MYU Tokyo

S & M 3894

Development of an Offline Tool Wear Detection Method Based on Audio Signals

Ching-hsiang Yang, Huan-kai Chau, and Wei-chen Lee*

Department of Mechanical Engineering, National Taiwan University of Science and Technology, No. 43, Keelung Road, Section 4, Taipei 106335, Taiwan

(Received August 22, 2024; accepted January 6, 2025)

Keywords: offline, tool wear detection, deep learning, audio, classification

Tool wear substantially affects efficiency, accuracy, and cost in the manufacturing industry. The objective of this research was to propose an innovative method to detect tool wear offline on the basis of a deep learning classification model using audio signals. Tool wear experiments were conducted, and an International Organization for Standardization (ISO) standard was used to categorize tool wear into three levels. Using offline signals and three-level categorization can help non-professionals who operate cutting machines easily determine the tool's condition. A mechanism was designed to collect the audio signals generated by a tool offline to avoid inconsistencies in manual operation. The collected signals were then converted into the frequency domain using the fast Fourier transform (FFT) to facilitate the observation of frequency variations of the signals, followed by the normalization and extraction of the frequency range. Data augmentation techniques were used to help generate more data to increase the robustness of the classification model, which was built on the basis of convolutional neural networks (CNNs). The results showed that the CNN model achieved an accuracy of 84.44% in classifying the wear of unseen tools, which outperformed the results obtained using the method in one of the previous research studies. In summary, we demonstrated that offline audio signals can be used for tool wear detection, providing a simple solution to detect tool wear without using the complicated online tool wear detection approach.

1. Introduction

As the use time of a cutting tool increases, the tool gradually becomes worn and can cause machining problems, such as low accuracy and high cutting force. Therefore, the development of a tool wear detection method is imperative in the manufacturing industry. The International Organization for Standardization (ISO) has proposed criteria(1) for tool wear, which include situations such as tool flank wear, chipping, cracks, and catastrophic failure. In recent years, more and more non-professionals operating cutting machines have become common. For example, some dental clinics have multi-axis engraving machines specializing in the manufacture of dental crowns and bridges. However, technicians operating these machines

*Corresponding author: e-mail: wclee@mail.ntust.edu.tw
https://doi.org/10.18494/SAM5324

ISSN 0914-4935

usually lack sufficient machining knowledge and do not know the tool wear conditions, leading to inappropriate sizes of dental products. As a result, it is necessary to have a method that will enable non-professionals to easily detect tool wear.

Many institutions and researchers have developed techniques to detect tool wear. Various methods have been proposed, such as image, force, vibration, and audio detection.[2,3] Furthermore, numerous studies have focused on the use of multiple sensors to improve the accuracy of the predictions.[4–6] However, the use of multiple sensors also increases the complexity of implementing such an approach on cutting machines. In contrast, audio-signal-based tool wear detection has several advantages, including noncontact, nonintrusive, simple, and stable measurements. The audio signals can be collected online or offline. Online collection means that the audio signals are collected directly while the tool is cutting. Several studies using audio signals have been dedicated to online detection.[7–12] However, online audio signal detection is challenging because the audio signals would be significantly affected by noises from many sources during the machining process. Furthermore, installing audio sensors on a cutting machine for online detection may not appeal to some users.

To address this issue, we proposed an offline tool wear detection system based on audio signals. Offline means that the audio signals from a tool are generated and recorded in a controllable environment, which makes the approach simple and reliable.

2. Materials and Methods

2.1 Overview

The proposed offline tool wear detection method based on audio signals generated by a tool is shown in Fig. 1. Initially, we conducted tool wear experiments and used the ISO standard[1] to categorize tool wear. Owing to the challenges in manually collecting audio data, we developed an audio collection device. To analyze the frequency changes associated with various wear levels, we applied the fast Fourier transform (FFT) to the collected audio signals. Subsequently, the frequency domain signal was normalized to standardize the magnitude range, and the most distinctive frequency range was extracted as a feature for deep learning. To enhance the model's robustness, the data augmentation technique was employed, efficiently generating a substantial amount of training data. On the basis of previous research[13–20] highlighting the success of convolutional neural network (CNN) models in audio classification, we also adopted the CNN approach for tool wear classification in this study.

