A transfer learning approach based tool wear detection in the turning process using vibration signals

Sudhan Kasiviswanathan and Sakthivel Gnanasekaran*

School of Mechanical Engineering, Vellore Institute of Technology, Chennai, India

Continuous monitoring of the cutting tool insert's condition is essential to enhance product quality and efficient machining process, by reducing the machine downtime. But the available tool condition monitoring approaches are often limited by coolant induced visibility loss in the cutting zone that reduces the feature reliability. This study proposes a transfer learning based deep learning method where the machining vibration signals are converted into visual representations and classified using ResNet 18, MobileNet V2, SqueezeNet, ShuffleNet, DenseNet 201, and EfficientNet B0 pretrained convolutional neural networks. This combination enables the model to learn deep wear profiles from vibration data without the manual feature extraction. Also, this method enhances signal strength, making it highly suitable for smart, scalable, and real world manufacturing environments. The effects of the proposed pretrained network hyperparameters, such as mini batch size, solver type, learning rate, and filter size, were studied and EfficientNet B0 was identified as the best performing network with a classification accuracy of 89.23% for tool condition monitoring tasks.

1 Introduction

The present smarter and effective production environments use the Industrial Internet of Things (IIoT) to assist real time data generation, machine monitoring, and predictive analytics, which enables intelligent and connected systems within the present environment. In this framework, tool condition monitoring (TCM) has gained value as a main component of predictive maintenance. As the future industrial revolution focuses on human centric approaches, intelligent TCM systems will be the key to this competitiveness. This has stimulated growing interest among researchers in the development of advanced TCMs, that includes collecting, organizing, and analysing experimental data for real time applications and increases the tool wear detection, also reduces downtime of the machine (Okada et al., 2011; Jardine et al., 2006). For instance, Wilcox et al. framed a tool wear monitoring system using self organizing maps (SOM) and adaptive resonance theory neural networks. This system studied the sensor data from accelerometers, microphones, and strain gauges, and combined the neural network output with taylor's tool life model to rank tool wear. The study showed that integrating neural networks with an expert system improves accuracy and makes it a reliable real time monitoring solution (Silva et al., 2006). Kadim et al. identified the tool wear during turning operations by measuring strain and vibration from

the cutting tool using a piezoelectric strain sensor and an accelerometer connected to a data acquisition card. As a result, a sum of 24 feature indicators of tool wear was extracted from the obtained raw signal, which includes time domain, frequency domain, time series model coefficients, and wavelet packet analysis features by a (2 × 3) SOM (Abid Al-Sahib and Bachaa, 2005). Sudhan Kasiviswanathan et al. reviewed the indirect TCM methods for turning operations. The study showed the importance of integrating IIoT and ML into TCM systems for achieving higher efficiency, enhancing performance, and to support sustainable intelligent manufacturing (Kasiviswanathan et al., 2024). Alonso et al. utilized a feedforward backpropagation network to assess cutting tool's flank wear, leveraging vibration signals and singular spectrum analysis (Alonso and Salgado, 2008). Similarly, Dutta et al. developed a real time TCM system that applies image texture analysis like gray level cooccurrence matrix, Voronoi tessellation, and discrete wavelet transforms, to extract features related to the cutting tool's flank wear state. A linear support vector machine regression model was then used to predict tool flank wear and achieved a lower prediction error (Dutta et al., 2016). The advancement and demand for condition monitoring led to the use of sophisticated methods like infrared thermography. On the other hand, such methods are expensive and demand skilled operators (Bagavathiappan et al., 2013). Earlier studies show that vibration signals are widely used for monitoring machine tool conditions. However, analysing them is difficult because of their rapidly changing behaviour and intricate structure. Researchers have used different methods, including histograms, statistical measures, support vector machines and fuzzy logic, to track tool conditions (Mohanraj et al., 2020). Additionally, feature selection techniques like Filter methods, wrapper methods, embedded methods, and decision trees have been employed. The selected features are then classified using various types of ML algorithms, namely, Navie Bayes, decision trees, Random Forest, and K nearest neighbour's families. While numerous studies have explored the use of ML algorithms for TCM, the performance and efficiency of these algorithms are greatly dependent on the standard of the extracted and selected features. Since the present data led diagnostic methods are based on AI principles, they become an essential tool for handling these large datasets. Geoffrey Hinton introduced a deep learning strategy known as a deep neural network (DNN), which consists of numerous neural layers arranged hierarchically to extract information from input data. This configuration is referred to as "deep" because its intervention on raw data at different levels progressively uncovers the structure of complex data sets and independently identifies the most significant attributes (Ekundayo and Ezugwu, 2025). This ability to learn the features, paired with current nonlinear regression functions, has led DL models to gain widespread popularity in areas like Natural Language Processing, object detection, image classification, and pattern recognition (Elhefnawy et al., 2022). The research conducted by Yang Fu et al. used deep belief networks (DBN) to automatically construct a feature space for cutting tool monitoring and utilized a greedy layer wise strategy for pretraining and back propagation for fine tuning. The performance of DBNs is compared with manually defined features from both time and frequency domains. The results demonstrate that DBN provide comparable feature characterization with a significantly higher modelling accuracy

(Fu et al., 2015). Verstraete et al. applied short time fourier transform, wavelet transform, and hilbert huang transform to identify the rolling element bearing faults by vibration signals and achieved reliable accuracy with limited data (Verstraete et al., 2017). Luis Enrique Escajeda Ochoa et al. conducted TCM for highspeed machining using the stacked sparse autoencoder method and it showed reliable tool wear prediction (Ochoa et al., 2019), while Pradeep Katta et al. study presents optimized DBN for analysing induction motor performance by using stacked restricted Boltzmann machines and trained with an Ant Colony algorithm which is designed to extract features from sensor based vibration signals and experimental results showed robust fault detection accuracy (Katta et al., 2024). From these cases it is identified, DNN were primarily used as classifiers without fully exploiting their feature learning potential. However, since 2015, researchers have begun using DNN for learning, selecting, and classifying features for comprehensive solutions by eliminating the necessity for explicit feature extraction and selection, DNN learn from provided input images, with corresponding labels as output, where this advanced stage of fault diagnosis was explored in recent studies. The research conducted by Guo et al. (2016) to monitor the condition of roller bearings by employing a deep CNN with two ensembles, one method emphasizes feature extraction and fault pattern recognition, while the other focuses on fault classification. The complexity of vibration signals makes the task of feature extraction a challenging one. To address these difficulties, researchers are now concentrating on developing an automated system that can classify data directly from raw signals, bypassing the need for explicit feature extraction. Recently, DL has shown its potential in tackling condition monitoring issues by automatically learning features from images to achieve accurate classifications. In DL methods, especially CNN, image features are learned automatically, allowing classification without the need for separate attribute extraction, selection, and labelling tools. Although DL techniques are primarily designed for image processing, capturing images while the tool is in operation is challenging, hazardous, and costly. According to the literature, the best method for gathering vibration signals from the machining area is to place the sensor near the cutting environment. For this research, several sensor placement trial runs were conducted and positioning it on the tool holder, near the cutting area, was identified as the optimal location for vibration signal acquisition. CNNs can learn features from these graphs and classify them based on predefined wear profile categories.

From the literature it is found that only few attempts were made on utilizing DL based pretrained networks for industrial TCM applications and no attempts were made for utilizing vibration signatures as an input to the pretrained networks. These networks are referred to as "pretrained" because they are previously trained on a large scale dataset, enabling them to learn rich features from the input data. This process of leveraging a pretrained model is called transfer learning, and it significantly reduces the time and computational resources needed for training on new tasks, when the available data is limited. The proposed research presents a novel approach to assess the tool's condition using vibration signatures by employing high performing pretrained DL models, such as SqueezeNet, its lightweight architecture allows it to process the raw data with minimal computation, which makes it

suitable for the TCM application (Guo et al., 2016), ShuffleNet, its group convolution and channel shuffling process allows it to achieve high classification efficiency by capturing localized texture variations in vibration signature images (X. Zhang et al., 2017), ResNet 18, enables stable feature learning of wear profiles and training process (Li et al., 2024), MobileNet V2, its inverted residual block structures process the data with high accuracy and efficient for industrial applications (Sandler et al., 2018), DenseNet 201, offers a dense features which improves the structures flow (Salim et al., 2023) and EfficientNet B0, uses compound scaling of input data resolution to achieve accuracy (Ali et al., 2025).

