Proposal of a Time Series Anomaly Detection Method Using Image Encoding Techniques

Ryo Sakurai[1], Takehisa Yairi[2]

[1,2] *Department of Aeronautics and Astronautics. University of Tokyo, Bunkyo-ku, Tokyo-to, 113-8654, Japan*
potion.ha.mazui@gmail.com

Abstract

Time series anomaly detection is considered to play a major role in many areas of society. Models using RNN, which are well suited for time series data, have been studied. However, models using RNN have the problem of high cost; image encoding approaches combining Gramian Angular Fields(GAF) and Autoencoder are less expensive than RNN. However, the accuracy in existing studies is not as good as RNN. In this paper, we propose a time-series anomaly detection framework that first focuses on the structural issues of GAF and the reconstruction accuracy of Autoencoder. Experiments were conducted to verify the effectiveness of the framework. The results showed that the approach focusing on the structural issues of GAF achieved a significant improvement in accuracy, while the approach focusing on improving the reconstruction accuracy of the Autoencoder network decreased the anomaly detection accuracy. The reason for the lower accuracy was found to be that the networks with higher reconstruction accuracy accurately reconstructed even the anomaly images, making anomaly detection based on L1 errors impossible. These results indicate that an approach that focuses on the structural problems of GAF is effective, while an approach that improves the reconstruction accuracy of Autoencoder is not necessarily effective.

1. Introduction

Time-series anomaly detection contributes in various ways to many applications, such as monitoring system failures, detecting external cyber attacks, and diagnosing diseases. Unsupervised learning has been studied in the field of time series anomaly detection, especially in combination with RNNs (Recurrent Neural Networks), including LSTM (Long Short Term Memory) (Graves, 2012), which is effective for time series analysis (Mogren, 2016). However, methods that combine RNNs and generative models are difficult to speed up because the RNN part cannot process input data in parallel, and the large number of memory units increases both memory usage and computational cost.

In contrast, a completely different approach is image coding analysis of time series data using Gramian Anguler Fields (GAF) (Wang & Oates, 2014). While most methods evaluate time series in terms of points, this method evaluates time series data in terms of a set of data, or range. Specifically, the entire time series is first divided into several smaller windows. Each window corresponds to a smaller time series. there are various methods for time series anomaly detection using GAF. The simplest and lightest model is a combination of GAF and Autoencoder (Hinton & Salakhutdinov, 2006). There are several advantages to studying this method. First, it is much lighter than RNN models; second, GAF images can be easily converted into models for anomaly classification, as they differ by anomaly type; and third, it can be used with RNNs for ensemble training, as it is a completely different approach. On the other hand, there are some challenges, and existing research has shown that the accuracy is inferior to methods using RNNs. In some cases, approaches such as truncating test data values were taken to improve accuracy (Garcia, Michau, Ducoffe, Gupta, & Fink, 2021). However, this is an improvement to the dataset, not to the framework using GAF, and cannot be applied to all frameworks using GAF. In this paper, we propose an unsupervised anomaly detection model that combines Gramian Angular Fields (GAF) and Autoencoder. Since this model does not have an RNN part, it is capable of parallel processing of time-series data. Furthermore, this paper considers that the low accuracy of anomaly detection, which is a problem in existing models, is caused by structural problems in the GAF and Autoencoder parts, and attempts an approach to improve the accuracy by removing them. The contributions of this paper are threefold: 1) Proposed anomaly detection framework combining GAF and Autoencoder 2) Identified a structural problem with Gramian Angular Fields in the anomaly detection problem, proposed a remedy for the problem, and conducted experiments to confirm the effectiveness of the remedy. 3) Experiments were conducted to determine the extent to which the Autoencoder portion of the anomaly detection method, which combines Gramian Angular Fields and Autoencoder, has an effect on accuracy.

