SEMARAK ILMU PUBLISHING

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Chlorophyll Intensity Forecasting in Diverse Plant Species: Advancing with Enhanced Convolutional Neural Networks

Annie Anak Joseph[1,*], Michelle Nashrin Bawai[1], Charlie Sia Chin Voon[2], Kho Lee Chin[1], Kuryati Kipli[1], Ngu Sze Song[1]

[1] Department of Electrical and Electronics Engineering, Faculty of Engineering, Universiti Malaysia Sarawak (UNIMAS), Kota Samarahan, Sarawak, Malaysia
[2] Faculty of Engineering, Computing and Science, Swinburne University of Technology Sarawak Campus, 93350 Kuching, Sarawak, Malaysia

ARTICLE INFO

Article history:
Received 27 November 2026
Received in revised form 20 February 2026
Accepted 15 April 2026
Available online 4 May 2026

Keywords:
Chlorophyll intensity; convolutional neural network (CNN); prediction; leaf spectra; hybrid; mean square error (MSE)

ABSTRACT

Chlorophyll pigment plays a crucial role in photosynthesis by absorbing and harnessing light energy, ultimately supporting the plant's overall health and growth by providing essential nutrients. A convenient assessment of chlorophyll content is essential in smart management agriculture. It is very important to measure chlorophyll intensity in a leaf accurately as it can indicate the plant health status in terms of its growth, photosynthetic capacity, and nutritional stress. Several attempts have been made to implements computer vision to enhance the precision agriculture techniques However, recent literatures about chlorophyll intensity prediction of various plant species using non-destructive methods are limited and yet existing methods for chlorophyll intensity forecasting are time-consuming. Hence, a rapid and straightforward convolutional neural network (CNN) algorithm was proposed to predict chlorophyll intensity of various plant species based on leaf reflectance spectra. The datasets were taken from ANGERS Leaf Optical Properties Database (2003). The proposed model consists of Hybrid CNN as a feature extractor and Support Vector Regression (SVR) network as a predictor. Hybrid CNN was designed by modifying the architectures of AlexNet and PNet. The performance of Hybrid CNN with SVR (CNN-SVR) was also compared with AlexNet, PNet, and SVR. Hybrid CNN exhibits superior performance when compared to established models, including AlexNet, PNet, and standalone SVR. Notably, the mean square error values for training and testing datasets stand at 0.1558 and 1.149, respectively. This substantiates the model's efficacy in predicting chlorophyll intensity accurately, underscoring its potential utility in advancing precision agriculture techniques.

1. Introduction

Every plant contains vital green pigments known as chlorophyll which can indicate the concentration of the elements for the synthesis of chlorophyll such as nitrogen, iron, and magnesium [1][12]. Chlorophyll is beneficial during photosynthesis to absorb sufficient light energy and provide

* *Corresponding author.*
E-mail address: jannie@unimas.my

https://doi.org/10.37934/araset.59.5.275290

nutrients to the plant. Hence, the evaluation of crop diseases and yield predictions relies significantly on the quantification of green pigments in the leaves [2]. The traditional method of monitoring plants using the naked eye is still inadequate to predict their health based on the greenness of leaves. This is due to the significant downsides that are very challenging, especially for the farmers who have a less basic understanding of the nutritional deficiency of the plants and diagnosing the disease symptoms [3]. Therefore, the farmers require an expertise in this field is to help them to monitor the plant health status by quantifying the chlorophyll intensity based on the visual information of the leaves. Various methods have been proposed previously to predict the intensity of chlorophyll based on the plant leaves. Conventional methods for chlorophyll intensity prediction are either destructive or non-destructive [2]. For the destructive method, the pigments are crushed and extracted from the leaf sample using solvent extraction and subsequent spectroscopic chlorophyll analysis. However, this method needs cautious disposal of the extraction solvent waste, the high number of samples required, and the estimation parameters are restricted and time consuming [4]. Meanwhile, for the non-destructive method, the chlorophyll intensity is quantified by using specialized devices such as chlorophyll fluorescence [4], spectrometer [5], and Soil Plant Analysis Development (SPAD) [6]. These methods are categorized as conventional non-destructive measurement devices. The measurement of chlorophyll intensity using these conventional methods are mainly done by analyzing the leaf surface without destroying the samples. Nonetheless, the devices required large samples to maintain culture sterility [4], as well as are still costly and difficult in generating reflectance spectrum data for both personal and commercial purposes. Hence, a straightforward CNN algorithm should be designed in order to help the agriculturists, researchers, and farmers' burden in predicting chlorophyll intensity of various plant species. Low-cost algorithm design is essential because training and testing based on reflectance spectrum as input data can be completed more quickly with a simpler deep learning architecture. The accuracy of the CNN algorithm in producing astute prediction values of the chlorophyll intensity should be the primary consideration during its development, rather than just its computational cost and time consumption. This is because the greater the accuracy of the CNN algorithm in training and testing process, the nearest the prediction of chlorophyll intensity to the actual content in a leaf. Consequently, a deep learning model with a hybrid architecture that combines CNN and SVR has been created to tackle the issue, with the particular goal of significantly impacting the agricultural field. The rest of the article is structured as follows. Section 2 explains the related works. Section 3 discuss the Methodology. Section 4 discusses the results and discussion. Section 5 is dedicated to the conclusion.

2. Related Works

There are two methods for determining the amount of chlorophyll in leaf samples: destructive and non-destructive methods. Both techniques typically use the absorption rate of the red spectral regions between 600 and 700 nm to calculate the chlorophyll intensity. Therefore, the red region is suitable for reflectance and transmittance spectra measurement [7]. Destructive measurement method is a laboratory analysis approach performed with organic solvents to extract the leaf samples collected in the crops. Destructive method can determine the most accurate results for the prediction of chlorophyll intensity. Unfortunately, this method is not environmentally friendly, slow procedure, expensive, and tiresome [7]. Due to this reason, non-destructive measurement method is an alternative approach as it is environmentally friendly, affordable for commercial purposes and less time-consuming. The most common approaches of measurement technique used for laboratory controlled procedures and research purpose of chlorophyll intensity are SPAD Chlorophyll Meter and spectrophotometer.

