Received: August 6, 2023. Revised: September 16, 2023. 610

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

# A Novel Shadow Removal Adversarial Network with Multi-scale ROI-based Deep Model for Classification of Fruit Diseases

**Kavitha Shanmugam[1]*** **Sarojini Kaliyamoorthi[1]**

[1]*Department of Computer Science, L.R.G. Government Arts College for Women, Tiruppur, Tamil Nadu, India.*
* Corresponding author's Email: kavithas3phd@gmail.com

**Abstract:** Due to the variety of fruit disorders, food production is highly impacted, resulting in economic degradation. So, proper disease recognition and classification system are vital to reduce agricultural losses and encourage economic growth. For this purpose, a multi-scale region-of-interest (ROI)-based deep learning (MRDL) model was developed, which localizes and classifies multiple fruit disorders based on the new attention principle and feature representation. But, the shadow regions in the fruit pictures influence the model efficiency and the picture dataset was highly dynamic while adding new classes for different disorders; because collecting diseased class labels for a single fruit picture was not simple. Hence in this article, a novel cycle-generative adversarial network (CycleGAN) model is integrated with the MRDL to generate shadow-free fruit pictures and increase the accuracy of classifying multiple classes of fruit infections. Initially, an attribute registration and shadow removal adversarial network (ArSAN) is proposed to generate various fruit pictures without shadow areas. In this ArSAN, the attributes of objects and shadows in fruit pictures are initially obtained. Then, those are labeled by the attribute enrolment scheme with the CycleGAN. Based on the labeled attributes, the filtering weights are set to all pictures in the real picture dataset. Also, a similarity finding scheme is used to discover the best-matched pictures for the generator network. As a result, multiple generators are trained to produce the shadow-free fruit picture dataset with diverse cleanness characteristics. Moreover, the newly generated pictures are passed to the MRDL model for localizing the infected fruit areas and classifying the multiple disease classes. Finally, the experimental outcomes reveal that the ArSAN-MRDL model attains a 93.8%, 93.78% and 93.76% accuracy on apple, citrus and tomato disorder picture sets, respectively contrasted to the MRDL, artificial neural network (ANN), 13-layer convolutional neural network (CNN), L2MXception, CycleGAN-DenseNet, and conditional GAN (cGAN)-VGG16 models.

**Keywords:** Fruit pathogens, Multi-scale deep learning, Localization, Region-of-interest, CycleGAN, Shadow elimination, Attribute enrolment.

## 1. Introduction

Fruit cultivation is India's economic foundation, employing 52% of the population and contributing 18.5% of GDP [1-2]. Global warming increases fruit illnesses, necessitating diagnosis and classification to boost crop productivity and wealth creation [3]. Advanced technology improves disease classification effectiveness [4-5]. Automatic systems use various image processing and machine learning algorithms to improve fruit visuals and identify infections [6-10]. However, evaluating each feature is difficult and ambiguous [11]. DL models have outperformed conventional algorithms in classifying and identifying fruit diseases.

Convolutional neural networks (CNN) have shown significant efficiency in identifying and classifying fruit illnesses [12]. However, most CNN frameworks generate single-class labels, making it difficult to accurately recognize diseased areas in a single picture [13].

A multi-scale deep learning (MDL) model-based semantic partition [14] was developed that uses a deep encoder-decoder network, adaptive receptive field module (ARFM), bottleneck block (BB), depth-wise residual blocks (DRBs), and SE-ResNet modules for disease feature maps. However, the model struggles to distinguish distinct fruit samples.

Also, classifying multiple diseases in a single fruit picture was challenging.

So, the MRDL framework [15] was designed to improve the interpretational characteristics of fruit illness pictures by adopting a new attention principle. It performs classification and localization, adjusting the receptive field scale to accommodate visual characteristics at different resolutions. The main difference between MDL and MRDL is that MDL uses SE-ResNet to discover infected areas, while MRDL uses SE-ResNet to discover fruit phenotypic patterns. The dataset is highly dynamic, as adding new class types for multiple infections makes it difficult to obtain infected class labels for fruits. Random occurrences of fruit illnesses can cause inadequate pictures, making it complex to generate shadow-free pictures from shadowed pictures due to varying quality factors.

Therefore, this article introduces a novel CycleGAN model with the MRDL for fruit infection recognition and classification. The key goal of this study is to augment the infected fruit pictures and remove the shadow areas from the fruit pictures. This model involves 3 stages: (i) picture augmentation, (ii) feature mining and (iii) classification. In the initial stage, an ArSAN is proposed, which introduces an attribute registration strategy to CycleGAN for creating appropriate pictures without shadow regions. The main contributions of this paper are,

1. First, the attributes like the number, cleanness and class of objects in the given fruit pictures are extracted along with the shadow characteristics.
2. Then, the filtering weights are allocated to all pictures in the actual picture dataset according to the mined attributes.
3. Also, a similarity finding scheme is applied to find the best-matched pictures for the generator network.
4. Moreover, this novel network is utilized to train many generators and create a shadow-free fruit picture dataset with various cleanness characteristics.
5. Further, the created picture dataset is fed to the MRDL model, which performs the second and third stages, i.e. feature mining and classification stages are conducted to localize the infected areas and classify the multiple types of fruit illnesses simultaneously.

Thus, this model can increase the accuracy of classifying multiple fruit illnesses by creating an adequate number of shadow-free fruit pictures. The remaining manuscript is prepared as follows: Section 2 covers the literature survey. Section 3 describes the ArSAN-MRDL model and section 4 analyzes its performance. Section 5 concludes the study.

## 2. Literature survey

Lightweight CNN [16] with a new data augmentation scheme called “9 Angle Crop” was presented to classify oil palm fresh fruit bunch based on their maturity levels. But the accuracy was not effective due to the limited training pictures. An enhanced Xception named L2MXception [17] was developed, which ensembles normalization terms of L2-norm and mean to categorize the peach illnesses. However, the precision and recall were less because of limited and poor-quality training pictures. The wavelength selection schemes [18] were developed for classifying the infected and healthy mango fruits. But its accuracy was less since it did not capture complex features from the diseased fruit pictures.

