



ResNet50-Driven Quality Inspection for Recorder Musical Instrument

Rizki Putra Prastio¹, Rodik Wahyu Indrawan², Vanesia Tasib³, Zhilaan Abdurrasyid Rusmawan⁴

^{1,2,3,4}Robotics and Artificial Intelligence Engineering, Faculty of Advanced Technology and Multidiscipline,
Universitas Airlangga, Surabaya, Indonesia

¹r.p.prastio@ftmm.unair.ac.id, ²rodik.w.i@ftmm.unair.ac.id, ³vanesiataсіб23@gmail.com, ⁴zhilaanabdurasyid@gmail.com

Abstract

The manufacturer of a recorder musical instrument requires high-quality product. The aim is to produce precise tones and an aesthetic look at customer satisfaction. A major challenge encountered by manufacturers is traditional visual inspection. Human error is a major factor, notably over extended work periods and the subjective judgment of quality control personnel. This paper reports on the development of a machine vision system for detecting abnormal patterns on the inner surface of a recorder musical instrument. An industrial-grade camera with a resolution of 1280×1024 , paired with industrial lighting, was utilized. Due to its tube-shaped construction of the object, the bright-field imaging technique is applied to illuminate the interior. ResNet50 was selected as a feature extractor due to its balance between accuracy and efficiency. In addition, a Neural Network served as the classifier. A total of 1,118 images were collected as training data and 304 images as testing data. The training and testing data were separate sets that were taken independently, preventing any risk of data leakage. The test results indicated that the model performed exceptionally well in classification, achieving an accuracy of 95.7%, precision of 95.45%, sensitivity of 96.07%, and specificity of 95.36%. Moreover, the area under the curve of the Receiver Operating Characteristic (ROC AUC) score in test data reached 0.9906, reflecting the model's ability to separate features from the two classes. These findings suggest that the proposed method offers an alternative to subjective visual inspection. Future research may examine diverse deep learning architectures to further enhance performance while achieving faster classification.

Keywords: defect detection; musical instrument; neural network; recorder; ResNet50

How to Cite: R. P. Prastio, R. W. Indrawan, V. Tasib, and Z. A. Rusmawan, "ResNet50-Driven Quality Inspection for Recorder Musical Instrument", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 9, no. 5, pp. 975 - 982, Sep. 2025.

Permalink/DOI: <https://doi.org/10.29207/resti.v9i5.7058>

Received: August 7, 2025

Accepted: August 25, 2025

Available Online: September 29, 2025

This is an open-access article under the CC BY 4.0 License

Published by Ikatan Ahli Informatika Indonesia

1. Introduction

In manufacturing industries, defects are unavoidable. The flaws can be found in various areas, including both the surface and the interior of the products [1]-[3]. In industrial settings, such as in the production of musical instruments, quality control is a critical stage. In this step, the quality control personnel perform inspections to identify flawed products to ensure customer satisfaction. The recorder is among the musical instruments subjected to visual inspection during the final production stage. The recorder is a popular, tube-shaped, wind musical instrument widely played by many people. This musical device is commonly used to introduce music to younger generations, although many professionals also play this instrument [4], [5]. Recorder was historically made from reeds or animal bones, featuring multiple holes to generate sound [6]. As a result of technological advancement, a modern recorder is made from plastic.

In the final stage of production processes, a visual inspection is required to evaluate the quality of the finished product. During this process, each manufactured item undergoes binary or multi-class classification, resulting in at least two categories: good and damaged. Good items will enter the sales stream, while damaged items will be directed towards recycling streams. Quality assessment for the recorder becomes more challenging due to its tube shape. In real-world factory conditions, the quality control workers must examine the product's interior through a narrow opening. This procedure aims to identify any abnormal patterns in the product. Additionally, due to the reflective surface, the worker is required to adjust lighting conditions or viewing direction. Thus, the defective patterns are more visible. This traditional visual inspection is prone to human errors, resulting in inconsistent accuracy and slow processing. Therefore, a machine vision system is required to address this problem. Numerous studies show a promising

performance of artificial intelligence-driven machine vision that enhances accuracy, accelerates inspection time, and automates the processes [7]-[18].

Image-based classification begins with extracting features from the input image, resulting in a feature vector. Then, a classification algorithm takes a feature vector as input to determine the class of the item being inspected. In recent years, the adoption of deep learning architectures for visual inspection in industrial settings has seen a dramatic increase. Numerous studies have explored various deep learning models across a wide range of applications, such as medicine [19]-[23], manufacturing process [24]-[28], food and agriculture [29]-[35], and also art [36]. The data presented in these studies highlight the remarkable performance achieved by ResNet50. Nevertheless, the use of ResNet50 for visual inspection in musical instruments is scarcely investigated. An underlying factor is that musical instruments are not essential for ordinary daily practices. Thus, the quantity of musical instruments is limited compared to products intended for daily use. Yet, a visual inspection system remains essential in the musical instrument industry. Earlier studies provide evidence that ResNet50 has the potential to be implemented in the quality control of musical instrument production.

This study contributes to addressing a real challenge faced by a manufacturer of the recorder musical instruments. Moreover, it evaluates the potential of ResNet50 for musical instrument visual inspection, which few researchers have investigated. ResNet50 serves as a feature extractor, and a neural network works as the classifier. The recorder examined in this research consists of three components, namely the head joint, middle joint, and foot joint. Yet, this research's primary focus is on identifying abnormalities within the interior of the foot joint. There are two classes as the output of the classifier: Not Good (NG) for defective surface and OK for good surface. In total, we have collected 1,118 images for the training data, comprising 568 images labeled as NG and 550 images labeled as OK. For testing, 304 images were acquired, consisting of 153 images for NG and 151 images for OK. It should be noted that the test data was independently collected to ensure the distinction between the test set and the training set.

