



Gear Fault Detection Based on Convolutional Neural Networks and Support Vector Machines



Yeqi Fei^{1,2*}, Ziyang Meng¹, Tianxi Chen³, Pengfei Zhao⁴

¹ School of Intelligent Manufacturing, Nanjing University of Science and Technology ZiJin College, 210046 Nanjing, China

² College of Mechanical and Electronic Engineering, Nanjing Forestry University, 210037 Nanjing, China

³ Taicang Xinhai Port Development Co., Ltd, 215427 Suzhou, China

⁴ Suzhou Allink Technology Co., Ltd., 215000 Suzhou, China

* Correspondence: Fei Yeqi (feiyeqi@njfu.edu.cn)

Received: 11-18-2024

Revised: 12-21-2024

Accepted: 12-25-2024

Citation: Y. Q. Fei, Z. Y. Meng, T. X. Chen, and P. F. Zhao, "Gear fault detection based on convolutional neural networks and support vector machines," *Int J. Knowl. Innov Stud.*, vol. 2, no. 4, pp. 259–270, 2024. <https://doi.org/10.56578/ijkis020405>.



© 2024 by the author(s). Published by Acadlore Publishing Services Limited, Hong Kong. This article is available for free download and can be reused and cited, provided that the original published version is credited, under the CC BY 4.0 license.

Abstract: As a critical component of mechanical transmission systems, gears play a vital role in ensuring industrial production runs smoothly. Undetected gear failures can lead to mechanical breakdowns, production interruptions, and even safety hazards. Therefore, an efficient gear fault detection method is essential for maintaining industrial continuity and safety. This paper proposes a hybrid model that integrates convolutional neural networks (CNN) and support vector machines (SVM) for gear fault detection. The model leverages CNNs to automatically extract key features from vibration signals, while SVMs enhance classification accuracy, resulting in a high-precision fault diagnosis system. On a publicly available gear fault dataset, the proposed model achieved an impressive accuracy of 0.9922, significantly outperforming single-classifier models. Moreover, the model exhibits a short training time, demonstrating its computational efficiency. This research provides an effective and automated approach to gear fault detection, offering significant potential for industrial applications.

Keywords: Deep learning; Gear fault detection; Convolutional neural networks (CNN); Support vector machines (SVM); One-dimensional signal processing; Fault diagnosis

1 Introduction

Gears, as core components of mechanical transmission systems, play an extremely important role in industrial production and daily life. Gears are widely used in fields such as automobiles, railways, helicopters, and construction machinery [1, 2]. In key equipment such as transmissions, the stable operation of gears is directly related to the efficiency and reliability of the entire gear transmission system. However, due to continuous operation and variable working environments, gears may fail due to wear, fractures, fatigue, and other reasons. If these failures are not detected and handled in time, they are highly likely to lead to mechanical failures, production interruptions, and even serious safety accidents [3, 4]. Therefore, efficient and accurate fault detection and diagnosis of gears are of vital importance for ensuring the continuous operation of industrial production and production safety.

In recent years, the number of studies using deep learning algorithms for gear fault detection has increased significantly. Traditional fault detection methods mainly rely on vibration analysis, acoustic analysis, and temperature monitoring techniques. These methods have a certain ability to identify and diagnose gear faults under specific conditions. However, these methods are often limited in practical applications due to the complexity of signal processing, the limitations of feature extraction, and the lack of adaptability to complex and variable working conditions. For example, vibration analysis may struggle to extract weak or early fault signals, while acoustic analysis and temperature monitoring are susceptible to environmental noise and temperature fluctuations [5, 6]. As industrial equipment continues to develop towards high speed, heavy load, high precision, and high reliability, the complexity of its operating environment and the requirements for fault detection are also increasing. Against this background, the limitations of traditional fault detection methods have become increasingly prominent. These methods often fail to accurately capture early or minor fault characteristics when dealing with complex and variable

signals. Moreover, for equipment operating under high-speed and heavy-load conditions, the accuracy and real-time performance of diagnosis also face severe challenges [7–9]. To overcome these limitations, these advanced technologies use more sophisticated signal processing and feature extraction methods to more accurately identify and predict gear wear conditions, providing new ideas and methods for gear fault detection.

SVM [10, 11], with their efficient and precise characteristics, has become an indispensable tool in the field of gear fault detection. By deeply analyzing vibration signals, SVM can sensitively capture the nonlinear and non-stationary characteristics of gears, thereby achieving precise fault diagnosis. In recent years, the combination of SVM with advanced signal processing technologies has significantly improved the accuracy of complex mechanical fault diagnosis. To address the challenge of feature extraction from non-stationary vibration signals, researchers have proposed various innovative methods. Zhang et al. [12] proposed a gear fault diagnosis method based on Frequency-Modulated Empirical Mode Decomposition (FM-EMD) and SVM to identify the dynamic state and fault types of gearboxes. This method achieved near-frequency component separation in gearbox fault diagnosis, reaching a classification accuracy of over 90%. Medina et al. [13] described two algorithms for feature extraction from Poincaré maps, used for fault classification of vibration signals from rolling bearings and gearboxes under different loads and speeds. Using multi-class SVM for fault classification, the results showed that under variable load/speed conditions, the classification accuracy for gear and bearing faults reached 99.30% and 100.00%, respectively. Chen et al. [14] proposed an intelligent gear fault analysis method based on Fractional Wavelet Transform (FRWT) and SVM. This method used FRWT for noise elimination and SVM for training and recognition, achieving an accuracy of 96.7%. To address the challenges of small samples and high-dimensional data in industrial scenarios, derivative algorithms and parameter optimization strategies of SVM have been widely studied. Wang and Hu [15] compared the performance of SVM and LS-SVM in regression problems. The experimental results showed that LS-SVM maintained a classification accuracy similar to SVM when processing large-scale datasets while reducing computational resource consumption, achieving an accuracy of over 85.00%. Shen et al. [16] introduced an intelligent gear fault diagnosis model based on EMD and multi-class Transductive Support Vector Machine (TSVM). The experimental results showed that even when the number of unlabeled samples was 50 times that of labeled samples, the proposed method still achieved a test accuracy of 91.62%, significantly improving fault diagnosis performance under small sample conditions. Han et al. [17] proposed a gear fault feature extraction and diagnosis method combining EMD, Particle Swarm Optimization Support Vector Machine (PSO-SVM), and Fractal Box Dimension. This method effectively identified different types of gear faults under varying load excitations and achieved high-precision fault diagnosis. Kang et al. [18] developed a hybrid architecture based on Continuous Hidden Markov Model (CHMM) and SVM to predict equipment degradation trends through state probability estimation, achieving a prediction accuracy of over 95.00% in gearbox fault prediction. The adaptability of SVM in dynamic industrial environments, such as variable speed and variable load conditions, has been validated through the following techniques. In the application of bearing fault detection and diagnosis under different speeds, Pule et al. [19] explored the combined use of Principal Component Analysis (PCA) and SVM. Through experiments, the combination of PCA and SVM effectively diagnosed bearing faults under varying speeds, achieving a classification accuracy of 97.4%. Jiang et al. [20] studied a rotating machinery fault diagnosis method based on multi-sensor information fusion using SVM and time-domain features. In three case studies, it was found that the peak factor was the most sensitive feature for identifying gear defects, and all three cases exhibited high diagnostic accuracy, with an average accuracy of 92%.

