IPTEK, The Journal of Engineering, Vol. 10, No. 03, 2024 (eISSN: 2807-5064)

Rice Identification Using Convolutional Neural Network with YOLOv7 algorithm and VGG16

Assad Resi Alfurqan[a], Detak Yan Pratama[a], Andi Rahmadiansah[a], Dyah Sawitri[a], Lizda Johar Mawarani[a*]

[a]*Department of Engineering Physics, Insitut Telnologi Sepuluh Nopember, Surabaya 60111, Indonesia*

Corresponding author : lizda@its.ac.id

Abstract

Rice is the most widely consumed food worldwide. The many types of rice cause various difficulties in the process of classifying rice varieties. The process of manually classifying rice varieties that rely on human power has drawbacks including the subjectivity of assessment between observers, limited physical capabilities, and longer observation times. In this research a rice variety classification system has been developed using the Convolutional Neural Network with the YOLOv7 and VGG16 algorithms. The rice varieties classified are basmati, IR64, and rojolele varieties. The model with the YOLOv7 algorithm is trained for object segmentation of rice grains and is used to create rice grain image datasets. The model with the VGG16 algorithm was trained by transfer learning and used for classifying rice grain varieties. The model with a learning rate hyperparameter of 0,000061, the ReLU activation function, the number of neurons 256 in the second classification layer, with the fine-tuning training method, has the best performance with an accuracy value of 100%. The best VGG16 model weight is used in application implementation. Identification of the type of rice with the application can be done on the image of a batch of homogeneous and heterogeneous rice grains with various arrangements.

Keywords: CNN; Object identification; Rice type; VGG16; YOLOv7

1. Introduction

Based on the National Socio-Economic Survey in 2013, rice is the staple food for more than 95% of Indonesian people. Around 21 million households in agricultural areas are also recorded as having used rice farming as their main source of income. In addition, rice is the food most consumed by humans throughout the world [1]. The large number of types of rice on the market also creates various difficulties in the process of determining various types of rice on a large scale [2]. Meanwhile, the type of rice is also one of the considerations for consumers when purchasing rice. Consumer concerns about the originality of this type of rice have led to certification of rice originality by existing institutions to avoid individuals who have the intention of selling rice that does not match its type. Therefore, it is necessary to have an originality evaluation method to identify types of rice so that various parties, both producers and consumers, are not harmed [3].

This is the background for rice authentication carried out by several institutions as one of the reasons for grouping rice. Control is carried out by human vision by employing trained inspectors. However, determining the type of rice using this method can still be said to be less than optimal [4]. The weakness of manual identification is greatly influenced by various factors, including: (1) the existence of subjectivity factors in assessments between observers; (2) limited physical capabilities if human workers work too long, so that observation results are less accurate; (3) the time required for observation tends to be longer. In connection with this problem, a new solution is needed so that the process of identifying rice types can be carried out more effectively and efficiently.

Based on previous research, neural network technology has been used for various applications related to rice. These include classifying rice grains using wavelet decomposition [5], evaluating the quality of rice grains [6], identifying rice grains through image analysis and classification based on sparse representation [7], identifying diseases in rice [8], rice quality assessment using digital and optical image processing methods [9], rice grain identification based on color features [10], rice image evaluation based on shape and texture [11], analysis of Myanmar rice kernels using feature extraction [12], and analysis of Thai rice kernel quality [1], [13].

One of the neural network technologies is the convolutional neural network (CNN) which is widely used to classify images. CNN is inspired by the way the human brain processes information, and is designed to recognize patterns in data through a learning process [14]. CNN uses a process called convolution, where data is converted into a matrix of numbers and passed through several neural network layers [15]. Each CNN layer contains a set of neurons that process input data, and each neuron is connected to the next layer of neurons. The outputs of the neurons in each layer are then combined into the output of that layer, which is then passed to the next layer. This process is repeated until the output of the network is reached. Figure 1 shows a schematic of the process in CNN.

