ISSN 2616-5767 Vol.9, Issue 1: 47-52, DOI: 10.25236/AJETS.2026.090106

Research on Predictive Maintenance of CNC Machine Tools Based on Deep Learning

Jing Pu

Xihua University, Chengdu, Sichuan, 610039, China

Abstract: *Predictive maintenance (PdM) represents a transformative approach in modern manufacturing, aiming to forecast equipment failures through data analysis rather than relying on scheduled or reactive maintenance. This research presents a comprehensive deep learning-based framework for the predictive maintenance of Computer Numerical Control (CNC) machine tools, which are critical and costly assets in precision manufacturing. The proposed system utilizes a multi-sensor data fusion strategy, acquiring real-time operational data including vibration, acoustic emission, spindle current, and temperature. A hybrid deep learning model is developed, integrating Convolutional Neural Networks (CNNs) for automatic feature extraction from high-dimensional sensor signals and Long Short-Term Memory (LSTM) networks to capture temporal dependencies and degradation trends. The model is trained on historical run-to-failure data to learn the complex mapping between multi-modal sensor inputs and the Remaining Useful Life (RUL) of critical components such as spindle bearings and ball screws. Experimental validation is conducted on a three-axis CNC milling machine under controlled operational loads. The results demonstrate that the proposed CNN-LSTM model achieves superior predictive accuracy compared to traditional machine learning benchmarks like Support Vector Regression and standalone neural networks. The system successfully identifies incipient fault conditions with a high degree of precision, providing early warnings significantly ahead of functional failure. This capability enables optimal maintenance scheduling, minimizes unplanned downtime, reduces maintenance costs, and extends the operational lifespan of CNC machinery. The study confirms the significant potential of deep learning in enhancing the intelligence and reliability of industrial predictive maintenance systems.*

Keywords: *Predictive Maintenance; Deep Learning; CNC Machine Tools; Convolutional Neural Network; Long Short-Term Memory; Industrial Artificial Intelligence; Prognostics and Health Management*

1. Introduction

The manufacturing industry is undergoing a profound transformation under the paradigm of Industry 4.0, characterized by cyber-physical systems, the Internet of Things (IoT), and data-driven decision-making. Within this context, maintenance strategies for industrial equipment have evolved from traditional reactive and preventive models towards more advanced, condition-based, and predictive approaches [1]. Predictive maintenance stands out as a pivotal technology, as it leverages the analysis of equipment condition data to predict when a failure might occur, thereby allowing maintenance to be performed just in time. This strategy is particularly crucial for Computer Numerical Control (CNC) machine tools, which are the backbone of discrete manufacturing for aerospace, automotive, and mold-making industries. The high capital investment, critical role in production lines, and severe financial consequences of unscheduled downtime associated with CNC machines make their reliability and availability paramount concerns. Unexpected failures of key components, such as the spindle, guideways, or ball screws, can lead to catastrophic scrap loss, prolonged production stoppages, and costly emergency repairs.

Traditional time-based preventive maintenance, while reducing some failure risks, often leads to unnecessary part replacements and resource wastage, as it does not account for the actual health condition of the equipment [2]. Condition-based maintenance, which relies on monitoring specific parameters like vibration levels, marks an improvement but typically depends on expert-defined thresholds and may lack prognostic capability. The advent of pervasive sensor technology and big data analytics has unlocked the potential for truly predictive maintenance. However, the complex, nonlinear, and temporal nature of machine tool degradation poses a significant challenge for conventional

ISSN 2616-5767 Vol.9, Issue 1: 47-52, DOI: 10.25236/AJETS.2026.090106

statistical and shallow machine learning models [3]. These models often require extensive manual feature engineering and struggle to model long-term dependencies in time-series data [4].

Deep learning, a subset of machine learning characterized by deep neural networks with multiple processing layers, offers a powerful solution to these challenges. Its ability to automatically learn hierarchical feature representations from raw, high-dimensional data makes it exceptionally suited for analyzing complex sensor signals. Architectures like Convolutional Neural Networks (CNNs) excel at extracting spatial features from signal data, while Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, are designed to model temporal sequences and long-range dependencies. This research posits that a synergistic deep learning model, integrating CNN and LSTM, can effectively capture both the spatial patterns in multi-sensor snapshots and the temporal evolution of the degradation process in CNC machines [5]. The primary objective of this study is to develop, implement, and validate such a hybrid deep learning framework for accurately predicting the Remaining Useful Life (RUL) of critical CNC machine tool components. By transitioning from a diagnosis of current faults to a prognosis of future health, this research aims to contribute to the development of more intelligent, resilient, and cost-effective manufacturing systems.

