International Journal of Precision Engineering and Manufacturing (2024) 25:2583–2598
https://doi.org/10.1007/s12541-024-01069-6

Online ISSN 2005-4602
Print ISSN 2234-7593

REGULAR PAPER

Encoding Time Series as Images for Anomaly Detection in Manufacturing Processes Using Convolutional Neural Networks and Grad-CAM

Young-Joo Hyun[1,2] · Youngjun Yoo[1] · Yoonseok Kim[1,2] · Taeheon Lee[1,2] · Wooju Kim[2]

Received: 15 February 2024 / Revised: 13 June 2024 / Accepted: 13 June 2024 / Published online: 28 June 2024

Abstract
This study aims to develop an artificial intelligence-based model for analyzing the condition and detecting anomalies by encoding time-series data from manufacturing processes as images. Deep learning has demonstrated the significance of data analysis and anomaly detection in the vision field, and Convolutional Neural Networks (CNN) models have shown exceptional performance and high applicability in image analysis. Based on this, our study intends to utilize image encoding techniques to perform anomaly detection on time-series data. Data such as force, vibration, and sound from equipment during the manufacturing process are collected and transformed into images using various methods, including Gramian Difference Angular Field, Gramian Summation Angular Field, Markov Transition Field, and Recurrence Plot (RP). The transformed image data is then trained and classified for equipment conditions using various CNN models. Finally, we adopt the RP image encoding method and ResNet50 model, which demonstrated the highest accuracy of 99.6%, and compare them to the top 5 models. Based on the high accuracy demonstrated by the top five models, our proposed approach has proven to have significant performance, exhibiting a high success rate of over 90% even when applied to actual data for CNC-machining process. Through this, we propose a process that utilizes the explainable AI Grad-CAM system to identify the feature layer area of the image and confirm the presence of anomalies. With the proposed process, workers can identify abnormal areas or segments of abnormal conditions in the transformed image graph. By providing evidence for state judgment, even inexperienced workers can easily check the condition of manufacturing equipment.

Keywords Anomaly detection · Encoding image · Grad-CAM · Transfer learning · Deep learning

1 Introduction

Equipment maintenance is crucial for maintaining productivity and quality in manufacturing processes. However, equipment failures or abnormal conditions can occur due to various factors in the manufacturing process, making it challenging for on-site workers to determine the equipment's condition accurately while carrying out the manufacturing process. Furthermore, replacing equipment too quickly can result in financial losses, while delaying replacements can cause complete equipment damage, leading to decreased productivity and increased costs. Therefore, anomaly detection, which can quickly and accurately prevent equipment abnormalities, significantly contributes to improving productivity and efficiency in manufacturing processes. Through anomaly detection, abnormal machine conditions can be promptly identified, and fault prevention and repair plans can be developed to minimize downtime and operational disruptions [1, 2].

In many industrial environments, equipment generally operates with vibrations and noise. Therefore, existing anomaly detection techniques for predictive maintenance analyze periodic time-series data generated during the operation of manufacturing equipment to distinguish between normal and abnormal states [3, 4]. Commonly used

✉ Youngjun Yoo
youdalj@kitech.re.kr

✉ Wooju Kim
wkim@yonsei.ac.kr

[1] Korea Institute of Industrial Technology (KITECH), 89, Yangdaegiro-gil, Ipjang-myeon, Seobuk-gu, Cheonan-si, Chungcheongnam-do, Republic of Korea

[2] Department of Information and Industrial Engineering, Yonsei University, Seoul, Republic of Korea

techniques include time-series analysis, complex statistical models, and deep learning models based on time-series data. Literature [5] summarizes Various time series modeling analyses, such as the ARMA model, for Vibration-Based Structural Health Monitoring (SHM). Blázquez-García, Ane, et al. [6] is give a structured and comprehensive state-of-the-art on outlier detection techniques in time series data. Manufacturing equipment has a specific bathtub curve that depends on the factors affecting their condition. Such time-series data has a trend that continues with an increasing or constant pattern. Time-series data analysis models, such as LSTM (Long Short-Term Memory) [7] or RNN (Recurrent Neural Network) [8], are used in machine learning to predict future changes in data based on these temporal and sectional characteristics [9]. Machine learning research and technology development are actively taking place in various fields. Among these, Convolutional Neural Networks (CNNs), which analyze and classify image-based data, have shown exceptional performance in vision tasks such as image recognition and video processing. With the proven significance of deep learning technology for image analysis, researchers are exploring various methods to apply time-series data to image analysis models, in addition to images and videos. Consequently, a technique of transforming time-series data, which is based on temporal order, into images has been developed [10–12].

In manufacturing processes, data is commonly collected from numerous sensors. However, simultaneously analyzing all sensor data for time-series data analysis presents challenges. Conversely, image analysis overcomes this challenge by amalgamating multiple sensor data into a single image for analysis.The main advantage of detecting anomalies in a 2D map, as opposed to 1D time-series data, lies in its enhanced interpretability. This intuitive visualization empowers domain experts to easily comprehend and track patterns, resulting in a clearer explanation of anomalies. In essence, it facilitates a more effective representation of characteristic features. Moreover, these images contain rich and diverse information, enabling a comprehensive understanding of the system's condition [13]. Consequently, the analysis of images that integrate multiple sensor data proves valuable for detecting anomalies in intricate systems like manufacturing processes [14].

This study transforms time-series data from manufacturing equipment into images using four different methods: Gramian Difference Angular Field (GADF), Gramian Summation Angular Field (GASF), Markov Transition Field (MTF), and Recurrence Plot (RP). By comparing and analyzing the suitability of different backbones based on the characteristics of each image transformation method, discrimination is improved. This verifies the effectiveness of using image encoding for anomaly detection technology and applying deep learning models widely used in the computer vision field, as well as the usefulness of processing data [11, 12].

In manufacturing processes, detecting anomalies and understanding the state of equipment beforehand are crucial for on-site workers. However, most analysis models only provide accuracy or classification results without a clear explanation of how individual algorithms are learned or why they produce specific outcomes. [15, 16] Therefore, even if the performance of anomaly detection and classification/analysis is excellent, it is difficult to understand the scope and extent of actual problems that may arise [17]. In summary, while machine learning algorithms can provide accurate predictions, it is often challenging to interpret the results, which can lead to a lack of trust and reliance on these methods. To address this issue, researchers are exploring explainable AI techniques that can provide transparent insights into the decision-making process of machine learning algorithms. By providing interpretable results, these methods can facilitate better decision-making by domain experts and increase trust in machine learning models for actual data applications [18].

As a result, on-site workers only receive information about equipment issues, without additional interpretation, making it difficult to trust the model. To address these issues, this study proposes an anomaly detection model that utilizes Grad-CAM (Gradient-weighted Class Activation Mapping), a promising approach that visually explains the classification results of CNN models in object detection and recognition, in addition to the CNN model. Through this approach, we aim to provide on-site workers with spatial and temporal information on the location and size of detected anomaly patterns, contributing to identifying and tracking the causes of anomaly patterns during the manufacturing process.

The paper is structured as follows. Section 2 presents the process of frequency analysis and interpretation of time-series data. Section 3 verifies the proposed methodology, while Sect. 3.1 explains the dataset and the backbone model used in this study. Section 4 presents the quantitative results of the proposed method. Finally, the conclusion is presented in Sect. 5.

2 Background

2.1 Conversion of Time Series Signals from the Time Domain to the Frequency Domain

In this study, we transformed the raw time-series data using FFT (Fast Fourier Transform) and a low-pass filter, as depicted in Fig. 1. By transforming the raw time-series data into frequency signals, we analyzed the equipment-generated vibrations' frequencies during the manufacturing process and detected anomalies or defects. After transforming the

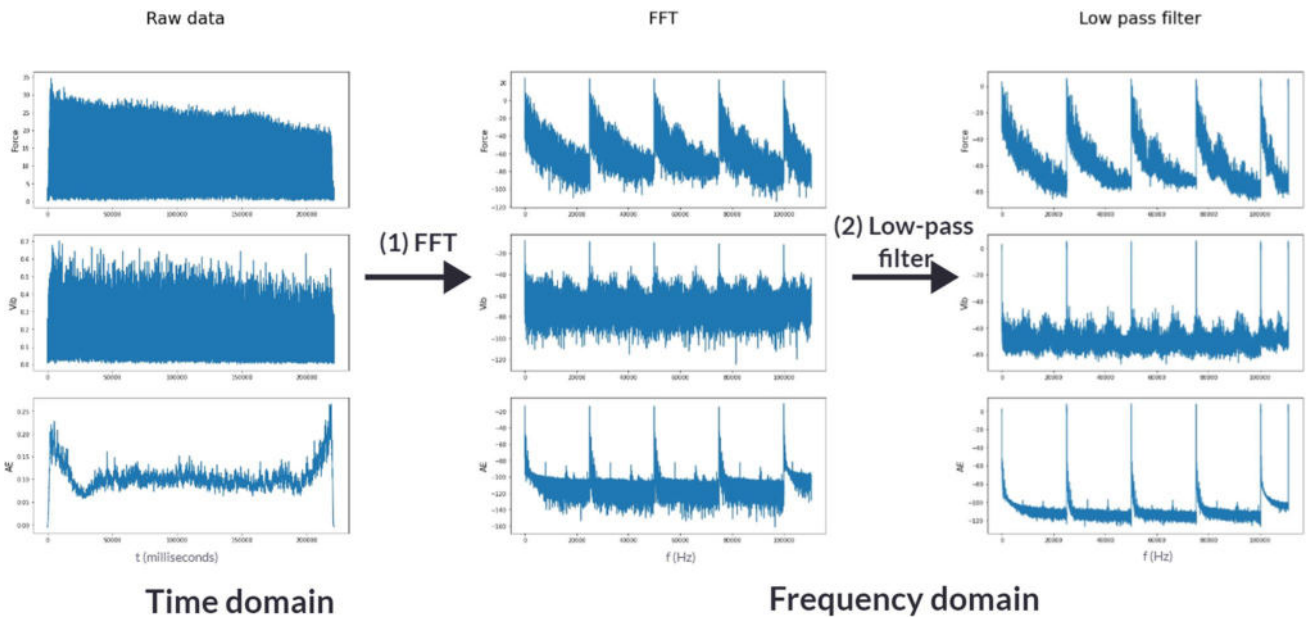

Fig. 1 Conversion of the time-series data to frequency domain through FFT (Fast Fourier transform) and Low-pass filter

continuous time-series data into frequency signals, we eliminated the unnecessary high-frequency components through a low-pass filter to reduce noise in the data and capture more detailed information. This process resulted in more meaningful analysis results for encoding image.

