

ارائه یک روش مؤثر برای به دست آوردن امضای آکوستیک

سودابه افشار^۱، سید جواد کاظمی تبار^۲، عطاالله ابراهیمزاده^۳

۱- کارشناسی ارشد ۲- استادیار، ۳- استاد، دانشگاه صنعتی نوشیروانی بابل

(دریافت: ۱۴۰۱/۱۰/۳۰، بازنگری: ۱۴۰۲/۰۱/۲۴، پذیرش: ۱۴۰۲/۰۱/۳۱، انتشار: ۱۴۰۲/۰۳/۰۱)

DOR: <https://dor.isc.ac/dor/20.1001.1.26762935.1402.14.1.3.3>

چکیده

تشخیص و طبقه‌بندی شناورهای دریایی از سیگنال‌های ساطع شده از آن‌ها، یک وظیفه مهم و با کاربردهای نظامی حیاتی، به‌ویژه در شرایط بحرانی و جنگ است. از زمان‌های قدیم تا به امروز، این وظیفه، به‌صورت سنتی توسط افسرهای سونار به کمک سیگنال‌های آکوستیکی دریافت شده توسط هیدروفون‌های سونار انجام می‌شد. امروزه، ارائه راهکارهایی که این وظیفه را به‌صورت خودکار و با کمترین دخالت انسان انجام دهند، یک ضرورت اجتناب‌ناپذیر است. در این پژوهش، با استفاده از روش‌های مختلف پیش‌پردازش و آماده‌سازی داده، ابعاد مجموعه داده تحت سناریوهای خاص گسترش داده می‌شود. با پردازش سیگنال، ویژگی‌های مختلفی را از مجموعه داده‌ای مربوط به سناریوهای مختلف استخراج می‌شود و توسط دانش یادگیری ماشین، الگوریتم‌های طبقه‌بندی متنوعی را به کار می‌رود. در نهایت، با محاسبه معیارهای ارزیابی مختلف، عملکرد سناریوهای مختلف بررسی و تحلیل می‌شود. نتایج این تحقیق، موفقیت‌آمیز بودن روش استخراج ویژگی ضرایب کپسترال فرکانسی مل (MFCC) را در هنگام به‌کارگیری الگوریتم‌های طبقه‌بندی مختلف، از جمله شبکه عصبی مصنوعی (ANN) و همچنین شکست روش شبکه عصبی پیچشی (CNN) را در هنگام آموزش توسط مجموعه داده‌های کوچک نشان می‌دهد. با مقایسه نتایج شبیه‌سازی معلوم شد با توجه به محدود بودن مجموعه داده مورد استفاده، روش ANN بهترین کارایی را ارائه است.

کلیدواژه‌ها: طبقه‌بندی شناورهای دریایی؛ هیدروفون‌های سونار؛ امضای آکوستیکی؛ گسترش داده‌ها؛ ضرایب کپسترال فرکانسی مل؛ شبکه عصبی مصنوعی؛ شبکه عصبی پیچشی.

An Effective Acoustic Signature Extraction Based on Optimization

S. Afshar, J. Kazemitabar^{*}, A. Ebrahimzadeh

Babol Noshirvani University of Technology

(Received: 2023/01/20 ; revised: 2023/04/13 ; Accepted: 2023/04/20 ; published: 2023/05/22)

Abstract

Identifying and classifying marine vehicles is of great significance in military applications. Traditionally, this task was performed by sonar officers based on acoustic signals received by sonar hydrophones. Nowadays, developing systems that perform this task automatically and with minimal human intervention is inevitable. In this study, using different augmentation methods, the size of this dataset is expanded under different scenarios. With the help of signal processing and machine learning knowledge, different features are extracted and next these features are classified using different classification algorithms. Finally, the performance of different scenarios is evaluated by calculating different measures. Results of this study show the success of the Mel-Frequency Cepstral Coefficient (MFCC) feature extraction method when using various classification algorithms, including Artificial Neural Network (ANN) and also the failure of Convolutional Neural Network (CNN) when training with small datasets. The overall comparison reveals that due to the small number of samples in the dataset, ANN presents the best performance.

Keywords: Marine Vehicle Classification; Sonar Hydrophone; Acoustic Signature; Data Augmentation; Mel-Frequency Cepstral Coefficients; Artificial Neural Network; Convolutional Neural Network.

*Corresponding Author E-mail: hoseini@mut.ac.ir

۱. مقدمه

می‌دهند و همچنین مقالاتی که تشخیص امضای آکوستیکی سایر اهداف زیر آب را بررسی می‌کنند، ارائه می‌شود.

در یک پژوهش، برای طبقه‌بندی سیگنال‌های آکوستیکی موجود در یک مجموعه داده واقعی که از ۱۱۰ نمونه داده متعلق به ۳ کلاس تشکیل می‌شود، از یک روش مبتنی بر دازش سیگنال دیجیتال برای پیش‌پردازش و استخراج ویژگی استفاده می‌کند. این مجموعه داده از ۴۲ نمونه از اصوات متعلق به قایق‌های سریع، ۲۶ نمونه از اصوات متعلق به کشتی‌های بازرگانی و ۴۲ نمونه از اصوات متعلق به نویز زیست‌محیطی دلفین تشکیل می‌شود. نتایج نشان می‌دهد، تبدیل فوریه زمان گسسته برای تجزیه و تحلیل سیگنال غیرایستا در مرحله پیش‌پردازش، سیگنال آکوستیکی به‌منظور تجزیه و تحلیل زمان - فرکانسی مناسب پنجره‌گذاری کارآمد است. همچنین، الگوریتم‌های شبکه عصبی مصنوعی (ANN) و ماشین بردار پشتیبان (SVM) را برای طبقه‌بندی و در هر دو روش به دقتی نزدیک به ۹۲ درصد دست می‌یابد [۵].

در یک پایان‌نامه در نظارت غیرفعال زیر آب به‌منظور بهبود دقت طبقه‌بندی خودکار کشتی‌های سطحی از نظر حجم و نوع، علاوه بر امضای آکوستیکی از امضاهای مغناطیسی و الکتریکی ساطع شده از کشتی‌های سطحی استفاده می‌کند. امضای الکتریکی کشتی نیز در اثر حرکت کشتی در میان آب، تجهیزات الکترونیکی پردازنده و محافظت کاتدی بدنه ایجاد می‌شود. بیشتر کشتی‌های سطحی برای جلوگیری از خوردگی بدنه از الکتریسیته استفاده می‌کنند که این امر در امضای الکتریکی ساطع شده از کشتی بیشترین مشارکت را دارد. این مرجع، از مجموعه داده‌ای از امضاهای آکوستیکی، مغناطیسی و الکتریکی که در سال ۲۰۱۳ از دو مکان مختلف در سوئد ضبط شده است، استفاده می‌کند. این مجموعه داده از ۱۱۷ ضبط ۳ تا ۶ دقیقه‌ای مربوط به کشتی‌های با ۴ حجم مختلف، یعنی کوچک، متوسط، بزرگ و غول‌آسا و متعلق به ۵ نوع مختلف، یعنی مسافربری، حامل‌های فله، کشتی‌های محتوی، نفت‌کش، کشتی‌های متنوع تشکیل می‌شود. این مرجع از این مجموعه داده برای طراحی دو طبقه‌بندی‌کننده ادغام سطح متوسط^۵ که یکی از آن‌ها کشتی‌های سطحی را از نظر حجم و دیگری کشتی‌های سطحی از نظر نوع طبقه‌بندی می‌کند، به کار می‌گیرد. نتایج طبقه‌بندی نشان می‌دهد، طبقه‌بندی‌کننده ادغام سطح متوسط در مقایسه با استفاده تنها از امضای آکوستیکی، عملکرد طبقه‌بندی را به ترتیب به میزان ۴/۲٪ و ۱۴٪ برای طبقه‌بندی حجم کشتی‌ها و نوع کشتی بهبود می‌دهد [۱].

شبکه عصبی ادراکی چندلایه (MLP)، عملیاتی‌ترین و متداول‌ترین مدل هوش مصنوعی برای طراحی طبقه‌بندی‌کننده

نظارت زیر آب^۱، یک زمینه مهم و با کاربردهای نظامی و غیرنظامی است. یک مسئله بسیار مهم در سامانه‌های نظارت زیر آب، مسئله تشخیص^۲ و طبقه‌بندی^۳ اهداف به‌وسیله سیگنال‌های ساطع شده از آن‌ها است. وقتی یک شناور دریایی همانند یک کشتی سطحی در میان آب سفر می‌کند، انواع گوناگونی از سیگنال‌ها را هم در محیط زیر سطح آب و هم در محیط بالای سطح آب ساطع می‌کند. یکی از این سیگنال‌ها، سیگنال آکوستیکی است. اکثر سامانه‌های نظارت زیر آب، از سیگنال‌های آکوستیکی برای اهداف مختلف، از جمله تشخیص و طبقه‌بندی شناورهای دریایی استفاده می‌کنند [۱]؛ از دلایل این امر می‌توان به این موارد اشاره کرد که سیگنال‌های رادیویی و نوری در محیط زیر آب به‌شدت تضعیف می‌شوند؛ امواج الکترومغناطیسی در محیط زیر آب نفوذ نمی‌کنند؛ هزینه تصویربرداری ماهواره‌ای بسیار زیاد است و این تصاویر نمی‌توانند حقایق زیر آب را به‌خوبی آشکار کنند؛ درحالی‌که آب محیطی به‌طور ویژه مناسب برای انتشار صدا است و امواج آکوستیکی می‌توانند تا کیلومترهای دور دست منتشر شوند [۴-۱]. درواقع، روش‌های آکوستیکی، روش‌های بسیار قدرتمندی هستند که به کمک آن‌ها می‌توان مجموعه داده‌های بزرگ و ضبط‌های زمان واقعی^۴ طولانی را آسان‌تر از سایر روش‌ها به دست آورد و پویای جمع‌آوری داده با استفاده از این روش مقرون‌به‌صرفه‌تر است [۲-۳].

زمانی که یک شناور دریایی همانند یک کشتی سطحی مانور داده می‌شود موتور، پروانه‌هایش و سایر عوامل مربوط به آن باعث ایجاد سیگنال آکوستیکی می‌شود. اگرچه این سیگنال آکوستیکی با انواع و اقسام نویز موجود در محیط آب همانند نویز ناشی از کشتی‌های دیگر، حیات موجودات دریایی، ارتعاشات کف دریا و... ترکیب می‌شود و سیگنال آکوستیکی که توسط هیدروفون‌های سونار دریافت می‌شود بسیار پیچیده است؛ اما سیگنال آکوستیکی ساطع شده از هر شناور دارای ویژگی‌های منحصر به فرد یا به عبارتی، امضای آکوستیکی مخصوص به خود است که امکان شناسایی فردی آن فراهم می‌کند [۵].

پیشینه طبقه‌بندی خودکار کشتی‌ها، زیردریایی‌ها و سایر شناورهای دریایی بر اساس امضاهای آکوستیکی آن‌ها، با توجه به این که این کار عمدتاً برای وزارت دفاع می‌شود، همواره عمومی نیست و مقالات عمومی کمی در این حوزه یافت می‌شود. در این بخش، مروری بر مقالاتی عمومی که این امر را مورد هدف قرار

¹ Underwater Surveillance

² Recognize

³ Classification

⁴ Real Time

به عنوان مجموعه داده آزمایش به کار می برد. این مرجع، برای استخراج ویژگی از هر دو مجموعه داده فوق الذکر از تبدیل هیلبرت هوانگ که برای پردازش سیگنال‌های غیرخطی و غیر ایستار بسیار مناسب است، استفاده می کند و انرژی نرمال شده مربوط به دامنه لحظه‌ای استخراج شده از ۴ تابع حالت ذاتی اول را به عنوان بردار ویژگی اتخاذ می کند. این مرجع همچنین، در مرحله طبقه بندی از الگوریتم ماشین بردار پشتیبان (SVM) استفاده می کند و نتایج طبقه بندی را با هنگامی که تبدیل موجک با تابع موجک مادر دابچیز^۵ با ۳ سطح تجزیه برای استخراج ویژگی به کار رود، مقایسه می کند و با توجه به دقت طبقه بندی بالاتر تبدیل هیلبرت هوانگ، برتری این تبدیل را بر تبدیل موجک نشان می دهد [۸].

