2.2.4.2. Verification of understanding requirements

Practical guidance – cross-domain and maritime

Authors: ALADDIN demonstrator project

Background

Understanding the state of the RAS and its operating environment is critical for system to operate safely while achieving its intended goals.

Regulations for condition monitoring systems can be found in the ISO standards 13372:2012 [1] and 26262-1:2018 [2] for machines and road vehicles, respectively. The data collection and management specification for automated vehicle trials can be found from [3].

This Body of Knowledge entry describes the verification of understanding requirements defined in ref[1]. The verification is demonstrated by testing the supervised learning and semi-supervised learning fault diagnostics methods developed in the project ALADDIN. In addition to the datasets collected from British Oceanographic Data Centre (BODC) [4], additional data collected in the field tests led by SOCIB and the National Oceanography Centre (NOC) are applied in this entry to test the algorithms' fault diagnostics performance, and to verify the understating requirements.

Stage Input and Output Artefacts

Required input/existing knowledge:
- List of all sensors on the RAS – the list should clearly label any redundant sensors on over-observed systems
- Signal output from all sensors, including both readings and the associated time stamps, synced for the RAS
- Knowledge of the dynamics of the RAS (desirable, especially for under-observed systems)
- Failure Mode Effect Analysis (FMEA) for the RAS (desirable), which can be completed according to the IEC 60812:2018 standards [5]
- Hazard and Operability study (HAZOP) for the RAS (desirable), which can be completed according to the IEC 61882:2016 standards [6]
- Formal description of the baseline RAS behaviour
- Metadata (e.g. calibration, system configuration, deployment reports or operator logs) to enable understanding of wider context
- The deployment data applied in this study needs clear labelling of the faults, ideally with fault starting and ending times indicated
- Training and validation datasets for baseline RAS behaviour – data from multiple RAS covering a wide range of normal missions as well as missions with faults would be ideal.

---

[1] https://www.york.ac.uk/assuring-autonomy/guidance/body-of-knowledge/implementation/2-2/2-2-1/2-2-1-2/

- Testing datasets (including datasets collected previously and the datasets to be collected by the scheduled field tests) for anomalous RAS behaviour – data of a wide range of anomalies would be ideal. Additionally, it is important to log the start and end times of the anomalous behaviours.

Assumptions

- Sensing requirements, as defined in ref[2], are met (i.e. the installed sensors are appropriate for the RAS)
- Understanding requirements, as defined in ref[3], are met
- The training, validation and testing datasets are collected by similar systems
- The incoming volume of data is manageable in real-time with the installed processing power
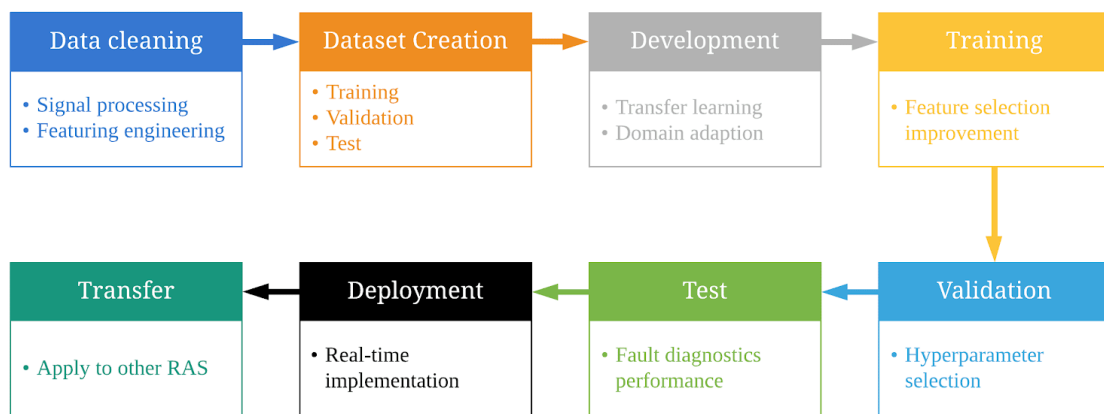
Stage input and output artefacts

Figure 1: Summary procedure to verify understanding requirements

1. **Data cleaning**: the data applied for the training, validation and test datasets, as well as the RAS real-time implementation data need to be cleaned through the following steps
   a. Signal processing and data treatment according to the data collection and management specification for automated vehicle trials [3]
   b. Feature engineering:
      i. Design additional virtual signals derived from dynamic models to better present the RAS's operating status and its environment
      ii. Data fusion to combine the virtual and actual signals
   c. Depending on the type of RAS and characteristics of the data, some application may need filtering of transient effects and measurement noise to reduce the computational cost and improve real-time performance
      i. In the case of Unmanned Aerial Vehicles, reserving full dynamic effects can be important

