

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی برق و رباتیک

رساله دکتری مهندسی برق - الکترونیک

تخمین ابعاد و سرعت خودروها با استفاده از ویدئوی دریافتی از دوربین کنار جاده‌ای

نگارنده: رسول عسگریان دهکردی

استاد راهنما

دکتر حسین خسروی

استاد مشاور

دکتر علیرضا احمدی فرد

آذر ۱۳۹۸

تخت‌ای ناپسیر، پیشکش به ساحت حضرت ولی عصر (عج)

و همچنین تقدیم به همه سربازان بی ادعای مافع حرم

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

از آن جا که جایگاه پدر و مادر خداوند محفوظ است، امیدوارم پروردگار مرا توفیق دهد تا فرزندی صالح برای پدر و مادرم باشم.

لازم است از برادر و خواهر عزیز، همچنین سایر اعضای خانواده ام به خاطر کمک ایشان تشکر کنم.

از سال ۱۳۹۴ که افتخار آشنایی با آقای دکتر حسین خسروی را دارم، ایشان برای من هم استاد درس هستند و هم استاد اخلاق. این

بزرگوار در هر موضوعی، دسی و غیر دسی، اینجانب را با مانت را بهمانی کردند و در هر شرایطی دست مرا گرفتند، امیدوارم خداوند در

هر شرایطی دستشان را بگیرد.

از آقای دکتر علیرضا احمدی فرد که استاد مشاور اینجانب بودند نیز تشکر می کنم، مصاحبت با ایشان همواره برای بنده شیرین و درس آموز است.

از کلیه بزرگوارانی که از دوران کودکی تا کنون به من چیزی آموخته اند، پاسکزارم.

تعمدنامه

اینجانب رسول عسگریان دهکردی دانشجوی دوره دکتری رشته مهندسی برق دانشکده مهندسی برق و رباتیک دانشگاه صنعتی شاهرود، نویسنده رساله تخمین ابعاد و سرعت خودروها با استفاده از ویدئوی دریافتی از دوربین کنار جاده‌ای تحت راهنمایی دکتر حسین خسروی 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.

در سال‌های اخیر در بسیاری از جاده‌ها، دوربین‌های جاده‌ای برای نظارت بر خودروها استفاده می‌شوند. با توجه به افزایش روزافزون این دوربین‌ها و همچنین تعداد بالای خودروهای در حال تردد، استفاده از انسان به‌عنوان کاربر مستقیم این دوربین‌ها در آینده با مشکلاتی همراه است. بر همین اساس تمایل بر این است که با استفاده از روش‌های بینایی ماشین و پردازش تصویر پارامترهایی مانند سرعت خودرو، کلاس خودرو و تعداد خودروهای عبوری به‌دست آید. روش‌هایی که تاکنون در این زمینه ارائه شده، مشکلات عمده‌ای از قبیل غیرخودکار بودن، عدم مقاومت نسبت به سایه، تک کاربرده بودن و محدودیت پایگاه داده‌های واقعی دارند.

هدف این رساله ارائه روشی برای تخمین سرعت و ابعاد خودروها است. این روش باید به‌گونه‌ای طراحی شود که اهداف کالیبراسیون تمام‌خودکار، مقاوم بودن نسبت به سایه، چند کاربرده بودن و ایجاد یک پایگاه داده جامع، برآورده شود. روش پیشنهادی برای تخمین سرعت و ابعاد به این‌صورت است که در قاب‌های ابتدایی با توجه به راستای حرکت خودروها، نقاط محوشدگی و صفحه فرضی جاده به‌دست می‌آید. سپس با شناسایی پیش‌زمینه و حذف سایه، محدوده هر خودرو تعیین و جعبه سه‌بعدی آن تشکیل می‌گردد. برای تعیین ضرایب متریک، چند خودرو از کلاس‌های رایج در چند قاب اول شناسایی شده و با توجه به ابعاد واقعی این خودروها برحسب متر و ابعاد معادل‌شان روی صفحه جاده برحسب پیکسل، ضرایب متریک محاسبه می‌شود. با این کار یکی از اهداف ما که کالیبراسیون تمام‌خودکار است، تکمیل می‌شود. در نهایت خودروهای عبوری بر صفحه فرضی تصویر شده و با ردیابی، سرعت و ابعاد دقیق آن‌ها محاسبه می‌گردد. ایجاد مقاومت در برابر سایه به‌عنوان یکی از اهداف رساله، منجر به تشخیص محدوده دقیق هر خودرو می‌شود و خطای کالیبراسیون و تخمین ابعاد و سرعت را کاهش می‌دهد. در طراحی الگوریتم حذف سایه علاوه بر دقت، سرعت پردازش نیز مهم است و سعی شده تا حد امکان این مرحله تاثیر منفی بر سرعت پردازش کلی نداشته باشد.

در حال حاضر مجموعه‌ای کامل شامل ویدئوهای دارای برچسب سرعت و تصاویر خودروها، خصوصاً در داخل ایران وجود ندارد. هدف دیگر رساله، ایجاد چنین پایگاهی است. لذا مجموعه‌ای با فیلم‌برداری هم‌زمان توسط دوربین عادی و لیزر از جاده‌ها تهیه گردید.

فرایند کالیبراسیون نیازمند شناسایی تعدادی خودروی رایج در ابتدای کار است، به کمک شبکه‌های عمیق و تثبیت تصویر، الگوریتمی قوی برای این منظور ارائه کرده‌ایم. با توجه به نصب دوربین‌های جاده در ارتفاع بالا و کوچک بودن

تصاویر خودرو، در الگوریتم پیشنهادی، از تکنیک‌های فراتفکیک‌پذیری برای افزایش وضوح خودرو و بهبود نرخ بازشناسی استفاده کرده‌ایم.

یکی دیگر از اهداف رساله، چند کاربرده بودن است. الگوریتم‌های حال حاضر در زمینه‌ی ترافیک، غالباً یا در زمینه تخمین سرعت هستند و یا در زمینه شناسایی نوع خودرو. امکان محاسبه ابعاد، وجود الگوریتم بازشناسی خودرو در ساختار رساله و همچنین وجود پایگاه داده از تصاویر و ویدئوها، این امکان را فراهم می‌کند که نوع خودروهای عبوری به کمک دو معیار "الگوریتم بازشناسی" و "ابعاد تخمین زده شده"، در ویدئو شناسایی شود. بنابراین علاوه بر تخمین ابعاد و سرعت، روشی برای تعیین نوع خودروهای عبوری به‌عنوان کاربرد ثانویه الگوریتم، ارائه شده است.

میانگین خطا در تخمین سرعت برابر با $1/2$ کیلومتر بر ساعت و در تخمین ابعاد برابر با 8% است که مقایسه این مقادیر با نتایج روش‌های دیگر حاکی از پاسخ بهتر روش پیشنهادی در تخمین سرعت و ابعاد است. سرعت پردازش روش پیشنهادی با استفاده از نرم‌افزار متلب ۲۰۱۶ بر روی یک رایانه با حافظه‌ی ۱۶ گیگابایت و پردازنده‌ی Core i7 با فرکانس ۲/۶ گیگاهرتز، حدود ۲/۱۶ قاب در ثانیه به‌دست آمد. نتایج حاصل از الگوریتم‌های بازشناسی و فراتفکیک‌پذیری پیشنهادی، همچنین استفاده از دو معیار برای شناسایی نوع خودروی عبوری در ویدئو، گویای پاسخ خوب همه این روش‌ها است.

کلمات کلیدی: تشخیص سرعت خودرو، تشخیص ابعاد خودرو، کالیبراسیون دوربین، نقطه محوشوندگی

لیست مقالات مستخرج از رساله

مقالات داخلی :

- ۱- تشخیص ابعاد و سرعت خودرو از ویدئوی دریافتی از دوربین کالیبره نشده، رسول عسگریان دهکردی و حسین خسروی، مجله مهندسی برق و کامپیوتر جهاد دانشگاهی
- ۲- الگوریتمی تمام خودکار و مقاوم در برابر سایه، برای تخمین سرعت و ابعاد خودروها از روی تصاویر دوربین‌های نظارتی جاده، رسول عسگریان و حسین خسروی، مجله بینایی ماشین
- ۳- تخمین سرعت و ابعاد خودرو به کمک کالیبراسیون مبتنی بر شناسایی تعدادی خودروی رایج توسط شبکه VGG، رسول عسگریان دهکردی، حسین خسروی، علیرضا احمدی فرد، مجله مهندسی برق دانشگاه تبریز

مقالات خارجی:

- 4- Vehicle Speed and Dimensions Estimation from On-Road Cameras using Common Vehicle Recognition by Bag of Word Method, Hossein Khosravi, Rasoul Asgarian Dehkordi, Alireza Ahmadyfard, IET image processing (Resubmission)
- 5- Vehicle Type Recognition Based on Dimension Estimation and Bag of Word Classification, Rasoul Asgarian Dehkordi, Hossein Khosravi, Journal of AI and Data Mining
- 6- Single Image Super Resolution Based on Sparse Representation Using Dictionaries Trained with Input Image Patches, Rasoul Asgarian Dehkordi, Hossein Khosravi, Alireza Ahmadyfard, IET image processing

۷- ارائه فرایندی ابتکاری برای شناسایی نوع خودروها (در حال نگارش)

مقالات کنفرانسی :

- ۸- عسگریان دهکردی، رسول و علیرضا احمدی فرد "افزایش تدریجی وضوح تصویر با یادگیری جزئیات از دیکشنری وصله‌های تصویر" دهمین کنفرانس بینایی ماشین و پردازش تصویر ایران، اصفهان، انجمن ماشین بینایی و پردازش تصویر ایران، دانشگاه صنعتی اصفهان، آذرماه ۱۳۹۶.
- ۹- عسگریان دهکردی، رسول و حسین خسروی "شناسایی خودکار ابعاد خودروها بر اساس ویدیوی دریافتی از دوربین کنار جاده‌ای" دهمین کنفرانس بینایی ماشین و پردازش تصویر ایران، اصفهان، انجمن ماشین بینایی و پردازش تصویر ایران، دانشگاه صنعتی اصفهان، آذرماه ۱۳۹۶.

۱۰- عسگریان دهکردی، رسول و حسین خسروی "روشی سریع و کارآمد برای حذف سایه خودروهای متحرک به‌منظور تخمین محدوده دقیق خودرو در تصاویر دریافتی از دوربین جاده‌ای" سومین کنفرانس پردازش سیگنال و سیستمهای هوشمند ایران، شاهرود، دانشگاه صنعتی شاهرود، آذرماه ۱۳۹۶.

فهرست مطالب

ن	فهرست جدول‌ها
س	فهرست شکل‌ها
ق	فهرست واژگان
۱	فصل اول : مقدمه
۲	۱-۱- بیان مسئله
۳	۲-۱- چالش‌ها
۵	۳-۱- عمده مشکلات روش‌های پیشین
۶	۴-۱- اهداف مد نظر رساله و مراحل روش پیشنهادی
۸	۵-۱- مسائل پیش‌روی روش پیشنهادی و راه حل ارائه شده برای هر کدام
۱۲	۶-۱- نوآوری‌های رساله
۱۴	۷-۱ ساختار رساله
۱۷	فصل دوم: مروری بر کارهای گذشته
۱۸	۱-۲- مقدمه
۱۸	۲-۲- آشکارسازی خودروهای متحرک و مقابله با سایه
۱۹	۱-۲-۲- مدل کردن پس‌زمینه و شناسایی اجسام متحرک
۲۰	۱-۲-۲-۱- شار نوری
۲۱	۲-۲-۱-۲- تفاضل دو قاب متوالی
۲۱	۳-۲-۱-۲- تفاضل پس‌زمینه
۲۳	۴-۲-۱-۲- مدل مخلوط گوسی
۲۵	۲-۲-۲- روش‌های حذف سایه

۳۲	۳-۲- روش‌های کالیبراسیون غیرخودکار و خودکار.....
۳۲	۳-۲-۱- روش‌های غیرخودکار.....
۳۵	۳-۲-۲- روش‌های خودکار.....
۴۱	۴-۲- شناسایی نوع خودرو.....
۴۲	۴-۲-۱- روش‌های مبتنی بر شبکه‌های عصبی عمیق.....
۴۵	۴-۲-۲- روش‌های مبتنی بر استخراج ویژگی و اعمال به کلاسه‌بند.....
۴۷	۴-۲-۵- فراتفکیک‌پذیری.....
۵۳	۴-۲-۶- نتیجه‌گیری.....
۵۷	فصل سوم: روش پیشنهادی برای کالیبراسیون تمام‌خودکار و تخمین ابعاد و سرعت
۵۸	۳-۱- مقدمه.....
۵۸	۳-۲- مراحل روش پیشنهادی برای تخمین ابعاد و سرعت.....
۶۰	۳-۲-۱- کالیبره کردن دوربین با توجه به جاده و خودروها (فاز اول).....
۶۱	۳-۲-۲- شناسایی دو نوع خودروی رایج به‌منظور محاسبه ضرایب متریک (فاز دوم).....
۶۲	۳-۲-۲-۱- مدل کردن پس‌زمینه و شناسایی پیش‌زمینه.....
۶۳	۳-۲-۲-۲- حذف سایه.....
۶۹	۳-۲-۲-۳- تشکیل جعبه خودرو.....
۷۱	۳-۲-۲-۴- تصویر کردن جعبه خودرو بر صفحه جاده و محاسبه ابعاد آن.....
۷۳	۳-۲-۲-۵- شناسایی خودروهای رایج به‌منظور به‌دست آوردن ضرایب متریک.....
۷۸	۳-۲-۳- تخمین سرعت و ابعاد واقعی خودرو (فاز سوم).....
۸۱	۳-۳- نتیجه‌گیری.....

فصل چهارم : ارائه ساختاری نوین برای بازشناسی تصاویر خودرو و پیشنهاد یک روش جدید برای افزایش تفکیک پذیری به منظور بهبود بازشناسی

۸۴-۱-۴-.....مقدمه ۸۴

۸۴-۲-۴- مکانیزم پیشنهادی برای بازشناسی تصویر..... ۸۴

۸۵-۱-۲-۴-.....مرحله آموزش ۸۵

۸۷-۲-۲-۴-.....مرحله آزمایش ۸۷

۹۰-۳-۴- پیشنهاد یک روش فراتفکیک پذیری جهت بهبود بازشناسی تصویر..... ۹۰

۹۱-۱-۳-۴- وصله بندی تصویر ورودی، ساخت واژه نامه و ایجاد یک تخمین اولیه از تصویر فراتفکیک-
پذیر شده..... ۹۱

۹۳-۲-۳-۴- پیشنهاد یک روش تکرارپذیر برای ساخت تصویر فراتفکیک پذیر شده..... ۹۳

۹۷-۴-۴- نتیجه گیری..... ۹۷

فصل پنجم : شناسایی نوع خودروی عبوری به کمک ابعاد و الگوریتم های بازشناسی

۱۰۰-۱-۵-.....مقدمه ۱۰۰

۱۰۲-۲-۵- بررسی مقدار شباهت خودروی عبوری به هر کلاس با مقایسه ابعاد (معیار دوم)..... ۱۰۲

۱۰۴-۳-۵- تعیین نوع خودروی عبوری به کمک معیارهای اول و دوم..... ۱۰۴

۱۰۵-۴-۵- نتیجه گیری..... ۱۰۵

فصل ششم: بررسی نتایج

۱۰۸-۱-۶-.....مقدمه ۱۰۸

۱۰۸-۲-۶- پایگاه داده های ویدئو و تصویر تهیه شده..... ۱۰۸

۱۱۰-۳-۶- نتایج روش پیشنهادی در تخمین ابعاد و سرعت..... ۱۱۰

۱۱۰-۱-۳-۶- بررسی زمان پردازش هر مرحله ی روش پیشنهادی..... ۱۱۰

۱۱۱-۲-۳-۶- بررسی پاسخ روش پیشنهادی هنگام وجود یا عدم وجود سایه..... ۱۱۱

۱۱۲-۳-۳-۶- مقایسه نتایج روش پیشنهادی با روش های دیگر..... ۱۱۲

۱۱۶	۴-۶- تخمین نوع خودروی عبوری.....
۱۱۷	۴-۶- ۱- بررسی دقت روش‌های بازشناسی.....
۱۱۸	۴-۶- ۲- تاثیر الگوریتم فراتفکیک‌پذیری پیشنهادی بر دقت بازشناسی.....
۱۱۹	۴-۶- ۳- اهمیت حذف سایه در تشخیص نوع خودرو.....
۱۲۰	۴-۶- ۴- شناسایی نوع خودرو به کمک ابعاد و الگوریتم‌های بازشناسی.....
۱۲۱	۴-۶- ۵- نتیجه‌گیری.....

فصل هفتم: نتیجه‌گیری و پیشنهادهایی برای ادامه کار ۱۲۳

۱۲۴	۷-۱- مقدمه.....
۱۲۴	۷-۲- جمع‌بندی روش‌های پیشنهادی در رساله.....
۱۲۶	۷-۳- ایرادات روش پیشنهادی.....
۱۲۸	۷-۴- پیشنهادهایی برای بهبود در آینده.....
۱۳۱	پیوست الف: روش کیفیات.....
۱۴۰	پیوست ب: روش یادگیری عمیق.....
۱۴۶	پیوست ج: ثبت‌کردن تصویر.....

مراجع ۱۵۳

فهرست جدول‌ها

۵۳	جدول ۱-۲: مقایسه بین روش‌های آشکارسازی خودرو.....
۵۴	جدول ۲-۲: مقایسه بین روش‌های خودکار و غیرخودکار کالیبراسیون.....
۵۴	جدول ۳-۲: مقایسه بین روش‌های بازیابی تصویر خودرو.....
۵۵	جدول ۴-۲: مقایسه بین روش‌های فراتفکیک‌پذیری.....
۸۹	جدول ۱-۴: تعیین نوع یک خودروی عبوری توسط شبکه CNN.....
۸۹	جدول ۲-۴: تعیین نوع یک خودروی عبوری توسط مکانیزم بازشناسی پیشنهادی.....
۱۰۴	جدول ۱-۵: تعیین نوع خودرو به کمک ابعاد.....
۱۰۵	جدول ۲-۵: تشخیص کلاس خودرو به کمک معیار بازشناسی و معیار ابعاد.....
۱۱۱	جدول ۱-۶: زمان مصرفی برای هر محاسبات هر مرحله از روش پیشنهادی
۱۱۲	جدول ۲-۶: بررسی خطای روش پیشنهادی در حالت وجود و عدم وجود سایه.....
۱۱۳	جدول ۳-۶: مقایسه میانگین خطای تخمین ابعاد و سرعت چند روش.....
۱۱۸	جدول ۴-۶: دقت روش‌های بازشناسی در طبقه‌بندی خودروها.....
۱۱۸	جدول ۵-۶: بررسی اثر فراتفکیک‌پذیری بر روی دقت بازشناسی.....
۱۱۹	جدول ۶-۶: بررسی اهمیت حذف سایه در تشخیص کلاس.....
۱۲۰	جدول ۷-۶: بررسی دقت در شناسایی خودروها برای ویدئوهای مورد آزمایش.....

فهرست شکل‌ها

- شکل ۱-۱: نمونه‌ای از تصویر دریافتی از دوربین‌های لیزری ناجا برای اندازه‌گیری سرعت..... ۲
- شکل ۲-۱: پاسخ یک روش شناسایی پیش‌زمینه در حضور سایه..... ۴
- شکل ۳-۱: نمونه تصویر یک پس‌زمینه ۵
- شکل ۴-۱: نمونه‌ای از روندنمای روش پیشنهادی..... ۸
- شکل ۱-۲: نتیجه روش RCNN در شناسایی نوع خودرو ۱۸
- شکل ۲-۲: تصاویری از پس‌زمینه و خودروهای مربوط به یک جاده..... ۲۰
- شکل ۳-۲: نتایج روش تفاضل پس‌زمینه [۱۸]..... ۲۲
- شکل ۴-۲: هیستوگرام تصاویر تفاضل و مقادیر آستانه بالا و پایین در هر رنگ..... ۲۳
- شکل ۵-۲: نتایج روش [۳۲] در تفکیک خودرو، جاده و سایه..... ۲۶
- شکل ۶-۲: محدوده رنگ متعلق به سایه و پس‌زمینه [۳۳]..... ۲۶
- شکل ۷-۲: نتایج حذف سایه روش لیون و همکاران [۳۴]..... ۲۷
- شکل ۸-۲: پاسخ روش [۳۸] در تعیین محدوده خودرو هنگام وجود سایه..... ۲۸
- شکل ۹-۲: مراحل روش لی و همکاران [۳۹] برای جداسازی پیش‌زمینه و پس‌زمینه..... ۲۹
- شکل ۱۰-۲: نمایش فرایند حذف سایه در روش زیائو [۴۰]..... ۳۰
- شکل ۱۱-۲: نمایش تغییرناپذیری لبه‌ها نسبت به تغییرات نوری..... ۳۰
- شکل ۱۲-۲: نمایش روندنمای روش [۴۱] در تعیین محدوده سایه و خودرو..... ۳۱
- شکل ۱۳-۲: نتایج حذف سایه روش [۴۱] در طول ۱۳ تکرار..... ۳۲
- شکل ۱۴-۲: تصاویر خطوط جاده برای تعیین نقطه محوشوندگی [۴۴]..... ۳۳
- شکل ۱۵-۲: نمایش راستای حرکت خودروها و طول خطوط جاده در روش یانگ [۴۶]..... ۳۳
- شکل ۱۶-۲: مراحل روش [۴۷] برای طبقه‌بندی خودروها..... ۳۴
- شکل ۱۷-۲: نمایش روندنمای و مستطیل انتخابی در روش تخمین سرعت ارائه شده در [۴۸]..... ۳۵
- شکل ۱۸-۲: نمایش خودرو و دوربین پلاک‌خوان [۵۲]..... ۳۶
- شکل ۱۹-۲: شناسایی پلاک‌ها به‌منظور کالبراسیون دوربین [۵۲]..... ۳۷
- شکل ۲۰-۲: مراحل استخراج کاراکترهای پلاک [۵۳-۵۴]..... ۳۷
- شکل ۲۱-۲: نمایش صفحه زمین و جاده و نحوه محاسبه جابه‌جایی در [۵۵]..... ۳۸
- شکل ۲۲-۲: نمایش وسعت دید دوربین جاده‌ای و تردد بالای خودروها در آزادراه [۵۶]..... ۳۹
- شکل ۲۳-۲: نمایش راستای نقاط محوشوندگی در [۱]..... ۴۱
- شکل ۲۴-۲: نمونه تصاویر خودروی به‌کار رفته در [۶۷]..... ۴۳
- شکل ۲۵-۲: آشکارسازی محدوده خودروها توسط روش [۷۰]..... ۴۴

شکل ۲-۲۶: ثبت تصویر از خودروها به منظور ایجاد مجموعه تصویر [۷۳]	۴۵
شکل ۲-۲۷: نمایش تقارن خودرو در [۷۴]	۴۵
شکل ۲-۲۸: نمایش یک نمونه‌ی واژه‌نامه در [۸۹]	۴۹
شکل ۲-۲۹: نمایش نحوه ساخت واژه‌نامه‌ها و نحوه بازسازی تصویر فراتفکیک پذیر شده در	
روش تسای [۹۳]	۵۱
شکل ۲-۳۰: نمایش واژه‌نامه‌های ساخته شده در روش ژو [۹۵]	۵۲
شکل ۲-۳۱: نمایش ارتباط وصله‌ها در هرم تصویر و نمایش نحوه بازسازی هر وصله در	
[۹۶]	۵۲
شکل ۳-۱: مراحل روش پیشنهادی برای تخمین ابعاد و سرعت	۶۰
شکل ۳-۲: نمایش راستای نقاط محوشدگی برای دو ویدئو	۶۱
شکل ۳-۳: مراحل فاز دوم تخمین ابعاد و سرعت	۶۲
شکل ۳-۴: شناسایی خودروی متحرک به عنوان پیش‌زمینه	۶۲
شکل ۳-۵: پاسخ روش [۹۸] در انتخاب محدوده در حضور سایه	۶۳
شکل ۳-۶: نمایش پیش‌زمینه و پس‌زمینه برای محدوده‌ای از یک جاده	۶۴
شکل ۳-۷: نمایش لبه‌های پس‌زمینه، پیش‌زمینه و مرز خودرو	۶۴
شکل ۳-۸: نمایش لبه‌های پیش‌زمینه و خودرو	۶۶
شکل ۳-۹: نمایش محدوده پیش‌زمینه (کادر قرمز) و محدوده تقریبی خودرو (کادر سبز)...	۶۶
شکل ۳-۱۰: نمایش خودروهای درون جاده برای یک محدوده زمانی، ضلع‌های A، B، C و D	
کادر محدوده‌ی خودرو	۶۷
شکل ۳-۱۱: نمایش محدوده تقریبی خودروها	۶۸
شکل ۳-۱۲: نمایش محدوده دقیق خودروها	۶۹
شکل ۳-۱۳: نمایش تصویر محدوده خودرو و لبه‌های آن	۷۰
شکل ۳-۱۴: تشکیل جعبه محیطی خودرو	۷۱
شکل ۳-۱۵: نمایش دوربین، صفحه تصویر و صفحه جاده دل خواه [۶۰]	۷۲
شکل ۳-۱۶: نتایج روش BOW در بازشناسی چند خودرو	۷۴
شکل ۳-۱۷: معماری VGG ۱۹	۷۵
شکل ۳-۱۸: نمودارهای accuracy و loss در طی فرایند آموزش شبکه VGG	۷۵
شکل ۳-۱۹: نتایج VGG ۱۹ در شناسایی نوع دقیق خودرو	۷۶
شکل ۳-۲۰: مراحل فاز سوم روش پیشنهادی در تخمین ابعاد و سرعت	۷۹
شکل ۳-۲۱: هیستوگرام‌های مربوط به بردارهای ابعاد و جابه‌جایی یک خودرو در مسیر	
حرکت	۸۱
شکل ۴-۱: مکانیزم بازشناسی پیشنهادی	۸۵
شکل ۴-۲: فرایند آموزش شبکه پیشنهادی	۸۶

۸۷	 شکل ۳-۴: نمایش تعدادی خودروی مرجع و تعدادی خودروی تثبیت شده
۸۸	 شکل ۴-۴: فرایند آزمایش مکانیزم پیشنهادی
۸۸	 شکل ۵-۴: تصویر ورودی و نمونه‌های تثبیت شده آن
۹۱	 شکل ۶-۴: روش پیشنهادی برای افزایش تفکیک پذیری
۹۳	 شکل ۷-۴: مراحل بخش دوم روش پیشنهادی برای تفکیک پذیری
	 شکل ۸-۴: نمایش تصویر یک خودرو و بزرگ شده‌ی آن توسط روش فراتفکیک پذیری
۹۶	 پیشنهادی
	 شکل ۵-۱: روندنمای مربوط به نحوه تعیین نوع خودروی عبوری در کنار تخمین ابعاد و
۱۰۱	 سرعت
۱۰۲	 شکل ۵-۲: برخی تشخیص‌های اشتباه توسط روش BOW
۱۰۸	 شکل ۶-۱: نحوه تهیه ویدئوهای مجموعه داده
۱۰۹	 شکل ۶-۲: نمایش سرعت و نمودار تردد خودروها در برخی ویدئوهای مجموعه داده
۱۱۰	 شکل ۶-۳: نمونه‌ای از تصاویر جمع‌آوری شده
۱۱۵	 شکل ۶-۴: مقایسه دقت در تشکیل جعبه محیطی خودرو
۱۱۷	 شکل ۶-۵: تصویر دوربین نظارتی سطح شهر (تهیه شده از اینترنت)
۱۱۹	 شکل ۶-۶: محدوده شناسایی شده برای خودرو قبل و بعد از حذف سایه
۱۲۷	 شکل ۷-۱: نمایش برخی مشکلات روش پیشنهادی
۱۲۸	 شکل ۷-۲: انتخاب بخشی از پس‌زمینه به‌عنوان خودرو
۱۲۹	 شکل ۷-۳: تشکیل جعبه محیطی خودرو به کمک شبکه‌های عمیق

علامہ اختصاری

BG	Background
BOW	Bag Of Word
CNN	Convolutional Neural Network
FG	Foreground
GMM	Gussian Mixture Model
HMM	Hidden Markov Model
HOG	Histogram Of Gradient
HR	High Resolution
ITS	Intelligent Transport System
LR	Low Resolution
MAP	Mean Average Precision
MLE	Maximum likelihood estimation
OMP	Orthogonal Matching Pursuit
PCA	Principal Component Analysis
ROI	Region Of Interest
RNN	Recurrent Neural Network
SIFT	Scale Invariant Feature Transform
SR	Super Resolution
SURF	Speeded Up Robust Features
SVD	Singular Value Decomposition
SVM	Support Vector Machine
VGG	Visual Geometry Group
VMMR	Vehicle Make and Model Recognition
VP	Vanishing Point
YOLO	You Look Only Once

واژه نامه

A

Accuracy

Automatic

B

Background

Bag Of Word

C

Calibration

Car Bounding Box

Classification

D

Dataset

Dimention

Downsamplin

E

Edge

F

Foreground

G

Gaussian Mixture Mode

Ground Truth

H

Hidden Markov Model

Histogram Of Gradient

I

Image Pyramids

Iteration

M

Maximum Likelihood Estimation

Mean Average Precision

N

Nearest Neighbor

O

Optical Flow

Object Class Recognition

P

Patch

Precesion

Projection

R

Recall

Registration

Recognition Rate

Region of Interest

Rigid

Road Side Camera

Resolution

S

Segmentation

Shadow

Speed

Super Resolution

Support Vector Machine

T

Tracking

U

Upsampling

V

Validation Set

Vanishing Point

نزدیک ترین همسایه

شار نوری

تشخیص کلاس شی

وصله

دقت

تصویر کردن

بازیافت

انطباق تصویر

نرخ بازشناسی

ناحیه مورد علاقه

سخت

دوربین کنار جاده ای

وضوح

قطعه بندی

سایه

سرعت

فراتفکیک پذیری

ماشین بردار پشتیبان

ردیابی

نمونه برداری افزایشی

مجموعه اعتبارسنجی

نقطه محوشوندگی

Vehicle Detection	آشکارسازی وسیله نقلیه
Vehicle Make and Model Recognition	شناسایی نوع و مدل وسیله نقلیه

فصل اول: مقدمه

۱-۱- بیان مسئله

مواردی همچون تخمین سرعت خودروها، شناسایی نوع خودرو و شمارش تعداد خودروها، از جمله نیازهای اساسی در نظارت و کنترل ترافیک است. روش‌های رایجی که در حال حاضر در این زمینه‌ها به کار می‌روند با محدودیت‌هایی همراه هستند. به عنوان مثال برای سنجش سرعت خودرو از روش‌هایی مانند رادار و لیزر استفاده می‌شود. این روش‌ها فقط قادر به شناسایی سرعت خودرو هستند و اطلاعاتی در مورد کلاس خودرو و نوع آن ارائه نمی‌دهند، علاوه بر این هزینه سخت‌افزاری بالایی نیز دارند. نمونه‌ای از تصاویر دوربین‌های لیزری در شکل ۱-۱ دیده می‌شود. همان‌طور که در شکل پیدا است، این دوربین‌ها تنها سرعت یک خودرو را که در معرض پرتو لیزر قرار دارد، اعلام می‌کنند.



شکل ۱-۱: نمونه‌ای از تصویر دریافتی از دوربین‌های لیزری ناجا برای اندازه‌گیری سرعت

امروزه تعداد دوربین‌های کنار جاده‌ای که به منظور نظارت و کنترل ترافیک به کار می‌روند، به سرعت در حال افزایش است. بررسی و تحلیل ویدئوهای ضبط شده‌ی این دوربین‌ها توسط کاربر انسانی صورت می‌گیرد که با افزایش روزافزون تعداد دوربین‌ها، هزینه‌ی زیادی دربر دارد. یکی از زمینه‌های مورد علاقه در بینایی ماشین و پردازش تصویر در سال‌های اخیر، استفاده از ویدئوی دوربین‌های نظارتی جاده برای تحلیل ترافیک و به دست آوردن سرعت، ابعاد، نوع و سایر پارامترهای مربوط به وسایل نقلیه و ترافیک

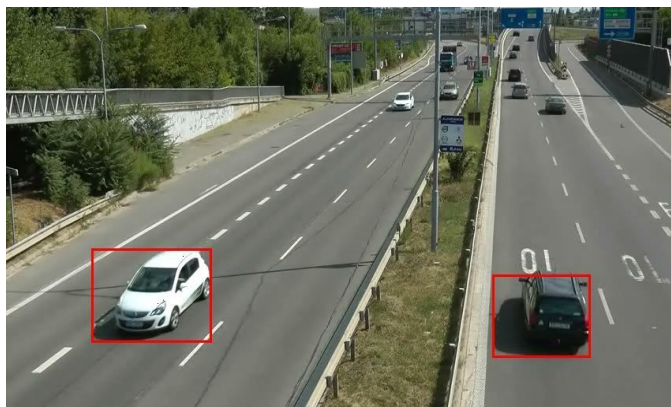
است. با توجه به استقرار این دوربین‌ها در اکثر جاده‌ها، استفاده از روش‌های بینایی ماشین هزینه چندان‌ی بر سامانه تحمیل نمی‌کند.

مساله اصلی مد نظر این رساله، ارائه روشی تمام‌خودکار برای تخمین سرعت و ابعاد خودروها است. تخمین سرعت خودرو مهم‌ترین نیاز سامانه‌های نظارتی است، اگر بتوان با استفاده از ویدئوی نظارتی سرعت خودروها را تخمین زد، نیاز به استفاده از دوربین‌ها و سخت‌افزارهای گران قیمت برطرف می‌شود. تشخیص ابعاد خودرو می‌تواند در کنار الگوریتم‌های بازشناسی، امکان شناسایی کلاس و نوع خودرو را به‌عنوان خروجی ثانویه‌ی الگوریتم، فراهم کند.

۱-۲- چالش‌ها

استفاده از ویدئوی دوربین‌های نظارتی برای تخمین کمیت‌های مد نظر چالش‌هایی را برای الگوریتم‌های بینایی ماشین ایجاد می‌کند که برخی عبارتند از:

۱. **شرایط آب و هوایی و نوری:** شرایط نور محیط با تغییر وضعیت خورشید از ابتدای صبح تا شب تغییر می‌کند. انعکاس نور توسط محیط و بدنه وسایل نقلیه باعث ایجاد خطا در تشخیص صحیح خودروها می‌شود. علاوه بر این در هنگام ظهر و تابش خورشید، به‌دلیل ایجاد سایه، سایه به همراه خودرو به‌عنوان پیش‌زمینه تشخیص داده می‌شود و در محاسبات سرعت، ابعاد و حتی تخمین نوع، خطای زیادی پیش می‌آید. تصویر شکل ۱-۲، پاسخ یک روش شناسایی پیش‌زمینه را در برابر سایه نشان می‌دهد.



شکل ۱-۲: پاسخ یک روش شناسایی پیش‌زمینه در حضور سایه

همچنین در صورت وجود ابر در آسمان ممکن است که محیط از حالت آفتابی به ابری و بالعکس تغییر حالت دهد که این موضوع مدل کردن پس‌زمینه و شناسایی دقیق پیش‌زمینه را با مشکل روبرو می‌کند.

۲. نویز^۱ در تصویر: چالش دیگری که ممکن است بوجود آید، نویز در تصویر است. تصاویری که توسط دوربین‌های ترافیکی جمع‌آوری شده، ممکن است دارای کیفیت بالایی نباشند و همچنین دوربین به دلیل وزش باد دچار لرزش شود. همین عوامل سبب افزایش اختلال در صحنه خواهد شد.

۳. افزایش تعداد خودروها و نزدیک بودن خودروها به یک‌دیگر: در این صورت ممکن است تشخیص دقیق یک خودرو با مشکل مواجه شود و امکان تحلیل رفتار آن وجود نداشته باشد.

۴. شتاب گرفتن خودرو و تغییر مسیر حرکت: اگر هدف تخمین سرعت خودروهای عبوری باشد، شتاب گرفتن خودرو منجر به افزایش خطا در تعیین سرعت دقیق آن می‌شود. همچنین اگر خودرو تغییر مسیر ناگهانی داشته باشد، ممکن است تشخیص آن با مشکل مواجه شود.

۵. درهم ریختگی (شلوغی) پس‌زمینه : در اکثر تصاویر ویدئویی که از بزرگراه‌ها به‌دست آمده، تصاویر به دلیل وجود اشیایی مانند درخت، تیر چراغ برق و غیره دارای درهم ریختگی است. وجود این اشیاء، آشکارسازی و استخراج وسیله نقلیه از روی صحنه را با چالش جدی روبرو

^۱Noise

می‌کند، مخصوصاً اگر چند خودرو هم‌زمان در صحنه باشند. تصویر شکل ۳-۱ پس‌زمینه را برای یک ویدئوی جاده‌ای نمایش می‌دهد. همان‌طور که مشخص است چندین تیر چراغ برق در دید دوربین قرار دارد که در صورت عبور خودرو می‌توانند باعث بروز انسداد شوند و شناسایی محدوده دقیق خودروها را با مشکل مواجه کنند. علاوه بر این چندین درخت در تصویر وجود دارد که در اثر وزش باد به حرکت درمی‌آیند و می‌توانند مدل کردن دقیق پس‌زمینه را با مشکلاتی مواجه کنند.



شکل ۳-۱: نمونه تصویر یک پس‌زمینه

۳-۱- عمده مشکلات روش‌های پیشین

روش‌های مرتبطی که تاکنون در زمینه تخمین ابعاد، سرعت و شناسایی نوع خودرو ارائه شده، دارای چند مشکل عمده هستند که تعدادی از این مشکلات عبارتند از:

۱. **غیرخودکار بودن:** اکثر تحقیقات در این زمینه قابلیت عمل کرد به صورت تمام‌خودکار را ندارند و نیاز به تنظیم دستی دارند.
۲. **عدم مقاومت نسبت به سایه:** سایه تشخیص محدوده دقیق خودرو را با مشکل مواجه می‌کند و خطای بسیاری از الگوریتم‌ها را افزایش می‌دهد.

۳. **تک کاربرده بودن** : اکثر روش‌ها یک کارایی خاص دارند و به‌عنوان مثال فقط سرعت و یا فقط نوع خودرو را شناسایی می‌کنند. همچنین سخت‌افزارهای موجود که برای پیاده‌سازی مورد استفاده قرار می‌گیرند، توان اجرای دو یا چند الگوریتم را ندارند. برای مثال نمی‌توان دو الگوریتم تخمین سرعت و شناسایی نوع خودرو را که کاملاً مجزا از هم هستند، هم‌زمان اجرا کرد.

۴. **عدم آزمایش بر روی داده‌های یکسان و واقعی** : پایگاه داده‌های عمومی فعلی (اعم از تصویر و ویدئو) برای این زمینه محدود هستند و پایگاه داده‌ی ایرانی مناسب برای این منظور وجود ندارد. بسیاری از روش‌های موجود برای سنجش دقت از داده‌های خودشان استفاده کرده‌اند که این داده‌ها غالباً اختصاصی هستند و در اختیار دیگران قرار ندارد. همچنین چه پایگاه‌های اختصاصی و چه پایگاه‌های عمومی دارای ایراداتی هستند، به‌عنوان مثال در ویدئوها تعداد خودروهای عبوری که دارای برچسب سرعت هستند، محدود است و یا تصاویر مورد استفاده برای بازشناسی از زاویه ایده‌آل و دید جلوی خودرو تهیه شده‌اند.

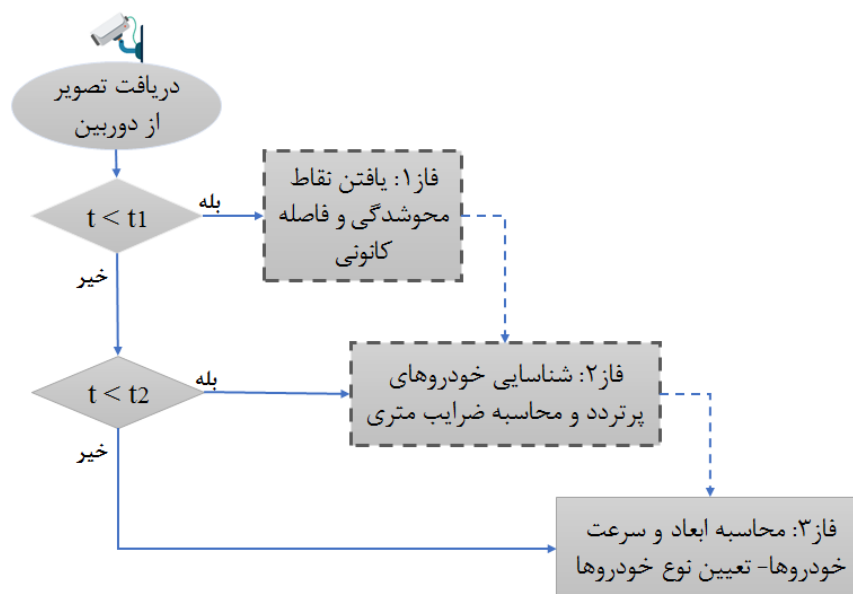
۱-۴- اهداف مد نظر رساله و مراحل روش پیشنهادی

اهداف مد نظر این رساله عبارتند از:

۱. ارائه روشی که در آن کالیبراسیون به‌طور تمام‌خودکار صورت گیرد و نیازمند تنظیم دستی نباشد.
۲. روش پیشنهادی تک کاربرده نباشد و علاوه بر تخمین سرعت و ابعاد، قابلیت شناسایی نوع خودروی عبوری را به‌عنوان خروجی ثانویه داشته باشد.
۳. روش ارائه شده در برابر سایه و تغییرات پس‌زمینه مقاوم باشد. به این دلیل که حذف سایه برای تمامی قاب‌ها صورت می‌گیرد، فرایند حذف سایه نباید زمان‌بر باشد.

۴. به این دلیل که پایگاه جامعی شامل تصاویر و ویدئوهای دارای برچسب خصوصاً در داخل کشور وجود ندارد، مجموعه‌ای برای این منظور جمع‌آوری شود.

برای دستیابی به اهداف ۱ تا ۳، مراحل روش پیشنهادی مطابق شکل ۱-۴ سازمان‌دهی شده است. این روش از سه فاز متوالی تشکیل شده است. در فاز اول و دوم در چند دقیقه اول ویدئوی دریافتی، پارامترهای پیش نیاز فاز سوم محاسبه می‌شوند. در فاز اول نقاط محوشوندگی و فاصله کانونی که در کالیبراسیون کاربرد دارند، محاسبه می‌شوند؛ زمان مورد نیاز برای این فاز (t_1) حدود ۱ دقیقه است. در فاز دوم پس از حذف پرسپکتیو و شناسایی چند خودروی رایج به کمک روش‌های بازشناسی، ابعاد پیکسلی خودروهای رایج محاسبه می‌شود و با توجه به این ابعاد و ابعاد متریک این خودروها، ضرایب متریک به دست می‌آید. (ضرایب متریک برای تبدیل اندازه‌گیری‌های پیکسلی به متری به کار می‌رود. با محاسبه این ضرایب، پارامترهای مربوط به فرایند کالیبراسیون تکمیل می‌شود). ابعاد پیکسلی بدون پرسپکتیو خودروهای رایج، امکان تعیین نوع دقیق خودروها در فاز سوم را نیز فراهم می‌کند؛ زمان مورد نیاز برای پایان فاز دوم (t_2) تقریباً ۴ دقیقه است. نهایتاً در فاز سوم که فاز دائمی الگوریتم است، ابعاد و سرعت خودروی عبوری و نوع آن (تعیین نوع وابسته به انتخاب کاربر است) مشخص می‌شود.



شکل ۴-۱: نمونه‌ای از روندنمای روش پیشنهادی

۱-۵- مسائل پیش‌روی روش پیشنهادی و راه‌حل ارائه شده برای

هر کدام

• کالیبراسیون

از اصلی‌ترین مسائل در رابطه با استفاده از دوربین‌های جاده‌ای برای تحلیل ترافیک، کالیبراسیون است. کالیبراسیون به معنای تخمین پارامترهای داخلی و خارجی دوربین است. پارامترهای داخلی، شامل فاصله کانونی (f_x, f_y) ، نقطه‌ی مرکز تصویر (o_x, o_y) و اعوجاج شعاعی دوربین است و پارامترهای خارجی مربوط به مکان (بردار T) و زاویه قرارگیری دوربین (ماتریس R) است. کالیبره کردن دوربین امکان مقابله با پرسپکتیو و اندازه‌گیری ابعاد واقعی اجسام موجود در صحنه را فراهم می‌کند. اکثر روش‌های موجود برای کالیبراسیون از نقاط محوشوندگی بهره می‌برند. خطوط موازی موجود در صحنه به دلیل وجود پرسپکتیو در تصویر، در نقاطی با یک‌دیگر برخورد می‌کنند که این نقاط، نقاط محوشدگی تصویر نامیده می‌شوند [۱].

اکثر روش‌هایی که تاکنون برای کالیبراسیون دوربین ارائه شده‌اند نیازمند تعیین برخی پارامترهای مربوط به جاده یا دوربین مانند فاصله بین خطوط جاده [۲] یا متوسط سرعت خودروها در جاده [۳] و یا ارتفاع دوربین از سطح جاده [۴و۵]، توسط کاربر هستند. به این صورت که برای استفاده از این سامانه ها باید متناسب با محل نصب دوربین، اطلاعاتی به سامانه داده شود و به‌عنوان مثال در صورتی که از سرعت متوسط خودروها استفاده شود، به‌دلیل تقریبی بودن سرعت متوسط تخمینی، در کالیبراسیون و سنجش سرعت خطا ایجاد می‌شود.

برای استفاده راحت‌تر از روش‌های تحلیل ترافیک، نیاز به روشی است که به‌طور تمام‌خودکار و بدون نیاز به تنظیم دستی امکان استفاده در جاده‌ها را داشته باشد. همچنین برای تحلیل دقیق‌تر رفتار خودروها در جاده روشی مد نظر است که اطلاعات جامع‌تری در مورد خودروها در اختیار قرار دهد [۶].

روش‌های تمام‌خودکاری که تاکنون در زمینه‌ی کالیبراسیون دوربین‌های جاده‌ای در دستگاه‌های حمل و نقل هوشمند ارائه شده است از اطلاعاتی مانند ابعاد پلاک خودروها، متوسط آماری ابعاد خودروها و ابعاد خودروی غالب برای کالیبره کردن و تبدیل اندازه‌گیری‌های پیکسلی به متریک استفاده می‌کنند [۷]. در صورت استفاده از ابعاد پلاک برای کالیبراسیون، وجود خطا در استخراج محل دقیق پلاک خودرو منجر به افزایش خطای کالیبراسیون می‌شود. استفاده از اطلاعات آماری متوسط ابعاد خودروها نیز نیازمند صرف زمان زیادی برای گذر تعداد قابل قبولی خودرو از دید دوربین است (منجر به افزایش زمان کالیبراسیون می‌شود)، علاوه بر این با خطا هم مواجه است. شناسایی خودروی غالب به‌منظور کالیبراسیون علاوه بر نیاز به صرف زمان زیاد برای گذر تعداد زیادی خودرو از دید دوربین، در صورتی که تنوع خودرو در یک کشور زیاد باشد به‌راحتی امکان‌پذیر نیست. علاوه بر این انجام کالیبراسیون با توجه به ابعاد فقط یک نوع خودرو (خودروی غالب) می‌تواند در تخمین سرعت و ابعاد سایر خودروها تأثیر منفی داشته و خطای کلی را افزایش دهد.

