

کاربرد تبدیل فوریه گسسته در زمان در شناسایی سیگنال‌های صوتی کشتی‌ها با استفاده از روش‌های کاهش بُعد و یادگیری توده‌ای

امیرحسین دوايي مرکزی^{1*}، میلاد نظر آهاری²

1- استاد، مهندسی مکانیک، دانشگاه علم و صنعت ایران، تهران

2- کارشناس ارشد، مهندسی مکانیک، دانشگاه علم و صنعت ایران، تهران

* تهران، صندوق پستی 1684613114، markazi@iust.ac.ir

اطلاعات مقاله

مقاله پژوهشی کامل
دریافت: 04 اردیبهشت 1394
پذیرش: 12 خرداد 1394
ارائه در سایت: 30 خرداد 1394
کلید واژگان:
سیگنال صوتی (اکوستیکی)
تبدیل فوریه گسسته در زمان
استخراج ویژگی
رده‌بند
یادگیری توده‌ای

چکیده

شناسایی و تفکیک کشتی‌ها توسط سیگنال‌های دریافتی توسط میکروفون‌های درون آبی (هیدروفون‌ها)، به‌ویژه در بندرهای تجاری پر رفت‌وآمد، می‌تواند کاربردهای زیادی در مدیریت ترافیک در بندرها داشته باشد. با این حال شناسایی و کلاسه‌بندی خودکار سیگنال‌های صوتی دریافتی از سیستم سونار غیرفعال به علت تغییرات زیاد مشخصه-های زمانی و فرکانسی سیگنال (حتی در شرایطی که سیگنال‌ها از یک منبع واحد دریافت شده باشند) مسأله‌ای چالش‌برانگیز است. در این پژوهش الگوریتم شناسایی سیگنال‌های صوتی مبتنی بر تبدیل فوریه، روش‌های کاهش بُعد متنوع (آنالیز مؤلفه‌های اصلی و هم‌خانواده‌های آن (6 روش) و آنالیز تفکیک‌کنندگی و هم‌خانواده‌های آن (3 روش)) و با بهره بردن از روش‌های یادگیری توده‌ای (4 مجمع مختلف) با سه رده‌بند پرسپترون چندلایه، شبکه عصبی احتمالاتی و ماشین بردار پشتیبان ارائه شده است. با انجام آزمون‌های کارایی متعدد و متنوع، کارایی روش‌های مختلف مورد ارزیابی قرار گرفته و مناسب‌ترین روش‌ها برای الگوریتم پیشنهادی گزینش شده‌اند. از این الگوریتم برای استخراج ویژگی، کاهش بُعد و کلاسه‌بندی صدای هشت کشتی استفاده شده است. الگوریتم پیشنهادی بر روی سیگنال‌های واقعی و نیز نسخه آغشته به نویز (با نسبت-های سیگنال به نویز 5-، 10- و 15-) آن‌ها تست شده است. نتایج نشان می‌دهند که نرخ کلاسه‌بندی صحیح الگوریتم پیشنهادی در نسبت‌های سیگنال به نویز 5-، 10- و 15- به‌ترتیب برابر 99/83، 97/06 و 83/56 درصد است.

Application of DWT for Ship's Acoustic Signal Identification Using Feature Extraction Methods and Ensemble Learning

Amir Hossein Davaie Markazi*, Milad Nazarahari

Department of Mechanical Engineering, Iran University of Science and Technology, Tehran, Iran
* P.O.B. 1684613114 Tehran, Iran, markazi@iust.ac.ir

ARTICLE INFORMATION

Original Research Paper
Received 24 April 2015
Accepted 02 June 2015
Available Online 20 June 2015

Keywords:

Audio (Acoustic) Signal
DWT
Feature Extraction
Classifier
Ensemble Learning

ABSTRACT

Identification and classification of signals which are heard by underwater microphones (hydrophones) can be used extensively in harbor traffic management, especially in economical harbors. However, automatic identification and classification of acoustic signals which are received by passive sonar system is a challenging problem, because of variation in temporal and frequency characteristics of signals (even they are received from a same source). In this paper, a novel method for classification of acoustic signals is presented, based on DWT as preprocessing, a diverse range of feature extraction methods (principal component analysis and its variations (6 methods) and discriminant analysis and its variations (3 methods)), and 4 ensemble learning methods with 3 classifiers (multilayer perceptron (MLP), probabilistic neural network (PNN) and support vector machine (SVM)). Performing a diverse range of performance tests, the performances of different methods are assessed and the best ones are chosen for the proposed method. The proposed method is used to extract features and classify acoustic signals of 8 ships. Using the proposed method, some real signals and their noisy version are classified. The accuracy of the proposed method in classification of test signals with Gaussian white noise with -5, -10 and -15 signal-to-noise ratio is obtained as 99.83%, 97.06% and 83.56%, respectively.

1- مقدمه

سیگنال‌های الکتریکی تبدیل می‌کنند. نویز ساطع شده توسط پروانه و موتور کشتی توسط هیدروفون‌ها دریافت شده و سپس به‌منظور بررسی اطلاعات دیداری و شنیداری آن توسط اپراتور سیستم سونار، مورد پردازش قرار می‌گیرد. اپراتور از این اطلاعات برای تصمیم‌گیری پیرامون شناسایی کشتی‌ها

فرآیند شناسایی کشتی‌ها به‌طور معمول توسط پردازش سیگنال‌های صوتی دریافتی با حسگرهای سیستم سونار غیرفعال انجام می‌شود. این حسگرها، هیدروفون‌ها و یا آرایه‌های هیدروفونی هستند که فشار صوت را به

بهره می‌برد. روش‌های سنتی برای کلاسه‌بندی سیگنال‌های سونار غیرفعال شامل بررسی مشخصه‌های سیگنال دریافتی، به هر دو صورت شنیداری و طیف فرکانسی، توسط اپراتور می‌شود. موضوعی که سبب می‌شود این کار به‌ویژه در عمل با دشواری فراوانی مواجه شود، پیچیدگی اصوات زیرآبی نایستادن آغشته به نویز است؛ همچنین، به‌طور معمول این اصوات شامل چندین مؤلفه از منابع مختلف مانند کاویتاسیون پروانه کشتی، ارتعاشات بدنه کشتی، ماشین‌آلات دوار داخلی و غیره هستند. نویز ساطع شده از کشتی در هنگام حرکت به‌طور عمده از ارتعاش ماشین‌آلات و کاویتاسیون پروانه ناشی می‌شود [1]. بیشترین این ماشین‌آلات صوتی در فرکانس دوران و هارمونیک‌های خود ایجاد می‌کنند. این نویز باند باریک، همراه با نویز باند پهن ناشی از کاویتاسیون پروانه، تا فواصل طولانی منتشر شده و به‌وسیله حسگرهای سیستم سونار غیرفعال دریافت می‌شوند.

بهره بردن از روش‌های هوشمند در مسأله شناسایی سیگنال‌های سونار، می‌تواند سبب افزایش سرعت شناسایی و نیز بهبود فرایند تصمیم‌گیری شود. در هنگام بهره بردن از یک سیستم شناسایی خودکار، استخراج ویژگی از سیگنال ساطع‌شده از کشتی گام کلیدی بشمار می‌رود. به علت پیچیده بودن محیط دریا، یافتن یک روش استخراج ویژگی مناسب که قادر به ارائه ماهیت هدف باشد، ضروری است. در ادامه، باید از یک رده‌بنده کارا به‌منظور کلاسه‌بندی سیگنال‌ها استفاده کرد.

به‌طور معمول از دو روش برای پیش‌پردازش سیگنال‌های صوتی زیرآبی استفاده می‌شود: تبدیل فوریه¹ و تبدیل موجک². برای مثال در [2]، پس از محاسبه تبدیل فوریه سیگنال‌ها، ضرایب تبدیل فوریه در بازه فرکانسی مورد نظر استخراج شده و در ادامه با اعمال الگوریتم پنجره برشی دوره³ [3]، اثرات نویز پس‌زمینه کاهش می‌یابد. سپس از روش‌های آنالیز مؤلفه‌های اصلی⁴ و آنالیز مؤلفه‌های اصلی غیرخطی برای استخراج ویژگی و از رده‌بند پرسپترون چندلایه⁵ برای کلاسه‌بندی اصوات استفاده شده است. چگونگی اثرگذاری بخش‌های مختلف روش پیشنهادی در [2] بر کارایی الگوریتم شناسایی، در [4] مورد بررسی قرار گرفته است. در پژوهش [5]، الگوریتمی به‌منظور تشخیص سیگنال‌های ناشناخته (سیگنال‌هایی که در هنگام آموزش رده‌بند استفاده نشده‌اند) ارائه شده است. استفاده از تبدیل فوریه و طیف توان حاصل از آن برای شناسایی نوع خودروها از روی صدایشان در [6] مورد بررسی قرار گرفته است. در این مقاله پس از محاسبه طیف توان از روی ضرایب تبدیل فوریه، از آنالیز مؤلفه‌های اصلی برای استخراج ویژگی استفاده شده است. در [7]، نویسندگان نخست با محاسبه تبدیل فوریه سیگنال‌ها، چگالی طیف توان متوسط⁶ هر سیگنال را محاسبه کرده و سپس با محاسبه تبدیل موجک طیف توان متوسط و نگه‌داشتن ضرایب در مقیاس‌های پایین، مؤلفه‌های با محتوای فرکانس بالا استخراج شده و به‌عنوان ورودی پرسپترون چندلایه در نظر گرفته می‌شوند. در [8]، همین نویسندگان از الگوریتم پنجره برشی دوره به‌جای تبدیل موجک برای استخراج مشخصه‌هایی با محتوای فرکانسی بالا در طیف توان متوسط استفاده کرده‌اند. استفاده از طیف نگاشت حاصل از تبدیل فوریه وابسته به زمان⁷ به‌عنوان مشخصه یک سیگنال در [9] ارائه شده است. در این مقاله، پس از محاسبه طیف نگاشت، از یک شبکه

عصبی خودسازمانده⁸ برای کاهش بُعد استفاده شده است؛ به‌نحوی که آرایش وزن‌های شبکه عصبی خودسازمانده بردار مشخصه مربوط به هر صوت را تشکیل می‌دهد. در نظر گرفتن ضرایب مدل پیش‌بین⁹ مربوط به طیف توان سیگنال‌ها به‌عنوان مشخصه در [11،10] انجام شده است. در [12] نخست داده‌های به‌کار رفته در استخراج مشخصه‌های طیفی معرفی شده و سپس الگوریتم پیشنهادی جهت کلاسه‌بندی شناورهای دریایی به کمک مدل پنهان مارکوف ارائه شده است.