---

[2] https://www.york.ac.uk/assuring-autonomy/guidance/body-of-knowledge/implementation/2-2/2-2-1/2-2-1-1/cross-domain-maritime/

[3] https://www.york.ac.uk/assuring-autonomy/guidance/body-of-knowledge/implementation/2-2/2-2-1/2-2-1-2/

ii. In the case for underwater gliders, the RAS can spend significant time in steady-state operations. Reducing/removing transient effects can beneficial

2. **Dataset creation**: creation of suitable datasets for the development of the anomaly detection system. Data augmentation (see Section 2.3.1) is required in many applications to achieve desired testing/deployment performance of the model
   a. Training dataset
   b. Validation dataset for hyperparameter selection and training
   c. Test dataset:
      i. Previous deployment data has not been seen by the model during training

3. **Development**: the fault diagnostics method with domain adaption to ensure faults can be inferred by the model in the target domain

4. **Training**: the anomaly detection system is trained using the previously prepared training dataset. Different signals can be selected as input during this stage to assess improvements in performance (prediction accuracy and computational cost)

5. **Validation**: the hyperparameters of the network (e.g. size of the Deep Neural Networks (DNNs)) are selected with a sensitivity analysis to improve prediction accuracy and computational cost and prevent overfitting on the validation dataset

6. **Test**: the ability of the fault diagnostics system in correctly identifying the fault types. See the summary of fault diagnostics performance on the test datasets in Figure 2

7. **Deployment**: The data collected in the FRONTIERS project in Mallorca, Spain, July 2021. (see Results section)

8. **Transfer:** apply the developed fault diagnostics methods to other RAS with metadata indicated in Body of Knowledge Section 2.2.1.2
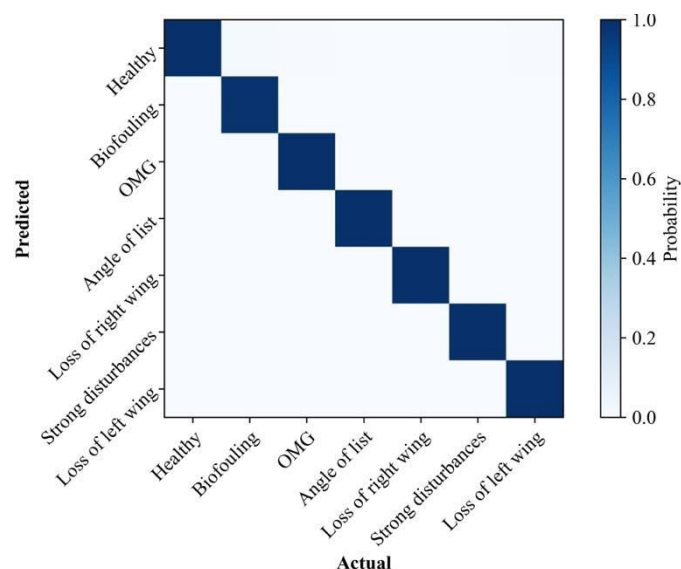

Figure 2: A summary of the fault diagnostics performance on the test datasets.