روی کرد اصلی این رساله ارائه روشی برای یافتن ابعاد و سرعت خودروها به طور تمام خودکار است. برای کالیبراسیون تمام خودکار نیاز است ابعاد متریک برخی اجسام موجود در صحنه در دست باشد. لذا برای این منظور در این رساله راه کاری پیشنهاد کردیم که در آن با استفاده از روش های بازشناسی نوع خودرو و معیارهای مناسب، چند خودرو با دقت بالا شناسایی می شوند و به کمک ابعاد واقعی این خودروها، پارامترهای کالیبراسیون به طور تمام خودکار محاسبه می شود.

● تشخیص نوع خودرو

شناسایی نوع دقیق خودرو به عنوان یک موضوع پر چالش، توجه بسیاری از محققین حوزه ITS^۱ را در چند سال گذشته به خود جلب نموده است. در حال حاضر اکثر سیستم های نظارتی از شماره پلاک خودرو برای شناسایی خودروهای متخلف استفاده می کنند. این احتمال وجود دارد که سیستم های ANPR^۲ نتوانند شماره پلاک دقیق را استخراج کنند و یا پلاک به هر دلیلی مخدوش شده باشد. بنابراین شناسایی نوع خودرو می تواند در شناسایی خودروهای مشکوک کمک شایانی برای پلیس فراهم آورد. شناسایی خودرو کاربردهای مختلف دیگری نیز دارد، به عنوان مثال آمارگیری از تعداد و انواع مدل های خودرو می تواند اطلاعات مناسبی در مورد سنگینی ترافیک (برای کنترل ترافیک) و فراوانی خودروها برای (فروشندهان خودرو و لوازم جانبی خودرو) همراه داشته باشد.

در روش پیشنهادی با شناسایی نوع چند خودرو در قاب های ابتدایی، پارامترهای مربوط به کالیبراسیون محاسبه شده و فرایند کالیبراسیون تکمیل می شود. برای بازشناسایی نوع خودرو در این رساله هم از روش های مبتنی بر استخراج ویژگی استفاده کردیم و هم ساختاری جدید و قوی، مبتنی بر شبکه های عصبی عمیق ارائه نمودیم.

یکی از اهداف افزایش تفکیک پذیری تصاویر، افزایش نرخ بازشناسی است. به دلیل فاصله دار بودن دوربین جاده ای از خودروها، ممکن است تصاویر دوربین جاده ای کیفیت مناسبی نداشته باشند. بنابراین

^۱ Intelligent Transport System

^۲ Automatic Number-Plate Recognition

افزایش تفکیک‌پذیری می‌تواند باعث بهبود نرخ بازشناسایی خودرو شود. از این‌رو در این رساله روشی جدید مبتنی بر بیان تُنک^۱ برای افزایش تفکیک‌پذیری تصاویر پیشنهاد کردیم. قابلیت تخمین ابعاد در روش پیشنهادی در کنار روش‌های شناسایی نوع خودرو، امکان شناسایی دقیق‌تر نوع خودروها را فراهم می‌کند؛ لذا در این رساله روشی ترکیبی برای شناسایی نوع خودروی عبوری به‌عنوان خروجی الگوریتم پیشنهاد کردیم.

● پایداری در برابر سایه

اکثر روش‌های موجود برای جداسازی خودروهای متحرک از جاده در ویدئو از مدل کردن پس‌زمینه و شناسایی پیش‌زمینه بهره می‌برند. در این رساله نیز از همین روش‌ها استفاده می‌شود. روش‌هایی که تاکنون برای شناسایی پیش‌زمینه ارائه شده‌اند در صورت بروز نیم‌سایه معمولاً فقط خودرو را به‌عنوان پیش‌زمینه در نظر می‌گیرند و مشکلی پیش نمی‌آید. اما در صورت بروز سایه کامل این روش‌ها مقاومت مناسبی ندارند و سایه خودرو را نیز به‌عنوان پیش‌زمینه و بخشی از خودرو در نظر می‌گیرند. در نظر گرفتن سایه به‌عنوان بخشی از خودرو، مشکلات فراوانی را برای سامانه‌های تشخیص نوع خودرو، سرعت خودرو و تردد شمار فراهم می‌کند [۷]، از طرف دیگر روش‌هایی که تاکنون برای مقابله با سایه خودروها ارائه گردیده، بسیار محدود است.

در این رساله هدف این است که برای کاهش خطا و ایجاد قابلیت استفاده در عمل، مقاومت در برابر سایه افزایش یابد به‌گونه‌ای که فرایند حذف سایه، بار محاسبات بالایی به الگوریتم اعمال نکند. لذا برای این منظور آزمایش‌های زیادی انجام گرفت و روشی سریع پیشنهاد شد.

● ویدئوها و تصاویر آموزش و آزمون

برای شناسایی خودروهای رایج نیاز به تهیه‌ی مجموعه‌ای از تصاویر مربوط به خودروها و آموزش این تصاویر به الگوریتم‌های بازشناسی است. همچنین برای آزمایش روش در تشخیص سرعت و ابعاد نیاز به

^۱ Sparse Representation

ویدئوهای جاده‌ای دارای برچسب است. متأسفانه در حال حاضر، پایگاه مناسبی از تصاویر خودروها و ویدئوهای جاده‌ای کشور ایران وجود ندارد. با توجه به این که برای آزمایش دقیق نیاز به ویدئوها و اطلاعات خودروهای ایرانی است، همچنین علاقه و هدف ما ایجاد یک مجموعه داده برای سایر پژوهش‌ها است، با کسب مجوزهای لازم برای فیلم‌برداری، مجموعه‌ای از تصاویر و ویدئوهای برچسب خورده تهیه کردیم.

۱-۶- نوآوری‌های رساله

در این رساله روشی تمام‌خودکار و مقاوم در برابر سایه برای محاسبه سرعت، ابعاد و شناسایی نوع خودرو ارائه شده است. با توجه به گستردگی مراحل، این روش در چندین بخش دارای نوآوری است که در ادامه معرفی می‌شوند.

۱. ایجاد مقاومت در برابر سایه یکی از اهداف این رساله است. برای این منظور الگوریتمی قوی و سریع ارائه شده است که فصل سوم بخش ۳-۲-۲ به‌طور کامل به تشریح این الگوریتم پرداخته است. این الگوریتم برتری قابل توجهی برای روش پیشنهادی نسبت به بسیاری از روش‌های حوزه ITS ایجاد کرده است.

۲. کالیبراسیون تمام‌خودکار یکی از مزیت‌های اصلی روش این رساله است. برای این منظور روشی جدید در بخش ۳-۲-۲-۵ پیشنهاد شده است. در این روش ابتدا با ردیابی خودروها در ویدئو به کمک الگوریتم‌های بازشناسی، نماینده‌هایی برای چند خودروی رایج شناسایی می‌شوند؛ سپس به کمک روش‌هایی دقیق، ضرایب متریک با توجه به ابعاد متریک و ابعاد پیکسلی (پس از حذف پرسپکتیو) این خودروها محاسبه می‌شوند و پارامترهای مورد نیاز برای کالیبراسیون تکمیل می‌شود.

۳. بخش ۳-۲-۳ که همان فاز سوم در تخمین ابعاد و سرعت است، بیان‌گر نوآوری دیگر این رساله برای کاهش هر چه بیشتر خطا در تخمین سرعت و ابعاد خودروهای متحرک است. در این

بخش با ردیابی تک-تک خودروها، هیستوگرام‌هایی برای ابعاد و سرعت آن ایجاد می‌شود و به کمک این هیستوگرام‌ها دقت در گزارش مقادیر ابعاد و سرعت افزایش می‌یابد. همچنین در سایر بخش‌های فصل سوم، بهبودها و نوآوری‌هایی صورت گرفته که منجر به کاهش خطا می‌گردد.

۴. یکی از محورهای کلیدی نوآوری شماره ۲، بازشناسی خودروهای رایج توسط الگوریتم بازشناسی است. همچنین یکی از خروجی‌های الگوریتم پیشنهادی، تشخیص نوع خودروی عبوری به کمک ابعاد و الگوریتم بازشناسی است. بنابراین دقت روش بازشناسی اهمیت زیادی دارد (خصوصاً این‌که زاویه‌ی دید دوربین از کنار جاده است و نه از روبروی خودروها). بر همین اساس در فصل چهارم بخش ۴-۲، ساختاری جدید برای بازشناسی تصاویر خودرو ارائه گردید. در این ساختار که در آن از تثبیت^۱ تصویر و شبکه‌های عصبی عمیق استفاده می‌شود، نحوه آموزش و آزمایش متفاوت از روش مرسوم است. ثبت تصاویر به کمک تبدیلات هندسی و نحوه آموزش و آزمایش در این ساختار، افزایش زیادی در قدرت این روش ایجاد می‌کند؛ به‌نحوی که حتی با استفاده از یک شبکه CNN ساده، قابلیت رقابت با بهترین الگوریتم‌های بازشناسی نوع خودرو را دارد.

۵. ارائه یک روش جدید برای افزایش تفکیک‌پذیری تصاویر با هدف افزایش دقت الگوریتم بازشناسی خودرو، نوآوری دیگر این رساله است. این روش در بخش ۴-۳ بیان شده است. مقایسه کمی و کیفی نتایج این روش با روش‌های دیگر مطرح در فراتفکیک‌پذیری بیان‌گر توان خوب آن است. در فصل ۶ جدول ۶-۵ نشان داده شده که اعمال این روش بر روی تصاویر خودرو، منجر به بهبود دقت در سه روش بازشناسی شده است. این موضوع بیان‌گر این است که این روش ضمن افزایش ابعاد تصویر، جزئیات تصویر را با دقت خوبی بازسازی می‌کند. این مطلب از این لحاظ نیز حائز اهمیت است که برای اولین بار تاثیر فراتفکیک‌پذیری روی دقت

^۱ Registration

بازشناسی خودرو مورد بررسی قرار می‌گیرد و سایر پژوهش‌ها این موضوع را بررسی نکرده‌اند. این درحالی است که با توجه به فاصله زیاد دوربین جاده‌ای از خودروها، افزایش تفکیک‌پذیری به کمک روشی مناسب می‌تواند در بهبود عمل‌کرد موثر باشد.

۶. نوآوری دیگر این رساله ارائه روشی ترکیبی برای شناسایی نوع خودروهای عبوری است. در این روش که در فصل ۶ ارائه شده است، توسط دو معیار وابستگی خودروهای عبوری به هر کلاس مشخص می‌شود. معیار اول مقایسه ابعاد پیکسلی بدون پرسپکتیو با مقادیر مرجع برای هر کلاس و معیار دیگر استفاده از الگوریتم بازشناسی است. استفاده از ابعاد و الگوریتم بازشناسی نوآوری شماره ۴، امکان تشخیص نوع خودرو از روی ویدئویی مشابه ویدئوی دوربین جاده‌ای را فراهم نموده است. این درحالی است که روش‌های بازشناسی که تاکنون ارائه شده‌اند از تصاویر با کیفیتی که از دید جلوی خودرو تهیه شده برای شناسایی استفاده کرده‌اند.

۷. با توجه به این‌که پایگاه مناسبی از تصاویر خودروها و ویدئوهای جاده‌ای خصوصاً برای کشور ایران وجود ندارد، دو پایگاه شامل مجموعه‌ای از تصاویر خودروها و مجموعه‌ای از ویدئوهای دارای برچسب جمع‌آوری گردید که جزئیات آن‌ها در فصل ۶ ارائه شده است.

۱-۷ ساختار رساله

در این فصل مبانی اولیه موضوع و مساله تحقیق، چالش‌ها، مشکلات روش‌های پیشین، اهداف و روندنمای روش پیشنهادی و نوآوری‌های این رساله بررسی شد. در فصل دوم این رساله، برخی تئوری‌ها و کارهای پیشین مرتبط بررسی می‌گردد. در فصل سوم ساختار روش پیشنهادی در تخمین ابعاد و سرعت به اجمال مورد بررسی قرار می‌گیرد و هر کدام از مراحل آن بیان می‌شود. در فصل چهارم روش پیشنهادی برای بازشناسی خودرو و همچنین روش پیشنهادی برای افزایش تفکیک‌پذیری تصاویر خودرو که منجر به بهبود دقت بازشناسی خودرو می‌شود، بررسی خواهد شد. در فصل پنجم روش ترکیبی که از ابعاد و الگوریتم بازشناسی برای تعیین نوع خودروهای رایج استفاده می‌کند، بیان می‌شود. در فصل

ششم بخش ۲-۶ پایگاه داده جمع‌آوری شده، معرفی می‌گردد. در بخش ۳-۶ نتایج روش پیشنهادی در تخمین ابعاد و سرعت از لحاظ زمان پردازش، تاثیر حذف سایه و مقایسه کمی و کیفی با سایر روش‌ها بررسی می‌گردد. در بخش ۴-۶ نتایج مباحث مربوط به تخمین نوع خودرو که شامل مقایسه دقت روش‌های بازشناسی، تاثیر الگوریتم فراتفکیک‌پذیری بر دقت بازشناسی، اهمیت حذف سایه در تشخیص نوع خودرو و شناسایی نوع خودرو در ویدئو به کمک ابعاد و الگوریتم‌های بازشناسی بیان خواهد شد. در نهایت در فصل هفتم نتیجه‌گیری رساله بیان شده و پیشنهادهایی برای ادامه کار ارائه می‌شود.

ساختار روش پیشنهادی به‌گونه‌ای است که با موضوعاتی از جمله شبکه‌های CNN^۱، روش BOW^۲ و تثبیت تصویر ارتباط دارد. لیکن جهت انسجام متن رساله، این مطالب به‌صورت پیوست بیان شده‌اند.

^۱ Convolutional Neural Network

^۲ Bag-Of-Words

فصل دوم: مروری بر کارهای گذشته

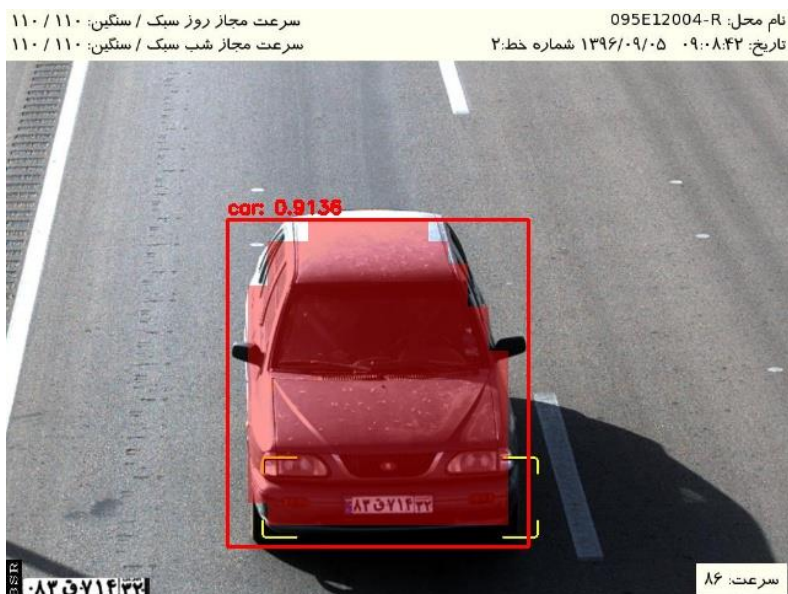
۲-۱- مقدمه

در ساختار روش پیشنهادی رساله، شناسایی محدوده‌ی دقیقی از خودرو دارای اهمیت است. بنابراین در این فصل ابتدا در مورد مدل‌سازی پس‌زمینه برای آشکارسازی خودروهای متحرک و روش‌های مقابله با سایه به تفصیل بحث می‌شود. از اهداف اصلی ما، ارائه روشی تمام‌خودکار است؛ لذا موضوع بعدی که در این فصل پوشش داده می‌شود، روش‌های خودکار و غیرخودکار کالبراسیون هستند. در روش رساله پیشنهاد شده که کالبراسیون به کمک شناسایی خودروها صورت گیرد، از این‌رو در ادامه این فصل در مورد بازشناسی خودرو و روش‌های فراتفکیک‌پذیری (اعمال این روش‌ها بر روی تصویر می‌تواند دقت بازشناسی خودرو را افزایش دهد)، بحث خواهد شد.

۲-۲- آشکارسازی خودروهای متحرک و مقابله با سایه

آشکارسازی دقیق محدوده‌ی خودروهای متحرک در روش پیشنهادی دارای اهمیت است و صحت عمل کرد الگوریتم تا حد زیادی به این مرحله وابسته است. برای شناسایی نوع اجسام اخیراً روش‌هایی مبتنی بر شبکه‌های عصبی عمیق ارائه شده که قادرند با دقت خوبی نوع اجسام را شناسایی کنند. به‌عنوان مثال شبکه‌های RCNN[۸] و YOLO[۹] حدود ۹۰ کلاس از اجسام را شناسایی می‌کنند که خودرو نیز در این کلاس‌ها قرار دارد. هرچند این شبکه‌ها قادرند خودرو را از زوایای گوناگون و با دقت بالا شناسایی کنند، اما دارای دو مشکل هستند: ۱. سرعت کم و ۲. عدم دقت کافی در تعیین مرز صحیح خودرو خصوصاً هنگام تابش خورشید

تصویر شکل ۲-۱ نتیجه تشخیص یک خودرو را برای شبکه RCNN نشان می‌دهد، همان‌طور که معلوم است محدوده دقیقی از خودرو مشخص نشده است.



شکل ۲-۱: نتیجه روش [۸] RCNN، برای شناسایی نوع خودرو

عدم تشخیص محدوده دقیق خودرو در روش پیشنهادی منجر به کاهش دقت این روش در تعیین ابعاد می‌شود و از این‌رو استفاده از این شبکه‌ها برای تعیین محدوده‌ی خودرو در روش پیشنهادی منتفی است. برای آشکارسازی خودروهای متحرک غالباً از روش‌های شناسایی پیش‌زمینه استفاده می‌شود. با توجه به ثابت بودن دوربین جاده‌ای و پاسخ خوب این روش‌ها بر روی ویدئو، بهترین راه برای شناسایی محدوده‌ی خودروها در الگوریتم پیشنهادی استفاده از همین روش‌ها است. همچنین برای این که بتوان در صورت تابش خورشید و بروز سایه، محدوده‌ی خودرو را به‌دست آورد، نیاز به ارائه روشی برای مقابله با سایه است. در ادامه روش‌های شناسایی پیش‌زمینه و مقابله با سایه بررسی می‌شود.

۲-۲-۱- مدل کردن پس‌زمینه و شناسایی اجسام متحرک

تصویر (الف) در شکل ۲-۲ تصویر پس‌زمینه و تصویر (ب) در این شکل تصویر پس‌زمینه همراه با خودرو (پیش‌زمینه) را برای یک جاده نشان می‌دهد. این تصاویر در یک محدوده زمانی کوتاه گرفته شده است.



ب: وجود خودرو به عنوان پیش‌زمینه در جاده



الف: تصویر پس‌زمینه

شکل ۲-۲: تصاویری از پس‌زمینه و خودروهای مربوط به جاده

تاکنون تکنیک‌های مختلفی برای آشکارسازی اشیاء متحرک ارائه شده است، در میان این روش‌ها استفاده از شار نوری^۱ [۱۰ و ۱۱]، تفاضل پس‌زمینه^۲ [۱۲] و روش مدل مخلوط گوسی^۳ کارا تر است. در ادامه به اختصار در مورد هر کدام از این روش‌ها بحث خواهد شد.

۲-۲-۱-۱- شار نوری

بعضی محققان برای حل مشکل آشکارسازی حرکت در تصویر این روش را ارائه داده‌اند [۱۰]. مدل شار نوری بر روی حرکات کوچک اشیاء متمرکز است و با اعمال روش شار نوری بر روی قاب‌های متوالی می‌توان حرکات موجود در صحنه را آشکار نمود [۱۳]؛ در اکثر موارد ویژگی‌های برجسته هر جسم مانند گوشه توسط این روش ردیابی می‌شود تا مکان جسم در هر قاب تعیین گردد. نقطه ضعف اصلی این روش زمان بالای پردازش آن است.

^۱Optical Flow

^۲ Background Subtraction

^۳Gaussian Mixture Model

۲-۱-۲-۲- تفاضل دو قاب متوالی

روش تفاضل دو قاب متوالی [۱۴-۱۷] ساده‌ترین روش در بخش‌بندی اجسام متحرک است. در این روش تفاضل دو قاب متوالی محاسبه می‌شود و تصویر حاصله با یک سطح آستانه بخش‌بندی می‌شود (رابطه ۱-۲):

$$I(x, y) = \begin{cases} 0 & |I_1(x, y) - I_2(x, y)| < \theta \\ 1 & \text{else} \end{cases} \quad (1-2)$$

در این رابطه، I_1 تصویر قاب قبلی، I_2 تصویر قاب کنونی و θ آستانه انتخابی است. هدف از اعمال آستانه کم کردن اثر نویز، تغییرات آب و هوا و اثر محیط است. هرچند این روش سرعت بالایی دارد، اما ضعف‌هایی دارد که عبارتند از:

- عدم تشخیص وسایل نقلیه ثابت.
- نتایج تحت تاثیر سرعت وسایل نقلیه است.
- در استخراج دقیق محدوده وسایل نقلیه ناتوان است.

۲-۱-۲-۳- تفاضل پس‌زمینه

در این روش برای آشکارسازی اشیاء پیش‌زمینه، تفاضل هر قاب از تصویر پس‌زمینه محاسبه می‌شود و بعد با استفاده از اعمالی مانند آستانه‌گذاری^۱، محل اجسام متحرک شناسایی می‌شود [۱۸-۲۵]:

$$I(x, y) = \begin{cases} 0 & |I_{in}(x, y) - I_B(x, y)| < \theta \\ 1 & \text{else} \end{cases} \quad (2-2)$$

I_{in} قاب ورودی، I_B تصویر پس‌زمینه و θ آستانه انتخابی هستند. در شکل ۳-۲ قاب ورودی، پس‌زمینه و نتایج آستانه‌گذاری مشاهده می‌شود.

^۱Thresholding



الف: قاب ورودی ب: پس زمینه ج: اجسام متحرک شناسایی شده

شکل ۲-۳: نتایج روش تفاضل پس زمینه [۱۸]

در [۲۶] روشی سریع و دقیق برای محاسبه سطح آستانه متغیر ارائه شده که به خوبی مقدار سطح آستانه را با تغییرات نور محیط وفق می دهد. مقدار سطح آستانه بر اساس هیستوگرام تصویر تفاضل پس زمینه $(B(x, y))$ و قاب قبل $(I_{i-1}(x, y))$ محاسبه می شود. از آن جا که کار بر روی تصاویر رنگی انجام می شود، ابتدا تصویر تفاضل، به صورت علامت دار در هر رنگ محاسبه می شود :

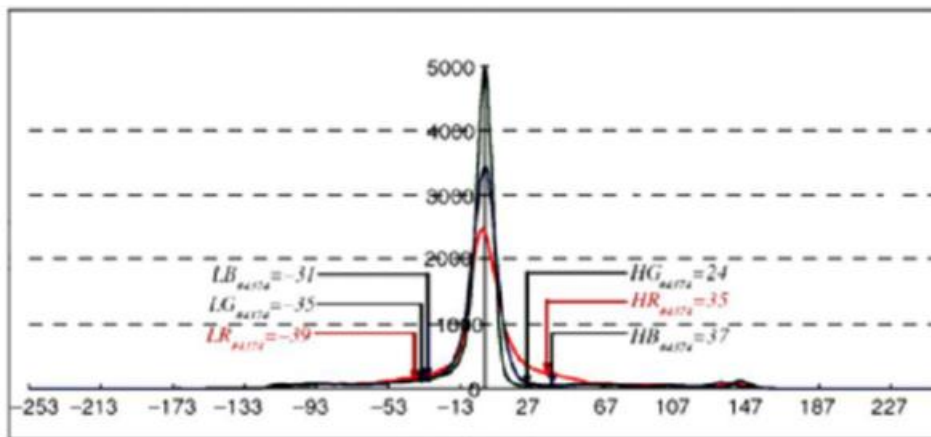
$$\begin{cases} \Delta R(x, y) = R_{i-1}(x, y) - R_B(x, y) \\ \Delta G(x, y) = G_{i-1}(x, y) - G_B(x, y) \\ \Delta B(x, y) = B_{i-1}(x, y) - B_B(x, y) \end{cases} \quad (3-2)$$

در این رابطه R_{i-1} ، G_{i-1} و B_{i-1} به ترتیب مولفه های قرمز، سبز و آبی تصویر I_{i-1} و R_B ، G_B و B_B مولفه های قرمز، سبز و آبی تصویر B هستند. سپس هیستوگرام تصاویر تفاضل به دست می آید و نرم می شود :

$$\begin{cases} H_{i-1,SR}[j] = \frac{\sum_{k=-2}^{k=2} H_{i-1,R}[j+K]}{5} \\ H_{i-1,SG}[j] = \frac{\sum_{k=-2}^{k=2} H_{i-1,G}[j+K]}{5} \\ H_{i-1,SB}[j] = \frac{\sum_{k=-2}^{k=2} H_{i-1,B}[j+K]}{5} \\ j = -253, \dots, -1, 0, 1, \dots, 253 \end{cases} \quad (4-2)$$

از آن جا که تصاویر تفاضل علامت دار هستند، هیستوگرام آن ها شبیه یک تابع گوسی با میانگین صفر است، زیرا اکثر پیکسل ها متعلق به پس زمینه هستند و در تصویر تفاضل نزدیک صفر می شوند. در هر

رنگ دو مقدار آستانه پایین و بالا که به ترتیب مثبت و منفی هستند از روی هیستوگرام آن محاسبه می‌شوند. معیار محاسبه این سطوح، تغییر در شیب منحنی هیستوگرام است و اولین نقاطی که در طرفین هر هیستوگرام این ویژگی را دارند به عنوان سطح آستانه بالا و پایین آن رنگ در نظر گرفته می‌شوند. هیستوگرام تصاویر تفاضل و مقادیر آستانه بالا و پایین رنگ‌ها در شکل ۴-۲ مشاهده می‌شود.



شکل ۴-۲: هیستوگرام تصاویر تفاضل و مقادیر آستانه بالا و پایین در هر رنگ [۲۶]

۴-۱-۲-۲ مدل مخلوط گوسی

در این روش هر پیکسل به صورت ترکیبی از چند توزیع گوسی مدل می‌شود. در [۲۷] استافر و همکاران برای تشخیص وسایل نقلیه عبوری در جاده‌ها از این روش استفاده کرده‌اند. اگر I_n قاب n ام ورودی باشد، تابع چگالی احتمال آن به صورت رابطه ۴-۵ فرض می‌شود:

$$P[I_n(x, y)] = \sum_{i=1}^K w_i(x, y) \times \eta[I_n(x, y), \mu_i(x, y), \Sigma_i(x, y)] \quad (4-5)$$

که در آن η یک تابع چگالی احتمال گوسی به صورت زیر است:

$$\eta(x, \mu, \Sigma) = \frac{1}{(2\pi)^2 |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)} \quad (4-6)$$

K تعداد توزیع‌ها است، w_i وزن لامین توزیع گوسی، μ_i میانگین لامین توزیع گوسی و Σ_i ماتریس کواریانس لامین توزیع گوسی است. ماتریس کواریانس به صورت $\Sigma_i = \sigma_i^2 I$ فرض می‌شود. هر پیکسل جدید با K توزیع گوسی موجود مقایسه می‌شود تا این که یک انطباق پیدا شود. تعریف انطباق این است که مقدار پیکسل کمتر از T برابر انحراف معیار با میانگین آن فاصله داشته باشد. پس از یافتن انطباق، مدل گوسی منطبق شده توسط روابط ۲-۷ تا ۲-۹ به روز می‌شود و میانگین و واریانس بقیه توزیع‌ها بدون تغییر باقی می‌ماند.

$$\mu_i(x, y) = (1 - \beta)\mu_i(x, y) + \beta I_n(x, y) \quad (۷-۲)$$

$$\sigma_i(x, y) = (1 - \beta)\sigma_i(x, y) + \beta [I_n(x, y) - \mu_i(x, y)]^T \times [I_n(x, y) - \mu_i(x, y)] \quad (۸-۲)$$

$$\beta = \alpha \cdot \eta [I_n(x, y), \mu_i(x, y), \sigma_i(x, y)] \quad (۹-۲)$$

همچنین برای تنظیم وزن توزیع‌های گوسی به این طریق عمل می‌شود :

$$w_i(x, y) = (1 - \alpha)w_i(x, y) + \alpha M_k \quad (۱۰-۲)$$

اگر انطباق کامل وجود داشته باشد $M_k = 1$ و در غیر این صورت $M_k = 0$ است. α نرخ به روز رسانی است. اگر هیچ کدام از K توزیع با پیکسل مد نظر انطباق نداشته باشد آن پیکسل متعلق به پیش‌زمینه تشخیص داده می‌شود. در نهایت مقدار پیکسل فعلی به پس‌زمینه و پیش‌زمینه دسته‌بندی می‌شود. برای این کار از این ایده که احتمال مشاهده (w) و تغییرات مقادیر (σ) آن‌ها در پیکسل‌های پس‌زمینه نسبت به پیش‌زمینه کمتر است، استفاده شده است. توجه شود که مقادیر مشاهده شده در طول زمان برای یک پیکسل در صورت وجود خودرو، به خاطر انعکاس متفاوت نور از قسمت‌های مختلف خودرو دارای تغییرات بیشتری نسبت به پس‌زمینه است. همچنین به خاطر تفاوت رنگ وسایل نقلیه مختلف، حتی در جاده‌های پرتردد احتمال مشاهده پس‌زمینه بیشتر است. برای دسته‌بندی پیکسل فعلی به پس‌زمینه یا پیش‌زمینه، ابتدا تمام توزیع‌های گاوسی پس از تطبیق با مقدار فعلی بر حسب σ/w مرتب

می‌شوند. از آن جا که توزیع‌های معرف پس‌زمینه، دارای σ/ω بیشتری هستند، تعداد $B = \operatorname{argmin}_b (\sum_{k=1}^b \omega_k > \Omega)$ توزیع با بیشترین σ/ω به‌عنوان پس‌زمینه و بقیه به‌عنوان پیش‌زمینه دسته‌بندی می‌شوند. پارامتر Ω آستانه وزن نامیده می‌شود.

روش ترکیب مدل‌های گوسی، تغییرات روشنایی در پس‌زمینه و همچنین حرکات ذاتی پس‌زمینه مثل حرکت برگ درختان را به خوبی مدل می‌کند. آزمایش‌های انجام شده بیان‌گر این است که این روش در شناسایی خودروهای متحرک موفق عمل می‌کند.

۲-۲-۲- روش‌های حذف سایه

همان‌طور که در فصل قبل اشاره شد استفاده از ویدیوی دوربین‌های نظارتی چالش‌های گوناگونی دارد، یکی از مهم‌ترین این چالش‌ها بروز سایه خودروها در اثر تابش نور خورشید است. حرکت سایه به همراه وسیله نقلیه حل این مسئله را مشکل‌تر می‌کند [۲۸]. ویژگی‌هایی که معمولاً برای شناسایی پیکسل‌های سایه‌ای استفاده می‌شوند عبارتند از :

۱. اطلاعات فضا‌های مختلف رنگی
۲. اطلاعات مکانی (استفاده از هندسه شی و جهت سایه)
۳. اطلاعات بافت
۴. روش‌های ناحیه محور، وصله محور و مبتنی بر مشخصات فیزیکی که از مفروضات مبتنی بر فیزیک اپتیک استفاده می‌کنند [۲۹-۳۱].

در [۳۲] تصویر رنگی پس‌زمینه و قاب ورودی از فضای RGB به فضای HSV منتقل شده‌اند. اگر تصویر ورودی I_n دارای مولفه‌های HSV به‌صورت I_n^S, I_n^V و I_n^H باشد و تصویر پس‌زمینه B_n دارای مولفه‌های B_n^S, B_n^V و B_n^H باشد، هر پیکسل پیش‌زمینه به شرط برقراری ۱-۲ سایه تلقی می‌شود :

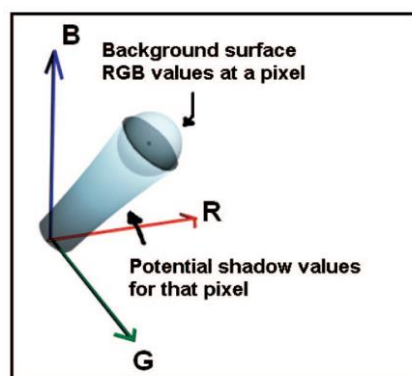
$$\tau_{v,1} < \frac{I_n^V(x,y)}{B_n^V(x,y)} < \tau_{v,2} \text{ AND } [|I_n^S(x,y) - B_n^S(x,y)| < \tau_S] \text{ AND } [|I_n^H(x,y) - B_n^H(x,y)| < \tau_H] \quad (۱۱-۲)$$

در این رابطه $\tau_{v,1}$ نور محیط را کنترل می کند، $\tau_{v,2}$ از تشخیص نویز به جای سایه جلوگیری می کند، τ_H و τ_S اختلاف مقادیر hue و saturation را کنترل می کنند. شکل ۵-۲ نتایج حذف سایه در این روش را نشان می دهد. مشکل این روش انتخاب مقدار آستانه ها و عدم تمایز بین خودروهای سیاه و سایه در بعضی موارد است.



شکل ۵-۲: نتایج روش [۳۲] در تفکیک خودرو، جاده و سایه

مارتل و همکارانش [۳۳] روش GSM^۱ را برای شناسایی سایه ارائه داده اند. در این روش ابتدا با استفاده از روش GMM اجسام متحرک شناسایی می شوند، سپس با استفاده از مدل گوسی که برای سایه به دست آورده اند، پیکسل های سایه ای را شناسایی می کنند. همچنین برای افزایش کارایی، مدل گوسی سایه را در طول زمان به روز می کنند. شکل ۶-۲ از این مرجع ناحیه رنگ مستعد سایه و پس زمینه را برای یک پیکسل نشان می دهد.



شکل ۶-۲: نمایش محدوده رنگ متعلق به سایه و پس زمینه در [۳۳]

^۱ Gaussian Mixture Shadow Model

لیون و همکاران [۳۴] از ویژگی‌های گابور برای مشخص کردن محدوده سایه استفاده کرده‌اند. در [۳۵] و [۳۶] نویسندگان روشی برای جداسازی سایه از اجسام متحرک در ویدئو ارائه داده‌اند. در این روش ابتدا مجموعه‌ای دارای برچسب توسط نویسندگان تهیه شده و ویژگی‌های رنگ و لبه (بافت) این مجموعه به کلاسه‌بند SVM آموزش داده شده است. نویسندگان این روش ادعا کرده‌اند که ویژگی‌های مورد استفاده نسبت به تغییرات نور مقاوم هستند و این روش برای آموزش نیازمند داده‌ی زیاد نیست. تصاویر شکل ۲-۷ نتیجه آزمایش این روش را نشان می‌دهد. مشکل این روش نیاز به تصاویر برچسب‌دار و عدم تفکیک دقیق نواحی جسم و سایه از یک‌دیگر در برخی موارد است.



(ب)



(الف)



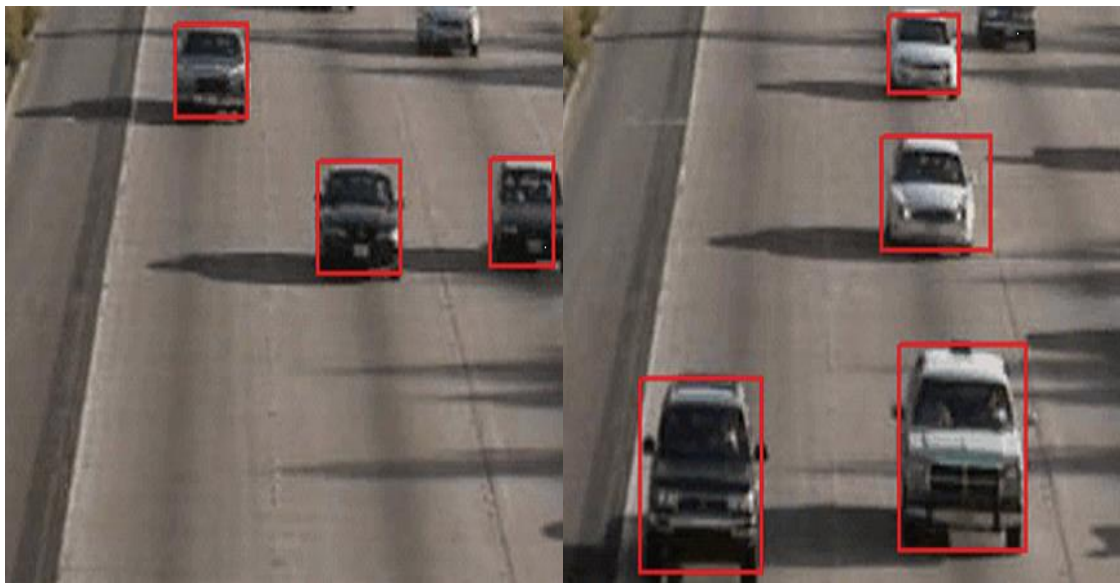
(د)



(ج)

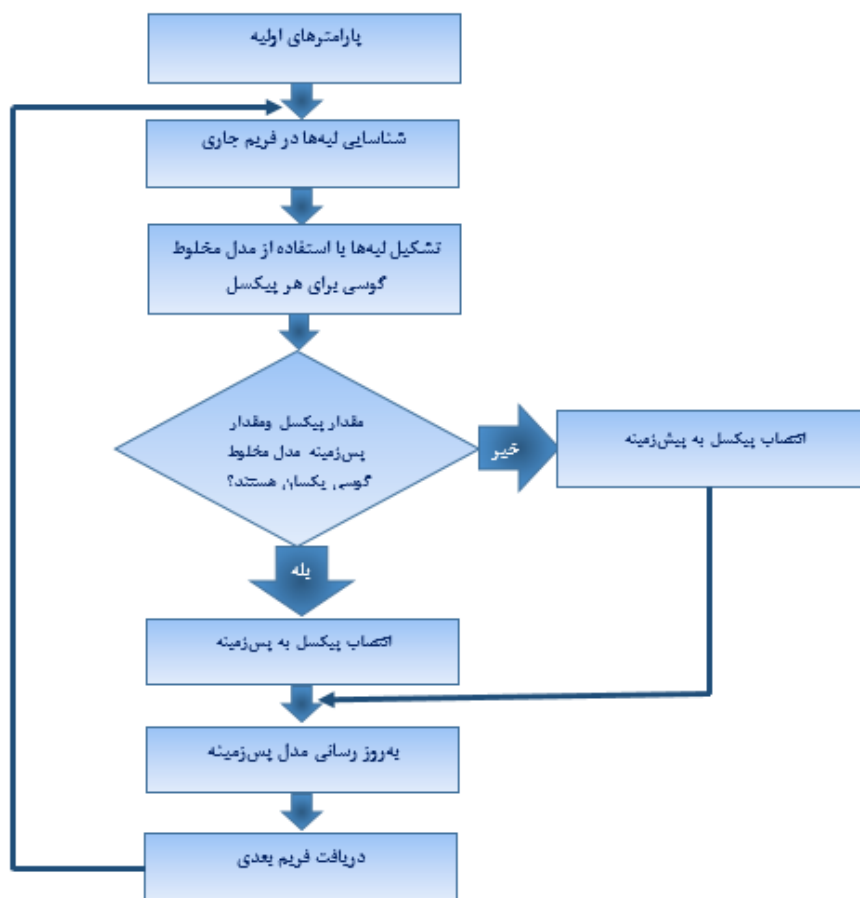
شکل ۲-۷: نتایج حذف سایه روش لیون و همکاران [۳۴]، (الف) و (ج): تصاویر ورودی، (ب) و (د): نتایج حاصل

لبه به عنوان یک ویژگی مقاوم می تواند در شرایط نورپردازی گوناگون مورد استفاده قرار گیرد؛ از این رو در صورت بروز سایه می تواند قابل اتکا باشد [۳۷]. در [۳۸] با استفاده از اطلاعات لبه ای خودروها و پس زمینه، همچنین استفاده از اطلاعات رنگ مربوط به خودروها و پس زمینه، مناطق سایه ای مشخص و با حذف این مناطق محدوده دقیق خودروها به دست آمده است. شکل ۸-۲ پاسخ این روش را در مقابل سایه نشان می دهد.



شکل ۸-۲: پاسخ روش [۳۸] در تعیین محدوده خودرو هنگام وجود سایه

لی و همکاران در [۳۹] مدل ترکیب گوسی را بر روی لبه های پس زمینه اعمال کردند و با تحلیل این مدل محل تغییرات لبه و مرزهای مربوط به پیش زمینه را شناسایی کرده اند. تصویر شکل ۹-۲ مراحل این روش را نشان می دهد. هرچند این روش از سرعت خوبی برخوردار است، در مواردی قادر به طبقه بندی درست پیکسل های پیش زمینه و پس زمینه نیست.



شکل ۲-۹: مراحل روش لی و همکاران [۳۹] برای جداسازی پیش‌زمینه و پس‌زمینه

مراحل روش زیائو و همکارانش [۴۰] عبارت است از :

۱. شناسایی محدوده‌ی خودرو و سایه به‌عنوان پیش‌زمینه

۲. لبه‌یابی از محدوده شناسایی شده

۳. حذف کانتور از لبه‌ها

۴. تعیین محدوده لبه‌های افقی و عمودی (محدوده‌ی خودرو)

۵. تعیین محدوده سایه

تصویر شکل ۲-۱۰ فرایندهای مراحل ۱ تا ۵ این روش را نشان می‌دهد. مشکل این روش این است

که قادر نیست مرز خودرو و سایه را به‌طور دقیق تعیین کند.



الف: تصویر ورودی ب: محدوده پیش‌زمینه ج: لبه‌های پیش‌زمینه د: مرز پیش‌زمینه و پس‌زمینه ه: لبه‌های خودرو بدون مرز



و: محدوده عمودی لبه‌های خودرو بدون مرز ز: محدوده افقی لبه‌های خودرو بدون مرز ح: محدوده لبه‌های خودرو بدون مرز ط: محدوده سایه

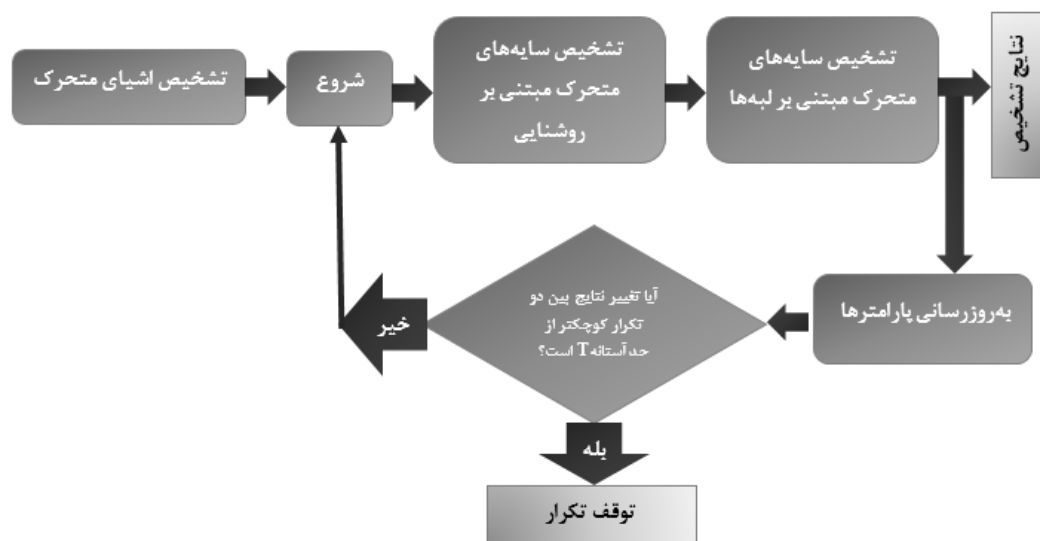
شکل ۲-۱۰: نمایش فرایند حذف سایه در روش زیائو [۴۰]

ژانگ و همکاران [۴۱] یک توزیع آماری مربع-کای برای سایه در نظر گرفته و پیکسل‌های هر جسم متحرک را به دو کلاس سایه و غیر سایه تقسیم‌بندی کردند. در این روش ادعا شده که لبه در تصویر، با شرایط نوردهی تغییر نمی‌کند. شکل ۲-۱۱ تصویر و لبه‌های آن‌ها را در شرایط نوردهی متفاوت نشان می‌دهد.



شکل ۲-۱۱: نمایش تغییرناپذیری لبه‌ها نسبت به تغییرات نوری [۴۱]

البته به هر حال لبه استخراج شده توسط روش‌های عادی ممکن است در بعضی از شرایط (مانند ناحیه همگن در تصویر) بسیار ضعیف باشد و نتواند برای تشخیص سایه استفاده شود. لذا ایشان، پارامتری به نام نرخ لبه را برای آشکارسازی لبه‌های متحرک به کار گرفته‌اند. نرخ لبه به کمک نسبت بین شدت یک پیکسل با پیکسل‌های همسایه آن محاسبه می‌شود. روندنمای این روش در شکل ۲-۱۲ نشان داده شده است.



شکل ۲-۱۲: نمایش روندنمای روش [۴۱] در تعیین محدوده سایه و خودرو

مطابق این روندنمای، به کمک یک فرایند تکرارپذیر ابتدا محدوده کاندیدا برای جسم متحرک با بررسی روشنایی تصویر شناسایی می‌شود و به کمک لبه‌یابی با نرخ لبه و مقایسه لبه‌های پیش‌زمینه و پس‌زمینه اصلاح می‌شود. سپس نرخ لبه با توجه به ناحیه‌ی جدید سایه و جسم اصلاح می‌گردد و این فرایند تا جایی ادامه می‌یابد که در تکرارها محدوده‌های مشخص شده برای سایه و پیش‌زمینه تغییر محسوسی نکند. شکل ۲-۱۳ نتیجه این روش را برای چند تکرار نشان می‌دهد. مشکل اصلی این روش زمان‌بر بودن آن است.

نتیجه تشخیص						
شماره تکرار	۱	۳	۵	۸	۱۰	۱۳

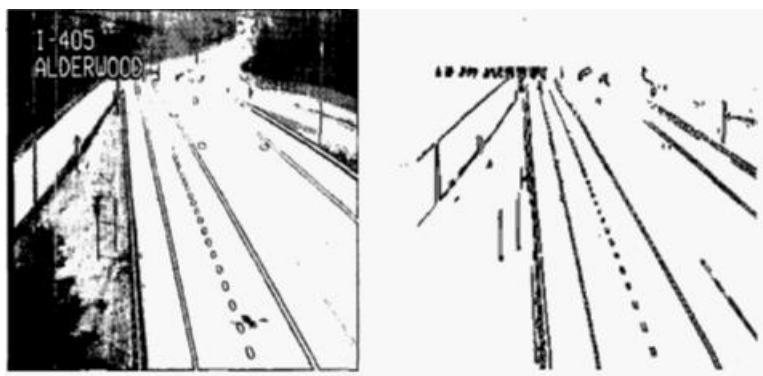
شکل ۲-۱۳: نتایج حذف سایه روش [۴۱] در طول ۱۳ تکرار

۲-۳- روش‌های کالبراسیون غیرخودکار و خودکار

روش‌هایی موجود در زمینه‌ی تحلیل ترافیک به‌طور کلی به دوگروه غیرخودکار (یا نیمه خودکار) و خودکار (تمام‌خودکار) تقسیم می‌شوند [۴۲]. در این بخش ابتدا روش‌های غیرخودکار مورد بررسی قرار می‌گیرد، سپس به بررسی روش‌های خودکار که روش پیشنهادی رساله نیز در این گروه قرار دارد، پرداخته می‌شود.

