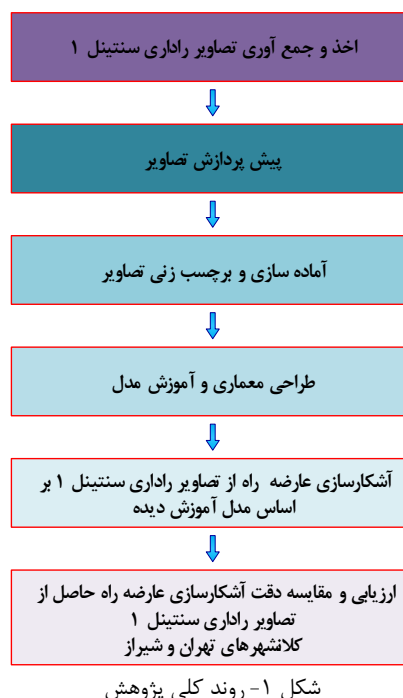
$$\sigma^0 = \text{Log}_{10} [DN^2/A_2] \quad (1)$$

که در آن A عامل کالیبراسیون است و از مقادیر سیگما در جدول جستجو (LUT) می‌توان آن را یافت.

لذا با در نظر گرفتن موارد یادشده، هدف این مقاله بهره‌گیری از روش‌های یادگیری عمیق به‌منظور ارزیابی و مقایسه دقت شناسایی و آشکارسازی عارضه راه درون‌شهری حاصل از تصاویر راداری سنتینل-۱ می‌باشد. درواقع این تحقیق به دنبال جواب این سؤال است که در شرایط یکسان از نظر تعداد نمونه‌های آموزشی، اعتبارسنجی و معماری و همچنین استفاده از تصاویر راداری سنتینل-۱ از شهرهای متفاوت، با چه دقتی می‌توان عارضه راه را در کلان‌های تهران و شیراز به‌عنوان مطالعه موردی شناسایی و آشکار نمود.

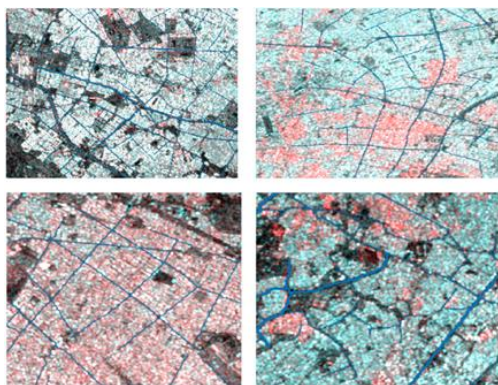
۲- مواد و روش

در این تحقیق از تصاویر راداری سنتینل-۱ کلان‌شهرهای تهران و شیراز برای تست و تصاویر چند شهر دیگر که در ادامه به آن اشاره خواهد شد برای آموزش و اعتبارسنجی شبکه جهت تشخیص و آشکارسازی عارضه راه استفاده می‌گردد. همچنین در پژوهش حاضر از مدل RDRCNN به‌منظور آموزش و آشکارسازی عارضه راه از تصاویر فوق بهره گرفته خواهد شد. به این صورت که مدل فوق روی تصاویر راداری سنتینل-۱ شهرهای تهران و شیراز پیاده‌سازی شده و نتایج باهم مقایسه خواهند گردید. به‌طورکلی روند شناسایی و آشکارسازی عارضه راه از تصاویر راداری سنتینل-۱ در شکل (۱) نشان داده شده است.



برچسب‌گذاری شده که به‌عنوان واقعیت زمینی^۲ از روی تصاویر برداشت شده‌اند با فرمت گرافیک شبکه قابل‌حمل^۳ (PNG) ذخیره‌سازی گردید.

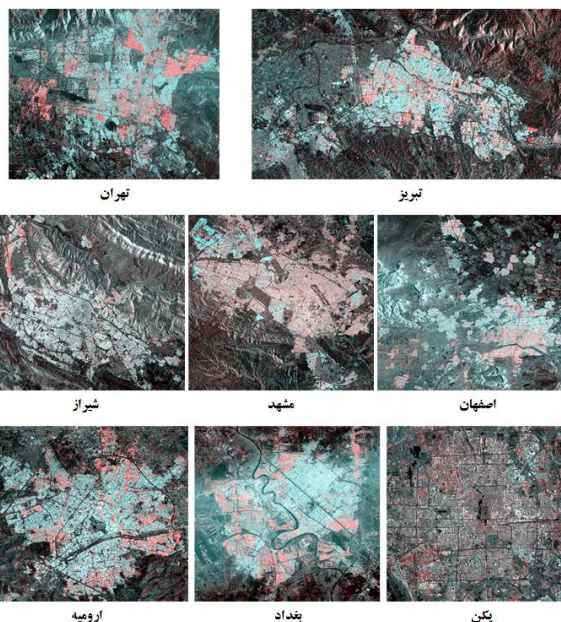
در شکل (۳) نمونه‌ای از تصاویر که پیکسل‌های مربوط به عارضه راه در آن برچسب‌زنی شده‌اند، ارائه شده است.



شکل ۳- نمونه‌هایی از تصاویر برچسب‌گذاری شده سنتینل-۱ مورد استفاده از شهرهای مختلف

طبق تئوری یادگیری عمیق اندازه قطعات تصویر روی نتایج نهایی مؤثر است [۴]؛ لذا پس از پیش‌پردازش و آماده‌سازی تصاویر و همچنین برچسب‌زنی پیکسل‌های مربوط به کلاس راه، با در نظر گرفتن قدرت تفکیک مکانی تصاویر و همچنین وضوح آن‌ها برای قابلیت تشخیص عارضه راه توسط الگوریتم، اقدام به تعیین اندازه قطعات تصاویر مورد استفاده گردید. در این راستا پس از انجام سعی و خطا نهایتاً با در نظر گرفتن مطالعه Zhang و همکاران، (۲۰۱۹) اندازه هرکدام از قطعات تصویر ۲۵۶ در ۲۵۶ تعیین گردید [۴]. به این صورت تمام ۸ فریم تصویر با تعریف دستور عملی به‌اندازه‌ی ۲۵۶ در ۲۵۶ پیکسل تبدیل شدند. بنابراین در این تحقیق از قطعات یکسان با اندازه‌ی سطر و ستون ۲۵۶×۲۵۶ پیکسل استفاده می‌گردد. در این میان از کل ۸ فریم کامل تصاویر از مناطق شهری، تعداد ۲۰۱۳ قطعه با اندازه‌های ۲۵۶×۲۵۶ پیکسل به دست آمد که پس جداسازی قطعات نامناسب، نهایتاً تعداد نمونه‌های آموزشی، نمونه‌های اعتبارسنجی و نمونه‌های تست برای شهرهای تهران و شیراز آماده‌سازی شدند (جدول ۲).

از آنجاکه طبق تحقیق Zhang و همکاران، (۲۰۱۹) تصاویر راداری چندقطبشی نسبت به تصاویر تک‌قطبشی نتایج مناسب‌تری در آشکارسازی و استخراج عارضه راه با استفاده از روش‌های یادگیری عمیق داشته است، بنابراین در این پژوهش از ترکیب رنگی دو قطبش VV-VH استفاده می‌شود [۴]. به این صورت که قطبش VV به رنگ قرمز مانتیور و قطبش VH به رنگ‌های سبز و آبی مانتیور اختصاص داده می‌شوند (شکل ۲).



