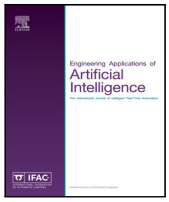
Engineering Applications of Artificial Intelligence 138 (2024) 109316

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai



Biofouling detection and classification in Tidal Stream Turbines through soft voting ensemble transfer learning of video images☆

Haroon Rashid [a], Mohamed Benbouzid [a,b,*], Yassine Amirat [c], Tarek Berghout [d], Hosna Titah-Benbouzid [a], Abdeslam Mamoune [a]

[a] University of Brest, UMR CNRS 6027, 29238 Brest, France
[b] Shanghai Maritime University, Logistics Engineering College, 201306 Shanghai, China
[c] ISEN Yncrea Ouest, L@bISEN, 29200 Brest, France
[d] University of Batna 2, Laboratory of Automation and Manufacturing Engineering, 05000 Batna, Algeria

ARTICLE INFO

Keywords:
Tidal stream turbine
Biofouling
Convolutional Neural Networks
Classification
Ensemble transfer learning
Soft voting ensemble

ABSTRACT

This study addresses the biofouling challenges in Tidal Stream Turbines (TSTs) to ensure their reliable and optimal operation. In this context, it is proposed an effective methodology employing a soft voting ensemble transfer learning-based approach for the detection and extent classification of biofouling. The proposed framework incorporates essential components such as data augmentation and pre-processing, including image resizing and data segmentation, forming a comprehensive video image-based approach. To overcome the constraint of limited data, experimental investigations were conducted, resulting in the acquisition of two datasets: one from the TST platform at Shanghai Maritime University (SMU) and the other from the tidal turbulence test facility at Lehigh University (LU). The three prominent convolutional neural network models, namely Visual Geometry Group (VGG), Residual Network (ResNet) and MobileNet, trained on these datasets, demonstrate precise detection and classification of turbine conditions, achieving an accuracy of 83% for the SMU dataset and 90% for the LU dataset. The noted disparity in accuracy for the SMU dataset is attributed to its smaller size, highlighting the significant impact of dataset scale on classification performance. This study provides valuable insight into the development of effective biofouling detection and classification strategies for TST systems.

1. Introduction

Renewable energy is crucial for reducing reliance on fossil fuels and addressing climate change. Tidal energy has emerged as a promising solution to tap into the ocean's current substantial energy potential. Unlike solar or wind power, which are susceptible to weather fluctuations, tidal energy stands out for its predictability, as tidal patterns can be reliably forecasted. This predictability facilitates effective planning and seamless integration into the power grid, enhancing grid stability and positioning tidal energy as a significant asset in the energy mix. Tidal Stream Turbines (TSTs) are pivotal in harnessing tidal energy but face significant challenges due to biofouling. Biofouling, the accumulation of marine organisms such as algae, barnacles, and mollusks on underwater surfaces, severely impacts the efficiency and reliability of TSTs (Titah-Benbouzid et al., 2023; Rashid et al., 2023c). Over 80% of a TST's blade surface can be affected, leading to performance degradation, increased maintenance costs, and reduced operational lifespan (Titah-Benbouzid and Benbouzid, 2017). This issue causes torque imbalances and increased fatigue loading, threatening the structural integrity of the turbines (Farkas et al., 2022). Additionally, biofouling alters blade surface characteristics, diminishing energy conversion efficiency (Maduka et al., 2023).

Effective prevention of biofouling requires the utilization of reliable detection processes. Detecting biofouling presents challenges due to its unpredictable occurrence and the uncertainties associated with the diverse and dynamic marine ecosystem conditions (Legg et al., 2015). Early detection becomes paramount in this context, as it allows for timely intervention before biofouling significantly impacts the efficiency of marine structures. The key objective is to prevent cumulative

☆ This work is supported by the PIA 3 CMQ Industries de la Mer Bretagne (IndMer), France.
* Corresponding author at: University of Brest, UMR CNRS 6027, 29238 Brest, France.
E-mail addresses: haroon.rashid@univ-brest.fr (H. Rashid), mohamed.benbouzid@univ-brest.fr (M. Benbouzid), yassine.amirat@isen-ouest.yncrea.fr (Y. Amirat), t.berghout@univ-batna2.dz (T. Berghout), hosna.titahbenbouzid@gmail.com (H. Titah-Benbouzid), Abdeslam.Mamoune@univ-brest.fr (A. Mamoune).

https://doi.org/10.1016/j.engappai.2024.109316
Received 17 May 2024; Received in revised form 31 July 2024; Accepted 10 September 2024
Available online 19 September 2024

(a) Initial stage of biofouling. (b) Advanced biofouling stage.

Fig. 1. Visualization of biofouling with significant impact.
Source: Courtesy of Prof. Yusaku Kyozuka (Kyozuka, 2018).

damage by addressing biofouling at its earlier stages as illustrated in Fig. 1. Furthermore, the most effective approach involves the use of environmentally friendly techniques for removing biofouling without causing damage to the protective coatings or paints applied to the structure (Hunsucker et al., 2019). This integrated strategy seeks to maintain optimal performance while minimizing the environmental impact of biofouling prevention measures.

Various methods have been developed to detect biofouling on TSTs (Chin et al., 2017; Rashid et al., 2023a; Zheng et al., 2019). For instance, sparse autoencoding combined with softmax regression has shown promise for real-time monitoring (Zheng et al., 2019), while depthwise separable convolutional neural networks (CNNs) have been utilized to reduce computational complexity while capturing local spatial relationships (Xin et al., 2021). Unsupervised learning techniques have also been explored for classifying ship-hull fouling conditions, offering a more adaptable and scalable solution for marine applications (Wang et al., 2006). Moreover, intelligent image recognition systems using deep learning and transfer learning have been proposed for automated marine fouling detection (Chin et al., 2017). Computer vision and machine learning techniques have also been used to develop effective and affordable monitoring systems for underwater structures (Liniger et al., 2022; Kim et al., 2022; Wang et al., 2006) investigate the use of unsupervised learning neural networks for classifying ship-hull fouling conditions, enhancing the efficiency of fouling detection and classification. Recent advancements include a novel algorithm for enhancing the identification of biofouling in underwater images and a segmentation model for improved feature extraction and fusion (Zhao et al., 2023).

Despite these advancements, there are significant gaps in the current methodologies. Most existing approaches rely heavily on extensive labeled datasets, which are often difficult and expensive to obtain in marine environments. Additionally, the accuracy and reliability of these methods can be hindered by the variability in underwater visibility and the dynamic nature of marine ecosystems (Hurtós et al., 2017; Rashid et al., 2023b). There is a lack of research focusing on early detection and precise classification of biofouling, which is critical for preventing significant impacts on turbine efficiency and maintenance (Song et al., 2019; Loxton et al., 2017; Amiruddin et al., 2024). Many current techniques do not provide the necessary early warning and accurate classification needed for effective maintenance strategies. Efficient biofouling detection also requires timely intervention to prevent significant impacts on turbine efficiency and maintenance costs. However, current techniques often fall short of providing early detection and accurate classification, which are crucial for effective maintenance and operational decision-making.

This study addresses these gaps by proposing a robust methodology employing a soft voting ensemble transfer learning-based approach for the detection and classification of biofouling on TST blades. The key contributions of this study are:

- **Enhanced detection accuracy:** Utilizes pre-trained models to overcome the challenge of limited labeled data, significantly improving the accuracy and efficiency of biofouling detection.
- **Comprehensive biofouling classification:** Implements an ensemble mechanism to classify biofouling into three distinct categories – clean, lightly fouled, and densely fouled – providing detailed insights into biofouling levels.
- **Improved model robustness:** Incorporates data augmentation and pre-processing techniques to enhance model robustness and generalization, ensuring better performance in diverse and dynamic marine conditions.

2. Datasets collection and pre-processing

The biofouling detection method proposed in this study undergoes testing with two distinct video image datasets, which are further elaborated below.

2.1. Shanghai maritime university dataset

The initial dataset, referred to as the Shanghai Maritime University (SMU) dataset, was gathered from the TST platform at the SMU. This platform features a 230W/8 pole pairs direct-drive permanent magnet synchronous generator-based TST, which operates within a water tunnel simulating waves and turbulences (Fig. 2) (Saidi et al., 2020).