2. Methods

2.1 Research stages

This research was carried out in several steps as depicted in Figure 1. It starts with a literature review that collects relevant recent studies related to a visual inspection system. Data collection and data splitting are then performed to prepare the data for experimentation. The feature extraction is then carried out, followed by the model development, training, and testing processes. The unsatisfactory result suggests that the model needs to be optimized, and then the training phase is repeated.

Once the result shows a satisfactory performance, the study summarizes the findings.

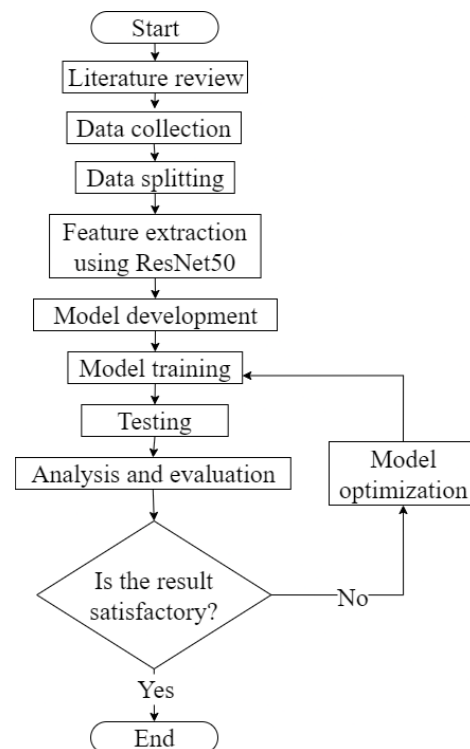


Figure 1. The flowchart of the research.

2.2 Data Collection Setup

A highly configurable industrial camera, resolution of 1280×1024 , was employed for image acquisition. The exposure time was set to 35 milliseconds (ms) for image capture. The lens was configured with a focal length of 16 mm and an f-number of f/22. This high f-number was chosen to achieve a greater depth of field (DOF). Given the object's length, it is important to preserve the image sharpness across the object. In other words, the image should be sharp from the closest point to the farthest point of the object from the camera. This implies that a long DOF is required in the image [37], [38]. The camera and lighting are arranged to replicate the inspection method employed by the factory workers. The object was positioned between the camera and the lighting. Both imaging devices, the camera and the lighting, were at an angle relative to the object. This setup follows the bright-field illumination technique, as described in [39]-[41]. This method enhances the visibility of surface patterns through reflected light. The reason behind this technique was based on the geometry of the object being inspected. The interior of the tube can only be inspected from one side, as depicted in Figure 2(a). It is important to highlight that the interior shape is curved, forming a tubular structure. Since the camera is only directed to one side, the rolling bar, as shown in Figure 2(b), is utilized to rotate the object. Consequently, this enables the image capture of the entire inner surface.

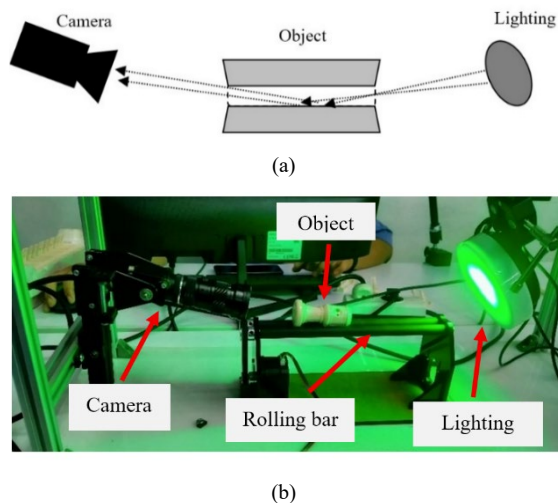


Figure 2. Set up of inspection. (a) Image acquisition diagram. (b) The experiment settings.

According to the experiment, the pattern seen in the image is amplified using green light. Figure 3 illustrates a sample image from the dataset, taken with the setup presented in Figure 2(b). As seen in Figure 3, the center region is noticeably brighter than the rest of the areas.

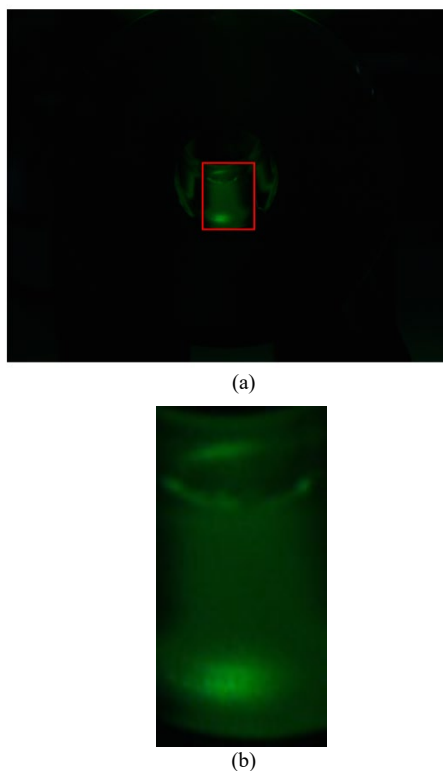


Figure 3. Example of the dataset. (a) The original captured image. (b) Cropped image based on ROI.

A red box indicates the Region of Interest (ROI) where the pattern is analyzed. The predefined ROI is cropped, as presented in Figure 3(b), and the feature extraction using ResNet50 is performed. In this study, the ROI size is 100×200 . According to the experiment, this is the optimal size for covering the entire pattern imprinted on the inner surface.

