Received: 15 September 2023 | Revised: 16 November 2023 | Accepted: 18 November 2023

IET Radar, Sonar & Navigation

DOI: 10.1049/rsn2.12518

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

Machine learning-based approach for maritime target classification and anomaly detection using millimetre wave radar Doppler signatures

Samiur Rahman | Aleksanteri B. Vattulainen | Duncan A. Robertson

School of Physics and Astronomy, University of St Andrews, St Andrews, Scotland, UK

Correspondence

Samiur Rahman, School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews, Scotland KY16 9SS, UK.
Email: sr206@st-andrews.ac.uk

Funding information

UK Research and Innovation, Grant/Award Number: EP/S032851/1

Abstract

The authors present multiple machine learning-based methods for distinguishing maritime targets from sea clutter. The main goal for this classification framework is to aid future millimetre wave radar system design for marine autonomy. Availability of empirical data at this frequency range in the literature is scarce. The classification and anomaly detection techniques reported here use experimental data collected from three different field trials from three different millimetre wave radars. Two W-band radars operating at 77 and 94 GHz and a G-band radar operating at 207 GHz were used for the field trial data collection. The dataset encompasses eight classes including sea clutter returns. The other targets are boat, stand up paddleboard/kayak, swimmer, buoy, pallet, stationary solid object (i.e. rock) and sea lion. The Doppler signatures of the targets have been investigated to generate feature values. Five feature values have been extracted from Doppler spectra and four feature values from Doppler spectrograms. The features were trained on a supervised learning model for classification as well as an unsupervised model for anomaly detection. The supervised learning was performed for both multi-class and 2-class (sea clutter and target) classification. The classification based on spectrum features provided an 84.3% and 80.1% validation and test accuracy respectively for the multi-class classification. For the spectrogram feature-based learning, the validation and test accuracy for multi-class increased to 93.3% and 88.7% respectively. For the 2-class classification, the spectrum feature-based training accuracies are 88.1% and 86.8%, whereas with the spectrogram feature-based model, the values are 95% and 94.1% for validation and test accuracies respectively. A one class support vector machine was also applied to an unlabelled dataset for anomaly detection training, with 10% outlier data. The cross-validation accuracy has shown very good agreement with the expected anomaly detection rate.

KEYWORDS

Doppler measurement, feature extraction, marine navigation, marine radar, sea clutter, support vector machines

1 | INTRODUCTION

Traditional marine navigation radars usually operate at frequencies up to X-band. There have been significant hardware component improvements in millimetre frequency bands (especially W-band) in recent years due to automotive and 5G applications. This technology development has also generated interest in autonomous marine vehicles using sensors operating at these frequencies. The advantage of increasing carrier frequency is that it enables the design of compact systems, with better range and cross-range resolution and higher Doppler sensitivity. Even though the range coverage decreases, it is still very much applicable to small- to medium-sized vessels.

One of the potential benefits of using millimetre wave radar is the higher Doppler shift and consequently better Doppler resolution for a given integration time which provides more detail in the Doppler domain. At present, most of the reports on Doppler analysis and target classification in the literature are at lower frequency bands. Spectral analysis and empirical observation-based modelling results have been comprehensively reported in refs. [1–5]. In all cases, the main experimental data sources are X-band radars, along with some analysis of C-band and Ku-band radar data.

In ref. [6], the X-band IPIX dataset [7] was used for Doppler spectrum-based small target detection. That work proposed target detection using two features based on either the Rayleigh distribution algorithm or the Entropy algorithm. The algorithms are based on the amplitude distribution in the Doppler domain and Doppler energy distribution levels in target and sea clutter bins. The validation of the method with the experimental data showed varying results. It was shown that the algorithms individually do not provide reliable detection performance, but better results might be obtained by merging them as their detection rates were sometimes contrasting for given data.

In ref. [8], the airborne X-band INGARA radar data [9] was used to develop a stationary wavelet transform (SWT)-based target detection method. Separating the relevant decomposition level and reconstruction was done to discriminate between target and clutter (sea-spikes in this case). Entropy-based calculations were done for this process. Very good detection performance was achieved, where prior knowledge of the target velocity gave a 3–7 dB improvement in the signal to interference ratio requirement. The radar dataset contains medium to high grazing angle data (15°–45°), from which medium grazing angle (30.5°–35.5°) data was used in this case.

A spatio-temporal domain joint filtering-based target detection technique was proposed in [10]. A 3D Fourier transform was used to transform the three-dimensional radar image sequence to the frequency-wavenumber domain. A sea-clutter suppressing filter was then applied. The method was tested on X-band marine radar data collected in the East China sea. An improvement of 8% in detection performance was shown over direct original image-based detection.

Deep learning-based approaches for target detection in sea clutter were explored in refs. [11, 12]. In ref. [11], a combined Yolov4 and Kalman filter-based target detection method was developed using X-band data [13]. A greater than 98% detection performance was reported for this method, showing a slightly better result than using only Yolov3 or Yolov4.

A Convolutional Neural Network and Dual-Perspective Attention (DPA)-based target detection technique was demonstrated in ref. [12], showing a 93.5% detection probability for 10^{-3} false alarm rate. Again, X-band radar data were collected for training the model. A one-class classifier operating in the time–frequency feature space using a normalised smoothed pseudo-Wigner-Ville distribution method. The technique was applied to the open source IPIX dataset and compared with detection methods based on amplitude, Doppler features or fractals, where the new method showed significant improvement in the detection rate (>20%).

In ref. [14], a three feature (normalised Hurst exponent, Doppler peak height and Doppler vector entropy)-based one class support vector machine (OCSVM) model was generated. This was applied to the IPIX dataset for verification, showing better detection probability of small targets than amplitude–frequency or a normalised Hurst exponent detector.

Fractal-based target detection methods have also been explored and reported in refs. [15–17]. The techniques were developed using radar receive data (again mainly open source IPIX) and showed promise for improving small target detection rates.

The literature review shows the shortage of reports on marine target classification beyond X-band and the lack of statistical analysis of categorised targets for algorithm development. The work presented here is part of a project (Sub-THz Radar sensing of the Environment for future Autonomous Marine platforms—STREAM), which intends to fill these gaps in the literature with the intended practical application being for marine autonomy.

As sea clutter arises from a very dynamic environment and is thus highly variable, creating a generalised model encompassing all the different sea conditions which does not also overlap with any target features is highly challenging. The pragmatic approach taken here is that there should be different thresholds or feature distributions for marine targets in sea clutter resulting from a given sea state. Instead of a single classification model, there can be a tree of models which can be used in real time depending on the sea condition. This requires more training of classification algorithms but is expected to be more effective and reliable in a real time scenario.

With this practical application in mind, this study focuses entirely on empirical millimetre wave radar data obtained from various field trials, as it is often observed that experimental trial data reveals features which are not produced by modelling. Millimetre wave radars suffer from atmospheric attenuation which can become significant at long ranges and will degrade the signal level and subsequently the signal-to-noise ratio (SNR). Similarly, millimetre wave radars will suffer more from absorption and backscatter due to rainfall compared to lower band radars (10 GHz or less) [18]. Even though the backscatter does not increase linearly with frequency due to Mie scattering [18], it is significantly higher than X-band. So, this study does not consider large, slow-moving vessels, which need long range surveillance. Instead, the intended end users would be small- to medium-sized vessels with fast manoeuvrability, where relatively short-range coverage of a few hundred metres is acceptable but the application requires fine spatial resolution and high Doppler sensitivity.

The advantage of higher Doppler sensitivity at millimetre wave bands is also used here for developing classification models. It should be noted that the method used is not strictly the conventional detection, track and identify process. The Doppler information can be used to directly determine the class and output an alarm accordingly (if classified as sea clutter then no alarm is raised). Tracking can then be started for the

target of interest. In terms of the Doppler statistical properties used for classification, the computational load has also been kept in mind for real time processing. For this reason, central moments and entropy values are used, which are both widely employed to evaluate signal characteristics and are not computationally very expensive. The anticipated novelty of this study is as follows:

- Classification by supervised learning of labelled low grazing angle low sea-state maritime targets along with sea clutter, instead of conventional detection. This target identification method may be useful for future smart autonomous systems to make decisions depending on the target type.
- Dataset of experimental Doppler signatures and statistical properties of sea clutter and marine targets at millimetre wave bands which is used for target discrimination.
- Anomaly detection by unsupervised learning of low grazing angle low sea-state maritime targets along with sea clutter using millimetre wave radar data.

Three millimetre wave radar systems (77, 94 and 207 GHz) have been used for extensive data collection of sea clutter and of various targets expected to be encountered in a maritime environment. All the datasets correspond to low grazing angle (1°–5°). Section 2 of the paper provides details of the three field trials corresponding to the datasets used for this study. The radar specifications are also given here. Section 3 describes the dataset used for classification and anomaly detection training, and Section 4 describing the methodology along with the definition of the Doppler domain features and their extraction processes. Section 5 shows and discusses the results of the classification training for both the multi-class (8-class) and the 2-class scenario. Section 6 demonstrates the OCSVM training method using unbalanced data for outlier/anomaly detection. An overall summary and discussion of intended future work are presented in the last section.

2 | FIELD TRIAL OUTLINE

2.1 | Locations

All the three trials were held in the UK, Figure 1a, Trial 1 was held at the Bruce Embankment, St Andrews, on the 15 December 2020. The trial location coordinates are 56°20′41″ N 2°48′06″ W. As seen in Figure 1b, a 94 GHz radar was used in this trial. In this trial, littoral sea clutter data was collected where the Douglas sea-state was around 0–1. Both incoming and receding wave data were collected. A protruding rock was also within the scene, which has been labelled as one of the targets (stationary solid object). The wind speed and temperature were quite stable throughout the data collection period at 13 km/h from the south and 8 °C. The grazing angle of the data was about 1°–3°.

Trial 2 was held on the 10th September, 2021 at the pool facility at the Sea Mammal Research Unit in St Andrews. Figure 1c shows the experimental setup where lots of static clutter can be seen around the pool. Careful antenna pointing (5° grazing angle) during data collection and static clutter removal during post processing were done to minimise the effects of the surrounding strong clutter return as much as possible. Data of three adult California sea lions engaged in different activities (swimming with different parts of the body exposed above the water, jumping) were collected [19]. The 77 GHz radar used in this trial had a wide beamwidth to provide a sufficient beam footprint in this short-range setup.