2. GRAMIAN ANGULER FIELDS

Gramian Anguler Fields (GAF) is a method for encoding time series as images. To construct a GAF image, consider a time series X with n components. Rescaling to the interval [-1, 1] is performed on the time series X. The $\tilde{X}_i$ is the new time series component obtained by rescaling, and the scaled time series is shown in the following equation.

$$\tilde{X} = \{\tilde{x_1}, \tilde{x_2}, \ldots, \tilde{x_n}\} \tag{1}$$

The rescaled time series is expressed in a polar coordinate system as in equation (2).

$$\begin{cases} \phi = \arccos{(\tilde{x}_i)}, -1 \leq \tilde{x}_i \leq 1, \tilde{x}_i \in \tilde{X} \\ r = \frac{t_i}{N}, t_i \in \mathbb{N} \end{cases} \tag{2}$$

Furthermore, by considering the triangular summation between each element represented in polar coordinates, it is possible to identify temporal correlations within different time intervals.There are two types of GAFs: Gramian Anguler Difference Fields (GADF), which use a sine function, and Gramian Anguler Summation Fields (GASF), which use a cosine function. GADF and GASF are defined as expressions (3),(4).

$$GADF = \begin{bmatrix} \sin{(\phi_1 - \phi_1)} & \cdots & \sin{(\phi_1 - \phi_n)} \\ \sin{(\phi_2 - \phi_1)} & \cdots & \sin{(\phi_2 - \phi_n)} \\ \vdots & \ddots & \vdots \\ \sin{(\phi_n - \phi_1)} & \cdots & \sin{(\phi_n - \phi_n)} \end{bmatrix}$$
$$= \sqrt{I - \tilde{X}^{2'}} \cdot \tilde{X} - \tilde{X}' \cdot \sqrt{I - \tilde{X}^2} \tag{3}$$

$$GASF = \begin{bmatrix} \cos{(\phi_1 + \phi_1)} & \cdots & \cos{(\phi_1 + \phi_n)} \\ \cos{(\phi_2 + \phi_1)} & \cdots & \cos{(\phi_2 + \phi_n)} \\ \vdots & \ddots & \vdots \\ \cos{(\phi_n + \phi_1)} & \cdots & \cos{(\phi_n + \phi_n)} \end{bmatrix}$$
$$= \tilde{X}' \cdot \tilde{X} - \sqrt{I - \tilde{X}^{2'}} \cdot \sqrt{I - \tilde{X}^2} \tag{4}$$

Two advantages exist for GAF. 1) In GAF, the time interval increases as the position moves from the upper left to the lower right of the Gram matrix, thus preserving the temporal dependence. Here $G_{i,j}$ represents the relationship between time i and j. 2)The main diagonal component, $G_{i,i}$, contains the original information in the Cartesian coordinate system before transformation to the polar coordinate system. Therefore, by using the main diagonal components, time series can be approximately reconstructed from the features learned by deep neural networks, etc.
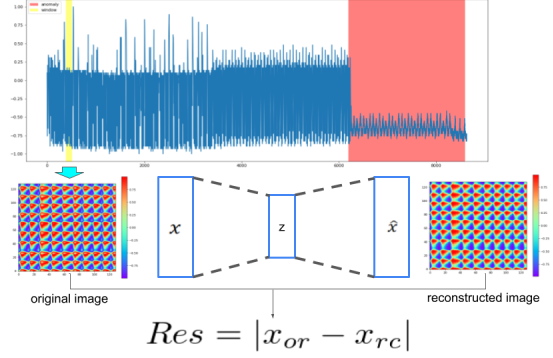

$$Res = |x_{or} - x_{rc}|$$

Figure 1. Algorithm of the proposed method

3. PROPOSED METHOD

3.1. Anomaly detection framework

When handling long time series in GAF, the size of the image data increases as a power of the size of the time series, as is clear from the formula (3). In addition, if a time series that is too long is input for time series anomaly detection, the problem arises that anomalies in localized areas may be missed. Therefore, in the proposed framework, as shown in Fig.reffig: algorithm of the proposed method, the target time series data is first decomposed into fine sub-time series with window size l, and each sub-time series is converted into image data by GAF. This image data is input to Autoencoder, and the reconstructed image is output from the network. The L1 norm between the reconstructed image and the original image is then taken, and if the L1 norm exceeds a certain threshold value, the partial time series is detected as an abnormal segment. Defining the residual $Resk$ of the kth sub-series as the L1 norm of the difference between the input x_{or} of size l^2 and its reconstruction x_{rc}, the following formula holds.

