The International Journal of Advanced Manufacturing Technology (2020) 109:953–974
https://doi.org/10.1007/s00170-020-05449-w

ORIGINAL ARTICLE

Review of tool condition monitoring in machining and opportunities for deep learning

G. Serin[1] · B. Sener[1] · A. M. Ozbayoglu[2] · H. O. Unver[1]

Received: 28 September 2019 / Accepted: 12 May 2020 / Published online: 10 July 2020

Abstract

Tool condition monitoring and machine tool diagnostics are performed using advanced sensors and computational intelligence to predict and avoid adverse conditions for cutting tools and machinery. Undesirable conditions during machining cause chatter, tool wear, and tool breakage, directly affecting the tool life and consequently the surface quality, dimensional accuracy of the machined parts, and tool costs. Tool condition monitoring is, therefore, extremely important for manufacturing efficiency and economics. Acoustic emission, vibration, power, and temperature sensors monitor the stability and efficiency of the machining process, collecting large amounts of data to detect tool wear, breakage, and chatter. Studies on monitoring the vibrations and acoustic emissions from machine tools have provided information and data regarding the detection of undesirable conditions. Herein, studies on tool condition monitoring are reviewed and classified. As Industry 4.0 penetrates all manufacturing sectors, the amount of manufacturing data generated has reached the level of big data, and classical artificial intelligence analyses are no longer adequate. Nevertheless, recent advances in deep learning methods have achieved revolutionary success in numerous industries. Deep multi-layer perceptron (DMLP), long-short-term memory (LSTM), convolutional neural network (CNN), and deep reinforcement learning (DRL) are among the most preferred methods of deep learning in recent years. As data size increases, these methods have shown promising performance improvement in prediction and learning, compared to classical artificial intelligence methods. This paper summarizes tool condition monitoring first, then presents the underlying theory of some of the most recent deep learning methods, and finally, attempts to identify new opportunities in tool condition monitoring, toward the realization of Industry 4.0.

Keywords Tool condition monitoring · Machining · Industry 4.0 · Deep multi-layer perceptron · Long-short-term memory · Convolutional neural network · Reinforcement learning

1 Introduction

The main objective of the machining process is to remove chips with the highest level of performance and maximize tool life. It is therefore extremely important to adjust the optimum process parameters during machining. It is also vital to monitor the tool conditions and detect any anomalies that can occur during machining to avoid hazardous conditions. During any machining operation, the cutting tool life directly affects the quality and cost of the process.

✉ H. O. Unver
hounver@etu.edu.tr

1 Department of Mechanical Engineering, TOBB University of Economics and Technology, Ankara, Turkey

2 Department of Computer Engineering, TOBB University of Economics and Technology, Ankara, Turkey

Through tool condition monitoring (TCM), problems such as tool wear, tool breakage, and chatter that can occur during machining can be eliminated. Hence, adverse effects from worn tools on the part quality, production time, and costs can be minimized. In addition, the high cost of a machine breakdown or workpiece breakage can be eliminated through TCM [1, 2].

TCM systems can be divided into hardware and software components. Hardware parts consist of various sensors that collect data such as temperature, vibration, acoustic emissions, and cutting force. In the software part of a TCM system, several processes such as signal preprocessing, feature extraction, feature selection, and decision-making are applied, as shown in Fig. 1. In a consistent and reliable TCM system, the software and hardware parts must be compatible and interact seamlessly. Hence, the tool life can be maximized, and machine tool failures can be avoided. Therefore, a reliable TCM system plays a crucial role in Industry 4.0 [2–4].

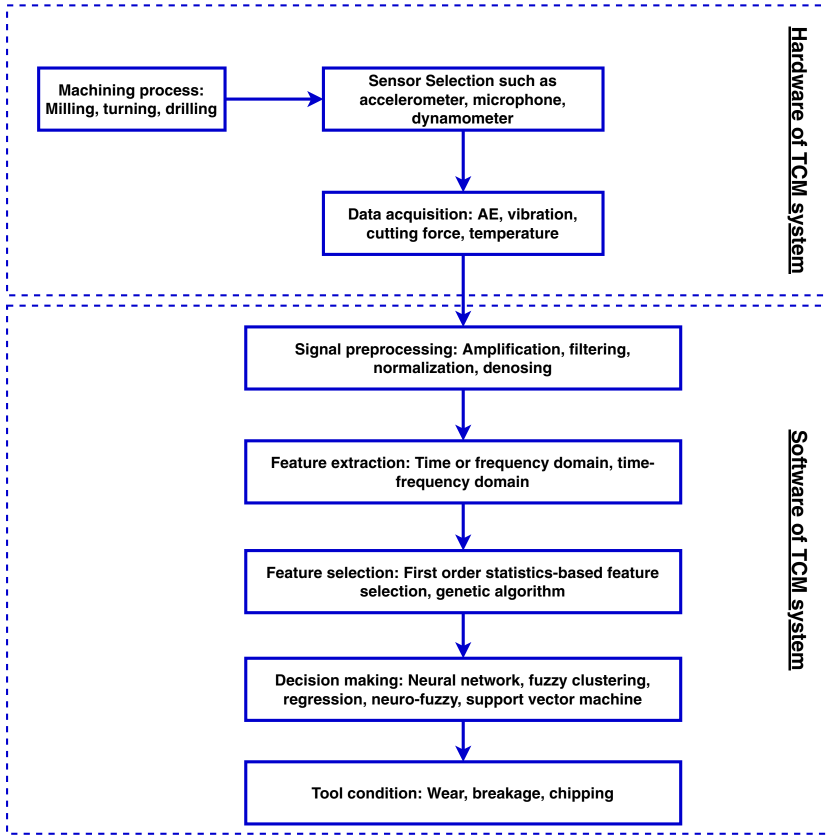

Fig. 1 Process flowchart of the TCM system

There are two methods, direct (offline) and indirect (online), used to monitor tool wear and breakage. For example, fiber optic sensors are directly applied to monitor the anomalies occurring in a cutting tool when applying optical methods such as a microscope or surface profiler. Direct measurement methods reveal the dimensional changes in the cutting tool and machined part with more precise results. However, the drawback of these methods is the deterioration in the accuracy and sensitivity of the optical gadget in the working environment of the machine tool. In addition, to conduct measures using a direct method, the operation must be continuously interrupted. This inversely affects the production time and quality. As opposed to direct methods, with indirect methods, the tool conditions such as the tool wear and tool breakage are observed using signals sourced from vibrations, temperature, cutting force, and acoustic emission, among other factors. Although the measurement accuracy is lower than direct methods, they have the advantages of easy installation and real-time monitoring without interrupting the process, as shown in Table 1 [3–5].

During indirect measurement, first, data is collected using various sensors such as accelerometers, microphones, torque meters, and dynamometers. Subsequently, amplification, data

Table 1 Types of tool condition monitoring

The types of TCM	Process parameter (input signal)	Transducer	Measurement (output signal)
Direct (offline)	Optical/vision	Optical instruments such as CCD camera or optic sensor	The concentration and size of the tool wear
	Electric resistance	Voltmeter	Junction resistance between tool and workpiece
	Displacement	Micrometer, pneumatic gauge, displacement transducer	The distance between tool and workpiece
	Acoustic emission	Acoustic emission transducer, microphone	Acoustic wave
Indirect (online)	Vibration	Accelerometer	Vibrations occurred in tools and machine tool
	Cutting force	Dynamometer or strain gauges	Cutting force of the main spindle during the machining
	Cutting temperature	Thermocouples	Temperature of the workpiece and the main spindle
	Electrical current	Amperemeter or dynamometer	Current or power consumption during the machining
	Surface roughness	CCD camera or fiber optic sensor	Surface roughness of the machined surface

filtering, labeling, and noise reduction methods are often used for signal preprocessing. Applying signal preprocessing methods correctly on the raw signal data will affect the success of tool condition monitoring and prediction algorithms. Required information should not be lost, but uncertainties and unnecessary information should be cleansed during preprocessing. If unnecessary noise exists in the data, the prediction model would not yield accurate results. In addition, data labeling requires extensive work to be applied consistently; otherwise, the learning process could be impaired significantly. Also, fast Fourier transform (FFT) and wavelet transform (WT) are usually preferred in order to convert available data to another form. There are several differences between FFT and WT methods. The most important difference is that the FFT method is localized in the frequency domain whereas the WT method is localized in both the frequency and time domains. Afterwards, feature extraction, feature selection, and the decision-making stages are applied to the data obtained by sensors to determine the tool condition [6, 7].

As mentioned above, optical instruments, accelerometers, microphones, dynamometers, and energy meters are often used to monitor tool condition in machining. However, the locations where the sensors are mounted on the machine tool can vary as shown in Fig. 2. Microphones receive the transient elastic waves (acoustic emission) generated by the rapid release of energy from sources in the material or local components. Therefore, the microphone is placed inside the machine tool cabinet. Since the microphone is not mounted to the spindle or fixture, it is not affected by the process residues such as cutting fluid or generated chip. However, it is prone to pick up any ambient sound and therefore demands more advanced prefiltering techniques. Energy meters also measure the current and voltage drawn by the spindle motor during machining. Hence, they are not affected from cutting residues as well, and they are very easy to install. On the other hand, dynamometers, CCD cameras, and accelerometers have to be mounted directly onto or very close to the spindle or fixture/workpiece assembly. Hence, they are more invasive to the process and can be affected by the dynamics and residues of the process. Although the measured data by them are usually much more reliable and need less preprocessing, their use in day-to-day operations for continuous monitoring can be limited.

To monitor the tool wear and breakage, direct digital image processing–based methods have been mainly used in previous studies owing to their low cost. As an example, Dutta et al. proposed a method for predicting progressive tool flank wear by applying machine vision during the turning operation. With their developed method, a grey level co-occurrence matrix (GLCM), Voronoi tessellation (VT), and discrete wavelet transform (DWT) are used to obtain data related to the feed marks, waviness, and roughness from the machine surface. They have used support vector machine (SVM) as the decision-making technique to accurately describe the tool condition [8].

With the implementation of classical artificial intelligence methods such as SVM, artificial neural network (ANN), and neuro-fuzzy as decision-making methods in tool condition monitoring, better results have been obtained particularly with the indirect methods. As these indirect methods do not interrupt the process, anomalies in cutting tools can be detected more efficiently compared with direct methods [3, 9, 10].

ANN is one of the statistical learning methods designed by taking inspiration from the working of the neurons in the human brain. The deep multi-layer perceptron (DMLP)

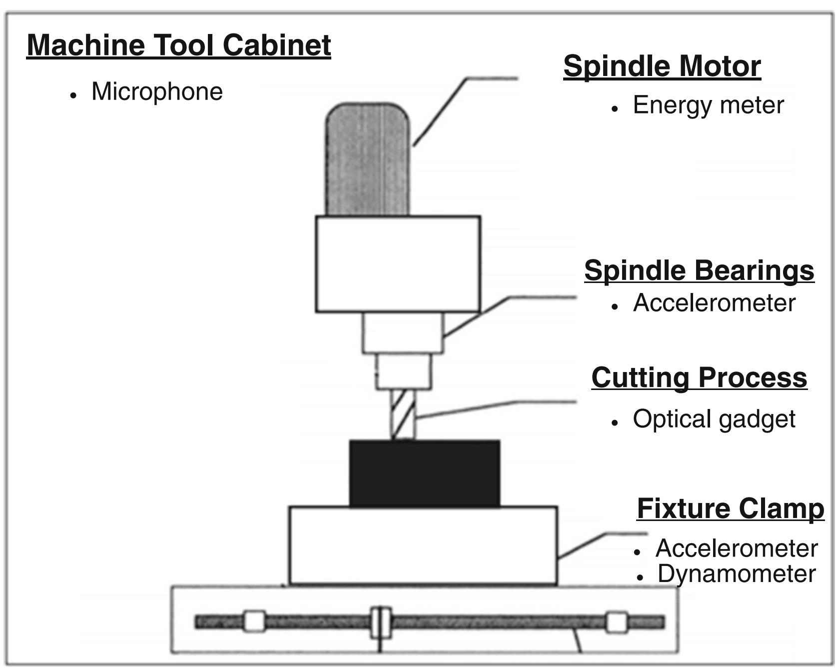

Fig. 2 Placement of various sensors in tool condition monitoring

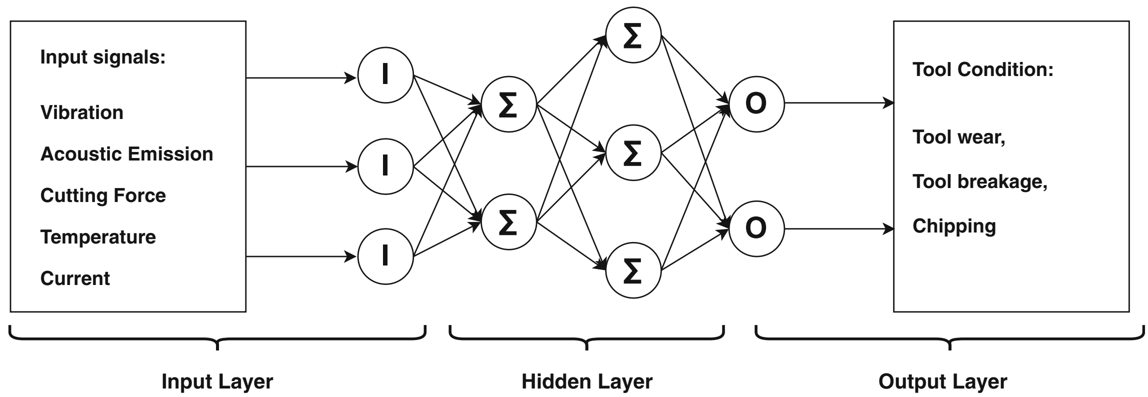

Fig. 3 Model of the artificial neural network with two hidden layers

method, shown in Fig. 3, which is a multiple-layer ANN, is also used to predict the tool condition. It consists of at least two hidden layers. As the use of sensors has become widespread, the size of the collected data has increased. As the number of data increases, the prediction models are being developed by using multi-layer ANNs instead of classical computational intelligence methods [11]. In this regard, Serin et al. have used DMLP neural networks to predict surface roughness and specific energy consumption during 5-axis milling [12]. Additionally, other researchers used SVM as an artificial intelligence technique in the decision-making stages of TCM [13–15].