۲. روش تحقیق

ابتدا از میان مجموعه داده‌های آکوستیکی عمومی و غیرفعال موجود در اینترنت، یک مجموعه داده که از سیگنال‌های آکوستیکی متعلق به شناورهای دریایی مختلف تشکیل می شود، به عنوان مجموعه داده‌ای که تشخیص خودکار کلاس متناظر به هر یک از سیگنال‌های موجود در آن، یا به عبارتی دیگر استخراج امضای آکوستیکی متناظر به هر یک از شناورهای آن، هدف اصلی این پژوهش است، اتخاذ می شود. در ادامه، روش‌هایی به منظور پیش پردازش و آماده سازی این مجموعه داده به کار می رود و چندین مجموعه داده جدید پیش پردازش شده ایجاد می شود. سپس، روش‌هایی برای استخراج ویژگی‌های کارآمد از مجموعه داده‌های پیش پردازش شده اعمال می شود. در نهایت، از چندین الگوریتم طبقه بندی برای طبقه بندی بردارهای ویژگی حاصل از فرایند استخراج ویژگی استفاده می شود.

داده پایه و عنصر اصلی همه الگوریتم‌های طبقه بندی مبتنی بر یادگیری ماشین و یادگیری عمیق است؛ به طوری که کارایی الگوریتم‌ها هم به ماهیت و مشخصه‌های مجموعه داده به کار رفته و هم به عملکرد الگوریتم‌های طبقه بندی بستگی دارد. مجموعه داده‌ها را از نظر پیچیدگی به دو دسته، یعنی مجموعه داده‌های ساده و مجموعه داده‌های پیچیده طبقه بندی می کنند که مجموعه داده متناظر با سیگنال‌های آکوستیکی شناورهای دریایی مختلف با توجه به پیچیدگی‌های محیط انتشار و انواع و اقسام نویز موجود در محیط زیر آب، به عنوان مجموعه داده پیچیده محسوب می شود؛ بنابراین برای دستیابی به یک طبقه بندی کننده خودکار موفق می بایست، یک مجموعه داده با ابعاد و کیفیت

است، همچنین از روش‌های فرایبتکاری برای آموزش این شبکه عصبی می توان استفاده کرد. ماهیت تصادفی الگوریتم‌های فرایبتکاری به آن‌ها کمک می کند تا از گیرکردن در کمینه‌های محلی جلوگیری کنند. نتایج طبقه بندی از نظر سرعت همگرایی و خطای طبقه بندی با سایر الگوریتم‌های فرایبتکاری همانند بهینه ساز گرگ خاکستری^۱ (GWO)، بهینه سازی مبتنی بر جغرافیای زیستی^۲ (BBO)، الگوریتم جستجوی داخلی^۳ (ISO) و روش‌های گروهی پردازش داده^۴ (GMDH) نشان می دهد که الگوریتم ازدحام ذرات سالپ در طبقه بندی مجموعه داده به کار رفته، سرعت همگرایی بیشتری دارد و به نتایج دقیق تری می انجامد [۶].

محدودیت‌های تبدیل فوریه کوتاه مدت، از جمله: این تبدیل، تنها قادر به تشخیص سیگنال‌هایی است که طولی متناسب با طول درجه زمانی به کار رفته در تبدیل دارند. همچنین، از آنجایی که این تبدیل از بخش بندی‌های زمانی با طول یکسان استفاده می کند، دقت طبقه بندی بسیار وابسته به طول بخش‌های زمانی است. علاوه بر این، تبدیل فوریه زمان کوتاه برای استخراج ویژگی‌های خاص مناسب نیست و بدون انجام تحلیل خاصی برای همه ویژگی‌های استخراج شده ارزش یکسانی قائل می شود [۷].

در یک پژوهش، به منظور طبقه بندی شناورها با استفاده از نویز انتشاری حاصل از آن‌ها از دو مجموعه داده به صورت جداگانه استفاده شد. بدین صورت که این مرجع اولین مجموعه داده را که از ۱۵ ضبط صوتی مربوط به ۳ کلاس زیر دریایی مختلف، یعنی کلاس A، B و C تشکیل می شود را با چندین بار گوش کردن به اصوات درهم موجود در یک پایگاه داده عمومی اینترنتی متعلق به نیرو دریایی آمریکا و جدا کردن بخش‌های مفید ایجاد می کند. در این مجموعه داده از هر کلاس ۵ ضبط صوتی وجود دارد که این مرجع، هریک ضبط‌ها را با فرکانس ۸۰۰۰ هرتز نمونه برداری می کند و از میان ضبط‌های متعلق به هر کلاس، یک ضبط صوتی را به عنوان داده آموزش و ۴ ضبط صوتی دیگر را به عنوان مجموعه داده آزمایش به کار می برد. این مرجع همچنین، از مجموعه داده دیگری استفاده می کند که از داده‌های واقعی مربوط به آب‌های کم عمق خلیج فارس استخراج شده است. این مجموعه داده از ۶۰ ضبط صوتی ۶ ثانیه‌ای با فرکانس نمونه برداری ۲۴۰ کیلوهرتز مربوط به ۲ کلاس شناور، یعنی شناور سبک و شناور سنگین تشکیل می شود. این مجموعه داده، یک مجموعه داده متوازن است که این مرجع از میان ۳۰ ضبط صوتی مربوط به هر کلاس، ۲۵ ضبط را به عنوان مجموعه داده آموزش و ۵ ضبط باقی مانده را

¹ Gray Wolf Optimizer

² Biogeography-based Optimizer

³ Interior Search Algorithm

⁴ Group Method of Data Handling

بخش صوتی کوتاه‌مدت از هر کلاس) در پایگاه‌داده اشاره‌شده است.

در جدول (۱)، بخش‌های جداشده از هر یک از پنج ضیطی که در ایجاد مجموعه داده به‌کاررفته در این پژوهش مشارکت دارند، به همراه کلاس متناظر به آن‌ها، مشخصه‌های آن‌ها، همانند سرعت حرکت زیردریایی مربوطه برحسب گره^۳، عمق قرارگیری آن برحسب پا^۴، فاصله آن از بویه سوناری رادیویی مصرفی برحسب یارد^۵ و مدت‌زمان هر یک از بخش‌های جدا شده ارائه‌شده است.

تمام این بخش‌های جداشده، همانند ضیط‌های اصلی موجود در این پایگاه‌داده در قالب mp3 و دارای یک کانال صوتی (مونو)^۶ و نرخ نمونه‌برداری ۲۲۰۵۰ هرتز هستند. قالب mp3، یک قالب فشرده‌سازی با اتلاف است که لایه‌هایی از ضیط‌صوتی که خارج از محدوده شنیداری انسان هستند را حذف می‌کند و لایه‌هایی که گوش انسان خوب می‌شنود را حفظ می‌کند. همچنین، با توجه به این که هر یک از بخش‌های جداشده فقط صدای مربوط به یک کلاس زیردریایی را در برمی‌گیرد، مسئله طبقه‌بندی موردتوجه این پایان‌نامه، یک مسئله طبقه‌بندی چند برچسبی نیست.

به‌منظور پردازش و طبقه‌بندی مجموعه داده جدول (۱)، زبان برنامه‌نویسی پایتون و ابزار پیاده‌سازی گوگل کولب انتخاب می‌شود.

۲-۲. پیش‌پردازش و آماده‌سازی مجموعه داده

مجموعه‌داده به‌کاررفته در این پژوهش که از بخش‌های جداشده از ضیط‌های صوتی موجود در پایگاه‌داده [۹] ساخته می‌شود؛ اگر به همان صورت که در جدول (۱) اشاره می‌شود، مورد استفاده قرار بگیرد، دارای دو مشکل اساسی است. اولاً، این مجموعه داده بسیار کوچک است و تنها از ۱۷ نمونه داده تشکیل می‌شود، همچنین باید برخی از نمونه‌ها به‌عنوان مجموعه داده آزمایش کنار گذاشته شوند. ثانیاً، هنگامی که هر یک از این نمونه‌ها بارگذاری می‌شوند، از قالب فشرده mp3 خارج می‌شوند و به یک دنباله یا سری زمانی از اعداد پشت‌سرهم تبدیل می‌شوند؛ با توجه به این که هر یک از این نمونه‌ها دارای نرخ نمونه‌برداری و تعداد کانال یکسان و درعین حال دارای مدت‌زمان متفاوتی است، ابعاد دنباله‌های اعداد ایجادشده متفاوت است؛ این در حالی است که اغلب کتابخانه‌های مربوط به یادگیری ماشین و یادگیری عمیق در پایتون مبتنی بر

مناسب و همچنین با تنوع کافی از کلاس‌ها اتخاذ گردد و پیش‌پردازشی مناسب بر آن اعمال گردد.

در این پژوهش، از پایگاه‌داده اینترنتی به‌کاررفته در مرجع [۸]، برای به‌دست‌آوردن امضاها و آکوستیکی مربوط به همان سه کلاس زیردریایی که آن مرجع مورد هدف قرار می‌دهد استفاده می‌شود. با مراجعه به تارنمای [۹]، بخش ضیط‌های آموزشی بویه سوناری رادیویی مصرفی^۱ 15P3، با گوش کردن مکرر به پنج ضیط‌صوتی اول که صدای مربوط به آن سه کلاس زیردریایی را در برمی‌گیرد، حذف بخش‌های حاوی صدای انسانی که محتوای هر یک ضیط را توضیح می‌دهد، جداسازی بخش‌های حاوی صدای زیردریایی‌ها، با برچسب‌گذاری و یادداشت‌کردن مشخصه‌های آن بخش‌ها، یک مجموعه داده از اصوات متناظر به آن سه کلاس زیردریایی ایجاد می‌شود؛ به‌طوری که در انتهای امر نتایج حاصل از این پژوهش با نتایج به‌دست‌آمده توسط مرجع [۸] قابل‌مقایسه است.

۲-۱. مجموعه داده

داده پایه و عنصر اصلی همه الگوریتم‌های طبقه‌بندی مبتنی بر یادگیری ماشین و یادگیری عمیق است؛ به‌طوری که کارایی الگوریتم‌ها هم به ماهیت و مشخصه‌های مجموعه داده به‌کاررفته و هم به عملکرد الگوریتم‌های طبقه‌بندی بستگی دارد. مجموعه داده‌ها را از نظر پیچیدگی به دو دسته، یعنی مجموعه داده‌های ساده و مجموعه داده‌های پیچیده طبقه‌بندی می‌کنند که مجموعه‌داده متناظر با سیگنال‌های آکوستیکی شناورهای دریایی مختلف، به‌عنوان مجموعه‌داده پیچیده محسوب می‌شود.

لازم به ذکر است که از آنجایی که کاویتاسیون تیغه‌های پروانه، منبع اصلی نویز ناشی هر وسیله نقلیه دریایی در حال حرکت در محیط آب است و حدود ۸۰ الی ۸۵ درصد نویز انتشاری ناشی از وسیله نقلیه دریایی را در محیط آب به خود اختصاص می‌دهد [۴]، مجموعه داده مورد استفاده در این پژوهش به کمک بخش‌هایی از ضیط‌های موجود که حاوی صدای کاویتاسیون یا ضربات پروانه^۲ باشند، ایجاد می‌شود. بدین ترتیب، در ایجاد مجموعه داده مربوط به این پژوهش، با توجه توضیحات شخصی که محتوای ضیط‌ها را توضیح می‌دهد، از بخش‌های حاوی صدای زیردریایی که به علت سرعت کم یا عمق زیاد تنها حاوی نویز ماشین‌آلات هستند، استفاده نمی‌شود و آن بخش‌ها کنار گذاشته می‌شوند. علاوه بر این، دلیل دیگر کنارگذاشتن بخش‌های حاوی صدای نویز ماشین‌آلات، تعداد بسیار کم آن‌ها (تقریباً یک الی دو

³ Knot

⁴ Foot

⁵ Yard

⁶ Mono

¹ Expendable Radio Sonobuoy Training Records

² Propeller Beats

در این تحقیق برای غلبه بر این دو مشکل، در مرحله پیش پردازش و آماده سازی داده ها با استفاده از روش های داده افزایی، تعداد نمونه های موجود در مجموعه داده گسترش داده می شود.

