Sādhanā (2021)46:130
https://doi.org/10.1007/s12046-021-01654-9

Tool wear classification based on machined surface images using convolution neural networks

M PHANI KUMAR, SAMIK DUTTA* and N C MURMU

Surface Engineering and Tribology Group, CSIR-Central Mechanical Engineering Research Institute, Durgapur 713 209, India
e-mail: s_dutta@cmeri.res.in

MS received 24 September 2020; revised 3 May 2021; accepted 29 May 2021

Abstract. Among several factors that are having a profound impact on the overall machining process efficiency, cutting tool wear is the most significant one. Monitoring and identification of cutting tool wear state well before to its failure is important to achieve superior machining quality and profitable production. With the recent advancements in computational hardware, significant amount of research is being carried out on using deep learning techniques, in specific, convolution neural networks (CNN) for developing cutting tool wear monitoring system. Although, few researchers reported the use of CNN as a pathway to tool wear classification problems with significant results, the fundamental methodology adopted by these techniques still needs to be investigated. Hence, in the present work, a deep CNN architecture is designed by choosing appropriate hyper-parameters and a CNN model is developed by selecting proper training parameters for cutting tool wear classification. Machined surface images acquired during turning operation performed on mild steel components under dry condition by uncoated carbide inserts as cutting tool are used as input data to the CNN model for predicting the tool condition. The proposed model, whose classification performance is independent of machining conditions, has capability to extract the features and classify the cutting tool among the two classes (i.e., unworn and worn classes). Accuracies of 96.3% and 99.9% are realized for classification of tool flank wear from raw and minimally pre-processed (contrast enhanced) machined surface images, respectively.

Keywords. Convolution neural network; machined surface images; tool wear classification; deep learning; hyper-parameters selection.

1. Introduction

Advent of Industry 4.0 in machining domain requires to transform conventional machining to intelligent machining. Intelligent machining should consist of various sensors, signal or image processing to obtain semantic features and artificial intelligence techniques for adaptive decision making for giving feedback to the process and systems for automatic control. Incorporation of intelligent machining enables more efficient machining by reducing machine tool downtime as well as optimizing the machining process. Process planning and process monitoring are two major verticals of intelligent machining and play an important role to incorporate automation [1]. Process monitoring involves usage of various sensors to collect useful dynamic information in terms of signals or images from the manufacturing process and then semantic decision is made as feedback to the process for subsequent control operation. It is an essential part to increase productivity, product quality and energy efficiency in machining, which ultimately reduces the product cost [2, 3]. As the quality of machined surface is highly dependent on condition of the cutting tool, observing its state is an important part in process monitoring. Productivity can be improved by timely identification of tool wear state well before its failure. Methods of monitoring tool wear can be broadly categorized into two streams: direct and indirect methods. In the former approach, cutting tool wear is measured directly using microscope or by machine vision techniques. However, use of microscope for in-situ monitoring of cutting tool wear is not a feasible solution in practise as it requires frequent removal of cutting tool which may cause misalignment and ultimately affects product quality. These disadvantages, upto certain extent can be overcome by knocking machine vision into practice. Although reported research [4, 5] presents application of machine vision for direct monitoring of cutting tool wear, inaccessibility of tool wear region by image acquisition system, in most of the machine tools, is a major limitation to practical applicability of this technique.

Indirect methods, on the other hand, rely on sensors, such as force [6], acoustic emissions [7], vibration [8], temperature [9], machine vision, to acquire corresponding signals

*For correspondence
Published online: 03 July 2021

for prediction of tool wear from the acquired signals using subsequent signal processing and feature extraction techniques. Tool wear estimation by processing machined surface images has an advantage as it is a non-contact and flexible technique for monitoring. Also, 2-D information from machined surface images can be extracted by incorporation of machine vision in monitoring. Texture analysis was mainly applied to extract semantic texture features for detecting tool condition [10–19]. Mannan *et al* [11] used radial basis function neural network (RBFNN) to predict tool flank wear using features extracted from machined surface images and sound signals by performing statistical texture analysis and discrete wavelet transform (DWT) for non-stationary signals, respectively. Kassim *et al* [12] predicted tool flank wear in turning and milling operations with the use of multi-layer perceptron neural network (MLPNN), a supervised machine learning technique. Input features to the network were extracted, in their research, from the machined surface images using connectivity-oriented Hough transform, i.e., a kind of geometrical texture analysis technique. In their another research, fractal features were extracted from turned surface images after performing fractal analysis and then the features were utilized in an unsupervised machine learning technique i.e., Hidden Markov model (HMM) for classification of tool condition [13]. Bhat *et al* attempted to classify tool condition in three states by extracting statistical texture features after performing gray level co-occurrence matrix (GLCM) based texture analysis on turned surface images. They explored the possibility of automated feature selection using Fisher's discriminant ratio (FDR). These features were then used in a supervised machine learning technique i.e. support vector machine (SVM) in [14] and unsupervised learning i.e. hidden Markov model (HMM) which were reported in [15], to classify the tool condition as unworn, semi-worn and worn states.

Use of machine learning enabled to utilize fusion of features extracted from various texture analysis techniques, so that variation in waviness or micro-texture, roughness and even changes in feed marks on the machined surfaces due to an increase in tool flank wear can be utilized in tool condition monitoring. For instance, Kassim *et al* [16] performed fusion of statistical and geometrical texture features extracted from turned surface images for tool condition classification by using Mahalanobis distance classifier. Dutta *et al* [17, 18] extracted features from turned surface images by using statistical, geometrical and signal processing based texture analyses to extract change in waviness, feed mark and roughness base features, as all these three are affected by an increase in tool wear. Then these features were utilized for prediction of tool flank wear using support vector machine-based regression (SVR) technique. In a similar manner, flank wears of coated and uncoated milling tools were performed in [19].

From the above-mentioned literature, it is observed that to analyze the tool condition using machined surface images, several complex or computationally expensive texture analyses techniques were used to extract features from the acquired images. Those features were then used to detect tool flank wear using machine learning techniques. A single set of extracted features may not be useful with good predictability for variations in machining process parameters. The effectiveness of the classifier in deriving the tool condition majorly depends on these extracted features. Two points are most important at the time of feature(s) extraction or selection in tool condition monitoring. Firstly, it should be kept in mind that the extracted features should reflect the changes in the cutting tool with high precision. Secondly, the features should be sensitive to the wide range of process parameters in tool condition detection through a single machine learning model, which is really a challenging task. In order to reduce the effort of identifying the most sensitive features and to gain wide applicability of these monitoring methods in industrial environment, researchers used data driven models to automate the feature selection process. In this context, automatic selection of parameters for image filtering, feature extraction, selection and decision making are time consuming and complex with shallow learning techniques, viz. artificial neural network (ANN), MLPNN, RBFNN, SVM, SVR, etc. Therefore, machine learning techniques mentioned above are less generalized solutions and its performance are not invariant to machining conditions. On the other hand, there is a way to avoid handcrafted feature selection to make the prediction model more generalized by using convolutional neural network (CNN) approach, a type of deep learning technique.

