

مروری بر روش‌های آشکارسازی اهداف در تصاویر پنکروماتیک سنجش از دوری با استفاده از الگوریتم‌های یادگیری عمیق

محمد ابراهیم عقیلی^۱، امیر آقابالائی^{۲*}

^۱ دانشجوی دکتری مخابرات سیستم - دانشکده مهندسی برق و کامپیوتر - دانشگاه تربیت مدرس
m.aghili@modares.ac.ir

^۲ استادیار - دانشگاه جامع امام حسین (ع)
remotesensing@ihu.ac.ir

(دریافت: بهمن ۱۴۰۲، تصویب: تیر ۱۴۰۳)

چکیده

طی سال‌های اخیر پیشرفت‌های سریعی در استفاده از یادگیری عمیق برای آشکارسازی اهداف در تصاویر پنکروماتیک سنجش از دوری با وضوح بالا انجام شده است. در این مقاله تحقیقات جاری این زمینه مرور شده و یافته‌های نوآورانه روش‌های موجود، بیان شده است. همچنین طیف وسیعی از آشکارسازی‌های مورد نظر شامل ساختمان، وسایل نقلیه، هواپیما و کشتی پوشش داده شده است. شبکه‌های عصبی عمیق بر خلاف روش‌های سنتی که با تکیه بر ویژگی‌های دست ساز عمل می‌کنند، یادگیری ویژگی سطح بالا و قابلیت‌های تمایز هدف را ارائه می‌دهند. آموزش سراسری، شبکه‌ها را قادر ساخته تا بازنمایی‌های مهمی را به طور مستقیم از تصاویر برای تشخیص دقیق بیاموزند. از جمله معماری‌های مهم و مورد بررسی می‌توان به Faster R-CNN، SSD و YOLO اشاره کرد که اغلب برای آشکارسازی اهداف، سفارشی‌سازی می‌شوند. علاوه بر استفاده مستقیم از تصاویر پنکروماتیک، از آن برای بهبود عملکرد تشخیص توسط ادغام با تصاویر دیگر مانند چندطیفی و مادون قرمز استفاده می‌شود. بهره‌بری از اطلاعات مکانی در تصاویر پنکروماتیک، صحت آشکارسازی اهداف را افزایش داده است. سازوکارهای توجه، اتصالات متراکم و ترکیب ویژگی‌های چندمقیاسی برای تقویت یادگیری ویژگی و بهبود نتایج در صحنه‌های پیچیده بررسی شده است. راهبردهای داده‌افزایی، یادگیری انتقالی، کاوش نمونه کم‌کیفیت و سازگاری مدل برای مدیریت تصاویر برچسب‌دار محدود و فرآیند آموزش شبکه استفاده شده‌اند. به طور کلی، روش‌های یادگیری عمیق پیشرفته به صحت‌های بالایی در آشکارسازی انواع مختلف اهداف دست یافته‌اند. اما چالش‌ها در مدیریت اهداف کوچک، متراکم و مبهم، باقی مانده است. یادگیری مستمر شبکه‌های عمیق برای اهداف جدید و تطبیق مدل‌ها، جهت‌گیری روش‌های اخیر می‌باشد. این بررسی بسیاری از روش‌های یادگیری عمیق آشکارسازی را پوشش داده و بینشی را برای هدایت تحقیقات آینده به سمت آشکارسازی‌های عملیاتی و قدرتمند برای تشخیص هدف در تصاویر پنکروماتیک با وضوح بالا ارائه می‌کند.

واژگان کلیدی: تصاویر پنکروماتیک، آشکارسازی هدف، یادگیری عمیق، ویژگی‌های مکانی، افزایش وضوح مکانی.

* نویسنده رابط

۱- مقدمه

با پیشرفت‌های سخت‌افزاری و نرم‌افزاری، تحول بنیادینی در سنجش‌ازدور در حال رخ دادن است. یکی از کاربری‌های مهم در این حوزه، آشکارسازی اهداف است که به سمت استفاده از یادگیری عمیق برای شناسایی خودکار اهداف در تصاویر پنکروماتیک^۱ با وضوح بالا، حرکت می‌کند [۱]. این بررسی مجموعه‌ی گسترده‌ای از مطالعات را در بر گرفته که بر آشکارسازی اهداف متنوع تمرکز دارند. تشخیص اهداف در تصاویر پنکروماتیک، زمینه خاص‌تری را نسبت به سایر تصاویر سنجش‌ازدور مانند RGB یا چندطیفی نشان داده- است [۲]. تصاویر پنکروماتیک معمولاً در شرایط تخصصی که به وضوح مکانی بالایی نیاز است اخذ و تجزیه و تحلیل می‌شوند. با جمع‌آوری انرژی فوتون‌ها در طیف مرئی و بخشی از مادون قرمز نزدیک، تصاویر پنکروماتیک به دقت بالا و محتوای بهینه‌شده اطلاعات مکانی دست یافته و مناسب آشکارسازی و تشخیص هدف می‌شوند [۳]. روند غالب در روش‌های مورد بررسی، ظهور شبکه‌های عصبی کانولوشن CNN^۲ است [۴]. این مدل‌های یادگیری عمیق به طور مداوم از رویکردهای سنتی مانند تبدیل ویژگی مقیاس‌ناسته SIFT^۳ [۵]، هیستوگرام گرادینان‌های جهت‌دار HOG^۴ [۶] و فیلترهای گابور^۵ [۷]، که فاقد قدرت کافی برای تفکیک هدف از پس‌زمینه هستند، عملکرد بهتری را نشان داده‌اند [۸]. همچنین در روش‌های سنتی امکان تعیین خودکار و دقیق پارامترها، مبتنی بر تصویر ورودی وجود ندارد. بهینه‌سازی خودکار شبکه‌های عمیق و تنظیم شبکه ازپیش-آموزش‌دیده^۶ برای تصاویر آموزشی جدید یا حتی تصویر با ماهیت جدید، مزیت‌های بزرگی را نسبت به روش‌های پیشین دارد [۹]. به طور مثال بهبود مدل تولید شده با نمونه‌های آموزشی جدید و یا استفاده از مدل ازپیش‌آموزش‌دیده روی تصاویر نوری و استفاده آن برای انواع دیگر تصاویر مانند پنکروماتیک، از جمله برتری روش‌های آشکارسازی اهداف با استفاده از شبکه‌های عمیق می‌باشد [۱۰].

در تحقیقات این حوزه، پیشرفت‌هایی مانند افزونه‌های توجه^۷ [۱۱]، ادغام ویژگی‌های چندمقیاسه [۱۲] و

راهبردهای یادگیری انتقالی^۸ [۱۳] برای رسیدن به الگوهای منحصربه‌فرد تصاویر ماهواره‌ای با وضوح بالا استفاده شده و مزایای قابل توجهی به دست آمده‌است [۱۴]. علاوه بر این موارد، روش‌های گروهی^۹ [۱۵] تأثیر بسیاری در افزایش دقت آشکارسازی داشته‌است. به این صورت که ادغام پیش‌بینی‌های چندین شبکه متنوع منجر به افزایش قابل توجه دقت و استحکام می‌شود [۱۶]. از جمله راهبردهای دیگر، کاوش در هم‌افزایی بین تصاویر پنکروماتیک و سایر تصاویر مانند چندطیفی و...، پتانسیل بهبود تشخیص را با گنجاندن منابع اطلاعاتی مکمل نشان داده‌است [۱۷].

فراتر از معماری‌های استاندارد شبکه‌های عمیق، محققان به طور فعال به بررسی پردازش‌های تخصصی برای آشکارسازی اهداف سنجش‌ازدوری پرداخته‌اند. این موارد شامل استفاده از pan-sharpening [۱۸] برای افزایش دقت مکانی، استفاده از عملیات ریخت‌شناسی^{۱۰} [۱۹] برای تفکیک نواحی پس‌زمینه و هدف است [۲۰]. همچنین روش‌های تشخیص ناهنجاری برای دسترسی به نواحی محتمل از اهداف به کار گرفته شده‌است [۲۱]. مجموعه تصاویر پنکروماتیک مختلفی از تصاویر هوایی و ماهواره‌ای در سنجنده‌هایی مانند Gaofen [۲۲]، Quickbird [۲۳] و WorldView [۲۴] است. مشخصه‌هایی مانند تنوع سنجنده-ها، دقت مکانی مختلف و گسترش تصاویر آموزشی ناهمگن، مدل‌های یادگیری عمیق با قابلیت تعمیم‌پذیری بالا را برجسته کرده‌اند [۲۵]. به طور کلی، این بررسی جامع پیشرفت‌های مهم در یادگیری عمیق برای آشکارسازی خودکار تصاویر پنکروماتیک را تحلیل کرده‌است. یادگیری مداوم، پیش‌بینی‌های قابل توضیح، برآورد عدم قطعیت و مدل‌های مستقل از سنجنده، مرزهای مهم تحقیقاتی هستند. با افزایش در دسترس بودن تصاویر پنکروماتیک از سکوه‌های نسل بعدی، یادگیری عمیق، مسیر بهره‌برداری کامل از اطلاعات مکانی-زمانی در حوزه‌ی هوش مصنوعی را هموار کرده‌است.

در این تحقیق واژگان کلیدی شامل: “Remote sensing”^{۱۰}، “Panchromatic”^{۱۱} و “Target detection”^{۱۲} در

۶ Pre-trained

۷ Attention module

۸ Transfer Learning

۹ Ensemble

۱۰ Morphology

۱ Panchromatic

۲ Convolutional Neural Network

۳ Scale-Invariant Features Transformation

۴ Histogram of Oriented Gradients

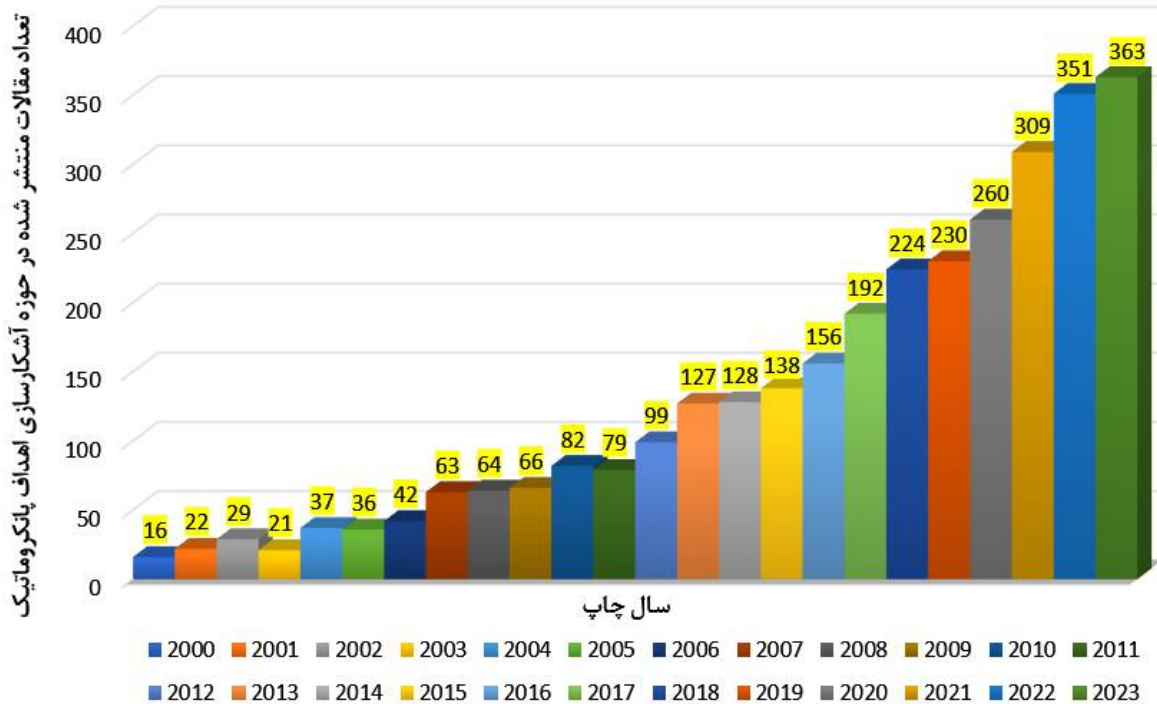
۵ Gabor filter

پایگاه اطلاعاتی گوگل اسکولار جست‌وجو شده‌است. گوگل اسکولار یک پایگاه داده عظیم و جامع از منابع علمی با در برگیری مقالات، کتاب‌ها، مجموعه مقالات کنفرانس، پایان‌نامه‌ها و موارد دیگر است. این پایگاه طیف گسترده‌ای از مقالات آشکارسازی اهداف را پوشش می‌دهد. این امر آن را به منبعی ارزشمند برای پژوهشگران، دانشجویان و متخصصانی تبدیل کرده‌است که به دنبال اطلاعات از منابع مختلف هستند. علاوه بر آن، قابلیت‌های جستجوی پیشرفته‌ای ارائه می‌دهد که به کاربران این امکان را داده که جستجوهای خود را دقیق‌تر کرده و مقالات مرتبط را با کارایی بیشتری پیدا کنند. آمار و ردیابی ارجاعات مقالات نشان می‌دهد که یک مقاله خاص چند بار توسط سایر پژوهشگران استناد شده‌است. این اطلاعات می‌تواند برای ارزیابی تأثیر و نفوذ یک مطالعه یا نویسنده مفید باشد. شکل ۱ تعداد مقالات حوزه آشکارسازی هدف در تصاویر پنکروماتیک را طی سال‌های ۲۰۰۰ تا ۲۰۲۳ نشان می‌دهد.