شکل ۲- تصاویر سنتینل-۱ مورد استفاده از شهرهای مختلف

پس از آماده‌سازی داده‌ها ابتدا تصاویر مورد نظر با فرمت فایل تصویری برچسب‌زده شده^۱ (Tiff) ذخیره‌سازی گردیده و در ادامه به‌منظور برچسب‌گذاری عارضه راه داخل تصاویر راداری از تصاویر گوگل ارث کمک گرفته شد. به این صورت که با روی هم‌گذاری هرکدام از تصاویر راداری روی تصویر متناسب با منطقه‌ی موردنظر روی گوگل ارث، نمونه‌برداری انجام گرفت. در این میان شیوه‌ی نمونه‌برداری به این صورت بود که تمام پیکسل‌هایی که متعلق به عارضه راه بودند برچسب راه به خود گرفتند. لذا برای این منظور یعنی تشخیص و شناسایی عارضه راه در داخل تصاویر از قطعه‌بندی معنایی استفاده می‌شود. در این راستا پس از برچسب‌گذاری پیکسل‌های مربوط به کلاس راه در تمام تصاویر مورد استفاده، فرمت فایل‌های

^۲ Ground True

^۳ Portable Network Graphic

^۱ Tagged Image File Format

جدول ۲- تعداد قطعات مناسب به‌دست‌آمده پس از جداسازی قطعات نامناسب

تصویر	آموزش	تست	اعتبارسنجی
تهران	۱۵۰۰	۱۳۵	۱۰۰
شیراز		۹۰	

به‌طور کلی پس از جداسازی و آماده کردن قطعات نهایی، از تصاویر شهرهای تهران و شیراز به‌عنوان داده‌های تست، در میان ۶ فریم تصویر باقی‌مانده تعداد ۱۰۰ قطعه تصویر که با دقت بالاتری برچسب‌گذاری شده بودند به‌عنوان داده‌های اعتبارسنجی و در نهایت بقیه‌ی قطعات باقی‌مانده یعنی ۱۵۰۰ قطعه هم به‌عنوان داده‌های آموزش به شبکه معرفی می‌شوند.

۲-۲-۲ مدل RDRCNN

مدل RDRCNN برای اولین بار توسط Gao و همکاران (۲۰۱۹) به‌منظور استخراج عارضه راه با استفاده از تصاویر با قدرت تفکیک مکانی بالا، پیشنهاد داده شد [۱۵]. معماری RDRCNN یک ساختار آموزش متقارن پایان به انتها جهت پیش‌بینی سطح پیکسل است. فرایند یادگیری پایان به انتها یک نوع فرایند یادگیری عمیق است که در آن همه پارامترها به‌جای گام‌به‌گام به‌طور مشترک آموزش داده می‌شوند؛ به این صورت که برخلاف رویکرد سنتی مستقیماً از ورودی به خروجی موردنظر می‌رود. معماری RDRCNN از نتایج حاصل از مدل‌های ResNet [۲۹]، U Net [۳۰] و Deeplab [۳۱] الهام گرفته شده است. RDRCNN از دو هسته مرکزی RCU و DPU و به دنبال آن از یک لایه پیچشی کامل تشکیل شده است [۱۵]. این معماری از سه قسمت طراحی شده است. قسمت اول برای استخراج ویژگی‌ها با استفاده از برخی RCUها با ساختار کوچک‌تر از طریق برخی از عملگرهای حداکثر تجمع ایجاد شده‌اند. قسمت دوم یعنی DPUs برای بزرگ کردن میدان دید (FOV) بدون از دست دادن وضوح تصویر با استفاده از واحدهای متصل متوالی در چند مقیاس است. قسمت سوم یک ساختار گسترده برای تولید نقشه آشکارسازی و استخراج راه است که به همان اندازه ورودی می‌باشد [۱۵]. لذا برای بهبود قابلیت بیان، عدم کاهش وضوح مکانی تصاویر در فضای ویژگی و جلوگیری از موانع و مشکلات

مربوط به گرادیان افزایشی و گرادیان کاهشی^۱ از RCU و DPU چندمقیاسی استفاده می‌شود [۱۵ و ۳۱]. در ادامه به توضیح RCU و DPU پرداخته خواهد شد.

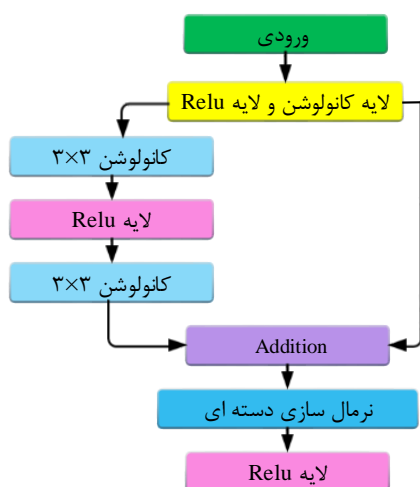
۲-۲-۱ ساختار واحد متصل باقی‌مانده (RCUs)

RCU یک شبکه عمیق است که در آن بخشی از داده‌های ورودی به لایه‌های بعدی پاس داده می‌شوند. این ویژگی، به این شبکه‌ها این امکان را می‌دهد که حقیقتاً عمیق باشند، اما در واقع، نوعی از شبکه‌های پیچشی عمیق بدون تأخیر صریح هستند. در واقع شبکه‌های عصبی باقیمانده این کار را با استفاده از اتصالات عرضی یا میان‌بر برای پریدن روی بعضی از لایه‌ها انجام می‌دهند.

ساختار شبکه فوق توسط معادله‌ی زیر قابل بیان است:

$$x_{l+1} = \omega_l + 1^{\sigma}(\omega_l x_l) + x_l \quad (2)$$

که در آن x_{l+1} بردارهای خروجی و x_l بردارهای ورودی لایه‌های موردنظر را نشان می‌دهند. همچنین تابع σ (.) نشان‌دهنده‌ی ReLU است. در واقع معادله‌ی فوق به‌وسیله‌ی یک اتصال میان‌بر و جمع عنصر انجام می‌شود و سپس دومین تابع فعال هم به آن اضافه می‌شود. هرچند He و همکاران، (۲۰۱۵) مطالب زیادی درباره تأثیرات ترکیبات مختلف RCU ارائه داده‌اند [۲۹]، اما در این مقاله، همچون مطالعه‌ی Gao و همکاران، (۲۰۱۹) از واحد متصل میان‌بر همان‌طور که در شکل (۴) نشان داده شده است، بدون تغییر استفاده می‌شود [۱۵].



شکل ۴- ساختار واحد متصل باقی‌مانده (RCUs)

^۱ exploding gradient and vanishing gradient

جدول ۴- ساختار واحد ادراک متسع شده (DPU)

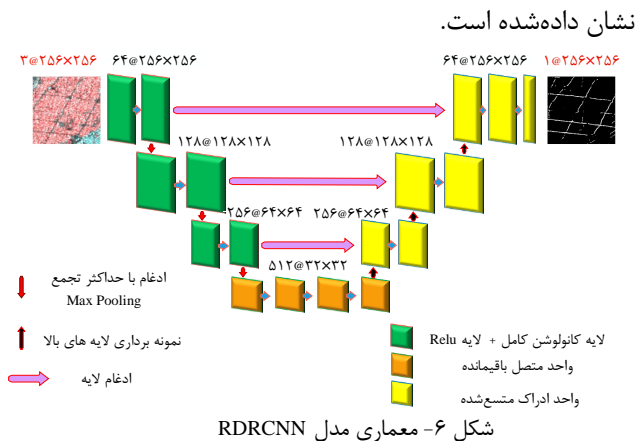
معیار	اندازه فیلتر	سایز تصویر	مقیاس متسع شده
DPU-۱	۳×۳	۵۱۲	۶
DPU-۲	۳×۳	۵۱۲	۱۲
DPU-۳	۳×۳	۵۱۲	۱۵
DPU-۴	۳×۳	۵۱۲	۲۴

همان‌طور که در شکل (۴) قابل‌مشاهده است برای جلوگیری از فاصله و شکاف بیش‌ازحد، یک لایه نرمال-سازی دسته‌ای به انتهای واحد پایه اضافه‌شده است [۳۲]. جزئیات بیشتر در مورد آن در جدول (۳) نشان داده‌شده است.

