

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی برق و رباتیک

پایان نامه کارشناسی ارشد مهندسی مخابرات سیستم

تشخیص و طبقه‌بندی خودروها با استفاده از صدای انتشاری آنها مبتنی بر
SVM

نگارنده:

فرهاد شهبازی بندانی

استاد راهنما:

آقای دکتر حسین مروی

شهریور ۱۳۹۷

شماره:
تاریخ:

باسمه تعالی



مدیریت تحصیلات تکمیلی

فرم شماره (۳) صورت جلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم / آقای: **فرهاد شهبازی بندانی** با شماره دانشجویی: **۹۴۱۰۶۱۴** رشته: **برق گرایش: مخابرات سیستم** تحت عنوان: **تشخیص و طبقه بندی خودروها با استفاده از صدای انتشاری آن ها مبتنی بر SVM** که در تاریخ **۹۷/۰۶/۰۷** با حضور هیئت محترم داوران در دانشگاه صنعتی شاهرود برگزار شد به شرح ذیل اعلام می گردد:

<input type="checkbox"/> مردود	<input type="checkbox"/> قبول (با درجه:)
<input type="checkbox"/> عملی	<input type="checkbox"/> نظری
نوع تحقیق:	

امضاء	مرتبۀ علمی	نام و نام خانوادگی	عضو هیئت داوران
			۱- استاد راهنمای اول
			۲- استاد راهنمای دوم
			۳- استاد مشاور
			۴- نماینده تحصیلات تکمیلی
			۵- استاد ممتحن اول
			۶- استاد ممتحن دوم

نام و نام خانوادگی رئیس دانشکده:

تاریخ و امضاء و مهر دانشکده:

تقدیم به پدر و مادر عزیزم

که موی سپید گشتند

تا روی سپیدگردم

تقدیر و تشکر

خداوند را شاکریم که به ما قدرت تعلیم را عنایت فرمود و نیروی عقل را عطا کرد تا با یاری آن پرده جهل را کنار زده و رازهای نهان طبیعت را آشکار سازیم.

از راهنمایی‌های بی‌شائبه استاد فرزانه، جناب آقای دکتر حسین مروی که در کمال سعه‌صدر، با حسن خلق و فروتنی، از هیچ کمکی در این عرصه بر من دریغ نمودند و زحمت راهنمایی این پایان‌نامه را بر عهده گرفتند، تشکر و قدردانی می‌نمایم.

لازم می‌دانم از اساتید بزرگوارم جناب آقای دکتر احمدی فرد، جناب آقای دکتر معروضی، جناب آقای دکتر گرایلو و جناب آقای دکتر جوان به پاس زحمات گران‌بهایشان در دوره کارشناسی ارشدم تشکر و قدردانی کنم.

در پایان از پدر و مادر عزیزم که همراهم همیشگی من در زندگی بوده کمال تشکر و قدردانی را دارم.

فرهاد شهبازی بندانی

شهریور ۱۳۹۷

تعهدنامه

اینجانب **فرهاد شهبازی بندانی** دانشجوی دوره کارشناسی ارشد رشته برق مخابرات سیستم دانشکده برق و رباتیک دانشگاه صنعتی شاهرود نویسنده پایان نامه تشخیص و طبقه‌بندی خودروها با استفاده از صدای انتشاری آن‌ها مبتنی بر SVM تحت راهنمایی آقای دکتر حسین مروی 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 مورداستفاده استناد شده است.
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود هست و مقالات مستخرج با نام «دانشگاه صنعتی شاهرود» و یا «Shahrood University of Technology» به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده‌اند در مقالات مستخرج از پایان نامه رعایت می‌گردد.
- در کلیه مراحل انجام این پایان نامه، در مواردی که از موجود زنده (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: ۱۳۹۷ / ۰۶ / ۰۷

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم افزارها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می‌باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی‌باشد.

چکیده

طراحی و تولید سیستم‌های تشخیص صوت، هدف تحقیقاتی بسیاری از مراکز علمی در نیم‌قرن اخیر است. یکی از اهداف انسان در تولید چنین سیستم‌هایی مسلماً توجه به این نکته است که ورود اطلاعات و اجرای دستورات به صورت صوتی باعث صرفه‌جویی در وقت و هزینه و بالا رفتن کیفیت زندگی می‌شود. موضوع مورد بحث در این پایان‌نامه، ارائه یک سیستم شناسایی و طبقه‌بندی هوشمند خودرو با پردازش صدای تولیدی آن است. سیستم ارائه‌شده بایستی بتواند از صدای خودرو، نوع آن را شناسایی و طبقه‌بندی نماید. اولین قدم در پیاده‌سازی این سیستم، جمع‌آوری پایگاه داده صوتی است. از این رو پایگاه داده را از ۶ کلاس خودرو سواری Toyota, Jeep, Ford Tempo, Ferrari 308, BMW 540i و yaris و Volvo Amazon که از هر کلاس ۵۱ نمونه صوتی در دسترس بوده، جمع‌آوری نموده‌ایم. عملکرد سیستم به این صورت است که از صدای ضبط‌شده هر خودرو، چهار ویژگی مهم MFCC، LPC، LPCC و LSF با طول‌های متفاوت ۱۳، ۲۶ و ۳۹ ضریب ویژگی در هر فریم را استخراج کرده و به‌عنوان ویژگی‌های مرجع خودرو مربوطه به طبقه بند SVM می‌دهد تا طبقه بند با این ویژگی‌ها آموزش ببیند. حال جهت ارزیابی، سیستم با دریافت یک صوت به‌عنوان ورودی ابتدا ویژگی‌های آن را استخراج کرده و به طبقه بند SVM می‌دهد سپس طبقه بند با مقایسه ویژگی‌های صوت ورودی با ویژگی‌های مرجع، نوع خودرو را شناسایی و طبقه‌بندی می‌کند. نتایج حاصل از پیاده‌سازی این سیستم با ۷۵ درصد داده‌ها جهت آموزش و ۲۵ درصد باقی‌مانده جهت ارزیابی و آزمایش طبقه بند نشان می‌دهد تکنیک LSF با دقت شناسایی میانگین ۸۵,۰۲٪ بهترین تکنیک استخراج ویژگی و روش MFCC نیز با دقت شناسایی میانگین ۸۲,۹۴٪ در رتبه بعد از آن قرار دارد. از طرفی تکنیک‌های استخراج ویژگی LPC و LPCC به ترتیب با دقت شناسایی میانگین ۵۵,۶۲٪ و ۵۸,۹٪ برای این سیستم شناسایی نمی‌توانند نتایج خوبی به همراه داشته باشند.

کلمات کلیدی: شناسایی و دسته‌بندی، استخراج ویژگی، SVM، MFCC، LPC، LPCC و LSF

فهرست مطالب

صفحه	عنوان
۱	فصل اول: مقدمه
۲	۱-۱ تاریخچه
۲	۲-۱ تعریف مسئله
۳	۳-۱ ضرورت انجام تحقیق
۳	۴-۱ ساختار سیستم شناسایی و طبقه‌بندی خودرو با پردازش صدای آن
۳	۱-۴-۱ نحوه کار سیستم شناسایی و طبقه‌بندی خودرو
۴	۵-۱ مزایای سیستم شناسایی و طبقه‌بندی خودرو با پردازش صدای آن
۵	۶-۱ ساختار پایان‌نامه
۷	فصل دوم: مروری بر کارهای پیشین
۸	۱-۲ مقدمه
۹	۲-۲ سیستم شناسایی و طبقه‌بندی پیشنهادی در کلیه مقالات
۹	۱-۲-۲ تهیه و آماده‌سازی پایگاه داده
۱۰	۲-۲-۲ استخراج ویژگی
۱۲	۳-۲-۲ طبقه‌بندی
۱۲	۳-۲ مروری بر کارهای انجام‌شده
۱۲	۱-۳-۲ تکنیک استخراج ویژگی از حوزه‌ی زمان
۱۵	۲-۳-۲ تکنیک استخراج ویژگی از حوزه‌ی فرکانس
۱۸	۳-۳-۲ تکنیک استخراج ویژگی از حوزه‌ی زمان - فرکانس
۲۱	۴-۲ نتیجه‌گیری

۲۳	فصل سوم: مبانی نظری
۲۴	۱-۳ مقدمه
۲۴	۲-۳ توصیف قسمت‌های مختلف سیستم پیشنهادی
۲۴	۱-۲-۳ پیش‌پردازش داده‌های صوتی
۲۵	۲-۲-۳ استخراج ویژگی
۲۶	۱-۲-۲-۳ روش استخراج ویژگی MFCC
۲۸	۲-۲-۲-۳ روش استخراج ویژگی LPC
۳۰	۳-۲-۲-۳ روش استخراج ویژگی LSF
۳۱	۴-۲-۲-۳ روش استخراج ویژگی LPCC
۳۲	۳-۲-۳ طبقه‌بندی
۳۲	۱-۳-۲-۳ طبقه‌بند SVM
۳۵	فصل چهارم: روش پیشنهادی و نتایج شبیه‌سازی
۳۶	۱-۴ مقدمه
۳۶	۲-۴ سیستم پیشنهادشده جهت بازشناسی هوشمند خودرو با پردازش صدای آن
۳۶	۱-۲-۴ جمع‌آوری پایگاه داده
۳۹	۲-۲-۴ پیش‌پردازش داده‌های صوتی
۴۰	۳-۲-۴ استخراج ویژگی
۴۱	۴-۲-۴ طبقه‌بندی
۴۲	۳-۴ نتایج حاصل از شبیه‌سازی انجام روش پیشنهادی
۴۲	۱-۳-۴ ماتریس درهم‌ریختگی
۴۶	۲-۳-۴ دقت بازشناسی روش‌های مختلف استخراج ویژگی با طول‌های متفاوت
۴۹	۴-۴ جمع‌بندی
۵۱	فصل پنجم: نتیجه‌گیری و پیشنهادها
۵۲	۱-۵ مقدمه
۵۲	۲-۵ نتیجه‌گیری
۵۲	۱-۲-۵ نتیجه‌گیری کلی از سیستم شناسایی هوشمند پیشنهادشده خودرو

۵-۲-۲ مقایسه عملکرد روش پیشنهادی با تعدادی از روش‌های مطرح شده قبلی ۵۳

۵-۳ خط سیر آینده ۵۶

فهرست مراجع ۵۷

فهرست شکل‌ها

- شکل ۱-۱ ساختار سیستم شناسایی و طبقه‌بندی خودرو با پردازش صدای آن ۳
- شکل ۱-۲ فراوانی مقالات و تحقیقات علمی در دو دهه‌ی اخیر ۸
- شکل ۲-۲ سیستم شناسایی و طبقه‌بندی پیشنهادی در کلیه مقالات ۹
- شکل ۱-۳ قسمت‌های مختلف سیستم پیشنهادی شناسایی و طبقه‌بندی خودرو با پردازش صدای آن ۲۴
- شکل ۲-۳ (چپ) سیگنال اصلی ورودی همراه با نویز (راست) سیگنال پردازش شده بدون نویز ۲۵
- شکل ۳-۳ قاب گذاری سیگنال صوتی ۲۶
- شکل ۴-۳ بلوک دیاگرام قسمت‌های مختلف الگوریتم استخراج ویژگی MFCC ۲۷
- شکل ۵-۳ بلوک دیاگرام قسمت‌های مختلف الگوریتم استخراج ویژگی LPC ۳۰
- شکل ۶-۳ بلوک دیاگرام روش استخراج ویژگی LPCC ۳۱
- شکل ۷-۳ تابع جداکننده، بردارهای پشتیبان و حاشیه اطمینان در یک طبقه بند SVM ۳۴
- شکل ۱-۴ بلوک دیاگرام سیستم پیشنهاد شده جهت بازشناسی هوشمند خودرو با پردازش صدای آن ۳۶
- شکل ۲-۴ بلوک دیاگرام استخراج ویژگی‌های مختلف از سیگنال‌های صوتی ۴۰
- شکل ۳-۴ دقت طبقه‌بندی صحیح هر کلاس خودروی پایگاه داده با روش‌های مختلف استخراج ویژگی (۱۳ ضریب) ۴۶
- شکل ۱-۵ مقایسه عملکرد سیستم بازشناسی هوشمند خودرو بر اساس ۴ روش استخراج ویژگی با ۳ طول مشخص ۵۲

فهرست جدول‌ها

- جدول ۴-۱ تصاویر به همراه نمودار سیگنال صوتی خودروهای مربوط به پایگاه داده صوتی ۳۷
- جدول ۴-۲ ماتریس درهم‌ریختگی سیستم بازشناسی هوشمند خودرو با روش استخراج ویژگی MFCC
(۱۳ ضریب) ۴۳
- جدول ۴-۳ ماتریس درهم‌ریختگی سیستم بازشناسی هوشمند خودرو با روش استخراج ویژگی LPCC
(۱۳ ضریب) ۴۳
- جدول ۴-۴ ماتریس درهم‌ریختگی سیستم بازشناسی هوشمند خودرو با روش استخراج ویژگی LPC
(۱۳ ضریب) ۴۴
- جدول ۴-۵ ماتریس درهم‌ریختگی سیستم بازشناسی هوشمند خودرو با روش استخراج ویژگی LSF
(۱۳ ضریب) ۴۵
- جدول ۴-۶ نتایج حاصل از شبیه‌سازی سیستم بازشناسی خودرو با روش‌های مختلف استخراج ویژگی
(۱۳ ضریب) ۴۷
- جدول ۴-۷ نتایج حاصل از شبیه‌سازی سیستم بازشناسی خودرو با روش‌های مختلف استخراج ویژگی
(۲۶ ضریب) ۴۷
- جدول ۴-۸ نتایج حاصل از شبیه‌سازی سیستم بازشناسی خودرو با روش‌های مختلف استخراج ویژگی
(۳۹ ضریب) ۴۸
- جدول ۵-۱ مقایسه عملکرد روش پیشنهادی با تعدادی از روش‌های مطرح‌شده در تحقیقات قبلی ۵۵

فصل اول:

مقدمه

۱-۱ تاریخچه

علم آکوستیک یکی از علوم قدیمی است که نخستین دانشمندان علم فیزیک نیز به آن پرداخته‌اند. در قرن هجدهم دالامبر^۱ و دانیل برنولی^۲ فرمول ریاضی مهم پخش صوت را در فیزیک و ریاضی کشف نمودند. پس از آن پیشرفت بسیاری در زمینه‌های متفاوت علم آکوستیک انجام گرفت و نظریه موجی بودن صوت توسط هویگنس^۳ پی‌ریزی گردید. با توجه به وسعت علم آکوستیک، هم‌اکنون شاخه‌های گوناگونی از این علم ایجاد گردیده که یکی از این شاخه‌ها تشخیص و تفکیک اصوات است. هدف از تشخیص صوت، ایجاد روشی هوشمند برای شناسایی اصوات بوده که اطلاعات صوتی را دریافت و سیگنال‌های صوتی مختلف را تفکیک و دسته‌بندی می‌نماید. اولین سیستم مبتنی بر فناوری تشخیص صوت در سال ۱۹۵۲ میلادی در آزمایشگاه بل طراحی شد.

۱-۲ تعریف مسئله

جهان در حال حرکت به سوی استفاده کردن از سیستم هوشمندی است که همانند انسان بتواند صدای محیط را درک کرده و تصمیمات صحیح و حساب‌شده اتخاذ نماید و به‌طور کلی تمامی فرایندهای انسانی را انجام داده و نقش عامل انسانی را حذف یا کم‌رنگ کند. از این‌رو ارائه‌ی سیستم هوشمندی که توانایی تشخیص صداهای محیط را داشته باشد و بر مبنای درک صحیح خود از اصوات موجود در محیط، پاسخ مناسبی به آن دهد بسیار ضروری به نظر می‌رسد. یکی از این سیستم‌های هوشمند، سیستم شناسایی و طبقه‌بندی خودرو با پردازش صدای آن است.

¹ D'alembert

² Daniel Bernoulli

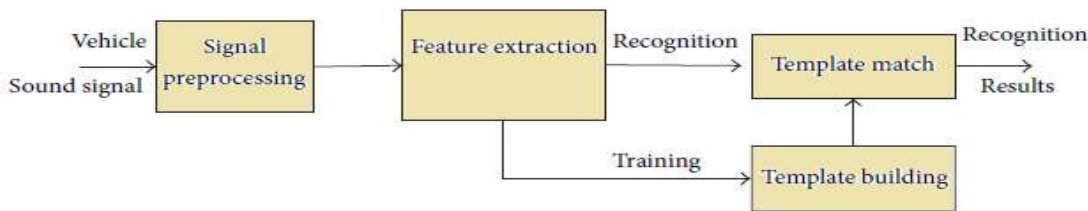
³ Huygens

۳-۱ ضرورت انجام تحقیق

تولید روزافزون خودرو و عرضه‌ی آن در بازار منجر به افزایش حجم ترافیک و تصادفات شده است که تأثیر مهمی روی اقتصاد و محیط‌زیست دارد. از این‌رو سیستم شناسایی و طبقه‌بندی خودروها به‌عنوان یک سیستم نظارت ترافیک هوشمند می‌تواند کمک شایانی در کنترل حجم ترافیک داشته باشد که به‌نوبه‌ی خود منجر به کاهش تصادفات و بهبود جریان ترافیک، زمان مسافرت، امنیت و آسایش کاربران جاده‌ها و اثرات زیست‌محیطی خواهد شد [۱]. علاوه بر این، سیستم شناسایی و طبقه‌بندی خودروها در زمینه‌های نظامی و غیرنظامی نظیر تخمین سرعت و موقعیت جغرافیایی خودرو، کاهش آلودگی صوتی، نظارت بر میدان جنگ و نظارت بر حمل‌ونقل و مرزهای کشور نیز کاربرد دارد [۲،۳]. حال با توجه به کارایی بسیار مهم این سیستم شناسایی و نیز به دلیل نبود تحقیقات علمی به‌طور جامع و کاربردی مرتبط با این زمینه در کشور ما، پژوهش در این زمینه از پردازش صوت را ضروری می‌سازد.

