SENSOR DEGRADATION DETECTION UISING VISUAL TIMESERIES AND DEEP CONVOLUTIONAL NEURAL NETWORKS

C. J. Wallace and S. D. J. McArthur
University of Strathclyde
Glasgow, UK
christopher.wallace@strath.ac.uk

ABSTRACT

Effective condition monitoring and prognostic analyses depend on accurate sensor data in order to estimate the health of an asset. Given the nature of many items of legacy equipment in the nuclear industry, this is often not the case, with degradation occurring due to the age and operation of sensors, communication routes as well as the asset itself. Detection of anomalous signals which do not reflect the behaviour of the asset is therefore essential, with the primary objective to repair or replace such components. This paper introduces a method of detection of sensor degradation using timeseries which have been converted to images, in order to leverage the powerful feature detection capabilities of modern deep convolutional neural networks. By converting 1-D time-series to 2-D representations via a Grammian Angular Field and using a small number of training examples, it is possible to train such a network to automatically identify features associated with faults. A case study is presented for a set of sensor types, demonstrating the capability of the model to generalise to previously unseen data from sensors of the similar type and identify faults at greater than 85% accuracy. The results demonstrate the benefits that can be derived from an unsupervised feature detection process for this type of problem and highlights the transferability of models trained on one sensor type and applied to previously unseen similar sensor types.

Key Words: Condition Monitoring, Sensor Degradation, Machine Learning

1. INTRODUCTION

Implementing condition monitoring and prognostic tools in order to manage industrial assets presupposes sufficiently accurate data that emergent faults can be detected but also that spurious sensor signals do not erroneously trigger unnecessary maintenance. The challenge in such implementations is therefore not only in the analysis of signals from the monitored asset itself but from the sensor infrastructure used to enable to monitoring. That is, the health of the sensor matters as much as the health of the asset being monitored, as described in [1]. Typically, in the nuclear industry this challenge is addressed through sensor self-checks or redundant measurements [2].

The maintenance burden caused by malfunctioning sensor infrastructure however can be significant, particularly where failure of a redundant channel can introduce a vulnerability in the control or operation of plant. Detection of such faults can be undertaken manually by a human operator or semi-automatically using software methods to process signals and identify either erratic measurements or signals exceeding pre-determined upper or lower limits. These approaches generally implement a set of statistical rules, either explicitly in software or implicitly through the review of data by an expert. This can be effective at identifying faults that meet particular conditions however the criteria for detection of such faults are relatively limited and often do not enable detection of more subtle faults that do not meet the pre-specified criteria, particularly when: - faults are intermittent, - faults do not evolve monotonically - faults manifest with a range of characteristics.

This is a peer-reviewed, accepted author manuscript of the following article: Wallace, C., & McArthur, S. (2021). Sensor degradation detection using visual timeseries and deep convolutional neural networks. In Proceedings of 12th Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies (NPIC & HMIT 2021) (pp. 304-312). American Nuclear Society. https://doi.org/10.13182/T124-34395

This paper describes a method of identifying degrading sensor contacts using a small number of labelled examples, which are transformed into a visual representation such that they can employ a Convolutional Neural Network (CNN) approach to feature generation and fault detection. The approach presented removes the bias inherent in many existing methods of sensor degradation detection by allowing the trained model to identify features relevant to degradation and separate them from those associated with system noise and operational actions.

2. APPROACH AND METHOD

2.1. 2-D REPRESENTATIONS OF UNIVARIATE TIME-SERIES

Previous work on the transformation of univariate timeseries to 2-dimensional images have demonstrated the powerful capability of such representations to encode the data such that temporal dependence is preserved [3]. A common approach is to employ Grammian Angular Fields (GAF) or Markov Transition Fields (MTF). Figure 1 shows an example of time-series windows transformed into 2-D representations using a Grammian Angular Summed Field (GASF).