در این مقاله، روشی برای کلاسه‌بندی سیگنال‌های صوتی زیرآبی هشت کشتی ارائه شده است؛ در این روش از تبدیل فوریه به‌عنوان روش پیش‌پردازش سیگنال‌ها، و از روش‌های استخراج ویژگی مختلف (آنالیز مؤلفه‌های اصلی و هم‌خانواده‌های آن (6 روش) و آنالیز تفکیک‌کنندگی خطی و هم‌خانواده‌های آن (3 روش)) برای تولید بردار ویژگی استفاده شده است. در انتها از روش‌های یادگیری توده‌ای¹⁰ (4 روش) با سه رده‌بند پرسپترون چندلایه، شبکه عصبی احتمالاتی¹¹ و ماشین بردار پشتیبان¹² برای کلاسه‌بندی سیگنال‌های کشتی‌ها استفاده شده است. وجه تمایز اصلی این پژوهش با تحقیقات مشابه را می‌توان به استفاده از روش‌های کاهش بُعد متنوع با ساختارهای پیچیده‌تر که تاکنون در شناسایی سیگنال‌های صوتی کشتی‌ها مورد بررسی قرار نگرفته بودند و نیز استفاده از روش‌های یادگیری توده‌ای به منظور ترکیب رده‌بندها با یکدیگر بیان کرد. نتایج به وضوح بهبود کارایی استفاده از این روش‌ها نسبت به روش‌های معمول را نشان می‌دهد.

ساختار این مقاله به این صورت است: در بخش 2 پایگاه داده مورد استفاده معرفی شده است. در بخش 3 روش‌های مورداستفاده در این پژوهش به‌طور مختصر تشریح شده‌اند. در بخش 4 الگوریتم شناسایی پیشنهادی معرفی شده و بخش‌های 5 و 6 به ترتیب به معرفی آزمون‌های انجام شده و بیان نتایج این آزمون‌ها اختصاص دارد. نتیجه‌گیری کار در بخش 7 ارائه شده است.

2- پایگاه داده

در این بخش توضیحاتی پیرامون پایگاه داده مورد استفاده ارائه می‌شود این پایگاه داده k نامیده می‌شود. طول سیگنال‌های موجود در پایگاه داده k تا 5 دقیقه است. این پایگاه داده شامل صدای ضبط شده هشت کشتی A, B, C, D, E, F, G و H همراه با نویز پس‌زمینه است. قابل ذکر است که برای هر یک از کشتی‌ها، چندین صدا که مبین سیگنال صوتی دریافتی از کشتی در سرعت و شرایط جوی متفاوت است، در پایگاه داده k موجود است. این اصوات به فرمت MP3، به‌صورت تک کانال و با نرخ نمونه‌برداری 22050 هرتز ذخیره‌شده بودند. با بررسی محتوای فرکانسی این اصوات مشخص شد که بیشینه محتوای فرکانسی این اصوات در بازه فرکانسی 0-1000 هرتز قرار دارد. از همین رو تمامی این اصوات با نرخ نمونه‌برداری 8000 هرتز و با فرمت AU ذخیره شدند.

به‌منظور ایجاد درک صحیحی از این سیگنال‌ها، طیف توان حاصل سیگنال‌های کشتی‌های A و B در شکل 1 نشان داده‌شده‌اند. همان‌طور که در شکل 1 مشاهده می‌شود، محتوای فرکانسی سیگنال‌ها شباهت زیادی به یکدیگر داشته و همین موضوع تفکیک میان این اصوات را با چالش مواجه

1- Fourier Transform
2- Wavelet Transform
3- Two-Pass Split Window (TPSW)
4- Principal Component Analysis (PCA)
5- Multilayer Perceptron (MLP)
6- Average Power Spectral Density (APSD)
7- Short-Time Fourier Transform

8- Self-Organizing Map
9- AutoRegressive Model
10- Ensemble Learning
11- Probabilistic Neural Network
12- Support Vector Machine

3-2- استخراج ویژگی

عملکرد یک رده‌بند به ارتباط میان تعداد نمونه‌ها، تعداد مشخصه‌ها و میزان پیچیدگی رده‌بند بستگی دارد. تا زمانی که چگالی احتمالاتی کلاس‌ها کاملاً مشخص باشد (یا به عبارت دیگر تعداد نمونه‌های مورد استفاده برای آموزش به اندازه کافی زیاد باشد که چگالی داده‌ها را به طور کامل مشخص کند)، با افزایش تعداد مشخصه‌ها احتمال کلاسه‌بندی اشتباه داده‌ها در یک قاعده تصمیم‌گیری افزایش نخواهد یافت. با این حال در عمل مشاهده می‌شود که در صورتی که تعداد نمونه‌های آموزش در مقایسه با تعداد مشخصه‌ها به نسبت اندک باشد، مشخصه‌های اضافی می‌تواند سبب کاهش کارایی رده‌بند شود. این رفتار پارادوکس گونه بانام پدیده اوج‌دهندگی⁴ شناخته می‌شود [15,14]. علت این موضوع را می‌توان این‌گونه بیان کرد: بیشتر رده‌بند‌های پارامتریک، پارامترهای ناشناخته‌ای را تخمین و از آن‌ها به‌عنوان پارامترهای صحیح چگالی‌های احتمالاتی کلاس‌ها بهره می‌برند. برای یک تعداد مشخص نمونه، با افزایش تعداد مشخصه‌ها (که متناظر با افزایش تعداد پارامترهای ناشناخته است)، کیفیت تخمین پارامترها کاهش می‌یابد. در نتیجه، کارایی رده‌بند‌هایی که از این پارامترها بهره می‌برند نیز دچار افت خواهد شد. تمامی رده‌بند‌های رایج، شامل شبکه‌های عصبی چندلایه، از این موضوع رنج می‌برند. روش‌های استخراج ویژگی، یک زیر فضای مناسب (به روش خطی یا غیرخطی) با بُعد m در فضای اصلی مشخصه‌ها با بُعد d ($m \leq d$) تولید می‌کنند. برای این کار روش‌های خطی و غیرخطی متفاوتی وجود دارد که در ادامه تشریح می‌شوند.

3-2-1- آنالیز مؤلفه‌های اصلی

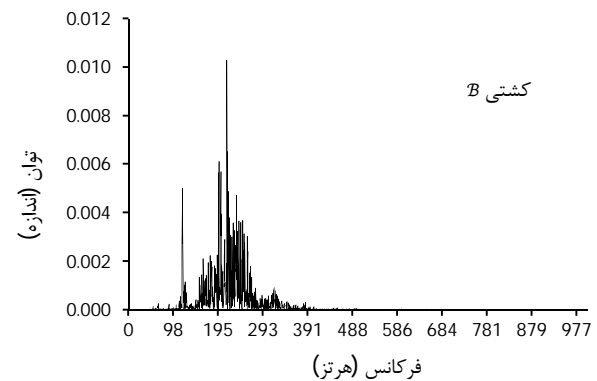
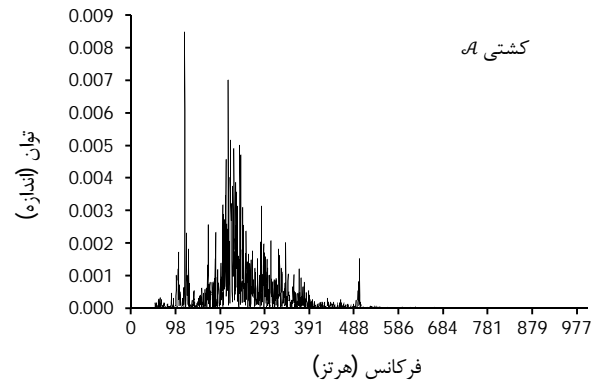
آنالیز مؤلفه‌های اصلی فضای داده‌های اصلی را بر اساس بردارهای ویژه ماتریس کوواریانس تشریح می‌کند. مقادیر ویژه متناظر با بردارهای ویژه، بیانگر انرژی مشخصه‌ها در راستای این بردارها است. فرض کنید تصویر کردن داده‌ها به m بردار ویژه با بزرگ‌ترین مقادیر ویژه متناظر موردنظر باشد، در این صورت آنالیز مؤلفه‌های اصلی یک روش مؤثر برای کاهش بُعد داده‌های اصلی با کمینه از دست رفتن اطلاعات به شمار می‌آید [17,16].

