

Detection of motorcycle helmets by traffic cameras in difficult situations using deep learning

M.Parvaneh^{ID}, P.Gholamnejad^{ID}*

*Assistant Professor, Shahid Sattari University of Aeronautical Sciences and Technology, Tehran, Iran

Received: 2024 /11/27, Revised: 2025/02/10, Accepted: 2025/03/13, Published: 2025/04/21

DOR: <https://dor.isc.ac/dor/20.1001.1.23224347.1404.13.1.7.2>

ABSTRACT

Reducing injuries and deaths caused by traffic accidents is always the concern of law enforcement officials and governments. In order to reduce injuries in places where the rate of traffic accidents is high due to motorcycle riders not wearing helmets, significant measures have been taken, including the presence of police officers. All these cases are done by human factors, which may reduce the quality of this monitoring and may not bring the desired result, such as the small number of employees and their fatigue. One of the driving violations is the non-use of helmets by motorcyclists. The solution presented in this research is to use deep learning algorithms to detect the use or non-use of helmets by motorcyclists. In addition to detection, the proposed system can analyze traffic data, including When violations are reduced or increased are used. In this research, three versions 416, 320 and spp of the YOLO_v3 deep neural network have been used to detect helmets, and considering that these networks have already been trained on COCO data, the first 53 layers of the network have been used as transfer learning and 53 layers Next, it was trained based on the data used in this research, and then the performance of these three networks was compared with each other. Finally, the proposed automatic helmet detection system performs the detection operation with 96.08% accuracy after being trained by the Helmet Detection dataset. Also, the model was compared with a number of previous works and more favorable results were obtained.

Keywords: Deep learning, neural network, computer vision, helmet.

تشخیص کلاه ایمنی موتورسواران توسط دوربین‌های ترافیکی در حالت‌های دشوار به کمک یادگیری عمیق

مسعود پروانه^{۱*}، پژمان غلام‌نژاد^۲

۱- کارشناسی ارشد، دانشگاه صنعتی مالک اشتر، تهران، ایران ۲- استادیار، دانشگاه علوم و فنون هوایی شهید ستاری، تهران، ایران

(دریافت: ۱۴۰۳/۰۹/۰۷، بازنگری: ۱۴۰۳/۱۱/۲۲، پذیرش: ۱۴۰۳/۱۲/۲۳، انتشار: ۱۴۰۴/۰۲/۰۱)

DOR: <https://dor.isc.ac/dor/20.1001.1.23224347.1404.13.1.7.2>



* این مقاله یک مقاله با دسترسی آزاد است که تحت شرایط و ضوابط مجوز Creative Commons Attribution (CC BY) توزیع شده است.

نویسندگان ©

ناشر: دانشگاه جامع امام حسین (ع)

چکیده

کاهش مصدومیت و مرگ‌ومیر ناشی از حوادث رانندگی همواره مورد توجه مسئولین انتظامی و دولت‌ها است. برای کاهش مصدومیت در مکان‌هایی که میزان وقوع حوادث رانندگی به علت عدم استفاده موتورسواران از کلاه ایمنی در آن زیاد است اقدامات قابل توجهی از جمله حضور افسر انتظامی صورت گرفته است. تمامی این موارد توسط عوامل انسانی صورت گرفته که ممکن است مواردی از قبیل تعداد اندک کارکنان و خستگی آن‌ها کیفیت این نظارت را کاهش داده و نتیجه دلخواه را به همراه نداشته باشد. یکی از موارد تخلفات رانندگی عدم استفاده از کلاه ایمنی توسط موتورسواران است، راهکار ارائه شده در این پژوهش بهره‌بردن از الگوریتم‌های یادگیری عمیق جهت تشخیص استفاده یا عدم استفاده کلاه ایمنی توسط موتورسواران است، سیستم پیشنهادی علاوه بر تشخیص می‌تواند در تحلیل داده‌های ترافیکی از جمله، در چه زمان‌هایی تخلفات کاهش یا افزایش می‌یابد مورد استفاده قرار گیرد. در این پژوهش جهت تشخیص کلاه ایمنی از سه نسخه ۴۱۶، ۳۲۰ و spp شبکه عصبی عمیق YOLO_v3 استفاده شده است و با عنایت به اینکه این شبکه‌ها از قبل بر روی دادگان COCO آموزش دیده‌اند از ۵۳ لایه اول شبکه به صورت یادگیری انتقالی استفاده شده و ۵۳ لایه دیگر بر اساس داده‌های مورد استفاده در این پژوهش مورد آموزش قرار گرفته و سپس عملکرد این سه شبکه با یکدیگر مقایسه شد. در نهایت سیستم پیشنهادی تشخیص خودکار کلاه ایمنی پس از آموزش توسط مجموعه داده Helmet Detection عملیات تشخیص را با مقدار mAP ۰.۸۹۶٪ انجام می‌دهد، همچنین مدل با تعدادی از کارهای پیشین مقایسه گردید و نتایج مطلوب‌تری حاصل شد.

کلیدواژه‌ها: یادگیری عمیق، شبکه عصبی، بینایی کامپیوتر، کلاه ایمنی

۱-مقدمه

با پیشرفت فناوری هوش مصنوعی^۱ و افزایش استفاده از این فناوری و کاربردهای آن، امروزه حوزه‌های مختلف هوش مصنوعی مانند بینایی کامپیوتر^۲ توانسته نقش پررنگی در هوشمندسازی قسمت‌های مختلف زندگی ایفا نماید. بینایی کامپیوتر در حوزه تولید محصولات مانند تشخیص عیوب محصول نهایی نقش بسزایی دارد، یا در بحث شناسایی زودهنگام بیماری‌ها و حتی درمان‌ها آن‌ها اقدامات مفیدی را انجام داده است. یکی دیگر از مسائلی که بینایی کامپیوتر می‌تواند نقش مفیدی را ایفا نماید، کاهش آسیب‌های جانی ناشی از تصادفات رانندگی است. با افزایش تعداد وسیله نقلیه، میزان تصادفات رانندگی نیز افزایش یافته است [۱]. همان‌طور که می‌دانیم در مرکز شهرها تعداد زیادی وسیله نقلیه از جمله موتورسیکلت‌ها در حال تردد هستند، کنترل این وسایل نقلیه از جمله رعایت استفاده از کلاه ایمنی^۳ توسط عوامل انسانی، می‌تواند هزینه مالی و صرف زمان زیادی را در برداشته باشد، همچنین استفاده از عوامل انسانی به علت‌هایی همچون خستگی می‌تواند دقت انجام صحیح امور محوله را کاهش دهد؛ اما با کمک هوش مصنوعی و بینایی کامپیوتر می‌توان به تشخیص تخلفات رانندگی و اقدامات پس از آن نظیر تذکر به فرد متخلف به‌عنوان یک عامل بازدارنده در کاهش حوادث رانندگی کمک کرد.

موتورسیکلت‌ها هنوز وسایل حمل‌ونقل اصلی در کشورهای در حال توسعه هستند. با این حال، در ترافیک جاده‌ای، آن‌ها کاربران آسیب‌پذیر جاده هستند و در صورت تصادف احتمال بیشتری دارد که باعث جراحت یا مرگ شوند. بر اساس گزارش وضعیت ایمنی جاده‌ای سازمان جهانی بهداشت [۲]، راکبان، موتورسواران و عابران پیاده نیمی از مرگ‌ومیرهای جهانی را در تصادفات جاده‌ای تشکیل می‌دهند. سیستم هوش مصنوعی می‌تواند از طریق دوربین‌های قرار گرفته در محل عبور و مرور وسایل نقلیه، ویدئو مربوط به آن را دریافت کرده و استفاده موتورسواران از کلاه ایمنی یا عدم استفاده آن از کلاه ایمنی را تشخیص دهد و پس از آن با استفاده از خروجی سیستم هوش مصنوعی به جریمه موتورسیکلت‌موردنظر یا ارسال پیامک هشدار به آن و حتی تحلیل‌های آماری بپردازد.