The remaining sections of this paper are organized as follows: In section 2, the monitoring techniques mostly focused on sensory technologies and related methods for signal processing and feature extraction are presented. In section 3, predominant deep learning methods and their fundamental architectures are presented in detail. Right after that, application of deep learning to fault detection and health monitoring is reviewed. This application area is similar to TCM sensory technologies and signal characteristics. In section 5, application scenarios for deep learning applied to various TCM problems are identified and conceptual architectures are presented. Finally, section 6 discusses the main conclusions of this review paper.

2 Monitoring techniques

With the TCM method, an accelerometer, a microphone, a thermocouple, and a dynamometer can be used as sensors to measure vibrations, acoustic emissions, temperature, and cutting force, respectively, as shown in Fig. 4. The signal outputs are extracted with the help of suitable sensors, generating analog signals. The transformation from the time-series to the frequency domain is achieved using methods such as FFT, WT, or short-time fast transform (STFT) during signal processing according to the desired signal type in the decision-making stage. Then, to efficiently apply artificial intelligence, the feature extraction process should be applied before the decision-making stage [16]. Feature extraction is a dimensionality reduction process where a large time-series dataset is reduced to smaller, more manageable groups for efficient processing. Feature selection is the process where prioritization of relevant features is done automatically or manually, so that the most effective ones are used. However, it is important that during the feature selection process no loss of information useful for further inference occurs.

2.1 Acoustic emission

Acoustic emission (AE) signals are measured using an acoustic emission transducer or microphone. This is a prevalent method among TCM techniques. AE signals consist of acoustic (elastic) waves, which are generated when the workpiece is exposed to plastic deformation after the cutting tool penetrates the workpiece during machining. They occur due to the rapid release of energy within a material as a result of plastic deformation [9, 17–19]. An example of the configuration of AE measurement is shown in Fig. 5.

According to [4] and [20], AE signals are divided into continuous and transient signals, which have different characteristics, as shown in Fig. 6. If the AE signal is continuous, it comes from the cutting tool and indicates tool wear. However, if the AE signal is transient, it indicates that the cutting tool is broken.

Kakade et al. also intensively worked on an AE-based TCM method [21]. They benefited from AE signals used to predict tool wear and chip formation generated during face milling. During an analysis, AE signals and the size of the cutting tool flank wear were measured at a fixed interval, and at the same time, chips were also collected synchronously for monitoring the tool wear and chip formation. Furthermore, Marinescu and Axinte conducted a study on the monitoring and detection of both tool failure and surface defects during the milling process using AE signals [22]. In this study, AE signals obtained during the milling are transformed from the time domain into the time-frequency domain using the STFT method to identify tool and workpiece malfunctions. The time-frequency domain represents a spectrogram of AE

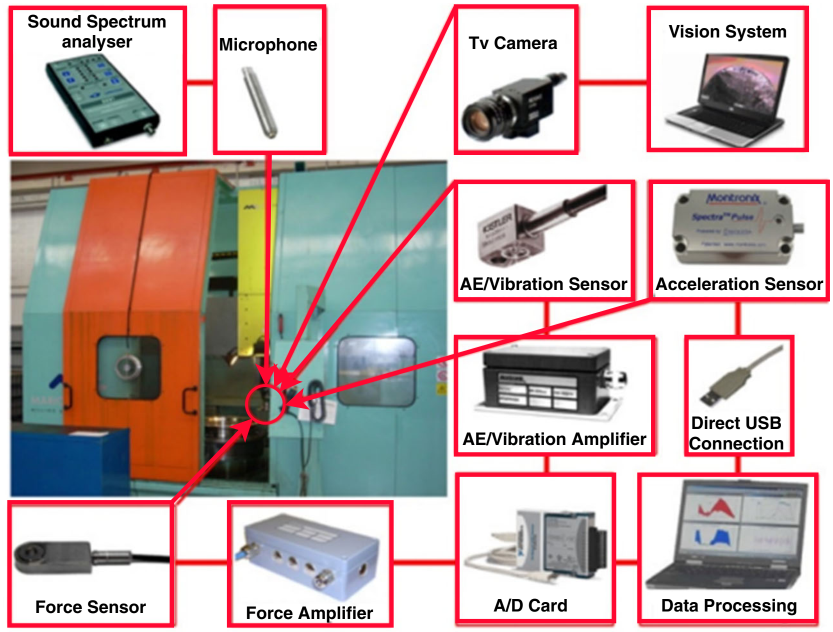

Fig. 4 Tool condition monitoring system [10]

signals. Here, the cutting time (T_c) for each tool was first determined. Then, STFT transformation was applied to the obtained time-series graphs for each tool, resulting in time-frequency matrices called spectrograms. A sample spectrogram obtained after an STFT transformation is shown in Fig. 7. In this spectrogram, the pattern of the STFT area representing each tooth engagement is extracted and combined on a dense spectrogram for further inference [22].

2.2 Vibration

In turning or milling operations, self-induced vibration, also known as chatter, often occurs. One of the most widely preferred ways to monitor the tool and process conditions is the use of vibration signals measured by an accelerometer. The signature of the tool path on the workpiece and unpleasant noises occur owing to excessive vibration. This situation

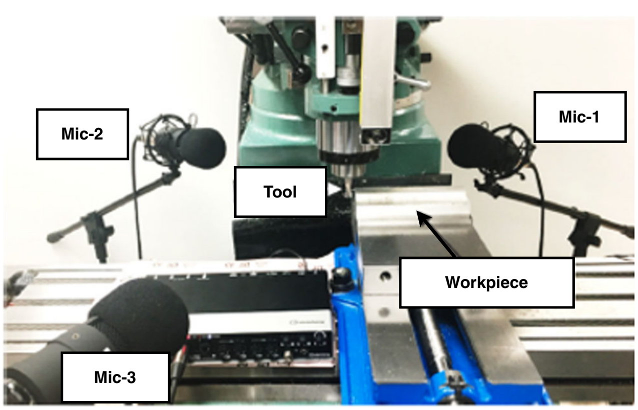

Fig. 5 Experimental AE setup for TCM [18]

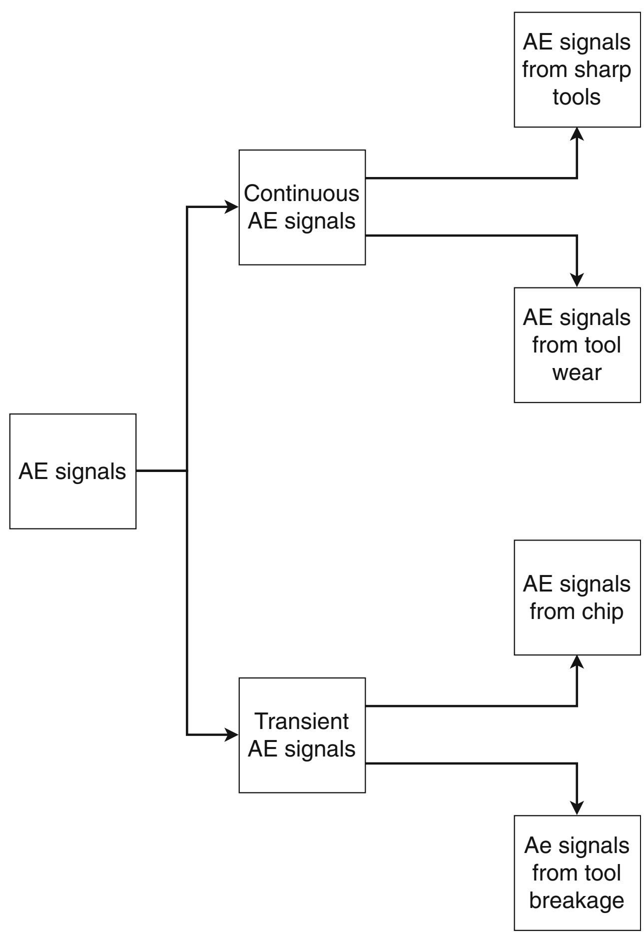

Fig. 6 Types of AE signals extracted during machining [20]

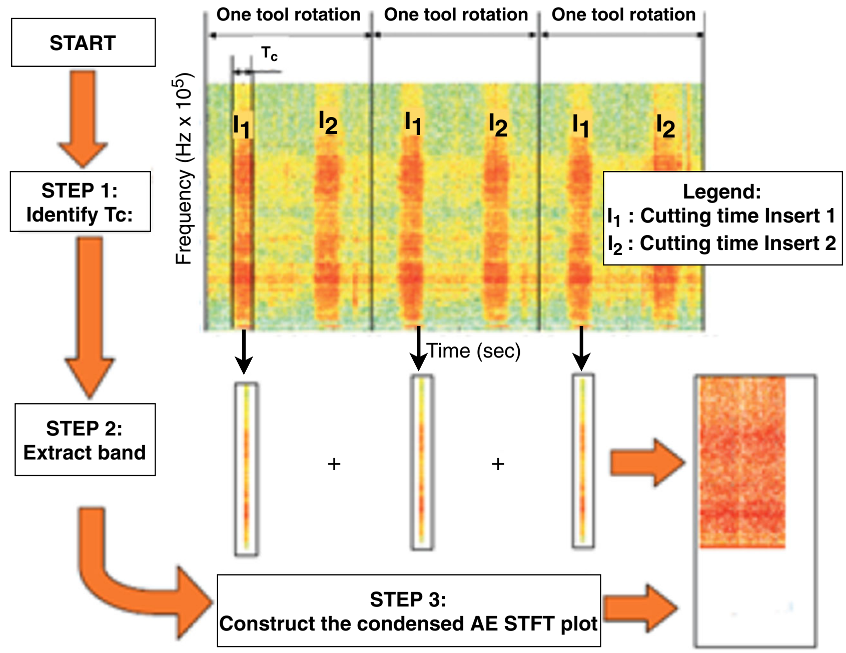

Fig. 7 Time-frequency domain graph (spectrogram) obtained using the STFT method [22]

causes deterioration in the surface quality, reduced tool life, and more tool wear. According to Teti et al., there are two vibration types, namely, dependent and independent cutting operations. A machine tool undergoes independent vibrations because of an unbalanced rotating component, inertia forces of the reciprocating parts, and a kinematic fault of the drives. In other words, factors not related to the cutting parameter can trigger an independent vibration. However, changes in the cutting parameters result in dependent vibrations [4, 10].

Antic et al. presented a novel texture-based tool wear monitoring method that consists of information related to the tool wear state monitored by sensor signals [23]. An STFT analysis was applied to the vibration signals obtained to transform the spectrograms from the time-series into the time-frequency domain. Here, the frequency domain represents the y-dimension, the time-series indicates the x-dimension, and the color distribution refers to the amplitude on a 2D textured image obtained using the STFT method. Spectrographs were then scanned using the appropriate filters in an MR8 filter bank, as shown in Fig. 8, to extract 2D texts for each predefined frequency band.

Table 2 summarizes the studies on vibration data and sensor fusion methods, in which multiple sensors are used to observe the conditions of the cutting tools. When the working

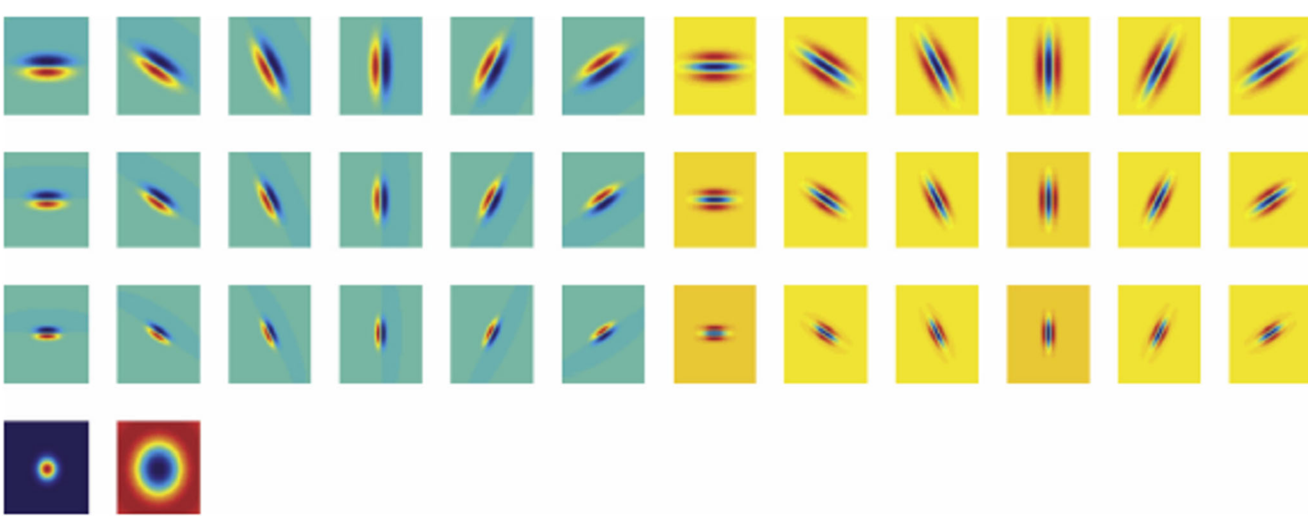

Fig. 8 Filter bank used to detect tool wear [23]

Table 2 Monitoring techniques using vibration signals

References	Types of machining operation	Measurement
[24]	Turning	Tool failure detection using sensor fusion
[25]	Milling process on stainless steel	Tool wear prediction using AE and vibration signals
[26]	Milling	Online tool condition monitoring using sensor fusion
[27]	Milling operation on S45C carbon steel	Tool condition prediction using vibration signals
[28]	Turning process on Inconel 718	Tool wear condition monitoring using sensor fusion system with cutting force and vibration signals
[29]	Milling process	Tool wear monitoring using accelerometer and dynamometer
[30]	Turning process on gray cast iron (FGL 250)	Tool condition monitoring using three-axis piezo-electric accelerometer and dynamometer

environment of a machine tool is considered, a TCM method using more than one sensor type is considered more reliable [31].

