Machine Learning-Based Detection of Cutting Tool End-of-Life Using Accelerometer Data

Žiga Kolar
ziga.kolar@ijs.si
Jožef Stefan Institute
Ljubljana, Slovenia

Alexandre Baran
alexandre.baran.pro@gmail.com
University Paris-Saclay
Paris, France

Jože Ravničan
joze.ravnican@unior.com
UNIOR, d.o.o.
Zreče, Slovenia

Matjaž Gams
matjaz.gams@ijs.si
Jožef Stefan Institute
Ljubljana, Slovenia

Abstract

Cutting tool wear is an important factor affecting machining quality, process reliability, and production efficiency. This paper presents a machine learning-based approach for detecting cutting tool end-of-life using accelerometer data collected during machining. The objective is to identify vibration patterns associated with late-stage tool degradation and to provide a wear-confidence indicator that can support tool replacement decisions. Tool-change timestamps recorded by the operator were used to define healthy and worn signal regions within each tool-life cycle. Several labeling strategies were investigated, including a simple 50/50 division and an Early/Late 30% strategy, where the ambiguous middle portion of each cycle was excluded. Time-domain and frequency-domain features were extracted from the vibration signals and used to train multiple classical machine learning classifiers. Model performance was assessed using cycle-based validation, with F1 score used for model selection and alert precision used to assess the reliability of high-confidence warnings. The best performance for each tool was achieved with different machine learning algorithms with the Early/Late 30% labeling strategy. The results show that accelerometer-based vibration features can provide useful information for detecting late-cycle tool wear. However, the proposed method is intended as a decision-support system rather than a fully automated tool replacement mechanism. Future validation in an industrial setting, with operator confirmation of true and false predictions, is required to assess practical reliability under real production conditions.

Keywords

signal processing, accelerometer, sensors, machine learning

1 Introduction

Cutting tool wear is one of the critical factors affecting the quality, productivity, and economic efficiency of machining processes. As a cutting tool gradually wears during operation, its ability to remove material effectively decreases, which can lead to poor surface finish, dimensional inaccuracies, increased cutting forces, excessive vibration, higher energy consumption, and potential damage to both the workpiece and the machine tool. Therefore, reliable detection of the point at which a cutting tool reaches the end of its useful life is essential for maintaining stable machining performance and avoiding unexpected process failures.

In conventional machining practice, tool wear is often assessed through manual inspection, fixed tool replacement intervals, or indirect judgement based on operator experience. Although these approaches are simple to implement, they have significant limitations. Manual inspection typically requires process interruption and is not suitable for continuous production, while fixed replacement intervals may result in either premature tool replacement or continued operation with an already worn tool. Premature replacement increases tooling costs, whereas delayed replacement can reduce product quality and lead to machine downtime. These limitations have motivated the development of automated tool condition monitoring systems capable of detecting tool degradation during machining.

Among the different sensing methods used for tool condition monitoring, accelerometer-based measurements are particularly suitable because tool wear is strongly reflected in the vibration behavior of the machining system. As the cutting edge deteriorates, the interaction between the tool and workpiece changes, producing variations in vibration amplitude, frequency content, and signal stability. Accelerometers provide a practical, non-invasive, and relatively low-cost means of capturing these dynamic changes. However, accelerometer signals are often complex, noisy, and affected by multiple process variables, making direct interpretation difficult using simple threshold-based methods.

Machine learning offers an effective approach for extracting useful information from accelerometer data and identifying patterns associated with different tool conditions. By learning from measured signal features or raw sensor data, machine learning models can classify tool states, detect abnormal vibration patterns, and estimate the point at which the cutting tool becomes worn out. Such data-driven methods are especially useful in machining environments where the relationship between sensor signals and tool wear is nonlinear and influenced by factors such as cutting speed, feed rate, depth of cut, tool geometry, and workpiece material.

This paper presents a machine learning-based approach for detecting cutting tool end-of-life using accelerometer data. Accelerometer signals collected during machining are processed and used to train models capable of identifying the transition from acceptable tool condition to worn-out tool condition. The objective is to develop a reliable method for detecting when the cutting tool should be replaced, thereby supporting predictive maintenance, reducing unnecessary tool changes, and preventing quality deterioration caused by excessive tool wear.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
Information Society 2026, Ljubljana, Slovenia
© 2026 Copyright held by the owner/author(s).
https://doi.org/10.70314/is.2026.skui.5

The main contribution of this work is the application of machine learning to accelerometer-based tool wear detection with a focus on identifying the end-of-life condition of the cutting tool. The proposed approach demonstrates how vibration data can be transformed into meaningful indicators of tool condition and used for automated decision-making in machining processes. By enabling timely detection of worn-out tools, the method contributes to improved process reliability, better product quality, and more efficient use of cutting tools.

