



## Article

# Remote Sensing Imagery Data Analysis Using Marine Predators Algorithm with Deep Learning for Food Crop Classification

Ahmed S. Almasoud <sup>1</sup>, Hanan Abdullah Mengash <sup>2</sup>, Muhammad Kashif Saeed <sup>3,\*</sup>, Faiz Abdullah Alotaibi <sup>4</sup>,  
Kamal M. Othman <sup>5</sup> and Ahmed Mahmud <sup>6</sup>

<sup>1</sup> Department of Information Systems, College of Computer and Information Sciences, Prince Sultan University, Riyadh 11586, Saudi Arabia; almasoud@psu.edu.sa

<sup>2</sup> Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia

<sup>3</sup> Department of Computer Science, Applied College, Muhayil, King Khalid University, Abha 61421, Saudi Arabia

<sup>4</sup> Department of Information Science, College of Humanities and Social Sciences, King Saud University, Riyadh 11437, Saudi Arabia

<sup>5</sup> Department of Electrical Engineering, College of Engineering and Islamic Architecture, Umm Al-Qura University, Makkah 21955, Saudi Arabia; kmothman@uqu.edu.sa

<sup>6</sup> Research Center, Future University in Egypt, New Cairo 11835, Egypt

\* Correspondence: mksaeed@kku.edu.sa

**Abstract:** Recently, the usage of remote sensing (RS) data attained from unmanned aerial vehicles (UAV) or satellite imagery has become increasingly popular for crop classification processes, namely soil classification, crop mapping, or yield prediction. Food crop classification using RS images (RSI) is a significant application of RS technology in agriculture. It involves the use of satellite or aerial imagery to identify and classify different types of food crops grown in a specific area. This information can be valuable for crop monitoring, yield estimation, and land management. Meeting the criteria for analyzing these data requires increasingly sophisticated methods and artificial intelligence (AI) technologies provide the necessary support. Due to the heterogeneity and fragmentation of crop planting, typical classification approaches have a lower classification performance. However, the DL technique can detect and categorize crop types effectively and has a stronger feature extraction capability. In this aspect, this study designed a new remote sensing imagery data analysis using the marine predators algorithm with deep learning for food crop classification (RSMMPA-DLFCC) technique. The RSMMPA-DLFCC technique mainly investigates the RS data and determines the variety of food crops. In the RSMMPA-DLFCC technique, the SimAM-EfficientNet model is utilized for the feature extraction process. The MPA is applied for the optimal hyperparameter selection process in order to optimize the accuracy of SimAM-EfficientNet architecture. MPA, inspired by the foraging behaviors of marine predators, perceptively explores hyperparameter configurations to optimize the hyperparameters, thereby improving the classification accuracy and generalization capabilities. For crop type detection and classification, an extreme learning machine (ELM) model can be used. The simulation analysis of the RSMMPA-DLFCC technique is performed on two benchmark datasets. The extensive analysis of the results portrayed the higher performance of the RSMMPA-DLFCC approach over existing DL techniques.

**Keywords:** remote sensing images; deep learning; crop classification; machine learning; computer vision



**Citation:** Almasoud, A.S.; Mengash, H.A.; Saeed, M.K.; Alotaibi, F.A.; Othman, K.M.; Mahmud, A. Remote Sensing Imagery Data Analysis Using Marine Predators Algorithm with Deep Learning for Food Crop Classification. *Biomimetics* **2023**, *8*, 535. <https://doi.org/10.3390/biomimetics8070535>

Academic Editors: Francesco Visentin, Junzhi Yu and Ali Leylavi Shoushtari

Received: 27 September 2023

Revised: 21 October 2023

Accepted: 31 October 2023

Published: 10 November 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Recent developments in remote sensing (RS) data and technologies deliver the ability of highly accessible, cheap and real time advantages [1]. In recent years, a massive quantity of global coverage RS images have been openly available [2]. In particular, Landsat 8 satellite offers high-resolution multispectral datasets including wealthy data on agricultural

vegetation development which is easily accessible. It allows us to examine the vegetation growth and forecast the changes over time from past to present [3]. RS is an effective data collection technology, and it is broadly employed in agriculture, for example, to monitor crop conditions, crop distribution, and to predict upcoming food production under various situations [4]. Though current agricultural RSs generally use sensors from satellite environments like Landsat and MODIS, they combine and integrate the data acquired from the aerial or ground-based sensors [5,6]. Even if satellite-borne sensors cover a larger range from a local to a national scale, precision agriculture needs remotely sensed data with high efficiency, knowledge, and high resolution to sufficiently study crop conditions, hence giving support to national food provision security.

Aerial or airborne RS that uses classical aerial photography taken from aircraft, light aircraft or unmanned aerial vehicles (UAVs) as its platform, and gets a higher ground resolution of a few centimeters than the satellite image resolution of a few to hundreds of meters provides two important advantages: Primarily, significant biochemical and biophysical variables can be calculated finely at most of the levels of an individual plant, and its images are without mixed pixel effects. Next, important phases of crop development can be finely noticed with the use of active and current crop height created by classical aerial triangulation technology [7]. Additionally, the highly accurate cropland mask, crop-specific categorization and circulation gained from airborne sensors provide extra training and validation data for satellite observation and additionally increase the respective outcome. Successful integration of various sensor sources, wavebands, and time-stamped RS images gives extensive feature data about crops [8]. Thus, it is a reasonable and significant study to discover the crop classification based on RS images.

Classical RS-based image classification procedures of ML were slowly used in the classification and detection of RS images. These models can be classified as supervised and unsupervised classes. The first holds minimum distance, maximum likelihood, and support vector machine (SVM). In this phase, the SVM is extensively applied in RS image classification, even though few problems exist. DL, referring to a deep neural network, is a type of ML technique, and because of its data expression and dominant feature extraction capability, it has been widely adopted. Over the years, the identification rate of DL on most classical identification processes has enhanced considerably [9]. Numerous studies have exhibited that DL can extract features from RS imagery and enhance the classifier performance.

This article develops a remote sensing imagery data analysis using the marine predators algorithm with deep learning for food crop classification (RSMMPA-DLFCC) method. The RSMMPA-DLFCC technique mainly investigates the RS data and determines the variety of food crops. In the RSMMPA-DLFCC technique, the SimAM-EfficientNet model is utilized for the feature extraction process. The MPA is applied for the optimal parameter selection to optimize the accuracy of SimAM-EfficientNet architecture. MPA, inspired by the foraging behaviors of marine predators, perceptively explores hyperparameter configurations to optimize the hyperparameters, thereby improving the classification accuracy and generalization capabilities. For crop type detection and classification, an extreme learning machine (ELM) model can be used. The simulation analysis of the RSMMPA-DLFCC method takes place on the UAV image dataset.

The rest of the paper is organized as follows. Section 2 provides the related works and Section 3 offers the proposed model. Then, Section 4 gives the result analysis and Section 5 concludes the paper.

## 2. Literature Review

Kwak and Park [10] examined self training with domain adversarial networks (STDAN) to classify crop types. The main function of STDAN is to integrate adversarial training for improving spectral discrepancy issues with self training in order to create novel trained data in the targeted field, utilizing present ground truth details. In [11], a unique structure based on deep CNN (DCNN) and the dual attention module (DAM) makes utilization

of the Sentinel 2 time series dataset which was projected for crop identification. Fresh DAM was applied to the removal of enlightened deep features using the advantages of spatial and spectral features of Sentinel 2 datasets. Reedha et al. [12] targeted the design of attention-related DL networks in a significant technique to state the earlier mentioned complications regarding weeds and crop detection with drone systems. The objective is to inspect visual transformers (ViT) and implement them in the identification of plants in UAV images. In [13], the results of accurate recognition were tested to associate the phenology of vegetation products by time series of Landsat8, digital elevation model (DEM), and Sentinel 1. Next, based on the agricultural phenology of crops, radar Sentinel1 and optical Landsat8 time-series data with DEM were used to enhance the performance classification.

Sun et al. [14] proposed a technique for attaining deduction of fine-scale crops by combining RS information from different satellite images by construction of chronological scale crop features inside the parcels employing Sentinel 2A, Gaofen-6, and Landsat 8. The authors adopted a feature-equivalent technique to fill in the missing values in the time series feature-building methods to prevent problems with unidentified crops. Li et al. [15] introduced a scale sequence object-based CNN (SS-OCNN) that identifies images at the object phase by taking segmented object crop parcels as the primary unit of analysis, therefore providing the limits between crop parcels that were defined precisely. Next, the segmented object was identified utilizing the CNN approach combined with an automated generating scale structure of input patch sizes.

Zhai et al. [16] examined the contribution of the data to rice planting area mapping. Specifically, the introduction of the red-edge band was to build a red-edge agricultural index derived from Sentinel 2 data. C band quad pol Radar sat 2 data was also utilized. The authors employed the random forest technique and finally collaborated with radar and optical data to plot rice-planted regions. In [17], the authors designed an enhanced crop planting structure to plot the structure for rainy and cloudy regions using collective optical data and SAR data. First, the author removed geo parcels from optical images with high dimensional resolution. Next, the authors made an RNN-based classification appropriate for remote detecting images on a geo parcel scale.

