iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.62493

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

CNN-Based Machine Tool Monitoring with STFT Image Analysis

Vishnu Priya Modupalli[1], Asha K N[2], Lavanya Santhosh[3], Naik Mehul Manjunath[4], Pavan N R[5], Pavithra S[6]
[1, 4, 5, 6]*Student, Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka*
[2, 3]*Assistant Professor, Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bengaluru, Karnataka*

Abstract**: In the realm of machine tool operation, effective condition monitoring holds predominant importance to ensure operational reliability and safety. Leveraging deep learning methodologies, particularly convolutional neural networks (CNN), for defect identification has gained significant attention. However, inherent challenges persist, including the extraction of salient features and potential information loss while extracting the features from raw vibration signals. In response, this study proposes an intelligent approach for condition monitoring in machine tools, integrating short-time Fourier transform and convolutional neural networks (STFT-CNN). The process entails using STFT to convert one-dimensional vibration signals into time-frequency pictures, which are then fed into the STFT-CNN model to acquire and identify fault features. Furthermore, the study explores optimizing STFT parameters, such as window type, window width, and translation overlap width, across various typical window functions to improve the effectiveness of the transformation process. Within the STFT-CNN model, the utilization of stacked double convolutional layers aims to augment the model's nonlinear expression capacity, thereby facilitating robust fault diagnosis capabilities in machine tool condition monitoring applications.
Keywords: Condition Monitoring, Machine Tools, CNN, STFT, Fault Detection.

I. INTRODUCTION

It's crucial to keep an eye on the condition of rotating machine elements. Considering that any issues with these components could be quite harmful. Such as a machine shutdown that would affect the entire production system. And higher maintenance costs[1]. Modern production equipment is getting more specialized and complex. This raises the cost of machine hours and subsequent investment expenses. Furthermore, these days machines are used in just-in-time and manufacturing lines. This indicates that a workpiece generated by a machine is immediately subjected to additional processing steps within the production chain. As a result, it is preferable to minimize the amount of equipment downtime brought on by unplanned breakdowns and maintenance intervals. A machine's entire line or even another business that depends on the produced workpiece may stop operating if one of its parts fails. Finding these flaws early on is therefore crucial to reducing the expenses associated with a failure[2]. Numerous sectors have used various strategies to enhance the dependability, accessibility, and security of contemporary industrial systems and applications, which are essential for operations. In order to cut off maintenance expenses[3]. Monitoring the tool conditions and detecting the irregularities that can occur during machining to avoid hazardous conditions is vital[4]. As a result. In the industrial sector, condition-based monitoring has become increasingly important. Conversely, though. Early maintenance choices are made through the use of condition monitoring and data collection.

Recently, there has been a lot of focus on rotating machinery (RM), condition monitoring (CM), and fault detection and diagnosis (FDD)[3]. Equipment used in industrial manufacturing frequently uses bearings. The proper functioning of bearings, which are the most prevalent parts of rotating machinery, has a direct impact on the effectiveness and safety of mechanical equipment[3]. Therefore, it is critical to do an early failure diagnostic on spinning machine parts. Deep learning techniques are the foundation of the majority of the present research on data-driven fault diagnosis, within the last ten years. In the realm of diagnosing machinery faults, CNN has been applied extensively and has shown positive outcomes. In general, four steps are involved in using CNN for fault diagnosis: data collection, model building, feature learning, and decision making.

A. Existing System

The cutting-edge approaches in monitoring conditions for machine tools are mostly employing deep learning techniques specifically one-dimensional (1D) CNN for fault diagnosis. These methods have produced good results in a number of studies, including the identification of structural deterioration and bearing issues. Even though they have been successful, there are significant limitations.

Firstly, 1D CNNs were originally designed for processing 2D images, hence their effectiveness in handling 1D signals, such as time-domain signals, is not fully taken advantage of. Secondly, applying 1D CNNs directly to time-domain signals often leads to the loss of critical fault feature information, resulting in inaccurate fault diagnosis.

