

تحلیل و ارزیابی روش تعیین عملکرد ماشین خودران در تقاطع‌ها در شرایط چراغ زرد با استفاده از هوش مصنوعی

مقاله علمی-پژوهشی

*شهریار افندی‌زاده (نویسنده مسئول)، استاد، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران، تهران، ایران

محمود احمدی نژاد، دانشیار، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران، تهران، ایران

امیرحسین داراب‌پور، دانشجوی دکتری، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران، تهران، ایران

حمید بیگدلی‌راد، دانشجوی دکتری، دانشکده مهندسی عمران، دانشگاه علم و صنعت ایران، تهران، ایران

*پست الکترونیکی نویسنده مسئول: zargari@iust.ac.ir

دریافت: ۱۴۰۴/۰۵/۰۵ - پذیرش: ۱۴۰۴/۱۰/۰۲

صفحه ۵۴-۳۵

چکیده

هدف اصلی این پژوهش، پیاده‌سازی یک سیستم هوش مصنوعی بر پایه یادگیری عمیق برای تشخیص و تعیین فاصله‌ی چراغ‌های راهنمایی تا خودروی خودران و عملکرد مناسب خودرو در شرایط زرد چراغ می‌باشد. نوآوری این پژوهش استفاده از مجموعه‌داده شهر تهران و پیاده‌سازی مدلی با دقت بیشتر از مدل‌های ساخته‌شده در مطالعات پیشین و همچنین استفاده از یک دوربین برای تشخیص فاصله خودرو تا چراغ راهنمایی می‌باشد. برای پیاده‌سازی مدل از مجموعه‌داده مختص چراغ‌های شهر تهران استفاده گردیده که مدل در چارچوب کتابخانه TensorFlow عمل می‌کند. برای ارزیابی عملکرد مدل در شرایط مختلف از جمله عرض تقاطع و سرعت وسیله نقلیه از تصاویر بخش آموزش استفاده شد که خودرو را قادر می‌سازد تا با استفاده از چهار سناریوی مطرح شده در این پژوهش تصمیم به عبور از تقاطع یا توقف در تقاطع بگیرد. تحلیل نتایج مدل با بررسی خروجی مدل همچون درستی، دقت، یادآوری، امتیاز F1 و سرعت مدل‌ها با نتایج مطالعات گذشته ارزیابی شده و نشان داد نتایج صحیح بوده و نسبت به مدل‌های موجود از دقت بالاتری برخوردار می‌باشد. همچنین مدل برتر ارائه شده در این پژوهش، دارای دقتی برابر ۹۶ و درستی برابر با ۹۸ درصد می‌باشد. این سیستم بر اساس داده‌های چراغ راهنمایی شهر تهران قادر است با خطای کمتر از یک درصد (۰/۸ درصد) فاصله‌ی خودرو تا چراغ راهنمایی را محاسبه کند که نشان دهنده دقت بالای مدل می‌باشد که می‌تواند یک عملکرد مناسب را در شرایط زرد چراغ ارائه دهد. از این مدل می‌توان به عنوان جایگزین انسان در تشخیص چراغ راهنمایی در خودروهای خودران در حالت زرد چراغ استفاده نمود.

واژه‌های کلیدی: خودرو خودران، چراغ راهنمایی، تشخیص فاصله، تک دوربین، یادگیری عمیق

۱- مقدمه

منجر شده‌اند. یکی از جنبه‌های حیاتی خودروهای خودران، تشخیص و تفسیر اطلاعات حاصل از محیط پیرامون آن‌ها می‌باشد (Afandizadeh Zargari et al., 2019). یکی از اجزای مهم در تصویربرداری محیط رانندگی، دوربین‌ها هستند. دوربین‌های نصب شده در خودروهای خودران، اطلاعات بصری محیط اطراف را به صورت پیوسته ضبط می‌کنند و این اطلاعات

با پیشرفت روزافزون فناوری‌های هوش مصنوعی و خودروهای خودران، صنعت خودروسازی به یک دوران نوآوری و تحولات عمده راه یافته است (Afandizadeh and Bigdeli Rad, 2021). خودروهای خودران با امکانات پیشرفته‌ای همچون سیستم‌های تشخیص محیطی، حسگرهای مختلف و سیستم‌های کنترل هوشمند، به افزایش امنیت و کارایی در حرکت رانندگی

گردد که یکی از نوآوری‌های این پژوهش ارائه مدلی با دقتی بیشتر از سایر مدل‌های ارائه شده در بخش مطالعات پیشین می‌باشد.

یکی از مسائل دیگری که در بحث ماشین‌های خودران مطرح است تشخیص فاصله تا عوامل متعدد مسیر می‌باشد از این رو مدلی که بتواند در مسیر، اشیا با اهمیت بالا را به درستی تشخیص دهد بخشی از فرآیند را به خوبی طی کرده و مدلی که بتواند تشخیص فاصله را تا آن شی داشته باشد به عنوان مدلی کارآمد در نظر گرفته می‌شود (Zargari and Rad, 2023). چرا که یکی از مهم‌ترین فاکتورها برای عملکرد مناسب در تقاطعات، فاصله‌ی وسیله‌ی نقلیه‌ی خودران تا چراغ‌راهنمایی می‌باشد، به این خاطر تلاش شده تا در این تحقیق، به تشخیص فاصله وسیله نقلیه‌ی خودران تا نزدیک‌ترین فانوس چراغ پرداخته شود. رویکرد این پژوهش استفاده از یک دوربین ساده به جای رویکردهای متفاوتی است که در گذشته به آن پرداخته شده است که این کار باعث افزایش هزینه در خودروهای خودران می‌شد. همچنین در خصوص عملکرد ماشین خودران در شرایط زرد چراغ، با استفاده از فرضیات مساله و نتایج بدست آمده بررسی خواهد گردید تا عملکرد مناسبی طبق سناریوهایی که در ادامه بیان خواهد شد اتخاذ گردد.

۲- پیشینه تحقیق

اولین مطالعات در تشخیص چراغ راهنمایی با استفاده از رنگ چراغ راهنمایی و در سال ۲۰۱۲ انجام شد. دیاز کابرا و همکاران در مورد یک تکنیک جدید برای تشخیص چراغ‌های معلق با استفاده از ویژگی‌های رنگی و تخمین فاصله آنها پژوهش کردند. این مقاله روشی را برای تشخیص چراغ‌های راهنمایی در تصاویر با استفاده از تکنیک‌های بینایی کامپیوتری توصیف می‌کند. این روش شامل مراحل مختلفی از جمله خوشه بندی رنگ، فیلتر کردن، برچسب گذاری است (Diaz-Cabrera et al., 2012). دیاز کابرا و همکاران در مطالعه دیگری تکنیکی برای تشخیص چراغ‌های راهنمایی و تخمین فاصله آنها با استفاده از یک دوربین ارائه دادند. این سیستم از خواص رنگ و خوشه‌بندی فازی برای تشخیص و از فیلترهای بیزی برای تخمین فاصله استفاده می‌کند. این سیستم در سناریوهای واقعی شهری در طول روز و شب آزمایش شد. این سیستم از فیلترهای بیزی مانند فیلتر کالمن و فیلتر ذرات برای تخمین فاصله بین چراغ‌راهنمایی و خودروی

به عنوان ورودی به سیستم‌های هوش مصنوعی مورد استفاده قرار می‌گیرند. یکی از وظایف مهم سیستم‌های هوش مصنوعی در خودروهای خودران، تشخیص چراغ‌های راهنمایی و کنترل ترافیک از فاصله‌ای امن است تا علاوه بر ایجاد راهبری امن برای خودرو، امکان واکنش سریع‌تر به تغییرات محیطی نیز ممکن شود (Gupta et al., 2021). رانندگی خودکار احتمالاً قلب حمل‌ونقل شهری در آینده خواهد بود. خودروهای خودران این پتانسیل را دارند که ایمنی مسافران را افزایش دهند و همچنین سفرهای جاده‌ای را کوتاه‌تر و لذت بخش‌تر کنند. همچنین تاثیر تقاطع‌های جاده‌ای بر تاخیرهای ترافیکی منجر به اتلاف عظیم منابع انسانی و طبیعی می‌شود (Afandizadeh et al., 2023).

مسئله مورد بحث در این پژوهش، بررسی چند مدل یادگیری عمیق برای تشخیص چراغ‌راهنمایی در حالت زرد و تشخیص فاصله آن تا خودرو خودران می‌باشد. با پیشرفت‌های اخیر در هوش مصنوعی (AI)، یادگیری ماشینی (ML) و یادگیری عمیق (DL)، کاربردهای مختلف این تکنیک‌ها برجستگی پیدا کرده و مورد توجه قرار گرفته‌اند. در هوش مصنوعی وسایل نقلیه خودران، سیستم‌های تشخیص بصری (VRS) تعبیه شده است که شامل طبقه‌بندی تصویر، تشخیص اشیا، بخش‌بندی و محلی‌سازی است. برای اینکه این خودروها به یک واقعیت کاربردی تبدیل شوند، باید به ادراک و شناخت مجهز شوند تا با سناریوهای متفاوت وفق پیدا کنند، به تصمیمات مناسب برسند، و مناسب و ایمن‌ترین اقدام را در هر زمان انجام دهند (Gupta et al., 2021). در این پژوهش، سعی شده به عملکرد مدل‌های مختلف پردازش تصویر برای تشخیص چراغ‌های راهنمایی در تقاطعات و تعیین بهترین مدل پرداخته شود که ضرورت این پژوهش به این خاطر است که ارائه مدل مناسب در شرایط زرد چراغ می‌تواند عملکرد مناسبی را به همراه داشته باشد. برای این کار ضروری است به فاکتورهای متعددی از جمله دقت و درستی، سرعت و خطای مدل پرداخته شود. مساله‌ای که در این مطالعه مورد بررسی قرار می‌گیرد عملکرد خودروی خودران در شرایط زرد چراغ می‌باشد با توجه به اهمیت استفاده از مدلی با سرعت و دقت مناسب برای تشخیص عملکرد در حالت زرد چراغ نیاز به ساخت مدلی با دقتی فراتر از مدل‌های پیشین ضروری به نظر می‌رسد که هدف اصلی پژوهش حاضر می‌باشد. برای حل این مساله در این پژوهش به ارائه پنج مدل TensorFlow و پردازش تصویر براساس رنگ پرداخته شده تا بهترین مدل انتخاب

مولر و همکاران روشی را برای استفاده از دید استریو برای تشخیص چراغ‌های راهنمایی شرح می‌دهند. این رویکرد شامل استفاده از یک دوربین استریو برای ثبت تصاویر، استفاده از یک آشکارساز رنگ برای تولید نامزدهای چراغ‌راهنمایی و استفاده از یک مولد فرضیه برای تخمین مکان و جهت چراغ‌های راهنمایی است. مولد فرضیه‌ها از دانش قبلی از آرایش معمولی چراغ‌های راهنمایی (سه لامپ که به صورت عمودی مرتب شده‌اند) برای تخمین مکان و جهت چراغ‌راهنمایی بر اساس بخش رنگ استفاده می‌کند. برای تخمین فاصله تا چراغ راهنمایی، آنها استفاده از یک تصویر نابرابری تولید شده توسط دوربین استریو را توضیح می‌دهند (Fregin et al., 2017).

مطالعه‌ای توسط کولکارنی و همکاران انجام شد در این مطالعه مدل مبتنی بر شبکه عصبی عمیق را برای تشخیص مطمئن چراغ‌های راهنمایی با استفاده از یادگیری انتقال پیشنهاد داده شد. این روش شامل استفاده از یک شبکه کانولوشنال مبتنی بر منطقه سریعتر مدل Inception V2 در TensorFlow برای یادگیری انتقال است. این مدل بر روی یک مجموعه داده حاوی تصاویر مختلف از علائم ترافیکی مطابق با سیگنال‌های ترافیکی هند که در پنج نوع طبقه متمایز می‌شوند، آموزش داده شد (Kulkarni et al., 2018).