### 3. The Proposed Model

This manuscript offered the development of automated food crop classification using the RSMMPA-DLFCC technique. The RSMMPA-DLFCC technique mainly investigates the RS data and determines different types of food crops. In the RSMMPA-DLFCC technique, three major phases of operations are involved, namely the SimAM-EfficientNet feature extractor, MPA-based hyperparameter tuning, and ELM classification. Figure 1 represents the entire process of the RSMMPA-DLFCC approach.

#### 3.1. Feature Extraction Using SimAM-EfficientNet Model

The RSMMPA-DLFCC technique applies the SimAM-EfficientNet model to derive feature vectors. A novel CNN called EfficientNet was launched by Google researchers [18]. The study uses a multi-dimensional hybrid method scaling model making them consider the speed and accuracy of the model even though the existing network has advanced considerably in speed and accuracy. Through compound scaling factors, ResNet raises the network depth to optimize the performance. By improving accuracy and ensuring speed, EfficientNet balances the network depth, width, and resolution. EfficientNet-B0 is the initial EfficientNet model. The most basic model B0 is: concerning resolution, layers, and channels, B1-B7 overall of 7 models adapted from B0.

Many existing attention modules generate 1D or 2D weights. Next, the weights created are extended for channel and spatial attention. Generally, the present attention module faces the two subsequent challenges. The former is the attention module could extract features through channel and space that results in the flexibility of attention weight. Moreover, CNN is influenced by a series of factors and has a complex structure. SimAM considers these spaces and channels in contrast to them. Without adding parameters, it presents

3D attention weights to the original network. Based on neuroscience theory, an energy function can be defined and, in turn, derive a solution that converges faster. This operation is executed in ten lines of code. An additional benefit of SimAM is that it prevents excessive adjustment to the network architecture. Hence, SimAM is lightweight, more flexible, and modular. In numerous instances, SimAM is better than the conventional CBAM and SE attention models. Figure 2 illustrates the architecture of SimAM-EfficientNet.

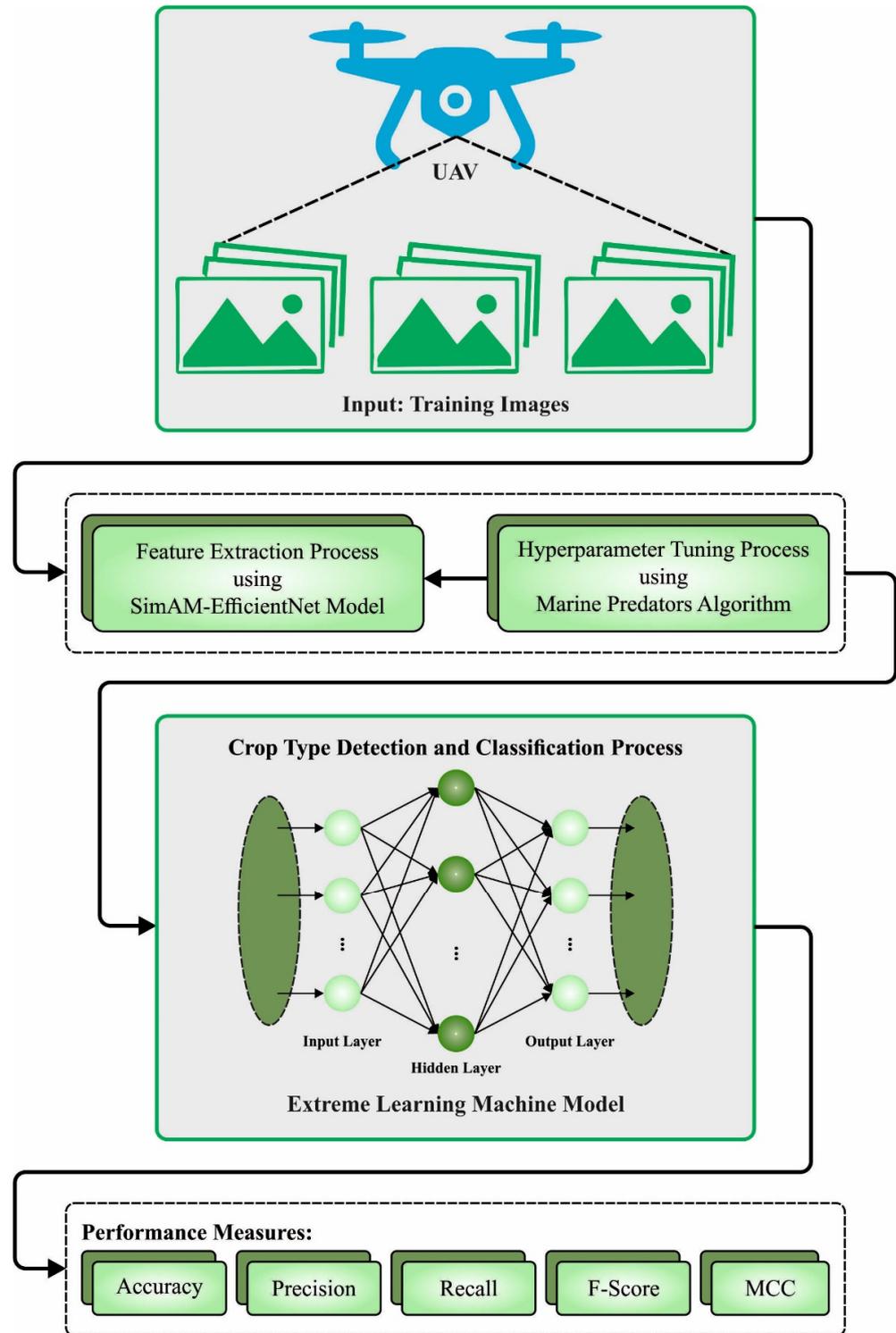
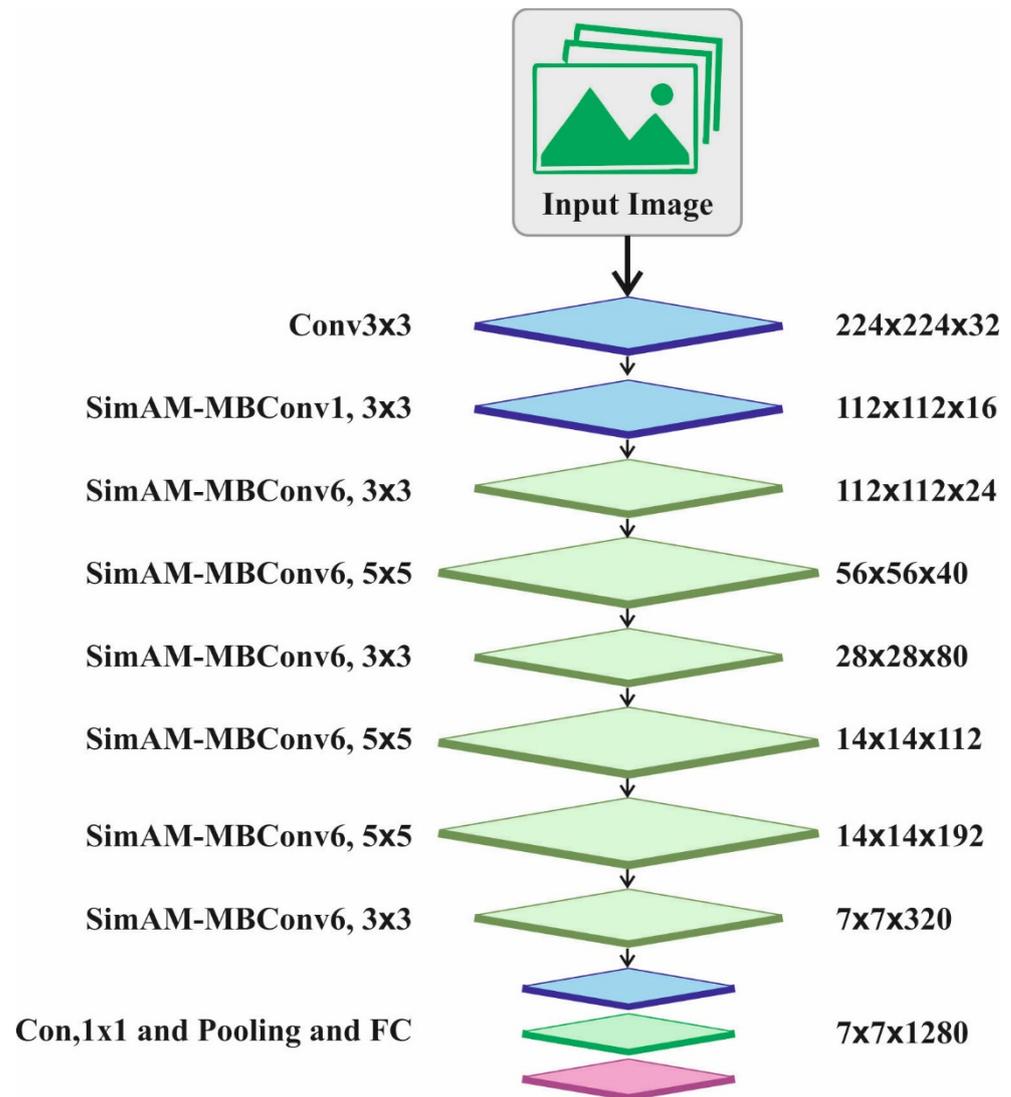


Figure 1. Overall process of RSMPPA-DLFCC algorithm.



