SAFENESS: A Semi-Supervised Transfer Learning Approach for Sea State Estimation Using Ship Motion Data

Xu Cheng, *Member, IEEE,* Guoyuan Li, *Senior Member, IEEE,* Robert Skulstad, and Houxiang Zhang, *Senior Member, IEEE*

***Abstract*—Autonomous vessels have been identified as a promising innovation in advancing marine transportation, providing an effective means to mitigate the risk of accidents, pollution incidents, and carbon dioxide emissions. Accurate sea state estimation (SSE) plays a critical role in facilitating onboard decision-making and optimizing operational efficiency for autonomous ships. Traditional SSE approaches relying on external sensors, such as wave buoys and wave radars, are limited by cost considerations. Model-based methods are highly relying on the understanding of human knowledge to ships. Data-driven models also provide promising solutions, but their generalization is low. To address this challenge, a semi-supervised transfer learning approach for SSE (SAFENESS) is proposed. The model is trained using sufficient data in the source ship and limited data from the target ship and finally applied to the target ship. A data alignment algorithm is utilized to use the limited data of the target ship. To enhance the learning capability of the framework, two attention mechanisms are proposed, and a multi-class adversarial discriminator is introduced that can align the distributions of different domains. The effectiveness of our approach is validated through comprehensive comparisons with eleven established transfer learning methods, demonstrating the superiority of our model. The competitiveness of the proposed attention modules is verified by comparing them with state-of-the-art attention modules. The significance of each component and the influence of key parameters have been thoroughly explored in the ablation and sensitivity analysis. Our method has potential applications in maritime safety, navigation, and operation optimization.**

***Index Terms*—Sea state estimation, autonomous ship, time series classification, transfer learning.**

I. Introduction

Autonomous ships represent a new class of marine vessels capable of operating without an onboard human crew [1]. Equipped with advanced navigation and control systems, these ships autonomously perceive their surroundings, plan routes, and execute actions [1]. The potential benefits of autonomous ships in the realm of marine transportation are extensive, including a reduction in accident and spill risks, decreased emissions and fuel consumption, increased cargo capacity, and enhanced operational flexibility [2], [3]. Furthermore, autonomous ships have the potential to improve the working conditions and welfare of seafarers. Fig. 1 illustrates the prototype autonomous ship of the Norwegian University of Science and Technology (NTNU).

The evolution of autonomous ships confronts a multitude of challenges. A foundational hurdle involves the accurate perception of the sea state [4]. In oceanography, the sea state is characterized by the overall wind, wave, and swell conditions at a specific open-sea location [5]. Four prevalent approaches exist for sea state estimation (SSE): wave-buoys, weather forecasts, meteorological satellites, and wave radars [6], [7], [8]. Each method, however, has its constraints. Wave buoys are limited by their placement locations, and weather forecasts are less effective due to delays in relaying critical information. Meteorological satellites often grapple with cloud interference. Wave radar, while satisfying real-time estimation needs for onsite sea states, is a costly solution requiring regular calibration [9].

At present, almost all ships are equipped with motion sensors, which allow us to conceptualize the ship as a large wave buoy, its motion response acting as an indicator of sea conditions [10]. In effect, the ship is inherently outfitted with a system for SSE [11]. SSE methodologies based on ship motion responses are typically classified as either model-based or model-free [12], [13].

Model-based approaches presuppose the existence of a response function to map the relationship between the sea state and ship motion. However, a significant challenge inherent in model-based methods is the necessity for accurate response function creation, which requires substantial human knowledge [9]. In contrast, in the era of big data, model-free methods are rapidly proliferating and exhibit substantial promise for SSE [9], [6]. These methods extract the relationship between the sea state and ship motion directly from the collected data, considerably improving estimation accuracy and eliminating the necessity for prior knowledge about the ship [14].

However, several significant challenges persist in estimating the sea state using model-free methods.

- **Ship type dependency**: SSE models, relying on ship motion data, inherently vary based on ship types. Due to distinct hydrodynamic characteristics among different vessels, models derived from specific ship types are not interchangeable. Consequently, applying a model-free SSE method trained on one ship type to another is impractical, which becomes particularly problematic when attempting to model newer ships with limited historical data from older, well-documented vessels.

Xu Cheng, Robert Skulstad, Guoyuan Li and Houxiang Zhang are with the Department of Ocean Operations and Civil Engineering, Norwegian University of Science and Technology, Aalesund, 6009 Norway.

Houxiang Zhang is the corresponding author (email: hozh@ntnu.no).

Fig. 1. Illustration of the prototype autonomous ship in the NTNU [15].

- **Ship load dependency**: A ship's performance in different sea conditions is contingent on its load. For example, a ship may be fully loaded when departing but empty upon arrival. Thus, SSE models constructed from ship motion data under various load conditions may not generalize well. Models trained for one load scenario may exhibit subpar performance when applied to a different load condition, necessitating improvements in model-free approaches to ensure applicability across varying loads.
- **Data availability dependency**: Model-free methods heavily rely on high-quality training data, which poses challenges in gathering comprehensive data covering a wide range of sea states, especially when aiming to amass sufficient volumes for training high-performance, data-driven models in practical settings. Additionally, maintaining consistency in data distribution between training and testing datasets is crucial for ensuring a model's effectiveness in real-world applications. Real-world data collection often falls short in terms of both quantity and quality. Although simulators or digital twins can generate diverse data, models trained exclusively on such simulated data may suffer performance degradation when applied to physical ships. Therefore, it becomes imperative to establish a robust connection between the virtual digital twin and the physical ship to effectively bridge the simulation-reality gap.

While addressing the aforementioned challenges, transfer learning emerges as a promising solution. Cheng et al. pioneered the development of the first SSE model using transfer learning techniques [8]. However, their approach relied on a convolutional neural network (CNN)-based Siamese network for similarity learning, yielding competitive results but introducing specific limitations: 1) Computational Complexity: The model's training process is computationally demanding and time-consuming. During training, their network computes distances between pairs of inputs, resulting in a quadratic increase in computational complexity as the dataset size grows. 2) Assumption of Domain Similarity: Cheng et al.'s model operates under the assumption that the source and target domains exhibit sufficient similarity. If the data distributions significantly differ between these domains, the transferred knowledge may not yield substantial benefits, potentially leading to a degradation in model performance.

This work extends the efforts to tackle the challenge of poor generalization exhibited by data-driven models, building upon the groundwork laid by Cheng et al. [8]. Our focus is on constructing a data-driven SSE model capable of seamless transfer between various ship types and consistent performance across different loading levels of the same ship. More precisely, we endeavor to develop an SSE model utilizing ample source ship data alongside limited target ship data. In this pursuit, we frame the SSE problem as a time series classification task, aligning with prior research in the field [9], [6], [8].

The main contributions of this work are as follows:

- This research presents a novel **S**emi-supervised tr**A**nsFer l**E**ar**N**ing approach for s**E**a **S**tate e**S**timation (**SAFENESS**). This innovative approach facilitates the transfer of knowledge from a source vessel, abundant in data, to a target vessel with limited data, transcending different ship types and varying load conditions. To achieve this, we introduce a data alignment algorithm, which aligns the source and target domains in both feature and label spaces, effectively mitigating domain discrepancies. Moreover, we incorporate two attention mechanisms to assess the importance of each sensor and time step for SSE. To discern between the source and target domains and to address the computational complexity and significant domain shift, we employ adversarial training.
- We evaluate our approach on two cases: transferring knowledge between different types of ships and between different loading levels of the same ship. Comparative results, which include eleven state-of-the-art transfer learning models, highlight the competitive nature of our proposed model. A comprehensive ablation and sensitivity analysis is conducted to examine the significance of each component and the impact of key parameters.

This research is structured as follows. A literature review on SSE and transfer learning is offered in Section II. The SAFENESS architecture is described in Section III. Section IV discusses the experiment results, and Section V describes our conclusions.

II. Related Work

This section reviews some related works on SSE and transfer learning, and provides some background knowledge and definitions that are relevant to the proposed approach. The section also discusses the advantages and disadvantages of different SSE methods and transfer learning techniques, and identifies the research gap that this paper aims to fill.