۲-۳. داده افزایی

داده افزایی، یک روش متداول برای افزایش مصنوعی تعداد داده، به ویژه زمانی است که داده به اندازه کافی موجود نیست. در داده افزایی، با اصلاح داده های موجود به میزان کوچک، به گونه ای که معنا و مفهوم داده ها عوض نشود، داده های جدیدی ایجاد می شود که برچسب مربوط به آن ها با برچسب متناظر با داده های اولیه یکسان است. این عمل، با توجه به این که از صرف هزینه های مربوط به جمع آوری و برچسب گذاری داده اجتناب می کند و مقاومت مدل را نسبت به برخی تغییرات بهبود می دهد، دارای مزیت است [۱۰].

در این پژوهش برای داده افزایی ۱۷ ضبط صوتی داده شرح داده شده در جدول (۱)، سناریوهای زیر مورد بررسی قرار می گیرد:

- در سناریوی اول، هر یک از ضبط ها از طریق بخش بندی زمانی به بخش هایی به طول مساوی، یعنی بخش های یک ثانیه ای با همپوشانی ۰,۲ ثانیه تفکیک می شود و با ذخیره سازی هر یک از آن ها به صورت یک پرونده صوتی جداگانه، گسترش داده می شوند. این داده های صوتی گسترش داده شده، با برچسب های مشابه با ضبط های اولیه برچسب گذاری می شوند. لازم به ذکر است که در این سناریو پیش از عمل بخش بندی، ۲ ثانیه آخر مربوط به هر یک از ضبط ها، به عنوان مجموعه داده از پیش دیده نشده آزمایش، کنار گذاشته می شود و به ۲ بازه بدون همپوشانی تفکیک می شود؛ سپس عمل بخش بندی با همپوشانی بر قسمت باقیمانده ضبط ها انجام می شود و مجموعه داده آموزش ایجاد می شود. این رویکرد، مصداقی از داده افزایی به شیوه انتقال زمانی است و در مرجع [۱۱] نیز به کار رفته است.
- این سناریو، مشابه سناریوی اول است؛ با این تفاوت که برای ایجاد مجموعه داده آموزش، هر یک از ضبط ها به بخش هایی به طول مساوی، یعنی ۲ ثانیه ای و با همپوشانی ۰,۲ ثانیه تفکیک می شود و پیش از عمل تفکیک، ۲ ثانیه آخر هر ضبط، به عنوان یک عضو از مجموعه داده از پیش دیده نشده آزمایش کنار گذاشته می شود.
- در سناریوی سوم، از هر کلاس یک ضبط به صورت تصادفی (یعنی، در مجموع سه ضبط) انتخاب می شود و به عنوان ضبط از پیش دیده نشده آزمایش کنار گذاشته می شود. از طریق بخش بندی زمانی ۱ ثانیه ای و بدون همپوشانی ضبط های کنار گذاشته شده، مجموعه داده از پیش دیده

محاسبات ماتریسی هستند و به یک بازنمایی برداری یا ماتریسی با ابعاد یکسان از داده های ورودی نیاز دارند.

جدول ۱. بخش های جدا شده از هر یک از ضبط ها و مشخصه آنها

شماره ضبط	کلاس بخش تفکیک شده	مشخصه ها	مدت زمان (s)
۱	B	حاوی صدای ضربات پروانه در حال کاویتاسیون (بخش ۱)	۱۱
	B	حاوی صدای ضربات پروانه در حال کاویتاسیون (بخش ۲)	۲۱
	B	حاوی صدای ضربات پروانه در حال کاویتاسیون و نویز ماشین آلات	۳۲
۲	A	حاوی صدای کاویتاسیون پروانه در سرعت ۷ گره و در عمق پریسکوپ (بخش ۱)	۹
	A	حاوی صدای کاویتاسیون پروانه در سرعت ۷ گره و در عمق پریسکوپ (بخش ۲)	۱۱
	A	حاوی صدای کاویتاسیون پروانه در سرعت ۶ گره و در عمق پریسکوپ (بخش ۱)	۱۱
	A	حاوی صدای کاویتاسیون پروانه در سرعت ۶ گره و در عمق پریسکوپ (بخش ۲)	۱۴
	A	حاوی صدای کاویتاسیون پروانه در سرعت ۸ گره و در عمق ۲۵۰ پا (بخش ۱)	۷
	A	حاوی صدای کاویتاسیون پروانه در سرعت ۸ گره و در عمق ۲۵۰ پا (بخش ۲)	۱۴
۳	C	حاوی صدای کاویتاسیون پروانه در سرعت ۳ گره، عمق پریسکوپ و فاصله ۵۰۰ یارد (بخش ۱)	۱۲
	C	حاوی صدای کاویتاسیون پروانه در سرعت ۳ گره، عمق پریسکوپ و فاصله ۵۰۰ یارد (بخش ۲)	۱۶
	C	حاوی صدای کاویتاسیون پروانه در سرعت ۳ گره، عمق ۱۰۰ پا	۲۵
۴	C	حاوی صدای ضعیف و نامنظم کاویتاسیون پروانه در سرعت ۲,۵ گره و در عمق پریسکوپ (بخش ۱)	۱۰
	C	حاوی صدای ضعیف و نامنظم کاویتاسیون پروانه در سرعت ۲,۵ گره و در عمق پریسکوپ (بخش ۲)	۵
۵	C	حاوی صدای کاویتاسیون پروانه در سرعت ۳ گره و در عمق نامعلوم	۲۰
	B	حاوی صدای کاویتاسیون پروانه در فاصله ۲۰۰ یارد، در سرعت ۶ گره و در عمق پریسکوپ	۲۴
	B	حاوی صدای کاویتاسیون پروانه در فاصله ۱۰۰۰ یارد، در سرعت ۶ گره و در عمق پریسکوپ	۲۵

به اصوات زیردریایی‌ها و برجسب‌های متناظر به آن‌ها، به همراه معماری شبکه و وزن‌هایی که کریستوفر چین در مرجع [۱۲] برای تشخیص تماس‌های نهنگ راست اقیانوس اطلس شمالی به کار می‌برد، برای تشخیص اصوات زیردریایی‌ها استفاده می‌شود.

بعد از داده‌افزایی، هر یک از ضبط‌های گسترش داده شده با توجه به این که اطلاعات طیفی مفید مربوط به مشخصه‌های کلاس آن‌ها در فرکانس زیر 4kHz واقع است، طبق قضیه نمونه‌برداری نایکوئیست با فرکانس نمونه‌برداری دوبرابر این فرکانس، یعنی 8kHz بارگذاری می‌شود و بدین ترتیب هر یک از ضبط‌ها به آرایه‌ای از اعداد پشت‌سرهم تبدیل می‌شوند.

۴-۲. استخراج ویژگی

در سناریوهای اول تا ششم، به‌منظور استخراج ویژگی از ضرایب کپسترال فرکانسی مل (MFCC) استفاده می‌شود و بدین ترتیب، یک نمایش پارامتری از ضبط‌ها به دست می‌آید. ضرایب کپسترال فرکانسی مل (MFCC)، یک روش تخصصی برای استخراج ویژگی از صداهای پر از مدولاسیون، همانند گفتار است [۱۱]. از آنجایی که اجزای مکانیکی وسایل نقلیه دریایی، معمولاً بزرگ هستند و این اجزا در اغلب مواقع، با سرعت کم می‌چرخند، بنابراین مشخصه‌های اساسی که وسایل نقلیه دریایی را توصیف می‌کنند تمایل دارند در محدوده فرکانس پایین باشند و همانند گفتار نیازمند وضوح بیشتر در محدوده فرکانس پایین و وضوح کمتر در فرکانس‌های بالا هستند [۱]. بدین ترتیب، اگر تعداد کل K بخش زمانی از بخش‌بندی سیگنال حاصل شود و $X[k]$ بیانگر طیف‌نگاشت یافته به مقیاس مل مربوط به k امین قاب پنجره‌گذاری شده باشد و $\log(X[k])$ بیانگر لگاریتم دامنه طیف مل باشد، با استفاده از تبدیل کسینوسی گسسته، ضرایب کپسترال فرکانسی مل به‌صورت زیر محاسبه می‌شوند:

$$c[n] = \sum_{k=1}^K \log_{10}(X[k]) \cos\left(\frac{n(k-0.5)\pi}{K}\right) \quad n = 1, 2, \dots, K$$

در رابطه فوق، $c[n]$ بیانگر n امین ضرایب کپسترال فرکانسی مل است.

در این پژوهش، در سناریوهای اول تا ششم، تعداد ۲۰ ضرایب MFCC از هر یک از آرایه‌های عددی حاصل از بارگذاری ضبط‌های یک یا دو تانیه‌ای گسترش داده شده استخراج می‌شود و میانگین هر ضرایب MFCC در طول بازه یک یا دو تانیه‌ای به‌عنوان یک ویژگی مطرح می‌شود. علاوه بر این، تمام این ۲۰ ویژگی استخراج شده به‌منظور دارا بودن میانگین صفر و واریانس یک استاندارد می‌شوند. بدین ترتیب، در هر یک از سناریوهای اول تا ششم در مرحله استخراج ویژگی، یک پرونده با قالب csv که نام

نشده آزمایش ایجاد می‌شود. همچنین، بقیه ضبط‌ها از طریق بخش‌بندی زمانی، به بازه‌های به طول ۱ ثانیه و با همپوشانی ۰.۲ ثانیه تفکیک می‌شوند و مجموعه داده آموزش را ایجاد می‌کنند.

۴. سناریوی چهارم، مشابه سناریوی سوم است؛ با این تفاوت که در این سناریو ضبط‌های کنار گذاشته شده آزمایش از طریق بخش‌بندی زمانی، به بازه‌های ۲ ثانیه‌ای و با همپوشانی ۱ ثانیه تفکیک می‌شوند و مجموعه داده آزمایش از پیش دیده نشده را ایجاد می‌کنند. همچنین برای ایجاد مجموعه داده آموزش، ضبط‌های باقیمانده به بخش‌های ۲ ثانیه‌ای و با همپوشانی ۰.۲ ثانیه تفکیک می‌شوند.

۵. سناریوی پنجم، مشابه سناریوی اول است؛ با این تفاوت که به هر عضو از مجموعه داده آزمایش، مقداری نویز تصادفی با توزیع گوسی افزوده می‌شود. لازم به ذکر است که نویز گوسی پیش از افزوده شدن در ضریب ۰.۲ ضرب می‌شود و بعد با سیگنال اصلی جمع می‌شود تا اثر نویز از اثر سیگنال اولیه کمتر باشد.

۶. سناریوی ششم، مشابه سناریو دوم است؛ به هر عضو از مجموعه داده آزمایش، مقداری نویز تصادفی با توزیع گوسی افزوده می‌شود. لازم به ذکر است که نویز گوسی پیش از افزوده شدن در ضریب ۰.۱۵ ضرب می‌شود و بعد با سیگنال اصلی جمع می‌شود تا اثر نویز از اثر سیگنال اولیه کمتر باشد. ۷. در سناریوی هفتم، اسپکتروگرام‌های مربوط به مجموعه داده گسترش داده شده حاصل از سناریوی دوم استخراج می‌شود و هر یک از اسپکتروگرام‌ها به‌صورت یک تصویر با قالب png ذخیره می‌شود.

