

# Online Fault Detection in Autonomous Ferries: Using Fault-type Independent Spectral Anomaly Detection

André Listou Ellefsen, Peihua Han, Xu Cheng, *Student Member, IEEE*, Finn Tore Holmeset, Vilmar Æsøy, and Houxiang Zhang\*, *Senior Member, IEEE*

**Abstract**—Enthusiasm for ship autonomy is flourishing in the maritime industry. In this context, data-driven Prognostics and Health Management (PHM) systems have emerged as the optimal way to improve operational reliability and system safety. However, further research is needed to enhance the essential actions relating to such a system. Fault detection is the first and most crucial action of any data-driven PHM system. In this study, we propose a fault-type independent spectral anomaly detection algorithm for marine diesel engine degradation in autonomous ferries. The benefits of the algorithm are verified on three fault-types where the nature of degradation differs. Both normal operation data and faulty degradation data have been collected from a marine diesel engine, using two different engine load profiles. These profiles aim to replicate real autonomous ferry crossing operations, environmental conditions the ferry may encounter. First, the data is subjected to a feature selection process to remove irrelevant and redundant features. Then, a multi-regime normalization method is performed on the data to merge the engine loads into one context. Finally, a variational autoencoder is trained to estimate velocity and acceleration calculations of the anomaly score. Generic and dynamic threshold limits are simultaneously established to detect the fault time step online. The algorithm achieved an accuracy of 97.66% in the final test when the acceleration was used as the fault detector. The results suggest that the algorithm is independent of fault-types with different nature of degradation related to the marine diesel engine.

**Index Terms**—Autonomous ferry, marine diesel engine, multi-regime normalization, online fault detection, prognostics and health management

## I. INTRODUCTION

**T**ODAY, ship autonomy is the most-sought research objective at the Norwegian University of Science and Technology in Aalesund [1], [2]. However, autonomous ships were considered to be a futuristic fantasy only six years ago [3]. Yet inland autonomous ferries carrying tiny crews primarily to make passengers feel safe will be in commercial use on the west coast of Norway in the very near future [4]. The industry,

\*Corresponding author.

André Listou Ellefsen, Peihua Han, Xu Cheng, Finn Tore Holmeset, Vilmar Æsøy, and Houxiang Zhang are with the Department of Ocean Operations and Civil Engineering, as part of the Mechatronics Laboratory, Norwegian University of Science and Technology, Aalesund, 6009, Norway, (e-mail: andre.ellefsen@ntnu.no; peihua.han@ntnu.no; xu.cheng@ntnu.no; fih@ntnu.no; vilmar.aesoy@ntnu.no; hozh@ntnu.no).

Manuscript received January 19, 2020; revised March 16, 2020; accepted May 1, 2020

as well as academics, anticipate that these ferries will improve both safety and profitability [5]. Maintaining, operating, and navigating the vessels without crew involvement will necessitate the use of highly automated systems and belonging sensor equipment, and degradation of such systems during operation poses a serious threat to operations [6].

Prognostics and Health Management (PHM) is the area of research with the greatest promise to manage maintenance operations for zero-downtime performance of autonomous ferries [6]. A data-driven PHM system goes far beyond traditional maintenance approaches, such as reactive maintenance and preventive maintenance, currently in use onboard ships [7]. Such a system use algorithms built on sensor measurements to perform automatic fault detection, fault isolation, fault classification, and associated remaining useful life (RUL) predictions to devise an ideal maintenance schedule that eliminates failures [8]. Autonomous ferries will transfer real-time operational sensor data to a remote control center to conduct the essential actions of a data-driven PHM system (see Figure 1). Thus, it will be possible to schedule maintenance operations to the next appropriate port of call. The ideal maintenance schedule will considerably enhance operational availability and reliability and system safety.

Anomaly detection techniques aim to discover deviations from normal operation data. In a data-driven PHM viewpoint, such deviations are symptoms of incipient faults [9]. Fault detection is the first and most crucial action of any data-driven PHM system. It should be performed automatically to detect the fault time step in degradation data. Then, this time step can be used to construct both labels for fault classification and run-to-failure targets for RUL predictions. Interest in spectral anomaly detection techniques has increased recently. These techniques try to produce the lower dimensional embedding of the input data where anomalies and normal operation data are generally distinct [10]. The reconstruction error at each time step between the input data and its low dimensional reconstruction is then used as an anomaly score to detect anomalies [10]. The principal components analysis method is one of the best-known traditional spectral anomaly detection techniques [11]. However, deep neural networks (DNNs) have recently shown superior performance for this purpose [9]. DNNs allow dimension reduction through several hidden layers with non-linear transformations, and hence, obtain more

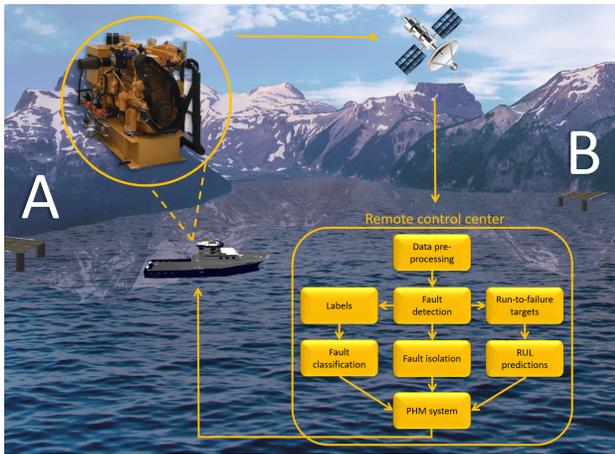


Fig. 1. Illustration of an autonomous ferry, crossing a fjord from dock A to B. Since there are limited amounts of crew members onboard, such ferries need to transfer real-time operational sensor data to a remote control center to conduct the essential actions of a data-driven PHM system. Then, maintenance operations can be scheduled to the next appropriate port of call.

abstract features to produce a better reconstruction of the input data.

The marine diesel engine is one of the most critical components onboard ferries since it has an important role in both propulsion and power generation [12]. It is subjected to rapid variations in operational loads, depending on both the task of operation and environmental conditions. In such complexity, the degradation phenomena cannot be presented directly for cutting-edge spectral anomaly detection algorithms since the sensor measurements are highly connected to the operational loads. Hence, a multi-regime normalization method has to be performed on the raw input data to present the degradation phenomena [13]. Additionally, the nature of degradation of typical fault-types associated with the marine diesel engine might be different from one another and significantly similar to normal operation data.

This paper proposes a fault-type independent spectral anomaly detection algorithm for marine diesel engine degradation in autonomous ferries. The variational autoencoder (VAE) is the selected DNN as it outperforms a feed-forward neural network (FNN) with one hidden layer, the traditional autoencoder (AE), and the long-short term memory (LSTM), in terms of reconstruction-based fault detection for maritime components in [9]. As similar to [14], a replicated autonomous ferry crossing operation is used to produce two engine load profiles. These profiles reflect different environmental conditions affecting the ferry. Both normal operation and faulty degradation data sets are collected from the two profiles, and a fault is introduced at an unknown time step in the degradation data sets. During the experiments, three fault-types with different nature of degradation are used for both validation and final test of the proposed algorithm. The complete algorithm is summarized as follows: First, the VAE is trained on pre-processed normal operation data. Second, the trained VAE is used to calculate the velocity and the acceleration of the anomaly score at each time step in faulty degradation data. Simultaneously, generic and dynamic threshold limits are established. Both the

calculations and the threshold limits change dynamically with time. This enables online fault detection as a fault is detected automatically once the velocity and acceleration calculations exceed the threshold limits.

