**Novel Machine Learning Techniques for Classification of Rolling Bearings**

**ABSTRACT:**

Rolling bearing faults frequently cause rotating equipment failure, leading to costly downtime and maintenance expenses. As a result, researchers have focused on developing effective methods for diagnosing these faults. In this paper, we explore the potential of Machine Learning (ML) techniques for classifying the health status of bearings. Our approach involves decomposing the signal, extracting statistical features, and using feature selection employing Binary Grey Wolf Optimization. We propose an ensemble method using voting classifiers to diagnose faults based on the reduced set of features. To evaluate the performance of our methods, we utilize several performance indicators. Our results demonstrate that the proposed voting classifiers method achieves superior fault classification, highlighting its potential for use in predictive maintenance applications.

**INTRODUCTION**

1

Rolling bearings are critical components in rotating machinery and are widely used in various industrial applications such as automotive, aerospace, and manufacturing. These components play a crucial role in the smooth operation of machines by reducing friction between moving parts. However, the reliability and performance of rolling bearings can significantly affect the overall health of machinery. Faults in bearings, such as surface damage, misalignment, and wear, can lead to severe consequences, including machinery breakdowns, costly repairs, unexpected downtime, and potential safety hazards. Therefore, early and accurate detection of bearing faults is essential for maintaining system efficiency, minimizing maintenance costs, and preventing catastrophic failures.

In recent years, researchers have concentrated on developing advanced techniques to diagnose bearing faults effectively. Traditional diagnostic methods, such as vibration analysis and signal processing techniques, have been widely used to monitor bearing conditions. These methods involve analyzing the vibration signals generated by bearings to detect anomalies. However, traditional approaches often face challenges due to the complexity of vibration signals, the presence of noise, and the requirement for expert knowledge to interpret results. With the advent of Machine Learning (ML) techniques, new opportunities have emerged for automating the fault diagnosis process, improving accuracy, and reducing dependency on manual expertise. ML algorithms can analyze large datasets, learn patterns, and make predictions, making them suitable for fault diagnosis applications.

This paper explores the potential of ML techniques for classifying the health status of rolling bearings. The proposed approach begins with signal decomposition, where raw vibration signals are processed to extract useful information. Following this, a set of statistical features is derived from the decomposed signals, providing a comprehensive representation of the bearing’s condition. To enhance the diagnostic performance, feature selection is performed using Binary Grey Wolf Optimization, which helps in selecting the most relevant features, thereby reducing computational complexity and improving classification accuracy.

**SCOPE OF THE PROJECT**

The scope of this project is centered around developing an efficient and automated system for diagnosing faults in rolling bearings, which are critical components in rotating machinery. The primary aim is to leverage Machine Learning (ML) techniques to improve the reliability and accuracy of fault detection, thereby reducing downtime and maintenance costs. The project involves a comprehensive approach that starts with signal decomposition and statistical feature extraction, followed by feature selection using Binary Grey Wolf Optimization. By focusing on optimizing feature sets, we can enhance the performance of classifiers, ensuring the model is both efficient and scalable. The proposed solution incorporates an ensemble learning method using voting classifiers, specifically Random Forest, XGBoost, and Support Vector Classifier (SVC), to deliver superior diagnostic accuracy. This ensemble approach capitalizes on the strengths of each individual classifier, providing a robust and resilient fault detection framework. Additionally, the system's capabilities extend to predictive maintenance, offering early detection of bearing issues, which can significantly prolong equipment lifespan and prevent unexpected failures. This project is scalable and adaptable, making it applicable to various types of rotating equipment across different industries, ultimately contributing to smarter maintenance strategies and operational efficiency.

**OBJECTIVE**

The primary objective of this project is to develop a robust and efficient diagnostic system for detecting faults in rolling bearings, which are critical components in many industrial machines. The system aims to leverage advanced Machine Learning (ML) techniques to improve the accuracy and reliability of fault diagnosis, thus minimizing the risks of unexpected equipment failures. By decomposing vibration signals, extracting key statistical features, and optimizing feature selection using Binary Grey Wolf Optimization, the project focuses on enhancing the classification accuracy for identifying bearing health conditions. Additionally, it proposes an ensemble approach using voting classifiers—specifically Random Forest, XGBoost, and Support Vector Classifier (SVC)—to ensure a high-performing, generalized model capable of distinguishing between various fault types with precision. The overarching goal is to integrate this system into predictive maintenance frameworks, enabling industries to proactively manage equipment health, optimize maintenance schedules, and ultimately reduce operational costs while maximizing machinery uptime. Through comprehensive performance evaluations using standard indicators, the system aims to set a new benchmark in bearing fault diagnostics, surpassing traditional methods and contributing to the field of smart maintenance solutions.