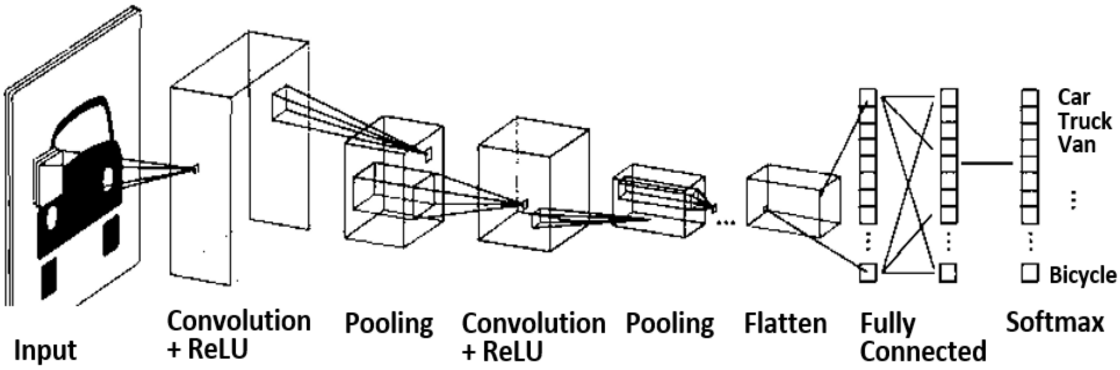

Figure 1. Scheme of CNN structure.

This research aims to design a system for identifying rice types using convolutional neural network technology using the YOLOv7 (You Only Look Once Version 7) and VGG16 (Visual Geometry Group 16 Layers) algorithms with the transfer learning method [16]. YOLO is one of the object detection algorithms in the computer vision domain for fast and accurate results. YOLO has many versions and the latest is version 7, namely YOLOv7. The YOLOv7 architecture is based on a single convolutional neural network (CNN) capable of processing the entire image in one forward-pass [17]. Object detection also uses IoU (intersection over union), a metric commonly used to evaluate the accuracy of object detection models [18]. This metric is a measure of overlap between the predicted bounding box and the ground truth box. IoU is calculated as the ratio of the intersection of the predicted bounding box and the ground truth box to the union of the predicted bounding box and the ground truth bounding box. Meanwhile, VGG16 is a CNN model for image classification developed by the Visual Geometry Group at the University of Oxford [19]. The VGG16 architecture is a deep neural network containing 16 layers, including 13 convolutional layers and 3 fully-connected layers. The architecture is characterized by the use of small convolutional filters (3x3) and a very deep architecture with a large number of filters. VGG16 uses a sequence of convolutional layers, pooling, and fully-connected to extract features from input images and make predictions.

The combined application of these two technologies is aimed at improving the rice identification process which is still done manually. YOLOv7 is used because of its efficient and fast detection model and VGG16 attracts with its ease of implementation. With this design, it is hoped that the results will be more optimal than previous research and can help rice producers and the public as rice consumers to be able to recognize types of rice more effectively and efficiently.

2. Method

2.1. Rice Image Capture for Datasets

The image dataset of a batch of rice grains is obtained by taking images of rice. The image samples for IR64, rojolele, and basmati rice each consist of 100 images of a batch of rice grains with a uniform resolution of 2992 x 2992 pixels, each of which contains random rice images of around 50 grains. The image of the rice batch was captured using a smartphone camera placed in a studio box set-up measuring 50 cm x 50 cm with a black background. As seen in the Figure 2, the studio box is also equipped with LED lights as a lighting component so that images can be captured with optimal quality.

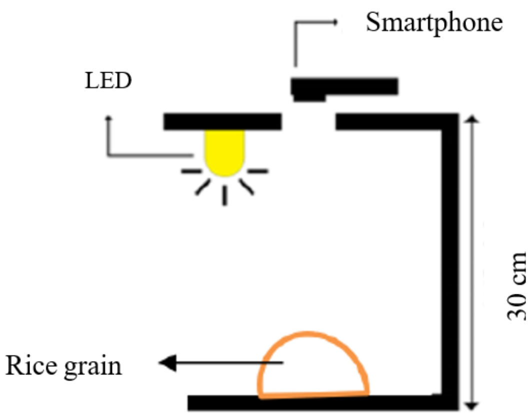

Figure 2. Illustration of the set up of equipment for taking image of rice.

