Unsupervised Fault Detection in Sensor Networks for Extraterrestrial Habitats

ZIXIN WANG, MOHAMMAD R. JAHANSHAHI, MOHSEN AZIMI and SHIRLEY J. DYKE

ABSTRACT

Various sensors, such as accelerometers, temperature sensors, and pressure sensors, are employed in deep space habitats to monitor operating conditions. The health management system (HMS) depends on accurate sensor data to make reliable decisions. However, the harsh environment of space, characterized by micrometeorite impacts, fire, radiation, dust, and other extreme conditions, can lead to sensor malfunctions. Consequently, the collected data may be corrupted by anomalies, significantly impairing the performance of the HMS. Detecting sensor anomalies and excluding faulty data from decision-making processes is therefore essential. Unsupervised learning has been widely adopted for anomaly detection, as it can train models using unlabeled datasets. This paper presents a novel unsupervised learning approach based on convolutional autoencoders (CAEs) to detect faults in temperature and pressure sensors within the habitat. The proposed method is thoroughly evaluated using a habitat simulator (HabSim), and its capabilities and limitations are discussed.

INTRODUCTION

To effectively monitor and manage the health state of a deep-space habitat, a variety of sensors, such as accelerometers, temperature sensors, and pressure sensors, are essential. However, the harsh conditions in extraterrestrial environments, such as radiation, dust, seismic activity, extreme temperatures, and micrometeorite impacts, can impair or

Zixin Wang, Ph.D. Candidate, Email: wang4591@purdue.edu. Lyles School of Civil and Construction Engineering, Purdue University, West Lafayette, IN 47906, USA
Mohammad R. Jahanshahi, Associate Professor, Email: jahansha@purdue.edu. Lyles School of Civil and Construction Engineering, Elmore Family School of Electrical and Computer Engineering (Courtesy), Purdue University, West Lafayette, IN 47906, USA
Mohsen Azimi, Assistant Professor, Email: azimi@me.msstate.edu. Michael W. Hall School of Mechanical Engineering, Mississippi State University, Mississippi State, MS 39762, USA
Shirley J. Dyke, Professor, Email: sdyke@purdue.edu. School of Mechanical Engineering, Lyles School of Civil and Construction Engineering, Purdue University, West Lafayette, IN 47906, USA

damage these sensors. It is important to note that high-quality sensor data is critical for the HMS to perform accurate fault detection, diagnosis, and robust decision-making.

Methods for sensor fault detection can be broadly categorized into model-based and data-driven approaches. Model-based methods require the manual development of mathematical models that describe the system, which is often time-consuming and demands extensive domain knowledge [1]. These models are used to predict expected sensor measurements, and the discrepancy between actual and predicted values, known as the residual, serves as a fault indicator for detecting sensor anomalies [2]. In contrast, data-driven methods do not rely on physics-based models; instead, they learn system behavior and capture intrinsic relationships among variables directly from measurement data. Data-driven approaches are typically divided into supervised and unsupervised learning. Supervised learning requires labeled data from both normal and faulty conditions for training. Unsupervised learning, trained exclusively on normal-state sensor data, has been widely adopted for sensor fault detection.

In this study, a novel sensor fault detection method for extraterrestrial habitats is developed using CAEs. The proposed approach is employed to detect faults in temperature and pressure sensors and is systematically evaluated using HabSim. Several illustrative examples are presented, including: (1) the effect of the temperature-pressure coupling on the performance of CAEs; and (2) the capability of CAEs to change the number of sensors without requiring network re-design or re-training.