Trial 3 was conducted at Coniston Water, Lake District (54°20′50.67″ N 3°4′48.62″ W) over three days (30 August 2022–1 September2022). Figure 1d shows the trial setup, where the 94 GHz and the 207 GHz radar data were obtained. This was an extensive trial, where many clutter and target data were collected. Even though the clutter here is not sea clutter as this is a freshwater lake, for brevity the term 'sea clutter' will be used here to define Trial 3 clutter data. As small- to medium-sized boats will often encounter clutter returns from lake/freshwater, the data is of equal importance in the context of the application, and also likely to be very similar to short range sea clutter in low sea states. The weather remained benign throughout the trial period (~10 km/h with occasional gusts increasing the speed to 15/20 km/h). The Douglas sea-state was roughly 0–1 for the whole duration. As the radars were placed on the shore, only incoming waves were observed. Target data were collected both opportunistically (boat, stand up paddleboard [SUP] etc.) as well as in a planned manner (swimmer, pallet). The grazing angle range was 1°–4°.

It should be noted that in all three trials, 24 GHz radar data were collected as well. The 24 GHz data have not been utilised in this study for two reasons. Firstly, this work focuses exclusively on the millimetre wave range (30–300 GHz). Additionally, the spatial resolution and/or antenna beamwidth of the 24 GHz radar are different from the other three instruments, leading to slightly different statistical properties.

2.2 | Radar systems

All three frequency modulated continuous wave (FMCW) Doppler radars used for this work were designed and built at the University of St Andrews. Table 1 provides the pertinent radar parameter values. The linearly polarised W-band 77 GHz radar FAROS-E was initially built as a demonstrator for drone detection [20], but in Trial 2 it was used with a pair of low gain horn antennas for the reason mentioned above. It can operate with either horizontal–horizontal (HH) or vertical–vertical (VV) polarisation. The W-band 94 GHz radar T-220 has very low noise [21], and operates with circular polarisation (CP), with odd bounce, right CP transmit and left CP receive. The G-band 207 GHz radar Theseus was built during this project [22], primarily for target and clutter phenomenology studies, but can also be used for other applications such as security. It should be noted that the antenna beam footprint and range resolution of the 94 and 207 GHz radars are comparable. The CRI for each radar were set to ensure similar unambiguous Doppler range for all the systems. All the radar data corresponding to this study were collected in staring mode.

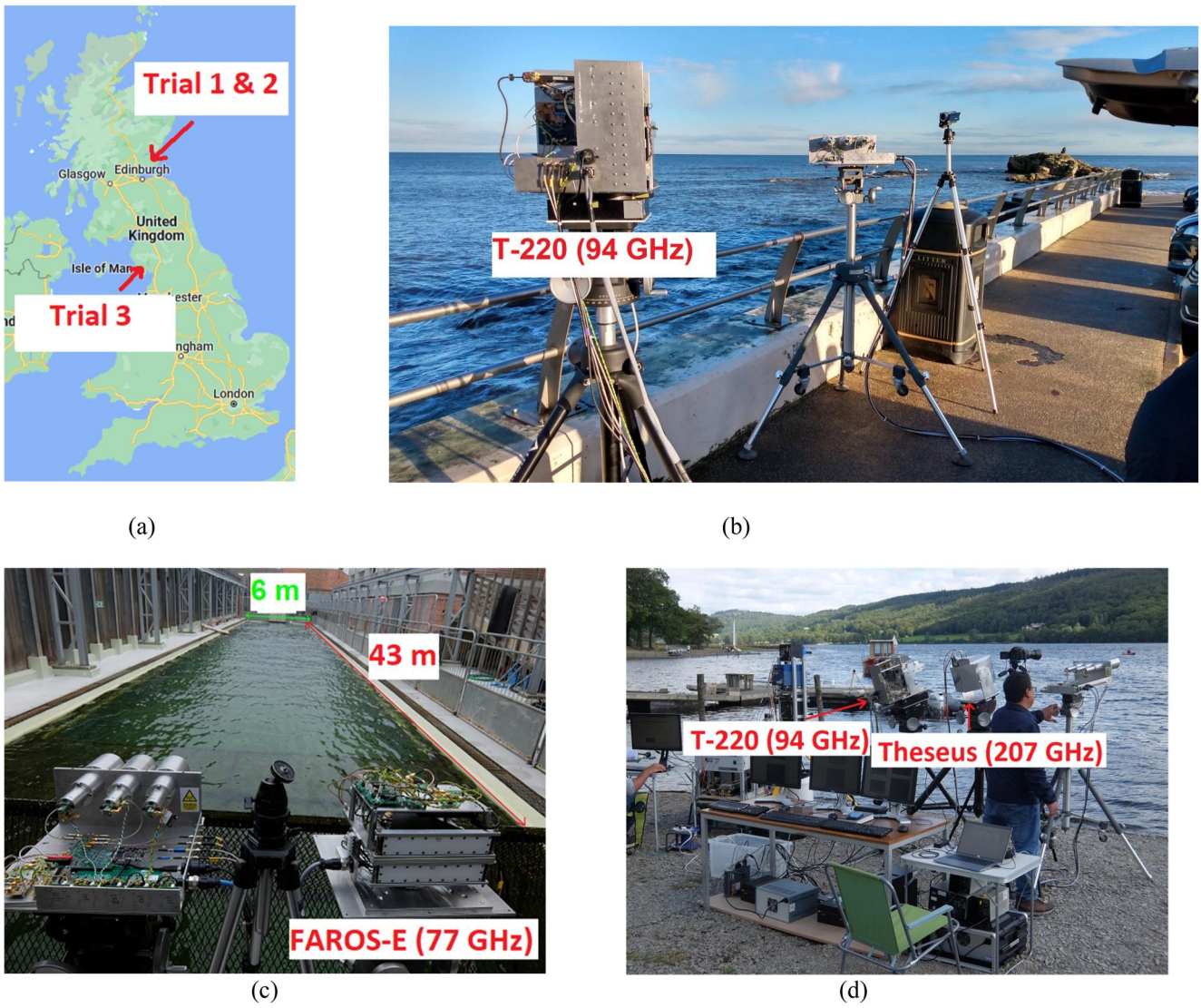

FIGURE 1 (a) Field trial locations, (b) Bruce Embankment trial setup (Trial 1), (c) SMRU trial setup (Trial 2), and (d) Coniston trial setup (Trial 3).

Parameter	FAROS-E	T-220	Theseus
Centre frequency	77 GHz	94 GHz	207 GHz
Modulation	FMCW	FMCW	FMCW
Antenna beamwidth (one-way)	13° az., 13° el.	0.92° az., 3° el.	2° az., 2° el.
Antenna gain	22.2 dBi	40.5 dBi	38.1 dBi
Polarisation	Linear (HH, VV)	Circular (odd bounce)	Linear (HH, VV)
Tx power	+25 dBm	+18 dBm	+12 dBm
Bandwidth/range resolution	750 MHz/20 cm	750 MHz/20 cm	2000 MHz/7.5 cm
Chirp repetition interval (CRI)	150.9 μs	122.34 μs	67.58 μs
Maximum unambiguous velocity	±6.46 m/s	±6.54 m/s	±5.18 m/s
Instrumented range	204.8 m	204.8 m	153.6 m

TABLE 1 77, 94 and 207 GHZ radar specifications.

3 | DOPPLER DATASET

The labelled dataset for classification training was created to incorporate diversity in terms of operating frequency, polarisation, wave direction, wind speed and range. Figure 2 shows images of all eight targets. It should be understood that the majority of the data from Trial 1 and 3 are of clutter. Also, the distribution of the other targets in the dataset is not equal. This simply points to the fact that in practice, it was not possible to obtain equal amounts of data for all the targets. Plots in Figures 3 and 4 demonstrate the nature of the datasets. The clutter returns from Trials 1 and 3 were mostly spikes from burst scattering rather than distributed returns [2]. This is due to the combination of calmer sea condition and clutter sensitivity at different frequencies. At these millimetre-wave frequencies, the clutter returns have been observed to be spikier in nature compared to lower bands. This was the case even during stronger waves due to wind gusts. In Figure 3a,b, range-time-intensity (RTI) plots show the spiky nature of the clutter returns. Both the plots were obtained from Trial 3, but the same trend was present at Trial 1 as well [22]. It is also seen that the 94 GHz radar returns are significantly stronger than the 207 GHz returns. Figure 3a–d show that the 207 GHz radar only picks up the brightest returns compared to 94 GHz. This is due to the greater hardware sensitivity of the 94 GHz radar, as it becomes more difficult to maintain the same radar link-budget at higher frequency bands. Whether this is also an effect of the different types of interaction between the sea surface and the radar signals at different frequencies is not verified yet and will require further study. However, the spike returns have very good SNR, as evident from Figure 3c,d. The duration of this burst scattering type spikes is about a few hundred milliseconds, as shown in Figure 3e. They tend to remain correlated over a large part of that duration, as shown in Figure 3f. This agrees with the observations of X-band burst scattering [1].

The RTI plots of different types of targets for the training dataset are shown in Figure 4. The target data have been obtained from all three trials, where Trial 2 was specifically only for sea lion data collection. Targets were all within short range (<100 m), providing high SNR as demonstrated in Figure 4. The high resolution at millimetre-wave frequencies allows the acquisition of detailed signatures which are advantageous for Doppler feature extraction. The labelled dataset for supervised learning is comprised of 564 values for each Doppler feature. The eight classes and the number of single feature values corresponding to each of those targets are as follows:

- Sea clutter (324)
- Swimmer (28)
- Sea lion (96)
- Boat (20)
- SUP/Kayak (16)
- Buoy (28)
- Stationary solid object (32)
- Pallet (20)

The SUP and kayak have been grouped into a single class, mainly because very few opportunistic examples of kayak data were collected.