**EXISTING SYSTEM:**

Traditional bearing fault diagnosis, researchers have relied on signal processing techniques to extract features from vibration signals. These techniques involve time-domain, frequency-domain, or time-frequency domain analyses. The extracted features are then used in machine learning classifiers like Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Decision Trees (DT), and Random Forest (RF) for fault classification.

Moreover, many existing systems use individual classifiers for fault diagnosis, which can lead to suboptimal results due to the limitations of single-model approaches. These methods often suffer from overfitting, poor generalization, or sensitivity to noisy data. The accuracy of these standalone classifiers is often insufficient for real-time predictive maintenance applications, especially when dealing with complex fault patterns in bearing systems.

**EXISTING SYSTEM DISADVANTAGES:**

* Limited Generalization
* Noise Sensitivity
* Overfitting and Underfitting
* Feature Dependence
* Model Complexity

**LITERATURE SURVEY**

**Title:** A Fault Diagnosis Method of Rolling Bearing Based on Improved Recurrence Plot and Convolutional Neural Network

**Author:** Xiaoping Liu, Lijian Xia, Jian Shi, Lijie Zhang, Linying Bai, Shao-Ping Wang

**Year:** 2023

**Description:** The recurrence plot (RP) method has been introduced into bearing fault diagnosis due to its capability of effectively analyzing nonlinear and nonstationary waveform signals in dynamic systems. However, the interference of noise increases the difficulty of RP-based fault diagnosis. To solve this problem, this article proposed a novel antinoise bearing fault diagnosis method based on improved RP and a convolutional neural network (CNN). First, different scales of approximation coefficients and detail coefficients were obtained and constructed for RP based on wavelet packet decomposition (WPD) on the vibrational signal. Meanwhile, redundant parts of each RP were removed according to its symmetry characteristics, and the remaining parts of these RPs were spliced into multiscale asymmetric RP (MARP) containing all coefficients. Then, a fault diagnosis model for rolling bearing was established with MARP as the input of the pretrained ResNet-34. Finally, the validity of the proposed fault diagnosis method was validated on the Paderborn bearing dataset. Experimental results showed that the proposed fault diagnosis method achieved an accuracy of 90% under Gaussian white noise with a signal-to-noise ratio (SNR) of above −6 dB.

**Title:** A Two-Stage Feature Selection Approach for Fruit Recognition Using Camera Images With Various Machine Learning Classifiers

**Author:** Tri Tran Minh Huynh; Tuan Minh Le; Long Ton That; Ly Van Tran; Son Vu Truong Dao

**Year:** 2022.

**Description**: Fruit and vegetable identification and classification system is always necessary and advantageous for the agriculture business, the food processing sector, as well as the convenience shops and hypermarkets where these products are sold. Therefore, it is necessary to build an effective automated tool to meet the needs of the market by boosting the outcome, in order to improve economic efficiency. In this paper, a two-stage model is proposed to recognize fruits using camera images. We employed a Densnet121 to get the features from the fruits dataset in the first module. In the second stage, we utilize a feature subset selection method to choose the most significant features for recognizing fruits from the images of the fruits. In this study, Adaptive particle - Grey Wolf Optimization (APGWO) has been applied for choosing the most pertinent features. The final subset feature has been used for recognizing fruits using several machine learning classifiers, namely K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), and Multilayer Perceptron (MLP). The proposed research’s experimental results are highly effective; the training time of proposed models is reduced to over 50%, and the classification accuracy reaches 99%.