$$Resk = \sum_{i=1}^{l^2} |x_{or}^i - x_{rc}^i| \tag{5}$$

If the value of the L1 norm exceeds the set threshold value, the window is set as the window containing the abnormal series position and anomaly detection is performed.

3.2. Proposal to improve accuracy

The anomaly detection method combining GAF and Autoencoder has several problems that cause low accuracy. In this section, we present the problems that cause low accuracy and discuss possible improvements.

3.2.1. Overlay of GADF and GASF

GADF and GASF each have their own problems: GADF results in an output GAF image that is all zeros when there are only two values in the window to be transformed. This is evident from the equation (3). This means that GADF cannot

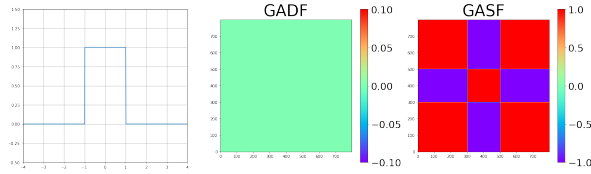detect anomalies in pulse or step format, as shown in Figure 2.

Figure 2. Image of pulse function transformed by GADF and GASF, where pulse are detected by GASF but not by GADF.

On the other hand, GASF is the same matrix even if the time series is inverted with respect to the x-axis. This is evident from the equation (4). In other words, if a value goes up or down in a certain window, the time series data has the exact opposite meaning, but the data in the GASF image is exactly the same. As a simple example of this problem, the case of a sinusoidal function is shown in Figure 3.
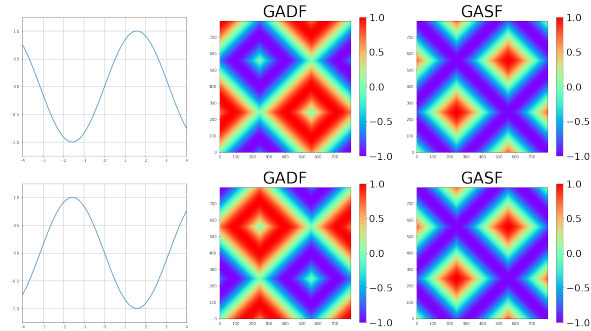

Figure 3. Images of transformed sine function and its inverted function by GADF and GASF. GADF reflects the inversion of the function in the image, while GASF does not.

These problems are believed to have a significant negative impact on anomaly detection by GAFs. Therefore, in this paper, we propose an improvement proposal called "superposition of GADF and GASF". Specifically, GADFs and GASFs are superimposed in the channel direction and used as inputs to Autoencoder. This improves accuracy by compensating each other for anomalies that cannot be detected by the individual GADFs and GASFs alone.

3.2.2. Normalized scaling of GAF images

When the GADF has only two extremes in the time series, near the maximum or minimum, the values in the GADF image are reduced to extremely small values. This can be easily explained in terms of the GADF latent space, which is illustrated in Fig. reffig:GADFspace. This figure shows that when the values in the time series are only at the extremes near the maximum or minimum, i.e., when the values of the series data are concentrated in the red circles in the four corners, the GADF values are concentrated near 0. In many cases,

windows with only extreme values in time-series data contain anomalous values, but as the value of GADF becomes smaller, the value of the L1 norm Resk in the formula (5) also becomes smaller. As a result, the residuals that indicate abnormal values become smaller, making it impossible to detect abnormalities.

