Detection of Rice Leaf Pests Based on Images with Convolution Neural Network in Yollo v8

Ahmad Fauzi[1]*, Kiki Ahmad Baihaqi[2], Anggun Pertiwi[3], Yudo Devianto[4], Saruni Dwiasnati [5],
Fakultas Ilmu Komputer, Universitas Buana Perjuangan Karawang[1], [2], [3]
Karawang, Indonesia
Fakultas Ilmu Komputer, Universitas Mercu Buana [4][5]
Jakarta, Indonesia
afauzi@ubpkarawang.ac.id[1], kikiahmad@ubpkarawang.ac.id[2], anggun.pertiwi@ubpkarawang.ac.id[3],
yudo.devianto@mercubuana.ac.id[4], saruni.dwiasnati@mercubuana.ac.id[5]

Abstract— Detection of rice leaf pests is important in agriculture because it can help farmers determine appropriate preventive measures. One method that can be used to detect rice leaf pests is digital image processing technology. In this research, proof of suitability for solving this case was carried out between the Convolutional Neural Network (CNN) algorithm which was run offline with R-CNN and YOLOv8 for detecting rice leaf pests. At the data preparation stage, images of rice leaves were taken from various sources with a total of 100 images taken from website data and 10 images taken from the research site. Next, preprocessing and data augmentation are carried out to improve image quality and increase data variation. At the model training stage, a training and evaluation process is carried out using two types of algorithms, namely R-CNN and YOLOv8. The accuracy of the testing results using the same data using Yolov8 obtained 87.0% accuracy and 79% precision, while using R-CNN the results obtained were 85% for accuracy and 75% for precision with data divided into 80 training data 20 validation data and 10 testing data. Labeling the dataset uses Makesensei which has been completely standardized, with the resulting parameters being the spots on rice leaves.

Keywords— rice leaf pests, digital image processing, convolution neural network, yolo v8

I. INTRODUCTION

Rice leaf pests are one of the problems often faced by farmers in rice cultivation. Rice leaf pests can cause damage to the leaves so that the growth of rice plants is disrupted and crop yields decrease. Therefore, detecting rice leaf pests is important in efforts to prevent and control these pests.

In the current digital era, digital image processing technology can be an alternative for detecting rice leaf pests. One of the methods used in digital image processing technology uses Yolo v8. Yolo v8 is a real-time object detection algorithm developed by the developer team from Ultralytics, which can be modified according to the needs of research and object detection. These objects can be even small objects[1].

Machine Learning is a technology that develops algorithms to analyze and understand the characteristics of data based on special algorithms. This allows the algorithm to produce output that is in accordance with the stated objectives. Designed for self-learning, Machine Learning technology does not require direct instruction from the user. Its development is rooted in various fields of science, including statistics, mathematics, and data mining, facilitating machines to learn from data without the need for reprogramming by admins or users. Machine Learning algorithms have become important in data processing, including activities such as data mining and image data processing [2]. Machine Learning is often applied in identifying or classifying diseases [3] and [4]. In recent years, significant progress in machine learning has been achieved in various sectors [5]. Convolutional Neural Network (CNN) is a method that falls into the category of popular Machine Learning methods, generally used in classification tasks using image data [6].

Research that has been carried out using Yollo V5 for safety helmet detection shows that Yollo V5 has time efficiency in detection and makes it easier to carry out the desired parameters. The research also explains that labeling is easy to use in training data, as well as having a fairly high level of accuracy reaches 90% [7].

Meanwhile, other research discusses the detection of fruit objects, which prioritizes fruit freshness detection with an accuracy of 92% and uses datasets sourced from websites, so it is not data taken directly by researchers [8]. There was other research, namely regarding rice leaf pest detection using CNN in 2018 with an accuracy of 78.2% with 5 convolution layers, with the lowest reaching 58.4% [9].

Disadvantages found in this study include difficulties in identifying pests in the early phases of an attack and the inability to effectively monitor large areas, which are exacerbated by environmental factors such as changes in weather, light intensity, and leaf background color that can hinder the pest detection process. Additionally, despite advances in technology such as remote sensing and methods based on Machine Learning, there are still challenges regarding accuracy and efficiency. For example, machine learning models often require extensive and varied training datasets to operate optimally, which can be difficult to obtain.

