Journal of Computers Vol. 36 No. 2, April 2025, pp. 1-13
https://doi.org/10.63367/199115992025043602001

A Normalized Failure Detection Model Using Deep Learning for Improving the Outcomes of Industrial Production

Changrui Hou and Xuanming Wan*

School of Intelligent Manufacturing , Liuzhou Railway Vocational Technical College,
Liuzhou, Guangxi, China
hcr1030@163.com, wxm@ltzy.edu.cn

Received 27 February 2024; Revised 21 March 2024; Accepted 21 March 2024

Abstract. Industrial production control monitors quality through fault diagnosis, debugging, and interactive controller improvements. The fault diagnosis levels are intense, pursuing a periodic cycle that requires varying operation and control factors. This article introduces a Deep Learning-based Normalized Failure Detection Model (DL-NFDM) to address the issues above. This model exploits deep learning methods such as Faster Region processing Convolutional Neural Networking (Faster-R-CNN) to identify repeated control errors over production intervals. The harmonized intervals are grouped as single entities to prevent additional control faults. In this learning process, the asynchronous control and operation factors avoid failures of control and fault diagnosis. The prioritized losses based on production intervals are segregated with the fault factor; hence, the production depreciation at different intervals is confined. This learning process is recurrent based on fault intensity, production decrease, and time-prolonging delay from different outputs. Therefore, the training is dissolved and reformed for further failure handling. The independent intervals are utilized in the consecutive fault detection instance for training the learning model. The proposed model's performance is validated using the metrics of fault detection, detection time, and production errors. The suggested model has a detection time of 0.231 seconds and an accuracy of 97.56%. The rate of production errors reduces as the number of iterations increases, and the rate at which faults are detected is approximately 95.61%. The average precision rate is 93.46%, and the rate at which they are detected is 93.46%.

Keywords: deep learning, fault detection, industrial production, CNN, quality control, Faster-R-CNN

1 Introduction

Mechanical, electrical, technological and informational electromagnetic subsystems and devices are commonplace in manufacturing processes. The manufacturing industry is achieving intelligent transformation through automation, artificial intelligence (AI), and the Internet of Things (IoT) [1]. There has been a steady rise in the importance of reliability concerns in determining the viability of many cutting-edge industrial systems [2]. Any disruption to the system's normal functioning can have far-reaching consequences, including diminished performance and complete collapse in the worst situations [3]. Faults might put lives at risk and lead to other disastrous outcomes. That's it's so important to detect problems as soon as possible so that you can do preventative maintenance and avoid problems [4].

Model-based, signal-based, knowledge-driven and hybrid approaches are the broad categories into which fault diagnostic techniques fall. In light of this, conventional approaches to fault diagnosis are often grounded in mechanisms, feature frequencies, or fault extraction of features [5]. Diagnosing the defect in industrial production with a complicated structure using standard subjective fault detection methods is challenging due to their reliance on practical experience and expert knowledge [6]. There has been progress and success in defect diagnosis, particularly in model-based and signal-based methods. The use of Deep Learning (DL) as the core of a combination of different extracting features approaches for defect detection is gaining popularity for quality control [7]. The new avenue of discovery for fault diagnosis and health administration is made possible by this effective instrument for data processing and mechanical system fault identification [8].

The support vector machine, or SVM, and artificial neural network (ANN) are two examples of classic ma-

* Corresponding Author

chine learning techniques that have successfully detected production errors in industrial settings [9]. EMD (Empirical Mode Decomposition) and Wavelets package decomposed were used together for obtaining information; ANN was used for early failure identification. There have been many attempts to identify bearing faults through unique methodologies, such as combining multiscale fuzzy measure entropy, infinite feature selection, and SVM analysis [10]. Modern failure detection approaches have their strengths and weaknesses, nevertheless. On the one hand, extracting features calls for expertise in a wide range of signal processing techniques and extensive hands-on building experience; on the other hand, the connection between the two processes is overlooked because the two processes are handled independently [11]. However, it comes to model training, a simple simulation is used to analyse the complicated conversion between these communication and health statuses, resulting in a model lacking evaluation capacity and generalization performance confronted with industrial outcome data [12].

The deep learning (DL) hypothesis in Science provided a new theory and technique of artificial intelligence, which inspired a tidal wave of study across several disciplines, including the Normalized Failure Detection Model fault. It was believed that DL was a machine learning technique, and its breakthroughs in the areas of reputation, audio, video, and text analysis were examined at length. DL has been proven to offer several study and implementation opportunities [13]. Multilayer nonlinear network training has learnt Possible characteristics of specimens, enhancing DL's capacity for classification or prediction. Neural trust network is one of the most researched DL techniques [14]. In addition to fixing a problem with standard vibration analysis, provided a potentially valuable intelligent tool for identifying issues, which is crucial for successful manufacturing output. An embedded kernel machine for extreme learning was suggested and used to drive the gearbox, rotor, and engine joint to provide effective failure detection through visualization [15]. Due to DL's increasing popularity in various detection applications, experts in the equipment industry are paying much more consideration to the DL-based article proposing a Deep, proactive learning-based Normalized Failure Detection Model (DL-NFDM) for fault detection during production for quality control. First, pairs of fault samples were fed into the same feature extraction network. Then the vibratory signal information was mapped to the low-dimensional characteristic space using a Long and Short-Term Memory (LSTM) network and CNN. After that, the characteristics from the extracted samples were compared using the connection assessment network, and the results were fed into an analysis network for fault detection.

The main objective of the article include:

1. The proposed approach uses the Faster-R-CNN deep learning model to detect cyclical control issues during manufacturing downtime. The Deep learning-based Normalized Failure Detection Model (DL-NFDM) for quality control is introduced in this article as a solution to these issues in industrial production.
2. The model uses DL techniques like Faster Region processing Convolutional Neural Network (Faster-R-CNN) to regulate control blunders that occur at regular intervals throughout production.
3. Fault detection, detection time, and production mistakes are used to verify the efficacy of the suggested model. The proposed model speeds up detection and increases precision.