جدول ۳- جزئیات ساختار واحد متصل باقی‌مانده (RCUs)

واحد	لایه	اندازه فیلتر
RC-۱	لایه کانولوشن و لایه Relu	۱×۱
RC-۲	لایه کانولوشن و لایه Relu	۳×۳
RC-۳	لایه کانولوشن	۳×۳
RC-۴	Addition	-
RC-۵	نرمال‌سازی دسته‌ای	-
RC-۶	لایه Relu	-

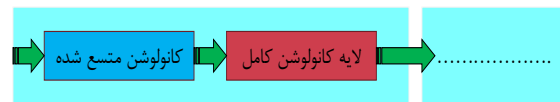
به‌طورکلی معماری شبکه RDRCNN در شکل (۶)



۲-۲-۱- ساختار واحد ادراک متسع شده (DPU)

به‌منظور پذیرنده بودن و تفکیک مکانی بالا از DPUs استفاده می‌شود [۳۳]. درواقع این معماری نه‌تنها وضوح مکانی را حفظ می‌کند بلکه ورودی دریافت شبکه را هم بزرگ‌تر در نظر می‌گیرد [۳۴].

همان‌طور که در شکل (۵) قابل‌مشاهده است یک DPU از یک‌لایه پیچشی متسع‌شده و یک لایه پیچشی کامل تشکیل‌شده است که پیچشی متسع‌شد از کرنل‌های خاص با وزن‌های کم برای بزرگ کردن FOV استفاده می‌کند و پیچشی کامل روابط همسایگی میان آن‌ها را حفظ می‌کند. هم اندازه کرنل و هم فاصله وزن‌های کم، با افزایش ضریب انبساط به‌صورت نمایی افزایش می‌یابند؛ بنابراین با افزایش ضریب DPU، FOV نیز به‌صورت نمایی افزایش می‌یابد [۱۵ و ۳۵].



شکل ۵- ساختار واحد ادراک متسع شده (DPU)

همان‌طور که در شکل (۶) مشاهده می‌شود، مربع سبزرنگ، لایه‌های کانولوشن کامل و یک لایه غیرخطی ReLU که از توابع فعال‌سازی غیرخطی هستند را نشان می‌دهد. درواقع توابع غیرخطی به مدل کمک می‌کنند تا نگاشت‌های پیچیده را میان ورودی‌ها و خروجی‌های شبکه به وجود آورد؛ به‌عبارت‌دیگر، این توابع به مدل این امکان را می‌دهند تا خود را با داده‌های پیچیده و غیرخطی وفق دهد. این موضوع برای داده‌های پیچیده‌ای، مانند تصاویر بسیار مهم است. مربع‌های نارنجی و زردرنگ به ترتیب نشان‌دهنده واحد متصل باقی‌مانده و واحد ادراک متسع شده می‌باشند که بخش‌های قبل به تشریح آن‌ها پرداخته شد.

فلش‌های قرمز رنگ نشان‌دهنده لایه‌های ادغام^۱ با حداکثر تجمع^۲ می‌باشد که روش max pooling از حداکثر مقدار بین خوشه‌های نورونی در لایه پیشین استفاده می‌کند. یک لایه ادغام معمولاً بعد از یک لایه پیچشی قرار می‌گیرد و از آن برای کاهش اندازه ویژگی نگاشت‌ها و پارامترهای شبکه می‌توان استفاده کرد. فلش‌های سیاه و بنفش‌رنگ به ترتیب به‌منظور نمونه‌برداری از لایه‌های بالا و ادغام لایه‌ها به کار برده می‌شود.

در این تحقیق، همچون مطالعات Yu و Koltun، (۲۰۱۵) و Gao و همکاران، (۲۰۱۹) چهار مقیاس برای DPU طراحی خواهد شد [۱۵ و ۳۳]. جزئیات پارامترهای مورداستفاده در جدول (۴) آورده شده است.

^۱ pooling
^۲ Max Pooling

۳- پیاده‌سازی و ارزیابی

ارزیابی یک شبکه، نیازمند ارزیابی دقت آن مدل بر روی کلاس و عارضه‌ی مورد مطالعه است. برای برچسب-گذاری عارضه و کلاس راه مورد نظر در این تحقیق از تصاویر راداری (سنتینل-۱) استفاده شد؛ به طوری که از انواع مختلف راه‌ها از نظر پهنا، اندازه و اشکال گوناگون در ۸ شهر مختلف استفاده گردید.

۳-۱- آموزش و بهینه‌سازی مدل

در تحقیق حاضر به منظور آموزش و پیاده‌سازی مدل-های مورد نظر از زبان برنامه‌نویسی پایتون و کتابخانه متن‌باز تنسورفلو^۱ استفاده گردید [۱۵ و ۳۶]. برای آشکارسازی عارضه راه، تابع ضرر^۲ با استفاده از وزن‌های مختلف مربوط به هر کلاس (راه و غیر راه) از معادله‌ی (۳) انجام می‌گیرد.

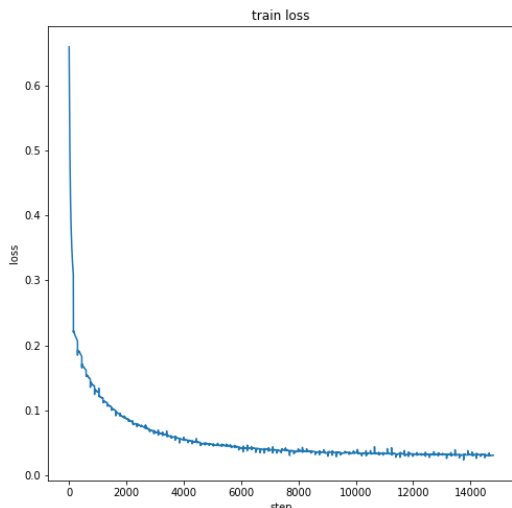
$$loss(y^{\wedge}, y) = -\frac{1}{N} \sum_{i=1}^N y^{(i)} \log y^{\wedge(i)} + (1 - y^{(i)}) \log(1 - y^{\wedge(i)}) \quad (3)$$

که در آن y^{\wedge} برابر با مقداری که شبکه تشخیص داده است. y برابر با مقدار واقعی (پیکسل‌های برچسب‌گذاری شده)

N تعداد کل تصاویر آموزشی است.

تابع ضرر یا تابع هزینه^۳ در واقع میزان خطا در هر بار اجرای شبکه‌ی عصبی را برای داده‌های آموزشی نمایش می‌دهد؛ به این صورت که شبکه با مشاهده مقدار ضرری که در هر بار اجرا داشته است، یاد می‌گیرد که چه مقدار بایستی وزن‌ها و انحراف را به‌نگام^۴ کند.

در این پژوهش فرایند آموزش مدل با بهینه‌سازی خطای آنتروپی متقاطع^۵ با استفاده از الگوریتم Adam انجام می‌شود که به‌طور خاص برای آموزش شبکه‌های عصبی عمیق طراحی شده است [۳۷]. نتایج حاصل از الگوریتم بهینه‌سازی Adam برای تصاویر راداری مربوط به مدل RDRCNN در شکل (۷) نشان داده شده است.



شکل ۷- منحنی تکرار بر روی مجموعه تصاویر (سنتینل-۱): خطای آموزش (محور افقی: تعداد تکرار، محور عمودی: اختلاف اندازه‌گیری)

به‌طور کلی بهترین و مناسب‌ترین مدل زمانی به دست می‌آید که در هنگام آموزش، تابع LOSS کمترین مقدار را داشته باشد.