The CNN [19] was used for classifying the degree of maturity of the different fruits. But its f-score was ineffective because of using images with poor illumination and backgrounds. To recognize the diseased and healthy apple fruits and leaves, a novel multi-scale dense categorization system [20] with Cycle-GAN was developed. But the accuracy was affected by the fruit pictures’ shadow backgrounds. A conditional GAN (cGAN) [21] with VGG16 was developed to create and classify healthy and diseased fruit pictures. But the background in the fruit pictures may impact the model accuracy.

A novel method [22] was developed to enhance the quality of apple fruit images for identifying and classifying apple diseases. Also, the sweetness of an apple was predicted based on the leaf and texture of the fruit. Then, the GLCM and HOG features were extracted and fused. The fused feature was fed to the support vector machine (SVM) classifier to categorize apple diseases. But the number of images was limited, resulting in poor accuracy, precision and recall.

Citrus disease identification using CNN-based feature extraction and softmax classifier [23] was developed using the hyperspectral images. But the misclassification rate was still high and accuracy was less because wind scar, insect damage, and scab diseases appeared spectrally similar.

A modified CNN [24] was developed for automated hawthorn fruit classification. First, the hawthorn images were pre-processed using Otsu thresholding to segment the hawthorn fruit from the background. Then, the image dataset was enhanced by the data augmentation and the hawthorn fruit images were classified by the modified CNN into

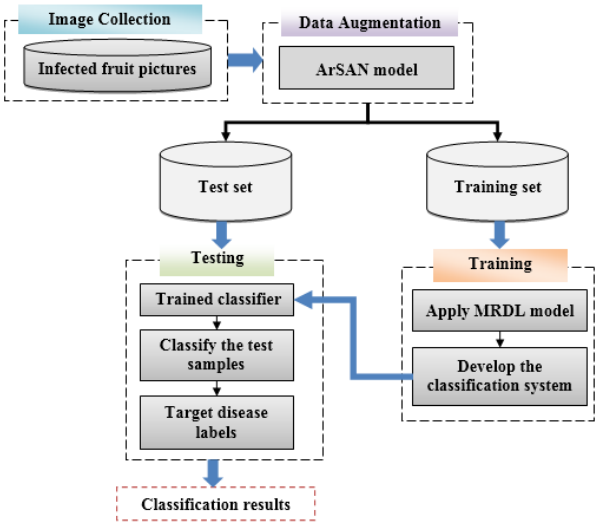


Figure. 1 Overview of ArSAN-MRDL-based fruit disease recognition and classification system

Table 1. Lists of notations

| Notations | Description |
|---|---|
| $i$ | Actual fruit picture |
| $N_i, C_i, F_i$ | Number, class and cleanness of objects in fruit image, respectively |
| $L_R^d, L_G^d, L_B^d$ | Weight of Red, Green and Blue channels for fruit images |
| $L_{R,G,B}^d$ | Range of the direct illumination of RGB color channels |
| $R_{R,G,B}$ | Reflectance |
| $S_i$ | Shadow characteristics of a given fruit image |
| $CV_i$ | Characteristic vector for fruit image |
| $N_q, C_q, F_q, S_q$ | Number, class, cleanness and shadow characteristics of objects in query image, respectively |
| $CV_q, CV_d$ | Characteristic vector for an input query image and created desired image, respectively |
| $F_d$ | Desired cleanness of objects |
| $W_{iq}$ | Filtering weight between $i^{th}$ picture and the input query picture |
| $A$ | Configuration vector |
| $W_{id}$ | Filtering weight between $i^{th}$ picture and the created desired picture |
| $\rho$ | Random variable |
| $\mathcal{G}, \mathcal{D}$ | Generator and discriminator networks, respectively |
| $\mathbb{E}$ | Expectation operator |
| $z$ | Random noise input |
| $\mathcal{G}(z)$ | Picture created by $\mathcal{G}$ network from $z$ |
| $\mathcal{D}(i)$ | Probability of $i$ belonging to the actual picture |
| $p_{image}(i)$ | Picture distribution |
| $p_z(z)$ | Data from latent space $z$ |

different kinds of diseases. But the classification accuracy was less due to the shadow and background features in the fruit pictures.

Gai et al. [25] developed an enhanced YOLOv4 with a DenseNet model to recognize cherry fruits. But this model was not appropriate for fruits with large volumes and accuracy may degrade when the quality of fruit pictures is poor.

Compared to the above-studied models, the proposed ArSAN-MRDL model can increase the number of training images by enhancing the visual quality of fruit pictures without shadow and background properties for effective classification of multiple fruit diseases.

## 3. Proposed methodology

In this section, the ArSAN-MRDL framework is explained. An overview of this study is illustrated in Fig. 1. Initially, fruit pictures having various infections are acquired from the publicly accessible datasets. Those acquired pictures are then applied to the ArSAN for increasing the number of pictures, which are split into training and test sets. The training sets are fed to the MRDL classifier [15] for learning and creating a learned classifier. Such a learned classifier model is further considered to categorize the test sets into the different types of fruit infections. The notations used in this study are given in Table 1.

### 3.1 Dataset description

1. Apple disease database [26]: It comprises a picture gallery of healthy and diseased apple fruits. The diseases targeted in this database are Blotch_Apple, Normal_Apple, Rot_Apple, and Scab_Apple. There are 116 Blotch_Apple pictures, 67 Normal_Apple pictures, 114 Rot_Apple pictures and 85 Scab_Apple pictures available for training. Similarly, 30 Blotch_Apple pictures, 24 Normal_Apple pictures, 38 Rot_Apple pictures and 28 Scab_Apple pictures available for testing.

2. Citrus disease database [27]: It contains the pictures of non-diseased and diseased citrus fruits. The diseases involved in this database are the blackspot, canker, scab, greening and melanosis. Every picture was tagged by the field specialist Dr. Basharat ALi Saleem for defining all categories like black spot, canker, greening, scab and non-diseased pictures.

