

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی برق و رباتیک
رساله دکتری مهندسی برق - الکترونیک

آشکارسازی بلادرنگ خطوط جاده و تشخیص موانع روبروی خودرو

نگارنده: علی اصغر فلاح جوشقانی

استاد راهنما

دکتر علی سلیمانی

استاد مشاور

دکتر حسین خسروی

اسفند ۱۴۰۰

تقدیم ہے:

بہ روح پاک پدر مہربانم، کہ چون شقایق داغ نبودش را بر سینہ دارم

قبل از هر چیز خداوند بزرگ را شاکرم که، همواره توجه و عنایتش را شامل حال بنده نموده است.

الکون که به یاری پروردگار و راهبانی اساتید بزرگوارم موفق به پایان این رساله شده‌ام و وظیفه خود دانسته که نهایت سپاسگزاری

را از تمامی عزیزانی که در این راه به من کمک کرده‌اند، به عل آورم:

در آغاز لازم است از پدر مرحوم، مادر و همسر عزیزم به خاطر کمک‌های بی‌دینشان تشکر کنم.

از استاد بزرگ جناب آقای دکتر علی سلیمانی که راهبانی این رساله را بر عهده داشته‌اند کمال تشکر را دارم.

از جناب آقای دکتر حسین خسروی که استاد مشاور این رساله بوده‌اند نیز قدردانی می‌نمایم.

از کلیه بزرگوارانی که از دوران کودکی تاکنون به من چهری آموخته‌اند، سپاسگزارم.

تعمدنامه

اینجانب **علی اصغر فلاح جوشقانی** دانشجوی دوره دکتری رشته **مهندسی برق** دانشکده مهندسی برق و رباتیک دانشگاه صنعتی شاهرود، نویسنده رساله **آشکارسازی بلادرنگ خطوط جاده و تشخیص موانع روبروی خودرو** تحت راهنمایی **دکتر علی سلیمانی** متعهد می‌شوم.

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است .

- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است .

- مطالب مندرج در پایان‌نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .

- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می‌باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « **Shahrood University of Technology** » به چاپ خواهد رسید .

- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان‌نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.

- در کلیه مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است .

- در کلیه مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می‌باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود . استفاده از اطلاعات و نتایج موجود در پایان‌نامه بدون ذکر مرجع مجاز نمی‌باشد.

چکیده

صنعت حمل و نقل یکی از صنایع فعال در جهان است و ایمنی در جاده‌ها از مهم‌ترین حوزه‌های تحقیق و توسعه در این صنعت است. سیستم تشخیص خطوط جاده و سیستم تشخیص موانع از جمله تجهیزات ایمنی هستند که در خودروهای هوشمند به کار می‌روند تا وسیله نقلیه را در بین خطوط جاده هدایت کنند و از انحراف خودرو و بروز تصادفات در برخورد با موانع جلوگیری کنند. تصاویر ویدیویی حاوی اطلاعات بسیار ارزشمندی درباره محیط اطراف خودرو هستند و نقش اساسی در تشخیص خطوط جاده و موانع ایفا می‌کنند. با توجه به اینکه در تشخیص خطوط جاده و موانع در تصاویر ویدیویی با یک محیط طبیعی روبه‌رو هستیم، عوامل متعددی مانند شرایط مختلف جوی، سایه ساختمان‌ها و درختان کنار جاده‌ها، سایه اتومبیل‌ها، وضعیت تابش نور خورشید در ساعات مختلف روز مسئله تشخیص خطوط جاده و موانع را با مشکلات متعددی مواجه می‌کنند. از طرف دیگر با توجه به حجم محاسبات زیاد جهت پردازش اطلاعات تصویر و همچنین نیاز به بلادرنگ بودن سیستم تشخیص خطوط جاده و مانع، پایین آوردن پیچیدگی حاصل از محاسبات پردازش‌های موجود و در نتیجه بالا بردن کارایی سیستم از جهت توان محاسباتی باید در نظر گرفته شود.

در این رساله، هدف تشخیص خطوط جاده و موانع روبه‌روی خودرو با استفاده از تصاویر ویدیویی گرفته شده از جاده و تعیین شرایط رانندگی در جاده است. بدین منظور از دو روش پیشنهادی متفاوت در این پژوهش استفاده نموده‌ایم. ابتدا با استفاده از یک رویکرد کلاسیک مبتنی بر الگوریتم‌های پایه‌ای پردازش تصویر و با به کارگیری لبه‌یابی کنی و همچنین الگوریتم تبدیل هاف، اقدام به شناسایی خطوط جاده با رویکرد بدون ناظر نموده‌ایم. از جمله نوآوری‌هایی که در این روش علاوه بر بلادرنگ بودن که به عنوان هدف اصلی در روش اجرایی صورت گرفته است، انجام یک ماسک‌گذاری پویا برای تشخیص ناحیه مورد علاقه در تصویر است. در روش پیشنهادی دوم، با استفاده از یک رویکرد مبتنی بر یادگیری عمیق و با به کارگیری یک شبکه مولد رقابتی نیمه‌نظارتی اقدام به شناسایی خطوط و در ادامه آن با به کارگیری یک توصیف‌گر هیستوگرام مبتنی بر گرادیان، اقدام به شناسایی موانع روبه‌روی خودرو نموده‌ایم. از جمله نکات حائز اهمیت در روش دوم، استفاده از یک شبکه عمیق مولد رقابتی نیمه‌نظارتی در فرایند قطعه‌بندی و همچنین پیچیدگی محاسباتی پایین‌تر آن نسبت به معماری‌های شبکه‌های عمیق شناخته‌شده نظیر AlexNet، ResNet50 و VGG19 است. همچنین در الگوریتم تشخیص مانع، با به کارگیری سیاست ماسک‌گذاری پویا، به افزایش سرعت اجرایی در تشخیص خودروهای

روبرویی دست یافته‌ایم. در واقع ماسک‌گذاری به کار گرفته شده، ایده‌ای بسیار ساده، در جهت افزایش سرعت پردازشی الگوریتم است.

نتایج حاصل از به کارگیری روش پیشنهادی اول بیان گر دقت تشخیص خطوط ۹۶/۷۸ درصد به صورت بلادرنگ بر روی پایگاه داده IROADS مورد استفاده در شرایط متفاوت محیطی است. همچنین نتایج حاصل از به کارگیری روش پیشنهادی دوم جهت تشخیص خطوط جاده، دارای دقت پیکسلی ۹۰/۷۸ درصد می‌باشد که بیان گر این نکته است که علاوه بر درصد نسبتاً مناسبی که روش اجرایی نسبت به رویکردهای نظارت شده دارد، از لحاظ پیچیدگی محاسباتی در تعداد پارامترهای آموزشی بکار گرفته شده، با ۲۱ میلیون پارامتر آموزشی، بهبود قابل توجهی نسبت به شبکه‌های عمیق نامبرده شده داشته است. در نهایت لازم به ذکر است که خروجی الگوریتم تشخیص موانع روبروی خودرو نیز با توجه به معیار کیفی که برای دقت عمل کردی آن تعریف شده است، از دقت ۹۲/۳ درصد در تشخیص موانع برخوردار است.

کلمات کلیدی: تشخیص خطوط جاده، تبدیل هاف، تشخیص موانع، یادگیری عمیق، شبکه‌های مولد رقابتی، یادگیری نیمه نظارتی

لیست مقالات مستخرج از رساله

مقالات ژورنالی:

1- Real-time Lane Detection Based on Image Edge Feature and Hough Transform, Ali Asghar Fallah, Ali Soleymani, Hossein Khosravi, Journal of Electrical and Computer Engineering Innovations (JECEI).

2- A Method for Automatic Lane Detection Using a Deep Network, Ali Asghar Fallah, Ali Soleymani, Hossein Khosravi, International Journal of Engineering.

مقالات کنفرانسی:

۳- علی اصغر فلاح و علی سلیمانی "تشخیص خطوط جاده براساس ترکیب ویژگی‌های رنگ و لبه تصویر" چهارمین کنفرانس پردازش سیگنال و سیستم‌های هوشمند ایران، تهران، دانشگاه صنعتی امیرکبیر، دیماه ۱۳۹۷.

فهرست مطالب

فصل اول مقدمه	۱
۱-۱ مقدمه	۲
۱-۱-۱ زیرساخت‌های هوشمند	۳
۲-۱ تشخیص خطوط جاده‌ای	۳
۳-۱ تشخیص موانع	۵
۴-۱ چالش‌های پیش‌روی در عملیات تشخیص خطوط جاده و تشخیص موانع	۵
۱-۴-۱ چالش‌های مربوط به شرایط محیطی	۵
۲-۴-۱ چالش‌های مربوط به پیاده‌سازی نرم‌افزاری	۶
۱-۲-۴-۱ پیچیدگی محاسباتی	۶
۲-۲-۴-۱ برچسب‌گذاری دادگان در روش‌های نظارت‌شده	۷
۳-۲-۴-۱ نیازمندی به مجموعه دادگان با تعداد نمونه‌های زیاد	۷
۴-۲-۴-۱ جامعیت الگوریتم‌های اجرایی	۷
۵-۱ ضرورت انجام تحقیق	۸
۶-۱ اهداف پژوهش	۹
۱-۶-۱ تشخیص خطوط جاده	۹
۲-۶-۱ تشخیص مانع	۱۰
۷-۱ ساختار رساله	۱۱
فصل دوم مروری بر پیشینه تحقیق	۱۳
۱-۲ مقدمه	۱۴
۲-۲ مروری بر پژوهش‌های پیشین در زمینه تشخیص خطوط جاده	۱۴
۱-۲-۲ تشخیص خطوط جاده بر اساس رویکرد کلاسیک مبتنی بر دستورات پایه‌ای پردازش تصویر	۱۵
۲-۲-۲ تشخیص خطوط جاده بر اساس رویکرد مبتنی بر یادگیری عمیق	۳۰
۳-۲ مروری بر پژوهش‌های پیشین در زمینه تشخیص موانع	۳۸
۴-۲ نتیجه‌گیری	۴۵
فصل سوم مبانی نظری	۴۷
۱-۳ مقدمه	۴۸
۲-۳ تبدیل هاف	۴۸

۴۹	۳-۳ لبه‌یابی در تصویر
۵۰	۳-۳-۱ الگوریتم لبه‌یابی کنی
۵۰	۳-۴ شبکه‌های عصبی
۵۲	۳-۵ شبکه‌های عصبی عمیق
۵۳	۳-۵-۱ معماری‌های اصلی در شبکه‌های عصبی عمیق
۵۳	۳-۵-۲ شبکه‌های عمیق کانولوشنی
۵۴	۳-۵-۲-۱ لایه‌های موجود در شبکه کانولوشن
۵۴	۳-۵-۲-۲ لایه کانولوشن
۵۶	۳-۵-۲-۳ لایه پولینگ
۵۷	۳-۵-۲-۴ لایه ReLu
۵۷	۳-۵-۲-۵ لایه تماماً متصل
۵۸	۳-۵-۳ شبکه‌های عمیق مولد رقابتی
۵۸	۳-۵-۴ الگوی یادگیری نیمه‌نظارتی
۵۹	۳-۵-۵ شبکه مولد رقابتی نیمه‌نظارتی (SGAN)
۶۲	۳-۶ توصیف‌کننده
۶۲	۳-۶-۱ هیستوگرام مبتنی بر گرادیان
۶۴	۳-۷ طبقه‌بند ماشین بردار پشتیبان
۶۵	۳-۸ معیارهای دقت شناسایی خطوط
۶۵	۳-۸-۱ معیار (IOU)
۶۶	۳-۸-۲ معیار دقت
۶۷	۳-۹ نتیجه‌گیری
۶۹	فصل چهارم روشهای پیشنهادی
۷۰	۴-۱ مقدمه
۷۰	۴-۲ الگوریتم پیشنهادی مبتنی بر رویکرد کلاسیک استفاده از دستورات پایه‌ای پردازش تصویر
۷۱	۴-۲-۱ عملیات پیش‌پردازش
۷۲	۴-۲-۲ انتخاب ناحیه مورد علاقه
۷۳	۴-۲-۳ تشخیص لبه
۷۵	۴-۲-۴ تبدیل هاف
۷۵	۴-۳ الگوریتم پیشنهادی مبتنی بر یادگیری

۱-۳-۴	شناسایی خطوط با استفاده از شبکه مولد رقابتی نیمه نظارتی	۷۶
۱-۱-۳-۴	شبکه تفکیک کننده (نظارت شده)	۷۶
۲-۱-۳-۴	شبکه تفکیک کننده (نظارت نشده)	۷۹
۳-۱-۳-۴	شبکه مولد	۷۹
۴-۱-۳-۴	هایپر پارامترهای مربوط به آموزش شبکه ها	۸۱
۱-۴-۱-۳-۴	شبکه مولد	۸۱
۲-۴-۱-۳-۴	شبکه تفکیک کننده (نظارت شده)	۸۱
۳-۴-۱-۳-۴	شبکه تفکیک کننده (نظارت نشده)	۸۱
۴-۴	شناسایی موانع	۸۲
۵-۴	الگوریتم کلی شناسایی مانع	۸۲
۱-۵-۴	استخراج ویژگی از تصاویر دادگان کمی با استفاده از الگوریتم HOG	۸۳
۲-۵-۴	آموزش طبقه بند ماشین بردار پشتیبان برای شناسایی نوع مانع مورد نظر	۸۴
۳-۵-۴	جستجوی مانع هدف در تصویر	۸۵
۱-۳-۵-۴	لغزش کشویی برای جستجوی مانع	۸۵
۲-۳-۵-۴	تعیین کلاس هر پنجره برای شناسایی مانع	۸۶
۳-۳-۵-۴	تعیین پنجره بهینه شناسایی مانع	۸۶
۴-۵-۴	تعیین موانع در محدوده روبروی خودرو	۸۷
۶-۴	نتیجه گیری	۸۸
	فصل پنجم نتایج	۸۹
۱-۵	مقدمه	۹۰
۲-۵	معرفی دادگان	۹۰
۱-۲-۵	پایگاه داده IROADS	۹۱
۲-۲-۵	پایگاه داده Aly's Method	۹۱
۳-۲-۵	پایگاه داده Tusimple	۹۲
۴-۲-۵	پایگاه داده تهیه شده داخلی	۹۲
۳-۵	نتایج حاصل از روش پیشنهادی اول (رویکرد کلاسیک مبتنی بر دستورات پایه ای پردازش تصویر)	۹۳
۱-۳-۵	معیار سنجش کیفی شناسایی خطوط در روش پیشنهادی اول (نرخ تشخیص)	۹۳
۲-۳-۵	نتایج بصری روش پیشنهادی اول	۹۴
۳-۳-۵	دقت روش پیشنهادی اول	۹۶

۹۸.....	۴-۵ نتایج حاصل از روش پیشنهادی دوم (رویکرد مبتنی بر یادگیری عمیق)
۹۹.....	۱-۴-۵ ارزیابی بصری روش پیشنهادی دوم
۹۹.....	۱-۴-۵ خروجی شبکه مولد
۱۰۰.....	۲-۴-۵ شناسایی خطوط با استفاده از قسمت تفکیک‌کننده
۱۰۲.....	۲-۴-۵ دقت روش پیشنهادی دوم (روش مبتنی بر یادگیری عمیق)
۱۰۳.....	۵-۵ شناسایی مانع
۱۰۵.....	۶-۵ نتیجه‌گیری
۱۰۷.....	فصل ششم جمع‌بندی
۱۰۸.....	۱-۶ مقدمه
۱۰۹.....	۲-۶ مزایای روش پیشنهادی اول
۱۰۹.....	۱-۲-۶ روش اجرایی بلادرنگ
۱۰۹.....	۲-۲-۶ پویایی انتخاب ناحیه مورد علاقه
۱۰۹.....	۳-۲-۶ کارایی بالا در انواع شرایط محیطی با توجه به دو پایگاه داده به کار گرفته شده
۱۱۰.....	۳-۶ معایب روش پیشنهادی اول
۱۱۰.....	۱-۳-۶ غیرخودکار
۱۱۰.....	۲-۳-۶ عدم وجود معیار کمی مطمئن
۱۱۰.....	۴-۶ مزایای روش پیشنهادی دوم
۱۱۰.....	۱-۴-۶ عدم نیاز به دادگان کاملاً برچسب‌گذاری شده
۱۱۱.....	۲-۴-۶ قابلیت بالا در توسعه
۱۱۱.....	۳-۴-۶ روش خودکار
۱۱۱.....	۴-۴-۶ پارامتریک و قابل ارزیابی
۱۱۱.....	۵-۶ معایب روش پیشنهادی دوم
۱۱۲.....	۶-۶ مزایای روش تشخیص مانع
۱۱۲.....	۱-۶-۶ سادگی، عدم نیاز به پردازش زیاد
۱۱۲.....	۷-۶ معایب روش تشخیص مانع
۱۱۲.....	۱-۷-۶ نیاز به تنظیمات دستی بر روی تصاویر مختلف جاده
۱۱۲.....	۸-۶ پیشنهاداتی برای بهبود در آینده
۱۱۴.....	مراجع

فهرست شکل‌ها

شکل ۱-۱: انواع خطوط جاده‌ای [۵]	۴
شکل ۲-۱: نمونه‌هایی از شرایط محیطی جاده [۵]	۶
شکل ۳-۱: ایده‌آل‌های به‌کارگیری رویکرد میانی	۹
شکل ۴-۱: دیاگرام کلی فرایند تشخیص خطوط جاده و موانع	۱۱
شکل ۱-۲: دیاگرام کلی تشخیص خطوط جاده‌ای	۱۴
شکل ۲-۲: مراحل تشخیص خطوط جاده در روش GOLD [۷]	۱۶
شکل ۳-۲: نتایج روش GOLD [۶]	۱۷
شکل ۴-۲: مکان نقطه محوشدگی [۱۰]	۱۹
شکل ۵-۲: تقسیم تصویر جاده به دو ناحیه [۱۶]	۲۲
شکل ۶-۲: انتخاب ناحیه مورد علاقه از تصویر ورودی [۱۹]	۲۲
شکل ۷-۲: انتخاب ناحیه مورد علاقه از تصویر ورودی [۲۰]	۲۳
شکل ۸-۲: فلوچارت روش پیشنهادی در مرجع [۲۱]	۲۴
شکل ۹-۲: تعیین ناحیه مورد علاقه توسط نقطه محوشدگی [۲۱]	۲۵
شکل ۱۰-۲: مقادیر کانال Y برای رنگهای مختلف در شرایط روشنایی مختلف [۲۱]	۲۵
شکل ۱۱-۲: مقادیر کانال Cb برای رنگهای مختلف در شرایط روشنایی مختلف [۲۱]	۲۶
شکل ۱۲-۲: فلوچارت روش پیشنهادی در مرجع [۲۴]	۲۸
شکل ۱۳-۲: بلوک دیاگرام روش پیشنهادی در مرجع [۲۷]	۲۹
شکل ۱۴-۲: شبکه عمیق به‌کار گرفته‌شده در روش پیشنهادی مرجع [۳۸]	۳۱
شکل ۱۵-۲: رویکرد الگوریتم پیشنهادی در پژوهش [۳۸]	۳۱
شکل ۱۶-۲: معماری شبکه در روش پیشنهادی مرجع [۳۹]	۳۲
شکل ۱۷-۲: ساختار کلی روش پیشنهادی در مرجع [۴۰]	۳۳
شکل ۱۸-۲: معماری الگوریتم پیشنهادی در مرجع [۴۱]	۳۳
شکل ۱۹-۲: ساختار معماری شبکه استفاده شده در مرجع [۴۲]	۳۵
شکل ۲۰-۲: فلوچارت روش پیشنهادی در مرجع [۴۴]	۳۶
شکل ۲۱-۲: معماری روش پیشنهادی مرجع [۴۵]	۳۷
شکل ۲۲-۲: فرایند پردازش برای هر فریم در روش پیشنهادی مرجع [۵۷]	۳۹
شکل ۲۳-۲: بلوک دیاگرام روش پیشنهادی در مرجع [۶۲]	۴۱

- شکل ۲-۲۴: روش پیشنهادی تشخیص موانع در مرجع [۶۹]..... ۴۳
- شکل ۲-۲۵: روش پیشنهادی تشخیص موانع در مرجع [۷۰]..... ۴۴
- شکل ۲-۲۶: تشخیص سایه خودرو به همراه سایه به عنوان مانع در روش پیشنهادی [۷۱]..... ۴۴
- شکل ۲-۲۷: تشخیص سایه خودرو به همراه سایه به عنوان مانع در روش پیشنهادی [۷۲]..... ۴۵
- شکل ۳-۱: نحوه محاسبه خط با استفاده از تبدیل هاف..... ۴۹
- شکل ۳-۲: ساختار یک شبکه عصبی ساده..... ۵۱
- شکل ۳-۳: شماتیک شبکه عمیق در مقابل شبکه عصبی ساده..... ۵۳
- شکل ۳-۴: مثالی از معماری شبکه کانولوشن..... ۵۴
- شکل ۳-۵: عملیات کانولوشنی و محاسبه یک مجموعه نقشه ویژگی..... ۵۵
- شکل ۳-۶: خروجی عملیات پولینگ بیشینه..... ۵۶
- شکل ۳-۷: تابع Relu..... ۵۷
- شکل ۳-۸: شبکه مولد رقابتی نیمه نظارتی..... ۶۰
- شکل ۳-۹: الف. تصویر اصلی ب. تصویر HOG..... ۶۴
- شکل ۳-۱۰: روش ماشین بردار پشتیبان خطی..... ۶۵
- شکل ۳-۱۱: مفهوم معیار IOU..... ۶۶
- شکل ۴-۱: بلوک دیاگرام تشخیص خطوط جاده با استفاده از رویکرد کلاسیک..... ۷۱
- شکل ۴-۳: الف. ماسک جهت انتخاب ناحیه مورد علاقه ب. تصویر دودویی ناحیه مورد علاقه..... ۷۳
- شکل ۴-۴: ماسکهای الگوریتم لبه‌یابی..... ۷۴
- شکل ۴-۵: اعمال الگوریتم کنی به تصویر..... ۷۵
- شکل ۴-۶: خروجی نهایی الگوریتم تشخیص خطوط جاده با رویکرد کلاسیک (روش پیشنهادی اول)..... ۷۵
- شکل ۴-۷: فرایند شناسایی خطوط با استفاده از شبکه مولد رقابتی نیمه نظارتی..... ۷۶
- شکل ۴-۸: شبکه Unet استفاده شده در تفکیک کننده..... ۷۸
- شکل ۴-۹: ساختار شبکه مولد به کار گرفته شده..... ۷۹
- شکل ۴-۱۰: بلوک دیاگرام شناسایی مانع در تصویر..... ۸۳
- شکل ۴-۱۱: تعدادی از تصاویر دادگان برای آموزش مدل تشخیص نوع مانع..... ۸۴
- شکل ۴-۱۲: محدوده اصلی جست و جوی خودرو با عملیات لغزش کشویی پنجره..... ۸۶
- شکل ۴-۱۳: شناسایی پنجره بهینه در هر فریم و پنجره‌های دارای همپوشانی..... ۸۷
- شکل ۴-۱۴: تعیین محدوده فرضی مقابل ماشین با توجه به ساختار تصاویر و محل قرارگیری دوربین..... ۸۷
- شکل ۵-۱: نمونه هایی از تصاویر پایگاه داده‌ها IROADS [۸۱]..... ۹۱

- شکل ۵-۲: نمونه‌هایی از تصاویر پایگاه داده‌ها Aly's Method [۸۲] ۹۲
- شکل ۵-۳: تعدادی از تصاویر پایگاه داده‌ها Tusimple ۹۲
- شکل ۵-۴: تصاویر پایگاه داده‌ها تهیه‌شده داخلی ۹۳
- شکل ۵-۵: تعدادی از نتایج آزمایشات روش پیشنهادی اول بر روی پایگاه داده‌ها IROADS با خروجی مطلوب ۹۵
- شکل ۵-۶: تعدادی از نتایج آزمایشات روش پیشنهادی اول بر روی پایگاه داده‌ها Aly's Method با خروجی مطلوب ۹۵
- شکل ۵-۷: تعدادی از تصاویر دارای خطا در تشخیص خطوط برای پایگاه داده‌ها IROADS ۹۶
- شکل ۵-۸: تعدادی از تصاویر دارای خطا در تشخیص خطوط برای پایگاه داده‌ها Aly's Method ۹۶
- شکل ۵-۹: تعدادی از تصاویر تولیدشده با به‌کارگیری شبکه مولد و تصاویر شبیه متناظر با آن‌ها در مجموعه داده اصلی ۱۰۰
- شکل ۵-۱۰: عمل‌کرد مطلوب شناسایی خطوط از به‌کارگیری شبکه عمیق مولد رقابتی نیمه‌نظارتی ۱۰۱
- شکل ۵-۱۱: تعدادی از خروجی‌های با عمل‌کرد نامطلوب در شناسایی خطوط در روش پیشنهادی مبتنی بر یادگیری عمیق ۱۰۱
- شکل ۵-۱۲: نمونه خروجی شناسایی خطوط با استفاده از ResNet50 ۱۰۱
- شکل ۵-۱۳: نمونه خروجی تشخیص موانع ۱۰۴

فهرست جدول‌ها

- جدول ۱-۲: تعدادی از پژوهش‌های شناسایی جاده و خطوط با استفاده از رویکرد مبتنی بر یادگیری عمیق ۳۷
- جدول ۱-۵: نتایج حاصل از روش پیشنهادی اول روی پایگاه داده‌ها IROADS ۹۷
- جدول ۲-۵: مقایسه روش پیشنهادی اول با تعدادی از پژوهش‌های مرتبط دیگر بر روی پایگاه داده‌ها IROADS ۹۷
- جدول ۳-۵: نتایج حاصل از روش پیشنهادی اول روی پایگاه داده‌ها Aly's Method ۹۸
- جدول ۴-۵: مقادیر دقت شناسایی مبتنی بر پیکسل (ACC) و معیار IOU ۱۰۲
- جدول ۵-۵: مقادیر دقت شناسایی تعدادی از پژوهش‌های مرتبط ۱۰۴

علائم اختصاری

Bag Of Word	BOW	کیف کلمات
Convolutional Neural Networks	CNN	شبکه‌های عصبی کانولوشن
Fully Convolutional Networks	FCN	شبکه‌های کانولوشنی تماما متصل
Gussian Mixture Model	GMM	مدل مخلوط گوسی
Hidden Markov Model	HMM	مدل مخفی مارکوف
Histogram Oriented Gradient	HOG	هیستوگرام مبتنی بر گردایان
Intelligent Transport System	ITS	سیستم حمل و نقل هوشمند
Intersection Over Union	IOU	اشتراک اجتماع
Inverse Prespective Maping	IPM	نقشه منظر معکوس
Long Short Term Memory	LSTM	حافظه کوتاه بلند مدت
Mean Average Precision	MAP	میانگین دقت
Maximum likelihood estimation	MLE	تخمین بیشترین شباهت
Multi Layer Perceptron	MLP	شبکه چندلایه پرسپترون
Principal Component Analysis	PCA	تجزیه و تحلیل مؤلفه های اصلی
Recurrent Convolutional Neural Networks	RCNN	شبکه‌های عصبی کانولوشنی بازگشتی
Region Of Interest	ROI	ناحیه مورد علاقه
Recurrent Neural Networks	RNN	شبکه‌های عصبی بازگشتی
Semi-supervised Generative Adversarial Networks	S-GAN	شبکه‌های مولد رقابتی نیمه‌نظاتی
Scale Invariant Feature Transform	SIFT	تبدیل فرم ویژگی مقیاس ثابت
Speeded Up Robust Features	SURF	ویژگی مقاوم سرعت یافته
Singular Value Decomposition	SVD	تجزیه مقادیر تکین
Support Vector Machine	SVM	ماشین بردار پشتیبان
Visual Geometry Group	VGG	گروه هندسی بصری
Vehicle Make and Model Recognition	VMMR	شناسایی مدل‌های خودرو
Vanishing Point	VP	نقطه محوشوندگی
You Look Only Once	YOLO	فقط یک نگاه کن

فصل اول

مقدمه

صنعت حمل و نقل، به دلایلی همچون انعطاف‌پذیری، دسترسی آسان و سهم بالا در جابه‌جائی مسافر و کالا در کشور، از اهمیت ویژه برخوردار بوده و لذا یکی از صنایع فعال در جهان است. از طرفی باتوجه به مشکلات مربوط به حمل و نقل جاده‌ای، همچون تصادفات، که از پرآمارترین اتفاقاتی است که در دنیا منجر به وقوع حادثه‌هایی بعضاً دلخراش می‌شود، ایمنی جاده‌ها به یکی از مهم‌ترین حوزه‌های تحقیقی و توسعه‌ای صنعت نامبرده تبدیل شده‌است. شناخت عوامل مهم در بروز مشکلات موجود، اولین مرحله‌ای است می‌تواند ما را در انجام مطالعات و توسعه سیستم‌های ایمنی به سمت هدف سوق دهد. معمولاً سه عامل انسان، جاده و وسیله نقلیه سه منبع اصلی تصادفات جاده‌ای به شمار می‌آیند. باید سعی داشت در هریک از زمینه‌هایی که عوامل اصلی وقوع این مخاطرات می‌شوند، با انجام مطالعات و اقدامات برنامه‌ریزی‌شده میزان وقوع آن‌ها را به کمترین حد ممکن رساند. از طرفی کاهش اینگونه مخاطرات، مستلزم به‌کارگیری قوانین و همچنین ایجاد زیرساخت‌های سخت افزاری و نرم افزاری است که بتوان با استفاده از آن‌ها اتفاقات احتمالی را به حداقل رساند.

مهم‌ترین عامل از بین این عوامل نامبرده، خطاهای انسانی از جمله انحراف خودروها از خطوط‌شان هست، که می‌تواند ناشی از خستگی، خواب‌آلودگی و کاهش سطح هوشیاری راننده باشد. بدین منظور ایجاد سیستم‌هایی همچون، سیستم تشخیص خطوط جاده‌ای و همچنین سیستم تشخیص موانع از مهم‌ترین عملیات‌های با ماهیت نرم‌افزاری هستند که در کنترل خودکار(رانندگی خودکار) مطرح می‌شود [۴-۱].

در واقع سیستم‌های تشخیص خطوط جاده‌ای و تشخیص موانع از جمله تجهیزات ایمنی نرم‌افزاری می‌باشند که در سیستم‌های ترافیکی، برای کاهش خطرات ناشی از انحراف خودروها از خطوط در جاده‌های شهری و بین شهری بسیار حائز اهمیت هستند و مورد توجه قرار گرفته می‌گیرند.

در این زمینه مطالعات بسیاری در حوزه بینایی ماشین انجام گرفته است و اغلب این مطالعات سعی در ارائه سیستم‌های با دقت بالا داشته‌اند که نتایج خوبی نیز به‌دست آورده‌اند. اما انجام چنین فرآیندی با توجه به نیازمندی‌های لازم برای پیاده‌سازی این سیستم‌ها، با چالش‌هایی مواجه است که سعی در حل آن‌ها باید داشت.

سیستم‌هایی که برای کاهش مخاطرات درون جاده‌های مطرح هستند، معمولاً با توجه به عوامل اصلی که مطرح شدند، در دو مفهوم اصلی بیان می‌گردند. ۱. زیرساخت‌های هوشمند ۲. خودروهای هوشمند

۱-۱-۱ زیرساخت‌های هوشمند

اغلب مقصود از زیرساخت هوشمند، امکاناتی است که در جاده‌ها برای کنترل ترافیک و ثبت و ضبط اطلاعاتی همچون سیستم‌های پلاک‌خوان، سیستم‌های تردد خودروها، سیستم‌های بیان ترافیک و اطلاعات آب و هوایی می‌توان اشاره نمود. مهم‌ترین وظیفه‌ای که در ایجاد زیرساخت‌های هوشمند وجود دارد، کنترل ترافیک و روان‌سازی جریان‌های جاده‌ای و به تبع آن افزایش سطح ایمنی، برای کاهش مخاطرات عوامل غیرانسانی است. این عمل با ذخیره‌سازی و ارسال داده‌های تهیه‌شده از طریق پردازش تصاویر ویدیویی و حسگرهای موجود در تقاطع‌ها و جاده‌های درون‌شهری و برون‌شهری صورت می‌گیرد.

۱-۱-۲ خودروهای هوشمند

آنچه که ما به‌عنوان تشخیص خطوط جاده‌ای و همچنین تشخیص موانع سعی در انجام آن داریم، در بررسی خودروهای هوشمند مطرح می‌شوند. در واقع هدف اصلی از ایجاد خودروهای هوشمند، کمک به راننده در کاهش مخاطرات جاده‌ای با منبع انسانی است. از جمله موارد دیگری که علاوه بر تشخیص خطوط جاده‌ای و تشخیص موانع برای خودروهای هوشمند می‌توان از آنان نام برد: سیستم‌های اضطراری کمک ترمز، تشخیص جهت حرکت بین خطوط، هشدار تصادف از جلو و تشخیص خواب‌آلودگی و حواس‌پرتی راننده می‌باشند. در واقع هدف اصلی ایجاد خودروهای هوشمند، گرفتن اطلاعات از محیط اطراف و رسیدن به درک درستی از موقعیت خودرو نسبت به محیط اطراف است تا درنهایت خودروی هوشمند به‌دست‌آید.

۱-۲ تشخیص خطوط جاده‌ای

خطوط موازی که در امتداد طول مسیر حرکت خودروها درون جاده‌ها با ساختارهای مشخصی قابل مشاهده‌اند، خطوط جاده نامیده می‌شوند. این خطوط می‌توانند به‌صورت پیوسته، گسسته و یا به‌صورت میخک‌های با رنگ

زرد مشخص باشند. شکل ۱-۱، انواع مختلف خطوط جاده را نشان می‌دهد. رنگ این خطوط سفید و گاهی زرد است [۵].



شکل ۱-۱: انواع خطوط جاده‌ای [۵]

تشخیص خطوط جاده عبارت است از شناسایی مرزهای خطوط جاده در یک تصویر جاده، بدون اینکه اطلاعات قبلی از مکان خطوط آن در دسترس باشد. البته این تعریف بیشتر برای روش‌های کلاسیک که با استفاده از الگوریتم‌های پردازش تصویر کار می‌کنند، مطرح می‌شود. در سیستم‌های تشخیص خطوط جاده که با استفاده از روش‌های مبتنی بر یادگیری همچون یادگیری عمیق نظارتی انجام می‌شوند، هدف، تشخیص خطوط به صورت تماماً خودکار است. لذا با استفاده از تصاویری که در آن‌ها خطوط جاده مشخص شده‌اند (تصاویر ماسک)، اقدام به ایجاد یک مدل کلی برای تخمین خطوط می‌نماییم. در سیستم‌های تشخیص خطوط نظارت نشده، فرایند شناسایی خطوط، در اغلب روش‌هایی که با رویکرد کلاسیک مبتنی بر الگوریتم‌های پایه‌ای پردازش تصویر انجام می‌شوند، با وجود سرعت بالایی که دارند، با مشکلاتی همچون نیازمندی سیستم‌ها به تنظیمات غیرخودکار اولیه، مواجه هستند.

هدف اصلی از ایجاد سیستم تشخیص خطوط جاده برای انجام واکنش بلادرنگ، همچنین اعلام هشدار به موقع به راننده خودرو در شرایطی است که انحراف از مسیر اصلی جاده، ممکن است رخ دهد. چرا که اکثر تصادفات درون جاده‌ای به دلیل عدم آگاهی راننده از شرایط محیطی جاده و یا غفلت وی نسبت به شرایط ممکن است. سیستم‌های تشخیص خطوط جاده به عنوان یک تکنولوژی مهم برای رانندگی مطمئن و کاهش تصادفات جاده‌ای پیشنهاد شده‌اند و وظیفه دارند با انتخاب درست خطوط مسیر و تعیین موقعیت مناسب برای خودرو از انحراف آن به سمت چپ یا راست جلوگیری کنند.

۳-۱ تشخیص موانع

تشخیص موانع یکی از موضوعات مورد علاقه پژوهش‌گران در زمینه پردازش تصویر و بینایی ماشین است. تشخیص

موانع جاده یک چالش اساسی در سیستم‌های حمل و نقل هوشمند، سیستم‌های دستیار راننده و خودروهای هوشمند است. سیستم تشخیص موانع جاده جهت جلوگیری از تصادفات جاده‌ای و اعلام هشدار به راننده خودرو برای عدم برخورد با موانع موجود در سطح جاده، طراحی می‌شود. در این سیستم‌ها موانع روبروی خودرو در خودروهای هوشمند برای کاهش تصادفات ناشی از برخورد با موانع جلوی خودرو پیشنهاد می‌شوند.

۴-۱ چالش‌های پیش‌روی در عملیات تشخیص خطوط جاده و تشخیص موانع

چالش‌های اصلی در تشخیص اهداف مورد نظر با توجه به رویکردهای مختلفی که برای انجام عملیاتی صورت می‌گیرند، به دو دسته کلی: ۱. چالش‌های مربوط به شرایط محیطی و ۲. چالش‌های مربوط به عملیات پیاده‌سازی نرم‌افزاری تقسیم می‌شوند.

۱-۴-۱ چالش‌های مربوط به شرایط محیطی

عواملی مانند انعکاس شدید نور خورشید که منتج به آفتاب شدید می‌شود، هوای ابری، هوای مه‌آلود، باران، سایه ساختمان‌ها که اغلب مربوط به جاده‌های داخل شهری است، همچنین سایه درختان و یا خودروها، تاریکی شب، نور داخل تونل‌ها و نور مستقیم چراغ خودروهای روبرویی باعث ایجاد شرایط مختلف برای جاده‌ها می‌شوند. همه موارد بیان شده باعث ایجاد چالش در یک ویژگی تحت مفهوم تغییر نور تصویر در سطح جاده‌ها می‌شود. این مورد اصلی‌ترین عاملی است که ممکن است باعث تشخیص‌های اشتباه در پردازش شرایط مختلف جاده در حالت‌های مختلف تصاویر جاده‌ها شود. در شکل ۱-۲ تعدادی از این موارد نشان داده شده‌اند.