METHODOLOGY

CNNs specialize in processing data that has a known grid-like topology [3]. Autoencoder is a type of unsupervised neural network that is utilized for data compression, dimensionality reduction, and feature extraction [4]. If autoencoders use convolution operation in at least one of their layers, they are referred to as CAEs. In this study, fully convolutional layers are utilized in CAEs, as shown in Figure 1, which is different from auto-associative neural networks (AANNs) that contain fully connected layers. Convolution, also known as cross-correlation, is an operation on two functions (i.e., the input and kernel). Suppose the input and kernel are two-dimensional, the output of the convolutional layer, which is referred to as the feature map, can be obtained by Eq. (1):

$$F(p,q) = (K * I)(p,q) = \sum_{a} \sum_{b} I(p+a, q+b) K(a,b) \tag{1}$$

where I and F are the input and output of the convolutional layer, respectively. K is the kernel that slides across the input to extract features. a and b are the row and column indices in the kernel K. p and q are the row and column indices in the output F. Convolution provides a means for working with inputs of variable size [3]. The weight-sharing and local receptive field strategies in CNN make the model less prone to overfitting [5]. The size of the output of the convolutional layer can be determined by the size of the input, kernel, padding, and stride using Eq. (2):

$$F_s = \frac{I_s - K_s + 2P_s}{S_s} + 1 \tag{2}$$

where F_s, I_s, K_s, P_s, S_s denote the size of the output, input, kernel, padding, and stride in the corresponding dimension (i.e., height or width), respectively.

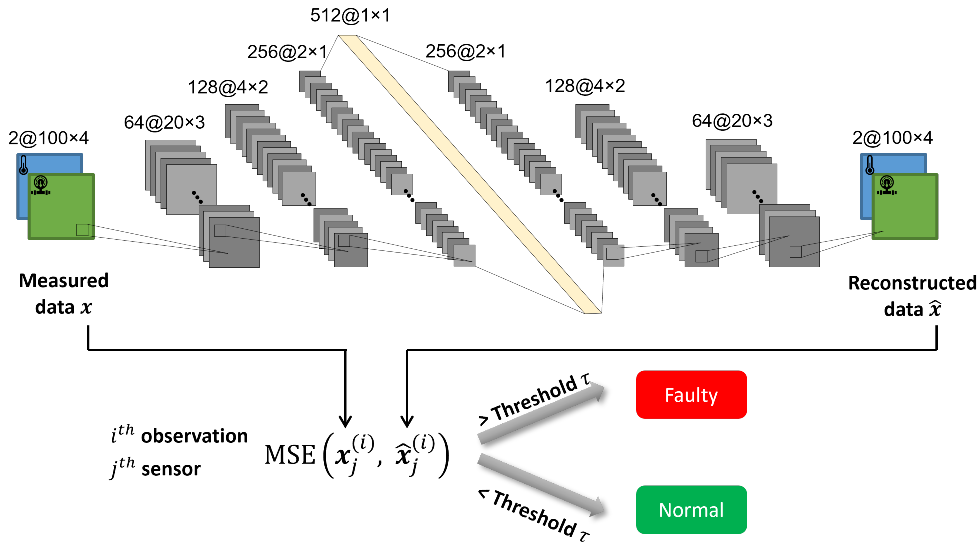

Figure 1. Schematic diagram of CAE architecture. The layer labels represent 'the number of channels @ the size of each channel.'

CAE consists of an encoder and a decoder. The encoder compresses the input data to extract high-level features. The decoder reconstructs the input data using the latent representation generated by the encoder. As shown in Figure 1, the height and width of the feature map decrease in the encoder, while the depth increases. To this end, the size of the kernel and stride decreases in the encoder, while the number of kernels increases. The decoder has a symmetric architecture of the encoder. The input of CAE is the measured data and the output is the reconstructed data. The CAE is trained on normal-state sensor data to minimize the mean squared error (MSE) between its input and reconstructed output. A faulty sensor can be detected if the MSE between the measured data and the reconstructed data is larger than a predefined threshold τ (Figure 1).

ILLUSTRATIVE EXAMPLES USING THE HABITAT SIMULATOR

HabSim is a plug-and-play environment developed in MATLAB Simulink to simulate a space habitat [6]. HabSim can simulate a space habitat in the nominal state, as well as hazardous states caused by physics-based disruption scenarios (e.g., micrometeorite impact) and phenomenological disruption scenarios (e.g., fire). In this study, we investigate the fault detection of temperature and pressure sensors in the interior environment by analyzing the data generated by HabSim in both the nominal state and hazardous states caused by micrometeorite impact and fire scenarios.

