IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 14, 2021, accepted July 1, 2021, date of publication July 12, 2021, date of current version July 21, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3096550

Detection of Mulberry Ripeness Stages Using Deep Learning Models

SEYED-HASSAN MIRAEI ASHTIANI[1], SHIMA JAVANMARDI[2], MEHRDAD JAHANBANIFARD[2,3], ALEX MARTYNENKO[4], AND FONS J. VERBEEK[2]

[1]Department of Biosystems Engineering, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad 9177948974, Iran
[2]Imaging and Bioinformatics Section, Leiden Institute of Advanced Computer Science (LIACS), Leiden University, 2333 CA Leiden, The Netherlands
[3]Naturalis Biodiversity Center, 2333 CR Leiden, The Netherlands
[4]Department of Engineering, Faculty of Agriculture, Dalhousie University, Truro, NS B2N 5E6, Canada

Corresponding author: Shima Javanmardi (s.javanmardi@liacs.leidenuniv.nl)

The work of Mehrdad Jahanbanifard was supported in part by the European Union's Horizon 2020 Research and Innovation Programme through the Plant.ID under the Marie Sklodowska-Curie Grant 765000.

ABSTRACT Ripeness classification is one of the most challenging tasks in the postharvest management of mulberry fruit. The risks of microbial contamination and human error in manual sorting are significant; it may result in quality degradation and wasting of processed products. Due to advanced developments in computer vision and machine learning, automated sorting became possible. This study presents the results of developing and testing a computer vision-based application using convolutional neural networks (CNNs) for the classification of mulberry fruit ripening stages. To reduce the training cost and improve the accuracy of classification, transfer learning was used to fine-tune the CNN models. The CNN models in the test include DenseNet, Inception-v3, ResNet-18, ResNet-50, and AlexNet. Transfer learning was used to fine-tune the models and improve the accuracy of classification. The AlexNet and ResNet-18 networks exhibited the best performance with 98.32% and 98.65% overall accuracy for classifying the ripeness of white and black mulberries, respectively. Moreover, the performance of the models did not change when the data sets of both genotypes were mixed. The ResNet-18 was able to classify both genotype and ripeness from 600 fruit images in 2.36 min with an overall accuracy of 98.03%, which was superior to other architectures. It indicates that the model could be used for precise classification of the ripening stages of mulberries and other horticultural products, as a part of an automated sorting system.

INDEX TERMS Convolutional neural network, computer vision, online detection, ripening classification, transfer learning.

I. INTRODUCTION

Mulberry (*Morus* spp., *Moraceae* family) is one of the fruit species, widely distributed from temperate to subtropical zones of the northern hemisphere to the tropical zones of the southern hemisphere [1], [2]. Among the 24 known *Morus* cultivars, white mulberry (*Morus alba* L.), black mulberry (*Morus nigra* L.), and red mulberry (*Morus rubra* L.) are the most cultivated species in the world [3]–[5]. They are an excellent source of many nutritive compounds such as vitamins, minerals, polysaccharides, fatty acids, and amino acids [3], [6], as well as phenolic compounds including carotenoids, anthocyanins, flavonoids, and phenolic acids [7], [8] with the health-promoting and pharmacological effects, such as anti-cancer, anti-cholesterol, anti-inflammation, anti-diabetic, anti-aging, antioxidant, anti-obesity and neuroprotection [9], [10]. In addition to high nutritional value and bioactivity, the low acidity and very sweet taste of white mulberry and the slightly acidic flavor and attractive dark color of black mulberry led to a rapid increase in their production and consumption [11], [12]. The fully ripened fruits are usually consumed either fresh or used as ingredients in marmalade, tea, vinegar, wine, juice, jam, ice cream, jelly, syrup, food colorant, natural dyes, dried fruits, and other food and cosmetic products [4], [5], [13].

Fruit ripening is accompanied by the color change because of the pigment concentration in the fruit skin [4], [6].

The associate editor coordinating the review of this manuscript and approving it for publication was Turgay Celik.

FIGURE 1. Distribution of white (a) and black (b) mulberries with different ripening levels on a branch (Adapted from https://www.trees.com/mulberry-trees).

During growth period (∼ 25 to 30 d), mulberry skin color changes from green (unripe) to white, red, purple-black (fully ripe) [14]. Mulberry is a non-climacteric product, therefore, its harvest in the suitable ripening stage is highly significant from a nutritional and economic perspective [11].

Mulberry harvest operation is usually carried out by spreading a sheet under a tree and shaking the branches mechanically or manually [15], [16]. Due to alternate ripening patterns (Fig. 1), harvest includes a mixture of fruits at various ripening levels [14], [17]. This inconsistency of the product negatively affects its commercial value and marketability [18]. On the other hand, fresh mulberry is available only in the short term and it is hardly commercialized. Fresh, fully ripened fruits decay fast after harvesting due to the soft structure, high moisture content, respiration, growth, and proliferation of microorganisms on the surface [2], [3], [10]. Although their shelf-life can be extended by up to six weeks at cold storage, they cannot be preserved for a long time at ambient temperature [19]. Therefore, further processing is required to extend the availability and utilization of health benefits. Unfortunately, fruits with a low ripening index are discarded and wasted during quality control [11], [14]. As a result, the classification of fruits from a ripening stage perspective is essential for initial sorting. At this stage, unripen or unmarketable fruits could be separated from the batch for the next processing, while the best quality fresh fruits are delivered to consumers [1], [18].

The popularity of mulberry fruits among consumers and processing companies increases due to their recently discovered nutritional and nutraceutical value [7], [20]. Since the nutritional and functional compounds decrease and/or increase during ripening, a proper classification approach would extend the range of applications of mulberries in food and pharmaceutical industries [4], [14], [21]. Due to the good performance of deep learning models in classifying agricultural products, in this paper we consider using the deep neural networks to classify the mulberry fruit according to their ripeness.

A. PROBLEM STATEMENT

Unfortunately, the mulberry industry has significant problems with mulberry classification. Usually, identification of the ripening degree of mulberry fruits is done by trained personnel through visual inspection of fruit color. This process is exhausting, time-consuming, subjective, and costly [22], [23]. Alternatively, identification of ripening stages can be done by chemical or physicochemical methods [24]. Although these methods have better accuracy, they are time-consuming, costly, destructive, and sometimes require complex analytical equipment. They allow quantifying total flavonoids, anthocyanins, and total soluble solids (TSS), but are limited to a certain amount of samples, which is not suitable for automatic sorting systems [25], [26].

Therefore, the need for an automated non-destructive sorting system to increase fruit utilization and supply high-quality mulberry products to consumers is indispensable. Recently, smart analytical tools such as spectroscopy and spectral imaging, electronic noses, computer vision have been utilized to evaluate ripening levels of fruits [27], [28]. Some of them use machine learning and pattern recognition techniques [26], [29], [30]. Table 1 summarizes the relevant research for various horticultural products using the traditional machine learning algorithms and handcrafted features. Despite the progress in the classification, these applications are limited because of the following reasons:

1) The complexity of these methods due to the manual choice of features [31].
2) Inability to differentiate the subtle differences between subordinate classes [24].

TABLE 1. Comparative analysis of classical machine vision techniques applied for ripening stage classification.

Product	Feature	# of classes	Data Analysis	Accuracy %	References
Red and white mulberries	Color Texture Shape	3	ANN SVM	98.26-99.13	Azarmdel *et al.* [17]
Cape gooseberry	Color	7	ANN DT SVM KNN	85.90-92.65	Castro *et al.* [22]
Tomato	Color	3	BPNN	99.31	Wan *et al.* [23]
Banana	Color Texture Shape	4	NB LDA SVM	82.60-100	Zhuang *et al.* [25]
Papaya	Color	3	RDF	94.3	Pereira *et al.* [26]
Apple	Color	4	Hybrid ANN-GA	97.88	Sabzi *et al.* [27]
Apple	Color	4	SVM KNN ANN-GA ANN-PSO ANN-FA	92.56-97.16	Pourdarbani *et al.* [28]
Apple	Color	3	MDA	100	Cárdenas-Pérez *et al.* [29]
Yellow and red sweet peppers	Color Shape	4	RF LR	89.5-97.3	Harel *et al.* [30]

MDA: multivariate discriminant analysis; BPNN: backpropagation neural network; RDF: random decision forest; ANN: artificial neural network; DT: decision tree; SVM: support vector machine; KNN: K-nearest neighbors; NB: naïve Bayes; LDA: linear discriminant analysis; GA: genetic algorithm; PSO: particle swarm optimization; FA: firefly algorithm; RF: random forest; LR: logistic regression.

B. NOVELTY

Despite recent advancements in computer vision, only one research has been conducted to classify the ripening stages of mulberries [17]. In this work, traditional handcrafted features and two traditional machine learning classifiers, i.e., ANN and SVM, have been used. The good performance of deep learning networks in many studies in the classification of agricultural products prompted us to examine these models to classify the mulberry fruit according to their ripeness. Our study is the first-time attempt to solve the challenge of accurate detection of mulberries ripening stage by using deep learning. This approach is more accurate, enabling the classification of 4 ripening stages, which is more challenging compared to the previous research (3 ripening stages). It was demonstrated that deep learning can automatically extract features without human intervention and accurately classify the ripening stage of mulberries. Another novelty of this paper is the calculation of the time required to complete the classification process, which plays a key role in the design of smart mulberries sorting system. This is in contrast to [17], which did not examine the classification time.