This paper is structured as follows: Section 2 reviews prior work and related research in tool wear detection and condition monitoring. Section 3 presents the proposed approach and the machine learning methods used for cutting tool wear detection. Section 4 reports and analyzes the experimental results. Finally, Section 5 concludes the paper and outlines directions for future work.

2 Related Work

Tool condition monitoring has been widely investigated to improve machining reliability, reduce downtime, and optimize tool replacement. Both signal-processing and artificial-intelligence methods have been applied using vibration, cutting force, acoustic emission, and other process signals.

Several studies demonstrate the usefulness of vibration for indirect tool-wear monitoring. Prasad and Babu reported a strong relationship between vibration displacement and flank wear in turning [11], while Dimla showed that time- and frequency-domain acceleration features, particularly spectral characteristics, are sensitive to tool condition [5]. Abu-Mahfouz used vibration features with neural networks to classify drill wear, with frequency-domain features providing particularly useful information [1]. Alonso and Salgado similarly found that medium- and high-frequency vibration components contain information related to flank wear [2]. More recently, Xue et al. combined vibration sensing, signal decomposition, dimensionality reduction, and weighted random forests for tool-wear recognition in full life-cycle milling experiments [15]. Esmaili et al. also showed that vibration energy, particularly in the 0–1 kHz range, can provide useful indicators of machining condition and degradation [6].

Machine-learning approaches have also been applied to other sensing modalities and more advanced monitoring architectures. Marani et al. predicted flank wear from cutting-force measurements using an adaptive neuro-fuzzy inference system [8], while Chehrehzad et al. combined time- and frequency-domain features with recurrent neural networks for real-time drilling-tool wear prediction on an industrial edge device [3]. For aluminum machining data, Hidden Markov Models have also been successfully applied to tri-axial accelerometer measurements [14].

Recent reviews confirm the increasing use of AI for tool condition monitoring across turning, milling, drilling, and grinding [10, 4, 9]. They emphasize the importance of sensor selection, preprocessing, feature extraction, and model design, while also identifying limitations such as the lack of common benchmark datasets. Hybrid approaches combining data-driven models with physical knowledge, expert knowledge, transfer learning, or heterogeneous architectures have additionally been proposed to improve generalizability and industrial reliability [12]. Huang et al. transformed milling vibration signals into time–frequency representations using STFT and applied a deep convolutional neural network for tool-wear prediction [7].

More recent and complementary studies further demonstrate the use of vibration signals for tool-condition and tool-life monitoring. Sharma and Min analyzed time- and frequency-domain vibration features for near-real-time tool-wear monitoring using an accelerometer mounted on the tool holder [13].

Overall, previous work shows that vibration signals combined with suitable feature extraction and machine-learning methods provide a practical basis for tool-wear monitoring. The present study builds on this work by investigating accelerometer-based detection of late-stage tool wear and tool end-of-life.

3 Methodology

Accelerometer data was recorded during machining for three cutting tools (T1, T3, and T5). The accelerometer was mounted on the cutting machine and sampled at 5000 Hz using tri-axial acceleration measurements. The continuous recordings were divided into signal windows of 9 seconds for tool T1, 6 seconds for tool T3 and 6 seconds for tool T5. The dataset contained 14, 14, and 7 complete tool-life cycles for tools T1, T3, and T5, respectively. Each cycle extended from the beginning of use of a tool to the operator-recorded tool-change event.

A reference signal was provided in which the duration of each tool was known in advance. This reference signal was used to identify and analyze similar signals. A template-matching method was used to locate regions in the input spectrogram that were similar to a smaller reference spectrogram. The method slides the template across the input spectrogram and computes a similarity score at each position using the normalized cross-correlation coefficient. This metric measures how well the intensity pattern of the template matches the underlying spectrogram region, while normalizing for brightness and contrast differences. Higher values (closer to 1) indicate stronger matches, while lower values indicate poor similarity. This method is commonly used for object detection, feature localization, and pattern recognition when the object appearance is consistent and not heavily distorted.

The main objective was to detect changes in vibration patterns associated with cutting tool replacement events. For each machining cycle, the operator recorded a timestamp indicating when the tool was changed, which was stored in a log of tool-change events. This timestamp information defined the labeling strategy used to construct the training dataset.