شکل ۱-۲: نمونه‌هایی از شرایط محیطی جاده [۵]

۱-۴-۲ چالش‌های مربوط به پیاده‌سازی نرم‌افزاری

۱-۴-۲-۱ پیچیدگی محاسباتی

یکی از چالش‌های اصلی که در تشخیص خطوط جاده و موانع وجود دارد، حجم محاسبات زیاد جهت پردازش اطلاعات و تعیین خطوط جاده و موانع است. در مواقعی که بخواهیم از روش‌های مبتنی بر یادگیری و به‌ویژه یادگیری عمیق استفاده کنیم، این محاسبات بیشتر هم خواهند شد. عمل‌کرد صحیح سیستم‌های تشخیص خطوط جاده و موانع نیازمند وجود دو ویژگی اساسی است: اول با توجه به حساسیت کاربردی، نیازمند دقت بالایی هستند و دوم اینکه باید از سرعت بالای اجرایی و به نوعی از ویژگی بلادرنگ بودن برخوردار باشند. لذا کاهش پیچیدگی محاسباتی و افزایش توان محاسباتی، مواردی هستند که در سیستم‌های تشخیص خطوط جاده و موانع در نظر گرفته می‌شوند. در نتیجه ارائه یک الگوریتم با بار محاسباتی مناسب می‌تواند در هزینه محاسباتی صرفه‌جویی نماید. روش‌هایی که با استفاده از رویکردهای کلاسیک پردازشی که تنها با استفاده از دستورات پایه‌ای پردازش تصویر اقدام به شناسایی خطوط جاده دارند، معمولاً از پیچیدگی محاسباتی پایین و مناسبی برخوردار هستند، اما دارای دقت مناسبی نیستند. از طرفی در روش‌های نظارت‌شده دقت بالای عملیاتی وجود دارد، اما پیچیدگی محاسباتی بالایی دارند. همچنین مشکل دیگری که در روش‌های نظارت‌شده وجود دارد، نیازمندی به برچسب‌گذاری دادگان آموزش است.

۱-۴-۲-۲ برچسب‌گذاری دادگان در روش‌های نظارت‌شده

در روش‌های مبتنی بر یادگیری که از الگوهای نظارت‌شده برای عملیات قطعه‌بندی تصاویر استفاده می‌شود، معمولاً

حاشیه‌گذاری تصاویر^۱ و تهیه ماسک‌های مربوط برای دادگان تهیه‌شده، از لحاظ زمانی فرایندی بسیار هزینه‌بر است. یکی از چالش‌های اصلی موجود در انجام روش‌های مبتنی بر یادگیری همین مورد است. از طرفی دیگر استفاده از مجموعه دادگان آماده و برچسب‌گذاری‌شده نیز محدودیت‌های خاص خود را دارد. برای مثال ممکن است یک مجموعه داده با توجه به فعالیتی که قصد انجام آن را داریم، فاقد کارایی کافی باشد.

۱-۴-۲-۳ نیازمندی به مجموعه دادگان با تعداد نمونه‌های زیاد

استفاده از الگوریتم‌های با رویکرد یادگیری عمیق، معمولاً با بهبود چشمگیری از لحاظ درصد دقت در روش اجرایی همراه است. اما به کارگیری این روش‌ها معمولاً نیازمند به در اختیار داشتن مجموعه دادگان با حجم بالا است. چرا که تضمین بهبود و کارایی رویکردهای یادگیری عمیق، بیشتر منوط به به کارگیری حجم بالایی از دادگان مورد استفاده است. در واقع بزرگترین چالشی که در به کارگیری این روش‌ها وجود دارد، همین مسئله داشتن دادگان کافی برای آموزش سیستم‌های مورد نظر است.

۱-۴-۲-۴ جامعیت الگوریتم‌های اجرایی

یکی از مهم‌ترین دیدگاه‌هایی که در شناسایی خطوط جاده وجود دارد، وجود جامعیت روش پیشنهادی برای انواع جاده‌های موجود است. بدین معنی که زمانی الگوریتم و روشی برای شناسایی خطوط با توجه به یک مجموعه داده طراحی می‌شود، قابلیت انجام روش بر روی یک مجموعه داده دیگر از اهمیت بالایی برخوردار می‌شود. یکی از مهم‌ترین چالش‌هایی که اکثر روش‌های پیشین دارند، عدم قابلیت به کارگیری و استفاده الگوریتم بر روی مجموعه دادگان دیگری غیر از دادگان موجود در پژوهش صورت گرفته است. به عبارت دیگر الگوریتمی که برای یک مجموعه داده مورد استفاده قرار گرفته است، نسبت به مجموعه دادگان دیگر توسعه

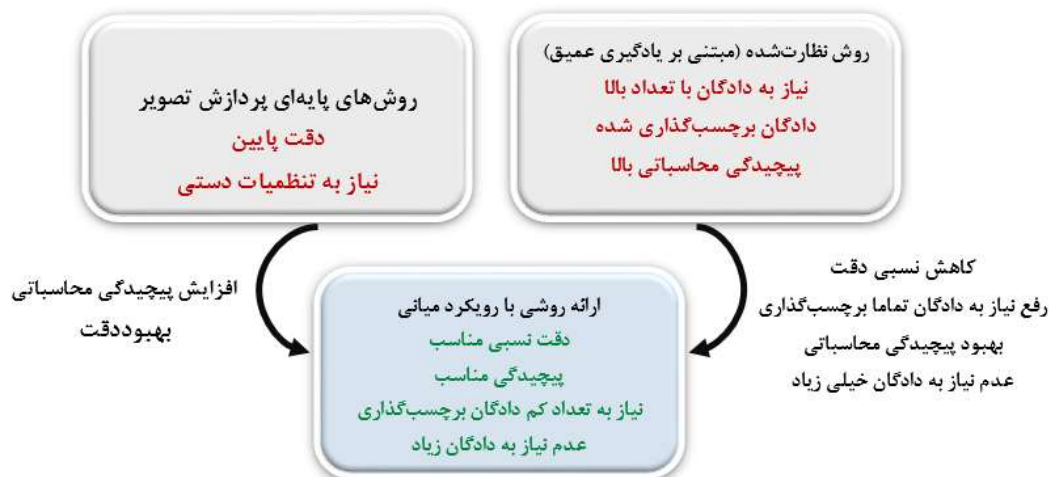
¹ Image Annotation

داده نشده است. البته حل چنین چالشی فارغ از رویکردی که مورد استفاده قرار می گیرد، نیازمند استفاده از مجموعه دادگان بسیار بالا، با در نظر گرفتن شرایط مختلف محیطی است.

۱-۵ ضرورت انجام تحقیق

تاکنون فعالیت های پژوهشی بسیاری در زمینه تشخیص خطوط و موانع انجام شده اند. اما با توجه به آنچه در بخش قبل به عنوان چالش های موجود در تشخیص خطوط و موانع بیان گردید، الگوریتم هایی که تا به حال ارائه شده اند، با ضعف هایی مواجه هستند. به طور کلی با توجه به رویکردهایی که در عملیات تشخیص خطوط و موانع وجود دارد، باید سعی نمود راهکار(هایی) ارائه داد که بتوان چالش های نامبرده را رفع نمود و یا بهبود نسبی داد. با توجه به آنچه که به عنوان چالش های موجود مطرح گردید، ارائه راه حلی که بتواند تمام ضعف های موجود در روش های پیشین را بهبود بخشد، تقریباً غیر ممکن است. اما اگر از یک سیاست مصالحه استفاده نماییم، شاید بتوان راه حلی به دست آورد، که بتواند ضعف های اصلی هریک از رویکردهای موجود را بهبود بخشد.

برای مثال در رویکردهایی که بر اساس الگوریتم های پایه ای پردازش تصویر اقدام به شناسایی خطوط می کنند، ضعف پایین بودن دقت روش اجرایی مطرح است، اما از لحاظ سرعت الگوریتمیک، بسیار مناسب هستند. از طرفی دیگر در رویکردهای مبتنی بر یادگیری نظارت شده، اگر بتوان سرعت اجرایی را با کاهش پیچیدگی محاسباتی ایجاد نمود، اتفاقی قابل قبول انجام شده است. حال برای روش های مبتنی بر یادگیری عمیق، ضعف نیاز به حجم بالای دادگان برچسب گذاری شده مطرح می شود. بهبود هریک از موارد بیان شده خود می تواند ارزشمند باشد. طبق تمام موارد بیان شده، شکل ۳-۱، ضرورت انجام تحقیق در موضوع مورد بحث، با یک رویکرد میانی را نشان می دهد.



شکل ۱-۳: ایده‌آل‌های به کارگیری رویکرد میانی

۱-۶ اهداف پژوهش

۱-۶-۱ تشخیص خطوط جاده

هدف ابتدایی از انجام این رساله، ایجاد یک سیستم تشخیص خطوط جاده، با استفاده از پردازش اطلاعات تصاویر ویدیویی تهیه شده توسط دوربین نصب‌شده در داخل خودرو است. بدین منظور با استفاده از الگوریتم‌های یادگیری عمیق مبتنی بر رویکرد نیمه‌نظارتی مدلی خودکار جهت تشخیص خطوط ارائه می‌دهیم. در مدل‌های نیمه‌نظارتی اغلب طبق آنچه در قبل بیان شد، به دلیل هزینه‌بر بودن تولید ماسک برای دادگان، تعداد کمی از نمونه‌ها دارای برجسته (ماسک خطوط جاده) بوده و تعداد زیادی از نمونه‌ها بدون برجسته هستند. در این روش‌ها دادگان بدون برجسته نقش بسیار مهمی را در آموزش مدل یادگیری ایفا می‌کنند. دو هدف اصلی مدنظر از به کارگیری رویکرد نیمه‌نظارتی در این رساله، را می‌توان موارد زیر دانست:

۱. بهبود دقت نسبت روش‌های نظارت‌نشده مبتنی بر الگوریتم‌های پایه‌ای پردازش تصویر

۲. بهبود پیچیدگی محاسباتی روش‌های نظارت‌نشده مبتنی بر یادگیری عمیق

به طور کلی روش‌های نیمه‌نظارتی همچون یک فرآیند یادگیری مبتنی بر آزمون هستند که بین ناظر و یادگیرنده وجود دارد.

۱-۶-۲ تشخیص موانع

در بخش دوم کار هدف شناسایی موانع روبروی خودرو است. مقصود از موانع روبروی خودرو در اینجا، صرفاً خودروهای دیگری هستند که در درون جاده می‌باشند. با توجه به ساختار تصاویری که وجود دارند، در بخش دوم با استفاده از فرایند مبتنی بر یادگیری اقدام به شناسایی موانع روبروی خودرو خواهیم نمود. بدین‌منظور ابتدا با شناسایی اشیای در هر فریم ویدیویی اقدام به شناسایی آن‌ها می‌نماییم. سپس با استفاده از یک فرضیه اصلی که با توجه به ساختار تصاویر داریم، اقدام به تعیین موانع روبرو خودرو خواهیم نمود.

فرضیه اصلی که با توجه به ساختار تصاویر برای بیننده کاملاً مشهود و بدیهی است، مکان قرارگیری دوربین درون

خودرو است. معمولاً با توجه به ساختار منبع فیلم‌برداری از لحاظ اینکه خود منبع در حرکت است یا خیر، از الگوریتم‌های مشخص ردیابی نظیر فیلترهای هم‌چون ذره، کالمن و غیره استفاده می‌شود. در تصاویری که ما در اختیار داریم، با توجه به اینکه خودرو در حرکت است، بنابراین ما پس‌زمینه ثابتی نخواهیم داشت. لذا استفاده از الگوریتم‌های متداول ردیابی با فیلترهای نامبرده، دقت کمی خواهند داشت. (ممکن است دوربین در حرکت کمی لرزش داشته باشد، اما مقدار آن ناچیز و در ردیابی تأثیر چندانی ندارد.) در نهایت با توجه به تخمینی که از محدوده روبروی خودرو با فرضیات تعیین‌شده خواهیم زد، اقدام به شناسایی موانع روبروی خودرو خواهیم نمود. دیاگرام کلی فرایند تشخیص خطوط و موانع در شکل ۱-۴ قابل مشاهده است.



شکل ۱-۴: دیاگرام کلی فرایند تشخیص خطوط جاده و موانع

۷-۱ ساختار رساله

در این فصل مقدمات موضوع و مساله تحقیق، چالش‌ها، ضرورت انجام تحقیق و اهداف پژوهش بررسی شد. در ادامه و در فصل دوم، به مروری بر کارهای انجام‌شده در زمینه تشخیص خطوط جاده و همچنین تشخیص موانع می‌پردازیم. در فصل سوم به بیان مبانی نظری به‌کار گرفته‌شده در انجام روال کلی کار خواهیم پرداخت. فصل چهارم، روش‌های پیشنهادی خود را به‌طور کامل توضیح می‌دهیم. بدین منظور، در ابتدا به بیان قسمت اول رساله که فرایند تشخیص خطوط جاده است خواهیم پرداخت و سپس در قسمت دوم، چگونگی تشخیص موانع روبروی خودرو را بیان می‌نماییم. در فصل پنجم نتایج حاصل از به‌کارگیری روش‌های پیشنهادی را بیان نموده و آن‌ها را با تعدادی از پژوهش‌های پیشین مورد مقایسه قرار می‌دهیم. نهایتاً در فصل ششم به بیان یک جمع‌بندی کلی از موضوع و فعالیت‌های انجام‌شده خواهیم پرداخت، مزایا و معایب روش‌های مطرح شده و پیشنهاداتی برای بهبود در آینده را توضیح خواهیم داد.

فصل دوم

مروری بر پیشینه تحقیق

۱-۲ مقدمه

در فصل قبل به مقدمه‌ای از موضوع مورد بحث پرداختیم. سیستم‌های تشخیص خطوط جاده‌ای و سیستم‌های تشخیص مانع را معرفی نمودیم. همچنین به چالش‌های موجود در فرایند عملیاتی و پیاده‌سازی آن‌ها پرداختیم. در این فصل پیشینه تحقیق تشخیص خطوط جاده و تشخیص موانع مورد بررسی قرار می‌گیرد و روش‌هایی که در این زمینه‌ها به کار گرفته شده‌اند مرور می‌شوند. همچنین الگوریتم‌های آن‌ها ارائه شده و نقاط ضعف و قوت آن‌ها مورد بررسی قرار می‌گیرند. با توجه به اینکه سیستم‌های تشخیص خطوط جاده و تشخیص موانع در کارهای قبلی به صورت جداگانه مورد بررسی قرار گرفته‌اند در این فصل پیشینه تحقیق دو سیستم را به صورت جداگانه مرور می‌کنیم. در انتها یک جمع بندی کلی و نتیجه گیری از روش‌های مطرح شده ارائه می‌کنیم.

۲-۲ مروری بر پژوهش‌های پیشین در زمینه تشخیص خطوط جاده

همانطور که در فصل اول بیان نمودیم، دو رویکرد اصلی در عملیات تشخیص جاده‌ای مطرح است. یکی تشخیص خطوط جاده‌ای با استفاده از الگوریتم‌های پایه‌ای پردازش تصاویر و دیگری الگوریتم‌های مبتنی بر یادگیری و علی‌الخصوص یادگیری عمیق که با رویکرد نظارت‌شده اقدام به تشخیص خطوط جاده‌ای می‌کنند. شکل ۱-۲ یک دیاگرام کلی از نحوه‌ی رویکرد عملیاتی و پژوهش‌هایی که در ادامه برای فرایند تشخیص خطوط جاده‌ای را بررسی می‌کنیم، نشان می‌دهد.



شکل ۱-۲: دیاگرام کلی تشخیص خطوط جاده‌ای

۲-۲-۱ تشخیص خطوط جاده بر اساس رویکرد کلاسیک مبتنی بر دستورات پایه‌ای پردازش تصویر

تشخیص خطوط جاده یکی از مهم‌ترین نیازهای خودروهای هوشمند به حساب می‌آید به این دلیل که در هدایت وسیله نقلیه، جلوگیری از تصادفات، هشدار انحراف از مسیر و کنترل عرضی خودرو کاربرد دارد. روش‌های مختلفی برای تشخیص خطوط جاده با استفاده از بینایی ماشین ارائه شده است.

در اکثر روش‌ها از ترکیب ویژگی‌های سطح پایین برای تشخیص خطوط جاده استفاده می‌شود. مهم‌ترین ویژگی که برای تشخیص خطوط جاده استفاده می‌شود، ویژگی لبه‌های تصویر است. برای تشخیص بهتر لبه‌ها در تصویر، اولین مرحله استفاده از تکنیک‌های پیش پردازش مناسب تصویر است سپس عملگر تشخیص لبه به کار گرفته می‌شود. اطلاعات مربوط به زاویه لبه‌های تصویر نیز به عنوان ویژگی مورد استفاده قرار می‌گیرد. از طرف دیگر با توجه به اینکه رنگ خطوط جاده سفید یا زرد است از ویژگی رنگ نیز استفاده می‌شود. ویژگی دیگری که می‌توان از آن استفاده نمود، شکل هندسی خطوط جاده است به این دلیل که شکل هندسی خطوط مستطیلی است می‌تواند در تشخیص خطوط در تصاویر موثر واقع شود. در ادامه تعدادی از روش‌هایی که با استفاده از ویژگی‌های سطح پایین خطوط جاده را تشخیص می‌دهند، مرور می‌کنیم.

سیستم GOLD^۱ یکی از معروف‌ترین سیستم‌های تشخیص خطوط جاده است که بر اساس تصویربرداری از جاده توسط دوربین نصب شده در خودرو انجام می‌شود [۶]. این سیستم قابلیت تشخیص خطوط جاده و نیز تشخیص موانع را دارد. تشخیص مسیر براساس تکنیک تطبیق الگو^۲ است که خطوط جاده را به دست می‌آورد و مکان‌یابی موانع در جلوی خودرو براساس پردازش تصاویر استریو انجام می‌گردد. به دلیل وجود اثر پرسپکتیو در تصویر عرض خطوط خط کشی جاده با فاصله از دوربین تغییر می‌کند. بنابراین ابتدا اثر پرسپکتیو تصویر برداشته می‌شود تا بتوان از تکنیک تطبیق الگو استفاده کرد. برای حذف اثر پرسپکتیو باید اطلاعات موقعیت، فاصله کانونی و زاویه دوربین را دانست. در این سیستم اثر پرسپکتیو با استفاده از نگاشت معکوس پرسپکتیو^۳ از

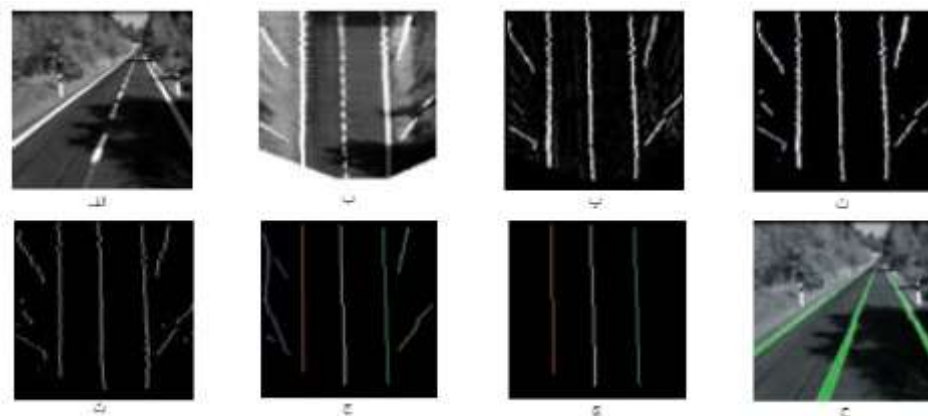
^۱ Generic Obstacle and Lane Detection

^۲ Pattern-matching technique

^۳ Inverse Perspective Mapping (IPM)

تصویر جاده حذف می‌شود. تصویر به‌دست آمده نمایی از جاده از دید چشم پرنده^۱ را نشان می‌دهد. خطوط جاده در این تصاویر به‌صورت موازی با یکدیگر بوده و از همسایگی‌های خود در چپ و راست روشن‌تر می‌باشند. بنابراین با جستجوی الگوهای افقی روشنایی تیره - روشن-تیره، توسط یک آشکارساز لبه افقی، قطعات خطوط مستقیم تصویر استخراج می‌شوند. در ادامه قطعات بهم نزدیک و یا دارای جهت مشابه را با یکدیگر ترکیب کرده، تا امکان ایجاد خطا توسط نویز یا انسداد حذف شود. در نهایت خطوطی که به یک مدل جاده با سه خط خط-کشی مطابقت بیشتری دارند به عنوان خطوط جاده انتخاب می‌شوند [۷]. مزیت‌های این روش مقاوم بودن الگوریتم در برابر تغییرات روشنایی، بافت‌های مختلف جاده، سایه‌ها و حرکت خودرو هستند.

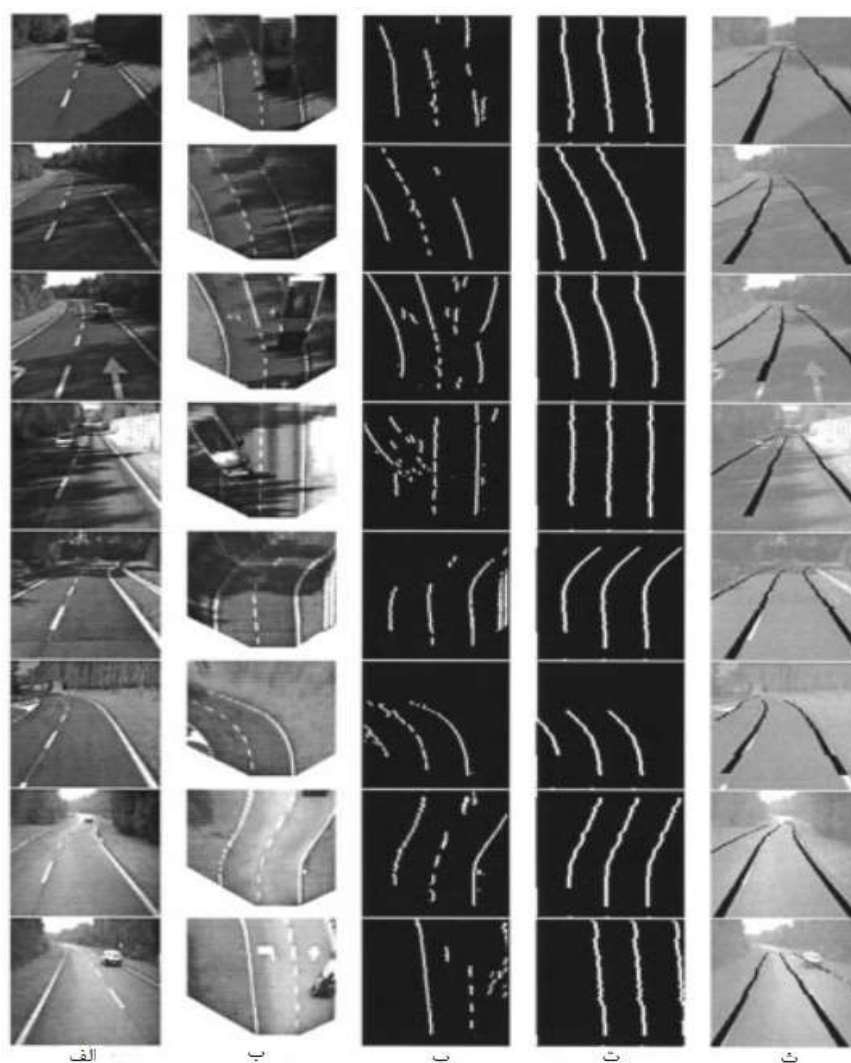
شکل ۲-۲ مراحل تشخیص خطوط جاده در روش GOLD را نشان می‌دهد. در قسمت (الف) تصویر جاده نشان داده شده است، در قسمت (ب) اثر پرسپکتیو تصویر حذف شده و در قسمت (پ) تصویر بهبود داده می‌شود تا در قسمت (ت) لبه‌های آن استخراج شود. در قسمت (ث) قطعات مستقیم استخراج شده و در قسمت (ج) این قطعات با هم ترکیب می‌شوند و در قسمت (چ) خطوطی که بیشترین تطبیق با الگوی جاده با سه خط را دارند به عنوان خطوط جاده انتخاب می‌شوند و در نهایت قسمت (ح) نتیجه نهایی بر روی تصویر جاده را نشان می‌دهد.



شکل ۲-۲: مراحل تشخیص خطوط جاده در روش GOLD [۷]

¹ Bird's eye view

در شکل ۲-۳ نتایج این روش بر روی تعدادی تصویر مختلف نشان داده شده است. در قسمت (الف) تصاویر اصلی نشان داده شده و در قسمت (ب) اثر پرسپکتیو از روی تصاویر برداشته شده است. در قسمت (پ) لبه‌های تصاویر جدا شده و در قسمت (ت) خطوط مستقیم با هم ترکیب شده‌اند. و در نهایت در قسمت (ث) خطوط تشخیص داده شده بر روی تصویر جاده نشان داده شده‌اند. این روش نسبت به نویز حساس است به‌عنوان مثال ممکن است لبه‌های سایه قوی‌تر از لبه‌های جاده باشند و تشخیص خطوط جاده را با مشکل روبه‌رو کنند.



شکل ۲-۳: نتایج روش GOLD [۶]

در [۸] برای تشخیص مسیر ابتدا تصویر رنگی RGB به تصویر خاکستری تبدیل می‌شود و سپس ناحیه مورد علاقه^۱ تعیین می‌گردد. در ادامه با استفاده از الگوریتم تشخیص لبه کنی^۲، لبه‌ها تشخیص داده می‌شوند و با کمک تبدیل هاف خطوط مسیر به دست می‌آیند. این روش بر روی سیستم عامل اندروید بر اساس کتابخانه OpenCV^۳ شبیه‌سازی شده است. نتیجه این پیاده‌سازی بلادرنگ نیست.

وانگ و هانگ در [۹]، ابتدا از یک مبدل رنگ RGB به خاکستری استفاده شده است. سپس از یک فیلتر دوبعدی برای تشخیص لبه استفاده می‌شود و با روش آستانه‌گذاری اُتسو تصویر حاصل باینری می‌شود و در نهایت از تبدیل هاف برای تشخیص خطوط جاده استفاده می‌شود. این روش نیازمند محاسبات زیادی در پردازش تصویر ویدیویی بلادرنگ است. برای استفاده بلادرنگ، این روش بر روی یک تراشه FPGA با ماژول متصل به دوربین FMC DV1 پیاده‌سازی شده که ماکزیمم فرکانس حدود ۱۰۸/۷ مگاهرتز دارد و برای کاربردهای عملی مناسب است. ضعفی که می‌توان برای موارد [۸] و [۹] بیان نمود، عدم پویایی انتخاب ناحیه مورد علاقه است.

در [۱۰]، الگوریتم تشخیص مسیر از روش تخمین نقطه محوشدگی^۴ استفاده می‌کند. اولین مرحله این الگوریتم تشخیص قطعات خط از تصاویر است. به دلیل اینکه تبدیل هاف از نظر محاسباتی سنگین است و خطوط را هم به فرم قطعه استخراج نمی‌کند، از الگوریتم EDLines جهت تشخیص قطعه خطوط استفاده می‌شود. این الگوریتم قطعه‌های خطوط را تشخیص می‌دهد و شامل مراحل تشخیص لبه و روش برازش خط با حداقل مربعات است. با توجه به شکل ۲-۴، زوایای خطوط سمت راست جاده در محدوده ۱۰۵ تا ۱۵۰ درجه و زوایای خطوط سمت چپ جاده در محدوده ۳۰ تا ۷۵ درجه است. در این مرحله خطوطی که زاویه آن‌ها در این محدوده‌ها نباشند حذف می‌شوند. این مرحله باعث می‌شود که زمان پردازش مراحل بعدی کم شود.

¹ Region of interest(ROI)

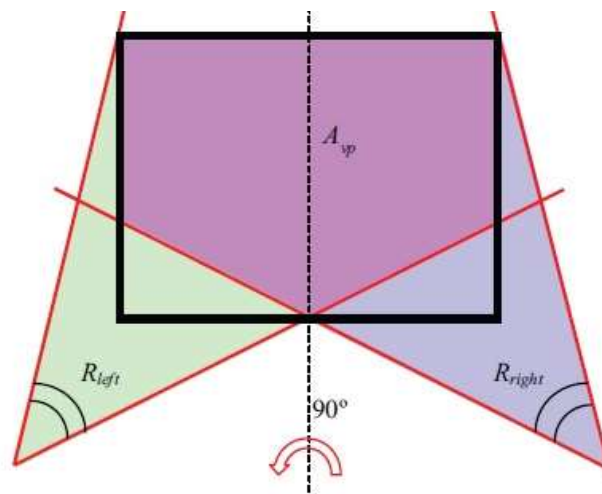
² Canny edge detection

³ Open Source Computer Vision Library

⁴ Vanishing point

سپس با استفاده از خطوط باقی مانده از مرحله قبل، مکان نقطه محوشدگی تخمین زده می شود. در شکل ۲-۴ مکان این نقطه (A_{vp})، نشان داده شده است که از محل تقاطع محدوده زاویه های خطوط سمت راست و چپ به دست می آید.

با توجه به اینکه خطوط خط کشی جاده همدیگر را در نقطه محوشدگی قطع کرده و همگرا می شوند، بنابراین خطوطی که آن نقطه را قطع نکنند حذف می شوند. در نهایت خطوط باقی مانده بر اساس زاویه شان کلاسه بندی^۱ می شوند. به طوری که اختلاف زاویه خطوط مربوط به یک کلاس از ۱۰ درجه کمتر است. سپس از هر کلاس خطی رسم می شود که یک نقطه از آن خط، نقطه محوشدگی و نقطه دیگر آن مرکز ثقل قطعه خطوط است. هر کلاس معرف یک خط خط کشی است. دو خط نزدیک به کناره های راست و چپ وسیله نقلیه، خطوط خط کشی مسیر خودرو هستند. این الگوریتم با دقت بالای ۹۰٪ کار می کند.



شکل ۲-۴: مکان نقطه محوشدگی [۱۰]

رحیمی در [۱۱]، سیستم هشدار انحراف از مسیری ارائه داده است، که امکان پیاده سازی بر روی پردازشگر ضعیف مانند تبلت و موبایل را دارد. در این سیستم بزرگراه ها و جاده های مستقیم و خلوت مورد بررسی قرار گرفته و فرض شده که تمامی جاده ها خطوط خط کشی دارند. با توجه به اینکه موقعیت دوربین تغییر نمی کند،

¹ Clustering

خطوط خط‌کشی راست و چپ جاده در نواحی مشخصی در تصاویر قرار دارند. بنابراین پردازش به این نواحی محدود می‌شود تا سرعت الگوریتم افزایش یابد. سپس با استفاده از الگوریتم تشخیص لبه کنی، لبه‌ها را شناسایی کرده و مولفه‌های بهم پیوسته استخراج می‌شوند. با استفاده از زاویه، موقعیت و شدت روشنایی مؤلفه‌های بهم پیوسته، خطوط خط‌کشی راست و چپ انتخاب می‌شود. برای تشخیص انحراف از مسیر، از حضور و عدم حضور خطوط خط‌کشی در فریم فعلی و فریم‌های قبلی آن استفاده می‌شود. این سیستم به صورت بلادرنگ (۴ فریم بر ثانیه) بر روی تبلت پیاده‌سازی شده است.

کاظمی [۱۲]، سیستم تشخیص و شناسایی جاده و خطوط موازی سطح جاده در تصاویر ویدیویی را ارائه می‌دهد. در این روش ابتدا ناحیه جاده از بقیه تصویر جدا شده و شناسایی خطوط جاده روی این ناحیه انجام می‌گیرد. با توجه به انواع جاده‌ها از یکی از دو ویژگی رنگ و یا لبه جاده‌ها برای شناسایی جاده استفاده می‌شود. سپس خطوط موازی کشیده شده روی جاده‌ها با استفاده از ویژگی‌های رنگ، لبه، شکل هندسی و امتدادشان شناسایی می‌شوند. تبدیل هاف برای علامت‌گذاری این خطوط به کاررفته است. در جاده‌های پیچ‌درپیچ با خطوط انحنادار ابتدا هیستوگرام و کنتراست تصویر تنظیم شده و بهبود می‌یابد سپس از دو ویژگی رنگ و شکل هندسی خطوط برای تشخیص آن‌ها استفاده می‌شود. این سیستم بر روی پایگاه‌های داده‌ی مختلفی پیاده‌سازی شده و نتایج مطلوبی دربرداشته است.

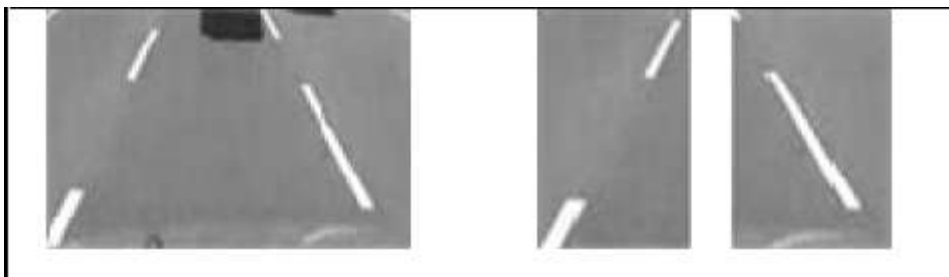
در [۱۳]، از اصول موازی کردن وظایف و داده‌ها در سیستم هشدار انحراف از مسیر جاده، با استفاده از موبایل انجام شده است. موبایل‌های چند هسته‌ای عمل‌کرد و مصرف توان را بهینه می‌کنند. در این مرجع از یک روش برنامه‌نویسی موازی جهت تشخیص خطوط و هشدار به راننده بر روی موبایل، با استفاده از یک سیستم چهار هسته‌ای ۱/۴ گیگا هرتز و ۴ RAM گیگابایتی استفاده شده است. در این روش برای تشخیص خطوط از روش استخراج رنگ استفاده شده است. برای تشخیص انحراف از مسیر، از روش جستجو کردن خط در تصویر استفاده می‌شود و در صورت نبود خط، سیستم به راننده هشدار می‌دهد. در این مرجع برای موازی‌سازی، تصویر به چهار زیر تصویر تقسیم می‌شود و هر زیر تصویر به یک هسته جهت پردازش فرستاده می‌شود. این الگوریتم بر روی

تصویر با رزولوشن 240×320 به صورت بلادرنگ پیاده سازی شده است. در این روش باید خطوط جاده به خوبی رنگ آمیزی شده باشند.

یو و همکاران [۱۴]، از تخمین نقطه محوشدگی جهت تشخیص خطوط جاده استفاده می کنند. در این روش یک دوربین در وسط قسمت جلوی خودرو نصب شده است و فرض می شود خطوط جاده برای فاصله کوتاهی از جلوی دوربین مستقیم است. در این روش ابتدا پاره خط های تصویر با استفاده از الگوریتم LSD^1 [۱۵] استخراج می شوند و قدرت هر پاره خط به صورت نسبت طول به عرض پاره خط محاسبه می شود، سپس با استفاده از دستورالعمل رأی گیری احتمالی از قدرت پاره خط های تشخیص داده شده، نقطه محوشدگی تخمین زده می شود. در ادامه خطوطی که از نقطه محوشدگی عبور می کنند مشخص می شوند و با توجه به زاویه و فاصله از وسط تصویر، خطوط جاده انتخاب می شوند.

در [۱۶]، ابتدا با استفاده از یک الگوریتم افزایش کیفیت، کیفیت تصویر اولیه افزایش می یابد تا خطوط جاده بهتر تشخیص داده شوند. سپس تصویر به دو ناحیه دلخواه تقسیم می شود به این صورت که خطوط سمت چپ و راست جاده در ناحیه های جداگانه قرار گیرند. در شکل ۲-۵ دو ناحیه دلخواه از تصویر جاده مشخص شده است. در ادامه دو ناحیه دلخواه انتخاب شده از حالت RGB به خاکستری تبدیل می شوند و برای کاهش نرخ خطای تشخیص لبه، فیلتر گوسین به تصاویر خاکستری اعمال می شود. سپس با استفاده از تشخیص لبه کنی، لبه های دو ناحیه انتخاب شده استخراج می شوند و به کمک تبدیل هاف خطوط سمت چپ و راست جاده انتخاب می شوند. این روش در مورد جاده های با پیچ های تند عمل کرد خوبی ندارد.

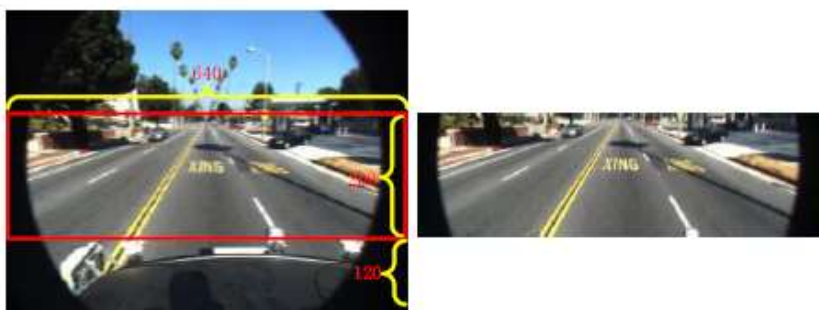
¹ Line Segment Detector



شکل ۲-۵: تقسیم تصویر جاده به دو ناحیه [۱۶]

در [۱۷]، ابتدا تصویر ورودی از حالت پرسپکتیو خارج می‌شود تا خطوط جاده به صورت موازی با یکدیگر قرار گیرند. سپس تصویر به دست آمده از حالت رنگی به حالت خاکستری تبدیل می‌شود و لبه‌های تصویر خاکستری با استفاده از الگوریتم تشخیص لبه [۱۸]، استخراج می‌شوند و با استفاده از عملگر فراخ شدن مورفولوژی نویزهای کوچک تصویر حذف می‌شوند. سپس با استفاده از تبدیل هاف خطوط جاده مشخص می‌شوند و با خطوط فریم قبلی تصویر مقایسه می‌شوند اگر متناسب بودند حفظ می‌شوند و در غیراین صورت خطوط بعدی تبدیل هاف بررسی می‌شوند.

