A Novel Class-Imbalanced Ship Motion Data-Based Cross-Scale Model for Sea State Estimation

Cheng, Xu; Wang, Kexin; Liu, Xiufeng; Yu, Qian; Shi, Fan; Ren, Zhengru; Chen, Shengyong

Published in:
IEEE Transactions on Intelligent Transportation Systems

Link to article, DOI:
10.1109/TITS.2023.3315674

Publication date:
2023

Document Version
Peer reviewed version

A Novel Class-Imbalanced Ship Motion Data-Based Cross-Scale Model for Sea State Estimation

Xu Cheng, *Member, IEEE*, Kexin Wang, Xiufeng Liu, Qian Yu, Fan Shi, *Member, IEEE*, Zhengru Ren, and Shengyong Chen, *Senior Member, IEEE*

***Abstract*— Sea state estimation (SSE) is significant to the development of autonomous ships, which can enhance the sustainable development of maritime transportation. Traditional model-based methods are limited by their drawbacks, such as high costs and inaccurate estimations. The deep learning model shows superior performance, but it requires that the sample quantity for each sea state should be almost the same. Since the occurrence probability of each state is different, the ships mainly work in low sea states, and the collected ship motion data for different sea states are highly imbalanced. This work proposes a novel class-imbalanced ship motion data-based cross-scale model for SSE. The model consists of three major components: a multi-scale feature learning module, a cross-scale feature learning module, and a prototype classifier module. The multi-scale and cross-scale feature learning modules are designed to learn abundant coarse and fine-level features from the ship motion data. The prototype classifier is utilized to overcome the limitation of the conventional softmax classifier to produce better estimates. Our research highlights our model's remarkable scalability and versatility with 30 publicly available datasets in time series classification, demonstrating superior performance over baseline methods in 21 cases. Notably, it outperformed ShapeNet by 5.72% and EDI by 26.3%. We further validated our model's proficiency using ship motion datasets, consistently surpassing eight state-of-the-art baselines and five class-imbalanced learning methods. Ablation and sensitivity studies, emphasize the critical role of each model component. Our findings underscore the model's robustness and its potential to advance time series classification in diverse domains.**

***Index Terms*— Autonomous ship, sustainable development, sea state estimation, class imbalance, time series classification.**

I. Introduction

THE rapid development of the marine economy presents huge pressure and impact on the environment [1], and sustainable development becomes an important issue to respond to the international call. But sustainable sea development faces many challenges, including pollution caused by ships and frequent ecological and environmental disasters. On the basis of ever-stricter environmental laws and the increasing demand for ship transportation with economic and social development, autonomous ships are more suitable for the strategic requirements of sustainable development. Nevertheless, autonomous ships are highly susceptible to sea states. Thus, a real-time and accurate estimation of sea state is needed to support ship driving, optimal path plan, and intelligent decision-making [2].

The sea state is the situation of wave and wind in the open sea for a certain location and moment [3]. A sea state is characterized by statistics, including the wave height, period, and spectrum. The sea state varies with time, as the wind and swell conditions change. Traditional techniques for identifying sea states include manual observations, wave-buoys, meteorological remote sensing satellites, and X-band radar [4]. The advantage of manual observations is that the data are highly persistent and are not dependent on external sensors, but the drawback is that they are very subjective and can be influenced by the observer's experience and biases. Wave buoys have the advantage of high reliability and comparability of the observed data, but the disadvantage is that they are weak in resistance to wind and require manual installation and placement, so they can only be used near the coast and for long-term use. Meteorological remote sensing satellites are not limited by geographical location, weather conditions, and human factors, and can cover remote geographical locations. However, its disadvantage is that it is susceptible to cloudy weather, while weather forecasts often have a time delay of several hours, which can limit their usefulness for real-time applications. Wave radars can provide accurate and real-time estimates of the sea state, but they are expensive to install and require frequent calibration, so they can only be equipped on a limited number of ships.

Due to the limitations of conventional methods, researchers have begun to consider using ship motion data to estimate the sea state. There are two main research directions in this area: model-based and model-free approaches [5], [6], [7]. The model-based approach involves building mathematical models of ship motion using domain knowledge and physical principles [8], [9], [10]. However, this approach is prone to produce incorrect estimates due to the complexity and randomness of the waves, and it requires a deep understanding of the relationship between sea states and ship motion [11], [12], [13]. In contrast, the model-free approach uses machine learning or deep learning techniques to automatically learn the

Manuscript received 12 December 2022; revised 2 May 2023 and 7 September 2023; accepted 7 September 2023. This work was supported by the National Natural Science Foundation of China under Grant 62306212, Grant 62020106004, and Grant 62272342. The Associate Editor for this article was P. Bellavista. *(Corresponding author: Fan Shi.)*

Xu Cheng, Kexin Wang, Fan Shi, and Shengyong Chen are with the School of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China (e-mail: shifan@email.tjut.edu.cn).

Xiufeng Liu is with the Department of Technology, Management, and Economics, Technical University of Denmark, 2800 Kongens Lyngby, Denmark.

Qian Yu is with the Department of Ocean Operations and Civil Engineering, Norwegian University of Science and Technology, 7034 Ålesund, Norway.

Zhengru Ren is with the Institute for Ocean Engineering, Shenzhen International Graduate School, Tsinghua University, Shenzhen 518071, China.

Digital Object Identifier 10.1109/TITS.2023.3315674

complex patterns and relationships between sea states and ship motion from the data. This approach provides a novel solution to estimate sea states without the need for extensive domain knowledge or complex mathematical models and has obtained significant attention recently [2], [14].

Model-free approaches can be grouped into two categories: shallow machine learning-based and deep learning-based. The difference between them is the way of knowledge learning. Shallow machine learning-based methods use signal analysis techniques and shallow neural networks, while deep learning-based methods extract features through their unique structure. In the literature, wavelet analysis, spectral analysis, and temporal analysis are popular machine learning-based methods for sea state estimation [4], [11]. The drawbacks of shallow machine learning-based methods are that manual feature extraction is time-consuming and labor-intensive, and it is more challenging to extract effective features if the ship motion data contains noise.

In contrast, deep learning-based methods use deep neural networks to automatically learn complex patterns and relationships from the ship motion data. These methods have been shown to produce more accurate estimates of the sea state compared to traditional methods and machine learning-based approaches [15]. The emergence of deep learning techniques, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), has enabled the development of powerful and effective models for sea state estimation (SSE). These models can learn complex features from the ship motion data without the need for manual feature engineering or domain knowledge and can produce accurate and reliable estimates of the sea state in real time.

However, there are still some challenges of using deep learning-based methods for SSE.

- Limited feature learning capability. Ship motion data typically consists of complex, fluctuating, and time-varying time series. The extraction of deep intrinsic features from raw data is a challenging task due to the presence of strong noise. Furthermore, ship motion data exhibit multi-scale characteristics under various sea states, adding complexity to the analysis. Numerous existing methodologies attempt to address these challenges by increasing the depth of networks to enhance their feature learning capabilities. However, such augmentation of neural network depth may inadvertently lead to overfitting, a situation in which the model becomes excessively specialized in the training data, resulting in sub-optimal performance.
- Imbalance ship motion data. According to [3], the likelihood of various sea states occurring differs significantly, with a higher concentration of occurrences in specific sea states, such as sea states 3 and 4. Furthermore, the probability distribution of sea state occurrence is not uniform across geographical locations. For instance, the global probability distribution contrasts with that of the North Atlantic region. It is also crucial to note that during extreme weather conditions, most vessels remain docked at port, resulting in an inability to gather data for extreme sea states. This data imbalance can adversely impact the model's performance, leading to results that are skewed toward sea states with a higher volume of available data.

To address the aforementioned challenges, this paper presents a novel deep learning model for SSE based on class-imbalanced ship motion data. To capture the multi-scale characteristics of the data, multi-scale and cross-scale feature learning modules are designed to extract abundant coarse and fine-level features from ship motion data, while facilitating information interaction among different scales. To tackle the class imbalance inherent in ship motion data, a prototype classifier is employed. The prototype classifier module is designed to overcome the limitations of conventional softmax classifiers, yielding improved estimates for imbalanced datasets.

In summary, the main contributions of this work are:

- A novel deep learning model is presented for SSE based on class-imbalanced ship motion data. The proposed model integrates a multi-scale and cross-scale feature learning module with a prototype classifier module to effectively address the feature representation and data imbalance.
- A prototype classifier module that incorporates the concept of clustering for class imbalance learning is introduced. The centroid of each cluster is learnable, and a distance-based loss function is employed to generate improved estimates tailored for class-imbalanced ship motion data.
- We conduct a comprehensive evaluation of our proposed model on a variety of public datasets and two ship motion datasets, demonstrating its good scalability, generalizability, and superiority over state-of-the-art baselines and imbalance learning methods.

