

Research Article

Rapid Detection of Hybrid Maize Parental Lines Using Stacking Ensemble Machine Learning

M. Aqil ¹, **M. Azrai** ¹, **M. J. Mejaya** ², **N. A. Subekti**³, **F. Tabri**¹, **N. N. Andayani**¹, **Rahma Wati**¹, **S. Panikkai**¹, **S. Suwardi** ¹, **Z. Bunyamin**¹, **E. Roy**¹, **M. Muslimin**¹, **M. Yasin**¹, and **E. Prakasa** ⁴

¹Indonesian Cereals Research Institute, Maros, Indonesia

²Indonesian Legumes and Tuber Crops Research Institute, Malang, Indonesia

³Indonesian Center for Food Crops Research and Development, Bogor, Indonesia

⁴National Research and Innovation Agency, Bandung, Indonesia

Correspondence should be addressed to M. Aqil; acchmad1973@gmail.com

Received 7 January 2022; Revised 25 March 2022; Accepted 4 April 2022; Published 26 April 2022

Academic Editor: Soumya Ranjan Nayak

Copyright © 2022 M. Aqil et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Hybrid maize seed production is a relatively complex task due to the coexistence of three distinct types of maize plants in the field: female, male, and contaminant/off-type plants. Female and contaminant/off-type plants' tassels should be removed immediately following flowering initiation, while male tassels should be retained to allow cross-pollination between male and female plants. Therefore, development of an intelligent tassel classification system is deemed critical for hybrid purity decision-making. The research's primary contribution is the integration of two widely used transfer learning architectures, Inception V3 and SqueezeNet, with stacking ensemble machine learning using four algorithms (logistic regression, support vector machine, random forest, and k-nearest neighbors) for rapid classification of tassel images. Tenfold cross-validation was used to evaluate the model performance. Cloud computing was also investigated using EfficientNet to compare the predictive performance of the models. The models' performance was assessed using four metrics: accuracy, AUC, precision, and recall. The results depicted an appropriate developed model that properly distinguished male, female, and contaminant plants. The integration of the model with machine learnings (logistic regression, SVM, random forest, and KNNs) enables rapid recognition of off-type plants even though it is operated by personnel with limited skills of seed technology on ideotype recognition. Among all the evaluated CNN architecture and stacking models, Inception V3-embedded images with logistic regression metaclassifier outperformed other models with accuracy of about 98%. SqueezeNet and EfficientNet provided comparable results for consistent tassel classification with slightly lower performance measures. The model was also subjected to a multidimensional scaling (MDS) analysis to investigate and comprehend misclassification. Male and female plants are clearly distinguished by MDS, but female and off-type/contamination plants are ambiguous. This indicates that the prediction errors were caused by highly similar data features among female and off-type images. The developed modern plant phenotyping model can be used to assist breeders/technicians in maintaining the quality of large-scale hybrid maize seed production activities in Indonesia.

1. Introduction

The Government of Indonesia has recently launched a national maize self-sufficiency acceleration program through various approaches, including high-yielding maize hybrid seeds, to increase domestic production. For instance, the Ministry of Agriculture supplied certified national superior maize hybrid variety seeds nationwide, although the implementation success rate remained constant at around

60% [1]. One of the problems faced in hybrid maize seeds is maintaining purity and quality from the parent seeds to produce hybrid seeds (i.e., first-generation seeds). Maize is a cross-pollinating crop that is easily contaminated by pollens from surrounding plants, later reducing seed purity and yield potential.

Furthermore, during the implementation of this national program, the national research agency called the Indonesian Agency for Agricultural Research and Development

(IAARD) was selected to execute this self-sufficiency acceleration program in collaboration with more than 30 private licensed seed producers. Hybrid maize seeds are formed from a cross between two desired maize parent lines, namely, the male and female lines. Generally, in the field, one row of male plants is sandwiched between four rows of female plants (1:4). However, during plant growth, contaminant plants usually emerged due to some constraints such as volunteer plants and adulteration of maize seeds during preparation [2]. When entering the flowering phase, the female parents' tassel should be removed to allow pollen from the male plants to fertilize the female lines, leading to hybrid seed generation. Besides, field inspections must be carried out carefully and regularly to ensure that all female tassels or off-type plants have been properly removed. This activity is usually done manually by a specialist or group of specialists with different knowledge levels on plant breeding. An additional problem faced in the production of hybrid seeds was in the large and scattered areas, making monitoring and execution of activities at critical stages, such as eliminating tassel and off-type plants, quite difficult.

In the last decade, artificial intelligence technology has experienced significant scale-up developments, including its application in field/on-site research. Machine learning technology has experienced improvements, especially in the accuracy of recognizing objects. The introduction of a new technique, which is nondestructive based on CNN, has raised the expectation for image recognition technology's acceleration [3]. CNN is a deep learning method from the development of Multilayer Perceptron (MLP) designed for two-dimensional processing data. CNN is included in the type of Deep Neural Network (DNN) because it comprises a deep level of recognition applied to an object. CNN is generally applied for image data processing, including the development of self-driving cars, distance recognition, disease diagnosis, product marketing, and others.

Currently, modern plant phenotyping of maize using machine learning and deep learning can differentiate tassels between male and female parents as well as other contaminating plants. No off-type plants are permitted to remain in the field during the two parents' flowering. Tassels exhibit a variety of morphological characteristics, including curvature and cohesiveness, and cohesiveness. The challenges faced in tassel detection include tassels that appear unsimultaneously, different in sizes and shapes of tassels, and other lines producing differences in colour and texture of the tassels, changes in light illumination during capturing, and the background of the shooting. Thus, the tassel classification should be conducted with care after the detection [4]. Image capture of plant disease for automatic detection of maize Northern Leaf Blight disease and tassel type were also developed with the use of a convolution neural network and integrated into smartphone application [5, 6]. Additionally, the estimated time and the number of tassel emergence in maize were managed automatically with the neural network regression approach. Furthermore, [7] reported deep learning based on the recurrent neural network through preselection data to estimate maize yields in the USA. The utilized constituents included weather data, soil quality, and

soil moisture collected at an interval of one hour. These data were processed using machine learning which predicted the spatial distribution of maize yields with an accuracy of above 80%. Machine vision-based technology was used to separate seeds based on parameters such as colour, texture, size, and shape characteristics based on photo images which were later applied to seed handling and seed sorting activities [8]. Other researches on object detection were introduced by involving hyperparameter setting for kernel counting on maize ear by using the CNN supervised learning method [9]. The results of the parameters through training revealed that the model is appropriate for calculating the number of seeds on an ear automatically with a depicted correlation level of more than 0.90.

The integration of deep learning and machine learning into prediction and classification tasks has attracted increasing attention in recent years. By combining diverse base machine learning estimators of various types, it is possible to create a more performant model, referred to as an ensemble model [10]. Ensemble methods outperform more established machine learning techniques referred to as base learners. Stacking is an ensemble approach that is frequently used to improve prediction accuracy by combining two or more independent machine learning algorithms into a metamodel [11]. Once it relates to prediction, stacking ensemble machine algorithms outperform single algorithms. The stacking of ensembles technique has been applied to a variety of agricultural problems, including prediction, classification, regression, and feature selection.

To address quality control concerns in maize seed production, the study sought to develop efficient tassel recognition methods capable of rapidly capturing various types of maize tassels during generative development. The detection model is capable of accurately identifying a variety of tassel parameters, including height, width, branches, another colour, glume colour, and other morphological characteristics of a particular line.

2. Materials and Methods

2.1. Study Area. The study to classify maize hybrid parental lines was conducted in 2020 in two different seasons, namely, rainy and dry seasons, at four different locations: Gowa, Bone, Maros, and Pangkep in South Sulawesi Province, Republic of Indonesia. Two lines were planted, including male parent of Nasa 29 variety (i.e., G102612 line) and female parent (i.e., MAL03 line). Planting system adopted was 1:4 ratio where one row of male plants was sandwiched to four rows of female plants (Figure 1). All lines/genotypes were bred and generated by the Ministry of Agriculture, the Republic of Indonesia, in collaboration with the International Maize and Wheat Improvement Centre (CIMMYT). Substantial variability was observed for all agronomic characteristics at various growing environment conditions across the country. During plant growth, contamination/off-type plants began to appear, necessitating their immediate removal from the fields.