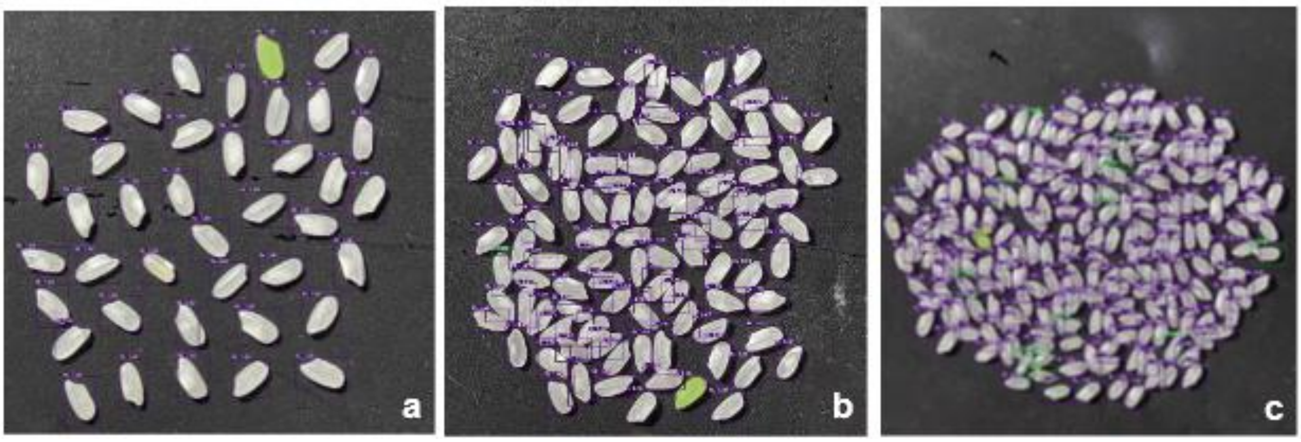

Figure 3. Arrangements of 50 rojolele rice grains (a) sparse; (b) dense; (c) very dense.

The batchs of rice grains were imaged in variety arrangement density (Figure 3) that is sparse, dense and very dense. An image is said to be sparse if there are 50 grains of rice in the background area of the rice image of 5 cm x 5 cm. An image is said to be dense if there are 100 grains of rice in the background area of the rice image of 5 cm x 5 cm. An image is said to be very dense if there are 200 grains of rice or more in the background area of the rice image of 5 cm x 5 cm. The variations were performed for each of the three types of rice and a mixture of the three.

The density of the image of a batch of rice grains can be expressed as the total surface area of the rice grains compared to the background area of the image of the batch of rice grains. So the density of the rice image can be determined by calculating according to equation (1)

$$R = \frac{(n * A)}{C} \tag{1}$$

where R is the density of the image of the rice grains batch, n is the number of rice grains, and C is the background area of the image of the batch of rice grains. The surface area of the rice image is calculated using the capsule-shaped rice approach. The surface area of the rice image can be calculated using equation (2).

$$A = (P - L) * L + \pi * L^2/4 \tag{2}$$

where A is the area of the rice grain image, P is the length of the rice grain, and L is the width of the rice grain.

Table 1. Image density of the rice grains at varying position densities.

Rice grains position	**Basmati**	**IR64**	**Rojolele**
Sparse	0.346	0.315	0.251
Dense	0.691	0.630	0.503
Very dense	1.383	1.260	1.006

Based on the length and width of each type of rice, the density level can be expressed as shown in Table 1. Basmati rice which is the longest has the highest density level. The image of a batch of rice grains is considered sparse if its density level is up to 0.346. The image of a batch of rice grains is considered dense if its density level is greater than 0.346 but less than 0.691. The image of a batch of rice grains is considered very dense if its density level exceeds 0.691. For the purpose of creating the VGG16 dataset, images of a batch of rice grains arranged in a sparse position were used.