چانگ‌زنگ و همکاران [۱۹]، با توجه به تصاویر ورودی، ناحیه وسط تصویر که مرتبط با سطح جاده است را به عنوان ناحیه مورد علاقه انتخاب می‌کنند. شکل ۲-۶ ناحیه انتخاب شده در این روش را نشان می‌دهد.



شکل ۲-۶: انتخاب ناحیه مورد علاقه از تصویر ورودی [۱۹]

سپس تصویر به دست آمده از حالت رنگی به خاکستری تبدیل می‌شود و لبه‌های تصویر توسط آشکارساز لبه کنی استخراج می‌شوند. در ادامه جهت لبه‌های تصویر محاسبه می‌شود و لبه‌های نویزی که جهت‌های غیر

متعارف دارند حذف می‌شوند. سپس از الگوریتم دسته‌بندی K-means برای انتخاب خطوط جاده استفاده می‌شود. در این مرجع از فیلتر کالمن برای دنبال کردن خطوط جاده استفاده می‌شود.

در [۲۰]، ابتدا تصویر از حالت رنگی به خاکستری تبدیل می‌شود سپس قسمت پایین تصویر به عنوان ناحیه مورد علاقه از تصویر جهت کاهش محاسبات انتخاب می‌شود. شکل ۲-۷ ناحیه مورد علاقه انتخاب شده از تصویر ورودی را نشان می‌دهد. سپس با استفاده از ترفندهای پیش پردازش تصویر مانند خوردگی^۱، فراخ‌شدن^۲ و هموارسازی^۳، اطلاعات تصویر بهبود یافته و نویزها حذف می‌شوند. سپس لبه‌های تصویر به کمک تشخیص لبه کنی شناسایی می‌شوند و تبدیل هاف جهت اتصال پاره‌خطها به یکدیگر استفاده می‌شود. در ادامه خطوط مشابه با یکدیگر گروه بندی می‌شوند و از آنها میانگین گرفته می‌شود سپس بهترین خطوط سمت چپ و راست تصویر به عنوان خطوط جاده انتخاب می‌شوند.



شکل ۲-۷: انتخاب ناحیه مورد علاقه از تصویر ورودی [۲۰]

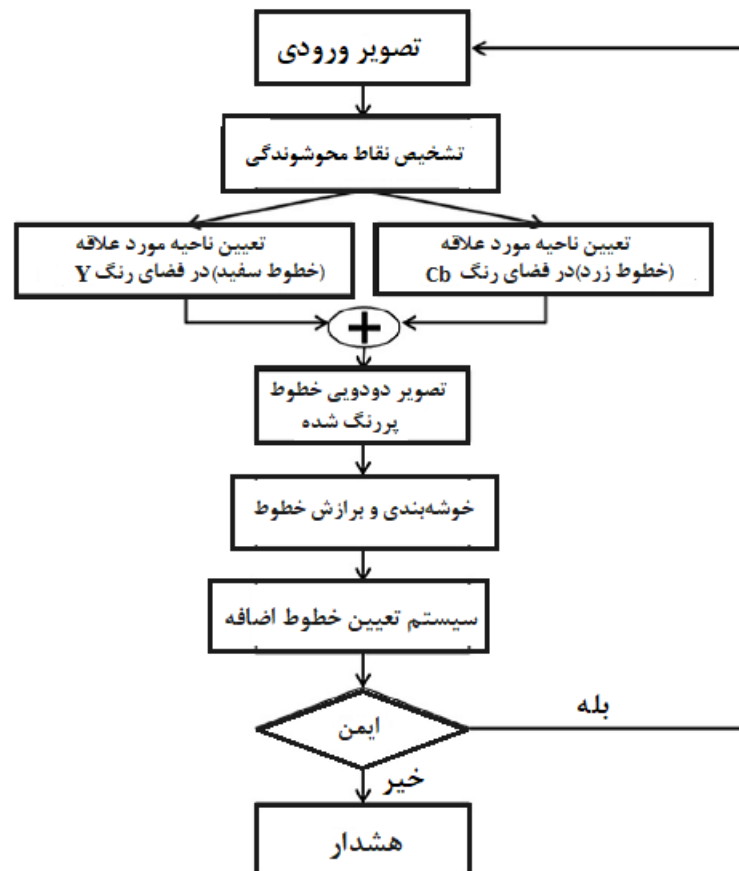
توسط سان و همکاران [۲۱]، یک روش تشخیص خطوط جاده بلادرنگ و مقاوم نسبت به تغییرات روشنایی ارائه شده است. فلوچارت روش پیشنهادی در شکل ۲-۸ نشان داده شده است. در این روش ابتدا با استفاده از تشخیص لبه کنی لبه‌های تصویر استخراج می‌شوند و به کمک تبدیل هاف پاره خطهای تصویر به‌دست می‌آیند و محل برخورد خطوط شناسایی شده محاسبه می‌شود و با استفاده از یک طرح رای‌گیری، مرکز ناحیه‌ای که بیشترین رای را دارد به عنوان نقطه محوشدگی انتخاب می‌شود.

¹ Erosion

² Warping

³ Smoothing

بعد از تعیین نقطه محوشدگی برای حذف نواحی غیرمفید در تعیین خطوط جاده، مطابق شکل ۲-۹ ناحیه پایین نقطه محوشدگی در تصویر به عنوان ناحیه مورد علاقه انتخاب می‌شود. با انتخاب ناحیه مورد علاقه زمان پردازش تصویر کاهش می‌یابد. برای اینکه الگوریتم پیشنهادی نسبت به تغییرات روشنایی مقاوم باشد از این نکته که خطوط با رنگ‌های مختلف ویژگی‌های مشخصی دارند استفاده می‌شود و خطوط زرد و سفید جاده به صورت جداگانه با توجه به طیف رنگی آنها در فضای رنگ YCbCr تعیین می‌شوند.

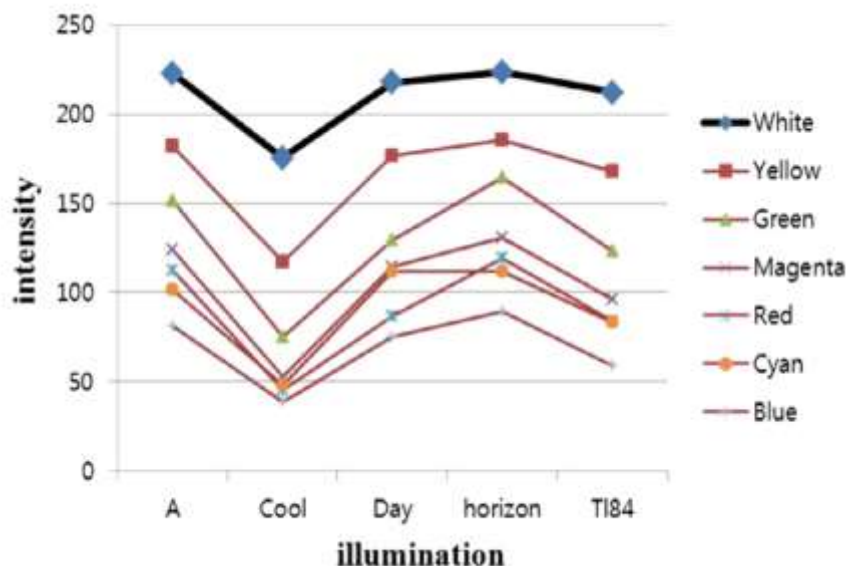


شکل ۲-۸: فلوچارت روش پیشنهادی در مرجع [۲۱]



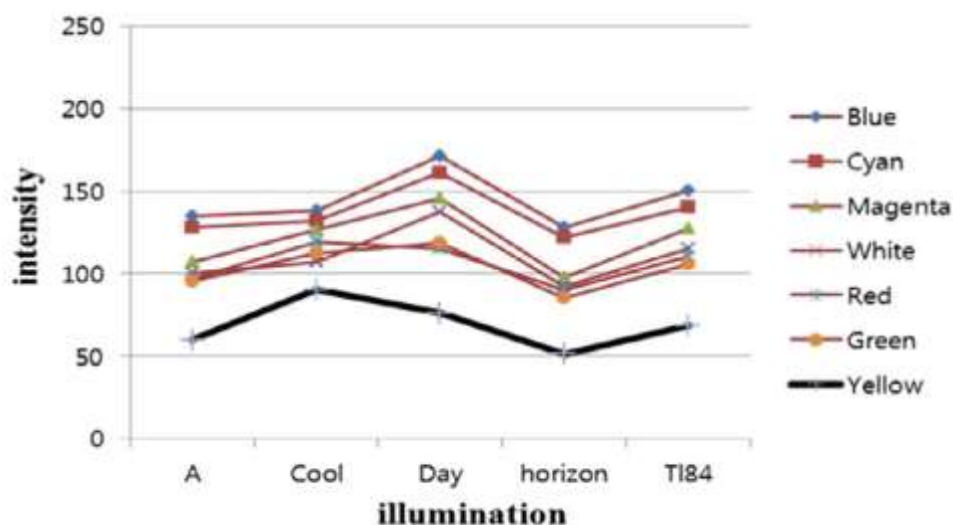
شکل ۲- ۹: تعیین ناحیه مورد علاقه توسط نقطه محوشدگی [۲۱]

با توجه به شکل ۲- ۱۰ مشاهده می‌شود مقادیر کانال Y در فضای رنگ YCbCr برای رنگ سفید در شرایط روشنایی مختلف نسبت به رنگ‌های دیگر بیشتر است بنابراین بر اساس این ویژگی به راحتی می‌توان خطوط سفید جاده را تعیین کرد. در مورد خطوط زرد رنگ در تصویر نیز با توجه به شکل ۲- ۱۱ مشخص است که مقادیر کانال Cb در فضای رنگ YCbCr برای رنگ زرد در شرایط روشنایی مختلف نسبت به رنگ‌های دیگر کمتر است بنابراین با استفاده از این ویژگی می‌توان به راحتی نواحی مربوط به خطوط زرد جاده را نیز مشخص نمود.



شکل ۲- ۱۰: مقادیر کانال Y برای رنگ‌های مختلف در شرایط روشنایی مختلف [۲۱]

سپس با تعیین مقدار یک برای ناحیه‌های مربوط به خطوط جاده و مقدار صفر برای بقیه نواحی تصویر، تصویر باینری ایجاد می‌شود. تصویر باینری نهایی با ترکیب تصویر مربوط به خطوط زرد و سفید به دست می‌آید. سپس با استفاده از روش دسته‌بندی اجزاء متصل به هم^۱ در تصویر باینری، ناحیه‌های خطوط مشخص شده و شماره گذاری می‌شوند و مرکز هر ناحیه محاسبه می‌شود. همچنین زاویه هر ناحیه و محل برخورد با محور Y نیز محاسبه می‌شود. ناحیه‌هایی که زاویه مشابه و محل تقاطع یکسانی داشته باشند با هم ترکیب شده و یک ناحیه را تشکیل می‌دهند و در نهایت خطوط چپ و راست جاده در تصویر تعیین می‌شوند.



شکل ۲-۱۱: مقادیر کانال Cb برای رنگ‌های مختلف در شرایط روشنایی مختلف [۲۱]

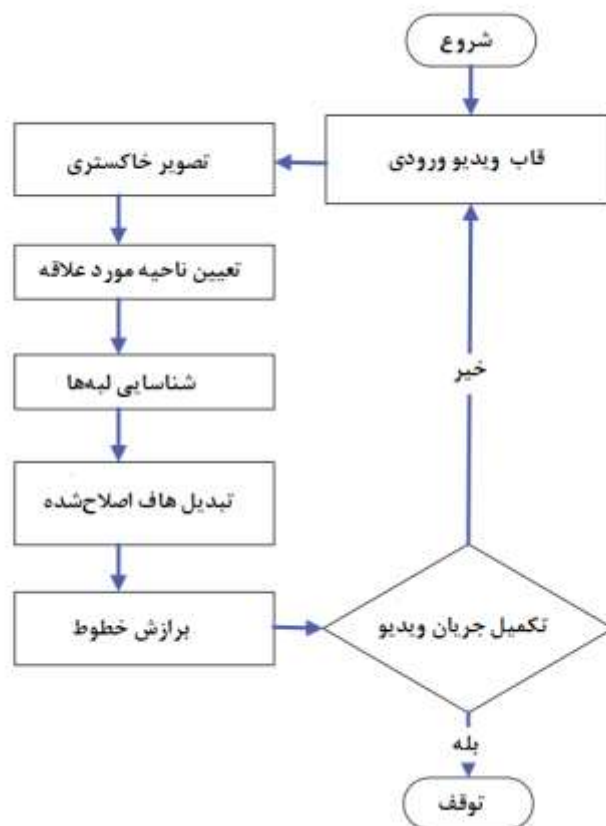
در [۲۲]، یک روش مقاوم تشخیص خطوط جاده بر مبنای تخمین نقطه محوشدگی پیشنهاد شده است که از فرایند رأی‌گیری احتمالی بر مبنای نقاط برخورد پاره‌خط‌های استخراج شده از تصویر ورودی استفاده می‌کند. در این روش ابتدا با استفاده از الگوریتم LSD [۱۵]، پاره‌خط‌های تصویر ورودی استخراج می‌شوند و قدرت هر پاره‌خط، نسبت طول هر پاره‌خط به عرض آن، محاسبه می‌شود. سپس نقطه محوشدگی خطوط با استفاده از یک فرایند رأی‌گیری احتمالی از قدرت پاره‌خط‌ها تخمین زده می‌شود. در ادامه پاره‌خط‌هایی که در نقطه محوشدگی همگرا نمی‌شوند حذف می‌شوند همچنین پاره‌خط‌هایی که جهت آنها در جهت خطوط جاده نیست

¹ Components connected together

نیز حذف می‌شوند و پاره‌خط‌های باقیمانده به عنوان پاره‌خط‌های کاندید خطوط جاده انتخاب می‌شوند و در نهایت پاره‌خط‌هایی که بیشترین شباهت را به خطوط جاده دارند به عنوان خطوط جاده انتخاب می‌شوند.

روش پیشنهادی در [۲۳]، از تبدیل هاف و فاصله اقلیدسی استفاده می‌کند. ابتدا تصویر از فضای RGB به فضای YCbCr منتقل می‌شود و با توجه به این نکته که سیستم بینایی انسان به روشنایی حساسیت بیشتری دارد فقط جزء Y تصویر برای تشخیص خطوط استفاده می‌شود. برای بهبود سرعت و دقت سیستم قسمت پایین تصویر که مربوط به جاده است به عنوان ناحیه مورد علاقه انتخاب می‌شود و برای افزایش کنتراست بین سطح جاده و خطوط جاده از برابری هیستوگرام در ناحیه مورد علاقه استفاده می‌شود و توسط آستانه‌گذاری اُتسو تصویر باینری به دست می‌آید. سپس ناحیه مورد علاقه به دو زیر ناحیه چپ و راست تقسیم می‌شود و برای هر زیر ناحیه لبه یاب کنی و تبدیل هاف به کار گرفته می‌شوند تا خطوط چپ و راست جاده به صورت جداگانه شناسایی شوند.

در [۲۴]، به دلیل نیاز به زمان زیاد جهت انجام محاسبات تبدیل هاف معمولی از تبدیل هاف اصلاح شده [۲۴] استفاده می‌شود. بلوک دیاگرام روش پیشنهادی در شکل ۲-۱۲ نشان داده شده است. در این روش ابتدا تصویر از حالت RGB به خاکستری تبدیل می‌شود سپس یک سوم قسمت پایین تصویر که شامل جاده است به عنوان ناحیه مورد علاقه انتخاب می‌شود و لبه‌های تصویر مشخص می‌شوند و با استفاده از تبدیل هاف بهبود یافته خطوط جاده تعیین می‌شوند.



شکل ۲- ۱۲: فلوچارت روش پیشنهادی در مرجع [۲۴]

لطفی و همکاران [۲۵]، ابتدا تصویر را از حالت پرسپکتیو حذف می‌کنند و تصویر جاده در جلوی خودرو از دید پرنده به دست می‌آید که خطوط جاده را به صورت موازی تبدیل می‌کند تا فرایند تشخیص خطوط را ساده‌تر کند. در ادامه برای تشخیص لبه‌های تصویر از دو فیلتر یک بعدی گوسی و لاپلاسین استفاده شده است سپس با استفاده از آستانه‌گذاری اُتسو تصویر باینری به دست می‌آید و به کمک تبدیل هاف ساده شده خطوط تصویر به دست می‌آیند. در ادامه با استفاده از یکسری خطوط افقی و نقاط برخورد آنها با خطوط تصویر و به کمک الگوریتم حداقل مربع میانگین^۱، خطوط اصلی جاده به دست می‌آیند.

در [۲۶]، ابتدا ناحیه زیر خط محوشدگی به عنوان ناحیه مورد علاقه انتخاب می‌شود سپس تصویر از حالت پرسپکتیو خارج می‌شود و در مرحله بعد با استفاده از آستانه‌گذاری و فیلترینگ روی تصویر به دست آمده،

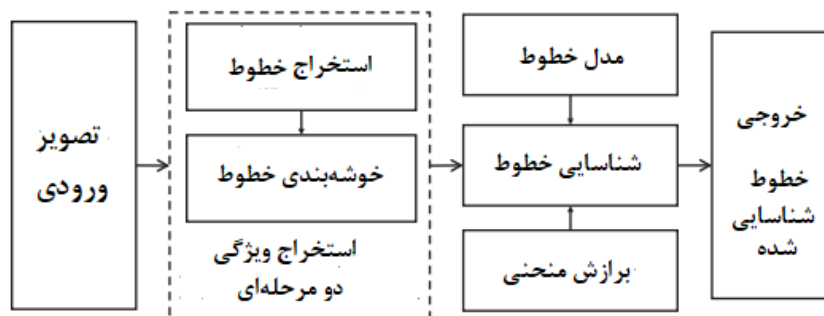
¹ Least Mean Square Algorithm

کاندیدهای خطوط جاده مشخص می‌شوند و در نهایت با استفاده از الگوریتم RANSAC پارامترهای مدل تخمین زده شده و خطوط جاده تعیین می‌شوند.

نیو و همکاران [۲۷]، از استخراج ویژگی دو مرحله‌ای جهت تشخیص خطوط جاده استفاده می‌کنند. در این روش برای استخراج ویژگی ابتدا تصویر به حالت خاکستری تبدیل می‌شود و با استفاده از الگوریتم لبه‌یاب کنی لبه‌های تصویر به دست آمده و سپس با استفاده از تبدیل هاف بهبود یافته پاره خط‌های تصویر استخراج می‌شوند. در مرحله دوم با استفاده از الگوریتم کلاسه‌بندی DBSCAN پاره خط‌های به دست آمده در مرحله اول کلاسه‌بندی می‌شوند. بلوک دیاگرام این روش در شکل ۲-۱۳ نشان داده شده است. بعد از استخراج ویژگی جهت تشخیص خطوط جاده از دو فرایند زیر استفاده می‌شود:

۱. گروه‌های کاندید که دارای احتمال تشابه بالا با استفاده از خصوصیات فضایی خطوط جاده دارند انتخاب می‌شوند.

۲. خطوط نهایی جاده به وسیله ترکیب با نقطه محوشدگی به دست می‌آیند.



شکل ۲-۱۳: بلوک دیاگرام روش پیشنهادی در مرجع [۲۷]

در [۲۸]، جهت تشخیص خطوط جاده ابتدا تصویر از حالت RGB به خاکستری تبدیل می‌شود سپس با استفاده از فیلتر گوسی تصویر هموار می‌شود و با استفاده از لبه‌یاب کنی لبه‌های تصویر به دست می‌آیند، در نهایت خطوط مستقیم با استفاده از تبدیل هاف محاسبه می‌شوند. سپس با اعمال محدودیت‌هایی خطوط جاده مشخص می‌شوند، به این صورت که ابتدا شیب خطوط به دست آمده محاسبه می‌شوند و خطوط با شیب کم که مربوط به خطوط افقی هستند حذف می‌شوند. سپس براساس شیب و نقطه میانی خطوط، آن‌ها به دو دسته

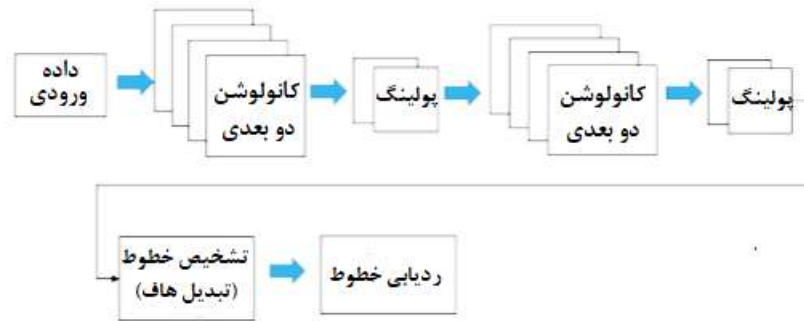
خطوط چپ و راست تقسیم می‌شوند و نزدیکترین خطوط از خط محور تصویر به عنوان کاندیدای خطوط چپ و راست انتخاب می‌شوند.

۲-۲-۲ تشخیص خطوط جاده بر اساس رویکرد مبتنی بر یادگیری عمیق

در پژوهش‌هایی که تشخیص خطوط جاده را با استفاده از روش‌های مبتنی بر یادگیری عمیق انجام می‌دهند، از نظر ماهیت کاربردی، می‌توان شناسایی خطوط و جاده را در سه سناریوی مختلف بیان نمود [۲۹]. اولین مورد شناسایی تنها مرز مسیر حرکت خود راننده است (یک راستا از جاده مورد نظر) [۳۰-۳۲]. دومین مورد شناسایی تعداد راستاهای حرکتی موجود در یک جاده است [۳۳-۳۵] و نهایتاً سومین مورد شناسایی تمام جاده موجود در تصویر است [۳۶ و ۳۷] (که ممکن است شامل دو باند حرکتی متفاوت باشد) که نسبت به دو مورد قبلی دارای چالش‌های بیشتری می‌تواند باشد. در ادامه به تعدادی از پژوهش‌های صورت گرفته مبتنی بر یادگیری عمیق خواهیم پرداخت.

ساتیش کومار و همکاران در [۳۸]، با استفاده از یک شبکه عصبی کانولوشن و الگوریتم لبه‌یابی سوبل در ادامه آن با استفاده از تبدیل هاف اقدام به شناسایی خطوط در وضعیت‌های متفاوت همچون روز، محیط‌های کم‌نور و شب نمودند. این افراد در فرایند تحقیقاتی خود از یک رویکرد نظارت‌شده با یک مجموعه داده ۴۰۰۰ تایی برچسب‌گذاری‌شده و همچنین برای بهبود روند آموزشی از یک عملیات افزایش داده‌ها^۱ برای افزایش نمونه داده استفاده نمودند. نتایج حاصل از پژوهش انجام گرفته دقت بالای روش ارائه شده آن‌ها در دوحالت به‌کارگیری افزایش داده‌ها و عدم به‌کارگیری آن در تمام وضعیت‌های آزمایشی را نشان می‌دهد. شکل‌های ۲-۱۴ و ۲-۱۵ کلیت روش پیشنهادی و همچنین خروجی عملیاتی روش به‌کار گرفته‌شده را برای آنان نشان می‌دهد.

¹ Data Augmentation



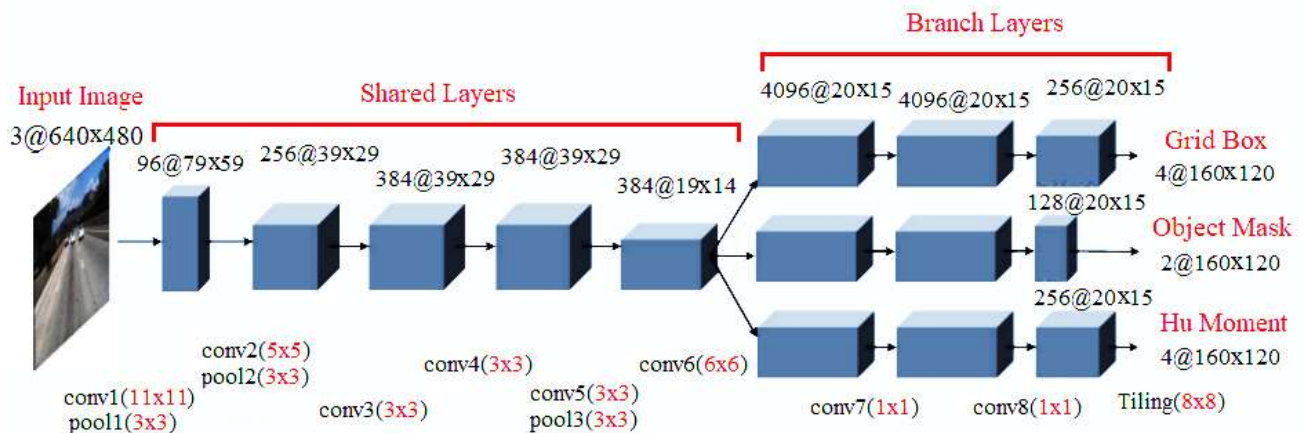
شکل ۲-۱۴: شبکه عمیق به کار گرفته شده در روش پیشنهادی مرجع [۳۸]



شکل ۲-۱۵: رویکرد الگوریتم پیشنهادی در پژوهش [۳۸]

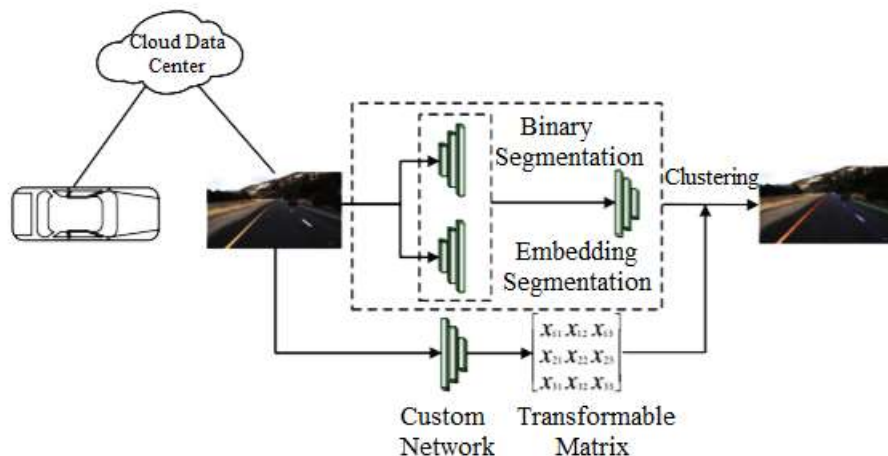
فعالیت دیگری که می‌توان به آن اشاره نمود، کاری است که هائو لیو و لیانگ لی [۳۹]، انجام دادند. این افراد در پژوهش خود اقدام به ارائه یک روش با عنوان 3TCNN نمودند. در روش اجرایی آن‌ها، ابتدا از یک شبکه کانولوشن با پنج لایه کانولوشن استفاده شده است. سپس در انتهای لایه پنجم یک انشعاب سه‌تایی برای فرایند تشخیص خطوط، شناسایی جاده و پیش‌بینی مسیر با استفاده از یک فرایند رگرسیون صورت می‌پذیرد. در واقع عملیات رگرسیون انجام گرفته در فعالیت انجام شده به منظور تخمین خطوط مسیر در زمان‌هایی است که خطوط جاده در تصویر به واسطه حرکت خودروها پوشانده می‌شود. از نکات قابل توجه روش ارائه شده در این

پژوهش، دقت برابر با ۹۴ درصد ارائه شده است. شکل ۲-۱۶ کلیت دیاگرام روش پیشنهادی و همچنین خروجی پژوهش مورد نظر نشان داده شده است.



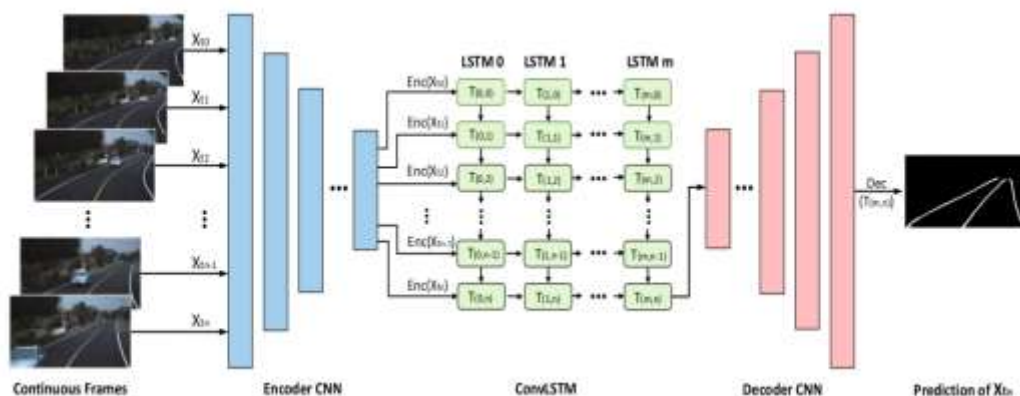
شکل ۲-۱۶: معماری شبکه در روش پیشنهادی مرجع [۳۹]

فعالیت دیگری که می‌توان به آن اشاره نمود، پژوهشی است که توسط وانگوی و همکارانش [۴۰]، انجام گرفت. این افراد در فعالیت خود از یک شبکه کانولوشن برای ایجاد یک مدل دوگانه قطعه‌بندی استفاده نمودند. ایده دیگری که محققان این پژوهش از آن استفاده نمودند، به‌کارگیری تکنیک محاسبات ابری در قطعه‌بندی و شناسایی یال‌ها برای بهبود از لحاظ کاهش محاسبات حاصل از عملیات آموزش است. با استفاده از شبکه عمیق کانولوشن و همچنین دو مدل قطعه‌بندی مشخص به‌کار گرفته‌شده در فعالیت پیش‌روی، در مرحله آخر کار شناسایی خطوط، با استفاده از ماتریس انتقال ایجادشده از مرحله قبل اقدام به شناسایی خطوط گردیده‌شد. از نکات پراهمیتی که می‌توان در فعالیت صورت گرفته به آن اشاره نمود، تنوع آزمایش در نمونه‌های با شرایط مختلف بوده است. همچنین می‌توان به دقت بالای ۹۷ درصدی روش در شناسایی خطوط در حالت تصاویر با شرایط محیطی ایده‌آل و دقت بالا در حالات آزمایشی چالشی اشاره نمود. شکل ۲-۱۷ کلیت کار مربوطه را نشان می‌دهد.



شکل ۲-۱۷: ساختار کلی روش پیشنهادی در مرجع [۴۰]

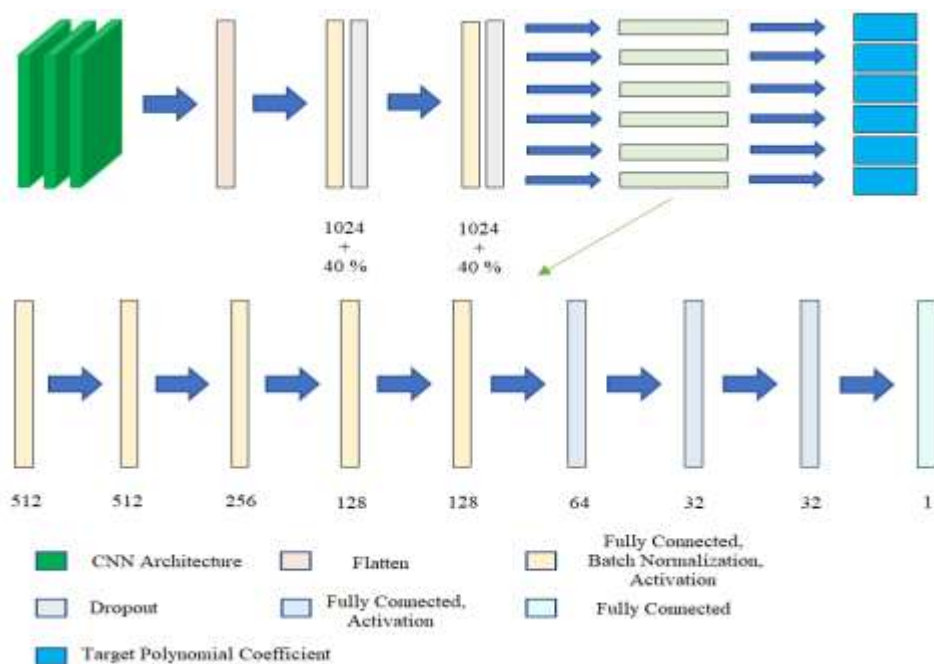
کین زو و همکارانش [۴۱]، در فعالیت خود با به کارگیری یک شبکه عمیق ترکیبی از شبکه‌های کانولوشن و RNN اقدام به شناسایی خطوط جاده نمودند. آن‌ها در فعالیت خود ابتدا با استفاده از یک شبکه کانولوشنی به به دست آوردن ویژگی‌های موجود در تصویر پرداختند. سپس با استفاده از شبکه LSTM که نوعی از شبکه‌های عمیق بازگشتی (RNN) هستند و با استفاده از ویژگی‌های استخراجی ارسال شده از شبکه‌ی کانولوشن به یادگیری سیستم و شناسایی خطوط جاده پرداختند. نکته حائز اهمیت در فعالیت این گروه استفاده از دو مجموعه داده با مقیاس و حجم بالا و آزمایش سیستم طراحی شده در شرایط چالشی ویژه بود که در هریک، نتایج به دست آمده نشان‌دهنده دقت بالای الگوریتم پیشنهادی را بیان می‌کند. شکل ۲-۱۸ معماری روش پیشنهادی را نشان می‌دهد.



شکل ۲-۱۸: معماری الگوریتم پیشنهادی در مرجع [۴۱]

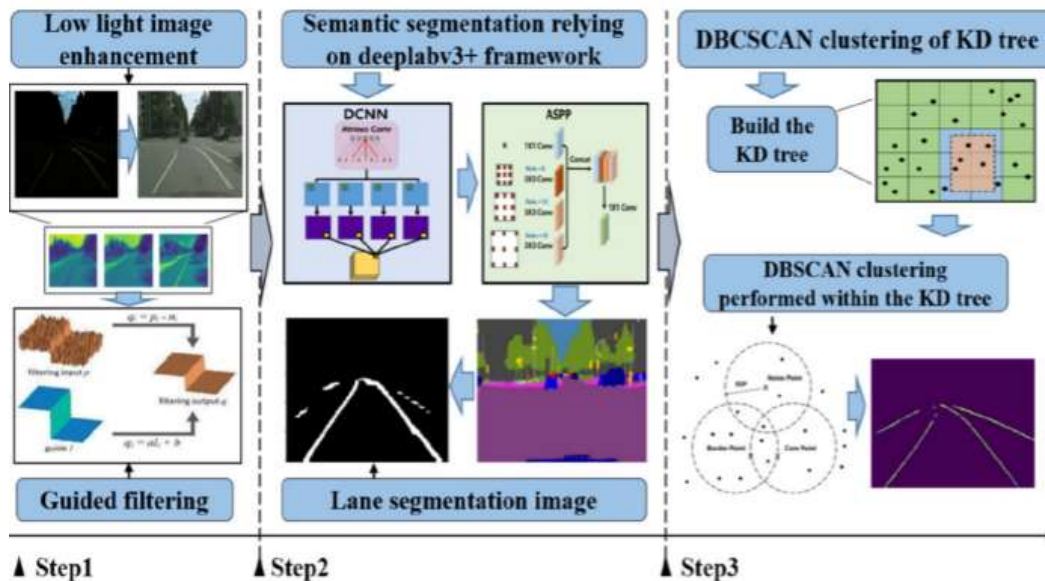
در [۴۲]، عثمان ایکسای و گوکمن با استفاده از یک شبکه عمیق ResNet50 برای عملیات شناسایی خطوط جاده از رویکرد قطعه‌بندی مبتنی بر پیکسل استفاده نمودند. آن‌ها در فعالیت خود از لایه‌های تماماً متصل، بعد از لایه‌های کانولوشن استفاده نمودند، به‌طوری که برای دو لایه اول تماماً متصل به دلیل جلوگیری از پدیده بیش‌برازش^۱ بعد هرکدام از آن‌ها یک لایه dropout با نرخ ۴۰٪ قرار دادند. از طرفی بعد از ساختار لایه‌ای نامبرده‌شده، از ۵ لایه تماماً متصل برای عملیات تخمین خطوط از چندجمله‌ای مرتبه دوم استفاده نمودند. شکل ۱۹-۲ معماری شبکه استفاده شده این فعالیت پژوهشی را نشان می‌دهد. یکی از نکات قابل توجه در پژوهش این افراد کاهش اندازه تصاویر ورودی مجموعه داده Tusimple [۴۳]، از اندازه ۱۲۰۰x۷۲۰ به اندازه ۱۶۰x۱۶۰ با حفظ رزولوشن تصاویر بود. از موارد قابل توجه دیگری که در فعالیت انجام شده می‌توان به آن اشاره نمود، جامعیت روش اجرایی با مقایسه با معماری‌های متفاوت از شبکه‌های کانولوشن همچون AlexNet، LeNet5، VGG16 و VGG19 می‌توان اشاره نمود. در نتایج حاصل، پارامتر IOU برای دادگان آزمون با استفاده از معماری موردنظر مقدار ۰٫۸۹، به‌دست آمده است.