سپس از بین مقالات چاپ شده، تعداد ۲۹ مقاله، مربوط به آشکارسازی اهداف توسط "Deep learning" انتخاب شده‌اند. مقصود از مرور انجام شده پاسخگویی به چند سوال کلیدی در خصوص آشکارسازی هدف در تصاویر پنکروماتیک مبتنی بر یادگیری عمیق است:

- مقدمات آشکارسازی هدف در تصاویر پنکروماتیک چیست؟
- چه نوع ساختارهایی از شبکه‌های عصبی برای آشکارسازی در این نوع تصویر وجود دارد؟
- ساختارهای موجود از نظر صحت و سرعت چه ویژگی‌هایی دارند؟
- آموزش شبکه عمیق نیاز به چه تمهیداتی دارد؟
- بهره‌بری از اطلاعات مکانی تصاویر پنکروماتیک چه تأثیری بر آشکارسازی اهداف در سایر تصاویر سنجش‌ازدوری دارد؟

مرور مقالات نشان داده‌است که یادگیری عمیق یک روش قدرتمند برای استخراج ویژگی‌های مکانی و آشکارسازی هدف در تصاویر پنکروماتیک است. با این حال، هنوز چالش‌هایی در این زمینه وجود دارد. چالش‌های مهم شامل نیاز به حجم زیاد تصویر برای آموزش شبکه‌های عمیق و تنوع آن با پس‌زمینه‌های مختلف برای تعمیم‌پذیری مدل آشکارسازی در صحنه‌های جدید است. افزایش تعداد سنجنده‌ها و تصاویر قابل دسترس می‌تواند به بهبود صحت آشکارسازی هدف در تصاویر پنکروماتیک کمک کند. از طرفی پیشرفت‌های سخت‌افزاری و طراحی شبکه‌های عمیق، قابلیت آشکارسازی بی‌درنگ را فراهم کرده‌اند.



شکل ۱- نمودار روند افزایشی تعداد مقالات آشکارسازی هدف در تصاویر پنکروماتیک بر گرفته از جست‌وجو در پایگاه اطلاعاتی گوگل اسکولار

۲- آشکارسازی هدف مبتنی بر یادگیری عمیق در تصاویر پنکروماتیک

پژوهش‌های قابل توجهی بر توسعه روش‌هایی برای شناسایی دقیق اهداف در تصاویر ماهواره‌ای پنکروماتیک متمرکز شده‌اند. این روش‌ها را به طور کلی می‌توان در چهار مرحله اصلی پیش‌پردازش تصویر، معماری‌های یادگیری عمیق، راهبردهای آموزش شبکه و پس‌پردازش آشکارسازی در نظر گرفت. در ادامه هر یک از این موارد مورد بررسی قرار گرفته‌اند.

۲-۱- پیش‌پردازش تصویر

مرحله پیش‌پردازش، گامی اساسی در بهبود کیفیت تصویر، کاهش نویز و جزئیات غیرضروری و برجسته‌سازی اهداف مورد نظر، پیش از تشخیص است [۲۶]. استفاده از فیلترهای غیرخطی [۲۷] و سراسری [۲۸] برای حذف نویز و در عین حال حفظ جزئیات مهم به ترسیم دقیق‌تر اهداف کمک می‌کند [۲۹]. بهبود کنتراست از طریق همسان‌سازی هیستوگرام^۱ یا برش^۲ برای گسترش دامنه پویایی^۳ شدت بازتاب موجب شده تا ویژگی‌ها و تفاوت‌های ظریف مربوط به هدف برجسته شوند [۳۰]. استفاده از عملیات مورفولوژیکی مانند opening و closing [۳۱] برای جداسازی اشیاء و حذف اشکالات کوچک، تشخیص نواحی محتمل از اهداف را تسهیل کرده‌است. ادغام تصویر پنکروماتیک با تصاویر چندطیفی از طریق pan-sharpening یا درون شبکه‌های عمیق به منظور استفاده از قدرت تفکیک مکانی بالای پنکروماتیک و اطلاعات طیفی باندهای چندطیفی، تصویر ورودی را غنی کرده‌است [۳۲]. اعمال تبدیل‌هایی مانند فوریه [۳۳] یا موجک [۳۴] بر روی تصویر برای برجسته‌سازی اهداف بر اساس ویژگی‌های حوزه فرکانس، موجب افزایش تمایز هدف از پس‌زمینه شده‌است. انتخاب مناسب روش‌های پیش‌پردازش بر اساس عواملی مانند میزان نویز تصویر و ویژگی‌های هدف اهمیت دارد. مرحله پیش‌پردازش با فیلتر کردن نویز، تقویت اهداف و ایجاد تصاویر ورودی غنی، به طور قابل توجهی دقت تشخیص را بهبود بخشیده‌است.

۲-۲- معماری‌های یادگیری عمیق

برای تصاویر پنکروماتیک، معماری‌های مختلفی از شبکه‌های عمیق معرفی شده‌اند. این طراحی‌های ساختاری شامل نوع، لایه‌ها، افزونه‌ها و توابع زیان آموزش شبکه است [۳۵] و [۳۶]. در بسیاری از موارد نیز، شبکه‌های طراحی شده برای سایر تصاویر با تغییرات اندکی برای تصویر پنکروماتیک استفاده شده‌اند. کمترین میزان تغییرات تنها در نحوه ورود تصویر به شبکه اعمال شده‌است. به طور مثال با تکرار کانال مربوط به تصویر پنکروماتیک و در نظرگیری آن به جای سه کانال مرتبط با تصویر RGB، از اطلاعات مکانی دقیق پانکروماتیک استفاده کرده و شبکه تمرکز بیشتری بر ویژگی‌های مکانی صحنه داشته‌است [۳۷].

۲-۲-۱- شبکه‌های عصبی کانولوشن

یکی از نوع‌های اولیه شبکه در این حوزه عمدتاً بر استفاده از CNN ساده برای استخراج ویژگی‌های سلسله مراتبی از تصویر ورودی و سپس لایه‌های کاملاً متصل^۴ برای طبقه‌بندی، متمرکز بوده‌است [۳۸]. با این حال، نقطه ضعف اولیه این شبکه، فقدان قابلیت مکان‌یابی بوده که در ساختارهای بعدی مرتفع شده‌است [۳۹]. شکل ۲ ساختار کلی شبکه‌های عصبی کانولوشن را نشان می‌دهد. در پژوهش [۴۰] به بررسی استفاده از ویژگی‌های یادگیری عمیق برای آشکارسازی پرداخته شده‌است. شناسایی اهداف کوچک مانند اتومبیل‌ها در تصاویر پنکروماتیک توسط آموزش CNN به مجموعه تصاویر برچسب‌گذاری شده بزرگ نیاز دارد. برای رفع این نیاز یک راه جایگزین، استفاده از ویژگی‌های عمیق از شبکه‌های از پیش‌آموزش دیده [۴۱] به عنوان توصیف‌گرهایی از صحنه مورد نظر است. بدین منظور از شبکه عمیق از پیش‌آموزش دیده CaffeNet [۴۲] بر روی مجموعه تصاویر ImageNet [۴۳] در مرحله استخراج ویژگی استفاده شده‌است. این ویژگی‌ها به یک طبقه‌بند ماشین بردار پشتیبانی (SVM)^۵ (خطی [۴۴]) برای تشخیص اهداف اتومبیل داده می‌شوند. این روش جایگزین مناسبی برای آموزش CNN از پایه ارائه داده و بدون نیاز به تصاویر آموزشی فراوان، آشکارسازی اهداف انجام شده‌است.

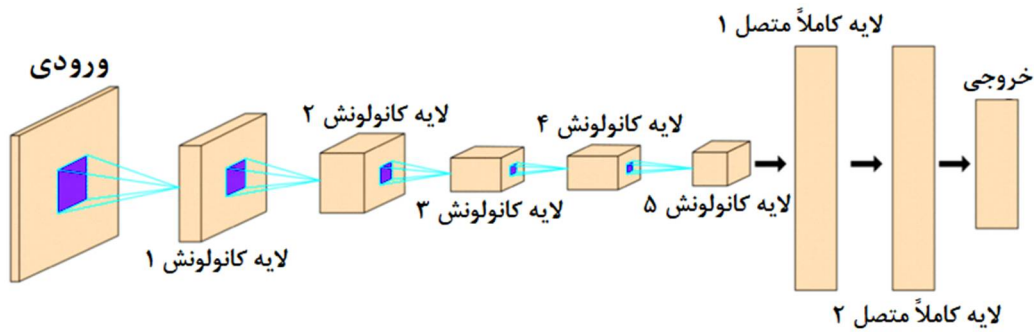
۴ Fully connected layers

۵ Support Vector Machine

۱ Histogram equalization

۲ Clipping

۳ Dynamic range



شکل ۲- ساختار کلی آشکارسازها با لایه‌های کانولوشن

شبکه Inception V3 [۴۹] ساخته شده‌اند و با استفاده از یک لایه MaxOut ترکیب می‌شوند.

دیگر پژوهش انجام شده [۹]، روشی برای استخراج خودکار اهداف ساختمان‌ها و جاده‌ها از ادغام تصویر پنکروماتیک با تصویر چندطیفی ارائه داده‌است. این روش ابتدا تصویر چندطیفی را با استفاده از pan-sharpening و فیلتر دو طرفه^۵ [۵۰] با بهبود لبه پیش‌پردازش می‌کند. یک شبکه عصبی CNN ویژگی‌های محلی پیکسل‌های همسایه را در تصویر آموزش می‌بیند. در نهایت، خوشه‌بندی fuzzy c-means [۵۱] ویژگی‌های طیف، شکل و حاصل از CNN را برای طبقه‌بندی هر شیء به عنوان هدف یا غیرهدف ترکیب می‌کند. محدودیت اصلی این شبکه، حذف برخی از مناطق حاوی هدف، به دلیل وجود درختان و سایه‌ها است. به منظور آشکارسازی اهداف هواپیما، پژوهش [۵۲] از یک شبکه عمیق سراسری استفاده کرده‌است تا ویژگی‌ها را مستقیماً از تصویر بیاموزد. معماری شبکه شامل لایه‌های کانولوشن و ادغام^۶ برای استخراج ویژگی، لایه‌های کاملاً متصل برای ترکیب ویژگی‌ها و لایه‌های خروجی برای پیش‌بینی کلاس و مکان هدف است. برای پیش‌بینی مکان هدف، تصویر به شبکه‌هایی تقسیم شده و بر حسب شبکه‌های تقسیم‌شده، مکان هدف به دست می‌آید.

در ادامه آشکارسازی اهداف مسکونی، پژوهش [۵۳] روشی برای آشکارسازی اهداف ساختمان از تصاویر ماهواره‌ای پنکروماتیک در کنار تصویر چندطیفی استفاده کرده‌است. شبکه آشکارساز، گروهی از سه مدل CNN برای آشکارسازی ساختمان‌ها است که هر کدام ورودی‌های مختلفی از باندهای چندطیفی و تصویر پنکروماتیک را می‌گیرند. ساختار لایه‌ای

روش ارائه شده در مطالعه [۸] سیستمی برای تشخیص خودکار اهداف نظامی مانند هواپیماها و تانک‌ها در تصاویر پنکروماتیک با رزولوشن بالا است. همان طور که ذکر شد ویژگی‌هایی مانند HOG و گابور برای مجموعه عظیمی از تصاویر پنکروماتیک به خوبی کار نکرده‌اند. سیستم پیشنهادی از جعبه‌های لبه^۱ [۴۵] و شبکه‌های عمیق CNN برای شناسایی اهداف در صحنه‌های پیچیده استفاده کرده‌است. جعبه‌های لبه با استفاده از اطلاعات مفید لبه در تصاویر، پیشنهادهایی برای اشیاء یا مناطق هدف احتمالی ایجاد می‌کنند. جعبه‌ی لبه برای هر هدف محتمل بر اساس خطوط مرزی^۲ [۴۶] و همبستگی لبه^۳ [۴۷]، امتیازی را محاسبه می‌کند. این امر به فیلتر کردن نواحی با احتمال پایین از وجود هدف کمک می‌کند. سپس نواحی فیلترشده به یک شبکه عمیق CNN برای استخراج و طبقه‌بندی خودکار ویژگی‌ها وارد می‌شوند. به طور کلی شبکه‌های CNN در برابر جابجایی‌ها و چرخش‌های جزئی هدف مقاوم هستند.