۸. در سناریوی هشتم، با استفاده از فیلتر میانگین متحرک^۱ و مجموعه داده گسترش داده شده حاصل از سناریوی دوم اسپکتروگرام‌هایی ایجاد می‌شوند که تضاد هر یک از آن‌ها، یک‌بار در راستای افقی و بار دیگر در راستای عمودی افزایش داده می‌شود و به‌جای ذخیره‌سازی اسپکتروگرام‌ها به‌صورت تصاویر جداگانه، مجموعه اسپکتروگرام‌ها و مجموعه برجسب‌ها، به‌صورت دو آرایه جداگانه ذخیره می‌شوند. روش افزایش تضاد به‌کاررفته در این سناریو از روشی که کریستوفر چین^۲ در مرجع [۱۲] برای افزایش تضاد اسپکتروگرام‌های مربوط به مجموعه داده چالش تشخیص تماس‌های نهنگ راست اقیانوس اطلس شمالی استفاده می‌کند، الهام گرفته شده است.

۹. در سناریوی نهم، از آرایه‌هایی که در سناریوی هشتم ذخیره می‌شوند، یعنی اسپکتروگرام‌های افزایش تضاد یافته متناظر

¹ Moving Average

² Christopher Chin

طبقه‌بندی استفاده می‌شود. از این پس، منظور از مجموعه داده آموزش همان مجموعه داده آموزش جدید است. در نهایت، پس از اتمام فرایند آموزش، در صورت دستیابی هم‌زمان به نتایج مطلوب هم بر مجموعه داده آموزش و هم بر مجموعه داده اعتبارسنجی، عملکرد الگوریتم‌های طبقه‌بندی بر مجموعه داده آزمایش ارزیابی می‌شود. در ادامه، ساختار طبقه‌بندی کننده پیشنهادی این پژوهش ارائه می‌شود:

به منظور طبقه‌بندی بردارهای ویژگی حاصل از سناریوهای اول تا ششم از چندین الگوریتم طبقه‌بندی با ابر پارامترهای مختلف استفاده می‌شود که در ادامه، هر الگوریتم و برخی از ابر پارامترهای مربوط به آن ارائه می‌شود:

شبکه عصبی مصنوعی (ANN) با لایه‌های حذف: از یک شبکه عصبی با دو لایه پنهان، به ترتیب دارای ۱۰۰ و ۴۰۰ نورون که همگی دارای تابع فعال‌سازی واحد یک‌سوساز خطی (ReLU) هستند، به همراه دو لایه حذف با نرخ حذف، به ترتیب برابر با ۰٫۲ و ۰٫۵، استفاده می‌شود. لایه حذف، روشی آسان برای جلوگیری از بیش‌برازش است. در این روش با توجه به نرخ حذف که خود یک ابر پارامتر است، به صورت تصادفی کسری از نورون‌های یک لایه غیرفعال می‌شود و آن دسته از نورون‌ها برای آموزش نادیده گرفته می‌شود. اگرچه این عمل، تا حدودی دقت را بر مجموعه آموزش کاهش می‌دهد؛ ولی از آنجایی که با بیش‌برازش مقابله می‌کند دارای مزیت است. به کارگیری لایه‌های حذف، این اطمینان را می‌دهد که شبکه برای طبقه‌بندی به نورون‌های خاص وابسته نیست. از لایه‌های حذف تنها در حین آموزش استفاده می‌شود و در حین ارزیابی استفاده نمی‌شود؛ چراکه طبقه‌بندی کننده می‌بایست در هنگام ارزیابی به‌طور بهینه کار کند؛ لذا نرخ حذف در حین ارزیابی به صفر تنظیم می‌شود [۱۱]. لایه خروجی این شبکه، دارای ۳ نورون (برابر با تعداد کلاس‌های خروجی یا تعداد شناورها) و تابع فعال‌سازی بیشینه هموار است.

تابع خطای آنتروپی متقاطع طبقه‌ای پراکنده^۳ به‌عنوان تابع خطای تابع هزینه در نظر گرفته می‌شود. از تابع بهینه‌ساز گرادیان کاهشی تصادفی (SGD) با نرخ یادگیری ۰٫۰۳ استفاده می‌شود و تعداد دفعات تکرار فرایند آموزش برابر با حداکثر ۳۵ بار در نظر گرفته می‌شود. علاوه بر این، به کمک از توابع فراخوان برگشتی^۴، میزان خطا بر مجموعه داده اعتبارسنجی را در طول تکرارهای مختلف آموزش نظارت می‌شود و در صورت عدم بهبود خطای طبقه‌بندی بر مجموعه داده اعتبارسنجی در طول ۵ تکرار، آموزش الگوریتم متوقف می‌شود و اگر این خطا در طول ۳ تکرار تغییری

پرونده‌های صوتی گسترده شده، مقادیر ۲۰ ضریب MFCC استاندارد شده مربوط به هر پرونده و کلاس متناظر با هر پرونده را در برمی‌گیرد ایجاد می‌شود.

برای استخراج ویژگی از اسپکتروگرام‌های حاصل از سناریو هفتم و هشتم، با به کارگیری شبکه‌های عصبی پیچشی (CNN) که از بهترین الگوریتم‌ها برای کار با تصاویر هستند، دیگر نیازی به ویژگی‌های مهندسی شده نیست و این معماری‌های مبتنی بر یادگیری عمیق به صورت خودکار ویژگی‌های موردنیاز را استخراج می‌کنند.

۲-۵. طبقه‌بندی کننده‌های پیشنهادی

هدف از ایجاد طبقه‌بندی کننده، ساخت یک مدل ریاضی است که بتواند از روی بردارهای ویژگی و برچسب‌های موجود در مجموعه داده آموزش، دانشی را استخراج کند و به کمک آن دانش استخراج شده قادر به پیش‌بینی مجموعه داده آزمایش باشد. در همه الگوریتم‌های طبقه‌بندی، پیش از آموزش هر الگوریتم لازم است مشخصه‌هایی از قبل تعیین شوند. این مشخصه‌ها که عملکرد الگوریتم‌ها به آن‌ها وابسته است، ابر پارامترهای^۱ الگوریتم طبقه‌بندی نامیده می‌شوند و در طول آموزش کاملاً ثابت هستند. تعیین ابر پارامترهای مختلف مربوط به یک الگوریتم طبقه‌بندی، به صورت تنها و جداگانه میسر نیست و تغییر هر یک ابر پارامتر بر مقدار بهینه سایر ابر پارامترها تأثیر می‌گذارد.

علاوه بر مسئله تعیین ابر پارامترهای یک الگوریتم طبقه‌بندی، مسئله دیگری که بر سر راه یک طبقه‌بندی کننده قرار می‌گیرد و همواره عملکرد آن را مورد تهدید قرار می‌دهد، مسئله بیش‌برازش است. در این حالت طبقه‌بندی کننده به جای استخراج دانش از مجموعه داده آموزش، آن‌ها را حفظ کرده و علی‌رغم عملکرد مطلوب بر مجموعه داده آموزش، به عملکرد بدی بر مجموعه داده آزمایش دست می‌یابد. یک روش مؤثر برای جلوگیری از بیش‌برازش و آموزش مؤثر الگوریتم‌های طبقه‌بندی این است که پیش از آموزش الگوریتم‌های طبقه‌بندی، بخشی از مجموعه داده آموزش و برچسب‌های متناظر به آن‌ها، به‌عنوان مجموعه داده اعتبارسنجی^۲ کنار گذاشته شود و در آموزش الگوریتم‌ها دخالت داده نشود و در هر تکرار از فرایند آموزش، با ارزیابی نتایج بر مجموعه داده اعتبارسنجی از بیش‌برازش مدل جلوگیری شود. در هر یک از سناریوهای ذکر شده، به صورت تصادفی ۲۰ درصد از مجموعه داده آموزش، به‌عنوان مجموعه داده اعتبارسنجی کنار گذاشته می‌شود و ۸۰ درصد باقیمانده، به‌عنوان مجموعه داده آموزش جدید در نظر گرفته می‌شود و برای آموزش الگوریتم‌های

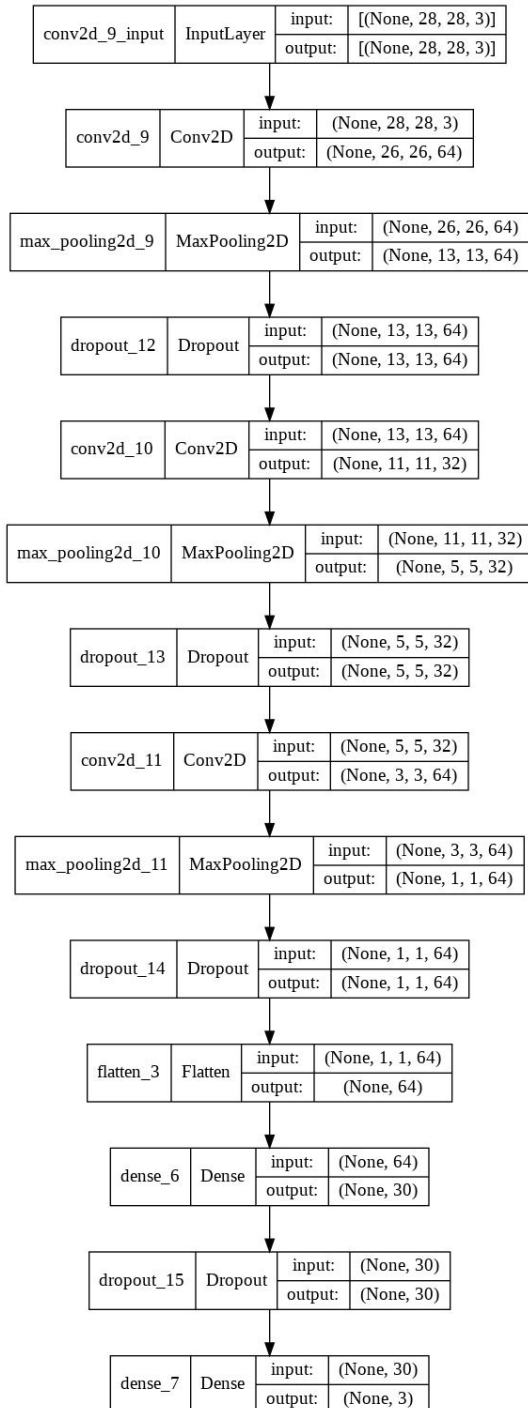
^۳ Sparse Categorical Cross Entropy

^۴ Call back

^۱ Hyper Parameter

^۲ Validation Dataset

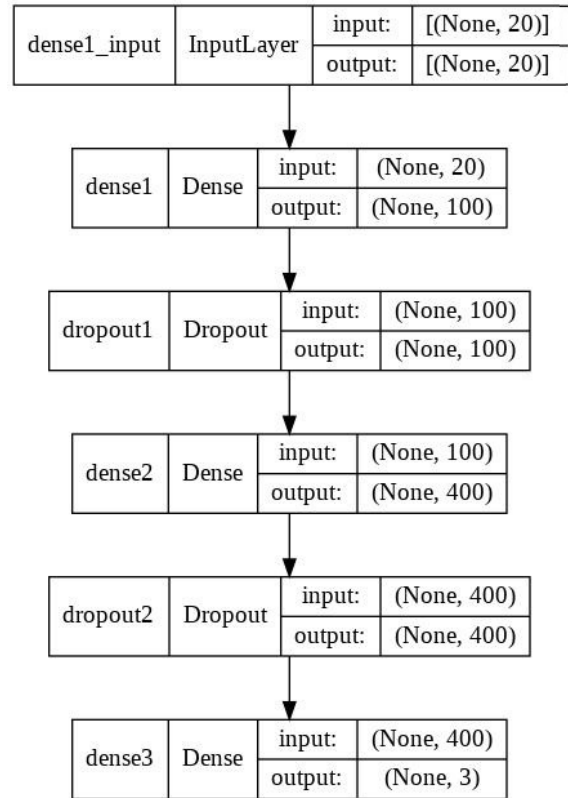
۴. ماشین بردار پشتیبان (SVM): از یک ماشین بردار پشتیبان با تابع هسته پایه شعاعی (RBF)، با مقدار گاما^۳ برابر با ۰/۰۲ و ضریب غرامت (C) برابر با ۵ استفاده می‌شود.