A. Sea State Estimation

Sea state, traditionally gauged by diverse sensors, can be monitored over vast oceans by satellites and unmanned aerial vehicles. However, their slow update rates and communication delays make them unsuitable for real-time applications. Although seabed and sea surface sensors can provide real-time measurements, the high costs of installation restrict their use

[16]. Standard incoherent wave radars gather sea state data by measuring wave surface backscatter and ship motion, while advanced coherent systems estimate wave height using the Doppler effect [17]. Despite their precision, these methods are expensive and can be disrupted by ship movement.

A current trend in sea state estimation (SSE) is to use a ship as a sensor [10], [6], [18]. This concept has led to two main SSE methodologies: model-based and data-driven. Model-based methods use a transfer function to map the sea state to the ship's motion [19], dividing further into frequency-based and time-based techniques [11], [20]. For instance, Ren et al. proposed a method to estimate the directional wave spectrum using a non-parametric approach, an L1 optimization, and a novel smoothness constraint [16].

Data-driven methods, on the other hand, extract hidden information from ship motion data. Machine learning and deep learning methods have been developed, with the key difference being the latter's reliance on learned features rather than human-made ones. Tu et al. developed the first machine learning method [9], while Cheng et al. introduced the first deep learning model [7]. More recently, Han et al. developed a hybrid model integrating the strengths of both methods [21].

Despite the success of data-driven methods, challenges persist. First, the ship type dependency issue: SSE models depend on ship motion data, and therefore their effectiveness varies with different ship types. This is problematic when a new ship lacks sufficient data for training. Second, load dependency: A ship's loading condition influences SSE models' performance, as varying cargo capacity and operational status affect a ship's stability across sea states. To address these issues, transfer learning might offer a viable solution, allowing the reuse of data-driven models across different scenarios.

B. Transfer Learning

Transfer learning, a technique facilitating the reuse of data-driven models across distinct domains or tasks, has been extensively investigated in various fields such as computer vision and natural language processing, demonstrating significant efficacy [22]. This method strives to acquire and transfer knowledge from a source domain or task to a designated target domain or task. However, the current body of literature reveals a paucity of studies applying transfer learning to the realm of SSE.

SSE typically utilizes time series data, effectively framing the application of transfer learning to SSE as a task of adapting transfer learning techniques to time series data. Several studies have ventured into this expansive realm. Fawaz et al. were pioneers in exploring the problem of time series classification using transfer learning techniques [23]. Gupta et al. delved into improving model performance on clinical time series data through transfer learning [24]. In the field of transportation systems, a transfer learning model has been proposed for driver model development [25]. Furthermore, Li et al. have innovated a transfer learning method based on a Gaussian process [26].

In the context of SSE, the first transfer learning model was created by Cheng et al., employing a Convolutional Neural Network (CNN)-based Siamese network for similarity

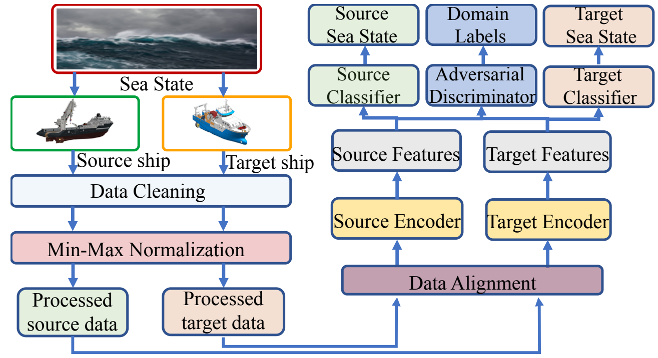


Fig. 2. Illustration of the proposed transferable framework for sea state estimation.

learning [8]. Despite achieving a competitive performance, this model required considerable training time due to its intricate structure.

Depending on whether there are labels in the source and target domains and the relationship between them, transfer learning can be classified as supervised, semi-supervised, or unsupervised. Unlike supervised learning, unsupervised transfer learning does not rely on labeled data that require substantial effort to annotate. However, unsupervised transfer learning may suffer from low performance due to lack of information about the target domain. Semi-supervised learning can serve as a compromise between supervised and unsupervised methods by using both labeled and unlabeled data from both domains.

In this paper, we explore semi-supervised transfer learning when there are sufficient data samples in the source domain and scarce data samples in the target domain. Because our focus is on autonomous ships in real-world environments, it is possible to obtain limited labeled data samples at a low cost. Moreover, improved encoder and adversarial discriminator are utilized for better knowledge transfer.

III. Semi-Supervised Adversarial Transfer Learning for SSE

This section describes the architecture and components of the proposed semi-supervised transfer learning approach for SSE (SAFENESS). The section explains how the approach uses a data alignment algorithm, two attention mechanisms, and an adversarial discriminator to achieve effective knowledge transfer from a source ship to a target ship. The section also presents the loss functions and optimization methods used in the approach.

A. Problem Description

In the context of semi-supervised transfer learning for SSE, our premise rests upon the source domain ship gathering extensive data paired with corresponding sea states. On the contrary, for the target domain ship, which may be a newly commissioned vessel, the available data is expected to be notably limited. This paucity of data presents an obstacle to the development of a high-performance, data-driven model, thereby highlighting the necessity for transfer learning.

In the source domain, the ensemble of samples is defined as $D_s = (X_s^i, y_s^i)_{i=1}^{n_s}$, wherein X_s and y_s symbolize the samples and their respective labels. The term n_s signifies the total number of samples. Meanwhile, the target domain harbors a more modest assortment of samples denoted as $D_t = (X_t^i, y_t^i)_{i=1}^{n_t}$, where X_t and y_t stand for the samples and corresponding labels in the target domain. The term n_t represents the total number of these samples. Given the relative scarcity of target domain data, it is anticipated that n_s will be significantly greater than n_t, indicating dissimilar probability distributions across the source and target domains.

The effectiveness of a deep learning model will likely diminish if a model trained using source domain data, X_s, is subsequently applied to target domain data, X_t. It is impracticable to solely utilize the limited data from X_t to train a deep learning model given the considerable data prerequisites for such a task. To address this predicament, we advocate for the development of a transfer learning model that can effectively navigate the aforementioned challenges.

TABLE I
NOTATIONS USED IN THIS WORK

No.	Para.	Meaning	No.	Para.	Meaning
1	B	Batch_size	14	MLP	multilayer perceptron
2	CNN	convolutional neural network	15	n_s	number of source samples
3	Conv1D	1D convolution	16	n_t	number of target samples
4	CA	channel attention	17	ReLU	rectified linear unit
5	D_s	source domain	18	SSE	sea state estimation
6	D_t	target domain	19	TA	temporal attention
7	E_s	Training epochs	20	X_s	source samples
8	E_{adv}	Training epochs	21	X_t	target samples
9	E_{tot}	Training epochs	22	y_s	source labels
10	H	learned features by Conv1d	23	y_t	target labels
11	L_{CE}	loss function for encoder	24	Φ	encoder
12	L_{adv}	loss function for adversarial discriminator	25	Ψ	classifier
13	L_{total}	total loss function	26	Ω	adversarial discriminator

B. Framework

To realize the outlined objective, several hurdles required to be addressed. Initially, the scarcity of information for the target domain presented a conundrum of optimally exploiting this limited data. Secondly, the task of effectually transferring knowledge from the source to the target vessel demanded careful consideration. Lastly, the determination of the necessary feature extraction for facilitating knowledge transfer was also crucial.

To surmount these challenges, we applied data alignment to the limited target ship data to maximize its utilization. Following that, adversarial learning was implemented on the aligned data with the intention of aligning the semantic information across both source and target vessels. Conclusively, we adopted a novel Convolutional Neural Network (CNN) structure equipped with two attention modules, designed specifically to extract semantic features of higher value. This composite approach was intended to effectively enable the transfer of SSE knowledge from one maritime context to another despite the disparity in data abundance.