**Figure 2.** Architecture of SimAM-EfficientNet.

The SimAM model defines an energy function and looks for important neurons. It adds regular terms and uses binary labels. At last, the minimal energy is evaluated by the following expression:

$$e_t^* = \left(4(\lambda + \sigma^2)\right) / \left((t - u)^2 + 2\sigma^2 + 2\lambda\right) \tag{1}$$

$$u_t = \frac{1}{M-1} \sum_{i=1}^{M-1} x_i, \sigma_t^2 = \frac{1}{M-1} \sum_{i=1}^{M-1} (x_i - u_t)^2 \tag{2}$$

where  $\mu_t$  and  $\sigma_t^2$  are the mean and variance of each neuron.  $t$  is the target neuron.  $\lambda$  indicates the regularization coefficient. Using  $M = H \times W$ , the neuron count on that channel is attained. Finally, the dissimilarity between neurons and peripheral neurons is associated with the energy used. The implication of all the neurons is evaluated by  $1/e^*$ . The scaling operator is used to refine the feature and it can be formulated as follows:

$$X = X \cdot \text{sigmoid}(1/E) \tag{3}$$

The *sigmoid* function is used to limit the size of the  $E$  value. In Equation (3),  $E$  group each  $e$  across the channel and spatial sizes.

EfficientNet-B0 has a total of nine phases. The initial phase is  $3 \times 3$  convolutional layers. The second to the eighth phases are MBConv, which is the building block of these network models. The last phase is made up of a pooling layer, a  $1 \times 1$  convolutional layer, and the FC layer. MBConv has five different parts. The initial part is a  $1 \times 1$  convolutional layer. The next part is a depth-wise convolution layer. The third part is the SE attention mechanism. The fourth part is a  $1 \times 1$  convolutional layer for reduction dimension. Lastly, the dropout layer lessens the over-fitting problem. After the first convolutional layer, the SimAM module was added to increase channel and spatial weights. The original EfficientNet comprises the SE attention mechanism.

The SimAM-EfficientNet is made up of seven SimAM-MBConv models, one FC layer, two convolution layers, and one pooling layer. At first, the images with  $224 \times 224 \times 3$  dimensions are ascended by the  $3 \times 3$  convolution layers. The dimensions of the images obtained with features are  $112 \times 112 \times 32$ . Next, the image features are extracted by the SimAM-Conv. The connection will be deactivated when both SimAM-Convs are the same, and the input will connect. The FC layer is utilized for classification and the original channel is restored after a  $1 \times 1$  point-wise convolutional layer.

### 3.2. Hyperparameter Tuning Using MPA

For the optimal hyperparameter selection process, the MPA is applied. The MPA is derived from the foraging tactics of the ocean predator [19]. MPA is a population-based metaheuristic approach. The optimization technique begins with the arbitrary solution.

$$X_0 = X_{\min} + rand(X_{\max} - X_{\min}) \tag{4}$$

where  $X_{\min}$  and  $X_{\max}$  denotes the lower and upper boundaries, and  $rand$  is a randomly generated integer in the range  $[0, 1]$ . In the MPA, Prey and Elite are two different matrices with similar dimensions. The optimum solution is selected as the fittest predator while creating the Elite matrix.

The finding of and search for prey is checked through these matrices.  $\vec{X}^I$  indicates the dominant predator vector,  $n$  is the searching agent, and  $d$ , the dimension. Both prey and predator are the search agents.

$$Elite = \begin{bmatrix} X_{1,1}^I & X_{1,2}^I & \dots & X_{1,d}^I \\ X_{2,1}^I & X_{2,2}^I & \dots & X_{2,d}^I \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ X_{n,1}^I & X_{n,2}^I & \dots & X_{n,d}^I \end{bmatrix}_{n \times d} \tag{5}$$

where the  $j^{th}$  dimension of  $i^{th}$  prey is represented as  $X_{i,j}$ . The optimization method is connected to both matrices. Predator uses these matrices for updating the position.

$$Prey = \begin{bmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,d} \\ X_{2,1} & X_{2,2} & \dots & X_{2,d} \\ X_{3,1} & X_{3,1} & \dots & X_{3,d} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ X_{n,1} & X_{n,2} & \dots & X_{n,d} \end{bmatrix}_{n \times d} \tag{6}$$

In the MPA, there are three stages discussed in detail.

Phase 1 occurs if  $< ((Max\_Iter)/3)$ .  $Iter$  and  $Max\_Iter$  denote the existing and maximal iteration counter.  $P$  shows the constant number with the value of 0.5. The

appropriate tactic is one where the predator should stop. In Equation (7) of stage 1, vector  $R_B$  portrays the Brownian motion and uniformly distributed random number in [0,1].

$$\begin{aligned} \overrightarrow{stepsize}_i &= \vec{R} \otimes \left( \overrightarrow{Elite}_i - \vec{R}_B \otimes \overrightarrow{Prey}_i \right) i = 1, \dots n \\ \overrightarrow{Prey}_i &= \overrightarrow{Prey}_i + P \cdot \vec{R} \otimes \overrightarrow{stepsize}_i \end{aligned} \tag{7}$$

Phase 2 realized if  $((Max\_Iter)/3) < Iter < ((2Max\_Iter)/3)$ . Once the prey movement is Lévy, then the predator movement should be Brownian. The prey is responsible for exploitation, and the predator is responsible for exploration. The multiplication of  $\vec{R}_L$  and  $Prey$  represent the prey movement, and the prey movement can be exemplified by adding the  $stepsize$  to the prey position. The  $\vec{R}_L$  vector is a random number representing Lévy motion. CF denotes an adaptive parameter.  $stepsize$  for the predator movement can be controlled by the CF.

$$\begin{aligned} \overrightarrow{stepsize}_i &= \vec{R}_L \otimes \left( \overrightarrow{Elite}_i - \vec{R}_L \otimes \overrightarrow{Prey}_i \right) i = 1, \dots n/2 \\ \overrightarrow{Prey}_i &= \overrightarrow{Prey}_i + P \cdot \vec{R} \otimes \overrightarrow{stepsize}_i \end{aligned} \tag{8}$$

$$\begin{aligned} \overrightarrow{stepsize}_i &= \vec{R}_B \otimes \left( \vec{R}_B \otimes \overrightarrow{Elite}_i - \overrightarrow{Prey}_i \right) i = \frac{n}{2}, \dots n \\ \overrightarrow{Prey}_i &= \overrightarrow{Elite}_i + P \cdot CF \otimes \overrightarrow{stepsize}_i \\ CP &= \left( 1 - \frac{Iter}{Max\_Iter} \right)^{(2 \frac{Iter}{Max\_Iter})} \end{aligned} \tag{9}$$

Phase 3 occurs If  $> ((2Max\_Iter)/3)$ . As the optimum strategy, the predator movement is Lévy.

$$\begin{aligned} \overrightarrow{stepsize}_i &= \vec{R}_L \otimes \left( \vec{R}_L \otimes \overrightarrow{Elite}_i - \overrightarrow{Prey}_i \right) i = 1, \dots n \\ \overrightarrow{Prey}_i &= \overrightarrow{Elite}_i + P \cdot CF \otimes \overrightarrow{stepsize}_i \end{aligned} \tag{10}$$

The factors including fish aggregating devices (FADs) or eddy formation may affect the predator strategy are called the FADs effect.  $r$  is a randomly generated value within [0,1].  $\vec{U}$  shows the binary vector with an array of 0 and 1.  $r1$  and  $r2$  depict the random indexes of prey matrices.  $\vec{X}_{min}$  and  $\vec{X}_{max}$  denote the lower and upper boundaries of the dimension.

$$\overrightarrow{Prey}_i = \begin{cases} \overrightarrow{Prey}_i + CP \left[ \vec{X}_{min} + \vec{R} \otimes \left( \vec{x}_{max} - \vec{X}_{min} \right) \right] \otimes \vec{U}, & r \leq FADs \\ \overrightarrow{Prey}_i + [FADs(1 - r) + r] \left( \overrightarrow{Prey}_i - \overrightarrow{Prey}_i \right), & r > FADs \end{cases} \tag{11}$$

The fitness selection is a major factor in the MPA technique. An encoded solution is used for evaluating the outcome of the solution candidate. The accuracy values are the foremost conditions used to design an FF.

$$Fitness = \max (P) \tag{12}$$

$$P = \frac{TP}{TP + FP} \tag{13}$$

where  $TP$  and  $FP$  represent the true and false positive values.