Convolutional neural network (CNN) was first developed by LeCun *et al* in 1998 [20]. It has the ability to update the parameters jointly and automatically. In CNN, weights i.e., parameters of image filtering, edge detection, feature extraction, feature selection and decision making are updated due to its inherent architecture. Though multi-level data representation architecture of CNN requires more computational time on training, but the advent of advanced computing hardware, like graphics processing unit (GPU), makes it possible [21]. Multiple layer based data representations for filtering and non-linear mapping, realizing various filtered output in a single layer due to the presence of local connection property, detection of similar local patterns due to sharing of weights, matching of similar local patterns through pooling and updates of a large number of weights or filter parameters, automatically, are some of the major beneficial features of CNN for extracting higher abstract level information from raw, unscaled or rotated images [22]. As machined surface images require several processing from pre-processing to feature extraction and decision making due to the non-stationary nature, those multiple level representations are indispensable to realize an end-to-end solution. Due to these advantages, there are already various involvement of CNN to learn precise information from non-stationary raw images in applications

viz. object recognition, image classification, content based image retrieval, human pose analysis to name a few [23]. Implementation of intelligent process monitoring in smart manufacturing is a challenge as accurate extraction of proper information from high volume, veracity, variety and high speed data is required, which can be possible through deep learning techniques, due to its more generalization capability [24].

Hence, in the present work, classification of fresh and worn tool from machined surface images is performed by using CNN. A CNN architecture is proposed here by careful selection of hyper parameters, learning rate, batch size and number of epochs for accurate and generalized classification of tool wear states. Surface images acquired during the turning operation are provided as an input to the designed CNN. It is observed that developed CNN model performance is independent of machining conditions i.e., the variation in machining parameters. Accuracies of 96.3% and 99.9% are realized for classification of tool flank wear from raw and minimally pre-processed machined surface images, respectively.

2. Experimental set-up

In this work, a workpiece of AISI 1050 steel with 400 mm length is turned in a computer numerically controlled (CNC) machine (Ace designers made LT20) by using uncoated carbide insert, SNMG 120408-QM in dry conditions. Turning operations are performed at nine machining conditions by varying the cutting speed, feed rate and depth of cut, as shown in table 1.

After machining with each machining condition, gray-level images of the turned surfaces are acquired using an image acquisition system consisting of a CMOS camera (Dalsa, GenieHM1024), Navitar7000, 2/3″ format macro zoom lens and fibre optic (FO) guided illumination system with DC regulated quartz halogen lamp (Make: Navitar).

Table 1. Machining conditions for training and testing.

Machining conditions	Feed rate (mm/rev)	Depth of cut (mm)	Cutting speed (m/min)
Machining conditions for creating training data			
1	0.2	1	150
2	0.24	1	150
3	0.24	1.5	150
4	0.24	2	150
5	0.28	1	110
6	0.28	1.5	150
Machining conditions for creating testing data			
7	0.2	1.5	150
8	0.24	1.5	110
9	0.28	1	150

Figure 1 depicts the experimental set-up engaged for acquiring machined surface images.

A digital video recording software (Streampix V5.6) is used to acquire uncompressed images in .bmp format with 430 frames per second, where the workpiece is rotated with 50 rpm. Machined surface images of size 200 × 200 pixels (3.6 mm × 3.6 mm) are acquired by with a lens aperture, i.e., F-stop value of 5.6, and focal length of 108 mm with dark-field front illumination.

In order to create a training data set to develop the CNN model, machined surface images acquired from six machining conditions, as depicted in table 1, are used. The model trained is tested using the images obtained from the remaining three machining conditions. This methodology is followed to check the generalization performance of the CNN model, i.e., to check the model dependency on the machining conditions.

In each of the machining experiment, turning operation is performed on the workpiece for a length of 200 mm and average tool flank wear ($VB_{average}$) at zone B is measured by using Leica S6D stereo zoom microscope to know its state. The cutting tool is considered unworn if $VB_{average}$ is less than or near to 100 μm and is considered as worn if the $VB_{average}$ is more than or near to 300 μm according to ISO 3685:1993 [25]. In each of the tool state i.e., unworn and worn, 1290 images of turned surfaces are acquired. At machining condition 3, $VB_{average}$, at the first and second instances are found as 73 μm and 102 μm. In this situation, 1290 images per instance are acquired and thus, a total of 2580 images are utilized for unworn tool class. Similar procedure has been followed with other machine conditions where such similar situation arises. The number of images acquired for each machining condition are presented in table 2.

Hence, a total of 28380 images are acquired for all the machining conditions. From these images, 20640 number of images (10320 images against each of the unworn and worn class) are separated for training and the remaining images, i.e., 7740 images (3870 images against each of the unworn and worn class) are used to test the model.

3. Methodology

The aim of this work is to classify the condition of cutting tool into two classes, i.e., unworn and worn using machined surface images. Machined surfaces resulting from unworn tool produces more specular reflection (i.e. less brightness) in dark-field front illumination system, because of their smoothness when compared to surfaces resulting from the worn tool [17, 18]. Also, the feed marks of machined surfaces resulting from worn tool are more discontinuous than that of the surfaces resulting from unworn or fresh tool. Due to these discontinuities, local patches of similar patterns are less in surfaces machined by worn tool than that of

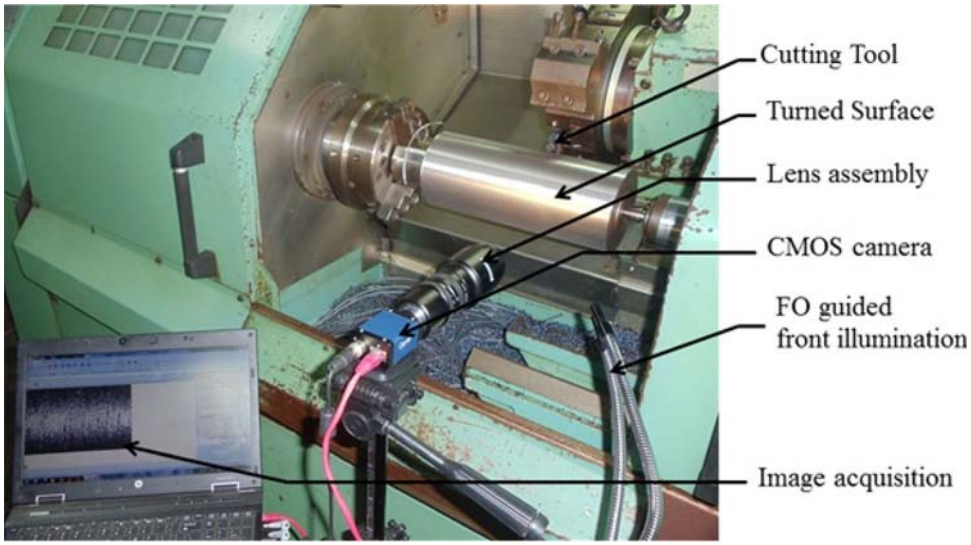

Figure 1. Experimental set-up for image acquisition.

