

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده : برق و رباتیک

گروه : الکترونیک

پایان نامه کارشناسی ارشد مهندسی برق-الکترونیک

آشکارسازی و شناسایی علائم راهنمایی و رانندگی با استفاده از شبکه‌های عصبی و انطباق
الگو

دانشجو : مسعود قویدل

استاد راهنما :

دکتر حسین خسروی

پایان نامه ارشد جهت اخذ درجه کارشناسی ارشد

ماه و سال انتشار : شهریور ۱۳۹۳

دانشگاه شاهرود

دانشکده : برق و رباتیک

گروه : الکترونیک

پایان نامه کارشناسی ارشد آقای مسعود قویدل

تحت عنوان: آشکارسازی و شناسایی علائم راهنمایی و رانندگی با استفاده از شبکه‌های عصبی و انطباق الگو

در تاریخ توسط کمیته تخصصی زیر جهت اخذ مدرک کارشناسی ارشد مورد ارزیابی و با درجه مورد پذیرش قرار گرفت.

امضاء	اساتید مشاور	امضاء	اساتید راهنما
	نام و نام خانوادگی :		نام و نام خانوادگی :
	نام و نام خانوادگی :		نام و نام خانوادگی :

امضاء	نماینده تحصیلات تکمیلی	امضاء	اساتید داور
	نام و نام خانوادگی :		نام و نام خانوادگی :
			نام و نام خانوادگی :
			نام و نام خانوادگی :
			نام و نام خانوادگی :

تقدیم

به پدر و مادر مهربانم

تشکر و قدردانی

سپاس خدای را که سخنوران، در ستودن او بمانند و شمارندگان، شمردن نعمت های او ندانند و کوشندگان، حق او را گزاردن نتوانند. و سلام و دورد بر محمد و خاندان پاک او، طاهران معصوم، هم آنان که وجودمان وامدار وجودشان است؛ و نفرین پیوسته بر دشمنان ایشان تا روز رستاخیز...

بدون شک جایگاه و منزلت معلم، اجل از آن است که در مقام قدردانی از زحمات بی شائبه‌ی او، با زبان قاصر و دست ناتوان، چیزی بنگاریم.

اما از آنجایی که تجلیل از معلم، سپاس از انسانی است که هدف و غایت آفرینش را تامین می کند و سلامت امانت‌هایی را که به دستش سپرده اند، تضمین؛ بر حسب وظیفه و از باب "من لم یشکر المنعم من المخلوقین لم یشکر الله عزّ و جلّ": ازپدر و مادر عزیزم، این دو معلم بزرگوام، که همواره بر کوتاهی و درستی من، قلم عفو کشیده و کریمانه از کنار غفلت‌هایم گذشته‌اند و در تمام عرصه های زندگی یار و یآوری بی چشم داشت برای من بوده اند؛ از استاد با کمالات و شایسته؛ جناب آقای دکتر خسروی که در کمال سعه صدر، با حسن خلق و فروتنی، از هیچ کمکی در این عرصه بر من دریغ ننمودند و زحمت راهنمایی این پایان‌نامه را بر عهده گرفتند، کمال تشکر و قدردانی را دارم.

همچنین برخورد واجب می‌دانم از تمامی اساتید گروه الکترونیک دانشگاه شاهرود تشکر و قدردانی لازم را نموده باشم.

تعهد نامه

اینجانب **مسعود قویدل** دانشجوی دوره کارشناسی ارشد رشته مهندسی برق الکترونیک گرایش دیجیتال دانشکده برق و رباتیک دانشگاه شاهرود نویسنده پایان نامه آشکارسازی و شناسایی علائم راهنمایی و رانندگی با استفاده از شبکه‌های عصبی و انطباق الگو تحت راهنمایی دکتر حسین خسروی متعهد می‌شوم .

- 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 این اثر به دانشگاه شاهرود می‌باشد 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 می‌گردد.
- 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 اخلاق انسانی رعایت شده است.

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزار ها و تجهیزاتی که 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.

چکیده

در این پایان‌نامه به آشکارسازی و شناسایی علائم راهنمایی و رانندگی می‌پردازیم. در مرحله اول، علامت ترافیکی را از میان اشیاء مختلفی که درون تصویر وجود دارد، استخراج کرده، سپس در مرحله دوم به شناسایی نوع دقیق آن می‌پردازیم. در مرحله آشکارسازی علامت، ابتدا به منظور کاهش حساسیت نسبت به تغییرات روشنایی، تصویر ورودی را از فضای رنگ RGB به فضای رنگ HSV تبدیل می‌کنیم و سپس با آستانه‌گذاری بر روی پارامترهای فام^۱ و اشباع، مناطق با رنگ قرمز و آبی را استخراج می‌کنیم.

هم‌چنین به منظور افزایش قدرت سیستم در استخراج مناطق با رنگ قرمز و آبی در شرایط متغیر روشنایی، از منطق فازی استفاده می‌کنیم. سپس با فیلترگذاری، از روی مناطق قرمز و آبی استخراج شده قسمت‌هایی که مثلث، دایره، هشت‌ضلعی و یا مستطیل هستند، استخراج می‌شوند. به منظور ارزیابی عملکرد سیستم آشکارسازی و نبود یک پایگاه داده از علائم راهنمایی و رانندگی موجود در ایران، اقدام به تهیه یک پایگاه داده از علائم موجود در ایران کردیم. این پایگاه داده شامل ۳۵۰ علامت در ۳۵ کلاس مختلف است، سیستم ما قادر به آشکارسازی ۹۷.۸۷٪ درصد از این علائم است. پس از استخراج علامت از میان تصویر، پس زمینه آن را حذف کرده و در روش اول بکارفته برای شناسایی دقیق علامت به روش پروفایل از علامت استخراج ویژگی می‌کنیم. سپس با استفاده ۸۴۰ علامت در ۴۰ کلاس مختلف، ماشین بردار پشتیبان را آموزش داده و با استفاده از ۴۱۶ علامت آن را مورد آزمایش قرار داده که به دقت شناسایی ۹۱.۱٪ رسیدیم.

در روش دوم که برای شناسایی دقیق علامت به کار برده‌ایم، به روش هیستوگرام گرادیان از علامت راهنمایی و رانندگی استخراج ویژگی می‌کنیم. سپس از ویژگی‌های استخراج شده برای آموزش شبکه عصبی استفاده می‌کنیم. مجموعه داده‌ای که برای آموزش و آزمایش شبکه عصبی استفاده شده است شامل ۱۲۸۰ علامت واقعی در ۴۰ کلاس مختلف است. ۶۷ درصد این مجموعه یعنی ۸۶۴ علامت را

^۱ Hue

به آموزش شبکه اختصاص داده و از ۳۳ درصد باقی مانده یعنی ۴۱۶ علامت برای آزمایش شبکه استفاده می کنیم. دقت طبقه بندی به دست آمده، ۹۹ درصد بر روی داده های آموزش و ۹۳.۸۶ درصد بر روی داده های آزمایش است. با توجه به این که این علائم در شرایط نوری، پس زمینه ها و مقیاس های مختلف قرار دارند و نیز دارای تاری و چرخش هستند. نتایج به دست آمده بیانگر توانایی روش پیشنهادی در آشکارسازی و تشخیص علائم است.

برای این که الگوریتم ارائه شده به صورت بلادرنگ عملیات آشکارسازی و شناسایی دقیق علامت را انجام دهد، برنامه اولیه را که با استفاده از نرم افزار متلب نوشتیم، به زبان C++ تبدیل کردیم. که به طور میانگین، به زمان آشکارسازی ۱۷۰ میلی ثانیه و زمان شناسایی دقیق ۱۲۱ میلی ثانیه برای هر تصویر رسیدیم.

واژه های کلیدی : فضای رنگ HSV، منطق فازی، ماشین بردار پشتیبان، ویژگی پروفایل، استخراج ویژگی، هیستوگرام گرادیان ، شبکه عصبی

لیست مقالات مستخرج از پایان نامه

- ۱- "شناسایی علائم راهنمایی و رانندگی در صحنه‌های واقعی با استفاده از ماشین بردار پشتیبان"، پذیرفته شده در ششمین کنفرانس فناوری اطلاعات و دانش، IKT 2014.
- ۲- "آشکارسازی و شناسایی علائم راهنمایی و رانندگی با استفاده از شبکه عصبی و منطق فازی"، ارسال شده به مجله هوش محاسباتی در مهندسی برق.

فهرست مطالب

فصل اول:

مقدمه.....	۱
۱-۱ مقدمه.....	۲
۲-۱ دسته‌بندی کلی علائم راهنمایی و رانندگی.....	۳
۳-۱ مراحل شناسایی خودکار علائم راهنمایی و رانندگی.....	۳
۴-۱ هدف پایان‌نامه.....	۶
فصل دوم: مروری بر کارهای انجام شده.....	۸
۱-۲ روش تقاطع خطوط مرور.....	۹
۲-۲ روشی برای استخراج ویژگی از علامت.....	۱۱
۳-۲ تقسیم‌بندی تصویر.....	۱۲
۴-۲ استخراج و شناسایی علامت از درون تصویر.....	۱۴
۱-۴-۲ استفاده از svm برای شناسایی علامت.....	۱۵
۲-۴-۲ چند روش برای استخراج ویژگی و شناسایی.....	۱۵
۳-۴-۲ استفاده از طبقه‌بند بیزین برای شناسایی علامت.....	۱۶
۴-۴-۲ چند رابطه برای آشکارسازی علامت و تقسیم‌بندی تصویر.....	۱۹
۵-۴-۲ استفاده از تطبیق الگو برای شناسایی علامت.....	۲۰
۶-۴-۲ استفاده از فضای رنگ ycbcr به جای فضای رنگ RGB.....	۲۲
فصل سوم: مباحث تئوری.....	۲۴
۱-۳ مدل رنگ HSI.....	۲۵
۲-۳ اجزاء متصل در یک تصویر باینری.....	۲۶
۳-۳ نرمال‌سازی هیستوگرام.....	۲۸

۳-۴ شبکه‌های عصبی.....	۳۱
۳-۴-۱ قاعده آموزش گرادیان نزولی.....	۳۳
۳-۴-۲ شبکه عصبی MLP.....	۳۴
۳-۵ منطق فازی.....	۳۵
۳-۵-۱ سیستم استنتاج فازی.....	۳۷
۳-۵-۲ کاربردهای سیستم فازی.....	۳۸
۳-۶ ماشین بردار پشتیبان.....	۳۸
۳-۷ کتابخانه Open CV.....	۴۲
فصل چهارم: روش پیشنهادی.....	۴۴
۴-۱ مقدمه.....	۴۵
۴-۲ تهیه یک پایگاه داده از علائم راهنمایی و رانندگی کشور ایران.....	۴۵
۴-۳ الگوریتم پیشنهادی برای آشکارسازی علامت راهنمایی و رانندگی.....	۴۷
۴-۳-۱ استفاده از منطق فازی برای آشکارسازی علامت.....	۵۲
۴-۳-۲ آشکارسازی علامت به صورت بلادرنگ با استفاده از C++.....	۵۷
۴-۴ الگوریتم پیشنهادی برای شناسایی نوع دقیق علامت.....	۵۷
۴-۴-۱ استفاده از طبقه‌بند ماشین بردار پشتیبان برای شناسایی دقیق علامت.....	۵۸
۴-۴-۲ نتایج تجربی بدست آمده از شناسایی دقیق با استفاده از ماشین بردار پشتیبان.....	۶۱
۴-۴-۳ استفاده از شبکه عصبی MLP برای شناسایی دقیق.....	۶۴
۴-۴-۴ نتایج تجربی بدست آمده از شناسایی دقیق با استفاده از شبکه عصبی MLP.....	۶۷
۴-۴-۵ شناسایی علامت به صورت بلادرنگ با استفاده از C++.....	۷۱
فصل پنجم: نتیجه‌گیری و پیشنهادات.....	۷۴
۵-۱ نتیجه‌گیری.....	۷۵

۲-۵ پیشنهادات..... ۷۶

مراجع..... ۷۷

فهرست شکل‌ها

- شکل ۱-۱: مجموعه‌ای از علائم راهنمایی و رانندگی که در کشور ایران بکار می‌رود..... ۵
- شکل ۱-۲: تصویر شامل علامت..... ۷
- شکل ۱-۳: مراحل انجام آشکارسازی و شناسایی علامت در این پایان‌نامه..... ۷
- شکل ۱-۲: محدود کردن ناحیه مورد جستجو [۷]..... ۹
- شکل ۲-۲: گوشه یابی از علائم [۷]..... ۱۰
- شکل ۲-۳: نواحی از علائم که خطوط مرور استخراج می‌شوند [۷]..... ۱۰
- شکل ۲-۴: استخراج خطوط مرور از علامت [۷]..... ۱۱
- شکل ۲-۵: دو تصویر که یکی از آن‌ها تماماً آبی است و دیگری تماماً قرمز است..... ۱۳
- شکل ۲-۶: هیستوگرام پارامترهای اشباع و فام در قسمت الف برای تصویر قرمز و قسمت ب برای تصویر آبی [۱۰]..... ۱۳
- شکل ۲-۷: دو ماسک برای استخراج ناحیه داخلی علامت [۱۰]..... ۱۵
- شکل ۲-۸: استخراج مناطق با رنگ قرمز و حذف نویزهای آن توسط فیلتر میانه [۳۰]..... ۲۱
- شکل ۲-۹: محاسبه میزان کجی علامت..... ۲۲
- شکل ۲-۱۰: جهت‌های شعاعی [۳۰]..... ۲۲
- شکل ۲-۱۱: (الف) فضای رنگ RGB و (ب) فضای رنگ YCbCr..... ۲۳
- شکل ۳-۱: مکعب رنگ RGB [۲۰]..... ۲۵
- شکل ۳-۲: مدل رنگ HSI [۲۰]..... ۲۶
- شکل ۳-۳: (الف) چیدمان پیکسل‌ها. (ب) پیکسل‌هایی که ۸ مجاورند. (ج) پیکسل‌هایی که m مجاورند [۲۰]..... ۲۸

- شکل ۳-۴: (الف) تابع صعودی که نشان می‌دهد چگونه چند مقدار به یک مقدار نگاشت می‌شوند. (ب) تابع تکنوای اکیداً صعودی که در هر دو جهت نگاشت یک به یک را موجب می‌شود..... ۳۰
- شکل ۳-۵: مدل سازی یک نرون..... ۳۲
- شکل ۳-۶: ساختار شبکه عصبی پرسپترون سه لایه..... ۳۵
- شکل ۳-۷: (الف) تابع عضویت دوزنقه‌ای و (ب) تابع عضویت زنگی شکل..... ۳۶
- شکل ۳-۸: مرزهای بیشماری که بین دو کلاس وجود دارند..... ۳۹
- شکل ۳-۹: مرزی که بیشترین فاصله را تا نمونه‌های مرزی دارد..... ۴۰
- شکل ۳-۱۰: محاسبه بهترین مرز بین دو کلاس..... ۴۱
- شکل ۴-۱: مراحل انجام آشکارسازی و شناسایی علامت راهنمایی و رانندگی..... ۴۶
- شکل ۴-۲: تعدادی از علائم راهنمایی و رانندگی موجود در پایگاه داده‌ای که اقدام به تهیه آن کردیم..... ۴۷
- شکل ۴-۳: دسته بندی کلی علائم راهنمایی و رانندگی از نظر رنگ و شکل هندسی..... ۴۸
- شکل ۴-۴: (الف) تصویر ورودی در فضای رنگ RGB و (ب) به ترتیب از راست به چپ سه کانال H, S و V این تصویر را در فضای رنگ HSV نشان می‌دهد..... ۴۹
- شکل ۴-۵: (الف) نقاط آبی و قرمز استخراج شده. (ب) پر کردن حفره‌های قسمت الف. (ج) حذف نویزهای موجود در تصویر ب با استفاده از عملیات مورفولوژیکی..... ۵۰
- شکل ۴-۶: الگوهایی که در مرحله آشکارسازی اجزاء متصل موجود در تصویر باینری را با آن‌ها مقایسه می‌کنیم..... ۵۰
- شکل ۴-۷: فرآیند پیدا کردن دایره، مثلث و مستطیل در تصویر باینری..... ۵۱
- شکل ۴-۸: نتایج آشکارسازی علامت..... ۵۲
- شکل ۴-۹: چند تصویر که به خوبی تقسیم‌بندی نشده است..... ۵۲
- شکل ۴-۱۰: تابع عضویت پارامتر فام..... ۵۳

- شکل ۴-۱۱: تابع عضویت پارامتر اشباع..... ۵۳
- شکل ۴-۱۲: استخراج مناطق با رنگ قرمز و آبی با استفاده از منطق فازی و رفع مشکل شکل (۴)- ۵۵..... (۹)
- شکل ۴-۱۳: قسمت الف شامل دو تصویر است که علائم درون آن استخراج شده است. و شکل ب که به علت دوری علامت از دوربین و نور کم، سیستم قادر به استخراج علامت از درون تصویر نیست..... ۵۶
- شکل ۴-۱۴: فرآیند حذف پس‌زمینه علامت آشکار شده..... ۵۷
- شکل ۴-۱۵: نحوه استخراج ویژگی برای تشخیص شکل کلی علامت. قسمت الف از سمت بالا شروع به حرکت کرده‌ایم تا به نوار علامت برسیم و در قسمت‌های ب تا د به ترتیب از چپ، راست و پایین به سمت نوار علامت حرکت شده است..... ۵۸
- شکل ۴-۱۶: قسمت الف نوار اطراف علامت دایره‌ای را نشان می‌دهد. و قسمت ب چهار بردار ویژگی استخراج شده از شکل الف را نشان می‌دهد..... ۵۹
- شکل ۴-۱۷: مراحل به دست آوردن قسمت درونی علامت..... ۶۰
- شکل ۴-۱۸: نحوه استخراج ویژگی برای شناسایی نوع دقیق علامت..... ۶۱
- شکل ۴-۱۹: نمونه‌ای از علائم که سیستم قادر به شناسایی آن‌ها بوده است. ردیف دوم علائمی است که با نویز گوسی آغشته شده‌اند..... ۶۲
- شکل ۴-۲۰: نمونه‌ای از علائم که سیستم قادر به شناسایی آن‌ها نبوده است..... ۶۳
- شکل ۴-۲۱: نا مناسب بودن روش‌های استخراج ویژگی که مبتنی بر تصویر باینری است، در این تصویر دیده می‌شود..... ۶۴
- شکل ۴-۲۲: نمودار تابع فعال سازی به ازاء سه مقدار مختلف برای پارامتر α ۶۷
- شکل ۴-۲۳: نمودار خطا طی ۱۵۰ دوره آموزش..... ۶۸
- شکل ۴-۲۴: ۴۰ کلاس از علائم که ما عملیات شناسایی را بر روی آن‌ها انجام داده‌ایم..... ۶۹

شکل ۴-۲۵: قسمت الف تعدادی از علائم که محتوای آن توسط سیستم شناسایی شده است و قسمت

ب تعدادی علامت که سیستم قادر به شناسایی آنها نبوده است.....۷۰

شکل ۴-۲۶: متعادل سازی هیستوگرام به منظور بهبود کیفیت تصاویر با شدت روشنایی زیاد یا شدت

روشنایی کم.....۷۱

فهرست جداول

- جدول ۱-۱: دسته‌بندی کلی علائم راهنمایی و رانندگی بر اساس شکل هندسی..... ۴
- جدول ۱-۲: خلاصه نتایج روش تقاطع خطوط مرور [۷]..... ۱۱
- جدول ۱-۳: چند تابع کرنل..... ۴۲
- جدول ۱-۴: قانون‌ها و محاسبه میانگین مراکز..... ۵۴
- جدول ۲-۴: مقایسه نتایج آشکارسازی..... ۵۶
- جدول ۳-۴: مقایسه کار انجام شده با سایر روش‌ها..... ۷۲

فصل اول:

مقدمه

۱-۱ مقدمه

علائم راهنمایی و رانندگی با دادن هشدارهایی به راننده موجب افزایش ایمنی حمل و نقل در سطح جاده‌ها و خیابانها می‌شوند. از این رو شناسایی خودکار علائم راهنمایی و رانندگی یکی از اجزاء مهم یک سیستم دستیار راننده به شمار می‌آید و همچنین در نسل آینده وسایل نقلیه می‌تواند به عنوان یک جزء مهم از وسایل نقلیه خودکار محسوب شود. این سیستم باید دارای سرعت و دقت بالایی باشد و در صحنه‌های طبیعی به صورت بلادرنگ علائم را شناسایی کند [۱].

اولین کار در زمینه شناسایی خودکار علائم راهنمایی و رانندگی به اواخر دهه ۱۹۶۰ بر می‌گردد اما پیشرفت‌های چشمگیر در این زمینه در دهه ۱۹۸۰ و ۱۹۹۰ وقتی که ایده‌ی سیستم‌های کمک راننده بر اساس بینایی کامپیوتر مطرح شد، صورت گرفت [۲]. از آن موقع به بعد روش‌های زیادی توسط گروه‌های تحقیقاتی ارائه شده است. از جمله این کارها، کار انجام شده توسط آزمایشگاه شناسایی تصویر دانشگاه KOBLENZ برلین است که یک سیستم شناسایی خودکار برای شناسایی علائم راهنمایی و رانندگی ارائه داد. بیش از ۶۰۰۰۰ تصویر از صحنه‌های ترافیکی برای آموزش این سیستم استفاده شده است و دقت شناسایی این سیستم در حدود ۹۰ درصد با صرف زمان ۲۰۰ میلی ثانیه است. در سال ۱۹۹۵ گروه تحقیقاتی دانشگاه جنوا در ایتالیا با صرف زمان ۵۰۰ میلی ثانیه و با دقت ۹۲ تا ۹۶ درصد علائم را شناسایی کرد [۳].

مشکلاتی در مسیر شناسایی خودکار علائم راهنمایی و رانندگی وجود دارند که مهم‌ترین آن‌ها به شرح زیر است [۴]:

- حرکت ولرزش اتومبیل موجب می‌شود تصاویری که توسط دوربین نصب شده بر روی اتومبیل گرفته می‌شوند دارای مقداری تاری باشند، برای رفع این مشکل نیاز به یک مرحله پیش‌پردازش و رفع تاری است.
- چون اغلب برای آشکارسازی علائم از رنگ آن‌ها استفاده می‌شود بنابراین میزان روشنایی که علائم دارند مهم است زیرا اگر روشنایی علائم از حدی کم تر باشد عمل شناسایی سخت

می شود. میزان روشنایی به عواملی همچون شرایط آب وهوا (بری، برفی و...)، سایه اشیاء اطراف علائم همچون ساختمان ها، در چه زمانی در طول شبانه روز قرار داریم و غیره بستگی دارد.

- کیفیت سیستم تصویربرداری مهم است زیرا اگر کیفیت سیستم تصویر برداری بالا باشد عمل شناسایی راحت تر انجام خواهد شد اما هزینه سیستم بالا می رود.
- خرابی بخشی از علامت، حضور شیء خاصی مانند شاخ و برگ درختان در مقابل علامت یا کم رنگ شدن آن در طول زمان نیز یکی دیگر از مشکلات است.

۱-۲ دسته بندی کلی علائم راهنمایی و رانندگی

علائم راهنمایی و رانندگی در شکل ها و رنگ های مشخصی هستند. مفهوم شکل ها در علائم عبور و مرور در جدول (۱-۱) نشان داده شده است.

رنگ ها در علائم عبور و مرور معنا و مفهوم خاصی دارند به عنوان مثال رنگ قرمز منع کننده یا ایست است. رنگ آبی راهنمای خدمات، حرکات مجاز، علائم اخباری و راهنمای مسیر در آزادراه ها است.

مجموعه ای از علائم راهنمایی و رانندگی که در کشور ایران بکار می رود در شکل (۱-۱) نشان داده شده است.

۱-۳ مراحل شناسایی خودکار علائم راهنمایی و رانندگی

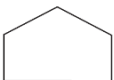
فرآیند شناسایی خودکار علائم راهنمایی و رانندگی شامل دو مرحله کلی است:

الف) آشکارسازی^۱

در مرحله آشکارسازی، باید علامت را از میان اشیاء مختلف موجود در تصویر پیدا کنیم که در این مرحله از دو ویژگی رنگ علائم و شکل هندسی آنها استفاده می شود زیرا علائم راهنمایی و رانندگی نسبت به سایر اشیاء موجود در تصویر دارای رنگ و شکل مشخص هستند.