Although SVM has achieved significant results in the field of gear fault detection, with the increasing complexity of mechanical systems and the diversification of industrial scenarios, a single SVM model gradually shows limitations in handling high-dimensional data and complex feature extraction. To address these challenges, deep learning technology, especially CNNs, provides new solutions for mechanical fault diagnosis. Traditional CNN is mainly used for two-dimensional image recognition, but one-dimensional convolutional neural networks (1D-CNN) have demonstrated strong performance in processing one-dimensional time series data, such as vibration signals. 1D-CNN can automatically extract key features from vibration signals and learn complex fault patterns through deep network structures, providing a more efficient and accurate method for mechanical fault diagnosis. In the field of mechanical fault diagnosis, 1D-CNN is used to extract features and classify faults from vibration signals of rotating machinery. Ma et al. [21] studied the problem of mechanical fault diagnosis based on compressed sensing and CNN. In experiments on planetary gearbox and bearing datasets, they used a 1D-CNN model to directly process compressed signals and performed feature extraction and fault classification using an improved LeNet-5 model. The results showed that this method achieved over 90% diagnostic accuracy on compressed signals with relatively short computation time. Wang et al. [22] proposed a method based on multi-sensor information fusion and 1D-CNN. The results showed that compared with traditional 1D or 2D-CNN and other fault classification methods, this model has a simpler structure, lower computational complexity, and an average accuracy of 99.83%. Peng et al. [23] proposed a deep CNN based on 1D residual blocks (Der-1DCNN). The diagnostic performance of this method was superior to four other advanced methods under different noise environments, with an accuracy rate above 95%, and even in a strong noise environment with an SNR of -16dB, the accuracy still reached 89.7%.

This method effectively improved the network's learning ability and noise resistance performance. Eren et al. [24] studied the problem of bearing fault diagnosis and proposed a method using a compact adaptive 1D-CNN classifier. This method learns optimal features directly from raw time-series sensor data, and experimental results showed that the classification accuracy on two benchmark datasets was 93.9% and 93.2%, respectively. Li et al. [25] studied the problem of gear pitting fault diagnosis under mixed operating conditions and proposed a method based on adaptive 1D separable convolution and residual connections. This method effectively reduced the number of model parameters and improved diagnostic performance by using separable convolution and residual connections. Experimental results showed that the diagnostic accuracy of gear pitting faults at different speeds reached 99.75%. Huang et al. [26] proposed a method to optimize network parameters by matching convolution kernel features with original signal features. Experimental results showed that this method achieved 100% classification accuracy in bearing vibration signal fault diagnosis and significantly improved computational efficiency. Kalay et al. [27] studied the problem of gear root crack diagnosis and proposed a method based on 1D-CNN. By establishing a dynamic model of asymmetric spur gears, they simulated the vibration responses of healthy and cracked gears. The experimental results showed that the classification accuracy of this method reached up to 99.251%, and when using asymmetric gear configurations, the classification accuracy increased by 12.8% compared with standard gears. Zhang et al. [28] studied the problem of gear crack depth diagnosis and proposed a model based on one-dimensional convolutional neural networks (CNN4GCDD). This model directly used raw vibration signals for diagnosis. Experimental results showed that CNN4GCDD achieved 100% accuracy on a single-speed dataset with 2048-point samples and 94.70% accuracy on a multi-speed dataset, outperforming the multi-layer LSTM model.

In summary, 1D-CNN has demonstrated strong performance and broad application prospects in the fields of mechanical fault diagnosis and network intrusion detection. In the field of mechanical fault diagnosis, 1D-CNN directly processes raw vibration signals without preprocessing, reducing computational costs and making it suitable for real-time diagnostic applications. In gear pitting fault diagnosis, bearing fault diagnosis, and high-speed train wheelset bearing fault diagnosis, 1D-CNN has exhibited high accuracy and low computational cost. In the field of network intrusion detection, 1D-CNN improves detection performance through the normalization of imbalanced data, significantly outperforming traditional machine learning methods. Future research will further explore the application of 1D-CNN under more complex working conditions and optimize network structures and parameters to further enhance model performance and generalization ability.

In recent years, hybrid models combining SVM and CNN have attracted widespread attention due to their excellent performance in complex data classification tasks. Studies have shown that by combining the efficient classification ability of SVM in small sample and high-dimensional feature spaces with the automatic feature extraction advantage of CNN, hybrid models can significantly improve model generalization ability and classification accuracy. In power system fault detection, Kumar et al. [29] proposed an SVM-CNN model that achieved an accuracy of 95.70% by integrating SVM's linear kernel partitioning with CNN's deep feature learning, outperforming individual SVM (88.20%) and CNN (91.25%) models. In fault detection and location of multi-terminal VSC-HVDC systems, Gnanamalar et al. [30] further improved accuracy to 99.87% and reduced response time to 2 milliseconds, demonstrating robustness in real-time scenarios. Similarly, Arévalo et al. [31] combined wavelet transform and differential protection technology in fault analysis of microgrid clusters, achieving 100% fault detection accuracy with a response time of less than 10 milliseconds, significantly outperforming traditional methods. Furthermore, addressing the data imbalance problem, Wu et al. [32] proposed an improved weighted loss function and physical feature fusion strategy in automotive radar target classification, increasing the F1 score and AUC to 0.90 and 0.99, respectively, effectively solving the challenge of pedestrian and vehicle classification. These findings indicate that the SVM-CNN model, through complementary optimization, can adapt to various data types such as images and time-series signals and shows potential in high-precision, low-latency scenarios, providing new solutions for cross-domain complex classification problems. However, existing studies mostly focus on single application scenarios, and further exploration is needed on how to optimize the hybrid model architecture to address broader cross-modal data fusion requirements.

To further explore the potential of combining CNN and SVM, this paper aims to develop a deep learning-based gear fault detection method by integrating the advantages of CNN and SVM to construct a novel hybrid model. This model has excellent feature learning capability, allowing it to automatically extract features from vibration signals and accurately classify faults.

