ارائه روشی در زمینه آشکارسازی اهداف در تصاویر ماهواره‌ای با استفاده از یادگیری عمیق و با رویکرد سنجش از دور و GIS

نادر بیرانوند[1*]، مهدی کیخایی [2]، رویین ممانی [3]

[1] پژوهشگر گروه علوم پایه، مرکز مطالعات زمینی و صنایع وابسته،دانشگاه امام علی (ع)،تهران ،ایران

[2] دانشجوی دکتری سنجش از دور و GIS دانشگاه تهران تهران، ایران

[3] کارشناسی ارشد عمران، دانشگاه شهید بهشتی ، تهران، ایران

تاریخ دریافت مقاله: ۱۳۹۹/۷/۱۵ **تاریخ پذیرش مقاله: ۱۳۹۹/۱۱/۱۲**

چکیده

آشکارسازی و شناسایی خودکار عوارض موجود در مناطق مختلف جغرافیایی با دقت بالا با توجه به اهداف گوناگون اعم از مدیریت شهری، اهداف نظامی و غیره از مسائل اساسی در شناخت پدیده‌های سطح محسوب می‌گردد. تکنیک‌های سنجش از دور و سیستم اطلاعات مکانی (GIS) به همراه تصاویر اخذ شده از سنجنده‌های ماهواره‌ای با تلفیق الگوریتم‌های یادگیری عمیق نمایان‌سازی عوارض موجود در تصاویر را به خوبی ارتقاء داده است. در این پژوهش به ارائه روشی نو در تلفیق همزمان لایه‌های CNN و Pooling موجود در الگوریتم مورد استفاده پرداخته شده که در نهایت منجر به کاهش چشمگیر زمان آموزش شبکه با استفاده از داده‌های جامع آموزشی با دقت بالا و در عین حال حجم زیاد گردید. در این تحقیق با بهره‌گیری از آموزش شبکه به وسیله داده‌های آموزشی به آشکارسازی کشتی‌های موجود در تصاویر ماهواره‌ای با ایجاد یک شبکه تماماً کانولوشنال FCN پرداخته شده‌است. جهت ارزیابی عملکرد و دقت الگوریتم استفاده شده در یافتن و آشکارسازی کشتی‌های موجود در تصاویر ماهواره‌ای، با اعمال این الگوریتم آشکارساز بر روی چندین تصویر ماهواره‌ای دیگر از معیارهای ارزیابی Precision، Recall و F1-Score استفاده شد که مقادیر آن‌ها به ترتیب برابر با ٪۱۰۰، ٪ ۹۷/۶۱ و ٪ ۹۸/۸۳ بوده‌است که نمایانگر دقت و قابل اعتماد بودن الگوریتم می‌باشد.

واژه‌های کلیدی: سنجش از دور، یادگیری عمیق، شبکه عصبی کانولوشن، آشکارسازی خودکار اهداف.

مقدّمه

شناسایی اهداف[۱] در تصاویر ماهواره‌ای با بهره‌گری از علم پردازش تصویر[۲] از دیرباز یکی از مهم‌ترین اهداف کاربردی در تمامی مراحل پیشرفت سیستم‌های کامپیوتری بوده است. در دهه اخیر با پیشرفت‌های رخ داده در زمینه روش‌های آموزش ماشین[۳] و همچنین ساخت سخت افزارهای قدرتمند جهت پردازش و یافتن چنین موضوعاتی گام‌های بلندی در این زمینه برداشته شده است (Napiorkowska & Milena, 2018, 2). با توجه به مشکلاتی از قبیل نیاز به دقت بالای تصاویر ورودی به سیستم‌های طراحی شده جهت پردازش تصویر و یافتن خودکار اهداف و نیاز به سیستم‌های کامپیوتری با سخت‌افزارهای بسیار خاص و گران‌قیمت در طبقه‌بندی چنین تصاویر، روش‌هایی از قبیل آموزش عمیق داده‌ها با استفاده از شبکه‌های عصبی امکان پردازش تصاویر با دقت و همچنین حجم بسیار بالا با کامپیوترهایی با سخت‌افزارهای متوسط را میسّر ساخت (Khan & Muhammad Jaleed, 2017: 4).در پژوهشی دیگر که با سرمایه‌گذاری و تحت نظارت مرکز پایش دریانوردی اروپا[۴] (LIMES) در سال ۲۰۱۰ انجام شد، ایده اولیه استفاده از سنجش از دور تصاویر ماهواره‌ای در یافتن کشتی‌های موجود در این کشور مطرح گردید که به منظور پایش کشتی‌های موجود و در راستای نظارت بر مسیرهای ترافیکی کشتی‌رانی، جلوگیری از صید غیرمجاز آبزیان و همچنین استخراج و حمل غیرقانونی فراورده‌های نفتی بود. مشکل اصلی روش اراوه شده در این پژوهش، وابستگی شدید سیستم پردازش تصویر طراحی شده به دقت و رزولوشن بالای تصاویر ورودی بود که سرعت پردازش تصاویر به شدت افزایش پیدا می‌کرد (Corbane & Christina, 2010: 3). در پژوهشی در سال ۲۰۲۰ از آموزش عمیق[۵] داده‌های ورودی در یافتن اهداف دریایی همچون کشتی‌ها استفاده شد که با توجه به دقت بسیار بالای آن به سرعت مورد پذیرش مرکز حفاظت و امنیت کشتیرانی جنوب شرق آسیا قرار گرفت. در روش ارائه شده از الگوریتم DCCN[۶] استفاده شده‌است که دقت محاسبات برابر با ۹۸٪/۷۵ بوده‌است (Stofa & Mohd, 2020: 6). شناسایی عوارض موجود در تصاویر، یکی از مسائل اساسی در حوضه تفسیر تصاویر به ویژه در تصاویر سنجش از دوری، به شمار می آید. یکی از روش های کارآمد و به روز در این زمینه، به کارگیری شیوه‌های یادگیری عمیق، جهت استخراج و تفسیر است. یک عارضه، مجموعه‌ای از الگوهای منحصر به فرد است که با عوارض مجاور خود متفاوت است، این تفاوت معمولا در یک یا چند ویژگی به طور همزمان اتفاق می افتد که می توان به تفاوت در شکل، رنگ و یا ویژگی‌های طیفی اشاره نمود. در این راستا، روش یادگیری عمیق با توانایی تحلیل مفاهیم انتزاعی سطح بالا، می‌تواند انتخاب مناسبی در این زمینه باشد (Farhadi & Kiani, 2019: 2). امروزه سیستم اطلاعات جغرافیایی نقشی محوری در مدیریت جامع شهری و زمینه‌های دیگر همچون عملیات‌های شناسایی، نظامی و ... به عهده دارد. محاسن استفاده از یک سامانه سنجش از دور در کنار سیستم GIS نسبت به سایر فناوری‌ها، کیفیت بالای تحلیل داده‌ها و امکان تجزیه و تحلیل آن‌ها با روش‌های پیشرفته، مدیریت حجم زیاد اطلاعات و کاهش زمان و هزینه از مهم‌ترین محاسن و ضرورت‌ها در وجود این سامانه، در به ثمر رسیدن اهداف مدیریت شهری می باشد. در چنین پژوهش‌هایی پس از تهیه نقشه پهنه‌بندی منطقه، بازدید و ارزیابی‌های میدانی انجام شده حاکی از آن است که سیستم اطلاعات جغرافیایی همراه با تحلیل‌های منطقه مورد بررسی با استفاده از روش‌های سنجش از دور ، به خوبی مکان‌های مناسب با هدف مورد نظر را استخراج نموده و رویکرد دانش مبنای سیستم