2.3 Cutting force

During the machining process, the cutting force will increase based on the selection of the cutting parameters and tool wear. As a cutting tool loses its sharpness, the cutting force will increase. When the increase in cutting force is above a critical level, it indicates a problem in the cutting process. Hence, by measuring the cutting force, the tool wear or breakage can be monitored [32]. To enhance the machining efficiency and product quality for aerospace components developed using an automated milling machining, a system based on the monitoring cutting forces was developed by Marinescu and Axinte [22]. The authors used the cutting force and acoustic emission signals together to detect and predict the occurrence of a cutting tool malfunction and surface deformation, as shown in Fig. 9.

Dimla and Lister carried out a study on an online tool wear monitoring system used during the turning process [33]. The force and vibration signals were measured using a dynamometer and accelerometer, respectively, and the data were then converted from the time-series to the frequency domain using FFT. Under excessive wear, an increase was also simultaneously seen in the cutting force, which can be interpreted as tool wear or breakage. Table 3 shows other studies

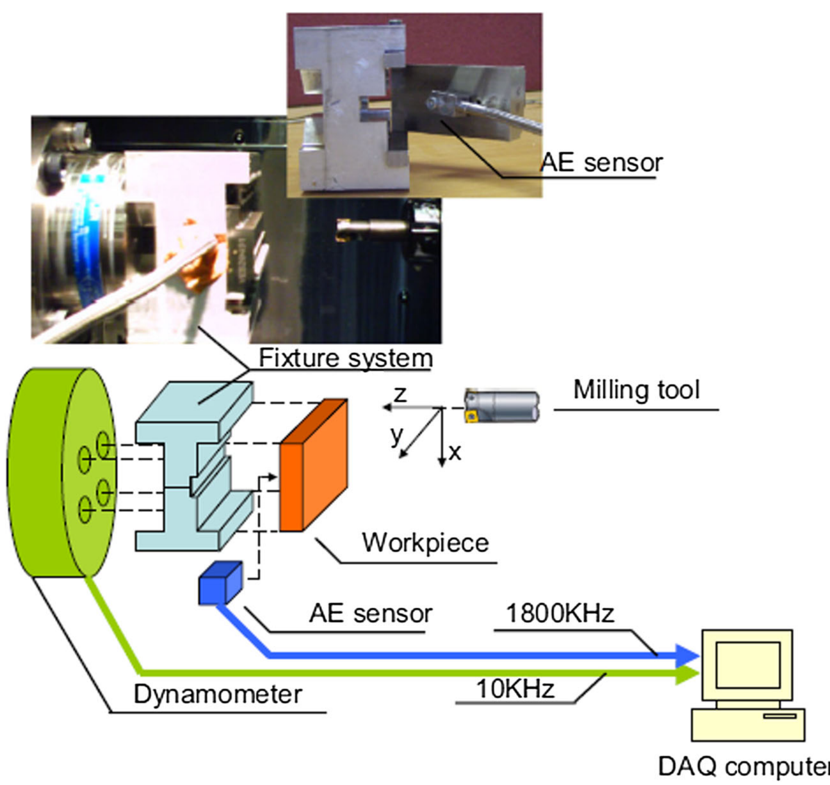

Fig. 9 TCM method using an AE sensor and a dynamometer [22]

Table 3 Monitoring techniques using cutting force

References	Types of machining operation	Measurement
[34]	Turning process on C45 steel bars	Tool condition monitoring using AE sensor and cutting force sensor
[35]	Milling on steel, aluminum, and Inconel workpieces	Tool failure and tool wear detection using Kistler dynamometer
[36]	Milling process on steel AISI-1018	Tool breakage detection using cutting force and power consumption
[37]	Milling operation on stainless steel HRC52	Tool wear sensing using multi-sensory fusion system with cutting force and vibration signals
[38]	Drilling on carbon fiber reinforced plastic	Tool condition monitoring using sensor fusion with thrust force, AE, vibration, and torque signals
[39]	Turning process on hardened JIS S45C carbon steel bar	Tool wear monitoring using a strain gauge

investigating the cutting force to establish a relationship between tool wear and breakage.

2.4 Other sensor types

There are other signal types such as the temperature, surface roughness, and energy that can be measured apart from the vibration, acoustic emission, and cutting force to monitor the tool condition. During the metal cutting processes, heat is generated because of energy released by the shearing of the material. When the resulting heat is too high, it will damage the cutting tool life and the surface quality of the workpiece. A high temperature is inevitable; hence, some researchers have also utilized the temperature of the cutting tool during the TCM process. The tool condition can also be estimated by monitoring the surface roughness. Often, an increased roughness is an indicator of a worn tool and excessive vibrations occurring during the process. Table 4 shows other studies related to temperature and surface roughness signals.

3 Deep learning methods and applications

The use of sensors has recently increased in many areas including self-driving cars, sentiment analysis, predictive maintenance, and fault diagnosis. Thus, objects can sense more signals such as vibrations, images, and acoustic emissions through the Internet of Things (IoT). Machine learning methods, which are a sub-branch of artificial intelligence (AI), have been frequently used in recent years since the concepts of IoT and cloud computing first emerged within Industry 4.0 [43]. The amount of data used to train a model using a machine learning method is extremely important. When the amount of training data is excessive, the prediction capability of traditional machine learning models such as Bayesian networks, linear regression, logistic regression, SVM, RBF, and single-layer ANN do not scale satisfactorily [15]. Nevertheless, as the amount of training data increases, multi-layer neural networks and deep learning methods exhibit superior performance for learning and prediction [44–46]. A comparative view of deep learning performance relative to some other AI methods based on the data size is shown in Fig. 10.

With the increasing use of sensors, the amount of data collected is gradually increasing, and the concept of big data is becoming predominant. Big data is a collection of larger amounts of more complex data using multiple sensors and represents the concept of 3Vs, namely, volume, variety, and velocity, as shown in Fig. 11. Deep learning models and in particular deep multi-layer neural networks utilize big data for their training and testing [47].

In classical computational intelligence methods, the features are extracted by the user to transmit the data to the prediction model. Feature extraction is critical in image

Table 4 Monitoring techniques using temperature, surface roughness, and cutting energy

References	Types of machining operation	Measurement
[40]	Boring process on AISI 1040 steel	Tool condition monitoring using surface roughness
[41]	Turning process AISI 52100 steel	Tool condition monitoring and prediction using surface roughness data together with neural network
[42]	Turning process on AISI 1117 steel	Tool condition monitoring with temperature signals

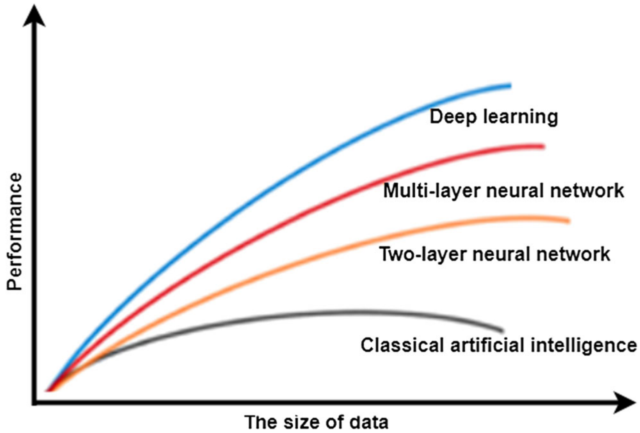

Fig. 10 Performance of selected machine learning techniques vs. size of data

pattern recognition for the performance of the prediction model. However, feature extraction is not needed in deep learning methods before the decision-making stage because deep learning methods implicitly extract the features themselves. During this stage, focus is given to the information required for the estimation model by ignoring unnecessary information. The fundamental structural difference between classical computational intelligence methods and deep learning methods is shown in Fig. 12 [44–46].

Deep learning, which is the newest and one of the most advanced machine learning methods, has been applied extensively to image recognition, sentiment analysis, and natural language processing since its early days of emergence. This method continues to find applications in many industries such as the automotive, medical, aviation, and finance. In addition, in the area of manufacturing, it is predicted that one of the potential areas of application of deep learning will be tool condition and process monitoring.

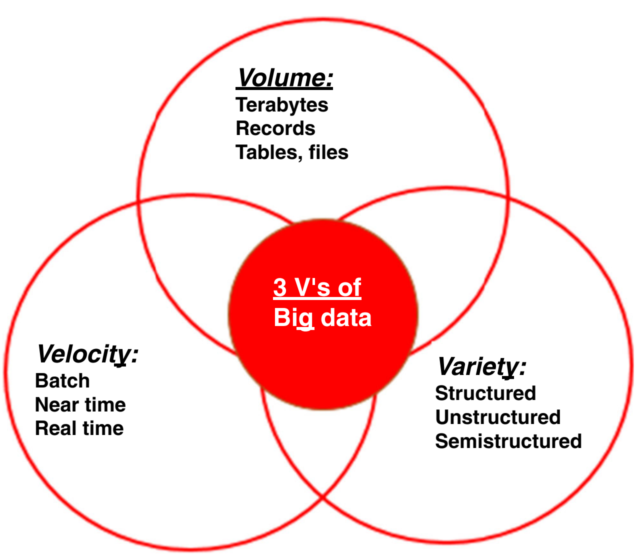

Fig. 11 Concept of 3Vs in big data

The deep multi-layer perceptron (DMLP) method is similar to the classical machine learning method applied in artificial intelligence. The most important difference with this method is that there is more than one hidden layer, and the method operates with big data. The most popular deep learning methods are convolutional neural networks (CNNs), recurrent neural networks (RNNs), and reinforcement learning (RL) [48, 49]. Major deep learning methods and their areas of application are presented in Fig. 13.

3.1 Convolutional neural network

The CNN, which is a major advanced deep learning method, has been proven in the fields of image recognition and classification. In this method, images are usually used as the input. The CNN algorithm extracts and evaluates the feature maps from the images by applying multiple filters hierarchically to the input data and as a result, it classifies objects in groups with similar patterns [50, 51]. To create a CNN model, three basic layer types are used, namely, a convolution (CONV) layer, a pooling (POOL) layer, and a fully connected layer, as shown in Fig. 14.

Equation (1) is used for the convolutional layers, as shown in Fig. 14. Here, f, c, w, s, and b respectively represent the input properties, output properties, filter, step size, and offset value. In addition, Φ indicates the activation function.

$$C_{xyz} = \Phi\left(\left(\sum_{j=0}^{l-1}\sum_{i=0}^{l-1}\sum_{k=0}^{d-1} W_{ijkz}\cdot f(sx+i)(sy+j)k\right) + b_z\right) \quad (1)$$

As mentioned above, the CNN method, which is a special type of multi-layer neural network, is frequently preferred in the areas of face recognition, gesture recognition, image classification, and image recognition to create prediction models using image-based data as input [50]. There have been numerous studies conducted on image processing using the CNN method. Within the scope of the ImageNet project, software programs compete every year in image recognition using large image-based datasets [52]. Krizhevsky et al. presented a complicated CNN-based prediction model that has five convolutional layers, three max-pooling layers, and three fully connected layers, as shown in Fig. 15 [53]. The convolutional layers in the AlexNet CNN–based model have filters with a size of 11 × 11, 5 × 5, and 3 × 3. They managed to reduce the rate of top 5 errors from 26 to 15.3%.

Recently, Zeiler and Fergus managed to reduce this ratio to 14.8 [54]. Their developed model is called ZFNet. In their model, 7 × 7-sized filters were applied in the first convolutional layer instead of an 11 × 11-sized filter. The logic behind this change is that a smaller filter size in the first convolution layer helps maintain much of the original pixel information in the input layer. Additionally, owing to the max-pooling layers, the features acquired by the convolution layers

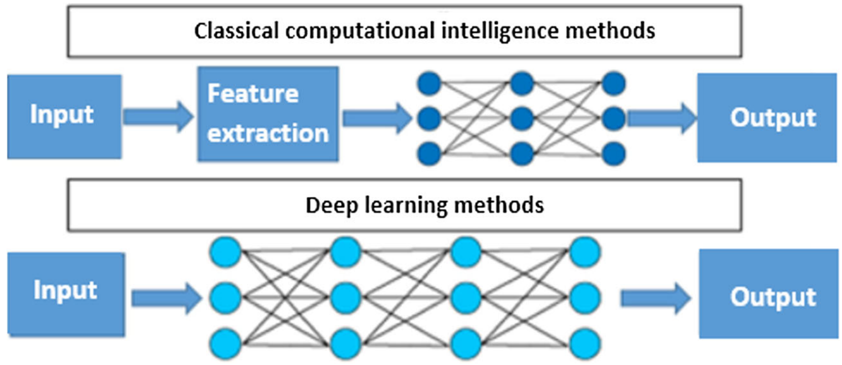

Fig. 12 Flowchart of the classical intelligence method versus deep learning

are reduced and summarized. The goal of this process is to down-sample an input such as an image and reduce its dimensionality. Their ZFNet CNN–based prediction model is shown in Fig. 16. Szegedy et al. developed the GoogLeNet prediction model, which has 22 layers [55]. The authors achieved a 6.7% error rate. ResNet, developed by He et al., has a deeper designed architecture than any other architecture developed thus far, with an error rate of 3.6%.