2.1 Soil Plant Analysis Development (SPAD) Chlorophyll Meter

SPAD chlorophyll meter is widely used for nondestructive approach to measure chlorophyll intensity in the plant tissue in vivo instantly [8]. The most common SPAD chlorophyll meter used in the agricultural research purpose is SPAD-502 (Konica-Minolta, Inc, Osaka, Japan). This handheld portable instrument is easy to use and fast in generating the measurement results [7]. Generally, the process of measuring chlorophyll intensity using SPAD-502 is started by placing a leaf on the measurement area. Then, both chlorophyll and non-chlorophyll are measured by using two light-emitting diode (LED) light sources which are implemented in the meter. A LED light source is specified the chlorophyll absorption peak of 650 nm while the other LED is centred at the absorption peak of 940 nm for non-chlorophyll content. The red transmitted light is emitted through the leaf sequentially using two LED light sources in the SPAD meter. Two silicon photodiode detectors are used to measure the leaf's infrared regions. The LED light sources with wavelengths of 650 and 940 nm are detected by the detectors. Later, in order to enable the microprocessor to detect the current, the light sources are converted into an electric current. Hence, the signal from the electric current is modified into SPAD readings of chlorophyll intensity [7]. The readings are converted into the absolute chlorophyll readings measured in the labs. The digital number of the output readings on SPAD chlorophyll meter and the lab measured chlorophyll intensity results for the same leaf samples are used as the variables for the conversion relationship [9]. The intensity of chlorophyll content increases with the amount of green pigments present in the leaf.

2.2 Spectrophotometer

Spectrophotometer is a non-destructive instrument that can measure the chlorophyll intensity in plant tissues in vivo [8] along with the digital image acquisition. According to [8], a leaf sample is laid to light from tungsten halogen, a light source generator through a probe. Tungsten halogen provides a large electromagnetic wavelength range of 360nm to 2400nm. The probe will snap the reflected light simultaneously and send it to the spectrometer. Later, the spectrometer measures the chlorophyll intensity and submit the results to the computer for visualization and quantification.

2.3 Random Forest

Random Forest (RF) is a machine learning algorithm that has been widely used in predicting leaf chlorophyll content (LCC) in various vegetation types. A thorough understanding of decision trees is also necessary in order to comprehend the RF working principle. A tree-like structure is used to represent the predictive model in decision trees [10]. Decision trees are supervised statistical models that are used to estimate the value of the target variable, in this case LCC, by using input features, or the predictor, in prediction tasks. In RF, each regression tree is grown from the root node to the terminal nodes which depending on the prediction conditions. Meanwhile, internal nodes which located between the beginning and ending nodes are considered as the decision points. To begin the tree's growth, a random subset of the larger training set's data is selected. The term "bootstrapping" refers to this method of expanding the forest's diversity [10]. Therefore, after a bootstrap sample is generated, an algorithm fits a prediction tree.

2.4 Support Vector Regression

A supervised machine learning method called Support Vector Regression (SVR) network was developed from the Support Vector Machine (SVM) method especially for regression or prediction tasks. The SVR algorithm is used to solve the regression problem by projecting the kernel functions for the training data into feature space. The sigmoid, polynomial, linear, and radial basis function (RBF) kernels are the most often utilized ones [11]. There are variety of deep learning models that had been demonstrated for their competency in predicting photosynthesis pigments such as anthocyanin, carotenoid, and chlorophyll [12-16]. Nevertheless, it is crucial to defend that each model comes with its own set of limitations. These limitations are varied from a model to another model which consist of its accuracy, architecture complexity, generalization to various plant species, and ability in capturing a specific pigment. For instance, the performance of AlexNet was consistent during training session either with large or limited amounts of datasets as it provided lower value of MSE. However, this model took longer time to complete its training session of predicting chlorophyll intensity [16]. Meanwhile for PNet, it had the lower risk of overfitting issue while having small value of mean absolute error (MAE), but in the case of predicting chlorophyll intensity only, PNet had less ability to focus more on this pigment. This was because chlorophyll content was overlapping with pigments [16]. VGGNet and ShallowNet were both less efficient in predicting the pigment as the models had the complex and simplest, respectively. This is due to the fact that a simple architecture may result in the algorithm being unable to correctly learn the training set, whereas a complex architecture carries a significant risk of overfitting issues. An enhanced CNN is created by merging two distinct architectures for the purpose of feature extraction after shortcomings in each model are identified. Hybrid CNN, an upgraded version of CNN with the hybrid of both AlexNet and PNet, is still thought to be a good architecture. SVR is the ideal network for this task, according to the readings, and it reduces the dimensionality of the data from the CNN model. Because of this, the strengths of both CNN and SVR are combined to improve the accuracy of predicting the chlorophyll intensity of different plants based on leaf spectra.

3. Methodology

In this section, MATLAB R2023a is used as a main software tool. The specifications of the workstation are stated in Table 1.

Table 1
Specifications of the Workstation Used

Specification	Description
Processor	AMD Ryzen 5 7520U
Clock Speed	2.8 – 4.3 GHz
GPU	AMD Radeon Graphics™
RAM	8GB
SSD	512GB
OS	Window 11 Home

The network system consists of four main phases which are dataset preparation, leaf spectra analysis, feature extraction, and prediction task.

3.1 Dataset Preparation

The preparation of dataset consists of actual measurement of chlorophyll intensity using conventional non-destructive technique and dataset acquisition of leaf spectra, which both were taken from ANGERS Leaf Optical Properties Database (2003). There were 552 datasets of leaf spectra in total which consisted of 43 different plant species. To make it easier to identify the plant species, each dataset was organized into a folder by name. Thirty percent of the datasets are applied as testing set to evaluate the performance of the trained model, and seventy percent of the datasets are used as training set to identify patterns and relationships between the features and target variable.