### 3.3. Classification Using ELM Model

The ELM algorithm is applied for the automated detection and classification of food crops. The ELM model is used to generate the weight between the hidden and the input layers at random, and during the training process, it does not need to be adjusted and only needs to set the number of HL neurons in order to attain an optimum result [20]. Assume  $N$  arbitrary sample  $(X, t)$ , where  $X_j = [x_{j1}, x_{j2} \dots x_{jn}]^T \in R^n, t_i = [t_{i1}, t_{i2} \dots t_{im}]^T \in R$  is formulated by

$$\sum_{i=1}^L \beta_i g(W_i \cdot X_j + b_i) = t_j, j = 1, \dots, N \tag{14}$$

The weight of  $i^{th}$  neurons in the input layer and HL is  $W_i = [w_{i1}, w_{i2} \dots w_{in}]^T$ , chosen at random. The resultant weight is  $\beta_i$ , and the learning objective is to obtain the fittest  $\beta_i$ . The  $j^{th}$  input vector is  $X_j$ . The inner product of  $W_i$  and  $X_j$  is  $W_i \cdot X_j$ . The bias of  $i^{th}$  HL neuron is  $b_i$ . The set non-linear activation function is  $g(x)$ . The output vector of the  $i^{th}$  neurons is  $g(W_i \cdot X_j + b_i)$ . The target vector attained from the  $j^{th}$  input vector is  $t_j$ . It can be represented in the matrix form:

$$H\beta = T$$

$$H(W_1, \dots, W_L, b_1, \dots, b_L, \dots, X_1, \dots, X_L)$$

$$= \begin{bmatrix} g(W_1 \cdot X_1 + b_1) & \dots & g(W_L \cdot X_N + b_1) \\ \vdots & \ddots & \vdots \\ g(W_1 \cdot X_N + b_1) & \dots & g(W_L \cdot X_N + b_1) \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix} \text{ and } T = \begin{bmatrix} T_1^T \\ \vdots \\ T_L^T \end{bmatrix} \tag{15}$$

The output of the HL node is  $H$ , the output weight is  $\beta$ , and the desired output is  $T$ . The following equation is used to get  $\hat{W}_i, \hat{\beta}_i, \hat{b}_i$  as follows:

$$\|H(\hat{W}_i, \hat{b}_i)\hat{\beta}_i - T\| = \min_{W, b, \beta} \|H(w_i, b_i)\beta_i - T\|, i = 1, \dots, L \tag{16}$$

As shown in Equation (17), this corresponds to minimalizing the loss function,

$$E = \sum_{j=1}^N \left( \sum_{i=1}^L \beta_i g(W_i \cdot X_j + b_i) - t_j \right)^2 \tag{17}$$

Since the HL offset and the input weight  $W_i$  are determined randomly, then the output matrix of HL is also defined. As shown in Equation (18), the training purpose is transmuted into resolving a linear formula  $H\beta = T$ :

$$\hat{\beta} = H^+T \tag{18}$$

where the optimum output weight is  $\hat{\beta}$ . The Moore–Penrose generalized the inverse of  $H$  matrix is  $H^+$ , and it is shown that the norm of the obtained solution is unique and minimal. Thus, ELM has better robustness and generalization.

### 4. Results Analysis

The proposed model is simulated using the Python 3.8.5 tool. The proposed model is experimented on PC i5-8600k, GeForce 1050Ti 4 GB, 16 GB RAM, 250 GB SSD, and 1 TB HDD. The food crop classification performance of the RSMFA-DLFC system is validated

on the UAV image dataset [21], comprising 6450 samples with six classes. For experimental validation, we have used 80:20 and 70:30 of training (TR)/testing (TS) set.

Figure 3 demonstrates the confusion matrices produced by the RSMPA-DLFCC technique under 80:20 and 70:30 of the TR phase/TS phase. The experimental values specified the efficient recognition of all six classes.

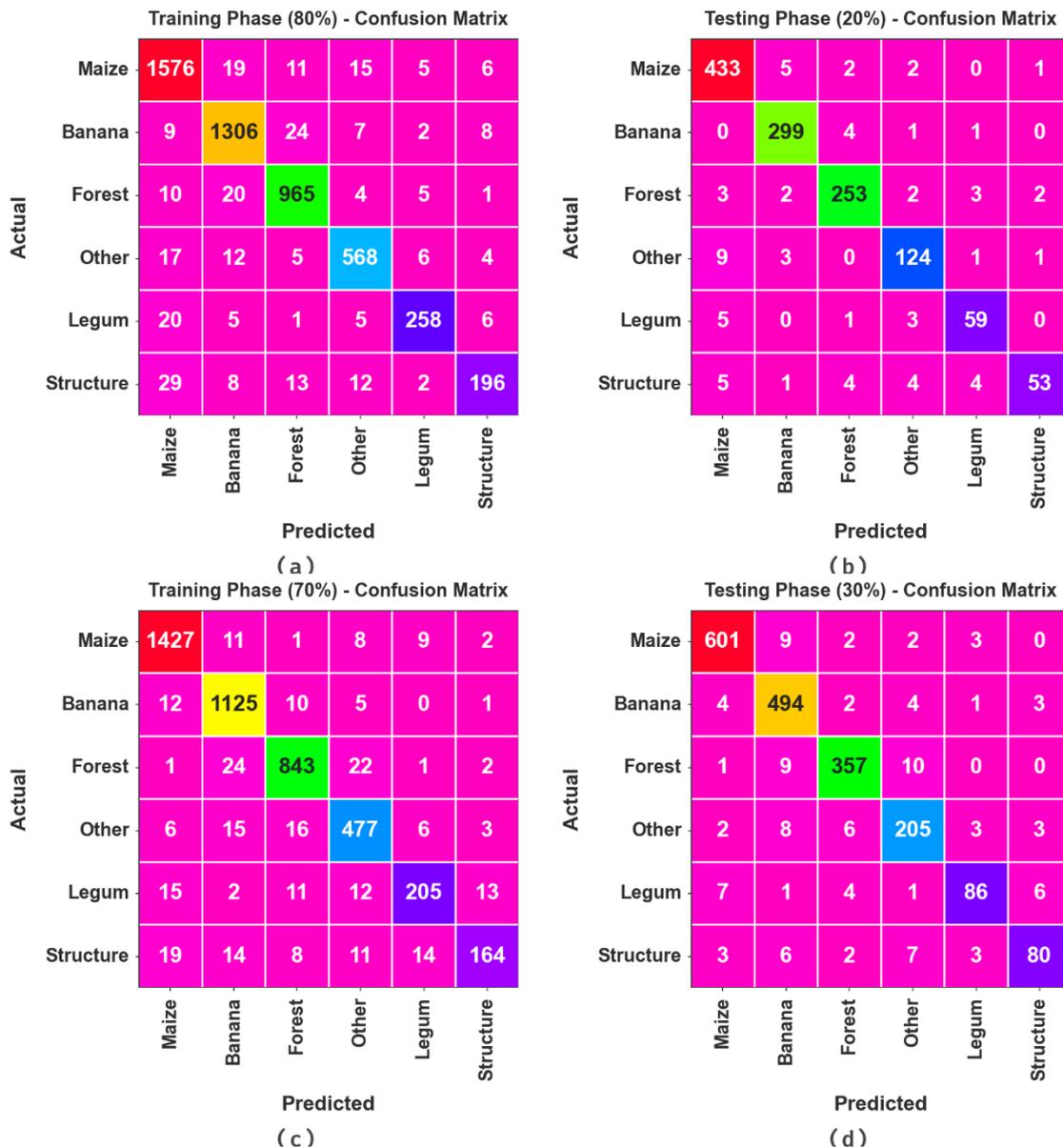
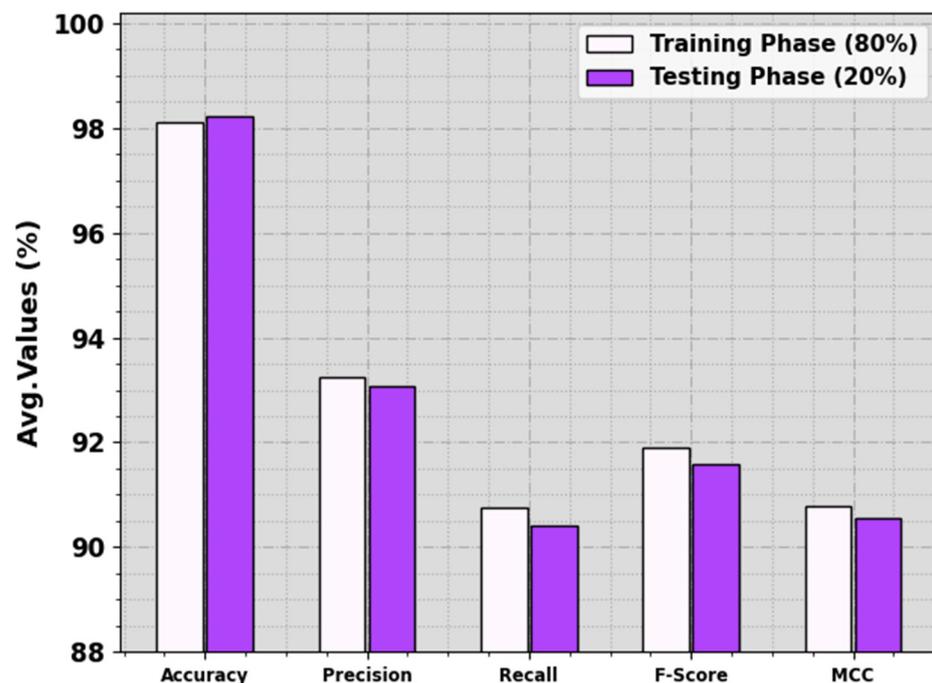


Figure 3. Confusion matrices of (a,b) 80:20 of TR phase/TS phase and (c,d) 70:30 of TR phase/TS phase.

In Table 1 and Figure 4, the food crop classification analysis of the RSMPA-DLFCC methodology is calculated at 80:20 of the TR phase/TS phase. The observational data specified that the RSMPA-DLFCC system properly categorizes seven types of crops. With 80% of the TR phase, the RSMPA-DLFCC technique offers an average  $accu_y$  of 98.12%,  $prec_n$  of 93.23%,  $reca_l$  of 90.76%,  $F_{score}$  of 91.89%, and MCC of 90.77%. Additionally, with 20% of TS phase, the RSMPA-DLFCC method offers an average  $accu_y$  of 98.22%,  $prec_n$  of 93.06%,  $reca_l$  of 90.42%,  $F_{score}$  of 91.57%, and MCC of 90.56%, respectively.

**Table 1.** Food crop classifier outcome of RSMMPA-DLFCC algorithm at 80:20 of TR phase/TS phase.

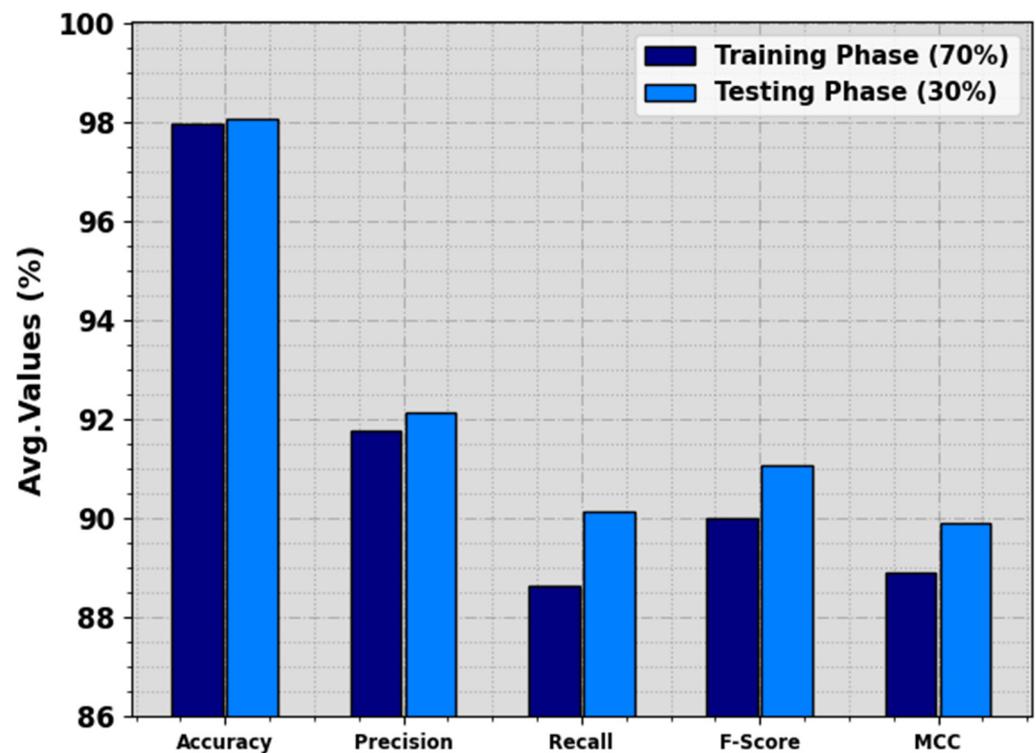
| Class          | $Accu_y$ | $Prec_n$ | $Reca_l$ | $F_{score}$ | MCC   |
|----------------|----------|----------|----------|-------------|-------|
| TR Phase (80%) |          |          |          |             |       |
| Maize          | 97.27    | 94.88    | 96.57    | 95.72       | 93.72 |
| Banana         | 97.79    | 95.33    | 96.31    | 95.82       | 94.32 |
| Forest         | 98.18    | 94.70    | 96.02    | 95.36       | 94.23 |
| Other          | 98.31    | 92.96    | 92.81    | 92.89       | 91.93 |
| Legum          | 98.90    | 92.81    | 87.46    | 90.05       | 89.51 |
| Structure      | 98.28    | 88.69    | 75.38    | 81.50       | 80.89 |
| Average        | 98.12    | 93.23    | 90.76    | 91.89       | 90.77 |
| TS Phase (20%) |          |          |          |             |       |
| Maize          | 97.52    | 95.16    | 97.74    | 96.44       | 94.55 |
| Banana         | 98.68    | 96.45    | 98.03    | 97.24       | 96.38 |
| Forest         | 98.22    | 95.83    | 95.47    | 95.65       | 94.53 |
| Other          | 97.98    | 91.18    | 89.86    | 90.51       | 89.39 |
| Legum          | 98.60    | 86.76    | 86.76    | 86.76       | 86.03 |
| Structure      | 98.29    | 92.98    | 74.65    | 82.81       | 82.47 |
| Average        | 98.22    | 93.06    | 90.42    | 91.57       | 90.56 |

**Figure 4.** Average of RSMMPA-DLFCC algorithm at 80:20 of TR phase/TS phase.

In Table 2 and Figure 5, the food crop classification analysis of the RSMMPA-DLFCC technique is calculated at 70:30 of TR Phase/TS Phase. The experimental values indicate that the RSMMPA-DLFCC technique appropriately categorizes seven types of crops. With 70% of the TR phase, the RSMMPA-DLFCC algorithm offers an average  $accu_y$  of 97.98%,  $prec_n$  of 91.79%,  $reca_l$  of 88.64%,  $F_{score}$  of 90.02%, and MCC of 88.90%, respectively. In addition, with 30% of TS phase, the RSMMPA-DLFCC system offers average  $accu_y$  of 98.07%,  $prec_n$  of 92.13%,  $reca_l$  of 90.13%,  $F_{score}$  of 91.06%, and MCC of 89.92%, correspondingly.

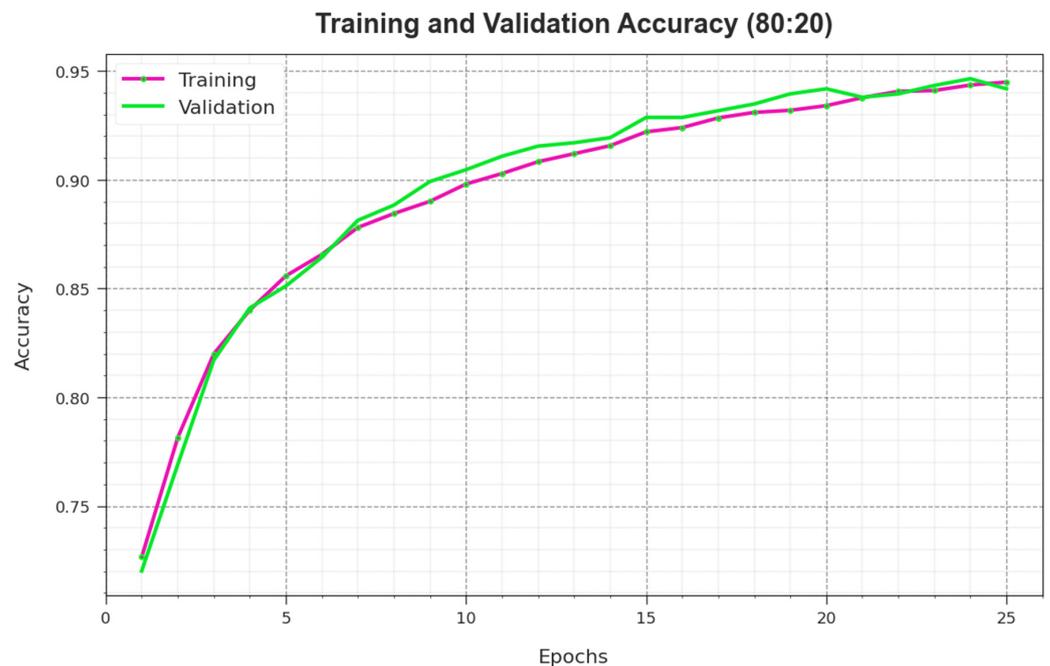
**Table 2.** Food crop classifier outcome of RSMPA-DLFCC algorithm at 70:30 of TR phase/TS phase.