تحقیق انجام شده در [۴۸] یک رویکرد نوآورانه برای آشکارسازی خودکار اهداف مسکونی برای کاربردهایی مانند برنامه‌ریزی شهری، تحلیل استفاده از زمین و ارزیابی بلایا ارائه شده‌است. این مرجع یک مدل شبکه عصبی کانولوشنی دو جریان (DSCNN)^۴ (ارائه می‌دهد که می‌تواند به طور مؤثر اطلاعات تصاویر پنکروماتیک و چندطیفی را برای شناسایی مناطق مسکونی ترکیب کند. این مدل شامل جریان‌های جداگانه‌ای برای پردازش تصاویر پنکروماتیک و چندطیفی است که به ترتیب بر استخراج ویژگی‌های شکل، بافت و ویژگی‌های طیفی تمرکز دارند. جریان‌ها بر اساس معماری

۴ Double Stream Convolutional Neural Network

۵ Bilateral

۶ Pooling

۱ EdgeBox

۲ Boundary contours

۳ Edge affinity

U-Net [۵۴] اصلاح شده، برای این شبکه در نظر گرفته شده است. ادغام خروجی‌های سه شبکه CNN برای به دست آوردن نتیجه نهایی از اهداف ساختمان انجام می‌شود. به طور خلاصه، این روش نشان داده است که استفاده از اطلاعات مکانی تصویر پنکروماتیک در کنار تصویر چندطیفی و رویکرد گروه‌سازی چند شبکه موجب تشخیص بهتر اهداف می‌شود. شبکه CNN عمیق چندوجهی [۱۷]، مبتنی بر شیء OMM-CNN^۱) طراحی شده، که برای استخراج و آشکارسازی اهداف ساختمان از تصاویر پنکروماتیک و چندطیفی استفاده می‌شود. در این شبکه چالش‌های ناشی از زمینه‌های متنوع تا حد زیادی برطرف شده است. خوشه بندی تکرارشونده (SLIC)^۲) [۵۵] و [۵۶] به طور همزمان تصاویر پنکروماتیک و چندطیفی را به superpixelها تقسیم کرده که با در نظرگیری اطلاعات طیفی-مکانی در الگوریتم SLIC، نویز کاهش داده می‌شود. شبکه عصبی چندوجهی از لایه‌های کانولوشن و ادغام خودمنطبق^۳ [۵۷] برای به حداکثر رساندن اطلاعات طیفی-مکانی استفاده می‌کند. خروجی‌ها با استفاده از superpixelهای بخش بندی شده^۴ بهبود یافته و دقت آشکارسازی شیء محتمل از هدف را افزایش داده است. روش COLOR^۵ معرفی شده در تحقیق [۵۸] یک چارچوب پیشگام برای تشخیص کارآمد فرودگاه و هواپیماها در تصاویر پنکروماتیک است. این روش دارای یک شبکه عمیق برای مکان‌یابی سریع فرودگاه و یک شبکه تشخیص هواپیما است. نوآوری کلیدی این روش در رویکرد داده‌محور و طراحی آن است که انطباق مداوم مدل با توزیع‌های مختلف و در حال تکامل تصاویر متنوع را ممکن می‌سازد.

۲-۲-۲- شبکه‌های پیشنهاددهنده ناحیه

این گونه از شبکه‌های عمیق آشکارساز با در نظرگیری دو تابع زیان فرآیند آشکارسازی را به دو مرحله تفکیک می‌کنند. در مرحله اول شبکه بر اساس تصاویر آموزشی تابع زیان مکان‌یابی هدف را بهینه‌سازی کرده تا بتواند در ابتدا

نواحی محتمل اهداف را در صحنه تشخیص دهد [۵۹] و [۶۰]. در مرحله دوم نواحی محتمل در بخش طبقه‌بندی شبکه، ارزیابی شده و نوع هدف تعیین می‌شود [۶۱]. این نوع شبکه آشکارساز به دلیل تفکیک آشکارسازی به دو مرحله و بهینه‌سازی دو تابع زیان، سرعت پایین‌تری نسبت به شبکه‌های یک مرحله‌ای دارند [۶۲]. شکل ۳ نشان‌دهنده ساختار کلی این گونه از آشکارسازها است.

یک روش مبتنی بر یادگیری عمیق جدید در پژوهش [۶۳] برای تشخیص خودکار کشتی ارائه شده است. این شبکه داری معماری کانولوشن به همراه تجزیه مقدارهای منفرد (SVD)^۶) [۶۴] برای یادگیری ویژگی است. این رویکرد در دو مرحله عمل می‌کند: شناسایی نواحیه محتمل کشتی و تأیید وجود کشتی. در مرحله اول، یک ساختار چندلایه کانولوشن، نواحی محتمل را برجسته کرده و ویژگی‌های سطح پایین با استفاده از SVD روی وصله‌های^۷ تصویر استخراج می‌شوند. در مرحله دوم، نواحی محتمل از هدف طبقه‌بندی می‌شوند. به منظور آشکارسازی اهداف کشتی، آشکارساز معرفی شده در [۶۵] نیز روشی جدید برای تشخیص این گونه اهداف در تصاویر پنکروماتیک ارائه شده است. اهداف کشتی مدنظر در تصاویر اندازه‌های بسیار متفاوتی دارند (از چند پیکسل تا هزاران پیکسل). اکثر روش‌ها فقط برای محدوده کوچکی از مقیاس‌های اندازه عمل می‌کنند. شبکه ارائه شده، بر اساس معماری Faster R-CNN [۶۶] ساخته شده است. این آشکارساز از یک شبکه CNN به عنوان استخراج‌کننده ویژگی، و برای رسیدگی به مقیاس‌های چندگانه، یک لایه فیلتر انتخابی سلسله مراتبی HSF^۸ جدید معرفی کرده است. لایه HSF نواحی محتمل اهداف کشتی را در شبکه ارائه می‌دهد.

رویکردی دیگر در آشکارسازی کشتی با استفاده از تشخیص ناهنجاری انجام شده [۶۵] و شبکه تحلیل مولفه اساسی تجمیع هرم مکانی (SPP-PCANet)^۹) را پیشنهاد می‌کند [۶۷]. برای استخراج نواحی محتمل، یک الگوریتم تشخیص ناهنجاری بر اساس توزیع گوسی چند متغیره

۶ Singular Value Decomposition

۷ Patch

۸ Hierarchical Selective Filtering

۹ Spatial Pyramid Pooling Principal Component Analysis Network

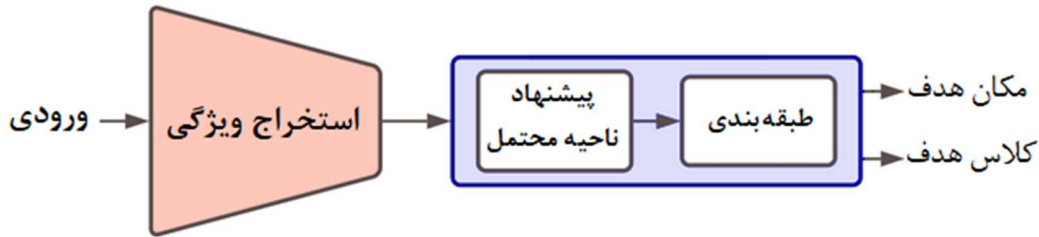
۱ Object-based Multi-Modal Convolution Neural Network

۲ Simple Linear Iterative Clustering

۳ Self-adaptive pooling

۴ Segmented

۵ Cycling, Offline Learning, and Online Representation



شکل ۳- ساختار کلی آشکارسازهای پیشنهاددهنده ناحیه

مطالعه انجام شده در [۷۴] نقش حیاتی pan-sharpening در بهبود آشکارسازی اهداف، به ویژه با استفاده از شبکه Faster R-CNN در تصاویر چندطیفی را بررسی کرده است. عملیات pan-sharpening با ادغام تصاویر چندطیفی با وضوح کم و تصاویر پنکروماتیک با دقت مکانی بالا، یک تصویر مصنوعی با تفکیک پذیری مکانی و طیفی بهبود یافته ایجاد می کند. هفت روش pan-sharpening مورد بررسی قرار گرفته اند که در میان آن ها، روش جایگزینی جزئی مولفه تطبیقی (PRACS) [۷۵] با کمترین اعوجاج، بهترین عملکرد را به نمایش گذاشته است. در مرحله آشکارسازی از شبکه عمیق Faster R-CNN انجام شده است که در آن تأثیر pan-sharpening کاملاً مشهود است. آشکارسازی اهدافی مانند کشتی ها، زمین های بسکتبال و زمین های بازی در تصاویر sharp شده نسبت به تصاویر اصلی به طور قابل توجهی بهبود یافته است. با وجود تأثیر قابل توجه pan-sharpening بر دقت تشخیص اهداف، انتخاب روش مناسب نقش بسیار مهمی ایفا می کند [۷۶]. بر اساس معیار صحت آشکارسازی، روش مدولاسیون شدت مبتنی بر فیلتر هموار (SFIM) [۷۷] بهترین عملکرد را نشان می دهد و پس از آن تجزیه و تحلیل مؤلفه اصلی (PCA) [۷۸] و PRACS قرار دارند. نکته جالب توجه این است که بالاترین کیفیت تصویر که توسط PRACS به دست آمده، همیشه با دقت تشخیص برتر همخوانی ندارد. اعوجاج مکانی، همانطور که در روش هایی مانند فیلتر کردن بالاگذر و SFIM مشاهده شده است، به عنوان عامل کلیدی مؤثر بر عملکرد تشخیص مطرح می شود. در اصل، پژوهش مذکور تأثیر قابل توجه pan-sharpening بر توانایی های شبکه Faster R-CNN برای آشکارسازی هدف در تصاویر چندطیفی را نشان داده است. در مطالعه [۷۹] نیز تأثیر pan-sharpening بر تشخیص هدف

استفاده شده و کشتی ها به عنوان ناهنجاری در برابر پس زمینه دریا تلقی می شوند. سپس وصله های محتمل به شبکه طراحی شده برای طبقه بندی، داده می شوند. PCANet از تبدیل PCA برای استخراج ویژگی استفاده کرده و بخش SPP امکان آموزش چندمقیاسی اهداف را فراهم می کند. در نهایت ویژگی های آموزشی استخراج شده از شبکه عمیق با استفاده از SVM طبقه بندی می شوند.

تحقیق انجام شده در [۶۸] با تمرکز بر آشکارساز اهداف ساختمان، سه شبکه Faster R-CNN [۶۶]، Mask R-CNN [۶۹] و YOLOv3 [۷۰] را مورد بررسی قرار داده است که از بین این شبکه های عمیق، شبکه Mask R-CNN بهترین نتیجه را کسب کرده است. این شبکه می تواند به صورت همزمان شناسایی و بخش بندی اهداف را انجام دهد. این مدل به طور خاص قادر است تا اهدافی مانند هواپیماها و ساختمان ها را شناسایی کند.

از تبدیل موجک گسسته (DWT) [۷۱] در پژوهش [۷۲] و یک شبکه عمیق متراکم باقیمانده (DRDN) [۷۳] برای آشکارسازی استفاده شده است. این روش شامل سه مرحله: بخش بندی دریا-خشکی، استخراج ناحیه محتمل کشتی با استفاده از DWT و طبقه بندی کشتی با DRDN. تصاویر پنکروماتیک جزئیات مکانی تصویر چندطیفی را افزایش می دهند. در مرحله تقسیم بندی دریا به خشکی، از تصاویر چندطیفی استفاده شده و شاخص تفاوت هنجار شده آب (NDWI) [۳] برای جداسازی مناطق دریا و خشکی محاسبه می شود. استخراج نواحی محتمل کشتی شامل اعمال DWT به تصویر مادون قرمز نزدیک بوده و مرحله طبقه بندی کشتی از یک DRDN با چهار بلوک متراکم استفاده می کند. این ساختار استفاده مجدد از ویژگی و یادگیری ویژگی سلسله مراتبی را تقویت می کند.

۵ Smoothing Filter-based Intensity Modulation

۶ Principal Component Analysis

۱ Discrete Wavelet Transform

۲ Deep Residual Dense Network

۳ Normalized Difference Water Index

۴ Partial Replacement Adaptive Component Substitution

در تصاویر پنکروماتیک بررسی شده است. سه الگوریتم شدت-رنگ-اشباع تعمیم یافته (GIHS)^۱ [۸۰]، افزودن موجک به تصویر پنکروماتیک (AWLP)^۲ [۸۱] و حداقل میانگین مربع خطا (MMSE)^۳ [۸۲] بر روی مجموعه تصویر QuickBird آزمایش شده که نشان می‌دهد MMSE موثرترین روش در تشخیص هدف با هشدارهای کاذب کمتر است.

پژوهش مطرح شده در [۸۳] کاربرد نوآورانه‌ای از الگوریتم‌های یادگیری عمیق، با تأکید اولیه بر مدل Faster R-CNN، برای شناسایی خودروها در تصاویر ماهواره‌ای با رزولوشن بالا را در شرایط ترافیک ارائه کرده است. در زمینه سیستم‌های حمل و نقل هوشمند، روش‌های سنتی تشخیص خودرو با محدودیت‌های ذاتی در صحت و استحکام مواجه هستند. این پژوهش با درک موفقیت یادگیری عمیق در وظایف تشخیص و آشکارسازی اهداف در تصویر، بر انگیزه استفاده از این روش‌ها برای بهبود آشکارسازی هدف خودرو تأکید کرده است. در بررسی کارهای مرتبط، نویسندگان اشاره کرده‌اند که روش‌های سنتی تشخیص خودرو اغلب به ویژگی‌هایی مانند شکل و حرکت متکی بوده و با چالش‌هایی مانند انسداد^۴ و وسایل نقلیه با حرکت آهسته روبرو هستند. در بخش استخراج ویژگی این روش، از شبکه عمیق AlexNet [۸۴] برای طبقه‌بندی تصویر استفاده شده است که به شناسایی مناطقی که حاوی خودروها هستند کمک می‌کند. سپس، مدل Faster R-CNN برای تشخیص خودرو با تنظیم پارامترهای خاص برای چالش‌های ناشی از تشخیص خودروهای کوچک به کار گرفته شده است. برای گسترش مجموعه تصویر آموزشی و بهبود عملکرد مدل، از داده‌افزایی استفاده شده است. علاوه بر آن، بر اجرای روش‌های هرس مدل^۵ و گسسته‌سازی^۶ ماتریس وزن‌ها و ماتریس گرادیان در فرآیند آموزش و بهبود شبکه عمیق تأکید شده است. این کار بدون به خطر انداختن ظرفیت آموزش شبکه و با افزایش سرعت محاسبات همراه است. همچنین حجم مدل به طور موثر فشرده شده است. کاربرد مدل در دنیای واقعی مانند یک تقاطع خیابان، اثربخشی آن را در تشخیص جریان ترافیک و خودروهای با سرعت بالا نشان می‌دهد.

