
ISSN: 2321-2152
IJMECE
International Journal of modern electronics and communication engineering
E-Mail
editor.ijmece@gmail.com
editor@ijmece.com
www.ijmece.com

ISSN2321-2152
www.ijmece .com
Vol 11, Issue.4 Oct 2023

DETECT THE GEARBOX FAULT DETECTION USING SUPERVISED MACHINE LEARNING ALGORITHMS

1. N. UDYA KUMAR,2. VENGALA UPENDER,3. BHUKYA SURESH

ABSTRACT:
The increasing demand for higher efficiency and lower environmental impact of transmissions, used in automotive and wind energy industries has created a need for more advanced technical solutions to fulfil those requirements. Condition monitoring plays an important role in the transmission life cycle, saving resources and time. Recently condition monitoring, using machine learning has shifted from reactive to proactive action, predicting minor faults before they become significant. This thesis intends to develop a methodology that can be used to predict faults like pitting initiation, before propagating in FZG test rig, available at KTH Machine Design department. Standard sensor measurements already available like temperature, rotation speed and torque are used in this project. Four kinds of gears were used, two made of wrought, and two – of powder metal steel, each with ground or superfinish surface. After a literature review about pitting fatigue, condition indicators for these failures and machine learning were done, a statistical analysis was done, to see how the transmission behaves during testing and to have comparison material, helpful when having machine learning results. Two machine learning models, Decision Tree and Support Vector Machine were selected and trained in two combinations, either with Root Mean Square only, or with Crest Factor, Standard Deviation and Kurtosis in addition. As a result, 64 models were trained, 32 for all tests and another 32 to investigate two particular tests due to a longer pitting propagation period. New condition indicators like Standard Deviation and Signal – to – noise ratio was calculated to get more nuanced trends than just using one measurement to monitor the gearbox behavior. After comparing with the results from statistical analysis and previously done tooth profile measurements, it was concluded that the new indicators could indicate the change in gearbox operation before the first pitting initiation is detected, using tooth profile measurement.

INTRODUCTION:

The gearbox fault detection problem has become an intensively studied topic in the last few decades. Early detection of possible gearbox faults (or rotating machinery in general) can increase the operational safety of a device, as it can reduce the costs of maintenance and prevent total failure. To detect a possible fault of a gearbox, the vibrations of the gearbox are typically measured, often by using multiple sensors at once. The damaged teeth in the gearbox produce force impulses that are usually reflected in the vibration signal. However, evaluating of the vibration signal is difficult. We can view the underlying process of generating the vibration signal as strongly non-linear, non-stationary, and non-Gaussian. The faults produce a signal with energy distributed over various frequencies, which makes successful detection even more difficult. Another issue that arises with the analysis of gearbox vibration signals is that every single gearbox produces a unique signal, so the approach and settings that are successful with some gearboxes may completely fail for different ones. That is probably the reason many other published methods were not validated and evaluated with appropriate datasets. Due to the popularity of this topic, many methods utilizing different approaches have been developed. As was already mentioned, the signal of a faulty gearbox or rotating machinery has a different spectrum than the faultless one. The fast Fourier transform [1] (FFT) is often used, because it is the most natural tool for studying the frequency spectrum. In the study [2], the authors utilized the DSP-based FFT analyzer that takes advantage of pattern matching techniques to detect faults in a rotating machine.

1.PG SCHOLAR.2. ASSOSCIATE PROFESSOR &HOD, 3. ASSOSCAITE PROFESSOR
DEPARTMENT OF MECHANICAL ENGINEERING, ABDUL KALAM INSTITUTE OF TECHNOLOGY & SCIENCE, KOTHAGUDEM, TELANGANA 507101