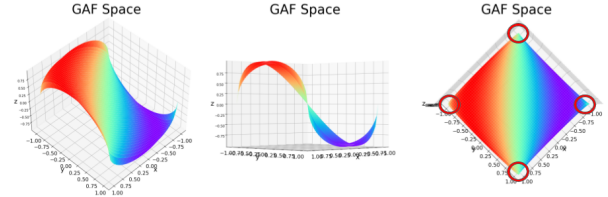Therefore, this paper proposes an improvement called "Nor-

Figure 4. GADF latent space. It can be seen that when the values of the series data are concentrated in the red circles in the four corners, the GADF values are concentrated around 0.

malized Scaling of GAF Images". Specifically, GADF image data is normalized to $[0, 1]$. This is expected to unify the residuals indicating abnormal values to the same scale as the other windows and improve accuracy.

3.2.3. Improved accuracy of Autoencoder network reconstruction

Not only GAF images, but also anomaly detection methods using Autoencoder have a problem that the accuracy of anomaly detection is reduced due to the low reconstruction accuracy of the Autoencoder network. Therefore, this paper proposes the following two improvements to improve the reconstruction accuracy of the Autoencoder network.

Complexity of Autoencoder network In existing papers, only simple networks were used for the Autoencoder portion. In this paper, we employ a network based on Unet (Ronneberger, Fischer, & Brox, 2015), which is a semantic segmentation model developed by Olaf et al. for biomedical applications. Recently, there is a model that is also used in Autoencoder and performs highly accurate reconstruction. The detailed structure of Unet used in this paper is described in the next chapter, "Experiments.

Adoption of SSIM to loss function The usual anomaly detection by Autoencoder fails to capture minute anomalies due to the blurring of the image, which results in a large restoration error even for normal data. Therefore, Bergmann et al. employed Structural SIMilarity (SSIM) as the loss function and succeeded in improving the accuracy of anomaly detection by Autoencoder by creating sharper reconstructed images (Bergmann, L 旦 we, Fauser, Sattlegger, & Steger, 2019). SSIM performs anomaly detection by calculating the similarity of "brightness," "contrast," and "structural information" in a small frame for a single image. In other words, it focuses on the similarity of the local structure of the image.This SSIM is

also used in this paper to test whether it is effective in improving accuracy in GAF.

4. EXPERIMENT

Experiments were conducted to verify whether the time-series anomaly detection framework proposed in the previous section actually works and whether the proposed improvements actually lead to improved accuracy.

4.1. Dataset

For the dataset, we will use the telemanon file on github published by Hundman et al in "Detecting Spacecraft Anomalies Using LSTMs and Nonparametric Dynamic Thresholding" (Hundman, Constantinou, Laporte, Colwell, & Soderstrom, 2018). This dataset is telemetry data obtained from the Soil Moisture Active/Passive Mission (SMAP) satellite launched by NASA and the Mars Science Laboratory (MSL) spacecraft. These data are divided into teacher data with no anomalies and test data with anomalies. The test data are labeled with anomaly intervals and are currently cited in over 500 papers.A summary of these datasets is shown in table 1. These data sets have multiple channels in each sequence. The

	SMAP	MSL	Total
Total anomaly sequences	69	36	105
Point anomalies (% tot.)	43(62%)	19(53%)	62(59%)
Contextual anomalies (% tot.)	26(38%)	17(47%)	43(41%)
Unique telemetry channels	55	27	82
Unique ISAs	28	19	47
Telemetry values evaluated	429,735	66,709	496,444

Table 1. Dataset information

telemetry value is the 0th channel, and the remaining channels are data about the commands sent. In this paper, only the 0th channel, the telemetry value, is used in the experiments.

4.2. Performance Evaluation Method

Abnormality can be defined in various ways. In this paper, an abnormality window is defined as a window that contains at least one abnormal series, and a normal window is defined as a window that consists only of normal series. The ROC-Curve and the AUC (Area Under the Curve) value based on the ROC-Curve are used as the evaluation index of the algorithm. In this paper, we calculate the AUC for each sequence for SMAP and MSL, and define the average value as the performance of the model.