Further, this study explored several conventional and current TCM methods used in real time environment and observed that effectiveness varies significantly. Traditional TCM approaches that use statistical data, frequency domain analysis, and customize features are effective, but their accuracy relies on the feature design. These methods also often have challenges on recording tool wear behaviour that changes over time. Image based monitoring methods are compatible for CNN based deep learning, but they have practical difficulties like coolant blocking near the cutting area, limited accessibility for tool workpiece interface, and decreased image resolution during rapid machining, which makes it harder to get continuous images while machining (Wang et al., 2021). Vibration signals based methods on the other hand provided a non intrusive, cost effective, and reliable way to monitor the tools condition (Y. Zhang et al., 2023). Traditional vibration based methods require high feature extraction, which may not possess in all wear conditions (Wang et al., 2021; Li et al., 2024). Even though vibration signals are commonly used to monitor tool wear, most studies still use manual features with traditional machine learning classifiers or train deep networks with small data sets (Wang et al., 2021). These approaches often fail due to the requirement for extensive preprocessing of signal data, and the application of transfer learning with pretrained CNN models, as well as the data conversion is limited. This survey shows that a hybrid deep learning method that combines the reliability of vibration sensing with the modern CNN is needed to transform one dimensional vibration data into two dimensional images.

This research mitigates these constraints by enhancing,

- A vibration based tool wear monitoring framework transforms one dimensional vibration signals into two dimensional GASF images, facilitating CNN based feature learning without necessitating manual feature engineering.
- A systematic transfer learning approach for tool wear classification under consistent experimental conditions was used to test the pretrained CNN architectures SqueezeNet, ShuffleNet, ResNet 18, MobileNet V2, DenseNet 201, and EfficientNet B0.
- A strategy for dividing data into batches that reduces signal similarity leakage and gives a reliable measure of the extent to which a model generalizes.
- A thorough study of hyperparameter optimization that observes how the size of the minibatch, the type of solver, the learning rate, and the size of the filter affect the classification of tool wear based on vibration images.

- Validation of a computationally efficient and scalable solution for real time monitoring of the condition of industrial tools using a small amount of training data.
- The suggested method uses existing visual feature representations and transforms them to work for wear recognition by combining transfer learning with pretrained CNNs.

2 Methodology

The stated methodology comprises several sequential stages, from signal acquisition to TCM prognostic model deployment, as shown in Figure 1, and is based on analyzing vibration signals using DL models.

- Tool insert selection: Based on microscopic and surface observations, for this study, four pattern tools are selected, namely, Good or No wear, Nose, Crater, and Flank wear.
- Vibration Signal: During machining operations, these selected tools are mounted to the tool holder one by one to obtain the vibration data using the accelerometer sensors fixed on the tool holder.
- Signal amplification: At this phase, the generated vibration signal will be filtered and amplified to enhance the signal to noise ratio.
- Analogue to Digital signal conversion: Using the ADC, the vibration signals are converted into digital signals for computing and frequency time (FT) domain vibration plot generation.
- Dataset preparation: For effective analysis, the training and testing datasets are split from the obtained 2D GASF image data. The training datasets are used to train the model, and the testing datasets are used to validate the model.
- Training phase: Using the training datasets, the proposed models were trained for their classification accuracy.
- Validation phase: The pretrained model's classification accuracy, precision, recall, and overall reliability were evaluated using the testing dataset.
- Prediction of tool wear: The decisions made by each proposed model on tool condition were monitored to present a reliable and non intrusive solution for real time tool wear prediction.

2.1 Transforming a 1D vibration plot into a 2D GASF image

The GASF method was utilized to convert vibration signals into images as it can record the time dependencies and encode global signal correlations in a structured two dimensional form. GASF captures the pairwise angular relationships between all time samples, which retains long range temporal information that is essential for identifying the difference between gradual and nonlinear tool wear progression. This differs from traditional time frequency representations like short time Fourier transform (STFT) and wavelet transform, which focus on localized frequency content.

FIGURE 1
Tool wear assessment methodology.
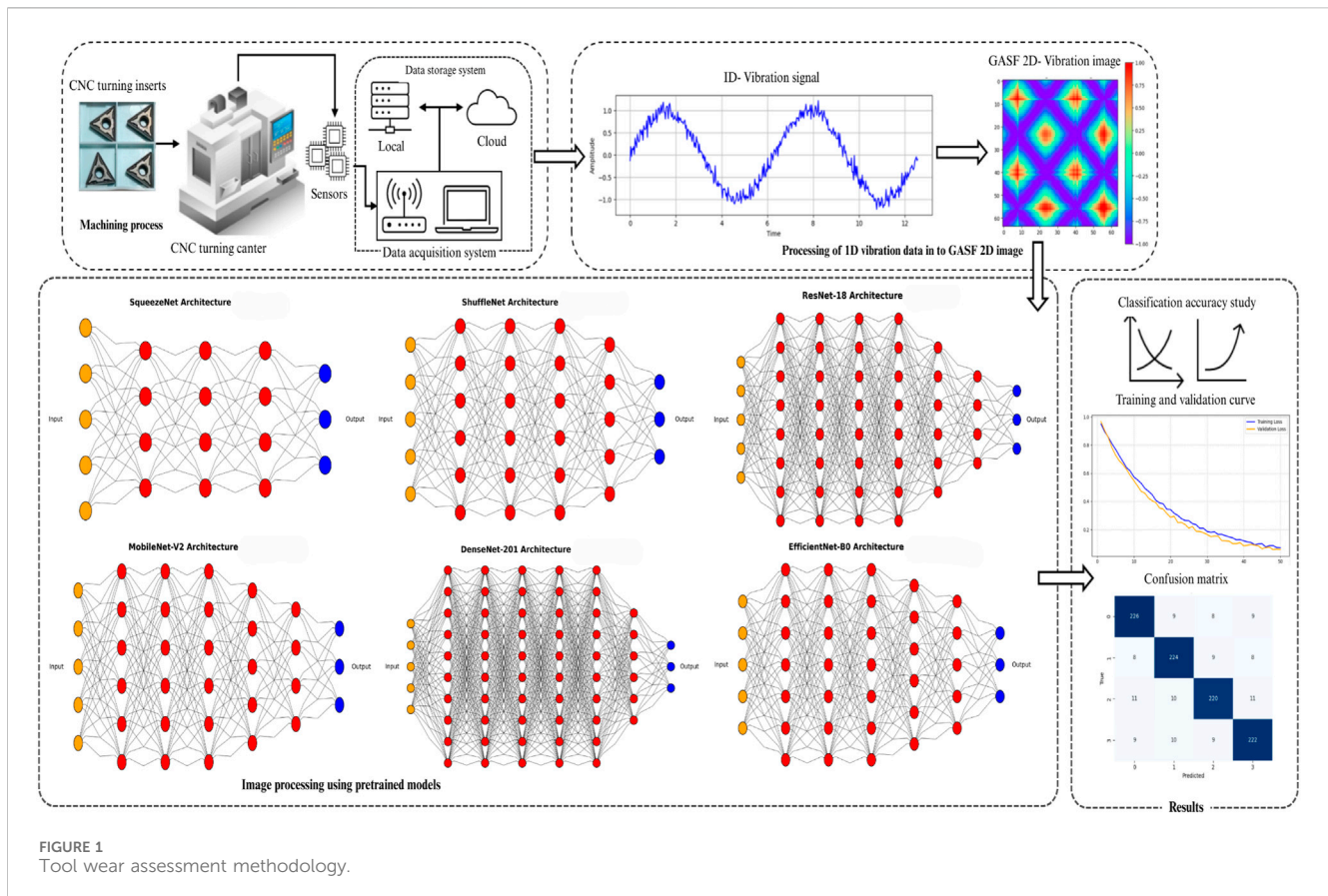
In GASF, the transformation of a normalized one dimensional signal into a polar coordinate system facilitates the encoding of temporal dynamics as spatial patterns, preserving both amplitude variation and temporal ordering. This representation is particularly beneficial in monitoring tool wear based on vibration plots because changes caused by wear often show up as small changes in signal correlation instead of clear spectral peaks. Additionally, GASF creates dense and smooth image textures that work well with convolutional neural networks. This makes it easy to learn features without having to do explicit time frequency decomposition or hand picking features.