¹ Overfitting



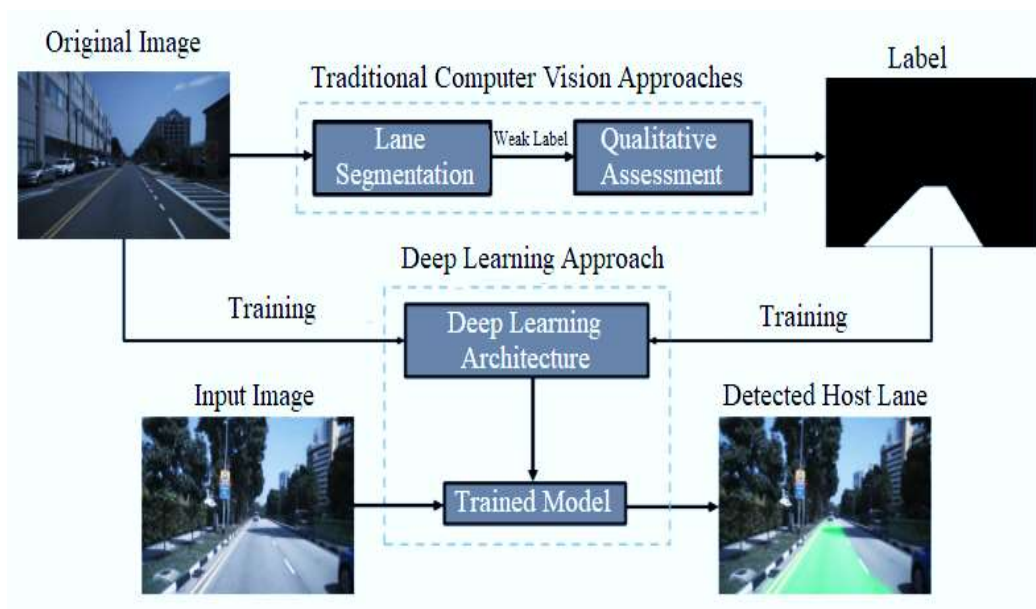
شکل ۲-۱۹: ساختار معماری شبکه استفاده شده در مرجع [۴۲]

شائون سانگ و همکاران [۴۴]، در فعالیت خود اقدام به یک شبکه عمیق تحت عنوان LLSS-Net نمودند. در معماری روش پیشنهادی این پژوهش‌گران از دو شبکه عمیق استفاده شده است. شبکه اول به منظور بهبود تصاویری که از لحاظ روشنایی دارای کیفیت پایینی هستند، عمل می‌کند و در شبکه دوم عملیات استخراج ویژگی به منظور انجام فرایند قطعه‌بندی معنایی برای شناسایی خطوط جاده صورت می‌گیرد. نهایتاً با به‌کارگیری عملیات خوشه‌بندی سریع، خطوط جاده به‌دست می‌آیند. در پژوهش انجام‌شده برای نشان دادن پایداری روش به‌کار گرفته‌شده، از دو مجموعه داده Tusimple و CitySpace [۴۵] برای آزمایش روش استفاده شده است. نتایج حاصل از به‌کارگیری روش نشان از دقت بالای روش اجرایی دارد. در شکل ۲-۲۰ دیگرام کلی روش اجرایی را نشان داده‌شده است.



شکل ۲-۲۰: فلوچارت روش پیشنهادی در مرجع [۴۴]

ریتاج یوسری وهمکاران در [۴۶]، با به کارگیری ترکیبی از الگوریتم‌های پایه‌ای پردازش تصویر به همراه شبکه‌های عمیق، اقدام به شناسایی جاده و خطوط نمودند. آن‌ها در فعالیت خود ابتدا با استفاده از یک مجموعه متوالی از دستورات پایه پردازش تصویر اقدام به تولید یک ماسک اولیه از تصاویر نمودند. سپس با به کارگیری شبکه‌های عمیق و آموزش مدل پیشنهادی اقدام به بهبود خروجی که همان آشکارسازی جاده و خطوط آن پرداختند. در فرایند مبتنی بر یادگیری عمیق از معماری‌های FCN همچون: SegNet ، SegNet اصلاح شده، Unet، ResUNet و ResUNet++ استفاده شده است. نتایج حاصل بر روی تصاویر شهری مجموعه داده به کار گرفته شده نشان از کارایی روش اجرایی در صحنه‌های متعدد با سناریوهای متفاوت دارد. همچنین در بین معماری‌های انجام شده، روش ResUNet++ نتیجه بهتری نسبت به سایر روش‌ها از خود نشان داده است. معماری کلی روش پیشنهادی در شکل ۲-۲۱ نشان داده شده است.



شکل ۲-۲۱: معماری روش پیشنهادی مرجع [۴۵]

از دیگر پژوهش‌هایی که با به‌کارگیری روش‌های مبتنی بر یادگیری عمیق اقدام به شناسایی جاده و خطوط جاده داشته‌اند، می‌توان به فعالیت‌های انجام‌شده در جدول ۲-۱ اشاره نمود.

جدول ۲-۱: تعدادی از پژوهش‌های شناسایی جاده و خطوط با استفاده از رویکرد مبتنی بر یادگیری عمیق

نویسندگان	کلیت روش اجرایی
فان چائو و همکاران [۴۷]	استفاده از شبکه عمیق کانولوشن، تبدیل هاف، الگوریتم کمترین مربعات خطا برای شناسایی خطوط جاده
نیون دایوی و همکاران [۴۸]	استفاده از شبکه عمیق LaneNet و خوشه‌بندی سریع در شناسایی خطوط جاده
ژی وانگ، ویکیانگ رن [۴۹]	استفاده از یک ساختار گذار-کدگشا با به‌کارگیری شبکه LSTM برای شناسایی خطوط جاده
ژین یو ژانگ و همکاران [۵۰]	استفاده از شبکه عمیق VGG مبتنی بر شبکه کانولوشن و استفاده از مکانیزم Self-Attention در شناسایی خطوط جاده
ژیان ژائو و همکاران [۵۱]	استفاده از Q-Learning+Faster-RCNN در شناسایی خطوط
المأمون و همکاران [۵۲]	استفاده از یادگیری عمیق و یک رمزگذار-رمزگشا بر اساس معماری SegNet
کاناگراج و همکاران [۵۳]	استفاده از یک شبکه عمیق مبتنی بر معماری LaneNet-5 برای شناسایی خطوط جاده
جیانگ ژانگ و همکاران [۵۴]	استفاده از یک شبکه کانولوشن دوگانه فضا_زمان بازگشتی، برای شناسایی خطوط در محیط‌های چالش‌برانگیز
لی مین هیوک و همکاران [۵۵]	استفاده از یک ساختار self Attention و به‌کارگیری شبکه عمیق رمزگذار رمزگشا در شناسایی خطوط

آنچه که به عنوان یک اصل در شناسایی خطوط با استفاده از شبکه‌های عمیق مطرح است، استفاده از یک ساختار کانولوشنی پایه در شبکه عمیق است.

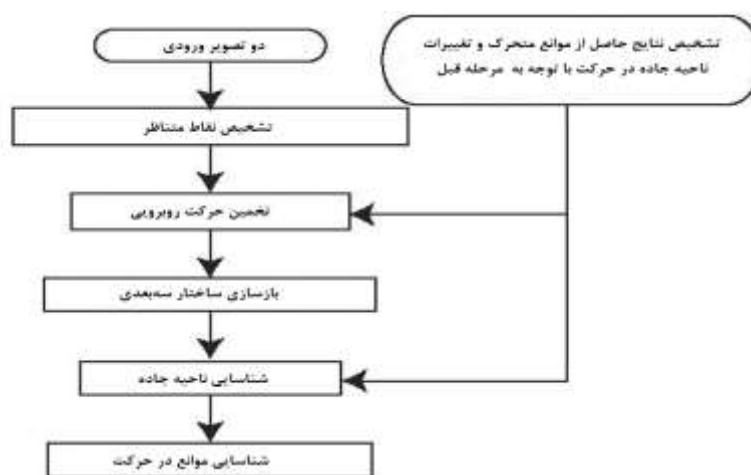
۲-۳ مروری بر پژوهش‌های پیشین در زمینه تشخیص موانع

تشخیص موانع یکی از سیستم‌های مورد نیاز در خودروهای هوشمند است، به این دلیل که در جلوگیری از تصادفات ناشی از برخورد با موانع جلوگیری می‌کند. روش‌های مبتنی بر بینایی مختلفی برای تشخیص موانع با استفاده از یک دوربین ارائه شده است. روش‌های مبتنی بر بینایی با یک دوربین به دو دسته روش‌های مبتنی بر حرکت و روش‌های مبتنی بر ظاهر تقسیم می‌شوند. در روش‌های مبتنی بر حرکت اندازه مانع باید خیلی بزرگ باشد و جابه‌جائی مانع زیاد باشد ولی در روش‌های مبتنی بر ظاهر، ویژگی‌های بافت، رنگ، قالب و لبه مورد استفاده قرار می‌گیرند. این روش‌ها دارای سرعت پردازش بالا، هزینه کمتر و قابلیت شناسایی موانع کوچک هستند. در ادامه تعدادی از روش‌های تشخیص موانع با استفاده از یک دوربین را مرور می‌کنیم.

در [۵۶]، یک روش تشخیص موانع برای روبات‌های متحرک ارائه می‌شود، که در آن هر پیکسل تصویر بر اساس رنگ آن، به طبقه مانع یا سطح زمین تقسیم می‌شود. در این روش ابتدا با استفاده از فیلتر گوسی نویز تصویر حذف می‌شود. سپس تصویر از فضای رنگ RGB به فضای HSI منتقل می‌شود. مرحله بعد یک ناحیه دوزنقه‌ای در جلوی ربات به عنوان سطح زمین مرجع انتخاب می‌شود. سپس پیکسل‌های تصویر با سطح مرجع مقایسه می‌شوند و به دو طبقه سطح زمین و مانع تقسیم می‌شوند. این روش به صورت بلادرنگ اجرا می‌شود و خروجی، تصویر مانع باینری با وضوح بالا است.

در [۵۷]، یک روش برای شناسایی موانع در حال حرکت بر روی جاده‌ها، با استفاده از یک دوربین مستقر روی خودرو پیشنهاد می‌شود. در این روش، فرض بر این است که دوربین کالیبره شده، یعنی پارامترهای داخلی دوربین شناخته شده است. در روش پیشنهادی، برای هر فریم، دو فریم متوالی مورد استفاده قرار می‌گیرد تا موانع در حال حرکت و ناحیه جاده در تصویر تشخیص داده شود. شکل ۲-۲۲ فرایند پردازش برای هر فریم را نشان می‌دهد. ابتدا، تشابه نقاط ویژگی بین دو تصویر ورودی شناسایی می‌شوند. در این روش نقاط ویژگی با

استفاده از آشکارساز گوشه هریس [۵۸]، به دست می‌آید و تشابه نقاط ویژگی به وسیله روش لوکاس-کاناد [۵۹]، محاسبه می‌شود. سپس، حرکت وسیله نقلیه با استفاده از تشابه نقاط ویژگی تخمین زده می‌شود. در ادامه، ساختار سه بعدی صحنه جاده با استفاده از تشابه نقاط ویژگی و حرکت تخمینی وسیله نقلیه بازسازی می‌شود. سپس، سطح زمین در فضای سه بعدی بازسازی شده، تخمین زده می‌شود و ناحیه جاده در تصویر تشخیص داده می‌شود. در نهایت، موانع در حال حرکت با ردیابی نقاط ویژگی روی فریم‌های متوالی تشخیص داده می‌شود. در نتیجه، موانع در حال حرکت و ناحیه جاده در فریم فعلی به دست می‌آیند. این نتایج به صورت بازگشتی برای تخمین حرکت خودرو و تشخیص ناحیه جاده در فریم بعدی استفاده می‌شود.



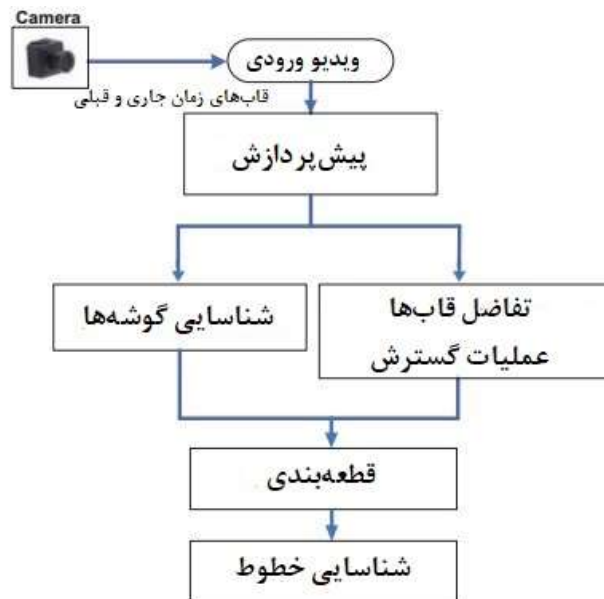
شکل ۲-۲۲: فرایند پردازش برای هر فریم در روش پیشنهادی مرجع [۵۷]

مولر و همکاران در [۶۰]، یک روش شناسایی عابرپیاده پیشنهاد کردند. در این روش ناحیه شناسایی به دو زیرناحیه، ناحیه فاصله نزدیک و ناحیه فاصله دور، تقسیم می‌شود. دو الگوریتم شناسایی مختلف در دو ناحیه به کار می‌روند. برای ناحیه با فاصله کم، جایی که مسیر حرکت مانع مشخص شده، مهم است، یک روش تقسیم-بندی حرکت با استفاده از نقاط موردنظر استفاده می‌شود. برای ناحیه با فاصله دور، جایی که حرکت مانع تشخیص داده شده مهم نیست، یک نگاشت پرسپکتیو معکوس اصلاح شده برای تشخیص موانع استفاده می‌شود.

در [۶۱]، یک الگوریتم جدید با استفاده از رنگ و بافت برای تشخیص موانع ارائه شده است. اطلاعات رنگ از لحاظ یادگیری و پردازش، دارای محاسبات کمی است. با این حال در بسیاری از موارد، رنگ به تنهایی اطلاعات کافی برای طبقه‌بندی فراهم نمی‌کند. در این حالت اطلاعات بافت می‌تواند عمل‌کرد طبقه‌بندی را بهبود دهد ولی هزینه زیادی تحمیل می‌کند. در این الگوریتم هردو ویژگی رنگ و بافت استفاده می‌شود اما زمانی ویژگی بافت نیاز است که ویژگی رنگ غیر قابل اعتماد باشد. در مرحله آموزش، ویژگی‌های بافت به‌طور خاص برای بهبود عمل‌کرد طبقه‌بند رنگ آموخته می‌شوند. الگوریتم مجموعه‌ای از ویژگی‌های بافت ساده را یاد می‌گیرد و تنها مؤثرترین ویژگی‌ها در مرحله طبقه‌بندی استفاده می‌شوند. بنابراین الگوریتم پیشنهادی درحالی‌که به اندازه کافی سریع برای اجرا در کامپیوتر است، نرخ طبقه‌بندی بسیار خوبی نیز دارد.

در [۶۲]، یک الگوریتم برای شناسایی موانع در سیستم دستیار پارکینگ پیشنهاد شده است. بلوک دیاگرام روش پیشنهادی به‌صورت شکل ۲-۲۳ می‌باشد. در مرحله پیش‌پردازش ابتدا ناحیه مورد علاقه از تصویر انتخاب می‌شود سپس عملیات پایین نمونه‌برداری برای افزایش سرعت پردازش انجام می‌شود. در ادامه برای تعیین موانع موجود، روش اختلاف فریم استفاده می‌شود. در این مرحله با استفاده از تصاویر دو فریم متوالی و آستانه‌گذاری آنها، اجزای مربوط به کاندیدهای موانع بالاتر از سطح زمین تقریباً مشخص می‌شوند. بعد، یک الگوریتم تشخیص گوشه قدرتمند به نام FAST [۶۳]، بر روی تصویر اصلی اعمال می‌شود. اگر فاصله بین گوشه‌های متعلق به یک جزء در تصویر اصلی پس از تبدیل^۱ IPM کشیده شود، این جزء به عنوان مانعی در نظر گرفته می‌شود که بالاتر از سطح زمین است. در صورتیکه موانع در منطقه هشدار شناسایی شوند، هشدار صوتی فعال خواهد شد.

¹ Inverse Perspective Mapping



شکل ۲-۲۳: بلوک دیاگرام روش پیشنهادی در مرجع [۶۲]

در [۶۴]، یک روش بلادرنگ برای تشخیص موانع در سیستم‌های کمک راننده پیشنهاد شده است که می‌تواند موانع را از سایه‌ها و علامت‌های سطح جاده جدا کند. در این روش ابتدا نقطه محوشدگی برای هر فریم محاسبه می‌شود و ناحیه پایین نقطه محوشدگی به عنوان ناحیه مورد علاقه انتخاب می‌شود. سپس نقاط ویژگی ناحیه مورد علاقه با استفاده از آشکارساز گوشه ^{1}KLT [۶۵]، محاسبه می‌شوند. در ادامه با استفاده از مدل ^{2}TCF و فیلتر اطمینان برای دو فریم متوالی موانع شناسایی می‌شوند و در نهایت یک فرایند بروزرسانی برای کاهش خطا و بروزرسانی الگوریتم وقتی خودرو حرکت می‌کند، به کار گرفته می‌شود.

در [۶۶]، یک الگوریتم جهت تشخیص موانع جاده با استفاده از یک دوربین مستقر در خودرو ارائه می‌شود. الگوریتم پیشنهادی در این مقاله دو روش آنالیز ویدیوهای دوبعدی و سه‌بعدی را باهم ترکیب می‌کند. ساختاری از تکنیک‌های حرکت و سرعت‌سنج دیداری برای تبدیل توالی تصویر دوبعدی به یک مدل ابر نقطه سه‌بعدی استفاده می‌شود. این مدل برای تخمین و شناسایی موانع با فاصله کم و متوسط از خودرو استفاده می‌شود. با استفاده از تکنیک‌های پردازش تصاویر دوبعدی مانند تشخیص لبه و بخش‌بندی رنگ، تخمین قابل قبولی از

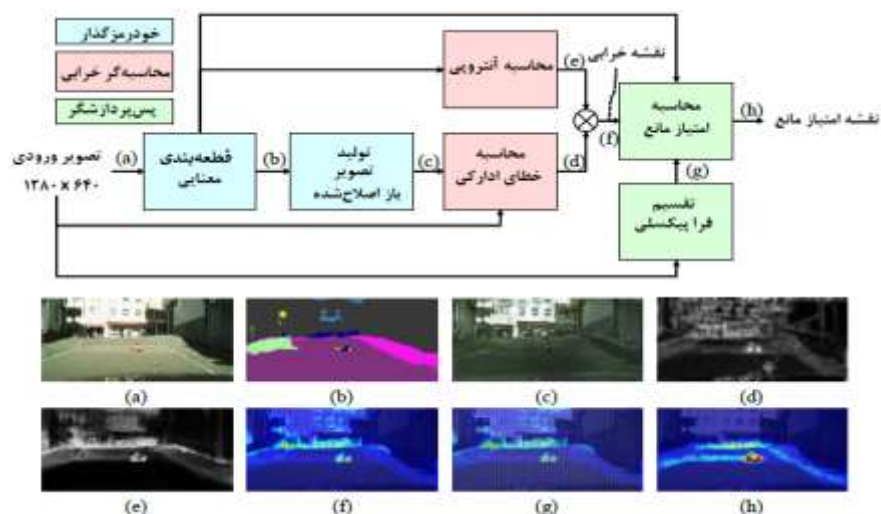
¹ Kanade-Lucas-Tomasi

² Two Consecutive Frames

اجزاء دور از خودرو ارائه می‌شود. ترکیبی از آن‌ها، هزینه‌های محاسباتی را کاهش داده و در نتیجه عملیات به- صورت بلادرنگ اجرا می‌شود.

در [۶۷]، یک روش تشخیص موانع جاده با استفاده از یادگیری عمیق پیشنهاد شده است. در این روش ابتدا با استفاده قطعه‌بندی معنایی، تصویر به ناحیه‌هایی که بیانگر اشیاء معنی‌داری هستند تقسیم می‌شود و برای پیکسل‌های متعلق به هر منطقه خاص یک برچسب اختصاص می‌یابد. قطعه‌بندی معنایی با استفاده از ابزارهای یادگیری عمیق مانند روش معمول استفاده از شبکه‌های عصبی کانولوشن انجام می‌شود. در این مقاله از الگوریتم طبقه‌بندی پیکسل SegNet استفاده شده است [۶۸]. خروجی الگوریتم SegNet تصویری است که به دوازده کلاس طبقه‌بندی شده است. در ادامه ناحیه مربوط به جاده تعیین می‌شود و با استفاده از الگوریتم جعبه محدود، محدوده نواحی مربوط به وسایل نقلیه و عابران پیاده مشخص می‌شود. سپس وسایل نقلیه و عابران پیاده‌ای که در ناحیه جاده هستند شناسایی می‌شوند.

اوگوشی و همکاران [۶۹]، با استفاده از رویکردی مبتنی بر یادگیری عمیق و با به‌کارگیری یک اتوانکدر اقدام به تشخیص موانع درون جاده در نماهای نرمال از طریق قطعه‌بندی معنایی نمودند. آن‌ها در روش پیشنهادی خود، از تصاویری که با استفاده از دوربین تعبیه‌شده درون خودرو جمع‌آوری گردیده بودند، اقدام به آموزش سیستم خود نمودند. سپس با استفاده از یک قطعه‌بند معنایی که از یک بخش کدگذار و یک بخش کدگشا برای تولید خروجی تصاویر بخش‌بندی‌شده استفاده شده بود، اقدام به شناسایی موانع نمودند. این افراد برای گسترش کارایی مدل خود، آن را بر روی مجموعه داده مشخصی که دارای موانع طبیعی بود، آموزش دادند. کلیت فرایند کار و دیاگرام کلی مدل پیشنهادی این پژوهش در شکل ۲-۲۴ قابل مشاهده است. نکته‌ای که در مورد فعالیت انجام‌شده مطرح است، انواع مختلف موانع موجود درون جاده است که صرفاً خودورهای درون جاده نیستند. در واقع موانعی هستند که از لحاظ اندازه ممکن است در تشخیص با چالش مواجه شوند. نتایج به‌دست آمده از پژوهش، بیان‌کننده دقت بالای روش پیشنهادی است.



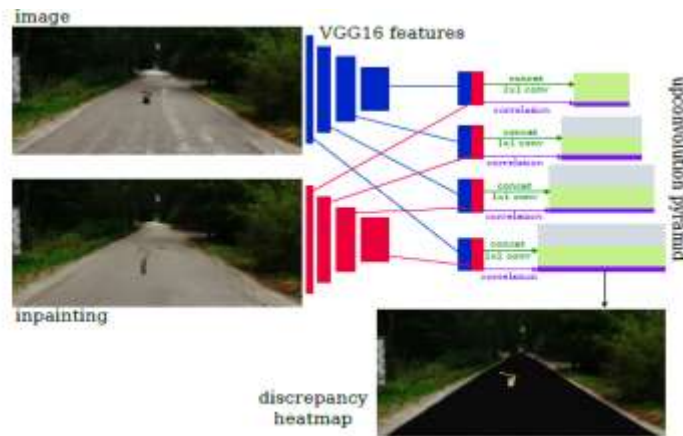
شکل ۲-۲۴: روش پیشنهادی تشخیص موانع در مرجع [۶۹]

لايز و همکاران وی [۷۰]، در فعاليت ديگري اقدام به شناسايي موانع غيرمعمول در جاده نمودند. هدف اين افراد شناسايي موانعي همچون قطعه سنگ، بطري نوشابه و در کل اشياي کوچک از اين قبيل درون جاده بود. رویکرد اصلي که آن‌ها در فعاليت خود به کار گرفته‌اند، از دو مرحله تشکيل شده است:

۱. ايجاد تصويري که در آن مانع درون تصوير پاک شده است و با استفاده از یک اسلايد مشخص نقاشي شده، تصوير بدون بازسازي شده است.

۲. تصوير اصلي با نسخه نقاشي شده مقايسه و ميزان شباهت و اختلاف آن‌ها مورد بررسي قرار داده می‌شود.

با استفاده از دو مرحله فوق، در صورتي که مانعي در تصوير وجود نداشته باشد، تصوير اصلي با تصوير نقاشي شده بسيار به همديگر نزديک هستند. بدین ترتيب با ارزيابي شبکه‌های موجود، مشخص می‌شود که مانع در تصوير وجود ندارد. از طرفي اگر در تصوير مانعي وجود داشته باشد، تصوير اصلي با تصوير نقاشي شده که مانع درون آن حذف شده، در قسمتي که مانع وجود دارد، اختلاف بالايی داشته که بيانگر مانع است. رویکرد اجرايي در شکل ۲-۲۵ قابل مشاهده است. ساختار اصلي که برای اجراي رویکرد پيشنهادهی از آن استفاده شده است، شبکه عميق VGG16 است.



شکل ۲-۲۵: روش پیشنهادی تشخیص موانع در مرجع [۷۰]

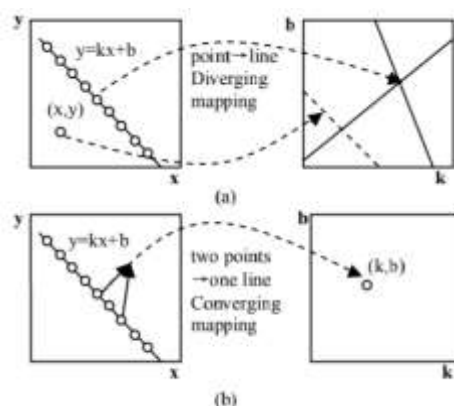
در [۷۱] با استفاده از شناسایی جاده و آنالیز اجسام متحرک (ردیابی اشیاء) اقدام به تشخیص موانع شده است. بدین منظور در مرحله اول، با استفاده از تبدیل wavelet اقدام به شناسایی جاده می‌شود. سپس در مرحله دوم فرایند، با آنالیز قسمت‌هایی که در حرکت هستند و با استفاده از مدل‌های بیزین اقدام به تشخیص موانع شده است. نکته قابل توجه در فعالیت انجام‌شده، قابل استفاده بودن روش پژوهش ارائه‌شده نسبت به شرایط مختلف محیطی با توجه به نتایج به‌دست‌آمده است. یکی از مشکلاتی که روش پیشنهادی پژوهش انجام‌شده دارد، شناسایی سایه خودرو همراه آن به‌عنوان مانع است. البته مشکل حاصل، به دلیل عدم در نظرگیری پردازش‌های همچون تشخیص

و حذف سایه است. در شکل ۲-۲۶، این مسئله مشخص است.



شکل ۲-۲۶: تشخیص سایه خودرو به همراه سایه به‌عنوان مانع در روش پیشنهادی [۷۱]

جینگ و همکاران [۷۲]، در فعالیت خود با استفاده از تبدیل هاف تصادفی [۷۲]، با توجه به ساختار الگوریتمی آن، اقدام به تشخیص موانعی همچون سرعت گیر و موانع طبیعی دیگر شده است. بدین منظور ابتدا با استفاده از فرایند لبه‌یابی اقدام به شناسایی لبه‌های تصویر می‌شود. سپس با استفاده از تبدیل هاف تصادفی، اقدام به شناسایی موانع شده است.



شکل ۲-۲۷: تشخیص سایه خودرو به همراه سایه به عنوان مانع در روش پیشنهادی [۷۲]

۲-۴ نتیجه‌گیری

در این بخش به ارائه پیشینه مطالعاتی از روش مورد نظر پرداختیم. در ابتدا به بیان روش‌های رویکرد کلاسیک مبتنی بر دستورات پایه‌ای پردازش تصویر پرداختیم. از مهم‌ترین امتیازات این روش‌ها، می‌توان به بلادرنگ بودن آن‌ها اشاره نمود. هرچند برای اجراهای مختلف با توجه به نوع داده ورودی، ممکن است نیاز به تنظیمات پیش از اجرا داشته باشند. از طرفی، نقطه ضعف دیگری که برای این روش‌ها می‌توان بیان نمود، عدم وجود یک معیار کمی و استاندارد، برای سنجش روش مورد نظر است. در واقع، معیار ارزیابی عمل‌کردی این روش‌ها بگونه‌ای نیست که بتوان مقداری کمی، از خود برنامه اجرایی به‌دست آورد. در مقابل، برای رویکرد مبتنی بر یادگیری عمیق، مهم‌ترین مزیتی که می‌توان عنوان نمود، دقت بالایی است که در مقایسه با روش‌های رویکرد اول دارند. از طرفی روش‌هایی هستند که با توجه به مقتضیات مورد نیاز برای اجرا، قابل ارزیابی با معیارهای کمی هستند

که از خود برنامه اجرایی قابل محاسبه می‌باشند. از مهم‌ترین چالش‌هایی هم که می‌توان برای آن‌ها بیان نمود، نیاز به دادگان زیاد و همچنین نیاز به دادگان برچسب‌گذاری شده می‌باشد.

فصل سوم

مبانی نظری

۳-۱ مقدمه

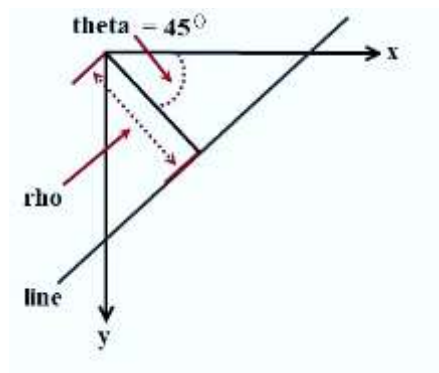
در فصل قبل به بررسی پیشینه‌ای از مطالعات انجام شده در زمینه تشخیص خطوط جاده‌ای و همچنین شناسایی موانع پرداختیم. در این فصل مفاهیم و مبانی نظری مورد استفاده در عملیات تشخیص خطوط و شناسایی موانع مورد بررسی قرار خواهد گرفت. بدین منظور در ابتدا دو مورد از الگوریتم‌های پایه‌ای پردازش تصویر مورد استفاده در روش پیشنهادی اول را بیان خواهیم نمود. سپس مبانی و مفاهیم مورد استفاده در روش پیشنهادی دوم را ارائه می‌دهیم. در قسمت پایانی این فصل نیز به بیان برخی مفاهیم مورد نیاز در شناسایی موانع، اشاره خواهیم نمود.

۳-۲ تبدیل هاف

تبدیل هاف یک تکنیک استخراج ویژگی می‌باشد که در پردازش تصاویر دیجیتال، بینایی ماشین و تجزیه و تحلیل تصویر به کار گرفته می‌شود. هدف از این تبدیل پیدا کردن شکل‌های مختلف، با ساختار ساده در تصویر است. توجه کنید که منظور از یک شکل ساده، شکلی است که بتوان آن را تنها با استفاده از تعداد کمی پارامتر نشان داد. در ابتدا تبدیل هاف برای شناسایی خطوط تصویر استفاده می‌شد اما به تدریج از آن برای تشخیص شکل‌های مختلف مانند دایره و بیضی هم استفاده شد. یک خط را می‌توان فقط به کمک دو پارامتر شیب و عرض از مبدا نشان داد. همچنین دایره با کمک ۳ پارامتر قابل نشان دادن است که عبارتند از مختصات مرکز و نیز شعاع R دایره. تبدیل هاف در یافتن این شکل‌ها در یک تصویر عمل کرد بسیار خوبی دارد. یکی از مهم‌ترین مزایای تبدیل هاف در یافتن خط و دایره در تصویر این است که نسبت به همپوشانی حساس نیست. به همین دلیل استفاده از این الگوریتم بسیار رایج است. ما از تبدیل هاف به منظور شناسایی خطوط استفاده می‌کنیم. در رابطه ۳-۱ روش استاندارد استخراج خط در یک تصویر با استفاده از تبدیل هاف بیان شده است. شکل ۳-۱ نحوه محاسبه خط را نشان می‌دهد. ما از تبدیل هاف برای تخمین مکان خطوط جاده و علامت‌گذاری آن‌ها استفاده می‌کنیم.

$$\rho = x \cos \theta + y \sin \theta$$

(۱-۳)



شکل ۳-۱: نحوه محاسبه خط با استفاده از تبدیل هاف

۳-۳ لبه‌یابی در تصویر

یکی از رایج‌ترین راه‌های تحلیل و شناسایی تصاویر برای انسان، استفاده از خطوط و مرزهای موجود در تصاویر است. در واقع سیستم بینایی انسان قبل از اینکه بخواهد ساختار رنگی درون تصاویر را بررسی کند، به خطوط بیان‌شده توجه دارد. لذا تفسیر لبه‌ها در تصویر قبل از هرگونه تحلیل دیگر، عملی منطقی است. در سیستم‌های پردازش تصاویر، شناسایی لبه‌ها بسیار اهمیت دارد. یکی از مهم‌ترین نتایجی که به‌عنوان هدف اصلی در لبه‌یابی تصاویر مطرح است، کاهش حجم پردازش داده‌ها است. نکته اصلی که باید به آن توجه نمود، حفظ ساختار ترکیبی تصویر از لحاظ مرزماند موجود در تصویر است. مرزماند، بیانگر شروع و اتمام یک بخش مشخصی از تصویر است. این شروع و اتمام، در واقع با همان مفهوم لبه قابل شناسایی هستند. در واقع، لبه به عنوان قسمتی که صفحات افقی و عمودی یک جسم در تصویر با همدیگر تلاقی دارند، در نظر گرفته می‌شود. الگوریتم‌های مختلفی از لبه‌یابی برای تصاویر وجود دارد. از مهم‌ترین الگوریتم‌هایی که می‌توان به آن اشاره نمود، لبه‌یابی کنی^۱ است که در ادامه مورد بررسی قرار خواهیم داد.

^۱ Canny Edge Detector

۳-۳-۱ الگوریتم لبه‌یابی کنی

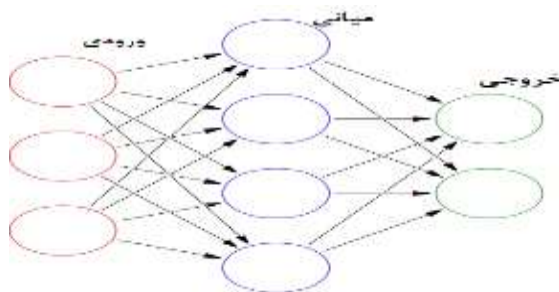
لبه‌یاب کنی در سال ۱۹۸۶ ایجاد شد و تا به امروز یک روش استاندارد و با دقت و کیفیت بالا محسوب می‌شود. این الگوریتم لبه‌یابی از سه بخش اصلی تضعیف‌کننده نویز، به‌دست‌آوردن نقاط کاندید لبه و حذف نقاط با احتمال پایین تشکیل شده است.

معیارهایی که در لبه‌یاب کنی مطرح هستند: ۱- کاهش دادن نرخ خطا بدین معنی که تا حد امکان هیچ لبه‌ای در تصویر نباید حذف شود و همچنین قسمتی که لبه نیست، نباید به‌عنوان لبه تشخیص داده‌شود. ۲- مکان تشخیص لبه واقعی باشد بدین معنی که در تشخیص لبه‌ها، هیچ پردازشی باعث جابجایی لبه از مکان واقعی خود در تصویر نباشد. در واقع تشخیص لبه کمترین فاصله ممکن از مکان واقعی خود را داشته باشد. ۳- تشخیص هر لبه به‌صورت یکتا باشد. (فقط یک پاسخ برای هر لبه) ۴- ضخامت لبه‌ها تا حد امکان کمترین مقدار خود باشند. یکی از امتیازات الگوریتم لبه‌یابی کنی، قدرت آن در تولید لبه‌های نازک تا حد یک بیکسل برای لبه‌های پیوسته است.

۳-۴ شبکه‌های عصبی

شبکه‌های عصبی مصنوعی با الهام از ساختار مغز انسان‌ها در هوش مصنوعی و بخصوص مسائل یادگیری ماشین به‌کار گرفته می‌شوند. معماری کلی این شبکه‌ها از اتصالات درونی مابین نرون‌ها ایجاد شده‌است. این نرون‌ها با ارتباطی که با یکدیگر دارند، مسائل با موضوعات متنوعی را حل می‌کنند. از جمله حوزه‌هایی که از شبکه‌های عصبی مصنوعی برای حل مسائل استفاده می‌گردد، می‌توان به کاربرد در علوم مختلف پزشکی، علوم مهندسی نظیر، صنایع، الکترونیک و علوم پایه همچون ریاضیات، آمار، اقتصاد و حتی هنر اشاره نمود. شبکه‌های عصبی مصنوعی بخشی از شاخه یادگیری ماشین محسوب می‌شوند که ایده اساسی کار آن‌ها همانطور که بیان شد، برقراری ارتباط درون‌ساختاری بین نرون‌هایشان است. در حل مسائل توسط این شبکه‌ها، اولین مرحله، عملیات

آموزش آن‌ها بواسطه الگوریتم‌های مختلف است. در شکل ۲-۳ ساختار کلی یک شبکه عصبی سه لایه مشخص است.



