International Journal of Advanced Engineering Research and Science (IJAERS)
Peer-Reviewed Journal
ISSN: 2349-6495(P) | 2456-1908(O)
Vol-12, Issue-12; Dec, 2025
Journal Home Page Available: https://ijaers.com/
Article DOI: https://dx.doi.org/10.22161/ijaers.1212.4

# A classification method for bearing surface defects based on acoustic emission technology and the YOLO-V11 algorithm

Xin Yu Guo, Liu Yi Yu, Yi Chen, Yan Zuo Chang*

Guangdong Provincial Key Laboratory of Petrochemical Equipment Fault Diagnosis, School of Energy and Power Engineering, Guangdong University of Petrochemical Technology, Maoming, Guangdong 525000, China
*Corresponding author: oshhhs1@gmail.com
[1]Email: 172876198@qq.com

Received: 06 Nov 2025,

Received in revised form: 03 Dec 2025,

Accepted: 08 Dec 2025,

Available online: 15 Dec 2025



***Keywords***— ***bearing fault diagnosis, YOLOV11, defect detection, image classification***

***Abstract***— *In contemporary petrochemical manufacturing, the identification of defects in a substantial quantity of bearings is frequently a necessity. Conventional methods of detection, namely manual inspection and acoustic emission detection, are often plagued by deficiencies in terms of efficiency and precision. The proposed methodology integrates acoustic emission technology with the open-source deep learning algorithm YOLO-V11, facilitating rapid detection of bearing faults. Initially, five types of bearings of the same model but different varieties are selected and installed on the same shaft segment, which is then connected to an acoustic emission detection device. Acoustic emission signals are obtained for each type of bearing according to the different types of bearing fault. These signals are then visualised in two dimensions to generate vibration images, which serve as input for the model and are used to train the YOLO-V11 model. The experimental findings demonstrate that the prediction accuracy and recall rate for various defects generally exceed 80%, thus substantiating the efficacy of the proposed method for industrial production in the diagnosis and classification of bearing defects.*

## I. INTRODUCTION

The advent of industrial production has led to a significant increase in the utilisation of bearings as vital components of mechanical equipment. These bearings have found application in a wide range of sectors, including machinery manufacturing, automobiles, wind power and new energy. As the primary rotating components of machinery, they must demonstrate resistance to friction and the high loads imposed by the equipment. In the context of actual production processes, it is inevitable that bearings will develop various defects, including scratches, wear, and cracks, due to the variable operating conditions to which they are subjected. These defects have the capacity to affect the performance and quality of the bearings to varying degrees, and consequently to have an indirect impact on the safety and stability of the entire equipment. In the absence of timely detection, the aforementioned issues have the potential to result in personnel casualties and economic losses during production. The traditional manual inspection process is largely dependent on the experience and judgement of the workers, which can be both inefficient and difficult to accurately identify the correct type of defect. Consequently, the necessity to expedite the accurate and efficient identification of bearing defects has become a matter of pressing concern.

Pan Tiancheng et al.[1]collected vibration signals from bearings in different defect states and then performed time-domain waveform analysis, frequency-domain analysis, and time-frequency domain analysis of the vibration signals to analyse the type of bearing faults. He Yuqi et al.[2]proposed a bearing fault diagnosis method based on multi-scale graph domain features. This method combines graph signal processing methods to effectively analyse signal propagation paths in bearing systems. It is not subject to traditional methods for data source limitations and improves the detection capability for hidden bearing faults. Wang Xinghe et al.[3]utilised spectral kurtogram to ascertain the centre frequency and bandwidth of the frequency bands containing fault information. They then employed envelope analysis to determine the bearing fault frequency, facilitating the analysis and diagnosis of bearing faults. In their seminal work, Zhang Xuhui et al.[4]pioneered a novel integration of fast spectral kurtogram with envelope order analysis, a methodology that has since become the gold standard for precise localization of bearing fault positions under strong interference. Nevertheless, these methodologies continue to exhibit deficiencies in terms of efficiency and precision, thereby impeding the timely identification of issues in industrial production. The advent of deep learning technology has precipitated a surge of interest in the development of defect detection methodologies for bearings, with these methodologies being predicated on the utilisation of deep learning models. As posited by several scholar[5-7], the capacity of deep learning technology to extract salient features of bearing defects is predicated on its ability to learn and analyse a range of bearing data, including, but not limited to, vibration signals. The employment of bespoke algorithm models facilitates the expeditious identification and prediction of defective bearing types. The utilisation of convolutional neural networks (CNN) and recurrent neural networks (RNN) serves to enhance the capacity to extract features from intricate data. Furthermore, deep learning models demonstrate a high degree of iterativeness, enabling targeted enhancements for a range of research objectives.