2.2. Dataset Creation with YOLOv7

After the images for each type were obtained, five images of a batch of rice grains for each type of rice were prepared to be used as the YOLOv7 training dataset. These images are then annotated using an application provided by the Roboflow site which provides dataset preparation services. Each grain of rice is manually annotated as a rice class with the available polygon tool. Each image of a batch of rice grains will produce an annotation file in YOLOv7 format which contains class information and polygon coordinates for each rice grain. Each file is arranged in training, validation, and testing groups with a ratio of 3:1:1 and uploaded to Google Drive for use in YOLOv7 training.

Training the YOLOv7 model for object segmentation is necessary because rice grains do not fall into the classes recognized by default by the YOLOv7 algorithm. The YOLOv7 algorithm is trained with 5 images per type of rice to produce a custom rice grain detection model. This approach is limited to segmenting rice grain objects. The classification of rice types has not been determined at this stage but will be carried out further at the type recognition stage with the VGG16 model. Next, the trained model will be used to crop the rice image so that an image of rice is produced per grain.

The object segmentation model with YOLOv7 that has been trained is used to recognize rice grains and perform cropping to produce images per grain of rice. This process produces a set of rice grain images as well as a distribution of confidence levels of the rice grains. A set of rice grain images with a confidence level above the threshold of 0.6 will be selected to produce a dataset for the classification of rice grain types at the next stage. By using the image of a batch of rice, 3 groups of rice types will be produced, each targeted at around 5000 grains of rice.

After the rice grain cropping process is carried out, approximately 5000 rice grains will be produced from each type of rice. Each type will be grouped into training/validation and testing datasets with a ratio of 80:20. The training/validation dataset will be used in the VGG16 model training process, while the testing dataset will be used to test and evaluate the performance of the VGG16 model training results.

2.3. Design a VGG16 Model for Classification

After the rice grain images were obtained from the previous stage, the VGG16 model was designed to classify the rice grain images according to their type. Before the VGG16 model was designed, pre-processing was carried out on the training, validation, and testing datasets. Further model variations were made based on the training method and the determination of its hyperparameters. The model with the best performance was then tested and used in the implementation of the application in the next stage. From the two models above, a model that can produce a higher accuracy value will be selected and developed. Hyperparameters are very important in deep learning algorithms because they describe the details of the training and have a direct impact on the output results of the model [20]. Therefore, the selection of hyperparameter values needs to be done to produce optimal accuracy values. Optimal search can be done with various optimization methods. The selected hyperparameter set will be used in the creation and training of the VGG16 model. The training methods tried were transfer learning without fine-tuning and with fine-tuning.

Next, the VGG16 model was trained using the Adam optimizer, a batch size of 64, and epochs of 50. Early stopping was applied to the model. Early stopping is used to stop training when the model no longer shows an increase in accuracy or a reduction in validation loss. Patience is also set to 10, which indicates the number of epochs to wait before stopping early if there is no further change in validation accuracy.

The VGG16 model is trained using grain images of three types of rice, each consisting of 4000 training and validation images. Then features are extracted from the rice images in the convolution layers of the VGG16 model. For training and evaluating the performance of the VGG16 model, the rice image dataset used in this study was randomly divided into training, validation, and testing datasets with a ratio of 68%, 12%, and 20% randomly. The training and validation datasets are only used during the model fitting process, while the testing dataset is used for the needs of evaluating the performance of model predictions on previously unknown samples. The number of images used for model design and testing for training, validation, and testing datasets is 3200, 800, and 1000, respectively.