4.3. Experimental Tasks and Results

This section describes the experimental tasks and results conducted in this thesis. In all experiments, the framework described in the previous section, "Framework for Anomaly Detection," is used. In addition, the networks shown in the table 2 are used for the Autoencoder except for the experiments using Unet. PReLU was used as the activation function, with a

Encoder	Input	–	–	$512 \times 512 \times 1$
	Conv1	2	2	$256 \times 256 \times 32$
	Conv2	2	2	$128 \times 128 \times 64$
FullyConnected	Flatten	–	–	$128 \times 128 \times 64$
	Bottleneck	–	–	512
	Dense2	–	–	$128 \times 128 \times 64$
Decoder	Mirrors the architecture of the Encoder			

Table 2. Architecture of Autoencoder

learning rate of 0.001, β 1 of 0.9, and Adam optimizer with L1 error as the loss function. Each window size l was set to 128, the number of epochs to 50, and the batch size to 128.

4.3.1. Performance evaluation of GADF and GASF overlays

Performance evaluation and comparison of three frameworks: 1) anomaly detection in GADF images only 2) anomaly detection in GASF images only 3) anomaly detection in GADF and GASF superimposed images
The results of the experiment are shown in the following table 3.

framework	SMAP	MLS
GADF	0.637	0.597
GASF	0.759	0.660
GADF + GASF	0.768	0.721

Table 3. Result of framework using overlay of GADF and GASF

The *GSDF + GASF* (GADF and GASF overlay method) performed best for both SMAP and MSL. The performance of GADF alone was considerably less accurate than that of GASF alone, but this is thought to be due to the shortcomings of GADF in detecting pulse- and step-type anomalies described in the previous section, which had a significant negative impact on the overall accuracy of the framework.

4.3.2. Performance evaluation of normalized scaling of GAF images

Performance evaluation and comparison will be made for the following two frameworks: 1) anomaly detection without normalized scaling of GAF images 2) anomaly detection with normalized scaling of GAF images. However, for these two frameworks, "GADF and GASF superposition" shall be used in both cases.The results of the experiment are shown in the following table 4.
Compared to *GSDF + GASF*, *GSDF + GASF + Min −*

framework	SMAP	MLS
GADF + GASF	0.768	0.721
GADF + GASF + Min-Max normalization	0.785	0.752

Table 4. Result of framework using Min-Max normalization

Maxnormalization (GADF and GASF overlay method plus

normalization) performed better for both SMAP and MSL.

4.3.3. Performance evaluation using Unet

Performance evaluation and comparison of the following two frameworks will be performed: 1) Anomaly detection using the network shown in the table 2 for the Autoencoder part, and 2) Anomaly detection using Unet for the Autoencoder part. The structure of the Unet used is shown in Figure 5. However, it is assumed that both "GADF and GASF overlay" and "normalized scaling of GAF images" are used for these two frameworks.
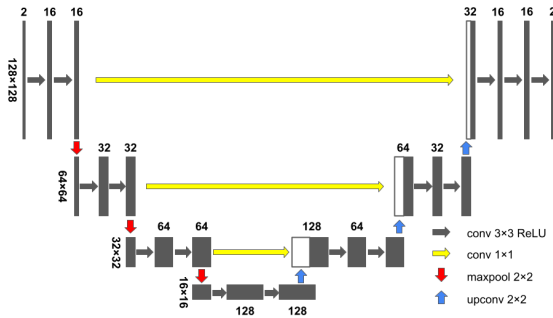

Figure 5. Structure of Unet

The results of the experiment are shown in the following table 5. Compared to $GSDF + GASF + Min - MaxNormalization$,

framework	SMAP	MLS
GADF + GASF + Min-Max normalization	0.785	0.752
GADF + GASF + Min-Max normalization + Unet	0.741	0.672