3.2 Leaf Spectra Analysis

Generally, within the wavelength range of 400 nm until 2450 nm of a leaf spectrum, it consists of information includes plant leaf physiological and chemical properties. Each leaf spectrum is divided into three regions: visible (VIS) (400–700 nm), near infrared (NIR) (700–1100 nm), and shortwave infrared (SWIR) (1100–2500 nm) on both ends of the spectrum. Chlorophyll and other photosynthetic pigments absorb heavily in the visible spectrum (VIS), according to [17]. In order to concentrate on the chlorophyll-specific spectra solely for the purpose of predicting chlorophyll intensity, all leaf reflectance spectra are analysed.

Initially, the wavelength range of each leaf reflectance spectrum of ANGERS was 400 nm to 2450 nm. The three distinct regions of plant leaf physiological and chemical properties in the ANGERS datasets are similar to those in [17]. Figure 1 presents the plant information regions. The VIS region of the ANGERS dataset has the highest chlorophyll content when compared to other regions, falling between 400 and 750 nm.

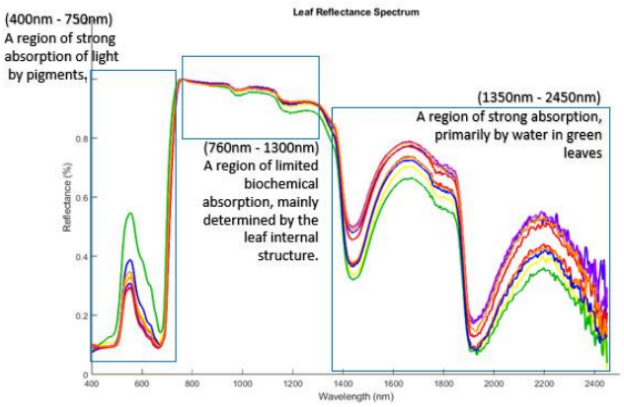

Fig. 1. Regions of VIS-NIR-SWIR of ANGERS Dataset

The leaf spectra of different plant species were measured in the 400–750 nm wavelength range. For this reason, the wavelength and reflectance values within the chlorophyll specific region are the only ones used for analysis and modification of each TXT file. Following that, two sets of leaf reflectance spectra were created: a training set (30%) and a testing set (70%). The feature extractor model used the resulting spectra as input data.

3.3 Feature Extraction

This is the most important stage because it involves carefully modifying the CNN model's architecture to enable each layer's parameters to match the leaf spectra that are being used as input data. A hybrid CNN is proposed to improve the prediction accuracy of chlorophyll intensity with less computational cost and learning time. The relevant information is then extracted from the leaf reflectance spectrum using a hybrid model.

3.4 Hybrid Convolution Neural Network

Two CNN models, AlexNet and PNet are hybrid to form an improved model. The stack diagrams for the PNet (b) and AlexNet (a) architectures are displayed in Figure 2. Originally intended for digital images found in 3D data. Meanwhile, PNet was designed by [16] for reflectance spectrum which is in 1D data. In this paper, the parameters in AlexNet was modified in order to be embedded into PNet. Based on these two models, a hybrid CNN for the 1D leaf reflectance spectrum was created.

The foundation of Hybrid CNN is AlexNet since it is a versatile and easily adjustable model that works with more than just image datasets. Meanwhile, PNet was also adopted for Hybrid CNN because of its simple algorithm that predicts plant pigments and yields the lowest mean square error values. As a result, using these models made it easier to extract important chlorophyll features quickly.

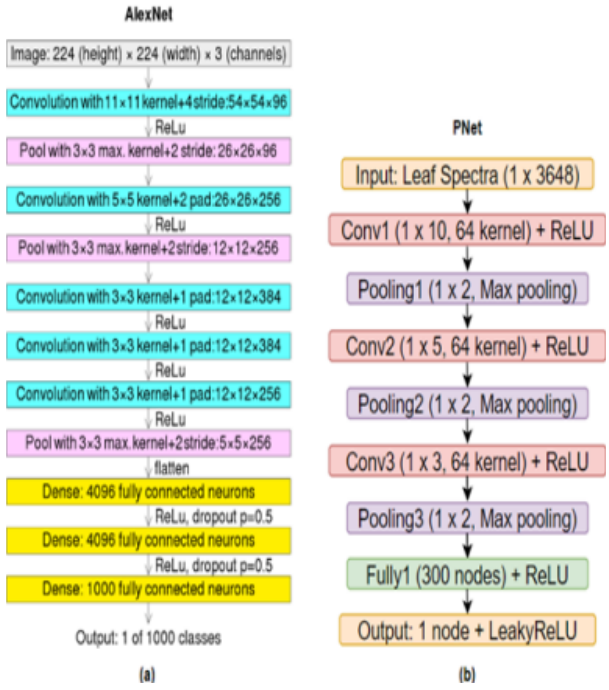

Fig. 2. (a) AlexNet Architecture (b) PNet Architecture

The AlexNet model's stack diagram, shown in Figure 2 (a), consists of one input layer, three hidden layers, three max pooling layers, three fully connected layers, five convolutional layers and one output layer. Each layer has parameters that can be changed. Consequently, a number of the AlexNet layer parameters are changed in order to incorporate PNet and the 1D algorithm. Four modification keys in AlexNet are taken by the parameters. Table 2 presents a tabular comparison of the parameters of the original and modified versions of AlexNet.

