Sensors and Materials, Vol. 38, No. 6 (2026) 3089–3124
MYU Tokyo

S & M 4491

Intelligent Determination of Optimal Grinding Wheel Dressing Timing Based on Acoustic Emission Signals

Yin-Wei Chao, Chun-Yen Chen, Yue-Feng Lin, Ming-Yi Tsai, and Kai-Jung Chen*

Department of Mechanical Engineering, National Chin Yi University of Technology,
Taichung 41170, Taiwan, R.O.C.

(Received March 4, 2026; accepted May 27, 2026)

Keywords: acoustic emission, abrasive grain exposure ratio, clustering analysis, dressing timing, ResNet

The grinding wheel condition critically affects surface integrity and process stability in precision grinding. During continuous grinding, abrasive grains exhibit multiple coexisting physical phenomena, including chip removal, self-sharpening, stuffing, and passivation, which cannot be adequately represented by conventional single-state monitoring approaches. In this study, we present an acoustic emission (AE)-based multi-label deep learning framework for intelligent grinding wheel condition recognition and dressing timing determination. AE signals collected during continuous surface grinding were transformed into time–frequency representations and labeled through the direct optical observation of abrasive grain conditions. A Residual Neural Network (ResNet)-based multi-label classifier was trained to simultaneously identify the four grinding phenomena. The proposed model achieved a macro-averaged F1-score of approximately 0.80, with area-under-curve values ranging from 0.73 to 0.81. Consecutive prediction behavior was further analyzed under continuous grinding conditions, showing that persistent AE-predicted degradation phenomena correspond to irreversible surface roughness deterioration beyond Ra = 0.4 μm. On the basis of these findings, an experimentally validated dressing timing criterion was established using consecutive multi-label predictions and surface roughness verification. The proposed approach provides a physically interpretable and data-driven solution for intelligent grinding process monitoring and dressing decision-making.

1. Introduction

Grinding is widely employed in manufacturing processes requiring high dimensional accuracy and superior surface quality, such as mold fabrication, precision machinery, and critical component production. As the primary cutting tool in grinding operations, the condition of the grinding wheel plays a crucial role in determining machining performance and process stability. During prolonged grinding, the abrasive grains of the wheel gradually undergo phenomena such as passivation, stuffing, and self-sharpening, which alter the cutting mechanism and adversely affect surface integrity. If dressing is not performed at an appropriate time, surface roughness deterioration, grinding burn, and process instability may occur. Conversely,

*Corresponding author: e-mail: hskchen5@ncut.edu.tw
https://doi.org/10.18494/SAM6311

ISSN 0914-4935
https://myukk.org/

excessive dressing results in the unnecessary material loss of both the grinding wheel and the workpiece. Therefore, establishing an effective and objective approach for determining the optimal dressing timing of grinding wheels remains a critical and challenging issue in grinding process control.

In recent years, advances in sensing technologies and intelligent manufacturing have motivated extensive research on grinding process monitoring to improve machining stability and surface quality. Among various sensing approaches, acoustic emission (AE) signals have been widely adopted for grinding condition recognition and grinding wheel deterioration monitoring owing to their high frequency bandwidth and strong sensitivity to material removal mechanisms and microscopic fracture events.[1–4] Previous studies have shown that AE signals can characterize abrasive–workpiece interactions and support the detection of wheel wear, passivation, stuffing, and abnormal conditions such as grinding burn.[5–7] In particular, Li *et al.* reported in *Sensors and Materials* that the time–frequency analysis of AE signals can effectively characterize the grinding status of alumina grinding wheels and is strongly correlated with surface roughness variations.[4]

With the rapid development of AI and machine learning, researchers have increasingly integrated time–frequency/frequency-domain representations with data-driven classifiers to automate grinding and dressing state identification, reducing reliance on subjective operator judgment and improving repeatability.[8–11] Recent studies have demonstrated that sensor-based monitoring systems enable the real-time observation of process conditions, tool wear, and surface integrity, thereby improving machining performance and reducing defects.[12] In particular, AE signals have been shown to be highly sensitive to variations in machining conditions and tool states, making them effective indicators for real-time monitoring and adaptive control.[13] Furthermore, the integration of AE sensing with machine learning techniques enables the extraction of complex and nonlinear signal patterns associated with tool wear and process degradation.[14] In addition, recent studies published in *Sensors and Materials* have demonstrated that AE-based sensing and sensor signal analysis can be effectively applied to fault diagnosis, wear condition monitoring, and machine tool state recognition, highlighting the growing importance of sensor-driven approaches in intelligent manufacturing applications.[15–17]

Nevertheless, most existing studies focus primarily on state recognition or wear assessment, while the direct translation of sensing results into actionable process control strategies remains limited. Therefore, there is a need to establish a physically interpretable and experimentally validated framework that bridges sensor-based signal analysis and practical decision-making, particularly for the optimal dressing timing of grinding wheels.

To address the aforementioned research gap, in this study, we aim to establish an intelligent method for determining the optimal dressing timing of grinding wheels based on AE signals, integrating sensing analysis with deep learning to develop practical decision-making criteria for shop-floor grinding operations. During the grinding process, AE signals are acquired and transformed into time–frequency representations to preserve spectral characteristics associated with different grinding states. A ResNet-RS270 deep learning model is then employed to classify the grinding conditions into four categories: stuffing, chip removal, self-sharpening, and

passivation. Furthermore, by analyzing consecutive signal patterns and validating the results with measured surface roughness, a concrete and operable criterion for optimal dressing timing is proposed. Unlike previous studies that primarily focus on state recognition or wear monitoring, the proposed approach translates sensing outcomes into actionable dressing decisions, thereby enhancing process stability and economic efficiency in grinding operations.

The remainder of this paper is organized as follows. In Sect. 2, we describe the experimental setup, grinding conditions, AE signal acquisition procedure, and data preprocessing and deep learning model architecture. In Sect. 3, we present the four-class grinding state classification results and model performance evaluation, followed by the validation of the proposed dressing timing criterion through consecutive signal pattern analysis and surface roughness measurements. Finally, in Sect. 4, we summarize the main findings of this study and discuss its potential applications in intelligent grinding monitoring systems and future research directions.

2. Materials and Methods

The overall experimental and analytical framework of this study is illustrated in Fig. 1. The objective was to establish a method of determining the optimal dressing timing of a grinding wheel based on AE signals. The proposed methodology consisted of three main stages, namely,

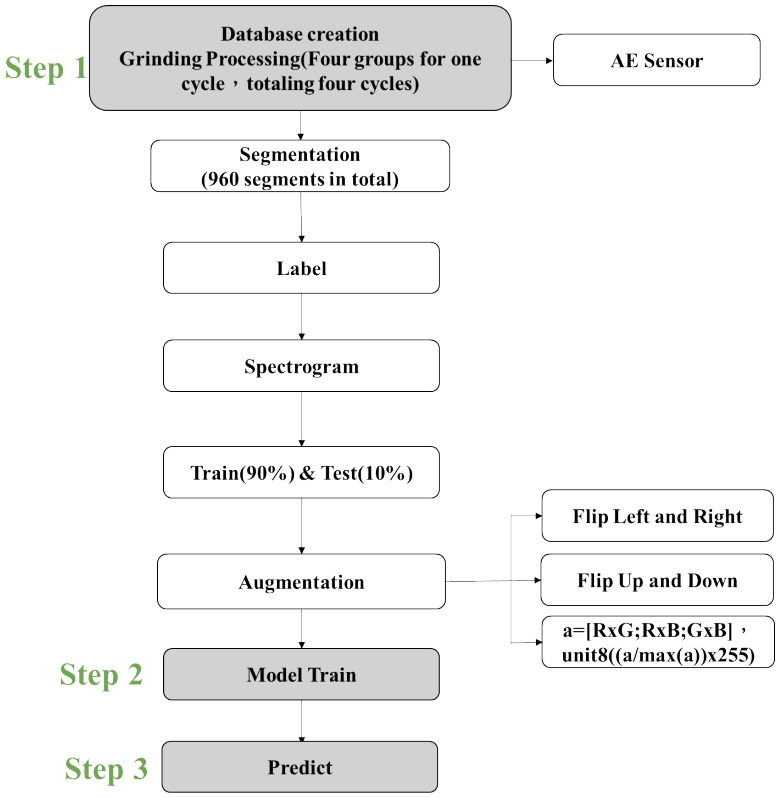


Fig. 1. (Color online) Overall experimental and analytical framework of this study.

database construction, model training, and validation, and was designed around the physical evolution of abrasive grain conditions during continuous grinding.

In the database construction stage, continuous grinding experiments were conducted under fixed machining parameters without intermediate dressing, allowing the grinding wheel to naturally undergo transitions among chip removal, self-sharpening, stuffing, and passivation states. Each grinding cycle corresponded to a cumulative material removal of 0.6 mm, resulting in a total grinding depth of 2.4 mm, thereby simulating wheel performance degradation during practical grinding operations. During grinding, AE signals generated by the interaction between the grinding wheel and the workpiece were continuously acquired as the primary sensing data.

In the training dataset construction stage, the acquired AE signals were segmented to extract effective grinding-related components, and each segment was treated as an independent sample. The signals were then transformed into frequency-domain and time–frequency representations. To ensure that the training dataset was physically meaningful, manual labeling was performed on the basis of direct observations of the actual abrasive grain conditions on the grinding wheel surface. High-depth optical microscopy and replica-based imaging techniques were employed to capture abrasive grain morphology. On the basis of whether the abrasive grains were fully exposed, exhibited self-sharpening fracture, were blocked by adhered chips, or showed passivated cutting edges, the corresponding AE samples were manually labeled into four grinding wheel states: chip removal, self-sharpening, stuffing, and passivation.

In the final prediction and validation stage, the labeled dataset was used to train a deep learning model for grinding wheel state classification, and model performance was evaluated using standard classification metrics. To further verify the physical consistency of the classification results, surface roughness measurements of the ground workpieces and replica-based observations of abrasive grain exposure were conducted. These physical measurements were used to assess whether the predicted grinding wheel states corresponded to actual abrasive behavior and surface quality evolution. Through this integrated workflow combining experimental observation, signal analysis, and physical validation, a practical framework for determining grinding wheel dressing timing was established.

2.1 Grinding system and materials

Grinding experiments were performed on a high-speed precision CNC surface grinding machine (ESG-1020CNC, Equiptop Hitech Corp.), as shown in Fig. 2. The machine is equipped with a Syntec MA11 numerical controller, providing stable and repeatable motion control during prolonged grinding operations. The main specifications of the grinding system relevant to machining stability and precision control are summarized in Table 1.

The grinding machine provides sufficient spindle power and stable rotational speed for continuous grinding experiments. To minimize mechanical interference during AE signal acquisition, the spindle vibration amplitude is maintained within 2 μm. In addition, an oil-circulation cooling system is implemented to ensure thermal stability under prolonged grinding conditions. The working table has a size of 250 × 500 mm^2, with maximum longitudinal and cross travels of 500 and 270 mm, respectively. The Z-axis is equipped with an optical scale

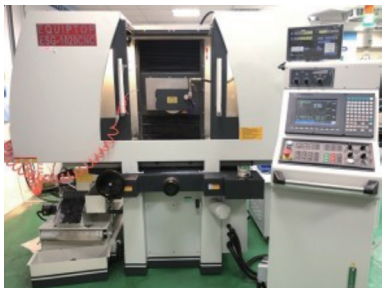


Fig. 2. (Color online) High-speed precision CNC surface grinding machine (ESG-1020CNC) used in grinding experiments.