2.1.1 Fast Fourier Transform

FFT is a method for transforming signals that represent amplitude and vibration intervals in the time domain into the frequency domain. This allows for the analysis of the periodicity of the frequency domain, which can be utilized for various signal analysis purposes, such as detecting or removing abnormal values like noise [19]. FFT is an efficient algorithm that quickly performs the DFT (Discrete Fourier Transform) and its inverse transformation on discrete input signals. It is widely used in many fields, ranging from algorithms for finding roots of partial differential equations in digital signal processing to other applications. DFT is a method of transforming a discrete-time signal measured in the time domain into the frequency domain [20]. Given a signal X with a series of consecutive values $X = x_1, x_2, \cdots, x_{n-1}$, the DFT is expressed as the Eq. (1).

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn}, \qquad k = 0, \cdots, N-1 \tag{1}$$

where the variable x_n represents the discrete values of the given signal, and there are N consecutive samples of the signal. DFT calculates N complex values in the frequency domain from N sample signals measured in the time domain using the equation mentioned above. It is mainly used for frequency analysis and can be inversely transformed to the original signal. DFT is widely used in digital signal processing, as well as various fields such as voice and video processing, data compression, and communication systems [21].

2.1.2 Pass Filter

The FFT plot (Fig. 1) shows a graph that was transformed by applying the DFT Eq. (1). to the data in this study. By transforming the time-domain data into the frequency-domain, it becomes possible to observe the periodicity of each feature. However, the vibration signal is relatively less clear compared to the force and acoustic emission (AE) signals. Therefore, this section proposes a method of applying an additional pass filter to improve the clarity of the vibration signal.

A pass filter is a type of filter that permits a specific frequency range to pass through in the frequency domain. Pass filters are categorized into three types: Band Pass Filters, Low Pass Filters, and High Pass Filters. Low Pass Filters enable signals with frequencies lower than the specified frequency to pass through, while High Pass Filters allow signals with frequencies higher than the specified frequency. Band Pass Filters allow only signals between two specific frequencies to pass through. By controlling these frequency components, pass filters are used for a variety of signal processing purposes as [22, 23]. For example, in voice signal processing, filtering can be applied to extract the desired frequency range of voice or music or to remove noise. High-pass filters are used in image processing to improve

the clarity of images since high-frequency components are typically noise [24].

Figure 2 shows the data of this study after applying a low-pass filter. Although Band Pass Filters and High Pass Filters showed decreased periodicity and discrimination, the data became clearer than before when a Low Pass Filter was applied. This improvement was observed not only in Force and AE but also in Vib, which previously had issues with periodicity. As a result, this study adopts frequency data analysis using Low-pass filters. Nevertheless, as signals can take various forms depending on the data and domain characteristics, it is necessary to apply an appropriate Pass filter based on the data and domain.

2.2 CNN-Based Image Classification Model

Convolutional Neural Networks (CNNs) are composed of convolution and pooling layers that help preserve local information in the input data during learning. In the convolution layer, a filter (kernel) is convolved with the input image to create a new feature map. The filter, made up of small parameters, is optimized during the learning process to extract features from the input image. In the pooling layer, the feature map is partitioned into smaller areas, and the maximum or average value is extracted from each area to create a new feature map. Through these processes, the CNN model can extract features that are robust to variations in location or size in the input image. The CNN repeatedly applies these convolution and pooling layers and finally adds a fully connected (FC) layer to perform classification tasks. The FC layer utilizes the features extracted from the previous layers to perform tasks such as classification or prediction of the input data. As a result, CNN demonstrates high performance in recognition and classification tasks for various data types such as images [25].

2.3 Explainable Artificial Intelligence

Explainable Artificial Intelligence (XAI) is a technique that aims to provide users with an explanation of how artificial intelligence (AI) models operate and make decisions in an understandable way. Unlike traditional models, XAI can clarify how a model functions, allowing users to understand and verify its results. This helps to build trust in the output of the

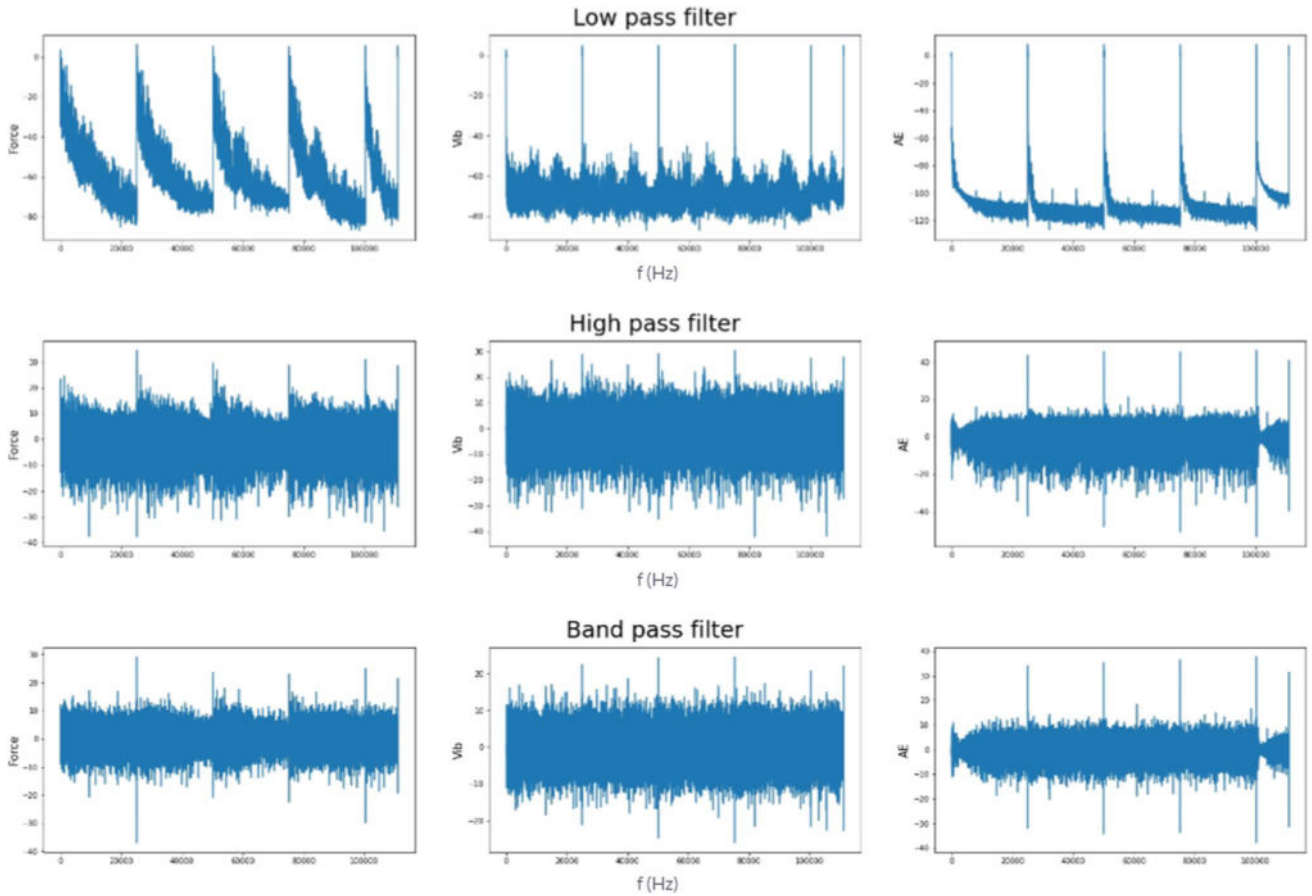

Fig. 2 Pass filter results of the vibration signals

model. XAI is primarily utilized in the field of deep learning, which employs large datasets and complex models. Due to the vast number of parameters in deep learning models, explaining how the model works can be a challenging task. To address this issue, XAI analyzes the internal decision-making processes of the model and provides visualizations. XAI is used in various fields to enhance the reliability and transparency of AI models, enabling users to interpret and improve the model's results by understanding its decision-making processes [26].