B. Proposed System

We suggest a unique method that uses the STFT to convert 1D signals into 2D time-frequency pictures in order to solve the inadequacies of the current system. By preserving both time and frequency domain information, this transformation makes it possible to depict more thorough fault features. When compared to raw time-domain data, time-frequency pictures have shown better defect diagnosis performance and are known to be more resilient in noisy conditions. Our proposed system, named STFT-CNN, integrates this transformation process with CNN for more accurate and efficient fault diagnosis. By feeding the generated time-frequency images into the CNN model, we aim to achieve improved fault detection with fewer learnable parameters, thus enhancing the overall effectiveness of condition monitoring for machine tools.

II. LITERATURE SURVEY

One-dimensional (1D) CNNs were effectively used for bearing problem diagnosis and detection because of their excellent efficacy in vibration signal processing [5, 6]. A 1D CNN-bearing defect diagnostic model working on time domain signals was studied by Zhang et al. [7]. Abdeljaber et al. [8] fed raw signals of time into a 1D CNN and applied it to real-time structural damage detection in bleachers. Su et al. [9] proposed ResNet to directly process the raw signal of time domain for fault diagnosis of a high-speed train bogie. Wang et al. [10] proposed a multi-attention one-dimensional convolutional neural network (MA1DCNN) to diagnose wheelset-bearing faults. Fast Fourier transform (FFT) was used by Zhao et al. [11] to convert 1D time domain signals into frequency domain images, which were then input into models for defect diagnosis such as BiLSTM, LeNet, AlexNet, ResNet18, and others. The discrete Fourier transform (DFT) was utilized by Janssens et al. [12] to convert signals from the time domain into the frequency domain, which was then fed into a CNN for problem identification.

Even with its use in fault diagnosis, the 1D CNN model still has the following shortcomings.

1) Given that CNN was first created to address the learning challenges associated with two-dimensional (2D) images, its benefits cannot be completely realized when 1D signals are used as the input.
2) The valuable fault feature information is lost when the time domain signal is processed directly using 1D CNN. The precise fault characteristics are outside the scope of the 1D CNN model.

Two-dimensional images are far more effective and efficient in diagnosing faults since they frequently carry a lot of fault information. Deep learning is capable of automatically extracting features from the pictures that describe the kind of deep-level bearing faults. To identify bearing fault states, a 2D shape conversion of the 1D vibration data is followed by image classification Deep learning is capable of automatically extracting features from the pictures that describe the kind of deep-level bearing faults. n[13]. In order to provide the model with statistical variables derived from vibration data, Bhadane et al. [14] constructed a 2D CNN for the purpose of classifying bearing defects. Hoang et al. [15] converted the original time domain signals into 2D gray-scale images based on the time series as input to CNN for fault diagnosis. Wang et al. [16] used FFT to segment the 1D raw signals, turn them into frequency domain signals, and then create 2D images from the frequency domain signals. Ultimately, the enhanced LeNet-5 model, which was trained on the 2D images, was able to quickly assess the bearing's reliability and project how long it would last. In order to diagnose faults, Wen et al. [17] suggested converting the original time domain signals into 2D grayscale images, which would then be entered into an upgraded LeNet-5 model.

In contrast to the 2D transformations found in the aforementioned literature, the STFT can be used to transform 1D signals and produce 2D time-frequency pictures. In addition to having more fault information, the time-frequency pictures also have information in the frequency and time domains. Compared with time series signals, time-frequency images are much easier to extract information in noisy environments, increasing the overall efficiency. Time-frequency domain inputs are notably superior to time-domain inputs, as has been shown in the study of defect diagnostics. The widespread usage of STFT in rotating machinery defect diagnosis highlights the technology's significance in real-world applications. Therefore, the time-frequency images are fed into the proposed CNN model for fault diagnosis, leading to better results achieved with significantly fewer learnable parameters. We used the STFT to generate 2D images from 1D signals, followed by fault diagnosis using a CNN. And furthermore, we construct a new network for bearing fault diagnosis based on STFT and CNN. The application of this combined approach shows promising results in real-world fault detection scenarios.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

III. THEORETICAL FUNDAMENTALS

A. Convolutional Neural Network

Traditional CNN is used in computer vision and is very good at extracting feature information from images. A CNN is a deep learning technique that is particularly well-suited for the examination of visual data. The layers that make up a CNN are often categorized into 3: Convolutional Layers, Pooling Layers, and Fully Connected Layers. The CNN's complexity rises as data moves through these layers, enabling it to detect progressively more abstract characteristics and greater areas of a picture. Figure 1 represents the general CNN structure.