4 | METHODOLOGY

The classification process in this study is predicated on leveraging high quality data acquired in the millimetre-wave band as this can provide higher range and Doppler resolution. The computational load however usually increases with

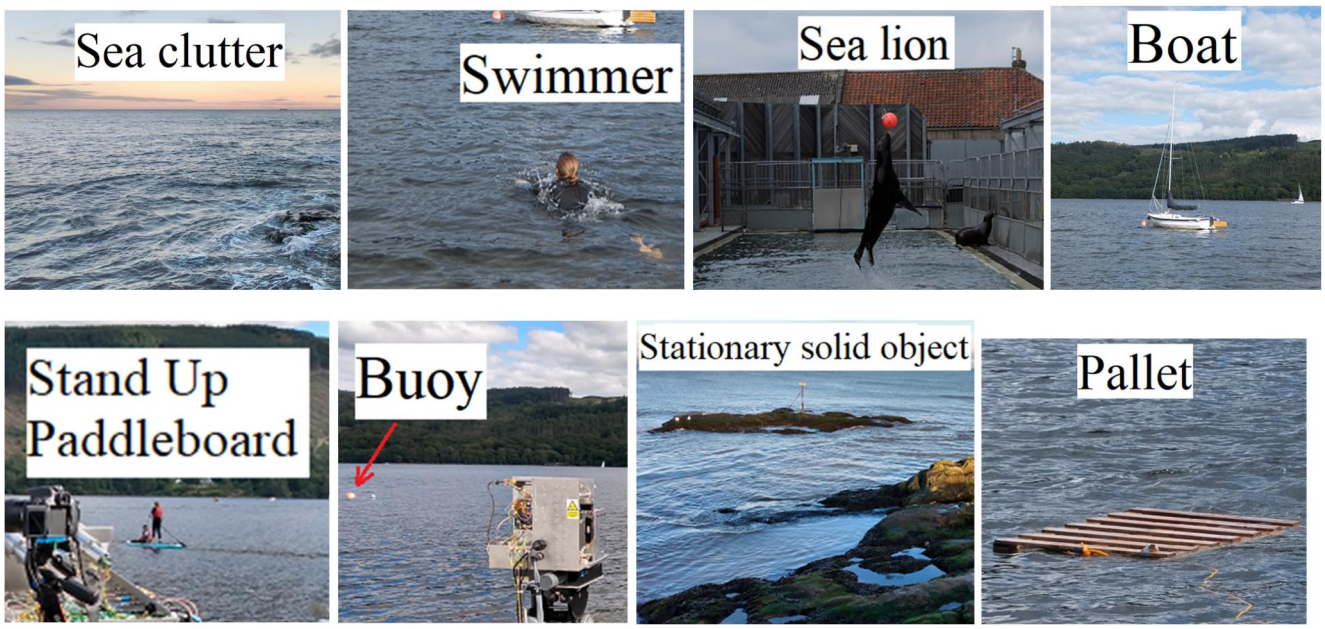

FIGURE 2 Targets (including sea clutter) used for supervised Doppler-based classification training.

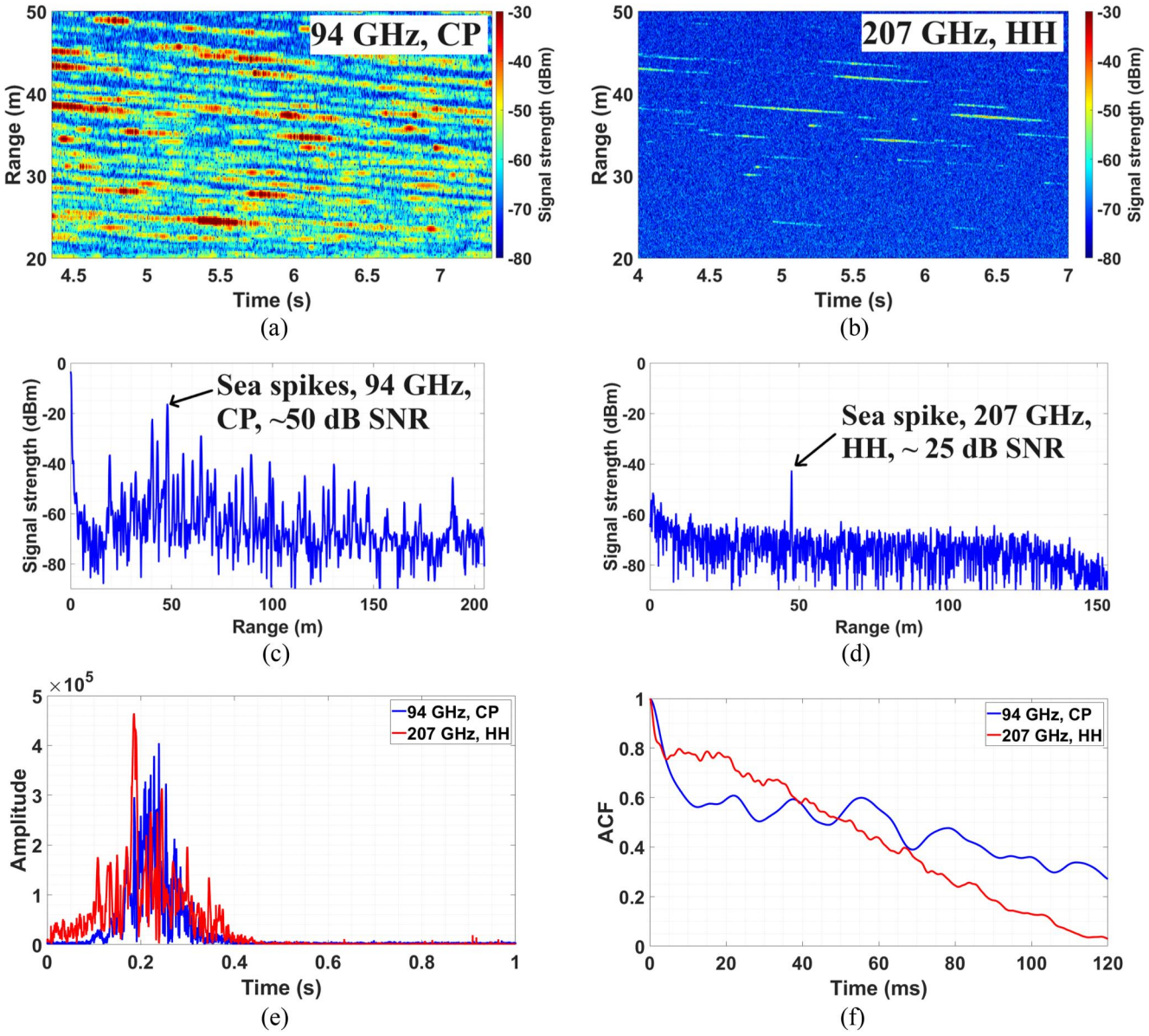

FIGURE 3 Examples of sea clutter datasets: (a) RTI plot at 94 GHz, (b) RTI plot at 207 GHz, (c) range profile at 94 GHz, (d) range profile at 207 GHz, (e) clutter range cell time history showing sea spikes at 94 and 207 GHZ, and (f) autocorrelation plots of the spikes at 94 and 207 GHz.

higher frequency radar data, and as the main anticipated use-case for these marine radars operating at millimetre wave band are small to medium-sized private or commercial vessels, a relatively low cost solution is required without degrading the sensor autonomy performance. So, in this study, the approach was twofold. Firstly, the use of feature extraction-based classification to garner the quantitative insight into Doppler characteristics, as reports on the Doppler properties of marine targets and clutter are very scarce. Secondly, to explore in detail various classification training models with different combinations of features. This undertaking should thus provide an extensive framework for using Doppler data in smart autonomous marine navigation. To obtain feature values, data files were processed by segmenting into 2 s slices. A longer integration time would be beneficial to obtain more reliable feature values, but it was a trade-off with a practical scanning system in mind, which may have an integration time of a fraction of a second. The motivation here is to have a broader understanding of the Doppler statistics and how they differ for various targets which can be better understood with a comparatively longer integration time. The information can then be used in the design of a practical autonomous sensor.

Algorithm 1 Doppler spread feature value.

1. **Input:** $S(i, j)$; $i = 1, 2, \ldots, I$; $j = 1, 2, \ldots, J$
   I = Doppler slice length
   J = Number of time slices
2. **Doppler data points:** $DDP(j)$
   While j < J

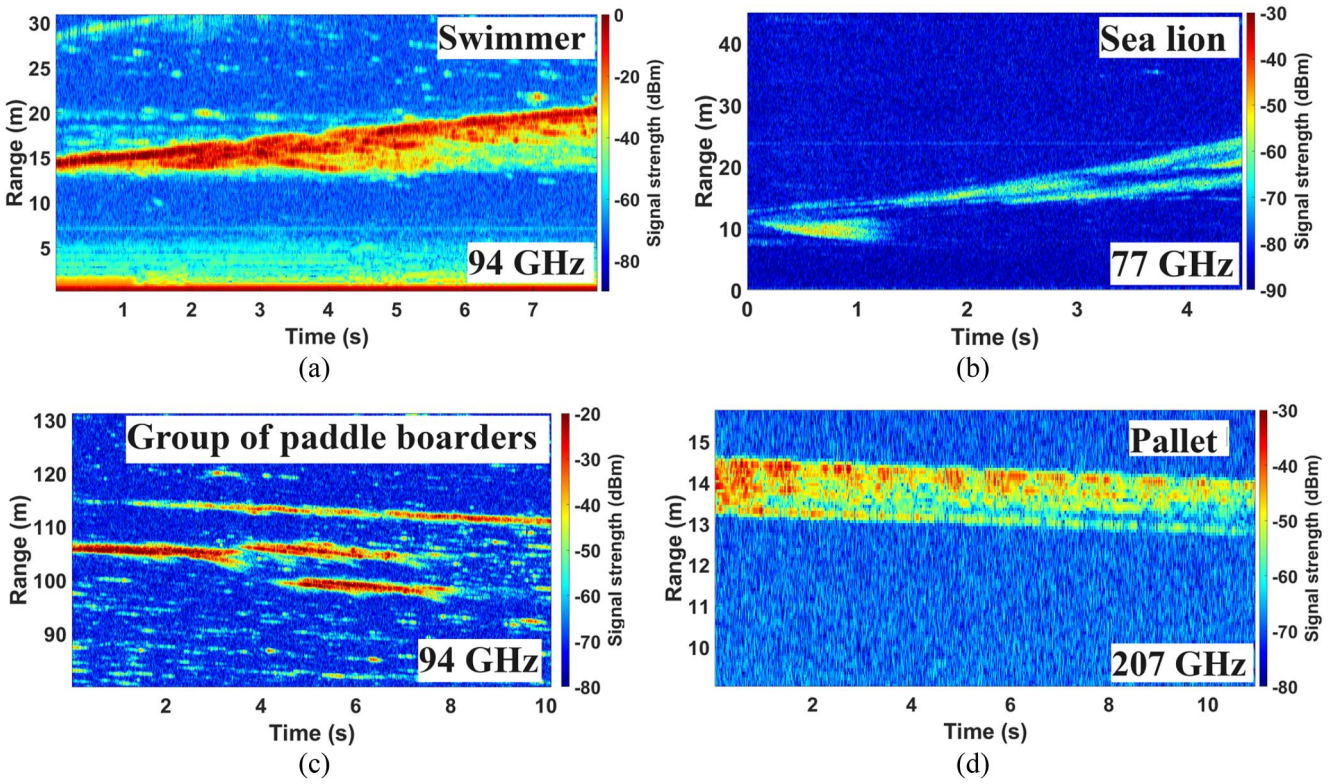

FIGURE 4 Example RTI plots of target datasets: (a) swimmer at 94 GHz, (b) sea lion at 77 GHz, (c) group of paddle boarders at 94 GHz, and (d) pallet at 207 GHz.

```
{While i < I
    { if S(i,j) >
    Threshold
    DDP(j) + = 1}}
```