2.3.1 Gradient-Weighted Class Activation Mapping

Gradient-weighted Class Activation Mapping (Grad-CAM) is a method of visualizing the features of an image that influence the results of a CNN model. It explains which parts of the image the CNN is focusing on to classify it into a particular class. The output of a CNN image classification model, which consists of several layers of convolutional and pooling layers, is explained using the feature maps inside the neural network. While these maps provide numerical results, Grad-CAM uses gradients of the model's output to visualize which parts of the input image had the most significant impact on the classification decision [17].

Grad-CAM visualizes the areas of the feature map that contribute to activating a particular class based on the overall image classification. To achieve this, Grad-CAM calculates the gradients of the feature map output from the last convolutional layer of the CNN and multiplies them with the class-specific gradients at that layer to calculate the importance of each location. The resulting important information is visualized as a heatmap of the same size as the input image. This heatmap provides a direct understanding of which parts of the input image the deep learning model is referencing when making classification decisions. Thus, Grad-CAM is used to interpret the results of image classification models and visualize the decision-making process of the model in an explainable format.

The following equation is utilized to generate the Grad-CAM map,

$$L^{c}_{gradCAM} = ReLU\left(\sum_{k=1}^{K} \sum_{i=1}^{H} \sum_{j=1}^{W} \alpha_{c,k}(i,j) A_k(i,j) \right). \quad (2)$$

The Eq. (2) represents the importance of each location in an input image to the classification of a specific class. In this equation, $L^{c}_{gradCAM}$ represents the Grad-CAM map for a particular class c, where $\alpha_{c,k}$ refers to the weight of the gradient of the feature map outputted from the last convolutional layer concerning class c. Additionally, $A_k(i,j)$ denotes the value of the $i-th$row and $j-th$ column in the feature map outputted from the last convolutional layer. Through the multiplication of $\alpha_{c,k}$ and $A_k(i,j)$, using the gradient for class c, the value for each pixel (i,j) is calculated, indicating the pixel's contribution to the classification of class c. By computing this value for all pixels, Grad-CAM generates a map that visually shows the regions of the CNN that contribute most to the activation of a specific class. This method improves the understanding of the CNN's operation and can help to optimize the model.

2.4 Encoding as Images from Time Series Data

GADF, GASF, MTF, and RP are all image encoding methods that convert time series data into images. GADF and GASF construct a correlation matrix of the Gramian angular field (x, y) to visualize frequency information and transform the data. GAF has achieved significant results in applications such as lightweight rolling bearing fault diagnosis [27] and monitoring workers' safety compliance [28]. MTF discretizes the time series data by partitioning it into bins, while RP uses a method to represent time series data as points in a 2D space to analyze its dynamic characteristics. In the field of exoplanet identification, MTF has proposed to analyze unevenly-sampled time series [29]. Additionally, MTF, in conjunction with deep learning techniques, has been applied to monitor workers' safety compliance [30]. In the shipbuilding industry, multisensor signals collected during the manufacture of anchor chains were analyzed through RP [31]. In the study of Earth's past climates, recurrence analysis was utilized to compare climate change patterns [32].

Each of these image encoding methods has unique features, and the proposed method aims to use them to develop an anomaly detection and classification model. In this section, we will explain the four image encoding methods that we used in this study: GADF, GASF, MTF, and RP. By fusing these encoded images, the resulting data provides a more comprehensive, multi-dimensional input for pattern recognition and machine learning algorithms, thereby enhancing our ability to conduct intricate data analyses. Furthermore, by employing image and data fusion techniques, damage detection challenges can be effectively transformed into pattern recognition tasks [33]. This approach is particularly valuable for processing multi-sensor data and detecting anomalies in complex systems, thus significantly aiding in assessing their health and operational status.

2.4.1 Gramian Angular Field (GAF)

The GAF method, proposed by Z. Wang in 2015, encodes time-series signals to images [34] as Fig. 3. In the industrial sector, there are repetitive manufacturing and production processes, and the data collected from them exhibit periodicity. GAF preserves the temporal relationships of the signal and separates the characteristic signal from the interference signal, encoding the time-domain signal as an image [27, 28]. GAF represents time-series in a 2D

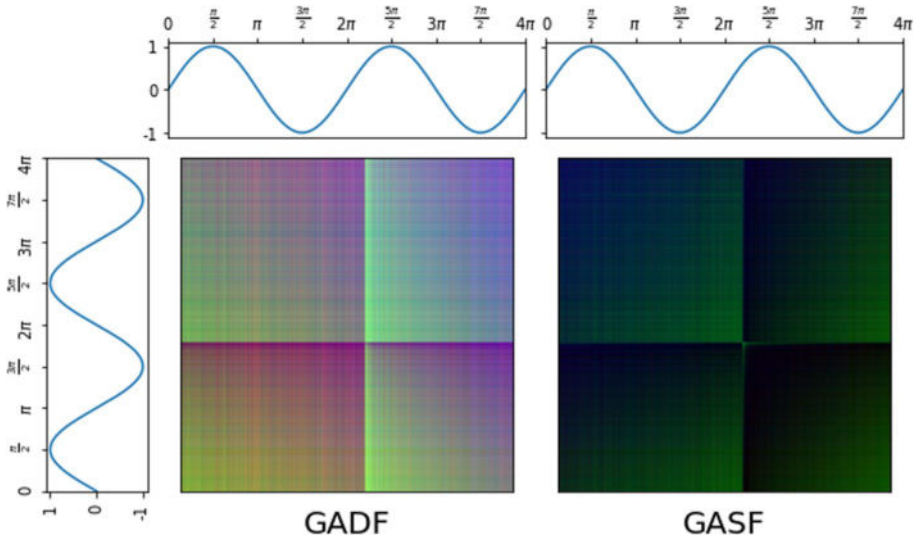

Fig. 3 Encoding results of gramian angular field

polar coordinate system that uses angles and distances between coordinates, instead of the vertical-based Cartesian coordinate system such as X-axis and Y-axis, thereby minimizing information loss of 1D time-series data. The resulting GAF images show temporal dependence as they move from the upper left to the lower right, and the main diagonal contains both the original data value and the angle information, allowing the original data to be reconstructed as follows:

$$\tilde{X}_i = \frac{(X_i - max(X)) + (X_i - min(X))}{max(X) - min(X)}. \tag{3}$$

Given a time-series X with real-valued observations x_n, $X = x_1, x_2, x_3, ..., x_n$ represents the observation values of the time-series. Encoding a GAF image, X is normalized and rescaled to $[-1, 1]$ using the min-max scaler in equation (3).

The angle ϕ is the arccos of x_i, and the diameter r is a constant represented by $\mathbb{N}$ to normalize the range of values in the polar coordinates based on the time index t_i as follows:

$$\begin{cases} \phi = \arccos(\tilde{x}_i), & -1 \le \tilde{x}_i \le 1, \quad \tilde{x}_i \in \tilde{X} \\ r = \frac{t_i}{N}, & t_i \in \mathbb{N}, \end{cases} \tag{4}$$

This transforms the time series X into polar coordinates, which can be used to identify the frequency of spectral data. When ϕ â^^ $[0, \pi]$, $\cos(\phi)$ is a bijective function, which creates a one-to-one correspondence in the polar coordinate system and has an inverse function when given any arbitrary time series [34].

GASF and GADF are determined by the sum and difference of trigonometric angles between each point in a time series. GASF uses cosine to calculate the sum, while GADF uses sine to calculate the difference as following equations,

$$\begin{aligned} GASF &= [cos(\phi_i + \phi_j)] \\ &= \tilde{X}' \cdot \tilde{X} - \sqrt{I - \tilde{X}^2}' \cdot \sqrt{I - \tilde{X}^2} \end{aligned} \tag{5}$$

$$\begin{aligned} GADF &= [sin(\phi_i - \phi_j)] \\ &= \sqrt{I - \tilde{X}^2}' \cdot \tilde{X} - \tilde{X}' \cdot \sqrt{I - \tilde{X}^2} \end{aligned} \tag{6}$$

2.4.2 Markov Transition Field (MTF)

MTF is a method of encoding time series data as images by connecting Markov chains with time series data [34] as Fig. 4. MTF calculates the transition probability of the Markov chain based on discretized time series data and converts it into the pixel values of an image. Encoding Images using MTF simplifies patterns in time series and can be applied to various fields such as image classification and pattern recognition [29] as

$$M = \begin{bmatrix} w_{ij} | \, x_1 \in q_i, \, x_1 \in q_j & \cdots & w_{ij} | \, x_1 \in q_i, \, x_n \in q_j \\ w_{ij} | \, x_2 \in q_i, \, x_1 \in q_j & \cdots & w_{ij} | \, x_2 \in q_i, \, x_n \in q_j \\ \vdots & \ddots & \vdots \\ w_{ij} | \, x_n \in q_i, \, x_1 \in q_j & \cdots & w_{ij} | \, x_n \in q_i, \, x_n \in q_j \end{bmatrix}. \tag{7}$$

To apply MTF to time-series data X, it is divided into Q quantiles, and the value of time index t_i is represented by x_i. x_i is assigned to quantile q_j (jâ^^ $[1, Q]$). A weighted adjacency matrix W of size $Q \times Q$ is constructed based on the first-order Markov chain method along the time axis. $w_{i,j}$ represents the transition probability from quantile qi to q_j where the data value belongs. The difference between two times $|i - j|$ represents the transition probability between quantiles and if the width of a quantile is large, it is aggregated to the quantile closest to the mean value. If the width of a quantile is small, the value is aggregated at the extreme