Table 1
Machine specifications relevant to machining stability and precision control.

Item	Specification
Machine model	ESG-1020CNC (Equiptop Hitech Corp.)
Controller	Syntec MA11 CNC system
Spindle motor power	3 HP
Spindle rotational speed	3450 rpm (60 Hz); 2850 rpm (50 Hz)
Maximum allowable wheel speed	3105 rpm (33 m/s)
Spindle vibration amplitude	≤2 μm
Spindle cooling system	Oil-circulation cooling
Table size	250 × 500 mm^2
Maximum longitudinal travel	500 mm
Maximum cross travel	270 mm
Z-axis feedback	Optical scale closed-loop control
Table feed speed	5–25 m/min (60 Hz); 5–20 m/min (50 Hz)

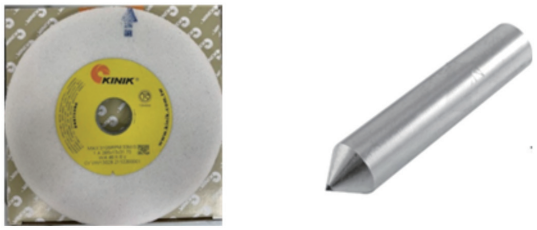
closed-loop feedback system, enabling precise and repeatable down-feed control throughout the grinding experiments.

A vitrified-bond white alumina grinding wheel (WA46) was employed as the cutting tool. The grinding wheel specifications were WA abrasive, grit size 46, grade K, and structure 8, with dimensions of φ205 × 13 × 31.75 mm^3 and a maximum allowable rotational speed of 3105 rpm (33 m/s), as summarized in Table 2. Prior to grinding, the wheel was dressed using a #10 single-point diamond dresser (DKB2E001B) to establish a consistent initial abrasive condition. The grinding wheel and dressing tool used in this study are shown in Fig. 3.

The workpiece material used in this study was medium-carbon steel S45C, in accordance with the Japanese Industrial Standards (JIS), as shown in Fig. 4. The chemical composition of S45C consists of 0.42–0.48 wt% carbon, 0.15–0.35 wt% silicon, and 0.60–0.90 wt% manganese, with maximum phosphorus and sulfur contents of 0.030 and 0.035 wt%, respectively, as listed in Table 3. S45C steel is widely used in mechanical components and tooling applications and is known to exhibit pronounced wheel loading, abrasive passivation, and self-sharpening behaviors during grinding, making it a suitable material for investigating grinding wheel condition evolution and AE-based condition monitoring.

Table 2
Specifications of the vitrified-bond white alumina grinding wheel (WA46).

Item	Specification
Abrasive	WA
Grit size	46
Grade	K
Structure	8
Bond type	Vitrified
Dimensions (mm)	φ205 × 13 × 31.75
Maximum speed	3105 rpm (33 m/s)



(a) (b)

Fig. 3. (Color online) (a) Vitrified-bond white alumina grinding wheel (WA46) and (b) #10 single-point diamond dresser (DKB2E001B) used in the experiments.

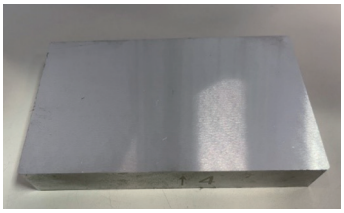


Fig. 4. (Color online) S45C medium-carbon steel workpiece used in grinding experiments.

Table 3
Chemical composition of S45C medium-carbon steel (JIS).

Steel grade	S45C
C (wt%)	0.42–0.48
Si (wt%)	0.15–0.35
Mn (wt%)	0.60–0.90
P (wt%)	≤0.030
S (wt%)	≤0.035

All grinding experiments were conducted under fixed grinding parameters to minimize process variability. The total material removal was set to 2.4 mm, with a down-feed of 0.01 mm per pass. The table feed speed was maintained at 10 m/min and the stroke length was fixed at 200 mm. The grinding wheel rotational speed was set to 2000 rpm, corresponding to a peripheral speed of 20.68 m/s. Surface roughness measurements were performed after every 0.6 mm of cumulative material removal to monitor the evolution of grinding wheel conditions during continuous grinding. The grinding interval of 1 mm was defined to ensure sufficient spatial

resolution for correlating surface roughness evolution with grinding wheel condition changes. The detailed grinding parameters are summarized in Table 4.

2.2 AE measurement and data acquisition

AE signals were employed to monitor grinding wheel conditions during the grinding process owing to their high sensitivity to microscopic deformation, grain fracture, frictional interactions, and material removal events occurring at the abrasive–workpiece interface. The applicability of AE sensing in grinding condition monitoring has been extensively validated in previous studies. Kwak and Ha[18] demonstrated that the root mean square (*RMS*) value of AE signals can be effectively used as input features for the neural-network-based diagnosis of grinding states such as normal cutting, chatter, and workpiece burn. Krishnan and Rameshkumar[19] further applied AE *RMS* features to construct a discrete hidden Markov model for grinding wheel condition prediction, successfully classifying wheel states into sharp, mildly worn, and fully worn categories. During wheel dressing, Dotto *et al.*[20] reported that AE signals synchronized with encoder feedback exhibit stable characteristics and can differentiate intact and damaged grinding wheels in specific frequency bands. More recently, Shen[21] demonstrated that fast-Fourier- transform-based frequency-domain features combined with appropriate window lengths provide robust indicators for grinding wheel wear monitoring. Wang *et al.*[22] also confirmed that AE signals are sensitive to material removal mechanisms in ultraprecision grinding, with distinct frequency-domain characteristics corresponding to different grinding depths. These findings collectively support the use of AE sensing as a reliable and physically meaningful approach for grinding wheel condition evaluation.

In this study, AE signals were acquired using a piezoelectric AE sensor (Type 8152C, Kistler Group), which provides a stable frequency response in the range of 50–400 kHz. The AE sensor and its frequency response characteristics are shown in Fig. 5.

The AE sensor was rigidly mounted on the grinding machine structure in close proximity to the grinding zone to ensure the efficient transmission of elastic stress waves generated during abrasive–workpiece interactions. A schematic of the experimental setup, including the grinding wheel, S45C workpiece, and AE sensor installation, is shown in Fig. 6.

The sensor output signals were transmitted to a data acquisition system consisting of a National Instruments CompactDAQ chassis (cDAQ-9174) equipped with a high-speed analog

Table 4
Grinding parameters used in the experiments.

Parameter	Value
Total material removal	2.4 mm
Down-feed per pass	0.01 mm
Table feed speed	10 m/min
Table stroke length	200 mm
Wheel rotational speed	2000 rpm
Wheel peripheral speed	20.68 m/s
Grinding interval	1 mm
Total machining duration	200 min

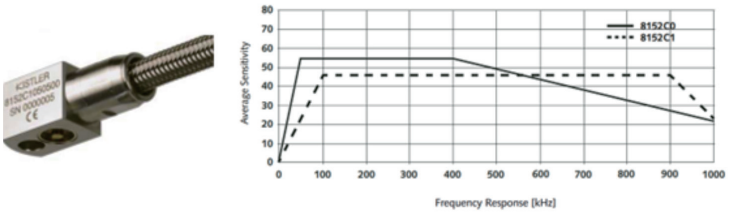


Fig. 5. (Color online) (a) Kistler-type 8152C AE sensor and (b) frequency response spectrum (C0).

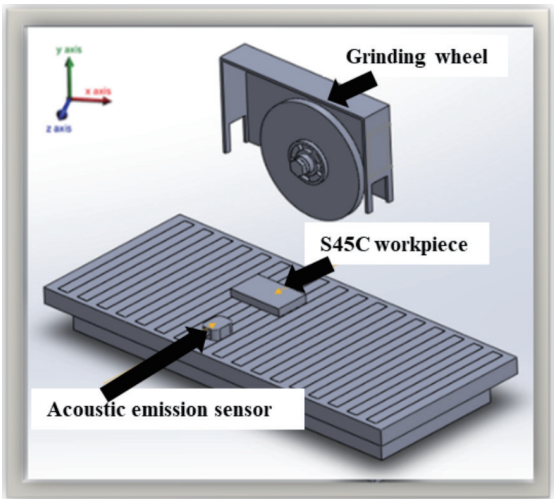


Fig. 6. (Color online) Experimental setup showing the grinding wheel, S45C workpiece, and AE sensor installation.

input module (NI-9223). The NI-9223 module supports simultaneous sampling with sufficient bandwidth for high-frequency AE signal acquisition. The key specifications of the data acquisition system are summarized in Table 5.

To determine an appropriate sampling frequency, preliminary experiments were conducted over a range of sampling rates between 50 and 400 kHz. Although higher sampling rates provided broader bandwidth coverage, excessive high-frequency noise contamination was observed near 400 kHz. Conversely, lower sampling rates led to insufficient spectral resolution for capturing characteristic AE components associated with abrasive cutting and grain fracture. Therefore, a root mean square error (*RMSE*)-based evaluation was performed to quantitatively assess signal stability and noise contribution under different sampling conditions. As illustrated in Fig. 7, the *RMSE* values stabilized and exhibited minimal variation near 200 kHz, indicating that dominant grinding-related AE features were sufficiently preserved while high-frequency noise was effectively suppressed. On the basis of this analysis, a sampling frequency of 200 kHz was selected for all subsequent AE measurements.

During grinding, continuous AE signals were recorded synchronously with the machining process. The raw AE signals were segmented into individual samples corresponding to effective grinding engagement intervals. Each signal segment was subsequently transformed into time–

Table 5
Specifications of the AE signal data acquisition system.

Item	Value
DAQ chassis	NI cDAQ-9174
Analog input module	NI-9223
Sampling mode	Simultaneous
Timing resolution	12.5 ns
Timing accuracy	50 ppm of sample rate
Input FIFO size	50 ppm of sample rate
Grinding interval	127 samples per slot

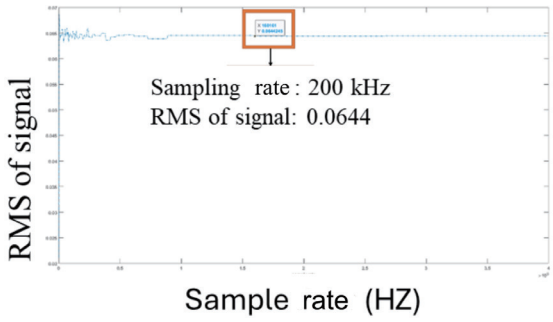


Fig. 7. (Color online) RMSE-based evaluation for determining the optimal AE sampling frequency.

frequency representations using short-time Fourier transform (STFT)-based spectral analysis, preserving both the temporal and spectral information of the AE signals. These time–frequency maps served as the primary input data for subsequent condition classification and model training.

To establish a supervised dataset for grinding wheel condition recognition, the segmented AE signals were manually annotated on the basis of the observed grinding behavior and direct inspection of grinding wheel surface conditions. Four representative grinding wheel states were defined: stuffing, chip removal, self-sharpening, and passivation. In total, 960 AE signal samples were extracted and labeled, forming the complete AE dataset used for subsequent deep-learning-based classification.

2.3 Surface characterization and abrasive exposure validation

To establish reliable ground truth labels for grinding wheel condition recognition, surface characterization was conducted by the direct optical observation of abrasive grain states on the grinding wheel surface after grinding. Unlike indirect inference methods based solely on signal features or replica-based techniques, in this study, we adopted the direct imaging of abrasive grains at fixed spatial locations on the grinding wheel as the final basis for manual annotation.