The main contributions of this paper include:

(1) Developing a novel hybrid framework that combines CNN and SVM for gear fault detection. This model utilizes CNN's automatic feature extraction capability to process raw vibration signals and leverages SVM's strong classification performance in high-dimensional spaces. The synergy of these two algorithms overcomes the limitations of traditional feature engineering and enhances diagnostic accuracy under complex operating conditions.

(2) Extensive experiments conducted on the gear fault dataset demonstrate that the hybrid model achieves outstanding accuracy (0.9922) and training efficiency (14.95 seconds). The CNN-SVM architecture outperforms

individual classifiers (SVM, RF, KNN) and hybrid variants (CNN-RF, CNN-KNN), proving its effectiveness in capturing multi-scale fault features and handling nonlinear decision boundaries.

The structure of this paper is arranged as follows: Section 2 provides a detailed introduction to the gear fault test data and preprocessing methods; Section 3 presents the architecture and working principle of the proposed CNN-SVM model; Section 4 showcases the experimental results and performance evaluation; and Section 5 summarizes the paper and discusses future research directions.

2 Material

This study utilizes the experimental dataset compiled by Zamanian [33] at Southern Methodist University, which is specifically designed for fault diagnosis research in gears within gearboxes. This dataset meticulously records acceleration signals under different gear conditions, specifically including no fault, a chipped tooth, and three consecutive worn teeth. The data originates from a carefully designed gearbox test platform, which consists of a 15-tooth pinion and a 110-tooth gear, jointly achieving a speed ratio of 7.33. During the test, the pinion rotates at a speed of 1420 revolutions per minute, and by observing the first major peak in the Fast Fourier Transform (FFT) of the signals obtained under different gear conditions, the Gear Meshing Frequency (GMF) is precisely calculated as 365 Hz. The data acquisition process employs a high-precision analog-to-digital converter (Advantech PCI-1710, 12-bit resolution, sampling rate of 100 kS/s) and a high-performance accelerometer (Analog Device, ADXL210JQC), ensuring a sampling frequency of 10 kHz, thereby accurately capturing subtle vibration variations during gear operation.

The experimental data is stored in .mat file format for standardized storage, facilitating subsequent data processing and analysis. These data can be loaded using the “load” command in MATLAB software. All measurement channel data are recorded in volts, and to ensure data accuracy, all signals have undergone detrending processing to effectively eliminate bias factors in the accelerometer. This dataset covers three different gear conditions, with acceleration signals recorded for a duration of 10 seconds under each condition. These abundant raw data provide a solid foundation for subsequent fault feature analysis, the establishment of accurate diagnostic models, and the evaluation of diagnostic algorithm performance.

3 Methodology

This study proposes an innovative gear fault detection method that integrates the advantages of CNN and SVM to construct a CNN-SVM hybrid model. Firstly, various preprocessing techniques were applied to the vibration signals for gear fault detection, including Normalization, Z-score Standardization, Regularization, and Direct Signal Embedding (DSE), to ensure the reliability of the experimental results. Then, the preprocessed data were fed into the proposed CNN-SVM model, fully utilizing the excellent feature extraction capability of CNN and the powerful classification capability of SVM to achieve high-accuracy gear fault classification.

To verify the effectiveness of the proposed model, two sets of experiments were designed for performance comparison. In the first set of experiments, the processed data were separately input into classifiers such as SVM, KNN, and RF for classification to evaluate the performance of individual classifiers. In the second set of experiments, a cross-validation method was used to compare the performance and accuracy of SVM, KNN, and RF with the output results of CNN, further verifying the superiority of the CNN-SVM hybrid model. Finally, based on the advantages of the model itself, an attempt was made to improve prediction accuracy through model fusion, aiming for better application results in the field of gear fault detection.

3.1 Data Preprocessing

In this study, the gearbox vibration signal dataset was partitioned, and after comparing multiple data preprocessing methods, the DSE method was ultimately determined as the optimal strategy, ensuring data preparation for model training and validation.

First, the vibration signal datasets of three different conditions (no fault, a chipped tooth, and three worn teeth) were loaded. Each dataset was divided into 1000 samples, where each sample was a single-channel sequence of length 100. Subsequently, each dataset was divided into a training set and a test set, with each training set containing 700 samples and each test set containing 300 samples. After partitioning, to ensure the randomness of sample distribution during training, the data were randomized. Then, the data were converted from array format to tensor format and stored. To meet the input requirements of the 1D-CNN, the training and validation sets were reshaped accordingly, and the data were encapsulated. Finally, two data loaders were obtained, one for model training and one for model validation. The training data loader contained 2100 samples, while the test data loader contained 900 samples. Through the above steps, the data were successfully prepared for the training and evaluation of the gearbox fault diagnosis model, and two data loaders were encapsulated to facilitate the model training and validation process.

3.2 CNN-SVM Model Architecture

This study proposes an intelligent diagnosis model based on the combination of CNN and SVM, aiming to efficiently extract features from mechanical vibration signals and achieve accurate classification. The model consists of multiple convolutional layers, pooling layers, activation layers, fully connected layers, and Dropout layers. Through multi-level feature extraction and compression, it can effectively capture complex patterns in vibration signals. As shown in Figure 1, it is referred to as the CNN-SVM Hybrid Model.

Figure 1 illustrates the structure of the Inception 1D Block layer. This structure receives data from the previous layer and processes it through multiple branches. It includes a 1×1 convolution kernel branch, a 3×1 convolution kernel branch, a branch composed of 3×1 and 5×1 convolution kernels, and a branch containing 3×1 , 5×1 , and 7×1 convolution kernels. These convolution kernels extract feature information at different scales. The output of each branch is processed by the ReLU activation function and batch normalization (BatchNorm). Finally, all branch outputs are concatenated along the channel dimension ($\text{dim}=1$), and the final output data is used for subsequent processing.