GASF is a deterministic and parameter light transformation that is sensitive to window size and decomposition parameters than other vibration to image encoding methods like recurrence plots or scalograms. This makes it stronger and easier to use in industry, especially when combined with transfer learning that uses pretrained CNN architectures. As a result, GASF is an excellent method to connect reliable vibration sensing with strong image based deep learning models for classifying tool wear.

The one dimensional vibration signals acquired from the accelerometer were converted into two dimensional GASF images to facilitate image based deep learning. The GASF encodes the temporal correlations of the signal into a structured matrix representation that is appropriate for CNN based analysis.

Step 1: Signal normalization, each vibration signal x(t) was first rescaled to the range [-1,1] using Equation 1

$$x_0^i = \frac{xi - \min(x)}{\max(x) - \min(x)} \times (Xmax - Xmin) + Xmin \qquad (1)$$

Where,

xi = original value
$\max(x) - \min(x)$ = minimum and maximum values of the feature in the dataset
$(Xmax - Xmin)$ = desired scaling range
x_0^i = normalized value

Step 2: Polar Encoding, the normalized signal was converted into angular form using Equation 2

$$\varphi = \arccos(\tilde{y}i), \, -1 \leq (\tilde{y}i) \leq 1, \, \tilde{y}i \in \tilde{Y} \qquad (2)$$

Where,

$\tilde{Y} = y1, y2., yn$ is the set of rescaled cutting force samples (after Min–Max normalization or similar).

The angular coordinate ϕ is obtained from the across mapping of $\tilde{y}i$

Step 3: 2D GASF image generation

The GASF matrix was computed using Equation 3

$$GASF = \begin{bmatrix} \cos(\phi1 + \phi1) & \cos(\phi1 + \phi2) & \cdots & \cos(\phi1 + \phi n) \\ \cos(\phi2 + \phi1) & \cos(\phi2 + \phi2) & \cdots & \cos(\phi2 + \phi n) \\ \vdots & \vdots & \ddots & \vdots \\ \cos(\phi n + \phi1) & \cos(\phi n + \phi2) & \cdots & \cos(\phi n + \phi n) \end{bmatrix} \qquad (3)$$

2.2 Transfer learning approach

To utilize the feature representations from large image datasets and to adapt these representations for the classification of tool wear based on vibration data transfer learning was applied. For this study, pretrained weights from the ImageNet dataset was used to initialize all CNNs used. We removed the original fully connected classification layers in each pretrained model and added a new task specific classification head which was generated for the tool wear classes. The convolutional backbone layers were initially frozen during training to retain the pretrained feature representations. Only the new classification layers were trained with the GASF images. Selective fine tuning was performed by unfreezing the upper convolutional blocks of each network, enabling restricted adaptation to the features extracted from vibration images. We used a low learning rate during the fine tuning process to lower the chance of losing all the pretrained weight. All the models used categorical cross entropy loss for training and were improved by solvers based on stochastic gradient descent. During the optimization study, hyperparameters such as learning rate, minibatch size, and solver type varied in a systematic way. This transfer learning approach allows models to converge promptly using fewer training data, and it further decreases overfitting.

3 Empirical research

The vibration data used for this study was obtained from standalone machining experiments which were conducted for this research. The experiments were performed with controlled cutting conditions along with a tri axial accelerometer mounted on the tool holder which allowed to continuously record the vibration signals while the machining. Progressive wear measurements were used to group the tool wear states, to make sure that signal acquisition and wear classification were consistently similar. This experimental design makes it possible to accurately test the proposed vibration based deep learning framework in real world machining situations.

3.1 Experimental facility configuration

The experimental setup for this study consisted of an industrial CNC turning center, a single axis accelerometer sensor, and a Data collection system (DAQ).

The single point cutting tool carbide Insert TNMG160404 was used along with the tool holder, and an EN8 carbon steel shaft with a length of 320 mm and a diameter of 50 mm was used as a workpiece and held in place by a hydraulic chuck on the CNC turning center. Vibration signals were obtained using a uniaxial piezoelectric accelerometer, which was securely mounted on the tool holder with a mount. These signals were processed and transformed into digital format by the signal conditioning module in LabVIEW software, which included an integrated analog to digital converter (ADC). The digitally converted vibration signal patterns are then transmitted to a personal system for storage. Figure 2 illustrates the experimental Configuration utilized for this research.
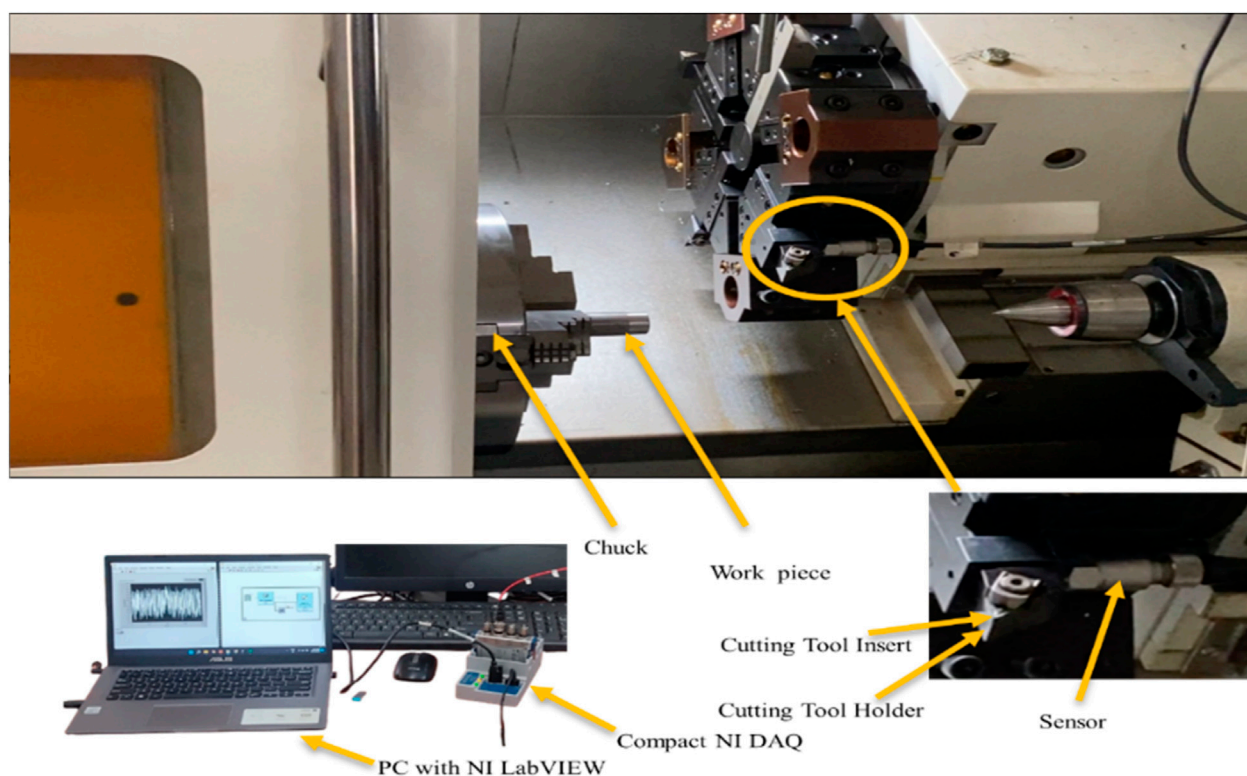

FIGURE 2
TCM experimental setup.