The convolution function is given as follows:

$$H_j^{l+1} = \sum_{i \in x_j} H_j^l * w_{ij}^{l+1} + b_j^{l+1} \quad (1)$$

where H_j^{l+1} denotes the jth feature map of the neuron at layer $l+1$, * denotes the convolution function, w_{ij}^{l+1} denotes the convolution kernel connecting the jth feature map of the neuron at layer l+1 and the ith feature map of the neuron at layer l, b_j^{l+1} denotes the bias, and x_j denotes the image of the input CNN.

There is a linear process in the convolution layer. A nonlinear activation function is introduced to the model to improve its classification performance. The Sigmoid function, Tanh function and ReLU function, which are frequently used are defined as follows:

$$f_{sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

$$f_{Tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3)$$

$$f_{ReLU}(x) = \begin{cases} x, x \geq 0 \\ 0, x < 0 \end{cases} = \max(0, x) \quad (4)$$

Pooling layers take part in reducing the network parameters, and it is given as follows:

$$H_j^{l+1} = \mathrm{f}(\beta_j^{l+1}\, down\, (H_j^l) + b_j^{l+1}) \quad (5)$$

where down(.) denotes a subsampling function, β denotes the multiplicative bias.

The two most popular pooling techniques are average and maximal pooling. While average pooling averages the window values and outputs them, maximum pooling produces the window's maximum value. We employ the greatest pooling layer in this study. Figure 2 displays a schematic of the Maximum pooling technique. In the example, the step size is two and the convolution kernel is two by two in size.

The feature data that was previously extracted is classified using the fully connected layer; this process is represented as follows:

$$y^k = f(w^k x^{k-1} + b^k) \quad (6)$$

where k is the *k-th* layer network, x^{k-1} is the input of the *(k-1)-th* fully connected layer, the y^k is the output of the *k-th* fully connected layer, w^k is the weight coefficient, b^k is the bias, and *f* is the classification function.

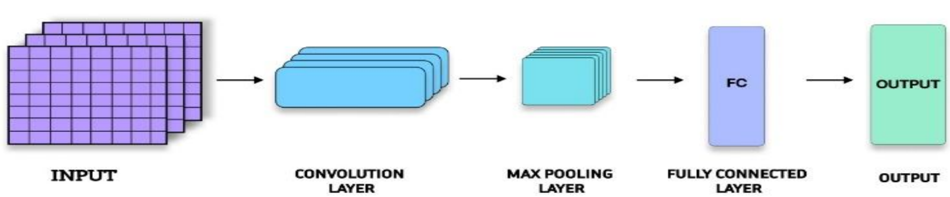


Fig.1 General CNN Structure

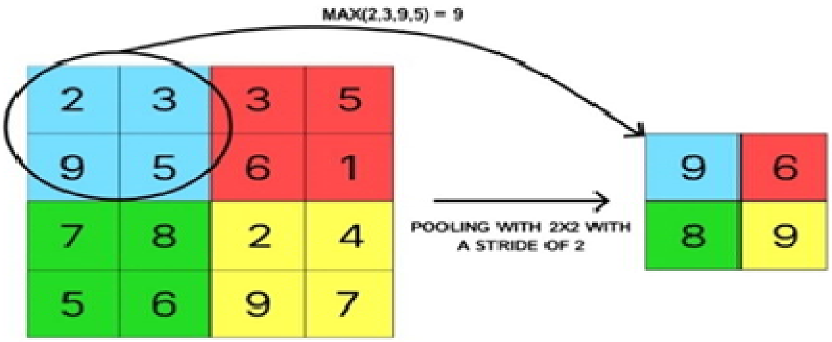


Fig.2 Maximum Pooling Method

B. *Time-frequency Analysis*

The time-frequency analysis technique is extensively employed in signal detection, equipment defect diagnostics, speech processing, and condition detection. In the 1940s, time-frequency analysis technology research was underway. Typical techniques for time-frequency analysis include S-Transform, STFT, and continuous wavelet transform (CWT). Time-frequency analysis displays the joint time-frequency properties of the signal by mapping 1D time domain signals to 2D time-frequency planes.