The surface morphology of abrasive grains was captured using a high-depth digital microscope (VHX-5000, Keyence), which provides sufficient depth of field and magnification to clearly resolve grain protrusion, chip adhesion, and cutting-edge conditions. When necessary, a

white-light interferometer (NewView™, ZYGO) was employed to support 3D surface morphology interpretation. Representative images of the observation instruments are shown in Figs. 8(a) and 8(b).

2.3.1 Physical mechanisms of abrasive grain phenomena

During grinding, abrasive grains are subjected to combined mechanical, thermal, and tribological interactions with the workpiece, leading to multiple physical phenomena that may occur independently or simultaneously. On the basis of the established grinding theory and direct surface observations, four representative abrasive grain states were defined in this study. A schematic illustration of these mechanisms is presented in Fig. 9.

(a) (b)

Fig. 8. (Color online) Surface characterization instruments used in this study: (a) VHX-5000 high-depth digital microscope and (b) white-light interferometer (NewView™, ZYGO).

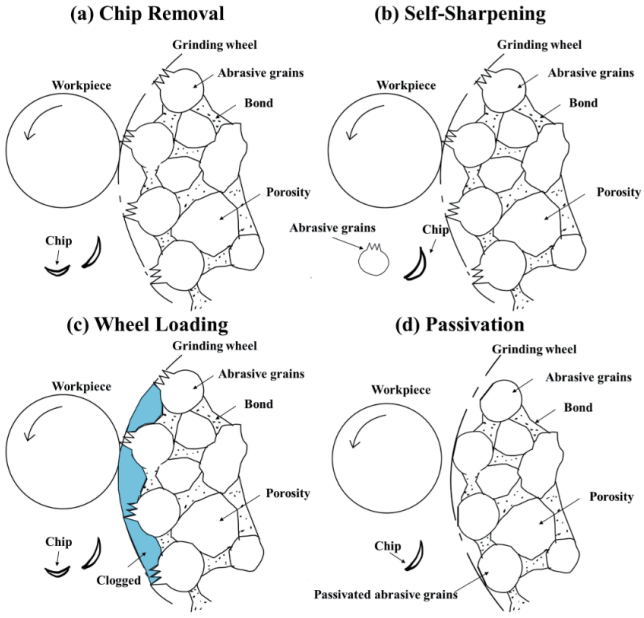


Fig. 9. (Color online) Schematic illustration of four representative abrasive grain phenomena during grinding: chip removal, self-sharpening, wheel loading and passivation.

I. Chip removal
   Abrasive grains actively engage the workpiece, generating chips that are effectively evacuated through wheel porosity. Grain cutting edges remain sharp and protruded, indicating stable cutting action.
II. Self-sharpening
   Micro-fracture or partial pull-out of abrasive grains occurs due to grinding forces, exposing new sharp cutting edges. This phenomenon restores cutting efficiency and often increases effective grain protrusion.
III. Passivation
   Abrasive grains experience progressive wear and edge rounding without fracture, resulting in smooth grain tips and reduced cutting ability. Passivated grains tend to plough or slide rather than cut.
IV. Stuffing/clogging
   Chips adhere to abrasive grains or fill wheel pores, obstructing chip evacuation. This leads to flattened surface morphology and suppressed grain protrusion.

These four phenomena are not mutually exclusive; multiple states may coexist within the same observation region or even on the same abrasive grain.

2.3.2 Direct-imaging-based multi-label annotation strategy

After each grinding interval, the grinding wheel was stopped and imaged at fixed, predefined locations on the wheel surface to ensure spatial consistency throughout the experiment. High-magnification images were repeatedly acquired at identical positions, enabling the direct comparison of abrasive grain evolution during continuous grinding.

Manual annotation was performed on the basis of the direct visual inspection of abrasive grain images. For each captured image, the presence or absence of each physical phenomenon was independently evaluated. A multi-label binary annotation scheme was adopted, as summarized in Table 6.

- Label = 1: the corresponding physical phenomenon is observed
- Label = 0: the corresponding physical phenomenon is not observed

Because multiple phenomena may occur simultaneously, each image may contain more than one positive label. For example, an image exhibiting both chip adhesion and rounded grain edges would be labeled as stuffing = 1 and passivation = 1.

Table 6
Definition of abrasive grain state labels used for manual annotation.

Label category	Physical meaning	Annotation rule
Chip removal	Effective chip evacuation with sharp grains	1 if clear chips and protruded edges are observed
Self-sharpening	Grain fracture or pull-out exposing new edges	1 if fresh sharp edges or fractured grains are visible
Passivation	Grain edge rounding or smoothing	1 if cutting edges appear blunt or polished
Stuffing/clogging	Chip adhesion or pore blockage	1 if chips adhere to grains or fill pores

Representative optical micrographs of abrasive grain states observed after grinding are shown in Fig. 10, where red-marked regions indicate characteristic features associated with each physical phenomenon.

2.3.3 Annotation consistency and reliability

To enhance annotation reliability and reduce subjectivity, a rule-based annotation protocol was established prior to labeling. All annotators followed identical visual criteria defined in Table 6, focusing on grain edge sharpness, chip morphology, and pore condition rather than global surface appearance.

In addition, a subset of images was independently annotated by multiple experts with experience in grinding and surface characterization. Discrepancies were reviewed jointly and consensus labels were assigned through discussion. This procedure ensured the inter-annotator consistency and improved robustness of the labeled dataset.

The resulting annotated dataset thus provides a physically interpretable, visually verifiable, and multi-label ground truth, forming a solid foundation for subsequent AE-based learning and classification.

2.4 Database construction and deep learning framework

To establish an intelligent framework for grinding wheel condition recognition and optimal dressing timing determination, a comprehensive database construction and deep-learning-based classification strategy was developed, as illustrated in Fig. 11. The proposed framework integrates AE signal acquisition, signal preprocessing, manual multi-label annotation, and deep neural network modeling into a unified workflow.

During continuous grinding experiments, raw AE signals were acquired at a sampling frequency of 200 kHz, as determined in Sect. 2.2. From the continuous AE data stream, signal

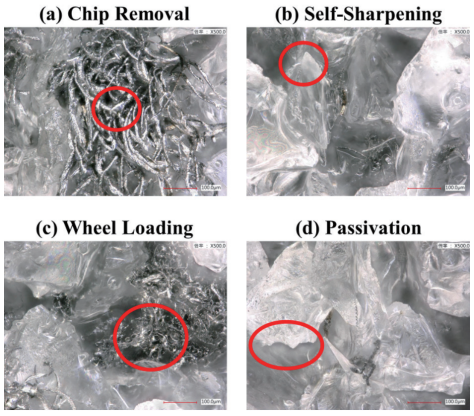


Fig. 10. (Color online) Representative optical micrographs of abrasive grain states captured directly from fixed locations on the grinding wheel surface after grinding (magnification: ×500). Red circles highlight characteristic regions corresponding to different physical phenomena.

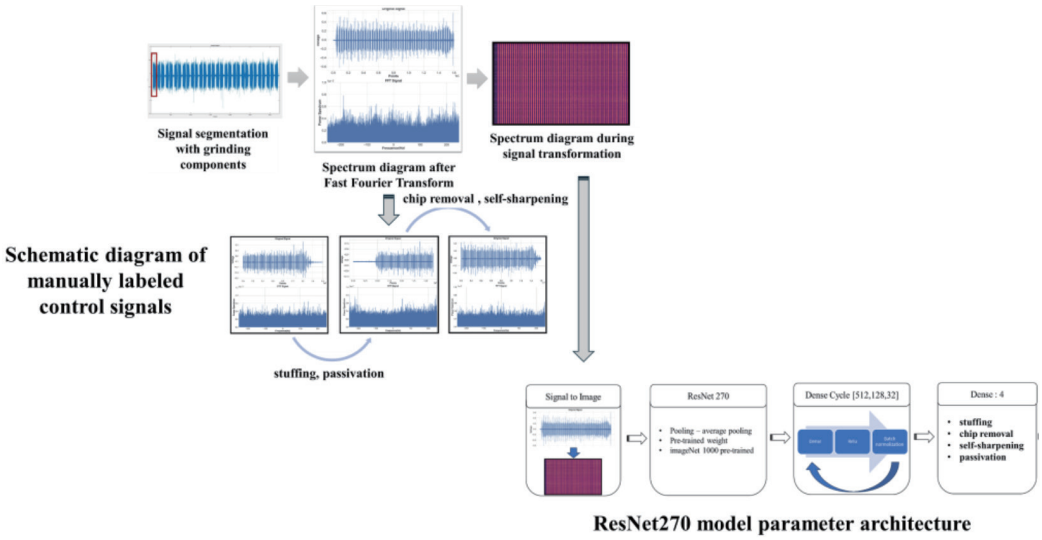


Fig. 11. (Color online) Overall framework of the proposed database construction and multi-label deep-learning-based grinding wheel condition recognition method, including AE signal acquisition, signal segmentation, STFT-based time–frequency transformation, manual multi-label annotation, and ResNet-based classification.

segments corresponding to effective grinding engagement periods were extracted on the basis of temporal continuity and amplitude characteristics, whereas idle and noncontact intervals were excluded. Each extracted segment represents a short-duration grinding event containing sufficient information on the abrasive–workpiece interaction. In total, 960 valid AE signal segments were obtained to construct the complete dataset used in this study.

To enhance the representational capability of AE signals for deep learning, each signal segment was transformed from the time domain into a time–frequency representation. Specifically, STFT-based spectral analysis was applied to preserve both the temporal evolution and frequency-domain characteristics of AE signals. STFT is defined as

$$X(t,f)=\sum_{n=-\infty}^{\infty} x(n)w(n-t)e^{-j2\pi fn}, \tag{1}$$

where $x(n)$ denotes the discrete AE signal, $w(\cdot)$ is the window function, and t and f represent the time and frequency indices, respectively. The magnitude spectrum $|X(t,f)|$ was used to construct time–frequency maps, which were subsequently normalized and converted into image-like inputs compatible with convolutional neural network architectures.

The sampling frequency was selected on the basis of signal quality and stability considerations using *RMSE* analysis. *RMSE* is defined as

$$RMSE=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2}, \tag{2}$$

where x_i represents the AE signal amplitude at sample i, $\bar{x}$ is the mean signal amplitude, and N is the total number of samples. A sampling frequency of 200 kHz was selected as it minimized *RMSE* while preserving characteristic frequency components associated with grinding wheel condition variations.

Unlike conventional single-label classification approaches, the grinding wheel condition labeling in this study adopted a multi-label annotation strategy. Ground truth labels were determined through the direct optical observation of abrasive grain conditions at fixed locations on the grinding wheel surface, as described in Sect. 2.3. Four representative physical phenomena were considered: stuffing, chip removal, self-sharpening, and passivation. Each phenomenon was independently annotated using a binary scheme, where the presence of a specific physical phenomenon was labeled as 1 and its absence as 0. As multiple phenomena may coexist simultaneously within the same grinding state, each AE signal segment was associated with a multi-label vector rather than a mutually exclusive class label, providing a more realistic representation of actual grinding wheel conditions.

For the deep learning model, a ResNet-RS270 architecture was employed as the feature extraction backbone. The network was initialized with ImageNet-pretrained weights to improve convergence stability and feature generalization. Global average pooling was applied to reduce sensitivity to localized noise while preserving global spectral patterns. Following the convolutional backbone, a dense layer sequence consisting of fully connected layers with dimensions [512, 128, 32], rectified linear unit activation functions, and batch normalization was implemented to enhance nonlinear representation capability and training robustness. The final output layer consisted of four sigmoid-activated neurons, corresponding to the four grinding wheel condition labels, enabling independent probability estimation for each phenomenon.