Table 2
Comparison of Parameters for Original Alexnet And Modified Alexnet

AlexNet	Original	Modified AlexNet		
Layers	Size	Activation Function	Size	-
Input	227 x 227 x 3	-	1 x 351	Activation Function
Conv1	11 x 11	ReLU	1 x 11	-
Pooling1	2 x 2	-	1 x 2	ReLU
Conv2	5 x 5	ReLU	1 x 5	-
Pooling2	2 x 2	-	1 x 2	ReLU
Conv3	3 x 3	ReLU	1 x 3	-
Conv4	3 x 3	ReLU	1 x 3	ReLU
Conv5	3 x 3	ReLU	1 x 3	ReLU
Pooling3	2 x 2	-	1 x 2	ReLU
Fully1	1 x 4096	ReLU	1 x 4096	-
Fully2	1 x 4096	ReLU	1 x 4096	ReLU
Fully3	1 x 1000	SoftMax	1 x 1	ReLU

1D leaf reflectance spectrum is applied as an input instead of 3D digital image. The wavelength range of leaf reflectance spectrum of chlorophyll-specific region is within 400 nm to 750 nm. Kernel size for all convolutional layers, pooling layers and fully connected layers are adjusted to 1 x n instead of n x n. For instance, 11 x 11 is a kernel size for the first convolutional layer but then, it is adjusted to 1 x 11 which is a series of 1D kernel size. Kernel size of the third fully connected layer is changed to 1 x 1 in order to replaced 1 x 1000. This is so because AlexNet's original purpose was to categorize pictures into 1000 classes. Furthermore, by using a 1 x 1 kernel size, the dimensionality can be reduced from 3D to 1D and the prediction of chlorophyll content can be the exclusive focus. Activation function is now Linear rather than SoftMax. The simple mathematical formulation of Linear is stated in Eq. (1) where $f(x)$ and x are indicating the output of the last fully connected layer the input of the last fully connected layer, respectively. Two convolutional layers of Modified AlexNet along with its pooling layers are replaced by the convolutional layers of PNet and its pooling layers is shown in Figure 3.

$$f(x) = linear = x \tag{1}$$

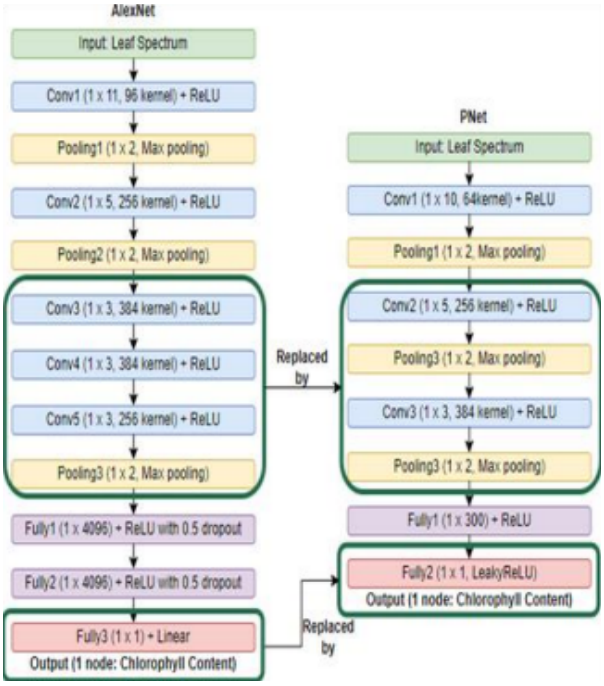

Fig. 3. Customisation of Modified AlexNet and PNet Architecture

As a result of the integration, both architectures became hybrid and presented as depicted in Figure 4. The performance of a CNN model can be affected by the amount of convolution layers. In most studies related to chlorophyll intensity prediction by using CNN, the performance of the models is getting less efficient when the number of convolutional layers is increasing [2,9,16,18]. This is because the features extracted from the leaf reflectance spectrum cannot specify the optimal chlorophyll content. Therefore, a CNN model with simple and straightforward architecture is proposed in this project in order to increase the model's performance efficiency in predicting the chlorophyll content. As the architecture of this Hybrid CNN is simple, it is easy to retune the hyperparameters during its training procedure. Furthermore, every convolutional layer is followed by a pooling layer because it has the ability to lessen the number of parameters. It can be done by minimising the feature dimensional at the output of every convolution layer. The minimisation has encouraged the computational cost to be lower [19-20].

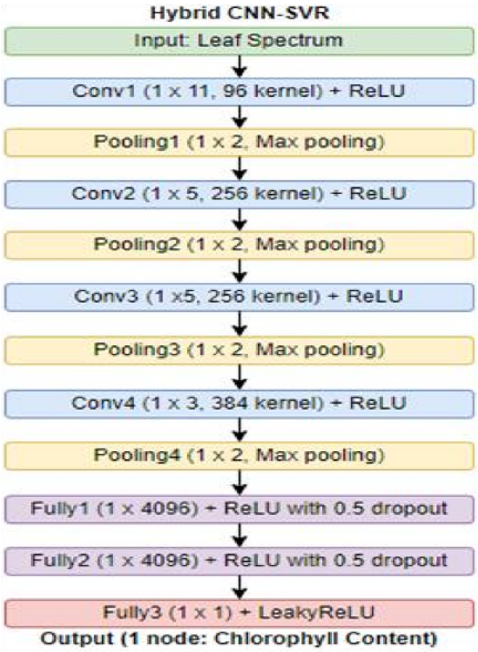

Fig. 4. Stack Diagram of Hybrid CNN Model

One input layer, four convolutional layers, four maxpooling layers with ReLU in each layer, three fully connected layers, and an output layer with one node for particular chlorophyll content features make up the architecture of the suggested model. Fully connected layers will extract and select pertinent features from the leaf reflectance spectrum. This model's output is substituted for an SVR network in the prediction task to produce a prediction of the amount of chlorophyll. Figure 5 illustrates the conceptual diagram of Hybrid CNN and then followed by SVR.

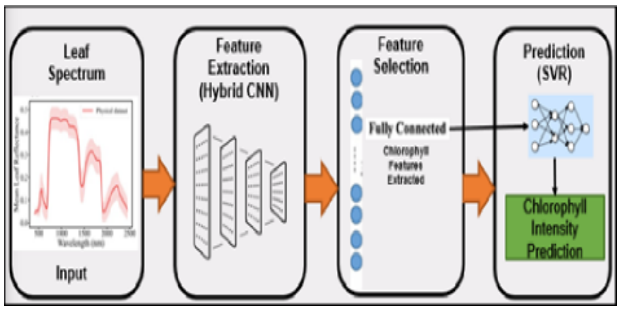

Fig. 5. Conceptual Diagram of Hybrid CNN

The SVR algorithm is employed to forecast the level of chlorophyll present in a nipa palm leaf. The SVR predictor model's overall structure is depicted in Figure 6 as it connects to the final fully connected layer.