3. Tomato disease database [28]: Pictures of 10 tomato infections from various open sources are chosen such as tomato malformed fruit, tomato blotchy ripening, tomato puffy fruit, tomato dehiscent fruit, tomato blossom-end rot, tomato sunscald, tomato virus disease, tomato gray mold,

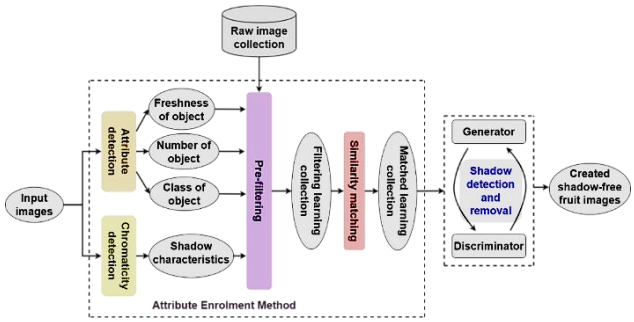


Figure. 2 Processes in the learning phase of ArSAN model

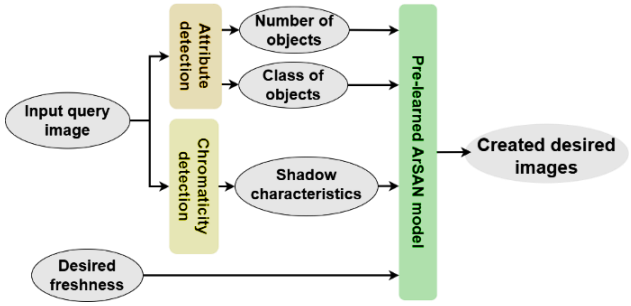


Figure. 3 Processes in the testing phase of ArSAN model

tomato ulcer disease and tomato anthracnose. Additionally, pictures of non-diseased tomatoes, a database of 11 categories of, with an overall of 286, tomato pictures are considered.

### 3.2 Picture augmentation based on ArSAN model

This ArSAN creates the desired fruit pictures with precise cleanness (freshness) characteristics depending on the given query fruit pictures. It adopts the method of attribute enrolling into CycleGAN to enhance learning efficiency. The structure of ArSAN involves 4 major components: (i) the attribute enrolling unit, which retrieves the proper characteristics of the objects in pictures and the shadow properties in the given fruit pictures by learning the correlation among color channels, (ii) the pre-filtering unit, which chooses suitable picture subgroups from the actual picture collection depending on the characteristics of the input query pictures, (iii) the similarity matching unit, which assigns the matching rules for the filtered learning pictures of CycleGAN and (iv) a generator-discriminator unit (also known as shadow detection and removal unit). The processes in the learning phase of the ArSAN model are presented in Fig. 2. The pre-filtering selects the pictures from the actual collection. The similarity matching is utilized to couple the pictures in the filtered learning collection. The picture creation network is learned depending on the matched learning collection.

Likewise, the processes in the test phase of the

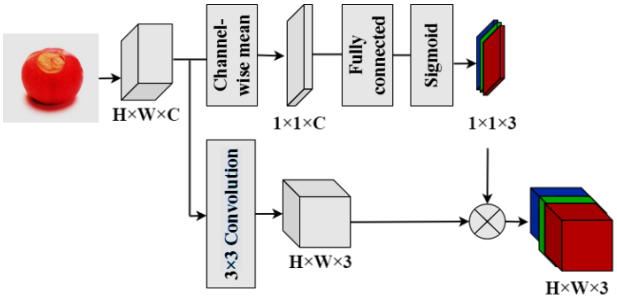


Figure. 4 Chromaticity detection network for shadow characteristics extraction

ArSAN are presented in Fig. 3. For an input query picture, the related desired picture is acquired by choosing an appropriate pre-learned network based on the number and class of objects in the query pictures including the desired cleanness and shadow (color intensity) characteristics of the objects.

#### 3.2.1. Attribute enrolment unit

This unit encompasses the attribute detection scheme, which detects the characteristics of fruit pictures such as cleanness, number and class of objects. For pictures in the original collection, the number $N_i$ and the class $C_i$ are obtained by the YOLO-v2 structure, whereas the cleanness $F_i$ of objects is determined using the AlexNet structure. AlexNet is chosen because of its basic architecture and less computation cost. Also, YOLO-v2 may accomplish a greater accuracy when sustaining a tolerable recognition speed. If many objects exist in a picture, the class of objects with the maximum numbers is called the class of the picture, whereas the cleanness of objects with the maximum numbers is called the cleanness of the picture.

Additionally, the shadow characteristics of the given pictures are captured by the chromaticity detection network (as shown in Fig. 4). This unit is used for weighting all color channels with a certain value as the absolute task and it may be represented by predicting $L_R^d, L_G^d$, and $L_B^d$ for all pictures. By predicting the weight of all channels in the fully connected unit, the range of the direct illumination of RGB color channels $\left(L_{R,G,B}^d\right)$ is predicted when being impacted by every other channel. This network is useful to learn the correlations among color channels since the convolutional unit can't effectively learn the correlations among channels.

By learning $L_{R,G,B}^d$ and the reflectance $\left(R_{R,G,B}\right)$ separately, it is possible that this system is learned as a stable prediction system for alterations in picture color. The predicted $L_{R,G,B}^d$ and $R_{R,G,B}$ values are characterized as the shadow characteristics $S_i$ of a given picture. According to these outcomes, the

characteristic vector $CV_i = (N_i, C_i, F_i, S_i)$ is created for all pictures in the actual picture collection.

Similarly, for an input query picture, the number $N_q$, the class $C_q$, the cleanness $F_q$ of objects and the shadow properties $S_q$ are determined. The characteristic vector of an input query picture is created as $CV_q = (N_q, C_q, F_q, S_q)$. As well, the characteristic vector of the created desired picture is obtained as $CV_d = (N_q, C_q, F_d, S_q)$ based on the characteristics of the input query picture, the desired cleanness $F_d$ and the shadow properties $S_q$.