۳-۲- معیارهای ارزیابی دقت مدل

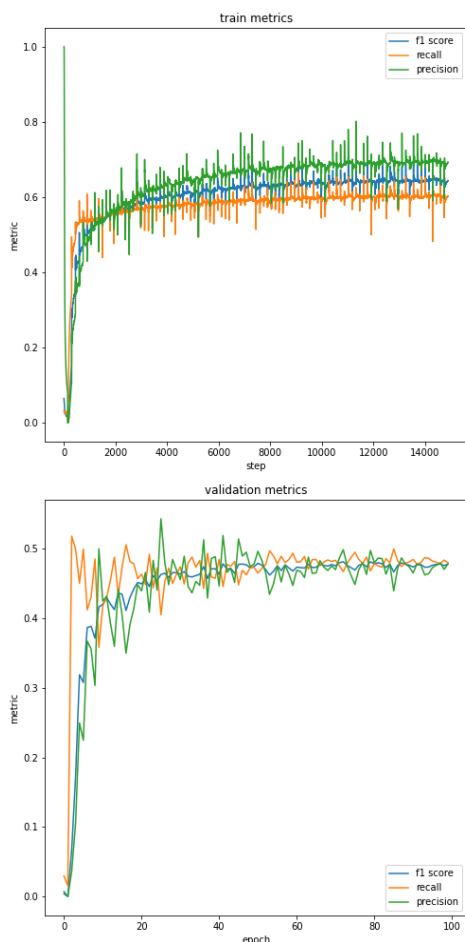
برای نمایش میزان دقت الگوریتم‌های مورد نظر از معیارهای صحت^۶، کامل بودن^۷، دقت کلی^۸ و امتیاز^۹ F_1 که ترکیبی از دو معیار صحت و کامل بودن می‌باشد، استفاده می‌گردد [۳۸-۳۹]. نتایج به‌دست آمده بعد از تحلیل به دو گروه داده مثبت و منفی تقسیم و نهایتاً شبکه برای هر هدف، مقادیری با عناوین FN، TN، TP و FP اختصاص داده می‌شود که بیانگر موارد زیر هستند:

- (۱) مثبت درست^{۱۰}: یک معیار مثبت است؛ یعنی راه‌هایی که به‌درستی توسط شبکه تشخیص داده شده است.
- (۲) مثبت اشتباه^{۱۱}: یک معیار منفی است؛ یعنی به نادرستی توسط شبکه، کلاس راه تشخیص داده شده است.
- (۳) منفی درست^{۱۲}: یک معیار مثبت است؛ یعنی کلاس غیر راهی که به‌درستی توسط شبکه رد شده است.

^۶ Precision
^۷ Recall
^۸ overall accuracy
^۹ F1 score
^{۱۰} True Positive
^{۱۱} False Positive
^{۱۲} True Negative

^۱ TensorFlow
^۲ loss function
^۳ cost function
^۴ update
^۵ binary cross entropy loss function

روند آموزش و اعتبارسنجی مدل RDRCNN با استفاده از معیارهای Recall, Precision, F₁ score برای تصاویر راداری سنتینل-۱ در شکل (۸) نشان داده شده است.



شکل ۸- نتایج شبکه با استفاده از متریک‌های ارزیابی صحت (تصاویر سنتینل-۱): بر روی داده‌های آموزشی (بالا) و داده‌های اعتبارسنجی (پایین)

با توجه به شکل (۸) هر چه روند متریک‌ها در طول دوره با معیارهای یادشده بیشتر باشد دقت مدل هم بیشتر خواهد بود.

۴- یافته‌ها و نتایج

پس از آموزش شبکه با مدل RDRCNN، به‌منظور آشکارسازی عارضه راه از تصاویر راداری سنتینل-۱ محدوده‌ی شهر تهران و شیراز و انجام سعی و خطا در روند آموزش شبکه، نهایتاً خروجی برای هرکدام از شهرهای موردنظر به دست آمد. برای نمایش خروجی‌های موردنظر از هرکدام از شهرهای تهران و شیراز، نتایج تعدادی از قطعات به‌عنوان نمونه از مناطق مختلف از شهرهای فوق به نمایش گذاشته خواهد شد. همان‌طور که در شکل (۹) قابل مشاهده

(۴) منفی اشتباه^۱: یک معیار منفی است؛ یعنی کلاس راهی که به نادرستی توسط شبکه، غیر راه شناسایی شده است.

(۵) درنهایت بر اساس معیارهای فوق یک ماتریس ابهام همانند جدول (۵) تشکیل خواهد شد:

جدول ۵- ماتریس ابهام

	راه	غیر راه
راه	TP	FP
غیر راه	FN	TN

لذا بر اساس تعاریف فوق روابط زیر مطرح خواهد بود: کامل بودن (R) را می‌توان در قالب تعداد مثبت‌های بازگردانده شده توسط مدل تعریف کرد. با کمک رابطه زیر و با استفاده از ماتریس ابهام، می‌توان آن را محاسبه کرد.

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

صحت (P) استفاده‌شده در بازیابی اسناد می‌تواند به‌صورت تعداد اسناد صحیح بازگردانده شده توسط مدل، تعریف شود. با کمک رابطه زیر و با استفاده از ماتریس ابهام، می‌توان آن را به‌سادگی محاسبه کرد.

$$Precision = \frac{TP}{TP + FP} \quad (۵)$$

امتیاز F₁، میانگین هارمونیک از صحت و کامل بودن را به ما می‌دهد. به بیان ریاضی، امتیاز F₁ میانگین وزن‌دار از صحت و کامل بودن (P و R) است. بهترین مقدار برای F₁، ۱ و بدترین مقدار ۰ است. با کمک رابطه زیر می‌توان امتیاز F₁ را محاسبه کرد.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (۶)$$

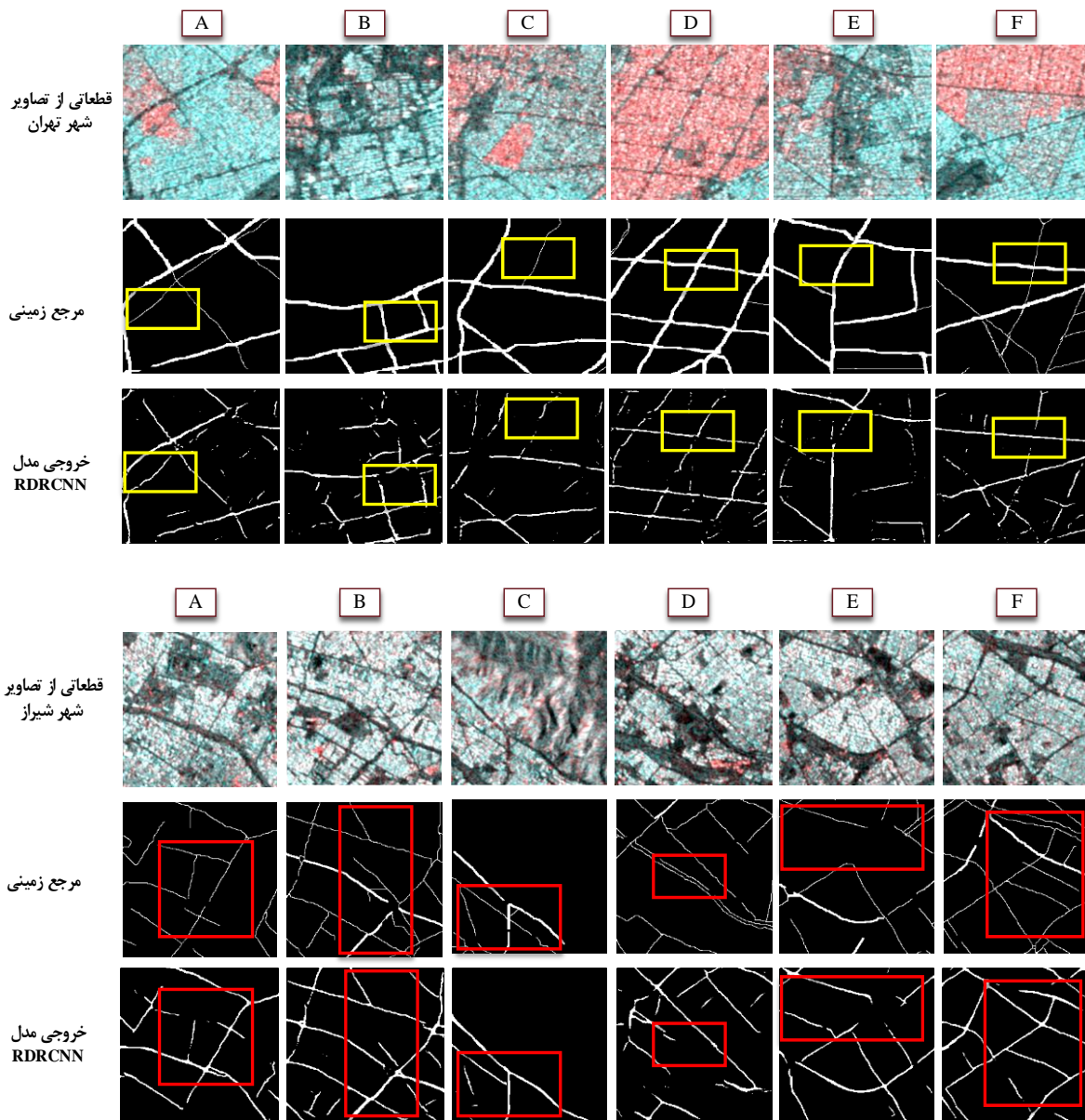
دقت کلی (OA) متداول‌ترین معیار کارایی برای الگوریتم‌های طبقه‌بندی است. می‌توان آن را در قالب تعداد پیش‌بینی‌های صحیح صورت گرفته به نسبت همه پیش‌بینی‌های صورت گرفته، تعریف کرد. با کمک رابطه زیر و با استفاده از ماتریس ابهام، می‌توان آن را محاسبه کرد.