Temperature and pressure interact with each other in the interior environment, which can be simulated in HabSim. For example, in the nominal state, both temperature and pressure remain in a small region. Moreover, both temperature and pressure decrease after the micrometeorite impact happens. The drop speeds of temperature and pressure are correlated, and they are both proportional to the size of the hole radius caused by the micrometeorite strike. Additionally, both temperature and pressure increase after the fire starts. The rise speeds of temperature and pressure are also correlated, and they are both proportional to fire intensity levels and spread rates. To investigate the effect of the coupling between temperature and pressure on the detection performance of CAE, the CAE models with the same architecture are trained and tested using the temperature-only data, the pressure-only data, and the combined temperature and pressure data, respectively.

The results obtained by the CAEs with and without the temperature-pressure coupling in different habitat scenarios are summarized in Table I. In the nominal state, CAE can achieve an F1 score of 1.00 either with or without the coupling. However, in micrometeorite impact and fire scenarios, the CAE with the coupling can produce higher F1 scores, as the coupling enhances the redundant information among sensors. It should be noted that even with the coupling, CAE performs better on the temperature data in micrometeorite impact scenarios, while it performs better on the pressure data in fire scenarios. This is because temperature and pressure have smaller decrease and increase rates in micrometeorite impact and fire scenarios, respectively.

TABLE I. Results obtained by the CAEs with and without the temperature-pressure coupling.

			True positive	True negative	False positive	False negative	Precision	Recall	F1 score
Nominal state	With coupling	Temperature	219	2421	0	0	1.00	1.00	1.00
		Pressure	222	2418	0	0	1.00	1.00	1.00
	Without coupling	Temperature	219	2421	0	0	1.00	1.00	1.00
		Pressure	222	2418	0	0	1.00	1.00	1.00
Meteorite impact	With coupling	Temperature	160	2416	21	43	0.88	0.79	0.83
		Pressure	103	2314	84	139	0.55	0.43	0.48
	Without coupling	Temperature	126	2412	25	77	0.83	0.62	0.71
		Pressure	92	2315	83	150	0.53	0.38	0.44
Fire	With coupling	Temperature	146	2354	71	69	0.67	0.68	0.68
		Pressure	213	2385	20	22	0.91	0.91	0.91
	Without coupling	Temperature	139	2345	80	76	0.63	0.65	0.64
		Pressure	179	2351	54	56	0.77	0.76	0.76

At the early stage of the commissioning of the habitat, the CAE models are trained using a default sensor configuration. For example, there are four temperature sensors and four pressure sensors in the interior environment. But in the future, additional sensors may be added to the habitat, or damaged sensors may be removed from the habitat. In these cases, the pre-trained data-driven models may not work under new sensor configurations. For instance, AANN utilizes fully connected layers in which every neuron in one layer is connected with every neuron in its adjacent layers using independent weights, which makes the trained AANN not compatible with different sizes of input. However, since CAE employs convolution operation by sliding a kernel with shared weights across the input, it can be adaptable to different sizes of the input. Therefore, unlike AANN, the CAE trained using the default sensor configuration can also be applied to new sensor configurations without re-designing the network architecture and re-training the neural network. In other words, CAE can continuously work regardless of removing or adding sensors, which is conducive to the resilient operation of space habitats.

The detection results under different sensor configurations, including the default sensor configuration with four temperature and four pressure sensors, as well as new sensor configurations with temperature sensor T2 removed or temperature sensor T5 added, are summarized in Table II. It can be observed that in the nominal state, the trained CAEs can be perfectly applied to new sensor configurations with an F1 score of 1.0. However, in micrometeorite impact scenarios, the performance of the CAE when applied to new sensor configurations is inferior to that when applied to the default sensor configuration. Moreover, the CAE trained with fire scenarios has better performance than the one trained with micrometeorite impact scenarios when applied to either default or new sensor configurations. This can be explained by the simpler data variation (i.e., monotonous increase) of the temperature and pressure in fire scenarios.