The initial dataset comprised 25 images extracted from the recorded videos of the TST. These images were subsequently categorized into three separate sets based on the degree of emulated biofouling they depicted: clean images, lightly fouled images, and densely fouled images (Qi and Wang, 2023). This dataset simulated biofouling textural characteristics by wrapping ropes around the turbine blades, as illustrated in Fig. 3.

2.2. Lehigh university dataset

The Lehigh University (LU) dataset experiments were conducted in the Tidal Turbulence Test facility (T^3F) located at LU in Pennsylvania,

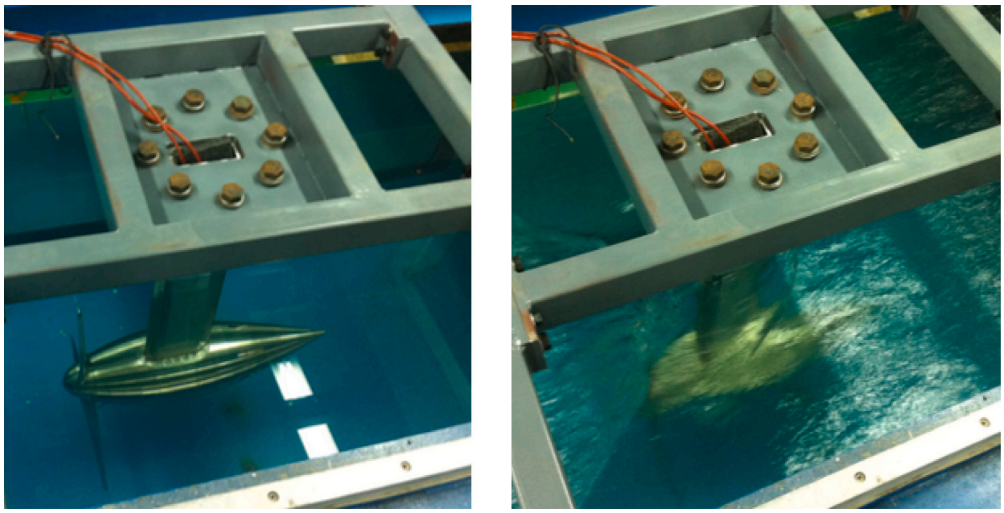


Fig. 2. Illustration of the tidal stream platform at the Shanghai Maritime University.

(a) Healthy blade. (b) Imbalanced blade with attachment. (c) Raw video image.

Fig. 3. Biofouling emulation illustration.

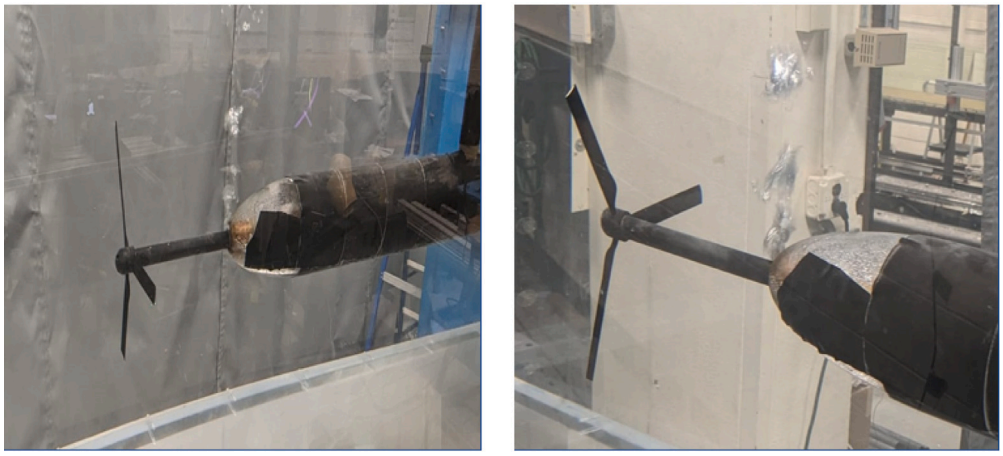


Fig. 4. Tidal Turbulence Test facility (T^3F) with tidal turbine.

(USA) as shown in Fig. 4. The water tunnel features an open surface test section with dimensions of 1.98 m in length, 0.61 m in width, and 0.61 m in height (Fig. 5) (Vinod and Banerjee, 2019). The flow speed in the tunnel is maintained using a 25HP single-stage axial flow pump with a variable frequency driver. The facility can achieve a maximum of 1 m/s inflow speed. However, all the experiments conducted in this work were run at an inflow speed of 0.83 m/s.

A 1:20 scaled, three-bladed tidal turbine with a rotor diameter (D) of 11“ (0.279 m) and a constant chord c = 0.56” (0.014 m) having SG6043 profiled blades was tested as shown in Fig. 6. The turbine design is the same as used in the previous study (Vinod et al., 2021) with a slightly shorter chord length. The rotor plane was fixed at 2D from the test section inlet with its axis along the tunnel center line. An Anaheim Automation (23MDSI) stepper motor provided precise control of the

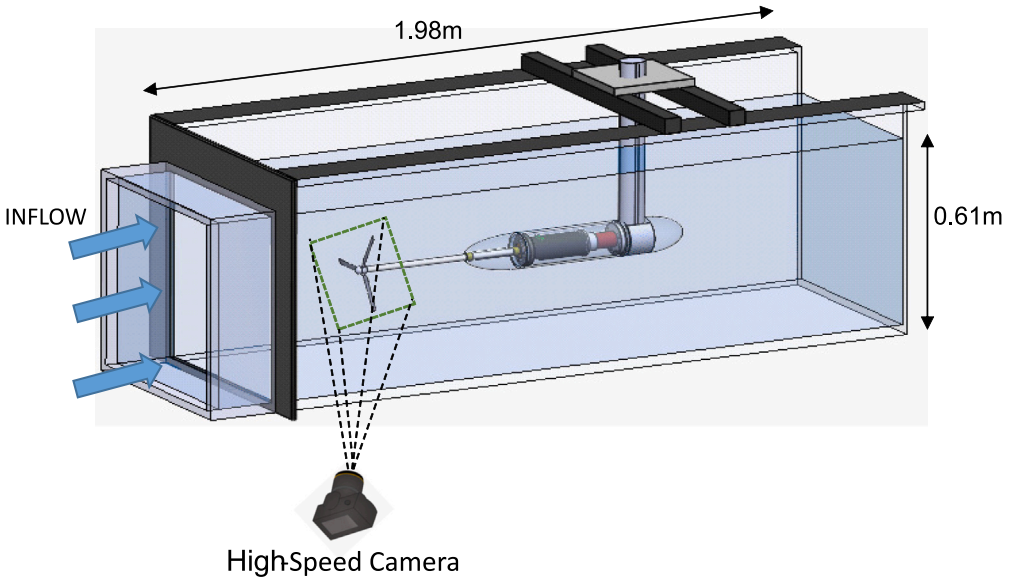


Fig. 5. Illustration of image data collection from the Lehigh University platform.

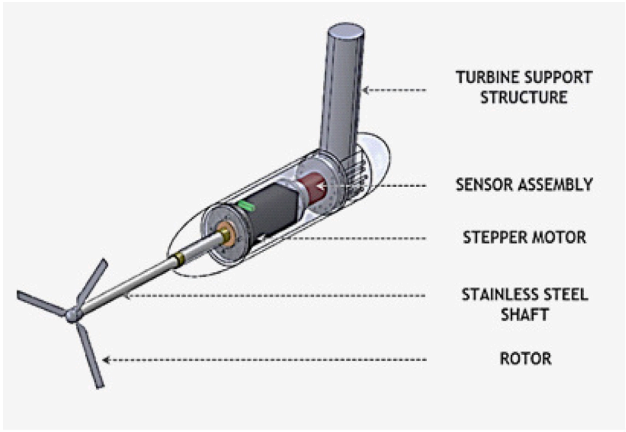


Fig. 6. 1:20 scale tidal stream turbine model used for acquiring the LU dataset.