Table 2. Number of acquired images against machining condition.

Machining condition	Number of machined surface images acquired	
	Unworn class	Worn class
1	1290	1290
2	1290	1290
3	2580	2580 ($VB_{average}$ are 299 μm and 340 μm)
4	2580	1290
5	1290	2580 ($VB_{average}$ are 291 μm and 330 μm)
6	1290	1290
7	1290	1290
8	1290	1290
9	1290	1290

the surfaces produced by unworn tool. Also, the non-stationary features becomes higher in the surface images with the increase in tool flank wear [18]. Attempts to detect these non-stationary features using non-linear function based DWT have already been mentioned in [26–30]. Therefore, local spatial similarity, change in feed marks and non-stationary feature information along with machine learning methods were used often in tool condition monitoring. CNN models, on the other hand, are capable of detecting local similarity and changes in feed marks through convolution, non-stationary information extraction by using non-linear activation function and multi-resolution processing due to multiple layer operations. So, with the use of CNN, accurate and generalized tool wear classification with an end-to-end solution suitable for industrial environment can be achieved. Hence, in the present work CNN is used for classifying the tool flank wear state.

3.1 *Architecture of CNN*

In the present work, CNN is used as a computing architecture for the tool wear classification. CNN automatically generates meaningful features for a specific task in an evolutionary way directly from huge amounts of raw data with minimal human interaction [31]. Figure 2 presents the architecture of the CNN used in this work for the tool wear classification from turned surface images. Each layer of CNN consists of convolution, activation function and pooling as the basic building blocks.

3.1.1 *Convolution layer:* The primary purpose of convolution operation is to extract features from the input image using a square shaped matrix called as filters or kernel or feature detectors. Using these filters, convolution operation captures the local dependencies in the original image. With a greater number of filters in the convolution, more image features get extracted and the network becomes better in recognizing patterns in test images. Convolution operation is performed with an activation function and some bias as expressed in Eq. (1) [32].

$$M^{l}_{i-s,j-s} = \varphi\left(\sum_{i,j=0}^{F,G} f_{i,j} M^{l-1}_{i,j}\right) \tag{1}$$

where, a filter, $f_{i,j}$, having filter size of $F \times G$, is convolved with an input matrix, $M^{l-1}_{i,j}$, present at $(l-1)^{th}$ layer, by a stride length, s. Then an activation function, φ, is used on the convolved output to extract higher abstract features, $M^{l}_{i-1,j-1}$ of layer l, considering no bias and zero

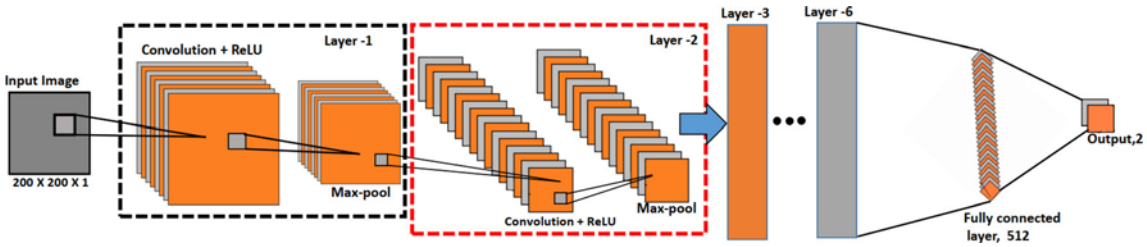

Figure 2. Architecture of CNN [24].

padding. In this work, stride length and filter size are considered as 1 and 3 ×3, respectively, for accurate identification of features.

Convolution is a linear operation, but machined surface images contain non-stationary information [18], so there is a need of some non-linear activation function to extract non-linear information. In CNN, sigmoid, tan hyperbolic, rectified linear unit (ReLU), to name a few, are used as non-linear activation functions. On the other hand, choice of the activation function is crucial to obtain computationally efficient and repeatable classification model. Sigmoid and tan hyperbolic functions are more computationally expensive in nature as compared to ReLU [33]. Therefore, in this work, ReLU is chosen as an activation function for all the convolution layers. ReLU function is expressed in Eq. (2).

$$\varphi(I) = Max(I, 0) \tag{2}$$

where I and $\varphi(I)$ are same as $\sum_{i,j=0}^{F,G} f_{i,j} M_{i,j}^{l-1}$ and $M_{i-s,j-s}^{l}$, respectively, as expressed in Eq. (1). ReLU chooses only the positive values and make negative input to zero, so its derivatives and the function itself both are monotonic. Number of filters, in convolution layer, is another hyper parameter to extract appropriate abstract level information for creating accurate and repeatable CNN model. Section 3.2 describes the selection process of appropriate number of filters.

3.1.2 *Pooling layer:* Pooling, also known as subsampling or down sampling is used to reduce the spatial size of the activation map without losing important information. Pooling operation makes the network invariant to small transformations, distortions and translations in the input image. Max, average and sum are the widely used pooling operations where maximum, mean and summation of the values resides in pooling window are performed, respectively. In general, a 2 × 2 pooling window with stride 1 is performed in CNN. Since the present work aims to classify the cutting tool state from the machined surface images, the difference between pixel values, due to excessive tool wear, over a small neighbourhood region should be detected and pixel value differences due to noise should be avoided. Therefore, max pooling operation is adopted in this work, as average or sum pooling can consider the influence of noise also.

3.1.3 *Dropout layer:* Use of more neurons, i.e., number of filters, in a deep network, sometimes affects the generalization performance of the model. Overfitting may occur due to presence of huge number of neurons [34]. Training time also increases due to high dimensional data. So, to avoid this situation, Srivastava *et al* [35] introduced dropout layer to drop some of the nodes (here 65% is chosen by hit and trial method), randomly, at the time of training to reduce overfitting and data dimensionality problem.