The rest of this work is organized as follows: Section II introduces the review of the related work. The details of the proposed model are presented in Section III. The evaluation of the proposed model is illustrated in Section IV. Section V gives the conclusion and future work of this paper.

II. Related Work

A. Model-Based Approach

Due to the limitations of traditional sea state estimation methods, there is a research trend that considers a ship as a big sensor [6], [7], i.e., wave buoy. More and more advanced methods for estimating sea states have been developed, usually including model-based and model-free methods. Among others, Nielsen et al. [8] propose the brute-force spectral approach, which uses a simple but efficient calculation. It provides wave spectrum estimation in seconds, which can be applied to real-time, shipboard control, and decision support systems. However, only waves in the same direction can be estimated, and its application is only limited to estimating the wave components in a certain frequency band. Subsequently, Ren et al. propose a novel model-based method for SSE using Bezier surface and L1 optimization [5]. However, there are several drawbacks to this method: the Response Amplitude Operator (RAO) cannot be calculated accurately, the vessel response may not be sensitive enough to the high-frequency tail of the wave spectrum, and the method cannot be applied to ships with forward speed. Han et al. [14] propose a vessel hydrodynamic seakeeping model based on onboard vessel motion measurements. They tried to improve the RAO accuracy by modifying key parameters, but the complex sea state and environmental loads are not considered sufficiently in their experiments, which may not be feasible in field practice. Recently, Dirdal et al. [16] propose a nonlinear Parametric Time Domain (PTPD) model for shipboard sensor arrays and an unscented Kalman filter algorithm to estimate wave direction and wave number. Their method addresses the disadvantages of the array being expensive, difficult to install, and fixed in place. The method is currently limited to the study of wave patterns, such as conventional harmonics, while the Unscented Kalman Filter (UKF) operation in more complex marine environments remains an issue.

B. Machine Learning Approach

In recent years, machine learning and deep learning have been applied in many fields with the popularity of artificial intelligence. However, the efforts in sea state estimation are still very limited. Machine learning and deep learning methods are data-driven methods that do not rely on prior knowledge and are also considered model-free methods. For the machine learning approaches, Han et al. [4] propose the following two methods for sea state estimation. The first method is to build a model based on the ship motion response and has been validated with real data. According to the authors, the model will be validated in the future using data from far from the coast, and that uncertainty will be considered for practical use of the model. The second method [17] is a hybrid approach that combines machine learning and the wave buoy analogy. If the training data does not contain information about the sea state, the machine learning will fail. In this case, the wave buoy analogy will be used to estimate the sea state. However, machine learning methods not only require a large amount of training data, but also often require manual feature identification, which is a tedious task.

C. Deep Learning Approach

Deep learning methods for SSE have received increasing attention in recent years, due to the strong feature extraction capability. Among others, Cheng et al. [18] propose a sea state estimation method based on deep neural networks, namely SeaStateNet. The method consists of a recurrent neural network with a long short-term memory, a convolutional neural network, and a fast Fourier transform block. It can be applied directly to process raw time series data with minimal manual feature design and only slight data pre-processing required. In addition, they propose SSENET [15], which is a network built on a superimposed convolutional neural network block, a channel attention module, and a feature attention module. It can estimate not only the wave height but also the wave direction. In addition, they propose SpectralNet [19], a spectral-based deep learning sea state estimation model. Unlike other methods, this model combines the spectrograms of each sensor into a new image by a short-time Fourier transform. Then, a two-dimensional convolutional neural network is built as a classifier to identify the ocean state.

TABLE I

DEFINITION OF SEA STATE [3]

Sea state	Description	Wave height	World wide probability	North Atlantic probability
0	Calm (glassy)	0	—	—
1	Calm (ripples)	0-0.1	11.2486	8.3103
2	Smooth	0.1-0.5	—	—
3	Slight	0.5-1.25	31.6851	28.1996
4	Moderate	1.25-2.5	40.1944	42.0273
5	Rough	2.5-4.0	12.8005	15.4435
6	Very rough	4.0-6.0	3.0253	4.2938
7	High	6.0-9.0	0.9263	1.4968
8	Very High	9.0-14.0	0.1190	0.2263
9	Extreme	> 14.0	0.0009	0.0016

Despite the increasing interest in deep learning-driven SSE, to the best of the author's knowledge, all existing methods necessitate a class-balanced dataset for model training, and their performance declines when the data is class-imbalanced. Consequently, this study aims to develop a deep learning model for SSE that effectively addresses the challenges posed by class-imbalanced data.

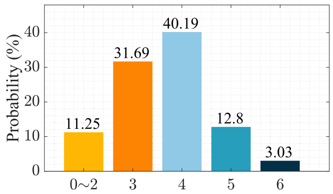
III. Class-Imbalanced Learning for Sea State Estimation

In this section, we describe the proposed model for class-imbalanced learning for sea state estimation. We first present an overview of the model, followed by a detailed description of each module.

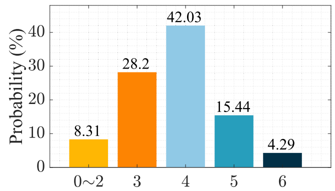
A. Problem Statement for SSE

SSE is the task of identifying the wave and wind conditions at a certain location and time on the open sea. It is an important problem for autonomous ships, as it can affect their navigation, safety, and performance. In this work, we formulate SSE as a time series classification problem, where the input is the ship motion data and the output is the sea state class. The ship motion data are collected from inertial measurement units (IMUs) installed on the ship, which measure the heave velocity, pitch angle, pitch velocity, and yaw angle of the ship. The sea state class is determined by the wave height, according to the definition in TABLE I. We consider seven sea states (0-6) in this work, as they account for more than 98% of the probability of occurrence in both worldwide and North Atlantic regions. For sea states $7 \sim 9$, the vessel is not suitable to go out to work. Additionally, we consider the first three sea states ($0 \sim 2$) to be similar, and thus we consider them as one state in this study.

Let $X \in \mathbb{R}^{N \times T}$ denote the ship motion data, where N is the number of input parameters (i.e., 4) and T is the length of the time series. Let $y \in \{0, \ldots, C\}$ denote the sea state class corresponding to X. It is important to acknowledge that the number of samples for each sea state class varies, further contributing to the complexity of the analysis. Our goal is to learn a function $f : X \rightarrow y$ that can accurately predict the sea state class from the class-imbalanced ship motion data. To achieve this goal, we propose a novel class-imbalanced ship motion data-based deep learning model for SSE, which is described in the following subsections.



(a) World wide



(b) North Atlantic

Fig. 1. Illustration of the imbalanced probability distribution of sea states for worldwide and north Atlantic.

B. Overview

Deep learning-based SSE is challenging due to the imbalanced distribution of sea states in the data, which means that some sea states are more frequent than others [20]. For example, in Fig. 1 we can find that the probability of the sea state is generally imbalanced, with the highest probability being sea state 4. This can lead to biased results that favor the majority classes (e.g., sea state 4) over the minority classes (e.g., sea state 1 or 6) [21]. Existing solutions to this problem can be categorized into data-level and algorithm-level methods [20]. Data-level methods involve preprocessing of the raw data, such as resampling or enriching the data to balance the class distribution [21], [22]. Algorithm-level methods involve modifying the training loss algorithm itself [23], [24] or tuning the model parameters during learning. However, both types of methods have some limitations such as requiring additional information about the data distribution or relying on specific hyper-parameters.

To address this challenge, we propose a novel class-imbalanced model for SSE that consists of three modules: data preprocessing, multi-scale dense neural network, and classifier. Fig. 2 shows the architecture of the proposed method. This model structure consists of four consecutive modules: data preprocessing, multi-scale feature learning (MSFL), cross-scale feature learning (CSFL), and prototype classifier (PC).

The data preprocessing module aims to reduce sensor noise present in raw ship motion data and enhance the overall data quality. This module also plays a crucial role in data normalization and segmentation into fixed-length windows, which can contribute to further noise reduction and diminished variability in the data. By applying these preprocessing techniques, the underlying structure and patterns in the dataset become more evident, facilitating more accurate and reliable modeling and analysis of ship motion data [22]. The data preprocessing in this work includes fixing outliers, correlation analysis, and data segmentation. Fixing outliers can reduce the influence on the model; correlation analysis to explore the relationship between time series from different sensors, Pearson correlation analysis is used in this study, and segmentation divides the time series with the sliding window technique. The sliding step is half of the window size, so half of the resulting data overlap, as shown in Fig. 3.

The multi-scale feature learning module is employed to extract fine-grained features from various scales of ship motion data. By capturing information at multiple resolutions, this module can effectively differentiate between similar sea states, providing a more comprehensive representation of the data. The multi-scale approach enables the model to recognize and exploit subtle distinctions in the data, ultimately leading to improved performance in SSE.

The cross-scale feature learning is specifically designed to study the interactions among different scales. By capturing and integrating information from multiple resolutions, it facilitates a more comprehensive understanding of the underlying data patterns. This module enables models to learn and exploit relationships between various scales, which can lead to enhanced performance in tasks that require recognizing complex patterns in the ship motion data.

