

کشف سیل در تصاویر اخذ شده توسط پهپاد با استفاده از معماری PSPNet و برآورد عدم قطعیت پیش‌بینی‌ها به کمک روش دورریز مونت کارلو

سید علی احمدی^{۱*}، علی محمدزاده^۲

^۱ دانشجوی دکتری سنجش از دور، دانشکده مهندسی نقشه‌برداری دانشگاه صنعتی خواجه نصیرالدین طوسی
s.a.ahmadi@email.kntu.ac.ir

^۲ استاد گروه فتوگرامتری و سنجش از دور، دانشکده مهندسی نقشه‌برداری دانشگاه صنعتی خواجه نصیرالدین طوسی
a_mohammadzadeh@kntu.ac.ir

(دریافت: آذر ۱۴۰۲، تصویب: دی ۱۴۰۲)

چکیده

سیلاب‌ها از شایع‌ترین و خطرناک‌ترین مخاطرات طبیعی هستند که در مقیاس وسیع بر جامعه تاثیر گذاشته و آسیب‌های مالی و جانی قابل توجهی را به آن وارد می‌کنند. استفاده از جدیدترین فناوری‌ها و نوآوری‌ها توسط مدیران و نیروهای امدادی سبب کاهش تاثیر مخرب سیل‌ها و صرفه‌جویی در هزینه‌ها می‌شود. پهپادهای مجهز به سنجنده‌های دقیق در کنار الگوریتم‌های پیشرفته بینایی ماشین و یادگیری عمیق می‌توانند به عنوان یک سکوی بالقوه برای فعالیت‌های نظارت، نقشه‌برداری، شناسایی و پهنه‌بندی سیلاب به صورتی کارآمد مورد استفاده قرار گیرند. در این مطالعه به منظور قطعه‌بندی معنایی تصاویر پهپادی با قدرت تفکیک مکانی بالا که پس از سیل از منطقه شهری اخذ شده‌اند، از معماری Pyramid Scene Parsing Network (PSPNet) به عنوان یک شبکه نوین، به همراه ResNeSt به عنوان رمزگذار استفاده شده‌است و در نهایت شبکه‌های مختلف با یکدیگر مقایسه شده‌اند. در این راستا، به جهت تفسیر بهتر و مطالعه قدرت، پایداری و عملکرد الگوریتم‌ها از روش Monte-Carlo Dropout (MCD) جهت برآورد عدم قطعیت مدل‌ها نیز استفاده شده‌است. نتایج مقایسه روش‌های مختلف نشان داد که با افزایش تعداد پارامترهای مدل و پیچیدگی شبکه، عملکرد شبکه در حین آموزش تحت معیار IoU تا ۱۰٪ و در زمان آزمایش تا ۳٪ بهبود پیدا کرده و قطعیت تصمیم‌گیری آن افزایش پیدا می‌کند. صحت (Accuracy) قطعه‌بندی معنایی تصاویر ۹۷٫۹۳٪ و معیار F1-Score تقریباً ۸۹٪ بوده‌است.

واژگان کلیدی: پهپاد، مدیریت بحران، یادگیری عمیق، کشف سیل، قطعه‌بندی معنایی، استخراج ساختمان

* نویسنده رابط

۱- مقدمه

سالانه هزاران نفر تحت تاثیر بلایای طبیعی جان خود را از دست می‌دهند. کشورها و مجامع بین‌المللی مبالغ زیادی را صرف بازسازی زیرساخت‌ها و ترمیم خسارات می‌کنند [۱]، [۲]. سیل‌ها یکی از پرتکرارترین و کشنده‌ترین بلایای طبیعی هستند که سبب وارد آمدن آسیب‌های سازه‌ای، اجتماعی، جانی و مالی بسیاری می‌شوند [۳]–[۶]. افزایش وقوع سیلاب طی سال‌های اخیر اغلب به گرمایش جهانی و تغییرات اقلیمی نسبت داده می‌شود [۴]، [۶]. با اینحال، آسیب‌های افزایش‌یافته ناشی از وقوع سیل‌ها با عوامل انسانی همچون رشد نمایی جمعیت، افزایش شهرنشینی بدون برنامه‌ریزی مناسب و مدیریت ناکارآمد سیلاب مرتبط می‌شود [۷]. به منظور بهبود فرآیندهای مدیریت بلایای طبیعی از جمله سیل‌ها، استفاده از آخرین پیشرفت‌ها و فناوری‌های روز لازم و ضروری می‌نماید.

در هنگام وقوع بحران‌های طبیعی همچون سیلاب‌ها، اطلاع از محدوده دقیق سیل، موقعیت ساختمان‌های آسیب‌دیده و مسیرهای آب‌گرفته، به نیروهای امدادی و مدیران حاضر در صحنه کمک می‌کند تا با برنامه‌ریزی بهتر، منابع و افراد را به شیوه بهینه مورد استفاده قرار دهند. از این رو، تهیه نقشه‌هایی که نشان‌دهنده موقعیت بیشترین آسیب‌های ممکن باشند برای مدیران و نیروهای امدادی دارای اهمیت به‌سزایی است [۸].

داده‌های سنجنش‌ازدوری به دلیل توانایی ارائه دید وسیع از مناطق آسیب‌دیده، امکان برآورد سریع‌تر تخریب‌های وارد شده را دارند [۹]. تصاویر ماهواره‌ای نوری (مثل Sentinel-2 یا Worldview-3)، داده‌های سنجنده‌های راداری با روزه مصنوعی^۱ (مثل Sentinel-1) و یا تصاویر هوایی و پهپادی عملکرد مفید خود را در مدیریت شرایط بحرانی، ارزیابی تخریب زیرساخت‌ها، ساختمان‌ها و جاده‌ها، و استخراج اطلاعات مورد نیاز به منظور تصمیم‌گیری سریع حین و پس از وقوع حوادث نشان داده‌اند [۱۰]–[۱۳].

داده‌های راداری امکان اخذ اطلاعات در شب و شرایط گوناگون آب و هوایی را دارند [۱۴]–[۱۸] و به دلیل جذب بالای امواج رادیویی در سطوح آب‌گرفته و عدم تاثیرپذیری از پوشش ابر در منطقه به راحتی برای استخراج نقشه پهنه‌بندی سیل در ساعات اولیه مورد استفاده قرار می‌گیرند؛ و به همین

دلیل در اغلب اوقات سریعترین داده موجود جهت برآورد وسعت حادثه در مقیاس وسیع هستند [۱۷]؛ با اینحال داده‌های رایگان سنجنده‌های راداری اغلب قدرت تفکیک مکانی متوسط و پایینی داشته و مناسب ارزیابی آسیب در سطح تک ساختمان نیستند [۱۹]. در مقابل تصاویر نوری ماهواره‌ای تفسیرپذیری ساده‌تری داشته و امکان ارائه قدرت تفکیک مکانی بهتر از یک متر را نیز دارند؛ اما تحت تاثیر عوامل مخربی چون ابر، سایه ابر، یا دود ناشی از شرایط جوی بعد از حادثه سیل، استفاده از آنها معمولاً با چالش‌هایی همراه است [۲۰]. بعلاوه، لزوماً مسیر و زمان عبور ماهواره از روی منطقه آسیب‌دیده همزمان با وقوع بیشینه آب‌گرفتگی‌ها نبوده و ارزیابی خسارت دقیق را محدود می‌کند [۲۱].

در مقایسه با داده‌های ماهواره‌ای که امکان ارائه اطلاعات از پهنه‌های وسیع‌تری را دارند، اخذ تصاویر با قدرت تفکیک مکانی بسیار بالا توسط پهپادها با هزینه کمتری امکان‌پذیر است [۱۲]. امکان تفسیر بصری سریع‌تر و دقیق‌تر محدوده حادثه توسط کاربر پهپاد و همچنین تمرکز بر محدوده خاصی در حین و پس از وقوع سیل از جمله مزایای داده‌های پهپادی است [۱۲]، [۲۱]. همچنین، در شرایط پس از بارش باران، در صورتی که همچنان پوشش ابر بر روی منطقه باشد، در حالی که سنجنده‌های فضایی امکان اخذ تصاویر نوری از سطح زمین را ندارند، سکوه‌های هوایی می‌توانند در زمان مناسب اقدام به تصویربرداری نمایند.

پهپادها (پرنده‌های هدایت‌پذیر از دور)، که به عنوان پرنده‌های بدون سرنشین^۲ (UAV) نیز شناخته می‌شوند، به طور فزاینده‌ای در حوزه‌های علمی گوناگون برای جمع‌آوری داده و انجام وظایفی که برای انسان‌ها دشوار، خطرناک یا غیرممکن است، مورد استفاده قرار می‌گیرند. در سال‌های اخیر، استفاده از فناوری پهپادی با تکیه بر سخت‌افزارهای کارآمدتر، تجاری‌سازی و سهولت دسترسی، به طور چشمگیری پیشرفت و توسعه داشته‌است. پهپادهای کوچک تجاری با امکان پرواز در ارتفاع‌های پایین، مجهز به دوربین‌های تصویربرداری با قدرت تفکیک مکانی بالا، با قابلیت تصویربرداری در زوایای گوناگون، در کنار الگوریتم‌های هوش مصنوعی، می‌توانند به عنوان یک ابزار قدرتمند جهت نظارت، نقشه‌برداری و شناسایی در مدیریت سیل به کار گرفته شوند [۲۲]–[۲۴]. با اینحال، علی‌رغم پیشرفت‌های شمرده‌شده در این حوزه، هندسه‌ی ناپایدارتر

۲ Unmanned Aerial Vehicle (UAV)

۱ Synthetic Aperture Radar (SAR)

تصویربرداری پهپادها، حجم و تعداد بالای تصاویر پوشش‌دار اخذشده، زوایای دید متغییر در تصاویر و جزئیات زیاد موجود در تصاویر استفاده عملی از این سکوها را محدود کرده و نیازمند توسعه و بهبود روش‌هایی جهت انجام بهینه پردازش‌ها می‌باشد.