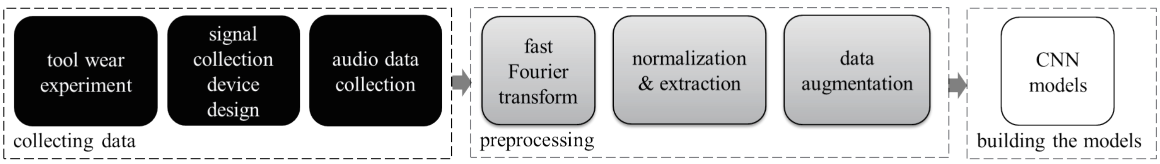


Fig. 1. Overview of tool wear detection method proposed in this research.

2.2 Tool wear experiment

To obtain tools with various levels of wear, we designed an experiment specifically aimed at generating such tools. In this experiment, we utilized the tools (part no. ZB-05-V1) from ARUM (Germany). This tool is made of tungsten carbide and features a ball-end cutting tool with a diamond-coated surface. The radius at the end of the tool is 1 mm. The tool was installed in a five-axis machining center (model no. CT-350) made by Tongtai (Taiwan), and the workpiece was made of low-carbon steel.

The machining plan of the workpiece to generate various levels of wear on the cutting tool is as follows. Each cutting path had a length of 50 mm and the spindle speed was set to 10000 rpm. To avoid any potential tool breakage caused by excessive cutting resistance, we limited the cutting depth to 0.3 mm and set the feed rate to 5 mm/min. Cutting fluid was used throughout the cutting process to reduce temperature and improve chip removal. The slot milling was adopted to generate uniform tool wear on both flanks of the tools. The machining path is illustrated in Fig. 2. The process started from the edge of the workpiece. As the cutting length reached 25 mm in the y-direction, a lateral movement of 1.2 mm was introduced in the x-direction. Finally, the cutting continued in the negative y-direction for another 25 mm, completing one machining pattern. We repeated the cutting of the workpiece using the same pattern at a nearby location until the desired level of tool wear was reached.

After each machining pattern was completed, a 2D vision measuring machine was used to examine the tool flank wear. We noticed a correlation between the wear area on the tool flank and the cutting length. As the cutting length increased, the wear area of the tool gradually increased as well. We referred to the tool deterioration phenomena described in ISO 8688-2:1989[1] and categorized the wear into three levels: "new" for tools without any cutting performed, "worn" for tools with uniform wear on the tool flank, and "failed" for tools with chipping. The definitions of tool wear levels are listed in Table 1. The "new" condition indicates that the tool is completely new and has not been cut. The "worn" and "failed" conditions represent the accumulated cutting lengths of 150 and 300 mm, respectively. We wanted to report

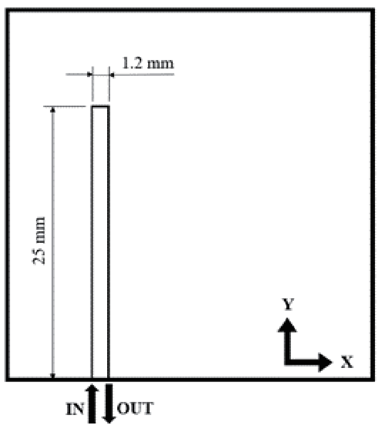
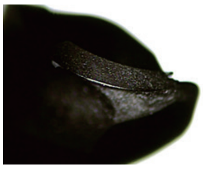
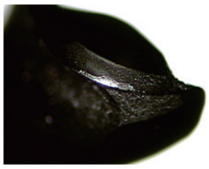


Fig. 2. Cutting pattern used to generate the tool wear.

Table 1
Definitions of various tool wear levels and their corresponding accumulated cutting lengths.

Tool-end image			
Accumulated cutting length	0 mm	150 mm	300 mm
ISO 8688-2: 1989	—	tool flank wear	chipping
Wear condition	new	worn	failed

the state of the tool as “new,” “worn” or “failed” so that non-professionals can easily understand whether the tool can still be used. If a cutting tool is in a “failed” state, it is at its end of life and cannot be used anymore.

