

تشخیص خطوط جاده با استفاده از معماری سبک‌وزن مبتنی بر ماژول توجه CBAM

رفعت معیری فر^۱، معصومه رضائی^{۲*}، نیک محمد بلوچ زهی^۳

^۱ دانشجوی ارشد گروه مهندسی کامپیوتر، دانشگاه سیستان و بلوچستان، زاهدان، ایران

^۲ استادیار گروه مهندسی کامپیوتر، دانشگاه سیستان و بلوچستان، زاهدان، ایران

^۳ استادیار گروه مهندسی فناوری اطلاعات، دانشگاه سیستان و بلوچستان، زاهدان، ایران

چکیده

مقاله پژوهشی

تاریخ دریافت:

۱۴۰۴/۱۰/۰۲

تاریخ پذیرش:

۱۴۰۵/۰۱/۱۴

کلیدواژه‌ها:

تشخیص خطوط جاده، سیستم‌های پیشرفته کمک‌راننده، رانندگی خودران، ماژول توجه کانالی - فضایی، تقطیر دانش، ایمنی حمل‌ونقل

نویسنده مسئول:

mrezaei@ece.usb.ac.ir

تشخیص خطوط جاده به‌عنوان عناصر حیاتی در سیستم‌های ادراک بصری خودروهای هوشمند، نقشی کلیدی در تضمین ناوبری ایمن ایفا می‌کنند و برای سیستم‌های رانندگی خودران و فناوری‌های پیشرفته کمک‌راننده ضروری هستند. با این حال، دستیابی به دقت بالا در تشخیص خطوط جاده در شرایط چالش‌برانگیز مانند سایه، ازدحام ترافیکی و محوشدگی خطوط جاده، در عین حفظ کارایی بلادرنگ، همچنان یک چالش اساسی به شمار می‌رود. در این پژوهش، یک معماری سبک‌وزن به‌منظور ارتقای عملکرد تشخیص خطوط جاده در محیط‌های پیچیده ارائه شده است که با تلفیق راهبردهای نوآورانه طراحی گردیده است. هسته اصلی این چارچوب، ادغام هدفمند ماژول توجه کانالی-فضایی در لایه‌های میانی شبکه ResNet است که امکان تقویت و پالایش ویژگی‌های متمایزکننده را فراهم می‌سازد. افزون بر این، با بهره‌گیری از راهبرد پیشرفته تقطیر دانش، مدل پیشنهادی به دقت تشخیص بالا همراه با کارایی محاسباتی بلادرنگ دست‌یافته است. ارزیابی روی مجموعه داده CULane نشان داد که روش پیشنهادی با امتیاز F1 معادل ۸۰/۲۰٪ و سرعت ۴۰۷ فریم بر ثانیه، به ترتیب بهبود ۰/۵٪ و افزایش سرعت ۱/۷۹٪ نسبت به مدل CLRKNet و روش ECBAM_ASPP داشته است. همچنین در مجموعه داده TuSimple، مدل پیشنهادی با دقت ۹۶/۹۶٪ کمترین خطای منفی کاذب ۱/۵۷٪ در میان روش‌های مقایسه شده، عملکرد برتری از خود نشان داده است. در مقایسه با روش‌های مبتنی بر توجه مانند ECBAM_ASPP و معماری‌های پرسرعتی همچون UFLD، روش پیشنهادی تعادل مطلوبی میان دقت و سرعت برقرار کرده و برای کاربردهای بلادرنگ در خودروهای خودران مناسب‌تر است.



: 10.22034/ABMIR.2026.24147.1209

E-ISSN: 2821-2037

/The Author 2026. Published by Yazd University This is an open

access article under the CC BY 4.0 License (<https://creativecommons.org/licenses/by/4.0/>).



۱- مقدمه

ویژگی محور به ویژگی‌هایی مانند گرادیان لبه، رنگ و شدت روشنایی وابسته‌اند و به کنتراست قوی نیاز دارند. هر دو رویکرد از مراحل نظیر استخراج منطقه موردعلاقه^۵ (ROI)، پیش‌پردازش تصویر، استخراج ویژگی‌ها و تطبیق خطوط پیروی می‌کنند [۸].

با ظهور یادگیری عمیق، روش‌های سنتی جای خود را به الگوریتم‌های پیشرفته‌تر مبتنی بر شبکه‌های عصبی عمیق داده‌اند که نیاز به مهندسی دستی ویژگی‌ها را حذف کرده و دقت و استحکام را بهبود می‌بخشند. تکنیک‌های مدرن تشخیص خطوط به چهار دسته اصلی تقسیم می‌شوند: مبتنی بر تقسیم‌بندی، مبتنی بر منحنی پارامتری، مبتنی بر نقاط کلیدی و مبتنی بر انکر [۹]. با وجود پیشرفت‌های قابل‌توجه، چالش‌هایی مانند عملکرد در شرایط پیچیده (نور کم، سایه، خطوط محوشده یا جاده‌های شلوغ) و نیاز به اجرای بلادرنگ روی سخت‌افزارهای محدود همچنان باقی است. مدل‌های بزرگ، مانند شبکه‌های عصبی پیچشی^۶ (CNNs)، با وجود دقت بالا در وظایف بینایی ماشین، به دلیل پیچیدگی محاسباتی و نیازهای حافظه‌ای گسترده، برای کاربردهای بلادرنگ در سیستم‌های با منابع محدود چالش‌برانگیز بوده و توسعه معماری‌های کارآمدتر را ضروری می‌سازد [۱۰].

در پاسخ به این چالش‌ها، مکانیزم‌های توجه^۷، نقش مهمی در بهبود عملکرد ایفا می‌کنند. پژوهش حاضر باهدف رفع این محدودیت‌ها، یک روش نوین مبتنی بر یادگیری عمیق برای تشخیص خطوط جاده‌ای ارائه می‌دهد که با ترکیب مدل‌های سبک‌وزن و مکانیزم‌های توجه بهبودیافته است.^۸ CBAM با ترکیب توجه کانالی^۹ و فضایی^{۱۰}، امکان تمرکز بر ویژگی‌های کلیدی تصویر را فراهم کرده و نویزهای غیرضروری را حذف می‌کند که این امر به بهبود دقت در شرایط دشوار مانند نور کم یا خطوط محوشده کمک می‌کند [۱۱]، و تعادل مطلوبی بین دقت و سرعت استنتاج برقرار می‌کند. نتایج آزمایش‌ها بر روی مجموعه داده‌های [۱۲]

حوزه حمل‌ونقل جاده‌ای به طور مداوم در حال پیشرفت است و فناوری‌های نوین، به‌ویژه در راستای بهبود ایمنی، کاهش اثرات زیست‌محیطی و افزایش کارایی وسایل نقلیه، نقش مهمی ایفا می‌کنند [۱]. طبق گزارش سازمان بهداشت جهانی^۱ (WHO)، تصادفات جاده‌ای سالانه جان بیش از ۱/۱۹ میلیون نفر را می‌گیرند و به‌ویژه برای گروه سنی ۵ تا ۲۹ سال یکی از علل اصلی مرگ‌ومیر به شمار می‌روند [۲]. در این راستا، فناوری‌های رانندگی خودکار و سیستم‌های پیشرفته کمک‌راننده^۲ (ADAS) به دلیل توانایی در کاهش تصادفات، بهبود جریان ترافیک و کاهش مصرف انرژی و انتشار گازهای گلخانه‌ای، موردتوجه گسترده قرار گرفته‌اند [۳]. خودروهای خودران به بخش جدایی‌ناپذیری از سیستم‌های حمل‌ونقل هوشمند^۳ (ITS) تبدیل شده‌اند و انتظار می‌رود در آینده با بهره‌گیری از انرژی‌های الکتریکی، به طور کامل عملیات رانندگی را بدون دخالت انسان مدیریت کنند. این فرایند نیازمند یکپارچگی فناوری‌های پیشرفته در حوزه‌های بینایی ماشین، سیستم‌های کنترلی، ارتباطات و اتوماسیون است تا ایمنی، کارایی و پایداری حمل‌ونقل بهبود یابد [۴]. تشخیص خطوط جاده‌ای یکی از ارکان کلیدی در توسعه فناوری‌های رانندگی خودکار و سیستم‌های پیشرفته کمک‌راننده محسوب می‌شود. این سیستم‌ها با شناسایی دقیق مسیر حرکت، از انحراف خودرو جلوگیری کرده و تصمیم‌گیری‌های ایمن‌تر و دقیق‌تری را امکان‌پذیر می‌سازند [۵]. [۶]. تشخیص سریع و قابل‌اعتماد خطوط جاده‌ای برای کاربردهایی مانند حفظ خط، تغییر مسیر و کنترل پیمایش تطبیقی^۴ حیاتی است و به پیشرفت اتوماسیون رانندگی کمک شایانی می‌کند [۷]. با این حال، روش‌های سنتی تشخیص خطوط که اغلب به دو دسته مدل‌محور و ویژگی محور تقسیم می‌شوند، محدودیت‌هایی دارند. رویکردهای مدل‌محور با استفاده از مدل‌های پیش‌تعریف‌شده خطوط، حساسیت به نویز را کاهش می‌دهند، درحالی‌که روش‌های

⁶ Convolutional Neural Networks

⁷ Attention

⁸ Convolutional Block Attention Module

⁹ Channel Attention Module

¹⁰ Spatial Attention Module

¹ World Health Organization

² Advanced driver-assistance system

³ Intelligent Transportation System

⁴ Adaptive cruise control

⁵ Region of interest

۲- مروری بر روش‌های پیشین

به دلیل پیشرفت در فناوری‌های یادگیری عمیق، محققان استراتژی‌های متعددی را توسعه داده‌اند که کار تشخیص خط را به طور قابل توجهی ساده، تسریع و بهبود بخشیده است [۱۸]. هم‌زمان، با گسترش روزافزون یادگیری عمیق و ظهور مداوم با مفاهیم جدید [۱۹]، حوزه تشخیص خط شاهد تخصصی شدن بیشتر و اصلاح رویکردهای خود بوده است. نگاهی به مسیرهای تحقیقاتی اصلی در این حوزه، روش‌های تشخیص خطوط مبتنی بر دوربین را می‌توان به دو دسته کلی تقسیم کرد: تشخیص خطوط دوبعدی و تشخیص خطوط سه‌بعدی [۲۰]. در این بخش، تمرکز بر روش‌های دوبعدی است که شامل چهار رویکرد اصلی می‌شوند: تقسیم‌بندی، انکر، نقطه کلیدی، و منحنی.

۲-۱ روش‌های تشخیص خطوط دوبعدی

روش‌های تشخیص خط دوبعدی برای ترسیم دقیق شکل و موقعیت خطوط در تصاویر تلاش می‌کنند و عمدتاً از چهار رویکرد متمایز استفاده می‌کنند [۹]:

- رویکردهای مبتنی بر تقسیم‌بندی: این روش‌ها تشخیص خطوط دوبعدی را به عنوان یک چالش طبقه‌بندی سطح پیکسل ارائه می‌کنند و ماسک‌های خط را تولید می‌کنند [۲۱]. با کاهش ویژگی‌های معنایی مؤثر و پردازش پسین، گروه‌های خط را بهبود می‌دهند، اما با هزینه‌های محاسباتی بالا مشخص می‌شوند. SCNN [۲۱] و RESA^۲ [۲۲] از الگوی تقسیم‌بندی معنایی برای طبقه‌بندی نمونه‌های خط جاده به عنوان کلاس‌های جداگانه در هر پیکسل استفاده می‌کنند. تطابق بین خط جاده و کلاس از طریق حاشیه‌نویسی تعیین می‌شود که انعطاف‌پذیری کمتری دارد. با این حال، این روش‌ها خطوط جاده را به عنوان نمونه‌های یکپارچه در نظر نمی‌گیرند و نیاز به پردازش پسین پرهزینه دارند [۲۳].
- رویکردهای مبتنی بر انکر: این روش‌ها به دلیل سادگی و کارایی‌شان در تشخیص خطوط دوبعدی استفاده می‌شوند و

CULane و [۱۳] TuSimple نشان می‌دهد که روش پیشنهادی عملکرد رقابتی و بلادرنگی ارائه می‌دهد.