The CNN method, which is frequently used in image recognition applications, is also used in gesture recognition [56–59], face recognition [60–62], scene labeling [63], and action recognition, as well [64, 65].

3.2 Recurrent neural network

The RNN method is usually used for sequential information or time-series data. A sequential order–based estimation model is used in several areas such as sentiment analysis, natural language processing, and finance. Recently, more accurate results were obtained by using the RNN method in translation and natural language processing. The most important reason for this is because the RNN method, which differs from the CNN method, does not translate each word separately, but interprets the sentence as a sequential order problem and evaluates the sentences in front of and behind it to be translated. In other words, with the RNN method, the output of the hidden layer is reused as the input, allowing better results to be obtained. As with other deep learning methods, the weights of the model are updated using the backpropagation through time algorithm. This algorithm utilizes errors obtained during the training of the RNN model to update the weights of the prediction model. As a disadvantage of the RNN method, it has problems with storing information that requires relatively longer time periods, and in order to achieve that, it stores extremely unnecessary information generated during training. Therefore, the long-short-term memory (LSTM) method, a specialized application of RNN, was developed to solve complex structured networks [15, 66, 67]. The general architecture of the LSTM method is shown in Fig. 17.

Most studies on gesture and image recognition have focused on CNN-based deep learning methods. Recently, however, researchers began preferring the LSTM method to develop prediction models on image recognition–based problems such as gesture and face recognition. Neverova et al. constructed an RNN-based model to predict human gestures [68]. They utilized multiple data types such as depth video, articulated pose, and speech. Tsironi et al. combined the CNN and LSTM methods to more successfully recognize gestures [69]. They combined the two methods because the CNN method has a better predictive ability on image-based data, whereas the LSTM method gives more successful results for larger data. The LSTM and RNN methods have achieved remarkable results in speech recognition, natural language processing, and sentiment analysis. The probability of the distribution over a sequence of words plays a vital role in these cases. Hochreiter and Schmidhuber studied an LSTM model for use in natural language processing based on an interest in the interactions between human language and computers [70].

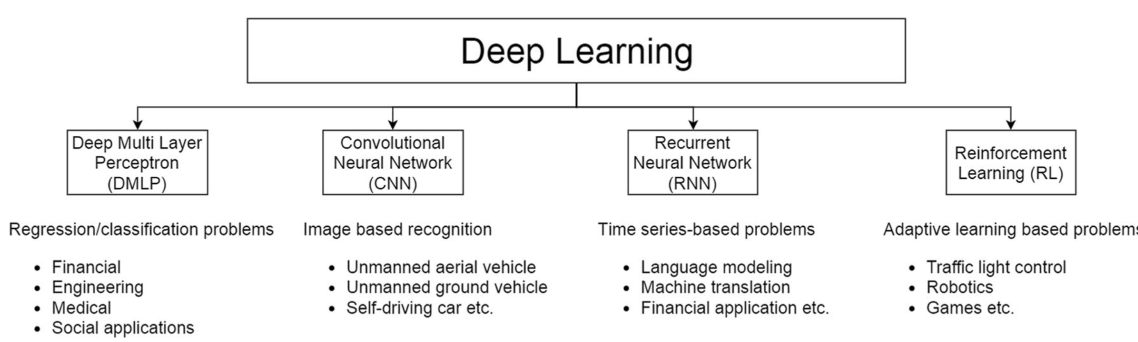

Fig. 13 Deep learning categories and applications

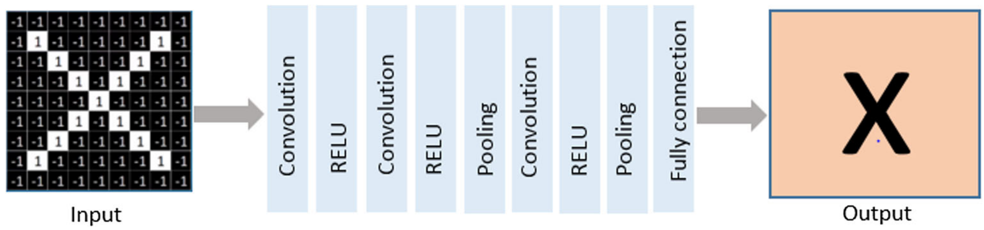

Fig. 14 Structure of the deep learning method

They used an LSTM-deep learning model instead of the RNN method because the LSTM method has the ability to forget redundant information during the analysis of long sentences, and therefore does not occupy too much memory. In addition, many studies on natural language processing have been conducted using the LSTM or RNN method [70–74]. In a review article compiled by Day and Lin, the resulting accuracies of sentiment analyses on Google Play consumers, conducted using three different methods, namely, naïve Bayes, SVM, and LSTM, were revealed [75]. As a result, the accuracy of the sentiment analysis when using the LSTM method was higher, as shown in Fig. 18.

3.3 Reinforcement learning

Reinforcement learning is another machine learning method, inspired by behaviorism, dealing with how software agents should be applied to achieve the highest reward-winning actions for an environment. The data used in reinforcement learning are not labeled, and thus the relationship between the input and output data is not given to the model. Therefore, this is a type of semi-supervised learning because it can learn the relationship between the input and output using a goal-oriented algorithm. The use of this method has recently increased, particularly in autonomous decision-making systems. The reinforcement learning method is modeled as a Markov decision process, as shown in Fig. 19. In other words, the most important feature of this system is to work with a reward and penalty system, and in this way, the behavior will generally achieve the highest reward. The reinforcement learning method consists of five basic elements [76, 77].

•An agent is trained using a goal-oriented algorithm and interacts with the environment.

•A state is the information obtained from the environment and is expressed as s_t.

•An award is the positive or negative feedback after an agent's interaction with the environment and is shown as r_t.

•A behavior is an agent's movement style based on the information obtained from the environment and is expressed as a_t.

•The environment is where the agent conducts an observation.

4 Applications of deep learning in fault detection and machine health monitoring

Data sources from vibration, acoustic, current, and cutting forces are collected using sensors, such as an accelerometer, AE sensor, energy meter, and dynamometer for predictive and preventive maintenance. It is possible to estimate any malfunctions that may occur in a machine tool by using any artificial intelligence method, such as deep learning [78, 79]. In this way, a fault can be detected during the machine health monitoring process, which helps with early identification of problems that may lead to bigger cost problems in the future. Predictive maintenance and preventive maintenance are provided to enhance the lifetime of the machine tools [80, 81].

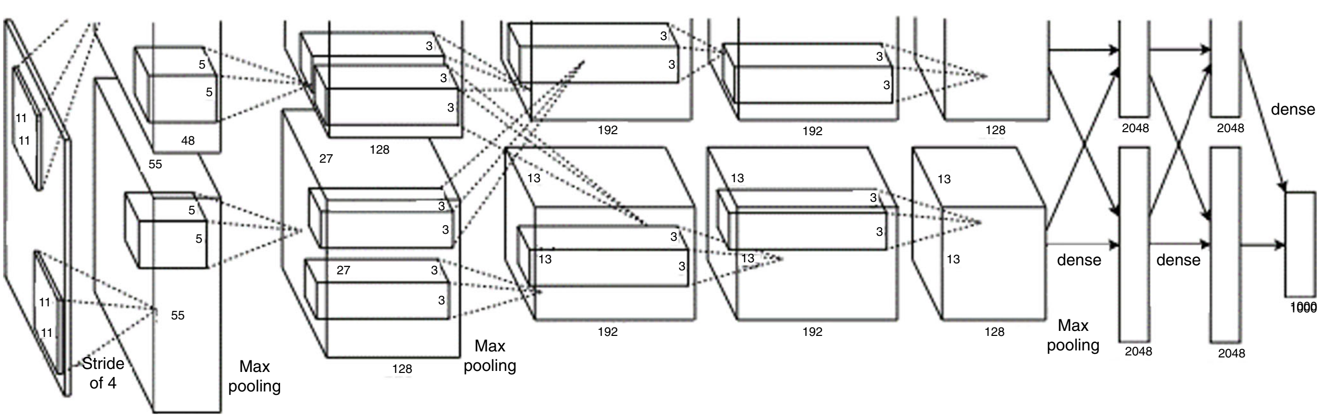

Fig. 15 Architecture of the AlexNet CNN–based model [53]

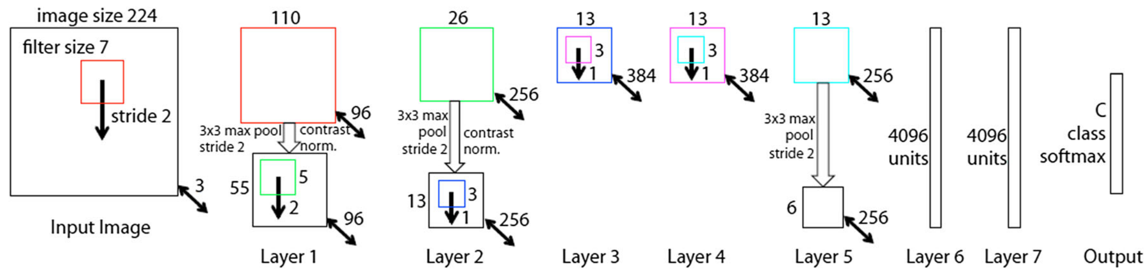

Fig. 16 Architecture of the ZFNet CNN–based model [54]

Classical machine learning methods that can be used for a fault diagnosis include SVM, naïve Bayes, and logistic regression methods. CNN, RNN, deep Boltzmann machine, and deep belief network are the deep learning methods used in this area so far [82]. For machine tools, the structures of the different machine health monitoring systems (MHMSs), including physical-based, conventional data-driven, and deep learning–based MHMSs, are shown in Fig. 20. In a deep learning–based MHMS architecture, CNN and LSTM methods are used together.

Li et al. have developed a deep learning–based fault diagnostic model using data augmentation for rotating machinery to reduce the maintenance cost and increase the operational safety and reliability [83]. They claimed that fault detection can be achieved with a data accuracy of up to 99% with less data through data augmentation. In addition, Fu et al. used a deep belief network to train the data collected as vibration signals during a milling operation. The authors mounted three accelerometers under the spindle housing to collect data from various cutting conditions. Thus, the deep learning method was used to determine the state of the machine tool [84]. Their experimental setup and sensor locations are shown in Fig. 21. Numerous other studies have shown a preference for deep learning methods in fault detection and machine health monitoring [85–91].

5 Opportunities and application scenarios

The use of deep learning methods in unmanned vehicles, natural language processing, financial sector, traffic lights, and energy consumption in buildings has increased in recent years. Within the scope of Industry 4.0, deep learning methods have also been penetrating the field of manufacturing research, and their potential application in this field is extremely high. Prior to the use of deep learning methods, tool wear was estimated mostly by applying classical methods such as ANN for the decision-making process of TCM [91, 92]. Figure 22 shows the classical structure of an ANN model used for TCM.

To increase the life of the cutting tool and the surface quality of work-part after machining, vibration, acoustic, and current data collected during milling, drilling, or turning process have been used along with classical machine learning methods in numerous studies [19, 22, 28, 93–97]. Furthermore, Griffin et al. conducted two novel studies on intelligent classification of anomalies and control simulation for grinding and drilling. They developed a neural network–based classifier to detect anomalies such as burn and chatter for grinding and increasing tool wear and malfunction for drilling. They also enhanced the real-time control system to prevent undesired surface anomalies for these two machining processes [98, 99].

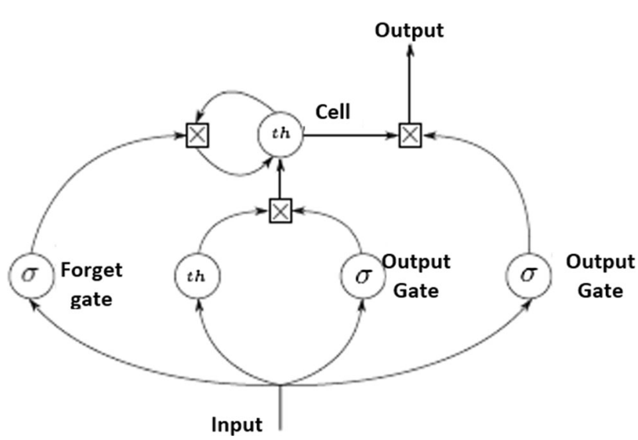

Fig. 17 Structure of long-short-term memory

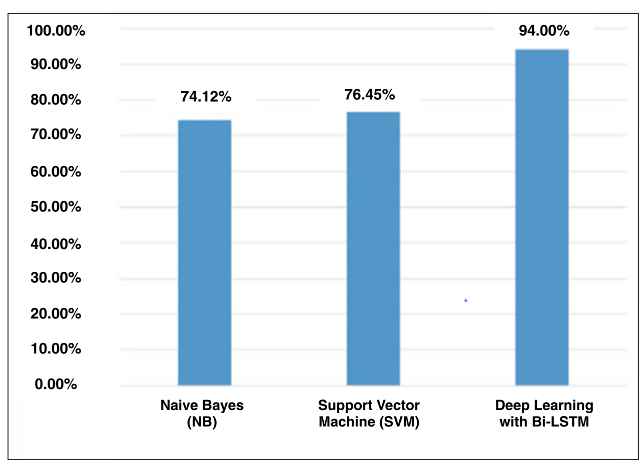

Fig. 18 Resulting accuracy of sentiment analyses using various methods [75]

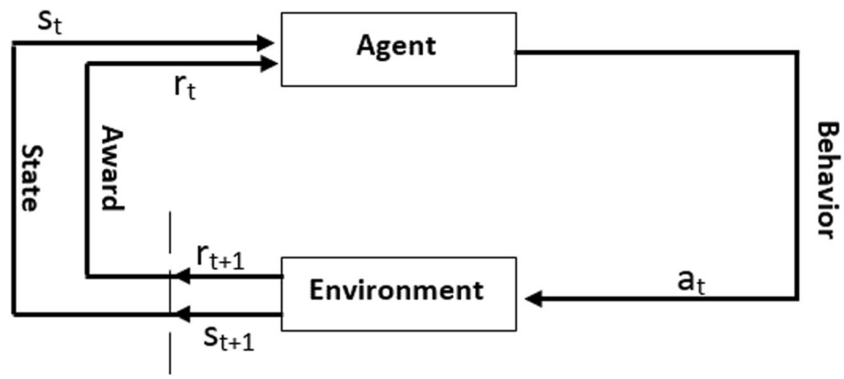

Fig. 19 Structure of the reinforcement learning method

Recently, the amount of data gathered in the shop floor has increased rapidly as the Internet of Things has become more widespread. Classical machine learning methods have not been very successful in the prediction process with big data. Therefore, deep learning methods such as DMLP, CNN, LSTM, and RL have begun to replace the classical machine learning methods. The trend of such applications can be delineated with the use of cloud technology related to the machine learning methods, as shown in Fig. 23. As a corollary, researchers have begun to believe they can obtain more accurate results when using deep learning applications.