3.2 Data collection

Data collection (DAQ) transforms real world physical signals into digital values for display, analysis, and storage. In this research, TCM was carried out by capturing vibration signals using a piezoelectric single axis accelerometer with a sensitivity of 10.22 mV/g. Accelerometers can capture a broad spectrum of vibrations, with their output voltage directly reflecting the intensity of the vibrations. In this study, the sensor was securely mounted on the arm of the cutting tool holder, positioned as close as possible to the cutting zone, using adhesive to maintain signal accuracy. An analog to digital converter (ADC) was used to digitise the collected analog signals. These digital vibration signals were then used to generate vibration plots, which were stored on a computer for further analysis. A TL based strategy was implemented to determine which pretrained model delivered the most effective performance for TCM.

3.3 Experimental process

3.3.1 Cutting tool insert selection

In this study, three commonly occurring types of tool wear conditions and no wear or good tool were considered for the experimentation as outlined in Table 1.

TABLE 1 Tool condition selection for experimentation.

Wear type	Wear description
No wear (or) good	Fresh cutting tool insert, an unused tool that has no wear
Nose wear	Wear occurred on the cutting edge of the insert due to friction between the workpiece and the cutting tool
Crater wear	A concave area formed on the rake face of the tool insert due to the sliding motion of the chip
Flank wear	Wear developed on the flank of the cutting edge because of friction between the side face of the cutting edge and the workpiece during machining

3.3.2 Experimentation

A new carbide turning insert with a 0.4 mm nose diameter was secured to a tool holder, which was fixed to a tool turret head, along with the acceleration sensor mounted on the cutting tool holder's arm. The factors for signal acquisition, such as the length of the sample, frequency of the sample, and the type of signal, were established in advance. Following the Nyquist sampling theorem, the frequency of the sample was set to 25 kHz, which is twice the observed frequency of 12.5 kHz. To begin the turning process, an EN8 steel shaft with a 50 mm diameter was positioned in the center of the three jaw hydraulic chuck. The machining parameters are given in Table 2, which were coded into the CNC turning center for each tool category. Upon commencing the machine, the DAQ system was powered on, and initial signals were discarded to reduce random variations. Vibration data were recorded from the mounted sensor. The signal collection factors were as follows:

- Length of sample: 8,192 steps
- Frequency of sample: 25 kHz
- Count of occurrences per condition: 86

3.4 Data processing

A dataset of images representing various tool factors was generated from the acquired vibration data signals for this study. A grouped partitioning approach was adopted to split the data set and to eliminate signal similarity from leak between training and testing sets. All vibration samples are collected from a single machining run, linked to one tool condition batch, were combined and categorized into either the training or testing set. Applying a 70/30 split at the batch level instead of the sample level performed well to prevent the training and testing data from overlapping. This approach also makes sure that the reported performance is a true indication of the model's ability to generalize Figure 3 shows how to turn one dimensional vibration plots into two dimensional GASF image data.

TABLE 2 Processing factors for experimentation.

| Machining sets				
Speed (rpm)	Feed (mm/rev)	Depth of cut (mm)	Tool Condition	Representation
1,000	0.2	1	No wear (or) good tool (NOT)	NOT-S1F0.2D1
1,000	0.2	1	Nose wear (NWT)	NWT-S1F0.2D1
1,000	0.2	1	Crater wear (CWT)	CWT-S1F0.2D1
1,000	0.2	1	Flank wear (FWT)	FWT-S1F0.2D1
1,000	0.2	1.5	No wear (or) good tool (NOT)	NOT-S1F0.2D1.5
1,000	0.2	1.5	Nose wear (NWT)	NWT-S1F0.2D1.5
1,000	0.2	1.5	Crater wear (CWT)	CWT-S1F0.2D1.5
1,000	0.2	1.5	Flank wear (FWT)	FWT-S1F0.2D1.5
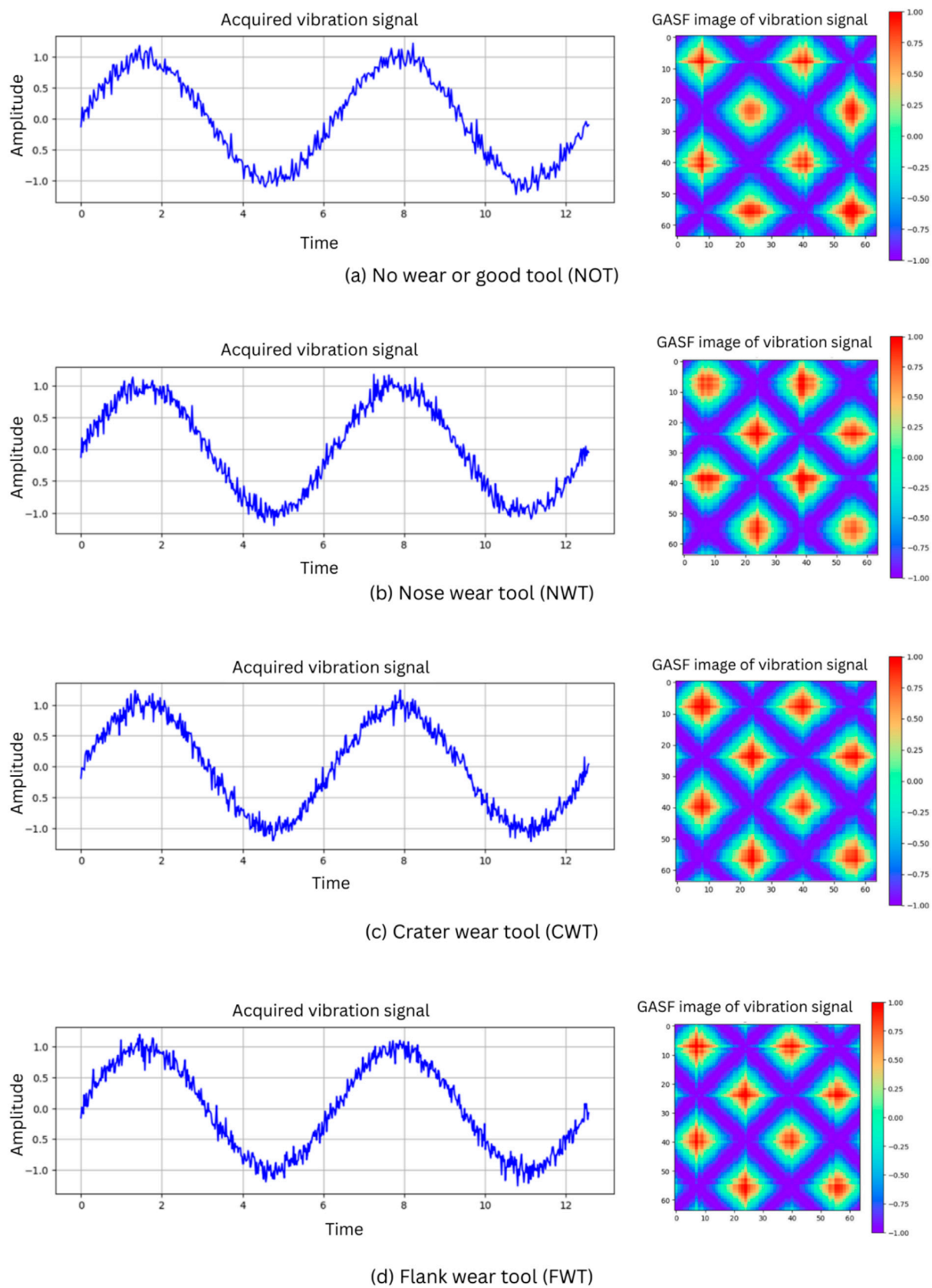
FIGURE 3
1D vibration data plots to 2D GSAF images of different wear tools (a) No wear or good tool (b) Nose wear tool (c) Crater wear tool (d) Flank wear tool.

A total of 688 images, generated from the set conditions, from the vibration signals. The input images underwent the required preprocessing steps to prepare them for use in the selected pretrained models:

- Resized: All images were resized to 224 × 224 pixels to match the input size required by the pretrained models.
- Normalization: Pixel values were normalized to the [0, 1] range to standardize input across samples and speed up convergence.
- Colour Channel Formatting: The single channel grayscale GASF images were replicated across three channels to conform with the RGB input format expected by the ImageNet pretrained models.
- Data Augmentation: To improve generalization, basic augmentation techniques were applied, including random rotations ±15°, horizontal flipping, and slight zooming. This helped expose the models to varied representations of tool wear patterns during training.