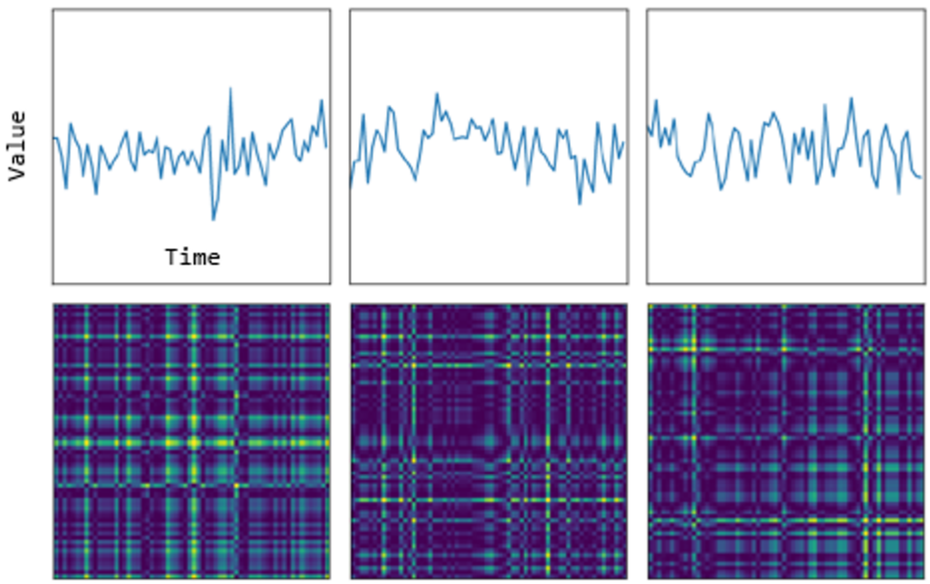

Figure 1. Three examples of time-series windows transformed into 2-D image representations

The process of transforming a univariate time-series by means of a GAF, as described in [3, 4] initially requires encoding the time-series X, with components x_i at time t_i and length N, to polar coordinates:

$$\begin{cases} \phi = arccos(x_i) - 1, -1 \leq x_i \leq 1, x_i \in X \\ r = \frac{t_i}{N}, t_i \in \mathbb{N} \end{cases} \tag{1}$$

before transforming the resulting vector to a GASF as:

$$GASF = [cos(\phi_i + \phi_j)] = X' \cdot X - \sqrt{I - X^2}' \cdot \sqrt{I - X^2} \tag{2}$$

where I is the unit row vector. The resulting GASF, a two-dimensional representation of the original vector X, has dimensions NxN. It is possible to reconstruct the original time-series as the main diagonal of the GASF, with the time index running from the upper left to lower right corner

2.2. CONVOLUTIONAL NEURAL NETWORK MODEL

Convolutional neural networks have been used extensively in recently years for a variety of tasks [5], in particular for image processing problems such as object detection, classification or post-processing. The essential feature of a convolutional neural network, compared to a more traditional Artificial Neuron Network (ANN) is the inclusion of a convolutional layer wherein one or more filters of a fixed size convolves with the feature data.

The results of this convolution allow the filters, of various sizes and configurations, to detect low-level characteristics from images such a colors, intensity or edges. With a large number of convolutional layers and filters of various sizes, it is possible to identify a range of feature types at different scales, which greatly improves the capability of the network to learn structures. The additional information generated by this process is 'pooled' using so called 'pooling layers' in order to condense the learned features into a more concise representation and allow for the model to learn these representations at different positions and rotations within images. Recent work [3, 6] has shown that this capability is not restricted to images of physical objects (e.g. photographs) but can work effectively on artificially generated representations of data [7–9]. These representations can include images generated by a range of means including recurrence plots, signature matrices as well as transforms such a GAF.

2.3. METHODOLOGY

This paper presents a method of combining the techniques described in order to provide improved capability for detection of sensor degradation. The core stages of the method are:

- Segmenting a univariate timeseries into overlapping windows
- Transformation of the labelled timeseries windows into 2-D images
- Training a CNN to accurately classify the images as 'healthy' or 'faulty' based on the labels
- Using the trained model to detect degradation on previously unseen sensor data

This approach is intended to remove the subjectivity and bias often associated with the detection of anomalous signals which lead to a restrictive set of features and allowing the creation of automatically generated features based on the content of the images, potentially improving the generalizability of the model. This process is shown in Figure 2.

3. IMPLEMENTATION CASE STUDY

The method described in the previous section was implemented for a sample set of industrial data derived from eight sensors on a nuclear power station, comprising four temperature sensors and four flow sensors. The contacts which relay the sensor measurements to the central data collection system have a known fault mechanism which occurs intermittently prior to failure and can manifest with a range of characteristics which are often not distinguishable from signal noise or normal operation. The risk of disruption or inadvertent damage to adjacent equipment during inspection precludes regular inspection and maintenance of the contact during normal operation.