در شرایطی که همبستگی میان متغیرهای مسئله خطی باشد، آنالیز مؤلفه‌های اصلی خطی انتخاب نخست خواهد بود. با این حال در شرایطی که مسئله دارای همبستگی غیرخطی باشد، بهره بردن از نسخه‌های غیرخطی می‌تواند سبب بهبود کارایی شود.

یکی از روش‌های اجرای آنالیز مؤلفه‌های اصلی غیرخطی، آنالیز مؤلفه‌های اصلی با استفاده از تابع کرنل⁵ است. این روش مبتنی بر تصویر داده‌ها در یک فضای مشخصه جدید \mathcal{F} توسط یک تابع غیرخطی Φ و سپس اعمال آنالیز مؤلفه‌های اصلی خطی در این فضای جدید است [18]. همچنین می‌توان با وارد کردن مبانی روش فازی به چهارچوب کاری آنالیز مؤلفه‌های اصلی، روش نوینی بانام آنالیز مؤلفه‌های اصلی فازی⁶ تولید کرد. در این روش کاهش بُعد با تخصیص یک درجه عضویت فازی به هر یک از داده‌ها و سپس محاسبه ماتریس کوواریانس داده‌های فازی شده، انجام می‌شود [19]. از سایر نسخه‌های آنالیز مؤلفه‌های اصلی، می‌توان به آنالیز مؤلفه‌های اصلی فازی با تابع کرنل [20] و احتمالاتی [21] اشاره کرد.

3-2-2- آنالیز تفکیک‌کنندگی

برخلاف آنالیز مؤلفه‌های اصلی، آنالیز تفکیک‌کنندگی از برچسب کلاس



شکل 1 طیف توان کشتی‌های A و B

می‌کند.

3- روش‌ها

یک سیستم شناسایی الگو به‌طور معمول شامل سه بخش است؛ پیش‌پردازش¹، استخراج ویژگی² و رده‌بند³. در ادامه هر یک از این موارد به‌طور خلاصه تشریح می‌شوند.

3-1- تبدیل فوریه

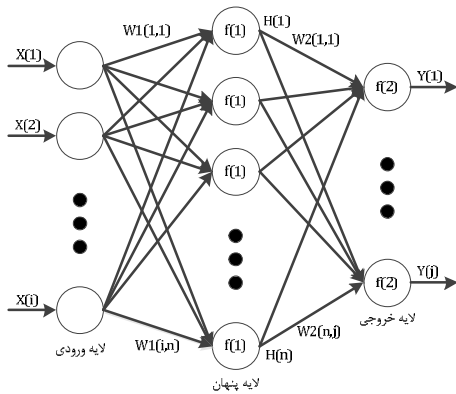
فرض می‌شود که نسخه گسسته در زمان سیگنال صوتی با طول N و نرخ نمونه‌برداری f_s با $x[j]$ نشان داده شود. محتوای فرکانسی سیگنال $x[j]$ در یک بازه زمانی مشخص را می‌توان با استفاده از تبدیل فوریه گسسته در زمان به‌صورت تابعی از فرکانس و با استفاده از ضرایب تبدیل فوریه $X[k]$ بیان کرد. تبدیل پارامترها بین حوزه زمان و فرکانس با استفاده از قضیه پارسوال آغاز می‌شود؛ قضیه پارسوال بیان می‌کند که حاصل جمع مجذور یک تابع با حاصل جمع مجذور تبدیل یافته آن برابر است:

$$\sum_{j=1}^{N-1} |x[j]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]X^*[k]| = \sum_{k=0}^{N-1} P[k] \quad (1)$$

که در آن $P[k]$ طیف توان بدون فاز (یا همان حاصل‌ضرب $X[k]$ و مزدوج آن)، و k اندیس فرکانس است. به‌طور معمول محتوای فرکانسی حاصل از تبدیل فوریه نسبت به فرکانس صفر متقارن است، در نتیجه در هنگام استفاده از طیف توان می‌توان کل آن (شامل فرکانس‌های مثبت و منفی) و یا تنها یک بخش از آن را مدنظر قرارداد [13].

4- Peaking Phenomenon
5- Kernel PCA
6- Fuzzy PCA

1- Preprocessing
2- Feature Extraction
3- Classifier



شکل 2 ساختار رده‌بند پرسپترون چندلایه

شبکه عصبی احتمالاتی شامل یک لایه ورودی، یک لایه الگو، یک لایه جمع‌بندی و یک لایه خروجی رقابتی است. این ساختار در شکل 3 نمایش داده شده است. نرون‌های لایه ورودی، تنها مشخصه‌های ورودی p را به نرون‌های لایه الگو منتقل می‌کنند، که در آن $x = [x_1, x_2, \dots, x_p]^T$ بیانگر تعداد مشخصه‌های استخراج شده است. در لایه الگو، خروجی نرون x_{ki} با استفاده از تابع گوسی چندبعدي (2) محاسبه می‌شود:

$$H_{ki}(x) = \frac{1}{(2\pi)^{d/2} \sigma^d} \exp\left[-\frac{(x - x_{ki})^T (x - x_{ki})}{2\sigma^2}\right] \quad (2)$$

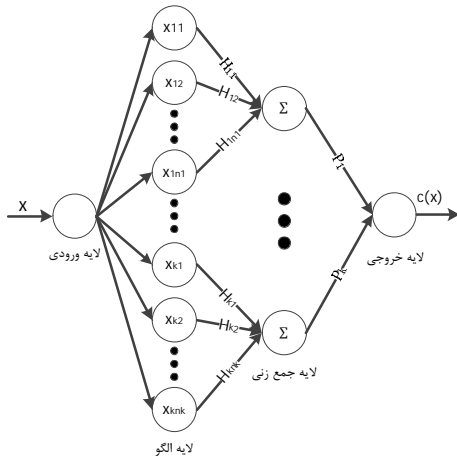
که در آن بردار نرون‌ها، σ پارامتر هموار کننده و d بُعد بردار الگوی است. نرون‌های لایه جمع‌بندی بیشینه شباهت³ بردار الگوی x را که با میانگین‌گیری از خروجی تمامی نرون‌های لایه الگو که به یک کلاس تعلق دارند به دست می‌آید، به کلاس k تخصیص می‌دهند:

$$P_k(x) = \frac{1}{n_k} \sum_{i=1}^{n_k} H_{ki}(x) \quad (3)$$

که در آن n_k تعداد کل نمونه‌های موجود در کلاس k است. نرون خروجی با استفاده از رابطه (4) و بر اساس قاعده تصمیم‌گیری بیز و با در نظر گرفتن فرض "یکسان بودن تابع هزینه متناظر با کلاسه‌بندی نادرست و نیز تابع احتمال تمامی کلاس‌ها" تعلق هر نمونه به کلاس صحیح را تعیین می‌کند:

$$c(x) = \operatorname{argmax}\{P_k(x)\}, k = 1, 2, \dots, m \quad (4)$$

که در آن m بیانگر تعداد کلاس‌ها در مجموعه آموزش و $c(x)$ کلاس تخمین‌زده شده برای الگوی x است [25].



شکل 3 ساختار رده‌بند شبکه عصبی احتمالاتی

داده‌ها برای استخراج مشخصه‌هایی با بیشینه قابلیت تفکیک‌کنندگی استفاده می‌کند. در آنالیز تفکیک‌کنندگی، تفکیک بین کلاس‌ها با جایگزین کردن ماتریس کوواریانس کلی در آنالیز مؤلفه‌های اصلی با یک معیار تفکیک‌پذیری دیگر همچون معیار فیشر که منجر به یافتن بردارهای ویژه $S_w^{-1} S_b$ (حاصل ضرب معکوس ماتریس کوواریانس داخل کلاسی، S_w ، در ماتریس کوواریانس بین کلاسی، S_b) می‌شود، انجام می‌گردد. این روش با نام آنالیز تفکیک‌کنندگی خطی شناخته می‌شود [16]. نسخه توسعه‌یافته آنالیز تفکیک‌کنندگی خطی، آنالیز تفکیک‌کنندگی تعمیم‌یافته [22] است. مزیت این روش نسبت به روش آنالیز تفکیک‌کنندگی خطی، بهره بردن از توابع کرنل غیرخطی به منظور یافتن تصویرهایی با بیشینه میزان تفکیک‌پذیری است.

یکی از معایب اصلی روش آنالیز تفکیک‌کنندگی خطی محدودیت آن در کاهش بُعد داده‌ها است؛ به این معنا که اگر فضای اصلی مشخصه‌ها d بُعدی باشد و نیز تعداد کلاس‌های مسأله c باشد، بُعد زیر فضای تولید شده حداکثر برابر $c-1$ خواهد بود. در شرایطی که $c \gg d$ ، این کاهش بُعد قابل توجه، سبب از دست رفتن مقدار زیادی از اطلاعات می‌شود. روش آنالیز تفکیک‌کنندگی غیر پارامتریک [23] راهکاری برای مقابله با این مشکل ارائه کرده است، به این معنا که در این روش برخلاف روش آنالیز تفکیک‌کنندگی خطی، تعداد مشخصه‌ها پس از کاهش بعد محدود به $c-1$ نبوده و بُعد زیر فضا پس از کاهش بُعد، می‌تواند هر مقدار دلخواه m در نظر گرفته شود.

