

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



رشته مهندسی برق الکترونیک

گرایش سیستم

پایان نامه کارشناسی ارشد

تشخیص و طبقه‌بندی بی درنگ نوع، رنگ و جهت حرکت خودروهای واقع بر سطح جاده در یک دنباله ویدئویی

احمد مصیبي

استاد راهنما: دکتر حسین خسروی

تیر ماه ۱۳۹۵



مدیریت تحصیلات تکمیلی

باسمه تعالی

شماره: ۱۴۰۶ - ۱۰۰

تاریخ: ۱۳۹۵/۰۴/۲۹

ویرایش: -----

فرم شماره ۷: صورتجلسه دفاع از پایان نامه تحصیلی دوره کارشناسی ارشد

با تأییدات خداوند متعال و با استعانت از حضرت ولی عصر (عج) ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد

خانم / آقای احمد مصیبی به شماره دانشجویی ۹۲۱۳۱۹۴ رشته مهندسی برق - الکترونیک

گرایش سیستم که در تاریخ ۹۵/۰۴/۲۹

تحت عنوان: تشخیص و طبقه بندی بی درنگ نوع، رنگ و جهت حرکت خودروهای واقع بر سطح جاده در یک دنباله ویدیویی

با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می گردد:

قبول (با درجه: بسیار خوب امتیاز ۱۸/۸۲)	☒ نظری	☐ عملی
☐ مردود	☐ دفاع مجدد	

۲- بسیار خوب (۱۸/۹۹ - ۱۸)

۱- عالی (۲۰ - ۱۹)

۴- قابل قبول (۱۵/۹۹ - ۱۴)

۳- خوب (۱۷/۹۹ - ۱۶)

۵- نمره کمتر از ۱۴ غیر قابل قبول

عضو هیأت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱- استاد راهنمای اول	مهندس	استاد	
۲- استاد راهنمای دوم	-	-	-
۳- استاد مشاور	-	-	-
۴- نماینده شورای تحصیلات تکمیلی	جبار کاشانی	استاد	
۵- استاد ممتحن اول	علیرضا احمدی	دانشیار	
۶- استاد ممتحن دوم	هادی کرمانلو	استاد	

نام و نام خانوادگی رئیس دانشکده:

تاریخ و امضاء و مهر دانشکده:

تقدیم بہ پدر و مادر عزیزم،

کہ در این مسیر من را باورد داشتندو...

باسپاس...

از زحمات و راهنمایی های استاد راهنمای بزرگوارم جناب آقای دکتر حسین خسروی که در مسیر همواره مرئیاری نمودند.

احمد مصیبي ۱۳۹۵

تعهد نامه

- اینجانب احمد مصیبی دانشجوی دوره کارشناسی ارشد رشته مهندسی برق الکترونیک گرایش سیستم دانشکده برق و رباتیک دانشگاه صنعتی شاهرود نویسنده پایان نامه تشخیص و طبقه‌بندی بی درنگ نوع، رنگ و جهت حرکت خودروهای واقع بر سطح جاده در یک دنباله ویدئویی تحت راهنمایی دکتر حسین خسروی متعهد می شوم.
- تحقیقات در این پایان نامه 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 پایان نامه 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 شاهرود می باشد و مقالات مستخرج با نام « دانشگاه شاهرود » و یا « Shahrood University » به چاپ خواهد رسید.
 - حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
 - در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
 - در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاقی انسانی رعایت شده است .

تاریخ : ۱۳۹۵/۴/۲۹

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 شاهرود می باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

- متن این صفحه نیز باید در ابتدای نسخه های تکثیر شده پایان نامه وجود داشته باشد .

چکیده

تشخیص نوع خودرو در جاده، به طور کلی نقش مهمی در کاربردهای مربوط به سیستم‌های حمل و نقل هوشمند جهت افزایش امنیت و سهولت تردد وسایل نقلیه در جاده‌ها ایفا می‌کند. روش‌های موجود برای تشخیص نوع خودرو، روش‌هایی هستند که با استفاده از حسگرهای مختلف، از جمله حسگرهای لیزری، راداری، دیداری و... اطلاعات مورد نیاز برای تشخیص را بدست آورده و آن‌ها را مورد بررسی قرار می‌دهند، اما در این میان استفاده از روش‌های دیداری و مبتنی بر تصاویر یا ویدئو، با توجه به اعتبار مناسب نتایج و نیز هزینه‌های کم، چه از نظر سخت افزار مورد نیاز و چه از نظر زمان پردازش، بیش‌تر از سایر موارد مورد توجه قرار گرفته‌اند. تشخیص خودرو می‌تواند در دو سطح صورت گیرد: این‌که نوع خودرو، تنها از نظر ابعاد و ویژگی‌های کلی ظاهری آن تعیین شود (تقسیم بندی به گروه‌های سواری، سنگین و ...) و یا این‌که به طور جزئی‌تر، مدل و شرکت سازنده خودرو هم تعیین شود. در این پایان نامه برای بخش تشخیص خودرو و نیز رنگ آن، از ویدئوهای نمای نزدیک تهیه شده از درگاه ورودی یک عوارضی استفاده شده است. این نوع ویدئوهای نمای نزدیک، به دلیل دشواری‌هایی که جهت جداسازی خودرو از تصاویر و نیز میدان دید محدودی که دارند، کم‌تر مورد توجه قرار گرفته‌اند. تشخیص نوع و مدل خودروهای عبوری از درگاه‌های عوارضی برای کاربردهای امنیتی، آمارگیری‌ها، خودکارسازی عملیات دریافت عوارض و... ضروری می‌نماید. در این پایان نامه الگوریتمی بلادرنگ برای تشخیص نوع و طبقه خودروهای عبوری از عوارضی‌ها ارائه شده است. متوسط درصد معیارهای حساسیت و دقت الگوریتم پیشنهادی در مرحله‌ی یافتن خودرو به ترتیب ۹۴/۷۵ و ۹۸/۸۳ درصد و میزان صحت در مراحل تشخیص کلاس کلی و مدل خودرو به ترتیب ۹۸ و ۹۶/۶۵ درصد بوده است. تشخیص رنگ نیز برای خودروهای سبک و پس از تعیین مدل انجام گرفته

است و برای تشخیص جهت حرکت خودرو از ویدئویی متفاوت که از نمای دور تهیه شده استفاده شده است.

کلید واژه‌ها :

تشخیص نوع خودرو، ویدئو، حمل و نقل هوشمند، بلادرنگ، HOG

مقالات مستخرج از پایان نامه :

1. A. Mosayebi and H. Khosravi, "*A SVM Based Approach for Real Time Detection and Classification of Vehicles at the Toll Gates Using Video Sequences*," presented at the Third National Conference and first International Conference on Applied Research in Electrical, Mechanical and Mechatronics Engineering, Malek-Ashtar University of Technology, Tehran, Iran, 2016.
2. A. Mosayebi and H. Khosravi, "*A Simple Approach for Real Time Speed Estimation of on Road Vehicles Using Video Sequences*," presented at the Third National Conference and first International Conference on Applied Research in Electrical, Mechanical and Mechatronics Engineering, Malek-Ashtar University of Technology, Tehran, Iran, 2016.

فهرست

۱	۱. فصل اول: مقدمه
۲	۱-۱ مقدمه
۳	۲-۱ روش دیداری برای تشخیص
۴	۳-۱ ورودی‌های سیستم مبتنی بر تصویر
۶	۴-۱ عملیات پردازشی
۶	۵-۱ انتخاب ROI
۷	۶-۱ استخراج ویژگی
۷	۷-۱ طبقه‌بندی
۸	۸-۱ ارزیابی
۱۱	۲. فصل دوم: پیشینه پژوهش
۱۲	۱-۲ مقدمه
۱۲	۲-۲ تشخیص و طبقه‌بندی نوع کلی خودرو
۱۷	۳-۲ تشخیص و طبقه‌بندی مدل خودرو
۲۴	۴-۲ تشخیص و جدا سازی وسیله نقلیه
۲۷	۵-۲ تشخیص رنگ خودرو
۲۹	۶-۲ تشخیص جهت خودرو
۳۳	۳. مبانی نظری
۳۴	۱-۳ مقدمه
۳۴	۲-۳ نادرهم بافی
۳۷	۳-۳ هیستوگرام گرادیان های جهت‌دار
۴۰	۴-۳ ماشین بردار پشتیبان
۴۶	۵-۳ ویژگی های تسريع شده مقاوم

۴۸.....	۳-۶ شار نوری
۵۳.....	۴. روش پیشنهادی
۵۴.....	۴-۱ مقدمه
۵۴.....	۴-۲ پیدا کردن ROI دربرگیرنده خودروها
۵۴.....	۴-۲-۱ استفاده از سایه زیر قسمت جلویی خودرو
۵۵.....	۴-۲-۲ تخمین پیش زمینه با استفاده از مدل مخلوط گاوسی ها
۵۶.....	۴-۲-۳ روش توصیفگر SURF متقارن
۵۷.....	۴-۲-۴ یافتن ROI به کمک هیستوگرام های افقی و عمودی تصویر پیش زمینه
۶۱.....	۴-۲-۵ تعیین پنجره توسط کاربر و استفاده از مرکز ثقل تصویر پیش زمینه داخل پنجره
۶۴.....	۴-۲-۶ استفاده از HOG و SVM
۶۶.....	۴-۲-۷ روش پیشنهادی برای تعیین ROI در بردارنده خودرو
۶۷.....	۴-۳ تشخیص کلاس کلی خودرو
۶۹.....	۴-۴ تعیین مدل خودرو و تشخیص رنگ
۶۹.....	۴-۴-۱ یافتن ناحیه جلوی خودروی سبک به طور دقیق تر
۷۴.....	۴-۴-۲ تعیین مدل خودروی سبک
۷۶.....	۴-۴-۳ تشخیص رنگ خودروی سبک
۷۸.....	۴-۵ تشخیص جهت حرکت خودرو
۸۵.....	۵. نتایج تجربی
۸۶.....	۵-۱ معرفی پایگاه داده
۸۶.....	۵-۱-۱ فرآیند انتخاب پایگاه داده و دسترسی به آن
۸۷.....	۵-۱-۲ مشخصات ویدئوهای تهیه شده
۸۸.....	۵-۲ نتایج عددی
۸۹.....	۵-۲-۱ نتایج مربوط به یافتن و تعیین کلاس کلی خودرو
۹۴.....	۵-۲-۲ نتایج مربوط به یافتن ROI جلوی خودروی سبک و تعیین مدل آن
۱۰۰.....	۵-۲-۳ تشخیص رنگ خودروی سبک

۴-۲-۵	تشخیص جهت حرکت خودرو	۱۰۰
۶	جمع بندی	۱۰۱
۲-۶	جمع بندی	۱۰۲
۳-۶	پیشنهادهای برای کارهای آینده	۱۰۳
۷	مراجع	۱۰۵

فهرست شکل‌ها

- شکل ۱-۲: دو نمونه استاندارد وسیله نقلیه یاد شده در روش [۲] و گوشه های هریس آنها
 الف) خودرو سواری ب) اتوبوس ۱۴
- شکل ۲-۲: نمونه‌هایی از تصاویر خودروهای مورد تشخیص در روش [۴] ۱۶
- شکل ۳-۲: روند آموزش سلسله مراتبی در روش [۸] ۱۸
- شکل ۴-۲: نمونه‌ای از جداسازی قسمت جلوی خودرو در روش [۹] ۱۹
- شکل ۵-۲: نقاط تقارنی، نواحی سایه زیر خودرو و جلوی خودرو یافته شده به کمک SSURF در روش [۱۰] ۲۰
- شکل ۶-۲: مراحل الگوریتم‌های الف) تشخیص و ب) تعیین مدل به کمک SSURF مربوط به روش [۱۰] ۲۲
- شکل ۷-۲: مراحل الف) مراحل دوگانه تعیین مدل در روش [۱۲] ب) ساختار مرحله اول در روش [۱۲] ساختار مرحله دوم در روش [۱۲] ۲۳
- شکل ۸-۲: مراحل کار در روش [۱۴] ۲۴
- شکل ۹-۲: نتایج تشخیص وسیله نقلیه به کمک الگوریتم Adaboost طبق روش [۱۵] ۲۵
- شکل ۱۰-۲: تصاویر الف) خودروها و ب) غیر خودروها در پایگاه داده روش [۱۷] ۲۶
- شکل ۱۱-۲: الف) روند آموزش در روش [۱۹] ب) روند آزمون در روش [۱۹] ۲۷
- شکل ۱۲-۲: یافتن قسمت های رنگی خودرو به کمک ماسک یابنده خودرو وابسته به موقعیت در روش [۲۰] ۲۸
- شکل ۱۳-۲: سه ناحیه از پیش تعیین شده و حالات ممکن برای حضور خودرو در آن نواحی در روش [۲۳] ۳۰
- شکل ۱۴-۲: نمودار حالات مربوط به مدل مخفی مارکوف در نظر گرفته شده در روش [۲۳] ۳۱
- شکل ۱۵-۲: الف) خودروی عبوری از خط مجازی در روش [۲۴] ب) میدان سرعت خودروی مورد الف) ۳۲
- شکل ۱-۳: ترکیب دو دامنه‌ی تشکیل دهنده‌ی یک قاب ۳۵
- شکل ۲-۳: تصاویر درهم بافته (الف) و نادرهم بافته (ب) ۳۶
- شکل ۳-۳: موقعیت سلول‌ها و بلوک‌ها در الگوریتم محاسبه‌ی HOG ۴۰
- شکل ۴-۳: سه مرز ممکن برای جداسازی داده‌ها به صورت خطی ۴۲
- شکل ۵-۳: مرزبندی‌های موجود در الگوریتم SVM در فضای دو بعدی ۴۲

شکل ۳-۶: (الف) داده‌های آبی و قرمز در فضای یک بعدی به صورت خطی جداپذیر نیستند (ب) جداپذیری خطی داده‌های دو گروه پس از انتقال به فضای دوبعدی با تابع درجه دوم ۴۵

شکل ۳-۷: نتیجه اجرای الگوریتم شار نوری به کمک مجموعه نقاط پراکنده بین دو قاب متوالی: نقاط قرمز مربوط به قاب قبلی و نقاط سبز مربوط به قاب بعد فعلی است ۵۰

شکل ۳-۸: نتیجه اجرای الگوریتم شار نوری چگال: تمام نقاط دارای حرکت بین دو قاب متوالی دنبال می‌شوند. ۵۱

شکل ۴-۱: سه نمونه تصویر که در آنها جداسازی سایه زیر خودرو به صورت یک ناحیه مستطیلی مجزا ممکن نیست: در ردیف پایین تصاویر سایه متناظر با تصاویر ردیف بالا نشان داده شده‌اند. ۵۵

شکل ۴-۲: استفاده از روش GMM برای تخمین پیش زمینه و یافتن خودرو ۵۶

شکل ۴-۳: نتیجه پیاده سازی روش [۱۰] بر روی دادگان پایان نامه: خطوط قرمز، واصل بین نقاط SURF (سبز رنگ) تقارنی هستند که به دلیل کیفیت نامناسب تصویر به درستی تعیین نشده‌اند. ۵۷

شکل ۴-۴: هیستوگرام‌های افقی و عمودی مربوط به یک تصویر دودویی پیش زمینه ۵۸

شکل ۴-۵: (الف) هیستوگرام افقی (ب) هیستوگرام افقی پس از اعمال فیلتر میانه با بازه‌های به طول ۳ (ج) هیستوگرام افقی پس از اعمال فیلتر میانگین با بازه‌های به طول ۲۱ و حذف مقادیر کمتر از $\alpha=0/05$ برابر مقدار بیشینه نمودار ۶۰

شکل ۴-۶: نتیجه نهایی الگوریتم: ROI در برگیرنده خودروی عبوری که به درستی به دست آمده است. ۶۰

شکل ۴-۷: انسداد قابل توجه خودروی جلویی با خودروی عقبی در حال حرکت ۶۱

شکل ۴-۸: سه ناحیه مستطیلی که در ابتدای الگوریتم توسط کاربر تعیین می‌شوند ۶۲

شکل ۴-۹: تصویر پیش زمینه: شرایط برای گرفتن تصویر در این صحنه مهیا است. دایره توپر سفید رنگ موقعیت مرکز ثقل ناحیه Centroid-ROI را نشان می‌دهد. ۶۳

شکل ۴-۱۰: تصویر ذخیره شده ۶۳

شکل ۴-۱۱: نمونه‌هایی از تصاویر جدا شده از قاب‌های ویدئو: در ردیف بالا تصاویر خودروها و در ردیف پایین تصاویر غیر خودروها قرار دارند. ۶۵

شکل ۴-۱۲: نمونه‌هایی از موارد زیاد خطا در یافتن خودرو به کمک HOG و SVM ۶۶

شکل ۴-۱۳: نمونه تصاویر خودروهای سبک و سنگین: در ردیف بالا تصاویر خودروهای سبک و در ردیف پایین تصاویر خودروهای سنگین قرار دارند. ۶۸

شکل ۴-۱۴: تصویر سایه حاصل از آستانه گذاری تصویر سطح خاکستری: پایین ترین قطعه خط افقی و مرکز آن، با رنگ خاکستری و مرز پایین ناحیه جلوی خودرو با رنگ سفید نشان داده شده‌اند. ۷۰

شکل ۴-۱۵: تصویر سایه: مرز سمت چپ ناحیه جلوی خودرو که در $x=87$ واقع شده با خط عمودی قرمز رنگ نمایش داده شده است. ۷۱

شکل ۴-۱۶: تصویر پیش زمینه: مرز سمت راست ناحیه جلوی خودرو که متناظر با مرز سمت راست وسیع ترین قله موجود در هیستوگرام عمودی است با یک خط عمودی قرمز رنگ نمایش داده شده است. ۷۲

شکل ۴-۱۷: تصویر گرفته شده از یک خودروی سبک: مستطیل سفید رنگ ناحیه جلوی خودرو را که الگوریتم پیشنهادی به دست داده است نشان می دهد. ۷۳

شکل ۴-۱۸: نمونه هایی از چهار مدل خودروی سبک در نظر گرفته شده در این پایان نامه. ۷۵

شکل ۴-۱۹: (الف) تصویر ناحیه جلوی خودرو (ب) لبه یابی شده تصویر الف : ناحیه مورد نظر برای انجام عملیات تشخیص رنگ با یک مستطیل قرمز رنگ و خط نیمه تصویر با خطی سبز رنگ نشان داده شده اند. ۷۷

شکل ۴-۲۰: گستره های hue مربوط به چهار گروه رنگ ۷۸

شکل ۴-۲۱: نمونه قابی ویدئوی تهیه شده برای تشخیص جهت ۷۹

شکل ۴-۲۲: (الف) تصویر دودویی پیشزمینه حاصل از تخمین با GMM (ب) تصویر دودویی پس از سایش ج) مستطیل محیطی حول خودرو کشیده شده است. ۸۰

شکل ۴-۲۳: خودرو در حال ورود است (الف) مرز پایین Car_ROI از خط L گذشته است (ب) نواحی مستطیلی R1 و R2 در بالا و پایین خط L در قاب ویدئو در هنگام رسیدن خودرو به خط L (پ) نواحی مستطیلی R1 و R2 در بالا و پایین خط L در تصویر دودویی در هنگام رسیدن خودرو به خط L (ت) نواحی مستطیلی R1 و R2 در بالا و پایین خط L در تصویر دودویی در هنگام عبور کامل خودرو ۸۲

شکل ۵-۱: نمونه قاب هایی از هریک از ویدئوهای تهیه شده از درگاه عوارضی باغچه ۸۷

شکل ۵-۲: نحوه قرارگیری دوربین نسبت به درگاه عوارضی در پایگاه داده تهیه شده ۸۸

شکل ۵-۳: نمونه تصاویر پایگاه داده روش [۲۷] : (الف) قاب نمونه از ویدئوی پایگاه داده (ب) یافتن ناحیه دوزنقه ای به عنوان شیشه جلوی خودرو ج) خودروهای پیدا شده ۹۳

فهرست جداول

جدول ۱-۲: نتایج بازشناسی روش [۷].....	۱۶
جدول ۱-۳: فرمول های چهار کرنل مهم مورد استفاده همراه با SVM و پارامترهای مهم آنها.....	۴۶
جدول ۱-۵: مدت زمان ویدئوها و نیز شرایط آب و هوا در ویدئوهای تهیه شده از عوارضی باغچه.....	۸۶
جدول ۲-۵: نتایج آزمون تشخیص وجود خودرو و تعیین کلاس کلی خودرو به شیوه گام به گام.....	۹۰
جدول ۳-۵: نتایج آزمون تشخیص وجود خودرو و تعیین کلاس کلی خودرو به شیوه گام به گام برای هریک از دو دسته خودروهای سبک و سنگین.....	۹۰
جدول ۴-۵: تعداد داده های آموزشی و نیز داده های اضافه شده از مراحل قبل به مجموعه آموزشی در هر مرحله.....	۹۱
جدول ۵-۵: نتایج تشخیص خودرو مربوط به روش [۱۵].....	۹۱
جدول ۶-۵: نتایج مربوط به روش [۱۰] برای یافتن نمای جلوی خودرو در سه نوع آب و هوای مختلف.....	۹۲
جدول ۷-۵: پارامترهای انتخاب شده برای استخراج HOG در روش [۲۷].....	۹۳
جدول ۸-۵: مشخصات به دست آمده با جستجوی شبکه ای برای طبقه بندهای تعیین مدل روش پیشنهادی.....	۹۵
جدول ۹-۵: نتایج حاصل از آزمون الگوریتم پیشنهادی تعیین مدل.....	۹۶
جدول ۱۰-۵: نتایج حاصل از آزمایش KNN با معیار فاصله اقلیدسی و به ازای مقادیر مختلف K.....	۹۷
جدول ۱۱-۵: درصد صحت و زمان اجرای الگوریتم تشخیص مدل خودرو در روش [۱۰].....	۹۷
جدول ۱۲-۵: نتایج مربوط به روش [۲۸].....	۹۸
جدول ۱۳-۵: مقایسه بهترین نتایج روش های یاد شده و روش پیشنهادی.....	۹۹

١. مقدمه

۱-۱ مقدمه

امروزه با توجه به پیشرفت روز افزون فناوری و جهت گیری کلی جوامع بشری به سمت آن و نیز نیاز روز افزون به خودکار سازی^۱ و افزایش سرعت و کارایی^۲ در کاربردهای مختلف، استفاده از فناوری در زمینه های گوناگون اجتناب ناپذیر است. نظارت بر حمل و نقل خودروها از جمله کاربردهایی است که از این قاعده مستثنی نیست، کنترل سرعت خودروها، کنترل ترافیک درون شهری و برون شهری، تشخیص جرائم رانندگی، دریافت عوارض در جاده ها و بسیاری موارد مشابه دیگر از جمله اموری هستند که نیاز به نظارت شبانه روزی و تلاش مداوم تعداد زیادی نیروی انسانی دارند. انجام امور نظارتی توسط انسان با خطا همراه بوده و مستلزم صرف هزینه های زیادی توسط دولت یا سایر ارگان های مسئول خواهد بود. سیستم های حمل و نقل هوشمند^۳، سیستم هایی شامل اجزاء مختلف هستند که سه علم رایانه، الکترونیک و مخابرات را برای تأمین امنیت و سهولت تردد وسایل نقلیه در جاده ها به کار می گیرند. این سیستم ها با توجه به مشکلات ذکر شده و نیز افزایش جمعیت، نیاز است تا بصورت خودکار درآیند. برای این کار سیستم های حسگری^۴ مختلف، اعم از لیزری، راداری، دیداری^۵ و... مورد استفاده قرار می گیرند. استفاده از حسگرهای الکترونیکی برای انجام اعمال نظارتی از جمله روش های قدیمی تر محسوب می شود که به دلیل هزینه و مشکلات زیاد نگه داری، روش های دیداری و مبتنی بر تصویر، در بسیاری از موارد به آنها ترجیح داده می شوند. تشخیص حضور خودروها در جاده ها، تعیین مدل آنها، شمارش آنها و به طور کلی نظارت بر رفت و آمد خودروها از این دست فعالیت ها هستند که به کمک سیستم های مبتنی بر تصویر قابل انجام هستند. روش های مبتنی بر تصویر، همان طور که از نامشان پیداست، عمل تشخیص را با انجام

^۱ Automation

^۲ Efficiency

^۳ Intelligent Transportation Systems (ITS)

^۴ Sensor

^۵ Vision based

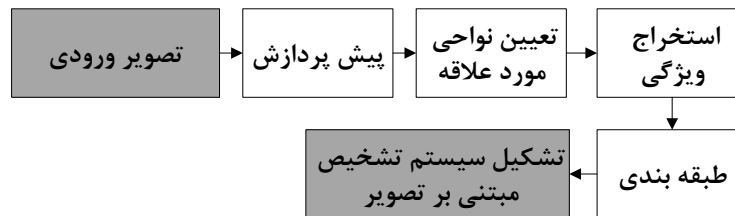
پردازش‌هایی بر روی تصاویر تهیه شده از خودروها انجام می‌دهند. تصاویر یاد شده بایستی تا حد امکان فاقد انواع نویز و ویدئوهای تهیه شده نیز بدون حرکت‌های ناخواسته باشند، که این‌ها نیازمند به کارگیری تجهیزات مناسب و نیز تصویربرداری به صورت مناسب است. در پردازش تصاویر خودرو روش‌های مختلفی به کار گرفته می‌شوند که هر کدام مزایا و معایبی دارند، در ادامه قالب کلی روش‌های دیداری بیان می‌شود.

۱-۲ روش دیداری برای تشخیص

روش‌های دیداری در تشخیص نوع خودرو، به روش‌هایی اطلاق می‌شوند که در آن‌ها از انواع مختلف دوربین‌های تصویربرداری اعم از نوری (عموماً رقومی^۱)، مادون قرمز^۲ و ... به عنوان حسگر استفاده می‌شود. در این روش‌ها یا دوربین مربوطه به صورت ثابت در محلی خاص تعبیه می‌شود و از خودروهای داخل جاده تصویر برمی‌دارد یا در برخی کاربردهای دیگر دوربین در داخل خودروی در حال حرکت تعبیه می‌شود و از خودروهای جلو یا عقب خودروی مربوطه تصویربرداری می‌کند. این تصاویر یا رشته‌های ویدئویی، به صورت خام یا پس از اعمال تغییرات، به عنوان ورودی، به سیستم تشخیصی وارد می‌شوند. در سیستم تشخیص، بر روی تصاویر مربوطه، عملیات پردازش تصویر مختلف اعمال می‌شود و تصاویر یا ویدیوها آماده می‌شوند، خودروها پیدا می‌شوند، نواحی خاصی از آن‌ها انتخاب و جداسازی می‌شوند، از آن نواحی ویژگی‌هایی استخراج می‌شود، آن گاه با استفاده از ویژگی‌های مناسب، تصاویر به تعدادی گروه از پیش تعیین شده طبقه‌بندی می‌شود، یعنی به نحوی ویژگی‌های خاص هر گروه از خودروها تعیین می‌شود و به این ترتیب یک سیستم تشخیص مبتنی بر تصویر شکل می‌گیرد. فلوچارت ۱-۱ این سیستم را بصورت نمودار بلوکی نشان می‌دهد.

^۱ Digital

^۲ Infrared



فلوچارت ۱-۱: تشکیل سیستم مبتنی بر تصویر

پس از تشکیل سیستم یاد شده، تصاویر جدیدی برای تعیین گروهی که به آن تعلق دارند و انجام عمل تشخیص، به سیستم وارد می‌شوند و با در نظر گرفتن معیارهای شباهت مختلف بین داده جدید و داده‌های گروه‌های قبلی، نوع وسیله‌ی نقلیه‌ی موجود در تصویر جدید ورودی مشخص می‌شود. لازم به ذکر است، تصاویر بسته به الگوریتم‌ها و روش‌های تشخیصی مورد استفاده و نیز بسته به کاربرد، می‌توانند کیفیت بالا یا پایینی داشته باشند.

۱-۳ ورودی‌های سیستم مبتنی بر تصویر

در مورد تصاویر ورودی موارد زیر قابل ذکر هستند:

- همان طور که در بخش قبل بیان شد ورودی‌های سیستم مربوطه هم می‌توانند به صورت رشته های ویدئویی و هم بصورت تک تصاویر باشند، که به کارگیری هریک از این دو دسته مزایا و معایبی دارد. استفاده از ویدئو، با توجه به خاصیت حرکت اشیا در آن، این امکان را فراهم می‌آورد تا بتوان پس زمینه^۱ و پیش زمینه^۲ را -در مقایسه با تک تصویر- راحت‌تر جداسازی کرد. از طرفی ویدئو نسبت به تک تصویر دارای تزاید اطلاعات است و همین می‌تواند موجب بالاتر رفتن هزینه‌های محاسباتی و ذخیره سازی شود.

^۱ Background

^۲ Foreground

- تصاویری که تهیه می‌شوند، بسته به کاربرد، می‌توانند در ابعاد و کیفیت‌های مختلف باشند و گستره‌ی وسیعی از تغییرات رنگ، سایه، نورپردازی (ساعات مختلف شبانه روز) و ... را شامل شوند.
- در برخی از کاربردها و روش‌های موجود، از تصاویر سطح خاکستری استفاده می‌شود، حال آن که برخی دیگر از روش‌ها هم هستند که از تصاویر رنگی استفاده می‌کنند و از رنگ به عنوان یک ویژگی مهم و مجزا کننده بهره می‌برند.
- روش‌های مختلف، نماهای مختلف خودرو، همچون نماهای جلو، عقب، کنار و مایل را مدنظر قرار داده‌اند. خودروها دارای قسمت‌های مختلفی هستند که برای کار تشخیص و استخراج ویژگی از آن‌ها استفاده می‌شود، به طور کلی تصاویری که تهیه می‌شوند بایستی حتماً نواحی خاص و حاوی ویژگی‌های متمایز کننده^۱ خودرو (مانند پلاک، لوگو، چراغ‌ها و...) را در بر گرفته باشند.
- تصاویر باید هم از نظر کمیت و هم از نظر تنوع مناسب باشند، به این معنا که هم تعداد کافی تصویر تهیه شده باشد، و هم از هر کلاس خودرو، داده‌هایی با تنوع شکلی کافی، وجود داده باشند.
- در مورد تهیه ویدئو، در صورتی که قرار است دوربین در تمام مدت تصویربرداری در یک محل خاص قرار گیرد بایستی دوربین تا حد ممکن در جای خود ثابت باشد و از حرکت‌های ناگهانی و شدید جلوگیری شود تا بتوان از ویژگی‌های مناسب تشخیص به کمک ویدئو بهره برد.
- تصاویر می‌توانند از فواصل مختلف و همچنین در مکان‌های مختلف (صحنه‌های متفاوت مانند اتوبان، ورودی پارکینگ‌ها و عوارضی‌ها، جاده‌های یک طرفه و دوطرفه و...) تهیه شده باشند و

^۱ Discriminative

نیز این امکان وجود دارد که بیش از یک خودرو در هر تصویر حضور داشته باشد و حتی با هم انسداد^۱ داشته باشند.