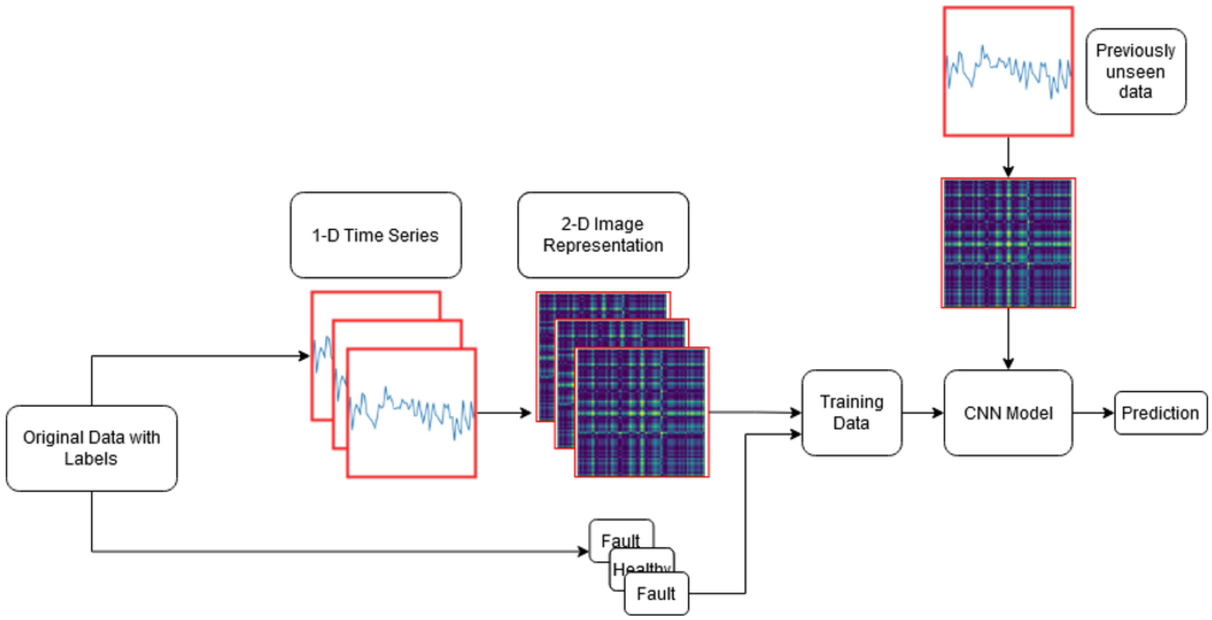

Figure 2. Transforming the existing labelled time-series into an image classification problem

3.1. DATA AND PREPROCESSING

The dataset was selected based on the availability of labelled examples of faults which occurred intermittently for a range of durations, with each sensor experiencing at least one period of behavior which indicated (and were subsequently confirmed) to be due to the known fault mechanism.

For testing purposes, data from one temperature sensor and one flow sensor were isolated and removed from the dataset for final testing. The training and validation data consisted of:

- Approximately three months of data (in total, for all sensors)
- Samples collected at 2-second resolution
- 19 discrete labelled intervals of behavior confirmed by a human expert to be due to degradation (later confirmed by inspection and replacement)

The data was subsequently segmented into 120-second windows, with consecutive windows shifted by one time measurement such that each data-point was captured within up to 120 data windows. This approach allowed for a significant increase in the number of samples and provided a range of contexts for the same degradation instance, for example in some data windows the characteristics of the data that manifest as a result of the degradation with occur at different time indexes and for varying durations. The resulting dataset consisted of 21,335 unique images with corresponding labels.

Figure 3 shows an example of two data windows converted to image representations for a trivial case of degradation which is clearly visible in the original time-series.

