Anomaly Detection and Fault Diagnostics for Underwater Gliders Using Deep Learning

Peng Wu
Department of Mechanical Engineering
University College London, London
London, UK
peng.wu.14@ucl.ac.uk

Catherine A. Harris
Marine Autonomous and Robotic Systems
National Oceanography Centre
Southampton, UK

Georgios Salavasidis
Marine Autonomous and Robotic Systems
National Oceanography Centre
Southampton, UK

Izzat Kamarudzaman
Marine Autonomous and Robotic Systems
National Oceanography Centre
Southampton, UK

Alexander B. Phillips
Marine Autonomous and Robotic Systems
National Oceanography Centre
Southampton, UK

Giles Thomas
Department of Mechanical Engineering
University College London, London
London, UK

Enrico Anderlini
Department of Mechanical Engineering
University College London, London
London, UK
e.anderlini@ucl.ac.uk

***Abstract*—Near Real-Time (NRT) anomaly detection and fault diagnostics for underwater gliders are challenging because satellite connections with limited bandwidth allow only decimated data to be sent back from the remote vehicle, whilst on-board systems are constrained by power and computational limits. Currently, anomaly detection and fault diagnostics for such vehicles require intensive monitoring from the operating pilots, which prohibits large scale deployments with multiple vehicles. This paper presents a system with NRT anomaly detection and fault diagnostics for multi-vehicle underwater glider fleets based on Bidirectional Generative Adversarial Networks with assistive hints. The unsupervised anomaly detection system is applied to assist in annotating deployment datasets to train a supervised fault diagnostics model. The results suggest that the anomaly detection system has successfully detected different types of anomalies, validated against model-based and rule-based approaches. In addition, the supervised fault diagnostics system has achieved high fault diagnostics accuracy on the test dataset.**

***Index Terms*—underwater glider, anomaly detection, generative adversarial networks, fault diagnostics**

I. Introduction

Underwater Gliders (UGs) (Fig. 1) are a type of Autonomous Underwater Vehicle (AUV) that are being used extensively for long-term observation of key physical oceanographic parameters [1]. They operate remotely at a low surge speed of approximately 0.3 $\mathrm{m\,s^{-1}}$, with deployments of several months [2]. However, developing Near Real-Time (NRT) anomaly detection and fault diagnostics systems for such vehicles remains challenging as decimated sensor data can only be transmitted off-board periodically during operations when the UG is on the surface.

As part of an ongoing collaboration, the authors have previously developed anomaly detection systems for UGs via different approaches. In [3], a simple but effective system was developed to detect the wing loss using the roll angle. In [4], system identification techniques were employed to detect changes in model parameters which further successfully deduced simulated and natural marine growth. Anderlini, et al. [5] further conducted a field test to validate a marine growth detection system for UGs using ensembles of regression trees. In [6], the use of a range of deep learning techniques was investigated to achieve over-the-horizon anomaly detection for UGs. In [7], an anomaly detection system based on an improved Bi-directional Generative Adversarial Network (BiGAN) was prototyped to enable generic anomaly detection for different types of anomalies.

For UGs operated over the horizon, some faults can only be revealed when the faulty UGs are recovered. Also, it is not clear when the faults developed. Some undetected faults can lead to critical failures and the loss of vehicle and/or data cargo. Therefore, it is essential to understand the actual causes of high anomaly scores during remote monitoring to allow operators to take appropriate mitigations to minimise

Corresponding author: Peng Wu (Email:peng.wu.14@ucl.ac.uk). This project is supported by the Assuring Autonomy International Programme, a partnership between Lloyd's Register Foundation and the University of York. C. A. Harris, G. Salavasidis, A. Lorenzo Lopez, I. Kamarudzaman and A. B. Phillips' contributions are also funded under the NERC/ISCF Oceanids programme. The results for simulated biofouling (unit 492) were obtained thanks to funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 731103.

Fig. 1. Slocum G2 underwater glider with Ocean Microstructure.

subsequent risks and maximise the successful delivery of the remainder of the deployment. This paper further compares the results acquired in [7] with other baseline approaches. In addition, a new supervised fault diagnostics method for UGs is proposed. The BiGAN-based anomaly detection system is applied to estimate when the faults are developed, such that the training dataset for the supervised fault diagnostics model can be accurately annotated. The results suggest that the BiGAN-based anomaly detection system has successfully detected different types of anomalies, in good agreement with model-based and rule-based approaches. The supervised fault diagnostics system has achieved high fault diagnostics accuracy on the available test dataset.