شکل ۲. نمودار جریان طبقه‌بندی کننده پیشنهادی در سناریوهای هفتم و هشتم

برای طبقه‌بندی مجموعه داده‌های حاصل از سناریوهای هفتم و هشتم، یک شبکه عصبی پیچشی با سه بلوک پیچشی استفاده

نکند، نرخ یادگیری در ضریب ۰.۵ ضرب می‌شود. لازم به ذکر است که توابع فراخوان ذکر شده و پارامترهای آن‌ها، نیز ابر پارامتر محسوب می‌شود.



شکل ۱. نمودار جریان اولین طبقه‌بندی کننده پیشنهادی

۱. K نزدیک‌ترین همسایه (KNN): از معیار فاصله اقلیدسی برای اندازه‌گیری فاصله استفاده می‌شود. در پیاده‌سازی‌های مختلف، مقادیر مختلفی برای K انتخاب می‌شود و نمودار دقت بر مجموعه داده‌های آموزش، اعتبارسنجی و آزمایش به ازای مقادیر مختلف K بین ۱ تا ۳۰ به تصویر کشیده می‌شود.

۲. درخت تصمیم (DT): از درخت تصمیم با معیار شاخص جینی استفاده می‌شود.

۳. شبکه عصبی مصنوعی (ANN) بدون لایه حذف: از یک شبکه عصبی بدون لایه حذف که دارای دو لایه پنهان با تعداد به ترتیب ۱۰۰ و ۸۰ نورون و یک لایه خروجی با ۳ نورون است، استفاده می‌شود که در آن تمامی نورون‌ها تابع فعال‌سازی واحد یک‌سوساز خطی (ReLU) دارند. تابع خطای آنتروپی متقاطع طبقه‌ای^۱ به‌عنوان تابع خطا در نظر گرفته می‌شود. از تابع بهینه‌ساز آدام^۲ با نرخ یادگیری ثابت ۰,۰۰۰۳ و حداکثر تعداد تکرار برابر با ۸۰ بار استفاده می‌شود.

^۱ Categorical Cross Entropy

^۲ Adam

^۳ Gamma

افزایش تضاد یافته در راستای افقی و عمودی مربوط به اصوات زیردریایی مربوط به سناریو نهم استفاده می‌شود. برای طبقه‌بندی مجموعه داده حاصل از سناریوی نهم، به جای ۳ لایه انتهایی شکل فوق، به ترتیب از یک لایه تمام متصل با تعداد ۵۰ نورون و تابع فعال‌سازی واحد یک‌سوساز خطی (ReLU)، یک لایه حذف با نرخ ۰/۵ و یک لایه خروجی تمام متصل با تعداد ۳ نورون و تابع فعال‌سازی بیشینه هموار استفاده می‌شود. علاوه بر این، وزن‌های مربوط به ۷ لایه اول مستقیماً انتقال داده می‌شود و در واقع این لایه‌ها غیرقابل آموزش هستند و وزن‌های لایه‌های باقی‌مانده توسط مجموعه داده حاصل از سناریوی نهم مجدداً تنظیم می‌شود.

۳. نتایج و بحث

در این بخش، ابعاد مجموعه داده‌های آموزش و آزمایش مربوط هر یک از سناریوهای مطرح شده، بخش پیش‌پردازش و آماده‌سازی مجموعه داده (مجموعاً نه سناریو) ارائه می‌شود و بعد از اعمال روش استخراج ویژگی و طبقه‌بندی مربوط به هر سناریو، نتایج مربوط به هر یک از پیاده‌سازی‌ها در قالب معیارهای ارزیابی مختلف همانند ماتریس پراکندگی، گزارش طبقه‌بندی و... ارائه می‌شود و در انتهای هر بخش، این نتایج به صورت مختصر تحلیل می‌شود.

۳-۱. پیاده‌سازی سناریوی اول

ابعاد مجموعه داده‌های آموزش و آزمایش حاصل از سناریوی اول که در آن، ۲ ثانیه آخر مربوط به هر یک از ضبط‌ها موجود، برای ایجاد مجموعه داده آزمایش کنار گذاشته می‌شوند و به ۲ بازه بدون همپوشانی تفکیک می‌شود و قسمت باقیمانده هر یک از ضبط‌ها برای ایجاد مجموعه داده آموزش، به بخش‌های یک‌ثانیه‌ای و با همپوشانی ۰/۲ ثانیه تفکیک می‌شود، در جدول (۲) ارائه می‌شود:

جدول ۲. ابعاد مجموعه داده‌های آموزش و آزمایش حاصل از سناریوی اول

ردیف	تعداد نمونه‌های آموزش	تعداد نمونه‌های آزمایش
کلاس A	۲۴۶	۱۲
کلاس B	۴۹۵	۱۰
کلاس C	۳۵۶	۱۲
جمع کل	۱۰۹۷	۳۴

برای مقایسه نتایج حاصل از این سناریو با نتایج به دست آمده توسط مرجع [۸]، در جدول (۳) به صورت خلاصه، درصد طبقه‌بندی صحیح هر کلاس (درصد یادآوری) با استفاده از

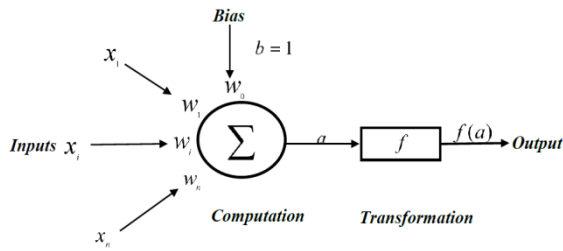
می‌شود که بلوک اول آن، از یک لایه پیچشی با ۶۴ فیلتر با ابعاد ۳*۳ و با توابع فعال‌سازی واحد یک‌سوساز خطی (ReLU)، یک لایه ادغام حداکثری با ابعاد ۲*۲ و یک لایه حذف با نرخ ۰/۲، بلوک دوم از یک لایه پیچشی با ۳۲ فیلتر با ابعاد ۳*۳ و با توابع فعال‌سازی واحد یک‌سوساز خطی (ReLU)، یک لایه ادغام حداکثری با ابعاد ۲*۲ و یک لایه حذف با نرخ ۰/۲، بلوک سوم از یک لایه پیچشی با ۶۴ فیلتر با ابعاد ۳*۳ و با توابع فعال‌سازی واحد یک‌سوساز خطی (ReLU)، یک لایه ادغام حداکثری با ابعاد ۲*۲ و یک لایه حذف با نرخ ۰/۲ تشکیل می‌شود. در این شبکه، در ادامه بلوک‌های پیچشی، یک لایه مسطح کننده برای تبدیل آرایه‌های چندبعدی حاصل از آخرین بلوک پیچشی به بردارهای یک‌بعدی، یک لایه تمام متصل با ۳۰ نورون و تابع فعال‌سازی واحد یک‌سوساز خطی (ReLU)، یک لایه حذف با نرخ ۰/۲ و در نهایت، یک لایه خروجی تمام متصل با تعداد ۳ نورون و تابع فعال‌سازی بیشینه هموار قرار می‌گیرد.

در آموزش این دو سناریو نیز از تابع بهینه‌ساز گرادیان کاهشی تصادفی (SGD) با نرخ یادگیری ۰/۰۳ استفاده می‌شود و تعداد دفعات تکرار فرایند آموزش برابر با ۳۵ بار در نظر گرفته می‌شود. علاوه بر این، به کمک از توابع فراخوان برگشتی^۱ در صورت عدم بهبود خطای طبقه‌بندی بر مجموعه داده اعتبارسنجی در طول ۵ تکرار، آموزش الگوریتم متوقف می‌شود و اگر این خطا در طول ۳ تکرار تغییری نکند، نرخ یادگیری در ضریب ۰/۵ ضرب می‌شود.

۲-۶. یادگیری انتقالی

در اغلب مواقع، آموزش الگوریتم‌های یادگیری عمیق به علت دارا بودن پارامترهای بسیار، به مجموعه داده‌های بزرگی احتیاج دارد. با این حال، دسترسی به داده‌های بزرگ همواره میسر نیست و روش‌هایی همانند داده‌افزایی نیز همواره پاسخگو نیست. در چنین مواقعی، می‌توان از روشی که یادگیری انتقالی نامیده می‌شود استفاده نمود. منظور از یادگیری انتقالی این است که از دانشی که در هنگام طراحی طبقه‌بندی کننده به کمک یک مجموعه داده آموزش با ابعاد بزرگ کسب می‌شود، در طبقه‌بندی یک مجموعه داده آموزش با ابعاد کوچک‌تر، ولی مرتبط به مجموعه داده اولیه استفاده نمود. به منظور طبقه‌بندی مجموعه داده حاصل از سناریوی نهم، از ساختار شبکه عصبی پیچشی و وزن‌هایی که کریستوفر چین در مرجع [۱۲]، برای حل چالش طبقه‌بندی تماس نهنگ‌های راست اقیانوس اطلس شمالی تارنمای کگل با بیش از ۴۰۰۰۰ داده آموزش به کار می‌گیرد و به کمک آن به دقت بالای ۹۶ درصد دست می‌یابد استفاده می‌شود و به کمک یادگیری انتقالی، از این دانش برای طبقه‌بندی اسپکتروگرام‌های

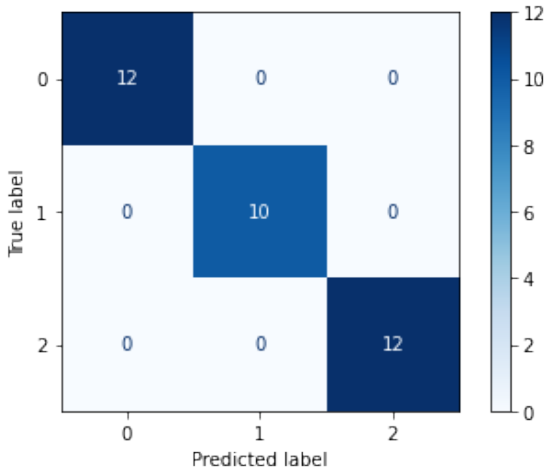
^۱ Call back



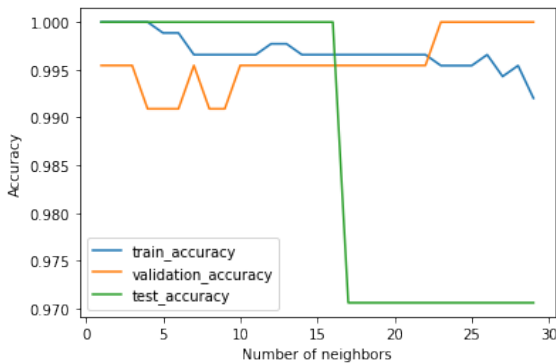
شکل ۴. ساختار نرون مورد استفاده در سناریوی اول

تابع فعال ساز به شکل زیر خروجی را به ورودی مرتبط می کند:

$$Y = f\left(\sum_{i=1}^n w_i X_i + b\right)$$



شکل ۵. ماتریس درهم آمیختگی مجموعه داده آزمایش سناریوی اول با استفاده از ANN با لایه های حذف



شکل ۶. نمودارهای دقت طبقه بندی داده های سناریوی اول با استفاده از KNN با مقادیر K بین ۱ تا ۳۰

جدول ۴. نتایج حاصل از آخرین تکرار فرایند آموزش سناریوی اول با الگوریتم ANN با لایه های حذف

تکرار	خطا بر مجموعه آموزش	دقت بر مجموعه آموزش	خطا بر مجموعه داده اعتبارسنجی	دقت بر مجموعه داده اعتبارسنجی	نرخ یادگیری
۲۷	۰/۰۰۷۳	۰/۹۹۸۹	۰/۰۰۸۸	۰/۹۹۵۵	۰/۰۳۰۰

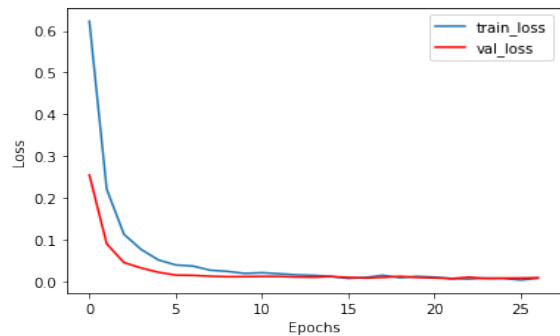
طبقه بندی کننده های فوق، به همراه میانگین وزن دار درصد طبقه بندی صحیح هر کلاس آورده شده است. لازم به ذکر است که ستون آخر جدول (۳) نتایج به دست آمده توسط مرجع [۸] است.