4 Application of pretrained models

The structured design of the CNN's architecture enables it to learn and extract the intended data from the inputs. These extracted data features during the convolutional process decide the performance of a network by generating weights and biases, which create a link between the input image and its features. The features like edges, textures, and shapes of the image are examined

by the model's filters, which aid in its overall decision making ability. To address the challenges faced by the conventional model and to utilize the features of DL for this study, six pretrained networks and the transfer learning (TL) approach are used. These pretrained models are trained to recognize general patterns and visual features, which makes them highly suited for tool wear classification tasks. The final classification layer of the model was replaced with the proposed models, where the initial layers are the same. The models are then fine tuned on the vibration datasets images converted by the GASF method. This method enables the model to quickly learn the specific patterns. Further, this section outlines the pretrained networks employed to evaluate the condition of a carbide cutting tool insert, and converting 1D vibration signals into 2D spectrograms enhances the ability of deep learning models to detect wear patterns. This approach is driven by the proven effectiveness of deep learning in enhancing feature recognition for industrial applications [26]. Figure 4 illustrates the architecture of the pretrained model utilized in this study.

4.1 Characteristics of the pretrained model used

To explore how well TL can be applied to tool condition monitoring, a selected mix of well known and modern pretrained convolutional neural networks was used. These models vary in design and complexity, which helps provide a balanced comparison of accuracy, efficiency, and suitability for industrial
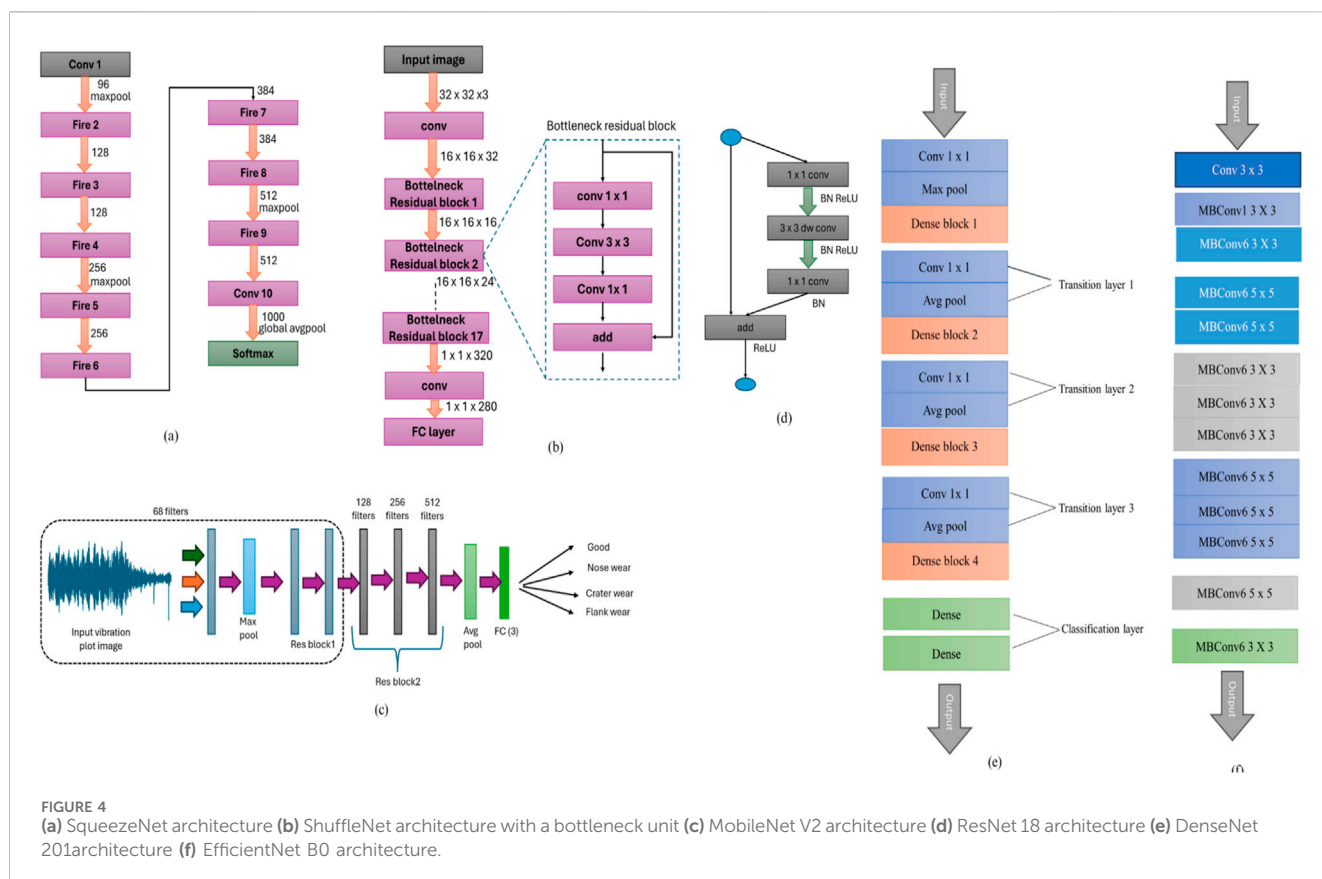

FIGURE 4
(a) SqueezeNet architecture **(b)** ShuffleNet architecture with a bottleneck unit **(c)** MobileNet V2 architecture **(d)** ResNet 18 architecture **(e)** DenseNet 201architecture **(f)** EfficientNet B0 architecture.

TABLE 3 Pretrained model characteristics.

Pretrained model	Characteristics
SqueezeNet	Lightweight model with "fire modules", reduces parameters while maintaining accuracy, ideal for low resource settings
ShuffleNet	Employs group convolutions and channel shuffling, highly efficient for edge and embedded device deployment
ResNet 18	Utilises residual connections to maintain gradient flow, efficient and reliable for moderate sized datasets
MobileNet V2	Incorporates inverted residuals and linear bottlenecks, optimised for speed and accuracy on mobile platforms
DenseNet 201	Densely connected layers promote feature reuse and stable gradient flow, they perform well in deep learning tasks with limited data
EfficientNetB0	Uses compound scaling and MBConv layers to achieve high accuracy with fewer parameters; suitable for real time industrial use

use. Each model was fine tuned using vibration signal plots from machining operations to assess its ability to classify tool wear effectively. A summary of each model and its role in this study is presented in Table 3.

5 Results and discussion

The Deep Learning Toolbox and TL package in MATLAB 2020a was used for the experiments. To ensure that no vibration samples from the same machining run were in both sets, a 70/30 train test split was executed at the tool batch level. Eight hyperparameter configurations were explored, and the key findings are summarized below.

5.1 Minibatch size effects

Minibatch size influences memory, training speed, and generalization. In this study, the 8,192 sample dataset was split into batches of 4, 8, and 16, revealing that the batch of 8 gave the best overall results. ResNet 18 reached 89.8% accuracy with this size, MobileNet V2 and ShuffleNet also peaked at batch 8, 87.9% and 86.1%, respectively, while SqueezeNet performed best with batch 4, 83.76% the DenseNet 201 achieved its best performance 89.4% and EfficientNet B0 also reached its top accuracy of 90.2% with a batch size of 8.

5.2 Solver type effects

The solver is an expansion algorithm that modifies model weights during training to reduce the loss function. This selection shows effects on the speed, accuracy, and generalization. Three optimisers were related in this study: Stochastic Gradient Descent with Momentum (SGDM), Adaptive Moment Estimation (ADAM), and Root Mean Square Propagation (RMSProp). SGDM produced the extreme accuracy for SqueezeNet 89.9%, ShuffleNet 86.7%, and ResNet 18 80.5%. MobileNet V2 performed best with Adam, 88.6%. DenseNet 201 performed best with RMSProp, achieving 88.9% accuracy and EfficientNet B0 reached its top performance of 87.3% using ADAM.