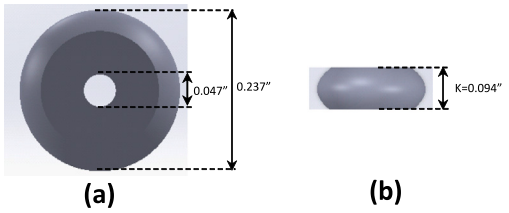


Fig. 7. Schematic view of attachment (beads) for the LU dataset: (a) Top view; (b) Side view.

rotor shaft's rotational speeds. The sensor assembly, located within the nacelle, included a torque (Interface Model-MRT2P) and thrust (SML-25) sensor to measure the loads on the rotor assembly. The nacelle having all the electronics, is sealed with pressurized air around 20 KPa gauge pressure to prevent ingress of water inside it.

To replicate the biofouling in a laboratory, we employed plastic beads as a surrogate for biological organisms commonly associated with fouling in aquatic environments. The plastic beads, outlined in detail in Fig. 7 provide a visual representation from both top and side views.

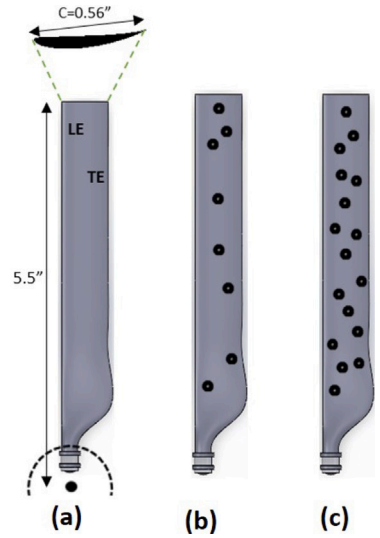


Fig. 8. Schematic of the beads and their distribution on the blade front face: (a) Clean; (b) Lightly fouled; (c) Densely fouled.

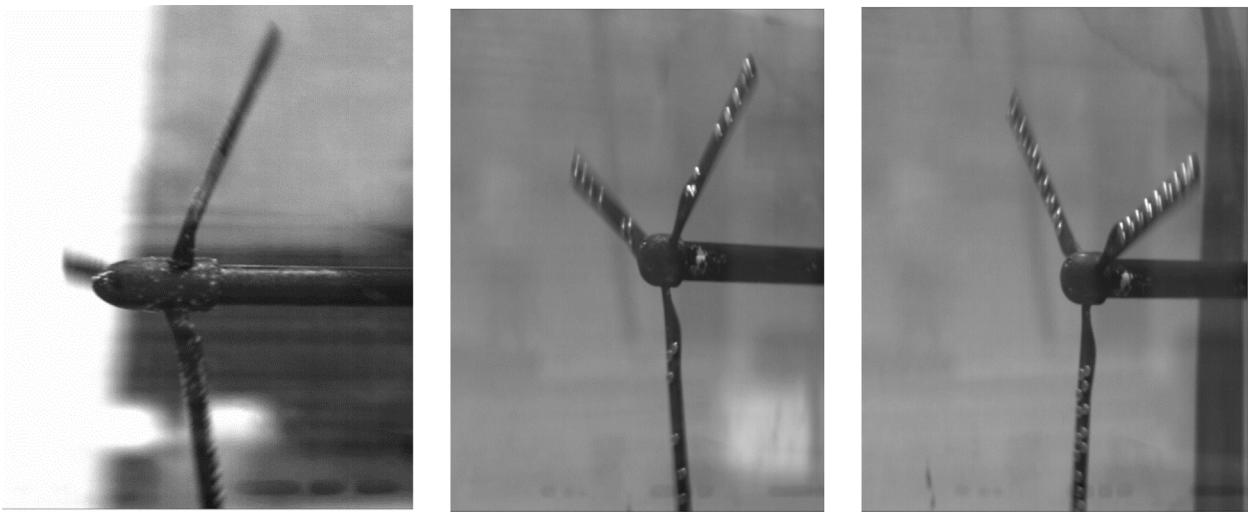
The bead distribution on each of the three blades for two different cases was considered, along with the clean blade as a baseline case. A brief discussion of each case is given below.

Clean blade: This category represents blades without any beads applied, serving as the baseline or control condition (Fig. 8a).

Lightly fouled blades: In this category, each of the three turbine blades has 8 plastic beads glued randomly to their surface. This configuration simulates a moderate surface roughness level due to light biofouling (Fig. 8b).

Densely fouled blades: This category represents a higher level of surface roughness, where each of the three turbine blades has 17 plastic beads glued randomly on their surface. This setup is used to simulate a more severe level of biofouling (Fig. 8c).

In Fig. 5 of our experimental setup, a camera is strategically installed to capture an image dataset. The purpose of this camera is to capture the images of plastic beads on the blades. The acquired image dataset becomes a crucial tool for assessing the varying degrees of fouling on the blades, as demonstrated in Fig. 9. Three distinct



(a) Clean blade. (b) Lightly fouled. (c) Densely fouled.

Fig. 9. LU dataset illustration of the emulation of biofouling.

sets of data have been meticulously collected, representing different conditions: clean blades (Fig. 9(a)), lightly fouled blades (Fig. 9(b)), and densely fouled (Fig. 9(c)) blades. This systematic approach allows us to quantitatively and qualitatively evaluate the extent of biofouling under different scenarios.

2.3. Data pre-processing

To ensure the accuracy and reliability of biofouling detection across varied environmental conditions, several pre-processing steps were applied to the datasets from both SMU and LU. These steps were crucial for standardizing the data and mitigating the effects of environmental factors such as lighting conditions and turbidity.

Image normalization: All images were normalized to a pixel value range of 0 to 1. This standardization process helps stabilize the model training by ensuring that all input images have consistent intensity values, which is essential for effective learning and comparison.

Lighting adjustment: To address the variations in lighting conditions across different datasets and image captures, histogram equalization and contrast adjustment techniques were employed. These methods enhance the visibility of features and correct any inconsistencies caused by varying light intensities during image acquisition.

Turbidity compensation: Given that turbidity can obscure details in underwater images, algorithms were utilized to reduce the impact of visual artifacts and enhance image clarity. This involved filtering out noise and improving contrast to make the biofouling features more distinguishable.

Image enhancement: Additional image enhancement techniques, such as sharpening and denoising, were applied to further improve the quality of the images. These enhancements helped in making subtle biofouling features more detectable and ensuring consistent data quality.

3. Soft voting ensemble transfer learning framework

The proposed soft voting ensemble transfer learning framework for biofouling detection and classification, depicted in Fig. 10, includes several key steps. It starts with data acquisition and pre-processing to ensure high-quality input. Transfer learning is then applied by fine-tuning pre-trained models on the biofouling dataset. Following this, multiple models are trained and validated. The core of the framework is the soft voting ensemble, which aggregates predictions from these models using weighted voting. The performance of the ensemble is assessed and optimized based on metrics such as accuracy and F1-score. Finally, the model is deployed for real-time biofouling detection, with ongoing monitoring for continued effectiveness. Detailed explanations of each step are provided below.

3.1. Data augmentation

To enhance the accuracy of our model, we implemented a technique known as data augmentation during the training phase. This process involves applying various transformations to the images in our training set. As illustrated in Fig. 11, these transformations include rotation, scaling, flipping, zooming, and brightening of the cropped input images. By subjecting each image to these diverse augmentation techniques, we aimed to create a more robust and varied dataset. The rationale behind this approach is to expose the model to a wider range of visual scenarios, simulating different conditions that it might encounter in real-world situations.

3.1.1. Dataset complexity considerations

It is important to recognize that while data augmentation plays a vital role in improving model performance, the complexity of the dataset itself also significantly influences the model's accuracy. In machine learning, the concept of Kolmogorov complexity, which measures the length of the shortest program that can generate a given dataset, suggests that datasets with excessive complexity can pose substantial challenges (Kabir and Garg, 2023). A highly complex dataset may include noise or irrelevant variations that can hinder the model's ability to learn meaningful patterns (Bolon-Canedo and Remeseiro, 2020).

In this study, alongside employing data augmentation, we also explored strategies to manage dataset complexity. By refining the dataset and reducing its complexity through techniques such as filtering and feature selection, we aimed to mitigate noise and focus the model's learning on the most relevant features. This approach helps in simplifying the learning task for the model, potentially leading to improved accuracy and better generalization. Thus, while augmentation enriches the training data, addressing dataset complexity ensures that the model can effectively learn and adapt to the core features of the data.