جدول ۳. درصد طبقه بندی صحیح هر کلاس در الگوریتم های پیشنهادی در سناریوی اول و روش مرجع [۸]

ردیف	MFCC -SVM	MFCC -ANN2	MFCC -DT	MFCC -KNN	MFCC -ANN	میانگین وزن دار
شناور A	۱۰۰	۹۲	۹۲	۱۰۰	۱۰۰	۶۲/۵
شناور B	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰
شناور C	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۶۲/۵
میانگین وزن دار	۱۰۰	۹۷	۹۷	۱۰۰	۱۰۰	۷۵

با توجه به نتایج ارائه شده در جدول فوق، تمامی الگوریتم های طبقه بندی پیشنهاد شده در این سناریو از این پژوهش قادر هستند که هر یک از کلاس ها را به خوبی تشخیص دهند و به میانگین وزن دار درصد طبقه بندی صحیح هر کلاس بالاتری نسبت به روش به کاررفته توسط مرجع [۸] دست یابند. از میان الگوریتم های پیشنهاد شده در این سناریو از این پژوهش، الگوریتم های شبکه عصبی مصنوعی (ANN) با لایه های حذف، K نزدیک ترین همسایه (KNN) و ماشین بردار پشتیبان به بهترین نتایج با میانگین وزن دار درصد طبقه بندی صحیح هر کلاس برابر با ۱۰۰ درصد می رسند.

شکل (۳) نمودار خطای داده های آموزش و اعتبارسنجی سناریوی اول با الگوریتم ANN با لایه های حذف را نشان می دهد. پرسپترون مورد استفاده در این سناریو در شکل (۴) توضیح داده شده است. شکل (۵) ماتریس درهم آمیختگی مجموعه داده آزمایش سناریوی اول را نشان می دهد. شکل (۶) نمودارهای دقت طبقه بندی داده های سناریوی اول با استفاده از KNN با مقادیر K بین ۱ تا ۳۰ را نمایش می دهد.



شکل ۳. نمودار خطای داده های آموزش و اعتبارسنجی سناریوی اول با الگوریتم ANN با لایه های حذف

با ارزیابی مدل آموزش دیده بر مجموعه داده آزمایش، نتایج زیر حاصل می گردد.

```
test_loss, test_acc = model.evaluate(X_test, Y_test)
```

```
2/2 [=====] - 0s 11ms/step - loss: 0.0204 - accuracy: 1.0000
```

شکل ۷. نتیجه ارزیابی مجموعه داده آزمایش سناریوی اول با طبقه بندی کننده ANN با لایه های حذف

گزارش طبقه بندی مجموعه داده آزمایش سناریوی اول با استفاده از ANN بدون لایه حذف

support	f1-score	recall	precision	
۱۲	۰/۹۶	۰/۹۲	۱/۰۰	A
۱۰	۰/۹۵	۱/۰۰	۰/۹۱	B
۱۲	۱/۰۰	۱/۰۰	۱/۰۰	C
۳۴	۰/۹۷			accuracy
۳۴	۰/۹۷	۰/۹۷	۰/۹۷	macro avg
۳۴	۰/۹۷	۰/۹۷	۰/۹۷	weighted avg

۳-۲. پیاده سازی سناریو دوم

ابعاد مجموعه داده های آموزش و آزمایش حاصل از سناریوی دوم که در آن، ۲ ثانیه آخر مربوط به هر یک از ضبط ها، برای ایجاد مجموعه داده آزمایش کنار گذاشته می شوند و قسمت باقیمانده هر یک از ضبط ها برای ایجاد مجموعه داده آموزش، به بخش های دو تاییه ای و با همپوشانی ۰/۲ ثانیه تفکیک می شود، در جدول (۴) ارائه می شود:

جدول ۵. ابعاد مجموعه داده های آموزش و آزمایش حاصل از سناریوی دوم

ردیف	تعداد نمونه های آموزش	تعداد نمونه های آزمایش
کلاس A	۲۱۶	۶
کلاس B	۴۷۰	۵
کلاس C	۳۲۶	۶
جمع کل	۱۰۱۲	۱۷

بدین ترتیب، در این سناریو ۱۷ نمونه داده موجود در مجموعه داده اولیه، به ۱۰۱۲ نمونه داده آموزش و ۱۷ نمونه داده آزمایش گسترش داده می شود که هر یک از نمونه های گسترش داده شده، دارای طول یکسان ۲ ثانیه است.

برای مقایسه نتایج حاصل از این سناریو با نتایج به دست آمده توسط مرجع [۸]، در جدول (۵) به صورت خلاصه، درصد طبقه بندی صحیح هر کلاس (درصد یادآوری) با استفاده از طبقه بندی کننده های فوق، به همراه میانگین وزن دار درصد طبقه بندی صحیح هر کلاس آورده شده است. لازم به ذکر است که ستون آخر جدول (۶) نتایج به دست آمده توسط مرجع [۸]

است. با توجه به نتایج ارائه شده در جدول (۵)، تمامی الگوریتم های طبقه بندی پیشنهاد شده در این سناریو قادر هستند که هر یک از کلاس ها را به خوبی تشخیص دهند و به میانگین وزن دار درصد طبقه بندی صحیح هر کلاس بالاتری نسبت به روش به کار رفته توسط مرجع [۸] دست یابند. همه الگوریتم های پیشنهاد شده در این سناریو به جز درخت تصمیم، به بهترین نتایج با میانگین وزن دار درصد طبقه بندی صحیح هر کلاس برابر با ۱۰۰ درصد دست می یابند. در کنار میانگین وزن دار درصد طبقه بندی صحیح هر کلاس بالاتر این سناریو، شیوه انتخاب مجموعه داده های آموزش، اعتبارسنجی و آزمایش این سناریو نسبت به روش به کار رفته توسط مرجع [۸] (انتخاب یکی از ضبط های مربوط به هر کلاس به عنوان داده های آموزش)، روش بهتری است.

جدول ۶. درصد طبقه بندی صحیح هر کلاس در الگوریتم های پیشنهادی در سناریوی دوم و روش مرجع [۴۱]

ردیف	MFCC -ANN	MFCC -KNN	MFCC -DT	MFCC -ANN2	MFCC -SVM	HHT-SVM
شناور A	۱۰۰	۱۰۰	۸۳	۱۰۰	۱۰۰	۶۲/۵
شناور B	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰
شناور C	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۶۲/۵
میانگین وزن دار	۱۰۰	۱۰۰	۹۴	۱۰۰	۱۰۰	۷۵

۳-۳. پیاده سازی سناریو سوم

در پیاده سازی سناریوی سوم، در اثر انتخاب تصادفی سه نمونه داده از میان ۱۷ نمونه داده موجود، به گونه ای که از هر کلاس، یک نمونه داده به عنوان مجموعه داده آزمایش انتخاب شود، نمونه های اشاره شده در جدول (۶)، به عنوان ضبط های آزمایش انتخاب می شوند.

جدول ۷. ضبط های آزمایش انتخاب شده در سناریوی سوم

شماره ضبط	کلاس بخش تفکیک شده	مشخصه ها	مدت زمان (S)
۲	A	حای صدای کاویتاسیون پروانه در سرعت ۷ گره و در عمق پریسکوپ (بخش ۲)	۱۱
۵	B	حای صدای کاویتاسیون پروانه در فاصله ۲۰۰ یارد، در سرعت ۶ گره و در عمق پریسکوپ	۲۴
۴	C	حای صدای کاویتاسیون پروانه در سرعت ۳ گره و در عمق نامعلوم	۲۰

ابعاد مجموعه داده های آزمایش و آموزش حاصل از سناریوی سوم که به ترتیب، از طریق بخش بندی زمانی ۱ ثانیه ای و بدون همپوشانی ضبط های آزمایش و بخش بندی زمانی ۱ ثانیه ای و با همپوشانی ۰,۲ ثانیه ای ضبط های آموزش ایجاد می شوند، در

۴-۳. پیاده‌سازی سناریوی چهارم

در پیاده‌سازی سناریوی چهارم، در اثر انتخاب تصادفی سه نمونه داده از ۱۷ میان نمونه داده موجود در جدول (۱)، به‌گونه‌ای که از هر کلاس، یک نمونه داده به‌عنوان مجموعه داده آزمایش انتخاب شود، نمونه‌های اشاره شده در جدول (۱۰) به‌عنوان ضابط‌های آزمایش انتخاب می‌شوند. به‌غیر از این سه نمونه داده، سایر ضابط‌های موجود در جدول (۱) به‌عنوان ضابط‌های آموزش در نظر گرفته می‌شوند.

جدول ۱۰. ضابط‌های آزمایش انتخاب شده در سناریوی چهارم

شماره ضبط	کلاس بخش تفکیک شده	مشخصه‌ها	مدت زمان (s)
۲	A	حاوی صدای کاویتاسیون پروانه در سرعت ۶ گره و در عمق پریسکوپ (بخش ۱)	۱۱
۱	B	حاوی صدای ضربات پروانه در حال کاویتاسیون و نویز ماشین‌آلات	۳۲
۳	C	حاوی صدای کاویتاسیون پروانه در سرعت ۳ گره، عمق پریسکوپ و فاصله ۵۰۰ یارد (بخش ۱)	۱۲

ابعاد مجموعه داده‌های آزمایش و آموزش که به ترتیب، از طریق بخش‌بندی زمانی ۲ ثانیه‌ای و با همپوشانی ۱ ثانیه‌ای ضابط‌های آزمایش و بخش‌بندی زمانی ۲ ثانیه‌ای و با همپوشانی ۰/۲ ثانیه‌ای ضابط‌های آموزش ایجاد می‌شوند، در جدول (۱۱) ارائه شده است.

جدول ۱۱. ابعاد مجموعه داده‌های آزمایش و آموزش حاصل از سناریوی چهارم

ردیف	تعداد نمونه‌های آموزش	تعداد نمونه‌های آزمایش
کلاس A	۲۳۰	۱۰
کلاس B	۳۶۹	۳۱
کلاس C	۳۳۵	۱۱
جمع کل	۹۳۴	۵۲

بدین ترتیب، در این سناریو ۱۷ نمونه داده موجود در مجموعه داده اولیه، به ۹۳۴ نمونه داده آموزش و ۵۲ نمونه داده آزمایش گسترش داده می‌شود که هر یک از نمونه‌های گسترش داده شده، دارای طول یکسان ۲ ثانیه است. نتایج در جدول (۱۲) جمع آوری شده است.