To prevent the genetic aspect from off-type, all two lines were planted separately in a 0.5 ha area and isolated distance



Mal 03 (Female)



Mal 03 (Female)



G102612 (Male)



G102612 (Male)



Various Off-type/contamination lines (Mr14, CLYN230, and CY7 lines)

FIGURE 1: Tassel/flower removal and types of parental lines existed in the field.

of approximately 300 m. Each genetic material was planted at a distance spacing of 70 cm × 25 cm with one plant per stand. Fertilizer was applied twice at 10 and 30 days after planting (DAP) with urea (46% N) and NPK (15% N, 15% P, and 15% K) fertilizers with doses of 300 and 400 kg/ha, respectively.

When entering the flowering phase, the female parents' tassel should be removed to allow pollen from the male plants to fertilize the female lines, leading to hybrid seed generation. Arbitrary tassel image datasets were obtained for

female parents, male parents, and contaminant plants. The contamination plants (neither male nor female plant) frequently existed as volunteer plants that should be quickly removed, such as Mr14, CLYN230, and CY7 lines. The dataset collected from the parental lines differed regarding the tassel types and was trained and validated with the deep learning CNN model. Images of the experimental sites were captured using an Inspire 2 drone equipped with a Zenmuse X5 Camera and 3-Axis Gimbal with 15 mmf/1.7 lenses, whereas the tassel image was captured using a high-

resolution CCD Digital Camera (Samsung). The captured maize tassel datasets were utilized for training and validation, as shown in Table 1. All tassel images were divided into 80% for model development using training datasets and the remaining 20% to validate tassel datasets. This method follows the majority of machine learning studies' recommendations that the validation set should be 10%–30% of the total dataset [12].

2.2. Model Development. The classification model for maize hybrid lines was developed using the Convolutional Neural Network (CNN) algorithm and was based on tassel-type features. CNN is the Multilayer Perceptron (MLP), which is designed to process two-dimensional data. Meanwhile, CNN is included in Deep Neural Network (DNN) analysis because of its high deep network that has been applied in several data image classification. Besides, CNN is a type of neural network commonly used in image data that can detect and recognize objects of an image. Generally, it does much from the usual neural network. Also, it consists of neurons that have weight, bias, and activation functions. However, the only difference can be seen in the architecture divided into two parts: Feature Extraction Layer (FEL) and Fully Connected Layer. FEL processed an image into features in the form of numbers, thus representing an image. Additionally, FEL consists of two parts, namely, the convolutional layer and the pooling layer.

Much of recent research on deep CNN has focused on increasing the accuracy of computer vision datasets, including improvement of network architecture. In this study, the classification of parental lines hybrid maize variety was developed by comparing several well-known CNN architectures, that is, Inception V3, SqueezeNet, and EfficientNet-B1. Inception V3 and SqueezeNet CNN analysis were done by using Orange visual programming. Additionally, EfficientNet-B1 execution was done in Google collaborator GPU deep learning service.

2.2.1. Ensemble Machine Learning. Transfer learning was combined with a deep learning model and several machine learning algorithms in this study to determine the type of parental lines (female line, male line, and off-type lines). Transfer learning was used to create the features, which were generated using the Inception V3 and SqueezeNet models. The models were built using logistic regression (LR), support vector machine (SVM), random forest (RF), and k-nearest neighbors (KNN) algorithms. By comparing the contribution of different models, the best model can be chosen as the base model for the stacking stage. The overall framework for the detection of maize parental lines images using ensemble machine learning is shown in Figure 2.

Image embedding was used for proceeding tassel classification on parental maize lines using Orange Software [13]. The embedding process will generate a vector representation of each image in a large number of features from a deep network. The Keras Python library provides an interface to the InceptionV3 and SqueezeNet tassel detection models. Inception V3 is among the CNN architectures that

Szegedy first introduced [14]. The study involved Inception V3 by considering that auxiliary classifiers did not contribute much during the training stage in the previous versions, and the possibility of fixing Inception V2 without drastically changing modules should be investigated further. As a solution, Inception V3 was involved in the study where all the improvements over Inception v2 were involved, such as the use of RMSProp optimizer, factorized 7×7 convolutions, batch norm in the auxiliary classifiers, and label smoothing (a type of regularizing component added to a loss formula which serves to prevent overfitting). Inception V3 represents an image as a set of 2048 features, which can then be further processed using supervised or unsupervised machine learning techniques. Figure 3 depicts the ensemble process using transfer learning based on the Inception V3 model.

Another model that was also assessed in the developed framework is SqueezeNet, which was done through transfer learning from the pretrained ImageNet datasets. SqueezeNet transforms raw images into their vector representation using a Deep Neural Network that was trained on millions of real-life prints (called image embedding). SqueezeNet is a deep model with microarchitecture to enable accurate image classification with fewer parameters [15]. It has a vital parameter, defined as a fire model that comprises a squeeze convolutional layer that has filters fitting into an expanded layer with a mix of convolutional field filters. It has fewer parameters and offers several advantages, such as a more efficient distributed training communication among servers and faster training time due to the small model architecture. SqueezeNet implemented three significant strategies for building efficient network architecture; that is, instead of 3×3 , SqueezeNet used 1×1 filters and expanded into convolution filters (called fire module). Also, a bypass system was employed to increase the filters per fire module. Max-pooling and global average pooling were involved before generating prediction/classification. Figure 4 depicts the ensemble process using transfer learning based on the SqueezeNet model. A fixed training/validation percentage of the sampling method was applied to assess the performance test and score of the tassel prediction by incorporating four popular machine learning methods: logistic regression, support vector machine, random forest, and k-nearest neighbors.

2.2.2. Machine Learning Algorithms. Machine learning has recently gained popularity in the field of pattern recognition, including image embedding, particularly in agricultural imaging. In this study, we examined a large number of features from a deep network via transfer learning of Inception V3 and SqueezeNet, as well as four popular machine learning models, that is, LR, RF, SVM, and KNN.

Logistic regression (LR) aimed to model the probability of an event occurring depending on the independent variables' values, in either numerical or categorical value. In machine learning, LR is used to classify data observations by estimating the probability that the observation is in a particular category. For example, the off-type plant's

TABLE 1: Maize tassel datasets used for the case study.

Trial field locations	Number	Lines	Training	Validation
Gowa_02	87	Female Mal 03	√	
Bone_02	274	Female Mal 03	√	
Pangkep_02	168	Female Mal 03		√
Gowa_01	102	Male-G102612	√	
Bone_01	267	Male-G102612	√	
Maros_01	160	Male-G102612	√	
Gowa_03	89	Others/contamination	√	
Bone_03	320	Others/contamination	√	
Pangkep_03	120	Others/contamination		√
Total	1587 images			

probabilities in the maize field can be predicted by various morphological information such as tassel shape, colour, glume, circularity, and silk colour. LR is also used to predict the effect of a series of variables on a binary response variable.

Support vector machine (SVM) is a type of machine learning to precisely classify different objects by drawing a decision boundary known as a hyperplane near the extreme points in the dataset. Essentially, SVM is a frontier that best segregates the two classes by creating a decision boundary that segregates the two classes. SVM can be used in multidimensional datasets, and the data points are referred to as vectors as they have coordinates within the space of data. In high-dimensional space, a function was used to transform the data from two-dimensional to three-dimensional features. A kernel trick was adopted to reduce the computational cost of a function that takes as inputs vectors in the original space and returns the vectors' dot product in the feature space (kernel function). Among the popular kernel types to transform the data into high-dimensional features are RBF, sigmoid, and polynomial. By selecting the appropriate kernel, high accuracy in classifying objects is permitted.

Random forests (RF) combine the simplicity of decision trees with various flexibility, resulting in a fast improvement in accuracy. RF makes a simple yet effective machine learning method for optimal classification or prediction. RF was created by creating a bootstrap dataset by randomly selecting the samples from the original dataset. Ensemble model through aggregations will improve the accuracy of RF and reduce the cost associated with storing and getting inferences from multiple models.

K-nearest neighbors (KNN) is a classification method for a particular dataset based on previously classified data. KNN is supervised learning, where the results of new query instances are classified based on the majority of the existing categories' proximity. Proximity can be thought of as the inverse of distance or inversely proportional to distance. The smaller the distance between two instances, the greater the "proximity" between the two cases. Thus, the k-nearest neighbors of an instance are defined as the k instance with the smallest distance/greatest proximity.