C. RELATED WORK

So far, the application of deep learning models in precision agriculture have shown advantages for automatic feature extraction and learning, transfer learning, quick adaptation to a new problem, dealing with heterogeneous big data, and obtaining higher accuracy and excellent performance [31], [32]. Convolutional neural networks (CNN) and their derivatives have shown to be among the most successful techniques in image classification and recognition [24], [33].

Recently, several studies have applied CNNs for ripeness classification by analyzing RGB images of fruits. For example, Zhang *et al.* [24] designed a CNN structure for fine-grained classification of banana maturity. The proposed CNN achieved a 95.6% classification accuracy, which was higher than conventional strategies such as Gabor + SVM, Wavelet + SVM, Wavelet + Gabor + SVM, and combined features + SVM approaches. Another CNN model was developed to classify five stages of tomato maturity based on skin color, i.e., green, light pink, pink, light red, and red [34]. The model was able to detect 100 images in less than 0.01 s with 91.9% accuracy. Halstead *et al.* [35] developed a robotic vision system for classifying the sweet pepper ripeness into three classes (unripe, partially ripe, and ripe) based on the parallel Faster R-CNN technique. The framework yielded a classification accuracy of 82.1%. Ge *et al.* [33] employed the Mask Region-CNN model to detect and classify different ripening levels (raw, pink, and ripe) of strawberries in farm conditions. Mohtar *et al.* [36] adopted an Inception-v3 model to classify six stages of ripening of mangosteen fruit with a classification accuracy of 91.9%. Liu *et al.* [37] proposed a modified densely-connected convolutional network (DenseNet), aiming to detect the maturity of tomatoes in complicated images. The detection rate of the improved DenseNet network was superior compared to the residual network (ResNet), DenseNet, and single-shot detector (SSD) frameworks. Huang *et al.* [38] developed a fuzzy Mask R-CNN model to classify the ripeness levels of cherry tomatoes into 4 categories. Their model was able to achieve an accuracy of 98%. In another study, Ramos *et al.* [39] attempted to classify the ripening stage of two grape cultivars.

In their work, they employed two CNN architectures containing 10 convolutional layers and VGG-19. The authors reported that changing the number of ripening classes from three to eight would improve the classification accuracy from 65.30 to 93.41%. To detect different maturity levels of date fruit, Faisal *et al.* [40] used three pre-trained architectures: VGG-19, Inception-v3, and NASNet. They achieved a correct classification rate greater than 0.99.

D. MAIN CONTRIBUTIONS AND PAPER STRUCTURE

Our study fills the gap in the knowledge, makes the following significant contributions:

1) To best of our knowledge and literature survey, this research is the first attempt to determine the ripening stages of mulberries using CNN-based deep learning architecture.
2) The performance of different CNN architectures including DenseNet, Inception-v3, ResNet-18, ResNet-50, and AlexNet has been evaluated for this classification problem.
3) To minimize the number of training images and reduce the training time, CNN models have been fine-tuned and optimized on our target data sets.

The remainder of this paper is structured as follows: In Section II, the materials and methods used in the study including the computer vision system and detection models, are presented and explained. In Section III, the results of the testing of proposed frameworks are reported. In Section IV, the main results are discussed. In Section V, the design of a computer vision-based sorting system for automatic detection of mulberry ripening stages is proposed. Finally, in Section VI, conclusions and suggestions for further studies are presented.

II. MATERIALS AND METHODS

A. DATA SET COMPOSITION

Fruits of two mulberry genotypes, i.e., black and white mulberry, grown under the same environmental conditions were collected at four successive ripening stages, i.e., unripe, semi-ripe, ripe, and overripe (Fig. 2). These samples were hand-harvested by an expert in the morning from a commercial orchard located in Shahriar (35° 36' 43'' N; 51° 07' 27'' E), Tehran Province, Iran, in the period May-June 2020. Fruits packed in plastic punnets were immediately transported to the experimental laboratory under refrigeration at ∼5 °C. Only healthy fruits without disease and mechanical damages were selected. In total, 1000 samples of mulberry (250 samples per ripening stage) were used for imaging under controlled conditions at 20 ± 1 °C and 60 ± 5% RH. The TSS value of the juice produced from the fruits at each ripening stage was measured using a hand-held refractometer (Master-53PT, Atago, Japan, ±0.2% accuracy). The measurements were performed ten times for each class. The TSS values of mulberry fruits at different ripening stages are presented in Table 2.

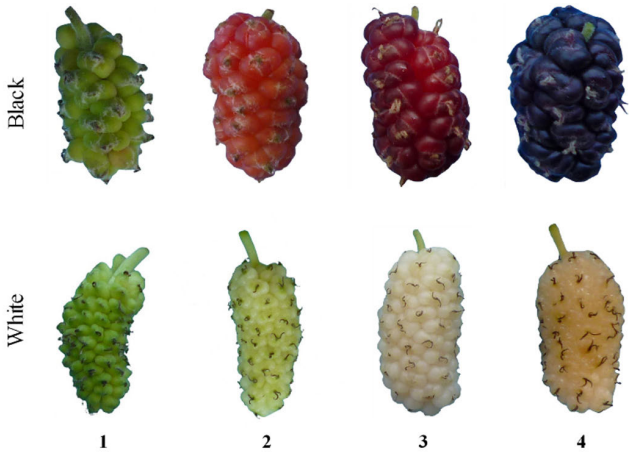

FIGURE 2. Different ripening stages of black and white mulberries: (1) unripe, (2) semi-ripe, (3) ripe, and (4) overripe.

TABLE 2. Changes in total soluble solids (TSS; °Brix) of mulberry fruits during different stages of ripening.

Ripening stage	Genotype	
	White	Black
Unripe	7.2 ± 0.4	6.7 ± 0.3
Semi-ripe	10.6 ± 0.6	9.4 ± 0.5
Ripe	13.9 ± 0.2	12.1 ± 0.8
Overripe	17.1 ± 0.7	14.8 ± 0.4

Mean TSS values and their standard deviation over 10 measurements

B. ACQUISITION OF IMAGES AND PREPROCESSING

Image acquisition was carried out by a digital camera (Nikon D3200 24.2 MP CMOS, Japan) placed at 25 cm from the samples in a 30^{L} cm × 30^{W} cm × 40^{H} cm illumination chamber, equipped with an 18 W circular fluorescent lamp with a color temperature of 6500 K. Each picture was acquired with the blue background of the chamber; no zoom nor flash were used. Images were stored in RGB format with a resolution of 300 dpi and an image size of 4320 × 3240. Image segmentation was accomplished by the modified unsupervised segmentation algorithm introduced by Aganj *et al.* [41]. This procedure includes converting RGB images into other color spaces including CIE Lab, HSV, and YCbCr [32]. Within these color spaces, the strongest contrast between the mulberries and background was obtained in YCbCr color space.

C. CNN MODELS

A typical structure of a deep CNN is composed of input and output layers, as well as multiple hidden layers.

The hidden layers of a CNN are generally made up of convolutional, pooling, activation, and fully connected layers and in some cases a Softmax layer [42], [43]. The CNNs have as a common characteristic that they can extract the features automatically from data and visualize the extracted features. For example, Fig. 3 shows visualization features of the first, intermediate, and last convolutional layers of

