Ocean Engineering 263 (2022) 112297

Contents lists available at ScienceDirect

Ocean Engineering

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

Development of a time series imaging approach for fault classification of marine systems

Christian Velasco-Gallego *, Iraklis Lazakis

Department of Naval Architecture, Ocean and Marine Engineering, University of Strathclyde, 100 Montrose Street, G4 0LZ, Glasgow, United Kingdom

ARTICLE INFO

Keywords:
Fault diagnosis
Time series imaging
Fault classification
Anomaly identification
Shipping industry
Marine systems
Prognostics and health management (PHM)

ABSTRACT

As part of any Prognostics and Health Management (PHM) system for the shipping industry, the determination of the current health of marine systems is fundamental. As such, diagnostic analytics is performed; a process that is typically constituted by fault detection, fault isolation, and fault identification. Although some efforts have been made to distinguish the faults and malfunctions (fault detection) that can occur in marine systems, the implementation of fault identification to provide a description of any considered fault type and its nature is still an unexplored area due to the lack of fault data. To overcome this, a methodology for the identification of anomalies in marine systems is presented in this paper. The proposed approach aims to analyse the implementation of time series imaging through the application of the first-order Markov chain in tandem with an analysis of both ResNet50V2 and Convolutional Neural Networks (CNNs) as part of the image classification task. To highlight the performance of this methodology, anomalies have been simulated considering the power parameter of a diesel generator. Results demonstrated the potential of time series imaging and image classification approaches, as the Markov-CNN achieved an accuracy of 95% when performing the fault classification task.

1. Introduction

There is an undeniable need to continue investing in technology within the maritime sector, as it has demonstrated its potential to enhance the safety conditions. The increasing level of information gathering and the enhancement of communication technologies on ships through the utilisation of sensors and Artificial Intelligence (AI) can enable better coordination between ships by enhancing the decision-making processes. An aspect that is fundamental in a sector whereby 75%–96% of accidents are attributed to human action owing to fatigue or bad judgement (Allianz Global Corporate and Specialty, 2012). Accordingly, an increase in investment is expected centred on four main groups that define the smart shipping industry: 1) smart port, 2) autonomous vessels, 3) on-board technologies, and 4) professional services technologies (London Economics et al., 2021).

Smart ports have already been an area of extensive research, one which relies on automation, big data, AI software systems, digital twins, and alternative energy. The port of Valencia, for example, demonstrates success in relation to digital twins. A method employed to dynamically track the port's lightning system (Wang et al., 2021b). By contrast, while significant advancements have since been perceived, autonomous ships are not yet as well established as smart ports technology due to challenges in technology development as well as both the administrative and safety requirements that are required for testing. However, it is expected that IMO level 3/4 autonomous ships will be completed in the next 5 or more years. Meanwhile, numerous efforts by academia have been performed in the analysis of such technologies. For instance, Bolbot et al. (2021) introduced a novel hybrid, semi-structured process for identifying and ranking hazardous scenarios, which was applied to assess the safety of an autonomous inland waterways ship at a preliminary design phase.

Another critical area for enabling smart shipping pertains to the on-board technologies, those which assist in safe navigation (Uyanik et al., 2021), ship performance (Bui and Perera, 2021; Farag and Ölçer, 2020), maintenance (Cheliotis et al., 2022; Han et al., 2021), connectivity (Bolbot et al., 2020), and alternative propulsion (Zhang et al., 2021) through the implementation of AI and vessel optimisation systems. Although a certain effort has been made to advance towards these on-board technologies, there is no real trend in the direction of technology development. All these preceding advancements towards the establishment of smart shipping as a technologically advanced industry also enable the development of professional services, which can

* Corresponding author.
E-mail address: christian.velasco@strath.ac.uk (C. Velasco-Gallego).

https://doi.org/10.1016/j.oceaneng.2022.112297
Received 29 March 2022; Received in revised form 1 August 2022; Accepted 11 August 2022
Available online 29 August 2022

incorporate novel technologies requiring well trained and qualified personnel.

Of all possible smart technologies being analysed in this respect, special attention is given to smart maintenance in this study, as further research in ship operations related to repairing and maintenance is needed. Although potential applications of AI within the shipping sector have been identified as performing the most efficient and economical maintenance practices (Department for Transport, 2019), and also present a positive impact on both the safety and security of the personnel, there is not yet a clear technological solution for such a matter. However, certain efforts have been perceived with regards to the development of Prognostics and Health Management (PHM) for shipping systems.

Although some efforts have been made to detect the faults and malfunctions (fault detection) that can occur in marine systems, the implementation of fault identification to provide a description of any considered fault type and its nature is still an unexplored area due to the lack of fault data. Most of the identified methodologies that introduced novel approaches for fault classification of marine systems presented an analysis of a version of a Support Vector Machine (SVM) model. With regards to Deep Learning (DL) methodologies, only two studies have been identified, which implemented either deep neural network or recurrent networks. However, analogous industries have widely analysed the utilisation of DL methodologies for fault classification (Li et al., 2019; Zhao et al., 2019b). Analogous industries have also explored the introduction of time series imaging in the fault classification process, an element that has presented promising results in identifying fault patterns that could not be perceived when analysing the original time series data. Accordingly, to overcome such a fact, a methodology for the identification of faults of marine systems is presented. The proposed approach aims to analyse the implementation of time series imaging through the application of the first-order Markov chain in tandem with an analysis of both ResNet50V2 and CNN architectures for image classification.

The following paragraphs are structured as follows. Section 2 presents a literature review with regards to the current research performed in both fault classification within the maritime industry and the application of time series imaging in analogous sectors. Section 3 describes the proposed methodology. Section 4 reflects on the results obtained after implementing the proposed methodology through a case study and a comparative analysis. Lastly, in Section 5 the conclusions and future work are outlined.

2. Literature review

Wang et al. (2020) presented a fault diagnosis framework constituted by an unsupervised (*k*-means algorithm) and supervised phase (Back Propagation (BP) neural network). To validate the performance of the framework, a case study on a marine diesel engine was performed, as it is considered a critical system. The established fault diagnosis scheme demonstrated high accuracy under both working and high-pressure oil pump wear exhaust valve leakage conditions, although the diagnostics of both the nozzle carbon deposition and piston ring damage conditions required an enhancement.

Cai et al. (2017) introduced another fault diagnosis framework for marine diesel engines. The first step was the structuring of the diesel engine system into subsystems to reduce the complexity of the analysis. Accordingly, the 1) fuel, 2) lubrication, 3) intake and exhaust, and 4) cooling systems were identified. Then, a classification model based on SVM was established to perform operating state monitoring and fault diagnosis. To finalise, the association rule mining algorithm was considered to analyse the relationships among the fault characteristics at distinct levels. A historical fault database was implemented for such a purpose. Hou et al. (2020) also proposed a fault diagnosis framework for a marine diesel engine. Specifically, the fuel oil supply system of the engine was considered. Analogous to Cai et al., (2017), the classification performance of the SVM model was analysed.

Senemmar and Zhang (2021) developed a new deep learning-based framework for fault detection, classification, and location identification simultaneously in shipboard power systems. A total of three distinct methodologies were introduced: 1) deep neural network, 2) gated recurrent unit, and 3) LSTM. To implement the case study, fault data from an 8-bus shipboard power system were simulated. A 99% accuracy was obtained, determining that the GRU-based model as the most effective DL model. The DNN model was the one that presented less accurate results.