3.2 *Training*

During training of CNN, cost function, which comprises of loss function and regularization, is required to minimize by updating of filter values or weights, $f_{i,j}$. The training starts with some initial weight values. Then it feeds the values layer after layer in forward direction to classify the input images and a loss is calculated by using a loss function, L. To calculate the loss function, target labels (here the unworn and the worn tool) of class 1 and class 2 are predefined against each input image. In this work Cross-entropy is used as loss function for binary classification, which is expressed in Eq. (3) [32].

$$\text{Cross - entropy} = -(y log(p) + (1 - y) log(1 - p)) \tag{3}$$

where y and p are the class label and prediction probability, respectively.

A regularization term, R, along with learning rate, l_r, are added in the cost function to penalize the higher deviations and to avoid overfitting problem [32, 34]. In this work, $L2$ regularization is performed. In the present problem minimization of an average cost function, C_B, over a batch of images, B, (Eq. (4)) is to be performed to update filter values or weights via each epoch.

$$C_B = \frac{1}{B} \sum_{input} (L + l_r R) \tag{4}$$

This is ultimately performed via back propagation approach [32] to obtain appropriate node parameters or

weights. For optimization, an appropriate optimizer should be used to obtain precise model of classification. In this work, stochastic gradient descent (SGD) technique for non-convex optimization is used as it can work on batches with less computational time. Hence, in CNN, appropriate selection of hyper-parameters, viz. activation functions, number of convolution and pooling layers, number of filters in each convolution layer, learning rate, optimizer, batch size, number of epochs etc. are important to realize an appropriate model.

In the present work, for all the convolution and dense layers, ReLU is chosen as the activation function because of its computationally inexpensive nature. Subsequently, Softmax function is chosen for binary classification. Choice of filters in first convolution layer is performed along with the number of convolution and pooling layers. After each convolution layer, one pooling layer is applied to extract proper information without noise by the use of Max pooling technique. After each pooling, the filter number is doubled to extract more abstract information. The strategy to select number of filters and layers is to minimize testing error with less computational time. To realize this, a total 24 computational experiments are performed by varying number of layers and number of filters, by keeping learning rate, batch size and number of epochs as 0.001, 30 and 20, respectively. Number of flatten layer, pair of dropout and dense layer are chosen as 1 and 2, respectively, after last pooling layer. The result of the computational experiments is summarized in table 3.

It can be depicted from table 3 that appropriate number of layers and filters are 6 and 8, respectively, due to a trade-off between computational time (0.1486 s) and test accuracy (0.9947 or 99.47%). Accordingly, the architecture of CNN adopted in the present work is shown in table 4.

On deciding the CNN architecture, computational experiments were performed to choose appropriate learning rate, batch size and number of epochs for training the selected architecture. Selection of proper learning rate is important to avoid under-fitting with less computational time. Keeping batch size and epoch as 60 and 30, respectively, learning rate is chosen by varying it among the values of 0.01, 0.0015, 0.002 and 0.0025. Further, selection of batch size is carried out by considering batch sizes of 30, 40 and 60. The range of batch size values are chosen such that they are divisible to the total number of training samples i.e., 20640, so that not a single image can fall outside the training data. Finally, number of epochs is chosen by varying it as 50, 100, 150, 200 and 250 with appropriate learning rate and batch size.

4. Results and discussion

In the present work, raw and pre-processed machined surface images are used to train the proposed CNN model. A minimal pre-processing using contrast limited adaptive histogram equalization (CLAHE) is performed to overcome inhomogeneous equalization issue. Comparison of accuracy between the classification models using raw and pre-processed images is performed here to check the robustness of the developed architecture.

Computational experiments for learning rate selection, batch size selection and epoch selection are first performed to realize appropriate training parameters. In selection of learning rate, a fixed batch size and epoch of 60 and 50, respectively, is used by varying learning rate from 0.001 to 0.0025 with a step of 0.0005. Pre-processed images are used as an input for these experiments.

Table 3. Summary of computational experiments to select layer and number of filters.

No. of. layers	No. of filters	Training loss	Training accuracy	Testing accuracy	Elapsed time (s)
3	2	0.2780	0.9875	0.5215	0.1286
	4	8.2650	0.5000	0.5	0.1377
	8	8.2539	0.5000	0.5	0.1347
	16	0.2066	1.0000	0.5947	0.1456
4	2	8.2512	0.5000	0.5	0.1315
	4	0.2567	0.9879	0.7770	0.1374
	8	0.2222	0.9987	0.4340	0.1336
	16	0.2756	0.9973	0.5175	0.1362
5	2	0.1695	0.9852	0.6659	0.1290
	4	0.1712	0.9976	0.9160	0.1276
	8	0.1636	1.0000	0.9821	0.1326
	16	0.1835	0.9999	0.9214	0.1502
6	2	0.2882	0.9138	0.7271	0.1250
	4	0.0471	0.9996	0.9407	0.1304
	8	**0.0724**	**0.9997**	**0.9947**	**0.1486**
	16	0.1157	0.9987	0.9831	0.1845

Bold values indicate as CNN parameters

Table 4. Architecture of CNN.

Layer no.	Operation	Activation function/Pooling strategy	Number of filters/nodes	Filter size/Pooling window
1.	Convolution	ReLU	8	3 × 3
2.	Pooling	Max	–	2 × 2
3.	Convolution	ReLU	16	3 × 3
4.	Pooling	Max	–	2 × 2
5.	Convolution	ReLU	32	3 × 3
6.	Pooling	Max	–	2 × 2
7.	Convolution	ReLU	64	3 × 3
8.	Pooling	Max	–	2 × 2
9.	Convolution	ReLU	128	3 × 3
10.	Pooling	Max	–	2 × 2
11.	Convolution	ReLU	256	3 × 3
12.	Pooling	Max	–	2 × 2
13.	Dropout	65% dropout		
14.	FC	ReLU	512	
15.	Dropout	65% dropout		
16.	FC	Softmax	2	

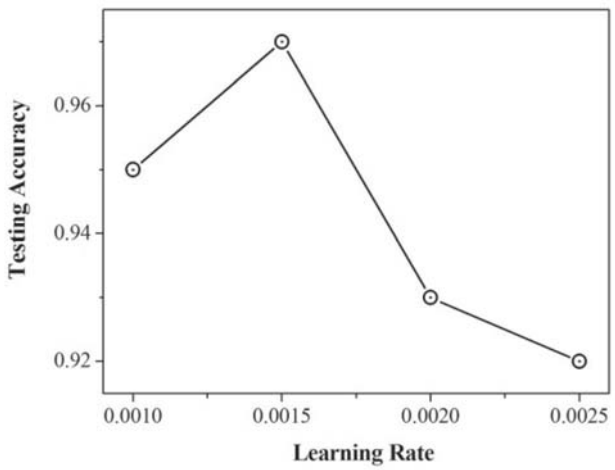

Figure 3. Variation of testing accuracy with learning rate.