The proposed algorithm is based on our already published fault detection algorithm in [9]. Our previous algorithm makes only offline fault detection possible. However, as opposed to our previous algorithm, the proposed algorithm in this study includes two principal improvements, that is, online and fault-type independent anomaly detection by utilizing generic and dynamic threshold limits. This study's main contributions are as follows:

- A fault-type independent spectral anomaly detection algorithm for marine diesel engine degradation in autonomous ferries is proposed.
- Generic and dynamic threshold limits are proposed to predict the fault time step online.
- The algorithm is independent of fault-types with different nature of degradation related to the marine diesel engine.

The overall organization of the paper is as follows. Section II introduces relevant and related work on spectral anomaly detection. Section III introduces the essential background on the VAE and the semi-supervised reconstruction framework. The experimental approach is explained in detail in section IV. Results and discussions are elaborated in section V. Finally, section VI concludes the paper and presents objectives for future work.

## II. RELATED WORK

Three different learning procedures exist for spectral anomaly detection algorithms: supervised, semi-supervised, and unsupervised. The availability and quality of the input data largely determine which learning procedure to choose for fault detection. Supervised learning involves training a supervised binary or multi-class classifier to differentiate normal operation data from faulty degradation data. This procedure is extremely powerful if predefined labels for both normal and faulty data points are available during the training stage.

G. Wu proposed a supervised FNN for fault detection of ship equipment in [15]. In [16], Xu et al. proposed an online fault diagnostics method based on convolutional neural networks (CNNs) and transfer learning. The proposed approach was trained in a supervised manner where a softmax output layer was used to classify faults related to both bearings and pumps. A supervised classifier was also used for fault detection in [17]. In this study, however, Sun et al. utilized an initial unsupervised learning procedure, before supervised fine-tuning, to do automatic feature extractions of rolling element bearings. Siegel et al. examined methods for detecting and disrupting arc faults in [18]. Both a binary and multi-class classifier were used during real-time classification experiments.

Even though the above studies have shown superior accuracy in terms of fault detection, there is a lack of labeled faults in the maritime industry [19]. This necessitates the use of semi-supervised or unsupervised learning, which does not require predefined fault labels. In the application of fault

detection, semi-supervised learning only uses normal operation data for training, while unsupervised learning has no previous knowledge of the input data where only intrinsic properties are used [20].

The sensors installed on autonomous ferries can be utilized to accumulate and collect normal operation data to use a semi-supervised learning framework. A VAE was used for anomaly detection in [9], [14]. In both studies, the maximum acceleration in faulty degradation data was estimated and used as the fault detector. However, utilizing the maximum acceleration makes only offline fault detection possible. This is because one would need the faulty degradation data in advance to determine the maximum acceleration.

The utilization of dynamic threshold limits can enable online fault detection. Park et al. [21] proposed an LSTM-based VAE anomaly detector for robot-assisted feeding. A varying state-based threshold value was used to detect anomalies. Thus, online anomaly detection is possible where the threshold value changes over the estimated state of task execution. Additionally, Hundman et al. [22] used non-parametric dynamic thresholds for spacecraft anomaly detection. Nevertheless, these studies apply the dynamic thresholds based on the raw anomaly score or a smooth version of it. For the marine diesel engine, such dynamic thresholds will reflect the nature of degradation of the specific fault-type used for fault detection. Therefore, different dynamic thresholds have to be created for different faults. This contradicts the fact that the goal of the improved fault detection algorithm in this study is to be fault-type independent.

### III. BACKGROUND

This section introduces the background theory on the VAE and the semi-supervised reconstruction framework.

#### A. Variational autoencoder

The VAE is a variant of the traditional autoencoder (AE) rooted in Bayesian inference [23]. It is composed of an encoder function  $z = q_{\theta_e}(z|x)$  and a decoder function  $r = p_{\theta_d}(x|z)$ . The encoder approximates the underlying probability distribution  $p_{\theta_d}(z)$ . Then, new data can be generated utilizing the decoder by sampling a set of latent variables  $z$  obtained from  $p_{\theta_d}(z)$ . By modeling the distribution of the latent variables instead of deterministic values, as conducted in the traditional AE, the VAE improves generalization since  $z$  are stochastic in nature [24]. Note that  $\theta_e$  and  $\theta_d$  are the biases and weights of the encoder and decoder, respectively. The VAE optimizes  $\theta_e$  and  $\theta_d$  by maximizing the variational lower bound  $J_{VAE}$  [23]:

$$J_{VAE}(\theta_e, \theta_d) = -D_{KL}(q_{\theta_e}(z|x) || p_{\theta_d}(z)) + E_{q_{\theta_e}(z|x)}[\log p_{\theta_d}(x|z)] \quad (1)$$

where  $D_{KL}$  is the Kullback-Leibler (KL) divergence. The KL divergence measures the similarity between the prior distribution of  $z$ ,  $p_{\theta_d}(z)$ , and the variational approximation  $q_{\theta_e}(z|x)$ . Maximizing  $J_{VAE}$  minimizes the KL divergence, hence pushing the approximated posterior  $q_{\theta_e}(z|x)$  towards the prior  $p_{\theta_d}(z)$ . The common choice of the prior distribution is

a Gaussian distribution,  $\mathcal{N}(\mu_z, \Sigma_z)$ , where a standard normal distribution  $\mathcal{N}(0, 1)$  is utilized. The second expression is the reconstruction log-likelihood of  $x$  with sampling from  $q_{\theta_e}(z|x)$  and referred to as the generative loss. The distribution of the second expression depends on the data type [10]. For real-valued input data, a Multivariate Gaussian is normally used.

The reconstruction log-likelihood needs to be calculated through Monte Carlo methods [23]. However, since these methods suffer from high variance and high computation resources, a reparameterization trick of  $z$  is used to obtain the gradients of the decoder in order to use the back-propagation algorithm. The random variable  $z \sim q_{\theta_e}(z|x)$  is replaced by a deterministic transformation, such that,  $z = \mu + \sigma\varepsilon$ ,  $\varepsilon \sim \mathcal{N}(0, 1)$  [10]. Thus, given a fixed input  $x$  and a variable  $\varepsilon$ , the total function is deterministic and continuous, meaning back-propagation can compute a gradient that will work for stochastic gradient descent [23]. Then, the encoder only needs to produce vectors of means  $\mu$  and standard deviations  $\sigma$  instead of vectors of real values.

#### B. Semi-supervised reconstruction framework

As in [9], the fault detection is conducted through a semi-supervised reconstruction framework, meaning only normal operation data is used for training the VAE. Consider  $\mathbf{x}_t = [x_1, \dots, x_n]_t$  as the input vector at time step  $t$ . To enable the VAE to reconstruct the normal operation data,  $\mathbf{x}_t$  is also used as the target  $\mathbf{y}_t$  for reconstruction at each  $t$ . In this way, the trained VAE is expected to produce relatively large reconstruction errors on unseen degradation data. Since the data gathered from the marine diesel engine is continuous sensor data, a fully connected output layer is attached to the VAE, where the mean squared error (MSE) is utilized to measure the reconstruction capability. Thus, the VAE minimizes the following loss function:

$$L_{VAE} = \frac{1}{n} \sum_{i=1}^n \|\hat{y}_i - y_i\|^2 \quad (2)$$

where  $n$  is the number of input features, and  $\hat{y}_i$  and  $y_i$  is the  $i_{th}$  reconstructed and target measurement, respectively.