2. Literature Review and Theoretical Framework

The evolution of predictive maintenance methodologies reflects the broader technological advancements in data acquisition and analysis techniques. Early approaches predominantly relied on statistical process control and time-series analysis methods, such as autoregressive integrated moving average models, which provided foundational frameworks but were limited in handling the complex, nonlinear degradation patterns characteristic of sophisticated machinery like CNC machine tools [6]. The introduction of traditional machine learning algorithms, including support vector machines and random forests, marked a significant step forward, enabling more sophisticated pattern recognition from condition monitoring data [7]. These methods, however, continued to depend heavily on expert knowledge for feature engineering and selection, creating a bottleneck in system development and limiting adaptability to different failure modes or machine configurations.

The emergence of deep learning has fundamentally altered this landscape by introducing models capable of automatic feature extraction from raw, high-dimensional data. Within the specific domain of industrial prognostics, several architectural paradigms have shown particular promise [8]. Convolutional Neural Networks have demonstrated exceptional capability in processing vibration spectrograms and other two-dimensional representations of sensor data, effectively identifying spatial patterns indicative of mechanical faults. Simultaneously, Recurrent Neural Networks, especially their variant Long Short-Term Memory networks, have proven adept at modeling temporal sequences, making them suitable for analyzing time-series sensor data where the order and historical context of measurements contain crucial information about degradation progression [9]. The fusion of these architectures represents a logical advancement, aiming to capture both the spatial features within individual data samples and the temporal dependencies across sequential observations.

In parallel with algorithmic developments, significant research attention has been directed toward multi-sensor data fusion strategies [10]. The underlying principle is that different physical phenomena associated with mechanical degradation—such as vibration, acoustic emission, thermal changes, and power consumption—provide complementary information. Data-level fusion involves the direct combination of raw signals, while feature-level fusion combines extracted characteristics, and decision-level fusion aggregates conclusions from multiple single-sensor models. The optimal fusion strategy remains an active area of investigation, balancing computational complexity with informational gain [11]. Furthermore, the formulation of the prediction target itself is critical. While binary classification (fault/no fault) and multi-class fault diagnosis are valuable, the regression-based prediction of Remaining Useful Life represents a more challenging but ultimately more useful paradigm for proactive maintenance planning, as it quantifies the time horizon for intervention [12].

Despite these advancements, a review of existing literature reveals several persistent gaps. Many studies validate models on standardized public datasets, such as those from bearing test rigs, which may not fully capture the complex operational environment and variable loading conditions of actual CNC machining centers. There is also a tendency to focus on a single component or failure mode, whereas industrial applications require systems capable of addressing multiple potential points of failure. Additionally, the practical challenges of model deployment, including computational resource requirements on the factory floor, real-time inference latency, and model adaptability to machine-to-

ISSN 2616-5767 Vol.9, Issue 1: 47-52, DOI: 10.25236/AJETS.2026.090106

machine variations, are often underexplored in theoretical research. This study addresses these gaps by developing and validating a comprehensive hybrid deep learning model on a fully operational CNC machine tool under realistic, variable loading conditions. It implements a multi-modal sensor fusion approach at the feature level and frames the output as a continuous RUL prediction, thereby bridging the divide between algorithmic innovation and practical industrial implementation. The theoretical framework underpinning this work integrates signal processing, deep neural network theory, and reliability engineering to create a cohesive.

3. Experimental Methods

The experimental methodology was designed to simulate real-world operational conditions and collect a comprehensive dataset for developing and validating the predictive maintenance model. The core of the experimental setup was a standard three-axis vertical CNC milling machine. To capture a holistic view of the machine's health, a multi-sensor monitoring system was deployed. Four types of sensors were strategically installed: tri-axial vibration accelerometers on the spindle housing and the Y-axis servo motor to capture mechanical oscillations and imbalances; an acoustic emission sensor near the spindle to detect high-frequency stress waves generated by incipient cracks or bearing defects; a current transducer on the spindle drive to monitor load variations and electrical anomalies; and thermocouples on the spindle bearing housings and the ball screw nuts to track temperature rises due to friction and wear. All sensors were calibrated and connected to a high-speed data acquisition system capable of synchronous sampling.

