

# مقایسه شبکه‌های عمیق Faster RCNN و RetinaNet

## جهت تشخیص خودرو در آب‌وهوای نامساعد

یاسر جمشیدی و راضیه سادات اخوت

شرایط آب‌وهوایی متفاوت چالش‌برانگیزتر از تشخیص وسایل نقلیه در حالت عادیست. در این مقاله الگوریتمی خودکار مبتنی بر روش یادگیری عمیق برای تشخیص و ردیابی وسایل نقلیه در شرایط آب‌وهوایی متغیر ارائه خواهد شد. از مهم‌ترین مزایای یادگیری عمیق می‌توان به یادگیری خودکار ویژگی‌ها، یادگیری چندلایه ویژگی‌ها، قدرت تعمیم بالا در شناسایی داده‌های جدید و در نتیجه دقت بالا در نتایج اشاره کرد.

هدف در این مقاله، تشخیص خودرو در شرایط نامساعد آب‌وهوایی است که به‌وسیله آموزش دو شبکه Faster RCNN و RetinaNet انجام می‌شود. مروری بر ادبیات این موضوع در بخش دوم ارائه می‌گردد. در بخش سوم به نحوه عملکرد سیستم تشخیص اشیا و روش‌های متعدد تشخیص به‌طور خلاصه اشاره شده است. معرفی پایگاه داده در بخش چهارم انجام شده و بخش پنجم به توضیح معیار ارزیابی و نحوه عملکرد آن می‌پردازد. نهایتاً در بخش ششم با آزمایش دو شبکه نام‌برده بر روی پایگاه داده وسایل نقلیه در آب‌وهوای متغیر به بررسی نتایج در تشخیص خودرو پرداخته شده است.

چکیده: تشخیص وسایل نقلیه و ردیابی آن، نقش مهمی در اتومبیل‌های خودران و سیستم‌های حمل‌ونقل هوشمند ایفا می‌کند. شرایط آب‌وهوایی نامساعد مانند حضور برف سنگین، مه، باران و گرد و غبار با کاهش دید دوربین، محدودیت‌های خطرناکی ایجاد کرده و بر عملکرد الگوریتم‌های تشخیصی استفاده‌شده در سیستم‌های نظارت بر ترافیک و برنامه‌های رانندگی خودکار تأثیر می‌گذارد. در این مقاله از شبکه عمیق تشخیص اشیا Faster RCNN با هسته ResNet50 و شبکه RetinaNet استفاده شده و دقت این دو شبکه جهت تشخیص خودرو در آب‌وهوای نامساعد مورد بررسی قرار می‌گیرد. پایگاه داده مورد استفاده، فایل DAWN می‌باشد که شامل تصاویر دنیای واقعی است و با انواع مختلفی از شرایط آب‌وهوایی نامطلوب جمع‌آوری شده‌اند. نتایج به‌دست‌آمده نشان می‌دهند که روش ارائه‌شده در بهترین حالت، دقت تشخیص را از ۰.۳٪ به ۷۵٪ افزایش داده و بیشترین میزان افزایش دقت نیز مربوط به شرایط بارانی می‌باشد. تمام پردازش‌ها به زبان پایتون و در گوگل کولب انجام شده است.

کلیدواژه: تشخیص شیء، تشخیص خودرو، یادگیری عمیق، سیستم‌های حمل‌ونقل هوشمند، پردازش تصویر در آب‌وهوای نامساعد.

## ۲- مروری بر کارهای گذشته

تشخیص خودرو، یکی از موضوعاتی است که به‌دلیل اهمیت آن، تاکنون زیاد بر روی آن کار شده است. به طور مثال Hassaballah و همکاران [۲] یک روش تشخیص وسیله نقلیه با چارچوب الگوی باینری محلی را پیشنهاد کرده‌اند. در واقع این افراد یک روش ساده و مؤثر تشخیص وسیله نقلیه را بر اساس ویژگی‌ها و هیستوگرام محلی خودرو پیشنهاد داده‌اند. Kumar Bhaskar و همکاران [۳]، یک الگوریتم منحصربه‌فرد را برای تشخیص وسیله نقلیه و ردیابی با استفاده از مدل مخلوط گوسی و روش‌های تشخیص شیء ایجاد کرده‌اند. مدل گوسی، وظیفه جداکردن پیش‌زمینه و پس‌زمینه از فریم تصویر با یادگیری پس‌زمینه یک صحنه را انجام می‌دهد. Wenqing Chu و همکاران [۴] نیز یک طرح جدید تشخیص وسیله نقلیه مبتنی بر شبکه‌های عصبی عمیق چندمنظوره و جداسازی منطقه مورد نظر را پیشنهاد داده‌اند. Luyang Zhang و همکاران [۵] با توجه به راندمان پایین تشخیص وسیله نقلیه در صحنه‌های واقعی، یک شبکه RetinaNet بهبودیافته را پیشنهاد می‌کنند. آنها یک ساختار پیچشی اکتاو و یک ساختار هرم وزنی جهت بهبود عملکرد RetinaNet برای تشخیص وسایل نقلیه معرفی کرده‌اند. از کانولوشن اکتاو به‌جای لایه کانولوشن سنتی استفاده کرده و به‌منظور بهبود کیفیت همجوشی ویژگی‌ها، یک ساختار شبکه هرمی ویژگی وزنی (WFPN)<sup>۱</sup> برای محدودکردن انتشار گرادینان‌ها بین سطوح مختلف

## ۱- مقدمه

پیش‌بینی مکان یا ردیابی وسایل نقلیه، موضوعی مهم در جاده‌هاست و از این رو تحقیقات زیادی برای ردیابی خودرو صورت گرفته که هم در شرایط عادی و هم در شرایط غیرعادی می‌توان از آنها و یا از ترکیب آنها استفاده نمود [۱]. کارایی تشخیص وسایل نقلیه به‌عنوان یک گام مهم در نظارت ترافیک یا نظارت هوشمند در نظر گرفته می‌شود. سیستم‌های تشخیص وسایل نقلیه باید دقیقاً اهداف ترافیک (مثلاً ماشین‌ها، موتورها و ...) را در کمترین زمان و با بالاترین دقت شناسایی کنند. برای تشخیص این اشیا، انواع مختلفی از حسگرها مثل دوربین و تشخیص نور در وسایل نقلیه مستقل به‌کار گرفته می‌شوند. در بین این انواع، کیفیت تصاویر دوربین کاملاً تحت تأثیر شرایط آب‌وهوایی مانند بارندگی سنگین، بارش برف سنگین، طوفان‌های برفی، گرد و غبار و شرایط نوری پایین قرار دارد. در چنین شرایطی، دید کافی برای دوربین‌های سیستم تشخیص وسایل نقلیه در جاده‌ها وجود نخواهد داشت؛ بنابراین تشخیص وسایل نقلیه در

این مقاله در تاریخ ۳ آبان ماه ۱۴۰۱ دریافت و در تاریخ ۱۰ آبان ماه ۱۴۰۲ بازنگری شد.