| Class          | $Accu_y$ | $Prec_n$ | $Recall$ | $F_{score}$ | MCC   |
|----------------|----------|----------|----------|-------------|-------|
| TR Phase (70%) |          |          |          |             |       |
| Maize          | 98.14    | 96.42    | 97.87    | 97.14       | 95.77 |
| Banana         | 97.92    | 94.46    | 97.57    | 95.99       | 94.61 |
| Forest         | 97.87    | 94.83    | 94.40    | 94.61       | 93.29 |
| Other          | 97.70    | 89.16    | 91.20    | 90.17       | 88.87 |
| Legum          | 98.16    | 87.23    | 79.46    | 83.16       | 82.29 |
| Structure      | 98.07    | 88.65    | 71.30    | 79.04       | 78.55 |
| Average        | 97.98    | 91.79    | 88.64    | 90.02       | 88.90 |
| TS Phase (30%) |          |          |          |             |       |
| Maize          | 98.29    | 97.25    | 97.41    | 97.33       | 96.08 |
| Banana         | 97.57    | 93.74    | 97.24    | 95.46       | 93.83 |
| Forest         | 98.14    | 95.71    | 94.69    | 95.20       | 94.05 |
| Other          | 97.62    | 89.52    | 90.31    | 89.91       | 88.57 |
| Legum          | 98.50    | 89.58    | 81.90    | 85.57       | 84.88 |
| Structure      | 98.29    | 86.96    | 79.21    | 82.90       | 82.11 |
| Average        | 98.07    | 92.13    | 90.13    | 91.06       | 89.92 |

**Figure 5.** Average of RSMPA-DLFCC algorithm at 70:30 of TR phase/TS phase.

To calculate the performance of the RSMPA-DLFCC methodology on 80:20 of TR Phase/TS Phase, TR and TS  $accu_y$  curves are defined, as shown in Figure 6. The TR and TS  $accu_y$  curves demonstrate the performance of the RSMPA-DLFCC technique over numerous epochs. The figure offers the details about the learning task and generalization capabilities of the RSMPA-DLFCC system. With a rise in epoch count, it is observed that the TR and TS

$accu_y$  curves attained are enhanced. It is noted that the RSMPA-DLFCC approach enriches testing accuracy that has the ability to identify the patterns in the TR and TS data.



**Figure 6.**  $Accu_y$  curve of RSMPA-DLFCC algorithm at 80:20 of TR phase/TS phase.

Figure 7 illustrates an overall TR and TS loss value of the RSMPA-DLFCC methodology on 80:20 of TR Phase/TS Phase over epochs. The TR loss shows the model loss acquired reduces over epochs. Mainly, the loss values are decreased as the model adapts the weight to diminish the predicted error on the TR and TS data. The loss analysis illustrates the level where the model is fitting the training data. It is evidenced that the TR and TS loss is progressively minimized and described that the RSMPA-DLFCC technique effectively learns the patterns revealed in the TR and TS data. It is also observed that the RSMPA-DLFCC methodology modifies the parameters for reducing the difference between the predicted and actual training labels.

The PR curve of the RSMPA-DLFCC approach on 80:20 of TR phase/TS phase, illustrated by plotting precision against recall as described in Figure 8, confirms that the RSMPA-DLFCC technique achieves improved PR values under all classes. The figure represents that the model learns to identify different class labels. The RSMPA-DLFCC achieves improved effectiveness in the recognition of positive samples with reduced false positives.

The ROC analysis, provided by the RSMPA-DLFCC system on 80:20 of TR phase/TS phase demonstrated in Figure 9, has the ability the differentiate between class labels. The figure shows valuable insights into the trade-off between the TPR and FPR rates over dissimilar classification thresholds and differing numbers of epochs. It introduces the accurately predicted performance of the RSMPA-DLFCC methodology on the classification of various classes.

In Table 3, detailed comparative results of the RSMPA-DLFCC technique are demonstrated with current models [22,23]. Figure 10 investigates a comparative analysis of the RSMPA-DLFCC with recent approaches in terms of  $accu_y$ . The experimental values highlighted that the RSMPA-DLFCC technique reaches an increased  $accu_y$  of 98.22%, whereas the SBODL-FCC, DNN, AlexNet, VGG-16, ResNet, and SVM models obtain decreased  $accu_y$  values of 97.43%, 86.23%, 90.49%, 90.35%, 87.70%, and 86.69%, respectively.

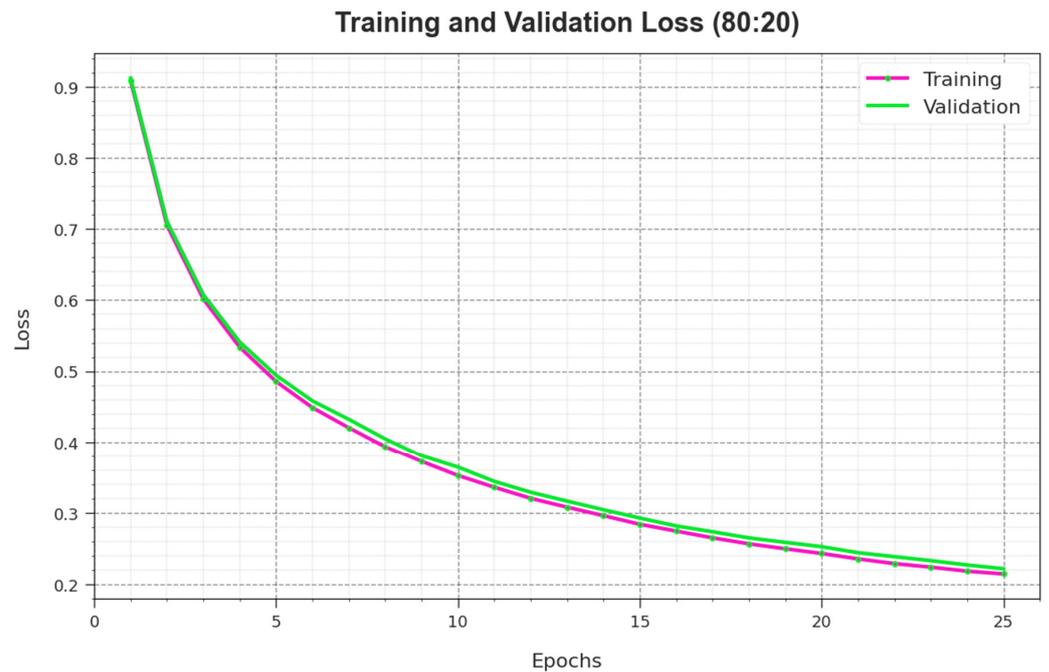


Figure 7. Loss curve of RSMPA-DLFCC algorithm at 80:20 of TR phase/TS phase.

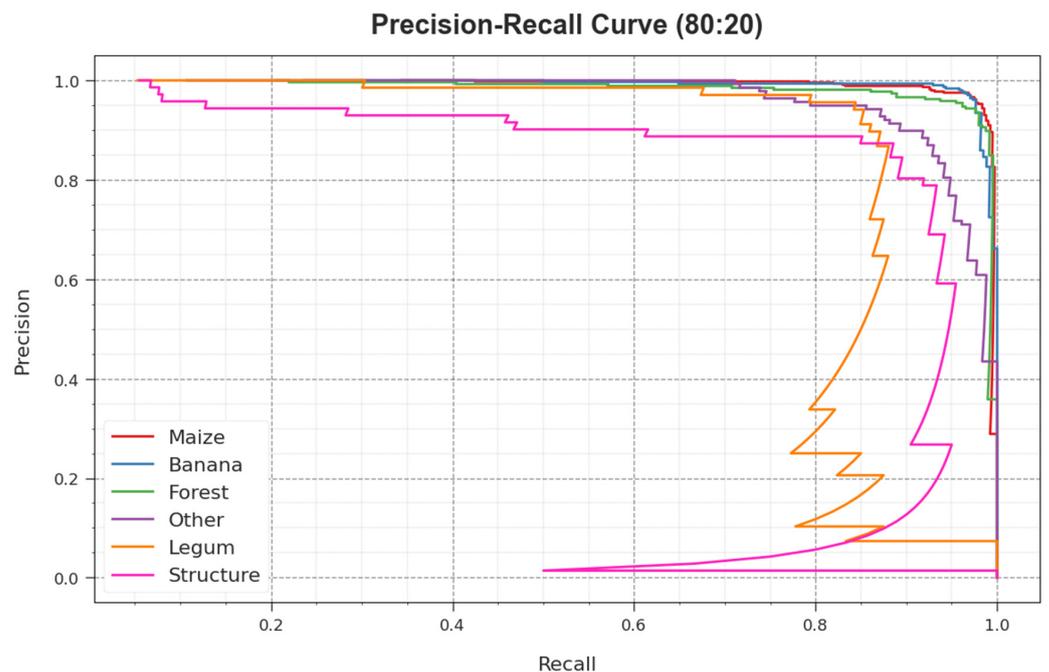


Figure 8. PR curve of RSMPA-DLFCC algorithm at 80:20 of TR/TS phase.