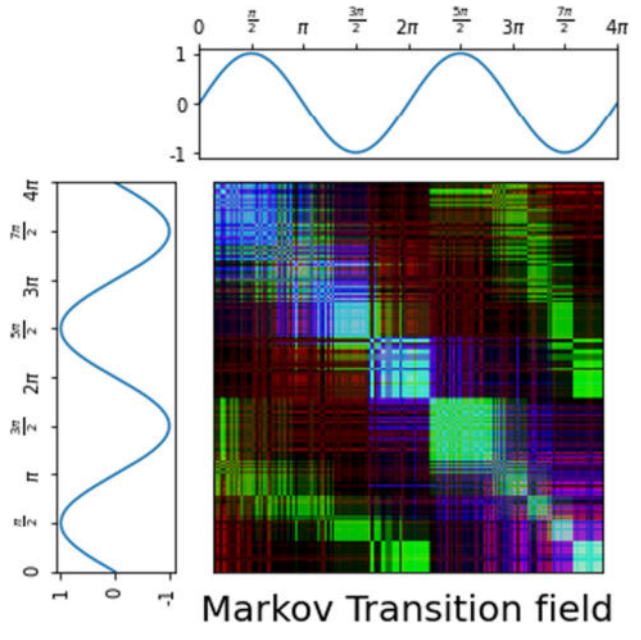

Fig. 4 Encoding results of Markov transition field

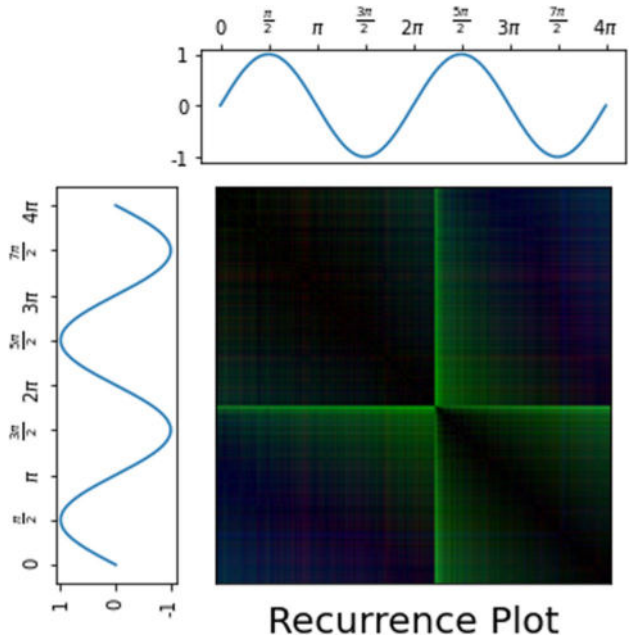

Fig. 5 Encoding results of gramian angular field

quantile. By normalizing the sum of each column of W to 1, a Markov transition matrix is constructed and W removes the time dependence on X's distribution and time index t_i. To construct MTF, each value is assigned to a quantile q_j (j âˆˆ [1, Q]) that corresponds to the time index t_i of the time series data value x_i. A weighted adjacency matrix W of size $Q \times Q$ is constructed based on the first-order Markov chain method along the time axis.

2.4.3 Recurrence Plot (RP)

RP is a method proposed by Eckmann et al. for encoding time series data to images as Fig. 5. It is a nonlinear analysis technique that explains the interactions of dynamic systems and measures time invariance [35]. RP converts sections in time series data where the same patterns are repeated into pixel values of an image and calculates the distance between the extracted sections. Based on this, a Recurrence Plot matrix is generated, which is transformed into an image using Eqs. (8) and (9), providing a visualization of the time-series data as following equations:

$$R_{i,j}(x) = \begin{cases} 1 & \text{if } ||\mathbf{x_i} - \mathbf{x_j}|| \leq \epsilon \\ 0 & \text{otherwise,} \end{cases} \tag{8}$$

$$R(i,j) = \Theta(\epsilon - ||x_i - x_j||), \tag{9}$$

where $R(i, j)$ represents the value of the i-th row and j-th column in the RP matrix, and $||x_i - x_j||$ represents the distance between time series data x_i and x_j, where ϵ is the distance threshold. As shown in Eq. (8), Θ is a threshold function that returns 1 if the distance is below the threshold and 0 otherwise.

RP is generated based on the distance matrix of time series data, allowing for the identification of interactions between closely located points in the data. The RP matrix is symmetrical along the diagonal, with the upper triangular area representing the interactions in the time series data. RP allows for the visualization of interactions in time series data, enabling the analysis of patterns in the data. This makes it possible to identify periodic patterns in the data and easily visualize the regularity of the time series data [31].

3 Proposed Methodology

The proposed method, as shown in Fig. 6, consists of several stages to detect anomalies in time series data. Firstly, the time-series data is transformed into the frequency domain using the process illustrated in Fig. 1. Subsequently, the frequency domain data is divided into specific wavelength intervals, and time-series images are generated using four different image generation algorithms. This methodology enables the integration of frequency data, including power, vibration, and noise, into a unified image format. By consolidating these data types, comprehensive analysis becomes feasible using information gathered from diverse perspectives. This integrated approach not only enhances the accuracy of anomaly detection but also substantiates the validity of defect detection processes [36]. These time-series images

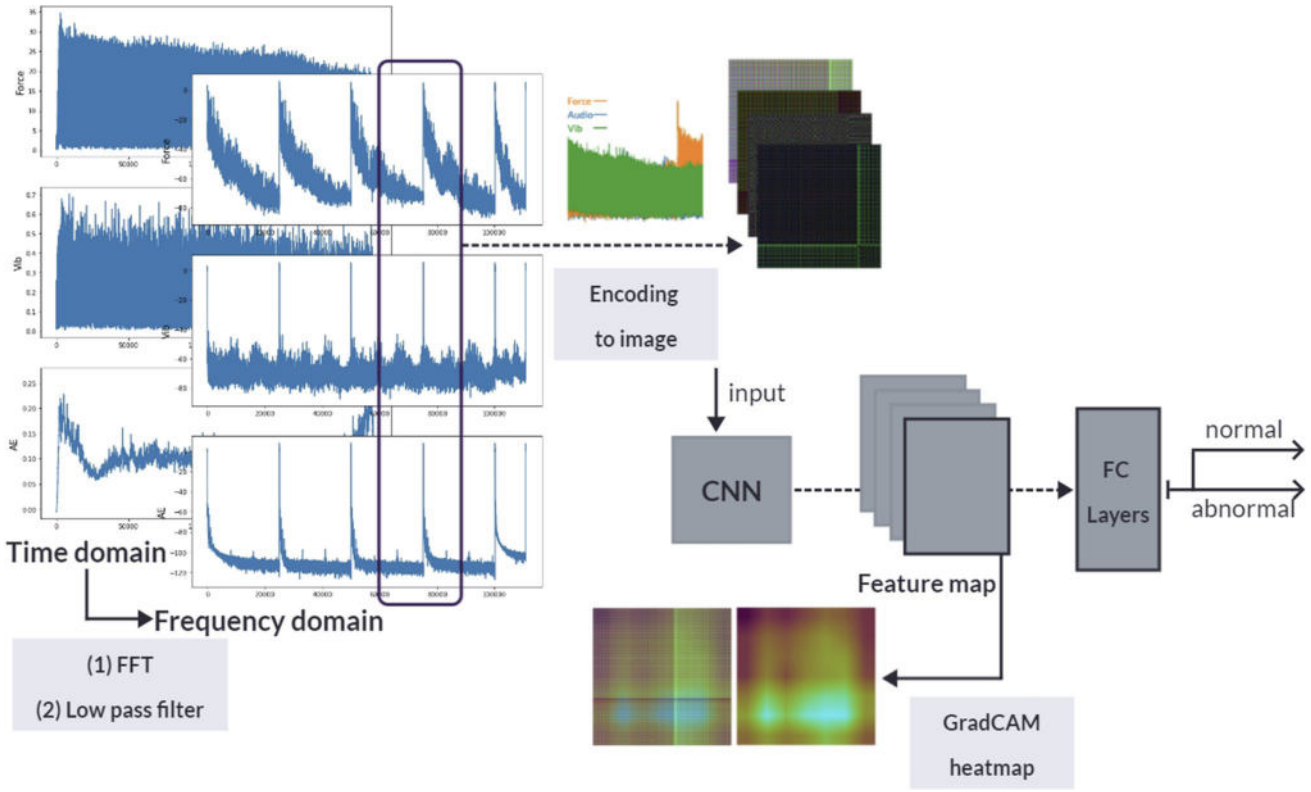

Fig. 6 Proposed process of the anomaly detection using encoding images and visualization of the activation heat-map

are used to train a CNN-based anomaly detection model that distinguishes between normal and abnormal data.

To evaluate the performance of the proposed method, we conducted two procedures. In the first procedure, we train the CNN model using time series images. In the second procedure, we explain the anomaly detection criteria by using the Grad CAM visualization results to provide model interpretation for the proposed method. The experimental results demonstrate that the proposed method achieves superior performance in detecting anomalies in time series data, and the model interpretation function using Grad CAM provides useful insights into the detection process.

Overall, our proposed method effectively detects anomalies in time series data through a well-defined process that includes transforming the data into the frequency domain, generating time-series images using various algorithms, and training a CNN model. The model interpretation function using Grad CAM enhances the transparency and interpretability of the proposed method, which could be useful in various practical applications.