Several labeling strategies were investigated. In the initial approach, each tool-life cycle was divided into two equal halves: the first 50% was labeled as healthy and the last 50% as worn. However, this strategy encompasses the intermediate portion of the cycle, where tool condition is inherently ambiguous. To mitigate this ambiguity, an Early/Late 30% labeling strategy was also examined. Under this strategy, the first 30% of each cycle is labeled as healthy, the intermediate 40% is discarded, and the final 30% is labeled as worn, thereby allowing the model to learn from unambiguous examples of early- and late-cycle behavior.

After the labels were selected, each signal was transformed into a set of numerical features. For each signal segment, a set of time-domain features was first calculated, including the mean, standard deviation, root mean square (RMS), maximum and minimum values, peak-to-peak amplitude, skewness, kurtosis, and crest factor. These features describe the overall vibration level, signal variability, impulsiveness, and asymmetry of the measured acceleration signal. In the context of cutting tool wear detection, such features are useful because tool degradation can lead to

Table 1: Comparison of the 50/50 and Early/Late 30% labeling strategies using cycle-based validation and a fixed alert threshold of 0.80.

		50/50 labeling			Early/Late 30% labeling		
Method	**Tool**	**Precision**	**Recall**	**F1**	**Precision**	**Recall**	**F1**
Gradient Boosting	T1	0.47	0.70	0.56	0.50	0.91	0.65
Gradient Boosting	T3	0.54	0.73	0.62	0.56	0.92	0.70
Gradient Boosting	T5	0.64	0.74	0.69	0.74	0.94	0.83
Extra Trees	T1	0.52	0.69	0.59	0.60	0.85	0.70
Extra Trees	T3	0.59	0.77	0.67	0.62	0.99	0.76
Extra Trees	T5	0.66	0.79	0.72	0.76	0.86	0.81
Random Forest	T1	0.55	0.75	0.63	0.64	0.85	0.73
Random Forest	T3	0.52	0.80	0.63	0.59	0.92	0.72
Random Forest	T5	0.60	0.78	0.68	0.69	0.86	0.77

increased vibration amplitude, stronger peaks, and changes in the statistical distribution of the signal.

In addition to time-domain features, frequency-domain features were extracted using Welch's method to estimate the power spectral density (PSD) of each signal segment. The PSD was used to compute spectral statistics, including spectral mean, spectral standard deviation, spectral skewness, and spectral kurtosis, which characterize how vibration energy is distributed across frequencies. The mean frequency and maximum PSD value were also included to capture dominant frequency behavior and the strongest vibration components. Finally, the spectrum was divided into five frequency bands, and the energy in each band was calculated. These band-energy features provide information about how the vibration energy changes across different frequency ranges, which is important because tool wear may not only increase the overall vibration level but also shift energy toward specific frequency bands.

Furthermore, a limit was imposed on the number of training samples extracted from each cycle, set to 200 samples per class. For instance, with this constraint, a maximum of 200 healthy samples and 200 worn samples are used from each cycle. This ensures that longer machining cycles do not disproportionately influence the training dataset and helps maintain a balanced representation across cycles.

4 Results

After the data was prepared, a cycle-based validation strategy was applied to the entire dataset to ensure robust model evaluation. This strategy selects one cycle as the test set and uses the remaining cycles for training. The procedure is repeated until each cycle has served as the test set once. This evaluation approach is particularly suitable for the present dataset, since it tests whether the learned patterns generalize across complete machining cycles rather than only across randomly selected samples from the same cycle.

Several classical machine learning models were evaluated in order to identify a suitable wear-confidence indicator for the monitored cutting tools. The models assessed included Random Forest, Extra Trees and Gradient Boosting classifiers. These models were selected because they are well suited for tabular feature-based classification tasks and can capture nonlinear relationships between the extracted signal features and the corresponding tool-wear condition.

The objective of this comparison was not solely to maximize classification accuracy, but also to identify models that provide a useful balance between detecting worn conditions and limiting false alerts. F1 score was therefore used as the primary metric for selecting the classifier for each tool because it jointly accounts for precision and recall. Alert precision was additionally considered an important operational metric, since false high-confidence warnings may lead to unnecessary tool replacement. An alert was generated when the predicted probability of the worn class was at least 0.80. The same fixed threshold of 0.80 was used for all models, tools, and labeling-strategy comparisons reported below.

According to the F1 criterion, the optimal classifier depended on the monitored tool. Random Forest achieved the highest F1 score for T1, Extra Trees for T3, and Gradient Boosting for T5. Alert precision was examined alongside F1 to assess the reliability of the resulting high-confidence warnings.