۴-۱ ساختار سیستم شناسایی و طبقه‌بندی خودرو با پردازش صدای آن

ساختار کلی سیستم شناسایی و طبقه‌بندی خودرو با پردازش صدای آن در شکل (۱-۱) نشان داده شده است [۳].



شکل ۱-۱ ساختار سیستم شناسایی و طبقه‌بندی خودرو با پردازش صدای آن [۳]

۱-۴-۱ نحوه کار سیستم شناسایی و طبقه‌بندی خودرو

به‌طور کلی نحوه کار سیستم شناسایی و طبقه‌بندی خودرو نشان داده‌شده در شکل (۱-۱) از دو مرحله اساسی تشکیل شده است:

الف) مرحله‌ی آموزش

سیگنال‌های صوتی مربوط به تعدادی خودرو در موقعیت‌های مختلف توسط حس‌گرهای صوتی ضبط و ویژگی‌های کلیدی آن‌ها طی مراحل مختلفی استخراج شده و در حافظه سیستم به‌عنوان ویژگی‌های مرجع ذخیره می‌گردد.

ب) مرحله‌ی آزمایش

سیگنال منبع صوت موردنظر و یا هر منبع صوت دیگر که در مقابل حس‌گر صوتی قرار می‌گیرد، به نمونه‌های مناسب جهت کار با سیستم شناسایی تبدیل شده و به دنبال آن ویژگی‌های کلیدی صوت طی روش‌های مختلف استخراج می‌گردد. سپس این ویژگی‌ها با ویژگی‌های منابع صوت مرجع جهت شناسایی و طبقه‌بندی خودرو مقایسه می‌شود.

۵-۱ مزایای سیستم شناسایی و طبقه‌بندی خودرو با پردازش صدای آن

نظارت ترافیک و تخمین پارامترهای مهم شهری آن یکی از شاخه‌های در حال ظهور بر پایه سیگنال صوتی است. این شاخه یکی از معیارهای هوشمند سازی سیستم حمل‌ونقل^۱ است. امروزه سیستم نظارت ترافیک مبتنی بر دوربین‌های تصویری است که دارای گستره‌ی وسیعی از کاربردهای متفاوت نظیر شمارش خودروها، اندازه‌گیری سرعت، شناسایی و طبقه‌بندی آن‌هاست اما نصب و راه‌اندازی، تعمیر و نگهداری این سیستم بسیار پرزحمت و هزینه‌بر است. از این رو یکی از مزایای سیستم شناسایی و طبقه‌بندی ارائه‌شده در این پایان‌نامه، استفاده از آن به‌عنوان یک سیستم نظارت ترافیک مبتنی بر سیگنال صوتی است. این روش علاوه بر این که تعمیر، نگهداری و نصب راحت و ارزانی داشته، نرخ شناسایی و انعطاف‌پذیری را نیز بهبود می‌بخشد که منجر به افزایش ایمنی و ضریب اطمینان پایین روش قبل می‌شود. در واقع صدای متفاوت خودروهای مختلف مانند اتومبیل، موتورسیکلت، اتوبوس و

^۱ Intelligent Transport Systems(ITSs)

کامیون اطلاعات مفیدی را به صورت آنلاین از وضعیت جاده‌ها و داده‌های آماری ترافیک در اختیار ما قرار می‌دهد [۱,۴].

۱-۶ ساختار پایان‌نامه

در فصل اول مقدمه‌ای در مورد سیستم شناسایی خودرو بیان شد. در فصل دوم به بررسی تحقیقات انجام‌شده در زمینه شناسایی خودرو با پردازش صدای آن و روش‌های مختلف برای انجام این کار پرداخته شده است. در فصل سوم مبانی نظری ارائه می‌شود. در فصل چهارم پایگاه داده صوتی معرفی می‌گردد، سپس روش پیشنهادی را ارائه کرده و سیستم شناسایی خودرو بر اساس آن پیاده‌سازی می‌شود. سپس نتایج شبیه‌سازی حاصل از پیاده‌سازی روش پیشنهادی را نشان می‌دهیم. در فصل پنجم نتیجه‌گیری کلی صورت گرفته و عملکرد آن با روش‌های موجود مقایسه خواهد شد و سپس پیشنهادهای ادامه کار ارائه می‌شود. در پایان، مراجع مورد استفاده در پایان‌نامه را ذکر می‌کنیم.

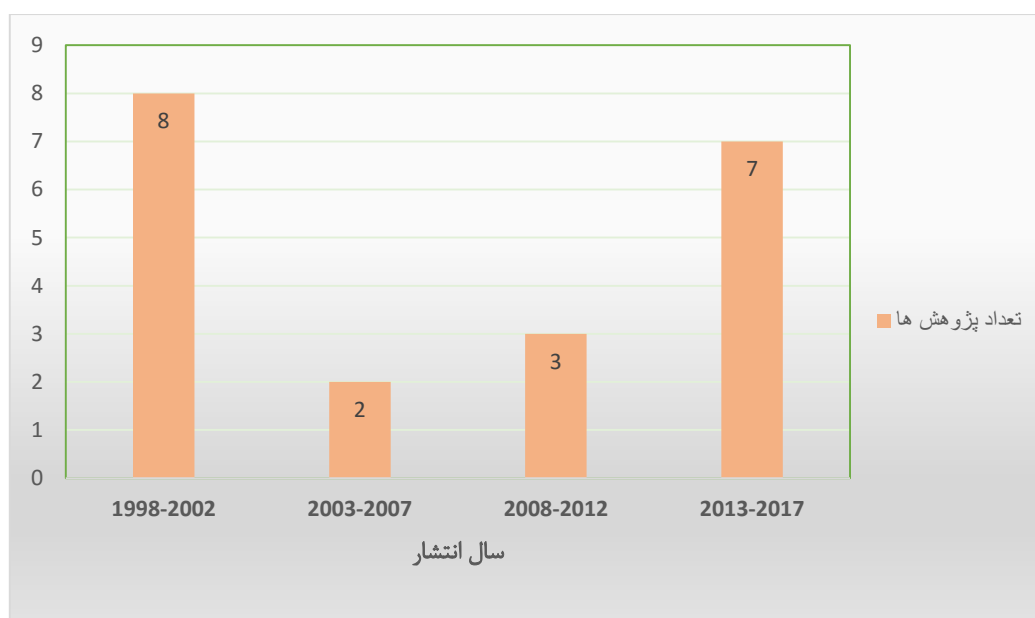
فصل دوم:

مروری بر کارهای پیشین

۱-۲ مقدمه

در این فصل با توجه به شناختی که از سیستم شناسایی و طبقه‌بندی خودرو، کاربردها و الزام وجود چنین مکانیزمی برای شناسایی اصوات داشته‌ایم، به شرح مختصری از تحقیقات انجام‌شده در ارتباط با این موضوع مهم خواهیم پرداخت و سپس نتیجه‌گیری پایان‌بخش این فصل است.

شکل (۱-۲) فراوانی مقالات و تحقیقات علمی در زمینه سیستم شناسایی و طبقه‌بندی خودروها با استفاده از پردازش صدای آن‌ها را در دو دهه‌ی اخیر نشان می‌دهد.

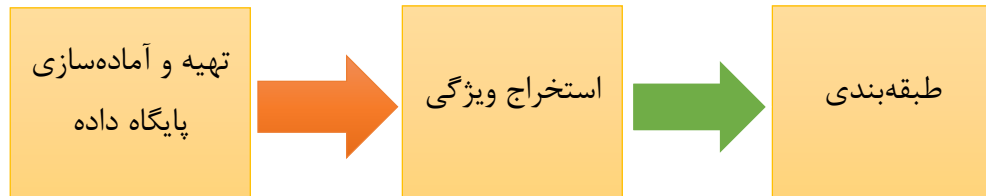


شکل ۱-۲ فراوانی مقالات و تحقیقات علمی در دو دهه‌ی اخیر

شکل (۱-۲) به‌خوبی نشان می‌دهد که در چند سال اخیر تحقیقات علمی سیر صعودی داشته که بیانگر اهمیت کاربرد این سیستم است.

۲-۲ سیستم شناسایی و طبقه‌بندی پیشنهادی در کلیه مقالات

در راستای تشخیص و طبقه‌بندی خودروها با پردازش صدای آن‌ها، تحقیقات علمی مختلفی در سال‌های اخیر انجام گرفته و برای آن سیستم‌هایی طراحی و پیشنهاد شده است. شکل (۲-۲) به خوبی یک سیستم شناسایی و طبقه‌بندی پیشنهادی در کلیه مقالات را نشان می‌دهد.



شکل ۲-۲ سیستم شناسایی و طبقه‌بندی پیشنهادی در کلیه مقالات

سیستم پیشنهادی نشان داده شده در شکل (۲-۲) از سه قسمت اساسی ساخته شده است:

۱-۲-۲ تهیه و آماده‌سازی پایگاه داده

در راستای تشخیص صدا این نکته حائز اهمیت است که پایگاه داده استاندارد وجود ندارد و در هر پژوهش، محقق با توجه به کاربرد آن، پایگاه داده موردنظر را تهیه و یا می‌سازد. پایگاه داده در تحقیقات صورت گرفته را می‌توان با توجه به دو معیار طبقه‌بندی نمود:

الف) کاربرد

در این نوع تحقیقات، خودروها به دودسته نظامی و غیرنظامی تقسیم می‌شوند.

ب) وزن

در این نوع تحقیقات خودروها به سه دسته سبک، نیمه سنگین و سنگین تقسیم می‌شوند.

پس از تهیه داده‌های صوتی موردنظر، باید پیش‌پردازش‌هایی از قبیل هم‌اندازه کردن طول سیگنال‌های صوتی، حذف مقدار DC سیگنال، از بین بردن قسمت‌های سکوت، حذف نویزهای پس‌زمینه، نمونه‌برداری، نرمالیزه کردن داده‌ها، کوانتیزه نمودن داده‌ها، قاب‌بندی و پنجره گذاری مناسبی بر روی پایگاه داده صورت گیرد تا داده‌های صوتی جهت عملکرد بهتر سیستم برای مرحله‌ی بعد آماده شوند.

۲-۲-۲ استخراج ویژگی

هر سیگنال صوتی خواص منحصر به فرد خود را داراست که با این خواص می‌توان سیگنال را شناسایی و تا حد قابل قبولی بازسازی کرد حال به این خواص، ویژگی‌های سیگنال گفته می‌شود. این ویژگی‌ها را می‌توان از سه حوزه‌ی زمان، فرکانس و زمان-فرکانس سیگنال صوتی استخراج کرد.

الف) ویژگی‌های حوزه‌ی زمان

این دسته ویژگی‌ها از سیگنال‌های صوتی در حوزه‌ی زمان استخراج می‌شوند. حال با توجه به این که سیگنال‌های صوتی نا ایستاد هستند و مشخصات آماری آن‌ها با گذشت زمان تغییر می‌کند، پس این ویژگی‌ها نیز تحت تأثیر تغییرات قرار می‌گیرند.

مهم‌ترین ویژگی‌هایی که در تحقیقات صورت گرفته شده از حوزه‌ی زمان استخراج شده‌اند عبارت‌اند از: ضرایب پیشگویی خطی^۱، انرژی زمان کوتاه^۲، تحلیل مؤلفه‌های اساسی^۳، پردازش سیگنال کد شده زمانی پیشرفته^۴، پردازش سیگنال کد شده زمانی^۵، نرخ عبور از صفر^۶، سری ولترا^۷ و ریشه میانگین مربعات^۸

¹ Linear Prediction Coding(LPC)

² Short Time Energy(STE)

³ Principal Components Analysis(PCA)

⁴ Improved Time Encoded Signal Processing(ITESP)

⁵Time Encoded Signal Processing(TESP)

⁶ Zero Crossing Ratio(ZCR)

⁷ Volterra series

⁸ Root Mean Square(RMS)

ب) ویژگی‌های حوزه‌ی فرکانس

این دسته ویژگی‌ها از طیف سیگنال‌های صوتی در حوزه‌ی فرکانس استخراج می‌شوند. در این حالت اطلاعات دقیقی از جزئیات زمانی مربوط به طیف سیگنال‌ها در دسترس نیست.

مهم‌ترین ویژگی‌هایی که در تحقیقات پیشین از حوزه‌ی فرکانس استخراج شده‌اند عبارت‌اند از: تبدیل فوریه گسسته^۱، تبدیل فوریه سریع^۲، ضرایب کپستروم^۳، ضرایب کپستروم مل فرکانسی^۴، ضرایب کپستروم پیشگویی خطی^۵، فرکانس‌های طیف خطی^۶، تخمین طیف توان^۷، وابستگی خطی هارمونیک‌ها^۸، مدل خود بازگشتی^۹، مدل خود بازگشتی متغیر بازمان^{۱۰}، میانگین طیفی^{۱۱}، انحراف از مرکز ثقل طیف^{۱۲} و فرمنت‌ها^{۱۳}

ج) ویژگی‌های حوزه‌ی زمان – فرکانس

این دسته ویژگی‌ها در واقع تلفیقی از ویژگی‌های حوزه‌ی زمان و فرکانس سیگنال‌های صوتی هستند. در این حالت با توجه به پردازش چند دقتی می‌توان به اطلاعات دقیقی از جزئیات زمانی مربوط به طیف سیگنال‌های صوتی دسترسی پیدا کرد.

مهم‌ترین ویژگی‌هایی که در تحقیقات پیشین از حوزه‌ی زمان – فرکانس استخراج شده‌اند عبارت‌اند از: تبدیل موجک گسسته^{۱۴} و تبدیل فوریه زمان کوتاه^{۱۵}

¹ Discrete Fourier Transform(DFT)

² Fast Fourier Transform(FFT)

³ Cepstral Coefficient(CC)

⁴ Mel Frequency Cepstral Coefficient(MFCC)

⁵ linear Predictive Cepstral Coefficient(LPCC)

⁶ linear Spectral Frequency(LSF)

⁷ Power Spectrum Estimation(PSE)

⁸ Harmonic Line Association(HLA)

⁹ Autoregressive Modeling(AR mod)

¹⁰ Time Varying Autoregressive(TVAR)

¹¹ Centroid Mean(CM)

¹² Centroid Standard Deviation(CSD)

¹³ Formant

¹⁴ Discrete Wavelet Transform(DWT)

¹⁵ Short Time Fourier Transform(STFT)

۳-۲-۲ طبقه‌بندی

جهت طبقه‌بندی ویژگی‌های مربوط به داده‌های صوتی هر خودرو، از یک طبقه بند مناسب استفاده شده است که طبقه بند با این ویژگی‌ها آموزش می‌بیند. حال در صورت دریافت هر نوع صوتی توسط حس‌گرها، ویژگی‌های آن صوت استخراج شده و برای شناسایی و طبقه‌بندی خودرو به طبقه بند داده می‌شود.

مهم‌ترین طبقه‌بندی‌هایی که در تحقیقات صورت گرفته‌شده از آن‌ها استفاده شده است، عبارت‌اند از: شبکه عصبی چندلایه^۱، شبکه عصبی تأخیر زمانی^۲، شبکه عصبی پس‌خور^۳، پرسپترون چندلایه^۴، شبکه عصبی پیش‌خور چندلایه^۵، شبکه عصبی احتمالی^۶، ماشین بردار پشتیبان^۷، K نزدیک‌ترین همسایه^۸، تخمین حداکثر احتمال^۹، رویکرد حداکثر فاصله^{۱۰}

۳-۲ مروری بر کارهای انجام‌شده

به‌منظور بررسی بهتر، کلیه‌ی مقالات مطالعه شده در زمینه سیستم شناسایی و طبقه‌بندی خودرو با پردازش صدای آن را با توجه به استفاده از نوع تکنیک استخراج ویژگی به سه دسته تقسیم می‌نماییم:

۱-۳-۲ تکنیک استخراج ویژگی از حوزه‌ی زمان

محققین در [۳] از الگوریتم‌های ITESP و TESP جهت استخراج ویژگی از سیگنال‌های صوتی دو نوع خودرو چرخ‌دار و ریلی استفاده کرده‌اند. نحوه‌ی کار الگوریتم ITESP به این صورت است که ابتدا بر روی هر سیگنال صوتی پیش‌پردازش‌های لازم صورت گرفته، سپس با فرکانس ۲۰ کیلوهرتز نمونه‌برداری می‌شود. حال دو پارامتر، طول دوره ZCR و تعداد مینیمم‌های مثبت در هر دوره از سیگنال صوتی را به‌عنوان یک زوج مرتب مطابق جدول استاندارد ۴۰ کاراکتری کدگذاری می‌کند و درنهایت به ازای

¹ Multilayer Neural Network(MLNN)

² Time Delay Neural Network(TDNN)

³ Backpropagation Neural Network(BPNN)

⁴ Multilayer Perceptron(MLP)

⁵ Multilayer Feedforward Neural

Network(MLFFNN)

⁶ Probabilistic Neural Network(PNN)

⁷ Support Vector Machine(SVM)

⁸ K nearest Neighbor(KNN)

⁹ Maximum Likelihood Estimation(MLE)

¹⁰ Maximum Distance Approach(MPP)

احتمال ظاهر شدن هر کاراکتر و یا دو کاراکتر متوالی یکسان به ترتیب ماتریس ویژگی یک‌بعدی S و یا دوبعدی A تشکیل می‌شود. حال اگر این الگوریتم از جدول استاندارد ۲۹ کاراکتری جهت کدگذاری استفاده کند و سپس به ازای احتمال ظاهر شدن هر کاراکتر، ماتریس ویژگی S را بسازد در این صورت، الگوریتم TESP نامیده می‌شود. جهت ارزیابی عملکرد سیستم، از دو الگوریتم یادشده به‌عنوان تکنیک‌های استخراج ویژگی با دو طبقه بند SVM و KNN استفاده شده است. نتایج شبیه‌سازی به‌خوبی نشان می‌دهد الگوریتم TESP با طبقه بند KNN عملکرد بهتری نسبت به همین الگوریتم با طبقه بند SVM دارد به‌طوری‌که نرخ شناسایی میانگین دو خودرو از ۵۱٪ به ۵۲٪ رسیده و نیز زمان پردازش موردنیاز جهت شناسایی ۰٫۱۹ ثانیه کاهش پیدا کرده است. از طرفی الگوریتم ITESP برای حالت ماتریس ویژگی A با طبقه بند SVM عملکرد بهتری نسبت به همین الگوریتم با طبقه بند KNN دارد به‌طوری‌که نرخ شناسایی میانگین دو خودرو از ۸۵٪ به ۸۷٪ رسیده و زمان پردازش جهت شناسایی نیز به ترتیب از ۲٫۵۵ ثانیه به ۲٫۳۷ ثانیه کاهش پیدا کرده است. این نتیجه برای الگوریتم ITESP در حالت استفاده از ماتریس ویژگی S نیز صادق است. برای الگوریتم ITESP در حالت ماتریس ویژگی A نسبت به S با طبقه بند KNN، نرخ شناسایی میانگین دو خودرو به دلیل استخراج ویژگی‌های بسیار منحصر به فرد صدای هر خودرو از ۸۰٪ به ۸۵٪ رسیده ولی زمان پردازش موردنیاز جهت شناسایی حدود ۱٫۷ ثانیه افزایش پیدا کرده است. این نتیجه برای همین الگوریتم با طبقه بند SVM نیز صادق است. یکی از مزایای استفاده از این الگوریتم‌ها این است که با پیاده‌سازی آن‌ها بر روی حسگرهای معمولی، قدرت محاسبات بالا رفته و زمان پردازش پایین می‌آید به‌طوری‌که در تکنیک‌های زمانی و فرکانسی مرسوم برای برآورده شدن همین انتظارات نیاز به حسگرهای گران‌قیمت داریم. عیب این روش‌ها در این مقاله محدودیت تعداد خودروهاست به‌طوری‌که نتایج شبیه‌سازی مربوط به دو نوع خودروی متفاوت یکی چرخ‌دار و یکی ریلی بوده که ممکن است با تنوع خودروها نتایج بهینه نشود.

