Computers and Electronics in Agriculture 179 (2020) 105842

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

# Data augmentation for automated pest classification in Mango farms

Kusrini Kusrini, Dr., First Author (Kusrini)[a,*], Suputa Suputa, Dr. (Suputa)[b], Arief Setyanto, Dr. (Arief)[a], I Made Artha Agastya (Artha)[a], Herlambang Priantoro (Herlambang)[c], Krishna Chandramouli, Dr. (Krishna)[d], Ebroul Izquierdo, Prof. (Ebroul)[d]

[a] Universitas AMIKOM Yogyakarta, Jl. Ringroad Utara, Condong Catur Depok, Sleman Yogyakarta 55583, Indonesia
[b] Department of Crop Protection Faculty of Agriculture Universitas Gadjah Mada, Bulaksumur Caturtunggal Depok, Sleman Yogyakarta 55281, Indonesia
[c] PT. Bank Mandiri, Plaza Mandiri, 6th Floor., Jl. Jend, Gatot Subroto Kav 36-38, Jakarta 12190, Indonesia
[d] Multimedia and Vision Research Group, School of EECS, Queen Mary Universityof London, Mile End Road, London E1 4NS, UK

ARTICLE INFO

*Keywords:*
Mango farms
Leaves infestation
Pest classification
Data augmentation
VGG-16 network

ABSTRACT

Mangos are native to South and Southeast Asian regions. They are one of the favorite fruits consumed globally, with an overall estimated consumption reaching up to 50.65 million metric tons in 2017. However, the production of mango is usually severely affected by pests that attack the fruit, stem, root or mango leaf. Addressing the need for an early stage automated or semi-automated pest identification system, the research presented in this paper proposes an advanced machine learning (ML) technique for analyzing large-scale mango fields and identification of the onset of biological threats using computer vision and deep-learning technologies. The ML technique presented in the paper extends the pre-trained VGG-16 deep-learning model to supplement the last layer with a fully connected network training of consisting of 2-layers. In addition, the research presented in the paper also considers the real-world operational conditions commonly faced by Indonesian farmers for collecting and processing visual information obtained from the Mango farms. The sparsity of the dataset availability for effectively training deep-learning network is addressed through the application of data augmentation process that is able to accurately recreate the conditions faced by the farmers. The overall accuracy of the proposed training solution achieved is 73% on the validation dataset and 76% for the testing data. The application of the augmentation transformation function leads to an improvement of 13.43% of accuracy on the testing data.

## 1. Introduction

The agricultural industry and in particular the farming community has constantly faced the threat of pests and environmental disruptions and is being considered a severe threat for food security and economic stability for both farmers and general public (Strange & Scott, 2005). Traditionally, such challenges are addressed through the local knowledge of farmers which has been passed down through generations and has paved way for mitigating some of the impact of pests. While the use of advanced scientific tools and solutions have influenced been largely adopted by various industrial sectors in the Indonesian region, the use of mobile computation and cloud deployment of deep-learning network models for the automation of agricultural services has not been fully exploited. Among the several types of agricultural plants which are affected by pests, infestation of leaves is regarded to have the maximum impact upon the food production. Among several plantation farms available in Indonesia, with an average quantity of Mango production reaching up to 2.2 Million metric tons out of the total global production of 50.65 Million metric tons as reported in 2017, the farming of Mango is considered one of the key economic factors that influences the Indonesian GDP. In order to ensure a steady volume of production, it is vital to ensure that the impact of possible diseases affecting the quality of the fruit is mitigated at an early stage. Addressing such a critical challenge in farming, the current practice adopted aims for the farmers to manually inspect and observe the presence of leaves infestation. However, following the increase in the growth of plantation areas of Mango within Indonesia, it is no longer feasible for farmers to undertake large-scale surveillance of Mango farms.

* Corresponding author at: Magister of Informatics, Universitas AMIKOM Yogyakarta, Jl. Ringroad Utara Condong Catur Depok, Sleman, Yogyakarta 55283, Indonesia.
*E-mail addresses:* kusrini@amikom.ac.id (K. Kusrini), puta@ugm.ac.id (S. Suputa), arief_s@amikom.ac.id (A. Setyanto), artha.agastya@amikom.ac.id (I.M.A. Agastya), lambanx@gmail.com (H. Priantoro), krishna.chandramouli@qmul.ac.uk (K. Chandramouli), ebroul.izquierdo@qmul.ac.uk (E. Izquierdo).

https://doi.org/10.1016/j.compag.2020.105842
Received 21 October 2019; Received in revised form 2 October 2020; Accepted 15 October 2020
Available online 16 November 2020

Subsequent to the developments reported in the field of cloud computing and smart handheld devices, there is an increasing interest in the development of modern tools and techniques for automating Integrated Pest Management (IPM) system and solution. As identified by Ha in 2014 (Ha, 2014), the objective of IPM is to design and develop a system for managing pests in agricultural production that employs multiple tactics in consideration of economic, environmental, ecological and human health impacts. For the farmers to successfully adopt the usage of such IPM systems, it is critical to ensure appropriate information from the wide-scale farming regions are collected and processed in a timely manner.

Therefore, addressing the broad scope of challenges in the cultivation of Mango within Indonesian farming community, the paper presents a data augmentation process for improving the quality of the dataset with real-world pests captured followed by the design of a deep-learning network for creating a baseline for the classification of multi-class pests. In addition, the paper also reports on the deployment of an overall holistic framework for the farmers to identify the pests from the field. The specific contributions of the paper include:

- to generate a dataset for leaves infestation that reflects the challenges faced by real-world constraints
- to extend the sparse dataset through the proposal of data augmentation techniques that improve the robustness of the proposed machine learning algorithm against overfitting
- to design and implement a fully connected deep-learning network extending VGG-16 network for multi-class classification of pest categories
- to evaluate the performance of the augmentation process with 3 different experimental runs
- to report on the developments of the mobile application-based cloud deployment of machine learning model for pest-classification in the farming region.

The paper is structured as follows. In Section 2, an overview of the literature review is presented followed by the description of the proposed data augmentation process for creating a large dataset that would suitably enable the application of deep-learning techniques in Section 3. In addition, the section also outlines the proposed framework for classifying 16-class Mango pest classification framework consisting of 15-category of pests and healthy leaves based on the proposed refinement to the VGG-16 network. The section concludes with an outline of the overall workflow that enables farmers to effectively exploit the cloud-based deployment of the proposed solution and receive in real-time the classified output for the category of pests that might be affecting the production of Mango. The experimental results of the data augmentation process and subsequently the classification output from the deep-learning network are presented in Section 4. In Section 4, a brief discussion and summary of the proposed approach is presented. The paper concludes with a remark on the proposed novelty and the distribution of data assets in Section 5 along with an outline of the future work.

## 2. Literature review

The scope of the research presented in the paper relates to the use of data augmentation process and the classification models for categorizing multi-class pests that affect Mango cultivation. Therefore, the literature review presented in this section has been appropriately categorised into two sub-sections. The first subsection outlines the reported techniques in the literature on the use of data augmentation process for enriching the sparse datasets, while the second subsection presents in detail the various machine learning and deep-learning algorithms that have been reported in the literature for pest classification.

### 2.1. Data augmentation

In order to build robust deep-learning models, it is critical to ensure the validation error during the training phase to be continually minimised along with training error. One of the approaches that has been successfully reported in the literature is data augmentation process (Shorten & Khoshgoftaar, 2019). The augmented data will represent a more comprehensive set of possible data points, thus minimizing the distance between the training and validation set, as well as any future testing sets. One of the common pitfalls in machine learning algorithms is for the algorithm to overfit on the training dataset and thus lose the ability to generalise the training model for new information that is presented upon which the network has not been trained. In order to address the robustness of the quality of training, several techniques have been published for improving the generalization ability of these models. The term 'Generalizability' refers to the performance difference of a model when evaluated on previously seen data (training data) versus data it has never seen before (testing data). Models with poor generalizability have overfitted the training data. One way to discover overfitting is to plot the training and validation accuracy at each epoch during training (Shorten & Khoshgoftaar, 2019).

In contrast, inspired by ManiFool (Paschali et al., 2019), present an augmentation process that is performed by a line-search manifold-exploration method which is hypothesised to learn the affine geometric transformations that had led to the misclassification on an image, while ensuring that it remains on the same manifold as the training data. This augmentation method populates any training dataset with images that lie on the border of the manifolds between two-classes and maximizes the variance the network is exposed to during training. The data augmentation process had been thoroughly evaluated on the challenging tasks of fine-grained skin lesion classification from limited data, and breast tumour classification of mammograms. However, such techniques have not been further explored in the context of agricultural domain for multi-class classification problems.

In comparison, the image-based data augmentation process aims to perform data transformations, that will result in the increase of problem specification which will be used to train the network for achieving generalisation. A brief review of the various image manipulations that are commonly adopted in the literature for the development of data augmentation strategies and policies like *flipping*, *color space*, *cropping*, *rotation*, *translation*, *noise injection* and *Colour space transformation.* Despite these publications, the challenge of pest recognition that affects Mango cultivation in the Indonesian region remains an open problem.