The experiment focused on accelerated degradation of the spindle bearing, a common critical point of failure. A batch of bearings was seeded with minor, controlled defects in the inner raceway to initiate the failure process. The CNC machine was then operated under a programmed cyclic load profile, simulating realistic machining cycles involving varying cutting forces and speeds. Data from all sensors was continuously collected at a predefined sampling rate throughout the entire run-to-failure lifetime of multiple bearing sets. This process was repeated to generate a sufficient dataset encompassing the complete degradation trajectory from healthy state to ultimate failure. The collected raw time-series data constituted the primary input for the model.

The data preprocessing pipeline involved several crucial steps. First, data synchronization and alignment were performed to ensure temporal coherence across all sensor channels. Subsequently, noise filtering was applied using wavelet transform techniques to remove high-frequency electrical noise and low-frequency drift without distorting the underlying fault signatures. The continuous data stream was then segmented into fixed-length time windows, each representing a snapshot of the machine's multi-sensor state at a given time. The corresponding label for each window was its Remaining Useful Life (RUL), defined as the number of operational hours left until a predefined failure threshold (e.g., a sharp increase in overall vibration amplitude) was reached. The RUL label was formulated as a continuous value, transforming the problem into a regression task.

The architecture of the proposed hybrid deep learning model, termed the CNN-LSTM network, was constructed as follows. The first stage consisted of multiple parallel one-dimensional convolutional layers, each processing input from a specific sensor modality or channel. These CNN layers were designed to perform automatic local feature extraction, identifying patterns like specific frequency components in vibration signals or transient spikes in acoustic emission. The outputs from these parallel CNN streams were then flattened and concatenated into a unified feature vector representing the spatial characteristics within each time window. This fused feature vector was then fed into a stacked LSTM network. The LSTM layers, with their memory cells and gating mechanisms, were responsible for learning the temporal dynamics and long-term dependencies across sequential time windows, effectively modeling how the extracted features evolved as the bearing degraded. The final layers consisted of fully connected dense layers that nonlinearly combined the high-level temporal features to output a single RUL prediction. For comparison, baseline models including a Support Vector Regression (SVR) model with handcrafted features (like root mean square, kurtosis, and spectral centroids), a standard Multi-Layer Perceptron (MLP), and standalone CNN and LSTM models were also developed and trained on the same dataset. All models were implemented using the TensorFlow framework and trained using the Adam optimizer, with mean squared error as the loss function. The dataset was split into training, validation, and testing sets in a chronological manner to prevent data leakage and ensure a realistic evaluation of prognostic performance.

ISSN 2616-5767 Vol.9, Issue 1: 47-52, DOI: 10.25236/AJETS.2026.090106

4. Results

The performance of the predictive models was rigorously evaluated on the unseen testing dataset. The primary evaluation metric was the Root Mean Square Error (RMSE) between the predicted RUL and the actual RUL, with lower values indicating higher accuracy. Secondary metrics included the Mean Absolute Error (MAE) and the Score function, which penalizes late predictions more heavily than early ones, reflecting the practical preference for conservative warnings. The quantitative results are summarized in the following tables, which compare the performance of the baseline models against the proposed CNN-LSTM hybrid.

Table 1. Overall Predictive Performance Metrics for Different Models

Model	RMSE (hours)	MAE (hours)	Score Function
Support Vector Regression (SVR)	18.74	15.22	245.6
Multi-Layer Perceptron (MLP)	14.56	11.87	178.3
Standalone CNN Model	12.31	9.45	132.1
Standalone LSTM Model	11.89	9.12	121.8
Proposed CNN-LSTM Hybrid	8.67	6.84	89.5

Table 1 clearly demonstrates the superiority of the deep learning approaches over the traditional SVR model. The hybrid CNN-LSTM model achieved the best performance across all metrics, with an RMSE of 8.67 hours, representing a significant improvement of over 25% compared to the best standalone deep model (LSTM) and more than 50% compared to the SVR baseline. This substantial reduction in error underscores the effectiveness of combining spatial feature extraction with temporal sequence modeling.

Table 2. Model Performance Across Different Phases of Bearing Degradation

Degradation Phase	CNN-LSTM RMSE (hours)	LSTM RMSE (hours)	CNN RMSE (hours)
Early (Healthy to Incipient Fault)	12.45	15.88	16.72
Middle (Progressive Degradation)	7.23	9.45	10.91
Late (Severe Fault to Failure)	6.32	8.12	9.05

A more granular analysis of performance across the degradation timeline, as shown in Table 2, reveals interesting insights. All models performed with less accuracy during the early phase, where fault signatures are subtle and buried in noise. However, the CNN-LSTM hybrid showed a relative advantage even in this challenging phase, likely due to the CNN's ability to detect weak spatial patterns. Its performance improved markedly during the middle and late phases as the fault features became more pronounced, and the LSTM component effectively leveraged the accumulated temporal history. This consistent accuracy across all phases is critical for reliable prognostics.