یاسر جمشیدی، گروه مهندسی برق، دانشکده فنی مهندسی، دانشگاه علم و فرهنگ، تهران، ایران، (email: yass2j@gmail.com).

راضیه سادات اخوت (نویسنده مسئول)، گروه مهندسی برق، دانشکده فنی مهندسی، دانشگاه علم و فرهنگ، تهران، ایران، (email: r.s.okhovat@usc.ac.ir).

درباره دلیل انتخاب این دو شبکه و همچنین نتایج به دست آمده توسط آنها بیشتر توضیح داده می شود.

### ۳- تشخیص اشیا

#### ۳-۱ یادگیری عمیق

یادگیری عمیق، زیرشاخه‌ای از یادگیری ماشین است که از لایه‌های متعدد تبدیلات خطی به منظور پردازش سیگنال‌های حسی مانند صدا و تصویر استفاده می کند [۱۸]. ماشین در این روش هر مفهوم پیچیده را به مفاهیم ساده‌تری تقسیم می کند و با ادامه این روند به مفاهیم پایه‌ای می رسد که قادر به تصمیم‌گیری برای آنهاست و بدین ترتیب، نیازی به نظارت کامل انسان برای مشخص کردن اطلاعات لازم ماشین در هر لحظه نیست. موضوعی که در یادگیری عمیق اهمیت زیادی دارد، نحوه ارائه اطلاعات است. ارائه دادن اطلاعات به ماشین باید به شیوه‌ای باشد که ماشین در کمترین زمان، اطلاعات کلیدی را که می تواند با استناد به آنها تصمیم بگیرد دریافت کند. یادگیری عمیق اغلب شامل یادگیری سلسله‌مراتبی است [۱۹]. شبکه عمیق نوعی شبکه عصبی مصنوعی است که با استفاده از تکنیک‌های ریاضی، ساختاری مشابه با ساختار مغز انسان به ما می دهد. هرچه تعداد نورون‌های مصنوعی در لایه پنهان بیشتر باشند مدل پیچیده‌تر می شود و به شبکه‌های عصبی که تعداد لایه‌های پنهان آنها زیاد است، شبکه عصبی عمیق گفته می شود. به وسیله شبکه‌های عصبی عمیق، مسائل بسیار پیچیده در زمینه پیش‌بینی و دسته‌بندی، ساده حل می شود و از طرفی پیشرفت تکنولوژی باعث شده که الگوریتم‌هایی برای بهینه‌سازی شبکه‌های عصبی معمولی به وجود آید تا بتواند تعداد لایه‌های نورونی شبکه‌های عصبی را از چند لایه به هزاران لایه برساند. شبکه عصبی عمیق، ارتباط داده‌های ورودی و خروجی را پیدا می کند. الگوریتم‌های یادگیری عمیق با معرفی راه‌حلی مؤثر و کارآمد، انقلابی را در جامعه بینایی کامپیوتر به ارمغان آورده است.

#### ۳-۲ انواع روش‌های تشخیص اشیا

تشخیص اشیا، روشی خودکار برای تعیین مکان اشیا مورد نظر در یک تصویر با توجه به پس‌زمینه است. در مدل‌های تشخیص شیء مبتنی بر یادگیری عمیق، معمولاً دو روش وجود دارد: الگوریتم‌های دسته اول، ابتدا بررسی منطقه‌ای انجام می دهند؛ بدین معنا که مناطقی که احتمال وجود اشیا در آنها بیشتر است انتخاب می شوند. این انتخاب یا توسط روش‌های قدیمی بینایی ماشین (مثل جستجوی انتخابی) انجام می شود یا توسط شبکه پیشنهادی منطقه‌ای بر مبنای یادگیری عمیق کار می کند. پس از به دست آمدن گروهی کوچک از پنجره‌ها، تشخیص اشیا به کمک معادله‌های متشکل از چند مدل رگرسیون و چند مدل طبقه‌بندی، انجام می شود. این کار شامل الگوریتم‌هایی از قبیل روش Faster RCNN [۲۰]، Fast RCNN [۲۱] و ... است. الگوریتم‌هایی که در این دسته قرار دارند معمولاً روش دوم‌مرحله‌ای نامیده می شوند. الگوریتم‌های دسته دوم تنها در مکان‌ها و اندازه‌های مشخص دنبال اجسام می گردند. این مکان‌ها و اندازه‌ها به گونه استراتژیکی انتخاب می شوند تا بیشتر حالات ممکن را پوشش دهند. الگوریتم‌های این گروه معمولاً عکس را به چند بخش با اندازه مشخص تقسیم می کنند. سپس در نظر می گیرند که در هر بخش، تعداد مشخصی اجسام با اشکال و اندازه‌های از پیش تعیین شده وجود دارد. الگوریتم‌های این دسته روش تک‌مرحله‌ای نامیده می شوند که YOLO [۲۲]، SSD [۲۳] و RetinaNet [۲۴] از مثال‌های این دسته هستند.