پژوهشگران در [۵] از دو تکنیک STE و ZCR جهت استخراج ویژگی از سیگنال صوتی خودرو بهره برده‌اند. در روش ارائه‌شده، ابتدا سیگنال صوتی خودرو با طول ۴۰ ثانیه ضبط و پیش‌پردازش‌های لازم بر روی آن صورت می‌گیرد. سپس با فرکانس ۸ کیلوهرتز نمونه‌برداری و با ۸ بیت کوانتیزه می‌شود. جهت ایستادن نمودن سیگنال صوتی، از اعمال قاب‌های ۲۰۴۸ نمونه‌ای با ۵۰ درصد هم‌پوشانی استفاده می‌کنیم. اکنون برای از بین بردن اثر ناپیوستگی ابتدا و انتهای قاب‌ها، یک پنجره همینگ در هر قاب ضرب می‌شود. در مرحله‌ی آخر با توجه به این‌که بیشترین مقادیر STE و ZCR در محدوده‌ی تمرکز سیگنال قرار دارد پس ویژگی‌های موردنظر نیز از قاب‌های مربوط به این محدوده‌ها استخراج می‌گردد. نتایج شبیه‌سازی به‌خوبی نشان می‌دهد هر صوت دارای تعدادی محدوده با تمرکز سیگنال است که این محدوده‌ها از قاب‌هایی شروع و به قاب‌هایی ختم می‌شود. از طرفی تعداد نمونه‌های سیگنال در هر محدوده نیز مشخص است. حال این پارامترها علاوه بر مقادیر STE و ZCR نیز می‌توانند به‌عنوان ویژگی‌های کلیدی هر صوت با تفرانس حداقل ۰,۱ و حداکثر ۰,۲ جهت شناسایی و طبقه‌بندی خودرو استفاده شوند.

یکی از مزایای این روش، استفاده نکردن از قاب‌های فاقد اطلاعات مفید است که خود منجر به کاهش زمان محاسبات می‌گردد. از طرفی با استفاده از این روش توانسته‌ایم علاوه بر دو تکنیک اصلی استخراج ویژگی، از مشخصات آماری محدوده‌های متمرکز بودن سیگنال نیز به‌عنوان پارامترهای کلیدی صوت بهره ببریم که به‌نوبه خود منجر به شناسایی و طبقه‌بندی دقیق‌تر خودرو می‌شود. مزیت دیگر این روش زمانی است که اگر صدای خودرو هنگام ضبط کردن با نویز ناشی از صدای محیط ترکیب شده باشد به دلیل انعطاف‌پذیر بودن این روش، آشکارسازی نوع خودرو هنوز کارآمد است.

۲-۳-۲ تکنیک استخراج ویژگی از حوزه‌ی فرکانس

پژوهشگران در [۶] از دو تکنیک HLA و PSE جهت استخراج ویژگی از صدای سه نوع خودروی نظامی بهره برده‌اند. این پژوهش به سفارش پژوهشکده ارتش ایالات متحده آمریکا انجام گرفته است. در این پژوهش بعد از استخراج ویژگی‌های HLA و PSE از سیگنال‌های صوتی جهت شناسایی و طبقه‌بندی به یک شبکه هوشمند BPNN داده خواهند شد. در این شبکه، لایه ورودی بستگی به ویژگی‌های به کار برده شده ۱۱ تا ۲۵۰ گره، لایه مخفی ۱۵ گره و لایه بیرونی سه گره دارد. ایده‌ی اصلی نویسندگان در این پژوهش استفاده از مشخصات آماری و لرزشی سیگنال است که جهت بالا بردن نرخ شناسایی دقیق سیستم، آن‌ها را با ویژگی‌های HLA و PSE ترکیب می‌کنند.

نتایج شبیه‌سازی نشان می‌دهد، دقت شناسایی تکنیک HLA با این طبقه بند ۸۰٪ تا ۹۲٪ است درحالی‌که اگر مشخصات آماری با این تکنیک ترکیب شود دقت شناسایی حدود ۲٪ بالا می‌رود. از طرفی دقت شناسایی روش PSE با همین طبقه بند ۱۰٪ تا ۹۴٪ است که با ترکیب مشخصات آماری و لرزشی با آن نرخ شناسایی تغییر نخواهد کرد. از این رو می‌توان نتیجه گرفت، تکنیک PSE به دلیل نرخ شناسایی میانگین پایین، عملکرد ضعیفی از خود نشان داده و کارآمد نیست. یکی از معایب این روش‌ها، تنوع پایین خودروهاست که می‌تواند عملکرد سیستم را تحت شعاع خود قرار دهد.

محققین در [۷] از تکنیک HLA جهت استخراج ویژگی از سیگنال‌های صوتی استفاده کرده‌اند. در این پژوهش، پایگاه داده شامل صدای ۱۸ خودرو نظامی سنگین و نیمه سنگین است که از لحاظ قدرت موتوری به ۶ کلاس کاری تقسیم می‌شوند.

حال پژوهشگران جهت مقایسه و بررسی بهتر، دو ایده جدید داده‌اند. در ایده اول، از صدای تک تک خودروها ویژگی HLA استخراج شده و سپس جهت شناسایی و طبقه‌بندی به یک شبکه BPNN داده می‌شود. دقت شناسایی میانگین در این مرحله ۸۸٪ است. در ایده دوم، خودروهای هم قدرت در یک کلاس جمع‌آوری شده و در مجموع ۶ کلاس تشکیل می‌دهند. حال از سیگنال‌های صوتی هر کلاس،

ویژگی HLA را استخراج کرده و سپس جهت شناسایی و طبقه‌بندی به همان شبکه BPNN می‌دهیم. دقت شناسایی میانگین سیستم در این مرحله ۹۵٪ می‌شود. در پایان می‌توان نتیجه گرفت در ایده‌ی دوم به دلیل ادغام نمودن خودروهای هم قدرت توانسته‌ایم تعداد کلاس‌ها رو پایین آورده و احتمال تعلق پذیری اصوات را به یک کلاس زیاد کنیم.

نویسندگان در [۸] طیف صدای احتراق موتور، سه مدل خودروی پژو ۲۰۶ اس دی، تندر ۹۰ و سوزوکی گراند ویتارا را مورد بررسی قرار داده‌اند. نتایج شبیه‌سازی نشان می‌دهد فرکانس‌های غالب صدای خودروی پژو ۲۰۶ نسبت به دو مدل دیگر برای محدوده‌ی شنوایی انسان خطرناک‌تر است. از این رو برای حل این مشکل می‌توان با به کار بردن قطعه‌ای به نام رزوناتور^۱ اثر آن را تا حد قابل قبولی کاهش داد. در پایان پژوهش به این نتیجه می‌رسیم که با تجزیه و تحلیل سیگنال صوتی خودروها در حوزه‌ی فرکانس می‌توان آن دسته از صوت‌هایی که برای حدود شنوایی انسان (۲۰۰۰ تا ۵۰۰۰ هرتز) مخاطره‌آمیز بوده را شناسایی کرد و صوت مورد نظر را جهت شناسایی و طبقه‌بندی به یک سیستم شناسایی خودرو داد. یکی از معایب این روش، تعداد محدود خودروهای مورد بررسی و انجام آزمایش‌ها در شرایط ایده آل و ایستا بوده و ممکن است در شرایط حرکتی و طبیعی نتایج بهینه نباشد و عملکرد سیستم را تحت تأثیر خود قرار دهد.

پژوهشگران در [۹] از تکنیک FFT جهت استخراج ویژگی از سیگنال‌های صوتی بهره برده‌اند. پایگاه داده شامل صدای دو نوع خودرو نظامی یکی چرخ‌دار و دیگری ریلی است. ایده‌ی اصلی پژوهش این است که از سیگنال‌های صوتی و لرزشی مربوط به هر خودرو FFT گرفته و طیف توانی آن‌ها را به دست آوریم. طیف توانی مربوط به دو نوع سیگنال صوتی و لرزشی را شامل دو قسمت باند باریک و باند پهن در نظر گرفته و بردارهای ویژگی مناسبی از این دو قسمت استخراج می‌کنیم. حال شناسایی و طبقه‌بندی به‌طور جداگانه با توجه به بردارهای ویژگی باند پهن و باند باریک صورت می‌گیرد. نحوه‌ی

^۱ Resonator

شناسایی و طبقه‌بندی با بردارهای ویژگی باند باریک به این صورت است که بردارهای ویژگی باند باریک سیگنال صوتی و لرزشی هر خودرو به‌عنوان ویژگی‌های مشترک آن خودرو محسوب شده که جهت شناسایی و طبقه‌بندی خودرو به‌طور جداگانه به سه طبقه بند SVM، KNN و MLE داده خواهند شد. از طرفی شناسایی و طبقه‌بندی با بردارهای ویژگی باند پهن نیز فرآیند مشابهی مانند باند باریک دارد. با توجه به نتایج شبیه‌سازی، نرخ شناسایی میانگین خودروها برای حالت باند باریک به ازای طبقه بندهای SVM، KNN (K=1) و MLE به ترتیب ۰/۹۵/۵، ۰/۹۲/۳۷ و ۰/۸۴/۹۳ بوده که طبقه بند SVM بیشترین دقت را بین سایر دارد. از طرفی نرخ شناسایی میانگین خودروها برای حالت باند پهن نیز به ازای طبقه بندهای SVM، KNN (K=1) و MLE به ترتیب ۰/۹۴/۴۷، ۰/۹۰/۶ و ۰/۸۵ است که دوباره طبقه بند SVM بیشترین دقت را بین سایر دارد. در پایان می‌توان نتیجه‌گیری کرد که در دو حالت باند پهن و باند باریک به‌طور جداگانه، طبقه بند SVM در رتبه نخست، طبقه بند KNN (K=1) در رتبه دوم و MLE در رتبه سوم بالاترین دقت شناسایی قرار دارند. حال با مقایسه دقت طبقه‌بندی‌های نظیر به نظیر در دو حالت باند پهن و باند باریک به این نکته پی می‌بریم که دقت شناسایی میانگین طبقه بندها در حالت باند باریک بالاتر است.

در [۱۰] از تکنیک‌های MFCC و فرمنت‌ها جهت استخراج ویژگی از سیگنال‌های صوتی خودرو استفاده شده است. پایگاه داده در این پژوهش شامل صدای سه نوع خودروی دوچرخ، سه‌چرخ و سنگین‌وزن با تعداد ۱۰۰ نمونه صوتی به ازای هر خودرو است. نحوه عملکرد سیستم به این‌گونه است که از سیگنال صوتی هر خودرو ویژگی‌های MFCC و فرمنت‌ها استخراج شده و سپس جهت شناسایی و طبقه‌بندی خودرو به یک طبقه بند KNN (K=5) داده می‌شود. نتایج شبیه‌سازی نشان می‌دهد نرخ شناسایی میانگین با تکنیک MFCC، ۰/۸۱/۳۳ و با تکنیک ویژگی فرمنت‌ها ۰/۷۰/۳۳ است. از این‌رو می‌توان نتیجه گرفت روش MFCC نسبت به روش فرمنت‌ها کارایی بهتری دارد.

نویسندگان در [۱۱] از تکنیک AR mod جهت استخراج ویژگی از سیگنال صوتی خودروها بهره برده‌اند. پایگاه داده در این پژوهش شامل صدای ۴ نوع خودروی سواری در محیط شهری است. عملکرد الگوریتم به این صورت است که از سیگنال‌های صوتی بعد از پیش‌پردازش‌های لازم با استفاده از تکنیک AR mod استخراج ویژگی انجام می‌گیرد. حال جهت طبقه‌بندی ویژگی مخصوص به هر خودرو از طبقه‌بند TDNN استفاده می‌شود. نتایج شبیه‌سازی نشان می‌دهد نرخ شناسایی و طبقه‌بندی میانگین با این طبقه‌بند برای نمونه‌های آموزش و آزمایش به ترتیب برابر ۹۶٪ و ۸۴٪ است.

در [۱۲] از تکنیک MFCC جهت استخراج ویژگی از سیگنال صوتی خودروها استفاده شده است. پایگاه داده شامل صدای ۴ نوع خودروی سنگین، متوسط، سبک و لرزشی است. عملکرد الگوریتم به این صورت است که از صدای خودروها با استفاده از تکنیک MFCC ویژگی‌ها استخراج شده و سپس جهت طبقه‌بندی به یک طبقه‌بند MLFFNN داده می‌شود. طبقه‌بند با ۷۵٪ داده‌ها آموزش و با ۲۵٪ داده‌ها آزمایش می‌شود. نتایج شبیه‌سازی نشان می‌دهد دقت طبقه‌بندی میانگین خودروها برابر با ۶۷,۴٪ است. از طرفی دو نوع خودرو متوسط و سبک به دلیل تشابه موتوری دارای میزان خطای شناسایی قابل توجهی نسبت به هم هستند.

۲-۳-۳ تکنیک استخراج ویژگی از حوزه‌ی زمان - فرکانس

در [۲] از تکنیک‌های حوزه‌ی زمان - فرکانس جهت استخراج ویژگی از سیگنال‌های صوتی ۵ نوع خودرو استفاده شده است. تکنیک‌های حوزه‌ی زمان شامل STE، ZCR و RMS بوده و تکنیک‌های حوزه‌ی فرکانس نیز شامل CM و CSD است. ایده‌ی اصلی این پژوهش مقایسه ویژگی‌های استخراج‌شده از دو حوزه‌ی زمان و فرکانس نسبت به هم است. حال با نمونه‌برداری از سیگنال صوتی سه خودرو از پایگاه داده با فرکانس ۸ کیلوهرتز و دو خودرو دیگر با فرکانس ۱۶ کیلوهرتز، ارزیابی و مقایسه ویژگی‌ها صورت می‌گیرد. نتایج نشان می‌دهد در مقایسه ویژگی‌ها در حالت نمونه‌برداری با فرکانس ۸ کیلوهرتز، مقادیر ویژگی‌های فرکانسی سه خودرو نسبت به مقادیر ویژگی‌های زمانی از هم دورتر بوده که منجر به

بالا رفتن دقت شناسایی می‌شود. از طرفی در مقایسه به هنگام نمونه‌برداری با فرکانس ۱۶ کیلوهرتز، مقادیر ویژگی‌های زمانی و فرکانسی به هم نزدیک‌تر شده و باعث خطای شناسایی خواهد شد.

محقق در [۱۳] از تکنیک‌های فرکانسی (روش اول) و زمانی همبسته^۱ (روش دوم) جهت استخراج ویژگی از سیگنال صوتی استفاده کرده است. در روش دوم سیگنال صوتی را بر اساس سری ولترا می‌نویسیم. از طرفی چون در این سری ضرایب همبسته است، بنابراین سیگنال پارامتری می‌شود. جهت شناسایی و طبقه‌بندی از طبقه‌بند SVM استفاده شده است. نتایج شبیه‌سازی نشان می‌دهد نرخ شناسایی روش دوم ۷۷/۳۹٪ و روش اول ۷۴/۶۱٪ است. حال با ترکیب ویژگی‌های تکنیک اول و دوم (ترکیب اطلاعات^۲) نتیجه بهبود پیدا کرده و به دقت ۸۴/۰۲٪ می‌رسد. از این رو می‌توان نتیجه گرفت دقت شناسایی سیستم در حالت ترکیب اطلاعات عملکرد بهتری نسبت به دو تکنیک استخراج ویژگی دیگر دارد. عیب این روش محدودیت در تعداد خودروهاست.

در [۱۴] از دو تکنیک فرکانسی (توزیع طیفی) و زمان - فرکانسی (تبدیل موجک) جهت استخراج ویژگی از سیگنال صوتی خودروها استفاده می‌شود. برای ارزیابی این دو تکنیک، دو معیار نرخ طبقه‌بندی صحیح^۳ و نسبت تفکیک‌پذیری^۴ مدنظر است. جهت شناسایی و طبقه‌بندی نیز از دو طبقه‌بند SVM و KNN استفاده می‌شود. نتایج شبیه‌سازی نشان می‌دهد نرخ طبقه‌بندی صحیح با تکنیک فرکانسی به ازای طبقه‌بند SVM عملکرد بهتری نسبت به طبقه‌بند KNN دارد و از طرفی نیز نسبت به تکنیک زمان - فرکانس کارآمدتر است. عیب این روش محدودیت در تعداد خودروها بوده و ممکن است این نتایج به ازای طبقه‌بندها و تکنیک‌های مختلف بهینه نباشد.