II. Method

Figure 2 shows the workflow of the anomaly detection and fault diagnostics for underwater gliders using deep learning. This workflow comprises two parts, i.e. unsupervised anomaly detection and supervised fault diagnostics. The unsupervised anomaly detection system is developed to alert the operators about the occurrence of an abnormal vehicle status that deviates from the normal baseline operating pattern. The developed unsupervised anomaly detection system is applied to assist in annotating the datasets with anomalies, as the exact times when the anomalies developed are unknown. With this approach, the training dataset for the supervised fault diagnostics can be accurately annotated.

A. Anomaly detection using BiGAN

Figure 3 shows the structure of the BiGAN [8]. The discriminator D, the generator G and encoder E are trained concurrently such that D maximises the probability of assigning the correct label to both tuples $(x, E(x))$ and $(G(z), z)$. To fool the discriminator, the encoder E and the generator G must learn to invert one another. The encoder E maps data x to its latent feature z. A trained BiGAN encoder E can be a feature representation for related semantic tasks. For instance, z can be a representation of data x. The discriminator D discriminates $(x, E(x))$ and $(G(z), z)$, which is different from the standard GAN [9]. Readers may refer to [7] for more details of the BiGAN-based anomaly detection system. The training objective of the BiGAN is:

$$\min_{G,E}\max_{D} V(D,E,G) = \mathbb{E}_{x\sim p_{data}(x)}\left[\log D\left(x, E(x)\right)\right] + \mathbb{E}_{z\sim p_z(z)}\left[\log\left(1 - D(G(z), z)\right)\right] \quad (1)$$

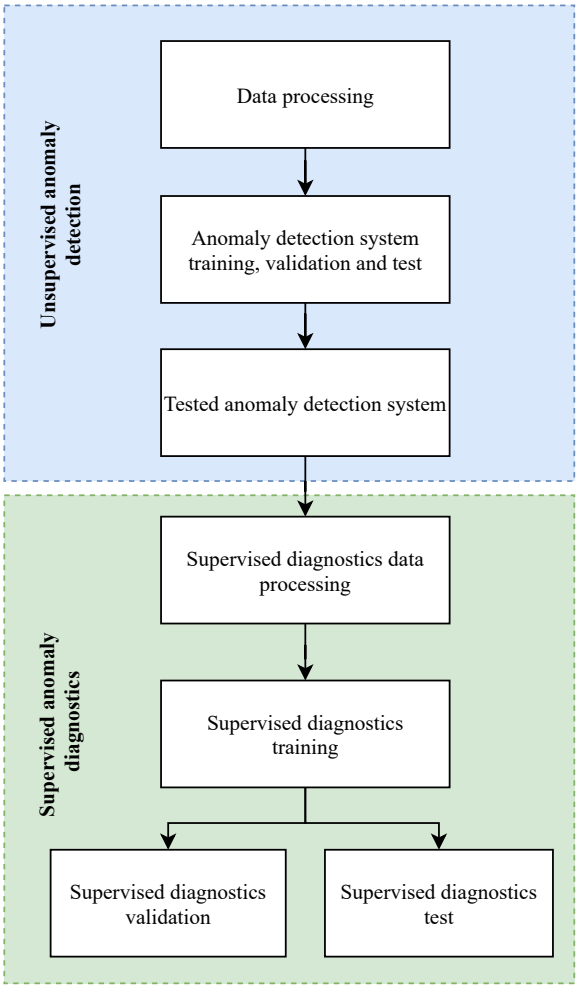

Fig. 2. Workflow of the anomaly detection and fault diagnostics for underwater gliders using deep learning.

In addition to the original BiGAN, assistive hints comprising data reconstruction and neural network feature errors have been added to guide the anomaly detection system training [7].

To fool the discriminator D, the encoder E and generator G must learn to invert one another [8]. The authors [7] use the reconstruction error between the input sample x and reconstructed data $\hat{x}$ to assist the BiGAN training, i.e. using the L^2 norm between x and $\hat{x}$ via E and G:

$$\mathcal{L}_{re} = \frac{1}{n_x}\left\|x - \hat{x}\right\|_2 \quad (2)$$

where $\hat{x} = G(E(x))$ and n_x is the number of input data elements.