Table 5. Test accuracy achieved by varying batch size and number of epochs with a learning rate of 0.0015.

No. of Epochs	Batch size	Test accuracy
50	30	0.75
	40	0.94
	60	0.97
100	30	0.97
	40	0.98
	60	0.93
150	30	0.98
	40	0.97
	60	0.92
200	**30**	**0.999**
	40	0.98
	60	0.97
250	30	0.98
	40	0.96
	60	0.95

Bold values indicate as CNN parameters

It can be depicted from figure 3 that maximum test accuracy is obtained as 0.97 at 0.0015 learning rate. Therefore, 0.0015 learning rate is selected for training. Now, to select batch size and number of epochs, both are varied as mentioned in section 3, and the results of test accuracy are shown in table 5.

It can be observed from the table 5 that the best test accuracy, i.e., 0.999 or 99.9% is achieved using 30 batch size and with 200 number of epochs, hence, these values are adopted for training the CNN model. After selection of the CNN architecture and the parameters, training is performed to obtain a classification model, where the classification accuracy and loss for pre-processed test images are found as 99.9% and 0.0021, respectively. The high accuracy has been achieved as because accurate noise filtering, segmentation, edge gradient detection, and then higher-level abstract information can be possible to extract due to the deep architecture of CNN. Further, for a comparison, raw images are also used to train the same CNN architecture using the same training parameters. A test accuracy of 96.3% and loss of 0.1548 are realized, when the CNN model is trained with raw images.

As an example, figure 4 presents the images of the surface machined using unworn and worn tool with machining conditions mentioned at sl. no.5 as presented in Table 2.

It is well known that machining with unworn tool results in surface with less roughness and surface roughness increases with an increase in tool wear.

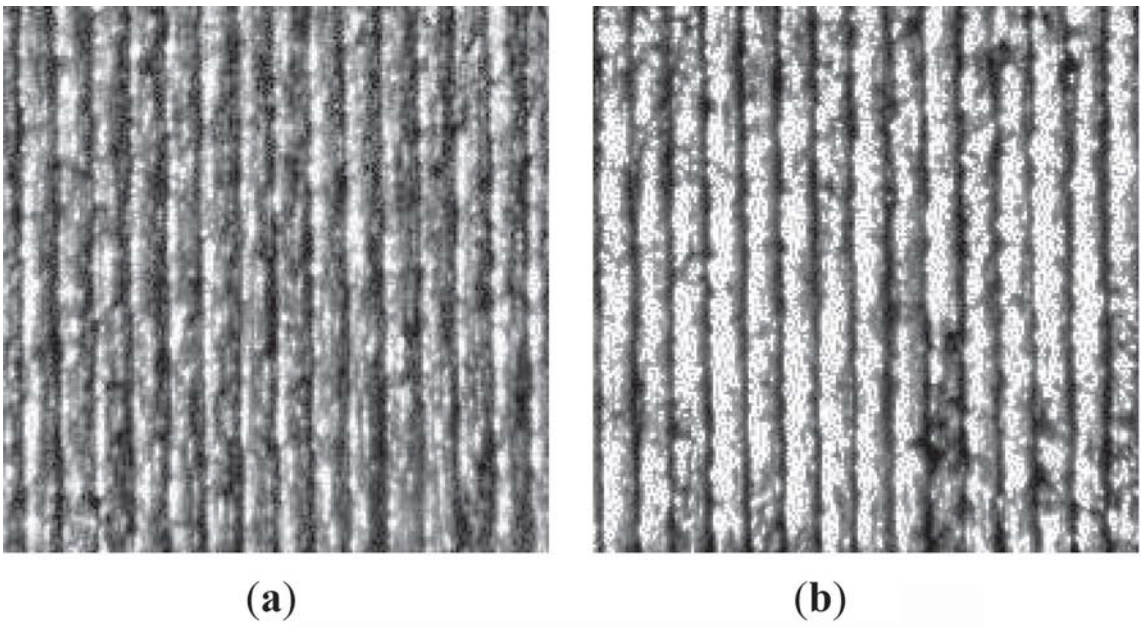

Figure 4. Turned surface images by (**a**) unworn tool, (**b**) worn tool using machining condition no. 5.

Surfaces resulting from unworn tool not only have low roughness but are more specular in nature causing less contrast when compared to the surfaces resulting from worn tool, and is also evident from figure 4. In addition, more continuous feed marks results in the surface machined by unworn tool than that by worn tool. Therefore, to distinguish between the unworn and worn tool from the machined surface images, two major attributes i.e., contrast and discontinuity in the feed marks may play the key roles. In this perspective, more abstract level information is required to extract those key attributes, accurately. These key attributes were extracted by performing image pre-processing, edge gradient detection and image texture analysis for features extraction to disctriminate between unworn and worn tool in most of the literature [12, 16]. However, in those algorithm, manual feature selections and machining condition dependent model are two major limitations. CNN, on the other hand, helps to overcome these limitations by waiving off the expertise dependent image pre-processing and feature selection. In the present work, extraction of the abstract level information through CNN, which is analogous to the image pre-processing and gradient extraction, is performed in subsequent convolutional and pooling layers. Contour maps of the activations extracted from subsequent convolutional and pooling layers for figure 4a, b are shown in figure 5 to present the enhancement of delineation between the machined surfaces generated by unworn and worn tool in layer by layer increment.

It can be observed from the contour maps (one of the filtered outputs for each of the images in figure 4a, b) depicted in figure 5a, b that almost similar information like input images can be seen, as very lower level features abstraction are happened in first convolution layer. After first convolution operation, pooling is performed to extract little higher abstract information, as seen in figure 5c, d, attributing towards more delineation. Comparing figure 5c with d, it can be easily visualized that more discontinuous higher value pixels are present in figure 5d than that in figure 5c, due to rougher suface produced using worn tool. In the next higher level i.e., in 2^{nd} convolution layer, edge segmentation started which can be depicted from figure 5e, f. After 2^{nd} pooling, the difference between the two classes of surfaces become more prominent, as depicted in figure 5g, h. Also, at the same layers in CNN, extraction of edge gradient based information becomes evident, which is shown in figure 6.