3.1 Backbone Model

Table 1 presents the hyperparameters used in the proposed model. The learning rate was set to 1e-05, the batch size was set to 32, and we trained the model for 20 epochs using the Adam optimization algorithm.

Table 1 Hyperparameters used in training

Patameter values	
Learning rate	0.00001
Batch size	32
Epoch	20
Optimization function	Adam

We trained the model on each cutter's train, test, and validation datasets, using 10 different CNN network models, including VGG, ResNet, DenseNet, etc., for each image encoding method. To be specific, we used 75% of the training dataset for training and the remaining 25% for validation.

4 Numerical Results and Discussion

4.1 Demonstration case 1: 2010 PHM Data Challenge dataset

The proposed methodology uses PHM dataset, which is a set of time-series sensor data generated in various industrial processes, measured from CNC machines [37] to analyze

images for anomaly detection and prediction of devices in the manufacturing process. The PHM dataset is commonly used in the field of Prognostics and Health Management (PHM) and is applied in various fields such as health condition monitoring, predictive maintenance, fault diagnosis, reconstruction, design, and optimization [38–40].

This study used the IEEE PHM 2010 dataset for model training and validation. The CNC machine was operated with a cutter spindle speed of 10,400 RPM and a feed rate of 1555 mm/min. The experiment involved using six distinct cutters (defined as C1, C2, C3, C4, C5, C6). The data were collected at a sampling rate of 50 KHz/channel while milling the material to a depth of 0.125 mm along the Y-axis and 0.2 mm along the Z-axis [37].

Features X, Y, Z(N) and X, Y, Z(g) denote the forces and vibration values applied to each X, Y, Z axis. They are converted into Force and Vib, respectively, by normalizing the vectors by dividing their size. Converting them into unit vector data reduces the scale difference between features, enhances the model's convergence speed, and improves its performance. The resulting multivariate time-series data is then transformed into frequency data using the Fast Fourier Transform (FFT) and low-pass filters as shown in Fig. 1.

The parameters of the low-pass filter are outlined in Table 2. *fc* denotes the cutoff frequency as a fraction of the sampling rate, while *b* represents the transition bandwidth, which is also based on the sampling rate. The length of the filter, *N*, is determined by the transition bandwidth and is calculated using the formula [4/*b*]. This calculation indicates that the filter length is inversely proportional to the transition bandwidth. Additionally, our calculations require that the filter length be an odd number. In this experiment, with *fc* set at 0.1, the cutoff frequency is 10% of the frequency range. Using these parameters, the low-pass filter effectively removes higher frequencies from the data signal. The design ensures minimal signal distortion within the passband while effectively attenuating unwanted frequencies in the stopband. This characteristic is crucial for preserving the integrity of the original signal's low-frequency components.

The proposed methodology utilizes the PHM dataset, a collection of time-series sensor data obtained from various industrial processes, specifically measured from CNC machines, for the purpose of analyzing images to detect anomalies and predict device behavior in the manufacturing process.For this study, an image dataset was compiled based on the PHM dataset in Table 3 for each cutter (C1, C2, C3, C4, C5, C6) of the CNC machine, resulting in a total of approximately 6,000 images per cutter and a combined dataset of 36,000 images. Among these, around 27,000 images were labeled as normal, while about 9000 images were labeled as abnormal. The sampling rate was set to 50kHz.

For anomaly classification, we divided the normal range and the wear (anomaly) range of the cutter and labeled the time series data as normal if it stayed within the normal range for a certain period and labeled it as abnormal if it went beyond the range. Figure 7 shows the time series graphs of the labeled normal and abnormal data and the transformed image data using four methods: GASF, GADF, MFT, and RP. (a) shows the graphs and transformed image data of the normal range, and (b) shows those of the abnormal range. Unlike Fig. 7a and b shows a decreasing trend for each feature.

Using the main diagonal from the top left to the bottom right as a reference, normal images exhibit intersecting lines corresponding to increasing patterns in GAF and RP images, while abnormal images, which show relatively consistent decreases, do not have intersecting lines. In the case of MTF, bright color mappings can be observed in the increasing portions of normal images, whereas specific intervals lack mapping values in abnormal images. These observations indicate that the characteristics of time-series data can be identified through different features such as color, dots, and lines in the corresponding positions of 2D images. Through these images, the aim is to detect anomalies in a commercialized 6 mm ball-nose tungsten carbide cutter using the PHM 2010 dataset. Through this, we detect anomalies in the 6 mm ball-nose tungsten carbide cutter that was commercially used for the PHM 2010 data measurement.

4.1.1 Case 1: Training Result

Table 4 shows the validation accuracies of each network model for GADF, GASF, MTF, and RP image representations, obtained from training the proposed model in this paper. The numbers have been rounded to six decimal places. Although each image representation method showed

Table 2 Parameters of low-pass filters

Feature	Representation	Value
fc	Cutoff frequency	0.1
b	Transition band	0.08
N	Filter length	50

Table 3 Data description for the IEEE PHM 2010 dataset

Feature	Representation	Physical quantity
X(N)	Force (N) in X dimension	Force
Y(N)	Force (N) in Y dimension	Force
Z(N)	Force (N) in Z dimension	Force
X(g)	Vibration (g) in X dimension	Vib
Y(g)	Vibration (g) in Y dimension	Vib
Z(g)	Vibration (g) in Z dimension	Vib
AE(V)	AE-RMS (V)	AE

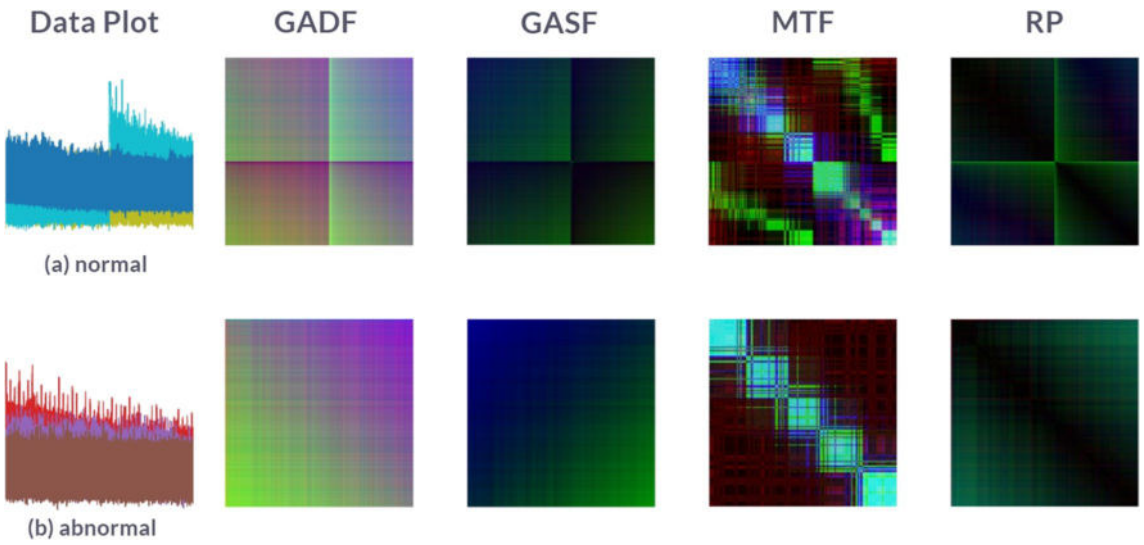

Fig. 7 Graph of normal and abnormal ranges

different performances depending on the network model used, most of the accuracies converged to a value close to 1.0, indicating relatively high accuracy in anomaly detection. The top 5 methods were RP using ResNet50 with an accuracy of approximately 99.6%, GASF using VGG16, GADF using MobileNet, GASF using ResNet50, and GADF using VGG19.

Figure 8 presents a visualization of the accuracy and loss rates for the top 5 models trained using the GADF image representation method. The loss rate indicates whether there are problems with underfitting or overfitting during training, where the loss rate does not decrease below a certain value or increases during training. As shown in Fig. 8, the loss rates of the trained models converged to zero, indicating effective learning. The validation accuracies and loss rates also suggest that overfitting was prevented, and the models achieved appropriate training.

4.2 Demonstration Case 2: Actual Data from CNC-Machining Process

Through Sect. 4.1.1. Training result, it was possible to confirm that the proposed approach in this study yielded significant results in the PHM 2010 dataset, as demonstrated above. Consequently, the applicability and significance of this methodology, utilizing real-data (Table 5) instead of the PHM 2010 dataset, can be expected. Therefore, additional data sets collected from actual field operations were obtained to further train the top five models.

The collected data consists of information obtained during milling and cutting processes in real manufacturing site. As shown in Fig. 9, it includes force and noise data, specifically recording the magnitude of forces applied along the XYZ axes. While these data exhibit similar characteristics to the PHM dataset, the data has lack of the vibration data in real manufacturing site.

4.2.1 Case 2: Training Result

Table 6 shows the comparison of accuracy among the top 5 models for Case 1: 2010 PHM Data Challenge dataset and Case 2: Actual data from CNC-machining process manufacturing field dataset. Case 1 demonstrates an accuracy close to 99%, while Case 2 exhibits a relatively lower overall accuracy of approximately 91%. This difference can be attributed to the actual data having only two time-series features, forces and noise, in comparison to the PHM dataset, which includes three features: forces, vibrations, and noise. Consequently, the limited characteristics of the actual data from CNC-machining process result in the lower accuracy.