۴-۱ عملیات پردازشی

در این مرحله از کار، تصاویر بصورت نرم افزاری تحت عملیات مختلف پردازشی قرار می‌گیرند تا برای مراحل بعدی آماده شوند. از جمله این عملیات می‌توان به لبه‌یابی^۲ تصویر، هموارسازی^۳ و اعمال سایر فیلترهای سرتاسری^۴ به تصویر اشاره نمود. این مجموعه عملیات، ویژگی‌های مهم موجود در تصویر را هرچه بیش‌تر در دسترس قرار می‌دهند یا تعداد آن‌ها را به حد مورد نیاز می‌رسانند. این مرحله از آن جهت حائز اهمیت است که کار استخراج ویژگی و انتخاب ناحیه مورد علاقه^۵ را در مراحل بعدی ساده‌تر می‌کند.

۵-۱ انتخاب ROI

عموماً نواحی‌ای که به عنوان ROI انتخاب می‌شوند قسمت‌هایی از خودرو را در بر می‌گیرند که نماینده شکل و کلاس کلی خودرو یا مدل آن و دارای ویژگی‌هایی متمایز کننده باشند. به عنوان نمونه‌هایی از این نواحی می‌توان قسمت جلو یا عقب خودرو که پلاک خودرو یا لوگوی آن را شامل می‌شوند یا قسمت‌های شامل چراغ‌ها را نام برد. این نواحی به روش‌های مختلف تشخیص داده شده و به صورت مستطیلی از تصویر جدا می‌شوند تا بر روی آن‌ها الگوریتم‌های تشخیصی اعمال شود و به این ترتیب هزینه‌ی محاسباتی کم‌تر شود.

^۱ Occlusion

^۲ Edge detection

^۳ Smoothing

^۴ Global filters

^۵ Region Of Interest (ROI)

۱-۶ استخراج ویژگی

هدف از این پایان نامه تشخیص نوع خودرو است و تشخیص هر چیزی با تعیین ویژگی‌های آن ممکن است، لذا نیاز است از نواحی خاصی از تصویر با توجه به کاربرد و ماهیت تصویر، ویژگی‌هایی استخراج شود. لازم به ذکر است که پیش پردازش‌هایی که در مراحل قبل تا این مرحله صورت می‌گیرند در بهبود، برجسته سازی، سهولت استخراج و به کارگیری این ویژگی‌ها نقشی اساسی و مهم دارند. هریک از روش‌های مختلف موجود، ویژگی‌های گوناگونی را مد نظر قرار داده‌اند، برخی تبدیلات خاصی را بکار برده‌اند و نتایج خروجی آن را به عنوان ویژگی در نظر گرفته‌اند و برخی از ویژگی‌هایی که بصورت مستقیم از تصویر به دست می‌آیند استفاده می‌کنند. نکته‌ی اصلی در انتخاب ویژگی آن است که اولاً ویژگی‌های انتخاب شده تا حد امکان از پایداری^۱ مناسبی برخوردار باشند، یعنی اغلب در تصاویر، موجود و قابل دستیابی باشد و ثانیاً این ویژگی‌ها متمایز کننده باشند، به این معنا که برای گروه‌های مختلف خودرو متفاوت باشد.

۱-۷ طبقه‌بندی

در این مرحله، خودروها به تعداد خاصی دسته که از پیش تعیین شده‌اند، گروه‌بندی می‌شوند. برای این کار، از ویژگی‌های بدست آمده در مرحله قبل و روش‌های مختلف طبقه‌بندی استفاده می‌شود. در استفاده از روش‌های مختلف طبقه‌بندی داده، موضوعات مختلفی مد نظر قرار می‌گیرند، از جمله این که روش انتخابی با نوع داده‌ها (ویژگی‌ها) تناسب داشته باشد، به این معنا که طبقه‌بند^۲ مورد استفاده متناسب با توزیع داده‌ها انتخاب شده باشد تا بتواند به خوبی داده‌های دسته‌های مختلف را از هم تمییز دهد. موضوع

^۱ Stability

^۲ Classifier

دیگری که اهمیت دارد این است که بسته به کاربرد یا کیفیت تصاویر یا عوامل دیگر، به چه میزان دقت در طبقه‌بندی نیاز است و بنابر همین، نوع طبق بند انتخاب شود. همچنین می‌توان عاملی چون هزینه‌ی زمانی الگوریتم طبقه‌بندی را نیز در این انتخاب مهم دانست، چرا که این موضوع می‌تواند در بلادرنگ یا غیربلادرنگ بودن روش مؤثر باشد. به بیان دیگر، بایستی بین میزان دقت و سرعت مورد نیاز برای کار مصالحه‌ای معقول وجود داشته باشد، لذا در برخی روش‌ها با توجه به همین موضوع، از ترکیبی از طبقه‌بندها استفاده می‌شود.

۸-۱ ارزیابی

بعد از طبقه‌بندی نوبت به ارزیابی کار می‌رسد، به این شکل که تعدادی داده با همان قالب داده‌های مربوط به مراحل قبلی تهیه می‌شوند (بخشی از تصاویر اولیه که قبلاً مورد استفاده قرار نگرفته‌اند)، سپس تحت همان فرآیندهای پردازشی یاد شده قرار می‌گیرند و همان ویژگی‌ها از آن‌ها استخراج می‌شود. بعد با معیارهای متداول، میزان شباهت داده یاد شده با گروه‌ها (کلاس‌های خودرو) تعیین و تعلق آن به یکی از گروه‌ها تعیین می‌شود. پس از این کار به صورت بانظارت^۱ یا بدون نظارت^۲ تعلق داده جدید ورودی به گروهی خاص تعیین می‌شود و با انجام مقایسه، دقت کار سنجیده و بصورت درصد بیان می‌شود. همچنین روش‌های دیگر، معیارهای شباهت دیگر و یا طبقه‌بندهای دیگر مورد آزمایش قرار می‌گیرند و نتایج آن آزمون‌ها هم جهت مقایسه با روش پیشنهادی ارائه می‌شوند.

در ادامه، در فصل دوم، روش‌های پیشین مربوط به تشخیص نوع خودرو، مدل خودرو و یافتن خودرو در تصاویر مورد بحث قرار می‌گیرند، همچنین به چند روش تشخیص رنگ و جهت خودرو اشاره می‌شود. در فصل سوم، مباحث نظری مورد استفاده در روش پیشنهادی شرح داده می‌شوند. در فصل چهارم، مراحل

^۱ Supervised

^۲ Unsupervised

مختلف طی شده برای انجام روش پیشنهادی بیان شده‌اند، در فصل پنجم، نتایج تجربی اجرای روش پیشنهادی و مقایسه با سایر روش‌ها انجام گرفته است و در نهایت فصل ششم، به نتیجه‌گیری و پیشنهاد برای کارهای آینده اختصاص یافته است.

۲. پیشینه پژوهش

۲-۱ مقدمه

همان طور که پیش از این گفته شد تشخیص خودرو یک فرآیند چند مرحله‌ای است؛ برای به انجام رساندن هریک از مراحل یاد شده یا بهبود آن‌ها روش‌های مختلفی ارائه شده‌اند که نواقص موجود در روش‌های پیشین امکان ظهور ایده‌های نو را به وجود می‌آورد. تصاویر ورودی می‌توانند از صحنه‌ها و نماهای مختلف، فواصل متفاوت و ... تهیه شده باشند و کار با تصاویر مختلف نیازمند پرداختن به ایده‌های مختلف است. بسته به نوع تصاویر موجود در پایگاه داده^۱ عملیات پردازشی مختلفی قابل انجام است. برخی روش‌های موجود، در زمینه تعیین و چگونگی یافتن ROI و برخی در خصوص استخراج ویژگی‌های مناسب به طرح ایده پرداخته‌اند. همچنین روش‌های مختلف از طبقه‌بندهای مختلفی بسته به نوع داده‌ها، دقت و سرعت مورد نیاز خود استفاده می‌کنند. برای تشخیص نوع خودرو دو رویکرد کلی وجود دارد: رویکرد اول تعیین کلاس کلی خودروها است (سواری، سنگین و ...) و رویکرد دوم تعیین مدل و شرکت سازنده خودروها است. مباحث مطرح شده در این فصل مروری است بر کارهایی که تاکنون انجام شده و شامل بخش‌های تشخیص و طبقه‌بندی نوع کلی خودرو، تشخیص و طبقه‌بندی مدل خودرو، تشخیص و جدا سازی وسیله نقلیه، تشخیص رنگ خودرو و تشخیص جهت حرکت خودرو است.

۲-۲ تشخیص و طبقه‌بندی نوع کلی خودرو

در این بخش روش‌هایی معرفی می‌شوند که تنها عمل تشخیص کلاس کلی وسیله‌ی نقلیه را انجام می‌دهند. در [۱] برای دسته‌بندی وسایل نقلیه به پنج گروه سواری، ون، سواری بدون صندوق^۲، اتوبوس و کامیون، تصاویر ثابتی از نمای کنار وسایل نقلیه تهیه شده است. در این روش دو-سوم بالای تصاویر

^۱ Database

^۲ Hatch-back sedan

لبه‌یابی شده که دربرگیرنده قسمت‌های مهم خودرو و حاوی اطلاعات تعیین کننده‌ای است که مورد استفاده قرار گرفته است. از تبدیل گابور^۱ برای استخراج ویژگی استفاده شده است و برای دسته‌بندی، بردار ویژگی ورودی به دسته‌ای تعلق می‌گیرد که کمترین فاصله اقلیدسی را با مرکز آن داشته باشد. در این روش استفاده از بانک فیلتر گبور سبب کاهش هزینه‌ی محاسباتی و نیز کاهش حافظه‌ی مورد نیاز شده است. علاوه بر این، نوعی روش نمونه‌برداری ارائه شده است تا با آن قسمت‌هایی از تصاویر جدا شده و مورد پردازش قرار گیرند و بدین وسیله هزینه‌های پردازشی کم‌تر شود. در این روش نرخ بازشناسی در حدود ۹۵٪ حاصل شده است. در [۲] از ویدئو برای طبقه‌بندی وسایل نقلیه به سه دسته سواری، اتوبوس و کامیون استفاده شده است، در این روش از وسایل نقلیه، از نمای کنار تصویربرداری شده است. ابتدا وسیله نقلیه عبوری از جاده با تفریق پیش زمینه^۲ از پس زمینه^۳ با استفاده از نوعی مدل مخلوط گوسی^۴ [۳] جداسازی می‌شود. پیش از اجرای الگوریتم، نمونه‌های استاندارد از هریک از سه دسته وسیله نقلیه یاد شده تهیه و به کمک گوشه‌یاب هریس^۵ گوشه‌های مهم موجود در آن‌ها تعیین می‌شود. در مرحله ارزیابی، گوشه‌های مهم هر وسیله نقلیه عبوری پیدا شده هم تعیین می‌شود و معیار فاصله‌ی هاسدورف^۶ بین گوشه‌های پیدا شده هر خودروی عبوری و گوشه‌های پیدا شده نمونه‌های استاندارد محاسبه می‌شود، خودروی عبوری به دسته‌ای تعلق خواهد داشت که کمترین فاصله هاسدورف را با آن نمونه‌ها داشته باشد. استفاده از گوشه‌ها در این روش سبب شده تا عواملی چون تنوع خودروها، شرایط نورپردازی متفاوت، انسداد و چرخش جزئی تأثیری بر نتایج نداشته باشند. دقت تفکیک برای سواری، کامیون و اتوبوس به

^۱ Gabor transform

^۲ Foreground

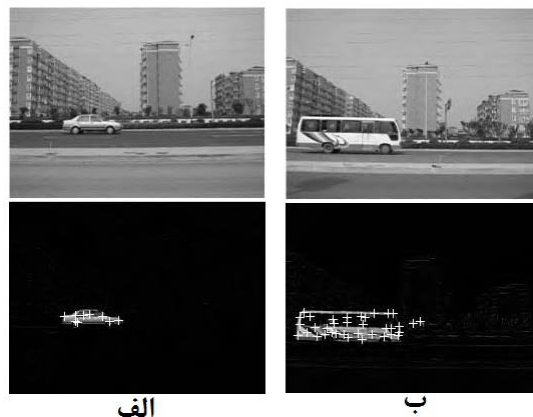
^۳ Background

^۴ Mixture of Gaussian Model (GMM)

^۵ Harris corner detector

^۶ Hausdorff

ترتیب ۱۰۰٪، ۹۵٪ و ۹۶٪ و زمان اجرای الگوریتم هم برای این سه دسته به ترتیب ۲۴۷، ۲۸۶ و ۲۶۳ میلی ثانیه است. در شکل ۱-۲ دو نوع خودرو و گوشه‌های هریس مربوط به آن‌ها آورده شده است.



شکل ۱-۲: دو نمونه استاندارد وسیله نقلیه یاد شده در روش [۲] و گوشه‌های هریس آن‌ها (الف) خودرو سواری (ب) اتوبوس

در [۴] الگوریتمی برای دسته بندی نمای روبه‌روی وسایل نقلیه عبوری به چهار دسته سواری، ون، کامیونت و اتوبوس ارائه شده است. در وهله اول برای یافتن وسیله نقلیه دو دسته داده با برجسب‌های مثبت و منفی تهیه می‌شود، داده‌های مثبت تصاویری حاوی نمای روبه‌روی خودروها و داده‌های منفی تصاویری فاقد خودرو هستند، سپس ویژگی‌های هار^۱ از این دو دسته تصویر استخراج می‌شود و یک مجموعه طبقه‌بند شبکه عصبی چند لایه پرسپترون^۲ با این ویژگی‌ها آموزش داده می‌شود. تصاویر مورد استفاده در این روش به وسیله دوربین‌های مداربسته نصب شده در ارتفاع و مشرف بر جاده تهیه شده‌اند. با دریافت هر تصویر، پنجره‌هایی بر روی ناحیه‌ای از تصویر لغزانده می‌شود و وجود یا عدم وجود وسیله نقلیه بررسی می‌شود، در صورت یافتن وسیله نقلیه پنجره حاوی آن جداسازی می‌شود. برای طبقه‌بندی خودروی یافت شده به چهار نوع یاد شده، از روشی موسوم به چندی‌سازی برداری^۳ استفاده

^۱ Haar features

^۲ Multiple Layer Perceptron (MLP)

^۳ Classified Vector Quantization (CVQ)

شده است که در آن به کمک سه الگوریتم آموزشی بدون نظارت نقشه خودسامانده^۱، گاز عصبی^۲ و K میانگین^۳ برای هریک از آن‌ها کتابچه رمزی^۴ از تصاویر استاندارد تهیه می‌شود. آن‌گاه به کمک کتابچه رمز و معیار نزدیکترین همسایگی، داده‌های ورودی دسته‌بندی می‌شوند. همچنین در الگوریتم CVQ گزینه رد^۵ هم در نظر گرفته شده است، به این معنی که پیش از اقدام به دسته‌بندی داده‌ها، احتمال این که داده به درستی طبقه‌بندی شود محاسبه می‌شود، اگر این احتمال کم باشد طبقه‌بندی برای این داده صورت نمی‌گیرد و به این ترتیب دقت بازشناسی روش بیش‌تر می‌شود. این روش با در نظر گرفتن نرخ رد ۸٪، به دقت بازشناسی ۹۵٪ رسیده است، خودروهای مورد تشخیص در این روش در شکل ۲-۲ آورده شده است. استفاده از پیکسل‌های تصویر به صورت خام به عنوان ویژگی، به دلیل وجود انواع نویز و تغییرات نور، نامناسب است، اما مانند آنچه [۵] انجام داده است، به کمک تحلیل مؤلفه‌های اصلی^۶ می‌توان پیکسل‌های تصویر را با بردارهایی با ابعاد کم‌تر که دربردارنده اطلاعات مهم تصویر هستند و کار با آن‌ها به مراتب به صرفه‌تر است جایگزین نمود. در این روش در مرحله آموزش تصاویری از نمای روبه‌روی پنج دسته وسیله نقلیه تهیه می‌شود، مولفه‌های اساسی پیکسل‌های هر تصویر بعنوان بردار ویژگی استخراج می‌شود و داده‌ها با استفاده از K-means به پنج دسته گروه‌بندی می‌شوند. در مرحله آزمون، خودرو با تفریق پس زمینه و حذف سایه در ویدئو، جداسازی می‌شود، محل پلاک خودرو تعیین می‌شود و با توجه به محل پلاک، ناحیه جلوی خودرو مشخص و جدا می‌شود. بردار ویژگی مربوط به تصویر جلوی خودرو نیز به کمک PCA استخراج می‌شود، مجموع مربعات فواصل آن بردار تا بردارهای هریک از پنج دسته محاسبه می‌شود و داده به گروهی تعلق می‌گیرد که کمترین فاصله را با بردارهای ویژگی آن دسته داشته باشد.

^۱ Self-Organizing Map (SOM)

^۲ Neural Gas (NG)

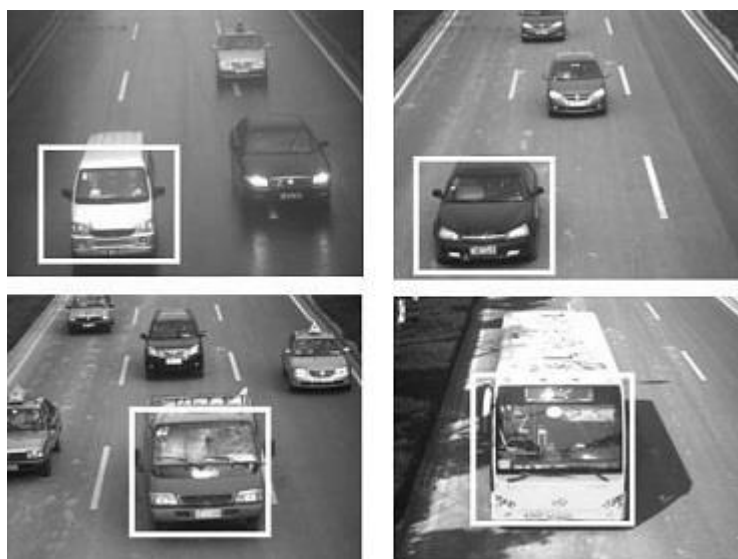
^۳ K-means

^۴ Codebook

^۵ Reject option

^۶ Principal Components Analysis (PCA)

دقت استخراج ناحیه جلوی خودرو در این روش ۸۰٪ و نرخ بازشناسی مربوط به تعیین دسته خودرو ۹۰٪ است. در [۶] از GMM برای تخمین پیش‌زمینه و جداسازی وسیله نقلیه عبوری استفاده شده است. ویژگی‌های هندسی همچون طول، عرض، نسبت طول و عرض، اندازه پلاک و ... از خودروی جدا شده استخراج می‌شوند. براساس این ویژگی‌ها، یک شبکه بیزین پویا^۱ [۷] شکل داده می‌شود و وسایل نقلیه به چهار دسته سواری، اتوبوس، مینی‌بوس و ناشناخته تفکیک می‌شوند، نتایج بازشناسی الگوریتم در مورد ۸۰ خودرو در جدول ۱-۲ آورده شده است.



شکل ۲-۲: نمونه‌هایی از تصاویر خودروهای مورد تشخیص در روش [۴]

جدول ۱-۲: نتایج بازشناسی روش [۷]

Vehicle type	Classificatio result				Total
	<i>Sedan</i>	<i>Bus</i>	<i>Microbus</i>	<i>Unknown</i>	
<i>Sedan</i>	45	0	6	0	51
<i>Bus</i>	0	7	2	0	9
<i>Microbus</i>	3	1	12	0	17
<i>Unknown</i>	0	0	0	3	3

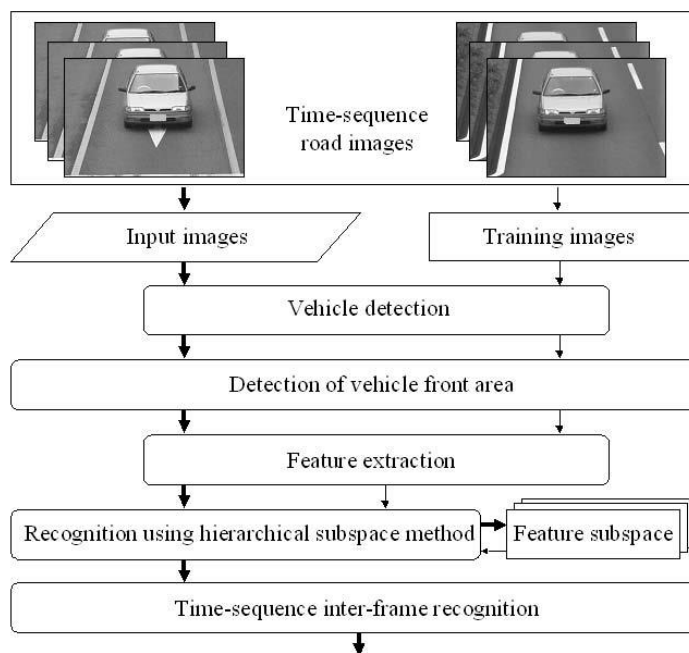
^۱ Dynamic Bayesian Network (DBN)

۳-۲ تشخیص و طبقه بندی مدل خودرو

در [۸] برای بررسی وارد شدن خودرو به ناحیه میدان دید دوربین و همچنین یافتن ناحیه جلوی خودرو، از تصویری موسوم به تصویر لبه تقارنی^۱ استفاده شده است. در این نوع تصویر، پیکسل‌های مربوط به سطح جاده مقادیری کم و مرزهای خودرو با جاده دارای مقادیر زیاد هستند. با توجه به همین موضوع و انجام پویش‌های افقی و عمودی، مرزهای بالا و پایین ناحیه جلوی خودرو تعیین و جداسازی می‌شود. طی یک فرآیند سلسله مراتبی ابتدا خودروهای با مدل‌های مشابه در گروه‌های جداگانه قرار می‌گیرند و سپس مدل آن‌ها به طور دقیق‌تر تعیین می‌شود. در مرحله اول پیکسل‌های ناحیه جلوی خودرو با توجه به میانگین و واریانس نرمال‌سازی می‌شوند و حاصل تصویری با ابعاد کوچکتر است که پیکسل‌های آن به عنوان بردار ویژگی در نظر گرفته می‌شوند. در مرحله دوم نواحی خاصی از تصویر ناحیه جلوی خودرو به عنوان ویژگی استخراج می‌شود. در هریک از این مراحل، از پیش، داده‌هایی برای آموزش تهیه شده‌اند و پس از استخراج ویژگی‌های یاد شده از آن‌ها، با یکی از روش‌های کاهش ویژگی به زیرفضایی^۲ با ابعاد کمتر برده می‌شوند. در هریک از دو مرحله، داده ورودی به کلاسی تعلق می‌گیرد که بیش‌ترین میزان شباهت را با آن داشته باشد، از طرفی، اگر شباهت داده‌ی ورودی با مدل‌های استاندارد آموزش داده شده از حدی کمتر باشد داده رد خواهد شد، در شکل ۳-۲ روند آموزش سلسله مراتبی مربوط به روش [۸] نشان داده شده است.

^۱ Symmetry edge image

^۲ Subspace



شکل ۲-۳: روند آموزش سلسله مراتبی در روش [۸]

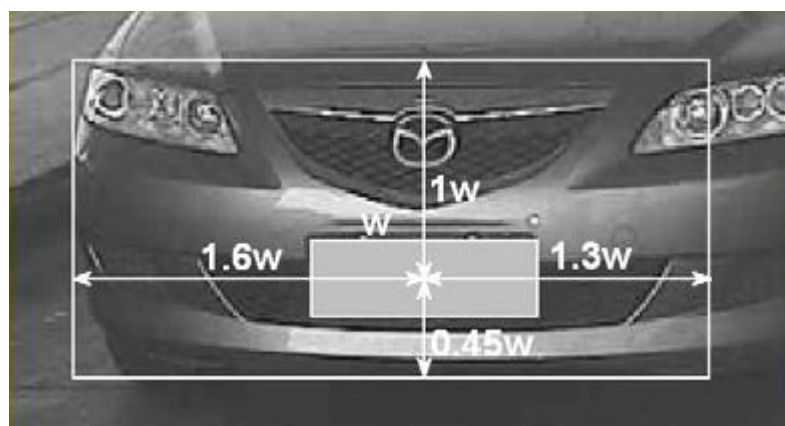
در این روش، دقت بازشناسی کلاس کلی خودرو بیش از ۹۹٪ و دقت بازشناسی مدل خودرو بیش از ۹۲٪ بوده است. در روش [۹] از تحلیل مجزا ساز خطی دو بعدی^۱ برای تشخیص مدل خودرو استفاده شده است. در این روش ابتدا پلاک خودروی عبوری پیدا و ناحیه جلوی خودرو با توجه به موقعیت و اندازه پلاک جداسازی می‌شود، شکل ۲-۴ نمونه‌ای از جداسازی قسمت جلوی خودرو در این روش را نشان می‌دهد. سپس تصویر گرادیان با لبه یاب سوبل^۲ به دست آمده و به کمک از تحلیل مجزا ساز خطی دو بعدی، برای هر تصویر یک ماتریس ویژگی ایجاد می‌شود. همین شیوه برای استخراج ویژگی از داده‌های آموزشی نیز انجام می‌شود. در نهایت با محاسبه فاصله داده آزمون از داده‌های هر مدل و استفاده از طبقه بند نزدیک‌ترین همسایگی^۳ مدل خودرو مشخص می‌شود. تعداد ۶۰۰ تصویر از ۲۰ مدل خودرو به عنوان

^۱ 2 Dimensional Linear Discriminant Analysis (2DLDA)

^۲ Sobel

^۳ Nearest neighbor

داده‌های آموزشی و ۸۱۴ تصویر هم برای آزمون استفاده شده است و نرخ بازشناسی ۹۴٪ حاصل شده است، این نکته قابل ذکر است که استفاده از تحلیل مجزا ساز خطی دو بعدی در مقایسه با PCA سبب افزایش نرخ بازشناسی به میزان ۳۰٪ شده است.



شکل ۲-۴: نمونه‌ای از جداسازی قسمت جلوی خودرو در روش [۹]

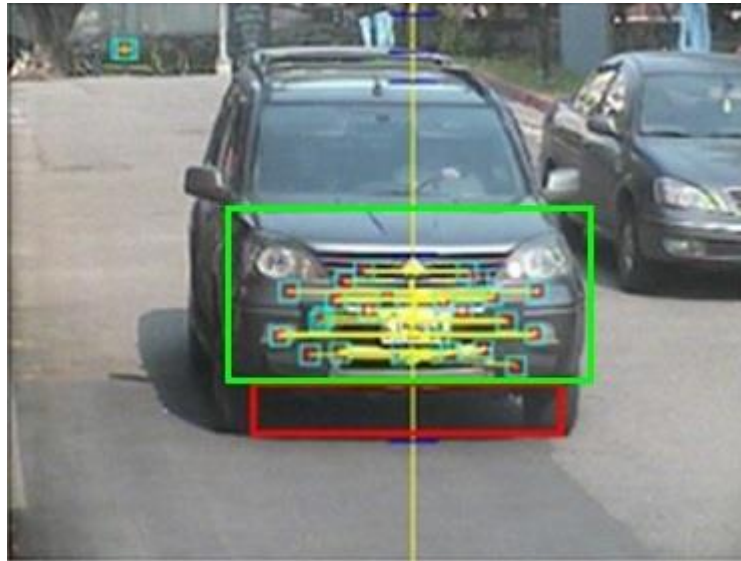
در [۱۰] از نوع تغییر یافته ویژگی‌های تسریع شده مقاوم^۱ موسوم به ویژگی‌های تسریع شده مقاوم متقارن^۲ برای یافتن ناحیه جلوی خودرو استفاده شده است و لذا نیازی به تخمین پیش‌زمینه، آموزش طبقه‌بندهای آبشاری^۳ و یا سایر روش‌های پرهزینه نیست. به کمک SSURF نقاط متقارن در صحنه پیدا می‌شوند، وقتی تعداد این نقاط تقارنی از حدی بیش‌تر شود و همچنین خط واسط این نقاط تقریباً یکی باشد، خودرو به صحنه وارد شده است و بایستی ناحیه جلوی آن استخراج شود. برای این کار ابتدا خط وسط خودرو به کمک نقاط تقارنی تعیین می‌شود، سپس سایه موجود در زیر خودرو با در نظر گرفتن میانگین و انحراف معیار شدت روشنایی تعدادی داده آموزشی (تعدادی تصویر تهیه شده از سایه‌های زیر خودرو) پیدا می‌شود. حد بالای سایه همان حد پایین جلوی خودرو است و حدود بالا، چپ و راست با

^۱ Speeded Up Robust Features (SURF)

^۲ Symmetrical SURF (SSURF)

^۳ Cascade classifiers

توجه با ابعاد سایه تعیین می‌شوند. مستطیل‌های قرمز و سبز رنگ در شکل ۲-۵ به ترتیب نواحی سایه زیر خودرو و جلوی خودرو را نشان می‌دهند و نقاط متصل شده به هم با خطوط افقی هم نقاط تقارنی یاد شده هستند.



شکل ۲-۵: نقاط تقارنی، نواحی سایه زیر خودرو و جلوی خودرو یافته شده به کمک SSURF در روش [۱۰]

برای تشخیص مدل خودرو، ناحیه جلوی خودرو به شبکه‌هایی^۱ تقسیم بندی می‌شود، برای هریک از این شبکه‌ها بردارهای هیستوگرام گرادیان‌های جهت‌دار^۲ و SURF استخراج شده و با استفاده از یک طبقه‌بند ماشین بردار پشتیبان^۳ که از قبل با داده‌های آموزشی آموزش داده شده است مورد دسته بندی قرار می‌گیرد. در نهایت هم با در نظر گرفتن آراء طبقه‌بندهای تمامی شبکه‌ها مدل خودرو معلوم می‌شود، در شکل ۲-۶ الف و ب به ترتیب مراحل الگوریتم‌های تشخیص و تعیین مدل نشان داده شده‌اند. دقت تشخیص خودرو و دقت بازشناسی هردو بیش از ۹۸٪ بوده‌اند و زمان پردازش هر قاب تصویر هم در حدود ۲۴ میلی‌ثانیه بوده است. در [۱۱] روشی مشابه با [۱۰] آورده ارائه شده است با این تفاوت که

^۱ Networks

^۲ Histogram of Oriented Gradients (HOG)

^۳ Support Vector Machine (SVM)

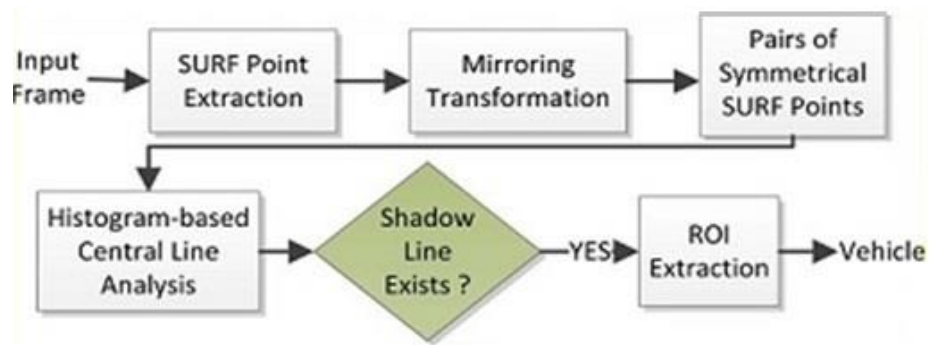
طبقه‌بندهای SVM استفاده شده در [۱۰] حذف شده‌اند و ابعاد تمام ویژگی‌های استخراج شده از همه شبکه‌ها با روشی تحت عنوان نمایش تُنک^۱ کم شده است و در نهایت با استفاده از معیار فاصله همینگ^۲ کار تعیین مدل خودرو به انجام رسیده است. در [۱۲] خودرو از میان تصاویر ثابت تهیه شده توسط دوربین‌های تعبیه شده بر فراز چهارراه‌ها پیدا می‌شود. در این روش ابتدا محل و اندازه پلاک خودرو با یافتن شماره پلاک خودرو تعیین می‌شود و با توجه به این دو معیار، ناحیه جلوی خودرو به دست می‌آید. برای استخراج ویژگی از این ناحیه، از هرم هیستوگرام گرادیان‌های جهت‌دار^۳ و تبدیل گابور استفاده شده است. تعیین مدل، در این روش طی دو مرحله‌ی طبقه‌بندی صورت می‌گیرد، در مرحله اول چهار طبقه‌بند MLP، KNN، SVM و RF^۴ با هریک از دو ویژگی یاد شده آموزش داده می‌شوند. آنگاه داده آزمون به مرحله اول وارد می‌شود، در این قسمت توابع خطا و هزینه محاسبه می‌شوند و با توجه به آن‌ها داده یا مورد دسته‌بندی قرار می‌گیرد یا رد می‌شود. اگر داده در مرحله اول رد شود به مرحله دوم وارد می‌شود، در مرحله دوم تعدادی طبقه‌بند MLP در نظر گرفته شده است، در این مرحله هم داده آزمون یا مورد دسته‌بندی قرار می‌گیرد یا رد می‌شود. مراحل دوگانه تعیین مدل خودرو در این روش در شکل ۲-۷ نشان داده شده‌اند. متوسط نرخ بازشناسی مربوط به ۲۱ مدل مختلف خودروی در نظر گرفته شده در این روش برابر ۹۸/۷٪ بوده است.

^۱ Sparse representation

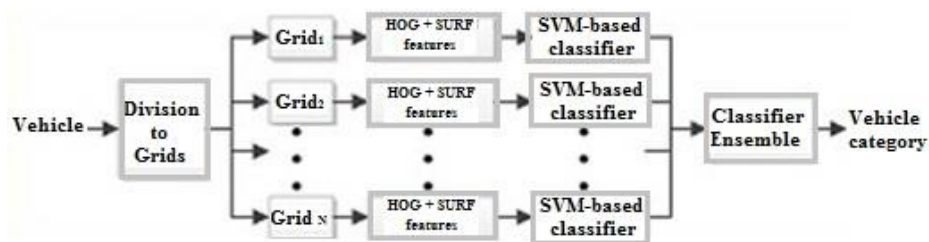
^۲ Hamming distance

^۳ Pyramid Histogram of Oriented Gradients (PHOG)

^۴ Random Forest



الف



ب

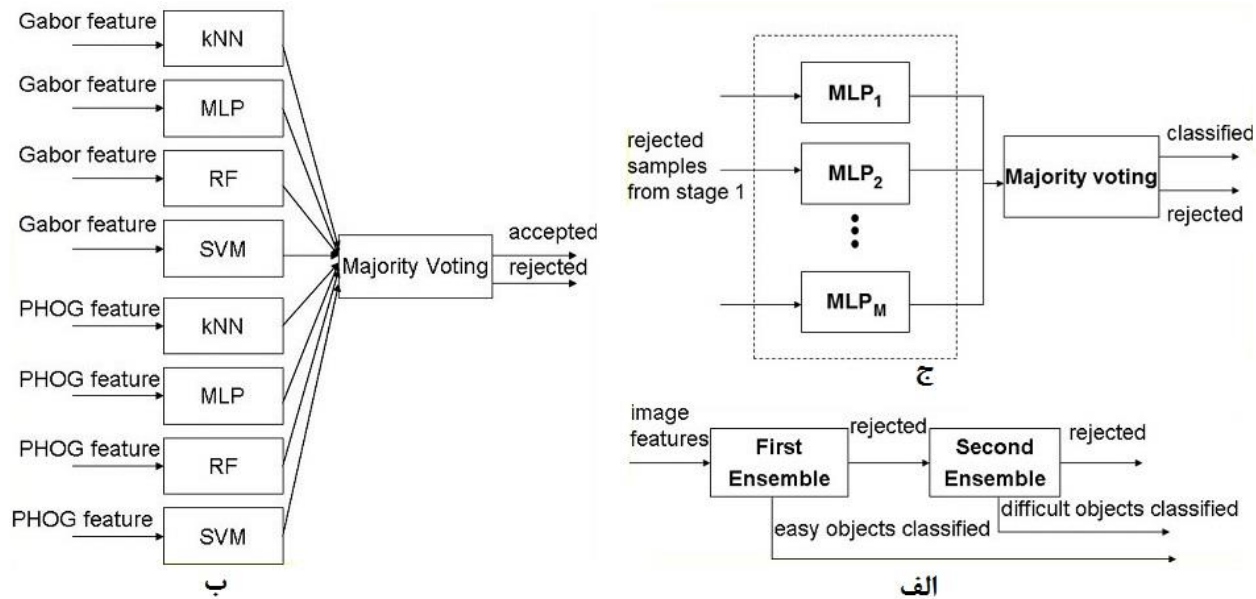
شکل ۲-۶: مراحل الگوریتم‌های الف) تشخیص و ب) تعیین مدل به کمک SSURF مربوط به روش [۱۰]

در [۱۳] هم روشی برای تعیین مدل خودروها با استفاده از تصاویر ثابت گرفته شده توسط دوربین‌های ترافیکی ارائه شده است. ابتدا ناحیه جلوی خودرو با استفاده از آشکارساز مبتنی بر قسمت^۱ پیدا می‌شود. آشکارساز مبتنی بر قسمت، مجموعه‌ای از چند بلوک در کنار هم در نظر گرفته شده است که در هریک از آن‌ها یک یا چند ویژگی استخراج می‌شود و مقادیر ویژگی‌ها در هریک از بلوک‌ها و نیز نسبت مقادیر آن‌ها، وجود یا عدم وجود یک شیء از پیش تعیین شده را مشخص می‌کند. پس از یافتن جلوی خودرو، محل چراغ‌ها و پلاک هم پیدا می‌شوند تا بتوان به کمک آن‌ها اعوجاج افکنشی^۲ احتمالی را اصلاح کرد. پس از چند مرحله اعمال فیلترهای مختلف و نرمال‌سازی، تصویری شامل لبه‌ها و بافت‌های تصویر به دست می‌آید که از آن به عنوان بردار ویژگی استفاده می‌شود. در نهایت مدل خودرو به کمک طبقه‌بندی

^۱ Part-based detector

^۲ Projective distortion

که از تعدادی شبکه عصبی تشکیل شده، تعیین می‌شود. نرخ بازشناسی مربوط به ۳۰ مدل خودروی در نظر گرفته شده در این روش، حدود ۹۹٪ است.

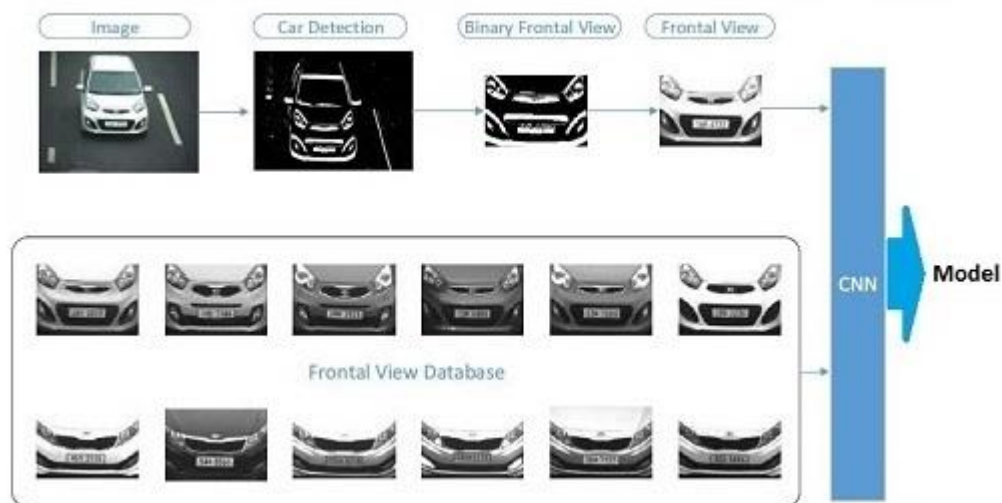


شکل ۲-۷: مراحل الف) مراحل دوگانه تعیین مدل در روش [۱۲] ب) ساختار مرحله اول در روش [۱۲] ساختار مرحله دوم در روش [۱۲]

در روش [۱۴] شبکه عصبی پیچشی^۱ جهت تعیین مدل خودرو به کار رفته است. در این روش ابتدا، یک CNN با تصاویر از پیش تهیه شده‌ی جلوی خودرو آموزش داده می‌شود، سپس برای آزمون، پیش‌زمینه با تفاضل‌گیری از دو فریم متوالی ویدئو به دست می‌آید، ناحیه جلوی خودرو با اعمال فیلتر تقارنی^۲ به تصویر پیش‌زمینه پیدا می‌شود و آن‌گاه مدل خودرو توسط CNN تعیین می‌شود. شکل ۲-۸ مراحل موجود در این روش را نشان داده است، نرخ بازشناسی مدل خودرو در این روش ۸۸٪/۴ بوده است.

^۱ Convolutional Neural Network (CNN)

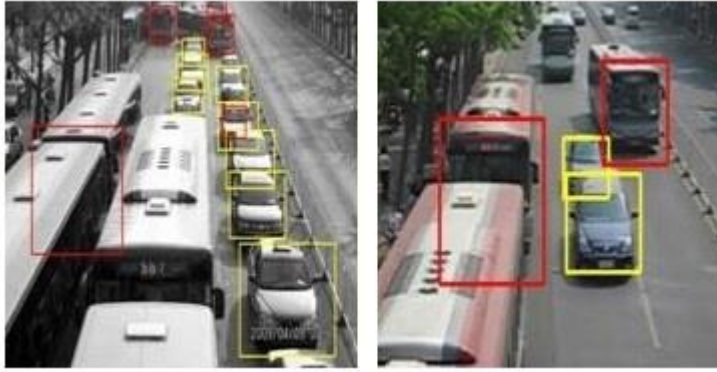
^۲ Symmetrical filter



شکل ۸-۲: مراحل کار در روش [۱۴]

۲-۴ تشخیص و جدا سازی وسیله نقلیه

در [۱۵] از الگوریتم آدابوست و طبقه‌بندهای آبخاری [۱۶] معرفی شده توسط ویولا و جونز برای یافتن خودرو در تصاویر ثابت و ویدئو استفاده شده است. مجموعه‌ای از ۹۰۰۰ تصویر به عنوان داده‌های آموزشی در نظر گرفته شده‌اند که از این تعداد ۳۰۰۰ تصویر مربوط به خودروهای سواری، ۱۰۰۰ تصویر مربوط به اتوبوس و مابقی به عنوان داده‌های منفی هستند. در این تصاویر سعی شده تا حد امکان حالت‌های مختلف برای تصاویر غیر خودرو و نیز حالت‌های انسداد برای تصاویر خودرو در نظر گرفته شده باشد. در شکل ۲-۹ نتایج اجرای الگوریتم نشان داده شده است، در این شکل مستطیل‌های قرمز رنگ اتوبوس و مستطیل‌های زرد رنگ خودروی سواری هستند.



شکل ۹-۲: نتایج تشخیص وسیله نقلیه به کمک الگوریتم Adaboost طبق روش [۱۵]

در [۱۷] از تحلیل چند-تفکیکی^۱ به کمک موجک هار^۲ برای یافتن وسیله نقلیه در تصاویر ثابت استفاده شده است. در این روش سایه زیر خودروی حاضر در صحنه پیدا می‌شود و با در نظر گرفتن نسبتی از ابعاد آن، ناحیه جلوی خودرو پیدا می‌شود. تعیین ROI جلوی خودرو برای تمام تصاویر پایگاه داده انجام می‌شود، ۵۲۹۳ تصویر برای آموزش طبقه‌بند SVM در نظر گرفته شده‌اند که از این میان ۲۳۹۰ تصویر مربوط به خودروها و ۲۹۰۳ تصویر مربوط به غیر خودرو هستند. برای آزمون هم ۴۱۵۷ تصویر خودرو و ۲۳۰۰ تصویر غیر خودرو در نظر گرفته شده است و نرخ بازشناسی ۹۷٪/۱ به دست آمده است. در شکل ۱۰-۲ الف و ب به ترتیب نمونه‌هایی از تصاویر جمع آوری شده مربوط به خودروها و غیر خودروها در این روش نشان داده شده‌اند. در [۱۸] با استفاده از GMM، پیش‌زمینه که متشکل از وسیله نقلیه و سایه آن است جداسازی می‌شود. آن‌گاه به کمک روشی مبتنی بر فضای HSV^۳ ناحیه سایه خودرو هم مشخص می‌شود. نواحی سایه خودرو و نیز کل پیش‌زمینه لبه‌یابی می‌شوند و تفریق لبه‌های این دو ناحیه، خودرو بدون سایه را به دست می‌دهد. در [۱۹] برای یافتن ناحیه جلوی خودرو با توجه به خاصیت توالی موجود در ویدئو، از مدل مخفی مارکوف^۴ استفاده شده است. در این روش ROI‌هایی در هر قاب ویدئو در نظر

^۱ Multi-resolution

^۲ Haar wavelet

^۳ Hue-Saturation-Value (HSV)

^۴ Hidden Markov Model (HMM)

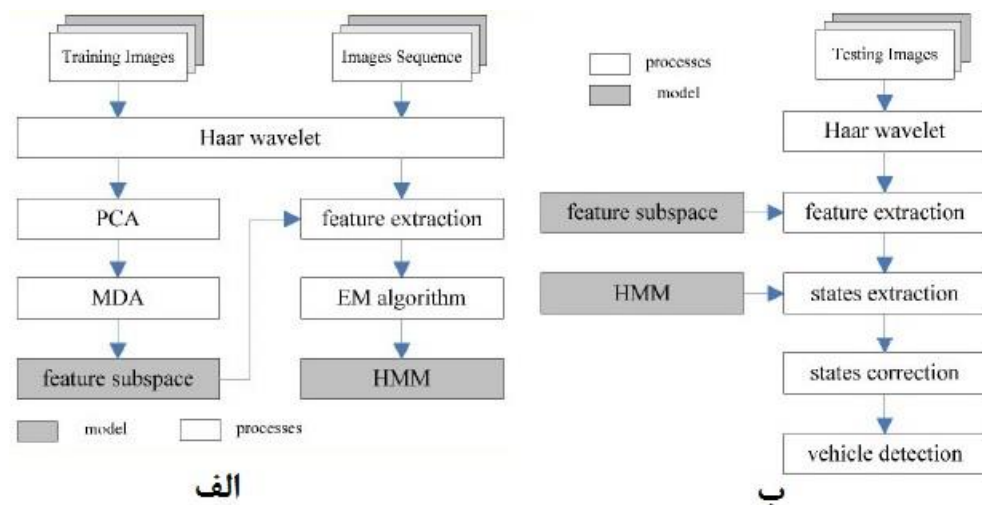
گرفته می‌شود، به کمک موجک هار از این ROI ها استخراج ویژگی می‌شود و PCA و تحلیل مجزا ساز چندگانه^۱ نیز برای کاهش بعد ویژگی‌ها به کار گرفته می‌شود. سپس با استفاده از یک HMM سه حالتی که با تصاویر آموزشی ایجاد شده هر ROI در یکی از سه دسته‌ی جاده، بالای خودرو یا بدنه خودرو جای می‌گیرد. همانطور که در شکل ۱۱-۲ الف نشان داده شده است ابتدا ویژگی‌ها از تصاویر آموزشی استخراج می‌شوند و به کمک PCA و MDA زیرفضایی با ابعاد کمتر ایجاد می‌شود. آن‌گاه در ویدئویی که برای آموزش در نظر گرفته شده از ROI یاد شده استخراج ویژگی می‌شود و ابعاد ویژگی‌ها با بردن آن‌ها به زیر فضا کاهش داده می‌شود. در نهایت با استفاده از الگوریتم EM^۲ یک HMM سه حالتی ساخته می‌شود. در مرحله آزمون همان طور که در شکل ۱۱-۲ ب نشان داده شده است از مدل ساخته شده در مرحله آموزش، برای یافتن جلوی خودرو استفاده می‌شود.



شکل ۱۱-۲ : تصاویر الف) خودروها و ب) غیر خودروها در پایگاه داده روش [۱۷]

^۱ Multiple Discriminate Analysis (MDA)

^۲ Expectation Maximization (EM)

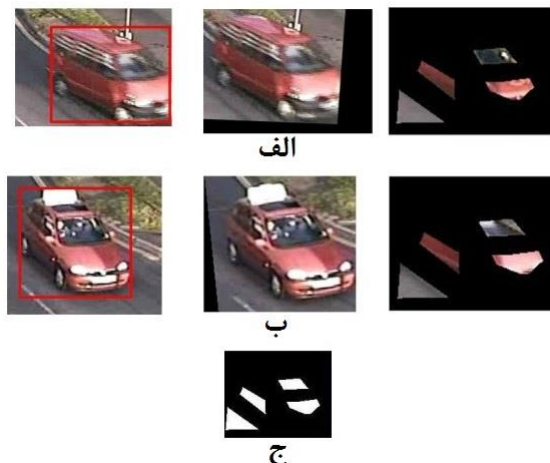


شکل ۲-۱۱: الف) روند آموزش در روش [۱۹] ب) روند آزمون در روش [۱۹]

۲-۵ تشخیص رنگ خودرو

روش [۲۰] پس از یافتن خودرو، همان طور که در شکل ۲-۱۲ نشان داده شده است، از مجموعه‌ای از چند یابنده خودرو وابسته به موقعیت^۱ برای یافتن قسمت‌های رنگی خودرو همچون کاپوت خودرو استفاده می‌کند. موارد الف و ب در این شکل دو خودرو را نشان می‌دهند که بسته به زاویه، به اندازه مناسب چرخش داده شده‌اند و قسمت‌های رنگی آن‌ها با مناسب‌ترین ماسک جدا شده است. همچنین مورد ج هم ماسک وابسته به موقعیت استفاده شده در موارد الف و ب را نشان می‌دهد.

^۱ Pose-dependent vehicle detector



شکل ۲-۱۲: یافتن قسمت‌های رنگی خودرو به کمک ماسک یابنده خودرو وابسته به موقعیت در روش [۲۰]

این روش حاصلضرب شدت روشنایی و مقدار S^1 در هر پیکسل از نواحی یاد شده را به عنوان قدرت رنگ معرفی می‌کند و نواحی مورد نظر با محاسبه دو مقدار متوسط قدرت رنگ^۲ و بیشینه قدرت رنگ و آستانه‌گذاری، به دو دسته رنگی یا غیر رنگی تقسیم می‌شوند. این روش همچنین معیار دیگری تحت عنوان "قدرت رنگ وزن دار شده با مقدار H^3 " را برای تقسیم بندی خودروهای رنگی به دو دسته قرمز و آبی در نظر می‌گیرد، این معیار در رابطه ۱-۲ آورده شده است، در این رابطه CS_i قدرت رنگ، Hue_i مقدار H در پیکسل i ام از پیکسل‌های کاپوت خودرو هستند، همچنین N تعداد پیکسل‌های کاپوت خودرو است.

$$CS \text{ weighted Hue} = \frac{\sum_{i=1}^N CS_i Hue_i}{\sum_{i=1}^N CS_i} \quad (1-2)$$

سپس خودروهای آبی یا قرمز تشخیص داده شده با در نظر گرفتن متوسط روشنایی پیکسل‌ها و متوسط قدرت رنگشان به دو دسته تاریک و روشن تقسیم‌بندی می‌شوند. خودروهای غیررنگی با توجه به شدت

^۱ Saturation

^۲ Chromatic Strength (CS)

^۳ CS weighted Hue

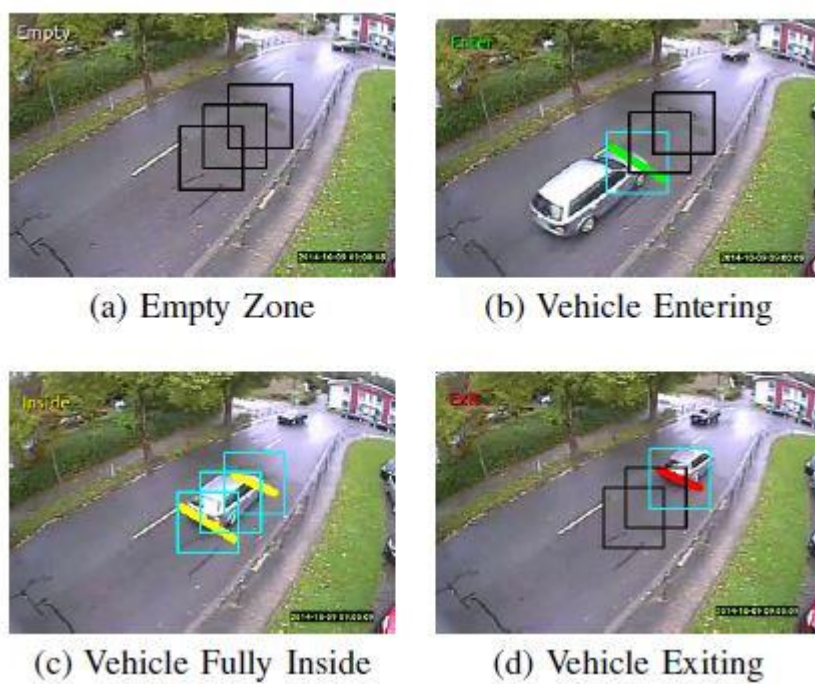
روشنایی به دو دسته روشن و مشکی تقسیم‌بندی می‌شوند و خودروهای غیررنگی روشن هم با توجه به واریانس روشنایی در دو دسته نقره‌ای و سفید جای می‌گیرند. روش [۲۱] به نحوی همان روش [۱۰] است که در بخش تشخیص و طبقه‌بندی مدل خودرو به آن پرداخته شد با این تفاوت که بخش تشخیص رنگ به آن اضافه شده است. در این روش پس از استخراج ناحیه جلوی خودرو، مشابه با روش [۲۰] یک فرآیند مرحله به مرحله برای تعیین رنگ در نظر گرفته شده است و با در نظر گرفتن معیار قدرت رنگ بیان شده در رابطه ۲-۲ خودروها به دو دسته رنگی و غیر رنگی تقسیم شده‌اند، در این رابطه R_{avg} ، G_{avg} و B_{avg} متوسط به ترتیب مولفه‌های قرمز، سبز و آبی پیکسل‌های ناحیه جلوی خودرو هستند. سپس با استخراج بردارهای ویژگی مختلف، در فضاها رنگ RGB، LAB و ... از داده‌های آموزشی و آموزش طبقه‌بندهای مختلف، خودروهای رنگی به چهار گروه قرمز، سبز، آبی و زرد و خودروهای غیررنگی هم به پنج گروه مشکی، سفید، نقره‌ای، آبی تیره و سبز تیره دسته‌بندی شده‌اند. در [۲۲] ابتدا پلاک خودرو پیدا می‌شود، ناحیه‌ای در کنار پلاک خودرو جدا می‌شود، سپس با توجه به متوسط مقدار H پیکسل‌های ناحیه یاد شده رنگ خودرو تشخیص داده می‌شود.

$$CS = \left(\left(\frac{R_{avg}}{255} - \frac{1}{3} \right)^2 + \left(\frac{G_{avg}}{255} - \frac{1}{3} \right)^2 + \left(\frac{B_{avg}}{255} - \frac{1}{3} \right)^2 \right) + (\max\{R_{avg}, G_{avg}, B_{avg}\} - \min\{R_{avg}, G_{avg}, B_{avg}\}) \quad (2-2)$$

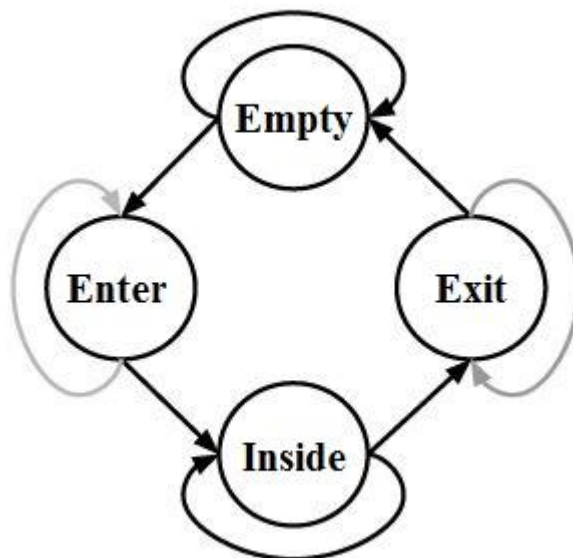
۲-۶ تشخیص جهت خودرو

در [۲۳] برای تشخیص عبور خودرو سه ناحیه از پیش تعیین شده در نظر گرفته شده است و برای هریک از این نواحی سه طبقه‌بند مبتنی بر ویژگی‌های هار با استفاده از تعداد زیادی نمونه آموزشی تهیه شده است و هریک از این طبقه‌بندها حضور یا عدم حضور خودرو در ناحیه متناظرشان را اعلام می‌کنند. یک مدل مخفی مارکوف هم برای تعیین ورود یا خروج خودرو ایجاد شده است که با توجه به حضور یا عدم

حضور در نواحی سه گانه یاد شده در نظر گرفته شده است. در شکل ۱۳-۲ سه ناحیه مربوطه و نیز حالات ممکن برای حضور خودرو در آن‌ها نشان داده شده است و در شکل ۱۴-۲ نمودار حالات مربوط به مدل مخفی مارکوف در نظر گرفته شده در این روش نشان داده شده است.



شکل ۱۳-۲: سه ناحیه از پیش تعیین شده و حالات ممکن برای حضور خودرو در آن نواحی در روش [۲۳]



شکل ۲-۱۴: نمودار حالات مربوط به مدل مخفی مارکوف در نظر گرفته شده در روش [۲۳]

در [۲۴] پیش‌زمینه به کمک روش GMM و سپس میدان سرعت^۱ در نواحی پیش‌زمینه به کمک الگوریتم شار نوری^۲ به صورت چگال^۳ استخراج می‌شوند. پیکسل‌های موجود در میدان سرعت و واقع بر یک خط مجازی در تصویر، بسته به سرعتشان و با در نظر گرفتن آستانه‌ها و قیدهایی به دو گروه عابر پیاده و خودرو تقسیم می‌شوند، جهت حرکتشان تشخیص داده شده و نیز تعداد هر گروه شمارش می‌شود. شکل ۲-۱۵ نمونه قاب ویدئوی مورد استفاده و میدان سرعت به دست آمده برای یک خودرو را نشان می‌دهد.

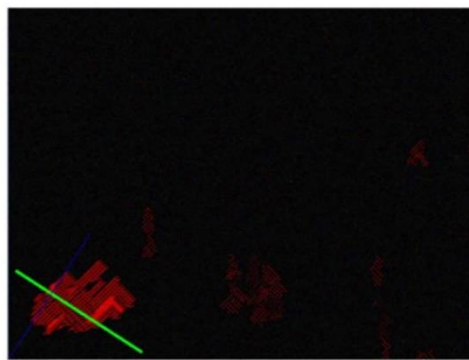
^۱ Velocity Field

^۲ Optical Flow (OF)

^۳ Dense



الف



ب

شکل ۲-۱۵: الف) خودروی عبوری از خط مجازی در روش [۲۴] ب) میدان سرعت خودروی مورد (الف)

۳. مبانی نظری

۳-۱ مقدمه

پیش از پرداختن به روش پیشنهادی نیاز است تا اندکی در مورد ابزارها و روش‌های موجود و با اهمیت، که به نحوی در روش پیشنهادی از آن‌ها بهره گرفته شده است، بحث شود. در این فصل، ابتدا پیش-پردازشی موسوم به نادرهم‌بافی^۱ مورد بررسی قرار می‌گیرد. سپس ویژگی HOG و طبقه‌بند ماشین بردار پشتیبان و نیز توصیفگر SURF به اختصار بررسی می‌شوند.

۳-۲ نادرهم‌بافی

برای درک عمل نادرهم‌بافی ابتدا بایستی مفهوم درهم‌بافی^۲ شرح داده شود. درهم‌بافی ابتدا در دهه‌ی ۱۹۳۰ میلادی در نمایش‌گرهای مبتنی بر لامپ پرتوی کاتدی^۳ همچون تلویزیون‌های آنالوگ، برای کاهش پهنای باند مورد نیاز برای ارسال و نیز حذف نویز فلیکر^۴ انجام شد. در واقع در سیستم تصویربرداری درهم‌بافته تصاویر به صورت خط به خط خوانده می‌شوند به این ترتیب که ابتدا خطوط افقی فرد تصاویر خوانده می‌شوند و در دامنه‌ای از داده^۵ ذخیره می‌شوند و سپس خطوط افقی زوج خوانده می‌شوند. به دامنه‌ی داده ناشی از خطوط فرد دامنه‌ی فوقانی و به دامنه‌ی داده ناشی از خطوط زوج دامنه‌ی تحتانی می‌گویند، ترکیب این دو دامنه یک فریم از تصویر را ایجاد می‌کند. با این کار دو دامنه‌ی مجزا به وجود می‌آید که هریک از آن‌ها دنباله‌ای از تصاویر با نرخ قابی^۶ برابر نصف نرخ قاب اصلی هستند و لذا ارسال آن‌ها به پهنای باند کم‌تری نیاز خواهد داشت، در برخی دوربین‌های تصویربرداری هم این اثر مشاهده می‌شود. در شکل ۳-۱ نمونه‌ای از دو دامنه‌ی تشکیل دهنده‌ی یک قاب نشان داده شده است.

^۱ De-interlacing

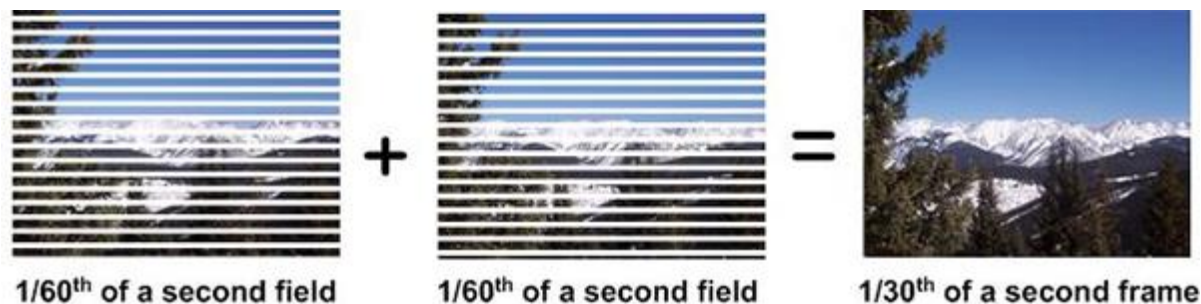
^۲ Interlacing

^۳ Cathode Ray Tube (CRT)

^۴ Flicker noise

^۵ Data Field

^۶ Frame rate



شکل ۳-۱: ترکیب دو دامنه‌ی تشکیل دهنده‌ی یک قاب

معایب درهم بافی:

به دلیل ماهیت پیوسته بودن نمایش‌گرهای آنالوگ، این دامنه‌ها که به صورت پشت سر هم می‌آیند در این نوع نمایش‌گرها از نظر دیداری به صورت مناسبی دیده می‌شوند اما در نمایش‌گرهای دیجیتال، این تصاویر همراه با اثری به نام پله‌ای شدن^۱ مشاهده می‌شود، این اثر باعث می‌شود خطوط مورب موجود در تصویر بصورت دندان‌دندانه دیده شوند. این اشکال هنگامی پدید می‌آید که از یک سوژه با سرعت بالا تصویر می‌گیریم، این اتفاق به این علت رخ می‌دهد که هنگام ضبط خطوط فرد و ذخیره شدن در دامنه‌ی فوقانی، سوژه در مکان الف قرار دارد، ولی هنگام ضبط دامنه تحتانی سوژه به علت گذشت زمان به مکان ب منتقل می‌شود. چون سرعت سوژه بالا است جابه‌جایی در زمان کوتاه بین ضبط دو میدان قابل توجه بوده و باعث می‌شود تا تصویر به صورت دندان‌دندانه دیده شود. اشکال بعدی اثر ارتعاش^۲ است که هنگامی به وجود می‌آید که از خطوطی افقی که در تصویر به سمت بالا یا پایین می‌روند، تصویربرداری می‌شود، در این حالت حرکت این خطوط با لرزش همراه خواهند بود.

^۱ Aliasing

^۲ Twitter

نادرهم بافی:

آثار ناخواسته‌ی یاد شده باعث می‌شوند مرز اشیاء موجود در تصاویر بصورت مبهم در بیاید و توصیف آن‌ها با استفاده از توصیفگرهایی همچون HOG مشکل شود. به همین علت بایستی برای رفع این آثار، بر روی ویدئوها عمل نادرهم‌بافی انجام گیرد. روش‌های موجود برای این کار به سه دسته‌ی کلی تقسیم می‌شوند:

(۱) ترکیب دامنه‌ها، که با جمع دامنه‌ها، متوسط‌گیری آن‌ها و... صورت می‌گیرد. (۲) گسترش دامنه‌ها که روشی است که در آن خطوط خالی در هریک از دامنه‌های افقی و عمودی یا حذف می‌شوند و رزولوشن کاهش پیدا می‌کند و یا خطوط خالی هر دامنه با خطوط غیر خالی آن پر می‌شوند. (۳) تشخیص حرکت که در آن علاوه بر به‌کارگیری دو روش پیشین، میزان و جهت حرکت‌های موجود در ویدئو هم تشخیص داده می‌شوند تا ترکیب یا گسترش دامنه‌ها بهتر انجام شود. در شکل ۲-۳ الف و ب به ترتیب تصاویر درهم‌بافته و نادرهم‌بافته نشان داده شده‌اند.



ب

الف

شکل ۲-۳: تصاویر درهم‌بافته (الف) و نادرهم‌بافته (ب)

۳-۳ هیستوگرام گرادیان‌های جهت‌دار

توصیفگر هیستوگرام گرادیان‌های جهت‌دار اولین بار در سال ۲۰۰۵ توسط دلال^۱ و تریگز^۲ به منظور تشخیص عابر پیاده و بر روی تصاویر ثابت معرفی شد. هدف اصلی از ارائه‌ی چنین توصیفگری، توصیف شکل ظاهری اشیاء موجود در تصویر به صورت محلی و با استفاده از شدت و جهت لبه‌های موجود در آن است. مراحل پیاده‌سازی الگوریتم HOG به صورت زیر می‌باشد:

(۱) ابتدا مقادیر گرادیان برای هر پیکسل تصویر محاسبه می‌شود. معمول‌ترین کرنل‌های مورد استفاده برای این کار، دو کرنل یک بعدی هستند که در رابطه ۳-۱ آورده شده‌اند، D_x کرنل افقی و D_y کرنل عمودی است.

$$D_x = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}, \quad D_y = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad (۱-۳)$$

حاصل این کار دو تصویر گرادیان افقی (I_x) و گرادیان عمودی (I_y) خواهد بود. سپس مقادیر زاویه‌ی گرادیان هم برای هر پیکسل تصویر با استفاده از رابطه ۳-۲ محاسبه می‌شود، این مقادیر بسته به این که گرادیان بی‌علامت^۳ مد نظر باشد یا گرادیان علامت‌دار^۴، می‌توانند به ترتیب بین ۰ و ۱۸۰ درجه یا ۰ و ۳۶۰ درجه باشند. بازه‌های یاد شده به تعداد دلخواه B (معمولاً ۹) بازه تقسیم می‌شوند.

$$\theta(x, y) = \arctan\left(\frac{I_y(x, y)}{I_x(x, y)}\right) \quad (۲-۳)$$

^۱ Navneet Dalal

^۲ Bill Triggs

^۳ Unsigned gradient

^۴ Signed gradient

(۲) سپس تصویر به نواحی ای تحت عنوان سلول^۱ به ابعاد $c \times c$ پیکسل تقسیم بندی می شود و در هر سلول، هیستوگرامی از مقادیر زاویه ای نرمال شده محاسبه می شود. با محدود کردن تعداد بازه ها ممکن است برخی پیکسل ها با وجود نزدیک بودن مقدارشان به دو بازه، تنها به یکی از آن ها رأی دهند، لذا برای پیش گیری از این امر، در محاسبه ی هیستوگرام، هر پیکسل به دو بازه با مقادیر نزدیک به مقدار خود رأی می دهد. در این مرحله به ازای هر سلول یک بردار B عضوی به نام c_i خواهیم داشت. برای این که اثر تغییرات روشنایی^۲ و تبا^۳ در تصویر حذف شود باید هیستوگرام های سلول ها به صورت محلی نیز نرمال شوند. برای این کار تعداد مشخصی از سلول های مجاور هم در یک بلوک^۴ به ابعاد $b \times b$ سلول گروه بندی می شوند، به طوری که هر بلوک با بلوک های مجاور خود هم پوشانی دارد. هیستوگرام های همه ی سلول های یک بلوک به صورت متوالی در کنار هم قرار داده می شوند و بردارهای v_i را می سازند. سپس این بردارها با استفاده از معیار نرمال سازی L_1 ^۵، طبق رابطه ۳-۳ نرمال می شوند. به جای معیار یاد شده از سایر معیارهای رایج مثل معیار نرمال سازی L_2 ^۶ یا معیار ریشه ی دوم L_1 ^۷ هم می توان استفاده نمود.

$$v_i \leftarrow \frac{v_i}{\sqrt{\|v_i\|^2 + \epsilon^2}} \quad (3-3)$$

در نتیجه هر بلوک یک بردار هیستوگرام $b \times b \times B$ عضوی خواهد داشت. فرض کنیم تصویر دارای N بلوک باشد (مقدار N به ابعاد تصویر بستگی دارد)، پس برای هر تصویر، N هیستوگرام بلوکی به طول $b \times b \times B$ خواهیم داشت. باز هم همه ی N هیستوگرام بلوکی را به طور متوالی کنار هم قرار می دهیم و نرمال

^۱ Cell

^۲ Illumination

^۳ Contrast

^۴ Block

^۵ L1-norm

^۶ L2-norm

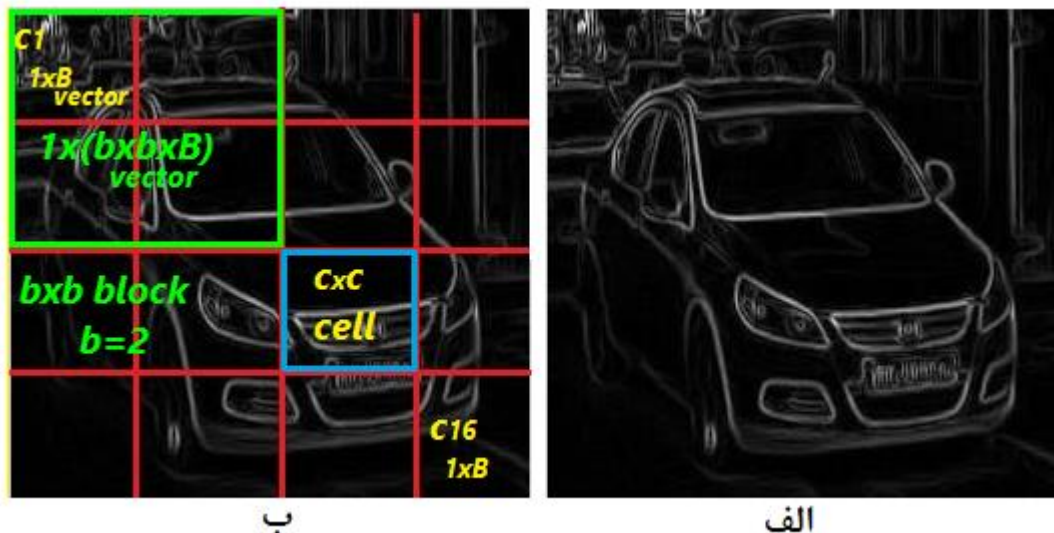
^۷ L1-square root (L1-sqrt)

می‌کنیم تا بردار HOG نهایی تصویر به ابعاد $b \times b \times B \times N$ با نام H به دست آید. نرمال‌سازی این بردار طبق رابطه ۳-۴ صورت می‌گیرد، در این رابطه τ یک آستانه است تا تأثیر مقادیر بزرگ گرادیان را حذف کند و می‌تواند برابر 0.2 اختیار شود و h_n هر یک از N عضو H است.

$$\begin{aligned} H &\leftarrow \frac{H}{\sqrt{\|H\|^2 + \epsilon^2}} \\ h_n &\leftarrow \min(h_n, \tau) \\ H &\leftarrow \frac{H}{\sqrt{\|H\|^2 + \epsilon^2}} \end{aligned} \quad (4-3)$$

شایان ذکر است که بلوک‌ها می‌توانند به دو شکل مستطیلی و دایره‌ای در نظر گرفته شوند. شکل ۳-۳ الف تصویر گرادیان و شکل ۳-۳ ب موقعیت سلول‌ها و بلوک‌ها را در حالت مستطیلی برای یک تصویر نوعی نشان می‌دهند. الگوریتم HOG بر روی سلول‌ها و بلوک‌های محلی محاسبات را انجام می‌دهد لذا HOG از جمله ویژگی‌هایی است که در برابر تبدیلات هندسی و نوری^۱ مقاوم است، اما در برابر تغییرات جهت اشیاء مقاوم نیست، لذا از HOG در صورتی می‌توان برای تشخیص شیء بهره برد که شیء در همه تصاویر با زاویه‌ی ثابت و یکسانی ظاهر شده باشد. بردارهای HOG را می‌توان به عنوان داده‌های ورودی SVM، شبکه‌ی عصبی و یا هر طبقه‌بند دیگری به کار برد.

^۱ Photometric



شکل ۳-۳: موقعیت سلول‌ها و بلوک‌ها در الگوریتم محاسبه‌ی HOG

۳-۴ ماشین بردار پشتیبان

الگوریتم اولیه‌ی ماشین بردار پشتیبان در سال ۱۹۶۳ میلادی توسط وِپنایک^۱ و چرونینکیش^۲ ابداع شد. این روش از جمله روش‌های نسبتاً جدید در میان روش‌های تشخیص الگو^۳ است که در سال‌های اخیر کارایی خوبی نسبت به روش‌های قدیمی‌تر برای طبقه‌بندی از جمله شبکه‌های عصبی مصنوعی^۴ از جمله پرسپترون^۵ نشان داده است. مبنای کاری طبقه‌بند SVM دسته‌بندی خطی داده‌ها است و در تقسیم‌بندی خطی داده‌ها سعی می‌شود خطی انتخاب شود که حاشیه‌ی اطمینان بیشتری با داده‌های مرزی داشته باشد. برای تفکیک داده‌هایی که جداپذیر خطی^۶ باشند می‌توان از این طبقه‌بند استفاده کرد اما در برخی موارد که داده‌ها جداپذیر خطی نیستند ممکن است بتوان به کمک توابع غیرخطی آن‌ها را به فضاهای با

^۱ Vladimir N. Vapnik

^۲ Alexey Ya. Chervonenkis

^۴ Pattern recognition

^۵ Artificial neural networks

^۶ Perceptron

^۷ Linearly separable

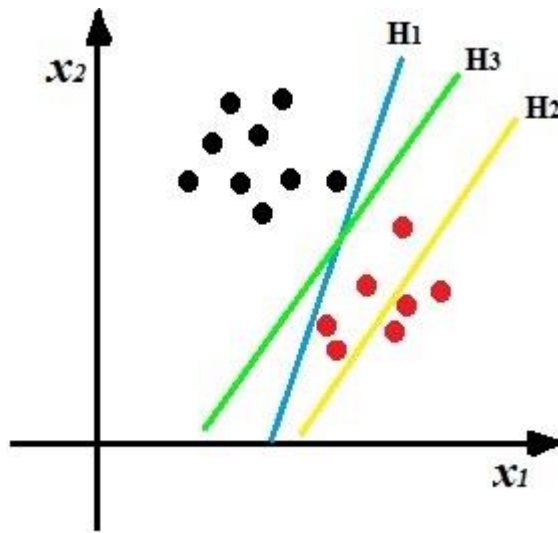
ابعاد بالاتر برد تا در آن جا به صورت خطی جداپذیر باشند. در ادامه به صورت مختصر به توضیح محاسبات ریاضی مربوط به این الگوریتم پرداخته شده است. فرض کنیم مجموعه‌ای به نام D داریم از n داده، می‌خواهیم آن‌ها را بوسیله‌ی یک ابرصفحه^۸ به دو دسته‌ی مجزا تقسیم کنیم به صورتی که داده‌های واقع در مرز بیش‌ترین فاصله‌ی ممکن را از ابرصفحه داشته باشند. به عنوان یک مثال در فضای دو بعدی، در شکل ۳-۴ سه مرز ممکن نشان داده شده‌اند، خط زرد رنگ (H_2) داده‌های قرمز و سیاه را به درستی از هم جداسازی نمی‌کند، خط آبی رنگ (H_1) داده‌ها را به درستی جدا کرده ولی بیش‌ترین حاشیه‌ی اطمینان را بین خود و داده‌های مرزی هر دو دسته ایجاد نکرده، اما خط سبز رنگ (H_3) علاوه بر جداسازی داده‌ها به صورت صحیح، بیش‌ترین حاشیه‌ی اطمینان را هم ایجاد کرده، لذا در الگوریتم SVM، هدف یافتن چنین مرزی است. هر ابرصفحه را می‌توان به صورت رابطه ۳-۵ نمایش داد، در این رابطه W بردار یک‌ه‌ی عمود بر ابرصفحه است و b یک عدد ثابت است.

$$\vec{W} \cdot \vec{X} - b = 0 \quad (5-3)$$

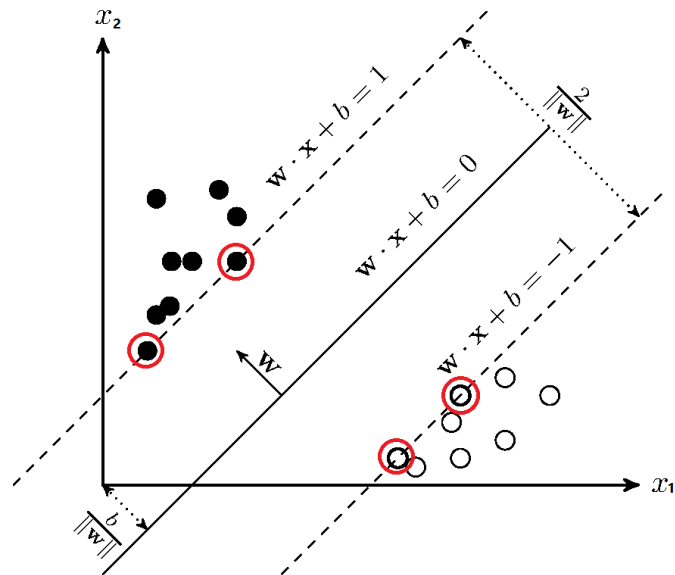
تعدادی از داده‌های هر دسته که در بیرونی‌ترین مکان و بر روی یک ابرصفحه قرار دارند بردارهای پشتیبان نامیده می‌شوند، این ابر صفحه‌ها را می‌توان به صورت رابطه ۳-۶ نشان داد. حول بردارهای پشتیبان در شکل ۳-۵ دایره‌هایی کشیده شده است، خطوط مرزی (ابرصفحه‌های مرزی در فضاهای با ابعاد بالاتر) هم با خط‌چین نمایش داده شده‌اند.

$$\vec{W} \cdot \vec{X} - b = 1 \quad , \quad \vec{W} \cdot \vec{X} - b = -1 \quad (6-3)$$

⁸ Hyper-plane



شکل ۳-۴: سه مرز ممکن برای جداسازی داده‌ها به صورت خطی



شکل ۳-۵: مرزبندی‌های موجود در الگوریتم SVM در فضای دو بعدی

اثبات می‌شود که فاصله‌ی دو صفحه‌ی مرزی برابر با $\frac{2}{\|W\|}$ است و برای این که بیش‌ترین حاشیه را داشته باشیم لازم است این فاصله بیشینه شود یا به صورت معادل $\|W\|$ کمینه شود. برای اینکه از ورود داده‌ها به حاشیه جلوگیری کنیم، شرایط موجود در رابطه ۷-۳ را اضافه می‌کنیم.

$$\forall x \in D_{-1} \quad \overrightarrow{W} \cdot \vec{x} - b \leq -1 \quad , \quad \forall x \in D_{+1} \quad \overrightarrow{W} \cdot \vec{x} - b \geq +1 \quad (7-3)$$

D_{-1} و D_{+1} به ترتیب مجموعه داده‌های دسته‌های اول و دوم هستند. اگر هریک از داده‌های گروه اول و دوم را به ترتیب با برچسب‌های $y_i = -1$ و $y_i = +1$ در نظر بگیریم به مسئله‌ی بهینه‌سازی رابطه ۸-۳ می‌رسیم.

$$\begin{aligned} & \min_{w,b} \{ \|W\| \} \\ & s. t \\ & \forall x_i \in D \quad , \quad i = 1, \dots, n \quad , \quad y_i (\overrightarrow{W} \cdot \vec{x} - b) \geq 1 \end{aligned} \quad (8-3)$$

بایستی W و b را بیابیم که مسئله‌ی بهینه‌سازی را حل کنند، با توجه به این موضوع که می‌توان به جای $\|W\|$ ، $\frac{1}{2} \|W\|^2$ را کمینه کرد، با استفاده از ضرایب نامنفی لاگرانژ^۱ می‌توان مسئله را به صورت رابطه ۹-۳ بیان کرد.

$$\min_{w,b} \left\{ \max_{\alpha} \left\{ \frac{1}{2} \|W\|^2 - \sum_{i=1}^n \alpha_i [y_i (\overrightarrow{W} \cdot \vec{x}_i - b) - 1] \right\} \right\} \quad (9-3)$$

این مسئله به روش برنامه‌ریزی درجه دوم^۱ حل می‌شود و تعدادی α به دست می‌آید، سپس به کمک آن‌ها، همان طور که در روابط ۱۰-۳ و ۱۱-۳ آمده، W و b و در نتیجه ابرصفحه‌ی بهینه پیدا می‌شود، N_{sv} در رابطه ۱۱-۳ تعداد بردارهای پشتیبان است.

¹ Non-negative Lagrange multipliers

$$w = \sum_{i=1}^n \alpha_i y_i x_i \quad (۱۰-۳)$$

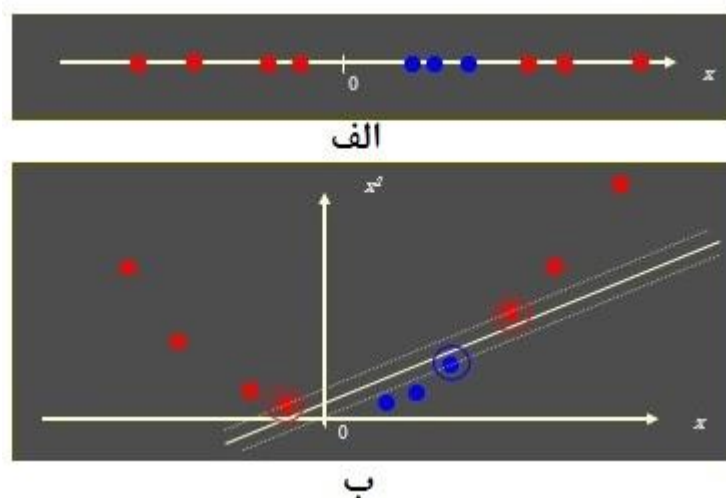
$$b = \frac{1}{N_{sv}} \sum_{i=1}^{N_{sv}} (w \cdot x_i - y_i) \quad (۱۱-۳)$$

از جمله اشکالات SVM می‌توان به این موارد اشاره کرد: تمام داده‌های مورد استفاده برای آموزش باید به صورت دستی برچسب زده شده باشند، پارامترهای به دست آمده برای یک دسته داده‌ی آموزشی به راحتی قابل تفسیر نیست و همچنین این که الگوریتم SVM به طور اختصاصی برای مسئله‌ی دوکلاسه طراحی شده و لذا برای حل یک مسئله‌ی چندکلاسه نیاز است به ازای هر جفت کلاس داده یک SVM آموزش داده شود. اشکال دیگری هم که در مورد SVM مطرح است این است که مرز جداکننده در SVM در مقایسه با روش‌های احتمالاتی ثابت است، یعنی اگر داده‌ی جدیدی برای آزمایش وارد شود حتماً با احتمال کامل عضو یکی از کلاس‌ها تشخیص داده می‌شود و این امر در برخی حوزه‌ها همچون پزشکی که در آن‌ها تصمیم‌گیری‌ها براساس نتایج احتمالاتی صورت می‌پذیرد، مطلوب نیست، اما برای رفع این مشکل در [۲۵] روشی ارائه شده که طبق آن علاوه بر تخصیص برچسب تعلق نمونه به یکی از دو گروه، میزان احتمال درستی برچسب تعلق داده شده به نمونه هم محاسبه می‌شود و بر همین اساس کار طبقه‌بندی چندکلاسه با SVM با دقت بیشتری قابل انجام است. در ابتدای بحث به این موضوع اشاره شد که در صورتی که داده‌ها به صورت خطی جداپذیر نباشند ممکن است بتوان آن‌ها را با توابع مختلف به فضاهایی با ابعاد بالاتر برد تا در آن‌جا به صورت خطی جداپذیر شوند و بتوان الگوریتم SVM را بر روی آن‌ها پیاده کرد، برای مثال می‌توان از کرنل‌های پرکاربردی همچون خطی، سیگموئید، تابع پایه شعاعی^۲ و

² Quadratic Programming (QP)

¹ Radial basis function (RBF)

چند جمله‌ای اشاره کرد. شکل ۳-۶ نشان می‌دهد چگونه ممکن است یک دسته داده که در فضای یک بعدی بصورت خطی جداپذیر نیستند، پس از انتقال به فضای دو بعدی به کمک تابع درجه دوم به صورت خطی جداپذیر می‌شوند.



شکل ۳-۶ : (الف) داده‌های آبی و قرمز در فضای یک بعدی به صورت خطی جداپذیر نیستند (ب) جداپذیری خطی داده‌های دو گروه پس از انتقال به فضای دوبعدی با تابع درجه دوم

در صورت استفاده از کرنل به همراه SVM نیاز است تا پارامترهای آن هم تعیین شوند، در جدول ۳-۱ فرمول‌های چهار کرنلی که پیش از این از آن‌ها نام برده شد و همچنین پارامترهای مهم آن‌ها برای داده نوعی $X=[x_i, x_j]^T$ آورده شده‌اند C مستقل از به کارگیری یا عدم به کارگیری کرنل و پارامتری است که بین خطا و بیش‌برازش^۱ مصالحه برقرار می‌کند، به این معنا که با افزایش C میزان خطای طبقه‌بند کم می‌شود اما اگر C بسیار بزرگ اختیار شود باعث خواهد شد که طبقه‌بند قدرت تعمیم^۲ خود را از دست بدهد و با ورود داده جدید خطا افزایش یابد.

^۱ Overfitting

^۲ Generality

جدول ۳-۱: فرمول‌های چهار کرنل مهم مورد استفاده همراه با SVM و پارامترهای مهم آنها

پارامترهای قابل تنظیم	فرمول کرنل	کرنل
C	$K(x_i, x_j) = x_i^T \cdot x_j$	خطی
C, γ , coef0	$K(x_i, x_j) = \tanh(\gamma \cdot x_i^T \cdot x_j + \text{coef0}), \gamma > 0$	سیگموئید
C, γ	$K(x_i, x_j) = e^{-\gamma(\ x_i - x_j\ ^2)}, \gamma > 0$	RBF
C, γ , coef0, degree	$K(x_i, x_j) = (\gamma \cdot x_i^T \cdot x_j + \text{coef0})^{\text{degree}}, \gamma > 0$	چند جمله‌ای

۳-۵ ویژگی‌های تسریع شده مقاوم

SURF، اولین بار توسط هربرت بی^۱ و چند تن دیگر در سال ۲۰۰۶ ارائه شد. این تبدیل عمدتاً برگرفته از تبدیل SIFT^۲ است اما از لحاظ زمانی بسیار به صرفه‌تر و نیز نسبت به تبدیل‌های هندسی و چرخش بسیار مقاوم‌تر است. در بینایی ماشین از این تبدیل به منظور تشخیص، طبقه‌بندی و نیز بازسازی سه بعدی اشیاء استفاده می‌شود. بازشناسی اشیاء به کمک این توصیفگر طی سه مرحله کلی انجام می‌گیرد:

(۱) یافتن نقاط مورد علاقه^۳ (۲) توصیف همسایگی محلی و (۳) تطبیق SURF برای یافتن نقاط کلیدی از ماتریس هسین^۴ استفاده می‌کند. آنچه توصیفگر SURF را نسبت به بسیاری دیگر از توصیفگرها سریع‌تر ساخته است، استفاده از تصویر انتگرال^۵ برای محاسبه ماتریس هسین است. رابطه ۳-۱۲ این ماتریس را نشان می‌دهد، در این رابطه درایه‌های ماتریس، مشتقات جزئی مرتبه دوم تصویر سطح خاکستری، p نشان دهنده مختصات نقطه مورد بررسی و σ انحراف معیار استاندارد هستند، مقادیر مختلف σ ، معادل با اندازه‌های مختلف پنجره‌ها در محاسبه مشتقات دوم است.

^۱ Herbert Bay

^۲ Scale-Invariant Feature Transform (SIFT)

^۳ Interest points

^۴ Hessian matrix

^۵ Integral image

$$H(p, \sigma) = \begin{pmatrix} L_{xx}(p, \sigma) & L_{xy}(p, \sigma) \\ L_{xy}(p, \sigma) & L_{yy}(p, \sigma) \end{pmatrix} \quad (۱۲-۳)$$

همان طور که گفته شد، تصویر انتگرال محاسبه ماتریس هسین را سریع تر می کند، مقدار هر پیکسل در مختصات $p(x,y)$ از این تصویر برابر است با جمع شدت های روشنایی پیکسل های با مقادیر عرض و ارتفاع کمتر از x و y . لذا چنین تصویری امکان انجام جمع ها و تفاضل های لازم برای محاسبه ماتریس هسین را مستقل از اندازه پنجره انتخابی و با سرعت بیش تری فراهم می آورد. نقاطی که در آن ها دترمینان ماتریس هسین بیشینه است به عنوان نقاط کلیدی انتخاب می شوند، همچنین مقیاسی که در آن نقطه کلیدی یافت شده است ذخیره می شود. توصیفگر مربوطه هم بر اساس نحوه توزیع پاسخ های موجک ها^۱ حول نقطه کلیدی و انتخاب بردارهایی با اندازه ۶۴ تایی (برای بهینه سازی هزینه زمانی) ایجاد می شود. برداری نیز برای تعیین جهت توصیفگر محاسبه می شود: ناحیه ای دایره ای حول نقطه کلیدی و با شعاعی متناسب با مقیاسی که نقطه کلیدی در آن یافت شده در نظر گرفته می شود و در آن ناحیه، پاسخ های موجک ها در جهت های x و y (باز هم متناسب با مقیاس مربوطه) و با استفاده از تصویر انتگرال محاسبه می شود. نقاطی از ناحیه دایره ای که قدرت موجک ها در آن ها غالب است در نظر گرفته می شوند و بردارهایی که ابتدای آن ها نقاط کلیدی و انتهای آن ها نقاط غالب است محاسبه می شوند: جهت غالب از برآیندگیری جهت بردارهای مذکور حاصل می شود. این جهت یافته شده با وجود تغییرات در جهت تصویر، باز هم قابل محاسبه هستند و همین باعث می شود این توصیفگر نسبت به چرخش مقاوم باشد.

^۱ Haar wavelet

۳-۶ شار نوری

به الگویی از حرکت ظاهری اشیاء، سطوح و لبه‌های موجود در صحنه که بر اساس حرکت نسبی بین ناظر (دوربین یا چشم) و صحنه شکل می‌گیرد، شار نوری^۱ گفته می‌شود. چنین مفهومی اولین بار توسط یک روانشناس آمریکایی به نام جیمز گیbson در دهه ۱۹۴۰ میلادی به منظور تشریح محرک‌های بصری حیوانات معرفی شد. مفهوم شار نوری در علوم مهندسی در حوزه‌های مختلف مرتبط با پردازش تصویر و بینایی ماشین از جمله تشخیص حرکت، بخش‌بندی اشیاء، تثبیت ویدئو، فشرده‌سازی ویدئو و ... می‌گنجد. مفهوم شار نوری در این حوزه، به حرکت ظاهری اشیاء بین دو قاب متوالی ویدئو که ناشی از حرکت شیء یا دوربین باشد اطلاق می‌شود. این الگو یک میدان برداری^۲ دو بعدی است که هر بردار آن جابه‌جایی نقاط بین دو قاب را نشان می‌دهد. در محاسبه OF فرض‌های زیادی در نظر گرفته می‌شود، که از آن میان دو فرض زیر حائز اهمیت بیش‌تری هستند:

(۱) شدت روشنایی پیکسل‌های شیء متحرک بین دو قاب متوالی تغییر نکند.

(۲) بین دو قاب متوالی، پیکسل‌های مجاور هم به یک میزان حرکت داشته باشند.

فرض شود شدت روشنایی پیکسل P در زمان t به صورت $I(x,y,t)$ در نظر گرفته شده باشد، اگر پس از گذشت زمان dt قاب بعدی وارد شود و پیکسل P به اندازه (dx,dy) جابه‌جا شده باشد، فرض اول به رابطه ۳-۱۳ منجر می‌شود.

$$I(x,y,t) = I(x+dx,y+dy,t+dt) \quad (۳-۱۳)$$

با در نظر گرفتن تقریب سری تیلور و حذف مشتقات مراتب بالا نهایتاً معادله شار نوری طبق رابطه ۳-۱۴ به دست می‌آید:

^۱ Optical Flow (OF)

^۲ Vector field

$$I_x \cdot u + I_y \cdot v + I_t = 0 \quad I_x = \frac{\partial I}{\partial x}, I_y = \frac{\partial I}{\partial y}, u = \frac{dx}{dt}, \quad (14-3)$$

$$v = \frac{dy}{dt}, I_t = \frac{\partial I}{\partial t}$$

f_x و f_y و نیز f_t قابل محاسبه‌اند، لذا معادله‌ای با دو مجهول u و v وجود دارد که برای حل آن روش‌های گوناگونی همچون روش همبستگی فاز^۱، روش‌های مبتنی بر بلوک^۲، روش‌های بهینه‌سازی گسسته و روش‌های دیفرانسیلی^۳ مطرح شده‌اند. از جمله روش‌های دیفرانسیلی هم می‌توان به روش لوکاس-کِنید^۴، روش هورن-شانک^۴، روش باکستن-باکستن^۵، روش تغییرات عمومی^۶ و روش بلک-چپسون^۷ اشاره نمود. در این بخش به روش لوکاس-کِنید برای حل معادله شار نوری می‌پردازیم. این روش یک همسایگی 3×3 حول پیکسل نوعی P در نظر می‌گیرد، طبق دومین فرض که بین دو قاب متوالی، پیکسل‌های مجاور هم بایستی به یک میزان حرکت داشته باشند، هر ۹ نقطه همسایگی هم بین دو قاب متوالی به یک میزان جابه‌جایی خواهند داشت، لذا به ازای این ۹ نقطه می‌توان f_x ، f_y و f_t را به دست آورد و ۹ معادله با ۲ متغیر داشت. در این صورت جواب دقیقی وجود ندارد، به همین دلیل روش لوکاس-کِنید با استفاده از روش کمینه مربع خطا به یک سیستم دو معادله - دو مجهول مانند آن چه در رابطه ۳-۱۵ آمده می‌رسد و با حل آن و یافتن u و v مختصات جدید P را می‌یابد.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \sum_i I_{x_i}^2 & \sum_i I_{x_i} I_{y_i} \\ \sum_i I_{x_i} I_{y_i} & \sum_i I_{y_i}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i I_{x_i} I_{t_i} \\ -\sum_i I_{y_i} I_{t_i} \end{bmatrix} \quad i = 1, 2, \dots, 9 \quad (15-3)$$

^۱ Phase correlation

^۲ Block-based methods

^۳ Differential methods

^۴ Horn-Schunck

^۵ Buxton-Buxton

^۶ General variational method

^۷ Black-Jepson

در صورتی که نقطه نوعی P بین دو قاب متوالی دارای جابه‌جایی بزرگی باشد، این الگوریتم دچار مشکل خواهد شد، برای حل این مشکل، روش لوکاس-کِنید از هرم تصویر^۱ کمک می‌گیرد. جابه‌جایی‌های بزرگ نقاط در سطوح پایین هرم تصویر، در سطوح بالای هرم تصویر به صورت جابه‌جایی‌های کوچک در می‌آیند و لذا با بالا رفتن در هرم تصویر، این مشکل قابل حل است. با الگوریتم شار نوری هم می‌توان مانند آن چه در شکل ۷-۳ نشان داده شده مجموعه نقاط پراکنده از تصویر را بین دو قاب متوالی دنبال کرد و هم می‌توان مانند شکل ۸-۳ تمامی نقاط دارای حرکت در قاب را دنبال نمود، که روش دوم "شار نوری چگال"^۲ نامیده می‌شود. نقاط پراکنده می‌توانند گوشه‌های مهم موجود در تصویر باشند که با الگوریتم‌های متنوع گوشه‌یابی یافت می‌شوند، در این پایان‌نامه گوشه‌یابی شی-توماسی^۳ برای یافتن نقاط پراکنده یاد شده استفاده شده است.



شکل ۷-۳: نتیجه اجرای الگوریتم شار نوری به کمک مجموعه نقاط پراکنده بین دو قاب متوالی: نقاط قرمز مربوط به قاب قبلی و نقاط سبز مربوط به قاب بعد فعلی است

^۱ Image pyramid

^۲ Dense Optical Flow

^۳ Shi-Tomasi



شکل ۸-۳: نتیجه اجرای الگوریتم شار نوری چگال: تمام نقاط دارای حرکت بین دو قاب متوالی دنبال می‌شوند.

۴. روش پیشنهادی

۴-۱ مقدمه

در این پروژه ویدئوهایی از خودروهای عبوری از درگاه عوارضی تهیه شده که تشخیص نوع و رنگ خودرو طی سه گام اصلی بر روی این ویدئوها انجام می‌شود، این مراحل عبارت‌اند از: (۱) پیدا کردن ناحیه مورد علاقه (ROI)، (۲) تشخیص کلاس کلی خودرو و (۳) تعیین مدل خودرو و تعیین رنگ خودرو. در این فصل هریک از مراحل یاد شده به صورت جداگانه بیان می‌شوند. همچنین برای تشخیص جهت حرکت خودرو، ویدئوهایی از نمای بالای خودروهای عبوری تهیه شده است که الگوریتم تشخیص جهت پیشنهادی بر روی آن‌ها اعمال می‌شود.

۴-۲ پیدا کردن ROI دربرگیرنده خودروها

برای تشخیص نوع و رنگ خودرو نیاز است که ابتدا خودروی عبوری پیدا و جداسازی شود. برای این کار بایستی ابتدا هنگامی که خودرو در حال عبور از صحنه است قابی از ویدئو که دربرگیرنده خودرو است تشخیص داده شود و پس از آن ناحیه‌ای از تصویر که خودرو در آن واقع است، به عنوان ROI جداسازی شود. برای تحقق این دو عمل، چند ایده طرح شده و مورد امتحان قرار گرفته‌اند که در این بخش، هریک از آن‌ها بررسی می‌شوند و سپس روش نهایی اتخاذ شده بیان می‌گردد.

۴-۲-۱ استفاده از سایه زیر قسمت جلویی خودرو

سایه موجود در زیر خودرو در تصاویر تهیه شده از نمای جلو و پشت خودرو می‌تواند به عنوان معیاری از وجود خودرو در صحنه مورد توجه قرار گیرد. این سایه از جمله تاریک‌ترین نواحی موجود در تصویر است لذا می‌توان با آستانه‌گذاری تصویر سطح خاکستری، تاریک‌ترین نواحی را به دست آورد. این آستانه‌گذاری می‌تواند با داده‌های آموزشی به دست آید، به این ترتیب که نمونه‌هایی از نواحی سایه زیر سپر خودروها

در تصاویر سطح خاکستری پایگاه داده به صورت دستی جداسازی شود، متوسط و واریانس شدت روشنایی مربوط به هریک از نمونه‌ها محاسبه شود و دو آستانه برای واریانس و شدت روشنایی نواحی سایه زیر خودرو در نظر گرفته شود. باید در نظر داشت می‌توان برای سایه زیر خودرو با توجه به مسقف بودن عوارضی محل فیلم برداری آستانه‌هایی ثابت و بدون در نظر گرفتن شرایط نورپردازی محاسبه نمود. قصد بر آن بود پس از آستانه‌گذاری، سایه‌ی زیر سپر خودرو به صورت ناحیه‌ای مستطیل شکل جداسازی شود اما به دلیل آن که در تصاویر تهیه شده از عوارضی، سایه‌های تاریک زیادی وجود دارند که با هم هم-پوشانی قابل توجهی دارند و جداسازی یک ناحیه مستطیلی شکل به صورت مجزا ممکن نبود، این ایده کنار گذاشته شد، نمونه‌هایی از تصاویری که در آن چنین اشکالی قابل مشاهده است در شکل ۴-۱ نشان داده شده‌اند.



شکل ۴-۱: سه نمونه تصویر که در آن‌ها جداسازی سایه زیر خودرو به صورت یک ناحیه مستطیلی مجزا ممکن نیست : در ردیف پایین تصاویر سایه متناظر با تصاویر ردیف بالا نشان داده شده‌اند.

۴-۲-۲ تخمین پیش‌زمینه با استفاده از مدل مخلوط گاوسی‌ها

مدل مخلوط گاوسی‌ها یک روش برای تخمین پیش‌زمینه به صورت افقی است. در این روش برای هر پیکسل از هر قاب ویدئو با توجه به سه مقدار قرمز، آبی و سبز تشکیل دهنده پیکسل تعداد مشخصی حالت و متناظرًا تعداد مشخصی تابع احتمال گاوسی در نظر گرفته می‌شود. سپس با توجه به اطلاعات پیکسل از قاب‌های متوالی پیشین، همچون میزان تغییرات رنگ پیکسل مربوطه، هریک از پیکسل‌های

قاب جدید در یکی از دو دسته پیش‌زمینه یا پس‌زمینه قرار می‌گیرند. به این ترتیب پیکسل‌های مربوط به شیء متحرک که دارای حرکت‌اند، به عنوان پیش‌زمینه و سایر پیکسل‌ها که ساکن‌اند به عنوان پس‌زمینه در نظر گرفته می‌شوند. از اطلاعات تصویر پس‌زمینه تخمین زده شده مربوط به قاب فعلی برای به دست آوردن تصویر پیش‌زمینه در قاب بعدی استفاده می‌شود، به همین دلیل اگر شیء متحرک (به عنوان مثال خودرو) در چند قاب متوالی از ویدئو به صورت تقریباً ساکن در آید، تصویر پس‌زمینه و در نتیجه تصویر پیش‌زمینه مناسبی به دست نخواهد داد. از آنجایی که ویدئوها از درگاه عوارضی در داخل ایران تهیه شده که در آن خودروها برای مدت زمان کوتاهی به منظور پرداخت عوارض در صحنه متوقف می‌شوند، GMM قادر نخواهد بود به خوبی پیش‌زمینه را مدل کند، به همین دلیل ایده کار با GMM هم مناسب نبود. در شکل ۲-۴ این ایراد نشان داده شده است، در این شکل مورد الف لحظه عبور یک خودرو را از درگاه نشان می‌دهد، مورد ب تصویر پس‌زمینه، مورد پ ماسک دودویی پیش‌زمینه و مورد ت هم تصویر پیش‌زمینه را نشان می‌دهد، مشاهده می‌شود که پس‌زمینه به دلیل توقف نسبتاً طولانی خودروی قبلی (کامیون) به درستی مدل نشده است و تصویر پیش‌زمینه هم به درستی به دست نیامده است.

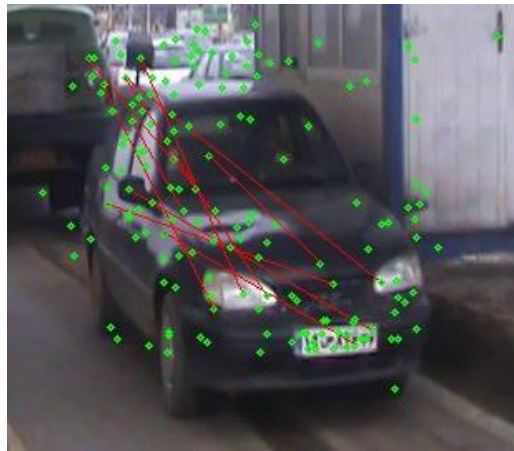


شکل ۲-۴: استفاده از روش GMM برای تخمین پیش‌زمینه و یافتن خودرو

۲-۴-۳ روش توصیفگر SURF متقارن

ناحیه جلوی خودرو نسبت به خط عمودی وسط آن متقارن است، همان گونه که در فصل قبل توضیح داده شد، توصیفگر SURF تعدادی نقطه کلیدی در تصویر پیدا می‌کند، در فصل دوم هم در [۱۰] به

روشی اشاره شد که در آن نقاط کلیدی SURF که متقارن بودند پیدا می‌شدند، با یافتن چنین نقاطی امکان آن بود تا بدون توسل به روش‌های تخمین پیش‌زمینه، ناحیه جلوی خودرو را پیدا نمود. در این روش به نحوی الگوهای شدت روشنایی حول هریک از نقاط کلیدی مورد بررسی قرار می‌گرفتند و جفت نقاط کلیدی که الگوهای شدت روشنایی حولشان به صورت آینه‌ای مشابه یکدیگر بودند به عنوان جفت نقاط تقارنی در نظر گرفته می‌شدند. اما با توجه به ماهیت تصاویر تهیه شده از عوارضی برای این پایان-نامه، که دارای نویز بودند و کیفیت متوسطی داشتند، خودروها به صورت مایل در صحنه ظاهر می‌شدند و لذا وجود تداخل و انسداد با خودروهای عقبی و خودروهای درگاه مجاور (که ناگزیر در ویدئوها حضور دارند) در آن‌ها اجتناب ناپذیر بود، تشخیص نقاط تقارنی به صورت درست و دقیق ممکن نبود. در شکل ۳-۴ نمونه‌ای از عدم امکان پیاده‌سازی درست روش مذکور نشان داده شده است و در آن مشاهده می‌شود که نقاط تقارنی به درستی تشخیص داده نشده‌اند.

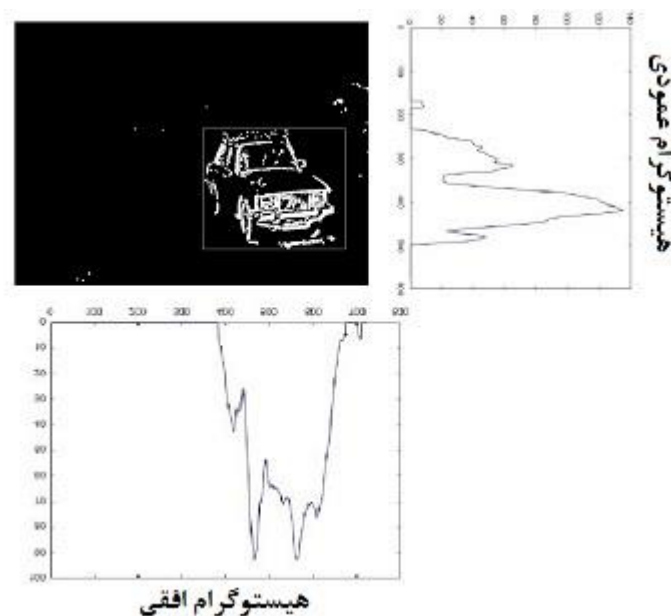


شکل ۳-۴: نتیجه پیاده‌سازی روش [۱۰] بر روی دادگان پایان‌نامه: خطوط قرمز، واصل بین نقاط SURF (سبز رنگ) تقارنی هستند که به دلیل کیفیت نامناسب تصویر به درستی تعیین نشده‌اند.

۴-۲-۴ یافتن ROI به کمک هیستوگرام‌های افقی و عمودی تصویر پیش‌زمینه

راه دیگری که برای تخمین تقریبی پیش‌زمینه وجود دارد محاسبه تفاضل دو قاب متوالی ویدئو و سپس آستانه‌گذاری آن است، تصویر پیش‌زمینه‌ی حاصل، شامل پیکسل‌هایی است که بین دو فریم دارای حرکت

هستند. می‌توان با استفاده از هیستوگرام‌های افقی و عمودی تصویر پیش‌زمینه، خودروی عبوری را یافت، روشی بر همین مبنا پیاده‌سازی شد که در این بخش به توضیح بیش‌تر در مورد آن می‌پردازیم. هیستوگرام عمودی تصویر دودویی، یک آرایه یک بعدی به طول ارتفاع قاب است که مقدار هر عضو آن برابر با تعداد پیکسل‌های سفید (پیش‌زمینه) موجود در هریک از سطرها است. هیستوگرام افقی هم به شکلی مشابه، آرایه‌ای است که مقدار هر عضو آن برابر با تعداد پیکسل‌های سفید (پیش‌زمینه) موجود در هریک از ستون‌ها است. شکل ۴-۴ هیستوگرام‌های افقی و عمودی مربوط به یک تصویر دودویی پیش-زمینه را نشان می‌دهد.



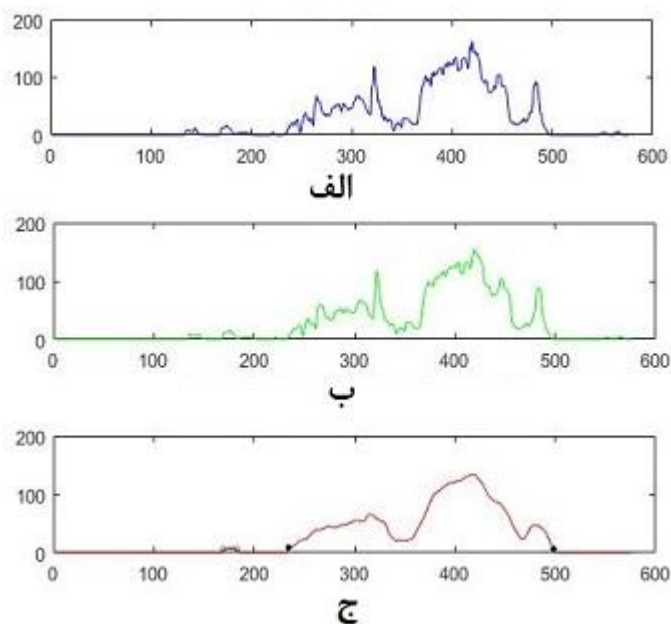
شکل ۴-۴: هیستوگرام‌های افقی و عمودی مربوط به یک تصویر دودویی پیش‌زمینه

سپس به هریک از دو هیستوگرام محاسبه شده، فیلترهای میانه و میانگین‌گیر متحرک اعمال می‌شود تا تغییرات شدید موجود در آن‌ها حذف و نمودارها هموار شوند، فیلتر میانگین‌گیر متحرک طبق رابطه ۴-۱ در هر بازه از هیستوگرام محاسبه می‌شود، در این رابطه x_{in} داده n ام از بازه m ام یک سیگنال یک بعدی

مثل هیستوگرام و $N+1$ طول بازه میانگین گیری است، همچنین فیلتر میانه متحرک هم با محاسبه داده میانی در هر بازه سیگنال هیستوگرام قابل محاسبه است.

$$\mu_i = \frac{1}{N+1} \sum_{n=i-\frac{N}{2}}^{i+\frac{N}{2}} x_{i_n} \quad (1-4)$$

آنگاه بیشینه نمودارها در نظر گرفته می شود و مقادیر کمتر از α برابر مقدار بیشینه صفر می شود. α مقداری کوچک در حد چند صدم است که در این جا برابر 0.05 در نظر گرفته شده است. این مقدار، متوسط مقادیر α ای است که با ازای آن ها بهترین نتایج جداسازی از بین 200 مورد امتحان شده به دست آمده است. همانطور که شکل ۴-۵ ج نشان می دهد، پس از این عمل، دو نمودار هیستوگرام خواهیم داشت که هریک متشکل از چندین قله جدا از هم است. از بین تمام قله های موجود در هریک از دو هیستوگرام، قله ای که بیش ترین مقدار کلی و همچنین بیش ترین مساحت را داراست به عنوان وسیع ترین قله در نظر گرفته می شود. وسیع ترین قله ها در هریک از دو هیستوگرام پیدا می شوند و محل های ابتدا و انتهای قله ها به عنوان مرزهای چپ و راست و بالا و پایین خودرو در نظر گرفته می شوند. این محل ها با نقاط مشکی رنگ در شکل ۴-۵ ج نشان داده شده اند، همچنین شکل ۴-۶ نتیجه نهایی الگوریتم را نشان می دهد.



شکل ۴-۵: الف) هیستوگرام افقی (ب) هیستوگرام افقی پس از اعمال فیلتر میانه با بازه‌های به طول ۳ (ج) هیستوگرام افقی پس از اعمال فیلتر میانگین با بازه‌های به طول ۲۱ و حذف مقادیر کمتر از $\alpha=0.05$ برابر مقدار بیشینه نمودار



شکل ۴-۶: نتیجه نهایی الگوریتم: ROI در برگرفته خودروی عبوری که به درستی به دست آمده است.

پس از پیاده‌سازی الگوریتم مواردی پیش آمد که در آن‌ها علاوه بر خودروی جلویی که باید مورد پردازش قرار گیرد، خودرویی که در پشت آن قرار دارد و با خودروی جلویی دچار انسداد است هم دارای حرکت-

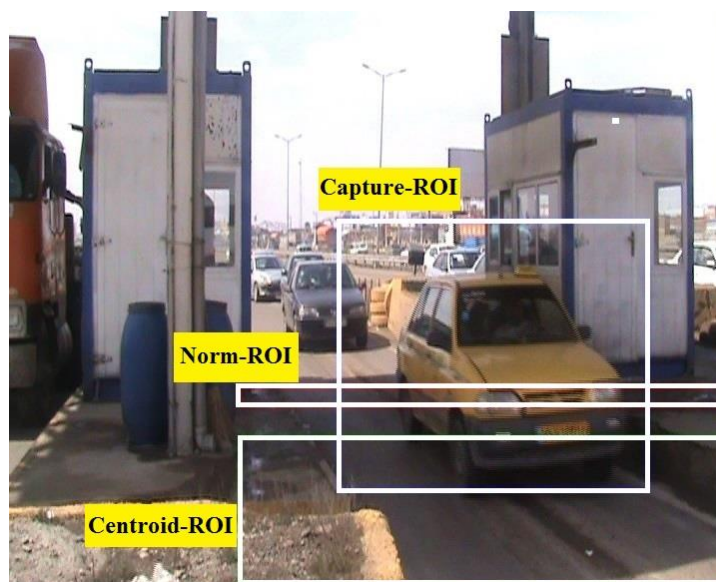
های قابل توجه بود و لذا الگوریتم قادر نبود مرزهای بالا و چپ مورد نظر و در نتیجه ROI مورد نظر را به درستی استخراج کند. نمونه‌ای از این وضعیت را می‌توان در شکل ۷-۴ مشاهده نمود.



شکل ۷-۴: انسداد قابل توجه خودروی جلویی با خودروی عقبی در حال حرکت

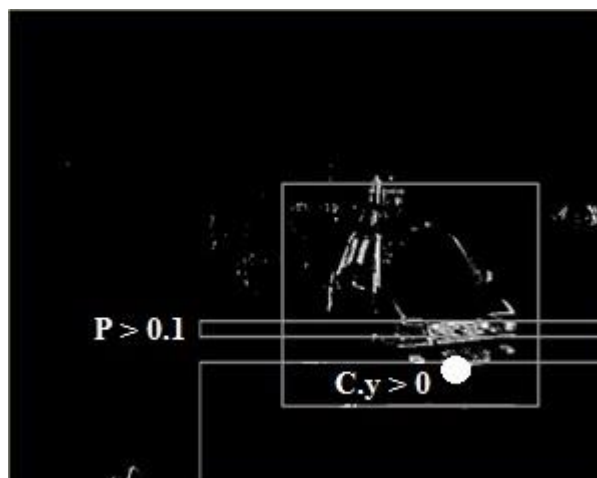
۴-۲-۵ تعیین پنجره توسط کاربر و استفاده از مرکز ثقل تصویر پیش‌زمینه داخل پنجره

با توجه به این که ویدئوها از درگاه عوارضی تهیه شده‌اند و خودروها حتماً از یک ناحیه خاص تصویر، یعنی خروجی عوارضی عبور می‌کنند، ایده بعدی که مد نظر قرار گرفت این بود که می‌توان با توجه به موقعیت مرکز ثقل در ناحیه خاصی از تصویر پیش‌زمینه، عبور خودروی جدید را تشخیص داد و ناحیه از پیش تعیین شده‌ای از تصویر را در هنگام حضور خودرو در آن به عنوان تصویر خودرو ضبط نمود. برای این کار سه ROI در نظر گرفته شد، این سه ناحیه در شکل ۸-۴ نشان داده شده‌اند.



شکل ۴-۸: سه ناحیه مستطیلی که در ابتدای الگوریتم توسط کاربر تعیین می‌شوند

الگوریتم به این صورت است که در هر قاب ورودی، درصد پیکسل‌های سفید پیش‌زمینه که داخل Norm-ROI قرار می‌گیرند محاسبه می‌شود، این مقدار را "P" می‌نامیم. مرکز ثقل پیکسل‌های پیش‌زمینه در داخل ناحیه Centroid-ROI نیز محاسبه می‌شود و ما آن را "C" و مختصات y آن را هم "C.y" می‌نامیم. براساس این اطلاعات، الگوریتم تصمیم می‌گیرد که آیا تصویر داخل Capture-ROI را ذخیره کند یا نه: اگر $P > 0.1$ و $C.y > 0$ باشد تصویر ضبط می‌شود، در شکل ۴-۹ صحنه‌ای مشاهده می‌شود که در آن شرایط یاد شده برقرار است و شکل ۴-۱۰ هم تصویر ذخیره شده را نشان می‌دهد. سیستم به این شکل عمل می‌کند که برای این که الگوریتم بتواند تصویر خودروی بعدی را ضبط کند بایستی دو شرط $C.y > T$ و $P < 0.1$ برقرار شوند، T یک آستانه است که توسط کاربر تعیین می‌شود. این روند در الگوریتم ۴-۱ آورده شده است.



شکل ۴-۹: تصویر پیش‌زمینه: شرایط برای گرفتن تصویر در این صحنه مهیا است. دایره توپر سفید رنگ موقعیت مرکز ثقل ناحیه Centroid-ROI را نشان می‌دهد.



شکل ۴-۱۰: تصویر ذخیره شده

مقدار دهی اولیه:

♦ مقدار Flag را برابر false قرار بده.

♦ T را مقداردهی کن.

برای هر قاب ورودی ویدئو:

• محاسبه کن:

- درصد پیکسل‌های سفید پیش‌زمینه داخل Norm-ROI $P =$
- مرکز ثقل پیکسل‌های پیش‌زمینه در داخل ناحیه Centroid-ROI $C =$

• اگر $(P > 0.1)$ و $(C.y > 0)$ و $(Flag = false)$

▪ ناحیه Capture-ROI از تصویر را ذخیره کن

▪ Flag را برابر true قرار بده

• اگر $(P < 0.1)$ و $(C.y > T)$ و $(Flag = true)$

▪ Flag را برابر false قرار بده

الگوریتم ۴-۱: شبه کد برای تشخیص عبور خودرو و ذخیره تصویر خودرو به کمک سه ROI تعیین شده

روش تفاضل قاب‌ها، در خصوص قسمت‌های پشت خودروهای سنگین، قادر نیست تصویر پیش‌زمینه را به خوبی مدل کند به همین خاطر، شرایط یاد شده در بالا برای گرفتن تصویر، هنگام عبور خودروهای سنگین ممکن بود چندین بار به صورت ناخواسته (هنگامی که خودرو در صحنه حضور ندارد) برآورده شوند و این نامطلوب است، به همین دلیل این روش هم ناقص بود.

۴-۲-۶ استفاده از HOG و SVM

در فصل سوم، ویژگی HOG و نیز طبقه بند SVM توضیح داده شدند. ایده‌ی بعدی که پیاده‌سازی شد این بود ابتدا یک ناحیه خاص از صحنه که خودروها در هنگام عبور در آن ظاهر می‌شدند، به طور مثال همان Capture-ROI که در بخش قبل معرفی شد، در نظر گرفته شود. سپس در هر قاب یکی از ویدئوهای تهیه شده، تصویر مربوط Capture-ROI ذخیره شد. هریک از تصاویر ذخیره شده یا شامل خودرو بودند یا فاقد آن، تصاویر ذخیره شده با همین معیار به دو دسته خودرو و غیر خودرو تقسیم شدند، در شکل ۴-۱۱ نمونه‌هایی از این تصاویر آورده شده‌اند.



شکل ۴-۱۱: نمونه‌هایی از تصاویر جدا شده از قاب‌های ویدئو: در ردیف بالا تصاویر خودروها و در ردیف پایین تصاویر غیر خودروها قرار دارند.

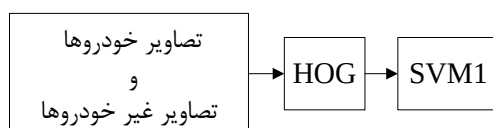
آن‌گاه تعدادی از این تصاویر خودروها و غیر خودروها انتخاب شدند و بردارهای HOG از تصاویر هریک از دو دسته استخراج شدند. سپس با استفاده از این بردارها یک طبقه‌بند SVM با کرنل خطی آموزش داده شد. در هنگام ارزیابی، بردار HOG مربوط به ناحیه Capture-ROI محاسبه شده و سپس SVM آموزش داده شده آن تصویر را به یکی از دو گروه خودرو یا غیر خودرو منتسب می‌کند که اگر خودرو باشد، یعنی ROI مربوط به خودرو پیدا شده است. طبق این روش برای دریافت نتایج بسیار دقیق نیاز بود تا تعداد بسیار زیادی داده خودرو و غیر خودرو تهیه شود، همچنین پس از اجرای الگوریتم، مشاهده شد که موارد زیادی وجود داشتند که قسمت عقب خودروهای سنگین مانند تریلرها و اتوبوس‌ها به دلیل وجود الگوهای مشابه با نمای جلوی خودروها به اشتباه خودرو تشخیص داده شدند، به همین دلیل از به کارگیری این ایده صرف نظر شد. در شکل ۴-۱۲ نمونه‌هایی از حالاتی که در آن‌ها چنین مشکلی به وجود آمده است نشان داده شده است، این تصاویر نشان می‌دهند که برای دستیابی به نتایج بهتر نیاز است تعداد تصاویر آموزشی و تنوع آن‌ها زیاد باشد.



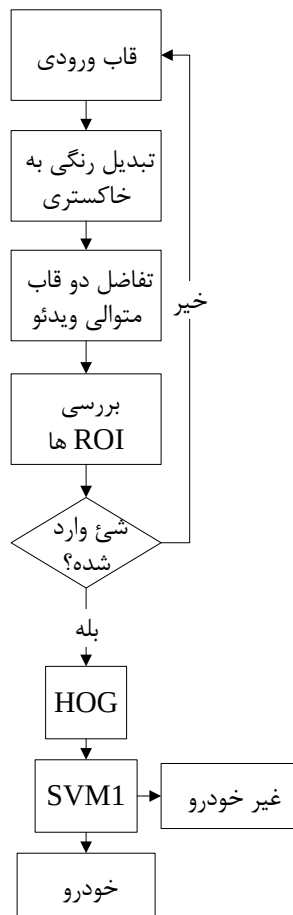
شکل ۴-۱۲: نمونه‌هایی از موارد زیاد خطا در یافتن خودرو به کمک HOG و SVM

۴-۲-۷ روش پیشنهادی برای تعیین ROI در بردارنده خودرو

ترکیب روش‌های بحث شده در دو بخش اخیر به عنوان روش نهایی پیشنهادی در نظر گرفته شد، به این ترتیب که ابتدا روش ارائه شده در بخش ۴-۲-۵ اعمال و تصویر احتمالی خودرو جدا و بردار HOG از آن استخراج می‌شود. سپس با استفاده از طبقه‌بند SVM یاد شده در بخش ۴-۲-۶ خودرو یا غیرخودرو بودن تصویر جدا شده تشخیص داده می‌شود. با این کار خطای تشخیص به میزان بسیار زیادی کاهش می‌یابد و از طرفی این روش به دلیل عدم پیچیدگی، سرعت اجرای قابل قبولی دارد، فلوچارت ۴-۱ فرآیند آموزش طبقه‌بند و فلوچارت ۴-۲ فرآیند آزمون در روش پیشنهادی برای یافتن خودرو را به نمایش می‌گذارند.



فلوچارت ۴-۱: فرآیند آموزش طبقه‌بند برای یافتن خودرو



فلوچارت ۴-۲: فرآیند آزمون در روش پیشنهادی برای یافتن خودرو

۴-۳ تشخیص کلاس کلی خودرو

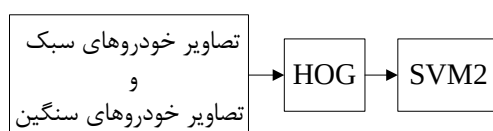
در این پایان نامه منظور از تشخیص کلاس کلی خودرو تعیین این است که خودرو پیدا شده یک خودروی سبک، مانند سواری، شاسی بلند، وانت و ... است و یا این که یک خودروی سنگین، مانند کامیون، کامیونت، اتوبوس، مینی بوس و ... است.

برای تشخیص کلاس کلی خودرو هم از بردارهای HOG به همراه طبقه‌بند SVM استفاده نمودیم. برای این کار ابتدا دو دسته داده شامل تعداد زیادی تصاویر خودروهای سبک و سنگین تهیه می‌شود که نمونه-

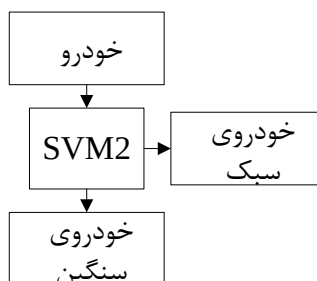
هایی از آن‌ها در شکل ۴-۱۳ نشان داده شده‌اند. از دو دسته تصویر تهیه شده قبلی بردارهای HOG استخراج و از آن‌ها برای آموزش یک طبقه‌بند SVM دیگر با کرنل خطی استفاده می‌شود. به کمک این SVM، تصاویری که به عنوان خودرو تشخیص داده شده‌اند به یکی از دو گروه خودروهای سبک و سنگین منتسب می‌شوند و به این ترتیب کار تشخیص کلاس کلی خودرو انجام می‌شود. فلوچارت ۴-۳ فرآیند آموزش طبقه‌بند و فلوچارت ۴-۴ فرآیند آزمون در روش پیشنهادی برای تعیین کلاس کلی خودرو را نشان می‌دهند.



شکل ۴-۱۳: نمونه تصاویر خودروهای سبک و سنگین: در ردیف بالا تصاویر خودروهای سبک و در ردیف پایین تصاویر خودروهای سنگین قرار دارند.



فلوچارت ۴-۳: فرآیند آموزش طبقه‌بند در روش پیشنهادی برای تعیین کلاس کلی خودرو



فلوچارت ۴-۴: فرآیند آزمون در روش پیشنهادی برای تعیین کلاس کلی خودرو

۴-۴ تعیین مدل خودرو و تشخیص رنگ

پیش از پرداختن به بحث اصلی این قسمت، باید بیان شود که در این پایان نامه به دلیل محدود بودن پایگاه داده، اولاً فقط تعیین مدل مربوط به خودروهای سبک و ثانیاً تنها چهار مدل خاص از میان خودروهای موجود در جاده‌های ایران که دارای فراوانی نسبتاً زیادی هستند مد نظر قرار گرفته است، همچنین قسمت تشخیص رنگ هم تنها برای خودروهای سبک در نظر گرفته شده است. برای آن که بتوان مدل و رنگ را تعیین کرد بایستی ناحیه جلوی خودروی سبک که حاوی اطلاعات مدل و رنگ است را به صورت دقیق‌تر از میان تصویر ذخیره شده اولیه جداسازی نمود، ابتدا به روش پیشنهادی برای جداسازی این ناحیه پرداخته شده است و سپس روش تعیین مدل و تشخیص رنگ ارائه شده است.

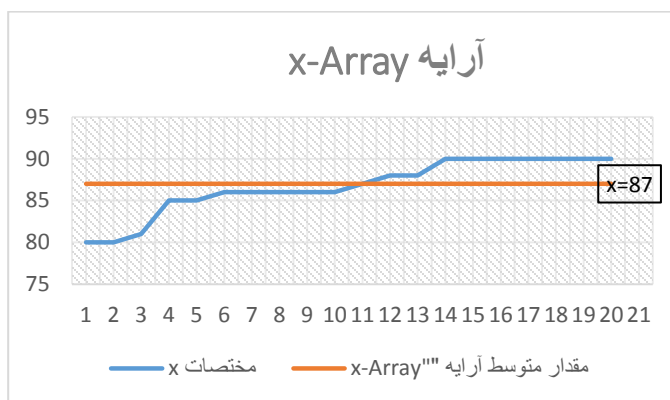
۴-۴-۱ یافتن ناحیه جلوی خودروی سبک به طور دقیق‌تر

ابتدا نیاز است تا تصویر سایه همانند آن چه در بخش ۴-۲-۱ ذکر شد به دست آید و برای حذف نویز، تصویر دودویی حاصل با استفاده از المان ساختار مربعی 3×3 مورد سایش قرار گیرد. در مرحله بعد به کمک عملگر سوبل، لبه‌های موجود در تصویر به دست می‌آیند، سپس به کمک تبدیل هاف احتمالاتی [۲۶]، قطعه خطوط افقی موجود در تصویر لبه‌یابی شده پیدا می‌شوند. پایین‌ترین قطعه خط افقی موجود انتخاب و نقطه مرکزی آن هم همان‌طور که در شکل ۴-۱۴ نشان داده شده، تعیین می‌گردد. نقطه مرکزی "Center" و مختصات y آن "Center.y" نامیده می‌شوند و خط افقی عبوری از "Center" به عنوان مرز پایین ناحیه جلوی خودرو در نظر گرفته می‌شود.

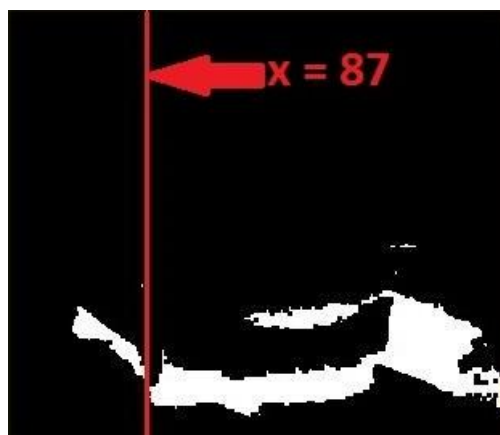


شکل ۴-۱۴: تصویر سایه حاصل از آستانه‌گذاری تصویر سطح خاکستری: پایین‌ترین قطعه خط افقی و مرکز آن، با رنگ خاکستری و مرز پایین ناحیه جلوی خودرو با رنگ سفید نشان داده شده‌اند.

پس از این کار بایستی مرزهای چپ و راست ناحیه جلوی خودرو معین شوند. برای پیدا کردن مرز سمت چپ، مجموعه‌ای از سطرهای تصویر سایه بین $y = \text{Center.y} - 20$ و $y = \text{Center.y} + 20$ در نظر گرفته می‌شوند. جمع پیکسل‌های سفید هریک از این سطرها محاسبه و "S" نامیده می‌شود. اگر S از حدی بزرگتر باشد آن‌گاه پیکسل‌های آن سطر، از چپ به راست مورد بررسی قرار می‌گیرد تا اولین پیکسل سفید پیدا شود، با پیدا شدن چنین پیکسلی، مختصات x آن ذخیره می‌شود. این عمل برای تمام سطرهای مجموعه یاد شده تکرار می‌شود و در نهایت آرایه‌ای از مختصات x اولین پیکسل‌ها ایجاد می‌شود که "x-Array" نامیده می‌شود. مقدار متوسط این آرایه به عنوان مرز سمت چپ ناحیه جلوی خودرو در نظر گرفته می‌شود. "x-Array" برای یک نمونه خودرو مربوط به شکل ۴-۱۴ در نمودار ۴-۱۵ نشان داده شده است و خط افقی قرمز رنگ مقدار متوسط این آرایه را نشان می‌دهد. همچنین مرز سمت چپ خودرو نیز با خط عمودی قرمز رنگ در شکل ۴-۱۵ نشان داده شده است.



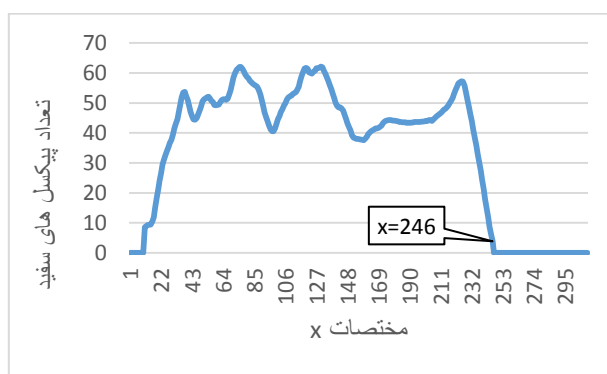
نمودار ۴-۱: آرایه "x-Array" و مقدار متوسطش که همان مرز سمت چپ خودرو است برابر ۸۷ است.



شکل ۴-۱۵: تصویر سایه: مرز سمت چپ ناحیه جلوی خودرو که در $x=87$ واقع شده با خط عمودی قرمز رنگ نمایش داده شده است.

برای تعیین مرز سمت راست از هیستوگرام عمودی تصویر پیش‌زمینه حاصل از تفاضل قاب‌های متوالی استفاده می‌شود و این عمل به دلیل ثابت بودن زاویه دوربین نسبت به جهت حرکت خودروها در پایگاه داده و این واقعیت که در پایگاه داده ما قسمت سمت راست خودروهای جلویی عبوری از درگاه عوارضی تداخلی با سایر خودروها نداشتند انجام گرفت. هیستوگرام عمودی تصویر پیش‌زمینه حاصل از تفاضل قاب‌های متوالی، به همان شکل که در بخش ۴-۲-۴ بیان شد، محاسبه می‌شود که "Hx" نامیده می‌شود. برای حذف تغییرات شدید موجود در "Hx"، فیلترهای میانگین‌گیر متحرک و میانگین‌گیر یک بعدی به صورت متوالی به آن اعمال می‌شود. بیشینه مقدار "Hx"، "MAX" نامیده شده و مقادیر کمتر از

0.05MAX در "Hx" صفر می‌شوند. پس از انجام این عملیات، "Hx" متشکل از مجموعه‌ای از چندین قله جدا از هم خواهد بود. همان طور که در نمودار ۲-۴ نشان داده شده است از بین تمامی قله‌ها، قله‌ی دارای بیش‌ترین مساحت و نیز بزرگ‌ترین بیشینه انتخاب و موقعیت مرز سمت راست قله مذکور به عنوان مرز سمت راست ناحیه جلوی خودرو در نظر گرفته می‌شود. در شکل ۴-۱۶، این مرز با یک خط عمودی قرمز رنگ نشان داده شده است.

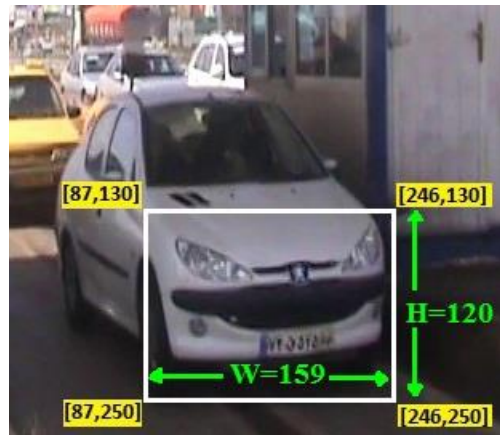


نمودار ۲-۴: Hx: مرز سمت راست وسیع‌ترین قله موجود در هیستوگرام عمودی که در $x=246$ واقع است.



شکل ۴-۱۶: تصویر پیش‌زمینه: مرز سمت راست ناحیه جلوی خودرو که متناظر با مرز سمت راست وسیع‌ترین قله موجود در هیستوگرام عمودی است با یک خط عمودی قرمز رنگ نمایش داده شده است.

فاصله بین مرزهای راست و چپ را "W" نامیده و خط $y = \text{Center.y} - 0.75W$ به عنوان مرز بالایی جلوی خودرو در نظر گرفته می‌شود تا در نهایت ROI جلوی خودرو همان‌طور که در شکل ۴-۱۷ نشان داده شده است، به دست آید، فرآیند یافتن ROI جلوی خودرو در الگوریتم ۴-۲ آورده شده است.



شکل ۴-۱۷: تصویر گرفته شده از یک خودروی سبک: مستطیل سفید رنگ ناحیه جلوی خودرو را که الگوریتم پیشنهادی به دست داده است نشان می‌دهد.

- مقدار دهی اولیه : T و آرایه x-Array
- استخراج سایه از تصویر.
- لبه‌یابی تصویر سایه به کمک لبه‌یاب سوبل.
- یافتن قطعه خطوط افقی موجود به کمک تبدیل هاف احتمالاتی.
- یافتن پایین‌ترین قطعه خط افقی موجود و نقطه مرکزی آن.
 - نقطه مرکزی : Center
 - مختصات y نقطه مرکزی : Center.y
 - مرز پایین ROI : $y = \text{Center.y}$
- برای $i = \text{Center.y} - 20$ تا $i = \text{Center.y} + 20$ عملیات زیر انجام شود:
 - محاسبه شود: جمع پیکسل‌های سطر iام $S =$
 - اگر $S > T$:
- پیکسل‌های سطر iام تصویر سایه از چپ به راست بررسی شوند و مختصات x اولین پیکسل سفید پیدا شده در آرایه x-Array ذخیره شود.
- مقدار متوسط آرایه x-Array به عنوان مرز سمت چپ ROI انتخاب شود.
- محاسبه هیستوگرام افقی تصویر پیش‌زمینه حاصل از تفاضل قاب‌های متوالی ویدئو.

- اعمال فیلترهای میانه و میانگین به هیستوگرام افقی.
- در نظر گرفته شود: مقدار بیشینه هیستوگرام افقی $= MAX$.
- حذف مقادیر کوچکتر از $0.05MAX$ در هیستوگرام افقی.
- یافتن وسیع‌ترین قله موجود در هیستوگرام افقی.
- مرز سمت راست وسیع‌ترین قله به عنوان مرز سمت چپ ROI انتخاب شود.
- در نظر گرفته شود: فاصله مرزهای چپ و راست $= W$.
- خط $y=Center.y-0.75W$ به عنوان مرز بالای ROI انتخاب شود.

الگوریتم ۲-۴: الگوریتم پیشنهادی برای یافتن ROI جلوی خودروی سبک

۲-۴-۴ تعیین مدل خودروی سبک

همان گونه که در فلوچارت ۴-۵ نشان داده شده است، از تصویر ROI به دست آمده ویژگی‌های HOG استخراج و به عنوان ورودی به SVMهای شماره ۳ تا ۶ داده می‌شود تا مدل خودرو تعیین شود. هریک از این SVMها یک طبقه‌بند دو کلاسه یک در برابر بقیه^۱ هستند که با استفاده از تصاویر آموزشی‌ای که به صورت دستی تهیه شده‌اند آموزش داده شده‌اند، ابعاد این تصاویر به 126×168 تغییر داده می‌شوند.

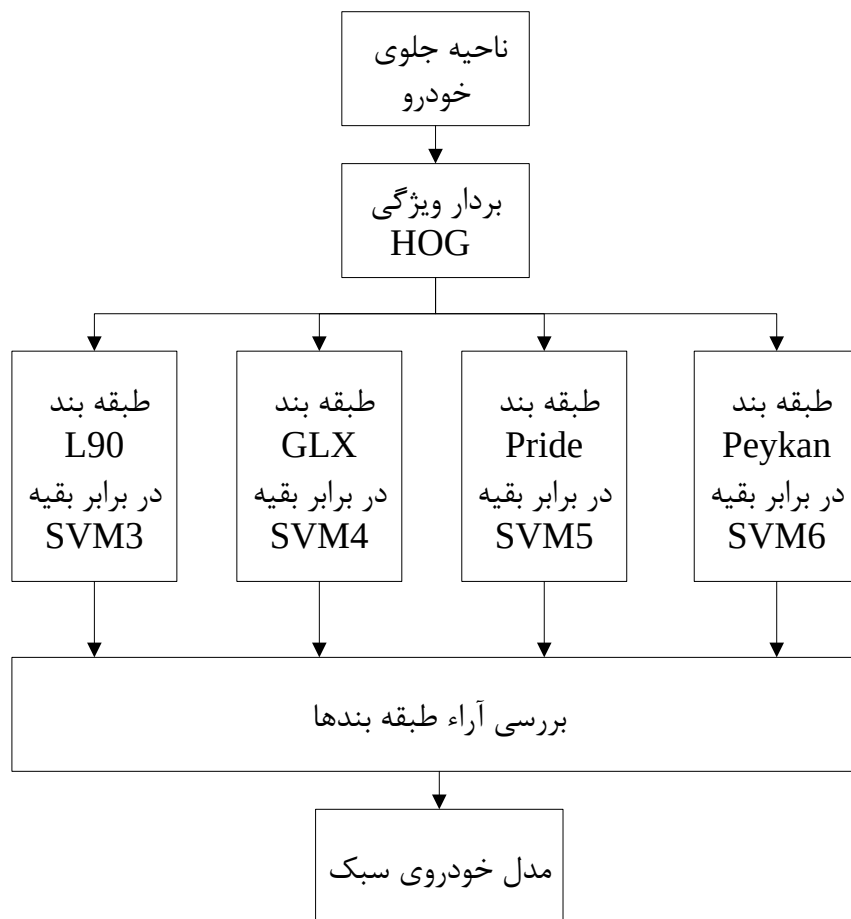
به عنوان مثال، از هریک از چهار مدل خودروی سبک مورد نظر، سه نمونه در شکل ۴-۱۸ آورده شده است، در سطرهای اول تا چهارم این شکل به ترتیب نمونه‌های L90، GLX، Pride و Peykan مشاهده می‌شوند. این چهار مدل به این دلیل انتخاب شده‌اند که نسبتاً در بین خودروهای موجود در جاده‌های ایران بیش‌ترین تعداد را دارند و نسبت به سایر مدل‌ها، بیش‌تر متمایز کننده هستند. طبقه‌بندی به این صورت انجام می‌گیرد که ابتدا ابعاد ROI به 126×168 تغییر داده می‌شود، سپس بردار ویژگی از آن استخراج شده و به هریک از طبقه‌بندها داده می‌شود. هر طبقه‌بند به داده ورودی‌اش یا برچسب مدل در نظر گرفته شده‌اش و یا برچسب غیره را اختصاص می‌دهد. در این میان، مدل خودرو را طبقه‌بندی تعیین می‌کند که برچسب مدل خودش را به داده اختصاص داده باشد و در عین حال مابقی طبقه‌بندها هم به

^۱ One vs Rest (OvR)

داده ورودی برچسب غیره را تخصیص داده باشند. البته ممکن است حالتی رخ دهد که بیش از یک طبقه - بند برچسب مدل خودش را به داده ورودی بدهد و برای جلوگیری از این مورد خاص میزان احتمال تعلق خودرو به مدل اختصاصی هر طبقه بند طبق آنچه در [۲۵] آمده محاسبه می شود و طبقه بندی که بیش - ترین احتمال را بدهد مدل خودرو را تعیین می کند. در نهایت اگر مقادیر احتمال ها هم همگی بیشینه و یکسان باشند داده ورودی رد می شود.



شکل ۴-۱۸: نمونه هایی از چهار مدل خودروی سبک در نظر گرفته شده در این پایان نامه.



فلوچارت ۴-۵: روند تعیین مدل خودروی سبک

۴-۴-۳ تشخیص رنگ خودروی سبک

برای تشخیص رنگ خودرو ابتدا بایستی قسمتی از خودرو که دارای رنگ است جداسازی شود، به همین جهت قسمت مربوط به کاپوت جدا می‌شود. برای این کار به کمک لبه‌یاب Canny لبه‌های موجود در تصویر ناحیه یافت شده جلوی خودرو استخراج می‌شوند، پنجره‌ای به طولی برابر عرض ناحیه جلوی خودرو و عرض مشخص در نظر گرفته می‌شود و آن‌گاه این پنجره از بالا تا نیمه تصویر لبه لغزنده می‌شود و در هر مکان، واریانس لبه محاسبه می‌شود و در نهایت محلی که تصویر لبه در آن کمترین واریانس را

داشته باشد به عنوان کاپوت خودرو انتخاب می‌شود، در شکل ۴-۱۹ این ناحیه با رنگ قرمز بر روی ناحیه جلوی یک خودرو نشان داده شده است.



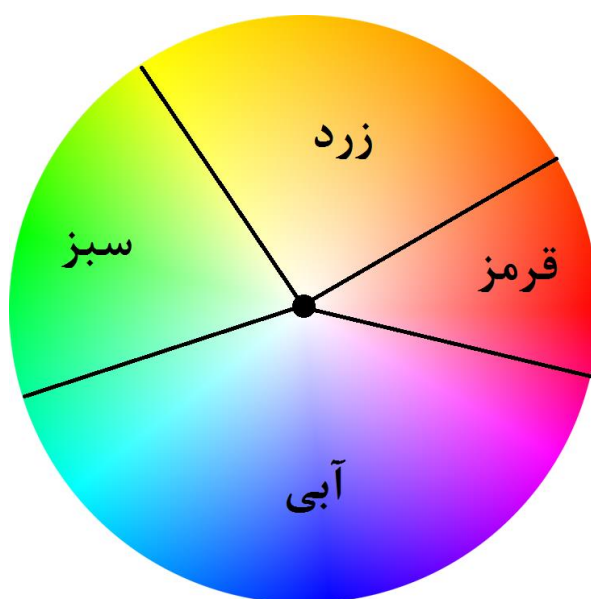
شکل ۴-۱۹: الف) تصویر ناحیه جلوی خودرو (ب) لبه‌یابی شده تصویر الف : ناحیه مورد نظر برای انجام عملیات تشخیص رنگ با یک مستطیل قرمز رنگ و خط نیمه تصویر با خطی سبز رنگ نشان داده شده‌اند.

پس از جداسازی قسمت رنگی کاپوت، برای تعیین رنگ، ابتدا مقادیر hue پیکسل‌های این ناحیه با استفاده از خوشه‌بند kmeans به $k=3$ دسته خوشه‌بندی می‌شوند، سپس پیکسل‌های مربوط به خوشه دارای بیش‌ترین عضو برای تعیین رنگ غالب در نظر گرفته می‌شوند. دسته بندی به این تعداد دسته به این علت صورت گرفته که بهترین نتایج جداسازی پیکسل‌های رنگ غالب به ازای این تعداد دسته حاصل شده است. در این پایان‌نامه تعیین رنگ خودرو به این صورت است که ابتدا به کمک معیار قدرت رنگ^۱ که در [۲۱] به کار گرفته شده خودرو در یکی از دو دسته رنگی یا غیر رنگی جای می‌گیرد. معیار قدرت رنگ یاد شده در رابطه ۴-۲ آورده شده است، در این رابطه R_{avg} ، G_{avg} و B_{avg} متوسط به ترتیب مولفه‌های قرمز، سبز و آبی پیکسل‌های رنگ غالب هستند.

$$CS = \left(\left(\frac{R_{avg}}{255} - \frac{1}{3} \right)^2 + \left(\frac{G_{avg}}{255} - \frac{1}{3} \right)^2 + \left(\frac{B_{avg}}{255} - \frac{1}{3} \right)^2 \right) + (\max\{R_{avg}, G_{avg}, B_{avg}\} - \min\{R_{avg}, G_{avg}, B_{avg}\}) \quad (2-4)$$

^۱ Chromatic Strength (CS)

اگر مقدار CS از یک مقدار آستانه خاص کم‌تر باشد، خودرو در دسته غیررنگی‌ها و اگر بیش‌تر از آن باشد، در دسته رنگی‌ها قرار می‌گیرد، این آستانه به کمک داده‌های آموزشی تعیین می‌شود. خودروهای رنگی با توجه به متوسط مقدار hue پیکسل‌های رنگ غالب به سه دسته قرمز، زرد، سبز و آبی تقسیم بندی می‌شوند و خودروهای غیررنگی هم با توجه به متوسط شدت روشنایی به دو دسته روشن و مشکی دسته بندی می‌شوند، لازم به ذکر است به دلیل کیفیت پایین تصاویر پایگاه داده این تعداد دسته رنگ در نظر گرفته شده است. گستره‌های hue مربوط به چهار گروه رنگ یاد شده در دایره رنگ در شکل ۴-۲۰ نشان داده شده‌اند.



شکل ۴-۲۰: گستره‌های hue مربوط به چهار گروه رنگ

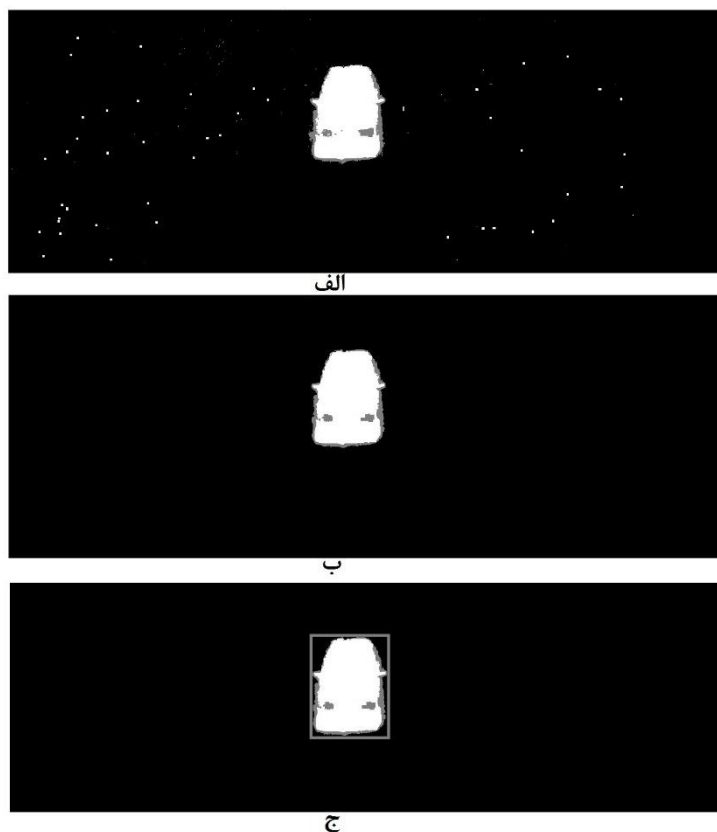
۴-۵ تشخیص جهت حرکت خودرو

همان طور که در مقدمه در بخش ۴-۱ بیان شد برای تشخیص جهت، ویدئویی از بالا و نمای روبه‌روی و پشت خودروهای عبوری تهیه شده است که در شکل ۴-۲۱ نمونه قابی از این ویدئو نشان داده شده است، در این ویدئو ترافیک و در نتیجه انسداد بین خودروها وجود ندارد. برای تشخیص جهت با استفاده از این

نوع ویدئو ابتدا خودروها و ماسک دودویی آن‌ها به کمک روش GMM از پس زمینه جدا می‌شوند. سپس تصویر دودویی مورد سایش قرار می‌گیرد تا نویز جزئی از بین برود و آن گاه کانتورهای بزرگ موجود در تصویر دودویی استخراج می‌شوند. در این حالت در هر قاب ویدئو تعدادی کانتور جدا از هم وجود دارند که مستطیل محیطی هر کدام قابل به دست آمدن است و آن را Car_ROI نامیده‌ایم، در شکل ۴-۲۲ تصویر دودویی پیش‌زمینه حاصل از تخمین با GMM نشان داده شده است.



شکل ۴-۲۱: نمونه قابی ویدئوی تهیه شده برای تشخیص جهت



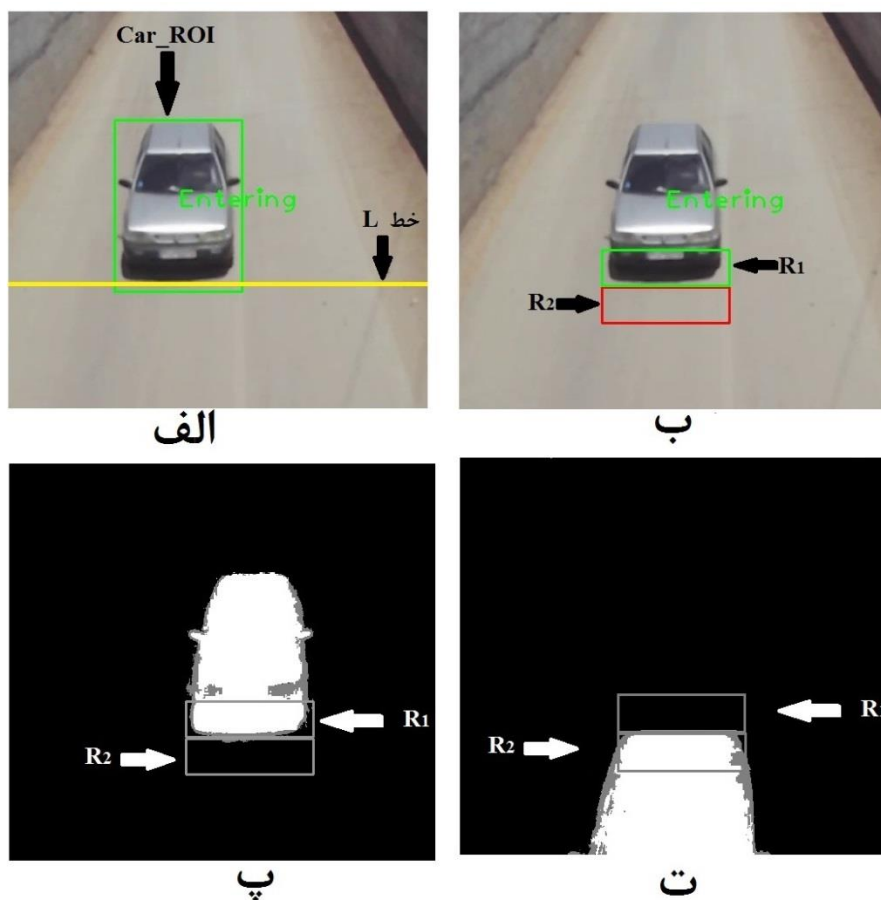
شکل ۴-۲۲: الف) تصویر دودویی پیش‌زمینه حاصل از تخمین با GMM (ب) تصویر دودویی پس از سایش (ج) مستطیل محیطی حول خودرو کشیده شده است.

پس از یافتن خودروهای عبوری به کمک کانتورها در هر قاب از ویدئو، نقاط کلیدی پیدا شده به کمک گوشه‌یاب Shi-Tomasi که از قاب پیشین دنبال شده‌اند و داخل Car_ROI خودرو قرار دارند، بررسی می‌شوند. تعداد نقاط کلیدی که مؤلفه y آن‌ها در موقعیت فعلی بزرگ‌تر از مؤلفه y آن‌ها در موقعیت قبلی است شمرده می‌شوند، این تعداد N_{enter} نامیده می‌شود و تعداد مابقی نقاط کلیدی داخل Car_ROI خودرو، هم N_{exit} نامیده می‌شود. اگر $N_{exit} < N_{enter}$ باشد خودرو در حال حرکت به سمت پایین و ورود است، در غیر این صورت خودرو در حال حرکت به سمت بالا و خروج است، به این ترتیب جهت خودروها تشخیص داده می‌شود. برای شمردن خودروهای ورودی و خروجی هم به این شکل عمل می‌شود

که: ابتدا صحنه به سه قسمت مساوی چپ، وسط و راست تقسیم می‌شود و شش پرچم^۱ برای ورود یا خروج از هریک از سه ناحیه صحنه در نظر گرفته می‌شود. حال در هر قاب بررسی می‌شود اگر خودرو در حال ورود و حرکت به سمت پایین باشد و مانند شکل ۴-۲۳ الف، مرز پایین Car_ROI از خط از پیش تعیین شده L به سمت پایین عبور کرده باشد، دو ناحیه مستطیلی به طولی برابر عرض Car_ROI و عرضی برابر با 0.06 ام ارتفاع تصویر در بالا و پایین خط L، به ترتیب به نام‌های R1 و R2 در نظر گرفته می‌شود، این دو ناحیه در شکل ۴-۲۳ ب و پ نشان داده شده‌اند. اگر درصد پیکسل‌های سفید موجود در R1 بیش‌تر از ۰/۷ و هم زمان درصد پیکسل‌های سفید موجود در R2 کم‌تر از ۰/۴ باشد، خودروی مربوطه به عنوان خودروی ورودی شمارش می‌شود و با توجه به موقعیت مرکز Car_ROI و قرارگیری آن در هریک از سه ناحیه یاد شده پرچم مربوطه تغییر وضعیت می‌دهد. پس از آن، هنگامی که درصد پیکسل‌های سفید موجود در R1 کم‌تر از ۰/۴ و هم زمان درصد پیکسل‌های سفید موجود در R2 بیش‌تر از ۰/۷ شود پرچم مربوط به آن ناحیه دوباره به وضعیت اولیه خود برمی‌گردد و خودروی بعدی که از آن ناحیه وارد شود، می‌تواند شمارش شود، این حالت در شکل ۴-۲۳ ت نشان داده شده است. اگر خودرو در حال خروج و حرکت به سمت بالا باشد و مرز بالای Car_ROI از خط از پیش تعیین شده L به سمت بالا عبور کرده باشد باز هم دو ناحیه مستطیلی به طولی برابر با عرض Car_ROI و عرضی برابر با 0.06 ام ارتفاع تصویر در بالا و پایین خط L، به ترتیب به نام‌های R1 و R2 در نظر گرفته می‌شود. اگر درصد پیکسل‌های سفید موجود در R2 بیش‌تر از ۰/۷ و هم زمان درصد پیکسل‌های سفید موجود در R1 کم‌تر از ۰/۴ باشد خودروی عبوری به عنوان خودروی خروجی شمارش می‌شود و با توجه به موقعیت مرکز Car_ROI و قرارگیری آن در هریک از سه ناحیه یاد شده پرچم مربوطه‌اش تغییر وضعیت می‌دهد. پس از آن هنگامی که درصد پیکسل‌های سفید موجود در R2 کم‌تر از ۰/۴ و هم زمان درصد پیکسل‌های سفید موجود در

^۱ Flag

R1 بیش‌تر از ۰/۷ شود پرچم مربوط به آن ناحیه دوباره به وضعیت اولیه خود برمی‌گردد و خودروی خروجی بعدی از آن ناحیه می‌تواند شمارش شود. روند تشخیص جهت و شمارش خودروها برای هریک از سه ناحیه و مربوط به پس از یافتن کانتورها در الگوریتم ۳-۴ بیان شده است.



شکل ۴-۲۳: خودرو در حال ورود است (الف) مرز پایین Car_ROI از خط L گذشته است (ب) نواحی مستطیلی R1 و R2 در بالا و پایین خط L در قاب ویدئو در هنگام رسیدن خودرو به خط L (پ) نواحی مستطیلی R1 و R2 در بالا و پایین خط L در تصویر دودویی در هنگام رسیدن خودرو به خط L (ت) نواحی مستطیلی R1 و R2 در بالا و پایین خط L در تصویر دودویی در هنگام عبور کامل خودرو

مقداردهی اولیه:

✓ ابتدا تعداد خودروهای ورودی را $N_{enter}=0$ و تعداد خودروهای خروجی را $N_{exit}=0$ قرار بده.

✓ قرار بده $Flag_{enter}=false$ و $Flag_{exit}=false$.

✓ خط افقی L و محل آن $L.y$ را در نظر بگیر.

• در هر قاب ویدئو:

▪ کانتورهای خودروها به دست آیند.

▪ نقاط کلیدی مربوط به قاب قبل و دنبال شده آن‌ها در قاب فعلی به دست آیند.
▪ برای هر کانتور:

○ مستطیل محیطی هر کانتور به دست آید و Car_ROI نامیده شود. مرز بالای این ناحیه y_{up} و مرز آن y_{down} نامیده شود.

○ تعداد نقاط کلیدی داخل Car_ROI که مؤلفه y آن‌ها در موقعیت فعلی، بزرگ‌تر از مؤلفه y آن‌ها در موقعیت قبلی است، N_{down} نامیده شود.

○ تعداد نقاط کلیدی داخل Car_ROI که مؤلفه y آن‌ها در موقعیت فعلی کوچک‌تر از مؤلفه y آن‌ها در موقعیت قبلی است، N_{up} نامیده شود.

❖ اگر $N_{up} < N_{down}$ و $y_{down} > L.y$

○ دو ناحیه $R1$ و $R2$ به طولی برابر طول Car_ROI و عرضی برابر با 0.06 ام ارتفاع تصویر در بالا و پایین خط L ، به ترتیب به نام‌های $R1$ و $R2$ در نظر گرفته شود.

○ درصد پیکسل‌های سفید داخل $R1$ و $R2$ به ترتیب $P1$ و $P2$ نامیده شوند.

♦ اگر $Flag_{enter}=false$ بود و $P1 > 0.7$ و $P2 < 0.4$

➤ N_{enter} را یکی زیاد کن.

➤ $Flag_{enter}=true$ قرار بده.

♦ اگر $Flag_{enter}=true$ بود و $P1 < 0.4$ و $P2 > 0.7$

➤ Flag_enter=false قرار بده.

❖ اگر $N_down < N_up$ و $y_up < L.y$

○ دو ناحیه R1 و R2 به طولی برابر طول Car_ROI و عرضی برابر با 0.06 ام ارتفاع تصویر در بالا و پایین خط L، به ترتیب به نام‌های R1 و R2 در نظر گرفته شود.

○ درصد پیکسل‌های سفید داخل R1 و R2 به ترتیب P1 و P2 نامیده شوند.

◆ اگر Flag_exit=false بود و $P1 < 0.4$ و $P2 > 0.7$

➤ N_exit را یکی زیاد کن.

➤ Flag_exit=true قرار بده.

◆ اگر Flag_exit=true بود و $P1 > 0.7$ و $P2 < 0.4$

➤ Flag_exit=false قرار بده.

الگوریتم ۳-۴: روند پیشنهادی برای تشخیص جهت و شمارش خودروها در هریک از سه ناحیه تصویر

۵. نتایج تجربی

در این فصل پایگاه داده مورد استفاده در این پایان نامه معرفی و نحوه تهیه آن بیان می شود، نتایج مربوط به بخش های مختلف الگوریتم پیشنهادی بیان و همچنین نتایج روش پیشنهادی با نتایج برخی روش های دیگر مقایسه می شود.

۵-۱ معرفی پایگاه داده

۵-۱-۱ فرآیند انتخاب پایگاه داده و دسترسی به آن

ویدئوهای مورد نیاز با دوربین شخصی از درگاه یک عوارضی تهیه شدند، برای این کار به عوارضی باغچه در چهل کیلومتری مشهد مراجعه شد. تصویربرداری در روز و از ساعت ۷ صبح الی ۱۲:۱۵ ظهر به صورت منفصل هر یک ساعت و در بازه های زمانی ۳۰ یا ۴۵ دقیقه ای انجام شد. در نهایت در حدود دو ساعت و پانزده دقیقه فیلم حاصل شد، در جدول ۵-۱ مدت زمان ویدئوها و نیز شرایط آب و هوایی آورده شده است. لازم به ذکر است عوارضی یاد شده مسقف بوده و خودروها با وجود آب و هوای مختلف همواره در سایه این سقف قرار می گرفتند. در شکل ۵-۱ الف تا چ قاب هایی از هریک از این ویدئوها به عنوان نمونه آورده شده است.

جدول ۵-۱: مدت زمان ویدئوها و نیز شرایط آب و هوا در ویدئوهای تهیه شده از عوارضی باغچه

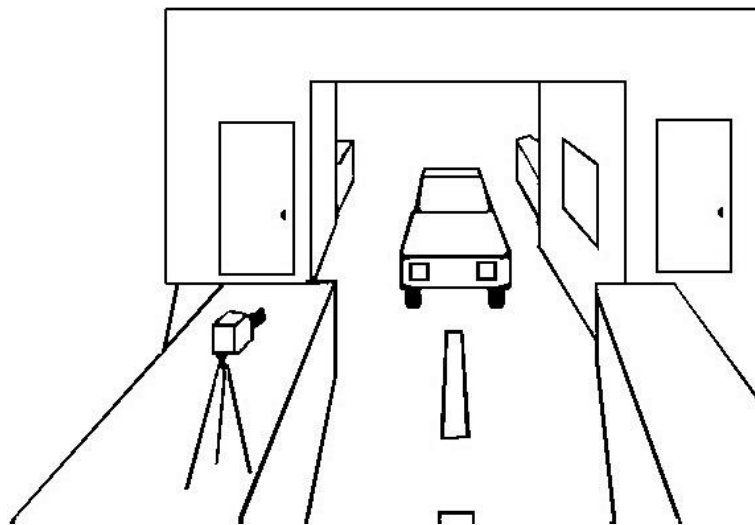
شماره ویدئو	دقیقه	شرایط آب و هوا	ساعت آغاز
ویدئوی اول (V1)	دقیقه ۱-۳۰	ابری (شکل ۵-۱ الف)	۷:۰۰
ویدئوی دوم (V2)	دقیقه ۱-۳۰	ابری (شکل ۵-۱ ب)	۸:۳۰
ویدئوی سوم (V3)	دقیقه ۱-۴	ابری (شکل ۵-۱ پ)	۹:۰۰
	دقیقه ۴-۱۰	نیمه آفتابی (شکل ۵-۱ ت)	
	دقیقه ۱۰-۱۵	آفتابی (شکل ۵-۱ ث)	
ویدئوی چهارم (V4)	دقیقه ۱-۳۰	آفتابی (شکل ۵-۱ ج)	۱۰:۱۵
ویدئوی پنجم (V5)	دقیقه ۱-۳۰	آفتابی (شکل ۵-۱ چ)	۱۱:۴۵



شکل ۵-۱: نمونه قاب‌هایی از هریک از ویدئوهای تهیه شده از درگاه عوارضی باغچه

۵-۱-۲ مشخصات ویدئوهای تهیه شده

برای ضبط ویدئوها از یک دوربین SONY DCR SR46 با کیفیت متوسط استفاده شد و دوربین همان گونه که در شکل ۵-۲ نشان داده شده، بر روی یک سه پایه و بر روی سکویی واقع در کنار معبر خودروها در فاصله ۱۰ الی ۱۵ متری دروازه عوارضی قرار داده شده بود.



شکل ۲-۵: نحوه قرارگیری دوربین نسبت به درگاه عوارضی در پایگاه داده تهیه شده

ویدئوهای حاصل رنگی و به فرمت MPG و با نرخ ۲۵ قاب بر ثانیه هستند و ابعاد هر قاب 720×576 است. قابل ذکر است که در ویدئوهای تهیه شده اثر ناخواسته درهم‌بافی که در بخش ۳-۲ به آن اشاره شد وجود داشت و به همین دلیل پیش از انجام عملیات پردازشی، به کمک نرم‌افزار iDealshare VideoGo 6 مورد عمل نادرهم‌بافی قرار گرفتند. ابعاد سه ROI در نظر گرفته شده در بخش ۴-۲-۵ و نشان داده شده در شکل ۴-۸ یعنی Capture-ROI، Norm-ROI و Centroid-ROI به ترتیب 310×270 ، 490×20 و 490×150 هستند.

۲-۵ نتایج عددی

در این قسمت، دقت بازشناسی مربوط به یافتن، تشخیص خودرو و تعیین مدل خودرو بیان و مقایسه‌هایی با برخی روش‌های دیگر انجام می‌شود. پیش از این بیان شد که پنج ویدئو به عنوان پایگاه داده در اختیار داریم که آن‌ها را "V1"، "V2"، "V3"، "V4" و "V5" نامیده‌ایم. برای ارزیابی الگوریتم‌های ارائه شده یک رویه گام به گام در پیش گرفته شد، به این ترتیب که طبقه‌بند SVM1 با استفاده از تصاویر آموزشی جمع‌آوری شده از ویدئوی V1 آموزش داده و سپس از ویدئوی V2 برای ارزیابی این طبقه‌بند استفاده شد.

سپس داده‌های مثبت-صحیح^۱ و منفی-کاذب^۲ یا همان خودروها به داده‌های مثبت مجموعه‌ی آموزشی اولیه و داده‌های منفی-صحیح^۳ و مثبت-کاذب^۴ یا همان غیر خودروها به داده‌های منفی مجموعه‌ی آموزشی اولیه اضافه گردید. این روند تا جایی ادامه یافت که از ویدئوی V4 هم برای ارزیابی طبقه‌بند استفاده شد و مثل قبل، داده‌های مناسب به مجموعه آموزشی کلی اضافه شدند. پس از این مرحله داده‌های مثبت موجود در مجموعه آموزشی نیز به دو دسته‌ی خودروهای سبک و سنگین تفکیک و از آن‌ها برای آموزش طبقه‌بند SVM2 استفاده شد. در مرحله آخر هم، از ویدئوی V5 برای ارزیابی هر دو طبقه‌بند SVM1 و SVM2 و نیز تعیین مدل توسط طبقه‌بندهای SVM3، SVM4، SVM5 و SVM6 استفاده شد.

۵-۲-۱ نتایج مربوط به یافتن و تعیین کلاس کلی خودرو

داده‌های آموزشی و آزمون در این مرحله، دارای ابعاد ۳۱۰×۲۷۰ هستند. برای ارزیابی و گزارش نتایج الگوریتم‌های این دو قسمت، از سه معیار حساسیت^۵ و دقت^۶ و صحت^۷ استفاده شده است، این سه معیار به ترتیب در روابط ۱-۵، ۲-۵ و ۳-۵ آورده شده‌اند.

$$TPR = TP / (TP + FN) \quad (۱-۵)$$

$$PPV = TP / (TP + FP) \quad (۲-۵)$$

$$Accuracy = \frac{T}{T+F} \quad (۳-۵)$$

T = تعداد تشخیص درست
F = تعداد تشخیص غلط

^۱ True-Positive (TP)

^۲ False-Negative (FN)

^۳ True-Negative (TN)

^۴ False-Positive (FP)

^۵ True-Positive Rate (TPR)

^۶ Positive Predictive Value (PPV)

^۷ Accuracy

نتایج ارزیابی گام به گام یاد شده در جدول ۵-۲ آورده شده است، همانطور که پیش از این هم بیان شد، تعیین کلاس کلی خودرو تنها در آخرین گام یعنی در هنگام استفاده از ویدئوی پنجم برای ارزیابی، انجام شده است و برای بررسی عملکرد طبقه‌بند SVM2 تنها از معیار صحت که نسبت تشخیص درست به کل تعداد خودروهاست استفاده شده است. همچنین بایستی در نظر داشت تعداد داده‌های TN برابر تعداد تمام قاب‌هایی از ویدئو است که خودرو در آن‌ها حضور نداشته و مقداری بزرگ است، لذا در جدول ۵-۲ درج نشده است، به همین خاطر به صورت تصادفی و به تعداد دلخواه از داده‌های TN حاصل از هر مرحله به داده‌های منفی مرحله بعد اضافه شده‌اند. در جدول ۵-۳ هم به طور جزئی‌تر در هر گام، تعداد داده‌های TP، FN و FP و نیز معیارهای حساسیت و دقت مربوط به مرحله تشخیص وجود خودرو (SVM1) برای هریک از دو دسته خودروهای سبک و سنگین آورده شده است.

جدول ۵-۲: نتایج آزمون تشخیص وجود خودرو و تعیین کلاس کلی خودرو به شیوه گام به گام

گام	ویدئوی آموزشی	ویدئوی ارزیابی	تعداد خودرو سبک	تعداد خودرو سنگین	تعداد خودرو در کل	TP	FP	FN	TPR SVM1	PPV SVM1	T تشخیص درست	F تشخیص نادرست	Accuracy SVM2
۱	V1	V2	۹۰	۵۱	۱۴۱	۱۳۴	۱	۶	%۹۵/۷	%۹۹/۳	-	-	-
۲	V1, V2	V3	۴۰	۲۹	۶۹	۶۵	۲	۲	%۹۷	%۹۷	-	-	-
۳	V1, V2, V3	V4	۷۶	۴۱	۱۱۷	۱۰۴	۰	۱۳	%۸۹	%۱۰۰	-	-	-
۴	V1, V2, V3, V4	V5	۱۸۳	۴۰	۲۲۳	۲۱۵	۲	۶	%۹۷/۳	%۹۹	۲۱۵	۴	%۹۸

جدول ۵-۳: نتایج آزمون تشخیص وجود خودرو و تعیین کلاس کلی خودرو به شیوه گام به گام برای هریک از دو دسته خودروهای سبک و سنگین

گام	تعداد خودرو سبک	تعداد خودرو سنگین	TP سبک	FP سبک	FN سبک	TPR SVM1 سبک	PPV SVM1 سبک	TP سنگین	FP سنگین	FN سنگین	TPR SVM1 سنگین	PPV SVM1 سنگین
۱	۹۰	۵۱	۸۷	۰	۳	%۹۶/۷	%۱۰۰	۴۷	۱	۳	%۹۴	%۹۷/۹
۲	۴۰	۲۹	۴۰	۰	۰	%۱۰۰	%۱۰۰	۲۵	۲	۲	%۹۲/۶	%۹۲/۶
۳	۷۶	۴۱	۶۴	۰	۱۲	%۸۴/۲	%۱۰۰	۴۰	۰	۱	%۹۷/۶	%۱۰۰
۴	۱۸۳	۴۰	۱۷۹	۰	۴	%۹۷/۸	%۱۰۰	۳۶	۲	۲	%۹۴/۷	%۹۴/۷

در جدول ۴-۵ تعداد داده‌های آموزشی هر مرحله و نیز داده‌های اضافه شده از مراحل قبل به مجموعه آموزشی آورده شده است.

جدول ۴-۵: تعداد داده‌های آموزشی و نیز داده‌های اضافه شده از مراحل قبل به مجموعه آموزشی در هر مرحله

گام	داده‌های مثبت	داده‌های منفی	FN اضافه شده	TP اضافه شده	FP اضافه شده	TN اضافه شده	خودروهای سبک	خودروهای سنگین
۱	۱۶۵	۱۷۴۸	-	-	-	-	-	-
۲	۳۰۵	۱۷۹۹	۶	۱۳۴	۱	۵۰	-	-
۳	۳۷۲	۱۸۷۷	۲	۶۵	۲	۷۶	-	-
۴	۴۸۹	۱۹۴۰	۱۳	۱۰۴	۰	۶۳	۲۱۲	۲۷۷

الگوریتم‌های پیشنهادی به زبان C++ در محیط ویژوال استودیو ۲۰۱۳ و به کمک کتابخانه Opencv 2.4.8 پیاده سازی شده‌اند. زمان اجرای مجموعه الگوریتم‌های پیشنهادی تشخیص وجود خودرو و تعیین کلاس کلی خودرو بر روی هر قاب ویدئو، با استفاده از یک لپ‌تاپ Core i7 با هسته اینتل 3610QM، ۲/۳ گیگاهرتز و ۸ گیگابایت حافظه در حدود ۹ میلی ثانیه بوده است. پیش از این بیان شد که در روش [۱۵] از الگوریتم Adaboost و ویژگی‌های Haar برای یافتن خودرو در تصاویر ثابت یا ویدئو استفاده شده است، در این روش مجموعه‌ای از ۹۰۰۰ تصویر، متشکل از ۳۰۰۰ تصویر خودروی سبک به ابعاد ۲۵×۳۰، ۱۰۰۰ خودروی سنگین به ابعاد ۴۰×۴۰ و ۵۰۰۰ تصویر غیر خودرو برای آموزش در نظر گرفته شده است. در جدول ۵-۵ نتایج اجرای الگوریتم این روش به همراه تعداد داده‌های آزمون در دو حالت تشخیص با استفاده از تصاویر ثابت و ویدئو آورده شده است.

جدول ۵-۵: نتایج تشخیص خودرو مربوط به روش [۱۵]

نوع خودرو	تعداد داده آزمون	دقت تشخیص در تصاویر ثابت	دقت تشخیص در ویدئو
سبک	۲۰۰۰	٪۸۴/۶	٪۷۱/۴
سنگین	۱۰۰۰	٪۷۸/۷	٪۶۷/۳
کل	۳۰۰۰	٪۸۲/۶	٪۷۰

این روش بر روی یک رایانه Core2 Duo با هسته اینتل ۲/۶۶ گیگاهرتز و ۳ گیگابایت حافظه اجرا شده است و قابلیت پردازش ۳۰ قاب ویدئو در ثانیه را داراست، یا به طور معادل زمان اجرای الگوریتم بر روی هر قاب در حدود ۳۳ میلی ثانیه بوده است. اگرچه زمان اجرای الگوریتم [۱۵] نسبت به روش پیشنهادی بیش‌تر است اما بایستی در نظر داشت امکانات سخت افزاری به کار گرفته شده برای اجرای الگوریتم پیشنهادی بهتر بوده است، و از آن مهم‌تر، در روش پیشنهادی نیازی به جستجو در کل قاب ویدئو وجود ندارد. از طرفی همان طور که در جدول ۵-۵ آورده شده است برای به کارگیری الگوریتم Adaboost به نسبت به تعداد داده‌های آموزشی بیش‌تری نیاز است، در بخش ۲-۴ و در شکل ۲-۹ نمونه‌ای از تصاویر استفاده شده در روش [۱۵] نشان داده شده است که به طور کلی با تصاویر پایگاه داده ما متفاوت است. استفاده از SURF متقارن در روش [۱۰] معرفی شد، جدول ۵-۶ نتایج مربوط به روش [۱۰] برای یافتن نمای جلوی خودرو در سه نوع آب و هوای مختلف را نشان می‌دهد، در این جدول منظور از دقت، نسبت تعداد خودروی درست شناسایی شده به تعداد کل خودروهای شناسایی شده است.

جدول ۵-۶: نتایج مربوط به روش [۱۰] برای یافتن نمای جلوی خودرو در سه نوع آب و هوای مختلف

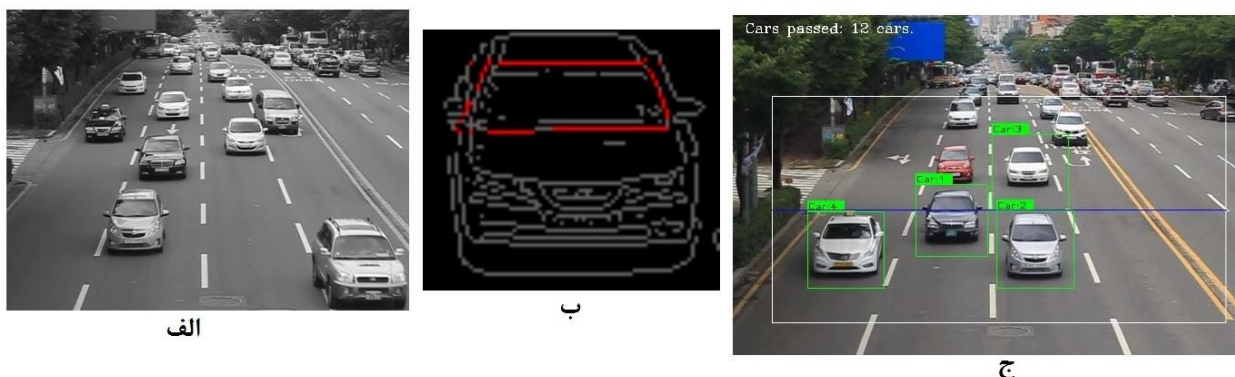
شرایط آب و هوایی	آفتابی	ابری	بارانی
تعداد تعداد خودروی درست شناسایی شده از ویدئوها	۱۶۱۳	۲۰۷۰	۱۷۶۱
تعداد کل خودروهای شناسایی شده از ویدئوها	۱۶۴۴	۲۰۹۴	۱۷۹۱
دقت	۹۸/۱۱٪	۹۸/۸۵٪	۹۸/۳۲٪

زمان لازم برای اجرای الگوریتم روش [۱۰] با استفاده از یک لپ‌تاپ Core i5 3337U و هسته اینتل ۱/۸ گیگاهرتز، ۴ گیگابایت حافظه، به ازای هر قاب ویدئو تقریباً برابر با ۲۲ میلی‌ثانیه است. اگرچه زمان اجرای این روش نسبت به روش پیشنهادی بیش‌تر است اما مزیت این روش نسبت به روش پیشنهادی عدم نیاز آن به تعریف ROI از پیش تعیین شده برای تشخیص ورود خودرو است، از طرفی با به کارگیری سخت

افزار قوی‌تر می‌توان عملکرد زمانی آن را بهبود داد. با وجود تمامی مزایای استفاده از روش [۱۰] به سبب دلایل بیان شده در بخش ۴-۲-۳ از این روش استفاده نشد. در روش [۱۳] عملکرد سیستم یافتن خودرو بسیار خوب است و دقت ۱۰۰٪ حاصل شده است، اما این روش بر روی تصاویر ثابت اجرا شده و زمان اجرای الگوریتم‌های یافتن و تشخیص نوع خودرو برای هر تصویر به ابعاد 1600×1200 مجموعاً در حدود ۱ ثانیه بوده است. روش [۲۷] به کمک لبه‌یابی و تبدیل هاف، نواحی دوزنقه‌ای شکل را به عنوان شیشه جلو خودرو می‌یابد و با توجه به آن ناحیه در بر دارنده خودرو را استخراج می‌کند. سپس وجود خودرو در این نواحی به کمک بردار ویژگی HOG و طبقه‌بند SVM تأیید یا رد می‌شود، شکل ۵-۳ نمونه تصاویر پایگاه داده استفاده شده در این روش را نشان می‌دهد. پارامترهای انتخاب شده برای استخراج HOG در این روش در جدول ۵-۷ آورده شده‌اند.

جدول ۵-۷: پارامترهای انتخاب شده برای استخراج HOG در روش [۲۷]

تعداد بازه	ابعاد بلوک	ابعاد سلول
۹	16×16	8×8



شکل ۵-۳: نمونه تصاویر پایگاه داده روش [۲۷]: (الف) قاب نمونه از ویدئوی پایگاه داده (ب) یافتن ناحیه دوزنقه‌ای به عنوان شیشه جلوی خودرو (ج) خودروهای پیدا شده

ویدئوهای مورد استفاده که به کمک یک دوربین مداربسته واقع در ارتفاع ۶ متری بالای یک جاده چهار بانده ضبط شده‌اند، دارای نرخ قاب ۳۰ قاب بر ثانیه هستند. ابعاد هر قاب آن‌ها برابر 640×480 بوده و با توجه به پارمترهای اختیار شده، بردارهای HOG استخراج شده 1×1764 بوده‌اند. ۶۰۰۰ تصویر از خودروهای بدون انسداد و ۴۰۰ تصویر دارای انسداد به عنوان نمونه‌های مثبت و نیز ۱۳۸۰ تصویر هم به عنوان نمونه‌های منفی برای استخراج ویژگی و آموزش SVM، از ویدئوها جدا شده است. این روش که در آن دقت بازشناسی برابر با ۹۸٪ بوده، بر روی یک رایانه Core2 Duo با هسته اینتل ۲/۶۶ گیگاهرتز، ۱ گیگابایت حافظه و با به کارگیری روش‌های پردازش موازی و به کمک GPU^۱ شرکت NVIDIA اجرا شده است و توانایی پردازش ۱۵ قاب در ثانیه را داراست. در روش پیشنهادی پایان‌نامه، برای پیاده‌سازی الگوریتم از پردازش موازی استفاده نشده است، اما از سخت افزار سطح بالاتری استفاده شده است، با تمام این توضیحات و در نظر گرفتن محدودیت‌های اعمال شده در روش پیشنهادی، همچون در نظر گرفتن پنجره ثابت، باز هم روش پیشنهادی در قیاس با روش [۲۷] از سرعت قابل قبولی برخوردار است.

۵-۲-۲ نتایج مربوط به یافتن ROI جلوی خودروی سبک و تعیین مدل آن

الگوریتم جداسازی ROI جلوی خودرو کاملاً دقیق است و به خوبی ناحیه جلوی تمامی ۱۷۹ خودروی سبک تشخیص داده شده از ویدئوی V5 را جدا می‌کند. همان طور که در بخش ۴-۴-۲ توضیح داده شد، برای تعیین مدل خودروی سبک، چهار طبقه‌بند SVM "یک نوع در برابر بقیه" در نظر گرفته شده است. برای آموزش این طبقه‌بندها به این شکل عمل می‌شود که تعدادی داده آموزشی از چهار ویدئوی اول جداسازی می‌شود و داده‌های مربوط به هریک از چهار نوع خودرو هم جدا می‌شوند. سپس هریک از SVMها با استفاده از این داده‌ها آموزش داده می‌شوند و طی یک فرآیند جستجوی شبکه‌ای^۲ و با استفاده

^۱ Graphics Processing Unit (GPU)

^۲ Grid Search

از تعدادی داده ارزیابی، کرنل و پارامترهایی که حاصلضرب سه معیار دقت^۱، حساسیت^۲ و صحت^۳ را بیشینه می‌کنند به عنوان پارامترها و کرنل SVM در نظر گرفته می‌شوند، در جدول ۵-۸ مشخصات اختیار شده برای هریک از چهار SVM آورده شده است. ۲۱ تصویر از نوع GLX، ۳۰ تصویر از نوع Pride، ۱۶ تصویر از نوع L90، ۱۵ تصویر از نوع Peykan و ۵۰ تصویر از سایر انواع خودروها برای آموزش SVMها در نظر گرفته شده است. سپس به کمک این مجموعه طبقه‌بند، مدل خودرو در یکی از پنج دسته L90، GLX، Pride، Peykan و یا غیره قرار می‌گیرد.

جدول ۵-۸: مشخصات به دست آمده با جستجوی شبکه‌ای برای طبقه‌بندهای تعیین مدل روش پیشنهادی

درجه	مقدار coef0	مقدار γ	مقدار C	کرنل	طبقه‌بند
۱	۰	۰/۱	۱	چند جمله‌ای	GLX در برابر بقیه
-	۰	۰/۲	۱	RBF	Pride در برابر بقیه
-	-	-	۱	خطی	L90 در برابر بقیه
-	-	-	۰/۲	خطی	Peykan در برابر بقیه

در نهایت، الگوریتم تعیین مدل بر روی ۱۷۹ تصویر خودروی سبک جدا شده از ویدئوی V5 (گام ۴ در جدول ۵-۲) مورد آزمون قرار می‌گیرد، نتایج حاصل از آزمون به همراه تعداد داده‌های آزمون مورد استفاده در جدول ۵-۹ آورده شده است.

^۱ Positive Predictive Value (PPV)

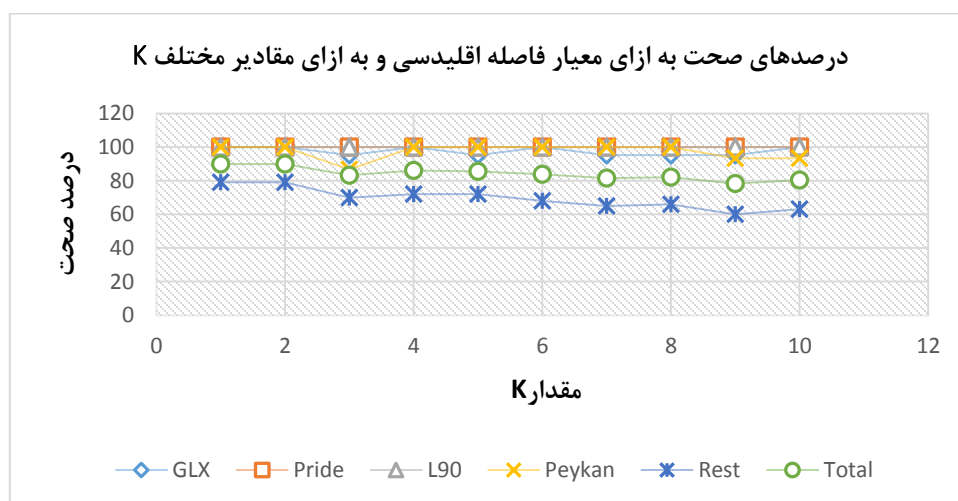
^۲ True-Positive Rate (TPR)

^۳ Accuracy

جدول ۵-۹: نتایج حاصل از آزمون الگوریتم پیشنهادی تعیین مدل

تعداد کل	تعداد تشخیص درست	تعداد تشخیص اشتباه	درصد صحت	داده
۲۰	۲۰	۰	۱۰۰٪	GLX
۲۸	۲۶	۲	۹۲/۸۶٪	Pride
۱۶	۱۶	۰	۱۰۰٪	L90
۱۵	۱۳	۲	۸۶/۶۷٪	Peykan
۱۰۰	۹۸	۲	۹۸٪	غیره
۱۷۹	۱۷۳	۶	۹۶/۶۵٪	در کل

زمان اجرای الگوریتم پیشنهادی برای هریک از بخش‌های جداسازی قسمت جلوی خودرو و تعیین مدل در حدود ۳ میلی‌ثانیه و در مجموع ۶ میلی‌ثانیه بوده است. برای مقایسه، طبقه‌بند KNN با مقادیر مختلف K و معیارهای مختلف فاصله از جمله اقلیدسی، بلوک شهری^۱ و کسینوسی نیز مورد آزمایش قرار گرفت، در نمودار ۵-۱ نتایج حاصل از آزمایش KNN برای هریک از ۵ مجموعه داده، با معیار فاصله اقلیدسی و به ازای مقادیر مختلف K نشان داده شده است، همچنین بهترین نتایج که به ازای K=1 حاصل شده است در جدول ۵-۱۰ آورده شده است.



نمودار ۵-۱: نتایج حاصل از آزمایش KNN با معیار فاصله اقلیدسی و به ازای مقادیر مختلف

^۱ City block

جدول ۵-۱۰: نتایج حاصل از آزمایش KNN با معیار فاصله اقلیدسی و به ازای مقادیر مختلف K

صحت کلی	صحت غیره	صحت Peykan	صحت L90	صحت Pride	صحت GLX	K
٪۸۹/۹۴	٪۷۹	٪۱۰۰	٪۱۰۰	٪۱۰۰	٪۱۰۰	۱

بهترین نتایج مربوط به معیارهای بلوک شهری و کسینوسی هم همانند نتایج جدول ۵-۱۰ بوده‌اند، مشاهده می‌شود نتایج مربوط به SVMها بهتر بوده است.

در روش [۱۰] برای تعیین مدل، هر مرتبه چهار ویژگی مختلف از هریک از شبکه‌های نشان داده شده در شکل ۲-۶ ب استخراج شده است، همچنین چهار روش مختلف برای آموزش طبقه‌بند خودرو در نظر گرفته شده‌اند، که در هر کدام از درصد خاصی از شبکه‌ها برای آموزش طبقه‌بند استفاده شده است، در جدول ۵-۱۱ درصد متوسط صحت حاصل و نیز زمان پردازش هر قاب ویدئو به ازای استفاده هریک از چهار ویژگی آورده شده است.

جدول ۵-۱۱: درصد صحت و زمان اجرای الگوریتم تشخیص مدل خودرو در روش [۱۰]

ویژگی	HOG	SURF	SIFT	SURF-SIFT
متوسط درصد صحت	۹۴/۲۹	۹۶/۸۲	۹۴/۷۸	۹۵/۸۳
زمان اجرا به ازای هر قاب	۴۲ میلی ثانیه	۴۸ میلی ثانیه	۷۷۰ میلی ثانیه	۶۳ میلی ثانیه

مشاهده می‌شود که زمان اجرای روش [۱۰] به مراتب طولانی‌تر از روش پیشنهادی است، همچنین دقت یافتن خودرو برابر با ۹۸/۴۸٪ بوده است. در روش [۱۳] میزان صحت تشخیص مدل خودرو به طور میانگین برابر ۹۹٪ بوده است، اما زمان اجرای آن طولانی است. در روش [۹] ۱۴۱۴ تصویر از ۲۰ مدل خودرو در اختیار بوده است که از این میان ۶۰۰ تصویر برای آموزش (۳۰ تصویر از هر مدل) و ۸۱۴ تصویر برای آزمون در نظر گرفته شده است. با استفاده از تحلیل مجزاساز خطی دو بعدی و طبقه‌بند نزدیک‌ترین

همسایگی به تعیین مدل خودروها پرداخته شده است که دقت بازشناسی ۹۴٪/۷ حاصل شده و زمان اجرای آن به کمک پردازنده‌ای با فرکانس ۲/۰۱ گیگاهرتز برابر ۳۷۱ میلی‌ثانیه بوده است. در [۱۴] ۳۲۱۰ تصویر برای آموزش ۱۰۷ مدل خودرو با استفاده از شبکه عصبی پیچشی (۳۰ تصویر از هر مدل) مورد استفاده قرار گرفته است و دقت بازشناسی ۸۸/۴٪ به دست آمده است. در روش [۱۳] به کمک تطبیق الگو با استفاده از مجموع مطلق تفاضل‌ها^۱ دقت بازشناسی در حدود ۹۵٪ بوده است و زمان اجرای آن بر روی هر تصویر با استفاده از پردازنده AMD با فرکانس ۲/۸۱ گیگاهرتز در حدود ۹ میلی‌ثانیه بوده است. در [۲۸] بر روی ویدئوهای یک دوربین مدار بسته^۲ کار شده است، ۱۰ تصویر به ازای هر یک از ۲۲ مدل خودرو برای آموزش در نظر گرفته شده و برای آزمون از یک ویدئوی مشابه با نرخ ۲۵ قاب بر ثانیه و شامل ۲۴ خودروی عبوری استفاده شده است. ابعاد هر قاب هم برابر ۷۵۶×۷۲۰ است. پس از یافتن ناحیه جلوی خودرو با توجه با موقعیت و اندازه پلاک، به کمک ویژگی‌ها و طبقه‌بندهای مختلف همانطور که در جدول ۵-۱۲ نشان داده شده است کار طبقه‌بندی انجام گرفته است که در بهترین حالت دقت بازشناسی در حدود ۹۵٪ به دست آمده است. در این جدول اعداد نوشته شده در برابر ویژگی‌ها ابعاد بردار ویژگی هستند. زمان اجرای این روش در حدود ۲۰۰ میلی‌ثانیه بوده است.

جدول ۵-۱۲: نتایج مربوط به روش [۲۸]

	LESH-128 ^۳	LESH-512	HOG-576
KNN(K=1)	۷۰/۸۳	۷۰/۸۳	۶۲/۵
KNN(K=3)	۷۰/۸۳	۷۰/۸۳	۶۲/۵
SVM (one vs one)	۷۰/۸۳	۹۱/۶۷	۵۸/۳۳
SVM (one vs all)	۷۵	۹۵/۸۳	۶۶/۶۷

^۱ Sum of Absolute Differences (SAD)

^۲ Closed-Circuit television (CCTV)

^۳ Local Energy Shape Histogram (LESH)

در نهایت برای جمع‌بندی و مقایسه کلی، مشخصات مختلف متناظر با بهترین نتایج روش‌های یاد شده و روش پیشنهادی در جدول ۵-۱۳ آورده شده است، در این جدول خانه‌های مربوط به مواردی که در روش‌ها ذکر نشده‌اند با خط فاصله پر شده‌اند.

جدول ۵-۱۳: مقایسه بهترین نتایج روش‌های یاد شده و روش پیشنهادی

روش پیشنهادی	روش [۲۸]	روش [۲۷]	روش [۱۴]	روش [۱۳]	روش [۱۰]	روش [۹]	مشخصات
۱۹۰	۱۹۶	-	-	-	۲۸۴۶	۶۰۰	تعداد تصاویر آموزشی
۱۷۹	-	-	-	-	۴۰۹۰	۸۱۴	تعداد تصاویر آزمون یا خودروهای موجود در ویدئوی آزمون
۳۶۹	-	۶۰۰۰	۳۲۱۰	۱۱۹۶	۶۹۳۶	۱۴۱۴	کل تصاویر
۴	۲۲	۳۶	۱۰۷	۳۰	۲۹	۲۰	تعداد مدل
۷۲۰×۵۷۶	۷۲۰×۵۷۶	-	۶۷×۶۷	۱۶۰۰×۱۲۰۰	-	۱۰۰×۵۰	ابعاد تصاویر
۱۷۰×۱۳۰	-	-	۶۷×۶۷	-	-	-	ابعاد ROI
۲۵	۲۵	-	-	-	-	-	نرخ قاب ویدئو fps
۳۰دقیقه	۲دقیقه	-	-	-	-	-	مدت ویدئو
HOG	LESH	پیکسل-های خام	پیکسل-های خام	پیکسل‌های لبه	SURF	پیکسل خام با 2DLDA	ویژگی
۱۷۶۴	۵۱۲	-	-	-	۶۴	۳۶×۲۰	ابعاد ویژگی
SVM	SVM	تطبیق الگو با SAD	CNN	شبکه عصبی	SVM	نزدیک‌ترین همسایگی	طبقه‌بند
۱۰۰	-	-	۱۰۰	۱۰۰	-	-	درصد دقت یافتن ROI
۹۵/۷	۹۵/۸۳	۹۵	۸۸/۴	۹۹	۹۶/۸۲	۹۴/۷	درصد دقت تشخیص مدل

پیشنهادی	روش [۲۸]	روش [۲۷]	روش [۱۴]	روش [۱۳]	روش [۱۰]	روش [۹]	مشخصات
۳	-	۲۳۶	-	۵۰۰	۲۲	-	زمان اجرای یافتن ROI ms
۳	-	۲۹۲	-	-	۴۸	-	زمان اجرای تشخیص مدل ms
۶	۲۰۰	۵۲۸	-	-	۷۰	۳۷۱	زمان اجرای مجموع ms

۵-۲-۳ تشخیص رنگ خودروی سبک

فرآیندهای قبلی به صورت متوالی و بدون وقفه به انجام می‌رسیدند اما برای تشخیص رنگ، به عنوان یک بخش مجزا، تعداد ۳۸۹ تصویر از نمای روبه‌روی خودرو تهیه شد و الگوریتم مربوطه بر روی آن‌ها اجرا شد. از این میان ۱۷ خودرو رنگی و مابقی غیر رنگی بوده‌اند. دقت مربوط به تعیین رنگی یا غیر رنگی بودن خودرو و تعیین رنگ خودروهای رنگی ۱۰۰٪ بوده است و دقت مربوط به تعیین مشکی و روشن بودن خودروهای غیررنگی هم ۹۹٪ و زمان اجرای الگوریتم در حدود ۴ میلی‌ثانیه به ازای هر تصویر بوده است.

۵-۲-۴ تشخیص جهت حرکت خودرو

ویدئویی به مدت حدود ۱۰ دقیقه با نرخ ۳۰ قاب در ثانیه تهیه شد که ابعاد هر قاب آن برابر ۱۲۸۰×۷۲۰ است. در این ویدئو ۱۹ خودرو از صحنه خارج و ۲۰ خودرو به صحنه وارد می‌شوند. تمام خودروها به درستی شمارش و جهت حرکت همگی به درستی تشخیص داده شد و زمان اجرای الگوریتم در حدود ۱۵ میلی‌ثانیه بود.

۶. جمع بندی

۶-۲ جمع بندی

در این پایان نامه روش های مختلفی در مورد جداسازی خودرو و تشخیص نوع آن مورد بررسی قرار گرفت که در هر یک بسته به نیاز و کاربرد و پایگاه داده تهیه شده اش رویکردی متفاوت از سایر روش ها در پیش گرفته شده بود. روش های مربوط به تشخیص نوع خودرو در سه دسته مرور شدند: تشخیص و طبقه بندی نوع کلی خودرو، تشخیص و طبقه بندی مدل خودرو و تشخیص و جداسازی وسیله نقلیه. همچنین روش های مربوط به تشخیص رنگ و جهت حرکت خودرو نیز در دو بخش مجزا مرور شدند. اساس کار در مراحل تعیین کلاس کلی خودرو و تعیین مدل خودروی سبک استفاده از HOG و SVM بود. با توجه به ماهیت تصاویر پایگاه داده تهیه شده یک روش پیشنهادی برای جداسازی ناحیه جلوی خودروی سبک معرفی شد که طی آن یک مجموعه عملیات شامل یافتن سایه زیر خودرو، لبه یابی، استفاده از تبدیل هاف احتمالاتی و یافتن هیستوگرام های افقی و عمودی تصویر پیش زمینه به انجام می گرفت. در این پایان نامه سیستم تشخیص مدل خودرو تنها برای چهار نوع خودروی L90، GLX، Pride و Peykan شکل داده شد چرا که این چهار مدل در بین خودروهای موجود در جاده های ایران نسبتاً بیشترین فراوانی و تمایز شکلی را با سایر مدل های موجود داشتند، از طرفی کمبود داده های آموزشی از انواع دیگر خودروها در پایگاه داده دلیل دیگری برای انتخاب انحصاری این چهار مدل بود. برای تشخیص مدل خودرو چهار SVM یک در برابر بقیه در نظر گرفته شد که خروجی آن ها علاوه بر برچسب گروه مربوطه، میزان احتمال و درستی برچسب تشخیص داده شده نیز بود. برای تعیین رنگ خودروهای سبک به دلیل محدودیت ناشی از کیفیت پایین تصاویر پایگاه داده خودروهای رنگی با توجه به مقدار Hue در چهار دسته قرمز، سبز، آبی و زرد و خودروهای غیر رنگی با توجه به شدت رنگ در دو دسته روشن و مشکی قرار گرفتند. در انتها برای بخش تشخیص جهت خودرو ویدئویی متفاوت تهیه شد، از GMM برای تخمین پیش زمینه استفاده شد، کانتور خودروها و مستطیل محیطی آن ها به دست آورده شد و با شار نوری جهت حرکت

خودرو تشخیص داده شد. هریک از مراحل یاد شده در مدت زمانی در حدود چند میلی ثانیه به اجرا در آمدند و بلادرنگ بودند. به طور کلی سعی شد به سبب کثرت روش‌های مختلف در زمینه تشخیص نوع خودرو با انواع دادگان موجود، با وجود تمامی محدودیت‌های موجود، دادگانی تهیه و بر روی آن کار شود که در ایران کمتر مورد توجه قرار گرفته است.

۳-۶ پیشنهادات برای کارهای آینده

پیشنهاد می‌شود که برای تهیه پایگاه داده از دوربین و تجهیزات با کیفیت‌تر استفاده شود و ویدئوهای طولانی‌تری تهیه شوند و دوربین به نحوی تعبیه شود که تنها نمای روبه‌روی خودروی جلویی قابل مشاهده باشد و انسداد بین خودروها قابل رویت نباشد. به این ترتیب با داشتن ویدئوهای بیش‌تر روش‌هایی که نیازمند تعداد زیادی داده آموزشی هستند (همچون روش Adabbost) هم قابل پیاده‌سازی هستند. همچنین دقت بیش‌تر در تشخیص رنگ خودروها امکان‌پذیر خواهد شد. از طرفی در صورت تهیه دادگان بیش‌تر، تشخیص سایر مدل‌های خودروهای سبک و نیز تشخیص مدل خودروهای سنگین ممکن خواهد شد.

٧. مراجع

- [1] Peijin, J., L. Jin, and X. Li. *Vision-based Vehicle Type Classification Using Partial Gabor Filter Bank*. in *Automation and Logistics, 2007 IEEE International Conference on*. 2007.
- [2] Yuren, D., L. Caiyun, and S. Wen. *A real-time vehicle recognition method based on video sequence images*. in *Electronic Measurement & Instruments, 2009. ICEMI '09. 9th International Conference on*. 2009.
- [3] McHugh, J.M., et al., *Foreground-Adaptive Background Subtraction*. *Signal Processing Letters, IEEE*, 2009. **16**(5): p. 390-393.
- [4] Zhang, B. and Y. Zhou. *Reliable vehicle type classification by Classified Vector Quantization*. in *Image and Signal Processing (CISP), 2012 5th International Congress on*. 2012.
- [5] Yu, P., et al. *Vehicle Type Classification Using PCA with Self-Clustering*. in *Multimedia and Expo Workshops (ICMEW), 2012 IEEE International Conference on*. 2012.
- [6] Yuqiang, L. and W. Kunfeng. *Vehicle classification system based on dynamic Bayesian network*. in *Service Operations and Logistics, and Informatics (SOLI), 2014 IEEE International Conference on*. 2014.
- [7] Kafai, M. and B. Bhanu, *Dynamic Bayesian Networks for Vehicle Classification in Video*. *Industrial Informatics, IEEE Transactions on*, 2012. **8**(1): p. 100-109.
- [8] Mingxie, Z., T. Gotoh, and M. Shiohara. *A Hierarchical Algorithm for Vehicle Model Type Recognition on Time-Sequence Road Images*. in *Intelligent Transportation Systems Conference, 2006. ITSC '06. IEEE*. 2006.
- [9] Hua, H., et al. *A 2DLDA Based Algorithm for Real Time Vehicle Type Recognition*. in *Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on*. 2008.
- [10] Jun-Wei, H., C. Li-Chih, and C. Duan-yu, *Symmetrical SURF and Its Applications to Vehicle Detection and Vehicle Make and Model Recognition*. *Intelligent Transportation Systems, IEEE Transactions on*, 2014. **15**(1): p. 6-20.
- [11] Chen, L.-C., et al., *Vehicle make and model recognition using sparse representation and symmetrical SURFs*. *Pattern Recognition*, 2015. **48**(6): p. 1979-1998.
- [12] Bailing, Z., *Reliable Classification of Vehicle Types Based on Cascade Classifier Ensembles*. *Intelligent Transportation Systems, IEEE Transactions on*, 2013. **14**(1): p. 322-332.
- [13] Hongsheng, H., S. Zhenzhou, and T. Jindong, *Recognition of Car Makes and Models From a Single Traffic-Camera Image*. *Intelligent Transportation Systems, IEEE Transactions on*, 2015. **16**(6): p. 3182-3192.
- [14] Gao, Y. and H.J. Lee. *Vehicle Make Recognition Based on Convolutional Neural Network*. in *Information Science and Security (ICISS), 2015 2nd International Conference on*. 2015.

- [15] Chunpeng, W., et al. *Detection of Front-View Vehicle with Occlusions Using AdaBoost*. in *Information Engineering and Computer Science*, 2009. ICIECS 2009. *International Conference on*. 2009.
- [16] Viola, P. and M. Jones. *Rapid object detection using a boosted cascade of simple features*. in *Computer Vision and Pattern Recognition*, 2001. CVPR 2001. *Proceedings of the 2001 IEEE Computer Society Conference on*. 2001.
- [17] Xuezhi, W., et al. *An improved wavelet feature extraction approach for vehicle detection*. in *Vehicular Electronics and Safety*, 2007. ICVES. *IEEE International Conference on*. 2007.
- [18] Bo, Q., et al. *Fast Detection of Vehicles Based-on the Moving Region*. in *Computer-Aided Design and Computer Graphics*, 2007 10th *IEEE International Conference on*. 2007.
- [19] Ming, Y., et al. *An HMM-Based Algorithm for Vehicle Detection in Congested Traffic Situations*. in *Intelligent Transportation Systems Conference*, 2007. ITSC 2007. *IEEE*. 2007.
- [20] Brown, L.M., A. Datta, and S. Pankanti. *Tree-based vehicle color classification using spatial features on publicly available continuous data*. in *Advanced Video and Signal Based Surveillance (AVSS)*, 2013 10th *IEEE International Conference on*. 2013.
- [21] Chen, L.C., et al. *Real-time vehicle color identification using symmetrical SURFs and chromatic strength*. in *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*. 2015.
- [22] Dong, Y., M. Pei, and X. Qin. *Vehicle Color Recognition Based on License Plate Color*. in *Computational Intelligence and Security (CIS)*, 2014 Tenth *International Conference on*. 2014.
- [23] Miller, N., et al. *A Hidden Markov Model for Vehicle Detection and Counting*. in *Computer and Robot Vision (CRV)*, 2015 12th *Conference on*. 2015.
- [24] Jiajun, L., X. Yi, and Y. Xiaokang. *Counting pedestrians and cars with an improved virtual gate method*. in *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*. 2010.
- [25] Chang, C.-C. and C.-J. Lin, *LIBSVM: A library for support vector machines*. *ACM Trans. Intell. Syst. Technol.*, 2011. **2**(3): p. 1-27.
- [26] Kiryati, N., Y. Eldar, and A.M. Bruckstein, *A probabilistic Hough transform*. *Pattern Recognition*, 1991. **24**(4): p. 303-316.
- [27] Van Pham, H. and B.-R. Lee, *Front-view car detection and counting with occlusion in dense traffic flow*. *International Journal of Control, Automation and Systems*, 2015. **13**(5): p. 1150-1160.
- [28] Saravi, S. and E.A. Edirisinghe. *Vehicle Make and Model Recognition in CCTV footage*. in *Digital Signal Processing (DSP)*, 2013 18th *International Conference on*. 2013.

Abstract

On road vehicle type recognition generally plays an important role in intelligent transportation applications, increasing security and ease of movement on roads. Already existing methods for vehicle type recognition are the ones that get the information required for recognition using different sensors, including laser, radar and vision sensors and study such information, but among all, vision based methods have been more attractive due to reliability of results and low hardware and time costs. Vehicle type recognition can be done in two levels: Recognizing type based on dimension and general appearance properties of vehicle (like classifying vehicles into classes of heavy, light and so forth) or more precisely recognizing vehicle make and model. In this thesis for vehicle type and color recognition sections videos of cars passing through a toll-gate are used. These kind of videos are less paid attention because of their limited field of view and difficulties of vehicle region of interest extraction. Type and model recognition of vehicles passing through toll-gates is necessary for security applications, statistical surveys, and automated toll collection and so on. A real-time algorithm is presented and the average recall and precision of the detection stage are 94.75% and 98.83% respectively and the accuracy of the vehicle general type and vehicle model recognition stages are 98% and 96.65% respectively.

Keywords: Vehicle Type Recognition, Video, Intelligent Transportation, Real-time, HOG



Shahrood University of Technology
Department of Electrical and Robotics Engineering

Real-time detection and classification of on-road vehicles type, color and movement direction in video sequences

Ahmad Mosayebi

Supervisor:

Dr. Hossein Khosravi

July 2016