پیشنهاد داده‌اند. Ju-Seok Shin و همکاران [۶] یک سیستم تشخیص خودرو در زمان واقعی با استفاده از طرح یادگیری عمیق برای کاهش نرخ مثبت کاذب پیشنهاد کرده‌اند. علاوه بر این، الگوریتم خود را در یک سیستم تعبیه شده برای تأیید زمان واقعی پیاده‌سازی کرده‌اند. نتایج تجربی نشان می دهند که میانگین زمان پردازش مازول‌های تشخیص خودرو در این روش حدود ۱۵ فریم در ثانیه است. Hai Wang و همکاران در [۷]، پنج الگوریتم اصلی یادگیری عمیق جهت تشخیص اشیا یعنی YOLOv3، RetinaNet، SSD، R-FCN، Faster RCNN را بر روی داده‌های خودرویی KITTI تست کرده و نتایج به دست آمده را از نظر زمان و دقت تشخیص تجزیه و تحلیل کرده‌اند. در سال ۲۰۱۷، Hulin Kuang و همکاران [۸]، یک سیستم تشخیص وسیله نقلیه مؤثر در شب با ترکیب استخراج منطقه مورد نظر و بهبود تصویر را برای تشخیص خودرو در شب ارائه داده‌اند که برای تشخیص وسایل نقلیه بیشتر به چراغ‌های قرمز عقب خودروها متمرکز شده‌اند. Sakaridis و همکاران [۹] یک مدل مبتنی بر شبکه عصبی (CNN) را برای ایجاد مه مصنوعی بر روی تصاویر وسایل نقلیه واقعی برای بررسی الگوریتم‌های مه‌زدایی در محیط‌های ترافیکی پیشنهاد کردند. مه، دید یک صحنه را به میزان قابل توجهی کاهش می دهد. Hodges و همکاران [۱۰] مدلی با یک شبکه مه‌زدایی برای اصلاح تصاویر کامل و یک شبکه جداکننده برای بهبود کارایی پارامترهای وزنی را برای افزایش عملکرد تشخیص وسیله نقلیه بر روی مجموعه داده تصاویر مه‌آلود و مه‌آلود مصنوعی پیشنهاد نموده‌اند. معماری شبکه مه‌سکن بر اساس معماری DehazeNet [۱۱] می باشد. Siyuan Li و همکاران [۱۲] در سال ۲۰۱۹ یک الگوریتم حذف باران از تصویر با معیار جدید را در مقیاس بزرگ، شامل تصاویر بارانی مصنوعی و واقعی از انواع مختلف باران در صحنه کنترل ترافیک و تشخیص خودرو ارائه کرده‌اند. برای ارزیابی بهتر در این مقاله از انواع باران‌های مصنوعی بر روی تصاویر برای آموزش و ارزیابی شبکه استفاده گردیده و بعد از آن از یک مرحله پیش‌پردازش برای حذف انواع باران از تصاویر استفاده می شود. به این صورت که یک تصویر ورودی را به اجزای فرکانس پایین و بالا تجزیه کرده و سپس فرکانس‌های رگه‌های باران را از طریق کدگذاری پراکنده از لایه فرکانس بالا جدا می کنند. روش‌های متعدد دیگری نیز برای حذف باران از تصویر ارائه شده‌اند که لزوماً برای تصاویر ترافیکی نیستند. به طور مثال Yi-Lei Chen و همکارش [۱۳] لایه‌های پس‌زمینه و رگه‌های باران را بر اساس طبقه‌بندی فرکانسی و حذف لبه‌های ضعیف تجزیه کردند. Lei Zhu و همکاران [۱۴] نیز یک روش حذف باران را بر اساس جداسازی رگه‌های باران که معمولاً طیف باریکی از جهات را در بر می گیرند، معرفی کرده‌اند. بدین ترتیب پس از حذف باران از تصاویر ترافیکی، نهایتاً با یک مدل شبکه عمیق مثل RetinaNet می توان وسایل نقلیه در تصاویر اصلاح شده را تشخیص داد. Sudha و همکاران [۱۵] یک الگوریتم ردیابی وسیله نقلیه چندگانه را با شرایط آب‌وهوایی مختلف مثل بارانی، شب و مه با استفاده از یادگیری عمیق پیشنهاد می کنند که با استفاده از YOLO v3 کار کرده و الگوریتم تشخیص پس‌زمینه از شیء تقویت شده است. نتایج آزمایش الگوریتم پیشنهادی با ده ویدئوی ورودی مختلف و دو مجموعه داده معیار KITTI و DETRAC آزمایش شده‌اند که دقت ۹۸٫۶٪ را می دهد. Xiaohong Han و همکاران [۱۶] یک مدل YOLO-v2 [۱۷] تغییر یافته معرفی نموده‌اند تا تشخیص تصاویر کوچک خودرو دارای مشکلات دقت کم و عملکرد ضعیف را بهبود بخشند. در این مقاله از دو شبکه Faster RCNN و RetinaNet برای تشخیص خودرو در شرایط بد آب‌وهوایی استفاده شده است. در ادامه

از یک شبکه اصلی و دو زیرشبکه تشکیل گردیده است [۲۶]. شبکه اصلی مسئول محاسبه یک نقشه ویژگی کانونولشن در کل تصویر ورودی است؛ یعنی از هر طبقه‌بندی‌کننده دلخواه می‌تواند استفاده کند. اولین زیرشبکه، طبقه‌بندی شیء را بر روی خروجی شبکه اصلی انجام می‌دهد و در واقع احتمال حضور شیء را در هر موقعیت مکانی پیش‌بینی می‌کند. زیرشبکه دوم رگرسیون کادر مستطیلی کانونولشنی را انجام می‌دهد. این دو زیرشبکه دارای طراحی ساده‌ای هستند. شکل ۲ معماری شبکه RetinaNet را نمایش می‌دهد.

#### ۴- پایگاه داده

پایگاه داده مورد استفاده در این مقاله DAWN است که شامل چهار پوشه مه‌آلود، بارانی، برفی و گرد و غبار می‌باشد. هر پوشه شامل پنج نوع شیء جهت تشخیص (خودرو (۸۲/۲۱٪)، اتوبوس (۲/۰۵٪)، کامیون (۸/۲۲٪)، موتورسیکلت و دوچرخه (۱/۳۶٪) و انسان (۶/۰۷٪)) و ۷۸۴۵ کادر محدودکننده یا مستطیلی است که به صورت فایل PASCAL VOC می‌باشد [۱]. فایل PASCAL VOC شامل شماره تصویر، اندازه تصویر، کلاس شیء (مثلاً خودرو یا دوچرخه) و ابعاد کادرهای مستطیلی شیء که در تصویر مشخص شده می‌باشد. لازم به ذکر است که به‌ازای هر تعداد شیء در هر تصویر، این مقادیر ثبت می‌شود. شکل ۳ نمونه‌های مختلف آبوهوایی پایگاه داده را نمایش می‌دهد.

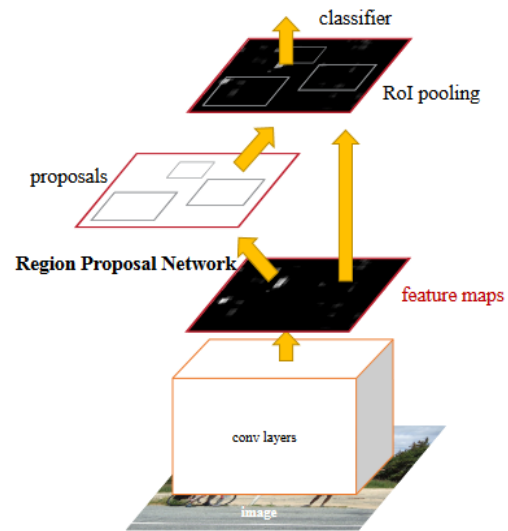
#### ۵- معیار ارزیابی آشکارسازی شیء

##### ۱-۵ تعریف Precision و Recall

برای محاسبه Precision و Recall نیاز است که TP، FP و FN برای تشخیص اشیا تعریف شود. برای محاسبه TP، FP، FN نیاز است IoU محاسبه شود. IoU سنسج هم‌پوشانی بین دو کادر مستطیلی به نام‌های کادر مستطیلی مبنا و کادر مستطیلی پیش‌بینی‌شده را ارزیابی می‌کند. با اعمال IoU می‌توان تشخیص داد که آشکارساز معتبر است یا نامعتبر. پیش‌بینی صحیح (TP) به این معناست که IoU بیشتر از حد آستانه است. این آستانه بسته به چالش قابل تنظیم بوده ولی ۰/۵ یک مقدار استاندارد است. به عنوان مثال بعضی رقابت‌ها نظیر MS COCO از  $mAP@0.5$  یا  $mAP@0.75$  استفاده می‌کنند. این بدان معناست که اگر IoU بیشتر از این مقدار آستانه باشد مثبت صحیح (TP) و اگر کمتر از آستانه باشد مثبت کاذب (FP) در نظر گرفته می‌شود.