There is not much research comparing the performance of the CNN and YOLOv8 algorithms in detecting rice leaf pests. Therefore, this research was conducted to compare the performance between the CNN algorithms and YOLOv8 in detecting rice leaf pests. It is hoped that the results of this

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v13i1.2008, Copyright ©2024
Submitted : October 30, 2023, Revised : December 27, 2023, Accepted : January 11, 2024, Published : February 15, 2024

research will provide useful information for farmers in choosing an effective and efficient rice leaf pest detection method.

The main difference between these two methods lies in the basic approach to detecting objects. CNN approaches such as Faster R-CNN tend to be more complex and require multiple processing steps, while YOLO treats object detection as a direct regression problem on the entire image using only one processing step. Both have their advantages and disadvantages, depending on the specific needs of the object detection task.

II. METHOD

Data in the form of pest images obtained via the website as a dataset, it cannot be directly processed for detection or can be indirectly detected by Yolo and the convolution neural network algorithm. The first stage carried out was to augment and then label the pest variants using the website of the labeling service provider. Data augmentation is a stage of image data processing, augmentation is the process of changing or modifying an image in such a way that the computer detects that the modified image is a different image, but humans can always know that the edited image is the same [10]. Improvement may occur. increasing the accuracy of the trained CNN model, because with this increase, the model receives additional data, which can help in creating a model that can generalize better [11].

Conventional feature extraction networks usually utilize standard convolution blocks to form complex convolutional neural networks. However, the uniform structure of these convolutional neural networks is less effective in extracting the characteristics of various rice pests and often loses important features after several layers of convolution. Therefore, we developed a new feature extraction network that integrates deformation convolution into Yolov8, which enables more efficient feature extraction of rice pests, especially for pests with various shapes, in the detection process.

Improvements made in this research include flipping the image horizontally, flipping the image with a tilt every 5 degrees [10], zooming randomly, with a maximum zoom of 50% of the image size, and also rotating the image randomly up to a maximum of 90. One type of enhancement that is commonly done is to rotate the image by a certain amount [11]. So you get a new variant of the same image. And enrich the image database.

A. Preprocessing Data

The pre-processing stage is quite an important stage in image processing. The preprocessing is carried out to optimize image quality, thereby making it easier and increasing the system's ability to identify objects. The preprocessing stages in this system are divided into two stages, namely original data and augmented data. On the original data, preprocessing is only carried out by resizing the image. Meanwhile, pre-processing for data augmentation is carried out by resizing and augmenting the data [12]. Data processing is important before the data is applied to machine learning. This step ensures that the data is of good quality when used to train the model. Image data generator techniques are used in processing the Alzheimer's Disease dataset. This technique is used to overcome small dataset sizes and produces additional tens of thousands of images [13].

The use of the CNN algorithm in this research is based on research conducted by Tsung-Ren Huang et al., who claim that the flow of the CNN process is the best. The concept can be seen in Figure 2.