شکل ۲-۳: ساختار یک شبکه عصبی ساده

در بیانی دقیق‌تر در مورد معماری یک شبکه عصبی می‌توان گفت که ساختارشان از تعدادی گره که نرون نامیده می‌شوند و همچنین یال‌هایی که ارتباط‌دهنده نرون‌ها هستند، تشکیل شده‌است. یال‌های موجود در شبکه دارای یک ارزش مقداری هستند که اصطلاحاً وزن‌های شبکه نامیده می‌شوند. وقتی بیان می‌کنیم یک شبکه آموزش داده شده است، در واقع وزن‌های آن برای انجام عملیات مورد نظرمان به مقدار ایده‌آل جهت حل مسئله با کم‌ترین خطای مورد نظر رسیده‌اند. نرون‌های درون ساختار شبکه در گروه‌هایی تحت عنوان لایه قرار می‌گیرند. معمولاً خروجی هر لایه به عنوان ورودی لایه دیگر استفاده می‌شود. شبکه‌ها معمولاً شامل لایه‌های ورودی، میانی و خروجی هستند. رابطه ۲-۳ نحوه محاسبه هر نرون در شبکه عصبی و رابطه ۳-۳ خروجی نرون را بیان می‌کنند.

$$n_i = b + \sum_i w_i x_i \quad (2-3)$$

$$o_i = f(n_i) = f(b + \sum_i w_i x_i) \quad (3-3)$$

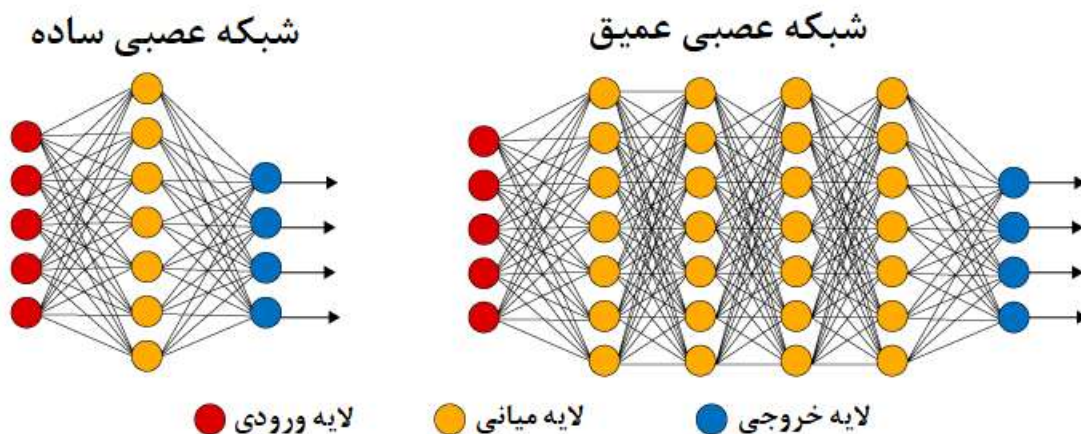
در واقع مقدار o_i وزن تولیدشده حاصل از نرون n_i در لایه میانی است. x_i ورودی شبکه و w_i وزن‌های شبکه که اتصال‌دهنده لایه ورودی به لایه میانی می‌باشند، هستند.

نکته قابل توجه، وجود لایه‌های میانی براساس کارکرد شبکه و یا ایده کاربردی کاربر تعیین می‌شوند که می‌توانند

وجود نداشته باشند یا بیشتر از یک لایه میانی متناسب با نیاز کاربر استفاده شوند. در صورتی که بیشتر از یک لایه میانی داشته باشیم، روابط ۲-۳ و ۳-۳ را می‌توانیم برای نرون‌ها و وزن‌های هر لایه مورد نظر به‌دست آوریم. زمانی که تعداد لایه‌های یک شبکه عصبی با توجه به شرایط مشخصی، همچون افزایش انتزاع (به این معنی که با افزایش تعداد لایه‌های میانی، ویژگی‌های جدیدتری به‌دست آوریم) بیشتر گردند (معمولا بیشتر از سه لایه برای شبکه که شامل یک لایه ورودی، یک لایه میانی و یک لایه خروجی است)، مفهومی پیچیده‌تر با نام شبکه‌های عصبی عمیق مطرح می‌شوند. در ادامه به توضیحی از این شبکه‌ها می‌پردازیم.

۳-۵ شبکه‌های عصبی عمیق

یادگیری عمیق ابزاری بسیار قدرتمند در یادگیری است که با به‌کارگیری شبکه‌های عصبی عمیق پیاده‌سازی می‌شوند از لحاظ قابلیت محاسباتی و کارایی، نسبت به شبکه‌های عصبی معمول و در کل الگوریتم‌های یادگیری مرسوم، امکانات بیشتری را برای ما فراهم می‌کنند. نکته دیگری که در مورد به‌کارگیری شبکه‌های عمیق قابل ذکر است، قابلیت و انعطاف‌پذیری بالای آن‌ها در کار با داده‌های با حجم بالا است. البته یکی از ملزومات استفاده از شبکه‌های عمیق نیاز به وجود مجموعه دادگان، با حجم بالای تعداد نمونه‌ها است. با توجه به اینکه در شبکه‌های عمیق با افزایش نمونه‌ها، فرایند آموزش بهتری نسبت به حالت معمول با استفاده از الگوریتم‌های معمول یادگیری داریم، افزایش دقت در این نوع یادگیری امری منطقی است. نکته مهم دیگری که در یادگیری شبکه‌های عمیق نسبت به الگوریتم‌های معمول یادگیری مطرح می‌شود، فرایند استخراج ویژگی در این نوع یادگیری است. یادگیری عمیق معمولا فرایند استخراج ویژگی را برخلاف الگوریتم‌های معمول یادگیری به‌صورت خودکار انجام می‌دهند لذا باعث بهینه بودن ویژگی‌های انتخابی برای آموزش می‌شود. شکل ۳-۳ یک شماتیک، از مفهوم شبکه عمیق در مقابل شبکه عصبی ساده را نشان می‌دهد. لایه‌های پنهان و لایه خروجی در شبکه عمیق، با توجه به نوع شبکه عمیق تعیین می‌شوند. برای مثال هریک از لایه‌های میانی، می‌توانند عملیات کانولوشن، پولینگ و یا سایر عملیات‌های موجود در یک شبکه کانولوشنی باشند. از طرفی لایه خروجی خود به‌تنهایی می‌تواند یک شبکه عمیق تماما متصل به‌عنوان یک طبقه‌بند باشد.



شکل ۳-۳: شماتیک شبکه عمیق در مقابل شبکه عصبی ساده

۳-۵-۱ معماری‌های اصلی در شبکه‌های عصبی عمیق

برای معماری‌های مختلف شبکه‌های عمیق، با افزایش تعداد لایه‌ها، پیچیدگی ساختاری آن‌ها بیشتر و لذا ویژگی‌های سطح بالاتری، در فرایند آموزش استخراج می‌شوند. از طرفی با افزایش پیچیدگی ساختاری عنوان شده، مدت زمان لازم برای آموزش نیز افزایش پیدا می‌کند.

از لحاظ نوع استفاده، شبکه‌های عمیق، قابلیت به کارگیری پارادایم‌های مختلف یادگیری را دارند. معمولاً سه پارادایم اصلی موجود در یادگیری، در انواع مختلف شبکه‌های عمیق وجود دارند. شبکه‌های عمیق کانولوشنی، شبکه‌های عمیق بازگشتی، شبکه‌های احتمالی و شبکه‌های مولد رقابتی چهار مورد از مهم‌ترین انواع شبکه‌های عمیق هستند. ما در این رساله از شبکه‌های عمیق کانولوشنی و شبکه‌های مولد رقابتی استفاده می‌کنیم که در ادامه آن‌ها را توضیح می‌دهیم.

۳-۵-۲ شبکه‌های عمیق کانولوشنی^۱

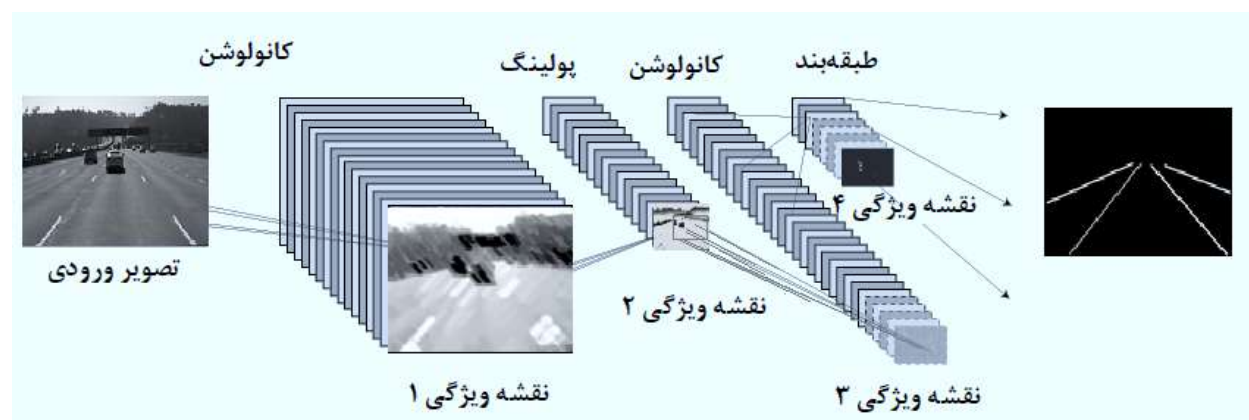
محوریت استفاده از شبکه‌های عمیق کانولوشنی بیشتر برای تصاویر است. در این شبکه‌ها معمولاً تصویر به عنوان ورودی به شبکه اعمال می‌شود و عملیات پردازشی به صورت مستقیم روی تصاویر و بر مبنای عملیات‌های

¹ Deep Convolution Networks

ماتریسی انجام می‌شود. در شبکه‌های عمیق کانولوشنی، لایه‌های کانولوشن^۱، پولینگ^۲، ReLu، نرمال‌سازی دسته‌ای^۳ و لایه تماماً متصل^۴ در کنار یکدیگر وجود دارند. مهم‌ترین فرایندهایی که اغلب از شبکه‌های عمیق کانولوشنی و به‌طور کلی معماری‌هایی که مبنای کانولوشن دارند، انجام می‌شوند، می‌توان به قطعه‌بندی تصاویر با رویکردهای موجود و همچنین طبقه‌بندی تصاویر اشاره نمود. از مهم‌ترین معماری‌های شبکه‌های عمیق کانولوشنی می‌توان به مواردی همچون LeNet، ResNet، AlexNet، VGGNet و GoogleNet اشاره نمود. البته هریک از این معماری‌ها خود به انواع مختلفی تقسیم‌بندی می‌شوند.

۳-۵-۲-۱ لایه‌های موجود در شبکه کانولوشن

همانطور که بیان شد، در شبکه‌های کانولوشن، ورودی شبکه تصویر است. شکل ۳-۴ معماری یک شبکه کانولوشن را نشان می‌دهد. در این بخش برای یک شبکه کانولوشن وظیفه عمل‌کردی هرکدام از لایه‌ها را توضیح می‌دهیم.



شکل ۳-۴: مثالی از معماری شبکه کانولوشن

۳-۵-۲-۲ لایه کانولوشن

مهم‌ترین لایه موجود در یک شبکه عمیق کانولوشن، لایه کانولوشن است. وظیفه‌ای که این لایه برعهده دارد،

¹ Convolution

² Pooling

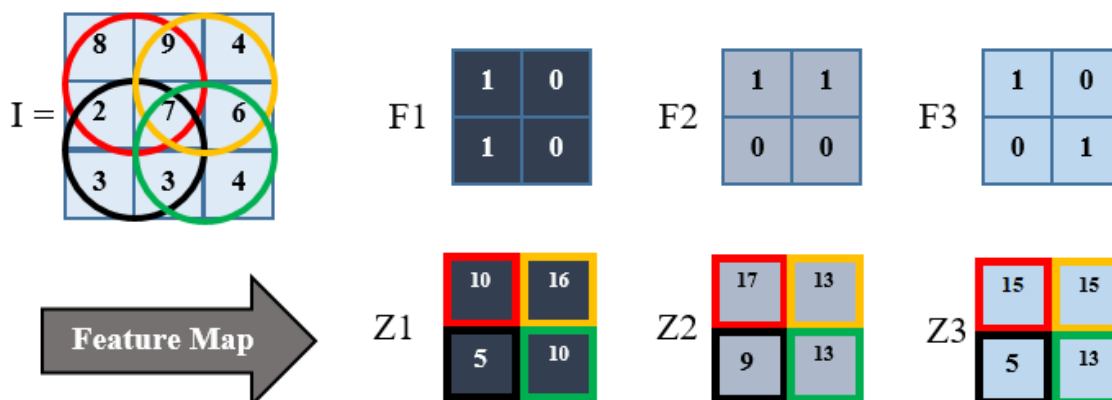
³ Batch Normalization

⁴ Fully Connected

استخراج ویژگی از تصویر ورودی است. برای استخراج ویژگی‌های موجود در تصویر، به لایه کانولوشن تعدادی فیلتر با یک اندازه مشخص^۱ و تعدادی پارامتر از جمله گام حرکت^۲ و همچنین سطح حاشیه صفر^۳ اعمال می‌شود. با ضرب نقطه‌ای یک پنجره از تصویر در هریک از فیلترها، یک استخراج ویژگی تحت عنوان نقشه ویژگی^۴ ایجاد می‌شود. رابطه ۳-۴ بیان‌کننده محاسبه یک نقشه ویژگی است. نحوه عملیات کانولوشنی و به‌دست آمدن یک مجموعه نقشه ویژگی در شکل ۳-۵ نمایش داده شده‌است.

$$z = I \cdot F \quad (3-4)$$

فرض کنیم I تصویر ورودی و F مجموعه فیلترهای مورد نظر باشند. از طرفی گام حرکت لغزشی برابر با ۱ و نهایتاً سطح حاشیه صفر برابر با مقدار ۰ باشد. بنابراین با فرض داشتن ۳ فیلتر ما سه نقشه ویژگی خواهیم داشت.



شکل ۳-۵: عملیات کانولوشنی و محاسبه یک مجموعه نقشه ویژگی

مسئله‌ای که در اینجا مهم است، اندازه نقشه ویژگی‌هایی است که تشکیل می‌شوند. به‌طور کلی با فرض بر اینکه اندازه ابعاد یک فیلتر مربعی برابر با f ، مقدار گام حرکت برابر با s ، اندازه ابعاد تصویر مربعی برابر n در m و نهایتاً سطح حاشیه صفر برابر با p باشد. اندازه یک نقشه ویژگی (Z_1, Z_2) از رابطه ۳-۵ به‌دست می‌آید.

¹ Kernel size

² Stride

³ Zero padding

⁴ Feature Map

$$\begin{cases} z_1 = \frac{(n - f + 2p)}{s} + 1 \\ z_2 = \frac{(m - f + 2p)}{s} + 1 \end{cases} \quad (5-3)$$

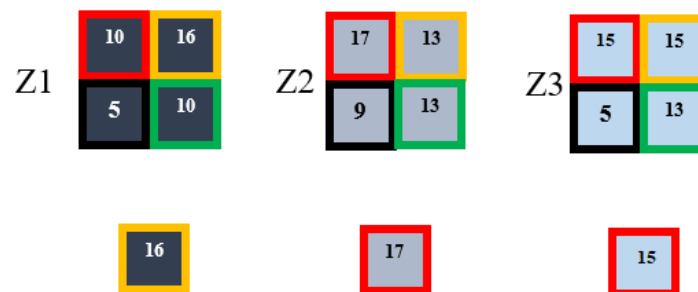
۳-۵-۲ لایه پولینگ

وظیفه اصلی پولینگ عملیات کاهش ویژگی است. با توجه به اندازه فیلتری که به پولینگ اعمال می‌شود، خروجی عملیات به صورت رابطه ۶-۳ است.

$$pool_{size} = \left\lfloor \frac{(z_1, z_2) - f}{s} \right\rfloor + 1 \quad (6-3)$$

در رابطه فوق (z_1, z_2) اندازه ابعاد نقشه ویژگی به دست آمده از رابطه ۵-۳ است. s گام حرکت و f نیز اندازه ابعاد فیلتر اعمال شده به عملیات پولینگ است.

خروجی یک عملیات پولینگ با توجه به نوع پولینگ مشخص می‌شود. معمولاً سه نوع عملیات پولینگ بیشینه^۱، پولینگ کمینه^۲ و پولینگ متوسط^۳ وجود دارد. در بین سه مورد پولینگ، پولینگ بیشینه بیشتر از بقیه مورد استفاده قرار می‌گیرد. شکل ۶-۳ خروجی حاصل از عملیات پولینگ بیشینه در شکل ۵-۳ را نشان می‌دهد.



شکل ۶-۳: خروجی عملیات پولینگ بیشینه

¹ Max Pooling

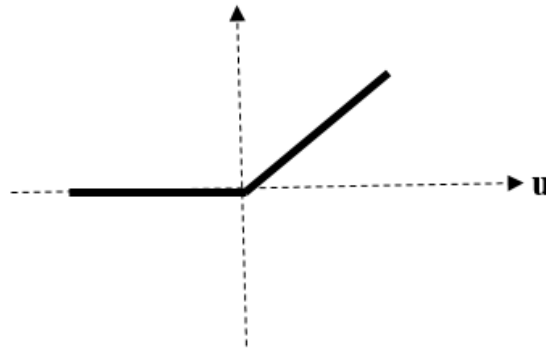
² Min Pooling

³ Avg Pooling

۳-۵-۲-۴ لایه ReLu

این لایه به خروجی هر نرون، یک تابع فعال‌ساز اعمال می‌کند. به‌طور معمول این لایه بعد از لایه کانولوشن قرار گرفته می‌شود. شکل ۷-۳ تابع Relu را نمایش و رابطه ۷-۳ تابع Relu را نشان می‌دهد.

$$f(u) = \begin{cases} 0 & x \leq 0 \\ u & x > 0 \end{cases} \quad (7-3)$$



شکل ۷-۳: تابع Relu

۳-۵-۲-۵ لایه تماماً متصل

این لایه معمولاً در انتهای یک شبکه کانولوشن قرار گرفته و عملیات طبقه‌بندی را انجام می‌دهد. در واقع لایه‌های نامبرده در قبل بیشتر در فرایند استخراج ویژگی استفاده می‌شدند اما لایه تماماً متصل برای فرایند طبقه‌بندی نهایی برروی تصویر موردنظر استفاده می‌شود. فرایند طبقه‌بندی با توجه به هدف کاربر می‌تواند تعیین‌کننده کلاس تصویر و یا عملیات قطعه‌بندی تصویر باشد.

آنچه در ۳-۵-۲ بیان شد، خلاصه‌ای از لایه‌های پایه‌ای در شبکه عمیق کانولوشن است. این لایه‌ها می‌توانند به ترتیب مدنظر و به دفعات مشخص، طبق صلاحدید کاربر مورد استفاده قرار گرفته شوند. این شبکه در عین تئوری ساده‌ای که دارد، از لحاظ عمل‌کردی بسیار قدرتمند است و همچنین به‌عنوان شبکه پایه در بسیاری از شبکه‌های عمیق مورد استفاده قرار می‌گیرد.

۳-۵-۳ شبکه‌های عمیق مولد رقابتی^۱

شبکه‌ی عمیق مولد رقابتی، نمونه‌ای از شبکه‌های عمیق است که در سال ۲۰۱۴ یوهان گودفلو [۷۳] معرفی نمود.

این شبکه از دو بخش اصلی مولد^۲ و تفکیک‌کننده^۳ تشکیل شده است که فرایند آموزش هرکدام از این شبکه‌ها با ایده تئوری بازی دو نفره عمل می‌کند [۷۳]. آموزش هریک از شبکه‌های نامبرده باید به گونه‌ای باشد که بالاترین دقت ممکن را داشته باشند. این رویکرد به این معنی است که هرکدام از شبکه‌های موجود دارای یک عمل‌کرد بهینه در ساختار خود هستند. بدین ترتیب می‌توان بیان نمود، عمل‌کرد رقابتی کل شبکه زمانی بهینه خواهد بود که هر کدام از شبکه‌های موجود دارای عمل‌کرد بهینه باشند. در این نوع شبکه، معمولاً فرایند آموزش به صورت نظارت‌نشده انجام می‌شود، اما می‌توان استفاده از شبکه را برای فرایندهای نظارت‌شده نیز انجام داد. از ساده‌ترین انواع شبکه‌های عمیق رقابتی، می‌توان شبکه عمیق خودرمزگذار^۴ را نام برد. همانطور که بیان گردید، شبکه‌های عمیق مولد رقابتی معمولاً به شکل نظارت‌نشده آموزش داده می‌شوند، اما نوع نظارت شده و البته نیمه‌نظارتی نیز در این نوع شبکه‌ها وجود دارد [۷۴]. شبکه عمیق مولد رقابتی نیمه‌نظارتی^۵ نوعی از شبکه‌های عمیق مولد رقابتی است که می‌توان برای فرایند طبقه‌بندی استفاده نمود. در ادامه این نوع شبکه را توضیح می‌دهیم.

۳-۵-۴ الگوی یادگیری نیمه‌نظارتی

در الگوهای^۶ یادگیری، با توجه به برچسب‌دار یا بدون برچسب بودن داده‌گان استفاده شده، می‌توان از یادگیری‌های نظارت‌شده و یا نظارت‌نشده استفاده نمود. قاعداً در فرایند آموزش سیستمی که در آن از یادگیری نظارت‌شده استفاده شود، دقت و عمل‌کرد نهایی سیستم بسیار بهتر خواهد بود. همچنین معیارهای جامع‌تری

¹ Deep Generative Adversarial Networks

² Generative

³ Discriminator

⁴ AutoEncoder

⁵ Semi-Supervised Deep Generative Adversarial Network

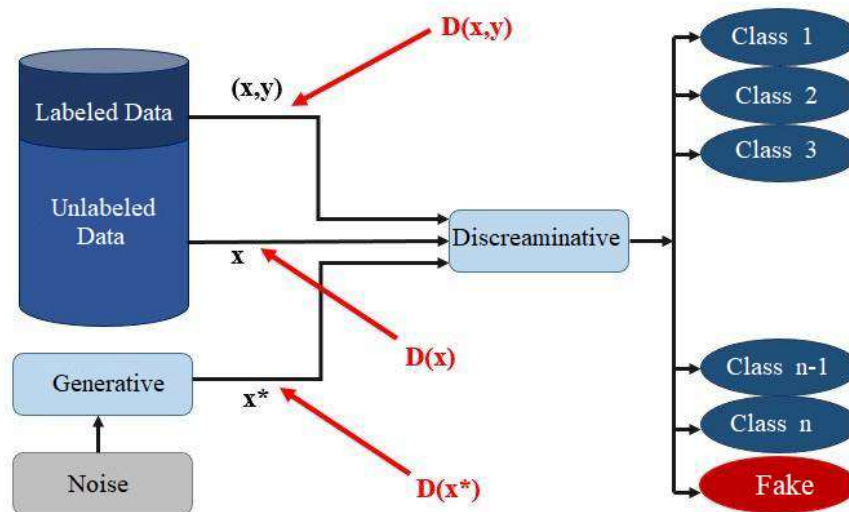
⁶ Paradigms

می‌توان برای ارزیابی عملکرد سیستم ارائه نمود. اما مسئله‌ای که مطرح است این است که فرایند برچسب‌گذاری برای دادگان بسیار زمان‌بر و حتی پر هزینه خواهد بود، لذا ایده استفاده از الگوی یادگیری نیمه‌نظارتی مطرح می‌شود که در آن مجموعه دادگانی که استفاده می‌شود ترکیبی از دادگان برچسب‌دار و فاقد برچسب می‌باشد و با توجه به الگوریتم موجود در مسئله نحوه استفاده از آن‌ها در فرایند آموزش سیستم به‌کار گرفته می‌شود. هرچند که نتایج حاصل از به‌کارگیری این روش‌های نیمه‌نظارتی آنقدر انتظارات را برآورده ننموده است، اما با توجه به سرعت بهتر نسبت به روش‌های تماماً نظارتی و دقت نسبی آن نسبت به رویکرد اجرایی، استفاده از آن رو به افزایش و البته زمینه پژوهشی برای بهتر شدن نتایج حاصل را بسیار گسترده نموده است.

۳-۵-۵ شبکه مولد رقابتی نیمه‌نظارتی^۱ (SGAN)

ساختار این شبکه در شکل ۳-۸ نشان داده شده است. این شبکه از دو بخش تفکیک‌کننده و مولد تشکیل شده است که هریک شبکه‌های عصبی هستند که طراح با توجه به ماهیت مسئله از آن‌ها استفاده می‌کند. تفاوت اصلی شبکه نیمه‌نظارتی با ساختار معمول شبکه مولد رقابتی، وجود دادگان برچسب‌دار در شبکه نیمه‌نظارتی است.

¹ Semi-Supervised Generative Adversarial Network



شکل ۳-۸: شبکه مولد رقابتی نیمه نظارتی

همانطور که در ساختار شبکه مشخص است، در یک مسئله با n کلاس، خروجی نهایی برابر با $n+1$ کلاس است که یکی از آن‌ها مربوط به کلاس دادگان تقلبی است که در شکل ۳-۸ با نام Fake مشخص شده است. از این شبکه اغلب برای فرایندهای طبقه‌بندی استفاده می‌شود. اما با انجام یک رویکرد مشخص، می‌توان برای بخش‌بندی تصاویر نیز از شبکه پیش‌روی استفاده نمود.

همانطور که بیان گردید، شبکه‌های مولد رقابتی از دو قسمت اصلی ایجاد شده‌اند. یکی شبکه مولد که با توجه به یک بردار نویز تصادفی داده‌های منطبق بر داده‌های اصلی ایجاد می‌کند و دیگری شبکه تفکیک‌کننده که وظیفه دارد، داده واقعی را از داده تقلبی تمایز دهد. روال کارکرد این شبکه بر طبق الگوریتم minmax و بر اساس یک تئوری بازی با توجه به رابطه ۳-۸ مشخص می‌شود.

$$\min_{(G)} \max_{(D)} V(D, G) = E_{x \sim p_{data}(x)} [\log(D(x))] + E_{z \sim p_z(z)} [\log(G(z))] \quad (۳-۸)$$

رابطه شماره ۳-۸ پایه اصلی تمام شبکه‌های مولد رقابتی است که ثابت بوده و با توجه به انواع شبکه‌های موجود از این نوع، روابط دیگر متفاوت هستند. در واقع رابطه شماره ۳-۸ کلیت ثابت برای تمام شبکه‌ها است. در این رابطه، $V(D, G)$ بیانگر تابع ارزش کلی است. برای تابع فوق مقدار $E_{x \sim p_{data}(x)}$ بیان‌کننده مقدار ارزش مورد انتظار برای قسمت نظارت‌نشده در شبکه تفکیک‌کننده است که از توزیع $p_{data}(x)$ بر روی دادگان بدون

برچسب واقعی به دست می‌آید. همچنین $E_{z \sim p_z(z)}$ مقدار ارزش مورد انتظار برای شبکه مولد است، که از توزیع $p_z(z)$ به دست می‌آید. این توزیع اغلب از نوع نرمال است. D تابع شبکه تفکیک کننده و G نیز تابع شبکه مولد هستند. از آنجایی که V ماهیت احتمالاتی دارد، بهینه ترین مقدار ممکن برای آن یک است.

در یک شبکه مولد رقابتی نیمه نظارتی، با توجه به ساختار الگوریتمی موجود، برای اینکه ما خطای بخش تفکیک کننده را کمینه کنیم، باید خطای سه بخش از ورودی های آن را کمینه سازیم. یکی خطای مربوط به بخش مولد، دوم خطای مربوط به دادگان بدون برچسب و نهایتاً خطای حاصل از دادگانی که بخش نظارت شده شبکه تفکیک کننده می باشند. رابطه شماره ۳-۹ نشان دهنده فرم کلی خطای بخش تفکیک کننده در شبکه مولد رقابتی نیمه نظارتی است.

$$L_D = -E_{x \sim p_{data}(x)}[\log(D(x))] - E_{z \sim p_z(z)}[\log(1 - D(G(z)))] + \gamma E_{x,y \sim p(x,y)}[(y, P(y|x, D))] \quad (9-3)$$

به طوری که در رابطه فوق، همانطور که بیان شد، تابع D بیانگر تابع شبکه تفکیک کننده بخش نظارت شده ای است که در فرایند قطعه بندی وجود دارد. این تابع در رابطه شماره ۳-۱۰ مشخص شده است. G معرف شبکه مولد، z بیانگر بردار نویز تصادفی نرمال، γ بیانگر ضریب متعادل سازی شبکه های تفکیک کننده و مولد و y خروجی مطلوب (تصویر ماسک مورد نظر) می باشند. همچنین $E_{x,y \sim p(x,y)}$ مقدار ارزش مورد انتظار نظارت شده شبکه تفکیک کننده هستند.

$$D(x) = [1 - P(y = fake|x)] \quad (10-3)$$

و خطای حاصل از بخش مولد شبکه که از رابطه ۳-۱۱ محاسبه می شود [۷۵].

$$L_G = E_{z \sim p_z(z)}[\log(1 - D(G(z)))] \quad (11-3)$$

با توجه به روابط ۳-۸ تا ۳-۱۱، می توان اذعان نمود که بهترین حالت عمل کردی ممکن برای شبکه، زمانی است که در مقادیر $p_z(z)$ و $p_{data}(x)$ باهم برابر باشند.

۳-۶ توصیف کننده^۱

توصیف کننده‌ها در یادگیری ماشین، یک بازنمایی از جزئیات ترکیب ساختاری دادگان دارند. بدین معنی که با استفاده از توصیف کننده‌ها می‌توان ویژگی‌های مهم ساختاری دادگان را برای پردازش‌های مورد نظر به دست آورد. توصیف کننده می‌تواند برای دادگان با فرمت‌های مهمی از قبیل متن، تصویر و سیگنال‌ها استفاده شود. از مهم‌ترین دلایل استفاده از توصیف کننده‌ها، استخراج ویژگی‌های پنهان و نگاشت فرمت‌های داده‌ای برای استفاده بهتر در پردازش‌های یادگیری است.

در پردازش تصاویر، توصیف کننده‌های متنوعی وجود دارد. از جمله مهم‌ترین مواردی که می‌توان نام برد: هیستوگرام مبتنی بر گرادیان^۲ [۷۶]، تبدیل فرم ویژگی مقیاس ثابت^۳ [۷۷]، ویژگی مقاوم سرعت یافته^۴ [۷۸] و کیف کلمات بصری^۵ [۷۹] می‌باشند. با توجه به ثابت نبودن پس‌زمینه تصاویر ویدیویی، استفاده از روش‌هایی همچون مدل مخلوط گوسی^۶ کارایی بهینه در شناسایی اشیاء را ندارند. ما در این رساله از توصیف کننده‌ها برای شناسایی اشیاء استفاده نموده‌ایم. با توجه به سادگی روش هیستوگرام مبتنی بر گرادیان، برای تشخیص نوع مانع از آن استفاده شده‌است. در ادامه به توضیح روش هیستوگرام مبتنی بر گرادیان می‌پردازیم.

۳-۶-۱ هیستوگرام مبتنی بر گرادیان

هیستوگرام مبتنی بر گرادیان، یک توصیف کننده ویژگی هست که معمولاً برای استخراج ویژگی از تصاویر مورد استفاده قرار می‌گیرد و به صورت گسترده‌ای در بینایی ماشین برای کارهایی تشخیص شی استفاده می‌شود. توانایی اصلی این توصیف کننده، شناسایی جهت لبه‌های موجود در تصویر است. این قابلیت با استفاده از گرادیان جهت تعیین می‌شود و به همین دلیل است که از قدرت بالایی برخوردار است. در عملیات تشخیص مانع، با توجه به اینکه تنها دو کلاس (ماشین و غیرماشین) داریم، الگوریتم HOG نسبت به الگوریتم‌های دیگر مناسبتر

¹ Descriptor

² Histogram Oriented of Gradient

³ Scale Invariant Feature Transform(SIFT)

⁴ Speeded-Up Robust Feature(SURF)

⁵ Visual Bag Of Word(VBOW)

⁶ Mixture Of Gaussian

است. الگوریتم‌هایی همچون SURF و SIFT و BOW زمانی که تعداد کلاس‌ها زیاد باشند، بهتر عمل می‌کنند. اما دارای پیچیدگی محاسباتی بالایی نسبت به HOG هستند. به همین دلیل در شناسایی کلاس مانع از HOG استفاده نمودیم.

اگر بخواهیم با توجه به آنچه در معرفی توصیف‌کننده HOG بیان شد، یک تعریف رسمی از آن داشته باشیم، می‌توانیم بیان کنیم: این توصیف‌کننده تعداد رخداد یک جهت در لبه‌های تصویر را بیان می‌کند. برای به‌دست آوردن تعداد رخدادهای یک جهت مشخص، مراحل زیر انجام می‌شود.

۱. تغییرات یک پیکسل از جهت‌های افقی و عمودی را با استفاده از روابط ۱۲-۳ به‌دست می‌آوریم.

$$\begin{cases} G_x = (I_{x+1,y} - I_{x-1,y}) \\ G_y = (I_{x,y+1} - I_{x,y-1}) \end{cases} \quad (12-3)$$

۲. مقدار اندازه جهت را برای هریک از پیکسل‌ها با استفاده از ۱۳-۳ محاسبه می‌کنیم.

$$\|A\| = \sqrt{G_x^2 + G_y^2} \quad (13-3)$$

۳. جهت گرادیان برای هر پیکسل، طبق رابطه ۱۴-۳ محاسبه می‌شود.

$$\tan \theta = \frac{G_y}{G_x} \quad (14-3)$$

۴. در آخرین مرحله، اندازه زاویه را برای هر پیکسل با استفاده از رابطه ۱۵-۳ به‌دست می‌آوریم.

$$\theta = \tan^{-1}\left(\frac{G_y}{G_x}\right) \quad (15-3)$$

خروجی الگوریتم HOG برای یک تصویر مشخص در شکل ۹-۳ قابل مشاهده است.



الف ب

شکل ۳-۹: الف. تصویر اصلی ب. تصویر HOG

۳-۷ طبقه‌بند ماشین بردار پشتیبان^۱

روش طبقه‌بندی SVM در سال ۱۹۶۳ برای مدل‌های خطی مطرح گردید و در سال ۱۹۶۵ برای جداسازی غیرخطی دادگان تعمیم داده شد. این روش به‌عنوان یکی از پرقدرت‌ترین الگوریتم‌های یادگیری ماشین جهت طبقه‌بندی دادگان محسوب می‌شود. طبقه‌بند ماشین بردار پشتیبان روشی جدید به حساب می‌آید که در سال‌های اخیر، نسبت به الگوریتم‌های قدیمی‌تری چون شبکه‌های MLP^۲ کارایی بیشتری از خود نشان داده است. رویکرد عملیاتی این روش جداسازی خطی دادگان است، به‌طوری که در آن با استفاده از یک فرایند بهینه‌سازی، خطی انتخاب می‌گردد که حاشیه اطمینان بیشینه شود. رابطه مرز جداسازی بین دو کلاس با استفاده از رابطه ۳-۱۶ مشخص می‌شود. همچنین برای بیشینه شدن حاشیه اطمینان بین دو کلاس، رابطه لاگرانژ ۳-۱۷ باید کمینه گردد [۸۰].

$$y = w \cdot X + b \quad (۳-۱۶)$$

$$\min_{w, b, \alpha} \left\{ \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i [y_i (w \cdot X_i - b) - 1] \right\} \quad (۳-۱۷)$$

برای اینکه رابطه ۳-۱۷ کمینه شود، باید قسمت دوم رابطه که عبارت حاصل سیگما است بیشینه شود. برای این کار باید α_i هایی را به دست آوریم که حاصل جمع را بیشینه کنند. به ازای هر نمونه یک α_i محاسبه می‌گردد

^۱ Support Vector Machine

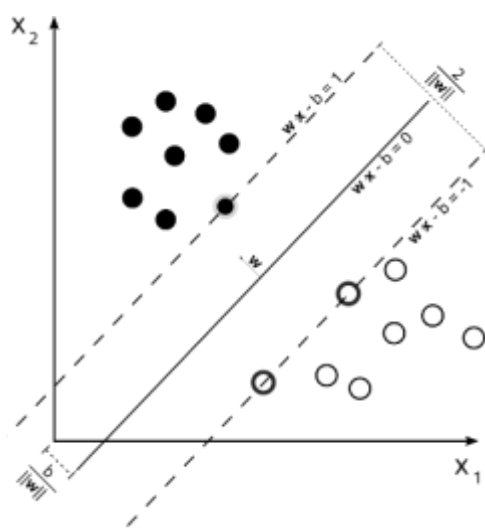
^۲ Multi Layer Perceptron

که در آن، α_i های مخالف صفر داده‌های مرزی و یا همان بردارهای پشتیبان هستند. پس از محاسبه این α_i ها، می‌توان مقادیر پارامترهای w و b را با استفاده از روابط ۱۸-۳ و ۱۹-۳ محاسبه کرد.

$$w = \sum_i \alpha_i y_i X_i \quad (18-3)$$

$$b = \frac{1}{N_{SV}} \sum_{i=1}^{N_{SV}} (w \cdot X_i - y_i) \quad (19-3)$$

در رابطه ۱۹-۳، متغیر i در محدوده تعداد بردارهای پشتیبان تعریف می‌شود [۸۰].



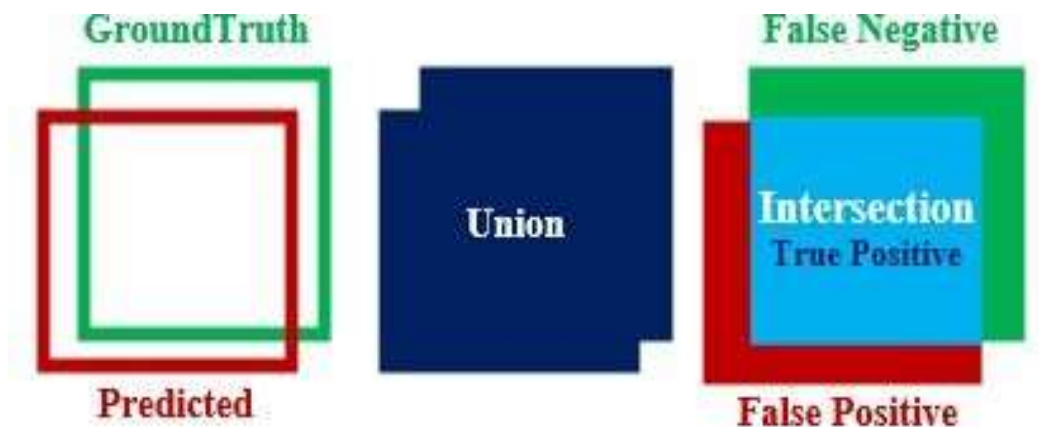
شکل ۳-۱۰: روش ماشین بردار پشتیبان خطی

۳-۸ معیارهای دقت شناسایی خطوط

۳-۸-۱ معیار $(IOU)^1$

معیار IOU یک معیار ارزیابی است که در عملیات قطعه‌بندی و تشخیص اشیاء استفاده می‌شود. شکل ۳-۱۱ مفهوم این معیار را بیان می‌کند و نحوه محاسبه این معیار در رابطه ۳-۲۰ تعریف شده‌است.