^۱ Detection

جدول ۱-۱: دسته‌بندی کلی علائم راهنمایی و رانندگی بر اساس شکل هندسی

مفهوم	شکل هندسی
علامت هشدار دهنده برای خطر	
علامت دستوری، بازدارنده	
علائم هشدار دهنده برای مدرسه و دبیرستان	
شروع و پایان حق تقدم مسیر	
معمولاً برای علائم راهنما	
معمولاً برای علائم دستوری و خدمات	
صرفاً برای رعایت حق تقدم	
معمولاً برای تابلوهای راهنما و هشدار برای خطرات ناشی از تعمیر و نگهداری	
صرفاً برای تابلوهای ایست	

ب) شناسایی^۱ دقیق

پس از استخراج علامت از درون تصویر، نوبت به دسته بندی دقیق آن می رسد که در این مرحله ابتدا می بایست پس زمینه علامت را حذف کرد و سپس از آن توسط روش های استخراج ویژگی^۱ استخراج ویژگی کرد و بوسیله طبقه بند^۲ مورد استفاده به شناسایی محتوای علامت پردازیم [۵].

^۱ Recognition



شکل ۱-۱: مجموعه‌ای از علائم راهنمایی و رانندگی که در کشور ایران بکار می‌رود.

در مرحله پیدا کردن علائم از درون تصویر از روش‌هایی همچون آستانه‌گیری رنگی در فضای رنگ HSV و تبدیل هاف^۳ استفاده می‌شود. مزیت آستانه‌گیری رنگی سادگی آن است و عیب آن عدم

^۱ Feature Extraction

^۲ Classifier

^۳ Hough Transform

شناسایی در شرایط متنوع نوری است. مزیت تبدیل هاف شناسایی علامت در حالت اعوجاج و نویزی است و عیب آن نیاز به حافظه زیاد به دلیل محاسبات پیچیده و سرعت پایین آن می‌باشد.

برای مشخص کردن نوع دقیق علامت از روش‌هایی همچون روش انطباق الگو^۱، ماشین بردار پشتیبان^۲ و شبکه‌های عصبی^۳ استفاده می‌شود. مزیت انطباق الگو پیاده‌سازی ساده آن است و عیب آن این است که اگر علامت مورد نظر، اندکی نقص داشته باشد، به خوبی تشخیص داده نخواهد شد. شبکه عصبی دارای این مزیت است که قدرت تشخیص در شرایط مختلف را دارد زیرا می‌توان در شرایط مختلف شبکه را آموزش داد و عیب آن این است که آموزش آن نیازمند تعداد زیادی داده آموزشی است [۶].

زمانی که از سیستم شناسایی خودکار علائم راهنمایی و رانندگی به عنوان یک سیستم کمک راننده استفاده می‌شود، می‌بایست پس از شناسایی علامت به گونه‌ای راننده را مطلع کرد. سیستم‌های متعددی که برای هشدار به راننده ایجاد شده اند عمدتاً بر مبنای هشدار دهنده صوتی، هشدار دهنده تصویری (استفاده از مانیتور) و کنترل کننده خودکار اتومبیل هستند. کنترل خودکار اتومبیل به این معنا که مثلاً با رسیدن اتومبیل به تابلوی کاهش سرعت، سرعت آن به صورت خودکار کاهش یابد.

۴-۱ هدف پایان‌نامه

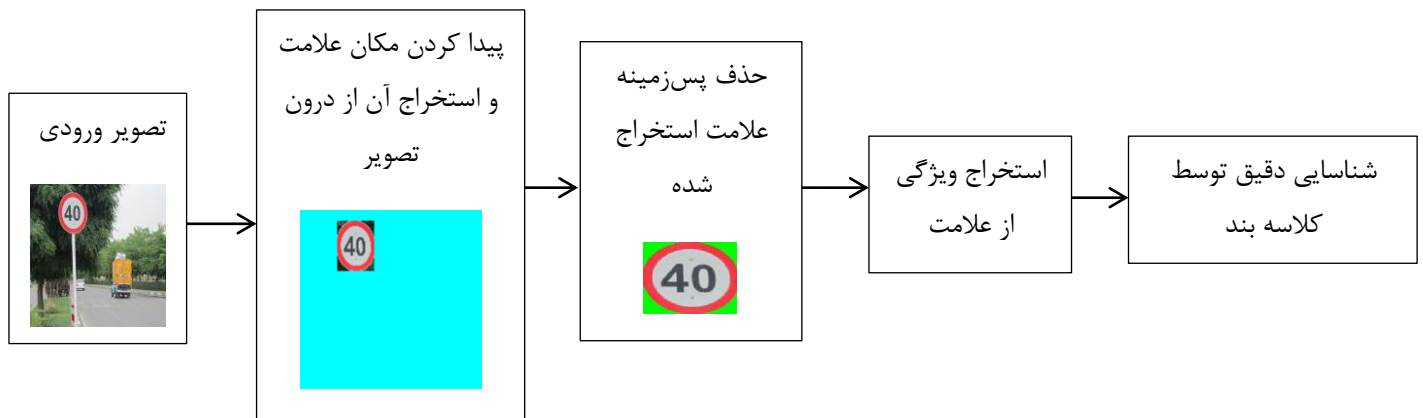
در این پایان‌نامه ابتدا علامت راهنمایی و رانندگی را از میان اشیاء مختلف درون تصویر استخراج می‌کنیم و سپس نوع دقیق علامت آشکار شده را مشخص می‌کنیم. تصاویر مورد استفاده در این پایان‌نامه همانند تصویر شکل (۱-۲)، تصاویری کامل از صحنه‌های واقعی است که شامل علامت راهنمایی است.

کار انجام شده در این پایان‌نامه به طور خلاصه در شکل (۱-۳) نشان داده شده است.

^۱ Template Matching
^۲ Support Vector Machine
^۳ Neural Networks



شکل ۱-۲: تصویر شامل علامت



شکل ۱-۳: مراحل انجام آشکارسازی و شناسایی علامت در این پایان نامه

فصل دوم:

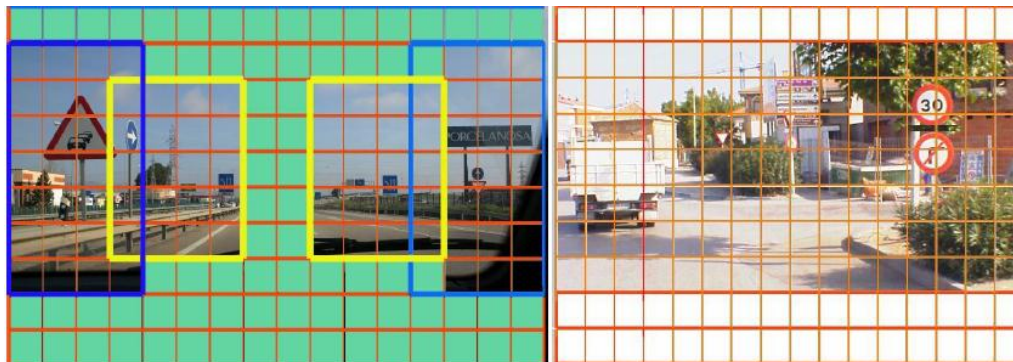
مروری بر کارهای انجام شده

در این فصل تعدادی از کارهای انجام شده در زمینه آشکارسازی و شناسایی علائم راهنمایی و رانندگی بیان می‌شود.

۲-۱ روش تقاطع خطوط مرور

روش تقاطع خطوط مرور نام روشی است که در [۷] برای استخراج ویژگی از علامت راهنمایی و رانندگی برای شناسایی نوع دقیق آن استفاده شده است. این روش در جستجوی خطوطی در تصویر است که به طوری مناسب در تقاطع با شکل هندسی قرار گرفته باشد. این خطوط در علائم مختلف منحصر به فرد هستند.

در این روش برای افزایش سرعت مرحله یافتن علامت، کل تصویر برای یافتن علامت جستجو نشده و فقط قسمت‌هایی که به احتمال زیاد علائم در آن مناطق تصویر قرار دارند جستجو می‌شود. ناحیه مورد جستجو برای مسیر شهری در شکل (۲-۱) قسمت الف و برای مسیر بین شهری در شکل (۲-۱) قسمت ب نشان داده شده است [۷].



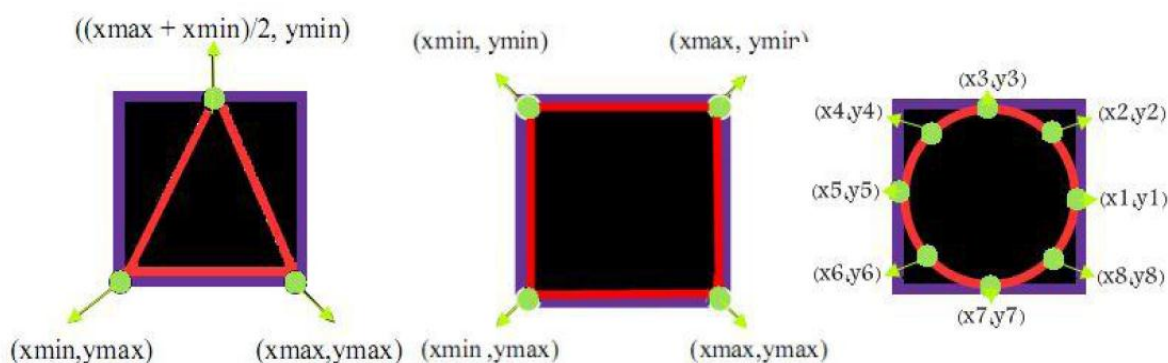
(ب) مسیر بین شهری

(الف) مسیر شهری

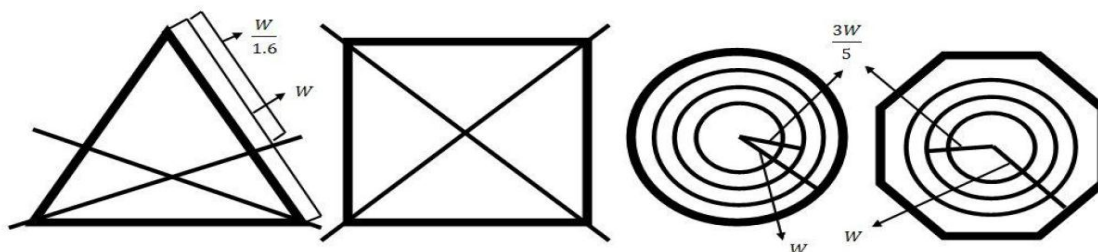
شکل ۲-۱: محدود کردن ناحیه مورد جستجو [۷].

سپس قبل از این که نوع دقیق علامت شناسایی شود بررسی می‌شود که علامت در کدام یک از دسته های مثلث، مربع و دایره قرار می گیرد. این دسته بندی بر مبنای گوشه یابی انجام می شود. به این صورت که با یافتن گوشه هایی خاص از شکل و مقایسه با ماسک هایی که از قبل تعریف شده است، شکل هندسی علامت مشخص می گردد. فرآیند گوشه یابی در شکل (۲-۲) نشان داده شده است.

بعد از یافتن گوشه‌هایی خاص از علامت فاصله اقلیدسی^۱ آن‌ها با گوشه‌های ماسک محاسبه می‌شود، کوچکترین مجموع فواصل اقلیدسی در ماسک‌ها نشان دهنده بیش‌ترین شباهت است. پس از این که نوع شکل (مثلث، دایره، مربع) مشخص شد. سپس باید نوع دقیق آن تشخیص داده شود. بدین منظور ابتدا تصویر باینری می‌شود. ایده اصلی این روش این است که تنها خطوطی از تصویر علائم که دارای اطلاعات مهمی هستند در نظر گرفته شود. نواحی از تصویر که این خطوط برای هر یک از علائم مستطیل، مثلث، دایره و هشت‌گوش استخراج می‌شود در شکل (۲-۳) نشان داده شده است. این خطوط، خطوط مرور نامگذاری می‌شوند. نحوه استخراج این خطوط در شکل (۲-۴) نشان داده شده است. پس از استخراج این خطوط از علامت، با خطوط ذخیره شده در بانک اطلاعاتی مقایسه می‌گردد. این خطوط قبلاً برای همه علائم به صورت نرمال شده در بانک اطلاعاتی ذخیره شده‌اند. فرآیند مقایسه به این صورت انجام می‌شود که خطوط استخراج شده از علامت با تک‌تک خطوط موجود در بانک اطلاعاتی اند بیتی می‌شود و کم‌ترین اختلاف نشان دهنده بیش‌ترین شباهت است.

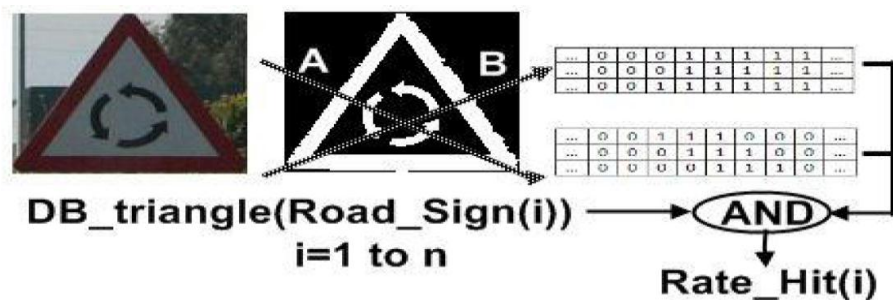


شکل ۲-۲: گوشه یابی از علائم [۷].



شکل ۲-۳: نواحی از علائم که خطوط مرور استخراج می‌شوند [۷].

^۱ Euclidean Distance



شکل ۲-۴: استخراج خطوط مرور از علامت. n تعداد علائم و $Rate_Hit(i)$ درصد تطبیق علامت با علامت مثلث i ام

موجود در بانک اطلاعاتی است [۷].

نتایج عملی این روش به طور خلاصه در جدول (۲-۱) زیر مشاهده میشود:

جدول ۲-۱: خلاصه نتایج روش تقاطع خطوط مرور [۷].

نوع علائم	تعداد شناسایی درست علائم	تعداد شناسایی نادرست علائم	میزان دقت
مثلث	۷۸	۷	٪۹۱
مستطیل	۵۴	۸	٪۸۷
دایره	۱۰۳	۲۸	٪۷۸
هشت گوش	۱۸	۷	٪۷۲
مجموع	۲۵۳	۵۰	٪۸۳

۲-۲ روشی برای استخراج ویژگی از علامت

در مرجع [۸] با فرض این که علامت در فضای رنگ^۱ RGB است و دارای ۳۰ ستون و ۳۰ سطر است،

۶۳ ویژگی به صورت زیر استخراج می شود:

سه ویژگی از روی سه کانال رنگ استخراج می شود، که عبارتند از MR, MG, MB. این سه ویژگی به

ترتیب برای کانال اول، دوم و سوم به صورت زیر محاسبه می شوند:

^۱ Color Space

$$MR = \frac{1}{256} \left(\frac{1}{900} \sum_{i=1}^{30} \sum_{j=1}^{30} r_{ij} \right) \quad (1-2)$$

$$MG = \frac{1}{256} \left(\frac{1}{900} \sum_{i=1}^{30} \sum_{j=1}^{30} g_{ij} \right) \quad (2-2)$$

$$MB = \frac{1}{256} \left(\frac{1}{900} \sum_{i=1}^{30} \sum_{j=1}^{30} b_{ij} \right) \quad (3-2)$$

که r_{ij} پیکسل سطر i ام و ستون j ام کانال قرمز و g_{ij} پیکسل سطر i ام و ستون j ام کانال سبز و b_{ij} پیکسل سطر i ام و ستون j ام کانال آبی است.

و برای به دست آوردن ۶۰ ویژگی دیگر، ابتدا تصویر را به یک تصویر خاکستری تبدیل کرده و سپس ۳۰ ویژگی عمودی و ۳۰ ویژگی افقی به صورت زیر تعریف می‌شود:

$$vh_i = \frac{1}{30} \sum_{j=1}^{30} (b'_{ij} > T) \quad , i = 1, 2, \dots, 30 \quad (4-2)$$

$$vh_j = \frac{1}{30} \sum_{i=1}^{30} (b'_{ij} > T) \quad , j = 1, 2, \dots, 30 \quad (5-2)$$

T مقدار آستانه است و به صورت زیر محاسبه می‌شود:

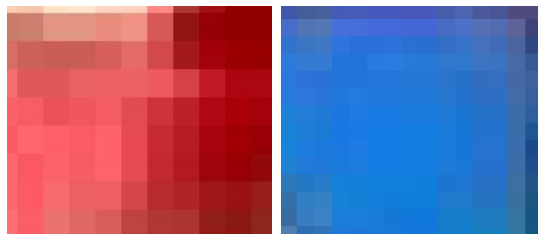
$$T = \frac{1}{900} \sum_{i=1}^{30} \sum_{j=1}^{30} b'_{ij} \quad (6-2)$$

۳-۲ تقسیم بندی تصویر

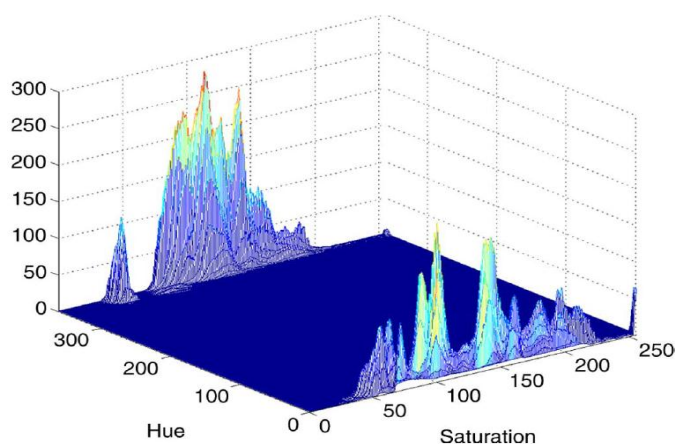
فضاهای رنگ متفاوتی برای تقسیم‌بندی تصاویری که در فضای آزاد قرار دارند، وجود دارد. مهم‌ترین مشکلاتی که در تقسیم‌بندی وجود دارد، تغییرات روشنایی و اعوجاج‌های درون تصویر است. اما مدل رنگ HSI در میان فضاهای رنگ مختلف به دلیل این که اطلاعات رنگ در پارامترهای اشباع^۱ و فام^۲ قرار دارد و از طرفی این پارامترها از پارامتر روشنایی مستقل هستند، از میان فضاهای رنگ مختلف بهترین نتیجه را در تقسیم بندی اولیه تصویر می‌دهد [۹]. به منظور بدست آوردن آستانه‌های مناسب اشباع و فام برای رنگ قرمز و آبی، دو تصویر که یکی از آن‌ها تماماً قرمز و دیگری تماماً آبی است را در نظر گرفته و هیستوگرام پارامترهای اشباع و فام برای هر یک از این تصاویر رسم می‌شود [۱۰]. دو

^۱ Saturation
^۲ Hue

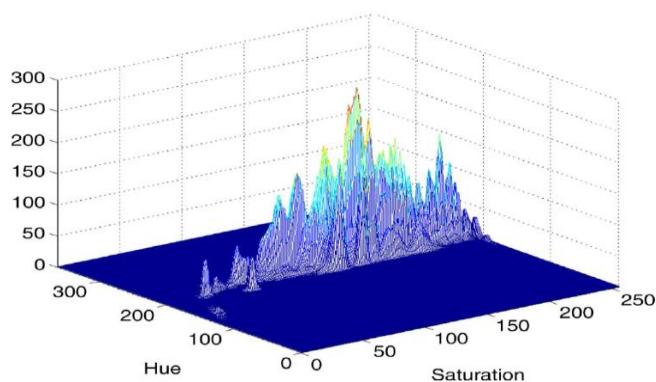
تصویر تماماً آبی و قرمز در شکل (۵-۲) نشان داده شده است. هیستوگرام‌ها در شکل (۶-۲) در قسمت الف برای تصویر قرمز و در قسمت ب برای تصویر آبی نشان داده شده است.



شکل ۵-۲: دو تصویر که یکی از آنها تماماً آبی است و دیگری تماماً قرمز است.



(الف)



(ب)

شکل ۶-۲: هیستوگرام پارامترهای اشباع و فام در قسمت الف برای تصویر قرمز و قسمت ب برای تصویر آبی [۱۰].

متاسفانه مولفه‌های فام و اشباع دارای اطلاعات کافی برای جداسازی قسمت‌های سفید تصویر نیست. بنابراین طبق رابطه‌ای که در [۱۱] آمده است، برای استخراج این مناطق از تصویر از رابطه زیر استفاده شده است.

$$f(R, G, B) = \frac{(|R-G|+|G-B|+|B-R|)}{3D} \quad (7-2)$$

که R,G,B به ترتیب مقادیر کانال‌های رنگ در تصویر RGB است. و D درجه استخراج است که بهترین نتیجه با در نظر گرفتن این پارامتر مساوی ۲۰ بدست می‌آید. اگر $f(R,G,B)$ بزرگتر از یک باشد نشان دهنده این است که پیکسل موردنظر رنگی است و اگر کوچکتر از یک باشد نشان‌دهنده این است که پیکسل موردنظر بدون رنگ یا سفید است. بعد از مرحله تقسیم‌بندی تصویر، پیکسل‌های تصویر بسته به این که متعلق به یکی از رنگ‌های موردنظر باشند، به اجزاء متصلی^۱ تقسیم می‌شوند. تقسیم‌بندی تصاویری که در شب گرفته می‌شوند مانند تقسیم‌بندی تصاویری است که در روز گرفته می‌شوند، با این تفاوت که برای استخراج مناطق با رنگ سفید نمی‌توان از رابطه (۷-۲) استفاده کرد. به دلیل این که در شب نور چراغ جلو وسیله نقلیه بر روی تابلو افتاده، از این رو علامت با رنگ سفید شبیه به رنگ زرد می‌شود. اما تفاوت این رنگ با رنگ زرد در مولفه اشباع آن‌ها خواهد بود و مولفه‌های فام آن‌ها مانند هم است.

۲-۴ استخراج و شناسایی علامت از درون تصویر

بعد از این که مناطقی از تصویر با رنگ مشخص استخراج شد، آن‌گاه نوبت به این می‌رسد که قسمت‌هایی از تصویر که علامت نبوده‌اند اما استخراج شده‌اند را حذف کرد، حال می‌بایست از ویژگی‌هایی نظیر ویژگی شکل هندسی و نسبت طول به عرض استفاده کرد و از میان اجزاء متصل مختلف، علامت را استخراج کرد. یا می‌توان اجزاء متصل بزرگ و اجزاء متصل کوچک تصویر را به عنوان نویز در نظر گرفت و آن‌ها را حذف کرد [۱۰].

^۱ Connected Component

به دلیل این که علائم راهنمایی و رانندگی چندضلعی‌های منظم و یا دایره‌ای شکل هستند، بنابراین نسبت مساحت مستطیلی که علامت را احاطه کرده است به مساحت علامت بزرگتر از ۱ است.

۲-۴-۱ استفاده از SVM^۱ برای شناسایی علامت

در مرجع [۱۰] شناسایی با استفاده از SVM انجام می‌گردد. در اکثر موارد، داده توسط یک تابع خطی جداپذیر نخواهد بود. یک راه حل این است که داده ورودی به فضای جدید نگاشت داده شود. کرنلی که توسط آن نگاشت انجام می‌گردد، کرنل گوسی است، این کرنل به صورت زیر تعریف می‌شود:

$$K(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \quad (۸-۲)$$

تابع تصمیم‌گیری برای یک بردار ورودی جدید به صورت زیر است:

$$f(x) = \text{sgn}(\sum_{i=1}^{N_s} \alpha_i y_i K(s_i, x) + b) \quad (۹-۲)$$

N_s تعداد بردارهای پشتیبان^۲ و s_i ها بردارهای پشتیبان هستند. در مرحله شناسایی، ناحیه داخلی علامت استخراج شده و به یک بلوک تصویر خاکستری ۳۱×۳۱ تبدیل می‌شود. شکل (۷-۲) نشان‌دهنده دو ماسک مستطیلی و دایره‌ای برای استخراج ناحیه داخلی علامت مورد نظر است.



شکل ۷-۲: دو ماسک برای استخراج ناحیه داخلی علامت [۱۰].