Additional examples from the dataset are shown in Figure 4. It can be seen that Figure 4(a) and (b) are both labeled as OK class, despite having noticeably distinct textures. A smooth texture in Figure 4(a) clearly classifies the image into the OK class. Meanwhile, based on the assessment from the quality control personnel, the image in Figure 4(b) is still accepted as OK despite the presence of a visible pattern. On the other hand, Figure 4(c) represents an NG image, where a noticeably strong pattern is evident. The provided examples underscore the difficulty in distinguishing those that pass and those that fail. The assessment carried out by workers can vary significantly due to subjective judgment. Figure 4(b) and Figure 4(c) may result in different judgments across workers. It is crucial for the model to provide highly discriminative image features between the two.

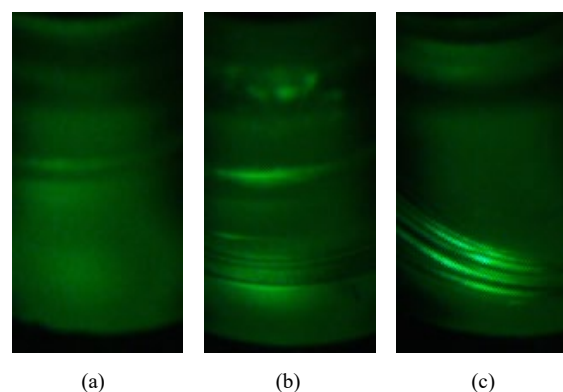


Figure 4. Example of cropped images. (a) The image labeled as OK has a smooth surface. (b) Image classified as OK despite a pattern being visible. (c) Image classified as NG.

2.3 ResNet50 and classifier

ResNet50 is a transfer learning model pre-trained on a large-scale dataset comprising up to 1,000 classes. As its name indicates, the architecture is composed of 50 trainable layers [42]-[44]. More precisely, this model is referred to as a deep network, enabling the model to learn more complex patterns. ResNet50 is expected to extract the fine features from the patterns, as displayed in Figure 4, which are challenging for quality control personnel. ResNet comes in several depths of layer, such as ResNet18, ResNet34, ResNet101, and ResNet152. Increasing the network's depth allows for the extraction of richer features. Yet, it demands greater computational resources and longer processing time. The choice of ResNet50 lies in its balance between performance and efficiency. The details of the ResNet50 layers are presented in Table 1 [45]-[47]. This study uses the default input size of 224×224 . Thus, the input image was resized to match the required size. The original ResNet50 model has a 1000-neuron fully connected layer, with a softmax activation function in its top layer. This is because the model was trained using the ImageNet dataset with 1000 classes. However, for the specific binary classification task in this study, the top layer is removed. Instead, a neural network (NN) for binary classification was created to achieve the goal. The ResNet50 is only used as a feature

extractor, resulting in a 2048 average-pooled flattened feature vector for each data. This vector is then passed to the NN classifier to obtain the prediction. The classifier architecture layers proposed in this research are given in Figure 5. The final layer of this classifier comprises two output neurons with a softmax activation function for binary classification.

Table 1. Layer Configuration of ResNet50.

Layer name	Output size	Configuration
Conv1	112×112	$7 \times 7, 64, \text{stride } 2$
Conv2	56×56	$3 \times 3 \text{ max pool, stride } 2$
Conv3	28×28	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
		$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
		$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
Conv4	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
Conv5	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$

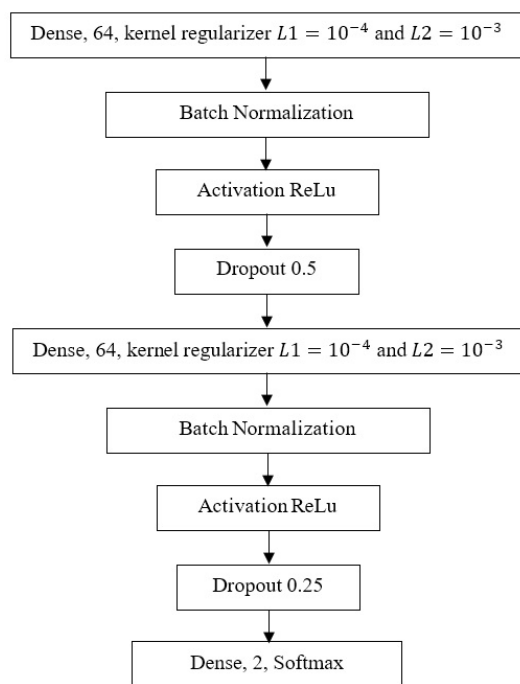


Figure 5. The architecture of NN classifier used in this research.

One of the crucial things is to avoid overfitting. In this situation, the model memorizes the training data, resulting in poor performance when classifying unseen data. To mitigate the risk of overfitting, two callback strategies were applied: reducing the learning rate on the plateau and early stopping. The model that learns well from the data typically will see its loss decrease over epochs. However, after a certain epoch is surpassed, the loss curve might slowly decrease as if it were a flat region or a plateau. To address this condition, the learning rate is reduced, allowing the model to discover a new minimum. The reduction of the learning rate will continue until the minimum learning rate is achieved. This technique has been demonstrated in [48]-[52]. Then, the epoch stops if the loss curve does

not exhibit significant improvement after multiple learning rate reductions. By implementing this approach, the time taken for the model's training is minimized. Moreover, it mitigates overfitting by early terminating the training process.