##### ۲-۵ محاسبه mAP

برای محاسبه Average Precision ابتدا نمودار  $PR^2$  رسم می‌شود. این منحنی روش مناسبی برای ارزیابی عملکرد آشکارسازی شیء است. اکنون که منحنی PR وجود دارد، AP یا دقت میانگین با محاسبه سطح زیر نمودار محاسبه می‌شود. آنگاه نهایتاً با میانگین‌گیری از AP برای دسته‌های مختلف mean AP به دست می‌آید. در مسابقات COCO، رنجی از مقادیر IoU (حد آستانه) در نظر گرفته می‌شود. ارزیابی COCO می‌تواند AP را به روش‌های مختلف یعنی  $AP@[0.5:0.95]$ ،  $AP@0.5$  و  $AP@0.75$  گزارش کند.  $AP@[0.5:0.95]$  یعنی مقدار آستانه IoU از ۰/۵ تا ۰/۹۵ با طول گام ۰/۰۵ به صورت ۱۰ آستانه تغییر داده شده است.



شکل ۱: Faster RCNN یک شبکه واحد و یکپارچه برای تشخیص اشیاست. مازول RPN به عنوان «پیش‌بینی» این شبکه یکپارچه عمل می‌کند [۱۳].

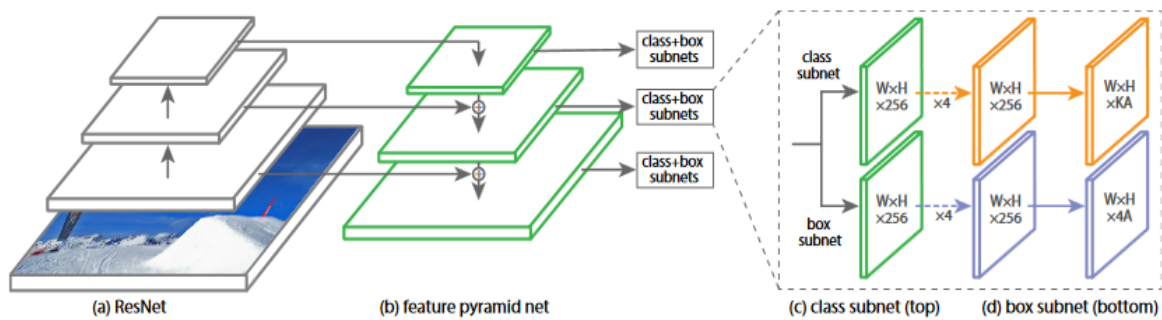
الگوریتم‌های این دسته نسبت به الگوریتم‌های دومرحله‌ای معمولاً سریع‌تر عمل می‌کنند؛ اما دقت کمتری دارند. اغلب برای استفاده از این الگوریتم‌ها از برنامه‌های دارای ویژگی شناسایی لحظه‌ای استفاده می‌کنند.

#### ۳-۳ شبکه کانونولشنی Faster RCNN

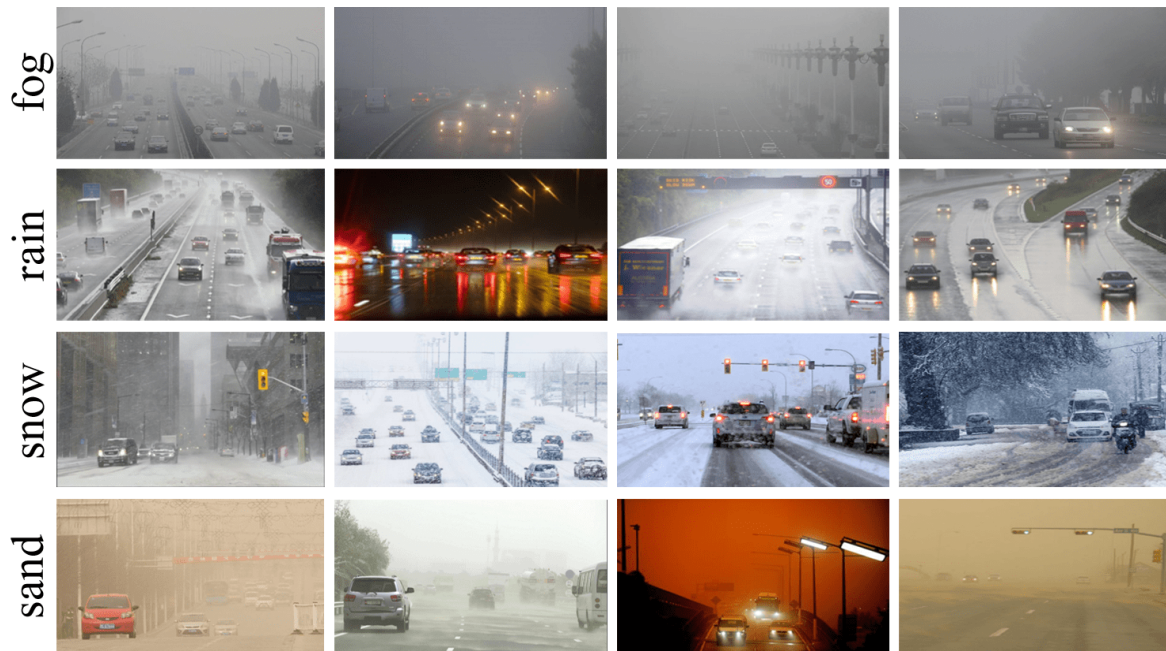
شبکه RCNN Faster، نسخه سوم خانواده RCNN است که در سال ۲۰۱۶ معرفی گردید و شکل ۱ چارچوب این شبکه را نمایش می‌دهد. Shaoqing Ren و همکاران، یک الگوریتم تشخیص اشیا را ارائه کردند که الگوریتم جستجوی گزینشی را حذف کرده و به شبکه اجازه می‌دهد تا پیشنهادها را منطقه را یاد بگیرد. شبیه به Fast RCNN، تصویر به‌عنوان ورودی به یک شبکه کانونولشن ارائه شده که یک نقشه ویژگی کانونولشن ارائه می‌دهد. به‌جای استفاده از الگوریتم جستجوی گزینشی بر روی نقشه ویژگی، برای شناسایی پیشنهادها منطقه از یک شبکه جداگانه برای پیش‌بینی پیشنهادها منطقه استفاده می‌شود. سپس پیشنهادها منطقه پیش‌بینی‌شده با استفاده از یک لایه ادغام ROI تغییر شکل می‌دهند که نهایتاً برای طبقه‌بندی تصویر در منطقه پیشنهادی و پیش‌بینی مقادیر افسست برای جعبه‌های مرزی استفاده می‌شود. Faster RCNN بسیار سریع‌تر از نسخه‌های قبلی خود است و بنابراین حتی می‌توان از آن برای تشخیص اشیا در زمان کم یا آنلاین استفاده کرد.