تعداد داده‌های آموزشی برای هر کلاس بین ۲۰ تا ۱۰۰ تغییر می‌کند، که به طور متوسط برای هر کلاس از ۵۰ داده آموزشی استفاده می‌شود. وقتی که ویژگی‌های موردنظر استخراج شده و بردارهای ویژگی ساخته می‌شوند، برخی از این بردارها، بردار پشتیبان هستند.

۲-۴-۲ چند روش برای استخراج ویژگی و شناسایی

^۱ Support Vector Machine

^۲ Support Vectors

مرجع [۸] مرحله آشکارسازی علائم را توسط تکنیک‌های پردازش تصویر مانند: آستانه‌گذاری رنگ، فیلتر گوسی و آشکارساز لبه canny انجام داده است و دسته بندی دقیق علامت را توسط شبکه عصبی انجام داده است. در مرجع [۱۲] آشکارسازی علامت بر اساس رنگ آن با استفاده از طبقه‌بند بیزین انجام شده است. سپس تبدیل SIFT^۱ برای استخراج یک مجموعه از ویژگی‌های ثابت به منظور شناسایی علامت به کار می‌رود. شناسایی با مقایسه ویژگی استخراج شده با ویژگی‌های ذخیره شده از علائم استاندارد انجام می‌شود.

مرجع [۱۳] آشکارسازی علائم راهنمایی و رانندگی را به این صورت انجام داده است که ابتدا تصویر ورودی به فضای رنگ خاکستری^۲ برده می‌شود، سپس لبه‌های تصویر با استفاده از آشکارساز لبه canny به دست می‌آیند. علت استفاده از این روش آشکارسازی لبه، دقت بالا نسبت به سایر روش‌های آشکارسازی لبه است. سپس با استفاده از عملیات مورفولوژیکی نویزهای موجود در تصویر حذف می‌گردد و حفره‌های موجود در تصویر پر می‌شوند. در مرجع [۱۴] یک الگوریتم آشکارسازی علائم دایره‌ای با استفاده از پارامتر روشنایی تصویر ورودی ارائه شده است.

در استفاده از شبکه عصبی برای شناسایی علائم می‌بایست از هر تصویر یک بردار ویژگی استخراج کنیم و به عنوان ویژگی ورودی به شبکه، از آن برای آموزش شبکه استفاده کنیم.

یک روش استخراج ویژگی استفاده از DCT^۳ یا تبدیل کسینوسی گسسته است. به این منظور ابتدا تصویر علامت که یک سیگنال دو بعدی است، تبدیل به یک سیگنال یک بعدی می‌شود.

یک بعدی کردن تصویر به این صورت انجام می‌شود که سطرهای تصویر را به صورت رشته ای کنار هم قرار داده و سپس یک DCT یک بعدی روی این رشته اعمال می‌شود و از صد ضریب اول DCT به عنوان ویژگی های ورودی استفاده می‌شود [۱۵].

۲-۴-۳ استفاده از طبقه‌بند بیزین برای شناسایی علامت

^۱ Scale Invariant Feature Transform

^۲ Gray

^۳ Discrete Cosine Transform

در مرجع [۱۶] شناسایی خودکار علائم راهنمایی و رانندگی با استفاده از طبقه‌بند بیزین انجام شده است. برای شناسایی علائم راهنمایی و رانندگی ابتدا یک بردار ویژگی مستقل از انتقال و مقیاس از تصاویر علائم استخراج گشته و سپس با استفاده از این بردارها طبقه‌بندهای بیزی آموزش داده می‌شوند. در نهایت این طبقه‌بندها برای طبقه‌بندی و شناسایی علائم به کار می‌روند. برای حصول دقت بیش‌تر شناسایی در دو مرحله انجام می‌شود یعنی ابتدا قالب کلی علامت تشخیص داده می‌شود و در مرحله بعد خود علامت شناسایی می‌گردد. آزمایش‌های انجام شده نشان می‌دهد که دقت روش پیشنهادی برای تصاویر بدون نویز ۹۶ درصد و برای تصاویر با نویز سفید با میانگین صفر و واریانس ۱۵ درصد حدود ۸۰ درصد است.

بردار ویژگی در این روش، از تصویر باینری به این صورت به دست می‌آید که در مرحله‌ی اول مختصات پیکسل‌های سیاه تصویر به صورت یک توزیع دوبعدی از داده‌ها در نظر گرفته می‌شوند و سپس ماتریس کواریانس آن محاسبه می‌گردد. در مرحله بعد دو مقدار ویژه و دو بردار ویژه ماتریس محاسبه می‌گردد. سپس تصویر بر اساس دو بردار ویژه محاسبه شده به چهار ناحیه تقسیم می‌شود. که ابتدای بردارهای ویژه در مرکز جرم تصویر در نظر گرفته می‌شود. حال برای هر چهار ناحیه به طور جداگانه ماتریس کواریانس، مقادیر و بردارهای ویژه محاسبه می‌شود. و چهار پارامتر مستقل از انتقال، دوران و مقیاس از آن‌ها استخراج می‌گردد. این پارامترها به شرح زیر تعریف می‌شوند:

نسبت ویژه: نسبت مقدار ویژه کوچک‌تر ناحیه به مقدار ویژه بزرگ‌تر ناحیه.

زاویه نرمال: زاویه تند بین بردار ویژه بزرگ‌تر ناحیه با بردار ویژه بزرگ‌تر شکل اصلی.

مرکز: مجذور فاصله بین مرکز جرم ناحیه با مرکز جرم شکل اصلی، تقسیم بر مقدار ویژه‌ی بزرگ‌تر شکل اصلی.

فشرده‌گی: نسبت مساحت ناحیه به مساحت مستطیل احاطه کننده ناحیه (مستطیل احاطه کننده برای هر ناحیه، کوچکترین مستطیلی است که تمام شکل را در بر می‌گیرد به طوری که اضلاع آن موازی با بردارهای ویژه شکل باشند).

برای شکل اصلی که چهار ناحیه را در بر می‌گیرد فقط پارامترهای نسبت ویژه و فشرده‌گی محاسبه می‌شوند. زیرا دو پارامتر دیگر صفر می‌باشند. برای چهار ناحیه کوچکتر هر چهار پارامتر محاسبه می‌شوند. بنابراین در کل هجده ویژگی از شکل به عنوان بردار ویژگی استخراج می‌شود.

در استفاده از طبقه‌بند بیزی فرض شده است که N کلاس C_i از علائم راهنمایی و رانندگی وجود دارند. در ساده‌ترین حالت C_i ها می‌توانند، هر یک از علائم راهنمایی و رانندگی باشند. هم‌چنین فرض بر این است که s علامت راهنمایی و رانندگی ورودی است. طبقه‌بند بیزی، s را به کلاس c منتسب می‌کند. به طوری که:

$$C = \arg_i \max P(C_i | S) \quad (10-2)$$

طبق رابطه بیز، معادله بالا برابر با رابطه زیر است:

$$C = \arg_i \max \frac{P(S|C_i)P(C_i)}{P(S)} = \arg_i \max P(S|C_i)P(C_i) \quad (11-2)$$

در رابطه بالا $p(c_i)$ احتمال پیشین کلاس c_i و $p(s|c_i)$ درجه شباهت s با کلاس c_i است. برای محاسبه درجه شباهت، فرض شده است که توزیع داده‌ها در کلاس‌های c_i از چگالی گوسی تبعیت می‌کند. با در نظر گرفتن این نکته که تصویر هریک از علائم راهنمایی و رانندگی با یک بردار ویژگی هجده مولفه‌ای نشان داده می‌شود توزیع موردنظر یک توزیع هجده بعدی است. اگر μ_i بردار میانگین و Σ_i ماتریس کواریانس مربوط به توزیع گوسی کلاس c_i باشد، درجه شباهت علامت s با کلاس c_i از رابطه (12-2) به دست می‌آید [16]:

$$P(S|C_i) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_i|}} e^{-(x_s - \mu_i)^T \Sigma_i^{-1} (x_s - \mu_i)/2} \quad (12-2)$$

که XS بردار ویژگی مربوط به علامت ورودی s و n برابر بعد بردار ویژگی یعنی هجده است. برای تخمین پارامترهای توزیع‌های گوسی باید تصاویر مختلفی از هر یک از علائم راهنمایی و رانندگی در اختیار باشد. دادگان مورد استفاده در این کار شامل ۱۱۳ تصویر استاندارد و بدون نویز از علائم راهنمایی و رانندگی است [۱۶].

۴-۴-۲ چند رابطه برای آشکارسازی علامت و تقسیم بندی تصویر

مرجع [۱۳] برای استخراج علامت دایره‌ای از میان اشیاء مختلفی که در تصویر باینری وجود دارد، پارامتر metric را به صورت زیر تعریف می‌کند.

$$Metric = 4\pi * Area/Perimeter^2 \quad (۱۳-۲)$$

که Area مساحت شی و Perimeter محیط شی است. در صورتی که برای یک شی $0.9 \leq Metric \leq 1$ باشد، این شی به عنوان علامت راهنمایی و رانندگی نگه داشته می‌شود. سپس از علامت آشکار شده، استخراج ویژگی می‌شود و با استفاده از ماشین بردار پشتیبان شناسایی می‌گردد.

در مرجع [۱۵] برای استخراج مناطق با رنگ قرمز به این صورت عمل می‌شود که اگر R_i, G_i و B_i کانال‌های رنگ تصویر ورودی باشند، آنگاه کانال‌های رنگ تصویر خروجی به صورت زیر خواهد بود:

$$R_o = \max(0, (R_i - G_i) + (R_i - B_i)), G_o = B_o = 0 \quad (۱۴-۲)$$

در مرجع [۱۷] برای استخراج ناحیه قرمز از رابطه زیر استفاده شده است.

$$REDNESS = R - \max(G, B) - a \times abs(G - B) \quad (۱۵-۲)$$

در مرجع [۱۸] رابطه‌ای که برای استخراج مناطق با رنگ قرمز ارائه شده است به صورت زیر است:

$$f_R(x) = \max(0, \frac{\min(x_R - x_G, x_R - x_B)}{s}) \quad (۱۶-۲)$$

که:

$$S = x_R + x_G + x_B \quad (۱۷-۲)$$

مرجع [۱۸] برای پیدا کردن علامت در یک تصویر باینری از یک ماسک لغزان استفاده کرده است. به این صورت که اگر (x,y) مختصات گوشه سمت چپ بالا ماسک باشد، و w و h به ترتیب عرض و ارتفاع این ماسک باشد آن گاه می توان با تعریف نسبت زیر و مقایسه آن با یک مقدار آستانه علامت را پیدا کرد:

$$\frac{\text{sum}(\text{img}(x,y,w,h))}{\text{area}(\text{img}(x,y,w,h))} > \Theta \quad (18-2)$$

Bounding Box به کوچکترین مستطیلی گویند که یک شی را احاطه کند. می توان در مرحله آشکارسازی از این مفهوم برای یافتن علامت استفاده کرد. برای ایجاد این مستطیل کافیست مختصات گوشه های جزء متصل در یک تصویر باینری محاسبه شود [۴].

پس از این که Bounding Box محاسبه شد از ویژگی های Bounding Box مانند طول، عرض و نسبت طول به عرض استفاده می شود و علامت استخراج می شود.

پس از این که علامت استخراج شد برای شناسایی دقیق آن ضریب همبستگی علامت آشکار شده با تک تک علائم موجود در بانک اطلاعاتی محاسبه می شود. بیشترین همبستگی نشان دهنده بیشترین شباهت علامت، با نمونه موجود در بانک اطلاعاتی است. رابطه ای که توسط آن ضریب همبستگی محاسبه می شود به صورت زیر است [۴]:

$$c(i,j) = \frac{\sum_m \sum_n (A_{mn} - A^-)(B_{mn} - B^-)}{\sqrt{\sum_m \sum_n (A_{mn} - A^-)^2 + (B_{mn} - B^-)^2}} \quad (19-2)$$

A ماتریس تصویر و B ماتریس الگو است. و A_{mn} میانگین A است.

۲-۴-۵ استفاده از تطبیق الگو برای شناسایی علامت

مرجع [۳۰] ابتدا تصویر را از فضای رنگ RGB به فضای رنگ HSV تبدیل می کند، سپس با استفاده از آستانه گذاری بر روی پارامتر فام، مناطق با رنگ قرمز را توسط رابطه (۲۰-۲) استخراج می کند.

$$L(x,y) = \begin{cases} 1 & 0.94 \leq H \leq 1 \text{ or } 0 \leq H \leq 0.1 \\ 0 & \text{otherwise} \end{cases} \quad (20-2)$$

سپس برای از بین بردن نویزهای موجود در تصویر به دست آمده، از یک فیلتر میانه^۱ با ابعاد ۵×۵ استفاده می‌شود. که طی این عملیات از تصویر شکل (۸-۲) قسمت الف تصویر شکل (۸-۲) قسمت ب به دست خواهد آمد.



(ب)

(الف)

شکل ۸-۲: استخراج مناطق با رنگ قرمز و حذف نویزهای آن توسط فیلتر میانه [۳۰].

سپس برای حذف اجزاء متصل موجود در تصویر که علامت راهنمایی و رانندگی نیستند، از ویژگی نسبت طول به عرض مستطیل محیطی^۲ جزء متصل استفاده می‌شود. این فیلتر در رابطه (۲۱-۲) نشان داده شده است.

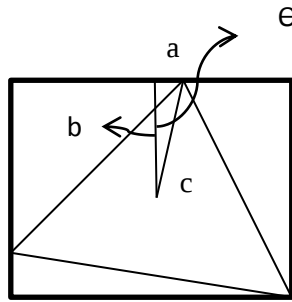
$$r_i = \begin{cases} r_i & \frac{3}{4} \leq \frac{\text{width}(b_i)}{\text{height}(b_i)} \leq \frac{4}{3} \\ 0 & \text{otherwise} \end{cases} \quad (21-2)$$

پس از آشکارسازی علامت میزان کجی آن توسط رابطه زیر محاسبه می‌شود:

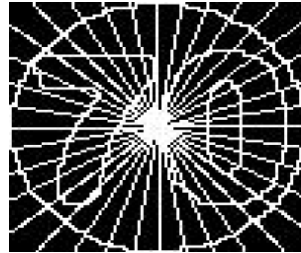
$$\theta = \tan^{-1} \left(\frac{a}{b} \right) \quad (22-2)$$

که پارامترهای θ ، a و b در شکل (۹-۲) نشان داده شده است. برای شناسایی نوع دقیق علامت از روش انطباق الگو استفاده می‌شود. به این صورت که ابتدا لبه‌های تصویر محاسبه می‌شود. سپس ۳۶ جهت شعاعی برای هر شکل در نظر گرفته می‌شود. جهت‌های شعاعی برای یک علامت در شکل (۲-۱۰) نشان داده شده است.

^۱ Median filter
^۲ Bounding Box



شکل ۲-۹: محاسبه میزان کجی علامت



شکل ۲-۱۰: جهت‌های شعاعی [۳۰].

سپس از روی جهت‌های شعاعی، برای هر علامت ۳۶ ویژگی به صورت زیر استخراج می‌شود:

$$Cr_{avg,i} = \frac{(d_1 + d_2 + \dots + d_m)}{3} \quad \text{where } i = 1, 2, \dots, 36 \quad (22-2)$$

که $\{d_1, d_2, \dots, d_m\}$ فاصله نقاط تقاطع تا مرکز هر خط است. پس از استخراج ویژگی از یک علامت، با جدول انطباق ویژگی که از قبل آماده شده است، مقایسه می‌گردد. مزیت این روش این است که به دلیل این که علائم راهنمایی و رانندگی در کشورهای مختلف، متفاوت است، به راحتی می‌توان جدول انطباق ویژگی را تغییر داد.

۲-۴-۶ استفاده از فضای رنگ YCbCr به جای فضای رنگ RGB

مرجع [۳۳] قبل از تقسیم بندی تصویر، ابتدا تصویر را از فضای رنگ RGB به فضای رنگ YCbCr تبدیل می‌کند. فضای رنگ RGB طبق رابطه (۲۳-۲) به فضای رنگ YCbCr تبدیل می‌شود.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.578 & 0.114 \\ -0.169 & -0.331 & 0.5 \\ 0.5 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (23-2)$$

در فضای رنگ YCbCr اطلاعات شدت روشنایی با مؤلفه Y نشان داده می‌شود. و اطلاعات رنگ به صورت دو مؤلفه Cb و Cr نمایش داده می‌شود. مؤلفه Cb تفاوت بین مؤلفه آبی و مقدار مرجع

است. و مؤلفه Cr تفاوت بین مؤلفه قرمز و مقدار مرجع است. شکل (۲-۱۱) قسمت الف تصویر در فضای رنگ RGB و قسمت ب فضای رنگ YCbCr این تصویر را نشان می‌دهد.



(ب)

(الف)

شکل ۲-۱۱: (الف) فضای رنگ RGB و (ب) فضای رنگ YCbCr.

سپس برای شناسایی علامت از میان اشیاء استخراج شده، از معیار خطای میانگین مربع^۱ (MSE) استفاده می‌شود. رابطه MSE به صورت زیر است.

$$MSE = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N [T(x, y) - I(x, y)]^2 \quad (2-24)$$

T تصویر الگو و I تصویر ورودی است. M و N ابعاد تصویر هستند. این معیار انطباق الگو بر اساس حداکثر شباهت تصویر ورودی با الگو است.

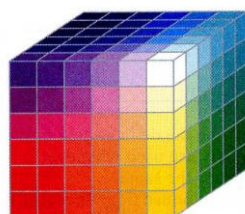
^۱ Mean square error

فصل سوم:

مباحث تئوری

۳-۱ مدل رنگ HSI

مدل رنگ HSI دارای سه کانال فام، اشباع و شدت روشنایی^۱ است. این مدل رنگ رابطه نزدیکی با نحوه دریافت رنگ توسط انسان دارد. و به علت جدا بودن پارامترهای فام و اشباع از پارامتر روشنایی، برای استخراج مناطقی از تصویر با رنگ مشخص روشنایی تأثیری کمی خواهد داشت. فضای رنگ HSI با در نظر گرفتن مکعب رنگ RGB (شکل (۳-۱)) در امتداد محور خاکستری آن (محوری که رئوس سیاه و سفید را به هم متصل می‌کند) فرمول‌بندی می‌شود، که منجر به تولید فضای رنگ HSI یا فضای رنگ هرمی شکل یا مخروطی شکل نشان داده شده در شکل (۳-۲) می‌شود. وقتی در این شکل در امتداد محور عمودی (خاکستری) پیش می‌رویم، اندازه صفحه دایره‌ای که عمود بر محور است تغییر می‌کند [۱۹].



شکل ۳-۱: مکعب رنگ RGB [۲۰].

پارامتر فام به صورت زاویه‌ای حول مقطع دایره‌ای رنگ، معمولاً با استفاده از محور قرمز به عنوان محور مرجع (صفر درجه) بیان می‌شود. مولفه‌ی مقدار روشنایی در امتداد محور مخروط اندازه‌گیری می‌شود. که در انتهای محور سیاه مقدارش صفر است و در انتهای محور سفید مقدارش یک است. تصویر گرفته شده توسط اکثر دوربین‌ها در فضای رنگ RGB است. بنابراین برای تبدیل فضای رنگ RGB به فضای رنگ HSI از روابط زیر استفاده می‌کنیم [۲۰].

مولفه‌ی فام به صورت زیر به دست می‌آید:

$$H = \begin{cases} \theta & B \leq G \\ 360 - \theta & B > G \end{cases} \quad (۳-۱)$$

که:

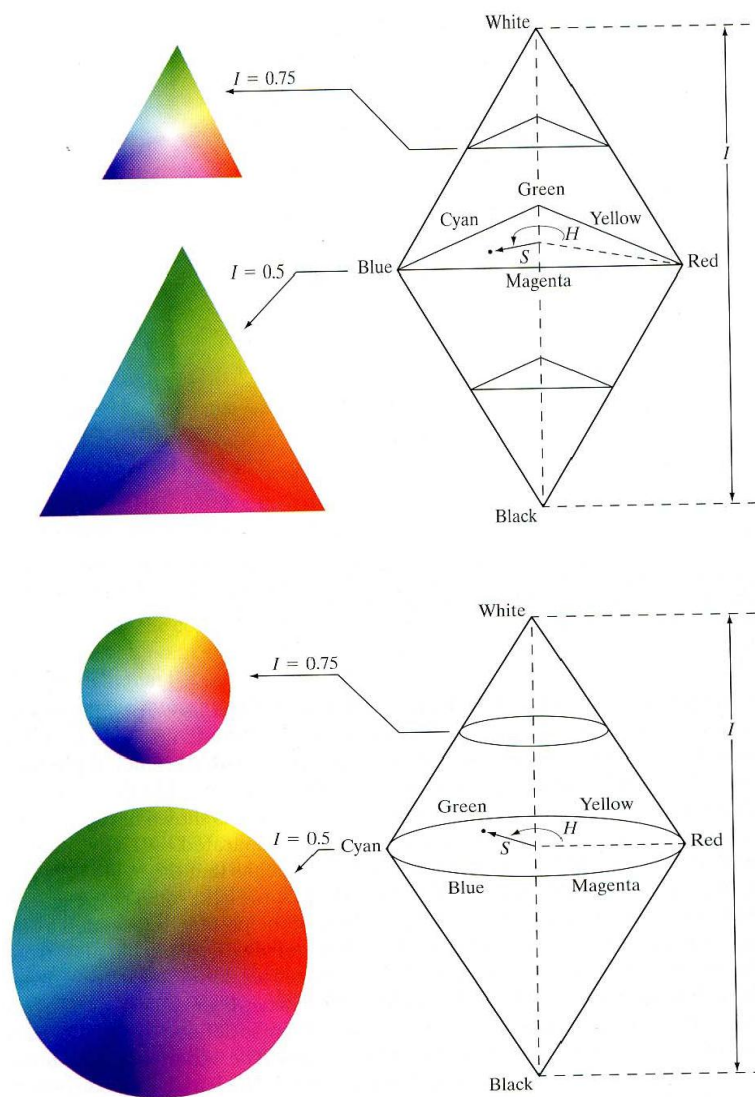
^۱ Intensity

$$\theta = \cos^{-1} \left\{ \frac{0.5[(R-G)+(R-B)]}{[(R-G)^2 + (R-G)(G-B)]0.5} \right\} \quad (2-3)$$

و مولفه‌ی اشباع و شدت روشنایی به ترتیب طبق رابطه‌ی (۳-۳) و (۴-۳) بدست می‌آیند.

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)] \quad (3-3)$$

$$I = \frac{1}{3}(R + G + B) \quad (4-3)$$



شکل ۲-۳: مدل رنگ HSI [۲۰].

۲-۳ اجزاء متصل در یک تصویر باینری

یک پیکسل P با مختصات (x,y) چهار همسایه افقی و عمودی دارد که مختصاتشان به صورت زیر

است:

$$(x+1,y), (x-1,y), (x,y+1), (x,y-1)$$

این مجموعه از پیکسل‌ها، ۴ همسایگان P نامیده می‌شوند، که به صورت $N_4(p)$ نمایش داده می‌شود. هر پیکسل یک واحد از (x,y) فاصله دارد. و مختصات چهار همسایه قطری p به شرح زیر است:

$$(x+1,y+1), (x+1,y-1), (x-1,y+1), (x-1,y-1)$$

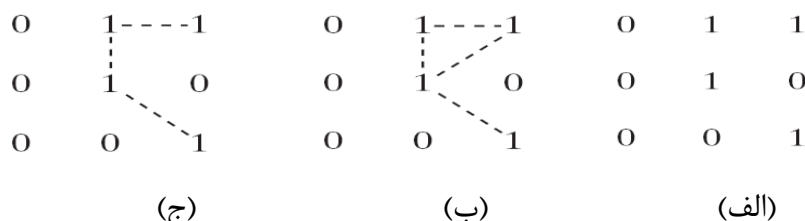
که با $N_D(p)$ نمایش داده می‌شوند. این نقاط به همراه ۴-همسایه، ۸-همسایه P نامیده شده و با $N_8(p)$ نشان داده می‌شوند. با فرض این که V مجموعه‌ای از مقادیر شدت روشنایی باشد که برای تعریف مجاورت استفاده می‌شود، در یک تصویر دودویی، اگر به مجاورتی از پیکسل‌ها با مقدار ۱ اشاره کنیم، آن‌گاه $V=\{1\}$ است. اما مجموعه V ممکن است که مقادیر بیشتری را شامل شود. برای مثال در مجاورتی از پیکسل‌ها با یک محدوده، مقادیر شدت روشنایی ممکن است بین صفر تا ۲۵۵ باشند. مجموعه V می‌تواند هر زیر مجموعه‌ای از این ۲۵۶ مقدار باشد. ما سه نوع مجاورت را در نظر می‌گیریم [۱۹]:

۱- دو پیکسل p و q با مقادیری از V ، ۴-مجاوریتی هستند اگر q در مجموعه‌ی $N_4(p)$ باشد.