2.4. Data Augmentation

The classifier is required to recognize various characteristics of the input data. In practical scenarios, the new input data may exhibit different characteristics from those found in the dataset. The model would result in poor performance if it were trained on only the existing dataset. This is because the model failed to correctly classify the new data since the features were unrecognized. To address this challenge, data augmentation is a vital phase prior to the training stage. Data augmentation enriches the data variability by applying several transformations such as scaling, rotation, flipping, contrast and brightness change, cropping, etc. These transformations increase the dataset size and introduce a wider range of features for the model to learn, improving generalization. Overfitting prevention is also a benefit of data augmentation.

However, the decision of which augmentation is necessary to take depends on the nature of the data itself. For the case experienced in this work, the image is prone to variations like contrast, brightness, sharpness, cropping, and their combination. In this research, the dataset was augmented 5 times and then divided into two subsets, 80% for training and 20% for validation.

2.5. Evaluation metrics

The model's capability in classification is assessed using a confusion matrix, a widely adopted tool in supervised learning for evaluating key performance metrics such as accuracy, precision, recall (sensitivity), and specificity [53], [54]. The confusion matrix is visualized in Figure 6. In this context, the NG class is treated as the positive and the OK label is considered negative. The True Positive (TP) is when the model correctly predicts the NG image as NG, while the True Negative (TN) is the OK image being correctly predicted as OK. On the other hand, the False Positive (FP) occurs when the OK image is classified as NG, and the False Negative (FN) occurs when the NG image is predicted as OK. Reducing the value of FP and FN is a priority objective for many researchers. In contrast, a high number of TP and TN is highly favored. The values of the metrics are determined with the equations described in Formulas 1 - 4 [20], [30], [31], [55]. Based on the given definitions, a model that performs excellently will have all its metrics converging to 1 or 100%.

Another important metric for evaluating a model's performance is the area under the Receiver Operating Characteristic curve, known as ROC AUC, which ranges from 0 to 1. This metric reflects the model's ability to distinguish between the two classes. A low

AUC indicates an inseparable distribution, causing the model to struggle to classify the sample. Thus, a higher AUC is desirable since the two different categories are well separated [56], [57]

		Predicted value	
		+	-
Actual value	+	TP	FN
	-	FP	TN

Figure 6. Confusion matrix

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (3)$$

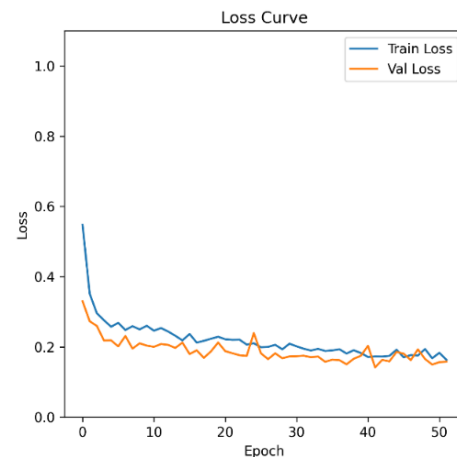
$$\text{Specificity} = \frac{TN}{TN+FP} \quad (4)$$

3. Results and Discussions

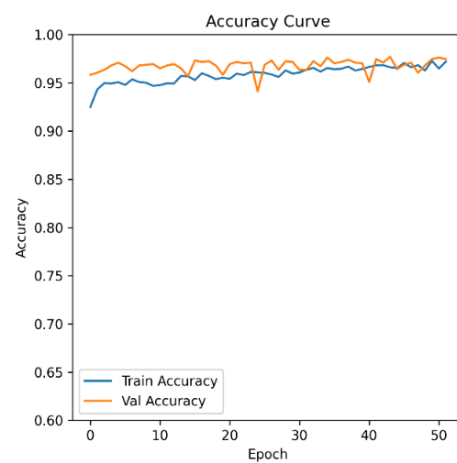
3.1 Training Stage

With the dataset provided in [58], it is divided into training and validation subsets, and the training phase was initiated. During the training, a set of data is given to the model, and the loss is measured. Here, the primary objective is to minimize the loss, or often referred to as error, by iterative learning. The results of the training stage are provided in Figure 7. It is clearly seen in Figure 7 (a) that the training loss and validation loss are relatively high at the beginning of the training and rapidly drop within the initial few epochs. Simultaneously, the accuracy curves experience a significant improvement, as shown in Figure 7 (b). This indicates that the model effectively learns the data from the training set. Although a slight gap remains between training and validation loss throughout the process, both curves finally converge at a value slightly below 0.2 when the training terminates at around epoch 50. Additionally, the accuracy curves reach around 0.97 or 97% at the end of the training phase. Figure 7(c) displays the declining learning rate throughout the training. As described in the method section, the algorithm will reduce the learning rate to find a new minimum if the loss decays very slowly. This result suggests that the model updates the learning rate multiple times, indicating stagnation in the loss curves at certain epochs. In this implementation, the learning rate is dynamically adjusted after five epochs without meaningful improvement in the loss curves. Meanwhile, the early termination of training is executed after ten epochs of persistent loss stagnation, in the situation where the minimum learning rate limit is achieved. Continuing the training beyond 50 epochs potentially causes the model to overfit. Thus, the early stopping callback terminates the training phase since no

further improvement is observed in training loss and validation loss.



(a)



(b)

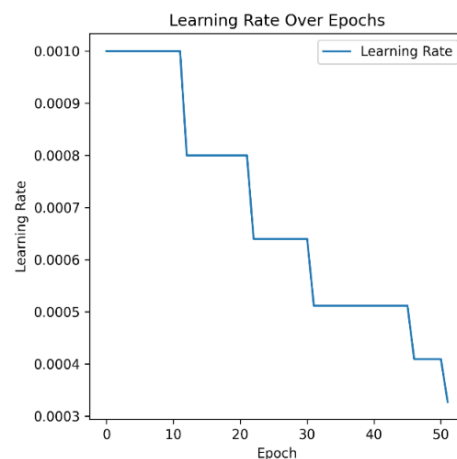


Figure 7. The training results. (a) Loss curve during training. (b) Accuracy curve. (c) The decrease in learning rate during training.

