KOMTEKINFO JOURNAL
LPPM Universitas Putra Indonesia YPTK Padang
Jl. Raya Lubuk Begalung, Padang, West Sumatera, Indonesia, Zip Code: 25221
Volume: 13, Issue: 2, Page: 95 - 101, June 30th, 2025, e-ISSN: 2502-8758
Available online at website: https://jkomtekinfo.org/ojs/index.php/komtekinfo

Lossy Image Compression Analysis using JPEG Algorithm on Rice Leaf Disease Dataset

Wahyu Saptha Negoro[1]✉, Ratih Adinda Destari[2], Asbon Hendra Azhar[3], Fhery Agustin[4]
[1234]Computer Science, Faculty of Engineering and Computer Science, Potensi Utama University, Medan, North Sumatra, Indonesia.
wahyusaptha1707@gmail.com

Abstract

The development of digital image processing technology has increased the need for more efficient image data storage and transmission, especially in the field of smart agriculture that utilizes digital images as a source of information. Rice leaf disease image datasets generally have quite large file sizes, requiring compression techniques to save storage capacity and speed up the data exchange process. This study aims to analyze the performance of lossy image compression using the JPEG algorithm on rice leaf disease datasets by evaluating file size efficiency and the quality of the compressed images. The research process includes collecting rice leaf disease image datasets, applying JPEG compression at several quality levels (quality factors), and measuring performance using the Compression Ratio (CR), Peak Signal-to-Noise Ratio (PSNR), and Mean Squared Error (MSE) parameters. The analysis results show that the lower the quality factor value used, the greater the compression ratio obtained, but the visual quality of the image decreases as indicated by the increase in the MSE value and the decrease in the PSNR value. Conversely, a higher quality factor is able to maintain image quality with the consequence of a larger file size. The findings of this study indicate that the JPEG algorithm is able to provide a good compromise between storage efficiency and visual image quality so that it remains suitable for use in rice leaf disease datasets, especially as a preprocessing stage in image processing and artificial intelligence systems for plant disease classification.

Keywords: JPEG, Lossy Compression, Rice Leaf Disease, *Quality Evaluation, Quality Factor*

1. Introduction

Rice (Oryza sativa L.) is a strategic food commodity that is a staple food source for more than half of the world's population. In Indonesia, rice plays an important role in maintaining national food security, so increasing productivity and quality of harvests is a top priority in agricultural sector development [1]. However, rice plant productivity still faces various challenges, one of which is the attack of leaf diseases such as Brown Spot, Leaf Blast, Bacterial Leaf Blight, Narrow Brown Spot, and several other diseases that can cause a significant decrease in crop yields if not detected early [2]. The development of these diseases is often difficult to observe manually because it requires expert personnel, a relatively long time, and a significant cost. Therefore, the use of digital image processing technology and artificial intelligence is a solution that is increasingly being applied in modern agriculture [3].

Advances in Computer Vision and Artificial Intelligence have made it possible to identify plant diseases automatically through digital image analysis. Various recent studies show that models based on Convolutional Neural Network (CNN), Vision Transformer (ViT), or a combination of both are able to produce a very high level of accuracy in classifying rice leaf diseases [4,5]. Apart from increasing the speed of identification, this approach is also able to help farmers make decisions more quickly so that losses due to the spread of disease can be minimized [6].

However, the increasing development of image-based disease detection systems has significantly increased the need for digital data storage. Rice leaf disease image datasets typically consist of thousands to tens of thousands of high-resolution images, requiring large storage capacities [7]. Furthermore, distributing datasets over the internet or using them on mobile devices becomes less efficient if the image size is too large. This presents a challenge in developing smart agricultural systems, which require both storage efficiency and data transmission speed [8].

One approach widely used to address these issues is digital image compression techniques. Image compression aims to reduce file size without losing information that is important for analysis purposes. In general, image compression is divided into two categories: lossless compression and lossy compression. In lossless compression, all image information can be completely recovered after the decompression process. Conversely, lossy compression removes some visual

Submitted: 20-03-2026 | Revised: 19-04-2026 | Accepted: 20-05-2026 | Published: 30-06-2026

information that is considered less important for human perception, thus producing a much smaller file size compared to lossless methods. Among various lossy compression algorithms, the JPEG algorithm remains the most widely used international standard to date [9]. JPEG utilizes the Discrete Cosine Transform (DCT) to convert spatial information into the frequency domain. Next, a quantization process is carried out so that high-frequency components that have less influence on visual perception can be reduced. The final stage, namely entropy coding, produces a smaller file size while maintaining visual quality at a certain level.