Table 3. Contribution Analysis of Different Sensor Modalities to the CNN-LSTM Model

Sensor Input Configuration	RMSE (hours)	Performance Change vs. Full Set
Vibration Only	11.54	33.10%
Acoustic Emission Only	13.27	53.10%
Spindle Current Only	15.89	83.30%
Temperature Only	21.45	147.40%
Vibration + Acoustic Emission	9.82	13.30%

To understand the value of multi-sensor data fusion, an ablation study was conducted, and the results are presented in Table 3. Using any single sensor modality led to significantly higher prediction errors. Temperature data alone yielded the poorest results, suggesting it is a lagging indicator. Vibration and acoustic emission data were the most informative individually. Crucially, the combination of vibration and acoustic emission already provided good results, but the inclusion of all four sensor types (the full set) yielded the lowest RMSE. This demonstrates that while certain sensors are primary indicators, the fusion of complementary information from diverse physical phenomena (mechanical vibration, stress waves, electrical load, thermal state) provides a more robust and accurate health assessment, allowing the model to cross-validate signals and improve prediction confidence.

Graphical results further illustrated the model's effectiveness. The trajectory of the predicted RUL from the CNN-LSTM model closely followed the actual RUL curve for multiple test bearings, maintaining a narrow and consistent prediction horizon. In contrast, the predictions from baseline models showed larger deviations, especially during transition periods between degradation phases. The model successfully generated early warnings, typically 30-50 operational hours before the final failure threshold, which is a practically useful lead time for scheduling maintenance interventions.

5. Discussion

The results unequivocally support the central hypothesis that a deep learning-based approach, specifically a hybrid CNN-LSTM architecture, can significantly enhance the accuracy of predictive maintenance for CNC machine tools. The marked performance gap between the proposed model and traditional methods like SVR highlights the limitations of manual feature engineering and linear models in capturing the complex, nonlinear degradation dynamics of mechanical systems. The deep learning models' ability to learn directly from raw or minimally processed sensor data is a major advantage, reducing dependency on domain expertise for feature selection and making the system more adaptable to different machine types or failure modes.

The superior performance of the CNN-LSTM hybrid over both the standalone CNN and LSTM models underscores the complementary strengths of these architectures. The CNN component acts as an intelligent, adaptive feature extractor. It discerns local patterns within a time window—such as the specific frequency band where bearing defect frequencies emerge from the vibration spectrum or the characteristic burst pattern in acoustic emission signals. This automated feature learning is more comprehensive and potentially more sensitive than a fixed set of handcrafted features. The LSTM component then provides the essential temporal context. Bearing degradation is not a series of independent events but a continuous process where the current state is intrinsically linked to its history. The LSTM's memory cells effectively track this progression, learning the rate of degradation and recognizing trajectories that lead to failure. This synergy allows the model to not just diagnose the current severity of a fault but to forecast its future evolution accurately.

The sensor fusion analysis offers critical practical insights. The poor performance of single-sensor models, particularly temperature, validates the need for a multi-sensor strategy. Different failure modes manifest differently across sensor modalities. For instance, a lubrication issue might cause a temperature rise before significant vibration changes, while a spall on a bearing raceway might generate a strong acoustic emission signal earlier than a current anomaly. A fusion-based system is therefore more robust and can provide early warnings for a wider variety of potential faults. It also adds redundancy, making the system less vulnerable to the failure of a single sensor.

However, the research also reveals challenges, primarily during the early degradation phase. The higher prediction errors in this phase indicate the difficulty of detecting incipient faults. Future work could explore more sophisticated data augmentation techniques, semi-supervised learning to leverage unlabeled data from normal operations, or the integration of transfer learning from simulated data or other machines to improve early-phase sensitivity. Another consideration is the computational cost and model interpretability. While the model is highly accurate, its "black-box" nature can be a barrier to adoption in safety-critical applications where engineers need to understand the rationale behind a prediction. Developing methods for explaining the model's decisions, such as attention mechanisms or saliency maps highlighting which sensor or time period most influenced a prediction, would be a valuable direction.