2.2.3. Cross-Validation Algorithms. K-fold cross-validation is a technique for determining the average success rate of a

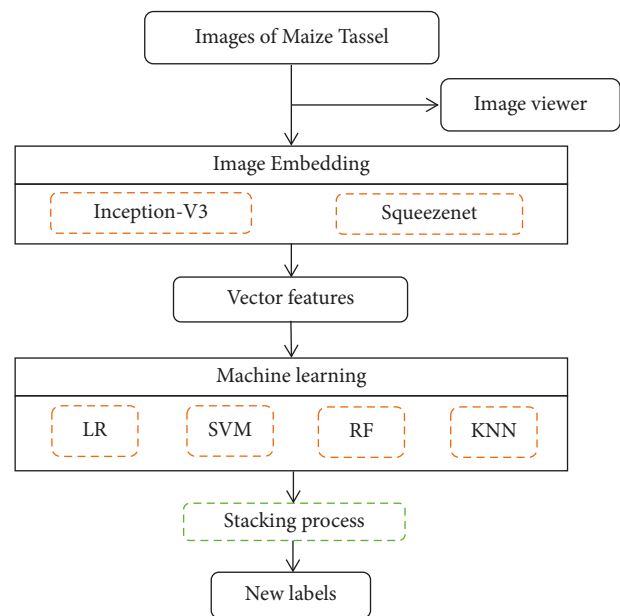


FIGURE 2: The overall framework for the detection of maize parental lines in this study.

system by performing redundancy on the input attributes and testing the system for several random input attributes. Cross-validation on a k -fold scale begins by dividing the desired number of n -fold datasets. The first fold occurs when the first part of the data is considered as validation data and the remaining as training data. Then, using that portion of the data, determine the accuracy, similarity, or proximity of a measurement result to the actual number or data. The second fold occurs when the second part of the data is treated as validation data, and the remaining data is treated as training data. Furthermore, accuracy is calculated by segmenting the data and continuing until it reaches the k -fold. Calculate the average precision of the k accuracy pieces mentioned previously. This average precision becomes the final precision. Figure 5 illustrates the overall process of k -folds.

2.2.4. Model Stacking. Stacking is a type of ensemble machine learning algorithm that utilizes metalearning algorithms to determine the optimal way to combine predictions from several base models. It is primarily used to train a new

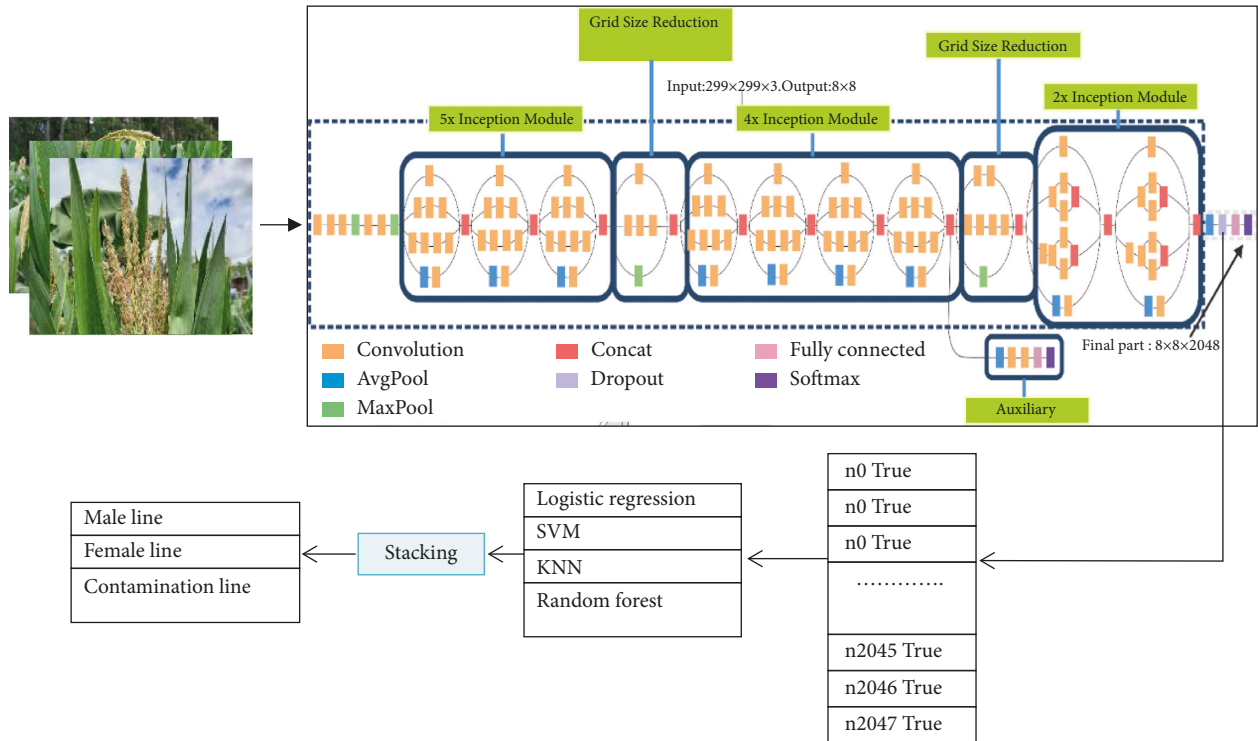


FIGURE 3: Proposed ensemble process using transfer learning based on the Inception V3 model.

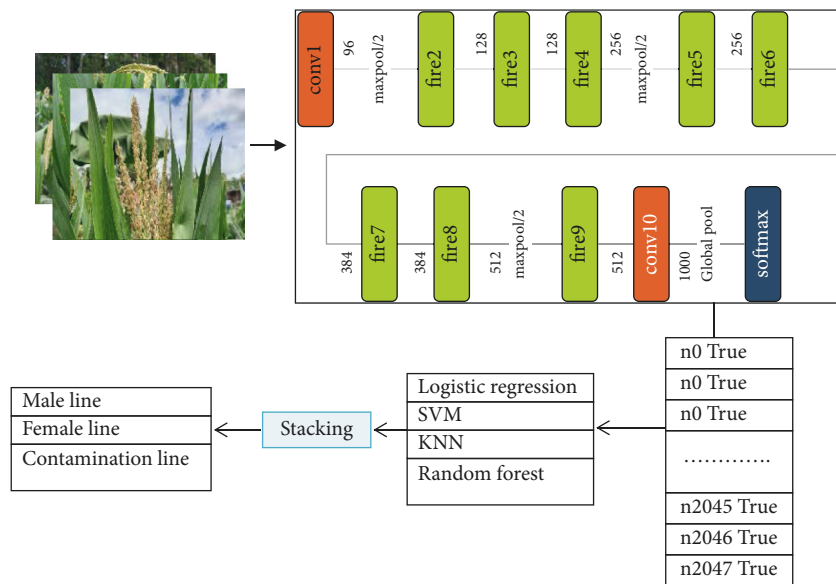


FIGURE 4: Proposed ensemble process using transfer learning based on the SqueezeNet model.

model using the output of multiple weak learners. Wolpert [16] introduced the stacking algorithm and generated a new direction in the field of combined models. In comparison to base models, the stacking model outperformed them. Sigletos et al. [17] conducted a thorough comparison of the bagging, boosting, and stacking algorithms and discovered that the stacking algorithm offers significant robustness benefits.

The model was developed over the sequence of a two-stage training procedure. The model stacking process is described as follows: (1) Using k-fold cross-validation, train the base model on the same datasets as the test model. (2) Using k-fold cross-validation, select the m base models that have significant performance for prediction. (3) Use the mean value of the base models' k-fold cross-validation as the new features. (4) Train a senior model using the new features

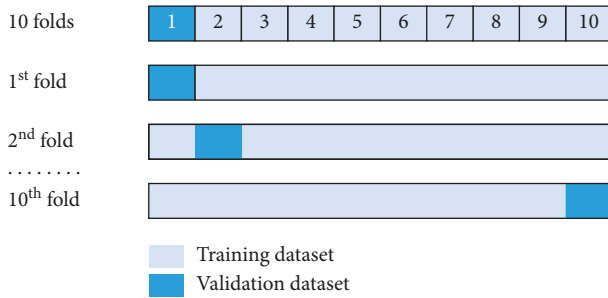


FIGURE 5: K-folds cross-validation.

selected in the previous step. While training with new features, the 5-fold cross-validation algorithm was used to ensure the robustness of the new features. Following that, LR was used as the metamodel to construct the model and make a final prediction in the subsequent stage. Figure 6 shows the overall flow of the model stacking process.

2.2.5. EfficientNet-B1. EfficientNet is a relatively new-scale neural architecture search to efficiently scale up CNN's size. The significant advantages of EfficientNet include higher accuracy, efficiency, and scaling of the CNN model [18]. This model is a compact model and may be combined with pretrained models for the purpose of determining the optimal scaling parameters for the computational process. This model also integrated with the TensorFlow Keras application ecosystem for executing the EfficientNet model. EfficientNet comprises five active modules that will be further combined to build a subblock in a skip connection. EfficientNet consisted of eight versions from EfficientNet-B0 to B8 (Figure 7). Model architecture is almost the same except for the number of feature maps used in the model version.