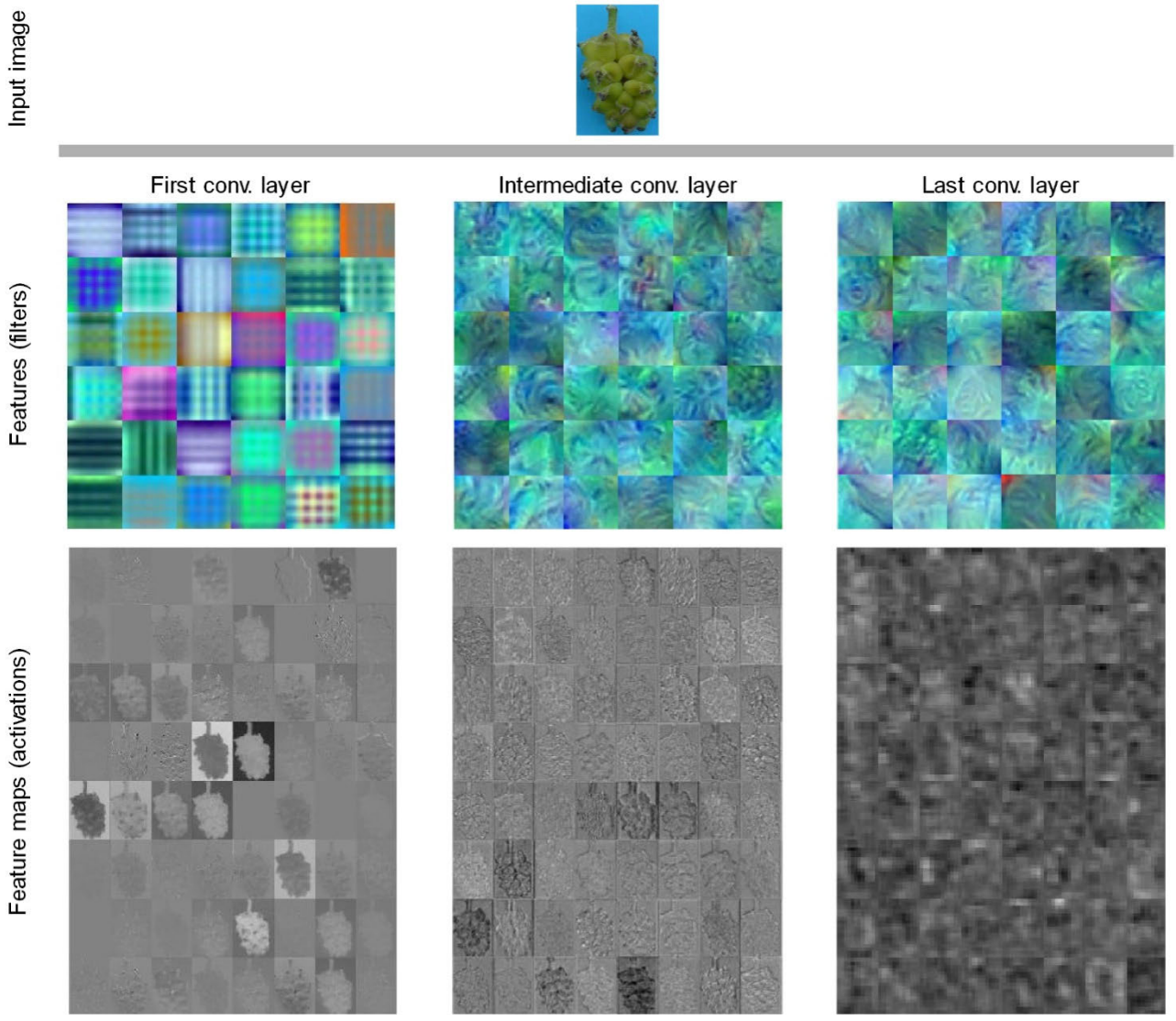

FIGURE 3. An example of visualization results of the ResNet-18 model. The figure shows the output features of the first, intermediate, and last convolutional layers at the top as well as feature maps of these features at the bottom.

the ResNet-18 model, as well as the activation of these features. The first convolutional layers mainly extract the primary features like colors and edges. The filters in the intermediate layers mostly contain texture information which is made of a combination of edges and colors. With the deepening of layers, their outputs become increasingly abstruse and less visually interpretable. Along with ResNet-18, four other well-known CNNs, such as AlexNet, Inception-v3, ResNet-50, and DenseNet, have been studied. These networks have been successfully used for a range of different image recognition tasks, such as leaf disease classification (ResNet by Deeba and Amutha [44] and Inception-v3 by Qiang *et al.* [45]), remote sensing (AlexNet, ResNet-34, ResNet-50, ResNet-101 ResNet-152, VGG-16, VGG-19 and DenseNet-121 by Rohith and Kumar [46]) and freshwater fish detection (DenseNet by Wang *et al.* [47]). Further, a short explanation of each CNN included in our study is provided hereafter. The workflow of our research with respect to testing of CNNs for the new task of classification of mulberry ripeness is illustrated in Fig. 4.

1) ALEXNET

AlexNet was introduced by Alex Krizhevsky [48] to compete in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC2012). The original design utilized two graphics processing units (GPUs) to speed up the training, but in this study, the single GPU processing version is used as it is more efficient with the newer GPUs. The AlexNet includes five convolutional and three fully-connected layers to process 227×227 pixels' images. In the first two fully connected layers, the Dropout regularization technique was applied to reduce the overfitting. Rectified Linear Units (ReLU) was utilized for all the hidden layers and Softmax for the output layer as the activation functions.

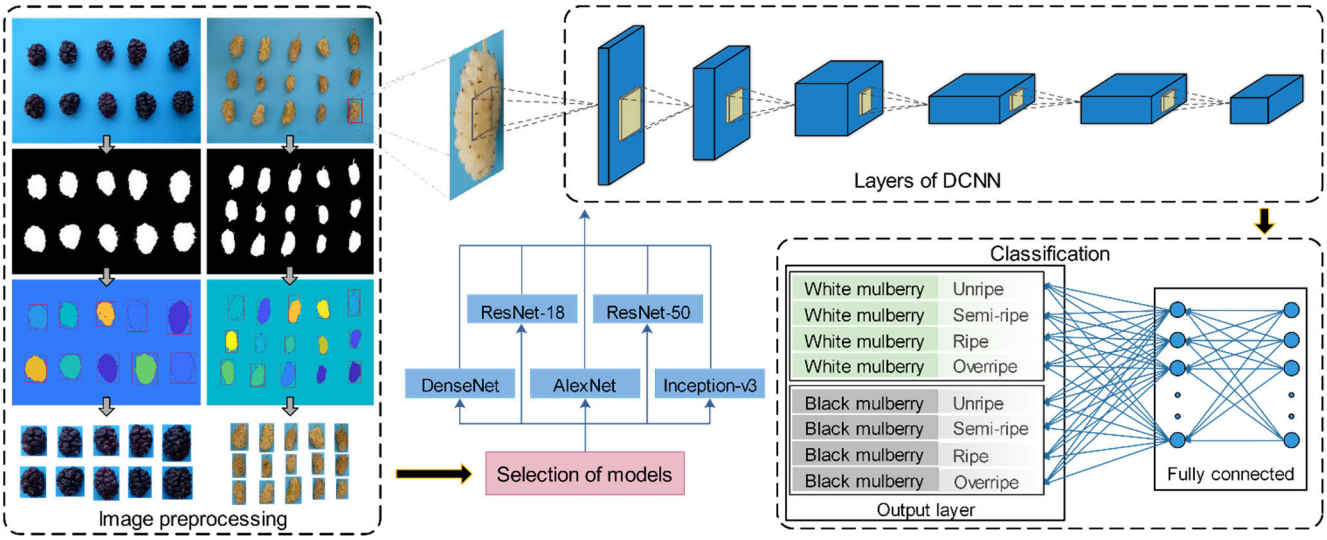

FIGURE 4. Schematic presentation of the workflow for classification of mulberry fruits.

2) RESNET

He *et al.* [49] introduced this CNN architecture to push the depth of convolutional networks to its limits. Due to a network-in-network (NIN) architecture, ResNet is theoretically capable of having an infinite depth without losing accuracy. In practice, it can have up to 152 layers by stacking ''residual blocks'' throughout the network. The NIN networks use blocks that have few convolutional layers but more complex structures (known as micro neural networks). These blocks help the whole network to extract better features by focusing on a smaller receptive field, instead of the usual convolutional networks, which scan the input image using linear filters [50]. ResNet has many stacked residual blocks, including a set of convolution and pooling layers. Although it has a similar architecture to AlexNet, it is about 20 times deeper due to the overcoming of the so-called degradation problem. ResNet has several implementations with different depths. In this study, ResNet-18 and ResNet-50 have been used.

3) INCEPTION-V3

In 2015, Google introduced a network called GoogLeNet, also known as inception-v1 [51] in order to achieve the performance of a deep network with a light-weight structure [52]. Inception-v1 has different kernel sizes (1×1, 3×3, 5×5) to extract feature maps in different scales, and by stacking them, the model can extract more features in total. This also reduces the parameters and accordingly reduces the computation [45]. Inception-v3 breaks down a large-scale convolution kernel into smaller convolution kernels (for instance breaks 3×3 kernels into two (1×3, 3×1) kernels). In this manner it contributes to further reduce network parameters and results it faster to run without sacrificing overall performance. At the same time this enables to extend the depth of the network [52].

4) DENSENET

In 2017, the idea of densely connected CNNs was proposed by Huang *et al.* [53]. This architecture is introduced to solve a notorious problem regarding very deep networks known as the ''vanishing-gradient''. The layers in DenseNet are connected to every other layer feed-forwardly and the feature-maps of each layer are used as inputs into all subsequent layers. This means that for a given network with n layers, there are $2n(n+1)$ direct connections between each layer and its subsequent layers, to compare with traditional CNNs with n layers which have n connections. In addition to mitigating the vanishing-gradient problem and reducing the number of parameters, DenseNet strengthens feature propagation and encourages feature reuse.

D. FINE-TUNING THE MODELS

A transfer learning approach was utilized to benefit from the pre-trained network by adjusting its parameters to our data set; this procedure is also known as fine-tuning. Fine-tuning is faster than training from scratch as a pre-trained network already has established weights. These weights are the result of learning over a data set (usually ImageNet) and help the network to train the features faster [42]. In order to realize this, the convolutional layers were frozen and the dense layer after those layers was trained. Furthermore, the last fully connected layer of networks was modified to have four output classes according to four levels of ripeness for either black or white mulberry respectively. If both data sets were combined, the last connected layer was modified to eight outputs for the genotype and ripeness detection.

Given that one of the most obvious approach to avoid overfitting is initializing all the weights of components in CNNs to a pre-trained model, ImageNet pre-trained CNNs have been used in all the experiments. It improves the generalization and the performance of the model [54].

TABLE 3. Specific parameters of the models in the evaluation.