¹ Intersection Over Union



شکل ۱۱-۳: مفهوم معیار IOU

شکل ۱۱-۳ سمت راست، بیان کننده کلیت معیار IOU است. همانطور که در شکل، قسمت وسط مشخص است، Union، بیان گر اجتماع تمام پیکسل‌های پیش‌بینی‌شده و پیکسل‌های صحیح موجود در ماسک است. پیش‌بینی ما از ماسک دارای دو ناحیه اصلی است، یکی اشتراک با مرجع صحیح که بخش اشتراک نامیده می‌شود (قسمت آبی رنگ در تصویر سمت راست) و دیگری FalsePositive (قسمت قرمز رنگ در تصویر سمت راست) که پیکسل‌های پیش‌بینی‌شده نادرست در عملیات تشخیص خطوط هستند. با توجه به آنچه که بیان شد، رابطه ۲۰-۳ معیار IOU را بیان می‌کند.

$$IoU = \frac{TP}{FP + TP + FN} \quad (۲۰-۳)$$

در رابطه ۲۰-۳ مقدار TP پیش‌بینی صحیح یا همان اشتراک است، FP برابر با پیش‌بینی نادرست که به عنوان FalsePositive به آن اشاره شد و نهایتاً FN هم قسمتی از خطوط جاده بوده که به اشتباه تشخیص داده نشده است.

۳-۸-۲ معیار دقت^۱

اما معیار دیگری که برای شناسایی خطوط می‌توان از آن نام برد، دقت است. این معیار بیشتر برای دقت قطعه‌بندی

^۱ Accuracy

در سطح پیکسلی مطرح می‌شود. تنها تفاوت معیارهای IOU و دقت، در نادیده گرفتن پیش‌بینی نادرست در معیار دقت است. رابطه ۲۱-۳ معیار دقت را بیان می‌کند.

$$ACC = \frac{N_{TP}}{N_{TP+FN}} \quad (۲۱-۳)$$

در این رابطه، متغیرهای N_{TP} و N_{TP+FN} به ترتیب تعداد نقاط (پیکسل‌های) پیش‌بینی شده به عنوان خطوط و پیکسل‌های خطوط موجود در مرجع صحیح (ماسک) می‌باشند.

۹-۳ نتیجه‌گیری

در این فصل به بیان مبانی نظری و مفاهیمی که در روش پیشنهادی خود به کار خواهیم گرفت اشاره نمودیم. در فصل بعد با بیان روش‌های پیشنهادی مورد نظر، چگونگی استفاده از هریک از آن‌ها به طور دقیق‌تری مشخص می‌شود.

فصل چهارم

روش‌های پیشنهادی

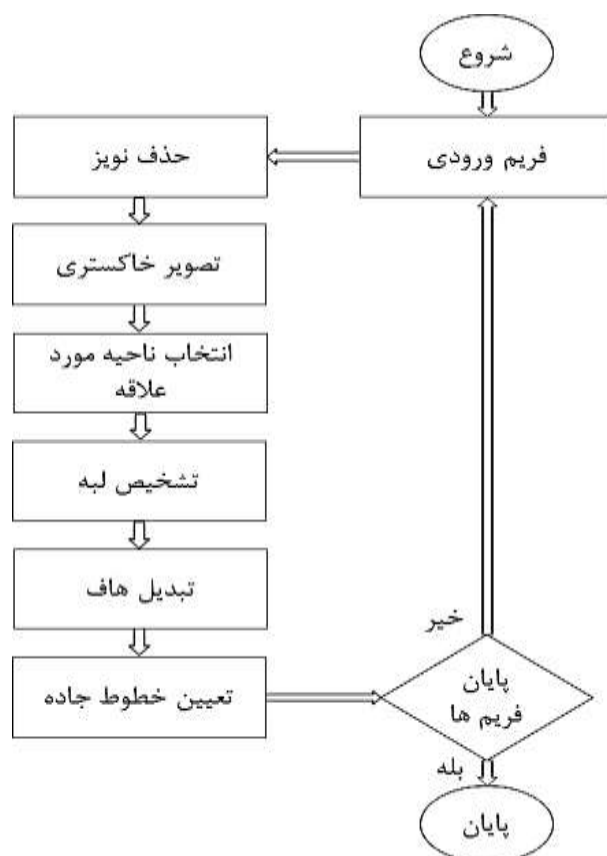
در این فصل، ابتدا الگوریتمی با رویکرد کلاسیک و مبتنی بر دستورات پایه‌ای پردازش تصویر برای شناسایی خطوط جاده بیان می‌کنیم. از ویژگی‌های اصلی این روش، سادگی، بلادرنگ بودن و عمل کرد خوب در شرایط محیطی مختلف است، هرچند نیازمند مجموعه تنظیمات اولیه، برای اجرا در مجموعه تصاویر ورودی با ویژگی‌های ساختاری متفاوت است. در روش پیشنهادی دوم، به تشریح مراحل فرایندهای شناسایی خطوط با استفاده از یک رویکرد مبتنی بر یادگیری می‌پردازیم. این روش، مبتنی بر یادگیری شبکه‌های عمیق بوده و از یک الگوی نیمه‌نظارتی برای شناسایی خطوط استفاده می‌کند. از مزایای این روش می‌توان به نیمه‌نظارتی بودن آن اشاره نمود که نسبت به روش‌های باناظر، نیاز به دادگان برچسب‌گذاری شده کمتری دارد. از مزایای روش پیشنهادی دوم نسبت به روش اول تشخیص خطوط، قابلیت ارزیابی روش بر مبنای معیارهای پارامتریکی است که می‌توان در خود برنامه آن‌ها را محاسبه و در نتیجه ارزیابی جامع نسبت به عملکرد روش داشت. هرچند از لحاظ سرعت زمان اجرایی و سادگی عملیاتی روش اول از یک روال الگوریتمی ساده‌تری برخوردار است. نهایتاً در بخش سوم، عملیات شناسایی موانع را بررسی می‌کنیم. در این بخش ابتدا یک روال یادگیری برای آموزش شناسایی خودروها به‌عنوان مانع در جاده ارائه می‌دهیم. از ویژگی‌های الگوریتم به‌کارگرفته‌شده، سادگی و همچنین دقت بالای آن می‌باشد.

۴-۲ الگوریتم پیشنهادی مبتنی بر رویکرد کلاسیک استفاده از دستورات پایه‌ای پردازش تصویر

بلوک‌دیگرام روش پیشنهادی اول برای تشخیص خطوط جاده در شکل ۴-۱ مشاهده می‌شود. در این روش پیشنهادی که مبتنی بر اجرای یک مجموعه متوالی از الگوریتم‌های پایه‌ای پردازش تصویر است، دو نکته مهم برای اجرای عملیات وجود دارد.

۱. بخش تعیین ناحیه مورد علاقه که با استفاده از یک ماسک‌گذاری به‌صورت پویا انجام می‌شود. لذا یکی از نوآوری‌هایی که برای این روش پیشنهادی مطرح است، همین مسئله است.

۲. دومین نکته حائز اهمیت در این روش پیشنهادی، که هدف اصلی فعالیت مورد نظر در این رساله است، بلادرنگ بودن روش است.



شکل ۴-۱: بلوک دیاگرام تشخیص خطوط جاده با استفاده از رویکرد کلاسیک

۴-۲-۱ عملیات پیش پردازش

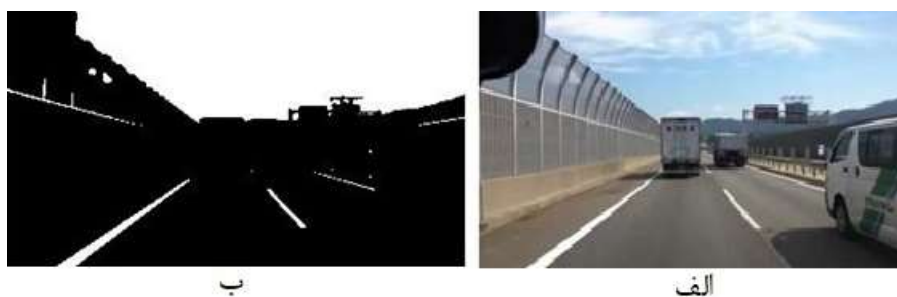
اولین قدم پس از دریافت هر فریم از تصویر ویدیویی تهیه شده توسط دوربین نصب شده در داخل خودرو، عملیات حذف نویزهای موجود در تصویر است. بدین منظور با استفاده از یک فیلتر گوسی با انحراف معیار استاندارد^۱ (سیگما) برابر با ۰,۵ و با دامنه فیلتر^۲ با مقدار پیش فرض خودکار^۳، نویز تصویر حذف می شود. با توجه به پارامترهایی که برای فیلتر حذف نویز استفاده شده است، ممکن است اثرات جانبی همچون مات شدن تصویر و به تبع آن از بین رفتن لبه ها به وجود بیاید. لذا مقادیر بکارگرفته شده، با سعی و خطا انجام شده است که چنین

^۱ Standard Deviation

^۲ Filter Domain

^۳ Auto

اثراتی به کمترین مقدار ممکن برسد. البته این نتایج زمانی که عملیات حذف نویز بر روی تصاویر خاکستری باشد، بیشتر مشهود است. لذا یکی از دلایل استفاده از فیلتر بر روی تصویر اصلی همین مسئله بوده است. پس از حذف نویز به منظور کاهش حجم محاسبات و افزایش سرعت پردازش، در ادامه فرایند تشخیص خطوط، تمام عملیات خود را بر روی تصویر خاکستری انجام می‌دهیم. پس از تغییر فضای عملیاتی کار به فضای خاکستری، با استفاده از روش آستانه‌گذاری اتسو و انتخاب یک مقدار آستانه مناسب، تصویر از حالت خاکستری به حالت دودویی تبدیل می‌شود. این عملیات به عنوان آخرین مرحله پیش‌پردازش است و خروجی آن به عنوان ورودی مرحله بعد و جهت انتخاب ناحیه مورد علاقه به کار گرفته می‌شود. نمونه‌ای از تصویر ورودی و خروجی مرحله پیش‌پردازش در شکل ۴-۲ قابل مشاهده است.



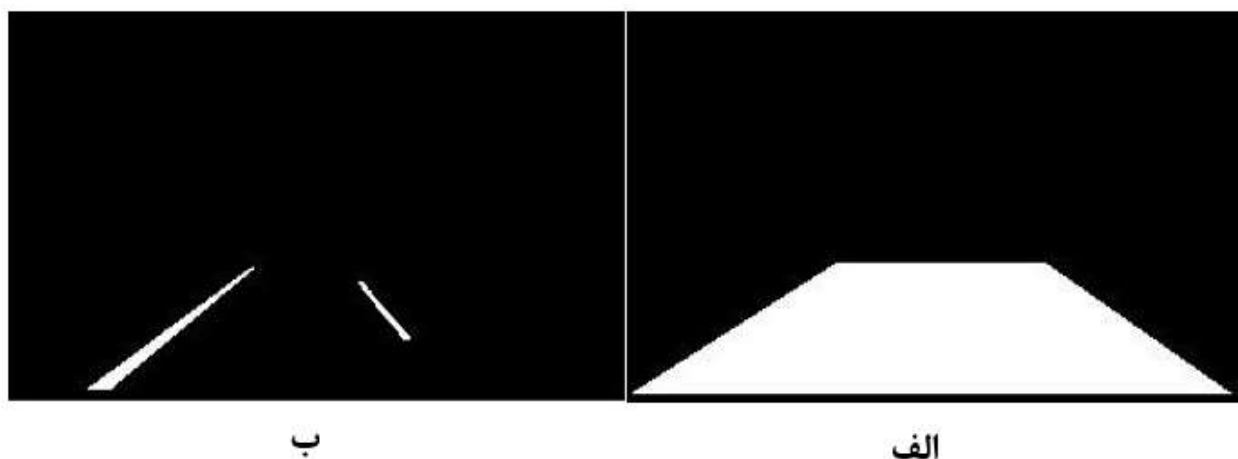
شکل ۴-۲: الف. نمونه تصویر اصلی ب. تصویر خروجی مرحله پیش‌پردازش

۴-۲-۲ انتخاب ناحیه مورد علاقه

پس از انجام عملیات پیش‌پردازش، یک ناحیه مورد علاقه وفقی از تصویر ورودی انتخاب می‌کنیم که شامل قسمت جاده در تصویر است و نواحی غیرمرتبط به جاده، مانند ناحیه مربوط به آسمان، از تصویر ورودی حذف می‌شوند.

باتوجه به اینکه دوربین جهت تهیه تصویر از جاده، داخل خودرو پشت شیشه جلو و وسط نصب شده است بنابراین در تصاویر به دست آمده از سطح جاده، خطوط جاده روبروی خودرو و در دوطرف آن قرار دارند. بدین منظور یک ناحیه چندضلعی پویا از تصویر روبروی خودرو که شامل خطوط جاده در دوطرف خودرو می‌باشد به عنوان ناحیه مورد علاقه انتخاب می‌شود. برای انتخاب ناحیه مورد علاقه از یک ماسک دوزنقه‌ای شکل استفاده

می‌کنیم در طراحی ماسک دوزنقه‌ای شکل، از اطلاعات نقطه محوشدگی تصویر استفاده می‌شود. به این صورت که پس از به‌دست آوردن نقطه محوشدگی، یک دوزنقه در ناحیه پایین نقطه محوشدگی ایجاد می‌شود که محدوده روبروی خودرو که شامل خطوط جاده هست را پوشش می‌دهد. در شکل ۳-۴ (الف)، ماسک طراحی شده برای نمونه تصویر ورودی نشان داده شده است که در آن قسمت مربوط به ناحیه مورد علاقه یک و بقیه تصویر صفر است. با اعمال این ماسک به تصویر دودویی ورودی، ناحیه مورد علاقه که شامل خطوط جاده در روبروی خودرو هست، انتخاب می‌شود. در شکل ۳-۴ (ب)، تصویر دودویی خروجی، پس از اعمال ماسک طراحی شده به تصویر دودویی ورودی نشان داده شده است.



شکل ۳-۴: الف. ماسک جهت انتخاب ناحیه مورد علاقه ب. تصویر دودویی ناحیه مورد علاقه

۳-۲-۴ تشخیص لبه

یکی از ویژگی‌های مفید و موثر در شناسایی اشیاء یا قسمت‌های موردنیاز در تصاویر، استفاده از لبه‌های تصویر است، بنابراین ضرورت طراحی آشکارسازهای مناسب از اهمیت بسیاری برخوردار است. لبه تصویر، نقاطی از تصویر است که در آن شدت روشنایی به‌طور ناگهانی تغییر کرده است. در حالت ایده‌آل با فرض اینکه مقدار شدت روشنایی فضای جاده یکنواخت و با شدت روشنایی خطوط روی آن متفاوت باشد، می‌توان مرز بین خطوط و جاده (لبه) را پیدا کرد. الگوریتم‌های زیادی در این زمینه وجود دارد که از آن جمله می‌توان به سوبل، کنی، رابرتز و پرویت اشاره کرد. در این رساله از الگوریتم لبه‌یابی کنی استفاده می‌کنیم. شناسایی لبه در

الگوریتم لبه‌یاب کنی شامل شش مرحله است. مرحله‌ی اول فیلتر کردن تصویر اولیه و حذف نویز است که برای این منظور از فیلتر گوسین با یک ماسک ساده استفاده می‌شود. مرحله دوم یافتن لبه‌های قوی با استفاده از دامنه‌گرادیان در هر نقطه است که برای این منظور از ماسک‌های لبه‌یابی و رابطه ۴-۱ استفاده می‌شود.

$$|G| = |G_x| + |G_y| \quad (۱-۴)$$

الگوریتم لبه‌یابی شامل دو ماسک، یکی عمودی و دیگری افقی است که در شکل ۴-۴ نشان داده شده‌اند.

-۱	۰	+۱
-۲	۰	+۲
-۱	۰	+۱

G_x

+۱	+۲	+۱
۰	۰	۰
-۱	-۲	-۱

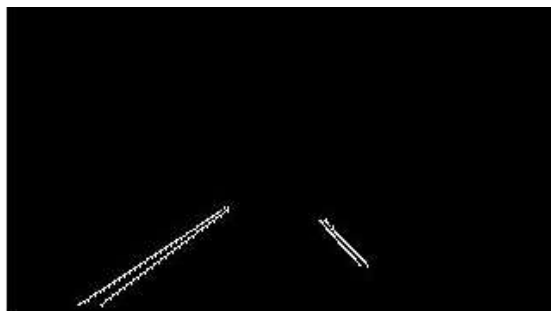
G_y

شکل ۴-۴: ماسک‌های الگوریتم لبه‌یابی

مرحله سوم به‌دست آوردن جهت لبه‌های تصویر به‌دست آمده با استفاده از ماسک‌های تعریف‌شده، به کمک رابطه ۴-۲ است.

$$theta = \arctan\left(\frac{G_y}{G_x}\right) \quad (۲-۴)$$

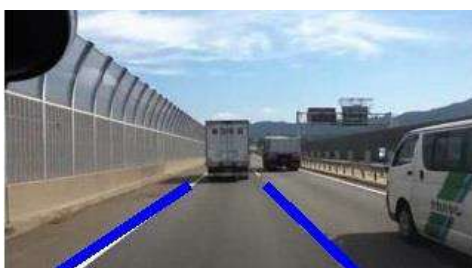
در مرحله چهارم جهت‌های به‌دست آمده در مرحله قبل، به یکی از چهار جهت ۰، ۴۵، ۹۰ و ۱۳۵ درجه تطبیق داده می‌شوند. مرحله پنجم سرکوب لبه‌های غیرحداکثر است. در این مرحله جهت لبه‌ها بررسی می‌شود و لبه‌هایی که نمی‌توان به عنوان لبه در نظر گرفت حذف می‌شوند. مرحله ششم، آستانه‌گیری در تشخیص لبه است برای این منظور دو مقدار آستانه بالا و پایین تعریف می‌شود، هر پیکسلی که دارای گرادیان بیشتر از مقدار آستانه بالا باشد به عنوان لبه پذیرفته می‌شود و در صورتی که مقداری کمتر از مقدار آستانه پایین باشد، رد می‌شود ولی اگر دارای مقداری بین این دو مقدار آستانه باشد در صورتی پذیرفته می‌شود که یکی از همسایه‌های آن پذیرفته شده باشند. با اعمال آشکارساز لبه کنی به تصویر باینری، لبه‌های تصویر را به‌دست می‌آوریم. در شکل ۴-۵ نتیجه اعمال الگوریتم لبه‌یاب کنی بر روی تصویر مرحله قبل قابل مشاهده است.



شکل ۴-۵: اعمال الگوریتم کنی به تصویر

۴-۲-۴ تبدیل هاف

در آخرین مرحله، با استفاده از تبدیل هاف خطوط اصلی جاده را شناسایی می‌کنیم. سپس نزدیکترین خطوط به خودرو از دو طرف به عنوان خطوط جاده در سمت چپ و راست در نظر گرفته می‌شود. تصویر نهایی خروجی در شکل ۴-۶ نشان داده شده است.



شکل ۴-۶: خروجی نهایی الگوریتم تشخیص خطوط جاده با رویکرد کلاسیک (روش پیشنهادی اول)

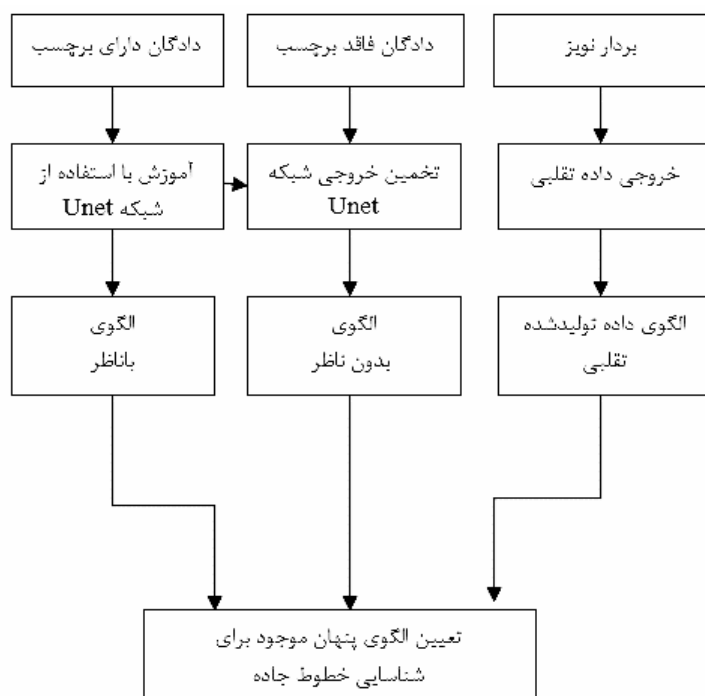
۴-۳ الگوریتم پیشنهادی مبتنی بر یادگیری

روش پیشنهادی دومی که قصد داریم با استفاده از آن شناسایی خطوط درون جاده را انجام دهیم، یک رویکرد مبتنی بر یادگیری است. بدین‌منظور، با به‌کارگیری یک شبکه عمیق و با استفاده از یک الگوی یادگیری نیمه‌نظارتی اقدام به آموزش سیستم شناسایی خطوط می‌نماییم. شبکه عمیق مورد استفاده در روش پیش‌رو، یک شبکه مولد رقابتی نیمه‌نظارتی است. معمولاً این شبکه برای طبقه‌بندی استفاده می‌شود. ما قصد داریم با تغییراتی در ساختار شبکه آن را برای هدف مورد نظرمان که قطع‌بندی است، استفاده نماییم. در واقع یکی از

نوآوری‌های این روش پیشنهادی استفاده از چنین ساختاری برای انجام هدف پیش‌رو است. در ادامه به توضیح روش پیشنهادی خواهیم پرداخت.

۴-۳-۱ شناسایی خطوط با استفاده از شبکه مولد رقابتی نیمه‌نظارتی

بلوک‌دیگرام فرایند قطعه‌بندی خطوط با استفاده از شبکه مولد رقابتی نیمه‌نظارتی در شکل ۴-۷ مشاهده می‌شود.



شکل ۴-۷: فرایند شناسایی خطوط با استفاده از شبکه مولد رقابتی نیمه‌نظارتی

در ادامه به بررسی چگونگی عملیات تشخیص خطوط می‌پردازیم. شبکه مولد رقابتی نیمه‌نظارتی که عملیات شناسایی خطوط را انجام می‌دهد، طبق دیگرام نشان داده شده در شکل ۴-۷ از ۳ بخش اصلی شبکه تفکیک‌کننده (نظارت‌شده)، شبکه تفکیک‌کننده (نظارت‌نشده) و شبکه مولد تشکیل شده است.

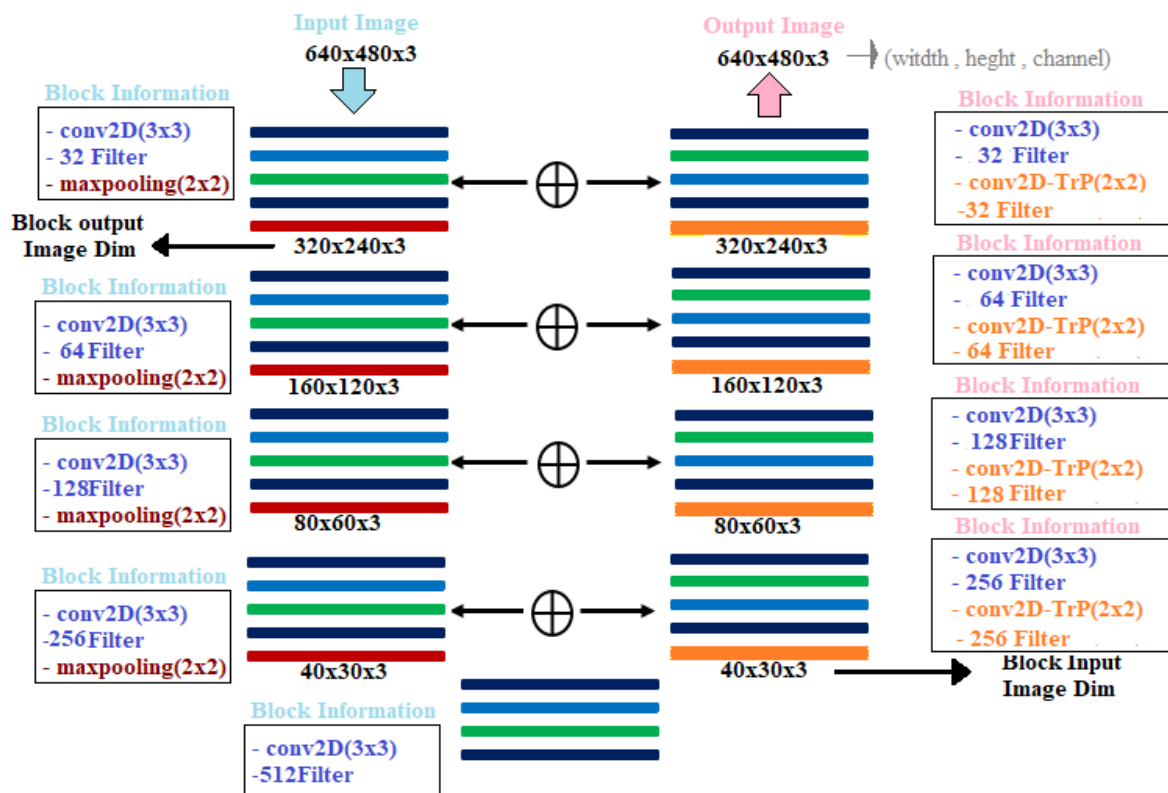
۴-۳-۱-۱ شبکه تفکیک‌کننده (نظارت‌شده)

در این قسمت از شبکه، فرایند آموزش تشخیص خطوط با استفاده از یک شبکه عمیق Unet به‌صورت نظارت‌شده انجام می‌شود. شبکه Unet، یک شبکه عمیق با ساختار پایه‌ای کانولوشن بوده، که دارای دو بخش

اصلی است. اولین بخش این شبکه، تعدادی لایه‌های کانولوشن هستند که طبق معمول فرایند استخراج ویژگی را انجام می‌دهند. بخش دوم در شبکه U-net، تعدادی لایه با عملیات معکوس قسمت اول است. بنابراین اگر به جای لایه تماماً متصلی که در انتهای شبکه کانولوشن معمول برای فرایند طبقه‌بندی تصاویر و یا قطعه‌بندی استفاده می‌شود (FCN)، از یک فرایند معکوس کانولوشن (با توجه به شرایط مقتضی، از تعداد مشخص فرایند کانولوشن معکوس) استفاده کنیم، می‌توانیم فرایند بخش‌بندی تصویر را اجرا نماییم. در واقع شبکه U-net به ترتیب از دو قسمت رمزگذار^۱ و رمزگشا^۲ تشکیل شده است که بخش رمزگذار آن اقدام به شناسایی ویژگی‌ها و استخراج آن‌ها نموده و بخش رمزگشای آن از طریق این ویژگی‌های استخراج‌شده اقدام به ایجاد تصویر قطعه‌بندی شده حاصل از تصویر اصلی از روی خروجی مطلوب که همان ماسک متناظر با تصاویر ورودی است، می‌کند. ساختار شبکه U-net به کار گرفته‌شده در شکل ۴-۸ قابل مشاهده است.

^۱ Encoder

^۲ Decoder



شکل ۴-۸: شبکه Unet استفاده شده در تفکیک کننده

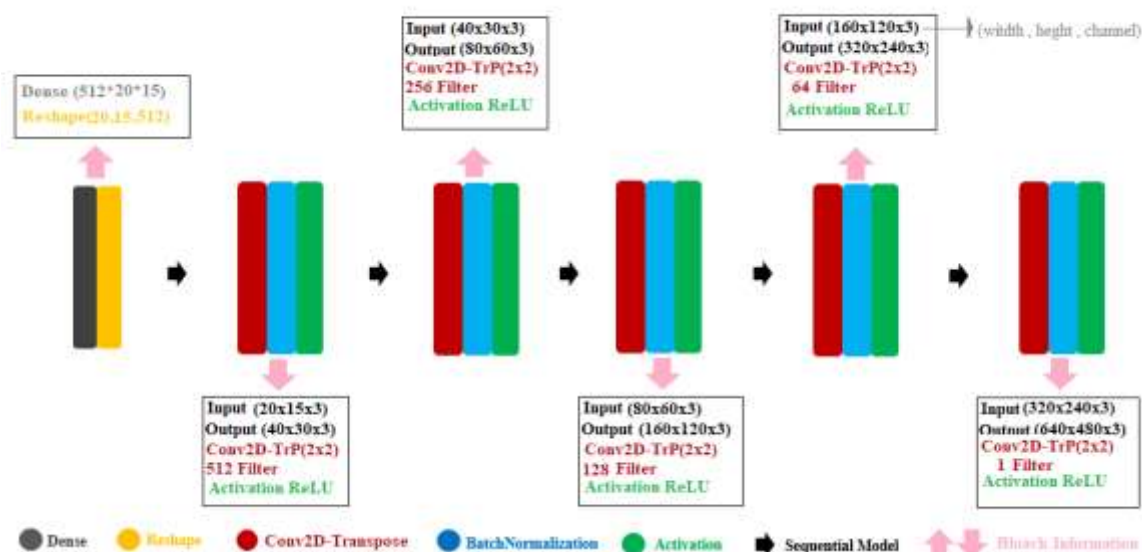
مطابق با آنچه که در شکل ۴-۸ مشاهده می شود، ساختار کلی تفکیک کننده، یک شبکه Unet با ۹ بلاک مشخص است. در قسمت ابتدایی (رمزگذار) شبکه که ۵ بلاک اول شبکه را تشکیل می دهند، به غیر از بلاک پنجم، در هر بلاک، عملیات کاهش ویژگی با استفاده از قرار دادن یک لایه پولینگ انجام می شود. هریک از بلاک های موجود در قسمت رمزگذار و همچنین قسمت دوم (رمزگشا) در شبکه Unet، دو لایه کانولوشن دارند. پارامترهای موجود در لایه کانولوشن، گام حرکت ۱ و حاشیه صفر با حالت same می باشند. عملیات پولینگ بیشینه با اندازه ۲x۲ انجام می شود. برای لایه های BatchNormalization و Relu هربار بعد از کانولوشن می آیند. در کانولوشن معکوس، پارامترهای موجود همانند آنچه در کانولوشن بیان گردید، می باشند. در این شبکه به منظور جلوگیری از بیش برآزش در هنگام آموزش، از dropout با نرخ ۰/۳ در ساختار شبکه استفاده شده است.

۴-۳-۱ شبکه تفکیک کننده (نظارت نشده)

این بخش از شبکه در واقع یک فرایند آزمون داخلی برای قسمت نظارت شده است. هدف اصلی قسمت نظارت نشده شبکه تفکیک کننده، کمک به قسمت نظارت شده برای آموزش بهتر است. چرا که هدف ما از به کارگیری چنین رویکردی، حل چالش کمبود داده‌ها برچسب گذاری شده است.

۴-۳-۱-۳ شبکه مولد

قسمت شبکه مولد با استفاده از یک بردار نویز تصادفی اقدام به تولید تصاویر تقلبی می‌کند. نکته‌ای که باید به آن اشاره نمود، عدم تأثیر و به کارگیری نتایج حاصل از قسمت مولد شبکه در آموزش بخش تفکیک کننده است. بدین معنی که در فرایند آموزش کل سیستم، وزن‌های بخش تفکیک کننده سیستم در هنگام آموزش بخش مولد بدون تغییر هستند. البته این به معنی عدم تأثیر شبکه مولد در فرایند کلی آموزش کل شبکه نیست. اتفاقاً بهینگی عمل کردی این قسمت، تأثیر بسیار مهمی در نتیجه کلی شبکه با هدف ارائه شده دارد. ما در قسمت مولد شبکه خود از یک رمزگشا با لایه‌های کانولوشن معکوس برای ساختن تصویر جعلی استفاده نموده‌ایم. شکل ۴-۹ ساختار شبکه مولد استفاده شده را نشان می‌دهد.



شکل ۴-۹: ساختار شبکه مولد به کار گرفته شده

در ساختار شبکه مولد به کار گرفته شده، همانطور که در شکل ۴-۹ مشخص است، ۵ بلاک اصلی وجود دارد.
مدل

کلی شبکه از نوع ترتیبی^۱ است. در هر لایه کانولوشن معکوس، از گام حرکت^۲ ۲ و همچنین حاشیه صفر^۳ با حالت same استفاده شده است. برای فعالسازی بعد از لایه BatchNormalization، از تابع Relu استفاده شده است. برای لایه کانولوشن نیز پارامترها به همان صورت که برای کانولوشن معکوس بیان شده، استفاده شده‌اند. در نهایت باید بیان کنیم به منظور جلوگیری از مشکل بیش‌برازش در آموزش شبکه، از لایه dropout با نرخ ۰/۳ استفاده نموده‌ایم.

^۱ Sequential
^۲ Stride
^۳ Zero Padding

۴-۱-۳-۴ هایپر پارامترهای مربوط به آموزش شبکه‌ها

۴-۱-۳-۴ شبکه مولد

در شبکه مولد، برای بهینه‌سازی آن از بهینه‌ساز آدام^۱ با نرخ یادگیری ۰/۰۱ استفاده شده است. دلیل استفاده از این بهینه‌ساز، سرعت بالای همگرایی آن نسبت به موارد دیگر همچون گرادیان نزولی تصادفی^۲ است. هایپر پارامتر دیگری که روند آموزش شبکه مولد را با آن بررسی می‌کنیم، خطای^۳ شبکه مولد است. برای خطا از Binary Cross Entropy استفاده شده است.

۴-۱-۳-۴ شبکه تفکیک‌کننده (نظارت‌شده)

برای شبکه تفکیک‌کننده قسمت نظارت‌شده، همچون شبکه مولد، از بهینه‌ساز آدام استفاده شده است. با توجه به این نکته که روند آموزش در این قسمت از شبکه تفکیک‌کننده با استفاده از معیارهای دقت و خطا بررسی می‌شود، ما برای دقت از دو معیار^۴ IOU و دقت (Accuracy) استفاده نموده‌ایم. همچنین برای خطا، از Binary Cross Entropy استفاده شده است.

۴-۱-۳-۴ شبکه تفکیک‌کننده (نظارت‌نشده)

ساختار شبکه تفکیک‌کننده نظارت‌نشده دقیقاً همان ساختار تفکیک‌کننده نظارت‌شده است. تنها تفاوت این قسمت، عدم وجود دادگان برچسب‌دار برای این شبکه است. بنابراین در این قسمت برای بررسی روند آموزش، دیگر معیارهای دقت نداریم. تنها معیار مورد بررسی همان خطا هست، که همچون قسمت‌های (الف) و (ب) از همان خطای بیان‌شده یعنی، Binary Cross Entropy استفاده شده است.

¹ Adam Optimizer

² Stochastic Gradient Descent(SGD)

³ Loss

⁴ Metric

۴-۴ شناسایی موانع

تشخیص مانع فرآیند استفاده از حسگرها، ساختارهای داده و الگوریتم‌ها برای تشخیص اشیای متحرک و غیر متحرکی است که باعث توقف حرکت می‌شوند. همانطور که در فصل اول بیان شد، یکی از مهم‌ترین دلایل استفاده از رانندگی خودکار، کاهش تصادفات و مخاطرات احتمالی در جاده‌ها است. یکی از اصلی‌ترین رویکردهایی که در کاهش مخاطرات احتمالی می‌توان در نظر گرفت، طراحی سیستم‌های شناسایی و هشدار موانع است. در واقع با طراحی سیستم‌های شناسایی مانع، می‌توان زودتر از وقوع یک برخورد احتمالی با مانع، آگاهی پیدا کرد و در نتیجه از آن جلوگیری کرد. از مهم‌ترین سیستم‌های شناسایی موانع که وجود دارند، سیستم شناسایی موانع جاده‌ای است. هدف اصلی شناسایی موانع جاده‌ای، طراحی سیستمی است که بتوان با استفاده از آن، موقعیت خودرو را نسبت به سایر خودروهای درون جاده مشخص نمود. بدین ترتیب می‌توان با ارائه راهکارهای مناسب، حرکت انواع خودروها را پیش‌بینی و از وقوع خطرات پیش‌رو جلوگیری نمود.

شناسایی خودروهای درون جاده در سیستم‌های ترافیکی، بر دو نوع اصلی است. ۱. شناسایی اجسام در حرکت و بررسی مواردی از قبیل نوع خودرو، سرعت خودرو و همچنین شناسایی پلاک که در این سیستم‌ها اغلب از دوربین‌های کنار جاده‌ای استفاده می‌شود. ۲. مشخص نمودن موانع روبروی خودرو که با منبع تصویر داخلی خود خودرو انجام می‌شود. هر کدام از موارد فوق بسته به هدف مورد نظر چالش‌های خود را دارند. هدف ما در این رساله شناسایی مانع (خودروهای درون جاده و روبروی خودرو) است. در ادامه به بیان الگوریتم پیشنهادی خواهیم پرداخت.

۴-۵ الگوریتم کلی شناسایی مانع

بلوک‌دیگرام شکل ۴-۱۰ الگوریتم شناسایی مانع را نشان می‌دهد. همانطور که در شکل مشخص است، تشخیص مانع دارای سه مرحله کلی است. ۱. مرحله استخراج ویژگی برای طبقه‌بندی نوع مانع (ماشین، غیرماشین) ۲. آموزش یک طبقه‌بند برای تشخیص مانع هدف. ۳. انجام عملیات جستجوی مانع مورد نظر با

کمک طبقه‌بند آموزش داده شده، در محدوده روبروی خودروی در حرکت. در ادامه هریک از موارد فوق را توضیح می‌دهیم.