The prototype classifier is a method designed to address the class-imbalance problem in ship motion data. By assigning a learnable prototype vector to each class, which represents its centroid in the feature space, the classifier can effectively capture the essence of each class. During the classification process, the distance between each sample and its corresponding prototype is computed, providing a basis for determining sea states.

C. Multi-Scale Feature Learning

Assuming that there are K scales in the multi-scale feature learning module and the ship motion data are denoted as $\boldsymbol{X} \in \mathbb{R}^{N\times T}$, the k^{th} scale feature representation can be represented as $\boldsymbol{X}^k \in \mathbb{R}^{N\times \frac{T}{2^{k-1}}\times F^k}$, where N is the number of input parameters, $\frac{T}{2^{k-1}}$ denotes the feature length of the k^{th} scale and F^k means the channel size.

In this work, the convolutional neural network (CNN) is used to learn local patterns along the time dimension. The kernel size is larger in the first layer and decreases in the subsequent layer. However, the flexibility of using only one CNN layer is constrained because the granularity of temporal correlations captured in the feature representation of two successive layers is sensitive to hyperparameter settings, such as kernel size and stride size. To improve CNN's flexibility in preserving underlying temporal dependencies at different time scales, we design two parallel CNN layers with kernel size 1×1 and a pooling layer with kernel size 1×2. After that, a point-wise addition operation is performed to generate the feature representations on each scale.

It should be noted that the learned multi-scale feature representations are versatile and comprehensive, allowing for the preservation of different kinds of temporal dependencies. Due to the fact that convolutional operations are carried out only along the time dimension, there is no interactivity between variables during learning processing.

D. Cross-Scale Feature Learning

First, the feature representation of each scale is fed into a CNN layer with the filter number of 128 and kernel size of 1×3. Next, we concatenate the feature representations from different scales to enable information exchange. The concatenated features are then sent to a CNN layer with the filter number of 128 and kernel size of 1×3 as well. As the CNN layers in the multi-scale feature learning module put their attention on the local area, this might lose the global dependencies in the learned feature representations.

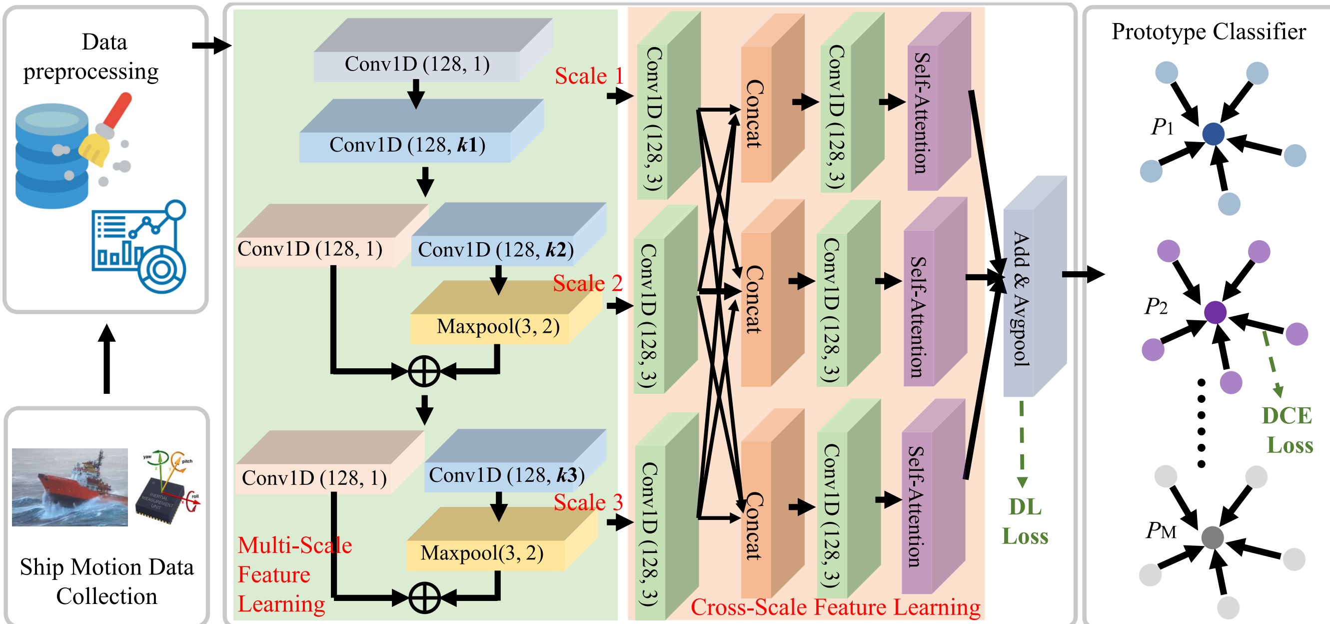


Fig. 2. Architecture of the proposed class-imbalanced model for SSE.

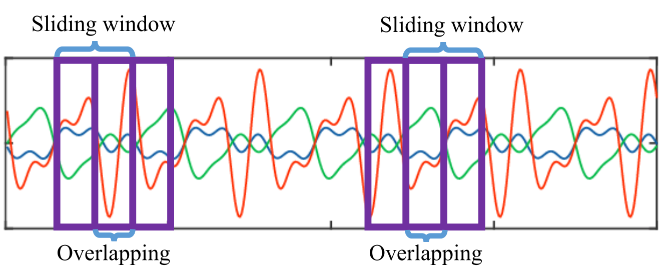


Fig. 3. Sliding windows to obtain samples.

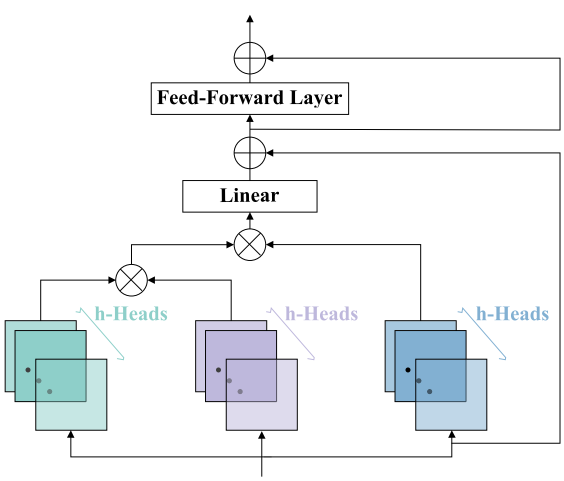


Fig. 4. Illustration of the self attention module.

To alleviate this issue, we employ a self-attention layer to capture the global temporal dependencies in the end.

The self-attention layer is inspired by the Transformer model proposed by Vaswani et al. [25], which uses a multi-head attention mechanism to compute the relevance of inputs to outputs. In our model, we use a scaled dot-product attention function, which is defined as

$$\text{Attention}(Q, K, V) = \text{softmax}(\frac{QK^T}{\sqrt{d_k}})V \tag{1}$$

where d_k is the dimension of the key vectors. The query vectors are from the previous layer, and the K and V are initialized as the same value as Q. In real-world applications, we simultaneously compute the attention on a collection of query vectors that are packed into a matrix Q, along with keys and values that are packed into matrices K and V.

We first transform Q, K, and V to d_k dimensional subspaces using linear transformations:

$$K_j = QW_j^k V_j = QW_j^v Q_j = QW_j^q \tag{2}$$

where $W_j^k, W_j^v, W_j^q \in \mathbb{R}^{d \times d_k}$ are trainable matrices. Then, we apply scaled-dot product attention to each subspace:

$$Attention(Q, K, V)_j = softmax\left(\frac{Q_j K_j^T}{\sqrt{d_k}}\right)V_j \tag{3}$$

We also use H parallel attention heads in order to jointly attend to the data from various representation sub-spaces at various points. We express the multi-head attention as follows:

$$MultiHead(Q, K, V) = Concat(\{\text{head}_j\}_{j=1}^{H})W^A \tag{4}$$

where $head_j = Attention(Q, K, V)_j$, $W^A \in \mathbb{R}^{Hd_k \times d}$ and $d_k = d/H$.

The output of the attention function is then passed through a feed-forward module, which consists of two linear transformations with a ReLU activation in between. The feed-forward module helps to learn non-linear dependencies among the features.