#### 3.2.2. Characteristic homogeneity-based picture pre-filtering

The filtering weight of all pictures in actual picture collection is determined according to its characteristic vector $CV_i$. The determination of filtering weight is given as:

$$W_{iq} = \|A \cdot (CV_i - CV_q)\| = \alpha_1 \times (N_i - N_q)^2 + \alpha_2 \times (C_i - C_q)^2 + \alpha_3 \times (F_i - F_q)^2 + \alpha_4 \times (S_i - S_q)^2 \quad (1)$$

$$W_{id} = \|A \cdot (CV_i - CV_d)\| = \alpha_1 \times (N_i - N_q)^2 + \alpha_2 \times (C_i - C_q)^2 + \alpha_3 \times (F_i - F_d)^2 + \alpha_4 \times (S_i - S_q)^2 \quad (2)$$

In Eq. (1), $W_{iq}$ refers to the filtering weight between $i^{th}$ picture and the input query picture, $A = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ denotes the configuration vector, which is assigned by users i.e., $\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$. In Eq. (2), $W_{id}$ indicates the filtering weight between $i^{th}$ picture and the created desired picture. The filtering weight of all pictures in the actual picture collection can differ with various input query pictures and desired cleanness.

Suitable pictures are filtered from the actual picture collection based on the filtering weight of all pictures and a few uncertainties (randomness). The randomness defines the random generation of the picture from the domain, which results in difficult to find the correlation between the points in the latent space to the created pictures. So, the mixture of filtering weight and uncertainty is to handle high quality and the variety of created desired pictures. Moreover, the filtered pictures are together creating a filtered learning collection.

**Algorithm: Similarity matching algorithm for healthy/infected fruit pictures**

**Input:** Input query picture, created desired picture and actual picture collection

**Output:** Correlation among pictures

1. Resize the picture dimension to $8 \times 8$ pixels;
2. Transform the resized RGB picture to a 64-level grayscale picture;
3. Determine the mean value of each luminance values of the precedent picture matrix;
4. Compare each element (pixel) of the picture matrix and the determined mean value;
5. Obtain the new binary matrix with 64 elements, where 1 denotes the intensity of the pixel is higher than or identical to the mean, and 0 or else;
6. Create the vector from the resulting binary matrix, initiating from the top left and going to the bottom right to get a 64-bit hash value;
7. Compare the resulting hash with other pictures hashes to determine the similarity based on the distance metric between the 2 hashes.

#### 3.2.3. Similarity matching-based learning collection pairing

The filtered pictures can be paired according to the similarity recognition. A perceptual hash algorithm is used for picture similarity recognition, which is described below:

#### 3.2.4. Generator unit

The network architecture of the generator is built depending on the CycleGAN network as depicted in Fig. 5. The layer configuration of generator is presented in Table 2.

Table 2. Layer configuration of generator in ArSAN model

| Layer | Convolution kernel dimension | No. of convolution kernels | Step length |
|---|---|---|---|
| Input layer | – | – | – |
| Layer 1 | $7 \times 7$ | 32 | 1 |
| Layer 2 | $3 \times 3$ | 64 | 2 |
| Layer 3 | $3 \times 3$ | 128 | 2 |
| Layer 4 – Layer 12 | $3 \times 3$ | 128 | 1 |
| Layer 13 | $3 \times 3$ | 64 | 12 |
| Layer 14 | $3 \times 3$ | 32 | 12 |
| Layer 15 | $7 \times 7$ | 3 | 1 |

Layer 0 is the input layer, whose input is the healthy/infected fruit pictures with a dimension of 256×256. From layer 1 to layer 3, 3 Convolution-InstanceNorm-ReLU units are implemented with varying step lengths like 1, 2 and 2, correspondingly. These layers aim to perform dimensional-reduction

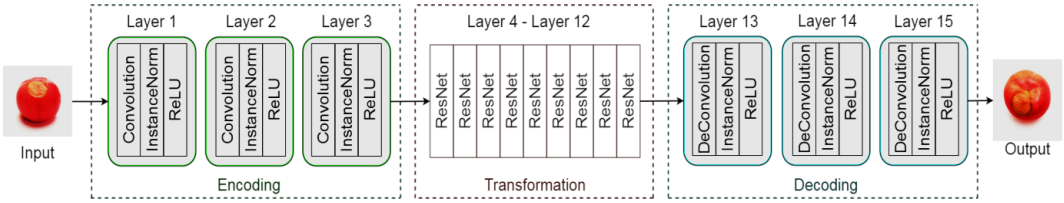


Figure. 5 Generator architecture in ArSAN model

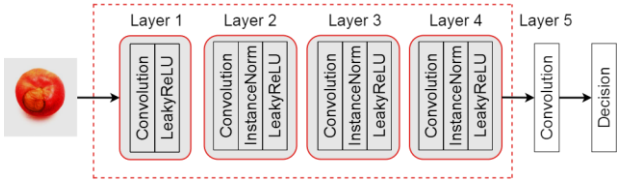


Figure. 6 Discriminator architecture in ArSAN model

Table 3. Layer configuration of discriminator in ArSAN model

| Layer | Convolution kernel dimension | No. of convolution kernels | Step length |
|---|---|---|---|
| Layer 1 | $4 \times 4$ | 64 | 2 |
| Layer 2 | $4 \times 4$ | 128 | 2 |
| Layer 3 | $4 \times 4$ | 256 | 2 |
| Layer 4 | $4 \times 4$ | 512 | 1 |
| Layer 5 | $4 \times 4$ | 1 | 1 |

sampling in the attribute space. From layer 4 to layer 12, nine continuous ResNet parts are built sequentially. Finally, 2 upsampling parts are applied such as transpose convolution to get the space and channel back to the entire dimension.

#### 3.2.5. Discriminator unit

The discriminator architecture of the ArSAN model is developed using the PatchGAN, which involves 5 layers as portrayed in Fig. 6.

The initial layer is the convolution-LeakyReLU. layer 2 to layer 4 includes 3 Convolution-InstanceNorm-ReLU layers. The 5th layer is the convolutional unit that creates a 1D outcome. The layer configuration of the discriminator is presented in Table 3.