3. **Output** : $DS = \sum_{j=1}^{J} DDP(j)$

Algorithm 2 Skewness feature value.

```
1. Input: S(i. j); i = 1,2,…,I; j = 1,2,…,J
2. Doppler slice skewness: DSS(j)
        While j < J
        {DSS(j) = Skewness [Dopp_slice]}]
3. Output: MS = mean [DSS(j)]
```

Algorithm 3 Kurtosis feature value.

```
1. Input: S(i. j); i = 1,2,…,I; j = 1,2,…,J
2. Doppler slice skewness: DSS(j)
        While j < J
        {DSS(j) = Kurtosis [Dopp_slice]}]
3. Output: MK = mean [DSS(j)]
```

Doppler data processing for the feature extraction process was performed by combining multiple range bins, instead of a single range slice. In this way, more information of a moving target Doppler signature is present within the same time window. This can also lead to the presence of multiple targets within a single Doppler spectrum or spectrogram. Hence, the number of range bins were selected in such a way which would account for slow moving targets (~1–2 m/s speed). For example, a target with 1 m/s speed would migrate 2 m within a 2 s segment. For the 94 GHz configuration where the range resolution is 20 cm, this would correspond to 10 range bins. During the Doppler processing, those 10 contiguous range bins would then be coherently added. Even though in this study the number of range bins were selected by manual observation of the raw data, in practice this can be automated by using a fixed range swath for each 2 s slice.

Algorithm 4 Entropy feature value.

```
1. Input: S(i. j); i = 1,2,…,I; j = 1,2,…,J
2. Doppler slice max value index: DMI(j)
        While j < J
        {DMI(j) = findpeak [Dopp_slice]}
3. Output: EN = approximate entropy
[DMI]
```

The classification structure is as follows-
For Doppler spectrum feature-based classification:

1. Perform fast Fourier transform (FFT) on radar time series data to obtain range profiles.
2. Select the number of contiguous range profiles corresponding to 2 s integration time.
3. Perform slow time FFT across a range bin or multiple range bins to obtain the Doppler spectrum. If multiple range bins are selected, the complex fast time FFT'd signals should first be coherently added.
4. Compute the feature values.
5. Feed the feature values into the trained classifier.
6. Obtain the target type prediction from the multi-class classifier or binary classification (clutter or target) from 2-class classifier.

For the spectrogram-based feature classification, the exact same steps are taken, except for the third step. Instead of performing a single FFT over the whole vector, a short time Fourier transform (STFT) is performed to generate the spectrogram. The STFT window length is 62 ms, kept short for better temporal resolution. During the STFT, a 99% overlap is used.

4.1 | Spectrum-based features

The Doppler spectrum-based training dataset corresponds to five feature values.

- Peak Doppler frequency
- Full width half maximum (FWHM)
- Second order central moment (standard deviation)
- Third order central moment (skewness)
- Fourth order central moment (kurtosis)

Figure 5 shows example spectra of the eight targets, which show different shapes, with their corresponding feature values. For example, the buoy and pallet have high kurtosis values due to their pointed shapes. The raised Doppler around zero for the stationary solid object is due to the waves hitting the rock.

4.2 | Spectrogram-based features

Figure 6 shows example spectrogram plots of all the targets. The spectrograms shown here correspond to a single data file in each case. Two second slices of these spectrograms were used during the feature extraction process. Four spectrogram feature values were obtained for each segment. The feature values are:

- Doppler spread
- Third order central moment (skewness)
- Fourth order central moment (kurtosis)
- Entropy

Feature ranking was performed (discussed in Section 4) to determine the impact of different features. The standard deviation in the case of the spectrogram-based training scored quite low, and so was discarded. This is due to all the classes (including sea clutter) having similar standard deviation values. The reason for this is that other targets also encompass return from the water's surface due to breaking waves and wakes generated by the target motion. Hence, parameters corresponding to overall shapes (i.e. third and fourth central moments) produce more distinguishable features. The entropy feature corresponds to temporal behaviour thus can increase the classification model performance when combined with the other features.

The Doppler spread was calculated simply by taking a Doppler slice and calculating the number of data points above a given threshold (~10 dB higher than the Doppler noise floor). This is done for all the slices within the 2 s timeframe and the total sum is then used as the feature value.

For extracting the skewness and kurtosis feature values from a spectrogram, the central moment values for each Doppler slice were calculated. For a single feature value, the mean of those values was taken.

It should be noted that a stationary solid object can be of different sizes and shapes of protruding rock. Meanwhile, this study does not focus on amplitude characteristics. The shape of a stationary target return in the Doppler domain would be very similar regardless of the target type.

For the entropy feature, at first the index corresponding to the maximum signal strength for each Doppler slice within the spectrogram frame was determined. A one-dimensional vector with those maximum index values was then created. The approximate entropy of the vector is then calculated, which is then used as the feature value. Approximate entropy quantifies the amount of randomness in temporal data and was calculated by the method described in ref. [23]. The algorithms to obtain these feature values are summarised in Algorithms 1–4, for a spectrogram matrix S. It should be noted that due to the sea lion data being collected in a confined pool, strong stationary clutter returns from the surroundings of the pool are present in the sea lion data, which can be seen in Figure 4 as the strong zero Doppler signal. During the feature extraction process, a high pass filter was used to suppress the zero Doppler return as in a realistic open water scenario these returns from around the pool would not be present.

Figure 7 illustrates the example plots of the statistical features obtained from the Doppler spectrograms for feature extraction. The discriminatory features are visually not always apparent for all cases, but some differences can be realised. For instance, in the bottom two plots, the maximum value index plot for a swimmer shows a periodic shape, corresponding to the Doppler feature seen in the spectrogram plot, whereas, the same plot corresponding to pallet has a very straight line. The impact of computational load was taken into account while forming the feature extraction method due to the intended real time operation. For example, singular value decomposition (SVD) was not explored as this process is usually computationally expensive.

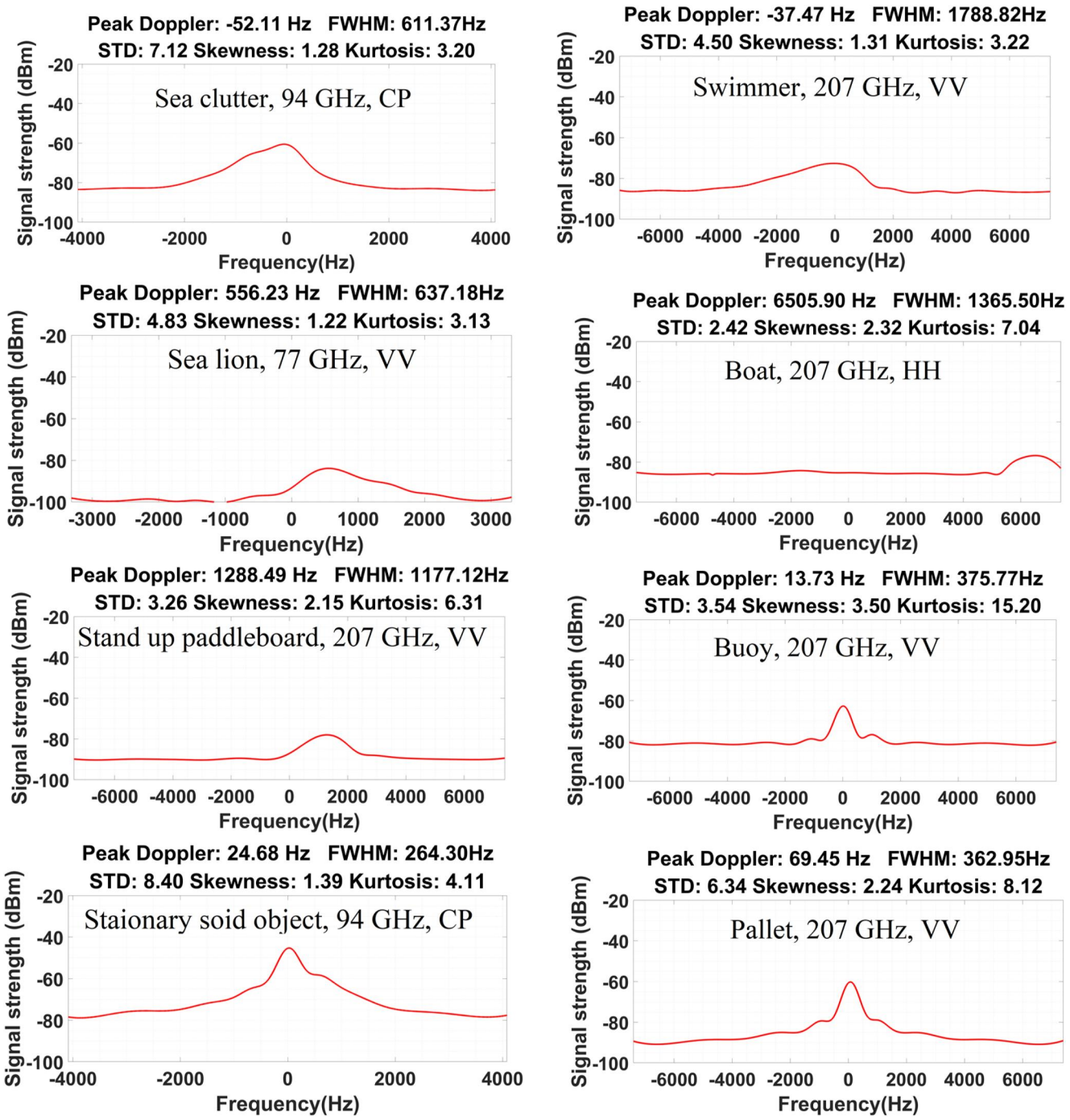

FIGURE 5 Example Doppler spectrum plots and the corresponding feature values of eight different targets.