۲- دو پیکسل p و q با مقادیری از V ، ۸-مجاوریتی هستند اگر q در مجموعه‌ی $N_8(p)$ باشد.

۳- دو پیکسل p و q با مقادیری از V ، m -مجاوریتی هستند اگر q در مجموعه‌ی $N_4(p)$ باشد. یا q در $N_D(p)$ باشد و مجموعه $N_4(q) \cap N_4(p)$ هیچ پیکسلی با مقادیری از V نداشته باشد.

چیدمانی از پیکسل‌ها در شکل (۳-۳) قسمت الف و پیکسل‌هایی که ۸ مجاورند در قسمت ب و حالت m -مجاوریتی با خطوط خط‌چین در قسمت ج برای $V=\{1\}$ نشان داده شده است.



شکل ۳-۳: (الف) چیدمان پیکسل‌ها. (ب) پیکسل‌هایی که ۸ مجاورند. (ج) پیکسل‌هایی که m مجاورند [۲۰].

با فرض این که S زیر مجموعه‌ای از پیکسل‌ها در یک تصویر باشد. دو پیکسل p و q ، متصل در S گفته می‌شوند اگر یک مسیر بین آن‌ها وجود داشته باشد که شامل تمامی پیکسل‌های S باشد. برای هر پیکسل p موجود در تصویر، مجموعه پیکسل‌هایی که متصل به این پیکسل باشند را یک جزء متصل^۱ در تصویر می‌گویند [۲۰].

۳-۳ نرمال‌سازی هیستوگرام

هیستوگرام یک تصویر دیجیتال با سطوح شدت روشنایی در بازه‌ی $[0, L-1]$ تابع گسسته‌ای به صورت $h(r_k) = n_k$ است که در آن r_k ، k امین مقدار شدت روشنایی و n_k تعداد پیکسل‌های تصویر با شدت روشنایی r_k است. نرمال‌سازی هیستوگرام با تقسیم هر جزء آن به تعداد کل پیکسل‌های تصویر MN معمول است (M و N ابعاد سطر و ستون تصویر هستند). بنابراین یک هیستوگرام نرمالیزه شده از رابطه‌ی $p(r_k) = \frac{n_k}{MN}$ به ازای $k = 0, 1, 2, \dots, L-1$ به دست می‌آید. در یک شکل تاریک، اجزای هیستوگرام در قسمت پایین (تاریک) مقیاس سطوح شدت روشنایی قرار دارند. به طور مشابه در یک شکل روشن، اجزای هیستوگرام در قسمت بالای (روشن) مقیاس سطوح روشنایی متمرکز شده‌اند. تصویری که پیکسل‌های آن همه سطوح ممکن شدت روشنایی را پوشش داده، و علاوه بر این توزیع یکنواخت داشته باشند، ظاهری با تباين بالا داشته و تغییرات سطوح خاکستری وسیع‌تر داشته و در نتیجه چنین تصویری، جزئیات بسیاری از سطوح خاکستری را نمایش می‌دهد و محدوده‌ی پویایی بالایی دارد. برای به دست آوردن تصویر با هیستوگرام نرمال، به صورت زیر عمل می‌کنیم.

^۱ Connected Component

با فرض پیوسته بودن سطوح شدت روشنایی تصویر و فرض این که r در بازه $[0, L-1]$ شدت روشنایی تصویر در حال پردازش باشد. برای هر r که در این شرایط صدق می کند. توجه خود را بر روی تبدیلی به صورت زیر متمرکز می کنیم:

$$s = T(r) \quad 0 \leq r \leq L - 1 \quad (5-3)$$

که به ازای هر پیکسل در تصویر ورودی با شدت روشنایی r ، یک خروجی با شدت روشنایی S می دهد. فرض می شود که :

(الف) $T(r)$ به صورت تکنوا در بازه $0 \leq r \leq L - 1$ صعودی است.

(ب) $0 \leq T(r) \leq L - 1$ به ازای $0 \leq r \leq L - 1$.

شرط (الف) تضمین کننده این موضوع است که مقادیر شدت روشنایی خروجی هیچ گاه از مقادیر شدت روشنایی ورودی متناظر کمتر نباشند. شرط (ب) تضمین می کند که دامنه های خروجی با ورودی یکی باشد. شکل (۴-۳) قسمت (الف)، تابعی را که شرایط (الف) و (ب) را برآورده می سازد، نمایش می دهد. در این جا می بینیم که ممکن است چند مقدار به یک مقدار نگاشت شوند و این دو حالت کماکان برقرار باشند. همچنان که شکل (۴-۳) قسمت (ب) نشان می دهد، شرایط تکنوای اکید بودن $T(r)$ باعث می شود که در نگاشت معکوس تنها یک مقدار داشته باشیم.

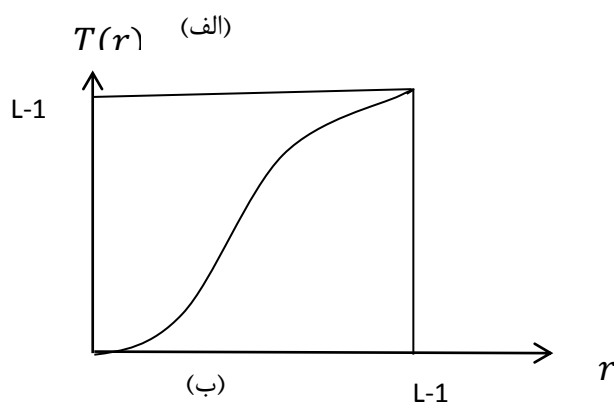
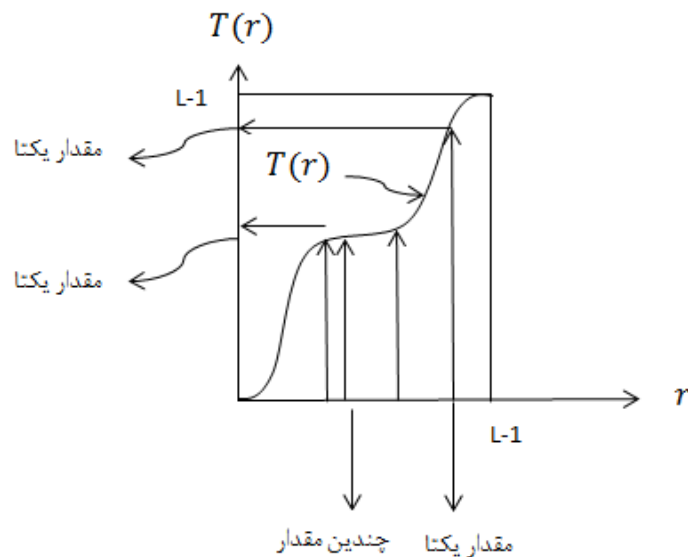
سطوح شدت روشنایی در یک تصویر را می توان به عنوان متغیرهای تصادفی در بازه $[0, L-1]$ در نظر گرفت. اگر $p_r(r)$ و $p_s(s)$ توابع چگالی احتمالی r و s باشند. آن گاه از تئوری احتمالات داریم که اگر $p_r(r)$ و $T(r)$ معلوم بوده و $T(r)$ در بازه مقادیر مورد بررسی، پیوسته و مشتق پذیر باشد، در این صورت تابع چگالی احتمالی متغیر S (نگاشت شده) را می توان از فرمول (۶-۳) یافت:

$$p_s(s) = p_r(r) \left| \frac{dr}{ds} \right| \quad (6-3)$$

یک تابع تبدیل با اهمیت در پردازش تصویر به صورت زیر است:

$$s = T(r) = (L - 1) \int_0^r p_r(w) dw \quad (7-3)$$

سمت راست این معادله به عنوان تابع توزیع تجمعی^۱ متغیر تصادفی r شناخته می‌شود. جهت یافتن $p_s(s)$ از معادله (۸-۳) استفاده می‌کنیم.



شکل ۳-۴: (الف) تابع صعودی که نشان می‌دهد چگونه چند مقدار به یک مقدار نگاشت می‌شوند. (ب) تابع تکنوای اکیداً صعودی که در هر دو جهت نگاشت یک به یک را موجب می‌شود.

$$\frac{ds}{dr} = \frac{dT(r)}{dr} = (L-1) \frac{d}{dr} \left[\int_0^r p_r(w) dw \right] = (L-1) p_r(r) \quad (۸-۳)$$

با جایگزینی این نتیجه به جای $\frac{dr}{ds}$ در معادله (۶-۳) و به خاطر مثبت بودن همه مقادیر داریم:

$$p_s(s) = p_r(r) \left| \frac{dr}{ds} \right| = p_r(r) \left| \frac{1}{(L-1)p_r(r)} \right| = \frac{1}{L-1} \quad 0 \leq s \leq L-1 \quad (۹-۳)$$

$p_s(s)$ یک تابع چگالی احتمال یکنواخت است.

^۱ Cumulative Distribution Function(CDF)

برای مقادیر گسسته، به جای توابع چگالی احتمال و انتگرال‌ها با احتمالات یا مقادیر هیستوگرام‌ها و جمع‌گذاری‌ها، سر و کار داریم. احتمال وقوع شدت روشنایی r_k در یک تصویر دیجیتال با رابطه (۳-۱۰) تقریب زده می‌شود:

$$p_r(r_k) = \frac{n_k}{MN} \quad k = 0, 1, 2, \dots, L-1 \quad (۱۰-۳)$$

که در آن MN تعداد کل پیکسل‌ها در تصویر، n_k تعداد پیکسل‌هایی که شدت روشنایی r_k دارند و L تعداد سطوح ممکن شدت روشنایی در تصویر است. فرم گسسته معادله (۳-۷) به صورت زیر است.

$$s_k = T(r_k) = (L-1) \sum_{j=0}^k p_r(r_j) = \frac{(L-1)}{MN} \sum_{j=0}^k n_j \quad k = 0, 1, \dots, L-1 \quad (۱۱-۳)$$

بنابراین یک تصویر خروجی با نگاشت هر پیکسل در تصویر ورودی با شدت روشنایی r_k به یک پیکسل متناظر با شدت روشنایی s_k در تصویر خروجی، با استفاده از معادله‌ی (۱۱-۳) حاصل می‌شود. تبدیل یا نگاشت $T(r_k)$ در این معادله را یک تبدیل نرمال‌سازی یا جبران‌سازی هیستوگرام می‌خوانند. که سبب می‌شود، تمام شدت‌های روشنایی، فراوانی تقریباً مشابهی داشته باشند [۱۹].

۳-۴ شبکه‌های عصبی

یک شبکه عصبی از واحدهای پردازنده ساده ساخته شده است و قابلیت ذخیره دانش تجربی را دارد. برخی از مهم‌ترین کاربردهای شبکه‌های عصبی عبارتند از: طبقه‌بندی با نظارت، طبقه‌بندی بدون نظارت یا همان خوشه‌بندی، تخمین تابع با خطای تقریباً صفر و پیش‌بینی سری‌های زمانی یعنی سری اطلاعاتی که با زمان تغییر می‌کند.

می‌توان شبکه‌های عصبی را بر اساس شیوه پردازش اطلاعات در آنها، به دو گروه شبکه‌های feed forward و شبکه‌های recurrent (که در آنها از فیدبک خروجی استفاده شده است) تقسیم بندی کرد. در شبکه‌های feed forward یک الگوی ورودی تنها یک خروجی تولید می‌کند و اطلاعات در یک جهت جاری می‌شوند، یعنی مسیر پسخور^۱ وجود ندارد، اما در شبکه‌های بازگشتی اطلاعات در چند

^۱ feedback

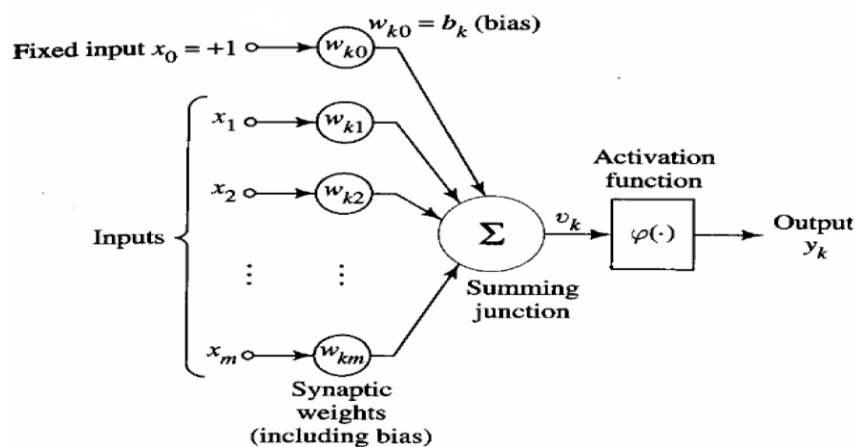
جهت جاری می شوند یعنی نرونها با استفاده از یک مسیر پسخور به خودشان یا نرونهای دیگر متصل می شوند.

نرون واحد اساسی پردازش اطلاعات در شبکه عصبی محسوب می شود. به این صورت که هر نرون ضرب ورودی در وزن ها و جمع آنها را انجام می دهد. مدل یک نرون در شکل (۳-۵) نشان داده شده است. همانطور که در این شکل مشاهده می شود پس از ضرب ورودی در وزن ها و جمع آنها، این مقدار یعنی v_k به عنوان ورودی تابع فعال سازی وارد این تابع می شود. نحوه محاسبه خروجی تابع فعال سازی به صورت زیر است [۲۱].

$$v_k = u_k + b_k \quad (۱۲-۳)$$

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (۱۳-۳)$$

$$y_k = \varphi(u_k + b_k) \quad (۱۴-۳)$$



شکل ۳-۵: مدل سازی یک نرون

با توجه به تنوعی که مقادیر وزن ها ممکن است داشته باشند نیاز به توابعی برای نرمال کردن آنها است. سه تابع فعال سازی در شکل زیر نشان داده شده است.

(۱) تابع فعال سازی پله

$$\varphi(v) = \begin{cases} 1 & v \geq 0 \\ 0 & v < 0 \end{cases} \quad (۱۵-۳)$$

(۲) تابع فعال سازی تکه ای خطی

$$\varphi(v) = \begin{cases} 1 & v \geq 0.5 \\ v & -0.5 < v < 0.5 \\ 0 & v \leq -0.5 \end{cases} \quad (۱۶-۳)$$

۳) تابع فعال سازی سیگموئید

$$\varphi(v) = \frac{1}{1+e^{-av}} \quad (۱۷-۳)$$

تابع فعال سازی سیگموئید نسبت به سایر توابع فعال سازی ترجیح داده می شود زیرا اولاً مشتق پذیر است و ثانیاً مشتق آن بر حسب خود تابع بدست می آید. مشتق تابع سیگموئید در رابطه (۱۸-۳) نشان داده شده است.

$$\varphi'(v) = a(1 - \varphi(v))\varphi(v) \quad (۱۸-۳)$$

مزیت استفاده از شبکه عصبی نسبت به سایر طبقه بندیها عبارت است از:

- ایجاد توابع مجزاساز غیرخطی. یعنی در صورتی که نمونه ها به صورت خطی جداپذیر نباشد، مرزهای غیرخطی بین آنها ایجاد خواهد شد.
- وقتی مثال زیادی برای آموزش داریم لازم نیست درگیر طراحی شویم و شبکه به خودی خود مرزهای بین کلاس ها را ایجاد خواهد کرد.
- در این روش نیازی به چگالی احتمالی شرطی کلاس ها نیست.
- فرآیند آموزش شبکه عصبی وقت گیر است اما بعد از این که آموزش تمام شد و وزن ها به دست آمدند فرآیند آزمایش شبکه سریع است.
- یک شبکه عصبی سه لایه به شرطی که تعداد گره های لایه میانی آن کافی باشد قادر به ساخت هر مرزی با هر پیچیدگی بین کلاس ها است.

۳-۴-۱ قاعده آموزش گرادیان نزولی

در این قاعده فرض می شود که تابع فعال سازی نرون ها یک تابع خطی است.

خطا در خروجی شبکه به صورت زیر محاسبه خواهد شد:

$$E(w) = E(w_1, w_2, \dots, w_m) = \frac{1}{2} \sum_{(x,d) \in D} (d - o(x))^2 \quad (۱۹-۳)$$

$$o(x) = w^t x \quad (۲۰-۳)$$

رابطه اصلاح وزن به صورت زیر است:

$$w_i(k+1) = w_i(k) + \Delta w_i \quad (۲۱-۳)$$

از طرفی:

$$\Delta w = -\eta(\nabla E(w)) = -\eta \left[\frac{\partial E}{\partial w_1}, \dots, \frac{\partial E}{\partial w_m} \right] \quad (۲۲-۳)$$

$$\Delta w_i = -\eta \frac{\partial E}{\partial w_i} \quad (۲۳-۳)$$

علامت منفی به این دلیل است که جهتی که خطا حداقل می‌شود، خلاف جهت مشتق است.

در نتیجه رابطه اصلاح وزن در این قاعده آموزش به صورت زیر خواهد بود [۲۱]:

$$w_i(k+1) = w_i(k) + \eta \sum_n (d_n - o_n) x_{in} \quad (۲۴-۳)$$

۳-۴-۲ شبکه عصبی MLP

ساختار یک پرسپترون سه لایه در شکل (۳-۶) نشان داده شده است. مشکل این شبکه این است که برای لایه میانی به دلیل این که مقادیر هدف یا تارگت تعریف نمی‌شود، نمی‌توان خطا را برای آن محاسبه کرد. بنابراین برای آموزش این شبکه و اصلاح وزن‌ها از روش پس‌انتشار خطا استفاده می‌شود. این روش، روش حداقل کردن مربع خطا است، که خطا در خروجی شبکه محاسبه می‌شود و این خطا به سمت ورودی انتشار داده می‌شود و وزن‌ها اصلاح می‌گردد [۲۵].

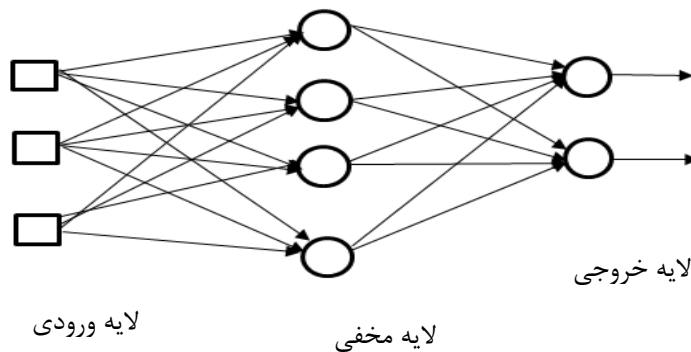
رابطه اصلاح وزن بین لایه میانی و لایه خروجی به صورت زیر خواهد بود [۲۲]:

$$\Delta w_{kj} = \eta \delta_k y_j = \eta (t_k - z_k) f'(net_k) y_j \quad (۲۵-۳)$$

که y_j خروجی نرون j ام لایه میانی و z_k خروجی نرون k ام لایه خروجی است. t_k خروجی هدف یا مقدار تارگت^۱ نرون k ام لایه خروجی است. η ضریب اصلاح وزن و w_{kj} وزن بین نرون j ام لایه میانی و نرون k ام لایه خروجی است. با توجه به این که تابع فعال سازی تمام نرون‌ها تابع سیگموئید است بنابراین مقدار $f'(net_k)$ به صورت زیر محاسبه می‌شود:

^۱ Target

$$f'(net_k) = net_k(1 - net_k) \quad (۲۶-۳)$$



شکل ۳-۶: ساختار شبکه عصبی پرسپترون سه لایه

رابطه اصلاح وزن بین لایه ورودی و میانی طبق رابطه زیر است [۲۲]:

$$\Delta w_{ji} = \eta x_i \delta_j = \eta \left[\sum_{k=1}^c w_{kj} \delta_k \right] f'(net_j) x_i \quad (۲۷-۳)$$

که:

$$f'(net_j) = net_j(1 - net_j) \quad (۲۸-۳)$$

$$\delta_k = (t_k - z_k) f'(net_k) \quad (۲۹-۳)$$

x_i ورودی نرون i ام لایه ورودی است. و w_{ji} وزن بین نرون i ام لایه ورودی و نرون j ام لایه میانی است.

اصلاح وزن شبکه را می‌توان به چند طریق انجام داد:

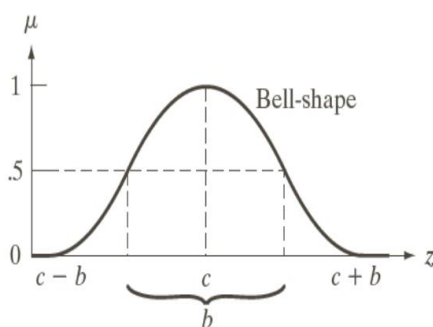
- ۱- اصلاح وزن بعد از دیدن هر نمونه آموزشی انجام شود.
- ۲- اصلاح وزن بعد از دیدن تمام نمونه‌های آموزشی اصلاح وزن انجام شود.
- ۳- نمونه‌های آموزشی به صورت تصادفی از مجموعه آموزش خوانده شده و به شبکه آموزش داده می‌شوند.
- ۴- بیشتر از نمونه‌هایی برای آموزش شبکه استفاده شود که شبکه نتوانسته آن‌ها را یاد بگیرد.

۳-۵ منطق فازی^۱

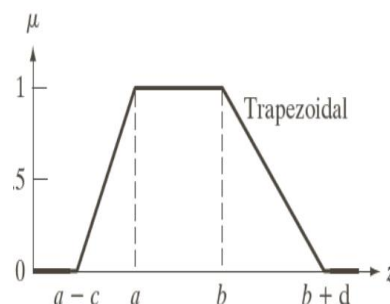
^۱ Fuzzy Logic

نظریه فازی برای اولین بار در سال ۱۹۶۵ ارائه شد. مجموعه‌های فازی یک فرمالیسم را در جهت پردازش اطلاعات غیردقیق فراهم می‌کنند.

با فرض این که C مجموعه‌ای از عناصر باشد، یک عنصر کلی از Z با z نشان داده می‌شود، یعنی $Z = \{z\}$. این مجموعه، مجموعه کلی خوانده می‌شود. مجموعه فازی A در Z با تابع عضویت $\mu_A(z)$ مشخص می‌شود که به هر جزء z یک مقدار حقیقی در بازه $[0,1]$ نسبت می‌دهد. مقدار $\mu_A(z)$ در Z درجه عضویت z در A را بیان می‌کند. هرچه مقدار $\mu_A(z)$ به واحد نزدیک‌تر باشد، درجه عضویت z در A بالاتر است. و هرچه مقدار $\mu_A(z)$ به صفر نزدیک‌تر باشد، عکس این امر صادق است. دو تابع عضویت در شکل (۷-۳) قسمت الف و ب نشان داده شده است [۱۹].



(ب)



(الف)

شکل ۷-۳: (الف) تابع عضویت ذوزنقه‌ای و (ب) تابع عضویت زنگی شکل.