E. Relation to Other MSFL and CSFL Module

The idea of the multi-scale and cross-scale feature learning has been used in many domains, such as time series [26], [27] and semantic segmentation [28], [29]. In the domain of time series, many works generate multiple feature scales through pooling and downsampling techniques or other techniques such as wavelet analysis. However, to the best of my knowledge, these works only produce multi-scale features, but the interaction of different scales is not widely explored. In the domain of semantic segmentation, multi-scale feature learning aims to capture features at different levels of granularity and resolution, while cross-scale feature learning aims to fuse features across different scales and enhance the feature representation. However, our work distinguishes itself from these papers in several aspects. First, we address the problem of sea state estimation from ship motion data, which is a novel and challenging application that poses specific requirements and difficulties for the network design. Unlike semantic segmentation, where the input is an image with rich visual information, our input is a time series of ship motion data with high noise, low resolution, and class imbalance. Therefore, we need a network structure that can effectively extract and fuse features from the ship motion data and produce accurate and robust estimates of the sea state. Second, we devise a novel cross-scale feature learning module that can effectively learn and fuse coarse and fine-level features from the ship motion data. We concatenate the feature representations from different scales to enable information exchange. The concatenated features are then sent to a convolutional network for feature learning. Moreover, a self-attention mechanism is utilized to enhance feature representations. Third, we integrate the proposed multi-scale, cross-scale network structure, and a prototype classifier to form a complete end-to-end model for class imbalanced sea state estimation. The prototype classifier module consists of a set of learnable prototypes that represent the centroids of each sea state class in the feature space, which can overcome the limitations of the conventional softmax classifier and produce better estimates for imbalanced datasets.

F. Prototype Classifier

Conventionally, a softmax classifier is employed to produce the probability for each class. Given $h(x) \in \mathbb{R}^{N \times F}$ and $W \in \mathbb{R}^{F \times k}$, which represent the features learned by the two previous modules and the parameters of softmax classifier, respectively, the classification operation can be formulated as follows:

$$l = W^T h(x) = [w_1^T h(x); w_2^T h(x); \cdots ; w_k^T h(x)] \qquad (5)$$

where w_i is the weight vector of the i-th class.

Then, a cross-entropy loss function, $\mathcal{L}$, is applied to optimize the learning by minimizing the discrepancy between the estimated probability and the ground-truth values, defined as:

$$\mathcal{L} = \sum_{i}^{k} \frac{N_i}{N} \mathcal{L}(D_i) \qquad (6)$$

where N_i is the training sample size of the i-th class, N is the total sample size and D_i is the samples belonging to the i-th class.

According to the Eq. (6), if the class imbalance exists and $N_1 \leq N_2 \leq \cdots \leq N_k$, we can derive $\mathcal{L}_1 \leq \mathcal{L}_2 \leq \cdots \leq \mathcal{L}_k$. The decision boundary will be biased towards the minority class. In this work, we seek another solution, called *prototype classifier*, in which each class has an equal number of prototypes, and the estimated probability of each class can be generated based on the class matching according to the Euclidean distance.

Essentially, prototype learning can be expressed as the clustering of the learned features. For the imbalanced ship motion data, the learned features by the multi-scale feature learning module and the cross-scale feature learning module are also class imbalanced. Through the prototype learning, the class imbalanced features are clustered into several clusters (or prototypes). Note that the number of prototypes for each class is equal and the classifier established using these prototypes can produce a fair classification towards the minority classes.

As discussed above, the ship motions in different sea states may have similar behaviors. Therefore, it is necessary to maximize the margin between classes to improve learning capability. Surprisingly, we found that prototype learning can not only solve the class imbalance problem but also help optimize the learned features, which makes them more separable.

Let $H(x_j)$ represent the learned latent features, the probability of the j-th class can be calculated as follows:

$$p\left(\mathbf{x} \in \mathbf{p}_j | \mathbf{x}\right) = \frac{\mathrm{e}^{-\alpha ED\left(H(x_j), \mathbf{p}_j\right)}}{\sum_{k=1}^{M} \mathrm{e}^{-\alpha ED\left(H(x_j), \mathbf{p}_k\right)}} \qquad (7)$$

where ED is the Euclidean distance between the latent feature and the prototype $\mathbf{p}_j$, and α is the steepness parameter.

G. Loss Function

There are two loss functions used in the model: distance-based cross-entropy loss and discriminant loss. The first one is used to optimize the parameters for prototype learning while the second one is used to enhance the intra-class compactness and obtain the discriminative learned features in the latent space. They are described as follows.

1) Distance-Based Cross-Entropy Loss (DCE): DCE is used for the prototype learning, which learns prototypes iteratively from the training data in an end-to-end fashion. This loss function is defined as

$$J_{\mathrm{DCE}}(\theta, \mathbf{P}) = -\sum_{j=1}^{M} \mathbf{1}\{j = y\} \log p\left(\mathbf{x} \in \mathbf{p}_j | \mathbf{x}\right) \qquad (8)$$

where $\mathbf{1}\{j = y\}$ means that if j is equal to the ground-truth label, its value will be set to 1; otherwise, 0. From this equation, we can know that the interclass distance is enlarged by minimizing the DCE.

2) Discriminant Loss: We use the discriminant loss (DL) function to obtain discriminative features for sea state estimation. It is a cross-entropy loss function with a softmax classifier at the end of the feature extraction module. We combine it with the DCE function to form a new loss function called DCEL, which balances between interclass and intraclass distances.

The DL function is defined as

$$J_{\mathrm{DL}}(\theta, \mathbf{P}) = -\sum_{c=1}^{M} y_{o,c} \log(p_{o,c}) \tag{9}$$

where θ are the parameters of the feature extraction module, $\mathbf{P}$ are the probabilities of each sea state class, and $y_{o,c}$ are the true labels. The DCEL function is defined as

$$J_{\mathrm{DCEL}}(\theta, \mathbf{P}) = J_{\mathrm{DCE}}(\theta, \mathbf{P}) + \gamma J_{\mathrm{DL}}(\theta, \mathbf{P}) \tag{10}$$

where J_{DCE} is the DCE function that measures the distance between the true and predicted labels. The DCEL function combines both DCE and DL to optimize the feature extraction and classification tasks.

The regularization coefficient γ controls the weight of DL in the DCEL function. A larger γ means more regularization and less overfitting. A smaller γ means more fitting and more details. The optimal value of γ depends on the data and the model complexity. We can use cross-validation and grid search to find the best value for γ. Algorithm 1 summarizes the main steps of our class-imbalanced ship motion data-based deep learning model for SSE.

Algorithm 1 Training Algorithm of the Class-Imbalanced Ship Motion Data-Based Deep Learning Model for SSE

Input: Ship motion data $X \in \mathbb{R}^{N\times T}$, sea state class $y \in \{0, \ldots, 6\}$, number of scales K, number of prototypes M, steepness parameter α, regularization coefficient γ

Output: Sea state class prediction $\hat{y}$

Preprocess X
Initialize parameters of MSFL, CSFL, and PC modules
repeat
    $X_1, \ldots, X_K \leftarrow MSFL(X)$
    $H(X) \leftarrow CSFL(X_1, \ldots, X_K)$
    $p_0, \ldots, p_6 \leftarrow PC(H(X))$
    Compute J_{DCE} and J_{DL} using Eqs. (8) and (9)
    Compute J_{DCEL} using Eq. (10)
    Update parameters using gradient descent
until *convergence*
$\hat{y} \leftarrow \arg\max p_0, \ldots, p_6$

IV. Experiments

The model is implemented using the deep learning framework, Pytorch (v.1.8.0). To find the optimal parameter settings for our model, we used a grid search method with different values for the number of layers, the filter size, the learning rate, and the regularization coefficient. We evaluated each parameter combination on the validation set and selected the one that achieved the highest F1 score. All experiments are performed on a server equipped with a 32 GB Tesla v100. The number of layers in the multi-scale features is set to 3 and the filter size is set to [11, 7, 5]. The learning rate of the Adam optimizer is set to $1e-4$. The number of learning epochs is set to 500. All experiments are repeated five times with different random seeds.

A. Datasets and Evaluation Metrics

1) Benchmark Datasets: The public time series classification dataset UEA [30] was selected for performance comparison. The UEA dataset contains data for various domains, such as human activity recognition, EEG/ECG classification, and numerical recognition. The dimensions and categories of each dataset are different. Due to the complexity of the UEA dataset, it is ideal to use it to evaluate the proposed model.

2) Ship Motion Datasets: The ship motion dataset is generated using the well-known Marine Systems Simulator (MSS) toolbox [31] for simulating dynamic operations in different sea states. According to TABLE I, the probability of Worldwide and North Atlantic sea states 0-6 accounts for more than 98%. Thus, we simulate these seven sea states, the first three of which are very similar, and they are merged into one sea state. To simulate the occurring probability of each sea state, we make the sample size of each sea state consistent with the probability. Since the highest probabilities occurred in sea states 3 and 4, the sample size was higher for these two sea states. Sea state 6 has the lowest probability of occurrence, resulting in the lowest number of samples. Considering that the occurring probability is in different locations, we generate two datasets: World-wide and North Atlantic based on their historic probabilities. In addition, according to [15], we use heave velocity, pitch angle, pitch velocity, and yaw angle as inputs. The ship motion data are partitioned into training, validation, and test sets with a ratio of 70% 10%, and 20%, respectively.