آشکارساز طراحی شده در مطالعه [۸۵]، نوآوری قابل توجهی را با اعمال اصلاحاتی به شبکه عمیق Mask R-CNN معرفی کرده و نتایج آشکارسازی را بهبود داده است. یکی از محدودیت‌های این شبکه عدم مسیر پایین به بالا^۷ در شبکه هرم ویژگی (FPN)^۸ [۸۶] بوده که این مورد اضافه شده است. شبکه‌های FPN با پردازش تصاویر در مقیاس‌های چندگانه در تشخیص اهداف عملکرد مناسبی دارند. این هرم ترکیبی از نقشه‌های ویژگی استخراج شده است که لبه‌ها و بافت‌های دقیق در پایه، نمایش‌های انتزاعی‌تر و ویژگی‌های سراسری در بالاترین سطح قرار می‌گیرند [۸۷]. در رویکردهای قبلی، اطلاعات صرفاً از بالا به پایین جریان می‌یافت، اما بررسی مسیر پایین به بالا نشان‌دهنده اهمیت آن در صحت آشکارسازی است. این مسیر با اعمال اتصالات جدید در شبکه انجام می‌شود. در واقع با این اصلاح، اهداف کوچک بهتر آشکار می‌شوند. در مرحله بعد، سازوکار توجه کانالی و مکانی برای تمرکز بر نواحی محتمل از کشتی و کاهش نویز به کار برده شده است. سازوکار توجه، یک روش در یادگیری عمیق است که به شبکه‌های عصبی اجازه می‌دهد تا بر روی ویژگی‌های خاصی از تصاویر تمرکز کنند. توجه مکانی بر روی ویژگی‌های مکانی مانند لبه‌ها و بافت‌ها و توجه کانالی بر روی ویژگی‌هایی مانند رنگ و روشنایی در تصویر تمرکز می‌کند. سپس از مدل ResNet-101 [۸۸] که با مجموعه تصویر ImageNet از پیش آموزش دیده شده است [۸۹]، به عنوان استخراج‌گر ویژگی استفاده شده است. شبکه به سرهای جداگانه برای تشخیص جعبه محدود کننده^۹ و طبقه‌بندی آن تقسیم می‌شود. در پژوهش [۹۰]، یک سیستم آشکارسازی هدف با پیچیدگی پایین، با الهام از بینایی حشرات معرفی شده است. با تقلید از پیکربندی چندچشمی^{۱۰} [۹۱]، فشرده‌سازی اطلاعات و سازوکار توجه، یک آشکارساز اهداف کوچک ارائه شده است. این سیستم چالش پس‌زمینه پیچیده را کاهش می‌دهد که از تأثیرات سازوکارهای توجه است.

مطالعه انجام شده در [۹۲] یک روش جدید برای تشخیص مناطق مورد نظر (ROIs)^{۱۱} [۹۳] در تصاویر پنکروماتیک با وضوح بالا ارائه داده است. برای این مقصود از

۸ Feature Pyramid Network

۹ Bounding box

۱۰ Multi-eye

۱۱ Region Of Interest

۱ Generalized Intensity-Hue-Saturation

۲ Additive Wavelet L to Pan

۳ Minimum Mean Square Error

۴ Occlusion

۵ Model Pruning

۶ Quantization

۷ Bottom-up

اصول توجه بصری در شناسایی و برجسته‌سازی نواحی محتمل هدف استفاده شده‌است. چندین مدل توجه بصری موجود، از جمله Itti [۹۴]، برجستگی بصری مبتنی بر گراف (GBVS) [۹۵]، باقی‌مانده‌های طیف^۲ [۹۶] و رویکردهای مبتنی بر موجک بررسی شده‌اند. برای پرداختن به این موضوع، یک مدل ادغام ویژگی چندمقیاسی تخصصی برای تصاویر پنکروماتیک ایجاد کرده‌اند. به این صورت که تصویر را در دو کانال شدت و جهت پردازش می‌کند. کانال شدت از یک روش جدید باقی‌مانده‌های طیف برای محاسبه برجستگی و کانال جهت‌گیری با استفاده از تبدیل موجک و استخراج ویژگی جهت، انجام می‌شود.

۲-۲-۳- آشکارسازهای یک مرحله‌ای

همان طور که در بخش قبلی ذکر شد، آشکارسازهای یک مرحله‌ای سرعت بالاتری نسبت به شبکه‌های پیشنهاددهنده ناحیه دارند. دلیل آن هم بهینه‌سازی همزمان ترکیب زیان-های حاصل از مکان‌یابی و نوع هدف است. اگرچه این نوع شبکه‌ها سرعت بالایی در آشکارساز اهداف دارند اما صحت آشکارسازی آن‌ها کاهش می‌یابد [۹۷]. در شکل ۴ ساختار آشکارسازهای یک مرحله‌ای آورده شده‌است.

یک آشکارساز یک مرحله‌ای و چندجعبه‌ای بهبودیافته (RSSD)^۳ برای تصاویر پنکروماتیک با وضوح بالا در پژوهش [۱۰] معرفی شده‌است. مقصود این آشکارساز دستیابی به تشخیص دقیق با سرعت بالا است. این شبکه از SSD به عنوان مبنا برای پیش‌بینی سریع مکان و نوع اهداف از روی ویژگی‌های تصویر استفاده می‌کند. شبکه SSD اغلب برای اهداف مورد نظر امتیاز اطمینان^۴ پایینی دارد. برای بهبود این امر، RefinedNet معرفی شده‌است که نواحی محتمل را از SSD گرفته، امتیازات اطمینان را برای تقویت اهداف بالقوه و کاهش تأثیر پس‌زمینه اصلاح می‌کند. همچنین از شبکه VGG-16 [۹۸] اصلاح‌شده برای استخراج ویژگی‌های چند-مقیاسه استفاده شده‌است.

در مطالعه [۹۹] یک رویکرد یک مرحله‌ای و مبتنی بر یادگیری عمیق برای تشخیص ساختمان توسط تصاویر پنکروماتیک و چندطیفی ارائه شده‌است. روش‌های مرسوم ابتدا تصاویر پنکروماتیک و چندطیفی را با استفاده از

تکنیک‌های pan-sharpening ترکیب می‌کنند و سپس روی تصاویر ادغام شده آشکارسازی انجام می‌شود. با این حال، این مقاله نشان داده‌است که چنین روش‌های ادغام در سطح تصویر برای تشخیص اهداف بهینه‌سازی نشده‌اند و ممکن است جزئیات مکانی ارزشمندی که در تصاویر پنکروماتیک برای شناسایی ساختمان‌ها وجود دارند را از دست بدهند. برای پرداختن به این مورد، یک شبکه ادغام نامتقارن دو جریانی (DAFNet)^۵ پیشنهاد شده‌است که ویژگی‌ها را مستقیماً از تصاویر پنکروماتیک و چندطیفی به طور جداگانه با استفاده از ResNet50 استخراج می‌کند. دو شبکه هرم ویژگی مستقل بر روی دو جریان تصویر برای دست آوردن ویژگی‌های چندمقیاسی ساخته شده‌اند. نشان داده شده‌است که این ادغام دوگانه FPN نسبت به ادغام همزمان مقیاس و نوع تصویر مؤثرتر است. هسته شبکه DAFNet یک افزونه ادغام ویژگی نامتقارن (AFF)^۶ است که ویژگی‌های پنکروماتیک را در حین ادغام با استفاده از اتصال پرشی برجسته می‌کند. این کار مانع از دست رفتن جزئیات مکانی حیاتی از تصاویر پنکروماتیک می‌شود.

۲-۲-۴- شبکه‌های سبک‌وزن

مقصود شبکه‌های عصبی سبک‌وزن دستیابی به صحت بالا برای تشخیص هدف در عین کاهش منابع محاسباتی و حافظه مورد نیاز است. این کار برای استقرار سیستم‌های تشخیص هدف در شرایط محدودیت سخت‌افزاری مورد توجه است. اصطلاح "سبک‌وزن" به معماری‌هایی اطلاق می‌شود که پارامترها و عملیات کمتری در مقایسه شبکه‌های بزرگ مانند ResNet یا VGGNet دارند. این امر از طریق روش‌هایی مانند کاهش اندازه فیلتر به دست می‌آید [۱۰۰]. کاهش پارامترها و محاسبات به این نوع از شبکه‌های آشکارساز اجازه می‌دهد تا سریع‌تر عمل کرده و از حافظه کمتری استفاده کنند و در عین حال صحت خوبی را داشته باشند. شبکه‌های سبک وزن معمولاً برای آشکارسازی اهداف کوچک مناسب هستند. همچنین می‌توانند به طور موثر ویژگی‌های مهمی را برای شناسایی اهداف، بدون نیاز به یک مدل بیش از حد پیچیده بیاموزند.

۴ Confidence score

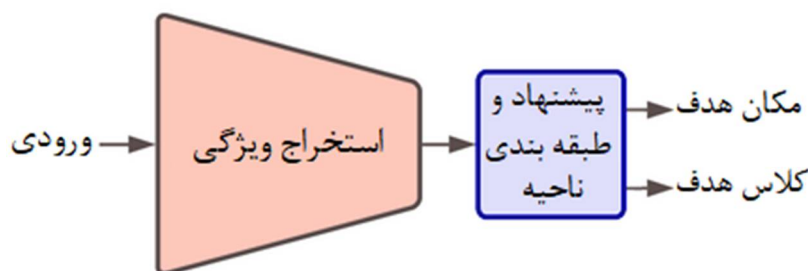
۵ Dual-stream Asymmetric Fusion Network

۶ Asymmetric Feature Fusion

۱ Graph Based Visual Saliency

۲ Spectrum residuals

۳ Refined Single-Shot multibox Detector



شکل ۴- ساختار کلی آشکارسازهای یک مرحله‌ای

تکرارشونده جدید، نواحی خشکی و دریایی را با دقت جدا می‌کند و هشدارهای کاذب را به حداقل می‌رساند. در مرحله بعد یک روش توزیع گوسی چند متغیره به طور موثر نواحی محتمل از هدف کشتی را در صحنه‌های پیچیده استخراج می‌کند. سپس یک مجموعه تصویر چند منبعی با ترکیب تصاویر پیش پردازش شده کشتی از تصاویر پنکروماتیک با تصاویر فرسرخ را ایجاد کرده‌اند که یادگیری ویژگی‌های کشتی را از طریق داده‌افزایی تسهیل می‌کند. در نهایت ترکیبی از ویژگی‌های سراسری تبدیل فوریه و یک CNN سبک‌وزن که برای سکوها‌های ماهواره‌ای طراحی شده‌است، جهت طبقه‌بندی نواحی محتمل استفاده می‌شود. این روش به طور قابل توجهی نیازهای ذخیره‌سازی و محاسباتی را در مقایسه با شبکه‌های عمیق استاندارد کاهش داده‌است. چارچوب پیشنهادی یک راه حل مناسب برای تشخیص کشتی با ترکیب تصاویر مادون قرمز و تصویر پنکروماتیک در شرایط محدودیت منابع است.

۲-۳- راهبردهای آموزش شبکه

شبکه‌های عمیق پیچیده با میلیون‌ها پارامتر برای دستیابی به عملکرد مناسب نیاز به راهبردهای آموزشی پیشرفته دارند. برای بهینه‌سازی عملکرد تشخیص، غلبه بر مسائلی مانند بیش‌برازش، خطای موجود در جعبه‌های محدودکننده تصویر و عدم تعادل کلاس بسیار مهم است. این مسائل را می‌توان با به کارگیری راهبردهای آموزشی دقیق، مانند داده‌افزایی، یادگیری انتقالی، کاهش نمونه کم کیفیت و سازگاری مدل برطرف کرد [۱۰۵]. علاوه بر این روش‌ها، چندین عامل دیگر وجود دارد که می‌تواند بر عملکرد شبکه‌های عمیق برای آشکارسازی هدف در تصاویر پنکروماتیک

در تحقیق [۱۰۱] رویکردی برای تشخیص و طبقه‌بندی خودکار کشتی‌ها در تصاویر پنکروماتیک با وضوح بالا ارائه شده‌است. شبکه عصبی سبک‌وزن با ساختار پیش‌رو^۱ برای طبقه‌بندی هر شیء تصویر به عنوان هدف کشتی یا غیر-کشتی استفاده می‌شود. در کنار آن الگوریتم ژنتیک، ترکیب بهینه ویژگی‌های ورودی و تعداد گره‌ها را در لایه پنهان شبکه تعیین می‌کند. این مطالعه پتانسیل استفاده از شبکه‌های عصبی سبک‌وزن به همراه بهینه‌سازی‌هایی توسط الگوریتم‌های ژنتیک برای شناسایی کشتی‌ها در تصاویر پنکروماتیک را نشان داده‌است.