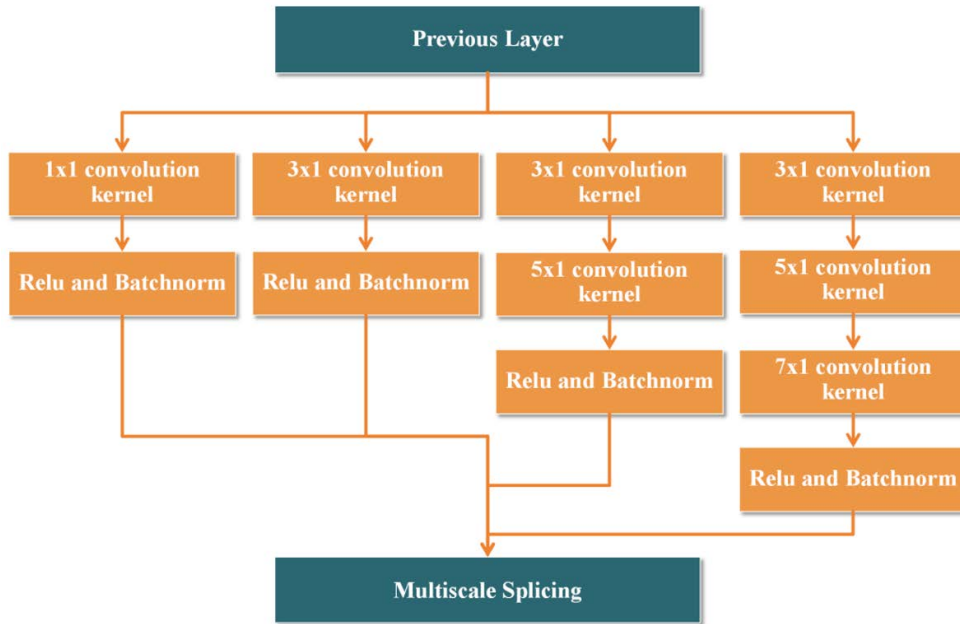


Figure 1. Inception 1D block layer structure

The complete CNN-SVM Hybrid Model is shown in Figure 2. I represents the number of input channels, O represents the number of output channels, C represents the convolution kernel size, S represents the convolution stride, and P represents the number of implicit zero paddings on both sides. The CNN model framework is built based on the PyTorch framework and mainly includes the following parts:

First, a Conv1D convolution layer is used to perform convolution operations on the input data. The convolution kernel size is 3, and the number of channels changes from 1 to 64. Padding of size 1 is used to maintain the feature map dimensions. Then, two Inception 1D Block layers are sequentially passed through. The first Inception 1D Block increases the number of channels from 64 to 516, while the second Inception 1D Block, due to having four branch outputs, has an input channel number of $128 \times 4 = 512$ and an output channel number of 1024.

The data then passes through a MaxPool1D layer with a pooling kernel size of 3 and a stride of 2 to reduce the spatial dimensions of the feature map. A Dropout layer with a probability of 0.5 is used to randomly drop some neurons to prevent overfitting. The Flatten layer is then used to flatten the multi-dimensional feature map into a one-dimensional vector. Finally, a view operation is used to extract a feature matrix, which is then used as the input to a SVM for training and obtaining results.

This hybrid model integrates the feature extraction capabilities of CNN with the classification advantages of SVM. CNN extracts data features automatically through convolutional and pooling layers, reduces the number of parameters, lowers the risk of overfitting, and improves training efficiency through parameter sharing. SVM, based on its structural risk minimization principle, improves accuracy, handles nonlinear problems with multiple kernel functions, and has high robustness, effectively dealing with noise and outliers. The output of CNN serves as the input for SVM, achieving an efficient combination of feature extraction and classification.

To verify the effectiveness of the InceptionNet1D model, this study adopted the experimental dataset provided

by Zamanian [33]. This dataset contains acceleration signals from a gearbox with helical gears, covering three different gear states: healthy state, a single chipped tooth, and three consecutive worn teeth. The data acquisition configuration and detailed descriptions have been provided in the related literature.

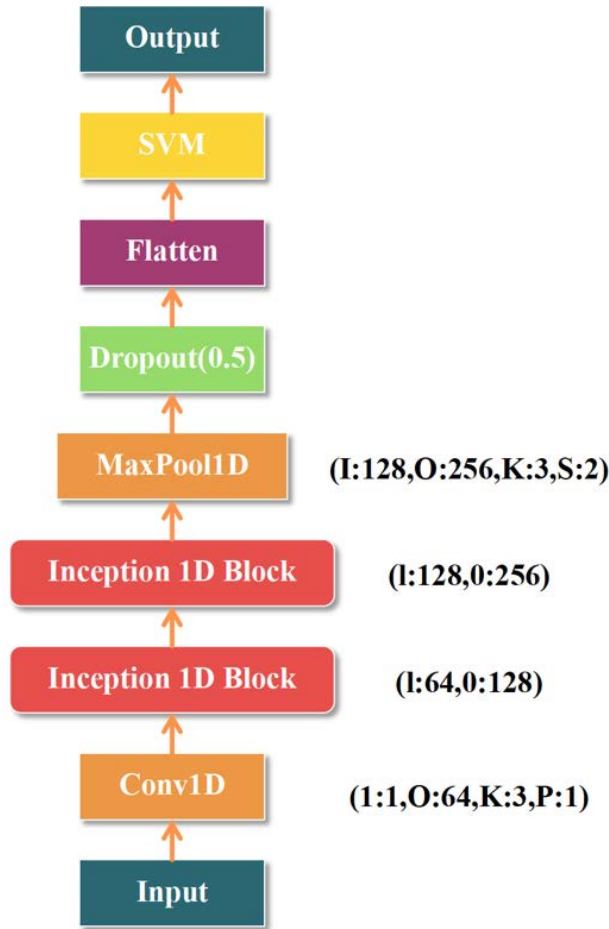


Figure 2. CNN-SVM hybrid model

4 Result Discussion

4.1 Experimental Setup

The gearbox fault diagnosis dataset used in this study contains vibration signals in three different fault states: no fault, a chipped tooth, and three worn teeth. Each state has 100,000 samples in its original state. These original signals were collected from a specific platform device system, which can simulate the operating conditions of gears in real-world working environments.

The experiments were conducted in a Windows 11 environment using a NVIDIA GeForce RTX 3060 Laptop GPU with 6GB memory on a PC. Python was used as the primary programming language, and PyTorch was used as the neural network learning framework to fully utilize parallel computing capabilities. The software environment for the experiments includes Python 3.10, deep learning framework PyTorch 2.1, and Scikit-learn 1.4.2 for data operations and machine learning algorithm implementation.

4.2 Selection of Preprocessing Methods

Data preprocessing plays a crucial role in model training, with the main objective of improving model accuracy, generalization ability, and training efficiency. Table 1 lists the performance evaluation results of different data preprocessing models. Yeo-Johnson power transformation is a data preprocessing technique designed to handle skewed distributions in data, making them closer to a normal distribution [34]. This transformation can be applied in regression analysis to improve model fitting performance. After adopting this method, the test set accuracy is 0.8889, with limited improvement in model performance. Z-score (standardization) is a commonly used data preprocessing method that converts data into a standard normal distribution based on its mean and standard deviation [35]. This eliminates the influence of different magnitudes, making the data comparable. However, the test set accuracy of

this method is only 0.3356, far lower than other methods, indicating that under the current dataset and model, standardization not only fails to improve model performance but also leads to a performance decline. Peak Value (peak extraction) focuses on extracting peak values in the data, assuming that peaks contain key data features to assist in model training and prediction [36]. The corresponding test set accuracy is 0.6467, which is relatively poor. It fails to fully extract useful information from the data to enhance model accuracy, suggesting that relying solely on peak information may not be sufficient to support efficient model learning. DSE directly uses raw data for training without complex preprocessing operations, preserving the original data information to the greatest extent and allowing the model to learn and extract useful features on its own. This method achieves a test set accuracy of 0.9800, obtaining the best model performance. This indicates that under the current dataset and model framework, the rich information contained in the original data can be effectively utilized by the model without excessive manual intervention or preprocessing, thereby achieving excellent performance.