¹ Time Correlated

² Information Fusion

³ Correct Classification Rate

⁴ Separability Ratio

در [۱۵] از دو تکنیک PCA و TVAR جهت استخراج ویژگی از سیگنال‌های صوتی خودروها استفاده شده است. پایگاه داده در این پژوهش شامل صدای شش خودرو نظامی است. جهت شناسایی و طبقه‌بندی خودروها از طبقه بند MLNN استفاده می‌شود. حال نرخ شناسایی این الگوریتم به ازای هر دو تکنیک ۸۳٪ تا ۹۵٪ است.

پیشنهاد نویسنده در [۱۶] استفاده از تکنیک DWT جهت استخراج ویژگی از سیگنال صوتی دو نوع خودروی نظامی است. نحوه عملکرد الگوریتم به این صورت است که ابتدا DWT با دو موجک هار^۱ و دابیشز^۲ ویژگی‌ها را از سیگنال صوتی خودرو استخراج می‌کند سپس ویژگی‌های استخراج شده با هر موجک به صورت جداگانه به یک طبقه بند MPP داده خواهد شد تا نوع خودرو را شناسایی و طبقه‌بندی کند. نتایج شبیه‌سازی نشان می‌دهد نرخ شناسایی الگوریتم به ازای موجک هار ۹۲/۹۸٪ و به ازای موجک دابیشز ۹۸/۲۵٪ است؛ بنابراین به این نتیجه می‌رسیم که موجک دابیشز ۴ نرخ شناسایی بهتری نسبت به موجک هار دارد.

محققین در [۱۷] از تکنیک DWT جهت استخراج ویژگی از سیگنال صوتی ۴ نوع خودرو نظامی بهره برده‌اند. نحوه عملکرد الگوریتم به این صورت است که سیگنال‌های صوتی ضبط شده را با فرکانس ۵ کیلوهرتز نمونه‌برداری کرده و سپس دیجیتال کنیم. حال سیگنال صوتی ایجاد شده را جهت ایستادن نمودن با ۵۰ درصد هم‌پوشانی قاب‌بندی می‌کنیم و در نهایت با استفاده از تکنیک DWT از هر قاب صوت، بردار ویژگی را استخراج می‌نماییم. جهت طبقه‌بندی ویژگی‌های مربوط به سیگنال صوتی خودروها از دو طبقه بند MLP و PNN استفاده می‌کنیم. نتایج شبیه‌سازی نشان می‌دهد نرخ شناسایی و طبقه‌بندی میانگین با طبقه بندهای MLP و PNN به ترتیب ۷۱٪ و ۷۳٪ است. پس می‌توان نتیجه گرفت طبقه بند PNN جهت شناسایی چهار خودرو مذکور کارآمدتر است.

¹ Haar

² Daubechies-4

در [۱۸] از تکنیک STFT جهت استخراج ویژگی از سیگنال صوتی خودروها استفاده شده است. پایگاه داده شامل صدای ۹ خودرو نظامی سبک و سنگین از نوع چرخ‌دار و ریلی است. عملکرد الگوریتم به این صورت است که از سیگنال صوتی خودروها با استفاده از تکنیک STFT استخراج ویژگی صورت می‌گیرد و سپس جهت طبقه‌بندی به یک طبقه بند MLNN داده می‌شود. حال نتایج شبیه‌سازی نشان می‌دهد اگر خودروها را به دو دسته‌ی سبک و سنگین تقسیم‌بندی کنیم نرخ شناسایی و طبقه‌بندی میانگین آن‌ها به ترتیب ۴۱٪ و ۸۷٪ است. از طرفی اگر به صورت خودروهای کشتی و چرخ‌دار تقسیم شوند در این حالت نرخ طبقه‌بندی میانگین نیز به ترتیب ۷۲٪ و ۷۱٪ خواهد شد.

۴-۲ نتیجه‌گیری

در اکثر مقالات مطالعه شده، شبیه‌سازی با طبقه‌بندی‌های متفاوت و روی تعداد محدودی خودرو صورت می‌گیرد که در صورت تنوع خودرو و طبقه‌بندی‌ها ممکن است این روش‌ها دیگر بهینه نباشد. از طرفی نرخ شناسایی خودروهای سنگین با الگوریتم‌های متفاوت نسبت به سایر خودروها بالاتر است زیرا این خودروها به دلیل تولید کردن سیگنال‌های صوتی و لرزشی به‌خوبی از دیگر خودروها قابل تمایز هستند. افزایش طول بردار ویژگی در تکنیک‌ها در بعضی مواقع علاوه بر زیاد شدن بار محاسباتی باعث کاهش نرخ شناسایی می‌شود. از این رو باید در انتخاب طول ویژگی‌ها دقت نمود. از طرفی در اکثر مقالات تکنیک زمان-فرکانسی به دلیل پردازش چند دقتی خود عملکرد بهتری نسبت به روش‌های زمانی و فرکانسی دارد. حال نیز از بین تمام طبقه‌بندی‌ها، طبقه‌بند SVM به‌خوبی در برابر نویز مقاوم بوده و معمولاً نرخ شناسایی و طبقه‌بندی بالایی دارد.

فصل سوم:

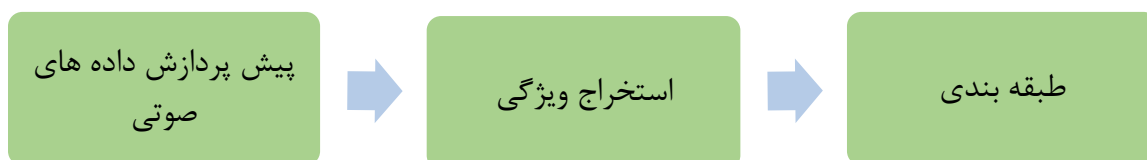
مبانی نظری

۱-۳ مقدمه

موتور خودرو و مکانیزم نیرو محرکه آن، دو منبع اصلی جهت تولید سیگنال صوتی خودروهاست. این سیگنال صوتی حاوی مقادیر زیادی اطلاعات مفید است که با استفاده از آن‌ها می‌توان نوع خودروها را شناسایی کرده و طبقه‌بندی نمود [۱۹].

۲-۳ توصیف قسمت‌های مختلف سیستم پیشنهادی

شکل (۱-۳) قسمت‌های مختلف سیستم پیشنهادی شناسایی و طبقه‌بندی خودرو با پردازش صدای آن را نشان می‌دهد.



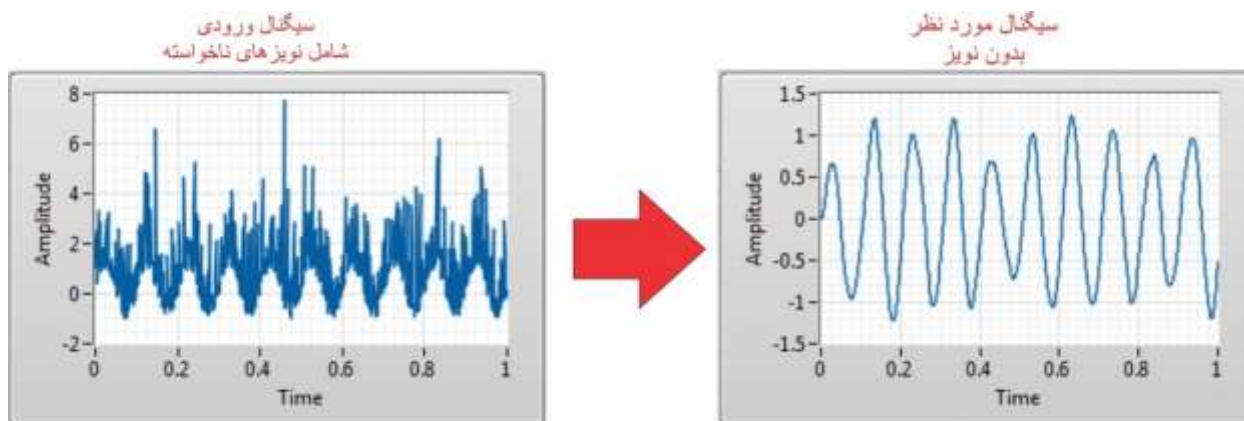
شکل ۱-۳ قسمت‌های مختلف سیستم پیشنهادی شناسایی و طبقه‌بندی خودرو با پردازش صدای آن

با توجه به شکل (۱-۳) این بخش از سه قسمت اساسی تشکیل شده است:

۱-۲-۳ پیش‌پردازش داده‌های صوتی

از سیگنال‌های صوتی جمع‌آوری شده که در واقع پایگاه داده مورد نیاز برای آموزش شبکه می‌باشند معمولاً به‌طور مستقیم و بدون هیچ‌گونه تغییر و پیش‌پردازی نمی‌توان برای آموزش شبکه استفاده کرد؛ زیرا با این کار، عملکرد شبکه به‌طور مستقیم تحت تأثیر قرار گرفته و شبکه به‌خوبی آموزش نمی‌بیند و دقت بازشناسی سیستم به‌شدت کاهش پیدا می‌کند. به همین منظور بر روی پایگاه داده قبل از اعمال به طبقات بعدی بازشناسی پیش‌پردازش‌هایی صورت می‌پذیرد. منظور از پیش‌پردازش به ترتیب عملیاتی شامل هم‌اندازه کردن طول سیگنال‌های صوتی، حذف DC سیگنال، از بین بردن قسمت‌های سکوت، حذف نویزهای پس‌زمینه، نمونه‌برداری از سیگنال جهت گسسته نمودن آن، نرمالیزه کردن نمونه‌ها به‌منظور از بین بردن داده‌های خارج از محدوده و درنهایت کوانتیزه کردن نمونه‌هاست.

شکل (۲-۳) حذف نویز از یک سیگنال صوتی را به‌عنوان یک نمونه پیش‌پردازش بر روی پایگاه داده صوتی نشان می‌دهد.



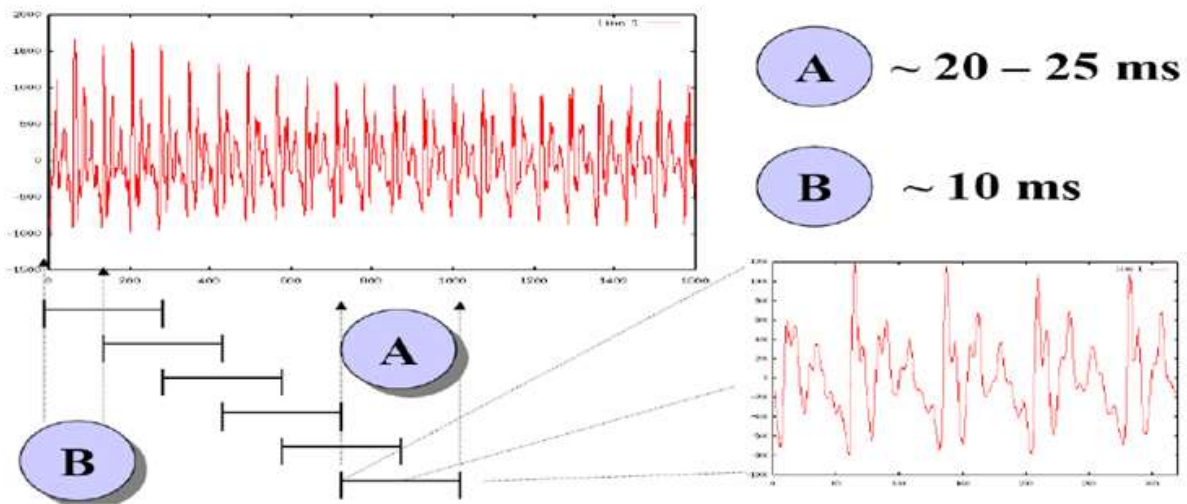
شکل ۲-۳ (چپ) سیگنال اصلی ورودی همراه با نویز (راست) سیگنال پردازش‌شده بدون نویز

همان‌طور که در شکل (۲-۳) مشاهده می‌شود، نویز سیگنال صوتی بعد از پیش‌پردازش لازم حذف شده است. در واقع هدف از این کار، کاهش زمان پردازش و افزایش دقت بازشناسی است.

۲-۲-۳ استخراج ویژگی

تمامی داده‌های صوتی که در قسمت قبل تولید شده‌اند، ناپیستان هستند یعنی مشخصات آماری آن‌ها با گذشت زمان تغییر می‌کند. از این‌رو برای تحلیل بهتر، سیگنال‌های صوتی را قاب‌گذاری نموده تا به‌صورت تقریباً ایستاد درآیند. برای این کار بین بردن اثر ناپیوستگی ابتدا و انتهای قاب‌ها در هنگام قاب‌گذاری که منجر به اعوجاج طیف آن‌ها می‌شود، تک‌تک قاب‌ها را در یک پنجره مناسب ضرب می‌کنند.

شکل (۳-۳) قاب گذاری یک سیگنال صوتی را نشان می دهد.



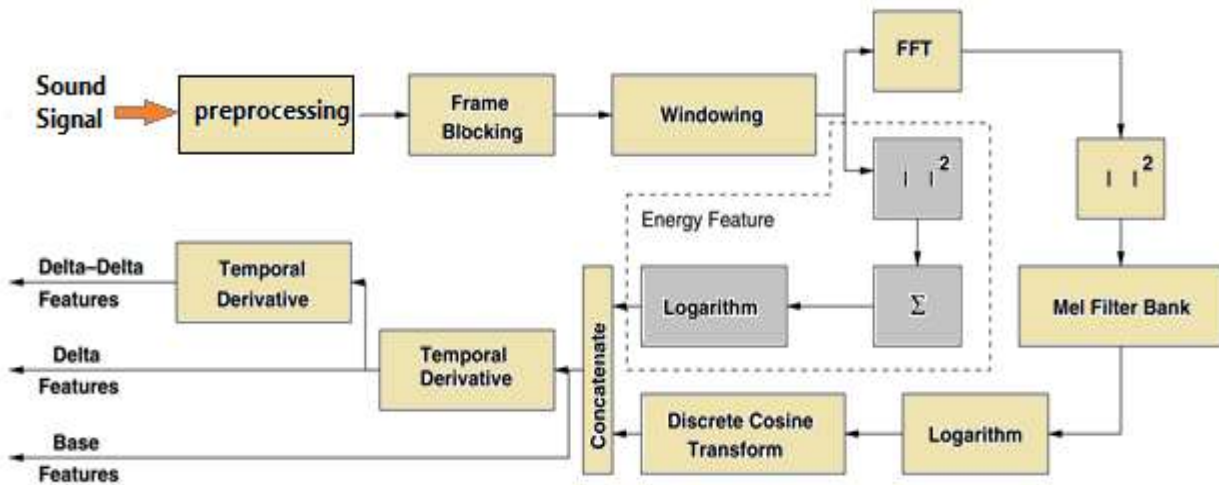
شکل ۳-۳ قاب گذاری سیگنال صوتی

شکل (۳-۳) یک سیگنال صوتی نا ایستان را نشان می دهد که جهت ایستان شدن، قاب گذاری شده است به طوری که قاب ها با هم ۵۰٪ هم پوشانی دارند. در مرحله بعد، از قاب های پنجره گذاری شده هر سیگنال صوتی تعدادی ویژگی مهم استخراج می گردد. در ادامه به تشریح چهار روش استخراج ویژگی استفاده شده در این پایان نامه می پردازیم:

۳-۲-۲-۱ روش استخراج ویژگی MFCC

روش استخراج ویژگی MFCC یک از پرکاربردترین روش های استخراج ویژگی در پردازش صوت است. این روش به دلیل استفاده کردن از فیلترهای Mel می تواند سیستم ادراک شنوایی انسان را به خوبی مدل کند. ویژگی های استخراج شده توسط این روش کمترین شباهت را به هم دارند. از این رو روش مطرح شده در محیط های تمیز دقت بازشناسی بالایی خواهد داشت. یکی از معایب روش استخراج ویژگی MFCC این است که در محیط های نویزی مقاومت کمی از خود نشان داده و دقت بازشناسی آن به شدت کاهش می یابد و از طرفی نیز به دلیل نادیده گرفتن فاز، نمایش محدودی را ایجاد می کند [۲۰، ۲۱].

شکل (۳-۴) بلوک دیاگرام روش استخراج ویژگی MFCC را نشان می‌دهد [۲۱].



شکل ۳-۴ بلوک دیاگرام قسمت‌های مختلف الگوریتم استخراج ویژگی MFCC [۲۱]

همان‌طور که در شکل (۳-۴) مشاهده می‌کنیم ابتدا سیگنال صوتی به منظور پیش‌پردازش‌های لازم وارد بلوک پیش‌پردازش می‌شود. در این بلوک عملیاتی هم چون حذف DC سیگنال و نرمالیزه کردن به مقدار میانگین صورت می‌پذیرد. سپس سیگنال پیش‌پردازش‌شده جهت ایستادن شدن وارد بلوک قاب‌بندی می‌شود تا به قاب‌های کوچک‌تری تقسیم‌بندی گردد. پس از آن سیگنال قاب‌بندی شده وارد بلوک پنجره‌گذاری می‌شود. در این بلوک تک‌تک قاب‌های سیگنال به منظور حذف اعوجاج ابتدا و انتهای آن‌ها در یک پنجره مناسب ضرب می‌شوند. سیگنال قاب‌بندی شده بعد از پنجره‌گذاری در دو مسیر قرار می‌گیرند. در مسیر اول که همان مسیر فوقانی است طیف قاب‌های سیگنال با استفاده از روش FFT استخراج شده و با عبور از بلوک بعد، اندازه آن‌ها محاسبه می‌گردد. سپس با عبور از بلوک فیلترهای Mel و در نهایت از دو بلوک لگاریتم‌گیر و تبدیل کسینوسی گسسته^۱ ویژگی‌های MFCC استخراج می‌گردد؛ اما در مسیر دوم از قاب‌های سیگنال صوتی، لگاریتم انرژی استخراج شده و به‌عنوان ویژگی‌های کمکی در کنار ویژگی‌های MFCC قرار می‌گیرد. حال این دو ویژگی در کنار هم ویژگی‌های MFCC پایه نامیده می‌شوند. معمولاً در کنار ویژگی‌های MFCC پایه از Delta MFCC و Delta Delta MFCC

^۱ Discrete Cosine Transform(DCT)

که همان مشتق‌های مرتبه اول و دوم ویژگی‌های MFCC پایه بوده نیز به‌عنوان ویژگی‌های اضافی به‌منظور بازشناسی قدرتمندتر استفاده می‌گردد.