The rest of the paper is as follows. Section 2 discusses previous research surveys. The section 3 describes the research methods. Section 4 describes experimental implementation settings and shows validation findings using an actual fault detection dataset. Section 5 concludes this article.

2 Related Work

Mouzenidis et al. [16] proposed Visual object identification is essential for many manufacturing processes, including robot navigation, quality assurance, and assembly of products. High-accuracy, resilient, and generalizable artificially intelligent object detection algorithms are essential in today's manufacturing environments (VAE-Faster-RCNN). The suggested method uses a VAE encoder-decoder network and a compelling attention layer to increase the resilience and generalization capacity of the initial Faster R-CNN technique. The recommended process yields a considerable efficiency 96.34%, as demonstrated by experimental results on two object identification datasets: the popular RGB-D Washington information and the QCONPASS database of industrialized objects introduced in the study.

Im et al. [17] proposed a comprehensive RODIS (Regional Object-Defect Inspection System) is a R-CNN based system for inspecting objects for defects in a given region. It is difficult to establish a manufacturing imaging system for the rear outer, and there is no precedence for a fully independent defect-inspection technique for the main outer. When the steel cutting industry relies on human inspection of the surface, cost and quality are both negatively impacted. Things discovered via trial and error when setting up RODIS in the field. In comparison with different models, Mask R-CNNs trained with Res-net-50-FPN perform exceptionally well, with average precision (AP) values of 71.63 (Object Detection) and 86.21 (Object Segmentation).

Lyu et al. [18] presented bearing fault diagnostics for motors in a high-noise industrial setting as a common problem, and the presented method is an effective solution. To address this problem, the paper proposes an innovative deep learning approach, the Rare-class Sample Generator-based smart bearing fault diagnosis method (RSG-SBFDM), based on the utilization of the remaining developing Unit, soft thresholding, and global context to resolve the complicated mapping connection between sound vibrations and various bearing faults. The suggested RSG combines soft threshold and international context working processes to reduce noise and extract features effectively. The proposed approach has been shown to have an average defect identification accuracy of 98% in experiments.

Zhao et al. [19] proposed a novel class unbalanced fault diagnostic framework for the bearing-rotor system using The NCVAE-AFL stands for the Normalizing Conditional Variational Auto-Encoder with Adjusting Focusing Losses. An NCVAE algorithm designed to enhance learning characteristics from data plays a crucial role in this diagnosing strategy. Meanwhile, the NCVAE model has a newly developed Adaptive Focus Loss (AFL) function that helps to balance the complexity of diagnosis across many different data classifications by focusing the training on a subset of data representing medical issues that are particularly challenging to categorize. The findings of the diagnostic show that the suggested diagnostics framework is more accurate (98.12%) and stable than the most recent methods for handling class-imbalanced failure data in mechanical structures.

Del et al. [20] proposed the auto encoders in the novel-object recognition process in depth. The present a basic framework for construction, including a localizer for pinpointing the most unambiguous signals and an identification layer for recognizing and segmenting the operational situations. Data from a test rig assesses the efficacy of four distinct implementations that use various deep-learning models. The results demonstrate that the automatic encoders perform better than the state-of-the-art standards and validate the efficiency of the architecture.

Jiang et al. [21] proposed state-of-the-art machine vision techniques to detect problems in the production line due to cyberattacks. In the paper, a novel Variation of Fuzzy Auto-encoder (VFA) approaches. Nonlinear transformation via deterministic functions can be used to its full potential with the help of fuzzy entropy and Euclidean vague similarity measurement, resulting in a completely realistic vision system. In conclusion, the suggested system successfully evaluates and classifies abnormalities in a highly complicated setting.

Abidi et al. [22] proposed the Jaya method is combined with Sea Lion Optimization (SLnO) to achieve optimal feature selection. The wide range over which the prediction values fall makes it challenging for machine learning and deep learning to get reliable results. Therefore, the decisions concerning the prediction network are made using a support vector machine (SVM). The SVM determines which network is suitable for making predictions over the relevant interval. Finally, a Recurrent Neural Network (RNN) is used to make the prediction. The weight of the RNN is optimized by a hybrid algorithm called J-SLnO. The suggested model is able to accurately predict the future condition of elements for service organizing, as demonstrated using an examination using two datasets.

Suawa et al. [23] proposed a Deeper Convolutional Neural Networks (DCNNs), LSTM, and CNN-LSTM are proposed for data-level sensor fusion for Brushless Direct Current (BLDC) motor failure diagnostics because they can automatically glean beneficial knowledge from incoming data. The results show that sound signals alone are more effective than vibrations alone. The model's precision is increased from 93.7% to 98.9% for the DCNN, a CNN-LSTM, and LSTM methods respectively after combining the data. This performance has never been achieved in analysing BLDC motor faults without extracting and fusing characteristics. These results show that deep learning may be applied to raw data from various sensors to produce valuable results without spending resources on feature extraction and data fusion.

Although various techniques are available for making maintenance forecasts, these methods all have their drawbacks. Therefore, a brand-new network based on DL should be created. Because DL is now widely used in many industry domains, experts in the industrial research field are paying a lot more attention to DL-based intelligent defect diagnosis. As a result, our analysis will centre on how to properly diagnose problems with spinning machinery. The use of deep neural networks for fault diagnostics will be a significant focus.