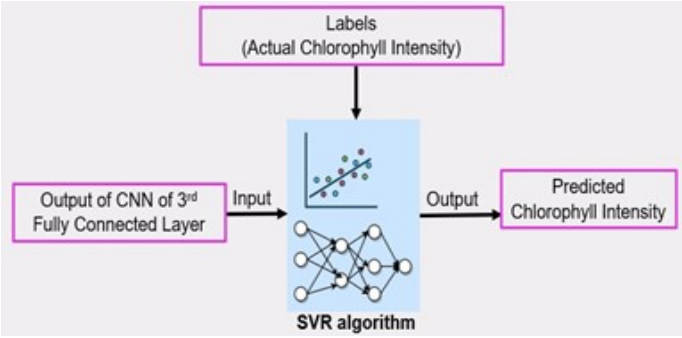

Fig. 6. General Structure of SVR Network

The input of SVR algorithm is obtained from the output of CNN's last fully connected layer. Meanwhile, the label applied is from the results of actual chlorophyll intensity which were taken by using SPAD-502 Plus Chlorophyll Meter. The label refers to the target variable which the algorithm is trying to predict. Using the ANGERS database, the predicted chlorophyll intensity of 43 plant species is obtained by pairing the input and the label. As a result, the prediction may be used to forecast a plant species' level of health. Figure 7 shows the general flowchart of the suggested Chlorophyll Intensity Prediction Model.

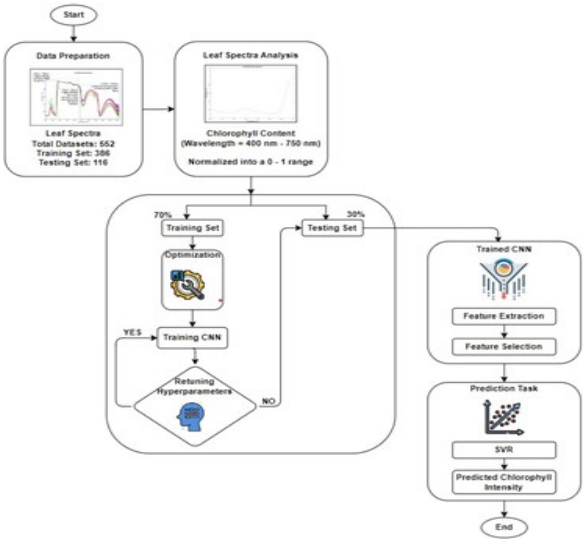

Fig. 7. Flowchart of Proposed Chlorophyll Intensity Prediction Model

4. Results and Discussion

In this paper, Hybrid CNN is trained with 386 leaf spectra to extract the chlorophyll content features. Raw datasets of leaf spectra are transformed into a high-dimensional feature representation to encode the relevant information. After the relevant features are extracted from hybrid CNN, the extracted features representation is passed to SVR network as its input data. This network uses the support vectors and a loss function to determine the suitable regression function that approaches the correlation between the features and chlorophyll intensity. Hybrid CNN is analysed by comparing its performance with three different optimizers including Stochastic Gradient Descent with Momentum (SGDM), Root Mean Squared Propagation (RMSprop), and Adaptive

Moment Estimation (Adam). Other than that, the performance of trained Hybrid CNN-SVR is compared and analyzed with the existing models which are AlexNet, PNet and SVR.

There are four different training settings were executed in order to determine the best hyperparameters for hybrid CNN in extracting the features of relevant chlorophyll content. Hyperparameters are the variables that significantly affect the model's training dynamic and generalization ability. The process of determining the optimal training setting involves experimenting with different values for each hyperparameter, taking into account factors such as computational cost, dataset size, and model complexity to ensure a smooth simulation run. The most significant hyperparameters to be tuned in hybrid CNN are learning rate, number of epochs, and batch-size. Learning rate decides the convergence speed and training behaviour at each iteration. The results of MSE value and training completion time for each training setting as shown in Table 3 were selected by performing the simulation for tenth times and the lowest MSE value was chosen. The optimal training setting was carefully selected based on the results of MSE value and training completion time. Thus, based on Table III, training setting 2 was the optimal settings for the hybrid model where it had Adam as its optimizer, learning rate of 0.01, maximum epochs of 50, validation frequency of 50, and minimum batch size of 55. Training settings 2 took 2 minutes and 1 second to be completed while resulting in 0.087 for its MSE value.

Table 3
Comparison Of Training Setting With Various Hyperparameters

Hyperparameters	Training Setting 1	Training Setting 2	Training Setting 3	Training Setting 4
Optimizer	Adam	Adam	Adam	Adam
Learning rate	0.001	0.01	0.01	0.01
Maximum Epochs	50	50	50	100
Minimum Batch Size	55	55	65	55
Validation Frequency	50	50	50	50
Epoch Completed	50 of 50	50 of 50	50 of 50	100 of 100
Mini-Batch RMSE	4.95	0.79	1.68	0.37
Mini-Batch Loss	12.20	0.30	1.40	0.07
Completion Time	1 min 26 sec	2 min 01 sec	3 min 20 sec	6 min 37 sec
MSE	0.897	0.087	0.313	0.331

The learning rate of 0.01 is appropriate for the proposed model to extract the important features of chlorophyll while allowing the model to converge the data efficiently without getting stuck in local minima. The maximum epochs of 50 also stroked a balance between not overfitting and not underfitting which led to better generalization. Meanwhile, the minimum batch size of 55 enabled more stable updates of the proposed model's parameters while encouraging an efficient convergence. The Mean Squared Error (MSE) holds significance as a metric for evaluating the efficacy of a hybrid Convolutional Neural Network (CNN) in the prediction of chlorophyll intensity. The results provide information about how close the expected and actual values of chlorophyll intensity are, as determined by the SPAD-502 Plus Chlorophyll Meter. The MSE value is more closely aligned with the real value when it was lower. Following the selection of training setting 2 as the optimal configuration with respect to different hyperparameters, the training efficacy of hybrid CNN was further examined using an additional gradient descent-based optimizer to confirm that Adam is the most optimal optimizer for Hybrid CNN. Hessian-based and gradient descent-based optimizers are the two types of optimizers used in deep learning [14]. The gradient descentbased method is used in this project because, compared to the hessian-based method, it has a smaller loss function to the model and offers a relevant training time because of its reasonable computational cost. The comparison of MSE values with various optimizers on Hybrid CNN is displayed in Figure 8.