Therefore, this paper primarily discusses the implementation of batch identification and classification of bearing defect types using acoustic emission and the YOLO-V11 model. Bearings of the same model but different defect types are installed on the same shaft segment. Acoustic emission technology is used to obtain one-dimensional signals of various bearings, which are then visualized in two dimensions to obtain waveform diagrams. Through training with the YOLO-V11 model, high-precision identification of bearing defect classification is achieved.

## II. YOLO-V11 MODEL

The YOLOv11 object detection algorithm, the latest release from Ultralytics, has been developed to achieve higher accuracy through enhanced feature extraction, optimised efficiency, and a reduction in parameters. The system is capable of supporting a variety of computer vision tasks, including object detection, segmentation, and classification. Furthermore, it is well-suited for a broad spectrum of deployment scenarios, ranging from edge devices to cloud platforms. In comparison with the preceding YOLOv8 model, YOLOv11 has undergone enhancement and optimisation in a number of areas. YOLOv11 is notable for its retention of the fast and accurate characteristics of the YOLO series. Furthermore, it innovatively introduces the C3K2 module and C2PSA module, with the objective of enhancing performance and flexibility.

The YOLOv11 model incorporates the C3k2 module, which replaces the C2f module present in the YOLOv8 architecture. The C3k2 module determines whether to use the C3k or Bottleneck structure by setting the parameter c3k. The integration of the C3k2 module serves to augment the model's capacity for feature extraction, while ensuring the preservation of efficiency.

A C2PSA module was incorporated subsequent to the SPPF layer, constituting a multi-head attention mechanism that serves to enhance the model's feature extraction capability. The C2PSA module has been developed to achieve a more powerful attention mechanism by embedding the PSA mechanism into the C2 structure.

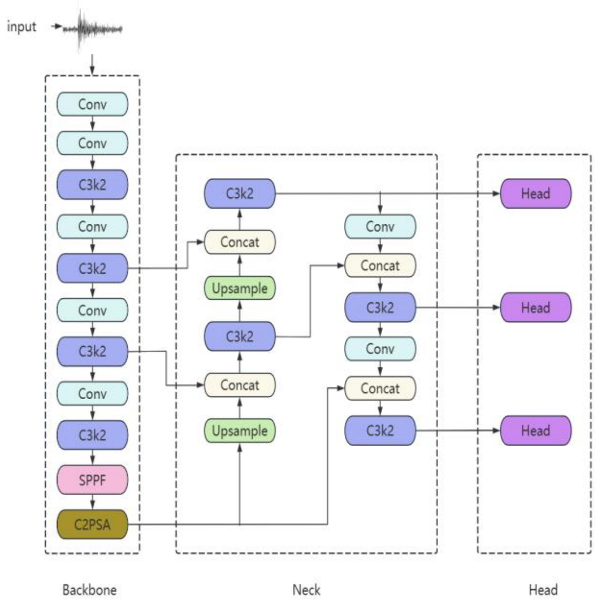


*Fig 1: YOLO-V11 structure diagram*

The classification detection head was modified by the introduction of two deep convolutions, which replaced the conventional convolutions that had previously been employed. This approach has been shown to significantly reduce both the parameter count and the computational load, while maintaining model performance.

YOLOv11 has modified the network depth and width of models for each version (e.g. YOLOv11n, YOLOv11s, YOLOv11m, YOLOv11l, YOLOv11x) in order to achieve enhanced balance and performance.

## III. EXPERIMENT AND ANALYSIS

### 3.1 Data Construction and Experimental Environment

The present paper is an attempt to collate a variety of signal parameters during the rotational operation of defective bearings in a laboratory setting. The generation of vibration signal graphs was achieved through the utilisation of visualisation techniques, culminating in the creation of a total of 6,968 vibration images in the PNG image format. The dataset was then divided into five categories, according to the type of defect. The following defects are to be noted:

-Bearing ball wear

-Bearing lacks balls

-Bearing outer cracking

-Bearing inner cracking

-No defect

The categorisation of the data was conducted in accordance with the following procedure. First, each category was proportionally divided into training sets, validation sets, and test sets. The division is illustrated in Table 1.

The operating system utilised in this experiment was Windows 11, with a central processing unit (CPU) of AMD Ryzen 7 7435H and a graphics processing unit (GPU) of NVIDIA GeForce RTX 4060 Laptop GPU. The Python version was 3.10, the PyTorch version was 2.2.1, and the Miniconda version was 25.7.

*Table 1 :Data Set Feature Distribution(Where, B_b_l indicates bearing ball missing, B_b_w indicates bearing ball wear, B_i_c indicates bearing inner crack, B_o_c indicates bearing outer crack)*

| Feature Category | B_b_l | B_b_w | B_i_c | B_o_c | Nor |
|---|---|---|---|---|---|
| Training set | 1500 | 1500 | 1500 | 600 | 80 |
| Validation set | 225 | 225 | 225 | 100 | 10 |
| Test set | 275 | 275 | 275 | 167 | 11 |

### 3.2 Results Analysis

The experiment selected YOLO-v11n as the training model, with the number of training epochs set to 250. The evaluation of the classification performance is undertaken by the present paper through the utilisation of the visualisation curves of the YOLO-V11 model, incorporating the metrics of precision (P), recall (R), and the F1 score, in accordance with the training outcomes. The visualisation curves are displayed in Fig 2. As illustrated in Figure 2, the trends of four indicators are presented: training loss (train/loss), validation loss (val/loss), Top1 accuracy (metrics/accuracy_top1), and Top5 accuracy (metrics/accuracy_top5). It is evident that the train/loss curve demonstrates a downward trend as the number of training epochs increases, gradually decreasing from an initial value of 1.6 to below 0.6. This finding suggests that during the training process, the model's calibration to the training data undergoes continuous enhancement, leading to a concomitant reduction in the training error. In a similar vein, the val/loss curve demonstrates a downward trend, with the model loss value decreasing from approximately 1.75 to below 0.5 as the training epochs rise. This finding indicates that the model's capacity to fit unseen data also improves, and the generalization error undergoes a gradual decrease. With regard to the accuracy curves, it is evident that as the number of training epochs increases, the Top1 accuracy rises steadily to above 0.8, signifying an enhancement in the model's capacity to predict single categories. Concurrently, the Top5 curve persists in its proximity to 1.00, thereby signifying that the model sustains an elevated degree of confidence in its predictions across a more extensive range.

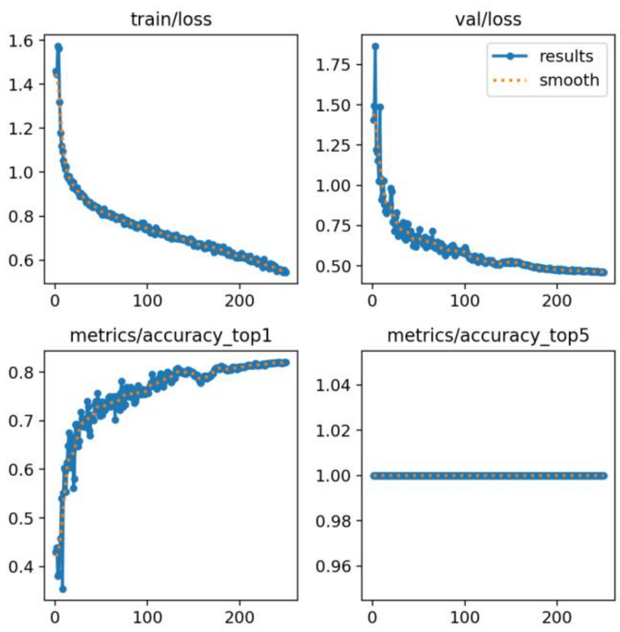


*Fig 2: Classification results of the YOLO-V11n model*

### 3.3 **Evaluation Metrics**

In order to further evaluate the accuracy of the model's classification, the experiment also adopted Precision (P), Recall (R), and F1-score as evaluation metrics.

Precision is defined as the ratio of correctly identified targets to all targets predicted by the model, thereby reflecting the model's accuracy in target recognition:

$$P = \frac{TP}{TP+FP} \tag{1}$$

The term 'recall' is defined as the ratio of correctly identified targets to all actual targets. This metric is indicative of the model's sensitivity to target recognition.

$$R = \frac{TP}{TP+FN} \tag{2}$$

The F1 score is the harmonic mean of precision and recall, which can more comprehensively evaluate performance:

$$F = \frac{2\ (P*R)}{(P+R)} \tag{3}$$

Where: True positives (TP) denote the number of actual targets correctly detected by the model, false positives (FP) indicate the number of instances detected as positive but are actually negative, and false negatives (FN) denote the number of actual positive instances not detected by the model. The calculation of these metrics is facilitated by the confusion matrix, a tabular representation employed for the evaluation of model classification performance. The model displays the correct and incorrect results of model classification intuitively, based on predicted categories and true categories. The confusion matrix is displayed in Fig 3. It is possible to calculate the precision, recall and F1 score for each category of defective bearings based on the dataset and confusion matrix that have been obtained. The results of this calculation are shown in Table 2.

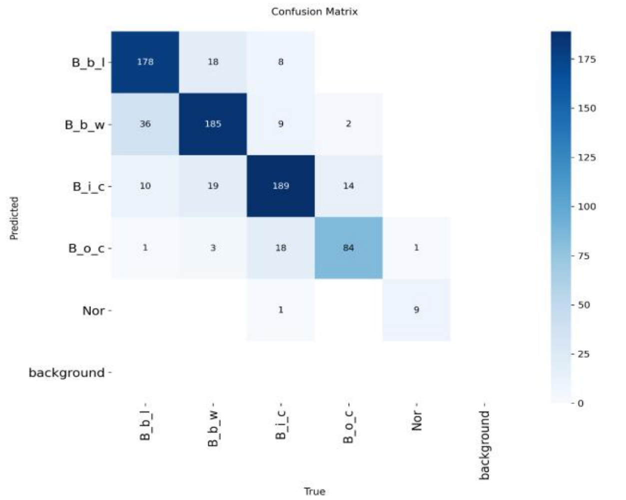


*Fig 3: Confusion matrix derived from the validation set samples*

*Table 2: Evaluation Indicators for Various Sample Types(Where, B_b_l indicates bearing ball missing, B_b_w indicates bearing ball wear, B_i_c indicates bearing inner crack, B_o_c indicates bearing outer crack)*

| Sample Category | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|
| B_b_l | 87.25 | 79.11 | 82.98 |
| B_b_w | 79.74 | 82.22 | 80.96 |
| B_i_c | 81.47 | 84.00 | 82.72 |
| B_o_c | 78.50 | 84.00 | 81.16 |
| Nor | 90.00 | 90.00 | 90.00 |

As demonstrated in Table 2, the precision and recall rates of samples from each defect category typically remain around 80%, with F1 scores consistently exceeding 80%. It is evident that high precision signifies the model's capacity to accurately categorise targets, exhibiting a reduced propensity for false positives. Concurrently, high recall implies the model's reduced likelihood of overlooking target categories, thereby diminishing false negatives. The experiments previously referenced demonstrate that this method can serve as a basis for detecting bearing defect categories.

## IV. SUMMARY

The paper is concerned with the problem of detecting defective categories of bearings. It proposes a method that combines YOLO-V11 models, which are based on deep learning, for the classification and recognition of defective bearings by analysing the vibration signal graphs of rotating bearings. Initially, vibration signals of various defective bearings were collected, and vibration images of different bearings were obtained through visualisation. These images were then employed to train the YOLO-V11n model for classification. The efficacy of the model training process was evidenced by the attainment of commendable performance metrics, including accuracy, recall rate, and F1 score. This outcome signified the successful execution of effective classification across a range of defect categories and the fulfilment of criteria for multi-fault identification. The model's high accuracy and recall rate ensured that it could accurately identify both normal and defective bearings, thereby reducing unnecessary downtime and cost losses due to misjudgments.

Despite the efficacy of the proposed methodology for the identification of bearing defect categories, it is important to acknowledge the limitations of the study. For instance, the dataset utilised in the experiments is imbalanced, with certain subsets containing a smaller number of data points. This has led to suboptimal training accuracy and reduced the reference value of the results. It is evident that the accuracy and recall rate for specific categories require enhancement. Furthermore, the training algorithm for the model can be refined to enhance detection rates. It is anticipated that, in the context of ongoing optimisation and enhancement of the model, this method will be adopted more extensively across a range of industrial scenarios.


## V. ACKNOWLEDGEMENTS

The successful completion of this research would not have been possible without the tremendous support of many teachers and peers. My supervisor, Yan Zuo Zhang, possesses profound knowledge and rigorous academic standards, providing meticulous guidance from topic selection to manuscript completion, guiding me in the right direction. My classmates and colleagues are intellectually vibrant, sparking numerous ideas through exchanges and discussions. My family provides silent support, offering me a warm haven. I am grateful to Guangdong University of Petrochemical Technology for providing excellent resources and platforms, enabling me to deepen my academic pursuits. I extend my thanks to all who have helped me; you are the beacons lighting my path forward.


## REFERENCES


[1] Pan Tiancheng, Chen Long, Chen Zhiqiang, Pu Chunlei, Ding Jing, & Li Lei. (2025). Rolling Mill Bearing Fault Diagnosis. Metallurgical Equipment, 4, 90 ~ 92, 106.

[2] He Yuqi, Zhang Bo, Su Chang, Zhang Wanhong, Zhang Hao, & Yin Aijun. (2025). Bearing Fault Diagnosis Method Based on Multi-scale Graph Domain Features. Noise and Vibration Control. Advance online publication. https://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CAPJ&dbname=CAPJLAST&filename=ZSZK2025091500D

[3] Wang Xinghe, Wang Hongjun, & Liu Guoqing. (2021). A rolling bearing fault diagnosis method based on spectral kurtosis. Equipment Management and Maintenance, 9, 151-153. https://doi.org/10.16621/j.cnki.issn1001-0599.2021.05.68

[4] Zhang Xuhui, Zhang Chao, Fan Hongwei, Mao Qinghua, & Yang Yiqing. (2021). Rapid Spectral Kurtosis Combined with Order Analysis for Rolling Bearing Fault Diagnosis. Journal of Vibration, 41(6), 1090 ~ 1095, 1235. https://doi.org/10.16450/j.cnki.issn.1004-6801.2021.06.007

[5] Wang Youdao & Zhang Peng. (2025). A review of domestic deep learning research in the field of bearing fault diagnosis. Industrial Innovation Research, 12, 88-90.

[6] Ciaburro, G., & Iannace, G. (2022). Machine-Learning-Based Methods for Acoustic Emission Testing: A Review. Applied Sciences, 12(20), 10476. https://doi.org/10.3390/app122010476

[7] Liu Minghui, Xiong Jianbin, Su Naiquan, Li Chunlin, Cen Jian, & Zhang Yuyu. (2023). A Review of Deep Learning-Based Fault Diagnosis for Petrochemical Equipment Bearings. Machine Tools and Hydraulics, 51(6), 171 ~ 180.

[8] Chen Qiangqiang, Dai Shaowu, Dai Hongde, & Nie Zijian. (2019). Review of Rolling Bearing Fault Diagnosis Methods. Instrument Technology, 9, 1 ~ 4, 42. https://doi.org/10.19432/j.cnki.issn1006-2394.2019.09.001

[9] Chang Wei. (2019). Research on Fault Diagnosis of Gear and Bearing in Coal Mine Machinery. Inner Mongolia Coal Economy, 15, 64. https://doi.org/10.13487/j.cnki.imce.014515

[10] Guo Yuan & Zhou Jun. (2024). Research Progress on Bearing Defect Detection Based on Machine Vision. Mechanical and Electrical Engineering, 41(5), 761 ~ 774.

[11] Li Kangning. (2025). Research on Application of Bearing Surface Defect Detection Based on Deep Learning [Master's Thesis, Shandong Jiaotong University]. https://doi.org/10.27864/d.cnki.gsjtd.2025.000206

[12] Liao Yunhu & Ji Guoyi. (2025). A New Method for Rolling Bearing Fault Diagnosis in Strong Noise Background. Noise and Vibration Control. Advance online publication. https://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CAPJ&dbname=CAPJLAST&filename=ZSZK2025091500F

[13] Wei Zishu, Chen Zhigang, Wang Yanxue, & Hasteer Madetihan. (2025). A lightweight bearing appearance defect detection algorithm based on SBSI-YOLO11. Journal of Guangxi Normal University (Natural Science Edition). Advance online publication. https://doi.org/10.16088/j.issn.1001-6600.2024121901

[14] Yan Zhiyong. (2008). Fault diagnosis of bearings in chemical slurry circulating pumps. Equipment Management and Maintenance, 6, 44-45.

[15] Yao Jingli, Cheng Guang, Wan Fei, & Zhu Deping. (2024). An Improved Lightweight Bearing Fault Detection Algorithm Based on YOLOv8. Computer Engineering and Applications, 60(21), 205 ~ 214.

[16] Zhong Xianyou. (2015). Time-Frequency Analysis Methods and Their Applications in Rotating Machinery Fault Diagnosis [Doctoral Dissertation, Wuhan University of Science and Technology].https://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CDFD&dbname=CDFDLAST2015&filename=1014389513.nh

[17] Shao, Y.-F., Jiang, P., Dong, Y., Li, W., & Zhang, W.-Q. (2024). AE-IRMamba: Low-Complexity Inverted Residual Mamba for Identification of Piezoelectric Ceramic and Optical Fiber Acoustic Emission Sensors Signals. *IEEE Sensors Journal*, *24*(21), 34549 ~ 34560. https://doi.org/10.1109/JSEN.2024.3457913