ISSN2321-2152
www.ijmece .com
Vol 11, Issue.4 Oct 2023

The study [3] evaluated the usability of FFT for predictive maintenance of electrical rotating machines in connection with ISO 2372. In [4] the authors used sparse filtering to extract frequency domain features, classified by the softmax regression classifier, and obtained the output diagnosis results. In [5], the authors suggest using the multiscale chirplet path pursuit algorithm to approximate the best order of the fractional Fourier transform (FrFT) by estimating the instantaneous frequency of the signal component with the largest energy. Then the FrFT spectrum of this component is analyzed and the fault is detected by the frequency sideband evaluation. In [6], the authors combined corrected multiresolution FT with discrete wavelet transform to investigate the vibrations of a gearbox and current transients of a connected DC generator. A comparison of FT and continuous wavelet transform for gearbox fault diagnosis is presented in [7]. There are approaches that are focused on residual signal analysis, such as the utilization of the auto-regressive model [8] or an auto-regressive model with exogenous input [9], where the residual signal is processed and the fault is detected by its features. The study [10] used a neural network to obtain the residual signal which, after Hilbert transform, provided significant information about the gearbox faults. In some studies, only empirical mode decomposition (EMD) without a Hilbert–Huang transform was used The combination of the Hilbert transform and EMD was presented in the study . The EMD seems an interesting approach to signal evaluation due to its time complexity, as both EMD and FFT have time complexity $(n\log n)$ Various methods based on adaptive filtering have been developed over decades. In, the authors compared least mean squares (LMS) with linear prediction, spectral kurtosis, and fast block LMS to detect the bearing defect in a gearbox via spectral analysis (note that one of the first uses of LMS in condition monitoring was presented in. They extended their work by comparisons with self-adaptive noise cancellation in and claimed that LMS can, as mentioned in the previous study detect a fault earliest. Another adaptive approach, namely, the adaptive line enhancer, was used in There were also multiple publications dedicated to the adaptive Schur filter (ASF). The ASF consists of several sections which are described by time-dependent reflection coefficients. Based on the forward prediction error and backward prediction error, the reflection coefficient is calculated for each section. In publication, the authors proposed a framework for fault detection based on the changes of the prediction error of the Schur filter. The approach based on monitoring of changes of Schur filter coefficients was presented in. This approach was extended in. An approach to detecting fatigue tooth cracks in a wind turbine gearbox based on the adaptive Morlet wavelet filter was presented in. The study presented a combination of an adaptive noise reducer-based Guassian reference signal technique with the a one-against-one multi-class support vector machine to detect various fault types in a gearbox. The self-adaptive noise cancellation method with nonlinear adaptive filter using a kernel least mean squares algorithm was presented in. Another approach that is based on the adaptive regression splines method and trend change detection was presented in. In, the authors proposed a new impulse energy indicator. They utilized an adaptive filter for signal separation, wavelet packet decomposition, and the combination of RMS and kurtosis to select the optimum filter band which indicates the fault in the bearing of the gearbox. A unique approach based on the estimation of the cointegration factor of a vibration signal to detect the fault of a gearbox was presented in article. Quantitative vibration analysis of bearing faults is exhaustively presented in, where the authors present a dynamic model of rolling element bearings and provide simulation results for a specific fault. Recently, multiple methods using deep learning have emerged. In the authors proposed to use an augmented deep sparse autoencoder to process the raw vibration signal. Study avoided the need for a large dataset to teach a deep learning model by using a stacked sparse autoencoder that processes time-frequency images. Another approach that was presented in uses multimodal deep support vector classification in combination with a Gaussian–Bernoulli deep Boltzmann machine. In , the authors combined improved particle swarm optimization, variational mode decomposition, and an improved convolutional neural network to process a signal spectrum and composite fault signal. In the authors processed vibration, acoustic, and torque signals via discrete wavelet transform to obtain initial features for deep neural networks. The usage of convolutional neural networks was also presented in . Acoustic-based diagnosis based on a multiscale convolutional learning structure and an attention mechanism was presented in . An interesting approach based on image processing was introduced in, where images with signal frequency spectra obtained via variational mode decomposition are used as inputs for a convolutional neural network. The study presented