3.2. MODEL TRAINING

Using the available data, an analysis pipeline was constructed in Python using the Tensorflow [10] framework to develop the CNN. The data was segmented and transformed as described earlier before being

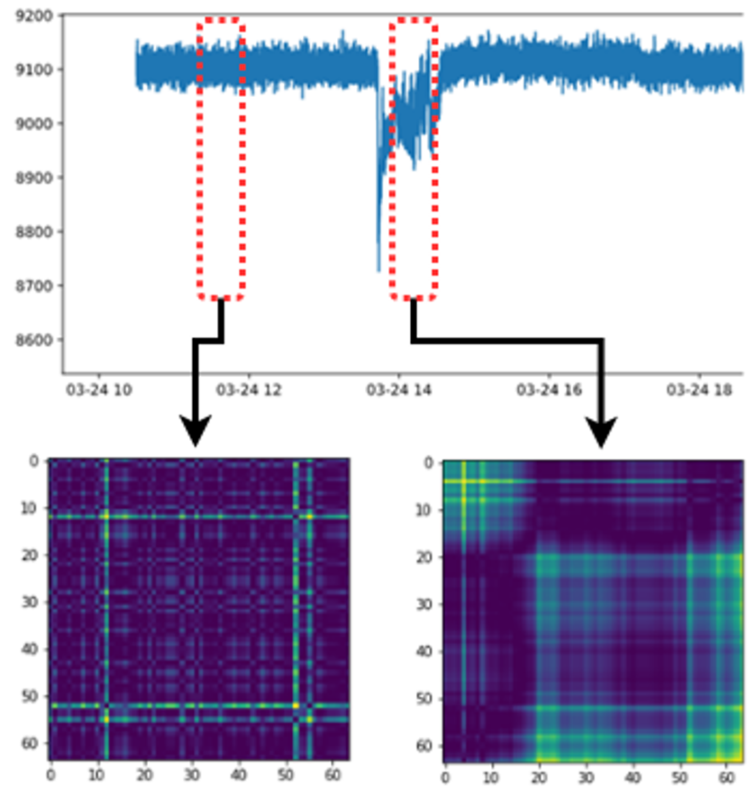

Figure 3. An example of a healthy data window (left) and a data window exhibiting degradation (right). Note the regular, well defined edges of the healthy data compared to the blurred image generated by the degraded data

randomly shuffled and split into a training set comprising 80% of the data and a validation set comprising 20% of the data. The data was then transformed to 2-D time-series images using PyTS [11]. A model structure was developed comprising:

- An input layer (each image)
- Two convolutional layers each connected to a pooling layer
- A dense output layer which compressed the output to a two-dimensional vector

Figure 4 shows a simplified schematic of this arrangement. The final output layer produced two values, corresponding to the probability of degradation being present and the probability of the data being healthy.

The model was trained until the learning rate against the validation data (which is not used from training) reduced to near zero, as shown in Figure 5.

3.3. MODEL TESTING

The trained model was tested against two previously unseen sensor sets with data pre-processed in the same manner as the training data, in order to produce a large number of data windows which were transformed into images before being used as inputs. For each data window, a 1x2 output vector was

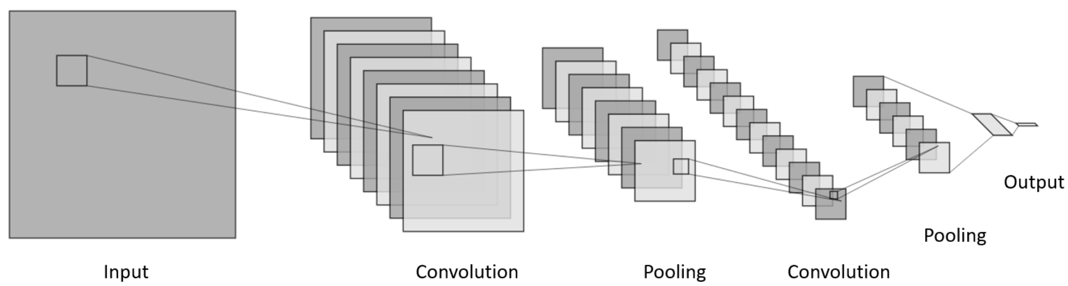

Figure 4. A simplifed view of the convolutional and pooling layers transforming an input "data window" image at the left to an output vector specifying the probability of 'degradation' or 'no degradation'