با وجود پیشرفت‌های اخیر، همچنان چالش‌هایی نظیر حفظ دقت در شرایط نوری دشوار و اجرای بلادرنگ روی سخت‌افزارهای محدود وجود دارد. در این پژوهش، با ادغام ماژول CBAM در معماری مبتنی بر تقطیر دانش، تلاش شده است این شکاف‌ها برطرف شود.

مشارکت‌ها و نوآوری‌های اصلی این پژوهش به شرح زیر است:

- ارائه یک معماری سبک‌وزن و بلادرنگ برای تشخیص خطوط جاده با استفاده از شبکه ResNet-34 [۱۴] به عنوان استخراج‌کننده ویژگی و بهره‌گیری از چارچوب تقطیر دانش [۱۵].
 - ادغام هدفمند ماژول توجه کانالی - مکانی در لایه‌های میانی شبکه پس‌زمینه به منظور تقویت ویژگی‌های مکانی و کانالی مرتبط با خطوط جاده، بدون افزایش چشمگیر بار محاسباتی
 - دستیابی به تعادل بهینه میان دقت و سرعت با ثبت امتیاز F1 معادل ۸۰/۲۰٪ در مجموعه داده CULane و سرعت استنتاج ۴۰۷ فریم بر ثانیه که بهبود ۰/۵۴٪ نسبت به روش CLRKNet [۱۶] و سرعت ۱/۷۹٪ برابر نسبت به روش ECBAM_ASPP [۱۷] را نشان می‌دهد.
 - کاهش چشمگیر خطای منفی - کاذب در مجموعه داده TuSimple با مقدار ۱/۵۷٪ که کمترین میزان در میان روش‌های مقایسه شده است و بیانگر توانایی بالا در تشخیص خطوط واقعی و جلوگیری از حذف آن‌ها است
- ساختار مقاله به این صورت است: در بخش ۲ مروری بر روش‌های پیشین بررسی شده است. بخش ۳ روش پیشنهادی به معرفی تقطیر دانش و ماژول CBAM اختصاص دارد. در بخش ۴ نتایج و بحث عملکرد مدل پیشنهادی ارزیابی شده و تحلیل نتایج ارائه می‌شود. در بخش ۵ نتیجه‌گیری پژوهش ارائه شده است.

¹ Spatial Convolutional Neural Networks

² Recurrent Feature-Shift Aggregator

به صورت رگرسیونی مدل‌سازی می‌کند. این روش‌ها به پردازش پسین برای گروه‌بندی نقاط نیاز دارند که از نظر محاسباتی پرهزینه است.

▪ **رویکردهای مبتنی بر منحنی:** این روش‌ها بر برازش خطوط با معادلات منحنی مختلف و پارامترهای خاص تمرکز دارند و تشخیص خط دوبعدی را به یک چالش رگرسیون پارامتر منحنی با شناسایی نقاط شروع و پایان تبدیل می‌کنند. یک نمونه خط به صورت مجموعه‌ای از پارامترهای منحنی نمایش داده می‌شود. LSTR [۳۴] شناسایی پارامترها را به صورت انتها-به-انتها^۴ با ترنسفورمر انجام می‌دهد. این روش‌ها استنتاج نسبتاً سریعی دارند، اما خطای یک پارامتر به‌طور کلی شکل خط را تحت تأثیر قرار می‌دهد.

روش‌های مبتنی بر بهبود فرایند یادگیری: علاوه بر طراحی معماری‌های جدید، بخشی از پژوهش‌ها بر بهبود فرایند آموزش شبکه‌های تشخیص خط تمرکز داشته‌اند. در این راستا، روشی با عنوان SAD^۶ [۳۵] معرفی شده است که هدف آن تقویت یادگیری نمایش در شبکه‌های تشخیص خط بدون افزایش هزینه استنتاج است. در این روش، نقشه‌های توجه استخراج‌شده از لایه‌های عمیق‌تر شبکه به‌عنوان سیگنال نظارتی درونی برای لایه‌های پیشین مورد استفاده قرار می‌گیرند تا انتقال اطلاعات زمینه‌ای و ساختاری صحنه بهبود یابد. این راهبرد بدون نیاز به برچسب‌گذاری اضافی، عملکرد مدل را در شرایط دشوار نظیر انسداد، نور نامناسب و خطوط مبهم ارتقا می‌دهد و امکان بهره‌گیری از شبکه‌های سبک با دقت بالا را فراهم می‌سازد. نتایج تجربی نشان می‌دهد که مدل‌های سبک مجهز به این مکانیزم می‌توانند با تعداد پارامتر کمتر و سرعت بالاتر، عملکردی رقابتی در مقایسه با روش‌های پیچیده‌تر ارائه دهند. این‌حال، این روش محدودیت‌هایی در یادگیری خطوط بسیار پیچیده دارد و حساس به کیفیت نقشه‌های توجه و ساختار داده‌های ورودی است. جدول (۱) مقایسه‌ای از روش‌های مختلف تشخیص خطوط و ویژگی‌های کلیدی هر رویکرد، شامل دقت، کارایی محاسباتی و پیچیدگی مدل، ارائه می‌دهد.

معمولاً از نقاط انکر خطی برای بازگشت به تغییرات موقعیتی نسبت به اهداف استفاده می‌کنند [۲۴، ۱۸]. انکرها به‌عنوان نقاط یا جعبه‌های فرضی اولیه برای شناسایی و تعیین موقعیت خطوط عمل می‌کنند و با استفاده از مدل‌های یادگیری ماشینی تنظیم و اصلاح می‌شوند. برای دورزدن محدودیت‌های انکرهای خطی، [۲۵] از یک فضای خط ذاتی استفاده کرده که کاندیدهای خط مختلفی تولید می‌کند. انکرهای ردیف طراحی شده به صورت ابتکاری پیکسل‌های ردیف را طبقه‌بندی می‌کنند [۲۶، ۲۷]. LaneATT [۱۸] از انکرهای خط برای یادگیری امتیاز اعتماد و جابه‌جایی مختصات محلی استفاده می‌کند. CLRNet^۱ [۲۸] در سال ۲۰۲۲ با پارامترهای انکر قابل یادگیری معرفی شد و از تخصیص بهینه ساده‌شده [۲۹] برای اختصاص پویا نزدیک‌ترین پیش‌بینی‌ها به مقادیر واقعی بهره می‌برد. UFLD^۲ [۲۳] ویژگی‌های سراسری را با مسطح کردن نقشه ویژگی‌ها ضبط کرده و طبقه‌بندی موقعیت خطوط را به صورت ردیفی یاد می‌گیرد. UFLDv2 [۳۰] نمایشی ردیفی - ستونی ارائه می‌دهد تا خطوط نزدیک به افقی را مدیریت کند. CondLaneNet^۳ [۲۶] یک نقشه حرارتی احتمال برای نقاط شروع خطوط یاد می‌گیرد و هسته‌های پیچشی پویا را استخراج می‌کند. Lane Former [۳۱] با استفاده از تکنیک‌های توجه مبتنی بر ترنسفورمر و تحلیل هم‌زمان ابعاد ردیف و ستون، شناسایی دقیق‌تر خطوط را در موقعیت‌های پیچیده ممکن می‌سازد.

▪ **رویکردهای مبتنی بر نقطه کلیدی:** این روش‌ها مدل‌سازی انعطاف‌پذیرتر و پراکنده‌تری از موقعیت‌های خط با تخمین مکان‌های نقطه ارائه می‌دهند و نقاط کلیدی متعلق به یک خط را با طرح‌های مختلف مرتبط می‌کنند [۳۰، ۱۸].

FOLOLane^۴ [۳۲] در سال ۲۰۲۱ خطوط را به‌عنوان نقاط کلیدی با الهام از تشخیص وضعیت بدن انسان شناسایی می‌کند. در سال ۲۰۲۲، GANet [۳۳] تغییرات موقعیتی نقاط کلیدی خطوط جاده را نسبت به نقاط شروع نمونه‌ها

⁴ Focus on local

⁵ End-to-End

⁶ Self-Attention Distillation

¹ Cross Layer Refinement Network

² Ultra-Fast Structure Aware Deep Lane Detection

³ Lane detection by conditional convolution

جدول (۱): مقایسه روش‌های تشخیص خط

روش	مدل	روش کار	مزایا	محدودیت
مبتنی بر تقسیم‌بندی	SCNN [۲۱]، RESA [۲۲]	طبقه‌بندی پیکسل	انسجام ساختاری بالا، دقت مناسب	هزینه محاسباتی بالا، نیاز به پس‌پردازش،
مبتنی بر انکر	LaneATT [۱۸]، CLRNet [۲۸]، UFLD/UFLDv2 [۲۳،۳۰]، CondLaneNet [۲۶]، LaneFormer [۳۱]	استفاده از نقاط/خطوط فرضی (انکر) و اصلاح با شبکه	سادگی، کارایی، امکان پیش‌بینی مختصات دقیق، انعطاف‌پذیری	محدودیت در انکرهای خطی، نیاز به طراحی مناسب انکرها
مبتنی بر نقاط کلیدی	FOLOLane [۳۲]، GANet [۳۳]	مدل‌سازی خطوط به‌صورت مجموعه نقاط کلیدی و اتصال آن‌ها	انعطاف‌پذیری بالا، بازسازی صریح ساختار خطوط	نیاز به پس‌پردازش برای گروه‌بندی نقاط، افزایش هزینه محاسباتی
مبتنی بر منحنی	LSTR [۳۴]	رگرسیون پارامترهای منحنی برای هر خط	سرعت استنتاج بالا، نمایش هندسی دقیق	حساس به خطا در تخمین پارامترها، شکل کلی خط ممکن است تحت تأثیر قرار گیرد
بهبود فرایند یادگیری	SAD [۳۵]	استفاده از نقشه‌های توجه لایه‌های عمیق برای هدایت یادگیری لایه‌های پیشین	بهبود عملکرد بدون افزایش هزینه استنتاج، مناسب خطوط مهم و شرایط دشوار	نیاز به طراحی مکانیزم توجه و محدودیت در یادگیری خطوط بسیار پیچیده، حساسیت به کیفیت نقشه‌های توجه

۳- روش پیشنهادی

مدل پیشنهادی با الهام از تقطیر دانش^۱ (چارچوب معلم - دانش‌آموز) و ساده‌سازی معماری در چارچوب مدل CLRKdNet طراحی شده که به‌دقت تشخیص بالا و کارایی بالاتر دست‌یافته و تعادل مطلوبی بین این دو معیار اساسی برقرار کرده است. این مدل به دلیل مصرف انرژی پایین و کارایی بالا برای سیستم‌های خودروبی ایده‌آل است. شکل (۱) فلوچارت روش پیشنهادی را نمایش می‌دهد.