3 System Methodology

3.1 Deep Learning based Normalized Failure Detection Model Structure

Fig. 1 shows the overall structure of the DL-NFDM process. Machines are used in several critical technological uses across various sectors. These include heating systems, mines, building sites, mineral extraction, and high-pressure environments. During breakdown, during the alignment, balancing, loosening or bearing and rotating machinery wear issues, it might have a domino effect that leads to the collapse of the entire machine on production. Research into vibrational computation is booming since it's important to identify healthy and unhealthy vibration signals collected from machinery. The mechanisms, specific frequencies, and feature extraction used in normalized identifying defects approaches are crucial to problem-solving. There is currently a lack of generalizable regulator design, no single solution for systems integration, and a focus on secondary signals rather than fundamental ones in hydraulic systems problem detection.

The collected signals are analysed using a data-driven approach to determine critical defect features. To assess the state of a machine, these methods can be tailored to the collected data and used to obtain private details. The extensive research on the effectiveness of CNN and LSTM, two DL-based methodologies, in rotational machineries like generators and compressors. The efficacy of these novel diagnostic approaches is highlighted. Ideas and guidance for extracting signals diagnosis research and implementation in manufacturing machinery for quality assurance. Using a combination of supervised and unsupervised learning, a DL network is an essential machine-learning technique for learning complex mathematical models at multiple layers. Instead of manually extracting fault features, deep learning algorithms may adaptively learn the data structure from the underlying information via a large number of abnormal modifications involving complex variability factors. The output of the feature extraction signal is given to the proposed DL-NFDM.

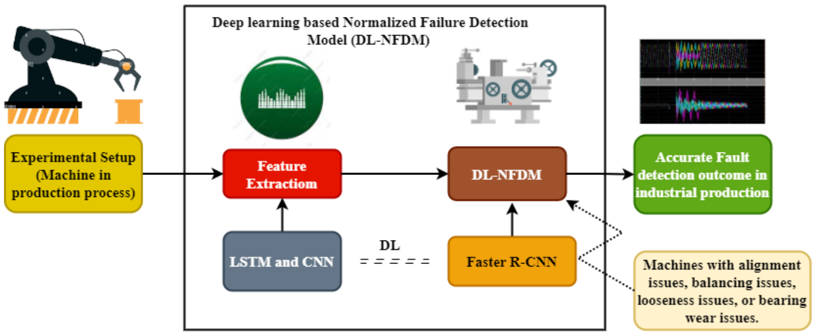

Fig. 1. Structure of normalized failure detection model using deep learning

An enhanced version of the Faster R-CNN method and a quality-focused variant of the technique that uses normalization on fault identification is suggested. Then, to better express the characteristics of fault identification in industrial production and to solve the challenges associated with many types of defects and parameter forms, characteristic structure networks the ResNet50 system, which uses a shape-variable CNN, can be used to gather imperfections of conceptual feature mappings. This paper's algorithm uses the more refined Region of Interest Alignment (ROI Align) technique in place of the more straightforward ROI Pooling methodology to better pinpoint the precise location of defects. The simulation is then fine-tuned to better focus on flaws while suppressing the distinguishing characteristics of a complicated backdrop using a refined concentration region prediction network.

As a result, this research offers a single-dimensional convolution neural network-long short-term recall (1D CNN-LSTM) model to facilitate the recognition of epileptic fits by evaluating vibration signals—initial processing and normalization of vibration signal data collected during industrial manufacturing. The industrial manufacturing vibrations sequencing data is then normalized, and a 1D convolutional neural network (CNN) is developed to obtain the characteristics. The LSTM layers process the extracted features to recover time-related elements further. Finally, multiple fully connected layers receive the resultant information for seizure epileptic detection.

3.2 Feature Extraction

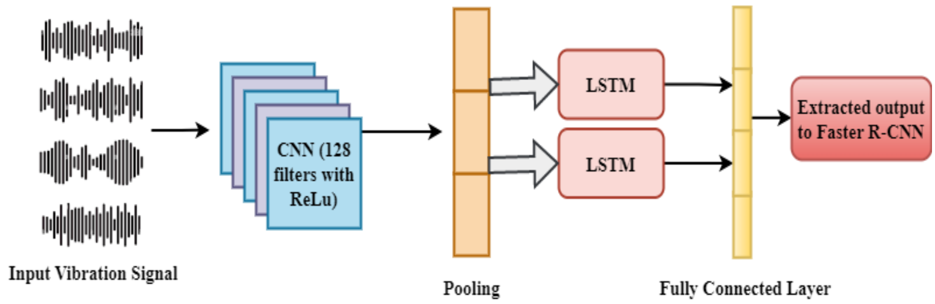

Fig. 2. Feature extraction

The recommended CNN-LSTM system has the following layers: input data, four convolutional layers, one pooling, LSTM, a FC (Fully Connected) layer, and a soft-max outcome. Fig. 2 depicts the network architecture in its finest form. To begin, the suggested model takes a vibration signal from the sensor attached to the industrial machine data as its source of information, with the input data having a shape of 178 1. To extract abstract characteristics from unprocessed signal data, the input is first processed by the first convoluted layer (Conv Layer1), which uses 128 1D convolutional kernels with a 3 1 form and a stride of 1. A ReLU (Rectified Linear Unit) activation layer comes next to the convolutional layer to add non-linearity to the suggested model potentially. A formal mathematical description of the convolutional process and the ReLU stimulation is shown in Equation 1.

$$x_j{}^{m-1} = \delta(\sum_{i=1}^{N-1} conv\left(w_{i,j}{}^{m}, y_i{}^{m-1}\right) + l_j{}^{m-1} \tag{1}$$

Where x_j^{m-1} is the j^{th} characteristic map of the $(m\text{-}1)^{th}$ layer, x_j^{m} is the j^{th} characteristic map of the m^{th} layer, and $w_{i,j}{}^{m}$ is the adaptable multilayer kernel. The overall size of the map of characteristics in the m^{th} layer has a smaller dimension than that in the $m\text{-}1^{st}$ layer, denoted by the notation N-1; conv1D denotes the1D convolution procedure without zero-padding; To prevent excess fitting, using the ReLU activation function, denoted by δ, and the distortion of the j^{th} characteristic map in the m^{th} layer, denoted by l_j^{m-1}.