**Title:** Application of Machine Learning in Epileptic Seizure Detection

**Author:** Ly V Tran, Hieu M Tran, Tuan M Le, Tri T M Huynh, Hung T Tran, Son V T Dao

**Year:** 2022.

**Description:** Epileptic seizure is a neurological condition caused by short and unexpectedly occurring electrical disruptions in the brain. It is estimated that roughly 60 million individuals worldwide have had an epileptic seizure. Experiencing an epileptic seizure can have serious consequences for the patient. Automatic seizure detection on electroencephalogram (EEG) recordings is essential due to the irregular and unpredictable nature of seizures. By thoroughly analyzing EEG records, neurophysiologists can discover important information and patterns, and proper and timely treatments can be provided for the patients. This research presents a novel machine learning-based approach for detecting epileptic seizures in EEG signals. A public EEG dataset from the University of Bonn was used to validate the approach. Meaningful statistical features were extracted from the original data using discrete wavelet transform analysis, then the relevant features were selected using feature selection based on the binary particle swarm optimizer. This facilitated the reduction of 75% data dimensionality and 47% computational time, which eventually sped up the classification process. After having been selected, relevant features were used to train different machine learning models, then hyperparameter optimization was utilized to further enhance the models' performance. The results achieved up to 98.4% accuracy and showed that the proposed method was very effective and practical in detecting seizure presence in EEG signals. In clinical applications, this method could help relieve the suffering of epilepsy patients and alleviate the workload of neurologists.

**Title:**  Signal-to-Image: Rolling Bearing Fault Diagnosis Using ResNet Family Deep-Learning Models

**Author:**  Guoguo Wu ,Xuerong Ji ,Guolai Yang 1,Ye Jia and Chuanchuan Cao

**Year:** 2023

**Description**: Rolling element bearings (REBs) are the most frequent cause of machine breakdowns. Traditional methods for fault diagnosis in rolling bearings rely on feature extraction and signal processing techniques. However, these methods can be affected by the complexity of the underlying patterns and the need for expert knowledge during signal analysis. This paper proposes a novel signal-to-image method in which the raw signal data are transformed into 2D images using continuous wavelet transform (CWT). This transformation enhances the features extracted from the raw data, allowing for further analysis and interpretation. Transformed images of both normal and faulty rolling bearings from the Case Western Reserve University (CWRU) dataset were used with deep-learning models from the ResNet family. They can automatically learn and identify patterns in raw vibration signals after continuous wavelet transform is used, eliminating the need for manual feature extraction. To further improve the training results, squeeze-and-excitation networks (SENets) were added to improve the process. By comparing results obtained from several models, we found that SE-ResNet152 has the best performance for REB fault diagnosis.

**Title:** Triboelectric nanogenerator-embedded intelligent bearing with rolling ball defect diagnosis via signal decomposition and automated machine learning

**Author**:Fangyang Dong , Hengyi Yang, Hengxu Du, Meixian Zhu, Ziyue Xi, Yulian Wang, Taili Du , Minyi Xu

**Year:** 2024**.**

**Description:** Smart fault diagnosis of bearings is of great significance due to their extensive applications on various occasions. Recently, self-powered sensing technology based on triboelectric nanogenerators promotes the development of intelligent bearings. However, the effective detection and recognition of the rolling element defects of bearings need to be investigated further. This study proposes a triboelectric sensor-embedded rolling bearing (T-bearing) to monitor the working conditions and conduct the defect diagnosis of rolling balls. The interdigitated copper electrode covered by polytetrafluoroethylene film is attached to the inner surface of the outer ring of a commercial bearing. Such a design not only directly forms the TENG with rolling balls to obtain the contact-sensing signals, but also successfully achieves the diagnosis of rolling ball defects with similar triboelectric signals through a novel analysis and prediction paradigm combining signal decomposition and automated machine learning. Finally, a recognition accuracy of 99.48% with five different conditions of bearing balls is reached, which is extremely superior to the highest accuracy of 78.34% without signal decomposition. Thus, this study provides a new strategy for the defect diagnosis and the intelligent application of tribo electricbearings.

**1.6 PROPOSED SYSTEM**

To address the limitations of traditional bearing fault diagnosis methods, we propose an advanced ensemble-based approach utilizing a voting classifier. Our system begins with signal decomposition and statistical feature extraction, followed by feature selection using Binary Grey Wolf Optimization (BGWO). Unlike conventional methods, we employ an ensemble of three powerful classifiers: Random Forest, XGBoost, and Support Vector Classifier (SVC). These models are combined using a voting strategy, which significantly enhances classification accuracy and robustness.