To further contextualize our approach, we provide a survey of various deep learning tools and techniques designed to handle data scarcity, addressing challenges, and offering solutions. Table 1 summarizes these tools and their respective references.

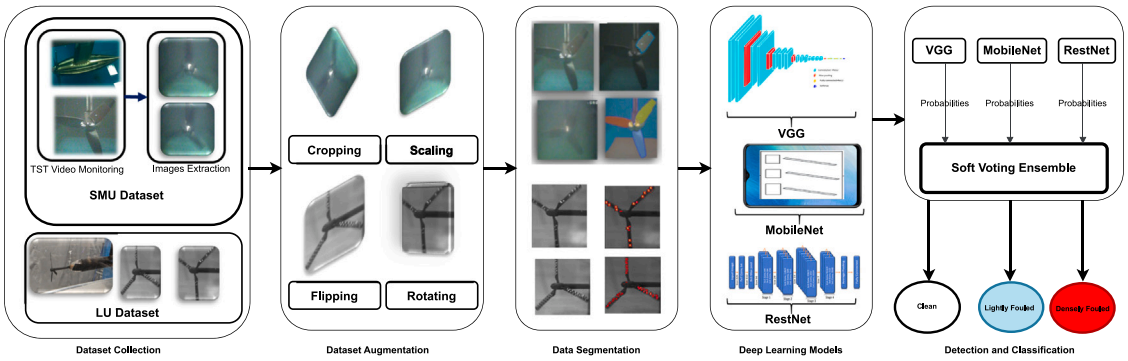


Fig. 10. Framework of the soft voting transfer learning-based approach for biofouling detection and classification.

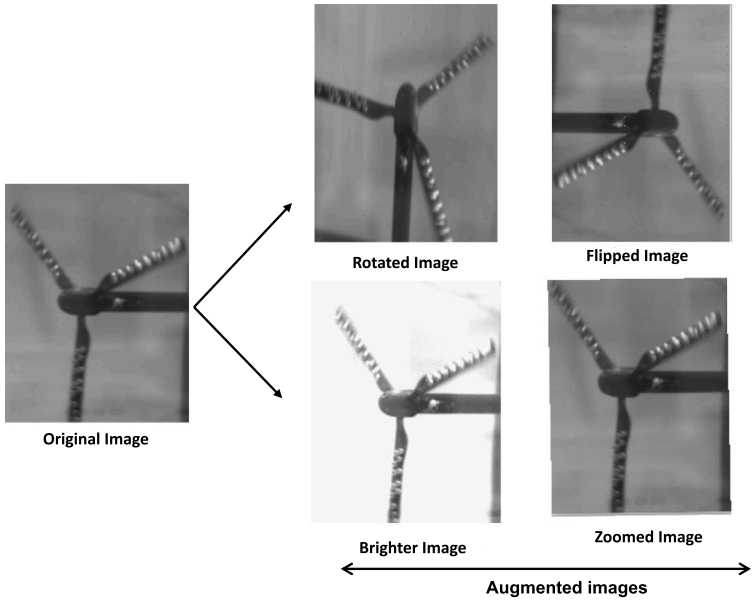


Fig. 11. Augmentation of images.

Table 1
Deep learning techniques handling data scarcity.

Technique	Addressed challenge	Reference
Data Augmentation	Limited dataset size	Shorten and Khoshgoftaar (2019)
Transfer Learning	Insufficient labeled data	Tan et al. (2018)
Generative Adversarial Networks (GANs)	Creating synthetic data	Creswell et al. (2018)
Few-Shot Learning	Learning from few examples	Wang et al. (2020)
Active Learning	Efficient data labeling	Ren et al. (2021)
Self-Supervised Learning	Using unlabeled data	Jing and Tian (2020)
Semi-Supervised Learning	Combining labeled and unlabeled data	Engelen and Hoos (2020)

3.2. Data segmentation and annotation

Data segmentation involves dividing input data into distinct segments or subsets based on specific criteria. In image processing, data segmentation refers to dividing an image into meaningful regions or segments. In this study, data segmentation is utilized to separate the regions of interest (fouling areas) from the image background or other irrelevant parts. The segmentation process in this study follows a systematic approach, as illustrated in Fig. 12, to enhance image interpretability and analysis.

The process begins with grayscale conversion, where images are transformed into grayscale to simplify subsequent operations and focus on intensity values. Following this, Otsu's thresholding is applied to automatically determine the optimal intensity levels, thereby creating a binary representation of the images (Xue and Titterington, 2011). Morphological operations are performed to refine the binary image further. These operations smooth edges and remove noise, improving the quality of the segmented image (Fallahdizcheh et al., 2023).

In the next step, regions connected to the image borders are removed to enhance segmentation accuracy and focus the analysis on the central areas. Connected component labeling is then employed to identify and label distinct regions within the binary image. A colored label image is created to facilitate better visualization, with different colors assigned to individual regions. To provide a clean delineation of boundaries, red rectangles are added around segmented regions, as shown in Fig. 13.

This approach is grounded in supervised learning, where annotated images serve as ground truth for training and validating the model. Annotations are provided by experts and represent the true classifications of different fouling levels. Experts with domain knowledge

Table 2
Selected CNN models and architectures for transfer learning.

Model type	Included architectures
ResNets	ResNet101, ResNet152, ResNet50, ResNet101V2, ResNet152V2, ResNet50V2
VGGs	VGG16, VGG19
MobileNets	MobileNet, MobileNetV2, MobileNetV3-Large, MobileNetV3-Small

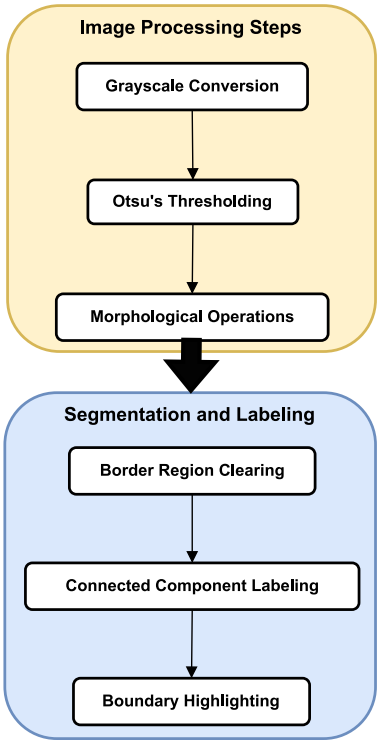


Fig. 12. Flowchart of the data segmentation approach.

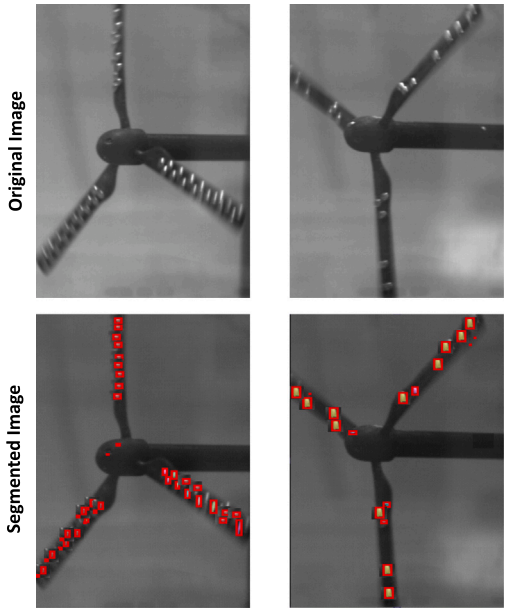


Fig. 13. Image segmentation.

ensure that the classification of fouling levels as clean, lightly fouled, and densely fouled reflects accurate and reliable assessments. These annotations guide the model in learning to recognize and differentiate between various levels of fouling, as the model is trained to identify these categories based on the provided annotations.

Biofouling levels are classified into clean, lightly fouled and densely fouled areas. Clean areas are regions with minimal or no fouling, lightly fouled areas are regions with some visible fouling but not heavily covered, and densely fouled areas are regions with extensive fouling. These distinctions help the model accurately classify new images by focusing on the annotated examples during training.