۲-۳-۱- روش‌های غیرخودکار

اکثر روش‌هایی که تاکنون در زمینه تحلیل ترافیک ارائه شده‌اند، قابلیت عمل کرد به‌طور تمام‌خودکار را دارا نیستند. سیستم ارائه شده در [۴۳] قادر به شمارش خودروها، اندازه‌گیری سرعت و ابعاد خودرو در هر قاب است. این سیستم بر اساس ردیابی و گروه‌بندی نقاط گوشه‌ای هر خودرو عمل می‌کند. برای استفاده از این سیستم کاربر می‌بایست مختصات چهار نقطه را به‌صورت دستی وارد کند، علاوه بر این برای اندازه‌گیری سرعت باید ناحیه‌ای برای ورود و خروج خودروها تعیین شود؛ به‌منظور ردیابی خودروها در این روش فیلتر کالمن به‌کار رفته است. کدی [۴۴] و همکارانش برای کالبره کردن دوربین به‌منظور تخمین سرعت از یک نقطه محوشدگی استفاده کرده‌اند و برای تعیین مختصات این نقطه محوشدگی از راستای خطوط جاده (تصاویر شکل ۲-۱۴) بهره برده‌اند. در این روش برای تبدیل پیکسل به متر، ضریب متریک با توجه به طول خطوط جاده بر حسب پیکسل در تصویر و طول واقعی خطوط جاده بر حسب متر، به‌دست آمده است.



شکل ۲-۱۴: تصاویر خطوط جاده برای تعیین نقطه محوشدگی [۴۴]

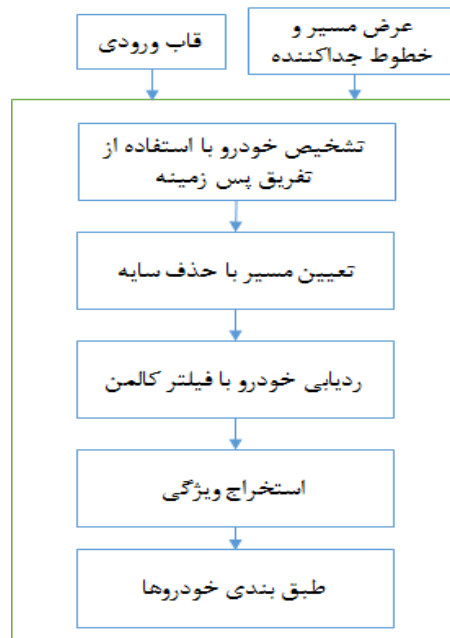
در [۴۵] برای شمارش تعداد خودروها و به دست آوردن سرعت در شب، چراغ‌های خودرو شناسایی و ردیابی می‌شوند؛ برای کالیبره کردن دوربین این روش نیازمند زاویه و ارتفاع دوربین از سطح جاده است. یانگ و همکارانش [۴۶] برای کالیبره کردن دوربین از دو نقطه محوشدگی بهره برده‌اند، نقطه محوشدگی اول را در راستای حرکت خودروها و نقطه دوم را عمود بر سطح جاده در نظر گرفته‌اند. مقیاس در این روش با توجه به ارتفاع دوربین از سطح جاده به دست آمده است. برای بررسی کمی و سنجش دقت در کالیبراسیون، این روش طول واقعی خطوط جاده را با طول تخمینی مقایسه کرده است. شکل ۲-۱۵ برخی تصاویر این روش را که متناسب با توضیحات بیان شده است، نشان می‌دهد.



شکل ۲-۱۵: نمایش راستای حرکت خودروها و طول خطوط جاده در روش یانگ [۴۶]

در [۴۷] روش جدیدی برای طبقه‌بندی خودروها ارائه شده، طرح کلی این روش در شکل ۲-۱۶ نمایش داده شده است. این روش برای تشخیص پس‌زمینه و پیش‌زمینه از روش تفریق پس‌زمینه بهره

می‌برد. پس از اجرای این الگوریتم در چند قاب اول با توجه به حرکت خودروها، مسیرها و خطوط جدا کننده هر مسیر مشخص می‌شود.

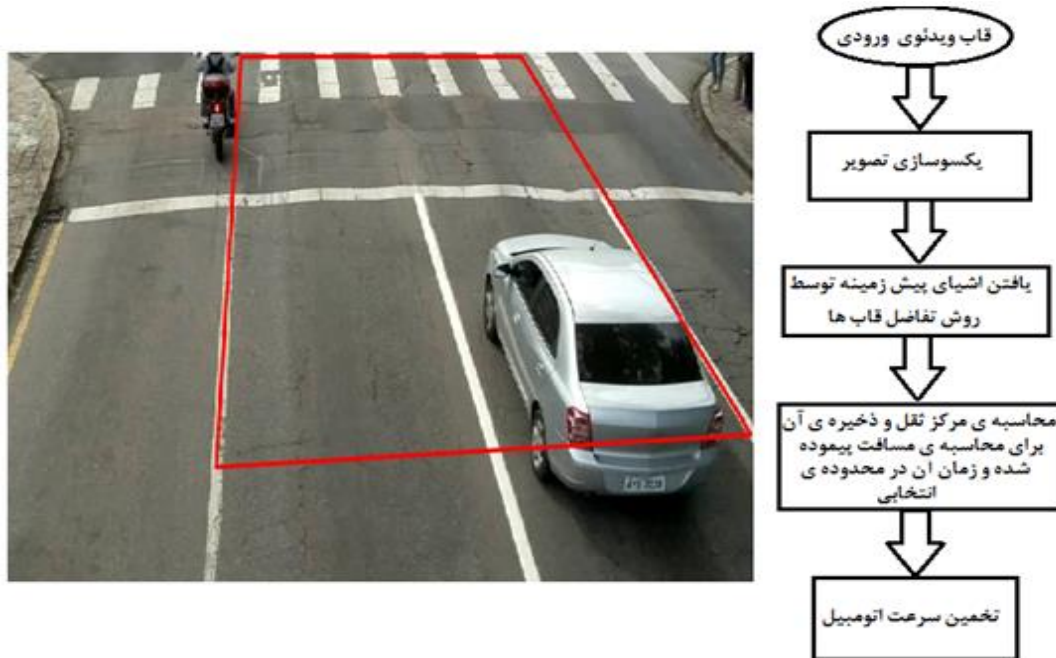


شکل ۲-۱۶: مراحل روش [۴۷] برای طبقه‌بندی خودروها

به‌منظور طبقه‌بندی خودروها در این روش از اطلاعات عرضِ متریک خودرو و اطلاعات مربوط به زاویه و شیب لبه‌های خودرو استفاده می‌شود. عرض خودرو بر حسب متر به‌طور تقریبی با توجه به عرض خطوط جاده تعیین شده است. با مقایسه عرض خودرو با مقادیر عرض مرجع هر کلاس، کلاس‌های محتمل برای خودرو تعیین می‌شود؛ سپس به کمک اطلاعات زاویه و شیب لبه‌ها، کلاس دقیق مشخص می‌شود.

آقایان و خسروی [۴۸]، روشی بلادرنگ برای تخمین سرعت خودروهای عبوری ارائه داده‌اند. در این روش به‌طور دستی، چهار نقطه از جاده که در دنیای واقعی، مستطیل است، به‌عنوان محل ورود و خروج خودروها در نظر گرفته می‌شود. سپس برای حذف پرسپکتیو، با نگاشت چهار نقطه در تصویر واقعی به چهار نقطه مستطیلی در تصویر نگاشت یافته، ماتریس تبدیل پیدا شده و پرسپکتیو حذف می‌شود. ضریب متریک در این روش با توجه به ابعاد واقعی این ناحیه بر حسب متر و ابعاد پیکسلی آن، به‌دست

می‌آید. تصویر (الف) در شکل ۲-۱۷ مراحل این روش و تصویر (ب) مستطیل انتخابی را برای یک جاده در این روش نمایش می‌دهد.



ب: مستطیل انتخابی در روش ارائه شده در [۴۸]

الف: روندنمای روش ارائه شده در [۴۸]

شکل ۲-۱۷: نمایش روندنمای و مستطیل انتخابی در روش تخمین سرعت ارائه شده در [۴۸]

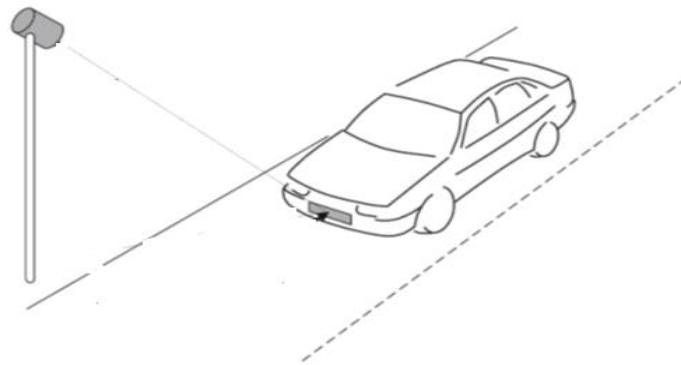
لووینسون و همکارانش [۴۹]، روشی ارائه کردند که در آن ابتدا خودرو آشکار می‌شود، سپس محل پلاک تعیین و به وسیله ردیاب لوکاس-کند [۵۰] ردیابی می‌شود. نهایتاً پس از حذف پرسپکتیو، جابه-جایی دقیق خودرو تعیین می‌شود. برای تخمین پارامترهای کالیبراسیون از یک الگوی مستطیلی روی جاده استفاده شده است.

۲-۳-۲- روش‌های خودکار

محدودیت اصلی روش‌های غیرخودکار نیاز به تنظیم دستی این روش‌ها است که این موضوع استفاده از این روش‌ها را در مکان‌های مختلف با محدودیت‌هایی مواجه می‌کند. علاوه بر این در [۵۱] نویسندگان روش‌های خودکار و غیرخودکار را مورد بررسی قرار داده‌اند و به این نتیجه رسیده‌اند که روش‌های

خودکار در مجموع خطای کمتری نسبت به روش‌های غیرخودکار دارند. تعداد روش‌هایی که قابلیت عمل کرد به‌طور تمام‌خودکار را دارند، محدود است؛ از طرفی این روش‌ها در حالت کلی قابلیت انعطاف-پذیری و کارایی بالاتری نسبت به روش‌های غیرخودکار دارند. در ادامه تعدادی از روش‌های خودکاری که تاکنون ارائه شده بررسی می‌شود.

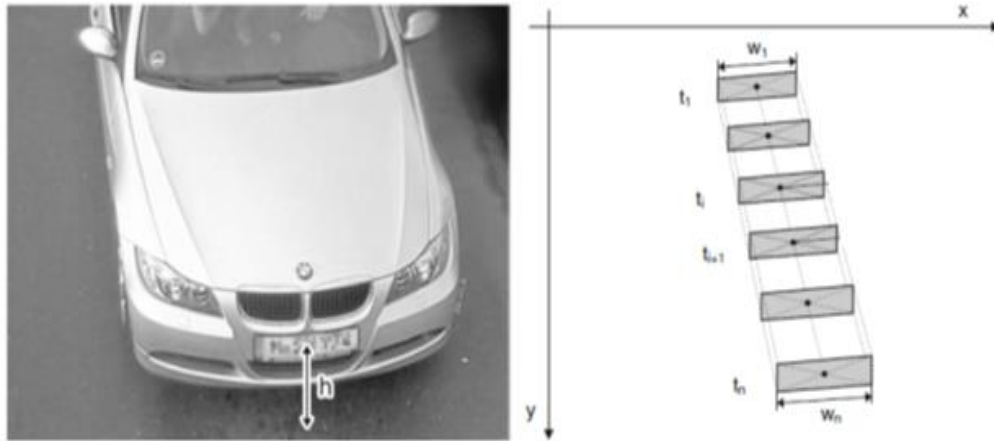
هدف اصلی در [۵۲] به‌دست آوردن سرعت خودروهای نقلیه به‌صورت خودکار با استفاده از دوربین-های پلاک‌خوان است. برای به‌دست آوردن سرعت خودروها معمولاً از دوربین‌های کنار جاده‌ای استفاده می‌شود که دید وسیعی نسبت به دوربین‌های پلاک‌خوان دارند. در این مرجع هدف استفاده از دوربین-های پلاک‌خوان به‌منظور تخمین سرعت خودروها در جاده است. روش پیشنهادی به این صورت است که ابتدا پارامترهای مربوط به مدل دوربین و جاده به‌صورت پارامتری درنظر گرفته می‌شود؛ سپس با توجه به اطلاعات متریک ابعاد پلاک خودروها، سعی می‌شود با شناسایی چندین پلاک، به کمک الگوریتم‌های بهینه‌سازی، پارامترهای دوربین مشخص شود و دوربین کالیبره گردد. تصویر ۲-۱۸ یک خودرو درون جاده به‌همراه پلاک آن و دوربین پلاک‌خوان را نمایش می‌دهد.



شکل ۲-۱۸: نمایش خودرو و دوربین پلاک‌خوان [۵۲]

برای کالیبره کردن دوربین و استفاده از الگوریتم تکاملی، در این روش از محدودیت‌هایی به‌عنوان فرضیه در بهینه‌سازی استفاده شده است. همان‌طور که اشاره شد از اطلاعات ابعاد پلاک برای بهینه‌سازی

استفاده می‌شود (شکل ۲-۱۹). بنابراین بروز خطا در شناسایی محدوده دقیق ابعاد پلاک و یا عدم همگرایی الگوریتم‌های بهینه‌ساز، منجر به خطا در محاسبه سرعت خودروها در این روش می‌گردد.



شکل ۲-۱۹: شناسایی پلاک‌ها به منظور کالیبراسیون دوربین [۵۲]

اسلامی و همکاران [۵۳-۵۴]، روشی ارائه کردند که در آن تعدادی خودرو در قاب‌های متوالی شناسایی می‌شود و مطابق شکل ۲-۲۰ مکان پلاک و کاراکترهای آن‌ها استخراج می‌گردد. در ادامه به کمک ابعاد پلاک و ابعاد کاراکترها، پارامترهای کالیبراسیون محاسبه می‌شوند و نهایتاً با ردیابی، سرعت خودرو به دست می‌آید.



ب: تصویر باینری پلاک



الف: پلاک ورودی



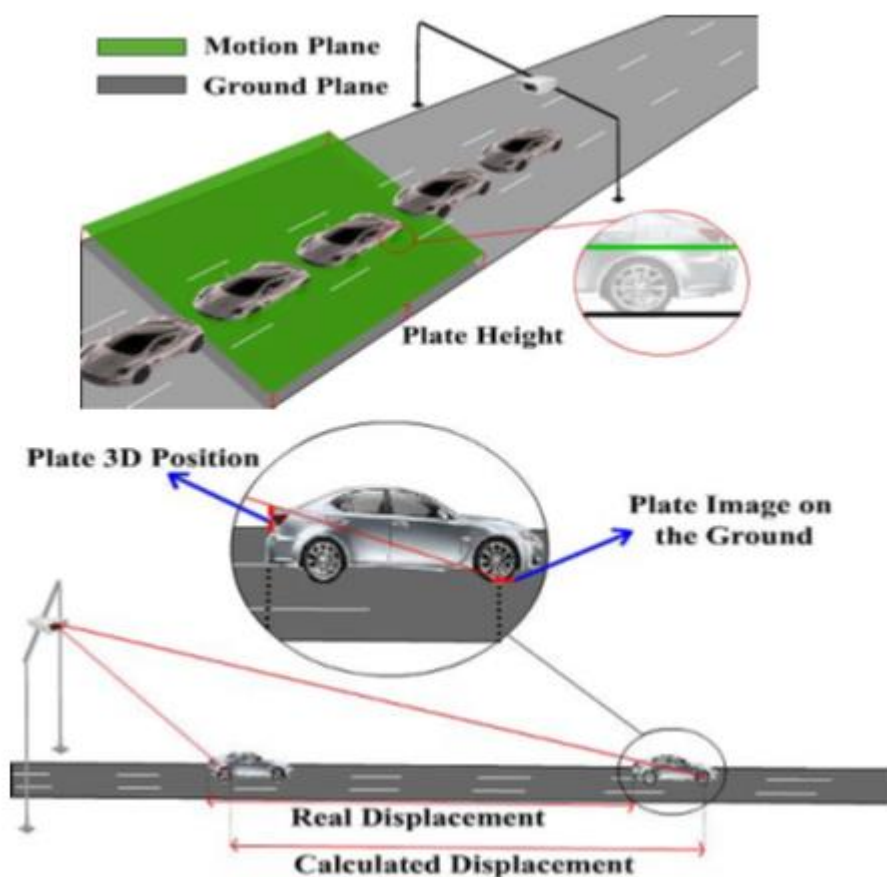
د: انتخاب محدوده هر کاراکتر



ج: باینری پرداخت شده

شکل ۲-۲۰: مراحل استخراج کاراکترهای پلاک [۵۳-۵۴]

روش فاموری و همکارانش [۵۵]، شامل دو فاز است. در فاز اول با شناسایی پلاک خودروها و با در نظر گرفتن حرکت نقطه مرکزی پلاک، صفحه‌ای فرضی که موازی صفحه زمین است، تعیین می‌شود. در فاز دوم با ردیابی و تصویر کردن پلاک خودرو بر روی این صفحه، سرعت خودرو محاسبه می‌شود. این روش برای آشکارسازی خودروها، حدود ۴۰۰۰ پلاک از خودروهای ایرانی را برای آموزش استفاده کرده است. شکل ۲-۲۱ نمونه‌های صفحه فرضی و صفحه زمین را برای این مرجع نمایش می‌دهد. مشکل عمده‌ای که برای روش‌های [۵۳] و [۵۵] وجود دارد و نویسندگان این مقالات نیز به این موضوع اشاره کرده‌اند، احتمال عدم شناسایی دقیق محدوده پلاک خودرو و ایجاد خطا در تخمین سرعت است.



شکل ۲-۲۱: نمایش صفحه زمین و جاده و نحوه محاسبه جابه‌جایی در [۵۵]

دانیل و همکارانش در [۵۶] روشی تمام‌خودکار برای تخمین سرعت وسایل نقلیه در آزادراه‌ها از طریق دوربین کنار جاده‌ای ارائه داده‌اند. با توجه به وسعت دید دوربین جاده‌ای و همچنین تردد بالای

خودروها در آزادراه‌ها، در این مقاله از اطلاعات آماری متریک طول متوسط خودروها به منظور کالیبراسیون و به دست آوردن سرعت خودروها استفاده شده است. رابطه ۲-۱۲ نحوه تخمین سرعت بر حسب کیلومتر بر ساعت در این روش را نشان می‌دهد :

$$v = \frac{3.6 \times \Delta_d}{\Delta_t} \quad \Delta_t = \frac{1}{FR} \quad (2-12)$$

در این رابطه Δ_d بیان گر مقدار جابه‌جایی خودرو بین دو قاب متوالی، FR بیانگر Frame Rate یا تعداد قاب‌های گرفته شده در یک ثانیه از طریق دوربین و Δ_t فاصله زمانی بین دو قاب است. شکل ۲-۲۲ از روش دانیل، وسعت دید و تردد بالای خودروها در آزادراه را نمایش می‌دهد.



شکل ۲-۲۲: نمایش وسعت دید دوربین جاده‌ای و تردد بالای خودروها در آزادراه [۵۶]

به منظور ساده‌سازی مساله، دانیل و همکاران از چندین فرضیه منطقی استفاده کرده‌اند. این فرضیه‌ها عبارتند از [۵۸-۵۶] :

- نرخ قاب دوربین‌های جاده‌ای بالاست.
- به دلیل ملاحظات قانونی و فیزیکی (برای خودروها)، سرعت خودروهای نقلیه خیلی زیاد نیست و خودروها جابه‌جایی آنی ندارند.
- ردیابی خودروها تقریباً در یک بعد است.
- طول وسایل نقلیه از یک توزیع شناخته شده پیروی می‌کند.

دوبسکا و همکارانش در [۱] روشی مناسب برای به دست آوردن نقاط محوشوندگی و فاصله کانونی که در کالیبراسیون کاربرد دارند، ارائه کرده‌اند. این روش به طور خلاصه از مراحل زیر تشکیل می‌شود:

(۱) برای یافتن نقطه محوشدگی اول، خودروها در چند قاب اول با استفاده از روش لوکاس-کند

[۵۰] ردیابی می‌شوند. مسیر حرکت خودروها به دست می‌آید. مختصات نقطه محوشدگی اول

با توجه به حرکت خودروها و در راستای مسیر حرکت مشخص می‌گردد.

(۲) راستای دومین نقطه محوشدگی باید به گونه‌ای باشد که موازی با سطح جاده و تقریباً عمود بر

راستای نقطه محوشدگی اول باشد. بنابراین برای به دست آوردن نقطه محوشدگی دوم ابتدا در

چند قاب لبه‌های پیش‌زمینه به دست می‌آید. سپس لبه‌هایی که در جهت نقطه محوشدگی اول

هستند، حذف می‌شود و مختصات نقطه محوشدگی دوم در راستای لبه‌های حذف نشده در نظر

گرفته می‌شود.

(۳) سومین نقطه محوشدگی باید در راستای عمود بر جاده باشد. دو نقطه محوشدگی اول و دوم

در راستای موازی با سطح جاده هستند. به همین دلیل نقطه محوشدگی سوم در راستای عمود

بر راستای نقاط محوشدگی اول و دوم در نظر گرفته می‌شود.

پس از طی سه مرحله بالا مختصات سه نقطه محوشدگی و فاصله کانونی مشخص می‌شود. شکل

۲-۲۳ راستای نقاط محوشدگی را برای دو ویدئو نمایش می‌دهد. در این شکل محورهای قرمز راستای

نقطه محوشدگی اول، محورهای سبز راستای نقطه محوشدگی دوم و محورهای آبی راستای نقطه

محوشدگی سوم را نشان می‌دهند. همچنین مرکز تصویر با دایره کوچک سبز رنگ و نقطه محوشدگی

اول با دایره قرمز مشخص شده‌اند، نقاط محوشدگی دوم و سوم خارج از کادر تصویر هستند و امکان

نمایش وجود ندارد.



شکل ۲-۲۳: نمایش راستای نقاط محوشوندگی در [۱]

ژاکوب و همکاران در [۵۹] که یک رساله دکتری است، روشی تمام‌خودکار برای تحلیل ترافیک ارائه داده‌اند؛ در [۶۰] بخشی از این رساله توسط ژاکوب، دوبسکا و همکاران به‌صورت یک مقاله ارائه گردیده است. در این روش با استفاده از الگوریتم ارائه شده در [۱] مختصات نقاط محوشدگی و فاصله کانونی در چند قاب به‌دست می‌آید. سپس با مدل کردن پس‌زمینه و شناسایی محدوده‌ی خودروی متحرک، جعبه محیطی خودرو تشکیل می‌شود و نهایتاً با اعمال یک نگاشت برای مقابله با پرسپکتیو و استفاده از ضریب متریک، ابعاد واقعی خودرو بر حسب متر به‌دست می‌آید. ضرب ضریب متریک در اندازه ابعاد بر حسب پیکسل منجر به به‌دست آوردن ابعاد بر حسب متر می‌شود؛ این ضریب در این روش به‌طور خودکار و با استفاده از اطلاعات آماری ابعاد خودروها تعیین می‌شود. برای تعیین سرعت در این روش، خودروها به‌طور جداگانه با استفاده از روش شارنوری ردیابی می‌شوند و مقدار جابه‌جایی و در نتیجه سرعت خودروها تخمین زده می‌شود.

۲-۴- شناسایی نوع خودرو

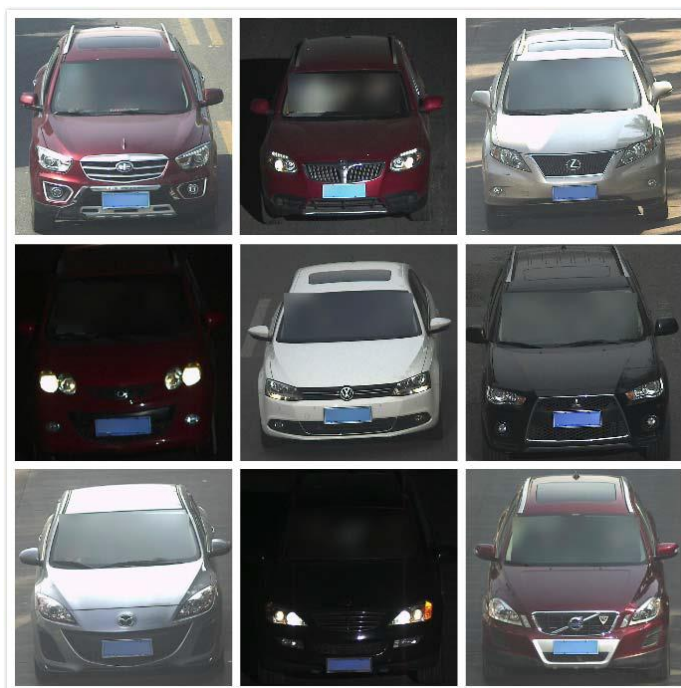
شناسایی نوع خودرو به‌عنوان یکی از زمینه‌های مهم در ITS، در چند سال گذشته توجه محققین خصوصاً در کشورهای پیشرفته را به خود جلب کرده است. در روش این رساله برای تکمیل کالیبراسیون نیازمند شناسایی تعدادی خودرو در ابتدای کار هستیم تا به کمک ابعاد متریک این خودروها پارامترهای مد نظر کالیبراسیون را استخراج کنیم. همچنین یکی از کاربردهای روش پیشنهادی شناسایی نوع خودروها،

به‌ویژه خودروهای رایج است. تاکنون روش‌های معتبری در زمینه شناسایی نوع خودرو ارائه شده که برای مجموعه خودشان دقت‌های خوبی به‌دست آورده‌اند. این روش‌ها تلاش می‌کنند با سنجیدن شباهت ظاهری خودرو با هر کلاس، نوع دقیق آن‌را شناسایی کنند. شباهت برخی بخش‌های خودروها به یک-دیگر (به‌عنوان مثال شباهت در جلو پنجره، لوگو و چراغ‌ها) یکی از عوامل ایجاد خطا در این الگوریتم‌ها است. یکی از مشکلات این پژوهش‌ها کمبود پایگاه داده‌های جامع برای خودروها است، تفاوت تنوع خودروها در هر منطقه با منطقه دیگر این مشکل را تشدید می‌کند. این موضوعات منجر به ایجاد محدودیت‌ها و فرضیه‌هایی در روش‌های ارائه شده در این زمینه گردیده است [۶۳-۶۱]. علاوه بر این طبق بررسی‌های انجام گرفته، اکثر این روش‌ها، شناسایی نوع را با توجه به مجموعه تصاویری انجام داده‌اند که از نمای نزدیک و معمولاً از جلوی خودرو گرفته شده است. این در حالی است که مکان قرارگیری دوربین‌های نظارتی معمولاً چند متر بالاتر از سطح جاده است و تصویری که به‌دست می‌آورند از نمای بالا و فاصله‌دار نسبت به خودرو است، که این موضوع منجر به کاهش کیفیت تصاویر خودرو می‌شود، شناسایی را با چالشی جدی روبرو می‌کند و خطای روش‌ها را در عمل افزایش می‌دهد [۶۶-۶۴]. روش‌های ارائه شده در زمینه بازشناسی تصاویر خودرو در یک دسته‌بندی کلی در دو گروه قرار می‌گیرند: ۱- روش‌های مبتنی بر شبکه‌های عصبی عمیق، ۲- روش‌های مبتنی بر استخراج ویژگی و اعمال به کلاسه‌بند.

۲-۴-۱- روش‌های مبتنی بر شبکه‌های عصبی عمیق

در [۶۷] به این موضوع اشاره شده که اگر دوربین مورد استفاده برای تهیه داده‌های آموزشی با دوربین مورد استفاده برای تهیه تصاویر آزمایش متفاوت باشد، درصد خطای روش‌ها افزایش می‌یابد. برای این‌که وابستگی به دوربین داده‌های آموزشی کاهش یابد، این روش از داده‌های وب که تصاویری با کیفیت مناسب از هر خودرو در اختیار قرار می‌دهند، استفاده کرده است. شبکه عصبی مورد استفاده در این پژوهش معماری شبیه معماری ALEXNET دارد که از پنج لایه کانولوشن و سه لایه pooling و سه

لایه تماماً متصل تشکیل شده است. برای آشکارسازی خودرو در تصاویر از [۶۸] Faster Rcnن استفاده شده که برای شناسایی اجسام به کار می‌رود و نیاز به مجموعه بزرگی برای آموزش دارد. این روش فقط برای شناسایی خودرو از دید جلو مناسب است، تصاویر شکل ۲-۲۴ برخی تصاویر به کار رفته در این پژوهش را نشان می‌دهد.



شکل ۲-۲۴: نمونه تصاویر خودروی به کار رفته در [۶۸]

یانگ و همکارانش [۶۹]، مجموعه داده‌ی بزرگی برای شناسایی نوع دقیق خودرو گردآوری کرده‌اند و با استفاده شبکه CNN روشی را پیشنهاد داده‌اند که قادر است خودروها را از نماهای مختلف شناسایی کند. یو و همکاران [۷۰]، از دو شبکه عصبی CNN استفاده کرده‌اند؛ شبکه اول برای آشکارسازی خودرو و شبکه دوم برای شناسایی نوع آن به کار رفته است. این روش دقت خوبی دارد اما مشکل آن این است که اگر خودروها بهم نزدیک شوند، در مواردی چند خودرو را یکی در نظر می‌گیرد. شکل ۲-۲۵ این موضوع را برای این روش نشان می‌دهد.



شکل ۲-۲۵: آشکارسازی محدوده خودروها توسط روش [۷۰]

ژاکوب و همکاران [۷۱]، روشی ارائه کردند که برای شناسایی نیاز به قرار دادن دوربین در بالای خودروها ندارد و می‌تواند از دوربینی که در مکان طبیعی خود قرار گرفته (کنار جاده) استفاده کند. داده‌های مورد استفاده در این روش حدود ۲۱۰۰۰ تصویر است که از زاویه مشابه زاویه دوربین کنار جاده‌ای تهیه شده‌اند. در این روش جعبه سه بعدی با توجه به نقاط محوشدگی تشکیل می‌شود و با توجه به این جعبه، تصویر بازشده^۱ خودرو به دست می‌آید و برای بهبود پاسخ به CNN اعمال می‌شود. روش [۷۲] نیز تا حدودی مشابه روش [۷۱] است. پایگاه داده جمع‌آوری شده در [۷۲] متشکل از ۱۱۶۰۰۰ تصویر است که از ویدئوی دوربین‌های نظارتی به دست آمده است. همین نویسندگان در [۷۳] روشی برای ایجاد یک مجموعه تصویر^۲ از خودروها ارائه کرده‌اند. از هر خودرو در طول مسیر حرکت سه تصویر گرفته شده و با توجه به اطلاعات اندازه به‌طور خودکار در کلاس مخصوص خود قرار گرفته است. البته پس از دسته‌بندی، نتایج زیاد دقیق نیست و خودروهای دسته‌بندی شده باید دستی اصلاح شوند. شکل ۲-۲۶ تصاویر اخذ شده از یک خودرو برای تشکیل مجموعه تصاویر را نمایش می‌دهد.

^۱ Unpack

^۲ Dataset



شکل ۲-۲۶: ثبت تصویر از خودروها به منظور ایجاد مجموعه تصویر [۷۳]

۲-۴-۲- روش‌های مبتنی بر استخراج ویژگی و اعمال به

کلاسه‌بند

شی و همکاران [۷۴]، با توجه به این که تصویر جلوی خودروها از تقارن برخوردار است (مطابق شکل ۲-۲۷)، ابتدا نقاط SIFT برای کل تصویر استخراج می‌کنند، سپس با بررسی نقاط متقارن همه اجسام متقارن موجود در تصویر را شناسایی می‌کنند و با انجام پردازش‌هایی خودرو را از بین این اجسام، جدا می‌کنند. در این روش پس از آشکار شدن خودرو، با پنجره‌گذاری روی تصویر حاصله و استفاده از ویژگی‌های HOG و SURF و اعمال به SVM، نوع دقیق خودرو شناسایی می‌شود. در این روش از ۲۰۴۸ تصویر به منظور آموزش و ۴۰۹۰ تصویر برای آزمایش استفاده شده است.



شکل ۲-۲۷: نمایش تقارن خودرو در [۷۴]

بیگلری و همکاران در [۷۵-۷۷] یک روش مبتنی بر زیر-بخش^۱ ارائه کرده‌اند که سعی می‌کند بخش‌های متمایز کننده را برای هر زیر گروه خودرو پیدا کند. آنها در [۷۵] اشاره کرده‌اند که استفاده از شبکه‌های عمیق CNN نیازمند محاسبات سنگین، حافظه و زمان زیاد است. به همین دلیل از ویژگی HOG برای توصیف هر بخش و از روش‌های مبتنی بر SVM برای کلاس‌بندی استفاده کرده‌اند. در این روش پس از آشکارسازی خودرو در هر تصویر، برای تعیین نوع به الگوریتمی اعمال می‌شود و امتیاز برای هر نوع خودرو محاسبه می‌گردد تا نوع آن تعیین شود. اگر امتیاز از حدی کمتر باشد به‌عنوان خودروی ناشناس تلقی می‌شود. پایگاه داده این روش متشکل از ۵۹۹۱ تصویر خودرو است. مزیت این روش این است که با تعداد کم داده، پاسخ خوبی (دقت ۹۷ درصد) به‌دست آورده است.

مونرو و مدین [۷۸]، از فاصله چراغ‌های جلوی خودرو با مرکز تصویر خودرو به‌عنوان یک بردار ویژگی در شناسایی نوع خودرو استفاده کردند. در [۷۹] هدف طبقه‌بندی خودروها در ۴ گروه، اعم از کامیون، اتوبوس، ون و سواری است. برای طبقه‌بندی خودروها در این روش از شبکه عصبی عمیق و ویژگی SIFT استفاده شده است؛ دقت این روش برای یک پایگاه داده شامل ۶۵۰۰ خودرو، ۹۷٪ گزارش شده است. در [۸۰] برای مقابله با محدودیت نور در شب از ویژگی‌های برجستگی و شکل چراغ‌های عقب خودرو و پلاک برای شناسایی استفاده شده، در این روش نویسندگان از KNN، SVM و درخت تصمیم برای کلاس‌بندی استفاده کرده‌اند. سرفراز و همکارانش [۸۱]، یک روش احتمالاتی مبتنی بر وصله‌بندی ارائه کرده‌اند که به‌طور خودکار مجموعه‌ای از وصله‌ها را برای کلاس‌های خودرو در فاز آموزش یاد می‌گیرد. سپس در فاز آزمایش کلاس خودرو را با توجه به مقدار شباهت وصله‌ها با مدل آموزشی تعیین می‌کند. در [۸۲] هدف شناسایی جسم مشخص و جداسازی آن از پس‌زمینه در یک صحنه است. روش اصلی کار این مقاله استفاده از تصاویر جسم مد نظر در اندازه‌های گوناگون و به‌دست آوردن روابط بین

^۱ Part-Base

این تصاویر به منظور آشکارسازی جسم در اندازه‌های متفاوت است، از این‌رو می‌توان از این روش برای شناسایی خودروها در جاده هم استفاده کرد.

۲-۵- فراتفکیک‌پذیری^۱

یکی از مهم‌ترین مفاهیم در زمینه‌ی پردازش تصویر، تفکیک‌پذیری است. تفکیک‌پذیری به معنای توانایی متمایزسازی جزئیات یک تصویر است. در بیشتر کاربردها از جمله حوزه‌های سنجش از دور، تصویربرداری پزشکی و دوربین‌های نظارتی، تصویر با تفکیک‌پذیری بالا^۲ مورد نیاز است. با افزایش تفکیک‌پذیری جزئیات بیشتری از صحنه تصویربرداری شده در دست است که این افزایش جزئیات می‌تواند باعث بهبود بازشناسی شود. با توجه به این‌که روش پیشنهادی برای کالبراسیون نیازمند شناسایی چند نوع خودرو است و همچنین یکی از کاربردهای روش پیشنهادی تعیین نوع خودروهای عبوری است، در ادامه فراتفکیک‌پذیری به‌عنوان یکی از مسائل تاثیرگذار در شناسایی نوع خودرو مورد بررسی قرار می‌گیرد و چند الگوریتم مشابه الگوریتم پیشنهادی برای فراتفکیک‌پذیری مرور می‌شوند.

برای گرفتن تصویر HR دستگاه تصویربردار نیاز به سخت‌افزارهای تصویری گران قیمت خواهد داشت که عملاً در کاربردهای واقعی مانند دوربین مداربسته و دوربین تلفن همراه، به‌دلیل افزایش هزینه نمی‌توان این سخت‌افزارها را به‌کار گرفت [۸۳ و ۸۴]. از این‌رو ارائه راه‌حلی بدون افزایش هزینه‌ی سخت‌افزار تصویربرداری همواره مورد توجه قرار گرفته است. روشی که می‌توان در این زمینه ارائه نمود استفاده از تکنیک‌های پردازش سیگنال بر روی تصاویر LR^۳ به‌منظور ارتقاء تفکیک‌پذیری این تصاویر است. از ویژگی مهم این روش‌ها عدم نیاز به سخت‌افزار اضافی و به صرفه بودن این روش‌ها از لحاظ هزینه است.

^۱ Super-Resolution

^۲ High-Resolution

^۳ Low-Resolution

روش‌هایی که در این زمینه مورد استفاده قرار می‌گیرند، تحت عنوان روش‌های فراتفکیک‌پذیری ارائه می‌شوند. روش‌های فراتفکیک‌پذیری می‌توانند به سه دسته کلی تقسیم‌بندی شوند: روش‌های مبتنی بر درون‌یابی^۱، روش‌های بازسازی^۲ به کمک چند قاب^۳ ویدئویی و روش‌های مبتنی بر یادگیری^۴. رایج‌ترین روش‌هایی که به‌منظور فراتفکیک‌پذیری مورد استفاده قرار می‌گیرند روش‌های مبتنی بر درون‌یابی هستند. این روش‌ها سرعت بالایی دارند اما قادر به بازسازی جزئیات نیستند و لبه‌های تصویر را هموار می‌کنند.

در روش‌های بازسازی مبتنی بر چند قاب، اطلاعات چندین قاب LR از یک صحنه برای ایجاد یک تصویر SR از همان صحنه استفاده می‌شود. این روش‌ها نیازمند این هستند که تعداد زیادی قاب با جابه‌جایی در حد کسری از پیکسل^۵ از یک صحنه در دسترس باشد که همین موضوع برای این روش‌ها محدودیت‌هایی ایجاد می‌کند [۸۶-۸۵].

در روش‌های مبتنی بر یادگیری سعی می‌شود با یادگیری روابط بین تصاویر LR و HR معادل در مجموعه داده آموزشی^۶، ساختاری برای ساخت تصویر SR ارائه شود. در این روش‌ها معمولاً تصاویر به وصله‌های^۷ هم‌اندازه تقسیم می‌شوند. مجموعه آموزشی می‌تواند از پایگاه‌های تصاویر تهیه شود و یا از خود تصویر ورودی استخراج گردد [۸۷].

روش‌هایی که از پایگاه تصاویر آموزشی برای یادگیری رابطه بین وصله‌های وضوح بالا و پایین استفاده می‌کنند، مبتنی بر ایده مطرح شده در مقاله فریمن [۸۸] هستند. این روش‌ها نیاز به تعداد زیادی جفت وصله تصویر LR و HR در مجموعه آموزشی خود دارند. یانگ و همکاران [۸۹] با به‌کارگیری آموزش

^۱ Interpolation

^۲ Reconstruction

^۳ Multi Frame

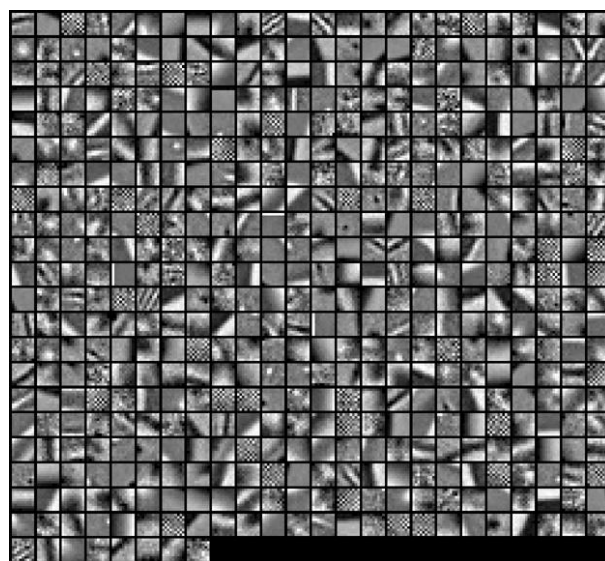
^۴ Learning Based

^۵ Sub Pixel

^۶ Training Dataset

^۷ Patch

واژه‌نامه^۱ و بازنمایی تُنک^۲ بر روی وصله‌های LR و HR به بازسازی وصله‌های SR می‌پردازند. در این روش تصاویر HR پایگاه داده نمونه‌برداری می‌شود تا مجموعه‌ی LR فراهم گردد. سپس با استفاده از وصله‌های تصاویر LR و HR دو واژه‌نامه مربوط به بازنمایی LR و HR ساخته می‌شود. فرضیه اصلی که در این روش در نظر گرفته می‌شود این است که هر یک از وصله‌های تصویر ورودی حداقل به یکی از وصله‌های موجود در مجموعه تصویر نزدیک است. با توجه به کوچک بودن اندازه وصله‌ها و تعداد زیاد تصاویر و تنوع این تصاویر در پایگاه داده این فرض غالباً برقرار است. بنابراین ضرایب تُنک برای وصله‌های تصویر ورودی به دست می‌آید و به واژه‌نامه مربوط به وصله‌های HR اعمال می‌شود و در نهایت وصله‌های SR ساخته می‌شود. تصویر شکل ۲-۲۸ یک نمونه واژه‌نامه HR را در این روش نشان می‌دهد. این واژه‌نامه ۵۱۲ اتم در ابعاد ۹×۹ دارد و برای آن ۱۰۰۰۰۰ وصله از تصاویر به کار رفته است.



شکل ۲-۲۸: نمایش یک نمونه‌ی واژه‌نامه در [۸۹]

در [۹۰] پس از ساخت واژه‌نامه‌های معادل HR/LR از روی وصله‌های تصاویر آموزشی، تصویر ورودی وصله‌بندی می‌شود و نزدیکترین اتم‌ها در واژه‌نامه LR برای هر وصله‌ی تصویر ورودی شناسایی می‌شوند. پس از شناسایی نزدیکترین اتم‌های LR برای هر وصله‌ی تصویر ورودی، اتم‌های معادل در

^۱ Dictionary

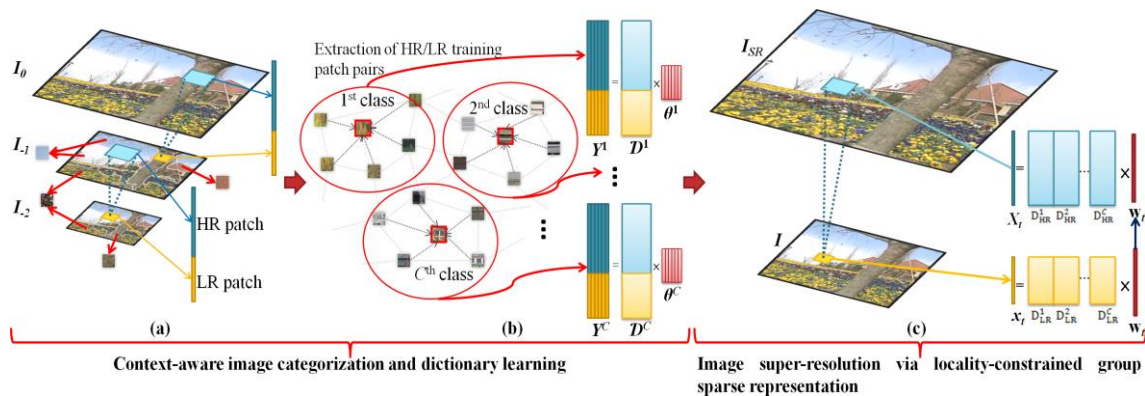
^۲ Sparse Representation

واژه‌نامه HR برای ساخت وصله‌های SR به کار می‌روند و تخمینی اولیه از تصویر SR به دست می‌آید. در این روش به منظور حذف اثر بلاکینگ و بهبود روشنایی وصله‌ها یک تخمین دیگر از تصویر SR با استفاده از روش درون‌یابی دوماکعبی ایجاد می‌شود. نهایتاً با میانگین‌گیری از دو تصویر تخمینی، تصویر SR نهایی ایجاد می‌شود.

زید و همکاران [۹۱] با نمونه‌برداری از تصاویر HR، تصاویر LR مورد نیاز برای تکمیل مجموعه داده خود را ساخته و در ادامه با حذف عناصر فرکانس پایین از مجموعه داده و به کارگیری الگوریتم PCA، واژه‌نامه‌های مورد نیاز برای ساخت تصویر SR را ایجاد کرده‌اند.

با توجه به این که روش‌های مبتنی بر پایگاه تصاویر آموزشی محدودیت‌هایی از جمله نیاز به مجموعه داده بزرگ و پیچیدگی محاسباتی دارند، اخیراً روش‌های مبتنی بر یادگیری رابطه وصله‌های LR و HR از اطلاعات خود تصویر ورودی مورد توجه قرار گرفته است. این روش‌ها بر این واقعیت تکیه می‌کنند که در یک تصویر طبیعی بسیاری از وصله‌ها به یک‌دیگر شبیه‌اند. ایرانی و همکاران [۹۲] این موضوع را به صورت آماری مورد تجزیه و تحلیل قرار داده‌اند و نتیجه گرفته‌اند که در یک تصویر ۹۰ درصد وصله‌ها بیش از ۹ وصله مشابه دارند. بنابراین بدون نیاز به پایگاه داده آموزشی و با استفاده از اطلاعات تصویر ورودی به ساخت تصویر SR پرداخته‌اند.

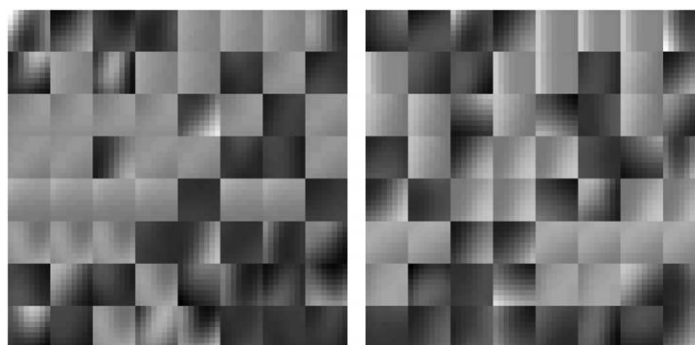
تسای و همکاران [۹۳] برای مشخص کردن وصله‌های مشابه در یک تصویر از ناحیه‌بندی استفاده کرده‌اند و برای هر بخش در تصویر LR و HR، واژه‌نامه‌های HR/LR را ساخته و در نهایت این واژه‌نامه‌ها را برای ساختن تصویر SR به کار برده‌اند. تصویر شکل ۲-۲۹ خلاصه نحوه عمل کرد این روش را در وصله‌بندی، ناحیه‌بندی، ساخت واژه‌نامه و بازسازی تصویر SR نشان می‌دهد. مشکل اصلی این روش پیچیدگی محاسباتی بالا و سرعت پایین آن است.