Tan et al. (2020) investigated the performance of the following one-class classifiers: One Class Support Vector Machine (OCSVM), Support Vector Data Description (SVDD), Global k-Nearest Neighbors (GKNN), Local Outlier Factor (LOF), Isolation Forest (IF), and Angle-Based Outlier Detection (ABOD). To that end, a real-data validated numerical simulator developed for a Frigate characterised by a combined diesel-electric and gas propulsion plant was utilised for a case study implementation. Based on the outlined results, the authors sorted the performance of the six analysed algorithms as follows: ABOD $>$ OCSVM $\approx$ SVDD $>$ GKNN $>$ IF $\approx$ LOF. Tan et al. (2021) presented an analogous comparative study, although the topic of study in this case was multi-label classification for simultaneous fault diagnosis. The comparative study consisted of analysing a total of five models: 1) Binary Relevance (BR), 2) Classifier Chains (CC), 3) multi-label k-nearest neighbour (MLKNN), 4) Binary Relevance k-nearest neighbour (BRKNN), and 5) multi-label twin support vector machine (MLTSVM). Analogous to Tan et al. (2020), a dataset generated from a real data validates simulator of a Frigate was considered for the performance of the case study. Based on the outlined results, it was determined that BR outperformed the remaining analysed methods.

Of all the methodologies implemented, four of the six identified studies presented an analysis of a version of SVM. The remaining two studies referred to DL approaches, in which the application of either deep neural networks or recurrent neural networks have been considered to some extent. However, although analogous industries have exploited the potential of powerful methods of image processing and time series imaging for fault detection and diagnostics (Zio, 2022), there is no evidence that such practices have been analysed and formalised within the shipping sector.

For instance, Fahim et al. (2021) proposed a self-attentive weight--sharing capsule network (WSCN) to perform both fault detection and classification in the transmission line domain. Prior to the implementation of WSCN, the authors encoded the time-series signal into an image by implementing the Gramian Angular Field (GAF) algorithm. The authors highlighted that transforming the time-series signal into an image is significant in revealing certain fault features and patterns that cannot be extracted from the original time-series signal. A Western-System-Coordinating-Council WSCC 9-bus and 3-machine test model modified with the series capacitor was analysed to determine the robustness of the self-attention WSCN.

Fahim et al. (2021b) introduced a unified unsupervised learning framework for short circuit fault analysis of a power transmission line. Analogous to Fahim et al. (2021), GAF was applied to transform the time-series oscillographs into images. A stacked denoising-autoencoder was integrated and modelled to guarantee the robustness of the framework against noise. Field data was considered for a case study with three types of fault classification results. Fahim et al. (2021c), Fahim et al. (2020), and Fahim et al. (2020b) also implemented GAF for image representation of sampled signals. Such images would then be considered as inputs of the proposed model.

Yao et al. (2020) proposed a framework for fault diagnosis with Full-scope Simulator based on the State Information Imaging (FDFSSII). FSFSSII aimed to construct a series of grey images that presented the operating transient (both normal and fault condition) according to the real time monitoring data. A case study based on the nuclear plant-wide fault diagnosis system was presented.

Kiangala and Wang (2020) developed a classification model based on time-series imaging and CNN. The image representation was obtained by implementing GAF. A case study based on data collected from the conveyor system was applied.

Of all the studies identified about time series imaging applied to fault classification, only two distinct approaches could be perceived: 1) GAF, and 2) FSFSSII. Specifically, all studies except one, which implemented the FSFSSII approach, considered GAF for encoding the signals into images. Thus, there is a need for exploring new methods for image representation from time series, as time series imaging has demonstrated their ability of discovering fault features and patterns that cannot be obtained from the original version of the time series. Accordingly, the main purpose of this study is to explore the first-order Markov chain model as a potential approach for performing time series imaging applied to fault classification. As a classifier, both a ResNet50V2 network and a CNN are analysed. ResNet50V2 has demonstrated its capability of extracting deep features (Rahimzadeh and Attar, 2020), whilst CNN is probably the most widely used method when dealing with images. Specifically, it has been widely implemented for image recognition tasks (Nisha and Meeral, 2021). To the best of the authors' knowledge, there is no evidence that either time series imaging or both ResNet50V2 and CNN have been considered to perform the fault classification task within the sector. Thus, the contribution of this paper can be summarised as follows:

- An image classification approach is considered for performing fault classification. Based on the available literature review presented, there is no evidence that such approaches have been considered within the shipping industry.
- The analysis of the first-order Markov chain model as a time series imaging technique. To the best of the author's knowledge, only GAF and FSFSSII methods have been considered when performing a fault classification task.
- The development of an overall framework that considers both the formalisation of a fault classification approach and a simulation module to address the lack of fault data and labelled data within the shipping industry.
- A validation process is performed, in which some widely used methods that have not been analysed for performing fault classification within the shipping industry are performed, such as 1D-CNN and GAF-CNN.

3. Methodology

Having explored the current research regarding fault classification within the shipping sector and determined the contribution introduced in this study based on the gaps identified, the proposed methodology, represented graphically in Fig. 1, is introduced. The first phase refers to the pre-processing of the input time series. Phase 2, named Anomalies Simulation, aims to simulate non-operational states to perform the classification task due to the lack of fault data within the shipping industry. The third phase refers to the encoding of time series sequences into images through the application of the first-order Markov chain model. Subsequently, in order to perform the fault classification task, the fourth phase is implemented, in which image classification is applied by applying the deep learning architectures ResNet50V2 and CNN. To validate the fault classification performance of the proposed approach, phase 5 is presented, in which a comparative study is also performed in order to evaluate the effectiveness of the proposed method with regards to other widely applied fault classification techniques.

3.1. Data pre-processing

As presented in analogous studies on marine machinery systems, the data preparation refers to the data imputation, steady states, and outliers' identification. The readers are referred to Velasco-Gallego and Lazakis (2022, 2021, 2020) for more details.

The data imputation step is performed by imputing the missing values with the method that outperforms within the comparative methodology, which is constituted by both univariate and multivariate imputation techniques. For identifying the distinct operational states, a novel approach based on both first-order Markov-chain and connected component analysis is implemented. With regards to the outliers' identification phase, the sequences considered by this study have been analysed heuristically and are based on the results obtained in the steady

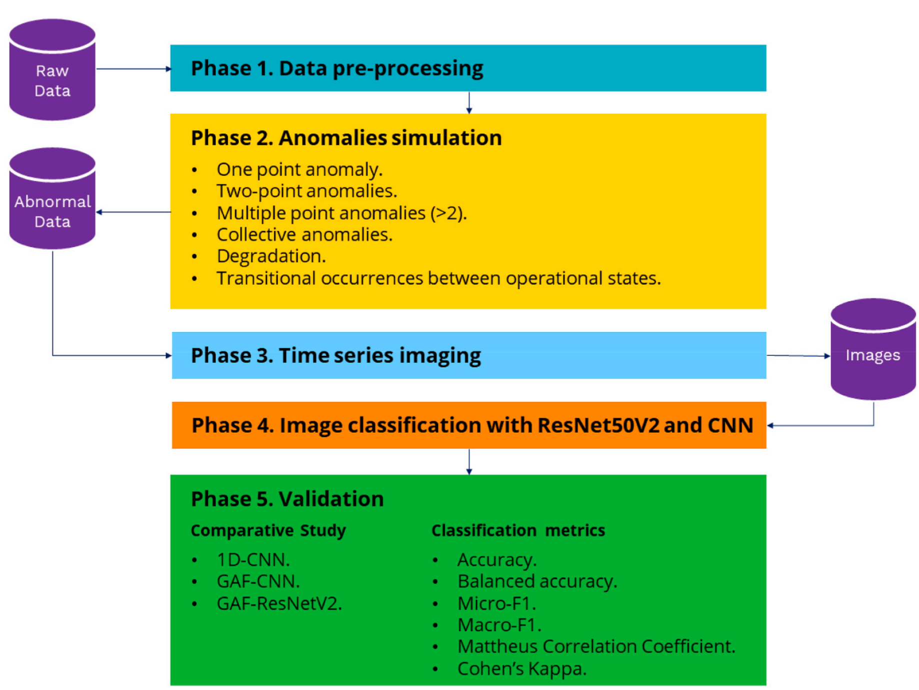

Fig. 1. Graphical representation of the proposed methodology.

states identification section in order to avert their occurrence. Therefore, it is assumed that all the sequences presented in this study as a case study do not contain abnormal instances. To finalise, data normalisation is also applied, and the sliding window algorithm is considered for the generation of sequences.