Parameters	Network				
	AlexNet	Inception-v3	ResNet-18	ResNet-50	DenseNet
Depth	8	48	18	50	201
Image size (pixel)	227×227	299×299	224×224	224×224	224×224
Solver (optimizer)	SGDM	ADAM	ADAM	ADAM	SGDM
Loss function	cross entropy	cross entropy	cross entropy	cross entropy	cross entropy
Batch size	64	64	64	64	64
Learning rate	0.001	0.001	0.001	0.001	0.001
Learning rate drop factor	0.1	0.1	0.1	0.1	0.1
Learning rate drop period	10	10	10	10	10
Momentum	0.9	0.9	0.9	0.9	0.9
Gradient threshold method	L2norm	L2norm	L2norm	L2norm	L2norm
Parameters	61 M	23.9 M	11.7 M	25.6 M	20 M

SGDM: Stochastic gradient descent with momentum; Adam: Adaptive momentum estimation; M: Millions

E. SOFTWARE AND HARDWARE PLATFORM

All these networks are implemented in the MATLAB Deep Learning Toolbox (MATLAB R2020b, Mathworks Inc.) and Python 3.6. The parameters for each network are summarized in Table 3. The training was done using a machine with AMD Ryzen 9 3900 12cores/24thread 3.7GHz CPU, 128 GB DDR4 RAM, and a GeForce RTX 2080Ti GPU card with 11GB memory. The machine was installed with 2 GPU's but for the experiments only one was used.

III. RESULTS

A. PERFORMANCE EVALUATION

The performance of the selected deep learning models for the recognition and classification of mulberry fruits was evaluated based on multiple indicators: training accuracy, validation accuracy, training loss, and validation loss in each epoch. Training accuracy is a measure of model correctness during the training phase, whereas validation accuracy is defined as the percentage of test data truly classified by the trained model. The cross-entropy error was used as the loss function. A training-validation strategy for training and testing the classifier's performance was developed. The image data set was split into two independent groups: 70% for training and 30% for validation of the trained model. The reason for splitting the data set into two subsets is that in small data sets, the additional split might lead to a smaller training set which may be exposed to overfitting [55]. To provide enough data for training, the validation set was used to assess the performance of the models. In this regard, 5-fold cross-validation was involved to tune model hyperparameters. As a general rule, the higher the number of iterations, the higher the detection accuracy. However, after a certain number of epochs, the accuracy of network recognition is not increasing and sometimes even reducing [56]. As a result, the best detection accuracy can be achieved by selecting an optimal number of epochs for training the CNN models [57]. In our experiments, the model accuracy and the loss function were monitored during the training process. Each of the experiments did run for a total of 100 epochs, where the number of epochs was defined as the number of times the network had to cycle through the data set. One hundred training epochs were enough since the validation accuracy and loss function demonstrated best results around 20 to 75 epochs. This is further discussed in detail in later subsections. Another criterion, considered in the model evaluation, was classification time, calculated as the time required to classify all the validation samples by the trained algorithm.

B. WHITE MULBERRY CLASSIFICATION RESULTS

In this section, a quantitative assessment of the deep learning models for the classification of white mulberry ripeness from images is presented. The classification performance of AlexNet, DenseNet, Inception-v3, and ResNet with 18 and 50 layers have been compared based on the same set of metrics. Figure 5a-d illustrates losses and accuracies of training and validation of the chosen deep learning models during the training procedure. As can be deduced from Fig. 5a and c, all models achieved high accuracy in both the training and the validation phase, respectively. **ResNet-50**achieved the maximum classification accuracy without overfitting in 60 epochs. After the 60th epoch, the curves of training accuracy and loss reached a plateau and no significant change was observed in both loss function and accuracy (Fig. 5a and b). As seen in Fig. 5c and d, significant fluctuation of validation accuracy and loss curves started after the 83rd epoch, which is the sign of overfitting.

The behavior of **Inception-v3** was similar to ResNet-50, converging at around epoch 60. In the case of **ResNet-18**, the loss and accuracy of the training set did stabilize after about 74 epochs, while the accuracy and loss of the validation set started to converge at around the same epoch with tiny fluctuations. The **DenseNet** reached its optimal performance at epoch 34. After this point, the training and validation accuracy and loss curves started to become constant. In **AlexNet,** the training procedure converges while attaining reasonable accuracy and loss after around 38 epochs (Fig. 5a and b). Similarly, the validation accuracy and loss curves saturated after the same epoch. It is worth noting that the very close results for training and validation accuracy indicate that the overfitting did not happen during the

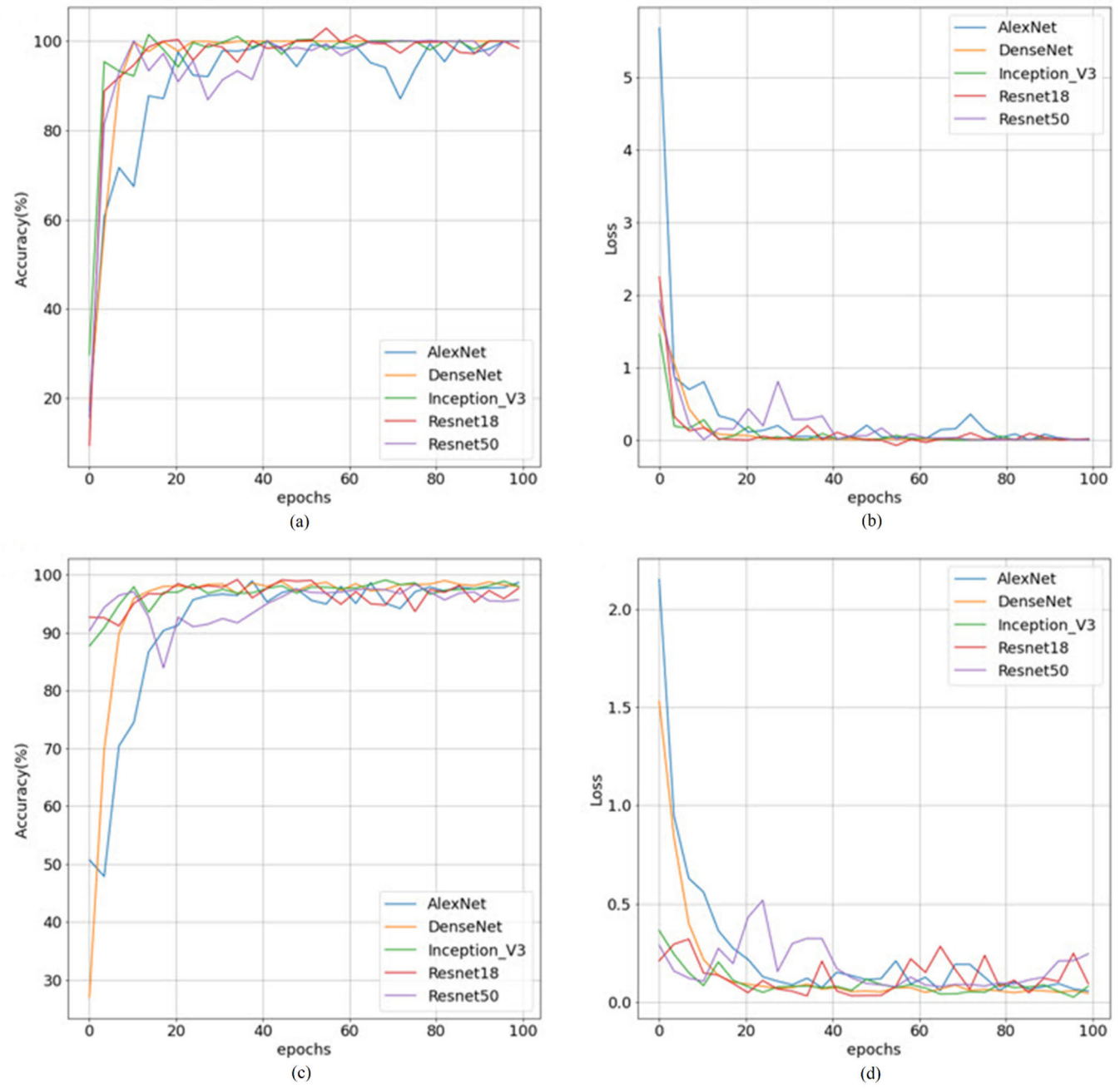

FIGURE 5. The behavior of the training (a, b) and validation (c, d) accuracies and losses of the white mulberry classification models.

training process [58]. The overall accuracy and loss of the five models for ripening classification of white mulberry are presented in Fig. 6a and b, respectively. Overall, all models achieved an accuracy of more than 96%, with DenseNet having the highest accuracy of 98.67% and the lowest loss of 0.0497. On the other hand, the highest loss of 0.182 and the lowest classification accuracy of 96.33% was obtained with ResNet-50. Figure 7 shows the classification time of the five CNN models. It follows that the DenseNet, despite having the highest classification accuracy, required the longest classification time compared to other models. AlexNet and ResNet-18 needed significantly less time for classification than others, with AlexNet faster than ResNet-18 by a very small margin. These results for classification time can be considered reasonable, since DenseNet contained the largest number of layers, whereas AlexNet had the fewest number of layers [57]. It can be concluded that among the compared networks, AlexNet was the best because of overall accuracy above 98%, very close to DenseNet. Although the loss was slightly higher, it was 26 times faster than DenseNet.