In the training process, a number of epochs are run containing the training and validation processes. For each epoch, the training accuracy, validation accuracy, training loss, and validation loss will be known. The value of the validation loss will continue to be observed for its tendency. If it is known that there is no improvement in the validation loss within a certain patience, then the training process will be stopped before the entire epoch is complete. As a result of the training process is a weight file with the best training accuracy along with the values of training accuracy, validation accuracy, training loss, and validation loss [21].

After the VGG16 model with the highest accuracy value was obtained, the model was tested again with the testing dataset. From the testing process, model performance can be evaluated based on the visualized confusion matrix. The visualized data contains information on TP (True Positive) values, FP (False Positive) values, TN (True Negative) and FN (False Negative) values. By using the confusion matrix, accuracy values, precision values, recall values and F1-score values can also be derived. From these metrics, information on rice classification performance can be known [22].

2.4. Application Implementation

From the VGG16 model that has been trained, weights will be obtained that can be used by applications. The application will be implemented in the Python programming language to test an image of a random batch of rice. Rice grains in the tested image will be identified through a detection process with the YOLOv7 model and a classification process with the VGG16 model. The rice grains that have been identified are then counted automatically to determine their overall performance.

An image of a batch of rice grains is selected to detect its type classification. For every change in the image of a batch of rice grains, the program will re-select. The image of rice can be a classification of one type of rice or a mixture of types of rice.

The rice detection process begins with image detection of rice grains. This step uses YOLOv7 with rice grain object recognition capabilities. This process will produce an array of sets of rice grains. For each grain of rice detected, the process of recognizing the type of rice will continue in the next step. For each grain of rice detected, the type of rice grain is determined. Determination of type using the VGG16 model. As a result, the confidence level of each grain of rice tested will be known.

From the confidence level value for each type of rice grain, the highest confidence level was selected as the selected type. If the confidence level is less than 60% then it is considered unclassified. The confidence level value of the rice grains for each type, as well as estimates of the types, are then saved. If the detected rice grains are not finished, the process will return to the previous step.

Confidence level data per grain of rice stored is processed to produce minimum, average, maximum and standard deviation values. Apart from that, by setting the highest confidence level above 0.6 as the selected type of rice, the total number of rice grains for each type can be determined. By testing a number of homogeneous and heterogeneous images, overall application performance can be determined.

3. Results and Discussion

3.1. Results of YOLOv7 Segmentation Model Training

From training, a YOLOv7 model is obtained with a certain weight. For square-shaped recognition, the precision value obtained is 0.86688 and the recall value obtained is 0.93303. As for the mAP (mean average precision) 0.5 value obtained is 0.94509 and the mAP 0.5-0.95 value obtained is 0.58023. Based on training data, the optimal value of mAp 0.5 square shape, mean average precision with IoU more than 0.5, was obtained at the 224th epoch.

After the training process is carried out, weights are obtained which need to be validated with a testing dataset which comes from 3 images of rice batchs consisting of around 150 grains of rice each. By comparing the performance metric curves from validation training and testing results, it can be seen that the relative performance figures have small differences. The difference can be seen in the precision-recall curve where the testing curve has decreased. The mAP50 value in each class decreased by around 0.02 in the training and testing datasets. The precision value decreases when the recall value gets higher. This shows that the false positive value increases when the false negative decreases.

3.2. Results of the VGG16 Model for Classification

The VGG16 model creation method using transfer learning is carried out without fine tuning and with fine tuning. From Table 2, it can be seen that the VGG16 model learning with transfer learning and fine tuning produces better accuracy values so that it is used in the next stage.

Table 2. Set hyperparameters and results.

Parameter	With fine tuning	Training time (s)
Learning rate	0.001	0.001
Activation function	relu	relu
First classification layer neurons	1072	1072
Second classification layer neurons	4096	4096
Output layer	3	3
Layer with fine tune	0	2
Epoch	50	50
Batch size	64	64
Training accuracy	0.93	0.97
Validation accuracy	0.921	0.956

Comparing the training results of the VGG16 model with Transfer Learning without fine tuning and with fine tuning, transfer learning with fine tuning gave higher training accuracy (0.97) and validation accuracy (0.956) so it was chosen for the next process.