3.2 Testing Stage

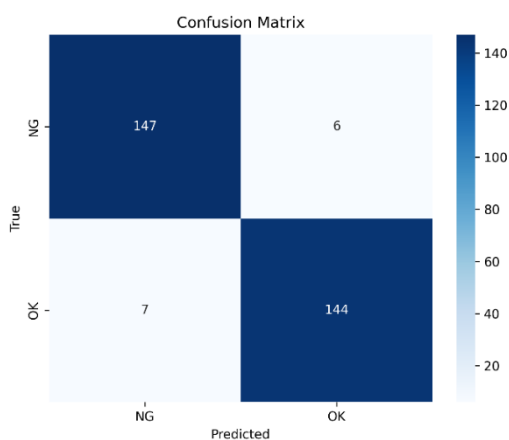
The model that has been trained and reached minimum loss, both in training data and validation data, is used to classify a set of unseen images. As previously mentioned, these images were taken independently from those used in the training stage. The results of the

testing phase are summarized in the confusion matrix shown in Figure 8 (a).

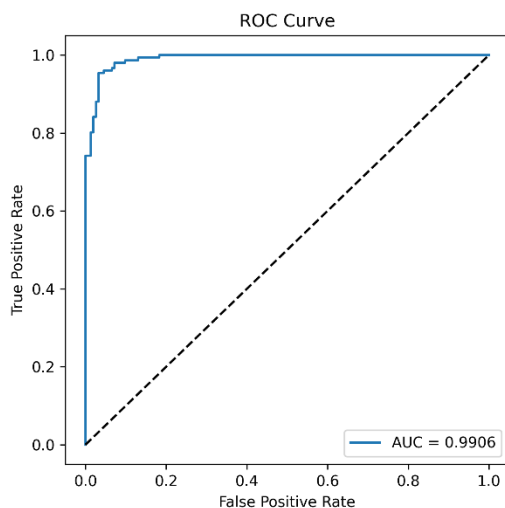
According to Formulas 1 - 4, the classification results are summarized in Table 2. It is seen that all metrics have very close values, ranging from 95% to around 96%. This suggests that the model's performance is balanced and effectively discriminates between NG and OK patterns. It is further confirmed by the ROC curve from the test data in Figure 8(b), which gives the AUC of 0.9906.

Table 2. The classification results.

Metrics	Value (%)
Accuracy	95.72
Precision	95.45
Sensitivity	96.07
Specificity	95.36



(a)



(b)

Figure 8. The testing results. (a) Confusion matrix. (b) ROC AUC from testing data.

The ROC curve confirms ResNet50's effectiveness in distinguishing feature patterns. These measurement outcomes show a considerable potential for real-world application. By the definition described in the evaluation metrics section, a lower FN leads to higher sensitivity. In this study, a lower FN has a significant

impact on the manufacturer's brand reputation. An FN occurs when a defective product (NG) is detected as OK, which may result in defective products reaching customers. Consequently, it leads to some complaints. Similarly, reducing FP leads to higher precision. It occurs when an OK object is classified as NG. Since a defective product will be recycled, lowering FP will gain the manufacturer's profit by minimizing unnecessary recycling. It is critical to maintain the level and the balance of the metrics. This machine vision can potentially be seamlessly integrated with a mechanical or robotics system that passes the object to the inspection area, enabling an automatic visual inspection.

4. Conclusions

This study has developed a visual inspection system for recorder manufacturing. It demonstrates the feasibility of implementing ResNet50 in musical instrument production, a domain rarely addressed in the literature. The system is designed to automate the detection of surface defects, previously carried out through manual inspection. The evaluation shows strong results with an accuracy of 95.7%, precision of 95.45%, sensitivity of 96.07%, and specificity of 95.36%. Additionally, the area under the curve or AUC is 0.9906. These findings suggest that the proposed model exhibits a strong capability to differentiate between the NG and OK images. High and balanced metrics values can influence the manufacturer's brand reputation and profit. By integrating the proposed system in the production process, it could refine the quality control process through human error reduction and faster inspection. Despite showing high performance, the technique explained in this paper requires multiple image captures to inspect the entire inner surface. Furthermore, this system is restricted to pattern assessment and cannot evaluate the product's shape and size. Future studies should explore some approaches that enable full inner-surface capture in a single step, while also incorporating the shape and size analysis. Moreover, the use of the latest deep learning architecture is recommended for exploration. Such improvements could shorten inspection time and enhance productivity.

Acknowledgements

The authors would like to express their gratitude to Universitas Airlangga for the financial support provided under the grant number 1005/B/UN3.FTMM/PT.01.03/2024.

References

- [1] G. Mao, Q. Zhixin, N. Jianwei, W. Yu, Z. Xiaojun, and T. Zhi, "Research on Weld Defect Detection of Multi-Scale Normalized Gray-Level Co-Occurrence Matrix," in *ICSMD 2023 - International Conference on Sensing, Measurement and Data Analytics in the Era of Artificial Intelligence, Proceedings, Institute of Electrical and Electronics Engineers Inc.*, 2023. doi: 10.1109/ICSMD60522.2023.10490684.