$$OA = \frac{TP + TN}{TP + TN + FN + FP} \quad (۷)$$

^۱ False Negative

است، خروجی حاصله از تصاویر سنتینل-۱ در هر دو شهر تهران و شیراز چندان متناسب با تصاویر واقعیت زمینی نیستند؛ که علت این امر هم ماهیت تصاویر راداری است. همان‌طور که در قطعات مختلف A-B-C-D-E-F هم مشاهده می‌شود و با کادربندی رنگی هم مشخص است، پیکسل‌هایی که در تصاویر واقعیت زمینی برچسب راه را به خود اختصاص داده‌اند در بعضی از قسمت‌ها به‌خوبی توسط مدل RDRCNN شناسایی نشده‌اند و برچسب غیر راه به آن‌ها اختصاص داده شده است و عکس این مورد خصوصاً در نتایج شهر تهران قابل‌مشاهده است که پیکسل‌هایی که در تصویر واقعیت زمینی به‌عنوان پیکسل راه برچسب‌زنی شده‌اند، به‌عنوان پیکسل راه شناسایی شده‌اند. به‌طور کلی در

نتایج هر دو شهر تهران و شیراز قابل‌مشاهده که راه‌های با پهنای عریض تا حد زیادی به‌درستی توسط مدل RDRCNN شناسایی و استخراج گردیده‌اند؛ اما در مورد راه‌های کم‌عرض همان‌طور هم که در کادربندی زردرنگ شهر تهران (قطعات A-F-C) و کادربندی قرمز رنگ شهر شیراز (قطعات C-D-E-F) هم مشخص است مدل مورد استفاده چندان مناسب عمل نکرده و در تشخیص آن‌ها خصوصاً در شهر شیراز با چالش مواجه بوده است. نکته قابل‌توجه این‌که در هر دو خروجی تهران و شیراز خصوصاً شهر تهران قابل‌مشاهده است که در تقاطع راه‌ها مدل در تشخیص راه با چالش مواجه بوده است (قطعات D-E-F شهر تهران و قطعات A-E-F شهر شیراز).



شکل ۹- نتایج حاصل مدل RDRCNN بر روی قطعات منتخب تصاویر سنتینل-۱ شهر تهران (بالا) شهر شیراز (پایین)

شده و نقشه راه شهر برای هرکدام از تصاویر و مدل‌های مورد استفاده آماده‌سازی شد. در شکل (۱۰) و (۱۱) به ترتیب تصاویر ورودی سنتینل-۱ از شهرهای تهران و شیراز به‌عنوان تصاویر تست، تصاویر برجسب‌زنی و نتایج حاصل از الگوریتم RDRCNN برای هرکدام از شهرهای یادشده نشان داده شده است.



(الف)



(ب)



(ج)

شکل ۱۰- (الف): تصویر ورودی از شهر تهران به‌عنوان داده تست، (ب): تصویر برجسب‌زنی شده به‌عنوان ورودی به مدل و (ج): نتایج حاصل از مدل RDRCNN

پس از آشکارسازی و شناسایی راه‌ها از تصاویر تست موردنظر (تهران و شیراز) با استفاده از مدل پیشنهادی در تحقیق، در مرحله‌ی بعد لازم است که نتایج به‌دست‌آمده با داده‌های مرجع (واقعیت زمینی) مقایسه و مورد ارزیابی قرار گیرند. لذا به‌منظور این کار همان‌طور که در بخش-های قبل به آن اشاره شد، نتایج تحقیق با تصاویر مرجع و استفاده از ماتریس خطا برای عوارض راه درست و نادرست شناسایی‌شده ارزیابی خواهد شد. جدول ارزیابی صحت برای شکل (۹) در جداول (۶) و (۷) آورده شده است.

جدول ۶- ارزیابی صحت عارضه راه آشکارسازی‌شده با استفاده از پارامترهای آماری مربوط به قطعات تصاویر راداری سنتینل-۱ شهر تهران

قطعات (تهران)	پارامترهای ارزیابی صحت			
	P	R	F1	OA
A	٪۵۱,۳۶	٪۵۸,۹۱	٪۵۴,۶۰	٪۹۲,۵۷
B	٪۵۲,۵۰	٪۵۷,۲۶	٪۵۴,۷۷	٪۹۱,۵۱
C	٪۵۱,۲۷	٪۵۷	٪۵۳,۹۸	٪۹۲,۱۱
D	٪۵۳,۶۱	٪۵۹,۳۶	٪۵۶,۳۳	٪۹۳,۹۱
E	٪۵۲	٪۵۸,۱۹	٪۵۴,۹۲	٪۹۳,۶۷
F	٪۵۱,۴۹	٪۵۷,۷۷	٪۵۳,۷۴	٪۹۴,۳۰

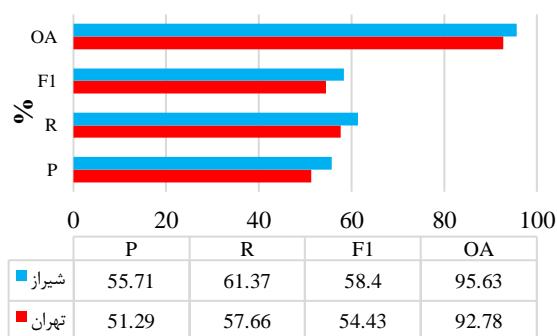
جدول ۷- ارزیابی صحت عارضه راه آشکارسازی‌شده با استفاده از پارامترهای آماری مربوط به قطعات تصاویر راداری سنتینل-۱ شهر شیراز

قطعات (شیراز)	پارامترهای ارزیابی صحت			
	P	R	F1	OA
A	٪۵۵,۱۲	٪۶۰,۸۷	٪۵۷,۸۵	٪۹۳,۴۱
B	٪۵۶,۸۸	٪۶۱,۰۹	٪۵۸,۹۰	٪۹۵,۲۲
C	٪۵۴,۱۶	٪۵۹,۶۲	٪۵۶,۷۵	٪۹۳,۸۳
D	٪۵۶,۴۱	٪۵۸,۳۲	٪۵۷,۳۴	٪۹۲,۷۶
E	٪۵۴,۳۷	٪۶۰,۶۱	٪۵۷,۳۲	٪۹۳,۵۹
F	٪۵۶,۱۸	٪۶۰,۲۵	٪۵۸,۱۰	٪۹۵,۲۴

همان‌طور که در جداول (۶) و (۷) هم قابل استناد است، مدل RDRCNN اندکی بهتر روی تصاویر شهر شیراز بهتر جواب داده است؛ اما به‌طور کلی مدل فوق جهت آشکارسازی عارضه راه از تصاویر راداری سنتینل-۱ شهرهای تهران و شیراز تقریباً با دقت برابری عمل کرده است و نتایج تمام پارامترهای ارزیابی صحت حکایت از این دارد که دقت تشخیص و شناسایی عارضه راه از تصاویر راداری شهرهای تهران و شیراز نزدیک به هم بوده‌اند.

پس از آموزش مدل‌ها و شناسایی عارضه راه از قطعات تصاویر، نهایتاً با همان مختصات ورودی قطعات تجمیع

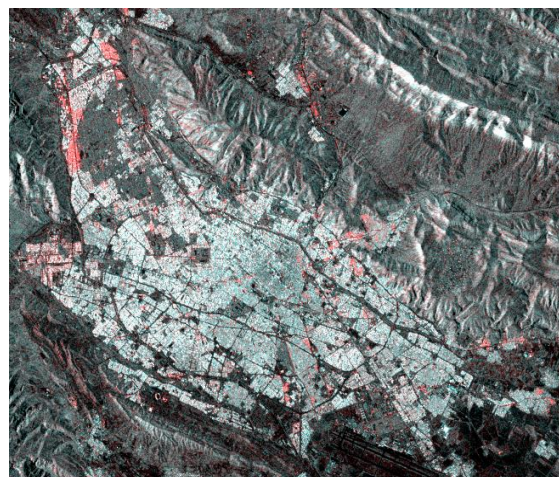
پیکسل‌های زیادی که مربوط به کلاس راه نیستند را به‌عنوان کلاس راه شناسایی کرده است. در این میان هرچند مدل مسیر و روند راه‌ها را شناسایی کرده اما در تقاطع‌ها و جاهایی که شبکه راه درون شهری پیچیده بوده دچار چالش شده به‌صورت بریده بریده پیکسل‌های راه را تشخیص داده است. از طرف دیگر همان‌گونه که در شکل (۱۱) نتایج حاصل از شهر شیراز قابل استناد است، مدل RDRCNN در تشخیص پیکسل‌های مربوط به کلاس راه در بسیاری از قسمت‌های تصویر با چالش مواجه بوده و نتوانسته باکیفیت مناسب آن‌ها را شناسایی کند. در این میان همان‌طور که در قسمت (ب) شکل (۱۱) مشاهده می‌شود اکثر راه‌های برجسبزی شده پهنای زیادی ندارند و کم‌عرض هستند؛ اما در خروجی‌های مدل (قسمت ج) با عرض و پهنای بیشتری آن‌ها را برآورد و تشخیص داده است. در نتایج شهر شیراز همانند نتایج شهر تهران اما مدل در تشخیص مسیر و روند راه‌ها موفق عمل کرده اما راه‌ها را به‌صورت بریده بریده شناسایی و آشکار کرده است. در این راستا در نتایج شهر شیراز هم قابل‌مشاهده است که پیکسل‌های غیر راه به‌عنوان کلاس راه شناسایی شده‌اند اما به‌شدت نتایج حاصل از خروجی‌های شهر تهران نیست. پس از تجمیع قطعات تصاویر و خروجی‌ها، جهت به دست آوردن میانگین دقت خروجی‌های موردنظر اقدام شد. در شکل (۱۲) میانگین دقت برای هرکدام از تصاویر در شهرهای تهران و شیراز قابل‌مشاهده است.



معیارهای ارزیابی صحت

شکل ۱۲- میانگین دقت عارضه‌ی راه آشکارسازی شده از شهرهای تهران و شیراز با استفاده از مدل RDRCNN

همان‌طور که در شکل (۱۲) قابل استناد است یافته‌های نهایی حاکی از آن است که مدل RDRCNN در آشکارسازی عارضه راه از تصاویر راداری سنتینل-۱ شهرهای تهران و شیراز تقریباً مشابه عمل کرده و در این



(الف)



(ب)



(ج)

شکل ۱۱- (الف): تصویر ورودی از شهر شیراز به‌عنوان داده تست، (ب): تصویر برجسبزی شده به‌عنوان ورودی به مدل و (ج): نتایج حاصل از مدل RDRCNN

همان‌طور که در شکل (۱۰) و قسمت (ج) قابل‌مشاهده است نتایج به‌دست‌آمده از شهر تهران از نظر بصری قابل قیاس با تصاویر برجسبزی شده (شکل ۱۱ قسمت (ب)) ورودی به شبکه نیست و مدل موردنظر

در تحقیق حاضر برای اولین بار به صورت هم‌زمان از تصاویر راداری سنتینل-۱ دو شهر مختلف (تهران و شیراز) با شرایط برابر از نظر داده‌های آموزشی، اعتبارسنجی و شبکه‌ی عصبی عمیق با معماری یکسان اقدام به آشکارسازی عارضه راه گردید. به‌طور کلی، نتایج حاصله حاکی از آن بود که مدل RDRCNN در هر دو شهر تهران و شیراز تقریباً یکسان عمل کرده و نتایج حاصل از تصاویر تست شهر شیراز هم از نظر بصری و هم از نظر متریک‌های ارزیابی صحت اندکی بهتر از نتایج حاصل از شهر تهران بوده است. در این میان در یافته‌های حاصل از هر دو شهر تعداد پیکسل‌هایی که مربوط به کلاس راه نبودند مدل آن‌ها را کلاس راه تشخیص داده است به‌وفور مشاهده می‌شود که علت این امر وجود واحدهای DPU و RCU موجود در معماری مدل RDRCNN است که در این راستا پیچیدگی بیشتر شبکه‌ی راه‌ها در شهر تهران متأثر از همین واحدها بوده و بیشتر از شهر شیراز پیکسل‌های غیر راه به‌عنوان کلاس راه تشخیص داده شده‌اند.

به‌طور کلی در این پژوهش برخلاف تحقیق Zhang و همکاران، (۲۰۱۹) که از مدل FCN جهت شناسایی و استخراج عارضه راه از تصاویر راداری سنتینل-۱ استفاده کرده بود [۴]، استفاده از مدل RDRCNN برای شهرهای تهران و شیراز چندان قابل آموزش نبود و نتایج نهایی دقت قابل قبولی از نظر بصری و متریک‌های آماری نداشت. در همین راستا شیخ‌قادری، (۱۴۰۰) در تحقیقی به این نتیجه دست‌یافت که استفاده از مدل FCN در کنار مدل RDRCNN به‌منظور آشکارسازی و استخراج راه از تصاویر سنتینل-۱ شهر تهران نتایج قابل قبولی نداشته و نتایج نهایی آن هم از نظر بصری و هم از نظر متریک‌های صحت-سنجی دقت بالایی را نشان نداده است [۴۱].

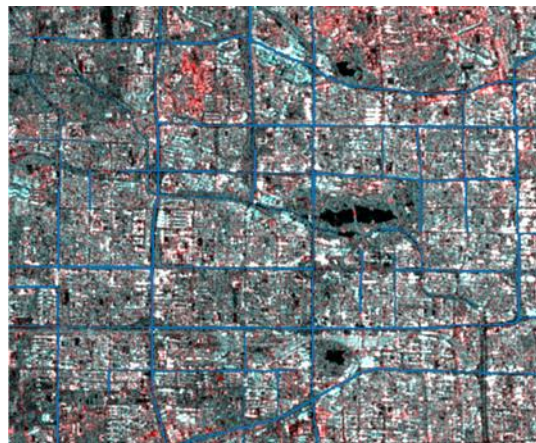
نکته‌ای که در این تحقیق وجود دارد و در پژوهش شیخ‌قادری، (۱۴۰۰) هم به آن اشاره شده است [۴۱]؛ اینکه در هنگام آموزش شبکه با تصاویر مختلف از شهرهای ایران، دقت آموزش پایین و خطای آن تا حد زیادی بالا بود؛ اما موقعی که تعداد ۲۰۰ قطعه تصویر با اندازه‌ی 256×256 برچسب‌گذاری شده از تصاویر راداری شهر پکن چین که دارای شبکه راه با پهنای عریض، منظم و پیچیده به روند آموزش اضافه شد، دقت آموزش به‌طور چشم‌گیری افزایش پیدا کرده است (شکل ۱۳).

راستا به‌صورت محدود در نتایج شهر شیراز معیارهای ارزیابی صحت اندکی بهبود یافته‌اند؛ اما به‌طور کلی یافته‌های نهایی از هر دو شهر فوق دقت بالایی را نشان نمی‌دهد.