In the convolution and pooling layers, extraction of abstract information similar to the extraction of edge gradients can be performed, simultaneously. One of the major advantages of CNN is simultaneous applications of more convolution filters in a single layer with optimized filter weights or coefficient. This is somewhat similar to the discrete wavelet transform (DWT) operation used in signal and image processing for extracting non-stationary features. However, mother wavelet and decomposition level selection are the major important steps in DWT for useful extraction non-stationary features from machine surface images[17, 18] to get the information of tool condition. Non-stationary features, on the other hand, can be extracted by using CNN also, without any endevour towards the selections of mother wavelet and decomposition level. This can be evident from figure 6, where contour maps of one of

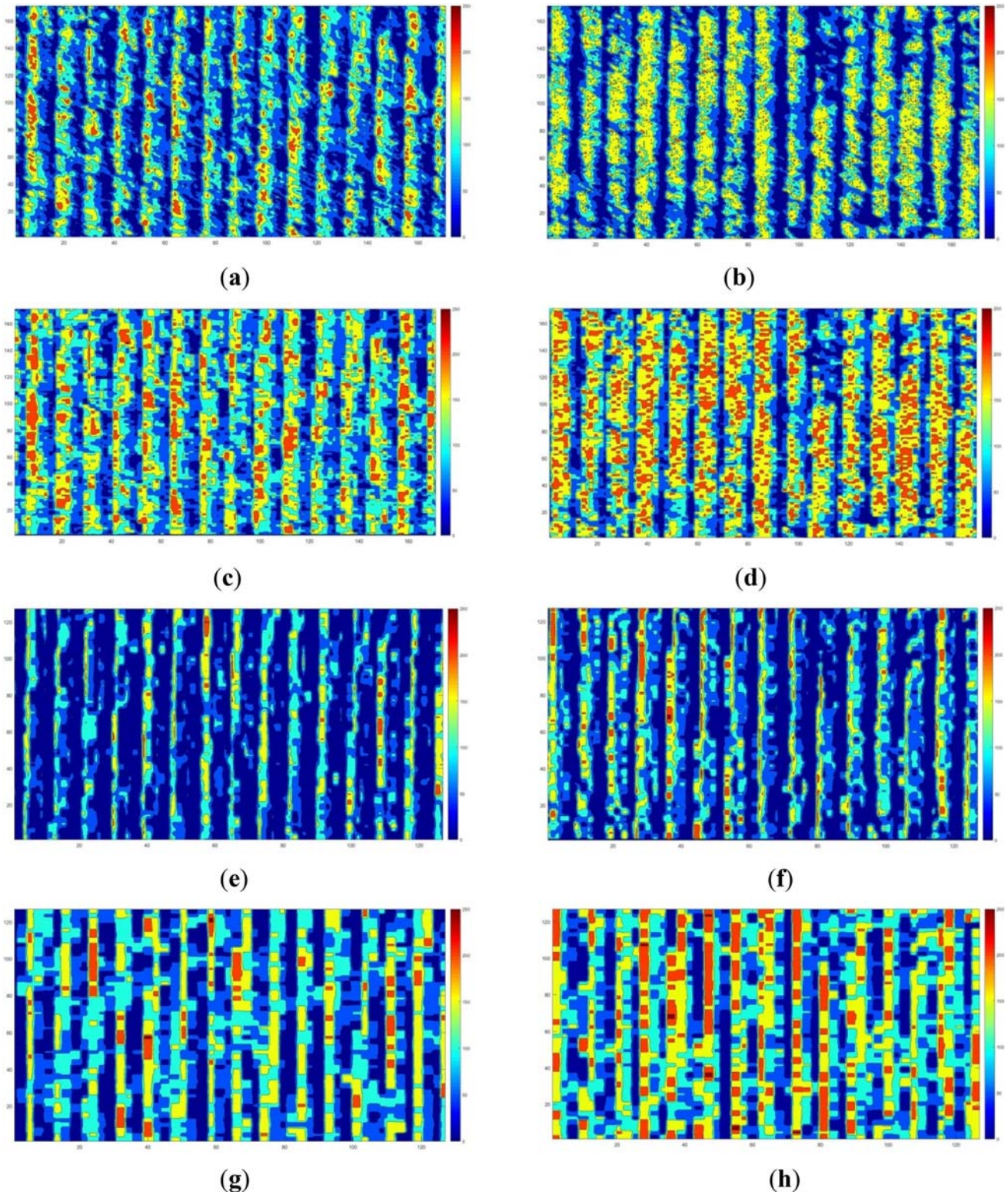

Figure 5. Contour maps of one of the filtered outputs from (**a**, **b**) 1st convolution layer, (**c**, **d**) 1st pooling layer, (**e**, **f**) 2nd convolution layer and (**g**, **h**) 2nd pooling layer of figure 4a, b, respectively, analogous to image pre-processing or filtering perspective.

the filtered outputs of 1st convolutional and 1st pooling layers for figure 4a, b are shown. In figure 6a, b, contour maps of one of the filtered outputs (1st convolutional layer) of figure 4a, b, respectively, are shown where edge gradient information can be depicted. Figure 6c, d are the outputs from the 1st pooling of figure 6a, b, respectively, where the