#### ۴-۳ شبکه RetinaNet

RetinaNet در سال ۲۰۱۷ توسط محققان از FAIR پیشنهاد شد که یک چارچوب یک‌مرحله‌ای مانند YOLO و SSD است. RetinaNet از شبکه پایه ResNet+FPN [۲۵] برای ایجاد یک هرم ویژگی کانونولشن چندمقیاسی غنی استفاده می‌کند. آشکارسازهای یک‌مرحله‌ای که در نمونه‌برداری منظم و مترامک از مکان‌های احتمالی اشیا استفاده می‌کنند، پتانسیل سریع‌تر و ساده‌تر بودن را دارند؛ اما به دلیل عدم تعادل کلاسی شدید که در طول آموزش با آن مواجه می‌شوند از دقت آشکارسازهای دومرحله‌ای عقب‌تر هستند. RetinaNet از یک تابع اتلاف جدید به نام focal loss برای رسیدگی به عدم تعادل کلاس در طول آموزش استفاده می‌کند. RetinaNet یک شبکه واحد و یکپارچه است که



شکل ۲: معماری شبکه RetinaNet یک مرحله‌ای از یک هسته FPN در بالای معماری ResNet برای تولید یک هرم ویژگی پیچیده و چندمقیاسی غنی استفاده می‌کند [۱۷].



شکل ۳: نمونه تصاویر مجموعه داده DAWN که چهار مورد شرایط بد آب‌وهوایی را نشان می‌دهد [۱].

روشنایی تصویر، چرخش تصویر و جابه‌جایی افقی و عمودی تصویر اشاره کرد. از طرفی نوع تغییر تصویر برای داده‌افزایی بسیار مهم است؛ به گونه‌ای که اگر از تبدیل‌های درستی استفاده نشود، نه تنها دقت تشخیص شبکه بیشتر نخواهد شد بلکه کمتر نیز می‌شود. مثلاً می‌توان نشان داد که چرخش تصویر با زاویه نادرست و همچنین جابه‌جایی عمودی به تشخیص خودرو کمکی نخواهد کرد. به عبارت دیگر اگر داده‌افزایی با یکی از این دو روش انجام شود، دقت تشخیص خودرو توسط شبکه آموزش داده‌شده با چنین داده‌هایی کاهش می‌یابد. در این مقاله از روش داده‌افزایی برای افزایش تعداد داده‌های آموزش استفاده شده و روش‌های مورد استفاده نیز تغییر در روشنایی تصویر و جابه‌جایی افقی تصویر هستند.

### ۶-۲ آموزش شبکه Faster RCNN

پایگاه داده DAWN به دو قسمت آموزش و تست تقسیم می‌شود و معمولاً ۸۰٪ داده‌ها برای آموزش شبکه و ۲۰٪ جهت تست مورد استفاده قرار می‌گیرند. سپس یک شبکه Faster RCNN از پیش‌آموزش داده‌شده به‌عنوان شبکه اصلی انتخاب می‌گردد و این شبکه توسط تصاویر آموزش داده می‌شود. با این روش، تشخیص شبکه در تصاویر مه‌آلود از ۲۶٪ در  $IoU = 0.5$  به ۷۰٪ افزایش پیدا کرده است. جدول ۱ نتیجه تست تصاویر مه‌آلود می‌باشد که این تست بر روی دیگر شرایط آب‌وهوایی نیز تکرار شده است. جدول ۲ مربوط به آب‌وهوای گرد و غبار، جدول ۳ مربوط به آب‌وهوای برفی و جدول ۴ مربوط به آب‌وهوای بارانی می‌باشد.

## ۶- نتایج و آزمون

### ۶-۱ پیش‌پردازش و داده‌افزایی

روش پیشنهادی در این مقاله برای پیش‌پردازش تصویر، <sup>۱</sup>AWB است که جهت بهبود کنتراست تصویر استفاده می‌شود و به شبکه برای تشخیص بهتر خودرو کمک خواهد کرد. نحوه کار سیستم AWB به این صورت خواهد بود که ابتدا تصویر ورودی را به سه کانال قرمز، آبی و سبز تبدیل کرده و سپس هر کانال تصویر به‌صورت جداگانه از یک فیلتر یکنواخت‌ساز هیستوگرام عبور داده می‌شود. مجدد با استفاده از فیلتر تصحیح گاما، نور و رنگ تصویر را بهبود داده و نهایتاً با فیلتر لاپلاسیان لبه‌های تصویر تقویت می‌شود.

شبکه‌های عصبی عمیق برای به‌دست‌آوردن نتایج خوب و جلوگیری از برازش بیش از حد به داده‌های آموزشی زیادی نیاز دارند؛ با این حال، اغلب به‌دست‌آوردن نمونه‌های آموزشی کافی بسیار دشوار است. خصوصاً برچسب‌زنی نمونه‌های آموزشی به زمان زیادی نیاز دارد و در مواردی بسیار پرهزینه است. داده‌افزایی، فرایند ایجاد نمونه‌های آموزشی جدید از نمونه‌های موجود است که به‌طور مؤثر انواع زیادی از عملیات تبدیل تصویر را پیاده‌سازی می‌کند. از جمله این تبدیلات می‌توان به تغییر در

1. Automatic White Balance
2. Histogram Equalization



جدول ۱: نتیجه تست شبکه FASTER RCNN برای تصاویر مه‌آلود.

شرایط محاسبه دقت	مقدار IoU	Average Precision
قبل از آموزش شبکه	۰٫۵	۰٫۲۶۲
قبل از آموزش شبکه	۰٫۷۵	۰٫۱۵۰
	[۰٫۵: ۰٫۰۵: ۰٫۹۵]	۰٫۱۴۷
بعد از آموزش شبکه	۰٫۵	۰٫۷۰۱
	۰٫۷۵	۰٫۳۷۴
	[۰٫۵: ۰٫۰۵: ۰٫۹۵]	۰٫۴۱۰

جدول ۲: نتیجه تست شبکه FASTER R-CNN برای تصاویر گرد و غبار.