The architecture of the proposed scheme for SSE is depicted in Fig. 2. The framework begins by instituting a cleaning process on the ship motion data, carried out identically across both source and target domains. Subsequently, data alignment is implemented, aiming to exploit the limited, yet critical, target information available. Upon alignment, the processed data is channeled through a specifically designed feature encoder to derive distinct features for the source and target domains. These features then dictate the refinement of the respective source and target encoders, guided by adversarial, source classification, and target classification losses. An in-depth explanation of each module within the proposed transfer learning framework is provided in the succeeding sections of this paper.

C. Data Alignment

As previously established, our model is designed to leverage a relatively sparse amount of information from the target ship. In an effort to enhance this scarce information, we employ the pairs of samples as suggested by [27]. This seemingly simple strategy proves to be extremely effective in creating a meaningful alignment between the source ship, characterized by an ample number of samples, and the target ship, characterized by a limited quantity. Four types of pairings are constructed. The first pairing comprises samples originating from ships sharing an identical class label. The second pairing also consists of samples possessing the same label, however, one sample originates from the source ship while the other is from the target ship. The third pairing is composed of samples from source ships belonging to different class labels. Finally, the fourth pair consists of samples from different class labels, where one is obtained from the source ship and the other from the target ship.

The choice of four pairings represents a carefully considered balance between data alignment effectiveness and computational efficiency. The rationale behind this decision is that increasing the number of pairs may not necessarily contribute to performance enhancement, but rather it may inadvertently introduce noise or redundancy into the data. Additionally, it is worth noting that expanding the number of pairs would invariably augment the computational complexity and prolong the training time of our model.

It is evident that the second pairing category has the fewest number of samples. To ensure consistent sample count across all pairing categories, we judiciously select samples from the first, third, and fourth pairs to match the sample size of the second pairing category.

D. Attention-Enabled Encoder

To capture richer and more relevant features, we introduce an attention-infused Convolutional Neural Network (CNN) as

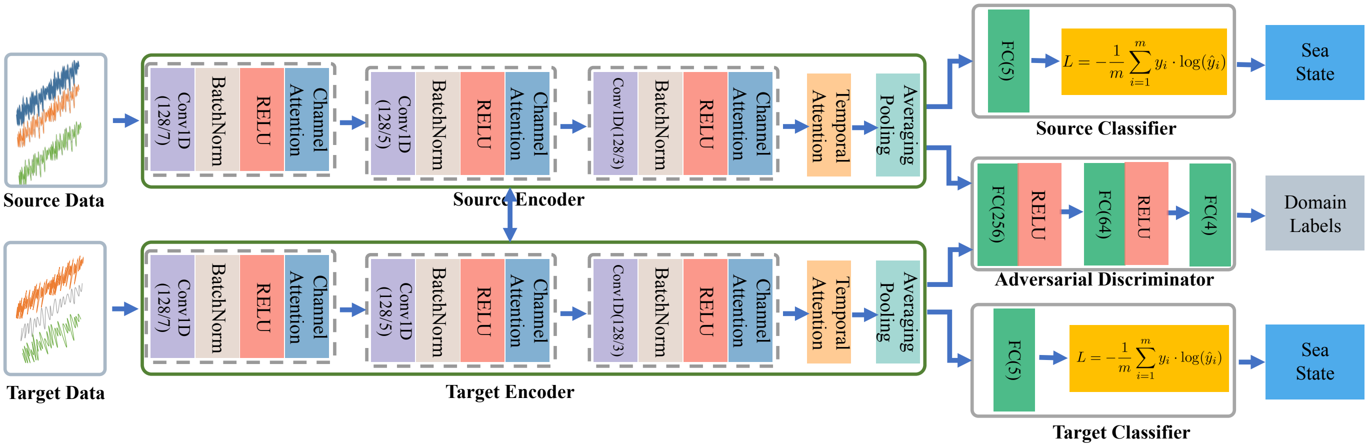


Fig. 3. Illustration of neural network.

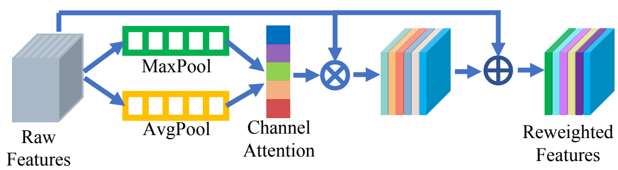


Fig. 4. Illustration of the CA module.

our feature encoder, as illustrated in Fig. 3. This encoder comprises three convolutional blocks and integrates two novel attention modules, namely Channel Attention (CA) and Temporal Attention (TA). Each convolutional block is composed of a 1D convolutional layer (Conv1D), a batch normalization layer (BatchNorm as shown in Fig. 3), a Rectified Linear Unit (ReLU) layer, and a CA module. It's important to note that the parameters for each Conv1D layer are indicated in brackets following the Conv1D label.

The Conv1D layer performs sliding convolutional operations along the temporal dimension of the input, subsequently producing a tensor of outputs [28]. This is followed by the batch normalization layer, which normalizes the output of the Conv1D layer to mitigate internal covariate shifts and enhance training stability [29]. Finally, the ReLU layer applies a non-linear activation function to the output of the batch normalization layer, thereby instilling non-linearity and sparsity within the network [30].

1) CA module: In this study, a 1D Convolutional Neural Network (Conv1D) serves as our feature extractor, deriving the feature $\mathbb{R}^{T\times C}$ from the raw data or feature map $\mathbb{R}^{T\times N}$. Here, C represents the number of feature channels. However, one of the limitations of traditional CNNs is that they treat each channel equally, thereby failing to highlight the essential information embedded within the learned features. To surmount this issue, a Channel Attention (CA) module is incorporated.

The primary function of the CA module is to allocate differential weights to the varying channels. The calculation process implemented by the CA module is illustrated in Fig. 4. Assuming the features learned by Conv1D to be $H = [x_1, x_2, ..., x_C] \in \mathbb{R}^{T\times C}$, we initially apply global average pooling and global max pooling to the learned features to comprehend the intra-relationships across different channels. The representations of these relationships are subsequently dispatched to two Multilayer Perceptrons (MLPs) with shared weights to calculate the attention weights. Lastly, the derived attention weights are utilized to calibrate the raw CNN features.

2) TA module: While the Channel Attention (CA) module emphasizes critical channels, it might neglect the temporal information within the time dimension. To counteract this issue, we introduce a Temporal Attention (TA) module. The calculation process employed by this module is depicted in Fig. 5.

Assuming the input to the TA module to be $H \in \mathbb{R}^{T\times C}$, it is first mapped to $\widetilde{H} \in \mathbb{R}^{T\times K}$ via a Multilayer Perceptron (MLP). A bidirectional Gated Recurrent Unit (BiGRU) is then employed to extract the temporal information. We leverage the principles from the CA module for the TA module to highlight the most relevant temporal features.

Furthermore, $\widetilde{H}$ is also directed through average pooling and max pooling operations, with the resulting pooled features being repeated along the time axis. These features then feed into two MLPs with shared weights. This process ensures that the key temporal characteristics within the input are duly highlighted, thereby improving the overall model performance.

E. Adversarial Discriminator

Let E_s and E_t denote the source and target encoders, respectively. A rudimentary approach to estimate sea states might involve the initialization of the target encoder E_t using the already trained source encoder E_s. This strategy is based on leveraging the insights and knowledge accumulated by the source encoder to bootstrap the target encoder, thereby enabling predictions about the sea state using the limited target data. However, this approach is not infallible, particularly when stark disparities exist between the source and target data. Therefore, to ensure accuracy in such situations, it becomes crucial to devise a methodology capable of minimizing the divergence between the source and target domains.