گیلاردی و همکاران در مورد یک سیستم تشخیص بی‌درنگ برای چراغ‌های راهنمایی عابر پیاده پژوهش کردند. این سیستم از شبکه‌های عصبی عمیق برای شناسایی و طبقه‌بندی چراغ‌های راهنمایی عابر پیاده استفاده می‌کند و محققان یک مجموعه داده عمومی از ۴۳۹۹ تصویر برچسب‌گذاری شده برای آموزش سیستم ایجاد کرده‌اند. این سیستم برای کمک به افراد کم بینا طراحی شده است که به طور ایمن از خیابان عبور کنند و می‌تواند در دستگاه‌های تلفن همراه تعبیه شود تا به آنها کمک کند تا در محیط‌های بیرونی حرکت کنند (Ghilardi et al., 2018).

مولر و همکاران در پژوهش دیگری در مورد رویکرد یادگیری عمیق برای تشخیص دقیق چراغ‌راهنمایی با استفاده از رویکرد تشخیص تک‌شات (SSD) مطالعه کردند. آنها انطباق‌هایی با رویکرد اصلی برای تشخیص اشیاء کوچک بدون نمونه‌برداری از خود لایه ارائه می‌کنند. آنها همچنین شبکه پایه اصلی را با یک شبکه Inception جایگزین می‌کنند (Müller et al., 2018). دو آن و همکاران روش پیشنهادی Center Net را ارائه می‌دهند که یک الگوریتم تشخیص اشیاء یک مرحله‌ای است. رویکرد آنها به ناتوانی Corner Net در ارزیابی داخل مناطق برش داده شده

خودران استفاده می‌کند. فاصله تخمینی نیز برای افزایش نتایج تشخیص، به استثنای تعداد زیادی از تشخیص‌های نادرست در مرحله ردیابی، استفاده می‌شود (Diaz-Cabrera et al., 2015).

اخباری و همکاران یک رویکرد بلادرنگ برای شناسایی و تشخیص سیگنال‌های چراغ‌راهنمایی در صحنه‌های ترافیکی پیچیده تنها بر اساس تکنیک‌های پردازش تصویر پیشنهاد کردند. سیستم پیشنهادی از ویژگی‌های رنگی صحنه برای تشخیص چراغ‌های راهنمایی در زمان واقعی استفاده می‌کند. این سیستم از چندین مرحله از جمله پیش پردازش، تشخیص افق تطبیقی، تقسیم‌بندی رنگ، طبقه‌بندی و اعتبارسنجی زمانی تشکیل شده است. اطلاعات رنگ چراغ‌راهنمایی به تغییرات شرایط نور مانند سایه‌ها، ابرها و انعکاس خورشید بسیار حساس است. این سیستم از فضای رنگی L^*a^*b برای استخراج اطلاعات رنگ سیگنال‌های چراغ‌راهنمایی در هر تصویر استفاده می‌کند. روش به دست آوردن مناطق چراغ‌راهنمایی نامزد شامل چندین مرحله از جمله تغییر فضای رنگ، تقسیم‌بندی رنگ و عملیات مورفولوژیکی است. اطلاعات رنگ چراغ‌راهنمایی به تغییرات شرایط نور مانند سایه‌ها، ابرها و انعکاس خورشید بسیار حساس است. پس از تست فضاهای رنگی مختلف، از فضای رنگی L^*a^*b استفاده می‌شود. دقت نهایی و یادآوری با استفاده از روش پیشنهادی به ترتیب ۹۵٪ و ۹۴/۷٪ است (Said et al., 2016).

نواک و همکاران یک رویکرد بلادرنگ برای تشخیص و طبقه‌بندی سیگنال‌های چراغ‌راهنمایی با استفاده از تکنیک‌های پردازش تصویر ارائه کردند. این سیستم از خواص رنگ صحنه برای تشخیص چراغ‌های راهنمایی و تکنیکی نوآورانه برای کاهش نیازهای محاسباتی استفاده می‌کند. آنها اهمیت تشخیص و درک چراغ‌راهنمایی برای سیستم‌های کمک راننده و چالش‌های موجود در تشخیص وضعیت چراغ‌راهنمایی را مورد بحث قرار می‌دهند. سیستم پیشنهادی از تکنیک‌های پردازش تصویر برای شناسایی و تشخیص سیگنال‌های چراغ‌راهنمایی در زمان واقعی استفاده می‌کند. این سیستم از چندین مرحله از جمله پیش پردازش، تشخیص افق تطبیقی، تقسیم بندی رنگ، طبقه‌بندی و اعتبارسنجی زمانی تشکیل شده است (Behrendt et al., 2017).

مروری بر چارچوب‌های یادگیری تقویتی عمیق برای کنترل سیگنال‌های ترافیکی ارائه می‌کند و چندین چارچوب یادگیری تقویتی عمیق نماینده را با توجه به فرآیندهای یادگیری، طراحی مدل شبکه عصبی عمیق و الگوریتم‌های بهینه‌سازی تجزیه و تحلیل می‌کند (Gregurić et al., 2020).

لی و همکاران در این پژوهش بر توسعه یک مدل برای کمک به تشخیص چراغ‌راهنمایی تلاش کردند. این روش شامل آموزش و آزمایش مدل‌های مختلف بر روی مجموعه داده LISA است که از تصاویر چراغ‌های راهنمایی در سناریوهای مختلف تشکیل شده است. محققان با شش مدل مختلف آزمایش کردند و دریافتند که مدل Faster RCNN بهترین عملکرد را دارد. نتایج نشان داد که مدل FasterRCNN-ResNet101 بالاترین عملکرد را در بین تمام مدل‌های آزمایش شده به دست آورد (Lee et al., 2022).

وانگ و همکاران در مقاله‌ای، یک روش جدید برای تشخیص همزمان چراغ‌های راهنمایی و رانندگی با استفاده از شبکه‌های عصبی عمیق و ویژگی‌های کانال یکپارچه ارائه کردند. این روش به منظور بهبود دقت و کارایی در تشخیص و پیگیری چراغ‌های راهنمایی در محیط‌های شهری پیچیده طراحی شده است. به عبارت دیگر، این روش به وسیله ترکیب داده‌های تصویری و ویژگی‌های کانال، توانسته است بهبود قابل توجهی در تشخیص و ردیابی چراغ‌های راهنمایی داشته باشد (Wang et al., 2022).

۳- روش شناسی تحقیق

در این پژوهش از دو روش تحقیق نرم‌افزاری و مدلسازی استفاده خواهد شد. دلیل استفاده از روش اول به این خاطر بود که تمامی مدل‌های هوش مصنوعی به وسیله برنامه‌نویسی ساخته می‌شوند. در این پژوهش نیز سعی شده با استفاده از زبان برنامه‌نویسی پایتون و در محیط Google Colab و با استفاده از پردازشگر گرافیکی A100 و با بهره‌گیری از کتابخانه TensorFlow مدل‌های مورد نظر ساخته شده و ارزیابی و اعتبارسنجی صورت گیرد. روش تحقیق دیگر مورد استفاده در این پژوهش مدلسازی خواهد بود که برای ساخت مدل‌ها با استفاده از شرایط موجود در کتابخانه Tensorflow از آن بهره گرفته شده است. در شکل ۱ فلوچارت روش پژوهش آورده شده است.

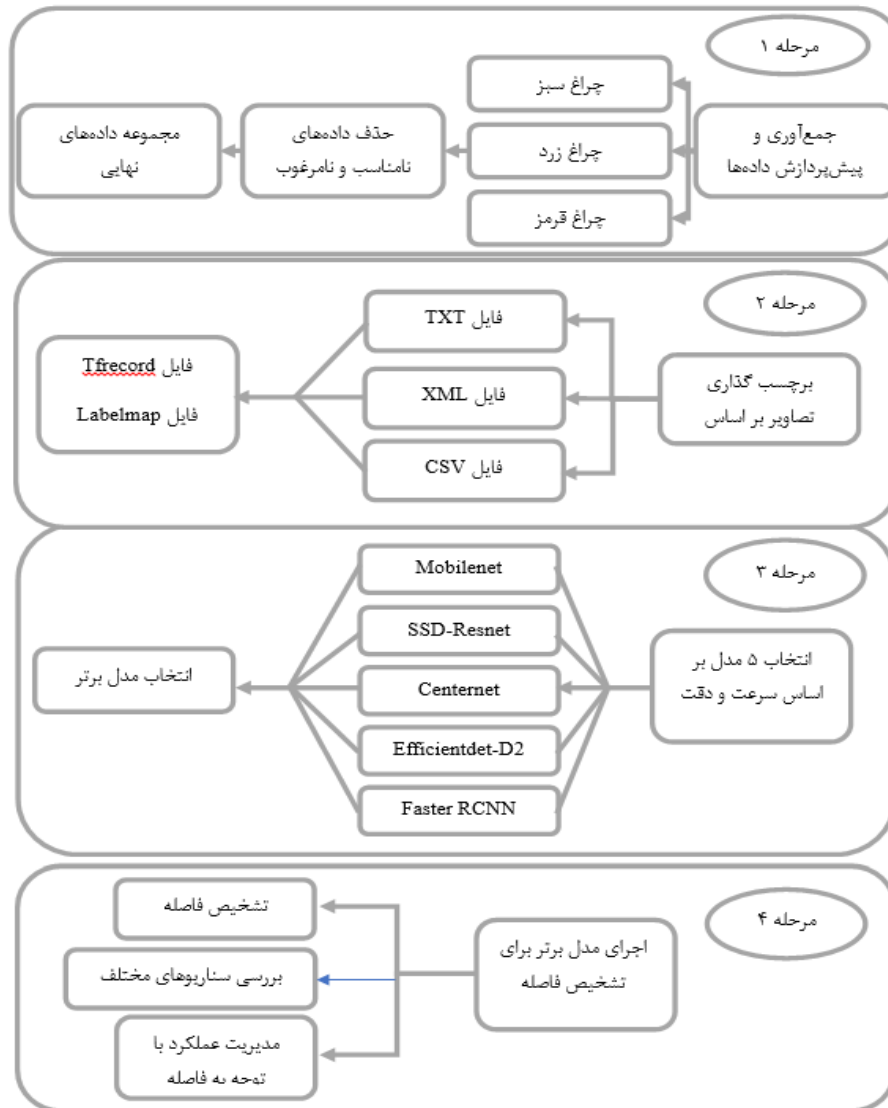
با کاوش در الگوهای بصری در هر منطقه پیشنهادی با حداقل هزینه می‌پردازد که این یک نقص مشترک همه رویکردهای یک مرحله‌ای است (Biswas et al., 2019).

اشنایدر و همکاران استفاده از رادار لیزری مدوله شده فرکانس برای اندازه‌گیری فاصله اجسام متحرک مورد بحث قرار دادند. این روش دقت بالایی را تا ۱۰ میکرومتر ارائه می‌دهد، اما اثرات دینامیکی مانند حرکت جسم یا اسکن پرتو لیزر می‌تواند باعث خطاهای اندازه‌گیری شود. برای غلبه بر این موضوع، نویسندگان مجموعه‌ای با دو دیود لیزری ایجاد کردند که به طور همزمان در فرکانس به سمت بالا و پایین تنظیم می‌شوند و سیگنال‌های هتروداين از هر دو تداخل‌سنج با یکدیگر ضرب می‌شوند. فرکانس میانی بدون اختلال، که متناسب با فاصله است، قابل ارزیابی است (Schneider et al., 2001).

مورولوچه و همکاران در مورد استفاده از استریوسکوپ، تکنیکی برای ایجاد تصاویر سه بعدی با استفاده از دو عکس گرفته شده در موقعیت‌های کمی متفاوت بحث می‌کنند. استریوسکوپ تکنیکی است که برای ایجاد توهم عمق در یک تصویر یا فیلم دو بعدی استفاده می‌شود. این کار با ارائه دو تصویر کمی متفاوت به هر چشم، شبیه‌سازی نحوه درک چشمان ما از عمق در دنیای واقعی است. سپس مغز این دو تصویر را با هم ترکیب می‌کند تا یک تصویر سه بعدی ایجاد کند (Mrovlje et al., 2008).