The use of a voting ensemble leverages the strengths of each classifier—Random Forest's ability to handle noisy data, XGBoost's gradient boosting efficiency, and SVC's precision in class boundaries. By integrating these models, our approach achieves superior fault classification performance.The results demonstrate that our ensemble method outperforms standalone classifiers, providing a reliable and efficient solution for early detection of bearing faults in rotating machinery.

**PROPOSED SYSTEM ADVANTAGES:**

* Improved Accuracy
* Robustness to Noise
* Enhanced Generalization
* Better Feature Selection
* Reduced Overfitting

**SYSTEM REQUIREMENTS**

**HARDWARE REQUIREMENTS**

The hardware requirements may serve as the basis for a contract for the implementation of the system and should therefore be a complete and consistent specification of the whole system. They are used by software engineers as the starting point for the system design. It should what the system do and not how it should be implemented.

* PROCESSOR : DUAL CORE 2 DUOS.
* RAM : 4GB DD RAM
* HARD DISK : 500 GB

**SOFTWARE REQUIREMENTS**

The software requirements document is the specification of the system. It should include both a definition and a specification of requirements. It is a set of what the system should do rather than how it should do it. The software requirements provide a basis for creating the software requirements specification. It is useful in estimating cost, planning team activities, performing tasks and tracking the teams and tracking the team’s progress throughout the development activity.

* Operating System : Windows 10
* Platform : Spyder3
* Programming Language : Python
* Front End : Spyder3

**SYSTEM ARCHITECTURE:**

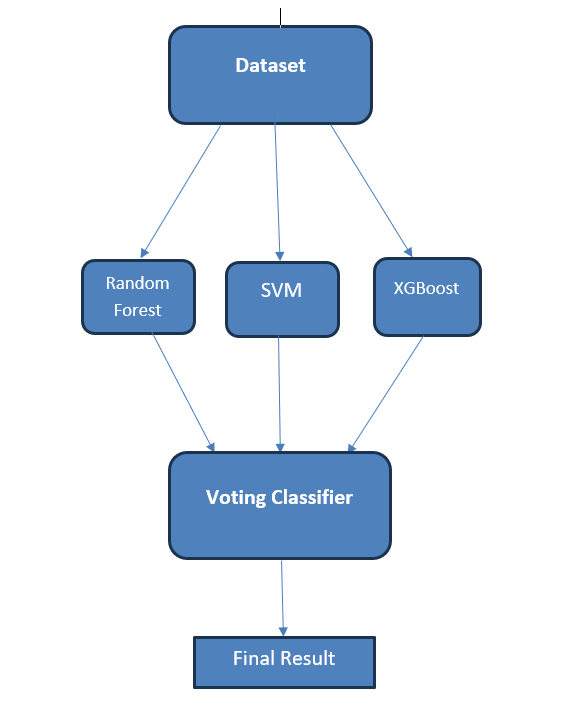


Fig : System Architecture

**REFERENCES**

[1] M. Syafrudin, G. Alfian, N. L. Fitriyani, and J. Rhee, “Performance Analysis of IoT-BasedSensor, Big Data Processing, and Machine Learning Model for Real-Time Monitoring System in Automotive Manufacturing,” Sensors, vol. 18, no. 9, Art. no. 9, Sep. 2018, doi: 10.3390/s18092946.

[2] F. Arellano-Espitia, M. Delgado-Prieto, V. Martinez-Viol, J. J. Saucedo-Dorantes, and R. A.

Osornio-Rios, “Deep-Learning-Based Methodology for Fault Diagnosis in Electromechanical Systems,” Sensors, vol. 20, no. 14, Art. no. 14, Jan. 2020, doi: 10.3390/s20143949.