Nonetheless, the GASF image encoding method and VGG16 model in Case 2 still achieve a significant accuracy of 91.8%. This high accuracy demonstrates the meaningful findings of this study, overcoming limitations posed by restricted experimental conditions and potential biases associated with manual label assignment. It signifies the relevance of the proposed methodology for real manufacturing site datasets and other domains involving time-series data.

4.3 Grad-CAM Result

The layer to which Grad-CAM is applied is selected after training the CNN model. The visualization of the activated regions differs depending on the feature map. This enables an understanding of which parts the model focuses on to recognize a particular class, and Grad-CAM's visualized

Table 4 Accuracy for the IEEE PHM2010 dataset

	VGG16	VGG19	ResNet50	InceptionResNetV2	Xception	MobileNet	DenseNet121	DenseNet169	DenseNet201	NASNetMobile
GADF	0.988176	**0.989865**	0.983108	0.753378	0.976351	**0.991554**	0.96875	**0.984797**	**0.980574**	**0.936655**
GASF	**0.994088**	0.976351	0.990709	0.751689	**0.969595**	0.98902	0.980574	0.961149	0.974662	0.857264
MTF	0.912162	0.905405	0.748311	**0.91723**	0.903716	0.941723	0.941723	0.94848	0.933277	0.918074
RP	0.984797	0.987331	**0.996622**	0.749155	0.965372	0.981419	**0.984797**	0.978885	0.962838	0.914696

activation map only provides useful information when the model accurately recognizes a specific class. Conversely, if the accuracy of the CNN model is not guaranteed, Grad-CAM's results visualizing the parts that are mainly activated for recognizing a particular class cannot be trusted. Therefore, efforts to improve the accuracy of the CNN model are necessary to ensure its reliability, and the model's performance is crucial in interpreting Grad-CAM results.

The selection criteria for the activation map in Grad-CAM are as follows: Grad-CAM generates an activation map using the back-propagated weights starting from the last convolution layer, which has a significant result. Therefore, in Grad-CAM, the last convolution layer is selected and an activation map that visualizes the most influential feature points for recognizing a particular class is chosen based on the score of the specific class.

Figure 10 visualizes the main areas for predicting image classes by each model based on their features through Grad-CAM. Figure 10a1 shows MobileNet of GADF, Fig. 10(b.1) shows VGG19, Fig. 10a2 shows ResNet of GASF, Fig. 10b2 shows VGG16, and finally, Fig. 10a3 shows ResNet of RP. These are the top five image transformation methods and classification models with high accuracy, as shown in Table 4.

For GADF, even for the same image, MobileNet and VGG19 show significant differences in the activation maps they visualize. The activation map of RP's ResNet50 with the highest accuracy is activated in relatively distinctive parts, indicating that it can correctly identify the image class. Therefore, Resnet50 can more clearly explain the image classes than other models.

5 Conclusion

In this study, we proposed an anomaly detection and classification process in manufacturing processes using Grad-CAM, one of the interpretable AI methodologies. Due to the recent Fourth Industrial Revolution and advances in hardware, there is an increasing demand for artificial intelligence and big data processing in the industry. The concept of predictive maintenance is receiving more attention, especially in the smart factory industry, where many processes are of interest and in progress.

5.1 Future Work and Potential Applications

This study aims to establish a foundation for DSS(Decision Support Systems) that significantly enhance predictive maintenance and operational efficiency of machinery in manufacturing environments through the use of artificial intelligence models. Literature [41] proposes a new DSS based on IoT and machine learning for predictive maintenance in

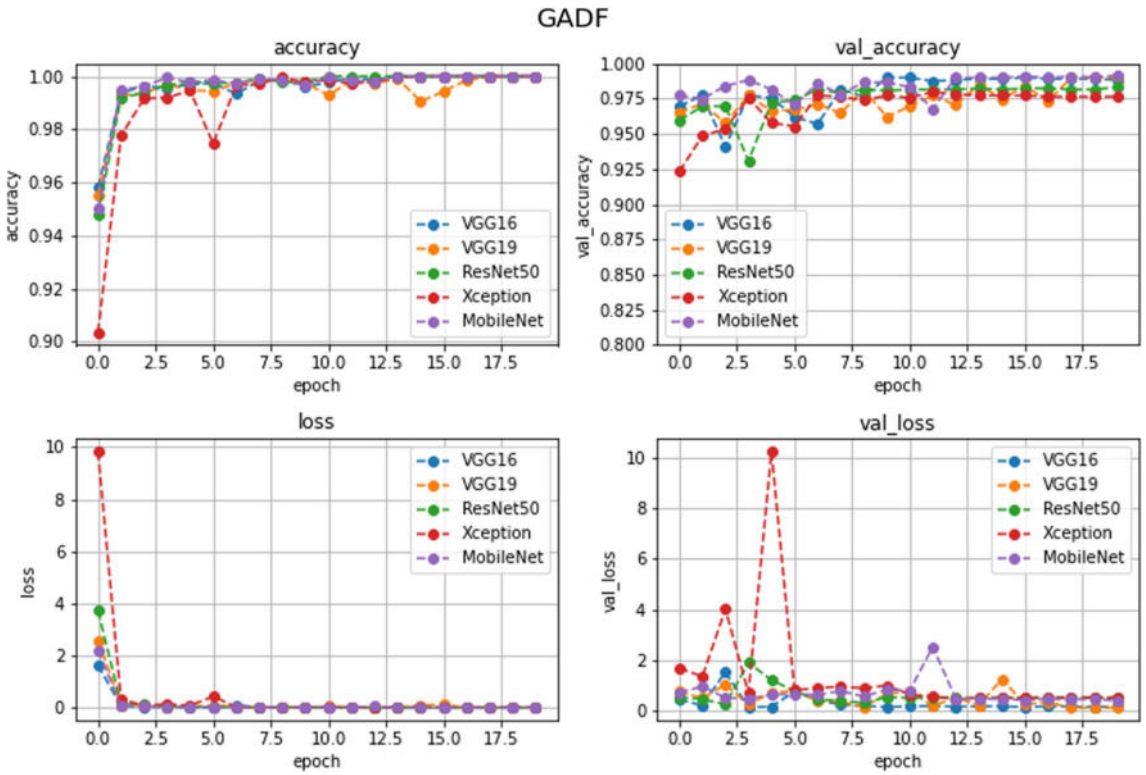

Fig. 8 GADF accuracy and loss

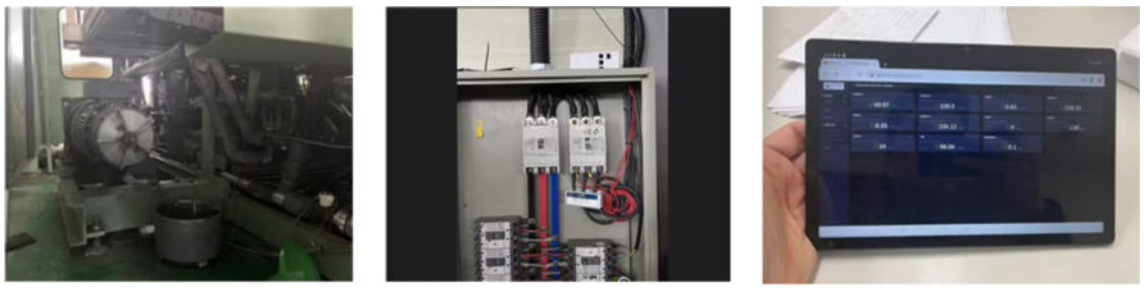

Fig. 9 Case2: CNC- machining process and IoT device of the manufacturing field data

Table 5 Data description from the actual data for CNC-machining process

Feature	Representation	Physical quantity
X(N)	Force (N) in X dimension	Force
Y(N)	Force (N) in Y dimension	Force
Z(N)	Force (N) in Z dimension	Force
AE(V)	AE-RMS (V)	AE

Table 6 Comparison of the accuracy for the case 1 and 2

		2010 PHM Data Challenge dataset	Actual data from CNC-machining process
RP	ResNet50	**0.996622**	0.915207
GASF	VGG16	0.994088	**0.918238**
	ResNet50	0.990709	0.915235
GADF	MobileNet	0.991554	0.915743
	VGG19	0.989865	0.915545