تفسیر چشمی و یا پردازش تعداد زیادی از تصاویر با قدرت تفکیک مکانی بالا با استفاده از روش‌های سنتی در سنجش از دور زمانبر و در عین حال در شرایط بحرانی، ناکارآمد است [۲۵]. اخیراً روش‌های بینایی ماشین مبتنی بر یادگیری عمیق همچون طبقه‌بندی [۲۶]، کشف عوارض [۲۷] یا تشخیص عوارض نتایج قابل توجهی در حل مسائل پیچیده پردازش تصویر ارائه کرده‌اند. برخلاف مدل‌های معمول یادگیری ماشین که نیاز به مهندسی و دستکاری ویژگی‌ها^۱ دارند، مدل‌های یادگیری عمیق با تکیه بر ساختار چند لایه قادرند تا فضای ویژگی پیچیده‌تری را ساخته و آنرا به صورت خودکار فرا گیرند [۲۸]، [۲۹]. این امر سبب افزایش صحت و بهبود عملکرد مدل‌ها می‌شود. در همین راستا، محققان سنجش‌ازدور نیز تمرکز خود را بر روی مدل‌های یادگیری عمیق معطوف کرده و در کاربردهای گوناگونی همچون طبقه‌بندی پوشش اراضی [۳۰]، درجه‌بندی میزان تخریب ساختمان‌ها پس از زلزله [۳۱] یا پیش‌بینی و پایش آتش‌سوزی در جنگل‌ها [۳۲]، از آنها استفاده کرده‌اند و بدین منظور داده‌های متنوعی را نیز توسعه داده‌اند [۳۳].

به عنوان مثال ژائو و همکارانش [۱۵] با استفاده از یک شبکه عمیق مبتنی بر UNet و اتکا بر داده‌های چند زمانه ماهواره سنتینل-۱ اقدام به پهنه‌بندی سیل در مناطق شهری نمودند. هومایرا و همکارانش [۳۴] با استفاده از شبکه‌های عصبی پیچشی اقدام به تشخیص سریع سیل در تصاویر مستخرج از شبکه‌های اجتماعی و رسانه‌ها کردند. با استفاده از شبکه‌های تمام‌پیچشی و استفاده از داده‌های اخذ شده توسط پهپاد، گبرهویت و همکارانش [۳۵] پهنه‌بندی سیل و طبقه‌بندی پیکسل‌های تصویر در چهار کلاس آب، ساختمان، گیاه و جاده را انجام دادند. همچنین یانگ و همکارانش [۳۶] اقدام به کشف عوارض ساختمانی آب گرفته در تصاویر

اخذشده توسط پهپاد کردند. با اینحال، بررسی مطالعات موجود در حوزه کشف سیل به کمک داده‌های سنجش از دوری (بخصوص تصاویر پهپادی) با استفاده از الگوریتم‌ها و مدل‌های یادگیری عمیق، نشان از عدم بلوغ کافی این حوزه و نیاز به انجام تحقیقات بیشتر در شرایط محیطی مختلف را نشان می‌دهد.

در میان کاربردهای گوناگون این مدل‌ها، قطعه‌بندی معنایی^۲ با هدف اختصاص دادن یک کلاس به هر پیکسل تصویر، یکی از مهمترین اهداف روش‌های یادگیری عمیق است [۳۷]. وجود عوارض متعدد در کلاس‌های مختلف و با ابعاد متفاوت در کنار پیچیدگی بافت تصویر از جمله مشکلات پیش روی روش‌های قطعه‌بندی معنایی بخصوص در تصاویر سنجش‌ازدوری هستند. اغلب روش‌های قطعه‌بندی معنایی مبتنی بر معماری‌های تمام‌پیچشی^۳ (FCN) هستند [۳۷]. استفاده از روش‌هایی که بر پایه‌ی شبکه‌های عصبی پیچشی^۴ (CNN) هستند [۳۸] باعث می‌شود تا امکان بررسی و اختصاص کلاس به هر پیکسل در تصاویر با ابعاد مختلف وجود داشته باشد. با اینحال عدم توجه به ساختارهای کلی^۵ در تصویر و تنها تمرکز بر ساختارهای محلی^۶ باعث کاهش دقت و افت عملکرد روش‌های معمول همچون FCN می‌شود. یکی از راهکارهای مفید در پاسخ به این مساله، استفاده از فنون ادغام هرم مکانی^۷ (SPP) [۳۹] است که به کمک ویژگی‌های مکانی تصویر در سطوح مختلف قدرت تفکیک مکانی، بتوان همزمان ساختارهای کلی و محلی را درک نمود. ترکیب توامان شبکه‌های پیچشی و فن ادغام هرم مکانی منجر به زایش معماری جدیدی با عنوان شبکه درک هرمی تصویر^۸ (PSPNet) شده‌است [۴۰] که قادر به استخراج سرنخ‌های محلی و کلی جهت تشخیص و طبقه‌بندی بهتر عوارض در تصویر می‌باشد.

لازم به ذکر است که ارتباطات مفهومی و زمینه‌ای در تمام سطح تصویر وجود دارند و به درک بهتر عوارض موجود در تصویر کمک می‌کنند (برای مثال حضور خودرو در خیابان و قایق در رودخانه). عدم قابلیت درک اطلاعات و ارتباطات زمینه‌ای احتمال تشخیص و طبقه‌بندی اشتباه را افزایش می‌دهد. همچنین وجود عوارض با ابعاد مختلف، از کوچک تا

۶ Local structures

۷ Spatial Pyramid Pooling (SPP)

۸ Pyramid Scene Parsing Network (PSPNet)

۱ Features

۲ Semantic segmentation

۳ Fully Convolutional Networks (FCN)

۴ Convolutional Neural Networks (CNN)

۵ Global structures

بزرگ، در یک صحنه از تصویر، شبکه پیچشی را به چالش وا می‌دارد. بدین صورت که درک و آموزش همزمان عوارض کوچک در کنار عوارض بزرگ برای شبکه دشوار می‌شود. شبکه PSPNet با تمرکز بر ارتباطات معنایی و زمینه‌ای در سطح تصویر امکان غلبه بر این مشکلات را دارد [۴۰].

در اغلب معماری‌های شبکه‌های یادگیری عمیق، بخشی از مدل مسئول استخراج ویژگی‌های عمیق از تصویر می‌باشد. به این بخش اصطلاحاً رمزگذار^۱ یا ستون فقرات^۲ مدل گفته می‌شود؛ چراکه درصد قابل توجهی از عملکرد مدل به توانمندی این بخش بستگی دارد. طی دهه گذشته معماری‌های رمزگذاری متعددی [۴۱] با ویژگی‌های گوناگون توسعه یافته‌اند. استخراج ویژگی‌های عمیق، تعداد پارامترها، استفاده از لایه‌های مختص توجه^۳ یا ترکیب ویژگی‌ها با مقیاس‌های مختلف از جمله عوامل تاثیرگذار در طراحی این معماری‌ها بوده‌اند که منجر به ظهور نسل‌های متنوعی همچون VGG, ResNet, Inception, EfficientNet یا MobileNet شده‌اند. با الهام گرفتن از فونونی همچون تقسیم-توجه^۴ و توجه به ویژگی‌ها^۵، محققان رمزگذارهای پیشرفته‌تری مثل ResNeSt را توسعه دادند که علاوه بر داشتن پارامترهای کمتر، قدرت درک ارتباطات میان ویژگی‌ها و توصیف تصاویر پیچیده‌تر را دارند [۴۲]. رمزگذارهای ResNeSt با ترکیب استراتژی‌های توجه به ویژگی‌ها و شبکه‌های چندشاخه‌ای (multi-path networks)، امکان درک همبستگی‌ها میان ویژگی‌های مختلف (cross-channel feature correlation)، در کنار حفظ اطلاعات مستقل را فراهم آوردند. این قابلیت‌ها امکان یادگیری بازنمایی‌های عمیق‌تری را فراهم می‌آورند که سبب بهبود عملکرد مدل می‌شوند. همچنین یادگیری سریع‌تر و همگرایی بهتر مدل، در کنار قابلیت تعمیم‌پذیری بالاتر مدل، از جمله ویژگی‌هایی هستند که استفاده از خانواده ResNeSt را توجیه می‌نمایند.

با اینحال، در کنار دقت بالا و عملکرد مناسب مدل‌های یادگیری عمیق، آنها همچنان به عنوان جعبه‌سیاه‌هایی شناخته می‌شوند که ریاضیات پیچیده داخل آنها تفسیرپذیر

نبوده و در نتیجه قابل اعتماد نیست [۴۳]. به همین دلیل، اغلب در کاربردهای حساسی همچون پزشکی یا مدیریت بحران، که افراد مسئول نیازمند اتخاذ تصمیماتی هستند که مستقیماً با جان انسان‌ها در ارتباط است، اعتماد کردن به نتایج این الگوریتم‌ها دشوار است. از این رو در سال‌های اخیر، روند تحقیقاتی در حوزه یادگیری عمیق به سمت توسعه هوش مصنوعی تفسیرپذیر^۶ یا قابل توضیح^۸ سوق پیدا کرده‌است [۲]، [۴۳]، [۴۴]. ارائه برآوردی از عدم قطعیت^۹ در پیش‌بینی‌های مدل یکی از روش‌هایی است که اعتمادپذیری نتایج را بالا می‌برد [۴۵]. روش یادگیری بیزی^{۱۰} (BL) [۴۶] یا روش دورریز مونت-کارلو^{۱۱} (MCD) [۴۷] از جمله روش‌هایی هستند که امکان برآورد عدم قطعیت نتایج شبکه را فراهم می‌آورند.

در این مطالعه، با تکیه بر معماری PSPNet و بهره‌گیری از رمزگذارهای پیشرفته خانواده ResNeSt، تصاویر پهپادی اخذ شده پس از سیلاب ناشی از طوفان هاروی در تگزاس آمریکا به منظور قطعه‌بندی معنایی عوارض مختلف مورد بررسی قرار گرفته و مقایسه دقیقی از آنها ارائه می‌شود. همچنین، با توجه به اهمیت ارائه برآورد عدم قطعیت بر روی نتایج قطعه‌بندی معنایی به منظور استفاده در تصمیم‌گیری‌های پس از حوادث طبیعی، برای اولین بار اقدام به ارائه نقشه‌های عدم قطعیت برای تصاویر پهپادی با قدرت تفکیک مکانی بالا نمودیم. بدین منظور از روش MCD به عنوان یک تخمین‌گر بیزی استفاده کردیم.

در ادامه ساختار این مقاله به شرح زیر است. در بخش دوم منطقه مطالعاتی و داده‌های مورد استفاده معرفی شده‌اند. سپس، به روش پیشنهادی در بخش سوم و چهارم زیربخش مربوطه به طور کامل پرداخته شده‌است. نتایج تحقیق و خروجی‌های بدست آمده در بخش چهارم ارائه شده‌اند و در بخش پنجم به بحث و بررسی در رابطه با نتایج پرداخته‌ایم. نهایتاً، در بخش انتهایی، نتیجه‌گیری حاصل از این مطالعه به همراه پیشنهاداتی جهت ادامه مسیر مطالعات پیش‌رو ارائه شده‌اند.

۸ Explainable AI (XAI)

۹ Uncertainty quantification

۱۰ Bayesian learning

۱۱ Monte-Carlo Dropout (MCD)

۱ Encoder

۲ Backbone

۳ Attention

۴ Split-Attention

۵ Channel-wise attention

۶ Representation

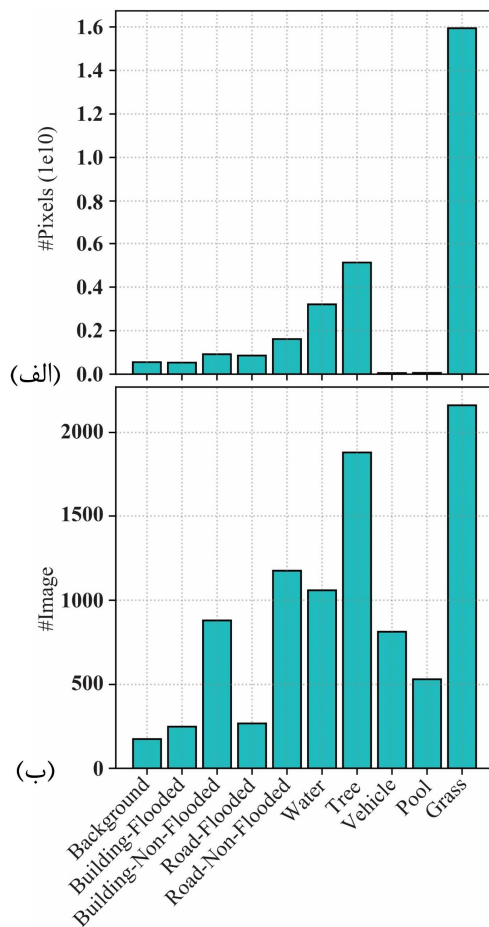
۷ Interpretable AI (IAI)

۲- منطقه مطالعاتی و داده‌ها

در شهریور سال ۱۳۹۶ (آگوست سال ۲۰۱۷)، بارش شدید باران در تگزاس و لوئیزیانا آمریکا، تحت تاثیر طوفان حاره‌ای هاروی^۱، باعث بروز سیل بی‌سابقه و مخربی طی ۵۰ سال اخیر شد که منجر به از دست رفتن جان بیش از ۱۰۰ انسان شد. تنها طی ۴ روز، بیش از ۱۰۰۰ میلی‌متر باران در مناطق تحت تاثیر بارید و این امر موجب جابه‌جایی بیش از ۳۰ هزار نفر و به زیر آب رفتن ۱۵۴ هزار ساختمان گردید. مطابق گزارش‌های مرکز ملی طوفان آمریکا، خسارات این حادثه معادل ۱۲۵ میلیارد دلار تخمین زده شدند که این حادثه را در جایگاه دومین حادثه پرهزینه تاریخ ایالات متحده قرار داد.

در این مطالعه از مجموعه داده FloodNet^۲ که پس از این حادثه جمع‌آوری شده است استفاده کرده‌ایم [۴۸]. داده FloodNet پس از بارش و در بازه زمانی ۸ تا ۱۳ شهریور ماه همان سال، توسط پهپاد DJI Mavic Pro، از مناطقی که متأثر از سیل بودند اخذ شد. این داده از دو جهت منحصر بفرد و ارزشمند است. اول اینکه تصاویر این مجموعه داده بلافاصله پس از حادثه و در زمان پاسخگویی^۳ به سیل‌زدگان توسط نیروهای امدادی برداشت شده است و نمایانگر شرایط واقعی پس از بحران می‌باشد. دوم اینکه تا آنجایی که نویسندگان این مطالعه اطلاع دارند، این داده تنها داده آماده شده برای قطعه‌بندی معنایی پس از حادثه سیل است که توسط پهپادهای کوچک اخذ شده است [۳۳]. به دلیل ارتفاع پایین پهپاد در زمان اخذ تصاویر (زیر ۲۰۰ متر بالای سطح زمین)، قدرت تفکیک مکانی تصاویر این مجموعه داده بسیار بالاست. این داده شامل ۲۳۴۳ تصویر رنگی با ابعاد ۳۰۰۰×۴۰۰۰ پیکسل است که در دسته‌های ۱۴۴۵، ۴۵۰ و ۴۴۸ تایی به ترتیب برای آموزش، اعتبارسنجی و آزمایش، تفکیک شده‌اند. در این مجموعه داده، ۹ کلاس مختلف تعریف شده است که عبارتند از: ساختمان، ساختمان سیل‌زده، جاده، جاده آب‌گرفته، آب (شامل رودخانه یا دریاچه)، درختان، خودرو، استخر و چمن. در تعریف کلاس‌ها، یک ساختمان زمانی سیل‌زده محسوب می‌شود که حداقل یکی از اضلاع آن با سیل در تماس باشد. با توجه به ابعاد مختلف عوارض و همچنین تعداد نمونه‌های مختلف در هر کلاس، FloodNet

یک مجموعه داده غیرهمگن به حساب می‌آید که فرآیند قطعه‌بندی معنایی را نیز دشوار می‌سازد. شکل ۱ توزیع ناهمگنی این مجموعه داده را به دو صورت نمایش می‌دهد. در شکل ۱-الف مجموع تعداد پیکسل‌های هر کلاس در تمامی تصاویر، و در شکل ۱-ب مجموع تصاویری که حاوی رخدادهایی از هر کلاس هستند، نشان داده شده‌اند. همچنین نمونه‌ای از تصاویر و داده‌های واقعیت زمینی مرتبط با آنها در شکل ۵ مشاهده می‌شود.



شکل ۱- نمایش ناهمگنی مجموعه داده FloodNet بر اساس الف) تعداد پیکسل‌های هر کلاس، و ب) تعداد تصاویر دارای هر کلاس.

۳- روش پیشنهادی

در این مطالعه، با استفاده از الگوریتم یادگیری عمیق و با تکیه بر رمزگذارهای مختلف، فرآیند قطعه‌بندی معنایی تصاویر پهپادی با قدرت تفکیک مکانی بالا انجام شده است.

^۳ Response phase

^۱ Hurricane Harvey

^۲ قابل دسترس در <https://github.com/BinaLab/FloodNet-Challenge-EARTHVISION2021>

همچنین به منظور افزایش اعتمادپذیری و کارآمدی روش‌ها در شرایط بحرانی و تصمیم‌گیری‌های حساس، عدم قطعیت پیش‌بینی‌های مدل نیز در قالب خروجی ارائه شده‌است. در این بخش ابتدا به معرفی الگوریتم یادگیری عمیق PSPNet پرداخته و پس از آن خانواده رمزگذارهای ResNeSt معرفی شده‌اند. در نهایت، روش برآورد عدم قطعیت MCD مورد بررسی قرار گرفته‌است. در زیربخش پایانی، معیارهای ارزیابی و اعتبارسنجی معرفی شده‌اند. شکل ۲ نمایشی تصویری از مراحل و اجزای مختلف روش را نشان می‌دهد.

۳-۱- مدل PSPNet

قطعه‌بندی معنایی تصاویر یکی از مهمترین زمینه‌های کاربرد مدل‌های یادگیری عمیق است که در آنها به هر پیکسل تصویر یک کلاس مشخص اختصاص پیدا می‌کند. در مقایسه با شبکه‌های عصبی عمیق، در مدل‌های قطعه‌بندی معنایی، لایه‌های تمام متصل^۱ با لایه‌های پیچشی جایگزین شده‌اند. عامل اصلی درک تصویر در شبکه‌های پیچشی، کرنل^۲ یا میدان درک^۳ است که معمولاً ابعاد ثابت و مشخصی دارد. کرنل‌ها معمولاً فقط توانایی درک ساختارهای محلی را دارند. افزایش ابعاد کرنل‌ها به منظور درک ساختارهای کلی در تصویر، منجر به افزایش شدید تعداد پارامترها و پیچیدگی شبکه می‌شود. بنابراین، برای حل این مسأله، در مطالعات مختلف روش‌های گوناگونی را جهت درک ساختارهای کلی در تصویر ارائه کرده‌اند. به عنوان مثال لیو و همکاران [۴۹] با استفاده از روش ادغام میانگین کلی^۴ سعی در افزایش میدان درک شبکه داشتند. در مقابل، با تمرکز بر ثابت نگه‌داشتن تعداد پارامترهای شبکه، چن و همکاران [۵۰] ادغام هرم مکانی حفره‌دار^۵ (ASPP) را ارائه کردند که در آن علی‌رغم افزایش ابعاد میدان درک، تعداد پارامترهای کرنل ثابت باقی می‌ماند.

در سال ۱۳۹۶ ژائو و همکارانش [۴۰] با استفاده از انتقال تصویر به سطوح مقیاسی چندلایه، مدل شبکه درک هرمی تصویر (PSPNet) را ارائه کردند. در این مدل ویژگی‌ها در چهار سطح مختلف از مقیاس‌های مکانی استخراج و با ویژگی‌های مستخرج از لایه رمزگذار ترکیب شده و پس از

گذشتن از یک لایه پیچشی، تصویر قطعه‌بندی شده را تولید می‌کنند. PSPNet یک روش موثر در درک مفاهیم کلی و محلی در تصاویر است و به لحاظ سرعت عملکرد و بار محاسباتی، افزایش چندانی نسبت به FCN توسعه یافته ندارد [۳۷]. با توجه به ویژگی‌های این شبکه انتظار می‌رود که امکان طبقه‌بندی پیکسل‌های تصاویر پهپادی که از زوایا و ارتفاعات مختلف اخذ شده‌اند با دقت کافی وجود داشته باشد. در شکل ۲ ساختار معماری PSPNet که شامل ویژگی‌های بدست آمده از رمزگذار، ادغام ویژگی‌ها، لایه‌های پیچشی با ابعاد کرنل مختلف، ادغام نهایی و لایه پیچشی است نشان داده شده‌است. در این ساختار، لایه رمزگذار با مدل‌های خانواده ResNeSt جایگزین شده‌است. در ادامه به معرفی این خانواده می‌پردازیم.

۳-۲- ResNeSt به عنوان رمزگذار

در معماری ResNet [۵۱] به منظور جلوگیری از فراموشی الگوها در فرآیند پس‌انتشار^۶، از اتصالات باقیمانده‌ای^۷ که ورودی را به خروجی لایه‌های پیچشی متصل می‌کنند استفاده شده‌است. از طرفی، امکان تصمیم‌گیری تطبیقی‌پذیر با توجه به ویژگی‌های ورودی به هر لایه، یکی از قابلیت‌های بلوک‌هایی همچون Squeeze-and-Excitation یا Selective-Kernel است که انعطاف‌پذیری بیشتری به مدل‌ها می‌دهد. واحد جدیدی که همزمان دارای اتصالات باقیمانده‌ای و بلوک تقسیم-توجه است، بلوک Residual-Split-Attention نامیده می‌شود و از چیدن تعدادی از آنها بر روی یکدیگر می‌توان به شبکه ResNeSt دست پیدا کرد [۴۲]. این ساختار امکان توجه به ویژگی‌ها در هر لایه را برای مدل فراهم می‌آورد. نسخه‌های مختلفی از ResNeSt با توجه به تعداد پارامترهای آنها بوجود آمده‌اند که بر اساس تعداد واحدهای تقسیم-توجه نامگذاری شده‌اند. در این مطالعه ۵ عضو از این خانواده با تعداد پارامترهای مختلف به عنوان رمزگذار مورد استفاده قرار گرفته‌اند. شبکه‌های رمزگذار وظیفه یادگیری و استخراج ویژگی‌های عمیق را دارند.

۵ Atrous Spatial Pyramid Pooling (ASPP)

۶ Back-propagation

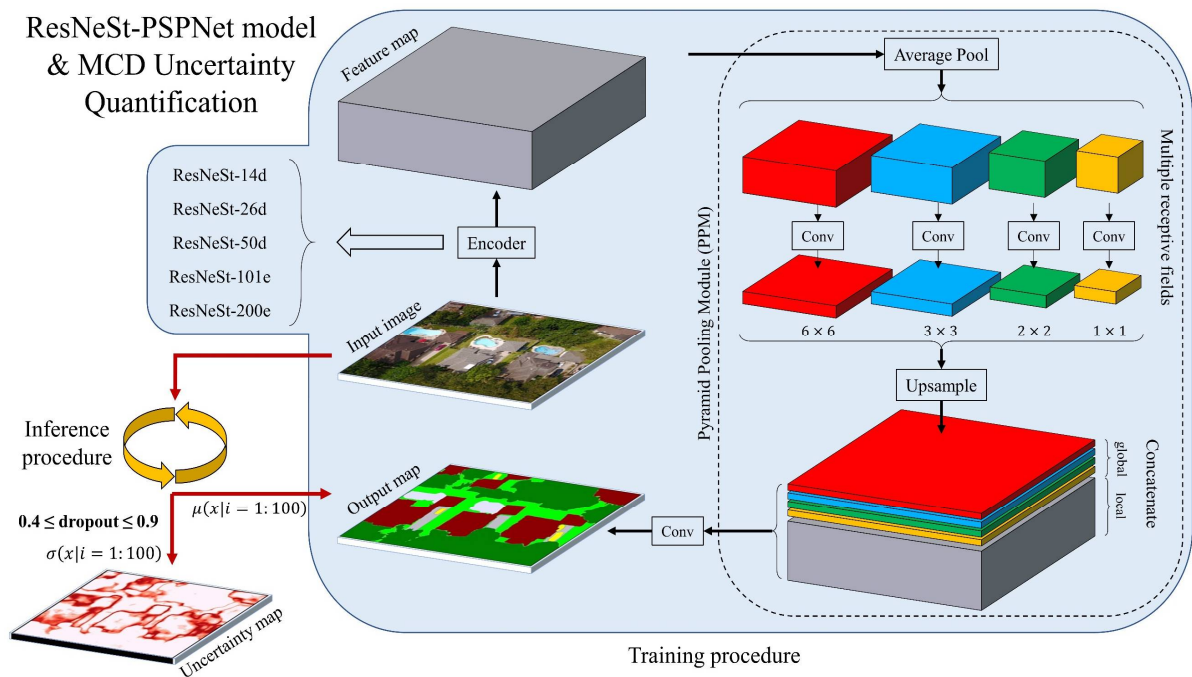
۷ Residual connections

۱ Fully Connected

۲ Kernel

۳ Receptive field

۴ Global Average Pooling (GAP)



شکل ۲- مراحل مختلف روش ارائه شده شامل رمزگذارهای ResNeSt، معماری PSPNet و فرآیند برآورد میزان عدم قطعیت نتایج

در رابطه با تصمیم‌های دارای عدم قطعیت باشد. به بیان دیگر، مقادیر Softmax نشان‌دهنده میزان تعلق یک پیکسل به کلاس ۱ نسبت به کلاس ۲ هستند؛ حال آنکه این میزان تعلق می‌تواند همراه با درصد‌های مختلفی از عدم قطعیت باشد؛ یعنی مدل می‌تواند نسبت به پیش‌بینی خودش مطمئن یا مردد باشد. به عنوان مثال، در مساله طبقه‌بندی انواع ساختمان‌ها در تصاویر هوایی، با ارائه تصویر یک خودرو به مدل، مقداری احتمالاتی (مجموع مقادیر تمام کلاس‌ها برابر با یک) که بیانگر میزان تعلق خودرو به هر یک از کلاس‌های ساختمانی است، توسط مدل پیش‌بینی می‌شود؛ اما آیا خودرو به کلاس‌های ساختمانی تعلق دارد؟! این مساله محققان را وادار نمود تا به دنبال راه حلی جهت ارائه میزان اطمینان یا عدم قطعیت مدل باشند.

نظریه احتمال بیز^۲ راهکاری ریاضیاتی برای محاسبه عدم قطعیت در مدل‌ها ارائه می‌کند. به بیان ساده‌تر، برآورد عدم قطعیت را می‌توان همان برآورد تابع توزیع نرمال، یا برآورد مقدار میانگین و انحراف معیار، ترجمه نمود. با اینحال، مدل‌های یادگیری عمیق بیزی اغلب ریاضیات پیچیده‌تری داشته و محاسبات سنگین‌تری را الزام می‌کنند.

جدول ۱ این اعضا و تعداد پارامترهای هر کدام را به تفکیک نشان می‌دهد.

یکی از روش‌هایی که به منظور تسریع فرآیند یادگیری مدل‌های عمیق استفاده می‌شود، انتقال یادگیری^۱ است [۵۲]. شیوه‌های گوناگونی برای انتقال یادگیری وجود دارد که یکی از پرکاربردترین آنها استفاده از وزن‌های بدست‌آمده از طریق آموزش شبکه بر روی یک مساله طبقه‌بندی (مثل طبقه‌بندی مجموعه داده‌ی ImageNet) می‌باشد. استفاده از وزن‌های از پیش آماده، علاوه بر بهبود عملکرد الگوریتم، حجم محاسباتی را به شدت کاهش می‌دهد. در این مطالعه از وزن‌های آموزش‌دیده خانواده ResNeSt به منظور انتقال یادگیری استفاده شده‌است.

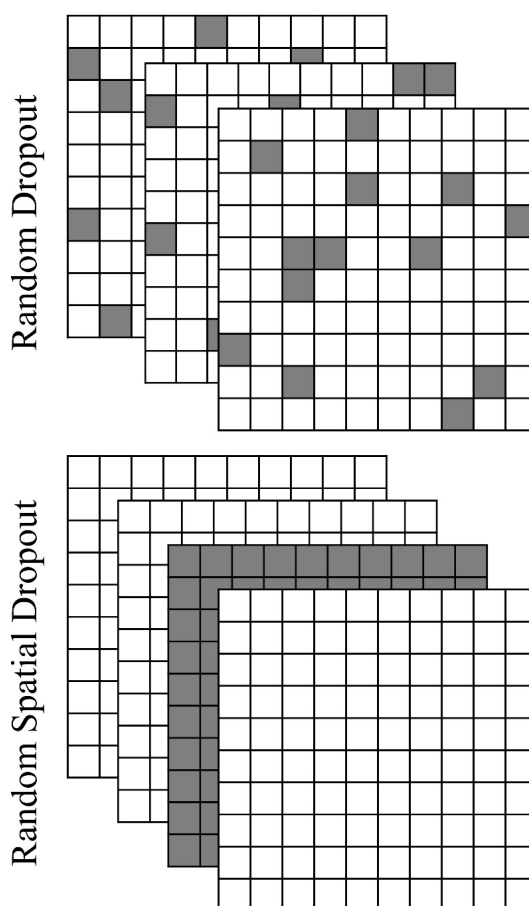
۳-۳- دورریز به عنوان یک تخمین‌گر بیزی

شبکه‌های یادگیری عمیق به خودی خود، به عنوان ابزارهایی برای حل مسائل مرتبط با حوزه یادگیری ماشین، امکان محاسبه عدم قطعیت مدل را ندارند. علاوه بر این، در مسائل طبقه‌بندی یا قطع‌بندی معنایی، اغلب به اشتباه، خروجی‌های احتمالاتی Softmax به عنوان قطعیت یا اعتمادپذیری مدل تفسیر می‌شوند. اما باید دانست که یک مدل حتی با ارائه مقادیر بالایی از Softmax می‌تواند

^۲ Bayesian Probability Theory

^۱ Transfer learning

گل و همکارانش [۴۷] نشان دادند که با استفاده از فن دورریز، جهت عدم اعمال برخی از وزن‌های کرنل‌های شبکه به صورت تصادفی، می‌توان به تخمین مناسبی از تابع توزیع نرمال پیش‌بینی‌ها دست پیدا کرد؛ که در نتیجه منجر به برآورد مناسبی از عدم قطعیت مدل می‌شود. به این روش، روش دورریز مونت-کارلو (MCD) نیز گفته می‌شود.



شکل ۳- نمایش مقایسه‌ی دورریز معمولی در برابر دورریز مکانی. پیکسل‌ها و لایه تیره نمایانگر مقادیر دور ریخته‌شده به صورت تصادفی هستند.

این روش با اینکه ساده است، نه تنها بار محاسباتی بیشتری را بوجود نمی‌آورد، بلکه دقت پیش‌بینی‌ها را نیز کاهش نمی‌دهد. در این روش، برخلاف حالت معمول که دورریز به عنوان یک تنظیم‌کننده^۱ در زمان آموزش شبکه استفاده می‌شد تا از بیش‌برازش^۲ جلوگیری کند، از آن در زمان پیش‌بینی و آزمایش استفاده می‌شود. در روش اولیه با خاموش کردن تعدادی از وزن‌های کرنل‌ها، و پیش‌بینی قطعه‌بندی معنایی برای یک تصویر به دفعات مکرر، امکان

محاسبه میانگین و انحراف معیار پیش‌بینی‌ها بدست می‌آید. با اینحال، از آنجا که وزن‌های کرنل‌ها همبستگی مکانی بالایی با یکدیگر دارند، خاموش کردن وزن‌ها به صورت تصادفی در یک لایه خیلی کارآمد نبود. از این رو، به جهت بهبود عملکرد دورریز و با الهام از مقاله تامسون و همکارانش [۵۳]، این عملیات را بر روی ویژگی‌ها انجام دادیم. در این حالت وزن‌های مربوط به یک کرنل به صورت کامل خاموش می‌شوند تا تاثیر همبستگی مکانی میان وزن‌ها از بین رفته و اثر خاموش شدن وزن‌ها به صورت کامل اعمال شود. به این روش دورریز مکانی یا Spatial Dropout گفته می‌شود. در شکل ۳ این مفهوم به صورت بصری نشان داده شده‌است.

۳-۴- ارزیابی و اعتبارسنجی

یکی از مراحل کلیدی در ارزیابی کیفیت و عملکرد یک مدل، استفاده از معیارهای اعتبارسنجی است. به همین منظور از داده‌های واقعی زمینی جهت مقایسه خروجی‌های مدل و برآورد آماره‌های کیفیت استفاده می‌شود. در ابتدا متغیرهایی که نشان‌دهنده تعداد پیکسل‌هایی هستند که درست یا اشتباه طبقه‌بندی شده‌اند، تعریف می‌شوند. سپس بر مبنای آنها، سایر معیارهای ارزیابی همچون Accuracy, Precision, Recall, F1-Score و IoU تعریف می‌شوند. روابط زیر به معرفی هر کدام از این معیارهای ارزیابی می‌پردازند:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (۱)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (۲)$$

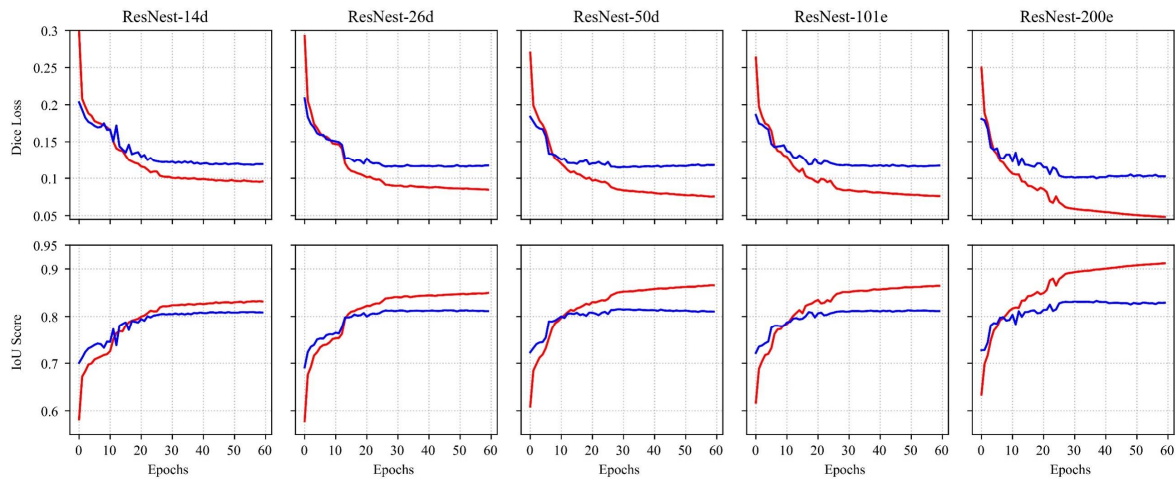
$$F1_{\text{score}} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (۳)$$

$$F1_{\text{score}} = \frac{2 \times TP}{2 \times TP + FP + FN}$$

$$\text{IoU} = \frac{TP}{TP + FP + FN} \quad (۴)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (۵)$$

در روابط بالا، متغیرهای TP، FP، FN و TN با توجه به ماتریس ابهام تعریف می‌شوند.



شکل ۴- نمودارهای پیش روند آموزش شبکه‌ها. هر ستون نمودار مربوط به یکی از رمزگذارها را نشان می‌دهد.

۴- نتایج تحقیق

در این بخش به ارائه نتایج بدست آمده از روش توسعه داده شده می‌پردازیم. در ابتدا جزئیات عملی پیاده‌سازی الگوریتم‌ها را بیان کرده و سپس نتایج حاصل را ارائه کرده ایم.

۴-۱- جزئیات پیاده‌سازی

در آموزش تمامی حالات مختلف مدل‌های یادگیری عمیق مورد استفاده در این مطالعه، از پارامترهای یکسان استفاده شده است تا مقایسه درست امکان پذیر باشد. بدین منظور، مدل‌ها با تابع بهینه کننده^۱ Adam و نرخ یادگیری^۲ ۰,۰۰۰۱ به تعداد ۶۰ دور^۳ آموزش داده شدند. همچنین در دور ۳۰ام نرخ یادگیری با ضریب ۰,۱ کاهش یافت تا یادگیری شبکه با روند بهتری صورت پذیرد.

با توجه به محدودیت حجم پردازش‌ها بر روی پردازنده گرافیکی، ابعاد تصاویر به ۵۱۲×۵۱۲ پیکسل کاهش یافت تا علاوه بر حفظ جزئیات تصویر تا حد مورد نیاز، از لبریز شدن حافظه گرافیکی در زمان آموزش نیز جلوگیری شود. همچنین تصاویر ورودی در قالب بسته‌های^۴ ۵ تایی وارد شبکه می‌شدند تا تعادلی میان تعداد پارامترهای شبکه و حافظه پردازنده گرافیکی برقرار شود. تمامی پیاده‌سازی‌های این مطالعه در زبان برنامه‌نویسی پایتون و توسط کتابخانه پایتورچ انجام گرفته است. به همین جهت از یک سرور مجهز به پردازنده گرافیکی Nvidia 2080Ti جهت انجام تمامی مراحل

آموزش و آزمایش مدل‌ها استفاده شده است. لازم به ذکر است که در آموزش اغلب مدل‌های یادگیری عمیق، مقدار حافظه پردازنده گرافیکی را می‌توان به عنوان گلوگاه^۵ اصلی آموزش در نظر گرفت و توان پردازشی CPU و RAM سیستم در اولویت‌های بعدی قرار دارند.

۴-۲- تابع زیان^۶

توابع زیان اختلاف میان خروجی پیش‌بینی شده و مقادیر واقعی را اندازه‌گیری می‌کنند. توابع زیان مختلفی متناسب با اهداف هر مساله تعریف می‌شوند. در این مقاله، با هدف قطعه‌بندی معنایی تصاویر هوایی که شامل عوارض با ابعاد مختلف هستند، از تابع زیان Dice [۵۴] استفاده کردیم. در واقع این تابع میزان همپوشانی قطعه‌های پیش‌بینی شده در مقایسه با قطعات واقعی را اندازه‌گیری می‌کند. رابطه مورد استفاده برای تابع زیان Dice به صورت زیر می‌باشد:

$$\text{Dice Loss} = 1 - \frac{2y\hat{y} + 1}{y + \hat{y} + 1}$$

$$\text{Dice Loss} \cong 1 - \frac{2 \times y \cap \hat{y}}{y \cup \hat{y}} \quad (6)$$

در رابطه بالا y مقدار واقعی و \hat{y} مقدار پیش‌بینی شده را نمایش می‌دهند. به بیان بهتر، این تابع زیان ۲ برابر مقدار همپوشانی دو قطعه را بر مجموع مساحت هر دو قطعه تقسیم می‌کند. هرچه این مقدار به صفر نزدیکتر باشد بهتر است.

۴ Batch size
۵ Bottleneck
۶ Loss function

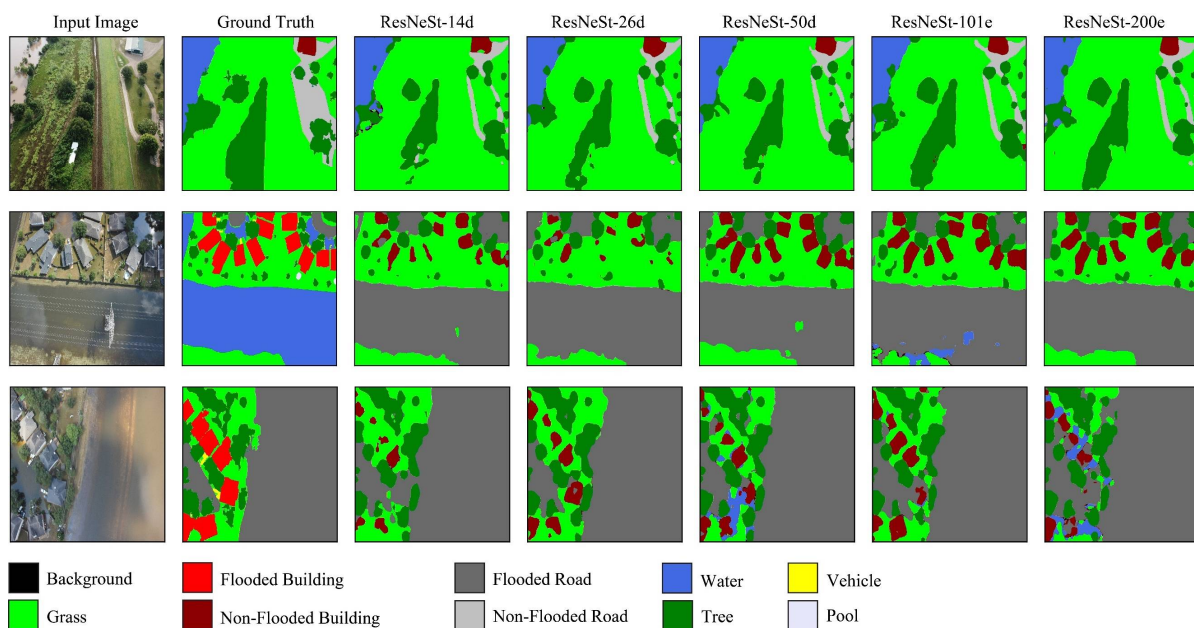
۱ Optimization function
۲ Learning rate
۳ Epoch

کاهش ایجاد نماید. این موضوع را می‌توان در ساختمان‌های استخراج شده در تصویر وسط یا آب‌گرفتگی‌های موجود در تصویر پایین مشاهده نمود. با اینحال، ارزیابی عملکرد الگوریتم‌ها بر روی تصاویر آزمایشی نشان دهنده عملکرد بسیار نزدیک روش‌ها به یکدیگر است به گونه‌ای که اختلاف مدل‌ها در معیارهای مختلف کیفیت حدود ۳٪ است. علاوه بر پیش‌بینی قطعه‌بندی معنایی تصاویر پهپادی، مقادیری در قالب عدم قطعیت نیز برای هر پیکسل برآورد شده‌است که نشان‌دهنده میزان اطمینان‌پذیری کلاس مربوط به هرکدام از پیکسل‌ها می‌باشد. شکل‌های ۶ و ۷ میانگین پیش‌بینی‌های حاصل از ۱۰۰ مرتبه اجرای مختلف روش MCD را با مقادیر دورریز ۰٫۴، ۰٫۵، ۰٫۶، ۰٫۷، ۰٫۸ و ۰٫۹ نمایش می‌دهند. به این ترتیب در هر اجرا، به منظور برآورد عدم قطعیت، ۴۰٪، ۵۰٪، ۶۰٪، ۷۰٪، ۸۰٪ یا ۹۰٪ از ویژگی‌هایی که به لایه قطعه‌بندی معنایی وارد می‌شوند به صورت تصادفی خاموش شده و پیش‌بینی بر اساس وزن‌های باقیمانده صورت می‌گیرد. نکته قابل توجه این است که برای مقادیر مختلف دورریز، میانگین خروجی‌ها که برآوردی از پیش‌بینی قطعی مدل‌ها است، با تخمین بسیار بالایی با یکدیگر برابر است که نشان‌دهنده ثبات پیش‌بینی‌ها در دفعات زیاد تکرار است. به بیان دیگر، رای حداکثری^۱ مدل‌ها در ۱۰۰ مرتبه اجرای مختلف، به مقدار پیش‌بینی بدون هرگونه دورریز نزدیک می‌شود.

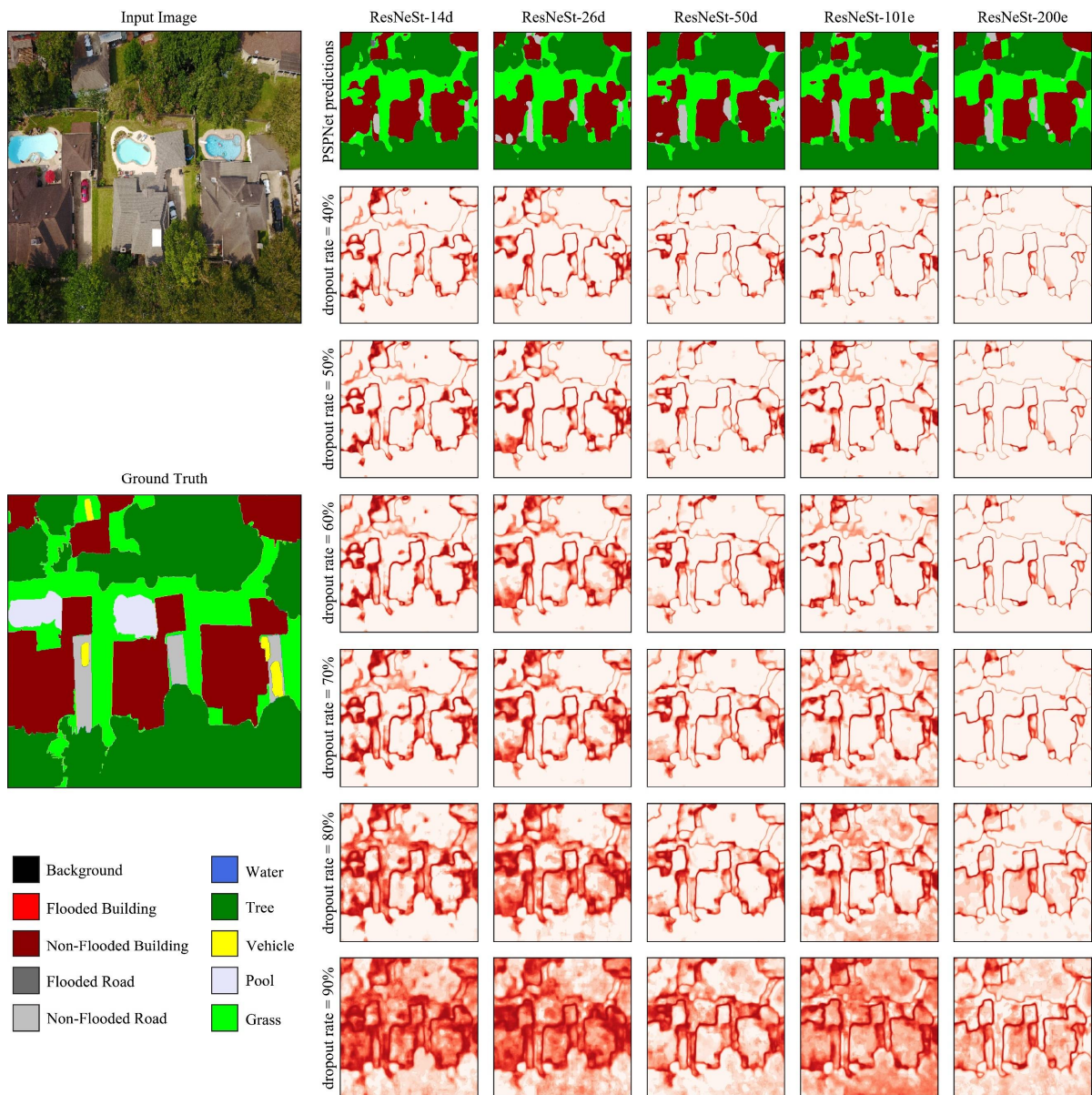
به منظور پایش روند آموزش شبکه‌ها، در کنار تابع زیان، از امتیاز IoU که هر دو بر روی داده‌های آموزشی و اعتبارسنجی برآورد می‌شدند، نیز استفاده کردیم. شکل ۴ نمودارهای پایش فرآیند آموزش شبکه را نشان می‌دهد.

۴-۳- خروجی‌ها و مقایسه نتایج

در این مطالعه از معماری PSPNet به عنوان شبکه اصلی و از ۵ عضو از خانواده ResNeSt به عنوان رمزگذارهای مختلف استفاده کردیم. تمامی رمزگذارها پیش از استفاده توسط مجموعه داده ImageNet آموزش دیده بودند و از وزن‌های بدست آمده از آن جهت بهبود سرعت و عملکرد الگوریتم استفاده کرده‌ایم. این رمزگذارها که در جدول ۱ نشان‌دهنده شده‌اند، در ساختار معماری شبکه با یکدیگر مشابه و در تعداد لایه‌ها و تعداد پارامترها متفاوت هستند. این تفاوت سبب یادگیری متفاوت هرکدام از آنها می‌شود؛ به گونه‌ای که در شکل ۴ می‌توان بهبود پلکانی فرآیند یادگیری از ResNeSt-14d تا ResNeSt-200e را مشاهده نمود. همچنین شکل ۵ خروجی هرکدام از ۵ شبکه مختلف را در کنار تصویر ورودی و نقشه واقعیت زمینی نمایش می‌دهد. همانطور که مشاهده می‌شود، شبکه مجهز به رمزگذار ResNeSt-200e با بیشترین تعداد پارامتر توانسته‌است بهتر از سایر مدل‌ها نتایج را پیش‌بینی نماید و تا ۱۰٪ در معیار IoU افزایش و ۵٪ در Loss



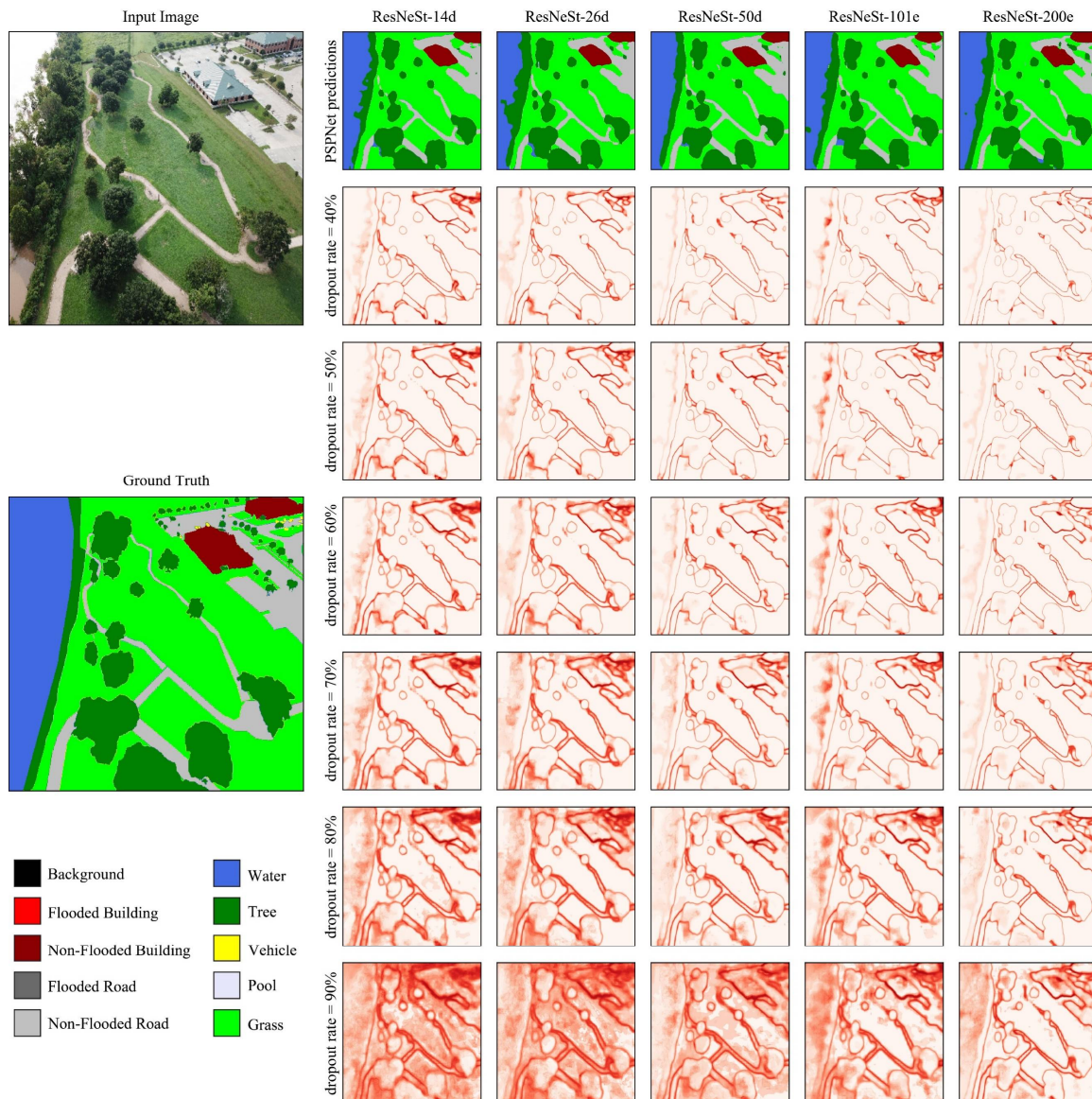
شکل ۵- مقایسه خروجی‌های مدل‌های مختلف در کنار تصویر ورودی و نقشه واقعیت زمینی



شکل ۶- نتایج حاصل از پیش‌بینی مدل‌ها با مقادیر دورریز مختلف

جدول ۱- ویژگی‌های شبکه‌های رمزگذار مختلف به همراه پارامترهای ارزیابی دقت مدل‌ها

معیار ارزیابی مدل‌ها						مدل PSPNet	پارامترها (میلیون)	شبکه رمزگذار
F1-Score	Precision	Recall	Accuracy	IoU Score	Dice loss			
۰٫۸۹۳۴	۰٫۸۹۳۸	۰٫۸۹۳۰	۰٫۹۷۸۷	۰٫۸۲۹۴	۰٫۱۰۸	ResNeSt-14d	۸	
۰٫۸۹۶۱	۰٫۸۹۶۳	۰٫۸۹۵۸	۰٫۹۷۹۲	۰٫۸۳۳۴	۰٫۱۰۵۱	ResNeSt-26d	۱۵	
۰٫۸۹۷۷	۰٫۸۹۸۰	۰٫۸۹۷۵	۰٫۹۷۹۶	۰٫۸۳۵۸	۰٫۱۰۳۲	ResNeSt-50d	۲۵	
۰٫۸۹۶۳	۰٫۸۹۶۵	۰٫۸۹۶۱	۰٫۹۷۹۳	۰٫۸۳۴۲	۰٫۱۰۴۴	ResNeSt-101e	۴۶	
۰٫۸۹۶۲	۰٫۸۹۶۴	۰٫۸۹۶۰	۰٫۹۷۹۳	۰٫۸۵۲۳	۰٫۰۹۰۳	ResNeSt-200e	۶۸	



شکل ۷- نتایج حاصل از پیش‌بینی مدل‌ها با مقادیر دورریز مختلف

همچنان در پیش‌بینی‌های خود مصمم است. نواحی قرمز رنگ در شکل‌های ۶ و ۷ نشان‌دهنده عدم قطعیت مدل هستند و با پررنگ‌تر شدن آنها، مقدار عددی عدم قطعیت نیز افزایش می‌یابد.

۵- بحث و بررسی

مجموعه داده‌ی FloodNet که در این مطالعه مورد استفاده قرار گرفته‌است، یکی از منحصر‌بفردترین داده‌های موجود در حوزه ارزیابی آثار وقوع بلایای طبیعی است. این داده که توسط پهباد تجاری، از ارتفاع پایین و بلافاصله پس از وقوع سیلاب اخذ شده‌است، ۹ کلاس پوشش اراضی را

با دقت در نتایج مشخص می‌شود که گرچه نتایج حاصل از شبکه‌ی دارای بیشترین تعداد پارامتر (ResNeSt-200e) با نتایج حاصل از شبکه‌ی دارای کمترین تعداد پارامتر (ResNeSt-14d) شباهت زیادی دارد؛ با اینحال مقادیر عدم قطعیت در این مدل‌ها بسیار متفاوت است. به عبارتی، شبکه‌ای که تعداد پارامتر بیشتری دارد، اطمینان‌پذیری بیشتری نیز داشته و به پیش‌بینی‌های خود مطمئن‌تر است. این امر با افزایش درصد دورریز وزن‌ها نیز مجدداً تایید می‌شود. به گونه‌ای که شبکه‌ای که پارامتر کمتری دارد، با افزایش میزان دورریز وزن‌ها، بیشتر دچار ابهام و عدم قطعیت می‌شود، حال آنکه شبکه قوی‌تر

۶- نتیجه‌گیری و پیشنهادات

در این مقاله با هدف قطعه‌بندی معنایی تصاویر اخذ شده توسط پهپاد پس از وقوع سیلاب در ایالت نگزاس در امریکا، از الگوریتم PSPNet به همراه رمزگذارهای از پیش آموزش‌دیده خانواده ResNeSt استفاده شده‌است. همچنین، از آنجایی که روش‌های مبتنی بر نظریه احتمالات بیز امکان برآورد درستی از عدم قطعیت مدل‌های یادگیری عمیق را فراهم می‌آورند؛ و از آنجایی که پیاده‌سازی این روش‌ها با دشواری‌های ریاضیاتی همراه است؛ جهت برآورد تخمینی از میزان عدم قطعیت پیش‌بینی‌های شبکه از روش دورریز مونت-کارلو استفاده کردیم. این روش که با تکرار فرآیند پیش‌بینی در دفعات مکرر (به عنوان مثال ۱۰۰ مرتبه)، امکان برآورد تخمینی از توزیع تصمیم‌گیری‌ها (میانگین و انحراف معیار) را فراهم می‌کند، در واقع نتایج حاصل از نظریه احتمالات بیز را شبیه‌سازی می‌نماید. از این رو، مقادیر عدم قطعیت حاصل از این روش قابل اعتماد می‌باشد. مقایسه رمزگذارهای مختلف با یکدیگر نشان داد که افزایش عمق و تعداد پارامترهای مدل در بهبود عملکرد مدل موثر بوده‌است. توجه به نتایج حاصل از مقادیر دورریز مختلف تایید می‌کند که افزایش پارامترهای مدل منجر به افزایش پایداری نتایج و همچنین کاهش عدم قطعیت آنها می‌شود.

پیشنهاد می‌شود در مطالعات آتی عدم قطعیت‌های حاصل از معماری‌های یادگیری عمیق مختلف با یکدیگر مقایسه شوند. همچنین با توجه به همپوشانی زیادی که تصاویر مجموعه داده FloodNet دارند، پیشنهاد می‌شود تا یک مرحله انتخاب تصویر انجام شود و کاهش تعداد نمونه‌های آموزشی به عنوان یکی از عوامل موثر در یادگیری مدل‌ها مورد بررسی قرار گیرد. همچنین، با توجه به نامتوازن بودن کلاس‌ها در این مجموعه داده، پیشنهاد می‌شود تا از روش‌های افزایش داده نیز استفاده شده و نتایج آن بررسی شود.

جهت قطعه‌بندی معنایی معرفی می‌نماید. پرواز پهپاد در ارتفاع‌های مختلف، تغییر زاویه تصویربرداری، تغییر شدت نور در تصاویر و وجود بازتاب ابر یا خورشید بر روی آنها از جمله ویژگی‌هایی هستند که این داده را تبدیل به یک داده چالشی برای مساله قطعه‌بندی معنایی می‌کنند. همچنین وجود دو کلاس مختلف برای آب‌های ناشی از سیل در مقابل رودخانه و دریاچه، تصمیم‌گیری برای الگوریتم‌ها را دشوار می‌سازد. به اضافه، همانطور که در شکل ۱ نشان داده شده‌است، پراکندگی کلاس‌ها در این مجموعه داده نامتوازن است و آموزش الگوریتم را پیچیده‌تر می‌سازد.

علاوه بر این، تمامی تصاویر برچسب‌گذاری شده در این مجموعه داده با دقت کافی تهیه نشده‌اند، چرا که کلاس اختصاص داده‌شده به برخی از پیکسل‌ها اشتباه است. هرچند که این موضوع امکان استفاده از این داده را برای مسائل نظارت‌شده ضعیف^۱ فراهم می‌آورد، با اینحال، قطعه‌بندی معنایی پیکسل‌ها را در مطالعه ما با مشکل همراه می‌سازد؛ زیرا بررسی دقت خروجی‌های مدل‌ها بر اساس داده‌های واقعیت زمینی دچار اشتباه می‌شود.

استخراج دقیق مرز ساختمان‌ها، تشخیص جاده‌های سیل‌گرفته، طبقه‌بندی دقیق پیکسل‌های درختان و برآورد مناسب عدم قطعیت‌ها از جمله نقاط قوت روش ارائه‌شده هستند. در مقابل، از جمله مشاهداتی که محدودیت‌های الگوریتم را نشان می‌دهند، عدم استخراج درست پیکسل‌های منتسب به خودروها و استخرها می‌باشد. کوچک بودن تصاویر ورودی به شبکه و تابع زیان نامناسب می‌تواند از جمله دلایلی باشند که منجر به بروز این نوع خطا می‌شوند. همچنین، با توجه به تعریف ساختمان آب‌گرفته (که سیل حداقل به یکی از اضلاع آن رسیده است)، طبقه‌بندی درست این گونه ساختمان‌ها نیز با مشکل مواجه شده‌است.

مراجع

- [۱] C. Lin, Y. Li, Y. Liu, X. Wang, and S. Geng, "Building damage assessment from post-hurricane imageries using unsupervised domain adaptation with enhanced feature discrimination," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–10, 2021.
- [۲] S. S. Matin and B. Pradhan, "Earthquake-Induced Building-Damage Mapping Using Explainable AI (XAI)," *Sensors*, vol. 21, no. 13, p. 4489, Jun. 2021, doi: 10.3390/s21134489.

^۱ Weakly supervised

- [۳] T. G. J. Rudner *et al.*, "Multi3Net: Segmenting Flooded Buildings via Fusion of Multiresolution, Multisensor, and Multitemporal Satellite Imagery," 2018, [Online]. Available: <http://arxiv.org/abs/1812.01756>
- [۴] M. N. Abdel-Mooty, A. Yosri, W. El-Dakhkhni, and P. Coulbaly, "Community Flood Resilience Categorization Framework," *International Journal of Disaster Risk Reduction*, vol. 61, p. 102349, 2021, doi: <https://doi.org/10.1016/j.ijdr.2021.102349>.
- [۵] O. M. Nofal and J. W. van de Lindt, "High-resolution flood risk approach to quantify the impact of policy change on flood losses at community-level," *International Journal of Disaster Risk Reduction*, vol. 62, p. 102429, 2021, doi: <https://doi.org/10.1016/j.ijdr.2021.102429>.
- [۶] U. Iqbal, P. Perez, W. Li, and J. Barthelemy, "How computer vision can facilitate flood management: A systematic review," *International Journal of Disaster Risk Reduction*, vol. 53, p. 102030, 2021, doi: <https://doi.org/10.1016/j.ijdr.2020.102030>.
- [۷] M. Zeleňáková, "Flood risk assessment and management in Slovakia," *WIT Transactions on Ecology and the Environment*, vol. 146, pp. 61–70, 2011.
- [۸] X. Wang and P. Li, "Extraction of urban building damage using spectral, height and corner information from VHR satellite images and airborne LiDAR data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, pp. 322–336, 2020.
- [۹] L. Dong and J. Shan, "A comprehensive review of earthquake-induced building damage detection with remote sensing techniques," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 84, pp. 85–99, Oct. 2013, doi: [10.1016/j.isprsjprs.2013.06.011](https://doi.org/10.1016/j.isprsjprs.2013.06.011).
- [۱۰] S. Cho, H. Xiu, and M. Matsuoka, "Backscattering Characteristics of SAR Images in Damaged Buildings Due to the 2016 Kumamoto Earthquake," *Remote Sens (Basel)*, vol. 15, no. 8, p. 2181, Apr. 2023, doi: [10.3390/rs15082181](https://doi.org/10.3390/rs15082181).
- [۱۱] M. Chini, N. Pierdicca, and W. J. Emery, "Exploiting SAR and VHR optical images to quantify damage caused by the 2003 Bam earthquake," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 1, pp. 145–152, 2008.
- [۱۲] N. Kerle, F. Nex, M. Gerke, D. Duarte, and A. Vetrivel, "UAV-based structural damage mapping: A review," *ISPRS Int J Geoinf*, vol. 9, no. 1, pp. 1–23, 2019, doi: [10.3390/ijgi9010014](https://doi.org/10.3390/ijgi9010014).
- [۱۳] S. Ghaffarian and N. Kerle, "Towards post-disaster debris identification for precise damage and recovery assessments from uav and satellite images," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, vol. 42, no. 2/W13, pp. 297–302, 2019, doi: [10.5194/isprs-archives-XLII-2-W13-297-2019](https://doi.org/10.5194/isprs-archives-XLII-2-W13-297-2019).
- [۱۴] B. Bauer-Marschallinger *et al.*, "Satellite-Based Flood Mapping through Bayesian Inference from a Sentinel-1 SAR Datacube," *Remote Sens (Basel)*, vol. 14, no. 15, p. 3673, 2022.
- [۱۵] J. Zhao *et al.*, "Urban-aware U-Net for large-scale urban flood mapping using multitemporal sentinel-1 intensity and interferometric coherence," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–21, 2022.
- [۱۶] C. Krullikowski *et al.*, "Estimating ensemble likelihoods for the Sentinel-1 based Global Flood Monitoring product of the Copernicus Emergency Management Service," *arXiv preprint arXiv:2304.12488*, 2023.
- [۱۷] E. Hamidi, B. G. Peter, D. F. Muñoz, H. Moftakhari, and H. Moradkhani, "Fast Flood Extent Monitoring With SAR Change Detection Using Google Earth Engine," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–19, 2023.
- [۱۸] S. Kundu, V. Lakshmi, and R. Torres, "Estimation of flood inundation and depth during Hurricane Florence using Sentinel-1 and UAVSAR data," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.
- [۱۹] A. F. S. Putri, W. Widyatmanti, and D. A. Umarhadi, "Sentinel-1 and Sentinel-2 data fusion to distinguish building damage level of the 2018 Lombok Earthquake," *Remote Sens Appl*, vol. 26, p. 100724, 2022.
- [۲۰] S. A. Ahmadi, A. Mohammadzadeh, N. Yokoya, and A. Ghorbanian, "BD-SKUNet: Selective-Kernel UNets for Building Damage Assessment in High-Resolution Satellite Images," *Remote Sens (Basel)*, vol. 16, no. 1, 2024, doi: [10.3390/rs16010182](https://doi.org/10.3390/rs16010182).
- [۲۱] L. Hashemi-Beni and A. A. Gebrehiwot, "Flood Extent Mapping: An Integrated Method Using Deep Learning and Region Growing Using UAV Optical Data," *IEEE J Sel Top Appl Earth Obs Remote Sens*, vol. 14, pp. 2127–2135, 2021, doi: [10.1109/JSTARS.2021.3051873](https://doi.org/10.1109/JSTARS.2021.3051873).

- [۲۲] L. Hashemi-Beni, J. Jones, G. Thompson, C. Johnson, and A. Gebrehiwot, "Challenges and Opportunities for UAV-Based Digital Elevation Model Generation for Flood-Risk Management: A Case of Princeville, North Carolina," *Sensors*, vol. 18, no. 11, 2018, doi: 10.3390/s18113843.
- [۲۳] H. Yao, R. Qin, and X. Chen, "Unmanned Aerial Vehicle for Remote Sensing Applications—A Review," *Remote Sens (Basel)*, vol. 11, no. 12, 2019, doi: 10.3390/rs11121443.
- [۲۴] V. V. Klemas, "Coastal and Environmental Remote Sensing from Unmanned Aerial Vehicles: An Overview," *J Coast Res*, vol. 31, no. 5, pp. 1260–1267, 2015, doi: 10.2112/JCOASTRES-D-15-00005.1.
- [۲۵] F. Yamazaki, Y. Yano, and M. Matsuoka, "Visual damage interpretation of buildings in Bam city using QuickBird images following the 2003 Bam, Iran, earthquake," *Earthquake Spectra*, vol. 21, no. 1_suppl, pp. 329–336, 2005.
- [۲۶] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: An overview," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6690–6709, 2019.
- [۲۷] Z.-Q. Zhao, P. Zheng, S. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans Neural Netw Learn Syst*, vol. 30, no. 11, pp. 3212–3232, 2019.
- [۲۸] G. Tsagkatakis, A. Aidini, K. Fotiadou, M. Giannopoulos, A. Pentari, and P. Tsakalides, "Survey of deep-learning approaches for remote sensing observation enhancement," *Sensors (Switzerland)*, vol. 19, no. 18, pp. 1–39, 2019, doi: 10.3390/s19183929.
- [۲۹] Y. Gu, Y. Wang, and Y. Li, "A survey on deep learning-driven remote sensing image scene understanding: Scene classification, scene retrieval and scene-guided object detection," *Applied Sciences (Switzerland)*, vol. 9, no. 10, 2019, doi: 10.3390/app9102110.
- [۳۰] X. Y. Tong *et al.*, "Land-cover classification with high-resolution remote sensing images using transferable deep models," *Remote Sens Environ*, vol. 237, no. July 2018, p. 111322, 2020, doi: 10.1016/j.rse.2019.111322.
- [۳۱] Y. Da, Z. Ji, and Y. Zhou, "Building damage assessment based on siamese hierarchical transformer framework," *Mathematics*, vol. 10, no. 11, p. 1898, 2022.
- [۳۲] M. Marjani, S. A. Ahmadi, and M. Mahdianpari, "FirePred: A hybrid multi-temporal convolutional neural network model for wildfire spread prediction," *Ecol Inform*, vol. 78, p. 102282, Dec. 2023, doi: 10.1016/j.ecoinf.2023.102282.
- [۳۳] M. Schmitt *et al.*, "There Are No Data Like More Data: *Datasets for deep learning in Earth observation*," *IEEE Geosci Remote Sens Mag*, pp. 2–36, 2023, doi: 10.1109/MGRS.2023.3293459.
- [۳۴] N. Humaira, S. Samadi, and N. C. Hubig, "DX-FloodLine: End-To-End Deep Explainable Pipeline for Real Time Flood Scene Object Detection from Multimedia Images," *IEEE Access*, 2023.
- [۳۵] A. Gebrehiwot, L. Hashemi-Beni, G. Thompson, P. Kordjamshidi, and T. E. Langan, "Deep Convolutional Neural Network for Flood Extent Mapping Using Unmanned Aerial Vehicles Data," *Sensors*, vol. 19, no. 7, 2019, doi: 10.3390/s19071486.
- [۳۶] K. Yang, S. Zhang, X. Yang, N. Wu, and others, "Flood Detection Based on Unmanned Aerial Vehicle System and Deep Learning," *Complexity*, vol. 2022, 2022.
- [۳۷] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.
- [۳۸] A. Khan, A. Sohail, U. Zahoor, and A. S. Qureshi, "A Survey of the Recent Architectures of Deep Convolutional Neural Networks," pp. 1–69, 2019, [Online]. Available: <http://arxiv.org/abs/1901.06032>
- [۳۹] S. K. He and X. Wang, "Spatial pyramid pooling in deep convolutional neural network for visual recognition," in *CVPR*, 2014.
- [۴۰] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2881–2890.
- [۴۱] O. Elharrouss, Y. Akbari, N. Almaadeed, and S. Al-Maadeed, "Backbones-review: Feature extraction networks for deep learning and deep reinforcement learning approaches," *arXiv preprint arXiv:2206.08016*, 2022.
- [۴۲] H. Zhang *et al.*, "Resnest: Split-attention networks," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 2736–2746.

- [۴۳] E. Tjoa and C. Guan, "A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI," *IEEE Trans Neural Netw Learn Syst*, vol. 32, no. 11, pp. 4793–4813, 2021, doi: 10.1109/TNNLS.2020.3027314.
- [۴۴] A. Singh, S. Sengupta, and V. Lakshminarayanan, "Explainable deep learning models in medical image analysis," *J Imaging*, vol. 6, no. 6, pp. 1–18, 2020, doi: 10.3390/JIMAGING6060052.
- [۴۵] M. Segú, A. Loquercio, and D. Scaramuzza, "A General Framework for Uncertainty Estimation in Deep Learning," *ArXiv*, 2019.
- [۴۶] W. J. Maddox, T. Garipov, P. Izmailov, D. Vetrov, and A. G. Wilson, "A simple baseline for bayesian uncertainty in deep learning," *ArXiv*, no. NeurIPS, pp. 1–25, 2019.
- [۴۷] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning," in *Proceedings of The 33rd International Conference on Machine Learning*, M. F. Balcan and K. Q. Weinberger, Eds., in *Proceedings of Machine Learning Research*, vol. 48. New York, New York, USA: PMLR, Nov. 2016, pp. 1050–1059. [Online]. Available: <https://proceedings.mlr.press/v48/gal16.html>
- [۴۸] M. Rahnemounfar, T. Chowdhury, A. Sarkar, D. Varshney, M. Yari, and R. R. Murphy, "FloodNet: A High Resolution Aerial Imagery Dataset for Post Flood Scene Understanding," *IEEE Access*, vol. 9, pp. 89644–89654, 2021, doi: 10.1109/ACCESS.2021.3090981.
- [۴۹] W. Liu, A. Rabinovich, and A. C. Berg, "Parsenet: Looking wider to see better," *arXiv preprint arXiv:1506.04579*, 2015.
- [۵۰] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE Trans Pattern Anal Mach Intell*, vol. 40, no. 4, pp. 834–848, 2017.
- [۵۱] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-Decem, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.
- [۵۲] J. Xia, B. Adriano, G. Baier, and N. Yokoya, "Building Damage Mapping Via Transfer Learning," *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 4841–4844, 2019, doi: 10.1109/igarss.2019.8900447.
- [۵۳] J. Tompson, R. Goroshin, A. Jain, Y. LeCun, and C. Bregler, "Efficient object localization using Convolutional Networks," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 648–656, 2014, [Online]. Available: <https://api.semanticscholar.org/CorpusID:206592615>
- [۵۴] S. Jadon, "A survey of loss functions for semantic segmentation," in *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 2020, pp. 1–7. doi: 10.1109/CIBCB48159.2020.9277638.