Figure 11 investigates a comparative analysis of the RSMPA-DLFCC system with recent techniques, with respect to  $prec_n$  and  $reca_1$ . The observational data highlighted that the RSMPA-DLFCC system attains a raised  $prec_n$  of 93.06%, while the SBODL-FCC, DNN, AlexNet, VGG-16, ResNet, and SVM methods obtain reduced  $prec_n$  values of 89.02%, 86.11%, 87.68%, 85.28%, 86.42%, and 87.99%, correspondingly. In addition, the RSMPA-DLFCC system attains  $reca_1$  values of 90.42% whereas SBODL-FCC, DNN, AlexNet, VGG-16, ResNet, and SVM systems get decreased  $reca_1$  values of 85.03%, 84.39%, 81.7%, 81.35%, 81.18%, and 83.61%, respectively. These experimental data indicated that the RSMPA-DLFCC methodology reaches the maximum food crop classification process.

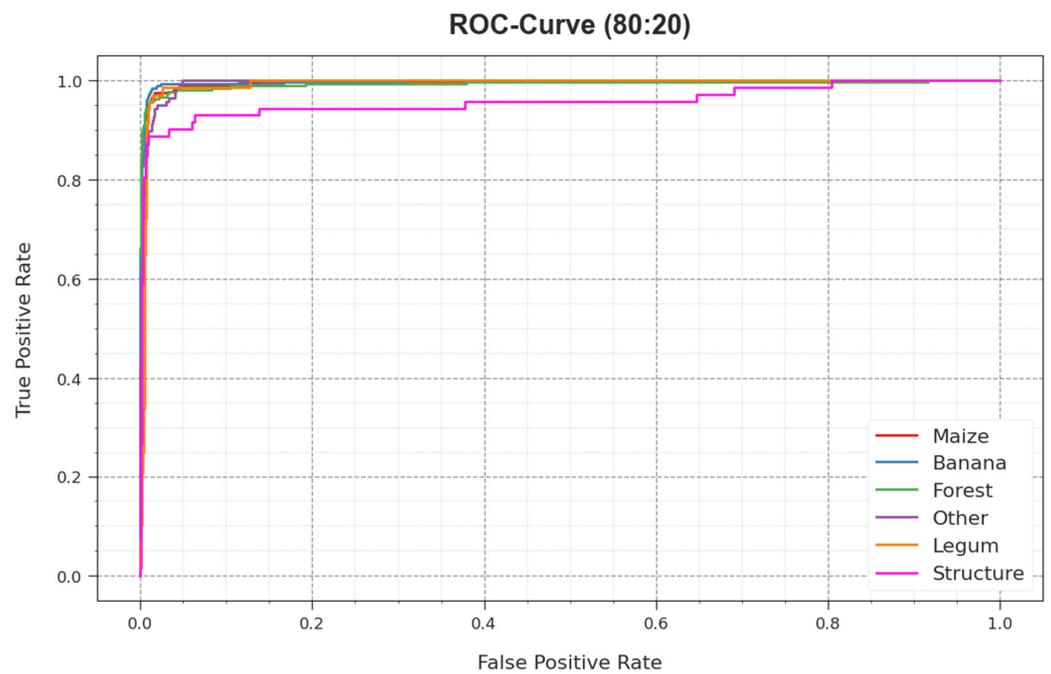


Figure 9. ROC curve of RSMPA-DLFCC algorithm at 80:20 of TR/TS phase.

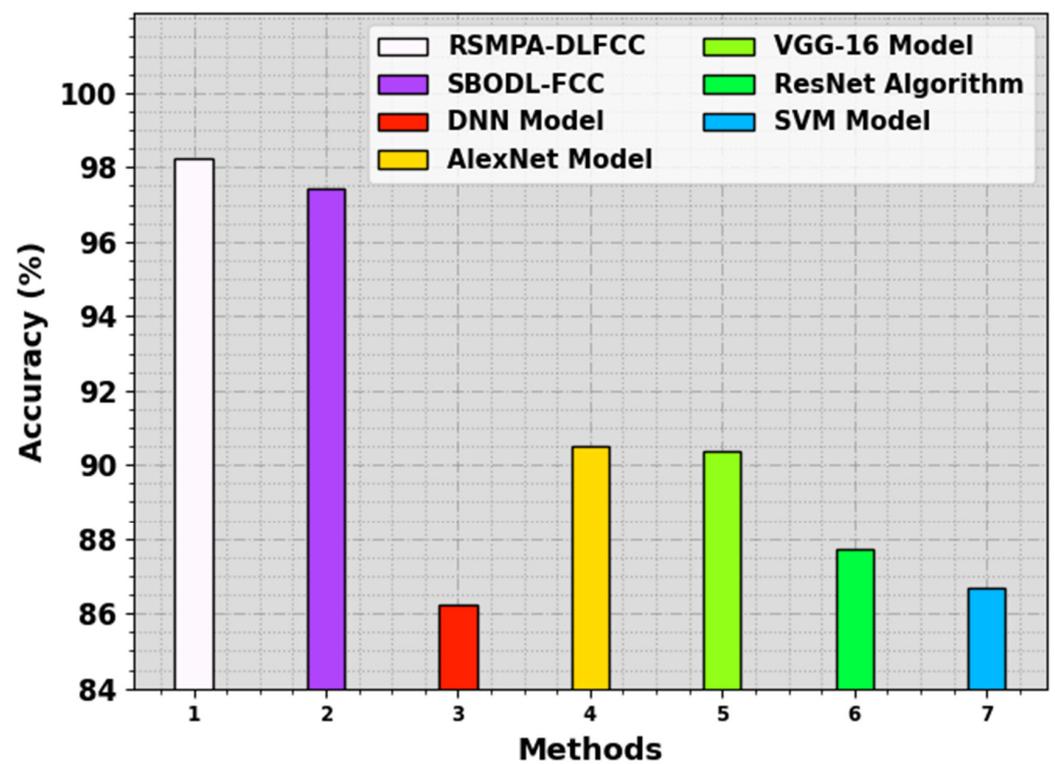


Figure 10. Accuracy Comparative outcome of RSMPA-DLFCC algorithm with other systems.

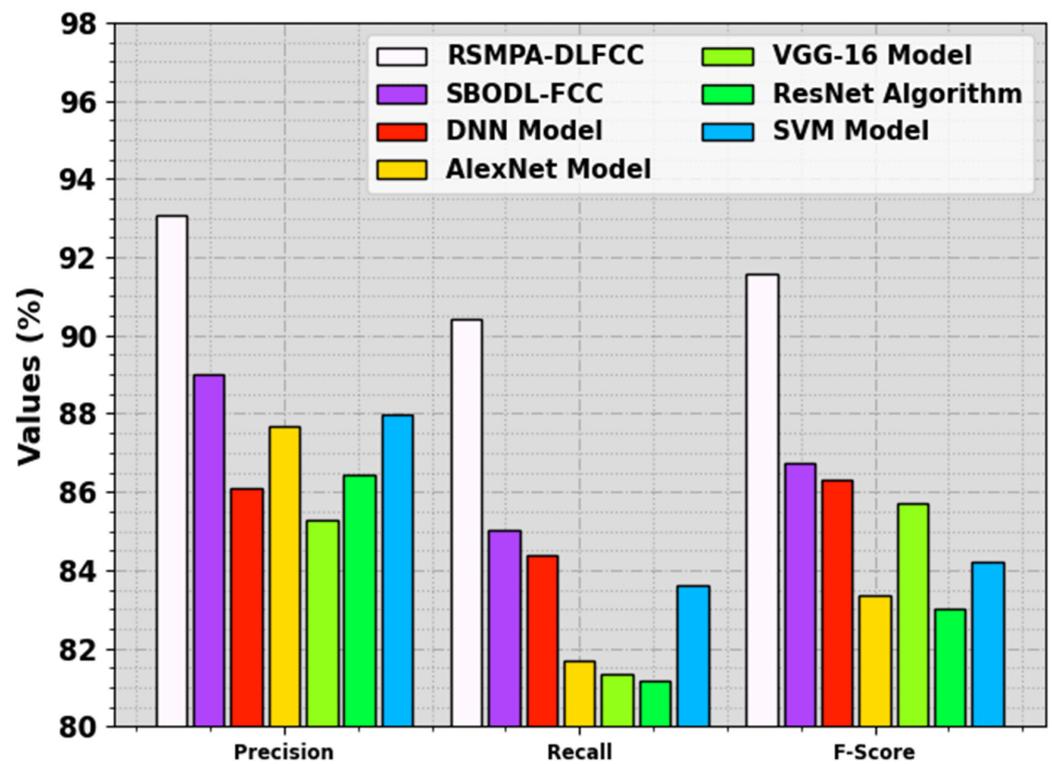


Figure 11. Comparative outcome of RSMPA-DLFCC algorithm with other systems.