Table 5. Result of framework using Unet

$GSDF + GASF + Min - MaxNormalization + Unet$ (GADF and GASF superposition method plus normalization and Unet) performed poorly for both SMAP and MSL. As an example, taking A-4, one of the sequences in SMAP, the network-based anomaly detection of the A-4 sequence in Table 2 resulted in perfect anomaly detection with $AUC = 1.0$, whereas Unet-based anomaly detection had a much lower accuracy of $AUC = 0.639$. The reason for the lower accuracy lies in the reconstruction of the anomaly window. The GADF image of the anomaly window by the network in Table 2 is shown in Figure 6, and the GADF image of the anomaly window by the unet The GADF image of the anomaly window by the Unet is shown in Figure 7.

As can be seen, the reconstructed image by the network in table 2 fails to reconstruct the anomaly image well and the difference is large, while the reconstructed image by Unet succeeds in reconstructing even the anomaly image and the difference is small. The difference is smaller. In addition, the L1 norm between images is smaller than the L1 norm of the previous image. The L1 norm between the images is 0.234 for the former and 0.037 for the latter. In other words, the

former does not reconstruct the anomaly image well, while the latter reconstructs it almost perfectly. From these results, it can be said that the decrease in accuracy in the framework employing Unet is due to the improved performance of the network, which is able to reconstruct abnormal images in the same way as normal images.
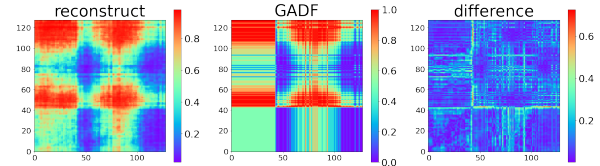

Figure 6. GADF image of anomaly part of A-4 using table2 network:from left to right, reconstruct image, original image, and difference between images
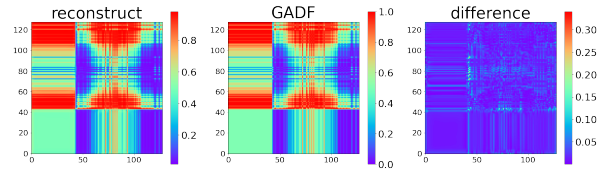

Figure 7. GADF image of anomaly part of A-4 using Unet:from left to right, reconstruct image, original image, and difference between images

4.3.4. Performance evaluation when SSIM is used for loss function

Performance evaluation and comparison will be performed for the following two frameworks: 1) anomaly detection using L1 error as the loss function 2) anomaly detection using SSIM as the loss function. For these two frameworks, both "GADF and GASF overlay" and "normalized scaling of GAF images" are assumed to be adopted.The results of the experiment are shown in the following table 6.

framework	SMAP	MLS
GADF + GASF + Min-Max normalization	0.785	0.752
GADF + GASF + Min-Max normalization + SSIM	0.795	0.675

Table 6. Result of framework using SSIM

Compared to $GSDF + GASF + Min - Maxnormalization$, $GSDF + GASF + Min - Maxnormalization + SSIM$ (GADF and GASF overlay method plus normalization and SSIM) increased SMAP accuracy but decreased performance at MSL The results show that the SMAP accuracy increased but the MSL performance decreased. First, we discuss the SMAP sequence A-4, which was also discussed in the previous section. In the A-4 sequence, $AUC = 1.0$ for both L1 norm and SSIM anomaly detection, indicating that perfect anomaly detection was achieved. The L1 norm in A-4 is shown in Table

7. In terms of the accuracy of the reconstructed images, it can be seen that in the normal area, the framework using SSIM is about halfway between the framework using Unet and the framework without Unet. In the abnormal images, the framework with SSIM performs as well as the one with L1 norm, i.e., the abnormality detection works.

framework	normal	anomaly
GADF + GASF + Min-Max normalization	0.201	0.234
GADF + GASF + Min-Max normalization + Unet	0.045	0.037
GADF + GASF + Min-Max normalization + SSIM	0.121	0.233