3-3- رده‌بندها و روش‌های یادگیری توده‌ای

در این پژوهش از پرسپترون چندلایه، شبکه عصبی احتمالاتی و ماشین بردار پشتیبان به‌عنوان رده‌بندهای روش‌های یادگیری توده‌ای استفاده شده است. ساختار پایه‌ای پرسپترون چندلایه از سه لایه تشکیل شده است؛ لایه ورودی، لایه پنهان و لایه خروجی. عملیات آموزش به‌منظور تعیین وزن‌های بین لایه‌های ورودی و لایه پنهان و نیز لایه پنهان و لایه خروجی انجام می‌شود. برای این منظور الگوریتم‌های متنوعی وجود دارد که شناخته‌شده‌ترین آن‌ها پس‌انتشار خطا¹ نامیده می‌شود [24]. این الگوریتم دارای دو فاز مجزا از هم است؛ نخست فاز پیشرو؛ در این فاز سیگنال ورودی به شبکه اعمال شده و اثر آن لایه به لایه محاسبه شده تا پاسخ شبکه حاصل شود. در طول این فاز وزن‌های سیناپسی ثابت هستند. فاز پس‌رو؛ در این فاز تفاضل بین پاسخ شبکه و پاسخ دلخواه (به‌عنوان خطا) محاسبه شده و این سیگنال خطا از سمت خروجی به سمت ورودی و به‌صورت لایه به لایه پس‌انتشار یافته و وزن‌ها به نحوی تصحیح می‌شوند که پاسخ شبکه را هر چه بیشتر به پاسخ دلخواه نزدیک کنند. شکل 2 ساختار یک شبکه پرسپترون چندلایه با یک لایه پنهان، n نرون، i ورودی، z خروجی را نشان می‌دهد. توابع $f(1)$ و $f(2)$ می‌توانند از انواع مختلفی همچون خطی و سیگموئید (تانژانت هیپربولیک) باشند.

شبکه عصبی احتمالاتی نخستین بار توسط اسپک² ارائه شده است. نشان داده می‌شود که در صورت وجود داده‌های کافی برای آموزش، شبکه عصبی احتمالاتی به رده‌بند بیز همگرا می‌شود [25, 26]. بر مبنای رویکرد بیز، خروجی شبکه عصبی احتمالاتی را می‌توان به‌عنوان تخمینی از احتمال میزان عضویت در یک کلاس در نظر گرفت. آموزش این رده‌بند بر پایه تخمین تابع چگالی احتمالاتی کلاس‌ها انجام می‌شود.

3- Maximum Likelihood

1- Error Backpropagation
2- Specht

منجر به اتخاذ تصمیم پیرامون کلاس ω_k می‌شود به نحوی که [28]:

$$\sum_{i=1}^L d_{i,k} = \max_{j=1}^c \sum_{i=1}^L d_{i,j} \quad (11)$$

3-3-2- ترکیب‌کننده پیوسته

ترکیب‌کننده‌های ساده با عدم قابلیت آموزش، میزان تعلق به کلاس ω_k را تنها با استفاده از خروجی رده‌بندهای مختلف حاضر در مجمع برای نمونه x و به صورت:

$$\mu_k(x) = \mathcal{F}[d_{1,k}(x), \dots, d_{L,k}(x)] \quad (12)$$

تعیین می‌کنند، که در آن \mathcal{F} تابع ترکیب‌کننده است. برچسب کلاس x همان اندیس $\mu_k(x)$ است. تابع ترکیب‌کننده \mathcal{F} را می‌توان به طرق مختلفی در نظر گرفت. مشهورترین انتخاب‌ها برای تابع \mathcal{F} عبارت‌اند از توابع میانگین، میانه، بیشینه، کمینه و حاصل ضرب.

3-3-3- بیز ساده

با اعمال رده‌بند D_i به مجموعه داده‌های آموزش، یک ماتریس اختلاط $c \times c$ به دست می‌آید. عضو $m(k, s)$ این ماتریس، $cm_{k,s}^i$ همان تعداد اعضای از مجموعه داده‌های آموزش است که برچسب کلاس حقیقی آن‌ها ω_k بوده درحالی‌که توسط D_i به آن‌ها کلاس ω_s تخصیص داده شده‌اند. منظور از N_s همان تعداد کل اعضای \mathbb{Z} از کلاس ω_s است. با در نظر گرفتن $cm_{k,s_i}^i/N_k$ به عنوان تخمینی برای $P(s_i|\omega_k)$ و نیز N_k/N به عنوان تخمینی از احتمال پیشین برای کلاس ω_s ، می‌توان میزان تعلق مشاهده x به کلاس ω_k را به صورت رابطه (13) بیان کرد:

$$\mu_k(x) \propto \frac{1}{N_k^{L-1}} \prod_{i=1}^L cm_{k,s_i}^i \quad (13)$$

3-3-4- میانگین وزن‌دار

یکی از انواع ترکیب‌کننده‌ها در خانواده میانگین وزن‌دار، بانام $c \times L$ وزن شناخته می‌شود. در این نوع ترکیب‌کننده، میزان تعلق برای کلاس ω_k به صورت رابطه (14) محاسبه می‌شود:

$$\mu_k(x) = \sum_{i=1}^L w_{i,k} d_{i,k}(x) \quad (14)$$

در رابطه (14) وزن‌ها برای هر کلاس به‌طور خاص تعیین می‌شوند. به‌طور معمول از رگرسیون خطی برای محاسبه وزن‌ها در این مدل استفاده می‌شود. با اندکی محاسبات ریاضی نشان داده می‌شود که پاسخ مسأله رگرسیون مقید از این دست به صورت رابطه (15) بیان می‌شود:

$$w = \Sigma^{-1} I (I^T \Sigma^{-1} I)^{-1} \quad (15)$$

که در آن $w = [w_1, \dots, w_L]$ بردار حاوی وزن‌ها، Σ ماتریس کوواریانس خطاهای تقریب رده‌بندها و I بردار L -عضوی با اعضای واحد است.

4- روش پیشنهادی

الگوریتم پیشنهادی از پنج مرحله تشکیل شده است، آماده‌سازی اصوات، پیش‌پردازش، کاهش بُعد و استخراج ویژگی، ماژول رده‌بندها و استفاده از روش‌های یادگیری توده‌ای. فلوچارت الگوریتم پیشنهادی در شکل 4 نشان داده شده است.

کلاس‌بندی توسط ماشین بردار پشتیبان با ایجاد ابر صفحات تفکیک‌کننده بهینه¹ انجام می‌شود. ابر صفحات تفکیک‌کننده بهینه حاشیه میان دونقطه مربوط به دو کلاس مجزا که کم‌ترین فاصله را از یکدیگر دارند بیشینه می‌کند. در نتیجه نامعادله (5) برای تمامی نقاط ورودی معتبر خواهد بود [16]:

$$y_i (w^T x_i + b) \geq 1, \text{ for all } x_i, i = 1, 2, \dots, l \quad (5)$$

رابطه (5) یک مسأله برنامه‌ریزی مرتبه دو بوده و برای حل آن باید از ضرایب لاگرانژ استفاده شود. نمونه‌های آموزشی که رابطه (5) را ارضا می‌کنند، بردارهای پشتیبان² نامیده می‌شوند. پس از حل مسأله، مقدار بهینه بایاس به صورت رابطه (6) تعیین می‌شود:

$$b^* = y_i - w^{*T} x_i \quad (6)$$

برای هر یک از بردارهای پشتیبان x_i تابع تصمیم‌گیری بهینه³ به صورت رابطه (7) حاصل می‌شود:

$$f(x) = \text{sgn} \left(\sum_{i=1}^l y_i \alpha_i^* x^T x_i + b^* \right) \quad (7)$$

که در آن α_i^* همان ضرایب لاگرانژ بهینه هستند [16]. در ماشین بردار پشتیبان غیرخطی، داده‌های آموزش با استفاده از تابع کرنل $K(x_i, x_j)$ به صورت غیرخطی به فضای مشخصه‌ای با بُعد بالا نگاشت شده و سپس در این فضا به صورت خطی از یکدیگر جدا می‌شوند [27]. توابع کرنل شامل توابع خطی، چندجمله‌ای (با رابطه (8)) و تابع با پایه شعاعی گوسی⁴ (با رابطه (9)) هستند:

$$K(x_i, x_j) = (x_i \cdot x_j + 1)^d \quad (8)$$

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (9)$$

که در آن d و σ به ترتیب مرتبه تابع چندجمله‌ای و پارامتر هموارکننده تابع با پایه شعاعی گوسی هستند. پس از انتخاب تابع کرنل، مسأله برنامه‌ریزی مرتبه دو تغییر کرده و پس از آموزش، تابع تصمیم‌گیری (10) حاصل می‌شود:

$$f(x) = \text{sgn} \left(\sum_{i=1}^l y_i \alpha_i^* K(x_i, x_j) + b^* \right) \quad (10)$$

برای مشاهده جزئیات هر یک از این رده‌بندها می‌توان به [24] و [16] مراجعه کرد.

روش‌های یادگیری توده‌ای از تنوع زیادی برخوردار هستند [28]. در این پژوهش از الگوریتم‌های (1) رأی اکثریت⁵، (2) ترکیب‌کننده پیوسته⁶ (با ترکیب‌کننده گوناگون)، (3) بیز ساده⁷ و (4) میانگین وزن‌دار⁸ استفاده شده است [28].