Table 3. Comparative outcome of RSMPA-DLFCC with other systems.

| Methods               | $Accu_y$ | $Prec_n$ | $Reca_l$ | $F_{score}$ |
|-----------------------|----------|----------|----------|-------------|
| RSMPA-DLFCC           | 98.22    | 93.06    | 90.42    | 91.57       |
| SBODL-FCC [22]        | 97.43    | 89.02    | 85.03    | 86.74       |
| DNN I [23]            | 86.23    | 86.11    | 84.39    | 86.29       |
| AlexNet Model [23]    | 90.49    | 87.68    | 81.7     | 83.36       |
| VGG-16 Model [23]     | 90.35    | 85.28    | 81.35    | 85.7        |
| ResNet Algorithm [23] | 87.7     | 86.42    | 81.18    | 83.02       |
| SVM Model [23]        | 86.69    | 87.99    | 83.61    | 84.21       |

## 5. Conclusions

This manuscript offered the development of automated food crop classification using the RSMPA-DLFCC technique. The RSMPA-DLFCC technique mainly investigates the RS data and determines different types of food crops. In the RSMPA-DLFCC technique, the SimAM-EfficientNet model is utilized for the feature extraction process. The MPA is applied for the optimum hyperparameter selection in order to optimize the accuracy of SimAM-EfficientNet architecture. The simulation analysis of the RSMPA-DLFCC method takes place on benchmark UAV image dataset. The widespread result analysis portrayed the higher performance of the RSMPA-DLFCC approach over existing DL models, with a maximum accuracy of 98.22%. In future work, real-time remote sensing data will be a priority, enabling the model to adapt dynamically to changing crop conditions and emerging threats. Moreover, future work can focus on the integration of multi-modal data sources, such as thermal imaging or hyperspectral data, and will broaden the scope of crop classification, providing a more comprehensive understanding of crop health and types. Finally, field tests can be performed to assess the real-world performance and accuracy of the RSMPA-DLFCC technique in diverse agricultural settings and will be essential for its practical deployment and validation.

**Author Contributions:** Conceptualization, A.S.A. and H.A.M.; methodology, H.A.M.; software, M.K.S.; validation, A.S.A., H.A.M. and M.K.S.; formal analysis, A.M.; investigation, K.M.O.; resources, A.M.; data curation, A.S.A.; writing—original draft preparation, A.S.A., H.A.M., M.K.S., K.M.O., F.A.A. and A.M.; writing—review and editing, H.A.M., F.A.A. and M.K.S.; visualization, F.A.A.; supervision, H.A.M.; project administration, M.K.S.; funding acquisition, H.A.M. and M.K.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through large group Research Project under grant number (RGP2/117/44), Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R114), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. Research Supporting Project number (RSPD2023R838), King Saud University, Riyadh, Saudi Arabia. This study is partially funded by the Future University in Egypt (FUE).

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** Data sharing is not applicable to this article as no datasets were generated during the current study.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest. The manuscript was written with contributions of all authors. All authors have given approval to the final version of the manuscript.

## References

- Joshi, A.; Pradhan, B.; Gite, S.; Chakraborty, S. Remote-Sensing Data and Deep-Learning Techniques in Crop Mapping and Yield Prediction: A Systematic Review. *Remote Sens.* **2023**, *15*, 2014. [[CrossRef](#)]
- Bouguettaya, A.; Zarzour, H.; Kechida, A.; Taberkit, A.M. Deep learning techniques to classify agricultural crops through UAV imagery: A review. *Neural Comput. Appl.* **2022**, *34*, 9511–9536. [[CrossRef](#)] [[PubMed](#)]
- Zhao, H.; Duan, S.; Liu, J.; Sun, L.; Reymondin, L. Evaluation of five deep learning models for crop type mapping using sentinel-2 time series images with missing information. *Remote Sens.* **2021**, *13*, 2790. [[CrossRef](#)]
- Orynbaikyzy, A.; Gessner, U.; Conrad, C. Crop type classification using a combination of optical and radar remote sensing data: A review. *Int. J. Remote Sens.* **2019**, *40*, 6553–6595. [[CrossRef](#)]
- de Azevedo, R.P.; Dallacort, R.; Boechat, C.L.; Teodoro, P.E.; Teodoro, L.P.R.; Rossi, F.S.; Correia Filho, W.L.F.; Della-Silva, J.L.; Baio, F.H.R.; Lima, M.; et al. Remotely sensed imagery and machine learning for mapping of sesame crop in the Brazilian Midwest. *Remote Sens. Appl. Soc. Environ.* **2023**, *32*, 101018. [[CrossRef](#)]
- Wang, L.; Wang, J.; Liu, Z.; Zhu, J.; Qin, F. Evaluation of a deep-learning model for multispectral remote sensing of land use and crop classification. *Crop J.* **2022**, *10*, 1435–1451. [[CrossRef](#)]
- Dash, R.; Dash, D.K.; Biswal, G.C. Classification of crop based on macronutrients and weather data using machine learning techniques. *Results Eng.* **2021**, *9*, 100203. [[CrossRef](#)]
- Kuang, X.; Guo, J.; Bai, J.; Geng, H.; Wang, H. Crop-Planting Area Prediction from Multi-Source Gaofen Satellite Images Using a Novel Deep Learning Model: A Case Study of Yangling District. *Remote Sens.* **2023**, *15*, 3792. [[CrossRef](#)]
- Suchi, S.D.; Menon, A.; Malik, A.; Hu, J.; Gao, J. Crop identification based on remote sensing data using machine learning approaches for fresno county, California. In Proceedings of the 2021 IEEE Seventh International Conference on Big Data Computing Service and Applications (BigDataService), Oxford, UK, 23–26 August 2021; pp. 115–124.
- Kwak, G.H.; Park, N.W. Unsupervised domain adaptation with adversarial self-training for crop classification using remote sensing images. *Remote Sens.* **2022**, *14*, 4639. [[CrossRef](#)]
- Seydi, S.T.; Amani, M.; Ghorbanian, A. A dual attention convolutional neural network for crop classification using time-series Sentinel-2 imagery. *Remote Sens.* **2022**, *14*, 498. [[CrossRef](#)]
- Reedha, R.; Dericquebourg, E.; Canals, R.; Hafiane, A. Transformer neural network for weed and crop classification of high resolution UAV images. *Remote Sens.* **2022**, *14*, 592. [[CrossRef](#)]
- Kordi, F.; Yousefi, H. Crop classification based on phenology information by using time series of optical and synthetic-aperture radar images. *Remote Sens. Appl. Soc. Environ.* **2022**, *27*, 100812. [[CrossRef](#)]
- Sun, Y.; Yao, N.; Luo, J.; Leng, P.; Liu, X. A spatiotemporal collaborative approach for precise crop planting structure mapping based on multi-source remote-sensing data. *Int. J. Remote Sens.* **2023**, 1–17. [[CrossRef](#)]
- Li, H.; Zhang, C.; Zhang, Y.; Zhang, S.; Ding, X.; Atkinson, P.M. A Scale Sequence Object-based Convolutional Neural Network (SS-OCNN) for crop classification from fine spatial resolution remotely sensed imagery. *Int. J. Digit. Earth* **2021**, *14*, 1528–1546. [[CrossRef](#)]
- Zhai, P.; Li, S.; He, Z.; Deng, Y.; Hu, Y. Collaborative mapping rice planting areas using multisource remote sensing data. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 5969–5972.

17. Sun, Y.; Luo, J.; Wu, T.; Zhou, Y.N.; Liu, H.; Gao, L.; Dong, W.; Liu, W.; Yang, Y.; Hu, X.; et al. Synchronous response analysis of features for remote sensing crop classification based on optical and SAR time-series data. *Sensors* **2019**, *19*, 4227. [[CrossRef](#)] [[PubMed](#)]
18. You, H.; Lu, Y.; Tang, H. Plant disease classification and adversarial attack using SimAM-EfficientNet and GP-MI-FGSM. *Sustainability* **2023**, *15*, 1233. [[CrossRef](#)]
19. Baştumur Kaya, C. A Novel Hybrid Method Based on the Marine Predators Algorithm and Adaptive Neuro-Fuzzy Inference System for the Identification of Nonlinear Systems. *Symmetry* **2023**, *15*, 1765. [[CrossRef](#)]
20. Zhou, S.; Tan, B. Electrocardiogram soft computing using hybrid deep learning CNN-ELM. *Appl. Soft Comput.* **2020**, *86*, 105778. [[CrossRef](#)]
21. Rineer, J.; Beach, R.; Lapidus, D.; O'Neil, M.; Temple, D.; Ujeneza, N.; Cajka, J.; Chew, R. Drone Imagery Classification Training Dataset for Crop Types in Rwanda. Version 1.0, Radiant MLHub. 2021. Available online: [https://mlhub.earth/data/rti\\_rwanda\\_crop\\_type](https://mlhub.earth/data/rti_rwanda_crop_type) (accessed on 13 June 2023).
22. Ahmed, M.A.; Aloufi, J.; Alnatheer, S. Satin Bowerbird Optimization with Convolutional LSTM for Food Crop Classification on UAV Imagery. *IEEE Access* **2023**, *11*, 41075–41083. [[CrossRef](#)]
23. Chew, R.; Rineer, J.; Beach, R.; O'Neil, M.; Ujeneza, N.; Lapidus, D.; Miano, T.; Hegarty-Craver, M.; Polly, J.; Temple, D.S. Deep neural networks and transfer learning for food crop identification in UAV images. *Drones* **2020**, *4*, 7. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.