مفهوم عضویت یا تعلق که در مجموعه‌های عادی بسیار آشنا است، در مجموعه‌های فازی معنای یکسانی ندارد. در مجموعه‌های عادی یک عنصر، عضو یک مجموعه متعلق هست یا نیست. در مجموعه‌های فازی همه عناصری که $\mu_A(z) = 1$ دارند، عضو کامل مجموعه هستند. و همه z هایی که $\mu_A(z) = 0$ دارند، عضو مجموعه نیستند. و همه z هایی که $\mu_A(z)$ آن‌ها بین دو مقدار صفر و یک است، اجزای نسبی مجموعه هستند. بنابراین یک مجموعه فازی یک زوج مرتب شامل مقادیر z و یک تابع عضویت که به هر z یک درجه عضویت نسبت می‌دهد، می‌باشد. یعنی:

^۱ Membership Function

$$A = \{z, \mu_A(z) | z \in Z\} \quad (3-30)$$

چند تعریف در مجموعه‌های فازی [۱۹]:

- ۱- مجموعه تهی: یک مجموعه فازی تهی است اگر و تنها اگر تابع عضویت آن در Z صفر باشد.
- ۲- معادل بودن: دو مجموعه فازی A و B در صورتی معادل هستند که $A = B$ اگر و تنها اگر

$$\mu_A(z) = \mu_B(z) \quad (\text{به ازای همه‌ی مقادیر } z \in Z)$$

- ۳- متمم (NOT) یک مجموعه فازی را با A^- یا $\text{NOT}(A)$ نمایش می‌دهند. و به صورت مجموعه‌ای که تابع عضویت آن به ازای همه‌ی $z \in Z$ به صورت زیر است، تعریف می‌کنند:

$$\mu_{A^-}(z) = 1 - \mu_A(z) \quad (3-31)$$

- ۴- زیر مجموعه: مجموعه فازی A زیر مجموعه‌ی مجموعه فازی B است، اگر و تنها اگر:

$$\mu_A(z) \leq \mu_B(z) \quad (3-32)$$

- ۵- اجتماع: اجتماع دو مجموعه فازی A و B را با $A \cup B$ نمایش می‌دهیم. مجموعه‌ی حاصل را محاسبه می‌کنیم که به ازای هر U عضویت زیر را دارد:

$$\mu_U(z) = \max[\mu_A(z), \mu_B(z)] \quad (3-33)$$

- ۶- اشتراک دو مجموعه فازی به صورت زیر است:

$$\mu_I(z) = \min[\mu_A(z), \mu_B(z)] \quad (3-34)$$

۳-۵-۱ سیستم استنتاج فازی

سیستم‌های فازی، سیستم‌های مبتنی بر دانش یا قواعد هستند. قلب یک سیستم فازی یک پایگاه دانش است، که از قواعد اگر-آن‌گاه فازی تشکیل شده است. به عنوان مثال اگر سرعت خودرو بالاست، آن‌گاه نیروی کمتری به پدال گاز وارد کنید. کلمات بالا و کم به وسیله توابع تعلق مشخص شده‌اند. فرض کنید می‌خواهیم کنترل‌کننده‌ای طراحی کنیم که سرعت خودرو را به طور خودکار کنترل کند. راه حل این است که رفتار رانندگان را شبیه‌سازی کنیم. بدین معنی که قواعدی را که راننده در حین حرکت استفاده می‌کند، به کنترل کننده خودکار تبدیل نماییم.

در صحبت‌های عامیانه راننده‌ها در شرایط طبیعی از ۳ قاعده زیر در حین رانندگی استفاده می‌کنند:

اگر سرعت پایین است، آن‌گاه نیروی بیشتری به پدال گاز وارد کنید.

اگر سرعت متوسط است، آن‌گاه نیروی متعادلی به پدال گاز وارد کنید.

اگر سرعت بالاست، آن‌گاه نیروی کم‌تری به پدال گاز وارد کنید.

انجام استدلال‌های فازی به عهده‌ی سیستم استنتاج فازی است. سیستمی که یک نداشت از ورودی

به خروجی را با استفاده از منطق فازی فرموله می‌کند به نام سیستم استنتاج فازی FIS شناخته می‌

شود. سیستم استنتاج فازی همچنین به نام سیستم مبتنی بر قواعد نیز نامیده می‌شود. زیرا این

سیستم‌ها از تعدادی عبارت «اگر – آنگاه» ساخته شده است.

سیستم‌های استنتاج فازی را می‌توان به سه کلاس «ممدانی (mamdani)»، «سوگنو (sugeno)» و

و «تاگاسی (Takagi)» تقسیم کرد.

۳-۵-۲ کاربردهای سیستم فازی

منطق فازی کاربردهای متعددی دارد. ساده‌ترین نمونه، یک سیستم کنترل دما یا ترموستات است، که

بر اساس قوانین فازی کار می‌کند. سال‌هاست که از منطق فازی برای کنترل دمای آب یا میزان کدر

شدن آبی که لباس‌ها در آن شسته شده‌اند در ساختمان‌های لباسشویی استفاده می‌شود.

منطق فازی در صنعت خودروسازی نیز کاربردهای فراوانی دارد. مثلاً سیستم ترمز در برخی از

خودروها از منطق فازی استفاده می‌کند. یکی از معروف‌ترین نمونه‌های به کارگیری منطق فازی در

سیستم‌های ترابری جهان، شبکه مونوریل (قطار تک‌ریل) توکیو در ژاپن است. سایر سیستم‌های

حرکتی و جابه‌جایی بار، مثل آسانسورها نیز از منطق فازی استفاده می‌کنند. سیستم‌های تهویه هوا

نیز به وفور منطق فازی را به کار می‌گیرند. همچنین از منطق فازی در سیستم‌های پردازش تصویر نیز

استفاده می‌شود.

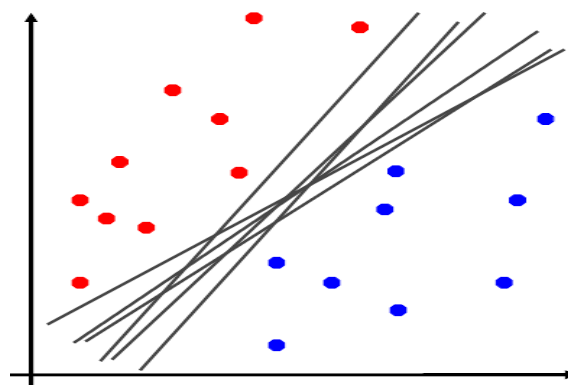
۳-۶ ماشین بردار پشتیبان

این شبکه در سال ۱۹۹۵ توسط Vapnik معرفی شد و بر پایه تئوری آموزش آماری دو کلاسه بنا گردید [۲۳]. بعدها در سال ۱۹۹۹ توسط Scholkopf و همکارانش گسترش یافت.

ماشین بردار پشتیبان یک طبقه‌بند با نظارت^۱ است. مزیت استفاده از ماشین بردار پشتیبان این است که اولاً بهترین مرز ممکن را بین دو کلاس ایجاد می‌کند یا مرزی که فاصله بین نمونه‌های مرزی تا آن مرز حداکثر شود. و ثانیاً نسبت به طبقه‌بندهایی مانند شبکه عصبی نیازمند داده آموزش کمتری است. اما عیبی که این طبقه‌بند دارد این است که قادر به جداسازی دو کلاس است که با در نظر گرفتن یک کلاس در مقابل بقیه این مشکل رفع می‌شود [۲۲].

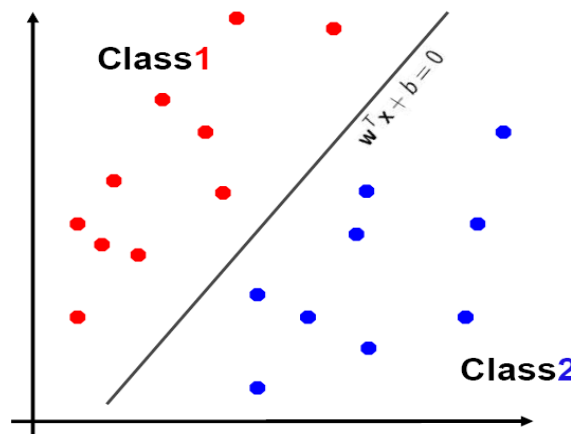
با توجه به شکل (۳-۸) خط‌های زیادی وجود دارد که قادر به جداسازی دو صفحه است اما بهترین مرز بین دو کلاس خطی است که بیشترین فاصله را تا الگوهای مرزی دارد. این خط در شکل (۳-۹) نشان داده شده است.

نمونه‌هایی که کمترین فاصله تا مرز را دارند بردارهای پشتیبان هستند. ماشین بردار پشتیبان، یک طبقه‌بند دو کلاسه است. برای یک مساله دو کلاسه داده‌های آموزشی به صورت $\{x_i, y_i\}$ بر چسب می‌خورند که $i=1, 2, \dots, l$ و $y_i \in \{-1, 1\}$ و $x_i \in \mathbb{R}^d$ است. y_i برای یک کلاس ۱ و برای کلاس دیگر ۱- است. d بعد بردار ویژگی و l تعداد نمونه‌های آموزش است.



شکل ۳-۸: مرزهای بیشماری که بین دو کلاس وجود دارند.

^۱ Supervised



شکل ۳-۹: مرزی که بیشترین فاصله را تا نمونه‌های مرزی دارد.

هر داده جدید درست طبقه‌بندی شده است به شرطی که در رابطه (۳۵-۳) صدق کند.

$$y_i(x_i \cdot w^T + b) - 1 \geq 0 \quad \forall i \quad (35-3)$$

w بردار عمود بر مرز مجزا ساز است. نقاطی که در معادله $x_i \cdot w^T + b = 1$ و معادله $x_i \cdot w^T + b = -1$ صدق می‌کنند همان بردارهای پشتیبان هستند. فاصله بین این دو خط برابر با $2/\|w\|$ است، هدف این است که این فاصله حداکثر شود به این منظور باید $\|w\|^2/2$ حداقل شود. طبق این و با توجه به معادله ۱ و تعریف ضرایب $\alpha_i \geq 0$ ، می‌بایست معادله لاگرانژین را که به صورت زیر تعریف می‌شود بهینه کنیم [۲۴]:

$$L_p = \frac{1}{2} \|w\|^2 - \sum_{i=0}^l \alpha_i y_i (x_i^T w + b) + \sum_{i=0}^l \alpha_i \quad (36-3)$$

حال اگر از رابطه بالا نسبت به w و b مشتق جزئی گرفته شده و مساوی صفر قرار داده شود، مقدار بهینه w و b به دست خواهد آمد که این کار در معادلات (۳۷-۳) و (۳۸-۳) نشان داده است.

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w_0 = \sum_{i=1}^l \alpha_i x_i y_i \quad (37-3)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=0}^l \alpha_i y_i = 0 \quad (38-3)$$

حال اگر مقدار w به دست آمده از رابطه (۳۷-۳) را در رابطه (۳۶-۳) قرار دهیم آن‌گاه:

$$L_p = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (39-3)$$

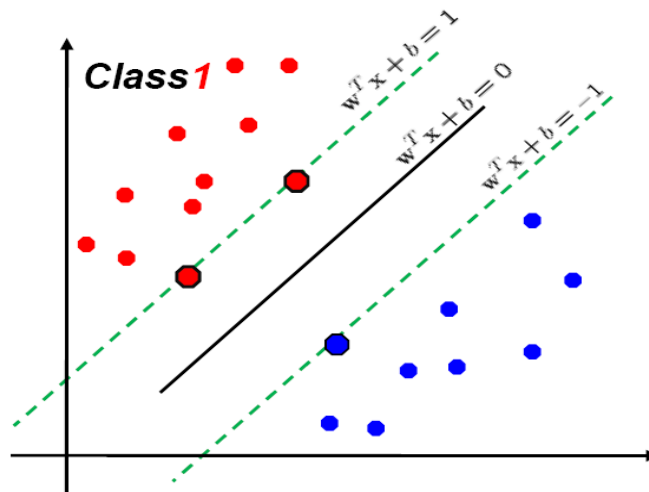
به ازای هر بردار پشتیبان یک b به دست خواهد آمد که با توجه به $y_i(w^T x_i + b) = 1$ میانگین به صورت زیر به دست خواهد آمد:

$$b = \frac{1}{N} \sum_{i=1}^N (y_i - x_i^T w_0) \quad (40-3)$$

N در این جا تعداد بردارهای پشتیبان است. پس از این که w و b محاسبه شد، بهترین مرز ممکن بین دو کلاس بدست خواهد آمد که این مرز در شکل (۳-۱۰) نشان داده شده است. آن گاه با تعریف تابع علامت به صورت زیر به راحتی می توان یک داده جدید را طبقه بندی کرد.

$$f(x) = \text{sgn}(x \cdot w^T + b) \quad (41-3)$$

برای داده جدید اگر $x \cdot w^T + b > 0$ باشد آن گاه مقدار تابع برابر با ۱ خواهد بود و این داده مربوط به کلاس یک خواهد بود و اگر $x \cdot w^T + b < 0$ آن گاه مقدار تابع $f(x)$ برابر با -۱ خواهد بود و داده مربوط به کلاس ۲ خواهد بود.



شکل ۳-۱۰: محاسبه بهترین مرز بین دو کلاس.

در مواردی که داده ها به صورت خطی جداپذیر نیستند توسط تابعی، داده ها به فضای دیگر منتقل می شوند که در فضای جدید امکان تفکیک خطی کلاس ها وجود دارد. تابعی که توسط آن داده ها را به فضای جدید نگاشت داده می شود تابع کرنل نام دارد. چند تابع کرنل در جدول (۳-۱) نشان داده شده است.

جدول ۳-۱: چند تابع کرنل

نوع طبقه‌بندی	تابع کرنل
تابع خطی از درجه p	$K(x_i, x_j) = (x_i^T x_j)^p$
تابع چند جمله‌ای از درجه p	$K(x_i, x_j) = (x_i^T x_j + 1)^p$
تابع گوسی	$K(x_i, x_j) = e^{-\frac{\ x_i - x_j\ ^2}{2\sigma^2}}$
تابع پرسپترون چند لایه	$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + \mu)$
تابع دیریکله برای مسائل شرایط مرزی	$K(x_i, x_j) = \frac{\sin((n + \frac{1}{2})(x_i - x_j))}{2\sin(\frac{x_i - x_j}{2})}$

تابع کرنل به صورت زیر تعریف می‌شود:

$$k(x_i, x_j) = \varphi^t(x_i)\varphi(x_j) \quad (۳-۴۲)$$

در این حالت معادله (۳-۳۷) و (۳-۳۹) به شکل زیر تغییر می‌کنند:

$$w = \sum_{i=1}^l \alpha_i y_i \varphi(x_i) \quad (۳-۴۳)$$

$$Lp = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j \varphi^t(x_i) \varphi(x_j) \quad (۳-۴۴)$$

۷-۳ کتابخانه Open CV

Open CV یا همان Open Computer Vision Library مجموعه‌ای از کتابخانه‌های برنامه‌نویسی پردازش تصویر است. و همانطور که از نامش پیداست یک کتابخانه متن باز است. این مجموعه بیشتر بر پردازش تصویر بلادرنگ^۱ تمرکز دارد. این کتابخانه ابتدا توسط شرکت اینتل ساخته شد. و هم اکنون توسط شرکت‌های Willow Garage و Itseez پشتیبانی می‌گردد. دلیل استفاده از یک

^۱ Real time

کتابخانه آماده این است که گاهی اوقات نوشتن برخی از توابعی که نیازمندیم در برنامه از آن‌ها استفاده کنیم، بسیار زمان‌بر است.

این کتابخانه، ابتدا به زبان C نوشته و توزیع می‌شد و از نسخه‌ی ۲ به بعد و زمانی که به شرکت Willow Garage سپرده شد، تغییرات زیادی در پیاده‌سازی کتابخانه به منظور پشتیبانی از C++ صورت گرفت و هم‌اکنون توابع این کتابخانه به هر دو زبان C و C++ و همچنین زبان پایتون در اختیار قرار دارد.

کتابخانه Open CV بر خلاف اغلب کتابخانه‌های پردازشی که تنها برای یک سیستم عامل قابل استفاده‌اند، در سیستم عامل‌های ویندوز، لینوکس، ISO و اندروید قابل استفاده است. که این مزیت بزرگی برای Open CV نسبت به سایر کتابخانه‌ها محسوب می‌شود.

اگر بخواهیم این کتابخانه را با نرم‌افزارهایی مانند نرم‌افزار متلب^۱ مقایسه کنیم، کار کردن با این کتابخانه کمی دشوارتر است، اما سرعت اجرای برنامه‌های آن از نرم‌افزار متلب غالباً بیشتر است. و فایل‌های اجرایی و خروجی آن بر روی هر کامپیوتری بدون نیاز به نصب برنامه اصلی آن قابل اجرا خواهد بود، این در حالی است که برنامه نوشته شده در نرم‌افزار متلب تنها بر روی سیستمی اجرا خواهد شد که اصل^۲ این برنامه بر روی آن سیستم نصب باشد. بنابراین نتیجه می‌گیریم که کتابخانه Open CV بسیار مناسب برای کارهای تجاری و کاربردی است. به عنوان مثال برخی از توابع این کتابخانه عبارتند از: پیدا کردن مرزهای تصویر، تغییر ابعاد تصویر، نرمال‌سازی روشنایی تصویر و انجام عملیات مورفولوژیکی بر روی تصویر.

^۱ Matlab
^۲ Source

فصل چهارم:

روش پیشنهادی

۴-۱ مقدمه

هدف از انجام این پایان‌نامه، آشکارسازی و شناسایی علائم ترافیکی است. آشکارسازی یعنی یافتن موقعیت علامت در تصویر و شناسایی، یعنی تشخیص نوع علامت از قبیل ورود ممنوع، سبقت آزاد یا محدودیت سرعت. در این فصل، الگوریتم پیشنهادی برای آشکارسازی و شناسایی نوع دقیق علائم ترافیکی در صحنه‌های واقعی را ارائه می‌کنیم. ابتدا به آشکارسازی علامت راهنمایی و رانندگی از درون تصویر می‌پردازیم. سپس شناسایی نوع دقیق علامت آشکار شده انجام می‌شود. و در نهایت به بیان نتایج تجربی و مقایسه نتایج به دست آمده با سایر کارهای انجام شده پرداخته می‌شود. مراحل آشکارسازی و شناسایی دقیق علامت که در این پایان‌نامه انجام شده است، به طور خلاصه در شکل (۴-۱) نشان داده شده است.

۴-۲ تهیه یک پایگاه داده از علائم راهنمایی و رانندگی کشور ایران

به علت عدم وجود یک پایگاه داده از علائم راهنمایی و رانندگی که از سطح خیابان‌ها و جاده‌های ایران از علائم تهیه شده باشد. اقدام به تهیه یک پایگاه داده از علائم راهنمایی و رانندگی موجود در ایران کردیم.

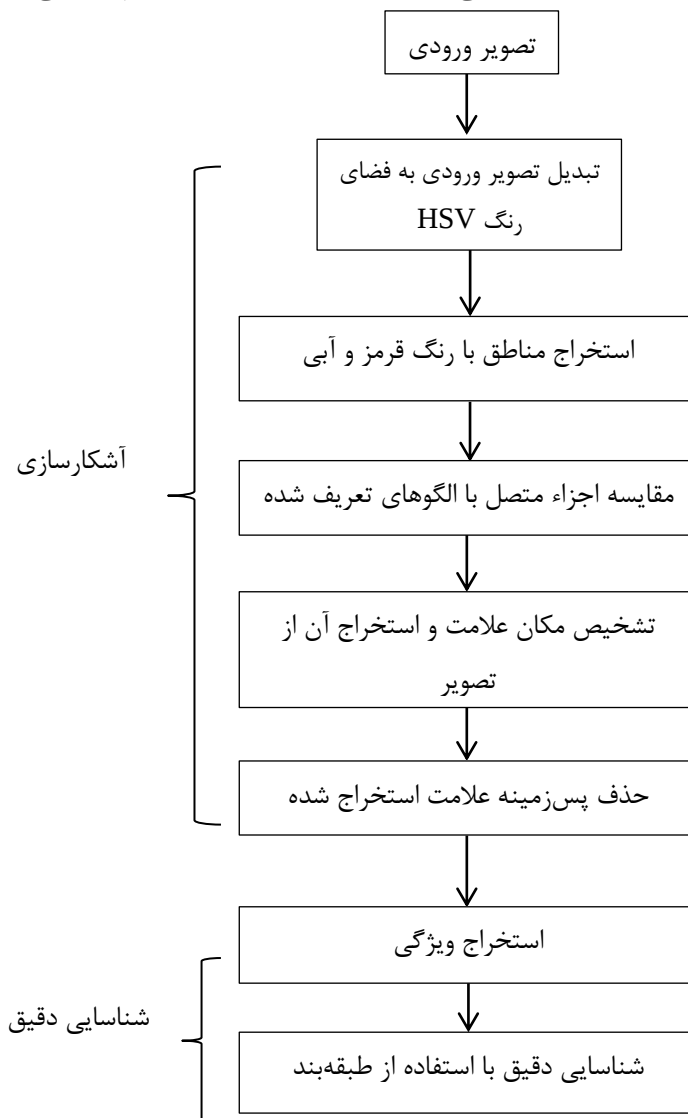
یک پایگاه داده از علائم راهنمایی و رانندگی باید دارای ویژگی‌های زیر باشد:

۱. حتی الامکان شامل تمام انواع از علائم راهنمایی و رانندگی باشد.
۲. تصاویر گرفته شده از علائم، از فواصل و زوایای مختلف و در مقیاس‌های متفاوت باشند.
۳. شرایط واقعی تصویربرداری از روی خودرو که غالباً با تاری و لرزش همراه است، در برخی از تصاویر پایگاه داده رعایت شود.
۴. در شرایط نوری مختلف از علائم تصویربرداری شود. مانند شرایط آب و هوای برفی، بارانی و آفتابی، یا علائمی که در سایه ساختمان‌ها یا درختان قرار دارند و دارای نور کم هستند. باید

در شب از علائم تصویر برداری شود. زیرا در شب، نور چراغ جلوی وسیله نقلیه بر روی علامت می افتد و رنگ قسمت های مختلف علامت، خصوصاً نواحی سفید را تغییر می دهد.

۵. از علائمی که در اثر گذشت زمان رنگ آن پریده و یا علامت رانندگی خراب یا کج شده است، تصویر برداری شود.

۶. کیفیت سیستم تصویر برداری خیلی بالا نباشد، زیرا هزینه سیستم بالا می رود.



شکل ۴-۱: مراحل انجام آشکارسازی و شناسایی علامت راهنمایی و رانندگی

تصاویری که ما تهیه کرده ایم، از سطح خیابان ها و جاده های شهری و بین شهری از علائم گرفته شده است. ویژگی این پایگاه داده این است، که اولاً تصاویر از فواصل و موقعیت های مختلف از علامت گرفته شده است. و ثانیاً در شرایط نوری و پس زمینه های مختلفی قرار دارند، همچنین برخی از

تصاویر، از روی خودرو در حال حرکت گرفته شده است. این پایگاه داده شامل ۳۵۰ علامت راهنمایی و رانندگی در ۳۵ کلاس مختلف است. هر علامت دارای ۲۰۰۰ سطر و ۲۵۰۰ ستون است، و دوربینی که توسط آن این پایگاه داده تهیه شده است ۵ مگاپیکسل بوده است. تعدادی از علائم موجود در این پایگاه داده در شکل (۲-۴) نشان داده شده است.



شکل ۲-۴: تعدادی از علائم راهنمایی و رانندگی موجود در پایگاه داده‌ای که اقدام به تهیه آن کردیم.

۳-۴ الگوریتم پیشنهادی برای آشکارسازی علامت راهنمایی و رانندگی

هدف از مرحله آشکارسازی این است که علامت را از میان اشیاء مختلفی که در تصویر وجود دارد پیدا کنیم.

علائم راهنمایی و رانندگی همانطور که در شکل (۳-۴) نشان داده شده است، دارای چند شکل و رنگ کلی هستند. رنگهای مورد استفاده شامل قرمز، آبی، زرد و سفید است و اشکال هندسی مورد استفاده

شامل مثلث، دایره، مستطیل، لوزی، پنج ضلعی و هشت ضلعی است. از همین دو خاصیت یعنی رنگ و شکل هندسی برای آشکارسازی علامت استفاده می‌کنیم.

در روش پیشنهادی ابتدا، به منظور کاهش حساسیت نسبت به تغییرات نور، تصویر ورودی را از فضای رنگ RGB به فضای رنگ HSV می‌بریم. فضای رنگ HSV به این دلیل که پارامتر روشنایی از پارامترهای فام^۱ و اشباع جداست یا اطلاعات رنگ در پارامترهای فام و اشباع است، روشنایی کم‌ترین تاثیر را بر روی استخراج مناطق با رنگ موردنظر ما خواهد داشت [۹].



شکل ۳-۴: دسته بندی کلی علائم راهنمایی و رانندگی از نظر رنگ و شکل هندسی.

آستانه پارامتر فام برای رنگ قرمز در دو بازه‌ی $0 < H < 0.07$ و $0.9 < H < 1$ و برای رنگ آبی $0.5 < H < 0.7$ و آستانه اشباع برای رنگ قرمز در محدوده‌ی $0.5 < S < 1$ و برای رنگ آبی $0.4 < S < 0.95$ است.