TABLE II. Results under different sensor configurations.

	Sensor configuration	True positive	True negative	False positive	False negative	Precision	Recall	F1 score
Nominal state	Default	441	4839	0	0	1.00	1.00	1.00
	Remove T2	384	4236	0	0	1.00	1.00	1.00
	Add T5	401	5539	0	0	1.00	1.00	1.00
Micrometeorite impact	Default	263	4730	105	182	0.71	0.59	0.65
	Remove T2	220	4086	133	181	0.62	0.55	0.58
	Add T5	182	5397	154	207	0.54	0.47	0.50
Fire	Default	359	4739	91	91	0.80	0.80	0.80
	Remove T2	332	4137	78	73	0.81	0.82	0.81
	Add T5	303	5469	71	97	0.81	0.76	0.78

CONCLUDING REMARKS

In this paper, an unsupervised learning approach based on CAEs is developed to detect sensor faults in smart extraterrestrial habitats. Several illustrative examples are demonstrated in different habitat scenarios simulated by HabSim, including the nominal state of the habitat, micrometeorite impact, and fire scenarios. The performance of CAE is systematically evaluated using HabSim. Based on the illustrative example results, the following conclusions can be drawn:

- Utilizing the temperature-pressure coupling can improve the detection performance of CAE, as the coupling enhances the redundant information among sensors. However, even with the coupling, CAE perform better on the temperature data in micrometeorite impact scenarios, while they perform better on the pressure data in fire scenarios. This is because temperature and pressure have smaller decrease and increase rates in micrometeorite impact and fire scenarios, respectively.

- CAE has the capability of being applied to new sensor configurations without re-designing the network architecture and re-training the neural network. In the nominal state, CAE can achieve an F1 score of 1.0 when either removing sensors or adding sensors. The CAE performs better in fire scenarios than in micrometeorite scenarios when applied to new sensor configurations, primarily due to the simpler data variations in fire events, characterized by a monotonous increase in temperature and pressure.

ACKNOWLEDGMENT

This research was funded by the National Aeronautics and Space Administration under award number 80NSSC19K1076.

REFERENCES

1. Spirkovska, L., D. Iverson, D. Hall, W. Taylor, A. Patterson-Hine, B. Brown, B. Ferrell, and R. Waterman. 2010. "Anomaly Detection for Next-Generation Space Launch Ground Operations," in *SpaceOps 2010 Conference*, AIAA, pp. 1–10, doi:10.2514/6.2010-2182.
2. Du, Z., X. Jin, and Y. Yang. 2009. "Fault diagnosis for temperature, flow rate and pressure sensors in VAV systems using wavelet neural network," *Applied Energy*, 86(9):1624–1631, ISSN 0306-2619, doi:10.1016/j.apenergy.2009.01.015.
3. Goodfellow, I., Y. Bengio, and A. Courville. 2016. *Deep Learning*, MIT Press.
4. Bank, D., N. Koenigstein, and R. Giryes. 2020. "Autoencoders," *arXiv e-prints*:arXiv:2003.05991, doi:10.48550/arXiv.2003.05991.
5. LeCun, Y., Y. Bengio, and G. Hinton. 2015. "Deep learning," *Nature*, 521(7553):436–444, ISSN 1476-4687, doi:10.1038/nature14539.
6. Azimi, M., A. Lund, Y. Fu, H. Montoya, L. Vaccino, R. Murali Krishnan, S. Rhee, L. Chebbo, A. Shahriar, Z. Wang, A. Maghareh, and S. J. Dyke. 2025. "HabSim: A Modular-Coupled Virtual Testbed for Simulating Extraterrestrial Habitat Systems," *AIAA Journal*, 63(2):376–388, doi:10.2514/1.J063845.