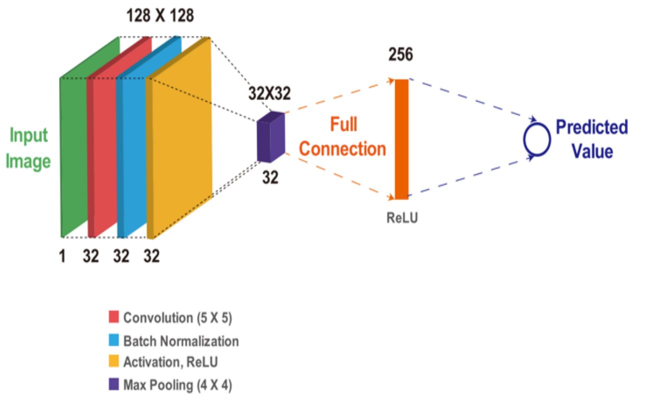

Figure 2. Layers in CNN

The input image is then trained and processed in the CNN algorithm through layers and a convolution process until the prediction results are obtained [14]. Once completed, it will go through the classification stage, which is discussed in the stages in this sub-chapter. CNN, or Convolutional Neural Network, provides a variety of transfer learning models commonly used in image processing research, including models such as VGG-16, ResNet-50, and DenseNet-201, which are often adopted by researchers in the field [15]. CNN has a good capacity to learn image representation [16]. The CNN model can produce feature maps sequentially by grouping simple patterns consisting of nodes and edges from the first layer of input data. By performing convolutional operations on the input layer using trainable kernels, this feature map is generated. Pooling and non-linear modifications operate together to support network convergence. After that, an estimate is made using the map features that have been processed.

CNN is a type of neural network architecture used for image processing and spatial pattern recognition. CNN can understand hierarchical features in images through convolutional, pooling, and fully connected layers. For object detection, CNN is often used with approaches such as (Region-based CNN) or known as Faster R-CNN. This approach involves complex processing stages, including feature extraction with convolution, region proposition, and final classification.

B. Yolo v8

Yolo is an algorithm in digital image processing technology, especially in the sector of deep learning. The stages carried out in Yolo in detecting objects are detecting objects by dividing the image into sxs-sized regions or grids, then detecting with each bounding box having 5 information values, and detection by means of each grid will predict the probability class value if it is predicted that there is an object in it. [17].

The first stage that Yolo carried out was Model Initialization, Preparing a predetermined neural network

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v13i1.2008, Copyright ©2024
Submitted : October 30, 2023, Revised : December 27, 2023, Accepted : January 11, 2024, Published : February 15, 2024

architecture for detecting leaf pest objects. Then Input Processing is Inserting images into the model. then Feature Extraction, which involves convolution layers to extract important features from the image. Then Predictive Processing Predictions are carried out for each cell (grid cell) in the image, including information about the bounding box and object class. Next, the Non-Maximum Suppression (NMS) process is reducing redundancy in detection by eliminating overlapping or competing predictions and the output is providing output in the form of bounding boxes and labels for detected objects along with their precision.

The architecture of Yolo version eight can be seen in Picture 3.

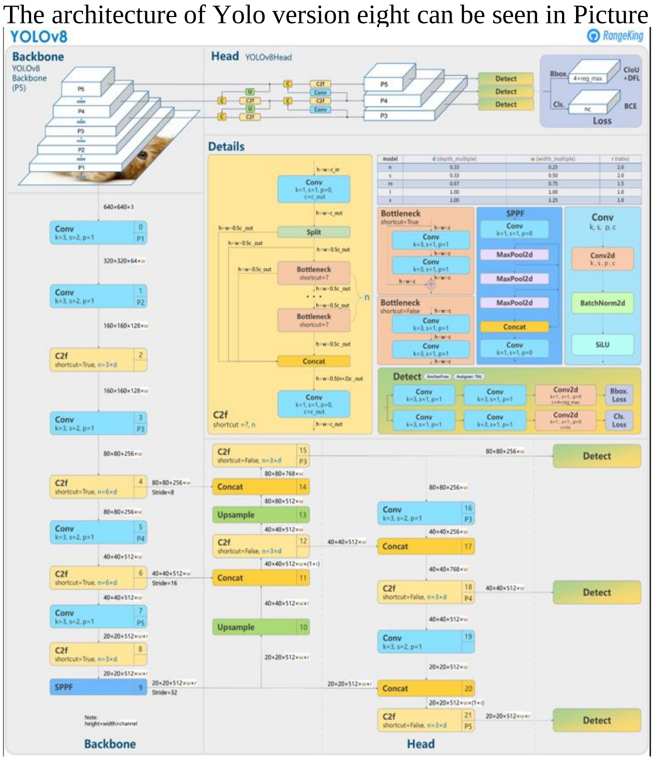

Figure 3. architecture yolo 8

YOLO is a different object detection approach. This approach designs object detection as a regression problem on the entire image, rather than as a series of sequential image processing tasks. YOLO divides the image into cells (grid cells) and outputs object and bounding box predictions from each cell at once. YOLO enables real-time object detection because it only requires one processing step to generate predictions.