5.3 Learning rate factor effects

The learning rate controls how much a model's weights are updated during training. Choosing the exact value is difficult for this process. A lesser value can slow down the process, while a higher value can end up in poor training of the model. The general used common values of 0.01, 0.001, and 0.0001 are chosen for the training phase. The highest accuracy of all the used networks was reached at a 0.0001 learning rate factor.

5.4 Filter size effects

For feature extraction, the filter size plays an important role by determining the receptive field. The model's execution, computational efficiency depends on the proper selection of filter size. Filter sizes of 1×1, 3×3, and 5×5 were examined. ResNet 18 was most accurate with 1×1 filters, 83.3%, whereas ShuffleNet, SqueezeNet, and MobileNet V2 all peaked with 5×5 filters, 87.4%, 87.1%, and 87%. DenseNet 201 used a filter size of 3×3 and achieved an accuracy of 88.5%. EfficientNet B0 also peaked at 90.2% with 3×3 filters.

5.5 Pretrained model comparative study with hyperparameter optimization

Using optimal hyperparameters, EfficientNet B0 achieved the highest classification accuracy of 89.23% while also requiring less computation time with 50 epochs, as shown in Table 4.

TABLE 4 Overall classification accuracy of the pretrained models.

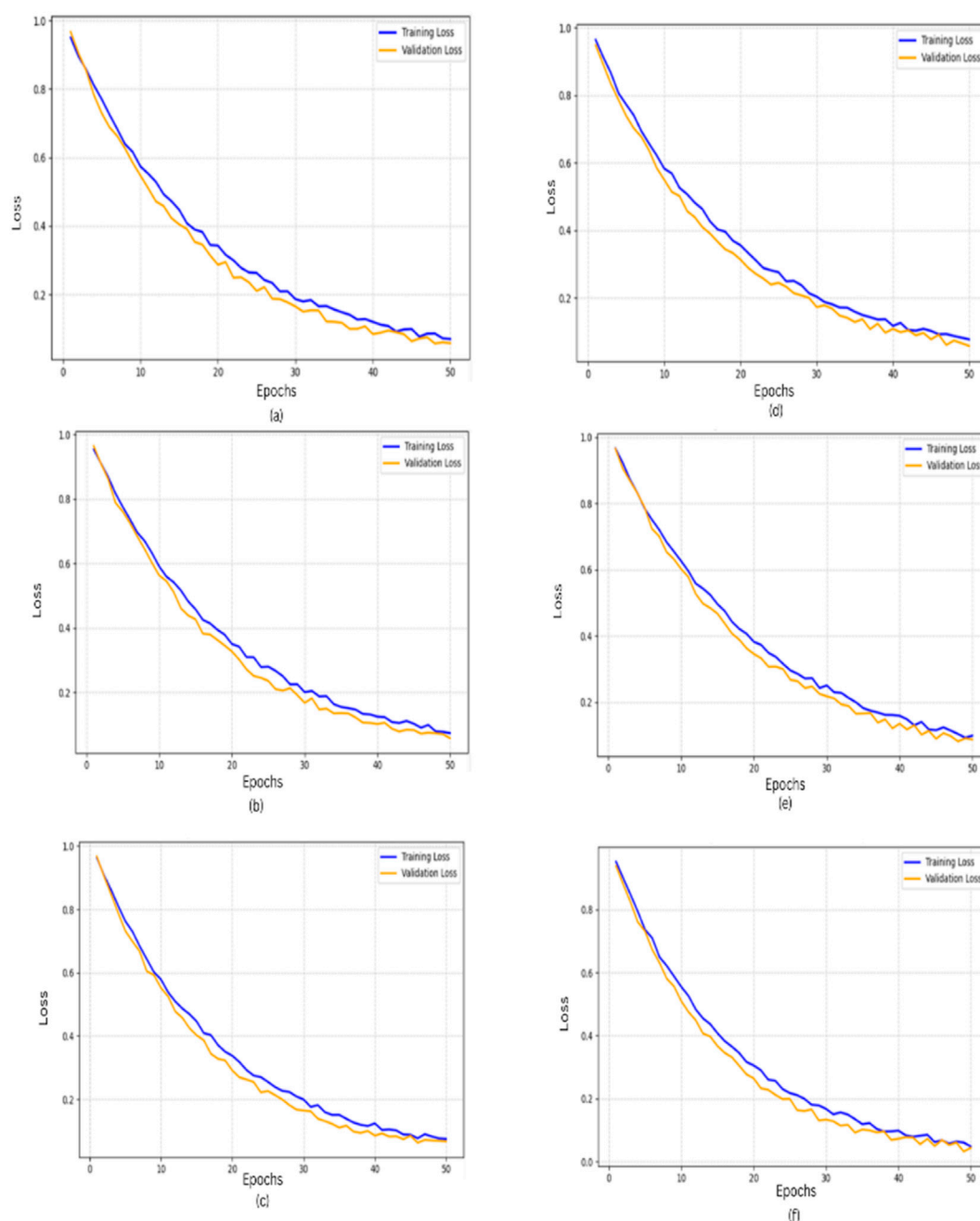| Hyperparameter	Minibatch size	Solver type	Learning rate factor	Filter size	Overall accuracy
SqueezeNet	83.76%	89.9%	0.0001	87.1%	86.92%
ShuffleNet	86.1%	86.7%	0.0001	87.4%	86.73%
ResNet 18	89.8%	80.5%	0.0001	83.3%	84.53%
MobileNet V2	87.9%	88.6%	0.0001	87%	87.83%
DenseNet 201	89.9%	88.9%	0.0001	88.5%	89.1%
EfficientNet B0	90.2%	87.3%	0.0001	90.2%	89.23%

FIGURE 5
Training and Loss Validation Curve at 50 Epochs **(a)** EfficientNet B0 **(b)** DenseNet 201, **(c)** MobileNet V2, **(d)** ResNet 18, **(e)** ShuffleNet and **(f)** SqueezeNet.