جیا و همکاران از یک شبکه ترکیبی ویژگی چند مقیاسی برای بهبود توانایی نمایش CNN بر روی اشیاء با ویژگی‌های مقیاس مختلف قبل از BiFPN استفاده کردند. آنها یک ماژول استخراج ویژگی پیشرفته برای این منظور طراحی کردند. این روش همچنین شامل انجام یک پیش‌پردازش ۱ بعدی برای به دست آوردن همبستگی بین کانال‌ها و کالیبراسیون مجدد ورودی اصلی توسط وزن می‌شد. لایه P6 آنها از Maxpool استفاده می‌کرد. در پژوهش آنها از یک مدل یادگیری عمیق به نام Efficientdet و شبکه‌های عصبی کانولوشنال استفاده شده است که این روش‌ها را به عنوان روش‌های اصلی کاربردهای یادگیری عمیق برای تشخیص اشیاء در تصاویر سنجش از دور ذکر می‌کنند (Jia et al., 2022).

موضوع گورگوریک و همکاران بر روی استفاده از یادگیری تقویتی عمیق (DRL) برای کنترل سیگنال ترافیک تمرکز دارد. یادگیری تقویتی عمیق ترکیبی از یادگیری تقویتی و شبکه‌های عصبی عمیق است که امکان تقریب عملکرد کیفیت کارآمد را بدون نیاز به ساختارهای داده مانند جدول فراهم می‌کند. این مقاله



شکل ۱. فلوچارت روش پژوهش

۳-۱- جمع‌آوری داده‌ها

اعتبارسنجی تقسیم می‌شوند. از داده‌های آموزش برای فرآیند یادگیری مدل، از مجموعه داده اعتبارسنجی برای پیش و نظارت بر مدل، تنظیم بهتر ابرپارامترها و همچنین در برخی موارد پایان دادن به یادگیری مدل استفاده شده و برای فرآیند آزمون مدل نیز از مجموعه داده آزمایش استفاده می‌شود. در این پژوهش نیز از این اصل استفاده شده است و داده‌ها به سه گروه آموزش، اعتبارسنجی و آزمایش تقسیم شده‌اند که به شرح زیر هستند.

۱. داده‌های آموزش شامل ۷۰ درصد داده‌های کل و ۱۴۴۷ تصویر
 ۲. داده‌های اعتبارسنجی شامل ۱۵ درصد داده‌های کل و ۲۹۰ تصویر

۳. داده‌های آزمایش شامل ۱۵ درصد داده‌های کل و ۲۹۰ تصویر.

برای آماده‌سازی داده و قرار دادن آن داخل مدل، تصاویر چراغ‌های راهنمایی از زوایای متفاوت در هر سه حالت چراغ‌راهنمایی (سبز، زرد و قرمز) تصویربرداری شد. برای اینکار ۲۰۲۷ تصویر از چراغ‌های راهنمایی در تقاطعات مختلف شهر تهران گرفته شد. تصاویر اولیه با ابعاد ۴۶۲۴ در ۲۶۰۴ پیکسل ثبت شده بود که این حجم برای آموزش به مدل فضای زیادی اشغال می‌کرد به همین خاطر ابعاد تصاویر کاهش یافت. ابعاد تصویر جدید ۱۶ مرتبه نسبت به ابعاد تصویر اولیه کاهش پیدا کرده و طول آن از ۴۶۲۴ پیکسل به ۱۱۵۶ پیکسل و عرض آن از ۲۶۰۴ پیکسل به ۶۵۱ پیکسل کاهش یافت.

روش‌های یادگیری ماشین، برای فرآیند مدل‌سازی، داده‌ها به سه بخش، مجموعه آموزش، مجموعه آزمایش و مجموعه

۳-۲- حذف داده‌های نامناسب برای آموزش

در تصاویر جمع‌آوری شده، تصاویری وجود دارد که ممکن است مدل را برای آموزش دچار مشکل کند، از جمله مواردی که می‌تواند تصویر را در دسته تصاویر نامناسب برای آموزش قرار داد می‌توان به موارد زیر اشاره کرد:

۱. خاموش بودن چراغ‌راهنمایی
۲. چراغ‌های راهنمایی چشمک زن

۳. چند رنگ چراغ همزمان روشن باشد.
 ۴. تمام چراغ‌راهنمایی به درستی مشخص نباشد.
 ۵. تصویر از کیفیت خوبی برخوردار نباشد و تار باشد.
 ۶. چراغ‌های راهنمایی باید از یک نوع باشد.
- تعداد تصاویر نهایی و تعداد جعبه‌های مرزی به تفکیک رنگ به ترتیب در جدول ۱ و ۲ نشان داده شده است. داده‌های تقسیم‌بندی شده به صورت تصادفی انتخاب شده است.

جدول ۱. تعداد تصویر به تفکیک دسته

مجموع	قرمز	زرد	سبز	
۱۴۴۷	۳۷۴	۲۳۲	۸۴۱	آموزش
۲۹۰	۷۵	۴۶	۱۶۹	آزمایش
۲۹۰	۷۶	۴۶	۱۶۸	اعتبارسنجی
۲۰۲۷	۵۲۵	۳۲۴	۱۱۷۸	مجموع

جدول ۲. تعداد جعبه‌های مرزی به تفکیک دسته

قرمز	زرد	سبز	
۸۶۰	۵۳۱	۱۹۳۷	آموزش
۱۷۱	۱۰۸	۳۸۷	آزمایش
۱۷۳	۱۰۷	۳۸۵	اعتبارسنجی
۱۲۰۴	۷۴۶	۲۷۰۹	مجموع

۳-۳- برچسب‌گذاری تصاویر

می‌توانند با اعمال عملیات کانولوشن بر روی تصاویر، ویژگی‌های مختصری از تصویر استخراج کنند و به این ترتیب، مسائلی همچون تشخیص اشیاء، تمایز طبقات و شناسایی الگوها را با دقت بالا حل کنند (Abdi et Afandizadeh et al., 2023).

یک شبکه‌ی عصبی کانولوشنی معمولاً یک ماتریس مرتبه ۳ را به عنوان ورودی خود می‌گیرد، به عنوان مثال، یک تصویر با ردیف‌های H، ستون‌های W و ۳ کانال (کانال‌های رنگی R، G، B). با این حال، ورودی‌های تانسور مرتبه بالاتر می‌توانند توسط شبکه‌ی عصبی کانولوشنی به روشی مشابه مدیریت شوند. سپس ورودی به صورت متوالی از طریق یک سری پردازش می‌گذرد. یک مرحله پردازش معمولاً لایه نامیده می‌شود که می‌تواند یک لایه کانولوشن، یک لایه ادغام، یک لایه عادی سازی و غیره باشد. (Wu et al., 2017)

در مرحله بعد برای استفاده از جعبه مرزی، نیاز هست مرز مستطیلی شکلی چراغ‌راهنمایی را نسبت به سایر قسمت‌ها جدا شود. برای این کار از وب سایت www.makesense.ai کمک گرفته شده است. همچنین باید خاطر نشان کرد جداسازی مرزها با برچسب‌های سبز و زرد قرمز همراه خواهد بود تا مدل درک درستی از تصویر و شرایط آن داشته باشد.

نکته بسیار مهم آنجاست که جعبه مرزی باید با دقت بسیار بالایی ترسیم شود تا باعث پایین آمدن خطای محاسباتی گردد. خروجی‌های نهایی حاصل از وب سایت makesense در سه حالت CSV، XML و txt خواهد بود. خروجی با فرمت CSV برای تمامی عکس‌ها یک فایل و خروجی در دو حالت XML و txt برای هر عکس یک فایل خواهد بود.

۳-۴- آماده سازی فایل Tffrecord

۳-۶- مدل ResNet

قبل از ResNet، راه‌های مختلفی برای مقابله با مشکل ناپدید شدن گرادینان وجود داشت. اما به نظر می‌رسد هیچ یک از راه‌حل‌های موجود یک بار برای همیشه مشکل را حل نمی‌کنند. ایده اصلی ResNet این است که به اصطلاح اتصال میانبر هویت را معرفی کرد که از یک یا چند لایه عبور می‌کند.

وقتی تعداد لایه‌های مدل را افزایش می‌دهیم، یک مشکل رایج در مدل‌های یادگیری عمیق مرتبط با آن وجود دارد که ناپدید شدن گرادینان نامیده می‌شود که باعث می‌شود گرادینان صفر یا خیلی بزرگ شود. بنابراین وقتی تعداد لایه‌ها را افزایش می‌دهیم، میزان خطای آموزش و آزمون نیز افزایش می‌یابد.

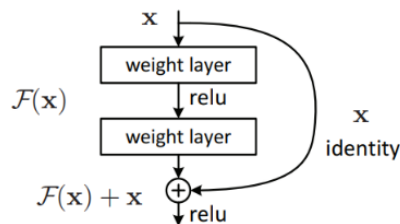
شبکه باقیمانده: به منظور حل مشکل گرادینان ناپدید شدن، این معماری مفهومی به نام بلوک‌های باقیمانده را معرفی کرد. در این شبکه از تکنیکی به نام اتصالات پرش استفاده می‌شود. اتصال پرش فعال‌سازی یک لایه را با پرش برخی از لایه‌ها در بین آن‌ها به لایه‌های بعدی متصل می‌کند. این یک بلوک باقی مانده را تشکیل می‌دهد. Resnet با چیدن این بلوک‌های باقیمانده در کنار هم ساخته می‌شود. رویکرد پشت این شبکه این است که به جای یادگیری لایه‌ها نقشه‌برداری زیربنایی، به شبکه اجازه می‌دهد تا با نگاشت باقیمانده مطابقت داشته باشد.

این فایل، فایل اولیه‌ی مدل خواهد بود و با استفاده از فایل‌هایی که از سایت makesense حاصل می‌شود ساخته شده است درست می‌شود. برای این کار بایستی فایل‌های موجود گفته شده با استفاده از فایل برنامه نویسی به فایل Tffrecord تبدیل شود. فایل Tffrecord فایل مناسب برای کتابخانه TensorFlow محسوب می‌شود تا با استفاده از آن مدل‌ها تشخیص شی که تمامی آن‌ها مدل‌های TensorFlow هستند به خوبی عمل کنند. فایل پایتون Tffrecord اطلاعات را از فایل CSV گرفته و با استفاده از تصاویر موجود ویژگی‌های جعبه مرزی هر عکس را داخل یک فایل کامل نگهداری می‌کند. همانطور که در قسمت‌های قبل گفته شده بود داده‌ها به صورت تصادفی تقسیم شده‌اند ولی باید توجه داشت که فایل CSV برای هر سه قسمت آموزش، اعتبارسنجی و آزمایش دارای اسم فایل در همان پوشه‌ها باشند چون در غیر این صورت نتایج خروجی اشتباهی خواهند داشت. در نهایت برای مدل سازی سه فایل Tffrecord ساخته می‌شود که برای سه قسمت آموزش، اعتبارسنجی و آزمایش خواهد بود.

۳-۵- شبکه‌ی عصبی کانولوشنی

شبکه‌های عصبی کانولوشنال یکی از انواع پرکاربرد شبکه‌های عصبی عمیق هستند که به خصوص برای پردازش تصویر و تشخیص الگوها به کار می‌روند. شبکه‌های عصبی کانولوشنال

دسته‌بندی برای تشخیص دسته‌های اشیاء (مثلاً خودرو، چرخ، یا موتور) و لایه مکان‌یابی برای مشخص کردن مختصات مربع محدوده‌های اشیاء (مانند مختصات گوشه‌ها یا نقاط مرکزی) مورد استفاده قرار می‌گیرند (Ren et al., 2015).



شکل ۲. بلوک Resnet

۳-۸- مدل Efficientdet

یکی از پیشرفته‌ترین الگوریتم‌های تشخیص اشیاء است که ساختاری ساده و عملکرد خوبی دارد. این مدل در هفت نسخه از D0 تا D6 است و وضوح، عمق و عرض مدل را می‌توان به طور همزمان با توجه به محدودیت‌های منابع برای برآوردن نیازهای تشخیص تحت شرایط مختلف، توسط آن مقیاس‌بندی نماید (Jia et al., 2022).