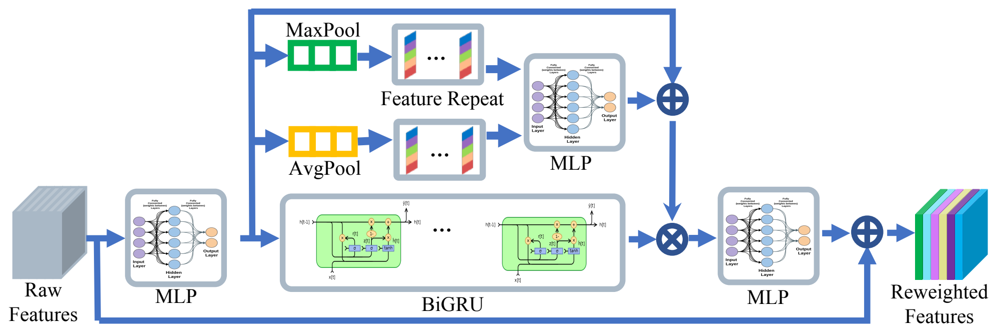


Fig. 5. Illustration of temporal attention module.

In this study, we put forward the proposition of employing adversarial learning for model training. The underlying principle of this approach is to train the network in a way that the generated output is indistinguishable from samples drawn from a source distribution for a discriminator [31]. This methodology facilitates the enhancement of the network's performance on a specific task by prompting it to produce output that closely resembles the source distribution. The objective here is to learn a mapping from the target domain to the source domain such that the distribution of the mapped target data mimics the source data distribution as closely as possible. This can aid in overcoming the issue of distribution discrepancies between the source and target data sets.

In the current work, a multi-class discriminator with four distinct outputs is deployed to categorize sample pairs into different classes. This innovative approach allows the network to generate output that aligns semantically and class-wise with the source distribution while simultaneously confounding the discriminator. The network and discriminator are trained in an adversarial fashion, with the overarching aim of enabling the network to yield output that the discriminator struggles to categorize accurately. Specifically, the objective is to bewilder the discriminator concerning pairs 1 and 2, and pairs 3 and 4. This strategy is predicated on the fact that pairs 1 and 2 share the same class label but originate from different distributions, while pairs 3 and 4, despite arising from distinct distributions, possess different class labels. Should the discriminator find it difficult to differentiate between pairs 1 and 2 or between pairs 3 and 4, it can be concluded that the adversarial learning method has achieved its intended goal.

Our methodology is distinct from other adversarial learning techniques in that it also utilizes information from the target distribution during the training phase. This aids in effectively managing scenarios where there is insufficient target data available for training. As illustrated in Fig. 3, the adversarial discriminator is composed of three fully connected (FC) blocks, which are responsible for the classification of the input sample pairs. It's worth noting that the first two FC blocks consist of an FC layer and a RELU layer, while the final block only contains an FC layer supplemented by a softmax layer (not depicted in Fig. 3). By employing this training methodology for the network and the discriminator, we can accomplish our objective of discovering a shared feature space that is both invariant to the domain and semantically aligned.

F. Training Process

In order to achieve the intended model performance, this study adopts a step-wise training methodology. The specifics of the algorithm are presented in **Algorithm** 1. The initial step (rows 1 to 8) involves training the source encoder and its associated classifier utilizing the ship motion data from the source ship. The optimization of the source encoder and its corresponding classifier is facilitated by the loss function defined as follows:

$$L_{CE} = -\frac{1}{m}\sum_{i=1}^{m} y_i \log(\hat{y}_i) = -\frac{1}{m}\sum_{i=1}^{m} y_i \log(\Psi(\Phi(x_s^i))) \tag{1}$$

where y_i is the true value for class i. Φ denotes the encoder. Ψ means the classifier. x_s^i is the samples of the i-th class from the source ship. $\Psi(\Phi(x_s^i))$ is the value of the i-th class predicted by the proposed SAFENESS.

Given the limited information available from the target ship, a data alignment algorithm and a multi-class adversarial discriminator have been proposed in this study to utilize this information. In contrast to the conventional adversarial discriminator, which only generates binary outputs (true/false), our work utilizes a multi-class approach due to the creation of four pairs to leverage the restricted target samples.

Hence, the second step (lines 9 to 17 in the algorithm) involves training the adversarial discriminator using the four constructed pairs. Throughout this training phase, the encoder and classifier remain frozen and are not updated. The cross-entropy (CE) loss function as defined in Eq. (2) is employed to optimize the model during this step. This process essentially serves to initialize the discriminator with data from the four different pairs.

$$L_{adv}^{'} = -\frac{1}{k}\sum_{i=1}^{k=4} y_p^i \log(\widehat{y_p^i}) = -\frac{1}{k}\sum_{i=1}^{k=4} y_p^i \log(\Omega(\Phi(x_p^i))) \tag{2}$$

where x_p^i and y_p^i mean the true value for class i of the pair data. Ω denotes the adversarial discriminator. $\Omega(\Phi(x_p^i))$ is the predicted value of the i-th class by the proposed SAFENESS. There are four pairs used, the k is thus set to 4.

Algorithm 1 Training algorithm

Input : Φ, Ψ, Ω, E_s, E_{adv} and E_{tot}, B, (X_s, y_s), (X_t, y_t), refer TABLE I for detail information.

Output: Well-trained model: Φ, Ψ, and Ω

1 **for** $e = 0;\ e < E_s;\ e = e + 1$ **do**
2   **for** $b = 0;\ b < B;\ b = b + 1$ **do**
3   b-th batch samples (X_s^b, y_s^b) from the source dataset
4         $y_{out}^{pre} \leftarrow \Psi(\Phi(X_s^b))$
5         $l_s \leftarrow L_{CE}(y_{out}^{pre}, y_s^b)$
6         $W_s \leftarrow W_s - \lambda_s \nabla l_s(W_s)$
7   **end**
8 **end**
9 **for** $e = 0;\ e < E_{adv};\ e = e + 1$ **do**
10   **for** $b = 0;\ b < B;\ b = b + 1$ **do**
11   Getting four pairs as described in Section III-C.
12   b-th batch samples (X_{ps}^b, y_p^b) and (X_{pt}^b, y_p^b) from the four paired data.
13         $y_{out}^{pre} \leftarrow \Omega[concat((\Phi(X_{ps}^b)), \Phi(X_{pt}^b)))]$.
14         $l_{adv} \leftarrow L'_{adv}(y_{out}^{pre}, y_p^b)$.
15         $W_{adv} \leftarrow W_{adv} - \lambda_{adv} \nabla l_{adv}(W_{adv})$
16   **end**
17 **end**
18 **for** $e = 0;\ e < E_{tot};\ e = e + 1$ **do**
19   **for** $b = 0;\ b < B;\ b = b + 1$ **do**
20   Getting the second and fourth pairs as described in Section III-C.
21   b-th batch samples $(X'^b_{ps}, y'^b_{ps}, y'^b_p)$ and $(X'^b_{pt}, y'^b_{pt}, y'^b_p)$ from the two paired data.
22         $y_{out}^{pre} \leftarrow \Omega[concat((\Phi(X'^b_{ps})), \Phi(X'^b_{ps})))]$.
23         $l_{adv} \leftarrow L_{adv}(y_{out}^{pre}, y'^b_p)$.
24         $l_s \leftarrow L_{CE}(y_{out}^{pre}, y'^b_{ps})$
25         $l_t \leftarrow L_{CE}(y_{out}^{pre}, y'^b_{pt})$
26         $l_{sum} = l_t + l_s + \eta * l_{adv}$
27         $W_s \leftarrow W_s - \lambda_{sum} \nabla l_{sum}(W_s)$
28   **end**
29   **for** $b = 0;\ b < B;\ b = b + 1$ **do**
30   Getting the four pairs as described in Section III-C.
31   b-th batch samples (X_{ps}^b, y_p^b) and (X_{pt}^b, y_p^b) from the four paired data.
32         $y_{out}^{pre} \leftarrow \Omega[concat((\Phi(X_{ps}^b)), \Phi(X_{pt}^b)))]$.
33         $l_{adv} \leftarrow L'_{adv}(y_{out}^{pre}, y_p^b)$.
34         $W_{adv} \leftarrow W_{adv} - \lambda_{adv} \nabla l_{adv}(W_{adv})$
35   **end**
36 **end**

In the third step of the process (lines 18 to 36 in the algorithm), there are two sub-steps. The initial sub-step (lines 19 to 28) involves freezing the discriminator and training the encoder and classifier using samples from the second and fourth pairs out of the four data pairs. In an attempt to confuse the discriminator between pairs 1 and 2, and pairs 3 and 4, the true label for the second pair is borrowed from the first pair, and the true label for the fourth pair is borrowed from the third pair. The loss function utilized in this sub-step is defined as follows:

$$\begin{aligned} L_{total} &= \eta L_{\text{adv}} + L_{\text{CE}}(X_s, Y_s) + L_{\text{CE}}(X_t, Y_t) \\ L_{\text{adv}} &= -y_p^1 log(\Omega(\Phi(x_p^2))) - y_p^3 log(\Phi(x_p^4)) \end{aligned} \tag{3}$$

where η is used to balance the classification and confusion. X_s and Y_s are the samples from the source ship. X_t and Y_t are the samples from the target ship. y_p^1 and y_p^3 are the true value of the first pair data and the third pair data, respectively. x_p^2 and x_p^4 are the samples from the second pair data and the fourth pair data.