Method

This work proposes a novel fault diagnostics deep learning model to diagnose faults for Marine Autonomous System via domain adaption and transfer learning (Figure 3). The proposed model (i.e. Marine Autonomous System Net (MASNet)) is applied to address the challenging fault diagnostics tasks for distinct types of underwater gliders that are under-observed and remotely operated in different regions and tasks by different institutions. Based on the improved Bidirectional Generative Adversarial Networks (BiGAN) developed in our previous study [7], we further extract invariant features from both the source and target domains, such that that model can detect unseen faults in the target domains with only a limited number of categories of data for training. The fault diagnostics results are evaluated against rule-based results. The MASNet show effectiveness in generalising invariant features present in the source and target domain datasets collected by distinct underwater gliders operated by different institutes in different regions hence achieving high fault diagnostics performance in the field test.
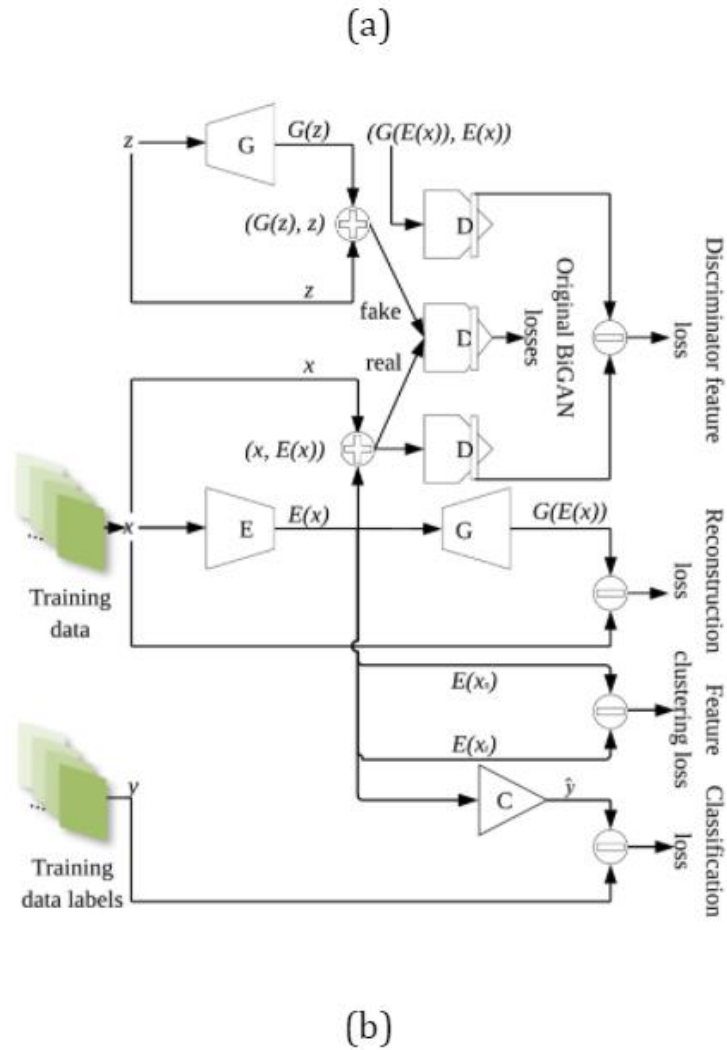

(a)

(a)

(b)
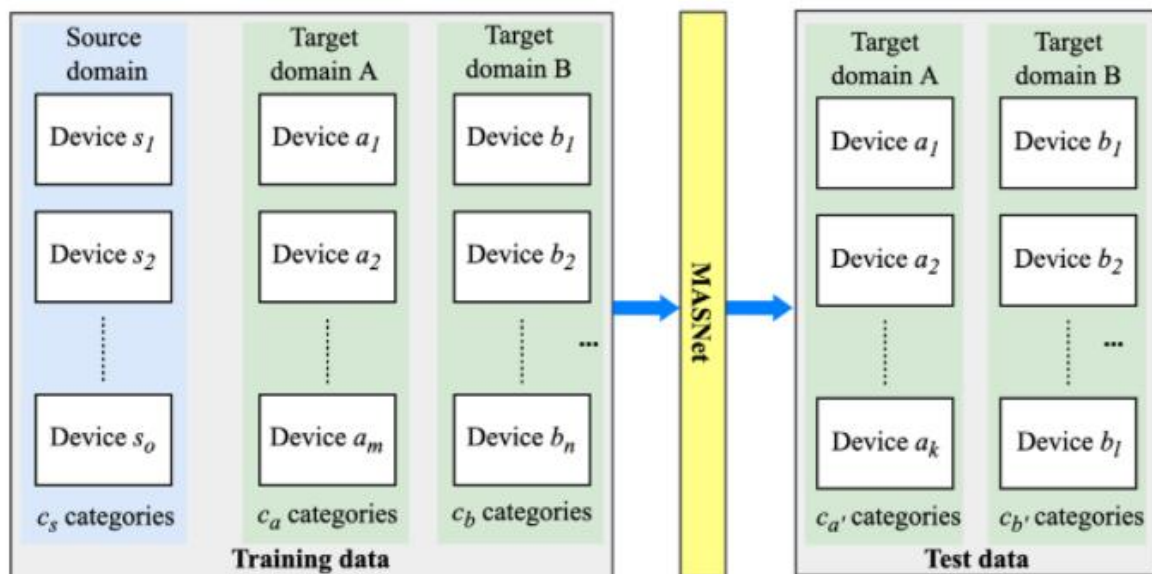
Figure 3: (a) The source domain data is collected by devices labelled as $s_1, s_2, ..., s_o$ in different missions and has c_s different categories. The target domain data is collected by devices of distinct types in different deployments and can be operated by different institutes with distinct settings. The proposed MASNet learns from the source domain and target domain data to classify the categories of the test data which may not be present in the training data. (b) Based on an improved BiGAN-based anomaly detection model proposed in our previous study [7], we add an additional classifier to the model. A feature clustering loss term is added to align the categorical features of the source and target domains, using the encoded latent information from the encoder E.