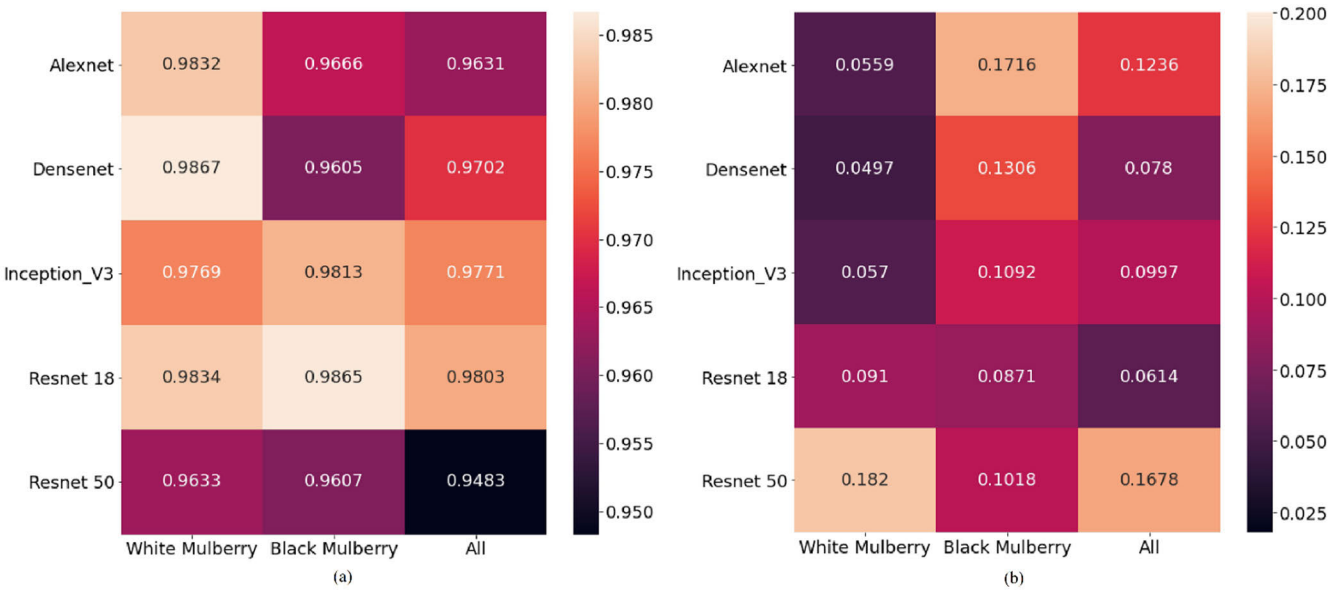

FIGURE 6. Comparison of overall accuracy (a) and loss (b) of proposed models.

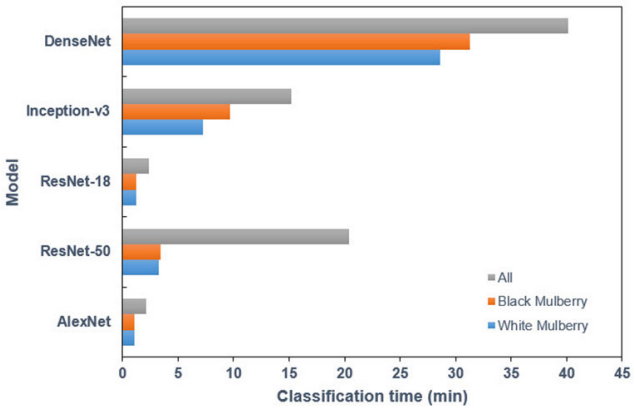

FIGURE 7. Classification time of proposed models.

C. BLACK MULBERRY CLASSIFICATION RESULTS

In this section, the performance of deep learning models for the classification of black mulberries ripening stage from images is evaluated. The graphical representation of the training and validation accuracy and loss of the classifiers for each experimental run is shown in Fig. 8a-d. In the case of the **ResNet-50** model, we observed that with the number of the training epochs the algorithm gradually converged and the best results were achieved for 71 epochs (Fig. 8a and b). At the beginning of the training, the classification accuracy of the algorithm was relatively low and then gradually improved. The high accuracy rate and low loss rate were achieved after about 20 epochs and convergence was achieved in approximately 70 epochs. With increasing the number of epochs, the training and validation losses decreased from about 2.6 to 0.0002 and 0.4 to 0.09, respectively. As displayed in Fig. 8a and c, the training and validation accuracy increased from about 23% to 100% and 86% to 98% with the progression in training epochs, respectively.

In **Inception-v3**, the training accuracy increased with the number of epochs and then stabilized after the 75th epoch. After this epoch, the validation accuracy began to decline. Likewise, the training loss became stable after 75 epochs and the validation loss began to increase (Fig. 8b and d), implying that the model performed better on the training data set than on the validation data set. The training and validation losses of the model reduced from almost 1.3 to 0.00008 and 0.6 to 0.1 at the end of the training, respectively.

For **DenseNet** architecture, the training accuracy reached saturation at 20 epochs, but the validation accuracy fluctuated between 95% and 99% starting from about the 30th epoch. The training loss for DenseNet rapidly decreased from the first to twenty epochs and then became steady, while the validation loss values rapidly reduced from the first to fifteen epochs and then reasonably stabilized although fluctuating between 0.06 and 0.15. As illustrated in Fig. 8a and b, the training of the AlexNet model stopped after 60 epochs as the training accuracy and loss started to plateau. After this epoch, the validation accuracy of the model decreased and the loss parameter increased (Fig. 8d). A comparison of the results in Fig. 6a and b depicts that all proposed CNN models can efficiently classify ripening stages of black mulberry. For all scenarios, the training loss and accuracy were higher or approximately equal to the validation loss and accuracy, indicating that the networks were able to generalize well without overfitting. The **ResNet-18**model showed the best performance in both accuracy and loss function. The accuracy of ResNet-18 was 98.65%, which is 2.6% higher than that of DenseNet. Although DenseNet performed the best in classifying ripening levels of white

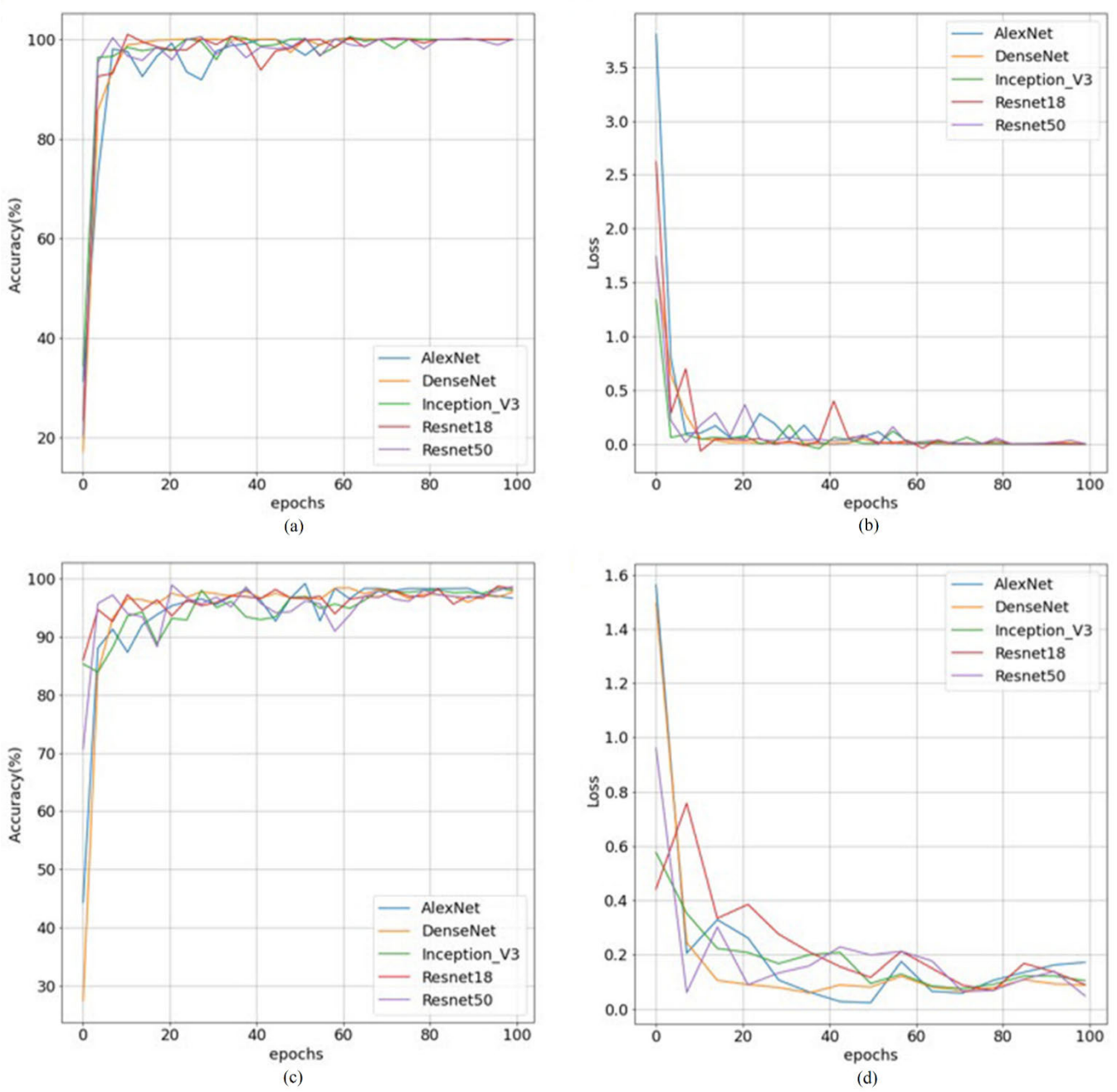

FIGURE 8. The behavior of the training (a, b) and validation (c, d) accuracies and losses of the black mulberry classification models.

mulberry, it was the relative worst in classifying black mulberry. Compared with the other CNN models, the **AlexNet** had the highest loss value of 0.1716, which is twice higher than that of ResNet-18. As shown in Fig. 7, classification time varied from 1.05 to 31.25 min with AlexNet requiring the shortest classification time and DenseNet requiring the longest classification time because of an extensive number of layers [56]. Although AlexNet was the champion in the classification time, the ResNet-18 required only 8 seconds more, which is negligibly small compared to the overall classification time. Therefore, by considering all the aforementioned results, ResNet-18 is considered the best option among all networks for classifying the ripening stage of black mulberry.