[3]R. Marino, C. Wisultschew, A. Otero, J. M. Lanza Gutierrez, J. Portilla, and E. de la Torre, “A

Machine-Learning-Based Distributed System for Fault Diagnosis With Scalable Detection Quality in Industrial IoT,” IEEE Internet Things J., vol. 8, no. 6, pp. 4339–4352, Mar. 10.1109/JIOT.2020.3026211. 2021,

[4] X. Liu, L. Xia, J. Shi, L. Zhang, L. Bai, and S. Wang, “A Fault Diagnosis Method of Rolling Bearing Based on Improved Recurrence Plot and Convolutional Neural Network,” IEEE Sens. J., vol. 23, no. 10, pp. 10767–10775, May 2023, doi: 10.1109/JSEN.2023.3265409.

[5] H. Tao, J. Qiu, Y. Chen, V. Stojanovic, and L. Cheng, “Unsupervised cross-domain rolling bearing fault diagnosis based on time-frequency information fusion,” J. Frankl. Inst., vol. 360, no. 2, pp. 14541477, Jan. 2023, doi10.1016/j.jfranklin.2022.11.004

[6] A. Alkaya and İ. Eker, “Variance sensitive adaptive 2, threshold-based PCA method for fault detection with experimental application,” ISA Trans., vol. 50, no. pp. 287–302, Apr. 2011, doi: 10.1016/j.isatra.2010.12.004.

[7] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, “KNN Model-Based Approach in Classification,” in On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE, R. Meersman, Z. Tari, and D. C. Schmidt, Eds., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2003, pp. 986–996. doi: 10.1007/978-3-540-399643\_62.

[8] G. Wu, X. Ji, G. Yang, Y. Jia, and C. Cao, “Signal to-Image: Rolling Bearing Fault Diagnosis Using ResNet Family Deep-Learning Models,” Processes, vol. 11, no. 5, Art. no. 5, May 2023, doi: 10.3390/pr11051527.

[9] F. Dong et al., “Triboelectric nanogenerator embedded intelligent bearing with rolling ball defect diagnosis via signal decomposition and automated machine learning,” Nano Energy, vol. 119, p. 109072, Jan. 10.1016/j.nanoen.2023.109072. 2024,

[10] S. Afrasiabi, M. Afrasiabi, B. Parang, and M. Mohammadi, “Real-Time Bearing Fault Diagnosis of Induction Motors with Accelerated Deep Learning Approach,” in 2019 10th International Power Electronics, Drive Systems and Technologies Conference (PEDSTC), Feb. 2019, pp. 155–159. doi: 10.1109/PEDSTC.2019.8697244.

[11] X. Li, W. Zhang, and Q. Ding, “Understanding and improving deep learning-based rolling bearing fault diagnosis with attention mechanism,” Signal Process., vol. 161, pp. 136–154, Aug. 2019, doi: 10.1016/j.sigpro.2019.03.019.

[12] G. Xu, M. Liu, Z. Jiang, D. Söffker, and W. Shen, “Bearing Fault Diagnosis Method Based on Deep Convolutional Neural Network and Random Forest Ensemble Learning,” Sensors, vol. 19, no. 5, Art. no. 5, Jan. 2019, doi: 10.3390/s19051088.

[13] L. Yuan, D. Lian, X. Kang, Y. Chen, and K. Zhai, “Rolling Bearing Fault Diagnosis Based on Convolutional Neural Network and Support Vector Machine,” IEEE Access, vol. 8, pp. 137395–137406, 2020, doi: 10.1109/ACCESS.2020.3012053.

[14] X. Li, Y. Yang, H. Pan, J. Cheng, and J. Cheng, “A novel deep stacking least squares support vector machine for rolling bearing fault diagnosis,” Comput. Ind., vol. 110, pp. 36–47, Sep. 2019, doi: 10.1016/j.compind.2019.05.005.

[15] A. Ucar, M. Karakose, and N. Kırımça, “Artificial Intelligence for Predictive Maintenance

Applications: Key Components, Trustworthiness, and Future Trends,” Appl. Sci., vol. 14, no. 2, Art. no. 2, Jan. 2024, doi: 10.3390/app14020898.

[16] S. Kaparthi and D. Bumblauskas, “Designing predictive maintenance systems using decision tree based machine learning techniques,” Int. J. Qual. Reliab. Manag., vol. 37, no. 4, pp. 659–686, Jan. 2020, doi: 10.1108/IJQRM-04-2019-0131.