شبکه عصبی Efficientdet یک مدل پیشرفته و کارآمد برای تشخیص اشیاء در تصاویر است که از ترکیب معماری Efficientdet و شبکه‌های عصبی دیگر برای تشخیص اشیاء با دقت بالا و بازدهی بالا استفاده می‌شود. معماری Efficientdet برای افزایش کارایی و دقت مدل‌های تشخیص اشیاء، از روش‌های بهینه‌سازی شبکه و اندازه‌گیری‌های مختلف استفاده می‌کند. اصول اصلی این معماری عبارت‌اند از:

مقیاس‌بندی مرکب: این معماری از روش مقیاس‌بندی مرکب استفاده می‌کند تا ابعاد مختلف شبکه را همزمان با هم تغییر دهد. به این ترتیب، اندازه و پیچیدگی شبکه می‌تواند با حفظ دقت افزایش یابد.

شبکه هرمی ویژگی دو جهته (BiFPN): لایه بهبودیافته‌ای از ویژگی شبکه هرمی است که از سطوح مختلف ویژگی‌ها برای تشخیص اشیاء استفاده می‌کند. اما BiFPN با افزودن مسیرهای جهت‌دار بین لایه‌ها، امکان انتقال اطلاعات بین انواع ویژگی‌ها را ارتقاء می‌دهد و باعث بهبود دقت و سرعت شبکه می‌شود.

ستون فقرات مدل: برای استخراج ویژگی‌ها از تصاویر، از یک نوع از شبکه‌های عصبی به نام Efficientdet استفاده می‌شود. این شبکه‌ها با استفاده از روش مقیاس‌بندی مرکب ابعاد مختلف شبکه را تغییر می‌دهند و با حفظ دقت، کارایی مدل را بهبود می‌بخشند (Tan et al., 2020).

۳-۹- مدل CenterNet

شبکه عصبی CenterNet یک مدل پیشرفته در حوزه تشخیص اشیاء و مکان‌یابی در تصاویر است. این مدل توسط پژوهشگران

مزیت افزودن این نوع اتصال پرش این است که اگر هر لایه‌ای به عملکرد معماری لطمه بزند، با تنظیم کردن از آن عبور می‌کند. بنابراین، این منجر به آموزش یک شبکه عصبی بسیار عمیق بدون مشکلات ناشی از ناپدید شدن گرادیان می‌شود.

شبکه عصبی ResNet یکی از معماری‌های پیشرفته شبکه‌های کانولوشنال است که برای حل مشکل کاهش عمق در شبکه‌های عمیق به کار می‌رود. این معماری از ایده‌های نوآورانه استفاده می‌کند تا به صورت موثر و با دقت بالا، شبکه‌های کانولوشنال با تعداد لایه‌های بسیار عمیق را آموزش دهد (He et al., 2016).

۳-۷- مدل Faster RCNN

ساختار Faster RCNN از سه قسمت اصلی تشکیل شده است: شبکه پیش‌پردازش تصویر برای استخراج ویژگی‌ها، لایه شبکه پیشنهادی منطقه (RPN) برای تولید اولیه محتمل‌ترین مناطق مرتبط با اشیاء و سپس لایه‌های دسته‌بندی و مکان‌یابی برای دقیق‌تر کردن تشخیص اشیاء که در هر منطقه پیش‌بینی شده است. شبکه پیش‌پردازش تصویر: در این مدل، از یک شبکه پیشنهادی منطقه مانند VGG16 یا ResNet استفاده می‌شود تا ویژگی‌های مختصر تصویر استخراج شوند. این ویژگی‌ها به عنوان ورودی به لایه شبکه پیشنهادی منطقه ارسال می‌شوند.

لایه شبکه پیشنهادی منطقه: این لایه، مسئول ایجاد محتمل‌ترین محدوده‌ها برای اشیاء در تصویر است. شبکه پیشنهادی منطقه با استفاده از جعبه‌های لنگر که اندازه‌ها و نسبت‌های مختلفی دارند، احتمال وجود اشیاء در هر محدوده را پیش‌بینی می‌کند. سپس با استفاده از این احتمال‌ها، محدوده‌هایی که در اشیاء احتمالاً وجود دارند، انتخاب می‌شوند و به لایه‌های بعدی برای دسته‌بندی و مکان‌یابی ارسال می‌شوند.

لایه‌های دسته‌بندی و مکان‌یابی: این لایه‌ها ورودی‌هایی را دریافت می‌کنند که توسط شبکه پیشنهادی منطقه انتخاب شده‌اند و وظیفه تشخیص دقیق‌تر اشیاء در هر محدوده را دارند. لایه

به دقت، درستی، یادآوری و امتیاز Fl مدل، مقدار خطای کل کمتر از سایرین و سرعت مدل اشاره کرد که مجموع این پارامترها به انتخاب یک مدل به عنوان مدل برتر منجر می‌شود. در قسمت بعد به بررسی عملکرد خودرو با استفاده از مدل برتر بهبود یافته پرداخته خواهد شد.

۳-۱۲- تشخیص فاصله با استفاده از یک دوربین

تشخیص فاصله‌ی خودرو تا چراغ‌راهنمایی این امکان را به ما می‌دهد تا تصمیم مناسب برای عبور از چراغ‌راهنمایی، توقف پشت آن و یا کاهش سرعت خودرو و عبور را بگیریم. برای تشخیص فاصله تا خودرو نیاز به موارد زیر می‌باشد:

تصویر مرجع: این تصویر شامل موارد زیر خواهد بود.

فاصله‌ی واقعی تا چراغ‌راهنمایی به متر

عرض واقعی چراغ به متر

طول واقعی چراغ به متر

طول چراغ در تصویر مرجع به پیکسل

فاصله کانونی دوربین

طول و عرض چراغ تشخیص داده شده به وسیله‌ی مدل بر اساس پیکسل

برای محاسبه‌ی فاصله‌ی ماشین خودران تا چراغ‌راهنمایی در ابتدا باید فاصله‌ی کانونی دوربینی که تصاویر را ضبط کرده است بدست آوریم، برای این کار باید از فرمول (۱) استفاده کرد:

(۱)

$$FL = (w * rd) / rw$$

FL فاصله کانونی دوربین، w عرض چراغ در تصویر مرجع به پیکسل، rd فاصله‌ی واقعی تا چراغ در تصویر مرجع به متر و rw عرض واقعی چراغ در تصویر مرجع به متر می‌باشد.

پس از محاسبه‌ی فاصله‌ی کانونی می‌توان فاصله‌ی انواع تصاویر ضبط شده را با استفاده از دوربین مشخص بدست آورد که فرمول آن به صورت رابطه ۲ است:

(۲)

$$D = (FL * rw) / wf$$

D فاصله‌ی ماشین خودران تا چراغ‌راهنمایی به متر، FL فاصله‌ی کانونی دوربین، rw عرض واقعی چراغ در تصویر مرجع به متر و wf عرض چراغ در فریم مورد بررسی به پیکسل می‌باشد که از مدل بدست می‌آید.

از دانشگاه کالیفرنیا در سال ۲۰۱۹ ارائه شد و به دلیل سادگی و دقت بالا در تشخیص اشیاء، مورد توجه بسیاری قرار گرفته است. ساختار شبکه عصبی $CenterNet$ برای تشخیص اشیاء از دو بخش اصلی تشکیل شده است: بخش کانولوشنی ستون فقرات برای استخراج ویژگی‌ها برای مکان‌یابی دقیق‌تر اشیاء در تصویر است.

بخش کانولوشنی ستون فقرات: در این بخش از یک شبکه پیش‌پردازش تصویر مانند $ResNet$ یا $Hourglass$ استفاده می‌شود تا ویژگی‌های مختلف تصویر استخراج شوند. این ویژگی‌ها به عنوان ورودی به بخش مکان‌یابی ارسال می‌شوند. بخش مکان‌یابی: این بخش وظیفه دارد مکان‌یابی دقیق‌تر اشیاء در تصویر را انجام دهد. برای این منظور، از یک شبکه عصبی ساده با لایه‌های کانولوشنی و لایه‌های فشرده استفاده می‌شود که بر روی ویژگی‌های ورودی اعمال می‌شود. خروجی این شبکه شامل نقاط مرکزی اشیاء و مختصات مستطیل محدوده‌های اشیاء است. مزیت اصلی $CenterNet$ در مقایسه با مدل‌های دیگر، کارایی بالا و سادگی معماری آن است که موجب کاهش زمان آموزش و پیچیدگی مدل می‌شود و به عنوان یک راه‌حل کارا برای تشخیص اشیاء در تصاویر محسوب می‌شود [۴۷].

۳-۱۰- مدل Mobilenet

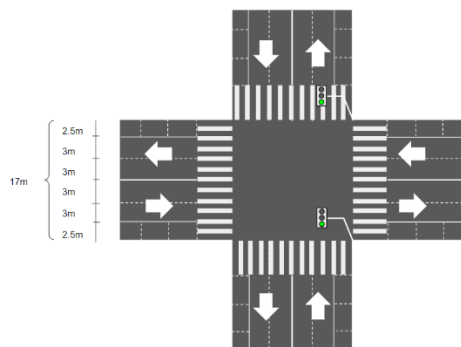
مدل عصبی $MobileNet$ یک معماری شبکه عصبی کانولوشنال موبایل است که به منظور کاهش تعداد پارامترها و پردازش موردنیاز در مدل‌های پیچیده‌تر برای استفاده در دستگاه‌های محدود منابع طراحی شده است. این معماری توسط محققین گوگل با هدف ارائه مدل‌های کارآمدتر برای استفاده در دستگاه‌های همراه و دیگر سیستم‌های محدود ارائه شده است. پارامترها: مقادیری که در فیلترها برای اجرای کانولوشن استفاده می‌شوند. پارامترها در فرآیند آموزش مدل بهینه‌سازی می‌شوند. میزان کاهش پارامترها: میزان کاهش تعداد پارامترها و حجم مدل از طریق تغییر ابعاد فیلترها یا تعداد لایه‌ها است.

۳-۱۱- انتخاب مدل برتر

پس از اجرای مدل و حصول نتایج مدل‌ها در قسمت‌های مختلف از جمله درستی و یادآوری مدل، مدلی از بین این پنج مورد به عنوان مدل برتر برای استفاده در تشخیص چراغ‌راهنمایی انتخاب می‌گردد. از جمله پارامترهای مهم در انتخاب مدل برتر می‌توان

چراغ مورد ارزیابی قرار می‌گیرد. همانطور که توضیح داده شد سناریوهای مورد ارزیابی با رویکردی از جانب خودروی خودران می‌باشد که در جهت مستقیم حرکت می‌کند و تنها چراغ‌های سبز زرد و قرمز را می‌بیند. در قسمت بعد مدل ابتدا رنگ سبز چراغ راهنمایی را تشخیص می‌دهد و منتظر زرد شدن چراغ می‌ماند تا تصمیم‌گیری کند. پس از زرد شدن چراغ در اولین فریمی که به مدل می‌رسد، مدل باید تشخیص رنگ زرد چراغ داده و فاصله‌ی خودرو تا فانوس مرجع را در اولین فریم پس از زرد شدن چراغ ارزیابی کند.

در شکل ۵ مشخصات تقاطع مورد بررسی نشان داده شده است عرض قسمت پارک حاشیه‌ای برابر با ۲/۵ متر و عرض هر یکی از لین‌های راه برابر با ۳ متر در نظر گرفته شده است که مجموع عرض یک سمت تقاطع برابر با ۱۷ متر می‌باشد.