3.2. Anomalies simulation

Due to the lack of fault data and complete degradation data, a common approach to validate the performance of fault diagnosis and prognosis frameworks is the generation of synthetic data through, for instance, the utilisation of physical model simulation methods (Zhang et al., 2022).

Accordingly, to validate the proposed methodology, a total of six data patterns that do not refer to normal steady operational conditions have been simulated based on the detection of such patterns in analysed marine systems datasets. These are identified as point anomalies: 1) only one point anomaly is presented in the sequence, 2) two-point anomalies are identified in the sequence, 3) multiple point anomalies ($>$2) can be observed in the sequence, 4) collective anomalies, 5) degradation sequences, and 6) transitional occurrences between operational states.

Point anomalies refer to those instances that differ from others with regards to their attributes (Ramchandran and Sangaiah, 2018). For instance, an abrupt change in a steady operational state is considered as a point anomaly. To generate such anomalies, two aspects need to be determined. The first refers to the anomalous ratio, which establishes the intensity of the abnormality. To that end, a ratio between pre-defined minimum and maximum intensity thresholds is selected at random. The second aspect aims to determine the position within the sequence that the anomaly occurs. This aspect is also determined randomly based on the size of the analysed sequence. Such an approach is conducted for the first three defined scenarios (only one point anomaly is presented in the sequence, two-point anomalies are identified in the sequence, and multiple point anomalies can be observed in the sequence). Thus, the difference between these scenarios rests on the number of point anomalies presented in the sequence.

The fourth category, collective anomalies, are usually formed due to a combination of numerous instances. An example of this is a high variability in the exhaust gas outlet temperature parameter of the turbocharger in a steady operational state context. To simulate these collective anomalies, noise is injected by considering different Gaussian distributions of various mean levels, such as those analogously performed by Zhao et al. (2019). Specific attention is given to degradation patterns due to their criticality in performing the prognosis stage adequately. Therefore, although considered a collective anomaly, it is analysed as another distinct category in this study. Accordingly, an exponential model with Brownian motion is considered to simulate degradation patterns due to its effective universality in machinery when reflecting the characteristics of accelerated fault degradation in engineering (Li et al., 2021). It is considered that the degradation process of an item of marine systems can be described as a stochastic process $X(t)$, $t \geq 0$, $X(t)$ being a condition indicator of the item being analysed at time t. A representation of $X(t)$, if an exponential model is considered, is described hereunder.

$$X(t) = \theta' \exp\left(\left(\beta' - \frac{\sigma^2}{2}\right)t + \sigma B(t)\right), \tag{6}$$

where θ' and β' are random parameters representing the individual differences of components, σ is a deterministic parameter representing the increasing random error, and $B(t)$ is a standard Brownian Motion, which represents the stochastic dynamics in the degradation process.

The last scenario refers to transitional occurrences between operational states that occur due to, for instance, environmental situations or variations in the operating condition (Theotokatos et al., 2020). The adequate identification of such states is of preeminent importance to ensure both computational efficiency and model effectiveness (Velasco-Gallego and Lazakis, 2022). To generate the distinct steady states within the sequence, such a sequence is divided into two sub-sequences. The instance from which the division is initiated is selected randomly. Then, the values of the instances of one of the sub-sequences are increased to create the distinct states. This increment is also selected at random by considering two intensity thresholds.

3.3. Time series imaging

As comprehensively described in the literature review section, only two methods have been identified for encoding original time series into images for performing fault classification, none of them being analysed within the shipping sector. These are GAF and FSFSSII. However, despite its distinct application, the first-order Markov chain has previously been considered for the application of time series imaging within the shipping industry. Specifically, it was implemented in tandem with connected component analysis for the identification of steady states (Velasco-Gallego and Lazakis, 2022). Moreover, although the process of time series imaging was not considered, the first-order Markov chain has been successfully applied for data imputation (Velasco-Gallego and Lazakis, 2021), for instance. Accordingly, this study also analyses the first-order Markov chain model by estimating the transition matrix for image representation.

To estimate such a matrix, the definition of the discrete time stochastic process is considered. A discrete time stochastic process, $(X_n)_{n\in\mathbb{N}}$, which takes values in a finite set S, is considered to have the Markov property if the probability distribution of X_{n+1} at time $n+1$ only hinges on the previous state X_n at time n, and not on all the past values of X_k for $k \leq n-1$. Thus,

$$\mathbb{P}(X_{n+1} = j | X_n = i_n,\ X_{n-1} = i_{n-1},\ \ldots,\ Z_0 = i_0) = \mathbb{P}(Z_{n+1} = j | Z_n = i_n) = p(i,j) \tag{1}$$

where $i_0,\ i_1,\ \ldots,\ i_n, j \in S$.The probability $p(i,j)$ indicates the likelihood that the previous state i is followed by the current state j. All the possible transition probabilities of a process can be collected in a rxr matrix, where each (i,j) entry P_{ij} is $p(i,j)$,

$$\boldsymbol{P} = \left(P_{ij}\right)_{1\leq i,\, j\leq r} = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,r} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ p_{r,1} & p_{r,2} & \cdots & p_{r,r} \end{pmatrix} \tag{2}$$

and that satisfies

$$0 \leq P_{ij} \leq 1, \quad 1 \leq i,j \leq r, \tag{3}$$

$$\sum_{j=1}^{r} P_{ij} = 1, \quad 1 \leq i \leq r. \tag{4}$$

3.4. Image classification with ResNet50V2 and Convolutional Neural Networks (CNNs)

One of the two networks considered for analysis to perform image classification in this study is the ResNet50V2 network, as it has demonstrated its capability of extracting deep features (Rahimzadeh and Attar, 2020). Furthermore, such a network has not been applied within the maritime industry for image classification purposes to the best of the authors' knowledge.

ResNet50V2 is a type of deep residual network proposed by He et al. (2016b). Deep residual networks, a.k.a. ResNets (He et al., 2016), consists of many stacked "Residual Units", which can be expressed in a general form as presented in Eq. (5):

$$\boldsymbol{y}_l = h(\boldsymbol{x}_l) + \mathscr{F}(\boldsymbol{x}_l,\ \mathscr{W}_l), \boldsymbol{x}_{l+1} = f(\boldsymbol{y}_l), \tag{5}$$

where x_l and x_{l+1} are input and output of the l-th unit, and $\mathscr{F}$ is a residual function. $h(x_l) = x_l$ is an identity mapping and f is a ReLU (Nair and Hinton, 2010) function. The essence of ResNets is to learn the additive residual function $\mathscr{F}$ respecting $h(x_l)$ by attaching an identify skip connection or "shortcut". The architecture of ResNet50 is presented in Fig. 2. ResNet50V2 is an enhancement of ResNet50 in which a new residual unit has been introduced to both facilitate easier training and enhance generalisation.

Due to the lack of fault data within the shipping industry, the ResNet50V2 network has been pretrained by utilising the popular dataset ImageNet, which presents more than 1000-class single labels (Russakovsky et al., 2015).

The second analysed method is the CNN, which is a type of feed-forward artificial Neural Networks (NNs) that is constituted by a feature extraction step and either a classification or a regression task. As the main objective of this study is to develop an approach for fault classification, only the classification task is considered in this context.