Suppose the wavelet mother function is *(t)* , and the scale factor and translation factor are *a* and *b*, respectively. The wavelet mother function is scaled and translated to obtain the subfunction, whose equation is shown as follows:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) a, b \in R, a > 0 \tag{7}$$

It is known as the continuous wavelet function basis because the scale factor and translation factor are continuously altered. The continuous wavelet transform is the expansion of the continuous signal *s(t)* in the wavelet basis. Here is how the CWT is defined:

$$WT(a,b) = \frac{1}{\sqrt{a}}\int_R s(t)\psi^*\left(\frac{t-b}{a}\right)dt \tag{8}$$

The S-Transform is defined as follows:

$$S(\tau,f) = \int_{-\infty}^{+\infty} s(t)\frac{|f|}{\sqrt{2\pi}}e^{-\frac{(t-\tau)^2 f^2}{2}} e^{-j2\pi f t}\,dt \tag{9}$$

where t is the time, f is the frequency, j is an imaginary unit, and s is the center of the Gaussian window function. In 1946, Dennis Gabor proposed the STFT. Assuming that the signal is smooth in a little amount of time is the fundamental idea. The signal is split up into manageable chunks using a window function, and STFT is applied to each chunk. After that, join all of the spectral analyses to create a time-frequency image. This can be stated as the operation process:

$$STFT(t,\omega) = \int_{-\infty}^{+\infty} s(\tau)g(\tau-t)e^{-j\omega t}\,d\tau \tag{10}$$

Its spectrum can be computed as follows:

$$|STFT(t,\omega)|^2 = \left|\int_{-\omega}^{\omega} s(\tau)g(\tau-t)e^{-j\omega\tau}\,d\tau\right|^2 \tag{11}$$

where *s(t)* denotes the signal, *g(t)* denotes the window function, *t* and *τ* denote the moment, *g(τ-t)* denotes the window function whose center is located at moment *t*, and *ω* denotes the frequency.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

IV. METHODOLOGY

A. Procedures of the proposed method

We developed an approach for fault diagnostics based on the previously mentioned theoretical foundations. The suggested fault diagnosis system based on CNN and STFT is flow-diagrammed in Figure 3. It is evident that the ideal STFT is sampling 1D vibration signals and conducting the following operations to create time-frequency pictures. The 2D CNN is then used to classify and identify the faults in the images. The following figure illustrates the specifics of the fault diagnosis process.

1) *Data preprocessing Stage:* Sensors are collecting the bearing vibration signals, and these vibration signals are being divided into sample sequences in order. Next, the optimal STFT is applied to convert the sample sequences into time-frequency pictures. To speed up the data processing, the time-frequency image data are being normalized.
2) *Model Training Stage:* The training set samples are inputted into the designed 2D-CNN model. The trained model can be obtained by continuously updating the weights iteratively to minimize as well as stabilize the loss function.
3) *Fault Diagnosis Stage:* The testing set samples are inputted to the trained model to obtain fault diagnosis results.

B. Details of the CNN model

The suggested CNN, which has four fully connected layers (FC), two maximum pooling layers (MP), one flatten layer, and five convolutional layers (C), is depicted in Figure 3. The original signals are converted into images and fed into the proposed CNN model to classify the images. In this work, the suggested CNN model is used to complete the fault diagnosis task.