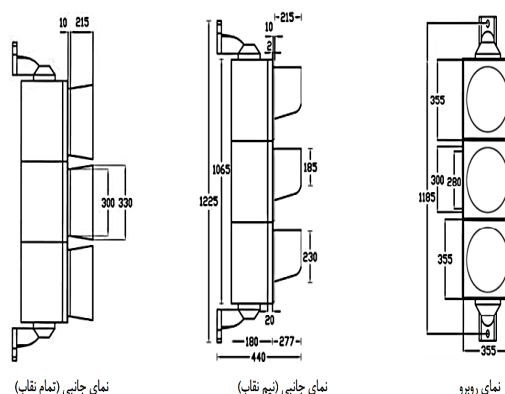


شکل ۵. مشخصات تقاطع مورد بررسی

۳-۱۴- سناریوهای موجود در بررسی عملکرد خودرو

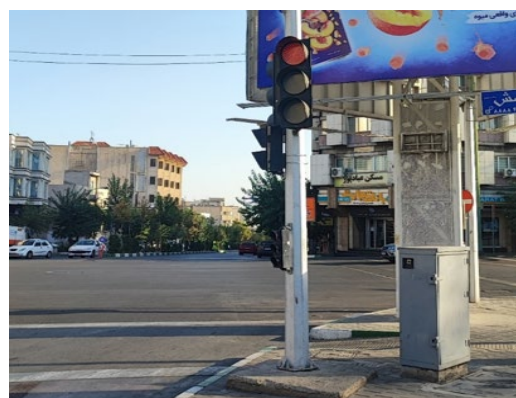
سناریوی اول: در سناریوی اول حالتی بررسی می‌شود که چراغ از سبز به زرد تغییر کرده و با توجه به سرعت وسیله نقلیه و فاصله‌ی آن تا انتهای تقاطع تصمیم به توقف می‌گیرد. سناریوی دوم: در سناریوی دوم حالتی بررسی می‌شود که وسیله نقلیه امکان عبور از چراغ راهنمایی را دارد و پس از سبز شدن چراغ می‌تواند تا قرمز شدن از تقاطع عبور کند. سناریوی سوم: در این سناریو حالتی بررسی می‌شود که فاصله‌ی خودرو تا فانوس ابتدای تقاطع در بازه‌ی بیشتر از ۳ برابر فاصله‌ی توقف کامل باشد. در این حالت وسیله نقلیه به طور خطی و برای افزایش ایمنی سرعت خود را کاهش می‌دهد تا در پشت تقاطع متوقف شود.

برای بدست آوردن ابعاد چراغ راهنمایی از مشخصات فنی تهیه، نصب، نگهداری و تعمیر چراغ‌های راهنمایی و رانندگی که در سال ۱۳۹۸ تنظیم شده است استفاده می‌شود. با توجه به شکل ۳ ابعاد چراغ راهنمایی که در این پژوهش از آن استفاده خواهد شد برای عرض مقدار ۳۵/۵ سانتی متر و برای ارتفاع مقدار ۱۰۱ سانتی متر در نظر گرفته شده است.



شکل ۳. ابعاد چراغ راهنمایی

برای محاسبه‌ی فاصله واقعی چراغ تا دوربین از یک تصویر معیار استفاده شده است که در شکل ۳ نشان داده شده است فاصله‌ی اندازه‌گیری شده برای اینکار به صورت افقی بوده و از ارتفاع چراغ و فاصله‌ی مورب فیثاغورسی تا چراغ صرف نظر شده است. فاصله‌ی که در شکل ۴ تا چراغ راهنمایی ثبت شده است ۷/۵ متر می‌باشد.



شکل ۴. تصویر مرجع برای تشخیص فاصله

۳-۱۳- عملکرد خودرو خودران در شرایط چراغ زرد

پس از محاسبه‌ی فاصله‌ی خودرو خودران تا چراغ راهنمایی چهار سناریوی مختلف برای عملکرد مناسب خودرو در شرایط زرد

خطای منظم‌سازی: هدف تکنیک‌های منظم‌سازی، مقابله با این امر با افزودن محدودیت‌هایی به فرآیند بهینه‌سازی، تشویق مدل به یادگیری نمایش‌های ساده‌تر و قوی‌تر است.

انحراف جعبه: این مرحله مربوط به پیش‌بینی دقیق مکان اشیاء در تصویر است. مدل با پیش‌بینی انحراف‌ها از موقعیت واقعی اشیاء، برای تعیین مرکز شیء و ترسیم جعبه‌های محدود کننده از آن‌ها استفاده می‌کند.

مقیاس جعبه: این مرحله مرتبط با پیش‌بینی اندازه و ابعاد دقیق جعبه‌های محدود کننده اشیاء است. مدل با تخمین مقیاس‌ها، به اشیاء با دقت بیشتری پیش‌بینی می‌کند.

مرکز شیء: در این مرحله، مدل سعی می‌کند مرکز دقیق شیء مورد نظر را در تصویر تشخیص دهد. این اطلاعات به تعیین مکان دقیق‌تر و برجسته‌تر اشیاء کمک می‌کند.

خطای کل: مجموع سه خطای قبلی خطای کل می‌باشد که نشان دهنده عملکرد مدل به طور کلی می‌باشد.

۴-۱- مدل Mobilenet

اولین مدل Tensorflow که در این مطالعه مورد ارزیابی مدل SSD Mobilenet می‌باشد نتایج مدل در قسمت تمرین در جدول ۳ نشان داده شده است. با توجه به ظرفیت سیستمی که مدل‌ها بر روی آن ساخته شده است پارامترهایی از جمله گام مشخص می‌گردد که برای مدل Mobilenet مقدار ۳۰۰۰۰ گام مشخص شده است که در هر مرحله مدل ۱۰۰ گام به جلو می‌رود و نتایج را یادداشت می‌کند.

سناریو چهارم: در این سناریو فاصله‌ی توقف خودرو بیشتر از فاصله تا فانوس نزدیک‌تر است بنابراین وسیله نقلیه باید سرعت خود را افزایش داده و از تقاطع عبور کند.

۴- نتایج و بحث

پنج مدل مختلف از مدل‌های Tensorflow در این مطالعه مورد ارزیابی قرار گرفته در این قسمت نتایج مدل‌ها در چندین بخش مختلف ارائه می‌شود اولین قسمت نتایج مرحله آموزش مدل شامل موارد زیر می‌باشد.

گام: مرحله چندم آموزش را نشان می‌دهد که به طور کلی مدل در هر گام نسبت به گام قبلی عملکرد بهتری را از خود باید نشان دهد.

زمان اجرا: زمان اجرای هر مرحله را به صورت تجمعی نشان می‌دهد.

میزان یادگیری: اندازه یا بزرگی گام را تعیین می‌کند که پارامترهای یک مدل در طی هر تکرار فرآیند بهینه‌سازی به روز می‌شوند.

خطای طبقه‌بندی: هدف خطای طبقه‌بندی اطمینان از این است که مدل به طور دقیق برچسب دسته را به هر شیء شناسایی شده اختصاص می‌دهد.

خطای محلی‌سازی: هدف خطای محلی‌سازی اطمینان از این است که مدل یاد می‌گیرد مختصات جعبه مرزی را پیش‌بینی کند که دقیقاً با مکان‌های حقیقت-زمینی مطابقت دارند.

جدول ۳. نتایج مدل Mobilenet در مرحله‌ی آموزش

گام	زمان اجرا	میزان یادگیری	طبقه‌بندی	خطای محلی	منظم‌سازی	خطای کل
۱۰۰	.	۰/۰۳۱۹۹۹	۰/۴۳۵۸۹۶	۰/۲۵۷۹۱۸	۰/۱۵۴۳۱	۰/۸۴۸۱۲۴
۶۱۰۰	۱۴/۱۱	۰/۰۷۷۸۸۱	۰/۱۲۶۷۳۷	۰/۱۳۲۲۶۹	۰/۱۲۲۷۵۳	۰/۳۸۱۷۵۹
۱۲۱۰۰	۲۸/۱۷	۰/۰۷۰۲۹۱	۰/۰۴۴۸۶	۰/۰۲۶۶۴۵	۰/۰۹۹۳۳۷	۰/۱۷۰۸۴۱
۱۸۱۰۰	۴۲/۲۶	۰/۰۵۸۲۷۴	۰/۰۱۴۸۳۹	۰/۰۱۶۴۶۱	۰/۰۸۳۱۱۴	۰/۱۱۴۴۱۴
۲۴۱۰۰	۵۶/۰۳	۰/۰۴۳۵۸۶	۰/۰۱۸۹۳۲	۰/۰۱۷۱۷	۰/۰۷۰۴۶۸	۰/۱۰۶۵۷
۳۰۰۰۰	۶۹/۵۲	۰/۰۲۸۶۱۹	۰/۰۰۳۳۱۵	۰/۰۰۸۲۷۳	۰/۰۶۱۹۶۹	۰/۰۷۳۵۵۷

دهنده زمان اجرای مدل به دقیقه می‌باشد که نشان می‌دهد مدل Mobilenet در زمان ۶۹/۵۲ دقیقه با ۳۰۰۰۰ گام اجرا شده است که زمان بسیار مناسبی است. مقدار خطای کل مدل در مرحله

یکی از پارامترهای بسیار مهم برای انتخاب مدل برتر زمان اجرای مدل و همچنین زمان آموزش مدل می‌باشد زمان اجرا در قسمت آموزش مدل به صورت تجمعی بوده و آخرین گام در مدل نشان

۴-۲- مدل Centernet

مدل دوم در این مطالعه که به کارگیری فرآیند آموزش و ارزیابی مدل Centernet را نمایندگی می‌کند، در قسمت تمرین با جزئیات در جدول ۴ به تفصیل مورد ارزیابی قرار گرفته است.

آموزش ۰/۰۷۳۵۵۷ می‌باشد که خطای بسیار پایین و قابل قبولی می‌باشد سایر خطاها دیگر مدل نیز به دلیل مقدار بسیار کوچک خطای کل نقش قابل توجهی در مدل ندارند.

جدول ۴. نتایج مدل Centernet در مرحله‌ی آموزش

گام	زمان اجرا	انحراف جعبه	مقیاس جعبه	مرکز شیء	خطای کل
۱۰۰	۰	۰/۳۸۰۴۶۵	۰/۱۵۳۱۶۵	۰/۶۴۰۶۴۸	۱/۱۷۴۲۷۸
۴۱۰۰	۵۸/۷۲	۰/۱۱۸۹۶۶	۰/۰۳۶۵۵	۰/۰۲۹۰۶۹	۰/۱۸۴۵۸۵
۸۱۰۰	۱۱۷/۳۷	۰/۰۸۴۲۷۹	۰/۰۵۱۹۰۴	۰/۰۹۶۲۶۲	۰/۲۳۲۴۴۵
۱۲۱۰۰	۱۷۵/۹۵	۰/۲۹۰۵۵۸	۰/۰۴۰۲۱۲	۰/۰۰۶۵۲	۰/۳۳۷۲۹
۱۶۱۰۰	۲۳۴/۳۷	۰/۱۸۸۲۱۹	۰/۰۵۵۲۳۴	۰/۱۹۷۴۹۵	۰/۴۴۰۹۴۹
۲۰۰۰۰	۲۹۰/۹۷	۰/۲۲۲۴۱۹	۰/۰۴۶۳۵۲	۰/۳۵۴۴۹۶	۰/۶۲۳۲۶۷

تخمین جعبه مرزی هر یک از طبقه‌های موجود دچار مشکل است و نمی‌تواند به خوبی عمل کند.

۴-۳- مدل Resnet

سومین مدل Tensorflow که در این مطالعه مورد ارزیابی مدل Resnet می‌باشد نتایج این مدل در قسمت تمرین در جدول ۵ نشان داده شده است برای مدل Resnet مقدار ۳۰۰۰۰ گام مشخص شده است.