2.3. Accuracy Assessment. Four popular measures were employed to determine the robustness of the models, that is, classification accuracy (CA), precision, recall, and AUC (equations (1) to (3) and AUC) as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}} \times 100\%, \quad (1)$$

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

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

True Negative (TN) values are data that are correctly classified as negative or false outputs. True Positive (TP) is data that is properly classified as positive or true output. False Positive (FP) is data that is classified incorrectly if the output is positive or true. False Negative (FN) is data that is classified incorrectly. Also, a confusion matrix was used for summarizing the performance of the tassel classification model.

The AUC is used to measure performance and separate them among different classification thresholds. Meanwhile,

the AUC probability ranked classification model randomly with the positive value ranked higher than the negative. The values of AUC range between 0 and 1. For instance, a model that is predicted as 100% wrong has an AUC value of 0.0, whereas the prediction of 100% correct AUC value is 1.0. AUC was preferably used for this research because it determines how well predictions are ranked rather than using their absolute values.

3. Results and Discussion

3.1. Agronomic Characteristics of Parental Lines. Maize seeds are produced by crossing male and female parental lines of the desired characteristics of the varieties. For instance, this study used the male and female parents of the commercial maize variety (i.e., Nasa 29) with specific tassel features. The field was planted with four rows of female plants and one row of male plants (1 : 4 proportion of male: female). During the onset of the flowering phase, the tassel of the female plants (Mal03) was removed to prevent cross-pollination from another source through dispersal agents. The pollen dust from the male plants (G102612) was used to pollinate the female, producing the first generation called hybrid. Field researchers carried out strict surveillance to ensure the female tassel was carefully, regularly, and completely removed. This activity was done manually and involved many field researchers and technicians with different expertise knowledge within five to ten days per hectare.

Proper detection of the type of maize tassel will significantly affect the purity of the seeds being produced. Prior to seed marketing in Indonesia, the national standard requires 98 percent purity and a 2% germination rate. Additionally, off-type/contamination tassel removal must consider several factors such as plant agronomic and environmental parameters. The male/G102612 line was developed from the recombination of 5 drought-tolerant lines introduced by CIMMYT in 2005. Additionally, the female/Mal03 line was developed from a base population that was resistant to downy mildew disease as part of the Asian Biotechnology Maize Network (AMBIONET) project, which was a collaboration between the CIMMYT and IAARD. The agronomic characters of the two parental lines and other contaminant lines are shown in Table 2.

The selection of appropriate lines should consider the genetic background and combine ability, either general combining ability or specific combining ability of the crossed lines. In general, the ideal parental pair for a hybrid has a good specific combining ability value, taking into account the tested lines of agronomic characters. Morphologically, the female/Mal03 parent was characterized by tassels with an upright and compact type with a length ranging from 230 to 260 mm and the number of branches ranging from 12 to 13 branches with spikelets in each branch ranging from 23 to 73 spikelets per plant. The compact and erect tassel character and semierect leaf type are expected to minimize obstacles in male plants' pollination so that the pollination occurs smoothly and more responsive toward high yield [19]. The choice of Mal03 as the female parent was due to the excellent

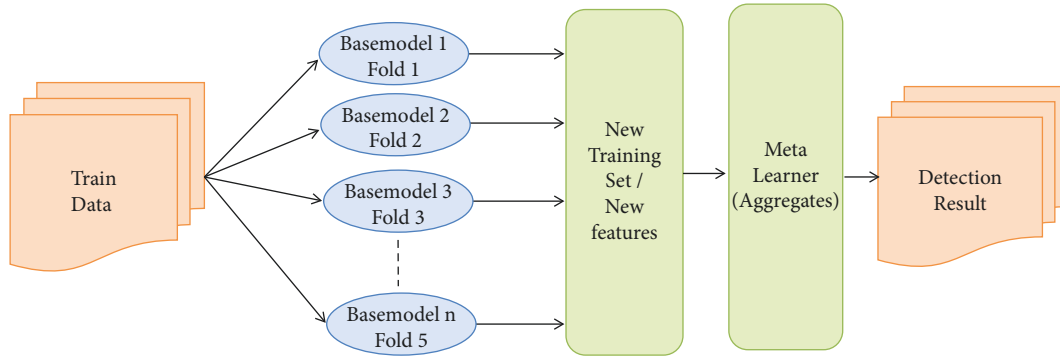


FIGURE 6: Model stacking (LR aggregating input).

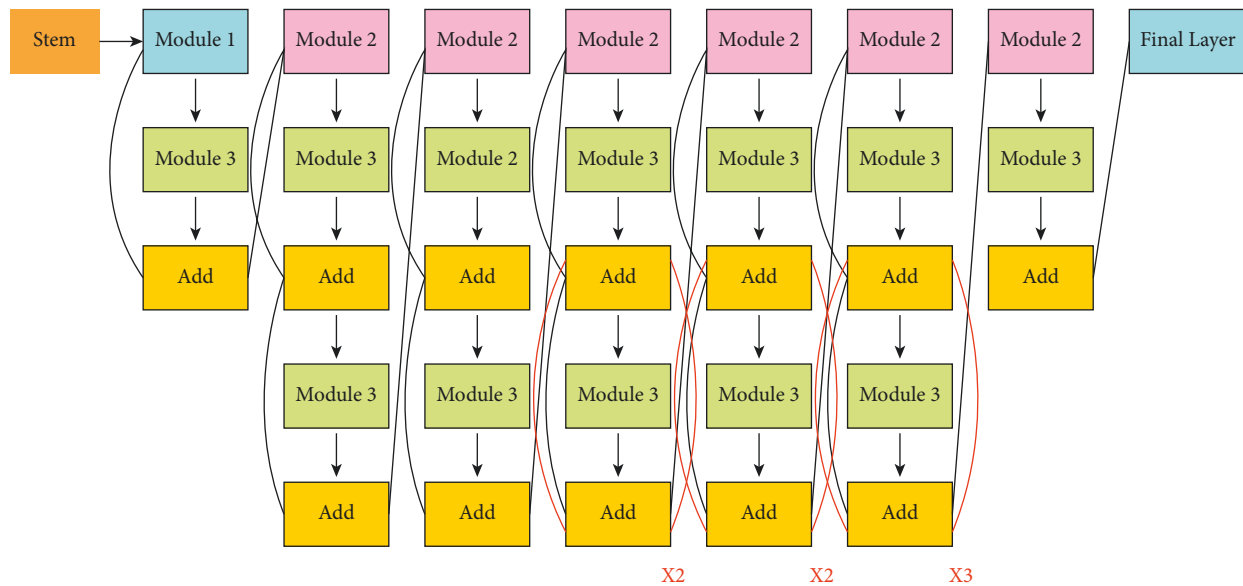


FIGURE 7: The model architecture of EfficientNet-B1.

cob appearance, enabling a good grain filling and increasing seed yield.

The male/G102612 parent has different morphological features compared to the female parent. The choice of G102612 as the male parent was mainly due to the longer shape ranging from 320 to 400 mm, with a higher number of branches and spikelets surrounding the tassel. Furthermore, the good tassel performance enables the production of huge amounts of pollen applied to the female plants through cross-pollination. Another merit of the male line was the compact tassel shape and its ability to slowly release pollens to enhance female plants' pollination. As for male performance, our previous research indicated that the G102612 line could release pollen up to 10 days after the initiation of the maize hybrid tassel.

3.2. Metamodel Comparison for Accurate Classifications. Three popular deep learning models were trained and validated in two cloud-based software services, namely, GPU-based Google collab and Orange Software. Notwithstanding, the model framework relies on detailed tassel morphological

information input, which is crucial for generating accurate classification. Two popular CNN models were also examined for our study as comparisons, that is, Inception model Version 3 and SqueezeNet using visual programming Orange 3.1 environment. InceptionV3 is pretrained on ImageNet datasets, and the embedding process makes use of activation from the model's penultimate layer, which represents the image as a vector. Transfer learning was involved by generating a vector representation of each image in a huge number of features from the Inception model. SqueezeNet is built on ImageNet using the pretrained model weights. The embedding is pre-SoftMax (flatten10) layer activation. As many as 2048 vector values (defined as N0 until N2047) were generated to allow visual assignment of tassels according to their classes (male, female, or off-type/contamination). The model has a total of 3,251,763 parameters that were further used to train and validate the tassel classification model.