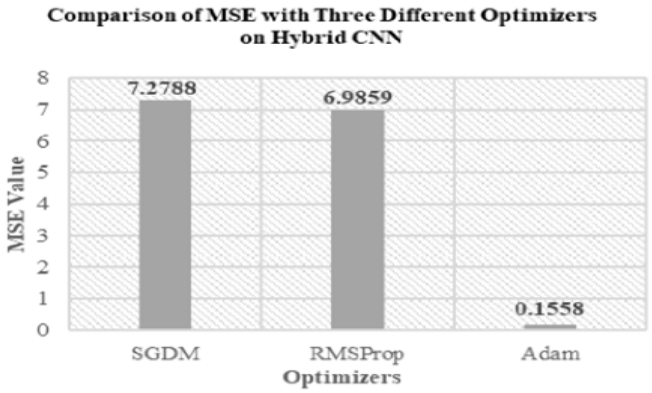

Fig. 8. Comparison of MSE with SGDM, RMSProp, and Adam Optimizers on Hybrid CNN

For the analysis, five smallest MSE values for each optimizer were collected and averaged to determine its performance indicator. Based on Figure 8, the optimizers were sorted from the oldest to the recent. It can be viewed that the different optimizers had significant impacts on the range of MSE values. SGDM and RMSProp optimizers had provided the worst performance as their MSE values were 7.2788 and 6.9859, respectively. Meanwhile for Adam optimizer, its MSE value appeared to be the lowest which was 0.1558 and it indicates the best performance for hybrid CNN among other optimizers. Adam optimizer can contribute to converge faster while having flexible learning rate in handling variety of gradients. As a result, Adam's selection as the best optimizer had an impact on the Hybrid CNN architecture.

Besides that, hybrid CNN-SVR is compared with the existing models including AlexNet, PNet, and SVR. Each model was performed using both training and testing set. In this paper, hybrid CNN-SVR shows the most efficient model in predicting chlorophyll intensity based on its MSE value and training completion time. Figure 9 shows the comparison of MSE for training and testing sets on four different models.

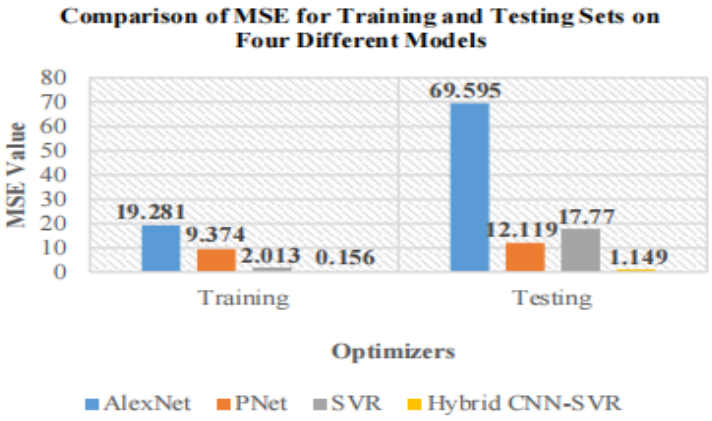

Fig. 9. Comparison of MSE for Training and Testing Sets on AlexNet, PNet, SVR, and Hybrid CNN

AlexNet exhibited the highest MSE values on both training (19.281) and testing sets (69.595), indicating a significant risk of overfitting due to the vast difference in MSE between the sets. PNet

showed less risk of overfitting but still had poor MSE values for a CNN model. SVR also demonstrated a high risk of overfitting based on its MSE values. The hybrid CNNSVR, however, achieved the lowest MSE values for both sets (0.156 for training and 1.149 for testing), indicating superior performance and reduced overfitting risk. AlexNet's complex architecture led to poor generalization, while PNet's simplicity resulted in underperformance due to its limited ability to capture diverse data patterns. Despite SVR's lower MSE with the training set, its overall performance remained subpar due to a high risk of overfitting.

Table 4 shows that PNet finished the training process in the shortest amount of time 20 seconds while AlexNet took the longest 1 hour, 32 minutes, and 36 seconds. This is because each model's complexity in extracting the features of chlorophyll content varies depending on how many layers it has. AlexNet had five convolutional layers while PNet had three convolutional layers. Nevertheless, the hybrid CNNSVR, which featured four convolutional layers, had the lowest mean square error (MSE) for the training set and required a reasonable amount of time 1 min 10 sec to finish the training process. This can be attributed to the architecture's efficiency, which enables faster training on datasets without images. Furthermore, the model avoided superfluous data and converged effectively due to appropriately adjusted hyperparameters. In this case, the further additional convolutional layers and too simple architecture can result in larger MSE. Thus, hybrid CNN-SVR with four convolutional layers has been shown as the best chlorophyll intensity prediction model.

Table 4
Comparison Of Training Completion Time With Four Different Models

Model	AlexNet	PNet	SVR	Hybrid CNN-SVR
Training Completion Time	1 hour 32 min 36 sec	20 sec	1 min 15sec	1 min 10 sec

Apart from that, scatter plots were used to assess the hybrid CNN-SVR's performance. The scatter plot comparison between PNet, AlexNet, SVR, and hybrid CNN-SVR for training and testing sets is shown in Figure 10. The relationship between the actual and predicted values of chlorophyll intensity is depicted in this scatterplot [16]. On each graph, the 1:1 line and the leaf spectrum sample were represented by the linear line and scatter plot, respectively. Because the leaf sample scatterplots for AlexNet, PNet, and SVR did not lie exactly on the linear line, the leaf sample plots were either overprediction or underprediction. In contrast to other models, the Hybrid CNN-SVR was the only one that nearly perfectly lied on the linear line. Therefore, the relationship between the two variables is poorer the more randomly distributed the scatter plots are. It was demonstrated by taking this analysis into account for hybrid CNN-SVR that the suggested model was appropriate for predicting chlorophyll intensity. As a result, the completion time was significant, offered the lowest MSE, and could predict the chlorophyll intensity of different plants at a lower computational cost.