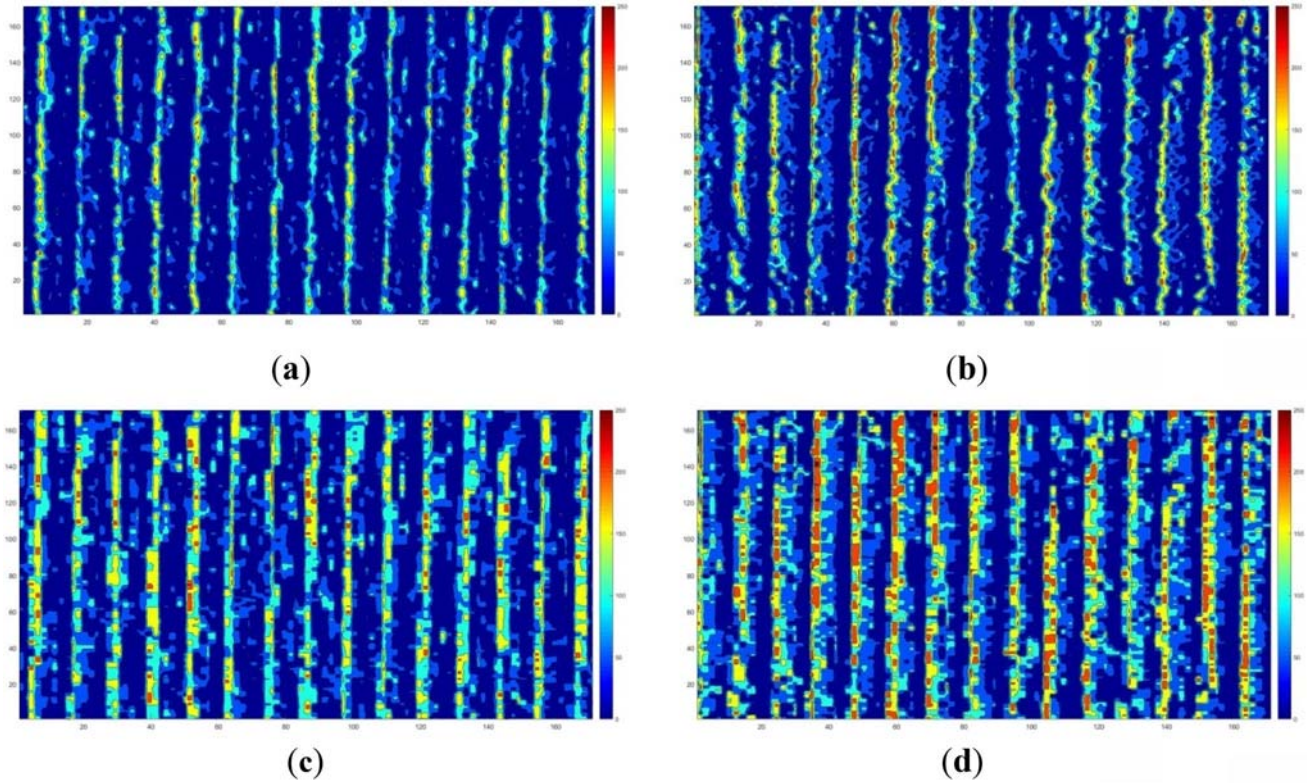

Figure 6. Contour maps of one of the filtered outputs for (**a**, **b**) 1st convolution layer, (**c**, **d**) 1st pooling layer, of figure 4a, b, respectively, analogous to gradient based image segmentation and DWT perspective.

contrast and discontinuity information from feed marks are more prominently classifiable, due to extraction of more localized information from pooling layer. Therefore, more accurate classification, independent of machining condition, between unworn and worn tool can be achieved by using CNN model, as difference between the higher abstract information of surfaces generated by unworn and worn tools becomes more prominent with increment in layers as mentioned in section 3.1. Finally, all the different abstract level information are used and generalized performance of the model enhances by using dropout layer.

Results obtained for classification by using CNN are compared with the classification by using support vector machine used in [14]. Here, 15 features from GLCM matrices of all the machined surface images were extracted. Then feature selection was performed by using Fisher's Discriminant Ratio (FDR) as performed in [14] for two equiprobable classes. According to the features' ranking features, Mean and Variance were selected. Then, values of these two features were extracted from the GLCM of all the pre-processed images kept for training to build a SVM classification model with linear kernel. Sequential minimal optimization (SMO) technique was used for hyperplane separation.

The obtained SVM classification model has 85% test accuracy, where plot against some of the test images are shown in figure 7. The points marked red in color are representing the mis classified images in figure 7.

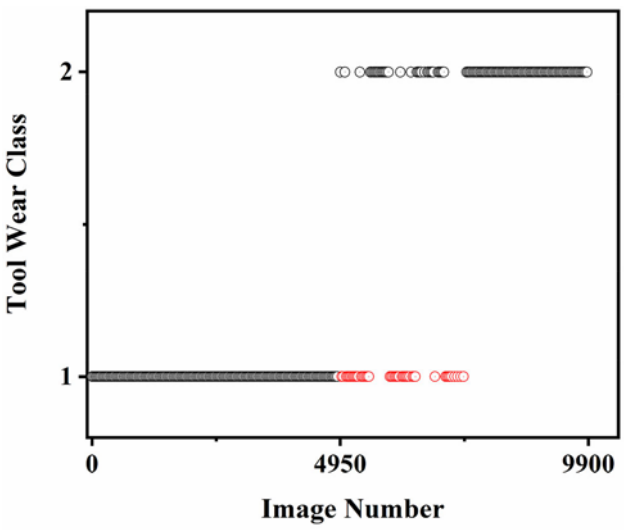

Figure 7. Plot of tool wear state versus test image number.

Although working of a CNN seems to be a blackbox, one can have a better understanding about the network by

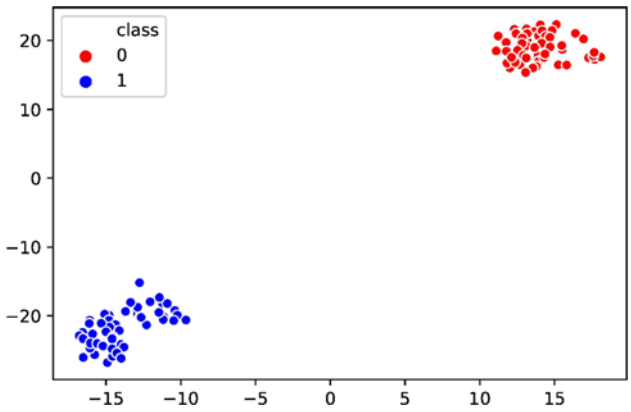

Figure 8. Visualization of the activations from fully connected layer by t-SNE.

visualizing the output activations from the layers. In order to analyse the ouptut activations, the t-distributed stochastic neighbour embedding (T-SNE) [36] algorithm was used. The classification efficiency of the proposed CNN model can be illustrated by visualizing the activations of the final fully connected layer with the help of T-SNE plot. A random smaple of 50 images from each of the classes are selected and procesed through the trained CNN network. For each image, the activations of the final fully connected layer are separated and visualized using T-SNE plot as shown in figure. Each point in figure represents an image sample and it is observed that samples of the same type are clustered together. Since the classification deals with machined surface images with worn (class 0) and unworn (class 1) tools and the same can be understood with the help of clusters formed. Therefore, we can understood, from figure 8, that the proposed CNN can classify the tool state, successfully, using the machined surface images.

5. Conclusions

In the present work, a CNN model is developed for classification of cutting tool flank wear in unworn and worn classes, by using machined surface images. Turning experiments are carried out by varying the cutting speed, feed rate and depth of cut using CNC controlled lathe machine. Turned surface images are acquired after each machining experiment and are used to train and test the developed CNN model. In order to realize a more generalized CNN model, appropriate hyperparameters and training parameters viz. number of layers and filters, learning rate, batch size, and epochs are chosen from the computational experiments. The developed CNN model is tested using the machined surface images acquired for the machining conditions which are completely different from the machining conditions at which training data set is obtained to find the generalization performance. It is observed that the developed model performance is independent of machining conditions and is capable of classifying the cutting tool state as worn or unworn. The proposed CNN architecture shows testing accuracies of 96.3% and 99.9% when tested with raw images and minimally pre-processed images of machined surfaces, respectively. Performance of the CNN model was compared with SVM model, where it can be concluded that the CNN model's performance is much better than SVM for generalized classification or parameter-independent classification of tool flank wear from turned surface images. The developed CNN model can be used to classify the cutting tool states from machined surface images in industrial environment.