3.3. Soft voting ensemble transfer learning

In this study, biofouling detection and classification are proposed through a soft voting ensemble approach utilizing well-known CNN models, namely MobileNets (Sandler et al., 2018), ResNets (Hu et al., 2024), and VGGs (Simonyan and Zisserman, 2014). The soft voting ensemble classifier integrates multiple models, with individual decisions being based on probability values to assign class labels to the data. Within the soft voting ensemble framework, predictions are weighted according to the significance of each classifier and subsequently merged to derive the sum of weighted probabilities (Fig. 10).

The three CNN models employed in this study consist of submodels outlined in Table 2. Initially trained on the ImageNet dataset, these models underwent further refinement through transfer learning and hyperparameter tuning to select the optimal hyperparameters and weights for each model.

3.3.1. VGG

VGG consists of two submodels, VGG-16 and VGG-19.

VGG-16: VGG-16 is a widely utilized convolutional neural network (CNN) model designed for image classification tasks (Lo et al., 2019). The model architecture includes 13 convolutional layers and 3 fully connected layers, all utilizing ReLU (Rectified Linear Unit) activation functions. The convolutional layers are organized in a sequence of blocks, each containing multiple 3×3 convolutional layers followed by max-pooling layers, which help in reducing spatial dimensions while retaining important features. VGG-16 is known for its depth and simplicity, and it utilizes a uniform architecture with small convolutional kernels (3×3) and pooling layers. It has a substantial parameter count of 138 million, which contributes to its high computational cost and memory requirements (Rezende et al., 2018). Trained on the ImageNet dataset, VGG-16's robustness allows it to perform effectively even with smaller, domain-specific datasets. The default input size for VGG-16 images is set at 224 by 224 pixels with three color channels, making it suitable for diverse image classification tasks (Likhar and Ridhorkar, 2024).

VGG-19: VGG-19 is an extension of the VGG-16 model with additional convolutional layers, totaling 16 convolution layers and 3 fully connected layers (Kasture et al., 2024). The increased depth allows VGG-19 to capture more complex features from images. Like VGG-16, VGG-19 uses ReLU activation functions and employs a similar architecture design with small convolutional kernels. The 16 convolutional layers are organized in a series of blocks that include 3×3 convolutions and max-pooling operations, which enhance its ability to extract hierarchical features. The model's depth contributes to improved performance in capturing intricate patterns but also increases computational demand. As with VGG-16, VGG-19 is pre-trained on the ImageNet dataset, and the 16 convolution layers are focused on feature extraction, while the 3 fully connected layers handle the classification task (Kordemir et al., 2024).

3.3.2. MobileNets

The MobileNets model is explicitly designed for efficient performance on mobile and embedded devices with constrained computational resources. First introduced in 2017, MobileNets has evolved through versions such as MobileNetV2 and MobileNetV3, each bringing improvements in efficiency and performance. MobileNet employs depthwise separable convolutions, which split the convolution operation into two separate layers: depthwise convolution and pointwise convolution. This design significantly reduces the number of parameters and computational complexity compared to traditional convolutional layers. MobileNetV2 introduced the concept of inverted residuals and linear bottlenecks, enhancing the model's efficiency and performance further. MobileNetV3 optimizes the architecture even more by integrating hardware-aware search and network scaling techniques, making it highly suitable for real-time applications on mobile devices (Quach et al., 2024). The model's compact size and processing speed make it advantageous for deployment in resource-constrained environments, and it performs well in scenarios requiring efficient and swift processing of image data.

3.3.3. ResNets

ResNet, short for Residual Network, represents a significant advancement in deep learning architectures by introducing the concept of residual learning. This approach addresses the challenges associated with training very deep neural networks by using residual blocks with skip connections. These skip connections allow the network to bypass one or more layers, facilitating the training of deeper networks and mitigating the vanishing gradient problem (He et al., 2016). Developed in 2015, ResNet's architecture includes various depths, such as ResNet-50, ResNet-101, and ResNet-152, which correspond to the number of layers in the network. The residual blocks in ResNet use identity mapping and learn residual functions, which improves the model's ability to learn complex representations while maintaining stable training dynamics. This architecture allows for effective performance even with extremely deep networks, making ResNet particularly powerful for tasks requiring robust feature extraction and high accuracy. The introduction of residual learning has greatly enhanced the performance of deep networks and facilitated their application in diverse and complex tasks.

3.4. Transfer learning

Implementing transfer learning involves several stages. Initially, the pixel values of input images are normalized to ensure they fall within the range of 0 to 1. Subsequently, modifications are made to the layers responsible for multi-class classification within the pre-trained models to accommodate binary classification. This adaptation is accomplished by appending a global average pooling layer (Hsiao et al., 2018) and a dropout layer (Khan Pathan and Rana, 2022) after the final convolutional layer of the pre-trained model. Following this, a fully connected layer with a single node is introduced, utilizing a sigmoid function for activation (Eq. (1)).

$$f(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

Here, x represents the input image.

For all models (VGG, MobileNet, and ResNet), the initial layers are frozen to retain the pre-trained weights. The original fully connected layers are replaced with a global average pooling layer, a dropout layer to prevent overfitting, and a fully connected layer with a single node using a sigmoid activation function for binary classification. The training process in this study entails 50 epochs using a specific learning rate. The dataset is trained using the mini-batch gradient descent method, with the selection of the optimal weights based on achieving the highest validation accuracy across all epochs. Optimal hyperparameters, including dropout rate, learning rates, batch size, and sub-models, are determined through hyperparameter tuning. The learning rate is fine-tuned to a lower value to ensure stable convergence and to prevent large updates that could disrupt the pre-trained weights. The dropout rate is optimized to balance regularization and model performance. The batch size is adjusted based on memory constraints and efficient training. The number of epochs is set to 50, with early stopping based on validation accuracy to avoid overtraining. The sub-model exhibiting the highest validation accuracy is selected from each model subset.

Table 3
Summary of datasets.

Dataset	Class distribution		
	Clean	Lightly fouled	Densely fouled
Original SMU	5	10	10
Augmented SMU	380	800	800
Original LU	100	558	542
Augmented LU	490	2525	2475

3.5. Evaluation metrics

A validation set was used in the training procedure to assess each sub-model's performance. For a given sub-model, the model that had the best accuracy (Eq. (2)) on the validation set was considered the optimal one.

$$Accuracy = \frac{TruePositive + TrueNegative}{Total} \tag{2}$$

Normalized by the total number of samples, this accuracy metric considers both the true positive and true negative predictions of the model. The classifier detection results of the validation set, with the four optimal models, include Accuracy (Eq. (2)), Precision (Eq. (3)), Recall (Eq. (4)), and F1-score (Eq. (5)).

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \tag{3}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \tag{4}$$

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

4. Results and discussion

In this section, we present the outcomes of our study to demonstrate the effectiveness of the proposed methods. Our validation process involves the utilization of two distinct datasets, both of which are categorized as multi-class datasets. The utilization of data augmentation increased the total number of images within the SMU dataset to 1980. Subsequently, these images were categorized into three distinct sets: 380 clean images, 800 lightly fouled images, and 800 densely fouled images. Following the implementation of data augmentation, the LU dataset has been expanded to 5490 samples from 1200. The summarized statistics for both the original dataset and the augmented dataset are presented in Table 3.

To ensure a fair representation of images from each set during the training and testing phases, the dataset was further divided in 80:20 ratio. This involved segregating each category of clean, lightly fouled, and densely fouled images into separate training and testing subsets, with 80% of the images allocated to training and 20% to testing. This balanced allocation across the three sets of images played a crucial role in preserving a fair and representative distribution of various conditions during both the training and evaluation procedures.

The soft voting ensemble classifier proposed in this study was coded using Python and TensorFlow. The experiments were conducted on a computer system equipped with an Intel 11th Gen Intel(R) Xeon(R) W-2223 CPU operating at 3.60 GHz and 128 GB RAM. Retraining and hyperparameter optimization followed the techniques outlined in Section 3. To identify the best hyperparameter values for each sub-model (Table 2), a hyperparameter tuning process was carried out.