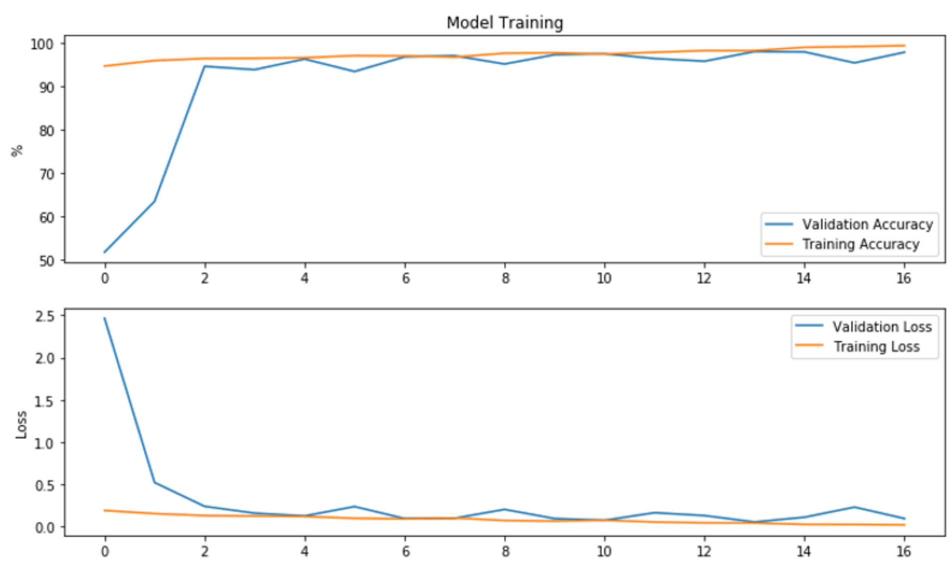

Figure 5. Model accuracy and loss for training (80%) and validation (20%) datasets

generated comprising estimates of 'degradation present' and 'no degradation'. In order to provide context for model performance, the accuracy was compared with previous work on these datasets using a more tradition Support Vector Machine (SVM) approach. This approach employed a set of pre-specified statistical features as inputs (range, skew, variance, mean, kurtosis, entropy) and the relevant labels. Work on this approached was discontinued due to low accuracy and an extremely large number of false positives which made the outputs of limited usefulness. Additionally, a Random Forest (RF) model from the Sci-Kit Learn package [12], using default parameters (100 trees) was used to provide a baseline model, using the statistical parameters generated for the SVM model.

3.4. RESULTS

Comparing the model described in this paper (Image-CNN) with the previous SVM-based approach and the RF model, there is clearly a significant benefit in both accuracy and generalizability as demonstrated by the performance against unseen sensors as shown in Figure 6.

	Model Accuracy		
Dataset	Image-CNN	SVM	Random Forest
Validation Data	**95%**	90%	85%
Unseen Temp Sensor	**87%**	42%	38%
Unseen Flow Sensor	**85%**	35%	34%

Figure 6. Testing results for the validation dataset and two previously unseen sensor datasets

The model described in this paper generalizes well to previously un-seen sensors and while the accuracy is reduced (up to 10%), this level of performance, based on the testing results is sufficient to limit the number of false positives.

Figure 7 shows an example of the model output, with the original input data (containing a recognizable example of degradation) shown in the upper plot while the lower plot shows the estimated probability of degradation at each time.