تشخیص استفاده یا عدم استفاده موتورسواران از کلاه ایمنی چالش‌های متعددی دارد از جمله آن‌ها می‌توان به موارد زیر اشاره نمود.

(۱) زاویه موتورسوار نسبت به دوربین: درواقع اهداف موردنظر در تصاویر دریافتی از دوربین می‌توانند در زاویه‌های مختلفی نسبت به دوربین قرار گیرند به طور مثال موتورسوار می‌تواند پشت به دوربین یا در مقابل دوربین باشد.

(۲) محیط نوری متفاوت: مورد دیگری که می‌تواند در تشخیص حالت موتورسوار تأثیرگذار باشد، قرارگرفتن محل موتورسوار و میزان نور آن قسمت است این میزان نور می‌تواند تحت تأثیر تاریکی شب و روشنایی روز قرار گیرد. در حالت نور کم که شامل حالت شب هم می‌شود تشخیص حالت موتورسواران دشوار است و حتی ممکن است سایه جسم دیگری بر روی موتورسواران مانع از واضح بودن تصویر دریافتی از آن شود.

(۳) فاصله موتورسوار تا دوربین: در بینایی کامپیوتر یکی از موارد مهم در تشخیص اشیا^۴ فاصله آن‌ها تا دوربین است. با افزایش فاصله شیء موردنظر عواملی همچون وضوح تصویر کاهش پیدا کرده و در نتیجه بر روی تشخیص نهایی آن شیء و حالت آن تأثیر می‌گذارد.

(۴) پوشش سروصورت موتورسوار: یکی دیگر از چالش‌های تشخیص استفاده یا عدم استفاده موتورسوار از کلاه ایمنی، پوشاندن یک قسمت از سروصورت یا تمامی آن است، در نتیجه تصویر کامل موتورسواران جهت تشخیص استفاده از کلاه ایمنی در دسترس نمی‌باشد.

هدف از سیستم تشخیص استفاده از کلاه ایمنی توسط موتورسواران، شناسایی سریع عدم استفاده از کلاه ایمنی و انجام اقدامات پیشگیرانه جهت جلوگیری از حوادث ناشی از تصادفات جاده‌ای است. سیستم موردنظر باید طبق الگوریتم‌های یادگیری و شبکه‌های عصبی عمیق، قادر باشد تا استفاده یا عدم استفاده موتورسواران از کلاه ایمنی را تشخیص داده و با قراردادن کادر اطراف آن، مکان دقیق یکی از دو حالت موردنظر را مشخص کند. ناظرین و مأمورین با استفاده از این سیستم می‌توانند اقداماتی نظیر جریمه را انجام دهند.

در مورد تشخیص کلاه ایمنی موتورسواران به‌وسیله یادگیری عمیق تحقیقات متعددی صورت گرفته است در این بخش خلاصه‌ای از مرتبط‌ترین تحقیقات انجام‌شده آورده شده است. مادوچاندا داسگوپتا و همکارانش [۳] در رویکرد پیشنهادی خود، معماری مبتنی بر شبکه عصبی پیچشی^۵ را برای تشخیص کلاه ایمنی موتورسواران پیاده‌سازی کردند، مدل پیشنهادی بر روی ویدیوهای ترافیکی ارزیابی شد و نتایج به‌دست‌آمده در مقایسه با سایر رویکردهای مبتنی بر شبکه عصبی پیچشی امیدوارکننده

^۴ Object Detection^۵ Convolutional neural network(CNN)^۱ Artificial Intelligence^۲ Computer Vision^۳ helmet

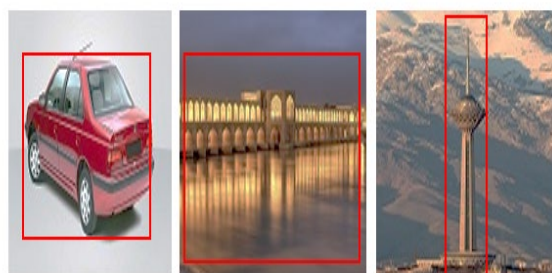
آورده شده است و بخش ۵ به تحلیل نتایج به‌دست‌آمده پرداخته است، در انتها در بخش ۶ نتیجه‌گیری عنوان گردیده است.

۲- کارهای مرتبط

۲-۱- تشخیص شیء

یکی از حوزه‌های اصلی هوش مصنوعی بینایی کامپیوتر است که از کاربردهای این حوزه تشخیص اشیا است، به طور مثال جهت اینکه بتوان تشخیص چهره را روی یک تصویر کلی اعمال کرد، ابتدا باید مکان‌ها و اندازه‌های هر چهره مشخص شود. تشخیص اشیا به‌طورکلی به دو زیرگروه تشخیص اشیا عمومی و تشخیص اشیا خاص [۱۲] تقسیم می‌گردد.

مطابق شکل (۱)، در تشخیص اشیا خاص، هدف یافتن اشیا یا افراد خاصی در تصویر می‌باشد، به‌طور مثال سی‌وسه‌پل یک شیء خاص می‌باشد. اما در تشخیص اشیا عمومی هدف یافتن گروهی مختلفی از اشیا می‌باشد، ازجمله آن‌ها می‌توان به هواپیما، گل، کلاه ایمنی اشاره کرد.



خودرو پژو پارس

سی‌وسه‌پل

برج میلاد

تشخیص اشیا خاص



کلاه ایمنی

کلاه ایمنی

هواپیما

هواپیما

تشخیص اشیا عمومی

شکل (۱): انواع تشخیص اشیا

بود. در ادامه تسبیح وارث و همکارانش [۴]، سامانه‌ای را جهت تشخیص خودکار کلاه ایمنی موتورسواران ارائه دادند و از ویدیوهای نظارتی ضبط‌شده توسط دوربین‌های نصب‌شده در کنار جاده‌ها استفاده کردند. آن‌ها در روش پیشنهادی خود از شبکه عصبی پیچشی مبتنی بر منطقه^۱ استفاده نمودند، تجزیه و تحلیل تجربی نشان می‌دهد که سیستم پیشنهادی آن‌ها دقت ۹۷/۶۹٪ را ارائه می‌دهد و از رقبای خود پیشی گرفت. همچنین الهام سلطانی‌کظمی [۵] همراه همکارانش یک سیستم تشخیص خودکار کلاه ایمنی موتورسواران را ارائه کردند، در این سیستم یک مدل تشخیص اشیا تک‌مرحله‌ای به نام YOLO_v5 پیشنهاد گردید. علاوه بر این از روش‌های تقویت داده‌ها مانند چرخش^۲ برای بهبود عملکرد مدل خود استفاده کردند و کارایی مدل با استفاده از میانگین دقت متوسط^۳، ۵۳/۷۷٪ بود. جفری آگورکو و همکارانش [۶] توسعه و ارزیابی یک مدل یادگیری عمیق YOLO_v5 به همراه یادگیری جمعی^۴ را برای کلاه ایمنی موتورسواران ارائه کردند، مدل پیشنهادی بر روی ۱۰۰ ویدیوی آزمایشی پیاده‌سازی گردید و به میانگین دقت متوسط، ۵۲/۶۷٪ دست پیدا کردند.

تحقیقات دیگری در مورد تشخیص خودکار کلاه ایمنی موتورسواران انجام گرفته است که از شبکه‌های مختلفی (Faster ResNet 50 + R-CNN [۷]، SSD [۸]، SSD MobileNet [۹]، توصیفگرهای HOG، HOP و LDB [۱۰] و ...) استفاده کرده‌اند. از دیگر تحقیقات جهت تشخیص کلاه ایمنی موتورسواران، سیستم تشخیص کلاه ایمنی موتورسواران مبتنی بر اینترنت اشیا^۵ [۱۱] است.