To evaluate the effect of the labeling strategy directly, the complete cycle-based validation procedure was repeated using both the 50/50 and Early/Late 30% labelling. The same feature extraction procedure, classifiers, validation folds, and decision threshold were used in both experiments. Table 1 compares precision, recall, and F1 score for the two labeling strategies. The Early/Late 30% strategy produced higher performance, with the largest improvements observed for tool T5.

The classification results show that the optimal model depends on the monitored tool. The highest F1 score for tool T1 was achieved using the Random Forest model, whereas the highest F1 scores for tools T3 and T5 were achieved using the Extra Trees and Gradient Boosting models, respectively.

For tool T1, the Random Forest classifier provided the best overall balance between precision and recall, reaching an F1 score of 0.73. Although Gradient Boosting achieved a slightly higher recall for this tool, its lower precision indicates a greater number of false alerts. Random Forest was therefore more suitable for T1, since it maintained high recall while also improving the reliability of the generated worn-tool alerts.

For tool T3, the Extra Trees classifier achieved the highest F1 score of 0.76. This result was mainly driven by its very high recall of 0.99, indicating that the model detected nearly all samples labeled as worn. Its precision was also higher than that of Gradient Boosting and Random Forest for the same tool, making it the most reliable choice among the evaluated models for T3.

For tool T5, Gradient Boosting produced the highest F1 score of 0.83, with a precision of 0.74 and a recall of 0.94. This was the strongest overall result among all tested tool-model combinations. Extra Trees achieved a slightly higher precision for T5, but its lower recall resulted in a marginally lower F1 score. Gradient Boosting therefore provided the best compromise between detecting worn samples and limiting false alerts for this tool.

Overall, the results indicate that ensemble-based classifiers are suitable for estimating tool-wear condition from the extracted

features. Recall values were consistently high across all tools and methods, showing that the models were generally effective at identifying worn-tool samples. However, the variation in precision demonstrates that the reliability of high-confidence alerts depends strongly on both the selected model and the specific tool being monitored. This confirms the need for tool-specific model selection rather than adopting a single classifier for all tools.

Based on the F1 scores and the alert-precision-oriented evaluation strategy, Random Forest was selected as the preferred model for T1, Extra Trees for T3, and Gradient Boosting for T5. These models provide the most suitable trade-off between detecting worn conditions and reducing false positive alerts under the applied cycle-based validation procedure.

These results should be interpreted as a preliminary wear-confidence indicator rather than as a basis for direct, automated tool-change decisions. The healthy and worn labels were derived from the relative position of each detected signal within a tool-life cycle, rather than from direct physical wear measurements. Consequently, the model output should be understood as the probability that a given signal belongs to the late-cycle, worn-labeled class. In a practical monitoring application, this information should be presented to the machine operator as a decision-support warning, while the final tool replacement decision should remain subject to operator verification or additional validation procedures.

5 Conclusion

This paper presented a machine learning-based approach for detecting cutting tool end-of-life using accelerometer data collected during machining. The main objective was to determine whether vibration-based features can provide a useful indication of late-stage tool wear and support tool replacement decisions. Tool-change timestamps recorded by the operator were used to define tool-life cycles, and different labeling strategies were considered to distinguish between healthy and worn tool states.

The results show that accelerometer signals contain relevant information for detecting changes associated with tool degradation. The Early/Late 30% labeling strategy was found to be suitable because it removes the ambiguous middle portion of each tool-life cycle and allows the models to learn from clearer examples of healthy and worn behavior. Time-domain and frequency-domain features extracted from the vibration signals enabled classical ensemble-based classifiers to identify worn-labeled samples with high recall across the evaluated tools.

The best-performing model depended on the specific tool being monitored. Random Forest achieved the highest F1 score for tool T1, Extra Trees achieved the best result for tool T3, and Gradient Boosting achieved the best result for tool T5. Classifier selection was based on F1 score, while alert precision was treated as an operational measure of the reliability of high-confidence worn-tool warnings. This indicates that a single model may not be optimal for all cutting tools and that tool-specific model selection can improve monitoring performance. The results also show that the proposed method is more appropriate as a wear-confidence indicator than as a direct automatic replacement mechanism, since the labels were derived from cycle position rather than from direct physical measurements of tool wear.

Overall, the study demonstrates that machine learning applied to accelerometer-based vibration features can support the detection of cutting tool end-of-life. The proposed approach can assist operators by providing high-confidence warnings when a tool is likely to be in a late wear stage, helping to reduce unnecessary tool changes while also avoiding continued machining with worn tools.