As in [9] and [14], the VAE is structured with two hidden layers  $(h_1, h_2)$  in the encoder,  $z$  units in the latent layer and two hidden layers  $(h_2, h_1)$  in the decoder. However, the number of hidden units in each layer differs from the previous studies as they are determined related to the number of input features  $n$ :

$$h_1 = \lfloor 1.2n \rfloor, \quad h_2 = \lfloor h_1/2 \rfloor, \quad z = \lfloor h_2/2 \rfloor \quad (3)$$

where  $\lfloor \cdot \rfloor$  is round down symbol.

### IV. EXPERIMENTAL STUDY

The following experimental study, uses Microsoft Windows 10, Java 8, “deeplearning4j” version 1.0.0-beta4 [25] as the deep learning library, and NVIDIA GeForce GTX 1060 6 GB as the graphics processing unit.

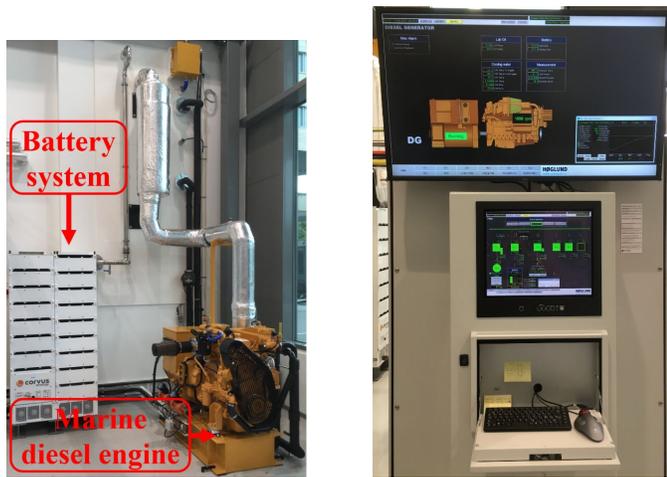


Fig. 2. The battery system, the marine diesel engine, and the automation system used for collecting the data sets.

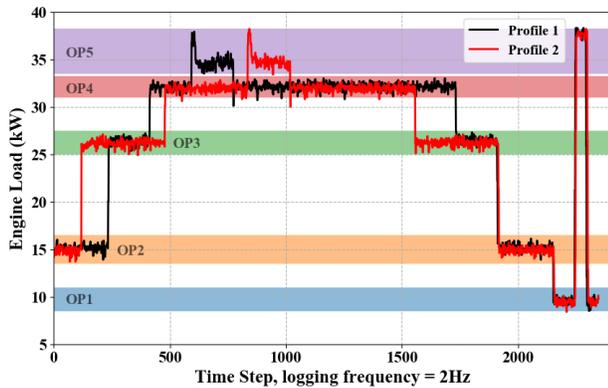


Fig. 3. Engine load profile 1 and 2.

#### A. Data sets

A hybrid power lab, established by the Department of Ocean Operations and Civil Engineering at the Norwegian University of Science and Technology in Aalesund, is used to collect the data sets. The lab intends to research ship autonomy. As seen in Figure 2, the lab includes a small marine diesel engine, a marine battery system, and a marine automation system to control the facilities. The power produced is supplied back to the power grid to simulate load variations in the system.

During the data collection process, the engine is driven by two different engine load profiles. As similar to [14], the two engine load profiles aim to replicate real-life autonomous ferry crossings on the west coast of Norway. First, the ferry is off-loading and on-loading vehicles before it leaves shore at a safe and constant velocity. Then, the ferry speeds to a suitable velocity. This velocity remains constant until it decreases safely. Finally, the ferry breaks just before it docks. In common, the two profiles are exposed to the same magnitudes of engine loads, but the length of each engine load differs to reflect different environmental conditions. Figure 3 shows the two engine load profiles, profile 1 and profile 2.

In this study, two fault-types are used for validation of the proposed algorithm. These are the air filter fault, the



Fig. 4. The restriction and bleed device used to provoke the air filter and turbo fault, respectively.

TABLE I  
THE SEVEN DATA SETS COLLECTED FROM THE HYBRID POWER LAB

| Data set               | Profile | Usage      | Seconds | Hz | Time steps |
|------------------------|---------|------------|---------|----|------------|
| Normal operation       | 1       | Training   | 1173    | 2  | 2346       |
| Normal operation       | 2       | Training   | 1173    | 2  | 2346       |
| Turbo degradation      | 1       | Validation | 1173    | 2  | 2346       |
| Turbo degradation      | 2       | Validation | 1173    | 2  | 2346       |
| Air filter degradation | 1       | Validation | 1173    | 2  | 2346       |
| Air filter degradation | 2       | Validation | 1173    | 2  | 2346       |
| Cooling degradation    | 1       | Final test | 1173    | 2  | 2346       |

clogging of the air filter, and the turbo fault, malfunction of the turbocharger. The air filter fault demonstrates the effect of a clogged air filter with the use of a restriction device, as seen in Figure 4. During the data collection process, this device is gradually adjusted from fully open to 90% closed to reduce the inlet flow of air to the turbocharger. The purpose of the turbo fault is to replicate efficiency reduction in the turbocharger. This is done by installing a bleed device on the charge air pipe between the turbocharger and the engine inlet manifold, as seen in Figure 4. Gradually bleeding of air during the data collection process results in reduced air pressure to the engine combustion process. A third fault-type is used for the final test of the proposed algorithm: a malfunction of the frequency-operated fan controlling the secondary cooling system in the engine. This fault, which appears in our previous work [14], is hereinafter referred to as the cooling fault. One normal operation data set, one turbo degradation data set, and one air filter degradation data set is collected from each profile. Additionally, one cooling degradation data set is collected from profile 1. Table I summarizes the seven data sets collected from the hybrid power lab.

#### B. Feature selection

All collected data sets include 47 input features from the hybrid power lab. As discovered in [14], features belonging to the battery system and the automation system are irrelevant for detecting faults in the marine diesel engine. When such features are removed, the VAE will provide a reconstruction process with higher degradation relevance. Additionally, fea-

TABLE II  
FEATURE SELECTION FOR THE MARINE DIESEL ENGINE

| Index | Description                                 | Unit      |
|-------|---|-----------|
| 1     | Boost pressure                              | bar       |
| 2     | Engine load                                 | kW        |
| 3     | Engine cooling water temperature            | °C        |
| 4     | Engine exhaust gas temperature              | °C        |
| 5     | Cooling water temperature out of the engine | °C        |
| 6     | Engine speed                                | rpm       |
| 7     | Diesel generator cooling water flow         | liter/min |
| 8     | Simulated propulsion load                   | kW        |
| 9     | Cooling fan speed controller                | rpm       |

tures with constant measurements are removed since these features provide no degradation information. The Pearson correlation analysis is also used to detect the linear relationship between the input features. If two input features have a high linear relationship, they likely contain redundant information. Then, expert human domain knowledge (HDK) is used to determine which of the redundant input features has less degradation relevance. Actually, in this study, the HDK is acquired from an engine chief engineer with 13 years of sailing experience and three years of experience with the development of a health monitoring system for rotating machinery. The redundant features are removed accordingly.