5 | CLASSIFICATION RESULTS

The labelled dataset with feature values was trained with different classifiers. Fivefold cross validation was used during training, and 5% of the data were separated for test accuracy calculation. In general, the classifier performance depends on the boundary layers of the features. Different conventional classifiers were used for training (Kernel Naïve Bayes [KNB], linear discriminant [LD], linear support vector machine [LSVM], quadratic support vector machine [QSVM]). It was found that QSVM consistently provided the best results.

5.1 | 8-class classification

Figure 8a shows the confusion matrix for the 8-class classification using spectrum-based features. It shows the pallet and sea clutter classes being predicted most accurately, where it

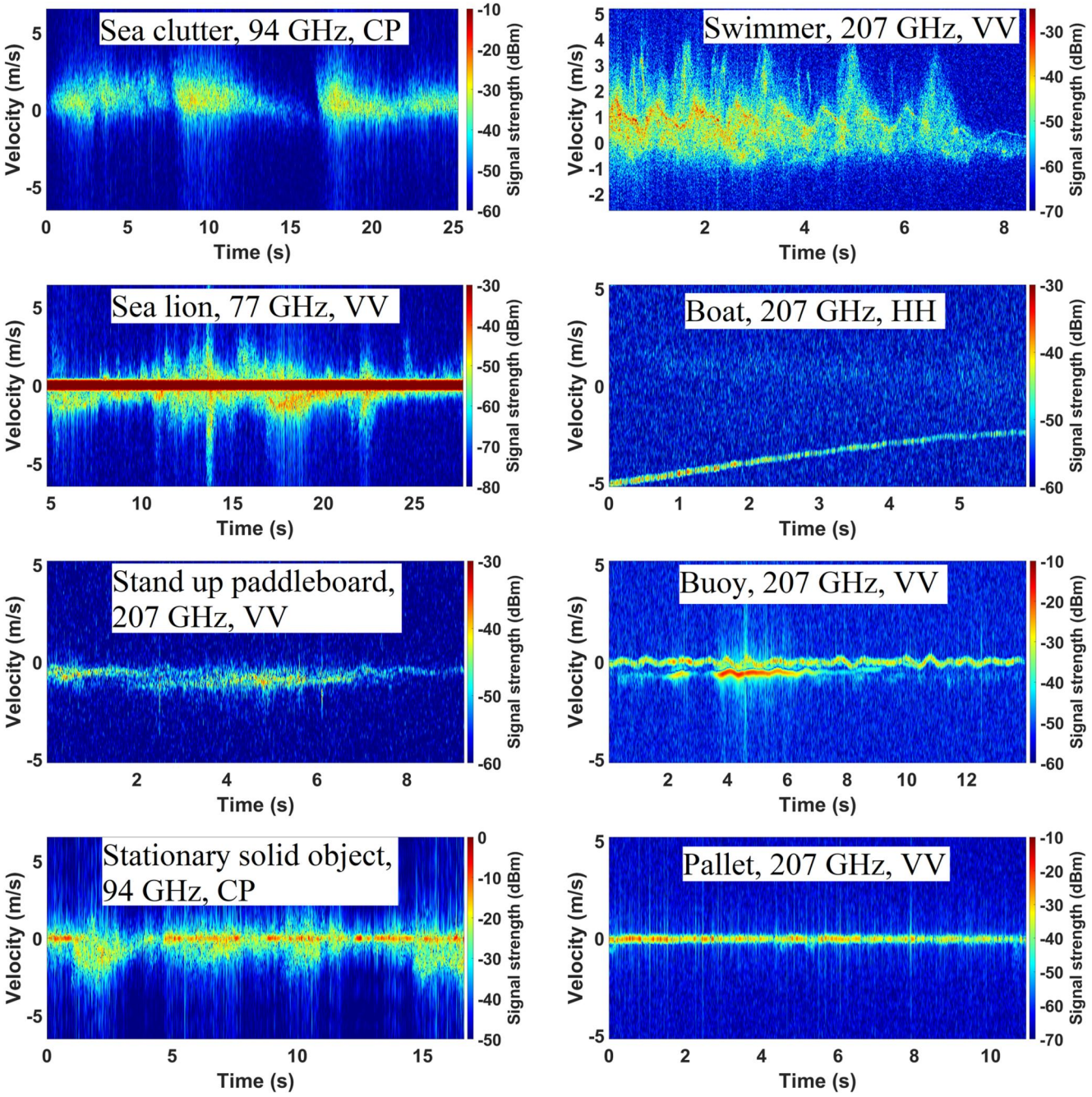

FIGURE 6 Example Doppler spectrogram plots of eight different targets.

should be noted that the training dataset contains more than 50% sea clutter data. Swimmer prediction performance is the worst, closely followed by boat, buoy and SUP. In half the cases, SUP and buoy targets have been predicted as clutter. Spectrogram feature-based classification improves the validation accuracy by 9%, seen in Figure 8c. In this case, the sea lion has a 100% prediction accuracy and sea clutter prediction is also high. Again, boat classification accuracy is very poor, just 50%. In both confusion matrices, it is seen that misclassification of different targets occurred mainly when those were confused with sea clutter. This is perhaps not surprising, as sea clutter is the most varying target of all due to its dynamic nature.

5.2 | 2-class classification

The 2-class classification training model shows a slightly better validation accuracy with both spectrum and spectrogram features, which is expected. As seen in Figure 8b,d, the accuracies increase to 88.1% for spectrum features and 95% for spectrogram features. These are not large improvements, especially

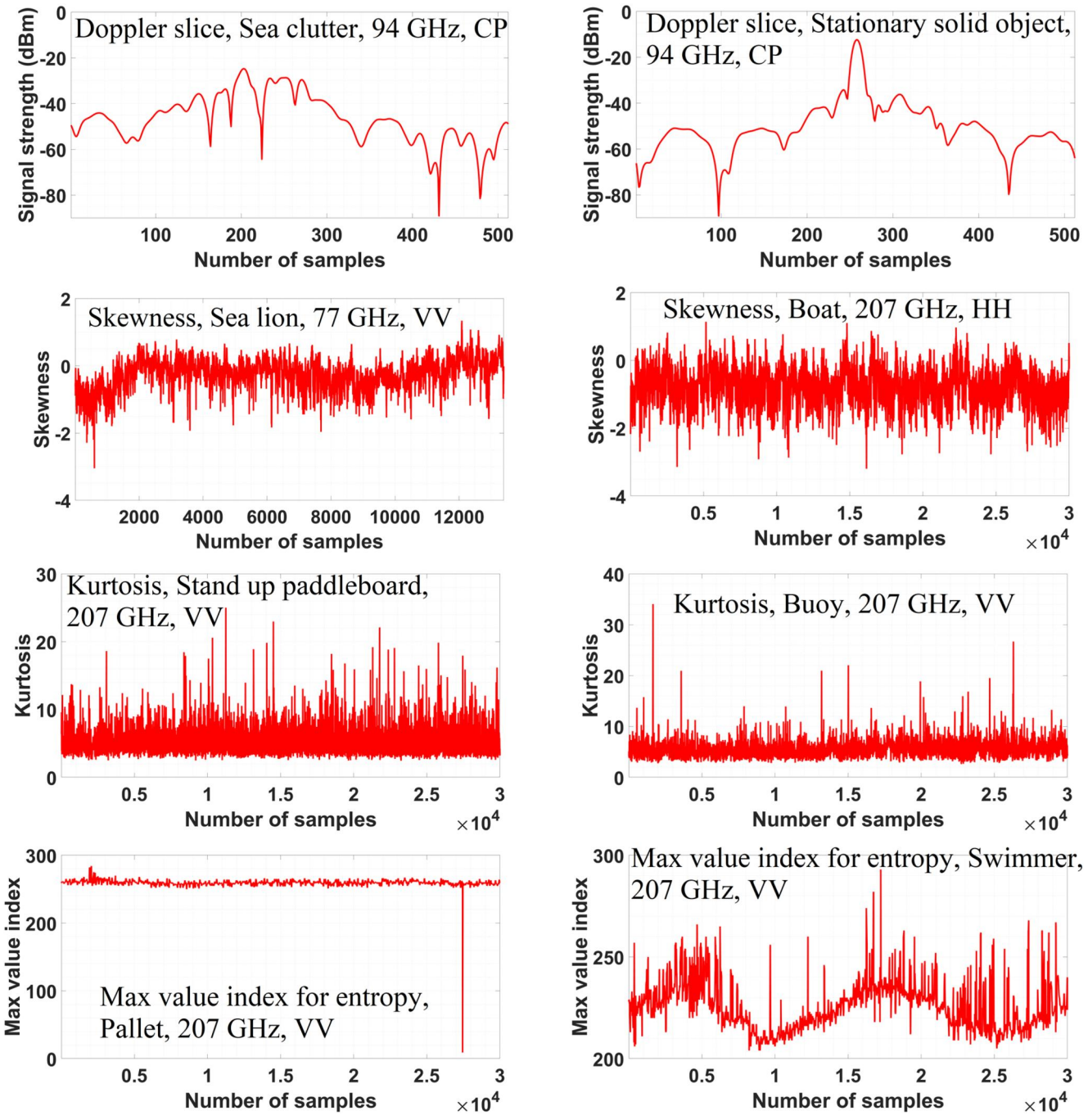

FIGURE 7 Example spectrogram feature plots of eight different targets.

in the case of spectrogram features. This is perhaps indicative of the overlap of target feature boundaries being greater for the sea clutter feature values than between other targets, as seen in Figure 8a,c.