Although the JPEG algorithm has been widely used in various multimedia applications due to its efficient data storage and transmission, its application to rice leaf disease image datasets still requires further study. JPEG is a lossy compression method that reduces image information through a quantization process, potentially eliminating high-frequency details, such as leaf surface texture, disease spot patterns, color changes, and other visual characteristics that are important in the process of feature extraction and Computer Vision-based classification. Several studies have shown that increasing the compression level can reduce image quality and impact classification accuracy if relevant visual information is also lost. In addition, in plant image analysis applications, the selection of the compression level is an important factor because it can affect the results of phenotypic characteristic measurements and disease detection. Therefore, an evaluation is needed regarding the limit of the JPEG compression level that is still able to maintain image quality so that disease information can still be optimally recognized by Computer Vision and Deep Learning models. [10]. Evaluation of the quality of compression results is generally carried out using several quantitative parameters, including Compression Ratio (CR) to measure the efficiency of file size reduction, Mean Squared Error (MSE) to determine the level of pixel error against the original image, Peak Signal-to-Noise Ratio (PSNR) to measure the quality of image reconstruction results, and Structural Similarity Index Measure (SSIM) to evaluate the level of visual structure similarity between the original image and the compressed image. The combination of these four parameters is able to provide a more comprehensive picture of the quality of the compression algorithm used [11].

In this study, the compression method used is the JPEG algorithm, which is one of the lossy image compression standards that is still most widely used today because it is able to provide a good balance between storage efficiency and visual quality. JPEG utilizes the Discrete Cosine Transform (DCT) to transform image data from the spatial domain to the frequency domain, then performs a quantization process on high-frequency coefficients so that the data size can be significantly reduced. The effectiveness of the JPEG algorithm makes it still relevant as a compression method in various applications, including the management of the rice leaf disease image dataset used in this study.

Recent studies have focused more on improving the accuracy of rice leaf disease classification using various Deep Learning architectures, such as CNN, MobileNet, EfficientNet, Vision Transformer, and CNN-ViT hybrid models [12,13,14]. Most of these studies used original datasets without evaluating the effect of image compression on the quality of the data used as input for the classification model. Thus, there is still a research gap regarding the performance analysis of JPEG compression algorithms on rice leaf disease datasets, especially in assessing the balance between storage efficiency and visual image quality [15].

Based on the description, this study proposes a lossy image compression analysis using the JPEG algorithm on a rice leaf disease dataset. The study focuses on the effect of Quality Factor (QF) variations on file size, Compression Ratio, MSE, and PSNR. The results of the study are expected to be a reference in determining the optimal compression level so that a balance between storage efficiency and image quality is obtained. In addition, the results of this study can also be a basis for further research that develops an Artificial Intelligence-based rice leaf disease classification system using compressed images as input data [16,17].

Most research in 2024–2025 is still focused on improving the accuracy of rice leaf disease classification using deep learning methods. Research specifically evaluating the impact of JPEG compression on the quality of rice leaf disease images before they are used in analysis and classification is still very limited. Therefore, research is needed that analyzes the relationship between compression level, storage efficiency, and visual image quality using objective parameters such as PSNR, MSE, and Compression Ratio.

2. Methods

2.1. Digital Image Processing

Digital image processing is a branch of computer science that studies techniques for acquiring, processing, analyzing, improving, and extracting information from digital images using computational algorithms. In contrast to analog image processing, all processes in digital images are carried out on numerical representations in the form of pixel matrices, thus enabling automatic and more accurate image manipulation. Digital image processing not only aims to improve the visual quality of an image, but also produces information that can be used by computer systems for various purposes, such as object recognition, segmentation, classification, and artificial intelligence-based decision making [18].

Advances in precision agriculture technology have increased the use of digital image processing in various agricultural activities. This technology allows for automatic identification of plant conditions, eliminating the need for time-consuming and labor-intensive manual inspections. Image processing can detect early disease symptoms through analysis of leaf color, texture, shape, and damage patterns.