Table 1. Performance evaluation of different data preprocessing models

Preprocessing Method	Yeo-Johnson	Z-score	Peak Value	Direct Signal Embedding
Test Set Accuracy	0.8889	0.3356	0.6467	0.9800

By analyzing the different preprocessing methods and their performance evaluation results in Table 1, it is clear that the DSE method should be adopted as the data preprocessing strategy in subsequent experiments. Further studies can explore the applicability and performance of the DSE method under different datasets and model architectures to enhance model performance and generalization ability, providing a solid theoretical and experimental foundation for model optimization and practical applications.

4.3 Experimental Comparison

4.3.1 Experimental comparison of machine learning models

The performance of SVM, RF, and KNN algorithms in gear fault classification prediction is evaluated.

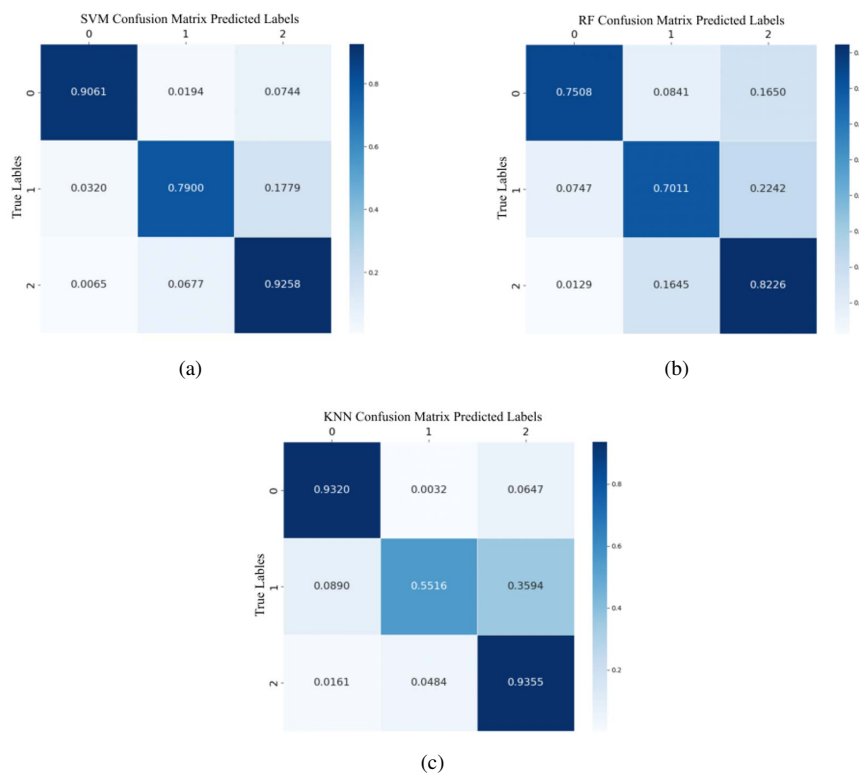


Figure 3. (a) Confusion matrix with SVM; (b) Confusion matrix with RF; (c) Confusion matrix with KNN

The confusion matrix is commonly used to visualize and quantify the performance of classification models. By presenting the true and predicted labels of each class, it helps to understand the accuracy and misclassification of the

model across different classes.

Figure 3 shows the confusion matrices of SVM, RF, and KNN algorithms in the gear fault classification task. SVM (subgraph (a) of Figure 3) perform well in classifying no-fault and three-worn-teeth categories, but the classification performance for one-tooth defect is relatively poor, with a large number of one-tooth defect samples misclassified as three-worn teeth. RF (subgraph (b) of Figure 3) fails to achieve an accuracy of 0.9000 in all three categories. KNN (subgraph (c) of Figure 3) perform well in classifying no-fault and three-worn-teeth categories, but the classification performance for one-tooth defect needs to be improved, with nearly half of one-tooth defect samples misclassified as fault-free or three-worn teeth. Table 2 presents the performance evaluation of different machine learning models. As seen in Table 2, SVM achieves the highest test set classification accuracy of 0.8767, followed by KNN with 0.8144, while RF has the lowest classification accuracy of 0.7600. This indicates that different fault types in the gear dataset can be classified using these models.