جدول ۱۲. درصد طبقه‌بندی صحیح هر کلاس در الگوریتم‌های پیشنهادی در سناریوی چهارم و روش مرجع [۸]

ردیف	ANN MFCC-	MFCC-KNN	MFCC-DT	MFCC-ANN2	MFCC-SVM	HHT-SVM
شناور A	۹۰	۱۰۰	۱۰۰	۸۰	۱۰۰	۶۲/۵
شناور B	۷۱	۴۸	۶۵	۶۵	۶۵	۱۰۰
شناور C	۱۰۰	۱۰۰	۴۵	۹۱	۱۰۰	۶۲/۵
میانگین وزن‌دار	۸۱	۶۹	۶۷	۷۳	۷۹	۷۵

جدول (۸) ارائه شده است. بدین ترتیب، در این سناریو ۱۷ نمونه داده موجود در مجموعه داده اولیه، به ۱۰۰۴ نمونه داده آموزش و ۵۵ نمونه داده آزمایش گسترش داده می‌شود که هر یک از نمونه‌های گسترش داده شده، دارای طول یکسان ۱ ثانیه است.

جدول ۸. ابعاد مجموعه داده‌های آزمایش و آموزش حاصل از سناریوی سوم

ردیف	تعداد نمونه‌های آموزش	تعداد نمونه‌های آزمایش
کلاس A	۲۵۵	۱۱
کلاس B	۴۲۹	۲۴
کلاس C	۳۲۰	۲۰
جمع کل	۱۰۰۴	۵۵

برای مقایسه نتایج حاصل از این سناریو با نتایج به‌دست‌آمده توسط مرجع [۸]، در جدول (۹) به‌صورت خلاصه، درصد طبقه‌بندی صحیح هر کلاس (درصد یادآوری) با استفاده از طبقه‌بندی کننده‌های فوق، به همراه میانگین وزن‌دار درصد طبقه‌بندی صحیح هر کلاس آورده شده است. لازم به ذکر است که ستون آخر این جدول نتایج به دست آمده توسط مرجع [۸] است.

جدول ۹. درصد طبقه‌بندی صحیح هر کلاس در الگوریتم‌های پیشنهادی در سناریوی سوم و روش مرجع [۸]

ردیف	ANN MFCC-	MFCC-KNN	MFCC-DT	MFCC-ANN2	MFCC-SVM	HHT-SVM
شناور A	۷۳	۹۱	۷۳	۷۳	۵۵	۶۲/۵
شناور B	۱۰۰	۱۰۰	۳۳	۹۶	۱۰۰	۱۰۰
شناور C	۹۰	۹۵	۵۰	۹۰	۹۵	۶۲/۵
میانگین وزن‌دار	۹۱	۹۶	۴۷	۸۹	۸۹	۷۵

با توجه به نتایج ارائه شده در جدول (۸)، تمامی الگوریتم‌های طبقه‌بندی پیشنهاد شده در این سناریو به‌جز درخت تصمیم قادر هستند که هر یک از کلاس‌ها را به‌خوبی تشخیص دهند و به میانگین وزن‌دار درصد طبقه‌بندی صحیح هر کلاس بالاتری نسبت به روش به‌کاررفته توسط مرجع [۸] دست یابند. ملاحظه می‌شود که نتایج حاصل از این سناریو نسبت به نتایج حاصل از سناریوی اول، از عملکرد پایین‌تری برخوردار است. علت این امر کوچک بودن مجموعه داده در دست، پیچیدگی و تغییرات زیاد آن است. به‌گونه‌ای که هنگامی که از هر کلاس یک ضابط به‌طور تصادفی به‌عنوان مجموعه داده آزمایش کنار گذاشته می‌شود، مشخصه‌های مربوط به ضابط‌ها به همان خوبی که دو ثانیه آخر ضابط‌ها به‌عنوان مجموعه داده آزمایش کنار گذاشته می‌شود، یاد گرفته نمی‌شود؛ ولی اگر مجموعه داده بزرگ‌تر و با تنوع بیشتری در دست باشد، اختلاف کمتری در نتایج حاصل می‌شود.

۳-۶. پیاده‌سازی سناریو ششم

ابعاد مجموعه داده‌های آموزش و آزمایش حاصل از سناریوی ششم مشابه ابعاد مجموعه داده‌های آموزش و آزمایش حاصل از سناریوی دوم است. تفاوت این سناریو با سناریوی دوم این است که در این سناریو به هر نمونه آزمایش حاصل از سناریو اول، مقداری نویز گوسی با توزیع نرمال (با میانگین صفر و انحراف معیار ۰/۱۵) اضافه می‌شود. نتایج حاصله در جدول (۱۴) جمع‌آوری شدند.

جدول ۱۴. درصد طبقه‌بندی صحیح هر کلاس

ردیف	ANN MFCC-	MFCC-KNN	MFCC-DT	MFCC-ANN2	MFCC-SVM
شناور A	۸۳	۸۳	۳۳	۸۳	۱۰۰
شناور B	۱۰۰	۱۰۰	۸۰	۱۰۰	۱۰۰
شناور C	۱۰۰	۱۰۰	۱۷	۱۰۰	۸۳
میانگین وزن‌دار	۹۴	۹۴	۴۱	۹۴	۹۴

با مقایسه نتایج ارائه‌شده در جدول (۱۳) با نتایج حاصل از سناریوی دوم، ملاحظه می‌شود که با افزودن نویز گوسی با میانگین صفر و انحراف معیار ۰/۱۵ به مجموعه داده آزمایش مربوط به سناریوی دوم در این سناریو، همه الگوریتم‌های به‌جز درخت تصمیم، افت عملکرد مدل ناچیزی در مقایسه با نتایج حاصل از سناریوی دوم دارند که این امر به علت تغییر و تنوع ایجادشده در مجموعه داده آزمایش است. همچنین با مقایسه نتایج حاصل از این سناریو با نتایج حاصل از سناریوی پنجم، ملاحظه می‌شود که تمامی الگوریتم‌ها به‌جز درخت تصمیم، بر این مجموعه داده عملکرد بهتری دارند. این امر به علت کوچک‌تر بودن نویز افزوده‌شده در این سناریو و همچنین تخریب کمتر ضبط‌های کیفیت پایین ساخته‌شده از پایگاه داده [۹] است.

۳-۷. پیاده‌سازی سناریو هفتم

از یادگیری عمیق، امروزه در جنبه‌های مختلف مهندسی بهره‌برداری می‌شود. از جمله برای تشخیص امضای دیجیتال نیز این روش پیشنهاد شده است [۱۴-۱۳]. ابعاد مجموعه داده‌های آموزش و آزمایش حاصل از سناریوی هفتم، مشابه با ابعاد مجموعه داده‌های آموزش و آزمایش حاصل از سناریوی دوم است. تفاوت این سناریو با سناریوی دوم این است که در این سناریو، اسپکتروگرام‌های مربوط به مجموعه داده گسترش داده شده حاصل از سناریوی دوم استخراج می‌شود و هر یک از اسپکتروگرام‌ها به صورت یک تصویر با قالب png ذخیره می‌شود. به منظور طبقه‌بندی اسپکتروگرام‌های به دست آمده از این سناریو از یک شبکه عصبی پیچشی (CNN) استفاده می‌شود. جدول (۱۵) درصد

با توجه به نتایج ارائه‌شده در جدول (۱۱)، الگوریتم‌های طبقه‌بندی شبکه‌های عصبی مصنوعی با لایه‌های حذف و ماشین بردار پشتیبان قادر هستند به میانگین وزن دار درصد طبقه‌بندی صحیح هر کلاس بالاتری نسبت به روش به کاررفته توسط مرجع [۸] دست یابند. در کنار میانگین وزن دار درصد طبقه‌بندی صحیح هر کلاس بالاتر در الگوریتم‌های اشاره‌شده از این سناریو از این پژوهش، شیوه انتخاب مجموعه داده‌های آموزش، اعتبارسنجی و آزمایش این سناریو نسبت به روش به کاررفته توسط مرجع [۸] (انتخاب یکی از ضبط‌های مربوط به هر کلاس به عنوان داده‌های آموزش)، روش بهتری است. با این حال ملاحظه می‌شود که نتایج حاصل از این سناریو نسبت به نتایج حاصل از سناریوی دوم، از عملکرد پایین‌تری برخوردار است. علت این امر، کوچک بودن مجموعه داده در دست، پیچیدگی و تغییرات زیاد آن است. به گونه‌ای که هنگامی که از هر کلاس یک ضبط به طور تصادفی به عنوان مجموعه داده آزمایش کنار گذاشته می‌شود، مشخصه‌های مربوط به ضبط‌ها به همان خوبی که دو ثانیه آخر ضبط‌ها به عنوان مجموعه داده آزمایش کنار گذاشته می‌شود، یاد گرفته نمی‌شود؛ ولی اگر مجموعه داده بزرگ‌تر و با تنوع بیشتری در دست باشد، اختلاف کمتری در نتایج حاصل می‌شود. همچنین با مقایسه نتایج حاصل از این سناریو از این پژوهش با نتایج حاصل از سناریوی سوم، ملاحظه می‌شود که نتایج حاصل از اکثر الگوریتم‌های مربوط به این سناریو، به علت انتخاب یک ضبط متفاوت از کلاس B، یعنی ضبط حاوی صدای ضربات پروانه در حال کوایتاسیون و نویز ماشین آلات به عنوان داده آزمایش به نتایج با عملکرد پایین‌تری دست می‌یابد. این امر نیز به علت کوچکی و عدم تنوع کافی مجموعه داده در دست، رخ می‌دهد.

۳-۵. پیاده‌سازی سناریو پنجم

ابعاد مجموعه داده‌های آموزش و آزمایش حاصل از سناریوی پنجم، مشابه ابعاد مجموعه داده‌های آموزش و آزمایش حاصل از سناریوی اول است که در جدول (۲) ارائه شده است. تفاوت این سناریو با سناریوی اول این است که در این سناریو به هر نمونه آزمایش حاصل از سناریو اول، مقداری نویز گوسی با توزیع نرمال (با میانگین صفر و انحراف معیار ۰/۲) اضافه می‌شود. نتایج در جدول (۱۳) خلاصه شده‌اند.

جدول ۱۳. درصد طبقه‌بندی صحیح هر کلاس

ردیف	ANN MFCC-	MFCC-KNN	MFCC-DT	MFCC-ANN2	MFCC-SVM
شناور A	۵۸	۵۸	۶۷	۵۰	۷۵
شناور B	۱۰۰	۱۰۰	۸۰	۱۰۰	۱۰۰
شناور C	۵۸	۹۲	۷۵	۵۰	۵۸
میانگین وزن‌دار	۷۱	۸۲	۷۴	۶۵	۷۶

ساختار مطمئن درون تصویر پیدا کنند، در امر طبقه‌بندی صدا، دارای نقصان هستند [۱۰].

۳-۸. پیاده‌سازی سناریو هشتم

در سناریوی هشتم مجموعه داده‌های آموزش و آزمایش با استفاده از مجموعه داده گسترش داده شده حاصل از سناریوی دوم و روش افزایش تضاد مشابه رویکرد به کاررفته توسط [۱۲]، یعنی با استفاده از فیلتر میانگین متحرک^۱، ایجاد می‌شوند و بدین ترتیب، اسپکتروگرام‌هایی که تضاد آن‌ها، یک‌بار در راستای افقی و بار دیگر در راستای عمودی افزایش داده می‌شود و به صورت دو آرایه جداگانه ذخیره می‌شوند، ایجاد می‌شوند. لذا ابعاد مجموعه داده‌های آموزش و آزمایش حاصل از سناریوی هشتم در واقع، دوبرابر مقادیر ارائه شده در جدول (۱۰) است. به منظور طبقه‌بندی آرایه‌های متناظر به اسپکتروگرام‌ها، از یک شبکه عصبی پیچشی (CNN) استفاده می‌شود.