In this study, digital image processing serves as the basis for the image compression process using the JPEG algorithm. Rice leaf disease datasets obtained from various sources will go through a preprocessing stage before being compressed. Next, the compressed images are evaluated using the Compression Ratio (CR), Mean Squared Error (MSE), and Peak Signal-to-Noise Ratio (PSNR) parameters to determine the balance between storage efficiency and image visual quality. This analysis is expected to provide recommendations for the optimal JPEG compression level so that the file size becomes smaller without losing important visual information, such as leaf color, surface texture, and disease spot patterns required in further analysis based on Computer Vision and Artificial Intelligence.

2.2. Image Compression

Image compression is a technique in digital image processing that aims to reduce the file size of an image while maintaining the visual quality and important information contained therein. The main purpose of image compression is to increase the efficiency of storage media usage, speed up the data transmission process over the network, and reduce bandwidth requirements without eliminating the main function of images as a medium for conveying information. As the Internet of Things (IoT), cloud computing, artificial intelligence, and Computer Vision technologies develop, the need for image compression techniques is becoming increasingly important because the volume of digital image data produced continues to increase significantly [19].

The general stages of the Lossy Compression Digital Image Compression Process can be seen in Figure 1 below.

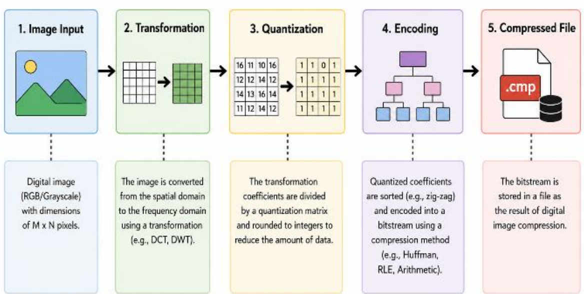


Figure 1. Digital Image Compression Process

In general, the compression process is carried out by utilizing the redundancy characteristics found in digital images. This redundancy includes coding redundancy, interpixel redundancy, and psychovisual redundancy. By eliminating or reducing this redundancy, data size can be reduced without losing information that has important value for users or computer systems. According to Dong and Pan (2023), most image data used in various modern applications is stored and transmitted in compressed form so that the compression process becomes an integral part of digital image processing systems [20].\

2.3. JPEG Algorithm (Joint Photographic Experts Group)

JPEG (Joint Photographic Experts Group) is a lossy compression type digital image compression standard that was developed to reduce image file size while maintaining visual quality that is acceptable to the human eye. The main principle of JPEG is to remove visual information that is less sensitive to the human visual system so that the file size is much smaller than the original image. Until now, JPEG is still the most widely used image compression format in multimedia applications, digital photography, web, and Computer Vision systems because it has a good balance between compression ratio and image quality [21].

2.3. Quality Evaluation

Quality evaluation is an important stage in image compression research because it is used to determine the extent to which the compression algorithm is able to reduce file size without losing important visual information. In lossy image compression, some image information is indeed lost so that quantitative measures are needed to measure the balance between compression efficiency and the quality of the reconstructed image. [21]. Evaluation of the quality of compressed images is generally carried out using objective metrics that compare the original image with the compressed image. The most widely used metrics are Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Compression Ratio (CR) because they are able to provide an overview of pixel errors, visual quality, similarity of image structure, and efficiency of file size reduction.

2.3.1 Compression Ratio (CR)

Compression Ratio shows how much the file size is reduced after the compression process.

$$CR=\frac{V_d+V_c}{V_c} \tag{1}$$

The higher the CR value, the higher the compression rate. However, increasing the compression ratio is usually accompanied by a decrease in visual quality.

Therefore, image compression research always seeks the best trade-off between file size and image quality.

2.3.2 Mean Squared Error (MSE)

Mean Squared Error (MSE) is a measure of the average squared error between the pixel values of the original image and the compressed image. The MSE value indicates the magnitude of the difference in pixel intensity after the compression process. The smaller the MSE value, the smaller the resulting error, thus improving the quality of the compressed image. Conversely, a high MSE value indicates that the compression process caused significant pixel changes.

The MSE formula is

$$\frac{1}{n} = \sum_{i=1}^{n}(Y_i - Y_i)^2 \quad (2)$$

According to a comprehensive review on compressed image quality evaluation, MSE remains a widely used basic indicator because it is simple and has a direct relationship with PSNR, although this metric does not fully reflect human visual perception.

2.3.3 Peak Signal-to-Noise Ratio (PSNR)

Peak Signal-to-Noise Ratio (PSNR) is the most commonly used metric to measure the quality of a compressed image. PSNR indicates the ratio of the maximum signal value (pixels) to the error level (noise) introduced by the compression process. PSNR values are expressed in decibels (dB).

$$PSNR = 10log_{10}\frac{R2}{MSE} \quad (3)$$

Interpretation:
40 dB: Excellent
30–40 dB: Good
20–30 dB: Fair
<20 dB: Poor

MAX = 255 for 8-bit images. MSE = Mean Squared Error.
The higher the PSNR value, the closer the compressed image quality is to the original image. In JPEG-based image compression research, the combination of these four metrics provides a comprehensive evaluation. MSE and PSNR measure numerical error, while Compression Ratio evaluates storage efficiency. This approach remains the best practice in current image compression research.

3. Results and Discussions

This study uses a quantitative experimental method that aims to analyze the performance of the JPEG (Joint Photographic Experts Group) algorithm in performing lossy image compression on a rice leaf disease dataset. The analysis is carried out by comparing file sizes before and after compression and evaluating image quality using the Compression Ratio (CR), Mean Squared Error (MSE), and Peak Signal-to-Noise Ratio (PSNR) parameters. The experimental method was chosen because it allows objective measurements of the effect of JPEG Quality Factor (QF) variations on image quality and storage efficiency. A similar approach is widely used in research on digital image processing and Computer Vision-based plant disease classification. The research flow can be seen in Figure 2 below.

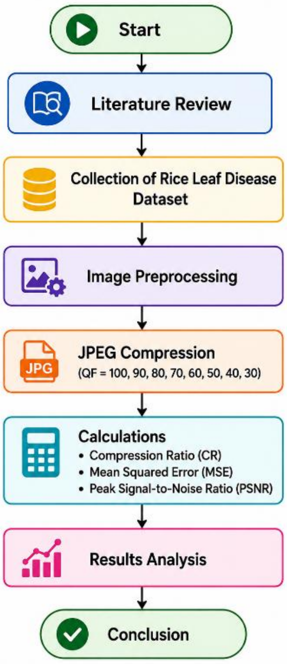


Figure 2. Research flow

The research phase began with a literature review on digital image processing, JPEG algorithms, and current research on rice leaf diseases. Next, dataset collection, preprocessing, compression using several Quality Factor values, and image quality evaluation were performed using four objective parameters.

3.1. Research Dataset

The dataset used in this study consists of rice leaf disease images taken directly from Pulau Keren Village. The dataset consists of Brown Spot, Leaf Blast, Bacterial Leaf Blight, and Healthy Leaf Blight, with a total of 420 images of rice diseases and healthy leaves. In this study, all images were converted to .jpg format to accommodate the compression process.

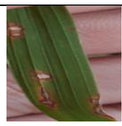
Brown Spot	Leaf Blast	Bacterial	Healthy
			

Figure 3. Research Dataset

In Figure 3 above is a dataset of rice diseases, where the sample data presented is 12 images.

3.2. Preprocessing Stages

The preprocessing stages aim to standardize data before compression. These include Data Acquisition (collecting rice leaf images), Data Selection (removing blurry, corrupted, and duplicate images), Resizing (aligning all pixels in the image), Normalization (ensuring all images are in RGB color format), and Storage (storing all images in JPG format). The preprocessing stages can be seen in Figure 4 below.

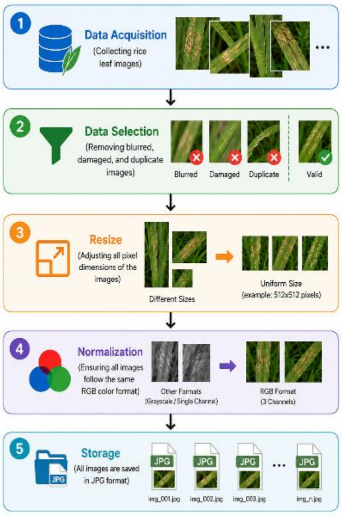


Figure 4. Preprocessing stages

3.3. JPEG compression

The stages of JPEG compression can be seen in Figure 5 below.

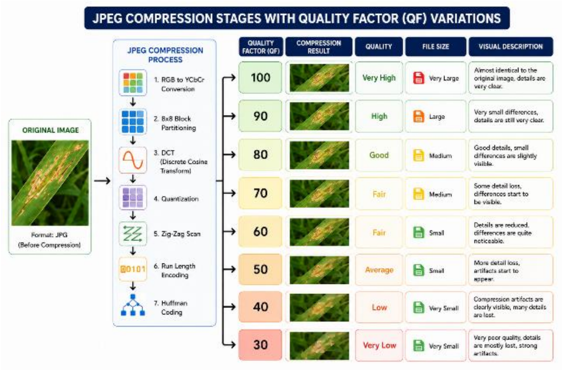


Figure 5. JPEG compression

The smaller the Quality Factor value, the smaller the file size, the Compression Ratio increases, the PSNR decreases, and the MSE increases. In this study, the Quality Factor used is 30% to see the quality that occurs in the image and is continued at the evaluation stage of the results. The original size obtained with QF 30% is 179.6171875 KB while the resulting Compression size is: 96.2314453125 KB.

3.4. Evaluation of Results

The evaluation was conducted using three parameters. The evaluation results used in this study can be seen in Figure 6 below.

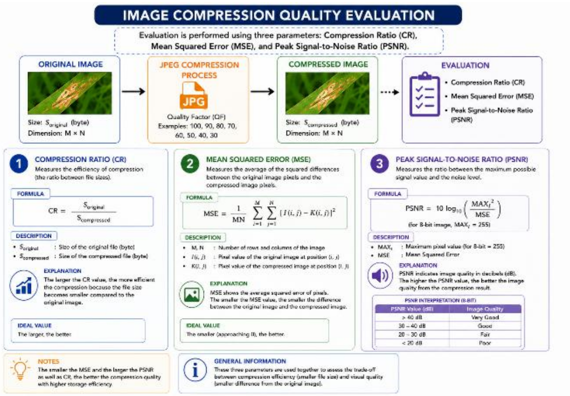


Figure 6. Evaluation of Results

The evaluation results used were Compression Ratio, MSE, and PSNR. Based on the evaluation results used and the research conducted in compressing the rice leaf disease dataset images, the results obtained are in Figure 7 below.

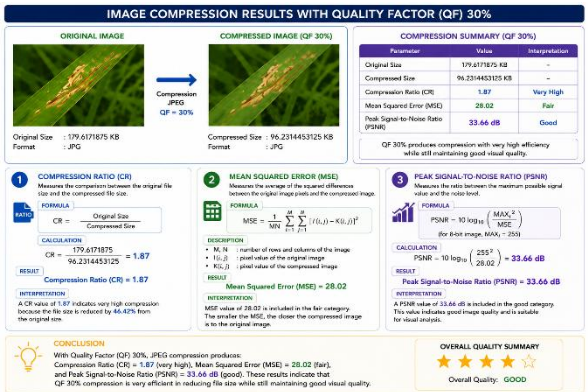


Figure 7. Compression Ratio, MSE, and PSNR Results

Based on the results obtained as shown in Figure 7 above, the QF 30% image results still provide good visual quality.

3.5. Data analysis

The data was analyzed using a quantitative descriptive approach with the following steps:

- Calculate file sizes before and after compression.
- Calculate Compression Ratio.
- Calculate MSE.

- Calculate PSNR.
- Compare all values for each Quality Factor (QF).
- Determine the optimum QF value based on the balance between compression efficiency and visual quality.

3.4. Displaying Histogram

At this stage, the researcher will display the difference in the histogram of the original image with the histogram of the compressed image, which can be seen in Figure 8 below.

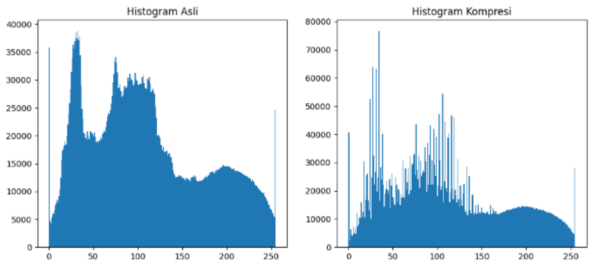


Figure 8. Histogram Comparison

Based on Figure 8 above, it shows the difference produced by a QF of 30% on the image. However, it does not damage the image quality.

3.5. Displaying Compression Results

At this stage, the researcher will display several compression results from QF (10, 30, 50, 70, and 90), and the differences can be seen. The compression results can be seen in Figure 9 below.

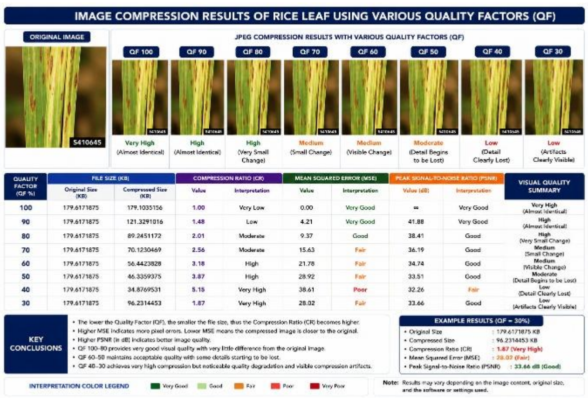


Figure 9. QF Compression Results (10,30,50,70,90)

Based on Figure 9 above, it shows a significant difference in QF results in terms of the quality and capacity of the resulting image. Therefore, it can be concluded that a QF of 30% is sufficient with fairly good quality, while the compression capacity is very high.

4. Conclusions

Based on the test results carried out by researchers from various QFs, the QF of 30% shows quite good image quality results and very high compression results, which are suitable for use on large datasets if the storage is not large.

Acknowledgements

The author would like to thank Potensi Utama University for its assistance in this research and to the farmers and village heads in Pulau Sejuk Village, Lima Puluh District, Batu Bara Regency, Medan, who have accepted and provided time and permission to be able to conduct research on rice leaf diseases. He also expressed his gratitude to the publisher of the scientific journal for their assistance in correcting typos and completing the manuscript.

Author Contributions Statement

Contributor Role Taxonomy (CRediT) to recognize individual author contributions, the contributions made can be seen in the following table.

Name of Author	C	M	So	Va	Fo	I	R	D	W
Wahyu Saptha Negoro	✓	✓	✓	✓	✓	✓	✓	✓	✓
Asbon Hendra Azhar		✓			✓	✓	✓	✓	✓
Fhery Agustin	✓		✓	✓			✓		✓

References

[1] H. M. Yusuf, S. A. Yusuf, A. H. Abubakar, M. Abdullahi, and I. H. Hassan, “A systematic review of deep learning techniques for rice disease recognition: Current trends and future directions,” Franklin Open, vol. 8, Art. no. 100154, 2024, doi: 10.1016/j.fraope.2024.100154.

[2] M. K. Barnwal, A. Kotasthane, N. Magculia, et al., “A review on crop losses, epidemiology and disease management of rice brown spot to identify research priorities and knowledge gaps,” European Journal of Plant Pathology, vol. 136, no. 3, pp. 443–457, 2013, doi: 10.1007/s10658-013-0195-6.

[3] M. M. Hasan, A. F. M. S. Uddin, M. R. Akhond, M. J. Uddin, M. A. Hossain, and M. A. Hossain, “Machine learning and image processing techniques for rice disease detection: A critical analysis,” International Journal of Plant Biology, vol. 14, pp. 1190–1207, 2023, doi: 10.3390/ijpb14040087.

[4] J. Lu, L. Tan, and H. Jiang, “Review on convolutional neural network (CNN) applied to plant leaf disease classification,” Agriculture, vol. 11, no. 8, Art. no. 707, 2021, doi: 10.3390/agriculture11080707.

[5] R. Mukherjee, A. Ghosh, C. Chakraborty, J. N. De, and D. P. Mishra, “Rice leaf disease identification and classification using machine learning techniques: A comprehensive review,”

Engineering Applications of Artificial Intelligence, vol. 139, pt. B, Art. no. 109639, 2025, doi: 10.1016/j.engappai.2024.109639.

[6] G. K. V. L. Udayananda, C. Shyalika, and P. P. N. V. Kumara, "Rice plant disease diagnosing using machine learning techniques: A comprehensive review," SN Applied Sciences, vol. 4, Art. no. 311, 2022, doi: 10.1007/s42452-022-05194-7.

[7] Z. Salman, A. Muhammad, M. J. Piran, and D. Han, "Crop-saving with AI: Latest trends in deep learning techniques for plant pathology," Frontiers in Plant Science, vol. 14, Art. no. 1224709, 2023, doi: 10.3389/fpls.2023.1224709.

[8] Y. Wang, U. S. Rajkumar Dhamodharan, N. Sarwar, et al., "A hybrid approach for rice crop disease detection in agricultural IoT system," Discover Sustainability, vol. 5, Art. no. 99, 2024, doi: 10.1007/s43621-024-00285-4.

[9] G. K. Wallace, "The JPEG still picture compression standard," Communications of the ACM, vol. 34, no. 4, pp. 30–44, 1991, doi: 10.1145/103085.103089.

[10] W.-L. Lau, Z.-L. Li, and K. W.-K. Lam, "Effects of JPEG compression on image classification," International Journal of Remote Sensing, vol. 24, no. 7, pp. 1535–1544, 2003, doi: 10.1080/01431160210142842.

[11] V.-I. Ungureanu, P. Negirla, and A. Korodi, "Image-compression techniques: Classical and 'region-of-interest-based' approaches presented in recent papers," Sensors, vol. 24, no. 3, Art. no. 791, 2024, doi: 10.3390/s24030791.

[12] S. Upadhyay and R. Prasad, "Efficient-ViT B0Net: A high-performance lightweight transformer for rice leaf disease recognition and classification," Journal of the Nigerian Society of Physical Sciences, 2024.

[13] R. K. Rachman, D. I. M. Setiadi, A. Susanto, K. Nugroho, and H. M. M. Islam, "Enhanced Vision Transformer and transfer learning approach to improve rice disease recognition," Journal of Computing Theories and Applications, 2024, doi: 10.62411/jcta.10459.

[14] T. W. Utami, M. Novita, and K. Latifa, "Implementation and comparative analysis of CNN and transfer learning models (EfficientNetB0, MobileNetV2, and ResNet50) for rice leaf disease detection based on digital images," Journal of Applied Informatics and Computing, vol. 9, no. 6, 2025, doi: 10.30871/jaic.v9i6.11616.

[15] A. Julianto and M. Jannah, "Evaluasi komparatif lightweight convolutional neural network untuk klasifikasi penyakit daun dan hama tanaman padi," Building of Informatics, Technology and Science (BITS), vol. 7, no. 3, pp. 1844–1853, 2025, doi: 10.47065/bits.v7i3.8698.

[16] S. Jamil, "Review of image quality assessment methods for compressed images," Journal of Imaging, vol. 10, no. 5, Art. no. 113, 2024, doi: 10.3390/jimaging10050113.

[17] A. Sharma, B. Rajesh, and M. Javed, "Detection of plant leaf disease directly in the JPEG compressed domain using transfer learning technique," in Advanced Machine Intelligence and Signal Processing, D. Gupta, K. Sambyo, M. Prasad, and S. Agarwal, Eds. Singapore: Springer, 2022, pp. 365–376, doi: 10.1007/978-981-19-0840-8_30.

[18] Z. Mahmood, "Digital image processing: Advanced technologies and applications," Applied Sciences, vol. 14, no. 14, Art. no. 6051, 2024, doi: 10.3390/app14146051.

[19] Y. Dong and W. D. Pan, "A survey on compression domain image and video data processing and analysis techniques," Information, vol. 14, no. 3, Art. no. 184, 2023, doi: 10.3390/info14030184.

[20] M. K. Ibraheem, A. V. Dvorkovich, and I. M. A. Al-khafaji, "A comprehensive literature review on image and video compression: Trends, algorithms, and techniques," International Journal of Information Systems and Informatics, 2024, doi: 10.18280/isi.290307.

[21] M. Dutta et al., "Rice leaf disease classification—A comparative approach using CNN, CAAR-U-Net, and MobileNet-V2 architectures," Technologies, 2024.

Biographies of Authors

	**Wahyu Saptha Negoro** is a full-time lecturer at the Faculty of Computer Science and Engineering, Potensi Utama University, Medan. He earned his Doctorate in Information Technology from Putra Indonesia University "YPTK", Padang. His research interests include Digital Image Processing and Computer Vision. He can be contacted via email at wahyusaptha1707@gmail.com
	**Asbon Hendra Azhar** is a full-time lecturer at the Faculty of Computer Science and Engineering, Potensi Utama University, Medan. He earned his Master's degree in Informatics Engineering from Putra Indonesia University "YPTK", Padang. His research interests include Decision Support Systems and Data Mining. He can be reached by email at asbon.upu@gmail.com
	**Fhery Agustin** is a He is a full-time lecturer at the Faculty of Computer Science and Engineering, Potensi Utama University, Medan. He earned his Master's degree in Informatics Engineering from Putra Indonesia University "YPTK", Padang. His research interests include Decision Support Systems and Data Mining. He can be reached by email at fheryagustin@gmail.com