در ادامه، مقاله به این صورت بخش‌بندی گردیده است: در بخش ۲ تحقیق‌های صورت گرفته و کارهای مرتبط در حوزه تشخیص کلاه ایمنی موتورسواران بیان‌شده است در ادامه، مقدمه‌ای در مورد تشخیص اشیا بیان می‌شود، علاوه بر این مقدمه‌ای در رابطه با یادگیری عمیق^۶ و آموزش شبکه مورد استفاده بیان گردیده است. در بخش ۳ معماری و جزئیات شبکه عصبی پیشنهادی که در این پژوهش جهت تشخیص کلاه ایمنی موتورسواران مورد استفاده قرار گرفته، بررسی شده، همچنین در همین بخش مجموعه داده مورد استفاده و نحوه آموزش آن‌ها بیان گردیده است، در بخش ۴ معیارهای ارزیابی مورد استفاده

^۱ Region Based Convolutional neural network(R-CNN)

^۲ rotation

^۳ Mean Average Precision (mAP)

^۴ Ensemble learning

^۵ Internet of things(IOT)

^۶ Deep learning

امروزه بسیاری از پژوهش‌ها در این بخش صورت گرفته است

[۱۶].

الگوریتم‌های تشخیص شیء دومرحله‌ای دقت بالایی دارد، اما زمان زیادی طول می‌کشد تا از الگوریتم جستجوی انتخابی^۴ در حین تشخیص یک منطقه از تصویر عبور کنند، در مقابل الگوریتم‌های تشخیص شیء یک‌مرحله‌ای سریع‌تر اما دقت کمتری دارند [۱۷].

۲-۳- یادگیری انتقالی^۵

یادگیری عمیق پاسخ بسیاری از مشکلات یادگیری ماشین در دو دهه گذشته بوده است. با این حال، با دو محدودیت قابل توجه همراه است: وابستگی به داده‌های برچسب‌گذاری شده گسترده و هزینه‌های آموزش. یادگیری انتقالی در یادگیری عمیق که به عنوان یادگیری انتقالی عمیق^۶ شناخته می‌شود، تلاش می‌کند تا با استفاده مجدد از دانش به‌دست‌آمده از داده‌های منبع، این اتکا بودن به داده و هزینه‌ها را کاهش دهد. بیشتر روش‌های یادگیری انتقالی عمیق، رویکردهای مبتنی بر شبکه یا مدل هستند. این روش‌ها وابستگی مدل‌های یادگیری عمیق به داده‌های آموزشی گسترده را کاهش داده و هزینه‌های آموزشی را به شدت کاهش می‌دهد [۱۸].

روش‌های یادگیری انتقالی به دو روش زیر تقسیم می‌شوند:

(۱) روش تنظیم دقیق^۷

(۲) روش مبتنی بر ویژگی^۸

روش تنظیم دقیق، با یک مدل پیش آموزش‌دیده^۹ شروع شده و سپس تمام پارامترهای مدل برای کار جدید که مدنظر است به‌روز می‌شود، به عبارتی دیگر در اصل کل مدل مجدداً آموزش می‌بیند. از طرف دیگر در روش مبتنی بر ویژگی، با یک مدل پیش آموزش‌دیده شروع شده و فقط وزن لایه‌های نهایی را که پیش-بینی‌ها از آن استخراج می‌گردد، به‌روز می‌گردد، به طور مثال از شبکه عصبی پیچشی که از قبل آموزش‌دیده به‌عنوان یک استخراج‌کننده ویژگی به‌صورت ثابت استفاده می‌شود و فقط لایه خروجی آن تغییر می‌یابد [۱۹].

و دادگان^۱ عمومی بیشترین استفاده را در تحقیقات به خود اختصاص داده‌اند که از جمله آن‌ها می‌توان به COCO [۱۳]، ImageNet [۱۴]، اشاره نمود که از مجموعه‌ای از دادگان عمومی تشکیل شده‌اند.

چالش‌های متفاوتی در تشخیص اشیاء وجود دارد که باعث می‌شود طراحی و پیاده‌سازی مدل برای تشخیص اشیاء با مشکل مواجه شود. یکی از این مشکلات تغییراتی می‌باشد که درون یک کلاس رخ می‌دهد به‌طور مثال مطابق شکل (۲) اعضاء موجود در دسته کلاه ایمنی می‌توانند در زمینه‌های مختلفی از جمله رنگ، جنس و اندازه با یکدیگر تفاوت‌هایی داشته باشند.



شکل (۲): نمونه تغییرات درون دسته کلاه ایمنی

باتوجه به اینکه دادگان تصاویر از دوربین‌ها استخراج می‌شوند، دادگان حاصل از دوربین از جنبه‌های مختلفی از جمله فاصله دوربین از شیء، مقدار نوری که به آن شیء رسیده است همچنین زاویه قرارگیری آن نسبت به دوربین، می‌تواند متفاوت باشد که در شکل (۲) برخی از آن‌ها نمایان شده است.

۲-۲- روش‌های تشخیص اشیاء

به‌طور کلی دو روش جهت تشخیص اشیاء وجود دارد [۱۵]:

(۱) روش‌های مبتنی بر تشخیص شیء دومرحله‌ای^۲

(۲) روش‌های مبتنی بر تشخیص شیء یک‌مرحله‌ای^۳

• روش‌های مبتنی بر تشخیص شیء دومرحله‌ای:

در الگوریتم‌های تشخیص شیء دومرحله‌ای در مرحله اول باید یک پیشنهاد برای تشخیص مناطقی که احتمال وجود شیء در آن است استخراج شود و در مرحله دوم موقعیت‌یابی و طبقه‌بندی منطقه موردنظر صورت گیرد [۱۶].

• روش‌های مبتنی بر تشخیص شیء یک‌مرحله‌ای:

در الگوریتم‌های تشخیص شیء یک‌مرحله‌ای به طور مستقیم موقعیت‌یابی و طبقه‌بندی منطقه موردنظر صورت می‌گیرد، در نتیجه نیازی به یک پیشنهاد جهت تشخیص مناطق نمی‌باشد