2.3 Audio data collection

The audio signals of a tool were generated by holding the tool and using the tool end to scratch an aluminum plate, as illustrated in Fig. 3. A linear motion module consisting of a hydraulic rod was developed to consistently make the scratch with the tool. The completed audio data collection device with the linear motion module is shown in Fig. 4. Using the device can reduce the effects of manual operation, such as varying the scratching force, tool holding angle, and scratching speed, during the data collection.

To use the device, we first fixed the tool on the cutting tool holder, pushed the tool holder to one side of the aluminum plate, and released the tool holder. Owing to the spring in the hydraulic rod, the tool holder slowly returned to its original location, causing the end of the tool to scratch on the aluminum plate. The audio signals were collected by the microphone next to the aluminum plate. The microphone that we used in this study was a condenser microphone (part no. AM310) made by AVerMedia (Taiwan). The settings for recording audio signals included a sampling rate of 48 kHz and a 16-bit sampling capacity.

Figure 5 shows the time-domain plots of audio signals corresponding to different levels of tool wear. The total time length was about 1 s. To avoid misclassifying non-tool-generated audio signals as tool-generated signals, additional audio signals, including operation noises from the mechanism and background noises, were collected and classified as “env.”

We used three tools to collect their audio signals while at the “new,” “worn,” and “failed,” levels. Twenty audio data were collected for each tool at one wear level. Thus, 180 raw data were collected for the three tools at all three levels. Furthermore, 60 “env” audio data were also collected to allow us to have 240 audio signals in total.

2.4 Data preprocessing

FFT was then employed to convert audio signals into frequency domain representations, allowing the observation of frequency distributions under different tool wear levels. The analysis

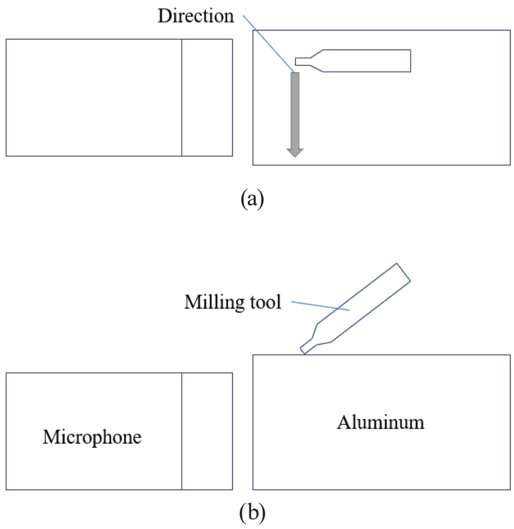


Fig. 3. (Color online) Audio signals generated by scratching an aluminum plate with the tool were collected using the microphone next to the aluminum plate. (a) Top and (b) front views.

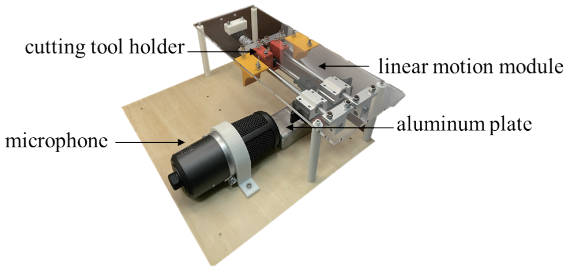


Fig. 4. (Color online) Completed audio data collection device.