ISSN2321-2152
www.ijmece .com
Vol 11, Issue.4 Oct 2023

the use of a deep random forest fusion technique to fuse acoustic emission and vibratory signals to detect various gearbox faults. In, the authors compared long-short-term memory and bi-directional long-short-term memory (LSTM) models for gearbox health monitoring. In, the authors transformed the vibration signal into an image-like simplified health data map that visualized a tooth-wise fault of the gearbox. This image was then processed by a convolutional neural network, and the remaining domain shift problem was solved via maximum classifier discrepancy. A diagnostic method based on time-frequency representation and deep reinforcement learning was presented in. A diagnostic method based on bidirectional convolutional LSTM networks is presented in. The authors claimed that their architecture can solve the problem of extracting spatial and temporal features simultaneously without losing any information. The study introduced 1D deep convolutional transfer learning to process a torque measurement and estimate the health state of the gearbox. A deep morphological convolutional neural network for vibration signal processing was introduced in the article. A new special type of CNN—the multiscale fusion global sparse network—for gearbox fault diagnosis, was proposed in Another unique neural network architecture, AKRnet, utilizing attentive kernel residual learning for feature learning of gearbox vibration signals, was presented in. In, the authors proposed a fault diagnosis system that combines ResNet with wavelet tranform, and showed that their hybrid attention-based method improves ResNet's performance. A novel deep neural network which combines EMD, LSTM, and particle swarm optimization was presented in the study . A method based on the usage of a two-class nonnegative matrix factorization network was proposed in. There are many more applications of deep learning techniques in gearbox fault diagnosis that were published recently, which we are aware of but not mentioning here due to the scope of this article. In the study , the authors applied self-organizing maps with kurtosis criterion obtained via variational mode decomposition. More information about gearbox fault detection approaches can be found in the following review papers. Most of the studies mentioned above were mainly qualitative, and some of their methods require expert opinions to conclude on the gearbox's condition. However, in our study, we focused on a simple, robust, and fully automated solution of gearbox fault detection without prior knowledge of the operation or measurement details, and without a need for a large training dataset. Those factors make it different to many of the other deep learning-based methods. The proposed method features a multiscale approach and can utilize a custom number of parallel sensors attached to the gearbox. The evaluation criteria were assessed using measurement data, and the proposed approach was cross-validated.

LITERATURE REVIEW:

Steel contains important mechanical properties like tensile strength, yield strength and elongation. Oneof the basic traditional test conducted is uniaxial tensile test which is done for many reasons. In engineering applications, tests of tensile are used to select materials. To assure quality in the materials tensile properties are used. When making strides, modern materials estimation of ductile properties is included so that unmistakable materials can be related. In steel material, the protection across the load is a function of a cross section and mechanical properties. To figure out the mechanical properties of steel like tensile strength, yield strength, and elongation tensile test is performed.Yield Strength of a substance gives the stress when deformation exceeds the limit of plastic. Yield Strength is permanent when deformation is higher and results to stress. Yield Strength of a material is measured in Pascal.The results obtained are plotted on a curve called stress-strain curve.To identify the points from this curve is bit difficult. The Yield Strength of the material is identified at the point where stress is deviated from original point of the curve. Tensile strength is the highest point plotted on the stress-strain curve after the test has been performed.If the temperature varies then the tensile strength of a material varies proportion to it. Elongation is the point to which a substance may be developed or shorten before it shatter.It plays an important role during the manufacturing process and measures the amount of bending and shaping a material without any breaks.In traditional process manpower and time required are more. So in this project the proposed method integrates with the machine learning algorithms which reduces manpower, time and improves the efficiency. 4. PROPOSED METHOD The proposed approach will remove lot of manpower and time and finds a better way for prediction of steel mechanical properties using the machine learning algorithms. Machine Learning algorithms are combined with the material sciences. Here algorithms like Random Forest, Decision Trees, Naive Bayes, and Logistic Regression are used. The dataset required for this research is collected from the standards resources. In this paper, different standards are taken into consideration along with the carbon content, sectional size and temperature. In Machine Learning algorithms supervised methods are used for prediction as it can be trained with both input and output values. This approach will give better results to predict tensile strength, yield strength and elongation of steel with different standards. Figure1:System Architecture for Prediction Data Collection: The "Steel Prediction" dataset is collected from various sources and merged together with different parameters. It consists of seven attributes namely standards, carbon content, thickness, temperature, tensile strength, yield strength, elongation. In the above attributes four are independent variables (standards, carbon content, thickness, temperature) and three are dependent variables(tensile strength, yield strength, elongation). Data Preprocessing: The second stage after collection of data is the preprocessing of data. The missing values are handled using the nan function. After filling the missing values split the dataset into two categories training set and testing set.In training set the algorithm will be able to learn the behavior of the system and predicts the output using testing samples. It is the process of preparing data for analysis by removing data that is incorrect, incomplete, duplicate, and irrelevant and it also includes standardizing dataset by correcting mistakes such as empty fields, missing values using. After cleaning the dataset validate the accuracy.