شکل ۲-۲۹: نمایش نحوه ساخت واژه‌نامه‌ها و نحوه بازسازی تصویر فراتفکیک‌پذیر شده در روش تسای [۹۳]

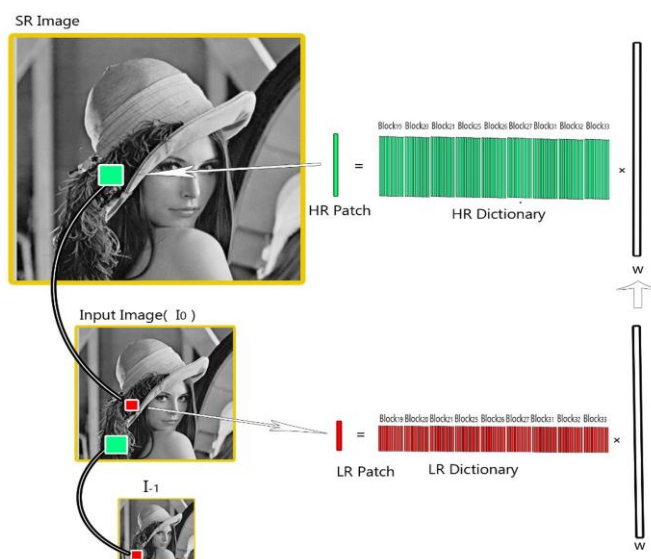
چن و همکاران [۹۴] ابتدا با روش درون‌یابی دومکعبی تصویر ورودی را افزایش ابعاد می‌دهند. سپس تصویر حاصله وصله‌بندی می‌شود و برای هر وصله آن با استفاده از وصله‌های همسایه یک واژه‌نامه HR به‌دست می‌آید. در ادامه اتم‌های این واژه‌نامه با روش درون‌یابی دومکعبی کاهش ابعاد می‌یابد و واژه‌نامه‌ی LR برای وصله معادل در تصویر ورودی ساخته می‌شود. در نهایت ضرایب تُنک برای هر وصله‌ی تصویر ورودی با توجه به واژه‌نامه‌ی آن محاسبه می‌شود و با اعمال این ضرایب به واژه‌نامه‌ی HR معادل، وصله SR معادل ایجاد می‌گردد.

ژو و همکاران [۹۵] روش جدیدی برای فراتفکیک‌پذیری ارائه کرده‌اند. در این روش ابتدا تعدادی از وصله‌های تصویر به‌صورت تصادفی استخراج می‌شود، سپس وصله‌های هموار (وصله‌های فرکانس پایین) کنار گذاشته می‌شود و وصله‌های فرکانس بالا برای ساخت واژه‌نامه‌ها مورد استفاده قرار می‌گیرد. در این الگوریتم از یک روش سریع مبتنی بر KSVD برای ساخت واژه‌نامه استفاده شده است. با توجه به کنار گذاشتن وصله‌های فرکانس پایین و استفاده از یک روش سریع برای ساخت واژه‌نامه، این روش از سرعت بالایی برخوردار است. تصاویر شکل ۲-۳۰ نمونه واژه‌نامه‌های ساخته‌شده برای دو تصویر را در این روش نمایش می‌دهد.



شکل ۲-۳۰: نمایش واژه‌نامه‌های ساخته شده در روش ژو [۹۵]

در [۹۶] روشی سریع برای فراتفکیک‌پذیری مبتنی بر خود یادگیری، توسط مکاری و احمدی‌فرد ارائه شده است. در این روش تصویر ورودی پس از دریافت وصله‌بندی می‌شود و برای هر یک از این وصله‌ها با استفاده از اطلاعات خود وصله و ۸ وصله مجاور آن یک واژه‌نامه تشکیل می‌شود. همچنین با استفاده از روش درون‌یابی دومتکعبی، نسخه کوچک شده‌ی تصویر ورودی به‌دست می‌آید و این تصویر کوچک نیز مشابه تصویر ورودی وصله‌بندی می‌شود و برای هر وصله آن واژه‌نامه ساخته می‌شود. در ادامه با مشخص کردن رابطه بین وصله‌های معادل در هرم تصویر، وصله‌های تصویر SR تخمین زده می‌شود. تصویر شکل ۲-۳۱ ساختار این روش برای بازسازی تصویر خروجی به کمک هرم تصویر و واژه‌نامه‌ها را نشان می‌دهد.



شکل ۲-۳۱: نمایش ارتباط وصله‌ها در هرم تصویر و نمایش نحوه بازسازی هر وصله در [۹۶]

۲-۶- نتیجه‌گیری

در این فصل به بررسی روش‌هایی که با بخشی از ساختار الگوریتم پیشنهادی مرتبط هستند، پرداخته شد. در ادامه روش‌های مطرح شده در هر بخش این فصل را جمع‌بندی می‌کنیم. در جدول ۲-۱ مقایسه بین روش‌های آشکارسازی خودرو آمده است.

جدول ۲-۱: مقایسه بین روش‌های آشکارسازی خودرو

معایب	مزایا	روش
+ سرعت پردازش پایین + عدم دقت بعضی از این روش‌ها در تعیین محدوده خودرو	+ قابلیت تمیز دادن خودروها در هنگام ایجاد هم‌پوشانی	مبتنی بر شبکه‌های عمیق
+ عدم قابلیت تمیز دادن خودروها در صورت نزدیکی به یکدیگر	+ سرعت پردازش بالاتر + عمل‌کرد مناسب بر روی ویدئوهای با دوربین ثابت	مبتنی بر مدل‌سازی پس‌زمینه

شبکه‌های عصبی عمیق به دلیل سرعت پردازش پایین و نیاز به سیستم قوی، برای آشکارسازی خودروها گزینه مناسبی نیستند. در مورد روش‌های مبتنی بر مدل‌سازی پس‌زمینه، آزمایش‌هایی انجام دادیم و مشخص شد روش‌های مبتنی بر مخلوط گوسی برای ویدئوهای جاده‌ای گزینه‌ی مناسبی هستند. در بین روش‌های موجود برای حذف سایه، پژوهش‌های بررسی شده بیان‌گر این هستند که روش‌های مبتنی بر لبه می‌توانند نسبت به سایر روش‌های مبتنی بر ویژگی‌های بافتی و رنگی، سرعت و دقت بهتری داشته باشند. دلیل این موضوع سریع بودن عمل‌گرهای استخراج لبه و پایداری لبه در شرایط نوری گوناگون است.

در جدول ۲-۲ مقایسه‌ای بین روش‌های کالیراسیون خودکار و غیرخودکار کالیراسیون صورت گرفته است. این مقایسه به وضوح بیان‌گر برتری روش‌های خودکار است.

جدول ۲-۲: مقایسه بین روش‌های خودکار و غیر خودکار کالبراسیون

نوع روش کالبراسیون	مزایا	معایب
روش‌های غیر خودکار	+ محاسبات کمتر	+ نیاز به تنظیم دستی متناسب با مکان نصب. + نیاز به تنظیم دستی مجدد در صورت تغییر زاویه دوربین در مواردی همچون وزش باد + خطای بیشتر
روش‌های تمام خودکار	+ عدم نیاز به تنظیم دستی و مجدد در هر مکان + خطای کمتر + قابلیت چند کاربرده بودن	+ محاسبات اولیه بیشتر

جدول ۲-۳ مقایسه بین روش‌های بازشناسی تصاویر خودرو را نمایش می‌دهد. در صورت وجود سیستم مناسب و داده کافی برای آموزش، روش‌های مبتنی بر CNN برتری دارند.

جدول ۲-۳: مقایسه بین روش‌های بازیابی تصویر خودرو

روش	مزایا	معایب
مبتنی بر CNN	+ دقت بیشتر	+ نیاز به سیستم قوی برای پیاده‌سازی + نیاز به داده زیاد جهت آموزش
مبتنی بر استخراج ویژگی و اعمال به کلاسه‌بند	+ لزوماً نیازی به وجود سیستم قوی جهت پیاده‌سازی نیست + امکان آموزش با داده‌های کمتر نسبت به روش CNN	+ نیاز به اجرای آزمایش‌های زیاد جهت انتخاب ویژگی‌ها و کلاسه‌بند مناسب + دقت کمتر نسبت به روش CNN

در جدول ۲-۴ مقایسه بین روش‌های فراتفکیک‌پذیری صورت گرفته است. یکی از کاربردها و اهداف الگوریتم‌های فراتفکیک‌پذیری، بهبود نرخ بازشناسی است. برای این منظور نیاز به بازسازی جزئیات تصویر با دقت بالا است. با توجه به مزایا و معایب مطرح شده در جدول، روش‌های مبتنی بر یادگیری گزینه خوبی برای استفاده در این رساله هستند.

جدول ۲-۴: مقایسه بین روش‌های فراتفکیک‌پذیری

روش	مزایا	معایب
مبتنی بر درونیابی	+ سرعت بالا	+ دقت کم در بازسازی
بازسازی مبتنی بر چند قاب	+ دقت زیاد	+ نیاز به تعدادی قاب با جابه‌جایی در حد کسری از پیکسل + سرعت کم
بازسازی مبتنی بر یادگیری	+ دقت زیاد	+ نیاز به مجموعه آموزشی (قابل تهیه از پایگاه‌های تصاویر و یا استخراج از خود تصویر)

فصل سوم: روش پیشنهادی برای کالیبراسیون تمام خودکار و تخمین ابعاد و

سرعت

۳-۱- مقدمه

خروجی اصلی روش مطرح در این رساله، سرعت و ابعاد خودروهای عبوری است. در این فصل به نحوه تخمین سرعت و ابعاد پرداخته می‌شود. روش پیشنهادی برای تخمین این پارامترها شامل سه فاز است؛ دو فاز گذرا که تنها یک‌بار اجرا می‌شوند و پارامترهای کالیبراسیون دوربین را به دست می‌دهند و یک فاز دائمی که با توجه به نتایج دو فاز قبلی، سرعت و ابعاد خودروها را محاسبه می‌کند.

برخی پارامترهای کالیبراسیون مثل فاصله کانونی و مرکز تصویر، در فاز اول محاسبه می‌شوند، اما برای تکمیل فرایند کالیبراسیون نیاز است ابعاد واقعی برخی اجسام صحنه مشخص باشد. لذا در فاز دوم، چند خودروی رایج (مثل پراید و پژو ۴۰۵) توسط روش‌های بازشناسی شناسایی شده و ضرایب متریک با توجه به ابعاد واقعی این خودروها محاسبه می‌شود. در فاز سوم پس از آشکارسازی خودروها، حذف سایه و حذف پرسپکتیو، ابعاد و سرعت واقعی برای هر خودرو محاسبه می‌شود.

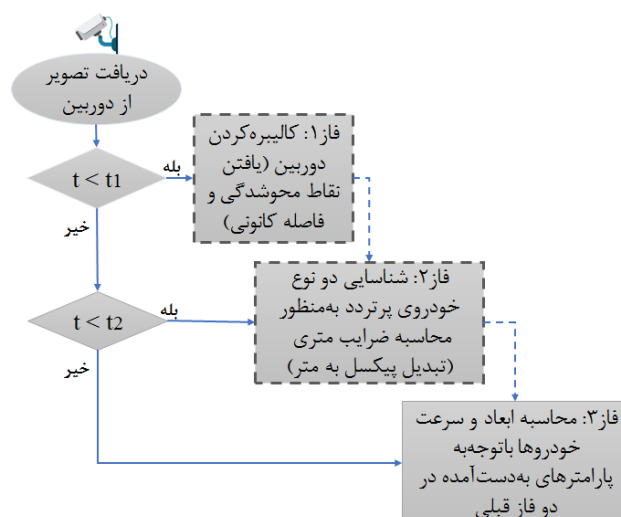
یکی از اهداف اصلی روش پیشنهادی ایجاد مقاومت در برابر سایه است، بدون حذف سایه تخمین پارامترهای مد نظر با خطای زیادی مواجه می‌شود. بلوک حذف سایه از مراحل اصلی رساله است که در این فصل راه کار مواجهه با این چالش بیان می‌شود.

۳-۲- مراحل روش پیشنهادی برای تخمین ابعاد و سرعت

تصویر شکل ۳-۱ مراحل روش برای تخمین ابعاد و سرعت را نمایش می‌دهد. با توجه به این شکل سه فاز پیش‌رو است. فاز اول و دوم در حقیقت نوعی پیش‌پردازش برای فاز سوم هستند. در فاز اول نقاط محوشدگی و فاصله کانونی دوربین با استفاده از مسیر حرکت و اطلاعات لبه‌ای خودروها در حدود ۱ دقیقه اول ویدئوی ورودی تعیین می‌شوند (t_1 تقریباً ۱ دقیقه است)، این پارامترها فقط یک‌بار در ابتدای کار محاسبه می‌شوند.

در فاز دوم، ابتدا محدوده‌ی خودروها با حذف اثرات نامطلوب نویز و سایه مشخص می‌شود و با استفاده از نقاط محوشدگی که در فاز قبل محاسبه شد، جعبه محیطی خودروها به‌دست می‌آید. برای مقابله با پرسپکتیو مختصات گوشه‌های جعبه خودرو روی صفحه فرضی زمین تصویر می‌شود و طول، عرض و ارتفاع خودروها برحسب پیکسل محاسبه می‌گردد. برای بیان این مقادیر برحسب متر، نیاز به ضرایب متریک است. لذا در این فاز با شناسایی چند نماینده برای دو کلاس رایج خودرو (بخش ۳-۲-۵) و مقایسه ابعاد واقعی این خودروها و ابعاد پیکسلی آن‌ها روی صفحه فرضی، ضرایب متریک به‌دست می‌آید. براساس داده‌های مورد آزمایش، فاز دوم حدود ۳ دقیقه طول می‌کشد (اختلاف بین t_2 و t_1 حدود ۳ دقیقه است).

پس از طی شدن فاز اول و دوم، مختصات نقاط محوشدگی و ضرایب متریک مشخص شده و می‌توان ابعاد و سرعت واقعی خودروها را تخمین زد. برای این منظور در فاز سوم، مانند فاز دوم ابتدا خودروها آشکار شده و حذف سایه صورت می‌گیرد. سپس مختصات معادل خودروها بر صفحه جاده به‌دست آمده و ابعاد پیکسلی بدون پرسپکتیو برای هر خودرو در هر قاب مشخص می‌شود. برای محاسبه سرعت خودرو نیاز است که خودرو بین قاب‌های متوالی ردیابی شود، علاوه بر این مقادیر محاسبه شده برای ابعاد یک خودرو در برخی قاب‌ها تفاوت اندکی با یک‌دیگر دارند. لذا برای افزایش دقت در تخمین ابعاد و سرعت هر خودرو، با ردیابی خودروها و با توجه به مقادیر ابعاد و جابه‌جایی خودرو در هر قاب، هیستوگرامی برای طول، عرض، ارتفاع و جابه‌جایی آن تشکیل می‌شود. با تحلیل این هیستوگرام‌ها، ابعاد و جابه‌جایی پیکسلی دقیق‌تر خودرو تعیین شده و با ضرب این مقادیر در ضرایب متریک مربوطه، مقادیر واقعی ابعاد و سرعت برای هر خودرو به‌دست می‌آید.



شکل ۳-۱: مراحل روش پیشنهادی برای تخمین ابعاد و سرعت

۳-۲-۱- کالیبره کردن دوربین با توجه به جاده و خودروها (فاز اول)

در الگوریتم پیشنهادی به منظور تخمین مختصات نقاط محوشدگی (VP_1 , VP_2 و VP_3) و فاصله کانونی (f) روش ارائه شده در [۱] به کار می‌رود. در این روش VP_1 و VP_2 با توجه به جهت حرکت خودروها به دست می‌آید، سپس f از رابطه (۳-۱) محاسبه می‌شود.

$$\begin{cases} VP_1 = (u_x, u_y), VP_2 = (v_x, v_y), P = (P_x, P_y) \\ f = \sqrt{-(VP_1 - P) \cdot (VP_2 - P)} \end{cases} \quad (3-1)$$

در این رابطه P مرکز تصویر و f فاصله کانونی است. با معلوم شدن f ، برای محاسبه سومین نقطه محوشدگی خواهیم داشت :

$$\begin{cases} VP_1' = (u_x, u_y, f), VP_2' = (v_x, v_y, f), \\ P' = (P_x, P_y, 0) \\ VP_3' = (VP_1' - P') \times (VP_2' - P') \end{cases} \quad (3-2)$$

با انتخاب دو مؤلفه اول VP_3' به عنوان مؤلفه‌های VP_3 ، مختصات نقطه محوشدگی سوم مشخص می‌شود. شکل ۳-۲ راستای نقاط محوشدگی را برای دو ویدئو نمایش می‌دهد. در این شکل‌ها محورهای

قرمز راستای نقطه محوشدگی اول، محورهای سبز راستای نقطه محوشدگی دوم و محورهای آبی راستای نقطه محوشدگی سوم را نشان می‌دهند. همچنین مرکز تصویر (نقطه P) با دایره کوچک سبز رنگ مشخص شده است. این پارامترها با استفاده از حدود یک دقیقه اول ویدئوی ورودی محاسبه می‌شوند. همان‌طور که مشخص است علی‌رغم این‌که به‌جز جاده، درختان و حاشیه جاده نیز در تصویر وجود داشته، این روش توانسته با دقت خوبی امتداد نقاط محوشدگی را تخمین بزند. در فصل ۲ بخش ۲-۳- عمل‌کرد این روش مرور شد، در صورت نیاز برای برای آشنایی بیشتر می‌توان به [۹۷] مراجعه نمود.

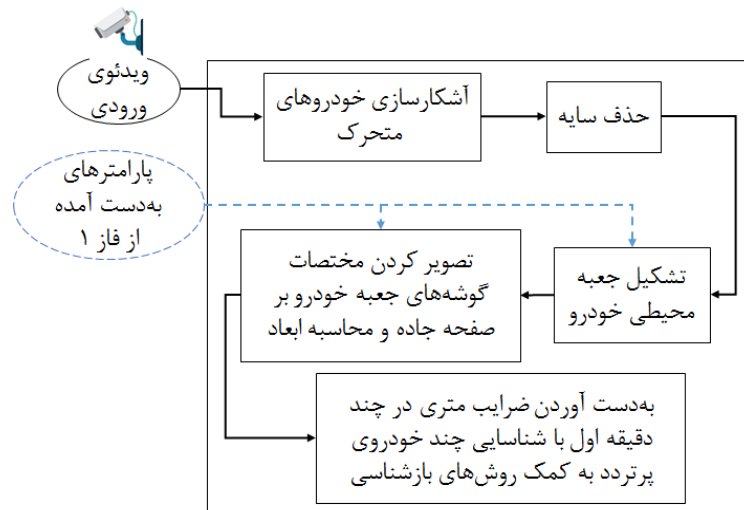


شکل ۲-۳: نمایش راستای نقاط محوشدگی برای دو ویدئو

۲-۲-۳- شناسایی دو نوع خودروی رایج به منظور محاسبه ضرایب

متریک (فاز دوم)

تصویر شکل ۳-۳ مراحل فاز دوم را نمایش می‌دهد. در این فاز هدف این است که پس از آشکارسازی خودروها، حذف سایه، مقابله با پرسپکتیو و همچنین تعیین نوع چند خودرو، ضرایب متریک محاسبه شود. فرایند مربوط به فاز دوم پس از حصول ضرایب متریک پایان می‌یابد.



شکل ۳-۳: مراحل فاز دوم تخمین ابعاد و سرعت

۳-۲-۱- مدل کردن پس‌زمینه و شناسایی پیش‌زمینه

با توجه به متحرک بودن خودروها، می‌توان برای شناسایی محدوده‌ی خودروها از روش‌های آشکارسازی پیش‌زمینه استفاده کرد. تاکنون روش‌های متعددی در این زمینه ارائه شده است. با توجه به انجام آزمایش‌های متعدد روی روش‌های مختلف؛ استفاده از روش ارائه شده در [۹۸] برای این منظور پیشنهاد می‌شود. این روش که یک روش وفقی بهبود یافته براساس مدل ترکیبی گاوسی (IGMM) است، سرعت بالایی دارد و مقاومت خوبی در برابر نیم سایه و نویز از خود نشان می‌دهد، ضمن این‌که در طول زمان به‌روز می‌گردد. تصویر شکل ۳-۴ پاسخ این روش را در برابر نیم سایه نشان می‌دهد.



شکل ۳-۴: شناسایی خودروی متحرک به‌عنوان پیش‌زمینه

۳-۲-۲- حذف سایه

روش IGMM در مقابل نیم سایه مقاومت مناسبی دارد، اما یکی از مشکلات اصلی روش‌های شناسایی پیش‌زمینه، پاسخ نامناسب این روش‌ها در برابر سایه کامل است. تصویر شکل ۳-۵ پاسخ روش IGMM را در برابر سایه کامل نشان می‌دهد، همان‌طور که مشخص است این روش قادر نیست خودرو را به تنهایی شناسایی کند و سایه را نیز همراه خودرو به‌عنوان پیش‌زمینه در نظر می‌گیرد.



شکل ۳-۵: پاسخ روش [۹۸] در انتخاب محدوده در حضور سایه

در صورتی که سایه نیز به عنوان بخشی از خودرو در نظر گرفته شود، الگوریتم در مراحل بعد برای به‌دست آوردن جعبه خودرو و محاسبات ابعاد و سرعت با مشکل مواجه می‌شود. در روش پیشنهادی قصد افزایش مقاومت در برابر سایه کامل است. با توجه به این که حذف سایه بخشی از الگوریتم اصلی است، مراحل حذف سایه باید تا حد امکان زمان‌بر نباشد. در ادامه، مراحل الف و ب، روش پیشنهادی برای حذف سایه را تشریح می‌کنند.

• الف : تعیین محدوده تقریبی هر خودرو با شناسایی لبه‌های آن

تصویر (الف) در شکل ۳-۶ محدوده معادل خودرو در تصویر پس‌زمینه (background)، تصویر (ب) محدوده تشخیص داده شده به‌عنوان پیش‌زمینه (foreground) در قاب ورودی و تصویر (ج) محدوده باینری پیش‌زمینه (تصویر mask) در قاب ورودی را نشان می‌دهد (نتایج حاصل از روش IGMM).



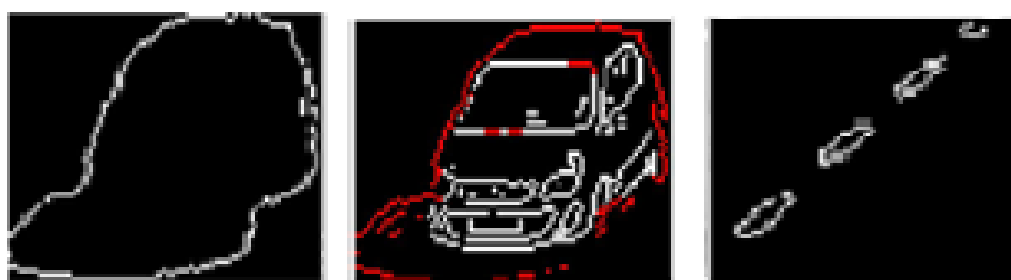
الف: محدوده معادل با
پیش‌زمینه در تصویر پس‌زمینه

ب: محدوده پیش‌زمینه در قاب
ورودی

ج: محدوده باینری خودرو و
سایه آن به‌عنوان پیش‌زمینه

شکل ۳-۶: نمایش پیش‌زمینه و پس‌زمینه برای محدوده‌ای از یک جاده

پس از شناسایی پیش‌زمینه و محدوده معادل آن در پس‌زمینه، نوبت به لبه‌یابی از این تصویر می‌رسد. تصویر حاصل از لبه‌یابی تصویر mask، کانتور (contour) نامیده می‌شود، این تصویر در حقیقت بیان‌گر مرز بین پس‌زمینه و پیش‌زمینه است. همچنین پس از لبه‌یابی از تصویر پس‌زمینه، تصویر به‌دست آمده در تصویر mask ضرب می‌شود و لبه پس‌زمینه (b_edge) نامیده می‌شود. در مورد تصویر پیش‌زمینه نیز تصویر حاصله لبه پیش‌زمینه (f_edge) نام‌گذاری می‌شود. شکل ۳-۷ تصاویر لبه پس‌زمینه، لبه پیش‌زمینه و کانتور را نشان می‌دهد.



الف: b_edge

ب: f_edge

ج: contour

شکل ۳-۷: نمایش لبه‌های پس‌زمینه، پیش‌زمینه و مرز خودرو

دقت در تصویر لبه پیش‌زمینه بیان‌گر این است که این تصویر علاوه بر اطلاعات لبه‌ای خودرو شامل بخش‌هایی از تصاویر کانتور و لبه پس‌زمینه نیز هست؛ این لبه‌ها در تصویر ۳-۷ (ب) با رنگ قرمز نمایش داده شده است. در صورتی که بتوان این اشتراک‌ها را از تصویر لبه پیش‌زمینه حذف کرد، می‌توان

محدوده تقریبی خودرو را بدون اثر سایه به دست آورد. بدین منظور ابتدا تصاویر کانتور و لبه پس زمینه با استفاده از عملیات مورفولوژی به اندازه یک پنجره 2×2 گسترش می یابد. پس از گسترش، برای به دست آوردن محدوده تقریبی خودرو ابتدا تصویر لبه پیش زمینه به منظور از بین بردن نویزهای احتمالی به فیلتر میانه اعمال می شود، سپس مطابق روابط ۳-۳ و ۴-۳ عمل می شود [۳۸]:

$$I(x, y) = \max(f_edge(x, y), b_edge(x, y)) - b_edge(x, y) \quad (3-3)$$

$$car_edge(x, y) = \max(I(x, y), contour(x, y)) - contour(x, y) \quad (4-3)$$

عملگر max در این روابط بیان گر انتخاب بیشترین مقدار برای یک پیکسل است. به عنوان مثال تصویر حاصل از رابطه $\max(f_edge, b_edge)$ تصویری است که در پیکسل هایی که تصویر لبه پس زمینه یا تصویر لبه پیش زمینه برابر با یک هستند، مقدار یک را دارد. بنابراین تصویر I، تصویری است که فقط شامل لبه های پیش زمینه است و لبه های پس زمینه از آن حذف گردیده، و تصویر car_edge (لبه خودرو) تصویری است که فقط شامل بیشتر لبه های خودرو است و لبه های پس زمینه، سایه و کانتور در آن وجود ندارد. تصویر (الف) در شکل ۳-۸ لبه پیش زمینه و تصویر (ب)، تصویر لبه خودرو را نشان می دهد. همان طور که مشخص است در تصویر لبه خودرو تاثیر لبه های سایه و پس زمینه از بین رفته و این تصویر شامل اکثر لبه های خودرو است. بنابراین می تواند برای تعیین محدوده تقریبی خودرو استفاده گردد. با توجه به این که با استفاده از روابط ۳-۳ و ۴-۳ می توان اثر لبه های اضافی بر روی محدوده ی خودرو را از بین برد، علت اصلی گسترش کانتور و لبه پس زمینه این است که اطمینان حاصل شود همه لبه های اضافی از محدوده ی خودرو حذف شوند.



ب: لبه‌های خودرو (car_edge)



الف: پیش‌زمینه (f_edge)

شکل ۳-۸: نمایش لبه‌های پیش‌زمینه و خودرو

تصاویر شکل ۳-۹ محدوده پیش‌زمینه و محدوده به‌دست آمده برای دو خودرو تا این مرحله را نمایش می‌دهد. بررسی تصاویر خروجی برای تعداد زیادی خودرو، بیان‌گر این است که هر چند سایه حذف شده است، اما برای تعداد کمی از خودروها بخش کوچکی از بدنه‌ی خودرو هم حذف شده که این موضوع منجر به اندکی کاهش دقت می‌گردد. در [۹۹] برای افزایش دقت در حذف سایه و به‌دست آوردن محدوده دقیق خودرو، از ویژگی رنگ در کنار ویژگی لبه استفاده شده است. استفاده از رنگ برای شناسایی محدوده‌های متعلق به خودرو و سایه نیازمند استفاده از روش‌های آماری است که سبب افزایش پیچیدگی و کاهش سرعت الگوریتم می‌شود. از این‌رو در مرحله بعد روشی به‌کار می‌رود که بدون محاسبات زیاد، محدوده دقیق‌تری از خودرو به‌دست آید.



شکل ۳-۹: نمایش محدوده پیش‌زمینه (کادر قرمز) و محدوده تقریبی خودرو (کادر سبز)

• ب : شناسایی محدوده دقیق خودرو

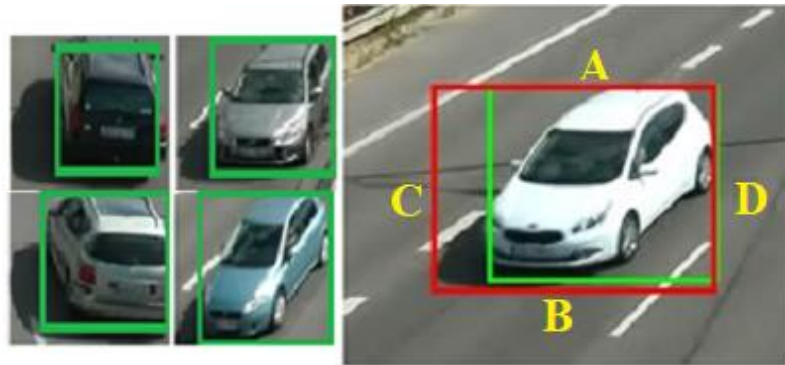
در مرحله قبل با حذف لبه‌های مشترک مرزی و پس‌زمینه از لبه‌های پیش‌زمینه، محدوده تقریبی تعیین شد. لبه‌هایی که طی فرایند مرحله قبل حذف شدند در تصویر (ب) شکل ۳-۷ با رنگ قرمز نمایش داده شده‌اند. با دقت در این تصویر و تصویر (ب) شکل ۳-۸ مشخص است که حذف بخشی از لبه‌های مرزی خودرو عامل کوچک شدن محدوده به‌دست آمده، نسبت به محدوده واقعی است. در صورتی که بتوان تشخیص داد سایه در کدام ناحیه خودرو قرار گرفته، می‌توان از حذف لبه‌های مرزی در نواحی بدون سایه جلوگیری کرد و محدوده دقیق‌تری به‌دست آورد.

محل نصب دوربین‌های جاده‌ای در امتداد حرکت خودروها و در راستای نقطه محوشدگی اول است. با مشاهداتی که روی چند ویدئوی کنار جاده‌ای داشتیم و با توجه به موقعیت دوربین و تردد بالای خودروها در جاده، مشخص شد موقعیت سایه برای تمامی خودروهای درون جاده در یک محدوده زمانی مشخص یکسان است. علاوه بر این سایه نمی‌تواند هم‌زمان جلو و عقب خودرو یا سمت راست و چپ خودرو قرار گیرد. تصاویر شکل ۳-۱۰ چند خودرو و سایه آن را برای یک جاده در یک محدوده زمانی کوتاه نشان می‌دهد. در صورتی که اضلاع کادر در برگیرنده پیش‌زمینه را مطابق این شکل در نظر بگیریم، برای تمامی خودروها ضلع A و D منطبق بر مرز خودرو و ضلع C و B منطبق بر مرز سایه هستند. بنابراین در فرایند حذف سایه و تعیین محدوده دقیق خودرو ضلع A و ضلع D نباید جابه‌جا شوند.



شکل ۳-۱۰: نمایش خودروهای درون جاده برای یک محدوده زمانی؛ ضلع‌های A، B، C و D: اضلاع کادر محدوده‌ی خودرو

تصویر شکل ۳-۱۱ محدوده تقریبی به‌دست آمده برای خودروها در مرحله قبل را نمایش می‌دهد. در صورتی که دقت شود مشخص است که ضلع A و D برای اکثر خودروها جابه‌جا نشده ولی اضلاع B و C برای همه خودروها از مرز سایه جابه‌جا شده و بر مرز خودرو منطبق شده است.



شکل ۳-۱۱: نمایش محدوده تقریبی خودروها

با توجه به مشابه بودن جهت سایه برای همه خودروها در یک زمان، عدم قرارگیری سایه به‌طور هم‌زمان در جلو و عقب یا چپ و راست خودرو، می‌توان با به‌دست آوردن کادر تقریبی برای چند خودرو، محدوده‌ی سایه را به‌دست آورد (جهت سایه را شناسایی کرد) و اضلاعی از کادر که طی فرآیند حذف سایه نباید جابه‌جا شوند را تشخیص داد. در حقیقت اضلاعی از کادر که بیشترین جابه‌جایی را داشته‌اند در سمتی از خودرو بوده‌اند که سایه و پس‌زمینه وجود داشته و با حذف لبه‌های مربوط به پس‌زمینه و مرز سایه از لبه‌های پیش‌زمینه و به‌دست آوردن محدوده تقریبی خودرو، این اضلاع بیشتر از بقیه اضلاع جابه‌جا شده‌اند و سایر اضلاع غالباً جابه‌جایی نداشته و یا خیلی کم جابه‌جا شده‌اند.

بنابراین به این ترتیب عمل می‌شود که با توجه به بخش قبل محدوده تقریبی خودرو به‌دست می‌آید، سپس برای تعیین اضلاع جابه‌جا شونده محدوده تقریبی سایه برای n خودروی قبل از این خودرو در نظر گرفته می‌شود و هر کدام از اضلاع که برای بیشتر از ۷۰ درصد از این n خودرو حداقل ۳ پیکسل جابه‌جایی داشته‌اند، به‌عنوان اضلاع جابه‌جا شونده منظور می‌شوند.

در آزمایش‌ها مشخص شد بررسی محدوده تقریبی سایه برای حدود ۱۰ خودرو قبل به‌منظور شناسایی جهت سایه کافی است. ممکن است شرایطی پیش آید که همه اضلاع مستطیل محیطی خودرو، بدون جابه‌جایی شناخته شوند (مثل اوایل صبح و زمانی که آسمان ابری است و سایه‌ای وجود ندارد)، تحت این شرایط محدوده‌ای که برای خودرو در مرحله (الف) به‌دست آمده به‌عنوان محدوده دقیق خودرو در نظر گرفته می‌شود. شکل ۳-۱۲ محدوده دقیق به‌دست آمده برای چند خودرو را نمایش می‌دهد.



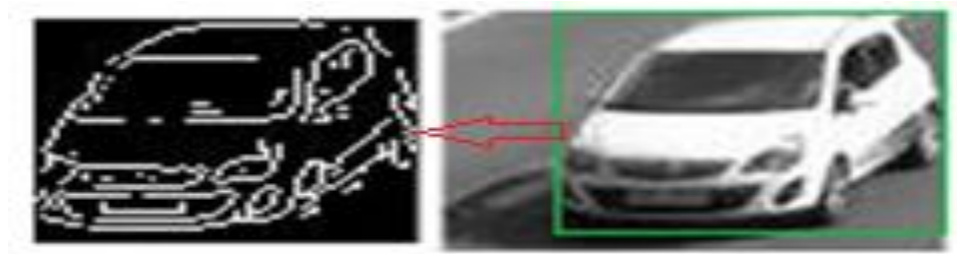
شکل ۳-۱۲: نمایش محدوده دقیق خودروها

هر چند ممکن است طی فرایند حذف سایه کمی خطا وجود داشته باشد، لیکن در روش پیشنهادی هدف اصلی از حذف سایه تشکیل جعبه محیطی خودرو با دقت بالا است. فرایند تشکیل جعبه به‌نحوی است که محدوده‌ای که تا این مرحله برای هر خودرو مشخص شده منجر به ایجاد جعبه‌ی محیطی با دقت کافی می‌شود.

۳-۲-۲-۳- تشکیل جعبه خودرو

برای تشکیل جعبه خودرو نیاز است خطوط مماس برای هر نقطه محوشدگی به‌دست آید. بدین منظور از محدوده دقیق به‌دست آمده برای هر خودرو (محدوده مشخص شده در بخش قبل) در تصویر پیش-زمینه لبه‌یابی می‌شود و لبه‌های مشترک با پس‌زمینه حذف می‌شوند (مانند رابطه ۳-۳). سپس در تصویر حاصل لبه‌های ضعیف‌تر و همچنین لبه‌هایی که جهت‌شان با راستای هیچ‌کدام از نقاط محوشدگی مطابقت خوبی ندارد، حذف می‌شوند. با این‌کار اثر لبه‌های اضافی و لبه‌های مربوط به مرز سایه و پس‌زمینه که ممکن است در محدوده‌ی خودرو قرار گرفته باشند، از بین می‌رود. شکل ۳-۱۳ محدوده

دقیق مشخص شده برای یک خودرو و تصویر لبه‌های درون محدوده بدون لبه‌های مزاحم را نشان می‌دهد.

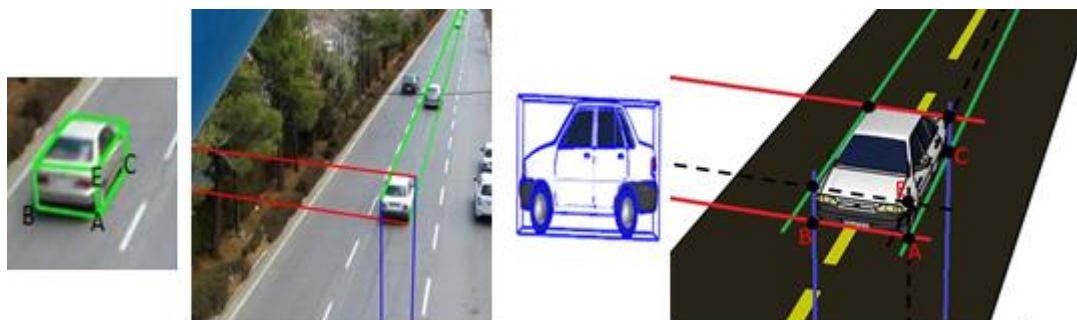


ب: لبه‌های خودرو بدون لبه‌های مزاحم

الف: محدوده دقیق خودرو

شکل ۳-۱۳: نمایش تصویر محدوده خودرو و لبه‌های آن

در ادامه در تصویر موجود از هر پیکسل لبه‌ای خودرو به تک تک نقاط محوشدگی خطی ترسیم می‌شود و با توجه به این خطوط، خطوط مماس و جعبه محیطی برای هر خودرو به دست می‌آید. بسته به نحوه‌ی قرارگیری خودرو در جاده، ترسیم جعبه می‌تواند کمی متفاوت باشد. جعبه باید به گونه‌ای باشد که خودرو را کامل در برگیرد. تصاویر شکل ۳-۱۴ خطوط مماس و جعبه تشکیل شده برای خودروها را نمایش می‌دهند. در این تصاویر خطوط سبز رنگ، مماس‌های ترسیم شده به سمت نقطه‌ی محوشدگی اول، خطوط قرمز رنگ مماس‌های ترسیمی به سمت نقطه دوم و خطوط آبی مماس‌های ترسیم شده به سمت نقطه‌ی سوم هستند. نقاط C, B, A و E مختصات مربوط به برخی از گوشه‌های جعبه خودرو هستند. از فواصل AE ، AB و AC به عنوان ارتفاع، عرض و طول خودرو در محاسبات بعدی استفاده می‌شود و نیاز به محاسبه مختصات سایر گوشه‌های جعبه نیست.



شکل ۳-۱۴: تشکیل جعبه محیطی خودرو

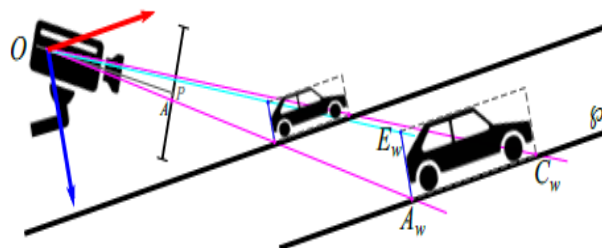
۳-۲-۲-۴- تصویر کردن جعبه خودرو بر صفحه جاده و محاسبه ابعاد آن

پس از به دست آوردن جعبه محیطی و ابعاد آن، مشاهده می شود که در اثر حرکت خودرو در راستای نقطه محوشدگی اول، اندازه خودرو کوچک تر می شود و ابعاد به دست آمده برای خودرو در طول زمان تغییر می کند. این در حالی است که باید ابعاد به دست آمده برای یک خودرو در قاب های مختلف یکسان باشد. برای حل این مشکل از صفحه جاده فرضی که موازی با جاده واقعی است استفاده می شود و با تصویر کردن خودرو بر صفحه جاده ابعاد آن بر حسب پیکسل محاسبه می شود.

صفحه جاده فرضی $\emptyset$ باید موازی با جاده واقعی باشد، بنابراین این صفحه، موازی با محورهای هم راستای نقاط محوشدگی اول و دوم است و می توان بردار نرمال آن را از ضرب خارجی این دو بردار به دست آورد. بردار نرمال در حقیقت موازی راستای نقطه محوشدگی سوم است. برای تصویر کردن روی صفحه جاده، مرکز صفحه تصویر $T = [p_x \ p_y \ f]$ و مکان دوربین $o = [p_x \ p_y \ 0]$ در نظر گرفته می شود (p_x و p_y مختصات مرکز تصویر است). همان طور که گفته شد، صفحه جاده فرضی موازی با جاده اصلی است لیکن فاصله بین شان معلوم نیست، لذا فعلاً صفحه جاده فرضی به صورت دل خواه موازی با راستای نقاط محوشدگی اول و دوم در نظر گرفته می شود. سپس ابعاد به دست آمده با استفاده از ضرایب متریک اصلاح می گردد. تصویر شکل ۳-۱۵ دوربین، صفحه تصویر و دو صفحه جاده دل خواه با فاصله های متفاوت نسبت به دوربین را نمایش می دهد. با در نظر گرفتن صفحه جاده $\emptyset$ ، نقاط A_w ، B_w و C_w با

تصویر کردن نقاط A ، B و C از صفحه تصویر بر روی صفحه جاده $\emptyset$ و نقطه E_w با تصویر کردن نقطه E بر بردار نرمال صفحه جاده (N) به دست می آیند.

$$\begin{cases} A_w = \emptyset \cap \overrightarrow{OA} & C_w = \emptyset \cap \overrightarrow{OC} & B_w = \emptyset \cap \overrightarrow{OB} \\ E_w = N \cap \overrightarrow{OE} \end{cases} \quad (5-3)$$



شکل ۳-۱۵: نمایش دوربین، صفحه تصویر و صفحه جاده دل خواه [۶۰]

پس از تصویر کردن نقاط؛ طول، عرض و ارتفاع خودرو بر روی صفحه جاده مشخص می گردد :

$$l_w = |A_w C_w| \quad w_w = |A_w B_w| \quad h_w = |A_w E_w| \quad (6-3)$$

لازم به ذکر است که این ابعاد، ابعاد واقعی متریک خودرو نیست. با استفاده از رابطه ۳-۶ می توان ابعاد خودرو را در هر قاب برای صفحه دل خواه تخمین زد؛ لیکن در صورتی که مقادیر به دست آمده برای ابعاد یک خودروی خاص باهم مقایسه شود، مشاهده می شود این مقادیر در برخی قاب ها تقریباً یکسان و در برخی دیگر، مقداری باهم تفاوت دارند. دلیل این تفاوت ها عواملی مانند عدم تشخیص صحیح خودرو به عنوان پیش زمینه و هم راستا نبودن دقیق خودرو با راستای نقاط محوشدگی در برخی از قاب ها است. این مسئله باید به نحوی سازمان دهی شود که برای یک خودرو، یک اندازه ثابت و دقیق برای تمامی قاب ها گزارش شود. با بررسی هایی که برای چند خودرو در چندین ویدئو انجام گرفت، مشخص شد در صورتی که مقادیر ابعاد پس از محاسبه در هر قاب، گرد شوند و هیستوگرام طول، عرض و ارتفاع یک خودرو برای قاب هایی که خودرو در دید دوربین قرار دارد به دست آید، بیشینه ی این هیستوگرام ها مربوط به طول، عرض و ارتفاع دقیق تر خودرو است.

برای تشکیل هیستوگرام ابعاد یک خودرو نیاز است خودرو در طول قاب‌هایی که در دید دوربین قرار دارد ردیابی شود. در روش پیشنهادی برای شناسایی مکان خودرو از ردیاب شار نوری [۵۰] استفاده می‌شود و با ردیابی نقطه A (شکل ۳-۱۴) برای هر خودرو، مکان آن در هر قاب تعیین می‌شود. به این ترتیب هرکدام از خودروهایی که در دید دوربین قرار می‌گیرند، ردیابی می‌شوند و با توجه به مقادیر به‌دست آمده برای ابعاد خودرو در طول مسیر حرکتش هیستوگرام‌هایی تشکیل می‌شود و با توجه به بیشینه این هیستوگرام‌ها، طول، عرض و ارتفاع خودرو در صفحه جاده دل‌خواه تعیین می‌شود.

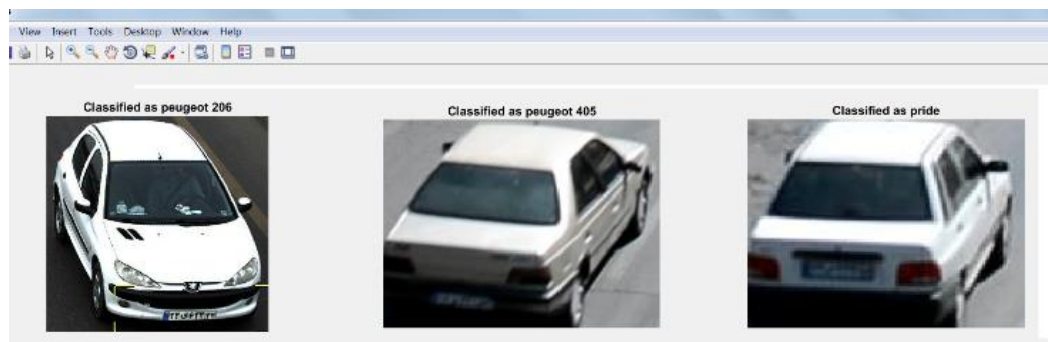
۳-۲-۲-۵- شناسایی خودروهای رایج به منظور به‌دست آوردن ضرایب متریک

همان‌طور که اشاره شد برای حصول ضرایب متریک در چند قاب اول نیاز است چند خودرو شناسایی شود. انتخاب الگوریتم مناسب برای بازشناسی در کاهش خطا موثر است، همچنین در صورتی که یکی از خروجی‌های روش پیشنهادی شناسایی نوع خودرو باشد، این الگوریتم بازشناسی در جایگاه یک معیار در کنار ابعاد برای تعیین نوع دقیق خودرو استفاده می‌شود. بنابراین قدرت روش بازشناسی با اهمیت است. با توجه به اهمیت موضوع، سه روش برای بازشناسی پیشنهاد گردید :

۱. روش اول استفاده از BOW^۱: روش BOW به‌عنوان یکی از روش‌های موفق در بسیاری از الگوریتم‌های بینایی ماشین برای شناسایی الگو مورد تأکید قرار گرفته است. از مزیت‌های این روش این است که با داده آموزشی کم نیز پاسخ مناسبی دارد. ساختار این روش در الگوریتم پیشنهادی به این ترتیب است که ابتدا تصویر هر خودرو پنجره‌گذاری می‌شود، سپس برای هر کدام از این پنجره‌ها ویژگی surf استخراج می‌گردد و با اعمال این ویژگی‌ها به الگوریتم K_Means برای کاهش فضای ویژگی و کاهش طول آن، نهایتاً برداری برای توصیف هر خودرو به‌دست می‌آید. در مرحله آموزش این بردار به الگوریتم SVM برای طبقه‌بندی گروه‌های مد

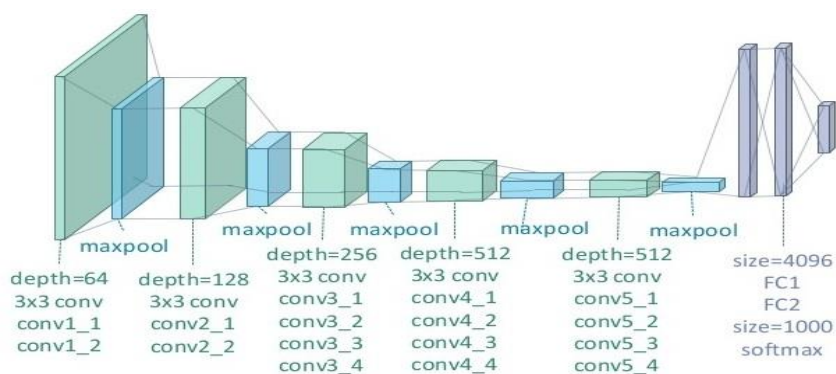
^۱ Bag Of Word

نظر خودرو آموزش داده می‌شود و در مرحله آزمایش نوع خودرو را شناسایی می‌کند. این الگوریتم بسته به میزان شباهت خودرو به هر کدام از کلاس‌ها یک امتیاز اختصاص می‌دهد و کلاسی که بیشترین امتیاز را داشته باشد به عنوان کلاس برنده معرفی می‌کند. توضیحات بیشتر در مورد این روش در پیوست (الف) آمده است. شکل ۳-۱۶، تصاویر اعمالی به این الگوریتم و کلاس شناسایی شده برای هر تصویر در متلب را نشان می‌دهد.