شرایط محاسبه دقت	مقدار IoU	Average Precision
قبل از آموزش شبکه	۰٫۵	۰٫۲۱۴
قبل از آموزش شبکه	۰٫۷۵	۰٫۱۳۶
	[۰٫۵: ۰٫۰۵: ۰٫۹۵]	۰٫۱۲۸
بعد از آموزش شبکه	۰٫۵	۰٫۵۷۰
	۰٫۷۵	۰٫۲۸۶
	[۰٫۵: ۰٫۰۵: ۰٫۹۵]	۰٫۳۵۰

جدول ۳: نتیجه تست شبکه FASTER R-CNN برای تصاویر برفی.

شرایط محاسبه دقت	مقدار IoU	Average Precision
قبل از آموزش شبکه	۰٫۵	۰٫۱۵۰
قبل از آموزش شبکه	۰٫۷۵	۰٫۰۹۲
	[۰٫۵: ۰٫۰۵: ۰٫۹۵]	۰٫۰۹۲
بعد از آموزش شبکه	۰٫۵	۰٫۵۲۵
	۰٫۷۵	۰٫۳۶۶
	[۰٫۵: ۰٫۰۵: ۰٫۹۵]	۰٫۳۲۰

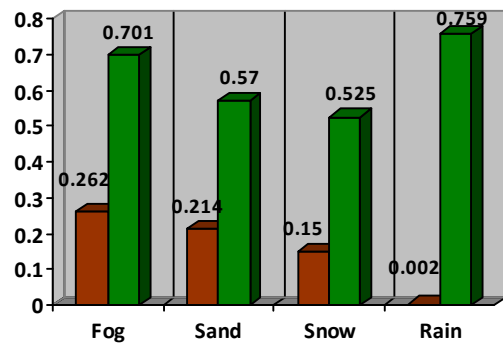
جدول ۴: نتیجه تست شبکه FASTER RCNN برای تصاویر بارانی.

شرایط محاسبه دقت	مقدار IoU	Average Precision
قبل از آموزش شبکه	۰٫۵	۰٫۰۰۲
قبل از آموزش شبکه	۰٫۷۵	۰٫۰۰۰
	[۰٫۵: ۰٫۰۵: ۰٫۹۵]	۰٫۰۰۱
بعد از آموزش شبکه	۰٫۵	۰٫۷۵۹
	۰٫۷۵	۰٫۶۵۶
	[۰٫۵: ۰٫۰۵: ۰٫۹۵]	۰٫۵۵۱

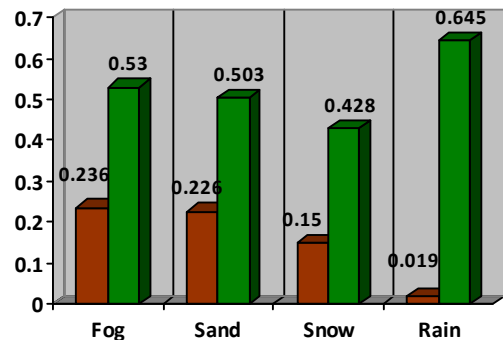
آموزش داده می‌شود. نتیجه آموزش و تست شبکه در هر آب‌وهوا در جداول ۵ تا ۸ قابل مشاهده است. در شکل ۵ می‌توانید نتیجه تست چهار آب‌وهوای مختلف را قبل و بعد از آموزش در شبکه RetinaNet مقایسه کنید. شکل ۶ مثالی از تشخیص خودرو در شرایط مختلف آب‌وهوایی بعد از آموزش شبکه می‌باشد.

### ۶-۴ مقایسه شبکه‌ها

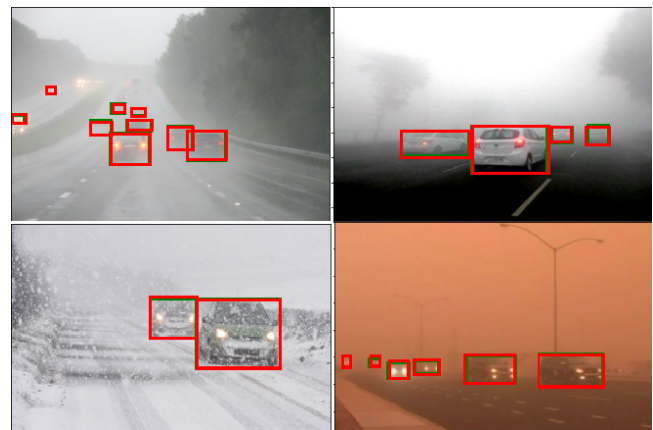
بعد از آموزش شبکه‌ها و دیدن نتایج در این بخش، نتایج به‌دست‌آمده را با نتایج [۱] مقایسه خواهیم کرد. در [۱] از شبکه‌های مختلف برای تشخیص خودرو در شرایط بد آب‌وهوایی استفاده شده و نکته مورد توجه آن است که این شبکه‌ها از پیش‌آموزش داده‌شده هستند و آموزش جدیدی توسط مجموعه داده DAWN که مختص تشخیص خودرو در شرایط بد آب‌وهوایی است داده نشده‌اند. علت این امر آن است که هدف در [۱]، علاوه بر تشخیص خودرو، ردگیری آن است و بنابراین تمرکز مقاله بر روی تشخیص خودرو نیست. با توجه به [۱]، بهترین نتایج جهت تشخیص



شکل ۴: نتیجه تست شبکه Faster RCNN که قسمت قرمز رنگ، قبل از آموزش شبکه و قسمت سبز رنگ، بعد از آموزش شبکه است.



شکل ۵: نتیجه تست شبکه RetinaNet که قسمت قرمز رنگ، قبل از آموزش شبکه و قسمت سبز رنگ، بعد از آموزش شبکه است.



شکل ۶: تصاویر تشخیص خودرو توسط شبکه Faster RCNN در شرایط مختلف آب‌وهوایی (کادر قرمز پیش‌بینی شبکه و کادر سبز کادر اصلی است).

در شکل ۴ می‌توانید نتیجه تست چهار آب‌وهوای مختلف را قبل و بعد از آموزش در شبکه Faster RCNN مقایسه کنید. آنچه که در شکل نشان داده شده است، دقت تشخیص خودرو توسط شبکه Faster RCNN در شرایط مختلف آب‌وهوایی و با انتخاب  $IoU = 0.5$  می‌باشد. همان‌طور که مشاهده می‌شود، افزایش دقت شبکه جهت تشخیص خودرو در شرایط بد آب‌وهوایی با آموزش شبکه بر روی چنین داده‌هایی به‌وضوح قابل تشخیص است. همچنین اختلاف دقت در شرایط مختلف آب‌وهوایی متفاوت است؛ اما در همه حالات‌ها به میزان قابل توجهی وجود دارد. مثلاً در شرایط بارانی، دقت شبکه پس از آموزش آن بر روی چنین داده‌هایی از ۰٫۲٪ به ۷۵٫۹٪ افزایش یافته است.