Feature importance scoring for the spectrum and spectrogram features was performed. The ReliefF algorithm was used as this incorporates multiclass categorical variable, which allows for dealing with multiclass problems [24]. Chi-Square algorithm is another suitable candidate for feature ranking and was investigated as well. The ranking order in that case changed only for the last two features. As the performance of Chi-Square is comparatively more dependent on the data sample size, only ReliefF findings are considered here. The ranking result can be seen in Figure 9. The three central moment values have ranked highest for spectrum features. For spectrogram features, the Doppler spread feature is by far the dominant one, whereas entropy and kurtosis have the lowest importance. Even though Table 2 shows that the classification performance decreases with feature dimensionality reduction, this ranking gives a very useful indication of which features can

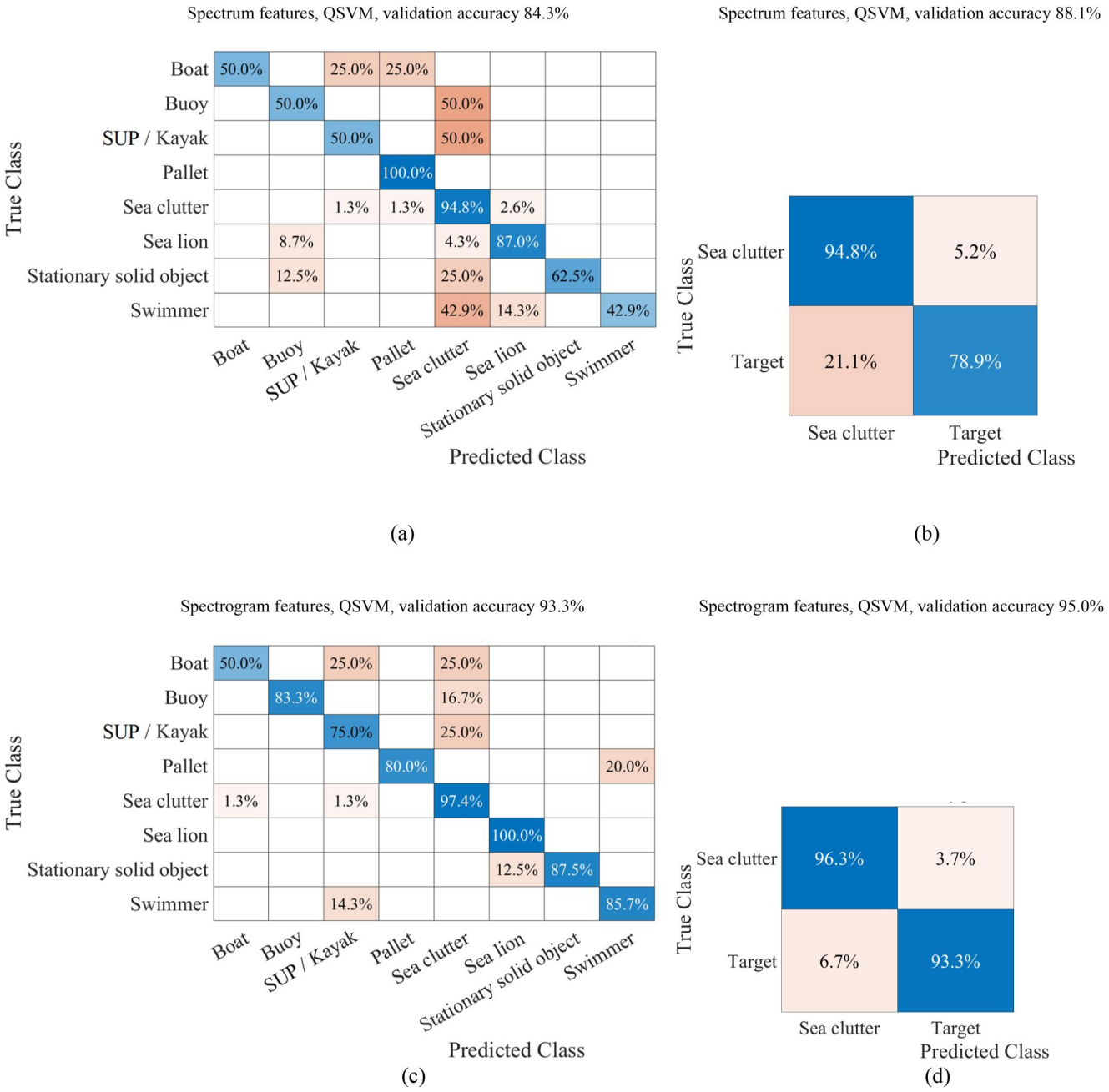

FIGURE 8 Confusion matrices of the QSVM training models showing the classification validation accuracies for 8-class and 2-class classification, (a) 8-class spectrum feature-based training, (b) 2-class spectrum feature-based training, (c) 8-class spectrogram feature-based training, and (d) 2-class spectrogram feature-based training.

be omitted if fewer features are used. This is necessary when the computational load becomes very restricted and multiple sensors are used—in which case the final decision would be made by sensor fusion so the accuracy of a single sensor can be relaxed. This ranking is also an important piece of information for the further study of Doppler characteristics, in terms of emphasising which statistical properties are most important. From Figure 10, the feature value distribution for sea clutter and targets can be visualised. It is evident that the Doppler spread features are more separated compared to the others. The overlap between clutter and target features in all cases in these 2D plots suggests that it is essential to use hyperplanes during classification training, even if fewer features are used. This justifies the training of different classifier models to obtain the optimised result as otherwise it would be very difficult to predict the best model. Classification training was conducted using various combinations of features with the different classifiers. The detailed summary of the obtained classification results is given in Table 2. QSVM showed best performance in all cases, whilst LD performed quite poorly.

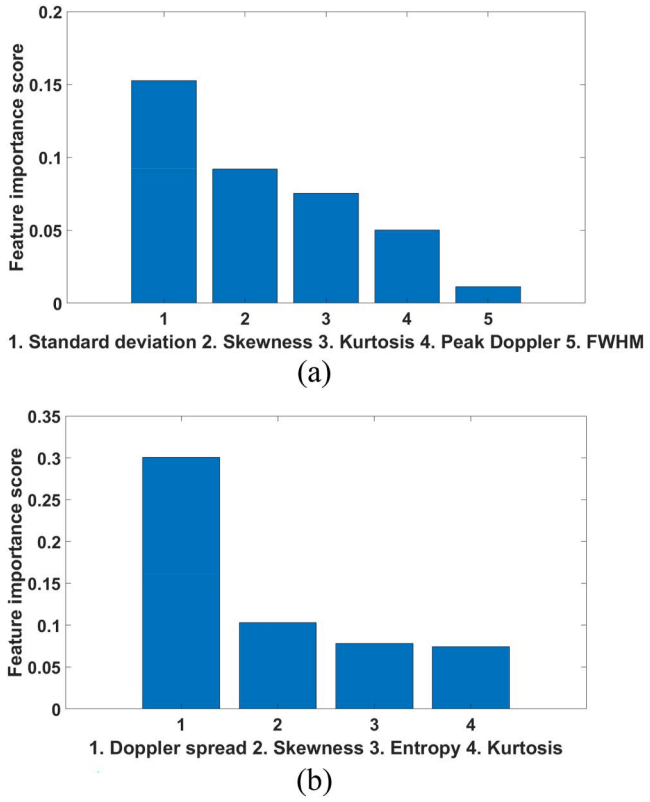

FIGURE 9 Feature importance score obtained by ReliefF algorithm: (a) spectrum features and (b) spectrogram features.

Quadratic Discriminant Analysis was also tried but it often failed to converge. It has been observed that the best performance is achieved when all the feature values are used for training. For instance, as the Doppler spread has been found to be the most dominant amongst the spectrogram features, training was then performed using only the Doppler spread feature but the classification accuracy reduced significantly. For the 2-class model, the validation accuracy dropped from 95% to 86.5%. This shows that even though the other features individually overlap, a QSVM hyperplane defines a better decision boundary with all four features combined. Table 3 gives an overview of the supervised QSVM training model classification performance. In all cases, the test accuracy is no lower than 5% below the validation accuracy which gives some confidence that the model is not being overfitted. It should be taken into account that even though the spectrum feature-based model shows lower accuracy, it has the advantage of lesser computational load, due to only one Fourier transform being required to obtain a spectrum compared to applying-a STFT for spectrogram plot generation.

6 | ANOMALY DETECTION

In practice, it can be anticipated that a system would encounter more sea clutter data than target returns in a marine environment. Thus, the sea clutter returns can be considered as the normative form of the data. For a sensor looking for a deviation from this, it would be useful to have a model which only flags suspected anomalies. The anticipation with the unsupervised training is that the feature values in the dataset would result in the training model learning what should be the normal or expected range of the feature values that correspond to sea clutter. The feature values from the targets would then lie outside that range, and hence would be identified as *n* anomalies. The autonomous system on the vehicle could then make a decision based on this information. To learn the normative form of a highly variable property such as sea clutter is quite challenging, as discussed earlier. It is highly unlikely that such a distinct form could be found encompassing all the different sea conditions. Hence, anomaly detection techniques explored here should be restricted to the given sea-state and grazing angle. Nonetheless, the results should show the efficacy of the method, which then can be duplicated for other sea-states. In this study, OCSVM method has been evaluated. The reason for selecting OCSVM is that it does not require a large dataset to train. The dataset used in this study is not considered as such. Also, this method allows for using feature values as training input.

6.1 | OCSVM-based anomaly detection

OCSVM is an unsupervised anomaly detection algorithm that uses high-dimension predictor spaces to construct a decision boundary [25]. To apply OCSVM, an unbalanced dataset was generated from the existing Doppler feature-based training data. This consisted of 320 sea clutter data points and 32 data points from different targets for each feature. This created an outlier proportion of 10% in the training data. The class labels were removed as this is an unsupervised learning technique. Although a Gaussian kernel is commonly used, a polynomial kernel was used on this occasion as it gave a slightly better prediction result. All the features were used during training in both cases (spectrum and spectrogram). After training, the model was then cross-validated to determine the number of anomaly observations. Figure 11 shows the histogram plots of the observations made by the cross-validated data. Here, a negative score corresponds to an anomaly. The model trained with spectrum features predicted slightly higher than 10% of the data as anomalous, whereas the spectrogram feature-based model predicts very close to 10% data to be outliers. In both cases the values are close to the actual number of anomalies fed to the training data. Figure 12 illustrates the distributions of anomalies in terms of the feature values. Different combinations of two features are shown here. Here, the zero contour region is the normative boundary, meaning that observations inside the zero contours (positive) are treated as sea clutter. The observations falling outside that boundary (negative contours) are determined as anomalies by the classifier, which in this context is the presence of a potential target. In practice, OCSVM performs this in a multi-dimensional hyperplane, but here different combinations of two features are shown to provide visualisation of the OCSVM decision boundaries. It

TABLE 2 Classification accuracy with different classification types, features and classifiers.

Classification type	Number of features	Classifier	Validation accuracy (%)
Spectrum features			
8-class	5	QSVM	**84.3**
8-class	5	LSVM	70.2
8-class	5	LD	76.6
8-class	5	KNB	70.2
8-class	3 (three central moments)	QSVM	76.6
8-class	3	LSVM	68.8
8-class	3	LD	69.5
8-class	3	KNB	63.1
2-class	5	QSVM	**88.1**
2-class	5	LSVM	66.7
2-class	5	LD	66.7
2-class	5	KNB	70.9
2-class	3	QSVM	79.4
2-class	3	LSVM	68.8
2-class	3	LD	68.1
2-class	3	KNB	77.3
Spectrogram features			
8-class	4	QSVM	**93.3**
8-class	4	LSVM	90.8
8-class	4	LD	81.6
8-class	4	KNB	89.4
8-class	2 (Doppler spread, skewness)	QSVM	89.4
8-class	2	LSVM	85.1
8-class	2	LD	71.6
8-class	2	KNB	84.4
8-class	1 (Doppler spread)	QSVM	67.4
8-class	1	LSVM	80.1
8-class	1	LD	55.3
8-class	1	KNB	80.9
2-class	4	QSVM	**95**
2-class	4	LSVM	69.5
2-class	4	LD	70.2
2-class	4	KNB	90.1
2-class	2	QSVM	92.9
2-class	2	LSVM	68.0
2-class	2	LD	72.3
2-class	2	KNB	90.1
2-class	1	QSVM	86.5
2-class	1	LSVM	67.4
2-class	1	LD	67.4
2-class	1	KNB	80.1

Note: The bold values correspond to the best validation accuracy for a given classification type.

has been observed that in all cases the anomalies (negative data points) are around 10% of the observation. This gives slightly more confidence in the importance of all the feature values during OCSVM training.

7 | CONCLUSION

Doppler feature-based machine learning techniques for automatic marine target classification and anomaly detection have been broadly explored in this study. Experimental data from three different target locations gathered with three radars operating in the millimetre wave frequency range have been used to create the training dataset. Statistical analysis such as central moments, FWHM, entropy and Doppler spread of the processed frequency domain data was done to obtain feature values for machine learning-based training. Five features from each Doppler spectrum plot and four features from spectrogram plot were extracted. Two training datasets consisting of 564 values for each feature for eight different classes were created, for both types of Doppler data. Conventional machine learning techniques were applied to the labelled datasets to generate classification models. The spectrum feature-based model showed a classification accuracy between 84% and 88%, and the spectrogram-based model an accuracy of 88%–95%. An anomaly detection technique was also explored using an unbalanced dataset with sea clutter being the normal scenario. An unsupervised OCSVM classifier-based model was produced for outlier detection. The cross-validation accuracy showed a good match with the actual percentage of contaminated data provided for training.

As far as the authors are aware, this is the first-time target classification in sea clutter with millimetre wave radar Doppler data has been reported. The results obtained in this study show promise for using radars at these frequencies for the intended application of sensing for marine autonomy. It is shown that the W-band and G-band radar returns in the Doppler domain contain useful information that can be used for target discrimination. This has initiated further study to explore the deep learning approach. Further field trials are planned to enhance the dataset not only in terms of the data volume, but also data diversity. Data at

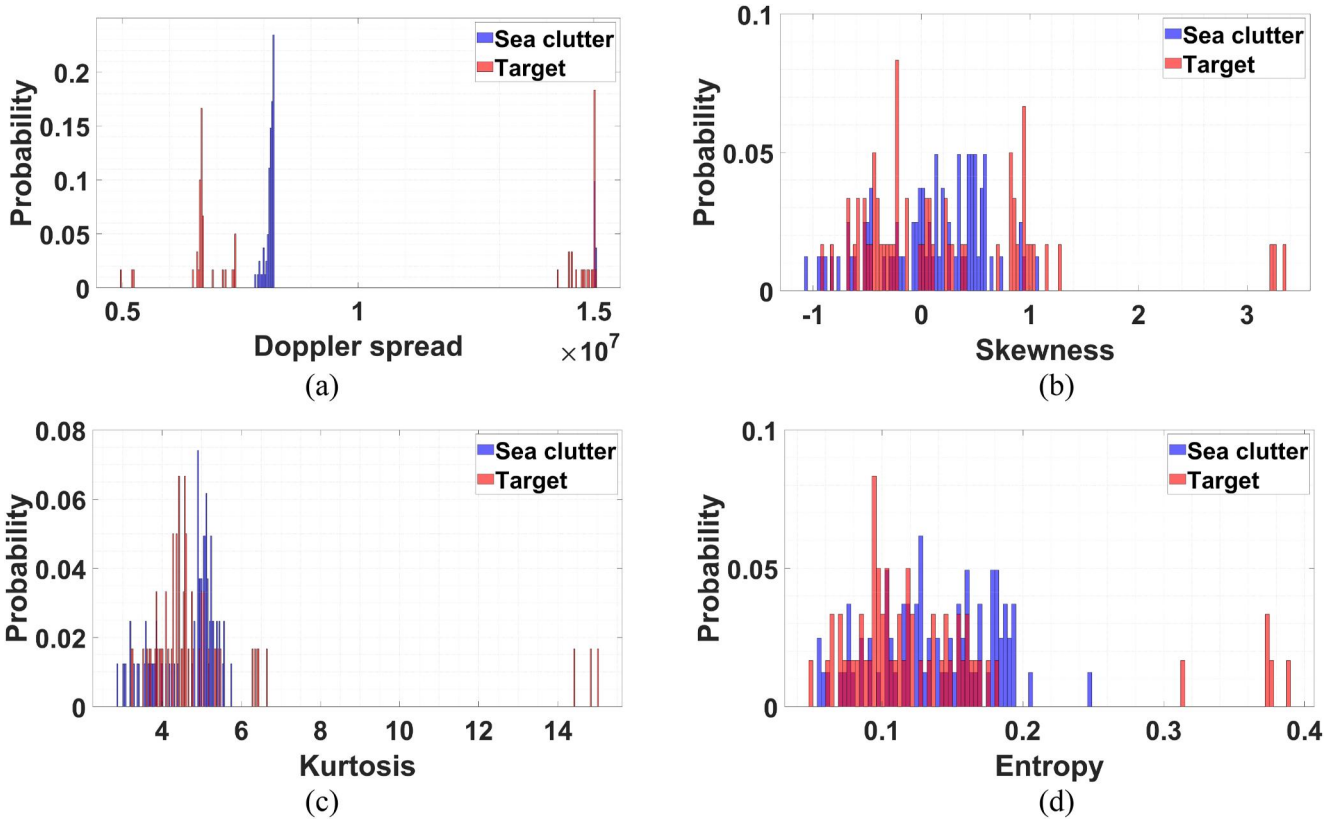

FIGURE 10 Overlaid histogram plots of the spectrogram features for sea clutter and target: (a) Doppler spread, (b) skewness, (c) kurtosis, and (d) entropy.

TABLE 3 QSVM training model performance parameters.

	8-class		2-class	
Feature source	**Validation accuracy (%)**	**Test accuracy (%)**	**Validation accuracy (%)**	**Test accuracy (%)**
Spectrum	84.3	80.1	88.1	86.8
Spectrogram	93.3	88.7	95	94.1

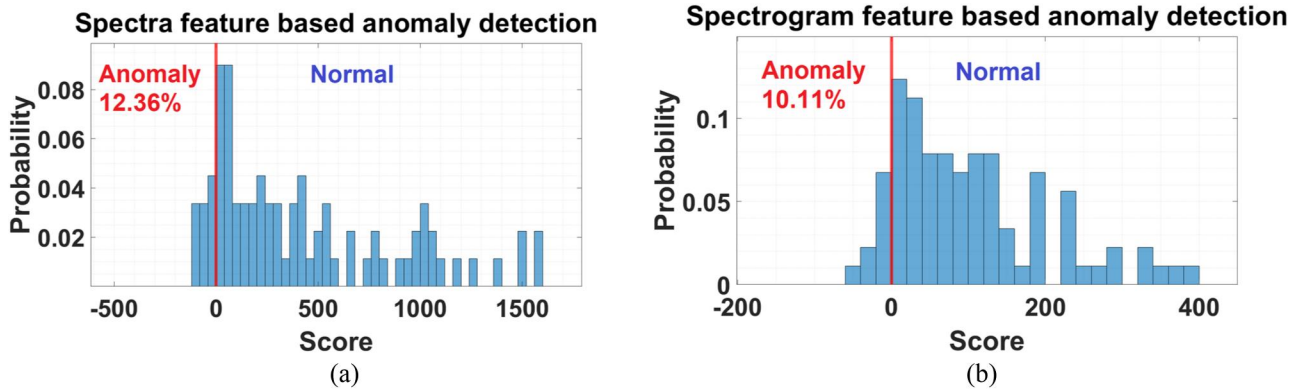

FIGURE 11 OCSVM anomaly predictor based on (a) spectrum features and (b) spectrogram features.

higher grazing angles and sea states would extend the current training dataset. This would create the option to generate similar models for those other sea conditions as well as benefitting the current model's performance. The larger dataset would also help to create a reliable deep learning-based classification framework.