چارچوب فشرده در پژوهش [۱۰۲] یک آشکارساز کشتی با استفاده از شبکه‌های عصبی عمیق و ماشین یادگیری افراطی (ELM)^۲ [۱۰۳] برای رسیدن به چالش‌های محاسباتی را پیشنهاد می‌کند. در مرحله آشکارسازی، دو رمزگذار خودکار حذف نویز موازی (SDA)^۳ فرکانس‌های مختلف تصویر را آموزش می‌بینند. فرکانس‌های پایین اطلاعات ضمنی کشتی را شامل شده و فرکانس‌های بالا جزئیات لبه و ساختار را نشان می‌دهند. خروجی‌های دو SDA با استفاده از یک ELM برای طبقه‌بندی ترکیب می‌شوند. ELM یک الگوریتم شبکه عصبی تک لایه کارآمد است که به طور تصادفی پارامترها را مقداردهی اولیه کرده و وزن خروجی را به صورت تحلیلی محاسبه می‌کند. این کار از تنظیم تکرارشونده جلوگیری کرده و با یادگیری بسیار سریع به تعمیم‌پذیری بالایی می‌رسد.

آشکارساز ارائه شده در [۱۰۴]، با هدف غلبه بر چالش‌های ناشی از صحنه‌های پیچیده و منابع محاسباتی محدود در ماهواره‌ها، روشی نوآورانه برای تشخیص هدف کشتی به شمار می‌آید. رویکرد پیشنهادی شامل یک فرآیند چهار مرحله‌ای است. در ابتدا یک الگوریتم بخش‌بندی

۱ Feed-forward

۲ Extreme Learning Machine

۳ Stacked Denoising Autoencoders

تأثیر بگذارد. این عوامل شامل انتخاب تابع زیان، بهینه‌ساز، نرخ یادگیری و اندازه دسته‌ای^۱ است. انتخاب ترکیب مناسبی از این عوامل می‌تواند عملکرد مدل را به میزان قابل توجهی بهبود بخشد.

۲-۳-۱- داده افزایی

به طور کلی چالش‌های آشکارسازی هدف به چند دسته مهم تقسیم می‌شوند. برای مقابله با این چالش‌ها از داده‌افزایی استفاده می‌شود تا شبکه برحسب حالت‌های مختلف و ممکن از هدف، آموزش دیده و تعمیم‌پذیری بالایی داشته باشد. چالش‌های مهم در آشکارسازی هدف عبارتند از: هدف چندمقیاسه، جهت‌گیری هدف، تأثیر محیط اطراف، انسداد، ابهام ظاهری و اهداف مرکب. در زیر هر یک از این موارد توضیح داده شده‌است:

- ✓ هدف چندمقیاسه: وجود هدف با اندازه‌ها و ابعاد مختلف در تصویر.
- ✓ جهت‌گیری هدف: زوایای مختلف از جهت قرارگیری هدف.
- ✓ تأثیر محیط اطراف: قرارگیری هدف در سایه - وجود اهداف به طور متراکم.
- ✓ انسداد: وجود تنها بخشی از هدف در تصویر مانند قرارگیری هدف در حاشیه‌های تصویر.
- ✓ ابهام ظاهری: شباهت زیاد هدف مطلوب با سایر اشیاء مشابه.
- ✓ اهداف مرکب: هدفی با ابعاد بزرگ (مانند آشکارسازی کارخانه‌ها) و وجود چندین زیرشیء^۲ درون آن.

داده‌افزایی راهبردی است که تولید نمونه‌های آموزشی جدید را با اعمال تبدیل‌هایی به نمونه‌های موجود ممکن می‌سازد. این کار به افزایش تعداد نمونه‌های مجموعه آموزشی و جلوگیری از بیش‌برازش کمک می‌کند. این تبدیل‌ها شامل: انتقال^۳، گردش^۴، چرخش^۵، مات‌سازی^۶، تغییر کنتراست/روشنایی^۷، برش^۸، مقیاس^۹ و نویز^{۱۰} به تصاویر اصلی است. این امر به طور قابل توجهی بیش‌برازش را کاهش می‌دهد و استحکام مدل را برای تغییرات ممکن از هدف بهبود

می‌بخشد [۱۰۶]. در شکل ۵ تبدیل‌های ذکر شده روی یک نمونه از اهداف پنکروماتیک اعمال شده‌است.

۲-۳-۲- یادگیری انتقالی

یادگیری انتقالی روش دیگری است که از شبکه‌های عمیق از پیش‌آموزش دیده برای مقداردهی اولیه وزن شبکه مورد نظر استفاده می‌شود. علاوه بر آن حفظ وزن‌های لایه‌ها و اضافه کردن لایه‌های جدید به شبکه نیز مدنظر است. یادگیری انتقالی به سرعت بخشیدن به روند آموزش و بهبود عملکرد مدل کمک می‌کند. شبکه‌هایی مانند ResNet و VGGNet [۱۰۷] که بر روی مجموعه تصاویر برچسب‌گذاری شده‌ی بزرگی مانند ImageNet از پیش آموزش دیده‌اند، بازنمایی‌های مهمی از ویژگی‌های عمومی تصویر را ارائه می‌دهند که می‌توانند به حوزه آشکارسازی منتقل شوند [۱۰۸]. تنظیم دقیق شبکه‌های از پیش‌آموزش دیده ویژگی‌ها را به طور موثر با حوزه جدید تطبیق می‌دهد و مقداردهی اولیه بسیار مناسبی از وزن‌های شبکه انجام می‌شود. این امر بر محدودیت در دسترس بودن تصاویر برچسب‌گذاری شده غلبه می‌کند.

۲-۳-۳- کاوش نمونه کم کیفیت

تمرکز آموزش روی نمونه‌های کم کیفیت، که در آن مدل در دوره‌های آموزشی^{۱۱} اولیه دچار اشتباه می‌شود، موجب توجه شبکه بر نقاط ضعف آن می‌شود. موارد نمونه‌های کم کیفیت شامل اهداف کوچک و منسدد، صحنه‌های متراکم و کنتراست کم است [۱۰۹].

۲-۳-۴- سازگاری مدل

تطبیق مداوم مدل بر روی تصاویر با و بدون برچسب جدید از مکان‌های مختلف و آب و هوا، مزایایی را برای تعمیم مدل نشان داده‌است. این روش مدل را قادر می‌سازد تا به طور تدریجی خود را در جریان تصاویر جدید برای مدیریت تغییرات توزیع آن‌ها بر روزرسانی کند [۱۱۰]. تعداد زیادی از تصاویر بدون برچسب را می‌توان به همراه برچسب‌های

۷ Contrast/brightness modification

۸ Cropping

۹ Scaling

۱۰ Noise

۱۱ Epochs

۱ Batch size

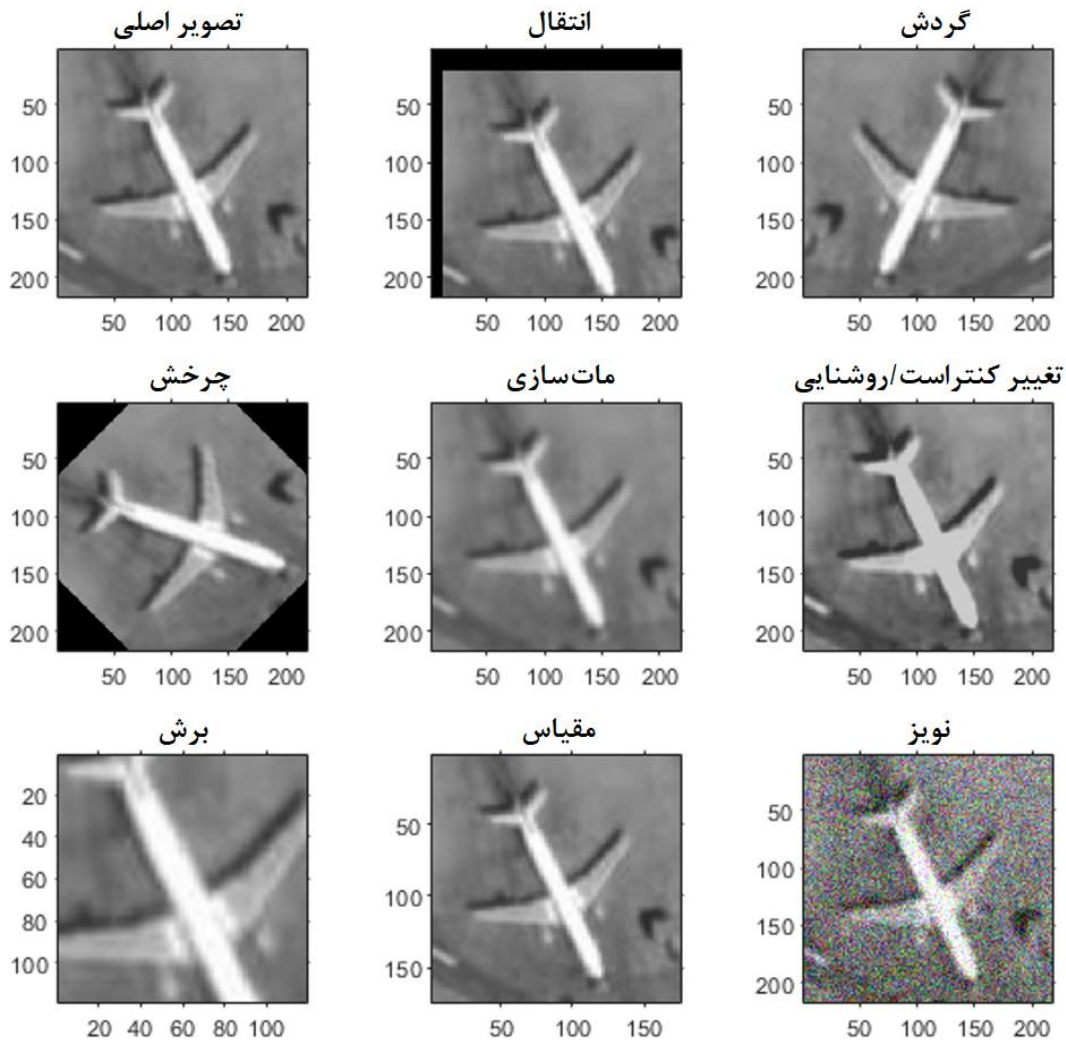
۲ sub-object

۳ Translation

۴ Flipping

۵ Rotation

۶ Blurring



شکل ۵- تبدیل‌های متداول در داده‌افزایی

غیرحداکثری^۲ است [۱۱۳]. این روش تشخیص‌های همپوشان را بر اساس اشتراک روی اجتماع^۳ [۱۱۴] در جعبه‌های محدودکننده و امتیازات اطمینان ادغام می‌کند. این روش به طور قابل توجهی شناسایی‌های متعدد از یک هدف را کاهش داده‌است. آستانه‌گذاری بر امتیازهای اطمینان شبکه، یک روش دیگر برای رد نواحی شبیه به هدف است. همچنین فیلتر کردن الگوهای مکانی هدف، به حذف اهداف کاذب کمک می‌کند. این فیلترسازی را می‌توان بر روی عواملی مانند اندازه هدف و نسبت ابعاد اعمال کرد. اطلاعات و دانش قبلی در مورد اشکال و مقیاس‌های هدف می‌تواند نتایج را اصلاح کند. پس‌پردازش یک راه مهم برای اصلاح

محدود از طریق روش‌های نیمه نظارتی مانند شبه-برچسب‌گذاری^۱ [۱۱۱] و کمیته‌سازی آنتروپی [۱۱۲] مورد استفاده قرار داد.

۲-۴- پس‌پردازش آشکارسازی

نتایج آشکارسازی اولیه و خام تولید شده توسط شبکه‌های عصبی عمیق اغلب حاوی تشخیص‌های تکراری برای یک هدف واحد هستند که با اهداف معتبر مطابقت ندارند. روش‌های پس‌پردازش معمولاً در خروجی‌های شبکه برای جمع‌بندی تشخیص‌ها و حذف چنین خطاهایی استفاده می‌شوند. یکی از پرکاربردترین روش‌ها، کاهش تأثیر

^۳ Intersection over union

^۱ Pseudo-labeling

^۲ Non-maximum suppression

آشکارسازی‌های شبکه است که قابلیت اطمینان تشخیص را تا حد زیادی بهبود می‌بخشد.

۳- مجموعه تصاویر مرتبط

تعدادی از مجموعه تصاویر حاوی تصاویر سنجش‌ازدور پنکروماتیک برای تسهیل تحقیقات در این زمینه ایجاد شده است. این مجموعه تصاویر معیارهایی مناسب برای توسعه و ارزیابی روش‌های آشکارسازی هدف ارائه می‌کنند. مجموعه تصاویر عمومی به همراه مشخصات، در جدول ۱ آورده شده اند. موفقیت الگوریتم‌های آشکارسازی هدف به طور اساسی به کیفیت و تنوع تصاویر آموزشی که در معرض آن قرار می‌گیرند بستگی دارد. خوشبختانه، محققان مجموعه تصاویر گسترده‌ای را به طور خاص برای تحلیل تصویر پنکروماتیک ایجاد کرده‌اند که هر کدام کاربری‌های مختلف دنیای واقعی ارائه می‌دهند.

مجموعه تصویر DOTA [۱۱۵] گواهی بر وسعت جهانی فناوری سنجش‌ازراه‌دور است. این مجموعه تصویر گسترده، ۲۸۰۶ تصویر پنکروماتیک با وضوح بالا را در بر می‌گیرد. اهداف مدنظر شامل هواپیما و کشتی است. تنوع مقیاس اهداف، مجموعه تصویر DOTA را به یک منبع ارزشمند برای تشخیص هدف چندمقیاسی تبدیل می‌کند. مجموعه تصویر NWPU VHR-10 [۱۱۶] بر بافت‌های پیچیده از پس‌زمینه متمرکز است. چالش اصلی این مجموعه تصویر تراکم اهداف و پیچیدگی حاصل از سایه‌ها و بازتاب‌ها است.