HDK is also used to remove inaccurate and unknown feature measurements concerning the marine diesel engine. For instance, the cooling water temperature to the engine is removed since it is considered an unknown parameter. This feature is affected by the outdoor temperature, and hence, it varies when data sets are collected at different dates and seasons. Fuel consumption is also removed from the data sets. While it is an important feature for detecting faults in the combustion process in the engine, the measurements obtained from the automation system were too inaccurate to be used in this study. Ultimately, nine input features, which are intended to reflect all degradation patterns in the marine diesel engine, remain in all data sets. Table II lists the final input features.

### C. Multi-regime operating conditions and normalization

As seen in Figure 3, the engine load changes drastically during the ferry crossing operation in both profiles. As a result, feature measurements are highly connected to the engine loads. This causes the feature measurements in the normal operation data to differ strongly between different engine loads. Thus, proper data pre-processing, in terms of multi-regime normalization, is necessary to present the actual normal operation phenomena for the VAE during the training phase [13].

Obviously, both profiles fall into five distinct operating conditions based on the engine load. First, the normal operation data sets in Table I are split into five data sets each based on the five operating conditions. Each feature in these data sets is then scaled with zero mean and unit variance (z-score) normalization:

$$\bar{x}_n^o = \frac{x_n^o - \mu^o}{\sigma^o} \quad (4)$$

where  $x_n$  is the input feature,  $n = 1, 2, \dots, 9$ , in operating condition  $o$ , and  $\mu$  and  $\sigma$  is the population mean and population

TABLE III  
HYPER-PARAMETERS

| Hyper-parameter        | Method/Value                |
|------------------------|-----------------------------|
| Activation function    | Rectified Linear Unit       |
| Learning rate          | $1 \cdot 10^{-3}$           |
| $l2$ regularization    | $1 \cdot 10^{-4}$           |
| Optimization algorithm | Stochastic Gradient Descent |
| Optimizer              | Adam                        |
| Weight initialization  | Xavier                      |

standard deviation of that feature. This yields five different normalization statistics, one for each operating condition. Finally, these normalization statistics are applied both to the raw normal operation data in the training phase and to the raw faulty degradation data in the anomaly detection. To apply different normalization statistics, the engine load is monitored at each time step.

### D. Training phase and anomaly detection

In the training phase, a VAE is established and trained on both normal operation data sets subjected to multi-regime normalization. An early stopping policy is utilized to reconstruct the normal operation data as precisely as possible by monitoring the average reconstruction error of all mini-batches. If the number of epochs with no decrease in the average reconstruction error is greater than four, the training phase is ended. Then, the VAE, in the epoch with the lowest average reconstruction error, is stored and used for anomaly detection. The mini-batch size is set to 128. The VAE is configured with hyper-parameters that provided great success for maritime components in [9]. These are shown in Table III.

In terms of time series data, it is practical to consider three categories of anomalies: point, collective, and contextual [11]. Point anomalies are single values that differ from previous values, collective anomalies are entire sequences of values that are anomalous, and contextual anomalies are single values

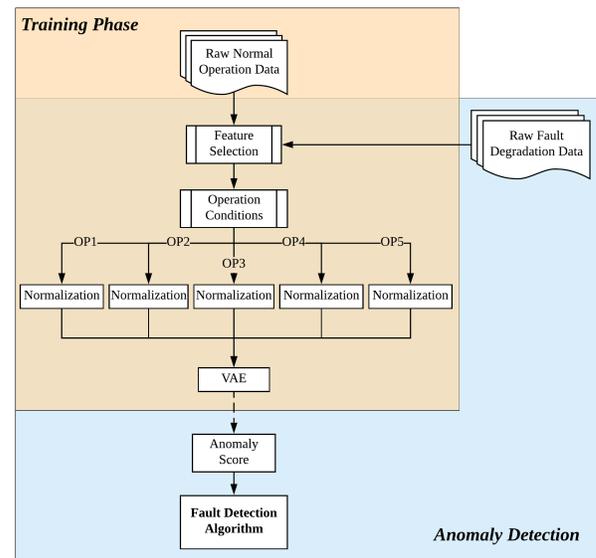


Fig. 5. A complete flowchart of the training phase and anomaly detection.

that are not different from previous values yet are anomalous concerning local values [22]. The nature of degradation in both the turbo and air filter fault is highly connected to the operating conditions and therefore they should be regarded as contextual anomalies. However, the VAE is only able to detect point anomalies. To detect contextual anomalies, the VAE has to be applied within a context [11]. We consider the five operating conditions as different contexts. Thus, the multi-regime normalization statistics also need to be applied to the faulty degradation data to merge the five different contexts into one context. In this way, the VAE can be used for anomaly detection and to estimate an anomaly score at each time step. Figure 5 shows a complete flowchart of the training phase and anomaly detection.

### E. Fault detection algorithm

1) *Online fault detection:* The anomaly score in faulty degradation data  $AS_d$  is estimated by using the trained VAE to calculate the MSE, Eq. 2, at each time step  $t$ . Then, the algorithm generates three sliding windows of length  $w$  to smooth  $AS_d$ :

$$w = \frac{T_d}{p} \quad (5)$$

where  $T_d$  is the total number of time steps in the faulty degradation data and  $p$  is an adjustable parameter.  $p$  determines the magnitude of smoothing conducted on  $AS_d$ . Hence, careful tuning of  $p$  is necessary since excessive smoothing might obscure important degradation trends. The three windows slide across  $AS_d$  for each  $t$ . A distance equivalent to  $w$  is used between each window. Simultaneously, the average anomaly score  $AS_{d,avg}$  is computed in each window. Additionally, the velocity  $v_d$  between windows 1 and 2 and between windows 2 and 3, and the acceleration  $a_d$  between the two velocities are calculated. Finally, the velocity fault time step  $\hat{f}_{t,v}$  and the acceleration fault time step  $\hat{f}_{t,a}$  are detected when  $v_d$  and  $a_d$  exceeds their dynamic threshold limits, respectively. The proposed algorithm is shown in Algorithm 1.

Large sensor measurement deviations compared to sensor measurements in normal operation data is a strong indication of an incipient fault [9]. These deviations can, of course, be detected by utilizing  $AS_d$  or  $AS_{d,avg}$  as the fault detectors. However, both  $AS_d$  and  $AS_{d,avg}$  will vary between different fault-types since they reflect the nature of degradation. Consequently, the corresponding threshold limits will be highly fault-dependent. The main goal of the proposed algorithm is to be fault-type independent.  $v_d$  will measure the rapidity in  $AS_{d,avg}$  and indicate if one or several sensor measurements have begun to diverge swiftly from normal operation data. However,  $a_d$  will measure increases and decreases in  $v_d$ . Due to latency in the marine diesel engine,  $a_d$  might be a better indication than  $v_d$  since there is an expected time delay before the faults will result in large sensor measurement deviations. Therefore,  $v_d$  and  $a_d$  are considered as more suitable fault detectors for the algorithm since the calculations are assumed to be similar between different fault-types. Consequently, generic and fault-independent threshold limits can be acquired. These limits are further elaborated in the following paragraph.