D. CLASSIFICATION RESULTS OF COMBINING BOTH GENOTYPES

The training and validation loss and accuracy curves of different models for the classification of both genotypes with the ripening stage are shown in Fig. 9a-d and explained in detail hereafter.

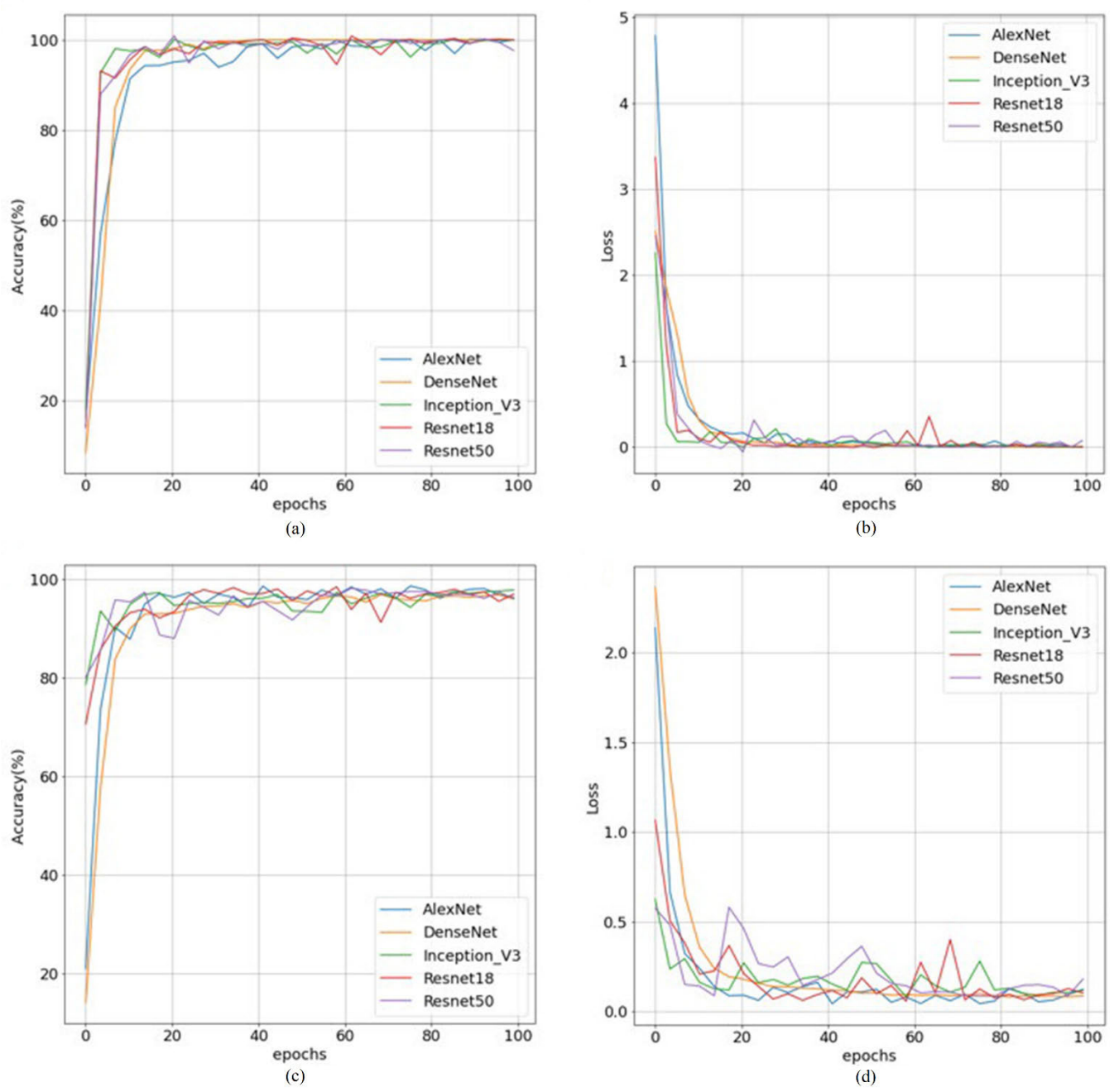

FIGURE 9. Training (a, b) and validation (c, d) accuracy and loss curves of the CNN models for classification of both genotypes.

ResNet-50: The best performance of the ResNet-50 model was reported around epoch 63 because of the greatest training accuracy and the lowest training loss (Fig. 9a and b). The training loss and accuracy saturated after 63 epochs, so the slope of both training plots was close to zero. The accuracy of the validation data set increased continuously up to epoch 63 but then started to decrease. In the same manner, the loss decreased continuously up to epoch 63 but then started to increase (Fig. 9c and d).

ResNet-18: The ResNet-18 model required a lower number of epochs (about 33) to reach the desired performance than the ResNet-50. Fig. 9b and d show strong fluctuation in the training and validation losses between the 46th and 78th epochs, indicating overfitting. However, after the 78th epoch, there were no excessive fluctuations, which tells that overfitting was considerably decreased.

Inception-v3: As shown in Fig. 9b and d, both training and validation loss of Inception-v3 decreased with the number of epochs and tended to flatten. The accuracy of the training and validation sets of the model reached optimal accuracy at about 30 epochs (Fig. 9a and c).

DenseNet: For the DenseNet model, it took around 43 epochs for the network to converge appropriately. It is visible from the trends in Fig. 9a and c, initially, the training and validation accuracy values rose sharply. Later, the growth was gradual and reached a plateau. Training and validation losses dropped consistently and converged, indicating a well-fitting model (Fig. 9b and d).

AlexNet: The training set loss in AlexNet rapidly declined in the first ten epochs and then slightly fluctuated as the number of epochs increased (Fig. 9b). From Fig. 9d it follows that the validation loss became relatively stable after 42 epochs. After the 42nd epoch, the fluctuations in the accuracy of the validation set were negligible: around 97% with 3% tolerance (Fig. 9c). No discrepancy between the training and validation accuracies implies that the proposed model had no evident overfitting.

ResNet-18 performed better than the other four models in our study, reaching an accuracy of 98.03% and loss of 0.0614, in contrast to ResNet-50, which had the lowest accuracy (94.83%) and the highest loss (0.1678). Fig. 6a and 6b illustrate the overall accuracy and loss of the proposed models for the classification of both genotypes based on their ripening stage. It follows that the classification capability of the proposed networks is generally maintained the similar, despite the combination of data sets and the increasing complexity of the problem. From Fig. 7 it follows that the classification time of all models for the mixed batch increased compared to other scenarios. For example, DenseNet required the longest classification time (40.14 min), while AlexNet had the shortest classification time (2.1 min). This is logical since the models perform two simultaneous tasks of classifying the genotype and degree of ripeness. Although AlexNet had the shortest classification time, the time taken by ResNet-18 to complete the classification process was only 15 sec longer. Given these results, ResNet-18 was selected as the best model for the classification of genotype and ripeness of mulberries.

The superiority of the ResNet-18 over other models is related to its well-designed topology and structure. One of the common problems among CNN models is performance saturation specifically in deep networks [59]. The ResNet-18 overcomes this issue by implementing an identity shortcut connection, which skips one or more layers and performs identity mapping of the layer than the original mapping [60]. Through the residual connections, all the inputs can forward propagate faster across the layers [49]. On the other hand, ResNet-18 architecture has a shallow depth which reduces the overfitting problem, parameters and overhead of computing resources [61]. In general, ResNets are easy to optimize and can easily obtain accuracy gains from considerably increased depth.

IV. DISCUSSION

Manual classification of mulberry fruit ripening stages at different points of the production chain, i.e., farmers, manufacturers, distributors, and retailers, is a challenging task. The introduction and application of novel technologies may help in solving the problem of real-time detection and classification of fruits according to ripening level. Our research has been a first step into developing a new intelligent sorting system, based on computer vision and deep learning techniques, to enable recognizing and classifying white and black mulberries into four classes, i.e., unripe, semi-ripe, ripe, and overripe. The potential of the deep learning technique to solve this problem is substantiated by the excellent performance of the CNN models that were evaluated. From the literature, only one attempt is known to classify the mulberry fruit based on the ripening stage [17]. In this study ripeness of red and white mulberry fruit were classified into three categories, i.e., unripe, ripe, and overripe. In addition to those three classes, the feasibility of classifying the semi-ripe class was also examined. For the classification task, the two traditional machine learning methods, i.e., ANN and SVM, were applied with handcrafted features. The results of this study are, unfortunately, difficult to compare with ours, because of differences in materials, techniques, data sets, as well as classification criteria.