Table 2. Performance evaluation of different machine learning models

Algorithm	SVM	RF	KNN
Test set accuracy	0.8767	0.7600	0.8144

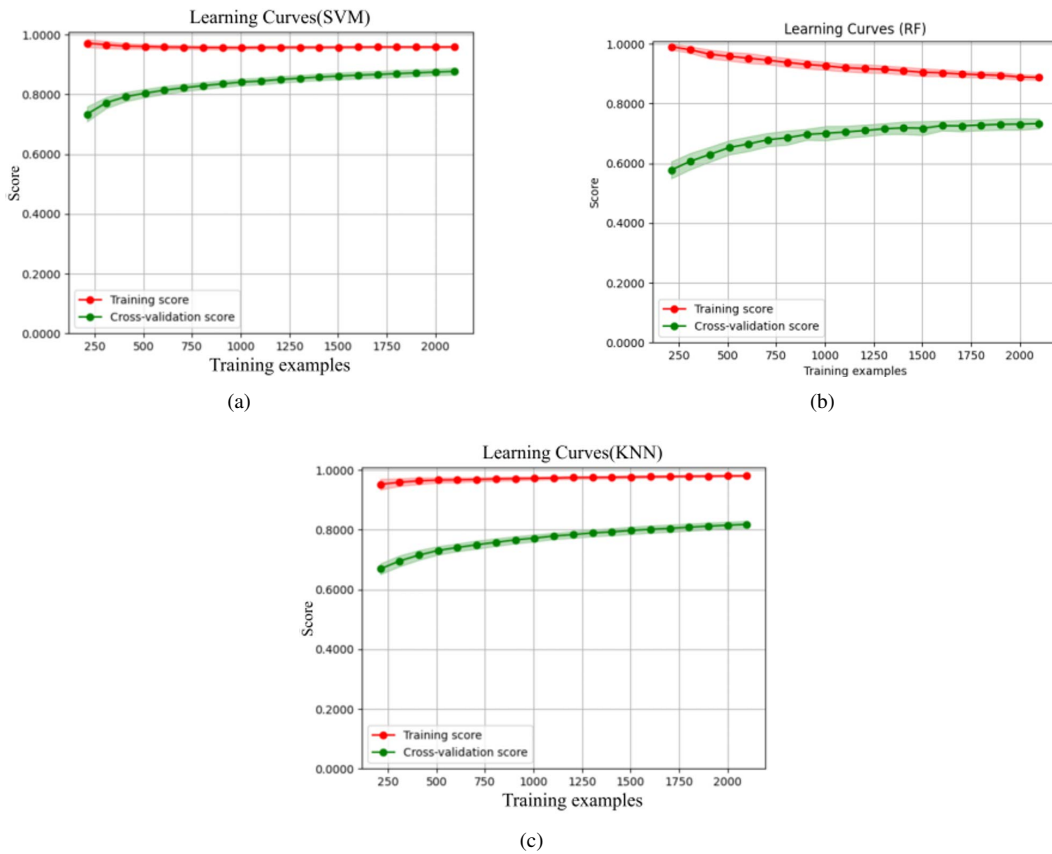


Figure 4. (a) Learning curves with SVM; (b) Learning curves with RF; (c) Learning curves with KNN

4.3.2 Experimental comparison of machine learning models with learning rate curve

In this section, the learning rate curve is used to evaluate and optimize SVM, RF, and KNN models, aiming to improve the training accuracy of the three models. During the training process, three different fault type datasets are first loaded and reshaped into a unified shape of (1000,100), then merged into the training set X. The labels Y are defined as [0,1,2], corresponding to the three fault types. Next, the SVM model is defined, with the RBF kernel function and OVR decision function shape. ShuffleSplit is used for cross-validation, dividing the dataset into training and test sets with 150 trials and a test set ratio of 30%. The learning_curve function calculates the training score and cross-validation score for different numbers of training samples and plots the learning curve. The number of training samples ranges from 10% to 100%, divided into 20 points. The learning curve shows the trend of training score and cross-validation score as the number of training samples increases, as well as their standard deviation range. The

training time is also recorded to evaluate the training efficiency of the model. Finally, through the learning curve, the performance and generalization ability of the model can be observed intuitively.

The learning curve analysis is shown in Figure 4, and the experimental results are presented in Table 3. From Figure 4, it can be seen that as the number of training samples increases, the training scores of the three models gradually stabilize, and the test scores also show a gradual increase and tend to stabilize. This indicates that the models are progressively learning the features of the data during training, and their performance on the test set also becomes stable, suggesting that the generalization ability of the models is improving. Table 3 shows that the best training score of the SVM model is 0.9718, the best cross-validation score is 0.8774, and the training time is 266.5980 seconds. Compared to other models, the SVM model has the highest score and the shortest training time, indicating that the SVM model has high generalization ability and learning efficiency on the current 1D vibration dataset. Specifically, the SVM model can effectively capture key information from the data and complete the training in a relatively short time, which makes it more advantageous in practical applications. Therefore, the SVM model is more suitable for the current data and can provide reliable model support for subsequent vibration data analysis and fault diagnosis.

Table 3. Performance evaluation of machine learning models after learning curve

Name	Leaning Curves (SVM)	Leaning Curves (RF)	Leaning Curves (KNN)
Training Time (s)	266.5980	494.3195	298.9411
Best Training Score	0.9593	0.8878	0.9811
Best Cross-Validation Score	0.8774	0.7332	0.8184

4.3.3 CNN model training results

Deep learning has demonstrated outstanding performance in gear fault detection in recent years. This section discusses the training effect of the CNN model.

Subgraph (a) and (b) of Figure 5 present the training records and curves of the CNN model. The training lasted for 150 epochs, and key indicators for each epoch were recorded. Subgraph (a) of Figure 5 shows that the training loss rapidly decreased to near 0, while the validation loss fluctuated initially and then stabilized. Subgraph (b) of Figure 5 shows that the training accuracy quickly rose to 1.0000 and remained stable, while the validation accuracy fluctuated initially and then stabilized, eventually reaching 0.9789. Overall, the training loss decreased, the validation loss stabilized, and both the training and validation accuracies were high, indicating that the model achieved the best learning effect on the training and validation sets.

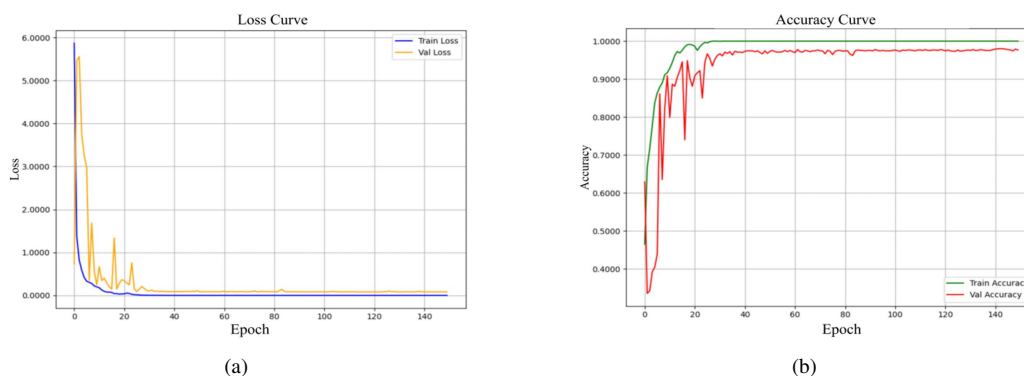


Figure 5. CNN model training performance

Although the CNN model exhibited high accuracy on both the training and validation sets, its generalization ability is still limited. To improve accuracy and generalization performance, a CNN+SVM combined model will be introduced in the future. This model performs better on multiple datasets and can handle complex classification tasks more effectively. The feature extraction capability of CNN combined with the machine learning classification ability aims to improve the accuracy and efficiency of gear fault detection. By training CNN to extract features and then using these features to train machine learning classifiers, it balances the complex feature representation of deep learning with the interpretability and flexibility of machine learning, aiming for the best detection results.

4.3.4 Experimental comparison of CNN with different models

By loading the previously trained CNN model weights, the features from the output layer (flatten layer) are extracted and used for traditional machine learning classifiers. The training time of each classifier is calculated, and

their performance is evaluated on the test set.

Subgraph (a) of Figure 6 shows the accuracy changes of the three models at different iteration counts. It can be seen that CNN+SVM maintains a high accuracy level at 0.9922, demonstrating good stability and superior classification performance. In contrast, the accuracy of the CNN+RF model fluctuates significantly and is overall lower than the CNN+SVM model; the CNN+KNN model has a significantly lower accuracy than the other two models, indicating poorer classification performance. Subgraph (b) of Figure 6 shows the change in training time for the three models with respect to iteration counts. Although the training time for CNN+SVM has increased, its accuracy advantage still makes it the better choice. In conclusion, CNN+SVM has a clear advantage in both accuracy and training time, especially in applications that require high classification performance.