### 2.2. Machine learning for multi-class pest classification

Following the recent innovations reported in the field of deep-learning and machine learning algorithms, a limited number of articles have been published addressing the challenge of pest detection based on the image processing.

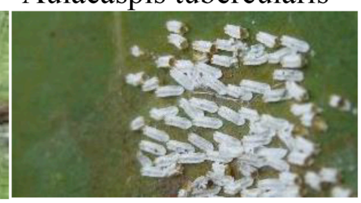
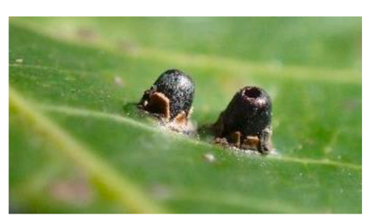
The cost of capturing visible scale images using low-cost visible sensors has been identified as a suitable for detecting pest. As such, the use of handheld devices including smart phones and tablets have been increasingly receiving acceptance among farmers for collection of information and subsequent processing (Zhang, Wu, You, & Zhang, 2017). The study of pests affecting the plants has been the areas of study for many researchers and in particular the use of images captured from smartphones has been accepted by farmers (Johannes, Seitz, & Se, 2017). Following the wide-scale data aggregation from the fields, the application of statistical tools and machine learning algorithm for the classification of multi-class pests has been reported in the literature including the use of Support Vector Machine (SVM) (Avendano, Ramos, & Prieto, 2017), Neural networks (Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, 2016), deep learning (Lu, Yi, Zeng, Liu, & Zhang, 2017) (Mohanty, Hughes, & Salathé, 2016) among others.

Subsequent to the increasing popularity of deep-learning network

**Fig. 1.** Example of Leaves infestation.

models that have been applied across other critical domains such as medical, object classification, the use of deep-learning algorithms has been proposed in the literature for the classification of pests (Mohanty et al., 2016) with results reported in a classification with a fairly good level of accuracy at 99.35% of the 26 types of plant diseases commonly affecting approximately 14 agricultural commodities. In particular, the use of Convolution Neural Network (CNN) for the classification of plant diseases affecting rice produce has been examined by (Lu et al., 2017). In

**Table 1**
Types of Pests affecting Mango farms in Indonesia.

| | | |
|---|---|---|
| Apoderus javanicus | Ischnaspis longirostris | Aulacaspis tubercularis |
|  |  |  |
| Mictis longicornis | Ceroplastes rubens | Neomelicharia sparsa |
|  |  |  |
| Cisaberoptus kenyae | Orthaga euadrusalis | Dappula tertia |
|  |  |  |
| Procon-tarinia matteiana | Dialeuropora decempuncta | Procon-tarinia rubus |
|  |  |  |
| Erosomyia sp | Valanga nigricornis | Icerya seychellarum |
|  |  |  |

the study authors report that, the type of disease was limited to 10 types of diseases that most often attacked rice and produced an accuracy of 95.48%. In the context of agriculture, the number of cultivated plants vary in their origin along with environmental impact on climate and other parameters. The evaluation of the pest and disease classification techniques as reported in the literature considers a limited type of classes that are detected. For example (Zhang et al., 2017) presented the research formulated to detect pests and diseases affecting only one type of apples, cucumbers, tulips, and rice.

The research presented in this paper aims at addressing the needs of the farmers with accessibility to real-time development and deployment of engineers other than for maintenance. In the next section, a detailed outline of the various techniques discussed is reported for the pest recognition system.

## 3. Proposed framework for pest recognition system

The proposed framework for pest recognition relies on the processing of real-world images captured with low-cost handheld devices from the Mango farms. One of the crucial requirements that is addressed in the research in the lack of resources for pre-processing the images captured by farmers. Thus, the image from the farm is processed as is and thus presents a set of unique challenges with complex background and partial occlusions and overlapping leaf structures upon other pests Therefore, addressing these challenges, the training of the pest recognition framework is carried out using data augmentation techniques such as noise, blur and contrast along with affine transformations. The rest of the section provides a detailed outline of the data generation process carried out in the paper for training the pest recognition framework.

### 3.1. Data generation

One of the key contributions of the research presented in the paper is the dataset development for training multi-class pest infestation network. As mentioned earlier, the pest infestation process is unique to the regional environment and the plantation that is being affected, the data collection was conducted through samples aggregated from mango plants throughout Indonesian archipelago. A range of samples that have been collected from the data collection is presented in Fig. 1, with the left most image representing the original image and the rest of the images in Figure represent some form of pest infestation.

The overall cultivation of the Mongo trees has been affected by a total of 181 pests including disease-causing pathogens and weeds out of which, 80% are commonly found in Indonesia (Suputa et al., 2015). The infestation of these pests tends to cause farm for various parts of the Mango cultivation including leaves, fruits, branches, stems and roots. Among these pests, 48 distinct types of pests have been identified to harm the leaves of the Mango trees globally. Therefore, in the research presented in the paper, 15 unique categories of pests that are identified to cause the most harm to Mongo cultivation are being studied and analyzed. The selected 15-pest categories also result in the structural deformity of the Mango leaves, facilitating the farmers to quickly contain the spread of the pest across the farm. Among the various challenges, researchers' carryout studies on pest infestation management often suffer from the limited amount of resources available for carrying out large-scale tests for automating the detection and categorization of pests' classes. These pests, *Apoderus javanicus, Aulacaspis tubercularis, Ceroplastes rubens, Cisaberoptus kenyae, Dappula tertia, Dialeuropora decempuncta, Erosomyia* sp.*, Icerya seychellarum, Ischnaspis longirostris, Mictis longicornis, Neomelicharia sparsa, Orthaga euadrusalis, Procontarinia matteiana, and Valanga nigricornis*, are commonly occurring in Indonesia and have been identified as a threat to the economic welfare among trading partner countries, such as Australia (Australian Government Department of Agriculture and Water Resources).

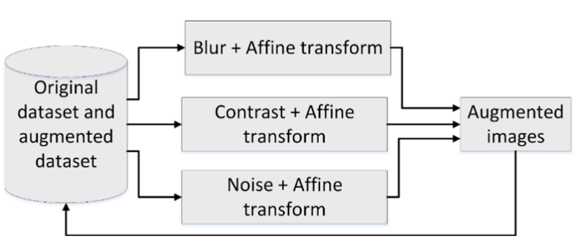



**Fig. 2.** Data augmentation workflow for training deep-learning network models.

As an instance, the population of *Apoderus javanicus* increases from August to September (Manjunath, 2018). There is high population density of *Aulacapsis tuberculari* during April to August (Salem, Mahmoud, & Ebadah, 2015). *Cisaberoptus kenyae* is recorded every year and the highest populations have been witnessed between January and August (Abou-Awad, Metwally, & Al–Azzazy, 2009). The most serious damage of D. tertia larvae often appears between June and July (Chang et al., 2018). High populations of Dialeuropora decempuncta is witnessed between March to June and low populations from October to January (Singh, Maheshwari, & Saratchandra, 2005). Icerya seychellarum exists every year, and the population increases from March and subsequently decreases from September (Mohamed, 2015).

The occurrence of following pests (*Erosomyia* sp.*, Ceroplastes rubens, Ischnaspis longirostris, Neomelicharia sparsa, Mictis longicornis, Orthaga euadrusalis, Procontarinia matteiana, Valanga nigricornis*) have not been formally recorded in the literature, but based on the observations in the mango farms, these pests are always found at each time of observation. Based on observations in Indonesia the existence of this pest is always found throughout the year.

There is a severe lack of visual examples from the suspicious destructor as this pest only appears occasionally but causes severe economic damage to the overall Mango cultivation. An overview of the different types of pests and associated number of image samples collected from the Indonesia has been tabulated in Table 1. These pests, *Apoderus javanicus, Aulacaspis tubercularis, Ceroplastes rubens, Cisaberoptus kenyae, Dappula tertia, Dialeuropora decempuncta, Erosomyia* sp.*, Icerya seychellarum, Ischnaspis longirostris, Mictis longicornis, Neomelicharia sparsa, Orthaga euadrusalis, Procontarinia matteiana, and Valanga nigricornis*, are commonly occurring in Indonesia and have been identified as a threat to the economic welfare among trading partner countries, such as Australia (Australian Government Department of Agriculture and Water Resources).

As an instance, the population of *Apoderus javanicus* increases from August to September (Manjunath, 2018). There is high population density of *Aulacapsis tuberculari* during April to August (Salem et al., 2015). *Cisaberoptus kenyae* is recorded every year and the highest populations have been witnessed between January and August (Abou-Awad at al., 2009). The most serious damage of D. tertia larvae often appears between June and July (Chang et al., 2018). High populations of Dialeuropora decempuncta is witnessed between March to June and low populations from October to January (Singh et al., 2005). Icerya seychellarum exists every year, and the population increases from March and subsequently decreases from September (Mohamed, 2015).