^۴ selective search algorithm

^۵ Transfer Learning

^۶ Deep Transfer Learning

^۷ Fine-tuning

^۸ feature extraction

^۹ pretrained

^۱ dataset

^۲ two-stage object detection

^۳ one-stage object detection

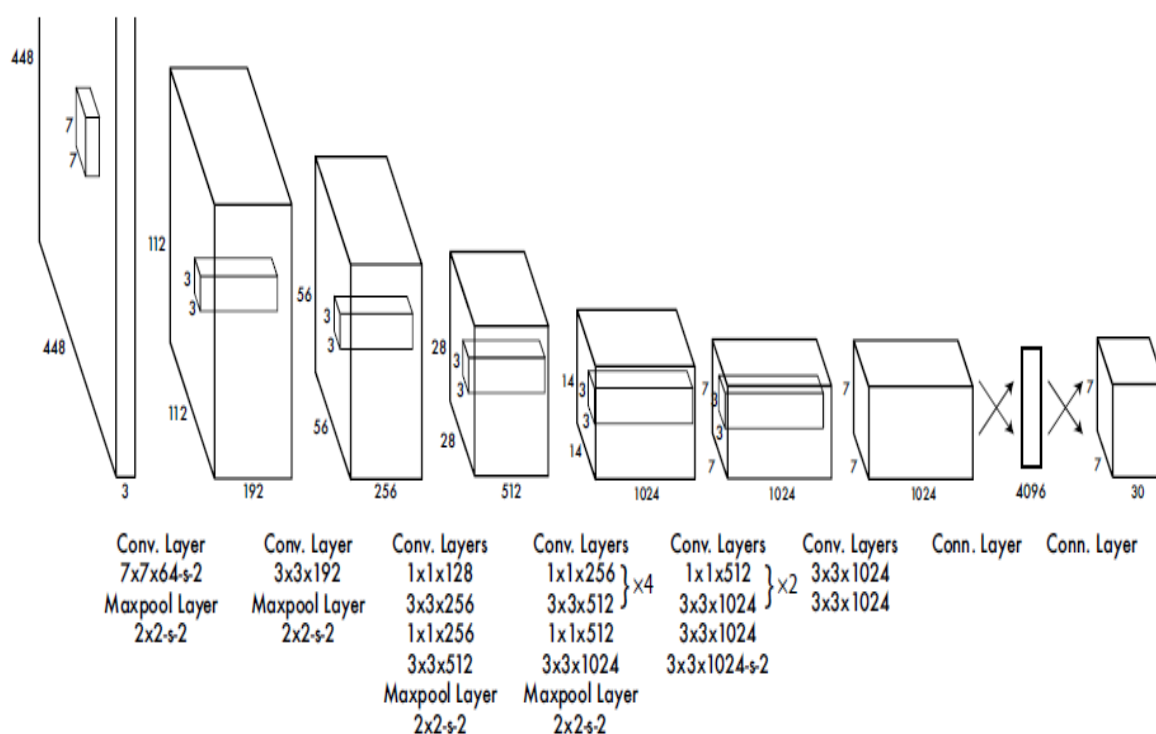
چگونه باهم تعامل دارند، سیستم بینایی انسان سریع و دقیق است و این امکان را می‌دهد تا کارهای پیچیده‌ای مانند رانندگی آگاهانه صورت بگیرد و الگوریتم YOLO بر اساس سیستم بینایی انسان طراحی شده است [۲۰]. در مقاله ارائه شده YOLO همان‌طور که در شکل (۳) مشاهده می‌شود. شبکه YOLO دارای ۲۴ لایه شبکه عصبی پیچشی می‌باشد که در انتها به دولایه کاملاً متصل^۳ ختم شده است، لایه‌های شبکه عصبی پیچشی 1×1 متناوب، فضای ویژگی‌ها را نسبت به لایه‌های قبلی کاهش می‌دهند. [۲۰]. از یک طرف عکس موردنظر وارد شبکه YOLO می‌شود و توسط یکسری شبکه عصبی پیچشی دست‌به‌دست می‌شود و هرکدام از این شبکه‌ها به یکسری ویژگی‌ها دست پیدا کرده و در نهایت در انتها توسط دولایه کاملاً متصل تشخیص می‌دهد تصویر متعلق به کدام دسته است.

۳-روش پیشنهادی

با عنایت به لزوم تشخیص استفاده یا عدم استفاده موتورسواران از کلاه ایمنی به صورت برخط^۱، همچنین سرعت سریع عبور آن‌ها از دید دوربین‌های ترافیکی از روش‌های مبتنی بر تشخیص شیء یک‌مرحله‌ای و به عبارتی از شبکه عصبی YOLO [۲۰] که کوتاه شده "You Only Look Once" به معنای "شما تنها یک بار نگاه می‌کنید" است، استفاده شده است. در ادامه جزئیات شبکه استفاده شده بیان می‌شود.

۳-۱- شبکه YOLO_V3

الگوریتم YOLO نخستین بار در سال ۲۰۱۶ در کنفرانس CVPR^۲ ارائه گردید. انسان‌ها به یک تصویر نگاه می‌کنند و فوراً می‌دانند که چه چیزهایی در تصویر وجود دارد، کجا هستند و



شکل (۳): معماری اولیه شبکه YOLO

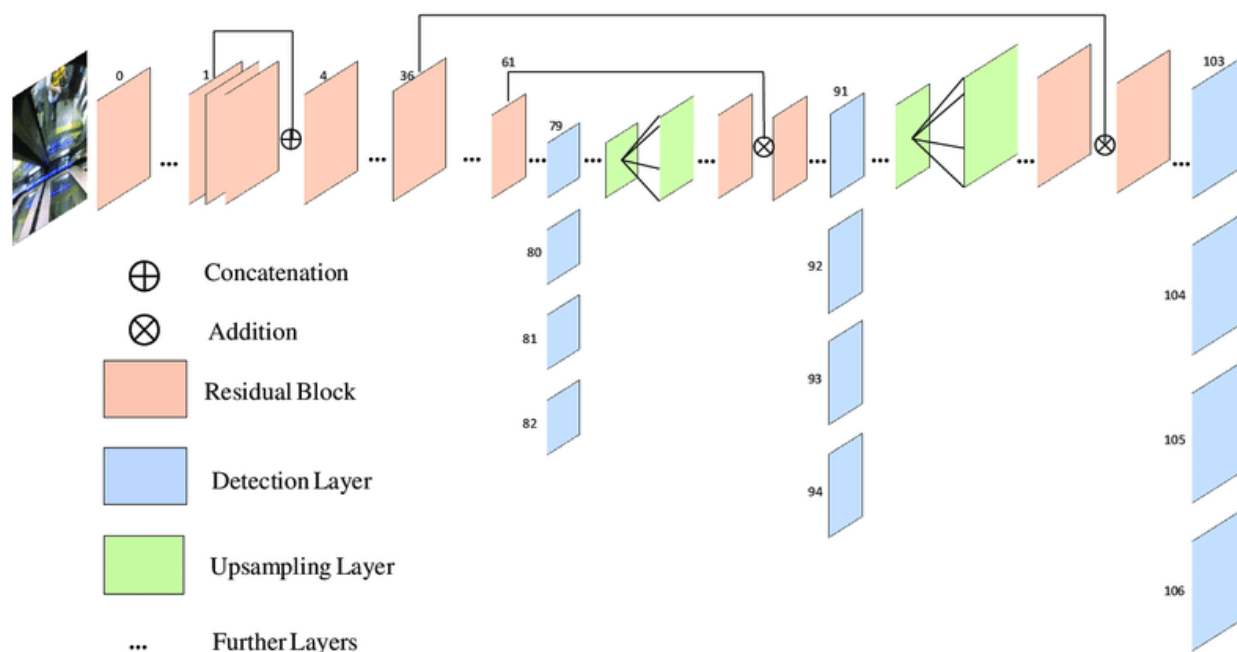
³ fully connected

¹ online

² Conference on Computer Vision and Pattern Recognition

یک ارزیابی شبکه اشیاء را پیش‌بینی می‌کند، این باعث می‌شود که YOLO بسیار سریع باشد، شبکه YOLO بیش از ۱۰۰۰ برابر سریع‌تر از شبکه عصبی پیچشی مبتنی بر منطقه و ۱۰۰ برابر سریع‌تر از مدل Fast آن است [۲۱].

شبکه YOLO چندین مزیت نسبت به سامانه‌های مبتنی بر طبقه‌بندی دارد. در زمان تست به کل تصویر نگاه می‌کند تا پیش‌بینی‌های آن بر اساس یک زمینه جهانی در تصویر مشخص شود، همچنین برخلاف سامانه‌هایی مانند شبکه عصبی پیچشی مبتنی بر منطقه که برای یک تصویر به هزاران الزام نیاز دارند، با



شکل (۴): معماری شبکه YOLO_v3 [۲۲]

همان‌طور که در شکل (۴) مشاهده می‌شود YOLO_v3 که در این پژوهش مورد استفاده قرار گرفته است [۲۲] نسبت به YOLO اولیه (شکل (۳)) از لحاظ ساختاری دارای تفاوت‌هایی است. در YOLO_v3 تصویر وارد شبکه شده و توسط یکسری شبکه‌های عصبی پیچشی دست‌به‌دست می‌شود و در نهایت یک ماتریس از پیش‌بینی اشیاء به عنوان خروجی داده می‌شود. در YOLO_v3 مطابق شکل (۴) سه پیش‌بینی داده می‌شود و هر کدام از این پیش‌بینی‌ها عکس را با یک اندازه^۱ مشخصی گرید بندی^۲ می‌کنند، علت این کار وجود اشیاء با اندازه‌های متفاوت در تصویر می‌باشد، در نتیجه هر کدام از این سه پیش‌بینی دارای دقت بالاتری باشد آن پیش‌بینی به عنوان پیش‌بینی نهایی انتخاب می‌گردد.