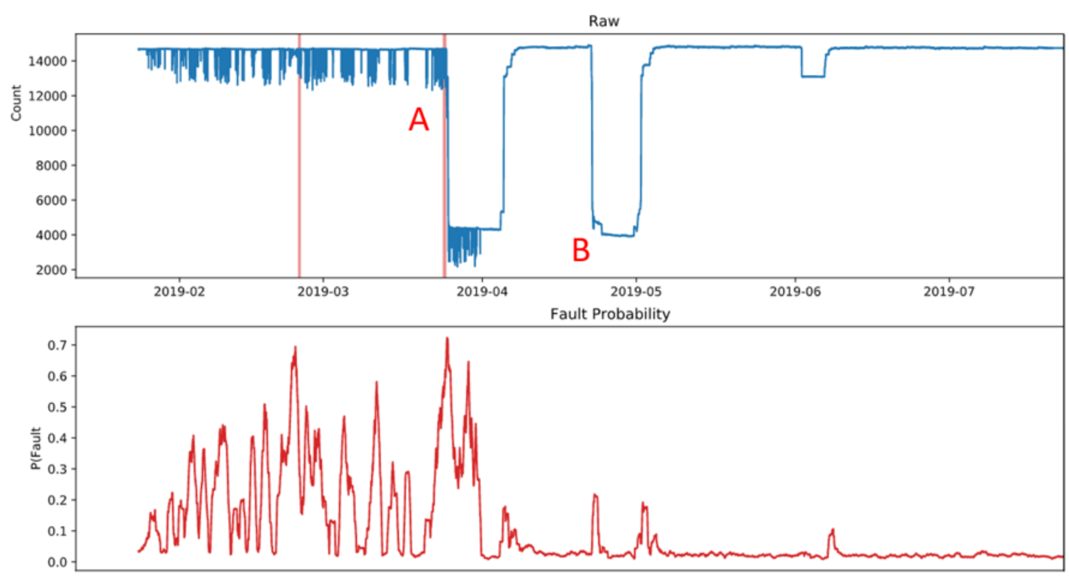

Figure 7. Example of model output, with P(fault) dropping after sensor repair

Note that the time-series in the region prior to the annotation 'A' shows intermittent high variance, which continues when the parameter drops to a new plateau before the variance decreases dramatically and the parameter value returns to its previous level. This corresponds to the repair of the sensor contact during a temporary outage before the system returns to steady state. Another short outage occurs near the annotation 'B', which causes a small increase in the estimate of degradation (a false positive) however this is short-lived and the probability of fault decreases quickly soon after.

The analysis tool has been run on an ad-hoc basis using a graphical interface for explorative analysis but work is underway to deploy an online monitoring system that continuously scans large volumes of data and alerts users to trends in potential degradation. Performance on a modern desktop computer (8 CPU cores with 32GB RAM) allows the analysis of dataset of around 16 million values (around one year of data) in around 90 seconds.

4. CONCLUSIONS

This paper has described a method of detecting sensor degradation using an approach designed to automatically generate detection features by using a small labelled set of examples. The model created was capable of accurately detecting degradation in training and validation data and generalized well to previously un-seen sensors with only a small drop in performance, unlike other models tested which saw significant drops in accuracy.

The performance demonstrated is adequate to ‘down-select’ the possible number of degraded sensor channels to a manageable number for an engineer to schedule maintenance without significant overhead. Testing of further models indicated improved performance through models developed for particular sensor types (e.g., a dedicated ‘flow model’ or a ‘temperature model’). Work is underway to determine the scalability of these models which although valuable in their accuracy, would each require examples of labelled failure data in order to create new models. Additionally, work is underway to determine the feasibility of using other system information or parameter correlations in order to dismiss a large fraction of the false positives and further reduce the operator burden.

ACKNOWLEDGMENTS

This work was funded by Bruce Power.

REFERENCES

[1] B. LIU, P. DO, B. IUNG, and M. XIE, “Stochastic Filtering Approach for Condition-Based Maintenance Considering Sensor Degradation,” *IEEE Transactions on Automation Science and Engineering*, **17**, *1*, 177 (2020).

[2] S. AUTHEN and J. HOLMBERG, “Reliability Analysis of Digital Systems in a Probabilistic Risk Analysis for Nuclear Power Plants,” *Nuclear Engineering & Technology*, **44**, *5*, 471 (2012).

[3] C. ZHANG. et al., “A Deep Neural Network for Unsupervised Anomaly Detection and Diagnosis in Multivariate Time Series Data,” *Proc. The AAAI Conference on Artificial Intelligence*, 2019.

[4] Z. WANG and T. OATES, “Imaging Time-Series to Improve Classification and Imputation,” *Proc. P Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI)*, 2015.

[5] Y. L. CUN, Y. BENGIO, and G. HINTON, “Deep Learning,” *Nature*, **521**, 436 (2015).

[6] I. MITICHE et al., “Imaging Time Series for the Classification of EMI Discharge Sources,” *Sensors*, **18** (2018).

[7] R. ZIMROZ and A. BARTKOWIAK, “Investigation on Spectral Structure of Gearbox Vibration Signals by Principal Component Analysis for Condition Monitoring Purposes,” *Proc. IOP Journal of Physics Conference Series*, 2011.

[8] D. WANG, Y. GUO, X. WU, J. NA, and G. LITAK, “Planetary-Gearbox Fault Classification by Convolutional Neural Network and Recurrence Plot,” *Applied Sciences*, **10**, *3* (2020).

[9] Y. HSUEH, V. R. ITTANGIHALA, H. C. W. WU, and C. KUO, “Condition Monitor System for Rotation Machine by CNN with Recurrence Plot,” *Energies*, **12**, *17* (2019).

[10] TENSORFLOW, *https://www.tensorflow.org/*.

[11] PYTS, *https://pyts.readthedocs.io/en/stable/*.

[12] SCIKIT-LEARN, *https://scikit-learn.org/*.