The first stage, feature extraction, is comprised of both convolutional layers and pooling layers. The convolutional layer is usually also referred to as the main block of CNN models. This consists of a set of filters, which are learnt throughout the training process, that convolve with the image and generate a feature map. Specifically, the filter slides over the entire image so that the dot product between each element of both the filter and the input can be estimated at every spatial position. To reduce the dimension of the resulting feature map, a pooling layer is usually introduced after the application of a convolutional layer. Although a loss of information can be perceived by applying such layers, they assist in averting overfitting and reducing the computational cost. The pooling task is performed by sectioning the input into non-overlapping rectangular subregions so that information from each subregion can be extracted. For this inquiry the max pooling layer is implemented.

The second stage refers to the classification task, implemented by the utilisation of fully connected layers. Such layers apply high-level logical operations by considering features from preceding layers. The output of the final layer is a n dimensional vector, n being the total number of classes being considered.

This step is performed by the implementation of the Python libraries Tensorflow and Keras.

3.5. Validation

To complement the validation by utilising simulated fault and non-operational data, a comparative study is implemented to determine the effectiveness of the proposed model based on widely used approaches. The first model considered is the 1D-Convolutional Neural Network (CNN) is also applied, as versions of such a model has presented promising results when dealing with time series data in analogous tasks, such as when predicting the Remaining Useful Life (RUL) (Yao et al., 2021).

In addition, to assess the performance of the first-order Markov chain model as a time series imaging method, the proposed methodology (Markov-ResNet50V2) and the CNN model (Markov-CNN) are modified to present the GAF as the time series imaging method (GAF-ResNet50V2, and GAF-CNN). To encode time series into images by implementing GAF, the *pyts* package is utilised (Wang and Oates, 2015).

Based on comprehensive reviews of classification metrics, such as the one performed by Grandini et al. (2020), and to adequately assess the models included in the comparative study, a total of six metrics has been selected: 1) accuracy, 2) balanced accuracy, 3) Micro F1, 4) Macro F1, 5) Mattheus Correlation Coefficient (MCC), and 6) Cohen's Kappa.

Prior to the definition of such metrics, the confusion matrix needs to be defined, as some of the metrics are computed based on such a concept. This matrix can be defined as a cross table that describes the number of occurrences between two rates (true/actual classification and predicted classification). A diagram representing a confusion matrix for multi-class classification is presented in Fig. 3.

Based on this concept, the first metric, accuracy, is defined. This is probably the most popular metric when addressing the multi-class

layer name	output size	50-layer
conv1	112x112	7x7, 64, stride 2
Conv2_x	56x56	3x3 max pool, stride 2
		$\begin{bmatrix} 1\text{x}1, 64 \\ 3\text{x}3, 64 \\ 1\text{x}1, 256 \end{bmatrix} \text{x } 3$
Conv3_x	28x28	$\begin{bmatrix} 1\text{x}1, 128 \\ 3\text{x}3, 128 \\ 1\text{x}1, 512 \end{bmatrix} \text{x } 4$
Conv4_x	14x14	$\begin{bmatrix} 1\text{x}1, 256 \\ 3\text{x}3, 256 \\ 1\text{x}1, 1024 \end{bmatrix} \text{x } 6$
Conv5_x	7x7	$\begin{bmatrix} 1\text{x}1, 512 \\ 3\text{x}3, 512 \\ 1\text{x}1, 2048 \end{bmatrix} \text{x } 3$
	1x1	average pool, 1000-d fc, softmax
FLOPs		$3.8\text{x}10^9$

Fig. 2. ResNet50 architecture for ImageNet (adapted from Rahimzadeh and Attar, 2020).

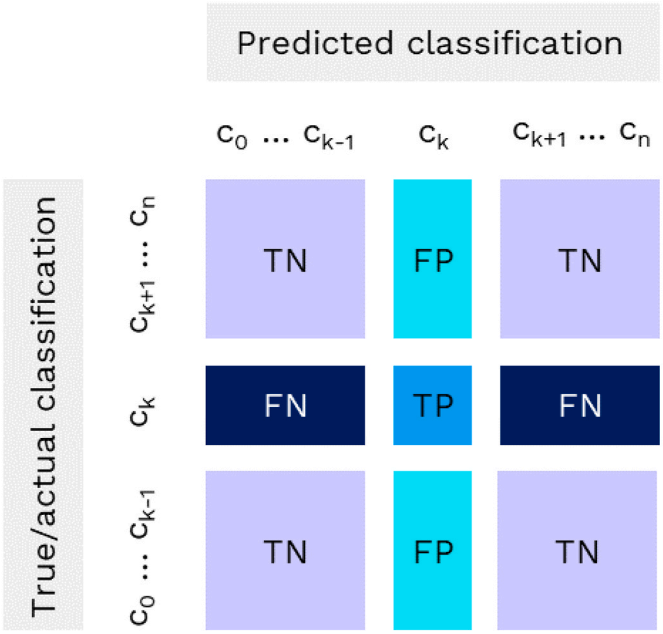

Fig. 3. Confusion matrix for multi-class classification with n classes. The estimation of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) is presented when considering a class k ($0 \leq k \leq n$).

classification task that considers all the elements of the confusion matrix (True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN)) as expressed in Eq. (7).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

The balanced accuracy is another widely used metric, the estimation of which is also related to the confusion matrix. This metric can be defined as an average of Recalls, as, firstly, an evaluation of the Recall for each class is performed and, subsequently, the obtained values are averaged to determine the balanced accuracy score. The Recall is the fraction of True Positive elements divided by the total number of the actual positives (see Eq. (8)).

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

The Micro F1-Score is computed by estimating Micro-Precision (Eq. (9)) and Micro-Recall (Eq. (10)). The Micro-averaging is presented in this case to avert differences between classes. As the Micro-Average Precision and Recall refer to the same values, the Micro-Average F1-Score is equal to both Micro-Average Precision and Recall, as the harmonic mean of two equal values is just the value (see Eq. (11)).

$$Micro\ Average\ Precision = \frac{\sum_{k=1}^{K} TP_k}{Grand\ Total} \tag{9}$$

$$Micro\ Average\ Recall = \frac{\sum_{k=1}^{K} TP_k}{Grand\ Total} \tag{10}$$

$$Micro\ F1 = \frac{\sum_{k=1}^{K} TP_k}{Grand\ Total} \tag{11}$$

Analogously, the Macro F1-Score is determined by estimating the Macro-Precision (Eq. (12)) and Macro-Recall (Eq. (13)). The Macro F1-Score is then estimated by determining the harmonic mean of Macro-Precision and Macro-Recall (see Eq. (14)).

$$Macro\ Average\ Precision = \frac{\sum_{k=1}^{K} Precision_k}{K} \tag{12}$$

$$Macro\ Average\ Recall = \frac{\sum_{k=1}^{K} Recall_k}{K} \tag{13}$$

$$Macro\ F1 = 2^* \left(\frac{Macro\ Average\ Precision^* Macro\ Average\ Recall}{Macro\ Average\ Precision^{-1} + Macro\ Average\ Recall^{-1}} \right) \tag{14}$$

The MCC is defined in terms of a confusion matrix C for K classes, as expressed hereunder.

$$MCC = \frac{c \times s - \sum_{k}^{K} p_k \times t_k}{\sqrt{\left(s^2 - \sum_{k}^{K} p_k^2\right)\left(s^2 - \sum_{k}^{K} t_k^2\right)}}, \tag{15}$$

where.

$c = \sum_{k}^{K} C_{kk}$ is the total number of correctly predicted elements.

$s = \sum_{i}^{K} \sum_{j}^{K} C_{ij}$ is the total number of elements.

$p_k = \sum_{i}^{K} C_{ki}$ is the number of times that class k was predicted.

$t_k = \sum_{i}^{K} C_{ik}$ is the number of times that class k truly occurred.