### ۶-۳ آموزش شبکه RetinaNet

در این بخش نیز یک شبکه RetinaNet از پیش‌آموزش داده‌شده به‌عنوان شبکه اصلی انتخاب می‌شود و این شبکه توسط تصاویر هر گروه

جهت تشخیص هرچه بهتر خودرو در تصاویر، بهترین روش استفاده از شبکه‌های عصبی عمیق در زمینه بینایی کامپیوتر می‌باشد. در این پژوهش سه کار جدید انجام شده است: اول پیش‌پردازش تصویر است که باعث افزایش کیفیت تصویر و در نتیجه افزایش دقت تشخیص خودروها در آب‌وهوای نامناسب می‌شود. دوم پیاده‌سازی روش‌های داده‌افزایی است که با استفاده از آنها آموزش شبکه با دقت بیشتری انجام می‌شود. سوم آموزش دو شبکه کاملاً عمیق با نام‌های Faster RCNN و RetinaNet است و هر دو بسیار عالی عمل می‌کنند. استفاده از شبکه‌های از پیش آموزش دیده شده، کمک بسزایی در بالابردن قدرت تشخیص و کم کردن دوره‌های آموزش کرده و نیاز شبکه به پایگاه داده خیلی قدرتمند را کمتر می‌کند؛ ولی با این حال تعداد بیشتر تصاویر کمک بیشتر به شبکه در جهت آموزش خواهد نمود. نهایتاً در این مقاله بهترین عملکرد برای شبکه Faster RCNN بوده است. می‌توان پیشنهادهایی جهت افق آینده تحقیق به شرح زیر ارائه نمود:

- ۱) استفاده از دیتابیس قوی‌تر با تعداد بیشتر تصاویر در جهت آموزش شبکه عصبی برای بالابردن دقت تشخیص خودرو
- ۲) بهتر نمودن کیفیت تصویر با روش پیش‌پردازش با فیلترهای متفاوت پردازش تصویر و آزمایش و بررسی بیشتر روی این موضوع
- ۳) استفاده از سایر شبکه‌های عمیق از پیش آموزش دیده و آزمایش روی این شبکه‌ها که به طور مثال می‌توان از شبکه YOLOv5 استفاده نمود و نتیجه را بررسی کرد.
- ۴) تغییر هسته شبکه‌های معرفی شده و تست این شبکه که نهایتاً عملکرد تشخیص را بهتر نماید.

## مراجع

- [1] M. Hassaballah, M. A. Kenk, K. Muhammad, and S. Minaee, "Vehicle detection and tracking in adverse weather using a deep learning framework," *IEEE Trans. on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4230-4242, Jul. 2020.
- [2] M. Hassaballah, M. A. Kenk, and I. M. El-Henawy, "Local binary pattern-based on-road vehicle detection in urban traffic scene," *Pattern Analysis and Applications*, vol. 23, no. 4, pp. 1505-1521, 2020.
- [3] P. Kumar Bhaskar and S. Yong, "Image processing based vehicle detection and tracking method," in *Proc. Int. Conf. on Computer and Information Sciences, ICCOINS'14*, 5 pp., Kuala Lumpur, Malaysia, 03-05 Jun. 2014.
- [4] W. Chu, Y. Liu, C. Shen, D. Cai, and X. S. Hua, "Multi-task vehicle detection with region-of-interest voting," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 432-441, Jan. 2018.
- [5] L. Zhang, et al., "Vehicle object detection based on improved retinanet," *J. of Physics: Conf. Series*, vol. 1757, no. 1, Article ID: 012070, 2021.
- [6] J. Shin, et al., "Real-time vehicle detection using deep learning scheme on embedded system," in *Proc. 9th Int. Conf. on Ubiquitous and Future Networks, ICUFN'17*, pp. 272-274, Milan, Italy, 4-7 Jul. 2017.
- [7] H. Wang, et al., "A comparative study of state-of-the-art deep learning algorithms for vehicle detection," *IEEE Intelligent Transportation Systems Magazine*, vol. 11, no. 2, pp. 82-95, Summer 2019.
- [8] H. Kuang, X. Zhang, Y. J. Li, L. L. H. Chan, and H. Yan, "Nighttime vehicle detection based on bio inspired image enhancement and weighted score-level feature fusion," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 4, pp. 927-936, Apr. 2017.
- [9] C. Sakaridis, D. Dai, and L. Van Gool, "Semantic foggy scene understanding with synthetic data," *Int. J. Comput. Vis.*, vol. 126, no. 9, pp. 973-992, Sept. 2018.
- [10] C. Hodges, M. Bennamoun, and H. Rahmani, "Single image dehazing using deep neural networks," *Pattern Recognit. Lett.*, vol. 128, pp. 70-77, Dec. 2019.
- [11] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "Dehazenet: an end-to-end system for single image haze removal," *IEEE Trans. on Image Processing*, vol. 25, pp. 5187-5198, 2016.

جدول ۵: نتیجه تست شبکه RETINANET برای تصاویر مه‌آلود.

شرایط محاسبه دقت	مقدار IoU	Average Precision
قبل از آموزش شبکه	۰٫۵	۰٫۲۳۶
بعد از آموزش شبکه	۰٫۷۵	۰٫۱۵۰
	[۰٫۵: ۰٫۷۵: ۰٫۹۵]	۰٫۱۴۲
	۰٫۵	۰٫۵۳۰
	۰٫۷۵	۰٫۲۴۷
	[۰٫۵: ۰٫۷۵: ۰٫۹۵]	۰٫۲۸۹

جدول ۶: نتیجه تست شبکه RETINANET برای تصاویر گرد و غبار.

شرایط محاسبه دقت	مقدار IoU	Average Precision
قبل از آموزش شبکه	۰٫۵	۰٫۲۲۶
بعد از آموزش شبکه	۰٫۷۵	۰٫۱۸۷
	[۰٫۵: ۰٫۷۵: ۰٫۹۵]	۰٫۱۶۱
	۰٫۵	۰٫۵۰۳
	۰٫۷۵	۰٫۲۵۶
	[۰٫۵: ۰٫۷۵: ۰٫۹۵]	۰٫۲۷۹

جدول ۷: نتیجه تست شبکه RETINANET برای تصاویر برفی.