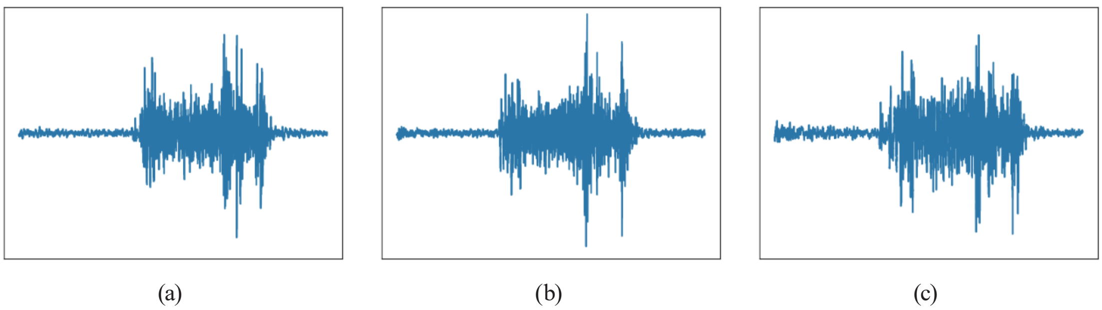


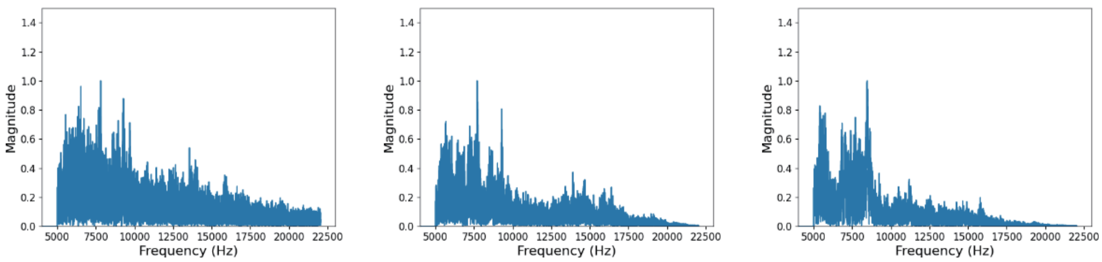
Fig. 5. (Color online) Time-domain plots of audio signals of tool wear level: (a) new, (b) worn, and (c) failed.

revealed that environmental audio signals lack significant resonance peaks beyond 5000 Hz, while tool-generated signals exhibit distinct peaks after this frequency. This distinction enables straightforward differentiation between the two types of audio signal. To ensure comparability, a frequency normalization technique was applied, normalizing audio signals above 5000 Hz to the range of zero and one. Figure 6 shows the frequency spectra of different tool wear levels after normalization, revealing three distinct frequency distribution patterns. For "new" tools, as the frequency increases, the magnitudes gradually decrease from low to high at a consistent rate. "Worn" tools exhibit larger magnitudes concentrated before 9265 Hz. In contrast, "failed" tools show a higher magnitude concentration before 7700 Hz. These characteristics provide valuable information for the successful classification of tool wear using audio signals.

Data augmentation techniques were utilized to obtain more data to train the CNN model and enhance its robustness. These techniques include time shifting, amplification, and noise addition. Time shifting randomly displaces the raw audio signal from −0.2 to +0.2 s. Amplification means multiplying the amplitude of the raw audio signal by a random number between 0.8 and 1.2. For noise addition, random noises with a normal distribution in the range of 0 to 0.0004 dB were added to the audio signals, considering the microphone's average noise value of approximately 0.0008 dB. These techniques were applied to each wear level, generating 65 additional data per level, leading to 260 additional data. The complete audio dataset consists of 500 audio data after the data augmentation process, as listed in Table 2.

2.5 Training the CNN model

In this study, we utilized CNN for tool wear classification. The data of 500 audio signals were divided into training, validation, and testing sets in an approximate 6:1:3 ratio. The training data comprised 300 audio data for model training, whereas 52 data were used for validation. The remaining 148 data were used to test the performance of the model.



(a) (b) (c)

Fig. 6. (Color online) Frequency spectra of tool wear level: (a) new, (b) worn, (c) failed after normalization.

Table 2
Total amount of data after data augmentation.

Category	No. of raw data	No. of augmented data	Summation
"new"	60	65	125
"worn"	60	65	125
"failed"	60	65	125
"env"	60	65	125
Total	240	260	500

The model took preprocessed audio signals as input. For feature extraction, two 1D convolutional layers with filter sizes of 16 and 32 were employed, activated by the ReLU activation function for nonlinear transformation. Each convolutional layer was followed by a max pooling layer to reduce the size of the feature data. A flattened layer was used to convert the feature data into a one-dimensional vector for the subsequent dense layer. A dropout layer was introduced after the dense layer to prevent overfitting. The output dense layer had the number of neurons corresponding to the classification levels.

3. Results and Discussion

3.1 Results of CNN model

We explored the best frequency range of audio signals for tool wear classification using the CNN model. Multiple CNN models were constructed using different frequency ranges from the training data, and their accuracy was evaluated on the test data. The results indicated that the CNN model achieved the highest accuracy of 99.32%, calculated using the confusion matrix shown in Fig. 7, when the frequency was between 7000 and 15000 Hz. As shown in Fig. 7, the model effectively distinguished between environmental and tool-generated audio signals.