شکل ۳-۱۶: نتایج روش BOW در بازشناسی چند خودرو

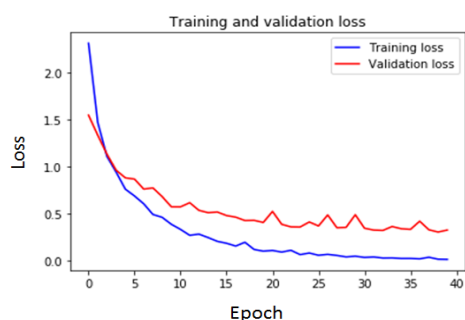
۲. روش دوم استفاده از معماری VGG : شبکه [۱۰۰] VGG یک شبکه عمیق است که توسط سیمونیان و زیسرمان پیشنهاد شد. این شبکه توانسته دقت بالایی برای پایگاه داده ImageNet که متشکل از حدود ۱۵ میلیون تصویر در ۱۰۰۰ کلاس است، به دست آورد. شبکه VGG در دو معماری مختلف با عنوان‌های VGG۱۶ و VGG۱۹ ارائه شده است. ابتدا شبکه VGG۱۶ پیشنهاد شد و بعدها با تغییراتی جزئی در شبکه VGG۱۶، شبکه VGG۱۹ مطرح گردید. شکل ۳-۱۷ معماری شبکه VGG۱۹ را نمایش می‌دهد.



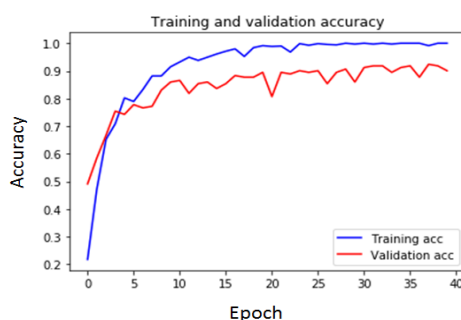
شکل ۳-۱۷: معماری VGG ۱۹

پس از تنظیم دقیق^۱ و آزمایش شبکه VGG۱۹ روی مجموعه تصاویر جمع‌آوری شده، دقت بالای ۹۰ درصد حاصل شد که این رقم بیان‌گر کارایی این شبکه است.

تصویر شکل ۳-۱۸ نمودار accuracy و loss را برای داده‌های اعتبار سنجی و آموزش در تعدادی از تکرارهای فرایند آموزش شبکه نشان می‌دهد. طی فرایند مربوط به این نمودارها شبکه VGG۱۹ برای کلاس‌های {پراید ۱۳۱، پژو ۴۰۵، پژو ۲۰۶، پژو پارس، پراید ۱۳۲، پیکان، ال ۹۰، سمند، پراید ۱۴۱، اتوبوس و کامیون، وانت و سایر} تنظیم دقیق شده است. همان‌طور که مشخص است دقت این روش پس از چند تکرار به حدود ۹۰٪ می‌رسد.



ب: نمودار Loss



الف: نمودار Accuracy

شکل ۳-۱۸: نمودارهای accuracy و loss در طی فرایند آموزش شبکه VGG

تصاویر شکل ۳-۱۹ برخی نتایج شبکه VGG۱۹ را نشان می‌دهد. کلاس پیش‌بینی شده و مقدار confidence برای هر تصویر در بالای آن گزارش شده است. همان‌طور که مشخص است

^۱ Fine Tune

علی‌رغم کیفیت نامناسب برخی تصاویر، این روش توانسته است درست دسته‌بندی کند. این در حالی است که اکثر مقالات ارائه شده در زمینه شناسایی نوع خودرو از تصاویر با کیفیتی که از فاصله نزدیک و نمای جلو یا عقب خودرو به دست آمده، استفاده می‌کنند.



شکل ۳-۱۹: نتایج VGG ۱۹ در شناسایی نوع دقیق خودرو

۳. روش سوم روشی است قوی که مبتنی بر تثبیت تصویر و شبکه‌های عصبی عمیق است. در مورد این روش که یکی از نوآوری‌های رساله است در فصل ۴ به تفصیل بحث می‌شود.

انتخاب بین این سه روش با توجه به دقت و سرعت مد نظر و سیستم مورد پیاده‌سازی صورت می‌گیرد. روش دوم از روش اول کمی قوی‌تر است و سیستمی قوی‌تر نسبت به روش اول نیاز دارد. روش سوم دقت بالاتری دارد و نیازمند سیستمی قوی‌تر نسبت به دو روش دیگر است.

روش‌هایی که تاکنون در زمینه شناسایی خودرو ارائه شده‌اند، یا از روی تصاویر، نوع خودرو را شناسایی می‌کنند و یا در صورت استفاده از ویدئو فقط با تحلیل یک قاب سعی در شناسایی نوع خودروها دارند. به دلیل وجود خطا در تخمین نوع، این امکان هست که نوع گزارش شده برای یک خودرو در قاب‌های مختلف یکسان نباشد. برای حل این مسئله در حین ردیابی خودرو، نوع آن را برای قاب‌هایی که در دید دوربین قرار دارد با الگوریتم بازشناسی، شناسایی کرده؛ آن‌گاه مطابق رابطه ۳-۷ کلاسی که در طول ردیابی بیشتر از سایر کلاس‌ها برنده شده به عنوان کلاس اصلی خودرو در نظر گرفته می‌شود :

$$C = \operatorname{argmax}(\sum_i \text{class}_i) \quad (7-3)$$

در رابطه ۷-۳، i بیان گر قاب‌هایی است که خودرو ردیابی می‌شود و class_i بیان گر بردار کلاس خودرو در هر قاب است. بررسی‌های انجام شده نشان گر این است که در این حالت خطا نسبت به زمانی که نوع خودرو با استفاده از فقط یک قاب شناسایی می‌شود، کاهش می‌یابد.

همان‌طور که تاکنون اشاره شده هدف این است که با شناسایی چند خودروی رایج ضرایب متریک محاسبه شود. برای این منظور کافی است که برای ۲ تا ۳ نوع خودروی رایج چند نماینده شناسایی شود. لذا در این جا به این ترتیب عمل می‌شود که برای هر کدام از انواع خودروی پراید ۱۳۱ و پژو ۴۰۵، n نمونه شناسایی می‌شود ($n > 4$)، سپس با توجه به ابعاد خودروها بر صفحه فرضی و ابعاد واقعی این خودروها، ضرایب متریک به دست می‌آید. با توجه به این که تردد در جاده‌های دارای دوربین جاده‌ای بالاست، شناسایی این تعداد نمونه برای کلاس‌های نام برده زیاد زمان بر نیست؛ مثلاً در ویدئوهای مورد بررسی در این رساله، این زمان حدود ۳ دقیقه است. پس از این که نماینده‌ها برای هر کلاس خودرو شناسایی شد، مطابق رابطه ۸-۳ برای یک ماتریس تشکیل می‌شود :

$$\begin{cases} \text{pride} = \begin{bmatrix} W_{\text{pride } 1} & \cdots & W_{\text{pride } n} \\ H_{\text{pride } 1} & \cdots & H_{\text{pride } n} \\ L_{\text{pride } 1} & \cdots & L_{\text{pride } n} \end{bmatrix} \\ \text{pride} = \begin{bmatrix} W_{405 } 1 & \cdots & W_{405 } n \\ H_{405 } 1 & \cdots & H_{405 } n \\ L_{405 } 1 & \cdots & L_{405 } n \end{bmatrix} \end{cases} \quad (8-3)$$

در این رابطه W ، H و L مقادیر ابعاد بر روی صفحه فرضی هستند. پس از تشکیل ماتریس‌ها، در هر ماتریس مقادیر طول، عرض و ارتفاعی که با سایر نمونه‌ها هم‌خوانی ندارد به عنوان مقادیر پرت حذف می‌شوند و با میانگین‌گیری از مقادیر باقیمانده، ابعاد کلاس مدنظر روی صفحه فرضی جاده مشخص می‌شود. دلیل این که مقادیر پرت در ماتریس‌ها حذف می‌شود این است که از اثر کاندیدهای اشتباه که اقلیت هم هستند بر ابعاد محاسباتی جلوگیری شود.

پس از به دست آوردن ابعاد خودروهای مدنظر روی جاده فرضی، نوبت به محاسبه ضرایب متریک

می‌رسد :

$$\left\{ \begin{array}{l} M_{L-pride} = \frac{L'_{pride}}{L_{pride}} \quad M_{W-pride} = \frac{W'_{pride}}{W_{pride}} \quad M_{H-pride} = \frac{H'_{pride}}{H_{pride}} \\ M_{L-p405} = \frac{L'_{p405}}{L_{p405}} \quad M_{W-p405} = \frac{W'_{p405}}{W_{p405}} \quad M_{H-p405} = \frac{H'_{p405}}{H_{p405}} \\ M_L = \frac{M_{L-pride} + M_{L-p405}}{2} \\ M_H = \frac{M_{H-pride} + M_{H-p405}}{2} \\ M_W = \frac{M_{W-pride} + M_{W-p405}}{2} \end{array} \right. \quad (9-3)$$

در رابطه ۹-۳، L', W', H' مقادیر ابعاد واقعی و H, W, L مقادیری است که با میانگین‌گیری از ماتریس-

ها، پس از حذف مقادیر پرت به دست آمده است. M_W, M_H, M_L ضرایب متریک طول، ارتفاع و عرض هستند؛ با ضرب این ضرایب در ابعاد هر خودرو بر روی صفحه فرضی، ابعاد واقعی متریک خودرو محاسبه می‌شود. لازم به ذکر است روش‌های پیشین، فقط یک ضریب متریک را برای محاسبات ابعاد و سرعت در نظر می‌گیرند، استفاده از ضرایب متریک جداگانه در روش پیشنهادی منجر به کاهش خطا در تخمین کمیت‌های مدنظر می‌شود.

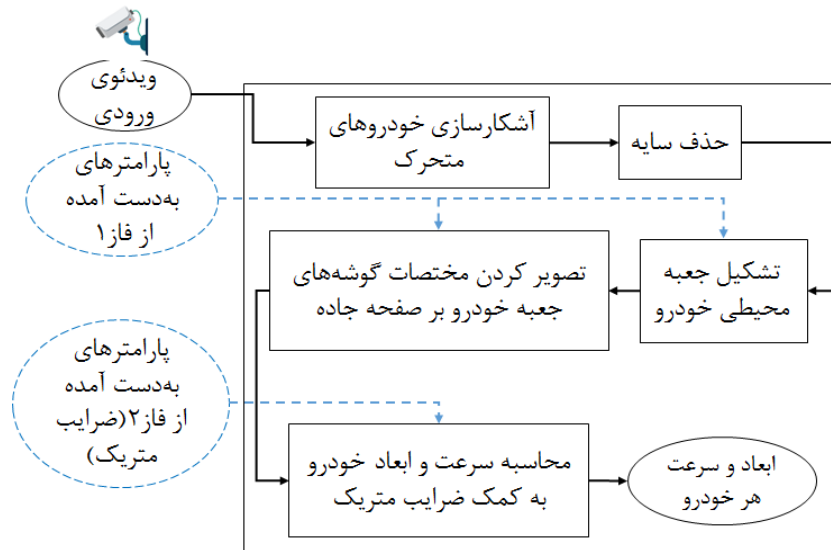
خودروهای پژو و پراید که به عنوان نمونه توسط الگوریتم بازشناسی شناسایی می‌شوند، خودروهایی

هستند که از نظر ابعادی با هم تفاوت دارند؛ تفاوت در ابعاد و پرتدد بودن خودروهای انتخابی منجر به کاهش میانگین خطای الگوریتم می‌شود.

۳-۲-۳- تخمین سرعت و ابعاد واقعی خودرو (فاز سوم)

پارامترهای اولیه مورد نیاز الگوریتم که عبارت‌اند از نقاط محوشدگی و ضرایب متریک در فازهای اول و دوم محاسبه می‌شوند. در فاز سوم از این پارامترها برای محاسبه ابعاد و سرعت خودرو استفاده می‌شود. شکل ۳-۲۰ مراحل این فاز را نشان می‌دهد، با توجه به این شکل و شکل ۳-۳، قسمت‌های شناسایی

خودروهای متحرک، حذف سایه، تشکیل جعبه خودرو و تصویر کردن جعبه خودرو بر صفحه جاده در فاز سوم مشابه همین قسمت‌ها در فاز دوم است، بنابراین از ذکر مجدد آن خودداری می‌شود.



شکل ۳-۲۰: مراحل فاز سوم روش پیشنهادی در تخمین ابعاد و سرعت

برای تخمین سرعت خودرو نیاز است تا جابه‌جایی آن بین دو قاب متوالی به دست آید (Δ_d). با توجه به نرخ قاب بالای دوربین‌ها و این حقیقت که اکثریت خودروها در جاده تغییر سرعت و جهت ناگهانی ندارند، Δ_d بین قاب‌های متوالی تقریباً یکسان است. لیکن برای افزایش دقت در گزارش سرعت و ابعاد خودروها، مانند بخش ۳-۲-۲ با ردیابی خودرو و مقداری گرد کردن این کمیت‌ها در هر قاب، برای ابعاد و Δ_d هیستوگرامی تشکیل می‌شود و نهایتاً با ضرب کردن ضرایب متریک در اندیسی که در آن هیستوگرام بیشینه است، مقادیر واقعی ابعاد و سرعت با دقت خوبی به دست می‌آید. رابطه ۳-۱۰ این فرایند را تشریح می‌کند:

$$\begin{cases} L' = M_L \times \text{argmax}(Hist_L) \\ H' = M_H \times \text{argmax}(Hist_H) \\ W' = M_W \times \text{argmax}(Hist_W) \\ \Delta'_d = M_L \times \text{argmax}(Hist_{\Delta_d}) \end{cases} \quad (۳-۱۰)$$

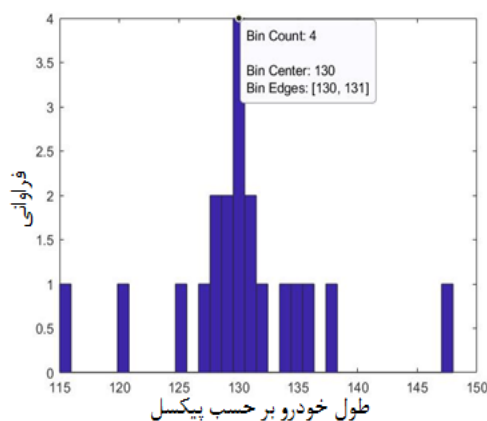
در این رابطه M_L ، M_W و M_H ضرایب متریک، $Hist_L$ ، $Hist_H$ ، $Hist_W$ و $Hist_{\Delta_d}$ هیستوگرام ابعاد و

جابه‌جایی و L' ، H' ، W' و Δ_d' مقادیری هستند که به‌عنوان ابعاد واقعی و جابه‌جایی بین دو قاب در نظر گرفته می‌شوند. از رابطه ۳-۱۱ نیز برای تخمین سرعت بر حسب کیلومتر بر ساعت استفاده می‌شود.

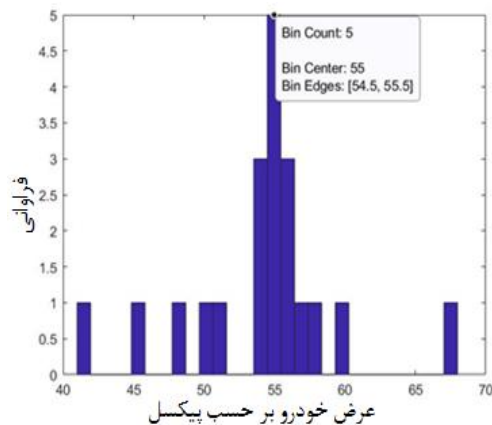
$$\begin{cases} v = \frac{3.6 \times \Delta_d}{\Delta_t} \\ \Delta_t = \frac{1}{FR} \end{cases} \quad (3-11)$$

در این رابطه FR بیان‌گر Frame Rate و Δ_t فاصله زمانی بین دو قاب است. در نهایت، V به‌عنوان سرعت و L' ، H' و W' به‌عنوان ابعاد واقعی هر خودرو گزارش می‌شود.

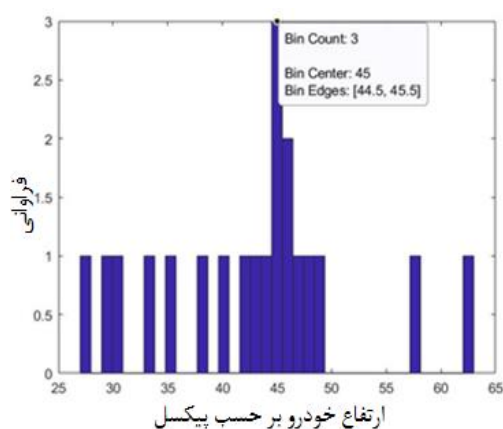
تصاویر شکل ۳-۲۱ نمونه هیستوگرام بردارهای ابعاد و جابه‌جایی یک خودرو که در ۲۰ قاب متوالی ردیابی شده است را نمایش می‌دهند. محور افقی در این نمودارها مربوط به مقدار کمیت بر روی صفحه فرضی و محور عمودی مربوط به تراکم هر مقدار در بردار مربوطه است. همان‌طور که مشخص است مقادیر هر کدام از این نمودارها حول یک عدد بیشترین تجمع را دارند. پرتکرارترین عدد در بردارهای طول، عرض، ارتفاع و جابه‌جایی به ترتیب عبارتند از ۱۳۰، ۵۵، ۴۵ و ۲۳، همچنین ضریب متری برای ویدئوی مورد استفاده که از پایگاه دوسکا است، برابر با ۰/۳۳ به دست آمد. به این ترتیب مقادیر متریک برابر با ۴/۲۹، ۱/۸۱، ۱/۴۸ و ۰/۷۵ حاصل می‌شود. با جای‌گذاری مقدار جابه‌جایی در رابطه ۳-۱۱ مقدار سرعت برابر با ۸۱ کیلومتر بر ساعت می‌شود، همچنین مقدار ground truth سرعت برابر با ۸۲ کیلومتر بر ساعت گزارش شده که مقایسه این دو مقدار حاکی از عمل‌کرد خوب روش پیشنهادی است. نمودارهای ترسیم شده نشان‌گر این است که مقادیر مربوط به ابعاد و سرعت در برخی قاب‌ها تفاوت قابل ملاحظه‌ای با مقدار دقیق آن دارد، لذا در صورتی که هر یک از این مقادیر لحظه‌ای یا میانگین آن‌ها به‌عنوان مقدار نهایی استفاده می‌شد، خطا مقداری افزایش می‌یافت.



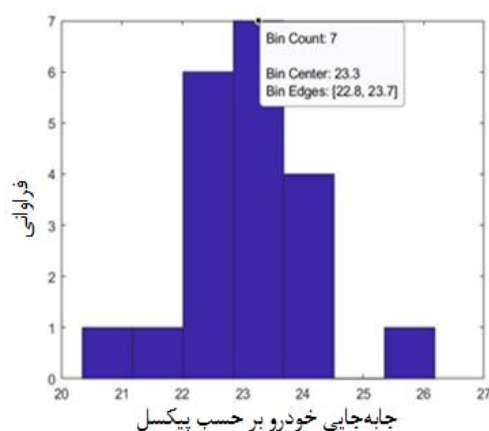
ب: طول خودرو



الف: عرض خودرو



د: ارتفاع خودرو



ج: جابه‌جایی خودرو

شکل ۳-۲۱: هیستوگرام‌های مربوط به بردارهای ابعاد و جابه‌جایی یک خودرو در طول مسیر حرکت

۳-۳- نتیجه‌گیری

در این فصل روش پیشنهادی رساله برای تخمین سرعت و ابعاد خودروهای عبوری که محصول اصلی رساله است، ارائه شد. روش بیان شده شامل سه فاز است که در فازهای اول و دوم پارامترهای مورد نیاز برای کالیبراسیون محاسبه می‌شود. در فاز سوم به کمک این پارامترها سرعت و ابعاد خودروهای عبوری تعیین می‌گردد. تمام خودکار بودن و مقاومت الگوریتم نسبت به سایه، از اهداف محقق شده در روش این فصل است. تخمین ابعاد، می‌تواند قابلیت‌هایی از جمله شناسایی نوع خودرو را فراهم آورد. این مهم، رسیدن به هدف چند کاربرده بودن را هموار می‌کند.

فصل چهارم: ارائه ساختاری نوین برای بازشناسی تصاویر خود و پیشهاد یک

روش جدید برای افزایش تحلیک پذیری به منظور بهبود بازشناسی

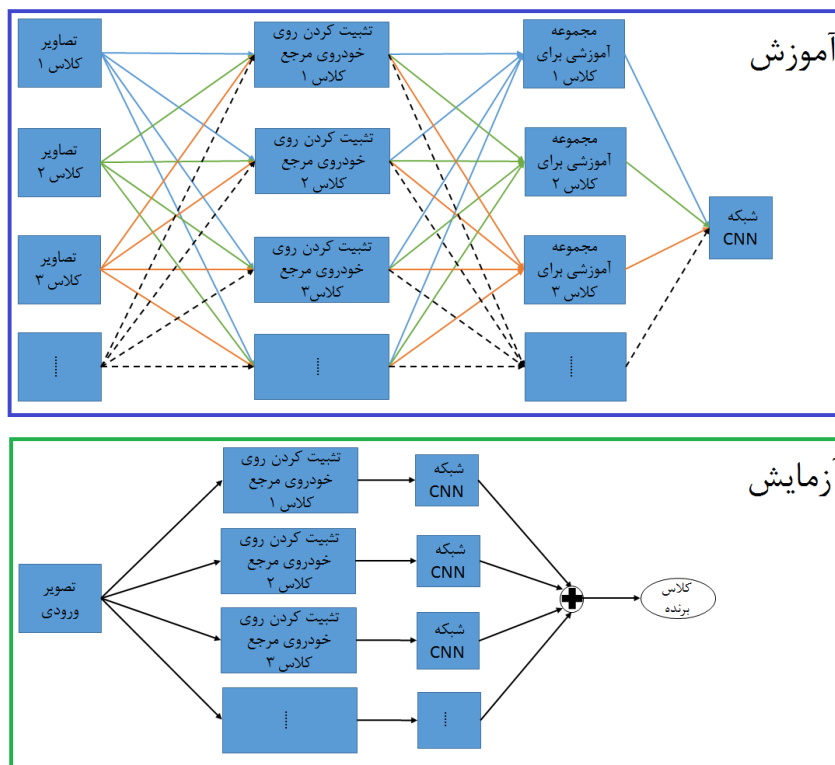
۴-۱- مقدمه

بازشناسی تصاویر خودرو در این رساله از دو جهت کاربرد دارد. جهت اول این که فرایند کالبراسیون خودکارِ روش پیشنهادی که در فصل قبل تشریح شد به کمک شناسایی خودروها تکمیل می‌شود. جهت دوم هم این که اگر شناسایی نوع خودرو به عنوان یک محصول خروجی مد نظر باشد، الگوریتم بازشناسی به عنوان یک معیار در کنار ابعاد برای شناسایی نوع دقیق خودرو استفاده می‌شود. بنابراین استفاده از روشی دقیق برای بازشناسی حائز اهمیت است. در این فصل به کمک شبکه‌های عصبی عمیق و تثبیت تصاویر، ساختاری جدید با فرایند آموزش و آزمایش متفاوت، برای شناسایی نوع خودرو ارائه کرده‌ایم.

یکی از اهداف اصلی الگوریتم‌های فراتفکیک‌پذیری بهبود نرخ بازشناسی است. دوربین‌های جاده‌ای چند متر بالاتر از سطح جاده نصب می‌شوند و ممکن است تصاویری که ضبط می‌کنند کیفیت مناسبی نداشته باشند. بنابراین اگر بتوان تفکیک‌پذیری تصاویر خودرو را افزایش داد، دقت بازشناسی افزایش می‌یابد. از این رو در بخش دوم این فصل روشی برای فراتفکیک‌پذیری مبتنی بر بیان تُنک معرفی کرده‌ایم.

۴-۲- مکانیزم پیشنهادی برای بازشناسی تصویر

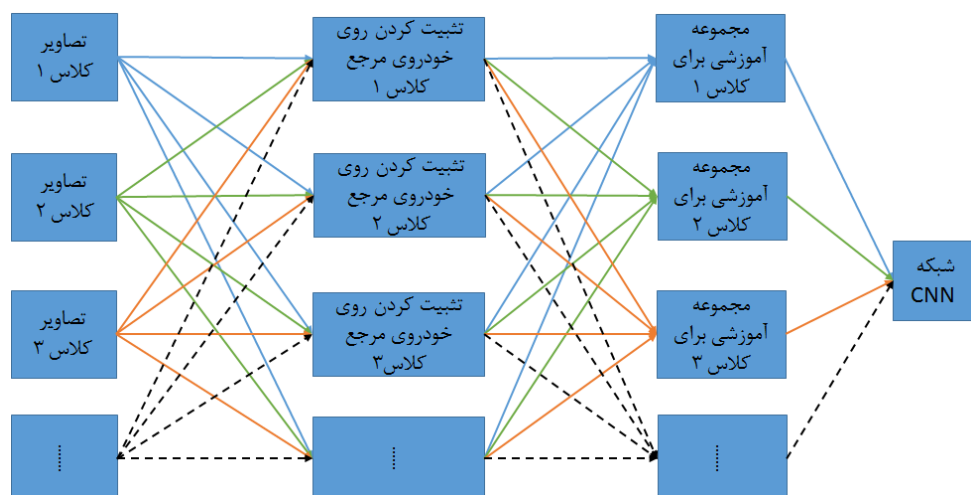
تصویر شکل ۴-۱، روندنمای روش بازشناسی پیشنهادی را نشان می‌دهد. همان طور که مشخص است نحوه آموزش و آزمایش این روش متفاوت از روش مرسوم در آموزش و آزمایش الگوریتم‌های بازشناسی است. در این جا هدف نشان دادن کارایی این مکانیزم است، لذا از یک شبکه CNN ساده (با ۵ لایه کانولوشن، لایه 3×3 pooling و ۶ لایه تماماً) متصل استفاده گردید. در ادامه نحوه آموزش و آزمایش به طور جداگانه بررسی می‌گردد.



شکل ۴-۱: مکانیزم بازشناسی پیشنهادی

۴-۲-۱- مرحله آموزش

تصویر شکل ۴-۲ نحوه آموزش مکانیزم بازشناسی را به طور جداگانه نشان می دهد. با توجه به این شکل، در مرحله آموزش تک-تک تصاویر هر کلاس بر روی تصاویر مرجع برای سایر کلاس ها به طور جداگانه تثبیت می شوند تا مجموعه های آموزشی برای هر کلاس ساخته شود. در این صورت اگر تصاویر اولیه هر کلاس M عدد باشد و تعداد کلاس ها N ، پس از تثبیت کردن، برای هر کلاس تعداد $M \times N$ عدد تصویر آموزشی ساخته می شود. پس از ساخت مجموعه های آموزشی، این مجموعه ها به یک شبکه CNN اعمال می شوند.



شکل ۴-۲: فرایند آموزش شبکه پیشنهادی

انواع مختلف خودرو، اعم از پژو، پراید و سایر گروه‌ها هر چند از نظر ظاهری با هم متفاوتند، اما از منظر استخراج ویژگی و اعمال به کلاسه‌بند شباهت‌هایی هم با یک‌دیگر دارند که این شباهت‌ها منجر به ایجاد خطا هنگام کلاسه‌بندی می‌شود [۱۰۱-۱۰۲]. در مرحله آموزش، به‌عنوان مثال تصاویر کلاس خودروی پراید ۱۳۱ روی تصویر مرجع کلاس‌های پژو ۴۰۵، سمند، پراید ۱۳۲ و دیگر کلاس‌ها تثبیت می‌شود، سپس در مجموعه تصاویر پراید ۱۳۱ قرار می‌گیرد. این کار برای همه کلاس‌ها انجام می‌شود تا کلیه مجموعه‌های آموزشی ساخته شود و در نهایت همه این مجموعه‌ها به شبکه CNN اعمال می‌شوند. تثبیت کردن تصاویر هر کلاس بر روی تصویر مرجع خود کلاس می‌تواند منجر به افزایش دقت کلاس-بندی و تثبیت کردن تصاویر یک کلاس بر روی تصویر مرجع سایر کلاس‌ها می‌تواند منجر به کاهش دقت کلاس‌بندی شود. اما در حقیقت به شبکه عمیق دیکته می‌شود (با توجه به قابلیت بالای این شبکه‌ها) که از ویژگی‌هایی استفاده کند که قادرند تمایز بیشتری بین کلاس‌ها قائل شوند (این موضوع به نحوه وزن‌دهی شبکه طی فرایند آموزش و مباحث فاصله بین کلاسی و درون کلاسی مربوط می‌شود. در پیوست (ب) در مورد شبکه‌های عمیق و نحوه وزن‌دهی آن بحث شده است). البته قطعاً این امر به‌طور کامل برآورده نمی‌شود، اما فرایند آزمایش شبکه، به نحوی متناسب با نحوه آموزش آن، سازمان-دهی شده که خطای الگوریتم کاهش قابل ملاحظه‌ای پیدا کند. تثبیت کردن دارای این مزیت است که منجر به افزایش تعداد داده‌های آموزشی نیز می‌شود. تصاویری که به‌عنوان مرجع برای هر کلاس انتخاب

می‌شوند بهتر است به‌طور واضح و از نمای مستقیم هر خودرو تهیه شده باشند. تصاویر سطر اول شکل ۳-۴ برخی تصاویر مرجع و تصاویر سطر دوم آن تعدادی تصاویر تثبیت شده را نمایش می‌دهند. پیوست (ج) به موضوع تثبیت تصاویر پرداخته است.



(ج) تصویر مرجع (نمونه سوم)



(ب) تصویر مرجع (نمونه دوم)



(الف) تصویر مرجع (نمونه اول)



(و) تصویر ثبت شده (نمونه سوم)



(ه) تصویر ثبت شده (نمونه دوم)



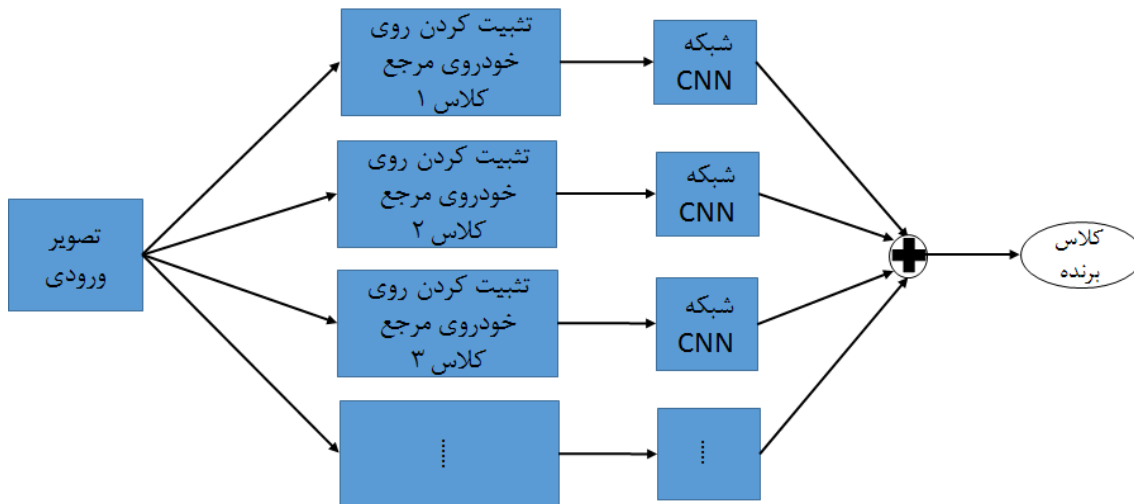
(د) تصویر ثبت شده (نمونه اول)

شکل ۳-۴: نمایش تعدادی خودروی مرجع و تعدادی خودروی تثبیت شده

۴-۲-۲- مرحله آزمایش

نحوه آزمایش روش بازشناسی در شکل ۴-۴ نمایش داده شده است. در این فرایند، ابتدا تصویر ورودی روی تصاویر مرجع تثبیت می‌شود، سپس تصاویر حاصله به شبکه CNN اعمال می‌شوند. خروجی شبکه CNN یک بردار امتیاز عددی است که هر کدام از داریه‌های آن متناظر با یکی از کلاس‌ها است، مجموع اعداد داریه‌های این بردار برابر با عدد ۱ است و داریه‌ی بزرگ‌تر همان کلاس برنده است. در نهایت بردار

خروجی همه این شبکه‌های CNN با یکدیگر جمع می‌شود و با توجه به بردار مجموع، کلاس برنده تعیین می‌شود.



شکل ۴-۴: فرایند آزمایش مکانیزم پیشنهادی

تصویر شکل ۴-۵ یک نمونه تصویر ورودی و نمونه‌های تثبیت شده آن را نمایش می‌دهد. در اینجا به منظور نشان دادن کارایی از ۶ کلاس خودرو که به هم شباهت ظاهری زیادی دارند استفاده شد، این کلاس‌ها عبارتند از پراید ۱۳۱، پراید ۱۳۲، پژو ۲۰۶، پژو ۴۰۵، پژو پارس و سمند.



شکل ۴-۵: تصویر ورودی و نمونه‌های تثبیت شده آن

جدول ۴-۱ امتیاز خروجی شبکه عصبی را برای تصویر ورودی شکل ۴-۵ بدون استفاده از مکانیزم

پیشنهادی و تثبیت کردن نمایش می‌دهد. با توجه به شکل، تصویر ورودی مربوط است به خودروی پراید ۱۳۱، اما با توجه به فاصله و زاویه دوربین و همچنین شباهت بسیار زیاد خودروهای پراید ۱۳۱ و ۱۳۲، شبکه به اشتباه بیشترین امتیاز را به خودروی ۱۳۲ اختصاص داده است.

جدول ۴-۱: تعیین نوع یک خودروی عبوری توسط شبکه CNN

نوع خودرو	پژو ۴۰۵	پژو ۲۰۶	پژو پارس	پراید ۱۳۱	پراید ۱۳۲	سمند
امتیاز	۰	۰	۰	۰/۰۱۳۱	۰/۹۸۶۹	۰

در جدول ۴-۲ نتایج مکانیزم پیشنهادی برای بازشناسی پراید ۱۳۱ شکل ۴-۵ نشان داده شده است. همان‌طور که مشخص است پس از تثبیت کردن و جمع امتیازها، کلاس ۱۳۱ برنده شده است. آزمایش این روش بر روی مجموعه‌ای از تصاویر خودروها، بهبود حدود ۱۰ درصدی را نسبت حالتی که همین شبکه CNN به تنهایی مورد استفاده قرار گرفت (روش عادی آموزش و آزمایش)، ایجاد نمود.

جدول ۴-۲: تعیین نوع یک خودروی عبوری توسط مکانیزم بازشناسی پیشنهادی

	۲۰۶	۴۰۵	پارس	۱۳۱	۱۳۲	سمند
بردار امتیاز حاصل از تثبیت روی پژو ۲۰۶	۰	۰	۰/۰۰۰۵	۰/۶۴۸۹	۰/۳۴۱۷	۰/۰۰۸۹
بردار امتیاز حاصل از تثبیت روی پژو ۴۰۵	۰	۰	۰	۰/۰۰۰۱	۰/۹۹۹۲	۰/۰۰۰۷
بردار امتیاز حاصل از تثبیت روی پژو پارس	۰/۰۰۸۷	۰	۰	۰/۲۲۷۲	۰/۲۶۳۴	۰/۵۰۰۷
بردار امتیاز حاصل از تثبیت روی پراید ۱۳۱	۰	۰	۰/۰۰۳۲	۰/۶۳۹۴	۰/۳۴۰۱	۰/۰۱۷۳
بردار امتیاز حاصل از تثبیت روی پراید ۱۳۲	۰	۰	۰	۰/۷۸۷۷	۰/۲۱۲۳	۰
بردار امتیاز حاصل از تثبیت روی سمند	۰	۰	۰	۰/۹۰۴۴	۰/۰۹۵۶	۰
مجموع امتیازها	۰/۰۰۸۷	۰	۰/۰۰۳۷	۳/۲۰۷۶	۲/۲۵۲۴	۰/۵۲۷۵

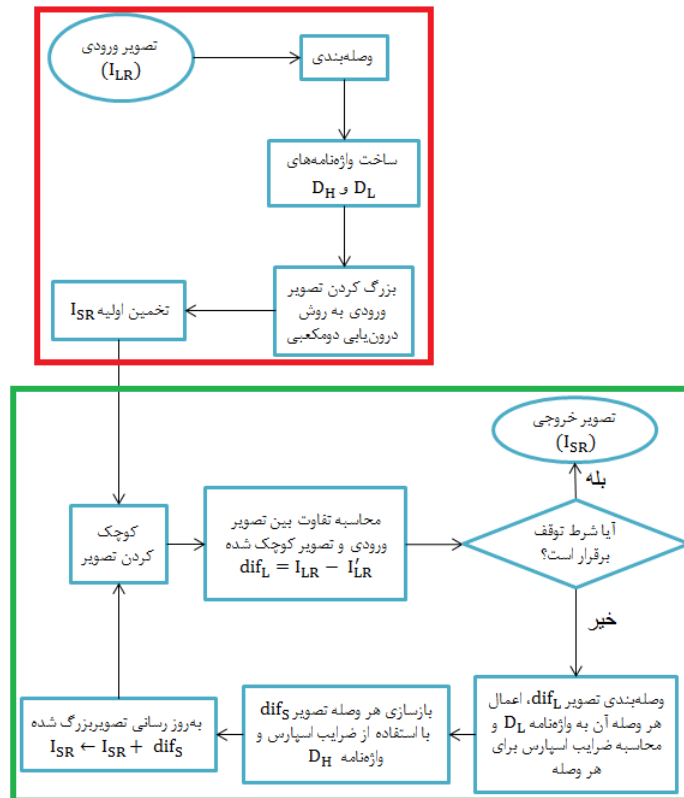
۴-۳- پیشنهاد یک روش فراتفکیک پذیری جهت بهبود

بازشناسی تصویر

همان‌طور که در فصل دوم بخش ۲-۵ بحث شد، افزایش تفکیک‌پذیری تصویر می‌تواند بهبود کارایی روش‌های بازشناسی را به‌دنبال داشته باشد. اعمال روش فراتفکیک‌پذیری پیشنهادی برای افزایش ابعاد تصویر به حدود $1/5$ برابر، منجر به بهبود حدود $1/7\%$ در دقت بازشناسی مکانیزم پیشنهادی گردید (در فصل نتایج بررسی شده است). تصویر شکل ۴-۶ مراحل این الگوریتم را نشان می‌دهد، این روش از دو بخش کلی تشکیل شده است. در بخش اول که با کادر قرمز رنگ در شکل مشخص شده، تصویر ورودی به وصله‌هایی با ابعاد یکسان تقسیم می‌شود. این وصله‌ها به بردارهای ستونی تغییر شکل می‌یابند تا به کمک روش [۱۰۳] KSVD، واژه‌نامه‌ی بازنمایی تُنک تصویر ورودی به‌دست آید. این واژه‌نامه HR تلقی شده و ستون‌های آن اتم‌های واژه‌نامه نامیده می‌شود. هر یک از اتم‌های واژه‌نامه HR با نمونه‌برداری رو به پایین^۱ به اتم متناظر در واژه‌نامه LR تبدیل می‌شود و از این‌جا واژه‌نامه LR ساخته می‌شود. در ادامه تصویر ورودی با استفاده از نمونه‌افزایی و درون‌یابی دومکعبی افزایش مقیاس می‌یابد، به‌این‌صورت یک تخمین اولیه از تصویر SR به‌دست می‌آید.

در بخش دوم که در شکل با کادر سبز رنگ مشخص شده، در یک الگوریتم تکرار و به کمک بازنمایی تُنک حاصل از واژه‌نامه‌های HR/LR، تخمین ارزش پیکسل‌های تصویر SR به‌روز می‌شود؛ الگوریتم در یک حلقه تکرار قرار می‌گیرد و در هر تکرار دقت تصویر تخمین زده شده افزایش می‌یابد.

^۱ down-sampling



شکل ۴-۶: روش پیشنهادی برای افزایش تفکیک پذیری واژه‌نامه

۴-۳-۱- وصله‌بندی تصویر ورودی، ساخت واژه‌نامه و ایجاد یک تخمین

اولیه از تصویر فراتفکیک پذیر شده

در بخش اول روش فراتفکیک‌پذیری پیشنهادی، تصویر ورودی پس از دریافت وصله‌بندی می‌شود، از روی وصله‌ها واژه‌نامه ساخته می‌شود و یک تخمین اولیه از تصویر SR با استفاده از روش درون‌یابی دومکعبی به دست می‌آید.

تصویر ورودی (I_{LR})، پس از دریافت به وصله‌های $m \times m$ تقسیم‌بندی می‌شود و از همه‌ی این وصله‌ها برای آموزش واژه‌نامه‌ی تفکیک‌پذیری بالا (D_H) استفاده می‌شود. برای این منظور ابتدا هر کدام از وصله‌ها به صورت یک بردار ستونی $m^2 \times 1$ تغییر شکل می‌یابد، سپس با در کنار هم قرار دادن این بردارهای ستونی، ماتریس بردار وصله‌های تصویر که $Data_H$ نامیده می‌شود، تشکیل می‌گردد. اگر تصویر دارای k وصله باشد ماتریس $Data_H$ ابعادی معادل $m^2 \times k$ خواهد داشت.

واژه‌نامه‌ی وصله‌های D_H , HR، با در نظر گرفتن تئوری بازنمایی تُنک باید از ماتریس $Data_H$ طوری ساخته شود که هر یک از وصله‌های موجود در ماتریس $Data_H$ (ستون‌های ماتریس) به کمک تعداد معدودی از ستون‌های این واژه‌نامه قابل بازسازی باشد. α_i بردار ضرایب تُنک برای بازسازی i امین وصله در $Data_H$ است. واژه‌نامه D_H دارای Z اتم است و هر اتم یک بردار $1 \times m^2$ می‌باشد، لذا ابعاد این واژه‌نامه $m^2 \times Z$ است که $Z \ll k$ است. یکی از متداول‌ترین روش‌ها برای ساخت واژه‌نامه D_H روش [۱۰۳] KSVD است. این روش مبتنی بر کمینه کردن رابطه (۴-۱) با قید تُنک بودن ضرایب بردار α_i (ها) عمل می‌کند:

$$\begin{cases} \min \|Data_H - D_H \alpha\|_2^2 \\ \text{s.t. } \|\alpha_i\|_0 \leq K, i = 1, 2, \dots, K \end{cases} \quad (۴-۱)$$

همان‌طور که اشاره شد تعداد ستون‌های واژه‌نامه (اتم‌ها) کمتر از تعداد ستون‌های ماتریس $Data_H$ هستند، اما تعداد سطرهای این ماتریس‌ها با هم برابرند؛ تعداد سطرهای این ماتریس‌ها برابر تعداد پیکسل‌های هر وصله‌ی تصویر است.

پس از به دست آوردن واژه‌نامه D_H ، باید واژه‌نامه LR (D_L) از روی آن به دست آید. بدین منظور مشابه روش گزارش شده در [۹۵] عمل می‌شود. به این صورت که ابتدا هر کدام از اتم‌های (بردارهای ستونی) $1 \times m^2$ به صورت وصله‌های $m \times m$ در می‌آید. سپس با روش درون‌یابی دومکعبی به یک ماتریس $n \times n$ ($n < m$) تبدیل می‌شود و نهایتاً این ماتریس $n \times n$ به صورت یک بردار $1 \times n^2$ تغییر شکل می‌دهد، این بردار به عنوان اتم وضوح پایین در ماتریس D_L جایگذاری می‌شود. ابعاد واژه‌نامه D_L برابر با $n^2 \times Z$ خواهد بود، به عبارت دقیق‌تر i امین اتم D_L نسخه وضوح پایین i امین اتم در واژه‌نامه D_H است.

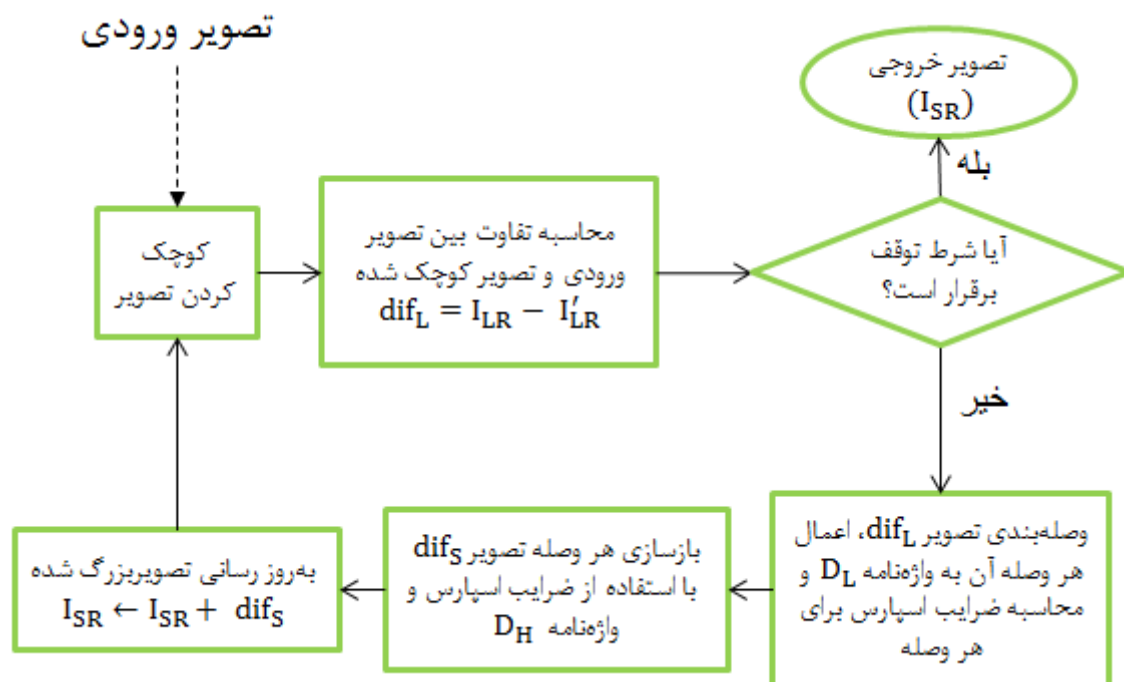
پس از ساخت واژه‌نامه‌ها، برای ایجاد یک تخمین اولیه از تصویر SR، تصویر ورودی با روش درون‌یابی دومکعبی افزایش ابعاد می‌یابد. این روش افزایش نمونه سرعت بسیار بالایی دارد اما به دلیل استفاده از درون‌یابی باعث هموار شدن لبه‌ها و جزئیات تصویر می‌شود. معمولاً این روش به تنهایی برای بزرگ‌نمایی به کار نمی‌رود و به عنوان بخشی از روش در الگوریتم‌های فراتفکیک‌پذیری استفاده می‌شود. در الگوریتم

پیشنهادی با استفاده از روش درونیابی دومکعبی یک تخمین اولیه برای تصویر SR به دست می آید، سپس در مراحل بعدی ارزش پیکسل‌های تصویر به دست آمده در جهت برگرداندن جزئیات به تصویر تغییر می یابد.