The occurrence of following pests (*Erosomyia* sp.*, Ceroplastes rubens, Ischnaspis longirostris, Neomelicharia sparsa, Mictis longicornis, Orthaga euadrusalis, Procontarinia matteiana, Valanga nigricornis*) have not been formally recorded in the literature, but based on the observations in the mango farms, these pests are always found at each time of observation. Based on observations in Indonesia the existence of this pest is always found throughout the year. The images depicted in the table represent the primary data that was captured from infested leaves from Mango farms. In order to highlight the visual characteristics of the pests, the collected images were cropped to indicate part of the pest specific characteristic. The images used in the experiments were taken real time from infested leaves, without any preprocessing.

Despite the availability of a significant image database, the amount

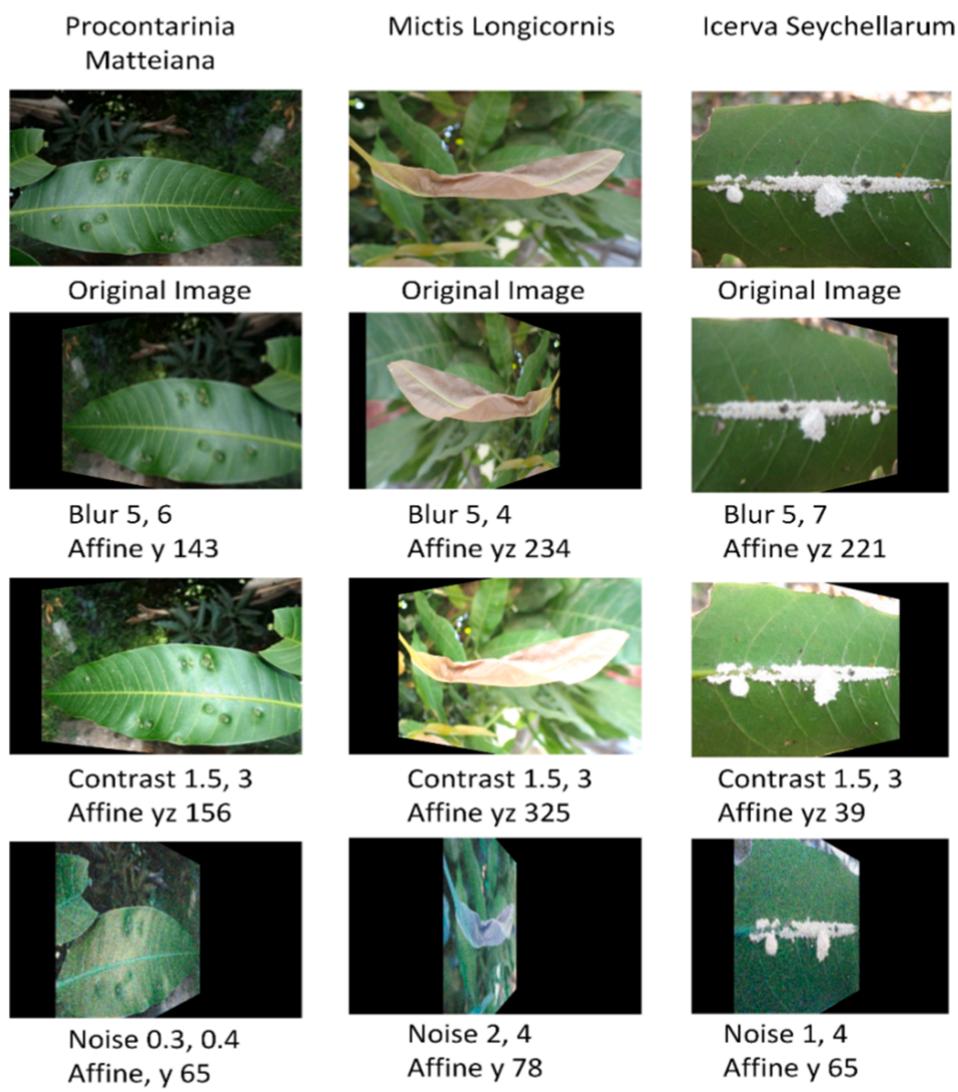



**Fig. 3.** Sample images in the augmented dataset.

of data is not enough to develop advanced classification models based on deep learning. In addition, due to the micro-differences between the various types of pest infestation the use of statistical machine learning techniques based on pixel-level hand-crafted features has not yielded high accuracy for the multi-class classification of pests. Therefore, the use of data augmentation techniques has been adopted for improving the quality of the data for the training of deep-learning network. The high-level outline of the image manipulation steps adopted for data augmentation in this paper is presented in Fig. 2. For each image collected from the Mango cultivation field, the samples are subjected to five mathematical operations namely (i) noise; (ii) blur; (iii) contrast and (iv) affine transformation. The novelty of the proposed approach is the construction of the sequential and non-sequential workflow in which each of the mathematical operators is performed in several combinations.

The implementation framework of the augmentation sequence is presented in Fig. 2. The original dataset from obtained from the Mango farm has been subjected to three distinct forms of augmentation process. The steps have been carefully chosen to ensure the resulting outcome is subsequently used to train the deep-learning network that is able to avoid overfitting to the given training dataset. In addition, it is also worth noting that the selection of the various data augmentation steps and sequences that have been considered are inspired from the data collection process as undertaken by the farmers, in which the collected samples have been found to be rotated, blurred and illumination changes.

- **Noise Addition:**

In order to achieve robust generalisation of the image dataset, the next step in the data augmentation process included in the introduction of noise levels based on Gaussian distribution as presented in (Das et al., 2016). In order to ensure the visibility of the random noise the random number is multiplied by a regularization constant. The mean parameter is randomly generated between {0.1, 0.2, 0.3, 0.4}, meanwhile the deviation parameter between {0,0.1, …, 0.5}

- **Blur:**

The next data augmentation process implemented in the paper relates to blur and often represents the lack of auto-focus functionality in the collection of data samples. Therefore, the blur parameters that are used to transform the images also follow the Gaussian distribution of the parameters. This implementation of the Gausian blur was carried out by kernel size 5 and standard deviation {2, 3, 4, 5, 6, 7}

- **Contrast:**

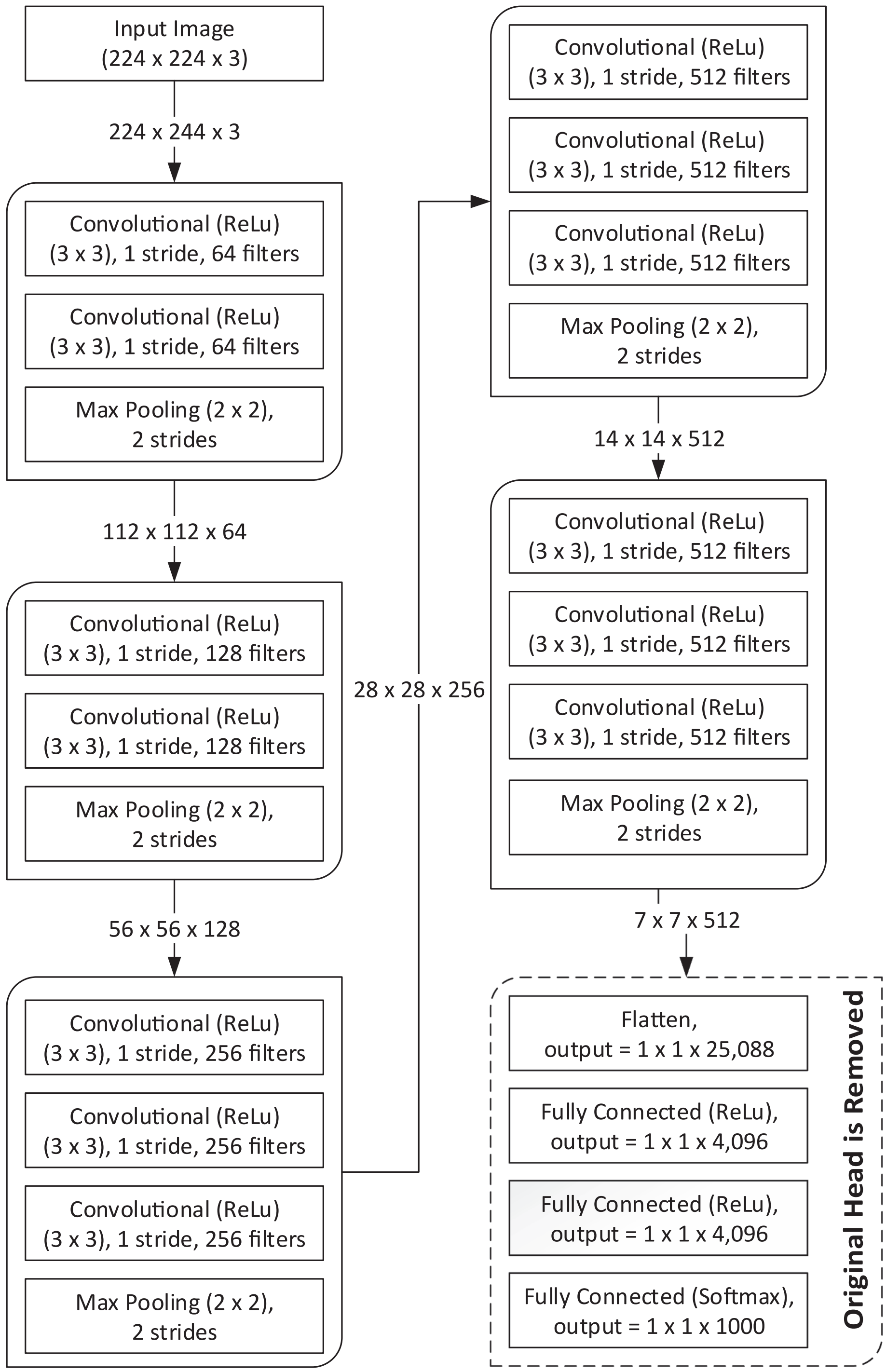



**Fig. 4.** VGG-16 Architecture with the original head is removed.

Contrast variations (Szeliski, 2010) are generated using contrast parameter {1, 1.5} and brightness {0,1,2,3,4,5}.