3) Evaluation Metrics: For the public dataset, the widely used metric "accuracy" was used since most of the datasets are class balanced. For the ship motion dataset, since the dataset is class imbalanced, the following three metrics, **recall**, **F1**, and **Matthews correlation coefficient (MCC)**, are utilized.

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{13}$$

where TP, FP, FN, and TN represent true positive, false positive, false negative, and true negative, respectively.

B. Comparisons With the State-of-the-Art Methods on UEA Dataset

We compare the proposed model with the following eight baseline methods:

- **EDI, DTWI, and DTWD**: They are three widely used benchmark classifiers for time series classification. They were named based on how to calculate the distance. For example, ED means the Euclidean distance and DTW represents dynamic time warping [30], [32].

TABLE II
PERFORMANCE EVALUATION WITH STATE-OF-THE-ART METHODS IN 30 PUBLIC TIME SERIES CLASSIFICATION DATASETS

Dataset	EDI	DTWI	DTWD	MLSTM-FCNs	WEASEL+MUSE	Negative samples	TapNet	ShapeNet	Ours
ArticularyWordRecognition	0.970	0.980	0.987	0.973	0.990	0.987	0.987	0.987	**0.993**
AtrialFibrillation	0.267	0.267	0.220	0.267	0.333	0.133	0.333	0.400	**0.667**
BasicMotions	0.676	**1.000**	0.975	0.950	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
CharacterTrajectories	0.964	0.969	0.989	0.985	0.990	0.994	**0.997**	0.980	**0.997**
Cricket	0.944	0.986	**1.000**	0.917	**1.000**	0.986	0.958	0.986	**1.000**
DuckDuckGeese	0.275	0.550	0.600	0.675	0.575	0.675	0.575	**0.725**	0.400
EigenWorms	0.549	N/A	0.618	0.504	0.890	0.878	0.489	0.878	**0.931**
Epilepsy	0.666	0.978	0.964	0.761	1.000	0.957	0.971	**0.987**	0.986
ERing	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133	**0.30**
EthanolConcentration	0.293	0.304	0.323	0.373	**0.430**	0.236	0.323	0.312	0.331
FaceDetection	0.519	N/A	0.529	0.545	0.545	0.528	0.556	**0.602**	0.589
FingerMovements	0.550	0.520	0.530	0.580	0.490	0.540	0.530	0.580	**0.670**
HandMovementDirection	0.278	0.306	0.231	0.365	0.365	0.270	0.378	0.338	**0.460**
Handwriting	0.200	0.316	0.286	0.286	**0.605**	0.533	0.357	0.451	0.624
Heartbeat	0.619	0.658	0.717	0.663	0.727	0.737	0.751	0.756	**0.785**
InsectWingbeat	0.128	N/A	N/A	0.167	N/A	0.160	0.208	**0.250**	0.179
JapaneseVowels	0.924	0.959	0.949	0.976	0.973	**0.989**	0.965	0.984	0.987
Libras	0.833	0.894	0.870	0.856	0.878	0.867	0.85	0.856	**0.939**
LSST	0.456	0.575	0.551	0.373	0.590	0.558	0.568	0.590	**0.635**
MotorImagery	0.510	N/A	0.500	0.510	0.500	0.540	0.590	0.610	**0.640**
NATOPS	0.850	0.850	0.883	0.889	0.870	0.944	0.939	0.883	**0.988**
PEMS-SF	0.705	0.734	0.711	0.699	N/A	0.688	0.751	0.751	**0.873**
PenDigits	0.973	0.939	0.977	0.978	0.948	0.983	0.980	0.977	**0.994**
Phoneme	0.104	0.151	0.151	0.110	0.190	0.246	0.175	0.298	**0.316**
RacketSports	0.868	0.842	0.803	0.803	**0.934**	0.862	0.868	0.882	0.915
SelfRegulationSCP1	0.771	0.765	0.775	0.874	0.710	0.846	0.652	0.782	**0.901**
SelfRegulationSCP2	0.483	0.533	0.539	0.472	0.460	0.556	0.550	0.578	**0.611**
SpokenArabicDigits	0.967	0.959	0.963	0.990	0.982	0.956	0.983	0.975	**0.996**
StandWalkJump	0.200	0.333	0.200	0.067	0.333	0.400	0.400	0.533	**0.667**
UWaveGestureLibrary	0.881	0.868	0.903	0.891	**0.916**	0.884	0.894	0.906	0.919
Average accuracy	0.585	0.668	0.651	0.621	0.691	0.669	0.657	0.699	**0.739**
Wins/Ties	0	1	1	0	6	2	2	5	**21**

- **MLSTM-FCN**: It is a novel deep-learning model which combines the LSTM and FCN. In addition, the squeeze-and-excitation attention module is utilized to enhance feature learning [33].
- **WEASEL-MUSE**: It is a famous model using the bag-of-pattern. To better obtain useful features, the model can perform feature selection adaptively and can be suitable to variable window lengths [34].
- **Negative samples**: This model introduced the concept of negative samples. And these negative samples are helpful for model training. Notably, the model uses SVM as a classifier [35].
- **TapNet**: It is a recently proposed model for time series classification. It uses the distance-based classifier rather than the conventional softmax classifier [36].
- **ShapeNet**: It is a recently proposed shapelet-based time series classification model. The model can embed shapelets of different lengths into a unified space for selection, and the model is trained using a customized cluster-wise triplet loss function [37].

TABLE II shows the experimental results. As shown, our model wins the baselines on 21 out of 30 datasets. Based on the average accuracy, our method performs the best, followed by ShapeNet. More specifically, our method improves by 5.72% over ShapeNet and by 26.3% over the least EDI. Another interesting finding is that our method performs particularly well on certain datasets, for example, on the ERing, AtrialFibrillation, and HandMovementDirection datasets, where it improves 125.6%, 66.8%, and 21.7% over the second best method, respectively. Nevertheless, this does not mean that our method outperforms all baseline methods on all datasets. For example, for the DuckDuckGeese and EthanolConcentration datasets, ours ranks 8th and 3rd, respectively.

Based on these experimental results, we can conclude that the classification performance of the proposed model is the best in general, thus proving its applicability to classification tasks. The remarkable performance is mainly due to the fact that our model is equipped with a specially designed multi-scale feature learning module and a cross-scale feature learning module, which allows it to achieve a stronger learning capability. In addition, the two loss functions make the learned features more differentiable, thus improving the classification ability.

In order to assess the statistical differences between our proposed method and baselines applied to the 30 UEA datasets, we conduct a post-hoc Nemenyi test using the "AVG rank" of each method across various datasets. To better visualize these differences, we present a critical difference (CD) diagram. This diagram organizes 9 multivariate time series classifiers in ascending order based on their "AVG rank," as illustrated in Fig. 5, using a confidence level of 0.95. According to the results of the post-hoc Nemenyi test, our proposed method shows statistically significant differences in comparison to

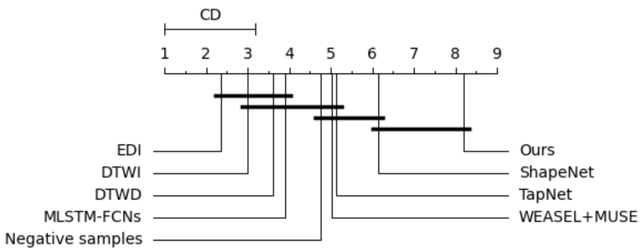


Fig. 5. CD diagram illustrates the comparative performance of nine multivariate time series classifiers across 30 UEA datasets, using a confidence level of 0.95. In the plot, a thick horizontal line indicates that there is no significant difference between a set of methodologies.

other methods, except for ShapeNet, under the specified level of statistical significance.

C. Baseline Comparison of Sea State Estimation

To fully verify the performance of the proposed model, we compare the state-of-the-art (SOTA) works in the domain of data-driven sea state estimation as well as some baseline models in the domain of time series classification. The details of the used baselines are as follows:

- **MLP** [38]: It is another baseline in the machine learning and deep learning community. In this work, we follow Miao's work but change the model with three layers. In each layer, the hidden number is set to 100.
- **CNN** [39]: It is a widely used baseline for deep learning based sea state estimation. We very the number of the filters in CNN and select the one with the best performance for comparison.
- **LSTM-FCNs**: It is a classic model for time series classification. It uses a parallel structure to learn temporal and spatial features simultaneously.
- **TapNet**: It is a recently proposed model for time series classification. It uses the distance-based classifier rather than the conventional softmax classifier.
- **SeaStateNet** [18]: It is the first deep learning based model for sea state estimation. It has three branches for learning temporal, spatial and frequency features, respectively.
- **SSENET** [15]: It is a deep densely connected network for estimating wave height and wave simultaneously.
- **SpectralSeaNet** [19]: It is a deep learning model using the spectrum image as input. Ship motion data needs to be converted to spectrum images by the fast Fourier transform for sea state estimation.
- **Hybrid model** [17]: It is a recently proposed hybrid model that combines data-driven and ship motion models. It takes the sea state estimation as a time series classification problem using ship motion data. For this comparison, we only need the data-driven model without the ship motion model.