3-3-1- رأی اکثریت

فرض شود که برچسب خروجی رده‌بندها به صورت بردارهای دودویی c -بُعدی $[d_{i,1}, \dots, d_{i,c}]^T \in \{0,1\}, i = 1, \dots, L$ داده شده‌اند، که در آن L تعداد رده‌بندها و c تعداد کلاس‌ها است. اگر رده‌بند D_i مشاهده‌ای را به کلاس ω_j مربوط کرده باشد، آنگاه $d_{i,j} = 1$ و در غیر این صورت $d_{i,j} = 0$ رأی اکثریت

1- Optimal Separating Hyper-planes (OSH)

2- Support Vector (SV)

3- Optimal Decision Function (ODF)

4- Gaussian radial basis function (GRBF)

5- Majority Vote

6- Continuous Combiner

7- Naive Bayes

8- Weighted Average

9- Confusion Matrix

4-3-3- کاهش بُعد و استخراج ویژگی

4-3-3-1- گام نخست: کاهش بُعد

در این گام مؤلفه‌های طیف توان به صورت جفت جفت (گام جمع‌زنی 2) با یکدیگر جمع شده و در یک بردار دیگر قرار می‌گیرند. باید توجه داشت که می‌توان به جای جمع کردن هر دو مؤلفه متوالی طیف توان، هر n مؤلفه متوالی را با یکدیگر جمع نمود و به این ترتیب کاهش بُعد بیشتری انجام می‌شود. سپس طیف توان ایجاد شده در مرحله قبل با گام 2 برگزینش می‌شود؛ به این معنا که از هر دو عضو متوالی تنها یکی انتخاب می‌شود.

4-3-3-2- گام دوم: استخراج ویژگی

در این گام با استفاده از روش آنالیز مؤلفه‌های اصلی، طیف توان از فضای اصلی با بُعد بالا به زیر فضایی با بُعد مناسب نگاشت می‌شود. برای این منظور خروجی گام نخست (کاهش بُعد)، با استفاده از روش آنالیز مؤلفه‌های اصلی به زیر فضایی با بُعد مناسب نگاشت می‌شود. باید توجه داشت که برای انتخاب بُعد زیر فضای ثانویه، می‌توان از بزرگی مقادیر ویژه متناظر با بردارهای ویژه بهره برد. می‌توان نشان داد که نسبت هر مقدار ویژه به حاصل جمع کلیه مقادیر ویژه برابر است با میزان واریانس داده‌ها در راستای مؤلفه اصلی متناظر در زیر فضای ثانویه؛ از همین رو با انتخاب تعداد کافی از بردارهای ویژه، می‌توان مقدار مناسبی از اطلاعات را در زیر فضای ثانویه حفظ کرد. پس از انتخاب تعداد مؤلفه‌های اصلی، بردار ویژه‌های متناظر با این تعداد مقادیر ویژه از ماتریس کوواریانس ماتریس توان استخراج شده و با ضرب آن‌ها در ماتریس توان، بُعد این ماتریس کاهش می‌یابد. پس از انجام این مراحل، داده‌ها برای ورود به رده‌بندها آماده هستند.

4-4- مازول رده‌بندها

پس از تولید مشخصه‌های مناسب برای اصوات، از سه رده‌بنده پرسپترون چندلایه، شبکه عصبی احتمالاتی و ماشین بردار پشتیبان برای کلاسه‌بندی سیگنال‌ها استفاده می‌شود.

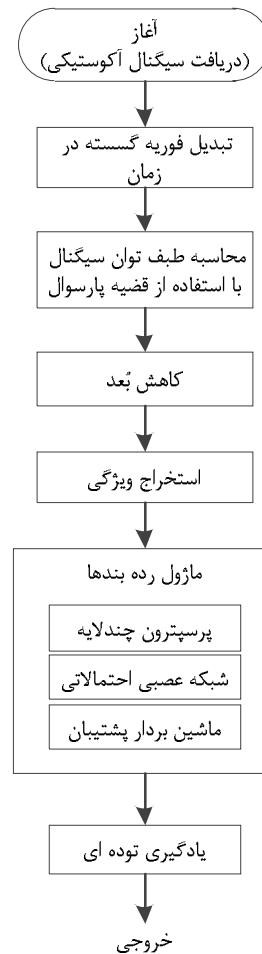
4-5- یادگیری توده‌ای

در انتها، خروجی سه رده‌بند مذکور به روش‌های یادگیری توده‌ای مختلف سپرده شده تا تصمیم نهایی پیرامون کلاس سیگنال‌ها اتخاذ شود.

5- آزمون‌های کارایی

جدول 1 پارامترهای موجود در الگوریتم پیشنهادی را که در تمامی آزمون‌های کارایی مشترک هستند (مگر خلاف آن ذکر شود) نشان می‌دهد. در تمامی روش‌های استخراج ویژگی، همواره کاهش بُعد از فضای اصلی به زیر فضای ثانویه به نحوی انجام شده است که واریانس داده‌های تصویر شده در زیر فضای ثانویه برابر 95 درصد باشد. برای آموزش شبکه پرسپترون چندلایه از الگوریتم پس‌انتشار ارتجاعی¹ و نیز از تابع با پایه شعاعی گوسی به عنوان تابع کرنل در ماشین بردار پشتیبان استفاده شده است. تعداد نرون‌های لایه پنهان در پرسپترون چندلایه و نیز مقادیر پارامتر هموار کننده در شبکه عصبی احتمالاتی و ماشین بردار پشتیبان در هنگام آموزش تمامی آزمون‌ها در یک بازه مشخص تغییر داده شده و رده‌بندهایی با بهترین ساختار برای مرحله آزمون انتخاب شده‌اند.

پس از انجام مرحله آموزش رده‌بندها، سیگنال‌های اصلی به نویز سفید



شکل 4 فلوجارت الگوریتم پیشنهادی

4-1- آماده‌سازی اصوات

گام نخست در آماده‌سازی اصوات بهنجار کردن آن‌ها است. این موضوع از آن جهت ضروری است که در هنگام ضبط اصوات، فاصله منابع صوت با هیدروفون متفاوت بوده و همین مسأله سبب می‌شود شدت (بلندی) اصوات با یکدیگر متفاوت باشد. پس از بهنجار کردن اصوات، میانگین سیگنال‌ها حذف می‌شود. در گام دوم، سیگنال‌های اولیه (که طول‌های مختلفی دارند) به بخش‌های هم طولی تقسیم می‌شوند. سپس هر یک از این بخش‌های تولیدشده در یک سطر از ماتریسی که بانام ماتریس ورودی شناخته می‌شود، قرار می‌گیرد.

4-2- پیش‌پردازش

در این بخش نخست تبدیل فوریه گسسته در زمان هر یک از سطرهای ماتریس ورودی با استفاده از الگوریتم‌ها تبدیل فوریه سریع محاسبه می‌شود. در محاسبه تبدیل فوریه، تبدیل به صورت تبدیل فوریه n -نقطه‌ای محاسبه می‌شود؛ همچنین فرکانس‌ها در فواصل f_s/n' خوانده می‌شوند، که در آن n' کوچک‌ترین عدد توان 2 بزرگ‌تر از طول سیگنال ورودی است. پس از محاسبه تبدیل فوریه گسسته در زمان، بر طبق قضیه پارسوال و با استفاده از رابطه (1)، حاصل ضرب ضرایب تبدیل فوریه $X[k]$ در مزدوجشان برای هر یک از سیگنال‌های ماتریس ورودی محاسبه شده تا طیف توان $P[k]$ سیگنال‌ها حاصل شود. سپس طیف توان هر یک از سطرهای ماتریس ورودی در سطر متناظر از ماتریس دیگری بانام ماتریس توان قرار داده می‌شود.

1- Resilient Back Propagation

جدول 2 روش‌های مختلف مورد استفاده در الگوریتم پیشنهادی

مرحله	روش‌ها
پنجره و همپوشانی	- استفاده نکردن از پنجره $A_1 =$ - استفاده کردن از پنجره $A_2 =$ - همپوشانی بخش‌های متوالی سیگنال‌ها با هم $A_3 =$
جمع‌زنی و برگزینش	- جمع‌زنی و برگزینش با گام $B_1 = 2$ - جمع‌زنی و برگزینش با گام‌های مختلف $B_2 =$
کاهش بعد	- آنالیز مؤلفه‌های اصلی $C_1 =$ - آنالیز مؤلفه‌های اصلی با وزن متغیر $C_2 =$ - آنالیز مؤلفه‌های اصلی احتمالاتی $C_3 =$ - آنالیز مؤلفه‌های اصلی فازی $C_4 =$ - آنالیز مؤلفه‌های اصلی با تابع کرنل $C_5 =$ - آنالیز مؤلفه‌های اصلی فازی با تابع کرنل $C_6 =$ - آنالیز تفکیک‌کنندگی خطی $C_7 =$ - آنالیز تفکیک‌کنندگی تعمیم‌یافته $C_8 =$ - آنالیز تفکیک‌کنندگی غیر پارامتریک $C_9 =$
یادگیری توده‌ای	- بدون روش یادگیری توده‌ای $D_1 =$ - روش یادگیری توده‌ای رأی اکثریت $D_2 =$ - روش یادگیری توده‌ای ترکیب‌کننده پیوسته $D_3 =$ - روش یادگیری توده‌ای بیز ساده $D_4 =$ - روش یادگیری توده‌ای میانگین وزن‌دار $D_5 =$

جدول 3 درصد کلاسه‌بندی صحیح سه رده‌بند پرسپترون چندلایه، شبکه عصبی احتمالاتی و ماشین بردار پشتیبان در آزمون‌های کارایی مربوط به روش‌های مختلف