---

۱ Target detection
۲ Machine learning
۳ Image processing
۴ Land and Sea Integrated Monitoring for European Security
۵ Deep learning
۶ Densely Connected Convolotional Network

اطلاعات جغرافیایی و سنجش از دور با ترکیب تجربه و دانش کارشناسان امر، باعث بهبود مدیریت شهری در راستای حفظ جان مردم در هنگام رخداد مخاطرات و بلایای طبیعی، کاهش خسارات، صرفه‌جویی در زمان و هزینه‌ها در فعالیت‌ها خواهد شد. از موارد بسیار مهم دیگری که لزوم توجه به این مکان‌یابی را بیشتر آشکار می‌سازد در مناطق مرزی و حساس است. با توجه به گستردگی وضعیت ترابری و حمل و نقل با پیشرفت های دنیای کنونی، شناسایی محل‌های مناسب برای بهود شرایط مدیریتی شهر و کشور، استفاده از اصول سنجش از دور در راستای صرفه جویی در زمان و هزینه و همچنین دقت قابل قبول آن امری ضروری است. اگرچه تکنولوژی و وسایل الکترونیکی بصری در عصر امروز ترقی و پیشرفت زیادی داشته است ولی هنوز اهمیت و نقش حضور فیزیکی در بسیاری از مناطق ملزم و ضروری است سامانه اطلاعات مکانی با تحلیل‌های مناسب برای مکان یابی نصب سامانه‌های پایشگر راداری و دیدبانی دریایی و هوایی یکی از مواردی است که کمک شایانی به مسئولان این حوزه می‌نماید (Keykhaee, Rangzan, 2016: 5). نتایج حاصل از تحلیل مناطق حساس با بهره‌گیری سنجش از دور در پژوهش‌های مختلف نشان می‌دهد یکی از مهم‌ترین و کاربردی‌ترین استفاده‌ها از این علم نوین در برنامه‌ریزی ها و مدیریت دفاعی کشور می‌باشد. در یافتن اهدافی که چارچوب خاصی برای تعیین اهمیت و ارزش معیارها وجود دارند، استفاده از روش های تصمیم گیری ناشی از شناسایی اهداف[1] مورد نیاز، از بهترین گزینه ها می باشد. در شناسایی اهداف با استفاده از یک هدف از پیش شناسایی و بررسی شده همچون یک هواپیما، کشتی یا هدف مورد نظر دیگری جهت آموزش داده‌های خام موجود که در واقع تصاویر هوایی یا ماهواره‌ای هستند جهت شناسایی اهدافی مشابه صورت می‌پذیرد.

پیشینه و مبانی نظری تحقیق:

بهره‌گیری از یادگیری عمیق در سنجش از دور

دورسنجی امروزه نقش مهمی را در شناسایی اهداف سطح زمین بدون نیاز به تماس مستقیم و روش‌های اکتشافی صحرایی ایفا می‌کند. توسعه روش‌های پردازش این داده‌ها در زمینه‌های مختلف آماده‌سازی داده‌ها و استخراج اطلاعات گوناگون مانند خصوصیات طیفی اجزاء سازنده تصویر و فراوانی آن‌ها در پیکسل‌های متشکله تصویر، تمهیدات لازم برای کاربرد این روش‌ها را بیش از پیش مهیا کرده‌است. از طرف دیگر بهره‌گیری مناسب از این روش‌های آنالیز همانند هر فن‌آوری جدیدی نیاز به بومی‌سازی آن دارد امید آنست در کشور عزیزمان نیز به لطف پژوهشگران گرانقدر در این زمینه پیشرفت‌های بسیاری در آینده‌ای نه چندان دور رخ خواهد داد (Oskooee,2015:5). یادگیری عمیق یکی از زیر شاخه‌های یادگیری ماشین است که هدف آن یاد گرفتن چکیده‌ای سطح بالا از داده‌ها با استفاده از معماری‌های سلسله مراتبی است و یک رویکرد در حال ظهور است که به طور گسترده در حوزه هوش مصنوعی مانند پردازش زبان طبیعی و بینایی ماشینی استفاده شده‌است. ابداع ایده اصلی یادگیری عمیق به سال‌های قبل باز می‌گردد اما به دلیل نبود سخت‌افزار قوی امکان پیاده سازی آن وجود نداشته و یا مشکل بوده است. امروزه سه دلیل عمده برای رونق گرفتن یادگیری عمیق وجود دارد:

۱) توانایی پردازش چیپ ها در سال های اخیر افزایش چشم‌گیری داشته است (مانند GPU ها)

۲) در هزینه تهیه و نگهداری سخت افزارهای محاسباتی کاهش قابل ملاحظه‌ای بوجود آمده است

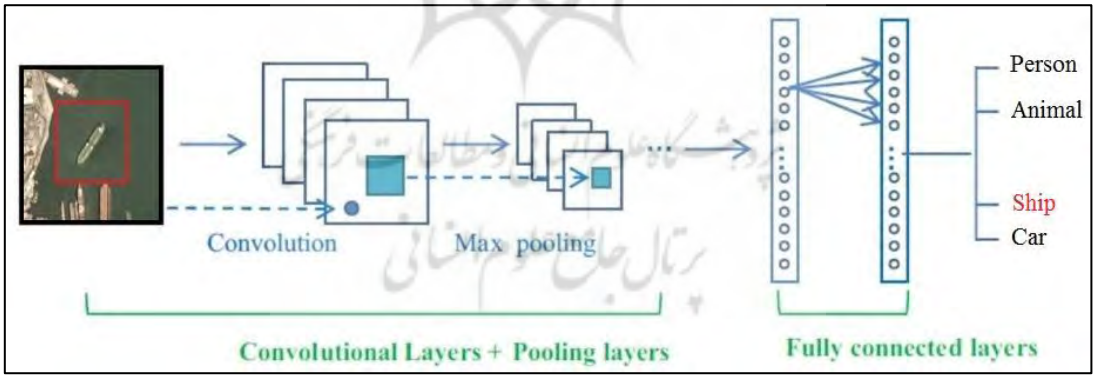
۳) پیشرفت های قابل توجه در الگوریتم‌های یادگیری ماشین به وجود آمده است و از آنجایی که شبکه‌های عمیق می‌توانند ویژگی‌های مناسبی از داده ورودی استخراج کنند، در تحقیقات اخیر نشان داده شده‌است که در وظایف بینایی ماشین موفق عمل کرده‌اند. در پژوهش های انجام شده در زمینه روش های یادگیری عمیق به طور گسترده‌ای توسط پژوهشگران استفاده شده‌اند و به بازدهی های بالایی دست یافته اند. الگوریتم های یادگیری عمیق به چهار دسته تقسیم می شوند: شبکه های عصبی

[1] Target detection

کانولوشن[۱]، ماشین بولتزمن محصور[۲]، خودرمزگذارها[۳] و کدینگ پراکنده[۴].
در پژوهشی به صورت جامع شبکه های عصبی کانولوشن بررسی شده و در ادامه نحوه استفاده از آن در پردازش تصاویر و پیدا کردن اهداف مورد نیاز در تصاویر ماهواره‌ای بیان شده است (Sezavar, 2015: 3). شاید بتوان بزرگترین مزیت یادگیری عمیق را در استخراج و یادگیری ویژگی‌ها به صورت خودکار دانست (Lotfollahi,2015: 3).

شبکه عصبی کانولوشن

شبکه‌های عصبی کانولوشن (CNN) یکی از مهم‌ترین روش های یادگیری عمیق هستند که در آنها چندین لایه با روشی قدرتمند آموزش می‌بینند (Lecun, 1998: 6). این روش بسیار کارآمد بوده و یکی از رایج‌ترین روش‌ها در کاربردهای مختلف بینایی کامپیوتر است. تصویر کلی یک معماری شبکه عصبی کانولوشن در شکل۱ نمایش داده شده‌است. بطور کلی, یک شبکه CNNاز سه لایه اصلی تشکیل می‌شود که عبارتند از : لایه کانولوشن، لایه ادغام[۵] و لایه تماماً متصل[۶] که این لایه‌های مختلف، وظایف متفاوتی را انجام می‌دهند. در شکل۱ یک معماری کلی از شبکه عصبی کانولوشن برای دسته بندی تصاویر (Krizhevsky, 2017: 8) بصورت لایه به لایه نمایش داده شده‌است. در هر شبکه عصبی کانولوشن دو مرحله برای آموزش وجود دارد: ابتدا مرحله پیش رونده[۷] و به عقب یا پس انتشار[۸]. در مرحله رو به جلو تصویر از داخل شبکه عبور داده می‌شود و بردار ویژگی آن استخراج می‌شود. با توجه به این که پیشتر مشخص شده که تصویر ورودی در چه کلاسی از کلاس‌های خروجی قرار دارد یک تابع loss تعریف شده و وزن‌های فیلترها طوری انتخاب می‌گردد که این تابع کمینه شود. به منظور انتخاب وزن‌های بهینه، از گرادیان در مرحله بازگشتی استفاده می‌گردد. گرادیان‌ها با استفاده از قانون بازگشتی در امتداد لایه‌ها محاسبه می‌شود و وزن‌ها و پارامترها با توجه به گرادیان بروز می‌شوند (Sezavar & Fersi, 2015: 3).



شکل ۱- شمای کلی معماری یک شبکه عصبی کانولوشن و لایه‌های آن (نویسندگان، ۱۳۹۹)

۱ Convolutional Neural Networks
۲ Restricted Boltzmann Machines
۳ Autoencoders
۴ Sparse Coding
۵ Pooling
۶ Fully Connected Layer
۷ Feed forward
۸ Backpropagation

در مرحله اول تصویر ورودی به شبکه تغذیه می‌شود و در این مرحله ضرب نقطه ای بین ورودی و پارامترهای هر نورون و نهایتا اعمال عملیات کانولوشن در هر لایه می‌باشد. سپس خروجی شبکه محاسبه می‌شود. در این جا به منظور تنظیم پارامترهای شبکه و یا به عبارت دیگر همان آموزش شبکه، از نتیجه خروجی جهت محاسبه میزان خطای شبکه استفاده می‌شود. به این منظور خروجی شبکه را با استفاده از یک تابع خطا[۱] با پاسخ صحیح مقایسه کرده و این طور میزان خطا محاسبه می‌گردد. در مرحله بعدی بر اساس میزان خطای محاسبه شده مرحله پس انتشار آغاز می‌گردد. در این مرحله گرادیانت هر پارامتر با توجه به قائده زنجیره[۲] محاسبه شده و تمامی پارامترها با توجه به تاثیری که بر خطای ایجاد شده در شبکه دارند تغییر پیدا می‌کنند. بعد از بروز آوری پارامترها مرحله بعدی پیش رونده آغاز می‌گردد. بعد از تکرار تعداد مناسبی از این مراحل آموزش شبکه پایان می‌یابد.

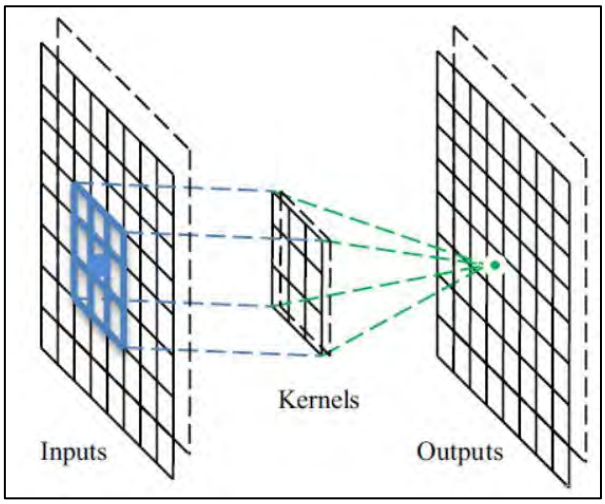
انواع لایه های شبکه CNN

در حالت کلی، یک شبکه عصبی کانولوشن یک شبکه عصبی سلسله مراتبی است که لایه‌های کانولوشنی آن بصورت یک در میان با لایه‌های ادغام بوده و بعد از آن‌ها تعدادی لایه تماما متصل وجود دارد.

لایه کانولوشن: در این لایه ها، شبکه CNN از شالوده‌های مختلف تصاویر[۳] برای شروع پایش تصویر ورودی[۴] و همینطور Feature Mapهای میانی استفاده می‌کند و این‌گونه Feature Map های مختلفی همانند آنچه در شکل۲ نشان داده شده‌است، ایجاد می‌گردد (۱۲). به بیان ساده تر در هر معماری CNN چندین لایه ی کانولوشن وجود دارد، در این لایه‌ها تصویر ورودی با فیلترهایی که ضرایب آنها قابل آموزش است، پایش می‌شود و این فیلترها روی تصویر حرکت داده می‌شوند. عمق فیلتر با عمق عکس برابر است و به ازای هر فیلتر مجزا یک صفحه ویژگی ایجاد می‌گردد. اگر از n فیلتر استفاده شود، n صفحه ویژگی به وجود می‌آید. وزن‌های هر فیلتر که w نامیده می‌شوند قابل آموزش هستند و در طول آموزش شبکه مرتبا بروز می‌شوند. بعد از عمل convolve حاصل با یک عدد بایاس جمع شده و در صفحه ویژگی ذخیره می‌شود. به عنوان مثال اندازه ی فیلترها می‌تواند ۳×۳ باشد. بواسطه فواید معرفی شده توسط عملیات convolution ، بعضی از مقالات تحقیقاتی مشهور از آن جهت جایگزینی لایه های تماما متصل استفاده کردند تا با این کار سرعت فرآیند یادگیری را افزایش دهند (Oquab & Bottou, 2017: 9) یکی از روش‌های جالب مدیریت لایه‌های کانولوشنی، روش NIN[۵] (Lin & Chen, 2013: 11) است که در آن ایده اصلی جایگزینی لایه کانولوشنی با یک شبکه عصبی پرسپترون کوچک است که شامل چندین لایه تماما متصل با توابع فعال سازی غیرخطی است. به این ترتیب فیلترهای خطی با شبکه‌های عصبی غیرخطی جایگزین می‌شوند. این روش باعث بدست آوردن نتایج خوبی در دسته بندی تصاویر می‌شود (Sezavar & Fersi, 2015: 3).

---

۱ Loss Function
۲ Chain Rule
۳ Kernel
۴ Convolve
۵ Network in Network



شکل۲- نمایش عمل کانولوشن بین تصویر ورودی و فیلتر ۳×۳ (Zeiler & 2013: 6)

انجام عملیات convolution سه فایده دارد:

۱) مکانیزم اشتراک وزن در هر Feature Map باعث کاهش شدید تعداد پارامترها می‌شود.

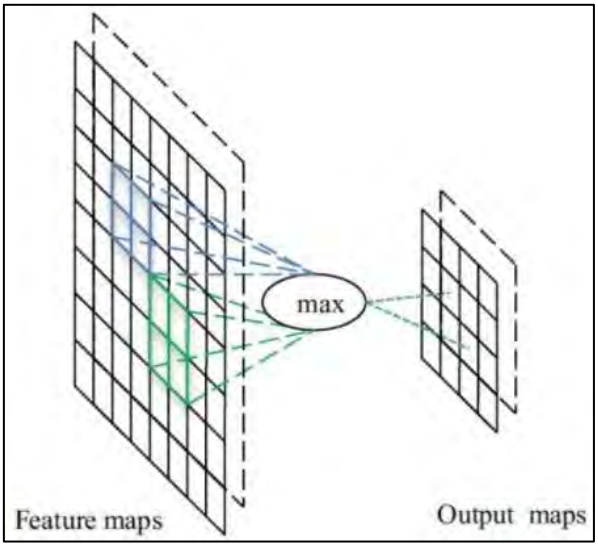
۲) اتصال محلی, ارتباط بین پیکسل های همسایه را یاد می‌گیرد.

۳) باعث تغییرناپذیری و ثبات نسبت به تغییر مکان شئ می‌گردد.

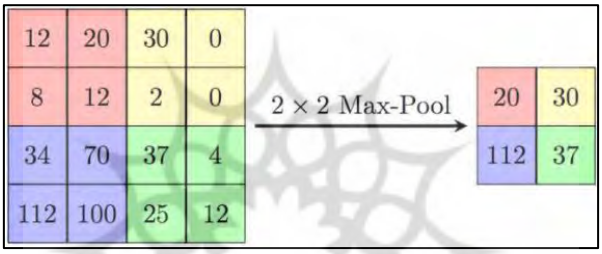
لایه‌های ادغام: بعد از این لایه معمولا لایه ادغام می‌آید. برای ادغام کردن چندین استراتژی وجود دارد که روش معمول آن استفاده از ادغام ماکزیمم است. در این روش باکس هایی معمولا با اندازه ی ۴×۴ انتخاب می‌شوند و روی تصویر حرکت داده می‌شوند و در هر چهار پیکسل، مقدار ماکزیمم آنها انتخاب شده و به لایه بعدی انتقال داده می‌شود. با این کار تعداد صفحات ویژگی ثابت مانده اما اندازه آن کاهش پیدا کرده است. در شکل۳ یک لایه ادغام را مشاهده می‌کنید که بر روی صفحه ی ۸×۸ اعمال شده و اندازه فیلتر ادغام ۲×۲ است. بنابراین ابعاد صفحه به ۴×۴ کاهش پیدا کرده‌است. معروف‌ترین نوع ادغام تابع ادغام حداکثر[۱] است که از خروجی لایه کانولوشن ماکزیمم می‌گیرد و در شکل۵ نحوه عملکرد لایه ادغام حداکثر نشان داده شده است (Zeiler & 2013: 6).

---

۱ Maxpooling

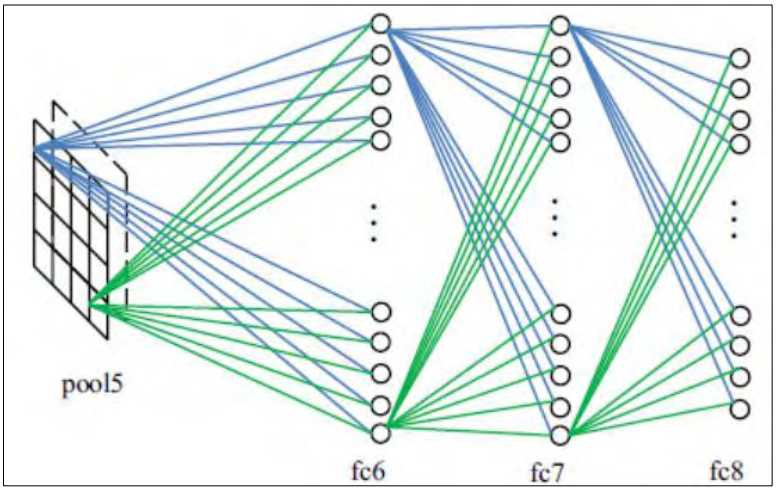


شکل۴- یک لایه ادغام ماکزیمم و کاهش ابعاد بردار ویژگی (Zeiler & 2013: 7)



شکل۵- عملیات ادغام ماکزیمم (Zeiler & 2013: 8)

لایه تماماً متصل: بعد از چندین لایه کانولوشن و ادغام، در نهایت ابعاد صفحات ویژگی بسیار کوچک شده اما تعداد آن‌ها زیاد می‌شود. بنابراین در آخر از لایه‌های تماما متصل استفاده می‌گردد. این لایه‌ها صفحه ویژگی دو بعدی را به بردار یک بعدی تبدیل می‌کنند. لایه‌های تماما متصل مانند یک شبکه عصبی قدیمی عمل می‌کنند و شامل ۹۰ درصد از پارامترهای CNN هستند (Zeiler & 2013: 6). لایه تماما متصل به ما اجازه می‌دهد تا نتیجه شبکه در قالب یک بردار با اندازه مشخص ارائه گردد. می‌توان از این بردار برای دسته‌بندی تصاویر (Zeiler & 2013: 6) استفاده نمود و یا اینکه جهت ادامه پردازش‌های بعدی مورد استفاده قرار داد (Girshick, 2014: 11). در شکل۶ لایه ی تماما متصل را مشاهده می‌کنید که بعد از لایه ادغام پنجم آمده‌است.



شکل۶- نمایش لایه های تماما متصل بعد از آخرین لایه ی ادغام (Zeiler & 2013: 8)

مواد و روش تحقیق:

در این پژوهش از تصاویر با رزولوشن بالا RGB جهت آشکارسازی کشتی‌های موجود استفاده شده است، که از ماهواره PlanetScope تهیه شده‌اند و همچنین جهت ایجاد لایه‌های شبکه CNN، یادگیری عمیق داده‌های ورودی و ارزیابی دقت آشکارسازی اهداف مورد نظر به عنوان داده خروجی از زبان برنامه نویسی پایتون[1] در محیط برنامه نویسی نرم افزار JetBrains PyCharm Pro 2020.3.2 استفاده شده‌است. JetBrains PyCharm یک IDE برنامه نویسی زبان پایتون بسیار قدرتمند است که با ابزار فوق حرفه ای خود مدیریت پروژه ها و برنامه نویسی به زبان پایتون را آسان و سریع می کند. زبان پایتون که یک زبان شی‌گرا و کاربردی است و خود یک مفسر[2] است توسط Guido van Rossum هلندی در سال ۱۹۹۱ ابداع شد. نرم افزار PyCharm باعث کدنویسی سریع تر، دقت بالاتر و برنامه نویسی هوشمند به زبان پایتون می شود.

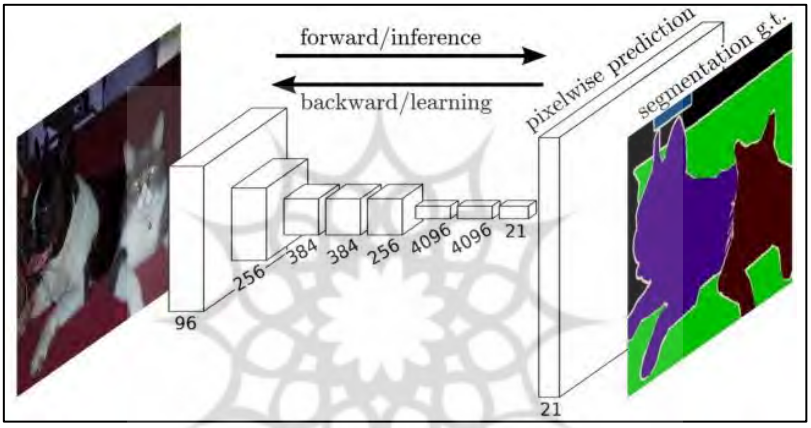
شبکه‌های عصبی کانولوشن، یکی از مهم ترین روش های یادگیری عمیق هستند که در آنها چندین لایه با روشی قدرتمند آموزش می‌بینند که پیشتر به معرفی آن پرداخته شد. این روش بسیار کارآمد، یکی از رایج‌ترین روش‌ها در کاربردهای مختلف بینایی ماشین است. شبکه های عصبی کانولوشنی دارای دو روش مشتق شده R-CNN و FCN است. R-CNN در پژوهش‌های مربوط به شناسایی عوارض، کاربرد دارد. روش FCN بیشتر در تحقیقات قطعه بندی معنایی تصاویر، لبه یابی و غیره مورد استفاده قرار می‌گیرد و به معماری کلی برای الگوریتم‌های نوین جهت شناسایی اشیاء ، تبدیل شده‌است (Farhadi & Kiani, 2019: 2).