مجموعه تصویر SpaceNet [۱۱۷] بالاتری وضوح مکانی را در بین مجموعه‌های تصویر مورد بررسی دارد که ارزش اطلاعات مکانی را، به خوبی در صحت آشکارسازی اهداف ساختمانی نشان می‌دهد. این مجموعه تصویر به منظور آشکارسازی ساختمان‌های تخریب شده طی حوادث طبیعی است. مجموعه تصویر 5M-Building [۶۸] بیشترین تعداد تصویر را بین سایر مجموعه‌های تصویر مورد بررسی، به خود اختصاص داده‌است. این مجموعه تصویر غول‌پیکر، بیش از ۵ میلیون ساختمان را شامل می‌شود که به دقت از تصاویر هوایی در شهرهای و مناطق مختلف استخراج شده‌اند.

حاشیه‌نویسی‌ها فراتر از جعبه‌های محدودکننده ساده ارائه شده‌اند و ویژگی‌های اضافی مانند نوع و اندازه ساختمان را در بر می‌گیرند. در نهایت، مجموعه تصویر CARPK [۱۱۸] تمرکز را به آشکارسازی به ظاهر پیش پا افتاده اما حیاتی وسایل نقلیه در پارکینگ‌ها معطوف می‌کند. این مجموعه تصویر شامل تصاویر هوایی با دقت ۰/۵ متر بر پیکسل است که خودروهای پارک شده تحت شرایط مختلف را شامل می‌شود. چالش آشکارسازی این مجموعه تصویر، در تشخیص و شمارش دقیق خودروها با شرایط انسداد توسط سایر خودروها، سایه‌ها و حتی پوشش گیاهی نهفته‌است. الگوریتم‌هایی که روی CARPK آموزش دیده‌اند نقش حیاتی در بهینه‌سازی مدیریت فضای پارکینگ و بهبود کارایی سیستم‌های حمل‌ونقل شهری ایفا می‌کنند.

جدول ۱- مجموعه تصاویر عمومی برای آشکارسازی هدف در تصاویر پنکروماتیک

مجموعه تصویر	سنجنده	اهداف	تعداد تصاویر	دقت مکانی (متر)	پیوند
DOTA	Gaofen 1 & 2, Jilin-1, Pakistan Remote Sensing Satellite 1 & 2, Land Satellite-2	هواپیما - کشتی	۲۸۰۶	۰/۵ - ۲	https://captain-whu.github.io/DOTA/dataset.html
NWPU VHR-10	Gaofen 1 & 2, Jilin-1, Pakistan Remote Sensing Satellite-1, QuickBird-2	هواپیما - کشتی	۸۰۰	۰/۵ - ۲	https://pan.baidu.com/s/1DWibgMXGbC1V5aAuN54JUA?pwd=1234
SpaceNet	WorldView-3	ساختمان	۱۲۰۸	۰/۳	https://spacenet.ai/spacenet-buildings-dataset-v2/
5M-Building	Gaofen-2	ساختمان	۱۰۰۰۰	۰/۸	https://xut2019.github.io/
CARPK	Gaofen-2	وسیله نقلیه	۱۴۳۱	۰/۵	https://laf1.github.io/LPN/

پنکروماتیک با وضوح بالا را فراهم می‌کنند. تصاویر پنکروماتیک و حاشیه‌نویسی‌های دقیق آن‌ها، محققان و توسعه‌دهندگان را قادر می‌سازد تا اهداف مختلف را آشکارسازی کنند. در این تحقیق، بررسی جامعی از

انتخاب مجموعه تصویر بهینه برای آشکارسازی اهداف، یک تصمیم حیاتی است که بر اساس نیازهای خاص گرفته می‌شود. این مجموعه‌های متنوع تصویر، اساس پیشرفت‌های مبتنی بر یادگیری عمیق و هوش مصنوعی در تحلیل تصاویر

آورده شده است. TP_i تعداد اهداف کلاس i ام که به درستی تشخیص داده شده اند، FP_i تعداد اشیاء i ام که به اشتباه به عنوان هدف تشخیص داده شده اند و n_T تعداد کلاس‌های اهداف است.

$$mAP = \frac{1}{n_T} \sum_{i \in n_T} \frac{TP_i}{TP_i + FP_i} \quad (1)$$

شبکه‌های عصبی کانولوشن در استخراج ویژگی‌های بصری از تصاویر سرآمد هستند که آن‌ها را به انتخاب‌های اولیه برای آشکارسازی اهداف تبدیل می‌کند. به عنوان مثال، شبکه کانولوشنی SVDNet با استفاده از تجزیه مقادیر تکینه، ویژگی‌های متمایزگری را برای آشکارسازی هدف در تصاویر پنکروماتیک استخراج می‌کند. آشکارساز SVDNet ادغام شبکه CNN با SVD را برای استخراج ویژگی‌های اساسی بر عهده دارد. روش‌های مبتنی بر شبکه پیشنهاددهنده ناحیه پیش از هر چیز، صحت را اولویت قرار می‌دهند. همان طور که بررسی شد، شبکه Faster R-CNN یک آشکارساز پیشنهاددهنده ناحیه و پرکاربرد است که ابتدا نواحی محتمل از هدف را به دست آورده و سپس هر ناحیه را طبقه‌بندی می‌کند. این رویکرد به قیمت سرعت آشکارسازی تمام می‌شود و تنها قادر است ۱/۶ فریم بر ثانیه را پردازش کند. به طور مشابه، SPP-PCANet یک شبکه پیشنهاددهنده ناحیه است که به صحت بالاتر ۷۶/۱ درصد، نسبت به آشکارساز Faster R-CNN می‌رسد. آشکارسازهای یک مرحله‌ای مستقیماً هم کلاس و هم موقعیت هدف را پیش‌بینی می‌کنند و تا حدودی صحت را برای افزایش سرعت قابل توجه، فدا می‌کنند. به عنوان مثال، شبکه SSD یک آشکارساز یک مرحله‌ای بوده و به طور قابل توجهی سریع‌تر از آشکارسازهای پیشنهاددهنده ناحیه و شبکه‌های متداول کانولوشن است. این آشکارساز با سرعت ۲۱/۹۶ فریم بر ثانیه و صحت ۵۱/۸۳ درصد به اوج سرعت پردازشی می‌رسد، در حالی که YOLO-v2 با ۲۳/۱۹ فریم بر ثانیه و صحت بسیار پایین ۳۱/۲۰ درصد کار می‌کند. اگرچه این سرعت‌ها بی‌شک چشمگیر هستند، اما صحت پایین‌تر نسبت به آشکارسازهای پیشنهاددهنده، نیاز به بررسی دقیق دارد. البته ساختارهای پیشرفته‌تر، مانند YOLO-v3 به صحت‌های بالایی دست یافته‌اند. به طور کلی،

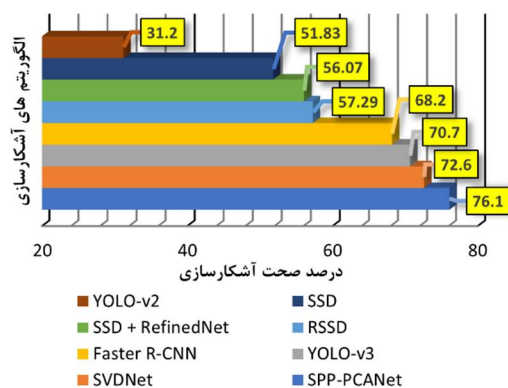
مجموعه‌های کلیدی تصویر ارائه شده است. با ادامه روند تکامل فناوری و افزایش تقاضا برای آشکارسازی دقیق و کارآمد اهداف، می‌توان انتظار داشت مجموعه‌های تصویر بیشتری در آینده ارائه شوند.

۴- مقایسه عملکرد روش‌های آشکارسازی

تعدادی از شبکه‌های عمیق کلیدی در آشکارسازی اهداف برای تصاویر سنجش از دور پنکروماتیک طبق جدول ۲ و شکل ۶ آورده و مقایسه شده‌اند.

جدول ۲- مقایسه شبکه‌های عمیق آشکارسازی هدف در تصاویر پنکروماتیک

روش	نوع	صحت آشکارسازی (mAP)	سرعت (فریم بر ثانیه)
SVDNet	شبکه عصبی کانولوشن	۷۲/۶	----
SPP-PCANet	پیشنهاددهنده ناحیه	۷۶/۱	----
Faster R-CNN	پیشنهاددهنده ناحیه	۶۸/۲	۱/۶
SSD	یک مرحله‌ای	۵۱/۸۳	۲۱/۹۶
SSD + RefinedNet	یک مرحله‌ای	۵۶/۰۷	۱۵/۸۲
RSSD	یک مرحله‌ای	۵۷/۲۹	۱۵/۷۶
YOLO-v2	یک مرحله‌ای	۳۱/۲۰	۲۳/۱۹
YOLO-v3	یک مرحله‌ای	۷۰/۷	----



شکل ۶- نمودار مقایسه صحت روش‌های آشکارسازی هدف در تصاویر پنکروماتیک

معیارهای ارزیابی صحت و سرعت آشکارسازی (تعداد فریم بر ثانیه) هر یک نیز ارائه شده است. در برخی موارد به دلیل عدم گزارش در مرجع مبوطه، سرعت آشکارساز ذکر نشده است. معیار متداول در بررسی صحت آشکارسازی، میانگین دقت متوسط (mAP)^۱ است که در زیر، رابطه آن

^۱ mean Average Precision

آشکارسازهای یک مرحله‌ای نسبت به روش‌های پیشنهاددهنده ناحیه سریع‌تر اما صحت کمتری دارند. مبادله بین سرعت، دقت و محدودیت‌های سخت‌افزاری، انتخاب آشکارساز مناسب را برای یک کاربری خاص تعیین می‌کند. به منظور بررسی بهتر روش‌های ذکر شده، نوع هدف و مجموعه تصاویر استفاده شده در آموزش و ارزیابی آشکارسازها در جدول ۳ آورده شده‌است. در ستون مجموعه تصاویر این جدول، سنجنده تصاویر پردازش شده نیز بیان شده‌است.

جدول ۳- نوع و مجموعه تصاویر استفاده شده در روش‌های آشکارسازی هدف

مجموعه تصاویر	نوع هدف	روش
DOTA (Gaofen 1) - NWPU VHR-10 (Gaofen 1)	کشتی	SVDNet
DOTA (Gaofen 1 & 2) - NWPU VHR-10 (Gaofen 1& 2)	کشتی	SPP-PCANet
DOTA (Gaofen 1 & 2) - NWPU VHR-10 (Gaofen 1& 2)	کشتی	Faster R-CNN
DOTA (Gaofen 1) - NWPU VHR-10 (Gaofen 1)	کشتی	YOLO-v3
DOTA (Gaofen 2) - NWPU VHR-10 (Gaofen 2)	هواپیما - کشتی	SSD
DOTA (Gaofen 2) - NWPU VHR-10 (Gaofen 2)	هواپیما - کشتی	SSD + RefinedNet
DOTA (Gaofen 2) - NWPU VHR-10 (Gaofen 2)	هواپیما - کشتی	RSSD
DOTA (Gaofen 2) - NWPU VHR-10 (Gaofen 2)	هواپیما - کشتی	YOLO-v2

طبق جدول‌های ۲ و ۳ از بین روش‌های آشکارسازی هدف کشتی، آشکارساز SPP-PCANet که یک الگوریتم پیشنهاددهنده ناحیه است، بالاترین میزان صحت را داشته‌است. دلیل صحت بالای این روش، استفاده از جمعیت هرم مکانی برای استخراج ویژگی، حذف هشدارهای کاذب و آشکارسازی چندمقیاسه است. در روش‌های آشکارسازی دو هدفی (کشتی-هواپیما)، روش RSSD رتبه اول را کسب کرده‌است. این روش، رویکرد یادگیری عمیق را با الگوریتم‌های سنتی تطبیق الگو ترکیب کرده‌است. الگوهای در نظر گرفته شده، برگرفته از تغییر شکل آماری اهداف و منطبق بر هندسه کشتی و هواپیما هستند. این روش با وجود صحت بالا، تنها نرخ آشکارسازی ۱۵/۷۶ فریم بر ثانیه را دارد.