نام آزمون	پرسپترون چندلایه	شبکه عصبی احتمالاتی	ماشین بردار پشتیبان
$A_1A_3B_1C_1D_1$ (آزمون پایه)	86/84	94/94	90/21
$A_2A_3B_1C_1D_1$ (پنجره همینگ)	71/07	60/62	14/33
$A_1A_3(25\%)B_1C_1D_1$	83/22	94/18	90/47
$A_1A_3(50\%)B_1C_1D_1$	82/20	96/12	94/43
$A_1A_3B_2(AS = 5, DS = 2)C_1D_1$	78/58	97/80	90/55
$A_1A_3B_2(AS = 15, DS = 2)C_1D_1$	81/78	96/96	91/03
$A_1A_3B_2(AS = 2, DS = 4)C_1D_1$	75/37	81/96	84/82
$A_1A_3B_1C_2D_1$	70/82	15/76	77/65
$A_1A_3B_1C_3D_1$	71/75	83/64	61/46
$A_1A_3B_1C_4D_1$	87/26	94/94	90/21
$A_1A_3B_1C_5D_1$	88/70	96/29	94/26
$A_1A_3B_1C_6D_1$	85/66	40/13	84/23
$A_1A_3B_1C_7D_1$	6/32	15/76	4/38
$A_1A_3B_1C_8D_1$	39/03	17/11	22/34
$A_1A_3B_1C_9D_1$	5/31	3/37	1/51

جدول 4 نتایج مربوط به ترکیب کردن خروجی سه رده‌بند پرسپترون

گوسی با نسبت سیگنال به نویز 5- آغشته شده و به‌عنوان سیگنال‌های آزمون استفاده می‌شوند. با توجه به اینکه بخش‌هایی از الگوریتم پیشنهادی، همانند وزن‌های اولیه رده‌بند پرسپترون چندلایه در هنگام آموزش و یا نویز سفید گوسی افزوده شده به سیگنال، در هر بار اجرا متفاوت از اجرای پیشین خواهد بود، به‌منظور اطمینان از صحت پاسخ‌ها و عدم وابستگی آن‌ها به شرایط اولیه آزمون کارایی، هر آزمون به‌صورت چهار مرحله مستقل از هم اجرا شده و نتایج ارائه‌شده میانگین این چهار اجرای مستقل خواهد بود.

جدول 2 روش‌های جایگزین برای هر یک از پنج مرحله الگوریتم پیشنهادی را نشان می‌دهد. همان‌طور که مشاهده می‌شود برای هر یک از روش‌ها یک نماد مشخص در نظر گرفته شده و آزمون‌های کارایی بر اساس این نمادها تولید و نام‌گذاری می‌شوند. بر اساس جدول 2، یک آزمون پایه به‌صورت $A_1A_3B_1C_1D_1$ پیشنهاد شده و چگونگی اثرگذاری هر یک از روش‌ها بر بهبود کارایی الگوریتم پیشنهادی در مقایسه با این آزمون پایه سنجیده می‌شوند. در آزمون پایه، در هنگام تقسیم سیگنال‌ها به بخش‌های هم طول از پنجره استفاده نشده، سیگنال‌ها با یکدیگر هم‌پوشانی نداشته، گام جمع‌زنی¹ و برگزینش² برابر 2 در نظر گرفته شده، از روش آنالیز مؤلفه‌های اصلی برای استخراج ویژگی استفاده شده و نتایج تنها مربوط به استفاده از رده‌بندها (بدون یادگیری توده‌ای) خواهد بود. به همین ترتیب سایر آزمون‌ها با تغییر قسمت‌های مختلف آزمون پایه ایجاد می‌شوند.

6- تشریح و بحث نتایج

جدول 3 نتایج آزمون‌های کارایی انجام‌شده با الگوریتم پیشنهادی با داده‌های آغشته به نویز با نسبت سیگنال به نویز 5- را نشان می‌دهد.

در آزمون $A_2A_3B_1C_1D_1$ از یک پنجره همینگ هم طول با سیگنال در هنگام جداسازی سیگنال‌های هم طول از سیگنال‌های خام استفاده شده است. در این آزمون، در هنگام جداسازی سیگنال‌ها، هر بخش در یک پنجره همینگ ضرب شده است تا از به وجود آمدن اثر گیبز در هنگام محاسبه تبدیل فوریه ممانعت به عمل آید. در آزمون‌های مربوط به بررسی اثر هم‌پوشانی سیگنال‌ها با هم، عبارت 25% (یا 50%) به معنای هم‌پوشانی دو سیگنال متوالی به‌اندازه 25% (یا 50%) طول سیگنال‌ها است. آزمون‌های مربوط به بررسی اثر گام جمع‌زنی و برگزینش در سه حالت انجام شده است. در حالت اول و دوم اثر افزایش گام جمع‌زنی به ازای گام برگزینش ثابت و در حالت سوم اثر افزایش گام برگزینش به ازای گام جمع‌زنی ثابت بررسی شده است.

جدول 1 مشخصه‌ها و پارامترهای الگوریتم پیشنهادی

مشخصه یا پارامتر	مقدار
تعداد مؤلفه‌های اصلی به مقداری باشد که	-
استخراج ویژگی	واریانس داده‌ها در زیر فضای ثانویه بیش از 95 درصد شود
الگوریتم آموزش شبکه عصبی	- پس‌انتشار خطای ارتجاعی
ساختار و تعداد نرون‌ها لایه	- با یک لایه پنهان
پنهان شبکه عصبی	تعداد نرون‌ها در بازه 5 تا 70
شبکه عصبی احتمالاتی	- پارامتر هموار کننده در بازه 0/005 تا 5
ماشین بردار پشتیبان	- پارامتر هموار کننده در بازه 1 تا 12

1- Averaging Step
2- Downsampling Step

جدول 4 درصد کلاسه‌بندی صحیح در آزمون‌های کارایی مربوط به روش‌های یادگیری توده‌ای مختلف

نام آزمون	نتایج روش یادگیری توده‌ای
$A_1A_3B_1C_1D_2$	97/72
$A_1A_3B_1C_1D_3$ (میانگین)	97/72
$A_1A_3B_1C_1D_3$ (بیشینه)	96/45
$A_1A_3B_1C_1D_3$ (کمینه)	87/01
$A_1A_3B_1C_1D_3$ (حاصل ضرب)	87/01
$A_1A_3B_1C_1D_3$ (میانه)	97/72
$A_1A_3B_1C_1D_4$	97/21
$A_1A_3B_1C_1D_5$	93/59

جدول 5 مقادیر مطلوب مشخصه‌ها یا پارامترهای قسمت‌های مختلف الگوریتم پیشنهادی

مشخصه یا پارامتر	مقدار
طول سیگنال و میزان هم‌پوشانی	- طول سیگنال‌ها 1 ثانیه - همپوشانی 50 درصد طول سیگنال‌ها
جمع‌زنی و برگزینش	- گام جمع‌زنی = 15 - گام برگزینش = 1
استخراج ویژگی	- روش آنالیز مؤلفه‌های اصلی با استفاده از تابع با پایه شعاعی با پارامتر هموار کننده 2 به‌عنوان تابع کرنل - تعداد مؤلفه‌های اصلی به مقداری باشد که واریانس داده‌ها در زیر فضای ثانویه بیش از 95 درصد شود
رده‌بندها	- پرسپترون چندلایه - شبکه عصبی احتمالاتی - ماشین بردار پشتیبان
یادگیری توده‌ای	- روش‌های رأی اکثریت و ترکیب‌کننده پیوسته با تابع ترکیب‌کننده میانگین ساده و میانه

ارزیابی قرار گرفته و پس از انتخاب بهترین روش، مقادیر بهینه سایر پارامترهای اثرگذار بر کارایی الگوریتم بررسی می‌شود.

شکل 5 نشان می‌دهد که کارایی روش‌های یادگیری توده‌ای رأی اکثریت و ترکیب‌کننده پیوسته با تابع ترکیب‌کننده میانگین و میانه در هر سه مقدار نسبت سیگنال به نویز با یکدیگر برابر هستند. به علت سادگی ساختار روش یادگیری توده‌ای رأی اکثریت، در ادامه از همین روش استفاده می‌شود.

شکل 6 کارایی الگوریتم به ازای مقادیر مختلف طول سیگنال را نشان می‌دهد. همان‌طور که مشاهده می‌شود کارایی الگوریتم در نسبت‌های سیگنال به نویز بالا (5-) به ازای طول‌های مختلف سیگنال تفاوت فراوانی ندارد، درحالی‌که با کاهش نسبت سیگنال به نویز، اختلاف معناداری میان نتایج مربوط به طول‌های مختلف مشاهده می‌شود. با توجه به شکل 6، افزایش طول سیگنال‌ها سبب بهبود کارایی الگوریتم پیشنهادی به‌ویژه در نسبت‌های سیگنال به نویز پایین می‌شود، از همین رو، طول سیگنال‌ها برابر 2 ثانیه در نظر گرفته می‌شود.