The convolutional and stimulation processes result in 128 map features of size 176 1. Then, the maximum-pooling layer is applied to the Conv Layer1 output. The following is an equation 2-based explanation of the 1D, a maximum pooling procedure.

$$\rho_i^{\alpha} = (\rho_i^{\alpha'} : \alpha \le \alpha' < \alpha + a) \tag{2}$$

Where ρ_i^{α} is the α^{th} neuronal in the i^{th} characteristic map preceding the maximum pooling operation, $\rho_i^{\alpha'}$ is the α'^{th} neuronal in the i^{th} information map after the procedure has been performed, and a is the size of the sharing window. The pooling window size and the distance between rows of windows are set to 2 in Pooling Layer 1. It may speed up the training process by decreasing the parameters needed to train the suggested model. 64 feature maps of size 88x1 are produced as a result of the pooling procedure. The next step is to use the three convolutional layers to extract additional high-level information that might aid the classification. ConvLayer2, ConvLayer3, and ConvLayer4 each include 128 kernels in the form of a 3x1 matrix, 516 kernels in the form of a 3x1 matrix in ConvLayer3, and 1023 kernels in the form of a 3 1 matrix in ConvLayer4. Like Conv Layer1, ReLU is used for nonlinear activation, and the convolution procedure is the same. Following the 1D convolutional layers, the 1023 feature maps acquired will be input into a single FC layer of 256 neurons, where dropout will be applied to the final output. Overfitting problems may be alleviated by dropout, and FCLayer1 can concatenate the result of the convolution layers and lower the size of feature maps to fit the input of LSTM layers. To circumvent the long-term dependence issue that plagues regular RNNs, the output features are sent via FC Layer1 before being fed into LSTM layers.

The LSTM cell consists of four gates: the cell's internal state entrance, the gate for forgetting, the inward data gate, and the resultant data gate. LSTM work together to remember what has come before and boost each other's capacities for extracting insights from vibration time series data. Each LSTM layer, Layer1 and Layer2, consists of 64 neurons. The features will first be processed by the LSTM layers, with the output features then supplied into the FC layers. There are 256 neurons in FC Layer2, 128 in FC Layer3, and 64 in FC Layer4. An output layer based on softmax is added to the suggested model to complete the recognition process. The proposed model's fine-grained configuration may be tailored to the requirements of a given seizure-related identification challenge.

3.3 Faster-R-CNN in DL-NFDM

Fig. 3 illustrates the structure faster R-CNN. Retrieved multiscale characteristic maps are fed into a focus sector proposal network with a fused recurrent focus module in the proposed enhanced Faster-R-CNN model. The model combines a refined path aggregation characteristic prism networks with a multilayer residual network, specifically ResNet50. The generated multi-scale maps of attributes are then loaded into the focal point with ROI Align to complete the defect diagnosis in manufacturing equipment.

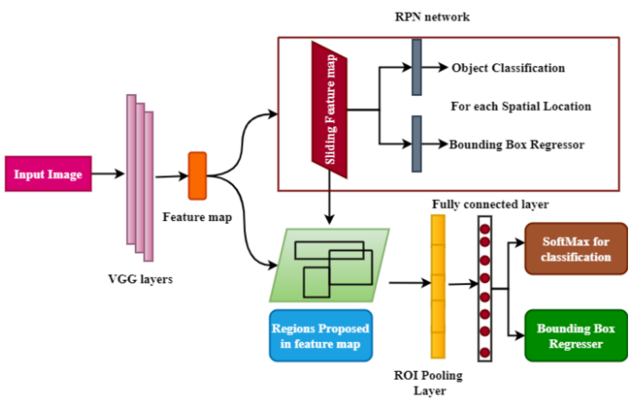

Fig. 3. Structure of Faster-R-CNN

Faster-R-CNN is a new advancement that has successfully integrated normalized flaw identification and detection into a single network for quality control, allowing for near real-time speeds and outstanding performance. This device has had a significant impact, inspiring a wide range of enhancements from other researchers. Faster R-CNN speeds up the detection process by breaking it into two phases. Initially, images are run via a characteristic harvester method and a VGG16 model to forecast boundary proposals in the Region Proposal network (RPN) step. Multilayer convolutional networks, such as VGG-16, ResNet50, and ResNet101, are widely employed for the faster R-CNN feature extraction layer. While deepening the network, the VGG16 network uses shorter linkages to solve fading and exploding gradients. This paper uses VGG 16 as the extracted image from LSTM and CNN for smoothening for upgraded quality outcomes in the production process, as shown in equation 3.

$$u_0 = \sum a u_i . x\left(u_0 + u_x\right) \tag{3}$$

Where u_0 the location within the convolution corresponds to the image location, and a is the attribute. The convolution kernel affects a region of the image denoted by u_i. Pixel positions in the compression u_0 other than are denoted by u_y.

The output of the feature map to the sliding window is represented by Equation 4.

$$u_0 = \sum au_i.x\left(u_0 + u_y + \varphi u_y\right) \quad \textbf{(4)}$$

Offsets allow for angular variation in the grid points of the initial convolution core. Because of this, the flexible convolution component φu_y can represent features more accurately. The flexible convolution improves the original convolution's characteristic extraction quality by allowing for a more accurate approximation of the target object's actual shape through a positional change in the sample point distribution. In this paper, Convolutional Networks to sliding windowing feature to improve its detection accuracy for irregularly shaped machine panels. Adding deformation characteristics to the standard layering of convolution is the goal of accurate convolution.