۴-۳-۲- پیشنهاد یک روش تکرارپذیر برای ساخت تصویر

فرا تفکیک پذیر شده

در بخش دوم هدف این است که کیفیت تصویر SR، که در مرحله قبل با استفاده از روش درونیابی دومکعبی به دست آمد، با استفاده از واژه‌نامه‌های D_L و D_H بهبود یابد. تصویر شکل ۴-۷ مراحل این فرایند را نشان می دهد.



شکل ۴-۷: مراحل بخش دوم روش پیشنهادی برای فرا تفکیک پذیری

در اجرای تکرار الگوریتم سعی بر این است که جزئیات از بین رفته در تصویر حاصل از افزایش ابعاد درونیابی دومکعبی به تصویر SR برگردد. برای این منظور در هر تکرار از الگوریتم، آخرین نتیجه تصویر SR با تصویر ورودی مقایسه می شود. در بار اول اجرای الگوریتم تخمین اولیه ی تصویر SR از روش

درون‌یابی دومکعبی جایگزین می‌شود، در تکرارهای بعد نسخه جدید SR در الگوریتم از روی نسخه‌ی قبلی ساخته می‌شود.

چون ابعاد تصویر ورودی با تصویر SR به‌دست آمده برابر نیست، تصویر SR ایجاد شده در تکرار قبل با استفاده از روش کاهش نمونه‌ها و درون‌یابی دومکعبی به ابعاد تصویر ورودی کوچک می‌شود، این تصویر با نماد I'_{LR} نمایش داده می‌شود.

تفاوت بین تصویر ورودی I_{LR} و I'_{LR} با dif_L نشان داده می‌شود :

$$dif_L = I_{LR} - I'_{LR} \quad (۲-۴)$$

در این رابطه، dif_L یک ماتریس هم اندازه I_{LR} و I'_{LR} است. هرچه تصویر فراتفکیک‌پذیری شده I_{SR} دقیق‌تر باشد، تفاوت بین دو تصویر I_{LR} و I'_{LR} کمتر خواهد بود. بنابراین مراحل بعدی الگوریتم باید به گونه‌ای سازمان‌دهی شود که اندازه عناصر ماتریس dif_L به صفر نزدیک‌تر شود. رابطه ۳-۴ به‌عنوان یک معیار ساده و مناسب برای بررسی اختلاف بازسازی و شرط توقف می‌تواند استفاده شود :

$$E = \sqrt{\frac{1}{MN} \sum_{i=0}^M \sum_{j=0}^N (dif_L(i, j))^2} < \beta \quad (۳-۴)$$

در این رابطه M و N بیان‌گر تعداد سطر و ستون‌های ماتریس تفاضل dif_L هستند. در صورتی که مقدار E عدد کوچکی باشد I_{SR} دارای دقت مناسبی است و الگوریتم می‌تواند متوقف شود و در غیر این‌صورت کار باید ادامه پیدا کند. با توجه به آزمایش‌های انجام گرفته مقدار مناسب برای β عددی حدود 10^5 به‌دست آمد و مشخص شد که الگوریتم پس از حدود ۲ تا ۳ تکرار به حدی کمتر از این مقدار خطا می‌رسد.

در صورت برقرار نبودن شرط ۳-۴، معادل SR ماتریس dif_L به کمک واژه‌نامه‌های آموزش داده شده از وصله‌های تصویر به‌دست می‌آید و به نتیجه تصویر تخمین زده شده اضافه می‌گردد. با توجه به این‌که تصویر dif_L هم اندازه تصویر ورودی I_{LR} است، به وصله‌های $n \times n$ که با هم دارای هم‌پوشانی هستند تقسیم‌بندی می‌شود. پس از این‌که وصله‌های $n \times n$ از dif_L استخراج شد، هر وصله با روبش

سطرها به صورت یک بردارستونی $n^2 \times 1$ درمی آید و با استفاده از روش OMP ضرایب تُنک برای هر بردار از روی واژهنامه‌ی D_L تعیین می‌شود.

وظیفه روش OMP کمینه کردن رابطه ۴-۴ است، دقت شود که در این رابطه واژهنامه موجود است و فقط باید ضرایب تُنک به دست آید؛ ولی در رابطه ۱-۴ هدف از حل معادله به روش KSVD، به دست آوردن واژهنامه است. الگوریتم OMP در حقیقت با ترکیب ستون‌های واژهنامه D_L سعی در بازسازی دقیق هر وصله دارد، جزئیات مربوط به روش OMP در [۱۰۳] بیان شده است.

$$\begin{cases} \min \|X - D_L \lambda\|_2^2 \\ \text{s.t. } \|\lambda\|_0 \leq K \end{cases} \quad (4-4)$$

در رابطه ۴-۴، X بردار ستونی شده یکی از وصله‌های ورودی و λ بردار ضرایب تُنک است که باید برای هر وصله به طور جداگانه به دست آید. نرم صفر λ باید کوچکتر از K باشد به این معنی که تعداد اتم‌هایی که در بازسازی وصله‌ها استفاده می‌شوند کمتر از K باشد.

همان‌طور که در بخش ۴-۳-۱ بیان شد اتم‌های واژهنامه D_L و D_H متناظر با یکدیگر هستند، بنابراین از این واژهنامه‌ها می‌توان برای تبدیل وصله‌های تصویر کوچک به بزرگ و بالعکس بهره برد. به عنوان مثال اگر وصله با اندازه $n \times n$ موجود باشد، می‌توان ضریب تُنک را برای آن با استفاده از واژهنامه‌ی D_L به دست آورد و با استفاده از واژهنامه‌ی D_H ، وصله با اندازه $m \times m$ که به صورت تقریبی بزرگ شده وصله‌ی با اندازه $n \times n$ است را ساخت. تصویر dif_S بزرگ شده تصویر dif_L با استفاده از واژهنامه‌ها است، بنابراین برای به دست آوردن این تصویر در هر تکرار الگوریتم کافی است ضرایب تُنک برای هر وصله $LR(patch_{LR})$ که در رابطه ۴-۴ به دست آمد، برای بازسازی وصله‌ی $SR(patch_{SR})$ به کار رود. رابطه ۴-۵ نحوه ساخت یک وصله در تصویر dif_S با استفاده از ضرایب تُنک را نشان می‌دهد :

$$patch_{SR} = D_H \times \lambda \quad (5-4)$$

پس به طور خلاصه برای تشکیل dif_S به این صورت عمل شد که :

۱. تصویر dif_L به وصله‌های $n \times n$ تقسیم شد، برای هر وصله بطور جداگانه بردار تُنک با استفاده از

واژه‌نامه‌ی D_L به‌دست آمد (مطابق رابطه ۴-۴).

۲. بردار تُنک به‌دست آمده برای هر وصله‌ی تصویر dif_L به واژه‌نامه D_H اعمال شد تا وصله معادل

در تصویر dif_S به‌دست آید (مطابق رابطه ۴-۵).

پس از به‌دست آوردن dif_S برای به‌روز رسانی I_{SR} خواهیم داشت :

$$I_{SR} \leftarrow I_{SR} + dif_S \quad (۴-۶)$$

مراحل ساخت dif_L ، dif_S و به‌روز رسانی I_{SR} تا برقراری شرط توقف ادامه می‌یابد، پس از برقراری

شرط توقف می‌توان تصویر I_{SR} را به‌عنوان نمونه فراتفکیک‌پذیرشده‌ی تصویر ورودی درنظر گرفت. با

توجه به این‌که در این‌جا هدف بهبود نرخ بازشناسی خودرو به کمک فراتفکیک‌پذیری است، همه تصاویر

خودرو اعم از تصاویر آموزش و آزمایش باید به الگوریتم فراتفکیک‌پذیر اعمال شوند. شکل ۴-۸ یک

تصویر ورودی و بزرگ‌شده‌ی آن توسط الگوریتم فراتفکیک‌پذیر پیشنهادی را نمایش می‌دهد.



ب: تصویر فراتفکیک‌پذیر شده



الف: تصویر ورودی

شکل ۴-۸: نمایش تصویر یک خودرو و بزرگ‌شده‌ی آن توسط روش فراتفکیک‌پذیری پیشنهادی

۴-۴- نتیجه‌گیری

با توجه به اهمیت بازشناسی نوع خودرو در ساختار روش‌های مطرح در رساله، در این فصل دو الگوریتم در رابطه با بازشناسی تصاویر مورد بررسی قرار گرفت. در الگوریتم اول، ساختار جدیدی برای شناسایی نوع خودرو مطرح شد. در این ساختار که از خواص شبکه‌های عمیق و تثبیت کردن تصویر بهره می‌برد، نحوه آموزش و آزمایش به گونه‌ای سازمان‌دهی شده که منجر به کاهش هر چه بیشتر خطا می‌شود. در الگوریتم دوم، با توجه به تاثیر مثبت افزایش تفکیک‌پذیری بر نرخ بازشناسی، الگوریتمی در زمینه فراتفکیک‌پذیری پیشنهاد شد. در این الگوریتم با استفاده از روش‌های مبتنی بر بیان تُنک، تفکیک‌پذیری تصویر به کمک یک روش تکرارپذیر بهبود می‌یابد.

فصل پنجم: شناسایی نوع خودروی عبوری به کمک ابعاد و الگوریتم های بازشناسی

۵-۱- مقدمه

همان‌طور که در فصل اول اشاره شده، یکی از خروجی‌های روش پیشنهادی رساله، نوع خودروی عبوری است. در این فصل هدف اصلی ارائه الگوریتمی ترکیبی به‌منظور شناسایی نوع خودروهای پرتدد از روی ویدئوی دریافتی از دوربین نظارتی کنار جاده‌ای است، به نحوی که مسیر برای دستیابی به یک الگوریتم عملی در این زمینه هموارتر شود. تفاوت اصلی این روش با سایر روش‌هایی که تاکنون در این زمینه ارائه شده‌اند این است که در این‌جا برای تعیین نوع خودروی پرتدد از دو معیار کاملاً مجزا استفاده می‌شود:

۱. معیار اول استفاده از روش‌هایی مانند [۶۲، ۷۵] و یا سه روش معرفی شده در بخش ۳-۲-۲-۵، برای سنجش میزان شباهت ظاهری است. مجموعه تصاویری که برای آموزش الگوریتم مربوط به این معیار استفاده می‌شود تصاویری است که با فیلم‌برداری از جاده‌ها از ارتفاع بالا به‌دست آمده است. هرچند افزایش تعداد داده‌های آموزشی می‌تواند منجر به بهبود عمل‌کرد این روش‌ها شود، اما به‌هرحال عواملی مانند وزش باد، تغییرات نوری محیط، کوچک بودن تصویر خودرو، شباهت ظاهری خودروها، تغییرات دستی ایجاد شده در ظاهر خودرو و نویز محیط در کاهش دقت این روش‌ها تاثیر دارد، لذا از معیار دوم به‌منظور کاهش خطا استفاده می‌شود.

۲. معیار دوم حذف پرسپکتیو و به‌دست آوردن ابعاد خودروهای عبوری و تعیین نوع خودرو به کمک ابعاد آن است. برای این منظور ابتدا در چند دقیقه به کمک معیار اول ابعاد کلاس‌های رایج خودرو بر روی صفحه فرضی جاده مشخص می‌شود، سپس این ابعاد برای تعیین نوع خودرو در قاب‌های آتی ویدئو استفاده می‌شود.

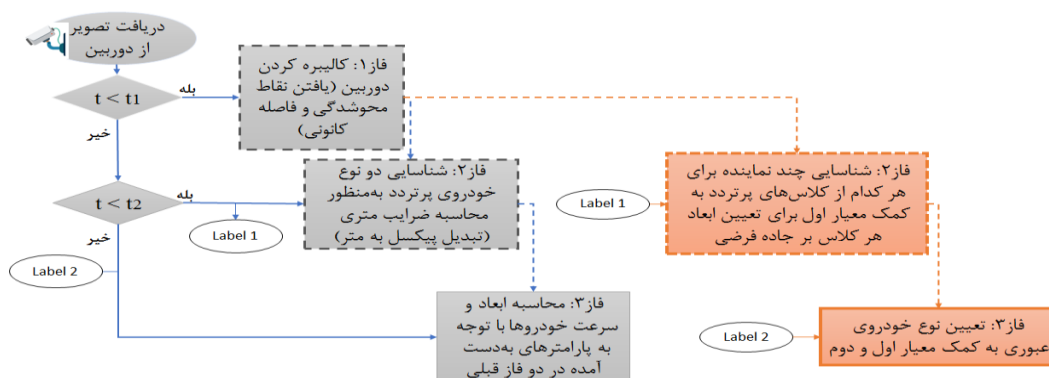
در این روش معیار دوم مکمل معیار اول است و خطای ناشی از معیار اول را کاهش می‌دهد. لذا این موضوع قدرت انتخاب را برای معیار اول تا حدی افزایش می‌دهد و لزوماً نیازی نیست که از روش‌هایی

پیچیده که به حافظه و قدرت بالا، زمان آموزش و مجموعه داده زیاد نیاز دارند استفاده شود. لذا روش بازشناسی ارائه شده در فصل ۴ و روش‌های دیگر بررسی شده در بخش ۳-۲-۲-۵، می‌تواند به عنوان معیار اول استفاده شود، سپس با معیار دوم دقت افزایش یابد.

تصویر شکل ۵-۱ روندنمای نحوه تعیین نوع خودروی عبوری در کنار نحوه تخمین ابعاد و سرعت (فصل سوم) نشان می‌دهد. مراحل مربوط به تعیین نوع خودروی عبوری با رنگ نارنجی نمایش داده شده است. تعیین نوع خودروی عبوری نیز شامل سه فاز است که فاز اول با فاز اول تخمین ابعاد و سرعت یکسان است و فازهای دوم و سوم آن به موازات فازهای تخمین ابعاد و سرعت صورت می‌گیرند. فاز دوم در تعیین نوع، مشابه فاز دوم در تخمین سرعت و ابعاد است با این تفاوت که در تعیین نوع باید برای تمامی کلاس‌های مد نظر چندین نماینده شناسایی شود و پس از حذف مقادیر پرت، ابعاد هر کلاس بر روی صفحه فرضی مشخص گردد. البته ممکن است بسته به تردد خودروها در جاده، زمان t_2 که زمان شناسایی چند نماینده برای انواع خودرو است، مقداری افزایش یابد. همچنین این‌جا لزومی به محاسبه ضرایب متریک نیست، هر چند می‌توان به کمک ابعاد متریک نیز نوع خودرو را شناسایی کرد. تنها مرحله‌ی منحصر به فرد در تعیین نوع خودرو، فاز سوم است. بنابراین با توجه به اشتراکات مباحث این فصل و فصل سوم، در این‌جا فقط در مورد مراحل فاز سوم بحث می‌شود. این مراحل عبارتند از :

۱. بررسی شباهت خودروها با مقایسه ابعاد (معیار دوم)، و ۲. تعیین نوع خودروی عبوری به کمک معیار

اول و دوم



شکل ۵-۱: روندنمای مربوط به نحوه تعیین نوع خودروی عبوری در کنار تخمین ابعاد و سرعت

۵-۲- بررسی مقدار شباهت خودروی عبوری به هر کلاس با

مقایسه ابعاد (معیار دوم)

هرچند با استفاده از روش‌های قوی برای معیار اول دقت بازشناسی افزایش می‌یابد، اما به‌هرحال هر کدام از روش‌های بازشناسی در مواردی قادر نیست نوع خودرو را درست گزارش کند. دلیل این موضوع شباهت ظاهری خودروها به یکدیگر است، برای مثال شباهت در جلو پنجره و چراغ‌های جلو. تصویر شکل ۵-۲ نتایج روش BOW (در صورتی که معیار اول باشد) را برای برخی موارد که این روش قادر نبوده نوع خودرو را درست تشخیص دهد نشان می‌دهد.



شکل ۵-۲: برخی تشخیص‌های اشتباه توسط روش BOW

بررسی جامع خطاهای معیار اول بیان‌گر این است که در مواردی کلاس تشخیص داده شده برای یک خودرو از لحاظ ابعاد و بدنه مشابه با کلاس واقعی خودرو است (مثل شباهت پراید ۱۳۱ و پراید ۱۳۲)، ولی در مواردی کلاس برنده با کلاس واقعی کاملاً غیر همسان است (مثل خودروی پراید شکل ۵-۲ که پژو ۴۰۵ تشخیص داده شده). برای همین برای افزایش دقت در گزارش نوع خودرو از معیار دوم در کنار معیار اول استفاده می‌شود.

پس از حذف پرسپکتیو و محاسبه ابعاد خودروی عبوری، نوبت به مقایسه ابعاد این خودرو با ابعاد مربوط به خودروهای پرتدد می‌رسد (مطابق شکل ۵-۱، ابعاد این خودروها بر صفحه جاده در فاز دوم مشخص شده است). لیکن به هر حال مقادیر به دست آمده برای هر خودرو، حول مقادیر اصلی آن کمی نوسان دارد. بنابراین برای بررسی وابستگی یک خودرو به یک کلاس خواهیم داشت :

$$\begin{cases} L_m - K \times L_m < L_w < L_m + K \times L_m \\ W_m - K \times W_m < W_w < W_m + K \times W_m \\ H_m - K \times H_m < H_w < H_m + K \times H_m \end{cases} \quad (۵-۱)$$

در این رابطه L_m, H_m, W_m مقادیر ابعاد مربوط به هر کلاس و L_w, H_w, W_w مقادیر مربوط به خودروی عبوری هستند. در صورتی که K عدد بزرگی انتخاب شود، کلاس‌های زیادی به عنوان کلاس اصلی خودرو انتخاب می‌شود و معیار دوم کارایی خود را از دست می‌دهد و اگر K عدد خیلی کوچکی انتخاب شود ممکن است خودرو در محدوده کلاس واقعی آن قرار نگیرد. بر این اساس با توجه به آزمایش‌های صورت گرفته؛ برای کاهش هر چه بیشتر خطا در معیار دوم، K برابر با $۰/۰۳$ انتخاب می‌شود به این معنی که تا ۳ درصد خطا برای هر مولفه‌ی ابعاد قابل قبول است.

بنابراین در صورتی که هر سه شرط رابطه ۵-۱ برای یک خودرو برقرار باشند، احتمال قرارگیری خودرو در آن کلاس وجود دارد. همچنین ممکن است که یک خودرو در بیش از یک کلاس قرار بگیرد. به عنوان مثال در خانواده پژو، ابعاد پژو ۴۰۵ و پژو پارس تقریباً یکسان است و در صورتی که خودروی عبوری پژو باشد در هر ۲ گروه قرار می‌گیرد. علاوه بر این احتمال کمی نیز وجود دارد که به دلیل عواملی همچون تغییر نوری محیط، ابعاد خودروی عبوری بیشتر از ۳ درصد با مقدار مرجع آن تفاوت داشته باشد. در این صورت معیار دوم خودرو را در دسته درست قرار نمی‌دهد، اما این موضوع در بخش بعدی به نحوی سازمان‌دهی شده که تاثیر منفی بر روی عمل کرد الگوریتم نداشته باشد. جدول ۵-۱ وابستگی ابعاد خودروی پراید سفید رنگ شکل ۵-۲ را به برخی از دسته‌ها نشان می‌دهد.

جدول ۵-۱: تعیین نوع پراید سفید رنگ شکل ۵-۲ به کمک ابعاد

نوع خودرو	پژو ۴۰۵	پراید ۱۳۱	سمند	پراید ۱۳۲	پژو پارس
امکان عضویت	×	✓	×	✓	×

علامت ✓ بیان گر مطابقت ابعاد خودروی مدنظر به ابعاد کلاس و علامت × بیان گر عدم مطابقت است.

۵-۳- تعیین نوع خودروی عبوری به کمک معیارهای اول و دوم

همان طور که در بخش قبل بیان شد، معیار اول در مواردی قادر نیست نوع خودروی عبوری را درست تشخیص دهد و گاهی خودروی عبوری را در خانواده های اشتباه قرار می دهد. همچنین گفته شد که استفاده از ابعاد هم در مواردی نمی تواند نوع دقیق خودرو را تعیین کند.

با بررسی های که بر روی نتایج معیار اول انجام گرفت مشخص شد تقریباً در تمام مواردی که این معیار به خطا رفته، نوع واقعی خودرو را به عنوان اولویت دوم خود در نظر گرفته است. ترکیب اطلاعات معیار اول و ابعاد برای تعیین نوع خودرو می تواند خطای کمتری از هر کدام از این روش ها داشته باشد. بنابراین برای تعیین نوع دقیق خودرو به این ترتیب عمل می شود که اولویت کلاس های برنده برای خودرو با توجه به امتیاز کسب شده توسط معیار اول به ترتیب تعیین می شود، آنگاه اولین اولیوی که شروط ابعاد را برآورده کرده باشد به عنوان نوع دقیق خودرو در نظر گرفته می شود. همچنین در صورتی که هیچ کدام از ۲ اولویت معیار اول، شرط ابعاد را برآورده نکند، کلاس خودرو همان کلاس تشخیص داده شده توسط معیار اول در نظر گرفته می شود (در این صورت احتمال دارد که معیار دوم بنا به دلایلی مانند تفاوت بیش از ۳ درصدی ابعاد یک خودروی عبوری با مقدار مرجع آن، به خطا رفته باشد که با در نظر گرفتن اولین اولویت معیار اول به عنوان کلاس برنده، از تاثیر منفی معیار دوم بر نتیجه الگوریتم جلوگیری می شود). جدول ۵-۲ تعیین نوع دقیق را برای خودروی پراید سفید رنگ شکل ۵-۲ نشان

می‌دهد. کلاس‌های کاندید برای روش BOW (با فرض استفاده به‌عنوان معیار اول) برحسب امتیاز به‌صورت صعودی به نزولی مرتب شده‌اند.

جدول ۵-۲: تشخیص کلاس پراید سفید رنگ شکل ۵-۲ به کمک معیار بازشناسی و معیار ابعاد

...	پژو پارس	پراید ۱۳۲	پراید ۱۳۱	پژو ۴۰۵	نوع خودرو
...	-۰/۱۵۳۷	-۰/۱۴۸۵	-۰/۱۴۷۱	-۰/۱۴۳۸	امتیاز BOW
...	×	✓	✓	×	انطباق ابعاد
×	×	×	✓	×	کلاس نهایی

همان‌طور که از جدول مشخص است، با ترکیب اطلاعات روش BOW و ابعاد، نوع خودرو به درستی شناسایی شده است. این درحالی است که معیار اول به تنهایی نتوانسته نوع خودرو را درست تعیین کند.

۵-۴- نتیجه‌گیری

روش‌هایی که تاکنون در زمینه بازشناسی خودرو ارائه شده است، غالباً روی تصاویر گرفته شده از نمای مستقیم و نزدیک خودروها آزمایش شده‌اند. در این فصل نحوه شناسایی نوع خودروهای عبوری به کمک دو معیار بررسی شد. معیار اول برای شناسایی نوع خودرو، همان روش‌های بازشناسی است. معیار دوم که مکمل معیار اول است، استفاده از ابعاد خودروها است. بررسی پاسخ روش‌های بازشناسی بیان‌گر این است که در اکثر مواردی که اولویت کلاس‌ها به اشتباه تعیین شده، کلاس اصلی خودرو، اولویت دوم بوده است. بنابراین در این فصل پس از تعیین احتمال هر کلاس توسط معیار اول برای خودروی عبوری، دو کلاسی که بیشترین احتمال را داشته‌اند، توسط معیار دوم مورد قضاوت نهایی قرار می‌گیرند. نتایج

بیان گر این است که استفاده از دو معیار، منجر به کاهش خطا در تعیین نوع خودروی عبوری در ویدئوی جاده‌ای می‌شود.

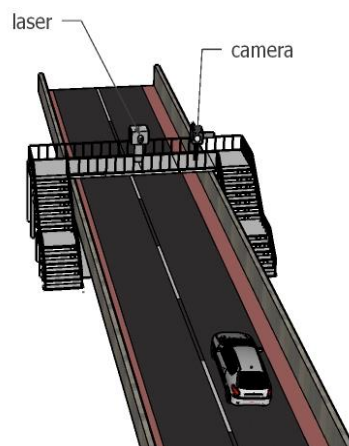
فصل ششم: بررسی نتایج

۶-۱- مقدمه

در ابتدای این فصل پایگاه داده جمع‌آوری شده که مجموعه‌ای جامع متشکل از ویدئوها و تصاویر است، بررسی می‌شود. در ادامه‌ی فصل نتایج روش پیشنهادی در تخمین ابعاد و سرعت در مقایسه با روش‌های دیگر بیان می‌شود و عمل‌کرد الگوریتم حذف سایه سنجیده خواهد شد. در بخش بعدی ابتدا روش بازشناسی پیشنهادی ارزیابی می‌شود و تاثیر الگوریتم حذف سایه بر الگوریتم‌های بازشناس بیان می‌شود. سپس اثر الگوریتم فراتفکیک‌پذیر پیشنهادی بر دقت بازشناسی قضاوت می‌شود و در نهایت نتایج الگوریتم فصل قبل که تعیین نوع خودروی عبوری در ویدئو توسط دو معیار است، بررسی می‌شود.

۶-۲- پایگاه داده‌های ویدئو و تصویر تهیه شده

پایگاه داده جمع‌آوری شده برای آزمایش الگوریتم از دو بخش تشکیل شده است : ۱- مجموعه‌ای از ویدئوها، ۲- مجموعه‌ای از تصاویر. ویدئوها با اخذ مجوز از پلیس و فیلم‌برداری از جاده‌های مختلف به‌دست آمده است. دوربین در این ویدئوها چند متر بالاتر از سطح جاده قرار گرفته و زاویه آن مشابه زاویه دوربین‌های نظارتی کنار جاده‌ای تنظیم شده است. تصویر شکل ۶-۱ نحوه تهیه ویدئوها را نمایش می‌دهد.



شکل ۶-۱: نحوه تهیه ویدئوهای مجموعه داده

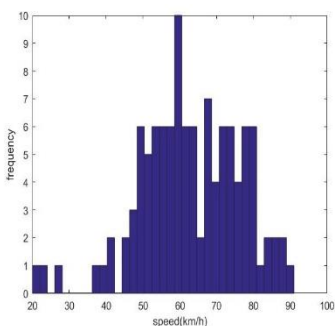
برای تخمین سرعت واقعی خودروها در این ویدئوها از یک دوربین لیزری ناجا استفاده شده، میانگین سرعت خودروها بسته به هر جاده متفاوت است. سطر اول تصاویر شکل ۶-۲، نمونه‌ای از تصاویر دوربین لیزری و سرعت ثبت شده توسط آن را نمایش می‌دهد. سطر دوم و سوم این شکل، نمودار تردد تعدادی خودرو بر حسب سرعت را برای دقایقی از ویدئوهای ضبط شده در محدوده شهر شاهرود نشان می‌دهد.



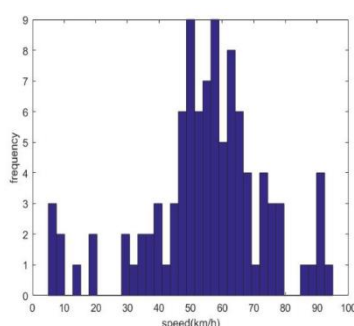
ب: سرعت ثبت شده توسط
دوربین لیزری (نمونه دوم)



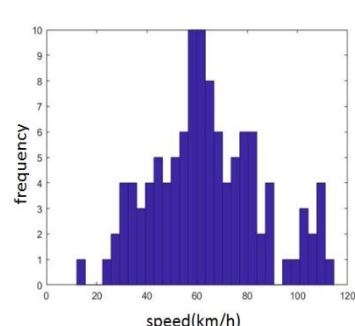
الف: سرعت ثبت شده توسط
دوربین لیزری (نمونه اول)



ه: نمودار تردد خودروها بر
حسب سرعت (نمونه سوم)



د: نمودار تردد خودروها بر
حسب سرعت (نمونه دوم)



ج: نمودار تردد خودروها بر
حسب سرعت (نمونه اول)

شکل ۶-۲: نمایش سرعت و نمودار تردد خودروها در برخی ویدئوهای مجموعه داده

مجموعه تصاویر جمع‌آوری شده حدود ۷۰۰۰ است که ۳۵۰۰ تصویر مربوط به نمای روبرو و ۳۵۰۰ تصویر مربوط به نمای پشت خودرو است. بخشی از تصاویر مربوط به جلوی خودرو تصاویری است که از پلیس تهیه شد و بخش دیگر تصاویری است که با استخراج یک تا دو تصویر برای هر خودرو، از فریم‌های نامتوالی ویدئوی جاده‌ای به‌دست آمد. تصاویر شامل گروه‌های زیادی از خودروها مختلف می‌شود لیکن

به طور منطقی اکثر این تصاویر مربوط به خودروهای سواری پرتردد در ایران است. شکل ۳-۶ تعدادی از این تصاویر را نشان می‌دهد.



شکل ۳-۶: نمونه‌ای از تصاویر جمع‌آوری شده

۳-۶- نتایج روش پیشنهادی در تخمین ابعاد و سرعت

به‌منظور ارزیابی روش پیشنهادی دو پایگاه داده مورد استفاده قرار گرفت. پایگاه داده اول همان پایگاه ارائه شده توسط دوبسکا [۶۰] است که شامل چند ویدئو است که از چند جاده تهیه شده‌اند. پایگاه دوم هم پایگاه جمع‌آوری شده توسط خود ما است. در ادامه نتایج روش پیشنهادی برای این دو پایگاه داده بررسی می‌شود.

۳-۶-۱- بررسی زمان پردازش هر مرحله‌ی روش پیشنهادی

ابعاد ویدئو در مجموعه داده دوبسکا، 480×854 و در مجموعه‌ی جمع‌آوری شده 540×960 است. با توجه به این‌که اندازه قاب‌ها در دو مجموعه‌ی آزمایش تفاوت زیادی باهم ندارند، می‌توان محدوده‌ی زمانی هر یک از مراحل روش پیشنهادی را برای هر قاب مطابق جدول ۳-۶ در نظر گرفت. این زمان‌ها برای یک رایانه با ۱۶ GB RAM و ۲/۶ GHZ CPU و با استفاده از نرم‌افزار متلب ۲۰۱۶ به‌دست آمده است. هر دو مجموعه، دارای چند ویدئو از جاده‌ها هستند که در برخی تردد خودروها زیاد و در برخی

کم است. در این جدول محدوده‌های پایین، میانگین زمان هر قاب برای ویدئوی با کمترین تردد و محدوده‌ی بالا، میانگین زمان هر قاب برای ویدئوی با بیشترین تردد است.

جدول ۶-۱: زمان مصرفی برای هر کدام از مراحل روش پیشنهادی بر حسب ثانیه

مرحله	زمان - تردد کم	زمان - تردد زیاد
آشکارسازی خودروها	۰/۰۵	۰/۰۸
حذف سایه	۰/۰۵	۰/۰۹
تشکیل جعبه و تصویر کردن	۰/۰۹	۰/۳۱
ردیابی	۰/۰۹	۰/۲۲
مجموع	۰/۲۸	۰/۷

۶-۳-۲- بررسی پاسخ روش پیشنهادی هنگام وجود یا عدم وجود سایه

همان‌طور که در فصول قبل اشاره شده، وجود و عدم وجود سایه از جمله عوامل موثر بر روی خطا، خصوصاً در تخمین ابعاد هستند. در جدول ۶-۲ محدوده خطای روش پیشنهادی برای حالت‌های «سایه و نیم سایه» و «بدون سایه یا سایه ضعیف» گزارش شده است. حداکثر خطا، مربوط به خودرویی است که الگوریتم در تخمین ابعاد و یا سرعت آن کمترین دقت را داشته و حداقل خطا مربوط به خودرویی است که الگوریتم برای تخمین پارامتر آن بیشترین دقت را داشته است. در پایگاه داده اول همه‌ی ویدئوها بدون سایه هستند ولی پایگاه داده دوم هر دو حالت با سایه و بدون سایه را شامل می‌شود. همچنین در پایگاه اول مکان دوربین و شرایط جاده‌ها مناسب‌تر بوده، به همین دلیل خطا برای این پایگاه در مجموع کمتر است. در تخمین سرعت چون از ردیاب استفاده می‌شود وجود یا عدم وجود سایه

تأثیر چندانی بر خطا ندارد. در مواردی که تفاوت سرعت تخمینی با مقدار ground truth عددی نزدیک به صفر بوده، برابر با صفر در نظر گرفته شده است.

جدول ۶-۲: بررسی خطای روش پیشنهادی در حالت وجود و عدم وجود سایه

		پایگاه داده اول		پایگاه داده دوم	
		سرعت km/h	ابعاد	سرعت km/h	ابعاد
سایه و نیم‌سایه	حداکثر خطا	۴/۹	۱۵	-	-
	حداقل خطا	۰	۱/۶۵	-	-
بدون سایه	حداکثر خطا	۴/۵	۹/۷	۴/۳	۹
	حداقل خطا	۰	۱	۰	۰/۴
میانگین		۱/۲	۸	۱/۱۳	۵/۱

۶-۳-۳- مقایسه نتایج روش پیشنهادی با روش‌های دیگر

به دلیل در دسترس نبودن پیاده‌سازی، پایگاه داده و همچنین نیاز به تنظیم دستی اکثر روش‌ها، امکان مقایسه منصفانه روش‌ها بر روی یک مجموعه‌ی مشخص وجود ندارد. مرجع [۵۱] برای مقایسه نتیجه روش‌ها با یک‌دیگر مقدار میانگین خطای تخمین سرعت هر کدام از روش‌ها را برای پایگاه داده‌ی مورد استفاده در همان مقاله، گزارش کرده است. جدول ۶-۳ مقدار خطای روش پیشنهادی و چند روش دیگر را نشان می‌دهد. لازم به ذکر است محاسبه میانگین خطا فقط برای خودروهایی صورت گرفته که توسط روش‌ها امکان شناسایی وجود داشته؛ بنابراین در صورتی که روشی نتوانسته خودرویی را تشخیص دهد، در محاسبه میانگین خطای روش، این خودرو نادیده گرفته شده است. روش‌های [۴۵، ۴۹، ۵۲، ۶۰] که در جدول مورد مقایسه گرفته‌اند، روش‌هایی هستند که در [۵۱] مورد تأکید بوده‌اند. علامت "-" در جدول بیان‌گر این است که الگوریتم قابلیت اندازه‌گیری کمیت مد نظر را نداشته است.

جدول ۶-۳: میانگین خطای تخمین ابعاد و سرعت چند روش

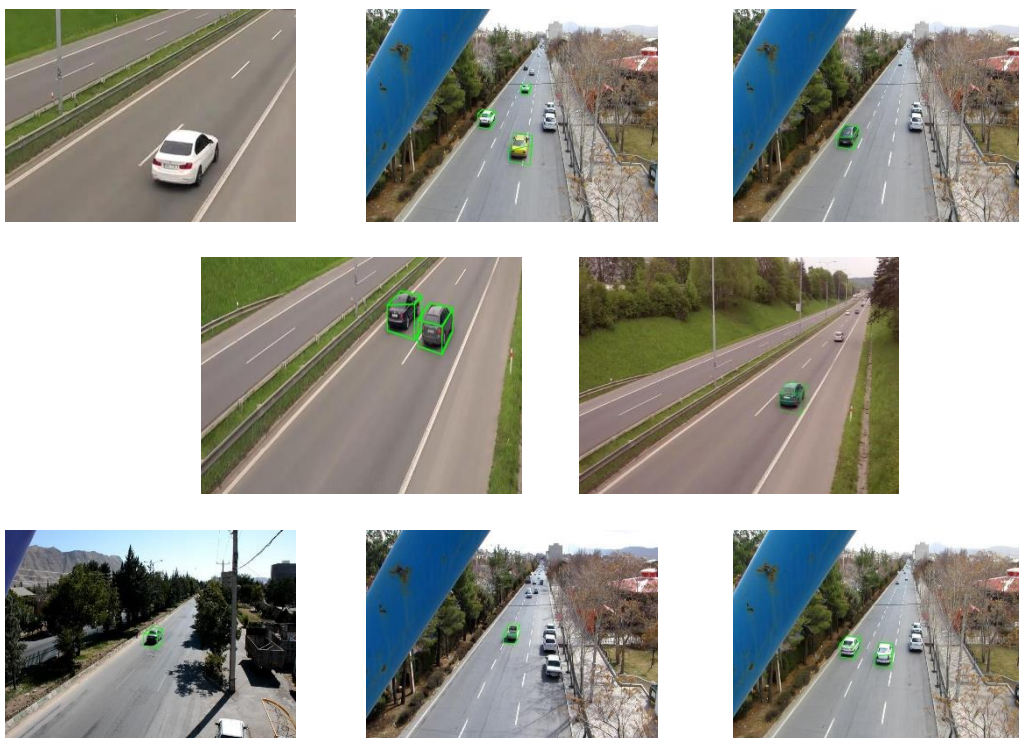
روش	نحوه تهیه ground truth	میانگین خطا در تخمین سرعت کیلومتر بر ساعت (km/h)	میانگین خطا در تخمین ابعاد درصد (%)
سینا [۴۵]	GPS	۳/۳	—
آقایان [۴۸]	GPS	۱/۳۹	—
لویسنون [۴۹]	حلقه‌های القایی	۱/۶۳	—
فیلیپاک [۵۲]	حلقه‌های القایی	۲/۳	—
اسلامی [۵۳]	رادار و لیزر	۱/۲۲	—
فاموری [۵۵]	لیزر	۱/۴۳	—
دوبسکا [۶۰]	پایگاه اول-GPS	۱/۵۶	٪۰.۶
دوبسکا [۶۰]	پایگاه دوم- لیزر	خطای بیش از ٪۲۰	خطای بیش از ٪۲۰
پیشنهادی	پایگاه اول-GPS	۱/۱۳	٪۵.۱
پیشنهادی	پایگاه دوم- لیزر	۱/۲	٪۰.۸

در میان الگوریتم‌های مورد مقایسه روش‌های [۴۵، ۴۸، ۴۹] برای کالیبراسیون نیازمند ورود پارامترهایی مانند ارتفاع، زاویه دوربین متناسب با محل نصب دوربین توسط کاربر هستند، وجود کمی خطا در اندازه‌گیری این پارامترها منجر به ایجاد خطای زیادی در کالیبراسیون می‌شود. علاوه بر این مواردی مانند ناهمواری و شیب سطح جاده می‌تواند در افزایش خطای این روش‌ها تأثیر زیادی داشته باشد.

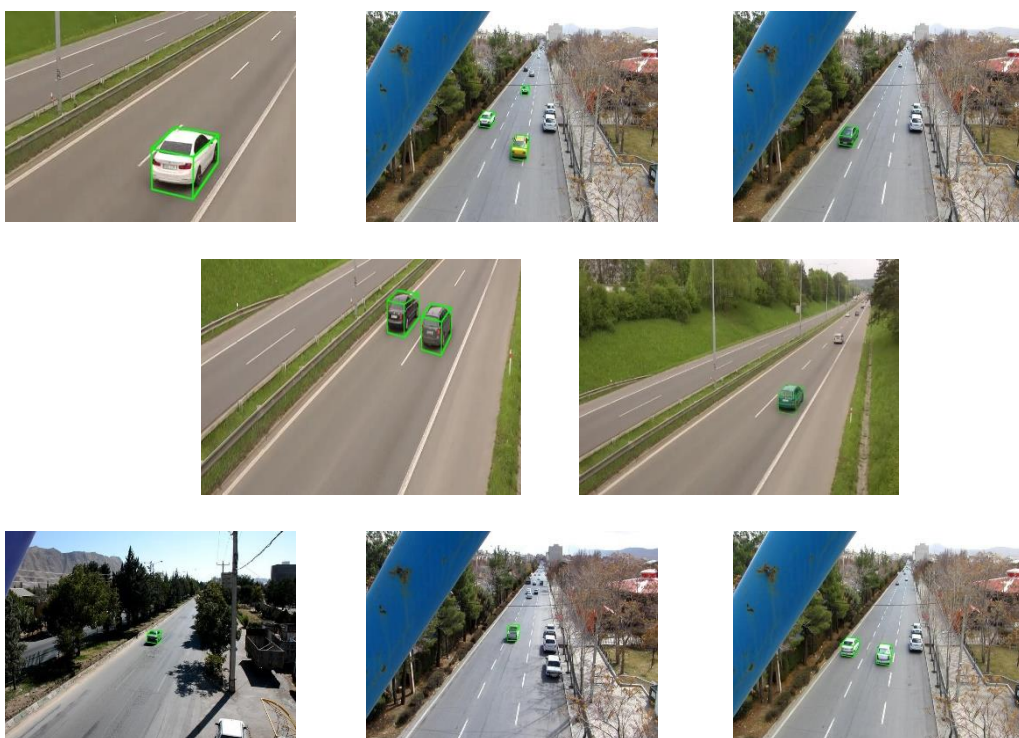
در [۵۲] که کالیبراسیون آن مبتنی بر ابعاد پلاک خودرو است، شناسایی محل دقیق پلاک خودرو در مواردی به راحتی امکان پذیر نیست و منجر به ایجاد خطا می‌شود. در [۵۳] کالیبراسیون مبتنی بر ابعاد و فونت پلاک است. هرچند در این روش با به کارگیری الگوریتم‌های مناسب دقت افزایش یافته، اما به هر حال طبق گفته نویسنده عدم شناسایی دقیق محدوده برخی پلاک‌ها خطای این روش را کمی

افزایش داده است. روش [۵۵] همانند روش‌های [۵۲] و [۵۳] نیازمند تعیین محدوده چند پلاک است، در نتیجه خطا در تعیین پلاک در خطای تخمین سرعت این روش تأثیر دارد.

برای روش [۶۰] و روش پیشنهادی، به دلیل شباهت ساختاری و در دست بودن پیاده‌سازی امکان مقایسه دقیق‌تر بر روی پایگاه داده یکسان فراهم است. شکل ۴-۶ نتایج روش پیشنهادی و روش دوبسکا را برای قاب‌هایی که روش دوبسکا پاسخ خوبی نداشته نشان می‌دهد. تصاویر گروه (الف) نتایج روش دوبسکا و تصاویر گروه (ب) نتایج روش پیشنهادی است. همان‌گونه که مشخص است در این تصاویر به دلیل تغییرات نوری محیط و بروز سایه، امکان شناسایی محدوده دقیق خودرو برای این روش فراهم نبوده و جعبه‌ی خودرو بزرگ‌تر از اندازه واقعی آن است؛ حتی در تصاویر خودروی سفید رنگ سطر اول، این روش قادر نبوده خودرو را شناسایی کند و برای آن جعبه‌ای تشکیل دهد. این در حالی است که روش پیشنهادی برای همه موارد پاسخ خوبی داشته که دلیل اصلی آن مدل کردن بهتر پس‌زمینه و ایجاد مقاومت در برابر سایه است.



الف: نتایج حاصل از به کارگیری روش دوبسکا



ب: نتایج حاصل از به کارگیری روش پیشنهادی

شکل ۴-۶: مقایسه دقت در تشکیل جعبه محیطی خودرو

مطابق جدول ۳-۶ روش دوبسکا [۶۰] برای پایگاه آزمایش اول خطای زیادی ندارد اما برای مجموعه آزمایش دوم به دلیل عدم مقاومت نسبت به سایه کارایی خود را از دست می‌دهد و خطای آن به شدت افزایش می‌یابد. این روش مشکلات دیگری هم دارد، به عنوان مثال با توجه به شکل ۴-۶، در مواردی قادر نیست خودروهای عبوری را شناسایی کند. همچنین دقت روش به کار رفته برای تشکیل جعبه خودرو و محاسبه سرعت و ابعاد در این روش پایین است. علاوه بر این، مراحل این روش محاسبات سنگینی دارد که منجر می‌شود سرعت پردازش روش کاهش یابد به گونه‌ای که میانگین زمان اجرای آن برابر با ۱/۳۶ قاب در ثانیه است.

با توجه به جدول ۳-۶، روش پیشنهادی پاسخ بهتری نسبت به سایر روش‌ها دارد. دلیل این موضوع، ارائه و استفاده از روش‌های بهینه و دقیق در کالیبراسیون، شناسایی خودروها، محاسبه پارامترها و افزایش مقاومت نسبت به سایه است. دلیل این که مقدار خطا برای پایگاه دوم نسبت به پایگاه اول افزایش یافته، تابش شدید نور خورشید بر سطح جاده و بدنه‌ی وسایل نقلیه است که تشکیل دقیق جعبه‌ی خودرو را کمی مشکل می‌کند و در نتیجه در محاسبه سرعت و ابعاد، مخصوصاً ارتفاع خودرو، مقداری خطا ایجاد می‌کند. لیکن در مجموع خطای روش برای هر دو پایگاه داده کم است که این موضوع بیان‌گر عمل کرد خوب در شرایط نوری گوناگون است. میانگین زمان مصرفی برای پردازش هر قاب ۰/۴۶ ثانیه است که در مقایسه با روش [۶۰] و با توجه به توان بالاتر روش پیشنهادی، می‌توان نتیجه گرفت این روش از لحاظ زمان اجرا هم در جایگاه خوبی قرار دارد.

۴-۶- تخمین نوع خودروی عبوری

شناسایی نوع خودرو هم در مرحله کالیبراسیون دوربین اهمیت دارد و هم در صورتی که نوع خودروی عبوری، به عنوان خروجی الگوریتم مد نظر باشد. با توجه به این که شناسایی وسایل پرتدد در عمل اهمیت بیشتری دارد، کلاس‌های مد نظر برای شناسایی ۹ کلاس از خودروهای پرتدد در نظر گرفته شد. این کلاس‌ها عبارتند از {پراید ۱۳۱، پراید ۱۳۲، پراید ۱۴۱، پژو ۴۰۵، پژو پارس، پژو ۲۰۶، پیکان،

ال ۹۰، سمند، وانت باری، اتوبوس-کامیون و سایر}. گروه‌های وانت باری، اتوبوس-کامیون و سایر، همگی گروه‌های متفرقه هستند که جهت کاهش خطا، به‌عنوان کلاس‌های اضافی به سیستم آموزش داده شدند. تصویر شکل ۵-۶، تصویر یکی از دوربین‌های نظارتی ترافیک است. بجز ۱ الی ۲ خودروی این تصویر، مابقی خودروهای موجود در دسته کلاس‌های مد نظر برای شناسایی قرار می‌گیرند. در ادامه دقت روش‌های بازشناسی تصویر بررسی می‌شود و عمل‌کرد روش ارائه شده در فصل پنجم که برای شناسایی نوع خودرو از روی ویدئو به کمک دو معیار است بیان خواهد شد.



شکل ۵-۶: تصویر دوربین نظارتی سطح شهر (تهیه شده از اینترنت)

۶-۴-۱- بررسی دقت روش‌های بازشناسی

روش‌های مورد استفاده برای بازشناسی هم برای کالیبراسیون و محاسبه ضرایب متریک در فصل سوم و هم به‌عنوان معیار اول در فصل پنجم کاربرد دارند. جدول ۴-۶ دقت این روش‌ها را نمایش می‌دهد. علی‌رغم فاصله زیاد دوربین از خودروها، این روش‌ها از پاسخ خوبی برخوردار هستند. با توجه به جدول، میانگین دقت مکانیزم بازشناسی پیشنهادی که در فصل چهار ارائه شد بیشتر از BOW و VGG است. این در حالی است که شبکه CNN مورد استفاده در این ساختار، شبکه‌ای نسبتاً ساده است. (تغییر در لایه‌های این شبکه‌ی CNN و آزمایش مکانیزم پیشنهادی بر روی پایگاه تصاویر خودروهای دیگر و

مقایسه نتایج آن با مراجع معتبر شناسایی نوع خودرو که در فصل ۲ بررسی شدند، بیانگر کارایی و دقت بالای آن در شناسایی نوع خودرو است).