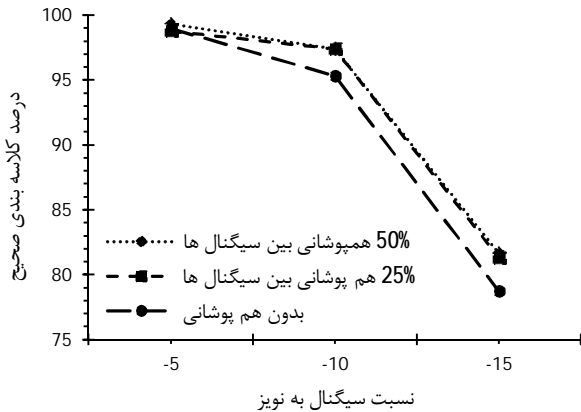
چندلایه، شبکه عصبی احتمالاتی و ماشین بردار پشتیبان با استفاده از چهار روش یادگیری توده‌ای را نشان می‌دهد. آزمون‌های مربوط به روش یادگیری توده‌ای با الگوریتم ترکیب‌کننده پیوسته برای پنج تابع ترکیب‌کننده میانگین ساده، حاصل ضرب، میانه، بیشینه و کمینه انجام شده است.

بر اساس نتایج موجود در جداول 3 و 4 می‌توان عوامل مؤثر بر بهبود کارایی الگوریتم پیشنهادی را شناسایی نمود؛

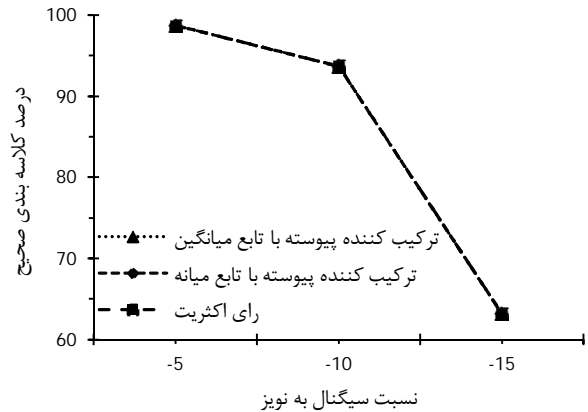
- نتایج جدول 3 نشان می‌دهد که ایجاد هم‌پوشانی بین سیگنال‌ها، سبب افزایش کارایی دو رده‌بند شبکه عصبی احتمالاتی و ماشین بردار پشتیبان می‌شود.
- نتایج جدول 3 نشان می‌دهد که افزایش گام جمع‌زنی از 2 به 15، سبب بهبود کارایی دو رده‌بند شبکه عصبی احتمالاتی و ماشین بردار پشتیبان می‌شود.
- نتایج جدول 3 نشان می‌دهد که استفاده از الگوریتم کاهش بُعد آنالیز مؤلفه‌های اصلی با استفاده از تابع با پایه شعاعی گوسی با پارامتر هموار کننده 2 به‌عنوان تابع کرنل سبب بهبود کارایی هر سه رده‌بند پرسپترون چندلایه، شبکه عصبی احتمالاتی و ماشین بردار پشتیبان می‌شود.
- نتایج جدول 4 نشان می‌دهد که بهره بردن از روش‌های یادگیری توده‌ای رأی اکثریت یا ترکیب‌کننده پیوسته با تابع ترکیب‌کننده میانگین ساده یا میانه سبب بهبود نتایج الگوریتم شناسایی به‌طور قابل توجهی می‌شود.

مشاهده می‌شود که با اعمال برخی تغییرات بر روی بخش‌های مختلف الگوریتم پیشنهادی، کارایی یک رده‌بند کاهش یافته درحالی‌که کارایی دو رده‌بند دیگر افزایش می‌یابد. از مزایای استفاده از روش‌های یادگیری توده‌ای این است که با ترکیب کردن نتایج هر سه رده‌بند، نقاط ضعف هر یک را به‌وسیله نقاط قوت دیگری پوشش داده و در نتیجه در حالت کلی کارایی افزایش می‌یابد. همچنین مشاهده می‌شود که روش‌های مبتنی بر آنالیز تفکیک‌کنندگی، نتایج قابل قبولی ارائه نمی‌دهند. علت این امر را می‌توان به حساسیت این روش نسبت به نویز نسبت داد؛ زیرا در شرایطی که داده‌ها آزمون آغشته به نویز نباشند، کارایی این روش‌ها نسبت به روش‌های مبتنی بر آنالیز مؤلفه‌های اصلی بالاتر خواهد بود (اگرچه نتایج آن در اینجا ذکر نشده است). همچنین با مقایسه نتایج جداول 3 و 4 مشاهده می‌شود که استفاده از روش‌های یادگیری توده‌ای سبب می‌شود کارایی الگوریتم پیشنهادی نسبت به بهترین رده‌بند در آزمون پایه در حدود 2/8 درصد افزایش یابد. در نتیجه با مدنظر قرار دادن هزینه محاسباتی پایین استفاده از روش‌های یادگیری توده‌ای، استفاده از این روش‌ها اکیداً توصیه می‌شود.

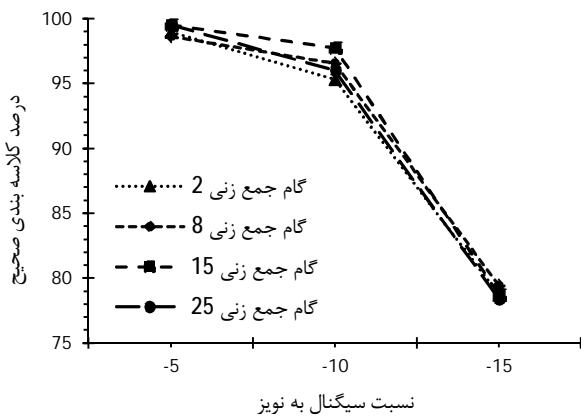
بر اساس نتایج حاصل شده، الگوریتمی شامل حالت بهینه هر یک از بخش‌ها تشکیل می‌شود. جدول 5 بخش‌های مختلف این الگوریتم را نشان می‌دهد. سپس کارایی الگوریتم پیشنهادی به ازای تغییر مقدار پارامتر یا نوع مشخصه هر از سطرها جدول 5 بررسی شده تا مقادیر مطلوب هر یک مشخص شوند. در این بخش کارایی الگوریتم پیشنهادی با سیگنال‌های آغشته به نویز گوسی با نسبت‌های سیگنال به نویز 5-، 10- و 15- ارزیابی می‌شود. باید توجه داشت که در ادامه، پس از یافتن حالت بهینه هر یک از پارامترها یا مشخصه‌ها، در آزمون‌های بعدی نیز از همین مقدار بهینه استفاده می‌شود تا در نهایت بتوان بالاترین کارایی الگوریتم پیشنهادی را مشخص کرد. برای این منظور نخست کارایی روش‌های یادگیری توده‌ای مختلف مورد



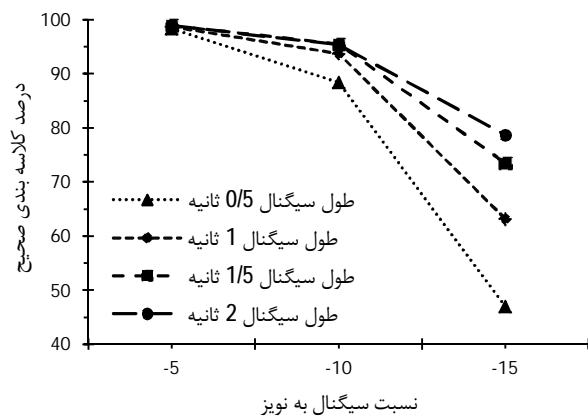
شکل 7 ارزیابی کارایی الگوریتم پیشنهادی به ازای مقادیر مختلف هم‌پوشانی بین سیگنال‌های متوالی پیشنهادی



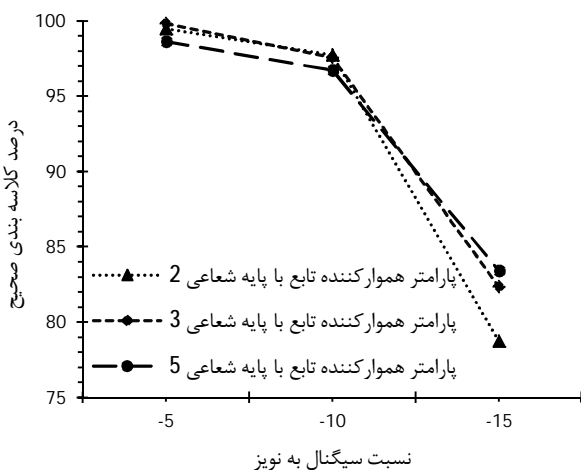
شکل 5 ارزیابی عملکرد روش‌های یادگیری توده‌ای مختلف بر کارایی الگوریتم



شکل 8 ارزیابی کارایی الگوریتم پیشنهادی به ازای گام‌های جمع‌زنی مختلف پیشنهادی



شکل 6 ارزیابی کارایی الگوریتم پیشنهادی به ازای طول‌های مختلف سیگنال پیشنهادی



شکل 9 ارزیابی کارایی الگوریتم پیشنهادی به ازای مقادیر مختلف پارامتر هموار کننده تابع با پایه شعاعی در روش آنالیز مؤلفه‌های اصلی با استفاده از تابع کرنل

شکل 7 کارایی الگوریتم شناسایی به ازای مقادیر مختلف هم‌پوشانی بین سیگنال‌ها را نشان می‌دهد. مشاهده می‌شود که همانند حالت قبل، ایجاد هم‌پوشانی بین سیگنال‌ها سبب بهبود کارایی الگوریتم پیشنهادی به‌ویژه در نسبت‌های سیگنال به نویز پایین می‌شود. با توجه به شکل 7 مشاهده می‌شود که کارایی الگوریتم به ازای هم‌پوشانی 25% و 50% شباهت زیادی به یکدیگر دارند، با این حال در نسبت سیگنال به نویز -5، کارایی الگوریتم به ازای هم‌پوشانی 50% بالاتر است.