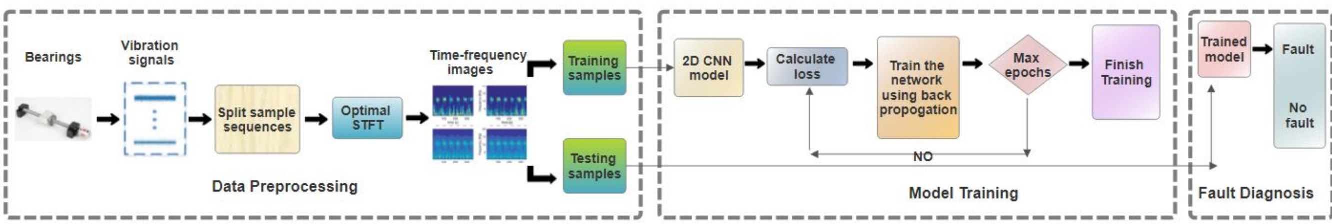


Fig 3. Flow Diagram of the Proposed Model

Table 1 displays the specific structural parameters for every layer in the CNN model. There are four components to the model. The first part consists of 32 convolutional kernels of size 5 ×5 followed by a 2 × 2 maximum pooling layers. The second part has a two-layer stack of 32 convolutional kernels of size 3 × 3 followed by a 2 × 2 maximum pooling layer. The third part has a two-layer stack of 32 convolutional kernels of size 3 × 3 followed by a flatten layer. Maximum pooling is applied after the first convolutional layer, while stacking is applied after the remaining two convolutional layers. The fourth part is a four-layer full connection layer with input dimensions of 256, 1024, 128, and 2 respectively.

Additionally, the benefits of this model can be summed up as follows:

1) A 5 × 5 convolution kernel is used in the first convolution layer to extract information from the time-frequency image's greater neighborhood range and produce superior features. In order to expand the receptive field, collect more data, and provide the network's later layers more information, huge convolutional kernels are employed in the first layer. Additionally, these kernels are more effective at suppressing high-frequency noise. [18]
2) Use two 3 × 3 convolution kernels instead of one 5 × 5 convolution kernel. Gaining more nonlinear expression capabilities requires two activation functions for each of the two 3 × 3 convolution layers. Less parameters can lower the computational effort while two layered convolutional layers can enhance the feature extraction capability.

TABLE I

STRUCTURAL PARAMETERS OF THE CNN MODEL

Layers	Parameters
C1	Conv2D(5 X 5 X 32)
MP1	MaxPool2D(2 X 2)
C2	Conv2D(3 X 3 X 32)
C3	Conv2D(3 X 3 X 32)
MP2	MaxPool2D(2 X 2)
C4	Conv2D(3 X 3 X 32)