LeakyReLU is chosen as the activation function of the discriminator and defined as:

$$LeakyReLU(I) = \begin{cases} i, & i \geq 0 \\ \rho i, & i < 0 \end{cases} \in R \quad (3)$$

In Eq. (3), $\rho$ is set as 0.2. During the learning task, the generator $\mathcal{G}$ aims to create pictures as actual as potential to deceive the discriminator $\mathcal{D}$, whereas $\mathcal{D}$ aims to differentiate the pictures created by $\mathcal{G}$ from the actual pictures as far as promising. The fitness factor is defined as:

$$\min_{\mathcal{G}} \max_{\mathcal{D}} (\mathcal{D}, \mathcal{G}) = \mathbb{E}_{i \sim p_{image}(i)} [\log \mathcal{D}(i)] + \mathbb{E}_{z \sim p_z(z)} \left[ log \left( 1 - \mathcal{D}(\mathcal{G}(z)) \right) \right] \quad (4)$$

In Eq. (4), $\mathbb{E}$ is the expectation operator, $i$ is the actual fruit picture, $z$ is the noise of the input $\mathcal{G}$ network, $\mathcal{G}(z)$ is the picture created by $\mathcal{G}$ network from a random noise input $z$ and $\mathcal{D}(i)$ is the probability of $i$ belonging to the actual picture, $p_{image}(i)$ is the picture distribution, $p_z(z)$ is the data from latent space $z$. Here, $\mathcal{D}$ is trained to maximize the probability $\log \mathcal{D}(i)$ and $\mathcal{G}$ is trained to minimize $log \left( 1 - \mathcal{D}(\mathcal{G}(z)) \right)$.

Thus, the ArSAN model creates the shadow-less fruit pictures efficiently according to the desired cleanness. After completing the ArSAN process, the created picture collections are concatenated with the actual sets to obtain the training and test samples.

The training samples are utilized to train the MRDL model [15] and the pre-learned MRDL model is later applied to classify the test samples into different fruit disease classes appropriately.

## 4. Experimental result

The performance of the ArSAN-MRDL framework is evaluated with the existing frameworks (such as MDL [14], MRDL [15], L2MXception [17], CycleGAN-DenseNet [20], cGAN-VGG16 [21], SVM [22], and modified CNN [24]) by executing all of them in MATLAB 2019a with the help of 3 distinct fruit picture sets (mentioned in Section 3.1).

To measure the performance of the proposed framework, the considered existing frameworks are also implemented and tested on the apple, citrus and tomato fruit picture datasets. Table 4 lists parameter settings for the ArSAN-MRDL and existing frameworks for fruit disease classification.

The following metrics are measured to compare the proposed and existing frameworks:

- Accuracy: It is the percentage of fruit pictures that were exactly classified among all the pictures analyzed.

$$Accuracy = \frac{True\ Positive\ (TP) + True\ Negative\ (TN)}{TP + TN + False\ Positive\ (FP) + False\ Negative\ (FN)} \quad (5)$$

Table 4. Parameter settings for existing and proposed ArSAN-MRDL framework

| Framework | Parameters | Range |
|---|---|---|
| SVM [22] | Kernel type | Linear |
| | Kernel degree | 2 |
| | Penalty | 0.1 |
| | Gamma | 0.01 |
| Modified CNN [24] | No. of convolutional layers | 8 |
| | Activation function | ReLU |
| | No. of pooling layers | 4 |
| | Learning rate | 0.0001 |
| | Batch size | 64 |
| | No. of epochs | 100 |
| L2MXception [17] | Learning rate | 0.001 |
| | Batch size | 64 |
| | Epoch | 50 |
| cGAN-VGG16 [21] | Learning rate | 0.002 |
| | Batch size | 64 |
| | Epochs | 100 |
| | Momentum | 0.9 |
| cycleGAN-DenseNet [20] | Learning rate | 0.0001 |
| | Batch size | 16 |
| | No. of epochs | 120 |
| | Dropout | 0.8 |
| | Decay rate | 0.95 |
| MDL [14], MRDL [15] and proposed ArSAN-MRDL | Learning rate | 0.0001 |
| | Batch size | 32 |
| | Momentum | 0.9 |
| | No. of epochs | 50 |
| | No. of iterations | 120 |

In Eq. (5), TP is the number of fruit pictures exactly categorized as healthy, TN is the quantity of fruit pictures exactly classified as related infection labels, FP is the quantity of unhealthy fruit pictures categorized as healthy and FN signifies the quantity of healthy fruit pictures categorized as diseased.

- Precision: It is calculated by

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

- Recall: It is calculated by

$$Recall = \frac{TP}{TP+FN} \quad (7)$$

- F-measure: It is determined by

$$F - measure = 2 \times \frac{Precision \cdot Recall}{Precision+Recall} \quad (8)$$

Fig. 7 portrays a few examples of the fruit pictures in various infection classes from the given datasets. Fig. 8 illustrates a few examples of the ArSAN model, i.e. shadow-less fruit pictures in various infection classes from the given datasets.