شبکه تمام کانولوشنال (FCN)

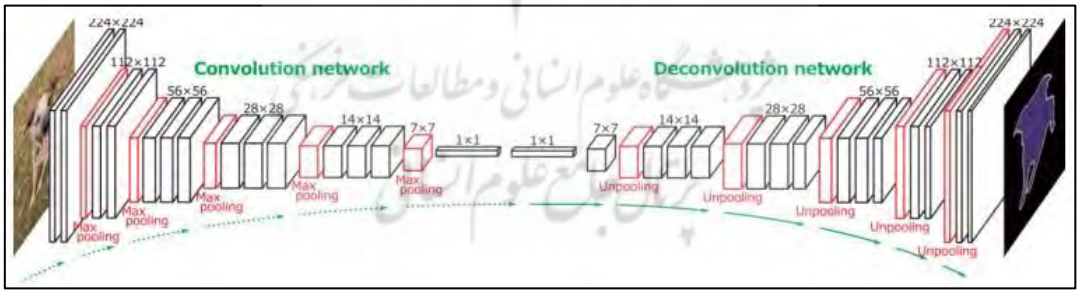
شبکه تمام کانولوشنال شبکه عصبی‌ای است که از لایه‌های کانولوشنی تشکیل شده که هیچ لایه تماماً متصل (FC) در انتهای شبکه وجود ندارد. یک شبکه CNN با لایه‌های تمام متصل به عنوان یک شبکه پایان به پایان قابل یادگیری با یک تمام کانولوشنی است. تفاوت اصلی شبکه تمام کانولوشنی در فیلترهای یادگیری شده در همه جای شبکه است. حتی لایه‌های تصمیم‌گیری در انتهای شبکه یک فیلتر قابل یادگیری هستند. شبکه تمام کانولوشنی تلاش می‌کند تا تصمیم‌گیری‌ها و بازنمایی‌ها را بر اساس ورودی‌های محلی مکانی یاد بگیرد. در واقع شبکه FCN مشتقی از CNN است که در اصل برای قطعه‌بندی

۱ Compiler
۲ Python

معنایی[۱] اولین بار به کار برده شد (Long & Shelhamer, 2015: 8). شبکه FCN روشی برای از نو طرح کردن مدل‌های CNNبه‌صورت شبکه‌های تمام کانولوشنی ارائه می‌کند و قادر به تولید خروجی‌هایی با اندازه متناسب به صورت بهینه است. گرچه FCN عمدتاً برای قطعه‌بندی معنایی ارائه شده‌است اما این تکنیک را می‌توان در کاربردهای دیگر نظیر دسته‌بندی تصاویر و لبه‌یابی نیز مورد استفاده قرار داد. به بیان دیگر می‌توان گفت که شبکه FCN بر خلاف شبکه CNN لایه‌های تمام متصل را در انتهای خود ندارد و می‌تواند خروجی مکانی را به کاربر بدهد. همچنین به همین دلیل، این شبکه قادر خواهد بود تا انواع تصاویر ورودی را بدون محدودیت ابعاد تصویر دریافت و پردازش کند. همچنین این شبکه‌ها قابلیت کانولوشون و دیکانولوشن را به ارمغان آورده‌اند که کاربردهای بی نظیری برای همانند بازسازی تصویر[۲]، حذف نویز[۳]، ترمیم تصویر[۴] و افزایش کیفیت تصویر[۵] و غیره دارد. در ادامه چندین شبکه معروف از FCN برای کاربردهای مختلف در شکل‌های ۶ و ۷ به نمایش گذاشته شده‌است. همچنین اخیرا شبکه‌های FCN دستاوردهای بسیار بزرگی در تشخیص اشیا در تصویر و ویدئو و تقسیم‌بندی نمونه‌ای[۶] بدست آورده است (Naji, 2018: 33)



شکل۶- نمونه‌ای از شبکه FCN برای قطعه‌بندی معنایی (Long & Shelhamer, 2015: 7)



شکل۷- نمونه‌ای از شبکه FCN به صورت لایه‌های کانولوشنی و دیکانولوشنی برای قطعه‌بندی معنایی (Noh & Hong, 2015: 3)

۱ Semantic Segmentation
۲ Noise cancelation
۳ Image repairing
۴ Inpainting
۵ Image Super-Resolution
۶ Instance Segmentation

پس از معرفی چند نمونه از الگوریتم‌های کاربردی یادگیری عمیق در بخش پیدا کردن اهداف زمینی و آنالیز تصاویر ماهواره‌ای با رویکرد سنجش از دور در پژوهشی جامع که در سال ۲۰۱۷ انجام شد، گام بلندی در زمینه بهبود دقت و سرعت روش‌های یادگیری عمیق در این زمینه برداشته شد و بسیاری از چالش‌های پژوهشگران که ناشی از جدید بودن این روش در زمینه پیداکردن اهداف مورد نظر در چنین تصاویری برطرف گردید (Zhu & Tuia, 2017: 13). در سال بعد از این پژوهش، مقالات و پژوهش‌های بسیاری در این زمینه انجام گردید که در یکی از مهم‌ترین پژوهش‌هایی که بر روی روش‌های مختلف الگوریتم‌های شبکه‌های عصبی کانولوشنی (RCN، Fast R-CNN، Faster R-CNN، FCN) انجام شد، منجر به ارایه یکی از بهترین روش‌های یادگیری عمیق در نوع خاصی از اهداف مورد نیاز انجامید. در این پژوهش که بر روی اهداف زمینی به شکل‌های عمومی یکسان همچون وسایل نقلیه عمومی و هلیکوپتر انجام شد به این نتیجه دست یافت که بهترین و سریع‌ترین روش برای چنین اهدافی استفاده از الگوریتم FCN می‌باشد (Cheng & Yang, 2018, 2).

محدوده مورد مطالعه:

یوکوهاما[۱] یکی از شهرهای ژاپن و مرکز استان کاناگاوا[۲] می‌باشد که در جنوب شرقی این کشور واقع شده است. این شهر با جمعیت ۳/۶ میلیون نفری خود پرجمعیت‌ترین شهر ژاپن پس از توکیو می‌باشد. این شهر پس از آن که در سال ۱۸۵۹ به عنوان یک بندر تجاری در نظر گرفته شده‌است، پس از آن به سرعت پیشرفت کرده‌است. این منطقه با توجه به تعدد عبور انواع کشتی‌ها از آن و همچنین در دسترس بودن نقشه‌ها و تصاویر آن انتخاب شده است که در شکل۸ مورد مشاهده می‌باشد.