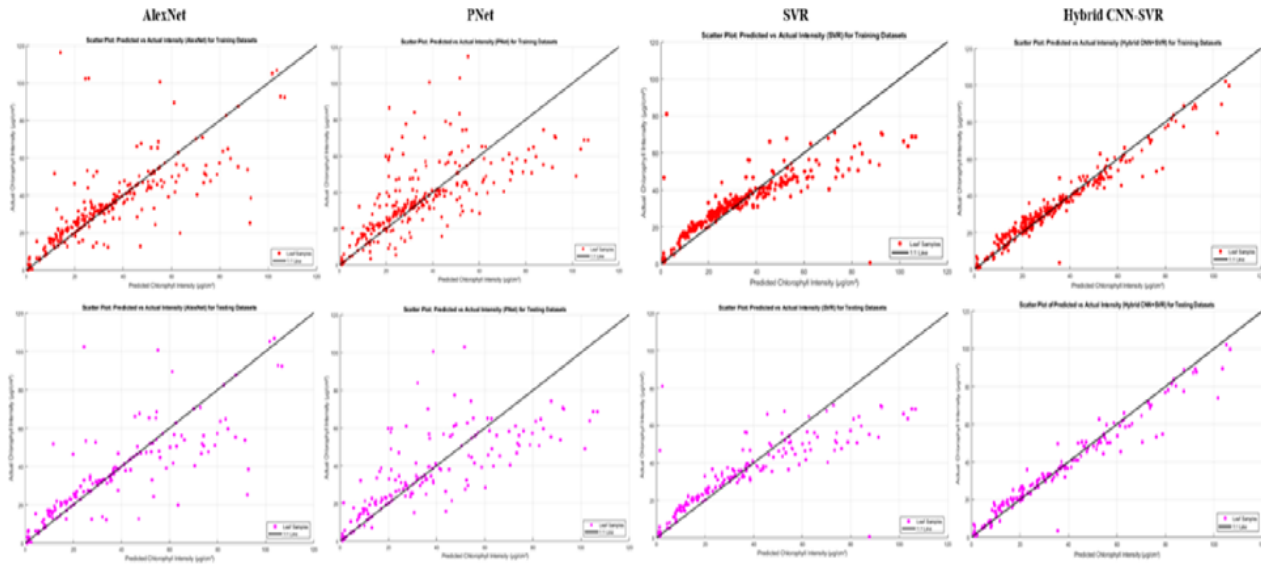

Fig. 10. The Scatter Plots of Actual and Predicted Chlorophyll Intensity for Training and Testing Sets on AlexNet, PNet, SVR, and Hybrid CNN-SVR

5. Conclusions

This study details the effective application of a nondestructive technique to estimate the amount of chlorophyll present in various plant species' leaves. The suggested approach has made a unique contribution to the design of an improved CNN architecture by merging the PNet and AlexNet CNN models and leveraging the SVR network to enhance the prediction task known as hybrid CNN-SVR. With MSE values of 0.156 and 1.149 for the training and testing sets, respectively, this model produced the lowest results. The proposed model successfully predicts chlorophyll intensity based on leaf spectra, utilizing the least complex architecture and optimal hyperparameters. The hybrid CNN-SVR is assessed using the MSE loss function, and the relationship between the observed and expected values of chlorophyll intensity is displayed using scatter plots. Hybrid CNN-SVR's performance is contrasted with that of other current models. As a result, hybrid CNN-SVR offers the following advantages over other models: First, its architecture, consisting of three fully connected and four convolutional layers, is simple and effective in predicting the intensity of greenery. Once trained, hybrid CNN-SVR can also be modified more easily without significantly compromising its performance. Second, because the hybrid CNN-SVR training session took less than five minutes to complete, its computational cost is regarded as reasonable. As a result, the learning curve for predicting chlorophyll intensity is shortened. Third, lowest MSE value is achieved, indicating how closely the actual value and the predicted value match. Lastly, hybrid CNN-SVR's performance can enhance the prediction model even with its small dataset of 552 leaf spectra from different plant species.

Acknowledgement

This research was not funded by any grant.

References

[1] Khoshrou, Mohsen Imanzadeh, Payam Zarafshan, Mohammad Dehghani, Gholamreza Chegini, Akbar Arabhosseini, and Behzad Zakeri. "Deep Learning Prediction of Chlorophyll Content in Tomato Leaves." In *2021 9th RSI International Conference on Robotics and Mechatronics (ICRoM)*, pp. 580-585. IEEE, 2021. https://doi.org/10.1109/ICRoM54204.2021.9663468

[2] Qi, Haixia, Zeyu Wu, Lei Zhang, Jianwen Li, Jingkang Zhou, Zou Jun, and Bingyu Zhu. "Monitoring of peanut leaves chlorophyll content based on drone-based multispectral image feature extraction." *Computers and Electronics in Agriculture* 187 (2021): 106292. https://doi.org/10.1016/j.compag.2021.106292

[3] Bhowmik, Shyamtanu, Anjan Kumar Talukdar, and Kandarpa Kumar Sarma. "Detection of disease in tea leaves using convolution neural network." In *2020 Advanced Communication Technologies and Signal Processing (ACTS)*, pp. 1-6. IEEE, 2020. https://doi.org/10.1109/ACTS49415.2020.9350413

[4] Yang, Jing, and Guanci Yang. "Modified convolutional neural network based on dropout and the stochastic gradient descent optimizer." *Algorithms* 11, no. 3 (2018): 28. https://doi.org/10.3390/a11030028

[5] Ibrahim, Shafaf, Noraini Hasan, Nurbaity Sabri, Khyrina Airin Fariza Abu Samah, and Muhamad Rahimi Rusland. "Palm leaf nutrient deficiency detection using convolutional neural network (CNN)." *International Journal of Nonlinear Analysis and Applications* 13, no. 1 (2022): 2008–6822.