3.2 Effects of preprocessing

We used various preprocessing techniques, including normalization and data augmentation. The effects of these preprocessing techniques on the CNN classification model were investigated here. As listed in Table 3, if we did not adopt the data augmentation technique, the accuracies were 88.89 and 90.28%. However, when data augmentation was adopted, the accuracies were improved to 97.97 and 99.32%. The normalization technique may increase or decrease the

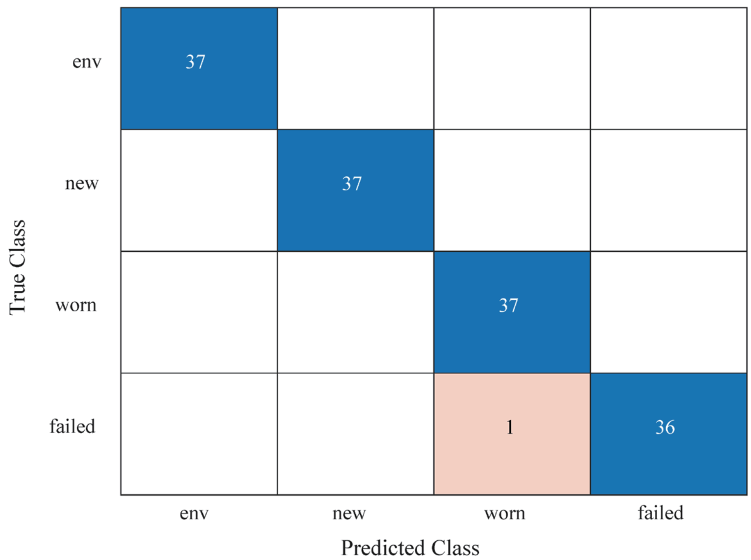


Fig. 7. (Color online) Confusion matrix for CNN model to classify the sample tools (accuracy = 99.32%).

Table 3
Effects of preprocessing on the accuracies of the CNN model: “Y” means that the technique was used and “N” means that the technique was not used.

Normalization	Data augmentation	Accuracy (%)
N	N	90.28
Y	N	88.89
N	Y	97.97
Y	Y	99.32

accuracy by approximately 1.4%, depending on whether the data augmentation technique was used. It can be realized that the data augmentation technique can help, but normalization may or may not help improve the accuracy of the CNN model in this research.

3.3 Out-of-sample test for CNN model

After obtaining the trained CNN model, we then applied it to evaluate the wear of nine tools whose audio signals were not collected previously. The nine unseen tools include three “new” tools, three “worn” tools, and three “failed” tools, as shown in Fig. 8. Ten audio data were collected for each tool, so 90 raw data were obtained in total. The confusion matrix for the CNN classification model is shown in Fig. 9. The model achieved an accuracy of 84.44% for classifying the tools. As shown in Fig. 9, the model performed well in distinguishing between the “new” and “worn” tools, but it misclassified two “worn” tools into “failed” tools. A significant error occurred when the model was used on the failed tools, where there was a 40% (12 out of 30) misclassification.

To understand the misclassifications made by the CNN model, we analyzed the corresponding frequency spectra. The frequency distributions of the audio signals generated by the “new” and “worn” unseen tools were similar to those obtained when using the samples for training. However, the frequency distribution of the “failed” tools differed significantly. Upon closer examination of the misclassified audio data, we found that among the three “failed” tools we used in the out-of-sample test, one exhibited chipping only, while the other two showed chipping and coating delamination. The audio signals from the tools with coating delamination had a distinct frequency distribution, combining the characteristics of the “worn” and “failed” tools. This may explain why the CNN model misclassified some audio signals from the “failed” tools as those from the “worn” tools.