در بیشتر کاربردها تعداد ویژگی‌های MFCC ۳۹ ضریب در هر قاب در نظر گرفته می‌شود، به این صورت که ۱۳ ضریب ویژگی MFCC پایه و از هر کدام از ویژگی‌های Delta MFCC و Delta Delta MFCC نیز ۱۳ ضریب ویژگی استخراج می‌گردد. از طرفی این تعداد ضرایب ویژگی را نمی‌توان در تمامی کاربردها عمومیت داد و بستگی به کاربرد، این تعداد ویژگی‌ها تغییر می‌کند [۲۲].

۲-۲-۲-۳ روش استخراج ویژگی LPC

تکنیک LPC یکی از روش‌هایی است که غالباً در استخراج ویژگی برای تحلیل سیگنال مورد استفاده قرار می‌گیرد. این روش یکی از روش‌های استخراج ویژگی از حوزه زمان به شمار می‌آید و از طرفی نیز، یک روش قوی‌تری برای تحلیل فایل‌های صوتی کد شده با کیفیت بهتر و نمونه‌هایی با نرخ بیت کمتر است [۲۳، ۲۴].

مزایای متعددی برای استفاده از روش LPC وجود دارد که از این بین می‌توان به موارد زیر اشاره نمود [۲۳، ۲۴]:

- روش LPC زمان محاسباتی کوتاه‌تر و کارایی‌تری برای پارامترهای سیگنال به دست می‌دهد.
 - این روش قادر است مشخصات مهم سیگنال ورودی را به‌خوبی استخراج کند.
- هدف اصلی این روش، تقریب درستی از نمونه صوتی دریافت شده به‌صورت ترکیبی خطی از نمونه‌های قبلی صوت است که آن را مطابق رابطه (۱-۳) نشان می‌دهیم [۲۳، ۲۴].

$$S(i) = a_1 * S(i - 1) + a_2 * S(i - 2) + \dots + a_p * S(i - p) \quad (1-3)$$

$S(i)$ دامنه‌ی نمونه در لحظه i ام و a_i ها ضرایب پیش‌بینی خطی هستند. معمولاً مقدار p عددی بین ۱۰ تا ۱۲ در نظر گرفته می‌شود.

در یک سیستم LTI می توان رابطه ی بین ورودی و خروجی را به صورت یک معادله تفاضلی مطابق رابطه (۲-۳) ارائه نمود.

$$\sum_{i=0}^p a(i) y(n-i) = \sum_{j=0}^q b(j) x(n-j) \quad (2-3)$$

در حالت کلی، LP^۱ تخمین خروجی در یک لحظه ی خاص به ازای خروجی های قبلی و ورودی زمان های مختلف مطابق رابطه (۳-۳) است. مدل تمام قطب (AR^۲)، مدل تمام صفر (MA^۳) و مدل صفر و قطب (ARMA^۴) به عنوان سه مدل مختلف برای رابطه (۳-۳) تعریف می شود.

$$y(n) = \sum_{j=0}^q b(j)x(n-j) - \sum_{i=1}^p a(i)y(n-i) \quad (3-3)$$

a(i) و b(j) ضرایب پیشگویی خطی، x(n) و y(n) به ترتیب ورودی و خروجی در لحظه n هستند.

در این پایان نامه، ما برای سیگنال صوت از مدل تمام قطب (AR) استفاده می کنیم. ایده اولیه روش LPC این است که نمونه فعلی سیگنال صوتی از ترکیب خطی نمونه های قبلی تشکیل می شود. پس با توجه به این نکته می توان مدل تمام قطب را به صورت رابطه (۴-۳) بیان نمود.

$$H(z) = \frac{1}{1 - \sum_{k=1}^p a_k z^{-k}} \quad (4-3)$$

H(z) تابع تبدیل سیستم تمام قطب و a_kها ضرایب پیشگویی خطی یا ضرایب LPC نامیده می شوند. در این روش با مینیمم کردن مجموع مربعات تفاضل (در یک بازه ی محدود) بین نمونه های صوت واقعی و مقادیر محاسبه شده یک مجموعه پارامتر منحصر به فرد با ضرایب پیش بینی شده به دست می آید. مقادیر ضرایب پیش بینی شده بلوک های سازنده LPC صوت را تشکیل می دهد [۲۳، ۲۴].

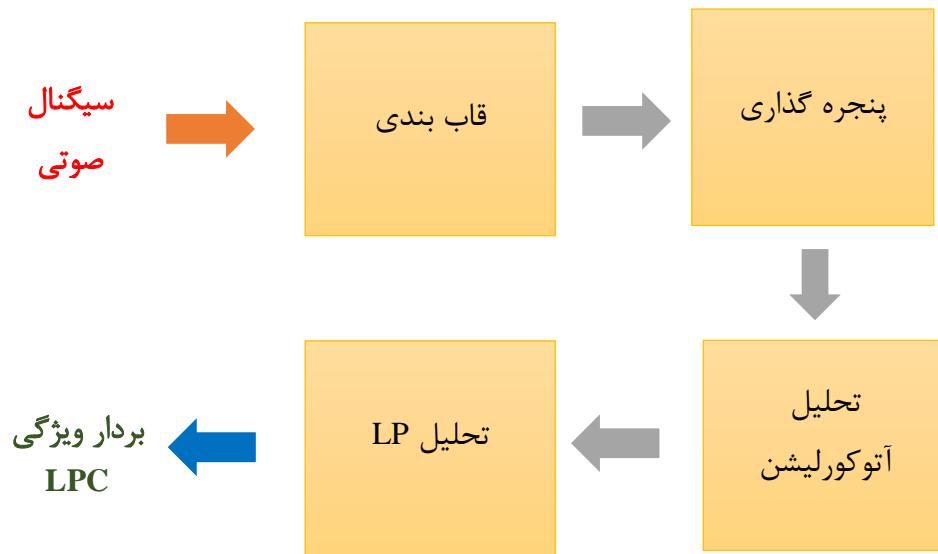
¹ Linear Prediction(LP)

² Autoregressive Model(AR)

³ Moving-Average Model(MA)

⁴ Autoregressive-Moving-Average Model(ARMA)

شکل (۵-۳) بلوک دیاگرام قسمت‌های مختلف الگوریتم استخراج ویژگی LPC را نشان می‌دهد [۲۵].



شکل ۵-۳ بلوک دیاگرام قسمت‌های مختلف الگوریتم استخراج ویژگی LPC [۲۵]

۳-۲-۲-۳ روش استخراج ویژگی LSF

LP یک روش کارآمد در بسیاری از کاربردهای صوت و گفتار نظیر بازشناسی، فشرده‌سازی و مدل کردن محسوب می‌شود [۲۶,۲۷]. مبتنی بر این روش، یک سیگنال صوتی را می‌توان مطابق رابطه (۵-۳) نمایش داد.

$$x_k = \sum_{i=1}^p a_i x_{k_i} \quad (۵-۳)$$

در رابطه فوق، k اندیس زمان و p درجه LP یا به عبارتی تعداد ضرایب LP به کار برده شده است. در این رابطه، ضرایب a_i از طریق مینیمم کردن خطای تخمین توسط دو روش کوواریانس و اتوکورلیشن به دست می‌آیند. رابطه (۵-۳) را می‌توان در حوزه فرکانس با تبدیل Z بازنویسی کرد [۲۸]. از این رو یک قطعه کوچکی از سیگنال صوتی می‌تواند خروجی یک فیلتر تمام قطب مطابق رابطه (۶-۳) باشد.

$$H(z) = \frac{1}{A(z)} = \frac{1}{1 - \sum_{i=1}^p (a_i z^{-i})} \quad (۶-۳)$$

$A(z)$ معکوس فیلتر تمام قطب است.

نمایش تکنیک LSF توسط Itakura معرفی گردید [۲۹]. این روش یک راه جایگزین برای نشان دادن ضرایب LP است. به منظور محاسبه ضرایب LSF چندجمله‌ای $A(z)$ به دو چندجمله‌ای مطابق با رابطه‌ها (۷-۳) و (۸-۳) تقسیم می‌گردد.

$$P(z) = A(z) + z^{p+1} A(z^{-1}) \quad (7-3)$$

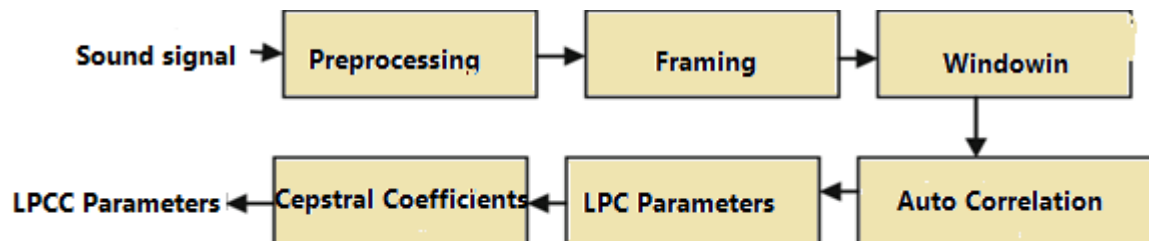
$$Q(z) = A(z) - z^{p+1} A(z^{-1}) \quad (8-3)$$

$P(z)$ یک چندجمله‌ای متقارن و $Q(z)$ یک چندجمله‌ای پادمقارن است. ریشه‌های $P(z)$ و $Q(z)$ ضرایب LSF را مشخص می‌کنند. LSF برای کاربردهای کونتیزاسیون و درون‌یابی بسیار مناسب است. از این رو روش LSF نسبت به روش LP توانا تر است تا یک سیگنال صوتی را به تعداد کم‌تری از ضرایب نگاشت دهد [۳۰].

۴-۲-۲-۳ روش استخراج ویژگی LPCC

تخمین پارامترهای مشترک سیگنال صوتی مانند فرکانس گام، انرژی قاب و فرمنت‌ها روش LPCC را به یکی از مهم‌ترین روش‌های استخراج ویژگی تبدیل کرده است. هدف از استخراج ویژگی، نشان دادن سیگنال صوتی با تعداد محدودی از معیارهای سیگنال است. با کمک LPC و روش اتوکورلیشن می‌توان ضرایب LPCC که بیشتر به ضرایب کپسترال ترجمه شده است را به دست آورد. روش LPCC در واقع مشتق ضرایب LPC است، از این رو LPC اولین گام برای یافتن LPCC است. به عبارتی LPCC به معنای LPC در دامنه کپستروم است [۳۱، ۳۲].

بلوک دیاگرام روش استخراج ویژگی LPCC در شکل (۶-۳) نشان داده شده است [۳۳].



شکل ۶-۳ بلوک دیاگرام روش استخراج ویژگی LPCC [۳۳]

مراحل محاسبه ضرایب LPCC عبارت‌اند از:

الف) پیش‌پردازش

بر روی سیگنال صوتی عملیاتی شامل هم طول کردن سیگنال، حذف مقدار DC، حذف نویز، نمونه‌برداری از سیگنال، نرمالیزه کردن و کوانتیزه نمودن داده‌ها صورت می‌گیرد.

ب) قاب بندی

بعد از مرحله پیش‌پردازش، سیگنال صوتی قاب بندی می‌شود. مدت‌زمان استاندارد همپوشانی بین دو قاب متوالی ۱۰ میلی‌ثانیه است.

ج) پنجره گذاری

قاب‌های ایجادشده در بلوک قبل، به‌منظور کاهش اثرات لبه در پنجره همینگ ضرب می‌شوند.

د) محاسبه LPC

در این مرحله بر روی قاب‌های عبور کرده از طبقات قبل تابع اتوکرولیشن اعمال می‌گردد. ضرایب LPCC با معرفی ضرایب کپستروم در پارامترهای LPC محاسبه می‌شوند. اگر بتوانیم بردار ضرایب پیش‌بینی را پیدا کنیم محاسبه ضرایب کپستروم آسان می‌شود.

۳-۲-۳ طبقه‌بندی

ویژگی‌های استخراج‌شده از داده‌های صوتی به دو دسته آموزش و آزمایش دسته‌بندی می‌شوند. جهت طبقه‌بندی و شناسایی خودرو، ویژگی‌های آموزش به یک طبقه بند مناسب داده می‌شود تا طبقه بند با این ویژگی‌ها آموزش ببیند. سپس طبقه بند با استفاده از ویژگی‌های آزمایش ارزیابی می‌شود.

۳-۲-۳-۱ طبقه بند SVM

طبقه بند SVM امروزه به‌عنوان یک ابزار بسیار قدرتمند جهت پیاده‌سازی الگوریتم‌های یادگیری محسوب می‌شود. از طرفی نیز با توجه به ساختار ساده، عدم نیاز به عملیات ریاضی پیچیده و سبک بودن بار محاسباتی آن عملکرد بهتری نسبت به روش‌های قدیمی‌تر نظیر شبکه عصبی پرسپترون از خود نشان داده است [۳۴].

طبقه بندهای SVM به طور کلی به دودسته Hard SVM و Soft SVM تقسیم می‌شوند. در طبقه بندهای Hard SVM حاشیه اطمینان باید به نحوی تعیین گردد که هیچ‌کدام از داده‌ها اشتباه طبقه‌بندی نشوند. از این رو مبنای کاری این طبقه بند دسته‌بندی کردن خطی داده‌ها است. در تقسیم خطی داده‌ها سعی بر انتخاب خطی آن‌هاست تا حاشیه اطمینان ماکزیمم گردد. معادله مرز بین دو کلاس را می‌توان به صورت رابطه (۹-۳) نوشت و برای تحقق هدف بیشینه شدن حاشیه اطمینان بین دو کلاس رابطه (۱۰-۳) باید کمینه گردد [۳۵].

$$y = w \cdot x + b \quad (9-3)$$

$$L_p = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n a_i [y_i (w \cdot x_i - b) - 1] \quad (10-3)$$

رابطه (۱۰-۳) مسئله لاگرانژ نامیده می‌شود. W بردار وزن بردار عمود بر ابرصفحه، a_i ها ضرایب لاگرانژ، b میزان انحراف و x_i بردار ورودی‌هاست.

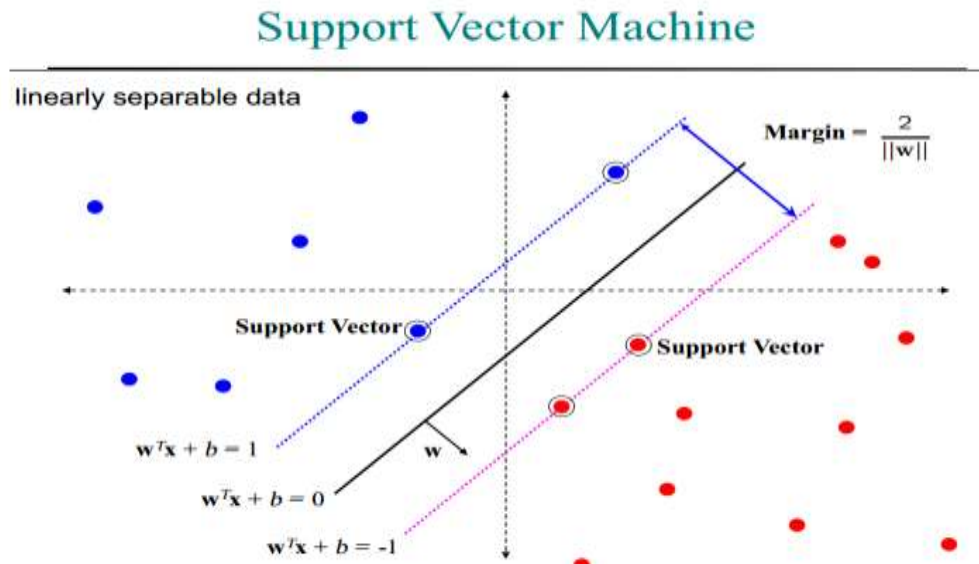
برای کمینه کردن رابطه (۱۰-۳) باید قسمت دوم رابطه یعنی همان سیگما بیشینه گردد. از این رو به دنبال α_i هایی هستیم که این قسمت را بیشینه نماید. به ازای هر کدام از داده‌ها یک α_i محاسبه می‌گردد که α_i های غیر صفر بیان‌کننده داده‌های مرزی یا همان بردارهای پشتیبان هستند. پس از محاسبه α_i ها می‌توان w و b را به ترتیب با استفاده از روابط (۱۱-۳) و (۱۲-۳) محاسبه نمود [۳۵].

$$w = \sum_{i=1}^{N_{sv}} a_i y_i x_i \quad (11-3)$$

$$b = \frac{1}{N_{sv}} \sum_{i=1}^{N_{sv}} (w \cdot x_i - y_i) \quad (12-3)$$

N_{sv} تعداد بردارهای پشتیبان است.

شکل (۷-۳) تابع جداکننده، بردارهای پشتیبان و حاشیه اطمینان را در یک طبقه بند SVM نشان می‌دهد [۳۵].