Then the VGG16 model was rebuilt with more optimal hyperparameters and fine tuning. At a learning rate of 0.000061, a relu activation function, and a number of neurons of 256, better performance was obtained. From the training process, it was found that in 22 epochs, the training accuracy was 0.982 and the validation accuracy was 0.973. From this model training, we obtained better training and validation accuracy values compared to previously trained models. The VGG16 model resulting from this training was then tested and used in the application.

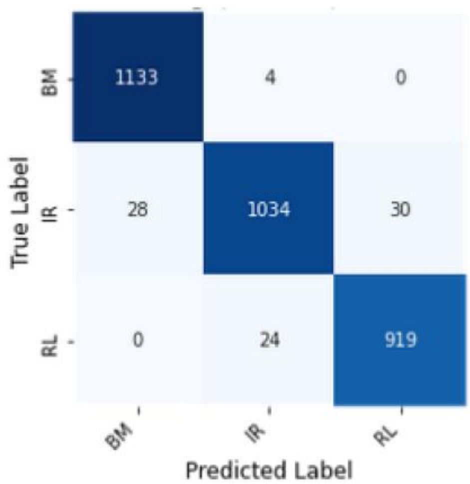

Figure 4. Confusing matrix model VGG16 with hyperparameter learning rate 0.000061, activation function ReLU, and number of neurons 256.

The performance of the VGG16 model for the overall classification needs of rice types is represented by the confusion matrix visualization shown in Figure 4. Based on the matrix, there were 1,133 basmati rice grains that were correctly classified as basmati rice, 28 basmati rice grains that were classified as IR64 rice grains, 4 IR64 rice grains that were classified as basmati rice grains, 1,034 IR64 rice grains that were successfully classified correctly, 30 IR64 rice grains that were classified as rojolele rice grains, 24 rojolele rice grains that were classified as IR64 rice grains, and 913 rojolele rice grains that were correctly classified as rojolele rice.

The values shown in Table 3 represent the model performance in classifying the three classes of rice types studied, namely basmati, IR64, and rojolele. Precision is the proportion of true positive predictions from all positive predictions made by the model. A high precision value indicates that the model is quite good at identifying true positives and does not produce false positive predictions. Recall is the proportion of true positive predictions from all actual positive events. A high recall value indicates that the model is quite good at identifying all positive events. F1-score is the harmonic mean of precision and recall. This value shows a fairly good balance between precision and recall.

Table 3. Performance metrics of the VGG16 model with learning rate hyperparameters 0.000061, activation function ReLU, and number of neurons 256.

Rice type	Precision	Recall	F1	Testing Accuracy
Basmati	97.59%	99.65%	98.61%	97.29%
IR64	97.36%	94.69%	96.01%	97.29%
Rojolele	96.84%	97.45%	97.15%	97.29%

Based on Table 3, the model has good performance in terms of precision, recall, and F1 score for the three classes of basmati, IR64, and rojolele. For the basmati type, the precision value is 97%, recall 99.65% and f1 98.61%. For the IR64 type, the precision value is 97.36%, recall 94.69% and f1 96.01%. For the rojolele type, the precision value is 96.84%, recall 97.45% and F1 97.15%. Based on these values, the tested model can be said to have good classification performance. High precision, recall, and F1 scores for each class indicate that this model is able to make accurate predictions and identify most true positives for each class.

3.3. Application Implementation Results

Application implementation is carried out by modifying the reference program. In the object cropping, object labeling, object area box marking sections, modifications are made so that the type of rice and its confidence level can be displayed on the image of the rice batch. The detection data is used to obtain a rice classification score which includes minimum, average, maximum values, standard deviation of the confidence level for each type of rice, as well as calculating the number of rice grains for each type.

Table 4. Performance metrics of rice identification with sparse position.

Rice Type	Precision (%)	Recall (%)	F1 (%)
Basmati	100	98.00	98.99
IR64	100	88.00	93.62
Rojolele	100	100	100

Table 5. Performance metrics of rice identification with dense position.