بر اساس نتایج، مشاهده می‌شود که مدل Centernet با اجرای مدل به مدت ۲۹۰/۹۷ دقیقه با ۲۰۰۰۰ گام آموزش، عملکرد خود را ارائه داده است. این زمان نسبت به مدل Mobilenet که با اجرای ۳۰۰۰۰ گام تقریباً ۵ برابر زمان می‌برد، نشان‌دهنده زمانبری بیشتر مدل Centernet نسبت به مدل پیشین است. این امر به‌عنوان یک نقطه منفی در ارتباط با مدل Centernet به‌طور کلی مورد مطالعه قرار می‌گیرد. مقدار خطای کل مدل در مرحله آموزش ۰/۶۲۳۲۶۷ می‌باشد که بخش اصلی آن را خطای انحراف جعبه و مرکز شیء تشکیل داده است که نشان می‌دهد مدل در

جدول ۵. نتایج مدل SSD Resnet در مرحله‌ی آموزش

گام	زمان اجرا	میزان یادگیری	طبقه‌بندی	خطای محلی	منظم‌سازی	خطای کل
۱۰۰	۰	۰/۰۱۴۶۶۶	۰/۷۷۰۴۳۵	۰/۴۸۸۴۹۶	۴۹/۱۵۶۱۷	۵۰/۴۱۵۱
۶۱۰۰	۷۵/۰۸	۰/۰۳۶۹۴۵	۰/۰۵۶۴۱۶	۰/۰۲۵۶۳۵	۹/۱۱۶۶۹۲	۹/۱۹۸۷۴۲
۱۲۱۰۰	۱۵۰/۲۷	۰/۰۲۳۸۰۱	۰/۱۲۶۸۲۲	۰/۰۱۷۳۱	۲/۰۸۴۰۹۸	۲/۲۲۸۲۳۱
۱۸۱۰۰	۲۲۵/۴۵	۰/۰۰۸۲۴۴	۰/۰۶۹۷۶۸	۰/۰۱۹۲۵۶	۰/۹۸۷۸۵۳	۱/۰۷۶۸۷۷
۲۴۱۰۰	۳۰۰/۶۶	۰/۰۰۰۱۵۱	۰/۰۳۸۶۲۷	۰/۰۰۷۹۳۹	۰/۸۴۵۹۶۶	۰/۸۹۲۵۳۱
۳۰۰۰۰	۳۷۴/۵۶	۰	۰/۰۴۸۷	۰/۰۰۸۸۷۸	۰/۸۴۵۶۵۲	۰/۹۰۳۲۳

منظم‌سازی مربوط می‌شود که نشان می‌دهد مدل دچار بیش‌برازش شده است.

۴-۴- مدل Efficientdet

چهارمین مدل Tensorflow که در این مطالعه مورد ارزیابی مدل Efficientdet می‌باشد نتایج این مدل در قسمت تمرین در جدول ۶ نشان داده شده است برای مدل Efficientdet مقدار ۲۰۰۰۰ گام مشخص شده است.

بر اساس نتایج، مشاهده می‌شود که مدل Resnet با اجرای مدل به مدت ۳۷۴/۵۶ دقیقه با ۳۰۰۰۰ گام آموزش، عملکرد خود را ارائه داده است. این زمان نسبت به مدل Mobilenet که با اجرای ۳۰۰۰۰ گام تقریباً ۶ برابر زمان می‌برد، نشان‌دهنده زمانبری بیشتر مدل Resnet نسبت به مدل پیشین است. همچنین زمان بیشتری را از مدل Centernet نشان می‌دهد که این امر به عنوان یک نقطه منفی در ارتباط با مدل Resnet در نظر گرفته می‌شود. مقدار خطای کل مدل در مرحله آموزش ۰/۹۰۳۲۳ می‌باشد. خطای اصلی مدل در این قسمت تماماً به خطای

جدول ۶. نتایج مدل Efficientdet در مرحله‌ی آموزش

گام	زمان اجرا	میزان یادگیری	طبقه‌بندی	خطای محلی	منظم‌سازی	خطای کل
۱۰۰	۰	۰/۰۰۴۱۶	۰/۶۹۰۲۸۸	۰/۳۹۲۷۵۲	۰/۰۳۴۶۸۵	۱/۱۱۷۷۲۵
۴۱۰۰	۲۵/۳۷	۰/۰۷۹۹۹۴	۰/۳۳۶۶۸۶	۰/۲۵۵۶۳۵	۰/۰۶	۰/۶۵۲۳۲۱
۸۱۰۰	۵۰/۹۲	۰/۰۷۹۹۳	۰/۲۳۰۰۹۱	۰/۰۷۹۰۲۴	۰/۰۶۶۵۹۸	۰/۳۷۵۷۱۳
۱۲۱۰۰	۷۶/۵۱	۰/۰۷۹۷۹۵	۰/۲۷۷۱۲	۰/۱۴۱۶۵۵	۰/۰۷۱۴۶۲	۰/۴۹۰۲۳۶
۱۶۱۰۰	۱۰۲/۲۶	۰/۰۷۹۵۸۸	۰/۲۰۹۹۱۸	۰/۰۴۳۳۹۹	۰/۰۷۲۲۷۵	۰/۳۲۵۵۹۳
۲۰۰۰۰	۱۲۷/۳۴	۰/۰۷۹۳۱۹	۰/۰۷۷۶۴۹	۰/۰۳۰۵۳۹	۰/۰۷۴۵۸	۰/۱۸۲۷۶۸

تصویر و جلوگیری از بیش‌برازش دچار مشکل است همچنین باید در قسمت اعتبارسنجی بررسی گردد که آیا این موارد در اعتبارسنجی نیز وجود دارد یا خیر.

۴-۵- مدل Faster RCNN

پنجمین مدل Tensorflow که در این مطالعه مورد ارزیابی مدل Faster RCNN نتایج مدل در قسمت تمرین در جدول ۷ نشان داده شده است برای مدل Faster RCNN مقدار ۲۰۰۰۰ گام مشخص شده است.

بر اساس نتایج، مشاهده می‌شود که مدل Efficientdet با اجرای مدل به مدت ۱۲۷/۳۴ دقیقه با ۲۰۰۰۰ گام آموزش، عملکرد خود را ارائه داده است. این زمان نسبت به مدل Mobilenet که با اجرای ۳۰۰۰۰ گام بیشتر زمان می‌برد اما این مدل نسبت به مدل Centernet و مدل Resnet زمان بسیار کمتری را نشان می‌دهد که این نشان دهنده‌ی عملکرد بهتر این مدل در این زمینه نسبت به دو مدل دیگر می‌باشد.

مقدار خطای کل مدل در مرحله آموزش ۰/۱۸۲۷۶۸ می‌باشد که بخش اصلی این خطا را خطای طبقه‌بندی و منظم‌سازی تشکیل می‌دهد که نشان می‌دهد مدل در پیش‌بینی طبقه موارد داخل

جدول ۷. نتایج مدل Faster RCNN در مرحله آموزش

خطای کل	خطای عینیت RPN	منظم سازی	خطای محلی RPN	خطای محلی BoxClassifier	خطای طبقه بندی BoxClassifier	میزان یادگیری	زمان اجرا	گام
۰/۲۳۸	۰/۰۰۹۱	۰	۰/۰۵۶	۰/۰۸۸	۰/۰۸۵	۰/۰۱۴	۰	۱۰۰
۰/۰۸۵	۰/۰۰۰۵	۰	۰/۰۰۵	۰/۰۶۲	۰/۰۱۷	۰/۰۳۹	۲۵/۴۵	۴۱۰۰
۰/۱۵۶	۰/۰۰۱۳	۰	۰/۰۰۶	۰/۰۹۶	۰/۰۵۲	۰/۰۳۳	۵۱/۰۸	۸۱۰۰
۰/۰۹۸	۰/۰۰۰۶۷	۰	۰/۰۰۵	۰/۰۵۰	۰/۰۴۳	۰/۰۲۴	۷۶/۷۲	۱۲۱۰۰
۰/۰۲۴	۰/۰۰۰۲	۰	۰/۰۰۲	۰/۰۰۹	۰/۰۱۳	۰/۰۱۳	۱۰۲/۵	۱۶۱۰۰
۰/۰۴۵	۰/۰۰۱۱	۰	۰/۰۰۲	۰/۰۲۶	۰/۰۱۵	۰/۰۰۴	۱۲۷/۵	۲۰۰۰۰

در تشخیص اشیاء و تصمیم‌گیری در مورد ویژگی‌های آن‌ها خواهد بود.

۴-۶- مقایسه‌ی نتایج مدل‌ها

۴-۶-۱- بر اساس مجموع خطای آموزش و اعتبارسنجی

با توجه به نتایج بدست آمده در بخش‌های قبل برای مقایسه مدل‌ها بر اساس خطای کل می‌توان از جدول ۸ و شکل ۵ استفاده کرد در این جدول نشان داده شده است بیشترین خطا در مرحله آموزش و اعتبارسنجی مربوط به مدل Resnet بوده و کمترین خطا در مرحله آموزش و اعتبارسنجی مربوط به مدل Faster RCNN می‌باشد. با این حال نباید پارامترهای دیگری برای انتخاب مدل برتر دخیل هستند.

بر اساس نتایج، مشاهده می‌شود که مدل Faster RCNN با اجرای مدل به مدت ۱۲۷/۵ دقیقه با ۲۰۰۰۰ گام آموزش، عملکرد خود را ارائه داده است. این زمان دقیقاً برابر با زمان اجرای مدل Efficientdet است اما این مدل نسبت به مدل Centernet و مدل Resnet زمان بسیار کمتری را نشان می‌دهد که این نشان‌دهنده‌ی عملکرد بهتر این مدل در این زمینه نسبت به دو مدل دیگر می‌باشد.

خطای کل مدل در مرحله آموزش مقدار ۰/۰۴۵ باشد که قابل قبول است. بخش اصلی خطای مدل شامل خطای محلی‌سازی و طبقه‌بندی می‌باشد که مدل در تشخیص محل جعبه‌های مرزی دچار مشکل می‌باشد که با توجه به مقدار آن بنظر مشکل جدی نمی‌آید و این نشان‌دهنده توانایی و دقت مدل Faster RCNN

جدول ۸. مقایسه‌ی خطای آموزش و اعتبارسنجی بین مدل‌ها

خطای کل در مرحله اعتبارسنجی	خطای کل در مرحله آموزش	مدل
۰/۲۲۸	۰/۰۷۴	SSD Mobilenet
۰/۲۹۸	۰/۶۲۳	Centernet HG
۰/۹۲۸	۰/۹۰۳	SSD Resnet101
۰/۳۰۸	۰/۱۸۳	Efficientdet D2
۰/۰۲۵	۰/۰۴۵	Faster RCNN

بیش‌برازش می‌باشد و در سایر داده‌ها به عنوان ورودی احتمالاً عملکرد مناسبی را نمی‌تواند از خود نشان دهد. از طرف دیگر مدل Mobilenet نتایج نزدیکی به مدل Faster RCNN دارد که می‌توان به عنوان یکی از گزینه‌های اصلی در کنار مدل Faster RCNN برای انتخاب مدل برتر باشد.

همانطور که در جدول ۸ نشان داده شده است مدل Resnet بیشترین خطا را نسبت به سایر مدل‌های دیگر دارد اما نکته قابل توجه آنجاست که در قسمت‌های قبل توضیح داده شد که اکثر خطای این مدل به دلیل خطای منظم‌سازی بوده به این خاطر این مدل نمی‌تواند به عنوان مدل برتر انتخاب گردد چرا که دچار

۴-۶-۲- درستی مدل‌ها

Efficientdet و مدل Mobilenet هر کدام با سه مورد پایتتر از سایرین به عنوان مدل‌های ضعیف شناخته می‌شوند. نکته دیگر این است که دو مدل دیگر Faster RCNN و Centernet عملکردی نزدیک به مدل Resnet داشته‌اند و اختلاف فاحشی مشاهده نشده است

جدول ۹ مقایسه‌ی بین درستی انواع مدل‌ها را نشان می‌دهد با توجه به جدول ۹ از ۶ مورد، ۴ مورد را مدل Resnet در حالی عملکرد بهتری داشته که در قسمت قبل توضیح داده شد که از همه‌ی مدل‌ها خطای بیشتری دارد. از طرف مقابل مدل

جدول ۹. مقایسه درستی مدل‌ها

Faster-RCNN	Efficientdet-D2	SSD-Resnet	Centernet	Mobilenet	IOU
۰/۸۰۳	۰/۶۱۷	۰/۸۲۷	۰/۷۴۱	۰/۶۲۹	۰/۵ و ۰/۹۵ کل
۰/۹۹۴	۰/۹۵۸	۰/۹۹۷	۰/۹۹۹	۰/۹۴۷	۰/۵ کل
۰/۹۵۶	۰/۷۴۰	۰/۹۶۰	۰/۹۱۷	۰/۷۰۰	۰/۷۵ کل
۰/۷۳۵	۰/۵۰۱	۰/۷۴۵	۰/۶۷۸	۰/۴۶۹	۰/۵ و ۰/۹۵ کوچک
۰/۸۶۶	۰/۷۰۵	۰/۸۸۵	۰/۸۲۸	۰/۸۱۳	۰/۵ و ۰/۹۵ متوسط
۰/۹۷۱	۰/۸۵۵	۰/۹۵۰	۰/۹۴۴	۰/۹۰۰	۰/۵ و ۰/۹۵ بزرگ

۴-۶-۳- یادآوری مدل‌ها

Mobilenet با وجود چندین تشخیص نادرست اما به طور کلی عملکرد بسیار مناسبی از خود در این قسمت نشان داده است.