شکل ۷-۳ تابع جداکننده، بردارهای پشتیبان و حاشیه اطمینان در یک طبقه بند SVM [۳۵]

حال در صورتی که داده‌ها به صورت خطی تفکیک پذیر نشوند از طبقه بندهای Soft SVM بهره می‌جوئیم. اما برای سادگی و حذف پیچیدگی این طبقه بند، داده‌های غیرقابل تفکیک شده را با استفاده از یک تابع انتقال از فضای غیرخطی به فضای خطی نگاشت می‌دهیم. لذا رابطه‌ی مرز به صورت رابطه تعمیم‌یافته (۱۳-۳) تبدیل می‌شود [۳۵].

$$g(x) = w^t Q(x) + b = \sum_{i \in SV} a Q(x_i)^t Q(x) + b = \sum_{i \in SV} a k(x_i, x) + b \quad (13-3)$$

$k(x_i, x)$ تابع کرنل است.

فصل چهارم:

روش پیشنهادی و نتایج شبیه‌سازی

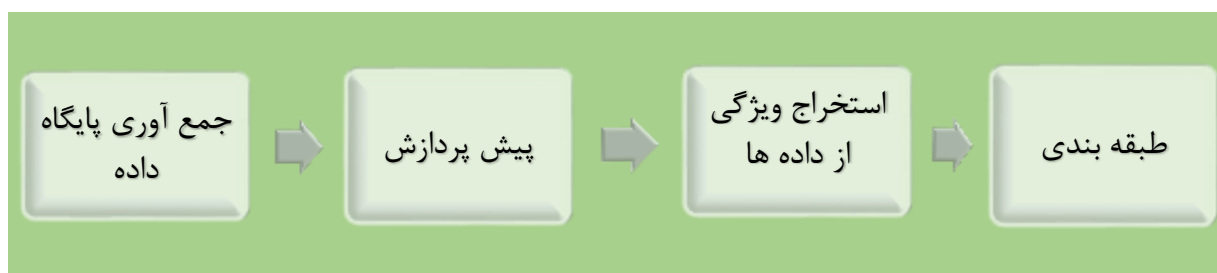
۱-۴ مقدمه

در این فصل ابتدا روش پیشنهادی جهت شناسایی و طبقه‌بندی خودرو با استفاده از صدای آن بیان می‌شود و سپس در پایان فصل، نتایج حاصل از انجام این روش پیشنهادی آورده خواهد شد.

۲-۴ سیستم پیشنهادشده جهت بازشناسی هوشمند خودرو با پردازش صدای آن

جهت شناسایی خودکار وسایل نقلیه با استفاده از پردازش صدای تولیدی هرکدام از آن‌ها نیازمند یک سیستم شناسایی و طبقه‌بندی هوشمند بر پایه سیگنال صوتی هستیم.

شکل (۱-۴) بلوک دیاگرام سیستم پیشنهادشده جهت بازشناسی هوشمند خودرو با پردازش صدای آن را نشان می‌دهد.



شکل ۱-۴ بلوک دیاگرام سیستم پیشنهادشده جهت بازشناسی هوشمند خودرو با پردازش صدای آن

با توجه به شکل (۱-۴) سیستم شناسایی هوشمند وسایل نقلیه از ۵ مرحله اساسی تشکیل شده است:

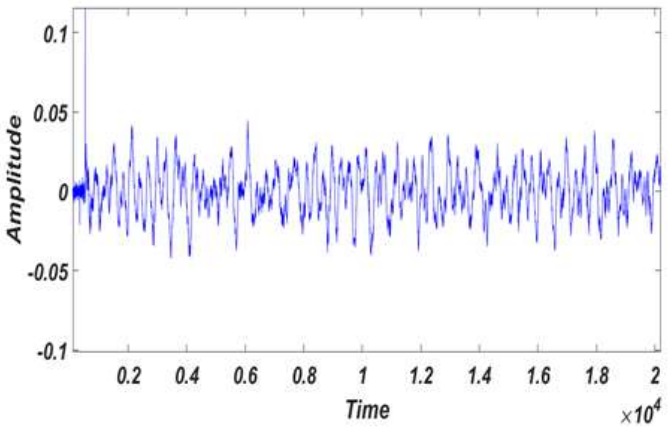

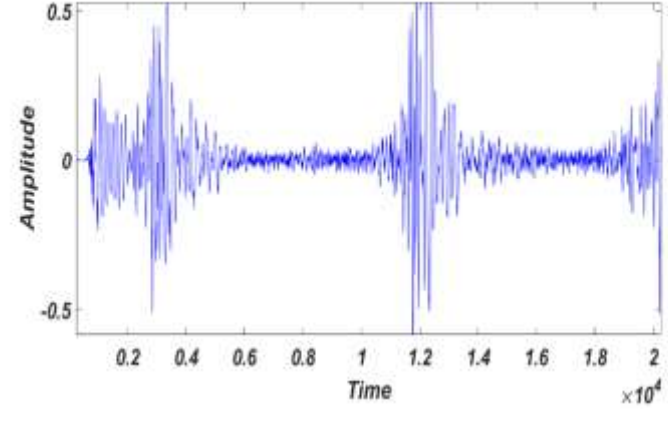

۱-۲-۴ جمع‌آوری پایگاه داده

اولین قدم در پیاده‌سازی این سیستم، جمع‌آوری پایگاه داده صوتی است. پایگاه داده را از سیگنال‌های صوتی ۶ کلاس خودرو سواری BMW 540i, Ferrari 308, Ford Tempob, Jeep, Toyota yaris و Volvo Amazon که از هر کلاس ۵۱ نمونه صوتی در اختیار بوده جمع‌آوری نموده‌ایم.

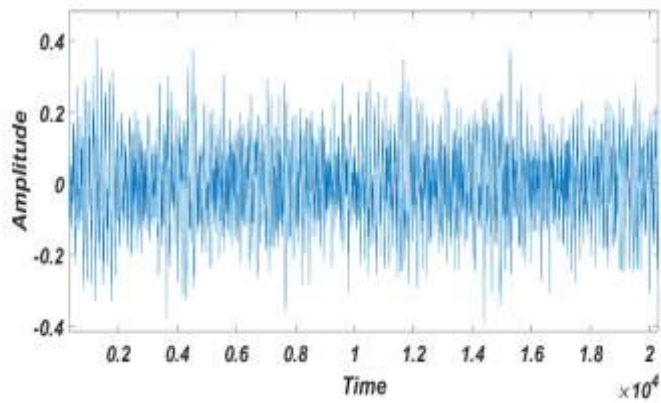
سیستم پیاده‌سازی شده با هر تعداد کلاس سازگار است، به این معنا که در صورت افزایش تعداد کلاس‌ها این سیستم باز هم به‌درستی کار خواهد کرد.

تصاویر خودروهای پایگاه داده صوتی به همراه نمودار سیگنال صوتی آن‌ها در جدول (۱-۴) گردآوری شده است.^۱

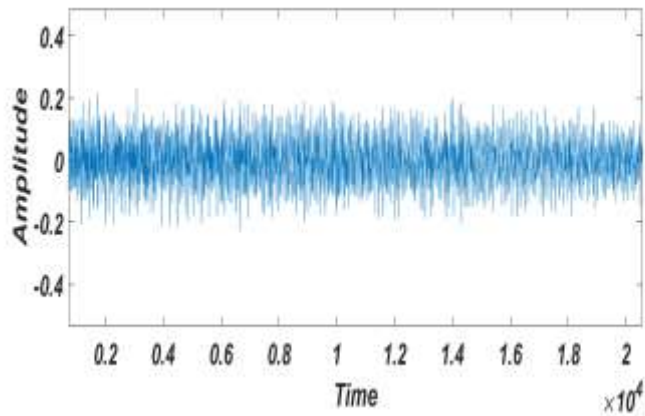
جدول ۱-۴ تصاویر به همراه نمودار سیگنال صوتی خودروهای مربوط به پایگاه داده صوتی

نمودار سیگنال خودرو مربوطه	تصویر خودرو
	 <p data-bbox="1053 1008 1228 1041">Toyota Yaris</p>
	 <p data-bbox="1053 1556 1220 1590">BMW 540i</p>

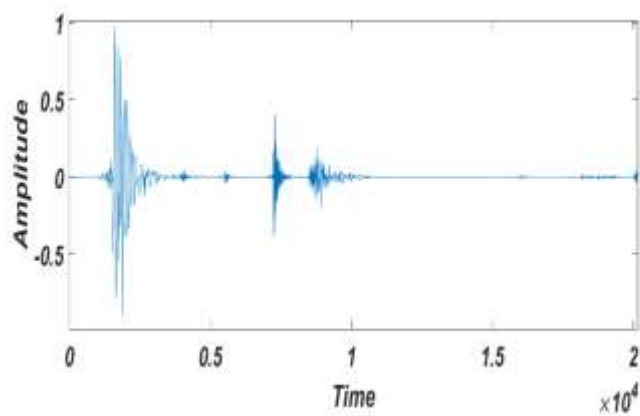
¹ URL <http://www.sounddogs.com>



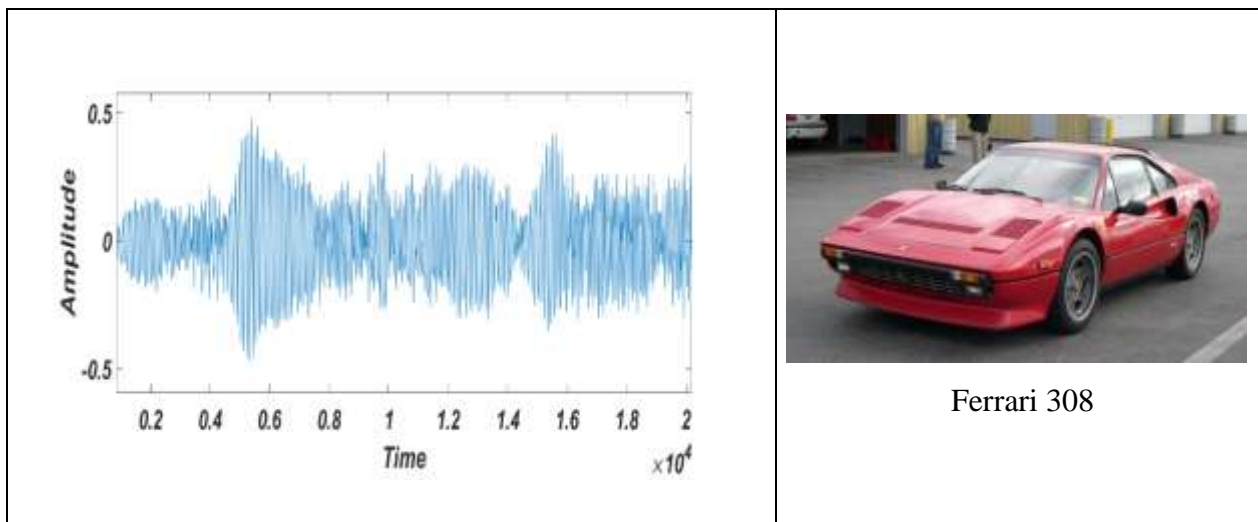
Volvo Amazon



Jeep Military



Ford Tempo



۲-۲-۴ پیش پردازش داده‌های صوتی

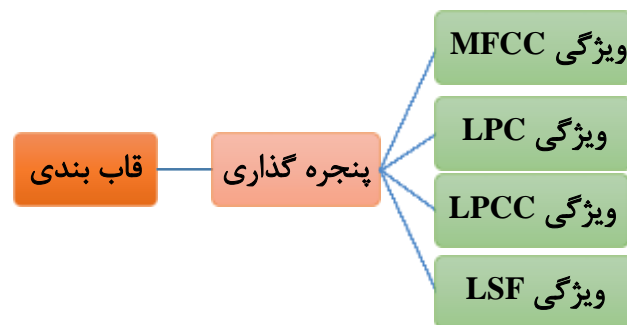
دومین مرحله در پیاده‌سازی سیستم شناسایی هوشمند خودرو، پیش‌پردازش داده‌های صوتی است. در واقع پایگاه داده استاندارد برای بازشناسی هوشمند خودرو وجود ندارد و آن‌ها را از طریق سایت‌های اینترنتی مختلف جمع‌آوری نموده‌ایم. به همین دلیل این داده‌ها می‌توانند اختلافات زیادی با یکدیگر از قبیل محیط ضبط صدای متفاوت، فرکانس نمونه‌گیری متفاوت، طول متفاوت، وجود قسمت‌های سکوت زیاد، دامنه سیگنال کم یا زیاد و ... داشته باشند. از این رو قبل از استفاده از پایگاه داده جمع‌آوری شده بر روی آن‌ها پیش‌پردازش‌هایی شامل حذف نویز، هم‌اندازه کردن طول نمونه‌ها، نمونه‌برداری، نرمالیزه سازی و کوانتیزه کردن داده‌ها صورت می‌گیرد.

جهت پیش‌پردازش ابتدا نویز پس‌زمینه را از تک‌تک سیگنال‌های صوتی حذف می‌کنیم سپس طول همه‌ی سیگنال‌ها را برابر ۲ ثانیه در نظر گرفته و از هریک با فرکانس ۱۰ کیلوهرتز نمونه‌برداری می‌شود و در پایان داده‌های صوتی ۲۰۰۰۰ نمونه‌ای خواهیم داشت. از طرفی جهت مقادیر نمونه‌های خارج از محدوده، تک‌تک مقادیر نمونه‌های یک داده صوتی را به بزرگ‌ترین مقدار نمونه آن تقسیم می‌کنیم تا عملیات نرمالیزاسیون انجام گیرد. حال جهت دیجیتال کردن نمونه‌های صوتی، همه‌ی داده‌های صوتی را با ۸ بیت کوانتیزه می‌کنیم.

۳-۲-۴ استخراج ویژگی

سومین مرحله در پیاده‌سازی سیستم شناسایی هوشمند خودرو، استخراج ویژگی از داده‌های صوتی است. در این پایان‌نامه از ۴ روش استخراج ویژگی MFCC، LPC، LPCC و LSF استفاده می‌شود.

شکل (۲-۴) بلوک دیاگرام استخراج ویژگی‌های مختلف از سیگنال صوتی را نشان می‌دهد.



شکل ۲-۴ بلوک دیاگرام استخراج ویژگی‌های مختلف از سیگنال‌های صوتی

همان‌طور که در شکل (۲-۴) مشاهده می‌کنیم ابتدا تک‌تک سیگنال‌های صوتی پایگاه داده بعد از عملیات پیش‌پردازش وارد بلوک قاب‌بندی می‌شوند. در این بلوک سیگنال‌های صوتی به قاب‌هایی به طول ۲۵ میلی‌ثانیه با ۱۰ میلی‌ثانیه هم‌پوشانی بین هر دو قاب تبدیل خواهند شد. سپس قاب‌های ایجادشده از بلوک پنجره‌گذاری عبور می‌کنند. در این بلوک همه‌ی قاب‌ها جهت از بین بردن اعوجاج‌های ابتدا و انتهای آن‌ها در پنجره همینگ ضرب خواهند شد. پس از آن، قاب‌های پنجره‌گذاری شده وارد طبقات استخراج ویژگی می‌شوند. در این طبقات با توجه به تعداد ویژگی‌های موردنیاز، از هر قاب ویژگی استخراج می‌گردد. در این پایان‌نامه جهت ارزیابی بهتر، ۴ روش استخراج ویژگی MFCC، LPC، LPCC و LSF با سه طول مختلف ۱۳، ۲۶ و ۳۹ ضریب در هر قاب را مورد بررسی قرار داده‌ایم. در پایان این مرحله، ویژگی‌های تمامی داده‌های صوتی پایگاه داده استخراج شده و جهت آموزش و آزمایش طبقه‌بند آماده می‌گردد.

۴-۲-۴ طبقه‌بندی

در این مرحله، ویژگی‌های تولیدشده از مرحله قبل به دو دسته آموزش و آزمایش تقسیم می‌شوند. داده‌های آموزش وظیفه آموزش دادن طبقه بند و داده‌های آزمایش نیز وظیفه‌ی آزمایش و ارزیابی طبقه بند را بر عهده دارند.

در این پایان‌نامه از ماشین بردار پشتیبان یا همان SVM به‌عنوان طبقه بند استفاده شده است. به دلیل توزیع غیرخطی داده‌های آموزش، کرنل خطی نمی‌تواند طبقه‌بندی را به‌درستی انجام دهد. از این رو از یک کرنل چندجمله‌ای با درجه ۴ مطابق رابطه (۱-۴) برای شبکه SVM استفاده شده است.

$$\text{kernel} = (\text{Gamma} * U' * V + \text{Coef})^4 \quad (1-4)$$

در رابطه فوق $\text{Gamma} = 1/\text{Number of Feature}$ ، u' و V به ترتیب ورودی و خروجی طبقه بند و Coef یک عدد ثابت است که ما آن را در این پایان‌نامه صفر در نظر گرفته‌ایم.

طبقه بند ابتدا با ۷۵ درصد کل داده‌های صوتی پایگاه داده آموزش می‌بیند و سپس با ۲۵ درصد کل داده‌های باقی‌مانده آزمایش می‌شود. در مرحله آزمایش هرچه دقت بازشناسی بیشتر باشد می‌توان گفت تکنیک‌های استخراج ویژگی با آن تعداد ضرایب ویژگی در هر قاب برای بازشناسی هوشمند خودرو مناسب‌تر است.

۳-۴ نتایج حاصل از شبیه‌سازی انجام روش پیشنهادی

در این بخش نتایج حاصل از پیاده‌سازی سیستم شناسایی هوشمند خودرو با استفاده از پردازش صدای آن ارائه می‌شود. لازم به ذکر است که تمامی پیاده‌سازی‌های انجام‌شده در نرم‌افزار MATLAB صورت پذیرفته است.

شبیه‌سازی‌های انجام‌شده در این پایان‌نامه در دو قسمت ارائه می‌گردد. ابتدا ماتریس درهم‌ریختگی^۱ هر چهار روش استخراج ویژگی ارائه شده است. سپس نتایج حاصل از شبیه‌سازی به ازای هر چهار روش استخراج ویژگی با سه طول متفاوت برای ویژگی‌ها نشان داده می‌شود.