**Algorithm 1** Algorithm for detecting the fault time step in faulty degradation data.

---

**Input:**  $T_d, AS_d, p, v_n, v_{lower}, v_{upper}, a_n, a_{lower}, a_{upper}$   
**Output:**  $\hat{f}_{t,v}, \hat{f}_{t,a}$   
*Initialization :*  
 $w \leftarrow T_d / p$   
 $v_{d,first} = \mathbf{true}$   
 $a_{d,first} = \mathbf{true}$   
*Generate three sliding windows of length  $w$  to slide across  $AS_d$  for each  $t$ .  $AS_{d,avg}$  is computed in each window. A distance equivalent to  $w$  is used between each window.*  
**for**  $t := 1$  to  $T_d$  **do**  
 $v_{d1} \leftarrow AS_{d,avg1} - AS_{d,avg2}$   
 $v_{d2} \leftarrow AS_{d,avg2} - AS_{d,avg3}$   
 $a_d \leftarrow v_{d1} - v_{d2}$   
**if** ( $v_{d,first} = \mathbf{true}$ ) **then**  
**if** ( $v_{d1} > v_n[t] + v_{upper}$  **or**  $v_{d1} < v_n[t] + v_{lower}$ ) **then**  
 $\hat{f}_{t,v} \leftarrow t - (w \cdot 1.5)$   
 $v_{d,first} = \mathbf{false}$   
**end if**  
**end if**  
**if** ( $a_{d,first} = \mathbf{true}$ ) **then**  
**if** ( $a_d > a_n[t] + a_{upper}$  **or**  $a_d < a_n[t] + a_{lower}$ ) **then**  
 $\hat{f}_{t,a} \leftarrow t - (w \cdot 2.5)$   
 $a_{d,first} = \mathbf{false}$   
**end if**  
**end if**  
**end for**  
**return**  $\hat{f}_{t,v}, \hat{f}_{t,a}$

---

2) *Generic and dynamic threshold limits:* In this study, the threshold limits are based on the velocity  $v_n$  and the acceleration  $a_n$  in the average anomaly score of normal operation data for both profiles. The procedure to measure both  $v_n$  and  $a_n$  is exactly the same as in Algorithm 1. Seven different  $p$  values, in the 30 to 90 range, are used during the experiments. In order to obtain the associated dynamic threshold limits, the minimum and maximum velocities of  $v_n$ ,  $v_{min}$ , and  $v_{max}$ , and the minimum and maximum accelerations of  $a_n$ ,  $a_{min}$  and  $a_{max}$ , are calculated for each  $p$  value in each profile. Then, a common set of upper and lower thresholds for both  $v_n$  and  $a_n$  are calculated based on the following formulas:

$$v_{upper} = \frac{|(v_{max,1} + v_{max,2}) - (v_{min,1} + v_{min,2})|}{2} \quad (6)$$

$$v_{lower} = -v_{upper} \quad (7)$$

$$a_{upper} = \frac{|(a_{max,1} + a_{max,2}) - (a_{min,1} + a_{min,2})|}{2} \quad (8)$$

$$a_{lower} = -a_{upper} \quad (9)$$

The common set of upper and lower thresholds for each  $p$  value are shown in Table IV. The limits will change dynamically through time when the upper and lower thresholds are added to  $v_n$  and  $a_n$ , as performed in Algorithm 1.

TABLE IV  
A COMMON SET OF UPPER AND LOWER THRESHOLDS FOR BOTH THE VELOCITY AND THE ACCELERATION

| $p$ | $v_{lower}$ | $v_{upper}$ | $a_{lower}$ | $a_{upper}$ |
|-----|-------------|-------------|-------------|-------------|
| 30  | -2.63       | 2.63        | -4.10       | 4.10        |
| 40  | -3.40       | 3.40        | -4.97       | 4.97        |
| 50  | -3.81       | 3.81        | -6.35       | 6.35        |
| 60  | -4.40       | 4.40        | -7.26       | 7.26        |
| 70  | -5.10       | 5.10        | -8.58       | 8.58        |
| 80  | -5.74       | 5.74        | -9.67       | 9.67        |
| 90  | -6.32       | 6.32        | -10.54      | 10.54       |

The generic and dynamic threshold limits are computed before they are applied in the fault detection algorithm. However, new engine load profiles are likely to be encountered in real-life data-driven PHM systems in autonomous ferries. Then, the computation complexity will increase since  $v_{min}$ ,  $v_{max}$ ,  $a_{min}$ , and  $a_{max}$  of the new profile have to be calculated and incorporated in Eqs. 6, 7, 8, and 9 before new fault detections can start.

## V. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this study, both velocity and acceleration calculations will be used as the fault detectors. The air filter and turbo degradation in both profiles will be used as the validation data sets for the proposed algorithm. The validation aims to discover the best performing fault detector and the most suitable  $p$  value. Seven different  $p$  values, in the 30-90 range, will be compared. A low  $p$  value might smooth the anomaly score too much and ignore significant degradation patterns. In contrast, a high  $p$  value might provide irrelevant spikes that would also affect the velocity and acceleration calculations. In the end, the final experiment will use the cooling degradation as the final test data set of the algorithm. This experiment aims to further test the fault-type independence of the algorithm.

### A. Validation

To validate both  $\hat{f}_{t,v}$  and  $\hat{f}_{t,a}$ , the true fault time step  $f_t$  has to be determined. Since both the air filter fault and the turbo fault in both profiles are provoked gradually during the data collection process,  $f_t$  can not be decided based a recorded time step. Thus,  $f_t$  is determined based on expert HDK. The boost pressure is the key feature to monitor for fault detection for both fault-types. As already mentioned, both faults-types are highly connected to the engine loads and subjected to different nature of degradation. Therefore,  $f_t$  is determined where the deviation in boost pressure between normal operation data and faulty degradation data in percentage is largest. The determined  $f_t$  for both fault-types in both profiles is shown in Table V.

Table VI shows  $\hat{f}_{t,v}$  and  $\hat{f}_{t,a}$  for each  $p$  value in both profiles for both fault-types. The accuracy evaluations,  $Acc_v$  and  $Acc_a$ , are based on the following formula:

$$Acc(\%) = \left(1 - \frac{\|\hat{f}_t - f_t\|}{T_d}\right) \cdot 100 \quad (10)$$

where  $Acc(\%)$  can be considered as the distance between the detection and  $f_t$ . In the following discussions, a satisfactory

TABLE V  
THE TRUE FAULT TIME STEP  $f_t$

| Fault-type | Profile | Largest deviation in boost pressure (%) | $f_t$ |
|------------|---------|---|-------|
| Air filter | 1       | 15.79                                   | 1670  |
|            | 2       | 10.53                                   | 1433  |
| Turbo      | 1       | 21.05                                   | 1431  |
|            | 2       | 21.05                                   | 1427  |

TABLE VI  
VALIDATION: THE TRUE FAULT TIME STEP  $f_t$  COMPARED TO THE DETECTED FAULT TIME STEP  $\hat{f}_t$

| Fault-type | Profile | $f_t$ | $p$ | $w$ | $\hat{f}_{t,v}$ | $Acc_v(\%)$ | $\hat{f}_{t,a}$ | $Acc_a(\%)$ |
|------------|---------|-------|-----|-----|-----------------|-------------|-----------------|-------------|
| Air filter | 1       | 1670  | 30  | 78  | 1255            | 82.31       | 1502            | 92.84       |
|            |         |       | 40  | 58  | 1278            | 83.29       | 1609            | 97.40       |
|            |         |       | 50  | 46  | 1289            | 83.76       | 1648            | 99.06       |
|            |         |       | 60  | 39  | 1549            | 94.84       | 1660            | 99.57       |
|            |         |       | 70  | 33  | 1566            | 95.57       | 1674            | 99.83       |
|            |         |       | 80  | 29  | 1706            | 98.47       | 1680            | 99.57       |
|            |         |       | 90  | 26  | 1709            | 98.34       | 1682            | 99.49       |
|            | 2       | 1433  | 30  | 78  | 1362            | 96.97       | 1428            | 99.79       |
|            |         |       | 40  | 58  | 1392            | 98.25       | 1445            | 99.49       |
|            |         |       | 50  | 46  | 1404            | 98.76       | 1458            | 98.93       |
|            |         |       | 60  | 39  | 1532            | 95.78       | 1483            | 97.87       |
|            |         |       | 70  | 33  | 1540            | 95.44       | 0               | 38.92       |
|            |         |       | 80  | 29  | 0               | 38.92       | 0               | 38.92       |
|            |         |       | 90  | 26  | 0               | 38.92       | 0               | 38.92       |
| Turbo      | 1       | 1431  | 30  | 78  | 731             | 70.16       | 693             | 68.54       |
|            |         |       | 40  | 58  | 771             | 71.87       | 745             | 70.76       |
|            |         |       | 50  | 46  | 786             | 72.51       | 752             | 71.06       |
|            |         |       | 60  | 39  | 794             | 72.85       | 1347            | 96.42       |
|            |         |       | 70  | 33  | 368             | 54.69       | 1362            | 97.06       |
|            |         |       | 80  | 29  | 1395            | 98.47       | 1374            | 97.57       |
|            |         |       | 90  | 26  | 1399            | 98.64       | 1381            | 97.87       |
|            | 2       | 1427  | 30  | 78  | 951             | 79.71       | 892             | 77.20       |
|            |         |       | 40  | 58  | 979             | 80.90       | 929             | 78.77       |
|            |         |       | 50  | 46  | 991             | 81.42       | 1329            | 95.82       |
|            |         |       | 60  | 39  | 1005            | 82.01       | 1347            | 96.59       |
|            |         |       | 70  | 33  | 1387            | 98.29       | 1361            | 97.19       |
|            |         |       | 80  | 29  | 1393            | 98.55       | 1371            | 97.61       |
|            |         |       | 90  | 26  | 1397            | 98.72       | 1378            | 97.91       |

TABLE VII  
VALIDATION: THE AVERAGE ACCURACY FOR EACH  $p$  VALUE

| $p$ | $w$ | Avg. $Acc_v(\%)$ | Avg. $Acc_a(\%)$ |
|-----|-----|------------------|------------------|
| 30  | 78  | 82.29            | 84.59            |
| 40  | 58  | 83.58            | 86.60            |
| 50  | 46  | 84.11            | 91.22            |
| 60  | 39  | 86.37            | <b>97.61</b>     |
| 70  | 33  | 86.00            | 83.25            |
| 80  | 29  | 83.60            | 83.42            |
| 90  | 26  | 83.65            | 83.55            |

accuracy is considered to be above 95%. For the air filter fault in profile 1,  $\hat{f}_{t,v}$  provides satisfactory accuracy by  $p$  values between 70 and 90, while  $\hat{f}_{t,a}$  provides satisfactory accuracy by  $p$  values between 40 and 90. On the other hand, for the air filter fault in profile 2,  $\hat{f}_{t,v}$  provides satisfactory accuracy by  $p$  values between 30 and 70, while  $\hat{f}_{t,a}$  provides satisfactory accuracy by  $p$  values between 30 and 60. As Table V reflects, the air filter fault in profile 2 is subjected to a lower deviation in boost pressure than the air filter fault in profile 1. As a consequence, the air filter fault in profile 2 is subjected to lower magnitudes of both velocity and acceleration calculations, and hence, requires smaller upper and lower thresholds. As Table IV shows, low  $p$  values result in smaller upper and lower thresholds. This issue reflects the difficulty of creating generic and dynamic threshold limits even for the same fault-type that is subjected to different environmental conditions in the form of different engine load

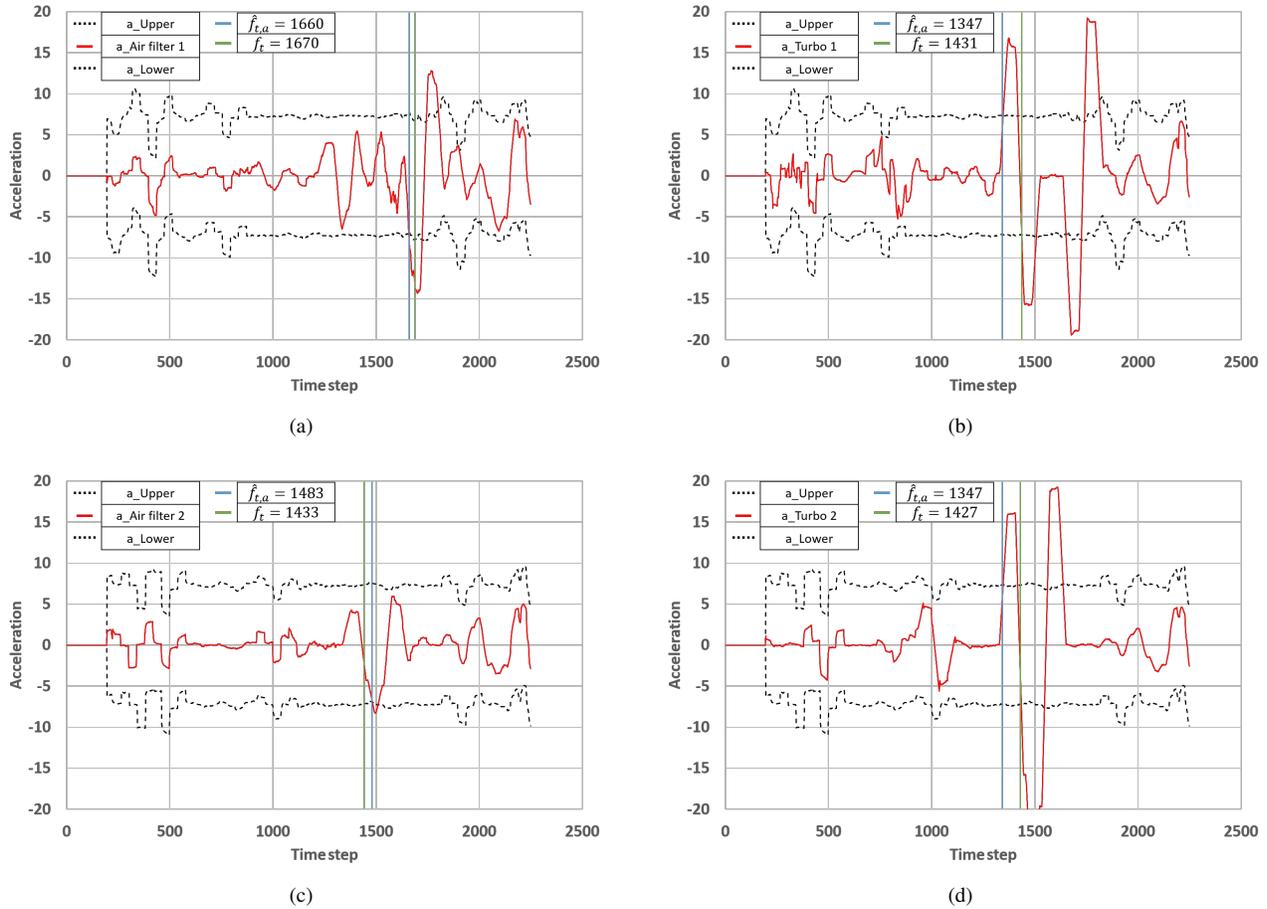


Fig. 6. Automatic fault detection where  $p = 60$  and the acceleration is used as the fault detector. (a) Air filter fault in engine load profile 1. (b) Turbo fault in engine load profile 1. (c) Air filter fault in engine load profile 2. (d) Turbo fault in engine load profile 2.

profiles.