این محدوده‌ها به این صورت به دست آمده است، که هیستوگرام کانال‌های فام و اشباع را برای یک تصویر تماماً قرمز و یک تصویر تماماً آبی (همانند دو تصویر نشان داده شده در شکل (۲-۵) از فصل دوم) رسم کرده‌ایم. و این محدوده‌ها را از روی این هیستوگرام‌ها به دست آورده‌ایم. استخراج مناطق با رنگ قرمز توسط رابطه (۴-۱) و استخراج مناطق با رنگ آبی توسط رابطه (۴-۲) صورت می‌گیرد.

$$I(x, y) = \begin{cases} 1 & (0 < H(x, y) < 0.07 \text{ or } 0.9 < H(x, y) < 1) \text{ and } 0.5 < S(x, y) < 1 \\ 0 & \text{otherwise} \end{cases} \quad (1-4)$$

^۱ Hue

$$I(x,y) = \begin{cases} 1 & 0.5 < H(x,y) < 0.7 \quad \text{and} \quad 0.4 < S(x,y) < 0.95 \\ 0 & \text{otherwise} \end{cases} \quad (۲-۴)$$

در شکل (۴-۴) قسمت الف تصویر ورودی در فضای رنگ RGB بوده و قسمت ب سه کانال H، S و V این تصویر را در فضای رنگ HSV نشان می‌دهد.

پس از این که تصویر ورودی را به فضای رنگ HSV تبدیل کردیم با آستانه‌گذاری بر روی پارامترهای فام و اشباع مناطق آبی و قرمز تصویر ورودی را استخراج می‌کنیم. همانطور که در شکل (۴-۵) قسمت الف مشاهده می‌شود همه اجزاء تصویر که به رنگ آبی و قرمز بوده‌اند استخراج شده‌اند. سپس حفره‌های نواحی استخراج شده را پر می‌کنیم، شکل (۴-۵) قسمت ب این تصویر را نشان می‌دهد. از عملیات مورفولوژیک فرسایش استفاده کرده و نویزهایی که درون تصویر وجود دارند را حذف می‌کنیم (شکل (۴-۵) قسمت ج).



(الف)



(ب)

شکل ۴-۴: (الف) تصویر ورودی در فضای رنگ RGB و (ب) به ترتیب از راست به چپ سه کانال H، S و V این تصویر را در فضای رنگ HSV نشان می‌دهد.



(ج)

(ب)

(الف)

شکل ۴-۵: (الف) نقاط آبی و قرمز استخراج شده. (ب) پر کردن حفره‌های قسمت الف. (ج) حذف نویزهای موجود در تصویر ب با استفاده از عملیات مورفولوژیکی.

سپس تک تک اجزاء متصل از درون تصویر باینری را استخراج می‌کنیم و با الگوهایی که از قبل آماده کرده‌ایم مقایسه می‌کنیم. هرکدام از اجزاء متصل که کم‌ترین اختلاف را با یکی از این الگوها داشته باشد به عنوان علامت ترافیکی، نگه داشته و بقیه قسمت‌های تصویر را حذف می‌کنیم. الگوهایی که اجزاء متصل را با آن‌ها مقایسه می‌کنیم در شکل (۴-۶) الف تا ج نشان داده شده است.



(ج)

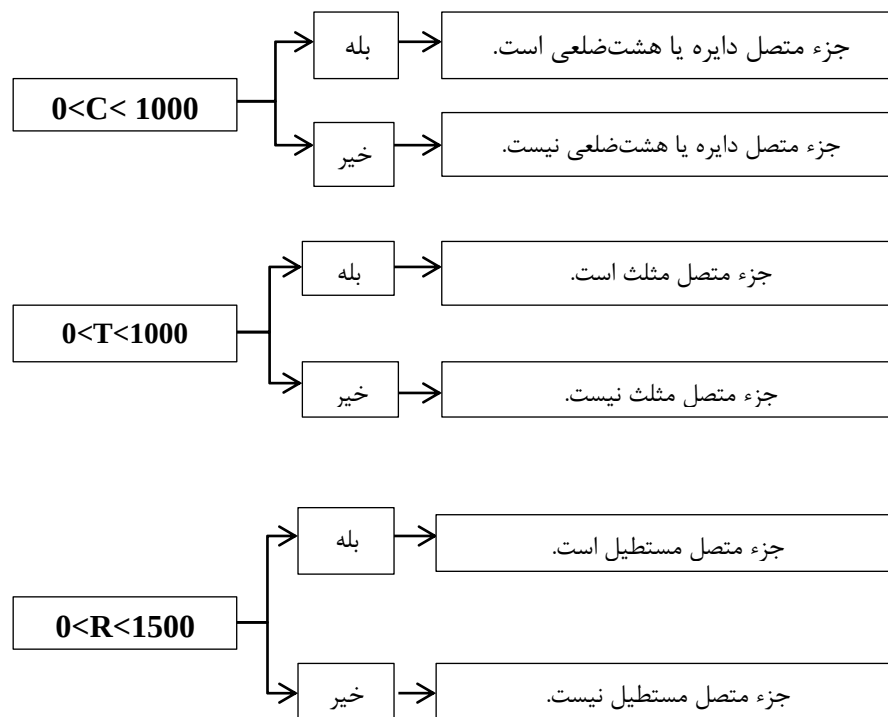
(ب)

(الف)

شکل ۴-۶: الگوهایی که در مرحله آشکارسازی اجزاء متصل موجود در تصویر باینری را با آن‌ها مقایسه می‌کنیم. فرآیند مقایسه به این صورت انجام می‌شود که ابعاد جزء متصل را به 100×100 تبدیل می‌کنیم. سپس تفاضل آن را با سه الگوی نشان داده شده در شکل (۴-۶) به دست می‌آوریم. ابعاد این سه الگو نیز 100×100 است. نتیجه تفاضل جزء متصل با هر یک از این الگوها خود یک تصویر باینری خواهد بود. حال اگر مجموع پیکسل‌های تصویر به دست آمده از تفاضل جزء متصل با الگوی دایره را C، مجموع پیکسل‌های تصویر به دست آمده از تفاضل جزء متصل با الگوی مثلث را T و مجموع پیکسل‌های تصویر به دست آمده از تفاضل جزء متصل با الگوی مستطیل را R بنامیم. فرآیند پیدا کردن دایره، هشت‌ضلعی، مثلث و مستطیل در تصویر باینری، در شکل (۴-۷) نشان داده شده است. در مرحله آشکارسازی، دایره و هشت‌ضلعی را در یک کلاس در نظر می‌گیریم.

در صورتی که جزء متصل دایره، هشت‌ضلعی، مثلث و مستطیل بود آن را در تصویر نگه داشته، در غیر این صورت آن را حذف می‌کنیم. نتیجه آشکارسازی علامت را در شکل (۴-۸) قسمت الف تا ج مشاهده می‌کنیم. برای ارزیابی عملکرد این الگوریتم ۳۳۰ تصویر، از علائم موجود در پایگاه داده که شامل مثلث، دایره، هشت‌ضلعی و مستطیل است را مورد آزمایش قرار دادیم، که سیستم قادر به آشکارسازی ۹۳.۹۳ درصد از علائم است. اگر بیش از یک علامت در یک تصویر موجود باشد، سیستم قادر به استخراج آن‌ها است.

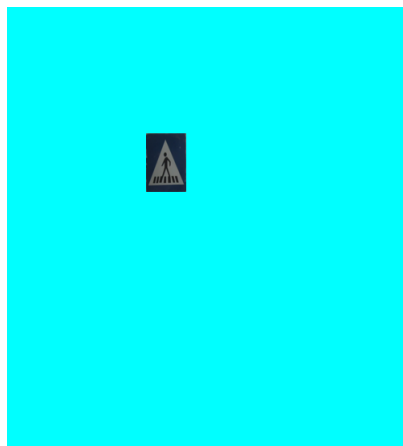
چند نمونه از تصاویری که سیستم قادر به استخراج علامت نبوده است، در شکل (۴-۹) نشان داده شده است. علت عدم استخراج علامت در این تصاویر این است که وقتی ناحیه قرمز و آبی تصویر را استخراج می‌کنیم، قسمت قرمز و آبی علامت به خوبی استخراج نمی‌شود. این موضوع در شکل (۴-۹) ردیف سمت راست تصویر، نشان داده شده است. حال برای رفع این مشکل به سراغ منطق فازی می‌رویم.



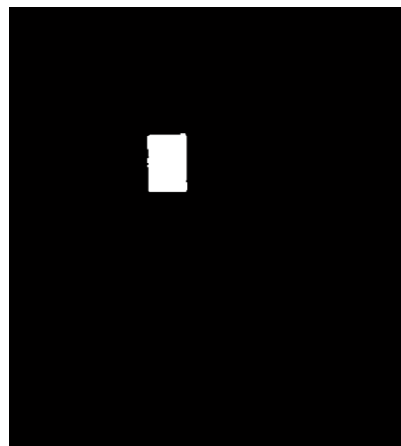
شکل ۴-۷: فرآیند پیدا کردن دایره، مثلث و مستطیل در تصویر باینری.



(ج)

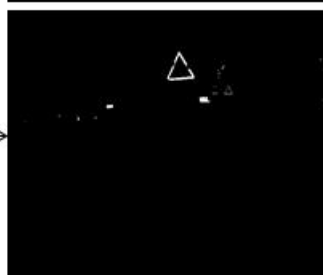
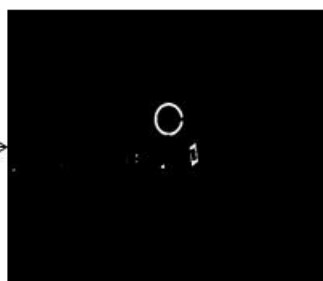


(ب)



(الف)

شکل ۴-۸: نتایج آشکارسازی علامت

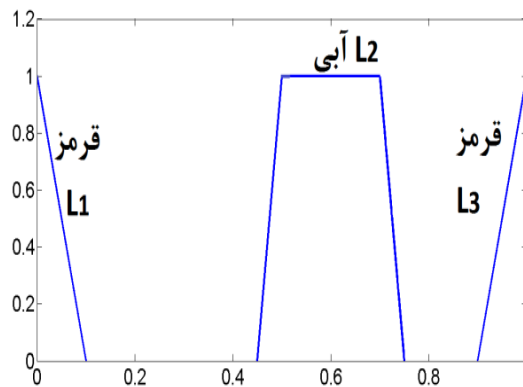


شکل ۴-۹: چند تصویر که به خوبی تقسیم‌بندی نشده است.

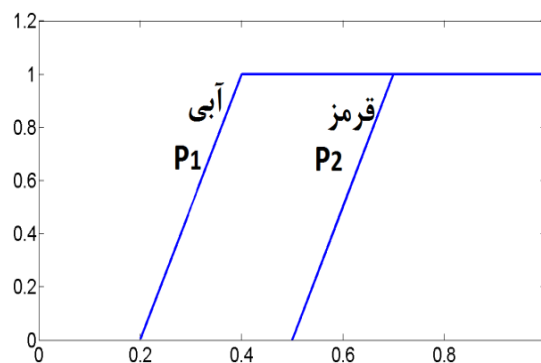
۴-۳-۱ استفاده از منطق فازی برای آشکارسازی علامت

در روش معمولی برای استخراج مناطق با رنگ قرمز و آبی آستانه پارامتر فام برای رنگ قرمز در دو بازه $0 < H < 0.07$ و $0.9 < H < 1$ و برای رنگ آبی $0.5 < H < 0.7$ و آستانه اشباع برای رنگ قرمز

در محدوده‌ی $0.5 < S < 1$ و برای رنگ آبی $0.4 < S < 0.95$ است. حال در این قسمت از منطق فازی برای استخراج مناطق با رنگ آبی و قرمز استفاده می‌کنیم. بدین منظور ابتدا تابع عضویت پارامتر فام را در شکل (۴-۱۰) و تابع عضویت پارامتر اشباع را در شکل (۴-۱۱) تعریف می‌کنیم.



شکل ۴-۱۰: تابع عضویت پارامتر فام



شکل ۴-۱۱: تابع عضویت پارامتر اشباع

سیستم فازی ما دارای دو ورودی است. میانگین مراکز را همانند جدول (۴-۱) به صورت ترکیب خطی از ورودی‌ها به دست می‌آوریم، در نتیجه سیستم فازی بوجود آمده از نوع سوگنو خواهد بود. برای هر پیکسل دو ورودی x_1 , x_2 داریم که مقدار پیکسل خروجی را توسط فرمول (۴-۳) بدست می‌آوریم [۳۱].

$$f(x_1, x_2) = \frac{\sum_{l=1}^M y^l (\prod_{i=1}^n \mu_{A_i^l}(x_i))}{\sum_{l=1}^M (\prod_{i=1}^n \mu_{A_i^l}(x_i))} \quad (۴-۳)$$

n تعداد ورودی‌ها است. که دو ورودی $x_1 = h$ و $x_2 = s$ را داریم و M تعداد قانون‌ها است. و y^L ها میانگین مراکز هستند. $\mu_{A_1^l}(x_1)$ تابع عضویت پارامتر فام و $\mu_{A_2^l}(x_2)$ تابع عضویت پارامتر اشباع است.

جدول ۴-۱: قانون‌ها و محاسبه میانگین مراکز.

		y'
$s \quad h$		
قانون‌ها	$P_1 \quad L_1$	$P_1 + L_1$
	$P_1 \quad L_2$	$P_1 + L_2$
	$P_1 \quad L_3$	$P_1 + L_3$
	$P_2 \quad L_1$	$P_2 + L_1$
	$P_2 \quad L_2$	$P_2 + L_2$
	$P_2 \quad L_3$	$P_2 + L_3$

با توجه به این که تابع عضویت پارامتر فام دارای سه زیرتابع است (شکل ۴-۱۰) و تابع عضویت پارامتر اشباع دارای دو زیرتابع است (شکل ۴-۱۱) بنابراین ۶ قانون خواهیم داشت، این ۶ قانون در جدول (۴-۱) نشان داده شده است. به عنوان مثال سطر اول این جدول بیانگر این است، که وقتی $0 \leq h \leq 0.1$ و $0.2 \leq s \leq 1$ است، آن گاه $l=1$ یعنی قانون اول را داریم و میانگین مرکز یعنی y^L مجموع این دو مقدار خواهد بود.

در شکل (۴-۹) مناطق با رنگ قرمز و آبی را به روش معمولی استخراج کردیم. همانطور که در این شکل مشاهده می‌شود به یک پیکسل که مقدار پارامتر فام و اشباع آن در بازه مورد نظر ما است مقدار ۲۵۵ اختصاص می‌دهیم و به یک پیکسل که مقادیر فام و اشباع آن خارج از بازه مورد نظر است مقدار صفر می‌دهیم. اما در شکل (۴-۱۲) استخراج مناطق با رنگ قرمز و آبی را با استفاده از منطق فازی انجام داده‌ایم و مشکل روش معمولی را برطرف کرده‌ایم، به این صورت که به ازای هر پیکسل دو ورودی فام و اشباع وارد سیستم فازی می‌شود و پیکسل خروجی مطابق فرمول (۴-۱) بر اساس

درجه عضویتی که به رنگ آبی یا قرمز دارد مقداری بین ۰ تا ۲۵۵ خواهد گرفت. که برای رسیدن به تصویر باینری از روی تصویری که خروجی سیستم فازی بوده است، این تصویر را با آستانه ۱۲۵ باینری می‌کنیم.

نمونه‌ای از تصاویر به همراه علامت استخراج شده آن در شکل (۴-۱۳) قسمت الف نشان داده شده است. نمونه‌ای از تصاویر که سیستم قادر به استخراج علامت نبوده در شکل (۴-۱۳) قسمت ب نشان داده شده است. دلایل عمده عدم استخراج علامت از میان تصویر، نور کم، دوری علامت از دوربین و خرابی و کامل نبودن علامت است.

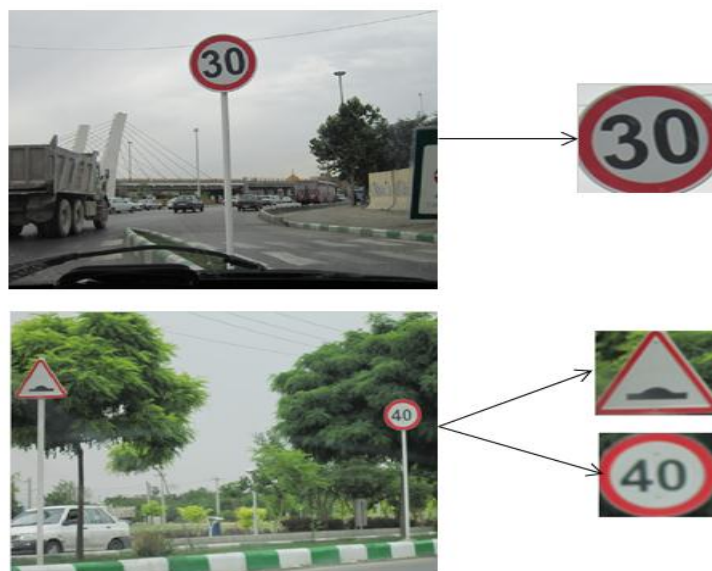
همان ۳۳۰ علامتی که در قسمت قبل مورد آزمایش قرار دادیم و به دقت ۹۳.۹۳ درصد رسیدیم، را با استفاده از منطق فازی مورد آزمایش قرار دادیم که به دقت ۹۷.۸۷ درصد رسیدیم. از این نتیجه به قدرتمند بودن روش فازی پی می‌بریم. به علت عدم وجود پایگاه داده مشترک، نتایج بخش آشکارسازی را نمی‌توان به خوبی با سایر کارهای انجام شده مقایسه کرد. اما به مقایسه این روش آشکارسازی با سایر روش‌های ارائه شده، بر روی پایگاه داده خودمان می‌پردازیم و نتایج را مقایسه می‌کنیم، این مقایسه در جدول (۴-۲) مشاهده می‌شود.



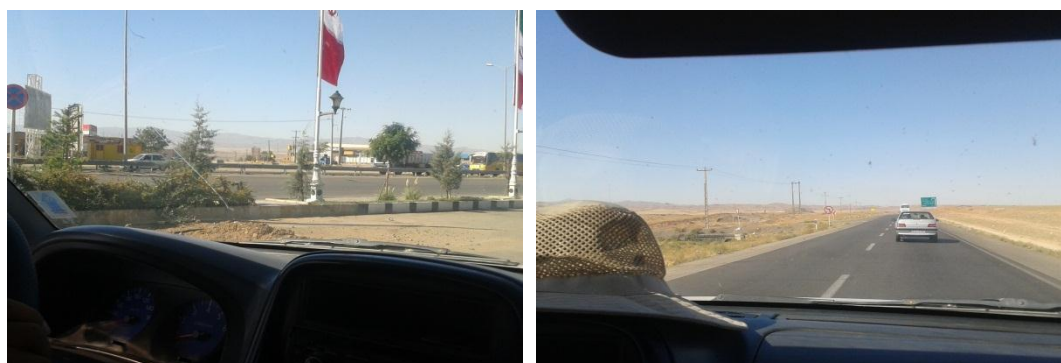
شکل ۴-۱۲: استخراج مناطق با رنگ قرمز و آبی با استفاده از منطق فازی و رفع مشکل شکل (۴-۹).

جدول ۴-۲: مقایسه نتایج آشکارسازی

روش آشکارسازی	تعداد تصاویر	دقت به دست آمده
روش پیشنهادی (معمولی)	۳۳۰	٪۹۳.۹۳
روش پیشنهادی (فازی)	۳۳۰	٪۹۷.۸۷
مرجع [۳۴]	۳۳۰	٪۹۳.۰۳
مرجع [۳۰]	۳۳۰	٪۹۵.۱۵



(الف)



(ب)

شکل ۴-۱۳: قسمت الف شامل دو تصویر است که علائم درون آن استخراج شده است. و شکل ب که به علت دوری

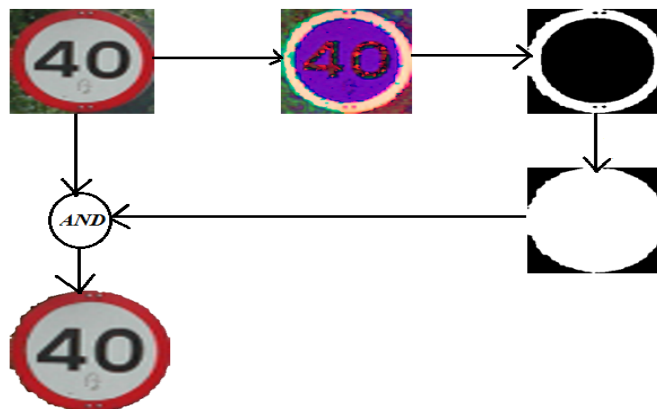
علامت از دوربین و نور کم، سیستم قادر به استخراج علامت از درون تصویر نیست.

۴-۳-۲ آشکارسازی علامت به صورت بلادرنگ با استفاده از C++

در حالت واقعی، یک سیستم آشکارسازی و شناسایی علائم راهنمایی و رانندگی بر روی وسیله نقلیه نصب می‌شود، بنابراین باید با سرعت بالایی علائم را آشکارسازی و شناسایی کند. برای این که الگوریتم ما به صورت بلادرنگ علامت را از میان اشیاء مختلف موجود در تصویر پیدا کند، برنامه اولیه را که با استفاده از نرم‌افزار متلب نوشتیم، با استفاده از نرم‌افزار Microsoft Visual Studio 2010 و کتابخانه Open CV به زبان C++ تبدیل کردیم. برنامه نوشته شده را روی یک کامپیوتر با مشخصات Core 2 CPU, 1.2GHz and 2GB RAM آزمایش کردیم، که به طور میانگین، به زمان آشکارسازی ۱۷۰ میلی ثانیه برای هر تصویر رسیدیم.

۴-۴ الگوریتم پیشنهادی برای شناسایی نوع دقیق علامت

برای شناسایی نوع دقیق علامت ابتدا می‌بایست پس‌زمینه علامت آشکار شده را حذف کنیم. برای حذف پس‌زمینه علامت آشکار شده ابتدا آن را به فضای رنگ HSV تبدیل می‌کنیم و با استفاده از منطق فازی همانطور که در بخش (۴-۳-۱) توضیح داده شد، مناطق با رنگ قرمز و آبی را استخراج می‌کنیم. (شبه مرحله آشکارسازی) سپس حفره موجود در تصویر را پر کرده، تصویر حاصل را با تصویر ورودی AND می‌کنیم در نتیجه این عملیات پس‌زمینه علامت حذف خواهد شد. این فرآیند در شکل (۴-۱۴) نشان داده شده است.

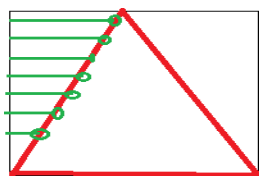


شکل ۴-۱۴: فرآیند حذف پس‌زمینه علامت آشکار شده.

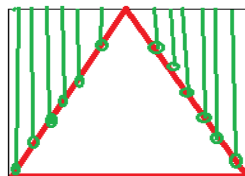
۴-۴-۱ استفاده از طبقه‌بند ماشین بردار پشتیبان برای شناسایی دقیق علامت

به دلیل این که تعداد کلاس‌های علائم راهنمایی و رانندگی زیاد است، به منظور بالا رفتن سرعت الگوریتم، قبل از این که به شناسایی نوع دقیق علامت بپردازیم، قالب کلی آن را مشخص می‌کنیم. قالب کلی علامت می‌تواند مثلث، دایره و مستطیل باشد. مسئله ما در مرحله اول یعنی شناسایی قالب کلی علامت، یک مسئله شناسایی ۳ کلاسه می‌شود. طبقه‌بندی که برای شناسایی شکل کلی علامت استفاده می‌کنیم ماشین بردار پشتیبان است.

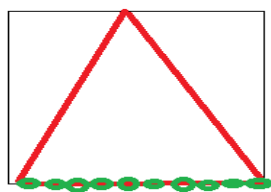
ماشین بردار پشتیبان، یک طبقه‌بند دو کلاسه است. برای یک مساله دو کلاسه داده‌های آموزشی به صورت $\{x_i, y_i\}$ بر چسب می‌خورند که $i=1, 2, \dots, l$ و $y_i \in \{-1, 1\}$ و $x_i \in \mathbb{R}^d$ است. y_i برای یک کلاس ۱ است و برای کلاس دیگر ۱- است. d بعد بردار ویژگی است و l تعداد نمونه‌های آموزش است. بردارهای ویژگی، یعنی x_i ها، فاصله لبه‌های مربع احاطه‌کننده علامت تا نوار علامت است [۱۰]. این نحوه استخراج ویژگی در شکل (۴-۱۵) قسمت الف تا ج برای شکل مثلث نشان داده شده است.