3.4 Comparison with previous research

In a previous study,[7] various feature extraction methods and classification models were tested to monitor the condition of face milling tools using audio signals. Their results showed that the highest accuracy was achieved by combining the discrete wavelet transform (DWT) with the support vector machine (SVM). On the basis of their approach, we applied the DWT to extract the first seven wavelet coefficients from audio signals and trained an SVM classification model. The SVM model achieved an accuracy of 80.40% in the test data and an accuracy of

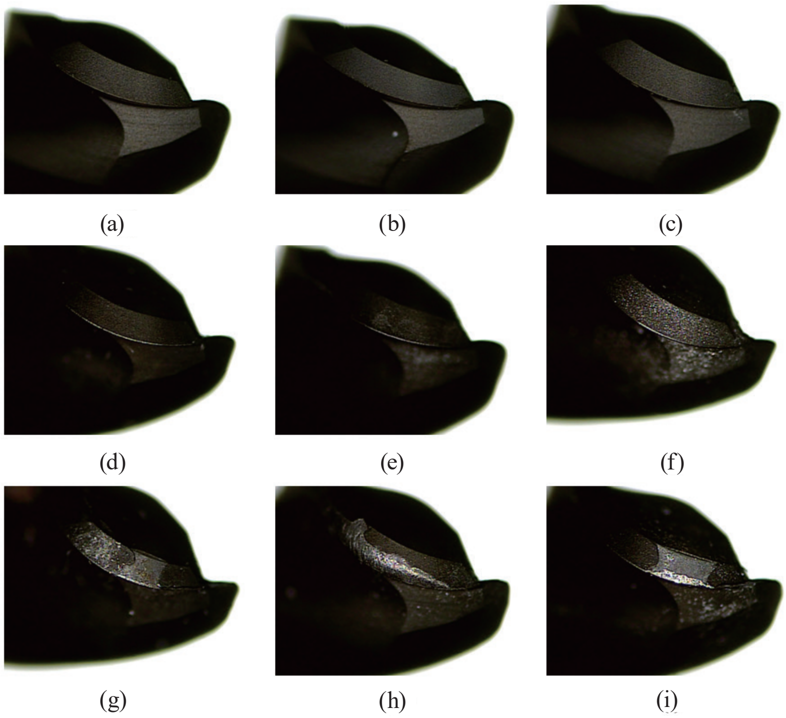


Fig. 8. The tool-end images were captured using the 2D vision measuring machine. (a) to (c), (d) to (f), and (g) to (i) are "new", "worn", and "failed" tools, respectively.

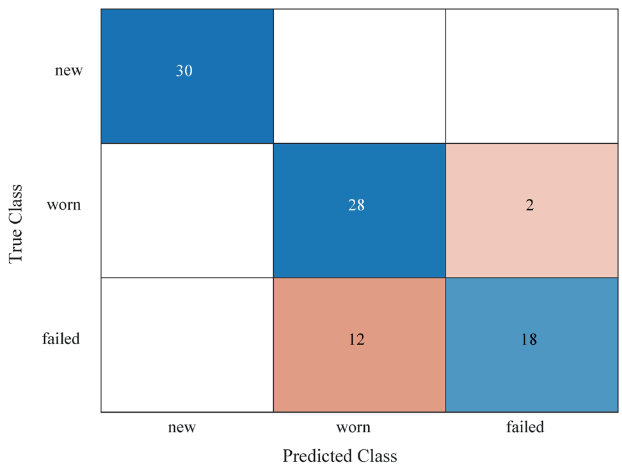


Fig. 9. (Color online) Confusion matrix of CNN model in classifying unseen tools (accuracy = 84.44%).

51.11% in the out-of-sample data. The results obtained using the proposed method and those obtained using the method of the previous research are compared in Table 4. These results demonstrate that our proposed method for tool wear classification outperforms the method in the previous approach.

Table 4
Comparison of accuracy between the proposed method and the method(7) in the previous research.

	Proposed method (%)	Previous method (%)
Test data	99.32	80.40
Out-of-sample data	84.44	51.11

4. Conclusions

An offline tool wear detection method based on audio signals was proposed in this paper. The preprocessed audio signals were used to train a CNN model. The results showed that the CNN model achieved an accuracy of 84.44% for the audio data in the out-of-sample test. Furthermore, we compared our method with an approach using the DWT and SVM, which achieved an accuracy of 51.11% using the same out-of-sample datasets. The method does not need to install sensors and collect data on the cutting machines, making it simple, especially for non-professionals seeking an easy-to-use solution to determine the tool wear approximately.

Acknowledgments

The authors thank the National Science and Technology Council, Taiwan, for financially supporting this research under Grant no. MOST 110-2221-E-011-111.