Results

Figure 4 details the fault diagnostics results of the proposed MASNet applied to the target domain data collected in the FRONTIERS project in collaboration with SOCIB. It is apparent that the MASNet has correctly detected the wing loss scenarios (in comparison with the rule-based method of using the mean roll angle difference between the ascents and descents shown in Figure 3. In the early stage of the test deployment (until mid of 2021-07-07), the MASNet has correctly labelled the glider's status as healthy. Note this data collected over this period has been applied to align the invariant features between the target and source domains. The three scattered points (angle of list) are the times when the pitch angle was manually changed. In the later stage of the test, the MASNet has correctly labelled the

angle of list status (simulated by adjusting balancing pills on the starboard and port sides). The final stage is mostly labelled as either healthy or angle of list, which is mainly due to multiple incorrect trim and ballast settings were applied over this period, making the vehicle's real status has never presented in normal scientific missions.
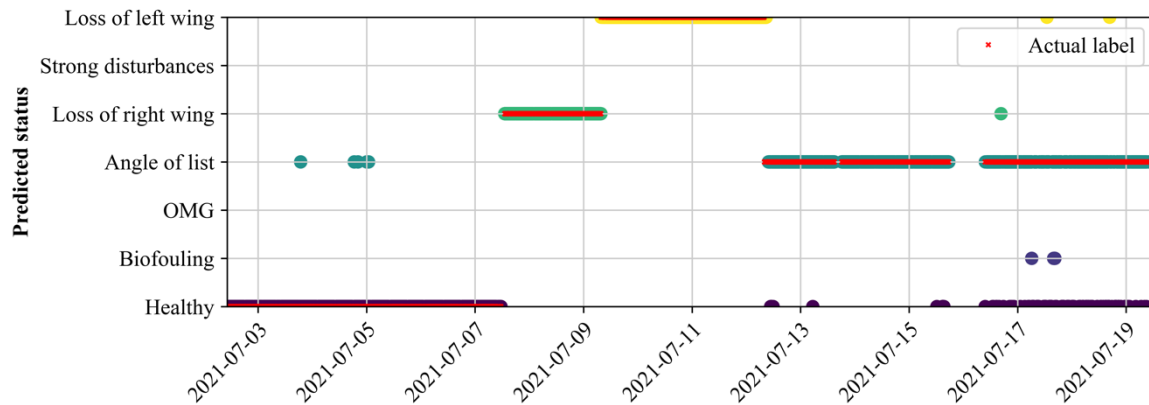

Figure 4: Fault diagnostics results: predicted operating status of the Slocum G2 UG operated by SOCIB in the FRONTIERS project test, in comparison with the actual labels.

Advantages of the approach

- The applied approach is generic and is designed with an expandable architecture and can be extended to any RAS.
- The training data can either be fully labelled or semi-labelled (via approaches such as MixMatch).
- RAS normal variances are allowed.

Limitations of the approach

- The training dataset needs to broadly capture the normal and faulty patterns of the RAS operating state and the operating environment, states that vary from the patterns existing in the training dataset will cause confusion in understanding.
- Fully labelled data may require extensive efforts but could provide better classification performance; semi-labelled data requires fewer labelling efforts but may be less accurate.
- Existing data from the source domain and is sufficient to ensure effective domain adaption.
- Training and validation require significant computational resources and time.
- A pre-trained fault diagnostics model requires the same sensor/feature list in deployment.

References

[1]   ISO, 13372:2012(en): Condition monitoring and diagnostics of machines, 2012.

[2]   ISO, 26262-1:2018(en): Road vehicles — Functional safety, 2018.

[3]   British Standards Institution, PAS 1881:2020 Assuring the safety of automated vehicle trials and testing – Specification, 2021.

[4]   British Oceanographic Data Centre, "Glider inventory," National Oceanography Centre , 2021. [Online]. Available: https://www.bodc.ac.uk/data/bodc_database/gliders/. [Accessed 10 September 2021].

[5]   IEC, 60812:2018: Failure modes and effects analysis (FMEA and FMECA), 2018.

[6]   IEC, 61882:2016: Hazard and operability studies (HAZOP studies) - Application guide, 2016.

[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]   P. Wu, C. A. Harris, G. Salavasidis, I. Kamarudzaman, A. B. Phillips, G. Thomas and E. Anderlini, "Anomaly Detection and Fault Diagnostics for Underwater Gliders Using Deep Learning," in IEEE/OES OCEANS 2021, accepted, San Diego, 2021.