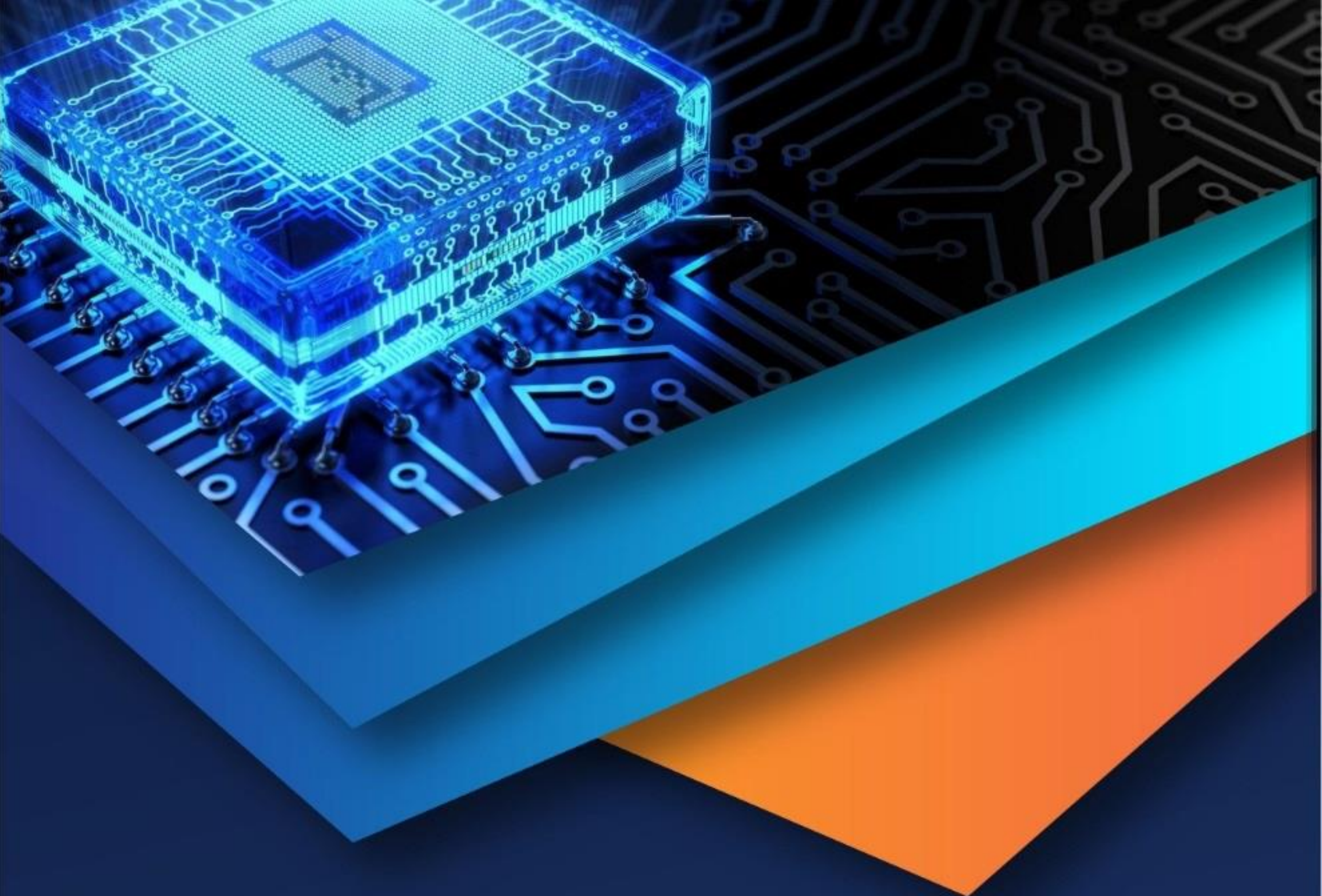
C5	Conv2D(3 X 3 X 32)
Flatten Layer	
FC1	Input Dimensions = 256
FC2	Input Dimensions = 1024
FC3	Input Dimensions = 128
FC4	Input Dimensions = 2

V. CONCLUSION AND FUTURE SCOPE

CNN for condition monitoring in machine tools holds promise in enhancing fault diagnosis efficiency and accuracy. By transforming one-dimensional vibration signals into two-dimensional time-frequency images and leveraging CNNs for fault feature acquisition, the system addresses inherent challenges in extracting salient features from raw vibration signals. Future enhancements of the project could involve further exploration of advanced feature extraction techniques, integration of multimodal data sources, and optimization of model generalization through transfer learning methods. Real-time monitoring capabilities, adaptive learning mechanisms, and deployment on edge devices or IoT platforms could enhance the system's ability to detect emerging faults and enable distributed condition monitoring across interconnected machine tool networks. Additionally, using unsupervised learning techniques for anomaly detection and conducting extensive validation testing on several datasets can provide insightful views into the robustness of the system and guide further improvements to ensure its effectiveness in industrial environments. Overall, the proposed approach represents a promising direction for advancing condition monitoring in the industrial sector.