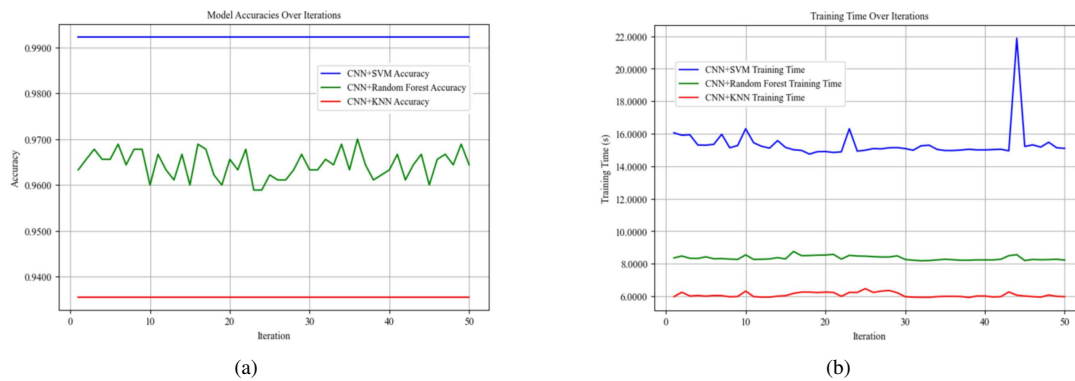


Figure 6. CNN with different models in the iterative process accuracy and training time comparison

5 Conclusion

This study developed a hybrid model based on CNN and SVM for gear fault detection. The model utilizes the powerful feature extraction capability of CNN and the efficient classification performance of SVM, achieving high-accuracy diagnosis of gear faults and reducing dependence on manual feature engineering. Experimental results show that the model achieved a high accuracy of 0.9922 on the gear fault dataset, significantly outperforming the single classifier models. Additionally, the model has a short training time, demonstrating its efficient training performance. Future work will focus on further optimization of model parameters and testing on more diversified datasets to enhance the model's robustness and adaptability, promoting its widespread application in real industrial environments.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflict of Interests

The author declares no conflicts of interest.

References

- [1] F. Karpat, A. E. Dirik, O. Dogan, O. C. Kalay, B. Korcuklu, and C. Yüce, "A novel Ai-based method for spur gear early fault diagnosis in railway gearboxes," in *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*, Istanbul, Turkey, 2020, pp. 1–6. <https://doi.org/10.1109/ASYU50717.2020.9259819>
- [2] X. Zhang, L. Wang, and Q. Miao, "Fault diagnosis techniques for planetary gearboxes under variable conditions: A review," in *2016 Prognostics and System Health Management Conference (PHM-Chengdu)*, Chengdu, China, 2016, pp. 1–11. <https://doi.org/10.1109/PHM.2016.7819889>
- [3] Y. G. Lei, J. Lin, Z. J. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," *Mech. Syst. Signal Process.*, vol. 35, no. 1-2, pp. 108–126, 2013. <https://doi.org/10.1016/j.ymssp.2012.09.015>
- [4] Y. Lei, J. Lin, Z. He, and Y. Zi, "Application of an improved kurtogram method for fault diagnosis of rolling element bearings," *Mech. Syst. Signal Process.*, vol. 25, no. 5, pp. 1738–1749, 2011. <https://doi.org/10.1016/j.ymssp.2010.12.011>
- [5] Z. Jiang, Q. Han, and X. Xu, "Fault diagnosis of planetary gearbox based on motor current signal analysis," *Shock Vib.*, vol. 2020, no. 1, p. 8854776, 2020. <https://doi.org/10.1155/2020/8854776>