TABLE III shows the experimental results. As shown in the table, our model achieves more competitive results than the baselines in F1 for the World Wide and North Atlantic datasets. LSTM-FCNs rank second, with the highest precision in the World Wide dataset and the highest recall in the North Atlantic dataset. CNNs and MLPs are ranked in the last two places for their limited learning ability. An interesting finding is that deep learning models specifically designed for SSE are not expected to perform well. For example, the performance of SpectralSeaNet and Hybrid is significantly worse than that of our model. These quantitative results show that our model can outperform existing SOTA methods.

TABLE III
PERFORMANCE COMPARISON WITH STATE-OF-THE-ART METHODS ON SHIP MOTION DATASETS

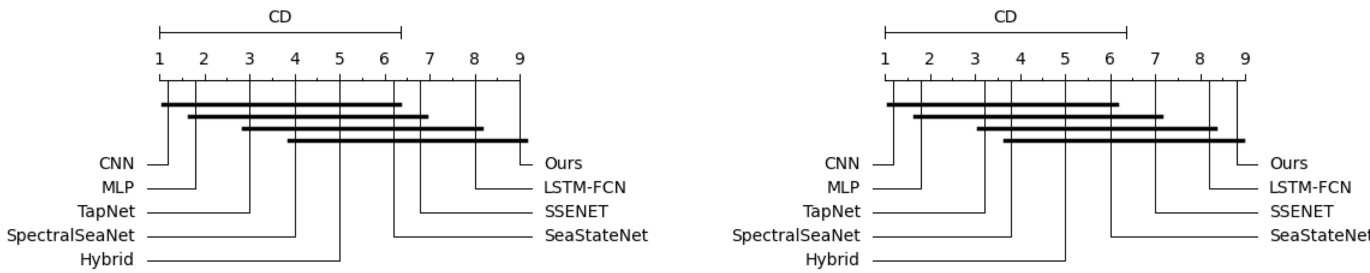
Methods	World Wide			North Atlantic		
	P	R	F1	P	R	F1
MLP	66.18	58.01	59.43	68.95	58.97	60.87
CNN	72.05	55.85	57.94	71.72	55.72	59.77
LSTM-FCNs	**80.16**	71.98	74.31	75.66	**75.47**	75.09
TapNet	72.48	62.14	64.31	73.76	63.25	65.13
SeaStateNet	77.13	68.97	71.57	76.23	70.12	72.43
SSENET	79.02	70.14	72.75	77.46	70.74	73.03
SpectralSeaNet	73.46	63.76	67.89	73.58	64.93	66.45
Hybrid	75.13	65.12	68.91	75.57	63.87	69.32
Ours	78.68	**78.22**	**78.38**	**77.01**	74.46	**75.46**

Analogous to the previous section, we also present the CD diagram (Fig.6) comparing our proposed method with STOA approaches when applied to the two sea state datasets. As depicted in Fig. 6, our proposed method demonstrates statistically significant differences when compared to CNN, MLP, and TapNet. However, there is little difference in comparison to SpectralSeaNet, Hybrid, SeaStateNet, SSENET, and LSTM-FCN across both datasets, given the specified level of statistical significance.

D. Comparisons With Class Imbalance Learning Methods

- **Focalloss**: It is a well-known loss function proposed for class imbalance learning [23]. In this work, we set the weight for each class based on its statistical information.
- **Class-balance**: It is a variant of focal loss by introducing the concept of effective number of samples [24]. We use the same setting as in [24].
- **WeightedCE**: It is a variant of the CE loss by assigning different weights to different classes. These weights can be used to overcome the class imbalance problem.
- **LDAM** (label-distribution-aware margin): It is a recent proposed loss function for class imbalance learning. LDAM can be used to replace the standard cross-entropy during training. We use the same setting as in [40].
- **BalanceSoftmax**: It is an elegant and unbiased variant of softmax for class imbalance learning. The method accommodates the label distribution shift between training and test data sets [41].

To further evaluate the proposed model, we compare it with four SOTA loss-based methods for class imbalance classification using the two ship motion datasets. For comparison, the proposed prototype-based classifier is replaced by the traditional softmax classifier. These loss functions are then employed to optimize the learning of the deep model. It is worth noting that after removing our prototype-based classifier, we also remove the two loss functions that accompany it, namely the DL loss and the DCE loss.



(a) World wide (b) North Atlantic

Fig. 6. CD diagram illustrates the comparative performance of eight methods, using a confidence level of 0.95. In the plot, a thick horizontal line indicates that there is no significant difference between a set of methodologies.

TABLE IV

PERFORMANCE COMPARISON WITH STATE-OF-THE-ART METHODS OF IMBALANCED LEARNING ON SHIP MOTION DATASETS

	World Wide			North Atlantic		
	P	R	F1	P	R	F1
FocalLoss	60.17	69.82	62.23	62.94	70.94	65.16
LDAMLoss	75.83	74.34	74.81	76.21	73.87	74.74
WeightedCE	75.67	74.38	74.91	68.72	73.34	70.58
ClassBalanced	60.23	69.26	61.00	63.31	69.79	63.57
BalancedSoftmax	**80.20**	72.06	73.86	73.53	65.49	67.71
Ours	78.68	**78.22**	**78.38**	**77.01**	74.46	**75.46**

TABLE IV presents the results, which shows that our model performs better in general. The results demonstrate the applicability of our prototype classifier for class imbalance classification. LDAMLoss ranks second, which shows good performance for class imbalance classification. Surprisingly, the ClassBalanced method does not achieve the performance as expected. The reason might be that it has many hyperparameters for tuning.

E. Ablation and Sensitivity Analysis

To evaluate the validity of each component in our proposed model, we create three variants as follows:

- **SA**: The self-attention module is removed, but the other modules are kept.
- **CS**: The cross-scale feature learning module is not used, but outputs of different scale are simply concatenated. Note that when the CS module is removed, SA will be removed together.
- **PC**: The prototype classifier is replaced by the softmax classifier.

The original model is called '**ORG**' in this section. The performance of these four models is presented in Fig. 7. We can observe that there is a large performance drop when the cross-scale feature learning module is removed (see CS bar). This confirms the importance of the cross-scale learning module and its self-attention module, which make a significant contribution to the performance of the proposed model. In addition, according to the results, the SA module is more important than the prototype classifier in the World Wide dataset, while only a slight difference is observed for the North Atlantic dataset.

To delve deeper into the distinctions between each class, particularly in terms of performance when dealing with smaller training samples, we present visualized results in Fig. 8. These results clearly illustrate a notable decrease in performance for sea states characterized by limited training samples. Additionally, it's evident from the results that the prototype classifier contributes to an enhancement in performance.

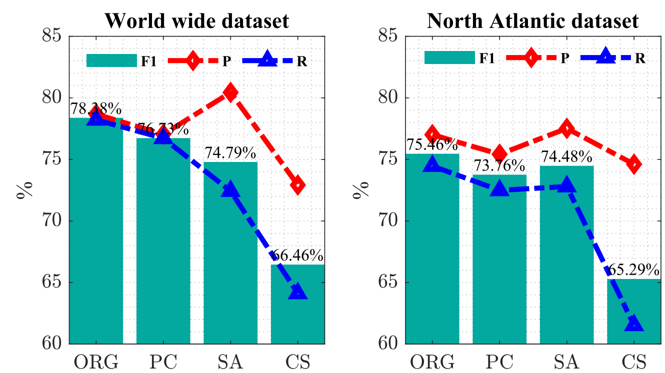


Fig. 7. Ablation study.

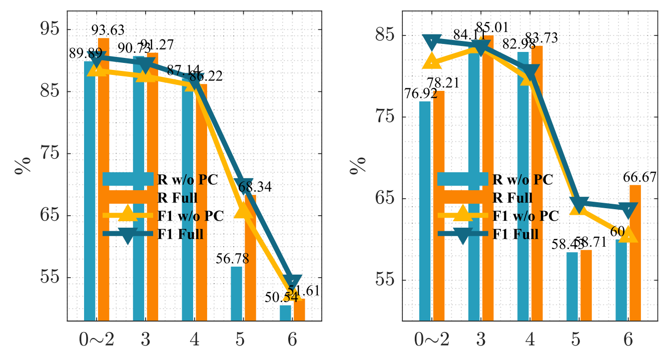


Fig. 8. Accuracy of each sea state with/without prototype classifier.

To evaluate the impact of the number of scales on the proposed model, we perform a sensitivity analysis by varying its value. The results are illustrated in Fig. 9. From the results, we can see that the model performance increases with the increase in the number of scales for both datasets. The highest performance is obtained when the scale number is set to 3 for both datasets.