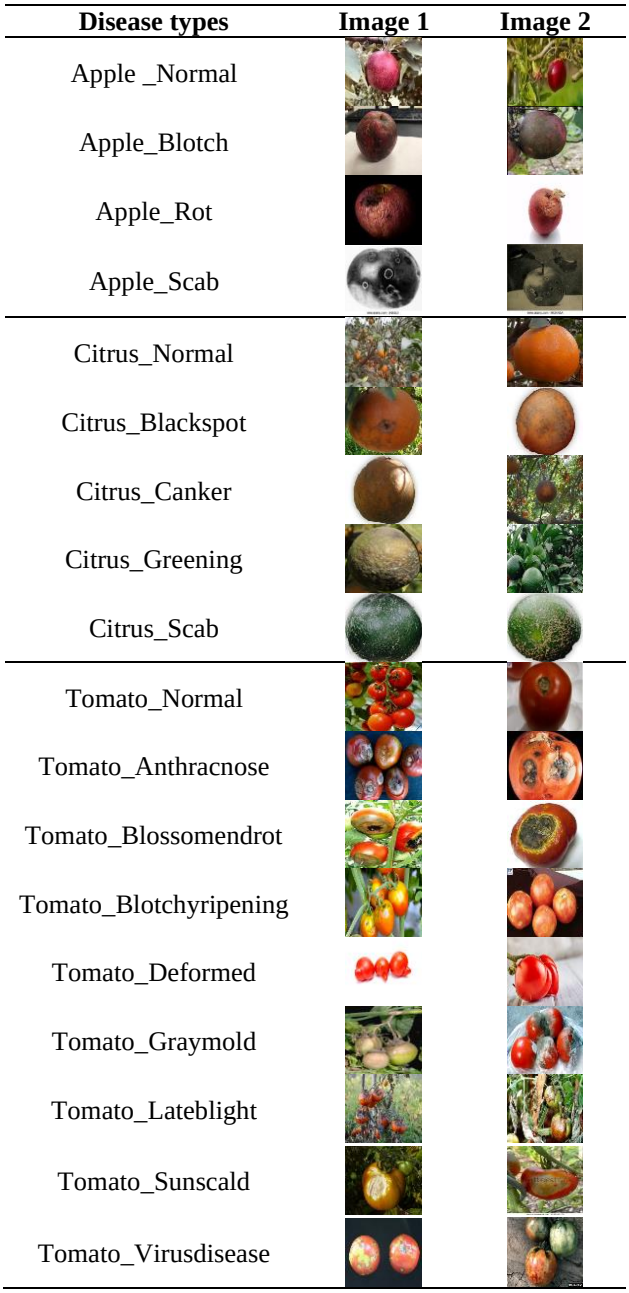


Figure. 7 Examples of various diseased fruit pictures applied in this study

Fig. 9 illustrates the precision of the different models tested using the apple, citrus and tomato fruit pictures. It exhibits that the precision of the ArSAN-MRDL on the apple fruit picture set is 17.18%, 12.08%, 10.42%, 7.84%, 3.33%, 2.63% and 1.8% higher than the SVM, L2MXception, modified CNN, MDL, MRDL, cGAN-VGG16 and CycleGAN-DenseNet, respectively. The precision of the ArSAN-MRDL model on the citrus fruit picture set is 18.16%, 13.14%, 11.4%, 8.69%, 3.82%, 2.47% and 1.46%, respectively higher than the SVM, L2MXception, modified CNN, MDL, MRDL, cGAN-VGG16 and CycleGAN-DenseNet. For the tomato fruit picture set, the precision of the ArSAN-MRDL is 16.62%, 11.7%,

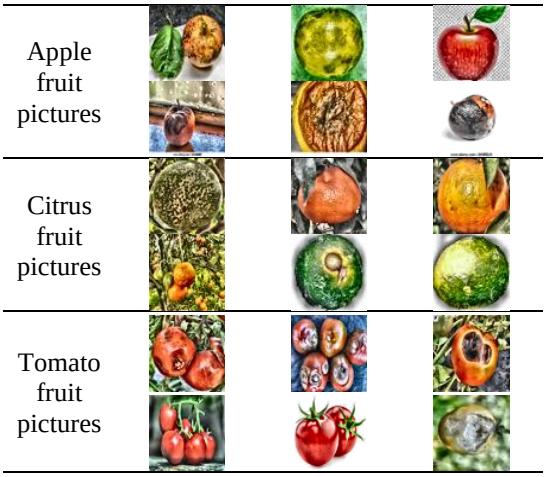


Figure. 8 Some examples of shadow-less diseased fruit pictures from ArSAN model

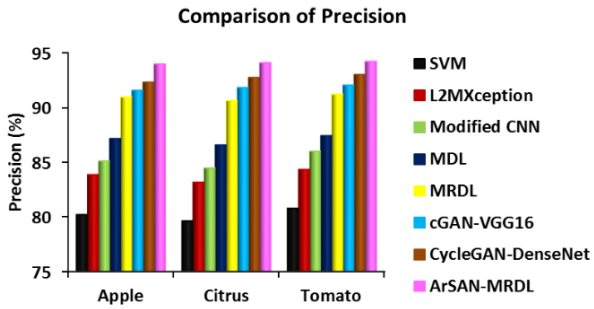


Figure. 9 Comparison of precision for ArSAN-MRDL and existing models on different fruit disease picture sets

9.55%, 7.77%, 3.29%, 2.36% and 1.28% higher than the SVM, L2MXception, modified CNN, MDL, MRDL, cGAN-VGG16 and CycleGAN-DenseNet, respectively.

Fig. 10 depicts the recall of the different models applied on the apple, citrus and tomato fruit pictures. It highlights that the recall of the ArSAN-MRDL model on the apple fruit picture set is 15.5%, 11.7%, 9.7%, 7.5%, 2.8%, 1.9% and 1.2% greater than the SVM, L2MXception, modified CNN, MDL, MRDL, cGAN-VGG16 and CycleGAN-DenseNet models, respectively. The recall of the ArSAN-MRDL model on the citrus fruit picture set is 14.85%, 12.34%, 10.01%, 7.94%, 3.03%, 2.14% and 1.18%, respectively greater than the SVM, L2MXception, modified CNN, MDL, MRDL, cGAN-VGG16 and CycleGAN-DenseNet. For the tomato fruit picture set, the recall of the ArSAN-MRDL model is 13.73%, 10.85%, 9.27%, 6.94%, 2.47%, 1.71% and 0.95%, respectively greater than the SVM, L2MXception, modified CNN, MDL, MRDL, cGAN-VGG16 and CycleGAN-DenseNet models.