نکته قابل توجه این که معیار دقت کلی یا OA، معیار دقیقی به‌منظور صحت سنجی نتایج نمی‌باشد؛ به این دلیل که معیار فوق تمام پیکسل‌های موردنظر در یک تصویر یعنی پیکسل‌های برچسب‌زنی شده و غیر برچسب‌زنی شده را در نظر می‌گیرد و این مورد باعث می‌شود که مقدار معیار فوق به‌صورت محسوسی افزایش پیدا کند. همان‌طور که از نتایج هم قابل استناد است در دو خروجی تصاویر سنتینل-۱ معیار OA مقدار بسیار بالایی نسبت به معیارهای دیگر دارد که در واقع علت این مقدار بالا، در نظر گرفتن تمام پیکسل‌ها و کلاس‌های مدنظر است؛ بنابراین معیارهای P ، R و F_1 از نظر علمی قابل استنادتر هستند و در تمام مراجع و اسناد معتبر که در زمینه‌ی شناسایی و آشکارسازی عارضه راه وجود دارد، از آن‌ها استفاده کرده‌اند. هرچند در بین معیارهای P ، R و F_1 هم معیار P یعنی صحت یا همان تعداد اسناد صحیح بازگردانده شده توسط مدل، در بعضی از مطالعات به‌عنوان یک معیار مناسب برای صحت سنجی نسبت به معیارهای دیگر به آن استناد کرده‌اند [۴ و ۱۵] یا اینکه در بعضی از مطالعات دیگر معیار F_1 به‌عنوان معیار مناسب در نظر گرفته شده است [۲۱ و ۴۰].

۵- بحث

در این مقاله با پیشنهاد یکی از مدل‌های یادگیری عمیق به تشخیص و آشکارسازی عارضه راه درون‌شهری کلان‌شهرهای تهران و شیراز با استفاده از تصاویر راداری سنتینل-۱ پرداخته شد. در این راستا پس از پیش‌پردازش تصاویر از ۸ شهر مختلف، ابتدا با استفاده روش قطعه‌بندی معنایی تمام پیکسل‌های مربوط به عارضه راه در تصاویر تصویر راداری برچسب‌گذاری شده و پس از آن تصاویر فوق به قطعات 256×256 تبدیل شدند. در مرحله‌ی بعد پس از جداسازی قطعات نامناسب به‌صورت یکسان در تمام تصاویر، تعداد قطعات مربوط به داده‌ی تست، آموزش و اعتبارسنجی انتخاب شدند. در این راستا داده‌ی تست در نهایت پس از مراحل فوق، با استفاده مدل RDRCNN اقدام به آموزش، ارزیابی و تشخیص و آشکارسازی عارضه راه درون‌شهری شهرهای تهران و شیراز با استفاده از تصاویر راداری سنتینل-۱ شد.



شکل ۱۳- نمونه‌ای از تصویر راداری سنتینل-۱ شهر پکن چین

پس از بررسی این موضوع، مشخص شد که پهنای عرض راه‌ها در این موضوع تأثیر چشم‌گیری دارد. در واقع اکثر شهرهای ایران پهنای عرض راه‌های آن کوچک است و این موضوع در تصاویر راداری به علت ماهیت تصویربرداری که دارند، نمایان‌تر هم می‌شود. لذا با توجه به نتایج تحقیق Zhang و همکاران، (۲۰۱۹) که در تصاویر مورد استفاده‌ی آن‌ها راه‌ها عرض و پهنای بزرگ‌تری داشتند [۴]، نتایج پژوهش شیخ‌قادری، (۱۴۰۰) و همچنین نتایج حاصل از خروجی‌های تصاویر راداری در این تحقیق برای شهرهای تهران و شیراز، قابل استناد است عرض و پهنای راه‌ها در نحوه‌ی آموزش و همچنین دقت نهایی آشکارسازی و استخراج عارضه راه از تصاویر راداری سنتینل-۱ تأثیر بسزایی دارد. در همین راستا Hong و همکاران، (۲۰۱۸) نیز به این مسئله اشاره کرده‌اند که آشکارسازی و استخراج راه با استفاده از روش‌های یادگیری عمیق در تصاویر ماهواره‌ای به شدت تحت تأثیر طیف، بافت‌ها و موقعیت‌های مختلف قرار می‌گیرد؛ در همین حال، ویژگی‌های هندسی جاده‌ها نظیر تغییر انحنای و عرض نیز چالش‌هایی را برای استخراج راه به ارمغان می‌آورد [۴۲]. لذا با توجه به مراجع و مستندات موجود، نتایج حاصل از این تحقیق که دقت بالایی برای خروجی‌های شهرهای تهران و شیراز نداشت؛ می‌تواند قابل توجه باشد. به‌طور کلی وجود پیکسل‌های مخلوط در تصاویر راداری همچنین نویز اسپیکل، ضرایب پس‌پراکندگی متفاوت تصاویر SAR سنتینل-۱ و کم‌عرض بودن راه‌ها در تصاویر تست (شهرهای تهران و شیراز) و داده‌های آموزشی، فرایند برچسب‌گذاری را با چالش و نهایتاً باعث پایین آمدن دقت مدل مورد استفاده، شده است. از طرفی هم با توجه به ماهیت و قدرت تفکیک مکانی پایین تصاویر راداری سنتینل-۱ و

مقایسه با تحقیقات پیشین، یافته‌های حاصله در این تحقیق به دلیل مشکلات ذکر شده در طول انجام فرایند و همچنین فقدان مطالعات لازم در زمینه‌ی آشکارسازی عارضه راه با این تصاویر، نمی‌توان با قاطعیت در مورد آن اظهار نظر کرد.

۶- نتیجه‌گیری

در این مقاله، یک شبکه عصبی پیچشی باقیمانده عمیق پالایش‌شده (RDRCNN) با یک واحد متصل باقی‌مانده (RCU) و یک واحد ادراک متسع شده (DPU) باهدف شناسایی و آشکارسازی راه از تصاویر سنجش‌ازدور راداری و ارزیابی و مقایسه‌ی دقت خروجی‌های نهایی در کلان‌شهرهای تهران و شیراز با شرایط برابر ارائه گردید. در این راستا پس از آماده‌سازی و برچسب‌گذاری تصاویر راداری سنتینل-۱ با روش قطعه‌بندی معنایی از ۸ شهر مختلف با ویژگی‌های مختلف از نظر هندسه، شکل و عرض راه‌ها، از مدل RDRCNN برای تشخیص و آشکارسازی راه‌های دوتا از مهم‌ترین شهرهای ایران اقدام شد. یافته‌های پژوهش نشان‌دهنده دقت پایین خروجی‌های هر دو شهر هم از نظر کیفی و هم از نظر کمی می‌باشد. در این میان استفاده از تصاویر با قدرت تفکیک مکانی متوسط راداری سنتینل-۱ به علت پایین بودن وضوح مکانی، وجود پیکسل‌های مخلوط، باز پراکنش‌های متفاوت و کم‌عرض بودن راه‌ها (در شهرهای ایران) مشکلاتی را در هنگام برچسب‌گذاری ایجاد کردند و همین موارد نه تنها دقت آموزش شبکه را پایین آوردند بلکه دقت نهایی خروجی‌های هر دو منطقه‌ی مورد مطالعه (تهران و شیراز) را تحت شعاع قرار دادند. به‌طور کلی با مقایسه‌ی نتایج حاصل از این پژوهش و تعداد محدود پژوهش‌های دیگر می‌توان به این قضیه به‌صورت تقریبی پی برد که شناسایی و آشکارسازی عارضه راه با استفاده از تصاویر راداری سنتینل-۱ می‌تواند کار چالش‌برانگیزی علی‌الخصوص در شهرهای ایران باشد؛ بنابراین پیشنهاد می‌شود در مطالعات آتی با بهره‌گیری از دیگر روش‌های مبتنی بر شبکه‌های عصبی عمیق به تشخیص و آشکارسازی عارضه راه از تصاویر سنتینل-۱ در دیگر شهرهای ایران پرداخته شود.

سپاسگزاری

نویسندگان این مقاله از آژانس فضایی اروپا به خاطر در اختیار قرار دادن تصاویر ماهواره‌ای راداری سنتینل-۱، نهایت تشکر و قدردانی را دارند.