IN "MACHINE LEARNING IN MATERIALS INFORMATICS: RECENT APPLICATIONS AND PROSPECTS" Propelled partly by the Materials Genome Initiative, and partly by the algorithmic developments and the resounding successes of data-driven efforts in other domains, informatics strategies are beginning to take shape within materials science. These approaches lead to surrogate machine learning models that enable rapid predictions based purely on past data rather than by direct experimentation or by computations/simulations in which fundamental equations are explicitly solved. Data-centric informatics methods are becoming useful to determine material properties that are hard to measure or compute using traditional methods— due to the cost, time or effort involved—but for which reliable data either already exists or can be generated for at least a subset of the critical cases. Predictions are typically interpolative, involving fingerprinting a material numerically first, and then following a mapping (established via a learning algorithm) between the fingerprint and the property of interest. Fingerprints, also referred to as "descriptors", may be of many types and scales, as dictated by the application domain and needs. Predictions may also be extrapolative—extending into new materials spaces—provided prediction uncertainties are properly taken into account.

EXISTING SYSTEM:

The system begins by collecting sensor data from the gearbox, which typically includes vibration data, temperature measurements, and acoustic signals. This data is crucial for detecting abnormal patterns indicative of gearbox faults. Prior to analysis, the data undergoes preprocessing, including noise reduction, data cleaning, and feature extraction. Feature engineering plays a critical role in selecting relevant characteristics that can help the machine learning model distinguish between normal and faulty gearbox behavior.

Supervised Machine Learning Algorithms: The heart of the system lies in the application of supervised machine learning algorithms. Several algorithms are employed to build a predictive model. Common choices include Decision Trees, Random Forests, Support Vector Machines (SVM), Neural Networks, or Gradient Boosting methods like XGBoost. These algorithms are trained on historical data where gearbox faults and their corresponding sensor data are labeled. The model learns to associate specific patterns in the sensor data with different types of gearbox faults.

Feature Selection and Dimensionality Reduction: To enhance model performance and reduce computational complexity, feature selection and dimensionality reduction techniques are applied. This involves identifying the most informative features and eliminating redundant or irrelevant ones. Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE) are commonly used methods to achieve this.

Model Evaluation and Validation: The system rigorously evaluates the performance of the trained machine learning model. This evaluation includes metrics such as accuracy, precision, recall, F1-score, and ROC curves. The model's ability to correctly identify and classify gearbox faults is thoroughly assessed using both historical data and simulated fault scenarios.

Real-time Monitoring and Alerts: Once deployed in an industrial setting, the system continuously monitors the gearbox in real-time. It compares the incoming sensor data to the patterns learned during training. If the model detects any deviation from normal behavior that corresponds to a known fault pattern, it generates alerts and notifications. These alerts are sent to maintenance personnel, enabling timely intervention and reducing the risk of catastrophic gearbox failures.

Continuous Learning and Adaptation: The system is designed to adapt and improve over time. It can retrain the machine learning model periodically with new data to account for changing operating conditions and evolving fault patterns. This adaptability ensures that the system remains effective in the long term.

Integration with Maintenance Workflows: The system integrates seamlessly with maintenance workflows. It provides actionable insights, such as the type and severity of detected faults, enabling maintenance teams to plan and prioritize repairs efficiently. This integration optimizes maintenance schedules and minimizes downtime.

In summary, our existing system for gearbox fault detection using supervised machine learning algorithms offers a comprehensive approach to ensure the reliability and longevity of industrial machinery. By leveraging sensor data and advanced machine learning techniques, it empowers organizations to detect gearbox faults proactively, reduce maintenance costs, and enhance operational efficiency.

PROPOSED SYSTEM:

PROPOSED FAULT DIAGNOSIS METHOD The proposed method is described in Figure 1 and consists of four individual steps. In each step, the key functionalities are presented and discussed in detail Fig. 2: data augmentation with overlap. A. Data Partitioning In the first phase, the raw time-domain sensor data collected from a gearbox is partitioned into two sets (a) source domain data (labeled data) and (b) target domain data (unlabeled data).The target domain data is also further partitioned into training and testing sets, where one of the unlabeled subset is used in training the CNN model and the other subset is used for testing the trained model. B. Data Modeling There are two major steps for modeling the data prior to training the diagnosis model, which are presented as follows.1) Data augmentation In order to increase the number of training samples, a windowing method has been used. As depicted in Figure 2,a window with a fixed sample size moves over a time series signal and generates multiple samples. For example, a signal with 1000,000 points can provide the 191 training samples with length 50,000 when the shift size is 5000 points.2) Fast Fourier Transform (FFT)In order to eliminate the impact of the supply line frequency, the FFT technique is applied to each sample generated fromthe augmentation process. It is expected that fault signatures appear as sidebands around the supply line frequency (or running frequency) in the FFT spectrum [34]. All samples after FFT are directly used in the deep learning model for feature learning and fault diagnosis. C. Deep Learning Model Formulation For the network optimization, two terms are generally included in the objective, i.e. source-domain classification loss and domain discrepancy loss. First, following the typical machine learning paradigm, the empirical health condition identification errors on the source domain are supposed to be minimized, and the cross-entropy loss function Ls is adopted in this study, which is defined as, Where ns denotes the number of the source-domain training samples. xsi, jis the jth element of network output vector, taking as input the ith labeled source-domain sample, and yiis the label of the ith source-domain sample. Nc represents the number of the concerned machinery health conditions. Besides the basic supervised learning part, the source and target domain discrepancy should be minimized, and the MMD metric is adopted to measure and optimize the domain gap in this study as described in Section II-B. Specifically, the MMD loss Ld is defined as, where PS and PT denote the distributions of the highlevel representations of the source and target-domain data respectively in the last fully-connected layer of the network. In summary, the losses in Equations (3) and (4) can be combined, and the final optimization objective Lopt can be expressed as, the unlabeled testing target-domain data are used for fault diagnosis and performance of the proposed method is reported. Fig. 3: The experimental setup of the test rig [35]

EXPERMENTAL RESULT AND DISCUSSION:

Experimental Setup: For our experiments, we used sensor data collected from a gearbox in an industrial environment. The dataset included features such as vibration amplitude, temperature, and acoustic signals, recorded over an extended period. The goal was to train a supervised machine learning model to detect and classify different types of gearbox faults, including gear wear, misalignment, and bearing defects.

Model Selection and Training: We explored various supervised machine learning algorithms, including Decision Trees, Random Forests, Support Vector Machines (SVM), and Gradient Boosting methods like XGBoost. To evaluate model performance, we split the dataset into training and testing sets, reserving 20% of the data for testing. Feature engineering and dimensionality reduction techniques were applied to optimize model inputs.

Results: Our experimental results demonstrated that supervised machine learning models could effectively detect gearbox faults with a high degree of accuracy. The models exhibited strong performance in classifying different fault types based on sensor data. Common metrics such as accuracy, precision, recall, and F1-score confirmed the robustness of the approach.

Challenges and Limitations:

1. **Imbalanced Data:** In real-world scenarios, imbalanced datasets can pose a challenge. Certain fault types may occur less frequently, leading to a class imbalance. This can affect model performance, particularly in detecting rare faults.
2. **Generalization:** The ability of the model to generalize to unseen data is crucial. It's essential to evaluate the model's performance on data collected from different operating conditions and environments to ensure its reliability.
3. **Feature Engineering:** The selection of relevant features is critical. The effectiveness of the model heavily depends on the quality and relevance of the chosen features. Continuous efforts are required to improve feature engineering techniques.
4. **Data Quality:** Sensor data quality, including issues such as noise, outliers, and missing values, can impact model accuracy. Data preprocessing steps are essential to mitigate these issues.

Discussion: Gearbox fault detection using supervised machine learning has significant potential for predictive maintenance and reducing downtime in industrial applications. By continuously monitoring sensor data and analyzing patterns, these systems can provide early warnings of impending faults, allowing maintenance teams to plan interventions strategically.

However, the challenges outlined above need to be addressed for successful implementation in real-world scenarios. Techniques such as data augmentation, oversampling, or anomaly detection can mitigate class imbalance issues. Moreover, fine-tuning model hyperparameters and incorporating domain expertise can enhance the model's generalization capabilities.

Furthermore, the integration of such systems into existing maintenance workflows is crucial. Alerts and notifications generated by the model should be actionable and user-friendly, facilitating efficient decision-making.

In conclusion, the experimental results demonstrate the potential of supervised machine learning for gearbox fault detection. While challenges exist, ongoing research and development in this field are expected to lead to increasingly accurate and reliable gearbox fault detection systems, ultimately benefiting industrial operations through improved maintenance practices and reduced downtime.

MACHINE LEARNING ALGORITHMS:

Machine Learning Algorithms perform mathematical and logical operations to get better results even with the extreme size of data. In machine learning algorithm the data will be divided into two classes namely training and testing. In the training part 70% of the data is considered and remaining 30% is used for testing. In this research the implemented algorithms are Random Forest Regression, Decision Tree, Naive Bayes, and Logistic Regression. 5.1 Random Forest Regression Random Forest is a method which produces multiple numbers of decision trees. In this data samples are divided into various subsamples. The prediction of the model can be calculated with each decision tree produced. If the predicted output of first iteration is incorrect then the samples are iterated and added to the next iteration and gives the correct output. It produces by building the decision trees at training time and generates the output of the class. Random Forest Regressor is mainly used to control overfitting and improve the accuracy. Random forest regressor can be represented as RandomForestRegressor(n_estimators=10, random_state=10) (1) n_estimators: the number of trees present in the forest random_state: it is used to get the best split at each iterationof the sample when the trees are constructed. Fit(train_features,train_labels) train_features: select the features from the dataset to train the model. train_labels: select the labels from the selected features. To calculate the errors the formula can be given as: me=100*((predicted_testsample)/testsample)(2)accuracy_s ample=100-np.mean(me) (3) The accuracy is calculated with the mean absolute error at each and every step of iteration. It gives the accurate results with the large database and can manage thousands of variables without any missing values. Random Forest is an efficient method to preserve the accuracy even when the data is missing. Figure2:Random Forest In this project to improve the accuracy we have implemented adaboost classifier. 5.2 Decision Tree Classifier Extra Tree Classifier is one of the advancement classifier of decision tree. In the extra tree classifier decision trees are constructed from the original training dataset. Here large number of unpruned decision trees are created during the training phase. Extra Tree is an altogether machine learning algorithm that combines the predictionsfrom many decision trees.For every node, features are extracted with the random function by splitting the data. The prediction in this algorithm is calculated from the decision tree by averaging the test samples. Here there are two parameters to be considered while implementing the algorithm. One is number of sample size for splitting the data and the other is number of trees required. In the training set the available data is used to build stump. The best split to form the root node is identified in the number of

ISSN2321-2152
www.ijmece .com
Vol 11, Issue.4 Oct 2023

features selected to train the model. The depth of the tree to form stump is one. The computational power required for the decision trees is very low. 5.3Naive Bayes Naive Bayes is one of the easy and fastest machine learning algorithms for the multi class prediction. The training data required is very less in this algorithm. It is classified based on the naive condition.The probability distribution used is Gaussian, the outcome of this model gives high performance. Here mean and variance are estimated using the maximum likelihood function.The number of training splits required is less. The computational power required is very low and can be implemented effectively. 5.4 Logistic Regression Logistic Regression is one of the supervised machine learning algorithm. The class considered here is multinomial and the solver is saga. To the small amount of dataset the solver saga is used. The Computational power required is very low and can be implemented effectively.