Fig. 4 shows the structure of the RPN network. The preconditioned network model is fed into the RPN as input during training, and its image feature maps that the pre-trained network model retrieved are output. RPN is a time-efficient alternative to more standard methods for coming up with ideas for flaw regions. The main benefit of Faster R-CNN is that it utilizes the RPN network to produce candidate images using the anchoring method before proceeding with metadata extraction. To boost detection accuracy and efficiency, a network is used to combine the processes of contender fault image selection, boundaries regression analysis, and categorization. Faster R-CNN is used to recognize regions proposed by the RPN, a fully convolutional network (FCN) that can be trained from beginning to end. Scale- and aspect-ratio-independent frame predictions are possible using RPN.

Learning strategies are provided to simultaneously maximize the box for offers while determining the desired score while retaining an established bid, which is made possible by further combining RPN and Faster-R-CNN into a network by sharing convolution features. The feature map is subjected to sliding using a sliding window. Sliding to a new location causes the generation of a one-dimensional vector, which outputs the desired probability via two fully connected layers. The goal probability comprises both the foreground likelihood of an event occurring and the background probability of the event not happening, as well as the bounding box regression parameters. Each anchor has its scale and aspect ratio, clustered in the middle of the sliding window. Fig. 4 illustrates the article's usage of the five scales and three aspect ratios. The RPN network proposes regions on the feature maps based on sliding windows of varying widths and dimensions, which are subsequently used as input by the network's classification and regression layers.

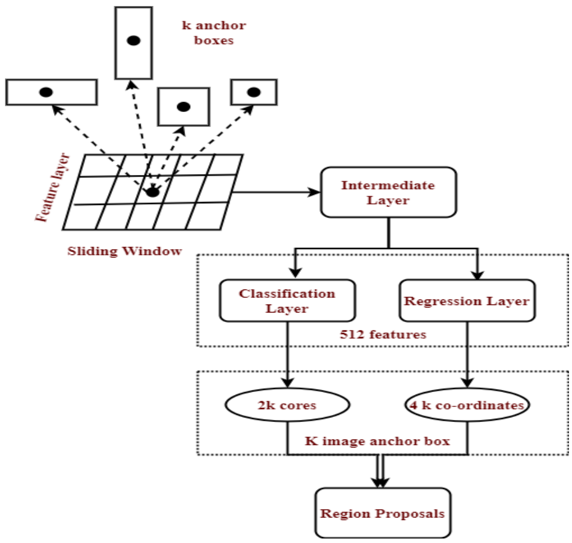

Fig. 4. Region proposal network

The suggested normalized fault region is checked for the presence of a fault detection target using a normalization function (Softmax) in the categorization layer. The regression analysis layer is responsible for calculating the offset of the proposed region bounding box regression, which is then applied to modify the anchoring points to obtain a precise emphasize prospective region. The RPN network's loss equation 5 is calculated using multi-task loss function, which considers both categorization loss and boundary regression loss, and is specified as.

$$Loss = \frac{1}{n_c}\sum_i CL\left(x_i.x_i^*\right) + \gamma \frac{1}{n_r} x_i^* R\left(y_i.y_i^*\right) \tag{5}$$

Where x_i is the number of fault detection images used in training, x_i^* is the probability that an image is a target, y_i is whether or not the image has a target (1 if the image has a target, 0 otherwise), y_i^* is the four-parameterized collaborate vector data of the anticipated extending box, R is the four-parameterized communicate vector of the designated bounding box, and γ is the batch size. The standard deviation of the balancing parameter is 1, and n_c represents the number of fault detection images; *Loss* is the outputs of the categorization and prediction layers are and, respectively. This can help the CNN in RPN better model shapes and handle variations in geometry. These suggestions are then utilized to extract features from the best feature maps, which are then put into the Fast R-CNN for use later in categorization and boundary-box extrapolation.

For areas the feature map suggests, employ Attention-RPN to sharpen the simulation's attention on fault characteristics while simultaneously suppressing features from complicated backgrounds. The improved network is better at pinpointing the precise position of flaws in the presence of challenging picture noises. Input into the Attention-RPN is seen once the feature extraction is complete in Fig. 3. The initial feature is obtained by performing a convolution of three by three on the supplied map of attributes. Targeting regression and categorization using Attention-RPN is then improved by convolving the feature map along with a convolutional blocks attentiveness module to acquire finer variables F″.

The quantization error is generated twice in the Faster R-CNN model since ROI Pooling was roughly used to combine varying-sized potential regions into a single feature map. Here are the measures to take:

- Using the input image as a reference, the proposed region is remapped to its original location in the attribute map. Composite integer values are used as the default rounding factor for the spots.
- A grid of 7x7 cells is generated from the obtained region. A quantization and rounding process is applied to the positions of the floating-point computation cells.

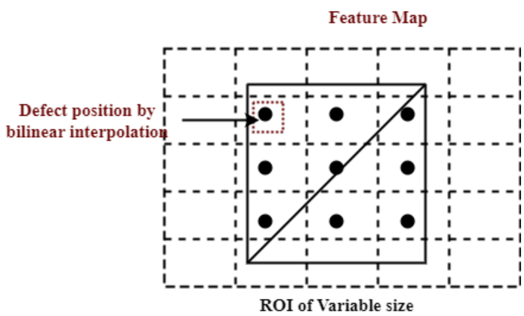

Fig. 5. ROI Align in the feature map

As an alternative to the inexact ROI Pooling, proposed using ROI Align to fix the problem of misaligned regions. Using a regional feature aggregation strategy, ROI Align differentiates itself from ROI Pooling beyond just measuring and pooling. Fig. 5 for visual evidence.

- The first step is to iterate over all possible candidate locations while avoiding quantization of the floating point coordinates used to map them—each cell in the potential region is of unknown size.
- Locate four different points within each cell to use as a sample. A bilinear interpolation method is used to calculate the floating numbers dimensions of observed values to determine the value of locations. The ROI output can then be obtained in a predetermined metric.

While Faster R-CNN is far quicker than Fast RCNN, its speed is still constrained by the first stage's CNN feature extraction and the second stage's expensive per-region calculation. The spatial information was necessary for precise object detection is encoded in a collection of position-sensitive score maps substituted for the fully linked layers in Faster-R-CNN. This research uses hybrid of the LSTM and CNN feature extraction networks and the Faster-R-CNN network to pinpoint the exact location of quality issues in industrial production with unprecedented precision and reliability. The training benefits of VGG 16 lie in the deep networks can create, whereas those of Faster-R-CNN lies in their ability to train the entire network.

4 Results and Discussion

The proposed model utilizes casting product image data for quality inspection dataset to identify the normalized fault detection in industrial manufacturing for quality control. https://www.kaggle.com/datasets/ravirajsinh45/real-life-industrial-dataset-of-casting-product [24]. This data set consists of casting products made in factories. A projectile can be launched by pouring a liquid substance into a mould including a hollow cavity of the proper size and allowing it to harden there before releasing it. Data collection to identify casting flaws. A flaw in the metal casting process is known as a casting fault. Defects in casting can come in various forms, including blow holes, pinholes, burrs, shrinkage flaws, mould material flaws, molten metal flaws, metallurgical flaws, and more. The casting profession is not a place for flaws. Every sector relies on its quality inspection division to weed out faulty goods. Fault detection, detection time, Accuracy and production errors are used to verify the effectiveness of the suggested model.

4.1 Fault Detection Rate (FDR)

The percentage of false detection for the proposed DL-NFDM with Faster R-CNN is displayed in Fig. 6. The model that was taught is next evaluated on a test set of images in industrial processing with scratch defects to confirm its efficacy in identifying weaknesses on actual objects and its capacity to extrapolate the flaw identification to varied material environments. More specifically, DL has allowed for the composition of low-level features from external and nonlinear modules to form more abstract abstractions in terms of categories or attributes and the acquisition of complicated functions with distributed representations of data features. Three other models, the VAE-Faster-RCNN, RODIS, and NCVAE-AFL, are compared to the one proposed here. An initial Faster R-CNN for multiscale defect identification, as shown in the Fig. 6, the false detection rate is 89.07%. Additionally, the detection value of VGG16 with a flexible convolution network paired with ROI Align is enhanced by 3.47%. This is on top of the 1.68% improvement achieved by Faster R-CNN + LSTM and CNN with deformable convolution. The attention paid to the FC layer modules resulted in a FDR that was 95.61% lower than the initial model. FDR is calculated based on Equation 6.

$$FDR = \frac{FI}{TI + FI + TIM + FIM} \tag{6}$$

Where FI is the false image detection, TI is the true image detection; TIM denotes the true image membrane detection and FIM denotes false image membrane detection.

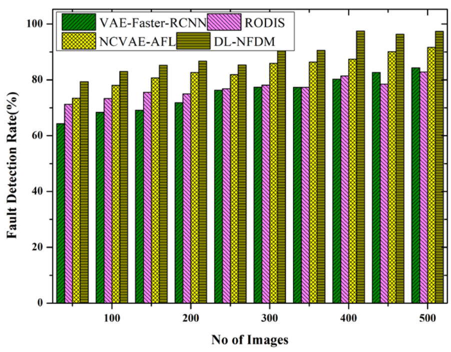

Fig. 6. Fault detection rate

4.2 Detection Time(s) of Feature Extraction Networks

Various image input sizes are used to evaluate the performance of the two primary extractions of features networks, CNN-Res Net 50, LSTM, CNN+LSTM, VGG 16, Mobile Net V2, and Manually labelling and MobileNet-V2. The leading characteristic extraction network used is ResNet-50, and an input image size of 512 by 512 pixels yields the best detection speed. Setting the dimensions and ratios of the foundation frames will allow for comparing the experimental information to verify the impact of anchor frame size and proportion on fault experiment outcomes. Experiments using ablation for image frames of various sizes and aspect ratios are compared in Table 1. Evidence suggests that the anchor frames of 5 sizes and 3 ratios are superior, as their total number of detected defects is higher than that of the other sizes and proportions. Anchor frames have a broader receptive field, leading to more precise target detection. A single fault is present, the text technique achieves the same level of Accuracy as hand marking (92%). In cases where relaxed and linked flaws are present, it reaches the same level of Accuracy as manual marking. During network instruction, a training record of the diminished value and precision will be saved once the training set has been prepared, and the log data can then be displayed and produced.

Table 1. Detection time(s) of feature extraction networks

Network	Input image	Detected faults	Detection time(s)
CNN-Res Net 50	1024x1024	256	0.242
LSTM	512x512	234	0.278
CNN+LSTM	1024x1024	267	0.231
VGG 16	1024x1024	213	0.265
Mobile Net V2	1024x1024	310	34
Manual labelling	512x512	342	40

4.3 Average Precision of Faster R-CNN

Fig. 7 compares the proposed model to specific industry standard models, including the VAE-Faster-RCNN, RODIS, and NCVAE-AFL, regarding average precision. In this case, the proposed model uses a faster CNN, which aids in providing precise defect diagnosis in the machinery used in industrial production. Furthermore, the proposed model's feature extraction uses ResNet 50 as its backbone network, facilitating the smoothening of more images. Several convolution layers are ablated to test how susceptible the model is to changes in that parameter. The experimental findings suggest that in the presence of ambiguous semantic information, introducing an LSTM + convolution layer can result in a rise in false positive samples, improving Accuracy and recall. The computation, training, and average detection period for just one picture will all increase as the number of flexible convolutional layers is raised. The research substitutes only the 3 x 3 transformations in stage 2 of LSTM with the deformed inversion to ensure that the movement of the flexibility inverted core may be acquired through superior characteristics and to simplify the model calculation. Average precision is calculated using equation 7 of the proposed model and is generally relatively high based on the confusion matrix.