Given the multi-label nature of the problem, binary cross-entropy (BCE) was adopted as the loss function:

$$L = -\frac{1}{K}\sum_{K=1}^{K}[y_k \log(\hat{y}_k)] + (1-y_k)\log(1-\hat{y}_k)], \tag{3}$$

where $K = 4$ represents the number of grinding wheel phenomena, y_k denotes the ground truth label, and $\hat{y}_k$ is the predicted probability for the kth phenomenon. This formulation allows the simultaneous prediction of multiple coexisting grinding wheel states.

The dataset was randomly divided into training and validation sets using a 9:1 ratio. To mitigate data imbalance and improve model generalization, data augmentation techniques were applied to the time–frequency images, including intensity scaling and minor temporal shifting. Model performance was evaluated using metrics suitable for multi-label classification, including accuracy, precision, recall, macro-averaged F1-score, area under the receiver operating characteristic curve (ROC), and area under curves (AUC).

Through this integrated database construction and deep learning framework, AE signal characteristics were effectively mapped to physically meaningful grinding wheel conditions, forming the foundation for the subsequent analysis of dressing timing criteria and intelligent grinding process monitoring.

2.5 Statistical analysis

To enhance the statistical rigor of the experimental results, additional statistical analysis was conducted to evaluate the significance of surface roughness evolution during continuous grinding. Specifically, one-way analysis of variance (ANOVA) was performed to compare the post-grinding surface roughness (*Ra*) values across different cumulative material removal stages (0.6, 1.2, 1.8, and 2.4 mm). Each group consisted of four repeated measurements obtained from independent standard blocks.

The ANOVA results indicate statistically significant differences among the groups ($p < 0.05$), confirming that surface roughness degradation is strongly associated with increasing cumulative material removal. Post hoc comparisons further reveal that the differences between early-stage (≤1.2 mm) and late-stage (≥1.8 mm) grinding conditions are particularly significant, supporting the existence of a transition from stable cutting to degradation-dominated grinding behavior.

In addition, confidence intervals and standard deviation values were analyzed to quantify the variability of surface quality. The increasing variance observed at higher cumulative removal levels further supports the transition from stable cutting conditions to degraded grinding states.

To further assess model generalization, *k*-fold cross-validation ($k = 5$) was conducted. The dataset was randomly partitioned into five subsets with balanced label distributions, and the model was iteratively trained and validated across different splits. The averaged evaluation metrics (accuracy, precision, recall, and F1-score) demonstrate consistent performance across folds, indicating that the proposed framework is not overfitted to a specific data partition.

These statistical and validation procedures provide quantitative support for both the observed surface roughness trends and the robustness of the proposed AE-based multi-label classification model.

3. Results

In this section, we present the experimental and analytical results obtained from the proposed AE-based multi-label framework for grinding wheel condition recognition and dressing timing determination. The results are organized to first establish the macroscopic machining-quality evolution during continuous grinding, using surface roughness as the primary quality indicator. Next, the statistical distribution of the four grinding wheel phenomena labeled from direct optical observations is summarized to clarify the multi-label nature of the dataset. The recognition performance of the proposed deep learning model is then reported using learning curves, ROC analysis, and label-wise evaluation metrics. Finally, consecutive multi-label predictions are correlated with surface roughness degradation to derive and validate the proposed dressing timing criterion.

3.1 Surface roughness evolution under continuous grinding

The evolution of surface roughness under continuous grinding was investigated to evaluate the degradation of grinding performance as a function of cumulative material removal. Four grinding stages were considered, corresponding to cumulative removal depths of 0.6, 1.2, 1.8,

and 2.4 mm. At each grinding stage, four standard blocks were processed under identical grinding conditions. Prior to grinding, the initial surface roughness of all standard blocks was measured and recorded. The grinding wheel was dressed before each grinding stage to ensure a consistent initial wheel condition.

Figure 12 presents the variation in arithmetic average surface roughness (*Ra*) as a function of cumulative material removal. For clarity, mean values with standard deviation are shown for both pre-grinding and post-grinding conditions. The detailed surface roughness data for individual standard blocks are summarized in Table 7.

Overall, the surface roughness before grinding remained relatively stable throughout all grinding stages, with mean *Ra* values confined within a narrow range of 0.108–0.124 μm. This indicates that the initial surface condition of the standard blocks was consistent and that the dressing process effectively restored the grinding wheel to a comparable starting state before each experiment.

In contrast, the surface roughness measured after grinding exhibited a pronounced increasing trend with increasing cumulative material removal. At a cumulative removal of 0.6 mm, the mean post-grinding *Ra* was 0.155 ± 0.027 μm. As the cumulative removal increased to 1.2 and 1.8 mm, the mean *Ra* increased to 0.263 ± 0.009 and 0.313 ± 0.013 μm, respectively. When the cumulative removal reached 2.4 mm, the post-grinding surface roughness rose sharply to 0.416 ± 0.014 μm.

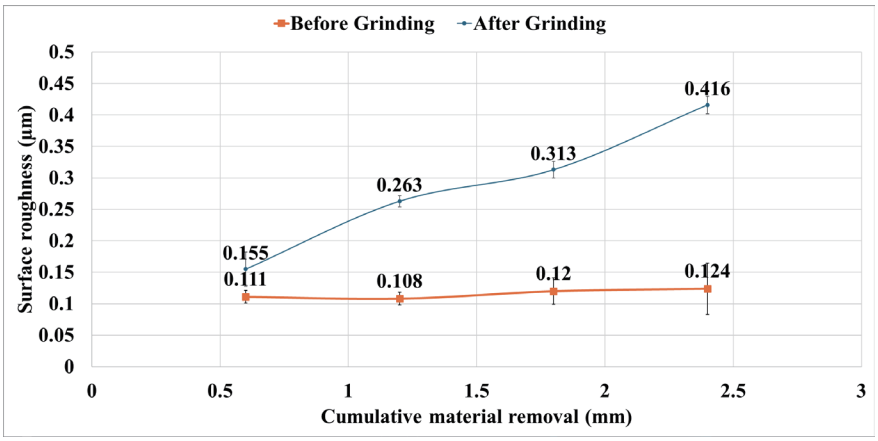


Fig. 12. (Color online) Evolution of surface roughness (*Ra*) as a function of cumulative material removal. Mean values with standard deviation are shown for surface roughness measured before and after grinding at cumulative removal depths of 0.6, 1.2, 1.8, and 2.4 mm.

Table 7
Surface roughness of standard blocks before and after grinding at different cumulative material removal levels.

Cumulative material removal (mm)	Surface roughness before grinding (μm)	Surface roughness after grinding (μm)
0.6	0.111 ± 0.010	0.155 ± 0.027
1.2	0.108 ± 0.010	0.263 ± 0.009
1.8	0.120 ± 0.021	0.313 ± 0.013
2.4	0.124 ± 0.041	0.416 ± 0.014

In addition to the increase in mean surface roughness, a gradual increase in variability was observed at higher cumulative removal levels, as reflected by the standard deviation values. This suggests that grinding performance became less uniform as grinding progressed, even though identical process parameters were maintained. The divergence between pre-grinding and post-grinding surface roughness values became increasingly significant beyond a cumulative removal of approximately 1.8 mm.

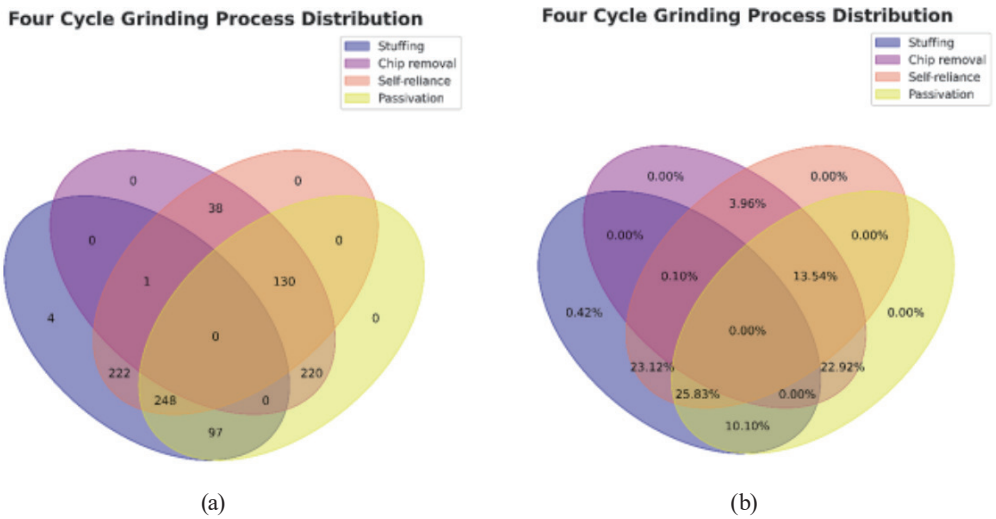
These results clearly demonstrate that surface roughness degradation accelerates with increasing cumulative grinding depth, providing quantitative evidence of progressive changes in grinding behavior during continuous grinding. The observed roughness evolution serves as an important experimental indicator for the subsequent analysis of grinding wheel condition transitions and their relationship with AE characteristics.

3.2 Multi-label recognition performance of grinding wheel phenomena

In this section, we evaluate the recognition performance of the proposed multi-label deep learning framework for grinding wheel condition monitoring. The analysis focuses on three aspects: (i) the intrinsic distribution characteristics of grinding wheel phenomena in the constructed dataset, (ii) overall model classification performance, and (iii) label-wise recognition behavior corresponding to different physical grinding mechanisms.

3.2.1 Distribution of grinding wheel phenomena

The statistical distribution of grinding wheel phenomena within the 960-sample AE dataset is illustrated in Fig. 13. The multi-label annotations were derived from the manual inspection of grinding wheel surface photographs captured after continuous grinding experiments. Abrasive grain conditions were directly observed and classified into stuffing, chip removal, self-sharpening, and passivation according to their physical characteristics. Because the labeling was



(a) (b)

Fig. 13. (Color online) (a) Count-based and (b) percentage-based Venn diagrams illustrating the multi-label distribution of grinding wheel phenomena.

based on the direct visual evaluation of abrasive grain states rather than AE signal features, the ground truth reflects the actual physical condition of the grinding wheel surface.

Owing to the multi-label annotation strategy adopted in this study, each AE signal sample may contain one or more simultaneously occurring physical phenomena, resulting in overlapping label distributions rather than mutually exclusive categories. This multi-label structure is consistent with the physical reality of grinding processes, where abrasive wear, loading, and grain renewal often occur concurrently.

Among the four grinding wheel phenomena, stuffing was the most frequently observed condition, appearing in 475 samples (49.5%), followed by passivation in 445 samples (46.4%), self-sharpening in 391 samples (40.7%), and chip removal in 261 samples (27.2%). Because multiple labels may be assigned to a single sample, the cumulative percentage exceeds 100%, reflecting the compound and transitional nature of real grinding wheel surface states.

Figure 13(a) presents the count-based Venn diagram depicting the co-occurrence structure among the four phenomena, whereas Fig. 13(b) shows the corresponding percentage-based distribution normalized by the total sample size ($N = 960$). A substantial overlap can be observed between stuffing and passivation, as well as between stuffing and self-sharpening, indicating that abrasive loading and grain dulling frequently coexist during continuous grinding. In contrast, purely single-phenomenon samples constitute only a limited proportion of the dataset, further supporting the non-discrete evolution of grinding wheel surface conditions.