From an implementation perspective, the success of this model paves the way for its deployment in an edge-cloud computing framework. The feature extraction and initial processing could be handled by edge devices on the shop floor for low-latency monitoring, while the complex LSTM-based prognosis could be performed on a central cloud server that aggregates data from multiple machines, enabling fleet-wide health management and comparative analytics.

6. Conclusion

This research has successfully developed and validated a deep learning-driven framework for the predictive maintenance of CNC machine tools. By formulating the problem as a Remaining Useful Life (RUL) regression task and employing a hybrid CNN-LSTM neural network architecture, the study demonstrates a significant advancement over conventional predictive maintenance methodologies. The model effectively leverages multi-sensor data fusion, automatically learning discriminative spatial-temporal features from raw vibration, acoustic emission, current, and temperature signals to accurately forecast the degradation trajectory of critical components like spindle bearings. Experimental results on a real CNC milling platform confirm that the proposed model achieves a substantial improvement in prediction accuracy, with lower root mean square error and more reliable early warnings compared to

ISSN 2616-5767 Vol.9, Issue 1: 47-52, DOI: 10.25236/AJETS.2026.090106

traditional machine learning and standalone deep learning models. The analysis further elucidates the critical importance of integrating complementary sensor data to build a robust prognostic system.

The implications of this work are substantial for the manufacturing industry. The ability to accurately predict failures with a sufficient lead time enables a true shift from schedule-based or reactive maintenance to a condition-based, predictive strategy. This translates directly into tangible benefits: minimization of unplanned and costly production stoppages, optimization of maintenance resources by performing interventions only when needed, reduction in spare parts inventory through better planning, prevention of secondary damage from catastrophic failures, and overall extension of machine tool service life. While challenges remain in further improving early fault detection and enhancing model interpretability, the findings firmly establish deep learning as a powerful and essential tool for building the intelligent, self-aware, and highly available manufacturing systems demanded by Industry 4.0. Future work will focus on extending the framework to other critical components, testing its generalizability across different machine models, and integrating it into a full-scale industrial digital twin for real-time lifecycle management.

References

[1] Li, A., Smith, B., & Johnson, C. Advancements in Predictive Maintenance: A Comprehensive Review [J]. Journal of Manufacturing Technology, 2021, 15(3): 45–62.

[2] Zhang, Y., Wang, Q., & Chen, L. Impact of Predictive Maintenance on Manufacturing Efficiency [J]. International Journal of Industrial Engineering, 2022, 28(1): 17–34.

[3] Gupta, R., Patel, S., & Singh, M. Enhancing Predictive Maintenance through Real time IoT Sensor Data Integration [J]. Journal of Industrial Internet of Things, 2021, 12(4): 112–128.

[4] Chen, H., Zhang, L., & Li, Q. Deep Learning Applications in Predictive Maintenance: Unveiling Subtle Patterns for Early Equipment Failure Detection [J]. Journal of Machine Learning Research, 2021, 18(3): 102–118.

[5] Wang, Y., & Liu, M. Improving Prediction Accuracy in Predictive Maintenance: A Comprehensive Study on Ensemble Methods [J]. Expert Systems with Applications, 2022, 29(4): 215–230.

[6] Li, Y., Wang, Y., Xu, H., & Hu, S. J. Cloud manufacturing: A new manufacturing paradigm [J]. Enterprise Information Systems, 2017, 11(5): 779–810.

[7] Lee, J., Kao, H. A., & Yang, S. Service innovation and smart analytics for Industry 4.0 and big data environment [J]. Procedia CIRP, 2014, 16: 3–8.

[8] Wang, D., Chen, G., Lu, C., & Wang, H. Big data analytics in predictive maintenance: A review [J]. IEEE Access, 2018, 6: 69002–69016.

[9] Ribeiro, A., Santos, T., Gomes, T., & Azevedo, A. Machine learning in predictive maintenance: A comprehensive survey [J]. Computers in Industry, 2019, 109: 184–206.

[10] Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. Digital twin-driven product design, manufacturing and service with big data [J]. The International Journal of Advanced Manufacturing Technology, 2018, 94(9-12): 3563–3576.

[11] Li, W., Miao, Q., Jiao, J., & Wang, Y. A survey on big data architectures and machine learning algorithms in smart city [J]. Future Generation Computer Systems, 2018, 89: 641–648.

[12] Duan, F., Wang, Y., & Wang, W. Machine learning models in manufacturing [J]. Journal of Manufacturing Systems, 2018, 48: 13–25.