5.1 Use of deep learning methods in TCM

The future of deep learning–based applications such as DMLP, CNN, LSTM, and RL in TCM is extremely promising. The collection of vibration, acoustic, cutting force, and current data prior to applying a deep learning method plays a vital role in TCM. After applying various preprocessing operations on the time-series data collected from the sensors such as an accelerometer, microphone, dynamometer, and energy meter, the data is used as input to various deep learning methods as depicted in Fig. 24. Deep learning–based TCM can be used for numerous machining operations, including turning milling, drilling, broaching, and grinding.

Some studies have shown that various sensor types are used for both single- and multi-teeth cutting operations for TCM [9, 100, 101]. However, the signal processing method of the sensory data, which will be used by the deep learning algorithm is of the essence. The data can be in the form of time-series, feature sets, frequency domain, time-frequency (spectrogram) or WT graphs. Each deep learning method will require a particular form of processed data to predict an objective of the TCM, successfully. In the next sub-sections, four exemplarity application scenarios of deep learning methods in order to achieve various goals of TCM are proposed as possible research avenues.

5.1.1 Scenario 1: Use of DMLP for tool wear regression

The DMLP method uses scalar data as input. To use this method in machining applications, it is necessary to extract

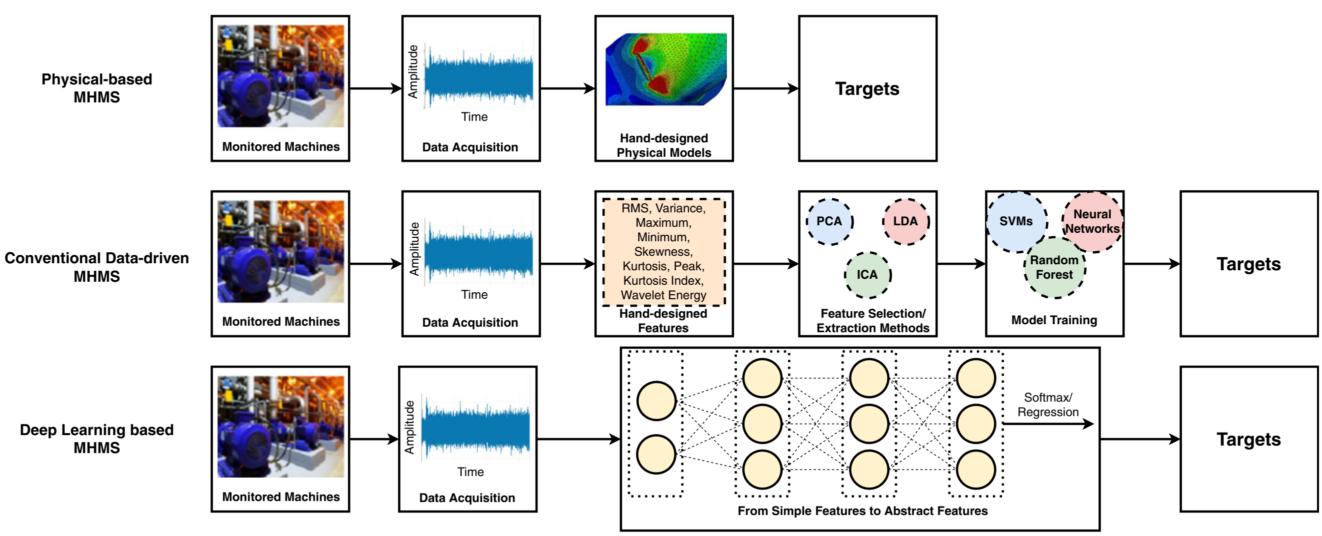

Fig. 20 Structures of three different MHMSs [82]

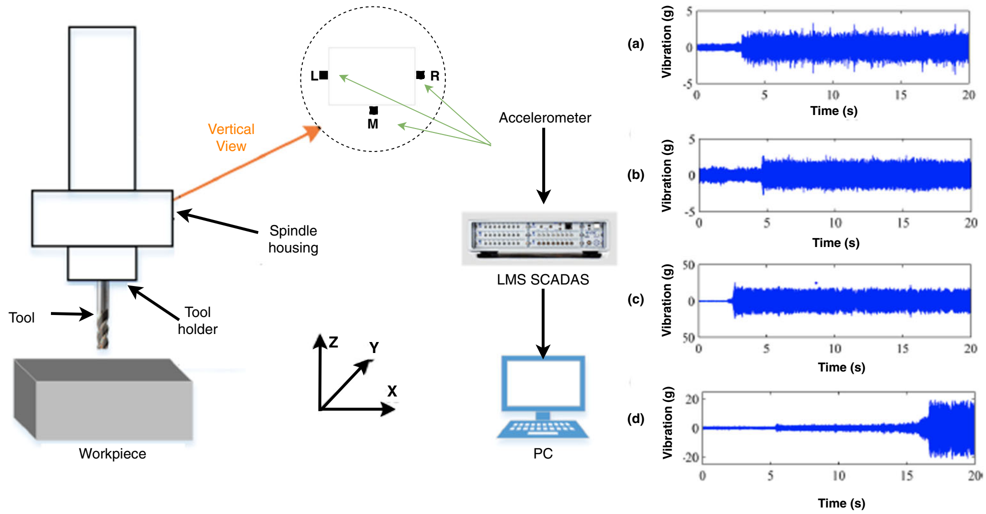

Fig. 21 Experiment setup used to collect vibration data [84]

the features from time-series data captured from sensors. These features are mean, variance, standard deviation, skewness, kurtosis, mean absolute deviation, mean frequency, median frequency, signal-to-noise ratio, and power spectrum deformation. Next, the most appropriate features are selected to reflect the characteristics of the data and redundant ones are eliminated. The selected features are introduced as input to the DMLP method for training. In addition, the microscope measured tool wear values are introduced to the algorithm as labels. As a result, when feature vectors are introduced as input to the trained DMLP-based prediction model, the amount of tool wear can be estimated. The concept architecture for this scenario is shown in Fig. 25.

5.1.2 Scenario 2: Use of CNN for tool wear classification

In the CNN method, spectrograms are used as input for model training. The spectrograms are obtained by applying STFT on the time-series data collected from sensors during the machining process. In the convolution layer,

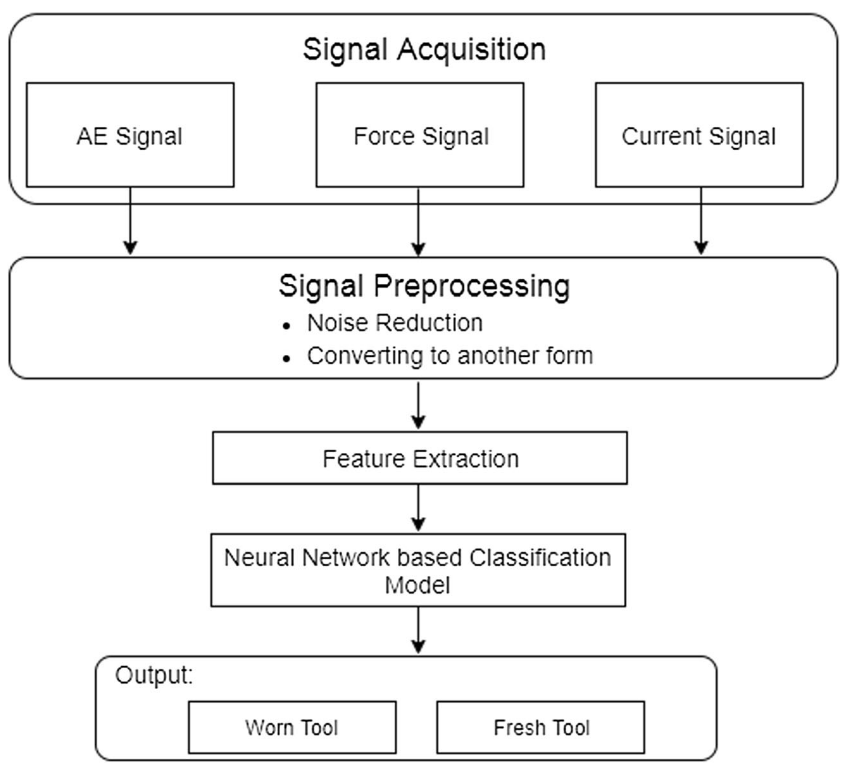

Fig. 22 Structure of a basic ANN model used for TCM

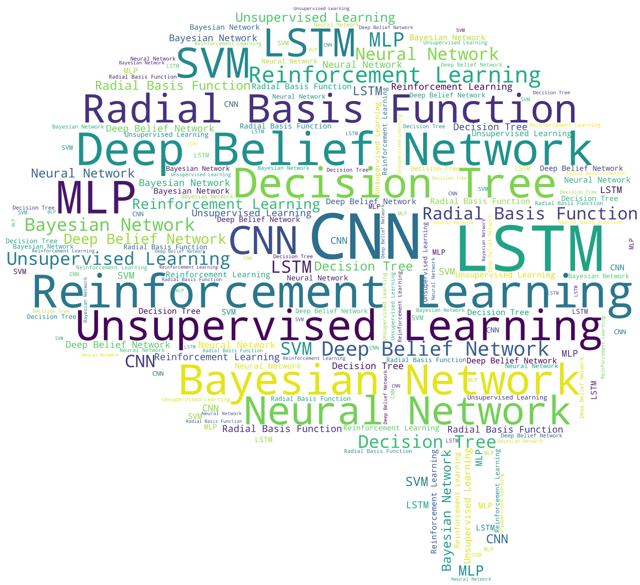

Fig. 23 Trend in applications using tool condition monitoring and machining

the kernel filter is applied on the spectrogram and consequently, features are extracted from the spectrograms. Then, in the pooling layer, the spatial size of the convolved features is reduced. The last layer, the fully connected layer, classifies the cutting tool wear according to the stage of the wear. In summary, to develop a CNN-based prediction model, spectrograms are introduced as input and the associated tool wear stage of each spectrogram is used as labels. Hence, the trained CNN–based prediction model can then classify tool wear during machining process with the concept architecture shown in Fig. 26.

5.1.3 Scenario 3: Use of LSTM to predict remaining useful life

In this scenario, the tool remaining useful life (RUL) can be predicted using the LSTM-ANN-based prediction model. In the LSTM method, time-series data captured from the sensors are used as input. Here, the trend of the data is estimated by applying filters on the time-series data, meanwhile the reconstruction loss value is calculated by the prediction model as an output of the LSTM-based prediction model. Afterwards, cutting parameters used during the machining process and the reconstruction loss are used as input to the ANN model. Meanwhile, the amount of tool wear measured by a microscope should be used to label the input data. Thus, the RUL of the cutting tool can be estimated using the LSTM-ANN prediction model with the concept architecture shown in Fig. 27.

5.1.4 Scenario 4: Use of RL for tool life extension with multi-sensor TCM

In the last scenario, the RL-based adaptive control system is proposed to extend tool life by avoiding chatter. In RL-based architectures, there are three main actors: the environment, interpreter, and agent. The main goal of this scenario is to increase tool life while controlling the machining process control parameters adaptively to avoid chatter.

The environment consists of the machining tool and a multi-sensor data collection platform. Here, data from various sources can complement each other for detection of different tool-related issues during machining. For this purpose, using industrial DAQ platform vendors (e.g., National Instrument, Honeywell, Rockwell) is a smart choice as they can accommodate various data collection modules on a standardized bus. This approach also eliminates potential data synchronization difficulties, as all multi-source data are automatically synchronized by the platform.

The interpreter is a prediction model based on SVM, which is trained using the data collected from a multi-sensor system of the environment. By applying FFT, data collected as time-series are converted to frequency domain, which is suitable for chatter detection analysis and accurate labeling. As chatter starts to form around dominant modes of the spindle-tool system during machining, increase in chatter frequencies should be observed. When this occurs, the data should be labeled as chatter, and the remaining data as no-chatter. After the data labeling process is completed, the data are introduced to the SVM model for training. The trained model can then classify incoming data as chatter or no chatter during online monitoring and generate a reward/penalty score for the agent.