Rice Type	Precision (%)	Recall (%)	F1 (%)
Basmati	87.25	89.00	88.12
IR64	98.89	89.00	93.68
Rojolele	98.97	96.00	97.46

Table 6. Performance metrics of rice identification with very dense position.

Rice Type	Precision (%)	Recall (%)	F1 (%)
Basmati	100	80.50	89.20
IR64	98.78	81.00	89.01
Rojolele	98.91	91.00	94.79

The implementation provides the identification performance metrics results as shown in Table 4 for the sparse position, Table 5 for the dense position, and Table 6 for the very dense position. It can be seen from the tables that in the sparse position, identification is easy to do. All types of rice show 100% precision. The closer the rice grain position, the easier the identification is, which is seen from the lower performance metrics.

In general, the success of identification by this system is represented by the accuracy. Classification of rice types can be done well. Table 7 shows that identification of rojolele rice has the highest accuracy even at a very tight density position, still above 90% and reaching 100% at a sparse position. Identification of Basmati rice provides lower accuracy and IR64 is the lowest but has almost the same value as Basmati. While mixed rice can be detected with fairly good accuracy.

Table 7. Accuracy in identifying rice types according to the density of rice grain positions.

Position Density	**Accuracy (%)**			
	Basmati	**IR64**	**Rojolele**	**Mix**
Sparse	98.00	88.00	100	96.67
Dense	88.76	88.12	95.05	89.01
Very Dense	80.50	80.20	90.10	65.56

The density of the rice grain position greatly affects identification performance. The tighter, the lower the accuracy value and even for mixed rice at a very tight position, the accuracy decreases sharply to 65.56%. Images that are too dense make detection difficult due to the large number of overlapping bounding boxes of predicted rice grains. A different approach is needed to detect and crop images.

4. Conclusions

Based on the research analysis that has been carried out, it can be concluded that the YOLOv7 algorithm and the CNN model, namely VGG16 with transfer learning, can be used to classify types of rice grains. The CNN model with VGG16 and transfer learning that has been created can be used to classify and automatically calculate types of rice grains. The optimal hyperparameters of the designed model are learning rate: 0.000061, activation function, number of neurons 256. Rice identification was successfully carried out with a high accuracy value of up to 100% on Rojolele rice with a sparse position density. The denser the rice grain position, the accuracy decreased to the lowest value of 65.56% in the identification of mixed rice types.

References

[1] A. Gudipalli, A. N. Prabha, And P. Reddy Ch, “A Review On Analysis And Grading Of Rice Using Image Processing,” *Arpn Journal Of Engineering And Applied Sciences*, Vol. 11, No. 23, Pp. 13550–13555, 2016.

[2] M. Śliwińska-Bartel, D. T. Burns, And C. Elliott, “Rice Fraud A Global Problem: A Review Of Analytical Tools To Detect Species, Country Of Origin And Adulterations,” Oct. 01, 2021.

[3] M. C. Custodio, R. P. Cuevas, J. Ynion, A. G. Laborte, M. L. Velasco, And M. Demont, “Rice Quality: How Is It Defined By Consumers, Industry, Food Scientists, And Geneticists?,” Oct. 01, 2019.

[4] T. G. Devi, P. Neelamegam, And S. Sudha, “Machine Vision Based Quality Analysis Of Rice Grains,” *Ieee International Conference On Power, Control, Signals And Instrumentation Engineering (Icpcsi-2017)*, Pp. 1052–1055, 2018.

[5] K. R. Singh And S. Chaudhury, “Efficient Technique For Rice Grain Classification Using Back-Propagation Neural Network And Wavelet Decomposition,” *Iet Computer Vision*, Vol. 10, No. 8, Pp. 780–787, Dec. 2016.

[6] Y. Ogawa, “Quality Evaluation Of Rice,” Elsevier Inc., Apr. 2016.