- **Affine Transformation:**

The affine transformation is used to simulate images rendered from different camera positions and projections.

Sample outcomes of the data augmentation process are presented in Fig. 3. Here, three examples from the original dataset has been selected upon which three data distortion operations, as presented in Fig. 2, have been applied. In total, for every image in the original dataset, additional 150 images have been generated through the presented augmentation process.

### 3.2. Proposed machine learning framework

Following the dataset generation step, the next stage in the processing relates to the training of the deep-learning network for the classification of different types of pests that affect Mango plants in Indonesia. In order to achieve this, the VGG-16 network architecture (Simonyan & Zisserman, 2014) has been updated to replace the last block containing the softmax classification with a fully-connected layer (FCL) and train the network with the image features extracted from the VGG-16. The weights of the VGG-16 network have been preserved from

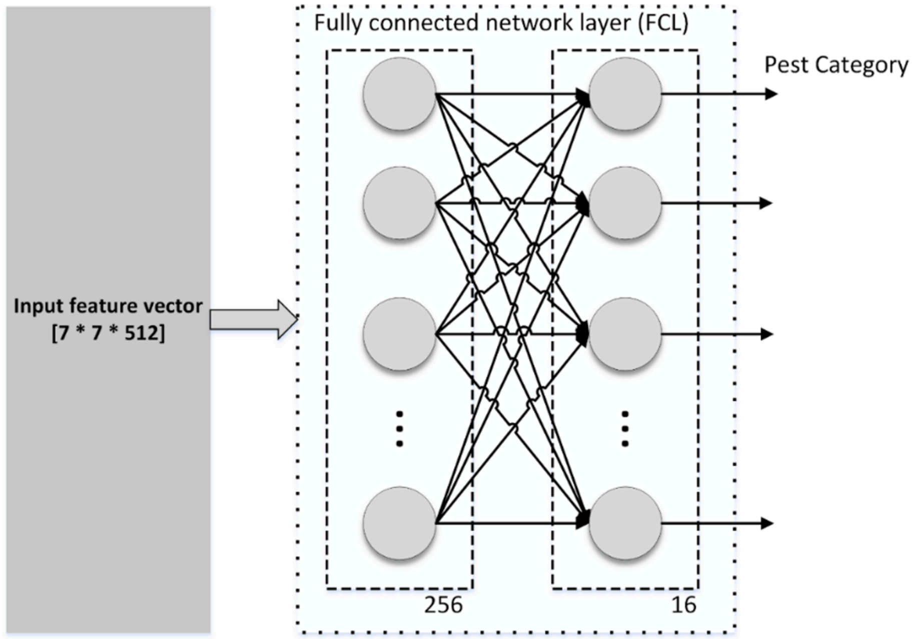



**Fig. 5.** The new head of VGG-16.

the pre-trained model. The extracted features are flattened and used as input to the FCL layer network which consists of 2-layers. The first layer is activated by the ReLU function and consists of 256 nodes, followed by the second layer consisting of 16 nodes activated by softmax. The output of the network classifies the pest model. The VGG-16 network architecture is presented in Fig. 4 followed by the proposed training framework depicted in Fig. 5.

### 3.3. Implementation specification

The overall implementation of the FCL network is achieved in TensorFlow using Keras library (Chollet François, 2015). The training of the Pest classification has been achieved through by establishing a robust visual correspondence between the various types of the pests. The training parameters utilized include a learning rate of $1 \times 10^{-5}$ and epoch of 50. The loss function has been calculated using binary cross-entropy model.

## 4. Result and discussion

The overall original dataset has been divided into three subsets namely (i) training; (ii) validation and (iii) testing with 60%, 20% and 20% respectively. As the quantity of datasets available is not enough to robustly train the deep-learning network models, the process of data augmented as presented in Section 3, is carried out. Therefore, the objective of the experimentation process is to validate both the