شکل 8 کارایی الگوریتم پیشنهادی را به ازای مقادیر مختلف گام جمع‌زنی نشان می‌دهد. در نگاه کلی به ازای مقادیر مختلف گام جمع‌زنی، کارایی الگوریتم پیشنهادی به تقریباً یکسان است، با این حال در نسبت‌های سیگنال به نویز -5 و -10، کارایی به ازای گام جمع‌زنی 15 نسبت به سایر مقادیر بالاتر است.

شکل‌های 9 و 10 کارایی الگوریتم پیشنهادی را به ترتیب به ازای مقادیر مختلف پارامتر هموار کننده تابع با پایه شعاعی و مرتبه تابع چندجمله‌ای در روش استخراج ویژگی آنالیز مؤلفه‌های اصلی با استفاده از تابع کرنل را نشان می‌دهند. با مقایسه شکل‌های 9 و 10 مشاهده می‌شود که در حالت کلی، استفاده از تابع چندجمله‌ای با رابطه (8) و با مرتبه 5 بهترین نتایج را در تمامی نسبت‌های سیگنال به نویز به دنبال دارد.

با انجام آزمون‌های کارایی و بررسی عملکرد روش‌های مختلف، و سپس یافتن مقدار بهینه پارامترهای اثرگذار بر روی هر روش، الگوریتم پیشنهادی

با مشخصه‌های کلی ذکر شده در جدول 5 و نیز مقادیر بهینه انتخاب شده بر اساس شکل‌های 5 تا 10، توانایی تفکیک سیگنال‌های صوتی هشت کشتی آغشته به نویز سفید گوسی با نسبت‌های سیگنال به نویز -5، -10 و

[4] W. S. Filho, J. M. Seixas, N. N. Moura, Preprocessing passive sonar signals for neural classification, *IET Radar, Sonar and Navigation*, Vol. 5, pp. 605-612, 2011.

[5] W. S. Filho, J. M. Seixas, L. P. Caloba, Enlarging neural class detection capacity in passive sonar systems, *IEEE International Symposium on Circuits and Systems-ISCAS 2002*, Vol. 3, pp. 105-108, 2002.

[6] H. Wu, M. Siegel, P. Khosla, Vehicle sound signature recognition by frequency vector principal component analysis, *IEEE Instrumentation and Measurement Technology Conference*, 1998.

[7] C. H. Chen, J. D. Lee, M. C. Lin, Classification of underwater signals using wavelet transforms and neural networks, *Mathematical and Computer Modelling*, Vol. 27, pp. 47-60, 1998.

[8] C. H. Chen, J. D. Lee, M. C. Lin, Classification of underwater signals using neural networks, *Tamkang Journal of Science and Engineering*, Vol. 3, pp. 31-48, 2000.

[9] B. P. Howell, S. Wood, S. Koksai, Passive sonar recognition and analysis using hybrid neural networks, *IEEE Proceedings Oceans*, Vol. 4, pp. 1917-1924, 2003.

[10] M. Farrokhrzooz, M. Karimi, Ship noise classification using probabilistic neural network and AR model coefficients, *Oceans 2005 - Europe* Vol. 2, 2005.

[11] M. Farrokhrzooz, M. Karimi, Marine vessels acoustic radiated noise classification in passive sonar using probabilistic neural network and spectral features, *Intelligent Automation and Soft Computing*, Vol. 17, pp. 369-383, 2011.

[12] H. Rezazadeh, H. Keyvani, B. Gholami, Ocean acoustic signal processing using Markov Model, in *12th Iranian Student Conference on Electrical Engineering*, Tabriz, 2009. (In Persian)

[13] A. V. Oppenheim, R. W. Schaffer, *Discrete-Time Signal Processing*, 3rd ed.: Pearson Higher Education, 2010.

[14] S. J. Raudys, A. K. Jain, Small sample size effects in statistical pattern recognition: recommendations for practitioners, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13 pp. 252-264, 1991.

[15] S. J. Raudys, V. Pikelis, On dimensionality, sample size, classification error, and complexity of classification algorithms in pattern recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 2, pp. 243-251, 1980.

[16] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification*, 2nd ed., New York: Wiley, 2001.

[17] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed.: Springer, 2002.

[18] B. Scholkopf, A. Smola, K. R. Muller, Nonlinear component analysis as a kernel eigenvalue problem, *Neural Computation (MIT Press Journal)*, Vol. 5, pp. 1299-1319, 1998.

[19] H. F. Pop, Principal components analysis based on a fuzzy sets approach, *Univ. Babeş-Bolyai, Informatica*, Vol. 46, pp. 45-52, 2001.

[20] W. Xiaohong, Z. Jianjiang, Fuzzy principal component analysis and its kernel based model, *Journal of Electronics (China)*, Vol. 24, pp. 772-775, 2007.

[21] M. E. Tipping, C. M. Bishop, Probabilistic principal component analysis, *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, Vol. 61, pp. 611-622, 1999.

[22] C. H. Park, H. Park, Nonlinear discriminant analysis using kernel functions and the generalized singular value decomposition, *SIAM Journal on Matrix Analysis and Applications*, Vol. 27, pp. 87-102, 2005.

[23] K. Fukunaga, J. M. Mantock, Nonparametric Discriminant Analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 5, pp. 671-678, 1983.

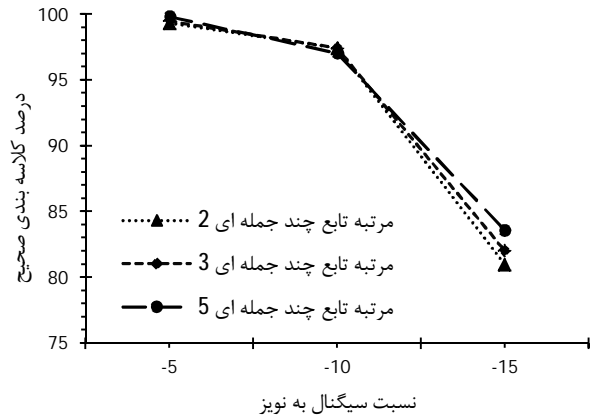
[24] S. Haykin, *Neural Networks, A Comprehensive Foundation*: Prentice Hall, 1999.

[25] K. Z. Mao, K. C. Tan, W. Ser, Probabilistic neural-network structure determination for pattern classification, *IEEE Transactions on Neural Networks* Vol. 11, pp. 1009-1016, 2000.

[26] D. F. Specht, Probabilistic neural network, *Neural Networks*, Vol. 3, pp. 109-118, 1990.

[27] C. Burges, A tutorial on support vector machines for pattern recognition, *Data Mining and Knowledge Discovery*, Vol. 2, No. 2, pp. 121-167, 1998.

[28] L. I. Kuncheva, *Combining Pattern Classifiers, Methods and Algorithms*, New Jersey: John Wiley & Sons, Inc, 2004.



شکل 10 ارزیابی کارایی الگوریتم پیشنهادی به ازای مقادیر مختلف مرتبه تابع چندجمله‌ای در روش آنالیز مؤلفه‌های اصلی با استفاده از تابع کرنل

15- را به ترتیب با دقت 99/83، 97/06 و 83/56 درصد دارا است.

7- نتیجه گیری

در این پژوهش روشی برای شناسایی سیگنال‌های صوتی هشت کشتی ارائه شده است. در این روش از تبدیل فوری به‌عنوان پیش‌پردازش، الگوریتم‌های متنوع کاهش بُعد و نیز روش‌های یادگیری توده‌ای گوناگون استفاده شده است. با انجام آزمون‌های کارایی متنوع و متعدد، چگونگی اثرگذاری هر یک از روش‌ها و پارامترهای اثرگذار در آن‌ها بر کارایی الگوریتم پیشنهادی مورد ارزیابی قرار گرفته‌اند. وجه تمایز اصلی این پژوهش با تحقیقات مشابه را می‌توان به استفاده از روش‌های کاهش بُعد متنوع با ساختارهای پیچیده‌تر که تاکنون در شناسایی سیگنال‌های صوتی کشتی‌ها مورد بررسی قرار نگرفته بودند و نیز استفاده از روش‌های یادگیری توده‌ای به منظور ترکیب رده‌بندها با یکدیگر بیان کرد. نتایج به وضوح بهبود کارایی استفاده از این روش‌ها نسبت به روش‌های معمول را نشان می‌دهد. نتایج نشان می‌دهند که در حالت بهینه، الگوریتم پیشنهادی قادر خواهد بود سیگنال‌های صوتی هشت کشتی را در هنگام آزمون و در شرایطی که به نویز سفید گوسی با نسبت سیگنال به نویز 5-، 10- و 15- آغشته شده‌اند به ترتیب با دقت 99/83، 97/06 و 83/56 درصد کلاسه‌بندی کند.

8- مراجع

[1] R. J. Urick, *Principles of Underwater Sound*: McGraw-Hill, 1983.

[2] W. S. Filho, J. M. Seixas, L. P. Caloba, Principal component analysis for classifying passive sonar signals, *IEEE International Symposium on Circuits and Systems*, 2001.

[3] P. A. A. Esquefi, L. W. P. Biscainho, V. Vdlimdki, M. Karjalainen, Removal of long pulses from audio signals using two-pass split-window filtering, *Journal of the Audio Engineering Society*, 2002.