جدول ۱۰ مقایسه‌ی بین یادآوری انواع مدل‌ها را نشان می‌دهد با توجه به جدول ۱۰ و همانطور که در قسمت‌های قبل به تک به تک توضیح داده شد می‌توان نتیجه گرفت کدام مدل عملکرد بهتری نسبت به سایر مدل‌ها در قسمت درستی داشته است از ۶ مورد ۵ مورد را مدل Resnet در حالی عملکرد بهتری داشته که در قسمت قبل توضیح داده شد که از همه‌ی مدل‌ها خطای بیشتری دارد. از طرف مقابل مدل Efficientdet در تمام موارد پایتتر از سایرین به عنوان مدل ضعیف شناخته می‌شود. نکته دیگر این است که مدل Faster RCNN عملکردی نزدیک به مدل Resnet داشته و اختلاف فاحشی مشاهده نشده است.

۴-۶-۷- عملکرد مدل Faster RCNN

با توجه به نتایج جدول ۱۲ می‌توان دریافت مدل Faster RCNN به طور کلی عملکرد بسیار مناسبی از خود در این قسمت نشان داده است. با توجه به نتایج بدست آمده و اهمیت رنگ سبز و زرد در مقایسه با رنگ زرد می‌توان دریافت که مدل دقتی تقریباً برابر با ۹۷ درصد و امتیاز FI ۹۸/۴ درصد در قسمت سبز و زرد دارا می‌باشد که بسیار مناسب است.

۴-۶-۸- عملکرد مدل Efficientdet

با توجه به نتایج جدول ۱۳ و ماتریس سردرگمی می‌توان دریافت مدل Efficientdet به طور کلی عملکرد بسیار بدی از خود در این قسمت نشان داده است. با توجه به نتایج بدست آمده و اهمیت رنگ سبز و زرد در مقایسه با رنگ زرد می‌توان دریافت که مدل دقتی تقریباً برابر با ۶۰ درصد و امتیاز FI، ۶۴ درصد در قسمت سبز و زرد دارا می‌باشد که اصلاً مناسب نیست و نباید به عنوان مدل برتر انتخاب گردد.

۴-۶-۵- عملکرد مدل Mobilenet

در جدول ۱۱ درستی، یادآوری، دقت و امتیاز FI مدل Mobilenet در هرکدام از رنگ‌های چراغ مشخص شده است دو رنگ با اهمیت بالا در این پژوهش رنگ سبز و زرد می‌باشد که باید به نتایج حاصل از مدل به شدت توجه کرد. نحوه محاسبه‌ی هر کدام از این موارد در قسمت‌های قبلی توضیح داده شده است. با توجه به نتایج جدول ۱۲ می‌توان دریافت مدل

جدول ۱۰. مقایسه یادآوری مدل‌ها

Faster-RCNN	Efficientdet-D2	SSD-Resnet	Centernet	Mobilenet	AR
۰/۳۹۰	۰/۳۲۶	۰/۳۹۲	۰/۳۴۹	۰/۳۴۰	۱ نمونه
۰/۸۳۹	۰/۶۷۶	۰/۸۵۳	۰/۸۰۴	۰/۷۲۴	۱۰ نمونه
۰/۸۳۹	۰/۶۷۹	۰/۸۵۳	۰/۸۰۴	۰/۷۲۸	۱۰۰ نمونه
۰/۷۷۲	۰/۵۹۹	۰/۷۹۱	۰/۷۴۲	۰/۶۰۲	۱۰۰ نمونه کوچک
۰/۹۰۲	۰/۷۵۲	۰/۹۱۱	۰/۸۶۵	۰/۸۴۶	۱۰۰ نمونه متوسط
۰/۹۷۱	۰/۹۰۷	۰/۹۵۰	۰/۹۵۷	۰/۹۱۴	۱۰۰ نمونه بزرگ

جدول ۱۱. درستی، یادآوری، دقت و امتیاز F1 با IOU@0.5 در مدل Mobilenet

امتیاز F1	دقت	یادآوری	درستی	
۰/۹۴۰۷۱۱	۰/۸۸۸۰۶۰	۰/۹۲۲۴۸۱	۰/۹۵۹۶۷۷	سبز
۰/۹۶۱۹۰۵	۰/۹۲۶۶۰۶	۰/۹۳۵۱۸۵	۰/۹۹۰۱۹۶	زرد
۰/۹۳۷۶۸۶	۰/۸۸۲۶۸۲	۰/۹۲۳۹۷۷	۰/۹۵۱۸۰۷	قرمز

جدول ۱۲. درستی، یادآوری، دقت و امتیاز F1 با IOU@0.5 در مدل Faster RCNN

امتیاز F1	دقت	یادآوری	درستی	
۰/۹۸۶۹۷۹	۰/۹۷۴۲۹۳	۰/۹۷۹۳۲۸	۰/۹۹۴۷۵۱	سبز
۰/۹۸۱۴۸۱	۰/۹۶۳۶۳۶	۰/۹۸۱۴۸۱	۰/۹۸۱۴۸۱	زرد
۰/۹۶۴۹۱۲	۰/۹۳۲۲۰۳	۰/۹۶۴۹۱۲	۰/۹۶۴۹۱۲	قرمز

جدول ۱۳. درستی، یادآوری، دقت و امتیاز F1 با IOU@0.5 در مدل Efficientdet

امتیاز F1	دقت	یادآوری	درستی	
۰/۵۵۰۸۴۸	۰/۶۲۲۷۳۹	۰/۶۲۲۷۳۹	۱/۰۰	سبز
۰/۷۲۹۴۱۲	۰/۵۷۴۰۷۴	۰/۵۷۴۰۷۴	۱/۰۰	زرد
۰/۵۵۰۸۴۸	۰/۳۸۰۱۱۷۰	۰/۳۸۰۱۱۷	۱/۰۰	قرمز

۴-۶-۹- مقایسه بر اساس سرعت و درستی متوسط COCO

با توجه به جدول ۱۴ سرعت هر یک از مدل‌هایی که تا الان مورد بررسی قرار گرفته است نشان داده شده با توجه به این نتایج می‌توان دریافت که سرعت مدل با افزایش درستی کاهش پیدا می‌کند و سرعتی که در این نشان داده شده است در واقع زمان مورد نیاز برای اجرا می‌باشد برای هر یکی از فریم‌هایی که مدل داده می‌شود. با این توضیحات می‌توان دریافت مدل Mobilenet با درستی کمتر از میانگین سایر مدل‌ها از سرعت

بسیار بالایی برای پردازش برخوردار است و با توجه به اینکه زمان تصمیم‌گیری وسیله نقلیه برای عملکرد مناسب ۳ ثانیه ارزیابی می‌شود می‌توان این مدل را بهترین مدل از زمان اجرا در نظر گرفت از طرف مقابل مدل Centernet و مدل Efficientdet کمترین سرعت ممکن را دارند و در عین حال از درستی پایینی نسبت به زمان اجرا برخوردار است.



شکل ۶. اعتبارسنجی تشخیص فاصله توسط مدل

جدول ۱۴. مقایسه‌ی سرعت و عملکرد مدل‌های مورد بررسی

مدل	سرعت (ms)	COCO mAP
CenterNet HourGlass104 512*512	۷۰	۴۱/۹
EfficientDet-D2 768*768	۶۷	۴۱/۸
SSD MobileNet V2 FPNLite 320*320	۲۲/۲	۲۲
SSD ResNet101 V1 FPN 640*640 (RetinaNet101)	۳۵/۶	۵۷
Faster R-CNN ResNet101 V1 640*640	۳۱/۸	۵۵

۴-۷- انتخاب مدل برتر

با نظر به خلاصه‌ای که از تمام بخش‌های پیشین ارائه شده است، بهترین مدل مرتبط با هر قسمت به شرح زیر مشخص می‌شود:

خطا مدل Faster RCNN

نتایج بصری مدل، مدل Resnet

درستی کل مدل، مدل Resnet و Faster RCNN

یادآوری کل مدل، مدل Resnet و Faster RCNN

ماتریس سردرگمی مدل Faster RCNN

درستی، یادآوری، دقت و امتیاز F1 مدل Faster RCNN

سرعت مدل، مدل Mobilenet

با توجه به نتایج گفته شده در قسمت قبل می‌توان دریافت بهترین مدل برای استفاده مدل Faster RCNN خواهد بود چرا که در اکثر موارد نتایج خوبی را از خود نشان داده است مدل جایگزین دیگری نیز می‌توان انتخاب کرد در صورتی که سرعت مدل به خوبی نبود این مدل جایگزین مدل Faster RCNN گردد. مدل جایگزین مدل Mobilenet خواهد بود که با توجه به سرعت بالایی که دارد و با توجه به سایر نتایج که خیلی بد نبوده است می‌تواند به عنوان مدل جایگزین برتر انتخاب گردد.

۴-۸- صحت‌سنجی تشخیص صحیح فاصله

در این قسمت یکی از تصاویری را که از قبل فاصله آن تا چراغ اندازه‌گیری شده است به مدل داده می‌شود تا مشخص گردد فاصله به درستی تشخیص داده شده است یا خیر. تصویری که در این قسمت به مدل داده شده فاصله‌ی ۱۰ متری داشته که همانطور که در شکل ۶ مشخص است با اختلاف ۸ سانتی‌متری و زیر یک درصد به خوبی تشخیص داده شده است.

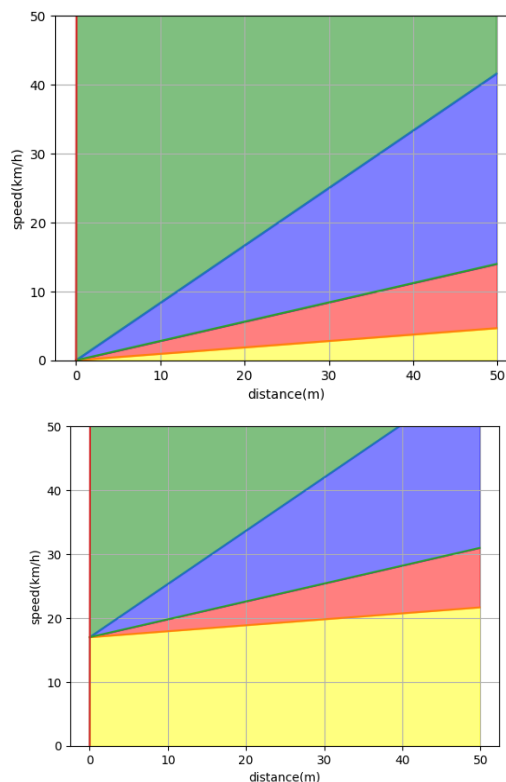
۴-۹- عملکرد خودرو نسبت به سرعت و فاصله از تقاطع

به طور کلی با توجه به روابطی که در قسمت قبل مورد بررسی قرار گرفت سه معادله خط بدست می‌آید که چهار سناریوی مورد بررسی در بین این خطوط قرار می‌گیرند که در شکل ۷ نشان داده شده است. شکل ۷ عملکرد خودرو را در دو شرایط عرض تقاطع با مقدار صفر و ۱۷ متر نشان می‌دهد که در شرایط عرض صفر بازه‌های رنگی مشخص شده هر یک بیانگر سناریوی مشخصی می‌باشد که در این شکل رنگ قرمز بیانگر سناریوی اول یعنی حالتی که خودرو باید متوقف شود، رنگ سبز بیانگر سناریوی دوم یعنی حالتی که خودرو باید از تقاطع عبور کند، رنگ آبی بیانگر سناریوی سوم یعنی خودرو باید سرعت خود را افزایش داده و از تقاطع عبور کند و رنگ زرد بیانگر سناریوی چهارم یعنی خودرو به آرامی باید متوقف شود.

