مجله کنترل
ISSN (print) 2008-8345
ISSN (online) 2538-3752
مقاله پژوهشی
جلد۱۸، شماره ۲، تابستان ۱۴۰۳، صفحه ۹۴-۸۵

تشخیص بلادرنگ نوع خودرو با استفاده از طبقه‌بند مبتنی بر شبکه کانولوشن و ادغام موجک هار

سید محمد جوادی مقدم[۱]، حسین غلامعلی نژاد[۲]

[۱] استادیار، دانشکده مهندسی، گروه کامپیوتر، دانشگاه بزرگمهر قاینات، smjavadim@buqaen.ac.ir

[۲] استادیار، دانشکده مهندسی، گروه کامپیوتر، دانشگاه بزرگمهر قاینات ، h_gholamalinejad@buqaen.ac.ir

دریافت: ۱۴۰۲/۰۴/۰۸ ویرایش: ۱۴۰۳/۰۳/۱۹ پذیرش: ۱۴۰۳/۰۵/۱۷

چکیده: در سال‌های اخیر، طبقه‌بندی بلادرنگ نوع خودرو، به دلیل کاربردش در کنترل و تحلیل ترافیک، از موضوعات جذاب و بسیار پرکاربرد بوده است. با توجه به محدودیت تعداد نمونه‌ای آموزشی برچسب‌دار باکیفیت، تغییرات در وضعیت زاویه‌ای خودرو و دوربین، کیفیت و وضوح تصویر، نور و شرایط آب و هوایی، دقت تشخیص از چالش های مهم در سیستم های تشخیص نوع خودرو می باشد. در این مقاله، یک شبکه‌ی کانولوشن بلادرنگ جدید برای تشخیص نوع خودروها با دقت بالا معرفی شده‌است. در ساختار عصبی پیشنهادی، لایه‌ی ادغام جدید با استفاده از تبدیل موجک هار علاوه برکاهش ابعاد نقشه‌ی ویژگی خروجی از لایه‌ی کانولوشن، قابلیت استخراج ویژگی را نیز دارد. نوآوری دیگر ساختاری، استفاده از بلوک‌های فشار-تحریک[۱] قبل از لایه‌های کانولوشن میانی می‌باشد که موجب بالارفتن دقت بازشناسی می‌شود. علاوه‌براین، در الگوریتم یادگیری پس‌انتشار، با اصلاح روش به‌روزرسانی وزن‌ها، از طریق تغییر ساختار بهینه‌سازها، پایداری بیشتر شبکه و بالاتررفتن دقت بازشناسی حاصل شده است. روش پیشنهادی روی دو مجموعه داده‌ی IRVD و مجموعه MIO-TCD ارزیابی شده‌است. نتایج ارزیابی نشان می‌دهند که ساختار پیشنهادی در مقایسه با ساختارهای کانولوشن رایج، از نظر معیارهای طبقه‌بندی، عملکرد بهتری داشته است.

کلمات کلیدی: شبکه کانولوشن، یادگیری عمیق، تشخیص خودرو، موجک هار.

Real-time vehicle type recognition using a convolution and Haar wavelet pooling based classifier

Seyyed Mohammad Javadi Moghaddam, Hossein Gholamalinejad

Abstract: Over the past few years, real-time classification of vehicle types has become an increasingly popular and important topic, given its wide range of applications in traffic control and analysis. Among the various methods available for classifying car types, convolutional neural networks (CNNs) have emerged as particularly appealing. In this article, we introduce a new real-time CNN architecture that is specifically designed to detect different types of cars. This innovative structure incorporates several unique features, including novel network architecture and structural elements, as well as an advanced learning method based on the back-propagation algorithm. One key aspect of our proposed method is its use of feature extraction at three different locations within the network, which allows for more accurate and efficient classification of car types. To evaluate the performance of our approach, we conducted experiments on two popular datasets (IRVD and MIO-TCD), and compared our results against those obtained using traditional CNN structures. Our

[1] SE Block: Squeeze-and-Excitation blocks

[Downloaded from joc.kntu.ac.ir on 2026-11-17]

evaluation results demonstrate that our proposed CNN architecture outperforms existing approaches, achieving superior classification accuracy across a range of criteria. Overall, our work represents a significant advance in the field of real-time car type classification, with broad implications for traffic management and analysis.

Keywords: Convolution Neural network, Deep learning, IRVD dataset, Vehicle type Recognition.

۱- مقدمه

امروزه تجزیه و تحلیل بصری در مورد نظارت بر ترافیک، یکی از موضوعات جالب توجه در بینایی ماشین به شمار می‌رود. طبقه‌بندی و تشخیص وسایل نقلیه به دلیل مواردی مثل تغییرات در وضعیت زاویه‌ای خودرو و دوربین، کیفیت و وضوح تصویر، نور و شرایط آب و هوایی، در زمره‌ی مسائل پیچیده و دشوار در بینایی ماشین به شمار می‌روند [۱]. سیستم‌های کنترل‌تردد هوشمند با هدف به حداقل رساندن مشکلات ترافیکی ایجاد شده‌اند. یکی از بهترین راه کارهای کنترل تردد هوشمند مبتنی بر هوش مصنوعی، استفاده از شبکه‌های عصبی کانولوشن است [۲]. در سال‌های اخیر، شبکه‌های عصبی کانولوشن در شاخه‌های مختلف علمی از قبیل بینایی ماشین، بازشناسی گفتار و سیستم‌های مبتنی بر هوش مصنوعی، موفقیت‌های چشم‌گیری داشته‌اند. قدرت بالای روش‌های طبقه-بندی مبتنی بر کانولوشن، سهولت استفاده و عدم نیاز به استخراج دستی ویژگی، از مزایای این روش‌ها است. یکی از کاربردهای شبکه‌های کانولوشن، طبقه‌بندی تصاویر است که در سال‌های اخیر ساختارهای مختلفی برای نیل به این هدف معرفی شده‌است. کاربرد برخی از این ساختارها، استفاده در طبقه‌بندی تصاویر پیچیده‌ای مانند تصاویر خودرو است.

شبکه‌های کانولوشن با چالش‌های زیادی مواجه هستند [۳]. یکی از چالش های اساسی، زیاد بودن تعداد تکرار لازم داده‌ها برای یادگیری شبکه یا تثبیت نسبی وزن‌های شبکه است. جهت‌دهی به فرآیند استخراج ویژگی در این شبکه‌ها می‌تواند فرآیند یادگیری را تسریع نموده و تعداد دفعات تکرار را کاهش دهد.

راهکارهای ارائه‌شده در این مقاله، با هدف جهت‌دهی به فرآیند استخراج ویژگی، هم در سطح ساختار طبقه بند و هم در سطح الگوریتم یادگیری، می باشد.

۲- کارهای پیشین

روش‌های شناسایی نوع خودرو را می‌توان از جنبه‌های مختلف تقسیم‌بندی کرد. روش‌های شناسایی که به استخراج دستی ویژگی مانند ویژگی‌های هندسی [۴]، HOG [۵]، و الگوریتم SURF [۶] نیاز دارند. نقطه ضعف این روش‌ها از دست رفتن برخی از ویژگی‌های مهم تصویر در فرایند استخراج ویژگی است.

امروزه رایج‌ترین تکنیک، روش‌های مبتنی بر روش یادگیری عمیق است که از مزیت قابل توجهی در استخراج ویژگی‌های سطح پایین و سطح بالای تصویر برای بهبود طبقه‌بندی شی و تشخیص برخوردار است [۷]. شبکه‌های عصبی مبتنی بر کانولوشن، از رایج‌ترین روش‌ها برای دسته‌بندی خودروها هستند. سون و همکارانش [۸] روش‌های جدیدی مبتنی بر این شبکه‌ها ارائه کردند. آن‌ها از شبکه‌ی کانولوشن بدون نظارت و از شبکه-های نیمه‌نظارتی برای استخراج ویژگی‌های خودروها استفاده‌کردند و به صحت ۸۸/۳۵ درصدی در بازشناسی نوع خودرورسیدند.

شبکه‌ی عصبی کانولوشن مبتنی بر پیشنهاد-ناحیه[1] توسط عثمان و همکارانش برای طبقه‌بندی خودروهاروش دیگری بود که ارائه شد [۹]. در این روش، هدف، طبقه‌بندی تصاویر خودرو با زوایای مختلف (نمای جلو، نمای جانبی، نمای عقب) بود. صحت ۹۹/۲۴ درصد، در این روش گزارش شده‌است.

آونگ [۱۰] یک رویکرد طبقه‌بندی نوع خودرو با استفاده از یک ویژگی پیشرفته پیشنهاد کرد که درآن ازتکنیک استخراج مبتنی بر شبکه عصبی کانولوشن فیلتر شده پراکنده با استراتژی پرش از لایه استفاده کرد. دقت این روش ۹۳٪ بود. جوشوا [۱۱] از ترکیب الگوریتم ماشین بردار پشتیبان و OCR برای طبقه‌بندی خودرو استفاده کرد و توانست صحت ۹۵ درصدی را حاصل کند. بورمن و همکاران [۱۲] ازالگوریتم کوانتیزاسیون بردار یادگیری برای شناسایی نوع خودرو استفاده کردند. دقت بدست آمده توسط این روش ۸۳٪ بود. خراکی [۱۳] از یک شبکه کانولوشن عمیق برای طبقه‌بندی خودروهای اورژانس در سیستم‌های ترافیک استفاده کرد و نشان داد که شبکه Densenet121 برای این موضوع مناسب می باشد.

هو و همکاران [۱۴]یک مدل طبقه‌بندی نوع خودرو با استفاده از یک روش استخراج ویژگی پیشرفته معرفی کرد. این روش از Soft max به‌عنوان طبقه بند استفاده می‌کند. این روش، با اعمال بر روی یک مجموعه‌ی داده-ی کوچک ۱۲۰۰ تایی از تصاویر نمای جلوی شش کلاس خودرو، صحت ۹۷/۸۰ درصدی را نتیجه داده است. غلامعلی نژاد [۱۵] یک روش ترکیبی با استفاده از شبکه کانولوشن و لایه های ادغام موجکی ارائه کرد و به دقت ۹۷/۱۲ درصد رسید. کار جدید دیگر توسط ژو و همکارانش معرفی شده-است که برای تشخیص وسایل نقلیه، از تکنیک شبکه در شبکه (NIN) استفاده کرده است [۱۶]. در این روش، از ادغام میانگین سراسری برای جلوگیری از بیش برازش استفاده می‌شود. صحت این مدل برای طبقه‌بندی

[1] R-CNN: Region-proposal CNN

[Downloaded from joc.kntu.ac.ir on 2026-11-17]

وسایل نقلیه، بر روی مجموعه‌ی داده‌ی اتومبیل استنفورد، ۹۷/۲۰ درصد گزارش شده‌است.

ژو و همکارانش یک روش مبتنی بر یادگیری عمیق با استفاده از فهرست فشرده سازی برای شناسایی نوع خودرو پیشنهاد داده است [۱۷]. در این روش، از یک شبکه‌ی عصبی کانولوشن هرمی چند مقیاس برای تشخیص استفاده شده‌است. نتایج گزارش شده در این مقاله نشان می‌دهد میزان تشخیص آن بر روی داده‌های آزمون جمع‌آوری شده، بیش از ۹۵ درصد است.

این مقاله با تغییر در سطح ساختار طبقه بند و هم در سطح الگوریتم یادگیری، سعی کرده است که صحت و دیگر معیارهای طبقه بندی را بهبود بخشد.

۳- روش پیشنهادی

معماری و اجزای راهکار پیشنهادی برای بازشناسی بلادرنگ نوع خودرو مبتنی بر کانولوشن، در ادامه بیان می‌شود.

۳-۱ به‌روزرسانی وزن‌ها

فرآیند یادگیری با نظارت در شبکه‌های عصبی مبتنی بر کانولوشن، توسط الگوریتم پس‌انتشار اجرا می‌شود. برای به‌روزرسانی وزن‌ها در این الگوریتم، روش‌های مختلفی برای همگرایی بهتر و جلوگیری از بیش‌برازش پیشنهاد شده‌است. در شبکه‌های عصبی کانولوشن، وزن‌های لایه‌ی کانولوشن، فرآیند تغییر اطلاعات بین لایه‌ها را به‌عهده دارند. این وزن‌ها که به‌صورت ماتریس در حافظه ذخیره می‌شوند، در هر لایه عموما شامل اطلاعات مشابهی هستند. حذف هدفمند این اطلاعات مشابه، می‌تواند به ساختن نقشه‌ی ویژگی‌های بهتر کمک کند. الگوریتم بهینه‌ساز پیشنهادی موسوم به بهینه‌ساز WGD برای حذف هدفمند اطلاعات مشابه موجود در لایه‌های کانولوشن است. برای تحقق این هدف، ما از ماتریس میانگین گرادیان در هر لایه استفاده کردیم. در هر لایه‌ی همگشت، ماتریس میانگین گرادیان کرنل‌ها محاسبه شده و از همه‌ی گرادیان کرنل‌های وزنی موجود در این لایه، تفریق می‌شود. رابطه‌ی آن، به‌صورت زیر است:

$$h_{k,x_n} = \nabla_{x_n} J\left(x_n(k)\right) - \mu x_n$$
$$\mu x_n = \frac{1}{p}\sum_{i=1}^{p} \nabla_{x_n} J\left(x_n(i)\right) \qquad (۱)$$

در رابطه ۱، x_n ماتریس کرنل‌های وزنی بلوک اصلی nام، k شماره‌ی کرنل وزنی، μx_n میانگین گرادیان‌های کرنل‌های وزنی در هر کدام از بلوک‌های اصلی است. تابع J تابع هزینه است. P تعداد ماتریس‌های وزن در بلوک اصلی xام است.

۳-۲ لایه‌ی ادغام DWT[1]

یکی از اصلی‌ترین لایه‌ها در شبکه‌های عصبی مبتنی بر کانولوشن، لایه‌ی ادغام است. وظیفه‌ی اصلی این لایه، کاهش ابعاد نقشه‌ی ویژگی‌های استخراج شده از لایه‌های کانولوشن است. ساختار پیشنهادی ادغام، بر پایه‌ی تبدیل موجک گسسته‌ی هار تدوین شده‌است. با استفاده از این تبدیل، لایه‌ی ادغام فرآیند کاهش ابعاد و استخراج ویژگی‌های مفید را اجرا می‌کند [۱۸]. اگر x را یک تصویر در نظر بگیریم، می‌توانیم تبدیل موجک گسسته‌ی دو بعدی با چهار فیلتر کانولوشن را استفاده کنیم؛ فیلتر پایین گذر f_{LL}، و فیلترهای بالاگذر f_{LH}, f_{HL}, f_{HH}. وقتی این فیلترها به ورودی x اعمال شوند، چهار تصویر به نام‌های $x_{LL}, x_{LH}, x_{HL}, x_{HH}$ تولید خواهند شد. ابعاد این تصاویر، نسبت به تصویر مرحله‌ی قبل، یک دوم است. در صورتی که از موجک هار استفاده کنیم، تعریف فیلترها به‌صورت زیر است:

$$\begin{pmatrix} f_{LL} = \begin{pmatrix} +1 & +1 \\ +1 & +1 \end{pmatrix} & f_{LH} = \begin{pmatrix} -1 & -1 \\ +1 & +1 \end{pmatrix} \\ f_{HL} = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} & f_{HH} = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \end{pmatrix} \qquad (۲)$$

و خروجی‌های این تبدیل گسسته با روابط زیر محاسبه خواهند شد:

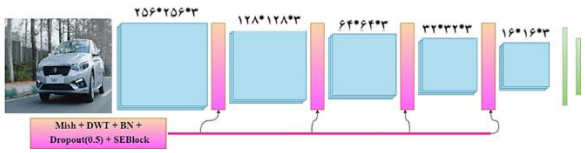
$$\begin{pmatrix} x_{LL} = (f_{LL} \otimes x) \downarrow_2 & x_{LH} = (f_{LH} \otimes x) \downarrow_2 \\ x_{HL} = (f_{HL} \otimes x) \downarrow_2 & x_{HH} = (f_{HH} \otimes x) \downarrow_2 \end{pmatrix} \qquad (۳)$$

در رابطه‌ی فوق، علامت $\otimes$ نشان دهنده‌ی عملگر کانولوشن و علامت $\downarrow_2$ به معنای نمونه‌گیری کاهشی با مرتبه‌ی ۲ است.

۳-۳ طبقه بند پیشنهادی

برای طراحی یک ساختار مناسب جهت طبقه‌بندی خودرو، دو مسئله در نظر گرفته شده است: سرعت شناسایی و صحت شناسایی. برای رسیدن به پاسخ سریع، ساختار مورد نظر باید تعداد پارامترهای آموزش پذیر کمی داشته باشد و برای تحقق هدف رسیدن به یک صحت بازشناسی بالا، نوع و مرتبه‌ی لایه‌ها باید به دقت انتخاب شوند.

در ساختار پیشنهادی، لایه‌های شبکه‌ی کانولوشن سنتی با بلوک‌های فشار-تحریک و لایه‌های ادغام موجکی و لایه‌های حذف تصادفی ترکیب شده است. شکل ۱ ساختار طبقه‌بند پیشنهادی را نشان می‌دهد.



شکل ۱: ساختار طبقه بند پیشنهادی

[1] Discrete Wavelet Transform

[Downloaded from joc.kntu.ac.ir on 2026-11-17]

در این ساختار، بعد از هر لایه‌ی کانولوشن از تابع فعال ساز Mish به همراه DWT برای استخراج ویژگی استفاده شده‌است. سپس از نرمال‌ساز دسته‌ای[1] که به اختصار BN نامیده می‌شود، بعد از هر کدام از لایه‌های DWT استفاده می‌شود. این کار با هدف کاهش زمان همگرایی وزن‌ها در مرحله‌ی یادگیری و رسیدن به راندمان بهتر انجام شده‌است. بلوک فشار-تحریک[2] بعد از هر لایه‌ی Dropout، با هدف استخراج ویژگی‌های بهتر از نقشه‌ی ویژگی خارج شده از لایه‌های پیشین، تعبیه شده‌است. ترکیب لایه‌ی کانولوشن بعلاوه‌ی ۶ لایه‌ی معرفی شده، ۵ بار در این ساختار تکرار شده‌است. یک لایه‌ی Dropout با احتمال انتخاب تصادفی ۰/۱ از داده‌ها برای تنظیم شبکه و جلوگیری از بیش برازش استفاده شده‌است. در انتها دو لایه‌ی تمام متصل، ماتریس نقشه‌ی ویژگی را به بردار تبدیل کرده و برای ورود به لایه‌ی طبقه‌بند اماده می‌کنند. از تابع فعالیت Softmax برای طبقه‌بندی استفاده شده‌است.

در ساختارپیشنهادی، استفاده ازتابع فعال‌سازی Mish، در استخراج ویژگی‌های دقیق‌تر به شبکه کمک می‌کند. در [۱۹, ۲۰] به صورت مبسوط در مورد این تابع فعالساز و تاثیر آن در عملکرد شبکه‌های عمیق بحث شده است. در این دو پژوهش بیان شده است که استفاده از این تابع فعالساز موجب افزایش مقدار صحت و دقت شبکه ها شده ولی زمان پردازش را هم افزایش داده است. علاوه بر این، در این ساختار، قبل از هر بلوک فشار-تحریک از یک حذف تصادفی با احتمال انتخاب پنجاه درصد و در لایه‌ی انتهایی یک حذف تصادفی با احتمال انتخاب ده‌درصد استفاده شده است.

از نظر تنظیم پارامترهای جانبی مورد استفاده در مرحله‌ی یادگیری، اصلی-ترین نوآوری صورت گرفته، پیشنهاد یک روش جدید برای به‌روزرسانی وزن‌ها در اجرای الگوریتم پس‌انتشار است. روش WGD روشی ساده و قابل استفاده در تمام بهینه‌سازهای مبتنی بر گرادیان نزولی است. در ساختار پیشنهادی از بهینه‌ساز بهبود یافته‌ی WGD-SGDM [۲۱] استفاده می‌کند.

٤- آزمایش‌ها و شبیه‌سازی‌ها

این بخش توان عملکردی ساختار پیشنهادی را مورد ارزیابی قرار می دهد. کلیه آزمایش‌های صورت گرفته در این بخش در یک سیستم با پردازنده Core i3-9100f با ۱۶گیگابایت حافظه، فرکانس کار ۳۶۰۰ مگاهرتز و پردازنده‌ی گرافیکی RTX Turbo 2080 در مد کاری بازی پیاده شده-است. زمان‌های گزارش شده، زمان واقعی لازم برای بازشناسی است و زمان بارگذاری داده‌ها از آن کم شده است. ابتدا، روش به‌روزرسانی وزن پیشنهادی WGD، برای بهبود عملکرد پس انتشارخطا مورد بررسی قرار می‌گیرد و سپس ساختار طبقه‌بند پیشنهادی، بررسی می‌شود.

٤-١ بررسی روش پیشنهادی WGD

در شبکه‌های مبتنی بر کانولوشن، در هر کدام از لایه‌های کانولوشن، تعدادی کرنل وزنی وجود دارد. در روش پیشنهادی WGD، ابتدا گرادیان برای هر کدام از این کرنل‌ها محاسبه می‌شود، سپس میانگین این گرادیان‌ها حساب می‌شود. در نهایت، از گرادیان‌های محاسبه شده، مقدار میانگین کسر می‌شود. خروجی این مرحله، برای استفاده در بهینه‌ساز مورد استفاده قرار می‌گیرد.

همان‌طور که قبلا بیان شد، این روش، به‌سادگی در هر بهینه‌سازی قابل پیاده‌سازی است. با توجه به‌این‌که رایج‌ترین بهینه‌سازهای مبتنی بر گرادیان نزولی، SGDM، Adam، Nadam و Radam هستند، الگوریتم پیشنهادی در این بهینه‌سازها بررسی شده‌است.

برای بررسی عملکرد این روش، از سه طبقه‌بند DenseNet121، ResNet18 و ResNet50 و آزمایش‌ها بر روی دو مجموعه‌ی داده‌ی CIFAR-100 و Animals-10 اجرا شده است. بالا بودن تعداد دسته‌ها علت استفاده از این دو مجموعه‌ی داده است. تاثیر روش WGD بر روی شناسایی داده‌های خودرویی در بخش بررسی ساختار پیشنهادی، گزارش شده است. برای تمام آزمایش‌های زیر، شرط توقف، اتمام تعداد مراحل تکرار یادگیری است که برابر با مقدار ۵۰ تنظیم شده¬است.

ابتدا به تحلیل آزمایش‌ها روی Animals-10 می‌پردازیم که نتایج آن¬ها در ادامه نشان داده شده‌است. مجموعه‌ی داده‌ی Animals-10 یک مجموعه‌ی پیچیده با تصاویر کیفیت بالا از حیوانات در ۱۰ دسته‌ی مختلف است [۲۲]. در این مجموعه، چالش‌هایی مثل پس زمینه‌های مختلف، زوایای دید مختلف، و دسته‌های بالانس نشده وجود دارد.

همان‌طور که در جدول‌های ۲ ، ۳ و ۴ مشاهده می‌شود، افزودن WGD به همه‌ی بهینه‌سازها موجب بالارفتن صحت بازشناسی و کاهش خطای شبکه شده است. بهترین نتیجه در بهینه‌ساز Radam بهبودیافته با روش WGD در شبکه‌ی DenseNet121 با مقدار ۹۰ درصد است. بیشترین میزان بهبود صحت بازشناسی در بهینه‌ساز Nadam بهبود یافته با روشWGD است که با طبقه‌بند عمیق ResNet50 حاصل شده و درصد بازشناسی را از۶۷ درصد به ۷۶ درصد رسانده¬است.

پارامتر ضریب کاپا کوهن پارامتری آماری است که برای اندازه‌گیری قابلیت اطمینان برای اقلام طبقه‌بندی کیفی استفاده می‌شود [۲۳]. لذا می-توان از آن به عنوان ابزاری برای نشان دادن قدرت ساختار طبقه‌بند بهره گرفت. با توجه به جدول ۱ بهترین بهبود در این پارامترها را در بهینه‌ساز Nadam بوده است که به مقدار ۶ درصد در شبکه‌ی DenseNwt121 و شبکه‌ی ResNet18 و حدود ۱۰ درصد در شبکه‌ی ResNet50 است. بالاترین درصد ضریب کاپا هم در شبکه‌ی DenseNet121 با بهینه‌ساز Radam بهبود یافته با روش WGD به‌دست آمده است.

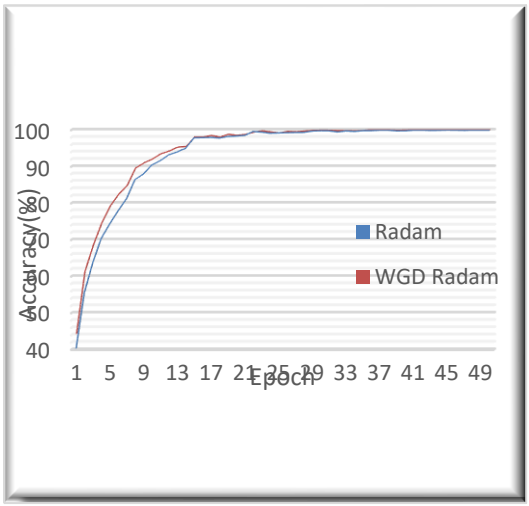
[1] Batch Normalization: BN
[2] Squeeze-and-excitation Blocks

[Downloaded from joc.kntu.ac.ir on 2026-11-17]

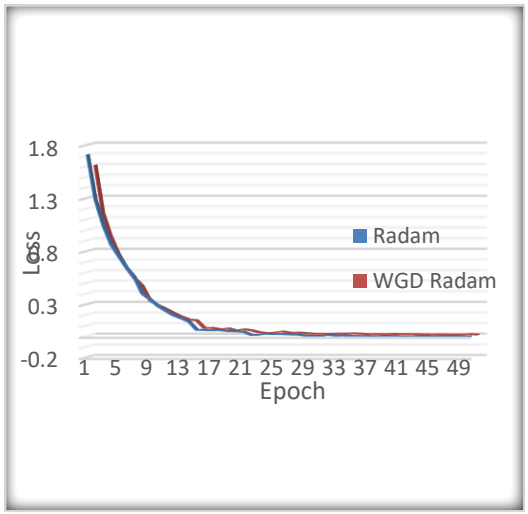
جدول ۱: نتایج طبقه بندی برای مجموعه‌ی Animals-10

بهینه ساز	شبکه	صحت	اتلاف[3]	فرخوانی مجدد[2]	دقت[1]	ضریب کاپا
SGDM	DenseNet121	۸۷/۰۰	۰/۵۰۷	۸۵/۸۴	۸۶/۵۱	۸۵/۲۰
	ResNet18	۸۱/۱۹	۰/۷۲۷	۷۹/۱۹	۸۰/۶۴	۷۸/۵۴
	ResNet50	۷۷/۰۳	۱/۹۷۴	۷۴/۵۷	۷۶/۳۲	۷۳/۷۹
WGD SGDM	DensNet121	۸۷/۶۵	۰/۴۵۳	۸۷/۱۷	۸۷/۳۷	۸۵/۹۵
	ResNet18	۸۱/۵۷	۰/۶۵۴	۷۹/۹۳	۸۱/۳۴	۷۹/۰۰
	ResNet50	۸۱/۳۸	۰/۸۰۳	۷۹/۴۴	۸۰/۶۶	۷۸/۷۶
Adam	DensNet121	۷۷/۴۱	۰/۶۸۰	۷۶/۳۴	۷۵/۹۹	۷۴/۳۲
	ResNet18	۷۲/۹۰	۰/۹۰۱	۷۰/۳۱	۷۱/۵۸	۶۹/۰۹
	ResNet50	۷۳/۹۷	۰/۷۹۴	۶۹/۶۳	۷۳/۸۴	۷۰/۱۹
WGD Adam	DensNet121	۷۹/۵۵	۰/۶۶۰	۷۷/۶۱	۷۸/۱۹	۷۶/۶۹
	ResNet18	۷۵/۳۱	۰/۸۵۳	۷۴/۲۰	۷۳/۱۵	۷۱/۹۹
	ResNet50	۷۶/۶۴	۰/۷۶۹	۷۴/۵۶	۷۶/۲۵	۷۳/۳۹
Nadam	DensNet121	۷۳/۶۶	۰/۷۹۷	۷۱/۱۷	۷۲/۷۹	۶۹/۹۶
	ResNet18	۶۹/۱۹	۰/۹۶۵	۶۷/۰۸	۶۶/۸۷	۶۵/۰۱
	ResNet50	۶۷/۲۰	۰/۹۴۹	۶۴/۷۳	۶۳/۸۰	۶۲/۷۵
WGD Nadam	DensNet121	۷۸/۳۶	۰/۶۲۱	۷۶/۵۷	۷۷/۰۸	۷۵/۳۶
	ResNet18	۷۵/۲۷	۰/۸۱۸	۵۲/۵۵	۷۴/۵۱	۷۱/۷۸
	ResNet50	۷۶/۰۷	۰/۷۴۲	۷۴/۱۴	۷۴/۵۰	۷۲/۷۶
Radam	DensNet121	۸۷/۶۹	۰/۷۶۰	۸۷/۴۰	۸۶/۷۷	۸۶/۰۱
	ResNet18	۸۲/۹۱	۱/۲۵۹	۸۱/۱۲	۸۲/۳۴	۸۰/۵۱
	ResNet50	۸۳/۴۱	۱/۰۹۷	۸۲/۱۹	۸۲/۳۷	۸۱/۱۲
WGD Radam	DensNet121	۹۰/۰۲	۰/۶۲۲	۹۰/۰۹	۸۹/۵۶	۸۸/۶۶
	ResNet18	۸۴/۷۱	۱/۱۱۰	۸۳/۶۹	۸۳/۹۷	۸۲/۶۰
	ResNet50	۵۸/۲۸	۰/۸۹۳	۸۴/۶۳	۸۴/۳۶	۸۴/۳۶

منحنی فرآیند یادگیری با استفاده از بهینه‌ساز Radam معمولی و بهبود-یافته با WGD در شکل ۲، نمایش داده‌شده است. همان‌طور که دیده می-شود، استفاده از WGD باعث افزایش درصد صحت بازشناسی و کاهش مقدار ضرر شده‌است.



(الف)



(ب)

شکل ۲: عملکرد شبکه‌ی DenseNet121 در فاز یادگیری روی مجموعه‌ی Animals-10 (الف) منحنی صحت (ب) منحنی خطا

این آزمایش‌ها روی مجموعه‌ی داده‌ی CIFAR-100 نیز انجام شده است. این مجموعه‌ی داده، حاوی تصاویری با کیفیت پایین و ابعاد کوچک است که در ۱۰۰ دسته، تقسیم بندی شده‌اند [۲۴]. نتایج آزمایش‌های این مرحله در جدول ۲ نشان داده‌شده‌است. همان‌طور که دیده می‌شود، علی‌رغم اینکه این مجموعه داده تصاویر کیفیت پایینی دارد، صحت ۷۰/۳۵ درصدی روی داده‌های آزمایش با بهینه‌ساز WGD Radam به‌دست آمده‌است.

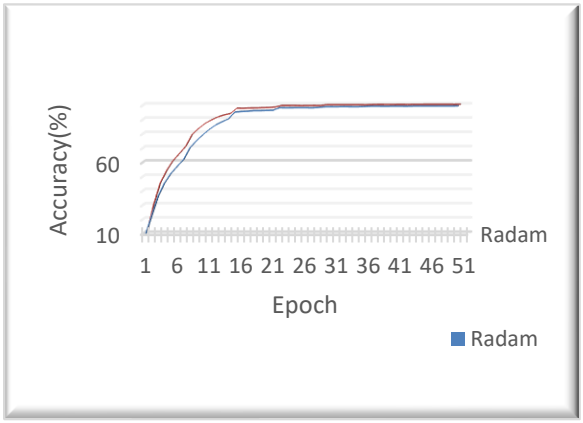
[1] Precision
[2] Recall
[3] Loss

[Downloaded from joc.kntu.ac.ir on 2026-11-17]

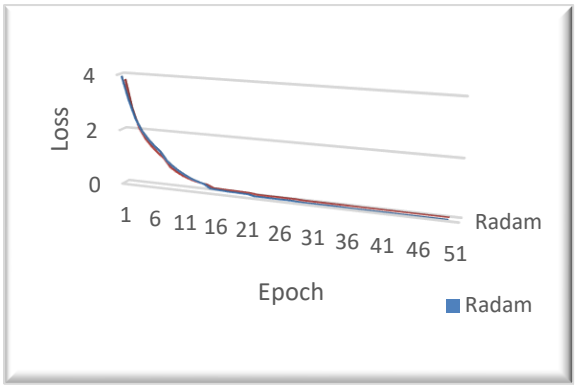
جدول ۲: نتایج طبقه بندی برای مجموعه‌ی CIFAR100

بهینه ساز	شبکه	صحت	اتلاف[3]	فرخوانی مجدد[2]	دقت[1]	ضریب کاپا
SGDM	DenseNet121	۸۷/۰۰	۰/۵۰۷	۸۵/۸۴	۸۶/۵۱	۸۵/۲۰
	ResNet18	۸۱/۱۹	۰/۷۲۷	۷۹/۱۹	۸۰/۶۴	۷۸/۵۴
	Resnet50	۷۷/۰۳	۱/۹۷۴	۷۴/۵۷	۷۶/۳۲	۷۳/۷۹
WGD SGDM	DenseNet121	۸۷/۶۵	۰/۴۵۳	۸۷/۱۷	۸۷/۳۷	۸۵/۹۵
	Resnet18	۸۱/۵۷	۰/۶۵۴	۷۹/۹۳	۸۱/۳۴	۷۹/۰۰
	Resnet50	۸۱/۳۸	۰/۸۰۳	۷۹/۴۴	۸۰/۶۶	۷۸/۷۶
Adam	DenseNet121	۷۷/۴۱	۰/۶۸۱	۷۶/۳۴	۷۵/۹۹	۷۴/۳۲
	Resnet18	۷۲/۹۰	۰/۹۰۱	۷۰/۳۱	۷۱/۵۸	۶۹/۰۹
	Resnet50	۷۳/۹۷	۰/۷۹۴	۶۹/۶۳	۷۳/۸۴	۷۰/۱۹
WGD Adam	DenseNet121	۷۹/۵۵	۰/۶۶۰	۷۷/۶۱	۷۸/۱۹	۷۶/۶۹
	Resnet18	۷۵/۳۱	۰/۸۵۳	۷۴/۲۰	۷۳/۱۵	۷۱/۹۹
	Resnet50	۷۶/۶۴	۰/۷۶۹	۷۴/۵۶	۷۶/۲۵	۷۳/۳۹
Nadam	DenseNet121	۷۳/۶۶	۰/۷۹۶	۷۱/۱۷	۷۲/۷۹	۶۹/۹۶
	Resnet18	۶۹/۱۹	۰/۹۶۵	۶۷/۰۸	۶۶/۸۷	۶۵/۰۱
	Resnet50	۶۷/۲۰	۰/۹۴۹	۶۴/۷۳	۶۳/۸۰	۶۲/۷۵
WGD Nadam	DenseNet121	۷۸/۳۶	۰/۶۲۱	۷۶/۵۷	۷۷/۰۸	۷۵/۳۶
	Resnet18	۷۵/۲۷	۰/۸۱۸	۵۲/۵۵	۷۴/۵۱	۷۱/۷۸
	Resnet50	۷۶/۰۷	۰/۷۴۲	۷۴/۱۴	۷۴/۵۰	۷۲/۷۶
Radam	DenseNet121	۸۷/۶۹	۰/۷۶۰	۸۷/۴۰	۸۶/۷۷	۸۶/۰۱
	Resnet18	۸۲/۹۱	۱/۲۵۹	۸۱/۱۲	۸۲/۳۴	۸۰/۵۱
	Resnet50	۸۳/۴۱	۱/۰۹۷	۸۲/۱۹	۸۲/۳۷	۸۱/۱۲
WGD Radam	DenseNet121	۹۰/۰۲	۰/۶۲۲	۹۰/۰۹	۸۹/۵۶	۸۸/۶۶
	Resnet18	۸۴/۷۱	۱/۱۱۰	۸۳/۶۹	۸۳/۹۷	۸۲/۶۰
	Resnet50	۵۸/۲۸	۰/۸۹۳	۸۴/۶۳	۸۴/۳۶	۸۴/۳۶

منحنی فرآیند یادگیری با استفاده از بهینه‌ساز Radam معمولی و نسخه‌ی بهبود یافته ی آن با WGD در شکل ۳ نشان داده شده است. در این منحنی‌ها هم، توانایی WGD در افزایش صحت بازشناسی و کاهش ضرر کاملا مشهود است.



(الف)



(ب)

شکل ۳: منحنی عملکرد شبکه‌ی DenseNet121 در فاز یادگیری روی مجموعه‌ی CIFAR-100 (الف) منحنی صحت و (ب) منحنی خطا

۴-۲ بررسی طبقه‌بند بهبود یافته‌ی پیشنهادی

برای بررسی ساختار پیشنهادی، عملکرد آن با تعدادی از شبکه‌های عصبی عمیق معروف شامل، VGG، ResNet، Inception v3، و DarkNet بر روی مجموعه داده‌های خودرویی IRVD و MIO-TCD، مجموعه داده ی کوید۱۹ و Animals-10 بررسی شده است.

آزمایش‌های IRVD

اولین مجموعه‌ی داده، IRVD است که توسط غلامعلی‌نژاد و همکاران گردآوری شده‌است [۲۵]. تصاویر این مجموعه‌ی داده در پنج گروه، اتوبوس، باری سنگین، باری متوسط، سدان وانت دسته‌بندی شده‌اند و دارای چالش‌های زیادی هستند. این مجموعه‌ی داده در شرایط نوری و آب و هوایی مختلفی مثل هوای بارانی، آفتابی، ابری، شب مهتابی و شب

[1] Precision
[2] Recall
[3] Loss

[Downloaded from joc.kntu.ac.ir on 2026-11-17]

تاریک گردآوری شده‌است. علاوه بر این موارد، شرایط مختلف جاده شامل جاده‌ی خیس و خشک، آسفالت سالم و معیوب، جاده‌های یک طرفه‌ی دو و سه خطی و جاده‌های دو طرفه نیز مورد توجه بوده‌است. تصاویر شب در دوحالت گرفته‌شده است؛ یکی با استفاده از نور پروژکتورهای LED و دیگری نور شب‌های مهتابی یا نور تابیده‌شده از خودروهای مجاور روی این خودرو. دوربین‌های مورد استفاده برای تصویربرداری، در ارتفاعات مختلف، از ۵۰ سانتی‌متر تا ۷ متر، نصب شده‌اند. بعد از فیلم‌برداری، قاب‌های شامل تصویر خودرو استخراج گردیده و مستطیل احاطه کننده‌ی خودرو، بریده شده‌است.

در جدول ۳ گزارش شده‌است. IRVDنتایج طبقه بندهای مختلف بر روی با توجه به جدول ۳، از نظر همه‌ی معیارهای مقایسه‌ای، طبقه‌بند پیشنهادی، رفتار بهتری داشته است. از نظر صحت بازشناسی، بهترین نتیجه با مقدار ۹۹/۵۹ درصد با شبکه‌ی پیشنهادی به‌دست آمده است. این شبکه بیش از یک پردازش DWTپنج میلیون پارامتر دارد. لازم به‌ذکر است از آن‌جاکه سفارشی است که در کد استاندارد ساختار پیشنهادی وارد می شود، برای و به زبان کودا تصویب نشده و در حال حاضر زمان اجرای آن روی GPU پردازنده‌ی گرافیکی بهینه نشده‌است و احتمال دارد با بهینه‌تر شدن کد زمان بازشناسی، از این مقدار نیز کم‌تر شود.DWTنوشتاری

جدول ۳: مقایسه‌ی روش پیشنهادی با برخی شبکه‌های همگشت معروف روی IRVD

مدل	تعداد پارامتر (میلیون)	صحت	فراخوانی مجدد	دقت	ضریب کاپا	زمان بازشناسی (ms) روی GPU
VGG11+ BN	۱۲	۹۹/۲۴	۹۸/۰۴	۹۹/۰۱	۹۸/۹۳	۰/۱۶۰
ResNet50	۲۳	۹۹/۱۱	۹۷/۹۷	۹۸/۵۵	۹۹/۰۱	۰/۶۳۹
ResNet152	۵۸	۹۸/۹۲	۹۸/۲۵	۹۸/۴۳	۹۸/۲۱	۱/۸۰۰
DarkNet53	۴۰	۹۹/۳۸	۹۹/۰۹	۹۸/۹۵	۹۹/۱۱	۰/۶۴۷
Inception-v3	۲۴	۹۹/۲۸	۹۹/۰۵	۹۸/۸۸	۹۸/۹۱	۱/۱۶۲
ساختار پیشنهادی	۵	۹۹/۵۹	۹۹/۱۴	۹۹/۳۷	۹۹/۲۵	۰/۰۴۷

روش‌پیشنهادی پنج میلیون پارامتر دارد که طبق جدول ۳، کم‌ترین مقدار است. با توجه به نتایج، می توان گفت الگوریتم پیشنهادی پیشنهادی با توجه به کم بودن تعداد پارامترهای آموزش پذیر، زمان بازشناسی بسیارکم و معیارهای بازشناسی، برای کاربردهای بلا درنگ مناسب است.

آزمایش‌های MIO-TCD

مجموعه‌ی داده‌ی دیگری که آزمایش شبکه روی آن اجرا شده، MIO-TCD است. این مجموعه، شامل ۷۸۶۷۰۲ تصویر در۱۱ دسته‌ی مختلف با چالش است که ۶۴۸۹۵۹ تصویر برای اهداف طبقه‌بندی و ۱۳۷۷۴۳ تصویر برای اهداف شناسایی محل اشیا در تصویر، آماده شده‌است. تصاویر در ساعات مختلف روز و ماه‌های مختلف سال از صدها دوربین رصد ترافیک مختلف در کانادا و ایالات متحده آمریکا گردآوری شده‌است. متاسفانه این مجموعه‌ی داده، کاملا تنظیم نشده است، به‌صورتی که حدود ۹۰ درصد داده‌ها در سه دسته‌ی car، background و pickup هستند.

نتایج اعمال این مجموعه‌ی داده به طبقه‌بندها در جدول۴ گزارش شده‌است. همان‌طور که در جدول ۴ مشاهده می‌شود، علی‌رغم این‌که ساختار پیشنهادی، کم‌ترین تعداد پارامتر را در بین این ساختارها دارد، برروی این مجموعه‌ی داده هم بهترین نتایج را به‌دست آورده‌است. با مقایسه‌ی نتایج جدول‌های ۳ و ۴، مشاهده می‌شود که نتایج طبقه‌بندها روی مجموعه‌ی MIO-TCD کمتر از نتایج حاصل از IRVD است. علت این کاهش مقدار در معیارها، بیشتر بودن تعداد کلاس‌های مجموعه‌ی MIO-TCD در مقایسه با IRVD است. کیفیت پایین تصاویر MIO-TCD نیز در این کاهش مقدار تاثیر گذار بوده‌است.

جدول ۴: مقایسه‌ی ساختار پیشنهادی با شبکه‌های دیگر روی MIO-TCD

مدل	تعداد پارامتر (میلیون)	صحت	فراخوانی مجدد	دقت	ضریب کاپا	زمان بازشناسی (ms) روی GPU
VGG11+ BN	۱۲۹	۹۴/۶۴	۸۱/۴۵	۸۴/۳۰	۹۲/۲۳	۰/۱۶۰
ResNet50	۲۳	۹۸/۶۸	۸۷/۰۶	۸۵/۸۲	۹۴/۸۴	۰/۶۳۹
Inception-v3	۲۴	۹۶/۱۸	۸۲/۷۸	۸۴/۷۸	۹۴/۰۴	۱/۱۶۲
ساختار پیشنهادی	۵	۹۶/۵۹	۸۴/۰۰	۹۰/۱۳	۹۴/۶۶	۰/۰۴۷

آزمایش‌های Animals-10

مجموعه‌ی داده‌ی حیوانات Animals-10، مجموعه‌ی دیگری است که برای بررسی ساختار پیشنهادی، از آن استفاده شده‌است. این مجموعه شامل ۲۶۱۸۳ تصویر از ده حیوان مختلف است [۲۲].

نتایج آزمایش‌های صورت گرفته روی این مجموعه‌ی داده، در جدول ۵، بیان شده‌است. همان‌طور که مشاهده می‌شود، ساختار پیشنهادی روی این مجموعه که شلوغ‌تر و پیچیده‌تر از مجموعه‌های دیگر بوده، باز هم نتایج خوب و قابل ملاحظه‌ای داشته است.

[Downloaded from joc.kntu.ac.ir on 2026-11-17]

جدول ۵: مقایسه‌ی روش پیشنهادی با برخی شبکه‌های همگشت روی Animals-10

مدل	تعداد پارامتر (میلیون)	صحت	فراخوانی مجدد	دقت	ضریب کاپا	زمان بازشناسی (ms) روی GPU
VGG11 + BN	۱۲۹	۸۵/۳۰	۸۵/۴۲	۸۴/۷۰	۸۴/۵۳	۰/۰۶۸
DenseNet121	۸	۷۲/۷۸	۷۲/۴۱	۸۲/۷۲	۶۸/۷۶	۰/۵۸۵
ResNet18	۱۱	۸۴/۰۶	۸۲/۶۱	۸۳/۴۹	۸۱/۸۳	۰/۰۶۳
ResNet50	۲۳	۸۳/۷۹	۸۱/۹۳	۸۲/۷۷	۸۱/۵۳	۰/۱۷۵
Inception-v3	۲۴	۸۵/۲۱	۸۴/۷۰	۸۴/۵۱	۸۳/۸۷	۰/۳۲۵
ساختار پیشنهادی	۵	۸۶/۸۵	۸۶/۱۶	۸۶/۴۸	۸۵/۰۳	۰/۰۹۲

آزمایش‌های COVUD-19 CT

یکی از مجموعه‌های داده‌ی تصویری، مجموعه‌ی COVID-19 CT است که در [۲۶, ۲۷] معرفی شده است. این مجموعه، شامل سه کلاس مختلف با ۱۹۶۸۵ تصویر است. نتایج اعمال طبقه‌بند پیشنهادی و سایر طبقه‌بندهای رایج در جدول۶ بیان شده‌است. طبق داده‌های این جدول، در بین ساختار-های عمیق رایج، بهترین نتایج توسط شبکه ی Inception-v3 ارائه‌شده-است که از نتایج حاصل از ساختار پیشنهادی ما کم‌تر است.

جدول ۶: مقایسه‌ی روش پیشنهادی با برخی شبکه‌های همگشت روی COVID19-CT

مدل	تعداد پارامتر (میلیون)	صحت	فراخوانی مجدد	دقت	ضریب کاپا	زمان بازشناسی (ms) روی GPU
VGG11 + BN	۱۲۹	۹۲/۷۳	۸۹/۴۲	۹۵/۳۵	۸۷/۸۷	۰/۰۵۳
DenseNet121	۸	۷۹/۱۱	۶۶/۰۹	۹۰/۱۷	۶۲/۳۷	۰/۵۸۸
ResNet18	۱۱	۹۳/۳۹	۸۹/۴۹	۹۵/۷۹	۸۸/۹۸	۰/۰۴۲
ResNet50	۲۳	۹۵/۹۸	۹۳/۹۸	۹۷/۰۱	۹۳/۴۰	۰/۲۰۵
Inception-v3	۲۴	۹۸/۱۱	۹۷/۶۱	۹۸/۴۹	۹۶/۹۴	۰/۳۸۶
GoogLeNet	۵	۹۷/۲۵	۹۶/۷۲	۹۷/۹۵	۹۶/۱۱	۰/۰۷۲
ساختار پیشنهادی	۵	۹۹/۰۳	۹۸/۹۱	۹۸/۷۱	۹۸/۴۳	۰/۰۶۹

۵- نتیجه گیری

در این مقاله، روشی نوین برای بازشناسی نوع خودروها معرفی شد. در این روش، علاوه بر پیشنهاد یک ساختار شبکه‌ی عصبی جدیدی، روشی نوینی برای به‌روزرسانی وزن‌ها در فرآیند یادگیری پس‌انتشار ارائه شده است. در ساختار عصبی پیشنهادی، لایه‌ی ادغام جدید علاوه بر کاهش ابعاد نقشه‌ی ویژگی خروجی از لایه‌ی کانولوشن، قابلیت استخراج ویژگی را نیز دارد. این لایه با استفاده از تبدیل موجک هار پیاده سازی شده است. نتایج آزمایش‌ها نشان داد که این لایه‌ی ادغام، نسبت به سایر لایه‌ها، تاثیر بهتری روی رفتار شبکه از نظر طبقه‌بندی دارد. نوآوری دیگر ساختاری، استفاده از بلوک‌های فشار-تحریک قبل از لایه‌های کانولوشن میانی می باشد. استفاده از این بلوک‌ها موجب بالارفتن صحت بازشناسی شد. چیدمان و معماری خاص این لایه‌ها، که به صورت ساختارپیشنهادی مورد بررسی و مطالعه قرار گرفته است، عملکرد خوب این ساختار را در کاربردهای طبقه‌بندی نشان می دهد.

نوآوری دیگر در الگوریتم یادگیری پس انتشار می باشد. با اصلاح روش به‌روزرسانی وزن‌ها، از طریق تغییر ساختار بهینه‌سازها، پایداری بیشتر شبکه و بالاتررفتن صحت بازشناسی حاصل شده است. این روش که به اختصار WGD نامیده شده، به آسانی قابل پیاده‌سازی در بهینه‌سازهای مختلف است. نتایج اعمال این نوآوری به شبکه‌های عمیق مختلف و روی مجموعه داده‌های تصویری مختلف، بالارفتن معیارهای طبقه‌بندی را نشان می دهد. نتایج نشان داد که روش پیشنهادی برای مجموعه داده‌ی IRVD زمان بازشناسی ۴۲ میلی‌ثانیه‌ای روی CPU و ۰/۰۴۸ میلی‌ثانیه‌ای روی GPU با صحت ۹۹/۵۹ درصدی را حاصل کرده است که هم از نظر صحت و هم از نظر زمان بازشناسی، در مقایسه با شبکه‌های کانولوشن رایج، بهترین حالت است. ساختار پیشنهادی بر روی مجموعه داده‌ی پرچالش MIO-TCD نیز صحت ۹۶/۵۹ درصدی را بدست آورده است.

با استفاده از این نوآوری در ساختار پیشنهادی، و انجام آزمایش¬های مختلف روی مجموعه‌ی داده‌های تصویری خودرویی IRVD و MIO-TCD ، مشخص می شود که ساختار پیشنهادی، نه تنها از نظر زمان بازشناسی، نسبت به سایر شبکه‌های عمیق رایج، عملکرد بهتری دارد، بلکه از نظر معیارهای چهارگانه‌ی طبقه‌بندی نیز نسبت به ساختارهای عمیق رایج وروی مجموعه‌ی داده‌های خودرویی، عملکرد بهتری داشته است.

٦- مراجع

[1] C. Zhang, H. Ho, W. H. Lam, W. Ma, S. Wong, and A. H. Chow, "Lane-based estimation of travel time distributions by vehicle type via vehicle re-identification using low-resolution video images," Journal of Intelligent Transportation Systems, pp. 1-20, 2022.

[2] N. Arora, Y. Kumar, R. Karkra, and M. Kumar, "Automatic vehicle detection system in different environment conditions using fast R-CNN,"

[Downloaded from joc.kntu.ac.ir on 2026-11-17]

efficient emergency vehicle classification in real-time traffic monitoring," IAES International Journal of Artificial Intelligence, vol. 11, no. 1, p. 110, 2022.

[14] X. Xu, J. Wu, J. Wang, Q. Qu, Z. Tan, and M. Luo, "Materials Identification of Polarized Pulse Laser Detection Based on Sparse Autoencoder and Softmax Classifier Framework," IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1-9, 2022.

[15] H. Gholamalinejad and H. Khosravi, "Vehicle classification using a real-time convolutional structure based on DWT pooling layer and SE blocks," Expert Systems with Applications, vol. 183, p. 115420, 2021.

[16] E. Zhu, M. Xu, and D. C. Pi, "Vehicle Type Recognition Algorithm Based on Improved Network in Network," Complexity, vol. 2021, 2021.

[17] Y. Zhou, "Vehicle image recognition using deep convolution neural network and compressed dictionary learning," Journal of Information Processing Systems, vol. 17, no. 2, pp. 411-425, 2021.

[18] H. Gholamalinejad and H. Khosravi, "Vehicle classification using a real-time convolutional structure based on DWT pooling layer and SE blocks," Expert systems with Applications, vol. 183, 2021.

[19] D. Misra, "Mish: A self regularized non-monotonic activation function," arXiv preprint arXiv:1908.08681, 2019.

[20] E. C. Seyrek and M. Uysal, "A comparative analysis of various activation functions and optimizers in a convolutional neural network for hyperspectral image classification," Multimedia Tools and Applications, pp. 1-32, 2023.

[21] H. Gholamalinejad and H. Khosravi, "Whitened gradient descent, a new updating method for optimizers in deep neural networks," Journal of AI and Data Mining, vol. 10, no. 4, pp. 467-477, 2022.

[22] W. Wang, P. Zhang, T. Lan, and V. Aggarwal, "Datacenter net profit optimization with deadline dependent pricing," in Information Sciences and Systems (CISS), 2012 46th Annual Conference on, 2012: IEEE, pp. 1-6.

[23] M. L. McHugh, "Interrater reliability: the kappa statistic," Biochemia medica: Biochemia medica, vol. 22, no. 3, pp. 276-282, 2012.

[24] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," 2009.

[25] H. Gholamalinejad and H. Khosravi, "IRVD: A Large-Scale Dataset for Classification of Iranian Vehicles in Urban Streets," Journal of AI and Data Mining, pp. -, 2020, doi: 10.22044/jadm.2020.8438.1982.

Multimedia Tools and Applications, vol. 81, no. 13, pp. 18715-18735, 2022.

[3] N. Arora and Y. Kumar, "Automatic vehicle detection system in Day and Night Mode: challenges, applications and panoramic review," Evolutionary Intelligence, pp. 1-19, 2022.

[4] M. Anandhalli, A. Tanuja, and P. Baligar, "Geometric invariant features for the detection and analysis of vehicle," Multimedia tools and applications, vol. 81, no. 23, pp. 33549-33567, 2022.

[5] R. S. El-Sayed and M. N. El-Sayed, "Classification of vehicles' types using histogram oriented gradients: comparative study and modification," IAES International Journal of Artificial Intelligence, vol. 9, no. 4, p. 700, 2020.

[6] A. Thomas, P. Harikrishnan, P. Palanisamy, and V. P. Gopi, "Moving vehicle candidate recognition and classification using inception-resnet-v2," in 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC), 2020: IEEE, pp. 467-472.

[7] H.-H. Jebamikyous and R. Kashef, "Autonomous vehicles perception (avp) using deep learning: Modeling, assessment, and challenges," IEEE Access, vol. 10, pp. 10523-10535, 2022.

[8] F. C. Soon, H. Y. Khaw, J. H. Chuah, and J. Kanesan, "Semisupervised PCA convolutional network for vehicle type classification," IEEE Transactions on Vehicular Technology, vol. 69, no. 8, pp. 8267-8277, 2020.

[9] Z. Huo, Y. Xia, and B. Zhang, "Vehicle type classification and attribute prediction using multi-task RCNN," in 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2016: IEEE, pp. 564-569.

[10] S. Awang, N. M. A. N. Azmi, and M. A. Rahman, "Vehicle type classification using an enhanced sparse-filtered convolutional neural network with layer-skipping strategy," IEEE Access, vol. 8, pp. 14265-14277, 2020.

[11] I. O. Joshua, M. O. Arowolo, M. O. Adebiyi, O. R. Oluwaseun, and K. A. Gbolagade, "Development of an Image Processing Techniques for Vehicle Classification Using OCR and SVM," in 2023 International Conference on Science, Engineering and Business for Sustainable Development Goals (SEB-SDG), 2023, vol. 1: IEEE, pp. 1-9.

[12] R. I. Borman, Y. Fernando, and Y. E. P. Yudoutomo, "Identification of Vehicle Types Using Learning Vector Quantization Algorithm with Morphological Features," Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), vol. 6, no. 2, pp. 339-345, 2022.

[13] A. Kherraki and R. El Ouazzani, "Deep convolutional neural networks architecture for an

[Downloaded from joc.kntu.ac.ir on 2026-11-17]

[26] "COVID-19 CT Scan Images." https://www.kaggle.com/azaemon/preprocessed-ct-scans-for-covid19?select=Original+CT+Scans (accessed.

[27] W. Ning et al., "iCTCF: an integrative resource of chest computed tomography images and clinical features of patients with COVID-19 pneumonia," 2020.

[Downloaded from joc.kntu.ac.ir on 2026-11-17]