۳-۱ تقطیر دانش

تقطیر دانش به‌عنوان یک تکنیک پیشرفته فشرده‌سازی مدل، با انتقال دانش از مدل معلم پیچیده به مدل دانش‌آموز سبک‌تر، امکان بهبود عملکرد و کاهش پیچیدگی محاسباتی را در کاربردهای بالاتر مانند تشخیص خطوط جاده فراهم می‌کند که این فرایند شامل دو بخش اصلی معلم و دانش‌آموز است. مدل معلم یک مدل پیچیده و حجیم است که معمولاً دارای معماری عمیق، تعداد زیادی

پارامتر و قابلیت پردازش چندمرحله‌ای است. وظیفه اصلی آن استخراج ویژگی‌های غنی و عمیق از داده‌ها و یادگیری بازنمایی‌های سطح بالا است. مدل معلم می‌تواند با استفاده از شبکه پس‌زمینه قدرتمند مانند ResNet-101 و ساختارهایی مانند (FPN)، ویژگی‌های محلی و جهانی تصویر را هم‌زمان شناسایی کند و پیش‌فرض‌های قابل‌یادگیری برای مسیرهای خطوط ایجاد نماید. ویژگی‌های کلیدی مدل معلم شامل توانایی استخراج ویژگی‌های پیچیده و چند مقیاسی از تصاویر، قابلیت انجام تعدیل و اصلاح‌های تکراری به‌منظور بهبود دقت پیش‌بینی‌ها و ارائه خروجی‌های دقیق و قابل‌اعتماد است که به‌عنوان مرجع اصلی برای آموزش مدل دانش‌آموز مورد استفاده قرار می‌گیرند.

در مقابل، مدل دانش‌آموز، نسخه‌ای سبک‌تر و بهبودیافته از مدل معلم است که برای کاربردهای با محدودیت زمانی و محاسباتی طراحی شده است. این مدل معمولاً دارای تعداد لایه‌ها و پارامترهای کمتر است و ساختاری ساده‌تر دارد تا سرعت اجرا افزایش یابد. هدف مدل دانش‌آموز این است که با تقلید رفتار و خروجی‌های

^۱Knowledge Distillation

تابع زیان کلی تقطیر دانش، مطابق رابطه (۱)، به صورت زیر تعریف می‌شود:

$$L_{dis} = W_{att}L_{att} + W_{prior}L_{prior} + W_{logit}L_{logit} \quad (1)$$

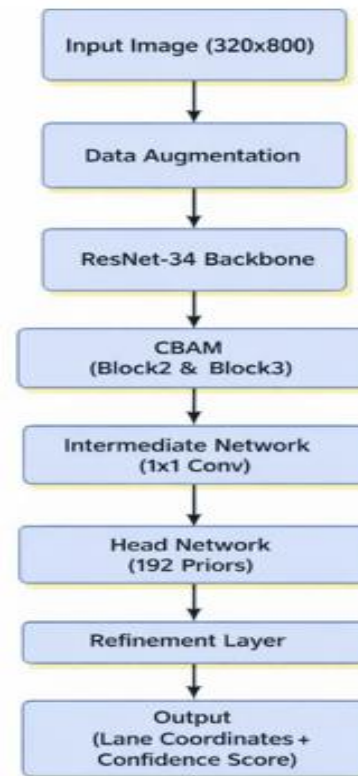
که در آن L_{dis} ، تابع زیان کل تقطیر دانش، L_{att} ، زیان تطبیق نقشه‌های توجه بین مدل معلم و دانش آموز، L_{prior} ، زیان تطبیق پیش فرض‌های اولیه خطوط، L_{logit} ، زیان تطبیق خروجی نهایی بین مدل معلم و دانش آموز و W_{att} ، W_{prior} ، W_{logit} ، ضرایب وزن دهی هر جز هستند که اهمیت نسبی آن‌ها را مشخص می‌کنند.

نوآوری اصلی در این پژوهش، ادغام ماژول توجه کانالی-فضایی در لایه‌های میانی مدل دانش آموز است که امکان تولید ویژگی‌های غنی را با سربار محاسباتی اندک فراهم می‌کند. این معماری پیشرفته در چارچوب سلسله‌مراتبی مبتنی بر شبکه‌های پس‌زمینه، میانی و سرطراحی شده و با اولویت‌بندی دقت در شرایط چالش‌برانگیز مانند سایه یا خطوط محو شده، تعادل مطلوبی بین دقت و سرعت برای کاربردهای کمک‌راننده و خودروهای خودران ایجاد می‌کند. مدل پیشنهادی بادقت بالا و حفظ کارایی بلادرنگ، پتانسیل قابل توجهی برای استقرار در سیستم‌های پیشرفته کمک‌راننده و وسایل نقلیه خودران دارا است.

▪ شبکه پس‌زمینه

شبکه ResNet [۱۴] یکی از پیشرفته‌ترین معماری‌های شبکه‌های عصبی عمیق است که در سال ۲۰۱۵ توسط شیوکی‌نگ رن و همکارانش در شرکت مایکروسافت معرفی شد. این معماری با استفاده از واحدهای باقیمانده و اتصالات کوتاه، مشکلات مربوط به افزایش عمق شبکه‌ها، به‌ویژه ناپدید شدن گرادینان‌ها، را برطرف کرده است. مفهوم کلید ResNet بر پایه یادگیری باقیمانده استوار است. شبکه ResNet-34 شامل ۳۴ لایه پیچشی در چهار بلوک است. تعداد کانال‌های خروجی ۶۴، ۱۲۸، ۲۵۶ و ۵۱۲ است. هر بلوک از دو لایه پیچشی با کرنل ۳×۳ تشکیل شده که پس از آن نرمال‌سازی دسته‌ای^۲ و تابع فعال‌سازی ReLU قرار می‌گیرد. اتصال میان‌بری خروجی بلوک را مستقیماً به ورودی اضافه می‌کند و به این ترتیب

مدل معلم، همان عملکرد را با هزینه محاسباتی کمتر ارائه دهد. برای رسیدن به این هدف، مدل دانش آموز از فرایندهای زیر استفاده می‌کند:



شکل (۱): فلوچارت روش پیشنهادی

تطبیق نقشه‌های توجه: نقشه‌های توجه تولید شده توسط مدل دانش آموز با نقشه‌های توجه مدل معلم مقایسه و تطبیق داده شدند. این تطبیق به مدل دانش آموز کمک کرد تا بتواند مناطق مهم در تصویر را دقیق‌تر شناسایی کرده و تمرکز بر ویژگی‌های کلیدی داشته باشد که در نهایت منجر به بهبود عملکرد تشخیص و تعمیم‌پذیری بهتر مدل می‌شود.

تطبیق پیش فرض‌های^۱ اولیه: پیش فرض‌های اولیه مدل معلم، مانند اطلاعات اولیه درباره موقعیت خطوط، به مدل دانش آموز منتقل شدند تا فرایند یادگیری سریع‌تر و دقیق‌تر شود.

تطبیق خروجی‌های نهایی: خروجی‌های نهایی مدل معلم با خروجی‌های مدل دانش آموز مقایسه شدند تا دقت پیش‌بینی‌ها بهبود یابد.

² Batch Normalization

¹ Priors

وابستگی‌های بین‌کانالی با صرفه‌جویی در پارامترها انجام شود. خروجی این ماژول، با تخصیص وزن‌های متفاوت به هر کانال، موجب تقویت ویژگی‌های معنایی کلیدی و کاهش اهمیت ویژگی‌های غیرضروری می‌شود.

- **توجه فضایی:** در این مؤلفه، ابتدا اطلاعات کانال‌ها با تجمیع میانگین و تجمیع بیشینه در امتداد محور کانالی فشرده می‌شوند و دو نقشه دوبعدی $(F_{max}^s, F_{avg}^s) \in F$ حاصل می‌گردد. این دو نقشه طبق رابطه‌ی (۳) به‌صورت کانال‌محور به یکدیگر الحاق شده و از یک‌لایه کانولوشن با فیلتر 7×7 عبور می‌کنند تا نقشه توجه مکانی $(M_s(F))$ تولید شود:

$$M_s(F) = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) = \sigma(f^{7 \times 7}([f_{avg}^s; f_{max}^s])) \quad (3)$$

دلیل استفاده از CBAM در روش پیشنهادی، افزایش توانایی مدل در تشخیص خطوط در شرایط چالش‌برانگیز مانند نور کم، سایه یا خطوط محو است. مزیت اصلی CBAM ناشی از توجه هم‌زمان کانالی و فضایی است که به شبکه امکان می‌دهد هم ویژگی‌های شاخص هر کانال و هم مناطق مهم تصویر را برجسته کند. برخلاف روش‌های پیچیده‌تر که ممکن است سرعت پردازش یا دقت را کاهش دهند، استفاده از CBAM در مدل پیشنهادی دقت را افزایش می‌دهد درحالی‌که تأثیر ناچیزی بر سرعت استنتاج^۲ دارد. این ویژگی، همراه با سبک بودن محاسبات، روش ما را برای کاربردهای بلادرنگ مانند سیستم‌های رانندگی خودکار بسیار مناسب می‌کند.

▪ شبکه میانی

شبکه میانی ویژگی‌های چند مقیاسی استخراج‌شده از شبکه پس‌زمینه را دریافت می‌کند و با استفاده از کانولوشن 1×1 تعداد کانال‌ها را به ۶۴ کاهش می‌دهد. این فشرده‌سازی ضمن حفظ اطلاعات معنایی مهم، بار محاسباتی را کاهش داده و سرعت استنتاج شبکه را افزایش می‌دهد. شبکه میانی قادر است ویژگی‌های سطح بالا و سطح پایین تصویر را با هم ادغام کند تا نمایشی غنی و جامع از تصویر ایجاد شود، به‌طوری‌که جزئیات محلی و

مسیر گرادیان کوتاه‌تر و پایدارتر می‌شود. در روش پیشنهادی به‌عنوان یک نوآوری برای تعیین نواحی موردتوجه، ماژول CBAM به بلوک دوم و سوم شبکه ResNet اضافه شده است تا ویژگی‌های کلیدی تقویت شوند که این امر توسط دو حالت زیر صورت می‌گیرد:

- **توجه کانالی:** با استفاده از ماژول، ویژگی‌های معنایی مهم در ابعاد کانال تقویت می‌شوند.
- **توجه فضایی:** با استفاده از ماژول، تمرکز مدل بر نواحی کلیدی تصویر (مانند خطوط جاده) افزایش می‌یابد.

۳-۱-۱- ماژول توجه کانالی - مکانی CBAM

برای تقویت تشخیص خطوط در شرایط پیچیده، ماژول توجه کانالی-مکانی CBAM در مراحل دوم و سوم ResNet-34 ادغام شد. CBAM با دو مکانیزم عمل می‌کند:

- **توجه کانالی:** این مؤلفه با به‌کارگیری تجمیع میانگین سراسری و تجمیع بیشینه سراسری بر روی نقشه‌های ویژگی، به مدل‌سازی وابستگی‌های بین کانالی پرداخته و نقشه توجه کانالی $(M_c(F))$ را تولید می‌کند. این فرایند موردنظر در رابطه (۲) به شرح زیر تعریف می‌شود:

$$M_c(F) = \sigma(MLP(AvgPool(F) + MLP(MaxPool(F)))M_c(F) = \sigma(w_1(w_0(f_{avg}^c)) + (w_1(w_0(f_{max}^c)))) \quad (2)$$

که در آن $(M_c(F))$ نقشه توجه کانالی (خروجی نهایی) است. همچنین F ، نقشه ویژگی ورودی، F_{cavg} ، میانگین‌گیری سراسری $(AvgPool(F))$ ، F_{cmax} ، بیشینه‌گیری سراسری $(MaxPool(F))$ ، W_0 ، وزن لایه اول، W_1 ، وزن لایه دوم و σ ، تابع سیگموئید است.

در این رابطه، (F_{avg}^c, F_{max}^c) به ترتیب بیانگر تجمیع میانگین و بیشینه هستند. همچنین W_0 و W_1 وزن‌های شبکه پرسپترون چندلایه^۱ (MLP) با یک‌لایه مخفی به ابعاد $R^{C \times 1 \times 1}$ بوده که در آن r نسبت کاهش است. نسبت کاهش تعیین می‌کند که تعداد کانال‌ها در لایه مخفی MLP چند برابر کمتر از تعداد کانال‌های ورودی باشد، تا پیچیدگی محاسبات کاهش یابد و یادگیری

² Inference speed

¹ Multilayer Perceptron

به‌راحتی در سیستم‌های بزرگ‌تر و کاربردهای بلادرنگ مانند رانندگی خودران و سیستم‌های پیشرفته کمک‌راننده ادغام شود و درعین‌حال، با تعداد ثابت پیش‌فرض‌ها و یک‌لایه پالایش، ترکیبی از سرعت، دقت و مقیاس‌پذیری را ارائه دهد. در نهایت، شبکه سر به‌عنوان بخش پایانی، با دریافت ویژگی‌های یکپارچه‌شده از شبکه میانی، به تولید پیش‌بینی نهایی می‌پردازد. این بخش با انجام عملیات رده‌بندی و بازگشت، وجود خطوط را تشخیص داده و مختصات دقیق و پارامترهای هندسی آن‌ها را پیش‌بینی می‌کند. خروجی شبکه سر شامل ۱۹۲ پیش‌فرض^۲ است که هر یک نمایانگر یک کاندیدای خط جاده است.

هر پیش‌فرض شامل مجموعه‌ای از مختصات نقاط نمونه‌برداری شده در امتداد محور عمودی تصویر به همراه یک نمره اطمینان^۳ برای بیان احتمال وجود خط است. به‌عبارت‌دیگر، برای هر یک از ۱۹۲ پیش‌فرض، مدل هم‌زمان مختصات هندسی خط و احتمال تعلق آن به کلاس خط معتبر را پیش‌بینی می‌کند. همکاری هماهنگ این سه بخش در کنار نوآوری‌های اعمال‌شده در شبکه پس‌زمینه، منجر به دستیابی به تعادل مطلوب بین دقت و سرعت شده است. این دستاورد نشان‌دهنده ایجاد همکاری مؤثر بین معماری بهبودیافته و فرایند انتقال دانش است که در نهایت منجر به ایجاد الگویی دقیق‌تر و کارآمد برای کاربردهای نیازمند پردازش بلادرنگ شده است. جزئیات معماری در شکل (۲) نشان داده شده‌اند. در این شکل ورودی شبکه یک تصویر رنگی (RGB) با ابعاد 320×800 پیکسل است که پس از تغییر اندازه از تصویر اصلی استخراج می‌شود. پیش از ورود به شبکه، تصویر تحت مراحل پیش‌پردازش شامل نرمال‌سازی مقادیر پیکسل‌ها در بازه $[0, 1]$ و استانداردسازی بر اساس میانگین و انحراف معیار مجموعه‌داده قرار می‌گیرد. همچنین در مرحله آموزش، تکنیک‌های افزایش داده نظیر چرخش جزئی، تغییر روشنایی و برش تصادفی به‌منظور بهبود تعمیم‌پذیری مدل اعمال می‌شوند.

اطلاعات معنایی حفظ شوند. جزئیات ساختاری شبکه میانی و ارتباط آنها با شبکه اصلی در جدول (۲) آورده شده است.

جدول (۲): جزئیات ساختاری شبکه میانی و ارتباط آن با بخش

های اصلی

بخش شبکه	لایه/ماژول	ابعاد خروجی	تعداد کانال	پارامترهای کلیدی
ورودی	تصویر RGB	$320 \times 800 \times 3$	۳	نرمال‌سازی
Backbone	ResNet-34Block1	200×80	۶۴	Con3x3
Backbone	ResNet-34Block2+ CBAM	100×40	۱۲۸	Attention
Backbone	ResNet-34Block3+ CBAM	50×20	۲۵۶	Attention
Backbone	ResNet-34Block4	25×10	۵۱۲	Residual
شبکه میانی	Conv1x1	-	۶۴	کاهش کانال
شبکه سر	192priors	-	-	مختصات + نمره
Refinement	Refine layer	-	-	اصلاح مختصات

■ شبکه سر

شبکه سر برای پیش‌بینی دقیق مختصات خطوط جاده طراحی شده و نقش کلیدی در استخراج ویژگی‌های هندسی خطوط دارد. این شبکه شامل ۱۹۲ پیش‌فرض است. این پیش‌فرض‌ها در حقیقت مختصات اولیه‌ای هستند که شبکه بر اساس آن‌ها موقعیت و شکل خطوط را پیش‌بینی و اصلاح می‌کند. تعداد ۱۹۲ انتخاب شد؛ زیرا آزمایش‌ها نشان دادند که این مقدار تعادل مناسبی بین پوشش کامل خطوط در تصویر و کارایی محاسباتی برقرار می‌کند؛ به‌طوری‌که تعداد کمتر از این مقدار باعث کاهش دقت و ازدست‌رفتن برخی خطوط می‌شود و تعداد بسیار بیشتر نیز بار محاسباتی را افزایش می‌دهد بدون آنکه بهبود محسوس در دقت ایجاد کند؛ بنابراین، مقدار ۱۹۲ به‌عنوان یک نقطه بهینه در این پژوهش انتخاب شد. برای بهبود و اصلاح مختصات اولیه، از یک‌لایه پالایش^۱ استفاده می‌کند که مقادیر اولیه مختصات و پارامترهای هندسی خطوط را دقیق‌تر می‌کند. این طراحی ساده و سبک باعث می‌شود شبکه

¹ Refine_layers

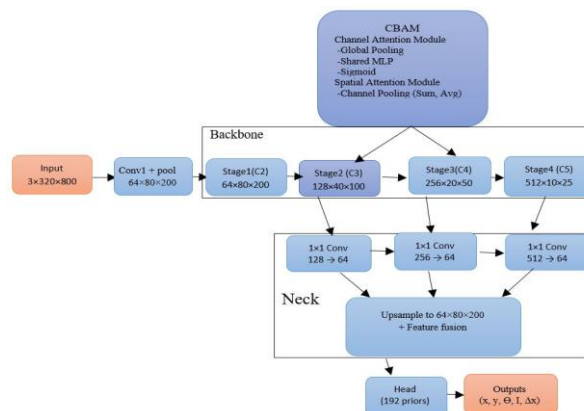
² Priors

گردآوری شده است. این مجموعه شامل ۱۳۳۲۳۵ فریم تصویری است که به سه بخش تقسیم شده است: ۸۸۸۰ فریم برای آموزش، ۹۶۷۵ فریم برای اعتبارسنجی و ۳۴۶۸۰ فریم برای آزمون. بخش آزمون این مجموعه شامل نه سناریو است که یک سناریوی نرمال و هشت سناریوی چالش‌برانگیز را در بر می‌گیرد؛ از جمله شرایط نوری نامطلوب، وجود سایه و پیچیدگی‌های هندسی مسیر که ارزیابی الگوریتم‌ها را به شرایط واقعی نزدیک‌تر می‌سازد در این مجموعه، تمرکز بر حاشیه‌نویسی دقیق چهار خط اصلی جاده است که به صورت دستی و با استفاده از روش اسپلاین مکعبی انجام شده است. این روش امکان ترسیم خطوط نرم و پیوسته را فراهم می‌کند، حتی در مواردی که خطوط به دلیل موانع یا شرایط محیطی مسدود شده‌اند. خطوط فرعی یا موانع مسیر مخالف حاشیه‌نویسی نشده‌اند تا تمرکز الگوریتم‌ها صرفاً بر خطوط کلیدی جاده باشد. این ویژگی‌ها CULane را به گزینه‌ای ایده‌آل برای ارزیابی عملکرد مدل‌های تشخیص خطوط در شرایط پیچیده شهری تبدیل کرده است. شکل (۳) نمونه‌هایی از مجموعه‌داده CULane در شرایط چالش‌برانگیز است.



شکل (۳): نمونه‌هایی از مجموعه‌داده CULane

مجموعه‌داده TuSimple بر صحنه‌های بزرگراهی تمرکز دارد و شامل ۳۲۶۸ تصویر آموزشی، ۳۵۸ تصویر اعتبارسنجی و ۲۷۸۲ تصویر آزمایشی است. رزولوشن تمامی تصاویر 720×1280 پیکسل است. در این مجموعه‌داده، خطوط جاده با استفاده از چندخطی‌ها نشانه‌گذاری شده‌اند، به طوری که در هر فریم حداکثر پنج خط و معمولاً چهار خط قابل مشاهده است ویژگی‌ای که باتوجه به قرارگیری خطوط در مرکز میدان دید، برای سامانه‌های رانندگی خودکار اهمیت بالایی دارد. همچنین، معیار دقت به‌عنوان شاخص اصلی ارزیابی در این مجموعه‌داده مورد استفاده قرار می‌گیرد. شکل (۴) نمونه‌ای از مجموعه‌داده TuSimple است.



شکل (۲): معماری مدل پیشنهادی (شبکه دانش‌آموز: شامل شبکه پس‌زمینه، ماژول CBAM، شبکه میانی و شبکه سر)

۴- نتایج

در این بخش، عملکرد مدل پیشنهادی که مبتنی بر ادغام مکانیزم توجه CBAM در چارچوب تقطیر دانش است، مورد ارزیابی قرار می‌گیرد. ارزیابی‌ها بر اساس دو مجموعه‌داده در این روش، دانش مدل معلم به مدل دانش‌آموز منتقل شده و هم‌زمان با بهره‌گیری از توجه کانالی و مکانی، ویژگی‌های معنادار تقویت می‌شوند. در روش پیشنهادی، دانش مدل معلم به مدل دانش‌آموز منتقل شده و هم‌زمان با بهره‌گیری از توجه کانالی و مکانی، ویژگی‌های معنادار ادامه، نتایج کمی مدل بر روی مجموعه‌داده‌های CULane و TuSimple ارائه و تحلیل می‌شود. ارزیابی‌های کمی بر اساس معیارهای استاندارد شامل F1-score، صحت، بازخوانی و نرخ فریم بر ثانیه انجام شدند تا عملکرد مدل پیشنهادی در مقایسه با مدل‌های پیشرفته سنجیده شود.

۴-۱ مجموعه‌داده

مجموعه‌داده‌های مورد استفاده در این پژوهش، CULane و TuSimple است که به دلیل تنوع بالای نمونه‌ها، نمایانگر سناریوهای واقعی رانندگی بوده و شامل تصاویر پیچیده با شرایط متنوع نوری و جوی است. این ویژگی‌ها امکان ارزیابی مدل را در موقعیت‌های مختلف فراهم می‌سازد. مجموعه‌داده CULane یکی از جامع‌ترین پایگاه‌های داده برای پژوهش در زمینه تشخیص خطوط جاده است که از ویدئوهای ضبط شده توسط دوربین‌های نصب‌شده روی شش خودروی مختلف در محیط‌های شهری پکن

پایه‌سازی مدل پیشنهادی با استفاده از زبان برنامه‌نویسی پایتون و چارچوب PyTorch انجام شد که به دلیل انعطاف‌پذیری و پشتیبانی قوی از عملیات یادگیری عمیق انتخاب شد. مدل پیشنهادی طی ۳۰ دوره و با اندازه دسته ۲۴ انجام شد. برای بهینه‌سازی از الگوریتم AdamW با نرخ یادگیری اولیه 4×10^{-4} و ضریب کاهش وزن 1×10^{-4} استفاده گردید. تابع زیان مورد استفاده ترکیبی از چهار مؤلفه اصلی بود: زیان IOU برای افزایش هم‌پوشانی پیش‌بینی با ناحیه هدف، زیان طبقه‌بندی منظور جداسازی دقیق خطوط از پس‌زمینه، زیان مختصات برای تخمین دقیق مکان خطوط، و زیان تقسیم‌بندی معنایی جهت بهبود شناسایی نواحی مرتبط با خطوط ضرایب این اجزاء به صورت تجربی انتخاب شدند تا تعادل مناسبی میان دقت وظایف مختلف (IOU، طبقه‌بندی، پیش‌بینی مختصات و تقسیم‌بندی معنایی) و کارایی محاسباتی برقرار گردد. این راهبرد منجر به بهبود عملکرد مدل بر روی مجموعه‌داده‌های TUSimple و CULane شد. برای ارزیابی عملکرد بلادرنگ، مدل روی RTX 4090 آزمایش شد که یک پلتفرم قدرتمند و مناسب برای کاربردهای پیشرفته خودرویی است. تنظیمات آموزش در جدول (۳) آورده شده است.

جدول (۳): تنظیمات آموزش مدل پیشنهادی

پارامتر	مقدار
Dataset	CULane, TUSimple
Epochs	۳۰
Batch size	۲۴
Optimizer	AdamW
Learning rate	4×10^{-4}
Weight decay	1×10^{-4}
Loss function	IoU + Classification + Coordinate + Semantic Segmentation
Hardware	NVIDIA RTX 4090
Framework	PyTorch (Python)

۴-۴ معیارهای ارزیابی

ارزیابی عملکرد مدل پیشنهادی با استفاده از معیارهای استاندارد و تحلیل کمی انجام شد. معیارهای کمی شامل صحت، دقت، بازخوانی و F1 هستند.



شکل (۴): نمونه‌هایی از مجموعه‌داده TUSimple

۴-۲ پیش‌پردازش داده

تصاویر ورودی به رزولوشن 320×800 پیکسل تغییر اندازه داده شدند. به منظور افزایش استحکام مدل و بهبود تعمیم‌پذیری آن، مجموعه‌ای از تکنیک‌های افزایش داده اعمال گردید که شامل موارد زیر بود: اعمال چرخش افقی با احتمال ۵۰٪ برای شبیه‌سازی زوایای دید مختلف، تنظیم روشنایی و کنتراست در محدوده ۰/۸۵ تا ۱/۱۵ و همچنین تغییر رنگ و اشباع در محدوده ± 15 واحد، هر یک با احتمال ۸۰٪، برای تقلید شرایط نوری متغیر. علاوه بر این، اثرات حرکتی با اعمال تار کردن با کرنل ۳ تا ۵ پیکسل و احتمال ۳۰٪ شبیه‌سازی شد. برای در نظر گرفتن تغییرات هندسی، تحولات افاین شامل چرخش ± 10 درجه، جابه‌جایی ± 0.1 و تغییر مقیاس در محدوده ۰/۹ تا ۱/۱ با احتمال ۸۰٪ استفاده گردید. این فرایندهای پیش‌پردازش به طور جمعی منجر به ایجاد مدلی شد که در برابر تغییرات گسترده در شرایط جاده‌ای، از جمله نوسانات نوری، شرایط جوی و تغییرات هندسی، عملکرد قوی‌تری از خود نشان می‌دهد. همچنین، این تکنیک‌ها با غنی‌سازی تنوع داده‌های آموزشی، به طور مؤثری از بیش‌برازش مدل جلوگیری کرده و قابلیت اطمینان آن را در مواجهه با سناریوهای واقعی بهبود بخشیدند.

۴-۳ فرایند آموزش

فرایند آموزش با استفاده از مجموعه‌داده‌های CULane [۱۲] و TUSimple [۱۳] برای آموزش اصلی و برای ارزیابی انجام شد. برای بهبود کارایی و کیفیت مدل، فرایند آموزش شامل چندین مرحله کلیدی بود که در ادامه توضیح داده می‌شود.

۴-۵ نتایج کمی

ارزیابی‌های کمی بر اساس معیارهای استاندارد شامل FI امتیاز، دقت، صحت، بازخوانی و نرخ فریم بر ثانیه انجام شدند تا عملکرد مدل پیشنهادی در مقایسه با مدل‌های پیشرفته سنجیده شود. این معیارها روی مجموعه داده‌های TUSimple و CULane محاسبه شد. نتایج در جداول زیر ارائه شده است. جدول (۴) نتایج مقایسه‌ای مدل پیشنهادی با دیگر مدل‌های پیشرفته را در مجموعه داده CULane نشان می‌دهد.

جدول (۴) مقایسه‌ای جامع میان روش پیشنهادی و تعدادی از پیشرفته‌ترین روش‌های تشخیص خط، از جمله SCNN, SAD, UFLDv2, UFLD, CondLaneNet, LaneATT و CLRLNet و CLRKNet ارائه می‌دهد.

معیار اصلی سنجش عملکرد، امتیاز FI به‌عنوان شاخص ترکیبی دقت و بازخوانی است و سرعت پردازش نیز بر حسب فریم بر ثانیه گزارش شده است. علاوه بر این، کارایی روش‌ها در سناریوهای متنوع محیطی شامل شرایط عادی، ترافیک متراکم، خیرگی نور، سایه، نبود خط واضح، وجود علائم سطح جاده، مسیرهای منحنی، تقاطع و شب به‌صورت درصد بیان شده است. بررسی عملکرد مدل در شرایط چالش‌برانگیز نشان‌دهنده پایداری و مقاومت آن است.

عملکرد مطلوب در شرایط عادی، محیط‌های پرتراffic و حضور سایه‌نشان می‌دهد که مدل در برابر پیچیدگی صحنه و تغییرات روشنایی مقاوم است و بیانگر قابلیت تعمیم مناسب مدل در وضعیت‌های نوری و هندسی گوناگون است. در سناریوهای فاقد خطوط واضح، مدل با اتکا به اطلاعات زمینه‌ای و ساختار فضایی تصویر، تشخیص قابل‌اعتمادی ارائه می‌دهد. به‌منظور تضمین هم‌خوانی کامل با شیوه گزارش‌دهی مطالعات پیشین و جلوگیری از هرگونه مقایسه ناهمگون، در دسته‌بندی Cross صرفاً شاخص مثبت کاذب به‌عنوان معیار ارزیابی لحاظ شده است و سایر شاخص‌ها مورد بررسی قرار نگرفته‌اند.

دقت: معیار دقت، نسبت نقاط خط حرکت است که به‌درستی شناسایی شده‌اند. یک پیش‌بینی زمانی صحیح در نظر گرفته می‌شود که بیش از ۸۵٪ نقاط پیش‌بینی شده در فاصله ۲۰ پیکسلی از خطوط واقعی قرار گیرند. این معیار توسط رابطه (۴) محاسبه می‌شود:

$$ACC = \frac{\sum_{clip} C_{clip}}{\sum_{clip} S_{clip}} \quad (4)$$

که در آن، Cclip تعداد نقاطی است که در محدوده ۲۰ پیکسلی از خط واقعی قرار دارند و Sclip تعداد کل نقاط خط در تصویر است. صحت (P): نسبت نمونه‌های مثبت که به‌درستی شناسایی شده‌اند به کل نمونه‌های پیش‌بینی شده به‌عنوان مثبت را نشان می‌دهد و توانایی مدل در شناسایی صحیح خطوط جاده را ارزیابی می‌کند (رابطه (۵)):

$$P = \frac{TP}{TP + FP} \quad (5)$$

بازخوانی (R): نسبت نمونه‌های مثبت واقعی است که به‌درستی شناسایی شده‌اند و توانایی مدل در پوشش کامل خطوط موجود در تصویر را می‌سنجد رابطه (۶) بیانگر این موضوع هست:

$$R = \frac{TP}{TP + FN} \quad (6)$$

امتیاز FI: میانگین هارمونیک دقت و بازخوانی است و معیاری جامع و متوازن برای ارزیابی عملکرد مدل ارائه می‌دهد که از رابطه (۷) محاسبه می‌شود:

$$FI = \frac{(Precision \times Recall) \times 2}{(Precision + Recall)} \quad (7)$$

نسبت‌های مثبت و منفی کاذب مطابق روابط (۸) و (۹) تعریف می‌شوند:

$$FP = \frac{F_{pred}}{N_{pred}} \quad (8)$$

$$FN = \frac{M_{pred}}{N_{gt}} \quad (9)$$

در این جا F_{pred} تعداد خطوط پیش‌بینی شده اشتباه، N_{pred} تعداد کل خطوط پیش‌بینی شده، M_{pred} تعداد خطوط واقعی که مدل شناسایی نکرده، N_{gt} تعداد کل خطوط واقعی است.

جدول (۴): نتایج مقایسه‌ای مدل پیشنهادی در مجموعه داده Culan

Method	Backbone	F1(%)	FPS	Normal	Crowd	Dazzle	Shadow	No line	Arrow	Curve	Cross	Night
SCNN [۲۱]	VGG16	۷۱/۶۰	۵۰	۹۰/۶۰	۶۹/۷۰	۵۸/۵۰	۶۶/۹۰	۴۳/۴۰	۸۴/۱۰	۶۴/۴۰	۱۹۹۰	۶۶/۱۰
LaneATT [۱۸]	ResNet18	۷۵/۱۳	۲۵۰	۹۱/۱۷	۷۲/۷۱	۶۵/۸۲	۶۸/۰۳	۴۹/۱۳	۸۷/۸۲	۶۳/۷۲	۱۰۲۰	۶۸/۵۸
LaneATT [۱۸]	ResNet34	۷۶/۶۸	۱۷۰	۹۲/۱۴	۷۵/۰۳	۶۴/۴۷	۷۸/۱۵	۴۹/۳۹	۸۸/۳۸	۶۷/۷۲	۱۳۳۰	۷۰/۷۲
CondLaneNet [۲۶]	ResNet18	۷۸/۱۴	۳۴۸	۹۲/۸۷	۷۵/۷۰	۷۰/۷۲	۸۰/۰۱	۵۲/۳۹	۸۹/۳۷	۷۲/۴۰	۱۳۶۴	۷۳/۲۳
CondLaneNet [۲۶]	ResNet34	۷۸/۷۴	۲۳۷	۹۳/۳۸	۷۷/۱۴	۷۱/۱۷	۷۹/۹۳	۵۱/۸۵	۸۹/۸۹	۷۳/۸۸	۱۳۸۷	۷۳/۹۲
UFLD [۲۳]	ResNet18	۶۸/۴۰	۶۲۰	۸۷/۷۰	۶۶/۰۰	۵۸/۴۰	۶۲/۸۰	۴۰/۲۰	۸۱/۰۰	۵۷/۹۰	۱۷۴۳	۶۲/۱۰
UFLD [۲۳]	ResNet34	۷۲/۳۰	۳۴۰	۹۰/۷۰	۷۰/۲۰	۵۹/۵۰	۶۹/۳۰	۴۴/۴۰	۸۵/۷۰	۶۹/۵۰	۲۰۳۷	۶۶/۷۹
SAD [۳۵]	ENet	۷۰/۸۰	۲۰۰	۹۰/۱۰	۶۸/۸۰	۶۰/۲۰	۶۵/۹۰	۴۱/۶۰	۸۴/۰۰	۶۵/۷۰	۱۹۹۸	۶۶/۰۰
SAD [۳۵]	ResNet101	۷۱/۸۰	۴۰	۹۰/۷۰	۷۰/۰۰	۵۹/۹۰	۶۷/۰۰	۴۳/۵۰	۸۴/۴۰	۶۵/۷۰	۲۰۵۲	۶۶/۳۰
UFLDv2 [۳۰]	ResNet18	۷۴/۷۰	۳۳۰	۹۱/۷۰	۷۳/۰۰	۶۴/۶۰	۷۴/۷۰	۴۷/۲۰	۸۷/۶۰	۶۸/۷۰	۱۹۹۸	۷۰/۲۰
UFLDv2 [۳۰]	ResNet34	۷۵/۹۰	۱۶۵	۹۲/۵۰	۷۴/۹۰	۶۵/۷۰	۷۵/۳۰	۴۹/۰۰	۸۸/۵۰	۷۰/۲۰	۱۸۶۴	۷۰/۶۰
CLRLNet [۲۸]	ResNet18	۷۹/۵۸	۲۷۵	۹۳/۳۰	۷۸/۳۳	۷۳/۷۱	۷۹/۶۶	۵۳/۱۵	۹۰/۲۵	۷۱/۵۶	۱۳۲۱	۷۵/۱۱
CLRLNet [۲۸]	ResNet34	۷۹/۷۳	۱۰۳/۱۵۶	۹۳/۴۹	۷۸/۰۶	۷۴/۵۷	۷۹/۹۲	۵۴/۱۰	۹۰/۵۹	۷۲/۷۷	۱۲۱۶	۷۵/۰۲
CLRLNet [۲۸]	ResNet101	۸۰/۱۳	۴۶	۹۳/۸۵	۷۸/۷۸	۷۲/۴۹	۸۲/۳۳	۵۴/۵۰	۸۹/۷۹	۷۵/۵۷	۱۲۶۲	۷۵/۵۱
CLRKNet [۱۶]	ResNet18	۷۹/۶۶	۴۵۰	۹۳/۳۴	۷۸/۱۹	۷۴/۶۲	۸۰/۴۶	۵۲/۶۱	۸۹/۹۸	۶۹/۱۷	۹۰۴	۷۶/۵۴
ECBAM_ASPP [۱۷]	ResNet34	۷۸/۴۵	۲۲۷/۳	۹۴/۰۸	۷۷/۱۷	۶۸/۵۶	۷۹/۲۰	۵۱/۱۸	۸۸/۷۲	۷۵/۴۱	۲۵۵۹	۷۳/۷۱
روش پیشنهادی	ResNet34	۸۰/۲۰	۴۰۷	۹۳/۶۲	۷۹/۱۰	۷۲/۶۲	۸۲/۵۹	۵۳/۶۷	۹۰/۵۳	۷۰/۲۷	۱۰۷۵	۷۵/۰۹

مقایسه در آن است که تشخیص دقیق و سریع خطوط جاده یکی از مهم‌ترین چالش‌های سامانه‌های بینایی ماشین در خودروهای هوشمند و خودران محسوب می‌شود. وجود عواملی مانند سایه‌ها، شلوغی ترافیکی، و حتی محوشدن یا از بین رفتن خطوط جاده باعث می‌شود که مدل‌ها در شرایط واقعی عملکردی متفاوت از آزمایشگاه داشته باشند؛ بنابراین تحلیل جامع نتایج این جدول می‌تواند نمایی روشن از نقاط قوت و ضعف روش‌های مختلف و همچنین جایگاه روش پیشنهادی در میان آن‌ها ارائه دهد. نخستین معیار مهم،

در مجموع، یافته‌ها نشان می‌دهد که روش پیشنهادی علاوه بر دستیابی به بالاترین دقت کلی، از نظر سرعت پردازش و پایداری در شرایط محیطی متنوع نیز عملکردی متعادل و رقابتی ارائه می‌دهد. این نتایج نشانگر قابلیت به‌کارگیری عملی مدل در سامانه‌های تشخیص خط و کاربردهای مرتبط با سامانه‌های کمک‌راننده پیشرفته است. جدول مقایسه‌ای ارائه‌شده عملکرد مجموعه‌ای از روش‌های مختلف تشخیص خطوط جاده را در شرایط متنوع و با معیارهای گوناگون بررسی می‌کند. اهمیت این



تکنیک‌های بهینه‌سازی تکمیل شود. روش پیشنهادی در مقایسه با مدل CLRkdNet در سناریوهای چالش‌برانگیز مانند سایه (۸۲/۵۹)، ازدحام (۷۹/۱۰)، سناریوی فلش (۹۰/۵۳)، بدون خط (۵۳/۶۷) و خطوط منحنی (۷۰/۲۷) عملکرد بهتری نسبت به CLRkdNet نشان داد. این بهبود به دلیل تقویت ویژگی‌های کانالی و فضایی در لایه‌های میانی توسط CBAM است که به مدل کمک می‌کند خطوط را در شرایط مختلف بهتر تشخیص دهد. با این حال، در سناریوی شب (۷۵/۰۹) و نور خیره‌کننده (۷۲/۶۲)، عملکرد اندکی ضعیف‌تر از CLRkdNet بود که ممکن است به دلیل از دست دادن برخی اطلاعات چند مقیاسی در مراحل میانی باشد و همچنین در مقایسه با مدل ECBAM_ASPP یک مدل دیگر در حوزه تشخیص خطوط که از CBAM استفاده کرده است، نتایج نشان می‌دهند که مدل ما، هم در دقت تشخیص و هم در سرعت استنتاج بهتر از مدل ECBAM_ASPP عمل کرده است. روش پیشنهادی با سرعتی حدود ۱/۷۹ برابر سریع‌تر از این روش عمل می‌کند و امتیاز F1 نیز از ۷۸/۴۵٪ به ۸۰/۲۰٪ بهبود یافته است. مقایسه با UFLD: اگرچه UFLD با ۶۲۰ FPS سریع‌ترین مدل است، اما امتیاز آن (۶۸/۴۰٪) کمتر از روش پیشنهادی است. این نشان می‌دهد که روش ما تعادل بهتری بین سرعت و دقت ارائه می‌دهد. مدل پیشنهادی با ثبت ۴۰۷ فریم بر ثانیه نسبت به CLRkdNet با ۴۵۰ فریم بر ثانیه کاهش جزئی در سرعت استنتاج نشان داده است. علت این افت کارایی، به‌کارگیری ماژول CBAM در لایه‌های میانی برای بهبود دقت است که موجب افزایش بار محاسباتی شده است. با این حال، سخت‌افزار پردازشی مورد استفاده در آزمایش‌ها توانسته است اثرات منفی ناشی از افزایش بار محاسباتی را به میزان قابل توجهی کاهش دهد و موجب گردد عملکرد مدل در سطح نسبتاً بالا ۴۰۷ فریم بر ثانیه به صورت پایدار حفظ شود.

در جدول (۵) امتیاز F1 و FPS روش‌های پیشرفته با شبکه پس‌زمینه یکسان بررسی شده است. این جدول مقایسه عملکرد روش‌های مختلف را بر روی مجموعه داده CULane نشان می‌دهد. مدل پیشنهادی با کسب F1 برابر ۸۰/۲۰٪ و سرعت ۴۰۷

شاخص F1 است که نشان‌دهنده دقت کلی و توازن میان نرخ تشخیص صحیح و نرخ خطا است.

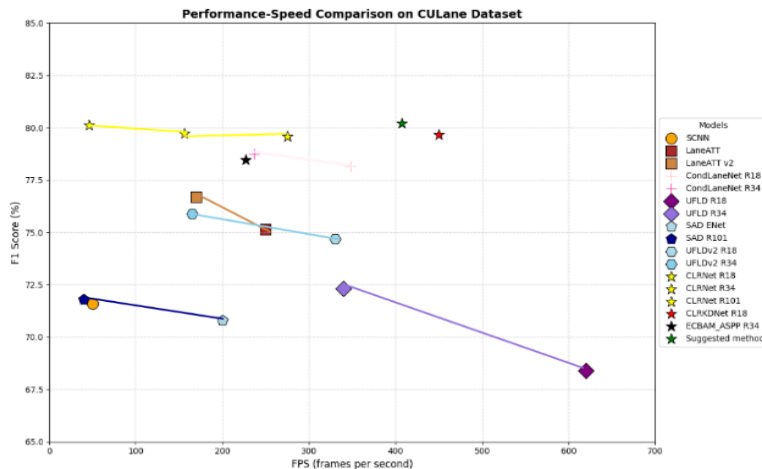
همان‌گونه که مشاهده می‌شود، روش پیشنهادی با مقدار ۸۰/۲۰ درصد بالاترین عملکرد را در میان تمامی روش‌های مقایسه شده به دست آورده است. این مقدار نسبت به مدل‌های قدیمی‌تر مانند SCNN که ۷۱/۶۰ درصد دقت دارند و مدل‌های جدیدتری همچون ECBAM_ASPP با دقت ۷۸/۴۵ درصد و CLRkdNet با دقت ۷۹/۶۶ درصد عملکرد بهتری دارد. این مسئله بیانگر آن است که معماری پیشنهادی توانسته عملکرد مؤثرتری در استخراج ویژگی‌ها و تشخیص خطوط در شرایط متنوع داشته باشد. در تحلیل شرایط محیطی مختلف، روش پیشنهادی در محیط‌های شلوغ با دقت ۷۹/۱۰ درصد و در شرایط وجود سایه با دقت ۸۲/۵۹ درصد بالاترین عملکرد را در میان تمامی مدل‌های مقایسه شده داشته است. این نتایج به‌ویژه در محیط‌های شهری و با ترافیک بالا که حضور وسایل نقلیه متعدد و سایه‌اندازی ناشی از ساختمان‌ها یا درختان می‌تواند چالش بزرگی برای سیستم‌های بینایی ماشین ایجاد کند، اهمیت فراوانی دارد. توانایی روش پیشنهادی در حفظ دقت بالا در این شرایط نشان‌دهنده پایداری معماری و بهره‌گیری صحیح از ویژگی‌های تصویری پایدار است. در مقایسه، بسیاری از روش‌های دیگر در همین شرایط کاهش محسوسی در دقت داشته‌اند و به‌ویژه مدل‌های قدیمی‌تر مانند UFLD یا SCNN نتایج ضعیف‌تری ارائه کرده‌اند. علاوه بر این، در تحلیل عناصر خاص جاده‌ای نیز نتایج به نفع روش پیشنهادی بوده است؛ به‌گونه‌ای که در شناسایی فلش‌ها، این مدل با دقت ۹۰/۵۳ درصد بالاترین عملکرد را در جدول ثبت کرده است. این مسئله از آن جهت اهمیت دارد که تشخیص صحیح علائم هندسی جاده، نقشی اساسی در هدایت خودروهای خودران و جلوگیری از انحراف مسیر دارد. در مجموع، می‌توان نتیجه گرفت که روش پیشنهادی علاوه بر توانایی در مدیریت شرایط دشوار محیطی، در شناسایی عناصر کلیدی جاده نیز برتری آشکاری نسبت به مدل‌های مطرح دارد. با وجود این، باید توجه داشت که عملکرد مدل در سایر شرایط خاص مانند رانندگی شبانه هنوز جای بهبود دارد و می‌تواند در آینده با به‌کارگیری داده‌های متنوع‌تر و

مناسب می‌سازد. همچنین شکل (۵) نتایج مقایسه‌ای مدل‌های تشخیص خط جاده را بر اساس معیار FPS و F1 در دیتاست CU-Lane به تصویر می‌کشد. در این میان، مدل پیشنهادی با دستیابی به ۴۰۷ FPS و امتیاز F1 معادل ۸۰/۲۰ درصد، تعادل مطلوبی بین سرعت و دقت ارائه می‌دهد. و در مقایسه، مدل UFLD با ۳۴۰ FPS و F1 معادل ۷۲/۳۰ را کسب کرده که در دقت از مدل پیشنهادی ضعیف تر است. مدل SCNN با ۵۰ FPS و امتیاز F1 معادل ۷۱/۶۰ درصد را ارائه می‌دهد و تمرکز بیشتری بر دقت در سرعت پایین دارد. مدل‌هایی مانند CondLaneNet با ۲۳۷ FPS و امتیاز F1 معادل ۷۸/۷۴ با CLRLNet با ۱۰۳/۱۵۶ FPS و امتیاز F1 معادل ۷۹/۷۳ درصد تعادل خوبی در سرعت و دقت نشان می‌دهند. در مجموع، مدل پیشنهادی با ترکیب سرعت و دقت قابل توجه، عملکرد مطلوبی در این مقایسه ارائه می‌دهد.

فریم بر ثانیه، عملکرد مطلوبی در هر دو معیار دقت و سرعت نسبت به سایر روش‌های مبتنی بر ResNet-34 داشته است. جدول (۵): مقایسه روش‌ها با شبکه پس‌زمینه مشترک (ResNet-34)

Method	Backbone	F1 (%)	FPS
LaneATT[۱۸]	ResNet-34	۷۶/۶۸	۱۷۰
CondLaneNet[۲۶]	ResNet-34	۷۸/۷۴	۱۳۷
UFLD[۲۳]	ResNet-34	۷۲/۳۰	۳۴۰
UFLDv2[۳۰]	ResNet-34	۷۵/۹۰	۱۶۵
CLRLNet [۲۸]	ResNet-34	۷۹/۷۳	۱۰۳/۱۵۶
ECBAM_ASPP[۱۷]	ResNet-34	۷۸/۴۵	۲۲۷/۳
Proposed (ours)	ResNet-34	۸۰/۲۰	۴۰۷

این نتایج بیانگر تعادل مناسب مدل پیشنهادی در حفظ دقت همراه با سرعت پردازش بالا است که آن را برای کاربردهای بلدانگ



شکل (۵): FPS در مقابل امتیاز F1 روش‌های پیشرفته در معیارهای CULane

عملکردی قابل مقایسه با پیشرفته‌ترین روش‌های موجود دارد، به‌گونه‌ای که مقدار F1 برای هر دو شبکه پس‌زمینه ResNet18 و ResNet34 برابر با ۹۷/۶۶ درصد و دقت آن‌ها به ترتیب ۹۶/۸۴ و ۹۶/۹۶ درصد گزارش شده است. این سطح از دقت عملکرد مدل پیشنهادی را در کنار روش‌هایی مانند CLRLNet و CLRLNet قرار می‌دهد. در ارزیابی خطاهای مثبت و منفی کاذب نیز مدل پیشنهادی عملکردی قابل توجه ارائه کرده است. مقدار FN شبکه ResNet34 برابر با ۱/۵۷ درصد بوده و کمترین

در جدول (۶)، عملکرد مدل پیشنهادی بر روی مجموعه داده TuSimple در کنار مجموعه‌ای از روش‌های مطرح تشخیص خطوط جاده مورد مقایسه قرار گرفته است. باتوجه به ماهیت این مجموعه داده که شامل تصاویر بزرگراهی با خطوط منظم و ساختارمند است، شاخص‌هایی همچون دقت، F1، FP و FN معیارهای دقیقی برای سنجش توان مدل‌ها در تشخیص صحیح خطوط جاده و جلوگیری از خطاهای حیاتی به شمار می‌آیند. نتایج نشان می‌دهد که مدل پیشنهادی از نظر شاخص‌های F1 و دقت

مقدار در میان تمامی روش‌های مقایسه شده محسوب می‌شود که نشان‌دهنده‌ی توان بالای مدل در تشخیص خطوط واقعی و جلوگیری از حذف خطوط حیاتی است. همچنین مقدار FP نیز هرچند اندکی بیشتر از روش‌هایی مانند CLRNet و

اما همچنان کمتر از مدل‌هایی همچون UFLD و UFLDv2 بوده و نشان می‌دهد که مدل پیشنهادی در جلوگیری از تشخیص خطوط نادرست عملکرد رضایت‌بخشی دارد.

جدول (۶): نتایج مقایسه‌ای مدل پیشنهادی در مجموعه داده TUSimple

RT (ms)	FN (%)	FP (%)	Acc (%)	F1 (%)	Backbone	Method
۲۰	۱/۸۰	۶/۱۷	۹۶/۵۳	۹۵/۹۷	VGG16	SCNN
۱/۷	۳/۹۲	۱۹/۰۵	۹۵/۸۲	۸۷/۸۷	ResNet18	UFLD
۳/۰	۳/۷۵	۱۸/۹۱	۹۵/۸۶	۸۸/۰۲	ResNet34	UFLD
۴/۰	۳/۰۱	۳/۵۶	۹۵/۵۷	۹۶/۷۱	ResNet18	LaneATT
۵/۹	۲/۹۲	۳/۵۳	۹۵/۶۳	۹۶/۷۷	ResNet34	LaneATT
۲/۹	۳/۸۰	۲/۱۸	۹۵/۴۸	۹۷/۰۱	ResNet18	CondLaneNet
۴/۲	۳/۸۲	۲/۲۰	۹۵/۳۷	۹۶/۹۸	ResNet34	CondLaneNet
۳/۲	۴/۶۱	۳/۰۶	۹۵/۶۵	۹۶/۱۶	ResNet18	UFLDv2
۵/۹	۴/۳۷	۳/۱۸	۹۵/۵۶	۹۶/۲۲	ResNet34	UFLDv2
۳/۷	۱/۹۲	۲/۲۸	۹۶/۸۴	۹۷/۸۹	ResNet18	CLRNet
۴/۹	۲/۰۸	۲/۲۷	۹۶/۸۷	۹۷/۸۲	ResNet34	CLRNet
۲/۴	۱/۹۲	۲/۴۳	۹۶/۸۰	۹۷/۷۸	ResNet18	CLRKNet
—	۴/۷	۱۹/۷	۹۵/۷۸	۸۷/۶۷	ResNet34	ECBAM_ASPP
۴/۲	۱/۷۰	۲/۹۳	۹۶/۸۴	۹۷/۶۶	ResNet18	مدل پیشنهادی
۳/۸	۱/۵۷	۳/۰۶	۹۶/۹۶	۹۷/۶۶	ResNet34	مدل پیشنهادی

از منظر سرعت پردازش، مدل پیشنهادی در مجموعه داده TuSimple نیز زمان پاسخ مطلوبی ارائه می‌دهد، به‌گونه‌ای که زمان اجرا برای ResNet18 و ResNet34 به ترتیب ۴/۲ و ۳/۸ میلی‌ثانیه گزارش شده است. این نتایج نشان می‌دهد که روش پیشنهادی با وجود بهره‌گیری از ماژول‌های تقویت ویژگی همچون CBAM، افزایش محسوسی در سربار محاسباتی ایجاد نکرده و همچنان سرعت لازم برای کاربردهای بلادرنگ را حفظ کرده است.

در مقایسه با روش‌های پیشرفته، مدل پیشنهادی در شاخص FN عملکردی بهتر از CLRNet و CLRKNet دارد. همچنین مقایسه با CondLaneNet نشان می‌دهد که مدل پیشنهادی رفتار ایمن‌تری در تشخیص خطوط دارد، زیرا هرچند CondLaneNet FP کمتری دارد، اما بالاتر آن می‌تواند به ازدست‌رفتن خطوط واقعی بینجامد. در مقایسه با روش‌هایی مانند UFLD و UFLDv2 نیز مدل پیشنهادی با وجود سرعت کمتر، به دلیل دقت

۵- نتیجه‌گیری

تشخیص دقیق خطوط جاده یکی از اجزای بنیادین در سامانه‌های رانندگی خودکار و سیستم‌های کمک‌راننده پیشرفته است، زیرا به طور مستقیم بر حفظ مسیر، پایداری حرکت خودرو و ایمنی تردد اثر می‌گذارد. با وجود پیشرفت‌های قابل توجه در روش‌های مبتنی بر یادگیری عمیق، دستیابی هم‌زمان به دقت بالا، پایداری در شرایط محیطی دشوار و کارایی بلادرنگ همچنان از چالش‌های مهم این حوزه به شمار می‌رود. در این پژوهش، یک چارچوب کارآمد برای



ترکیب راهبرد تقطیر دانش با مکانیزم توجه، موجب بهبود کیفیت بازنمایی ویژگی‌ها، افزایش تعمیم‌پذیری مدل و ایجاد تعادل مناسب میان دقت و سرعت شده است. این نتایج نشان می‌دهد که روش پیشنهادی ظرفیت مناسبی برای به‌کارگیری در سامانه‌های خودروهای هوشمند دارد. با این حال، علی‌رغم عملکرد مطلوب روش پیشنهادی در اغلب سناریوهای استاندارد و چالش‌برانگیز، نتایج نشان داد که در شرایط نوری شب و صحنه‌های دارای نور خیره‌کننده، عملکرد مدل نسبت به CLRKNet افت جزئی دارد. این مسئله عمدتاً به کاهش بخشی از اطلاعات چند مقیاسی در فرایند فشرده‌سازی ویژگی‌ها در شبکه میانی مربوط می‌شود که می‌تواند تشخیص خطوط را در شرایط با وضوح کم، دشوارتر سازد. استفاده از مجموعه داده‌های شبرنگ و روش‌های افزایش داده هدفمند برای شبیه‌سازی شرایط نوری ضعیف می‌تواند پایداری مدل را در محیط‌های کم‌نور بهبود دهد. همچنین بهینه‌سازی ساختار CBAM متناسب با ویژگی‌های مکانی خطوط در تصاویر شب، بهره‌گیری از مکانیزم‌های توجه سبک‌تر با قابلیت حفظ هم‌زمان جزئیات محلی و زمینه سراسری، و ارزیابی معماری پیشنهادی بر روی سخت‌افزارهای نهفته نظیر پلتفرم‌های NVIDIA Jetson می‌تواند گامی مهم در جهت افزایش کارایی عملی مدل در خودروهای واقعی باشد. نتایج این پژوهش نشان می‌دهد که ادغام تقطیر دانش با سازوکار توجه، چارچوبی مؤثر برای طراحی مدل‌های دقیق کاربردهای بلادرنگ تشخیص خطوط جاده ارائه می‌دهد.

References

- [1] E. Szumska, «Electric vehicle charging infrastructure along highways in the EU», *Energies*, vol. 16, no. 2, pp. 895, Jan. 2023.
- [2] World Health Organization, *Compendium of WHO and Other UN Guidance in Health and Environment, 2024 Update*. Geneva, Switzerland: World Health Organization, Jul. 2024.
- [3] X. Ma, W. Ouyang, A. Simonelli, and E. Ricci, «3D object detection from images for autonomous driving: A survey», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 5, pp. 3537–3556, 2023.

تشخیص خطوط جاده ارائه شد که بر پایه تقطیر دانش و ادغام ماژول توجه کانالی-مکانی CBAM طراحی شده است. لازم به ذکر است که تقطیر دانش، با هدف اصلی کوچک‌سازی و بهینه‌سازی مدل نهایی (Student) از طریق تقلید رفتار یک مدل بزرگ‌تر (Teacher) صورت می‌گیرد؛ این امر برای کاربردهایی که نیازمند کارایی بلادرنگ و اجرای روی سخت‌افزار محدود هستند، ایده‌آل است. این رویکرد با انتقال دانش (TL) که معمولاً برای تسریع یادگیری در یک کار جدید با استفاده از دانش منتقل شده از یک وظیفه پیش‌آموزش‌دیده متفاوت به کار می‌رود، متمایز است. هدف اصلی این پژوهش، دستیابی به یک مدل دانش‌آموز سبک و دقیق برای تشخیص خطوط جاده بود، نه انتقال دانش به یک وظیفه کاملاً جدید. در فرایند آموزش، علاوه بر خروجی نهایی مدل معلم، برخی بازنمایی‌های درونی شبکه نیز برای هدایت یادگیری مدل دانش‌آموز مورد استفاده قرار گرفتند تا انتقال اطلاعات ساختاری مؤثرتر انجام شود. از سوی دیگر، به‌کارگیری ماژول توجه در لایه‌های میانی شبکه به برجسته‌سازی ویژگی‌های مرتبط با خطوط جاده و تمرکز بهتر مدل بر نواحی مهم تصویر کمک کرده است. این طراحی امکان استخراج ویژگی‌های معنادار را فراهم ساخته و در عین حال ساختار سبک مدل را برای پردازش بلادرنگ حفظ کرده است. نتایج ارزیابی بر روی مجموعه داده‌های CULane و TuSimple نشان داد که روش پیشنهادی عملکردی پایدار و رقابتی ارائه می‌دهد. دستیابی به امتیاز FI برابر با ۸۰/۲۰٪ در CULane و دقت ۹۶/۹۶٪ در TuSimple بیانگر آن است که

- [4] A. Eskandarian, ed, *Handbook of Intelligent Vehicles*, vol. 2, pp. 165–232. London, U.K.: Springer, 2012.
- [5] A. Gurghian, T. Koduri, S. V. Bailur, K. J. Carey, and V. N. Murali, «DeepLanes: End-to-end lane position estimation using deep neural networks», in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 38–45, 2016.
- [6] B. He, R. Ai, Y. Yan, and X. Lang, «Accurate and robust lane detection based on dual-view convolutional neural network», in *2016 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1041–1046, Jun. 2016.
- [7] D. Liang, Y. C. Guo, S. K. Zhang, T. J. Mu, and X. Huang, «Lane detection: A survey with new



- results," J. Comput. Sci. Technol, vol. 35, no. 3, pp. 493-505, May 2020.
- [8] C. Liu, X. Li, Q. Liu, F. Yang, Z. Li, and M. Li, "A review of vision-based road detection technology for unmanned vehicles," in 2023 IEEE Intelligent Vehicles Symposium (IV), pp. 1-8, Jun. 2023.
- [9] F. Ma, X. Yan, G. Zhao, X. Xu, Y. Liu, J. Ma, and M. Liu, "Every dataset counts: Scaling up monocular 3D object detection with joint datasets training," in 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 11574-11580, Oct. 2024
- [10] F. Ma, S. Wang, and M. Liu, "An automatic multi-lidar extrinsic calibration algorithm using corner planes," in 2022 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 235-240, Dec. 2022.
- [11] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in Proceedings of the European Conference on Computer Vision (ECCV), pp. 3-19, 2018.
- [12] X. Pan, "CULane: A large-scale dataset for lane detection," GitHub Pages. Available: <https://xingangpan.github.io/projects/CULane.html>
- [13] TuSimple, "TuSimple benchmark," GitHub, 2017. Available: <https://github.com/TuSimple/tusimple-benchmark>
- [14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, 2016.
- [15] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," arXiv preprint arXiv:1503.02531, Mar. 2015
- [16] W. Qi, G. Zhao, F. Ma, L. Zheng, J. Ma, and M. Liu, "CLRKDNet: Speeding up lane detection with knowledge distillation," in 2024 IEEE 27th International Conference on Intelligent Transportation Systems (ITSC), pp. 679-686, Sep. 2024.
- [17] X. Gu, Q. Huang, and C. Du, "Lane Detection Based on ECBAM_ASPP Model," Sensors, vol. 24, no. 24, p. 8098, Dec. 2024
- [18] L. Tabelini, R. Berriel, T. M. Paixao, C. Badue, A. F. De Souza, and T. Oliveira-Santos, "Keep your eyes on the lane: Real-time attention-guided lane detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, , pp. 294-302, 2021
- [19] G. Zhao, F. Ma, W. Qi, Y. Liu, M. Liu, and J. Ma, "CurbNet: Curb detection framework based on LiDAR point cloud segmentation," IEEE Trans. Intell. Transp. Syst, pp. , Feb. 2025
- [20] C. Chen, L. Liu, S. Wan, X. Hui, and Q. Pei, "Data dissemination for industry 4.0 applications in internet of vehicles based on short-term traffic prediction," ACM Trans. Internet Technol, vol. 22, no. 1, pp. 1-8, Oct. 2021
- [21] X. Pan, J. Shi, P. Luo, X. Wang, and X. Tang, "Spatial as deep: Spatial cnn for trafficscene understanding," in Proc. AAAI Conf. Artif. Intell, vol. 32, no. 1, pp. , 2018.
- [22] T. Zheng et al., "Resa: Recurrent feature-shift aggregator for lane detection," in Proc. AAAI Conf. Artif. Intell, vol. 35, no. 4, pp. 3547-3554, 2021
- [23] Z. Qin, H. Wang, and X. Li, "Ultra fast structure-aware deep lane detection," in Computer Vision-ECCV 2020: 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part XXIV 16 , pp. 276-291, Aug. 2020.
- [24] X. Li, J. Li, X. Hu, and J. Yang, "Line-cnn: End-to-end traffic line detection with line proposal unit," IEEE Trans. Intell. Transp. Syst, vol. 21, no. 1, pp. 248-258, Jan. 2019
- [25] D. Jin, W. Park, S.-G. Jeong, H. Kwon, and C.-S. Kim, "Eigenlanes: Data-driven lane descriptors for structurally diverse lanes," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit, pp. 17163-17171, 2022
- [26] L. Liu, X. Chen, S. Zhu, and P. Tan, "Condlanenet: a top-to-down lane detection framework based on conditional convolution," Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3773-3782, 2021.
- [27] S. Yoo et al., "End-to-end lane marker detection via row-wise classification," Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops, pp. 1006-1007, Jun. 2020.
- [28] T. Zheng et al., "CLRNet: Cross layer refinement network for lane detection," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit, pp. 898-907, 2022
- [29] Z. Ge, "Yolox: Exceeding yolo series in 2021," arXiv preprint arXiv:2107.08430, 2021.
- [30] Z. Qin, P. Zhang, and X. Li, "Ultra fast deep lane detection with hybrid anchor driven



- ordinal classification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 5, pp. 2555-2568, Jun. 2022.
- [31] J. Han, X. Deng, X. Cai, Z. Yang, H. Xu, C. Xu, and X. Liang, "Laneformer: Object-aware row-column transformers for lane detection," in *Proc. AAAI Conf. Artif. Intell.*, vol. 36, no. 1, pp. 799-807, Jun. 2022.
- [32] Z. Qu, H. Jin, Y. Zhou, Z. Yang, and W. Zhang, "Focus on local: Detecting lane marker from bottom up via key point," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, pp. 14122-14130, 2021.
- [33] J. Wang et al, "A keypoint-based global association network for lane detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, pp. 1392-1401, 2022.
- [34] R. Liu, Z. Yuan, T. Liu, and Z. Xiong, "End-to-end lane shape prediction with transformers," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, pp. 3694-3702, 2021.
- [35] Y. Hou, Z. Ma, C. Liu, and C. C. Loy, "Learning lightweight lane detection CNNs by self attention distillation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, pp. 1013-1021, 2019.

Road Lane Detection Using a Lightweight Architecture Based on the CBAM Attention Module

Rafat Moaiery Far¹, Masoumeh Rezaei^{2*}, Nik-Mohammad Balouchzahi³

¹ Master's student, Department of Computer Engineering, University of Sistan and Baluchestan, Zahedan, Iran

² Assistant Professor, Department of Computer Engineering, University of Sistan and Baluchestan, Zahedan, Iran

³ Assistant Professor, Department of Information Technology, University of Sistan and Baluchestan, Zahedan, Iran

Article Information

Original Research Paper

Received:

2025 December 23

Accepted:

2026 April 03

Keywords:

Road Lane Detection, Advanced Driver-Assistance Systems, Autonomous Driving, Convolutional Block Attention Module, Knowledge Distillation, Transportation Safety

Corresponding Author*:

mrezaei@ece.usb.ac.ir

Abstract

Lane detection is a critical perception task in intelligent vehicle systems, playing a fundamental role in ensuring safe navigation for autonomous driving and advanced driver-assistance systems (ADAS). However, achieving high detection accuracy under challenging conditions—such as shadows, traffic congestion, and faded lane markings—while maintaining real-time performance remains a significant challenge. In this study, a lightweight architecture is proposed to enhance lane detection performance in complex environments through the integration of innovative strategies. The core of the proposed framework lies in the targeted incorporation of the Convolutional Block Attention Module (CBAM) into the intermediate layers of ResNet, enabling effective refinement of discriminative feature representations. Furthermore, by leveraging an advanced knowledge distillation strategy, the model achieves both high detection accuracy and real-time computational efficiency. Experimental evaluation on the CULane dataset demonstrates that the proposed method achieves an F1-score of 80.20% with a processing speed of 407 frames per second, representing improvements of 0.54% in accuracy and 1.79% in speed compared to CLRKNet and ECBAM_ASPP, respectively. Furthermore, on the TuSimple dataset, the proposed model attains an accuracy of 96.96% while achieving the lowest false-negative rate of 1.57% among the compared methods. Compared with attention-based approaches such as ECBAM_ASPP and high-speed architectures such as UFLD, the proposed method achieves a superior balance between detection accuracy and computational efficiency, making it more suitable for real-time deployment in autonomous vehicles.

 : 10.22034/ABMIR.2026.24147.1209

E-ISSN: [2821-2037](https://doi.org/10.22034/ABMIR.2026.24147.1209)

/The Author 2026. Published by Yazd University This is an open access article under the CC BY 4.0 License (<https://creativecommons.org/licenses/by/4.0/>).