To build models for predicting parental lines, we used traditional machine learning and ensemble methods (stacking model) in combination with pretrained Inception V3 and SqueezeNet. Four popular machine learning techniques were used in this study: logistic regression, support

TABLE 2: Description of male and female parental lines used in the study.

Variable	Male parent/G102612	Female parent/Mal03
Pedigree	Extracted from drought-tolerant CIMMYT inbred lines	Extracted from downy mildew tolerant lines
Days to tasseling (days after planting)	55	50
Tassel weight (g)	35.40–43.31	20.3–23.5
Tassel area (mm ²)	4105–13,000	3653–5664
Perimeter (mm)	210–1,752	124–688
Tassel length (mm)	320–400	230–260
Spike length (mm)	243–310	135–162
Number of branches	13–15	12–13
Number of spikelets	41–74	27–73
Tassel compactness	Compact	Compact
Tassel shape	Drooping	Erect
Glume colour	Green	Green
Another colour	Green-creamy	Greyed-red

vector machine, random forest, and k-nearest neighbor. Similar to the EfficientNet datasets, 80% of datasets were used for training the model and the remaining 20% for validation. The tuning of hyperparameters will be included in order to obtain optimal prediction, as shown in Table 3. Cross-validation was applied to derive predictions from training data and to train higher-level machine learning models. Tenfold cross-validation was adopted to evaluate all nine unstacking examined models.

The EfficientNet B1 used in this study is part of the TensorFlow and/or Keras applications ecosystem of pre-trained models to adapt the transfer learning paradigm and to design and implement the prediction framework by considering the parameters optimization. TensorFlow is a popular software library used by engineers to study deep computation [24]. Integrating TensorFlow with Keras can execute the model faster in an efficient model-building approach.

The input shape was set to 240×240 -pixel image resolution, followed by advanced data augmentation schemes such as flip, rotation, zoom, and shear, which were applied prior to model training. EfficientNet B1 utilizes a stochastic depth regularization to drop out neurons and the entire path of a network. The hyperparameter model used in the simulation included an optimal learning rate of 0.001; the number of epochs was 50 with three classification classes: Mal03, G102612, and contamination plant. The loss function used was the neural network cross-entropy loss which was optimized using the Adam optimizer. A SoftMax activation dropout and average pooling were adopted to smoothen out the last output dimension from the EfficientNet with noisy-student weight. The noisy student is a semisupervised learning technique that uses a larger or equal-sized student model and adds noise to the student during training [25]. The model has 7,890,051 parameters, with a composition of 7,828,003 trainable parameters and 62,048 nontrainable parameters. A summary of the model parameter of blocks 1A and 7b is shown in Table 4. These long blocks will have a varying number of subblocks with the increase of the model version.

The output was presented as model loss and accuracy, respectively, in Figure 8, which were later used to assess the impact of parameter setting on the model performance. The model was first tested against the training datasets to check the accuracy of tassel classification. The model gave a good precision in matching features with an approximation of 99.8% of the tassels correctly classified. Simultaneously, the model validation indicated that the magnitude of the model's response to the input change might cause a less significant reduction in the model accuracy of 94.9%. Adjustment of the dropout layer parameter to an optimum value of 0.5 provided a maximum increase in model accuracy for distinguishing features.

Meanwhile, the loss value of the model during training and validation processes remains stagnant at 0.15. A few errors in classification may be attributed to the model's difficulties to quickly identify tassels with similar morphological shapes, particularly female and off-type/contaminant categories. Most tassels were vulnerable to contamination from the surrounding maize plant due to exposure in the field.

Four indices, accuracy, AUC, precision, and recall, were used to assess the model performances. Inception V3 (LR with ridge regularization type/Model 4) and SVM with radial basis function kernel/RBF (Model 2) showed a performance rate of 0.975 or 97.50% and 0.967 or 96.70% correctly classified on test data. KNN and RF performed worse than LR and SVM, although the accuracy level was still under acceptable level, that is, 95.30% and 91.50%, respectively. Model improvement was also assessed by incorporating a newer CNN architecture, SqueezeNet, as a light image embedding widget to generate 1000 vector values (defined as N0 until N999) assigned to the tassel class. The model has a total of 1,587,000 parameters, the lowest parameter numbers as compared to the EfficientNet and Inception V3. SVM and LR SqueezeNet (Models 6 and 8, resp.) perform better at predicting tassels than KNN and RF models (Table 5).

Two additional models were evaluated for accuracy using ensemble methods: Inception V3 (LR aggregate) and SqueezeNet (SVM aggregate). Since LR, SVM, and KNN

TABLE 3: Setting of ML model parameters.

No	Method/model	Hyperparameter settings	References
1	Logistic regression	Regularization = ridge (L2); cost strength (C) = 1	[20]
2	Support vector machine	Kernel = radial basis function (RBF), tolerance = 0.001, degree = 3, cost = 1.00, and iteration = 100	[21]
3	Random forest	N_estimator = 70, criterion = entropy, and limit depth = 3	[22]
4	KNN	Neighbors = 5, weight = uniform (all neighbors equally weighted), and metric = Euclidean	
5	Ensemble model (stacking)	Final estimator = logistic regression; classifiers = random forest, support vector machine, and KNN	[23]

TABLE 4: Model summary of EfficientNet B1.

Layer (type)	Output shape	Param #	Connected to
input (InputLayer)	(None, 240, 240, 3)	0	
stem_conv (Conv2D)	(None, 120, 120, 32)	864	input_2[0][0]
stem_bn (BatchNormalization)	(None, 120, 120, 32)	128	stem_conv[0][0]
stem_activation (Activation)	(None, 120, 120, 32)	0	stem_bn[0][0]
block1a_dwconv (DepthwiseConv2D)	(None, 120, 120, 32)	288	stem_activation[0][0]
block1a_bn (BatchNormalization)	(None, 120, 120, 32)	128	block1a_dwconv[0][0]
block1a_activation (Activation)	(None, 120, 120, 32)	0	block1a_bn[0][0]
block1a_se_squeeze (GlobalAvera)	(None, 32)	0	block1a_activation[0][0]
block1a_se_reshape (Reshape)	(None, 1, 1, 32)	0	block1a_se_squeeze[0][0]
block1a_se_reduce (Conv2D)	(None, 1, 1, 8)	264	block1a_se_reshape[0][0]
block1a_se_expand (Conv2D)	(None, 1, 1, 32)	288	block1a_se_reduce[0][0]
block1a_se_excite (Multiply)	(None, 120, 120, 32)	0	block1a_activation[0][0] block1a_se_expand[0][0]
block1a_project_conv (Conv2D)	(None, 120, 120, 16)	512	block1a_se_excite[0][0]
block1a_project_bn (BatchNormal)	(None, 120, 120, 16)	64	block1a_project_conv[0][0]
<i>Continue to block 1b until block 7a (script not shown)</i>			
block7b_expand_conv (Conv2D)	(None, 8, 8, 1920)	614400	block7a_project_bn[0][0]
block7b_expand_bn (BatchNormali)	(None, 8, 8, 1920)	7680	block7b_expand_conv[0][0]
block7b_expand_activation (Acti)	(None, 8, 8, 1920)	0	block7b_expand_bn[0][0]
block7b_dwconv (DepthwiseConv2D)	(None, 8, 8, 1920)	17280	block7b_expand_activation[0][0]
block7b_bn (BatchNormalization)	(None, 8, 8, 1920)	7680	block7b_dwconv[0][0]
block7b_activation (Activation)	(None, 8, 8, 1920)	0	block7b_bn[0][0]
block7b_se_squeeze (GlobalAvera)	(None, 1920)	0	block7b_activation[0][0]
block7b_se_reshape (Reshape)	(None, 1, 1, 1920)	0	block7b_se_squeeze[0][0]
block7b_se_reduce (Conv2D)	(None, 1, 1, 80)	153680	block7b_se_reshape[0][0]
block7b_se_expand (Conv2D)	(None, 1, 1, 1920)	155520	block7b_se_reduce[0][0]
block7b_se_excite (Multiply)	(None, 8, 8, 1920)	0	block7b_activation[0][0] block7b_se_expand[0][0]
block7b_project_conv (Conv2D)	(None, 8, 8, 320)	614400	block7b_se_excite[0][0]
block7b_project_bn (BatchNormal)	(None, 8, 8, 320)	1280	block7b_project_conv[0][0]
block7b_drop (FixedDropout)	(None, 8, 8, 320)	0	block7b_project_bn[0][0]
block7b_add (Add)	(None, 8, 8, 320)	0	block7b_drop[0][0] block7a_project_bn[0][0]
top_conv (Conv2D)	(None, 8, 8, 1280)	409600	block7b_add[0][0]
top_bn (BatchNormalization)	(None, 8, 8, 1280)	5120	top_conv[0][0]
top_activation (Activation)	(None, 8, 8, 1280)	0	top_bn[0][0]
global_average_pooling2d_1 (Glo)	(None, 1280)	0	top_activation[0][0]
dropout_4 (Dropout)	(None, 1280)	0	global_average_pooling2d_1[0][0]
dense_4 (Dense)	(None, 1024)	1311744	dropout_4[0][0]
dense_5 (Dense)	(None, 3)	3075	dense_4[0][0]
Layer (type)	Output shape	Param #	Connected to
Input (InputLayer)	(None, 240, 240, 3)	0	

outperform all other models, they were used as the base models in stacking running. Stacking works by deducing individual learners' biases in relation to the training set [26].