- [2] Y. F. Tang and B. Zheng, "Thread Surface Defect Detection Based on Image Processing," in *IEEE Information Technology and Mechatronics Engineering Conference, ITOEC*, Institute of Electrical and Electronics Engineers Inc., 2025, pp. 1423–1427. doi: 10.1109/ITOEC63606.2025.10967740.
- [3] J. Meng and S. Wen, "Detection of Steel Surface Defects Based on Improved YOLOv8n Algorithm," in *2024 2nd International Conference on Algorithm, Image Processing and Machine Vision (AIPMV)*, IEEE, Jul. 2024, pp. 8–12. doi: 10.1109/AIPMV62663.2024.10692098.
- [4] J. Esclapés, A. Gómez, and A. Ibañez, "Flow. A socially responsible 3d printed one-handed recorder," *Int J Environ Res Public Health*, vol. 18, no. 22, Nov. 2021, doi: 10.3390/ijerph182212200.
- [5] D. LASOCKI, R. EHRLICH, N. TARASOV, and M. PETRI, "INTRODUCTION," in *The Recorder*, Yale University Press, 2022, pp. XI–XIV. doi: 10.2307/j.ctv335kw78.5.
- [6] S. Sankararaman, "Analysis of musical notes of flute: A recurrence-based phase space approach," *Chinese Journal of Physics*, vol. 83, pp. 113–122, Jun. 2023, doi: 10.1016/j.cjph.2023.02.004.
- [7] N. Toradmal, V. Bairagi, and D. Jadhav, "Feature Extraction of Unprinted Fabric for Defect Detection Using Gabor Filter, HOG and SVM," in *2024 International Conference on Emerging Smart Computing and Informatics, ESCI 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/ESCI59607.2024.10497237.
- [8] C. Wu, X. Zou, and Z. Yu, "A Detection Method for Wood Surface Defect Based on Feature Fusion," in *2022 4th International Conference on Frontiers Technology of Information and Computer, ICFTIC 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 876–880. doi: 10.1109/ICFTIC57696.2022.10075158.
- [9] T. Van Le, J. D. Joshua, T. Kim, J. Lee, S. H. Kim, and J. Y. Lee, "Surface defect detection using distributed features," *Eng Appl Artif Intell*, vol. 151, p. 110655, Jul. 2025, doi: 10.1016/j.engappai.2025.110655.
- [10] Z. Ren, "Surface Defect Detection and Classification Methods for Aluminum Profiles Based on CNN," in *Proceedings - 2024 17th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/CISP-BMEI64163.2024.10906210.
- [11] R. Fu, Y. Zhang, D. Novák, A. Strauss, and M. Cao, "A lightweight convolutional neural network-based model and system for defect detection and navigation on bridge road surface," *Advances in Engineering Software*, vol. 208, Oct. 2025, doi: 10.1016/j.advengsoft.2025.103972.
- [12] F. Ding et al., "Research on ultrasonic non-destructive detection method for defects in GFRP laminates based on machine learning," *Ocean Engineering*, vol. 327, May 2025, doi: 10.1016/j.oceaneng.2025.120972.
- [13] P. Natho, S. Boonying, P. Bonguleaum, N. Tantidontanet, and L. Chamuthai, "An enhanced machine vision system for smart poultry farms using deep learning," *Smart Agricultural Technology*, vol. 12, Dec. 2025, doi: 10.1016/j.atech.2025.101083.
- [14] K. Pawikhum, Y. Yang, L. He, and P. Heinemann, "Development of a Machine vision system for apple bud thinning in precision crop load management," *Comput Electron Agric*, vol. 236, Sep. 2025, doi: 10.1016/j.compag.2025.110479.
- [15] W. Han, "Research on Quality Monitoring Platform and Resource Scheduling Algorithm Based on Machine Vision," in *Procedia Computer Science*, Elsevier B.V., 2025, pp. 1064–1071. doi: 10.1016/j.procs.2025.04.060.
- [16] M. F. Mogro, F. A. Jacome, G. M. Cruz, and J. R. Zurita, "View of Sorting Line Assisted by A Robotic Manipulator and Artificial Vision with Active Safety," *Journal of Robotics and Control*, vol. 5, no. 2, 2024, doi: 10.18196/jrc.v5i2.20327.