These statistical characteristics confirm that grinding wheel condition evolution is not a stage-wise discrete process but rather a continuous and overlapping transition among multiple physical mechanisms. Therefore, modeling the problem as a multi-label recognition task is both physically justified and computationally appropriate.

3.2.2 Model classification performance

The manually annotated AE dataset was randomly divided into training and testing subsets using a 9:1 ratio. The deep learning model was trained for 100 epochs, and its learning behavior is illustrated in Fig. 14.

As shown in Fig. 14(a), the Macro-F1 loss for both training and testing sets decreases rapidly during the early epochs and gradually converges to stable values, indicating effective optimization and stable model convergence. No significant divergence between training and testing curves is observed, suggesting that overfitting is well controlled. Correspondingly, Fig. 14(b) demonstrates that the Macro-F1 score increases steadily and stabilizes at approximately 0.80 for the testing set, confirming satisfactory generalization performance under multi-label conditions.

To further evaluate the discriminative capability of the model for each grinding wheel phenomenon, receiver operating characteristic (ROC) analysis was performed. The ROC curves for stuffing, chip removal, self-sharpening, and passivation are presented in Fig. 15. The mean AUC values for stuffing, chip removal, self-sharpening, and passivation are 0.81, 0.80, 0.73, and 0.81, respectively.

These results indicate that the proposed AE-based time–frequency representation combined with a ResNet-based architecture provides stable discrimination capability for complex and

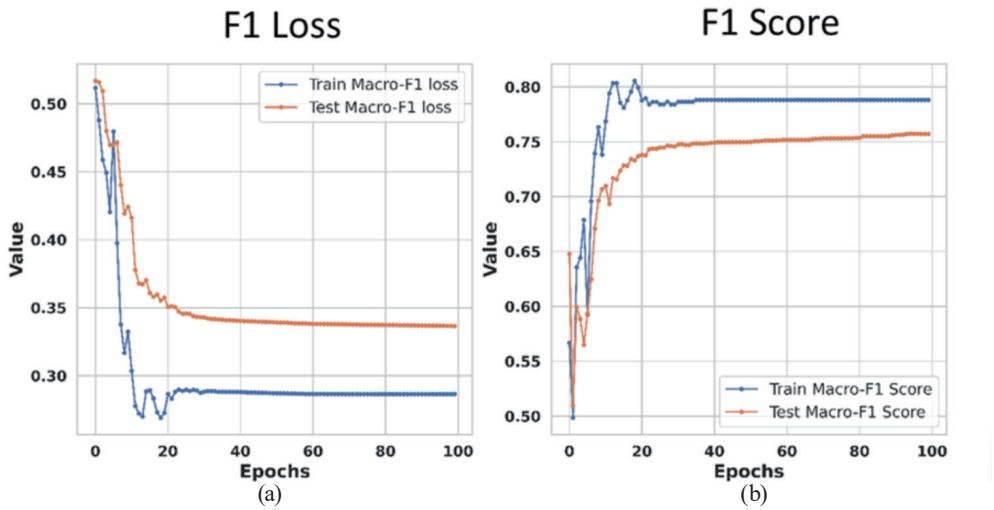


Fig. 14. (Color online) Macro-F1 (a) loss and (b) score curves of training and testing.

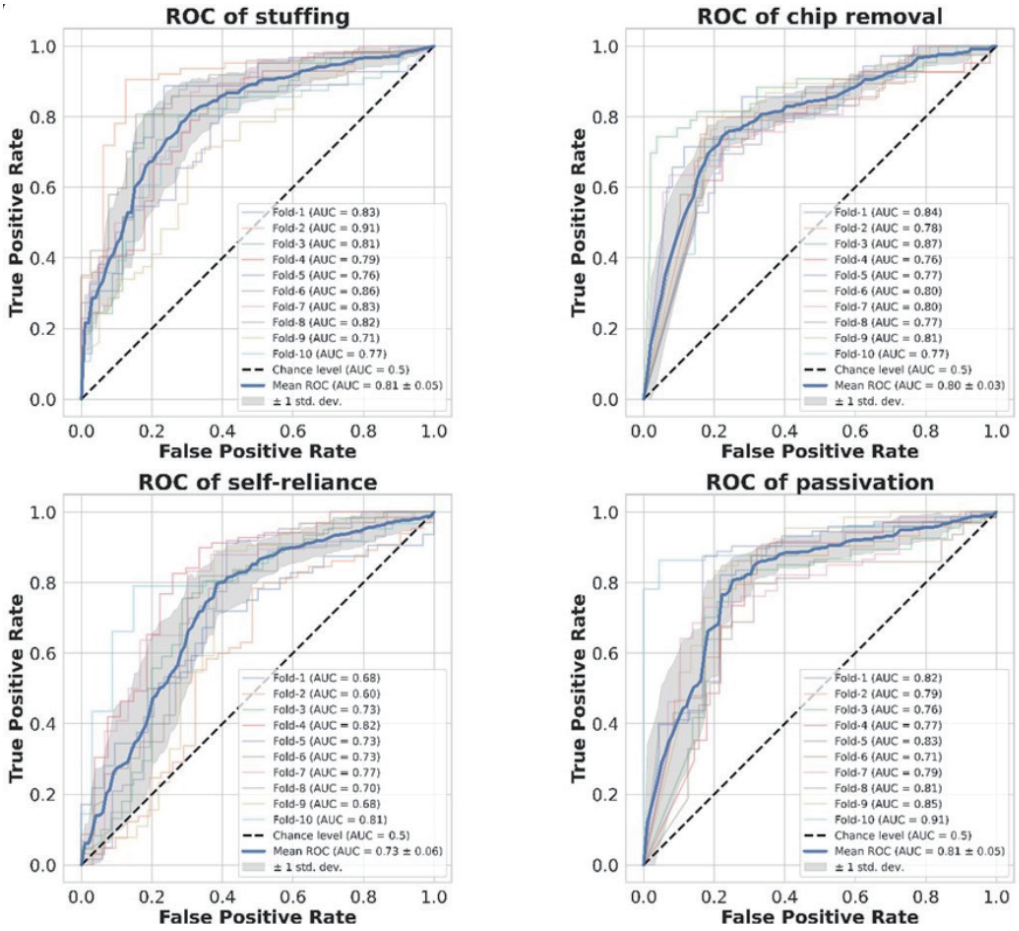


Fig. 15. (Color online) ROC curves for four grinding wheel phenomena.

overlapping grinding wheel phenomena. Although self-sharpening exhibits a comparatively lower AUC (0.73), its performance remains acceptable considering the transient and mixed characteristics of abrasive grain fracture events.

Table 8
Label-wise performance metrics (AUC, accuracy, recall, precision, and F1-score)

Performance	Phenomena			
	Stuffing	Chip removal	Self-sharpening	Passivation
AUC	0.81	0.80	0.73	0.81
Accuracy	0.78	0.78	0.76	0.80
Recall	0.83	0.70	0.82	0.82
Precision	0.81	0.74	0.81	0.89
F1-score	0.82	0.72	0.82	0.86

A quantitative comparison of performance metrics, including accuracy, recall, precision, and F1-score for each label, is summarized in Table 8. Overall, the model achieves balanced performance across all four grinding wheel phenomena, demonstrating the feasibility of AE-based multi-label recognition for continuous grinding processes.

3.2.3 Label-wise recognition characteristics

To further understand the recognition behavior of individual grinding wheel phenomena, label-wise metrics were analyzed on the basis of Table 8.

For stuffing, the model achieves a recall of 0.83 and a precision of 0.81, indicating strong sensitivity to stuffing-related AE signatures while maintaining a low false-positive rate. This result is consistent with the distinct AE characteristics generated by chip accumulation and pore blockage.

Chip removal exhibits comparatively lower recall (0.70) and precision (0.74). This reduced performance can be attributed to its transitional nature, as chip removal frequently coexists with self-sharpening or early-stage stuffing, resulting in less distinct AE signatures. Self-sharpening achieves balanced recall and precision values of 0.82 and 0.81, indicating that the model effectively captures AE patterns associated with abrasive grain fracture and renewal events despite their short duration and overlapping characteristics.

Passivation demonstrates the highest precision (0.89) among all labels, suggesting strong prediction confidence when identifying passivated states. Its recall of 0.82 indicates that most passivation-related samples are correctly detected, although some confusion may occur under stuffing-dominant conditions due to the physical similarity between abrasive dulling and loading mechanisms.

Overall, the label-wise analysis confirms that the proposed multi-label recognition framework not only achieves satisfactory overall performance but also produces physically interpretable classification behavior consistent with known grinding wheel wear mechanisms. This provides a reliable foundation for the subsequent development of intelligent dressing timing strategies.

3.3 Determination of optimal dressing timing based on consecutive multi-label predictions

While the results in Sects. 3.1 and 3.2 demonstrate that the proposed model can accurately recognize individual grinding wheel phenomena from AE signals, grinding wheel dressing in

practical machining environments is not triggered by isolated events. Instead, dressing decisions are typically made on the basis of the accumulation and persistence of wheel degradation during continuous grinding.

Therefore, in this section, we focus on analyzing consecutive prediction behavior obtained from continuous grinding experiments and relating these predictions to measured surface roughness evolution. By evaluating the temporal distribution of predicted grinding wheel phenomena across multiple grinding intervals, a practical dressing timing criterion is established on the basis of sustained degradation patterns rather than single-sample classification results.

In this section, our objective is to bridge the gap between multi-label recognition results and actionable dressing decisions, using consecutive AE-based predictions and surface roughness measurements collected under realistic grinding conditions.

3.3.1 Consecutive prediction behavior during continuous grinding

Continuous grinding experiments were conducted using the machining parameters listed in Table 4. To balance prediction reliability and computational efficiency under long-duration grinding conditions, AE signals were acquired at fixed cumulative material removal intervals of 0.6 mm rather than continuously throughout the entire process.

For each grinding interval, only AE signals corresponding to effective wheel–workpiece contact were retained, excluding idle or noncontact periods. As a result, each grinding interval yielded 30 valid AE signal segments, where the first segment corresponds to wheel entry and the last segment corresponds to wheel exit. This segmentation strategy ensures that the extracted AE data represent the overall grinding behavior of an entire machined surface, rather than isolated transient events.

For each grinding interval,

- only AE signals corresponding to effective wheel–workpiece contact were retained;
- each interval produced 30 valid AE signal segments;
- the first segment corresponds to wheel entry;
- the last segment corresponds to wheel exit.

Each AE segment was independently processed by the trained multi-label classifier to determine the occurrence of four grinding wheel phenomena: stuffing (S), chip removal (C), self-sharpening (R), and passivation (P). The procedure was repeated for eight consecutive grinding intervals, resulting in a total cumulative material removal of 4.8 mm. After each interval, the arithmetic average surface roughness Ra_t of the machined surface was measured.

Let

$$\hat{y}_i^{(k)} \in \{0,1\} \tag{4}$$

denote the predicted label of phenomenon k for the ith AE segment, where

$$k \in \{S,C,R,P\}. \tag{5}$$

The occurrence ratio of phenomenon k at grinding interval t is defined as

$$R_t^{(k)} = \frac{1}{30}\sum_{i=1}^{30}\hat{y}_i^{(k)}. \tag{6}$$

Binary occurrence ratios were adopted instead of probability averaging to emphasize the frequency of physical phenomena appearance, which provides a more interpretable and robust indicator for dressing decision-making in practical grinding operations.

Tables 9–16 show the predicted grinding wheel phenomena and the corresponding measured surface roughness values after each of the eight consecutive grinding intervals. Figure 16 illustrates the evolution of the predicted occurrence ratios $R_t^{(k)}$ together with the measured surface roughness Ra_t.

At the initial grinding intervals, chip removal and self-sharpening exhibited relatively high occurrence ratios, while stuffing and passivation remained low, indicating effective material removal with active abrasive cutting edges. During the intermediate intervals, the occurrence ratios of stuffing and passivation gradually increased, suggesting progressive chip accumulation and abrasive edge wear. In the later intervals, stuffing and passivation became dominant phenomena, while chip removal decreased substantially.

The measured surface roughness showed a consistent increasing trend with cumulative grinding depth. When stuffing and passivation ratios increased significantly, the corresponding surface roughness values approached and exceeded 0.4 μm, indicating a degradation in grinding performance.

Table 9
Predicted grinding wheel phenomena and measured surface roughness after the 1st grinding interval.

No.	Appearence			
	stuffing	chip removal	self-reliance	passivation
1	1	0	0	0
2	0	1	0	1
3	1	0	1	0
4	0	1	0	1
5	1	0	1	0
⋮	⋮	⋮	⋮	⋮
22	1	0	0	1
23	0	1	1	0
24	1	0	1	1
25	1	0	1	1
26	1	0	1	1
27	0	1	1	1
28	1	0	1	1
29	0	1	0	1
30	1	0	0	1

Number one grinding process $Ra \approx 0.1$

Table 10
Predicted grinding wheel phenomena and measured surface roughness after the 2nd grinding interval.

No.	Appearence			
	stuffing	chip removal	self-reliance	passivation
1	1	0	0	1
2	0	1	0	1
3	1	0	1	0
4	1	0	1	1
5	0	1	1	0
⋮	⋮	⋮	⋮	⋮
22	1	0	1	1
23	0	1	0	1
24	1	0	1	0
25	0	1	1	1
26	1	0	1	1
27	1	0	0	1
28	1	0	1	0
29	1	0	1	1
30	1	0	1	0

Number two grinding process $Ra \approx 0.1$

Table 11
Predicted grinding wheel phenomena and measured surface roughness after the 3rd grinding interval.

No.	Appearence			
	stuffing	chip removal	self-reliance	passivation
1	1	0	0	1
2	0	1	0	1
3	1	0	1	1
4	0	1	0	1
5	1	0	1	0
⋮	⋮	⋮	⋮	⋮
22	0	1	1	1
23	0	1	1	0
24	1	0	1	0
25	0	1	0	1
26	1	0	1	0
27	1	0	1	1
28	0	1	1	1
29	1	0	1	1
30	1	0	1	1

Number three grinding process $Ra \approx 0.2$

3.3.2 Dressing timing criterion

On the basis of the consecutive multi-label prediction results, an experimentally validated dressing timing criterion was established by correlating AE-based grinding wheel phenomenon predictions with post-grinding surface roughness evolution. Unlike experience-based or

Table 12
Predicted grinding wheel phenomena and measured surface roughness after the 4th grinding interval.

No.	Appearence			
	stuffing	chip removal	self-reliance	passivation
1	0	1	1	1
2	0	1	0	1
3	1	0	1	0
4	0	1	1	1
5	1	0	1	1
⋮	⋮	⋮	⋮	⋮
22	0	1	0	1
23	1	0	1	0
24	0	1	0	1
25	1	0	1	0
26	0	1	0	1
27	1	0	1	0
28	0	1	0	1
29	1	0	1	0
30	0	1	0	1

Number four grinding process $Ra \approx 0.2$

Table 13
Predicted grinding wheel phenomena and measured surface roughness after the 5th grinding interval.

No.	Appearence			
	stuffing	chip removal	self-reliance	passivation
1	0	1	0	1
2	1	0	1	1
3	1	0	1	0
4	0	1	0	1
5	1	0	1	0
⋮	⋮	⋮	⋮	⋮
22	0	1	1	1
23	1	0	1	1
24	0	1	1	1
25	1	0	1	1
26	1	0	0	1
27	0	0	1	1
28	1	0	0	1
29	0	0	1	1
30	1	0	0	1

Number five grinding process $Ra \approx 0.3$

heuristic thresholds commonly adopted in conventional grinding practice, the proposed criterion is derived directly from continuous grinding experiments, where the AE-based model serves as the primary decision mechanism, and surface roughness measurements are used as the physical validation of wheel performance degradation.

Table 14
Predicted grinding wheel phenomena and measured surface roughness after the 6th grinding interval.

No.	Appearence			
	stuffing	chip removal	self-reliance	passivation
1	1	0	0	1
2	0	1	0	1
3	1	0	1	0
4	0	1	0	1
5	1	0	1	0
⋮	⋮	⋮	⋮	⋮
22	1	0	1	0
23	0	1	0	1
24	1	0	1	0
25	0	1	0	1
26	1	0	1	0
27	0	1	0	1
28	1	0	1	0
29	0	1	0	1
30	1	0	1	0

Number six grinding process $Ra \approx 0.3$

Table 15
Predicted grinding wheel phenomena and measured surface roughness after the 7th grinding interval.

No.	Appearence			
	stuffing	chip removal	self-reliance	passivation
1	1	0	0	1
2	1	0	1	0
3	1	0	1	1
4	0	1	0	1
5	1	0	1	0
⋮	⋮	⋮	⋮	⋮
22	1	0	1	0
23	0	1	0	1
24	1	0	1	1
25	1	0	0	1
26	1	0	1	1
27	1	0	0	1
28	1	0	1	1
29	1	1	0	1
30	1	0	1	1

Number seven grinding process $Ra \approx 0.4$

According to the manufacturer's recommendation for the vitrified-bond WA46 grinding wheel used in this study, wheel dressing is required when the post-grinding surface roughness exceeds

$$Ra < 0.4\ \mu\text{m}, \tag{7}$$

Table 16
Predicted grinding wheel phenomena and measured surface roughness after the 8th grinding interval.

No.	Appearence			
	stuffing	chip removal	self-reliance	passivation
1	1	0	0	1
2	0	1	0	1
3	1	0	1	1
4	1	0	0	1
5	1	0	1	1
⋮	⋮	⋮	⋮	⋮
22	1	0	1	1
23	1	0	1	1
24	1	0	1	0
25	1	0	0	1
26	1	0	1	1
27	1	0	0	1
28	1	0	1	1
29	1	0	0	1
30	1	0	1	1

Number eight grinding process $Ra \approx 0.4$

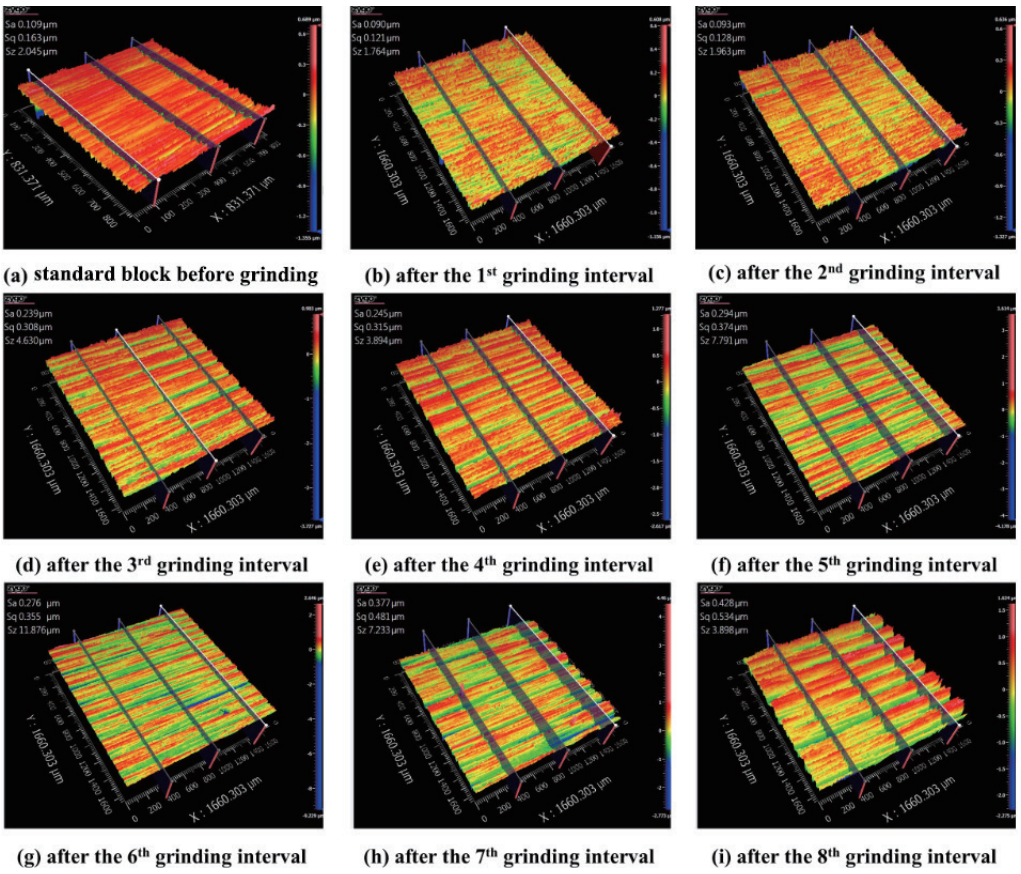


Fig. 16. (Color online) Variations in predicted occurrence ratios of grinding wheel phenomena and corresponding measured surface roughness (*Ra*) values over eight consecutive grinding intervals.

to prevent unstable machining conditions, surface burning, and further quality deterioration. This threshold was therefore adopted not as a control variable, but as a surface quality constraint for validating the effectiveness of the AE-based dressing timing criterion.

Experimental observations revealed a strong and consistent relationship between predicted grinding wheel degradation phenomena and surface roughness deterioration. During the initial grinding intervals, the AE-based model predominantly predicted chip removal and self-sharpening, indicating sufficient active cutting edges and effective material removal. Under these conditions, the measured surface roughness remained stably below the target threshold of $Ra > 0.4$ μm.

As grinding progressed, stuffing gradually increased owing to chip accumulation within wheel pores, while passivation emerged as abrasive grains lost cutting capability. These degradation-related phenomena were accompanied by a clear reduction in effective chip removal and a monotonic increase in surface roughness. This trend demonstrates that AE-based predictions capture the underlying physical mechanisms governing wheel performance degradation.

To quantitatively describe this degradation behavior using AE-based predictions, a degradation index was defined as the combined occurrence ratios of stuffing and passivation at grinding interval t:

$$D_t = R_t^{(s)} + R_t^{(p)}, \tag{8}$$

where $R_t^{(S)}$ and $R_t^{(P)}$ denote the predicted occurrence ratios of stuffing and passivation, respectively.

To mitigate interval-to-interval fluctuations and capture sustained degradation trends, a sliding window average of the degradation index was further introduced:

$$\bar{D}_t = \frac{1}{W} \sum_{j=t-W+1}^{t} D_j, , \tag{9}$$

where W represents the number of consecutive grinding intervals considered in the window.

Accordingly, the dressing timing condition can be mathematically expressed as

$$\bar{D}_t > \theta \text{ and } Ra_t > 0.4 \text{ μm}, \tag{10}$$

where θ is an empirically determined threshold indicating the sustained dominance of degradation-related phenomena.

More importantly, the continuous grinding experiments revealed a critical irreversible degradation behavior that provides strong physical justification for this criterion. Specifically, once the AE-based multi-label model predicted the persistent occurrence (≥five consecutive grinding intervals) of either

- (stuffing, passivation) or
- (stuffing, self-sharpening, passivation),

all subsequent grinding passes failed to restore the surface roughness to the acceptable range of $Ra < 0.4$ μm.

In the present experiments, this irreversible transition was observed after the seventh grinding interval, after which surface roughness consistently exceeded the target threshold regardless of further grinding. This result confirms that the grinding wheel had entered a cumulative degradation stage, where effective cutting edges were no longer sufficient to recover surface quality through continued grinding alone.

At this stage, further grinding only aggravated surface damage and performance recovery was achievable only through wheel dressing. Therefore, the final dressing timing criterion adopted in this study is summarized as follows. When the AE-based multi-label model predicts the consecutive occurrence (≥five grinding intervals) of either (stuffing, passivation) or (stuffing, self-sharpening, passivation), the grinding wheel is judged to have entered an irreversible degradation state, and immediate dressing is required.

3.4 Sensitivity analysis of threshold and window size

The original dressing timing criterion proposed in this study is based on the persistence of degradation-related phenomena over consecutive grinding intervals (≥ five intervals), as described in Sect. 3.3. To further evaluate the robustness of this rule-based criterion from a signal-based perspective, a sensitivity analysis was conducted on the degradation threshold (θ) and the sliding window size (W), as defined in Eqs. (8)–(10).

In this context, θ represents the threshold applied to the degradation index $D(k)$, indicating the dominance level of degradation-related phenomena, while W denotes the number of consecutive grinding intervals used in the sliding average of $D(k)$. These parameters provide an alternative interpretation of the original persistence-based rule by quantifying the magnitude and temporal consistency of degradation.

Different values of θ (0.5, 0.6, and 0.7) and W (3, 5, and 7 intervals) were examined to analyze their effects on dressing decision outcomes based on the experimental dataset. The results show that increasing θ delays the detection of degradation, as higher thresholds require a stronger dominance of degradation-related phenomena before triggering a decision. Conversely, lower θ values lead to earlier detection but may result in premature dressing decisions during transitional grinding stages.

Similarly, the window size W affects the temporal stability of the decision-making process. Smaller window sizes ($W = 3$) provide faster responses to changes in grinding conditions but are more susceptible to short-term fluctuations in AE predictions. Larger window sizes ($W = 7$) produce smoother and more stable decision behavior but may delay the identification of sustained degradation patterns.

Importantly, despite these variations, the onset of irreversible degradation—characterized by the sustained dominance of stuffing and passivation—was consistently identified within a limited interval range across all tested parameter combinations. The variations in dressing trigger points under different parameter settings were observed to remain within one grinding interval, indicating that the proposed dressing criterion is robust with respect to reasonable variations in θ and W.

These findings demonstrate that the original persistence-based rule (≥ five consecutive intervals) is not only physically meaningful but also consistent with signal-based interpretations using thresholding and temporal smoothing. Therefore, θ and W serve as analytical tools for evaluating the stability of the decision-making process rather than replacing the original dressing criterion.

As illustrated in Fig. 17, increasing θ and W results in a delayed but more stable detection of degradation, highlighting the trade-off between responsiveness and robustness. The trends are derived from experimental observations. From these observations, the selected parameters represent a practical balance that ensures the reliable detection of sustained degradation while minimizing the risks of over-dressing and under-dressing in practical grinding operations.

3.5 Summary of key results

In this study, we present an integrated experimental and data-driven framework for the intelligent determination of grinding wheel dressing timing based on AE signals and multi-label deep learning. The key results obtained from continuous grinding experiments, surface roughness measurements, and AE-based prediction analysis can be summarized as follows.

First, continuous grinding experiments demonstrated a clear and progressive degradation of surface quality as cumulative material removal increased. While the pre-grinding surface roughness of standard blocks remained stable across all grinding stages, the post-grinding surface roughness exhibited a monotonic increase with cumulative grinding depth. In particular, when the cumulative material removal exceeded approximately 2.4 mm, the surface roughness rose sharply and exceeded $Ra = 0.4$ μm, indicating a transition from stable grinding to degraded machining performance. This result establishes surface roughness evolution as a reliable physical indicator of grinding wheel condition deterioration.

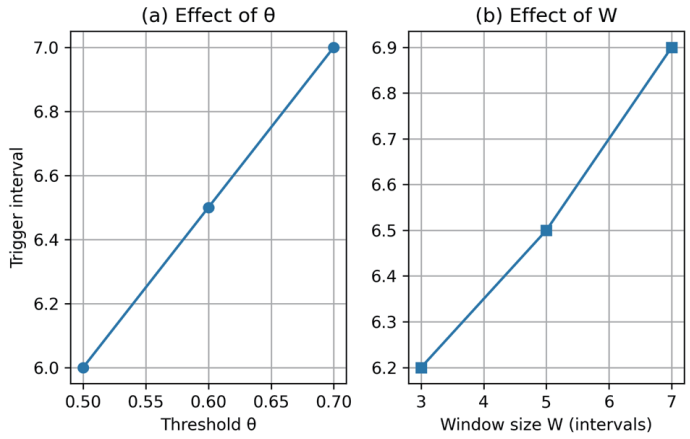


Fig. 17. (Color online) Sensitivity analysis of (a) degradation threshold θ and (b) sliding window size W based on the dressing trigger interval. Increasing θ delays the detection of degradation, while larger W values improve stability but introduce response delay. The variations in trigger points remain within a limited interval range, indicating the robustness of the proposed criterion.

Second, a physically grounded multi-label AE database was constructed through the direct optical observation of abrasive grain states on the grinding wheel surface. Unlike conventional single-label approaches, the proposed annotation strategy allows multiple grinding wheel phenomena—stuffing, chip removal, self-sharpening, and passivation—to be simultaneously assigned to a single AE signal sample. The statistical analysis of the 960-sample dataset revealed substantial overlap among these phenomena, confirming that grinding wheel condition evolution is not a discrete, stage-wise process but a continuous and overlapping transition of multiple physical mechanisms. This finding justifies the necessity of adopting a multi-label recognition framework for realistic grinding condition monitoring.

Third, the proposed ResNet-based deep learning model achieved stable and balanced multi-label recognition performance across all four grinding wheel phenomena. The model converged reliably without overfitting and achieved a macro-averaged F1-score of approximately 0.80 on the test dataset. Label-wise evaluation showed that stuffing and passivation were identified with high precision and recall, while chip removal and self-sharpening exhibited slightly lower but still acceptable performance owing to their transitional and coexisting characteristics. Overall, these results demonstrate that AE-based time–frequency representations combined with deep learning can effectively capture the complex and overlapping physical behaviors of abrasive grains during grinding.

Fourth, consecutive multi-label prediction analysis revealed a strong temporal relationship between predicted grinding wheel phenomena and surface roughness degradation. During early grinding intervals, chip removal and self-sharpening dominated, corresponding to effective cutting action and stable surface quality. As grinding progressed, the occurrence ratios of stuffing and passivation increased steadily, while chip removal decreased. This shift in predicted phenomena was accompanied by a consistent increase in measured surface roughness, confirming that the AE-based model accurately reflects the underlying physical degradation of the grinding wheel.

Finally, an experimentally validated dressing timing criterion was established by integrating consecutive AE-based multi-label predictions with surface roughness measurements. The key finding is that once the AE-based model predicts the persistent occurrence ($\geq$ five consecutive grinding intervals) of degradation-related phenomenon combinations—specifically (stuffing, passivation) or (stuffing, self-sharpening, passivation)—the grinding wheel enters an irreversible degradation stage. Under this condition, subsequent grinding passes are unable to restore surface roughness to the acceptable range of $Ra < 0.4$ μm, regardless of continued grinding. This result confirms that the proposed criterion is not merely heuristic but is physically validated through irreversible surface quality deterioration.

In summary, we demonstrated in this study that combining AE signal analysis, multi-label deep learning, and consecutive prediction behavior enables the reliable identification of grinding wheel degradation and optimal dressing timing. The proposed framework transforms AE-based condition recognition into actionable dressing decisions, providing a physically interpretable, data-driven, and experimentally validated solution for intelligent grinding process control.

4. Discussion

In this study, we proposed an AE-based multi-label deep learning framework for grinding wheel condition recognition and optimal dressing timing determination. The discussion focuses on the physical interpretability of the proposed approach, the advantages of multi-label modeling over conventional classification strategies, and the implications of consecutive prediction behavior for intelligent grinding process control.

4.1 Physical interpretation of AE-based multi-label predictions

The reliability of AE signals is a critical factor in grinding condition monitoring. In this study, signal consistency was ensured by maintaining identical experimental conditions, including fixed machine configuration, workpiece geometry, clamping conditions, and sensor installation location. These controls minimize variations in structural damping and signal transmission. Although changes in machine structure or clamping conditions may affect AE signal characteristics in practical applications, the proposed framework primarily relies on the pattern recognition of time–frequency features rather than absolute signal amplitude, which enhances robustness against such variations. For significantly different setups, model recalibration or transfer learning may be required.

From a physical perspective, the observed overlap between stuffing and passivation phenomena reflects their inherent coupling during the grinding process. When chip adhesion occurs (stuffing), wheel pores become progressively blocked, reducing chip evacuation efficiency. This results in increased friction and sliding at the abrasive–workpiece interface, which accelerates abrasive grain edge rounding and promotes passivation. Therefore, stuffing and passivation frequently coexist and should be interpreted as a physically meaningful transitional state rather than labeling ambiguity. Although visual similarity may introduce minor subjectivity in manual annotation, the labeling process in this study was based on rule-based criteria and expert consensus, ensuring high consistency and physical interpretability.

The experimental results further support this interpretation. During early grinding intervals, chip removal and self-sharpening were dominant, indicating effective cutting action and stable surface quality. As grinding progressed, stuffing and passivation gradually increased owing to chip accumulation and abrasive wear, leading to a reduction in effective chip removal. This transition was accompanied by a consistent increase in measured surface roughness.

The strong correlation between AE-predicted phenomena and surface roughness evolution demonstrates that AE signals capture not only instantaneous cutting events but also the cumulative degradation behavior of the grinding wheel. In particular, the frequent co-occurrence of stuffing and passivation indicates a transitional state in which material removal efficiency deteriorates despite continued grinding. These results highlight the importance of interpreting AE-based predictions from a physical-process perspective rather than treating them as purely abstract classification outputs.

Overall, the AE-based multi-label framework provides physically interpretable predictions that are consistent with established grinding mechanisms, offering a reliable basis for condition monitoring and subsequent process decision-making.

4.2 Irreversible degradation behavior and consecutive prediction analysis

A key finding of this study is the identification of irreversible grinding wheel degradation behavior through consecutive AE-based multi-label predictions. While transient occurrences of stuffing or passivation may arise from normal process fluctuations, the experimental results demonstrate that the sustained dominance of degradation-related phenomena corresponds to irreversible deterioration in surface quality.

Specifically, when the combinations of (stuffing, passivation) or (stuffing, self-sharpening, passivation) persist for five or more consecutive grinding intervals, the grinding wheel is no longer capable of restoring surface roughness to the acceptable range of $Ra < 0.4$ μm through continued grinding alone. In the present experiments, this irreversible transition occurred after the seventh grinding interval, after which all subsequent grinding passes resulted in surface roughness values exceeding the target threshold.

This behavior indicates that the grinding wheel has entered a cumulative degradation stage characterized by severe chip loading, pore blockage, and a reduction in the number of effective cutting edges. At this stage, further grinding not only fails to improve surface quality but may also exacerbate surface damage. Importantly, the proposed AE-based framework is able to detect this degradation trend earlier through consecutive prediction behavior, prior to the occurrence of severe surface roughness deterioration.