F. Discussion

This work proposed a deep learning model for SSE with the aim of directly learning the imbalanced ship motion

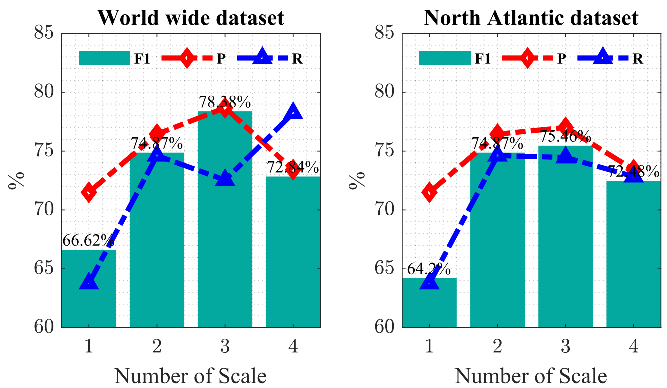


Fig. 9. Sensitivity analysis.

data. To the best of our knowledge, this work is the first deep-learning model to address the class imbalance problem for ship motion data for sea state estimation. The comparison with 30 public datasets shows the scalability and generality of the proposed model.

In this paper, two key experimental settings were used for sea state estimation using ship motion datasets. First, we only used four parameters for the modeling according to the previous work [15]. Second, to test the performance of class imbalance learning models, we generate the smallest quantity for each sea state. For example, for sea state 6, there are only around 200 samples in both World Wide and North Atlantic datasets. Under such settings, we compared our model with several SOTA baselines and class-imbalanced learning methods. Since our model is specifically designed to estimate sea states using imbalanced data, while other methods are not designed for this specific purpose. This is the main reason for their inferior performance. Compared with these class-imbalanced learning methods, our model is more like a "hard" operation on features than a "soft" operation on the adjustment of model learning during the training processing. Thus, our model is more effective than these learning-based methods.

In this study, although the performance of our model is the best, close to 80%, this accuracy is still relatively low. There is still some room for improvement before practical use. For example, the model can be further improved by means of data augmentation for rare sea states. The overlapping between different sliding windows can be increased, such as sea state 6, which is much less than sea state 3. In this way, the sample size will become closer. This will be left for our future work.

Another limitation that needs to be addressed for this work is how to handle incomplete ship motion data for SSE. Ship motion data can be very sparse due to a lot of unobservable data, which can affect the performance of our model. In this paper, we assume that the ship motion data are complete and do not contain missing values. However, this may not be realistic in some scenarios, and we should consider how to deal with incomplete data for SSE. One possible way to deal with incomplete data is to use imputation methods to fill in the missing values before applying our model. Another possible way is to use latent factor models to directly learn from the incomplete data without imputation. Latent factor models can use low-rank matrix factorization or tensor factorization to decompose incomplete data into a product of latent factors that capture the hidden structure and features of the data. We refer the reader to [42], [43], [44], [45], [46], and [47] for some recent studies that use latent factor models for incomplete data analysis. We plan to explore these methods and extend our model for SSE using incomplete ship motion data in our future work.

V. Conclusion

Today the world economy heavily relies on maritime transportation, especially during the COVID-19 epidemic. Economic development has stagnated due to the epidemic, so autonomous ships are an effective solution to the crisis. To develop autonomous ships, it is necessary, for example, to determine the sea state. In the real world, ships mainly work in low sea state, which leads to the absence of collecting high sea state data. This challenge limits the learning capability of deep learning models for sea state estimation. In this work, we proposed a novel deep learning model based on class-imbalanced ship motion data for SSE. The model learns multi-scale features on the ship motion data and then uses the obtained features to build a distance-based classifier, which can effectively improve classification performance compared to the conventional softmax classifier. A thorough evaluation of the proposed model has been conducted, and the experimental results show the model's effectiveness and superiority over baseline methods.

References

[1] Y.-W. Du and Y.-C. Wang, "Evaluation of marine ranching resources and environmental carrying capacity from the pressure-and-support perspective: A case study of Yantai," *Ecol. Indicators*, vol. 126, Jul. 2021, Art. no. 107688.

[2] X. Cheng, G. Li, P. Han, R. Skulstad, S. Chen, and H. Zhang, "Data-driven modeling for transferable sea state estimation between marine systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 2561–2571, Mar. 2022.

[3] T. I. Fossen, *Handbook of Marine Craft Hydrodynamics and Motion Control*. Hoboken, NJ, USA: Wiley, 2011.

[4] P. Han, G. Li, S. Skjong, B. Wu, and H. Zhang, "Data-driven sea state estimation for vessels using multi-domain features from motion responses," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2021, pp. 2120–2126.

[5] Z. Ren, X. Han, A. S. Verma, J. A. Dirdal, and R. Skjetne, "Sea state estimation based on vessel motion responses: Improved smoothness and robustness using Bézier surface and L1 optimization," *Mar. Struct.*, vol. 76, Mar. 2021, Art. no. 102904.

[6] M. Mittendorf, U. D. Nielsen, H. B. Bingham, and G. Storhaug, "Sea state identification using machine learning—A comparative study based on in-service data from a container vessel," *Mar. Struct.*, vol. 85, Sep. 2022, Art. no. 103274.

[7] H. Majidian, L. Wang, and H. Enshaei, "Part. B: A review of the real-time sea-state estimation, using wave buoy analogy; a decuple benchmark and future outlook," *Ocean Eng.*, vol. 266, Dec. 2022, Art. no. 111020.

[8] U. D. Nielsen, A. H. Brodtkorb, and A. J. Sørensen, "A brute-force spectral approach for wave estimation using measured vessel motions," *Mar. Struct.*, vol. 60, pp. 101–121, Jul. 2018.

[9] A. H. Brodtkorb, U. D. Nielsen, and A. J. Sørensen, "Sea state estimation using vessel response in dynamic positioning," *Appl. Ocean Res.*, vol. 70, pp. 76–86, Jan. 2018.

[10] U. D. Nielsen, A. H. Brodtkorb, and A. J. Sørensen, "Sea state estimation using multiple ships simultaneously as sailing wave buoys," *Appl. Ocean Res.*, vol. 83, pp. 65–76, Feb. 2019.

[11] F. Tu, S. S. Ge, Y. S. Choo, and C. C. Hang, "Sea state identification based on vessel motion response learning via multi-layer classifiers," *Ocean Eng.*, vol. 147, pp. 318–332, Jan. 2018.

[12] Y. Shu et al., "Path planning for ships assisted by the icebreaker in ice-covered waters in the northern sea route based on optimal control," *Ocean Eng.*, vol. 267, Jan. 2023, Art. no. 113182.

[13] X. Chen, X. Wu, D. K. Prasad, B. Wu, O. Postolache, and Y. Yang, "Pixel-wise ship identification from maritime images via a semantic segmentation model," *IEEE Sensors J.*, vol. 22, no. 18, pp. 18180–18191, Sep. 2022.

[14] X. Han, B. J. Leira, and S. Sævik, "Vessel hydrodynamic model tuning by discrete Bayesian updating using simulated onboard sensor data," *Ocean Eng.*, vol. 220, Jan. 2021, Art. no. 108407.

[15] X. Cheng, G. Li, A. L. Ellefsen, S. Chen, H. P. Hildre, and H. Zhang, "A novel densely connected convolutional neural network for sea-state estimation using ship motion data," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 5984–5993, Sep. 2020.

[16] J. A. Dirdal, R. Skjetne, J. Roháč, and T. I. Fossen, "Online wave direction and wave number estimation from surface vessel motions using distributed inertial measurement arrays and phase-time-path-differences," *Ocean Eng.*, vol. 249, Apr. 2022, Art. no. 110760.

[17] P. Han, G. Li, X. Cheng, S. Skjong, and H. Zhang, "An uncertainty-aware hybrid approach for sea state estimation using ship motion responses," *IEEE Trans. Ind. Informat.*, vol. 18, no. 2, pp. 891–900, Feb. 2022.

[18] X. Cheng, G. Li, R. Skulstad, S. Chen, H. P. Hildre, and H. Zhang, "Modeling and analysis of motion data from dynamically positioned vessels for sea state estimation," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, May 2019, pp. 6644–6650.

[19] X. Cheng, G. Li, R. Skulstad, H. Zhang, and S. Chen, "SpectralSeaNet: Spectrogram and convolutional network-based sea state estimation," in *Proc. 46th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2020, pp. 5069–5074.

[20] L. Wang, S. Xu, X. Wang, and Q. Zhu, "Addressing class imbalance in federated learning," 2020, *arXiv:2008.06217*.

[21] Y. Xie, M. Qiu, H. Zhang, L. Peng, and Z. Chen, "Gaussian distribution based oversampling for imbalanced data classification," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 2, pp. 667–679, Feb. 2022.

[22] C.-L. Liu and P.-Y. Hsieh, "Model-based synthetic sampling for imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 8, pp. 1543–1556, Aug. 2020.

[23] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2999–3007.

[24] Y. Cui, M. Jia, T.-Y. Lin, Y. Song, and S. Belongie, "Class-balanced loss based on effective number of samples," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 9260–9269.

[25] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 6000–6010.

[26] Z. Cui, W. Chen, and Y. Chen, "Multi-scale convolutional neural networks for time series classification," 2016, *arXiv:1603.06995*.

[27] J. Wang, Z. Wang, J. Li, and J. Wu, "Multilevel wavelet decomposition network for interpretable time series analysis," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2018, pp. 2437–2446.

[28] T. Pissas, C. S. Ravasio, L. D. Cruz, and C. Bergeles, "Multi-scale and cross-scale contrastive learning for semantic segmentation," in *Proc. 17th Eur. Conf. Comput. Vis. (ECCV)*, Tel Aviv, Israel. Cham, Switzerland: Springer, Oct. 2022, pp. 413–429.

[29] L. Huan, X. Zheng, S. Tang, and J. Gong, "Learning deep cross-scale feature propagation for indoor semantic segmentation," *ISPRS J. Photogramm. Remote Sens.*, vol. 176, pp. 42–53, Jun. 2021.

[30] A. Bagnall et al., "The UEA multivariate time series classification archive, 2018," 2018, *arXiv:1811.00075*.

[31] T. Perez and T. I. Fossen, "A MATLAB toolbox for parametric identification of radiation-force models of ships and offshore structures," *Model., Identificat. Control, Norwegian Res. Bull.*, vol. 30, no. 1, pp. 1–15, 2009.

[32] M. Shokoohi-Yekta, J. Wang, and E. Keogh, "On the non-trivial generalization of dynamic time warping to the multi-dimensional case," in *Proc. SIAM Int. Conf. Data Mining*, Jun. 2015, pp. 289–297.

[33] F. Karim, S. Majumdar, H. Darabi, and S. Harford, "Multivariate LSTM-FCNs for time series classification," *Neural Netw.*, vol. 116, pp. 237–245, Aug. 2019.

[34] P. Schäfer and U. Leser, "Multivariate time series classification with WEASEL+MUSE," 2017, *arXiv:1711.11343*.

[35] J.-Y. Franceschi, A. Dieuleveut, and M. Jaggi, "Unsupervised scalable representation learning for multivariate time series," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 4650–4661.

[36] X. Zhang, Y. Gao, J. Lin, and C.-T. Lu, "TapNet: Multivariate time series classification with attentional prototypical network," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, no. 4, pp. 6845–6852.

[37] G. Li, B. Choi, J. Xu, S. S. Bhowmick, K.-P. Chun, and G. L.-H. Wong, "ShapeNet: A shapelet-neural network approach for multivariate time series classification," in *Proc. AAAI Conf. Artif. Intell.*, 2021, vol. 35, no. 9, pp. 8375–8383.

[38] H. Miao, Y. Guo, G. Zhong, B. Liu, and G. Wang, "A novel model of estimating sea state bias based on multi-layer neural network and multi-source altimeter data," *Eur. J. Remote Sens.*, vol. 51, no. 1, pp. 616–626, Jan. 2018.

[39] J. N. Alfsen, "IMU-based sea state estimation using convolutional neural networks for DP vessels," M.S. thesis, NTNU, Trondheim, Norway, 2020.

[40] K. Cao, C. Wei, A. Gaidon, N. Arechiga, and T. Ma, "Learning imbalanced datasets with label-distribution-aware margin loss," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 1567–1578.

[41] J. Ren et al., "Balanced meta-softmax for long-tailed visual recognition," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 4175–4186.

[42] M. Chen, C. He, and X. Luo, "MNL: A highly-efficient model for large-scale dynamic weighted directed network representation," *IEEE Trans. Big Data*, vol. 9, no. 3, pp. 889–903, Jun. 2023.

[43] Y. Song, M. Li, Z. Zhu, G. Yang, and X. Luo, "Nonnegative latent factor analysis-incorporated and feature-weighted fuzzy double c-means clustering for incomplete data," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 10, pp. 4165–4176, Oct. 2022.

[44] M. Shang, X. Luo, Z. Liu, J. Chen, Y. Yuan, and M. Zhou, "Randomized latent factor model for high-dimensional and sparse matrices from industrial applications," *IEEE/CAA J. Autom. Sinica*, vol. 6, no. 1, pp. 131–141, Jan. 2019.

[45] D. Wu and X. Luo, "Robust latent factor analysis for precise representation of high-dimensional and sparse data," *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 4, pp. 796–805, Apr. 2021.

[46] J. Chen, R. Wang, D. Wu, and X. Luo, "A differential evolution-enhanced position-transitional approach to latent factor analysis," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 7, no. 2, pp. 389–401, Apr. 2023.

[47] D. Wu, X. Luo, Y. He, and M. Zhou, "A prediction-sampling-based multilayer-structured latent factor model for accurate representation to high-dimensional and sparse data," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Sep. 9, 2022, doi: 10.1109/TNNLS.2022.3200009.

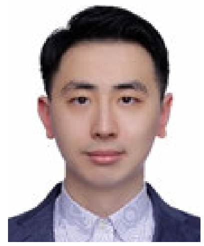
Xu Cheng (Member, IEEE) received the Ph.D. degree in engineering from the Department of Ocean Operations and Civil Engineering, Intelligent Systems Laboratory, Norwegian University of Science and Technology (NTNU), Ålesund, Norway, in June 2020. From June 2020 to March 2022, he was a Post-Doctoral Fellow and a Researcher with the Department of Manufacturing and Civil Engineering, Gjøvik, Norway. From April 2022 to May 2023, he worked at Smart Innovation Norway as a permanent researcher. From June 2023, he worked for the Tianjin University of Technology as a full Professor. He has applied for and coordinated more than five projects supported by the Norwegian Research Council (NFR), the EU, and industry. He has published more than 30 papers as the first and coauthor in his research interests, including data analysis and artificial intelligence in maritime operations, time series analysis, and predictive maintenance of wind turbines.

Kexin Wang received the B.Eng. degree in computer science and technology from the Tianjin University of Commerce, Tianjin, China, in 2021. She is currently pursuing the M.Sc. degree with the Tianjin University of Technology. Her research interests include data analysis, time series analysis, and predictive maintenance.

Xiufeng Liu received the Ph.D. degree in computer science from Aalborg University, Denmark, in 2012. He was a Post-Doctoral Researcher with the University of Waterloo and a Research Scientist with IBM, Canada, from 2013 to 2014. He is currently a Senior Researcher with the Department of Technology, Management and Economics, Technical University of Denmark. His research interests include smart meter data analysis, data warehousing, energy informatics, and big data.

Qian Yu received the B.Eng. degree in naval architecture and ocean engineering from the Jiangsu University of Science and Technology, Jiangsu, China, in 2011, and the master's degree in ship design from the Ocean Operations and Civil Engineering Department, Norwegian University of Science and Technology, Norway, in 2018. Her research interests include ship structural analyses, CFD, and data analysis.

Fan Shi (Member, IEEE) received the Ph.D. degree from Nankai University, Tianjin, China, in 2012. From June 2018 to August 2019, he was a Research Scholar with West Virginia University. He is currently a Professor with the School of Computer Science and Engineering, Tianjin University of Technology, Tianjin. His research interests include machine vision, pattern recognition, and optics.



Zhengru Ren received the B.Eng. degree in ocean engineering from the Dalian University of Technology, Dalian, China, in 2012, and the M.Sc. and Ph.D. degrees in marine technology from the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 2015 and 2019, respectively. From 2019 to 2021, he was a Post-Doctoral Research Fellow with the Centre for Research-Based Innovation on Marine Operations (SFI MOVE), Department of Marine Technology, NTNU. Since 2022, he has been an Assistant Professor with the Institute for Ocean Engineering, Tsinghua University, China. He has published more than 50 scientific papers in international journals and conferences and received 13 granted patents. His research interests include nonlinear control theory, sensor fusion, intelligent and digitalized ocean engineering, and offshore wind energy. In 2021, he received the National Science Foundation for Excellent Young Scholars (Overseas).

Shengyong Chen (Senior Member, IEEE) received the Ph.D. degree in computer vision from the City University of Hong Kong. From 2006 to 2007, he received a fellowship from the Alexander von Humboldt Foundation of Germany and was with the University of Hamburg, Germany. From 2008 to 2012, he was a Visiting Professor with Imperial College London and the University of Cambridge, U.K. He is currently a Full Professor with the Tianjin University of Technology and the Director of the Engineering Research Center of Learning-Based Intelligent System (Ministry of Education). He has been conducting research on vision sensors for robotics for more than 20 years. He has published over 300 scientific papers in international journals and conferences, including 80 papers in IEEE TRANSACTIONS. He also published more than ten books in the past years and applied more than 100 patents. He received the National Outstanding Youth Foundation Award of NSFC. He organized about 20 international conferences and serves as an Associate Editor for three international journals, such as IEEE TRANSACTIONS ON CYBERNETICS.