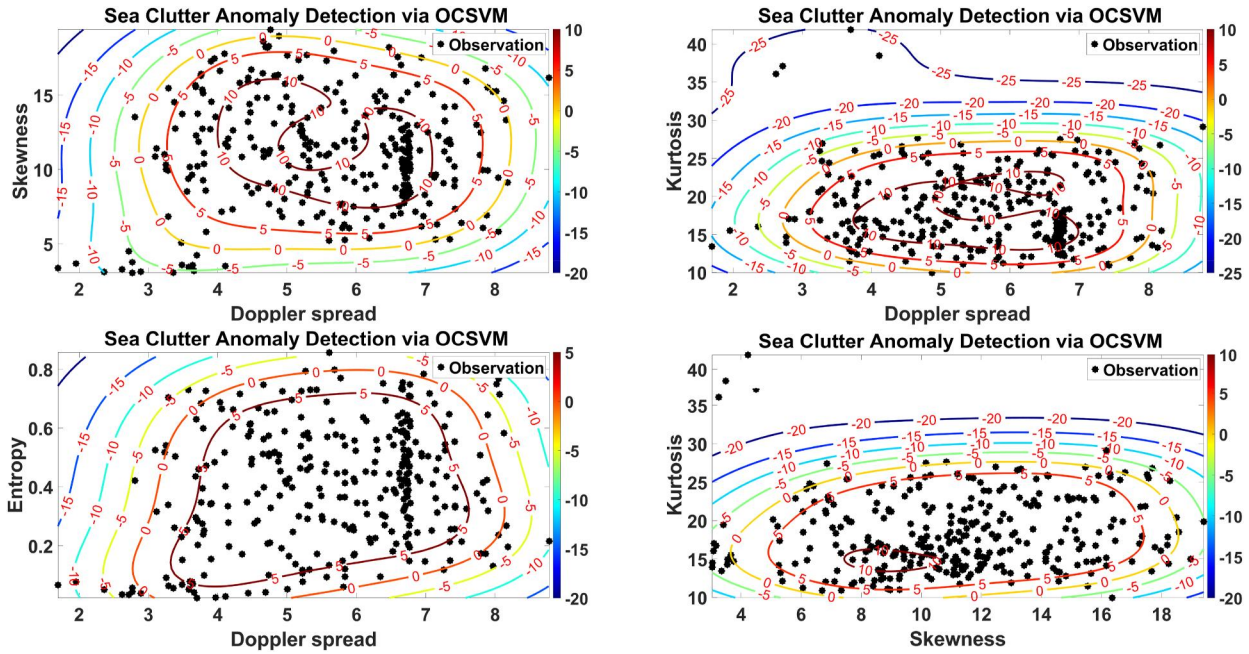

FIGURE 12 Contour plots of OCSVM anomaly detection feature values (spectrogram) where negative values correspond to anomalies.

AUTHOR CONTRIBUTIONS

Samiur Rahman: Data curation; formal analysis; investigation; methodology; validation; writing—original draft; writing—review and editing. **Aleksanteri B. Vattulainen**: Data curation; investigation; validation; visualisation; writing—review and editing. **Duncan A. Robertson**: Conceptualisation; data curation; funding acquisition; project administration; supervision; writing—review and editing.

ACKNOWLEDGEMENTS

The authors acknowledge Dr Douglas Gillespie, Ryan Milne and Simon Moss from the Sea Mammal Research Unit, St Andrews, for their support with Trial 2. Ethical approval for Trial 2 was granted by the University of St Andrews Bioethics committee. The authors also thank the staff at the Raymond Priestley Centre, Coniston, for their help over the duration of the Trial 3. This work was supported by the UK Engineering and Physical Sciences Research Council under grant EP/S032851/1.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data will be openly available in University of St Andrews public repository (Pure)- https://doi.org/10.17630/2bb28a35-11a7-4d86-8464-75288e6b66ce

ORCID

Samiur Rahman https://orcid.org/0000-0002-5477-4218
Aleksanteri B. Vattulainen https://orcid.org/0000-0003-1898-600X
Duncan A. Robertson https://orcid.org/0000-0002-4042-2772

REFERENCES

1. Raynal, A., Doerry, A.: Doppler characteristics of sea clutter. Albuquerque and Livermore. https://www.osti.gov/servlets/purl/992329/ (2010). Accessed 26 April 2023
2. Ward, K., Tough, R., Watts, S.: Sea Clutter: Scattering, the K Distribution and Radar Performance. Institution of Engineering and Technology (2013). https://doi.org/10.1049/PBRA025E
3. Watts, S., et al.: Doppler spectra of medium grazing angle sea clutter; part 1: characterisation. IET Radar, Sonar Navig. 10(1), 24–31 (2016). https://doi.org/10.1049/iet-rsn.2015.0148
4. Watts, S., et al.: Doppler spectra of medium grazing angle sea clutter; part 2: model assessment and simulation. IET Radar, Sonar Navig. 10(1), 32–42 (2016). https://doi.org/10.1049/iet-rsn.2015.0149
5. Rosenberg, L.: Parametric modeling of sea clutter Doppler spectra. IEEE Trans. Geosci. Rem. Sens. 60, 1–9 (2022). https://doi.org/10.1109/TGRS.2021.3107950
6. Wang, X., Liu, J., Liu, H.: Small target detection in sea clutter based on Doppler spectrum features. In: CIE International Conference of Radar Proceedings. Institute of Electrical and Electronics Engineers Inc. (2006). https://doi.org/10.1109/ICR.2006.343355
7. McMaster IPIX Radar. http://soma.mcmaster.ca/ipix/. Accessed 10 May 2023
8. Duk, V., Rosenberg, L., Ng, B.W.H.: Target detection in sea-clutter using stationary wavelet transforms. IEEE Trans. Aerosp. Electron. Syst. 53(3), 1136–1146 (2017). https://doi.org/10.1109/TAES.2017.2667558
9. Stacy, N.J.S., Burgess, M.P.: Ingara: the Australian airborne imaging radar system. In: International Geoscience and Remote Sensing Symposium (IGARSS), IEEE, pp. 2240–2242 (1994). https://doi.org/10.1109/IGARSS.1994.399703
10. Wen, B., Wei, Y., Lu, Z.: Sea clutter suppression and target detection algorithm of marine radar image sequence based on spatio-temporal domain joint filtering. Entropy 24(2), 250 (2022) https://doi.org/10.3390/E24020250
11. Dai, Y., et al.: Moving target detection in sea clutter based on deep learning methods. In: 2021 6th International Conference on Signal and Image Processing, ICSIP 2021, Institute of Electrical and Electronics Engineers Inc., pp. 113–117 (2021) https://doi.org/10.1109/ICSIP52628.2021.9688747

12. Wang, J., Li, S.: Maritime radar target detection in sea clutter based on CNN with dual-perspective attention. Geosci. Rem. Sens. Lett. IEEE 20, 2023–2025 (2023). https://doi.org/10.1109/LGRS.2022.3230443
13. Ningbo, L., et al.: Sea-detecting X-band radar and data acquisition program (in English). J. Radars 8(5), 656–667 (2019) https://doi.org/10.12000/JR19089
14. Xu, S., et al.: Floating small target detection in sea clutter by one-class SVM based on three detection features. In: 2019 International Applied Computational Electromagnetics Society Symposium-China, ACES 2019, Institute of Electrical and Electronics Engineers Inc. (2019) https://doi.org/10.23919/ACES48530.2019.9060796
15. Hu, J., et al.: Target detection within sea clutter: a comparative study by fractal scaling analyses. Fractals 14(3), 187–204 (2006) https://doi.org/10.1142/S0218348X06003210
16. Xu, Z., et al.: Target detection within sea clutter based on multifractal detrended fluctuation analysis. In: Advanced Engineering Forum. Trans Tech Publications Ltd., pp. 259–262 (2012). https://doi.org/10.4028/WWW.SCIENTIFIC.NET/AEF.4.259
17. Jayaprakash, A., Reddy, G.R., Prasad, N.S.S.R.K.: Small target detection within sea clutter based on fractal analysis. In: Procedia Technology. Elsevier, pp. 988–995(2016). https://doi.org/10.1016/J.PROTCY.2016.05.217
18. Richard, V.W., Kammerer, J.E., Bruce Wallace, H.: Rain backscatter measurements at millimeter wavelengths. IEEE Trans. Geosci. Rem. Sens. 26(3), 244–252 (1988). https://doi.org/10.1109/36.3027
19. Rahman, S., et al.: Millimetre wave radar signatures of sea lions. In: International Conference on Radar Systems (RADAR 2022), Edinburgh. Institution of Engineering and Technology (IET) (2023) https://doi.org/10.1049/ICP.2022.2284
20. Rahman, S., Robertson, D.A.: FAROS-E: a compact and low-cost millimeter wave surveillance radar for real time drone detection and classification. In: Proceedings International Radar Symposium. IEEE Computer Society, pp. 1–6 (2021) https://doi.org/10.23919/IRS51887.2021.9466175
21. Robertson, D.A., Brooker, G.M., Beasley, P.D.L.: Very low-phase noise, coherent 94GHz radar for micro-Doppler and vibrometry studies. In: Ranney, K.I., Doerry, A. (ed.) Proc. SPIE 9077, Radar Sensor Technology XVIII, International Society for Optics and Photonics, May 2014, p. 907719 https://doi.org/10.1117/12.2053015
22. Vattulainen, A.B., Rahman, S., Robertson, D.A.: G-band FMCW Doppler radar for sea clutter and target characterization. In: Proc. SPIE 12108, Radar Sensor Technology XXVI, SPIE, May 2022, pp. 292–300. https://doi.org/10.1117/12.2618497
23. Pincus, S.M.: Approximate entropy as a measure of system complexity. Proc. Natl. Acad. Sci. USA 88(6), 2297–2301 (1991) https://doi.org/10.1073/PNAS.88.6.2297
24. Robnik-Šikonja, M., Kononenko, I.: Theoretical and empirical analysis of ReliefF and RReliefF. Mach. Learn. 53(1–2), 23–69 (2003). https://doi.org/10.1023/A:1025667309714/METRICS
25. Cristianini, N., Shawe-Taylor, J.: An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press (2000). https://doi.org/10.1017/CBO9780511801389

How to cite this article: Rahman, S., Vattulainen, A.B., Robertson, D.A.: Machine learning-based approach for maritime target classification and anomaly detection using millimetre wave radar Doppler signatures. IET Radar Sonar Navig. 1–17 (2023). https://doi.org/10.1049/rsn2.12518