جدول ۴-۶: دقت روش‌های بازشناسی در طبقه‌بندی خودروها

نوع خودرو	پراید ۱۳۱	پژو پارس	پژو ۲۰۶	پژو ۴۰۵	پراید ۱۳۲	پیکان	ال ۹۰	سمند	پراید ۱۴۱	اتوبوس	وانت	سایر	میانگین
BOW	۹۰	۹۲	۱۰۰	۸۴	۸۵	۹۶	۹۲	۸۹/۶	۹۱	۸۴	۸۸	۸۷/۵	۸۹/۹
VGG	۹۲	۹۴	۱۰۰	۸۶	۸۴	۹۴	۹۶	۸۸	۹۴	۸۶	۸۶	۸۵	۹۰/۴
پیشنهادی	۹۵	۹۶	۹۸	۹۵	۹۳	۹۹	۹۷	۹۶	۹۷	۸۳	۸۵	۸۴	۹۳/۱

۶-۴-۲- تاثیر الگوریتم فراتفکیک‌پذیری پیشنهادی بر دقت بازشناسی

جدول ۶-۵ میانگین دقت روش‌های بازشناسی را قبل و بعد از اعمال الگوریتم فراتفکیک‌پذیر فصل چهارم نشان می‌دهد. بعد از اعمال الگوریتم ابعاد تصاویر از 300×300 به 400×400 افزایش یافته است. با توجه به این جدول، فراتفکیک‌پذیری دقت روش‌ها را افزایش داده است.

جدول ۶-۵: بررسی اثر فراتفکیک‌پذیری بر روی دقت بازشناسی

نام روش	قبل از اعمال فراتفکیک‌پذیری	پس از اعمال فراتفکیک‌پذیری
BOW	۸۹/۹	۹۱
VGG	۹۰/۴	۹۲/۲
روش بازشناسی پیشنهادی	۹۳/۱ ✓	۹۴/۸ ✓

۶-۴-۳- اهمیت حذف سایه در تشخیص نوع خودرو

حذف سایه برای تشخیص دقیق ناحیه خودرو علاوه بر افزایش دقت در تعیین ابعاد و سرعت، تاثیر زیادی نیز در افزایش دقتِ روشِ بازشناسی دارد. تصویر شکل ۶-۶ محدوده با سایه و بدون سایه را برای یک دستگاه پژو ۴۰۵، و جدول ۶-۶ نتیجه شناسایی با الگوریتم BOW را برای هر دو تصویر نشان می‌دهد. همان‌طور که مشخص است حذف سایه منجر شده BOW درست عمل کند (تعداد تصاویر آموزش داده شده به BOW، ۵۰۰۰ تصویر بوده است).



ب: پس از حذف سایه



الف: قبل از حذف سایه

شکل ۶-۶: محدوده شناسایی شده برای خودرو قبل و بعد از حذف سایه

جدول ۶-۶: بررسی اهمیت حذف سایه در تشخیص کلاس

نوع خودرو	امتیاز BOW قبل از حذف سایه	امتیاز BOW پس از حذف سایه
پراید ۱۳۱	-۰/۱۰۴۶	-۰/۱۱۷۵
پژو پارس	-۰/۱۱۷۷	-۰/۱۱۷۹
پژو ۲۰۶	-۰/۱۱۷۷	-۰/۱۱۷۴
پژو ۴۰۵	-۰/۱۱۷۷	-۰/۱۰۷۵
پراید ۱۳۲	-۰/۱۱۸۱	-۰/۱۱۹۵
پیکان	-۰/۱۱۷۷	-۰/۱۱۸۷
ال ۹۰	-۰/۱۱۷۷	-۰/۱۱۹۰
سمند	-۰/۱۱۷۷	-۰/۱۱۹۰
پراید ۱۴۱	-۰/۱۱۷۷	-۰/۱۱۹۰

۶-۴-۴- شناسایی نوع خودرو به کمک ابعاد و الگوریتم‌های بازشناسی

سایر روش‌های ارائه شده در زمینه شناسایی نوع خودرو از روی تصاویر نوع را تعیین می‌کنند. به دلیل عمل کرد روش پیشنهادی بر روی ویدئو، یکی نبودن مجموعه‌های آزمایش و عدم دسترسی به پیاده‌سازی سایر روش‌ها، امکان مقایسه منصفانه با پژوهش‌های دیگر وجود ندارد. جدول ۶-۷ نتایج را برای شناسایی خودرو در ویدئوهای آزمایش نشان می‌دهد. ویدئوهای آزمایش تحت همان شرایط ویدئوهای مربوط به تهیه تصاویر آموزشی تهیه شده است. در این جدول منظور از اجماع قاب‌ها، استفاده از رابطه ۳-۷ (ردیابی و تعیین نوع خودرو به کمک چند قاب) برای تعیین نوع خودرو است.

جدول ۶-۷: بررسی دقت در شناسایی خودروها برای ویدئوهای مورد آزمایش

روش بازشناسی	میانگین دقت بازشناسی برای ویدئوهای آزمایش
اجماع قاب‌ها + BOW	۷۳/۲ %
اجماع قاب‌ها + BOW + معیار ابعاد	۷۷ %
اجماع قاب‌ها + VGG	۷۴/۸ %
اجماع قاب‌ها + VGG + معیار ابعاد	۷۸/۱ %
اجماع قاب‌ها + روش پیشنهادی	۷۶/۵ %
اجماع قاب‌ها + روش پیشنهادی + معیار ابعاد	۷۹/۷ %

در [۷۵] اشاره شده که افزایش مجموعه‌های آزمایش منجر به افزایش خطا می‌شود، در این جا هم این رابطه برقرار است. همچنین در صورتی که کلاس‌های آموزشی زیاد شود کارایی معیار دوم کاهش می‌یابد و احتمال افزایش خطا وجود دارد. اما به هر حال با شناسایی کلاس‌های رایج خودرو نیاز اصلی سامانه‌های نظارتی ترافیک برطرف می‌شود. علاوه بر این همان‌طور که از جدول ۶-۷ مشخص است، استفاده از دو معیار منجر به کاهش خطا گردیده و دقت خوبی بر روی ویدئو فراهم آورده است.

۵-۶- نتیجه گیری

در این فصل نتایج روش‌های پیشنهادی مورد بررسی قرار گرفت. برای آزمایش روش پیشنهادی از دو مجموعه داده استفاده شد. مجموعه داده دویسکا و مجموعه‌ای که خود ما جمع‌آوری کردیم. مجموعه داده تهیه شده شامل ویدئوهای دارای برچسب سرعت و شامل تصاویر دسته‌بندی شده از خودروها است. یکی از مزایای این مجموعه این است که از زاویه مشابه زاویه دوربین جاده‌ای و در شرایط نوری مختلف محیطی تهیه شده که البته این موضوع برای الگوریتم‌ها چالش‌هایی ایجاد می‌کند.

خطای روش پیشنهادی برای مجموعه جمع‌آوری شده در تخمین ابعاد برابر با ۸ درصد و در تخمین سرعت برابر با ۱/۲ کیلومتر بر ساعت به دست آمد که مقایسه با روش‌های دیگر حاکی از خطای کمتر این روش است. از جمله دلایل کاهش خطا در این روش، ایجاد مقاومت در برابر سایه و شناسایی بهتر پیش‌زمینه است.

دقت الگوریتم بازشناسی پیشنهادی (علی‌رغم استفاده از یک شبکه CNN ساده) برای تصاویر پایگاه داده برابر با حدود ۹۳ درصد به دست آمد. همچنین اعمال الگوریتم فراتفکیک‌پذیر بر روی این تصاویر، منجر به بهبود ۱/۷ درصدی در دقت بازشناسی گردید، این بهبود بیان‌گر عمل‌کرد خوب الگوریتم فراتفکیک‌پذیر پیشنهادی است. تحلیل نتایج حاکی از این است که الگوریتم حذف سایه منجر به افزایش دقت بازشناسی هم می‌شود. استفاده از معیار دوم در کنار معیار اول برای شناسایی نوع خودرو، دقت نزدیک به ۸۰ درصد را حاصل نمود و نسبت به حالتی که یک معیار به تنهایی استفاده شود، ۳ درصد بهبود ایجاد کرد.

فصل هفتم: نتیجه‌گیری و پیشنهادهایی برای ادامه کار

۷-۱- مقدمه

در حال حاضر بسیاری از جاده‌ها و خیابان‌های شهری مجهز به دوربین‌های نظارتی جاده هستند. یکی از موضوعات روز دنیا استفاده از ویدئوی این دوربین‌ها برای تحلیل رفتار خودروها و تخمین مواردی مانند سرعت و نوع خودروها است. در این صورت نیاز به استفاده از ناظر انسانی و سخت‌افزارهای گران قیمت مانند لیزر کاهش می‌یابد. اکثر روش‌های موجود در این زمینه با مشکلاتی همچون نیاز به کالیبراسیون دستی، تک کاربرده بودن، عدم مقاومت به تغییرات نوری محیط و عدم آزمایش بر روی داده‌های یکسان و واقعی مواجه هستند. لذا در انجام این رساله با هدف کاهش این مشکلات، الگوریتمی معرفی گردید که قادر به کالیبراسیون تمام‌خودکار است؛ می‌تواند سرعت، ابعاد و نوع خودروها را شناسایی کند؛ همچنین در برابر تغییرات نوری محیط و سایه پاسخ خوبی دارد. علاوه بر این، مجموعه داده‌ای جامع، منطبق بر شرایط دوربین‌های جاده‌ای تهیه گردید. در ادامه این فصل ابتدا عمل‌کرد روش‌های پیشنهادی به طور خلاصه جمع‌بندی می‌شود، سپس برخی ایرادهای روش بیان می‌شود و نهایتاً پیشنهادهایی برای ادامه کار مطرح خواهد شد.

۷-۲- جمع‌بندی روش‌های پیشنهادی در رساله

در این رساله روشی خودکار برای محاسبه سرعت و ابعاد خودروها ارائه شد. قابلیت محاسبه ابعاد و وجود الگوریتم‌های بازشناسی خودرو در بدنه الگوریتم، امکان شناسایی نوع خودرو را نیز فراهم آورد. در فصل سوم نحوه کالیبراسیون و محاسبه ابعاد و سرعت ارائه شد. الگوریتم این فصل شامل ۳ فاز است که در فاز اول و دوم، پارامترهای کالیبراسیون به‌طور تمام‌خودکار و به کمک شناسایی تعدادی نماینده برای دو کلاس رایج خودرو، محاسبه می‌شوند. با تکمیل پارامترهای کالیبراسیون امکان نگاشت اندازه‌گیری‌های پیکسلی از فضای تصویر به اندازه‌گیری‌های متری در فضای واقعی فراهم می‌شود. پس از پایان فازهای اول و دوم، در فاز سوم با آشکارسازی خودرو، حذف سایه، تشکیل جعبه خودرو، حذف پرسپکتیو و ردیابی خودروها؛ ابعاد و سرعت خودروها با دقت خوب محاسبه می‌شود. نتایج الگوریتم ارائه شده در

این فصل، به صورت اجمالی در فصل ششم از نظر زمان اجرای مراحل، تاثیر حذف سایه و مقایسه کمی و کیفی با سایر روش‌ها بررسی شد. این نتایج بیان‌گر عمل‌کرد دقیق‌تر و مقاومت بیشتر این روش در صورت ایجاد سایه و تغییرات نوری است.

شناسایی خودروهای رایج برای کالبراسیون تمام‌خودکار در فصل سوم، وابسته به الگوریتم‌های بازشناسی نوع خودرو است. در فصل چهارم ساختاری قوی مبتنی بر شبکه‌های عمیق و تثبیت تصویر برای بازشناسی خودرو ارائه گردید. همچنین با توجه به تاثیر مثبت افزایش تفکیک‌پذیری بر نرخ بازشناسی، الگوریتمی برای افزایش تفکیک‌پذیری تصویر در این فصل ارائه شد. در فصل ۶، نتیجه ساختار بازشناسی و الگوریتم فراتفکیک‌پذیری برای بازشناسی تصاویر بررسی گردید. مقایسه نتایج ساختار بازشناسی پیشنهادی با روش‌های دیگر بیان‌گر دقت بالاتر آن و نتایج روش فراتفکیک‌پذیری نشان دهنده تاثیر این الگوریتم در بهبود بازشناسی است.

در فصل پنجم به کمک قابلیت تعیین ابعاد و الگوریتم بازشناسی فصل ۴، روشی ترکیبی برای شناسایی نوع خودروهای عبوری در ویدئو معرفی شد. در انتهای فصل ۶، نشان داده شد که استفاده از معیار ابعاد در کنار الگوریتم‌های بازشناسی، منجر به افزایش دقت در شناسایی نوع خودرو در ویدئو می‌شود. همچنین نشان داده شد که الگوریتم پیشنهادی برای حذف سایه، علاوه بر بهبود دقت در تخمین ابعاد، منجر به بهبود دقت روش‌های بازشناسی هم می‌شود.

در فصل ششم ابتدا پایگاه داده جمع‌آوری شده که شامل مجموعه‌ای از ویدئوها و تصاویر دارای برچسب است معرفی شد. ویژگی این مجموعه این است که از زاویه و ارتفاع مشابه زوایه و ارتفاع دوربین جاده‌ای تهیه شده و همچنین برای اولین بار است که چنین مجموعه کاملی شامل ویدئوها و خودروهای ایرانی جمع‌آوری می‌شود. در ادامه فصل ۶ نتایج الگوریتم‌های ارائه شده در فصل‌های ۳ تا ۵ مورد بررسی قرار گرفت.

۷-۳- ایرادات روش پیشنهادی

هر چند روش پیشنهادی نسبت به سایر روش‌های مورد بررسی برتری دارد، اما ایراداتی نیز دارد که برخی آن‌ها عبارتند از :

۱. یکی از مشکلاتی که برای این روش و روش‌های دیگر تشخیص ابعاد وجود دارد، ایجاد خطا در صورت نزدیکی بیش از حد خودروها به یکدیگر است. در صورتی که جعبه خودروها با هم هم‌پوشانی پیدا کند، این جعبه‌ها با یکدیگر ادغام می‌شوند و به صورت یک جعبه بزرگ در نظر گرفته می‌شوند. البته در روش پیشنهادی با توجه به حذف سایه، ابعاد جعبه بزرگ‌تر از اندازه واقعی نیست و این مشکل تا حد زیادی مرتفع می‌شود، اما همچنان احتمال هم‌پوشانی وجود دارد. این خطا از روش مورد استفاده برای شناسایی پیش‌زمینه ناشی می‌شود؛ این روش همه اجسام متحرک را شناسایی می‌کند و قادر نیست بین این اجسام تمایز قائل شود و در صورتی که دو خودرو به هم نزدیک شوند، این دو خودرو را یکی تلقی می‌کند. ادغام جعبه‌ها با یکدیگر امکان محاسبه خطا برای یک خودرو را در برخی از قاب‌ها سلب می‌کند. البته در توضیحات مربوط به جدول ۶-۳ گفته شد که اگر امکان شناسایی خودرو به هر دلیلی برای یک روش فراهم نباشد، این خودرو در محاسبه خطا لحاظ نمی‌شود. این عدم شناسایی محدوده دقیق خودرو بیشتر بر تخمین ابعاد اثر دارد و تاثیر کمتری بر خطای تخمین سرعت دارد، زیرا ممکن است امکان ردیابی خودرو فراهم باشد. از جمله ایرادات دیگر این روش چند تکه در نظر گرفتن یک خودرو بر اثر وجود عواملی همچون درخت و یا تیر چراغ برق در جاده است، این موضوع منجر می‌شود که الگوریتم یک خودرو را به عنوان مثال دو خودرو در نظر بگیرد. شکل ۷-۱ این مشکلات را نشان می‌دهد.



ج: یکی گرفتن خودروها به
دلیل نزدیکی به هم (نمونه ۲)

ب: یکی گرفتن خودروها به
دلیل نزدیکی به هم (نمونه ۱)

الف: دو تکه گرفتن خودرو در
اثر وجود مانع در صحنه

شکل ۷-۱: نمایش برخی مشکلات روش پیشنهادی

۲. الگوریتم [۱] که برای تشکیل نقاط محوشدگی مورد استفاده قرار گرفت، در صورتی می‌تواند نقاط محو را درست به دست آورد که زاویه مناسبی (مشابه زاویه دوربین‌های نظارتی کنار جاده) داشته باشد. لذا محدودیت زاویه دوربین در روش پیشنهادی حائز اهمیت است. به عنوان مثال اگر دوربین در محل پیچ جاده نصب شده باشد، الگوریتم نمی‌تواند درست عمل کند. زیرا این روش از راستای حرکت خودروها برای تخمین نقاط محو استفاده می‌کند و در جاده‌های دارای انحناء، تخمین خوبی از راستای حرکت به دست نمی‌آورد.

۳. هر چند در این روش سعی شد مقاومت نسبت به تغییرات نوری تا حد امکان افزایش یابد، اما به هر حال مواردی مانند بازتاب نور از بدنه خودرو و سطح جاده، شباهت رنگی پس‌زمینه و خودرو، نویز دوربین و اعوجاج لنز، منجر به افزایش خطا در هر قاب و نهایتاً افزایش خطای کلی خصوصاً در تخمین ابعاد می‌شود. همچنین در صورتی که تغییرات پس‌زمینه زیاد باشد ممکن است به اشتباه قسمتی از جاده به عنوان پس‌زمینه در نظر گرفته شود. به عنوان مثال در شکل ۷-۲ به دلیل وزش باد و حرکت سیم برق و درختان، به اشتباه بخش از پس‌زمینه، خودرو تلقی شده است. این خطا ناشی از روش مورد استفاده برای شناسایی پیش‌زمینه است.



شکل ۷-۲: انتخاب بخشی از پس زمینه به عنوان خودرو

۴. موضوع حائز اهمیت دیگر این است که در پایگاه داده جمع آوری شده تعداد خودروهایی که دارای سرعت بالا هستند کم است (این موضوع برای تمامی پایگاه های داده موجود برقرار است). این موضوع منجر می شود که نتوان تحلیل دقیقی از عمل کرد الگوریتم در سرعت های بالا داشت.
۵. در روش پیشنهادی سرعت نسبت به روش های مشابه افزایش یافت، اما به هر حال سرعت این روش در نرم افزار متلب بلادرنگ نیست.
۶. اگر هدف شناسایی نوع خودرو باشد، با افزایش تعداد کلاس های مد نظر برای شناسایی، کارایی معیار دوم که شناسایی بر اساس ابعاد است، کاهش می یابد. دلیل این موضوع شباهت ابعادی کلاس های خودرو به یکدیگر است.

۷-۴- پیشنهادهایی برای بهبود الگوریتم در آینده

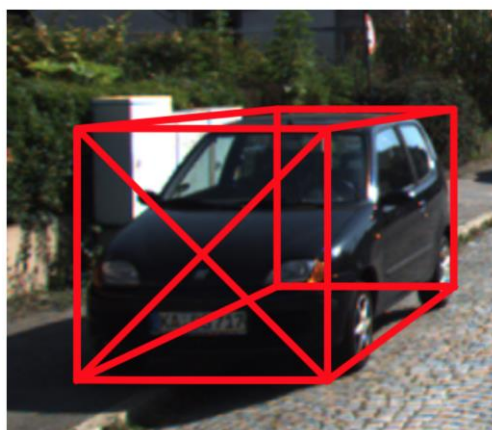
برای کاهش ایرادها و بهبود این روش می توان موارد ۱ تا ۶ را مورد بررسی قرار داد :

۱. همان طور که اشاره شد ممکن است به دلیل نزدیکی خودروها و یا دید کم دوربین، خودروها با یکدیگر هم پوشانی پیدا کنند و یا حتی بعضی خودروها دیده نشوند. برای حل این مشکل

می‌توان در بزرگراه‌ها و یا مسیرهای مهم، از چند دوربین استفاده کرد و با هم‌زمان کردن دوربین‌ها، رفتار تعداد بیشتری خودرو را تحلیل کرد [۱۰۴].

۲. در جاده‌های دارای قوس و پیچ و خم که انتخاب یک نقطه به‌عنوان نقطه محوشدگی منجر به افزایش خطا می‌شود، می‌توان با تقسیم جاده به چند ناحیه، برای هر ناحیه یک نقطه محوشوندگی استخراج کرد؛ به‌گونه‌ای که همه نقاط استخراجی در فاصله نزدیک به هم باشند. در این حالت انطباق جعبه تشکیلی با خودرو افزایش پیدا می‌کند و خطا کاهش می‌یابد.

۳. پیشرفت شبکه‌های عصبی عمیق و افزایش قدرت سیستم‌های پردازنده، امکان استفاده راحت‌تر از این شبکه‌ها را فراهم می‌کند. در صورت آموزش کامل، شبکه‌های عمیق این قابلیت را دارند که برای انتخاب محدوده دقیق خودرو، حتی در صورت وجود سایه مورد استفاده قرار بگیرند. همچنین این شبکه‌ها این قابلیت را دارند که برای تشکیل دقیق جعبه محیطی خودرو در آینده استفاده شوند. این مهم می‌تواند امکان کالیبراسیون دقیق‌تر را فراهم آورد. تصویر شکل ۷-۳ تشکیل جعبه خودرو به کمک شبکه‌های عمیق را نشان می‌دهد [۱۰۵].



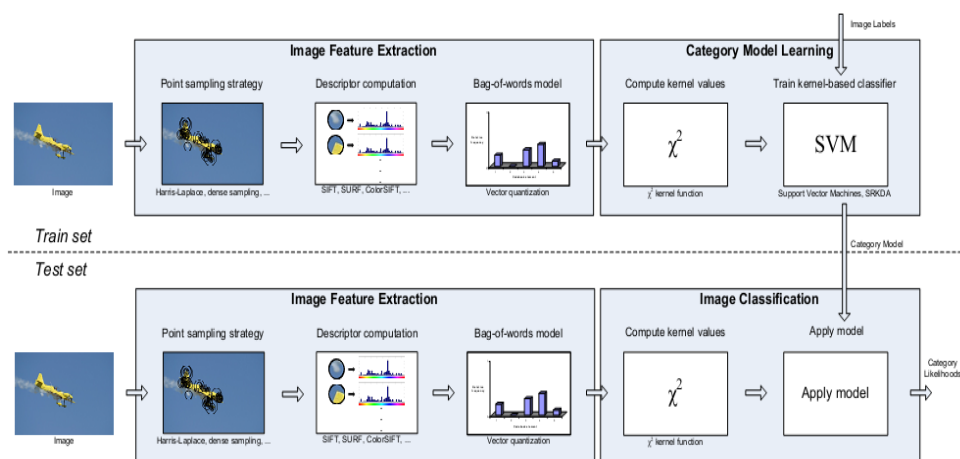
شکل ۷-۳: تشکیل جعبه محیطی خودرو به کمک شبکه‌های عمیق

۴. پژوهش‌هایی هستند که با استفاده از ناحیه‌بندی یا روش‌های دیگر، سطح جاده‌ها را شناسایی می‌کنند. می‌توان به کمک این روش‌ها جاده‌ها را شناسایی کرد و خطاهایی مانند خطای رخ داده در شکل ۷-۲ را از بین برد.
۵. تحلیل بهتر از تنوع و رفتار خودروها نیازمند پایگاه داده بزرگ‌تر است. برای این منظور می‌توان با اطلاع‌رسانی و برگزاری چالش جمع‌آوری داده (مانند چالش [۱۰۶] dota)، از طریق سازمانی همچون دانشگاه، پایگاه کنونی را گسترش داد.
۶. پیاده‌سازی روش پیشنهادی بر روی GPU به کمک opencl و یا cuda، می‌تواند منجر به افزایش سرعت آن گردد و امکان استفاده بلادرنگ را فراهم کند.

پوست الف: روش کیفی کلمات

۱- مقدمه

یکی از روش‌هایی که برای دسته‌بندی تصاویر مورد استفاده قرار می‌گیرد، روش کیف کلمات^۱ [۱۰۷] است. این روش از عمل‌کرد و کارایی بالایی در حوزه دسته‌بندی تصاویر برخوردار است و به همین دلیل محبوبیت بالایی نیز دارد. در استفاده از این روش برای دسته‌بندی تصاویر همان‌طور که در شکل ۱ نشان داده شده، سه مرحله اصلی، استخراج ویژگی تصویر، یادگیری مدل دسته‌بندی و طبقه‌بندی تصویر تشکیل وجود دارد. مرحله استخراج ویژگی خود از سه مرحله مختلف تشکیل شده است که در مرحله اول نقاط کلیدی تصویر که حاوی اطلاعات محلی تصویر هستند، شناسایی می‌شوند. سپس با استفاده از توصیف‌گرها اطراف این نقاط بررسی می‌شود. با توجه به این که هر تصویر حاوی چندین نقطه کلیدی است، با استفاده از توصیف‌گرها نقاط کلیدی به صورت بردارهای عددی نمایش داده می‌شوند. روش‌هایی که هدف آن‌ها توصیف ویژگی‌ها است، به توصیف‌گر ویژگی معروف هستند. در مرحله نهایی، کتاب رمز برای تولید مدل کیف کلمات بصری تشکیل می‌شود.



شکل ۱: مراحل استخراج ویژگی و دسته‌بندی تصویر [۱۰۷]

این روش ابتدا برای پردازش زبان‌های طبیعی و بازیابی اطلاعات معرفی شده بود و از آن برای دسته‌بندی متن استفاده می‌شد. اکنون در حوزه دسته‌بندی تصاویر و یا شناسایی اشیاء نیز از این

روش استفاده می‌شود، حتی با استفاده از روش کیف کلمات می‌توان کلیپ‌های ویدئویی با موضوعات مختلف را نیز دسته‌بندی نمود.

در این روش متون به صورت مجموعه‌ای بدون ترتیب از کلمات در نظر گرفته می‌شوند. برای توصیف مدل کیف کلمات در دسته‌بندی متن فرض کنیم دو جمله الف و ب را داریم :

الف: مجید به تماشای فیلم علاقه دارد. مریم نیز همینطور.

ب: مجید به تماشای مسابقات فوتبال نیز علاقه دارد.

بر اساس جملات فوق فرهنگ لغات استخراج شده عبارت است از :

{ "مجید", "تماشای", "مسابقات", "فوتبال", "علاقه", "دارد", "مریم", "نیز", "به", "همینطور", "فیلم" }

که شامل ۱۱ لغت مجزا است. بر اساس این فرهنگ لغات جملات را با بردارهای ۱۱ عضوی به

صورت زیر در نظر می‌گیریم.

$$\{1,1,0,0,1,1,1,0,1,1,1\}$$
$$\{1,1,1,1,1,1,0,1,1,0,0\}$$

که هر عضو موجود در بردار نشان دهنده تعداد متناظر با آن لغت در فرهنگ کلمات^۱ است. به این

نوع نمایش نمودار میله‌ای^۲ اطلاق می‌شود. همان‌گونه که نشان داده شده، ترتیب کلمات در جملات،

تاثیری در هیستوگرام آن نخواهد داشت.

برای استفاده از تکنیک کیف کلمات در حوزه تصاویر، می‌توان کاراکترها و جملات را به صورتی

تعریف نمود که هر تصویر کیفی از لغات بصری در نظر گرفته شود، به نحوی که ترتیب کلمات در این

کیف حائز اهمیت نیست. این لغات می‌توانند خصوصیات یا ویژگی‌های بصری باشند. یکی از

محبوب‌ترین ویژگی‌هایی که مورد استفاده قرار می‌گیرد، ویژگی SIFT^۳ است.

^۱ Vocabularies Book

^۲ Histogram

^۳ Scale Invariant Feature Transform

۲- کلمات بصری^۱

برای تعریف کلمات بصری سه مرحله استخراج ویژگی، توصیف ویژگی‌ها و تولید کتاب رمز یا فرهنگ کلمات بصری بایستی انجام شود. در مرحله استخراج ویژگی، شناسایی نقاط یا محل‌های کلیدی به‌عنوان عناصر مهم تصویر (کلمات بصری)، انجام می‌شود. نقاط کلیدی^۲ حاوی اطلاعات محلی تصویر هستند و به صورت خودکار می‌توان آن‌ها را شناسایی نمود. پس از این‌که نقاط کلیدی در تصویر شناسایی گردید، باید بتوان این نقاط را توصیف نمود. هر تصویر مجموعه‌ای از چندین نقطه کلیدی است و توصیف یک تصویر به این معنی است که چگونه بتوان این نقاط را به‌وسیله بردارهای عددی نمایش داد. توصیف ویژگی‌ها به وسیله توصیف‌گر^۳ ویژگی انجام می‌شود. یک توصیف‌گر خوب باید بتواند چرخش، شدت، مقیاس و تغییرات نسبی را تا حدی پشتیبانی کند. یکی از مشهورترین توصیف‌گرهای ویژگی، [۱۰۸] SIFT است. در این روش هر نقطه کلیدی به یک بردار ۱۲۸ عضوی تبدیل می‌شود و هر تصویر توسط مجموعه‌ای از بردارهای با طول یکسان توصیف می‌شود (ترتیب این بردارها مهم نیست).

مرحله نهایی برای تولید مدل کیف لغات ساخت کتاب رمز است. به عبارت دیگر تبدیل نقاط کلیدی به کلمات رمز (مشابه لغات در متن) که باعث تولید کتاب رمز (مشابه فرهنگ کلمات) می‌شود. یک کلمه رمز شامل چندین نقطه کلیدی مشابه است. یک روش ساده برای انجام این کار استفاده از یکی از الگوریتم‌های خوشه‌بندی^۴ مانند K-MEAN برای خوشه‌بندی تمامی بردارها است، سپس مراکز خوشه به عنوان کلمات رمز تعریف می‌شوند. تعداد مراکز خوشه اندازه کتاب رمز را تشکیل می‌دهد. سپس هر نقطه کلیدی تصویر به یکی از خوشه‌ها تخصیص می‌یابد و پس از این مرحله می‌توان هر تصویر را به‌وسیله هیستوگرام‌های کلمات رمز نشان داد. بدین ترتیب می‌توان روش کیف

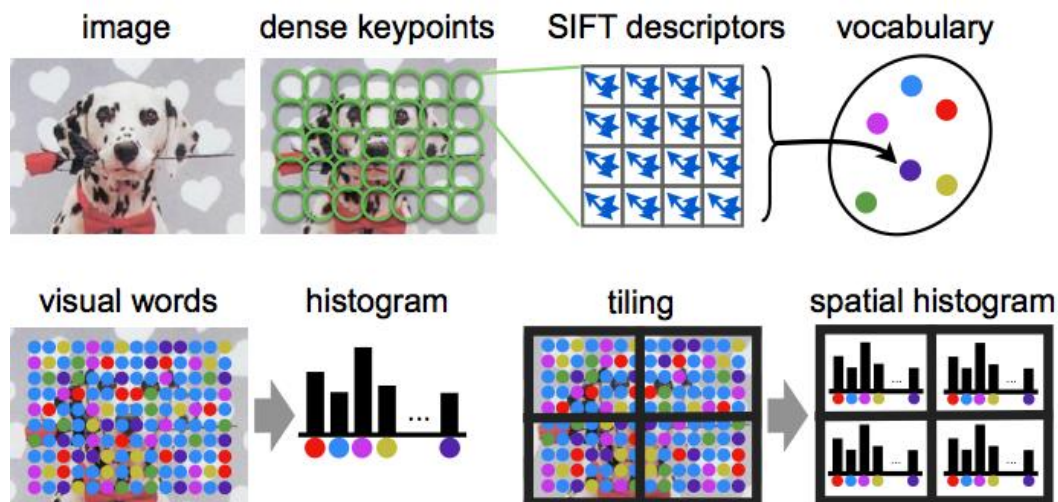
^۱ Visual Words

^۲ Key Points

^۳ Descriptor

^۴ Clustering

کلمات را برای استفاده در حوزه تصویر به کار برد.



شکل ۲: مراحل پیاده‌سازی روش کیف کلمات [۱۰۷]

الف- تعیین نقاط کلیدی

اولین مرحله از استخراج ویژگی عبارت است از شناسایی نقاط یا محل‌های کلیدی در تصویر که برای در نظر گرفتن به عنوان عناصر اصلی تصویر (لغات تصویری) مناسب هستند. این مرحله شامل اقداماتی است که در طی آن ویژگی‌های مفید برای شناسایی تصاویر از آن‌ها استخراج می‌گردد که در واقع این ویژگی‌ها، تصویر را ارائه می‌دهند. دو مثال از روش‌های یافتن نقاط کلیدی، آشکارساز نقطه گوشه Harris و آشکارساز دیفرانسل گوسین^۱ می‌باشند [۱۰۹].

• آشکارساز Harris

گوشه‌ها نقاطی از تصویر هستند که تغییرات درجات خاکستری در همسایگی آن‌ها شدید است و بنابراین به خوبی از نقاط همسایه خود متمایز هستند. مطابق با تحقیقات صورت پذیرفته در سال ۲۰۰۰، اپراتور استخراج گوشه [۱۱۰] Harris بالاترین تکرارپذیری و تمایز را در مقایسه با دیگر اپراتورهای استخراج گوشه دارد. در این الگوریتم به منظور استخراج گوشه‌ها، برای هر یک از

^۱ Difference of Gaussian

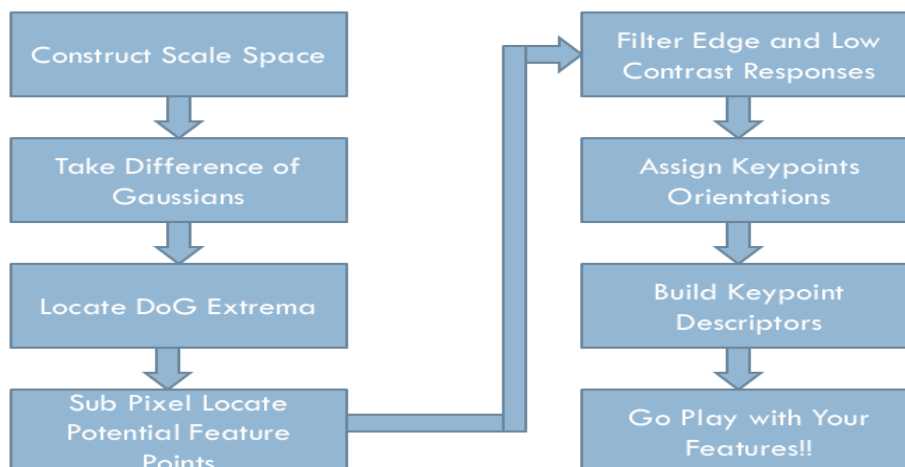
پیکسل‌های تصویر با استفاده از درجات خاکستری آن‌ها، معیاری برآورد می‌شود که نشانه‌ای از وجود یک گوشه است. در حالت استاندارد، نقاطی از تصویر که این معیار در همسایگی ۸ تایی آن‌ها دارای یک ماکزیمم محلی است و این ماکزیمم از یک مقدار آستانه بیشتر است، به‌عنوان گوشه استخراج می‌شود.

ب- توصیف ویژگی از تصاویر

برای توصیف نقاط کلیدی تصاویر معمولاً از توصیف‌گر [۱۰۸] SIFT استفاده می‌شود که کارایی آن در زمینه دسته‌بندی تصاویر و تشخیص اشیاء به‌خوبی اثبات گشته است. پس از یافتن نقاط کلیدی تصویر، با استفاده از روش SIFT، ویژگی‌های SIFT به وسیله بردارهایی با طول ۱۲۸ توصیف می‌گردد.

• الگوریتم SIFT

این الگوریتم یک روش تناظریابی مبتنی بر ویژگی است که در سال ۲۰۰۴ به‌منظور انجام فرآیند تشخیص الگو در تصاویر مرئی ارائه شد و از دو مرحله اصلی شامل استخراج ویژگی و ایجاد توصیف‌گر، تشکیل شده است. ویژگی‌های به‌دست آمده از تصویر توسط این الگوریتم نسبت به تغییرات در مقیاس تصویر و چرخش آن پایدار و نسبت به تغییرات در زاویه دید و شدت نور تا اندازه زیادی مقاوم هستند. نام تبدیل مقیاس غیروابسته به ویژگی به این دلیل است که الگوریتم داده‌های تصویر را به ویژگی‌های محلی تبدیل می‌کند که نسبت به تغییرات مقیاس، چرخش، زاویه دید و شدت نور مقاوم هستند. می‌توان مراحل الگوریتم SIFT را به صورت شکل ۳ نمایش داد.



شکل ۳: روندنمای الگوریتم SIFT [۱۰۸]

ج- تولید فرهنگ کلمات بصری

پس از استخراج بردارهای ویژگی SIFT، با استفاده از الگوریتم‌های خوشه‌بندی مانند K-MEAN این بردارها خوشه‌بندی می‌شوند. پس از خوشه‌بندی، مراکز خوشه که همان لغات تصویری هستند، مشخص شده و باید برای هر تصویر معین گردد که توصیف‌گرهای موجود در هر تصویر مربوط به کدام یک از مراکز خوشه هستند. پس از اتمام این مرحله برای هر تصویر یک هیستوگرام کیف لغات خواهیم داشت. این هیستوگرام مشخص کننده میزان موجود از هر مرکز خوشه در هر یک از تصاویر است، در واقع هر تصویر مانند یک متن در نظر گرفته شده و مراکز خوشه یا لغات تصویری به عنوان کلمات موجود در متن در نظر گرفته می‌شوند.

• خوشه‌بندی K-Mean

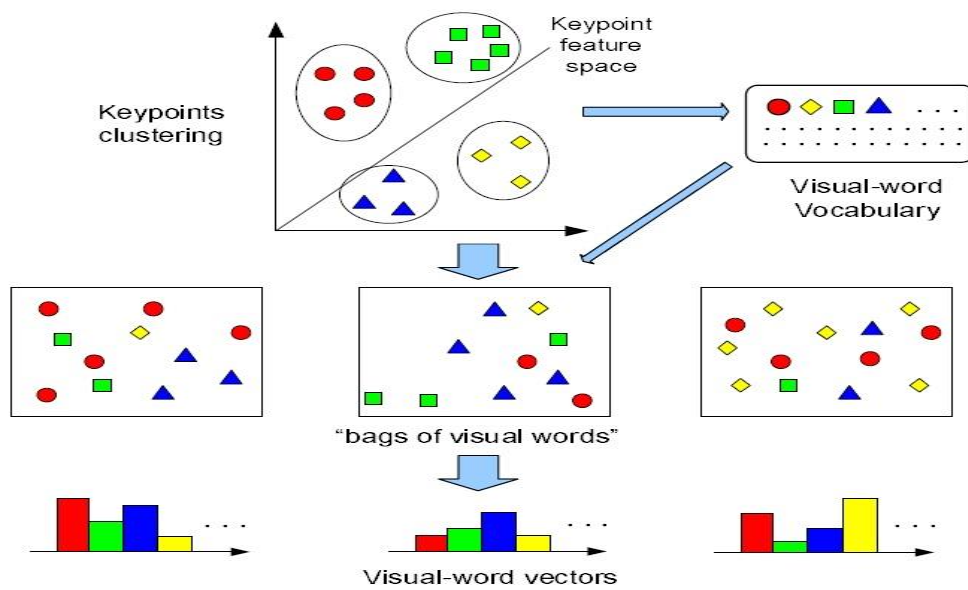
خوشه‌بندی به عملیات تقسیم داده‌های ناهمگن به تعدادی از زیر مجموعه‌ها یا گروه‌های همگن گفته می‌شود. در واقع خوشه‌بندی جهت گروه‌بندی راهی برای یافتن ساختار داده‌های پیچیده فراهم می‌نماید. الگوریتم K-Mean [۱۱۱] یکی از پرکاربردترین الگوریتم‌های خوشه‌بندی است. حرف K در اسم این الگوریتم نشان دهنده هدف این الگوریتم است که بیان گر پیدا کردن تعداد ثابتی از خوشه‌ها بر اساس نزدیکی نقاط داده به هم است. این روش علی‌رغم سادگی آن یک روش پایه برای بسیاری از روش‌های

خوشه‌بندی دیگر مانند خوشه‌بندی فازی محسوب می‌شود. برای این الگوریتم شکل‌های مختلفی بیان شده است ولی همه آن‌ها دارای روالی تکراری هستند. در نوع ساده‌ای از این روش ابتدا به تعداد خوشه‌های مورد نیاز نقاطی به صورت تصادفی انتخاب می‌شود. سپس داده‌ها با توجه به میزان نزدیکی یا شباهت به یکی از این خوشه‌ها نسبت داده می‌شوند و بدین ترتیب خوشه‌های جدیدی حاصل می‌شود. با تکرار همین روال می‌توان در هر تکرار با میانگین‌گیری از داده‌ها، مراکز جدیدی برای آن‌ها محاسبه کرد و مجدداً داده‌ها را به خوشه‌های جدید نسبت داد. این روند تا زمانی ادامه پیدا می‌کند که دیگر تغییری در داده‌ها حاصل نشود. تابع ۱ به عنوان تابع هدف مطرح است.

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (1)$$

که $\| \cdot \|$ معیار فاصله بین نقاط و c_j مرکز خوشه j ام است. مراحل چهارگانه که در ادامه بیان شده‌اند، مراحل پایه برای این روش محسوب می‌شود :

- I) در ابتدا K نقطه به عنوان نقاط مراکز خوشه‌ها انتخاب می‌شوند.
 - II) هر نمونه داده به خوشه‌ای که مرکز آن خوشه کمترین فاصله تا آن داده را دارد، نسبت داده می‌شود.
 - III) پس از تعلق تمام داده‌ها به یکی از خوشه‌ها، برای هر خوشه با میانگین‌گیری از نقاط متعلق به آن خوشه، یک نقطه جدید به عنوان مرکز محاسبه می‌گردد.
 - V) مراحل ۲ و ۳ تکرار می‌شوند تا زمانی که دیگر هیچ تغییری در مراکز خوشه‌ها حاصل نشود.
- تصویر شکل ۴، خوشه‌بندی لغات بصری را نشان می‌دهد.



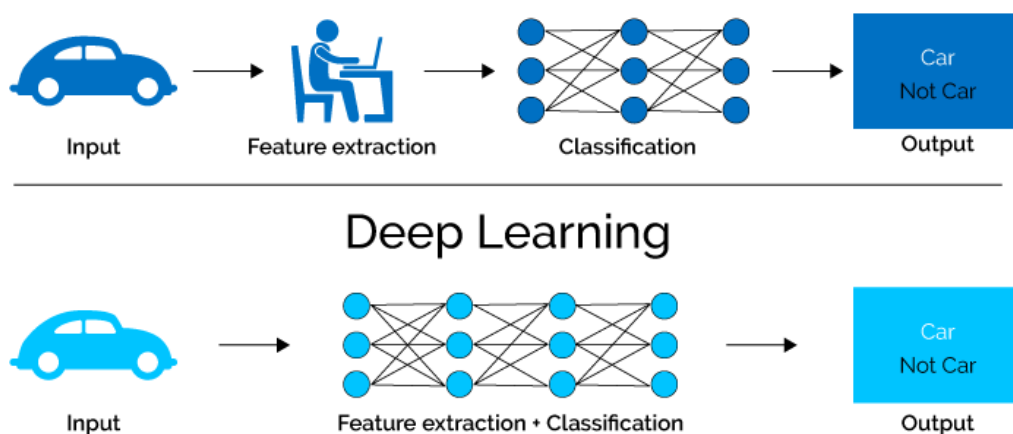
شکل ۴: خوشه‌بندی لغات بصری [۱۰۷]

پوست ب: روش یادگیری عمیق

۱- مقدمه

امروزه کاربرد شبکه‌های عصبی مصنوعی بسیار گسترده است. شبکه‌های عصبی مصنوعی، شبیه‌سازی فعالیت و پروسه‌ی یادگیری مغز انسان بر روی کامپیوتر هستند. شبکه‌های عصبی یکی از راه‌های هوشمندسازی کامپیوترها و به عبارتی یادگیری و بینایی ماشین هستند. در ادامه مقداری در مورد این شبکه‌ها بحث می‌شود، مطالب بیان شده در این بخش برگرفته از مرجع [۱۱۲] هستند.

در شبکه‌های عصبی مصنوعی غیر عمیق، ورودی شبکه، از نوع بردار است. یعنی باید ابتدا روی داده‌ها تغییراتی اعمال شود و ویژگی‌هایی استخراج گردد و نهایتاً این ویژگی‌ها به صورت بردار به شبکه عصبی اعمال شود. این فرآیند با فرآیند یادگیری مغز انسان مقداری تفاوت دارد، چون مغز داده را بدون تغییر و تحلیل می‌گیرد و فرآیند یادگیری یا آزمایش را تکمیل می‌کند. از این‌رو اعمال جایگزینی برای شبکه‌های عصبی مصنوعی موجود مورد مطالعه و تحقیق قرار گرفت و نتیجه‌ی آن، شبکه‌های عصبی عمیق است. شبکه‌هایی که از فرآیند یادگیری عمیق بهره می‌برند. هرچند ایده‌ی اولیه‌ی یادگیری عمیق سال‌ها پیش ارائه شد، ولی به دلیل ضعیف بودن سیستم‌های سخت‌افزاری موجود، امکان پیاده‌سازی وجود نداشت. با پیشرفت‌های صورت گرفته در زمینه‌ی سخت‌افزاری، در حال حاضر پیاده‌سازی این الگوریتم‌ها محقق شده است. تصویر شکل ۱ ساختار الگوریتم‌های عمیق را در مقایسه با سایر الگوریتم‌های بازشناسی که تاکنون ارائه شده‌اند، نشان می‌دهد.



شکل ۱: مقایسه ساختار یادگیری عمیق با ساختارهای قبلی بازشناسی [۱۱۲]

الگوریتم‌های یادگیری عمیق در دسته‌ی الگوریتم‌های یادگیری ماشین هستند. طی سال‌های اخیر، یادگیری عمیق در حوزه بینایی ماشین به‌صورت گسترده مورد مطالعه قرار گرفته است و به همین دلیل تعداد زیادی از روش‌های مرتبط با آن بوجود آمده است. این روش‌ها را بر اساس روش پایه‌ای که از آن مشتق شده‌اند می‌توان به ۴ دسته ۱. Convolutional Neural Networks، ۲. Restricted Boltzmann Machines، ۳. Autoencoders و ۴. Sparse Coding تقسیم‌بندی نمود. در میان این ۴ گروه، گروه اول از همه مهم‌تر است و بیشترین کاربرد را دارد، از این‌رو در ادامه این گروه مورد بررسی قرار می‌گیرد.