The effective performance variations observed in Table 4 across the pretrained networks can be defined by their architectural features compared to the characteristics of GASF images. EfficientNet B0 attains highest accuracy because of its compound scaling method, which balances the network's depth, width, and resolution. This balance allows the model to gain both global and local wear related features from GASF representations. EfficientNet B0 implements MBConv Mobile Inverted Bottleneck Convolution blocks and squeeze excitation modules to optimize the best use of the features and make it more sensitive to minor variations. These features are most useful for vibration based GASF images, where wear patterns are shown by minor variations in space and intensity (Ali et al., 2025). Lightweight models like SqueezeNet and ShuffleNet, apply aggressive parameter reduction approaches like fire modules and depthwise group convolutions these methods make it challenging for the models to learn detailed wear signatures which are present in 2D image data (Sandler et al., 2018). Deeper networks such as DenseNet 201 can represent more information, though they are more likely to overfit when there is not sufficient data due to the amount of feature connections increases. Their learning curves show
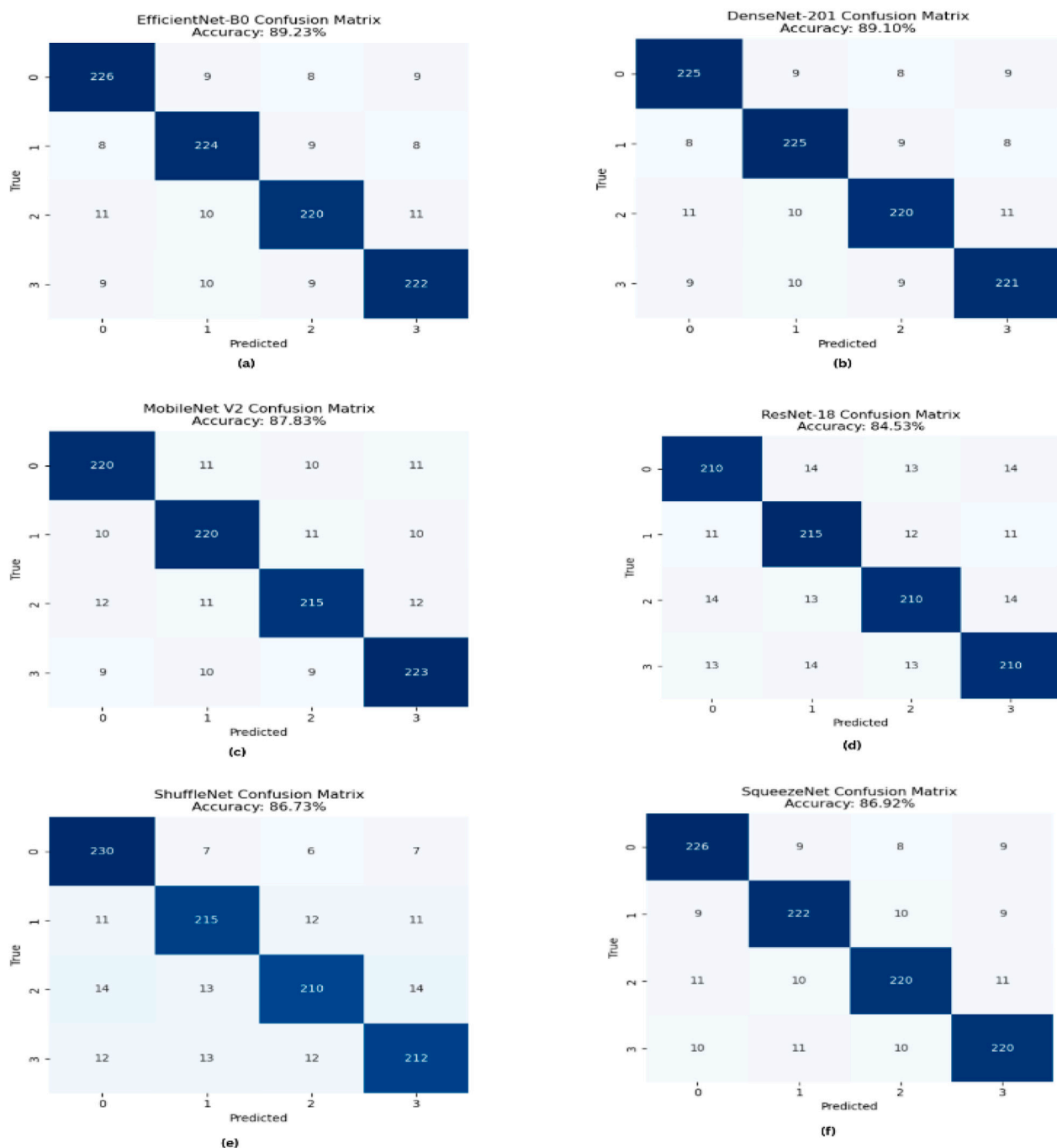
FIGURE 6
Confusion matrix analysis of the pretrained model **(a)** EfficientNet B0, **(b)** DenseNet 201, **(c)** MobileNet V2, **(d)** ResNet 18, **(e)** ShuffleNet and **(f)** SqueezeNet.

that they stabilize more slowly than EfficientNet B0 (Terzioğlu et al., 2025). ResNet 18 and MobileNet V2 performed moderately well using both residual and inverted residual structures (Ali et al., 2025). However, they did not employ the compound scaled feature extraction pipeline that makes EfficientNet B0 significantly more accurate and effective. The confusion matrix indicates that EfficientNet B0 can generalize better tool batches. To evaluate the model behaviour, training, and validation loss curves were analysed as shown in Figure 5. For the uniform comparison, all models were

trained for 50 epochs. But the loss curves for DenseNet 201, ShuffleNet, and SqueezeNet showed a down trend which become slightly unstable toward the end of training. This behavior indicates that these models benefited from extended epochs to attain full convergence. Preliminary tests with longer training times of 80 and 100 epochs showed that the accuracy improvement was usually less than 1% and did not change the ranking of model performance. EfficientNet B0 was always found to be the best network. The stopping criterion of 50 epochs was kept the same for all

TABLE 5 Comparison of attainment accuracy with other classifiers.

Classifiers	Classification accuracy in %
BayesNet	83.34
Decision tree	78.2
Naviebayes	85.28
K Star	78
Proposed method	89.23

architectures to make sure that the results were stable and representative.

EfficientNet B0 showed the most stable and balanced learning, indicating strong generalisation. Its residual connections helped maintain gradient flow, making it well suited for small datasets. In contrast, other models showed fluctuations or early convergence. Confusion matrix analysis confirmed EfficientNet B0 accuracy in distinguishing clear wear states, though it was slightly less effective with intermediate cases as shown in Figure 6.

5.6 Comparison with existing classifiers

Latest studies in tool wear class have employed a variety of ML and DL techniques to enhance projecting accuracy and strength. These approaches use diverse sensor data types, including vibration signals, acoustic emissions, and images, each involving separate feature extraction and classification strategies. Despite stated accuracy findings varying broadly across dissimilar studies due to variations in datasets, empirical setups, and evaluation protocols, making direct correlation of these metrics demanding and potentially deceiving (Wang et al., 2021). Therefore, this work points out a qualitative review of key methodologies and their respective strengths to aid the proposed approach. To enable a meaningful performance comparison, several baseline representative methods were applied and assessed them alongside with the proposed model using the same dataset and experimental settings (Y. Zhang et al., 2023; Li et al., 2024). The features consist of standard time domain descriptors such as RMS, peak value, crest factor, kurtosis, and skewness, as well as frequency domain characteristics that come from FFT analysis. This shows the common work in TCM, where the quality of handcrafted features has significant impacts on performance (González et al., 2022; Wang et al., 2021). The proposed pretrained CNN models were trained for GASF imaging, which makes it easier to extract the features. Table 5 provides a comparison of methods between traditional feature engineered classifiers and deep learning based feature learners, not a direct comparison using the same inputs. This distinction clarifies that the high performance of CNN models is due to their capability to learn discriminative representations directly from transformed vibration images, which reduces reliance on manual feature engineering (Y. Zhang et al., 2023; Wang et al., 2021).

Among the means tested, EfficientNe B0 showed the best performance. Traditional classifiers such as Decision Trees and Naive Bayes achieved accuracies of 78.2% and 85.28%, respectively, but were outperformed by the proposed model, which attained an accuracy of 89.23% using vibration signal plots combined with TL.

The generated dataset had 688 GASF images, but various design choices reduced the risk of overfitting. Transfer learning substantially decreased data requirements as pretrained CNNs possess broad, generalizable feature representations that were learned from the large ImageNet dataset. The final layers of the networks were tuned to perform efficiently with vibration based images, which makes the method effective with limited data. The tool batch level was used to split the data into 70/30 parts. This ensured that all of the test samples came from machining runs that had never been seen before. This method stops leaks and forces the models to go beyond just local signal similarities. Third, applying random rotations, flips, zooming, and normalization to the data set made the training samples diverse and made the model most reliable. The training and validation curves were not diverged and further processed without overfitting. The dataset was sufficient for testing the viability of the proposed method.

6 Conclusion

This study presents six pretrained DL models, SqueezeNet, ShuffleNet, ResNet 18, MobileNet V2, DenseNet 201, and EfficientNet B0, to assess the condition of single point cutting tool's wear through vibration signal plot images. The study focused on four definite wear states of the cutting tool, namely, no wear (indicating a good tool), nose wear, crater wear, and flank wear of cutting tool inserts. The utilized pretrained model, based on a CNN architecture, provides a complete monitoring of the cutting tool's condition by including feature extraction, selection and classification into an integrated framework, which gives an effective classification of the vibration plots. The experimental findings show that all the proposed models learned complex features and resulted in consistent classification for tool condition monitoring. The hyperparameters train test split ratio, optimizer, learning rate, and batch size were optimized for accuracy for the model's requirements. Of all the models EfficientNet B0 achieved the highest classification accuracy of 89.23% closely followed by DenseNet 201 with 89.1%, MobileNet V2 at 87.83% with less computation time SqueezeNet and ShuffleNet acchived the accuracy of 86.92% and 86.73%. ResNet 18 was the least performing model with an accuracy of 84.53%. Due to exceptional accuracy, minimal computational complexity, and proficiency in managing complex feature learning and achieving higher classification accuracy on this comparative analysis, EfficientNet B0 is recommended for real time monitoring of the condition of the cutting tool. Applications of a pretrained model in an industrial environment can achieve an accuracy and high oversight and reduce machine downtime, which can increase productivity. This approach aligns well with the modern industrial requirement where real time monitoring and predictive maintenance are crucial.