Our approach is based on deep learning frameworks, i.e., AlexNet, DenseNet, ResNet-18, ResNet-50 and Inception-v3, which do not require manual definition of features. The performance of the trained CNN models in detecting the mulberry type and ripeness demonstrated high potential with accuracies between 95 and 98%. Moreover, the size of the sample set in our study resulted in an increase of the level of classification reliability. Considering the fact that ResNet-18 and AlexNet have fewer layers compared to the other architectures, the classification accuracy achieved with these shallow networks was surprisingly good in comparison with deeper networks such as DenseNet. Although deeper models can improve the classification accuracy [62], increasing the number of layers raises challenges of degradation, computational cost, internal covariate shifts, and vanishing gradients [42]. Low sample size in a given data set in deep networks can lead to overfitting. In this case, shallow networks can be a possible solution [63]. The methodologies presented in this paper result in a significantly better classification, adding to the knowledge on ripening detection. From an economic perspective, potentially it is an extremely cost-effective system because it is based on images from low-cost digital cameras without the need for sophisticated imaging equipment, such as hyperspectral imaging, laser backscattering imaging, multispectral imaging, fluorescence imaging, magnetic resonance imaging, etc.

V. SCALE-UP AND INTEGRATION OF TECHNOLOGY

In Figure 10 a proposed design of an automatic mulberry sorting machine is depicted; it consists of a mulberry feeding unit, conveying units, imaging equipment, and sorting mechanism. Mulberry fruits poured in the container are directed to the flat conveyors which are equipped with hinged small shallow bowls. The fruits on the conveyor belt continuously pass through the image capturing area with a camera and illumination system mounted exactly perpendicular to the

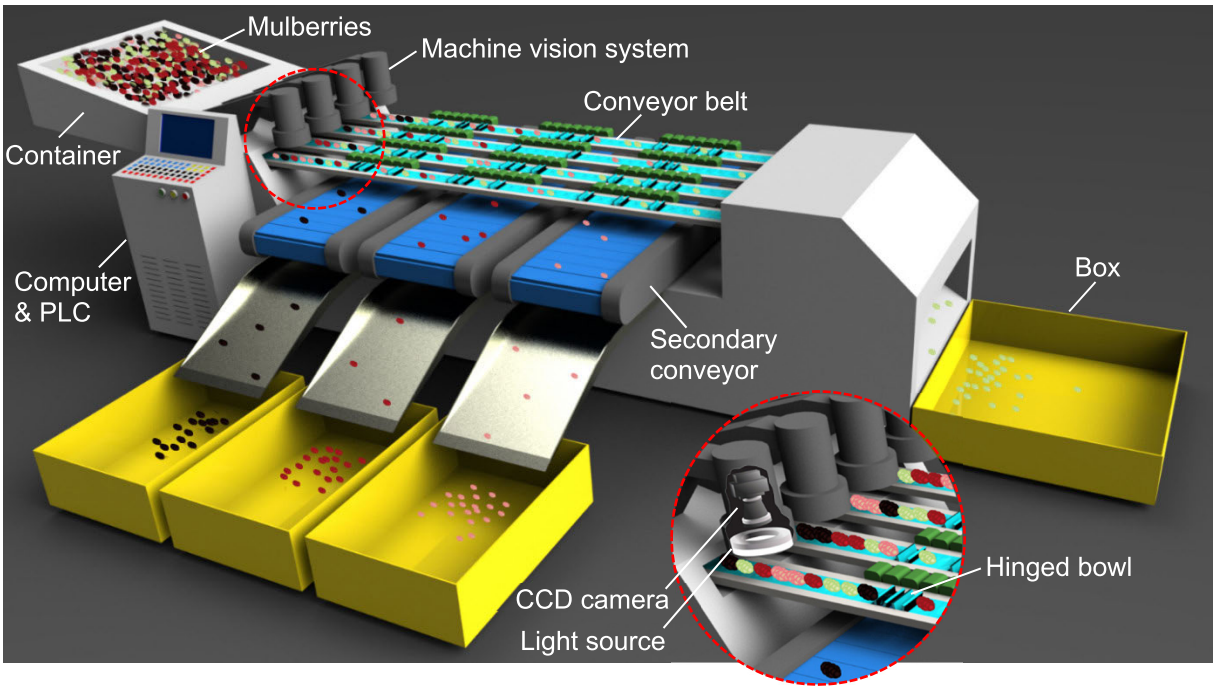

FIGURE 10. The conceptual design of an automatic mulberry sorting system.

center of the conveyor. Data from the camera are transferred to a computer in real-time to determine the fruit ripening class. The capacity of the system is increased due to multiple classification lines. When mulberry fruit reaches the exact location, determined by the speed of the conveyor, a signal is sent by the Programmable Logic Controller (PLC) [64] to open the bowl, which will deliver fruit to the second conveyor to be discharged into a specified box.

To prevent bruising damage, mulberry fruits, classified as overripe with the least mechanical strength are falling into the box at the end of the first conveyor without dropping on the second conveyor. From the product quality perspective, the factors affecting the mechanical behavior of the fruit, including the trajectory of falling, the material of the conveyor, and static and dynamic forces applied to the fruit, need to be carefully evaluated before constructing such a system.

VI. CONCLUSION

This study presents a comparison of different CNN-based models for the classification of two genotypes of mulberry, namely white and black, according to their ripening stages. The models that were evaluated include DenseNet, Inception-v3, ResNet-18, ResNet-50 and AlexNet. These models have been tested on a data set of 2000 fruit images (1000 per genotype), where the model was trained using 70% of the data set. The performance analysis has been done by comparing three performance metrics: accuracy, loss, and classification time. The major contributions of this study are as follows:

(1) Although all CNN architectures achieved high accuracy, the AlexNet model outperformed the other models in the classification of white mulberries with an accuracy of 98.32%, loss of 0.0559, and classification time of about 1 min.
(2) The experimental results demonstrated that the ResNet-18 model seems to be more reliable for the classification of black mulberry ripening with the best accuracy, minimal loss, and short classification time (98.65%, 0.0871, and 1.2 min, respectively).
(3) The overall performance of the ResNet-18 was the best when the data sets of both genotypes were combined. The model was neither overfitted nor underfitted. The recognition of the fruit genotype and classification of the ripening stage of 600 testing samples showed that the overall accuracy, loss rate, and processing time were 98.03%, 0.0614, and 2.36 min, respectively.

The results of the current study could be extended to the classification of more than four ripeness stages. This study also provides the foundation for the design of an automated sorting machine for mulberry fruit. Moreover, the deep learning frameworks applied in this study can serve as a template for other types of horticultural commodities.

ACKNOWLEDGMENT

(Seyed-Hassan Miraei Ashtiani and Shima Javanmardi contributed equally to this work.)

REFERENCES

[1] O. Saracoglu, "Phytochemical accumulation of anthocyanin rich mulberry (Morus laevigata) during ripening," *J. Food Meas. Characterization*, vol. 12, no. 3, pp. 2158–2163, Sep. 2018.
[2] H. Chen, H. Gao, X. Fang, L. Ye, Y. Zhou, and H. Yang, "Effects of allyl isothiocyanate treatment on postharvest quality and the activities of antioxidant enzymes of mulberry fruit," *Postharvest Biol. Technol.*, vol. 108, pp. 61–67, Oct. 2015.

[3] C. Nguyen and H. Nguyen, "The quality of mulberry juice as affected by enzyme treatments," *Beverages*, vol. 4, no. 2, p. 41, May 2018.

[4] A. Jelled, R. B. Hassine, A. Thouri, G. Flamini, H. Chahdoura, A. E. Arem, J. B. Lamine, A. Kacem, Z. Haouas, H. B. Cheikh, and L. Achour, "Immature mulberry fruits richness of promising constituents in contrast with mature ones: A comparative study among three tunisian species," *Ind. Crops Products*, vol. 95, pp. 434–443, Jan. 2017.

[5] E. M. Sánchez-Salcedo, P. Mena, C. García-Viguera, J. J. Martínez, and F. Hernández, "Phytochemical evaluation of white (*Morus alba* L.) and black (*Morus nigra* L.) mulberry fruits, a starting point for the assessment of their beneficial properties," *J. Funct. Foods*, vol. 12, pp. 399–408, Jan. 2015.

[6] H. Lou, Y. Hu, L. Zhang, P. Sun, and H. Lu, "Nondestructive evaluation of the changes of total flavonoid, total phenols, ABTS and DPPH radical scavenging activities, and sugars during mulberry (*Morus alba* L.) fruits development by chlorophyll fluorescence and RGB intensity values," *LWT-Food Sci. Technol.*, vol. 47, no. 1, pp. 19–24, Jun. 2012.

[7] L. Huang, Y. Zhou, L. Meng, D. Wu, and Y. He, "Comparison of different CCD detectors and chemometrics for predicting total anthocyanin content and antioxidant activity of mulberry fruit using visible and near infrared hyperspectral imaging technique," *Food Chem.*, vol. 224, pp. 1–10, Jun. 2017.

[8] Á. Calín-Sánchez, J. J. Martínez-Nicolás, S. Munera-Picazo, Á. A. Carbonell-Barrachina, P. Legua, and F. Hernández, "Bioactive compounds and sensory quality of black and white mulberries grown in spain," *Plant Foods Hum. Nutrition*, vol. 68, no. 4, pp. 370–377, Dec. 2013.