Future work should focus on validating the method in a real industrial environment. In particular, model predictions should be compared with operator feedback, physical inspection of tool wear, and quality measurements of the machined parts. Additional data from more tools, materials, cutting conditions, and production cycles would also help improve model generalization. Further improvements could include adaptive thresholds and integration with other sensor modalities to increase the reliability of the monitoring system under practical machining conditions.

Acknowledgements

The authors acknowledge funding support from the company UNIOR for the GREMO LIGHTWEIGHT project. The authors also acknowledge the funding from the Slovenian Research and Innovation Agency (ARIS), Grant (PR-10495) and Basic core funding P2-0209.

References

[1] Issam Abu-Mahfouz. 2003. Drilling wear detection and classification using vibration signals and artificial neural network. *International Journal of Machine Tools and Manufacture*, 43, 7, 707–720.

[2] FJ Alonso and DR Salgado. 2008. Analysis of the structure of vibration signals for tool wear detection. *Mechanical systems and signal processing*, 22, 3, 735–748.

[3] Mohammadreza Chehrehzad, Gamze Kecibas, Cemile Besirova, Ugur Uresin, Mumin Irican, and Ismail Lazoglu. 2024. Tool wear prediction through ai-assisted digital shadow using industrial edge device. *Journal of Manufacturing Processes*, 113, 117–130.

[4] Lorenzo Colantonio, Lucas Equeter, Pierre Dehombreux, and Francois Ducobu. 2021. A systematic literature review of cutting tool wear monitoring in turning by using artificial intelligence techniques. *Machines*, 9, 12, 351.

[5] DE Dimla Snr. 2002. The correlation of vibration signal features to cutting tool wear in a metal turning operation. *The International Journal of Advanced Manufacturing Technology*, 19, 10, 705–713.

[6] Parisa Esmaili and Loredana Cristaldi. 2024. Health indicator analysis in terms of condition monitoring on brownfield cnc milling machines using triaxial accelerometer. *IEEE Sensors Letters*, 8, 7, 1–4.

[7] Zhiwen Huang, Jianmin Zhu, Jingtao Lei, Xiaoru Li, and Fengqing Tian. 2021. Tool wear monitoring with vibration signals based on short-time fourier transform and deep convolutional neural network in milling. *Mathematical Problems in Engineering*, 2021, 1, 9976939.

[8] Mohsen Marani, Mohammadjavad Zeinali, Jules Kouam, Victor Songmene, and Chris K Mechefske. 2020. Prediction of cutting tool wear during a turning process using artificial intelligence techniques. *The International Journal of Advanced Manufacturing Technology*, 111, 1, 505–515.

[9] Roberto Munaro, Aldo Attanasio, and Antonio Del Prete. 2023. Tool wear monitoring with artificial intelligence methods: a review. *Journal of Manufacturing and Materials Processing*, 7, 4, 129.

[10] Danil Yu Pimenov, Andres Bustillo, Szymon Wojciechowski, Vishal S Sharma, Munish K Gupta, and Mustafa Kuntoğlu. 2023. Artificial intelligence systems for tool condition monitoring in machining: analysis and critical review. *Journal of Intelligent Manufacturing*, 34, 5, 2079–2121.

[11] Balla Srinivasa Prasad and M Prakash Babu. 2017. Correlation between vibration amplitude and tool wear in turning: numerical and experimental analysis. *Engineering Science and Technology, an International Journal*, 20, 1, 197–211.

[12] Busra Tan Saatci, Mustafa Ulas, and Turan Gurgenc. 2026. Hybrid ai systems for tool wear monitoring in manufacturing: a systematic review. *Applied Sciences*, 16, 1, 208.

[13] Deept Sharma and Sangkee Min. 2026. Vibration-based feature analysis and intelligent tool wear prediction for cnc machines: an industry 4.0 approach. *International Journal of Precision Engineering and Manufacturing-Smart Technology*, 4, 2, 57–66.

[14] Antonio G Vallejo Jr, Juan A Nolazco-Flores, Rubén Morales-Menéndez, L Enrique Sucar, and Ciro A Rodríguez. 2005. Tool-wear monitoring based on continuous hidden markov models. In *Iberoamerican Congress on Pattern Recognition*. Springer, 880–890.

[15] Zhongling Xue, Liang Li, Youling Wu, Yinfei Yang, Wentao Wu, Yuhang Zou, and Ni Chen. 2023. Study on tool wear state recognition algorithm based on spindle vibration signals collected by homemade tool condition monitoring ring. *Measurement*, 223, 113787.