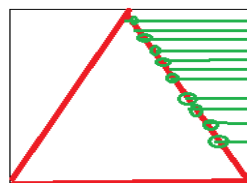
(ب)



(الف)



(د)



(ج)

شکل ۴-۱۵: نحوه استخراج ویژگی برای تشخیص شکل کلی علامت. قسمت الف از سمت بالا شروع به حرکت کرده‌ایم تا به نوار علامت برسیم و در قسمت‌های ب تا د به ترتیب از چپ، راست و پایین به سمت نوار علامت حرکت شده است.

یا به عنوان مثال همانطور که در شکل (۴-۱۶) مشاهده می‌کنیم از چهار طرف مرز مربع حرکت می‌کنیم تا به اولین پیکسل سفید برسیم، تعداد پیکسل‌هایی که جلو می‌رویم تا به اولین پیکسل

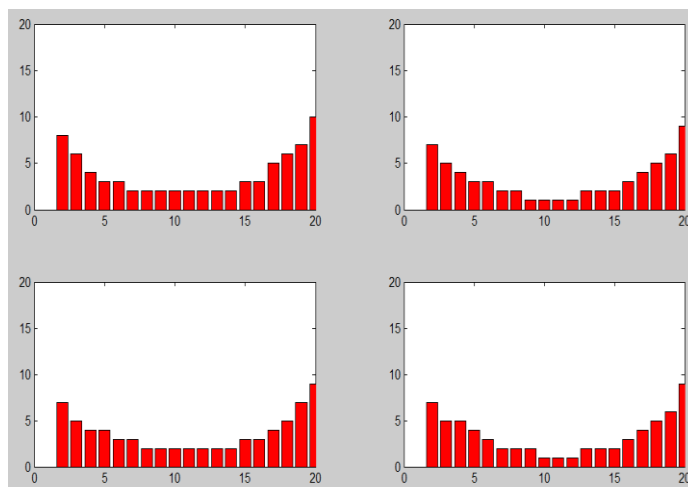
سفید برسیم یک ویژگی از بردار ویژگی محسوب می‌شود. چون ابعاد تصویر 20×20 است، بنابراین ابعاد هر بردار ویژگی 20×1 است، که با قرار دادن این چهار بردار ویژگی در کنار هم یک بردار ویژگی 80×1 خواهیم داشت.

به دلیل این که برای توصیف هر شکل چهار بردار ویژگی استخراج می‌شود این روش کاملاً به خرابی قسمتی از علامت مقاوم است. هر داده جدید درست طبقه‌بندی شده است به شرطی که در رابطه زیر صدق کند.

$$y_i(x_i w^T + b) - 1 \geq 0 \quad \forall i \quad (4-4)$$



(الف)



(ب)

شکل ۴-۱۶: قسمت الف نوار اطراف علامت دایره‌ای را نشان می‌دهد. و قسمت ب چهار بردار ویژگی استخراج شده از شکل الف را نشان می‌دهد.

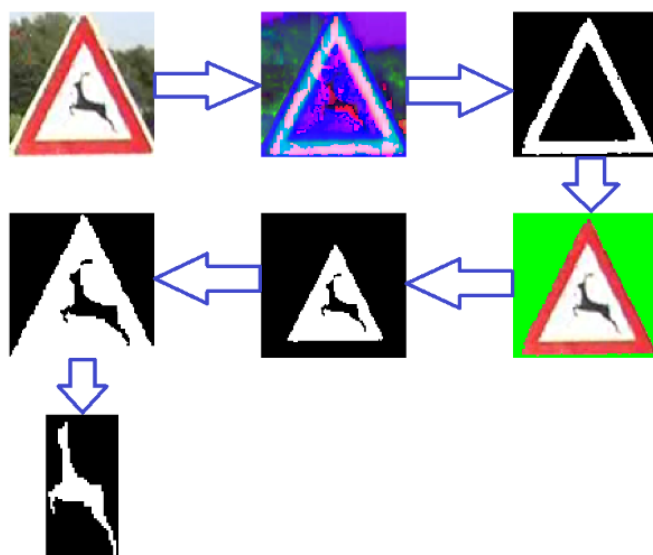
w بردار عمود بر مرز مجزا ساز است. نحوه به دست آوردن w و b در فصل قبل قسمت ماشین بردار پشتیبان توضیح داده شد. پس از به دست آوردن این دو پارامتر، آن گاه با تعریف تابع علامت به صورت زیر به راحتی می‌توان یک داده جدید را طبقه بندی کرد.

$$f(x) = \text{sgn}(x \cdot w^T + b) \quad (5-4)$$

برای داده جدید اگر $x \cdot w^T + b > 0$ باشد آن گاه مقدار تابع برابر با ۱ خواهد بود و این داده مربوط به کلاس یک خواهد بود. و اگر $x \cdot w^T + b < 0$ آن گاه مقدار تابع $f(x)$ برابر با -۱ خواهد بود و داده مربوط

به کلاس ۲ خواهد بود. با توجه به اینکه ما در اینجا سه کلاس داریم از تکنیک یک کلاس در مقابل بقیه استفاده کرده ایم.

بعد از این که شکل کلی علامت را تشخیص دادیم، یعنی مثلث، دایره یا مستطیل. نوبت به شناسایی نوع دقیق آن می رسد. برای دسته بندی دقیق علامت ابتدا ناحیه داخلی علامت را از درون تصویر استخراج می کنیم. برای استخراج ناحیه داخلی علامت ابتدا تصویر ورودی را به فضای رنگ HSV تبدیل می کنیم، پس زمینه آن را حذف می کنیم و ناحیه داخلی آن را استخراج می کنیم. این فرآیند در شکل (۴-۱۷) نشان داده شده است.



شکل ۴-۱۷: مراحل به دست آوردن قسمت درونی علامت

بعد از این که قسمت درونی علامت را استخراج کردیم نوبت به استخراج ویژگی می رسد. استخراج ویژگی در این قسمت دقیقاً مانند مرحله قبل انجام می شود یعنی از گوشه های مستطیل احاطه کننده قسمت داخلی علامت حرکت می کنیم تا به علامت برسیم. این موضوع در شکل (۴-۱۸) نشان داده

شده است. در هر یک از چهار قسمت نشان داده شده در شکل (۴-۱۸) یک بردار ویژگی استخراج می‌کنیم، که با قرار دادن این چهار بردار در کنار هم یک بردار ویژگی بدست می‌آید که معرف شکل است.



شکل ۴-۱۸: نحوه استخراج ویژگی برای شناسایی نوع دقیق علامت

به دلیل این که ماشین بردار پشتیبان قادر به جداسازی دو کلاس است و مساله ما چند کلاسه است، برای طبقه‌بندی یک داده جدید، یک کلاس را در مقابل بقیه کلاس‌ها در نظر می‌گیریم، یعنی بررسی می‌کنیم که داده جدید متعلق به کلاس یک است یا بقیه کلاس‌ها، اگر متعلق به بقیه کلاس‌ها بود، کلاس یک را کنار می‌گذاریم و بررسی می‌کنیم که مربوط به کلاس دو است یا بقیه کلاس‌ها، این روند را آنقدر ادامه می‌دهیم تا داده طبقه‌بندی شود. مزیت استفاده از ماشین بردار پشتیبان این است که اولاً بهترین مرز ممکن را بین دو کلاس ایجاد می‌کند و ثانیاً نسبت به طبقه‌بندیهایی مانند شبکه عصبی نیازمند داده آموزشی کم‌تری است. اما عیبی که این طبقه‌بند دارد این است که قادر به جداسازی دو کلاس است که ما در این این‌جا با در نظر گرفتن یک کلاس در مقابل بقیه این مشکل را رفع کرده‌ایم [۲۵].

۴-۴-۲ نتایج تجربی بدست آمده از شناسایی دقیق با استفاده از ماشین بردار

پشتیبان

الگوریتمی که در این قسمت از پایان‌نامه ارائه شد، مستقل از انتقال، مقیاس و تا حدودی مستقل از چرخش است. در این کار به شناسایی ۴۰ کلاس از علائم راهنمایی و رانندگی پرداختیم. پایگاه داده مورد استفاده در این بخش برگرفته از [۲۶] است، که توسط دوربین نصب شده بر روی خودرو از

فواصل مختلف از علائم تصویر برداری شده است. تعداد نمونه‌های آموزش برای هر کلاس ۲۱ عدد است، که در مجموع از ۸۴۰ علامت برای آموزش ماشین بردار پشتیبان استفاده شده است. و از ۴۱۶ علامت برای آزمایش عملکرد این سیستم استفاده شده است، که به منظور ارزیابی توانایی الگوریتم در شرایط نویزی، برخی از تصاویر این مجموعه با نویز گوسی آغشته شده‌اند. دقت شناسایی سیستم ما برای این پایگاه داده ۹۱.۱ درصد است. چند نمونه از علائمی که درست شناسایی شده است را در شکل (۴-۱۹) مشاهده می‌کنیم و چند نمونه از علائمی که اشتباه شناسایی شده است را در شکل (۴-۲۰) مشاهده می‌کنیم.

به دلیل این که در این قسمت از رنگ علائم برای حذف پس‌زمینه استفاده می‌شود، اگر میزان نور علائم از حدی کم‌تر باشد همانطور که در شکل (۴-۲۰) قسمت الف و ب مشاهده می‌شود، سیستم قادر به شناسایی علامت نیست یعنی سیستم نمی‌تواند قسمت قرمز را در تصویر پیدا کند و قسمت درونی آن را استخراج کند. و همچنین اگر درصد نویز اضافه شده به علائم همانطور که در شکل (۴-۲۰) قسمت ج و د نشان داده شده است، زیاد باشد، مشکل قسمت الف و ب تکرار می‌شود و سیستم قادر به شناسایی آن‌ها نخواهد بود.



شکل ۴-۱۹: نمونه‌ای از علائم که سیستم قادر به شناسایی آن‌ها بوده است. ردیف دوم علائمی است که با نویز گوسی آغشته شده‌اند.



(د)

(ج)

(ب)

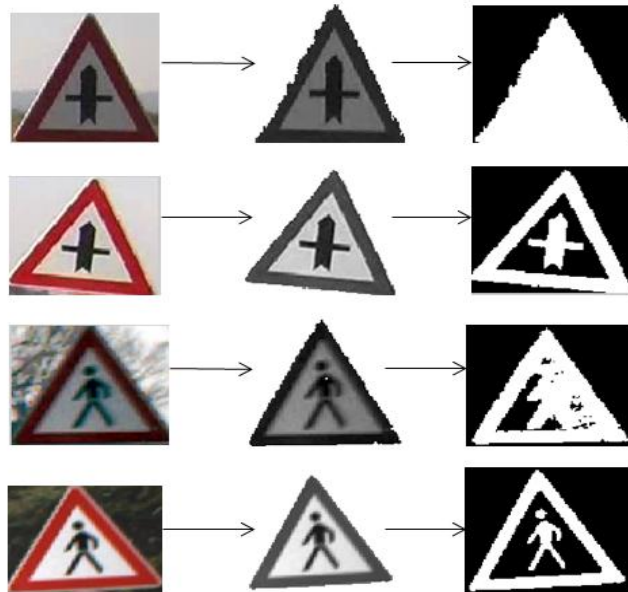
(الف)

شکل ۴-۲۰: نمونه‌ای از علائم که سیستم قادر به شناسایی آن‌ها نبوده است.

با توجه به عدم وجود پایگاه داده استاندارد و در دسترس، نتایج بدست آمده را نمی‌توان به خوبی با محدود کارهایی که در داخل کشور انجام شده است مقایسه کرد. در [۲۷] با استفاده از شبکه عصبی و در [۱۶] با استفاده از طبقه بند بیزی عمل شناسایی بر روی تصاویر اسکن شده از کتاب راهنمایی و رانندگی انجام شده است. در [۷] عمل شناسایی را با استفاده از روش انطباق الگو بر روی تصاویری شبیه به تصاویر مورد استفاده در این پایان‌نامه انجام داده است که به دقت ۸۴.۵ درصد رسیده است. این در حالی است که ما با استفاده از ماشین بردار پشتیبان عمل شناسایی را بر روی تصاویر نویزی انجام داده‌ایم و به دقت ۹۱.۱ درصد رسیده‌ایم. با استفاده از این روش عمل شناسایی را بر روی پایگاه داده خودمان انجام دادیم که به دقت ۹۲ درصد رسیدیم.

اما مشکلی که این روش دارد، این است که استخراج ویژگی از علامت مبتنی بر تصویر باینری است. این روش استخراج ویژگی را در [۳۵] برای استخراج ویژگی از علامت به کار بردیم.

همچنین اگر از روش استخراج ویژگی ارائه شده توسط [۷] که روش تقاطع خطوط مرور نام دارد یا روش استخراج ویژگی ارائه شده توسط [۲۹] که از پیکسل‌های تصویر باینری به عنوان ویژگی ورودی به شبکه عصبی استفاده کرده است، استفاده کنیم، پس از حذف پس زمینه علامت و خاکستری کردن آن، از دو تصویر در یک کلاس که در شرایط نوری مختلفی قرار دارند، همانطور که در شکل (۴-۲۱) مشاهده می‌شود بردارهای ویژگی کاملاً متفاوتی استخراج خواهد شد.



شکل ۴-۲۱: نا مناسب بودن روش‌های استخراج ویژگی که مبتنی بر تصویر باینری است، در این تصویر دیده می‌شود. برای باینری کردن تصویر خاکستری، علاوه بر روش معمول باینری کردن، روش آستانه گیری آتسو که بهترین آستانه ممکن را می‌دهد استفاده شد، ولی باز هم نتیجه تفاوتی نکرد. در نتیجه این روش‌ها، روشی مناسب برای استخراج ویژگی نیستند. زیرا علائم راهنمایی و رانندگی ممکن است در شرایط نوری مختلفی قرار داشته باشند. حال برای رفع این مشکل در مرحله بعد از روش استخراج ویژگی هیستوگرام گرادیان که مبتنی بر لبه‌های تصویر است، استفاده می‌کنیم. همچنین به علت این‌که تعدادی کافی علامت از هر کلاس داریم از شبکه عصبی برای شناسایی استفاده می‌کنیم.

۴-۴-۳ استفاده از شبکه عصبی MLP برای شناسایی دقیق

در این قسمت تصویری که پس زمینه آن حذف شده است، را به یک تصویر خاکستری تبدیل می‌کنیم و به روش هیستوگرام گرادیان از آن استخراج ویژگی می‌کنیم. فرآیند استخراج ویژگی به روش هیستوگرام گرادیان به صورت زیر است:

۱- ابتدا تصویر ورودی را به ۱۶ بلوک مربعی غیر هم‌پوشان تقسیم می‌کنیم.

۲- در هر پیکسل از هر بلوک با استفاده از عملگر سوپل گرادیان را در دو جهت x و y

محاسبه می‌کنیم. این محاسبات در فرمول‌های (۴-۶) تا (۴-۱۱) نشان داده شده است.

$$g_x(x, y) = f(x + 1, y - 1) + \quad (۶-۴)$$

$$2f(x + 1, y) + f(x + 1, y + 1) - f(x - 1, y - 1) - 2f(x - 1, y) - f(x - 1, y + 1)$$

$$g_y(x, y) = f(x - 1, y + 1) + \quad (۷-۴)$$

$$2f(x, y + 1) + f(x + 1, y + 1) - f(x - 1, y - 1) - 2f(x, y - 1) - f(x + 1, y - 1)$$

$$g(x, y) = [g_x, g_y]^T \quad (۸-۴)$$

$f(x, y)$ مقدار پیکسل در نقطه (x, y) است.

۳- برای هر پیکسل فاز و اندازه گرادیان را با استفاده از فرمول‌های (۹-۴) و (۱۰-۴) بدست می‌آوریم.

$$P(x, y) = \tan^{-1}\left(\frac{g_y}{g_x}\right) \quad (۹-۴)$$

$$A(x, y) = \sqrt{g_x^2 + g_y^2} \quad (۱۰-۴)$$

۴- زوایای هر بلوک را به ۱۶ بین کوانتیزه می‌کنیم. سپس برای هر بلوک ۱۶ ویژگی مطابق با ۱۶ فاز با استفاده از فرمول (۱۱-۴) به دست می‌آوریم.

$$F_\theta = \sum_{x_\theta, y_\theta} A(x, y) \quad (۱۱-۴)$$

F_θ ویژگی مربوط به زاویه θ است. (x_θ, y_θ) مختصات با زاویه θ است. با توجه با این که هر تصویر را به ۱۶ بلوک تقسیم کرده‌ایم و در هر بلوک ۱۶ ویژگی استخراج شده است، بنابراین به ازای هر تصویر $۱۶ \times ۱۶ = ۲۵۶$ ویژگی استخراج می‌شود.

مزیت استفاده از روش هیستوگرام گرادیان که در این قسمت برای استخراج ویژگی استفاده شده است این است که اولاً برای انجام محاسبات سریع است و ثانیاً ویژگی‌های مبتنی بر گرادیان یا لبه نسبت به سایر ویژگی‌ها، ویژگی‌های قدرتمندی هستند، زیرا در شرایط نوری مختلف اطلاعات لبه باقی خواهد ماند. بنابراین مشکل قسمت قبل در استخراج ویژگی حل می‌شود.

پایگاه داده‌ای که در این بخش برای آموزش و آزمایش شبکه عصبی استفاده می‌شود برگرفته از مجموعه داده‌ای است که توسط [۲۶] بوسیله دوربین نصب شده بر روی خودرو از فواصل مختلف از علائم تصویر برداری کرده‌اند. این مجموعه شامل ۱۲۸۰ علامت در ۴۰ کلاس مختلف است. ۶۷ درصد این مجموعه یعنی ۸۶۴ علامت را به آموزش شبکه اختصاص می‌دهیم. و از ۳۳ درصد باقی مانده یعنی ۴۱۶ علامت برای آزمایش شبکه استفاده می‌کنیم.

از هر یک از ۸۶۴ علامتی که در مجموعه آزمایش قرار دارند با استفاده از روش هیستوگرام گرادیان استخراج ویژگی می‌کنیم. در نتیجه ۸۶۴ بردار ویژگی خواهیم داشت که هر یک از این بردارها دارای ۲۵۶ ویژگی است. طبقه بند مورد استفاده، شبکه عصبی MLP است.

تعداد نرون‌های لایه ورودی این شبکه به همراه بایاس ۲۵۷ نرون است. تعداد نرون‌های لایه مخفی ۲۶ عدد و تعداد نرون‌های لایه خروجی به علت این که ۴۰ کلاس از علائم راهنمایی و رانندگی داریم، ۴۰ نرون است. تابع فعال‌سازی نرون‌ها در لایه مخفی و لایه خروجی تابع سیگموید است. این تابع را در رابطه (۱۲-۴) می‌بینیم.

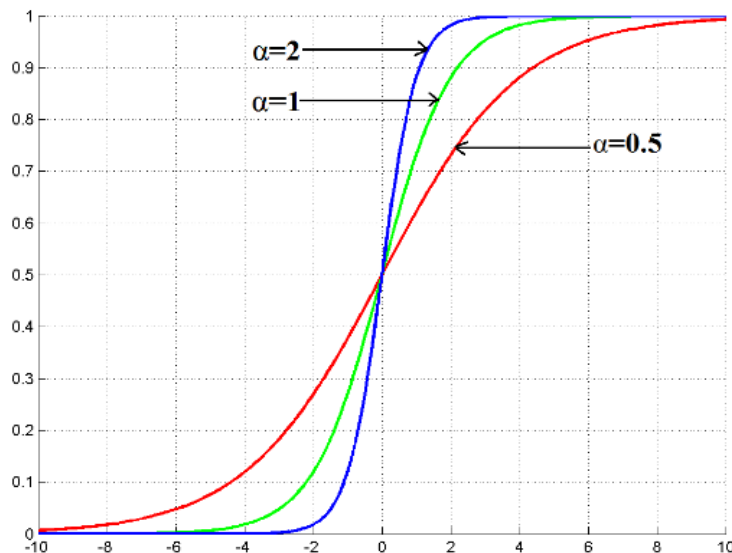
$$\varphi(v) = \frac{1}{1 + \exp(-\alpha v)} \quad (12-4)$$

با مدنظر قرار دادن نرون j ام در لایه مخفی، v_j مجموع ورودی‌هایی است که به این نرون وارد می‌شود (رابطه (۱۳-۴)) و $\varphi(v_j)$ مقدار خروجی این نرون است. x_i مقدار نرون i ام لایه ورودی است. w_{ji} وزن بین نرون j ام لایه مخفی و نرون i ام لایه ورودی است. محاسبه ورودی و خروجی برای نرون‌های لایه خروجی به همین صورت است. مقدار پارامتر α در محاسبات ما یک است.

$$v_j = \sum_{i=0}^{256} w_{ji} x_i \quad (13-4)$$

نمودار تابع سیگموید به ازای چند مقدار مختلف α در شکل (۲۲-۴) نشان داده شده است. همانطور که در این شکل مشاهده می‌شود با افزایش پارامتر α شکل این نمودار به تابع پله نزدیک می‌شود. روشی که ما توسط آن اصلاح وزن را انجام می‌دهیم روش پس انتشار خطا است. روش کار به این

صورت است که در مرحله آموزش یک نمونه آموزشی را در ورودی شبکه قرار می‌دهیم و سپس خطا را از تفاضل خروجی به دست آمده با خروجی اصلی محاسبه می‌کنیم. این خطا را به سمت ورودی انتشار داده و وزن‌ها از لایه ورودی به لایه مخفی و از لایه مخفی به لایه خروجی را اصلاح می‌کنیم.



شکل ۴-۲۲: نمودار تابع فعال سازی به ازاء سه مقدار مختلف برای پارامتر α

تمام نمونه‌های آموزش یعنی ۸۶۴ نمونه را در ۱۵۰ دوره آموزشی در ورودی شبکه قرار می‌دهیم، شبکه بعد از دیدن هر نمونه اصلاح وزن را انجام می‌دهد. در پایان هر دوره آموزش خطای کل را توسط رابطه‌ی (۴-۱۴) به دست می‌آوریم.

$$E = \frac{1}{2N} \sum_n \sum_j \left(d_j(n) - o_j(n) \right)^2 \quad (۴-۱۴)$$

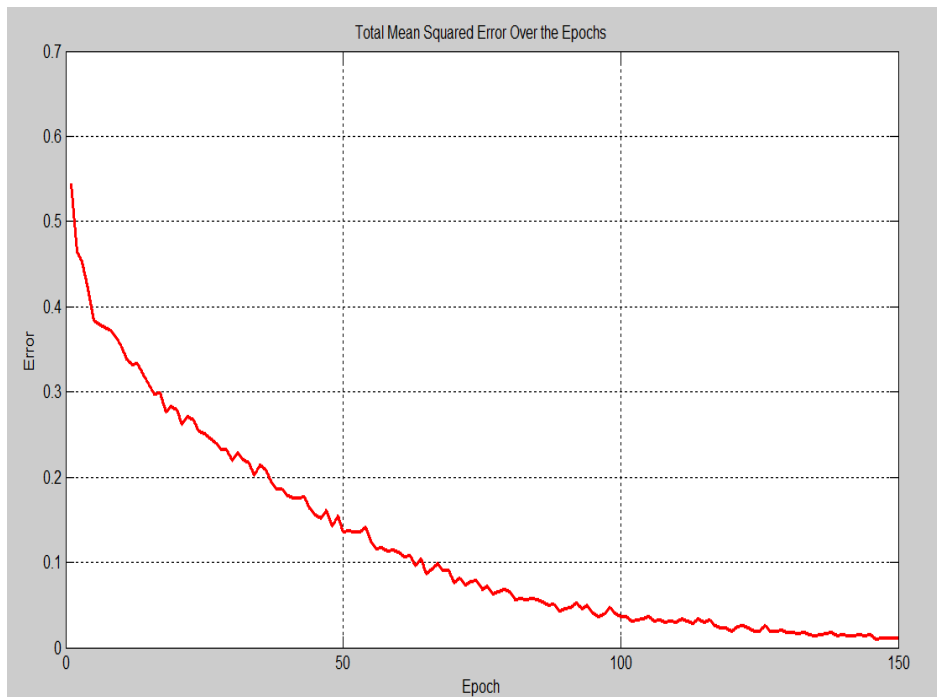
نمودار این خطا طی ۱۵۰ دوره آموزشی در شکل (۴-۲۳) نشان داده شده است. در این نمودار محور افقی نشان دهنده‌ی تعداد دوره آموزشی و محور عمودی نشان دهنده میزان خطا است.