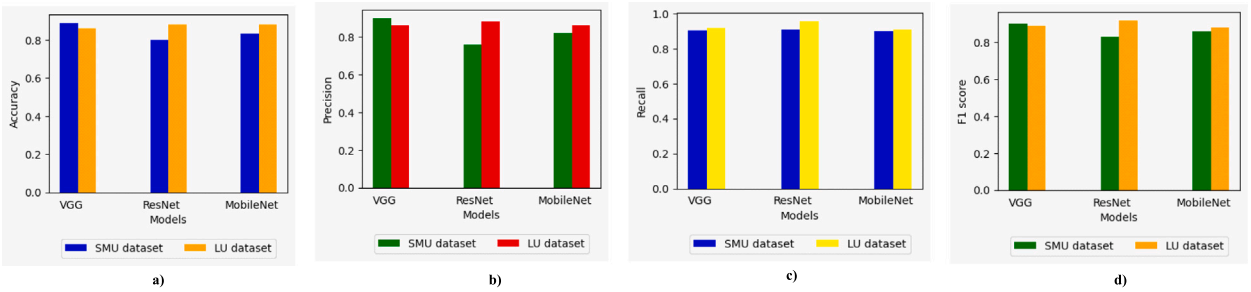


Fig. 14. Different metrics comparison of the classifiers for both datasets: (a) Accuracy; (b) Precision; (c) Recall; (d) F1-score.

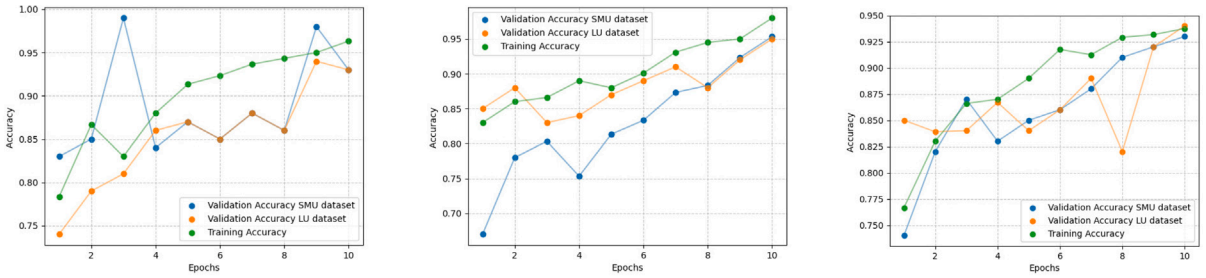


Fig. 15. Learning curves for used models.

Initially, hyperparameters were randomly chosen from the predefined search space, and subsequently, each sub-model underwent training using the selected hyperparameter values.

Fig. 14(a) illustrates the comparison of the accuracies of the MobileNet, VGG, and ResNet models. VGG has the highest accuracy of 0.89 and RestNet has the lowest accuracy of 0.84 at 20% validation data. In Fig. 14(b), the precision of three classifiers is compared. VGG achieves the highest precision at 0.90, while ResNet demonstrates the poorest performance among all classifiers, with a precision of 0.76. In Fig. 14(c), the recall of the employed classifiers is presented. The proposed classifier achieves a maximum recall of 0.91, while RestNet exhibits the highest performance. MobileNet performs slightly lower, attaining a recall of 0.90. In Fig. 14(d), which displays the F1-scores of the classifiers in the study, RestNet outperforms the other models with a maximum F1-score of 0.92. On the other hand, MobileNet shows the lowest F1-score, registering at 0.83.

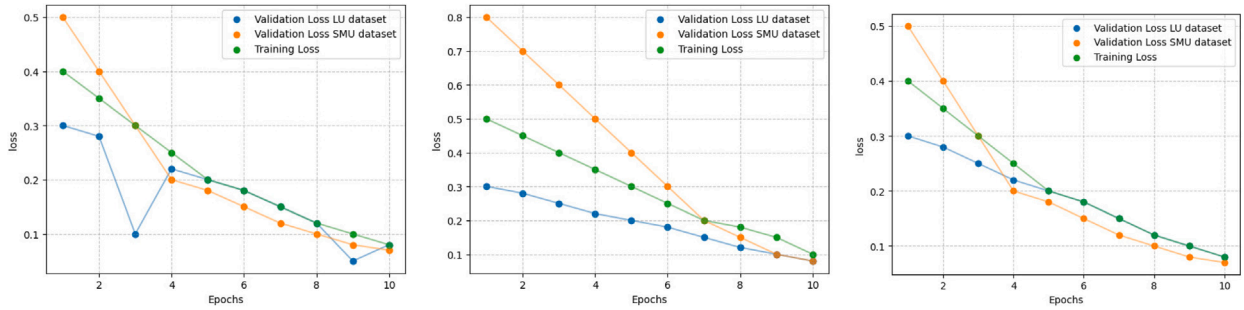
Fig. 15(a), Fig. 15(b), and Fig. 15(c) show the learning curves for MobileNet, RestNet, and VGG classifiers, respectively. These curves show that the training and validation accuracy for both datasets used in this study increased with the number of epochs. The learning curves indicate the models' convergence over training epochs. The training accuracy and validation accuracy increase steadily, showcasing the models' ability to learn effectively from the augmented datasets.

Fig. 16(a), Fig. 16(b), and Fig. 16(c) show the loss curves for MobileNet, RestNet, and VGG classifiers, respectively. These curves show that both validation and training loss decreased as the number of epochs increased. The reduction in training and validation loss further confirms the models' improved performance and generalization capability.

Confusion matrix results for TSTs biofouling detection using the soft voting ensemble classifier for the SMU and LU datasets are shown in Figs. 17(a) and 17(b), respectively. For the SMU and LU datasets, the reported classification accuracy of clean, lightly fouled, and densely fouled occurrences is 83% and 90%, respectively. The importance of dataset scale on classification performance was emphasized by attributing the SMU dataset's inferior accuracy to its smaller size in comparison to the LU dataset.

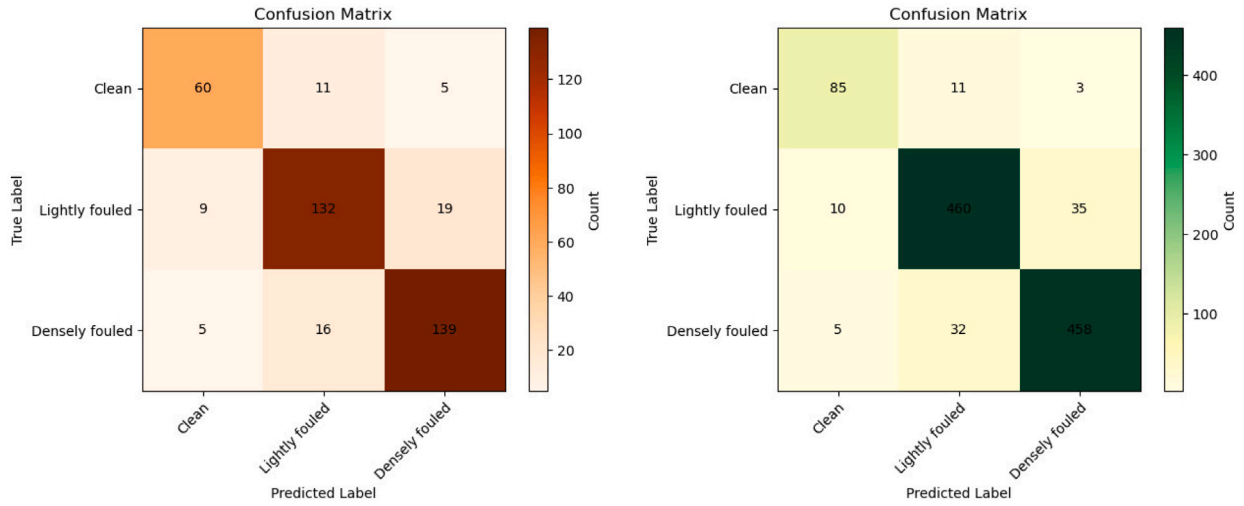
The learning and loss curves (Figs. 15 and 16) indicate the models' convergence over training epochs. The training accuracy and validation accuracy increase steadily, showcasing the models' ability to learn effectively from the augmented datasets. Specifically, for the SMU dataset, the training curve showed a steady decrease in loss with a corresponding increase in accuracy, which is indicative of effective learning. However, the validation curve initially mirrored the training curve but eventually leveled off, suggesting that the model had reached its generalization capacity with the given dataset size. In contrast, for the LU dataset, being larger, the model continued improving for a longer duration before the validation curve leveled off, indicating that the larger dataset helped in achieving better generalization and reduced the risk of overfitting. The reduction in training and validation loss further confirms the models' improved performance and generalization capability. To mitigate overfitting, several strategies were employed, including early stopping to prevent the model from learning noise in the training data, regularization techniques such as L2 regularization and dropout layers, and various data augmentation methods to artificially increase the size and variability of the training dataset. These measures collectively ensured that the models performed well on unseen data.