۵- نتیجه‌گیری و رویکردهای آینده

در این تحقیق پیشرفت‌های اخیر در استفاده از روش‌های یادگیری عمیق برای آشکارسازی اهداف در تصاویر سنجش‌ازدور پنکروماتیک بررسی شد. آشکارسازی هدف در

این تصاویر شامل شناسایی و مکان‌یابی اهداف خاص مانند کشتی، هواپیما و غیره است. پیش‌پردازش تصاویر، استخراج ویژگی و الگوریتم‌های طبقه‌بندی از جمله مراحل کلیدی در این فرآیند هستند. توسعه معماری‌های پیچیده همراه با مجموعه‌های تصویر معیار، منجر به بهبودهای قابل توجهی در صحت، سرعت و کارایی شده‌است. از نقطه نظر معماری، آشکارسازهای دو مرحله‌ای طبقه‌بندی دقیق ناحیه محتمل را ارائه می‌کنند اما کندتر هستند، در حالی که مدل‌های یک مرحله‌ای سرعت تشخیص بی‌درنگ را امکان‌پذیر می‌کنند. سازوکارهای توجه به افزایش تمرکز بر مناطق هدف و بهبود تفکیک‌پذیری کمک کرده‌اند. شبکه‌های سبک‌وزن به محدودیت‌های محاسباتی سخت‌افزار پرداخته‌اند. ساختارهای پیچیده‌تر مانند شبکه‌های عصبی کانولوشن، در صورت آموزش مناسب، از صحت بالاتری برخوردارند اما سرعت پردازش کمتری دارند. ساختارهای سبک‌تر با استفاده از سازوکارهای توجه، سرعت پردازش بالاتری دارند اما به طور معمول دارای صحت پایین‌تری هستند. انتخاب ساختار مناسب بستگی به نیازها و محدودیت‌های سخت‌افزاری دارد. راهبردهای آموزشی دقیق مانند داده‌افزایی، یادگیری انتقالی، کاوش نمونه کم کیفیت و سازگاری مدل موجب بهینه‌سازی عملکرد و توانایی تعمیم می‌شوند. تقویت داده‌های آموزشی نقش اساسی در آموزش و بهبود عملکرد مدل‌های یادگیری عمیق ایفا می‌کنند. افزایش داده‌های آموزشی توسط روش‌های انتقال، گردش، چرخش و... موجب می‌شود تا مدل عمیق قابلیت تعمیم‌پذیری بالایی پیدا کند و بتواند در مواجهه با شرایط مختلف، نتایج صحیحی در آشکارسازی داشته باشد.

روش‌های پس‌پردازش مانند کاهش تأثیر غیرحداکثری، آستانه‌گذاری و فیلتر کردن، به اصلاح پیش‌بینی‌های خام و اولیه شبکه توجه دارند. از آن جایی که تصاویر پنکروماتیک فاقد اطلاعات طیفی هستند، شبکه‌های عمیق با استفاده از سایر تصاویر طیفی در یادگیری ویژگی‌های مکانی و ساختاری قدرتمند برای تبعیض هدف از پس‌زمینه مؤثر بوده‌اند. اطلاعات مکانی غنی در تصاویر پنکروماتیک می‌تواند در تفسیر و درک بهتر محتوای تصویر کمک کند. این اطلاعات شامل توپولوژی و روابط مکانی بین اهداف و پس‌زمینه است. استفاده از این اطلاعات در کنار داده‌های طیفی می‌تواند به بهبود نتایج آشکارسازی در سایر انواع تصاویر سنجش از دور کمک کند. به نظر می‌رسد رویکردهای تحقیق آینده شامل

حاصل شده‌است، فرصت‌های زیادی برای افزایش عملکرد، کاربرد و مقیاس‌پذیری از طریق نوآوری در معماری شبکه، رویکرد ادغام اطلاعات و روش‌های آموزش مؤثر وجود دارد. علاوه بر موارد ذکر شده، برای افزایش شفافیت و قابلیت اعتماد سیستم‌های آشکارسازی هدف، می‌توان از روش‌های هوش مصنوعی قابل تفسیر استفاده کرد. رایانش ابری نیز می‌تواند به افزایش قدرت محاسباتی مورد نیاز برای پردازش تصاویر پنکروماتیک با وضوح بالا کمک کند.

ادغام تصاویر پنکروماتیک با تصاویر دیگر مانند ابرطیفی برای تکمیل وضوح مکانی با امضاءهای طیفی اهداف باشد. سایر رویکردهای امیدوارکننده شامل روش‌های نیمه‌نظارتی و یادگیری با نمونه‌های محدود برای به حداقل رساندن الزامات جعبه‌های محدودکننده و انطباق مدل برای مدیریت جریان مداوم تصاویر جدید هستند. آشکارسازی دقیق در صحنه‌های متراکم و تشخیص اهداف بسیار کوچک همچنان چالش برانگیز است. با وجود این که پیشرفت‌های قابل توجهی

مراجع

- [1] W. Han *et al.*, "Methods for Small, Weak Object Detection in Optical High-Resolution Remote Sensing Images: A survey of advances and challenges," *IEEE Geoscience and Remote Sensing Magazine*, vol. 9, no. 4, pp. 8-34, 2021, doi: 10.1109/MGRS.2020.3041450.
- [2] S. J. Wang *et al.*, "Target detection of remote sensing images based on deep learning method and system," in *Proceedings of the 3rd International Conference on Advanced Information Science and System*, 2021, pp. 1-7.
- [3] F. Hou, Y. Zhang, Y. Zhou, M. Zhang, B. Lv, and J. Wu, "Review on infrared imaging technology," *Sustainability*, vol. 14, no. 18, p. 11161, 2022.
- [4] K. Xu, S. Wang, Y. Jin, Q. Che, and B. Zhou, "Object detection-oriented style transfer network for panchromatic remote sensing image," *Journal of Applied Remote Sensing*, vol. 17, no. 2, pp. 026503-026503, 2023.
- [5] P. Lu, Y. Ding, and C. Wang, "Multi-small target detection and tracking based on improved YOLO and SIFT for drones," *International Journal of Innovative Computing, Information and Control*, vol. 17, no. 1, pp. 205-224, 2021.
- [6] T. Nie, X. Han, B. He, X. Li, H. Liu, and G. Bi, "Ship detection in panchromatic optical remote sensing images based on visual saliency and multi-dimensional feature description," *Remote Sensing*, vol. 12, no. 1, p. 152, 2020.
- [7] L. Bo, X. Xiaoyang, W. Xingxing, and T. Wenting, "Ship detection and classification from optical remote sensing images: A survey," *Chinese Journal of Aeronautics*, vol. 34, no. 3, pp. 145-163, 2021.
- [8] M. J. Khan, A. Yousaf, N. Javed, S. Nadeem, and K. Khurshid, "Automatic target detection in satellite images using deep learning," *Journal of Space Technology*, vol. 7, no. 1, pp. 44-49, 2017.
- [9] F. Huang, Y. Yu, and T. Feng, "Automatic extraction of impervious surfaces from high resolution remote sensing images based on deep learning," *Journal of Visual Communication and Image Representation*, vol. 58, pp. 453-461, 2019.
- [10] B. Hou, Z. Ren, W. Zhao, Q. Wu, and L. Jiao, "Object detection in high-resolution panchromatic images using deep models and spatial template matching," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 2, pp. 956-970, 2019.
- [11] S. Xiong, Y. Tan, Y. Li, C. Wen, and P. Yan, "Subtask attention based object detection in remote sensing images," *Remote Sensing*, vol. 13, no. 10, p. 1925, 2021.
- [12] Z. Zakria, J. Deng, R. Kumar, M. S. Khokhar, J. Cai, and J. Kumar, "Multiscale and direction target detecting in remote sensing images via modified YOLO-v4," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 1039-1048, 2022.
- [13] J. Chen, J. Sun, Y. Li, and C. Hou, "Object detection in remote sensing images based on deep transfer learning," *Multimedia Tools and Applications*, pp. 1-17, 2022.
- [14] D. Yan *et al.*, "Improved method to detect the tailings ponds from multispectral remote sensing images based on faster R-CNN and transfer learning," *Remote Sensing*, vol. 14, no. 1, p. 103, 2021.
- [15] Y. Zhang, K. Fu, H. Sun, X. Sun, X. Zheng, and H. Wang, "A multi-model ensemble method based on convolutional neural networks for aircraft detection in large remote sensing images," *Remote Sensing Letters*, vol. 9, no. 1, pp. 11-20, 2018.
- [16] X. Ji, L. Tang, T. Lu, and C. Cai, "DBENet: Dual-Branch Ensemble Network for Sea-Land Segmentation of Remote Sensing Images," *IEEE Transactions on Instrumentation and Measurement*, 2023.

- [17] Y. Chen, L. Tang, X. Yang, M. Bilal, and Q. Li, "Object-based multi-modal convolution neural networks for building extraction using panchromatic and multispectral imagery," *Neurocomputing*, vol. 386, pp. 136-146, 2020.
- [18] V. Tarverdiyev, I. Erer, N. H. Kaplan, and N. Musaoglu, "Target Detection in Multispectral Images via Detail Enhanced Pansharpening," in *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium, 2022*: IEEE, pp. 1544-1547.
- [19] E. R. Dougherty, *Digital image processing methods*. CRC Press, 2020.
- [20] Q. Wang *et al.*, "Simultaneous extracting area and quantity of agricultural greenhouses in large scale with deep learning method and high-resolution remote sensing images," *Science of The Total Environment*, vol. 872, p. 162229, 2023.
- [21] M. E. Aghili, M. Imani, and H. Ghassemian, "Clustering based background learning for hyperspectral anomaly detection," *The Egyptian Journal of Remote Sensing and Space Sciences*, vol. 26, no. 3, pp. 477-489, 2023/12/01/2023, doi: <https://doi.org/10.1016/j.ejrs.2023.06.001>.
- [22] Q. Liu, Y. Tian, L. Zhang, and B. Chen, "Urban Surface Water Mapping from VHR Images Based on Superpixel Segmentation and Target Detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 5339-5356, 2022.
- [23] Q. Yuan, Y. Wei, X. Meng, H. Shen, and L. Zhang, "A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 3, pp. 978-989, 2018.
- [24] S. Song *et al.*, "Intelligent object recognition of urban water bodies based on deep learning for multi-source and multi-temporal high spatial resolution remote sensing imagery," *Sensors*, vol. 20, no. 2, p. 397, 2020.
- [25] Y. Long, Y. Gong, Z. Xiao, and Q. Liu, "Accurate object localization in remote sensing images based on convolutional neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 5, pp. 2486-2498, 2017.
- [26] S. Ghaffarian, J. Valente, M. Van Der Voort, and B. Tekinerdogan, "Effect of attention mechanism in deep learning-based remote sensing image processing: A systematic literature review," *Remote Sensing*, vol. 13, no. 15, p. 2965, 2021.
- [27] J. Astola and P. Kuosmanen, *Fundamentals of nonlinear digital filtering*. CRC press, 2020.
- [28] Y. Rao, W. Zhao, Z. Zhu, J. Lu, and J. Zhou, "Global filter networks for image classification," *Advances in neural information processing systems*, vol. 34, pp. 980-993, 2021.
- [29] Y. Shi, H. Cui, Y. Yin, H. Song, Y. Li, and P. Gamba, "Transfer Learning With Nonlinear Spectral Synthesis for Hyperspectral Target Detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-17, 2023.
- [30] S. K. Köse, S. Ergünay, B. Ott, P. Wellig, and Y. Leblebici, "Target detection with deep learning in polarimetric imaging," in *Target and Background Signatures IV*, 2018, vol. 10794: SPIE, pp. 212-220.
- [31] P. Ghamisi *et al.*, "New frontiers in spectral-spatial hyperspectral image classification: The latest advances based on mathematical morphology, Markov random fields, segmentation, sparse representation, and deep learning," *IEEE geoscience and remote sensing magazine*, vol. 6, no. 3, pp. 10-43, 2018.
- [32] Y. Liu, M. Chang, and J. Xu, "High-resolution remote sensing image information extraction and target recognition based on multiple information fusion," *IEEE access*, vol. 8, pp. 121486-121500, 2020.
- [33] K. Yu *et al.*, "A Parallel Algorithm for Hyperspectral Target Detection Based on Weighted Alternating Direction Method of Multiplier," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023.
- [34] Y. Zhang and G. Hong, "A wavelet integrated image fusion approach for target detection in very high resolution satellite imagery," in *Signal Processing, Sensor Fusion, and Target Recognition XIV*, 2005, vol. 5809: SPIE, pp. 330-340.
- [35] 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*, vol. 9, no. 10, p. 2110, 2019.
- [36] A. Kaul and M. Kumari, "A literature review on remote sensing scene categorization based on convolutional neural networks," *International Journal of Remote Sensing*, vol. 44, no. 8, pp. 2611-2642, 2023.
- [37] Z. Dong, M. Wang, Y. Wang, Y. Zhu, and Z. Zhang, "Object detection in high resolution remote sensing imagery based on convolutional neural networks with suitable object scale features," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 3, pp. 2104-2114, 2019.