References

1 ISO 8688-2: 1989 (2016) Int Stand, Int Orga Stand. https://www.iso.org/standard/16092.html
2 M.-K. Liu, Y.-H. Tseng, and M.-Q. Tran: Int. J. Adv. Manuf. Technol. **103** (2019) 3361. https://doi.org/10.1007/s00170-019-03686-2
3 L. Móricz, Z. J. Viharos, A. Németh, A. Szépligeti, and M. Büki: Measurement **163** (2020) 108025. https://doi.org/10.1016/j.measurement.2020.108025
4 T. Benkedjouh, N. Zerhouni, and S. Rechak: Int. J. Adv. Manuf. Technol. **97** (2018) 3311. https://doi.org/10.1007/s00170-018-2018-6
5 H. Zheng and J. Lin: Proc. 2019 3rd Int. Conf. Robotics and Automation Sciences (IEEE, 2019) 63–68. https://doi.org/10.1109/ICRAS.2019.8809070
6 S. Bagri, A. Manwar, A. Varghese, S. Mujumdar, and S. S. Joshi: J. Manuf. Processes **71** (2021) 679. https://doi.org/10.1016/j.jmapro.2021.09.055
7 C. K. Madhusudana, H. Kumar, and S. Narendranath: Int. J. Syst. Assur. Eng. Manage. **8** (2017) 1643. https://doi.org/10.1007/s13198-017-0637-1
8 Z. Li, X. Liu, A. Incecik, M. K. Gupta, G. M. Królczyk, and P. Gardoni: J. Manuf. Processes **79** (2022) 233. https://doi.org/10.1016/j.jmapro.2022.04.066
9 A. M. Alzahrani, R. Liu, and J. R. Kolodziej: Proc. Annu. Conf. PHM Society (2018) 509. https://doi.org/10.36001/phmconf.2018.v10i1.509
10 A. Kothuru, S. P. Nooka, and R. Liu: J. Manuf. Sci. Eng. **140** (2018) 111006. https://doi.org/10.1115/1.4040874
11 Z. Li, R. Liu, and D. Wu: J. Manuf. Processes **48** (2019) 66. https://doi.org/10.1016/j.jmapro.2019.10.020
12 F. J. Alonso and D. R. Salgado: Proc. Inst. Mech. Eng., Part B: J. Eng. Manuf. **219** (2005) 703. https://doi.org/10.1243/095440505X32634
13 K. Palanisamy, D. Singhania, and A. Yao: arXiv preprint arXiv: 2007.11154 (2020). https://doi.org/10.48550/arXiv.2007.11154
14 S. Hershey, S. Chaudhuri, D. P. W. Ellis, J. F. Gemmeke, A. Jansen, R. C. Moore, M. Plakal, D. Platt, R. A. Saurous, B. Seybold, M. Slaney, R. J. Weiss, and K. Wilson: Proc. 2017 IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP 2017) 131–135. https://doi.org/10.1109/ICASSP.2017.7952132

15 F. Demir, M. Turkoglu, M. Aslan, and A. Sengur: Appl. Acoust. **170** (2020) 107520. https://doi.org/10.1016/j.apacoust.2020.107520
16 M. Scarpiniti, R. Parisi, and Y.-C. Lee: Appl. Sci. **14** (2024) 90. https://doi.org/10.3390/app14010090
17 Z. Chen and C. Liu: Proc. 2021 4th Int. Conf. Robotics, Control and Automation Engineering, (2021) 156–160. https://doi.org/10.1109/RCAE53607.2021.9638811.
18 J. Sun, Y. Wang, and X. Chen: 2019 IEEE/ACIS 18th Int. Conf. Computer and Information Science (2019) 1–6. https://doi.org/10.1109/ICIS46139.2019.8940257
19 P. Li, T. Song, and J. Hu: Proc. SPIE 12719, 2nd Int. Conf. Electronic Information Technology (2023) 127194M. https://doi-org.ntust.idm.oclc.org/10.1117/12.2685823
20 Y. Wang, Z. Chu, I. Ku, E. C. Smith, and E. T. Matson: 2022 6th IEEE Int. Conf. Robotic Computing (IEEE, 2022) 186. https://doi.org/10.1109/IRC55401.2022.00039