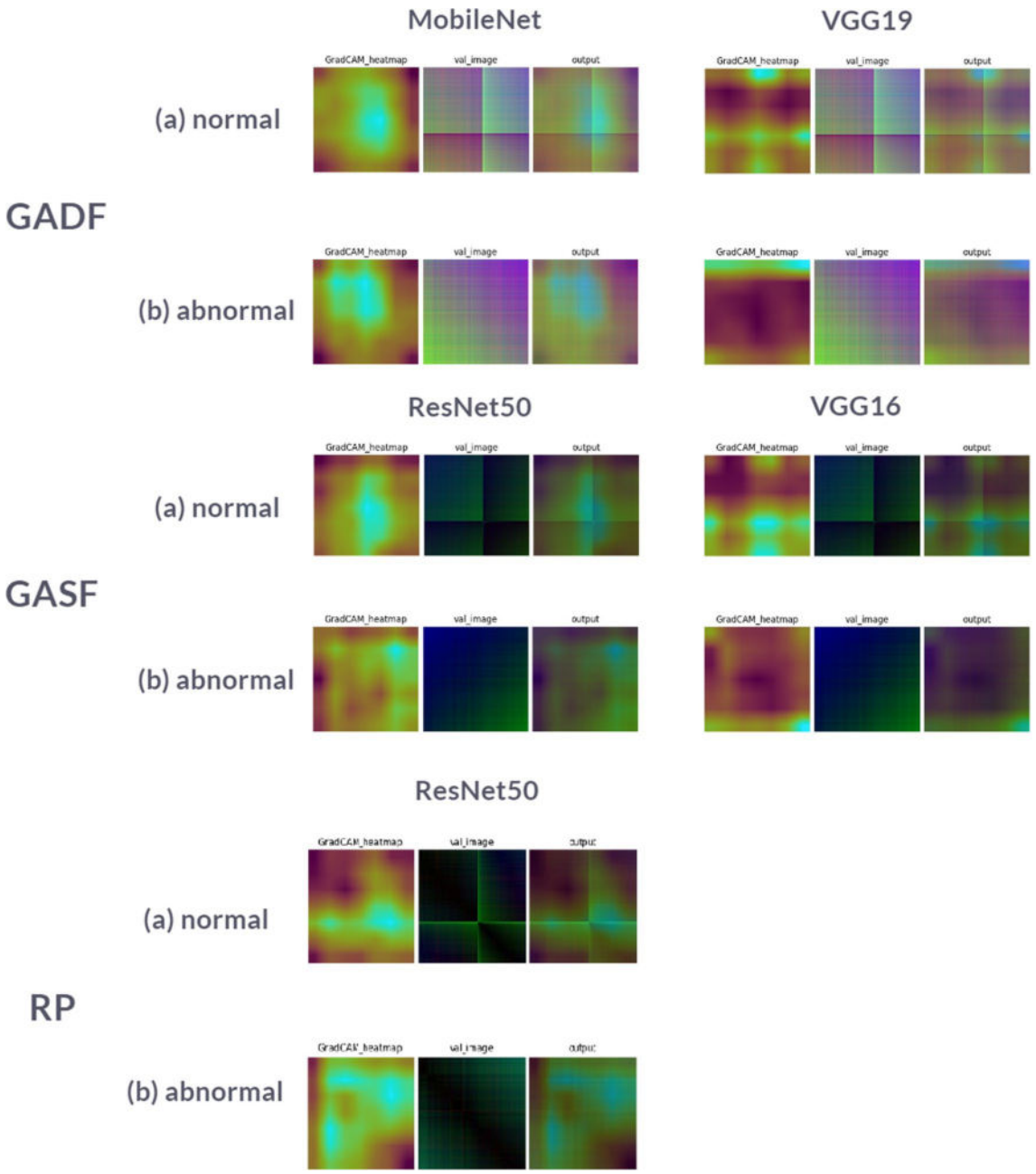

Fig. 10 Visualizing the activation map of a model using Grad-CAM

an Industry 4.0 environment. This system combines real-time data collected from IoT devices with machine learning algorithms, enabling maintenance personnel and operators to optimize processing quality and predict potential equipment failures, thereby effectively planning maintenance. As a result, manufacturers can reduce costs, enhance productivity, and minimize equipment downtime. Moreover, this technology plays a crucial role in Industry 4.0 by optimizing maintenance schedules and providing real-time alerts about operational risks.

Future research will further develop these AI standard models to explore their applicability in industrial fields. The integration of more sophisticated sensor technologies will improve the accuracy and speed of data collection and monitoring. Additionally, advancements in machine learning and deep learning technologies are expected to enhance the predictive capabilities of the models and more accurately identify potential issues. Additionally, the application of interpretable AI in this study enhances the model's reliability and understanding by providing workers with additional

visual information that increases interpretability. By utilizing this to support decision-making in manufacturing, this research contributes to the paradigm of ZDM (Zero Defect Manufacturing) [42]. As a result, the decision support provided by explainable AI in this study will play a crucial role in achieving zero-defect manufacturing in the industry. Therefore, our research is expected to contribute to solving real-world problems and bringing innovation to the industrial sector.

5.2 Limitations of the Proposed Method

In this paper, we utilized the PHM 2010 Dataset and actual data from CNC-machining process for predictive maintenance and conducted anomaly detection through image-based time series data analysis using a convolutional neural network (CNN), which showed a high accuracy rate of 90% for most cases, surpassing previous time-series data analysis. We selected the top five models and image visualization methods among the trained CNN models and performed a comparison and analysis of the activation regions using Grad-CAM .[43] Our research presents promising advancements for predictive maintenance in the manufacturing sector, however, there are several limitations that require further discussion.

Firstly, we acknowledge that the proposed approach demands a substantial amount of labeled data for the effective training of the CNN models. The requirement for extensive data collection and labeling, which is often time-consuming and costly, poses a significant challenge, particularly in diverse manufacturing settings. Additionally, a notable limitation of our study is that the dataset used was confined to a single type of machine. This restriction diminishes the broader applicability of our findings, as the model trained on this dataset may not perform adequately with data from other types of machinery or equipment, thus limiting its usability across various industrial domains. These issues underscore the imperative for ongoing research to address these challenges and to enhance the scalability and adaptability of our approach in varied manufacturing environments as [44].

Author Contributions Conceptualization: Youngjun Yoo. and Wooju Kim.; methodology: Youngjun Yoo. and Young-Joo, Hyun.; Software: Young-Joo, Hyun, Validation, Young-Joo, Hyun. and Youngjun Yoo.; Formal analysis: Young-Joo, Hyun, Taeheon Lee and Yoonseok Kim.; Investigation: Young-Joo, Hyun; Data curation: Yoonseok Kim and Taeheon; Writing-original draft preparation: Young-Joo, Hyun, Youngjun Yoo and Wooju Kim; Visualization: Youngjun Yoo; Supervision, Youngjun Yoo. and Wooju Kim; Project administration: Youngjun Yoo; Funding acquisition: Youngjun Yoo. All authors have read and agreed to the published version of the manuscript.

Funding This study has been funded with the support of the Ministry of SMEs and Startups as "Development of Intelligent SHWIS (AI-Smart Human Work Interactive Interface System) AR technology that provides AR Inspection" - RS-2022-00140809.

Availability of Data and Materials Not applicable.

Declarations

Ethics Approval and Consent to Participate No participation of humans takes place in this implementation process.

Competing interests The author declare that they have no competing interests.

References

1. Md, A. Q., Jha, K., Haneef, S., Sivaraman, A. K., & Tee, K. F. (2022). A review on data-driven quality prediction in the production process with machine learning for industry 4.0. *Processes, 10*, 1966. https://doi.org/10.3390/pr10101966
2. Kang, J., Lim, C., Maeng, H., & Park, K. (2024). Diagnosis of high-speed ball-bearing spindles by data mining of dynamic responses from various rotating elements. *International Journal of Precision Engineering and Manufacturing*. https://doi.org/10.1007/s12541-024-01007-6
3. Tagawa, Y., Maskeliūnas, R., & Damaševičius, R. (2021). Acoustic anomaly detection of mechanical failures in noisy real-life factory environments. *Electronics, 10*(19), 2329. https://doi.org/10.3390/electronics10192329
4. Amruthnath, N., & Gupta, T. (2018). A review on data-driven quality prediction in the production process with machine learning for industry 4.0. In *2018 5th international conference on industrial engineering and applications (ICIEA), IEEE*, pp. 355–361. https://doi.org/10.1109/IEA.2018.8387124.
5. Tee, K. F. (2018). Time series analysis for vibration-based structural health monitoring: A review. *Structural Durability and Health Monitoring, 12*(3), 129–147. https://doi.org/10.3970/sdhm.2018.04316
6. Blázquez-García, A., Conde, A., Mori, U., & Lozano, J. A. (2021). A review on outlier/anomaly detection in time series data. *ACM Computing Surveys (CSUR), 54*(3), 1–33. https://doi.org/10.1145/3444690
7. Kim, H., & Shon, T. (2022). Industrial network-based behavioral anomaly detection in AI-enabled smart manufacturing. *The Journal of Supercomputing, 78*(11), 13554–13563. https://doi.org/10.1007/s11227-022-04408-4
8. Chen, Z., Liu, Y., & Liu, S. (2017). Mechanical state prediction based on LSTM neural netwok. *In 2017 36th Chinese control conference (CCC), IEEE*, pp. 3876–3881. https://doi.org/10.23919/ChiCC.2017.8027963.
9. Mahalakshmi, G., Sridevi, S., & Rajaram, S. (2016). A survey on forecasting of time series data. In *2016 International conference on computing technologies and intelligent data engineering (ICCTIDE'16), IEEE*, pp. 1–8. https://doi.org/10.1109/ICCTIDE.2016.7725358.
10. Yoo, Y., & Jeong, S. (2022). Vibration analysis process based on spectrogram using gradient class activation map with selection process of CNN model and feature layer. *Displays, 73*, 102233. https://doi.org/10.1016/j.displa.2022.102233
11. Wang, Z., & Oates, T. (2015). Imaging time-series to improve classification and imputation. arXiv preprint , arXiv:1506.00327
12. Liao, Y., Qing, X., Wang, Y., & Zhang, F. (2023). Damage localization for composite structure using guided wave signals with Gramian angular field image coding and convolutional neural