- [۳۸] نیما فرهادی، عباس کیانی، و حمید عبادی، "توسعه مدلی مبتنی بر تشدید گرادیان در شبکه‌های کانولوشنی عمیق به منظور شناسایی اهداف در تصاویر سنجش ازدوری"، *نشریه علمی علوم و فنون نقشه برداری*، دوره ۱۱، شماره ۱، صفحات ۳۵-۵۰، ۱۴۰۰.
- [39] R. P. Loibl, "Target Detection using Convolutional Neural Networks," 2018.
- [40] M. Dahmane, S. Foucher, M. Beaulieu, F. Riendeau, Y. Bouroubi, and M. Benoit, "Object detection in pleiades images using deep features," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2016: IEEE, pp. 1552-1555.
- [41] X. Zhang, Y. n. Zhou, and J. Luo, "Deep learning for processing and analysis of remote sensing big data: A technical review," *Big Earth Data*, vol. 6, no. 4, pp. 527-560, 2022.
- [42] Y. Jia *et al.*, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the 22nd ACM international conference on Multimedia*, 2014, pp. 675-678.
- [43] J. Deng, W. Dong, R. Socher, L. J. Li, L. Kai, and F.-F. Li, "ImageNet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 20-25 June 2009 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.
- [44] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge: Cambridge University Press, 2000.
- [45] C. L. Zitnick and P. Dollár, "Edge boxes: Locating object proposals from edges," in *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V* 13, 2014: Springer, pp. 391-405.
- [46] W. Yuan, W. Ran, X. Shi, and R. Shibasaki, "Multi-Constraint Transformer based Automatic Building Extraction from High Resolution Remote Sensing Images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023.
- [47] B. Kalantar, S. B. Mansor, A. A. Halin, H. Z. M. Shafri, and M. Zand, "Multiple moving object detection from UAV videos using trajectories of matched regional adjacency graphs," *IEEE Transactions on geoscience and remote sensing*, vol. 55, no. 9, pp. 5198-5213, 2017.
- [48] Y. Tan, S. Xiong, and Y. Li, "Automatic extraction of built-up areas from panchromatic and multispectral remote sensing images using double-stream deep convolutional neural networks," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 11, pp. 3988-4004, 2018.
- [49] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818-2826.
- [50] J. Wei, Y. Liu, L. Li, W. Xie, S. Zhao, and Z. Zhao, "Improved YOLO X with Bilateral Attention for Small Object Detection," in *2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC)*, 2023: IEEE, pp. 1-6.
- [51] X. Ji, L. Huang, B.-H. Tang, G. Chen, and F. Cheng, "A Superpixel Spatial Intuitionistic Fuzzy C-Means Clustering Algorithm for Unsupervised Classification of High Spatial Resolution Remote Sensing Images," *Remote Sensing*, vol. 14, no. 14, p. 3490, 2022.
- [52] L. Shuxin, Z. Zhilong, and L. Biao, "A plane target detection algorithm in remote sensing images based on deep learning network technology," in *Journal of Physics: Conference Series*, 2018, vol. 960, no. 1: IOP Publishing, p. 012025.
- [53] G. Prathap and I. Afanasyev, "Deep learning approach for building detection in satellite multispectral imagery," in *2018 International Conference on Intelligent Systems (IS)*, 2018: IEEE, pp. 461-465.
- [54] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III* 18, 2015: Springer, pp. 234-241.
- [55] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC Superpixels Compared to State-of-the-Art Superpixel Methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274-2282, 2012, doi: 10.1109/TPAMI.2012.120.
- [56] S. Crommelinck, R. Bennett, M. Gerke, M. Koeva, M. Yang, and G. Vosselman, "SLIC superpixels for object delineation from UAV data," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, pp. 9-16, 2017.
- [57] L. Zhang *et al.*, "Structure-feature based graph self-adaptive pooling," in *Proceedings of The Web Conference 2020*, 2020, pp. 3098-3104.

- [58] Y. Zhong, Z. Zheng, A. Ma, X. Lu, and L. Zhang, "COLOR: cycling, offline learning, and online representation framework for airport and airplane detection using GF-2 satellite images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 12, pp. 8438-8449, 2020.
- [59] نیما فرهادی، عباس کیانی، و حمید عبادی، "شناسایی اهداف در تصاویر سنجش از دوری با قدرت تفکیک بالا با استفاده از روش‌های یادگیری عمیق"، *سنجش از دور و GIS ایران*، دوره ۱۱، شماره ۱، صفحات ۳۳-۴۸، ۱۳۹۸.
- [60] Q. Zheng, L. Zheng, Y. Bai, H. Liu, J. Deng, and Y. Li, "Boundary-aware network with two-stage partial decoders for salient object detection in remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-13, 2023.
- [61] Q. Lin, L. Xia, S. Li, and W. Chen, "Two-stage local attention network for salient object detection in remote sensing images," *IET Image Processing*, vol. 17, no. 3, pp. 849-861, 2023.
- [62] P. Bharati and A. Pramanik, "Deep learning techniques—R-CNN to mask R-CNN: a survey," *Computational Intelligence in Pattern Recognition: Proceedings of CIPR 2019*, pp. 657-668, 2020.
- [63] Z. Zou and Z. Shi, "Ship Detection in Spaceborne Optical Image With SVD Networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 5832-5845, 2016, doi: 10.1109/TGRS.2016.2572736.
- [64] K. Kanatani, "Singular Value Decomposition," in *Computer Vision: A Reference Guide*: Springer, 2021, pp. 1174-1177.
- [65] Q. Li, L. Mou, Q. Liu, Y. Wang, and X. X. Zhu, "HSF-Net: Multiscale Deep Feature Embedding for Ship Detection in Optical Remote Sensing Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 12, pp. 7147-7161, 2018, doi: 10.1109/TGRS.2018.2848901.
- [66] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, vol. 28, 2015.
- [67] N. Wang, B. Li, Q. Xu, and Y. Wang, "Automatic ship detection in optical remote sensing images based on anomaly detection and SPP-PCANet," *Remote Sensing*, vol. 11, no. 1, p. 47, 2018.
- [68] Z. Lu, T. Xu, K. Liu, Z. Liu, F. Zhou, and Q. Liu, "5M-Building: A large-scale high-resolution building dataset with CNN based detection analysis," in *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, 2019: IEEE, pp. 1385-1389.
- [69] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961-2969.
- [70] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.
- [71] G. Othman and D. Q. Zeebaree, "The applications of discrete wavelet transform in image processing: A review," *Journal of Soft Computing and Data Mining*, vol. 1, no. 2, pp. 31-43, 2020.
- [72] H. Wang and F. Miao, "Building extraction from remote sensing images using deep residual U-Net," *European Journal of Remote Sensing*, vol. 55, no. 1, pp. 71-85, 2022.
- [73] L. Chen, W. Shi, C. Fan, L. Zou, and D. Deng, "A novel coarse-to-fine method of ship detection in optical remote sensing images based on a deep residual dense network," *Remote Sensing*, vol. 12, no. 19, p. 3115, 2020.
- [74] F. Chen, S. Lou, and Y. Song, "Improving Object Detection of Remotely Sensed Multispectral Imagery Via Pan-sharpening," in *Proceedings of the 2020 9th International Conference on Computing and Pattern Recognition*, 2020, pp. 136-140.
- [75] J. Choi, K. Yu, and Y. Kim, "A new adaptive component-substitution-based satellite image fusion by using partial replacement," *IEEE transactions on geoscience and remote sensing*, vol. 49, no. 1, pp. 295-309, 2010.
- [76] A. Sekrecka, M. Kedzierski, and D. Wierzbicki, "Pre-processing of panchromatic images to improve object detection in pansharpened images," *Sensors*, vol. 19, no. 23, p. 5146, 2019.
- [77] J. Liu, "Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details," *International Journal of Remote Sensing*, vol. 21, no. 18, pp. 3461-3472, 2000.
- [78] P. Chavez, S. C. Sides, and J. A. Anderson, "Comparison of three different methods to merge multiresolution and multispectral data- Landsat TM and SPOT panchromatic," *Photogrammetric Engineering and remote sensing*, vol. 57, no. 3, pp. 295-303, 1991.
- [79] A. Garzelli, L. Capobianco, and F. Nencini, "On the effects of pan-sharpening to target detection," in *2009 IEEE international geoscience and remote sensing symposium*, 2009, vol. 2: IEEE, pp. II-136-II-139.
- [80] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, "A new look at IHS-like image fusion methods," *Information fusion*, vol. 2, no. 3, pp. 177-186, 2001.

- [81] X. Otazu, M. González-Audícana, O. Fors, and J. Núñez, "Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 10, pp. 2376-2385, 2005.
- [82] A. Garzelli, F. Nencini, and L. Capobianco, "Optimal MMSE Pan Sharpening of Very High Resolution Multispectral Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 1, pp. 228-236, 2008, doi: 10.1109/TGRS.2007.907604.
- [83] Q. Tan, J. Ling, J. Hu, X. Qin, and J. Hu, "Vehicle Detection in High Resolution Satellite Remote Sensing Images Based on Deep Learning," *IEEE Access*, vol. 8, pp. 153394-153402, 2020, doi: 10.1109/ACCESS.2020.3017894.
- [84] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.
- [85] X. Nie, M. Duan, H. Ding, B. Hu, and E. K. Wong, "Attention mask R-CNN for ship detection and segmentation from remote sensing images," *Ieee Access*, vol. 8, pp. 9325-9334, 2020.
- [86] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2117-2125.
- [87] Z. Li, E. Li, T. Xu, A. Samat, and W. Liu, "Feature alignment fpn for oriented object detection in remote sensing images," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1-5, 2023.
- [88] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.
- [89] Y. Ma, S. Chen, S. Ermon, and D. B. Lobell, "Transfer learning in environmental remote sensing," *Remote Sensing of Environment*, vol. 301, p. 113924, 2024.
- [90] F. Huang, L. Xu, M. Li, and M. Tang, "High-resolution remotely sensed small target detection by imitating fly visual perception mechanism," *Computational and mathematical methods in medicine*, vol. 2012, 2012.
- [91] Q. Cheng, S. Zhang, S. Bo, D. Chen, and H. Zhang, "Augmented reality dynamic image recognition technology based on deep learning algorithm," *IEEE Access*, vol. 8, pp. 137370-137384, 2020.
- [92] L. Zhang, K. Yang, and H. Li, "Regions of interest detection in panchromatic remote sensing images based on multiscale feature fusion," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 12, pp. 4704-4716, 2014.
- [93] S. Karim, Y. Zhang, S. Yin, I. Bibi, and A. A. Brohi, "A brief review and challenges of object detection in optical remote sensing imagery," *Multiagent and Grid Systems*, vol. 16, no. 3, pp. 227-243, 2020.
- [94] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 20, no. 11, pp. 1254-1259, 1998.
- [95] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," *Advances in neural information processing systems*, vol. 19, 2006.
- [96] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *2007 IEEE Conference on computer vision and pattern recognition*, 2007: Ieee, pp. 1-8.
- [97] S. Gui, S. Song, R. Qin, and Y. Tang, "Remote Sensing Object Detection in the Deep Learning Era—A Review," *Remote Sensing*, vol. 16, no. 2, p. 327, 2024.
- [98] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.
- [99] Z. Huang *et al.*, "Building Detection From Panchromatic and Multispectral Images With Dual-Stream Asymmetric Fusion Networks," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 3364-3377, 2023.
- [100] Z. Zhao *et al.*, "Lightweight Target Detection in High Resolution Remote Sensing Images," in *International Conference on Autonomous Unmanned Systems*, 2022: Springer, pp. 3252-3260.
- [101] C. Corbane, F. Marre, and M. Petit, "Using SPOT-5 HRG Data in Panchromatic Mode for Operational Detection of Small Ships in Tropical Area," *Sensors*, vol. 8, no. 5, pp. 2959-2973, 2008.
- [102] J. Tang, C. Deng, G.-B. Huang, and B. Zhao, "Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine," *IEEE transactions on geoscience and remote sensing*, vol. 53, no. 3, pp. 1174-1185, 2014.

- [103] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489-501, 2006.
- [104] N. Wang, B. Li, X. Wei, Y. Wang, and H. Yan, "Ship detection in spaceborne infrared image based on lightweight CNN and multisource feature cascade decision," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 5, pp. 4324-4339, 2020.
- [105] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.
- [106] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of big data*, vol. 6, no. 1, pp. 1-48, 2019.
- [107] U. Muhammad, W. Wang, S. P. Chattha, and S. Ali, "Pre-trained VGGNet architecture for remote-sensing image scene classification," in *2018 24th International Conference on Pattern Recognition (ICPR)*, 2018: IEEE, pp. 1622-1627.
- [108] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," in *Artificial Neural Networks and Machine Learning-ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings, Part III 27*, 2018: Springer, pp. 270-279.
- [109] L. Shen, Y. Sun, Z. Yu, L. Ding, X. Tian, and D. Tao, "On Efficient Training of Large-Scale Deep Learning Models: A Literature Review," *arXiv preprint arXiv:2304.03589*, 2023.
- [110] W. Yu, J. Li, Z. Wang, and Z. Yu, "Boosting SAR Aircraft Detection Performance with Multi-Stage Domain Adaptation Training," *Remote Sensing*, vol. 15, no. 18, p. 4614, 2023.
- [111] Y. Li, J. Xue, and K. Lu, "Semi-supervised Object Detection for Remote Sensing Images via Pseudo Labeling and Consistency Learning: RS-PCL for SSOD in Remote Sensing Images," in *Proceedings of the 15th International Conference on Digital Image Processing*, 2023, pp. 1-7.
- [112] D. Wittich, "Deep domain adaptation by weighted entropy minimization for the classification of aerial images," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*; 5, 2, vol. 5, no. 2, pp. 591-598, 2020.
- [113] S. Qiu, G. Wen, Z. Deng, J. Liu, and Y. Fan, "Accurate non-maximum suppression for object detection in high-resolution remote sensing images," *Remote Sensing Letters*, vol. 9, no. 3, pp. 237-246, 2018.
- [114] H. Choi, H.-J. Lee, H.-J. You, S.-Y. Rhee, and W.-S. Jeon, "Comparative analysis of generalized intersection over union," *Sensors and Materials*, vol. 31, no. 11, pp. 3849-3858, 2019.
- [115] G.-S. Xia *et al.*, "DOTA: A large-scale dataset for object detection in aerial images," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 3974-3983.
- [116] B. Cheng, Z. Li, B. Xu, X. Yao, Z. Ding, and T. Qin, "Structured object-level relational reasoning CNN-based target detection algorithm in a remote sensing image," *Remote Sensing*, vol. 13, no. 2, p. 281, 2021.
- [117] A. Van Etten, D. Lindenbaum, and T. M. Bacastow, "Spacenet: A remote sensing dataset and challenge series," *arXiv preprint arXiv:1807.01232*, 2018.
- [118] M.-R. Hsieh, Y.-L. Lin, and W. H. Hsu, "Drone-based object counting by spatially regularized regional proposal network," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 4145-4153.