در این سناریو نیز به علت اندازه کوچک مجموعه داده آموزش و زیاد بودن تعداد پارامترهایی که باید یاد گرفته شوند، مدل یادگیری توزیع واقعی داده‌ها شکست می‌خورد و با توجه به نتایج این سناریو، مدل در یکی از کمینه‌های محلی خود که مربوط به پیش‌بینی داده‌ها بر اساس درصد حضور داده‌ها در مجموعه داده آزمایش است، گیر می‌کند.

علاوه بر نکات ارائه شده در تحلیل نتایج سناریوی هفتم، باید اشاره شود که اگرچه اندازه مجموعه داده آموزش به کاررفته در سناریوی هشتم، دوبرابر اندازه مجموعه داده به کاررفته در سناریوی قبل است؛ ولی این اندازه نیز برای یادگیری پارامترهای بسیار موجود در طبقه‌بندی کننده شبکه عصبی پیچشی با معماری پیشنهادی کافی نیست. علاوه بر این، مشکل اندازه کوچک مجموعه داده آموزش را نمی‌توان تنها با استفاده از داده‌افزایی حل نمود؛ چراکه داده‌افزایی نمی‌تواند تعداد مشاهدات دنیای واقعی را افزایش دهد و فقط با افزایش تعداد داده‌ها کمک می‌کند که مدل نسبت به برخی تبدیلات همانند انتقال زمانی، افزودن نویز و ... مقاوم باشد. به عنوان مثال، بین ۱۰۰ مشاهده واقعی از یک کلاس و ۱۰۰ نمونه داده از یک کلاس، همانند ۱۰۰ مشاهده واقعی از صدای یک زردریایی خاص و ۱۰۰ نمونه داده از صدای یک زردریایی خاص تفاوت وجود دارد.

۳-۹. پیاده‌سازی سناریو نهم

مجموعه داده‌های آموزش و آزمایش مربوط به سناریوی نهم، همان آرایه‌های ذخیره شده در سناریوی هشتم هستند. تفاوت سناریوی نهم با سناریوی هشتم در طبقه‌بندی کننده و ترنند

حضور هر شناور را در مجموعه داده آموزش مربوط به این سناریو ارائه می‌دهد

جدول ۱۵. درصد حضور شناورهای مختلف در مجموعه داده آموزش سناریوی هفتم (دوم)

ردیف	تعداد نمونه‌های آموزش	درصد حضور در مجموعه داده آموزش
شناور A	۲۱۶	۲۱/۳۴
شناور B	۴۷۰	۴۶/۴۴
شناور C	۳۵۶	۲۱/۳۲
جمع کل	۱۰۱۲	۱۰۰

یک نکته در رابطه با روش به کاررفته در این سناریو از این پژوهش، این است که اسپکتروگرام‌ها با تصاویر معمولی متفاوت هستند. از جمله این تفاوت‌ها می‌توان به این موضوع اشاره کرد که محور افقی در اسپکتروگرام‌ها، محور زمان و محور عمودی، محور فرکانس است؛ در حالی که در تصاویر معمولی چنین چیزی صدق نمی‌کند. علاوه بر این، در تصاویر معمولی تصور می‌شود که تمامی پیکسل‌های مجاور به یک شی تعلق دارد؛ ولی در اسپکتروگرام‌ها پیکسل‌ها به لحاظ مکانی گروه‌بندی نمی‌شوند. اگرچه استخراج اسپکتروگرام داده‌های صوتی و به کارگیری شبکه‌های عصبی پیچشی، رویکردی است که در تاریخچه تحقیق بسیار به کاررفته است؛ ولی به این معنا نیست که حتی در صورت وجود مجموعه داده آموزش با اندازه مناسب، این روش بهترین رویکرد است. اشکالات بسیاری در هنگام در نظر گرفتن اسپکتروگرام‌ها به عنوان تصاویر وجود دارد؛ از جمله این که عملکرد مدل بسیار به نرم‌افزار وابسته است و روش محاسبه اسپکتروگرام بر نتایج تأثیر بسیار می‌گذارد. علاوه بر این، یک مشکل ذاتی که در اسپکتروگرام‌ها وجود دارد این است که وقایع صوتی گسسته در اسپکتروگرام تفکیک نمی‌شوند، در عوض همه آن‌ها باهم جمع می‌شوند. این بدین معناست که نمی‌توان ادعا کرد که فرکانس خاص مشاهده شده در یک اسپکتروگرام، به یک کلاس خاص متعلق است؛ بلکه دامنه‌های متعلق به آن فرکانس در اثر تعاملاتی پیچیده بین امواج صوتی همانند حذف فاز (در هنگام برهم‌نهی دو موج مختلف با فرکانس یکسان و فاز مخالف) ایجاد می‌شود و این موضوع قابلیت تفکیک را در هنگام مشاهده اسپکتروگرام دشوار می‌کند. نکته دیگر در رابطه با شبکه‌های عصبی پیچشی که باید به آن اشاره شود، این است که صدا به صورت توالی است و در اسپکتروگرام نیز هر مقدار بعد از مقدار دیگر می‌آید؛ ولی از آنجایی که شبکه‌های عصبی پیچشی این موضوع را نادیده می‌گیرند و با نادیده گرفتن زمان و فرکانس تلاش می‌کنند تا یک

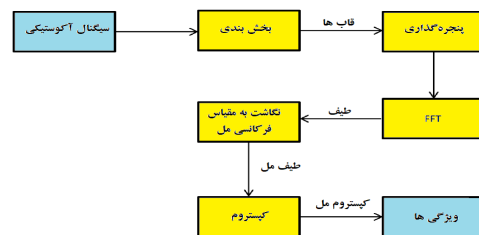
¹Moving Average

فرکانسی مل (MFCC) را در هنگام به کارگیری الگوریتم‌های طبقه‌بندی مختلف، از جمله شبکه عصبی مصنوعی (ANN) و همچنین شکست روش شبکه عصبی پیچشی (CNN) را در هنگام آموزش توسط مجموعه داده‌های کوچک نشان می‌دهد.

۵. مرجع‌ها

- [1] Axelsson, O. "Automatic Classification of Surface Ships Based on Signals from Passive Underwater Sensors"; M.S. Thesis, Lulea University of Technology, Sweden, 2018.
- [2] Malfante, M.; Mars, J. I.; Della Mura, M.; Gervaise, C. "Automatic Fish Sounds Classification"; J. Acoust. Soc. Am. 2018, 143, 2834-2846. <https://doi.org/10.1121/1.5036628>.
- [3] Neupane, D.; Seok, J. "A Review on Deep Learning-Based Approaches for Automatic Sonar Target Recognition"; Electronics 2020, 9, 1972. <https://doi.org/10.3390/electronics9111972>.
- [4] Bach, N. H.; Vue, L. H.; Nguyen, V. D. "Classification of Surface Vehicle Propeller Cavitation Noise using Spectrogram Processing in Combination with Convolution Neural Network"; Sensors 2021, 21, 3353. <https://doi.org/10.3390/s21103353>.
- [5] Leal, N.; Leal, E.; Sanchez, G. "Marine Vessel Recognition by Acoustic Signature"; ARPN J. Eng. Appl. Sci. 2015, 10, 9633-9639.
- [6] Khishe, M.; Mohammadi, H. "Passive Sonar Target Classification using Multi-Layer Perceptron Trained by Salp Swarm Algorithm"; Ocean Eng. 2019, 181, 98-108. <https://doi.org/10.1016/j.oceaneng.2019.04.013>
- [7] Shadojahromi, M.; Bagheri, V.; Rostami, H.; Keshavarz, A. "New Solution for Process and Classifying Inactive Sonar Signal with Echo Sound Features"; Third Int. Conf. and first Int. Practical Research in Electronic and Mechanics and Mechatronics, 2015 (In Persian).
- [8] Sayyadi, H.; Seyf, M.; Abniki, A. "Classifying and Identifying Underwater Purposes Based on Emitted Sound Signals"; Sea Engineering Publication, 2018 (In Persian). <https://doi.org/20.1001.1.17357608.1396.13.26.4.2>.
- [9] San Francisco Maritime National Park Association website
- [10] Rodriguez Molinuevo, J. "Right Whale up-call Detection using Deep Classifiers Over Underwater Noisy Recordings"; M.S. Thesis, Rovira i Virgili University, Spain, Oct. 2019.
- [11] Gimse, H. "Classification of Marine Vessels using Sonar Data and a Neural Network"; M.S. Thesis, Norwegian University of Science and Technology Apr. 2017, Norway.
- [12] Github; Chin, Ch.; Right-Whale-Convolutional-Neural-Network
- [13] Mimura, K.; Mimura, K.; Nakamura, K.; Takao, K.; Yasukawa, K.; Kato, Y. "Automated Detection of Hydrothermal Emission Signatures from Multi-Beam Echo Sounder Images Using Deep Learning"; IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 2023, 16, 2703-2710. <https://doi.org/10.2343/geochemj.2.0419>.
- [14] Khishe, M. "DRW-AE: A Deep Recurrent-Wavelet Autoencoder for Underwater Target Recognition"; IEEE J. Oceanic Eng. 2022, 47, 1083-1098. <https://ieeexplore.ieee.org/document/9852135>.

یادگیری به کاررفته در این سناریو است. در سناریوی نهم، از معماری شبکه و وزن‌هایی که کریستوفر چین در مرجع [۱۲] برای تشخیص تماس‌های نهنگ راست اقیانوس اطلس شمالی به کار می‌برد و برای تشخیص اصوات زیردریایی‌ها استفاده می‌شود. بدین صورت که در این سناریو با تغییر ۳ لایه انتهایی شبکه به کاررفته توسط مرجع [۱۲] و انتقال مستقیم وزن‌های مربوط به ۷ لایه اول، وزن‌های لایه‌های باقی‌مانده توسط مجموعه داده حاصل از سناریوی هشتم (نهم) مجدداً تنظیم می‌شود. بلوک دیاگرام روش مورد استفاده در این سناریو در شکل ۸ توضیح داده شده است.



شکل ۸. بلوک دیاگرام محاسبه MFCC

در این سناریو نیز با وجود به کارگیری یادگیری انتقالی، مدل یادگیری توزیع واقعی داده‌ها شکست می‌خورد و مدل در یکی از کمینه‌های محلی خود که مربوط به پیش‌بینی داده‌ها بر اساس درصد حضور داده‌ها در مجموعه داده آزمایش است، گیر می‌کند. در حالت کلی، یادگیری انتقالی در شرایطی سبب بهبود دقت مدل می‌شود که اولاً توزیع مدل آموزش‌دیده و داده‌های آموزش مطابقت داشته باشد، ثانیاً اندازه داده‌های آموزش حداقل باید به اندازه‌ای باشد که مدل با پارامترهای بسیار زیر برآزش^۱ نکند. علاوه بر این، توجه به این نکته نیز حائز اهمیت است که اصلاح، حذف یا اضافه کردن لایه‌های میانی در هنگام به کارگیری یادگیری انتقالی، از آنجایی که بر معماری مدل تأثیر می‌گذارد، ممکن است سبب عدم یادگیری خوب ویژگی‌ها شود؛ چراکه در شبکه‌های عصبی پیچشی دقیقاً مشخص نیست کدام لایه ویژگی‌ها را استخراج می‌کند.

۴. نتیجه‌گیری

در این پژوهش، با استفاده از روش‌های مختلف پیش‌پردازش و آماده‌سازی داده، ابعاد مجموعه داده تحت سناریوهای خاص گسترش داده شد. با پردازش سیگنال، ویژگی‌های مختلفی از مجموعه داده‌ای مربوط به سناریوهای مختلف استخراج شد و توسط دانش یادگیری ماشین، الگوریتم‌های طبقه‌بندی متنوعی را به کار رفت. در نهایت، با محاسبه معیارهای ارزیابی مختلف، عملکرد سناریوهای مختلف بررسی و تحلیل شد. نتایج این تحقیق، موفقیت‌آمیز بودن روش استخراج ویژگی ضرایب کپسترال

^۱ Underfit