۱-۳-۴ ماتریس درهم‌ریختگی

این ماتریس در واقع مبین میزان بازشناسی صحیح و اشتباه برای هر کلاس است. این ماتریس یک ماتریس مربعی با سایز $n * n$ بوده که n تعداد کلاس‌ها است. میزان دقت کلاس‌بندی صحیح هر کلاس بر روی درایه‌های قطر اصلی و کلاس‌بندی‌های ناصحیح بر روی درایه‌های دیگر قرار گرفته‌اند. برای یافتن دقت بازشناسی از روی ماتریس درهم‌ریختگی می‌توان از رابطه (۲-۴) استفاده نمود.

$$\text{دقت بازشناسی (\%)} = \frac{\sum_1 \text{تعداد کلاس‌ها} \text{ قطر اصلی}}{\sum_1 \text{تعداد کلاس‌ها} \text{ ماتریس}} \quad (2-4)$$

همان‌طور که در رابطه فوق مشاهده می‌شود دقت بازشناسی از طریق تقسیم حاصل جمع درایه‌های روی قطر اصلی به حاصل جمع تمامی درایه‌های ماتریس به دست می‌آید.

ماتریس‌های درهم‌ریختگی مربوط به روش‌های استخراج ویژگی MFCC، LPCC، LPC و LSF (با ۱۳ ضریب در هر قاب) به ترتیب در جدول‌های (۲-۴) تا (۵-۴) آورده شده‌اند.

^۱ Confusion Matrix

جدول ۲-۴ ماتریس درهم‌ریختگی سیستم بازشناسی هوشمند خودرو با روش استخراج ویژگی MFCC (۱۳ ضریب)

	BMW	Ferrari	Ford	Jeep	Police	Volvo
BMW	86.79	1.89	1.88	5.66	0	3.78
Ferrari	7.40	88.89	0	0	0	3.70
Ford	4.55	6.82	72.73	6.82	0	9.09
Jeep	0	0	0	100	0	0
Police	0	0	0	14	86	0
Volvo	0	0	1.97	37.25	0	60.78

با توجه به جدول (۲-۴) می‌توان مشاهده نمود که سیستم بازشناسی با روش استخراج ویژگی MFCC ماکزیمم دقت بازشناسی ۱۰۰٪ را در کلاس ۴ و می‌نیمم دقت بازشناسی ۶۰,۷۸٪ را در کلاس ۶ نتیجه می‌دهد. دلیل می‌نیمم بودن دقت بازشناسی در کلاس ۶ را می‌توان شباهت زیاد این کلاس با کلاس ۴ دانست.

جدول ۳-۴ ماتریس درهم‌ریختگی سیستم بازشناسی هوشمند خودرو با روش استخراج ویژگی LPCC (۱۳ ضریب)

	BMW	Ferrari	Ford	Jeep	Police	Volvo
BMW	52.83	0	39.62	0	7.54	0
Ferrari	3.70	51.85	31.49	3.70	1.85	7.40
Ford	0	22.73	59.06	6.81	11.36	0
Jeep	7.54	0	0	84.90	7.54	0
Police	0	0	0	0	100	0
Volvo	3.92	0	0	33.33	17.64	45.09

همان‌طور که در جدول (۳-۴) مشاهده می‌شود سیستم بازشناسی با روش استخراج ویژگی LPC همان‌طور که در جدول (۳-۴) مشاهده می‌شود سیستم بازشناسی با روش استخراج ویژگی LPC ماکزیمم دقت بازشناسی ۱۰۰٪ را در کلاس ۵ و می نیمم دقت بازشناسی ۴۵,۰۹٪ را در کلاس ۶ نتیجه داده است. دلیل می نیمم بودن دقت بازشناسی در کلاس ۶ را می توان شباهت زیاد این کلاس با کلاس ۴ دانست.

جدول ۴-۴ ماتریس درهم‌ریختگی سیستم بازشناسی هوشمند خودرو با روش استخراج ویژگی LPC (۱۳ ضریب)

	BMW	Ferrari	Ford	Jeep	Police	Volvo
BMW	35.85	0	56.6	1.88	3.77	1.88
Ferrari	5.55	81.48	11.11	0	0	1.85
Ford	0	2.27	59.1	13.63	22.72	2.27
Jeep	1.88	15.1	0	11.32	47.16	24.52
Police	0	0	0	0	100	0
Volvo	15.68	7.84	1.97	3.92	17.64	52.94

با توجه به جدول (۴-۴) می توان مشاهده نمود که سیستم بازشناسی با روش استخراج ویژگی LPC ماکزیمم دقت بازشناسی ۱۰۰٪ را در کلاس ۵ و می نیمم دقت بازشناسی ۱۱,۳۲٪ را در کلاس ۴ نتیجه داده است. دلیل می نیمم بودن دقت بازشناسی در کلاس ۴ را می توان شباهت زیاد این کلاس با کلاس‌های ۵ و ۶ با استفاده از این روش استخراج ویژگی دانست.

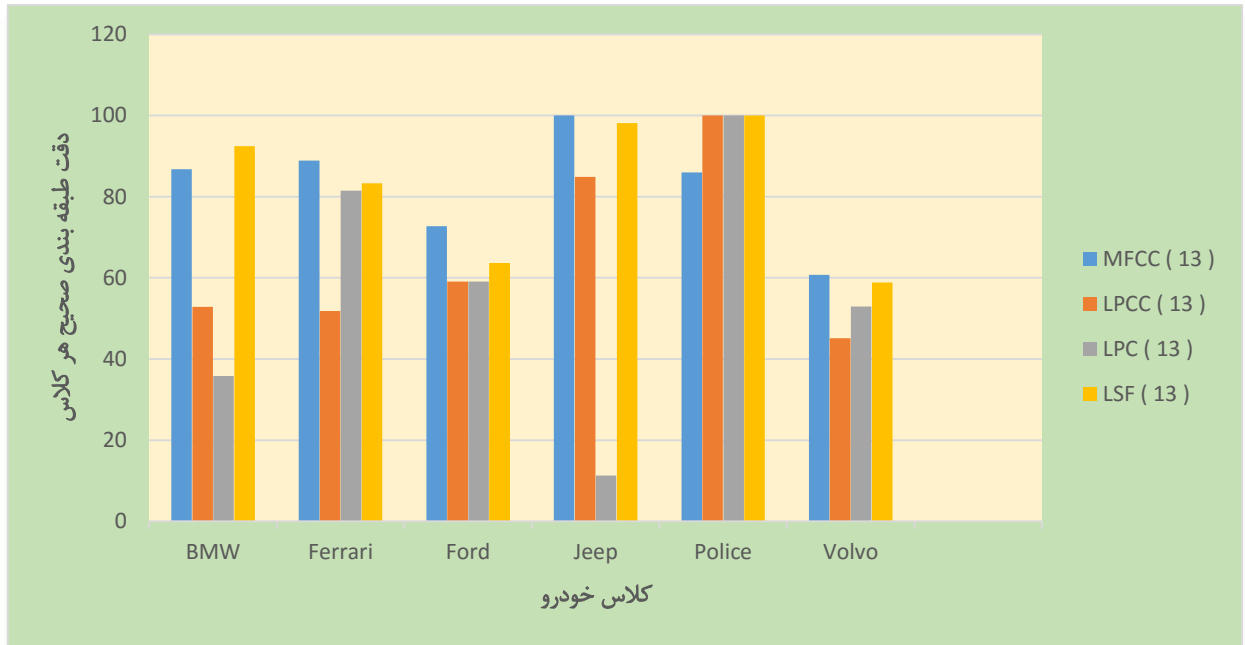
جدول ۴-۵ ماتریس درهم‌ریختگی سیستم بازشناسی هوشمند خودرو با روش استخراج ویژگی LSF (۱۳ ضریب)

	BMW	Ferrari	Ford	Jeep	Police	Volvo
BMW	92.45	1.88	0	1.88	1.88	1.88
Ferrari	3.70	83.33	5.55	3.70	1.85	1.85
Ford	2.27	20.45	63.63	9.09	0	4.54
Jeep	0	0	0	98.11	0	1.88
Police	0	0	0	0	100	0
Volvo	0	0	0	41.17	0	58.82

همان‌طور که در جدول (۴-۵) مشاهده می‌شود، سیستم بازشناسی با روش استخراج ویژگی LSF ماکزیمم دقت بازشناسی ۱۰۰٪ را در کلاس ۵ و می‌نیمم دقت بازشناسی ۵۸٫۸۲٪ را در کلاس ۶ نتیجه داده است. دلیل می‌نیمم بودن دقت بازشناسی در کلاس ۶ را می‌توان شباهت زیاد این کلاس با کلاس ۴ دانست.

با مقایسه این ۴ ماتریس مشاهده می‌شود روش استخراج ویژگی LSF در جایگاه اول، روش MFCC در جایگاه دوم، روش LPCC در جایگاه سوم و درنهایت روش LPC در جایگاه چهارم ماکزیمم دقت بازشناسی میانگین قرار دارند.

شکل (۳-۴) دقت طبقه‌بندی صحیح هر کلاس خودروی پایگاه داده را با روش‌های مختلف استخراج ویژگی با ۱۳ ضریب در هر قاب نشان می‌دهد.



شکل ۳-۴ دقت طبقه‌بندی صحیح هر کلاس خودروی پایگاه داده با روش‌های مختلف استخراج ویژگی (۱۳ ضریب) با توجه به شکل (۳-۴) مشاهده می‌شود دقت طبقه‌بندی سیستم شناسایی هوشمند خودرو با سه تکنیک استخراج ویژگی LPC، LPCC و LSF برای شناسایی خودروی Police برابر ۱۰۰٪ است. پس می‌توان نتیجه گرفت این خودرو دارای ویژگی‌های صوتی منحصر به فردی است که کمترین شباهت را به ویژگی‌های سیگنال‌های صوتی سایر کلاس‌ها دارد.

۲-۳-۴ دقت بازشناسی روش‌های مختلف استخراج ویژگی با طول‌های متفاوت

در این زیر بخش، شبیه‌سازی‌ها بر اساس دقت بازشناسی ۴ روش استخراج ویژگی موجود با طول‌های متفاوت ۱۳، ۲۶ و ۳۹ ضریب ویژگی در هر قاب صورت پذیرفته است.

در جدول‌های (۴-۶) تا (۴-۸) نتایج حاصل از شبیه‌سازی مربوط به ۴ روش مختلف استخراج ویژگی با طول متفاوتی از ویژگی‌ها در هر قاب ارائه گردیده است.

جدول ۶-۴ نتایج حاصل از شبیه‌سازی سیستم بازشناسی خودرو با روش‌های مختلف استخراج ویژگی (۱۳ ضریب)

	Bin	Accuracy
MFCC (12 MFCC + Energy)	13	83.6066
LPC	13	57.7049
LPCC	13	66.5574
LSF	13	84.2623

با توجه به جدول (۶-۴) عملکرد سیستم با روش استخراج ویژگی LSF با دقت بازشناسی ۸۴,۲۶۲۳٪

در رتبه اول و با روش استخراج ویژگی LPC با دقت بازشناسی ۵۷,۷۰۴۹٪ در رتبه آخر قرار دارد.

جدول ۷-۴ نتایج حاصل از شبیه‌سازی سیستم بازشناسی خودرو با روش‌های مختلف استخراج ویژگی (۲۶ ضریب)

	Bin	Accuracy
MFCC (25 MFCC + Energy)	26	82.2951
LPC	26	52.7869
LPCC	26	56.3934
LSF	26	85.2459

با توجه به جدول (۷-۴) مشاهده می‌گردد عملکرد سیستم با روش استخراج ویژگی LSF با دقت بازشناسی ۸۵,۲۴۵۹٪ در بالاترین سطح و با روش استخراج ویژگی LPC با دقت بازشناسی ۵۲,۷۸۶۹٪ در رتبه آخر قرار دارد.

جدول ۸-۴ نتایج حاصل از شبیه‌سازی سیستم بازشناسی خودرو با روش‌های مختلف استخراج ویژگی (۳۹ ضریب)

	Bin	Accuracy
MFCC (38 MFCC + Energy)	39	82.9508
LPC	39	56.3934
LPCC	39	53.7705
LSF	39	85.5738

همان‌طور که در جدول (۸-۴) مشاهده می‌شود عملکرد سیستم شناسایی هوشمند خودرو با روش استخراج ویژگی LSF با دقت بازشناسی ۸۵,۵۷۳۸٪ در رتبه اول و با روش استخراج ویژگی LPCC با دقت بازشناسی ۵۳,۷۷۰۵٪ در رتبه آخر قرار دارد.

با مقایسه این سه جدول مشاهده می‌شود، سیستم بازشناسی هوشمند خودرو با روش LSF با ویژگی‌هایی به طول ۱۳ و ۲۶ و ۳۹ ضریب در هر قاب در بین چهار روش استخراج ویژگی موجود بهترین عملکرد را دارد.

۴-۴ جمع‌بندی

در ابتدای این فصل، روش پیشنهادی جهت شناسایی و طبقه‌بندی هوشمند خودروها با پردازش صدای آن‌ها را بیان کردیم. به این منظور ابتدا پایگاه داده صوتی مورد استفاده در این پایان‌نامه معرفی شد و سپس از چهار روش MFCC، LPC، LSF و LPCC جهت استخراج ویژگی از سیگنال‌های صوتی پایگاه داده بهره بردیم. برای مرحله طبقه‌بندی نیز از یک طبقه‌بند SVM با یک کرنل چندجمله‌ای درجه ۴ استفاده نمودیم.

در پایان فصل نیز نتایج شبیه‌سازی حاصل از پیاده‌سازی روش پیشنهادی و مقایسه نتایج ناشی از به‌کارگیری روش‌های مختلف استخراج ویژگی با تعداد ضرایب ویژگی متفاوت با استفاده از طبقه‌بند SVM را به‌طور کامل مورد بحث و بررسی قرار دادیم.

در فصل بعدی که آخرین فصل پایان‌نامه است نتیجه‌گیری کلی در مورد سیستم شناسایی هوشمند پیشنهاد شده خودرو با استفاده از پردازش صوت و مقایسه عملکرد آن با روش‌های انجام شده قبلی و همچنین مسیر آینده برای ادامه و بهبود این طرح ارائه خواهد شد.

فصل پنجم:

نتیجه‌گیری و پیشنهادها

۱-۵ مقدمه

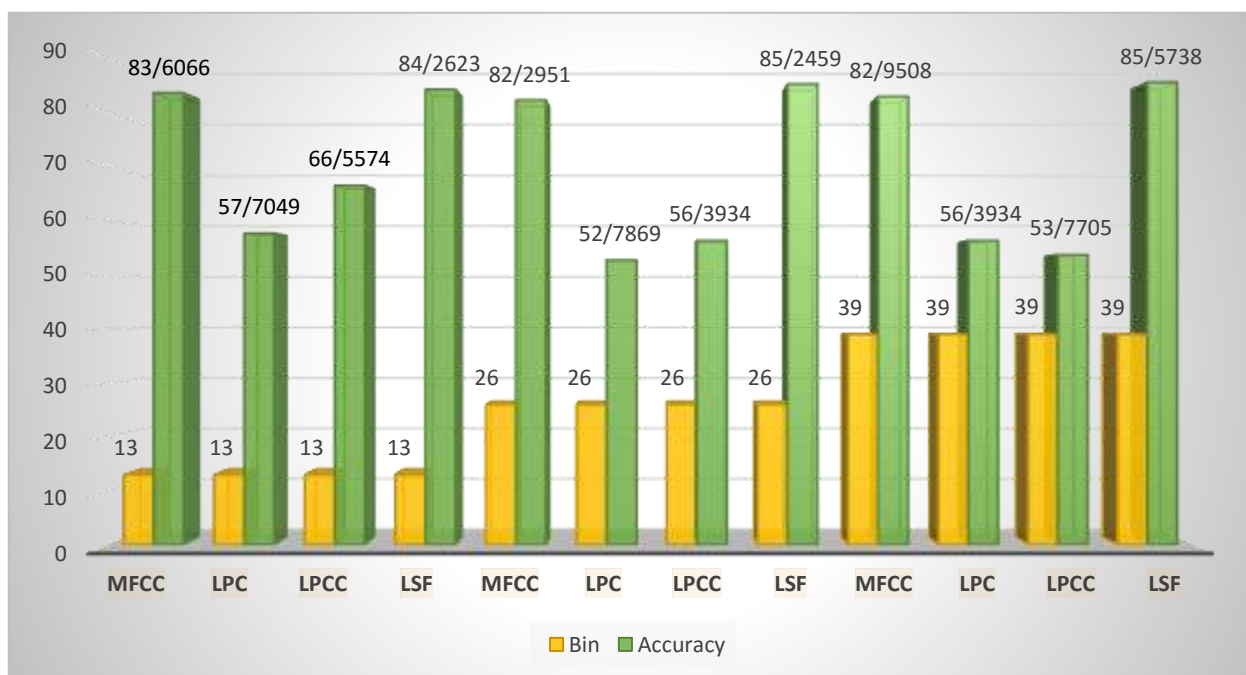
در این فصل ابتدا نتیجه‌گیری صورت می‌گیرد و در نهایت چند پیشنهاد به‌عنوان خط سیر آینده برای ادامه و بهبود سیستم پیشنهاد شده در پایان‌نامه ارائه می‌گردد.

۲-۵ نتیجه‌گیری

در این بخش ابتدا یک نتیجه‌گیری کلی از سیستم شناسایی هوشمند پیشنهاد شده خودرو صورت می‌گیرد و سپس عملکرد این سیستم پیشنهادی با روش‌های انجام شده در تحقیقات قبلی مقایسه می‌شود.

۱-۲-۵ نتیجه‌گیری کلی از سیستم شناسایی هوشمند پیشنهاد شده خودرو

مهم‌ترین نتایج شبیه‌سازی حاصل از پیاده‌سازی سیستم شناسایی هوشمند خودرو در نرم‌افزار MATLAB جهت تحلیل و مقایسه کلی عملکرد این سیستم بر اساس ۴ روش استخراج ویژگی موجود با ۳ طول ویژگی مشخص در شکل (۱-۵) نشان داده شده است.