C. Clarification

Image classification is one of the basic concepts in computer vision. It classifies images based on their message, which can be analyzed after translation. Image classification is also key to image detection, image segmentation, visual object tracking, behavior analysis, and many other high-level visual tasks. There are also a large number of image classification applications, such as face recognition, video content analysis, traffic scene recognition, content-based image retrieval, and automatic photo classification. Image classification describes the entire image through handcrafted features, or feature learning, and uses classifiers to categorize objects in the image. Therefore, how to extract features becomes important. Before deep learning was widely used, there was a word classification called “Bag of Words”. This method was introduced in natural language, where word bags function as a feature. For images, this “Bag of Words” method requires creating a dictionary. The simplest model of the Bag of Words framework can be a process of low-level feature extraction, feature coding, and classifier design [18], [19].

In contrast, image classification, based on deep learning, can replace the process of manually designing or selecting image features with hierarchical characterization through supervised or unsupervised learning.

Convolution Neural Networks have achieved incredible achievements in recent years. CNN uses image pixels as input, and it retains almost all the information from the original image. Feature extraction and high-level abstraction through convolution make the model output the result of image recognition. This I/O-based learning method has been a great success and is widely used [20].

Before 2012, although the main procedure was categorized into three parts, as mentioned before, usually the complete process of building an image recognition model included low-level feature learning, feature coding, spatial feature restriction, and classifier design. The flow can be seen in Figure 3, and more or less, the depiction of the system is as follows:

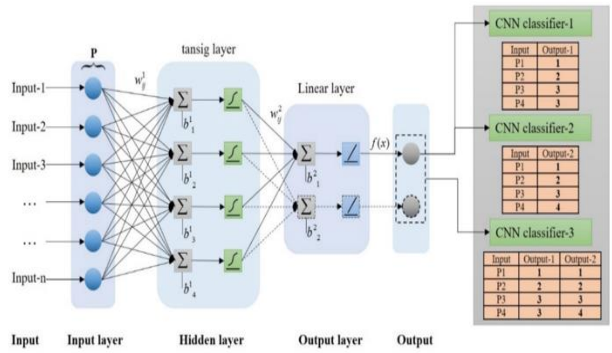

Figure 4. Classification flow illustration

Then the results of the classification will produce accuracy and precision, which will be the resulting parameters of the algorithm's performance in this case of rice leaf pests [21]. The formula for calculating it can be seen in formulas 1, 2, and 3.

$$\text{Akurasi} = \frac{\text{TP}}{\text{TP+TN+FP+F}} \text{ X } 100\% \quad (1)$$

$$\text{Presisi} = \frac{\text{TP}}{\text{TP+FP}} \text{ x } 100\% \quad (2)$$

$$\text{Loss} = \frac{\text{FP+FN}}{\text{TP+TN+FP+F}} \quad (3)$$

Explanation :

1. True Positive (TP): Data that is correctly classified by the system as a positive value (true).
2. True Negative (TN): Data that is correctly classified by the system as a negative value (false).
3. False Positive (FP): The data is wrong (negative) but is classified as true data (positive).
4. False Negative (FN): Data is correct (positive) but classified as wrong data (negative).

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v13i1.2008, Copyright ©2024
Submitted : October 30, 2023, Revised : December 27, 2023, Accepted : January 11, 2024, Published : February 15, 2024

III. RESULTS AND DISCUSSION

The results of the training and validation of the images were formed into a dataset, and then the results obtained from the detection of the two types of disease were taken as the theme in this paper.

This research uses the large-scale benchmark dataset IP102 Rice Pests published by CVPR in 2019 for pest identification as experimental data, formats it according to the VOC data set format, and uses Img labels to label categories, and coordinate pest information in the image. The data set contains nine types of rice pests, namely rice leafroller, rice leafworm, Asiatic rice borer, rice yellow borer, rice gall midge, brown planthopper, white-backed planthopper, rice water beetle and rice leafhopper. According to the number of images in each category, the training set and test set are divided in a ratio of 8:2. In this paper, an image enhancement method is used to improve the generalization capabilities of the model, including horizontal flip and horizontal rotation. Meanwhile, the data used as testing data was taken from agricultural land where the research was conducted. There are only two categories of diseases that can be diagnosed. So the class division only uses two classes. Namely bacterial and leaf blast.

TABLE 1. Detection Result

Image	type of plant disease
	leaf bacteria
	Leaf blast
	Leaf bacteria

After it is known that it is detected with a confidence level of 0.5, you can see a graph of the F1-Score, which shows that blight is easier to detect and classify than leaf blast. This can be seen in Figure 5, namely the curve resulting from the F1 score.

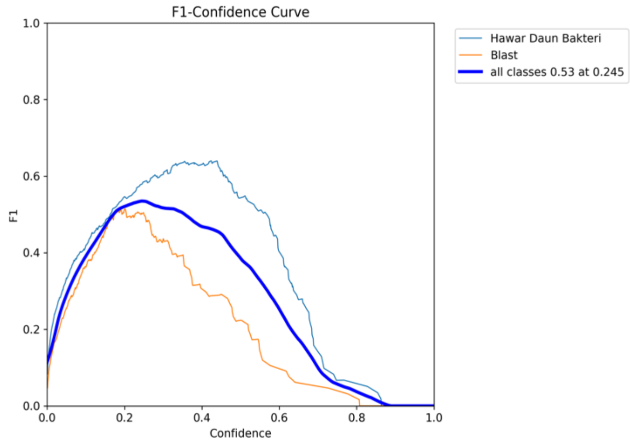

Figure 5. Kurva F1 Score

Meanwhile, the results from Precision Confidence show that all types of disease reached 0.79, with the results can be seen in the curve depicted in Figure 6.

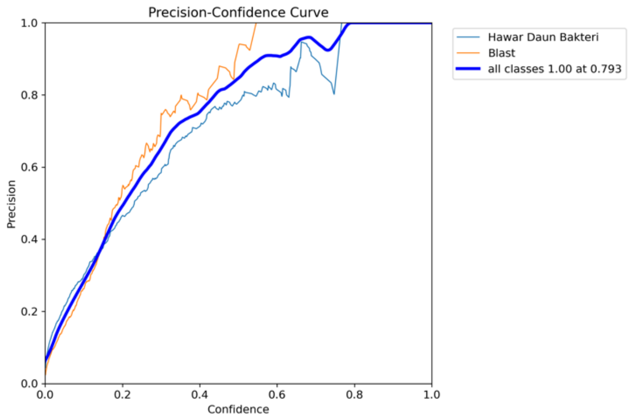

Figure 6. Precision confidance curve

Meanwhile, the precision recall results for bacterial leaf blight were 0.58, and for blast, they reached 0.45, which can be seen in Figure 7.

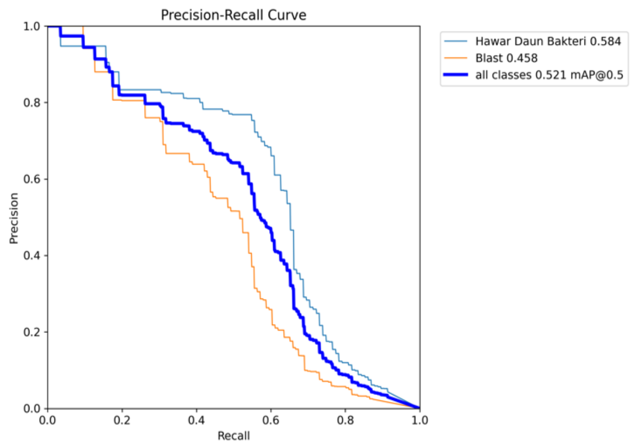

Figure 7. Precission recall

Apart from displaying precision, in this case we also display

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v13i1.2008, Copyright ©2024
Submitted : October 30, 2023, Revised : December 27, 2023, Accepted : January 11, 2024, Published : February 15, 2024

loss from various training and testing results, starting from train loss to validation loss. Which is depicted in figure 8.

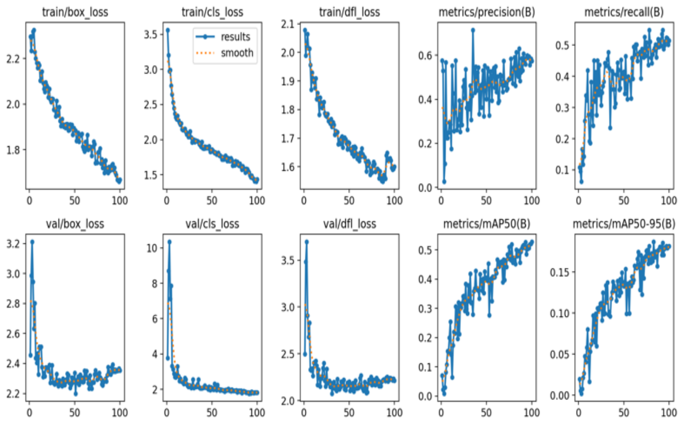

Figure 8. Loss grafik

seen in Figure 8 that the loss graph shows a good decrease during training but increases again in the validation data, this could indicate overfitting which is shown in the figure during training and validation. This happened because the parameters in this study were too broad, resulting in an accuracy that was below the average of 90%..

IV. Conclusion

The research results for the classification of rice leaf diseases had an accuracy of 87.0%, a loss of 0.1208, and a precision of 79%. for Yolo while the results from CNN are close to 85% with a precision of 78%. These results were obtained using augmented data with the best parameters. namely Optimizer RMSprop, learning rate 0.01, batch size 32, and epoch 100. The expected accuracy results can occur with a combination of datasets taken from the website and test data taken from the research area. This study does not compare the algorithm to other cases. just compare it with the dataset that has been taken from the research object on agricultural land as test data. So in this research it was found that Yolo V8 is more suitable which divides the image into cells (grid cells) and outputs object and bounding box predictions from each cell at once. YOLO enables real-time object detection because it only requires one processing step to generate predictions.

References

[1] Q. Wang, J. Tu, Y. Luo, and F. Liu, "Target Detection of Myrica rubra Based on Convolutional Neural Network and Transfer Learning," in Proceedings - 2021 International Conference on Electronic Information Technology and Smart Agriculture, ICEITSA 2021, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 485–490. doi: 10.1109/ICEITSA54226.2021.00098.

[2] D.-N. Le, V. S. Parvathy, D. Gupta, A. Khanna, J. J. P. C. Rodrigues, and K. Shankar, "IoT enabled depthwise sepa- rable convolution neural network with deep support vector machine for COVID-19 diagnosis and classification," Interna- tional Journal of Machine Learning and Cybernetics, vol. 12, pp. 3235–3248, Nov. 2021.

[3] M. Arabahmadi, R. Farahbakhsh, and J. Rezazadeh, "Deep Learning for Smart Healthcare—A Survey on Brain Tumor Detection from Medical Imaging," Sensors, vol. 22, p. 1960, Mar. 2022.

[4] P. K. Chahal, S. Pandey, and S. Goel, "A survey on brain tumor detection techniques for MR images," Multimedia Tools and Applications, vol. 79, pp. 21771–21814, Aug. 2020.

[5] Z. Qu, C. Cao, L. Liu, and D.-Y. Zhou, "A Deeply Supervised Convolutional Neural Network for Pavement Crack Detection With Multiscale Feature Fusion," IEEE Transactions on Neural Networks and Learning Systems, vol. 33, pp. 4890–4899, Sept. 2022.

[6] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, and Y. Fan, "A deep learning model integrating FCNNs and CRFs for brain tumor segmentation," Medical Image Analysis, vol. 43, pp. 98–111, Jan. 2018.

[7] Y. Guan, W. Li, T. Hu, and Q. Hou, "Design and Implementation of Safety Helmet Detection System Based on YOLOv5," in Proceedings - 2021 2nd Asia Conference on Computers and Communications, ACCC 2021, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 69–73. doi: 10.1109/ACCC54619.2021.00018.

[8] F. Paraijun et al., "Implementasi Algoritma Convolutional Neural Network Dalam Mengklasifikasi Kesegaran Buah Berdasarkan Citra Buah," vol. 11, no. 1, 2022, doi: 10.33322/kilat.v11i1.1458.

[8] P. Tejaswini, P. Singh, M. Ramchandani, Y. K. Rathore, and R. R. Janghel, "Rice Leaf Disease Classification Using Cnn," in IOP Conference Series: Earth and Environmental Science, Institute of Physics, 2022. doi: 10.1088/1755-1315/1032/1/012017.

[10] T.-R. Huang, S.-M. Hsu, and L.-C. Fu, "Data Augmentation via Face Morphing for Recognizing Intensities of Facial Emotions," IEEE Trans Affect Comput, vol. 14, no. 2, pp. 1228–1235, Apr. 2023, doi: 10.1109/TAFFC.2021.3096922.

[11] A. Mikolajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in 2018 International Interdisciplinary PhD Workshop (IIPhDW), IEEE, May 2018, pp. 117–122. doi: 10.1109/IIPHDW.2018.8388338.

[12] M. F. Naufal, "Analisis Perbandingan Algoritma SVM, KNN, dan CNN untuk Klasifikasi Citra Cuaca," Jurnal Teknologi Informasi dan Ilmu Komputer, vol. 8, no. 2, p. 311, 2021, doi: 10.25126/jtiik.2021824553.

[13] C. B. Gonçalves, J. R. Souza, and H. Fernandes, "CNN architecture optimization using bio-inspired algorithms for breast cancer detection in infrared images," Comput. Biol. Med., vol. 142, p. 105205, 2022.

[14] M. Ahmad et al., "Industry 4.0 technologies and their applications in fighting COVID-19 pandemic using deep learning techniques," Comput. Biol. Med., vol. 145, p. 105418, 2022.

[15] Y. Qin, "A Cancer Cell Image Classification Program : Based on CNN Model," in 2021 2nd International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT), IEEE, Oct. 2021, pp. 140–144. doi: 10.1109/AINIT54228.2021.00037.

[16] S. S. Dambal, M. K. Doddananjedevaru, and S. B. Gopalakrishna, "Premature Ventricular Contraction Classification Based on Spiral Search - Manta Ray Foraging and Bi-LSTM," Int. J. Intell. Eng. Syst., vol. 15, no. 6, pp. 1–10, 2022.

[17] K. Patel, V. Patel, V. Prajapati, D. Chauhan, A. Haji, and S. Degadwala, "Safety Helmet Detection Using YOLO V8," in 2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN), IEEE, Jun. 2023, pp. 22–26. doi: 10.1109/ICPCSN58827.2023.00012.

[18] C. Ma, S. Xu, X. Yi, L. Li, and C. Yu, "Research on Image Classification Method Based on DCNN," in 2020 International Conference on Computer Engineering and Application (ICCEA), IEEE, Mar. 2020, pp. 873–876. doi: 10.1109/ICCEA50009.2020.00192.

[19] H. Tang, "Image Classification based on CNN: Models and Modules," in 2022 International Conference on Big Data, Information and Computer Network (BDICN), IEEE, Jan. 2022, pp. 693–696. doi: 10.1109/BDICN55575.2022.00134.

[20] Y. Hao, "Convolutional Neural Networks for Image Classification," in 2021 2nd International Conference on Artificial Intelligence and Computer Engineering (ICAICE), IEEE, Nov. 2021, pp. 342–345. doi: 10.1109/ICAICE54393.2021.00073.

[21] A. Andreas, C. X. Mavromoustakis, H. Song, and J. M. Batalla, "CNN-Based Emotional Stress Classification using Smart Learning

Dataset," in 2022 IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics), IEEE, Aug. 2022, pp. 549–554. doi: 10.1109/iThings-GreenCom-CPSCom-SmartData-Cybermatics55523.2022.00107.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v13i1.2008, Copyright ©2024
Submitted : October 30, 2023, Revised : December 27, 2023, Accepted : January 11, 2024, Published : February 15, 2024