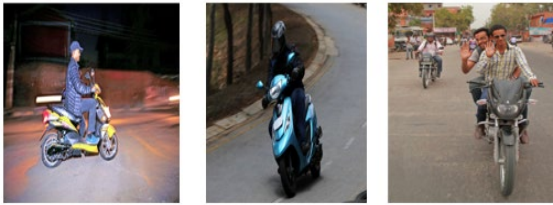
YOLO_v3 همانند دیگر نسخه‌های شبکه YOLO بر روی داده‌گان COCO آموزش دیده است و شامل پنج معماری مختلف است که در شکل (۵) نمایش داده شده است.

YOLOv3-320 mAP=51.5 , FPS=45	YOLOv3-416 mAP=55.3 , FPS=35
YOLOv3-608 mAP=57.9 , FPS=20	YOLOv3-tiny mAP=33.1 , FPS=220
YOLOv3-spp mAP=60.6 , FPS=20	

شکل (۵): معماری‌های مختلف YOLO_v3

^۱ size
^۲ grade

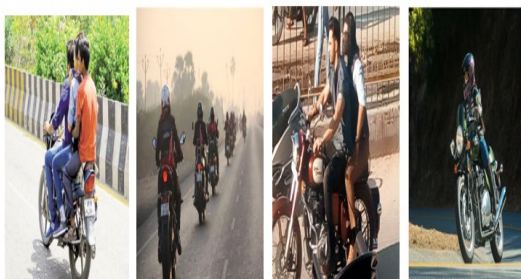
(۲) پوشش سروصورت موتورسوار: یکی دیگر از چالش‌های الگوریتم‌های تشخیص اشیاء، وجود مانع در تشخیص تصاویر



شکل (۶): تصاویر موتورسواران در حالت‌های مختلف نوری هدف می‌باشد که باعث می‌شود کل یا قسمتی از تصویر در دید دوربین قرار نگیرد. در مورد موتورسواران می‌تواند این مانع رخ دهد. به‌طور مثال این مانع می‌تواند به علت پوشاندن سروصورت یا قرار گرفتن موتورسواری در پشت یک خودرو رخ دهد. همان‌طور که در شکل (۷) مشاهده می‌کنید در دادگان Helmet Detection به این مورد توجه شده است.



شکل (۷): تصاویر موتورسواران در حالت با مانع (۳) فاصله موتورسواران تا دوربین: در دادگان Helmet Detection به فاصله تصاویر تا دوربین توجه شده و از تصاویر مختلف از نظر فاصله استفاده گردیده است و شکل (۸) نمونه‌هایی از این تصاویر می‌باشد.



شکل (۸): تصاویر موتورسواران در فاصله‌های مختلف (۴) زاویه موتورسواران نسبت به دوربین: با عنایت به اینکه زاویه قرار گرفتن موتورسواران نسبت به دوربین اشکال مختلفی از کلاه ایمنی را نشان می‌دهد، در دادگان موجود این موضوع مورد توجه قرار گرفته است همان‌طور که در شکل (۹) مشاهده می‌شود از تصاویر با زاویه‌های مختلف در دادگان استفاده شده

در این تحقیق سه نسخه معماری YOLOv3-416، YOLOv3-320 و YOLOv3-spp استفاده شده است، قسمتی از معماری‌های نامبرده شده به‌وسیله مجموعه داده Helmet Detection که در بخش ۲-۳ معرفی خواهد شد، مجدد آموزش دیده‌اند و کارایی آن‌ها با یکدیگر مقایسه شده است، نتایج مقایسه این سه معماری در بخش ۵ عنوان گردیده است.

۲-۳- معرفی مجموعه داده

یکی از چالش‌های مهم در پیاده‌سازی الگوریتم‌های مبتنی بر هوش مصنوعی یا به طور اختصاصی یادگیری عمیق، در دسترس نبودن مجموعه داده کافی جهت آموزش شبکه عصبی است، جهت انجام این تحقیق از دادگان Helmet Detection [۲۳] با مقدار تغییر استفاده شده است، به علت اینکه عکس‌های استفاده شده در این دادگان شامل افرادی است که از وسیله‌هایی مختلفی مانند دوچرخه، موتور، اسکوتر و... استفاده کرده، لذا با توجه به اینکه این تحقیق صرفاً برای تشخیص استفاده یا عدم استفاده موتورسوار از کلاه ایمنی می‌باشد، تنها تصاویر مربوط به موتورسواران از آن استخراج گردیده است و از طرفی چون عملیات برچسب‌زنی به خوبی صورت نگرفته بود مجدداً توسط ابزار labelImg [۲۴] عملیات برچسب‌زنی صورت گرفت.

همچنین با عنایت به اینکه بر اساس سه مورد (۱) زاویه موتورسوار نسبت به دوربین (۲) محیط نوری متفاوت (۳) فاصله موتورسوار تا دوربین، سیستم پیشنهادی باید توانایی تشخیص اشخاص در دو حالت با کلاه ایمنی یا فاقد کلاه ایمنی را داشته باشد، تغییراتی در برخی عکس‌ها بر اساس سه مورد فوق همانند تغییر شرایط نوری و ... صورت گرفت تا از هر سه شرایط در حالت‌های مختلف دادگانی وجود داشته باشد.

توجه به شرایط تصویربرداری یک عامل اساسی در جمع‌آوری تصاویر است. به طور مثال، زاویه مختلف دوربین نسبت به اشیاء می‌تواند در دقت تشخیص آن شیء تأثیرگذار باشد، دقت تشخیص صورت یک فرد که دوربین در مقابل چهره او است نسبت به دقت تشخیص صورت یک فرد که دوربین در کنار صورت او است بیشتر است. در مورد تصاویر استفاده شده در دادگان که در این تحقیق استفاده شده است موارد زیر قابل توجه است.

(۱) شرایط نوری: با عنایت به اینکه در دنیای واقعی اشیاء در شرایط مختلف نوری قرار می‌گیرند، داده‌ها در شرایط مختلف نوری جمع‌آوری شده است. البته باید توجه داشت که دوربین‌های ترافیکی در حالت کم‌نور یا در هنگام شب باید مجهز به حالت دید در شب باشند یا شیء موردنظر در نور کافی قرار داشته باشد و مطابق شکل (۶) در مجموعه داده Helmet Detection تصاویر در شرایط مختلف نوری قرار دارند.

این تصاویر به مدل داده می‌شود و باید طوری اشیاء توسط مدل آموزش ببینند که محدود اشیاء و کلاس آن را تشخیص دهد.



شکل (۱۱): محاسبه مقدار خطا

مطابق شکل (۱۱) در ابتدای آموزش، مدل اجرا می‌شود و قرار است مدل محدود سبز رنگ را که چهره یک بازیکن است یاد بگیرد اما به اشتباه شبکه محدود قرمز رنگ را مشخص می‌کند و باید مقدار این خطا به دست بیاید. به عبارتی باید مقدار دقت با پارامتری اندازه گرفته شود تا بیان شود به چه اندازه‌ای مدل دقیق کار کرده است و این دقت توسط شاخصی تحت عنوان mAP^3 تعیین می‌گردد.

در میان مجموعه داده‌های مختلف جهت تشخیص شیء که مورد استفاده قرار گرفته است و در بین جامعه علمی رایج‌ترین و مرسوم‌ترین معیاری که برای اندازه‌گیری دقت تشخیص اشیاء استفاده می‌شود، AP است [۲۵]. این معیار برای یک دسته از داده محاسبه می‌شود و لازم است برای کلیه دسته‌ها محاسبه شود در نتیجه باید از میانگین آن‌ها نسبت به تمامی کلاس‌ها استفاده کرد که به میانگین آن mAP می‌گویند [۲۶].

جهت ورود به بحث mAP دانستن چند مفهوم الزامی است. اولین مفهوم IOU^4 می‌باشد همان‌طور که در شکل (۱۲) مشاهده می‌شود مدل باید خودرو را با محدوده مستطیل سبز مشخص کند اما آنچه شبکه پیش‌بینی می‌کند محدوده مستطیل قرمز رنگ است. می‌توان گفت شبکه تا حدودی درست تشخیص داده اما به‌طور کامل این تشخیص رخ نداده است و در محاسبه IOU آن قسمتی که هر دو مستطیل در آن همپوشانی^۵ دارد در صورت محاسبه کسر IOU قرار می‌گیرد و در مخرج محاسبه کسر IOU اجتماع^۶ هر دو مستطیل قرار خواهد گرفت، به عبارت دیگر

است.

۳-۳- شبکه استفاده شده

شبکه‌ای که در این پژوهش به عنوان شبکه پیش



شکل (۹): تصاویر موتورسواران در زاویه‌های مختلف

آموزش داده شده استفاده گردیده است، شبکه YOLO نسخه ۳ می‌باشد. از آنجایی که در این شبکه از مجموعه داده COCO برای آموزش استفاده شده است لذا تعداد تصاویر موجود در دادگان Helmet Detection برای آموزش لایه‌های آخر شبکه عصبی کافی می‌باشد و تنها جهت استخراج ویژگی‌های سطح پایین از شبکه YOLO استفاده گردیده است.

۴- معیار ارزیابی

در کلیه دادگان مورد استفاده جهت تشخیص اشیاء توسط شبکه YOLO تعدادی تصویر وجود دارد که در هر تصویر، تمامی اشیاء که جهت تشخیص باید استفاده شوند لازم است محدوده آن‌ها توسط Bounding Box باهدف برچسب‌زنی مشخص گردد. همان‌طور که در شکل (۱۰) دیده می‌شود عملیات برچسب‌زنی توسط Bounding Box برای دودسته قایق^۱ و شخص^۲ صورت گرفته است.



شکل (۱۰): برچسب زنی اشیاء توسط Bounding Box [۲۵]

³ mean Average Precision

⁴ Intersection over Union

⁵ overlap

⁶ union

¹ boat

² person

AP برای کلاس K است.

۵- تحلیل نتایج

در این تحقیق از سه معماری YOLOv3-416، YOLOv3-320 و YOLOv3-spp استفاده شده و معماری‌های عنوان شده توسط مجموعه داده Helmet Detection آموزش دیده و ارزیابی گردیده است، لازم به ذکر است که آموزش در بستر رایگان گوگل (google colab [۳۳]) صورت گرفته است، همچنین در جدول (۱) هایپر پارامترهای استفاده شده به صورت خلاصه آورده شده است.

جدول (۱): نمونه‌ای از هایپر پارامترهای آموزش

آستانه IOU	نرخ یادگیری	تصاویر آموزش	تصاویر ارزیابی
0.02	0.001	289	96

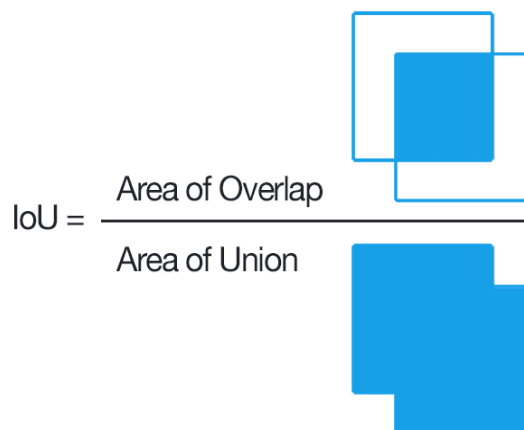
با توجه به جدول (۱)، از میان تصاویر دادگان Helmet Detection، تعداد ۲۸۹ تصویر برای آموزش و ۹۶ تصویر برای ارزیابی در نظر گرفته شده است. لازم به ذکر است در هر کدام از این تصاویر بین ۱ تا ۸ فرد دارای کلاه ایمنی یا فاقد کلاه ایمنی یافت می‌شود که در فرآیند آموزش یا ارزیابی از آن‌ها استفاده شده است. حد آستانه IOU برابر با ۰٫۲ قرار داده شده است و نرخ یادگیری هم ۰٫۰۰۱ در نظر گرفته شده است. با عنایت به اینکه YOLO_v3 دارای ۱۰۶ لایه می‌باشد، در هر سه معماری YOLOv3-416، YOLOv3-320، YOLOv3-spp و از ۵۳ لایه اول^۷ آن‌ها به صورت یادگیری انتقالی استفاده شده است و به اصطلاح این لایه‌ها Freeze شده‌اند و در ادامه ۵۳ لایه دیگر بر اساس دادگان Helmet Detection مورد آموزش قرار گرفته است.

جدول (۲): ماتریس ابهام^۸

برچسب پیش‌بینی شده		برچسب اصلی	
		مثبت (دارای کلاه ایمنی)	منفی (فاقد کلاه ایمنی)
مثبت (دارای کلاه ایمنی)	مثبت (دارای کلاه ایمنی)	174	8
	منفی (فاقد کلاه ایمنی)	13	158

در جدول (۲) ماتریس ابهام برای هر دو کلاس که شامل کلاه ایمنی یا فاقد کلاه ایمنی هستند ارائه شده است.

اشتراک دو مستطیل تقسیم بر اجتماع آن‌ها مقدار IOU می‌باشد و عددی است بین صفر و یک، در نتیجه هنگامی که عدد یک باشد یعنی دقیقاً شیء مورد نظر را درست تشخیص داده است و هنگامی که عدد صفر باشد یعنی مدل کاملاً اشتباه تشخیص داده است و در بسیاری از موارد عدد ۰٫۵ را به عنوان آستانه^۱ در نظر می‌گیرند یعنی بیشتر از ۰٫۵ قبول است و کمتر از ۰٫۵ خوب



شکل (۱۲): تشریح مفهوم IOU [۲۷]

نمی‌باشد.

با توجه به مفهوم IOU [۲۷] سه مفهوم TP^2 ، FP^3 و FN^4 به وجود می‌آید [۲۸] و می‌توان از روی این سه مفهوم به مفاهیم بازخوانی^۵ [۲۹] و دقت^۶ [۳۰] دست پیدا کرد.

مطابق فرمول زیر با استفاده از دو مفهوم بازخوانی و دقت می‌توان AP را به دست آورد.

$$AP = \frac{\sum_{k=0}^{K-1} [\text{Recall}(K) - \text{Recall}(K+1)]}{\text{Precision}(K)} \quad (1)$$

در فرمول ۳ [۳۱]، $\square$ برابر با آستانه است و با توجه به معیار mAP به صورت زیر محاسبه می‌شود.

$$mAP = \frac{1}{n} \sum_{k=1}^n AP_k \quad (2)$$

در فرمول ۲ [۳۲]، n تعداد کلاس اشیاء می‌باشد و AP_k معیار

¹ threshold

² True Positive

³ False Positive

⁴ False Negative

⁵ Recall

⁶ Precision

⁷ darknet-53

⁸ Confusion Matrix

جدول (۳): نتایج مربوط به آموزش معماری‌های مختلف

شبکه	دقت	بازخوانی	mAP
YOLOv3-416	0.88	0.90	86.72
YOLOv3-320	0.92	0.96	96.03
YOLOv3-spp	0.93	0.96	96.08



معیارهای مورد استفاده برای ارزیابی قبلاً در بخش ۴ عنوان گردیده است، مطابق نتایج به دست آمده از معماری‌های مختلف YOLOv3 در جدول (۳)، YOLOv3-spp بهترین عملکرد را داشته و توانسته است به مقدار mAP ۹۶,۰۸ برسد. همچنین بهره بردن از یادگیری انتقالی و استفاده از دانش شبکه آموزش دیده بر روی دادگان COCO، توانسته نتایج قابل قبولی را جهت مسئله جدید در اختیار قرار دهد.

در YOLOv3-spp هدف بهبود عملکرد تشخیص شیء یک مرحله‌ای YOLOv3 با افزودن یک لایه SPP در انتهای لایه ۵۳ است تا فرآیند استخراج ویژگی کارآمدتر گردد. بر اساس تحقیقات انجام شده توسط هی و دوستانش [۳۴] مدل SPP-net قادر به تولید خروجی با طول ثابت، بدون توجه به اندازه ورودی است و همین عامل باعث می‌شود تا فرآیند استخراج ویژگی بهبود یابد، گرچه ممکن است بهبود در عملکرد YOLOv3-spp نسبت به دیگر مدل‌های YOLOv3 در حد چند دهم یا حتی کمتر باشد.

در شکل (۱۳) نمونه‌هایی از نتایج سیستم پیشنهادی بر روی دادگان آزمون Helmet Detection را مشاهده می‌کنید. در این تصاویر افراد در زاویه، فاصله و میزان شدت نور مختلف حضور دارند و شبکه توانسته در تمامی حالت‌های مختلف عنوان شده حتی با پوشاندن چهره نتایج مطلوبی را ارائه دهد همچنین در این شکل "A" به معنای استفاده موتورسوار از کلاه ایمنی و "b" به معنای عدم استفاده موتورسوار از کلاه ایمنی می‌باشد و در این سیستم تعداد بسیار محدودی از تصاویر توسط سیستم پیشنهادی جزء هیچ کدام از کلاس‌ها تشخیص داده نشدند که نمونه آن در آخرین تصویر شکل (۱۳) قابل مشاهده است در این تصویر دو فرد سوار بر موتور هستند که سیستم پیشنهادی یکی از آن‌ها را درست تشخیص داده (استفاده موتورسوار از کلاه ایمنی) و برای دیگری کلاسی در نظر نگرفته است در واقع از میان ۹۶ تصویر که شامل ۳۵۳ فرد موتور سوار با کلاه ایمنی یا فاقد کلاه ایمنی هستند، تنها ۲۱ تصویر به اشتباه کلاس آن‌ها تشخیص داده شده یا کلاس آن‌ها به طور کلی تشخیص داده نشده است.

شکل (۱۳): نتایج سیستم تشخیص کلاه ایمنی بر روی داده‌های آزمون

جدول (۴): نتایج پیاده‌سازی دیگر روش‌ها و روش ارائه شده در این پژوهش

ردیف	پژوهشگران	مدل استفاده شده	سال ارائه مدل	دادگان مورد استفاده	معیار و نتایج ارزیابی
(۱)	آرمسترانگ آبواه و همکارانش [۳۵]	Few-ShotData sampling + YOLOv8	۲۰۲۳	2023 NVIDIA AI CITY CHALLENGE	mAP:58.61
(۲)	الهام سلطانی کاظمی [۳۶]	Genetic Algorithm-Enhanced YOLOv5	۲۰۲۳	2023 NVIDIA AI CITY CHALLENGE	mAP:66.67
(۳)	مادوچاندا داسگوپتا و همکارانش [۳]	CNN	۲۰۱۹	جمع‌آوری شده از سطح اینترنت	Accuracy:91.08
(۴)	سوتیکنو و همکارانش [۱۰]	Histogram of Oriented Phase and Gradient- Local Difference Binary (HOPG-LDB)	۲۰۲۱	JSC1	AP: 71.21
				JSC2	AP: 66.63
(۵)	مدل پیشنهادی	Yolo3	۲۰۲۴	Helmet Detection سفارشی شده	mAP:96.08

جهت اطمینان از مناسب بودن عملکرد مدل و دادگان مورد استفاده، در این پژوهش مدل‌های عنوان شده در جدول (۴) بر روی دادگان Helmet Detection اعمال و نتایج مطابق جدول (۵) حاصل گردید. مطابق این جدول مدل پیشنهادی عملکرد بهتری دارد و نسخه YOLOv3-spp به دلیل افزودن یک لایه SPP در انتهای لایه ۵۳، فرآیند استخراج ویژگی را کارآمدتر کرده است

مطابق نتایج به دست آمده در جدول (۴) مدل ارائه شده در این پژوهش به نسبت دیگر مدل‌های ارائه شده عملکرد مناسبی دارد. با توجه به اینکه دادگان Helmet Detection مجدداً در این پژوهش با دقت بالایی برچسب‌زنی گردیده است، همچنین تغییراتی در برخی عکس‌ها همانند تغییر شرایط نوری و ... صورت گرفت، لذا یادگیری مدل و عملکرد مناسب آن بر روی داده‌های آزمون مشاهده می‌گردد.

جدول (۵): مقایسه نتایج روش‌های مختلف ارائه شده

ردیف	پژوهشگران	مدل استفاده شده	سال ارائه مدل	دادگان مورد استفاده	معیار و نتایج ارزیابی
(۱)	آرمسترانگ آبواه و همکارانش [۳۵]	Few-ShotData sampling + YOLOv8	۲۰۲۳	2023 NVIDIA AI CITY CHALLENGE	mAP:67.18
(۲)	الهام سلطانی کاظمی [۳۶]	Genetic Algorithm-Enhanced YOLOv5	۲۰۲۳	2023 NVIDIA AI CITY CHALLENGE	mAP:74.23
(۳)	مادوچاندا داسگوپتا و همکارانش [۳]	CNN	۲۰۱۹	جمع‌آوری شده از سطح اینترنت	Accuracy:93.27
(۴)	سوتیکنو و همکارانش [۱۰]	Histogram of Oriented Phase and Gradient- Local Difference Binary (HOPG-LDB)	۲۰۲۱	JSC1	AP: 83.74
				JSC2	AP: 79.56
(۵)	مدل پیشنهادی	Yolo3	۲۰۲۴	Helmet Detection سفارشی شده	mAP:96.08

کلاه ایمنی را شناسایی کند. با توجه به اینکه یکی از مهم‌ترین مسائل جهت کاهش مصدومیت حوادث رانندگی تحلیل میزان استفاده و عدم استفاده از کلاه ایمنی توسط موتورسواران می‌باشد، سیستم پیشنهادی در این پژوهش می‌تواند در زمینه پیشگیری از وقوع تخلفات رانندگی که منجر به حوادث می‌شود از طریق استفاده از تحلیل داده، کمک شایانی نموده و دولت‌ها با دانش بیشتری در زمینه کاهش وقوع تخلفات برنامه‌ریزی نمایند.

جهت تشخیص بلادرنگ استفاده یا عدم استفاده موتورسواران از کلاه ایمنی، از دوربین‌های ترافیکی که بیشتر مواقع تصاویر در

۶- نتیجه‌گیری

کاهش مصدومیت، همچنین مرگ و میر ناشی از حوادث رانندگی یکی از اهداف تمامی دولت‌ها و بخش‌های انتظامی کشورها می‌باشد زیرا می‌تواند هزینه زیادی را به دنبال داشته باشد، در نتیجه همواره در راستای این هدف کارهایی صورت گرفته است. سیستم تشخیص استفاده از کلاه ایمنی توسط موتورسواران می‌تواند به صورت بلادرنگ و زنده از طریق دوربین‌های ترافیکی، محل‌های عبور و مرور و وسایل نقلیه همانند چهارراه‌ها و اتوبان‌ها را مورد پایش قرار داده و استفاده یا عدم استفاده موتورسواران از

- DOI:10.1109/CICT48419.2019.9066191.
- [4]. M. A. , M. B. A. , T. M. , Z. M. S. A. A. Tasbeeha Waris, "CNN-Based Automatic Helmet Violation Detection of Motorcyclists for an Intelligent Transportation System," *Hindawi*, 2022, DOI:10.1155/2022/8246776.
- [5]. E. Soltanikazemi, A. Aboah, E. Arthur and B. K. Hatuwal, "Fine-Tuning YOLOv5 with Genetic Algorithm For Helmet Violation Detection," *arXiv*, 2023.
- [6]. V. C. K. A.-N. , P. A. L. A. O. A. Geoffery Agorku, "Real-Time Helmet Violation Detection Using YOLOv5 and Ensemble Learning," *arXiv*, 2023.
- [7]. K. S. , R. H. Sherin Eliyas, "Helmet, Violation, Detection Using Deep Learning," *European Journal of Molecular & Clinical Medicine*, 2019.
- [8]. M. N. D. G. Vignesh Raj A G, "Helmet Detection using Single Shot Detector (SSD)," *IEEE*, 2021, DOI:10.1109/ICESC51422.2021.9532985.
- [9]. A. A. T. Tanupriya Choudhury, "A Deep Learning Approach to Helmet Detection for Road Safety," *NISCAIR*, 2020, DOI:10.56042/jsir.v79i06.39579.
- [10]. A. H. Sutikno, "Improving Detection Performance of Helmetless Motorcyclists Using the Combination of HOG, HOP, and LDB Descriptors," *INASS*, p. 428, 2021, DOI:10.22266/ijies2022.0228.39.
- [11]. N. R. A. G. A. B. K. S. R. Kurkute, "IOT Based Smart System for the Helmet Detection," in *International Conference on Sustainable Computing in Science, Technology & Management (SUSCOM-2019)*, 2019, DOI:10.2139/ssrn.3356793.
- [12]. P. Z.-T. X. W. Zhong-Qiu Zhao, "Object Detection with Deep Learning: A Review," *IEEE*, 2019, DOI:10.1109/TNNLS.2018.2876865.
- [13]. M. M. S. B. L. B. R. G. Tsung-YiLin, "Microsoft COCO: Common Objects in Context," *arxiv*, 2014.
- [14]. W. D. R. S. L.-J. L. K. L. L. F.-F. Jia Deng, "ImageNet: A large-scale hierarchical image database," *IEEE*, 2009, DOI:10.1109/CVPR.2009.5206848.
- [15]. K. D. K. R. R. J. D. A. M. B. Aditya Lohia, "Bibliometric Analysis of One-stage and Two-stage Object Detection," 2021.
- [16]. R. Z. W. Lixuan Du, "Overview of two-stage object detection algorithms," *ICSP 2020*, 2020, DOI:10.1088/1742-6596/1544/1/012033.
- [17]. R. S. C. Hang Zhang, "Review on One-Stage

آن به صورت واضح نمی‌باشد استفاده شده است. در این پژوهش اقدام به ساخت یک مجموعه داده‌دگان و برچسب‌زنی آن‌ها با استفاده از دادگان Helmet Detection شد. تصاویر موجود در دادگان دارای شرایطی از جمله شرایط نوری مختلف (همانند روز و شب)، پوشش سروصورت موتورسوار، فاصله مختلف موتورسواران تا دوربین و زاویه موتورسواران نسبت به دوربین (بالا، کنار و...) است، سیستم ارائه شده در این پژوهش توانایی شناسایی بلادرنگ استفاده یا عدم استفاده موتورسواران از کلاه ایمنی در این شرایط را دارد.

به دلیل شناسایی سریع، در این پژوهش از یک روش مبتنی بر تشخیص شیء یک مرحله‌ای تحت عنوان YOLOv3 استفاده شده که دارای ۵ نسخه مختلف است. در ادامه نتیجه گرفتیم که دادگان COCO شامل یک سری از دسته‌ها است که به دادگان Helmet Detection نزدیک نیستند، در نتیجه تعدادی از لایه‌های شبکه به دلیل مناسب نبودن آن‌ها جهت مسئله تشخیص کلاه ایمنی دوباره تحت آموزش قرار گرفت. دقت نسخه YOLOv3-416 مقدار $(mAP)/.۸۶/۷۲$ ، نسخه YOLOv3-320 مقدار $(mAP)/.۹۶/۰۳$ و نسخه YOLOv3-spp مقدار $(mAP)/.۹۶/۰۸$ است. دو نسخه YOLOv3-320 و YOLOv3-spp نتایج قابل قبولی را ارائه داده است؛ اما نسخه YOLOv3-spp به دلیل افزودن یک لایه SPP در انتهای لایه ۵۳، فرآیند استخراج ویژگی را کارآمدتر کرده است، در نتیجه پیشنهاد می‌شود از نسخه YOLOv3-spp استفاده گردد.

تشخیص استفاده یا عدم استفاده موتورسواران از کلاه ایمنی می‌تواند چالش‌های متفاوتی را به دنبال داشته باشد. برای مثال در تاریکی مطلق در هنگام شب، تشخیص دشوار یا حتی غیرممکن است، راه حل این موضوع استفاده از دوربین‌های دید در شب است؛ اما تصویرهای دریافتی از این دوربین‌ها با تصاویر عادی متفاوت است، در نتیجه مجموعه داده‌دگان جدید متناسب با این شرایط باید تولید شود.

۷-منابع

- [1]. R. G. R. a. G. M. Kh Tohidul Islam, "Recognition of Traffic Sign Based on Bag-of-Words and Artificial Neural Network," *MDPI*, 2017, DOI:10.3390/sym9080138.
- [2]. who, "Global status report on road safety 2018," <https://www.who.int/publications/i/item/9789241565684>, 2018.
- [3]. O. B. , S. C. Madhuchhanda Dasgupta, "Automated Helmet Detection for Multiple Motorcycle Riders using CNN," in *2019 IEEE Conference on Information and Communication Technology (CICT)*, 2019 ,

- [34]. X. Z. S. R. a. J. S. Kaiming He, "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," *IEEE*, 2015, DOI:10.1109/TPAMI.2015.2389824.
- [35]. "Real-time Multi-Class Helmet Violation Detection Using Few-Shot Data Sampling Technique and YOLOv8," in *Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2023.
- [36]. A. D. ., B. ., I. E. T. Elham Soltanikazemi, "Real-Time Helmet Violation Detection in AI City Challenge 2023 with Genetic Algorithm-Enhanced YOLOv5," in *2023 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, 2023.
- Object Detection Based on Deep Learning," *EAI/EU*, 2022, DOI:10.4108/eai.9-6-2022.174181.
- [18]. H. R. A. K. R. Mohammadreza Iman, "A Review of Deep Transfer Learning and Recent Advancements," *MDPI*, 2023, DOI:10.3390/technologies11020040.
- [19]. pytorch, "FINETUNING TORCHVISION MODELS," https://pytorch.org/tutorials/beginner/finetuning_torchvision_models_tutorial.html, 2023.
- [20]. S. D. G. F. Joseph Redmon, "You Only Look Once: Unified, Real-Time Object Detection," *IEEE*, 2016, DOI:10.1109/CVPR.2016.91.
- [21]. "YOLO: Real-Time Object Detection," <https://pjreddie.com/darknet/yolo/>, 2023.
- [22]. W. L. L. ., L. YUAN DAI, "Efficient Foreign Object Detection between PSDs and Metro Doors via Deep Neural Networks," *IEEE*, 2020, DOI:10.1109/ACCESS.2020.2978912.
- [23]. "Helmet Detection," <https://www.kaggle.com/datasets/andrewmvd/helmet-detection>, 2023.
- [24]. <https://github.com/heartexlabs/labelImg>, 2023.
- [25]. S. L. N. A. B. d. S. Rafael Padilla, "A Survey on Performance Metrics for Object-Detection Algorithms," *IEEE*, 2020, DOI:10.1109/IWSSIP48289.2020.
- [26]. "Common objects in Context," <https://cocodataset.org/#detection-eval>, 2023.
- [27]. "Intersection over Union (IoU) for object detection," <https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/>, 2024.
- [28]. "Confusion Matrix," <https://wiki.cloudfactory.com/docs/mp-wiki/metrics/confusion-matrix>, 2024.
- [29]. "Recall score," 2024, <https://wiki.cloudfactory.com/docs/mp-wiki/metrics/recall>.
- [30]. "Precision score," <https://wiki.cloudfactory.com/docs/mp-wiki/metrics/precision>, 2024.
- [31]. "Average Precision," <https://hasty.ai/docs/mp-wiki/metrics/average-precision>, 2023.
- [32]. "Mean Average Precision (mAP) Explained: Everything You Need to Know," <https://www.v7labs.com/blog/mean-average-precision>, 2023.
- [33]. G. Colab, ""Google Colab", https://colab.research.google.com/notebooks/intro.ipynb?utm_source=scs-index, 2023.