- [17] S. M. Kiran and D. N. Chandrappa, "Plant Leaf Disease Detection Using Efficient Image Processing and Machine Learning Algorithms," *Journal of Robotics and Control (JRC)*, vol. 4, no. 6, pp. 840–848, 2023, doi: 10.18196/jrc.v4i6.20342.
- [18] Y. Liu, Y. Zheng, T. Wei, and Y. Li, "Lightweight algorithm based on you only look once version 5 for multiple class defect detection on wind turbine blade surfaces," *Eng Appl Artif Intell*, vol. 138, Dec. 2024, doi: 10.1016/j.engappai.2024.109422.
- [19] A. Macharla and G. Pradeepini, "A novel framework for the diagnosis of Parkinson's disease using transfer learning with RESNET50 and SVM classifier," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 32, no. 2, pp. 877–886, 2023, doi: 10.11591/ijeecs.v32.i2.pp877-866.
- [20] R. M. Abd-Elghaffar, M. El-Zalabany, H. E. D. Moustafa, and M. El-Seddek, "Classification of focal liver disease in egyptian patients using ultrasound images and convolutional neural networks," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 27, no. 2, pp. 793–802, Aug. 2022, doi: 10.11591/ijeecs.v27.i2.pp793-802.
- [21] H. Imaduddin, I. C. Utomo, and D. A. Anggoro, "Fine-tuning ResNet-50 for the classification of visual impairments from retinal fundus images," *International Journal of Electrical and Computer Engineering*, vol. 14, no. 4, pp. 4175–4182, Aug. 2024, doi: 10.11591/ijece.v14i4.pp4175-4182.
- [22] A. Rath, B. S. P. Mishra, and D. K. Bagal, "ResNet50-based Deep Learning model for accurate brain tumor detection in MRI scans," *Next Research*, vol. 2, no. 1, p. 100104, Mar. 2025, doi: 10.1016/j.nexres.2024.100104.
- [23] D. Anggriandi and A. Sunyoto, "Multi-Class Classification of Skin Diseases Using ResNet50," in *2023 6th International Conference on Information and Communications Technology, ICOIACT 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 349–354. doi: 10.1109/ICOIACT59844.2023.10455953.
- [24] W. Prastiwinarti, M. K. Delimayanti, H. Kurniawan, Y. P. Pratama, H. Wendho, and R. Adi, "Efficient packaging defect detection: leveraging pre-trained vision models through transfer learning," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 34, no. 3, pp. 2096–2106, Jun. 2024, doi: 10.11591/ijeecs.v34.i3.pp2096-2106.
- [25] Q. Yu, Y. Zhang, J. Xu, Y. Zhao, and Y. Zhou, "Intelligent damage classification for tensile membrane structure based on continuous wavelet transform and improved ResNet50," *Measurement (Lond)*, vol. 227, Mar. 2024, doi: 10.1016/j.measurement.2024.114260.
- [26] K. S. Vishal, S. K. Srinidhi, U. S. Shashank, and R. P. Singh, "Enhanced Small Object Detection in Remotely Sensed Images using Hybrid ResNet50 and Feature Pyramid Network," in *Procedia Computer Science*, Elsevier B.V., 2025, pp. 1813–1822. doi: 10.1016/j.procs.2025.04.433.
- [27] L. Tan and Y. Hu, "Accurate Detection and Classification of Surface Defects in Electric Porcelain Insulators Based on Deep Learning Intelligent Algorithms," in *Procedia Computer Science*, Elsevier B.V., 2025, pp. 582–588. doi: 10.1016/j.procs.2025.04.248.
- [28] X. Zou, C. Wu, H. Liu, and Z. Yu, "Improved ResNet-50 model for identifying defects on wood surfaces," *Signal Image Video Process*, vol. 17, no. 6, pp. 3119–3126, Sep. 2023, doi: 10.1007/s11760-023-02533-y.
- [29] N. Sagar et al., "Precision detection of grapevine downy and powdery mildew diseased leaves and fruits using enhanced ResNet50 with batch normalization," *Comput Electron Agric*, vol. 232, May 2025, doi: 10.1016/j.compag.2025.110144.
- [30] N. A. Sehree and A. M. Khidhir, "Olive trees cases classification based on deep convolutional neural network from unmanned aerial vehicle imagery," *Indonesian Journal of Electrical Engineering and Computer Science*,