For the turbo fault in profile 1,  $\hat{f}_{t,v}$  provides satisfactory accuracy by the  $p$  values 80 and 90, while  $\hat{f}_{t,a}$  provides satisfactory accuracy by  $p$  values between 60 and 90. Similarly, for the turbo fault in profile 2,  $\hat{f}_{t,v}$  provides satisfactory accuracy by  $p$  values between 70 and 90, while  $\hat{f}_{t,a}$  provides satisfactory accuracy by  $p$  values between 50 and 90. Also as seen in Table V, the turbo fault in both profiles are subjected to a deviation of 21.05%, almost twice the deviation compared to the air filter fault in profile 2. This results in larger magnitudes of both velocity and acceleration calculations. Thus, the turbo fault in both profiles provides the highest accuracies by high  $p$  values and corresponding large upper and lower dynamic threshold limits.

To determine the best performing fault detector and the most suitable  $p$  value for both fault-types, the average velocity and acceleration accuracy for each  $p$  value is calculated, as shown in Table VII. When  $p = 60$ , the acceleration provides the highest average accuracy of 97.61%. Therefore, the acceleration is considered the most fault-independent fault detector. Figure 6 shows the acceleration calculations and the corresponding dynamic threshold limits when  $p = 60$  for both fault-types in both profiles. It is worth mentioning that the acceleration calculations and the dynamic threshold limits are

not plotted before the entire sliding window operation is active. In other words, the initial 195 time steps are plotted as zeros ( $w(60) \cdot 5 = 195$ ).

### B. Final test

The main intention of the final test of the proposed algorithm is to further test its independence towards different fault-types. The cooling degradation data in profile 1 is used for this purpose as this fault exhibits a totally different nature of degradation compared to both the air filter fault and the turbo fault. Thus, it can be considered to be new field data that the algorithm has never seen before. To evaluate the fault detection, the true fault time step  $f_t$  for the cooling fault is also determined based on expert HDK. When the cooling water temperature increases 85 °C,  $f_t$  is determined to be 1713.

As discovered in the validation, the acceleration is the best performing fault detector when  $p = 60$ . These settings are therefore used in the final test. As Table VIII shows, the algorithm detects the cooling fault with an accuracy of 97.66%. Also noted, both in the validation and the final test the trend is that the acceleration provides early detections, i.e.  $\hat{f}_{t,acc} < f_t$ , when  $p = 60$ . However, early detections with a corresponding high accuracy are considered as valid detections since there is an expected time delay in the marine

TABLE VIII  
FINAL TEST: THE TRUE FAULT TIME STEP  $f_t$  COMPARED TO THE DETECTED FAULT TIME STEP  $\hat{f}_t$  FOR COOLING DEGRADATION DATA

| Fault-type | Profile | $f_t$ | $p$ | $w$ | $\hat{f}_{t,a}$ | $Acc_a(\%)$ |
|------------|---------|-------|-----|-----|-----------------|-------------|
| Cooling    | 1       | 1713  | 60  | 39  | 1658            | 97.66       |

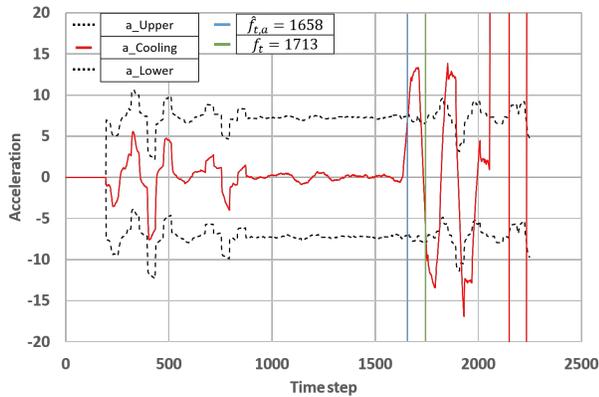


Fig. 7. Automatic fault detection where  $p = 60$  and the acceleration is used as the fault detector for cooling degradation data.

diesel engine before the faults will result in large sensor measurements deviations. Figure 7 shows the acceleration calculations and the corresponding dynamic threshold limits for the fault detection of the cooling degradation data. The final test proves that the algorithm is fault-type independent.

## VI. CONCLUSION AND FUTURE WORK

This paper has analyzed and proposed a fault-type independent spectral anomaly detection algorithm for marine diesel engine degradation in autonomous ferries where a VAE is used as the DNN. To do so, three fault-types with different nature of degradation have been used during the experiments. Both normal operation data and faulty degradation data have been collected from two different engine load profiles. These profiles aim to replicate real autonomous ferry crossing operations that might affect the ferry.

In the validation of the proposed algorithm, the acceleration has proven to be the most fault-independent fault detector, providing an average accuracy of 97.61%. Additionally, the acceleration achieved an accuracy of 97.66% in the final test of the algorithm. Thus, the algorithm has proved its independence of fault-types with different nature of degradation related to the marine diesel engine.

In this study, the engine loads were divided into five distinct operating conditions manually to do multi-regime normalization. However, if new operating conditions are encountered in real-life systems, this process has to be automated. For instance, through unsupervised clustering algorithms, such as the K-Means algorithm. One has to remember that fault detection is only the first action to be performed in a real-life data-driven PHM system. However, the detected fault time steps obtained from the faulty degradation data can be used to automatically label the data to account for both fault classification and RUL predictions. Also, due to the

VAE's generative characteristics, it is possible to derive the reconstruction of the data to analyze the underlying cause of the fault to do fault isolation. Our future work will include these crucial actions.

## ACKNOWLEDGMENT

This work was supported by the Norwegian University of Science and Technology within the Department of Ocean Operations and Civil Engineering under project no. 90329106. The authors would like to thank Digital Twins For Vessel Life Cycle Service and the Research Council of Norway, grant no. 280703.