The agent comprises of an LSTM-based control system, which can adjust the cutting parameters to extend tool life. Similar to scenario 3, LSTM will need pretraining for prediction of tool life. During online monitoring, the agent receives a reward/penalty score from the interpreter, multi-sensor data, and amount of tool wear. As a control action to the environment, it changes necessary cutting parameters to reduce chatter during machining thus extending tool life. Over time, LSTM can learn and evolve its memory to avoid chatter during machining and adapt to extend tool life. The concept architecture of RL-based multi-sensor TCM is shown in Fig. 28.

6 Conclusions

TCM has attracted significant attention from both academia and industry for the past few decades. As the cutting tool's condition is vital for the stability, quality, and economics of the process, research on TCM will continue to advance by using emerging AI methods and sensory technologies.

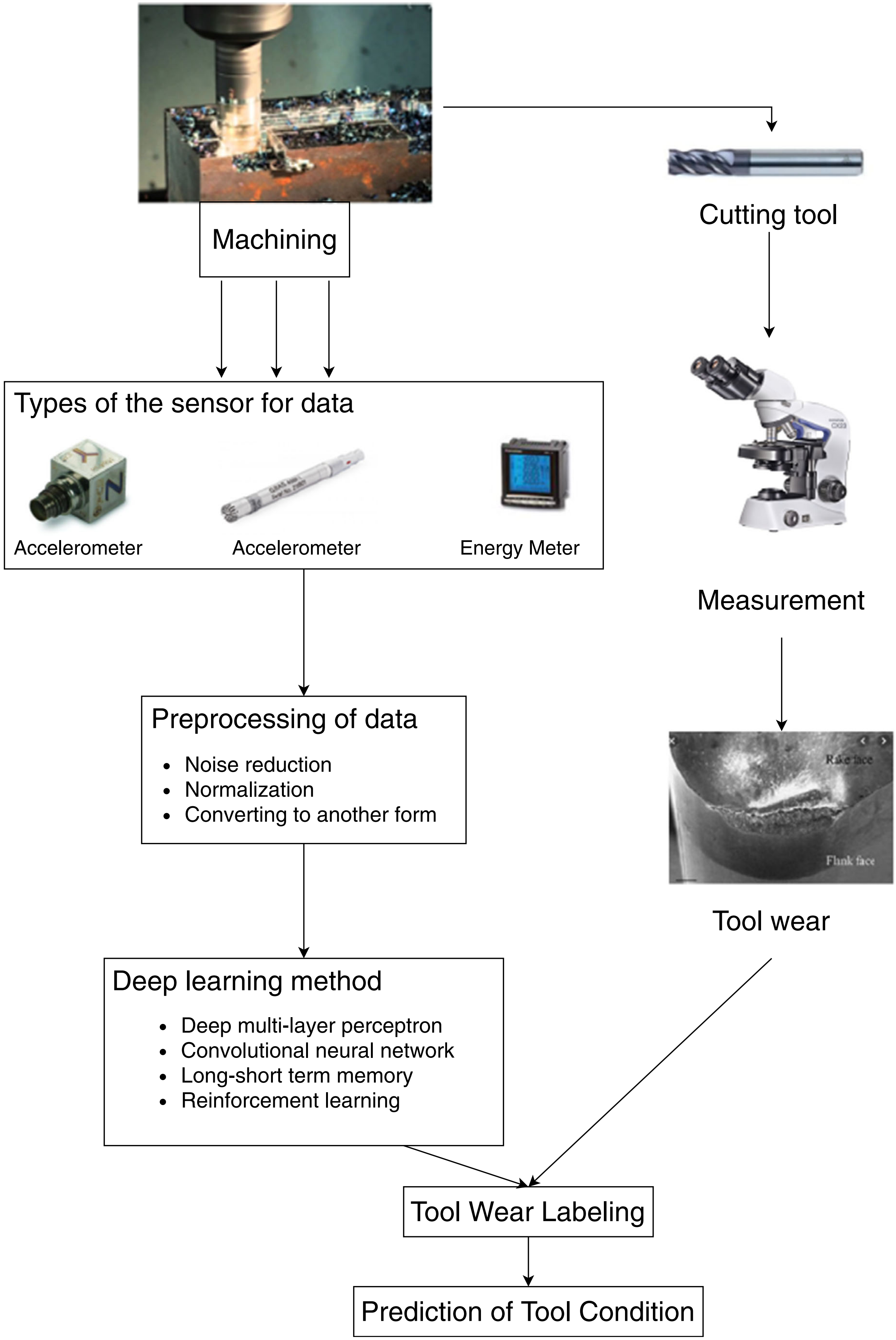

Fig. 24 System architecture for TCM using deep learning methods

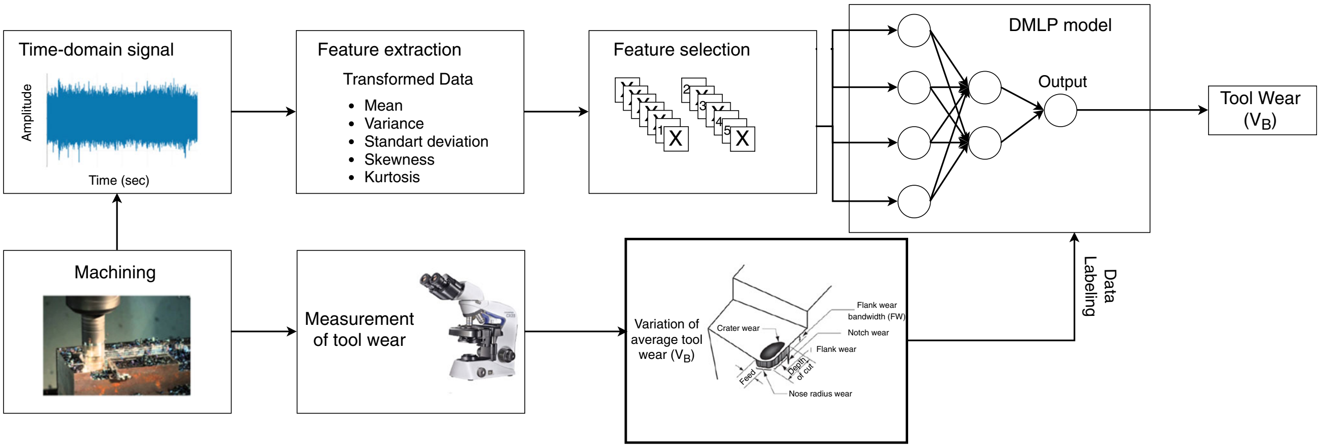

Fig. 25 Concept architecture of tool wear prediction using DMLP

TCM can be utilized in various metal removal processes such as turning, milling, drilling, and broaching, and grinding [100, 102]. The sensing technology in TCM is based on two approaches: direct measurement which relies on visible, measurable wear, and indirect measurements, which relies on artifacts of the process, such as vibration, acoustic emission, and current. Direct methods have the advantage of being more accurate; however, cutting forces cause stronger side effects, creating more research and implementation opportunities for indirect methods. Other reasons that indirect measurement techniques have been receiving greater attention lately are due to the fact that indirect sensory technology is advancing rapidly, it is not invasive to the process as much as direct measurement, and its cost is reducing continuously.

During early research of TCM utilizing classical AI methods, the quality and amount of data were far less than what it is today. During this era, classical techniques such as ANN, fuzzy clustering, and SVM were sufficient to handle limited data released from the process but lacked the accuracy of direct methods, and the ability to adapt to changing

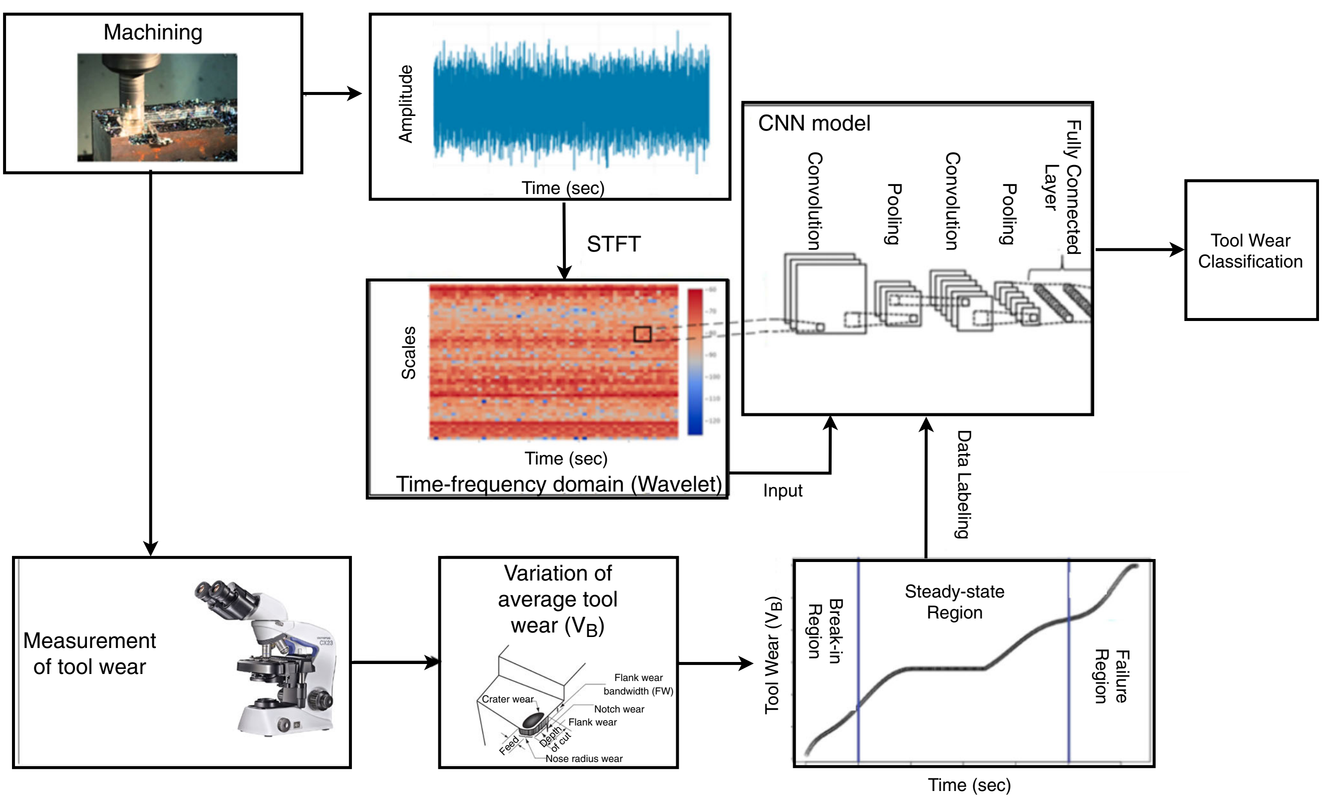

Fig. 26 Concept architecture of tool wear classification using CNN

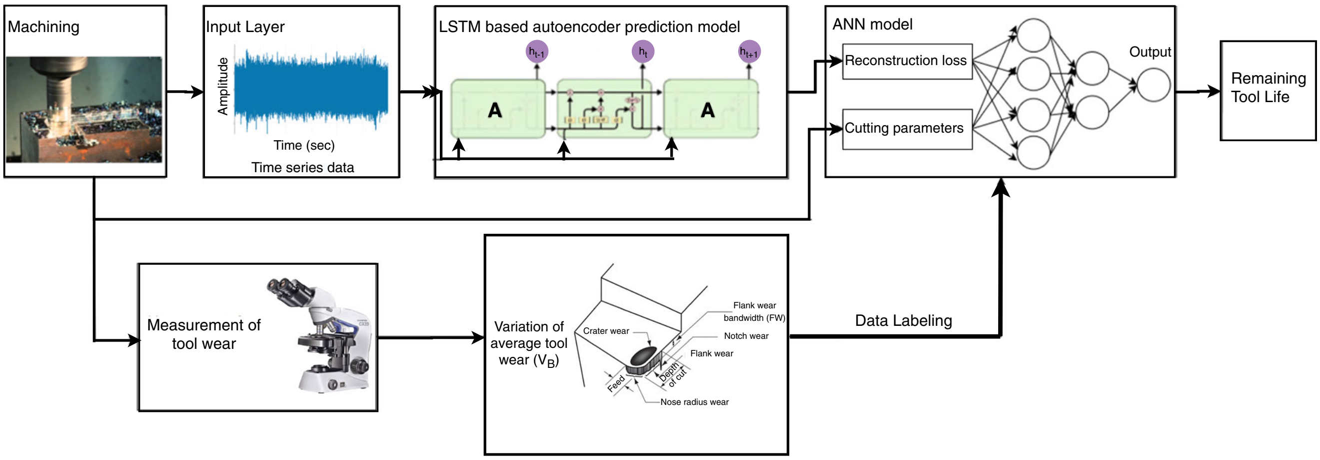

Fig. 27 Concept architecture of RUL prediction using LSTM

conditions and operational parameters during machining, easily. Thus, most remained as diagnostic solutions for cutting cases set by specific cutting conditions, rather than capable of providing prognostic information about generalized process conditions [102].

Since the proliferation of sensory technology at low cost and emergence of the Internet of Things (IoT) paradigm, the data explosion in manufacturing led to the realization of Big Data at the shop floor level. As classical AI methods remain inadequate for handling and processing Big Data, deep learning algorithms which have been showing significant success in application areas that can generate massive data (e.g., autonomous driving, medical image processing, finance, and robotics) present broad opportunities as well as many challenges for TCM.