عرض ۱۷ متر شرایطی مشابه با عرض صفر متر تقاطع دارد با این تفاوت که در این حالت عرض تقاطع برابر با ۱۷ متر در نظر گرفته شده است که همانطور که مشخص است با افزایش عرض تقاطع سرعت مورد نیاز برای عبور افزایش می‌یابد و سرعتی کمتر از ۱۸/۲ کیلومتر بر ساعت باید سرعت به آرامی کاهش داده شود.

با توجه به روش استفاده شده در این پژوهش در استفاده از دوربین برای تشخیص فاصله به وسیله فاصله کانونی دوربین می‌توان گفت استفاده از یک دوربین به دلیل دقت بالایی که در تشخیص فاصله دارد به درستی می‌تواند جایگزین روش‌های دیگر برای تشخیص فاصله گردد. اعتبارسنجی عملکرد تشخیص فاصله تا چراغ نشان دهنده خطای کمتر از یک درصد (۰/۸ درصد) بود که این امر نشان دهنده عملکرد بسیار مناسب برای تشخیص فاصله می‌باشد.

با توجه به یافته‌های این پژوهش این ادعا را می‌توان کرد که برای اولین بار از مدلی با دقت بالاتر از سایر مدل‌ها، برای تشخیص چراغ در این پژوهش ارائه شد که این مدل در تخمین طبقه چراغ‌راهنمایی و مکان آن با دقتی برابر ۹۶ و درستی ۹۸ درصد همراه می‌باشد. این مدل به دلیل دقت و سرعت بالا این مدل کاربردهایی از جمله جایگزینی به جای رانندگی سنتی را دارا می‌باشد. با استفاده از مدل پیشنهادی در این پژوهش پیشنهاد می‌شود سایر حالات در تقاطعات از جمله گردش به چپ، وجود عابرین پیاده و سطح سرویس‌های مختلف ترافیکی مورد بررسی قرار گیرد. پیشنهاد می‌شود تاثیر مدل پیشنهادی در کاهش تصادفات و عبور موثرتر از چراغ‌راهنمایی در شرایط زرد چراغ در ماشین‌های خودران نسبت به روش رانندگی سنتی بررسی گردد. همچنین می‌توان بررسی عملکرد ماشین‌های خودران نسبت به یکدیگر در صورت استفاده از این مدل اشاره کرد.



شکل ۷. عملکرد خودرو نسبت به سرعت و فاصله

۵- نتیجه گیری

در این پژوهش به بررسی مفاهیم کلی در ارتباط با مدل‌های یادگیری عمیق در تشخیص اشیا و شیوهی عملکردی آنها پرداخته شد با جمع آوری داده‌های چراغ‌راهنمایی از تقاطعات مختلف شهر تهران چندین مدل مختلف برای انتخاب به عنوان مدلی در جهت عملکرد خودرویی خودران در حالت زرد چراغ ساخته شده. این پژوهش توانسته است یک مدل جدید را معرفی کند که عملکرد بهتری نسبت به مدل‌های موجود در ادبیات پیشنهادی ارائه دهد. این مدل بر روی مجموعه‌ی چراغ‌های راهنمایی شهر تهران مورد ارزیابی قرار گرفته است و نه تنها برای شرایط منطقه‌ای خاص بهینه‌سازی شده است بلکه دارای دقت بسیار بالایی می‌باشد. نتایج بر اساس اعتبارسنجی‌های متعددی از جمله خطای مدل در دو حالت اعتبارسنجی و آموزش، اعتبارسنجی بصری، نتایج درستی و یادآوری مدل، دقت، امتیاز F1 و ماتریس‌های سردرگمی مورد ارزیابی قرار گرفته است. دقت این مدل ۹۶ درصد و درستی ۹۸ درصد می‌باشد که در مقایسه با سایر مدل‌ها عملکرد بهتری را از خود نشان داده است. همچنین از دیگر دستاوردهای این پژوهش استفاده از داده‌های مخصوص چراغ‌های راهنمایی شهر تهران بود که تا به حال از آن استفاده نشده است. با توجه به این که چراغ‌راهنمایی ابعاد ثابت و مشخصی دارد و

۶- مراجع

- Abdi, A., Mosadeq, Z., & Bigdeli Rad, H. (2020). Prioritization of factors affecting suburban road safety by fuzzy hierarchical analysis. *Journal of Transportation Research*.
- Afandizadeh, S., Aziz Jalali, D., & Bigdeli Rad, H. (2023). Optimal routing for shared autonomous vehicles feeder services in urban networks. *Journal of Transportation Research*.
- Afandizadeh, S., & Bigdeli Rad, H. (2021). Developing a model to determine the number of vehicles lane changing on freeways by Brownian motion method. *Nonlinear Engineering*, 10(1), 450-460.
- Afandizadeh Zargari, S., Bigdeli Rad, H., & Shaker, H. (2019). Using optimization and metaheuristic method to reduce the bus headway (Case study: Qazvin Bus Routes). *Quarterly Journal of Transportation Engineering*, 10(4), 833-849.
- Afandizadeh, S., Ourmazdi Khoramshahi, S., & Bigdeli Rad, H. (2023). Evaluation of road diet strategy as one of the methods of demand management for non-motorized vehicles and pedestrians in the cities center. *Road*.

- Jia, J., Fu, M., Liu, X., & Zheng, B. (2022). Underwater object detection based on improved efficientdet. *Remote Sensing*, 14(18), 4487.
- Kulkarni, R., Dhavalikar, S., & Bangar, S. (2018, August). Traffic light detection and recognition for self-driving cars using deep learning. *In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, IEEE. 1-4.
- Lee, J. Y., Ng, H., Yap, T. T. V., Goh, V. T., & Tong, H. L. (2022, December). Traffic Light Recognition Assistance for Color Vision Deficiency Using Deep Learning. *In International Conference on Computer, Information Technology and Intelligent Computing (CITIC 2022)*, Atlantis Press, 289-300.
- Mrovlje, J., & Vrancic, D. (2008, October). Distance measuring based on stereoscopic pictures. *In 9th International Ph. D., Workshop On Systems and Control: Young Generation Viewpoint*, Vol. 2, 1-6.
- Müller, J., & Dietmayer, K. (2018, November). Detecting traffic lights by single shot detection. *In 2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, IEEE. 266-273.
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, 28.
- Rzaev, E., Khanaev, A., & Amerikanov, A. (2021, May). Neural Network for Real-Time Object Detection on FPGA. *In 2021 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, IEEE. 719-723.
- Said, A. F., Hazrati, M. K., & Akhbari, F. (2016, October). Real-time detection and classification of traffic light signals. *In 2016 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, IEEE. 1-5.
- Schneider, R., Thümmel, P., & Stockmann, M. (2001). Distance measurement of moving objects by frequency modulated laser radar. *Optical Engineering*, 40(1), 33-37.
- Tan, M., Pang, R., & Le, Q. V. (2020). Efficientdet: Scalable and efficient object detection. *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 10781-10790.
- Wang, K., Tang, X., Zhao, S., & Zhou, Y. (2022). Simultaneous detection and tracking using deep learning and integrated channel feature for ambient traffic light recognition. *Journal of Ambient Intelligence and Humanized Computing*, 1-11.
- Wu, J. (2017). Introduction to convolutional neural networks. National Key Lab for Novel Software Technology. *Nanjing University*. China, 5(23), 495-496.
- Zargari, S. A., & Rad, H. B. (2023). Development of a gray box system identification model to estimate the parameters affecting traffic accidents. *Nonlinear Engineering*, 12(1), 20220218.
- Behrendt, K., Novak, L., & Botros, R. (2017, May). A deep learning approach to traffic lights: Detection, tracking, and classification. *In 2017 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 1370-1377.
- Biswas, D., Su, H., Wang, C., Stevanovic, A., & Wang, W. (2019). An automatic traffic density estimation using Single Shot Detection (SSD) and MobileNet-SSD. *Physics and Chemistry of the Earth, Parts A/B/C*, 110, 176-184.
- Diaz-Cabrera, M., Cerri, P., & Medici, P. (2015). Robust real-time traffic light detection and distance estimation using a single camera. *Expert Systems with Applications*, 42(8), 3911-3923.
- Diaz-Cabrera, M., Cerri, P., & Sanchez-Medina, J. (2012, September). Suspended traffic lights detection and distance estimation using color features. *In 2012 15th International IEEE Conference on Intelligent Transportation Systems*, IEEE, 1315-1320.
- Duan, K., Bai, S., Xie, L., Qi, H., Huang, Q., & Tian, Q. (2019). Centernet: Keypoint triplets for object detection. *In Proceedings of the IEEE/CVF International Conference on Computer Vision*, 6569-6578.
- Fregin, A., Müller, J., & Dietmayer, K. (2017, June). Three ways of using stereo vision for traffic light recognition. *In 2017 IEEE Intelligent Vehicles Symposium (IV)*, IEEE. 430-436.
- Ghilardi, M. C., Simoes, G., Wehrmann, J., Manssour, I. H., & Barros, R. C. (2018, July). Real-time detection of pedestrian traffic lights for visually-impaired people. *In 2018 International Joint Conference on Neural Networks (IJCNN)*, IEEE. 1-8.
- Gregurić, M., Vujić, M., Alexopoulos, C., & Miletić, M. (2020). Application of deep reinforcement learning in traffic signal control: An overview and impact of open traffic data. *Applied Sciences*, 10(11), 4011.
- Gupta, A., Anpalagan, A., Guan, L., & Khwaja, A. S. (2021). Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues. *Array*, 10, 100057.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770-778.

Analysis and Evaluation of the Method of Determining the Performance of Self-Driving Cars at Intersections in Yellow Light Conditions Using Artificial Intelligence

Shahriar Afandizadeh, Professor, School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran.

Mahmoud Ahmadinejad, Associate Professor, School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran.

Amirhossein Darabpour, M.Sc., Student, School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran.

Hamid Bigdeli Rad, Ph.D., Candidate, School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran.

E-mail: zargari@iust.ac.ir

Received: September 2025- Accepted: February 2026

ABSTRACT

The main goal of this research is to implement an artificial intelligence system based on deep learning to detect and determine the distance between traffic lights and the self-driving car and the proper performance of the car in yellow light conditions. To implement the model, the data set specific to the lights of Tehran city was used, and the model works in the framework of the TensorFlow library. To evaluate the performance of the model in different conditions, including the width of the intersection and the speed of the vehicle, the images of the training section were used, which enables the car to decide to cross the intersection or stop at the intersection using the four scenarios presented in this research. The analysis of the model results by checking the output of the model such as correctness, accuracy, recall, F1 score and speed of the models were evaluated with the results of past studies and showed that the results are correct and have higher accuracy than the existing models. Also, the best model presented in this research has an accuracy of 96% and an accuracy of 98%. Based on the traffic light data of Tehran city, this system is able to calculate the distance of the car to the traffic light with an error of less than one percent (0.8 percent), which shows the high accuracy of the model that can provide a proper performance in yellow light conditions.

Keywords: Self-Driving Car, Traffic Signal, Distance Detection, Single Camera, Deep Learning