شکل ۱-۵ مقایسه عملکرد سیستم بازشناسی هوشمند خودرو بر اساس ۴ روش استخراج ویژگی با ۳ طول مشخص

با توجه به شکل (۵-۱) مشاهده می‌شود سیستم شناسایی هوشمند خودرو با روش LSF در تمامی حالت‌ها از سایر روش‌های استخراج ویژگی موجود دقت بازشناسی بالاتری دارد. از این رو می‌توان این روش استخراج ویژگی را بهترین روش برای این سیستم شناسایی دانست. این سیستم با استفاده از روش استخراج ویژگی MFCC نیز دقت بازشناسی بالا و نزدیک به LSF دارد، حال می‌توان از روش MFCC در جایگاه دوم برای سیستم شناسایی هوشمند خودرو استفاده کرد. از طرفی دو روش استخراج ویژگی LPC و LPCC در تمامی حالات نسبت به دو روش LSF و MFCC دقت بازشناسی خیلی کمتری دارند. به همین دلیل این دو روش به منظور استخراج‌کننده ویژگی برای سیستم شناسایی هوشمند خودرو مناسب نیستند.

نتیجه دیگر این است که زیاد بودن تعداد ضرایب ویژگی همیشه باعث افزایش دقت بازشناسی نمی‌شود و حتی گاهی به دلیل حجیم شدن طول ویژگی‌ها باعث کاهش دقت بازشناسی می‌گردد [۳۶]. به همین دلیل باید در انتخاب تعداد ضرایب ویژگی هر قاب برای به دست آوردن دقت بازشناسی ماکزیمم، طول مناسب که در این سیستم ۱۳ بوده را در نظر گرفت.

۵-۲-۲ مقایسه عملکرد روش پیشنهادی با تعدادی از روش‌های مطرح‌شده قبلی

در هنگام مقایسه عملکرد سیستم شناسایی هوشمند پیشنهادشده در این پایان‌نامه با روش‌های مطرح‌شده در تحقیقات قبلی، باید به چند نکته زیر توجه خاص داشته باشیم:

الف) شرایط ضبط صدای خودروها در سایر تحقیقات نسبت به هم متفاوت است.

ب) نوع پیش‌پردازش‌های صورت گرفته بر روی پایگاه داده در هر یک از مقالات متفاوت است.

ج) تعداد خودروها و نوع آن‌ها از لحاظ وزن و کاربرد در سایر تحقیقات باهم تفاوت دارد.

د) نوع روش استخراج ویژگی و طبقه‌بندی که در سایر مقالات استفاده شده، باهم فرق دارد.

بنابراین به این نتیجه می‌رسیم که نمی‌توان یک مقایسه کامل و صحیحی از یک الگوریتم ارائه‌شده در یک پژوهش با یک پژوهش دیگر داشته باشیم و تنها نتیجه مقایسه‌ها، به صورت تقریبی و عینی است.

- دقت شناسایی روش پیشنهادی با تکنیک‌های LSF و MFCC با طبقه بند SVM نسبت به تکنیک FFT با همین طبقه بند در [۱۳] حدود ۱۰٪ بیشتر بوده و عملکرد بهتری داشته است.
- دقت شناسایی روش پیشنهادی با تکنیک MFCC با طبقه بند SVM در مقایسه با همین تکنیک استخراج ویژگی با طبقه بند MLFFNN در [۱۲] حدود ۱۶٪ بیشتر است.
- هم‌چنین دقت طبقه‌بندی صحیح الگوریتم پیشنهادی با تکنیک MFCC با طبقه بند SVM نسبت به همین تکنیک استخراج ویژگی با طبقه بند KNN (K=5) در [۱۰] تقریباً باهم برابر است و می‌توان نتیجه گرفت، هر دو طبقه بند جهت شناسایی و طبقه‌بندی با این تکنیک استخراج ویژگی مناسب هستند.
- دقت شناسایی روش پیشنهادی با تکنیک‌های فرکانسی با طبقه بند SVM در مقایسه با تکنیک FFT با طبقه بندهای SVM, MLE و KNN (K=1) در [۹] حدود ۱۰٪ کمتر بوده و عملکرد ضعیف‌تری داشته‌اند. یکی از علت‌های بالا بودن دقت شناسایی در [۹] محدود بودن تعداد کلاس‌های موجود است.
- دقت شناسایی روش پیشنهادی با تکنیک‌های LSF و MFCC با طبقه بند SVM نسبت به تکنیک HLA با طبقه بند BPNN در [۶] تقریباً برابر بوده و در مقایسه با تکنیک PSE با همین طبقه بند، بسیار بیشتر است.
- دقت طبقه‌بندی در روش پیشنهادی با تکنیک LPC با طبقه بند SVM در مقایسه با تکنیک TESP با همین طبقه بند در [۳] تقریباً برابر بوده و در مقایسه با تکنیک ITESP با همین طبقه بند حدود ۲۵٪ پایین‌تر است.

جدول (۵-۱) به صورت جامع تر، عملکرد روش پیشنهادی را در مقایسه با تعدادی از روش های مطرح شده در تحقیقات قبلی نشان می دهد.

جدول ۵-۱ مقایسه عملکرد روش پیشنهادی با تعدادی از روش های مطرح شده در تحقیقات قبلی

نام پژوهش	روش استخراج ویژگی	نوع طبقه بند	تعداد کلاس ها	نرخ طبقه بندی (درصد)
مقاله [۱۳]	FFT	SVM	۲ نوع خودرو غیرنظامی	۷۴,۶۱
مقاله [۱۲]	MFCC	MLFFNN	۴ نوع خودرو غیرنظامی	۶۷,۴
مقاله [۱۰]	MFCC	KNN (K=5)	۳ نوع خودرو غیرنظامی	۸۱,۳۳
مقاله [۹]	FFT	MLE	۲ نوع خودرو نظامی	۸۴,۹۳
		SVM		۹۵,۵
		KNN (K=1)		۹۲,۳۷
مقاله [۶]	HLA	BPNN	۳ نوع خودرو نظامی	۸۰ - ۹۲
مقاله [۳]	TESP	SVM	۲ نوع خودرو غیرنظامی	۵۱
پایان نامه	LSF (۱۳ ضریب)	SVM	۶ کلاس خودرو سواری	۸۴,۲۶
	MFCC (۱۳ ضریب)			۸۳,۶
	LPCC (۱۳ ضریب)			۶۶,۵۵
	LPC (۱۳ ضریب)			۵۷,۷

در پایان می توان نتیجه گرفت، تکنیک های استخراج ویژگی از حوزه ی فرکانس با طبقه بندهای SVM، KNN و ANN بیشترین نرخ تفکیک پذیری و طبقه بندی صحیح را در بین سایر تکنیک های استخراج ویژگی و طبقه بندها دارند.

۵-۳ خط سیر آینده

به منظور ادامه کار و بهبود سیستم شناسایی هوشمند نوع خودرو چند پیشنهاد ارائه می‌گردد:

❖ افزایش تعداد کلاس خودروهای مورداستفاده در پایگاه داده

یکی از پیشنهادهای ادامه کار افزایش تعداد کلاس خودروهای پایگاه داده اعم از نظامی و غیرنظامی است تا بتوان یک پایگاه داده کارآمد و کاربردی جهت شبیه‌سازی واقعی‌تر داشته باشیم.

❖ اعمال نویز به محیط به منظور بررسی میزان اثر نویز بر دقت شناسایی

در این پایان‌نامه، داده‌های صوتی موجود در پایگاه داده در یک محیط نسبتاً تمیز ضبط شده‌اند از این رو عملکرد سیستم در این حالت تا حدودی قابل قبول است. حال می‌توان با اعمال نویز به محیط ضبط صوت خودروها، میزان اثر نویز را بر دقت شناسایی و عملکرد سیستم ارزیابی کرد.

❖ بررسی روش‌های استخراج ویژگی دیگر

جهت بهبود عملکرد سیستم می‌توان از روش‌های مختلف استخراج ویژگی دیگر استفاده کرد و سپس سیستم را مورد ارزیابی دقیق‌تر قرار دهیم.

❖ استفاده از طبقه بندهای دیگر

یکی دیگر از پیشنهادهای ادامه کار استفاده از طبقه بندهای مختلف در سیستم پیشنهادی است تا بتوان به ازای آن‌ها عملکرد سیستم را بهبود بخشید.

فهرست مراجع

- [1] Borkar, P., & Malik, L. G. (2013). **Review on vehicular speed, density estimation and classification using acoustic signal**. International Journal for Traffic & Transport Engineering, 3(3).
- [2] Shah, M. L., & Mehta, P. D. (2014, May). **Analysis of acoustic signals for vehicle classification of four wheeler models using feature extraction methods**. In Recent Advances and Innovations in Engineering (ICRAIE), 2014 (pp. 1-4). IEEE.
- [3] Wang, Y., Wu, X., Li, X., & Zhou, J. (2014). **Vehicle type recognition in sensor networks using improved time encoded signal processing algorithm**. Mathematical Problems in Engineering, 2014.
- [4] Erb, S. (2007). **Classification of vehicles based on acoustic features**. na.
- [5] Gao, H., Xue, L., & Zhang, C. (2013). **Vehicle Class Recognition Based on Vehicle's Sound Signal**. In ICTIS 2013: Improving Multimodal Transportation Systems-Information, Safety, and Integration (pp. 898-903).
- [6] Wellman, M. C., Srour, N., & Hillis, D. B. (1997, July). **Feature extraction and fusion of acoustic and seismic sensors for target identification**. In Peace and Wartime Applications and Technical Issues for Unattended Ground Sensors (Vol. 3081, pp. 139-146). International Society for Optics and Photonics.
- [7] Succi, G. P., Pedersen, T. K., Gampert, R., & Prado, G. (1999, July). **Acoustic target tracking and target identification: recent results**. In Unattended Ground Sensor Technologies and Applications (Vol. 3713, pp. 10-22). International Society for Optics and Photonics.
- [8] مالکی، رضوانی میثم، تحلیل آکوستیک موتور احتراق داخلی خودرو های ۲۰۶ اس دی، تندر ۹۰ و سوزوکی گراند ویتارا، دومین همایش ملی مدیریت آلودگی هوا و صدا، دی ماه ۱۳۹۲، شماره AQM2013-1107، صفحه ۶، دانشگاه صنعتی شریف، تهران
- [9] Li, D., Wong, K. D., Hu, Y. H., & Sayeed, A. M. (2002). **Detection, classification, and tracking of targets**. IEEE signal processing magazine, 19(2), 17-29.
- [10] Bhave, N., & Rao, P. (2011, March). **Vehicle engine sound analysis applied to traffic congestion estimation**. In Proc. of International Symposium on CMMR and FRSM2011.

- [11] Nooralahiyan, A. Y., Kirby, H. R., & McKeown, D. (1998). **Vehicle classification by acoustic signature**. *Mathematical and Computer Modelling: An International Journal*, 27(9-11), 205-214.
- [12] George, J., Cyril, A., Koshy, B. I., & Mary, L. (2013). **Exploring sound signature for vehicle detection and classification using ANN**. *International Journal on Soft Computing*, 4(2), 29.
- [13] Biernacki, P. (2014, September). **Acoustic information fusion for vehicles identification**. In *Methods and Models in Automation and Robotics (MMAR), 2014 19th International Conference On* (pp. 711-715). IEEE.
- [14] Aljaafreh, A., & Dong, L. (2010, April). **An evaluation of feature extraction methods for vehicle classification based on acoustic signals**. In *Networking, Sensing and Control (ICNSC), 2010 International Conference on* (pp. 570-575). IEEE.
- [15] Eom, K. B. (1999). **Analysis of acoustic signatures from moving vehicles using time-varying autoregressive models**. *Multidimensional Systems and Signal Processing*, 10(4), 357-378.
- [16] Choe, H. C., Karlsen, R. E., Gerhart, G. R., & Meitzler, T. J. (1996, March). **Wavelet-based ground vehicle recognition using acoustic signals**. In *Wavelet Applications III (Vol. 2762, pp. 434-446)*. International Society for Optics and Photonics.
- [17] Maciejewski, H., Mazurkiewicz, J., Skowron, K., & Walkowiak, T. (1997, September). **Neural networks for vehicle recognition**. In *Proceeding of the 6th International Conference on Microelectronics for Neural Networks, Evolutionary and Fuzzy Systems* (p. 5).
- [18] Thammakhoun, N. B., & Lang, S. W. (1999). **Long range acoustic classification**. SANDERS ASSOCIATES INC NASHUA NH.
- [19] Wu, H., Siegel, M., & Khosla, P. (1998, May). **Vehicle sound signature recognition by frequency vector principal component analysis**. In *Instrumentation and Measurement Technology Conference, 1998. IMTC/98. Conference Proceedings. IEEE(Vol. 1, pp. 429-434)*. IEEE.
- [20] Zhao, X., & Wang, D. (2013, May). **Analyzing noise robustness of MFCC and GFCC features in speaker identification**. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on* (pp. 7204-7208). IEEE.
- [21] Majeed, S. A., Husain, H., Samad, S. A., & Idbeaa, T. F. (2015). **Mel frequency cepstral coefficients (MFCC) feature extraction enhancement in the application of**

speech recognition: A comparison study. Journal of Theoretical & Applied Information Technology, 79(1).

[22] Sadeghi, M., & Marvi, H. (2017, December). **Optimal MFCC features extraction by differential evolution algorithm for speaker recognition.** In Intelligent Systems and Signal Processing (ICSPIS), 2017 3rd Iranian Conference on (pp. 169-173). IEEE.

[23] Prabhakar, O. P., & Sahu, N. K. (2013). **A survey on: Voice command recognition technique.** International Journal of Advanced Research in Computer Science and Software Engineering, 3(5).

[24] Desai, N., Dhameliya, K., & Desai, V. (2013). **Feature extraction and classification techniques for speech recognition: A review.** International Journal of Emerging Technology and Advanced Engineering, 3(12), 367-371.

[25] Vanik, R. B., & Dudy, A. (2014). **Linear Predictive Coding Algorithm with its Application to Sound Signal Compression.** International Journal of Engineering and Innovative Technology, 3, 170-173.

[26] Mielikainen, J., Honkanen, R., Huang, B., Toivanen, P. J., & Lee, C. (2010). **Constant coefficients linear prediction for lossless compression of ultraspectral sounder data using a graphics processing unit.** Journal of Applied Remote Sensing, 4(1), 041774.

[27] Eduati, F., Corradin, A., Di Camillo, B., & Toffolo, G. (2010). **A Boolean approach to linear prediction for signaling network modeling.** PloS one, 5(9), e12789.

[28] Oppenheim, A. V. (1999). **Discrete-time signal processing.** Pearson Education India.

[29] Itakura, F. (1975). **Line spectrum representation of linear predictor coefficients of speech signals.** The Journal of the Acoustical Society of America, 57(S1), S35-S35.

[30] Raman, P., & Beex, A. L. (2015, December). **Using LSF features for speaker verification in noise.** In Signal and Information Processing (GlobalSIP), 2015 IEEE Global Conference on (pp. 697-701). IEEE.

[31] Sharma, U., Maheshkar, S., & Mishra, A. N. (2015, February). **Study of robust feature extraction techniques for speech recognition system.** In Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE), 2015 International Conference on (pp. 654-658). IEEE.

- [32] Sumithra, M. G., & Devika, A. K. (2012, January). **A study on feature extraction techniques for text independent speaker identification**. In Computer Communication and Informatics (ICCCI), 2012 International Conference on (pp. 1-5). IEEE.
- [33] Nehe, N. S., & Holambe, R. S. (2008, November). **New feature extraction methods using DWT and LPC for isolated word recognition**. In TENCON 2008-2008 IEEE Region 10 Conference (pp. 1-6). IEEE.
- [34] Pradhan, A. (2012). **Support vector machine-A survey**. International Journal of Emerging Technology and Advanced Engineering, 2(8), 82-85.
- [35] Ma, Y., & Guo, G. (Eds.). (2014). **Support vector machines applications** (pp. 23-26). New York: Springer.
- [36] Astapov, S., Preden, J. S., & Suurjaak, E. (2012, October). **A method of real-time mobile vehicle identification by means of acoustic noise analysis implemented on an embedded device**. In Electronics Conference (BEC), 2012 13th Biennial Baltic(pp. 283-286). IEEE.

Abstract

The design and production of voice recognition systems is the research goal of many scientific centers in the last half-century. One of the human goals in the production of such systems is of course the fact that the entry of information and the implementation of instructions in audio in addition to saving time and cost in different ways also increases the quality of life. The subject matter of this thesis is to provide an intelligent vehicle identification and classification algorithm with its voice processing which is now very important for reasons of use in military and civilian systems. The proposed algorithm should be able to use the car's voice identify and classify its type. The first step in implementing this algorithm is to collect the database. In this thesis, the database has been compiled from the 6 classes of BMW 540i, Ferrari 308, Ford Tempo, Jeep, Police and Volvo Amazon each of which contains 51 samples. The performance of the algorithm is to extract the four important parameters of the MFCC, LPC, LPCC and LSF with the length of the characteristics of 13, 26, and 39 coefficients per frame of the recorded voice of the vehicles and as corresponding vehicle reference features Give to the SVM classifier to be trained with these features. In order to evaluate the algorithm the system extracts the characteristics of an audio as an input and give it into the SVM classifier then classifies the type of vehicle by comparing the input sound characteristics with the reference characteristics. The results of the implementation of this algorithm show that the LSF technique with the detection accuracy of 85.02% of the best feature extraction technique and the MFCC method is also with accuracy of 82.94% it is ranked second. On the other hand the feature extraction techniques of the LPC and LPCC respectively with accuracy of 55.62% and 58.9% for this Identification algorithm can not be good results and their use is not useful.

Keywords: Identification and classification, Feature extraction, SVM, MFCC, LPC, LPCC and LSF



Faculty of Electrical and Robotics Engineering
M.Sc. Thesis in Communication Systems Engineering

**Detection and classification of vehicles by using vehicle sound
signal based SVM**

By:

Farhad Shahbazi Bandani

Supervisor:

Dr. Hosein Marvi

August 2018