Finally, the last metric considered is the Cohen's Kappa, which is similar to MCC when a muti-class classification task is being considered. Cohen's Kappa metric (K) can be described as follows:

$$K = \frac{c \times s - \sum_{k}^{K} p_k \times t_k}{s^2 - \sum_{k}^{K} p_k \times t_k} \tag{16}$$

4. Results

A case study is presented in this section in order to validate the performance of the proposed methodology. Accordingly, a Diesel GenSet (DG), which is used for auxiliary purposes onboard an Aframax size tanker ship, is considered. This is one out of three DGs that are utilised onboard the ship to provide all the onboard electrical supply that is required when the ship is either under way or when the ship is alongside loading/offloading cargo. The DG employed in this case study is a four-stroke in-line engine comprised of a total of 6 cylinders. Specifically, the DG power parameter is discussed further in this paper due to its criticality. However, any other parameter could be considered (e.g., exhaust gas inlet and outlet temperatures, cooling water inlet and outlet temperatures, and turbocharger lube oil pressure and temperature) but, due to the size of the paper and results, only the power parameter is presented in this study. Furthermore, although a univariate analysis is being presented as a case study due to the lack of fault data, the methodology has been structured to consider both a univariate and a multivariate approach.

More than 66,000 instances are analysed for such a parameter, in total. These instances have been collected in a 1-min frequency. A graphical representation of these can be perceived in Fig. 4. Moreover, the descriptive statistics are also presented in Table 1.

As it can be perceived in Fig. 6, raw data collected from marine machinery usually contains both non-operational states and variations in the operating conditions, which can adversely alter the fault diagnosis analysis. Accordingly, the steady states' identification phase is applied as part of the pre-processing step. In total, 81 operational sequences are determined. Each of these sequences are further analysed to establish their respective quality, and thus either accept or reject them for the training, validation, and test stages. Furthermore, as part of the preparation step, data normalisation is applied so that each sequence lies between 0 and 1 values. The sliding window algorithm is subsequently applied to successively section each of the identified operational sequences into subsequences. Various configurations have been performed to determine the most appropriate parameters of such an algorithm.

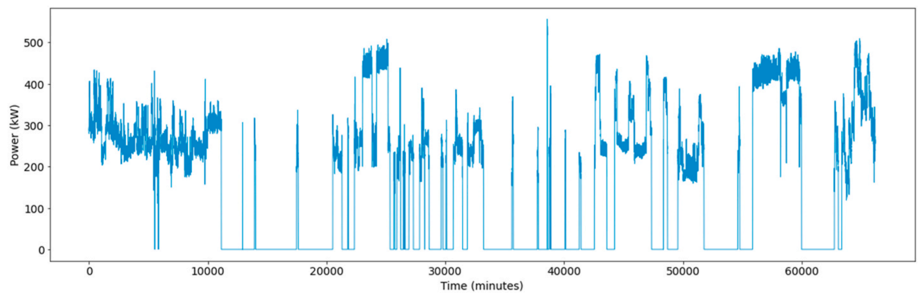

Fig. 4. Time series plot of the cooling air temperature monitored parameter.

Table 1
Descriptive statistics of the monitored parameter.

	Mean	Std.	Min.	25%	50%	75%	Max.
Power (kW)	151.67	159.15	0.0	0.0	177.95	273.30	555.93

After heuristically evaluating each of them, it is determined that the most optimal value for the window size parameter is 250 for this study.

Once the data preparation step is finalised, distinct patterns are simulated as introduced in the methodology section. These patterns refer to point anomalies: 1) only one point anomaly is presented in the sequence, 2) two-point anomalies are identified in the sequence, 3) multiple point anomalies (>2) can be observed in the sequence, 4) collective anomalies, 5) degradation sequences, and 6) transitional occurrences between operational states. A graphical representation of the distinct patterns simulated can be perceived in Fig. 6. Fig. 5 presents the sequence that has been altered to simulate such patterns. As perceived in Fig. 6(a)–6(c), point anomalies are presented in the form of spikes and refer to those instances that differ from others with regards to their attributes. Fig. 6(d) refer to collective anomalies, which are a combination of instances that present high variability in a steady operational state context. The degradation pattern is presented in Fig. 6(e) and refers to an anomaly that presents an exponential increase. Finally, the changes between operational states have also been simulated and captured. An example of two steady operational states in a sequence is presented in Fig. 6(f).

Subsequently, the time series are encoded into images so that the image classification task can be performed. As the first-order Markov chain model was implemented at this stage, the number of states that the transition matrix contains needs to be adequately estimated. Accordingly, more than ten values have been considered. These are within the range 40–150. Values lesser than 40 were not considered as the minimum dimensions that the image could contain was 32×32, as transfer learning was implemented, and, specifically, the ImageNet dataset was utilised as part of the pre-training process. Of all possibilities, the number of states is set to 50, as values greater than 50 do not facilitate a significant enhancement in the accuracy of the model and the risk of over-fitting increases.

Figs. 7 and 8 provide output examples of the time series imaging phase. Although the provided example images present a dimension of 10×10 for a better visual interpretation, the images used as input in the image classification phase present a resolution of 50×50, as stated in the preceding paragraph. Fig. 7 refers to an image generated from a normal operational sequence. Consequently, a clear diagonal can be observed, as the instances within an operational steady state do not vary in a significant manner. Therefore, the subsequent state of an instance usually adopts the state of the preceding instance or one around it, thus creating this diagonal. For instance, when considering normal operational sequences, if the current instance relates to state 2, it is highly probable that the subsequent instance will refer to state 2 or a near state, such as state 1 and 3. A similar characteristic can be perceived when the sequence presents a point anomaly (Fig. 8 (a)). However, the diagonal is shorter when a normal operational image is considered, as the states are defined based on a different range of values; this range having an abnormal value greater than the range of the normal operational sequence. Such an aspect does not apply when the number of point anomalies in a sequence is increased, as the relationship between the current and the preceding state is distorted, thus intrinsically disrupting the diagonal perceived in normal images. This applies similarly to collective anomalies and degradation images, as the huge number of abnormal values modified the steady context. With regards to the transition occurrences between operation steps, the scenarios are slightly different. For this case, a total of two diagonals can be observed, each of them referring to an operational steady state. Also, isolated pixels can be perceived, which refers to the transition state that occurs between steady operational states. By applying the image classification phase, it is expected that the deep learning method can learn such characteristics and identify the different operational, non-operational, and fault patterns presented in the distinct defined categories.

As part of the image classification phase, the ResNet50V2 network is analysed following the architecture described in Fig. 2. Moreover, as part of the comparative study, a traditional CNN architecture has also been considered. After a heuristic evaluation and the analysis of analogous studies (Almutairi et al., 2021; Yao et al., 2021), the analysed CNN architecture for image classification is comprised of two convolutional layers with 192 filters and kernel size of 3×3. The pooling operation presents a 2×2 dimension. After the feature extraction stage, a total of

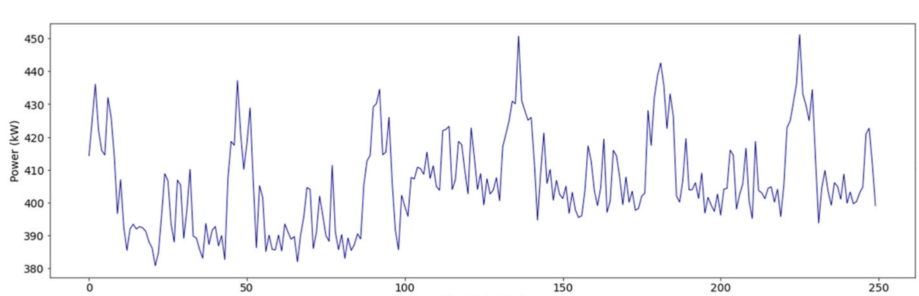

Fig. 5. Original sequence that has been altered to simulate the distinct abnormal operational sequences.

