2023 IEEE Conference on Artificial Intelligence (CAI) | 979-8-3503-3984-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/CAI54212.2023.00104

Multimodal and multi-view predictive maintenance: A case study in the oil industry

Tomás Souper
Intelligent Systems
Fraunhofer AICOS
Porto, Portugal
tomas.pereira@fraunhofer.pt

Duarte Oliper
Intelligent Systems
Fraunhofer AICOS
Porto, Portugal
duarte.pereira@fraunhofer.pt

Vitor Rolla
Intelligent Systems
Fraunhofer AICOS
Porto, Portugal
vitor.rolla@fraunhofer.pt

***Abstract*—This paper explores comprehensive prognostics and health management in the oil and gas industry. It provides a proactive approach to equipment maintenance by detecting potential problems and predicting faults before they occur. Distillers are crucial oil and gas industry components, requiring regular maintenance and monitoring to maintain optimal performance and prevent unplanned maintenance. Deep learning models have shown promising results for predictive maintenance, and multimodality can bring generalization capabilities to these models. This study proposes a multimodal (and multi-view) approach for predictive maintenance in an oil and gas industry distiller dataset. The goal is to demonstrate that this approach can achieve more liable and generalizable models for predictive maintenance than state-of-the-art neural networks when training on a medium-scale dataset.**

***Index Terms*—Predictive Maintenance, Multimodality, Multi-view, Industrial Applications**

I. Introduction

Comprehensive Prognostics and Health Management (CPHM) gained immense popularity in the industrial sector, particularly in the oil and gas industry. By monitoring the equipment's operating conditions, CPHM can detect potential problems and provide early warning signals to prevent unplanned maintenance events. As a result, the oil and gas industry has increasingly adopted CPHM to reduce costs, improve reliability, and enhance safety in its operations.

A distiller is a crucial component in the oil and gas industry, used to separate crude oil into various products such as gasoline, diesel, and jet fuel. The role of a distiller is to heat and vaporize crude oil and then condense the vapor back into liquids. Distillers are subjected to extreme operating conditions, such as high temperatures and pressures, making them susceptible to wear and tear. Therefore, distillers require regular maintenance, monitoring, and inspection to maintain optimal performance and prevent unplanned maintenance. Predictive maintenance techniques, such as vibration analysis and thermography, can help identify potential faults in a distiller. Additionally, monitoring equipment health data, such as temperature, pressure, and flow rates, can provide early warning signals of any issues, allowing for proactive maintenance to avoid costly and time-consuming downtime. Overall, monitoring the health of distillers is crucial for the safe and efficient operation of oil and gas companies.

Predictive Maintenance (PdM) techniques are designed to identify potential failures in a piece of equipment by analyzing monitoring data, aiming for early detection and issue resolution. PdM can reduce material and downtime costs and optimize processes [1]. It differs from reactive and preventive maintenance strategies, as it involves recognizing and addressing potential failures before their occurrence rather than relying on waiting for a failure to occur or following a fixed maintenance schedule.

Modern PdM tasks often take advantage of deep learning models such as Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) neural networks in sensor data for the prediction of asset malfunction in advance, as well as regression of the Remaining Useful Life (RUL) of the equipment [2]. With advances in deep learning, researchers have developed models that can effectively handle various modalities and learn from them simultaneously. Multimodality can allow models to perform better and provide more generalized results than those relying on a single modality. Multimodality is frequently mentioned interchangeably with multi-view approaches, which consist in translating the same data as different data modalities/views for the model to interpret. This results in improved generalization [3]. With this translation in mind, Gramian Angular Field (GAF) is a promising technique for image extraction from time series [4]. GAFs derived from sensor data have served as input to Convolutional Neural Network (CNN) algorithms - such as RESNET - to improve the performance of deep learning algorithm tasks [5].

This study proposes a multimodal and multi-view approach for predictive maintenance in an oil and gas industry distiller dataset introduced in a previous study [6]. The objective is to prove that this approach can attain more accurate and generalizable models for PdM than state-of-the-art neural networks when training on a medium-scale dataset. LSTM will serve as the benchmark.

II. Methodology

The distiller dataset includes multivariate time series data of monitoring sensors with an hourly resolution and operator

Identify applicable funding agency here. If none, delete this.

979-8-3503-3984-0/23/$31.00 ©2023 IEEE
DOI 10.1109/CAI54212.2023.00104

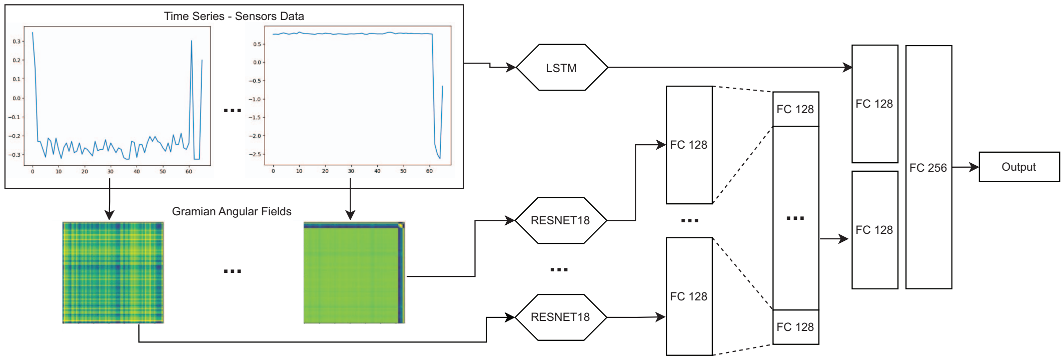

Fig. 1. Multimodal architecture. On the left side, we have a plot of the original time series data and their representation in the image space using the GAF algorithm. Each data modality is inputted into different models. The LSTM model processes the time series data of multiple sensors. Each sensor's image data undergoes a single pre-trained RESNET-18 model, with fine-tuning applied to the final layer. The embeddings extracted from each image are combined and subsequently reduced in size before concatenating with those obtained from the LSTM model. In this way, we can represent the data in a multimodal fashion.