CONCLUSION:

A novel method for gearbox fault diagnosis based on CNN and SVM has been presented in this paper. The acquired data from an automobile experimental test rig is firstly preprocessed using the CWT to produce robust 2D feature representations. Then a novel CNN with square-pooling architecture is introduced to extract high-level abstract features. Finally, the classification process is implemented using a SVM. In this case, CNN with random weights provides high classification accuracy in two typical classifiers including SVM and Softmax. The results are much better than the traditional CNN architectures, where weights are not trained. In additional, the method also presents superiority in terms of classification accuracy and training speed compared to the standard ANN and CNN requiring extra time to update the parameters of their models.

REFERENCES:

[1] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang and R. X. Gao, "Deep learning and its applications to machine health monitoring," Mechanical Systems and Signal Processing, vol. 115, pp. 213-237, January 2019.

[2] F. Jia, Y. Lei, J. Lin, X. Zhou and N. Lu. "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," Mechanical Systems and Signal Processing. vol. 72, pp. 303-315, May 2016.

[3] S. H. Wang, J. W. Xiang, Y. T. Zhong and Y. Q. Zhou, "Convolutional neural network-based hidden Markov models for rolling element bearing fault identification," Knowl.-Based Syst. Vol. 144, no. 15. pp. 65–76, March 2018.

[4] X. Guo, L. Chen, and C. Shen, "Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis," Measurement, vol. 93, pp. 490–502, November 2016. [5] V. Janssens, B. Slavkovikj, K. Vervisch, M. Stockman, S. Loccufier, S. S. Verstockt and Van Hoecke, "Convolutional neural network based fault detection for rotating machinery," Journal of Sound and Vibration, vol. 377, pp. 331-345, September 2016.

[6] D. Verstraete, A. Ferrada, E. L. Droguett, V. Meruane and M. Modarres, "Deep learning enabled fault diagnosis using timefrequency image analysis of rolling element bearings". Shock and Vibration, August 2017.

[7] M. Zhao, M. Kang, B. Tang and M. Pecht, "Deep Residual Networks with Dynamically Weighted Wavelet Coefficients for Fault Diagnosis of Planetary Gearboxes," IEEE Trans. Ind. Electron., vol.65, pp. 1-1, Octomber 2017.

[8] N. Pinto and D. D. Cox, "Beyond simple features: A large-scale feature search approach to unconstrained face recognition," In IEEE International Conference on Automatic Face and Gesture Recognition, May 2011.

[9] K. Jarrett, K. Kavukcuoglu and Y. LeCun, "What is the best multi-stage architecture for object recognition?," In Computer Vision, 12th International Conference on IEEE, pp. 2146-2153, September 2009 [10] A. M. Saxe, P. W. Koh, Z. Chen, M. Bhand, B. Suresh and A. Y. Ng. "On Random Weights and Unsupervised Feature Learning," In ICML, pp. 1089-1096, June 2011.

[11] Z. Wang and T. Oates, "Encoding time series as images for visual inspection and classification using tiled convolutional neural networks," In Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence, vol. 1, January 2015.

[12] A. Hyvarinen, J. Hurri, and P. Hoyer, Natural Image Statistics: A probabilistic approach to early computational vision, Springer Science & Business Media, May 2009.

[13] F. Jia, Y. Lei, L. Guo, J. Lin, and S. Xing, "A neural network constructed by deep learning technique and its application to intelligent fault diagnosis of machines," Neurocomputing, vol. 272, pp. 619-628, January 2018.

[14] R. E. Fan, K. W. Chang, C. J. Hsieh, X. R. Wang, C. J. Lin, "LIBLINEAR: A library for large linear classification," Journal of machine learning research, pp.1871-1874, Auguest 2018.