۴-۴-۴ نتایج تجربی بدست آمده از شناسایی دقیق با استفاده از شبکه عصبی

MLP

مجموعه داده‌ای که برای مرحله شناسایی دقیق استفاده کرده‌ایم شامل ۱۲۸۰ علامت در ۴۰ کلاس مختلف است. این کلاس‌ها در شکل (۴-۲۴) نشان داده شده است. از ۸۶۴ علامت یعنی ۶۷٪ داده‌ها

برای آموزش شبکه عصبی و از ۳۳٪ باقی مانده یعنی ۴۱۶ علامت برای آزمایش شبکه استفاده کردیم. دقت به دست آمده برای نمونه‌های آموزش ۹۹ درصد و برای نمونه‌های آزمایش ۹۳.۸۶ درصد است. همانند مرحله قبل، برخی از نمونه‌های آزمایش با نویز گوسی آغشته شده‌اند.



شکل ۴-۲۳: نمودار خطای طی ۱۵۰ دوره آموزش.

نمونه‌ای از علائم که سیستم قادر به شناسایی دقیق آن‌ها بوده در شکل (۴-۲۵) قسمت الف و چند نمونه از علائمی که سیستم قادر به شناسایی آن‌ها نبوده در شکل (۴-۲۵) قسمت ب نشان داده شده است.

برای تصاویر خیلی روشن یا خیلی تاریک می‌توان قبل از این که عمل آشکارسازی یا حذف پس‌زمینه علامت را انجام دهیم، ابتدا با استفاده از تکنیک متعادل‌سازی هیستوگرام روشنایی تصویر را متعادل کنیم، این کار به این دلیل که نواحی بسیار روشن و تاریک تصویر را به حالت نوری متعادل تبدیل می‌کند به بهتر آشکار شدن علامت کمک می‌کند. در تصویر RGB بر روی هر یک از سه کانال رنگ عملیات متعادل‌سازی هیستوگرام به صورت مجزا انجام می‌شود. این فرآیند در شکل (۴-۲۶) نشان داده شده است.



شکل ۴-۲۴: ۴۰ کلاس از علائم که ما عملیات شناسایی را بر روی آن‌ها انجام داده‌ایم.

با توجه به عدم وجود پایگاه داده استاندارد و در دسترس، نتایج بدست آمده را نمی‌توان به خوبی با معبود کارهایی که در داخل کشور انجام شده است مقایسه کرد. [۲۷] با استفاده از شبکه عصبی و [۱۶] با استفاده از طبقه بند بیزی عمل شناسایی را بر روی تصاویر اسکن شده از کتاب راهنمایی و

رانندگی انجام داده‌اند، در حالی که کار انجام شده در این پایان‌نامه بر روی تصاویر واقعی که از دوربین نصب شده بر روی وسیله نقلیه به‌دست آمده، انجام شده است.



(الف)

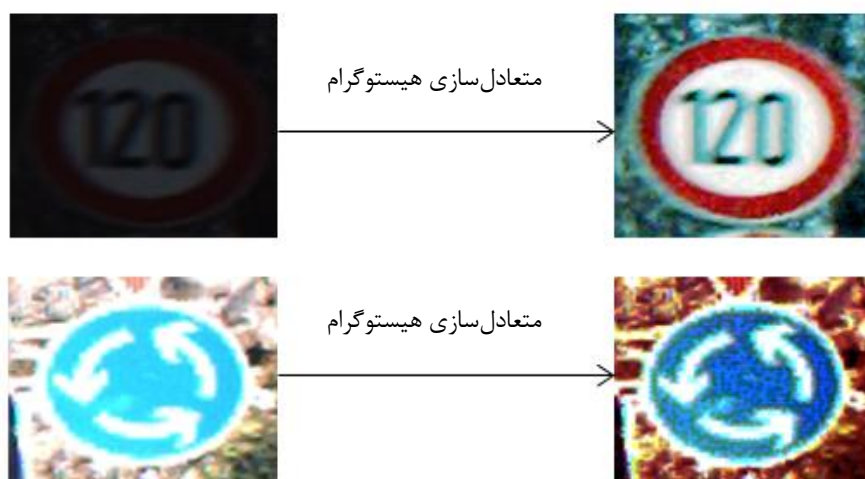


(ب)

شکل ۴-۲۵: قسمت الف تعدادی از علائم که محتوای آن توسط سیستم شناسایی شده است و قسمت ب تعدادی

علامت که سیستم قادر به شناسایی آن‌ها نبوده است.

مراجع [۲۷]، [۱۶]، [۲۸] و [۷] عمل شناسایی را با فرض این‌که علامت از درون تصویر استخراج شده است انجام داده‌اند. اما روش ارائه شده در این پایان‌نامه ابتدا علامت را از تصویر استخراج کرده، سپس به شناسایی آن می‌پردازد. مرجع [۲۸] عمل شناسایی را بر روی تصاویری شبیه به تصاویر مورد استفاده در این پایان‌نامه انجام داده است، که به دقت ۹۱.۰۴ درصد رسیده است.



شکل ۴-۲۶: متعادل سازی هیستوگرام به منظور بهبود کیفیت تصاویر با شدت روشنایی زیاد یا شدت روشنایی کم. مرجع [۷] با استفاده از روش انطباق الگو به دقت ۸۴.۵ درصد رسیده است. این در حالی است که مجموعه داده استفاده شده در این پایان نامه همانطور که چند نمونه از آن در شکل (۴-۲۴) قسمت الف مشاهده می شود شامل تصاویر تار، در شرایط نور کم و نور زیاد، بی کیفیت و نویزی است که سیستم ما قادر به شناسایی ۹۳.۸۶ درصد از ۴۱۶ علامت مورد آزمایش است. همچنین علائم پایگاه داده خودمان را توسط این روش شناسایی مورد آزمایش قرار دادیم که به دقت ۹۶ درصد رسیدیم.

۴-۴-۵ شناسایی علامت به صورت بلادرنگ با استفاده از C++

برای این که الگوریتم ما به صورت بلادرنگ عملیات شناسایی علامت را انجام دهد، مانند مرحله آشکارسازی، برنامه اولیه را که با استفاده از نرم افزار متلب نوشتیم، با استفاده از نرم افزار Visual Studio 2010 و کتابخانه Open CV به زبان C++ تبدیل کردیم، که بر روی یک کامپیوتر با مشخصات Core 2 CPU, 1.2GHz and 2GB RAM آزمایش کردیم، که به طور میانگین، به زمان شناسایی ۱۲۱ میلی ثانیه برای هر تصویر رسیدیم.

نتایج به دست آمده در این پایان نامه را به طور خلاصه با سایر کارهای انجام شده در جدول (۴-۳) مقایسه می کنیم.

جدول ۴-۳: مقایسه کار انجام شده با سایر روش‌ها.

مرحله آشکارسازی	میزان دقت	طبقه‌بند	پایگاه داده	روش
ندارد.	علائم بدون نويز: ۹۴٪ علائم با نويز گوسی: ۷۰٪	شبکه عصبی	علائم اسکن شده از کتاب راهنمایی و راهنمایی	مرجع [۲۷]
ندارد.	۹۱.۰۴٪	شبکه عصبی	۳۱۳ نمونه آموزش و ۱۱۴ نمونه آزمایش.	مرجع [۲۸]
دارد.	۸۵.۳۸٪	انطباق الگو با استفاده از فاصله Hausdorff	۳۳۷ علامت آزمایش	مرجع [۳۲]
دارد.	۹۰٪	انطباق الگو با استفاده از ضریب همبستگی	۲۰۴ علامت آزمایش	مرجع [۳۳]
ندارد.	۸۳٪	انطباق الگو	۳۰۳ علامت آزمایش	مرجع [۷]
ندارد.	علائم بدون نويز: ۹۶٪ علائم با نويز گوسی: ۸۰٪	طبقه‌بند بیزین	علائم اسکن شده از کتاب راهنمایی و راهنمایی	مرجع [۱۶]
دارد.	۹۱.۱٪	ماشین بردار پشتیبان	۸۴۰ داده آموزش و ۴۱۶ داده آزمایش (پایگاه داده مرجع ۲۶)	روش ارائه شده

روش ارائه شده	۸۶۴ داده آموزش و ۴۱۶ داده آزمایش (پایگاه داده مرجع ۲۶)	شبکه عصبی MLP	٪۹۳.۸۶	دارد.
روش ارائه شده	پایگاه داده خودمان	ماشین بردار پشتیبان	٪۹۲	دارد.
روش ارائه شده	پایگاه داده خودمان	شبکه عصبی MLP	٪۹۶	دارد.

فصل پنجم:

نتیجه گیری و پیشنهادات

۵-۱ نتیجه‌گیری

در این پایان‌نامه به شناسایی خودکار علائم راهنمایی و رانندگی پرداختیم. یک سیستم شناسایی علائم راهنمایی و رانندگی شامل دو مرحله آشکارسازی و شناسایی دقیق است. در مرحله آشکارسازی می‌بایست علامت راهنمایی و رانندگی از میان اشیاء مختلفی که در تصویر وجود دارد استخراج شود. که در این پایان‌نامه در این مرحله، از ویژگی رنگ و شکل هندسی علائم که نسبت به سایر اشیاء موجود در تصویر متمایز است، برای استخراج علامت استفاده شد. به منظور کاهش حساسیت نسبت به تغییرات نور ابتدا تصویر علامت را به فضای رنگ HSV تبدیل کردیم و ثانیاً از منطق فازی برای استخراج مناطق با رنگ آبی و قرمز استفاده کردیم. از مقایسه روش معمولی و روش فازی به قدرتمند بودن روش فازی برای استخراج علامت از درون تصویر پی بردیم.

به علت عدم وجود پایگاه داده‌ای که از علائم راهنمایی و رانندگی موجود در ایران تهیه شده باشد. اقدام به تهیه یک پایگاه داده از علائم راهنمایی و رانندگی موجود در ایران کردیم. این تصاویر از سطح خیابان‌ها و جاده‌ها از علائم گرفته شده است. ویژگی این پایگاه داده این است، که اولاً تصاویر از فواصل و موقعیت‌های مختلف از علامت گرفته شده است. و ثانیاً در شرایط نوری و پس‌زمینه‌های مختلفی قرار دارند. همچنین برخی از تصاویر این پایگاه داده از روی خودروی در حال حرکت تهیه شده است. این پایگاه داده شامل ۳۵۰ علامت راهنمایی و رانندگی در ۳۵ کلاس مختلف است.

پس از این‌که علامت را از درون تصویر استخراج کردیم قبل از شناسایی محتوای علامت، ابتدا پس‌زمینه علامت استخراج شده را حذف کرده، سپس برای شناسایی محتوای علامت از طبقه‌بند ماشین بردار پشتیبان و شبکه عصبی استفاده کردیم. در استفاده از ماشین بردار پشتیبان فرآیند شناسایی به این صورت انجام شد که پس از حذف پس‌زمینه علامت، ناحیه داخلی علامت را استخراج کرده و نوع دقیق علامت را توسط این طبقه‌بند شناسایی کردیم. نتایج بدست آمده یعنی نرخ شناسایی ۹۱.۱ درصد روی داده‌های آزمایش که شامل داده‌های نویزی و داده‌هایی در شرایط نوری مختلف است، نشان دهنده قدرتمندی این روش در شناسایی علائم راهنمایی و رانندگی است.

در استفاده از شبکه عصبی MLP پس از حذف پس‌زمینه علامت استخراج شده، با استفاده از روش هیستوگرام گرادیان استخراج ویژگی کردیم، این روش استخراج ویژگی از این روش استفاده شد که به خوبی اطلاعات جهتی را به ما می‌دهد و از طرفی یکی از مشخصه‌های خوب تصاویر علائم، همین تفاوت در اطلاعات جهتی است. سپس طبقه‌بند MLP را با استفاده از ۸۶۴ علامت آموزش دادیم و با استفاده از ۴۱۶ علامت تار، بی‌کیفیت و در شرایط نور کم و نور زیاد آزمایش کردیم که به دقت ۹۹ درصد بر روی داده‌های آموزش و ۹۳.۸۶ درصد بر روی داده‌های آزمایش رسیدیم. از نتایج به‌دست آمده به قدرتمندی روش ارائه شده در شناسایی و آشکارسازی علائم راهنمایی و رانندگی در صحنه‌های واقعی پی می‌بریم.

به منظور به دست آوردن الگوریتم بلادرنگ در آشکارسازی و شناسایی دقیق علامت، برنامه نوشته شده در نرم‌افزار متلب را به زبان C++ تبدیل کردیم، که به زمان پردازش ۱۷۰ میلی‌ثانیه در مرحله آشکارسازی و زمان پردازش ۱۲۱ میلی‌ثانیه در مرحله شناسایی دقیق رسیدیم.

۵-۲ پیشنهادات

- تکمیل کردن پایگاه داده، یعنی هم تعداد کلاس‌ها و هم تعداد علامت از هر کلاس افزایش داده شود.
- ارائه روشی که قادر به آشکارسازی و شناسایی علائم راهنمایی و رانندگی که دارای شکل هندسی لوزی و پنج‌ضلعی و رنگ زرد و سفید هستند، باشد.
- از روشهایی مانند تبدیل هاف که مبتنی بر رنگ علائم برای آشکارسازی نیست، استفاده شود. زیرا در شرایط نوری خیلی کم استخراج مناطق با رنگ مشخص امکان پذیر نیست.
- شناسایی علامتی که مقدار قابل توجهی از آن خراب است و یا دارای زاویه چرخش زیاد است.
- ترکیب چند نوع ویژگی برای این‌که بتوان تمایز بیشتری بین نمونه‌ها از کلاس‌های مختلف ایجاد کرد. همچنین از روش‌های کاهش ویژگی برای کاهش ابعاد بردار ویژگی استفاده گردد.

- [1] V.Andrey, K.Hyun Jo, "Automatic Detection and Recognition of Traffic Signs using Geometric Structure Analysis", *SICE-ICASE International Joint Conference*, pp.1451-1456, Busan , Korea, 2006.
- [2] A.Ruta, Y.Li, and X.Liu, "Real-time traffic sign recognition from video by class-specific discriminative features", *The Journal of the Pattern Recognition Society*, Volume 43, Issue 1, pp. 416-430, January 2010.
- [3] Nguwi Y., Kouzani A., 2006, "A Study on Automatic Recognition of Road Signs", in *Proc. IEEE Conference on Cybernetics and Intelligent Systems*, Bangkok, Thailand, pp. 1-6.
- [4] L.Bruno, G.Parla and C.Celauro, "Improved Traffic Signal Detection and Classification via Image Processing Algorithms", *5th International Congress – Sustainability of Road Infrastructures*, Volume 53, pp.811-821, October 2012.
- [5] ZHANG Hanxizi, WANG Bo, ZHENG Zhihui, DAI Yujie, "A Novel Detection and Recognition System for Chinese Traffic Signs", *Proceedings of the 32nd Chinese Control Conference*, pp.8102-8107, July 26-28, 2013, Xi'an, China.
- [6] Karla Brkić, "An overview of traffic sign detection methods", In Press.
- [7] ا.لهونی فرد، ع.چاله چاله، س.ع.رضوی ابراهیمی، "روش جدید تشخیص الگو برای سیستم هشدار دهنده شناسایی علائم راهنمایی و رانندگی"، *دومین کنفرانس ملی تصادفات جاده‌ای، سوانح ریلی و هوایی*، ایران، زنجان، ۱۳۹۰.
- [8] A.Lorsakul, J.Suthakorn, "Traffic Sign Recognition for Intelligent Vehicle/Driver Assistance System Using Neural Network on OpenCV", *The 4th International Conference on Ubiquitous Robots and Ambient Intelligence*, 2007, URAI.
- [9] U.Lydia Jau, Ch.Siong I and G.Weng Ng, " A comparison of RGB and HSI color segmentation in real – time video images: A preliminary study on road sign detection", *International Symposium on Information Technology*, Volume 4, pp.1-6, Aug 2008.
- [10] S.Bascon, S.Arroyo, P.Jimenez, H.Moreno and F.Ferreras, "Road-Sign Detection and Recognition Based on Support Vector Machines", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 8, No. 2, pp.264-278, June 2007.
- [11] H.Liu, D.Liu, and J.Xin, "Real time recognition of road traffic sign in motion image based on genetic algorithm", in *Proc. 1st. Int.Conf.Match.Learn. Cybern*, pp.83-86, Nov.2002.

- [12] A. Farag, A. Abdel-Hakim, "Detection, Categorization and Recognition of Road Signs for Autonomous Navigation", *Proceedings of Acivs 2004 (Advanced Concepts for Intelligent Vision Systems)*, pp.125-130, Brussels, Belgium.
- [13] R.Azad, B.Azad, I.Tavakoli Kazerooni, "Optimized Method for Iranian Road Signs Detection and Recognition System", *International Journal of Research in Computer Science*, Volume 4, Issue 1, pp. 19-26, January 2014.
- [14] Y. Aoyagi and T. Asakura, "A study on traffic sign recognition in scene image using genetic algorithms and neural networks" *22ndIEEE Int. Conf. Ind. Electron*, vol. 3, pp. 1838–1843, Aug. 1996, Taipei, Taiwan.
- [15] Ch.Shelke, P.Karde, "Traffic Sign Recognition", *International Journal Of Computational Engineering Research*. Vol. 2, Issue. 8, pp. 47-52, December 2012.
- [۱۶] ص.دارنگ، ح.دهقانی، م.شاهسونی، "شناسایی علائم راهنمایی و رانندگی با استفاده از طبقه‌بند بیزی"، هفتمین کنفرانس ماشین بینایی و پردازش تصویر/ایران، تهران، ۱۳۹۰.
- [17] L.Estevez and Kehtarnavaz, "A real time histographic approach to road sign recognition", *Image Analysis and Interpretation, Proceedings of the IEEE Southwest Symposium on*, pp.95-100, April1996.
- [18] F. Zaklouta , B.Stanciulescu, "Real-time traffic sign recognition in three stages" , *The Journal of the Robotics and Autonomous Systems*, In Press.
- [۱۹] رافائل سی. گونزالز، ریچارد ای. وودز؛ مترجمین: مجتبی لطفی زاد، امیر مسعود عمویی، سعید میر قاسمی، "پردازش دیجیتالی تصاویر"، ویراست سوم، کتاب نیاز (اتشارات پیوند نو)، تهران، ایران، ۱۳۹۰.
- [20] R.Gonzalez, R.Woods, "*Digital Image Processing*", Second Edition, USA, New Jersey, Tom Robbins, 2002.
- [21] Haykin.Simon," *Neural Networks and Learning Machines*", Third Edition, USA, New Jersey, Tom Robbins, 2008.
- [22] R.Duda, P.Hart, D.Stork, "*Pattern Classification*", Second Edition, Wiley, 2000.
- [23] V. Vapnik, "*The Nature of Statistical Learning Theory*", New York, Springer-Verlag, 1995.
- [24] M.Martinez-Ramon , "*Support Vector Machines for Antenna Array Processing and Electromagnetics*", Universidad Carlos III deMadrid, Spain, Morgan & Claypool ,USA,2006.
- [25] Beeman, Dave,"Multi-layer Perceptrons (Feed-forward Nets), Gradient Descent and Back Propagation", University of Colorado, 2001.

[25] S. Maldonado Bascon, J. Acevedo Rodríguez, S. Lafuente Arroyo, A. Fernández Caballero, F. Lopez-Ferreras, "An optimization on pictogram identification for the road-sign recognition task using SVMs", *Computer Vision and Image Understanding Journal*, Volume.114, Issue.3, pp.373-383, March 2010.

[26] M.A.Garcia-Garrido, M.A.Sotelo, E.Martín-Gorostiza, "Fast Traffic Sign Detection and Recognition Under Changing Lighting Conditions", *IEEE Intelligent Transportation Systems Conference*, pp.811-816, Canada, 2006.

[۲۷] م.بحرانی، م.جم راد، "شناسایی علائم راهنمایی و رانندگی با استفاده از شبکه عصبی"، دومین کنفرانس ماشین بینایی و پردازش تصویر/ایران، تهران، ۱۳۸۱.

[۲۸] ح.خسروی، م.شهرکی، "شناسایی هوشمند علائم ترافیکی با استفاده از ویژگی‌های مبتنی بر گرادیان"، هشتمین کنفرانس ماشین بینایی و پردازش تصویر، ایران، زنجان، ۱۳۹۲.

[29] Hechi.Ahmed, Mtibaa.Abdellaif, "Automatic Detection and Recognition of Road Sign for Driver Assistance System", *Electrotechnical Conference(MELECON), 16 th IEEE Mediterranean*, pp.888-891, 2012.

[30] Yong.Huang, Ren. Cao, "Fast Road Signs Recognition Using Contour Features", *6th International Congress on Image and Signal Processing (CISP)*, pp.101-106, 2013.

[۳۱] لی وانگ؛ مترجمین: محمد تشنه لب، نیما صفاریپور، داریوش افیونی، "سیستم‌های فازی و کنترل فازی"، چاپ ششم، انتشارات دانشگاه صنعتی خواجه نصیرالدین طوسی، تهران، ایران، ۱۳۸۹.

[32] S.Huang, C. Gao, S.Meng, Q.Li, C.Chen, and C. Zhang, "Circular Road Sign Detection and Recognition based on Hough Transform", *5th International Congress on Image and Signal Processing (CISP)*, pp.1214-1218, 2012.

[33] A.Hechri, A. Mtibaa, "Lanes and Road Signs Recognition for Driver Assistance System", *IJCSI International Journal of Computer Science Issues*, Vol. 8, Issue 6, No 1, pp.402-408, November 2011.

[34] S.Sallah, F.Hussin, and M.Yusoff, "Road Sign Detection and Recognition System for Real-Time Embedded Applications", *International Conference on Electrical, Control and Computer Engineering*, pp.213-218, Malaysia, June 21-22, 2011.

[۳۵] م.قویدل، ح.خسروی، "شناسایی علائم راهنمایی و رانندگی در صحنه‌های واقعی با استفاده از ماشین بردار پشتیبان"، ششمین کنفرانس فناوری اطلاعات و دانش، ایران، شاهرود، ۱۳۹۳.

Abstract

The purpose of the current research is automatic road sign detection and recognition. In first step we extract traffic sign from within various objects in the image. In the second step we recognize the type of the sign. In order to reduce illumination variation effect on detection step, we convert input image of RGB color space to HSI space. Then we extract red and blue regions of the input image by thresholding on hue and saturation parameters. Furthermore for better extraction of red and blue parts, we use fuzzy logic. In the next phase, we extract parts of the image that are triangle, circle, rectangle and octagon and send them to the recognition module.

We produced a data-set of road signs that exist in Iran, containing 350 road signs in 35 different classes. Detection accuracy that we attained is 97.87 percent on this data-set.

For exact classification of the sign, background of the detected sign is removed and then two approaches are examined. In the first approach, some features are extracted based on profile of the sign image. Support vector machine were trained on 840 signs of 40 classes and 416 signs used for testing. Classification accuracy of 91.1 percent on test data was achieved.

In the second approach, some features are extracted based on gradient histogram and neural network is used for classification. To evaluate the proposed method, a data set of 1280 signs in 40 classes used. 67 percent of them, 864 signs assigned for training and 33 percent, 416 signs, used for test. Classification accuracy that we attained is 99 percent on train data and 93.86 on test data. Notice that our signs are in various light conditions, various backgrounds and various scales. Furthermore they are blur and have some rotation. Experimental results show the effective performance of the proposed method.

In order to attain real time algorithm for detection and recognition of road signs, we converted early Matlab code to C++ code. We achieved a mean processing time of 170 ms per image in detection step and mean processing time of 121 ms per image in recognition step. These make the proposed method almost real-time.

Keywords: HSI color space, Fuzzy logic, Support Vector Machine, Profile Feature, Feature extraction, Gradient histogram, Neural Network



Shahrood University
Faculty of Electrical and Robotics Engineering

Traffic Sign Detection and Recognition Using Neural Networks and Template Matching

Thesis
Submitted in Partial Fulfillment of the
Requirements for the Degree of Master of Science (M.Sc.)
in Electronic Engineering

By:
Masoud Ghavidel

Supervisor:
Dr. Hossein Khosravi

Summer 2014