## REFERENCES

- [1] X. Cheng, G. Li, A. L. Ellefsen, S. Chen, H. P. Hildre, and H. Zhang, "A novel densely connected convolutional neural network for sea state estimation using ship motion data," *IEEE Transactions on Instrumentation and Measurement*, pp. 1–1, 2020.
- [2] R. Skulstad, G. Li, T. I. Fossen, B. Vik, and H. Zhang, "Dead reckoning of dynamically positioned ships: Using an efficient recurrent neural network," *IEEE Robotics Automation Magazine*, vol. 26, no. 3, pp. 39–51, Sep. 2019.
- [3] E. Jokioinen, "Remote and autonomous ships - the next steps: Introduction," *Rolls-Royce, Buckingham Gate, London: The Advanced Autonomous Waterborne Applications (AAWA)*, pp. 4–14, 2016.
- [4] O. Levander, "Autonomous ships on the high seas," *IEEE Spectrum*, vol. 54, no. 2, pp. 26–31, 2017.
- [5] L. Kretschmann, H.-C. Burmeister, and C. Jahn, "Analyzing the economic benefit of unmanned autonomous ships: An exploratory cost-comparison between an autonomous and a conventional bulk carrier," *Research in transportation business & management*, vol. 25, pp. 76–86, 2017.
- [6] A. L. Ellefsen, V. Æsøy, S. Ushakov, and H. Zhang, "A comprehensive survey of prognostics and health management based on deep learning for autonomous ships," *IEEE Transactions on Reliability*, vol. 68, no. 2, pp. 720–740, 2019.
- [7] K. E. Knutsen, G. Manno, and B. J. Vartdal, "Beyond condition monitoring in the maritime industry," *DNV GL Strategic Research & Innovation Position Paper*, 2014.
- [8] P. W. Kalgren, C. S. Byington, M. J. Roemer, and M. J. Watson, "Defining phm, a lexical evolution of maintenance and logistics," in *2006 IEEE Autoestcon*, Sept 2006, pp. 353–358.
- [9] A. L. Ellefsen, E. Bjørlykhaug, V. Æsøy, and H. Zhang, "An unsupervised reconstruction-based fault detection algorithm for maritime components," *IEEE Access*, vol. 7, pp. 16 101–16 109, 2019.
- [10] J. An and S. Cho, "Variational autoencoder based anomaly detection using reconstruction probability," *SNU Data Mining Center - Special Lecture on IE*, vol. 2, no. 1, 2015.
- [11] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, vol. 41, no. 3, p. 15, 2009.
- [12] J. A. P. Rubio, F. Vera-García, J. H. Grau, J. M. Cámara, and D. A. Hernandez, "Marine diesel engine failure simulator based on thermodynamic model," *Applied Thermal Engineering*, vol. 144, pp. 982–995, 2018.
- [13] O. Bektas, J. A. Jones, S. Sankararaman, I. Roychoudhury, and K. Goebel, "A neural network filtering approach for similarity-based remaining useful life estimation," *The International Journal of Advanced Manufacturing Technology*, vol. 101, no. 1-4, pp. 87–103, 2019.
- [14] A. L. Ellefsen, X. Cheng, F. T. Holmeset, S. Ushakov, V. Æsøy, and H. Zhang, "Automatic fault detection for marine diesel engine degradation in autonomous ferry crossing operation," in *2019 IEEE International Conference on Mechatronics and Automation (ICMA)*, Aug 2019, pp. 2195–2200.
- [15] G. Wu, "Fault detection method for ship equipment based on bp neural network," in *2018 International Conference on Robots & Intelligent System (ICRIS)*. IEEE, 2018, pp. 556–559.
- [16] G. Xu, M. Liu, Z. Jiang, W. Shen, and C. Huang, "Online fault diagnosis method based on transfer convolutional neural networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 2, pp. 509–520, Feb 2020.

- [17] J. Sun, C. Yan, and J. Wen, "Intelligent bearing fault diagnosis method combining compressed data acquisition and deep learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 1, pp. 185–195, Jan 2018.
- [18] J. E. Siegel, S. Pratt, Y. Sun, and S. E. Sarma, "Real-time deep neural networks for internet-enabled arc-fault detection," *Engineering Applications of Artificial Intelligence*, vol. 74, pp. 35 – 42, 2018.
- [19] A. S. Zymaris, Ø. Å. Alnes, K. E. Knutsen, and N. M. Kakalis, "Towards a model-based condition assessment of complex marine machinery systems using systems engineering," in *Proc. 3rd Eur. Conf. Prognostics Health Manage. Soc.*, 2016, pp. 1–15.
- [20] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," *arXiv preprint arXiv:1901.03407*, 2019.
- [21] D. Park, Y. Hoshi, and C. C. Kemp, "A multimodal anomaly detector for robot-assisted feeding using an lstm-based variational autoencoder," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1544–1551, 2018.
- [22] K. Hundman, V. Constantinou, C. Laporte, I. Colwell, and T. Soderstrom, "Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 2018, pp. 387–395.
- [23] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.
- [24] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.
- [25] "Eclipse deeplearning4j development team, deeplearning4j: Open-source distributed deep learning for the JVM," *Apache Software Foundation License 2.0*, <http://deeplearning4j.org>, 2020.



**André Listou Ellefsen** received his Master degree in Subsea Technology from the Norwegian University of Science and Technology (NTNU), Trondheim, Norway, in 2016. He is currently pursuing the Ph.D. degree with NTNU, Aalesund, Norway, as part of the Mechatronics Laboratory, Department of Ocean Operations and Civil Engineering. His current research interests include artificial intelligence, deep learning, decision support, predictive maintenance, prognostics and health management, and digital twins.



**Peihua Han** received his Bachelor and Master degree in Department of Architecture and Civil Engineering from Zhejiang University, China, in 2019. He is currently pursuing the Ph.D. degree with Norwegian University of Science and Technology (NTNU), Aalesund, Norway, as part of the Mechatronics Laboratory, Department of Ocean Operations and Civil Engineering. His current research interests include fault diagnosis and prognostics, predictive maintenance, machine learning, and uncertainty qualification.



**Xu Cheng** received his Master degree in Computer Science and Technology from Zhejiang University of Technology, Hangzhou, China, in 2015. He is currently working at NTNU, Aalesund, Norway, as part of the Mechatronics Laboratory within the Department of Ocean Operations and Civil Engineering, as a Ph.D. candidate. His current research interests include sensitivity analysis, neural network, ship motion modeling.



**Finn Tore Holmeset** has a background as a marine chief engineer with more than 13 years seagoing experience onboard different ship types and holding a professional letter as an automation mechanic as well. He received a Master degree in Management of Demanding Maritime Operations from the Norwegian University of Science and Technology (NTNU), Aalesund, Norway, in 2018. His current work is to develop the machinery lab at NTNU Aalesund and support various research work projects ongoing at the institute.



**Prof. Dr. Vilmar Æsøy** graduated from NTNU in 1989, and continued his research on natural gas fueled marine engines at NTNU/MARINTEK to 1997. In 1996 he received his PhD degree for his research on natural gas ignition and combustion through experimental investigations and numerical simulations. During the research period 1989–1997 he was engaged in several large R&D projects developing gas fueled engines and fuel injection systems for the diesel engine manufacturers, Wärtsilä and Bergen Diesel (Roll-Royce). From 1998 to 2002, he worked as R&D manager for Rolls-Royce Marine Deck Machinery. Since 2002 he has been employed in teaching at Aalesund University College, developing and teaching courses in marine product and systems design on bachelor and master level. From January 2010 he received the "green ship machinery" professorship. His special research interest is within the field of energy and environmental technology, with focus on combustion engines and the need for more environmental friendly and energy efficient systems.



**Prof. Dr. Houxiang Zhang** (IEEE Member 2004–IEEE Senior Member 2012) received his Ph.D. degree on Mechanical and Electronic Engineering from Robotics Institute, Beihang University in 2003. From 2004, he worked as Postdoctoral fellow, senior researcher at the Institute of Technical Aspects of Multimodal Systems (TAMS), Department of Informatics, Faculty of Mathematics, Informatics and Natural Sciences, University of Hamburg, Germany. In Feb. 2011, he finished the Habilitation on Informatics at University of Hamburg. Dr. Zhang joined the NTNU, Norway in April 2011 where he is a Professor on Mechatronics. Dr. Zhang has engaged into two main research areas: 1) Biological robots and modular robotics, especially on biological locomotion control, 2) Virtual prototyping in demanding marine operation. He has applied for and coordinated more than 20 projects supported by Norwegian Research Council (NFR), German Research Council (DFG), and industry. In these areas, he has published over 160 journal and conference papers as author or co-author. Dr. Zhang has received four best paper awards, and four finalist awards for best conference paper at International conference on Robotics and Automation.