REFERENCES

[1] Choudhary, Anurag, Tauheed Mian, and Shahab Fatima. "Convolutional neural network based bearing fault diagnosis of rotating machine using thermal images." Measurement 176 (2021): 109196.
[2] Schwendemann, Sebastian, Zubair Amjad, and Axel Sikora. "A survey of machine-learning techniques for condition monitoring and predictive maintenance of bearings in grinding machines." Computers in Industry 125 (2021): 103380.
[3] AlShorman, Omar, Muhammad Irfan, Nordin Saad, D. Zhen, Noman Haider, Adam Glowacz, and Ahmad AlShorman. "A review of artificial intelligence methods for condition monitoring and fault diagnosis of rolling element bearings for induction motor." Shock and Vibration 2020 (2020): 1-20.
[4] Serin, Gokberk, Batihan Sener, A. Murat Ozbayoglu, and Hakki Ozgur Unver. "Review of tool condition monitoring in machining and opportunities for deep learning." The International Journal of Advanced Manufacturing Technology 109, no. 3 (2020): 953-974.
[5] Wang, Xin, Dongxing Mao, and Xiaodong Li. "Bearing fault diagnosis based on vibro-acoustic data fusion and 1D-CNN network." Measurement 173 (2021): 108518.
[6] Senanayaka, Jagath Sri Lal, Huynh Van Khang, and Kjell G. Robbersmyr. "Toward self-supervised feature learning for online diagnosis of multiple faults in electric powertrains." IEEE Transactions on Industrial Informatics 17, no. 6 (2020): 3772-3781.
[7] Zhang, Wei, Gaoliang Peng, Chuanhao Li, Yuanhang Chen, and Zhujun Zhang. "A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals." Sensors 17, no. 2 (2017): 425.
[8] Abdeljaber, Osama, Onur Avci, Serkan Kiranyaz, Moncef Gabbouj, and Daniel J. Inman. "Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks." Journal of sound and vibration 388 (2017): 154-170.
[9] Su, Liyuan, Lei Ma, Na Qin, Deqing Huang, and Andrew H. Kemp. "Fault diagnosis of high-speed train bogie by residual-squeeze net." IEEE Transactions on Industrial Informatics 15, no. 7 (2019): 3856-3863.
[10] Wang, Huan, Zhiliang Liu, Dandan Peng, and Yong Qin. "Understanding and learning discriminant features based on multiattention 1DCNN for wheelset bearing fault diagnosis." IEEE Transactions on Industrial Informatics 16, no. 9 (2019): 5735-5745.
[11] Zhao, Zhibin, Tianfu Li, Jingyao Wu, Chuang Sun, Shibin Wang, Ruqiang Yan, and Xuefeng Chen. "Deep learning algorithms for rotating machinery intelligent diagnosis: An open source benchmark study." ISA transactions 107 (2020): 224-255.
[12] Janssens, Olivier, Viktor Slavkovikj, Bram Vervisch, Kurt Stockman, Mia Loccufier, Steven Verstockt, Rik Van de Walle, and Sofie Van Hoecke. "Convolutional neural network based fault detection for rotating machinery." Journal of Sound and Vibration 377 (2016): 331-345.
[13] Zhang, Shen, Shibo Zhang, Bingnan Wang, and Thomas G. Habetler. "Deep learning algorithms for bearing fault diagnostics—A comprehensive review." IEEE Access 8 (2020): 29857-29881.
[14] Bhadane, Mukesh, and K. I. Ramachandran. "Bearing fault identification and classification with convolutional neural network." In 2017 International Conference on Circuit, Power and Computing Technologies (ICCPCT), pp. 1-5. IEEE, 2017.
[15] Hoang, Duy-Tang, and Hee-Jun Kang. "Rolling element bearing fault diagnosis using convolutional neural network and vibration image." Cognitive Systems Research 53 (2019): 42-50.
[16] Wang, Qibin, Bo Zhao, Hongbo Ma, Jiantao Chang, and Gang Mao. "A method for rapidly evaluating reliability and predicting remaining useful life using two-dimensional convolutional neural network with signal conversion." Journal of Mechanical Science and Technology 33 (2019): 2561-2571.
[17] Wen, Long, Xinyu Li, Liang Gao, and Yuyan Zhang. "A new convolutional neural network-based data-driven fault diagnosis method." IEEE Transactions on Industrial Electronics 65, no. 7 (2017): 5990-5998.
[18] Yang, Jingli, Tianyu Gao, Shouda Jiang, Shijie Li, and Qing Tang. "Fault diagnosis of rotating machinery based on one-dimensional deep residual shrinkage network with a wide convolution layer." Shock and Vibration 2020 (2020): 1-12.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY