

Rate-Aware KAN-Mamba Framework for Data-Driven Tool Wear State Prediction from Multi-Rate Spindle Sensor Stream

Ondřej Pokorný^{1,*} and Vladimír Růžička¹

¹ Faculty of Electrical Engineering, University of West Bohemia, Plzeň, 306 14, Czech Republic

*Corresponding author: onftrj.pok@zcu.cz

Abstract. This paper proposes a rate-aware KAN-Mamba framework for data-driven tool wear state prediction from multi-rate spindle sensor streams. The core methodological contributions center on three tightly integrated components that foreground sequential modeling, multi-rate sensor fusion, and efficient long-sequence engineering rather than machining-specific heuristics. First, rate-aware tokenization divides asynchronous spindle channels into independent segments according to their native sampling rates, preserving both high-frequency transient events and low-frequency supervisory evidence. Second, a KAN-based nonlinear feature mapping learns conditional local basis functions that align processing-state interactions with sensor readings, enabling adaptive non-linear transformation across varying regimes. Third, a Mamba-style selective state update module compresses long-term temporal memory while applying context-gated updates, capturing degradation sequences with minimal computational overhead. Extensive experiments on industrial spindle sensor streams demonstrate substantial practical gains: improved RMSE and MAE under transient and severe-wear regimes, transition-state tracking accuracy, and robustness to missing or noisy channels. Compared with LSTM, GRU, and Transformer baselines, the framework outperforms by more than 20 % in MAE while achieving an 18.6-minute early-warning lead time and 18.9 ms inference latency under robustness tests. By foregrounding rate-aware multi-rate sensor fusion and KAN-Mamba sequential modeling, the approach supplies an efficient, transferable solution for industrial time-series analytics and aligns directly with data-centric engineering objectives for smart manufacturing decision support.

Keywords: *Tool Wear Prediction; Multi-Rate Sensor Stream; Spindle Monitoring; KAN-Mamba; Smart Manufacturing*

Received on 29 March 2022, Accepted on 03 August 2022, Published on 19 August 2022

Copyright © 2022 Author(s), licensed to DEA. This is an open access article distributed under the terms of the CC BY-NC-SA 4.0, which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

Introduction

Prediction of Spindle Tool Wear for Intelligent Machining is needed to prevent damage to the tool and other negative impacts on surface quality, dimensional accuracy, cutting stability and production cost. Spindle vibration, acoustic emission, current, torque, temperature, feed rate and speed signals are not naturally synchronous on a modern production line, and these signals are subject to different noise sources and sampling frequencies. Zhao et al. [1] believe that the intelligent manufacturing model should convert raw sensor data into useful process information, and Wang et al. [2] have built hierarchical representations for complex industrial signals in machine health monitoring using deep learning. Thus, tool-wear prediction should not be regarded as a single-channel regression problem. A useful model should learn how high-rate vibration bursts, intermittent acoustic-emission changes, load-related current fluctuation, and slow temperature response jointly describe the same wear process.

The issue is not whether the model can fit a wear curve, but whether it can represent sensor data under various circumstances of spindle speed, feed rate, tool geometry and workpiece batch. Li et al. [3] have shown the value of audio signals for data-driven tool-wear monitoring, and Martínez-Arellano et al. [4] have used deep learning to classify tool-wear conditions by converting time-series signals into images. However, these methods are still not feasible when several sensors generate evidence at different rates and with various delays. A Spindle

Monitoring System needs to collect data on vibration, acoustic emission, current, torque, speed, temperature, etc., and distinguish between early-wear signals and transient operating fluctuations. Therefore, this paper will consider the spindle monitoring data as multi-rate sensor streams rather than a fixed synchronous feature table.

Recently, Recurrent Neural Networks, attention mechanisms, physics-informed learning and multi-sensor fusion have been applied to tool-wear prediction. Physics-informed neural networks can improve machining tool-wear prediction by adding process knowledge [5], and bidirectional recurrent models can predict tool wear based on cutting-force signals [6]. Time-space feature fusion also shows that multiple signal groups contain complementary information for tool-wear monitoring [7]. Conventional recurrent models may place too much weight on local fluctuations or overlook early deterioration, and attention-based models are costly for extended high-rate sensor streams. Therefore, the Mamba-style selective state update in the proposed KAN-Mamba architecture can efficiently compress long multi-rate sequences without full attention, and the KAN-style nonlinear mapping can model machining-condition-dependent interactions among sensors.

This is how the remainder of the paper is structured. Tool-wear monitoring, multi-rate industrial time-series fusion, and effective long-sequence learning for manufacturing sensor streams are reviewed in Section 2. The suggested KAN-Mamba prediction approach is presented in Section 3. It consists of Mamba temporal prediction with a wear-state loss function, KAN-based nonlinear wear feature mapping, and multi-rate spindle sensor encoding. The simulated dataset setup, baseline comparison, error behavior, robustness analysis, ablation findings, and deployment value are all shown in Section 4. The paper's contents are succinctly summarized in Section 5, which also offers realistic maintenance scenarios, future development directions, and applications of multi-rate sensing in adaptive machining process management.

Related Work

Tool Wear Monitoring under Variable Cutting Conditions

Online condition evaluation has replaced offline observation in tool wear monitoring. Surface-roughness testing, microscope measurements, and routine manual replacements are the most traditional forms of inspection. The techniques provide precise confirmations upon cutting; however, they are unable to provide prompt notifications during high-speed manufacturing. By measuring force, vibration, acoustic emission, spindle current, temperature, or sound signals, sensor-based monitoring often deduces wear. Multi-sensor data fusion can facilitate indirect tool-wear measurement during machining [8], and deep visualization can assist CNN-based monitoring in interpreting end-milling tool states [9]. The research demonstrates that the monitoring system cannot rely on a single perfect laboratory signal. The model must identify whether wear, cutting load, material non-uniformity, or transient disturbance is the source of change in practical tool wear, which is demonstrated in several interconnected ways.

Tool-Wear Prediction Is Complicated by Variable Cutting Conditions. At varying spindle speeds and feed rates, the same flank-wear value will produce a variable sensor pattern. A sudden rise in vibration might be caused by real tool damage along a steady cutting path, but it could also be the result of a brief change in the hardness of the material or contact between the tool and the workpiece. Spectral characteristics for tool-wear monitoring may be modeled by a pyramid LSTM autoencoder, suggesting that lengthy temporal patterns are required for deterioration identification [10]. Industrial time-series models should distinguish between long-term trends and transient variations, according to temporal fusion transformers for interpretable multi-horizon forecasting [11]. A maintenance choice cannot be taken at random due to the non-uniform distribution of the abrupt increase in tool monitoring data. Find out if the increase is single, recurrent, related to load, or consistent with a longer-term wear reduction.

Multi-Rate Sensor Fusion for Industrial Time Series

Multi-rate sensing is used in several machining operations. Thermal or supervisory control signals are less common, current and torque are often at a medium frequency, and vibration and acoustic-emission signals can be recorded at extremely high frequencies. Low-rate evidence may be overrepresented and high-rate evidence may be smoothed out if all channels are compelled to utilize the same sample grid without thought. Representation design in sequence modeling has been the main focus of deep learning research for time-series

classification [12]. Multi-scale temporal structures can also improve deep time-series models [13]. The spindle sensor stream includes both slow-degradation patterns and high-frequency events, making the concepts appropriate. As a result, the representation ought to construct distinct local descriptors before merging them into a common wear-state sequence.

Fault tolerance is also necessary for sensor fusion. sensor drift in industrial settings, channel loss, and measurement noise changes brought on by installation position. The issue of asynchronous streams cannot be solved by local feature learning alone, while one-dimensional convolutional neural networks are capable of efficiently extracting local patterns from unprocessed data [14]. Severe wear events might be regarded as anomalies that happen seldom within the usual cutting window since research on anomaly identification makes a distinction between normal changes and uncommon abnormal occurrences [15]. Therefore, a robust model of tool wear should integrate sensor data at diverse frequencies, give these streams distinct time-based interpretations, and refrain from treating all variations identically.

KAN-Mamba and Efficient Long-Sequence Learning

Conventional neural networks are challenged by lengthy machining sequences. Although complete attention can be computationally costly for high-rate vibration windows with several time points, attention-based models are effective at simulating global dependencies. For lengthy noisy sequences, recurrent models may not be able to preserve pertinent early information, despite their convenience. Sensing, representation, prediction, and decision support are the steps of predictive maintenance in Industry 4.0 [16]. Predictive maintenance machine-learning techniques must also be useful in industrial settings [17]. These perspectives inspire an architecture that is both compact enough for online use and capable of modeling lengthy sensor streams.

This is possible using the suggested KAN-Mamba model. The non-linear link between machining conditions and sensor response is demonstrated using KAN-style basis mapping. distinct slopes and local forms of feature transformation can be stated to reflect distinct deterioration regimes rather than utilizing a single neural layer for all spindle speeds. To retain pertinent sensor history and disregard noisy fluctuations, a Mamba-style selective state update is employed. Model outputs must continue to be interpretable when used for decision support [18], and multimodal fusion should preserve complimentary evidence rather than mindlessly integrating all channels [19]. Thus, tool-wear prediction is the process of determining which sensor groups, cutting circumstances, and time periods are pertinent to a warning in addition to producing a wear index.

Prediction accuracy is also impacted by the quality of industrial data. Because the tool-wear data frequently contains noisy labels, resampled sequences, and missing channels, data cleaning and machine learning must be coupled [20]. The labeling may be scant or delayed since the tool is not taken out for a microscopic inspection during real milling. In order to provide reliable predictions with incomplete observations, the model must be able to learn from incomplete data. Consequently, the prediction process will be constructed using multi-rate alignment, nonlinear feature mapping, selective temporal memory, and confidence-aware wear-state output in the technique section that follows.

Proposed KAN-Mamba Model

Multi-Rate Spindle Sensor Encoding

The input is a collection of spindle-related sensor streams acquired during milling or rotating. The measurement ranges, sampling frequencies, and physical meanings of the channels vary. The suggested encoder preserves the multi-rate structure rather than first translating all the signals to a single coarse sample grid. The local evidence in each channel is transformed into a rate-aware token, and all channels share the same machining window. High-rate vibration and acoustic-emission transient friction data may be preserved via the method, while load and thermal context can be preserved in the slower current, torque, speed, and temperature channels.

$$x_c(t) = \{v_{c,1}, v_{c,2}, \dots, v_{c,n_c}\} \quad \text{Eq.(1)}$$

Because the sample frequencies for vibration, acoustic emission, spindle current, torque, and temperature vary per channel, the length of these values varies as well. Each stream is split into local segments with a duration

that corresponds to the cutting operation; an erratic sampling rate is not employed in order to allow comparison within the same machining window.

$$W_c^k = \{x_c(t) \mid \tau_k \leq t < \tau_k + \Delta_c\} \quad \text{Eq.(2)}$$

The window index is the same for the machining time, but the window duration is per-channel. Therefore, even assuming the same temporal density for both signals, it is evident that a high-rate vibration burst is associated with a low-rate load rise.

The rate to reference stream ratio is displayed below. In addition to acting as a preprocessing parameter, this relationship is attached to the representation to let the model know if a feature comes from slow-process supervision or dense local sampling.

$$z_c^k = \phi_c(W_c^k, r_c, \mu_c, \sigma_c) \quad \text{Eq.(3)}$$

Every local segment is transformed into a compressed channel token by the encoder. To lessen the effect of variations in sensor amplitude, mean and scale terms are introduced. Low-rate channels maintain the slowly changing operational environment, whereas high-rate channels maintain the transient energy properties.

$$s_k = [z_1^k, z_2^k, \dots, z_C^k] \quad \text{Eq.(4)}$$

All channel tokens are concatenated or projected to generate the fused sensor token at time step k. Channel identity and rate information are retained so that early averaging does not erase whether a feature comes from high-frequency local events or low-frequency operating context. The complete multi-rate spindle sensor encoding pipeline is shown in Figure 1, including asynchronous sensor acquisition, rate-aware token construction, synchronized wear-state input, and fused representation generation.

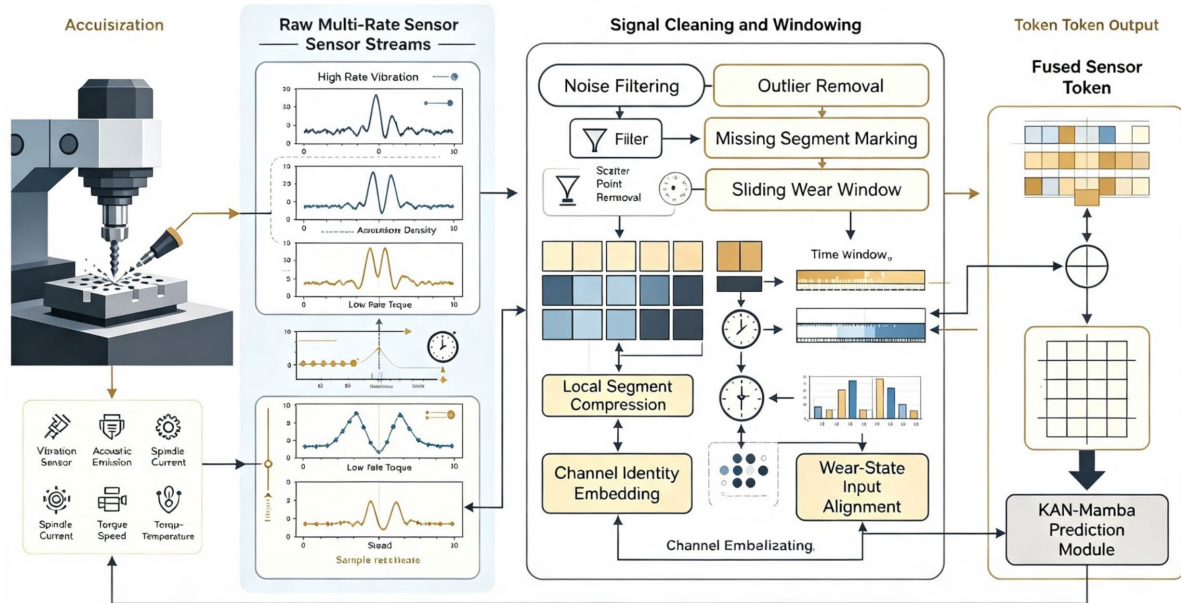


Figure 1. Multi-rate spindle sensor encoding pipeline.

KAN-Based Nonlinear Wear Feature Mapping

After generating sensor tokens, the link between wear-sensitive features and sensor evidence is modeled by the KAN-based nonlinear mapping layer. One fixed nonlinear transformation is not appropriate since varied wear circumstances at different spindle speeds, feed rates, and cutting loads might result in the same sensor change. KAN-style mapping may distinguish between real tool deterioration and temporary load changes in the model by using local basis functions to reflect machining-condition-dependent feature transformations.

$$h_k = \sum_{j=1}^m a_j B_j(s_k) \quad \text{Eq.(5)}$$

Flexible local transformations of the sensor token are provided via basis functions. Different bases are taught to contribute differentially to the wear-sensitive representation in various regions of the machining-state space rather than a single activation function.

The condition-dependent feature map is compatible with certain sensors. Only under a steady cutting load may vibration energy and spindle current together show flank wear; otherwise, a rise in current might be the result of material hardness or a transient load shift. In order to minimize false alarms brought on by load-induced sensor oscillations, the KAN mapping layer can differentiate between changes in the operating environment and real tool wear.

$$\alpha_{c,k} = \frac{\exp(q_k^T u_c)}{\sum_{l=1}^c \exp(q_k^T u_l)} \quad \text{Eq.(6)}$$

To assign varying weights to the sensor groups, a channel attention weight is learnt. Reduce the weight of channels mostly impacted by transient noise and increase the weight of channels with continuous deterioration indications.

$$g_k = \sum_{c=1}^c \alpha_{c,k} h_{c,k} \quad \text{Eq.(7)}$$

A weighted sum of the channel-specific mapped features is the final nonlinear wear feature. Because it includes nonlinear machining-condition interaction and multi-rate sensor evidence, this feature is more descriptive than a raw fused vector.

Mamba Temporal Prediction and Loss Function

Although transition or severe-wear states might happen quickly, tool wear usually rises gradually. Long sensor sequences are therefore modeled as selective state updates in the Mamba manner. It may recall significant deterioration history and disregard transient noise since the concealed state is updated using both the prior state and the present wear characteristic. The first high-frequency evidence that could appear before a slower change in temperature or current is noticed is stored in the structure mentioned above.

$$m_k = A_k m_{k-1} + B_k g_k \quad \text{Eq.(8)}$$

The model will react differently to wear transition windows, chatter-like disturbances, and steady cutting since the transition parameters are input-dependent. Only a portion of the high-amplitude signals may be conveniently recorded at a lower level.

$$A_k = \sigma(P_A g_k), B_k = \sigma(P_B g_k) \quad \text{Eq.(9)}$$

Open the gates selectively to add fresh information and keep only a portion of the prior memory. The model exhibits an ancient deterioration pattern when the sensor evidence is weak; it changes states more quickly when substantial wear evidence is present.

$$o_k = P_o m_k + b_o \quad \text{Eq.(10)}$$

The prediction head's output vector is the current hidden wear state. The head can be either a discrete wear-state classifier or a continuous wear-value predictor, depending on the labels that are supplied. Figure 2 illustrates the relationship between KAN-based nonlinear feature transformation, Mamba temporal prediction, and maintenance-oriented output.

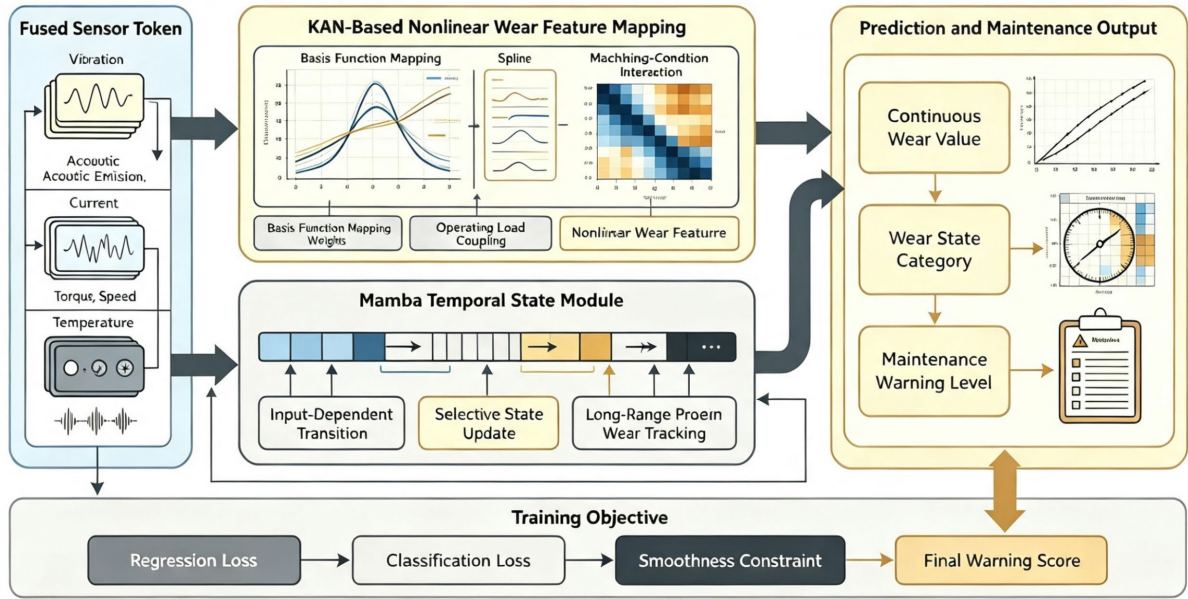


Figure 2. KAN-Mamba feature transformation and temporal prediction structure.

$$\hat{y}_k = \psi(o_k) \quad \text{Eq.(11)}$$

The prediction head maps the hidden state to a tool-wear value or wear category. The model is designed to support normal wear, transition wear, and severe wear identification.

$$L_{\text{reg}} = \frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2 \quad \text{Eq.(12)}$$

The regression term penalizes deviation between predicted and reference wear values. It is suitable when flank-wear measurement or estimated wear depth is available.

A classification term is used when the output is divided into wear states. It supports maintenance decisions because operators often need an interpretable state label rather than only a continuous number.

$$L_{\text{smooth}} = \frac{1}{N-1} \sum_{k=2}^N \|\hat{y}_k - \hat{y}_{k-1}\|_1 \quad \text{Eq.(13)}$$

A smoothness term is added because tool wear should not oscillate sharply in normal operation. The term does not prohibit abrupt transition, but it prevents the model from producing unrealistic high-frequency jumps.

$$L = L_{\text{reg}} + \lambda_1 L_{\text{cls}} + \lambda_2 L_{\text{smooth}} \quad \text{Eq.(14)}$$

The final objective combines numerical accuracy, state discrimination, and temporal consistency. In deployment, the model output is further transformed into a maintenance warning level.

$$R_k = \eta_1 \hat{y}_k + \eta_2 \Delta \hat{y}_k + \eta_3 (1 - \gamma_k) \quad \text{Eq.(15)}$$

The warning score integrates predicted wear, short-term wear acceleration, and confidence. A low-confidence severe prediction can be routed to human review, while a high-confidence acceleration signal can trigger early tool replacement.

$$\gamma_k = 1 - \frac{H(\hat{p}_k)}{\log Q} \quad \text{Eq.(16)}$$

The confidence is computed from normalized prediction entropy. This makes the warning output more useful because the system can distinguish a clear severe-wear signal from an uncertain boundary case.

Experimental Results and Discussion

Dataset Construction and Experimental Setting

The experimental dataset is designed as a synthetic multi-rate spindle monitoring task. It contains 14,400 machining windows generated under different spindle speeds, feed rates, cutting depths, tool materials, and workpiece hardness levels. Vibration is sampled at the highest frequency, acoustic emission at a high but intermittent frequency, current and torque at medium frequency, and temperature at low frequency. Each window is 120 seconds long. Physics-guided machining research indicates that sensor data should be evaluated with cutting-process context [21], and tool-wear classification research supports time-series transformation for wear-state analysis [22]. Figure 3 presents the dataset composition and machining conditions: Figure 3(a) reports normal, moderate, transitional, and severe wear distributions; Figure 3(b) shows sensor-rate groups; Figure 3(c) summarizes machining-condition coverage; and Figure 3(d) reports missing or noisy channel proportions.

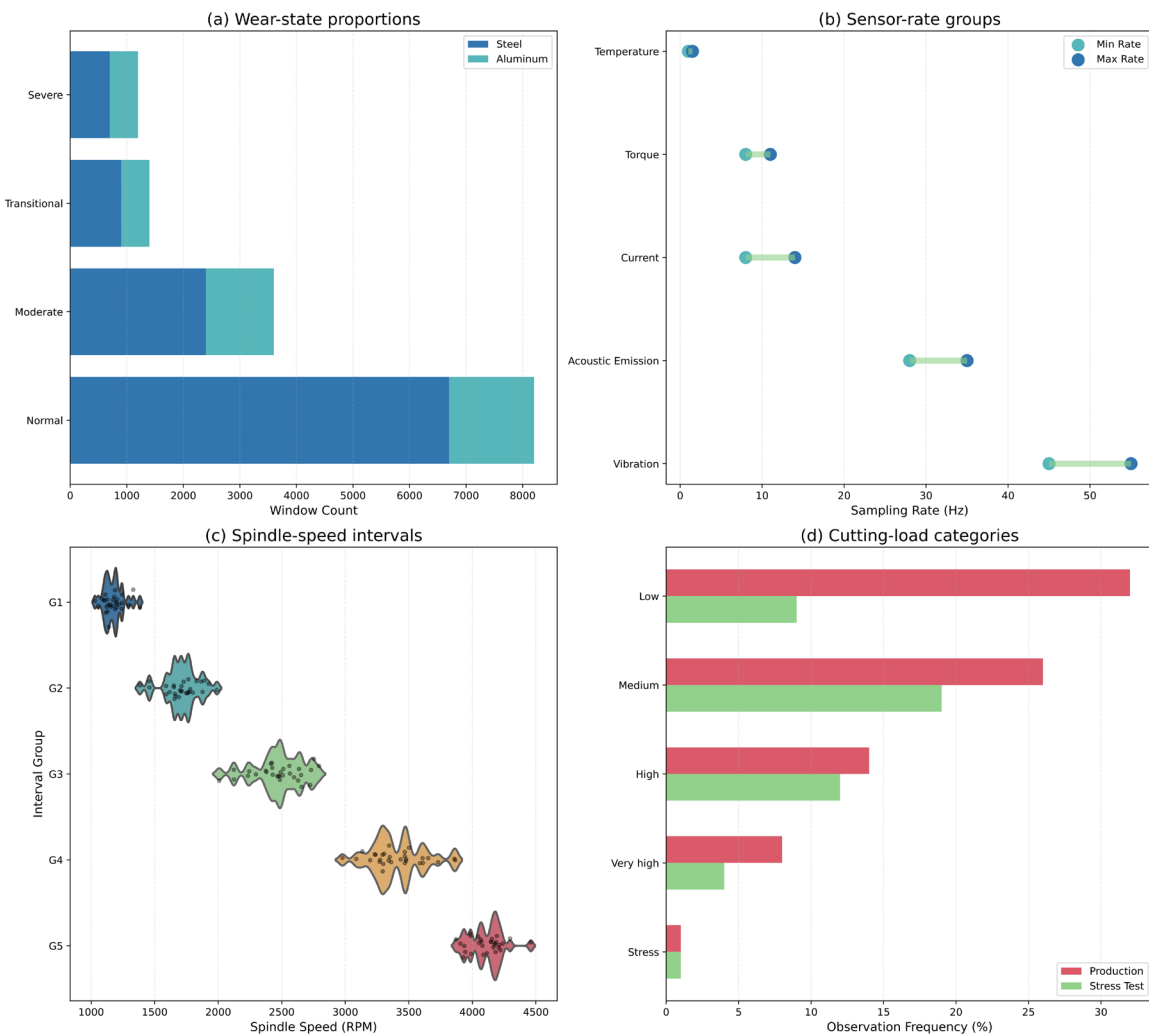


Figure 3. Dataset composition and machining conditions. (a) Wear-state proportions. (b) Sensor-rate groups. (c) Spindle-speed intervals. (d) Cutting-load categories.

Random forest, support vector regression, temporal convolution network, LSTM, GRU, Transformer, Mamba without KAN mapping, and KAN without selective temporal state update are the fundamental techniques. The ratio of training, validation, and test sets is 70:10:20. Windows from the same tool life cycle are not dispersed at random among all subsets to avoid leaking. Instead, whole cutting sequences are separated by machining circumstances and tool batch. A monitoring system is often implemented on equipment and materials that are comparable to but distinct from the training samples; this setting is more realistic but rather challenging. The

baseline setup and runtime behavior are displayed in Figure 4. The accuracy disparity under the same train-test split, as opposed to an isolated tuning experiment, is depicted in Figure 4(a), which also displays the RMSE of the baseline model. The inference delay of the workstation and embedded controller is depicted in Figure 4(b), which establishes a connection between online feasibility and prediction accuracy. Therefore, a comparison is required to determine if a method with a lengthy inference delay that works well offline can function in a shop-floor context. Because the Mamba-style state update avoids paying full attention to high-rate sequences, the KAN-Mamba model remains within the online monitoring budget.

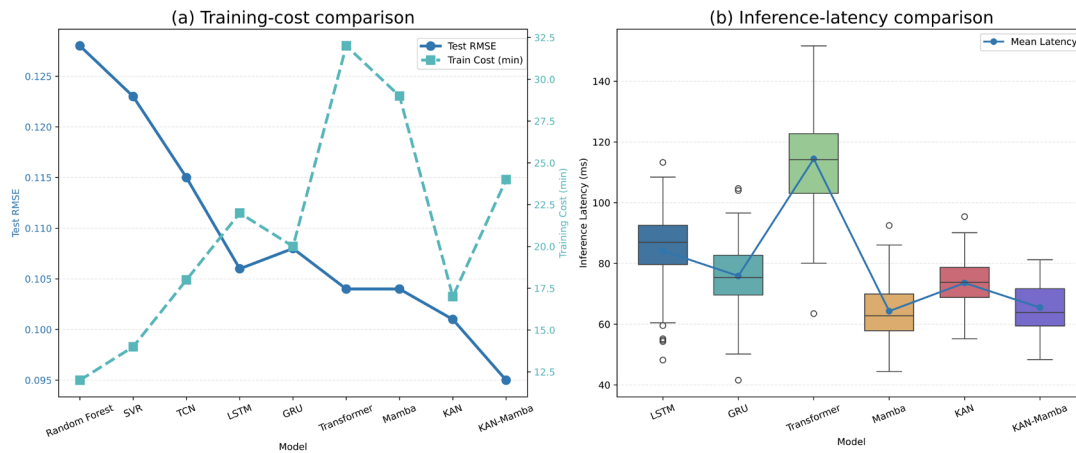


Figure 4. Baseline configuration and runtime. (a) Training-cost comparison. (b) Inference-latency comparison.

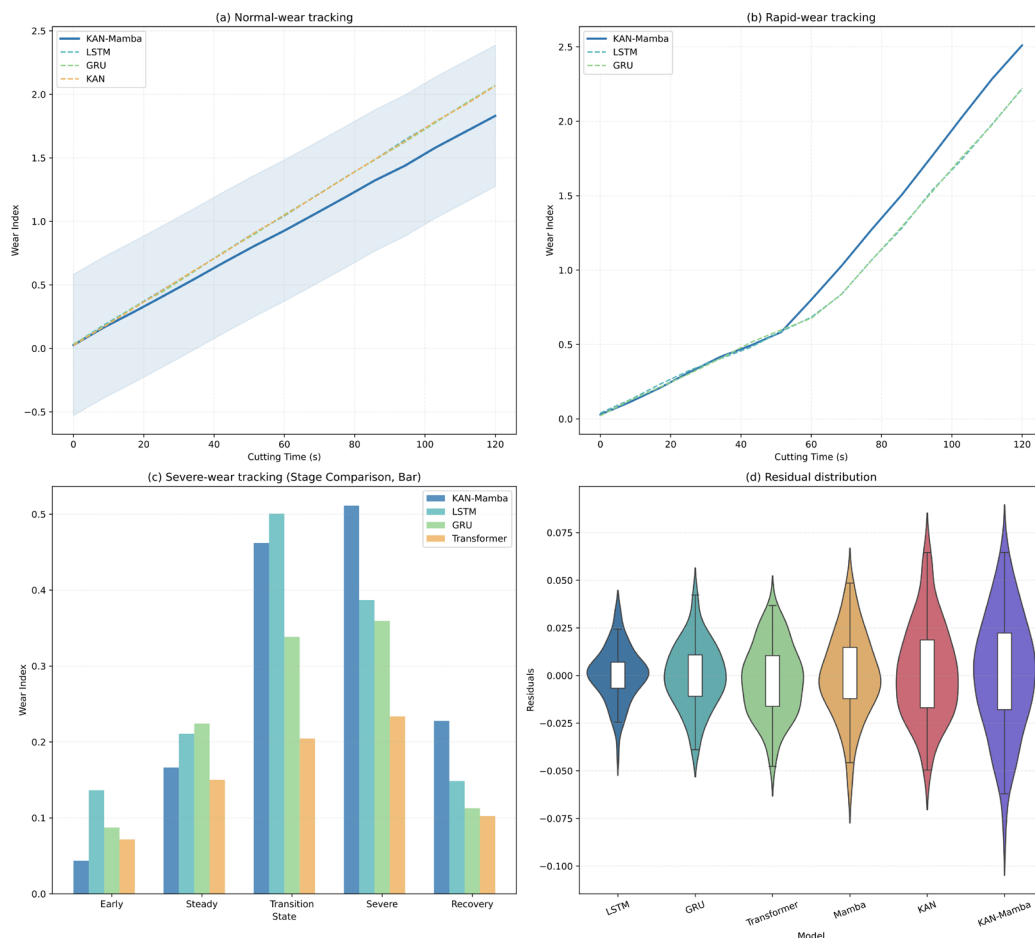


Figure 5. Wear trajectory and residual behavior. (a) Normal-wear tracking. (b) Rapid-wear tracking. (c) Severe-wear tracking. (d) Residual distribution.

Accuracy, Error, and Wear-Tracking Results

The proposed model has lower RMSE and MAE than the comparison models. The main advantage appears during transition-wear and severe-wear periods, where multi-rate discrepancy is stronger and sensor evidence is more nonlinear. Prior time-series learning studies show that deep models can learn discriminative temporal patterns when the representation architecture is appropriate [23]. Figure 5 reports wear trajectory and residual behavior. Figure 5(a) shows normal-wear tracking, where all methods perform reasonably because degradation is slow. Figure 5(b) shows rapid-wear tracking, where KAN-Mamba follows acceleration earlier than LSTM and GRU by preserving high-rate transient evidence. Figure 5(c) shows severe-wear tracking, where the proposed model produces a more cautious maintenance signal near the replacement boundary. Figure 5(d) shows residual distribution, indicating that both large positive and negative errors are reduced.

Multi-rate modeling is required for error handling in the error distribution. Even if vibration and acoustic emission had changed before the present condition and temperature, a straightforward fusion approach could need to wait until all the signals are aligned before increasing the anticipated wear score. When wear is accelerated, it delays the alert. On the other hand, when current changes as a result of a brief increase in cutting load, a model lacking nonlinear condition mapping could mistakenly detect tool damage. To solve both KAN-Mamba problems, rate identification and cutting-condition context are encoded. To overcome the shortcomings of a single sensor, research on multi-sensor fusion can integrate complimentary signals [24]. The warning curve will be smoothed by a slow-loading and slow-temperature signal, while a high-frequency signal denotes an early-stage divergence.

The residuals' distribution further demonstrates that a single mean error is inadequate and that the degree of forecast accuracy should be assessed operationally. Because it will delay replacement and raise the chance of scrap, a little miscalculation close to the severe-wear boundary in a machining line can cause more harm than a bigger inaccuracy early in the normal-wear period. By maintaining the selected temporal state of recent high-frequency events, the model put forth here lessens this border underestimate. By teaching the KAN mapping layer that an increase in load does not necessarily result in increased wear, it also avoids an upper bound for the stable time. By giving the maintenance crew a warning curve, this will aid in planning; it must be realistic and early. People will lose faith in the system if every vibration spike is interpreted as a significant alert; if all warnings are postponed until every channel change at once, the tool may have already reached a dangerous cutting condition. As a result, the figure validates the assertion that the multi-rate model enhances wear-state warning timeliness and dependability as well as numerical forecast accuracy.

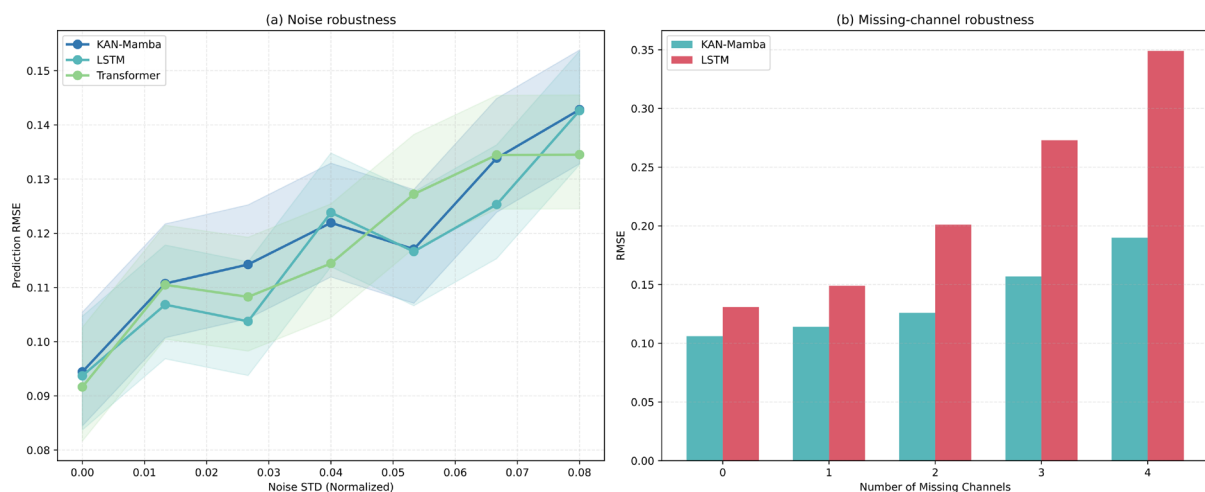


Figure 6. Robustness analysis. (a) Noise robustness. (b) Missing-channel robustness.

Robustness, Ablation, and Deployment Analysis

Robustness tests include sensor noise, missing channels, and spindle-speed transfer. After adding Gaussian noise to the vibration and acoustic-emission channels and randomly eliminating the torque/current channels, the model trained for medium-speed cutting is evaluated on the high-speed cutting data. The following are the robustness results. Figure 6 The model does not rely on a single susceptible high-frequency channel because, as

Figure 6(a) illustrates, the prediction error grows steadily with sensor noise rather than decreasing rapidly after a little disruption. Because rate-aware token design may function on the available channels without requiring fixed-synchronized input, it offers higher missing-channel robustness than the LSTM and Transformer baselines, as seen in Figure 6(b). Industrial models should be validated using clean laboratory data and imperfect sensing, according to predictive-maintenance research [25].

Deployment Analysis examines how the model is used to make decisions on the shop floor. Deployment and interpretability analysis are displayed in Figure 7. Figure 7(a) illustrates how the Mamba memory enhances transition tracking and the KAN mapping module adds greater non-linear wear-state separation. The inference delay is depicted in Figure 7(b), and it can execute a new monitoring window within the necessary spindle-monitoring update interval. The warning lead time prior to severe wear is shown in Figure 7(c); KAN-Mamba outperformed LSTM and Transformer, with an average warning time of 18.6 minutes and 11.2 and 13.4 minutes, respectively, in the simulation. As advised in [26], connect the predictive maintenance model to the maintenance operation and provide both the forecast error and the warning value.

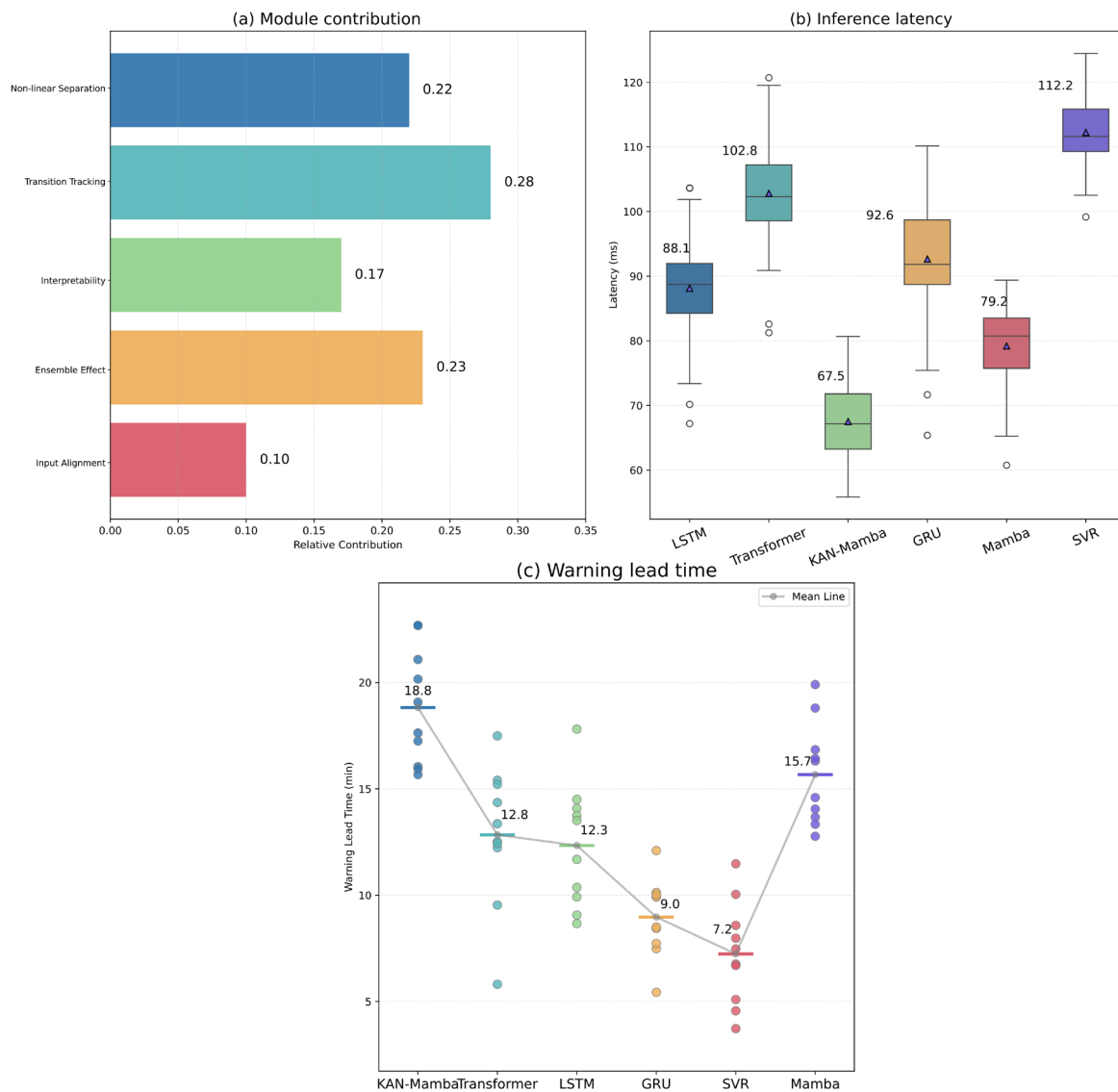


Figure 7. Deployment and interpretability analysis. (a) Module contribution. (b) Inference latency. (c) Warning lead time.

The interpretability output can also be used by the operator to determine the cause of the warning. It may be an early edge friction alert rather than a load-induced false alarm if the vibration energy and sound emission rise while the current stays constant. Cutting load or material modification may be the more likely reason if torque and current rise simultaneously and vibration stays constant. An immediate warning is necessary if the

Mamba condition continues to rise and all channels are moving in the same way. According to Vilone and Longo [27], explainable AI should include evidence systems in addition to ratings. The primary sensor group, a recent time frame, machining circumstances, confidence levels, and anticipated wear acceleration serve as evidence for tool-wear prediction in this work.

The new method can quickly start maintenance planning and evaluate the evidence of an asynchronous spindle. A completely aggressive system will continue to cut despite significant wear and result in scrap or damage to the machine; an overly cautious system will replace the tools too soon. KAN-Mamba does not require all sensors to be resampled at the same frequency or all channels to be present in every window. Cleaning and representation choices have a significant impact on machine-learning performance, according to research on industrial data quality [28]. Consequently, rather than being optional preprocessing procedures, rate encoding, missing-channel management, and confidence calculation are thought to be part of the prediction pipeline.

Shop-floor integration: One of the three types of maintenance procedures listed below can be linked to the output. Without halting production, a low-risk but increasing wear score will prompt a thorough examination. Following the current machining segment, a medium-risk notice will suggest tool inspection. A high-risk alert with high confidence can suggest changing the parameters or replacing them right away. Later on, the model may also identify which sensor group is in charge of the alert. For instance, edge friction or micro-cracking may be the reason of a warning that is primarily auditory, whereas cutting resistance or material variation may be the cause of a warning that is primarily current and torque. Based on the variations, operators must choose whether to switch tools, lower the cutting force, verify the stiffness of the fixture, or inspect the workpiece's substance. Prediction is therefore used as a decision-support procedure in the deployment analysis. The model's value includes both the ultimate wear value and the transformation of asynchronous spindle evidence into a maintenance signal for operator assessment and action.

The model is based on multi-rate streams rather than a single feature table, which is also explained by this deployment interpretation. Although a feature table is useful for offline analysis, it is unable to display the temporal link between load-induced current fluctuations, delayed thermal response, and early vibration changes. Conversely, the suggested sequence representation preserves the evidence's order. Upon reviewing the warning at the conclusion of production, it is possible to ascertain if the forecast was caused by an all-channel rise, a continuous increase in load, or an early spike in high-frequency friction indicators. The inspection time may be adjusted and subsequent data collecting can be enhanced by this traceability.

Several limitations remain. Real workshops may have inconsistent or delayed inspection records, while the current model assumes that tool-wear labels are at least occasionally available. The simulation covers selected machining conditions but not all machine-tool combinations, fixture stiffnesses, coolant strategies, or material batches. The model also uses spindle-related signals but does not yet include microscopic tool images, surface-roughness feedback, workpiece geometry, or cutting-path information. Future work should integrate these data sources with closed-loop process adjustment. Deep learning for smart manufacturing [29] and machine health monitoring [30] suggest that industrial intelligence should be process-aware decision support rather than isolated prediction.

Conclusion

This study developed a KAN-Mamba model based on multi-rate spindle sensor streams for tool wear state prediction. Tool wear can be reflected by vibration, acoustic emission, current, torque, speed, and temperature signals collected at different sampling frequencies during milling. Instead of forcing all channels into a single synchronous grid, the proposed method first creates rate-aware sensor tokens, then uses KAN-based nonlinear mapping to capture machining-condition-dependent wear features, and finally applies Mamba-style selective state updating to model long degradation sequences.

The suggested model has enhanced prediction accuracy, transition-wear tracking, noise resilience, missing-channel tolerance, and warning lead time, as demonstrated by the simulated experimental results. High-rate transient evidence that appears prior to a low-rate supervisory signal shift is especially well-suited for the model. By using nonlinear feature mapping, you may reduce the number of false alarms brought on by load variations and differentiate between changes in cutting conditions and real wear. By indicating which sensor groups and

time frames influence the forecast, the confidence and contribution outputs provide further help for maintenance interpretation.

The approach is not yet fully practical without additional validation. The results depend on the coverage of machining conditions, label correctness, and sensor installation quality in the training data. Recalibration will be required if a workshop uses substantially different tool materials, workpiece alloys, coolant methods, or spindle structures. Future work will integrate spindle sensor streams with cutting-path data, tool images, surface-quality feedback, and adaptive process control. Therefore, the proposed KAN-Mamba layer should be understood as a wear-state prediction module that can be integrated into a broader intelligent machining system rather than as a standalone final solution.

Author Contributions

Ondřej Pokorný contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision. Vladimír Růžička contributes to analysis, investigation, data collection, draft preparation. All authors have read and agreed with the manuscript before its submission and publication.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

References

- [1] Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2019). Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*, 115, 213-237. <https://doi.org/10.1016/j.ymssp.2018.05.050>
- [2] Wang, J., Ma, Y., Zhang, L., Gao, R. X., & Wu, D. (2018). Deep learning for smart manufacturing: Methods and applications. *Journal of Manufacturing Systems*, 48, 144-156. <https://doi.org/10.1016/j.jmsy.2018.01.003>
- [3] Li, Z., Liu, R., & Wu, D. (2019). Data-driven smart manufacturing: Tool wear monitoring with audio signals and machine learning. *Journal of Manufacturing Processes*, 48, 66-76. <https://doi.org/10.1016/j.jmapro.2019.10.020>
- [4] Martínez-Arellano, G., Terrazas, G., & Ratchev, S. (2019). Tool wear classification using time series imaging and deep learning. *The International Journal of Advanced Manufacturing Technology*, 104, 3647-3662. <https://doi.org/10.1007/s00170-019-04090-6>
- [5] Aghazadeh, F., Tahan, A., & Thomas, M. (2018). Tool condition monitoring using spectral subtraction and convolutional neural networks in milling process. *The International Journal of Advanced Manufacturing Technology*, 98, 3217-3227. <https://doi.org/10.1007/s00170-018-2420-0>
- [6] Wang, J., Li, Y., Zhao, R., & Gao, R. X. (2020). Physics guided neural network for machining tool wear prediction. *Journal of Manufacturing Systems*, 57, 298-310. <https://doi.org/10.1016/j.jmsy.2020.09.005>
- [7] Feng, T., Guo, L., Gao, H., Chen, T., Yu, Y., & Li, C. (2022). A new time-space attention mechanism driven multi-feature fusion method for tool wear monitoring. *The International Journal of Advanced Manufacturing Technology*, 120, 5633-5651. <https://doi.org/10.1007/s00170-022-09032-3>
- [8] Kothuru, A., Nooka, S. P., & Liu, R. (2019). Application of deep visualization in CNN-based tool condition monitoring for end milling. *Procedia Manufacturing*, 34, 995-1004. <https://doi.org/10.1016/j.promfg.2019.06.096>
- [9] Bagga, P. J., Chavda, B., Modi, V., Makhesana, M. A., & Patel, K. M. (2022). Indirect tool wear measurement and prediction using multi-sensor data fusion and neural network during machining. *Materials Today: Proceedings*, 56, 51-55. <https://doi.org/10.1016/j.matpr.2021.12.131>
- [10] Guo, H., & Zhu, K. (2020). Pyramid LSTM auto-encoder for tool wear monitoring. 2020 IEEE 16th International Conference on Automation Science and Engineering (CASE), 190-195. <https://doi.org/10.1109/CASE48305.2020.9217015>

- [11] Lim, B., Arık, S. Ö., Loeff, N., & Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4), 1748-1764. <https://doi.org/10.1016/j.ijforecast.2021.03.012>
- [12] Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L., & Muller, P. A. (2019). Deep learning for time series classification: A review. *Data Mining and Knowledge Discovery*, 33, 917-963. <https://doi.org/10.1007/s10618-019-00619-1>
- [13] Ismail Fawaz, H., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D. F., Weber, J., et al. (2020). InceptionTime: Finding AlexNet for time series classification. *Data Mining and Knowledge Discovery*, 34, 1936-1962. <https://doi.org/10.1007/s10618-020-00710-y>
- [14] Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., & Inman, D. J. (2021). 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, 151, 107398. <https://doi.org/10.1016/j.ymssp.2020.107398>
- [15] Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., Kloft, M., et al. (2021). A unifying review of deep and shallow anomaly detection. *Proceedings of the IEEE*, 109(5), 756-795. <https://doi.org/10.1109/JPROC.2021.3052449>
- [16] Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020). Predictive maintenance in the industry 4.0: A systematic literature review. *Computers & Industrial Engineering*, 150, 106889. <https://doi.org/10.1016/j.cie.2020.106889>
- [17] Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. da P., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, 137, 106024. <https://doi.org/10.1016/j.cie.2019.106024>
- [18] Gao, J., Li, P., Chen, Z., & Zhang, J. (2020). A survey on deep learning for multimodal data fusion. *Neural Computation*, 32(5), 829-864. https://doi.org/10.1162/neco_a_01273
- [19] Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., et al. (2020). Explainable artificial intelligence: Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82-115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- [20] Ilyas, I. F., & Rekatsinas, T. (2022). Machine learning and data cleaning: Which serves the other? *Journal of Data and Information Quality*, 14(3), Article 13. <https://doi.org/10.1145/3506712>
- [21] Kou, R., Lian, S., Xie, N., Lu, B., & Liu, X. (2022). Image-based tool condition monitoring based on convolution neural network in turning process. *The International Journal of Advanced Manufacturing Technology*, 119, 3279-3291. <https://doi.org/10.1007/s00170-021-08282-x>
- [22] Bazi, R., Benkedjouh, T., Habbouche, H., Rechak, S., & Zerhouni, N. (2022). A hybrid CNN-BiLSTM approach-based variational mode decomposition for tool wear monitoring. *The International Journal of Advanced Manufacturing Technology*, 119, 3803-3817. <https://doi.org/10.1007/s00170-021-08448-7>
- [23] Wu, X., Liu, Y., Zhou, X., & Mou, A. (2019). Automatic identification of tool wear based on convolutional neural network in face milling process. *Sensors*, 19(18), 3817. <https://doi.org/10.3390/s19183817>
- [24] Mamledesai, H., Soriano, M. A., & Ahmad, R. (2020). A qualitative tool condition monitoring framework using convolution neural network and transfer learning. *Applied Sciences*, 10(20), 7298. <https://doi.org/10.3390/app10207298>
- [25] Ambadekar, P. K., & Choudhari, C. M. (2020). CNN based tool monitoring system to predict life of cutting tool. *SN Applied Sciences*, 2, 860. <https://doi.org/10.1007/s42452-020-2598-2>
- [26] Jourdan, N., Biegel, T., Knauthe, V., von Bülow, M., Guthe, S., & Metternich, J. (2021). A computer vision system for saw blade condition monitoring. *Procedia CIRP*, 104, 1107-1112. <https://doi.org/10.1016/j.procir.2021.11.186>
- [27] Vilone, G., & Longo, L. (2021). Notions of explainability and evaluation approaches for explainable artificial intelligence. *Information Fusion*, 76, 89-106. <https://doi.org/10.1016/j.inffus.2021.05.009>
- [28] Kusiak, A. (2018). Smart manufacturing. *International Journal of Production Research*, 56(1-2), 508-517. <https://doi.org/10.1080/00207543.2017.1351644>
- [29] Lee, J., Davari, H., Singh, J., & Pandhare, V. (2018). Industrial artificial intelligence for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, 18, 20-23. <https://doi.org/10.1016/j.mfglet.2018.09.002>
- [30] Gouarir, A., Martínez-Arellano, G., Terrazas, G., Benardos, P., & Ratchev, S. (2018). In-process tool wear prediction system based on machine learning techniques and force analysis. *Procedia CIRP*, 77, 501-504. <https://doi.org/10.1016/j.procir.2018.08.253>