- [6] L. Zhang and Q. Yang, "Investigation of the design and fault prediction method for an abrasive particle sensor used in wind turbine gearbox," *Energies*, vol. 13, no. 2, p. 365, 2020. <https://doi.org/10.3390/en13020365>
- [7] H. Chang, P. Borghesani, and Z. Peng, "Investigation on the relationship between macropits and wear particles in a gear fatigue process," *Wear*, vol. 484, p. 203724, 2021. <https://doi.org/10.1016/j.wear.2021.203724>
- [8] K. Feng, W. A. Smith, P. Borghesani, R. B. Randall, and Z. Peng, "Use of cyclostationary properties of vibration signals to identify gear wear mechanisms and track wear evolution," *Mech. Syst. Signal Process.*, vol. 150, p. 107258, 2021. <https://doi.org/10.1016/j.ymsp.2020.107258>
- [9] K. Feng, J. C. Ji, Q. Ni, and M. Beer, "A review of vibration-based gear wear monitoring and prediction techniques," *Mech. Syst. Signal Process.*, vol. 182, p. 109605, 2023. <https://doi.org/10.1016/j.ymsp.2022.109605>
- [10] P. Hu, C. Zhao, J. Huang, and T. Song, "Intelligent and small samples gear fault detection based on wavelet analysis and improved CNN," *Processes*, vol. 11, no. 10, p. 2969, 2023. <https://doi.org/10.3390/pr11102969>
- [11] S. H. Yue, P. Li, and P. Y. Hao, "SVM classification: Its contents and challenges," *Appl. Math. Chin. Univ.*, vol. 18, pp. 332–342, 2003. <https://doi.org/10.1007/s11766-003-0059-5>
- [12] C. Zhang, Z. X. Peng, S. Chen, Z. X. Li, and J. G. Wang, "A gearbox fault diagnosis method based on frequency-modulated empirical mode decomposition and support vector machine," *Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci.*, vol. 232, no. 2, pp. 369–380, 2018. <https://doi.org/10.1177/0954406216677102>
- [13] R. Medina, J. C. Macancela, P. Lucero, D. Cabrera, R. V. Sánchez, and M. Cerrada, "Gear and bearing fault classification under different load and speed by using Poincaré plot features and SVM," *J. Intell. Manuf.*, vol. 33, no. 4, pp. 1031–1055, 2022. <https://doi.org/10.1007/s10845-020-01712-9>
- [14] H. F. Chen, Y. Q. Sun, Z. Y. Shi, and J. C. Lin, "Intelligent analysis method of gear faults based on FRWT and SVM," *Shock Vib.*, vol. 2016, p. 1582738, 2016. <https://doi.org/10.1155/2016/1582738>
- [15] H. F. Wang and D. J. Hu, "Comparison of SVM and LS-SVM for regression," in *2005 International Conference on Neural Networks and Brain*, Beijing, China, 2005, pp. 279–283. <https://doi.org/10.1109/ICNNB.2005.1614615>
- [16] Z. J. Shen, X. F. Chen, X. L. Zhang, and Z. J. He, "A novel intelligent gear fault diagnosis model based on EMD and multi-class TSVM," *Measurement*, vol. 45, no. 1, pp. 30–40, 2012. <https://doi.org/10.1016/j.measurement.2011.10.008>
- [17] D. Y. Han, N. Zhao, and P. M. Shi, "Gear fault feature extraction and diagnosis method under different load excitation based on EMD, PSO-SVM and fractal box dimension," *J. Mech. Sci. Technol.*, vol. 33, pp. 487–494, 2019. <https://doi.org/10.1007/s12206-019-0101-z>
- [18] J. S. Kang, X. H. Zhang, J. M. Zhao, and D. C. Cao, "Gearbox fault prognosis based on CHMM and SVM," in *2012 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, Chengdu, China, 2012, pp. 703–708. <https://doi.org/10.1109/ICQR2MSE.2012.6246327>
- [19] M. Pule, O. Matsebe, and R. Samikannu, "Application of PCA and SVM in fault detection and diagnosis of bearings with varying speed," *Math. Probl. Eng.*, vol. 2022, p. 5266054, 2022. <https://doi.org/10.1155/2022/5266054>
- [20] L. L. Jiang, H. K. Yin, X. J. Li, and S. W. Tang, "Fault diagnosis of rotating machinery based on multisensor information fusion using SVM and time-domain features," *Shock Vib.*, vol. 2014, p. 418178, 2014. <https://doi.org/10.1155/2014/418178>
- [21] Y. F. Ma, X. S. Jia, H. J. Bai, G. Z. Liu, G. L. Wang, C. M. Guo, and S. C. Wang, "A new fault diagnosis method based on convolutional neural network and compressive sensing," *J. Mech. Sci. Technol.*, vol. 33, no. 11, pp. 5177–5188, 2019. <https://doi.org/10.1007/s12206-019-1007-5>
- [22] J. X. Wang, D. Z. Wang, and X. H. Wang, "Fault diagnosis of industrial robots based on multi-sensor information fusion and 1D convolutional neural network," in *2020 39th Chinese Control Conference (CCC)*, Shenyang, China, 2020, pp. 3087–3091. <https://doi.org/10.23919/CCC50068.2020.9189568>
- [23] D. D. Peng, Z. L. Liu, H. Wang, Y. Qin, and L. M. Jia, "A novel deeper one-dimensional CNN with residual learning for fault diagnosis of wheelset bearings in high-speed trains," *IEEE Access*, vol. 7, pp. 10 278–10 293, 2019. <https://doi.org/10.1109/ACCESS.2018.2888842>
- [24] L. Eren, T. Ince, and S. Kiranyaz, "A generic intelligent bearing fault diagnosis system using compact adaptive 1D CNN classifier," *J. Signal Process. Syst.*, vol. 91, no. 2, pp. 179–189, 2019. <https://doi.org/10.1007/s11265-018-1378-3>
- [25] X. Y. Li, J. L. Li, C. Y. Zhao, Y. Z. Qu, and D. He, "Gear pitting fault diagnosis with mixed operating conditions based on adaptive 1D separable convolution with residual connection," *Mech. Syst. Signal Process.*, vol. 142, p. 106740, 2020. <https://doi.org/10.1016/j.ymsp.2020.106740>
- [26] S. Z. Huang, J. Tang, J. Y. Dai, and Y. Y. Wang, "Signal status recognition based on 1DCNN and its feature extraction mechanism analysis," *Sensors*, vol. 19, no. 9, p. 2018, 2019. <https://doi.org/10.3390/s19092018>

- [27] O. C. Kalay, E. Karpat, A. E. Dirik, and F. Karpat, "A one-dimensional convolutional neural network-based method for diagnosis of tooth root cracks in asymmetric spur gear pairs," *Machines*, vol. 11, no. 4, p. 413, 2023. <https://doi.org/10.3390/machines11040413>
- [28] S. H. Zhang, J. H. Zhou, E. H. Wang, and S. Pirttikangas, "CNN4GCDD: A one-dimensional convolutional neural network-based model for gear crack depth diagnosis," in *2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, Hangzhou, China, 2022, pp. 1138–1142. <https://doi.org/10.1109/CSCWD54268.2022.9776142>
- [29] N. M. G. Kumar, S. S. Priya, V. Varghese, K. Renuga, J. P. Patra, and N. Nishant, "An improved method for fault detection electric power system protection using SVMCNN model," in *2024 International Conference on Electronics, Computing, Communication and Control Technology (ICECCC)*, Bengaluru, India, 2024, pp. 1–6. <https://doi.org/10.1109/ICECCC61767.2024.10593976>
- [30] A. J. Gnanamalar, R. Bhavani, A. S. Arulini, and M. S. Veerajuu, "CNN-SVM based fault detection, classification and location of multi-terminal VSC-HVDC system," *J. Electr. Eng. Technol.*, vol. 18, no. 4, pp. 3335–3347, 2023. <https://doi.org/10.1007/s42835-023-01391-5>
- [31] P. Arévalo, A. Cano, D. Benavides, and F. Jurado, "Fault analysis in clustered microgrids utilizing SVM-CNN and differential protection," *Appl. Soft Comput.*, vol. 164, p. 112031, 2024. <https://doi.org/10.1016/j.asoc.2024.112031>
- [32] Q. S. Wu, T. Gao, Z. C. Lai, and D. Z. Li, "Hybrid SVM-CNN classification technique for human-vehicle targets in an automotive LFM CW radar," *Sensors*, vol. 20, no. 12, p. 3504, 2020. <https://doi.org/10.3390/s20123504>
- [33] A. H. Zamanian, "Experimental dataset for gear fault diagnosis," Ingram School of Engineering, Texas State University, United States, 2014. <https://doi.org/10.13140/RG.2.2.28152.44802/2>
- [34] S. Weisberg, "Yeo-Johnson power transformations," Department of Applied Statistics, University of Minnesota, United States, 2001.
- [35] M. Yuzukirmizi, H. Arslan, and O. Duman, "Multivariate statistical methods for detection of spur gear faults," *Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci.*, vol. 229, no. 14, pp. 2586–2598, 2015. <https://doi.org/10.1177/0954406214561050>
- [36] J. Igba, K. Alemzadeh, C. Durugbo, and E. T. Eiriksson, "Analysing RMS and peak values of vibration signals for condition monitoring of wind turbine gearboxes," *Renewable Energy*, vol. 91, pp. 90–106, 2016. <https://doi.org/10.1016/j.renene.2016.01.006>