The second sub-step (lines 29 to 36 in the algorithm) involves training the discriminator while the encoder and classifier are held constant. During this sub-step, we employ all four pairs of data to train the discriminator in the same manner as we did in the second step of the algorithm. This approach ensures that our adversarial learning framework is capable of robustly distinguishing between different classes of data pairs, even when they come from differing source and target distributions.

IV. Experiment

This section presents the experiment settings and results of the proposed approach. The section introduces the datasets, evaluation metrics, baseline methods, and implementation details used in the experiment. The section then reports and analyzes the performance of the proposed approach in two cases: transferring knowledge between different types of ships and between different load levels of the same ship. The section also conducts an ablation study and a sensitivity analysis to examine the significance of each component and the impact of key parameters in the approach.

A. Experiment settings

1) Data source: The data employed in this research was simulated by the Offshore Simulator Centre AS (OSC), a recognized training platform in the field of offshore operations [32]. In order to demonstrate the effectiveness of the proposed model, we established two distinct scenarios: the first scenario involves knowledge transfer between two different types of vessels, and the second scenario involves knowledge transfer between the same type of vessel but under different loading conditions. These scenarios were designed to thoroughly test the proposed model's ability to handle diverse conditions and transfer knowledge effectively.

- **Different types of ships**: Two research vessels (hereafter named RV_G and RV_Z) owned by the Norwegian University of Science and Technology (NTNU) are used to simulate the zigzag motion in the five different sea states. The RV_G is smaller than RV_Z. There are only three thrusters in RV_G but there are six thrusters in RV_Z. The knowledge transfer from RV_G to RV_Z is labeled as G→Z, and RV_Z to RV_G is labeled as Z→G.
- **Same type of ships**: The motion behavior of a ship in different sea conditions will be different under different load conditions, so we conduct the experiment of knowledge transfer between the same type of ship but with different loadings. The research vessel RV_G is chosen

as for the simulation. One case is that the RV_G is empty, and another is the RV_G is with loading of 100 tons. The knowledge transfer from empty loading to full loading is labeled as E→F, and full loading to empty loading is labeled as F→E.

In this study, we consider nine parameters in X_s and X_t: roll, yaw, pitch, surge velocity, sway velocity, heave velocity, roll velocity, pitch velocity, and yaw velocity. The sea states are defined based on the observed wave height, with ten categories ranging from calm to extreme conditions [5]. These sea states also have associated probabilities of occurrence, which have been outlined in the existing literature.

However, for the purposes of this work, we limit our consideration to sea states 0 through 5. These six states represent approximately 96% of the sea conditions that are typically encountered. Given the similarity between sea states 0 and 1, we choose to consolidate these into a single state for the ease of analysis.

2) Metric: The metric **F1** is used to evaluate the models. The definitions of **F1** is presented as follows:

$$Precision = \frac{TP}{TP+FP} \quad Recall = \frac{TP}{TP+FN} \tag{4}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{5}$$

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

All models are implemented by Pytorch and all experiments are repeated five times. To be fair, we select samples from the target ships according to different random seeds each time.

B. Comparison with simple transfer learning methods

Fine-tuning, direct transfer learning, and learning from scratch are three common simple transfer learning methods widely utilized in the literature: 1) Fine-tuning: This method involves taking a pre-trained model, typically a deep neural network, and further training it on a specific target task. Instead of retraining the entire model from scratch, earlier layers (often the lower ones) are frozen, while only the later layers' weights are updated to adapt to the new task. 2) Direct Transfer Learning: Similar to fine-tuning, direct transfer learning modifies a pre-trained model for a target task. However, it can involve more significant alterations, such as replacing or modifying specific layers in the pre-trained model to better align with the new task's requirements. 3) Learning from Scratch: In contrast to transfer learning, learning from scratch entails training a model entirely anew, starting with random initial weights and without relying on pre-trained knowledge. The model is trained from the ground up on the target task using available data.

Based on our experimental results shown in Fig.6, it is evident that our proposed method consistently achieves the highest level of performance. Fine-tuning emerges as the second-best approach, showcasing notable improvements in various transfer tasks. Specifically, we observe enhancements of 21.52%, 30.38%, 14.07%, and 7.66% in the G→Z, Z→G, E→F, and F→E transfer tasks, respectively. The superior

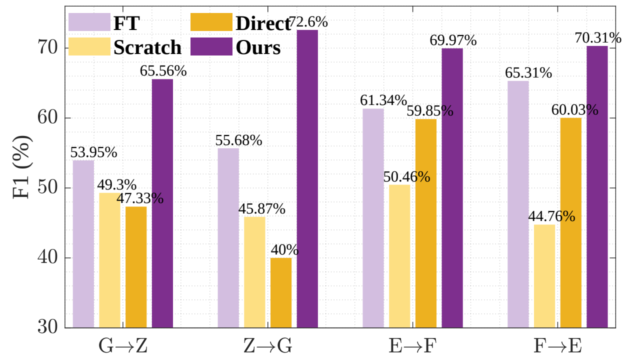


Fig. 6. Comparison of simple transfer learning methods.

performance of fine-tuning can be attributed to its ability to leverage both source data and limited information from the target domain. In contrast, the other two methods lack this capability, which accounts for their comparatively lower performance levels.

In the context of direct transfer learning, it is evident that knowledge transfer between ships of the same type yields superior accuracy compared to transfers between different types of ships. Specifically, the performance in tasks G→Z and Z→G lags behind that in tasks E→F, and F→E. This discrepancy aligns with expectations, as the distribution difference between ships of the same type is inherently smaller than that between ships of different types.

Regarding learning from scratch, it is noteworthy that the achieved performance is nearly identical in both cases, particularly in terms of average accuracy. This observation is intriguing, given that this method is susceptible to overfitting due to the limited information available for model training. It is plausible that this overfitting challenge constitutes a primary reason for the absence of a significant performance advantage in this approach.

C. Comparison with state-of-the-art deep transfer learning methods

To evaluate our model, we compare it to eight state-of-the-art approaches, implemented as follows: **1) DAN** (Deep Adaptation Network) is a classical method for transfer learning (i.e., domain adaptation). We replace the backbone with our proposed encoder. The learning rate is set to 3e-3 [33]. **2) DAAN** (Dynamic Adversarial Adaptation Network) can dynamically learn domain-invariant features [34]. The learning rate is set to 0.01. **3) DANN** (Domain-adversarial neural network) is used for tasks in which training and testing data have similar but different distributions [35]. **4) DeepCoral** is a method that integrates the CORAL and deep networks. DeepCoral has more powerful non-linear learning capabilities and works seamlessly with deep networks [36]. **5) BNM** (Batch Nuclear-norm Maximization) improves both discriminability and diversity by using batch nuclear-norm maximization on the output matrix [37]. **6) CADA** (Contrastive Adversarial Domain Adaptation) is used for predicting cross-domain remaining useful life. It is built on the adversarial domain adaptation architecture with a contrastive loss [31]. **7) FADA** (Few-shot Adversarial Domain Adaptation) provides a framework for addressing the problem of transfer learning when there are sufficient labeled data in the source domain but limited labeled data in the target domain [27]. **8) STLSSE** is a recently proposed model for SSE based on transfer learning [8].