[7] T. Y. Kuo, C. L. Chung, S. Y. Chen, H. A. Lin, And Y. F. Kuo, “Identifying Rice Grains Using Image Analysis And Sparse-Representation-Based Classification,” *Comput Electron Agric*, Vol. 127, Pp. 716–725, Sep. 2016.

[8] Y. Lu, S. Yi, N. Zeng, Y. Liu, And Y. Zhang, “Identification Of Rice Diseases Using Deep Convolutional Neural Networks,” *Neurocomputing*, Vol. 267, Pp. 378–384, Dec. 2017.

[9] Z. Parveen, M. A. Alam, And H. Shakir, "Assessment Of Quality Of Rice Grain Using Optical And Image Processing Technique," *2017 International Conference On Communication, Computing And Digital Systems (C-Code)*, Pp. 265–270, 2017.

[10] A. A. Aznan, R. Ruslan, I. H. Rukunudin, F. A. Azizan, And A. Y. Hashim, "Rice Seed Varieties Identification Based On Extracted Colour Features Using Image Processing And Artificial Neural Network (Ann)," *Int J Adv Sci Eng Inf Technol*, Vol. 7, No. 6, Pp. 2220–2225, 2017.

[11] B. Lurstwut And C. Pornpanomchai, "Image Analysis Based On Color, Shape And Texture For Rice Seed (Oryza Sativa L.) Germination Evaluation," *Agriculture And Natural Resources*, Vol. 51, No. 5, Pp. 383–389, Oct. 2017.

[12] T. N. Wah, P. E. San, And T. Hlaing, "Analysis On Feature Extraction And Classification Of Rice Kernels For Myanmar Rice Using Image Processing Techniques," *International Journal Of Scientific And Research Publications (Ijsrp)*, Vol. 8, No. 8, Aug. 2018.

[13] P. Kongsawat, S. Chivapreecha, And T. Sato, "Quality Assessment Of Thai Rice Kernels Using Low Cost Digital Image Processing System," *2018 International Workshop On Advanced Image Technology, Iwait*, Pp. 1–4, 2018.

[14] A. Patil And M. Rane, "Convolutional Neural Networks: An Overview And Its Applications In Pattern Recognition', Smart Innovation, Systems And Technologies," Pp. 21–30, 2021.

[15] Á. Zarándy, C. Rekeczky, P. Szolgay, And L. O. Chua, "Overview Of Cnn Research: 25 Years History And The Current Trends," *Proceedings -Ieee International Symposium On Circuits And Systems*, Pp. 401–404, 2015.

[16] F. Zhuang *Et Al.*, "A Comprehensive Survey On Transfer Learning," *Proceedings Of The Ieee*, Vol. 109, No. 1, Pp. 43–76, Jan. 2021.

[17] C.-Y. Wang, A. Bochkovskiy, And H.-Y. M. Liao, "Yolov7: Trainable Bag-Of-Freebies Sets New State-Of-The-Art For Real-Time Object Detectors," 2022.

[18] M. A. Rahman And Y. Wang, "Optimizing Intersection-Over-Union In Deep Neural Networks For Image Segmentation," *Springer International Publishing Ag*, Pp. 234–244, 2016.

[19] K. Simonyan And A. Zisserman, "Very Deep Convolutional Networks For Large-Scale Image Recognition," *3rd International Conference On Learning Representations, Iclr 2015 - Conference Track Proceedings*, Pp. 1–14, Sep. 2014.

[20] L. Yang And A. Shami, "On Hyperparameter Optimization Of Machine Learning Algorithms: Theory And Practice," *Neurocomputing*, Vol. 415, Pp. 295–316, Nov. 2020.

[21] Q. Wang, Y. Ma, K. Zhao, And Y. Tian, "A Comprehensive Survey Of Loss Functions In Machine Learning," *Annals Of Data Science*, Vol. 9, No. 2, Pp. 187–212, Apr. 2022.

[22] D. M. W. Powers And Ailab, "Evaluation: From Precision, Recall And F-Measure To Roc, Informedness, Markedness & Correlation," 2020.