۲- شبکه‌های عصبی کانولوشنی^۱

شبکه‌های عصبی کانولوشن (CNN) یکی از مهم‌ترین روش‌های یادگیری عمیق هستند که در آن‌ها چندین لایه با روشی قدرتمند آموزش می‌بینند. این روش بسیار کارآمد بوده و یکی از رایج‌ترین روش‌ها در کاربردهای مختلف بینایی کامپیوتر است. به‌طور کلی، یک شبکه CNN از سه لایه اصلی تشکیل می‌شود که عبارتند از: لایه کانولوشن، لایه Pooling و لایه تماماً متصل^۲. لایه‌های مختلف وظایف مختلفی را انجام می‌دهد. در هر شبکه عصبی کانولوشن دو مرحله برای آموزش وجود دارد. مرحله Feed forward و مرحله پس انتشار^۳. در مرحله اول تصویر ورودی به شبکه تغذیه می‌شود و این عمل چیزی جز ضرب نقطه‌ای بین ورودی و پارامترهای هر نورون و نهایتاً اعمال عملیات کانولوشن در هر لایه نیست؛ سپس خروجی شبکه محاسبه می‌شود. در این‌جا به‌منظور تنظیم پارامترهای شبکه و یا به عبارت دیگر همان آموزش شبکه، از نتیجه خروجی جهت محاسبه میزان خطای شبکه استفاده می‌شود. برای این کار خروجی شبکه با استفاده از یک تابع خطا^۴ با پاسخ صحیح مقایسه می‌شود و مقدار خطا به‌دست می‌آید. سپس مرحله پس انتشار آغاز می‌شود. در این مرحله مشتق هر پارامتر با توجه به قاعده

^۱ Convolutional Neural Networks

^۲ Fully Connected Layer

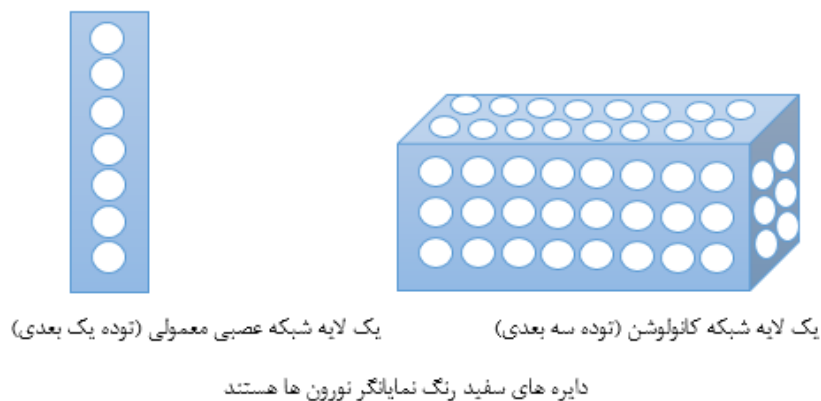
^۳ Back Propagation

^۴ Loss Function

Chain rule محاسبه شده و تمامی پارامترها با توجه به تاثیری که بر خطای ایجاد شده در شبکه دارند، تغییر می‌کنند. این مراحل ادامه می‌یابد و پس از تعداد مناسبی از تکرار، مراحل آموزش شبکه تمام می‌شود. در ادامه به‌طور مختصر نقش هر کدام از لایه‌های شبکه CNN بررسی می‌شود.

الف- لایه کانولوشن در شبکه‌های CNN

لایه کانولوشن هسته اصلی تشکیل دهنده شبکه عصبی CNN است و ساختار آن را می‌توان به‌صورت یک ساختار سه بعدی از نورون‌ها تفسیر کرد. در شبکه‌های عصبی معمولی هر لایه شبیه یک لیست (همانند یک مستطیل) از نورون‌ها است که هر نورون خروجی خاص خود را تولید می‌کند و نهایتاً یک لیست از خروجی‌ها که متناظر با هر نورون است، حاصل می‌شود. اما در شبکه عصبی کانولوشن بجای یک لیست ساده با یک لیست سه بعدی (یک مکعب) مواجه هستیم که نورون‌ها در سه بعد آن مرتب شده‌اند. در نتیجه خروجی این مکعب نیز یک توده سه بعدی خواهد بود. تصاویر شکل ۲ این مفاهیم را بهتر بیان می‌کند.

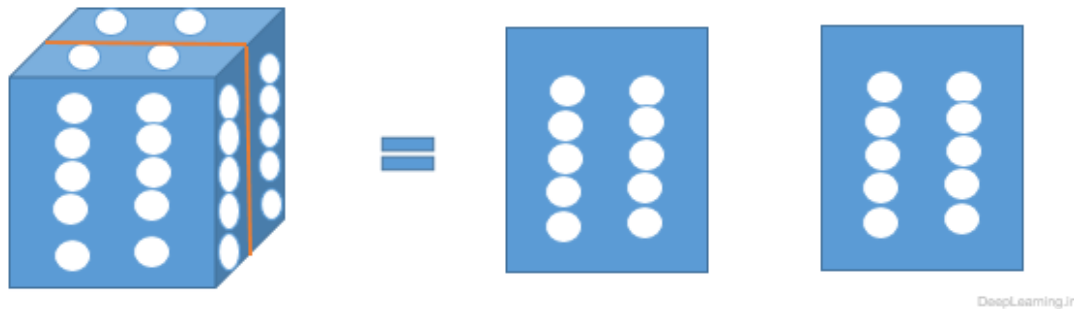


DeepLearning.ir

شکل ۲: ساختار شبکه کانولوشنی در مقایسه با شبکه عصبی معمولی [۱۱۲]

پارامترهای لایه کانولوشن شامل مجموعه‌ای از فیلترهای قابل یادگیری هستند. هر فیلتر از لحاظ مکانی کوچک بوده اما در امتداد عمق توده ورودی ادامه پیدا می‌کند. به زبان ساده‌تر می‌توان این‌طور گفت که یک توده سه بعدی وجود دارد. این توده‌ی سه بعدی دارای یک طول و عرض و یک عمق است. اگر فرض شود که عمق برابر با x باشد، به این معنی است که x برش (قاج) از توده وجود دارد. به‌عنوان

مثال اگر عمق برابر با ۱ باشد، یک ماتریس ساده (آرایه دوبعدی) وجود دارد که دارای طول و عرض است. حال اگر عمق برابر با ۲ باشد یعنی دو ماتریس وجود دارد (در هر عمق یک ماتریس)، شکل ۳ بیان گر این موضوع است.

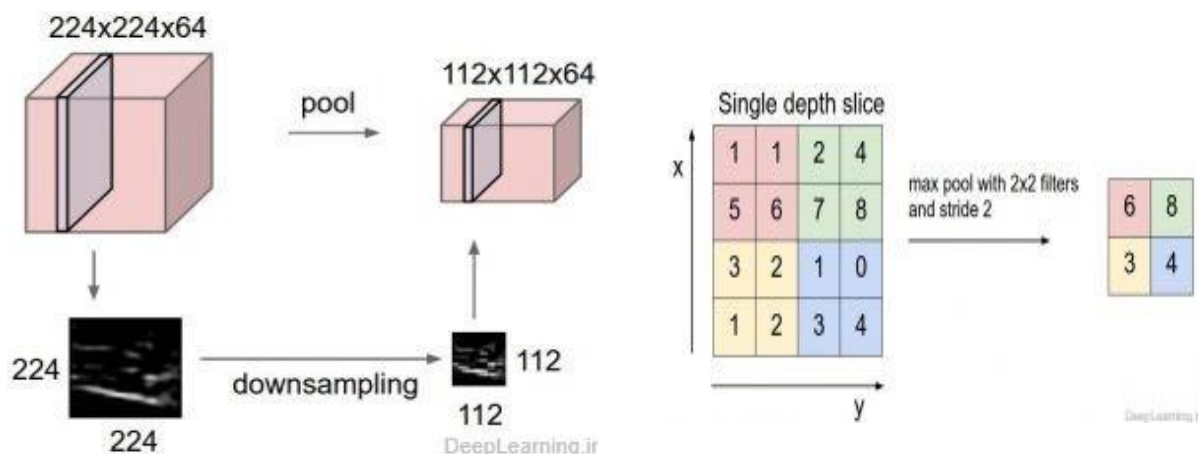


شکل ۳: برش شبکه عصبی عمیق در راستای عمق [۱۱۲]

ب- لایه Pooling

قرار دادن یک لایه Pooling بین چندین لایه کانولوشنی پشت سر هم در یک معماری کانولوشن امری رایج است. کارکرد این لایه کاهش ابعاد تصویر ورودی، جهت کاهش تعداد پارامترها و محاسبات در داخل شبکه و کنترل overfitting است. لایه Pooling به صورت مستقل بر روی هر برش عمقی از توده ورودی عمل کرده و آن را با استفاده از عملیات بیشینه گیری^۱ از لحاظ مکانی تغییر اندازه می دهد. رایج ترین شکل استفاده از این لایه به صورت استفاده این لایه با فیلترهایی با اندازه 2×2 به همراه stride (گام ۲) است که هر برش عمقی در ورودی را با حذف ۲ عنصر از عرض و ۲ عنصر از ارتفاع کاهش داده و باعث حذف ۷۵٪ مقادیر موجود در آن برش عمقی می شود (لازم به ذکر است بُعد عمق بدون تغییر باقی می ماند). تصویر شکل ۴ نحوه فرایند MAX Pooling را نشان می دهد.

^۱ MAX



شکل ۴: عملیات pooling

ج- لایه تماماً متصل

نورون‌هایی که در یک لایه تماماً متصل قرار دارند با تمام نورون‌های موجود در لایه قبلی ارتباط دارند، دقیقاً مانند همان چیزی که در شبکه‌های عصبی معمولی دیده می‌شود. بنابراین می‌توان از ضرب ماتریسی و سپس جمع نتیجه حاصله با بایاس، جهت محاسبه خروجی تمامی نورون‌ها در یک لحظه، استفاده کرد. به‌طور خلاصه تمامی قوانین مطرحی در شبکه‌های عصبی معمولی در این بخش صادق است.

پوستج: شیت کردن تصویر

۱- مقدمه

معمولاً در پردازش تصویر نیاز است تا اطلاعات دو یا چند تصویر را با یکدیگر ترکیب نمود. برای این منظور نیاز است تا یک رابطه منطقی بین نقاط این تصاویر پیدا کرد. هدف از تثبیت تصاویر^۱، یافتن تبدیل هندسی بین دو یا چند تصویر است. در حالت کلی باید یک تناظری یکتا بین یک نقطه از تصویر و نقطه‌ای دیگر از تصویر ثانویه به نحوی پیدا نمود که هر دو نشان دهنده یک نقطه از صحنه باشند. تثبیت تصاویر یکی از مسائل مشهور در حوزه پردازش تصویر و بینایی ماشین است و روش‌های زیادی برای آن ارائه شده است. استفاده از تثبیت تصاویر ما را قادر به مقایسه و ادغام داده‌های به دست آمده از مقیاس‌ها و زاویه‌های مختلف می‌سازد.

۲- تثبیت تصاویر با استفاده از نقاط ویژگی

یکی از پرکاربردترین و عملی‌ترین روش‌های تثبیت تصاویر، استفاده از نقاط ویژگی است که این روش خود پنج مرحله دارد [۱۱۳]:

(۱) استخراج نقاط دارای ویژگی‌های محلی

(۲) ایجاد توصیف متناسب با هر نقطه

(۳) انطباق توصیفات و به دست آوردن نقاط متناظر

(۴) محاسبه ماتریس تبدیل

(۵) تثبیت تصویر و درون‌یابی پیکسل‌ها

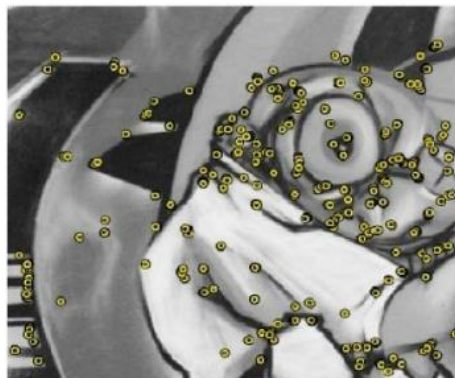
طبق مراحل فوق ابتدا با استفاده از روش‌های موجود نقاط دارای ویژگی استخراج می‌شود. سپس با استفاده از داده‌های موجود در اطراف نقطه مشخص شده در مرحله قبل، نقطه مورد نظر توصیف می‌شود و این توصیف با توصیفات نقاط تصویر دیگر مقایسه شده تا تناظری بین نقاط تصویر اول با تصویر ثانویه

^۱ Image Registration

ایجاد گردد. حال با استفاده از این نقاط متناظر می‌توان تخمینی مناسب از تابع تبدیل به‌دست آورد و به کمک این تابع تبدیل، تصویر اولیه را به تصویر ثانویه منطبق کرد. در ادامه در مورد این مراحل به اختصار بحث می‌شود.

الف - ویژگی‌های محلی

به طور کلی می‌توان گفت که یک ویژگی محلی شامل یک الگوی تصویری است که نسبت به همسایه‌های کناری خود متفاوت است. معمولاً این تفاوت‌ها شامل، شدت روشنایی، رنگ و بافت است [۱۱۴]. در شکل ۱ مثال‌هایی از نقاط دارای ویژگی‌های محلی در یک تصویر خاکستری آمده است. در این تصویر از آشکارساز گوشه هریس به‌عنوان یک آشکارساز ویژگی محلی استفاده شده است.



شکل ۱: اهمیت گوشه‌ها، یک تصویر با نقاط ویژگی که با الگوریتم آشکارساز گوشه به‌دست آمده است [۱۱۴]

● دسته‌بندی آشکارسازهای نقاط دارای ویژگی‌های محلی

در یک دسته‌بندی کلی روش‌های آشکارسازی نقاط ویژگی عبارتند از [۱۱۴]:

(۱) روش‌های آشکارسازی مبتنی بر انحنای کانتور

(۲) روش‌های آشکارسازی مبتنی بر شدت روشنایی

(۳) آشکارسازی مبتنی بر رخدادهای مکانی

(۴) روش‌های آشکارسازی مبتنی بر رنگ

۵) روش‌های آشکارسازی مبتنی بر مدل

۶) روش‌های مقاوم نسبت به نقطه دید

۷) روش‌های مبتنی بر بخش‌بندی

۸) روش‌های مبتنی بر یادگیری ماشین

● توصیف‌گر نقاط ویژگی

گام مهم بعدی پس از استخراج نقاط ویژگی در تثبیت تصاویر، توصیف نقاط ویژگی است. هدف از این بخش به‌دست آوردن یک توصیف کامل و جامع از نقاط دارای ویژگی است به‌نحوی که بتوان معیارهای شباهت را برای آن به‌کار برد. این توصیف‌گرها باید اطلاعاتی را در مورد شکل و بافت ساختار محلی در اختیار قرار دهد. محتوای اطلاعات و تغییرناپذیری آن دو ویژگی مهم توصیف‌گرهای محلی هستند. این دو ویژگی تمایز و نیرومندی توصیف‌گر را تعیین می‌کند. در حالت کلی بیشتر توصیف‌گرها نسبت به اطلاعاتی که حمل می‌کنند تغییرناپذیر هستند. مقدار این اطلاعات متناسب با الگوریتم مورد استفاده برای آشکارسازی است، زیرا این نقاط با توجه به تغییرات سیگنال تصویر به‌دست آمده است.

ب- الگوریتم‌های تطبیق نقاط مورد علاقه

همان‌گونه که قبلاً نیز اشاره شد، آشکارسازی نقاط مورد علاقه اولین مرحله در تثبیت تصاویر است. بعد از آن نیاز به محاسبه توصیف‌گر متناسب با هر ویژگی است. حال برای به‌دست آوردن تناسب نظیر به نظیر این نقاط ویژگی باید از معیارهای اندازه‌گیری و مقایسه استفاده نمود. این معیارها در دوره‌های زمانی متفاوت متناسب با توصیف‌گرها تغییر کرده است. در ابتدا برای تطابق از تفاضل‌گیری پیکسل‌های اطراف نقطه مورد نظر برای سنجش میزان شباهت استفاده می‌شد. امروزه بیشتر از معیارهایی مانند فاصله ماهالانوبیس و فاصله همینگ است، استفاده می‌شود [۱۱۱].

ج- محاسبه ماتریس تبدیل

در علم بینایی ماشین هر دو تصویر، توسط یک ماتریس تبدیل به یک‌دیگر مرتبط است. چرخش و

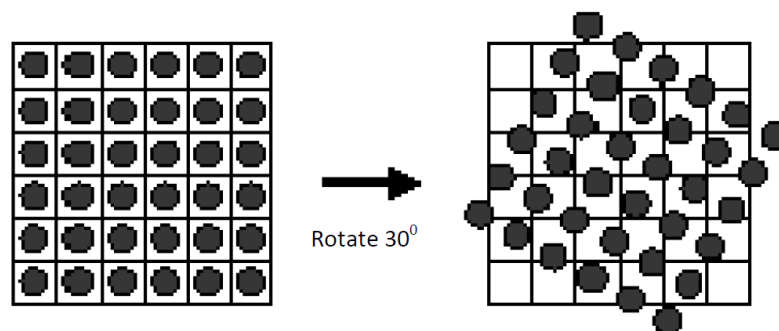
انتقال و حتی تغییر مقیاس یک دوربین را می‌توان با استفاده از این ماتریس نشان داد. از لحاظ ریاضی در فضای دو بعدی این ماتریس دارای ابعاد 3×3 بوده که به صورت رابطه ۱ تعریف می‌گردد:

$$\begin{pmatrix} h11 & h12 & h13 \\ h21 & h22 & h23 \\ h31 & h32 & h33 \end{pmatrix} \quad (1)$$

در عمل پارامتر $h33$ را یک در نظر می‌گیرند، لذا برای محاسبه ماتریس فوق حداقل ۴ نقطه مورد نیاز است، زیرا نقاط دارای دو بعد هستند که ۸ معادله را تولید می‌کند. از این ۸ معادله می‌توان ۸ مجهول را محاسبه نمود. روش‌های مورد استفاده برای حل این معادله عبارتند از حداقل میانگین مربعات خطا، الگوریتم اجماع نمونه‌های تصادفی [۱۱۱].

د- انطباق تصویر و درون‌یابی

پس از این که تابع تبدیل مورد نظر محاسبه گردید، باید تصویر اولیه را به تصویر ثانویه انتقال داد. اما ممکن است این انتقالات باعث شود پیکسل‌ها در نقاطی نگاشت شود که دارای مختصات صحیح نیست. تصویر ۲ این موضوع را نشان می‌دهد.



شکل ۲: مکان جدید پیکسل‌ها در اثر چرخش تصویر [۱۱۳]

برای حل این مشکل، ابتدایی‌ترین راه استفاده از روش نزدیک‌ترین همسایه است که در مختصات صحیح، اطلاعات رنگ نزدیک‌ترین پیکسل نگاشت شده به این نقطه را جایگزین اطلاعات پیکسل کنونی می‌کند. این روش دارای قدرت مناسبی در تبدیلات بزرگ نیست و باعث خطای زیاد می‌گردد. روش

دیگری که بیشتر مورد استفاده قرار می‌گیرد، روش درون‌یابی خطی است. در این روش برای محاسبه مقدار یک پیکسل در یک مختصات صحیح از ترکیب خطی نقاط تصویر در اطراف آن پیکسل استفاده می‌شود.

۳- تثبیت تصویر دو خودرو

تصاویر شکل ۳ دو دستگاه خودروی پژو را نشان می‌دهند، همان‌طور که مشخص است این دو خودرو از نظر زاویه‌ی تصویربرداری و ظاهر با هم تفاوت‌هایی دارند. هدف این است که خودروی مشکی رنگ بر خودروی سفیدرنگ تثبیت شود.

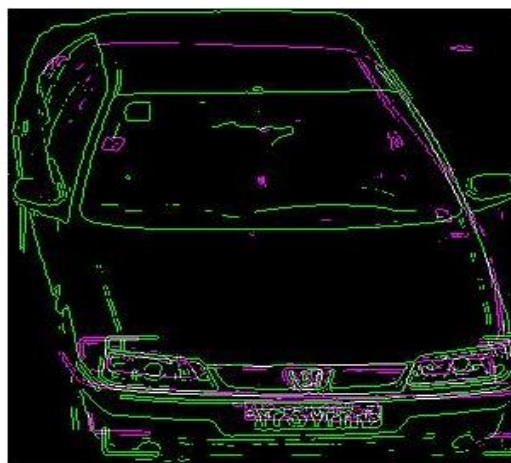


شکل ۳: خودروهای مورد نظر برای تثبیت

شکل ۴ تصاویر دو خودرو و لبه‌های آن‌ها را به‌طور هم‌زمان، قبل و بعد از تثبیت کردن نمایش می‌دهد. تصاویر سمت راست مربوط به قبل از تثبیت کردن و تصاویر سمت چپ مربوط به پس از تثبیت کردن است. بخش‌های سبز رنگ مربوط به خودروی ثابت، و بخش‌های صورتی رنگ مربوط به خودروی مد نظر برای تثبیت است. مقایسه تصاویر سمت راستی با سمت چپی، بیان‌گر انطباق بیشتر در تصاویر سمت چپی است.



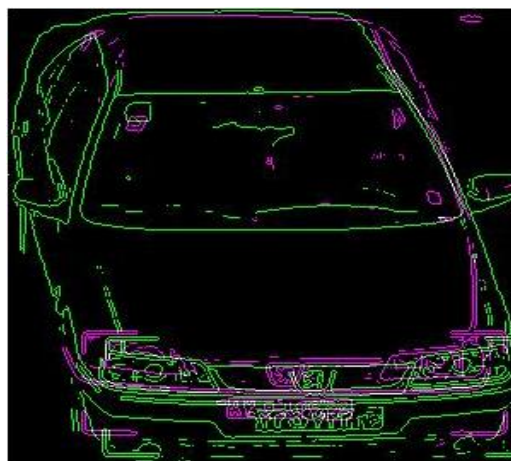
ب: بعد از تثبیت



الف: قبل از تثبیت



د: بعد از تثبیت



ج: قبل از تثبیت

شکل ۴: نتایج قبل و بعد از تثبیت

- [1] Dubská, M., Herout, A., Juránek, R., & Sochor, J. (2014). Fully automatic roadside camera calibration for traffic surveillance. *IEEE Transactions on Intelligent Transportation Systems*, 16(3), 1162-1171.
 - [2] Dawson, D. N., & Birchfield, S. T. (2013). An energy minimization approach to automatic traffic camera calibration. *IEEE Transactions on Intelligent Transportation Systems*, 14(3), 1095-1108.
 - [3] Schoepflin, T. N., & Dailey, D. J. (2003). Dynamic camera calibration of roadside traffic management cameras for vehicle speed estimation. *IEEE Transactions on Intelligent Transportation Systems*, 4(2), 90-98.
 - [4] Wang, K., Huang, H., Li, Y., & Wang, F. Y. (2007). Research on lane-marking line based camera calibration. In *IEEE International Conference on Vehicular Electronics and Safet*, (pp. 1-6).
 - [5] Pai, T. W., Juang, W. J., & Wang, L. J. (2001). An adaptive windowing prediction algorithm for vehicle speed estimation. In *Proceedings of IEEE Intelligent Transportation Systems*, (pp. 901-906).
- [۶] ارفع الرفیعی، سعید، (۱۳۹۵): "محاسبه‌ی تمام خودکار پارامترهای داخلی و خارجی دوربینهای کنار جادهای به منظور به‌دست آوردن ابعاد واقعی وسایل نقلیه"، پایان نامه کارشناسی ارشد. گروه الکترونیک دیجیتال. دانشکده مهندسی برق و رباتیک، دانشگاه صنعتی شاهرود.
- [7] He, H., Shao, Z., & Tan, J. (2015). Recognition of car makes and models from a single traffic-camera image. *IEEE Transactions on Intelligent Transportation Systems*, 16(6), 3182-3192.
 - [8] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. *Proceedings of the IEEE international conference on computer vision*, (pp. 2961-2969).
 - [9] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, (pp. 779-788).
 - [10] Bainbridge-Smith, A., & Lane, R. G. (1997). Determining optical flow using a differential method. *Image and Vision Computing*, 15(1), 11-22.
 - [11] Liu, F., & Picard, R. W. (1998). Finding Periodicity in Space and Time. In *ICCV*, (pp. 376-383).
 - [12] Ridder, C., Munkelt, O., & Kirchner, H. (1995). Adaptive background estimation and foreground detection using kalman-filtering. In *Proceedings of International Conference on recent Advances in Mechatronics*, (pp. 193-199).
 - [13] Wixson, L. (2000). Detecting salient motion by accumulating directionally-consistent flow. *IEEE transactions on pattern analysis and machine intelligence*, 22(8), 774-780.
 - [14] Fernandez-Caballero, A., Fernandez, M. A., Mira, J., & Delgado, A. E. (2003). Spatio-temporal shape building from image sequences using lateral interaction in accumulative computation. *Pattern Recognition*, 36(5), 1131-1142.
 - [15] Xie, L., et al. (2005). Robust vehicles extraction in a video-based intelligent transportation systems. In *Proceedings of International Conference on Communications, Circuits and Systems*, (vol. 2).
 - [16] Park, K., Lee, D., & Park, Y. (2007). Video-based detection of street-parking violation. In *IPCV*, (pp. 152-156).

- [17] Nguyen, P. V., & Le, H. B. (2008). A multi-modal particle filter based motorcycle tracking system. In Pacific Rim International Conference on Artificial Intelligence, (pp. 819-828).
- [18] Daigavane, M. P., & Bajaj, P. (2010). Real time vehicle detection and counting method for real time vehicle detection and counting method for unsupervised traffic video on highways unsupervised traffic video on highways. IJCSNS, 10(8), 112.
- [19] Ha, D. M., Lee, J. M., & Kim, Y. D. (2004). Neural-edge-based vehicle detection and traffic parameter extraction. Image and vision computing, 22(11), 899-907.
- [20] Chen, C. J., Chiu, C. C., Wu, B. F., Lin, S. P., & Huang, C. D. (2004). The moving object segmentation approach to vehicle extraction. In IEEE International Conference on Networking, Sensing and Control, (vol. 1, pp. 19-23)
- [21] Javed, O., & Shah, M. (2002). Tracking and object classification for automated surveillance. In European Conference on Computer Vision (pp. 343-357)
- [22] Chen, B., Lei, Y., & Li, W. (2004). A novel background model for real-time vehicle detection. In Proceedings 7th International Conference on Signal Processing, (vol. 2, pp. 1276-1279)
- [23] Sánchez, A., Nunes, E. O., & Conci, A. (2012). Using adaptive background subtraction into a multi-level model for traffic surveillance. Integrated Computer-Aided Engineering, 19(3), 239-256.
- [24] Gupte, S., Masoud, O., Martin, R. F., & Papanikolopoulos, N. P. (2002). Detection and classification of vehicles. IEEE Transactions on intelligent transportation systems, 3(1), 37-47.
- [25] Yen, J. C., Chang, F. J., & Chang, S. (1995). A new criterion for automatic multilevel thresholding. IEEE Transactions on Image Processing, 4(3), 370-378.
- [26] Chiu, C. C., Ku, M. Y., & Liang, L. W. (2009). A robust object segmentation system using a probability-based background extraction algorithm. IEEE Transactions on circuits and systems for video technology, 20(4), 518-528.
- [27] Stauffer, C., & Grimson, W. E. L. (2000). Learning patterns of activity using real-time tracking. IEEE Transactions on pattern analysis and machine intelligence, 22(8), 747-757.
- [28] Huerta, I., Amato, A., Roca, X., & González, J. (2013). Exploiting multiple cues in motion segmentation based on background subtraction. Neurocomputing, (vol.100, pp. 183-196).
- [29] Del Carmen, D. J. R., & Cajote, R. D. (2016). Moving shadow detection and removal for video-based traffic monitoring. In 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), (pp. 1-7)
- [30] Manikoth, N., Loce, R., Bernal, E., & Wu, W. (2012). Survey of computer vision in roadway transportation systems. In Visual Information Processing and Communication, (vol. 8305, p. 830).
- [31] Tian, B., et al. (2011). Video processing techniques for traffic flow monitoring: A survey. In 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), (pp. 1103-1108)
- [32] Cucchiara, R., Grana, C., Piccardi, M., & Prati, A. (2003). Detecting moving objects, ghosts, and shadows in video streams. IEEE transactions on pattern analysis and machine intelligence, 25(10), 1337-1342.

- [33] Martel-Brisson, N., & Zaccarin, A. (2007). Learning and removing cast shadows through a multidistribution approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(7), 1133-1146.
 - [34] Leone, A., Distant, C., & Buccolieri, F. (2005). A texture-based approach for shadow detection. In *IEEE Conference on Advanced Video and Signal Based Surveillance*, (pp. 371-376)
 - [35] Joshi, A. J., & Papanikolopoulos, N. P. (2008). Learning to detect moving shadows in dynamic environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(11), 2055-2063.
 - [36] Joshi, A. J., & Papanikolopoulos, N. (2008). Learning of moving cast shadows for dynamic environments. In *2008 IEEE International Conference on Robotics and Automation*, (pp. 987-992)
- [۳۷] معلم، پیمان، و عسگریان دهکردی، ر. (۱۳۹۶). "ردیابی تصویری سریع، مطمئن و مقاوم نسبت به انسداد با کمک یک مدل تقسیم‌شده مبتنی بر لبه"، نشریه مهندسی برق و مهندسی کامپیوتر ایران، ۱-ب، ۳۷-۴۵.

- [38] Lin, C. T., et al. (2010). An efficient and robust moving shadow removal algorithm and its applications in ITS. *EURASIP Journal on Advances in Signal Processing*, 20(1), 945130.
- [39] Li, Y., Li, Z., Tian, H., & Wang, Y. (2011). Vehicle detecting and shadow removing based on edged mixture Gaussian model. *IFAC Proceedings Volumes*, 44(1), 800-805.
- [40] Xiao, M., Han, C. Z., & Zhang, L. (2007). Moving shadow detection and removal for traffic sequences. *International Journal of Automation and Computing*, 4(1), 38-46.
- [41] Zhang, W., Fang, X. Z., Yang, X. K., & Wu, Q. J. (2007). Moving cast shadows detection using ratio edge. *IEEE Transactions on Multimedia*, 9(6), 1202-1214.
- [42] Chen, Z., & Huang, X. (2019). Pedestrian Detection for Autonomous Vehicle Using Multi-Spectral Cameras. *IEEE Transactions on Intelligent Vehicles*, 4(2), 211-219.
- [43] Beymer, D., McLauchlan, P., Coifman, B., & Malik, J. (1997). A real-time computer vision system for measuring traffic parameters. In *Proceedings of IEEE computer society conference on computer vision and pattern recognition*, (pp. 495-501).
- [44] Cathey, F. W., & Dailey, D. J. (2005). A novel technique to dynamically measure vehicle speed using uncalibrated roadway cameras. In *IEEE Proceedings. Intelligent Vehicles Symposium*, (pp. 777-782)
- [45] Sina, I., et al. (2013). Vehicle counting and speed measurement using headlight detection. In *2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, (pp. 149-154).
- [46] You, X., & Zheng, Y. (2016). An accurate and practical calibration method for roadside camera using two vanishing points. *Neurocomputing*, (vol.204, pp. 222-230).
- [47] Hsieh, J. W., et al. (2006). Automatic traffic surveillance system for vehicle tracking and classification. *IEEE Transactions on Intelligent Transportation Systems*, 7(2), 175-187.

- [۴۸] آقایان، امیر. ، و خسروی، حسین. (۱۳۹۶). "تخمین بلادرنگ سرعت خودرو از طریق دوربین به کمک ردیابی مرکز ثقل و پیاده‌سازی آن روی برد توسعه‌ی XU4" دهمین کنفرانس بینایی ماشین و پردازش تصویر ایران، اصفهان، انجمن ماشین بینایی و پردازش تصویر ایران، دانشگاه صنعتی اصفهان.
- [49] Luvizon, D. C., Nassu, B. T., & Minetto, R. (2014). Vehicle speed estimation by license plate detection and tracking. In 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), (pp. 6563-6567)
- [50] Tomasi, C., & Detection, T. K. (1991). Tracking of point features. Tech. Rep. CMU-CS-91-132, Carnegie Mellon University
- [51] Sochor, J., et al. (2019). "Comprehensive Data Set for Automatic Single Camera Visual Speed Measurement. IEEE Transactions on Intelligent Transportation Systems, 20 (5), 1633-1643.
- [52] Filipiak, P., Golenko, B., & Dolega, C. (2016). NSGA-II based auto-calibration of automatic number plate recognition camera for vehicle speed measurement. In European Conference on the Applications of Evolutionary Computation, (pp. 803-818).
- [53] Eslami, H., Raie, A., & Faez, K. (2016). Precise vehicle speed measurement for law enforcement applications based on calibrated camera with parallel standard patterns. IET Computer Vision, 10(5), 398-406.
- [۵۴] اسلامی، حامد، و راعی، ابوالقاسم. ، فائز، کریم. (۱۳۹۳). "استفاده از اطلاعات پلاک خودرو جهت کالیبراسیون دوربین های ثبت تخلف و مکانیابی خودرو در صحنه به صورت مقاوم"، دومین کنفرانس بین المللی بازشناسی الگو و تحلیل تصویر، رشت، انجمن ماشین بینایی و پردازش تصویر ایران، گیلان.
- [55] Famouri, M., Azimifar, Z., & Wong, A. (2019). A novel motion plane-based approach to vehicle speed estimation. IEEE Transactions on Intelligent Transportation Systems, 20(4), 1237-1246.
- [56] Dailey, D. J., Cathey, F. W., & Pumrin, S. (2000). An algorithm to estimate mean traffic speed using uncalibrated cameras. IEEE transactions on intelligent transportation systems, 1(2), 98-107.
- [۵۷] حیدری، وحید، و احمد زاده ، محمد رضا.(۱۳۹۱) "تشخیص و دسته‌بندی وسایل نقلیه بر اساس پردازش تصویر برای یک سیستم نظارتی ترافیک"، پایان‌نامه کارشناسی ارشد.گروه مخابرات.دانشکده مهندسی برق.دانشگاه صنعتی اصفهان.
- [۵۸] حنیفه لو، زینب، و منجمی، سید امیر حسن.(۱۳۹۲). "آشکارسازی و ردیابی خودروهای متخلف عبوری از چراغ قرمز به کمک روش‌های پیشرفته پردازش تصویر" پایان نامه کارشناسی ارشد.گروه هوش مصنوعی.دانشکده مهندسی کامپیوتر .دانشگاه اصفهان.
- [59] Sochor, J. (2014). Traffic analysis from video. Diplomová práce, Brno University of Technology, Faculty of Information Technology .
- [60] Dubská, M., Herout, A., & Sochor, J. (2014). Automatic Camera Calibration for Traffic Understanding. In BMVC, 4(6), 8.
- [61] Chen, Z., et al. (2019). Multi-view vehicle type recognition with feedback-enhancement multi-branch CNNs. IEEE Transactions on Circuits and Systems for Video Technology, 29 (9), 2590-2599.

- [62] Hu, Q., Wang, H., Li, T., & Shen, C. (2017). Deep CNNs with spatially weighted pooling for fine-grained car recognition. *IEEE Transactions on Intelligent Transportation Systems*, 18(11), 3147-3156.
- [63] Soon, F. C., Khaw, H. Y., Chuah, J. H., & Kanesan, J. (2019). PCANet-based convolutional neural network architecture for a vehicle model recognition system. *IEEE Transactions on Intelligent Transportation Systems*, 20(2), 749-759.
- [64] Guindel, C., Martin, D., & Armingol, J. M. (2018). Fast joint object detection and viewpoint estimation for traffic scene understanding. *IEEE Intelligent Transportation Systems Magazine*, 10(4), 74-86.
- [65] Huang, Y., Wu, R., Sun, Y., Wang, W., & Ding, X. (2015). Vehicle logo recognition system based on convolutional neural networks with a pretraining strategy. *IEEE Transactions on Intelligent Transportation Systems*, 16(4), 1951-1960.
- [66] Psyllos, A. P., Anagnostopoulos, C. N. E., & Kayafas, E. (2010). Vehicle logo recognition using a sift-based enhanced matching scheme. *IEEE transactions on intelligent transportation systems*, 11(2), 322-328.
- [67] Wang, J., Zheng, H., Huang, Y., & Ding, X. (2018). Vehicle type recognition in surveillance images from labeled web-nature data using deep transfer learning. *IEEE Transactions on Intelligent Transportation Systems*, 19(9), 2913-2922.
- [68] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, (pp. 91-99).
- [69] Yang, L., Luo, P., Change Loy, C., & Tang, X. (2015). A large-scale car dataset for fine-grained categorization and verification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 3973-3981).
- [70] Yu, S., Wu, Y., Li, W., Song, Z., & Zeng, W. (2017). A model for fine-grained vehicle classification based on deep learning. *Neurocomputing*, (vol.257, pp. 97-103).
- [71] Sochor, J., Herout, A., & Havel, J. (2016). Boxcars: 3d boxes as cnn input for improved fine-grained vehicle recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 3006-3015).
- [72] Sochor, J., Špaňhel, J., & Herout, A. (2019). Boxcars: Improving fine-grained recognition of vehicles using 3-d bounding boxes in traffic surveillance. *IEEE Transactions on Intelligent Transportation Systems*, 20(1), 97-108.
- [73] Sochor, J., & Herout, A. (2015). Unsupervised Processing of Vehicle Appearance for Automatic Understanding in Traffic Surveillance. In *2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, (pp. 1-8).
- [74] Hsieh, J. W., Chen, L. C., & Chen, D. Y. (2014). Symmetrical SURF and its applications to vehicle detection and vehicle make and model recognition. *IEEE Transactions on intelligent transportation systems*, 15(1), 6-20.
- [75] Biglari, M., Soleimani, A., & Hassanpour, H. (2018). A cascaded part-based system for fine-grained vehicle classification. *IEEE Transactions on Intelligent Transportation Systems*, 19(1), 273-283.
- [76] Biglari, M., Soleimani, A., & Hassanpour, H. (2017). Part-based recognition of vehicle make and model. *IET Image Processing*, 11(7), 483-491.

[۷۷] بیگلری، محسن، و سلیمانی، علی، و حسن پور، حمید. (۱۳۹۶). "شناسایی نوع و مدل وسیله نقلیه با استخراج خودکار بخش‌های مشترک" مجله بینایی ماشین و پردازش تصویر ایران، ۱، ۲۹-۳۹.

- [78] Munroe, D. T., & Madden, M. G. (2005). Multi-class and single-class classification approaches to vehicle model recognition from images, (pp.1-11).
- [79] Huttunen, H., Yancheshmeh, F. S., & Chen, K. (2016). Car type recognition with deep neural networks. In 2016 IEEE Intelligent Vehicles Symposium, (pp. 1115-1120)
- [80] Boonsim, N., & Prakoonwit, S. (2017). Car make and model recognition under limited lighting conditions at night. *Pattern Analysis and Applications*, 20(4), 1195-1207.
- [81] Sarfraz, M. S., & Khan, M. H. (2011). A Probabilistic Framework for Patch based Vehicle Type Recognition. *Visapp*.
- [82] Dollár, P., Appel, R., Belongie, S., & Perona, P. (2014). Fast feature pyramids for object detection. *IEEE transactions on pattern analysis and machine intelligence*, 36(8), 1532-1545.
- [83] Cristani, M., Cheng, D. S., Murino, V., & Pannullo, D. (2004). Distilling information with super-resolution for video surveillance. In *Proceedings of the ACM 2nd international workshop on Video surveillance & sensor networks*, (pp. 2-11)
- [84] Lin, F. C., Fookes, C. B., Chandran, V., & Sridharan, S. (2005). Investigation into optical flow super-resolution for surveillance applications.
- [85] Irani, M., & Peleg, S. (1991). Improving resolution by image registration. *CVGIP: Graphical models and image processing*, 53(3), 231-239.
- [86] Farsiu, S., Robinson, M. D., Elad, M., & Milanfar, P. (2004). Fast and robust multiframe super resolution. *IEEE transactions on image processing*, 13(10), 1327-1344.
- [87] Li, X., & Orchard, M. T. (2001). New edge-directed interpolation. *IEEE transactions on image processing*, 10(10), 1521-1527.
- [88] Freeman, W. T., Jones, T. R., & Pasztor, E. C. (2002). Example-based super-resolution. *IEEE Computer graphics and Applications*, 22(2), 56-65.
- [89] Yang, J., Wright, J., Huang, T. S., & Ma, Y. (2010). Image super-resolution via sparse representation. *IEEE transactions on image processing*, 19(11), 2861-2873.
- [90] Rasti, P., et al (2017). A new low-complexity patch-based image super-resolution. *IET Computer Vision*, 11(7), 567-576.
- [91] Zeyde, R., Elad, M., & Protter, M. (2010). On single image scale-up using sparse-representations. In *International conference on curves and surfaces*, (pp. 711-730).
- [92] Glasner, D., Bagon, S., & Irani, M. (2009). Super-resolution from a single image. In *2009 IEEE 12th International Conference on Computer Vision*, (pp. 349-356)
- [93] Tsai, C. Y., et al. (2012). Context-aware single image super-resolution using locality-constrained group sparse representation. In *2012 Visual Communications and Image Processing*, (pp. 1-6)
- [94] Chen, C., & Fowler, J. E. (2012). Single-image super-resolution using multihypothesis prediction. In *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, (pp. 608-612)

- [95] Zhu, Z., Guo, F., Yu, H., & Chen, C. (2014). Fast single image super-resolution via self-example learning and sparse representation. *IEEE Transactions on Multimedia*, 16(8), 2178-2190.
- [96] Mokari, A., & Ahmadyfard, A. (2016). Fast single image SR via dictionary learning. *IET Image Processing*, 11(2), 135-144.
- [97] C.Daniilidis. (Accessed: September 2018).How to Compute Intrinsic from Vanishing Points :[video file] retrieved from: <https://www.coursera.org/lecture/robotics-perception/how-to-compute-intrinsic-from-vanishing-points-jnaLs>.
- [98] Zivkovic, Z. (2004). Improved adaptive Gaussian mixture model for background subtraction. In *ICPR*, (vol.2, pp. 28-31).
- [99] Yang, M. T., Lo, K. H., Chiang, C. C., & Tai, W. K. (2008). Moving cast shadow detection by exploiting multiple cues. *IET Image Processing*, 2(2), 95-104.
- [100] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *Computer Science - Computer Vision and Pattern Recognition*, arXiv preprint arXiv:1409.1556.
- [101] Chen, C., Cai, X., Zhao, Q., Lv, L., & Shu, H. (2017). Vehicle type recognition based on multi-branch and multi-layer features. In *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, (pp. 2038-2041).
- [102] Ding, Z., & Mo, W. (2016). Vehicle feature recognition based on the frontal lamp pattern descriptor. In *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, (pp. 375-379).
- [103] Elad, M. (2010). Sparse and redundant representations: from theory to applications in signal and image processing. Springer Science & Business Media.
- [۱۰۴] عسگریان دهکردی، رسول، و معلم، پیمان، و معمار مقدم، علیرضا. (۱۳۹۳) "الگوریتمی قطعه- مبنا برای ردیابی بلادرنگ تصویری با مقاومت مناسب نسبت به انسداد و تغییرات هدف"، دومین کنفرانس بین المللی بازشناسی الگو و تحلیل تصویر، رشت، انجمن ماشین بینایی و پردازش تصویر ایران، دانشگاه گیلان.
- [105] Xia, G.S., Bia, X., Diang, J., et al. (Accessed: 2018). a large-Scale Dataset for Object Detection in Aerial Image(DOTA). retrieved from: <https://captain-whu.github.io/DOTA/dataset.html>
- [106] Mousavian, A., Anguelov, D., Flynn, J., & Kosecka, J. (2017). 3d bounding box estimation using deep learning and geometry. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (pp. 7074-7082).
- [107] Eraqi, H. (Accessed: March 2017).Bag of Visual Words for Image Classification. retrieved from: <http://heraqi.blogspot.com/2017/03/BoW.html>
- [108] Lowe, D. G., (Accessed: March 2004). Method and apparatus for identifying scale invariant features in an image and use of same for locating an object in an image.patent ducument from university of birtish coloumbia. retrieved from: <https://patents.google.com/patent/US6711293B1/en>
- [۱۰۹] عسگریان دهکردی، رسول، و معلم، پیمان. (۱۳۹۳). " ردیابی تصویری سریع و وفقی هدف مبتنی بر روزرسانی مدل منظری جهت بهبود مقاومت نسبت به انسداد".پایان نامه کارشناسی ارشد. گروه الکترونیک دیجیتال. دانشکده برق و کامپیوتر دانشگاه اصفهان.

- [110] Harris, C. G., & Stephens, M. (1988). A combined corner and edge detector. In Alvey vision conference, 15(50), 10-5244.
- [۱۱۱] لاچینی، رضا، و احمدزاده، محمد رضا. (۱۳۹۲). "تشخیص اشیا متحرک در تصاویر هوایی دوربین متحرک". پایان‌نامه کارشناسی ارشد. گروه هوش مصنوعی. دانشکده برق و کامپیوتر. دانشگاه صنعتی اصفهان.
- [۱۱۲] غلامعلی نژاد، حسین. (مهرماه ۱۳۹۶)، پیاده سازی شبکه عمیق در متلب: ساختن شبکه یادگیری عمیق ساده برای کلاسه بندی. دریافت شده از: https://shahaab-co.ir/dnn_matlab_classification_1/
- [113] Mahadevan, V., & Vasconcelos, N. (2009). Spatiotemporal saliency in dynamic scenes. IEEE transactions on pattern analysis and machine intelligence, 32(1), 171-177.
- [114] Zitova, B., & Flusser, J. (2003). Image registration methods: a survey. Image and vision computing, 21(11), 977-1000.

Abstract

In recent years, road cameras are used in most of the roads to monitor the vehicles. Due to the increasing number of these cameras and large number of vehicles, using human resources for these cameras, makes some problems in the future. So, it is desired to use machine vision and image processing methods to estimate some parameters like vehicle speed, vehicle class, and the number of passing vehicles. Previously proposed methods have some limitations like non-automatic nature, shadow sensitivity, single-purpose, and use of small and non-real datasets.

In this thesis we propose a method for speed and dimensions estimation of the vehicles. This method must meet some objectives like: fully automatic calibration, shadow resistance, multi-purpose, and provide us with a comprehensive dataset.

The proposed method is conducted as follows: In the first frames, vanishing points and the hypothetical road surface are obtained based on the direction of vehicles' movement. A GMM based foreground detection and a new shadow removal algorithm is applied to determine the bounding box of each vehicle and construct its 3D surrounding box.

To determine the metric coefficients, some popular vehicles are recognized in the first frames. Then, metric coefficients are calculated based on the real dimensions of the vehicles (in meter) and their corresponding dimensions on the road surface (in the pixel). These steps, satisfies one of our goals, which is fully-automatic calibration. Finally, the passing vehicles are projected on the hypothetical road surface, and their precise speed and dimensions are calculated by tracking.

Robustness of the algorithm against shadow leads to the accurate bounding box detection of each vehicle and reduces the errors of calibration, and speed and dimensions estimation. In the shadow removal algorithm, in addition to the accuracy, the processing time is important. So, we proposed a fast algorithm that its processing time is almost negligible.

Currently, there is no standard dataset of videos including speed tags and the vehicles' images, especially in Iran. Another objective of this thesis is to provide a dataset including such information. For that, we prepared such dataset by simultaneous capturing of the roads with IP and laser camera.

The calibration process needs to recognize some popular vehicles at first. Here a powerful algorithm is presented by deep networks and image registration. According to the installation of the road cameras in high altitudes, vehicle images are so small. A super-resolution algorithm is used to produce high resolution images and improve the recognition rate.

The other objective of this thesis is to be multi-purpose. Most of the existing ITS algorithms are about speed estimation or vehicle type recognition. The dimensions estimation and vehicle recognition in this thesis, along with the provided dataset of the images and videos make it possible to recognize the passing vehicles by two criteria: "recognition algorithm" and "the estimated dimensions". Hence, as the secondary application of the algorithm, a method is presented to determine the type of passing vehicles, in addition to the estimation of the dimensions and speed of the vehicles.

The average errors of speed and dimensions estimation are 1.2Km/h and 8%, respectively. The results show the superiority of the proposed method against previous methods in both criteria.

The processing speed of the proposed method is about 2.16 FPS using MATLAB 2016 on a computer with 16GB RAM and a 2.6 GHz Core i7 processor.

Keywords: vehicle speed estimation, vehicle dimension estimation, vehicle type recognition, camera calibration, vanishing point.



Shahrood University of Technology
Faculty of Electrical and Robotic Engineering

PHD Thesis

**Vehicle Speed and Dimension Estimation Using Videos
Captured by Roadside Cameras**

By:

Rasoul Asgarian Dehkordi

Supervisor:

Dr. Hossein Khosravi

Advisor:

Dr. Alireza Ahmadi Fard

November 2019