TABLE II
COMPARISON WITH SOTA TRANSFER LEARNING METHODS (F1 (%))

Methods	Different type of ships			Same type of ships		
	G→Z	Z→G	Average	E→F	F→E	Average
DAN	54.33	55.47	54.90	62.67	62.36	62.52
DAAN	53.43	56.49	54.96	58.85	60.57	59.71
DANN	53.61	56.37	54.99	58.68	60.57	59.63
DeepCoral	53.43	54.99	54.21	59.03	60.80	59.92
BNM	34.62	35.64	35.13	36.55	35.94	36.25
CADA	20.34	20.34	20.34	20.38	19.00	19.69
FADA	54.88	66.77	60.83	52.43	67.23	59.83
STLSSE	61.23	63.54	62.39	65.38	68.97	67.18
ours	**65.56**	**72.60**	**69.08**	**69.97**	**70.31**	**70.14**

The results are displayed in TABLE II. It is evident that our proposed method consistently achieves the best results across all four cases. When considering the knowledge transfer between different types of ships, Z→G attains a higher F1 accuracy than G→Z for most methods. STLSSE delivers the second-best performance, followed by FADA. In comparison to STLSSE, our method offers improvements of 7.07%, 14.26%, and 10.72% in the G→Z, Z→G, and average cases respectively.

When assessing knowledge transfer between the same type of ship but with different loadings, STLSSE again ranks second-best for both E→F and F→E scenarios. Compared to STLSSE, our proposed method shows improvements of 7.02%, 1.94%, and 4.41% in the G→Z, Z→G, and average cases respectively. Interestingly, CADA's performance appears to be equivalent to random guessing, given there are five sea states and its F1 score is approximately 20%.

D. Comparison with different encoders

TABLE III
ENCODER PERFORMANCE COMPARISON (F1 (%))

Methods	Different type of ships			Same type of ships		
	G→Z	Z→G	Average	E→F	F→E	Average
CNN	52.52	50.24	51.38	62.07	67.46	64.77
LSTM	43.45	48.44	45.95	46.09	48.35	47.22
GRU	40.44	44.71	42.58	40.41	42.69	41.55
MLP	21.27	35.75	28.51	28.12	28.58	28.35
MLF	63.70	68.03	65.87	69.10	67.78	68.44
FCN	64.24	69.77	67.01	66.93	65.03	65.98
DenseNet	63.22	66.95	65.09	62.07	64.43	63.25
ours	**65.56**	**72.60**	**69.08**	**69.97**	**70.31**	**70.14**

To illustrate the performance of the proposed encoder, we compared it to seven baselines, implemented as follows: **1) CNN**: A one layer CNN was utilized. The number of filters was 128 with size of 7. **2) LSTM**: A one layer LSTM was employed with a hidden size selected from {8, 16, 32, 64}. The best performance was chosen for comparison. **3) GRU**: A one layer GRU was employed with a hidden size selected from {8, 16, 32, 64}. The best performance was chosen for comparison. **4) MLP**: A three-layer MLP was utilized. The hidden size was set to 500 in each layer. The dropout layer was applied between different layers. **5) MLSTM_FCN (MLF)**: A parallel structure integrating LSTM and FCN was utilized [38]. The LSTM had one layer with hidden size of 8 and FCN has three layers with filters (128, 7), (256, 5) and (128, 3). **6) FCN**: A three-layer FCN was employed. The numbers of filters in each layer were 128, 256, and 128. The filter sizes were 7, 5, and 3 in each layer. **7) DenseNet**: The DenseNet was adapted from [6].

The results are presented in TABLE III. When considering knowledge transfer between different types of ships, our proposed method demonstrated the best performance. FCN obtained the second-best performance among the compared methods. For knowledge transfer between the same type of ship but with different loadings, our method again achieved the best performance, while MLSTM_FCN obtained the second-best performance. Notably, MLP performed poorly in both cases, indicating its limited capability to learn effective features. It is interesting to observe that DenseNet, despite its complex structure, did not yield the best performance in either case. These results highlight the importance of designing an encoder that is suitable for the task at hand, rather than simply relying on the most complicated architecture.

E. Comparison with different attention mechanisms

Four widely-used attention mechanisms were employed to further illustrate the performance of the proposed attention mechanism for sea state knowledge transfer between ships.

The details of the attention modules used are as follows: **1) SE** is the most classical attention module. It can learn the importance of each channel in CNN and then give more attention to these influential information [39]. **2) GC** provides a pioneering method for learning long-range relationship from the data. This method can achieve high accuracy but with significantly fewer computations [40]. **3) CA-SFCN** offers a good solution for capturing the long-range dependencies of the time series data and learning the interactions between different variables [41].

In this comparison, we removed all of the attention models we proposed and inserted the attention mechanisms into the position of CA. As illustrated in Fig. 7, our proposed method had the best performance, as compared to the other three attention mechanisms. From the results, we also determined that SE was second-best, followed by GC and CA-SFCN (which was the worst). This may have been due to the following aspects. First, our proposed method can learn features from the channel and temporal dimensions at the same time, while other models (e.g., SE) can only learn the channel information. Second,

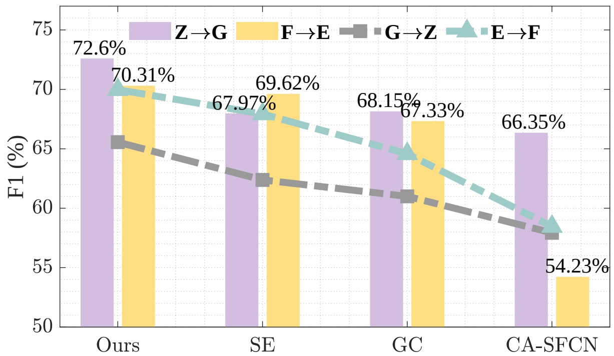


Fig. 7. Comparison of different attention mechanisms.

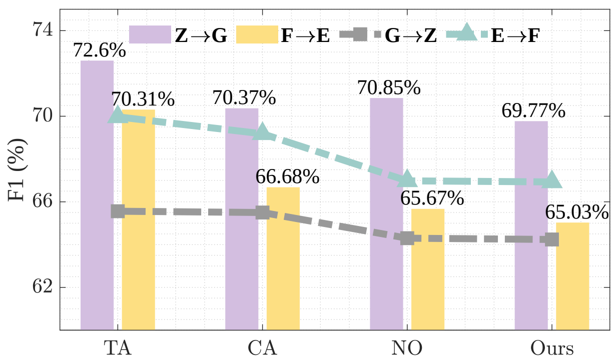
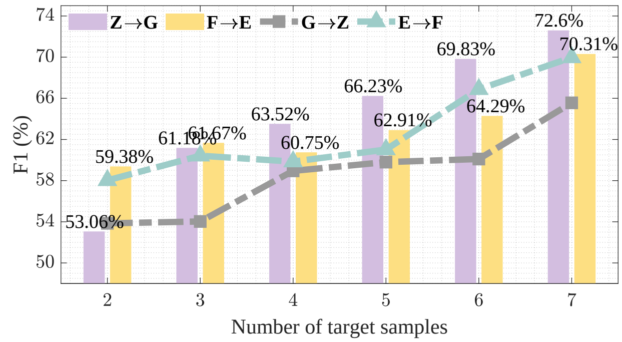
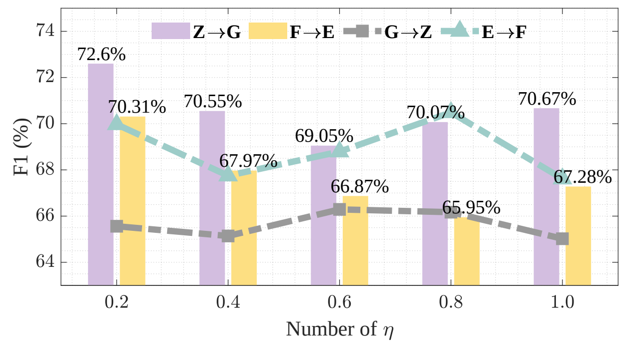


Fig. 8. Ablation analysis.