The comparative analysis of the evaluation metrics (Fig. 14) reveals that VGG consistently outperforms the other models in terms of accuracy and precision, suggesting its robustness in detecting and classifying biofouling conditions accurately. However, RestNet's higher recall and F1-score highlight its balanced performance across different evaluation metrics, indicating its reliability in scenarios where both precision and recall are crucial. Furthermore, the confusion matrices (Fig. 17) provide insight into the classification performance of the models on the SMU and LU datasets. The disparity in performance between the datasets underscores the importance of dataset size and diversity. The larger LU dataset enabled better generalization and higher classification accuracy, whereas the smaller SMU dataset resulted in relatively lower accuracy. This emphasizes the necessity of extensive and varied training data for improving model performance and achieving reliable biofouling detection in real-world applications.



(a) MobileNet (b) RestNet (c) VGG

Fig. 16. Loss curves for used models.



(a) SMU (b) LU dataset

Fig. 17. Confusion matrix for biofouling classification.

4.1. Results analysis

The experiments conducted demonstrate the efficacy of the soft voting ensemble classifier for biofouling detection, highlighting the model's ability to generalize across different scenarios. Despite the computational challenges encountered during training, the study's methodology effectively addresses these issues, leading to robust performance. The real-time biofouling detection potential is particularly noteworthy, considering the advancements in processing speed and latency achieved through the proposed techniques.

4.2. Comparison with existing methods

Compared to existing methods, the soft voting ensemble classifier shows a marked improvement in accuracy and generalization. Traditional approaches often rely on single models, which may not capture the diverse characteristics of biofouling. In contrast, the ensemble method uses the strengths of multiple sub-models, resulting in a more comprehensive understanding and detection capability. Additionally, the hyperparameter tuning process ensures optimal performance, further distinguishing this approach from less adaptive methods.

4.3. Practical implications

The ability to detect biofouling in real time has significant practical implications, particularly for industries reliant on marine equipment and infrastructure. Early and accurate detection can lead to timely maintenance, reducing downtime and associated costs. The study's findings suggest that the proposed methodology can be effectively implemented in real-world scenarios, provided that the computational resources are appropriately managed.

Addressing the key technical challenges in real-time biofouling detection involved ensuring rapid data processing and minimal latency, handling continuous data streams without bottlenecks, and balancing the trade-off between accuracy and speed. The proposed methodology utilized optimized algorithms and model architectures that are computationally less intensive, multi-threading and concurrent processing techniques for efficient data stream handling, and employed hardware acceleration to reduce latency. Computational challenges during model training, such as high memory usage, prolonged training times, and CPU limitations, were addressed through efficient memory management via batch training and garbage collection, distributed training to share the computational load, early stopping and automated hyperparameter tuning to minimize unnecessary computation, and hardware optimization. The significant disparity between dataset sizes posed problems of generalization and bias, which were mitigated by using data augmentation techniques to increase the size of the smallest dataset (SMU), by using transfer learning to pre-train models on the largest dataset (LU) and refining them on the smallest dataset, and by ensuring balanced sampling during training to reduce bias and improve robustness.

5. Conclusion

To ensure the reliable and optimal performance of TSTs, it is of prime importance to treat and mitigate biofouling. For this purpose, it has been proposed a soft-voting ensemble transfer learning-based approach for the detection and the classification of biofouling extent. The proposed video image-based framework integrates critical steps such as data augmentation, data pre-processing, and data segmentation for enhanced performance. Overcoming the scarcity of datasets, we conducted experiments and acquired two datasets: the first from the TSTs platform at the SMU (China), and the second from experiments conducted at the Tidal Turbulence Test facility at LU in Pennsylvania (USA). Using three well-known convolutional neural network models, namely, VGG, ResNet, and MobileNet, trained on these datasets, the proposed soft-voting ensemble classifier demonstrated accurate detection and classification of the TST conditions, achieving an accuracy of 83% for the SMU dataset and 90% for the LU dataset. The lower accuracy observed in the SMU dataset was attributable to its smaller size, highlighting the influence of dataset scale on estimation performance. The achieved results clearly highlight the effectiveness of the proposed video image-based ensemble transfer learning framework for robust biofouling detection and extent estimation in TST systems. Prospective investigations should be oriented towards improving biofouling detection and classification accuracy for prognosis purposes. To this end, it is suggested to explore recurrent expansion algorithms that have recently been introduced to explore deeper representations than ordinary deep networks, dealing with data unavailability, complexity, and drift (Berghout and Benbouzid, 2023).

CRediT authorship contribution statement

Haroon Rashid: Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mohamed Benbouzid:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Yassine Amirat:** Writing – review & editing, Visualization, Validation, Methodology, Formal analysis. **Tarek Berghout:** Writing – review & editing, Visualization, Validation. **Hosna Titah-Benbouzid:** Writing – review & editing, Visualization, Validation, Methodology. **Abdeslam Mamoune:** Writing – review & editing, Visualization, Validation, Supervision, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors would like to thank the Collège Doctoral de Bretagne (ED SPI.bzh) and the University of Brest for their support for the international mobility of Haroon Rashid to Lehigh University (USA). They also acknowledge Prof. Arindam Banerjee for his outstanding advice and supervision during the research stay at the Atlantic Marine Energy Center of Lehigh University, Bethlehem, PA (USA).

References

Amiruddin, W., Hakim, M.L., Firdhaus, A., Purnamasari, D., Tuswan, T., Samuel, S., Mursid, O., 2024. Prediction of the impact of biofouling roughness on a full-scale planing boat performance using CFD. Ocean Eng. 301, 117457.

Berghout, T., Benbouzid, M., 2023. What are recurrent expansion algorithms? Exploring a deeper space than deep learning. In: Computer Sciences & Mathematics Forum, Vol. 7, No. 1. MDPI, p. 10.

Bolon-Canedo, V., Remeseiro, B., 2020. Feature selection in image analysis: a survey. Artif. Intell. Rev. 53 (4), 2905–2931.

Chin, C., Si, J., Clare, A., Ma, M., 2017. Intelligent image recognition system for marine fouling using softmax transfer learning and deep convolutional neural networks. Complexity 2017, http://dx.doi.org/10.1155/2017/5730419.

Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., Bharath, A.A., 2018. Generative adversarial networks: An overview. IEEE Signal Process. Mag. 35 (1), 53–65.

Engelen, J.E.V., Hoos, H.H., 2020. A survey on semi-supervised learning. Mach. Learn. 109 (2), 373–440.

Fallahdizcheh, A., Laroia, S., Wang, C., 2023. Sequential active contour based on morphological-driven thresholding for ultrasound image segmentation of ascites. IEEE J. Biomed. Health Inf. 27 (9), 4305–4316. http://dx.doi.org/10.1109/JBHI.2023.3286869.

Farkas, A., Degiuli, N., Martić, I., Barbarić, M., Guzović, Z., 2022. The impact of biofilm on marine current turbine performance. Renew. Energy 190, 584–595.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778.

Hsiao, T.-Y., Chang, Y.-C., Chiu, C.-T., 2018. Filter-based deep-compression with global average pooling for convolutional networks. In: 2018 IEEE International Workshop on Signal Processing Systems. SiPS, pp. 247–251. http://dx.doi.org/10.1109/SiPS.2018.8598453.

Hu, Y., Deng, L., Wu, Y., Yao, M., Li, G., 2024. Advancing spiking neural networks toward deep residual learning. IEEE Trans. Neural Netw. Learn. Syst.

Hunsucker, K.Z., Gardner, H., Lieberman, K., Swain, G., 2019. Using hydrodynamic testing to assess the performance of fouling control coatings. Ocean Eng. 194, 106677.

Hurtós, N., Palomeras, N., Carrera, A., Carreras, M., 2017. Autonomous detection, following and mapping of an underwater chain using sonar. Ocean Eng. 130, 336–350.

Jing, L., Tian, Y., 2020. Self-supervised visual feature learning with deep neural networks: A survey. IEEE Trans. Pattern Anal. Mach. Intell. 43 (11), 4037–4058.