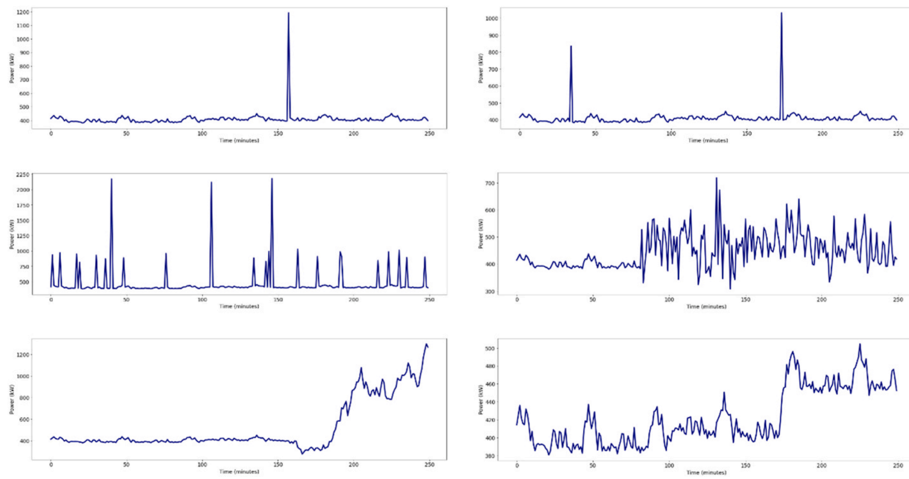

Fig. 6. Simulated sequence with (a) a point anomaly, (b) two-point anomalies, (c) multiple point anomalies, (d) collective anomalies, (e) degradation, and (f) transition occurrences between steady operational states.

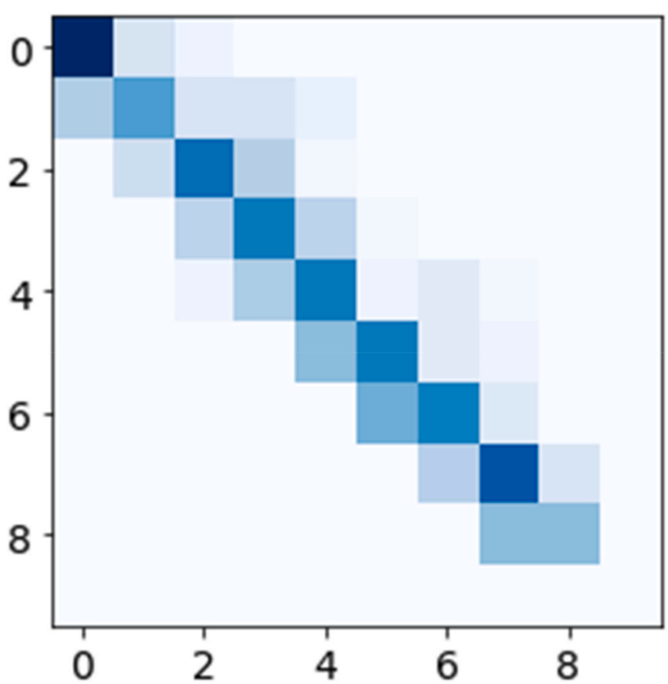

Fig. 7. Example of a normal sequence encoded into an image.

three fully connected layers with 64, 128, and 192 hidden units are defined. The same parameters have been considered for the 1D-CNN architecture to ensure the adequate performance of the comparative study. With regards to the models that consider GAF as a time series approach (GAF-ResNet50V2, and GAF-CNN), the input shapes of the analysed networks are equal to the models that consider the first-order Markov chain as a time series method (Markov-ResNetV2, and Markov-CNN). Thus, the dimensions of the GAF images are equal to the Markov ones: 50 × 50.

The results of the classification task after the training process in tandem with the comparative analysis are summarised in Table 2. As it can be perceived, the models have been ordered by their performance based on the accuracy score. The model that led to the most accurate results is the Markov-CNN. It presents a performance enhancement of a 2% when considering the second most accurate model, and a 23% when considering the least accurate model. Although the performance enhancement with regards to 1D-CNN is not significant, Markov-CNN is a turning point in the consideration of time series imaging approaches for performing fault classification tasks. The proposed approach, Markov-ResNet50V2, present nearly identical results as 1D-CNN, which, once again, shows the potential of time series imaging approaches when dealing with both time series data and fault classification tasks. However, results may suggest that the implementation of architectures such as ResNets may not be appropriate when dealing with the characteristics of case studies such as the one presented in this study or their complexity. As perceived in Figs. 7 and 8, the images outlined are less comprehensive than the ones considered in the computer vision tasks that these types of architectures were designed for. Accordingly, as precedingly stated, the utilisation of CNNs is sufficient to achieve a high calibre performance. However, a more comprehensive analysis of more sophisticated image versions obtained from either multivariate or higher order Markov chain models need to be performed to sustain such a fact. In addition, with regards to the transfer learning task performed in the proposed methodology, it can be perceived that its contribution in the performance enhancement was insignificant. Once again, it is probably due to the characteristics of the case study performed, as large number of images were simulated. However, its contribution may be significant when dealing with small sets of real-world faults, in which the amount of data available is limited. Accordingly, further research needs to be addressed to also sustain such a fact.

After analysing the third most accurate model, a significant drop can be observed in the accuracy performance, yielding a decrease of the accuracy score of more than 10%. This suggests that the proposed time series imaging approach outperforms GAF, which has been widely utilised in analogous fault classification studies.

Therefore, it can be perceived that there is a need to further analyse time series imaging approaches and image classification models due to their promising results when dealing with fault classification tasks. To ensure the enhancement opportunities based on the results obtained from this study, future work guidelines are presented hereunder.

- The main challenge presented when performing this study was the lack of fault data. Accordingly, anomalous data needed to be simulated. However, to enhance the validation stage of this study and determine potential pitfalls that need to be addressed, the implementation of real-world case studies including fault data is of preeminent importance.
- The first-order Markov chain was implemented in this study as a time series imaging approach. Thus, a univariate approach was presented. Further validation efforts to perform multivariate analysis is required. In this sense, an initial validation has been performed by the authors as a part of a multivariate analysis by stacking all the

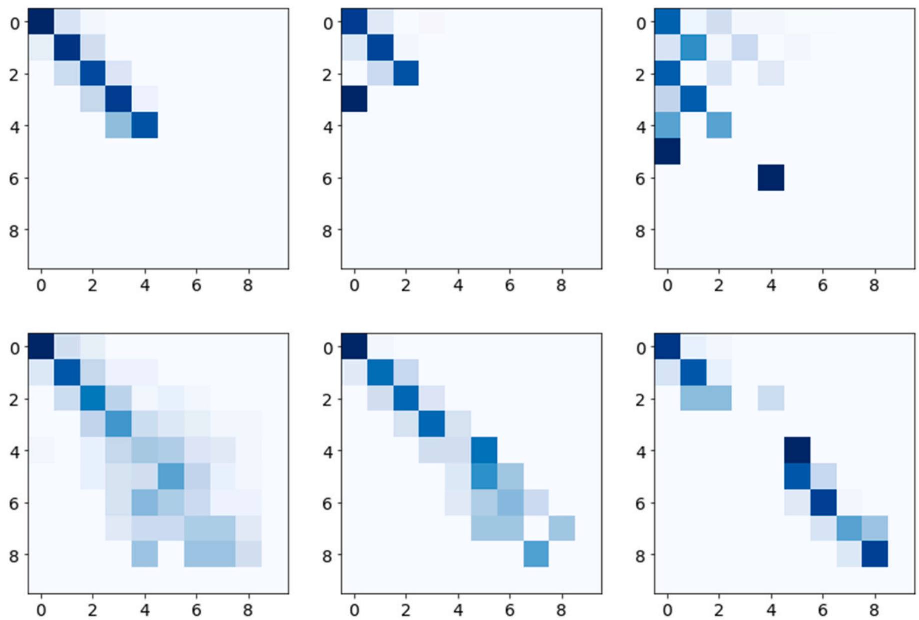

Fig. 8. Example of images generated that contain (a) a point anomaly, (b) two-point anomalies, (c) multiple point anomalies, (d) collective anomalies, (e) degradation, and (f) transition occurrences between steady operational states.