Acknowledgement

Authors are thankful to the Director, CSIR-CMERI, Durgapur and CAMM, CSIR-CMERI, Durgapur for motivation and support.

References

[1] Gao R, Wang L, Teti R, Dornfeld D, Kumara S, Mori M and Helu M 2015 Cloud-enabled prognosis for manufacturing. *CIRP Ann.* 64: 749–772
[2] Wang L, Adamson G, Holm M and Moore P 2012 A review of function blocks for process planning and control of manufacturing equipment. *J. Manuf. Syst.* 31: 269–279
[3] Peng T and Xu X 2014 Energy-efficient machining systems: a critical review. *Int. J. Adv. Manuf. Technol.* 72: 1389–1406
[4] Wu X, Liu Y, Zhou X and Mou A 2019 Automatic identification of tool wear based on convolutional neural network in face milling process. *Sensors* 19: 3817
[5] Lins R G, de Araujo P R M and Corazzim M 2020 In-process machine vision monitoring of tool wear for Cyber-Physical Production Systems. *Robot. Comput. Integr. Manuf.* 61: 101859
[6] Terrazas G, Martínez-Arellano G, Benardos P and Ratchev S 2018 Online tool wear classification during dry machining using real time cutting force measurements and a CNN approach. *J. Manuf. Mater. Process.* 2: 72
[7] Liu M-K, Tseng Y-H and Tran M-Q 2019 Tool wear monitoring and prediction based on sound signal. *Int. J. Adv. Manuf. Technol.* 103: 3361–3373
[8] Hui Y, Mei X, Jiang G, Tao T, Pei C and Ma Z, 2019, Milling Tool Wear State Recognition by Vibration Signal Using a Stacked Generalization Ensemble Model. *Shock Vib.* 2019
[9] Abou-El-Hossein K and Kops N 2017 Investigation on the use of cutting temperature and tool wear in the turning of mild steel bars. *J. Mech. Eng. Sci.* 11: 3038–3045
[10] Dutta S, Pal S K, Sen R, Dutta S, Pal S K and Sen R 2014 Digital image processing in machining. Springer, Berlin
[11] Mannan M A, Kassim A A and Jing M 2000 Application of image and sound analysis techniques to monitor the condition of cutting tools. *Pattern Recognit. Lett.* 21: 969–979

[12] Kassim A A, Mian Z and Mannan M A 2004 Connectivity oriented fast Hough transform for tool wear monitoring. *Pattern Recognit.* 37: 1925–1933

[13] Kassim A A, Mian Z and Mannan M A 2006 Tool condition classification using Hidden Markov Model based on fractal analysis of machined surface textures. *Mach. Vis. Appl.* 17: 327–336

[14] Bhat N N, Dutta S, Vashisth T, Pal S, Pal S K and Sen R 2016 Tool condition monitoring by SVM classification of machined surface images in turning. *Int. J. Adv. Manuf. Technol.* 83: 1487–1502

[15] Bhat N N, Dutta S, Pal S K and Pal S 2016 Tool condition classification in turning process using hidden Markov model based on texture analysis of machined surface images. *Measurement.* 90: 500–509

[16] Kassim A A, Mannan M A and Mian Z 2007 Texture analysis methods for tool condition monitoring. *Image Vis. Comput.* 25: 1080–1090

[17] 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. *Precis. Eng.* 43: 34–42

[18] Dutta S, Pal S K and Sen R, 2016, Tool condition monitoring in turning by applying machine vision. *J. Manuf. Sci. Eng.* 138

[19] Dutta S, Pal S K and Sen R 2018 Progressive tool condition monitoring of end milling from machined surface images. *Proc. Inst Mech. Eng. Part B J. Eng. Manuf.* 232: 251–266

[20] LeCun Y, Bottou L, Bengio Y and Haffner P 1998 Gradient-based learning applied to document recognition. *Proc. IEEE.* 86: 2278–2324

[21] Farabet C, Martini B, Akselrod P, Talay S, LeCun Y and Culurciello E, 2010 Hardware accelerated convolutional neural networks for synthetic vision systems, In: *Proc. 2010 IEEE Int. Symp. Circuits Syst., IEEE*, pp. 257–260

[22] LeCun Y, Bengio Y and Hinton G 2015 Deep learning. *Nature.* 521: 436–444

[23] Guo Y, Liu Y, Oerlemans A, Lao S, Wu S and Lew M S 2016 Deep learning for visual understanding: A review. *Neurocomputing.* 187: 27–48

[24] Wang J, Ma Y, Zhang L, Gao R X and Wu D 2018 Deep learning for smart manufacturing: Methods and applications. *J. Manuf. Syst.* 48: 144–156

[25] Iso S 1993 Tool Life Testing with Single Point Turning Tools. *ISO.* 3685: 1993

[26] Chang S I and Ravathur J S 2005 Computer vision based non-contact surface roughness assessment using wavelet transform and response surface methodology. *Qual. Eng.* 17: 435–451

[27] Grzesik W and Brol S 2009 Wavelet and fractal approach to surface roughness characterization after finish turning of different workpiece materials. *J. Mater. Process. Technol.* 209: 2522–2531

[28] Josso B, Burton D R and Lalor M J 2001 Wavelet strategy for surface roughness analysis and characterisation. *Comput. Methods Appl. Mech. Eng.* 191: 829–842

[29] Josso B, Burton D R and Lalor M J 2002 Frequency normalised wavelet transform for surface roughness analysis and characterisation. *Wear.* 252: 491–500

[30] Morala-Argüello P, Barreiro J and Alegre E 2012 A evaluation of surface roughness classes by computer vision using wavelet transform in the frequency domain. *Int. J. Adv. Manuf. Technol.* 59: 213–220

[31] Zhao R, Yan R, Chen Z, Mao K, Wang P and Gao R X 2019 Deep learning and its applications to machine health monitoring. *Mech. Syst. Signal Process.* 115: 213–237

[32] Goodfellow I, Bengio Y, Courville A, Goodfellow I, Bengio Y and Courville A, 2016, Deep learning, MIT press

[33] Shang W, Sohn K, Almeida D and Lee H, Understanding and improving convolutional neural networks via concatenated rectified linear units, In: *Int. Conf. Mach. Learn.*, pp. 2217–2225

[34] Hinton G E and Salakhutdinov R R 2006 Reducing the dimensionality of data with neural networks. *Science (80-)* 313: 504–507

[35] Srivastava N, Hinton G, Krizhevsky A, Sutskever I and Salakhutdinov R 2014 Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15: 1929–1958

[36] van der Maaten L and Hinton G 2008 Visualizing data using t-SNE. *J. Mach. Learn. Res.* 9: 2579–2605