Kabir, H., Garg, N., 2023. Machine learning enabled orthogonal camera goniometry for accurate and robust contact angle measurements. Scient. Rep. 13 (1), 1497.

Kasture, K.R., Patil, W.V., Shankar, A., 2024. Comparative analysis of deep learning models for early prediction and subtype classification of ovarian cancer: A comprehensive study. Int. J. Intell. Syst. Appl. Eng. 12 (7s), 507–515.

Khan Pathan, S.M., Rana, M.M., 2022. Investigation on classification of motor imagery signal using bidirectional LSTM with effect of dropout layers. In: 2022 International Conference on Advancement in Electrical and Electronic Engineering. ICAEEE, pp. 1–5. http://dx.doi.org/10.1109/ICAEEE54957.2022.9836415.

Kim, B., Kim, H., Han, S., Park, D., 2022. Inspection of underwater hull surface condition using the soft voting ensemble of the transfer-learned models. Sensors 22 (12), 4392. http://dx.doi.org/10.3390/s22124392.

Kordemir, M., Cevik, K.K., Bozkurt, A., 2024. A mask R-CNN approach for detection and classification of brain tumours from MR images. Comput. Meth. Biomech. Biomed. Eng. Imag. Visual. 1–11.

Kyozuka, Y., 2018. Observation of biofouling on two test plates with narrow gap in hirado strait, nagasaki. In: Proceedings of the 2018 IEEE OCEANS - MTS/IEEE Kobe Techno-Oceans. OTO, IEEE, Kobe, Japan, pp. 1–6.

Legg, M., Yücel, M., De Carellan, I.G., Kappatos, V., Selcuk, C., Gan, T., 2015. Acoustic methods for biofouling control: A review. Ocean Eng. 103, 237–247.

Likhar, K., Ridhorkar, S., 2024. Enhancing skin cancer detection: A comparative analysis of models with VGG-16, VGG-19, and inception V3. Int. J. Intell. Syst. Appl. Eng. 12 (10s), 502–514.

Liniger, J., Jensen, A.L., Pedersen, S., Sørensen, H., Mai, C., 2022. On the autonomous inspection and classification of marine growth on subsea structures. In: OCEANS 2022. Chennai, IEEE, pp. 1–7. http://dx.doi.org/10.1109/OCEANSChennai45887.2022.9775295.

Lo, W.W., Yang, X., Wang, Y., 2019. An xception convolutional neural network for malware classification with transfer learning. In: 2019 10th IFIP International Conference on New Technologies, Mobility and Security. NTMS, IEEE, pp. 1–5.

Loxton, J., Macleod, A., Nall, C.R., McCollin, T., Machado, I., Simas, T., Vance, T., Kenny, C., Want, A., Miller, R., 2017. Setting an agenda for biofouling research for the marine renewable energy industry. Int. J. Mar. Energy 19, 292–303.

Maduka, M., Schoefs, F., Thiagarajan, K., Bates, A., 2023. Hydrodynamic effects of biofouling-induced surface roughness–Review and research gaps for shallow water offshore wind energy structures. Ocean Eng. 272, 113798.

Qi, F., Wang, T., 2023. A semantic segmentation method based on image entropy weighted spatio-temporal fusion for blade attachment recognition of marine current turbines. J. Mar. Sci. Eng. 11 (4), 691, 1–14.

Quach, L.-D., Quoc, K.N., Quynh, A.N., Ngoc, H.T., Nghe, N.T., 2024. Tomato health monitoring system: Tomato classification, detection, and counting system based on YOLOv8 model with explainable MobileNet models using grad-CAM++. IEEE Access.

Rashid, H., Benbouzid, M., Amirat, Y., Berghout, T., Titah-Benbouzid, H., Mamoune, A., 2023a. Biofouling detection and extent classification in tidal stream turbines via a soft voting ensemble transfer learning approach. In: IECON 2023- 49th Annual Conference of the IEEE Industrial Electronics Society. pp. 01–06. http://dx.doi.org/10.1109/IECON51785.2023.10312201.

Rashid, H., Benbouzid, M., Titah-Benbouzid, H., Amirat, Y., Berghout, T., Mamoune, A., 2023b. Mapping a machine learning path forward for tidal stream turbines biofouling detection and estimation. In: IECON 2023- 49th Annual Conference of the IEEE Industrial Electronics Society. pp. 1–6. http://dx.doi.org/10.1109/IECON51785.2023.10312077.

Rashid, H., Benbouzid, M., Titah-Benbouzid, H., Amirat, Y., Mamoune, A., 2023c. Tidal stream turbine biofouling detection and estimation: A review-based roadmap. J. Mar. Sci. Eng. 11 (5), 908. http://dx.doi.org/10.3390/jmse11050908.

Ren, P., Xiao, Y., Chang, X., Huang, P.-Y., Li, Z., Gupta, B.B., Chen, X., Wang, X., 2021. A survey of deep active learning. ACM Comput. Surv. 54 (9), 1–40.

Rezende, E., Ruppert, G., Carvalho, T., Theophilo, A., Ramos, F., Geus, P.d., 2018. Malicious software classification using VGG16 deep neural network's bottleneck features. In: Information Technology-New Generations: 15th International Conference on Information Technology. Springer, pp. 51–59.

Saidi, L., Benbouzid, M.E.H., Diallo, D., Amirat, Y., Elbouchikhi, E., Wang, T., 2020. Higher-order spectra analysis-based diagnosis method of blades biofouling in a PMSG driven tidal stream turbine. Energies 13 (11), 2888, 1–18.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C., 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 4510–4520.

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. J. Big Data 6 (1), 1–48.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Song, S., Shi, W., Demirel, Y.K., Atlar, M., 2019. The effect of biofouling on the tidal turbine performance. In: Applied Energy Symposium.

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., Liu, C., 2018. A survey on deep transfer learning. In: Artificial Neural Networks and Machine Learning–ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings, Part III 27. Springer, pp. 270–279.

Titah-Benbouzid, H., Benbouzid, M.E.H., 2017. Biofouling issue on marine renewable energy converters: A state of the art review on impacts and prevention. Int. J. Energy Conver. 5 (3), 67–78.

Titah-Benbouzid, H., Rashid, H., Benbouzid, M., 2023. Biofouling issue in tidal stream turbines. In: Design, Control and Monitoring of Tidal Stream Turbine Systems. IET, pp. 181–204.

Vinod, A., Banerjee, A., 2019. Performance and near-wake characterization of a tidal current turbine in elevated levels of free stream turbulence. Appl. Energy 254, 113639.

Vinod, A., Han, C., Banerjee, A., 2021. Tidal turbine performance and near-wake characteristics in a sheared turbulent inflow. Renew. Energy 175, 840–852.

Wang, P.-F., Lieberman, S., Ho, L., 2006. Unsupervised learning neural network for classification of ship-hull fouling conditions. In: The 2006 IEEE International Joint Conference on Neural Network Proceedings. pp. 4601–4604. http://dx.doi.org/10.1109/IJCNN.2006.247089.

Wang, Y., Yao, Q., Kwok, J.T., Ni, L.M., 2020. Generalizing from a few examples: A survey on few-shot learning. ACM Comput. Surv. 53 (3), 1–34.

Xin, B., Zheng, Y., Wang, T., Chen, L., Wang, Y., 2021. A diagnosis method based on depthwise separable convolutional neural network for the attachment on the blade of marine current turbine. Proc. Inst. Mech. Eng. I 235 (10), 1916–1926. http://dx.doi.org/10.1177/0959651820937841.

Xue, J.-H., Titterington, D.M., 2011. t -Tests, F -tests and Otsu's methods for image thresholding. IEEE Trans. Image Process. 20 (8), 2392–2396. http://dx.doi.org/10.1109/TIP.2011.2114358.

Zhao, W., Han, F., Qiu, X., Peng, X., Zhao, Y., Zhang, J., 2023. Research on the identification and distribution of biofouling using underwater cleaning robot based on deep learning. Ocean Eng. 273, 113909.

Zheng, Y., Wang, T., Xin, B., Xie, T., Wang, Y., 2019. A sparse autoencoder and softmax regression based diagnosis method for the attachment on the blades of marine current turbine. Sensors 19 (4), 826. http://dx.doi.org/10.3390/s19040826.