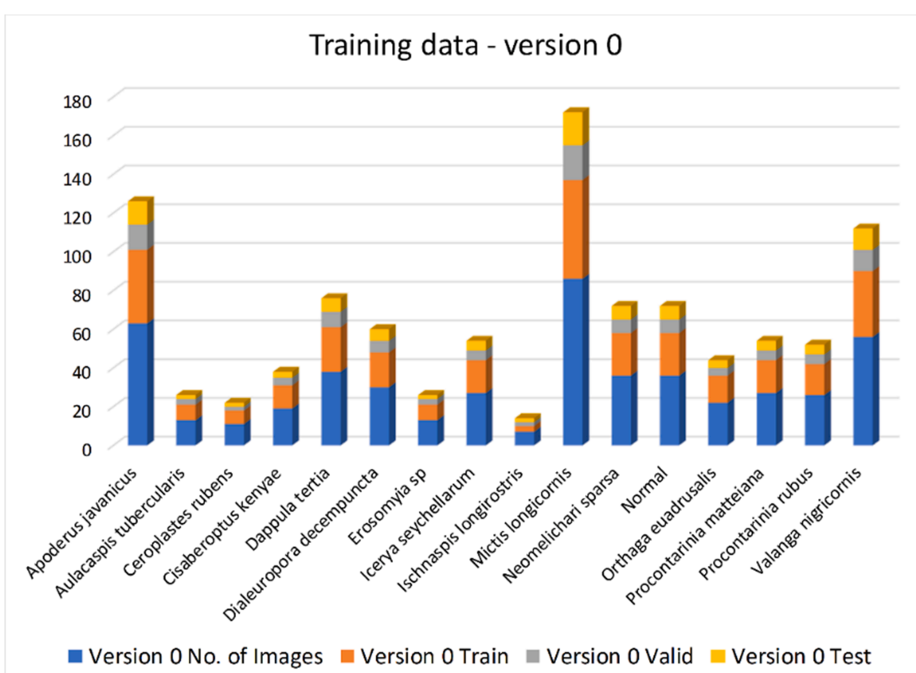



**Fig. 6.** Dataset distributed used for experimentation, version 0.

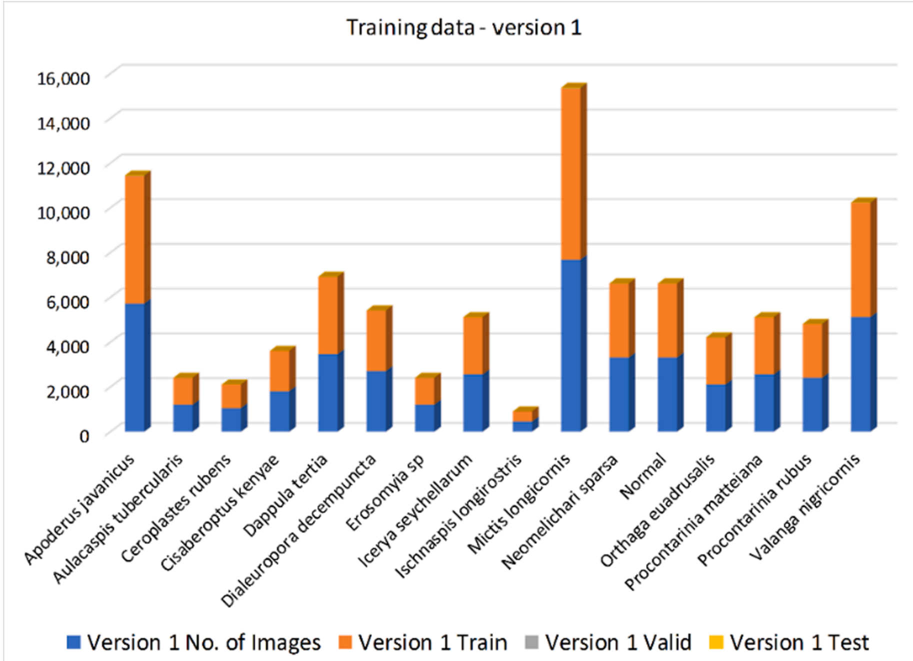



**Fig. 7.** Dataset distribution used for experimentation, version 1.

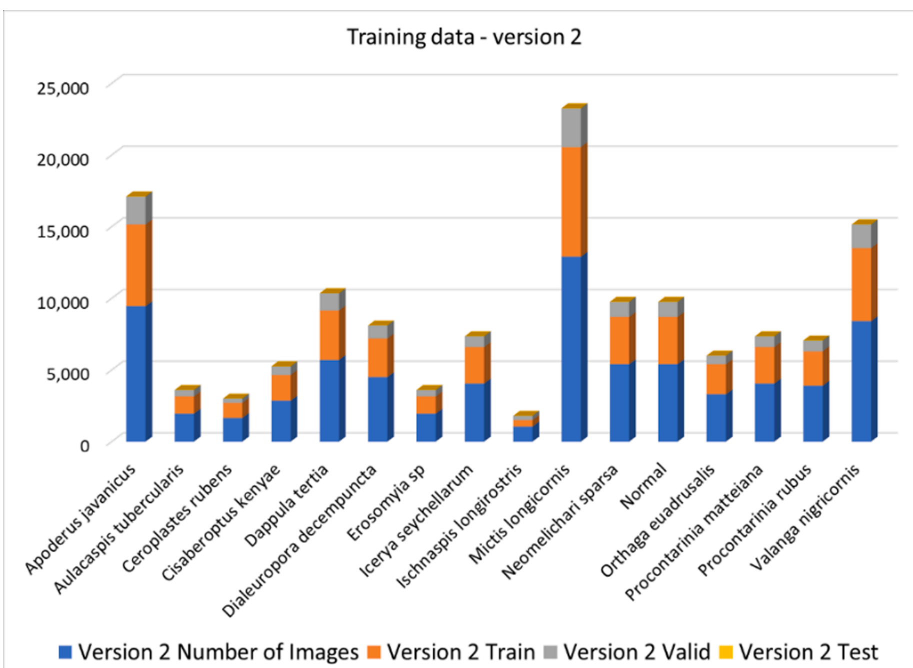



**Fig. 8.** Dataset distribution used for experimentation, version 2.

performance of the augmentation process and the proposed VGG-16 network improvement with the FCL network for training on both original and augmented dataset for improved classification process.

### 4.1. Experimental setup

In order to evaluate the impact of data augmentation, we carry out three experimental scenarios, as below:

- Version 0: the dataset consists of 510 original images as captured from the mango cultivation farms in Indonesia. The overall image dataset is divided into 3 parts with 310 images as training data, 103 images as validation data, 97 images as testing data. The objective of the experiment is to calculate the baseline performance of the system when the proposed network is trained using the limited number of images, validated and tested with originally captured images without any augmentation. The distribution of the dataset is presented in Fig. 6.
- Version 1: the image dataset consists of 46.500 samples as training data following the application of the data augmentation process. The objective of the scenario is to evaluate the performance of the network training while the validation and testing sequences

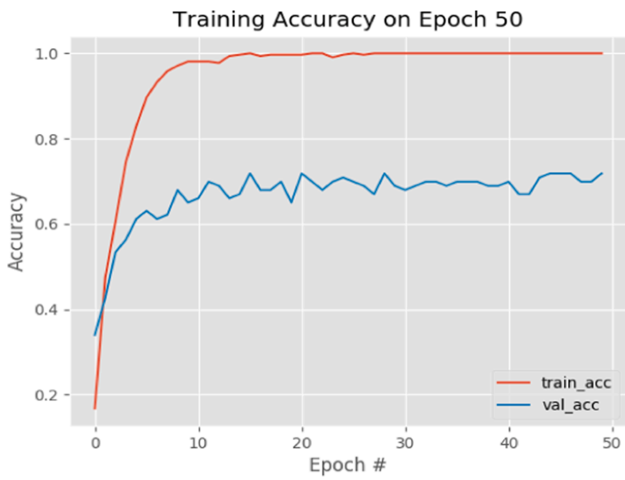



**Fig. 9.** Training process for version 0.

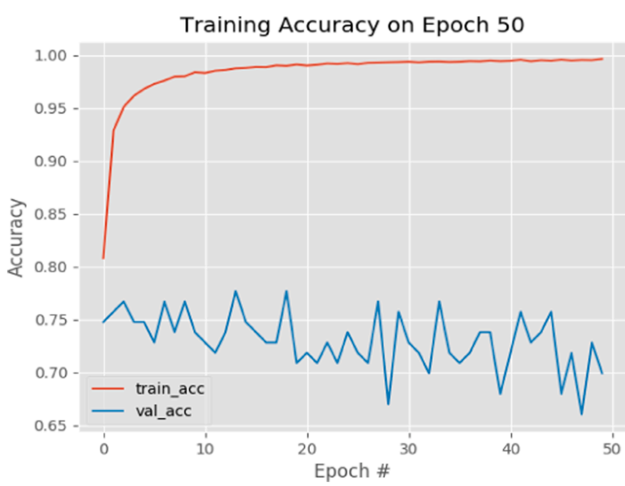



**Fig. 10.** Training process for version 1.

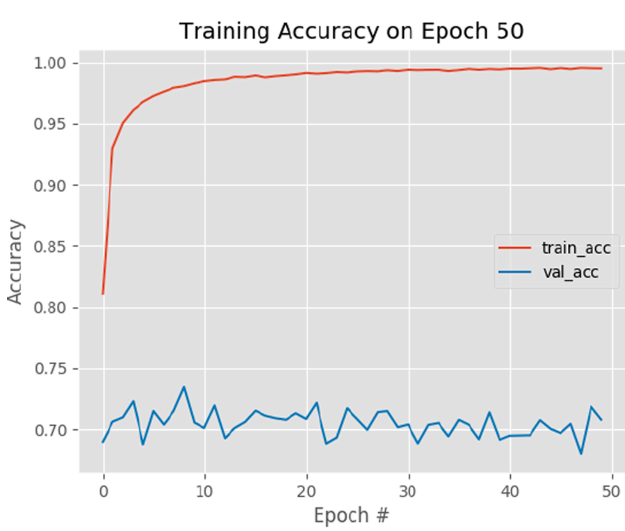



**Fig. 11.** Training process for version 2.

represent the original dataset without data augmentation. The data augmentation process is carried out using the framework presented in Fig. 2, and consists of noise addition, blur, contrast and affine transformation operation. The overall distribution of the dataset is presented in Fig. 7.

- Version 2: The dataset consists of 62.047 images in total. It consists of 46.500 training images as a result of the augmentation process of 510 original training images. The validation data consists of 15.450 images as a result of the augmentation for the original validation data. We keep the testing dataset without any modification to keep the original data from the field. The distribution of the dataset is presented in Fig. 8.

**Table 2**
Performance of the proposed network model for 16-class Mango pest classification.

| Experiment | Validation Accuracy | Testing Accuracy |
|---|---|---|
| Version 0 | 70% | 67% |
| Version 1 | 75% | 68% |
| Version 2 | 71% | 74% |

### 4.2. Experiment result

The training of the VGG-16 network with FCL network layer is presented in this section. The training of the FCL network layer is carried out for 50 epochs. The graphical representation of the training accuracy and the validation accuracy for each experimental run is presented in Fig. 9, Fig. 10 and Fig. 11 respectively. For the training of version 0, the training accuracy reaches a saturation at approximately 18 epochs, while the validation accuracy oscillates between 0.63 and 0.70 starting from 15 epochs.

Following the training of the version 0, the experimental setup of version 1, with training performed with the augmented data is presented in Fig. 10. As opposed to the use of validation data from the augmentation process, the experimental run uses the original data and ground truth to evaluate the training outcomes. As expected, the validation results have not been observed to be stabilised, with each the progression of each epochs. Despite the presence of the original data within the augmented dataset, the validation accuracy oscillates between 0.75 and 0.70 with a global minimum achieved at epoch 47 resulting in less than 67% accuracy. On the other hand, the training accuracy saturates around epoch 15, resulting in the zero slope for the training accuracy plot.

Finally, the training of the third experimental run is presented in Fig. 11. The validation model uses the augmented dataset like the ones used in the training of the FCL network with the VGG-16 network features extracted from the pre-trained model. The training outcomes resemble the version 0 model, which also uses a homogenous data sources for the training and validation process. The oscillations in the accuracy of validation dataset is minimal and is approximately 0.70 with 0.3 tolerance. The training accuracy saturates around 15 epochs, like the version 0 and version 1 training models.