شکل ۴-۱۰: بلوک‌دیگرام شناسایی مانع در تصویر

آنچه به‌عنوان نوآوری در تشخیص مانع میتوان به آن اشاره نمود، تعیین ناحیه مشخص در تصویر با استفاده از ماسک‌گذاری پویا است که باعث کاهش پیچیدگی روش تشخیص مانع می‌شود در ادامه مراحل توضیح داده شده‌اند.

۴-۵-۱ استخراج ویژگی از تصاویر دادگان کمی با استفاده از الگوریتم HOG

با استفاده از مجموعه دادگان که برای دو کلاس خودرو و غیرخودرو در اختیار داریم، ابتدا اقدام به استخراج ویژگی برای تصاویر، با استفاده از الگوریتم HOG می‌نماییم. پایگاه داده‌ای که در اختیار داریم، شامل ۸۷۹۲ تصویر برچسب‌گذاری شده با عنوان (Cars) و ۸۹۶۸ تصویر برچسب‌گذاری شده با عنوان (Non-Cars)

هستند. اندازه هر نمونه برابر با 64×64 و با فضای رنگی RGB است. تعدادی از تصاویر مربوط به دو کلاس در شکل ۴-۱۱ نشان داده شده است.



شکل ۴-۱۱: تعدادی از تصاویر دادگان برای آموزش مدل تشخیص نوع مانع

۴-۵-۲ آموزش طبقه‌بند ماشین بردار پشتیبان برای شناسایی نوع مانع مورد نظر

پس از به‌دست‌آوردن ویژگی‌هایی که برای دادگان دو کلاس محاسبه نمودیم، اقدام به آموزش یک طبقه‌بند ماشین بردار پشتیبان برای شناسایی نوع مانع هدف می‌کنیم. بدین منظور از ویژگی‌های استخراج شده از تصاویر دو کلاس ماشین و غیرماشین، که با استفاده از توصیف‌کننده HOG به‌دست آمدند، استفاده می‌کنیم. به‌طوری که با استخراج ویژگی هریک از تصاویر پایگاه داده موجود یک نگاشت انجام داده و عملیات آموزش بر روی ویژگی‌های نمونه‌ها انجام می‌شود.

نکات مهم در آموزش مدل طبقه‌بند ماشین بردار پشتیبان، به شرح زیر هستند:

۱. ابتدا دادگان استخراج شده از تصاویر که همان ویژگی‌های HOG هستند، را نرمال‌سازی می‌کنیم.
۲. برای فرایند آموزش ۸۰ درصد دادگان را به‌عنوان داده آموزش انتخاب می‌کنیم و سپس عملیات آموزش طبقه‌بند را انجام می‌دهیم.
۳. در فرایند آموزش، در هر تکرار ۲۰ درصد از دادگان آموزش، برای انجام اعتبارسنجی استفاده می‌شود.
۴. فرایند آموزش با ۲۰۰۰ تکرار برای به‌دست آوردن مدل نهایی طبقه‌بند انجام می‌شود.

۴-۵-۳ جستجوی مانع هدف در تصویر

با توجه به آنچه در قسمت قبل بیان شد، جستجوی مانع روبروی خودرو، صرفاً برای پیدا کردن ماشین‌های داخل

جاده و در واقع روبروی خودرو است. روش جستجوی مانع که در این رساله در نظر گرفته‌ایم، یک روش کلاسیک و با دقت بسیار خوب است. البته با توجه به حجم برخی محاسبات، پیچیدگی نسبتاً زیادی دارد. ایده‌ای که در این روش برای بالا بردن سرعت الگوریتم در نظر گرفته شده است، در کنار تعیین یک محدوده به عنوان ناحیه هدف برای جستجوی مانع، استفاده از یک ماسک‌گذاری پویا به منظور افزایش سرعت الگوریتم است. در واقع با این رویکرد، پیچیدگی محاسباتی سیستم کاهش و در نتیجه سرعت افزایش پیدا می‌کند.

۴-۵-۳-۱ لغزش کشویی برای جستجوی مانع

برای این مرحله، با قرار دادن یک ناحیه مشخص در تصاویر موجود و با در نظر گرفتن پنجره‌هایی با اندازه ثابت، اقدام به جستجو و طبقه‌بندی پنجره‌ها خواهیم نمود. بدین منظور طبق مراحل زیر عمل می‌شود:

۱. محدوده مورد انتظار برای جستجوی مانع هدف در تصویر را انتخاب می‌کنیم. در شکل ۴-۱۲ محدوده‌ای که انتظار می‌رود خودرو در آن مشاهده شود، نشان داده شده است.

۲. یک پارامتر همپوشانی برای عملیات جستجو اعمال می‌کنیم. این رویکرد با وجود ازدیاد فضای جستجوی مسئله، باعث می‌شود از یک شی در حرکت، حالات مختلفی داشته باشیم و با استفاده از آن بهینه‌ترین حالت ممکن را پیدا کنیم.



شکل ۴-۱۲: محدوده اصلی جست و جوی خودرو با عملیات لغزش کشویی پنجره

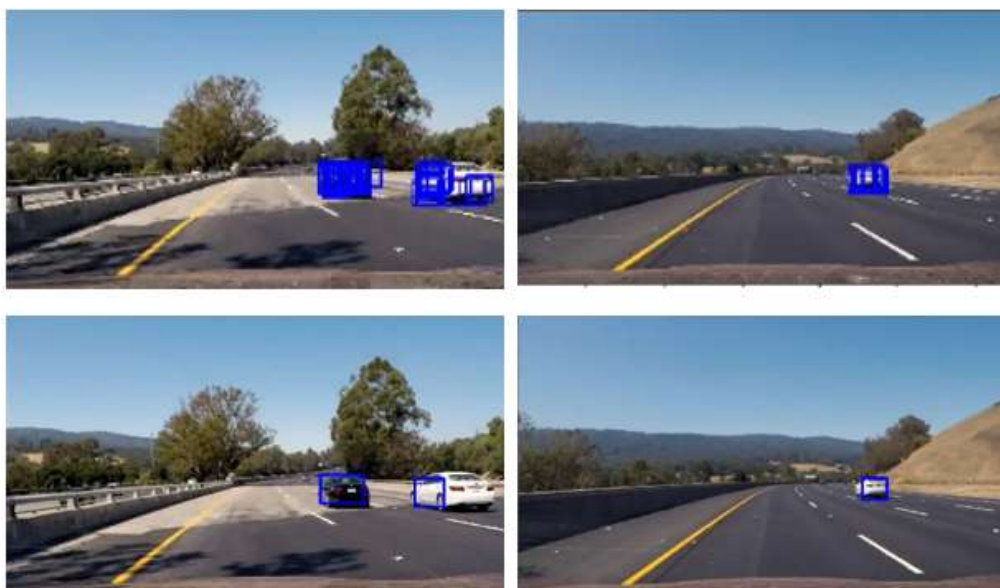
۴-۵-۳-۲ تعیین کلاس هر پنجره برای شناسایی مانع

بدین منظور، برای هر کدام پنجره، یک عملیات تغییر اندازه و نگاشت به یک اندازه 64×64 انجام می‌دهیم. عملیات

استخراج ویژگی HOG را انجام داده و با استفاده از مدل آموزش داده شده، تعیین می‌کنیم که آیا پنجره تعیین شده جز هدف است یا خیر.

۴-۵-۳-۳ تعیین پنجره بهینه شناسایی مانع

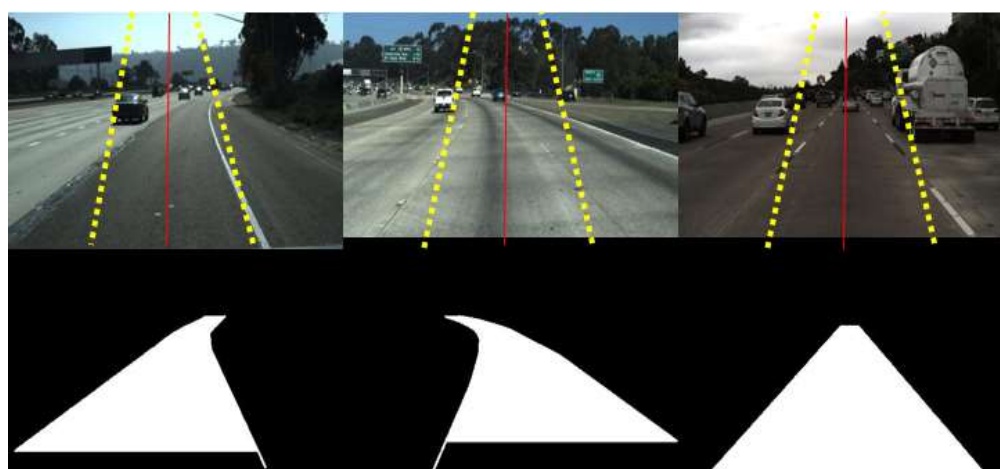
در فرایند لغزش پنجره‌ها در یک فریم، ممکن است برای هر مانع بیشتر از یک پنجره به دست آورده شود. لذا پس از لغزش، پنجره‌هایی که اشتراک دارند (آن‌هایی که با کلاس مانع هدف شناسایی شده‌اند) بررسی و پنجره بهینه با بالاترین امتیاز تعیین می‌شود. بدین معنی که بهترین پنجره، از بین تمام پنجره‌هایی که یک مانع در آن قرار گرفته است، انتخاب می‌شود.



شکل ۴-۱۳ شناسایی پنجره بهینه در هر فریم و پنجره‌های دارای همپوشانی

۴-۵-۴ تعیین موانع در محدوده روبروی خودرو

محل قرارگیری دوربین در تصویر، همواره در مرکز آن قرار دارد. این نکته در شکل ۴-۱۴ مشخص است. بنابراین با کوچکتر کردن محدوده تصویر، موانعی را که تعیین شده‌اند، برای محدوده تعیین شده انتخاب می‌کنیم.



شکل ۴-۱۴: تعیین محدوده فرضی مقابل ماشین با توجه به ساختار تصاویر و محل قرارگیری دوربین

۴-۶ نتیجه‌گیری

در این فصل به ارائه دو روش پیشنهادی جهت تشخیص خطوط و روشی برای شناسایی موانع پرداختیم. در ابتدا رویکرد کلاسیک با به‌کارگیری الگوریتم‌های پایه‌ای پردازش تصویر و سپس رویکردی مبتنی بر یادگیری عمیق و استفاده از شبکه مولد رقابتی نیمه‌نظارتی را توضیح دادیم و در انتها نیز به بیان چگونگی شناسایی موانع درون جاده و روبروی خودرو پرداختیم. بهبودهای صورت‌گرفته در هریک از قسمت‌های روش‌های پیشنهادی تشخیص خطوط و همچنین تشخیص موانع را بیان نمودیم. در فصل بعد به ارائه نتایج حاصل از به‌کارگیری هرکدام از بخش‌های بیان شده خواهیم پرداخت.

فصل پنجم

نتایج

۵-۱ مقدمه

در فصل قبل دو روش پیشنهادی، یکی با رویکرد کلاسیک مبتنی بر دستورات پایه‌ای پردازش تصویر و دیگری رویکردی مبتنی بر یادگیری عمیق و با استفاده از شبکه مولد رقابتی نیمه‌نظارتی برای شناسایی خطوط، ارائه نمودیم. در پایان نیز به شناسایی موانع روبروی خودرو با استفاده از تحلیل اجسام متحرک با استفاده از شناسایی آن به روش مدل مخلوط گوسی پرداختیم.

برای فصل پیش‌روی، در سه قسمت به بیان خروجی‌های روش‌های پیشنهادی و همچنین نتایج کمی و کیفی حاصل از روش‌های به‌کارگیری شده خواهیم پرداخت. بدین منظور، در ابتدا برای قسمت‌هایی که نیاز به بررسی نتایج داریم، به تعریف و بیان پارامترهایی همچون دقت عمل‌کردی پرداخته و با جزئیات کامل توضیح می‌دهیم. در انتها یک جمع‌بندی از عمل‌کرد روش‌های پیشنهادی انجام‌شده، خواهیم داشت.

در عملیات پیاده‌سازی، برای قسمت اول از محیط نسخه Matlab 2018b با استفاده از سیستم "2.5GHz intel core i7 Processor with 1600 DDR4 8G RAM" و برای قسمت‌های دوم و سوم کار، که روش پیشنهادی مبتنی بر یادگیری عمیق و همچنین تشخیص موانع هستند، از برنامه‌نویسی پایتون در محیط گوگل کولب استفاده شده است.

۵-۲ معرفی دادگان

پایگاه‌داده‌های مختلفی برای تشخیص خطوط جاده وجود دارد و بیشتر محققان در این زمینه از پایگاه‌های داده متفاوتی استفاده کرده‌اند. در این تحقیق برای اینکه بتوان نتایج روش پیشنهادی را با روش‌های دیگر مقایسه کرد و شرایط مختلف جاده‌ها در نظر گرفته شود، سه پایگاه داده جامع استاندارد استفاده کردیم. از طرف دیگر باتوجه به اینکه پایگاه داده داخلی عامی جهت بررسی نتایج وجود ندارد، یک پایگاه داده داخلی با نصب دوربین در درون خودرو تهیه نمودیم.

۵-۲-۱ پایگاه داده IROADS

این پایگاه داده‌ها [۸۱] شامل ۷ مجموعه پایگاه داده‌ها بوده که مجموعاً دارای ۴۷۰۰ رشته تصویر است. این پایگاه داده‌ها یک پایگاه کامل و جامع بوده به‌طوری که تقریباً تمام شرایط ممکن در جاده‌ها شامل روشنایی روز، تاریکی شب، نور بیش از اندازه خورشید، نور داخل تونل‌ها، بارندگی در روز و شب و جاده‌های برفی را دربر می‌گیرد. ابعاد هر فریم در این پایگاه داده ۶۴۰×۳۶۰ است. نمونه‌هایی از تصاویر این پایگاه داده در شکل ۵-۱ مشاهده می‌شود.



شکل ۵-۱: نمونه‌هایی از تصاویر پایگاه داده‌ها IROADS [۸۱]

۵-۲-۲ پایگاه داده Aly's Method

این پایگاه داده‌ها [۸۲] شامل ۴ مجموعه پایگاه داده‌ها به نام‌های Washington1, Cordova2, Cordova1 و Washington2 است. هرکدام از مجموعه‌ها دارای ۱۰ ویدیو با طول ۱ ثانیه بوده که شامل ۳۰ فریم هستند و در مجموع شامل ۱۲۰۰ تصویر از جاده در فضای شهری است. این پایگاه داده در اکثر تحقیقات انجام شده مورد استفاده قرار گرفته است و از جامعیت مناسبی برخوردار است. همچنین تصاویر این پایگاه داده‌ها دارای ساختمان‌ها، درختان و وسایل نقلیه متعددی بوده و علاوه بر خطوط مورد نظر برای شناسایی، خطوط متفرقه زیادی دارد. اندازه هر فریم در این پایگاه داده ۶۴۰×۳۶۰ است. شکل ۵-۲ نمونه‌هایی از تصاویر این پایگاه داده را نشان می‌دهد.



شکل ۵-۲: نمونه‌هایی از تصاویر پایگاه داده‌ها Aly's Method [۸۲]

۵-۲-۳ پایگاه داده Tusimple

این مجموعه داده شامل ۳۶۲۶ نمونه آموزش و ۲۷۸۲ نمونه آزمون است. هر نمونه شامل ۱ ثانیه ویدیو (۲۰ فریم) هست که متناظر با فریم ۲۰ ام یک ماسک خطوط وجود دارد. اندازه تصاویر برابر با 1280×720 است. این مجموعه داده را در قسمت تفکیک‌کننده نظارت‌شده شبکه عمیق استفاده نمودیم. همچنین از دادگانی که خودمان تهیه نمودیم، به‌عنوان مجموعه داده بدون برچسب در بخش نظارت‌نشده استفاده نمودیم.



شکل ۵-۳: تعدادی از تصاویر پایگاه داده‌ها Tusimple

۵-۲-۴ پایگاه داده تهیه‌شده داخلی

این پایگاه داده‌ها شامل تعداد ۳۰ ویدیو با زمان‌های متفاوتی است که از تعدادی جاده‌های بین‌شهری و شهری

تهیه شده است. از ویدیوهای در دسترس، ۵۸۰۰ تصویر استخراج شده است. اندازه هر فریم از ویدیوها برابر با ۷۲۰×۱۲۰۰ است. از قسمتی از این پایگاه داده‌ها، برای آموزش سیستم در روش پیشنهادی مبتنی بر یادگیری به‌عنوان دادگان بدون برچسب استفاده نموده‌ایم. تعدادی از تصاویر این مجموعه داده در تصویر ۴-۵ قابل مشاهده است.



شکل ۴-۵ : تصاویر پایگاه داده‌ها تهیه‌شده داخلی

۳-۵ نتایج حاصل از روش پیشنهادی اول (رویکرد کلاسیک مبتنی بر دستورات پایه‌ای پردازش تصویر)

۱-۳-۵ معیار سنجش کیفی شناسایی خطوط در روش پیشنهادی اول (نرخ تشخیص)

با توجه به اینکه در پایگاه داده‌های IROADS و Aly's Method خروجی مطلوب متناظر برای ارزیابی فرایند تشخیص خطوط، وجود ندارد، لذا جهت ارزیابی روش پیشنهادی از یک سنجش کیفی استفاده نمودیم. معادله بیان‌شده در رابطه ۱-۵ کلیت عملیات شناسایی خطوط را در روش پیشنهادی اول مورد سنجش و ارزیابی قرار می‌دهد.

$$DetectionRate(DR) = \frac{CorrectedRecognitionFrame(C)}{TotalFrame(T)} * 100 \quad (۱-۵)$$

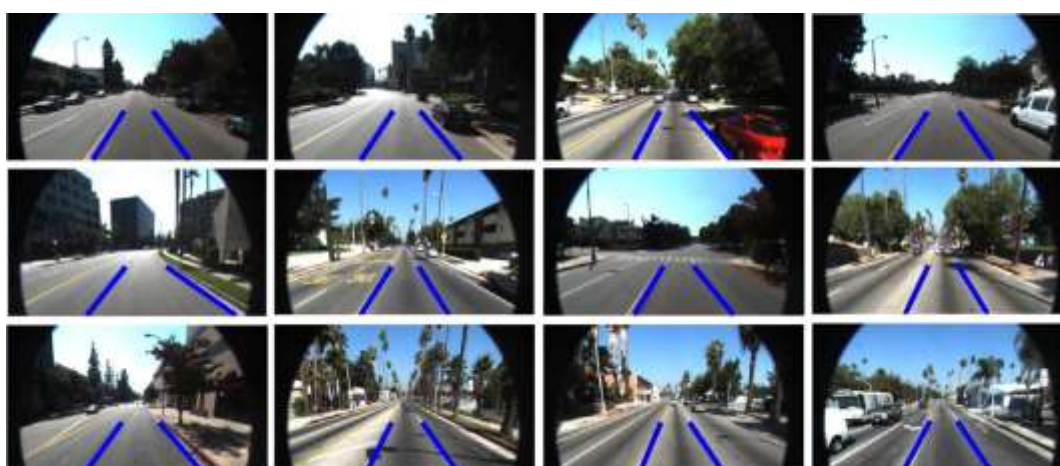
در رابطه ۵-۱ مقدار پارامتر C برابر با تعداد فریم‌هایی است که در آن‌ها، خطوط جاده در دوطرف خودرو به‌صورت صحیح تشخیص داده شده‌اند. لذا برای چنین فریم مورد مشخصی امتیاز ۱ لحاظ می‌شود و از طرفی دیگر اگر به‌ازای هر فریم، الگوریتم پیشنهادی نتواند هر دو خط درون جاده را تشخیص دهد و یا در تشخیص هریک از خطوط دارای خطا باشد، امتیاز صفر برای آن فریم در نظر گرفته می‌شود. پارامتر T نشان‌دهنده مجموع تعداد فریم‌هایی است که در آزمایش تشخیص خطوط مورد بررسی قرار داده‌ایم.

۵-۳-۲ نتایج بصری روش پیشنهادی اول

شکل‌های ۵-۵ و ۵-۶ به‌ترتیب تعدادی از نتایج حاصل از اعمال الگوریتم پیشنهادی بر روی پایگاه داده IROADS و پایگاه داده Aly's Method را نشان می‌دهند. همانطور که در این شکل‌ها مشخص است برای تمامی تصاویر موجود، خطوط جاده به‌درستی تشخیص داده شده‌اند. نکته‌ای که در اینجا می‌توان به آن اشاره کرد این است که درستی تشخیص یک پارامتر کیفی است. بدین معنی که تصویر از نظر بیننده دقیقاً خطوط دو طرف خودرو را به درستی تشخیص داده است. نکته‌ی دیگری که در نتایج شکل ۵-۵ و ۵-۶ دیده می‌شود، عمل کرد کیفی الگوریتم برای شناسایی خطوط جاده در تصاویر جاده‌ها با شرایط محیطی مختلف از جمله، داخل تونل با روشنایی کم، هوای برفی و بارانی، هوای صاف دارای روشنایی مناسب و نهایتاً جاده‌های دارای سایه است.



شکل ۵-۵: تعدادی از نتایج آزمایشات روش پیشنهادی اول بر روی پایگاه داده‌ها IROADS با خروجی مطلوب

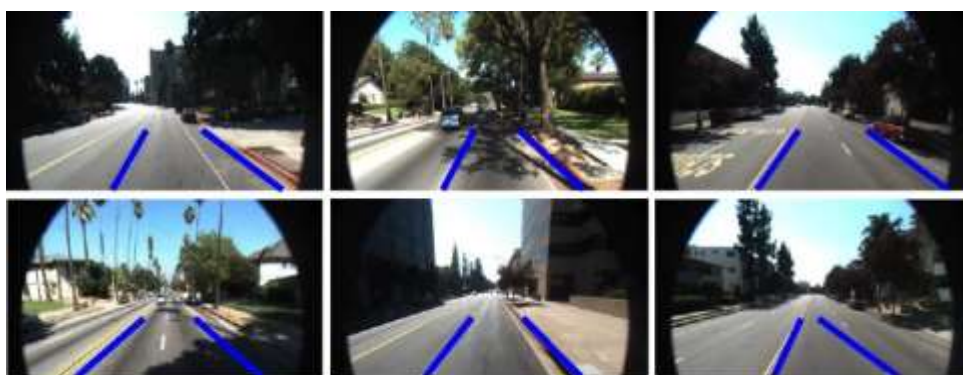


شکل ۵-۶: تعدادی از نتایج آزمایشات روش پیشنهادی اول بر روی پایگاه داده‌ها Aly's Method با خروجی مطلوب

شکل‌های ۵-۷ و ۵-۸ به ترتیب تعدادی از نتایج پیاده سازی الگوریتم پیشنهادی تشخیص خطوط جاده بر روی پایگاه داده‌ها IROADS و پایگاه داده‌ها Aly's Method که دارای خطا می‌باشند را نشان می‌دهند. همانطور که در این شکل‌ها مشاهده می‌شود، برای تمام موارد شناسایی خطوط، حداقل یکی از دو خط جاده در دوطرف خودرو دارای خطا است. این خطاها در مکان تخمین خط و زاویه تشخیص خط قابل مشاهده هستند.



شکل ۵-۷: تعدادی از تصاویر دارای خطا در تشخیص خطوط برای پایگاه داده‌ها IROADS



شکل ۵-۸: تعدادی از تصاویر دارای خطا در تشخیص خطوط برای پایگاه داده‌ها Aly's Method

۵-۳-۳ دقت روش پیشنهادی اول

حال با توجه به آنچه که با عنوان معیارهای تشخیص درست خطوط و یا خطا در تشخیص خطوط معرفی شدند، نتایج حاصل از روش پیشنهادی اول را در ادامه بررسی می‌کنیم.

پایگاه داده IROADS در مجموع دارای ۲۸۲۸ فریم است. این فریم‌ها همانطور که در شکل‌های ۵-۵ و ۵-۷ مشاهده شدند، در شرایط محیطی و آب و هوایی متفاوتی بودند. نتایج حاصل از الگوریتم پیشنهادی برای این پایگاه داده به تفکیک شرایط محیطی در جدول ۵-۱ مشاهده می‌شود. باتوجه به جدول ۵-۱ میانگین نرخ تشخیص خطوط برای تصاویر پایگاه داده‌ها IROADS ۹۶/۷۸ درصد بوده و زمان متوسط مورد نیاز برای پردازش هر فریم ۲۸ میلی ثانیه است. بنابراین الگوریتم پیشنهادی به‌صورت بلادرنگ تشخیص خطوط جاده را با دقت بالا انجام می‌دهد.

جدول ۵-۱: نتایج حاصل از روش پیشنهادی اول روی پایگاه داده‌ها IROADS

مجموعه داده IROADS	مجموع تعداد فریم	تعداد فریم‌های دارای خطا در تشخیص	نرخ تشخیص (درصد)	تعداد فریم بر ثانیه
DayLight	۹۰۳	۳۲	۹۶/۴۵	۳۵/۲
RainyDay	۱۰۴۹	۷	۹۹/۳۳	۳۵/۶
SnowyDay	۵۶۹	۳۱	۹۴/۵۵	۳۶/۲
Tunnel	۳۰۷	۲۱	۹۳/۱۶	۳۵/۸
مجموع	۲۸۲۸	۹۱	۹۶/۷۸	۳۵/۷

اما در مقایسه با فعالیت‌های پژوهشی دیگری که بر روی مجموعه داده IROADS انجام گرفته شده است، روش پیشنهادی ارائه شده در این رساله از نتایج قابل قبولی برخوردار است. جدول ۵-۲ مقایسه روش پیشنهادی در این رساله را با تعدادی از پژوهش‌های انجام شده نشان می‌دهد.

جدول ۵-۲: مقایسه روش پیشنهادی اول با تعدادی از پژوهش‌های مرتبط دیگر بر روی پایگاه داده‌ها IROADS

روش	معیار ارزیابی	DayLight ۹۰۳	RainyDay ۱۰۴۹	SnowyDay ۵۶۹	Tunnel ۳۰۷	مجموع ۲۸۲۸
Son's Method[۸۳]	نرخ تشخیص (درصد)	۹۵/۷۹	۸۸/۲۷	۸۶/۶۴	۹۳/۵۱	۹۰/۸۱
Cao's Method[۸۴]	نرخ تشخیص (درصد)	۹۶/۱۲	۹۳/۴۲	۹۲/۴۴	۹۴/۱۴	۹۴/۱۷
Hu's Method[۸۵]	نرخ تشخیص (درصد)	۹۶/۲۳	۹۴/۷۵	۹۳/۶۷	۹۴/۴۶	۹۴/۹۹
روش پیشنهادی	نرخ تشخیص (درصد)	۹۶/۴۵	۹۹/۳۳	۹۴/۵۵	۹۳/۱۶	۹۶/۷۸

طبق آنچه در جدول ۵-۲ قابل مشاهده است، روش پیشنهادی در این رساله نسبت به ۳ پژوهش صورت گرفته در [۸۳-۸۵] در اکثر شرایط نتایج بهتری در شناسایی خطوط ارائه داده است.

پایگاه داده Aly's Method در مجموع دارای ۱۲۲۵ فریم است. این فریم‌ها همانطور که در شکل‌های ۵-۶ و ۵-۸ مشاهده شدند، در شرایط جاده‌های مختلف درون شهری بودند. نتایج حاصل از اعمال الگوریتم پیشنهادی روی پایگاه داده Aly's Method به تفکیک حالت‌های مختلف جاده‌های شهری در جدول ۵-۳ مشاهده می‌شود.

باتوجه به جدول ۳-۵ میانگین نرخ تشخیص خطوط برای تصاویر پایگاه داده‌ها Aly's Method ۸۹/۴۷ درصد بوده و زمان متوسط مورد نیاز برای پردازش هر فریم ۲۹/۵ میلی ثانیه است. بنابراین الگوریتم پیشنهادی به صورت بلادرنگ تشخیص خطوط جاده را با دقت مناسب انجام می‌دهد. دلیل کاهش دقت تشخیص روش پیشنهادی بر روی پایگاه داده‌ها Aly's Method نسبت به پایگاه داده IROADS این است که تصاویر مربوطه به پایگاه داده‌ها Aly's Method از جاده‌های درون شهری بوده و این تصاویر دارای چالش‌های بیشتری مانند وجود سایه ساختمان‌ها و تقاطع‌ها هستند.

جدول ۳-۵: نتایج حاصل از روش پیشنهادی اول روی پایگاه داده‌ها Aly's Method

مجموعه داده Aly's Method	مجموع تعداد فریم	تعداد فریم‌های دارای خطا در تشخیص	نرخ تشخیص (درصد)	تعداد فریم بر ثانیه
Cordova1	۲۵۰	۴۸	۸۰/۸	۳۴/۳
Cordova2	۴۰۶	۲۳	۹۴/۳۳	۳۴/۱
Washington1	۳۳۷	۵۵	۸۳/۶۸	۳۳/۴
Washington2	۲۳۲	۳	۹۸/۷۱	۳۳/۶
مجموع	۱۲۲۵	۱۲۹	۸۹/۴۷	۳۳/۸۵

۴-۵ نتایج حاصل از روش پیشنهادی دوم (رویکرد مبتنی بر یادگیری عمیق)

در روش پیشنهادی دوم، با توجه به اینکه ما برای عملیات پیاده‌سازی از یک رویکرد نیمه‌نظارتی استفاده می‌کنیم،

می‌توانیم با توجه به برچسب‌های موجود، معیارهای کمی، برای سنجش عملکرد، روش خود داشته باشیم. بدین معنی که با استفاده از خروجی مطلوب و خروجی به‌دست آمده در برنامه، ارزیابی عملیات تشخیص خطوط را با استفاده از پارامترهای تعریف‌شده، در خود برنامه محاسبه کنیم. در فصل سوم دو معیار اصلی که برای ارزیابی در اکثر فرایندهای مربوط به قطعه‌بندی مورد استفاده قرار می‌گیرند را بیان کردیم، در این قسمت برای ارزیابی نتایج روش پیشنهادی از آن‌ها استفاده می‌کنیم. سپس نتایج حاصل از شناسایی خطوط را بیان می‌کنیم.

۵-۴-۱ ارزیابی بصری روش پیشنهادی دوم

۵-۴-۱-۱ خروجی شبکه مولد

بخش مولد شبکه رقابتی نیمه نظارتی، با تولید مجموعه داده و شرکت در قسمت اصلی آموزش شبکه همراه با خروجی بخش نظارت نشده قسمت تفکیک کننده، نقش مهمی در همگرایی شبکه به یک مقدار پایدار دارد. شکل ۵-۹ تعدادی از تصاویر تولید شده توسط شبکه مولد را در حین فرایند آموزش نشان می دهد. تصاویر a,b,c,g,h,i تصاویری هستند که توسط شبکه مولد ایجاد شده اند. برای تصاویر a,b,c عمل کرد شبکه مولد در تولید تصاویر، در وضعیت بسیار مناسبی است که تصاویر d,e,f نشان دهنده این موضوع هستند. این تصاویر، در مجموعه داده ها استفاده شده وجود دارند. از طرفی برای تصاویر g,h,i تولید شده نسبت به آنچه به عنوان تصاویر شبیه در موارد j,k,l آمده اند، از کیفیت نسبی پایین تری برخوردار هستند، اما با این وجود کلیت عمل کرد شبکه مولد، متناسب با هدف مسئله خوب است.



شکل ۵-۹: تعدادی از تصاویر تولیدشده با به کارگیری شبکه مولد و تصاویر شبیه متناظر با آن‌ها در مجموعه داده اصلی

۵-۴-۱-۲ شناسایی خطوط با استفاده از قسمت تفکیک کننده

فرایند اصلی استخراج خطوط جاده، در روش پیشنهادی دوم، با استفاده از شبکه مولد رقابتی نیمه نظارتی، در بخش تفکیک کننده نظارت شده آن انجام می شود. ما برای شناسایی خطوط در قسمت نظارت شده از یک شبکه عمیق Unet استفاده نمودیم. در ادامه تعدادی از خروجی های شبکه بر روی تعدادی از تصاویر مورد نظر را مشاهده می کنیم. شکل های ۵-۱۰ و ۵-۱۱ به ترتیب نتایج مطلوب و نتایج دارای خطا از نحوه شناسایی خطوط برای الگوریتم پیشنهادی است. همچنین شکل ۵-۱۲ خروجی شناسایی خطوط با کمک یک مدل نظارت شده با استفاده از شبکه عمیق با معماری ResNet50 را نشان می دهد. همانطور که مشخص است، کیفیت خروجی در

روش پیشنهادی ما با وجود رویکرد نیمه نظارتی، نسبت به روش نظارت شده با استفاده از شبکه عمیق با معماری ResNet50 از عمل کرد خوبی برخوردار بوده است.



شکل ۵-۱۰: عمل کرد مطلوب شناسایی خطوط از به کارگیری شبکه عمیق مولد رقابتی نیمه نظارتی



شکل ۵-۱۱: تعدادی از خروجی های با عمل کرد نامطلوب در شناسایی خطوط در روش پیشنهادی مبتنی بر یادگیری عمیق



شکل ۵-۱۲: نمونه خروجی شناسایی خطوط با استفاده از ResNet50

۵-۴-۲ دقت روش پیشنهادی دوم (روش مبتنی بر یادگیری عمیق)

نتایج حاصل از دقت روش پیشنهادی دوم در جدول ۵-۴ مشاهده می‌شود. این دقت مختص به بخش نظارت‌شده شبکه برای دادگان آزمون، که دارای برچسب (ماسک) متناظر می‌باشند، است که مقادیر دقت پیکسلی ACC و میانگین IOU در آموزش شبکه برای این دو بخش محاسبه می‌گردد. منظور از دقت پیکسلی در اینجا، کلاسی است که پیکسل‌ها براساس احتمال خروجی تابع sigmoid خواهند گرفت. دقت پیکسلی که با Accuracy مشخص شده است، همانطور که در ۵-۴-۱ بیان گردید، نسبت تعداد نقاط (پیکسل‌های) درست تعیین شده به کل تعداد پیکسل‌های خطوط موجود در تصویر ماسک متناظر (مرجع صحیح) است. با توجه به جدول، اختلاف نسبت درصد دقت پیکسلی روش پیشنهادی نسبت به سایر روش‌ها اشاره شده، اختلاف ناچیزی است، چنانکه شکل‌های ۵-۱۰ و ۵-۱۲ خروجی روش پیشنهادی در مقایسه با روش ارائه شده در شبکه ResNet50 قابل مشاهده است. اما مقدار کمی معیار IOU نسبت به روش‌های مقایسه شده اختلاف نسبتاً قابل توجهی با روش‌های نظارت‌شده دارد. از طرفی نیز میزان پیچیدگی محاسباتی روش پیشنهادی با توجه به پارامترهای کل شبکه ارائه شده، به طور محسوسی کاهش یافته است.

جدول ۵-۴: مقادیر دقت شناسایی مبتنی بر پیکسل (ACC) و معیار IOU

روش	پارادایم	dataSet تصاویر آموزش برچسب‌دار	ACC	IOU	پیچیدگی پارامترهای آموزش
AlexNet	نظارت‌شده	TuSimple ۳۶۳۶	۹۶/۱۱	۰/۸۴	۶۰ میلیون
VGG16	نظارت‌شده	TuSimple ۳۶۳۶	۹۶/۳۳	۰/۸۵	۱۳۸ میلیون
LaneNet	نظارت‌شده	TuSimple ۳۶۳۶	۹۵/۳۸	۰/۸۱	۱۰۰ میلیون
ResNet50	نظارت‌شده	TuSimple ۳۶۳۶	۹۶/۳۹	۰/۸۶	۴۶ میلیون
روش پیشنهادی S-GAN	نیمه نظارتی	TuSimple ۱۲۱۰	۹۰/۷۳	۰/۶۷	۲۱ میلیون

در جدول ۵-۴، همانطور که مشخص است، روش پیشنهادی با تعداد دادگان برچسب‌دار کمتر (در حدود یک‌سوم) نسبت به روش‌های دیگر و همچنین پارامترهای آموزشی کمتر در مدل پیشنهادی نسبت به آن‌ها از لحاظ دقت پیکسلی (Accuracy) در وضعیت مناسبی قرار دارد. چنانچه همانطور که در شکل‌های ۵-۱۰ و ۵-۱۲ نیز مشاهده گردید، عملیات شناسایی خطوط دو روش بسیار به همدیگر نزدیک هستند.

۵-۵ شناسایی مانع

در شناسایی مانع، به دلیل عدم وجود ماسک مورد نظر برای سنجش پارامتریک، به شکل بصری اقدام به بررسی دقت حاصل از روش پرداخته‌ایم. برای هر تصویر با بررسی و شمارش تعداد تصاویری که موانع درست تشخیص داده شده‌اند به هر تصویر امتیاز ۱ و یا صفر می‌دهیم. شکل ۵-۱۳ تعدادی از موانع تشخیص داده شده در محدوده مورد نظر را نشان می‌دهد. ما این بخش ۴۶۰ تصویر (در غالب ۲۳ ویدیو) داشتیم که دقت تشخیص مانع با توجه به رویکرد انتخابی برابر با ۸۱ درصد به‌دست آمده است. در مورد درصد به‌دست آمده ذکر این نکته حائز اهمیت است که نرخ تشخیص خطای مانع با وجود مقدار بسیار پایینی که در هریک از فریم‌ها دارد، با حتی یک تشخیص نادرست احتمالی (برای مثال قسمتی از یک ماشین خارج از محدوده روبرویی خودرو تشخیص داده شود) مقدار آن فریم برابر با صفر قرار داده می‌شود. لذا مقدار درصد به‌دست آمده، در صورت محاسبه تعداد موانع موجود در هر فریم، بسیار بالاتر از مقدار ذکر شده خواهد بود. ما با استفاده از معیار تشخیص صحیح خودرو در فریم، به دقت ۹۲/۳ درصد رسیدیم. در جدول ۵-۵ تعدادی از فعالیت‌هایی که فرایند تشخیص موانع در آن‌ها انجام شده است، بیان گردیده شده‌اند.