In Fig. 11, the f-measure of various models implemented on the different fruit picture sets is illustrated. It observes that the f-measure of the ArSAN-MRDL model on the apple fruit picture set is

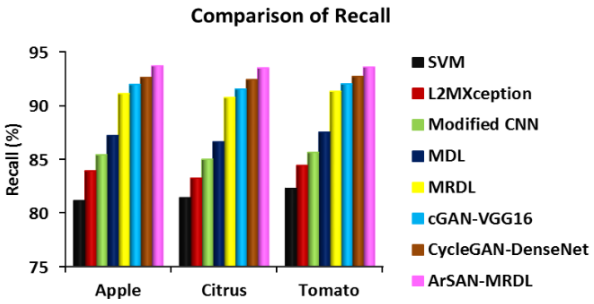


Figure. 10 Comparison of recall for ArSAN-MRDL and existing models on different fruit disease picture sets

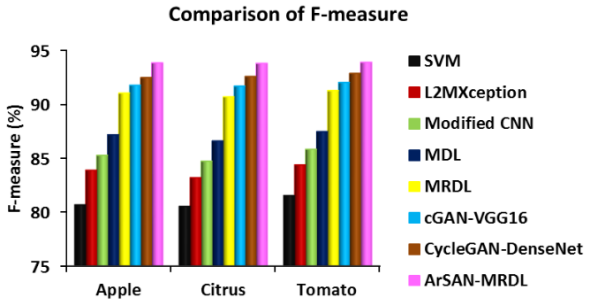


Figure. 11 Comparison of f-measure for ArSAN-MRDL and existing models on different fruit disease picture sets

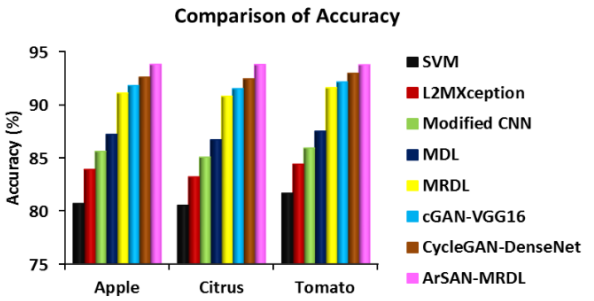


Figure 12. Comparison of accuracy for ArSAN-MRDL and existing models on different fruit disease picture sets

16.32%, 11.87%, 10.06%, 7.65%, 3.09%, 2.27% and 1.48% larger than the SVM, L2MXception, modified CNN, MDL, MRDL, cGAN-VGG16 and CycleGAN-DenseNet models, respectively. The f-measure of the ArSAN-MRDL model on the citrus fruit picture set is 16.48%, 12.74%, 10.7%, 8.31%, 3.42%, 2.3% and 1.32%, respectively larger than the SVM, L2MXception, modified CNN, MDL, MRDL, cGAN-VGG16 and CycleGAN-DenseNet models. For the tomato fruit picture set, the f-measure of the ArSAN-MRDL model is 15.16%, 11.27%, 9.4%, 7.35%, 2.88%, 2.03% and 1.11% larger than the SVM, L2MXception, modified CNN, MDL, MRDL, cGAN-VGG16 and CycleGAN-DenseNet models.

Fig. 12 portrays the accuracy of the different models on multiple fruit picture sets. It addresses that the accuracy of the ArSAN-MRDL model on the apple fruit picture set is 16.23%, 11.8%, 9.57%, 7.57%, 2.96%, 2.18% and 1.3% better than the SVM,

L2MXception, modified CNN, MDL, MRDL, cGAN-VGG16 and CycleGAN-DenseNet models, respectively. The accuracy of the ArSAN-MRDL model on the citrus fruit picture set is 16.47%, 12.72%, 10.26%, 8.17%, 3.28%, 2.48% and 1.46% better than the SVM, L2MXception, modified CNN, MDL, MRDL, cGAN-VGG16 and CycleGAN-DenseNet models, respectively. For the tomato fruit picture set, the accuracy of the ArSAN-MRDL model is 14.82%, 11.09%, 9.12%, 7.15%, 2.36%, 1.76% and 0.86% better than the SVM, L2MXception, modified CNN, MDL, MRDL, cGAN-VGG16 and CycleGAN-DenseNet models, respectively.

Thus, it realizes that the ArSAN-MRDL model enhances the performance of recognizing and classifying the variety of fruit diseases. This is due to the adoption of ArSAN for augmenting the number of pictures by creating more shadow-less fruit pictures, which supports the effective learning of the MRDL classifier.

## 5. Conclusion

In this paper, the ArSAN-MRDL model was developed for fruit disease classification. The ArSAN model was used to create shadow-free fruit pictures with multiple infections based on the different characteristics. Then, the MRDL model was trained to localize affected areas and classify them into target illness labels. At last, the test findings proved that the ArSAN-MRDL model on apple, citrus and tomato disorder picture sets reached an accuracy of 93.8%, 93.78% and 93.76%, respectively, in contrast with the classical models. Thus, the overall accuracy of the ArSAN-MRDL model is about 2.87% superior to the MRDL model for multiple fruit disease classification.

## Conflict of interest

The authors declare no conflict of interest.

## Author contributions

Conceptualization, methodology, software, validation, Kavitha Shanmugam; formal analysis, investigation, Sarojini Kaliyamoorthi; resources, data curation, writing—original draft preparation, Kavitha Shanmugam; writing—review and editing, Kavitha Shanmugam; visualization, supervision, Sarojini Kaliyamoorthi.

## References

[1] J. Rout, "Role of Women and their Involvement in Indian Agriculture: A Comparative Study among Different Districts of Odisha", *Orissa Economic Journal*, Vol. 49, No. 1 & 2, pp. 233-236, 2017.

[2] B. Kumari, H. Ayad, C. Fatih, A. Matuka, P. Mishra, and G. K. Vani, "Impact of Agricultural Exports Segmentation on Economic Growth in India", *Indian Journal of Economics and Development*, Vol. 18, No. 1, pp. 163-168, 2022.

[3] G. S. Malhi, M. Kaur, and P. Kaushik, "Impact of Climate Change on Agriculture and Its Mitigation Strategies: A Review", *Sustainability*, Vol. 13, No. 3, pp. 1-21, 2021.

[4] C. Wan, İ. Kahramanoğlu, and V. Okatan, "Application of Plant Natural Products for the Management of Postharvest Diseases in Fruits", *Folia Horticulturae*, Vol. 33, No. 1, pp. 203-215, 2021.

[5] Z. Iqbal, M. A. Khan, M. Sharif, J. H. Shah, M. H. U. Rehman, and K. Javed, "An Automated Detection and Classification of Citrus Plant Diseases using Picture Processing Techniques: A Review", *Computers and Electronics in Agriculture*, Vol. 153, pp. 12-32, 2018.