In addition, the discriminator's layer output before its output layer is extracted as an additional hint provided by the discriminator:

$$\mathcal{L}_{fe} = \frac{1}{n_f}\left\|f(x, z') - f(\hat{x}, z')\right\|_2 \quad (3)$$

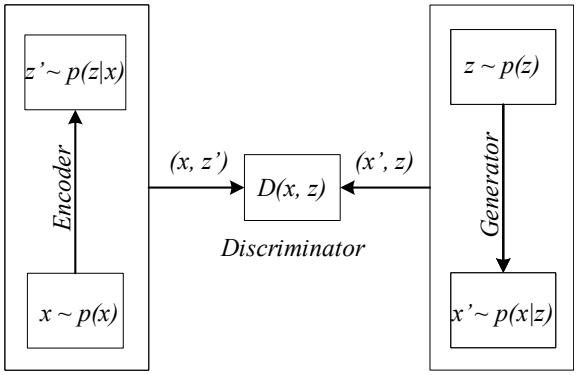

Fig. 3. Structure of the BiGAN [8].

where $z' = E(x)$ and n_f is the feature layer's number of neurons.

Combining $\mathcal{L}_{re}$ and $\mathcal{L}_{fe}$, the assistive hint loss function is [7]:

$$\mathcal{L}_{hint} = \frac{\kappa}{n_x} \|x - \hat{x}\|_2 + \frac{1}{n_f} \|f(x, z') - f(\hat{x}, z')\|_2 \quad (4)$$

where κ is an adjustable hyperparameter. In this study, $\kappa = 2$. It is worth noting that in the validation and test phases, the residual of $\mathcal{L}_{hint}$ is the anomaly score that indicates the degree of anomalies.

Figure 4 details the unsupervised workflow of the proposed anomaly detection system using BiGAN for UGs. The model is trained using full time-series datasets recovered from baseline deployments with no anomalies. Synthetic sensor faults are injected into the training data to validate the model. In the test phase, decimated NRT data will be applied to test the actual model performance for NRT condition monitoring. The reconstructed sensor data and the original input query data are compared. The system annotates the sensors with high anomaly score contributions to alert the pilot. At time step j, for the i^{th} sensor's reading, the reconstruction error $\delta_{i,j}$ is the absolute value of the difference between the enquiry data $x_{i,j}$ and the BiGAN reconstruction $\hat{x}_{i,j}$:

$$\delta_{i,j} = |x_{i,j} - \hat{x}_{i,j}| \quad (5)$$

which will enable the visualisation of anomalies on sensor readings.

B. Supervised fault diagnostics

The supervised fault diagnostics system aims to indicate the UGs' operating status when NRT data is transmitted via satellite connection, i.e. whether the vehicles present normal conditions or have developed a certain type of fault. Given a training dataset $\{(x_i, y_i)\}_{i=1}^{N}$ with N training samples such that x_i is the i^{th} data patch ($a \times b$ matrix, where a is the number of signals, b is the number of time steps in one data patch) and y_i is its corresponding label (i.e. fault class), the supervised learning algorithm seeks a function $g : X \rightarrow Y$ that predicts the unknown class $y \in Y$ of an observation $x \in X$, where X and Y are the input and output space of

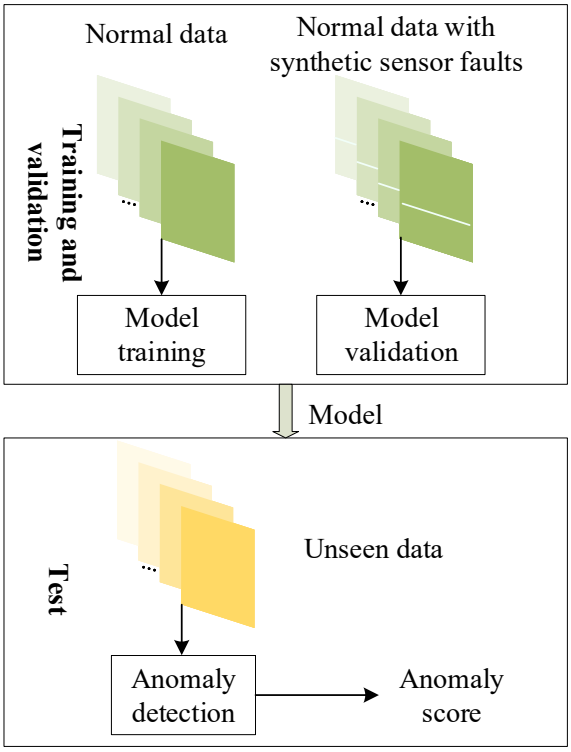

Fig. 4. Workflow of the BiGAN-based anomaly detection for UGs [7].

g, respectively. For training sample (x_i, y_i), a loss function $\mathcal{L}_s(y_i, \hat{y}_i)$ is minimised to find the function g, where $\hat{y}_i$ is the predicted label of x_i.