شکل۸– موقعیت جغرافیایی منطقه مورد مطالعه (نویسندگان، ۱۳۹۹)

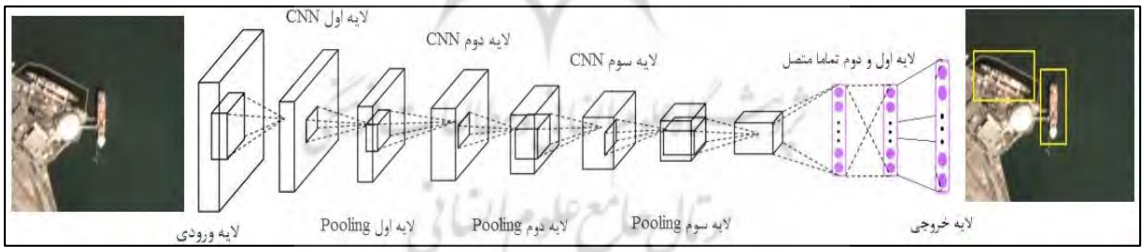
یافته‌ها:

۱ Yokohama
۲ Kanagawa

هدف از استفاده از تصاویر ماهواره‌ای در زمینه پیدا کردن محل دقیق کشتی‌ها در صورت امکان (موجود بودن تصاویر با کیفیت از منطقه مورد نظر) و همچنین ساده‌سازی پیدا کردن محل تقریبی و یا مسیر حرکت کشتی‌ها (در صورت عدم موجود چنین تصاویر با رزولوشن مناسب) جهت فائق آمدن بر مشکلات مختلفی که در ادامه به آن‌ها اشاره شده، می‌باشد. خودکارسازی امر مکانیابی کشتی‌ها برای بسیاری از اهداف در زمینه‌های مختلف می‌تواند مورد بهره‌وری قرار گیرد که از مهم‌ترین آن‌ها می‌توان به موارد زیر اشاره نمود. پایش و نظارت بر کشتی‌های تجاری عبوری از آبراهه‌های بین‌المللی و محلی با رویکردهای متفاوت اعم از کنترل واردات و صادرات کشور، مکانیابی دقیق و شناسایی کشتی‌های جنگی غیرخودی در زمینه طرح‌ریزی برنامه دفاعی با رویکرد پدافند عامل و غیرعامل، پایش و کنترل کشتی‌های حامل مواد نفتی در زمینه نظارت بر قاچاق سوخت و فراورده‌های نفتی، نظارت بر صید آبزیان و مقابله با صید بی‌رویه و غیرمجاز که مهم‌ترین عامل نابودی اکوسیستم‌های محیط‌های آبی می‌باشد. تصاویر ماهواره‌ای استفاده شده در این پژوهش تصاویری از خلیج یوکوهاما در ماه ژانویه سال ۲۰۲۰ میلادی[1] می‌باشد که در آن تنوع نوع کشتی‌ها و همچنین تنوع اندازه و ابعاد کشتی‌ها در نظر گرفته شده‌است و از ماهواره تصویربرداری PlanetScope در قالب ۴۰۰۰ تصویر مرجع شده RGB به ابعاد ۸۰×۸۰ پیکسل با دقت ۳ متر برای هر پیکسل جهت تهیه تصاویر کلاس استفاده شده‌است که در نهایت به پیکسل‌های «دارای کشتی» و یا «عدم وجود کشتی» طبقه‌بندی می‌شوند. هر تصویر ۸۰×۸۰ پیکسل استفاده شده حاوی ۱۹۲۰۰ کاراکتر عددی بوده که ۶۴۰۰ کاراکتر ابتدایی آن تشکیل دهنده رنگ قرمز، ۶۴۰۰ کاراکتر بعدی تشکیل دهنده رنگ سبز و ۶۴۰۰ کاراکتر پایانی، تشکیل دهنده رنگ آبی را شامل می‌شود.

از تعداد ۴۰۰۰ تصویر کلاس موجود، تعداد ۲۸۰۰ تصویر (معادل ۷۰٪ کل داده‌ها) جهت آموزش[2] شبکه و تعداد ۱۲۰۰ تصویر باقی مانده (معادل ۳۰٪ کل داده‌ها) جهت ارزیابی[3] شبکه استفاده شده است.

لایه‌های شبکه عصبی به کار برده شده، به سه لایه اصلی تقسیم می‌شود که هر لایه شامل زیرلایه‌های متفاوتی جهت ورود اطلاعات، پردازش شبکه و داده‌ها و خروجی مورد نظر می‌باشد: سه لایه CNN، سه لایه Pooling و دو لایه تماماً متصل. لایه‌های موجود در شکل۹ نشان داده شده است.



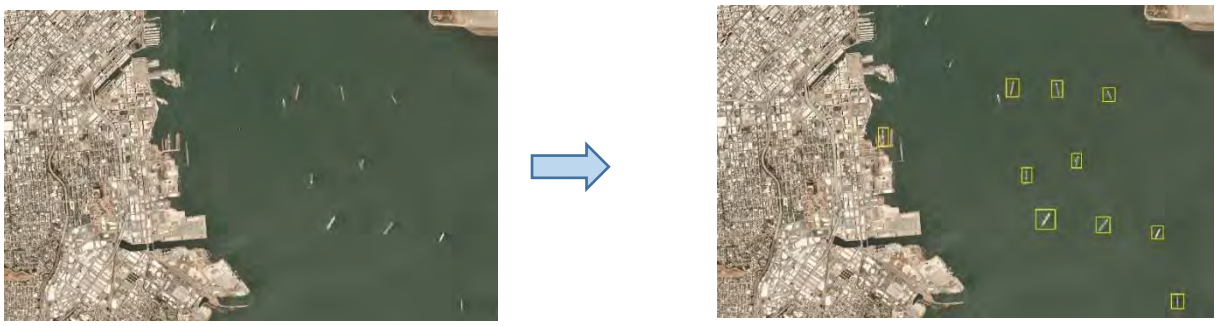
شکل۹- لایه‌های موجود در شبکه عصبی CNN به کار برده شده در فرآیند شناسایی کشتی‌ها (نویسندگان، ۱۳۹۹)

در ابتدا لایه‌های شبکه عصبی CNN با استفاده از الگوریتم شبکه تماماً کانولوشنال (FCN) در محیط برنامه‌نویسی پایتون Python ایجاد شده و شبکه با استفاده از داده‌های آموزشی در نظر گرفته شده، مورد آموزش قرار می‌گیرد، در ادامه با ورود تصاویر مورد نظر جهت آشکارسازی محل کشتی‌های موجود، کشتی‌های استخراج شده در تصویر توسط داده‌های ارزیابی در نظر گرفته شده مورد ارزیابی قرار می‌گیرد. در ایجاد لایه‌های شبکه CNN در الگوریتم پیشنهادی از کتابخانه‌های pytorch، tflearn، matplotlib، numpy، tensorflow-gpu، keras مربوط به یادگیری عمیق در زبان برنامه‌نویسی پایتون استفاده شده‌است. در شکل۱۰ آشکارسازی کشتی‌های موجود در تصویر مورد نظر، و در شکل۱۱ کشتی‌های یافته‌شده توسط الگوریتم

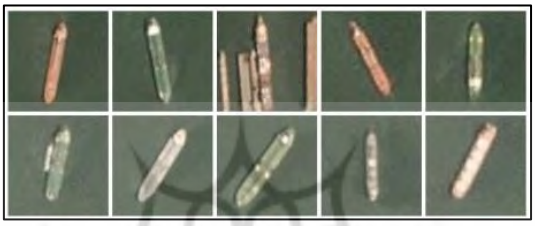
---

۱ January 2020
۲ Train
۳ Evaluation

مورد استفاده نشان داده شده‌است. سیستم مورد استفاده برای آموزش و ارزیابی شبکه دارای کارت گرافیکی Geforce GTX 1080 Ti با حافظه ۸ گیگابایتی و پردازنده Intel Core i7 – 9700K بوده‌است.