During the initial training stage, the three models generated new features, and cross-validation was used to demonstrate objective accuracy. The final ensemble model is validated

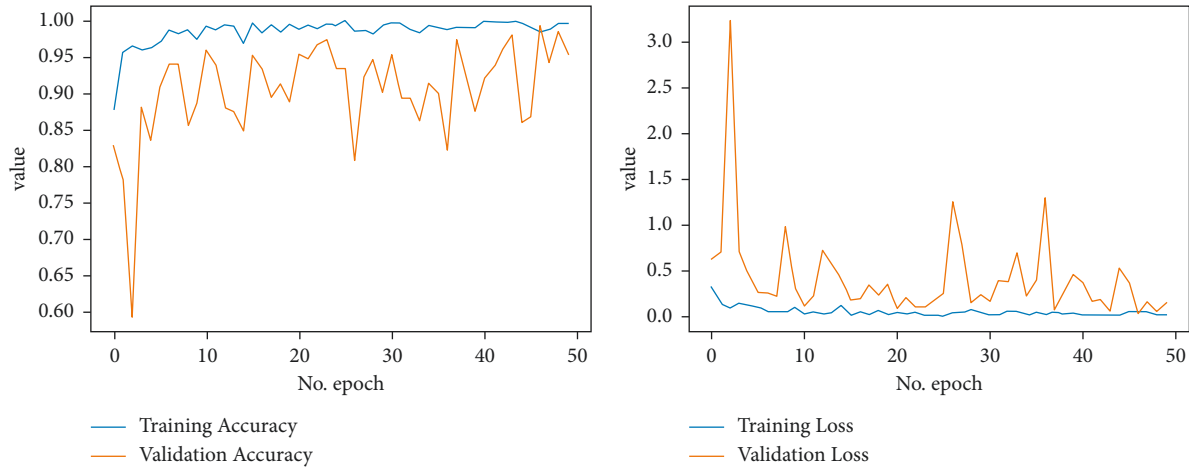


FIGURE 8: Variations in model accuracy and loss at the training and testing analysis.

TABLE 5: Model metrics assessment without stacking model.

Model		Metrics			
		Accuracy	AUC	Precision	Recall
Model 1	Efficient net V1	0.949	0.962	0.948	0.949
Model 2	Inception V3 (SVM)	0.967	0.980	0.970	0.974
Model 3	Inception V3 (RF)	0.915	0.980	0.906	0.908
Model 4	Inception V3 (LR)	0.975	0.996	0.975	0.975
Model 5	Inception V3 (KNN)	0.953	0.992	0.954	0.950
Model 6	SqueezeNet (SVM)	0.972	0.997	0.975	0.974
Model 7	SqueezeNet (RF)	0.910	0.982	0.917	0.918
Model 8	SqueezeNet (LR)	0.969	0.996	0.963	0.962
Model 9	SqueezeNet (KNN)	0.950	0.990	0.945	0.944

using LR and SVM algorithms. The model selection process is critical to the performance of the stacking model. All base models should perform admirably. Following model selection, there are no additional limitations on the newly created features. The stacking model and the other models' evaluation results are summarized in Table 6.

Table 6 indicated that the Inception V3 stacking model with LR (Model 10) produced the most accurate results. Additionally, the SqueezeNet stacking model with SVM outperformed the KNN and RF combination, indicating that the combination of stacking significantly improved the model's performance. Combining Inception V3 deep learning and LR as metalearner, stacking ensemble produced the best classification performance. Additionally, it demonstrated slightly improved classification performance for the datasets when SqueezeNet-SVM metalearners were used. This result demonstrated that stacking LR models can significantly improve prediction performance without requiring additional parameter tuning. In terms of performance, the stacking model outperformed the others in terms of accuracy, AUC, precision, and recall (Figure 9). Likewise, EfficientNet B1 showed a lower performance, although the cores are still comparable with approximately 4% difference with the best model. By fine-tuning or changing the B version of EfficientNet from B1 to a higher version, it is possible to explore a deeper network and increase the model's accuracy. Model ensemble has been shown to be

superior to traditional ML in making predictions [27]. Additionally, [28] reported that incorporating logistic regression, support vector machine, and neural network approaches into Fuzzy Dempster-Shafer (FDS) analysis improves the classification accuracy of paddy rice images significantly.

A simple corresponding plot of prediction errors was created for each of the evaluation datasets in order to investigate the effect of the various tassel traits on classification ability (Figure 10). The prediction errors indicate how perplexed the model is while predicting the tassel type. The plot indicates that the majority of misclassification occurs in female and contamination lines, whereas male lines exhibit less misclassification. Breeding of hybrid maize is essential for developing new plant varieties with higher yield potential and diseased resistance. Tassel morphology is considered one of the most essential traits for generating high purity maize hybrid [29]. Several lines were observed with slightly similar tassel morphological characteristics, including Mal03, CLYN 230, and CY7 (parental lines of other varieties) variety, which potentially makes misclassification. The male parent has also shown slightly similar attributes to the Mr14 line (parental line of other varieties). Further, the CNN model is susceptible to the slight change of tassel input. Tassel that emerged mainly occurred between 50 and 54 days after planting under a tropical lowland environment. However, the tassel that emerged within 3-4 days earlier

TABLE 6: Model metrics assessment with stacking model.

Model	Model	Metrics			
		Accuracy	AUC	Precision	Recall
Model 1	Efficient net V1	0.949	0.962	0.948	0.949
Model 2	Inception V3 (SVM)	0.967	0.980	0.970	0.974
Model 3	Inception V3 (RF)	0.915	0.980	0.906	0.908
Model 4	Inception V3 (LR)	0.975	0.996	0.975	0.975
Model 5	Inception V3 (KNN)	0.953	0.992	0.954	0.950
Model 6	SqueezeNet (SVM)	0.972	0.997	0.975	0.974
Model 7	SqueezeNet (RF)	0.910	0.982	0.917	0.918
Model 8	SqueezeNet (LR)	0.969	0.996	0.963	0.962
Model 9	SqueezeNet (KNN)	0.950	0.990	0.945	0.944
Model 10	Inception V3 (stacking model)	0.980	0.998	0.98	0.979
Model 11	SqueezeNet (stacking model)	0.974	0.992	0.966	0.967