جدول ۵-۵: مقادیر دقت شناسایی تعدادی از پژوهش‌های مرتبط

پژوهش انجام شده	روش اجرایی / نظارت شده، نظارت نشده	دقت
جانگ بائه [۸۶]	استفاده از شبکه عمیق RCNN / نظارت شده	۸۸/۲
الجعفری و همکاران [۸۷]	به کارگیری تشخیص ویژگی GIST به علاوه ردیابی با استفاده از فیلتر کالمن / نظارت نشده	۹۴/۹۸
نویتا هاسیبوآن و همکاران [۸۸]	به کارگیری روش تشخیص YOLO با استفاده یا یادگیری عمیق به منظور شناسایی نوع خودرو / نظارت شده	۸۷/۹۸
روش پیشنهادی (تعداد قاب‌ها با تشخیص نادرست)	به کارگیری HOG و طبقه‌بند SVM، به علاوه تعیین محدوده مورد روبرویی خودرو با استفاده از ماسک‌گذاری پویا / نظارت نشده	۸۱
روش پیشنهادی (تشخیص صحیح خودرو)	به کارگیری HOG و طبقه‌بند SVM، به علاوه تعیین محدوده مورد روبرویی خودرو با استفاده از ماسک‌گذاری پویا / نظارت نشده	۹۲/۳



شکل ۵-۱۳: نمونه خروجی تشخیص موانع

۵-۶ نتیجه‌گیری

در این بخش به بیان نتایج حاصل از به‌کارگیری روش‌های پیشنهادی پرداختیم. در روش اول که با استفاده از یک

رویکرد کلاسیک اقدام به شناسایی خطوط داشتیم، نتیجه حاصل از به‌کارگیری روش با توجه به معیاری که تعریف نمودیم، در حد قابل قبولی بود. در روش دوم که با توجه به رویکرد یادگیری عمیق به شناسایی خطوط با استفاده از الگوی نیمه‌نظارتی پرداختیم، به سه نتیجه کلی رسیدیم:

۱. روش مبتنی بر یادگیری عمیق، برای ما نسبت به روش رویکرد کلاسیک این امکان را فراهم می‌آورد که بررسی دقیق‌تری برای عملیات خود داشته باشیم.

۲. روش مبتنی بر یادگیری عمیق که به صورت نیمه‌نظارتی اقدام به شناسایی خطوط نموده است، از لحاظ پیچیدگی محاسباتی نسبت به معماری‌های مرسوم که برای شناسایی خطوط عمل می‌کنند، با توجه به تعداد پارامترهای نیازمند آموزش در شبکه به‌کار گرفته‌شده، دارای وضعیت بهتری است. این مسئله در جدول مقایسه با روش‌های دیگر نشان داده‌شد.

۳. دقت روش مبتنی بر یادگیری عمیق، که در روش پیشنهادی دوم بیان گردید، نسبت به روش پیشنهادی اول بهتر و نسبت به رویکردهای نظارت‌شده موجود در شبکه‌های نظارت‌شده، کمی دقت پایین‌تری داشته است.

در شناسایی موانع که با استفاده از روش شناسایی HOG بعلاوه طبقه‌بند ماشین بردار پشتیبان برای شناسایی نوع موانع هدف انجام گردید، دقت عمل‌کردی روش بر طبق آنچه که در نتایج خروجی مشخص شد، عمل‌کرد مطلوبی از نظر شناسایی موانع روبرو خودرو نشان می‌دهد. اگر چه روش شناسایی موانع باوجود داشتن سرعت مناسب، همچنان جا برای بهبود دارد.

فصل ششم

جمع بندی

شناسایی خطوط جاده‌ای یکی از مهم‌ترین سیستم‌هایی است که در فرایند رانندگی خودکار مطرح می‌شود. این فرایند به منظور مشخص‌سازی خطوط جاده جهت کنترل حرکت ماشین در یک خط مشخص در جاده صورت می‌گیرد و می‌تواند برای جلوگیری از تصادفات مؤثر باشد. شنایایی خطوط امروزه با استفاده از روش‌های مختلفی مورد بررسی قرار می‌گیرد. مهم‌ترین لازمه‌ای که در تعیین خطوط جاده برای فرایند رانندگی خودکار خودروها در کنار برخط بودن روش‌های پیشنهادی مطرح می‌شود، دقت بالایی است که باید روش‌های ارائه‌شده آن را تضمین کنند. یکی از جدیدترین رویکردهایی که امروزه برای تعیین خطوط ارائه شده‌اند، استفاده از الگوریتم‌های یادگیری برای آموزش سیستم‌های موجود در جهت تعیین یک روند تماماً خودکار برای شناسایی خطوط است. روش‌های خودکار مبتنی بر یادگیری که اغلب برای شناسایی خطوط جاده استفاده شده‌اند، اغلب از نوع نظارت‌شده هستند. این روش‌ها نیازمند یک مجموعه داده برچسب‌گذاری می‌باشند که خود هم باعث محدودیت در روند توسعه سیستم شناسایی خطوط می‌شود و نکته مهم‌تر اینکه ایجاد یک مجموعه داده استاندارد دارای برچسب، بسیار زمان‌بر است.

در این رساله برای تشخیص خطوط جاده ما اقدام به ارائه دو روش پیشنهادی، یکی مبتنی بر رویکردهای کلاسیک موجود در پردازش تصویر، که استفاده از الگوریتم‌های پایه‌ای است و دیگری یک روش خودکار و مبتنی بر یادگیری نیمه‌نظارتی با استفاده از شبکه‌های عصبی عمیق جهت استخراج خطوط جاده نموده‌ایم. بدین منظور چهار مجموعه داده متفاوت برای فرایند شناسایی خطوط در دو روش پیشنهادی به‌کارگیری شده است. نتایج حاصل از روش پیشنهادی اول که به صورت رویکرد کلاسیکی الگوریتم‌های پایه‌ای پردازش تصویر انجام شد، از لحاظ دقت بصری در وضعیت مناسبی برقرار بود. این روش دارای دقت بصری ۹۶ درصد است. در روش انجام شده دوم که با استفاده از رویکرد مبتنی بر یادگیری عمیق انجام گرفت، با استفاده از یک شبکه عمیق مولد رقابتی نیمه‌نظارتی اقدام به شناسایی خطوط نمودیم. مطابق با آنچه مورد انتظار ما بود، دقت حاصل از به‌کارگیری روش پیشنهادی نسبت به روش پیشنهادی اول، از لحاظ معیارهای ارزیابی کمی، بهبود نسبی

یافت. از طرفی روش پیشنهادی اول نسبت به شرایط محیطی متفاوت، نیاز به تنظیمات برخی از پارامترهای اصلی نظیر، تنظیم اولیه حد آستانه روشنایی، تنظیم پارامترهای مربوطه در لبه‌یابی کنی، تنظیم پارامترهای مربوط به الگوریتم تبدیل هاف، تنظیم پارامترهای مربوط به رفع نویز و غیره در الگوریتم اجرایی دارد. بنابراین فاقد توانایی خودکار در شناسایی خطوط جاده است. در ادامه به بیان برخی دیگر از معایب و مزایای روش‌های پیشنهادی و همچنین روش تشخیص مانع خواهیم پرداخت.

۶-۲-۲ مزایای روش پیشنهادی اول

۶-۲-۱ روش اجرایی بلادرنگ

روش‌های مبتنی بر رویکردهای کلاسیکی پردازش تصویر با توجه به پیچیدگی محاسباتی کمی که دارند، برای به‌کارگیری در سیستم‌هایی که سرعت عمل‌کرد اجرایی در آن‌ها از اهمیت بالایی برخوردار می‌باشند، بسیار مناسب هستند. چرا که در این الگوریتم‌ها، فرایند یادگیری که چالش اصلی در مورد پیچیدگی محاسباتی دارد، مطرح نیست. گرچه با ظهور پردازنده‌های قدرتمند، سرعت الگوریتم‌های مبتنی بر یادگیری نیز می‌توانند به سمت الگوریتم‌های بلادرنگ رود. در روش پیشنهادی اول، مهم‌ترین ویژگی موجود در الگوریتم بلادرنگ بودن آن هست که به‌عنوان هدف اصلی به آن دست یافته‌شده است.

۶-۲-۲ پویایی انتخاب ناحیه مورد علاقه

در این روش با استفاده از یک ماسک‌گذاری پویا، اقدام به مشخص نمودن ناحیه مورد علاقه با شناسایی و حذف نقاط محوشوندگی نمودیم. بدین ترتیب پیچیدگی محاسباتی سیستم نیز کاهش پیدا نموده است.

۶-۲-۳ کارایی بالا در انواع شرایط محیطی با توجه به دو پایگاه داده به‌کار گرفته شده

روش پیشنهادی اول با توجه به نتایج بررسی شده که برای دو پایگاه داده‌ها IROADS و Aly's Method انجام شد، در شرایط محیطی متفاوت دارای عمل‌کرد بسیار مناسبی بود. پایگاه داده‌ها IROADS دارای چهار شرایط آب و هوایی متفاوت بود که دقت عمل‌کردی برای هر چهار حالت از لحاظ معیار مورد بررسی بسیار قابل قبول

بود و میانگین دقت ۹۶/۷۸ درصد حاصل شد. پایگاه داده‌ها Aly's Method هم دارای تصاویر مربوط به جاده-های درون شهری با شرایط محیطی مختلف بود که دقت روش پیشنهادی ۸۹/۳۹ درصد به‌دست آمد.

۳-۶ معایب روش پیشنهادی اول

۱-۳-۶ غیر خودکار

روش‌هایی که مبتنی بر الگوریتم‌های پایه‌ای پردازش تصویر هستند، معمولاً با تغییرات در پایگاه دادگان مورد نظر، نیاز به تنظیمات مجدد برای اجرا بر روی آنان دارند. لذا برای به‌دست آوردن یک جامعیت کامل در گرفتن جواب، نیازمند به‌کارگیری مجموعه دادگان بسیار زیاد، برای رسیدن به یک حالت پایدار و قابل اطمینان در تمام حالات و شرایط محیطی ممکن است.

۲-۳-۶ عدم وجود معیار کمی مطمئن

همان‌طور که در روش پیشنهادی اول مشاهده کردیم، معیاری که برای ارزیابی و سنجش عملکردی روش به‌کار گرفتیم، به‌صورت کمی در داخل خود برنامه قابل محاسبه نبود. لذا باید به‌صورت بصری و بررسی تک‌تک نمونه‌های آزمون و ارزیابی بصری، اقدام به ارزیابی غیرپارامتریک روش مورد نظر داشته باشیم.

۴-۶ مزایای روش پیشنهادی دوم

۱-۴-۶ عدم نیاز به دادگان کاملاً برچسب‌گذاری شده

روش پیشنهادی دوم برای تشخیص خطوط جاده که مبتنی بر یادگیری عمیق بود، مشکل نیاز به برچسب-گذاری

تمام دادگان را نداشت. بدین معنی که برای فرایند آموزش، نیازمند محدودیت اینکه همه دادگان لزوماً باید دارای برچسب باشند نبود که چنین مسئله‌ای خود باعث صرفه‌جویی در زمان است.

۶-۴-۲ قابلیت بالا در توسعه

یکی از نتایجی که مزیت اول برای روش مبتنی بر یادگیری عمیق و نیمه‌نظارتی در پی دارد، قابلیت توسعه روش برای فرایند شناسایی خطوط در انواع دادگان است. بدین منظور با داشتن تعداد کمی از دادگان برچسب-گذاری شده از چندین نوع پایگاه داده‌ها و در کنار آن‌ها استفاده از دادگانی که صرفاً در آنان ماهیت وجود جاده در تصویر غالب باشد، (منظور این است که صرفاً بدانیم تصویری از جاده داریم) میتوانیم با طراحی یک معماری مناسب سیستمی جامع با عملکرد مناسب برای انواع مجموعه دادگان مختلف و حتی توسعه سیستم‌های تجاری داشته باشیم.

۶-۴-۳ روش خودکار

روش پیشنهادی دوم با توجه به ماهیت یادگیری که در ساختار الگوریتمی خود دارد، باعث می‌شود که روش با آموزش صحیح، قابلیت خودکار بودن برای انواع شرایط محیطی را داشته باشد. بدین معنی که ساختار کلی روش در نهایت منتج به مدلی می‌شود که تمام نمونه‌های موجود در خود را دارد و برای هدف مورد نظر پاسخگو است.

۶-۴-۴ پارامتریک و قابل ارزیابی

با توجه به اینکه در روش پیشنهادی مبتنی بر یادگیری عمیق برای ارزیابی عملکرد از معیارهای IOU و Accuracy استفاده کردیم، لذا نتایج حاصل با چنین معیارهایی دارای قابلیت اطمینان بالاتری نسبت به روش‌های غیرپارامتریک هستند.

۶-۵ معایب روش پیشنهادی دوم

روش‌های مبتنی بر یادگیری عمیق، باوجود دقت بالایی که برای کاربر فراهم می‌کنند، همچنان از لحاظ نیاز به استفاده از دادگان با حجم بالا ضعف می‌برند. از طرفی برای روش انجام‌شده در این رساله نیز چنین مشکلی باوجود اینکه برای دادگان برچسب‌گذاری شده تا حدی برطرف شده بود، اما باز هم در کل پا برجا است .

اختلاف پارامتریک معیار اصلی شناسایی و استخراج در فرایندهای قطعه‌بندی، از جمله شناسایی خطوط جاده‌ای، معمولاً IOU است. اختلاف این پارامتر برای روش‌های کاملاً نظارت‌شده، نسبت به روش اجرایی ما باوجود کاهش بسیار چشمگیر پیچیدگی محاسباتی روش، در حد بالایی است. گاهی در برخی سیستم‌هایی که حساسیت در دقت روش نسبت به سرعت اجرای آن بیشتر باشد، چنین روشی ممکن است از لحاظ عمل‌کردی با مشکل در کارایی روبرو باشد.

۶-۶ مزایای روش تشخیص مانع

۶-۶-۱ سادگی، عدم نیاز به پردازش زیاد

روش شناسایی که با استفاده از ترکیب الگوریتم‌های شناسایی ویژگی HOG و همچنین طبقه‌بندی SVM انجام شد، نیاز به محاسبات زیادی برای یادگیری مدل تشخیص نوع مانع ندارد. در واقع یادگیری این روش از لحاظ پیچیدگی محاسباتی بسیار پایین است. از طرفی نیز این روش از لحاظ اجرایی بسیار ساده است.

۶-۷ معایب روش تشخیص مانع

۶-۷-۱ نیاز به تنظیمات دستی بر روی تصاویر مختلف جاده

اصلی‌ترین مشکلی که این روش دارد، عدم کارایی خودکار برای ساختارهای متنوع تصویر در شناسایی خودرو است. بدین معنی که اگر ساختار کلی یک مجموعه همچون ساختار محیطی جاده داده تغییر کند، نیاز به تعیین دستی برخی پارامترها، همچون محدوده‌های شناسایی خودرو در تصویر و همچنین حد آستانه روشنایی محیط دارد. برای خودکارسازی می‌توان از الگوریتم‌های یادگیری عمیق استفاده نمود. هرچند، این مسئله خود باعث افزایش پیچیدگی محاسباتی روش می‌شود.

۶-۸ پیشنهاداتی برای بهبود در آینده

با توجه به آنچه که به‌عنوان معایب فعالیت انجام‌شده مطرح گردید، می‌توان موارد زیر را به‌عنوان فعالیت‌هایی جهت بهبود در آینده نام برد.

۱. تغییر ساختار شبکه مولد رقابتی نیمه نظارتی را به طوری که از بخش مولد آن برای تولید داده استفاده شود و با دادگان تولید شده، به بخش نظارت نشده سیستم کمک کرد، در این صورت می توان با کمی تغییر در ساختار کلی این شبکه از لحاظ مفهومی، کاهش پیچیدگی محاسباتی را برای شبکه محقق نمود.
۲. بهبود دقت پارامتریک رویکرد نیمه نظارتی انجام شده، خود می تواند به عنوان یک چالش مهم در مطالعات بعدی مطرح باشد.
۳. روش پیشنهادی دوم که در این رساله انجام گردید، دارای یک قابلیت بسیار بالا با توجه به ماهیت ساختاری شبکه های مولد رقابتی نیمه نظارتی در استفاده از انواع پایگاه داده های متفاوت از یک جنس است. مقصود این است که اگر بتوان از دادگان با ماهیت مثلا جاده از منابع مختلف استفاده نمود، می توان یک سیستم توسعه یافته بسیار قدرتمند ایجاد نمود. البته لازمه دیگر همانطور که بیان گردید، افزایش دقت عمل کردی روش های نیمه نظارتی است.
۴. برای روش تشخیص موانع، طبق آنچه در معایب روش بیان شد، می توان با استفاده از الگوریتم های مبتنی بر یادگیری عمیق، فرایند تشخیص موانع را برای محیط های مختلف به صورت خودکار انجام داد.

- [1] N. Garnett, et al., "3D-LaneNet: End-to-end 3D multiple lane detection," in Proceedings of the IEEE International Conference on Computer Vision, pp. 2921-2930, 2019.
- [2] P. Cudrano, et al., "Advances in centerline estimation for autonomous lateral control," in IEEE Intelligent Vehicles Symposium (IV), pp. 1415-1422, 2020.
- [3] B. Huval, et al., "An empirical evaluation of deep learning on highway driving," arXiv: 1504.01716, 2015.
- [4] S. Lee, et al., "VPGNet: Vanishing point guided network for lane and road marking detection and recognition," in Proceedings of the IEEE International Conference on Computer Vision, pp. 1947-1955, 2017.
- [5] S. Habib and M. Hannan, "Lane Departure Detection and Transmission using Hough Transform Method," *Przegląd Elektrotechniczny*, pp. 141-146, 2013.
- [6] M. Bertozzi and A. Broggi, "GOLD: A Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection," *IEEE Transaction on Image Processing*, vol. 7, no. 1, pp. 62-81, 1998.
- [7] M. Bertozzi and A. Broggi, "Vision-Based Vehicle Guidance," *Computer Vision*, vol. 30, pp. 49-55, 1997.
- [8] X. X. Wei and L. Meng, "A Method to Implementation of Lane Detection Under Android System Based on OpenCV," *Intelligent Technologies and Engineering Systems*, Springer, pp. 115-121, 2013.
- [9] W. Wang and X. Huang, "An FPGA Co-Processor for Adaptive Lane Departure Warning System," *IEEE International Symposium on Circuits and Systems (ISCAS2013)*, pp. 1380-1383, 2013.
- [10] B. Benligiray, C. Topal and C. Akinlar, "Video-Based Lane Detection Using a Fast Vanishing Point Estimation Method," *IEEE International Symposium on Multimedia*, pp. 348-351, 2012.
- [۱۱] ز. رحیمی افضل، "طراحی یک سیستم هشدار انحراف از جاده و پیاده سازی آن بر روی تبلت با سیستم عامل اندروید،" *دانشگاه صنعتی اصفهان*، ۱۳۹۳.
- [۱۲] م. کاظمی، "تشخیص خطوط جاده بر اساس تحلیل تصاویر صحنه های رانندگی،" *دانشگاه صنعتی نوشیروانی بابل*، ۱۳۹۳.
- [13] D. Abdullah and M. M. Al-hafidh, "Developing Parallel Application on Multi-core Mobile Phone," (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, vol. 4, no. 11, pp. 89-93, 2013.
- [14] J. H. Yoo, D. H. Kim and S.-K. Park, "A New Lane Detection Method Based on Vanishing Point Estimation with Probabilistic Voting," *IEEE International Conference on Consumer Electronics(ICCE)*, pp. 204-205, 2015.
- [15] R. Grompone von Gioi, J. Jakubowicz, J.-M. Morel and G. Randall, "LSD: A fast line segment detector with a false detection control," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 4, pp. 722-732, 2010.

- [16] J.-H. Cho, E. Tsogtbaatar, S.-H. Kim, Y.-M. N. P.-M.-L. Jang and S.-B. Cho, "Improved lane detection system using Hough transform with super-resolution reconstruction algorithm and multi-ROI," IEEE International Conference on Electronics, Information and Communications (ICEIC) , 2014.
- [17] T. Duong, C. Pham, T. H.-P. Tran, T. T. Duong, C. C. Pham and T. H.-P. Tran, "Near Real-Time Ego-Lane Detection in Highway and Urban Streets," IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia), 2016.
- [18] R. Jiang, M. Terauchi, R. Klette, S. Wang and T. Vaudrey, "Low Level Image Processing for Lane Detection and Tracking," Arts and Technology, vol. 30, pp. 190-197, 2010.
- [19] H. Changzheng, H. Jin and Y. Chaochao, "An Efficient Lane Markings Detection and Tracking Method based on Vanishing Point Constraints," in Proceedings of the 35th Chinese Control Conference, Chengdu, 2016.
- [20] C. Y. Low, Z. H and S. A. Mazlan, "Simple Robust Road Lane Detection Algorithm," in 5th International Conference on Intelligent and Advanced Systems (ICIAS), Kuala Lumpur, 2014.
- [21] J. Son, H. Yoo, S. Kim and K. Sohn, "Real-time Illumination Invariant Lane Detection for Lane Departure Warning System," Elsevier Expert Systems with Applications, pp. 1816-1824, 2015.
- [22] J. H. Yoo, S.-W. Lee, S.-K. Park and D. H. Kim, "A Robust Lane Detection Method Based on Vanishing Point Estimation Using the Relevance of Line Segments," IEEE Transactions on Intelligent Transportation Systems, pp. 3254-3266, 2017.
- [23] P. N. Bhujbal and S. P. Narote, "Lane Departure Warning System Based on Hough Transform and Euclidean Distance," 2015 Third International Conference on Image Information Processing (ICIIP), pp. 370-373, 2015.
- [24] S. C. Yi, Y. C. Chen and C. H. Chang, "A lane detection approach based on intelligent vision," Computers and Electrical Engineering 42, pp. 23-29, 2015.
- [25] O. G. Lotfy, A. A. Kassem, E. M. Nassief, H. A. Ali, M. R. Ayoub, M. A. El-Moursy and M. M. Farag, "Lane departure warning tracking system based on score mechanism," IEEE 59th International Midwest Symposium on Circuits and Systems (MWSCAS), pp. 1-4, 2016.
- [26] Y. Lee and H. Kim, "Real-time lane detection and departure warning system on embedded platform," IEEE 6th International Conference on Consumer Electronics-Berlin (ICCE-Berlin), pp. 1-4, 2016.
- [27] J. Niu, J. Lu, M. Xu, P. Lv and X. Zhao, "Robust Lane Detection using Two-stage Feature Extraction with Curve Fitting," Pattern Recognition 59, pp. 225-233, 2016.
- [28] H. Guan, W. Xingang, W. Wenqi, Z. Han and W. Yuanyuan, "Real-Time Lane-Vehicle Detection and Tracking System," Chinese Control and Decision Conference (CCDC), pp. 4438-4443, 2016.
- [29] Liang D, Guo YC, Zhang SK, Mu TJ, Huang X. Lane detection: a survey with new results. Journal of Computer Science and Technology. 35:493-505, 2020.

- [30] Oliveira G L, Burgard W, Brox T. Efficient deep models for monocular road segmentation. In Proc. the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 2016.
- [31] Liu X, Deng Z, Yang G. Drivable road detection based on dilated FPN with feature aggregation. In Proc. the 29th IEEE International Conference on Tools with Artificial Intelligence, November 2017.
- [32] Mamidala R S, Uthkota U, Shankar M B, Antony A J, Narasimhadhan A V. Dynamic approach for lane detection using Google street view and CNN. In Proc. the 2019 IEEE Region 10 Conference, October 2019.
- [33] Lee S, Kim J, Yoon J S, Shin S, Bailo O, Kim N, Lee T, Hong H S, Han S, Kweon I S. VPGNet: Vanishing point guided network for lane and road marking detection and recognition. In Proc. the 2017 IEEE International Conference on Computer Vision, October 2017.
- [34] Li J, Mei X, Prokhorov D V, Tao D. Deep neural network for structural prediction and lane detection in traffic scene. IEEE Transactions on Neural Networks and Learning Systems, 2017.
- [35] Hou Y, Ma Z, Liu C, Loy C C. Learning lightweight lane detection CNNs by self attention distillation. In Proc. The 2019 IEEE/CVF International Conference on Computer Vision, October 2019.
- [36] He B, Ai R, Yan Y, Lang X. Accurate and robust lane detection based on dual-view convolutional neural network. In Proc. the 2016 IEEE Intelligent Vehicles Symposium, June 2016.
- [37] Bai M, M'attys G, Homayounfar N, Wang S, Lakshmikanth S K, Urtasun R. Deep multi-sensor lane detection. In Proc. the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 2018.
- [38] Satti, Satish Kumar, K. Suganya Devi, Prasenjit Dhar, and P. Srinivasan. "A machine learning approach for detecting and tracking road boundary lanes." ICT Express 7, no. 1, 2021.
- [39] Lee DH, Liu JL. End-to-End Deep Learning of Lane Detection and Path Prediction for Real-Time Autonomous Driving. arXiv preprint arXiv: 2102.04738, 2021.
- [40] Wang W, Lin H, Wang J. CNN based lane detection with instance segmentation in edge-cloud computing. Journal of Cloud Computing. 9:1-0, 2020.
- [41] Zou Q, Jiang H, Dai Q, Yue Y, Chen L, Wang Q. Robust lane detection from continuous driving scenes using deep neural networks. IEEE transactions on vehicular technology. 69(1):41-54, 2019.
- [42] EKŞİ, Osman Tahir, and Gökhan GÖKMEN. "Comparing of Some Convolutional Neural Network (CNN) Architectures for Lane Detection." Balkan Journal of Electrical and Computer Engineering 8, no. 4: 314-319
- [43] Shirke, Suvarna, and R. Udayakumar. "Lane datasets for lane detection." In 2019 International Conference on Communication and Signal Processing (ICCSP), pp. 0792-0796. IEEE, 2019.
- [44] Song, Shuang, Wei Chen, Qianjie Liu, Huosheng Hu, Tengchao Huang, and Qingyuan Zhu. "A novel deep learning network for accurate lane detection in low-light environments." Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering 236, no. 2-3 (2022): 424-438.

- [45] Cordts, Marius, Mohamed Omran, Sebastian Ramos, Timo Scharwächter, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. "The cityscapes dataset." In CVPR Workshop on the Future of Datasets in Vision, vol. 2. 2015.
- [46] Yousri, Retaj, Mustafa A. Elattar, and M. Saeed Darweesh. "A deep learning-based benchmarking framework for lane segmentation in the complex and dynamic road scenes." *IEEE Access* 9 (2021): 117565-117580.
- [47] Chao, Fan, Song Yu-Pei, and Jiao Ya-Jie. "Multi-lane detection based on deep convolutional neural network." *IEEE Access* 7 (2019): 150833-150841.
- [48] Neven, Davy, Bert De Brabandere, Stamatios Georgoulis, Marc Proesmans, and Luc Van Gool. "Towards end-to-end lane detection: an instance segmentation approach." In 2018 IEEE intelligent vehicles symposium (IV), pp. 286-291. IEEE, 2018.
- [49] Wang, Ze, Weiqiang Ren, and Qiang Qiu. "Lanenet: Real-time lane detection networks for autonomous driving." *arXiv preprint arXiv:1807.01726* (2018).
- [50] Zhang, Xinyu, He Huang, Weiming Meng, and Dean Luo. "Improved lane detection method based on convolutional neural network using self-attention distillation." *Sens. Mater* 32 (2020): 4505-4516.
- [51] Zhao, Zhiyuan, Qi Wang, and Xuelong Li. "Deep reinforcement learning based lane detection and localization." *Neurocomputing* 413 (2020): 328-338.
- [52] Al Mamun, A., Poh Ping Em, and J. Hossen. "Lane marking detection using simple encode decode deep learning technique: SegNet." *Int. J. Electr. Comput. Eng* 11, no. 4 (2021): 3032-3039.
- [53] Kanagaraj, Nitin, David Hicks, Ayush Goyal, Sanju Tiwari, and Ghanapriya Singh. "Deep learning using computer vision in self driving cars for lane and traffic sign detection." *International Journal of System Assurance Engineering and Management* 12, no. 6 (2021): 1011-1025.
- [54] Zhang, Jiyong, Tao Deng, Fei Yan, and Wenbo Liu. "Lane detection model based on spatio-temporal network with double convolutional gated recurrent units." *IEEE Transactions on Intelligent Transportation Systems* (2021)..
- [55] Lee, Minhyeok, Junhyeop Lee, Dogyoon Lee, Woojin Kim, Sangwon Hwang, and Sangyoun Lee. "Robust lane detection via expanded self attention." In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 533-542. 2022.
- [56] I. Ulrich and I. Nourbakhsh, "Appearance Based Obstacle Detection with Monocular Color Vision," *AAAI/IAAI*, pp. 866-871, 2000.
- [57] K. Yamaguchi, T. Kato and Y. Ninomiya, "Moving Obstacle Detection using Monocular Vision," *IEEE Intelligent Vehicles Symposium*, p. 288–293, 2006.
- [58] C. Harris and M. Stephens, "A Combined Corner and Edge Detector," *Proc. Alvey Vision Conf*, pp. 147-151, 1988.
- [59] B. D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," *Int. Joint Conference on Artificial Intelligence*, pp. 674-679, 1981.
- [60] G. Ma, D. Muller, S. B. Park, S. M. Schneiders and A. Kummert, "Pedestrian Detection Using a Single Monochrome Camera," *IET Intelligent Transport Systems*, vol. 3, pp. 42-56, 2009.

- [61] C. Viet and I. Marshall, "An Efficient Obstacle Detection Algorithm Using Colour and Texture," World Academy of Science, Engineering and Technology, pp. 132-137, 2009.
- [62] Y. Lin, C. Lin, W. Liu and L. Chen, "A Vision-Based Obstacle Detection System for Parking Assistance," 8th IEEE Conference on Industrial Electronics and Applications (ICIEA), pp. 1627-1630, 2013.
- [63] E. Rosten, R. Porter and T. Drumond, "Fast and better: a machine learning approach to corner detection," IEEE Transaction on Pattern Analysis and Machine Intelligence, pp. 1-15, 2010.
- [64] B. Jia, R. Liu and M. Zhu, "Real-Time Obstacle Detection with Motion Features Using Monocular Vision," Springer-Verlag Berlin Heidelberg, 2014.
- [65] J. Bouguet, "Pyramidal Implementation of the Affine Lucas Kanade Feature Tracker Description of the Algorithm," Intel Corporation, 2001.
- [66] I. Nadav and E. Katz, "Off-road Path and Obstacle Detection using Monocular Camera," International Conference on the Science of Electrical Engineering (ISCEE), 2016.
- [67] N. Deepika and V. V. Sajith Variyar, "Obstacle Classification and Detection for Vision Based Navigation for Autonomous Driving," Advances in Computing, Communications and Informatics (ICACCI), 2017 International Conference on. IEEE, pp. 2092-2097, 2017.
- [68] V. Badrinarayanan, A. Kendall and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Robust Semantic Pixel-Wise Labelling," arXiv preprint arXiv:1511.00561, 2015.
- [69] Ohgushi, Toshiaki, Kenji Horiguchi, and Masao Yamanaka. "Road Obstacle Detection Method Based on an Autoencoder with Semantic Segmentation." In Proceedings of the Asian Conference on Computer Vision. 2020.
- [70] Lis, Krzysztof, Sina Honari, Pascal Fua, and Mathieu Salzmann. "Detecting Road Obstacles by Erasing Them." arXiv preprint arXiv:2012.13633 (2020).
- [71] Demonceaux, Cédric, Alex Potelle, and Djemâa Kachi-Akkouche. "Obstacle detection in road scene based on road detection and motion analysis." IFAC Proceedings Volumes 37, no. 8 (2004): 442-447.
- [72] Lv, Jing, Yuanyuan Shang, and Hui Ding. "Road Obstacle Detection Based on Randomized Hough Transform." In 2nd International Conference on Teaching and Computational Science (ICTCS 2014). 2014.
- [73] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., "Generative adversarial nets", Advances in neural information processing systems, (2014).
- [74] Odena, Augustus. "Semi-supervised learning with generative adversarial networks." arXiv preprint arXiv:1606.01583, 2016.
- [75] Souly, Nasim, Concetto Spampinato, and Mubarak Shah. "Semi supervised semantic segmentation using generative adversarial network." In Proceedings of the IEEE international conference on computer vision, pp. 5688-5696. 2017.
- [76] Chi Qin, L. A. I., and Soo Siang Teoh. "An efficient method of HOG feature extraction using selective histogram bin and PCA feature reduction." Advances in Electrical and Computer Engineering 16, no. 4 (2016): 101-108.

- [77] Jinxia, Liu, and Qiu Yuehong. "Application of SIFT feature extraction algorithm on the image registration." In IEEE 2011 10th International Conference on Electronic Measurement & Instruments, vol. 3, pp. 177-180. IEEE, 2011.
- [78] Juan, Su, Xu Qingsong, and Zhu Jinghua. "A scene matching algorithm based on surf feature." In 2010 International Conference on Image Analysis and Signal Processing, pp. 434-437. IEEE, 2010.
- [79] Yang, Jun, Yu-Gang Jiang, Alexander G. Hauptmann, and Chong-Wah Ngo. "Evaluating bag-of-visual-words representations in scene classification." In Proceedings of the international workshop on Workshop on multimedia information retrieval, pp. 197-206. 2007.
- [80] Cortes, Corinna, and Vladimir Vapnik. "Support-vector networks." Machine learning 20, no. 3 (1995): 273-297.
- [81] Rezaei, Mahdi, and Mutsuhiro Terauchi. "iROADS Dataset (Intercity Roads and Adverse Driving Scenarios)." Enpeda Image Sequence Analysis Test Site-EISATS (2014).
- [82] M. Aly, "Real Time Detection of Lane Markings in Urban Streets," IEEE Symposium on Intelligent Vehicles, pp. 7-12, 2008.
- [83] J. Son, H. Yoo, S. Kim, and K. Sohn, "Real-time illumination invariant lane detection for lane departure warning system," Expert Systems with Applications, vol. 42, pp. 1816-1824, 2015.
- [84] J. Cao, C. Song, S. Song, F. Xiao, and S. Peng, "Lane detection algorithm for intelligent vehicles in complex road conditions and dynamic environments," Sensors, vol. 19, 2019.
- [85] J. Hu, S. Xiong, Y. Sun, J. Zha, and C. Fu, "Research on Lane Detection Based on Global Search of Dynamic Region of Interest (DROI)," Applied Sciences, vol. 10, 2020.
- [86] Kim, JongBae. "Vehicle Detection Using Deep Learning Technique in Tunnel Road Environments." Symmetry 12, no. 12 (2020): 2012.
- [87] El Jaafari, Ilyas, Mohamed El Ansari, Lahcen Koutti, Ayoub Ellahyani, and Said Charfi. "A novel approach for on-road vehicle detection and tracking." Int. J. Adv. Comput. Sci. Appl. 7, no. 1 (2016): 594-601.
- [88] Hasibuan, Nisma Novita, Muhammad Zarlis, and Syahril Efendi. "Detection and Tracking Different Type of Cars With YOLO model combination and deep sort algorithm based on computer vision of traffic controlling." Sinkron: jurnal dan penelitian teknik informatika 6, no. 1 (2021): 210-221.

Abstract

The transportation industry is one of the most active industries in the world and road safety is one of the most important areas of research and development in this industry. Lane detection system and obstacle detection system are among the safety equipment used in smart vehicles to navigate the vehicle between the road lines and prevent vehicle deviation and accidents in dealing with obstacles. Video images contain very valuable information about the environment around the car and play a key role in detecting road lines and obstacles. Due to the fact that we are faced with a natural environment in recognizing road lines and obstacles in video images, there are several factors such as different weather conditions, the shadow of buildings and trees along the road, the shadow of cars, the state of sunlight at different hours of the day and the light of headlights of oncoming cars at night, make it difficult to detect road lines and obstacles. On the other hand, due to the large volume of computations for processing image information and also the need for real-time detection of road lines and obstacles, reducing computational complexity and increasing computing power should be considered for the proposed system.

In this thesis, the aim is to identify road lines and obstacles in front of the vehicle using video images taken from the road, and to determine the driving conditions on the road. For this purpose, we have used two different proposed methods in this study. First, we used a classical approach based on basic image processing algorithms, and using edge detection and Hough transform algorithm, to identify road lines with an unsupervised approach. One of the innovations in this method, in addition to being real-time, which has been done as the main goal in the executive method, is to perform a dynamic masking to identify the area of interest in the image. In the second proposed method, which is based on a deep learning approach, we have identified a line using a Semi-Supervised Generative Adversarial Network, and then using a Histogram Oriented Gradient descriptor, we identify the obstacles in front of the car. Among the important points in the second method is the use of a deep Semi-Supervised Generative Adversarial Network in the segmentation process, as well as its lower computational complexity compared to known deep network architectures such as ResNet50, AlexNet and VGG19. Also, in the obstacle detection algorithm, by applying the dynamic masking policy, we have achieved a faster execution speed in detecting oncoming vehicles. In fact, the masking used is a very simple idea to increase the processing speed of the algorithm.

The results of applying the first proposed method indicate the accuracy of 96.78% lane detection in real time on the IROADS database used in different environmental conditions. Also, the results of using the second proposed method for detecting road lines have a pixel accuracy of 90.78%, which indicates that in addition to the relatively good percentage that the implementation method has over the supervised approaches, in terms of computational complexity in the number of training parameters used, with 21 million training parameters, there has been a significant improvement over the aforementioned deep networks. Finally, it should be noted that the output of the obstacle detection algorithm in front of the car, according to the quality criterion defined for its operational accuracy, has an accuracy of 92.3% in obstacle detection.

Keywords: Lane Detection, Obstacle Detection, Hough Transform, Deep Learning, Generative Adversarial Networks, Semi-Supervised Learning.



Shahrood University of Technology
Faculty of Electrical and Robotic Engineering

Real-time Lane Detection and Recognition of Obstacles in Front of the Vehicle

By:

Ali Asghar Fallah joosheqani

Supervisor:

Dr. Ali Soleimani

Advisor:

Dr. Hossein Khosravi

February 2022