The function g is modelled with a neural network, as shown in Fig. 5. The original input x ($a \times b$ matrix) is flattened as the input of the neural network. Three fully connected layers (followed by their own batch normalization, LeakyReLU activation and dropout layers) forward propagate the features to a fully connected layer activated by SoftMax to output the predicted class. A cross-entropy loss function is applied to the output of the last fully connected layer.

III. Datasets

Table I details the 11 Slocum G2 deployment datasets used in this study [7], [10]. The first two datasets collected by units 345 and 397 are applied to train the BiGAN-based anomaly detection system as those are normal deployments without any anomalies. The other datasets used for testing include one healthy deployment of unit 419, and eight deployments with anomalies including biofouling, Ocean Microstructure Gliders (OMG, which have a large sensor appendage mounted on the exterior of the hull), angle of list, loss of wings and strong disturbances. The detailed data processing procedures can be found in [7].

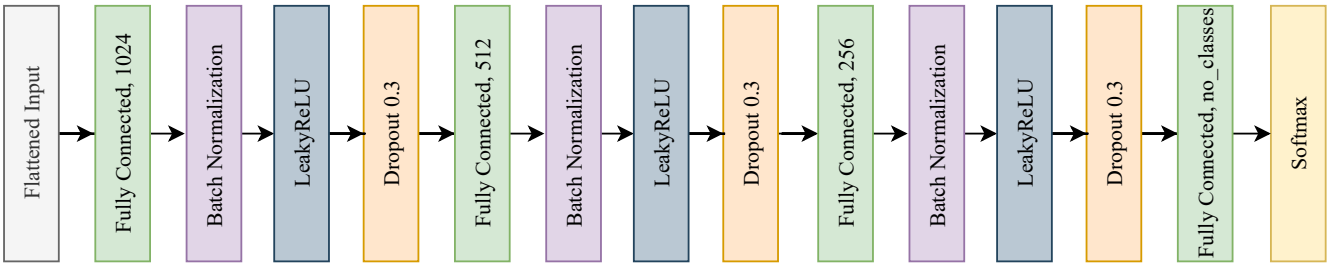

Fig. 5. Neural network configuration for the supervised learning.

TABLE I
THE DATASETS APPLIED FOR ANOMALY DETECTION SYSTEM TRAINING AND TESTING.

Glider ID	Glider status	Anomaly detection
unit 345	Healthy	Training
unit 397	Healthy	Training
unit 419	Healthy	Testing
unit 399	Biofouling	Testing
unit 423	OMG	Testing
unit 424	OMG	Testing
unit 194	Angle of list	Testing
unit 304	Loss of right wing	Testing
unit 345	Strong disturbances	Testing
unit 436	Loss of left wing	Testing
unit 492	Simulated biofouling	Testing

IV. RESULTS

In this section, the anomaly detection results will be presented in comparison with model-based and rule-based results acquired in [6], [11]. The anomaly detection results will then be applied to annotate the dive cycles of the deployments to train a supervised learning model for fault diagnostics. It should be noted that the supervised learning approach is an attempt to classify faults with limited training datasets. The limited dataset size and the bias towards healthy baseline conditions could potentially lead to a biased supervised learning model. To ensure generality, the training data would require numerous additional raw datasets with increased diversity.

A. Anomaly detection

Figure 6 shows a sample of the validation process where only the rudder angle signal is artificially manipulated to its minimum value (-0.52 rad) and the anomaly detection has successfully annotated this anomaly. After learning the distribution $x \sim p(x)$ of the training dataset, the model can output a high anomaly score that describes the degree of an anomaly.

Figure 7 shows the anomaly detection results against the model-based and rule-based ones. It is worth noting that only seven deployments are compared here due to the availability of baseline results for the remaining deployments. The BiGAN-based anomaly detection system has correctly shown anomaly score trends in good agreement with the model-based and rule-based methods for the seven deployments. Note that for some of the deployments (e.g. unit 436: wing loss), only one baseline is presented, since the missing model-based method did not correctly show the detection metric trend.