### 4.3. Experimental results

Following the training of the proposed network architecture with VGG-16 and FCL, the overall performance of the network upon testing is formulated in Table 2. Based on the experiment, it is shown that the version 2 experiments which uses augmented images for training, validation and testing yields the best performance. The improvement in the accuracy of the experiment is attributed to the overall learning distinguishability of features learnt by the network. Both validation and testing accuracies represent how well the model can generalize or predict an unseen data. As opposed to the use of validation data, the testing data represents the images captured from the real-world data. The comparative performance of validation accuracy in version 2 experiments showcases the importance of data augmentation process in training the network model. As mentioned earlier, the version 2

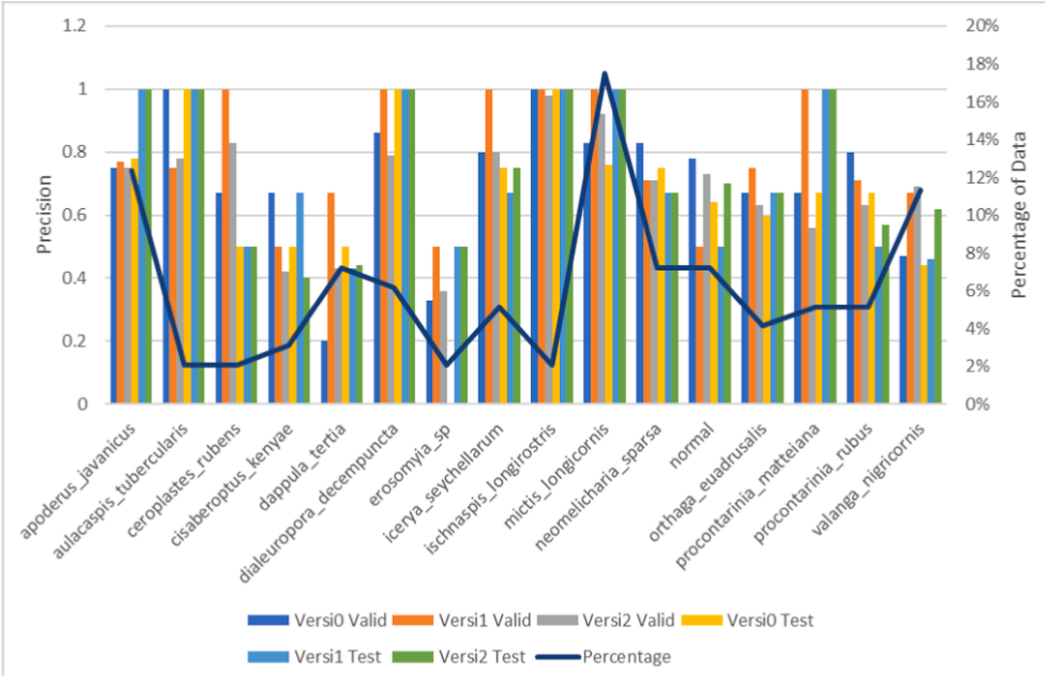



**Fig. 12.** Precision of the proposed network architecture for version 0, version 1 and version 2 trained network.

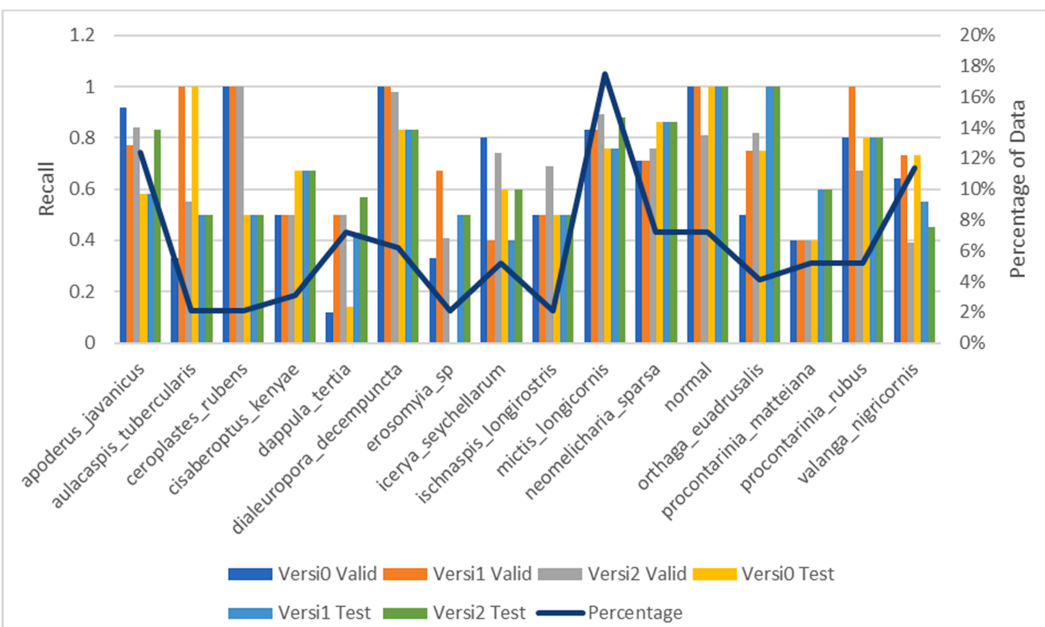



**Fig. 13.** Recall of the proposed network architecture for version 0, version 1 and version 2 trained network.

experiment uses the augmented image sequences for both training and validation, while the version 0 and version 1 experiment run rely on the use of original data for the validation. We attribute the increase in performance to the generalisability of the proposed network for learning distinguishable features available through the data augmentation process, which was not available in version 0 and version 1 runs.

In addition, the overall comparison of version 0, version 1 and version 2 experimental results for precision, recall and F1-measure is presented in Figs. 12 and 13 and 14. The overall training of the VGG-16 based proposed FCL network was carried out using the Titan V GPU with 12 GB Memory, capable of processing 640 of tensor cores. The training server was operated with 9th Generation Intel processor 9900 K consisting of 8 cores. The training model was stored on the SSD M.2 hard disk. The overall training period for each of the experimental dataset version 0, version 1 and version 2 is 74 s, 2,175 s and 7, 116 s respectively.

In addition to the performance assessment of the classification framework on the cumulative outcome of the proposed data augmentation framework, a detailed assessment of each data augmentation function is carried out. The objective of this evaluation is to evaluate the overall contribution of the selected data augmentation function namely noise, blur, contrast and affine transform towards improving the classification performance of the proposed framework. To achieve this objective, a set of 18 different data sets were generated, with an