Models	Random Forest			LSTM			RESNET			Multimodality		
(Window, Overlap)	(6,1)	(30,3)	(60,6)	(6,1)	(30,3)	(60,6)	(6,1)	(30,3)	(60,6)	(6,1)	(30,3)	(60,6)
Failures					**F1-Score**							
6	40	62	**86**	66	72	61	57	76	62	47	64	54
10	76	**82**	81	63	61	63	60	55	49	63	60	63
15	44	56	59	55	78	77	49	71	68	45	78	**80**
20	65	65	58	62	69	**70**	58	59	64	59	66	67
					Precision							
6	54	85	98	63	93	96	52	73	58	51	79	**100**
10	68	**77**	72	54	54	60	55	54	55	55	55	60
15	42	52	51	47	64	62	49	**72**	65	43	66	67
20	**67**	65	54	55	56	60	54	57	57	51	57	60

TABLE I
F1-SCORE AND PRECISION (%)

annotations of maintenance periods with a daily resolution. For PdM, twenty failures are selected with at least a ten-day functioning period without failures prior. The binary label for classification was defined through a five-day threshold, following domain experts' recommendations.

We conducted experiments across three distinct subsets, each of which was characterized by a varying number of failures - 6, 10, 15, and 20. For each of these subsets, the data was chronologically partitioned into train and test sets, with a ratio of 75/25. To allow for modeling and GAF image extraction, the time series data and labels were segmented into windows of 6, 30, and 60 hours - with an overlap of 1, 3, and 6 hours, respectively.

In modern PdM approaches, Random Forest and LSTMs use a single data modality - sensor-value data. One of the contributions of this work is to introduce multi-view to tackle the PdM problem, by transforming this time series data into an image representation, then using computer vision models to leverage this representation. Figure 1 shows an overview of the proposed method. To analyze the performance of the proposed method, we compare four different models: a standalone Random Forest model as a baseline, a standalone LSTM model, standalone RESNET models, and the Multimodality approach (both the LSTM and the RESNETs). A grid search was performed to optimize some of the hyperparameters: the size of the embedding to create the joint representation, the number of hidden layers in the LSTM, and its hidden size.

III. Results and Insights

In Table I, the recorded metrics for each experiment include the F1-score and precision. The F1-score serves as a means of assessing the models' overall performance, encompassing

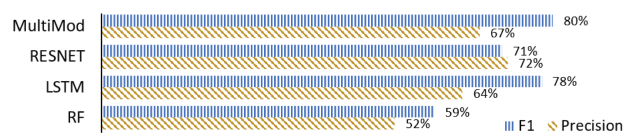

Fig. 2. Best results for fifteen failures regardless of window sequence and overlapping.

both positive and negative classes. In contrast, precision plays a critical role in meeting specific client requirements, as it aids in minimizing false positives alarms.

As introduced by [6], this dataset has different origins of failure. Therefore, generalization is needed when predicting these different types of failures. LSTM is better when more data is available, with fifteen and twenty failures. RF is better with fewer data but prone to overfitting, especially with only six and ten failures. At fifteen failures, multimodality shows the best tradeoffs between generalization and performance, see Figure 2. In this scenario, we used images to capture a particular behavior that brought the best forecasts together with the time series.

References

[1] M. Pech, J. Vrchota, and J. Bednář, "Predictive Maintenance and Intelligent Sensors in Smart Factory: Review," *Sensors*, vol. 21, no. 4, p. 1470, Jan. 2021, number: 4 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/1424-8220/21/4/1470

[2] O. Serradilla, E. Zugasti, and U. Zurutuza, "Deep learning models for predictive maintenance: a survey, comparison, challenges and prospect," Oct. 2020, arXiv:2010.03207 [cs]. [Online]. Available: http://arxiv.org/abs/2010.03207

[3] J. Zhao, X. Xie, X. Xu, and S. Sun, "Multi-view learning overview: Recent progress and new challenges," *Information Fusion*, vol. 38, pp. 43–54, Nov. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253516302032

[4] Z. Wang and T. Oates, "Imaging Time-Series to Improve Classification and Imputation," May 2015, arXiv:1506.00327 [cs, stat]. [Online]. Available: http://arxiv.org/abs/1506.00327

[5] P. Pfeiffer, J. Lahann, and P. Fettke, "Multivariate Business Process Representation Learning utilizing Gramian Angular Fields and Convolutional Neural Networks," Jun. 2021, arXiv:2106.08027 [cs, stat]. [Online]. Available: http://arxiv.org/abs/2106.08027

[6] D. Oliper, V. Rolla, and T. Souper, "Predictive maintenance approaches with free-text labels: A case study in the oil industry," in *Artificial Intelligence and Business Process Management*. AAAI, 2023, misc. [Online]. Available: https://ai4bpm.com/posters/AI4BPM_paper_3073.pdf