(a) Influence of the number of target samples



(b) Influence of η

Fig. 9. Sensitivity analysis.

both GC and CA-SFCN adopted a structure similar to self-attention, which can capture spatial information but performs poorly when learning temporal information.

F. Ablation Analysis

In order to further illustrate the rationality of the model designed here, we conducted an ablation study to evaluate the importance of each component. As the most complicated part is the encoder, we focused on the influence of each component therein. There were three variants built for the comparison, implemented as follows: **1) TA**: the temporal attention module was not used in the encoder. **2) CA**: all channel attention modules were removed from the encoder. **3) NO**: There were no any attention modules in the encoder.

The results are shown in Fig. 8. Our proposed model showed a better performance than the three variants, illustrating the rationality of the proposed model. The results show that the accuracy drop was significant when the CA module was removed, a condition similar to NO, which did not use any attention modules. Conversely, the accuracy did not decrease significantly when TA is removed. However, it is worth noting that we used three channels of attention, but only one TA of attention. Thus, it was difficult to verify if CA was more important than TA.

G. Sensitivity Analysis

A sensitivity analysis was conducted to better understand the influence of the number of target samples and η. We varied these important variables but kept other parameters unchanged. The variation range for the number of target samples and η is $\{2, 3, 4, 5, 6, 7\}$ and $\{0.2, 0.4, 0.6, 0.8, 1.0\}$, respectively. The results are shown in Fig. 9(a) and Fig. 9(b).

In the four cases, it is clear that the accuracy increased with the increasing number of target samples (see Fig. 9(a)). This was reasonable because if more samples were provided by the target, there would be more information for the proposed model to learn. However, we could not draw the same conclusion from Fig. 9(b). From the results alone, the effect of η was not obvious. Section III-F shows that we trained our model step-by-step, and η mainly affected one loss function of one step in the training process. Other loss functions also affected the training results. In other words, the step-by-step training method may have resulted in a coupling effect because the model was affected by multiple loss functions.

H. Discussion

Given the observed limited generalizability of data-driven models in the realm of sea state estimations, this study introduces a novel approach employing a transfer learning-based, data-driven model for said estimations. This research is predicated on the assumption that a substantial volume of data pertaining to a multitude of sea states can be amassed by the source vessel. However, the data accessible for utilization by the target vessel remains significantly constrained.

In order to substantiate the efficacy of the proposed model, it was juxtaposed with a range of methodologies under two distinct settings: (1) transfer between varying ship types and (2) transfer between different load conditions within the same ship type. Experimental findings suggest that our model yields substantial enhancements over state-of-the-art transfer learning methodologies, such as learning from scratch, direct transfer, unsupervised, and semi-supervised techniques. The relatively inferior performance of the learning from scratch, direct transfer, and unsupervised techniques can primarily be attributed to the ineffective utilization of the limited data derived from the target ship.

Our model surpasses semi-supervised methods primarily due to the innovative feature encoder that incorporates two novel attention modules. More specifically, in comparison with the transfer learning method specifically tailored for sea state estimation (i.e., STLSSE), our approach diverges in three critical ways:

- The data alignment technique employed for the utilization of limited target information is distinctly different, with our approach demonstrating a more straightforward and effective design.
- To optimize the knowledge transfer from the source ship to the target ship, adversarial learning is preferred over similarity learning as implemented in STLSSE. This alternative approach is applied to the aligned data to synchronize the semantic information between the source and target vessels.
- To extract useful features from ship motion data, a unique feature encoder is employed, marking a key departure from the methodology used in STLSSE.

Despite the promising results, there are certain limitations to the proposed model:

1) The model necessitates a minimal amount of information from the target ship. However, this may not always be feasible in real-world situations - for instance, if the model were to be applied to a newly launched ship without any existing historical data. A potential solution might be the execution of the model that enables knowledge transfer between two distinct types of ships, without the requirement for information from the target ship.

2) A further constraint of the proposed method is that it only considers sea state codes 0 to 5, with codes 0 and 1 being consolidated. Despite the more severe sea states constituting only 4% of the total, they hold substantial significance in certain circumstances, such as decision support, hull condition monitoring, and control systems.

3) The third limitation pertains to the application of our method to complex sea states, which may include multiple wave systems or nonlinear wave interactions. In this study, we assumed that the sea state could be characterized by a single wave system with a dominant wave direction and height. However, real-world sea states might be more intricate and dynamic, involving multiple wave systems with varying directions and heights, or nonlinear wave interactions such as wave breaking or wave-current interactions. These complex sea states may present challenges for our approach, as they might require more descriptive features or parameters, or additional data to accurately capture their characteristics. Therefore, applying our approach to complex sea states remains an open problem in our research. A potential solution might involve the use of a more sophisticated model that can manage multiple wave systems or nonlinear wave interactions. For instance, a directional spectrum model could be employed that can estimate the directional wave spectrum from ship motion data. Alternatively, a nonlinear model could be utilized that can account for the nonlinear effects of waves on ship motion.

In conclusion, the proposed model demonstrates a significant enhancement over other contemporary methodologies. It presents a valuable alternative for SSE, particularly during exigent circumstances such as epidemics when there is an elevated demand for ships.

V. CONCLUSIONS

This research explored how to construct a transferable SSE model that could be used to ensure the safety of ship operations. This study proposed SAFENESS, a model that can be trained using sufficient data from a source domain and limited information from a target domain. To achieve this goal, we utilized an algorithm for data alignment that employs limited data from the target domain and sufficient data from the source domain. To maximize knowledge transfer between the source and target domains, we employed an adversarial learning-based framework. Two attention mechanisms were proposed to enhance the learning capabilities of the framework. The model proposed in this research was verified via two cases. The first was one in which the ship types of the source and target domains were different, and the other was one in which the same ships were used, but with different loads. Through comprehensive comparisons of our proposed method to state-of-the-art approaches, it was determined that our model achieved superior performance. Future work will put more attention to these sea states. Moreover, we will also explore physically-informed SSE to leverage the strength of this model, and also compare the model performance between physical-model based methods and data-based methods.

ACKNOWLEDGMENT

Special thanks go to Offshore Simulator Centre AS (OSC) for their data.

REFERENCES

[1] A. BahooToroody, M. M. Abaei, O. V. Banda, P. Kujala, F. De Carlo, and R. Abbassi, "Prognostic health management of repairable ship systems through different autonomy degree; from current condition to fully autonomous ship," *Reliability Engineering & System Safety*, vol. 221, p. 108355, 2022.

[2] T. R. Walker, O. Adebambo, M. C. D. A. Feijoo, E. Elhaimer, T. Hossain, S. J. Edwards, C. E. Morrison, J. Romo, N. Sharma, S. Taylor *et al.*, "Environmental effects of marine transportation," in *World seas: an environmental evaluation*. Elsevier, 2019, pp. 505–530.

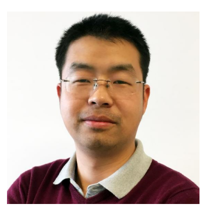
[3] S. Pribyl, "Autonomous vessels in the era of global environmental change," in *Autonomous Vessels in Maritime Affairs: Law and Governance Implications*. Springer, 2023, pp. 163–184.

[4] 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," *Marine Structures*, vol. 76, p. 102904, 2021.

[5] T. I. Fossen, *Handbook of marine craft hydrodynamics and motion control*. John Wiley & Sons, 2011.

[6] 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 Transactions on Instrumentation and Measurement*, 2020.