6.1 Future work

Current studies on tool wear classification have used various machine learning and deep learning methods with separate sensor inputs, but reported accuracies are often not directly comparable due to variations in datasets and setups. This work shows the effectiveness of using vibration signal plots with pretrained CNN

models. However, the current approach uses a single data processing pipeline. Future work will explore alternate techniques, such as wavelet decomposition, statistical features, and raw signal input to 1D CNNs or LSTMs, to improve robustness and provide a more comprehensive evaluation.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

SK: Investigation, Writing – review and editing, Supervision, Writing – original draft, Software, Validation, Visualization, Data curation, Resources, Methodology, Formal Analysis, Conceptualization, Project administration. SG: Methodology, Data curation, Visualization, Project administration, Conceptualization, Investigation, Funding acquisition, Validation, Formal Analysis, Writing – review and editing, Supervision.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abid Al-Sahib, K., and Bachaa, A. M. (2005). Tool wear monitoring in turning operation using vibration and strain measurement with neural network. *Nabeel.* 783–792. doi:10.1115/imece2005-80699

Ali, H., Shifa, N., Benlamri, R., Farooque, A. A., and Yaqub, R. (2025). A fine tuned EfficientNet-B0 convolutional neural network for accurate and efficient classification of apple leaf diseases. *Sci. Rep.* 15. 125732. doi:10.1038/s41598-025-04479-2

Alonso, F. J., and Salgado, D. R. (2008). Analysis of the structure of vibration signals for tool wear detection. *Mech. Syst. Signal Process.* 22 (3), 735–748. doi:10.1016/j.ymssp.2007.09.012

Bagavathiappan, S., Lahiri, B. B., Saravanan, T., Philip, J., and Jayakumar, T. (2013). Infrared thermography for condition monitoring - a review. *Infrared Phys. Technol.* 60, 35–55. doi:10.1016/j.infrared.2013.03.006

Dutta, S., Pal, S. K., and Sen, R. (2016). "*On-Machine tool prediction of flank wear from machined surface images using texture analyses and support vector regression,*" Elsevier, 43. 34–42. doi:10.1016/J.PRECISIONENG.2015.06.007

Ekundayo, O. S., and Ezugwu, A. E. (2025). Deep learning: historical overview from inception to actualization, models, applications and future trends. *Appl. Soft Comput.* 181, 113378. doi:10.1016/J.ASOC.2025.113378

Elhefnawy, M., Ahmed, R., and Mohamed, S. O. (2022). Fault classification in the process industry using polygon generation and deep learning. *J. Intelligent Manuf.* 33, 1531–1544. doi:10.1007/s10845-021-01742-x

Fu, Y., Zhang, Y., Qiao, H., Li, D., Zhou, H., and Leopold, J. (2015). Analysis of feature extracting ability for cutting state monitoring using deep belief networks. *Procedia CIRP* 31, 29–34. doi:10.1016/j.procir.2015.03.016

González, D., Alvarez, J., Sánchez, J. A., Godino, L., and Pombo, I. (2022). Deep learning-based feature extraction of acoustic emission signals for monitoring wear of grinding wheels. *Sensors* 22. 6911. doi:10.3390/S22186911

Guo, X., Chen, L., and Shen, C. (2016). Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. *Measurement* 93, 490–502. doi:10.1016/J.MEASUREMENT.2016.07.054

Jardine, A. K. S., Lin, D., and Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech. Syst. Signal Process.* 20, 1483–1510. doi:10.1016/J.YMSSP.2005.09.012

Kasiviswanathan, S., Gnanasekaran, S., Thangamuthu, M., and Rakkiyannan, J. (2024). Machine-Learning- and internet-of-things-driven techniques for monitoring tool wear in machining process: a comprehensive review. *J. Sens. Actuator Netw.* 13, 53. doi:10.3390/jsan13050053

Katta, P., Karunanithi, K., Raja, S. P., Ramesh, S., Vinoth John Prakash, S., and Joseph, D. (2024). Optimized deep belief network for efficient fault detection in induction motor. *Adv. Distributed Comput. Artif. Intell. J.* 13, e31616. doi:10.14201/adcaij.31616

Li, Y., Zhao, Z., Fu, Y., and Chen, Q. (2024). A novel approach for tool condition monitoring based on transfer learning of deep neural networks using time–frequency images. *J. Intelligent Manuf.* 35, 1159–1171. doi:10.1007/s10845-023-02099-z

Mohanraj, T., Shankar, S., Rajasekar, R., Sakthivel, N. R., and Pramanik, A. (2020). Tool condition monitoring techniques in milling Process-a review. *J. Mater. Res. Technol.* doi:10.1016/j.jmrt.2019.10.031

Ochoa, E., Enrique, L., Ruiz Quinde, I. B., Sumba, J. P. C., Guevara, A. V., and Morales-Menendez, R. (2019). New approach based on autoencoders to monitor the tool wear condition in HSM. *IFAC-PapersOnLine* 52, 206–211. doi:10.1016/j.ifacol.2019.09.142

Okada, M., Hosokawa, A., Tanaka, R., and Ueda, T. (2011). Cutting performance of PVD-coated carbide and CBN tools in hardmilling. *Int. J. Mach. Tools Manuf.* 51 (2), 127–132. Pergamon. doi:10.1016/J.IJMACHTOOLS.2010.10.007

Salim, F., Saeed, F., Basurra, S., Qasem, S. N., and Al-Hadhrami, T. (2023). DenseNet-201 and xception pre-trained deep learning models for fruit recognition. *Electron. Switz.* 12, 3132. doi:10.3390/electronics12143132

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L. C. (2018). "MobileNetV2: inverted residuals and linear bottlenecks," in *Proceedings of the IEEE computer society conference on computer vision and pattern recognition, January*. IEEE Computer Society, 4510–4520. doi:10.1109/CVPR.2018.00474

Silva, R. G., Wilcox, S. J., and Reuben, R. L. (2006). Development of a system for monitoring tool wear using artificial intelligence techniques. *Proc. Institution Mech. Eng. Part B J. Eng. Manuf.* 220 (8), 1333–1346. doi:10.1243/09544054JEM328

Terzioğlu, H., Gölcük, A., Shakarji, A. M. A., and Al-Bayati, M. Y. (2025). Comparative analysis of deep learning-based feature extraction and traditional

classification approaches for tomato disease detection. *Agron.* 15. 1509. doi:10.3390/AGRONOMY15071509

Verstraete, D., Ferrada, A., Droguett, E. L., Meruane, V., and Modarres, M. (2017). Deep learning enabled fault diagnosis using time-frequency image analysis of rolling element bearings. *Shock Vib.* 2017, 1–17. doi:10.1155/2017/5067651

Wang, Q., Wang, H., Hou, L., and Yi, S. (2021). Overview of tool wear monitoring methods based on convolutional neural network. *Appl. Sci.* 11, 12041. doi:10.3390/app112412041

Zhang, X., Zhou, X., Lin, M., and Sun, J. (2017). "ShuffleNet: an extremely efficient convolutional neural network for Mobile devices," in *Proceedings of the IEEE computer society conference on computer vision and pattern recognition, July*. IEEE Computer Society, 6848–6856. doi:10.1109/CVPR.2018.00716

Zhang, Y., Qi, X., Wang, T., and He, Y. (2023). Tool wear condition monitoring method based on deep learning with force signals. *Sensors*, 23, 4595. doi:10.3390/S23104595

Nomenclature

CNC   Computer Numerical Control

DL   Deep Learning

TCM   Tool Condition Monitoring

ART   Adaptive Resonance Theory

SOM   Self Organizing Maps

ML   Machine Learning

DTs   Decision Trees

IoT   Internet of Things

SVM   Support Vector Machine

DAQ   Data Collection System

TL   Transfer Learning

CNN   Convolutional Neural Network

FC   Fully Connected Layer

CONV   Convolution