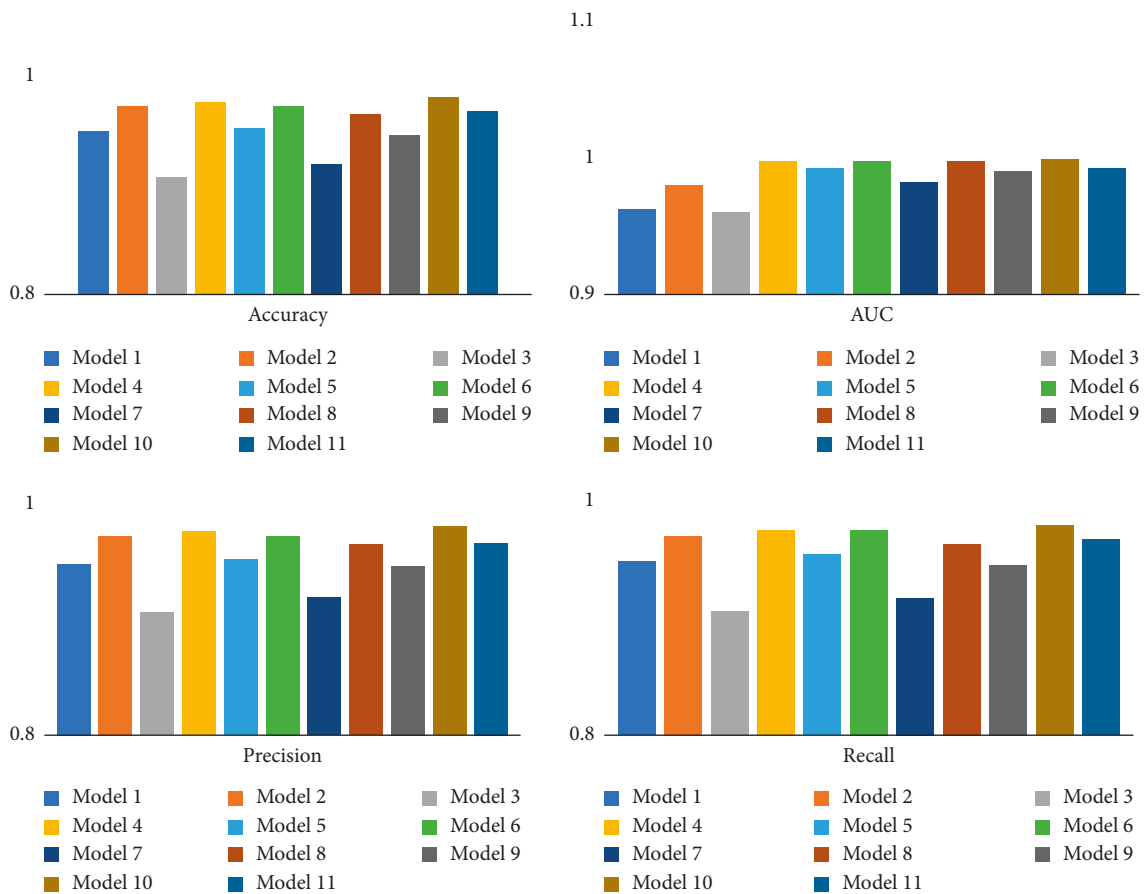


FIGURE 9: Comparison measures to determine the robustness of the models.

produced a different pattern of pollen shed and hanging off of the tassel. After emerging for 7–10 days, the tassel will gradually be dried off, which could be attributed to environmental conditions. Illumination and background images were factors that affected the proper detection of the tassel at random moments. Under natural conditions, information extraction from the plant would produce substantial amounts of noise such as the surrounding leaf, skylight, and sunshine illumination. Furthermore, [30] explained that computer vision advancement allows the machine to identify and detect tassel location with a challenging

accuracy influenced by background information. Besides, [31] also reported that the overlapping technique used in this study was one commonly used in the detection of the maize tassel worldwide using global regression [32–34] as well as local regression [35–38].

3.3. Interactive Data Visualization. The CNN method is a complex model, and many researchers have not fully understood the knowledge of network operations and behaviour in achieving good performance. In many cases, trial

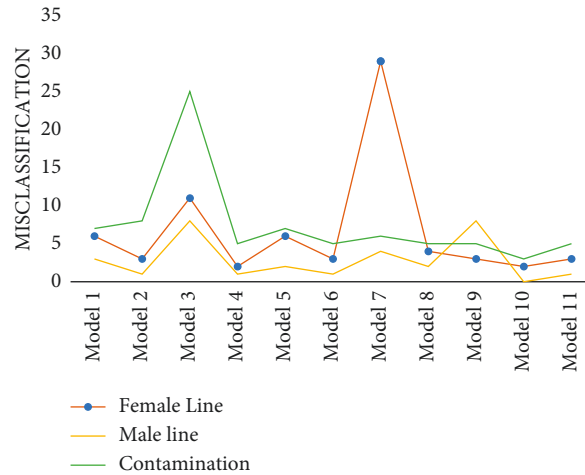


FIGURE 10: Plot of misclassification instances associated with the evaluation datasets.

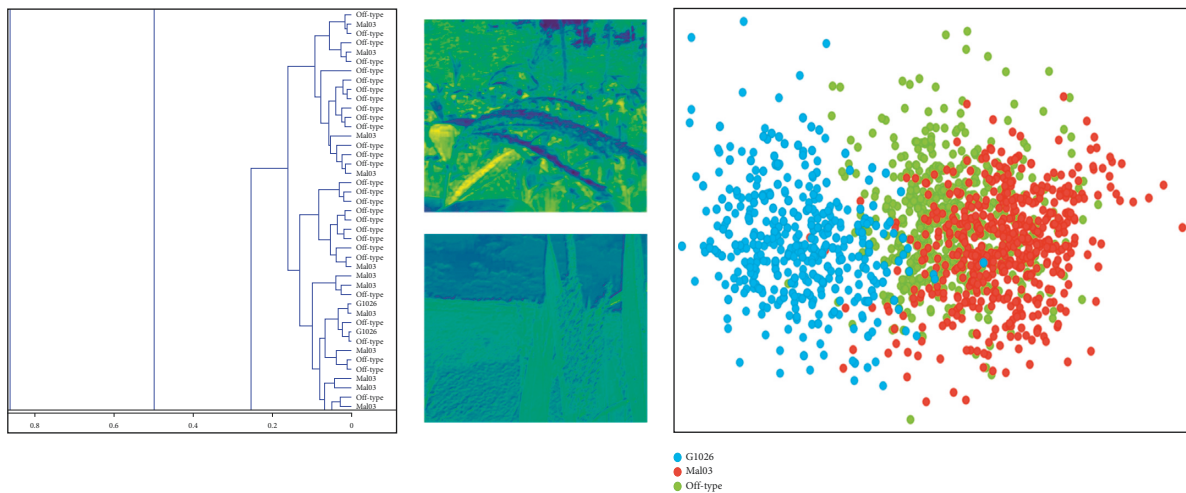


FIGURE 11: Dendrogram and multidimensional data projection MDS analysis of tassel.

and error methods have been deployed to ascertain the tuning parameters and network architecture to produce optimal results where needed. During the training process, a layer visualization was performed to observe how the model identified the inputs that were entered into the model. The weights of the convolution process were visualized to estimate how well the model can be trained. A good training model usually has a smooth and continuous filter, whereas the overfitting model will display a pattern with lots of noise [39]. The visualizations became more complex by encoding a more profound process of extracting features that were actively performed. Furthermore, a connected layer was mainly used to reconnect all nodes and determine which node was the most correlated with a particular class. Although the training visualization process has been carried out regarding the differences in tassels' morphology, the types of tassels that are off type are sometimes complicated to analyse even by using visual observation.

To explore and gain deeper information on misclassification and make sense of it, a multidimensional scaling

(MDS) analysis was applied to the model. Firstly, all vector parameters of the deep learning model were fed into a hierarchical clustering widget to generate a tassel dendrogram to classify each input parameter's similarity according to their target class. The Euclidian distance was applied to calculate the spaces between the embedded image parameters and display the result by constructing a connecting dendrogram. For further exploration, an MDS analysis was involved in projecting multidimensional data to a 2D space to allow a deeper understanding of the dataset's underlying pattern. The workflow of the dendrogram and multidimensional data projection is shown in Figure 11.

The highlighted MDS plot is shown in Figure 7 and derived from three different maize lines/genotypes classes and has various vector features related to the line/genotype class. There are three groups, blue for male plants (G102612), red for female plants (Mal03), and green for off-type plants. MDS separates the G102612 and Mal03 plants very well, and the separation is less pronounced between Mal03 and off-type plants (Figure 7). This indicates that the prediction

errors mainly occurred between Mal03 and off-type plants due to the highly similar data features, making it difficult for the model to accurately diagnose these two classes in several datasets. This unsupervised data is very useful to help plant breeders to get insight into the morphological aspects and additional information from big data processing of plant features (over 3,251,763 parameters).

Several studies indicated that the problematic light condition affects the overall classification of the model being tested [40]. Several False Negative cases were found in the classification process, especially in tassels that had similar shapes. Breeders or technicians assigned for detasseling activities in the field were generally made up of specialists, fully equipped with technical knowledge to differentiate off-type lines not based only on tassel characters but also on other supporting factors such as leaf shape, type, leaf angle, and hair colour. Another essential method used to improve the accuracy of tassel-type classification was rearranging the input image in the relevant region of interest. It enforced the network to identify the line-off-type from the entirely appropriate, focused area. Additionally, a maximum time is warranted at a fixed interval for capturing the tassel before its rupture and also to prevent pollens from off-type lines to adulterate the hybrid maize seeds produced.