[17] B. Qi, Y. Li, W. Yao, and Z. Li, “Application of EMD Combined with Deep Learning and

Knowledge Graph in Bearing Fault,” J. Signal Process. Syst., vol. 95, no. 8, pp. 935–954, Aug.

2023, doi: 10.1007/s11265-023-01845-z.

[18] F. Dao, Y. Zeng, and J. Qian, “A novel denoising method of the hydro-turbine runner for fault signal based on WT-EEMD,” Measurement, vol. 219, p. 113306, Sep. 2023, doi: 10.1016/j.measurement.2023.113306.

[19] Y. Hu, Y. Ouyang, Z. Wang, H. Yu, and L. Liu, “Vibration signal denoising method based on CEEMDAN and its application in brake disc unbalance detection,” Mech. Syst. Signal Process.,

vol. 187, p. 109972, Mar. 2023, doi: 10.1016/j.ymssp.2022.109972.

[20] R. Abdelkader, A. Kaddour, A. Bendiabdellah, and Z. Derouiche, “Rolling Bearing Fault Diagnosis Based on an Improved Denoising Method Using the Complete Ensemble Empirical Mode Decomposition and the Optimized Thresholding Operation,” IEEE Sens. J., vol. 18, no. 17, pp. 71667172, Sep. 2018, doi: 10.1109/JSEN.2018.2853136.

[21] Z. Wu and N. E. Huang, “Ensemble empirical mode decomposition: a noise-assisted data analysis method,” Adv. Adapt. Data Anal., vol. 01, no. 01, pp. 1–41, Jan. 2009, doi: 10.1142/S1793536909000047.

[22] Y. Li, Y. Li, X. Chen, J. Yu, H. Yang, and L. Wang, “A New Underwater Acoustic Signal Denoising Technique Based on CEEMDAN, Mutual Information, Permutation Entropy, and Wavelet Threshold Denoising,” Entropy, vol. 20, no. 8, Art. no. 8, Aug. 2018, doi: 10.3390/e20080563.

[23] Y. Yang, P. Fu, and Y. He, “Bearing Fault Automatic Classification Based on Deep Learning,” IEEE Access, vol. 6, pp. 71540–71554, 2018, doi: 10.1109/ACCESS.2018.2880990.

[24] J. R. Quinlan, “Induction of decision trees,” Mach. Learn., vol. 1, no. 1, pp. 81–106, Mar. 1986, doi: 10.1007/BF00116251.

[25] L. Breiman, “Random Forests,” Mach. Learn., vol. 45, no. 1, pp. 5–32, Oct. 2001, doi:

10.1023/A:1010933404324.

[26] T. T. M. Huynh, T. M. Le, L. T. That, L. V. Tran, and S. V. T. Dao, “A Two-Stage Feature Selection Approach for Fruit Recognition Using Camera Images With Various Machine Learning

Classifiers,” IEEE Access, vol. 10, pp. 132260132270, 2022, doi: 10.1109/ACCESS.2022.3227712.

[27] L. V. Tran, H. M. Tran, T. M. Le, T. T. M. Huynh, H. T. Tran, and S. V. T. Dao, “Application of Machine Learning in Epileptic Seizure Detection,” Diagnostics, vol. 12, no. 11, Art. no. 11, Nov. 2022, doi: 10.3390/diagnostics12112879.

[28] S. V. T. Dao, Z. Yu, L. V. Tran, P. N. K. Phan, T. T. M. Huynh, and T. M. Le, “An Analysis of Vocal Features for Parkinson’s Disease Classification Using Evolutionary Algorithms,” Diagnostics, vol. 12, no. 8, Art. no. 8, Aug. 2022, doi: 10.3390/diagnostics12081980.

[29] T. Le, T. Pham, and S. Dao, “Using Machine Learning to Predict the Defaults of Credit Card

Clients,” in Fintech with Artificial Intelligence, Big Data, and Blockchain, P. M. S. Choi and S. H. Huang, Eds., in Blockchain Technologies. , Singapore: Springer, 2021, pp. 133–152. doi:

10.1007/978-981-33-6137-9\_4.

[30] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey Wolf Optimizer,” Adv. Eng. Softw., vol. 69, pp. 4661, Mar. 2014, doi:10.1016/j.advengsoft.2013.12.007.