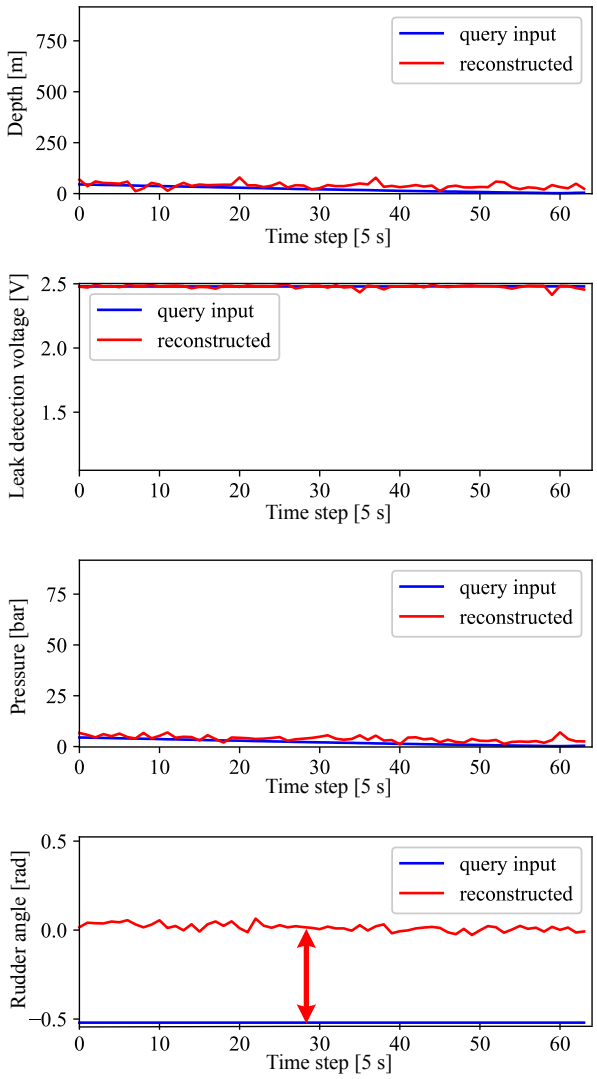

Fig. 6. A verification sample with the rudder angle signal manually set to its minimum value while the other signals are unchanged. For the rudder angle sensor, the signal reconstructed by the BiGAN is distributed around 0 and matches the actual sensor reading, suggesting that the model has learned the distribution of the training data.

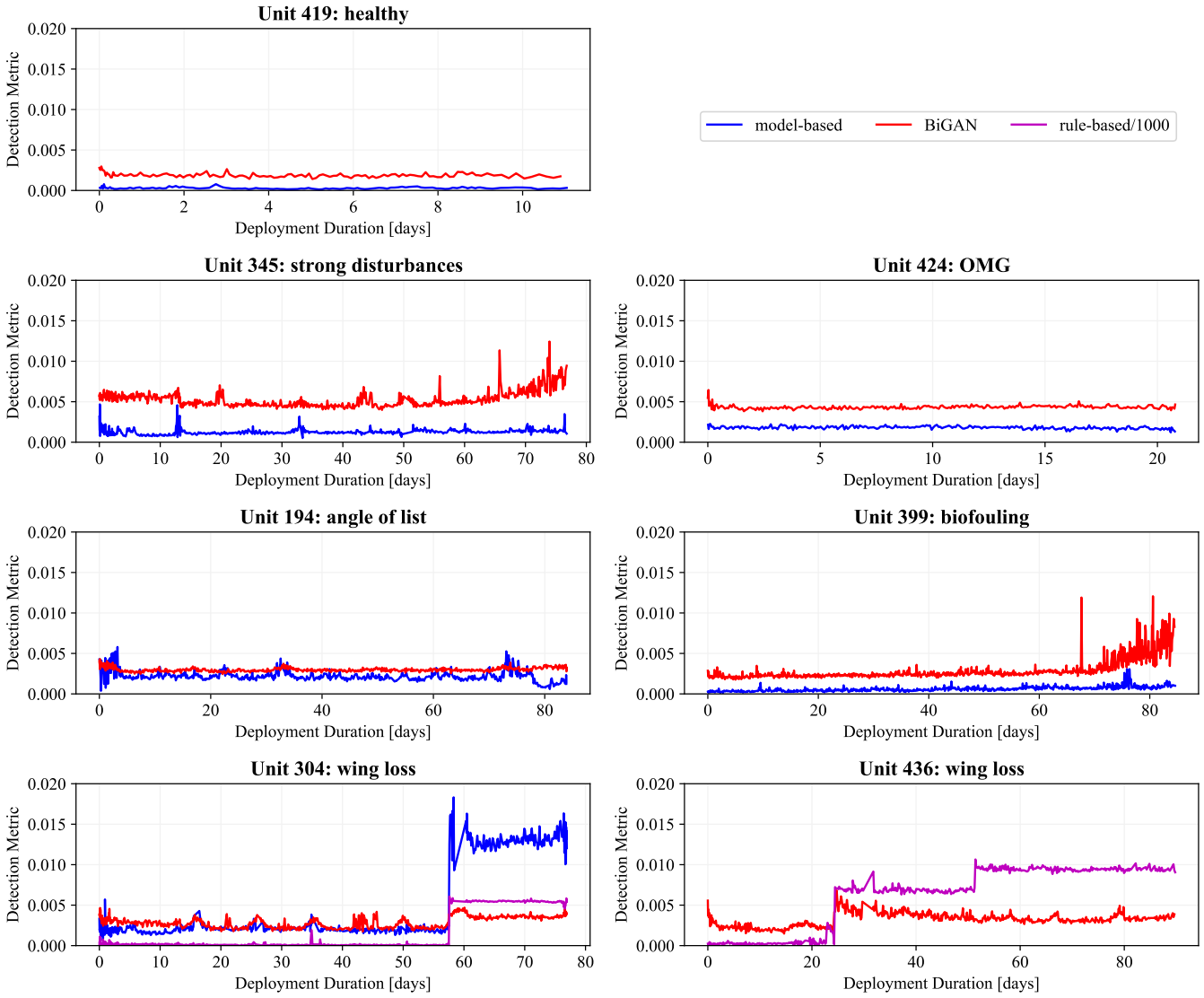

Fig. 7. Anomaly detection results using BiGAN, compared with model-based and rule-based approaches.

B. Fault diagnostics

Table II details the representative cycles selected from the datasets with the aid of the anomaly detection results. The chosen dive cycles are applied to the training, validation and test datasets for the supervised learning method. For each deployment, 50% of the selected cycles are randomly sampled to generate the training datasets, 25% of the cycles are applied to generate the validation dataset, and the remaining 25% of the cycles are applied to generate the test dataset. For each dive cycle, 100 data patches with 64 time steps are randomly sampled for both the training and test datasets. Note that there are 15 signals included in the original datasets. Therefore, the training, validation and test datasets include 179,200, 89,600 and 89,600 15×64 matrices with corresponding labels, respectively.

TABLE II
DATASET ANNOTATION USING UNSUPERVISED ANOMALY DETECTION RESULTS.

Glider ID	Glider status	Class label	Selected cycles
unit 345	Healthy	0	[0, 400]
unit 397	Healthy	0	[0, 400]
unit 419	Healthy	0	[0, 120]
unit 399	Biofouling	1	[800, 1050]
unit 423	OMG	2	[0, 200]
unit 424	OMG	2	[0, 270]
unit 194	Angle of list	3	[0, 700]
unit 304	Loss of right wing	4	[550, 700]
unit 345	Strong disturbances	5	[0, 620]
unit 436	Loss of left wing	6	[250, 650]
unit 492	Biofouling (simulated)	1	[0, 75]

The supervised fault diagnostics model is trained on the training dataset for 10 epochs, with the Adam optimiser and

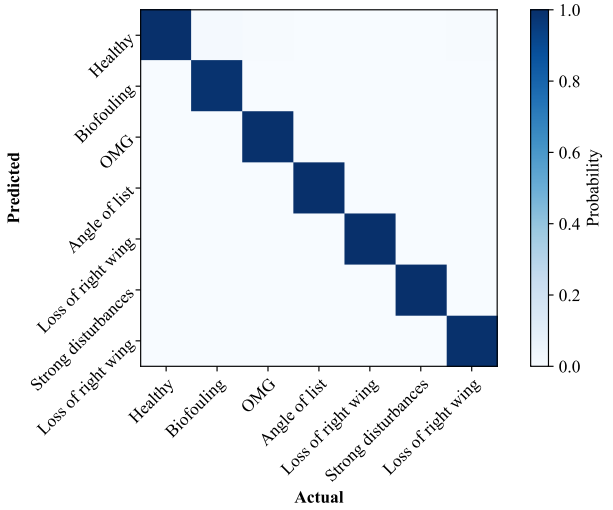

Fig. 8. Confusion matrix of the supervised fault diagnostics results on the test dataset.

a learning rate of 1×10^{-5}. The fault diagnostics accuracy on the validation dataset is 99.76%. Subsequently, the trained neural network is applied to the test dataset to detect and classify the anomalies. The overall accuracy of the model is 99.67% on the test dataset. Figure 8 shows the confusion matrix of the supervised fault diagnostics results on the test dataset, suggesting the model has achieved high fault diagnostics accuracy for the 6 types of faults considered, as well as healthy operating status.

Although high fault diagnostics accuracy has been achieved over the test dataset, the fault diagnostics model is trained with datasets collected from only a few deployments. Training the model with such a small dataset can lead to an overfitted model. An overfitted model could memorise features specific to an individual deployment or a particular vehicle, which could lead to degraded fault diagnostics performance in real applications. Due to limited data available, further investigations were not conducted. In future work, the training dataset will be enriched. In addition, automatic segmentation methods will be developed to annotate the deployment datasets.

V. Conclusions

This work has further extended the BiGAN-based anomaly detection system developed in [7] to assist with the annotation of UG deployment datasets. The performance of the BiGAN-based anomaly detection system has been compared with rule-based and model-based methods. The annotated deployment datasets were applied to train a supervised learning fault diagnostics model. With the limited availability of training samples of the different anomalies, the supervised learning model has achieved an accuracy of 99.67% in fault diagnostics. In further work, supervised and semi-supervised models will be developed based on larger datasets with better diversity.

References

[1] D. L. Rudnick, "Ocean Research Enabled by Underwater Gliders," *Annual Review of Marine Science*, vol. 8, no. 1, pp. 519–541, 2016.

[2] S. Wood, "Autonomous Underwater Gliders," in *Underwater Vehicles*, A. Inzartsev, Ed. IntechOpen, 2009, ch. 26, pp. 499–524.

[3] E. Anderlini, C. A. Harris, G. Salavasidis, A. Lorenzo, A. B. Phillips, and G. Thomas, "Autonomous detection of the loss of a wing for underwater gliders," in *2020 IEEE/OES Autonomous Underwater Vehicles Symposium (AUV)(50043)*. IEEE, 2020, pp. 1–6.

[4] E. Anderlini, G. Thomas, S. C. Woodward, D. A. Real-Arce, T. Morales, C. Barrera, and J. Hernández-Brito, "Identification of the dynamics of biofouled underwater gliders," in *2020 IEEE/OES Autonomous Underwater Vehicles Symposium (AUV)(50043)*. IEEE, 2020, pp. 1–6.

[5] E. Anderlini, D. A. Real-Arce, T. Morales, C. Barrera, J. J. Hernández-Brito, A. B. Phillips, and G. Thomas, "A marine growth detection system for underwater gliders," *IEEE Journal of Oceanic Engineering*, 2021.

[6] E. Anderlini, G. Salavasidis, C. A. Harris, P. Wu, A. Lorenzo, A. B. Phillips, and G. Thomas, "A remote anomaly detection system for slocum underwater gliders," *Ocean Engineering*, vol. 236, p. 109531, 2021.

[7] P. Wu, C. A. Harris, G. Salavasidis, A. Lorenzo-Lopez, I. Kamarudzaman, A. B. Phillips, G. Thomas, and E. Anderlini, "Unsupervised anomaly detection for underwater gliders using generative adversarial networks," *Engineering Applications of Artificial Intelligence*, vol. 104, p. 104379, 2021.

[8] J. Donahue, P. Krähenbühl, and T. Darrell, "Adversarial feature learning," *arXiv preprint arXiv:1605.09782*, 2016.

[9] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[10] BODC, "Glider inventory," 2019. [Online]. Available: https://www.bodc.ac.uk/data/bodc_database/gliders/

[11] Z. Bedjia-Johnson, P. Wu, D. Grande, and E. Anderlini, "Smart anomaly detection for slocum underwater gliders with a variational autoencoder with long short-term memory networks," *Applied Ocean Research*, 2021, submitted.