Table 7. L1 norm for A-4 sequence in each framework

Next, we discuss the MSL sequence P-15, which showed $AUC = 0.909$ for anomaly detection by L1 norm, $AUC = 0.873$ for anomaly detection by Unet, and $AUC = 0.575$ for anomaly detection by SSIM, showing a significant loss of accuracy in the SSIM framework. The L1 norm in P-15 is shown in Table 8. Unlike the other two frameworks, the framework using SSIM in the anomaly image has a smaller L1 norm in the anomaly image. In other words, the anomaly detection does not work because the anomaly image is also reconstructed.

framework	normal	anomaly
GADF + GASF + Min-Max normalization	0.090	0.115
GADF + GASF + Min-Max normalization + Unet	0.023	0.027
GADF + GASF + Min-Max normalization + SSIM	0.073	0.064

Table 8. L1 norm for P-15 sequence in each framework

From this, it can be said that the reason for the decreased accuracy in the framework employing SSIM is that, as in the case of Unet, the performance of the network has improved so much that abnormal images can be reconstructed in the same way as normal images. On the other hand, it is difficult to explain the reason why the judgment was successful in sequence A-4 but not in sequence P-15. Although it is only an estimation, we think that each framework has its own abnormal images that it is good at reconstructing, and this may have caused the difference in performance.

5. Conclusion and Future Work

In this paper, we proposed a GAF-based anomaly detection framework and a method for improving accuracy, and conducted experiments to verify the effectiveness of the proposed method. The proposed improvements of "overlay of GADF and GASF" and "normalized scaling of GAF images" improved the accuracy of GAF, since "GADF results in zero values when only two values exist in the window to be transformed," "GASF results in the same matrix even if the time series is inverted about the x axis," and "GADF results in the image values being concentrated around zero when the time series has only extreme values near the maximum or minimum. On the other hand, the "Improvement of anomaly detection accuracy by improving the reconstruction accuracy of

the Autoencoder network" did not work well in both Unet and SSIM due to the problem that the network accuracy improved to the extent that anomaly images could be reconstructed. However, there were some sequences in which the anomaly detection accuracy was partially improved, so this method is not completely useless. For future prospects, the first step is to analyze the different trends in anomaly image reconstruction for each framework, which will facilitate the search for the best framework for each data set. Therefore, it is necessary to conduct experiments on various data sets to determine the trends. Also, consider networks that do not allow reconstruction of anomalous images. For example, by introducing an attention mechanism, we can create a network that can learn only features of normal images.

References

Bergmann, P., Löwe, S., Fauser, M., Sattlegger, D., & Steger, C. (2019). Improving unsupervised defect segmentation by applying structural similarity to autoencoders. In *Proceedings of the 14th international joint conference on computer vision, imaging and computer graphics theory and applications.* SCITEPRESS - Science and Technology Publications. doi: 10.5220/0007364503720380

Garcia, G. R., Michau, G., Ducoffe, M., Gupta, J. S., & Fink, O. (2021, feb). Temporal signals to images: Monitoring the condition of industrial assets with deep learning image processing algorithms. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 236(4), 617–627. doi: 10.1177/1748006x21994446

Graves, A. (Ed.). (2012). *Supervised sequence labelling with recurrent neural networks.* Springer Berlin Heidelberg.

Hinton, G. E., & Salakhutdinov, R. R. (2006, July). Reducing the dimensionality of data with neural networks. *Science*, 313(5786), 504-507. doi: 10.1126/science.1127647

Hundman, K., Constantinou, V., Laporte, C., Colwell, I., & Soderstrom, T. (2018, jul). Detecting spacecraft anomalies using LSTMs and nonparametric dynamic thresholding. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery &amp data mining.* ACM. doi: 10.1145/3219819.3219845

Mogren, O. (2016). *C-rnn-gan: Continuous recurrent neural networks with adversarial training.*

Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-net: Convolutional networks for biomedical image segmentation.*

Wang, Z., & Oates, T. (2014). Encoding time series as images for visual inspection and classification using tiled convolutional neural networks..