شرایط محاسبه دقت	مقدار IoU	Average Precision
قبل از آموزش شبکه	۰٫۵	۰٫۱۵۰
بعد از آموزش شبکه	۰٫۷۵	۰٫۱۸۷
	[۰٫۵: ۰٫۷۵: ۰٫۹۵]	۰٫۰۹۱
	۰٫۵	۰٫۴۲۸
	۰٫۷۵	۰٫۲۸۵
	[۰٫۵: ۰٫۷۵: ۰٫۹۵]	۰٫۲۶۷

جدول ۸: نتیجه تست شبکه RETINANET برای تصاویر بارانی.

شرایط محاسبه دقت	مقدار IoU	Average Precision
قبل از آموزش شبکه	۰٫۵	۰٫۰۱۹
بعد از آموزش شبکه	۰٫۷۵	۰٫۰۱۰
	[۰٫۵: ۰٫۷۵: ۰٫۹۵]	۰٫۰۱۰
	۰٫۵	۰٫۶۴۵
	۰٫۷۵	۰٫۲۸۰
	[۰٫۵: ۰٫۷۵: ۰٫۹۵]	۰٫۳۷۷

جدول ۹: مقایسه دقت تشخیص در شبکه‌های مختلف.

Methods	Fog	Rain	Snow	Sand
YOLOv3_R [۷]	۳۸٫۹۵	۴۵٫۵۰	۵۰٫۳۳	۴۱٫۶۵
RetinaNet (proposed)	۵۲٫۰۰	۶۴٫۵۰	۴۲٫۸۰	۵۰٫۳۰
Faster RCNN (proposed)	۷۰٫۱۰	۷۵٫۹۰	۵۲٫۵۰	۵۷٫۰۰

خودرو در شرایط بد آب‌وهوایی مربوط به شبکه از پیش آموزش داده شده YOLO است که در جدول ۹ می‌توانید نتایج به دست آمده را مقایسه کنید. برای نمونه در آب‌وهوای مه‌آلود در شبکه از پیش آموزش داده شده YOLO در مقاله مرجع به دقت ۳۸٫۹۵ رسیده است؛ اما در این مقاله با آموزش شبکه Faster RCNN به دقت ۷۰٫۱ رسیده ایم.

## ۷- نتیجه گیری

با توجه به بررسی‌های انجام شده در این مقاله و آزمایش‌های مختلف

- [22] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *Proc. of the IEEE Conf. Computer Vision and Pattern Recognition, CVPR'16*, pp. 779-788, Las Vegas, NV, USA, 27-30 Jun. 2016.
- [23] W. Liu, et al., "SSD: single shot multibox detector," in *Proc. of the 14th European Conference, Computer Vision, ECCV'16*, pp. 21-37, Amsterdam, The Netherlands, 11-14 Oct. 2016.
- [24] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. of the IEEE Int. Conf. on Computer Vision*, pp. 2980-2988, Venice, Italy, 22-29 Oct. 2017.
- [25] T. Y. et al., "Feature pyramid networks for object detection," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 2117-2125, Honolulu, HI, USA, 21-26 Jul. 2017.
- [26] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 318 - 327, Feb. 2017.
- [12] S. Li, et al., "Single image deraining: a comprehensive benchmark analysis," in *Proc. of the IEEE/CVF Conf. Comput. Vis. Pattern Recognit*, pp. 3838-3847, Long Beach, CA, USA, 15-20 Jun. 2019.
- [13] Y. L. Chen and C. T. Hsu, "A generalized low-rank appearance model for spatio-temporally correlated rain streaks," in *Proc. of the IEEE Int. Conf. on Computer Vision*, pp. 1968-1975, Sydney, Australia, 1-8 Dec. 2013.
- [14] L. Zhu, C. W. Fu, D. Lischinski, and P. A. Heng, "Joint bi-layer optimization for single-image rain streak removal," in *Proc. of the IEEE In. Conf. on Computer Vision*, pp. 2526-2534, Venice, Italy, 27-29 Oct. 2017.
- [15] D. Sudha and J. Priyadarshini, "An intelligent multiple vehicle detection and tracking using modified vibe algorithm and deep learning algorithm," *Soft Computing*, vol. 24, no. 22, pp. 17417-17429, Nov. 2020.
- [16] X. Han, J. Chang, and K. Wang, "Real-time object detection based on YOLO-V2 for tiny vehicle object," *Procedia Computer Science*, vol. 183, pp. 61-72, 2021.
- [17] M. Shafiee, B. Chywl, F. Li, and A. Wong, Fast YOLO: A Fast You Only Look Once System for Real-Time Embedded Object Selection in Video, arXiv preprint arXiv:1709.05943, 2017.
- [18] S. R. Sree, S. B. Vyshnavi, and N. Jayapandian, "Real-world application of machine learning and deep learning," in *Proc. of the Int. Conf. on Smart Systems and Inventive Technology, ICSSIT'19*, pp. 1069-1073, Tirunelveli, India, 27-29 Nov. 2019.
- [19] M. A. Ponti, L. S. F. Ribeiro, T. S. Nazare, T. Bui, and J. Collomosse, "Everything you wanted to know about deep learning for computer vision but were afraid to ask," in *Proc. of the 30th SIBGRAPI Conf. on Graphics, Patterns and Images Tutoriais, SIBGRAPI-T'17*, pp. 17-41, Niteroi, Brazil, 17-18 Oct. 2017.
- [20] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, pp. 137-149, Jun. 2017.
- [21] R. Girshick, "Fast R-CNN," in *Proc. of the IEEE Int. Conf. on Computer Vision, ICCV'15*, pp. 1440-1448, Santiago, Chile, 07-13 Dec. 2015.

**یاسر جمشیدی** در سال ۱۳۸۸ مدرک کاردانی الکترونیک را از آموزشکده فنی قزوین و مدرک کارشناسی الکترونیک را از دانشگاه غیر انتفاعی دیلمان در سال ۱۳۹۲ و مدرک کارشناسی ارشد سیستم های الکترونیک دیجیتال را در سال ۱۴۰۱ از دانشگاه علم و فرهنگ دریافت نموده است. زمینه‌های مورد علاقه ایشان پردازش تصویر، هوش مصنوعی و برنامه‌نویسی Embedded هستند.

**راضیه سادات اخوت** در سال ۱۳۸۵ مدرک کارشناسی مهندسی برق را از دانشگاه صنعتی شریف و مدارک کارشناسی ارشد و دکترای مهندسی برق را به ترتیب در سال‌های ۱۳۸۷ و ۱۳۹۳ از دانشگاه علم و صنعت دریافت نمود. نام‌برده از سال ۱۳۹۹ تاکنون عضو هیأت علمی دانشگاه علم و فرهنگ است و زمینه‌های مورد علاقه ایشان پردازش سیگنال، پردازش تصویر، هوش مصنوعی و پیاده‌سازی مدارهای دیجیتال بر روی FPGA هستند.