networks. *Composite Structures*. https://doi.org/10.1016/j.compstruct.2023.116871
13. Yang, C. L., Chen, Z. X., & Yang, C. Y. (2019). Sensor classification using convolutional neural network by encoding multivariate time series as two-dimensional colored images. *Sensors, 20*(1), 168. https://doi.org/10.3390/s20010168
14. Deshpande, A. M., Minai, A. A., & Kumar, M. (2020). One-shot recognition of manufacturing defects in steel surfaces. *Procedia Manufacturing, 48*, 1064–1071. https://doi.org/10.1016/j.promfg.2020.05.146
15. Shin, S. J., Lee, J. H., Jadhav, S., & Kim, D. B. (2024). Material-adaptive anomaly detection using property-concatenated transfer learning in wire arc additive manufacturing. *International Journal of Precision Engineering and Manufacturing, 25*(2), 383–408. https://doi.org/10.1007/s12541-023-00924-2
16. Seo, J., Kim, W., & Lee, J. (2024). Unsupervised manufacturing fault detection based on self-labeled training of fingerprint image constructed from time-series data. *International Journal of Precision Engineering and Manufacturing*. https://doi.org/10.1007/s12541-023-00947-9
17. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision,* pp. 618–626. https://doi.org/10.1109/ICCV.2017.74
18. Kim, S. W., Kong, J. H., Lee, S. W., & Lee, S. (2022). Recent advances of artificial intelligence in manufacturing industrial sectors: A review. *International Journal of Precision Engineering and Manufacturing*. https://doi.org/10.1007/s12541-021-00600-3
19. Brigham, E. O., & Morrow, R. E. (1967). The fast Fourier transform. *IEEE Spectrum, 4*(12), 63–70. https://doi.org/10.1109/MSPEC.1967.5217220
20. Briggs, W. L., & Henson, V. E. (1995). The DFT: An owner's manual for the discrete Fourier transform. *Society for Industrial and Applied Mathematics*. https://doi.org/10.1137/1.9781611971514
21. Ding, C., Wang, Z., Ding, Q., & Yuan, Z. (2022). Convolutional neural network based on fast Fourier transform and Gramian angle field for fault identification of HVDC transmission line. *Sustainable Energy, Grids and Networks, 32*, 100888. https://doi.org/10.1016/j.segan.2022.100888
22. Christiano, L. J., & Fitzgerald, T. J. (2003). The band pass filter. *International Economic Review, 44*(2), 435–465. https://doi.org/10.1111/1468-2354.t01-1-00076
23. Ahn, D., Park, J. S., Kim, C. S., Kim, J., Qian, Y., & Itoh, T. (2001). A design of the low-pass filter using the novel microstrip defected ground structure. *IEEE Transactions on Microwave Theory and Techniques, 49*(1), 86–93. https://doi.org/10.1109/22.899965
24. Dogra, A., & Bhalla, P. (2014). Image sharpening by gaussian and butterworth high pass filter. *Biomedical & Pharmacology Journal, 7*(2), 707–713. https://doi.org/10.13005/bpj/545
25. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84–90. https://doi.org/10.1145/3065386
26. Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI-Explainable artificial intelligence. *Science Robotics, 4*(37), eaay7120. https://doi.org/10.1126/scirobotics.aay7120
27. Cui, J., Zhong, Q., Zheng, S., Peng, L., & Wen, J. (2022). A lightweight model for bearing fault diagnosis based on Gramian angular field and coordinate attention. *Machines, 10*(4), 282. https://doi.org/10.3390/machines10040282
28. Hong, S., Yoon, J., Ham, Y., Lee, B., & Kim, H. (2023). Monitoring safety behaviors of scaffolding workers using Gramian angular field convolution neural network based on IMU sensing data. *Automation in Construction, 148*, 104748. https://doi.org/10.1016/j.autcon.2023.104748
29. Bugueno, M., Molina, G., Mena, F., Olivares, P., & Araya, M. (2021). Harnessing the power of CNNs for unevenly-sampled light-curves using Markov Transition Field. *Astronomy and Computing, 35*, 100461. https://doi.org/10.1016/j.ascom.2021.100461
30. Zhang, R., Zheng, F., & Min, W. (2018). Sequential behavioral data processing using deep learning and the Markov transition field in online fraud detection. arXiv preprint arXiv:1808.05329.
31. Chen, Y., Su, S., & Yang, H. (2020). Convolutional neural network analysis of recurrence plots for anomaly detection. *International Journal of Bifurcation and Chaos, 30*(01), 2050002. https://doi.org/10.1142/S0218127420500029
32. Westerhold, T., Marwan, N., Drury, A. J., Liebrand, D., Agnini, C., Anagnostou, E., & Zachos, J. C. (2020). An astronomically dated record of Earth's climate and its predictability over the last 66 million years. *Science, 369*(6509), 1383–1387. https://doi.org/10.1126/science.aba6853
33. Bendjenna, H., Meraoumia, A., & Chergui, O. (2018). Pattern recognition system: From classical methods to deep learning techniques. *Journal of Electronic Imaging, 27*(3), 033008–033008. https://doi.org/10.1117/1.JEI.27.3.033008
34. Wang, Z., & Oates, T. (2015). *Imaging time-series to improve classification and imputation.* arXiv preprint arXiv:1506.00327.
35. Eckmann, J. P., Kamphorst, S. O., & Ruelle, D. (1995). Recurrence plots of dynamical systems. *World Scientific Series on Nonlinear Science Series A, 16*, 441–446. https://doi.org/10.1209/0295-5075/4/9/004
36. Tian, G., Yang, C., Lu, X., Wang, Z., Liang, Z., & Li, X. (2023). Inductance-to-digital converters (LDC) based integrative multi-parameter eddy current testing sensors for NDT &E. *NDT & E International*. https://doi.org/10.1016/j.ndteint.2023.102888
37. Li, X., S. Lim, B., H. Zhou, J., Huang, S., J. Phua, S., C. Shaw, K., & J. Er, M. (2021). Fuzzy Neural Network Modelling for Tool Wear Estimation in Dry Milling Operation. In *Annual conference of the PHM society, 1*(1). http://papers.phmsociety.org/index.php/phmconf/article/view/1403
38. Cai, W., Zhang, W., Hu, X., & Liu, Y. (2020). A hybrid information model based on long short-term memory network for tool condition monitoring. *Journal of Intelligent Manufacturing, 31*, 1497–1510. https://doi.org/10.1007/s10845-019-01526-4
39. Sun, H., Zhang, J., Mo, R., & Zhang, X. (2020). In-process tool condition forecasting based on a deep learning method. *Robotics and Computer-Integrated Manufacturing, 64*, 101924. https://doi.org/10.1016/j.rcim.2019.101924
40. Nor, A. K. M., Pedapati, S. R., Muhammad, M., & Leiva, V. (2022). Abnormality detection and failure prediction using explainable Bayesian deep learning: Methodology and case study with industrial data. *Mathematics, 10*(4), 554. https://doi.org/10.3390/math10040554
41. Rosati, R., Romeo, L., Cecchini, G., Tonetto, F., Viti, P., Mancini, A., & Frontoni, E. (2023). From knowledge-based to big data analytic model: A novel IoT and machine learning based decision support system for predictive maintenance in Industry 40. *Journal of Intelligent Manufacturing, 34*(1), 107–121. https://doi.org/10.1007/s10845-022-01960-x
42. Caiazzo, B., Di Nardo, M., Murino, T., Petrillo, A., Piccirillo, G., & Santini, S. (2022). Towards Zero Defect Manufacturing paradigm: A review of the state-of-the-art methods and open challenges. *Computers in Industry, 134*, 103548. https://doi.org/10.1016/j.compind.2021.103548
43. Kim, J., & Kim, J. M. (2020). Bearing fault diagnosis using grad-CAM and acoustic emission signals. *Applied Sciences, 10*(6), 2050. https://doi.org/10.3390/app10062050

44. Xia, C., Pan, Z., Fei, Z., Zhang, S., & Li, H. (2020). Vision based defects detection for Keyhole TIG welding using deep learning with visual explanation. *Journal of Manufacturing Processes, 56*, 845–855. https://doi.org/10.1016/j.jmapro.2020.05.033

Young-Joo Hyun She is a M.S. in the Department of Industrial Engineering at Yonsei University, and researcher in the Smart Manufacturing R&D Department at Korea Institute of Industrial Technology (KITECH). Her research interests include anomaly detection in computer vision and machine learning.

Youngjun Yoo He is a senior researcher in the Smart Manufacturing R&D Department at Korea Institute of Industrial Technology (KITECH). Before the KITECH, he worked Samsung Heavy Industries for smart ships. He got my Ph. D. degree at Pohang University of Science and Technology. His research subjects are inspection equipment using AI and IoT equipment, data analysis, and smart manufacturing.

Yoonseok Kim He is a M.S. in the Department of Industrial Engineering at Yonsei University, and researcher in the Smart Manufacturing R&D Department at Korea Institute of Industrial Technology (KITECH). His research interests include robotics engineering, control engineering, wireless communication and mobile services. Technically, He is interested in machine learning, reinforcement learning, optimization.

Taeheon Lee He is a M.S. in the Department of Industrial Engineering at Yonsei University, and researcher in the Smart Manufacturing R&D Department at Korea Institute of Industrial Technology (KITECH). His research interests include manufacturing and user behavior analytics. He is interested in machine learning and performance analysis.

Wooju Kim He is a professor in the Department of Industrial Engineering at Yonsei University. He received the Ph.D. degree in operations research from KAIST, South Korea, in 1994. His main research interests include natural language processing, anomaly detection, reliable knowledge discovery, big data intelligence, machine learning, and artificial intelligence.