[6] A. Cravero, S. Pardo, S. Sepúlveda, and L. Muñoz, "Challenges to Use Machine Learning in Agricultural Big Data: A Systematic Literature Review", *Agronomy*, Vol. 12, No. 3, pp. 1-34, 2022.

[7] K. Mochida, S. Koda, K. Inoue, T. Hirayama, S. Tanaka, R. Nishii, and F. Melgani, "Computer Vision-Based Phenotyping for Improvement of Plant Productivity: A Machine Learning Perspective", *GigaScience*, Vol. 8, No. 1, pp. 1-12, 2019.

[8] A. K. R. Chávez, J. A. Franco, A. A. F. Jaramillo, L. M. C. Medina, R. G. G. González, and Q. H. Escobedo, "Machine Learning for Plant Stress Modeling: A Perspective towards Hormesis Management", *Plants*, Vol. 11, No. 7, pp. 1-22, 2022.

[9] M. J. Feldmann, M. A. Hardigan, R. A. Famula, C. M. Lopez, A. Tabb, G. S. Cole, and S. J. Knapp, "Multi-Dimensional Machine Learning Approaches for Fruit Shape Phenotyping in Strawberry", *GigaScience*, Vol. 9, No. 5, pp. 1-17, 2020.

[10] Y. D. D. Sean, D Smith, V. S. P. Bitra, V. Bera, and N. Umar, "Development of Computer Vision System for Fruits", *Current Journal of Applied Science and Technology*, Vol. 40, No. 36, pp. 1-11, 2021.

[11] S. Ghazal, W. S. Qureshi, U. S. Khan, J. Iqbal, N. Rashid, and M. I. Tiwana, "Analysis of Visual Features and Classifiers for Fruit Classification Problem", *Computers and Electronics in Agriculture*, Vol. 187, pp. 1-9, 2021.

[12] M. Tripathi, "Analysis of Convolutional Neural Network Based Picture Classification Techniques", *Journal of Innovative Picture Processing*, Vol. 3, No. 02, pp. 100-117, 2021.

[13] P. Dhiman, V. Kukreja, P. Manoharan, A. Kaur, M. M. Kamruzzaman, I. B. Dhaou, and C. Iwendi, "A Novel Deep Learning Model for Detection of Severity Level of the Disease in Citrus Fruits", *Electronics*, Vol. 11, No. 3, pp. 495, 2022.

[14] T. Ilyas, A. Khan, M. Umraiz, Y. Jeong, and H. Kim, "Multi-Scale Context Aggregation for Strawberry Fruit Recognition and Disease Phenotyping", *IEEE Access*, Vol. 9, pp. 124491-124504, 2021.

[15] S. Kavitha and K. Sarojini, "Multi-Scale Region-Of-Interest based Deep Learning for Fruit Disease Identification and Classification", *ARPN Journal of Engineering and Applied Sciences*, Vol. 18, No. 4, pp. 358-368, 2023.

[16] G. N. Elwirehardja and J. S. Prayoga, "Oil Palm Fresh Fruit Bunch Ripeness Classification on Mobile Devices using Deep Learning Approaches", *Computers and Electronics in Agriculture*, Vol. 188, pp. 1-13, 2021.

[17] N. Yao, F. Ni, Z. Wang, J. Luo, W. K. Sung, C. Luo, and G. Li, "L2MXception: An Improved Xception Network for Classification of Peach Diseases", *Plant Methods*, Vol. 17, No. 1, pp. 1-13, 2021.

[18] A. Raghavendra, D. S. Guru, and M. K. Rao, "Mango Internal Defect Detection Based on Optimal Wavelength Selection Method using NIR Spectroscopy", *Artificial Intelligence in Agriculture*, Vol. 5, pp. 43-51, 2021.

[19] M. Rodriguez, F. Pastor, and W. Ugarte, "Classification of Fruit Ripeness Grades using a Convolutional Neural Network and Data Augmentation", In: *Proc. of IEEE Conf. of Open Innovations Association*, pp. 374-380, 2021.

[20] Y. Tian, E. Li, Z. Liang, M. Tan, and X. He, "Diagnosis of Typical Apple Diseases: a Deep Learning Method based on Multi-Scale Dense Classification Network", *Frontiers in Plant Science*, pp. 1-12, 2021.

[21] J. J. Bird, C. M. Barnes, L. J. Manso, A. Ekárt, and D. R. Faria, "Fruit Quality and Defect Picture Classification with Conditional GAN Data Augmentation", *Scientia Horticulturae*, Vol. 293, pp. 1-16, 2022.

[22] M. Kumar, Y. Pal, S. M. P. Gangadharan, K. Chakraborty, C. S. Yadav, H. Kumar, and B. Tiwari, "Apple Sweetness Measurement and Fruit Disease Prediction Using Image Processing Techniques Based on Human-Computer Interaction for Industry 4.0", *Wireless Communications and Mobile Computing*, Vol. 2022, pp. 1-12, 2022.

[23] P. K. Yadav, T. Burks, Q. Frederick, J. Qin, M. Kim, and M. A. Ritenour, "Citrus Disease Detection Using Convolution Neural Network Generated Features and Softmax Classifier on Hyperspectral Image Data", *Frontiers in Plant Science*, Vol. 13, pp. 1-17, 2022.

[24] R. Azadnia, S. Fouladi, and A. Jahanbakhshi, "Intelligent Detection and Waste Control of Hawthorn Fruit Based on Ripening Level Using Machine Vision System and Deep Learning Techniques", *Results in Engineering*, Vol. 17, pp. 1-13, 2023.

[25] R. Gai, N. Chen, and H. Yuan, "A Detection Algorithm for Cherry Fruits Based on the Improved YOLO-v4 Model", *Neural Computing and Applications*, Vol. 35, pp. 13895-13906, 2023.

[26] https://www.kaggle.com/kaivalyashah/apple-disease-detection

[27] https://data.mendeley.com/datasets/3f83gxmv57/2

[28] Q. Wang, F. Qi, M. Sun, J. Qu, and J. Xue, "Identification of Tomato Disease Types and Detection of Infected Areas Based on Deep Convolutional Neural Networks and Object Detection Techniques", *Computational Intelligence and Neuroscience*, pp. 1-16, 2019.