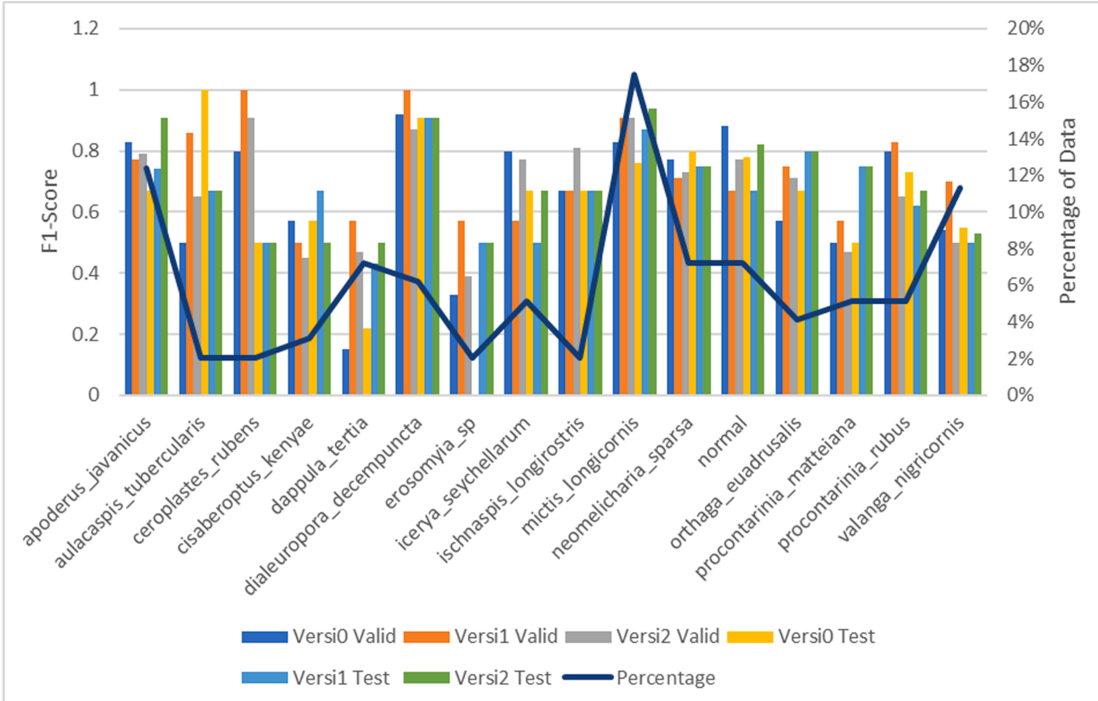



**Fig. 14.** F1-Measure of the proposed network architecture for version 0, version 1 and version 2 trained network.

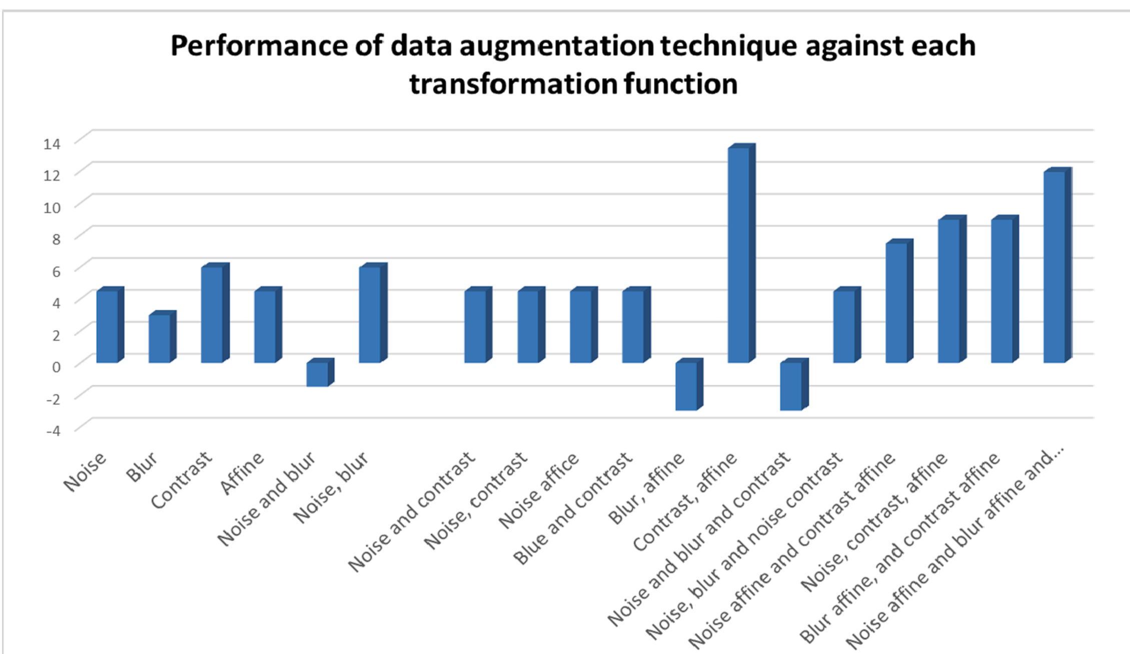



**Fig. 15.** F1 measure for blur, noise and contrast evaluation of data augmentation process.

exhaustive combination of all four data augmentation functions. Each dataset set has been processed through the classification framework with the training of the network carried out using the data augmented images. The testing images used for the evaluation remain the same as version 2 dataset. The proposed scenario achieved the best performance at 76% of accuracy on contrast and affine transformation experiments. In Fig. 15, a percentage comparison of each data augmentation function against the classification accuracy is evaluated against the baseline classification performance when applied without the data augmentation process. The analysis of results highlights the disproportional influence of the data augmentation functions in enhancing the classification performance. The application of contrast and affine transform results in an overall improvement of 13.43% in the classification accuracy. However, the application of noise and blur has resulted in the decrease of classification accuracy by 1.49%. Similar decrease in classification accuracy of 2.98% is also noted for the application of blur and affine transforms along with the application of noise, blur and contrast. The detailed evaluation of the result indicates the positive outcome of the two data augmentation functions namely contrast and affine, which has led to the overall improvement of the classification accuracy. Furthermore, the

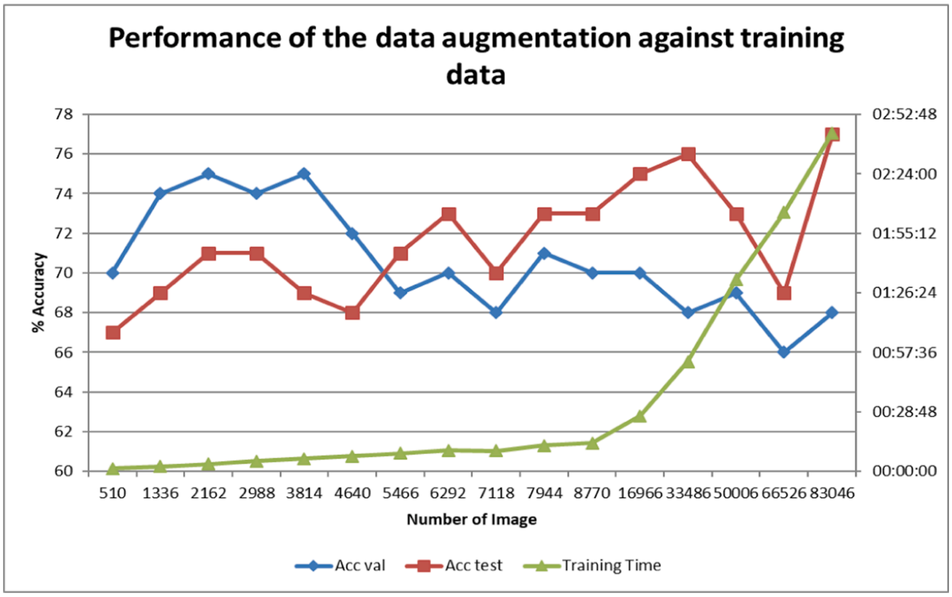



**Fig. 16.** Performance of data augmentation against training data.

application of the contrast function upon the images are limited to the scope of natural light in which the images are expected to be captured. This is achieved by the use of multiplication and addition transformation function consisting in total of 12 filters with α ranging from 1 to 1.5 and β ranging between 0 and 5 was applied on the original image. Similarly, the application of the affine geometric transform is carried out using three dimensional rotational across x, y and z axis. The rotational transform is applied throughout the 360 degrees in both y and z axis. The final dataset has been filtered for transpose images as they do not add value to valuable learning features. The implementation of the data augmentation functions was carried out using OpenCV Library (OpenCV Library).

Following the determination of disproportional influence of data augmentation function upon the multi-class pest classification framework, an additional experiment has been carried out that maps the classification performance against the quantity of the training data. As noted in the previous experiment, the application of contrast and affine transform across y- and z-axis has yielded an improvement of 13.43% improvement against the baseline evaluation without the use of data augmentation framework. Thus, in this experiment, our aim is to evaluate the quantity of the training data required to achieve the best performance in the multi-class pest classification. Therefore, to achieve this objective, the angle of rotation in y- and z-axis is systematically carried out for each of the 310 images from the training dataset. The training data was generated progressively by systematic variation in the number of rotations applied across two axes with a maximum of 100 different

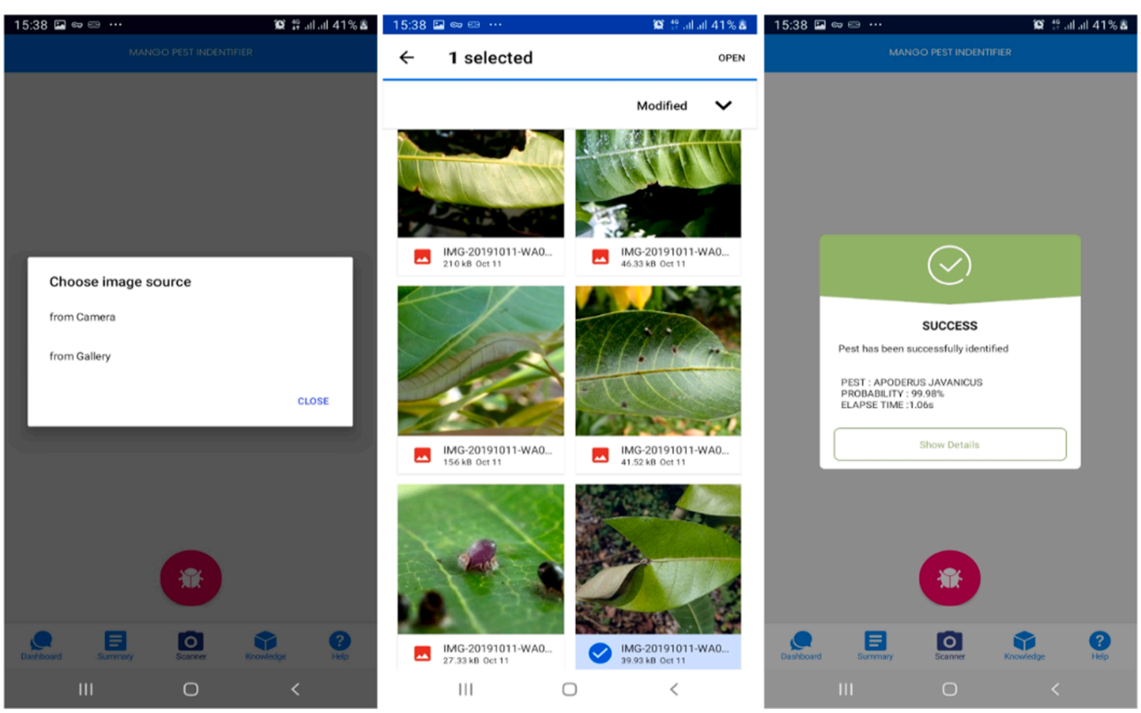



**Fig. 17.** Mobile (Android) application screen as used by Mango Farmers.

**Fig. 18.** Segmented image regions consisting of pests.

combinations applied across 360 degrees. For each of the transformation, the resultant dataset was filtered against transpose images as they do not add any additional value. The F1-measure for each training dataset is presented in Fig. 16. The classification performance saturates at 77%. Following the exhaustive list of 100 different angle variations applied on both y- and x-axes, further increase in the training data has saturated the performance of the multi-class pest classification framework. The graphical model represents the variations of the pest classifier which peaks at 83,046 training samples. The computational time required for each dataset is also presented along the z-axis.

## 5. Summary discussion and future work

Following the results presented, Orysomia sp has resulted in 0% precision and recall on the original dataset testing (version 0). However, the classification performance of the said class improves upon the application of the augmentation process, as verified in version 1 and version 2 runs of the experiments. Similarly, the overall classification of the multi-class pests has shown to have improved by the proposed VGG-16 based FCL network model, upon the application of the augmentation process. The overall multi-class classification of 16-class classification model has yielded in an increase of 13.43% accuracy in comparison to the baseline performance of the proposed approach. The increase in the performance is intuitively attributed to the use of data augmentation process with the combination of contrast and affine transform. Although in the literature, there are reports of using data augmentation process for improving the classification performance from 0.2% to 4.6% (Kobayashi, Tsuji, & Noto, 2019), to the best of our knowledge, such techniques have not been reported for multi-class pest classification. The process of the augmentation as proposed with the cascaded approach of geometric transformations of the pest infected data has yielded improved the robustness of the training model. Following the evaluation of the various models, the best performing model has been selected for the integration into the mobile platform, which has been used by the Mango farmers in Indonesia for the identification of pest categories. The implementation of the various user interactive screens has been presented in Fig. 17. In order to reduce the operational cost of the overall proposed framework, the evaluation of the testing framework has been achieved on the CPU system (with the training carried out on the GPU units). The cloud-based deployment of the CPU-only Tensorflow with Keras library has resulted in the computational time between 2 s to 2.99 s for the classification of the input image and provide a response to the mobile application.

Following the review of the results, an extended approach is also being considered for enhancing the quality of the data augmentation process. In this approach, the appearance of the pest regions is segmented as foreground along with the structural deformity experienced by the leaves. The segmented regions of the pest are subjected to data augmentation framework presented in the paper. The training of the deep-learning network is carried out with the superimposed augmented pest images as foreground against the naturally appearing background regions. An initial outcome of such an approach is presented in Fig. 18. The training process to be carried out on such dataset will facilitate the deep-learning models to distinguish between the foreground pest and the background images, thus leading to the improvement in performance accuracy. In addition, we will also evaluate the performance of the network training against the overfitting as presented in Figs. 9–11. The approach will be further invested as a part of our ongoing research activity.

## 6. Conclusion

In this paper, three contributions have been presented. The study of augmentation process for increasing the limited availability of pest infected dataset complemented by the proposed architecture for the training of multi-class pest classification model. The proposed classification framework extends the VGG-16 framework and extracted the deep-learning features from the network. The extracted features are further trained through a 2-layer fully connected network for achieving the classification outcome. The systematic evaluation of the proposed approaches was achieved based on three different datasets. The combination of these datasets includes the classification on the original data without augmentation process resulting in 67% of the overall accuracy, while the evaluation on the augmented data process has resulted in 76% overall accuracy with the application of contrast and affine transform. The additional experimental results also indicate the impact of different data augmentation functions on the classification performance. The final contribution provides an outline of the overall data workflow process integrated as a mobile application that can be directly used by the Mango farmers in Indonesia. The future work will include a detailed analysis of the deep-learning features extracted from multiple-pre-trained models and evaluate the quality of these features. In addition, the network architecture models will be further developed towards improving the performance of the classification model.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The research activity leading to the publication has been partially funded by the Ministry of Research, Technology and Higher Education Indonesia. Also, the authors would like to thank NVIDIA Corporation for the donation of the GPU used in this study.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2020.105842.

## References

Abou-Awad, B., Metwally, A., Al–Azzazy, M., 2009. Ecological, biological and control studies on the leaf coating and webbing mite cisaberoptus kenyae Keifer (Eriophyoidea: Eriophyidae) in Egypt. Acarines: J. Egypt. Soc. Acarol. 3 (1), 65–71.
Australian Government Department of Agriculture and Water Resources. 2015.
Avendano, J., Ramos, P.J., Prieto, F.A., 2017. A system for classifying vegetative structures on coffee branches based on videos recorded in the field by a mobile device. Expert Syst. Appl. 88, 178–192.
Chollet François, 2015. Keras: The Python Deep Learning library. Keras.Io. https://doi.org/10.1086/316861.
Chang, Mingshan, Luo, Ji, Wu, Yaojun, Wen, Juan, 2018. The occurrence rule and control technology of Dappula tertia Templeton, a pest of Eucalyptus In Jean-Paul Laclau. *Managing Eucalyptus plantations under global changes*. Le Corum, Montpellier-France.
Das, Medhi, Karsh, Laskar, 2016. Image Splicing Detection using Gaussian or Defocus Blur. 2016 International Conference on Communication and Signal Processing (ICCSP) 1237–1241. https://doi.org/10.1109/ICCSP.2016.7754350.

Ha, T.M., 2014. A review on the development of integrated pest management and its integration in modern agriculture. Asian J. Agric. Food Sci.

Johannes, A., Seitz, M., Se, B., 2017. Automatic plant disease diagnosis using mobile capture devices. Land.Technik AgEng 138, 200–209.

Kobayashi, K., Tsuji, J., Noto, M., 2019. Evaluation of Data Augmentation for Image-Based Plant-Disease Detection. In: Proceedings - 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018, 2206–2211. https://doi.org/10.1109/SMC.2018.00379.

Lu, Y., Yi, S., Zeng, N., Liu, Y., Zhang, Y., 2017. Identification of rice diseases using deep convolutional neural networks. Neurocomputing 267, 378–384.

Manjunath, J., 2018. Bio-ecology of Mango Leaf Twisting Weevil (Apoderus transquebaricus). International Journal of Pure & Applied Bioscience, 6(6), 375–382. https://doi.org/10.18782/2320-7051.6642.

Mohanty, S.P., Hughes, D., Salathé, M., 2016. Using Deep Learning for Image-Based Plant Disease Detection.

Mohamed S., Ghada, 2015. POPULATION DYNAMICS OF THE SEYCHELLARUM MEALYBUG, Icerya seychellarum (WESTWOOD) (HEMIPTERA: MARGARODIDAE) ON THE ORNAMENTAL PLANT. Journal of Plant Protection and Pathology 6 (3), 481–498. https://doi.org/10.21608/jppp.2015.53310.

OpenCV Website. 2014.

Paschali, M., Simson, W., Roy, A.G., Göbl, R., Wachinger, C., Navab, N., 2019. Manifold exploring data augmentation with geometric transformations for increased performance and robustness. Lect. Notes Comput. Sci. (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-030-20351-1_40.

Salem, H.A., Mahmoud, Y.A., Ebadah, I.M.A., 2015. Seasonal abundance, number of generations and associated injuries of the white mango scale, Aulacaspis tubercularis (Mangifera) (Newstead) (Homoptera: Diaspididae) attacking mango orchards. Res. J. Pharm. Biol. Chem. Sci. 6 (4), 1373–1379.

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. J. Big. Data 6 (1). https://doi.org/10.1186/s40537-019-0197-0.

Simonyan, K., Zisserman, A., 2014. A8_ Large-Scale Image Recognition. Very Deep Convolutional Networks for Large-Scale Image Recognition. https://doi.org/10.2146/ajhp170251.

Singh, R.N., Maheshwari, M., Saratchandra, B., 2005. Biocoenology and control of whiteflies in sericulture. Insect Sci., 12(6), 401–412. https://doi.org/10.1111/j.1744-7917.2005.00051.x.

Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., Stefanovic, D., 2016. Deep neural networks based recognition of plant diseases by leaf image classification. Computat. Intell. Neurosci. 2016, 1–11.

Strange, R.N., Scott, P.R., 2005. Plant disease: a threat to global food security. Annu. Rev. Phytopathol. 43 (1), 83–116.

Suputa, Cahyani, Kustaryati, A., Hasyim, A., Hasanah, I. U., Ratnaningrum, A. C., … Ma'Rufah, A. A. (2015). Pedoman Pengenalan dan Pengendalian Organisme Pengganggu Tumbuhan pada Tanaman Mangga (2nd ed.). Jakarta: Direktorat Perlindungan Tanaman Hortikultura.

Szeliski, R., 2010. Computer Vision: Algorithms and Applications. 1st ed. Heidelberg: Springer-Verlag.

Zhang, S., Wu, X., You, Z., & Zhang, L., 2017. Leaf image based cucumber disease recognition using sparse representation classification. Comput. Electron. Agric., 134, 135–141. https://doi.org/10.1016/j.compag.2017.01.014.