- In academia, cost of experiments is often high, and measurements required for labeling are intense. In such

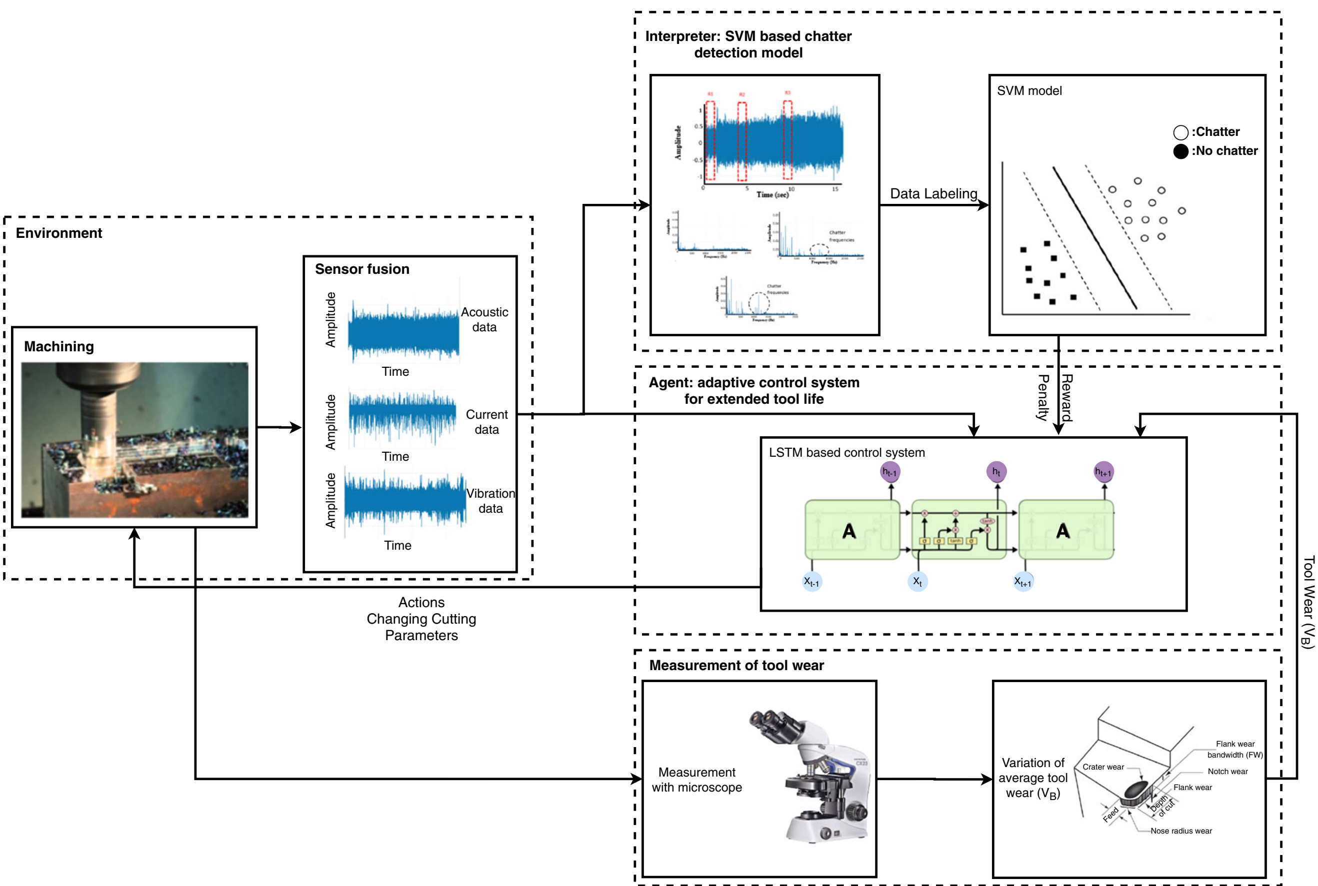

Fig. 28 Concept architecture of tool life extension using RL

conditions, regression to granularly labeled data could be used for production of synthetic data at a lower time scale, which would satisfy the large dataset requirement of DL [12]. As in scenario 1, labeled data segments can be increased by applying second-order regression to the wear curve, generating synthetic data by segmenting them into smaller partitions.

- A major challenge is to choose the right DL method for the process of concern. Although in theory it is possible to apply any DL method to any machining process, the probability of successes depends on the information content and quality of the processed signal which will feed the DL method. Hence, choice of the feature selection/extraction and signal transformation method is of the essence for a high performance system [3], warranting extensive research on efficacy of a variety of signal features and signal transformation/filtering methods such as STFT, WT, Kalman filtering, and Hilbert-Huang transform.
- Among metal cutting methods, interrupted processes such as milling and broaching are more complex due to their periodic forcing nature. However, they can provide more detailed imprints on the state of the tool or mode of wear in their signal over their non-stationary time-frequency profile [22]. This richer information can be detected by the DL successfully, if the right signal transformation technique is used. However, in some cases such as in scenario 2, fine segmentation of time-series data could increase the data size even more, which would create processing and storage bottlenecks.
- Predominant DL methods, such as DMLP, CNN, and LSTM, rely on accurate labeling of Big Data collected from a variety of operational conditions. This require cumbersome direct measurements and expert knowledge labeling for high-quality training, thus placing these methods in the supervised learning category. At the shop floor level there will be vast amounts of collected data from various operational conditions, hence unsupervised and semi-supervised approaches may provide greater opportunity, for enhancing the self-learning capability of DL. This can be achieved by coupling DL methods with unsupervised methods such as k-means clustering and unleash the adaptability of DL methods, where they can continuously update their training during regular operational monitoring. An example could be started by using scenario 2, where the CNN-based architecture can be combined with k-means clustering leading to a semi-supervised architecture that can handle streaming of unlabeled new data, after pretraining.
- Another case of semi-supervised learning can be inherently materialized within the suggested RL architecture as in scenario 4. In this case, a fully connected network like SVM is used as the utility function for RL and the appropriate reward/penalty can be assigned to decisions of SVM to avoid chatter. Hence, SVM will learn and broaden its training over time under various cutting conditions and stability of the operation would be improved, which, in turn reduces tool wear. Furthermore, the robustness of such a scenario could be improved by utilizing multi-sensor data by applying data fusion algorithms [103].
- Another solution to reduce the tremendous task of collecting and labeling large datasets from a variety of cutting conditions could be the application of the transfer learning (TL) approach to DL. By many accounts, TL is defined as adaptation of knowledge for a task to another related task, which could improve the generalization of DL methods and reduce training cost [104]. Such approaches have been used in image recognition fields, where certain groups of pretrained image networks (e.g., ImageNet, AlexNet, VGG, GoogLeNet, ResNet) are provided to worldwide researchers to be used as a base network to transfer to other specific applications of image recognition. This approach may not be helpful much to accomplish TL between process such as from turning to milling, as cutting control parameters and cutting tool form can drastically change. However, it could be utilized for transferring process-specific knowledge within a major type of process group, such as from end milling to profile milling or from one workpiece material to another. A great opportunity here is to build open pretrained networks such as “Turn-Net,” “Mill-Net,” or “Drill-Net.”
- The architecture presented in scenario 3 presents LSTM coupled with ANN in order to provide prognostic information on the tool condition, and improve its remaining useful life (RUL). As additional parameters related to material characteristics such as specific cutting force or tool-related data such as tool diameter or rake angle are used as additional inputs to ANN adaptability of the entire system can be increased rapidly by transferring the base learning from the LSTM network.
- In the near future, there will probably be more implementations of deep learning models on TCM due to the advancements in edge computing, edge AI, and cloud edge–partitioned decision support systems (DSSs) for machining processes. As covered in this review, numerous studies have already been published in the literature and with the increasing influence of Industry 4.0 and Internet of Things diffusion to the factory floor, further research exploring these exciting opportunities is expected.

Funding information This study was funded by The Scientific and Technological Research Council of Turkey (TUBITAK) through project grant no. 118M414.

References

1. Byrne G, Dornfeld D, Inasaki I, Ketteler G, König W, Teti R (1995) Tool condition monitoring (TCM) - the status of research and industrial application. CIRP Ann Manuf Technol 44(2):541–567
2. Zhang C, Yao X, Zhang J, Jin H (2016) Tool condition monitoring and remaining useful life prognostic based on a wireless sensor in dry milling operations. Sensors (Switzerland) 16(6)
3. Roth JT, Djurdjanovic D, Yang X, Mears L, Kurfess T (2010) Quality and inspection of machining operations: tool condition monitoring. J Manuf Sci Eng Trans ASME 132(4):0410151–04101516
4. Siddhpura A, Paurobally R (2013) A review of flank wear prediction methods for tool condition monitoring in a turning process. Int J Adv Manuf Technol 65(1–4):371–393
5. Sick B (2002) On-line and indirect tool wear monitoring in turning with artificial neural networks: a review of more than a decade of research. Mech Syst Signal Process 16(4):487–546
6. Dan L, Mathew J (1990) Tool wear and failure monitoring techniques for turning-a review. Int J Mach Tools Manuf 30(4):579–598
7. Ghani JA, Rizal M, Nuawi MZ, Ghazali MJ, Haron CHC (2011) Monitoring online cutting tool wear using low-cost technique and user-friendly GUI. Wear 271(9–10):2619–2624
8. Dutta S, Pal SK, Sen R (2016) Tool condition monitoring in turning by applying machine vision. J Manuf Sci Eng Trans ASME 138(5):1–17
9. Ambhore N, Kamble D, Chinchanikar S, Wayal V (2015) Tool condition monitoring system: a review. Mater Today Proc 2(4–5):3419–3428
10. Teti R, Jemielniak K, O'Donnell G, Dornfeld D (2010) Advanced monitoring of machining operations. CIRP Ann Manuf Technol 59(2):717–739
11. Rao CHS, Rao DN, Rao RNS (2006) Online prediction of diffusion wear on the flank through tool tip temperature in turning using artificial neural networks. Proc Inst Mech Eng B J Eng Manuf 220(12):2069–2076
12. Serin G, Gudelek MU, Ozbayoglu AM, Unver HO (2017) Estimation of parameters for the free-form machining with deep neural network. In: 2017 IEEE International Conference on Big Data (Big Data), pp 2102–2111. https://doi.org/10.1109/BigData.2017.8258158
13. Widodo A, Yang BS (2007) Support vector machine in machine condition monitoring and fault diagnosis. Mech Syst Signal Process 21(6):2560–2574
14. Saimurugan M, Ramachandran KI, Sugumaran V, Sakthivel NR (2011) Multi component fault diagnosis of rotational mechanical system based on decision tree and support vector machine. Expert Syst Appl 38(4):3819–3826
15. Haykin SS (1999) Neural networks : a comprehensive foundation. Prentice Hall
16. Lauro CH, Brandão LC, Baldo D, Reis RA, Davim JP (2014) Monitoring and processing signal applied in machining processes - a review. Meas J Int Meas Confed 58:73–86
17. Kothuru A, Nooka SP, Liu R (2018) Audio-based condition monitoring in milling of the workpiece material with the hardness variation using support vector machines and convolutional neural networks, ASME 2018 13th Int. Manuf. Sci. Eng. Conf. MSEC 2018, vol. 4, no. November 2018, pp 1–9
18. Kothuru A, Nooka SP, Liu R (2018) Application of audible sound signals for tool wear monitoring using machine learning techniques in end milling. Int J Adv Manuf Technol 95(9–12):3797–3808
19. Olufayo O, Abou-El-Hossein K (2015) Tool life estimation based on acoustic emission monitoring in end-milling of H13 mould-steel. Int J Adv Manuf Technol 81(1–4):39–51
20. Li X (2002) A brief review: acoustic emission method for tool wear monitoring during turning. Int J Mach Tools Manuf 42(2):157–165
21. Kakade S, Vijayaraghavan L, Krishnamurthy R (1994) In-process tool wear and chip-form monitoring in face milling operation using acoustic emission. J Mater Process Technol 44(3–4):207–214
22. Marinescu I, Axinte DA (2008) A critical analysis of effectiveness of acoustic emission signals to detect tool and workpiece malfunctions in milling operations. Int J Mach Tools Manuf 48(10):1148–1160
23. Antić A, Popović B, Krstanović L, Obradović R, Milošević M (2018) Novel texture-based descriptors for tool wear condition monitoring. Mech Syst Signal Process 98:1–15
24. Du Kim J, Choi IH (1996) Development of a tool failure detection system using multi-sensors. Int J Mach Tools Manuf 36(8):861–870
25. Wu D, Jennings C, Terpenny J, Kumara S, Gao RX (2018) Cloud-based parallel machine learning for tool wear prediction. J Manuf Sci Eng Trans ASME 140(4):1–10
26. Zhang XY, Lu X, Wang S, Wang W, Li WD (2018) A multi-sensor based online tool condition monitoring system for milling process. Procedia CIRP 72:1136–1141
27. Chen SL, Jen YW (2000) Data fusion neural network for tool condition monitoring in CNC milling machining. Int J Mach Tools Manuf 40(3):381–400
28. Mali R, Telsang MT, Gupta TVK (2017) Real time tool wear condition monitoring in hard turning of Inconel 718 using sensor fusion system. Mater Today Proc 4(8):8605–8612
29. Kious M, Ouahabi A, Boudraa M, Serra R, Cheknane A (2010) Detection process approach of tool wear in high speed milling. Meas J Int Meas Confed 43(10):1439–1446
30. Rmili W, Ouahabi A, Serra R, Leroy R (2016) An automatic system based on vibratory analysis for cutting tool wear monitoring. Meas J Int Meas Confed 77:117–123
31. Liang SY, Hecker RL, Landers RG (2002) Machining process monitoring and control: the state-of-the-art. ASME Int Mech Eng Congr Expo Proc 126(May 2004):599–610
32. Abellan-Nebot JV, Romero Subirón F (2010) A review of machining monitoring systems based on artificial intelligence process models. Int J Adv Manuf Technol 47(1–4):237–257
33. Dimla DE, Lister PM (2000) On-line metal cutting tool condition monitoring. II: tool-state classification using multi-layer perceptron neural networks. Int J Mach Tools Manuf 40(5):769–781
34. Bombiński S, Błazejak K, Nejman M, Jemielniak K (2016) Sensor signal segmentation for tool condition monitoring. Procedia CIRP 46:155–160
35. Nouri M, Fussell BK, Ziniti BL, Linder E (2015) Real-time tool wear monitoring in milling using a cutting condition independent method. Int J Mach Tools Manuf 89:1–13
36. Cho S, Asfour S, Onar A, Kaundinya N (2005) Tool breakage detection using support vector machine learning in a milling process. Int J Mach Tools Manuf 45(3):241–249
37. Wang J, Xie J, Zhao R, Zhang L, Duan L (2017) Multisensory fusion based virtual tool wear sensing for ubiquitous manufacturing. Robot Comput Integr Manuf 45:47–58
38. Caggiano A, Angelone R, Napolitano F, Nele L, Teti R (2018) Dimensionality reduction of sensorial features by principal component analysis for ANN machine learning in tool condition monitoring of CFRP drilling. Procedia CIRP 78:307–312
39. Rizal M, Ghani JA, Nuawi MZ, Haron CHC (2013) Online tool wear prediction system in the turning process using an adaptive