4. Conclusions

The classification of lines was performed using deep learning and ensemble machine learning combinations. This pair of cutting-edge tools is capable of distinguishing between tassel morphological characteristics. Three popular CNN models were examined for the classification of hybrid maize parental lines based on their tassel characteristics: Inception V3, SqueezeNet, and EfficientNet. To summarize, stacking with advanced deep learning as the base learner and logistic regression as the metalearner may be considered the optimal classification model for tassel classification and removal during seed production. The accuracy of the stacking model Inception V3-logistic regression is 98 percent. It also proves that model stacking is effective for high-dimensional features. Similarly, SqueezeNet and EfficientNet B1 demonstrated slightly lower performance, despite the fact that their scores are comparable. To further investigate, a multidimensional scaling analysis was performed by projecting multidimensional data to a two-dimensional space. This analysis revealed that prediction errors occurred primarily between female and off-type plants due to the highly similar data features, making it difficult for the model to accurately diagnose these two classes across multiple datasets. The ensemble model is advantageous for assisting plant breeders in gaining insight into morphological aspects and additional information derived from big data processing of plant characteristics. Further research is required to incorporate machine learning into smartphones to enable in-field and real-time tassel classification.

Data Availability

All data used to support the findings of this work are available upon request from the corresponding author.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

The research was financially supported by the Ministry of Agriculture, Republic of Indonesia (Grant no. 052-524111-2020). Various genetic materials were provided by CIMMYT Mexico. Thanks are due to the technicians of Bajeng Experimental Station, Indonesia, for their support.

References

- [1] K. A. Sulaiman and I. K. Kariyasa, *Cara Cepat Swasembada Jagung*, Jakarta, Indonesia.
- [2] P. S. Setimela and P. Kosina, *Strategies for Strengthening and Scaling Up Community-Based Seed Production*, CIMMYT, Mexico, 2006.
- [3] D. C. Cires, U. Meier, J. Masci, and L. M. Gambardella, "IJCAI11-210.pdf," in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, pp. 1237–1242, Barcelona, Spain, July 2003.
- [4] H. Lu, Z. Cao, Y. Xiao, B. Zhuang, and C. Shen, "TasselNet: counting maize tassels in the wild via local counts regression network," *Plant Methods*, vol. 13, no. 1, p. 79, 2017.
- [5] C. DeChant, T. Wiesner-Hanks, S. Chen et al., "Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning," *Phytopathology*, vol. 107, no. 11, pp. 1426–1432, 2017.
- [6] M. Neumann, L. Hallau, B. Klatt, K. Kersting, and C. Bauckhage, "Cell phone image-based plant disease classification," *Advances in Environmental Engineering and Green Technologies*, vol. 23, pp. 295–322, 2015.
- [7] Z. Jiang, C. Liu, B. Ganapathysubramanian, D. J. Hayes, and S. Sarkar, "Predicting county-scale maize yields with publicly available data," *Scientific Reports*, vol. 10, Article ID 14957, 2020.
- [8] A. D. d. Medeiros, M. D. Pereira, T. F. S. N. Soares, B. G. Noronha, and D. T. Pinheiro, "Computer vision as a complementary method to vigour analysis in maize seeds," *Journal of Experimental Agriculture International*, vol. 25, pp. 1–8, 2018.
- [9] S. Khaki, H. Pham, Y. Han, A. Kuhl, W. Kent, and L. Wang, "Convolutional neural networks for image-based corn kernel detection and counting," *Sensors (Switzerland)*, vol. 20, pp. 1–16, 2020.
- [10] C. Zhang and Y. Ma, *Ensemble Machine Learning: Methods and Applications*, Springer, Berlin, Germany, 2012.
- [11] H. Byeon, "Exploring factors associated with the social discrimination experience of children from multicultural families in South Korea by using stacking with non-linear algorithm," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 5, pp. 125–130, 2021.
- [12] T. Han, A. Siddique, K. Khayat, J. Huang, and A. Kumar, "An ensemble machine learning approach for prediction and optimization of modulus of elasticity of recycled aggregate concrete," *Construction and Building Materials*, vol. 244, Article ID 118271, 2020.
- [13] P. Godec, M. Pančur, N. Ilenič et al., "Democratized image analytics by visual programming through integration of deep models and small-scale machine learning," *Nature Communications*, vol. 10, p. 4551, 2019.

- [14] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, December 2016.
- [15] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," arXiv, 2016.
- [16] D. H. Wolpert, "Stacked generalization neural networks 5 241–59," 1992.
- [17] G. Sigletos, G. Paliouras, C. D. Spyropoulos, and M. Hatzopoulos, "Combining information extraction systems using voting and stacked generalization," *Journal of Machine Learning Research*, vol. 6, 2005.
- [18] M. Tan and Q. V. Le, "EfficientNet: rethinking model scaling for convolutional neural networks," in *Proceedings of the 36th International Conference on Machine Learning*, Long Beach, CA, USA, June 2019.
- [19] N. Upadyayula, H. S. da Silva, M. O. Bohn, and T. R. Rocheford, "Genetic and QTL analysis of maize tassel and ear inflorescence architecture," *Theoretical and Applied Genetics*, vol. 112, no. 4, pp. 592–606, 2006.
- [20] D. W. Hosmer Jr., S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, John Wiley & Sons, Hoboken, NJ, USA, 2013.
- [21] U. Gawande, M. Zaveri, and A. Kapur, "A novel algorithm for feature level fusion using SVM classifier for multibiometrics-based person identification," *Applied Computational Intelligence and Soft Computing*, vol. 2013, Article ID 515918, 11 pages, 2013.
- [22] Z. Zhou, *Ensemble Methods: Foundations and Algorithms*, Chapman & Hall/CRC, Boca Raton, FL, USA, 2012.
- [23] S. Džeroski and B. Ženko, "Is combining classifiers with stacking better than selecting the best one?" *Machine Learning*, vol. 54, pp. 255–273, 2004.
- [24] M. Abadi, A. Agarwal, P. Barham et al., "TensorFlow: Large-scale machine learning on heterogeneous distributed systems," arXiv, 2016.
- [25] Q. Xie, M. T. Luong, E. Hovy, and Q. V. Le, "Self-training with noisy student improves imagenet classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, June 2020.
- [26] "Anon ensemble methods: bagging, boosting and stacking | by joseph rocca | towards data science,".
- [27] M. Shahhosseini, G. Hu, and S. V. Archontoulis, "Forecasting corn yield with machine learning ensembles," *Frontiers of Plant Science*, vol. 11, 2020.
- [28] T. C. Lei, S. Wan, S. C. Wu, and H. P. Wang, "A new approach of ensemble learning technique to resolve the uncertainties of paddy area through image classification," *Remote Sensor*, vol. 12, pp. 1–23, 2020.
- [29] F. Rácz, S. Hidvégi, S. Záborszky, M. Pál, and C. Marton, "Pollen production of new generation inbred maize lines," *Cereal Research Communications*, vol. 34, pp. 633–636, 2006.
- [30] H. Lu, Z. Cao, Y. Xiao, Z. Fang, Y. Zhu, and K. Xian, "Fine-grained maize tassel trait characterization with multi-view representations," *Computers and Electronics in Agriculture*, vol. 118, pp. 143–158, 2015.
- [31] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627–1645, 2010.
- [32] A. B. Chan, Z. S. J. Liang, and N. Vasconcelos, "Privacy preserving crowd monitoring: counting people without people models or tracking," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, Salt Lake City, UT, USA, June 2008.
- [33] M. V. Giuffrida, M. Minervini, and S. Tsafaris, "Learning to count leaves in rosette plants 1," in *Proceedings of the Proceedings of the Computer Vision Problems in Plant Phenotyping*, Swansea, UK, September 2015.
- [34] L. Fiaschi, U. Koethe, R. Nair, and F. A. Hamprecht, "Learning to count with regression forest and structured labels," in *Proceedings of the International Conference on Pattern Recognition*, Tsukuba, Japan, November 2012.
- [35] C. Zhang, H. Li, X. Wang, and X. Yang, "Cross-scene crowd counting via deep convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, June 2015.
- [36] C. Arteta, V. Lempitsky, and A. Zisserman, "Counting in the wild," in *Proceedings of the European Conference on Computer Vision*, Amsterdam, Netherlands, October 2016.
- [37] C. Arteta, V. Lempitsky, J. A. Noble, and A. Zisserman, *Interactive Object Counting BT-Pixel-Level Encoding and Depth Layering for Instance-Level Semantic Labeling*, Springer, Berlin, Germany, 2014.
- [38] J. P. Cohen, G. Boucher, C. A. Glastonbury, H. Z. Lo, and Y. Bengio, "Count-ception: counting by fully convolutional redundant counting," in *Proceedings of the 2017 IEEE International Conference on Computer Vision Workshop*, Venice, Italy, October 2017.
- [39] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proceedings of the European Conference on Computer Vision*, Zurich, Switzerland, September 2014.
- [40] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1635–1650, 2010.