- [1] Boyko, A., & Funkhouser, T. (2011). Extracting roads from dense point clouds in large scale urban environment. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(6), S2-S12.
- [2] Mena, J. B. (2003). State of the art on automatic road extraction for GIS update: a novel classification. *Pattern recognition letters*, 24(16), 3037-3058.
- [3] Matkan, A. A., Hajeb, M., & Sadeghian, S. (2014). Road extraction from lidar data using support vector machine classification. *Photogrammetric Engineering & Remote Sensing*, 80(5), 409-422
- [4] Zhang, Q., Kong, Q., Zhang, C., You, S., Wei, H., Sun, R., & Li, L. (2019). A new road extraction method using Sentinel-1 SAR images based on the deep fully convolutional neural network. *European Journal of Remote Sensing*, 52(1), 572-582.
- [5] Wang, W., Yang, N., Zhang, Y., Wang, F., Cao, T., & Eklund, P. (2016). A review of road extraction from remote sensing images. *Journal of traffic and transportation engineering (english edition)*, 3(3), 271-282.
- [6] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- [7] He, Y., Kavukcuoglu, K., Wang, Y., Szlam, A. and Qi, Y., 2014, Unsupervised feature learning by deep sparse coding, In *Proceedings of the 2014 SIAM International Conference on Data Mining*, pp. 902-910. Society for Industrial and Applied Mathematics
- [8] Diehl, P.U., Neil, D., Binas, J., Cook, M., Liu, S.C. and Pfeiffer, M., 2015, Fast-classifying, high-accuracy spiking deep networks through weight and threshold balancing. In *Neural Networks (IJCNN), 2015 International Joint Conference on IEEE*, pp. 1-8.
- [9] Mnih, V.; Hinton, G.E. Learning to Detect Roads in High-Resolution Aerial Images. In *Computer Vision ECCV 2010*. Number 6316 in *Lecture Notes in Computer Science*; Daniilidis, K., Maragos, P., Paragios, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 210–223.
- [10] Li, Y., Peng, B., He, L., Fan, K., Li, Z., & Tong, L. (2019). Road extraction from unmanned aerial vehicle remote sensing images based on improved neural networks. *Sensors*, 19(19), 4115.
- [11] Xia, W., Zhang, Y. Z., Liu, J., Luo, L., & Yang, K. (2018). Road extraction from high resolution image with deep convolution network—A case study of GF-2 image. In *Multidisciplinary Digital Publishing Institute Proceedings (Vol. 2, No. 7, p. 325)*.
- [12] Lian, R., Wang, W., Mustafa, N., & Huang, L. (2020). Road extraction methods in high-resolution remote sensing images: A comprehensive review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 5489-5507.
- [13] Xu, Y., Xie, Z., Feng, Y., & Chen, Z. (2018). Road extraction from high-resolution remote sensing imagery using deep learning. *Remote Sensing*, 10(9), 1461.
- [14] Henry, C., Azimi, S. M., & Merkle, N. (2018). Road segmentation in SAR satellite images with deep fully convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 15(12), 1867-1871
- [15] Gao, L., Song, W., Dai, J., & Chen, Y. (2019). Road extraction from high-resolution remote sensing imagery using refined deep residual convolutional neural network. *Remote Sensing*, 11(5), 552.
- [16] Ren, Y., Yu, Y., & Guan, H. (2020). DA-CapsUNet: A Dual-Attention Capsule U-Net for Road Extraction from Remote Sensing Imagery. *Remote Sensing*, 12(18), 2866.
- [17] Lin, Y., Wan, L., Zhang, H., Wei, S., Ma, P., Li, Y., & Zhao, Z. (2021). Leveraging optical and SAR data with a UU-Net for large-scale road extraction. *International Journal of Applied Earth Observation and Geoinformation*, 103, 102498.
- [18] Wang, S., Yang, H., Wu, Q., Zheng, Z., Wu, Y., & Li, J. (2020). An improved method for road extraction from high-resolution remote-sensing images that enhances boundary information. *Sensors*, 20(7), 2064.
- [19] Gao, X., Sun, X., Zhang, Y., Yan, M., Xu, G., Sun, H., ... & Fu, K. (2018). An end-to-end neural network for road extraction from remote sensing imagery by multiple feature pyramid network. *IEEE Access*, 6, 39401-39414.
- [20] Hong, Z., Ming, D., Zhou, K., Guo, Y., & Lu, T. (2018). Road extraction from a high spatial resolution remote sensing image based on richer convolutional features. *IEEE Access*, 6, 46988-47000.
- [21] Abdollahi, A., Pradhan, B., & Alamri, A. (2020). VNet: An End-to-End Fully Convolutional Neural Network for Road Extraction From High-Resolution Remote Sensing Data. *IEEE Access*, 8, 179424-179436.
- [22] Wu, S., Du, C., Chen, H., Xu, Y., Guo, N., & Jing, N. (2019). Road extraction from very high resolution images using weakly labeled OpenStreetMap centerline. *ISPRS International Journal of Geo-Information*, 8(11), 478.

- [23] Sun, T., Chen, Z., Yang, W., & Wang, Y. (2018). Stacked u-nets with multi-output for road extraction. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 202-206).
- [24] Zhang, Z., Liu, Q., & Wang, Y. (2018). Road extraction by deep residual u-net. *IEEE Geoscience and Remote Sensing Letters*, 15(5), 749-753.
- [25] Wei, Y., Wang, Z., & Xu, M. (2017). Road structure refined CNN for road extraction in aerial image. *IEEE Geoscience and Remote Sensing Letters*, 14(5), 709-713.
- [26] Wulamu, A., Shi, Z., Zhang, D., & He, Z. (2019). Multiscale road extraction in remote sensing images. *Computational intelligence and neuroscience*, 2019.
- [27] Shan, B., & Fang, Y. (2020). A cross entropy based deep neural network model for road extraction from satellite images. *Entropy*, 22(5), 535.
- [28] Lee, J. S., Grunes, M. R., & De Grandi, G. (1999). Polarimetric SAR speckle filtering and its implication for classification. *IEEE Transactions on Geoscience and remote sensing*, 37(5), 2363-2373.
- [29] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. arXiv, 2015; arXiv:1512.03385.
- [30] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.
- [31] Chen, L.-C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully ConnectedCRFsarXiv,2016;arXiv:1606.00915.
- [32] Ioffe, S., & Szegedy, C. (2015, June). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International conference on machine learning (pp. 448-456). PMLR.
- [33] Yu, F., & Koltun, V. (2015). Multi-scale context aggregation by dilated convolutions. arXiv preprint arXiv:1511.07122.
- [34] Krähenbühl, P., & Koltun, V. (2011). Efficient inference in fully connected crfs with gaussian edge potentials. *Advances in neural information processing systems*, 24, 109-117.
- [35] Hamaguchi, R., Fujita, A., Nemoto, K., Imaizumi, T., & Hikosaka, S. (2018, March). Effective use of dilated convolutions for segmenting small object instances in remote sensing imagery. In 2018 IEEE winter conference on applications of computer vision (WACV) (pp. 1442-1450). IEEE.
- [36] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Zheng, X. (2016). Tensorflow: A system for large-scale machine learning. In 12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16) (pp. 265-283).
- [37] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
- [38] Akçay, H. G., & Aksoy, S. (2008). Automatic detection of geospatial objects using multiple hierarchical segmentations. *IEEE Transactions on Geoscience and Remote Sensing*, 46(7), 2097-2111.
- [39] Li, Y., Sun, X., Wang, H., Sun, H., & Li, X. (2012). Automatic target detection in high-resolution remote sensing images using a contour-based spatial model. *IEEE Geoscience and Remote Sensing Letters*, 9(5), 886-890.
- [40] He, H., Yang, D., Wang, S., Wang, S., & Li, Y. (2019). Road extraction by using atrous spatial pyramid pooling integrated encoder-decoder network and structural similarity loss. *Remote Sensing*, 11(9), 1015.
- [41] Sheikghaderi, S.H. (2021). Accuracy Assessment of Road Extraction Methods Based on Deep Learning Algorithms Using Radar and Optical Satellite Image. Master Thesis. Kharazmi University of Tehran.
- [42] Hong, Z., Ming, D., Zhou, K., Guo, Y., & Lu, T. (2018). Road extraction from a high spatial resolution remote sensing image based on richer convolutional features. *IEEE Access*, 6, 46988-47000.