[6] Iqbal, Muhammad Shakaib, Hazrat Ali, Son N. Tran, and Talha Iqbal. "Coconut trees detection and segmentation in aerial imagery using mask region-based convolution neural network." *IET Computer Vision* 15, no. 6 (2021): 428-439. https://doi.org/10.1049/cvi2.12028

[7] Zhang, Runfei, Peiqi Yang, Shouyang Liu, Caihong Wang, and Jing Liu. "Evaluation of the methods for estimating leaf chlorophyll content with SPAD chlorophyll meters." *Remote Sensing* 14, no. 20 (2022): 5144. https://doi.org/10.3390/rs14205144

[8] Zhang, Huichun, Yufeng Ge, Xinyan Xie, Abbas Atefi, Nuwan K. Wijewardane, and Suresh Thapa. "High throughput analysis of leaf chlorophyll content in sorghum using RGB, hyperspectral, and fluorescence imaging and sensor fusion." *Plant Methods* 18, no. 1 (2022): 1-17. https://doi.org/10.1186/s13007-022-00892-0

[9] Prilianti, Kestrilia Rega, Edi Setiyono, Oesman Hendra Kelana, and Tatas Hardo Panintingjati Brotosudarmo. "Deep chemometrics for nondestructive photosynthetic pigments prediction using leaf reflectance spectra." *Information Processing in Agriculture* 8, no. 1 (2021): 194-204. https://doi.org/10.1016/j.inpa.2020.02.001

[10] Shah, Syed Haleem, Yoseline Angel, Rasmus Houborg, Shawkat Ali, and Matthew F. McCabe. "A random forest machine learning approach for the retrieval of leaf chlorophyll content in wheat." *Remote Sensing* 11, no. 8 (2019): 920. https://doi.org/10.3390/rs11080920

[11] Yu, Kang, Georg Leufen, Mauricio Hunsche, Georg Noga, Xinping Chen, and Georg Bareth. "Investigation of leaf diseases and estimation of chlorophyll concentration in seven barley varieties using fluorescence and hyperspectral indices." *Remote Sensing* 6, no. 1 (2013): 64-86. https://doi.org/10.3390/rs6010064

[12] Prilianti, Kestrilia R., Ivan C. Onggara, Marcelinus AS Adhiwibawa, Tatas HP Brotosudarmo, Syaiful Anam, and Agus Suryanto. "Multispectral imaging and convolutional neural network for photosynthetic pigments prediction." In *2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, pp. 554-559. IEEE, 2018. https://doi.org/10.1109/EECSI.2018.8752649

[13] Sonobe, Rei, Hiroto Yamashita, Harumi Mihara, Akio Morita, and Takashi Ikka. "Estimation of leaf chlorophyll a, b and carotenoid contents and their ratios using hyperspectral reflectance." *Remote Sensing* 12, no. 19 (2020): 3265. https://doi.org/10.3390/rs12193265

[14] Prilianti, Kestrilia Rega, SyaifuL Anam, Tatas Hardo Panintingjati, and Agus Suryanto Brotosudarmo. "Non-destructive Photosynthetic Pigments Prediction using Multispectral Imagery and 2D-CNN." *International Journal of Computing* 20, no. 3 (2021): 391-399. https://doi.org/10.47839/ijc.20.3.2285

[15] Prilianti, Kestrilia Rega, Syaiful Anam, Tatas Hardo Panintingjati Brotosudarmo, and Agus Suryanto. "Real-time assessment of plant photosynthetic pigment contents with an artificial intelligence approach in a mobile application." *Journal of Agricultural Engineering* 51, no. 4 (2020): 220-228. https://doi.org/10.4081/jae.2020.1082

[16] Shi, Shuo, Lu Xu, Wei Gong, Bowen Chen, Biwu Chen, Fangfang Qu, Xingtao Tang, Jia Sun, and Jian Yang. "A convolution neural network for forest leaf chlorophyll and carotenoid estimation using hyperspectral reflectance." *International Journal of Applied Earth Observation and Geoinformation* 108 (2022): 102719. https://doi.org/10.1016/j.jag.2022.102719

[17] Ge, Yufeng, Abbas Atefi, Huichun Zhang, Chenyong Miao, Raghuprakash Kastoori Ramamurthy, Brandi Sigmon, Jinliang Yang, and James C. Schnable. "High-throughput analysis of leaf physiological and chemical traits with VIS–NIR–SWIR spectroscopy: a case study with a maize diversity panel." *Plant methods* 15, no. 1 (2019): 1-12. https://doi.org/10.1186/s13007-019-0450-8

[18] Khoshrou, Mohsen Imanzadeh, Payam Zarafshan, Mohammad Dehghani, Gholamreza Chegini, Akbar Arabhosseini, and Behzad Zakeri. "Deep Learning Prediction of Chlorophyll Content in Tomato Leaves." In *2021 9th RSI International Conference on Robotics and Mechatronics (ICRoM)*, pp. 580-585. IEEE, 2021. https://doi.org/10.1109/ICRoM54204.2021.9663468

[19] Qazi, Emad Ul Haq, Abdulrazaq Almorjan, and Tanveer Zia. "A one-dimensional convolutional neural network (1d-cnn) based deep learning system for network intrusion detection." *Applied Sciences* 12, no. 16 (2022): 7986. https://doi.org/10.3390/app12167986

[20] Justus Daniel, Brennan John, Bonner Stephen, Mcgough Andrew. "Predicting the Computational Cost of Deep Learning Models." In *2018 IEEE International Conference on Big Data (Big Data)*, pp. 3873-3882. IEEE, 2018. https://doi.org/10.1109/BigData.2018.8622396