[7] 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 *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 6644–6650.
[8] 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 Transactions on Intelligent Transportation Systems*, 2021.
[9] 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 Engineering*, vol. 147, pp. 318–332, 2018.
[10] U. D. Nielsen and J. Dietz, "Estimation of sea state parameters by the wave buoy analogy with comparisons to third generation spectral wave models," *Ocean Engineering*, vol. 216, p. 107781, 2020.
[11] A. H. Brodtkorb, U. D. Nielsen, and A. J. Sørensen, "Sea state estimation using vessel response in dynamic positioning," *Applied Ocean Research*, vol. 70, pp. 76–86, 2018.
[12] 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," *Marine Structures*, vol. 85, p. 103274, 2022.
[13] 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 Engineering*, p. 111020, 2022.
[14] T. Kawai, Y. Kawamura, T. Okada, T. Mitsuyuki, and X. Chen, "Sea state estimation using monitoring data by convolutional neural network (cnn)," *Journal of Marine Science and Technology*, vol. 26, no. 3, pp. 947–962, 2021.
[15] B.-O. H. Eriksen and M. Breivik, "Short-term asv collision avoidance with static and moving obstacles," *arXiv preprint arXiv:1907.04877*, 2019.
[16] 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," *Marine Structures*, vol. 76, p. 102904, 2021.
[17] X. Tian, P. Wang, X. Li, X. Wu, W. Lu, C. Wu, Z. Hu, H. Rong, H. Sun, A. Wang *et al.*, "Design and application of a monitoring system for the floatover installation," *Ocean Engineering*, vol. 150, pp. 194–208, 2018.
[18] U. D. Nielsen, A. H. Brodtkorb, and A. J. Sørensen, "Sea state estimation using multiple ships simultaneously as sailing wave buoys," *Applied Ocean Research*, vol. 83, pp. 65–76, 2019.
[19] R. E. Mounet, U. D. Nielsen, A. H. Brodtkorb, E. A. Tannuri, and P. C. de Mello, "Simultaneous sea state estimation and transfer function tuning using a network of dynamically positioned ships," *Applied Ocean Research*, vol. 129, p. 103367, 2022.
[20] R. Pascoal, L. P. Perera, and C. G. Soares, "Estimation of directional sea spectra from ship motions in sea trials," *Ocean Engineering*, vol. 132, pp. 126–137, 2017.
[21] 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 Transactions on Industrial Informatics*, 2021.
[22] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.
[23] H. I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, "Transfer learning for time series classification," in *2018 IEEE International Conference on Big Data (Big Data)*. IEEE, 2018, pp. 1367–1376.
[24] P. Gupta, P. Malhotra, J. Narwariya, L. Vig, and G. Shroff, "Transfer learning for clinical time series analysis using deep neural networks," *Journal of Healthcare Informatics Research*, vol. 4, no. 2, pp. 112–137, 2020.
[25] C. Lu, F. Hu, D. Cao, J. Gong, Y. Xing, and Z. Li, "Transfer learning for driver model adaptation in lane-changing scenarios using manifold alignment," *IEEE Transactions on Intelligent Transportation Systems*, 2019.
[26] Z. Li, J. Gong, C. Lu, and J. Xi, "Importance weighted gaussian process regression for transferable driver behaviour learning in the lane change scenario," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 12 497–12 509, 2020.
[27] S. Motiian, Q. Jones, S. M. Iranmanesh, and G. Doretto, "Few-shot adversarial domain adaptation," *arXiv preprint arXiv:1711.02536*, 2017.
[28] N. Dua, S. N. Singh, and V. B. Semwal, "Multi-input cnn-gru based human activity recognition using wearable sensors," *Computing*, vol. 103, pp. 1461–1478, 2021.
[29] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *International conference on machine learning*. pmlr, 2015, pp. 448–456.
[30] W. Jung, D. Jung, B. Kim, S. Lee, W. Rhee, and J. H. Ahn, "Restructuring batch normalization to accelerate cnn training," *Proceedings of Machine Learning and Systems*, vol. 1, pp. 14–26, 2019.
[31] M. Ragab, Z. Chen, M. Wu, C. S. Foo, C. K. Kwoh, R. Yan, and X. Li, "Contrastive adversarial domain adaptation for machine remaining useful life prediction," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5239–5249, 2020.
[32] "Offshore simulator centre," http://www.offsim.no/.
[33] M. Long, Y. Cao, J. Wang, and M. Jordan, "Learning transferable features with deep adaptation networks," in *International conference on machine learning*. PMLR, 2015, pp. 97–105.
[34] C. Yu, J. Wang, Y. Chen, and M. Huang, "Transfer learning with dynamic adversarial adaptation network," in *2019 IEEE International Conference on Data Mining (ICDM)*. IEEE, 2019, pp. 778–786.
[35] Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation," in *International conference on machine learning*. PMLR, 2015, pp. 1180–1189.
[36] B. Sun and K. Saenko, "Deep coral: Correlation alignment for deep domain adaptation," in *European conference on computer vision*. Springer, 2016, pp. 443–450.
[37] S. Cui, S. Wang, J. Zhuo, L. Li, Q. Huang, and Q. Tian, "Towards discriminability and diversity: Batch nuclear-norm maximization under label insufficient situations," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 3941–3950.
[38] F. Karim, S. Majumdar, H. Darabi, and S. Harford, "Multivariate lstm-fcns for time series classification," *Neural Networks*, vol. 116, pp. 237–245, 2019.
[39] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7132–7141.
[40] Y. Cao, J. Xu, S. Lin, F. Wei, and H. Hu, "Gcnet: Non-local networks meet squeeze-excitation networks and beyond," *arXiv preprint arXiv:1904.11492*, 2019.
[41] Y. Hao and H. Cao, "A new attention mechanism to classify multivariate time series." in *IJCAI*, 2020, pp. 1999–2005.



Xu Cheng (Member, IEEE) received the Ph.D. degree in engineering from the Intelligent Systems Laboratory, Department of Ocean Operations and Civil Engineering, 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. Since April 2022, he has been a Permanent Researcher with Smart Innovation Norway, Halden, Norway. He is currently a Full Professor at the Tianjin University of Technology, Tianjin, China. 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 60 articles as first and co-author in his research interests, including data analysis and artificial intelligence in maritime operations, time series analysis, and predictive maintenance of wind turbines.

Guoyuan Li (Senior Member, IEEE) received the Ph.D. degree in computer science from the Department of Informatics, Institute of Technical Aspects of Multimodal Systems, University of Hamburg, Hamburg, Germany, in 2013. In 2014, he joined the Department of Ocean Operations and Civil Engineering, Intelligent Systems Laboratory, Norwegian University of Science and Technology (NTNU), Ålesund, Norway. He is currently a Professor of ship intelligence at NTNU. He has published more than 80 articles in the areas of his research interests, including modeling and simulation of ship motion, autonomous navigation, intelligent control, optimization algorithms, and locomotion control of bioinspired robots.

Robert Skulstad received his M.Sc. degree at the department of Engineering Cybernetics and Ph.D. degree at the department of Ocean Operations and Civil Engineering from the Norwegian University of Science and Technology (NTNU), Norway, in 2014 and 2021, respectively. He now works as a post-doctoral fellow at the department of Ocean Operations and Civil Engineering at NTNU in Ålesund. His research interests include the use of machine learning for motion prediction and control in maritime operations.

Houxiang Zhang (Senior Member, IEEE) is a full Professor at the Department of Ocean Operations and Civil Engineering, Faculty of Engineering, Norwegian University of Science and Technology (NTNU).

Dr. Zhang received his Ph.D. degree on Mechanical and Electronic Engineering in 2003. From 2004, he worked as Postdoctoral fellow, senior researcher at the Institute of Technical Aspects of Multimodal Systems (TAMS), Department of Informatics, Faculty of Mathematics, Informatics and Natural Sciences, University of Hamburg, Germany. In Feb. 2011, he finished the Habilitation on Informatics at University of Hamburg. Dr. Zhang joined the NTNU, Norway in April 2011 where he is a Professor on Mechatronics. From 2011 to 2016, Dr. Zhang also hold a Norwegian national GIFT Professorship on product and system design funded by Norwegian Maritime Centre of Expertise. In 2019, Dr. Zhang has been elected to the member of Norwegian Academy of Technological Sciences.

Dr. Zhang has engaged into two main research areas including control, optimization and AI application especially on autonomous vehicle; and marine automation, digitalization and ship intelligence. He has applied for and coordinated more than 30 projects supported by Norwegian Research Council (NFR), German Research Council (DFG), EU, and industry. In these areas, he has published over 200 journal and conference papers as author or co-author. Dr. Zhang has received four best paper awards, and five finalist awards for best conference paper at International conference on Robotics and Automation.