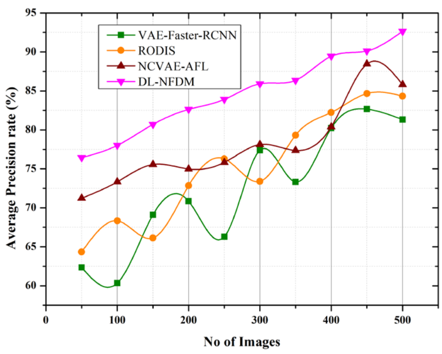

Fig. 7. Average precision of Faster R-CNN

$$Avg\ Precision = \frac{tp}{tp + fn} \tag{7}$$

4.4 Production Error

The Faster-R-CNN architecture is used to build network models; the divided dataset is used as training samples; and VGG16, CNN, and LSTM are used to train the pre-trained models of the Faster-R-CNN. Training the contact network fault identification and fault detection network model with the quicker R-CNN model reduces the production error seen in Fig. 8 for the total number of epochs. The production model's error rate is the percentage of times its predictions are off compared to the gold standard. In classification models, the term error rate is frequently used. The convolutional features are shared using the end-to-end training method. The Faster-R-CNN is used to optimize the model, with the maximum amount of iterations set to 1,000, the starting rate of learning set to 0.0001, the momentum coefficient set to 0.10, and cease training the function of loss converges. The number of iterations reached 40,000, with VGG16, ResNet101, and DenseNet121 as the backbone networks.

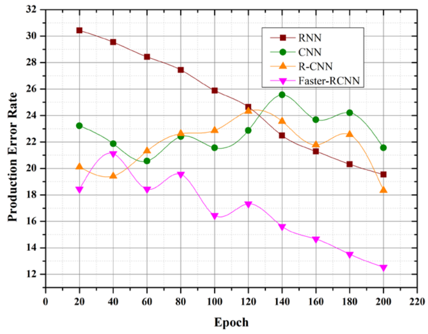

Fig. 8. Production errors

4.5 Accuracy of the Proposed Model

Fig. 9 displays the proposed model's accuracy rate. The number of times a model mistakenly recognizes the fault diagnosis of the background as cracks are used to determine the reliability based on erroneous detections. The percentage of visuals in which the subject incorrectly identifies a mark on the exterior of the equipment panel. In particular, when the surface of the background traditional representation of the image that needs to be identified is identical to that of the set used for training, the favourable rate of identification for genuine marks reaches 95.5%, completely validating the utility of the CNN strategy characterized in this study as an information enhancement method. The suggested model has an accurate detection rate of 98.51 per cent on the test set. The identification rate drops when the surface of the background diverges from the original training set, implying that the model's detection skill drops as well. Accuracy is calculated based on confusion matrix true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) are the four possible outcomes shown in Equation 8

$$Acc = \frac{Tp + TN}{Tp + Tn + Fp + Fn} \tag{8}$$

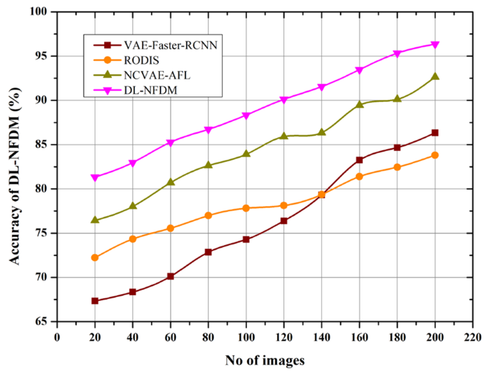

Fig. 9. Accuracy of the proposed model

5 Conclusion

This article presents a Normalized Failure Detection Model (DL-NFDM) based on Deep Learning to solve the problems in industrial machinery production. This model uses deep learning techniques like Faster-R-CNN to detect control faults that occur at regular intervals throughout the show. Harmonized intervals are bundled together to prevent further control errors. Proactive DL for fault detection during production for quality control is receiving much attention from specialists in the equipment business due to DL's growing popularity in several detection applications. First, defect sample pairs were fed into a single feature extraction network. Then an LSTM network and a CNN were used to map the vibratory signal information to the low-dimensional characteristic space. The collected samples' attributes were then compared using the connection assessment network, with the resulting data being put into an analysis network for defect identification. The severity of faults, the decline in output, and the delays in receiving various outcomes all play a role in this ongoing learning process. In the case of consecutive defect detection, independent intervals are used for model training. The proposed model's performance in terms of identifying defects, discovery time, precision exactness, and manufacturing losses to confirm its efficacy. The digital contrast between retrieved geometrical information with a 3D computer-aided design (CAD) model for fault evaluation can aid future study topics for fault diagnosis in industries.

Acknowledgments

This research was funded by Liuzhou Railway Vocational and Technical College Model Innovation Workshop (Liu Tie Xiao Fa No. [2022] 60); Liuzhou Railway-Liuzhou OVM Intelligent Equipment Collaborative Innovation Centre (Liu Tie Xiao Fa No. [2021] 31) and 2024 Guangxi Universities Young and Middle-aged Teachers' Scientific Research Basic Ability Enhancement Project (Fractional Order PID-based Quadcopter Control System Design and Implementation (No. 2024KY1461)).

References

[1] L. Rojas, Á. Peña, J. Garcia, AI-Driven Predictive Maintenance in Mining: A Systematic Literature Review on Fault Detection, Digital Twins, and Intelligent Asset Management, Applied Sciences 15(6)(2025) 3337. https://doi.org/10.3390/app15063337

[2] J.S. Kim, K.N. Choi, S.W. Kang, Infrared thermal image-based sustainable fault detection for electrical facilities, Sustainability 13(2)(2021) 557. https://doi.org/10.3390/su13020557

[3] Z. Wu, C. Jiao, J. Sun, L. Chen, Tire defect detection based on faster R-CNN, in: Proc. 2020 International Conference on Robotics and Rehabilitation Intelligence, 2020. https://doi.org/10.1007/978-981-33-4932-2_14

[4] H. Zhu, T. Peng, Y. Dai, C. Zhou, B. Sun, Fault detection of electrolyzer plate based on improved Mask R-CNN and infrared images, Measurement Science and Technology 33(8)(2022) 085405. https://doi.org/10.1088/1361-6501/ac5b29
[5] M. Li, Z. Zhang, L. Lei, X. Wang, X. Guo, Agricultural greenhouses detection in high-resolution satellite images based on convolutional neural networks: Comparison of faster R-CNN, YOLO v3 and SSD, Sensors 20(17)(2020) 4938. https://doi.org/10.3390/s20174938
[6] K. Zhao, Y. Wang, Q. Zhu, Y. Zuo, Intelligent Detection of Parcels Based on Improved Faster R-CNN, Applied Sciences 12(14)(2022) 7158. https://doi.org/10.3390/app12147158
[7] W. Zhang, C. Li, G. Peng, Y. Chen, Z. Zhang, A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load, Mechanical systems and signal processing 100(2018) 439-453. https://doi.org/10.1016/j.ymssp.2017.06.022
[8] M. Lucke, M. Chioua, C. Grimholt, M. Hollender, N.F. Thornhill, Integration of alarm design in fault detection and diagnosis through alarm-range normalization, Control Engineering Practice 98(2020) 104388. https://doi.org/10.1016/j.conengprac.2020.104388
[9] W. Zhang, G. Peng, C. Li, Y. Chen, Z. Zhang, A new deep learning model for fault diagnosis with good anti-noise and domain adaptation ability on raw vibration signals, Sensors 17(2)(2017) 425. https://doi.org/10.3390/s17020425
[10] X. Zhao, J. Yao, W. Deng, M. Jia, Z. Liu, Normalized conditional variational auto-encoder with adaptive focal loss for imbalanced fault diagnosis of bearing-rotor system, Mechanical Systems and Signal Processing 170(2022) 108826. https://doi.org/10.1016/j.ymssp.2022.108826
[11] A. Glaeser, V. Selvaraj, S. Lee, Y. Hwang, K. Lee, N. Lee, S. Lee, S. Min, Applications of deep learning for fault detection in industrial cold forging, International journal of production research 59(16)(2021) 4826-4835. https://doi.org/10.1080/00207543.2021.1891318
[12] J. Long, Y. Qin, Z. Yang, Y. Huang, C. Li, Discriminative feature learning using a multiscale convolutional capsule network from attitude data for fault diagnosis of industrial robots, Mechanical Systems and Signal Processing 182(2023) 109569. https://doi.org/10.1016/j.ymssp.2022.109569
[13] W. Li, X. Zhong, H. Shao, B. Cai, X. Yang, Multi-mode data augmentation and fault diagnosis of rotating machinery using modified ACGAN designed with new framework, Advanced Engineering Informatics 52(2022) 101552. https://doi.org/10.1016/j.aei.2022.101552
[14] S. Tang, Y. Zhu, S. Yuan, An adaptive deep learning model towards fault diagnosis of hydraulic piston pump using pressure signal, Engineering Failure Analysis 138(2022) 106300. https://doi.org/10.1016/j.engfailanal.2022.106300
[15] S. Dhote, C. Vichoray, R. Pais, S. Baskar, P.M. Shakeel, Hybrid geometric sampling and AdaBoost based deep learning approach for data imbalance in E-commerce, Electronic Commerce Research 20(2020) 259-274. https://doi.org/10.1007/s10660-019-09383-2
[16] P. Mouzenidis, A. Louros, D. Konstantinidis, K. Dimitropoulos, P. Daras, T. Mastos, Multi-modal Variational Faster R-CNN for Improved Visual Object Detection in Manufacturing, in: Proc. 2021 IEEE/CVF International Conference on Computer Vision, 2021. https://doi.org/10.1109/ICCVW54120.2021.00292
[17] D. Im, J. Jeong, R-CNN-Based Large-Scale Object-Defect Inspection System for Laser Cutting in the Automotive Industry, Processes 9(11)(2021) 2043. https://doi.org/10.3390/pr9112043
[18] P. Lyu, K. Zhang, W. Yu, B. Wang, C. Liu, A novel RSG-based intelligent bearing fault diagnosis method for motors in high-noise industrial environment, Advanced Engineering Informatics 52(2022) 101564. https://doi.org/10.1016/j.aei.2022.101564
[19] X. Zhao, J. Yao, W. Deng, M. Jia, Z. Liu, Normalized conditional variational auto-encoder with adaptive focal loss for imbalanced fault diagnosis of bearing-rotor system, Mechanical Systems and Signal Processing 170(2022) 108826. https://doi.org/10.1016/j.ymssp.2022.108826
[20] F. Del Buono, F. Calabrese, A. Baraldi, M. Paganelli, F. Guerra, Novelty detection with autoencoders for system health monitoring in industrial environments, Applied Sciences 12(10)(2022) 4931. https://doi.org/10.3390/app12104931
[21] W. Jiang, A machine vision anomaly detection system to industry 4.0 based on variational fuzzy autoencoder, Computational Intelligence and Neuroscience 2022(2022) 1945507. https://doi.org/10.1155/2022/1945507
[22] M.H. Abidi, M.K. Mohammed, H. Alkhalefah, Predictive maintenance planning for industry 4.0 using machine learning for sustainable manufacturing, Sustainability 14(6)(2022) 3387. https://doi.org/10.3390/su14063387
[23] P. Suawa, T. Meisel, M. Jongmanns, M. Huebner, M. Reichenbach, Modeling and fault detection of brushless direct current motor by deep learning sensor data fusion, Sensors 22(9)(2022) 3516. https://doi.org/10.3390/s22093516
[24] Kaggle, casting product image data for quality inspection. <https://www.kaggle.com/datasets/ravirajsinh45/real-life-industrial-dataset-of-casting-product> (accessed 18.11.2023).