neuro-fuzzy inference system. Appl Soft Comput J 13(4):1960–1968
40. Venkata Rao K, Murthy BSN, Mohan Rao N (2013) Cutting tool condition monitoring by analyzing surface roughness, work piece vibration and volume of metal removed for AISI 1040 steel in boring. Meas J Int Meas Confed 46(10):4075–4084
41. Özel T, Karpat Y (Apr. 2005) Predictive modeling of surface roughness and tool wear in hard turning using regression and neural networks. Int J Mach Tools Manuf 45(4–5):467–479
42. Korkut I, Acir A, Boy M (Sep. 2011) Application of regression and artificial neural network analysis in modelling of tool-chip interface temperature in machining. Expert Syst Appl 38(9): 11651–11656
43. Onal AC, Berat Sezer O, Ozbayoglu M, Dogdu E (2018) MIS-IoT: modular intelligent server based Internet of Things framework with big data and machine learning, in Proceedings - 2018 IEEE International Conference on Big Data. Big Data 2019:2270–2279
44. Smola A, Vishwanathan SVN, Clara S (2008) Introduction to machine learning, 1st edn. Cambridge University Press, United Kingdom
45. Russell SJ, Norvig P, Canny JF, Malik JM, Edwards DD (1995) Artificial intelligence: a modern approach, 3rd edn. Pearson Education Inc., USA
46. I. Goodfellow, Y. Bengio, and A Courville, (2016) Deep learning. Available: https://www.deeplearningbook.org/.
47. Singh N, Singh DP, Pant B (2018) A comprehensive study of big data machine learning approaches and challenges. Proc - 2017 Int Conf Next Gener Comput Inf Syst ICNGCIS 2017:58–63
48. Soniya SP, Singh L (2016) A review on advances in deep learning. 2015 IEEE Work Comput Intell Theor Appl Futur Dir WCI 2015: 1–6
49. Deshmukh PS (2018) Travel time prediction using neural networks: a literature review, 2018 Int. Conf. Information, Commun. Eng. Technol. ICICET 2018, no. xi, pp 1–5
50. Bhandare A, Bhide M, Gokhale P, Chandavarkar R (2016) Applications of convolutional neural networks. Int J Comput Sci Inf Technol 7(5):2206–2215
51. Ramos S, Gehrig S, Pinggera P, Franke U, Rother C (2016) Detecting unexpected obstacles for self-driving cars: fusing deep learning and geometric modeling. IEEE Intelligent Vehicles Symposium
52. Pak M, Kim S (2018) A review of deep learning in image recognition, Proc. 2017 4th Int. Conf. Comput. Appl. Inf. Process. Technol. CAIPT 2017, vol. 2018-Janua, pp 1–3
53. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. Adv Neural Inf Proces Syst. https://doi.org/10.1145/3065386
54. Zeiler MD, Fergus R (2014) Visualizing and understanding convolutional networks, Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8689 LNCS, no. PART 1, pp 818–833
55. Szegedy C, Liu W, Jia YQ, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 07-12-June, pp 1–9
56. Molchanov P, Gupta S, Kim K, Kautz J (2015) Hand gesture recognition with 3D convolutional neural networks. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, vol 2015-October, pp 1–7
57. Molchanov P, Yang X, Gupta S, Kim K, Tyree S, Kautz J (2016) Online detection and classification of dynamic hand gestures with recurrent 3D convolutional neural networks. In: Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp 4207–4215
58. X. Chai, Z. Liu, F. Yin, Z. Liu, and X. Chen, Two streams recurrent neural networks for large-scale continuous gesture recognition, Proc Int Conf Pattern Recogn, vol. 0, pp. 31–36, 2016.
59. Wang P, Li W, Liu S, Gao Z, Tang C, Ogunbona P (2017) Large-scale isolated gesture recognition using convolutional neural networks. https://doi.org/10.1109/ICPR.2016.7899599
60. Khalajzadeh H, Mansouri M, Teshnehlab M (2014) Face recognition using convolutional neural network and simple logistic classifier, in Advances in Intelligent Systems and. Computing 223: 197–207
61. Zhang X, Peng M, Chen T (2016) Face recognition from near-infrared images with convolutional neural network, 2016 8th Int. Conf Wirel Commun Signal Process WCSP 2016:1–5
62. Ramaiah NP, Ijjina EP, Mohan CK (2015) Illumination invariant face recognition using convolutional neural networks. In: 2015 IEEE Int. Conf. Signal Process. Informatics, Commun. Energy Syst. SPICES 2015, pp 1–4
63. Pinheiro PO, Com RC (2014) B1b_CRFasRNN 32
64. Wang P, Cao Y, Shen C, Liu L, Shen HT (2015) Temporal pyramid pooling based convolutional neural networks for action recognition. IEEE Trans Circ Syst Video Technol. https://doi.org/10.1109/TCSVT.2016.2576761
65. Simonyan K, Zisserman A Two-stream convolutional networks for action recognition in videos. Adv Neural Inf Proces Syst
66. Yang Q, He Z, Ge F, Zhang Y (2017) Sequence-to-sequence prediction of personal computer software by recurrent neural network. In: Proceedings of the International Joint Conference on Neural Networks, vol 2017-May, pp 934–940
67. Sharma P, Singh A (2017) Era of deep neural networks: a review, 8th Int. Conf Comput Commun Netw Technol 2017
68. N. Neverova, C. Wolf, G. Paci, G. Sommavilla, G. W. Taylor, and F. Nebout, A multi-scale approach to gesture detection and recognition, Proc IEEE Int Conf Comput Vis, pp. 484–491, 2013.
69. Tsironi E, Barros P, Wermter S (2016) Gesture recognition with a convolutional long short-term memory recurrent neural network, ESANN 2016 - 24th Eur. Symp. Artif. Neural Networks, pp 213–218
70. Hochreiter S, Schmidhuber J (Nov. 1997) Long short-term memory. Neural Comput 9(8):1735–1780
71. A. Graves, A. Mohamed, and G. Hinton (2013) Speech recognition with deep recurrent neural networks, IEEE International Conference on Acoustics, Speech and Signal Processing. https://doi.org/10.1109/ICASSP.2013.6638947
72. Sundermeyer M, Ney H, Schluter R (2015) From feedforward to recurrent LSTM neural networks for language modeling. IEEE Trans Audio Speech Lang Process 23(3):517–529
73. Mayer NM (2017) Echo state condition at the critical point. Entropy 19(1)
74. Ororbia A, Mikolov T, Reitter D (2017) Learning simpler language models with the delta recurrent neural network framework. Arxiv:1–27
75. Day MY, Da Lin Y (2017) Deep learning for sentiment analysis on Google play consumer review. In: Proc. - 2017 IEEE Int. Conf. Inf. Reuse Integr. IRI 2017, vol 2017-Janua, pp 382–388
76. Arbib MA (1995) The handbook of brain theory and neural networks. MIT Press
77. Sutton RS, Barto AG (2017) Reinforcement learning an introduction. The MIT press
78. Aghazadeh F, Tahan A, Thomas M (2018) Tool condition monitoring using spectral subtraction and convolutional neural networks in milling process. Int J Adv Manuf Technol 98(9–12): 3217–3227
79. Chen Y, Jin Y, Jiri G (2018) Predicting tool wear with multi-sensor data using deep belief networks. Int J Adv Manuf Technol 99(5–8):1917–1926

80. Shao SY, Sun WJ, Yan RQ, Wang P, Gao RX (2017) A deep learning approach for fault diagnosis of induction motors in manufacturing. Chin J Mech Eng (English Ed) 30(6):1347–1356
81. Riera-Guasp M, Antonino-Daviu JA, Capolino GA (2015) Advances in electrical machine, power electronic, and drive condition monitoring and fault detection: state of the art. IEEE Trans Ind Electron 62(3):1746–1759
82. Zhao R, Yan R, Chen Z, Mao K, Wang P, Gao RX (2019) Deep learning and its applications to machine health monitoring. Mech Syst Signal Process 115:213–237
83. Li X, Zhang W, Ding Q, Sun JQ (2018) Intelligent rotating machinery fault diagnosis based on deep learning using data augmentation. J Intell Manuf
84. Fu Y, Zhang Y, Gao H, Mao T, Zhou H, Sun R, Li D (2019) Automatic feature constructing from vibration signals for machining state monitoring. J Intell Manuf 30(3):995–1008
85. Liu H, Li L, Ma J (2016) Rolling bearing fault diagnosis based on STFT-deep learning and sound signals. Shock Vib. https://doi.org/10.1155/2016/6127479
86. Jia F, Lei Y, Guo L, Lin J, Xing S (Jan. 2018) A neural network constructed by deep learning technique and its application to intelligent fault diagnosis of machines. Neurocomputing 272:619–628
87. Chen Z, Deng S, Chen X, Li C, Sanchez RV, Qin H (2017) Deep neural networks-based rolling bearing fault diagnosis. Microelectron Reliab 75:327–333
88. Deng S, Cheng Z, Li C, Yao X, Chen Z, Sanchez RV (2017) Rolling bearing fault diagnosis based on deep Boltzmann machines. In: Proceedings of 2016 Prognostics and System Health Management Conference, PHM-Chengdu 2016
89. Li C, Sánchez RV, Zurita G, Cerrada M, Cabrera D (2016) Fault diagnosis for rotating machinery using vibration measurement deep statistical feature learning. Sensors (Switzerland) 16(6)
90. Guo X, Chen L, Shen C (Nov. 2016) Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. Meas J Int Meas Confed 93:490–502
91. Fuan W, Hongkai J, Haidong S, Wenjing D, Shuaipeng W (2017) An adaptive deep convolutional neural network for rolling bearing fault diagnosis. Meas Sci Technol 28(9)
92. Dornfeld DA, DeVries MF (1990) Neural network sensor fusion for tool condition monitoring. CIRP Ann Manuf Technol 39(1): 101–105
93. Hsieh WH, Lu MC, Chiou SJ (2012) Application of backpropagation neural network for spindle vibration-based tool wear monitoring in micro-milling. Int J Adv Manuf Technol 61(1–4):53–61
94. Srinivasan A, Dornfeld D, Bhinge R (2016) Integrated vibration and acoustic data fusion for chatter and tool condition classification in milling, in International Symposium on Flexible Automation, ISFA 201, pp 263–266
95. Quiza R, Figueira L, Davim JP (2008) Comparing statistical models and artificial neural networks on predicting the tool wear in hard machining D2 AISI steel. Int J Adv Manuf Technol 37(7–8):641–648
96. Silva RG, Wilcox SJ, Reuben RL (2006) Development of a system for monitoring tool wear using artificial intelligence techniques. Proc Inst Mech Eng B J Eng Manuf 220(8):1333–1346
97. Rangwala S, Dornfeld D (1990) Sensor integration using neural networks for intelligent tool condition monitoring. J Manuf Sci Eng Trans ASME 112(3):219–228
98. Griffin JM, Doberti AJ, Hernández V, Miranda NA, Vélez MA (2017) Multiple classification of the force and acceleration signals extracted during multiple machine processes: part 1 intelligent classification from an anomaly perspective. Int J Adv Manuf Technol 93(1-4):811–823
99. Griffin JM, Doberti AJ, Hernández V, Miranda NA, Vélez MA (2017) Multiple classification of the force and acceleration signals extracted during multiple machine processes: part 2 intelligent classification from an anomaly perspective. Int J Adv Manuf Technol 92(9-12):3207–3217
100. Rehorn AG, Jiang J, Orban PE (2005) State-of-the-art methods and results in tool condition monitoring: a review. Int J Adv Manuf Technol 26(7-8):693–710
101. T. Mohanraj, S. Shankar, R. Rajasekar, N.R. Sakthivel, and A. Pramanik, Tool condition monitoring techniques in milling process—a review, J Mater Res Technol, In Press, 2019.
102. Zhou Y, Xue W (2018) Review of tool condition monitoring methods in milling processes. Int J Adv Manuf Technol 96(5-8): 2509–2523
103. Duro JA, Padget JA, Bowen CR, Kim HA, Nassehi A (2016) Multi-sensor data fusion framework for CNC machining monitoring. Mech Syst Signal Process 66:505–520
104. Li X, Hu Y, Li M, Zheng J (2020) Fault diagnostics between different type of components: a transfer learning approach. Appl Soft Comput 86(105950)