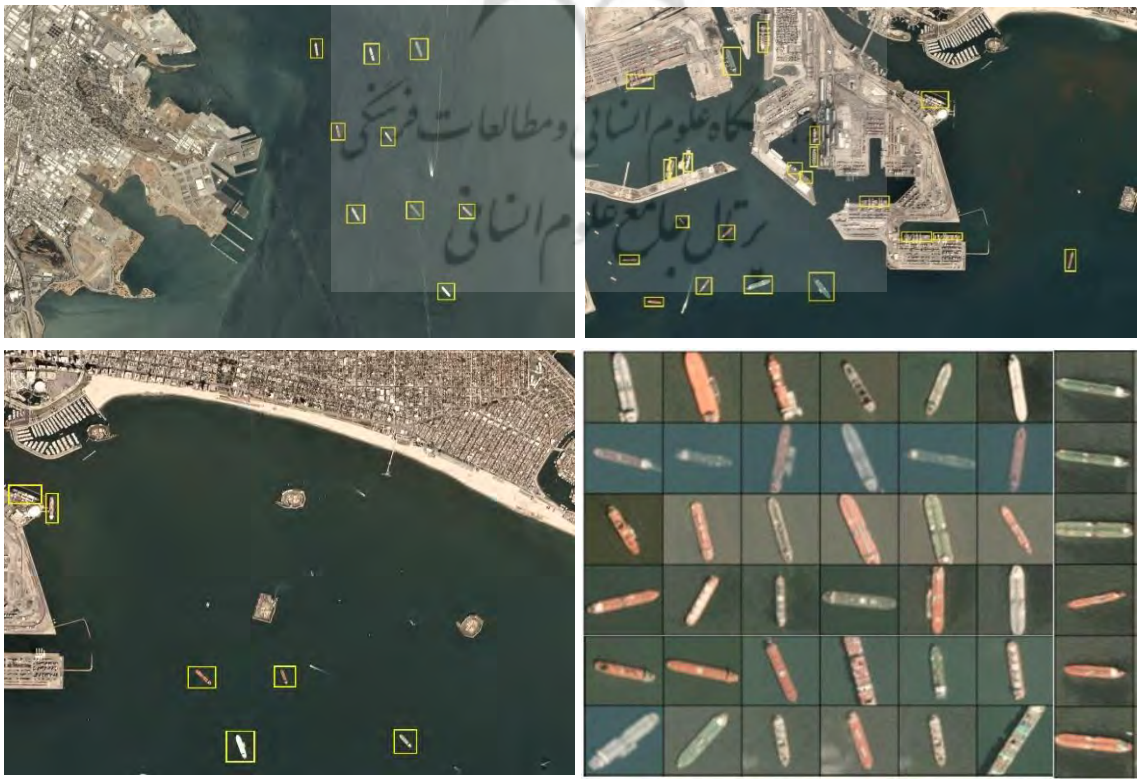


شکل ۱۰– آشکارسازی کشتی‌های موجود در تصویر از منطقه مطالعاتی (نویسندگان، ۱۳۹۹)



شکل ۱۱– کشتی‌های یافته‌شده توسط الگوریتم FCN مورد استفاده (نویسندگان، ۱۳۹۹)

جهت صحت‌سنجی و ارزیابی دقت الگوریتم مورد استفاده و آموزش یافته به وسیله تصاویر آموزشی در نظر گرفته شده در آشکارسازی کشتی‌های موجود، از سه تصویر اخذ شده دیگر از منطقه مطالعاتی ذکر شده و همچنین کشتی‌های آشکارسازی شده در این تصاویر در شکل۱۲ نشان قابل مشاهده می‌باشد.



شکل۱۲- شناسایی کشتی‌های موجود با بهره‌گیری از روش پیشنهادی (نویسندگان، ۱۳۹۹)

برای ارزیابی عملکرد الگوریتم استفاده شده در این پژوهش از حساسیت و تشخیص‌پذیری که دو شاخص مهم برای ارزیابی آماری عملکرد نتایج آزمون‌های طبقه‌بندی باینری (دودویی یا دوحالته) هستند، استفاده شده‌است. زمانی که بتوان داده‌ها را به دو دسته مثبت «دارای کشتی» و منفی «عدم وجود کشتی» تقسیم کرد، عملکرد الگوریتم موردنظر که اطلاعات را به این دو دسته تقسیم می‌کند با استفاده از شاخص‌های حساسیت و تشخیص‌پذیری قابل اندازه‌گیری و توصیف است. پارامتر تشخیص‌پذیری را اصطلاحاً دقت[۱] حساسیت را نیز اصطلاحا صحت[۲] می‌نامند. دسته بندی داده های حاصل از نتایج:

۱- مثبت صحیح (TP[۳]): هدف موردنظر درست شناسایی شده است.

۲- مثبت کاذب (FP[۴]): هدف موردنظر اشتباه شناسایی شده است.

۳- منفی صحیح(TN[۵]): هدف موردنظر به درستی رد شده است.

۴- منفی کاذب (FN[۶]): هدف موردنظر به اشتباه رد شده است.

Predicted \ Actual	(Positives)	(Negatives)
(Positives)	TP	FP
(Negatives)	FN	TN

صحت (Precision)

وقتی که مدل نتیجه را مثبت پیشبینی می‌کند، این نتیجه تا چه اندازه درست است؟ زمانی که ارزش FP بالا باشد، معیار صحت، معیار مناسبی خواهد بود. فرض کنید، مدلی برای تشخیص کشتی‌ها موجود باشد و این مدل Precision پایینی داشته باشد، نتیجه این امر این است که این مدل، بسیاری از کشتی‌ها را به اشتباه شناسایی می‌نماید. در واقع صحت، معیاری است که مشخص می‌کند دسته‌بند، به چه اندازه در شناسایی تمام کشتی‌های موجود موفق بوده‌است. این معیار در واقع نسبت مقداری موارد صحیح طبقه‌بندی شده توسط الگوریتم از یک کلاس مشخص، به کل تعداد مواردی که الگوریتم چه به صورت صحیح و چه به صورت غلط، در آن کلاس طبقه‌بندی کرده است که از رابطهٔ (۱) محاسبه می‌شود:

$$Precision = \frac{TP}{TP + FP}$$

رابطهٔ (۱)

پوشش یا حساسیت (Recall):

در نقطه مقابل این پارامتر، ممکن است در مواقعی دقت تشخیص کلاس منفی حائز اهمیت باشد. این معیار به معنی نسبتی از موارد منفی است که آزمایش آنها را به درستی به عنوان نمونه منفی تشخیص داده‌است. این پارامتر به صورت زیر محاسبه می‌شود. زمانی که ارزش FN بالا باشد، این معیار مناسب خواهد بود. فرض شود مدلی برای شناسایی محل دقیق کشتی‌های جنگی موجود در یک منطقه حساس با رویکرد تشخیص کشتی‌های غیرخودی طراحی شده باشد. اگر این مدل معیار Recall پایینی داشته باشد چه اتفاقی خواهد افتاد؟ این مدل کشتی‌های زیادی که غیرخودی هستند را خودی در نظر می‌گیرد و این موضوع بروز یک فاجعه بزرگ در امر دفاعی است. نسبت موارد صحیح طبقه‌بندی شده توسط الگوریتم از یک کلاس به تعداد موارد موجود در کلاس مذکور می‌باشد که از رابطهٔ (۲) زیر محاسبه می‌گردد:

---

۱ Precision
۲ Recall
۳ True Positive
۴ False Positive
۵ True Negative
۶ False Negative

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

معیار ارزیابی F1-Score یا F-measure :

این معیار، معیاری مناسب برای ارزیابی دقت یک الگوریتم تشخیص است. این معیار Precision و Recall را با هم در نظر می‌گیرد و در بهترین حالت، یک و در بدترین حالت صفر می‌باشد:

رابطهٔ (۳)
$$F1-measure = 2\left(\frac{Recall \times Precision}{Recall + Precision}\right)$$

مقادیر این معیارها برای الگوریتم استفاده شده در این پژوهش به شرح جدول۱ زیر می‌باشد:

جدول۱- نتیجه ارزیابی دقت الگوریتم پیشنهادی در آشکارسازی کشتی‌های موجود در تصاویر مورد نظر

F1-measure	Recall	Precision
۹۸/۸۳	۹۷/۶۱	۱۰۰

منبع: مطالعات میدانی نویسندگان، ۱۳۹۹

نتیجه‌گیری و ارائه پیشنهادات:

همان‌طور که در ابتدای این پژوهش به تحقیق‌های پیشین در زمینه شناسایی اهداف موجود در تصاویر ماهواره‌ای اشاره شد، یکی از بهترین و دقیق‌ترین روش‌ها در این زمینه روش شبکه‌های عصبی کانولوشن (CNN) می‌باشد. در روش ارائه شده در این پژوهش در زمینه شناسایی کشتی‌های موجود در تصاویر می‌توان با توجه به ابعاد آن‌ها، کشتی‌های آشکارشده توسط الگوریتم را به دو دسته «کشتی‌های نفتکش، باربری و نظامی» و «کشتی‌های صیادی، تفریحی و ...» تقسیم نمود و همچنین با توجه به روش پیشنهاد شده در قسمت پیشنهادات به شناسایی نوع خاص کشتی مورد نظر با توجه به درجه اهمیت آن نیز پرداخت. با بهره‌گیری از روش تماماً کانولوشن ارائه شده در این پژوهش (FCN) با استناد به نتایج شناسایی کشتی‌های موجود در تصاویر، معیارهای ارزیابی Precision ، Recall و F1-measure ، نشان‌دهنده دقیق بودن و قابل اعتماد بودن این روش در شناسایی چنین اهدافی می‌باشد. لازم به ذکر است دستیابی به نتایجی با چنین دقت تنها در صورت آموزش شبکه ایجاد شده با استفاده از داده‌های آموزشی کافی امکان‌پذیر می‌باشد.

پیشنهاد می‌گردد در پژوهش‌های آتی در این زمینه از الگوریتم‌های دیگر همچون RCN ، Fast-RCN ، ... و یا با استفاده از انتقال آموزشی الگوریتم (Transfer Learning) و همچنین با استفاده از شبکه‌های از پیش آموزش دیده‌ای همچون CIFAR10Net، GoogLeNet، VGG و ... به آشکارسازی اهداف مورد نظر پرداخته شود. در بخش تصاویر ورودی نیز می‌توان با بهره‌گیری از تصاویر چندطیفی یا ابرطیفی، به تشخیص نوع دقیق کشتی شناسایی شده با توجه به تفاوت در طیف بازگشتی از عرشه کشتی بر اساس آلیاژ و مصالح به کار برده شده به موجب کاربری کشتی‌های نفتکش، نظامی و ... پرداخته شود.

منابع و مآخذ:

1. Cheng, G., Yang, C., Yao, X., Guo, L. and Han, J., 2018. When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs. IEEE transactions on geoscience and remote sensing, 56(5), pp.2811-2821.
2. Corbane, C., Najman, L., Pecoul, E., Demagistri, L. and Petit, M., 2010. A complete processing chain for ship detection using optical satellite imagery. International Journal of Remote Sensing, 31(22), pp.5837-5854.
3. Farhadi N, Kiani, and Ebadi., 2019, May. Target detection from high-resolution remote sensing images using deep learning methods. Iranian Remote Sensing & GIS, Vol.11, No. 1, 33-48.
4. Girshick, R., Donahue, J., Darrell, T. and Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587).
5. https://deeplearning.ir/ ,2020 December.
6. Keykhaee, Rangzan, Taghizade., 2016, April. Potential finding of surveillance sites of combat units using GIS and fuzzy logic in Nasrabad area of Isfahan. Journal of Military Science and Technology, pp 109-128.
7. Khan, M.J., Yousaf, A., Javed, N., Nadeem, S. and Khurshid, K., 2017. Automatic target detection in satellite images using deep learning. J. Space Technol, 7(1), pp.44-49.
8. Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2017. Imagenet classification with deep convolutional neural networks. Communications of the ACM, 60(6), pp.84-90.
9. LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., 1998. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), pp.2278-2324.
10. Lin, M., Chen, Q. and Yan, S., 2013. Network in network. arXiv preprint arXiv:1312.4400.
11. Long, J., Shelhamer, E. and Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3431-3440).
12. Lotfollahi, M,. Shirali, R,. Jafari, M,. 2015, June. Application of Deep Learning in Computer Network Traffic Detection. 22nd Annual National Conference of Iranian Computer Association.
13. Naji, M,. 2018 December,. Image Fusion using Deep Learning, Master of degree thesis, Supervisor: Dr. Ali Aghagolzadeh, Babol Noshirvani University of Technology.
14. Napiorkowska, M., Petit, D. and Marti, P., 2018, July. Three applications of deep learning algorithms for object detection in satellite imagery. In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium (pp. 4839-4842). IEEE.
15. Noh, H., Hong, S. and Han, B., 2015. Learning deconvolution network for semantic segmentation. In Proceedings of the IEEE international conference on computer vision (pp. 1520-1528).
16. Oquab, M., Bottou, L., Laptev, I. and Sivic, J., 2015. Is object localization for free?-weakly-supervised learning with convolutional neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 685-694).
17. Oskooee M,. 2015, Hyperspectral telemetry, benefits and requirements of development and its future in Iran. Iranian Remote Sensing & GIS, Vol.6, No. 1, 22-35.
18. Sezavar, A., Fersi, H,. Mohamadzadeh, S,. 2015, June. Image retrieval using deep learning. 4th International Conference on Applied Research in Computer Engineering and Signal Processing. Vol.3, No. 1, 28-42.

19. Stofa, M.M., Zulkifley, M.A. and Zaki, S.Z.M., 2020, July. A deep learning approach to ship detection using satellite imagery. In IOP Conference Series: Earth and Environmental Science (Vol. 540, No. 1, p. 012049). IOP Publishing.
20. Zeiler, M.D., 2013. Hierarchical convolutional deep learning in computer vision (Doctoral dissertation, New York University).
21. Zhu, X.X., Tuia, D., Mou, L., Xia, G.S., Zhang, L., Xu, F. and Fraundorfer, F., 2017. Deep learning in remote sensing: A comprehensive review and list of resources. IEEE Geoscience and Remote Sensing Magazine, 5(4), pp.8-36.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

Provide a method for targets detection in satellite imagery using deep learning with remote sensing and GIS approach

Abstract

Automatic detection of features in different areas according to various objectives, including urban management, military objectives, etc., are one of the most up-to-date and important applications of machine learning today. Combining the Global Geographic System (GIS) with images taken from satellite sensors and finally using deep learning methods, which is one of the main branches of machine learning, is a great help to the visible subject. Made the effects in the images using remote sensing science.. At the beginning of this research, the various layers in the proposed algorithm have been comprehensively presented and introduced, and then a new method has been presented in the simultaneous combination of CNN and pooling layers in the algorithm used, which finally It led to a significant reduction in network training time using comprehensive training data with high accuracy and at the same time high volume, which in the end, after entering the fully connected layer to extract and identify the desired goals with acceptable accuracy along with cost-effectiveness. Save time. In this research, using network training through training data, ships in satellite images are detected by creating a fully convoluted FCN network. In order to evaluate the performance and accuracy of the algorithm used in finding and detecting ships in satellite images, by applying this detection algorithm on several other satellite images, Precision, Recall and F1-Score evaluation criteria were used. The values were equal to 100%, 97.61% and 98.83%, respectively, which indicates the accuracy and reliability of the algorithm.

Keywords: Remote Sensing, Deep Learning, Convolution Neural Network, Automatic Target De

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی