

# شبکه سریع و کم‌وزن برای شناسایی خطوط جاده با استفاده از معماری موبایل‌نت و توابع هزینه متفاوت

پژمان گودرزی، میلاد حیدری و مهدی حسین‌پور

کمک راننده، سیستم هشدار خروج خط (LDW) یا سیستم تشخیص خط<sup>۱</sup> است. این سیستم با اعلام هشدار در مواقعی که خودرو از مسیر خود منحرف می‌شود و احتمال تصادف وجود دارد، از تصادف و صدمات آن جلوگیری می‌کند. همچنین LDW به عنوان سیستم پایه بسیاری از سیستم‌های دیگر عمل می‌نماید (سیستم‌های جلوگیری از تصادف، مسیریابی و ...). موارد گفته‌شده پیشین همچون افزایش ایمنی مسافران و همچنین ضرورت تشخیص دقیق با هزینه محاسباتی پایین به عنوان انگیزه‌های اصلی ارائه کار پیش رو هستند.

به‌طور کلی روش‌های مربوط به تشخیص خط را می‌توان در دو دسته طبقه‌بندی کرد؛ دسته اول روش‌های مبتنی بر استخراج ویژگی‌ها به صورت دستی و دسته دوم روش‌های مبتنی بر شبکه‌های عصبی عمیق.

در دسته اول بیشتر از ویژگی‌های رنگ [۱] تا [۳] (به دلیل تمایز بالای رنگ و کنتراست خطوط و جاده) و لبه‌ها [۴] و [۵] به عنوان ویژگی‌های جداکننده استفاده می‌شود. با توجه به مشکل عدم جدایی مؤلفه‌های رنگ و روشنایی در فضای رنگی RGB در برخی کارها از فضاهای رنگی دیگر مثل HSI [۶] و YCbCr [۲] استفاده شده است. استفاده از الگوریتم‌های لبه‌یابی Canny [۴] و Sobel [۵] و یا الگوریتم‌های مدل‌کردن خطوط مانند Hough Transform [۶] برای تخمین پارامترهای مدل خطی از دیگر روش‌های دسته اول در حوزه شناسایی خطوط جاده می‌باشد. از مشکلات اصلی این روش در مواجهه آنها با جاده‌های دارای انحنا و پیچ است. برای حل این مشکل از مدل‌های غیرخطی همچون parabolic [۷] و cubic B-spline [۸] استفاده گردیده است. نمونه دیگری از روش‌های این دسته، استفاده و تغییر پرسپکتیو فریم ورودی است. استفاده از الگوریتم IPM [۹] (با نام چشم پرنده هم شناخته می‌شود) و دوربین استریو [۱۰] نمونه‌ای از این روش‌هاست.

روش‌های سنتی در تشخیص خطوط جاده در برابر شرایط مختلف از جمله تغییرات روشنایی، بازتاب نور خورشید از روی سطح جاده، ترک‌های جاده، نور کم، بازتاب نور خودروهای عبوری در شب، بارندگی، مه و ... دارای ضعف و افت کارایی هستند [۱۱].

در سال‌های اخیر، استفاده از شبکه‌های عصبی عمیق تبدیل به روش غالب در اکثر کاربردها گردیده است [۱۲] تا [۲۰]. برخلاف روش‌های دسته اول که محقق را ملزم به انتخاب و تشکیل مجموعه‌ای از ویژگی‌های متمایز به صورت دستی می‌کند، در شبکه‌های عصبی عمیق، استخراج ویژگی به صورت خودکار و با استفاده از یک مجموعه داده آموزشی انجام می‌شود.

مرجع [۲۱] اولین نمونه استفاده از شبکه عصبی عمیق برای تشخیص

چکیده: با استفاده از سیستم تشخیص خط در خودروهای خودران می‌توان موقعیت نسبی خودرو را نسبت به دیگر خودروها و همچنین احتمال خروج از خط و حتی امکان تصادف را بررسی کرد. در این مقاله، یک رویکرد تشخیص خط کم‌وزن و سریع برای تصاویر برگرفته‌شده از دوربین تعبیه‌گردیده در شیشه جلویی خودروها ارائه شده است. بیشتر روش‌های موجود، مسئله تشخیص خط را به صورت کلاس‌بندی در سطح پیکسل در نظر می‌گیرند. این روش‌ها با وجود داشتن قدرت تشخیص بالا، از دو ضعف داشتن پیچیدگی محاسباتی بالا و عدم توجه به اطلاعات محتوایی کلی تصویر منحصربه‌فرد خطوط (در نتیجه در صورت وجود مانع، امکان تشخیص ندارند) رنج می‌برند. روش پیشنهادی پیش رو با بهره‌گیری از روش انتخاب بر اساس ردیف، وجود خطوط در هر ردیف را بررسی می‌کند. همچنین استفاده از معماری موبایل‌نت باعث به‌دست‌آمدن نتایج خوب با تعداد پارامترهای یادگیری کمتر شده است. استفاده از سه تابع مختلف به عنوان توابع هزینه با اهداف متفاوت، باعث به‌دست‌آمدن نتایج عالی و در نظر گرفتن اطلاعات محتوایی کلی منحصربه‌فرد خطوط در کنار اطلاعات محلی شده است. آزمایش‌های انجام‌گرفته بر روی مجموعه تصاویر ویدئویی TuSimple نشان از عملکرد مناسب رویکرد پیشنهادی از لحاظ کارایی و مخصوصاً از لحاظ سرعت دارد.

کلیدواژه: تشخیص خط، خودروی خودران، موبایل‌نت، یادگیری عمیق.

## ۱- مقدمه

هر ساله تصادفات باعث ایجاد مشکلات مالی، جانی و زیست‌محیطی در سراسر دنیا می‌شود. عوامل مختلفی همچون عدم توجه به جلو، خروج از خط، خواب آلودگی و ... از دلایل ایجاد تصادفات هستند. خروج از خط غیرعمدی یکی از این موارد است که در کسری از ثانیه باعث ایجاد ریسک تصادف و ازدست‌رفتن جان مردم می‌شود. امروزه برای کاهش و جلوگیری از تصادفات و همچنین برای افزایش ایمنی خودروها از سیستم‌های مختلفی موسوم به سیستم‌های پیشرفته کمک راننده (ADAS) استفاده می‌شود. پتانسیل بالای اقتصادی و ارتباط مستقیم با جان افراد باعث تبدیل این موضوع به یکی از موضوعات داغ هم در حوزه پژوهش و هم در حوزه صنعت شده است. یکی از سیستم‌های پیشرفته

این مقاله در تاریخ ۱۹ شهریور ماه ۱۴۰۱ دریافت و در تاریخ ۱۷ اردیبهشت ماه ۱۴۰۲ بازنگری شد.

پژمان گودرزی (نویسنده مسئول)، پژوهشکده فناوری اطلاعات، پژوهشگاه ارتباطات و فناوری اطلاعات، تهران، ایران، (email: pgoudarzi@itrc.ac.ir).

میلاد حیدری، دانشکده فناوری اطلاعات و مهندسی کامپیوتر، دانشگاه شهید مدنی آذربایجان، تبریز، ایران، (email: m.heydari@azaruniv.ac.ir).

مهدی حسین‌پور، پژوهشکده فناوری اطلاعات، پژوهشگاه ارتباطات و فناوری اطلاعات، تهران، ایران، (email: mhossein@itrc.ac.ir).

تأثیری در بار محاسباتی شبکه نخواهد داشت.

لازم به ذکر است که انتخاب روش مذکور، تنها برای تشخیص و دقت انتخاب نشده است؛ بلکه زمان اجرا و توانایی ایجاد خروجی بیشتر در زمان مشخص نیز مد نظر قرار گرفته و نتایج ارائه شده در بخش آزمایش‌ها به برتری روش مذکور در این بخش، نسبت به کارهای قبلی اشاره دارند.

ادامه مقاله به شرح زیر است: در بخش ۲، معماری موبایل‌نت و تفاوت آن با معماری‌های دیگر بیان می‌گردد. در بخش ۳، روش پیشنهادی با جزئیات توضیح داده می‌شود. بخش ۴ توضیح آزمایش‌ها و نتایج به دست آمده را ارائه و با دیگر روش‌ها مقایسه می‌کند. بخش ۵ نیز مربوط به نتیجه‌گیری است.

## ۲- مفاهیم پایه

در این بخش مفاهیم پایه از جمله معماری شبکه موبایل‌نت و همچنین تفاوت لایه‌های استفاده شده در این معماری با لایه‌های کانولوشنی معمولی شرح داده می‌شود. ابتدا لایه‌های کانولوشنی استاندارد و سپس، کانولوشن عمق-جدپذیر که در معماری موبایل‌نت استفاده گردیده است، به همراه توابع فعال‌سازی شرح داده می‌شود.

### ۲-۱ کانولوشن استاندارد

در کانولوشن استاندارد، عملیات فیلترکردن و ادغام به صورت یک‌جا و تنها در یک مرحله انجام می‌شود. بدین صورت که با لغزاندن (کانوال) فیلترها روی تصویر ورودی به همراه همه کانال‌هایش، عملیات فیلترکردن و سپس با جمع صفحات حاصل، عملیات ادغام انجام می‌شود. در نتیجه، تعداد پارامترهای یک لایه کانولوشنی استاندارد برای تصویر ورودی با اندازه و تعداد کرنل (که تعداد کانال‌های خروجی لایه بعد را می‌سازد) که دارای اندازه باشد، بدون در نظر گرفتن بایاس و عملیات padding در (۱) آورده شده است

$$\text{num\_of\_params} = (W - 2 \times \left\lfloor \frac{K}{2} \right\rfloor) \times (H - 2 \times \left\lfloor \frac{K}{2} \right\rfloor) \times \text{Input\_Channels} \times (K \times K) \times \text{Output\_Channels} \quad (1)$$

در این معادله، نماد  $\lfloor \cdot \rfloor$  نشان‌دهنده عملگر کف در ریاضیات است.

### ۲-۲ کانولوشن عمق-جدپذیر

در موبایل‌نت در عوض کانولوشن استاندارد از نوع خاصی از کانولوشن به نام کانولوشن عمق-جدپذیر<sup>۱</sup> استفاده می‌شود. در این نوع کانولوشن هم دو مرحله فیلترکردن و ادغام وجود دارد؛ اما این دو مرحله در این نوع از هم جدا بوده و کمی متفاوت نسبت به نسخه استاندارد انجام می‌شود. در کانولوشن عمق-جدپذیر، ابتدا کانولوشن عمقی<sup>۲</sup> و سپس کانولوشن نقطه‌ای<sup>۳</sup> اعمال می‌شود.

کانولوشن عمقی معادل همان فیلترکردن در کانولوشن استاندارد است؛ به‌جز اینکه تنها یک کرنل برای هر کانال ورودی استفاده می‌شود. به عبارتی برای هر کانال از یک کرنل برای عملیات کانوالو استفاده می‌شود و در خروجی این مرحله به تعداد کانال ورودی، کانال خروجی خواهیم داشت. در نتیجه تعداد پارامترهای خروجی این مرحله برابر است با

خطوط جاده بود. در این مقاله، ابتدا با استفاده از CNN ویژگی‌های زمین استخراج شده و سپس از الگوریتم RANSAC برای عملیات خوشه‌بندی استفاده گردیده است. معماری شبکه استفاده شده در این مقاله از سه لایه کانولوشنی، دو لایه کاهش ابعاد و نهایتاً از سه لایه تماماً متصل تشکیل شده است.

برای بهبود دقت در بسیاری از کارها تعداد لایه‌ها و پارامترهای یادگیری، افزایش داده شده که این کار باعث افزایش بار محاسباتی شبکه می‌شود. در [۲۰] الگوریتم SAD ارائه گردیده که در این مقاله از نقشه توجه لایه‌های قبلی برای افزایش دقت و قدرت الگوریتم استفاده شده است. به عبارتی، این مقاله بدون داشتن داده برچسب‌دار، بیشتر و فقط با استفاده از اطلاعات داخلی خود شبکه، قدرت شبکه را افزایش داده است. با این کار بدون نیاز به افزایش پارامترهای یادگیری که در نتیجه آن بار محاسباتی شبکه افزایش می‌یابد، دقت شبکه افزایش داده شده است. این روش با اینکه دارای دقت بالایی در شناسایی خطوط جاده است، اما همچنان روشی با بار محاسباتی بالاست.

یکی دیگر از مشکلات تشخیص خط، عدم تشخیص خط در مواقعی است که مانعی مابین خطوط و دوربین وجود داشته باشد و یا اینکه بنا به دلایلی همچون انعکاس نور شدید و غیره، خطوط به خوبی قابل تشخیص نباشند. این مشکل در الگوریتم‌هایی که عملیات تقسیم‌بندی معنایی را در سطح پیکسل انجام می‌دهند بیشتر به چشم می‌خورد. به عبارتی، این الگوریتم‌ها تنها به اطلاعات پیکسلی اهمیت می‌دهند و به اطلاعات کلی تصویر منحصر به فرد خطوط همچون ارتباط خطوط با هم، موقعیت عمودی خطوط و غیره اهمیت نمی‌دهند. برای حل این مشکل، [۱۹] مکانیسم انتقال پیام را مابین سطرها و ستون‌ها در داخل یک لایه پیشنهاد کرده است. با وجود این، روش مذکور دارای بار محاسباتی بالایی است که استفاده از آن را در کاربردهای دنیای واقعی با مشکل مواجه می‌کند.

در این مقاله با الهام گرفتن از معماری الگوریتم [۱۸] و همچنین با استفاده از الگوریتم موبایل‌نت نسخه ۳ [۲۲]، یک روش با بار محاسباتی پایین برای حل مشکلات اشاره شده در بالا ارائه گردیده است. همچنین در این مقاله از توابع هزینه مختلف با کاربری‌های متفاوت استفاده شده است. برای افزایش قدرت الگوریتم از یک شاخه جداگانه کمکی که همانند الگوریتم‌های سطح پیکسلی عمل می‌کند، استفاده گردیده که این شاخه در زمان تست و ارزیابی استفاده نخواهد شد.

به طور خلاصه نوآوری‌های استفاده شده در این مقاله عبارتند از (۱) در این مقاله برای رفع مشکل عدم تشخیص در مواردی که انسداد وجود داشته باشد، برخلاف روش‌های تقسیم‌بندی معنایی در سطح پیکسل که تنها به اطلاعات محلی اهمیت می‌دهند، از اطلاعات کلی تصویر استفاده می‌شود. برای استفاده از این اطلاعات از توابع هزینه خاص که هر کدام بر اساس ویژگی‌های خطوط طراحی شده است، استفاده می‌شود (برای مثال ارتباط ساختاری خطوط). همچنین از لایه تماماً متصل در لایه آخر استفاده گردیده که دامنه استفاده از اطلاعات را به کل تصویر افزایش می‌دهد. (۲) برای کاهش بار محاسباتی شبکه عصبی از معماری موبایل‌نت نسخه ۳ استفاده شده است. این معماری با طراحی خاص لایه‌های کانولوشنی در عین اینکه پارامترهای یادگیری شبکه را کاهش می‌دهد، قدرت و دقت قابل قبولی را ارائه می‌کند. (۳) استفاده از یک شاخه جداگانه برای افزایش پایداری و همچنین افزایش سرعت یادگیری شبکه از دیگر مواردی می‌باشد که در این کار استفاده شده است. این شاخه که همانند الگوریتم‌های سطح پیکسلی عمل می‌کند، در زمان آموزش استفاده می‌شود و در زمان استنتاج و تست شبکه استفاده‌ای ندارد. در نتیجه، هیچ

1. Depth-Wise Separable  
2. Depth-Wise  
3. Point-Wise

### ۳-۲ شبکه موبایلنت

در معماری موبایلنت به جز لایه اول که از یک کانولوشن استاندارد بهره گرفته شده است، در لایه‌های بعدی از کانولوشن عمق-جدپذیر استفاده شده است. استفاده از این نوع کانولوشن باعث کاهش فوق‌العاده در تعداد پارامترهای یادگیری شده و باعث ایجاد یک شبکه کارآمد و سبک در مقابل دیگر معماری‌ها گردیده است. به عبارتی، موبایلنت مقدار محاسبات را از طریق بهینه‌سازی کانولوشن پایین می‌آورد. به عنوان مثال، شبکه محبوب ResNet-۱۸ دارای ۱۱ میلیون پارامتر است؛ در حالی که موبایلنت دارای تنها ۴/۲ میلیون پارامتر است. در این معماری به جای لایه Pooling از لایه کانولوشن با گام ۲ برای کاهش ابعاد استفاده شده است. در این شبکه بعد از هر لایه از یک لایه BatchNorm و تابع فعال‌ساز ReLU استفاده شده که این معماری از لحاظ دقت هم کارآمد است. برای مثال بر روی مجموعه داده ImageNet معماری موبایلنت دقت بالاتری نسبت به معماری GoogleNet با ۶/۸ میلیون پارامتر دارد. برای اطلاعات بیشتر از معماری موبایلنت به [۲۳] مراجعه کنید.

### ۴-۲ شبکه موبایلنت نسخه ۳

در این مقاله برای استخراج ویژگی از شبکه MobileNetV3 Large استفاده شده و در شکل ۱ معماری این شبکه آمده است. شبکه موبایلنت ورژن ۳ همانند نسخه اولیه خود از کانولوشن‌های عمقی و نقطه‌ای استفاده می‌کند. همچنین هر بلوک همانند نسخه ۲ این شبکه از ورودی و خروجی تنگ<sup>۱</sup> که مابین آنها از لایه‌های با ابعاد بالا برای استخراج بهتر ویژگی‌ها استفاده شده و همچنین از اتصالات باقیمانده<sup>۲</sup> نیز مابین ورودی و خروجی استفاده گردیده است. علاوه بر موارد یادشده که در نسخه‌های پیشین این شبکه موجود است، نسخه ۳ از چلاندن<sup>۳</sup> و فشردن<sup>۴</sup> داده‌ها نیز استفاده کرده است (شکل ۱-ب). برای اطلاعات بیشتر به [۲۲] مراجعه کنید.

### ۵-۲ تابع فعال‌ساز غیرخطی h-swish

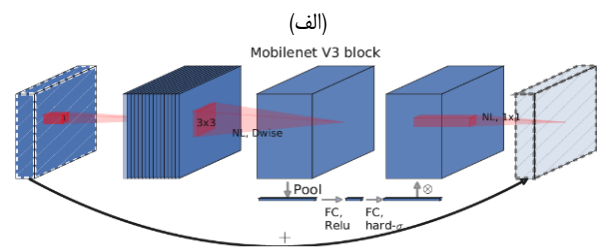
در موبایلنت نسخه ۳ در نیمی از موارد از تابع فعال‌ساز h-swish به جای ReLU استفاده شده است. معادله این فعال‌ساز در (۴) آمده است

$$h-swish[x] = x \frac{ReLU_6(x+3)}{6} \quad (4)$$

### ۳- روش پیشنهادی

در این بخش، روش پیشنهادی به همراه جزئیات به طور کامل تشریح می‌شود. روش پیشنهادی به دنبال رفع دو مشکل کاهش هزینه محاسباتی و امکان پیش‌بینی محل خط جاده حتی در صورت وجود انسداد (به عنوان نمونه وجود خودرو بر روی خط جاده) است. شکل ۲ نمودار بلوکی روش پیشنهادی را نشان می‌دهد. در این دیاگرام و برای نحوه انجام کار از [۱۸] و [۲۰] الهام گرفته شده است؛ اما از لحاظ معماری، لایه‌های استفاده‌شده در معماری و همچنین ورودی و خروجی هر لایه با این منابع متفاوت است. همان طور که ملاحظه می‌کنید بعد از دریافت تصویر، ابتدا عملیات استخراج ویژگی انجام می‌شود. برای استخراج هرچه بهتر ویژگی‌ها از

Input	Operator	exp size	#out	SE	NL	s
$224^2 \times 3$	conv2d	-	16	-	HS	2
$112^2 \times 16$	bneck, 3x3	16	16	-	RE	1
$112^2 \times 16$	bneck, 3x3	64	24	-	RE	2
$56^2 \times 24$	bneck, 3x3	72	24	-	RE	1
$56^2 \times 24$	bneck, 5x5	72	40	✓	RE	2
$28^2 \times 40$	bneck, 5x5	120	40	✓	RE	1
$28^2 \times 40$	bneck, 5x5	120	40	✓	RE	1
$28^2 \times 40$	bneck, 3x3	240	80	-	HS	2
$14^2 \times 80$	bneck, 3x3	200	80	-	HS	1
$14^2 \times 80$	bneck, 3x3	184	80	-	HS	1
$14^2 \times 80$	bneck, 3x3	184	80	-	HS	1
$14^2 \times 80$	bneck, 3x3	480	112	✓	HS	1
$14^2 \times 112$	bneck, 3x3	672	112	✓	HS	1
$14^2 \times 112$	bneck, 5x5	672	160	✓	HS	2
$7^2 \times 160$	bneck, 5x5	960	160	✓	HS	1
$7^2 \times 160$	bneck, 5x5	960	160	✓	HS	1
$7^2 \times 160$	conv2d, 1x1	-	960	-	HS	1
$7^2 \times 960$	pool, 7x7	-	-	-	-	1
$1^2 \times 960$	conv2d 1x1, NBN	-	1280	-	HS	1
$1^2 \times 1280$	conv2d 1x1, NBN	-	k	-	-	1



شکل ۱: شبکه MobileNetV3 Large [۲۲]. (الف) لایه‌ها به کاررفته در معماری و (ب) اجزای داخلی هر لایه بلوکی در این معماری.

$$depth\_wise\_params =$$

$$(W - 2 \times \left\lfloor \frac{K}{2} \right\rfloor) \times (H - 2 \times \left\lfloor \frac{K}{2} \right\rfloor) \times \quad (2)$$

$$Input\_Channels \times (K \times K)$$

کانولوشن نقطه‌ای معادل همان ادغام در کانولوشن استاندارد است. برخلاف کانولوشن استاندارد که در مرحله ادغام عملیات ساده جمع و ساخت یک نقشه ویژگی می‌باشد، در اینجا مرحله ادغام شامل انجام کانولوشن  $1 \times 1$  به تعداد کانال خروجی است. به عبارتی با توجه به اندازه کانولوشن این مرحله که برابر با  $1 \times 1$  است، تعداد پارامترهای کانولوشن نقطه‌ای برابر است با

$$point\_wise\_params =$$

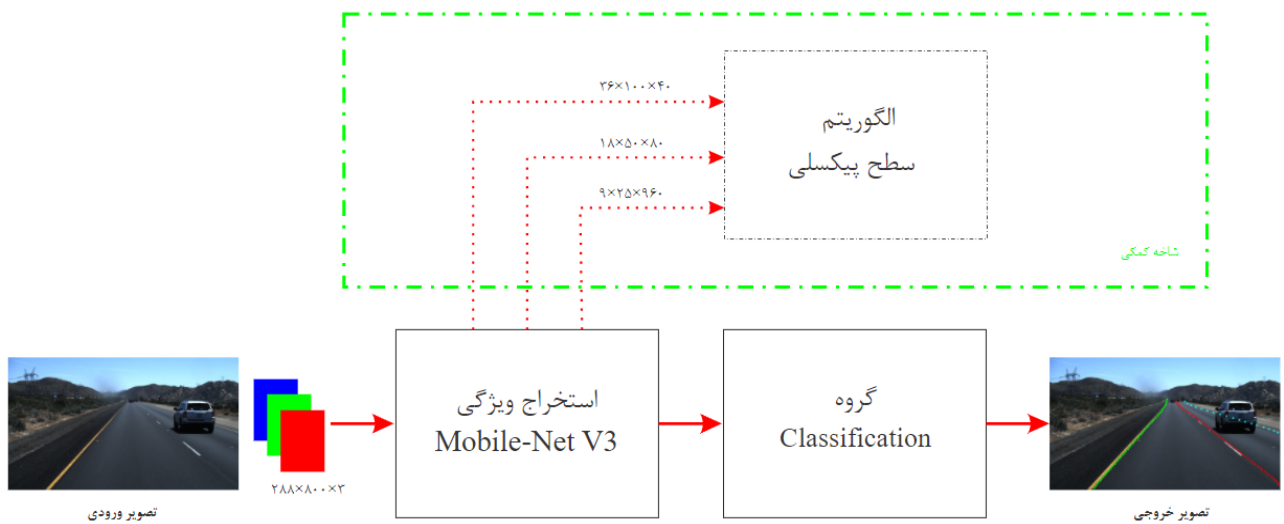
$$(W - 2 \times \left\lfloor \frac{K}{2} \right\rfloor) \times (H - 2 \times \left\lfloor \frac{K}{2} \right\rfloor) \times \quad (3)$$

$$Input\_Channels \times (1 \times 1) \times Output\_Channels$$

در نتیجه تعداد کل پارامترهای یادگیری برای کانولوشن عمق-جدپذیر با مجموع تعداد پارامترهای دو مرحله کانولوشن عمقی و کانولوشن نقطه‌ای برابر است. برای درک بهتر میزان تفاوت پارامترهای یادگیری در کانولوشن استاندارد و کانولوشن عمق-جدپذیر مثال زیر ارائه می‌شود.

فرض کنید که تصویر ورودی دارای عرض، طول و کانال به ترتیب ۱۲، ۱۲ و ۳ باشد و تعداد ۲۵۶ کرنل با سایز  $5 \times 5$  برای عملیات کانال استفاده شود. تعداد کل پارامترها بدون در نظر گرفتن بایاس برای کانولوشن استاندارد برابر با ۱۲۲۸۸۰۰ خواهد بود؛ در حالی که تعداد پارامترها برای ورودی مورد نظر برای کانولوشن عمقی و نقطه‌ای به ترتیب برابر با ۴۸۰۰ و ۴۹۱۵۲ است که مجموع آنها برابر با ۵۳۹۵۲ می‌باشد. به وضوح تفاوت تعداد پارامترها در این دو نوع کانولوشن دیده می‌شود.

1. Bottleneck
2. Residual
3. Squeeze
4. Excite



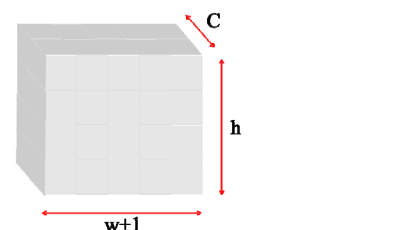
شکل ۲: نمودار بلوکی روش پیشنهادی.

پیکسل، کلاس خروجی تخصیص داده می‌شود. به عبارتی در صورتی که پیکسل مورد نظر جزو خطوط جاده تشخیص داده شود، کلاس خط را دریافت می‌کند. این روش‌ها با وجود قدرت بالا در تشخیص با توجه به اینکه برای هر پیکسل به صورت جداگانه تصمیم می‌گیرد، دارای بار محاسباتی بالاست. به عنوان مثال در صورتی که ابعاد تصویر برابر با و به ترتیب برای عرض و طول تصویر باشد و هدف تشخیص خط جاده باشد، خروجی کلاس‌بندی مورد نظر برابر با خواهد بود. برای تشخیص وجود یا نبود خط در محل پیکسل مورد نظر یک مورد بر ابعاد تعداد خطوط افزوده شده است.

در روش پیشنهادی در عوض پیکسل به دنبال تشخیص خط در هر ردیف هستیم و به عبارت دیگر، خروجی در روش پیشنهادی به صورت شبکه‌بندی خواهد بود. بدین صورت که ارتفاع تصویر را به تعداد مشخصی ردیف تقسیم می‌کنیم و همچنین هر ردیف به تعدادی ستون شبکه‌بندی می‌شود. با توجه به اینکه مقادیر ردیف و ستون‌های مشخص شده، بسیار کمتر از عرض و طول تصویر خواهد بود، در نتیجه شاهد کاهش شدید بار محاسباتی شبکه در روش پیشنهادی هستیم. به عبارتی اگر تعداد ردیف‌ها و ستون‌ها را به ترتیب  $w$  و  $h$  بنامیم و تعداد خطوط قابل تشخیص جاده را  $C$  در نظر بگیریم، خروجی کلاس‌بندی روش پیشنهادی برابر خواهد بود با  $C \times h \times (w+1)$  که البته در این خروجی برای تشخیص وجود و یا عدم وجود خط در هر ردیف، یک بعد اضافه در نظر گرفته شده است (شکل ۳).

در اینجا هیچ عملیات قطعه‌بندی انجام نمی‌شود. همان طور که در دیاگرام شکل ۲ نیز مشخص شده است کل تصویر به شبکه موبایل‌نت داده می‌شود. باید توجه داشت که تنها از قسمت استخراج ویژگی شبکه موبایل‌نت استفاده شده است. سپس خروجی این بخش به کلاس‌بندی با شبکه مشخص شده در جدول ۱ به منظور کلاس‌بندی نهایی داده می‌شود. خروجی به فرم  $C \times h \times (w+1)$  است. به عبارتی هر سلول در خروجی یک ناحیه از تصویر به ابعاد  $W/w$  (عرض ناحیه) در  $H/h$  (طول ناحیه) را نشان می‌دهد ( $H$  و  $W$  به ترتیب نشان‌دهنده عرض و طول کل تصویر و  $h$  و  $w$  نشان‌دهنده تعداد تقسیمات و همچنین خروجی شبکه هستند). سلول در شبکه مورد نظر دارای بعد  $C$  است که تعداد خطوط تشخیص است و هر بعد (یا عمق) نشان می‌دهد که آیا خط مورد نظر موجود است یا نه.

این روش در عین سادگی، باعث کاهش شدید بار محاسباتی می‌شود.



شکل ۳: لایه خروجی روش پیشنهادی (الهام‌گرفته شده از [۱۸]).

قسمت استخراج ویژگی مدل آماده موبایل‌نت نسخه ۳ استفاده شده است. در بخش ۲ و در زیربخش‌های جداگانه، معماری و نحوه کار این شبکه توضیح داده شد و همچنین علت پایین بودن بار محاسباتی این شبکه با توجه به نوع کانولوشن استفاده شده در آن بیان گردید. ویژگی‌های استخراج‌شده از این شبکه به عنوان ورودی به بخش کلاس‌بندی جدا (به منظور تشخیص ویژگی‌های مخصوص خطوط) داده می‌شود. برای استفاده از اطلاعات کلی تصویر منحصربه‌فرد خطوط به جای خروجی پیکسلی (همانند الگوریتم‌های تقسیم‌بندی معنایی) از لایه تماماً متصل استفاده شده است. همچنین همان طور که در بالای شکل ۲ دیده می‌شود (مستطیل سبزرنگ) از سه لایه قسمت استخراج ویژگی برای ایجاد شاخه کمکی استفاده گردیده است. این شاخه تنها در زمان آموزش و به منظور کمک به روند آموزش شبکه استفاده می‌شود که همانند الگوریتم‌های تقسیم‌بندی معنایی عمل کرده و خروجی تابع هزینه آن برای جهت‌دهی به کل شبکه و افزایش سرعت همگرایی شبکه استفاده می‌شود. این شاخه در زمان استنتاج حذف شده و بار محاسباتی برای شبکه ایجاد نمی‌کند. ایده استفاده از این شاخه از روش [۲۰] الهام گرفته شده است.

در ادامه، علت پایین بودن بار محاسباتی روش پیشنهادی و چگونگی رفع عدم شناسایی در صحنه‌های دارای انسداد در بخش‌های جداگانه بیان می‌گردد و همچنین توابع هزینه استفاده شده در روش پیشنهادی شرح داده می‌شود.

### ۳-۱ کاهش بار محاسباتی با تشخیص ردیفی و استفاده از بلوک استخراج ویژگی موبایل‌نت

در روش‌های تشخیص خط مبتنی بر تقسیم‌بندی پیکسلی به هر

نهایی استفاده می‌شود (شکل ۲). استفاده از لایه‌های تماماً متصل امکان جابه‌جایی و ارتباط ویژگی‌ها در کل تصویر را به شبکه می‌دهد. به عبارتی برخلاف روش‌های سطح پیکسلی که در آنها تنها اطلاعات محلی در طول شبکه جریان می‌یابد، در روش پیشنهادی با استفاده از لایه‌های تماماً متصل امکان استفاده از اطلاعات کل تصویر فراهم شده است. در جدول ۱ جزئیات و لایه‌های استفاده‌شده در این لایه‌ها آمده که منظور از Conv-bn-relu یعنی استفاده از لایه کانولوشنی که به دنبال آن از یک لایه نرمال‌سازی<sup>۱</sup> به منظور افزایش پایداری شبکه و از یک تابع فعال‌ساز ReLU استفاده شده است. همچنین InvertedResidual نشان‌دهنده بلوکی از نوع کانولوشن عمق-جداپذیر است که از دو نوع کانولوشن عمقی (dw) و نقطه‌ای (pw) تشکیل شده است. در لایه آخر نیز از دو لایه تماماً متصل که با FC نشان داده شده، استفاده گردیده است.

روش دیگری که به منظور استفاده از اطلاعات ساختاری خطوط و کلی تصویر منحصربه‌فرد خطوط و رفع مشکل عدم تشخیص در مواجهه با انسداد در این کار در نظر گرفته شده است، بهره‌گیری از توابع هزینه مختلف متناسب با موضوع تشخیص خط است. در ادامه، توابع هزینه مورد استفاده در روش پیشنهادی که از [۱۸] الهام گرفته شده است، شرح داده می‌شود.

### ۳-۳ توابع هزینه

توابع هزینه، نقش بسزایی در کنترل روند آموزش شبکه دارند. برای بهره‌گیری از اطلاعات ساختاری موجود در تصویر از توابع هزینه مختلفی استفاده می‌شود. توابع هزینه استفاده‌شده با توجه به ویژگی‌های ساختاری و رفتاری خط طراحی گردیده است. در ادامه به بررسی هر کدام از این توابع پرداخته می‌شود.

#### ۳-۳-۱ تابع هزینه برای عملیات کلاس‌بندی

تابع هزینه استفاده‌شده برای کلاس‌بندی در (۵) آمده و برای این تابع هزینه از الگوریتم آنتروپی متقاطع<sup>۲</sup> استفاده شده است

$$Loss_{cls} = \sum_{i=1}^C \sum_{j=1}^h Cross\_Entropy(P_{i,j,w+1}, T_{i,j,w+1}) \quad (5)$$

در این رابطه،  $P$  مقدار پیش‌بینی‌شده یا خروجی سیستم،  $T$  مقدار هدف یا خروجی واقعی،  $C$  تعداد خطوط پیش‌بینی،  $h$  تعداد ردیف‌های انتخابی و  $w$  تعداد ستون‌های انتخابی است. با توجه به اینکه هدف، تعیین مکان خط در یک ردیف است، یک ستون اضافه برای وجود یا عدم وجود خط در ردیف در نظر گرفته شده است. مقدار خروجی تابع آنتروپی متقاطع مابین صفر و یک و الگوریتم به دنبال کاهش این مقدار است. این تابع، تابع هزینه اصلی روش پیشنهادی محسوب می‌شود.

#### ۳-۳-۲ تابع هزینه متناظر با ساختار خطوط جاده

در محاسبه هزینه نهایی روش پیشنهادی از تابع هزینه دیگری نیز که در آن از ساختار خط الهام گرفته شده است، استفاده می‌شود. خطوط جاده ساختارهایی هستند که در راستای طول کشیده شده‌اند؛ در نتیجه موقعیت خطوط در ردیف‌های کنار هم نزدیک است. بر این اساس تابع هزینه دوم به صورت (۶) فرموله می‌شود

$$Loss_{str} = \sum_{i=1}^C \sum_{j=1}^{h-1} \|P_{i,j,w} - P_{i,j+1,w}\| \quad (6)$$

1. Batch Normalization
2. Cross Entropy

جدول ۱: اطلاعات لایه‌های کاهش بعد و لایه تماماً متصل برای بخش کلاس‌بندی روش پیشنهادی.

نام لایه	کرنل	بُعد (خروجی، ورودی)
Conv-bn-relu	(۱،۱)	(۹۶۰،۲۵۶)
InvertedResidual (dw)	(۳،۳)	(۲۵۶،۲۵۶)
InvertedResidual (pw)	(۱،۱)	(۲۵۶،۲۵۶)
Conv-bn-relu	(۱،۱)	(۲۵۶،۶۴)
Conv-bn-relu	(۵،۵)	(۶۴،۶۴)
Conv-bn-relu	(۱،۱)	(۶۴،۳۲)
Conv-bn-relu	(۳،۳)	(۳۲،۳۲)
InvertedResidual (dw)	(۳،۳)	(۳۲،۳۲)
InvertedResidual (pw)	(۱،۱)	(۳۲،۳۲)
Conv-bn-relu	(۱،۱)	(۳۲،۱۰)
FC		۲۰۴۸
FC		$C \times h \times (w+1)$ خروجی

از طرف دیگر، روش پیشنهادی برای کاهش هرچه بیشتر بار محاسباتی، معماری شبکه موبایل‌نت نسخه ۳ را به خدمت گرفته است. این شبکه با بهره‌گیری از کانولوشن‌های خاص، باعث کاهش بار محاسباتی گردیده و همچنین با طراحی خاص خود، کارایی و دقت را در سطح قابل قبولی نگه داشته است.

در روش پیشنهادی، پس از دریافت تصویر از بلوک‌های استخراج ویژگی، شبکه موبایل‌نت استفاده گردیده تا ویژگی‌های تصویر به خوبی شناسایی شود. خروجی قسمت استخراج ویژگی، تنسوری با طول و عرض  $9 \times 25$  و بعد ۹۶۰ است. موارد گفته‌شده در این بخش، علت پایین بودن بار محاسباتی را به خوبی نشان می‌دهد. در بخش آزمایش‌ها مقایسه زمان مصرفی شبکه با دیگر روش‌ها نشان داده شده و در ادامه، چگونگی برطرف‌شدن مشکل عدم تشخیص در حضور انسداد بررسی می‌شود.

### ۳-۳ برطرف‌کردن مشکل عدم تشخیص در مواجهه

#### با انسداد

الگوریتم‌های رایج تشخیص خط که مبتنی بر تقسیم‌بندی پیکسلی هستند، در مواقعی که خودرو و یا شیئی در مسیر دید پیکسل‌های خط و دوربین قرار می‌گیرد، با توجه به اینکه این الگوریتم‌ها تنها به اطلاعات محلی و پیکسلی توجه دارند، نمی‌توانند موقعیت خط را تشخیص دهند. این اتفاق در مواقعی همچون وجود نور شدید و یا ناپدیدبودن رنگ قسمتی از خط نیز رخ می‌دهد. به عبارتی نیاز به اطلاعات منحصربه‌فرد خط از کل تصویر و همچنین ارتباط پیکسل‌های خط و ارتباط ساختاری بین خطوط مختلف برای تخمین موقعیت خط، احساس می‌شود. باید توجه کرد که برطرف‌نمودن این مشکل به منظور افزایش دقت انجام نمی‌شود؛ بلکه از این روش برای استفاده در کاربردهایی دیگر همچون تشخیص موقعیت خودروها در داخل لاین (به عنوان نمونه در صورتی که نقطه تماس چرخ‌ها داخل خط تشخیصی کناری و خط تقریب زده شده توسط این روش در انسداد باشد، یعنی داخل لاین مربوط است) و یا تشخیص تعویض لاین و ... استفاده می‌شود.

در روش پیشنهادی برای حل این مشکل، دو نوع راه حل استفاده شده است. بعد از استخراج ویژگی توسط موبایل‌نت در روش پیشنهادی، ابتدا از یک بلوک شامل چندین لایه کانولوشنی برای استخراج ویژگی‌های منحصربه‌فرد خط، افزایش قدرت الگوریتم و کاهش بعد استفاده می‌شود. سپس در قسمت آخر شبکه از دو لایه تماماً متصل برای کلاس‌بندی



شکل ۴: نمونه‌هایی از مجموعه داده TuSimple.

استفاده می‌گردد و در زمان استنتاج کنار گذاشته می‌شود. ورودی این شاخه از سه لایه انتخابی از بخش استخراج ویژگی است. لایه‌هایی که در ترتیب بلوک‌بندی شبکه موبایل‌نت در اندیس‌های ۶، ۱۰ و ۱۶ قرار دارند و دارای ابعاد تنسور به ترتیب  $36 \times 100$  و بعد ۴۰،  $18 \times 50$  و بعد ۸۰ و  $9 \times 25$  و بعد ۹۶۰ هستند. ابتدا نقشه‌های ویژگی استخراج شده هم‌اندازه می‌شوند (با عملیات interpolate و به صورت bilinear). همانند جدول ۲، بعد از هر خروجی از چند لایه کانولوشنی استفاده شده و در نهایت خروجی‌ها با هم ترکیب می‌شوند. خروجی این شاخه همانند الگوریتم‌های سطح پیکسلی است.

#### ۴- آزمایش‌ها

در این بخش، کارایی روش پیشنهادی توسط آزمایش‌های مختلف مورد ارزیابی قرار می‌گیرد. در ادامه، ابتدا مجموعه داده ویدئویی که در آزمایش‌ها مورد استفاده قرار گرفته است، معرفی و معیارهای ارزیابی مورد استفاده در آزمایش‌ها تبیین می‌شود. سپس نتایج آزمایش‌ها انجام گرفته و مقایسات با سایر روش‌ها مورد بحث و بررسی قرار می‌گیرد. در نهایت، نمونه‌هایی از خروجی روش پیشنهادی در مواجهه با شرایط مختلف مخصوصاً انسداد نشان داده شده است.

#### ۴-۱ مجموعه داده مورد استفاده در آزمایش‌ها

برای ارزیابی روش پیشنهادی، یکی از مجموعه داده‌های بزرگ و معروف در زمینه تشخیص خطوط جاده یعنی مجموعه داده TuSimple [۲۴] استفاده شده است. در جدول ۳ اطلاعات دقیق این مجموعه داده آمده است.

داده‌های تصویری مورد استفاده در این مجموعه داده از تصاویر داخل بزرگراه‌ها هستند. تصاویر توسط دوربین جلوی خودرو جمع‌آوری شده است. شکل ۴ نمونه‌هایی از تصاویر این مجموعه داده ویدئویی را نشان می‌دهد. باید توجه داشت که چون در این مقاله، کل تصویر به صورت یک‌جا به شبکه داده می‌شود، نیازی به ساخت دیتابیس جداگانه نیست. عملیات ارزیابی دقیقاً همانند [۱۸] انجام شده است.

#### ۴-۲ معیارهای ارزیابی

تعاریف معیارهای ارزیابی مخصوصاً دو معیار اصلی ارزیابی سیستم یعنی دقت<sup>۱</sup> و تعداد فریم قابل محاسبه توسط شبکه در یک ثانیه (FPS) در زیر تعریف می‌شود.

1. Accuracy

جدول ۲: لایه‌های شاخه کمکی.

تعداد	نام لایه	کرنل	بُعد (خروجی، ورودی)
خروجی نقشه ویژگی اول ( $36 \times 100$ و بعد ۴۰)			
۱	Conv-bn-relu	(۱،۱)	(۴۰،۱۲۸)
۳	Conv-bn-relu	(۳،۳)	(۱۲۸،۱۲۸)
خروجی نقشه ویژگی دوم ( $18 \times 50$ و بعد ۸۰)			
interpolate با scale_factor برابر با ۲ و مدل bilinear			
۱	Conv-bn-relu	(۱،۱)	(۸۰،۱۲۸)
۲	Conv-bn-relu	(۳،۳)	(۱۲۸،۱۲۸)
خروجی نقشه ویژگی سوم ( $9 \times 25$ و بعد ۹۶۰)			
interpolate با scale_factor برابر با ۴ و مدل bilinear			
۱	Conv-bn-relu	(۱،۱)	(۹۶۰،۱۲۸)
۱	Conv-bn-relu	(۳،۳)	(۱۲۸،۱۲۸)
عملیات ترکیب خروجی‌ها و ساخت خروجی نهایی			
۱	Conv-bn-relu	(۳،۳)	(۳۸۴،۲۵۶)
۱	Conv-bn-relu	(۳،۳)	(۲۵۶،۱۲۸)
۲	Conv-bn-relu	(۳،۳)	(۱۲۸،۱۲۸)
۱	conv	(۳،۳)	(۱۲۸، (C+1) × H × W)

جدول ۳: اطلاعات دقیق مجموعه داده TuSimple.

تعداد فریم	آموزشی	اعتبارسنجی	داده تست	داده وضوح تصویر	تعداد خطوط
۶۴۰۸	۳۲۶۸	۳۵۸	۲۷۸۲	$1280 \times 720$	کمتر از ۵

که در این رابطه،  $\| \cdot \|_1$  نشان‌دهنده اندازه‌گیری فاصله نرم  $L_1$  است. به عبارتی این معادله، فاصله موقعیت خط در دو ردیف متوالی را در نظر می‌گیرد. مقدار این عبارت باید نزدیک به صفر باشد؛ حتی در مواردی که جاده دارای انحنا باشد، موقعیت خطوط در دو ردیف چسبیده، نزدیک به هم است.

#### ۳-۳-۳ تابع هزینه تقسیم‌بندی پیکسلی شاخه کمکی

در روش پیشنهادی از یک شاخه کمکی استفاده شده است. این شاخه کمکی تنها در زمان آموزش شبکه و به منظور افزایش سرعت همگرایی شبکه و همچنین استفاده از دقت الگوریتم سطح پیکسلی استفاده می‌شود. در زمان استنتاج از این شاخه استفاده نمی‌شود و در نتیجه تأثیری در بار محاسباتی شبکه نخواهد داشت. محاسبه مقدار هزینه در این شاخه کمکی همانند الگوریتم‌های سطح پیکسلی است. در این شاخه نیز از تابع هزینه آنتروپی متقاطع استفاده شده است.

تابع هزینه کلی که از مجموع توابع هزینه معرفی شده به دست آمده، در (۷) معرفی شده است

$$Loss_{all} = Loss_{cls} + \alpha Loss_{str} + \beta Loss_{seg} \quad (7)$$

که  $Loss_{seg}$  نشان‌دهنده تابع هزینه شاخه کمکی و  $\alpha$  و  $\beta$  به عنوان ضرایب تأثیر توابع هزینه می‌باشد. در روش پیشنهادی، مقادیر هر دوی ضرایب، یک در نظر گرفته شده است. در ادامه به بررسی دقیق‌تر شاخه کمکی پرداخته می‌شود.

#### ۴-۳ شاخه کمکی

همان‌طور که قبلاً گفته شد برای افزایش سرعت همگرایی و افزایش پایداری شبکه از یک شاخه مجزا که همانند الگوریتم‌های سطح پیکسلی عمل می‌کند، استفاده شده است. این شاخه تنها در زمان آموزش شبکه

جدول ۴: بررسی نتایج روش پیشنهادی بر روی TUSIMPLE.

دقت (درصد)	Top2 (درصد)	Top3 (درصد)	FPS
۹۵٫۲۰	۹۷٫۶۱	۹۸٫۰۵	۱۱٫۱۲

جدول ۵: بررسی و مقایسه روش‌های مختلف بر روی TUSIMPLE.

روش‌ها	معیار دقت (درصد)	FPS
[۱۷] (Resnet-۱۸)	۹۲٫۶۹	۱٫۲۶
[۱۷] (Resnet-۳۴)	۹۲٫۸۴	۰٫۶۳
[۱۵]	۹۶٫۳۸	۱٫۷۱
[۱۶]	۹۶٫۳۹	۰٫۲۸
[۱۹]	۹۶٫۵۳	۰٫۲۴
[۲۰]	۹۶٫۶۴	۲٫۴۰
[۱۸] (Resnet-۱۸)	۹۵٫۸۷	۹٫۲۱
[۱۸] (Resnet-۳۴)	۹۶٫۰۶	۵٫۵۲
روش پیشنهادی	۹۵٫۲۰	۹٫۷۱

دقت به تعداد تشخیص‌های درست سیستم به تعداد کل فریم‌ها اطلاق می‌شود. به عبارتی، زمانی که ردیف و ستون تشخیص داده شده به عنوان خطوط جاده در خروجی شبکه با مقدار واقعی یکی باشد، به عنوان نمونه تشخیصی درست در نظر گرفته می‌شود. رابطه (۸)، فرمول محاسبه این معیار را نشان می‌دهد

$$accuracy = \frac{\sum_{i=1}^N Correct\_prediction}{N} \quad (8)$$

که در این رابطه،  $N$  تعداد کل نمونه‌ها یا سلول‌های شبکه‌بندی شده و  $Correct\_prediction$  نشان‌دهنده سلول‌های شبکه‌بندی در فریم‌هاست که درست تشخیص داده شده‌اند.

فریم بر ثانیه یا FPS به تعداد فریم‌هایی گفته می‌شود که شبکه، توانایی محاسبه و ایجاد خروجی در یک ثانیه را دارد. با توجه به اهمیت بار محاسباتی الگوریتم‌ها در دنیای واقعی و همچنین امکان اجرای آنها به صورت بلادرنگ، از این معیار ارزیابی مهم استفاده شده است.

معیار دقت در بسیاری از موارد، قدرت الگوریتم را به درستی نشان نمی‌دهد و بسیار سخت‌گیرانه ارزیابی می‌کند. برای ارزیابی دقیق قدرت الگوریتم از معیار  $Topk$  نیز استفاده شده است. معیار  $Topk$  همانند معیار دقت محاسبه می‌شود؛ با این تفاوت که سلول‌هایی که با اختلاف تنها  $k$  ستون با ستون واقعی خطوط، توسط روش پیشنهادی سلول خط انتخاب شوند هم به عنوان سلول درست در نظر گرفته می‌شوند. برای ارزیابی روش پیشنهادی از  $Top2$  و  $Top3$  به همراه معیار دقت استفاده شده است. منظور از  $Top2$  این است در صورتی که سیستم سلولی را به عنوان سلول خط شناسایی کرد و موقعیت این سلول با اختلاف حداکثر دو سلول از سلول خط واقعی قرار داشت، همچنان پاسخ سیستم درست ارزیابی شود. همچنین روال مشابهی برای  $Top3$  است؛ با این تفاوت که حداکثر ۳ سلول از سلول خط واقعی قرار داشته باشد.

### ۳-۴ جزئیات پیاده‌سازی

تصاویر، همانند مقداری که در مجموعه داده تعیین شده، به ۵۶ ردیف تقسیم گردیده است. با توجه به اینکه تنها دانستن لاین اصلی به همراه دو لاین کناری برای بسیاری از موارد کافی است، در روش پیشنهادی تعداد خطوط جاده ۴ و همچنین تعداد ستون‌ها برای شبکه‌بندی هر ردیف ۵۰ در نظر گرفته شده است.

برای کاهش بار محاسباتی همانند [۱۸]، تصاویر به وضوح  $288 \times 80$  تغییر داده شده‌اند. برای آموزش شبکه از بهینه‌ساز Adam [۲۵] و از زمان‌بند CosineAnnealingWarmRestarts [۲۶] برای تنظیم نرخ یادگیری استفاده شده است. مقدار اولیه نرخ یادگیری و تعداد دسته‌ها به ترتیب ۱ و ۱۶ در نظر گرفته شده و برای جلوگیری از آموزش بیش از حد شبکه از روش افزایش داده<sup>۱</sup> استفاده گردیده است. از عملیاتی همچون چرخش و شیف‌ت به بالا، پایین، چپ و راست استفاده شده است.

تمامی مراحل آزمایش و تست بر روی پردازنده Core i5-10400F و با استفاده از کتابخانه pytorch انجام شده است.

#### 1. Data Augmentation

### ۴-۴ نتایج آزمایش

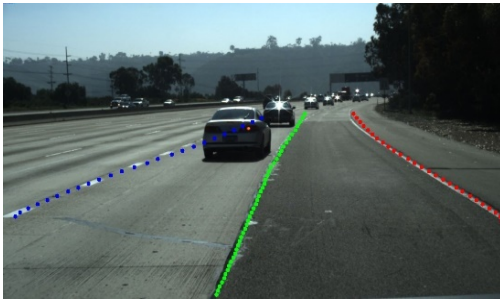
جدول ۴ خروجی روش پیشنهادی را بر روی مجموعه داده تست TUSimple نشان می‌دهد. همان طور که ملاحظه می‌کنید، دقت روش پیشنهادی در سطح خوبی قرار دارد. توجه داشته باشید که؟؟؟ در معیار  $Top2$  به رقم ۹۷٫۶٪ رسیده که رقم بسیار بالایی است و نشان از قدرت بالای این روش دارد. همچنین روش پیشنهادی، قابلیت محاسبه تعداد فریم بالایی را حتی در سیستم مبتنی بر CPU نه‌چندان قدرتمند مورد آزمایش نشان داده است.

### ۴-۵ مقایسه با روش‌های دیگر

در این بخش، روش پیشنهادی با روش‌های مختلف دیگر از لحاظ دقت و زمان اجرا مقایسه می‌شود. برای مقایسه، ۶ منبع مختلف و روش‌های [۱۵] تا [۲۰] برای انجام مقایسات استفاده شده است. برای هر دو روش [۱۸] و [۱۷] پایه Resnet-۱۸ و همچنین پایه Resnet-۳۴ در نظر گرفته شده و در جدول ۵ مقایسه روش‌ها آمده است. با توجه به جدول، روش پیشنهادی توانسته که نتیجه قابل قبولی را از منظر دقت در میان روش‌های مختلف به دست آورد و با اختلاف اندکی نسبت به [۲۰] قرار گیرد. نقطه قوت روش پیشنهادی که آن را در بین روش‌های کاملاً مناسب برای کاربردهای دنیای واقعی کرده است، بار محاسباتی پایین آن است. بار محاسباتی پایین باعث مناسب‌بودن این روش برای کاربردهای بلادرنگ دنیای واقعی شده است. استفاده از موبایل‌نت به همراه معماری خاص ردیفی روش پیشنهادی باعث کاهش بار محاسباتی گردیده است.

### ۴-۶ نتایج کیفی

وجود مانع در مسیر دید دوربین و خطوط جاده، وجود شرایط روشنایی متفاوت، وجود لکه و ناهمواری بر روی جاده و ... از مشکلات شناسایی خطوط جاده هستند. این مشکلات در روش‌های مبتنی بر کلاس‌بندی سطح پیکسلی به دلیل عدم اهمیت به اطلاعات کلی تصویر منحصربه‌فرد خطوط، بیشتر به چشم می‌خورد. با استفاده از اطلاعات کلی تصویر منحصربه‌فرد خطوط و ارتباط ساختاری خطوط می‌توان موقعیت خطوط در تصویر را پیش‌بینی کرد. روش پیشنهادی با به‌کارگیری لایه‌های تماماً متصل در لایه آخر و با استفاده از توابع هزینه مختلف، امکان شناسایی خطوط با وجود مشکلات یادشده مخصوصاً انسداد را فراهم کرده است. شکل ۵ نمونه‌هایی از خروجی روش پیشنهادی را در برخورد با مشکلات مطرح‌شده نشان می‌دهد. وجود انواع جاده با رنگ‌های تیره و روشن، شرایط روشنایی مختلف، وجود مانع بزرگ در خط دید و همچنین انحنای جاده در این تصاویر دیده می‌شوند.



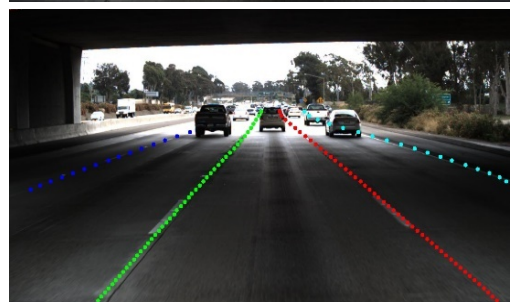
(د)



(الف)



(ه)



(ب)



(ج)

شکل ۵: نمونه‌هایی از خروجی روش پیشنهادی در مواجهه با شرایط مختلف (الف) تا (ه). در هر مورد تصویر بالا، تصویر اصلی و تصویر پایین، خروجی روش پیشنهادی است.

باید توجه داشت که در این مقاله، روش محاسبه خطا همانند [۱۸] و نه به صورت پیکسلی است. به عبارتی از الگوریتم محاسبه خطا و دقت بر اساس ناحیه مستطیلی درست و نه پیکسل درست در نظر گرفته شده است (ناحیه‌ای به بزرگی  $W/w$  برای عرض و  $H/h$  برای طول). در واقع، مستطیلی که دارای بالاترین همپوشانی (بالای ۵۰٪) و همچنین نزدیک‌ترین به سمت چپ تصویر داشته باشد، به عنوان مستطیل خط انتخاب شده است (چون تعداد خطوط مشخص است (در این مقاله تعداد  $C$  برابر چهار، یعنی تشخیص چهار خط هم‌زمان انتخاب شده است) و حتی اگر یک خط شامل چند مستطیل باشد اولین مستطیل از چپ به عنوان مستطیل محل خط انتخاب می‌شود (با توجه به اصول شماره‌گذاری پیکسل‌های تصویر که از سمت چپ شماره‌گذاری می‌شوند). این عملیات هم در مرحله آموزش شبکه و هم در مرحله تست و ارزیابی استفاده شده است. در رابطه با خطوط در مواجهه با انسداد باید در نظر داشته باشیم که الگوریتم مذکور از روی خطوط شناسایی شده در مناطق عدم انسداد، تخمینی از محل قرارگیری خطوط را فراهم می‌آورد. می‌توان از