Table 2
Classification metrics results for performance evaluation of the multi-fault classification task.

Model	Accuracy	Balanced Accuracy	Micro F1	Macro F1	MCC	Cohen's Kappa
Markov-CNN	0.95	0.95	0.95	0.94	0.94	0.94
1D-CNN	0.93	0.94	0.93	0.93	0.92	0.92
Markov-ResNet50V2	0.93	0.93	0.93	0.93	0.91	0.91
GAF-CNN	0.83	0.84	0.83	0.83	0.81	0.81
GAF-ResNet50V2	0.72	0.72	0.72	0.71	0.67	0.67

individual transition matrices (one for each parameter) to present them as the input of the different image classification models being analysed. The authors expect to present such results in subsequent studies. Moreover, the consideration of other multivariate approaches, such as multivariate Markov chains, are being studied.

- The study of other time series imaging approaches and classification models is also suggested. In this study, a comparative study has been performed in which a total of two time series imaging approaches and four classification models have been assessed. However, due to an increased interest in PHM within the shipping industry and other comparable sectors, such as manufacturing and aerospace, numerous state-of-the-art methods are being introduced that need to be considered to continue advancing the enhancement of fault classification tasks within the shipping industry. Moreover, other elements that both complement the models developed and enhance transparency and performance need to be also considered. Examples of these are the implementation of explainable artificial intelligence or the consideration of evolutionary algorithms for applying hyper-parameters optimisation.

5. Conclusions

Fault classification is a preeminent phase of the fault diagnosis module; a module that aims at the identification of the failure modes and their causes so that a relationship between the monitoring data and the fault condition can be established. Despite its importance, this is still an unexplored area within the shipping industry.

As demonstrated in this paper, the methodology presented, which is comprised of a time series imaging approach based on the implementation of the first-order Markov chain, and an image classifier, such as the ResNet50V2 and CNN architectures, exposed its applicability for performing the fault classification task when identifying anomalies of marine systems.

Due to the lack of fault data and data availability within the sector, it was necessary to simulate anomalies. In total, six distinct anomalies, such as point anomalies and collective anomalies, were considered. Moreover, to validate the proposed methodology, a case study on a diesel generator of a tanker ship was presented. Specifically, the power parameter was analysed. In addition, a comparative study of other potential models was introduced. These were recognised as 1D-CNN, GAF-CNN, and GAF-ResNet50V2. Results demonstrated that Markov-CNN outperformed the remaining analysed methods by achieving an accuracy of 95%, and thus suggesting the potential of time series imaging and image classification approaches for the performance of the fault classification task. To continue analysing the potential of these, further research needs to be performed. The validation of the methods with real-world fault data, the application of multivariate analysis, and the consideration of other novel techniques are within the future research agenda and are suggested for other researchers that are keen to contribute towards the application of Smart Maintenance within the shipping sector.

CRediT authorship contribution statement

Christian Velasco-Gallego: Conceptualization, Methodology, Software, Formal analysis, Visualization, Validation, Writing – original draft, Writing – review & editing. **Iraklis Lazakis:** Conceptualization, Methodology, Resources, Validation, Writing – review & editing, Supervision.

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

The data that has been used is confidential.

References

Allianz Global Corporate & Specialty, 2012. Safety and Shipping. From Titanic to Costa Concordia, pp. 1–33.

Almutairi, M., Nikitas, N., Abdeljaber, O., Avci, O., Bocian, M., 2021. A methodological approach towards evaluating structural damage severity using 1D CNNs. Structures 34, 1–12. https://doi.org/10.1016/j.istruc.2021.10.029.

Bolbot, V., Theotokatos, G., Wennersberg, L.A., Faivre, J., Vassalos, D., Boulougouris, E., Rødseth, Ø.J., Andersen, P., Pauwlyn, A.-S., Coillie, A.V., 2021. A novel risk assessment process: application to an autonomous inland waterways ship. Proc. Inst. Mech. Eng. O J. Risk Reliab. 1–23. https://doi.org/10.1177/1748006X211051829.

Bolbot, V., Theotokatos, G., Boulougouris, E., Vassalos, D., 2020. A novel cyber-risk assessment method for ship systems. Saf. Sci. 131, 1–14. https://doi.org/10.1016/j.ssci.2020.104908.

Bui, K.Q., Perera, L.P., 2021. Advanced data analytics for ship performance monitoring under localized operational conditions. Ocean Eng. 235, 1–14. https://doi.org/10.1016/j.oceaneng.2021.109392.

Cai, C., Weng, X., Zhang, C., 2017. A novel approach for marine diesel engine fault diagnosis. Cluster Comput. 20, 1691–1702. https://doi.org/10.1007/s10586-017-0748-0.

Cheliotis, M., Lazakis, I., Cheliotis, A., 2022. Bayesian and machine learning-based fault detection and diagnostics for marine applications. Ships Offshore Struct. 1–14. https://doi.org/10.1080/17445302.2021.2012015.

Fahim, S.R., Sarker, S.K., Muyeen, S.M., Sheikh, M.R.I., Das, S.K., Simoes, M., 2021. A robust self-attentive capsule network for fault diagnosis of series-compensated transmission line. IEEE Trans. Power Deliv. 36, 3846–3857. https://doi.org/10.1109/TPWRD.2021.3049861.

Fahim, S.R., Muyeen, S.M., Sarker, Y., Sarker, S.K., Das, S.K., 2021b. An Agreement Based Dynamic Routing Method for Fault Diagnosis in Power Network with Enhanced Noise Immunity, 2021 31st Australasian Universities Power Engineering Conference, pp. 1–5. https://doi.org/10.1109/AUPEC52110.2021.9597762.

Fahim, S.R., Niloy, S., Shatil, A.H., Hazari, M.R., Sarker, S.K., Das, S.K., 2021c. An unsupervised protection scheme for overhead transmission line with emphasis on situations during line and source parameter variation. Int. Conf. Robot. Electr. Signal Process. Tech. 758–762. https://doi.org/10.1109/ICREST51555.2021.9331170.

Fahim, S.R., Sarker, Y., Sarker, S.K., Sheikh, M.R.I., Das, S.K., 2020. Self-attention convolutional neural network with time series imaging based feature extraction for transmission line fault detection and classification. Elec. Power Syst. Res. 187, 1–12. https://doi.org/10.1016/j.epsr.2020.106437.

Fahim, S.R., Islam Sarker, M.R., Arifuzzaman, M., Hosen, M.S., Sarker, S.K., Das, S.K., 2020b. A Novel Approach to Fault Diagnosis of High Voltage Transmission Line – A Self Attentive Convolutional Neural Network Model, 2020 IEEE Region 10 Symposium, pp. 1329–1332. https://doi.org/10.1109/TENSYMP50017.2020.9230660.

Farag, Y.B.A., Ölçer, A.I., 2020. The development of a ship performance model in varying operating conditions based on ANN and regression techniques. Ocean Eng. 198, 1–12. https://doi.org/10.1016/j.oceaneng.2020.106972.

Grandini, M., Bagli, E., Visani, G., 2020. Metrics for multi-class classification: an overview. Comput. Sci. 1–17. ArXiv, pp. https://arxiv.org/pdf/2008.05756.pdf.