The incorporation of occurrence ratios and sliding-window aggregation enhances the robustness of the analysis by suppressing interval-to-interval noise and emphasizing sustained degradation patterns. As a result, consecutive multi-label prediction provides a reliable early indicator of grinding wheel failure modes that cannot be mitigated without dressing.

Overall, in this study, we demonstrated that the persistence of AE-predicted degradation phenomena can be directly linked to irreversible process outcomes, thereby establishing a quantitative, physically interpretable, and experimentally validated criterion for optimal dressing timing in continuous grinding operations.

4.3 Advantages and practical implications of the proposed framework

Compared with conventional grinding wheel monitoring approaches, the proposed AE-based multi-label framework offers several notable advantages. First, it provides physically interpretable outputs by directly associating predicted labels with observable grinding wheel phenomena. This contrasts with conventional black-box indicators that lack direct correspondence to underlying physical mechanisms.

Second, the use of consecutive prediction behavior, rather than single-point thresholds, enables the early identification of irreversible degradation trends. This capability is particularly valuable in industrial environments, where delayed dressing decisions may result in unstable machining conditions, surface burning, or increased scrap rates.

Third, the framework is readily applicable to practical grinding systems. AE sensors can be installed in a nonintrusive manner and operate in real time, making the approach suitable for online condition monitoring. In addition, the multi-label formulation allows flexible extension to

incorporate additional grinding phenomena or adapt to different abrasive–workpiece combinations.

From a practical standpoint, the experimentally derived dressing criterion aligns with established machining experience while providing a quantitative, data-driven justification. By linking AE-based predictions to surface roughness constraints, the proposed framework effectively bridges the gap between sensor-level signal interpretation and quality-oriented process control.

Note that while the persistence-based dressing criterion (≥ five consecutive intervals) was derived under specific experimental conditions (WA46 grinding wheel and S45C workpiece), the underlying principle—namely, the sustained dominance of degradation-related phenomena and the inability to recover surface quality—is generally applicable. Therefore, the framework is transferable, although the specific threshold may require calibration for different materials, wheel types, or process conditions.

Overall, the proposed method demonstrates how sensor-based signal analysis, when combined with physically grounded interpretation and data-driven modeling, can transform raw monitoring data into actionable process decisions. This capability highlights the potential of the framework for integration into sensor-based intelligent manufacturing systems and real-time adaptive process control.

4.4 Limitations and future work

The selection of sampling frequency plays a critical role in capturing AE characteristics associated with different grinding phenomena. In this study, a sampling frequency of 200 kHz was selected on the basis of *RMSE*-based signal stability analysis, balancing spectral resolution and noise suppression.

It is recognized that certain high-frequency components associated with abrasive grain fracture during self-sharpening may exceed this bandwidth. This limitation may partially explain the relatively lower AUC value observed for the self-sharpening category. Future work will consider higher-bandwidth acquisition systems to capture ultrahigh-frequency components and further improve the recognition accuracy of transient grinding events.

4.5 Discussion summary

In summary, we demonstrated that AE-based multi-label deep learning provides a robust and physically interpretable approach for grinding wheel condition monitoring. The proposed framework successfully captures overlapping grinding phenomena, identifies irreversible degradation behavior through consecutive predictions, and establishes an experimentally validated dressing timing criterion.

By integrating AE signal analysis, multi-label classification, and surface quality evaluation, the method offers a reliable foundation for intelligent dressing decision-making in continuous grinding processes. The results suggest strong potential for industrial implementation and future extension to adaptive grinding control systems.

5. Conclusion

In this study, an AE-based sensing and data-driven framework was developed for grinding wheel condition recognition and optimal dressing timing determination during continuous grinding. By integrating high-frequency AE signal acquisition, physically interpretable multi-label annotation, and deep-learning-based signal classification, the proposed approach establishes a direct link between sensor-derived information and grinding process decision-making.

Unlike conventional monitoring methods that rely on single-state classification or empirical thresholds, the proposed framework demonstrates that AE signals can effectively capture the coexistence and evolution of multiple grinding wheel phenomena. Through time–frequency representation and multi-label modeling, key physical mechanisms—chip removal, self-sharpening, stuffing, and passivation—are simultaneously identified, enabling a more realistic and physically consistent interpretation of grinding behavior.

The experimental results confirm that AE-based sensing is highly sensitive to the transition from stable cutting conditions to degradation-dominated states. In particular, the persistence of degradation-related phenomena corresponds to irreversible deterioration in surface quality. This finding enables the establishment of an experimentally validated dressing timing criterion based on consecutive multi-label predictions, thereby transforming sensor signals into actionable process control decisions.

Although the specific dressing threshold depends on the abrasive–workpiece combination and operating conditions, the underlying sensing framework is inherently transferable. The integration of AE sensing, physically grounded interpretation, and data-driven modeling provides a scalable methodology for intelligent monitoring across diverse grinding systems.

Overall, we demonstrated that sensor-based signal analysis, when combined with machine learning and process knowledge, can effectively bridge the gap between raw sensing data and practical manufacturing decisions. The proposed framework offers strong potential for real-time implementation in intelligent machining environments and contributes to the advancement of sensor-driven smart manufacturing technologies.

References

1 T. W. Liao, C. F. Ting, J. Qu, and P. J. Blau: Int. J. Mach. Tools Manuf. **47** (2007) 580. https://doi.org/10.1016/j.ijmachtools.2006.05.008
2 T. W. Liao: Eng. Appl. Artif. Intell. **23** (2010) 74. https://doi.org/10.1016/j.engappai.2009.09.004
3 I. Inasaki and K. Okamura: CIRP Annals **34** (1985) 277. https://doi.org/10.1016/S0007-8506(07)61772-7
4 K. Y. Li, Y. F. Lin, M. Y. Tsai, I. C. Chiu, and J. Y. Chen: Sens. Mater. **33** (2021) 3513. https://doi.org/10.18494/sam.2021.3578
5 J. Badger, S. Murphy, and G. E. O'Donnell: Int. J. Mach. Tools Manuf. **125** (2018) 11. https://doi.org/10.1016/j.ijmachtools.2017.11.007
6 Z. Wang, P. Willett, P. R. DeAguiar, and J. Webster: Int. J. Mach. Tools Manuf. **41** (2001) 283. https://doi.org/10.1016/S0890-6955(00)00057-2
7 P. Lezanski: J. Mater. Proc. Technol. **109** (2001) 258. https://doi.org/10.1016/S0924-0136(00)00808-6
8 D. F. G. Moia, I. H. Thomazella, P. R. Aguiar, E. C. Bianchi, C. H. R. Martins, and M. Marchi: J. Braz. Soc. Mech. Sci. Eng. **37** (2015) 627. https://doi.org/10.1007/s40430-014-0191-6

9 D. González, J. Alvarez, J. A. Sánchez, L. Godino, and Pombo: Sensors **22** (2022) 14. https://doi.org/10.3390/s22186911
10 T. Kon, H. Mano, H. Iwai, Y. Ando, A. Korenaga, T. Ohana, K. Ashida, and Y. Wakazono: Lubricants **12** (2024) 12. https://doi.org/10.3390/lubricants12030100
11 T. Jessel, C. Byrne, M. Eaton, B. Merrifield, S. Harris, and R. Pullin: Int. J. Adv. Manuf. Technol. **130** (2024) 1107. https://doi.org/10.1007/s00170-023-12700-7
12 A. Shokrani, H. Dogan, D. Burian, T. D. Nwabueze, P. Kolar, Z. R. Liao, A. Sadek, R. Teti, P. Wang, R. Pavel, and T. Schmitz: CIRP J. Manuf. Sci. Technol. **51** (2024) 263. https://doi.org/10.1016/j.cirpj.2024.05.001
13 M. Dado, P. Koleda, F. Vlasic, and J. Salva: Appl. Sci.-Basel **15** (2025) 18. https://doi.org/10.3390/app15126659
14 L. Abu, D. H. Huo, and Z. P. Liu: Int. J. Adv. Manuf. Technol. (2026) 20. https://doi.org/10.1007/s00170-026-17960-7
15 J. I. Z. Chen and W. C. Lo: Sens. Mater. **35** (2023) 4597. https://doi.org/10.18494/sam4545
16 C. K. Huang, C. H. Chen, K. Y. Li, and S. J. Wei: Sens. Mater. **36** (2024) 3817. https://doi.org/10.18494/sam5175
17 L. Wu, K. C. Wang, H. Gao, L. Qiang, and C. H. Yang: Sens. Mater. **36** (2024) 2765. https://doi.org/10.18494/sam4878
18 J. S. Kwak and M. K. Ha: J. Mater. Proc. Technol. **147** (2004) 65. https://doi.org/10.1016/j.jmatprotec.2003.11.016
19 P. S. Krishnan and K. Rameshkumar: Mater. Today: Proc. **46** (2021) 9168. https://doi.org/10.1016/j.matpr.2019.12.428
20 F. R. L. Dotto, P. R. Aguiar, F. A. Alexandre, L. Simões, W. N. Lopes, D. M. D'Addona, and E. C. Bianchi: Procedia CIRP **79** (2019) 298. https://doi.org/10.1016/j.procir.2019.02.070
21 C. H. Shen: Appl. Acoust. **196** (2022) 10. https://doi.org/10.1016/j.apacoust.2022.108863
22 S. Wang, G. Y. Sun, Q. L. Zhao, and X. D. Yang: J. Manuf. Process. **90** (2023) 94. https://doi.org/10.1016/j.jmapro.2023.02.009

About the Authors

Yin-Wei Chao received his B.S. and M.S. degrees from National Chin-Yi University of Technology, Taiwan, in 2021 and 2023, respectively. His research interests are in artificial neural network and smart manufacturing. (kraker00622@gmail.com)

Chun-Yen Chen received his B.S. and M.S. degrees from National Chin-Yi University of Technology, Taiwan, in 2013 and 2015, respectively. He has been a Ph.D. student in the same university since 2021. His research interests are in precision machining and smart manufacturing. (jack1baycomtw@gmail.com)

Yue-Feng Lin received his B.S., M.S., and Ph.D. degrees from National Cheng-Kung University, Taiwan, in 2008, 2009, and 2016, respectively. From 2017 to 2018, he was a postdoctoral researcher at National Chung Hsing University, Taiwan. Since 2018, he has been an assistant professor at National Chin-Yi University of Technology, Taiwan. His research interests include the machining and finishing of difficult-to-cut materials. (yflin@ncut.edu.tw)

Ming-Yi Tsai received his B.S. degree from National Taiwan University of Science and Technology and his M.S. degree from National Chung Hsing University, Taiwan, in 1998. He received his Ph.D. degree from National Taiwan University, Taiwan, in 2006. He is a professor at the Department of Mechanical Engineering, National Chin-Yi University of Technology. His research interests include the grinding and polishing of difficult-to-cut materials. (mytsai@ncut.edu.tw)

Kai-Jung Chen is an accomplished individual with a bachelor's degree in engineering (2010) and a master's degree in biomechanical engineering (2012) from National Cheng Kung University in Taiwan, and a Ph.D. degree in engineering (2019) from the University of Liverpool, UK. From 2014 to 2019, he played a pivotal role in the Biomechanical Engineering Group at the University of Liverpool, focusing on precision ophthalmic equipment development. Since 2020, he has been a faculty member in the Department of Mechanical Engineering at National Chin-Yi University of Technology, Taiwan. His diverse research interests include computational biomechanics, medical assistive device development, big data and AI technology, and semiconductor and brittle material processing. (hskchen5@ncut.edu.tw)