- Computing, BigComp'18*, pp. 649-652, Shanghai, China, 15-17 Jan. 2018.
- [11] S. P. Narote, P. N. Bhujbal, A. S. Narote, and D. M. Dhane, "A review of recent advances in lane detection and departure warning system," *Pattern Recognition*, vol. 73, pp. 216-234, Jan. 2018.
- [12] J. Tang, S. Li, and P. Liu, "A review of lane detection methods based on deep learning," *Pattern Recognition*, vol. 111, Article ID: 107623, Mar. 2021.
- [13] X. Wu, D. Sahoo, and S. C. Hoi, "Recent advances in deep learning for object detection," *Neurocomputing*, vol. 396, pp. 39-64, 5 Jul. 2020.
- [14] Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu, "Object detection with deep learning: a review," *IEEE Trans. on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3212-3232, Nov. 2019.
- [15] D. Neven, B. D. Brabandere, S. Georgoulis, M. Proesmans, and L. V. Gool, "Towards end-to-end lane detection: an instance segmentation approach," in *Proc. IEEE Intelligent Vehicles Symp. (IV)*, pp. 286-291, Changshu, China, 26-30 Jun. 2018.
- [16] M. Ghafoorian, C. Nugteren, N. Baka, O. Booi, and M. Hofmann, "EL-GAN: embedding loss driven generative adversarial networks for lane detection," in *Proc. of the European Conf. on Computer Vision Workshops*, 14 pp., 2018.
- [17] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834-848, Apr. 2017.
- [18] Z. Qin, H. Wang, and X. Li, "Ultra fast structure-aware deep lane detection," in *Proc. of the Computer Vision-ECCV*, pp. 276-291, 2020.
- [19] X. Pan, J. Shi, P. Luo, X. Wang, and X. Tang, "Spatial as deep: spatial CNN for traffic scene understanding," in *Proc. of the AAAI Conf. on Artificial Intelligence*, pp. 7276-7283, New Orleans, LA, USA, 2-7 Feb. 2018.
- [20] Y. Hou, Z. Ma, C. Liu, and C. C. Loy, "Learning lightweight lane detection CNNs by self attention distillation," in *Proc. of the IEEE/CVF Int. Conf. on Computer Vision*, pp. 1013-1021, Seoul, Korea (South), 27 Oct.-2 Nov. 2019.
- [21] J. Kim and M. Lee, "Robust lane detection based on convolutional neural network and random sample consensus," in *Proc. of the Int. Conf. on Neural Information Processing*, pp. 454-461, Montreal, Canada 8-13 Dec. 2014.
- [22] A. Howard, et al., "Searching for mobilenetv3," in *Proc. of the IEEE/CVF Int. Conf. on Computer Vision*, pp. 1314-1324, Seoul, Korea (South), 27 Oct.-2 Nov. 2019.
- [23] A. G. Howard, et al., "Mobilenets: efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.
- [24] Tusimple. *Lane Detection Benchmark*, 2017. Available: <https://github.com/TuSimple/tusimple-benchmark>
- [25] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, arXiv preprint arXiv:1412.6980, 2014.
- [26] I. Loshchilov and F. Hutter, *SGDR: Stochastic Gradient Descent with Warm Restarts*, arXiv preprint arXiv:1608.03983, 2016.

**پژمان گودرزی** تحصیلات خود را در مقطع کارشناسی در رشته الکترونیک از دانشگاه صنعتی شریف و در مقاطع کارشناسی ارشد و دکترا در مخابرات سیستم و مهندسی برق از دانشگاه صنعتی اصفهان به ترتیب در سال‌های ۱۳۷۴، ۱۳۷۷ و ۱۳۸۳ به پایان رسانده است و هم‌اکنون دانشیار پژوهشکده فناوری اطلاعات پژوهشگاه ارتباطات و فناوری اطلاعات می‌باشد. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: تخصیص منابع در شبکه، بهینه‌سازی، ارسال ویدئو بر روی شبکه، ارزیابی کیفیت خدمات، محاسبات نرم، خودروهای خودران و نظریه بازی.

**میلاhd حیدری** مدرک کارشناسی مهندسی فناوری اطلاعات را از دانشگاه شهید مدنی تبریز دریافت نموده است. نام‌برده در حال حاضر در واحد توسعه نرم‌افزار شرکت موتورسازان تراکتورسازی ایران به فعالیت اشتغال دارد. زمینه‌های تحقیقاتی مورد علاقه نامبرده عبارتند از: هوش مصنوعی، خودروهای خودران و مهندسی نرم‌افزار.

**مهدی حسین‌پور** مدرک کارشناسی ارشد مهندسی مخابرات را از دانشگاه صنعتی امیرکبیر (پلی‌تکنیک) دریافت نموده است. نام‌برده در حال حاضر در آزمایشگاه پژوهشکده فناوری اطلاعات پژوهشگاه ارتباطات و فناوری اطلاعات به فعالیت اشتغال دارد. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: توسعه نرم‌افزار، حکمرانی داده، ارسال ویدئو بر روی شبکه، ارزیابی کیفیت خدمات و خودروهای خودران.

الگوریتم‌های تشخیص محل تماس خودرو (مثلاً با استفاده از نقطه تماس پایین چرخ‌ها) و همچنین تخمین یادشده، محل خودرو را نسبت به لاین‌ها بررسی نمود و فعالیت‌هایی نظیر تعویض لاین را تشخیص داد.

## ۵- نتیجه‌گیری

در این مقاله، روش کم‌وزن تشخیص خطوط جاده که مبتنی بر تشخیص ردیفی است، ارائه شد. روش پیشنهادی با به‌کارگیری معماری موبایل‌نت نسخه ۳ توانست ویژگی‌های مورد نیاز را برای تشخیص خطوط با کمترین بار محاسباتی استخراج کند. برای رفع مشکل عدم تشخیص خطوط در مواجهه با انسداد تصویر از ویژگی‌های کلی تصویر منحصربه‌فرد خطوط و اطلاعات ساختاری خطوط استفاده شد. بدین منظور از دو راهکار استفاده از توابع هزینه مختلف و استفاده از ویژگی‌های کلی تصویر منحصربه‌فرد خطوط بهره گرفته شده و در تمام لایه‌های شبکه به کاهش بار محاسباتی توجه ویژه گردیده است. برای افزایش پایداری شبکه و همگرایی سریع‌تر شبکه از شاخه کمکی استفاده شده است. نتایج، سرعت بالای شبکه و امکان استفاده در کاربردهای بلادرنگ دنیای واقعی را نشان می‌دهند. همچنین این شبکه از دقت کافی نیز برخوردار است. برای پژوهش‌های پیش رو می‌توان به استفاده از مجموعه داده‌های دیگر و بزرگ‌تر اشاره کرد. می‌توان توابع هزینه‌های مختلف مبتنی بر دیگر ویژگی‌های ساختاری خطوط جاده را در نظر گرفت. استفاده از معماری‌های دیگر برای افزایش دقت نیز می‌تواند از دیگر کارهای آینده در این زمینه باشد.

## ۶- سپاسگزاری

مؤلفین بر خود لازم می‌دانند که از حمایت‌های مالی پژوهشگاه ارتباطات و فناوری اطلاعات در حین اجرای فعالیت تشکر کنند.

## مراجع

- [1] S. T. Ying, T. S. Jeng, and V. Chan, "HSI color model based lane-marking detection," in *Proc. IEEE Intelligent Transportation Systems Conf.*, pp. 1168-1172, Toronto, Canada, 17-20 Sept. 2006.
- [2] H. Y. Cheng, B. S. Jeng, P. T. Tseng, and K. C. Fan, "Lane detection with moving vehicles in the traffic scenes," *IEEE Trans. on Intelligent Transportation Systems*, vol. 7, no. 4, pp. 571-582, Dec. 2006.
- [3] M. A. Sotelo, F. J. Rodriguez, L. Magdalena, L. M. Bergasa, and L. Boquete, "A color vision-based lane tracking system for autonomous driving on unmarked roads," *Autonomous Robots*, vol. 16, pp. 95-116, 2004.
- [4] P. C. Wu, C. Y. Chang, and C. H. J. P. R. Lin, "Lane-mark extraction for automobiles under complex conditions," *Pattern Recognition*, vol. 47, no. 8, pp. 2756-2767, Aug. 2014.
- [5] C. Mu and X. Ma, "Lane detection based on object segmentation and piecewise fitting," *Indonesian J. of Electrical Engineering and Computer Science*, vol. 12, no. 5, pp. 3491-3500, May 2014.
- [6] J. Niu, J. Lu, M. Xu, P. Lv, and X. Zhao, "Robust lane detection using two-stage feature extraction with curve fitting," *Pattern Recognition*, vol. 59, pp. 225-233, Nov. 2016.
- [7] J. C. McCall and M. M. Trivedi, "Video-based lane estimation and tracking for driver assistance: survey, system, and evaluation," *IEEE Trans. on Intelligent Transportation Systems*, vol. 7, no. 1, pp. 20-37, Mar. 2006.
- [8] P. Y. Hsiao, C. W. Yeh, S. S. Huang, and L. C. Fu, "A portable vision-based real-time lane departure warning system: day and night," *IEEE Trans. on Vehicular Technology*, vol. 58, no. 4, pp. 2089-2094, May 2008.
- [9] M. Fu, X. Wang, H. Ma, Y. Yang, and M. Wang, "Multi-lanes detection based on panoramic camera," in *Proc. 11th IEEE Int. Conf. on Control & Automation, ICCA'14*, pp. 655-660, Taichung, Taiwan, 18-20 Jun. 2014.
- [10] J. G. Kim, J. H. Yoo, and J. C. Koo, "Road and lane detection using stereo camera," in *Proc. IEEE Int. Conf. on Big Data and Smart*