Han, P., Ellefsen, A.L., Li, G., Æsøy, V., Zhang, H., 2021. Fault Prognostics using LSTM networks: application to marine diesel engine. IEEE Sensor. J. 1–8. https://doi.org/10.1109/JSEN.2021.3119151.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778. https://doi.org/10.1109/CVPR.2016.90.

He, K., Zhang, X., Ren, S., Sun, J., 2016b, 2016. ECCV 2016. Lecture Notes in Computer Science. In: Computer Vision – ECCV, vol. 9908. Springer, Cham, pp. 630–645. https://doi.org/10.1007/978-3-319-46493-0_38. Identity Mapping in Deep Residual Networks.

Hou, L., Zhang, J., Du, B., 2020. a fault diagnosis model of marine diesel engine fuel oil supply system using PCA and optimized SVM. J. Phys. Conf. 1576, 1–9. https://doi.org/10.1088/1742-6596/1576/1/012045.

Kiangala, K.s., Wang, Z., 2020. An effective predictive maintenance framework for conveyor motors using dual time-series imaging and convolutional neural network in an industry 4.0 environment. IEEE Access 8, 121033–121049. https://10.1109/ACCESS.2020.3006788.

Li, Z., Wang, Y., Wang, K., 2019. A deep learning driven method for fault classification and degradation assessment in mechanical equipment. Comput. Ind. 104, 1–10. https://doi.org/10.1016/j.compind.2018.07.002.

Li, Y., Huang, X., Ding, P., Zhao, C., 2021. Wiener-based remaining useful life prediction of rolling bearings using improved Kalman filtering and adaptive modification. Measurement 182, 1–17. https://doi.org/10.1016/j.measurement.2021.109706.

London Economics, NLA International, Marine South East, 2021. Consultancy Research into the UK Maritime Technology Sector, 1-151.

Nair, V., Hinton, G.E., 2010. Rectified Linear Units Improve Restricted Boltzmann Machines. ICML'10: Proceedings of the 27th International Conference on International Conference on Machine Learning, pp. 807–814.

Nisha, S.S., Meeral, M.N., 2021. 9 – Applications of Deep Learning in Biomedical Engineering. Handbook of Deep Learning in Biomedical Engineering. Techniques and Applications, pp. 245–270. https://doi.org/10.1016/B978-0-12-823014-5.00008-9.

Rahimzadeh, M., Attar, A., 2020. A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. Inform. Med. Unlocked 19, 1–9. https://doi.org/10.1016/j.imu.2020.100360.

Ramchandran, A., Sangaiah, A.K., 2018. Chapter 11 – Unsupervised Anomaly Detection for High Dimensional Data – and Exploratory Analysis. Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications. Intelligent Data-Centric Systems, pp. 233–251. https://doi.org/10.1016/B978-0-12-813314-9.00011-6.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L., 2015. ImageNet large scale visual recognition challenge. Int. J. Comput. Vis. 115, 211–252. https://doi.org/10.1007/s11263-015-0816-y.

Senemmar, S., Zhang, J., 2021. Deep Learning-Based Fault Detection, Classification, and Locating in Shipboard and Power Systems, 2021 IEEE Electric Ship Technologies Symposium, pp. 1–6. https://doi.org/10.1109/ESTS49166.2021.9512342.

Tan, Y., Zhang, J., Tian, H., Jiang, D., Guo, L., Wang, G., Lin, Y., 2021. Multi-label classification for simultaneous fault diagnosis of marine machinery: a comparative study. Ocean Eng. 239, 1–11. https://doi.org/10.1016/j.oceaneng.2021.109723.

Tan, Y., Tian, H., Jiang, R., Lin, Y., Zhang, J., 2020. A comparative investigation of data-driven approaches based on one-class classifiers for condition monitoring of marine machinery system. Ocean Eng. 201, 1–12. https://doi.org/10.1016/j.oceaneng.2020.107174.

Theotokatos, G., Stoumpos, S., Bolbot, V., Boulogouris, E., 2020. Simulation-based investigation of a marine dual-engine. J. Mar. Eng. Technol. 19, 1–13. https://doi.org/10.1080/20464177.2020.1717266.

Uyanik, T., Karatuğ, C., Arslanoğlu, Y., 2021. Machine learning based visibility estimation to ensure safer navigation in strait of Istanbul. Appl. Ocean Res. 112, 1–10. https://doi.org/10.1016/j.apor.2021.102693.

Velasco-Gallego, C., Lazakis, I., 2020. Real-time data-driven missing data imputation for short-term sensor data of marine systems. A comparative study. Ocean Eng. 218, 1–23. https://doi.org/10.1016/j.oceaneng.2020.108261.

Velasco-Gallego, C., Lazakis, I., 2021. A novel framework for imputing large gaps of missing values from time series sensor data of marine machinery systems. Ships Offshore Struct. 1–11. https://doi.org/10.1080/17445302.2021.1943850.

Velasco-Gallego, C., Lazakis, I., 2022. A real-time data-driven framework for the identification of steady states of marine machinery. Appl. Ocean Res. 121, 1–12. https://doi.org/10.1016/j.apor.2022.103052.

Wang, Z., Oates, T., 2015. Encoding Time Series as Images for Visual Inspection and Classification Using Tiled Convolutional Neural Networks. AAAI Workshop, pp. 1–7.

Wang, S., Wang, J., Wang, R., 2020. A novel scheme for intelligent fault diagnosis of marine diesel engine using the multi-information fusion technology. IOP Conf. Ser. Mater. Sci. Eng. 782, 1–12. https://doi.org/10.1088/1757-899X/782/3/032022.

Wang, K., Hu, Q., Zhou, M., Zun, Z., Qian, X., 2021. Multi-aspect applications and development challenges of digital twin-driven management in global smart ports. Case Stud. Transport Pol. 9, 1298–1312. https://doi.org/10.1016/j.cstp.2021.06.014.

Yao, Y., Wang, J., Xie, M., Hu, L., Wang, J., 2020. A new approach for fault diagnosis with full-scope simulator based on state information imaging in nuclear power plant. Ann. Nucl. Energy 141, 1–9. https://doi.org/10.1016/j.anucene.2019.107274.

Yao, D., Li, B., Liu, H., Yang, J., Jia, L., 2021. Remaining useful life prediction of roller bearings based on improved 1D-CNN and simple recurrent unit. Measurement 175, 1–14. https://doi.org/10.1016/j.measurement.2021.109166.

Zhang, P., Lozano, J., Wang, Y., 2021. Using Flettner Rotors and Parafoil as alternative propulsion systems for bulk carriers. J. Clean. Prod. 317, 1–14. https://doi.org/10.1016/j.jclepro.2021.128418.

Zhang, P., Gao, Z., Cao, L., Dong, F., Zou, Y., Wang, K., Zhang, Y., Sun, P., 2022. Marine systems and equipment Prognostics and health management: a systematic review from health condition monitoring to maintenance strategy. Machines 10, 1–53. https://doi.org/10.3390/machines10020072.

Zhao, Z., Cerf, S., Birke, R., Robu, G., Bouchenak, S., Mokhtar, S.B., Chen, L.Y., 2019. Robust Anomaly Detection on Unreliable Data. In: 49th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), pp. 630–637. https://doi.org/10.1109/DSN.2019.00068.

Zhao, D., Wang, T., Chu, F., 2019b. Deep convolutional neural network based planet bearing fault classification. Comput. Ind. 107, 59–66. https://doi.org/10.1016/j.compind.2019.02.001.

Zio, E., 2022. Prognostics and Health Management (PHM): where are we and where do we (need to) go in theory and practice. Reliab. Eng. Syst. Saf. 218, 1–16. https://doi.org/10.1016/j.ress.2021.108119.