- vol. 27, no. 1, pp. 92–101, Jul. 2022, doi: 10.11591/ijeecs.v27.i1.pp92-101.
- [31] A. S. Hatem, M. S. Altememe, and M. A. Fadhel, "Identifying corn leaves diseases by extensive use of transfer learning: a comparative study," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 29, no. 2, pp. 1030–1038, Feb. 2023, doi: 10.11591/ijeecs.v29.i2.pp1030-1038.
- [32] P. Bohlol, S. Hosseinpour, and M. Soltani Firouz, "Improved food recognition using a refined ResNet50 architecture with improved fully connected layers," *Curr Res Food Sci*, vol. 10, Jan. 2025, doi: 10.1016/j.crf.2025.101005.
- [33] F. Tempola, R. Wardoyo, A. Musdholifah, Rosihan, and A. K. Nugroho, "Implementation of Deep Learning With Resnet50 Modification for Clove classification," in *Proceeding - IEEE 9th Information Technology International Seminar, ITIS 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, doi: 10.1109/ITIS59651.2023.10419960.
- [34] G. Singh, K. Guleria, and S. Sharma, "Advanced Fruit Sorting: Pre-trained ResNet50 Model for Rotten and Fresh Fruit Classification," in *2024 4th Asian Conference on Innovation in Technology (ASIANCON)*, IEEE, Aug. 2024, pp. 1–5, doi: 10.1109/ASIANCON62057.2024.10837782.
- [35] "Sadimantara2024-View of The Classification of Aflatoxin Contamination Level in Cocoa Beans using Fluorescence Imaging and Deep learning".
- [36] I. A. M. Diones, W. Castillo, P. N. Santiago, M. P. Geronimo, J. A. MacAso, and S. D. C. Domingo, "Performance of ResNet50 in Classifying Artificial Intelligence Generated Arts," in *Proceedings of the International Colloquium on Signal Processing and Its Applications, CSPA*, Institute of Electrical and Electronics Engineers Inc., 2025, pp. 361–366, doi: 10.1109/CSPA64953.2025.10933073.
- [37] M. R. Rai and J. Rosen, "Depth-of-field engineering in coded aperture imaging," *Opt Express*, vol. 29, no. 2, p. 1634, Jan. 2021, doi: 10.1364/oe.412744.
- [38] X. Qian et al., "Towards Single-Lens Controllable Depth-of-Field Imaging via Depth-Aware Point Spread Functions," *IEEE Trans Comput Imaging*, 2025, doi: 10.1109/TCI.2025.3544019.
- [39] M. Gerges and X. Chen, "Adaptive Lighting for Curved and Nonuniform Objects in Optomechanical Inspection Systems," *IEEE/ASME Transactions on Mechatronics*, vol. 27, no. 6, pp. 5792–5802, Dec. 2022, doi: 10.1109/TMECH.2022.3189344.
- [40] Q. Sun, K. Xu, H. Liu, and J. Wang, "Unsupervised surface defect detection of aluminum sheets with combined bright-field and dark-field illumination," *Opt Lasers Eng*, vol. 168, Sep. 2023, doi: 10.1016/j.optlaseng.2023.107674.
- [41] V. Marin-Lizarraga, R. Rodríguez-García, J. L. García-Cordero, D. May-Arrijo, C. De-la-Peña, and L. D. Patiño-López, "Simultaneous bright-field and fluorescence lensless imaging with high excitation light extinction for microfluidics applications," *Opt Lasers Eng*, vol. 185, Feb. 2025, doi: 10.1016/j.optlaseng.2024.108724.
- [42] Z. T. Omer and A. H. Abbas, "Image anomalies detection using transfer learning of resnet-50 convolutional neural network," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 27, no. 1, pp. 198–205, Jul. 2022, doi: 10.11591/ijeecs.v27.i1.pp198-205.
- [43] D. Thekedath and R. R. Sedamkar, "Detecting Affect States Using VGG16, ResNet50 and SE-ResNet50 Networks," *SN Comput Sci*, vol. 1, no. 2, Mar. 2020, doi: 10.1007/s42979-020-0114-9.
- [44] C. Sahaya Pushpa Sarmila Star, T. M. Inbamalar, and A. Milton, "Automatic semantic segmentation of breast cancer in DCE-MRI using DeepLabV3+ with modified ResNet50," *Biomed Signal Process Control*, vol. 99, Jan. 2025, doi: 10.1016/j.bspc.2024.106691.
- [45] K. Hanapi, A. Setyanto, and A. Dwi Hartanto, "Comparison of ResNet50 and SqueezeNet1.1 for Plastic Bottles and Cans Classification," in *Proceeding - 6th International Conference on Information Technology, Information Systems and Electrical Engineering: Applying Data Sciences and Artificial Intelligence Technologies for Environmental Sustainability, ICITISEE 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 525–529, doi: 10.1109/ICITISEE57756.2022.10057634.
- [46] A. Noel, W. Vincent, and J. E. Piou, "Resnet50 to Detect Landslides, Damaged Infrastructures and Crumbled Houses from Haiti 2010 and 2021 Earthquakes," in *Proceedings of the 2023 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology, IAICT 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 214–219, doi: 10.1109/IAICT59002.2023.10205657.
- [47] S. D. Bhatt, H. B. Soni, H. R. Kher, and T. D. Pawar, "Automated System for Lung Nodule Classification Based on ResNet50 and SVM," in *2022 3rd International Conference on Issues and Challenges in Intelligent Computing Techniques, ICICT 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, doi: 10.1109/ICICT55121.2022.10064515.
- [48] J. Park, D. Yi, and S. Ji, "A novel learning rate schedule in optimization for neural networks and it's convergence," *Symmetry (Basel)*, vol. 12, no. 4, Apr. 2020, doi: 10.3390/SYM12040660.
- [49] H. Hazimeh and N. Ponomareva, "Mind the (optimality) Gap: A Gap-Aware Learning Rate Scheduler for Adversarial Nets," 2023. [Online]. Available: <https://github.com/google-research/>
- [50] W. An, J. Zhang, J. Liu, K. Shao, K. Bian, and J. Zhou, "Learning Rate Decay Scheduler Using the Loss Operator as Generalization," in *2024 IEEE 5th International Conference on Pattern Recognition and Machine Learning, PRML 2024*, Institute of Electrical and Electronics Engineers Inc., 2024, pp. 70–75, doi: 10.1109/PRML62565.2024.10779896.
- [51] M. Cui, Y. Li, X. Wang, J. Wang, and H. Rong, "Rapid Reconstruction of Acoustic Holograms Using Improved IU-Net with Adaptive Learning Rate Scheduling," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3496923.
- [52] L. Wen, L. Gao, X. Li, and B. Zeng, "Convolutional Neural Network with Automatic Learning Rate Scheduler for Fault Classification," *IEEE Trans Instrum Meas*, vol. 70, 2021, doi: 10.1109/TIM.2020.3048792.
- [53] J. Erban, P. É. Portier, E. Egyed-Zsigmond, and D. Nurbakova, "Confusion Matrices: A Unified Theory," *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3507199.
- [54] J. Xu, Y. Zhang, and D. Miao, "Three-way confusion matrix for classification: A measure driven view," *Inf Sci (N Y)*, vol. 507, pp. 772–794, Jan. 2020, doi: 10.1016/j.ins.2019.06.064.
- [55] D. Valero-Carreras, J. Alcaraz, and M. Landete, "Comparing two SVM models through different metrics based on the confusion matrix," *Comput Oper Res*, vol. 152, Apr. 2023, doi: 10.1016/j.cor.2022.106131.
- [56] H. Zhu, S. Liu, W. Xu, J. Dai, and M. Benbouzid, "Linearithmic and unbiased implementation of DeLong's algorithm for comparing the areas under correlated ROC curves," *Expert Syst Appl*, vol. 246, Jul. 2024, doi: 10.1016/j.eswa.2024.123194.
- [57] A. R. Rachakonda and A. Bhatnagar, "ARatio: Extending area under the ROC curve for probabilistic labels," *Pattern Recognit Lett*, vol. 150, pp. 265–271, Oct. 2021, doi: 10.1016/j.patrec.2021.06.023.
- [58] Rizki Putra Prastio, Rodik Wahyu Indrawan, Vanesia Tasib, and Zhilaan Abdurasyid Rusmawan, "Recorder Defect Detection," Mendeley Data. [Online]. Available: <https://data.mendeley.com/datasets/hxtcnt9jr/1>