[9] Y. You, N. Li, X. Han, J. Guo, Y. Zhao, G. Liu, W. Huang, and J. Zhan, "Influence of different sterilization treatments on the color and anthocyanin contents of mulberry juice during refrigerated storage," *Innov. Food Sci. Emerg. Technol.*, vol. 48, pp. 1–10, Aug. 2018.

[10] Q. Han, H. Gao, H. Chen, X. Fang, and W. Wu, "Precooling and ozone treatments affects postharvest quality of black mulberry (*Morus nigra*) fruits," *Food Chem.*, vol. 221, pp. 1947–1953, Apr. 2017.

[11] S. Nayab, K. Razzaq, S. Ullah, I. A. Rajwana, M. Amin, H. N. Faried, G. Akhtar, A. S. Khan, Z. Asghar, H. Hassan, and A. Naz, "Genotypes and harvest maturity influence the nutritional fruit quality of mulberry," *Scientia Horticulturae*, vol. 266, May 2020, Art. no. 109311.

[12] D. Donno, A. K. Cerutti, I. Prgomet, M. G. Mellano, and G. L. Beccaro, "Foodomics for mulberry fruit (*Morus spp.*): Analytical fingerprint as antioxidants' and health properties' determination tool," *Food Res. Int.*, vol. 69, pp. 179–188, Mar. 2015.

[13] G. K. Rohela, P. Shukla, Muttanna, R. Kumar, and S. R. Chowdhury, "Mulberry (*Morus spp.*): An ideal plant for sustainable development," *Trees, Forests People*, vol. 2, Dec. 2020, Art. no. 100011.

[14] J. Yang, X. Liu, X. Zhang, Q. Jin, and J. Li, "Phenolic profiles, antioxidant activities, and neuroprotective properties of mulberry (*Morus atropurpurea* Roxb.) fruit extracts from different ripening stages," *J. Food Sci.*, vol. 81, no. 10, pp. C2439–C2446, Oct. 2016.

[15] A. Assirelli, F. Stagno, A. Cocchi, S. Sirri, A. Saviane, D. Giovannini, and S. Cappellozza, "Innovative system for mulberry fruit harvesting," *J. Berry Res.*, vol. 9, no. 4, pp. 615–630, Nov. 2019.

[16] N. Tabakoglu and H. Karaca, "Effects of ozone-enriched storage atmosphere on postharvest quality of black mulberry fruits (*Morus nigra* L.)," *LWT-Food Sci. Technol.*, vol. 92, pp. 276–281, Jun. 2018.

[17] H. Azarmdel, A. Jahanbakhshi, S. S. Mohtasebi, and A. R. Muñoz, "Evaluation of image processing technique as an expert system in mulberry fruit grading based on ripeness level using artificial neural networks (ANNs) and support vector machine (SVM)," *Postharvest Biol. Technol.*, vol. 166, Aug. 2020, Art. no. 111201.

[18] J. Yang, H. Wen, L. Zhang, X. Zhang, Z. Fu, and J. Li, "The influence of ripening stage and region on the chemical compounds in mulberry fruits (*Morus atropurpurea* Roxb.) based on UPLC-QTOF-MS," *Food Res. Int.*, vol. 100, pp. 159–165, Oct. 2017.

[19] G. Hojjatpanah, M. Fazaeli, and Z. Emam-Djomeh, "Effects of heating method and conditions on the quality attributes of black mulberry (*Morus nigra*) juice concentrate," *Int. J. Food Sci. Technol.*, vol. 46, no. 5, pp. 956–962, May 2011.

[20] H. Li, Z. Yang, Q. Zeng, S. Wang, Y. Luo, Y. Huang, Y. Xin, and N. He, "Abnormal expression of bHLH3 disrupts a flavonoid homeostasis network, causing differences in pigment composition among mulberry fruits," *Hortic. Res.*, vol. 7, no. 1, p. 83, 2020.

[21] Y. Lee and K. T. Hwang, "Changes in physicochemical properties of mulberry fruits (*Morus alba* L.) during ripening," *Sci. Hortic.*, vol. 217, pp. 189–196, Mar. 2017.

[22] W. Castro, J. Oblitas, M. De-la-Torre, C. Cotrina, K. Bazán, and H. Avila-George, "Classification of cape gooseberry fruit according to its level of ripeness using machine learning techniques and different color spaces," *IEEE Access*, vol. 7, pp. 27389–27400, 2019.

[23] P. Wan, A. Toudeshki, H. Tan, and R. Ehsani, "A methodology for fresh tomato maturity detection using computer vision," *Comput. Electron. Agricult.*, vol. 146, pp. 43–50, Mar. 2018.

[24] Y. Zhang, J. Lian, M. Fan, and Y. Zheng, "Deep indicator for fine-grained classification of banana's ripening stages," *EURASIP J. Image Video Process.*, vol. 2018, no, 1, p. 46, 2018.

[25] J. Zhuang, C. Hou, Y. Tang, Y. He, Q. Guo, A. Miao, Z. Zhong, and S. Luo, "Assessment of external properties for identifying banana fruit maturity stages using optical imaging techniques," *Sensors*, vol. 19, no. 3, p. 2910, 2019.

[26] L. F. S. Pereira, S. Barbon, Jr., N. A. Valous, and D. F. Barbin, "Predicting the ripening of papaya fruit with digital imaging and random forests," *Comput. Electron. Agric.*, vol. 145, pp. 76–82, Feb. 2018.

[27] S. Sabzi, Y. Abbaspour-Gilandeh, G. García-Mateos, A. Ruiz-Canales, J. M. Molina-Martínez, and J. I. Arribas, "An automatic non-destructive method for the classification of the ripeness stage of red delicious apples in orchards using aerial video," *Agronomy*, vol. 9, no. 2, p. 84, 2019.

[28] R. Pourdarbani, S. Sabzi, D. Kalantari, J. Paliwal, B. Benmouna, G. García-Mateos, and J. M. Molina-Martínez, "Estimation of different ripening stages of Fuji apples using image processing and spectroscopy based on the majority voting method," *Comput. Electron. Agric.*, vol. 176, Sep. 2020, Art. no. 105643.

[29] S. Cárdenas-Pérez, J. Chanona-Pérez, J. V. Méndez-Méndez, G. Calderón-Domínguez, R. López-Santiago, M. J. Perea-Flores, and I. Arzate-Vázquez, "Evaluation of the ripening stages of apple (Golden Delicious) by means of computer vision system," *Biosyst. Eng.*, vol. 159, pp. 46–58, Jul. 2017.

[30] B. Harel, Y. Parmet, and Y. Edan, "Maturity classification of sweet peppers using image datasets acquired in different times," *Comput. Ind.*, vol. 121, Oct. 2020, Art. no. 103274.

[31] L. Zhou, C. Zhang, F. Liu, Z. Qiu, and Y. He, "Application of deep learning in food: A review," *Compr. Rev. Food Sci. Food Saf.*, vol. 18, no. 6, pp. 1793–1811, 2019.

[32] S. Javanmardi, S. H. M. Ashtiani, F. J. Verbeek, and A. Martynenko, "Computer-vision classification of corn seed varieties using deep convolutional neural network," *J. Stored Prod. Res.*, vol. 92, May 2021, Art. no. 101800.

[33] Y. Ge, Y. Xiong, and P. J. From, "Instance segmentation and localization of strawberries in farm conditions for automatic fruit harvesting," *IFAC-PapersOnLine*, vol. 52, no. 30, pp. 294–299, 2019.

[34] L. Zhang, J. Jia, G. Gui, X. Hao, W. Gao, and M. Wang, "Deep learning based improved classification system for designing tomato harvesting robot," *IEEE Access*, vol. 6, pp. 67940–67950, Oct. 2018.

[35] M. Halstead, C. McCool, S. Denman, T. Perez, and C. Fookes, "Fruit quantity and ripeness estimation using a robotic vision system," *IEEE IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 2995–3002, Oct. 2018.

[36] I. A. Mohtar, N. S. S. Ramli, and Z. Ahmad, "Automatic classification of mangosteen ripening stages using deep learning," in *Proc. 1st Int. Conf. Artif. Intell. Data Sci. (AiDAS)*, Sep. 2019, pp. 44–47.

[37] J. Liu, J. Pi, and L. Xia, "A novel and high precision tomato maturity recognition algorithm based on multi-level deep residual network," *Multimedia Tools Appl.*, vol. 79, pp. 9403–9417, Apr. 2020.

[38] Y.-P. Huang, T.-H. Wang, and H. Basanta, "Using fuzzy mask R-CNN model to automatically identify tomato ripeness," *IEEE Access*, vol. 8, pp. 207672–207682, 2020.

[39] R. P. Ramos, J. S. Gomes, R. M. Prates, E. F. S. Filho, B. J. Teruel, and D. dos Santos Costa, "Non-invasive setup for grape maturation classification using deep learning," *J. Sci. Food Agric.*, vol. 101, no. 5, pp. 2042–2051, 2021.

[40] M. Faisal, M. Alsulaiman, M. Arafah, and M. A. Mekhtiche, "IHDS: Intelligent harvesting decision system for date fruit based on maturity stage using deep learning and computer vision," *IEEE Access*, vol. 8, pp. 167985–167997, 2020.

[41] I. Aganj, M. G. Harisinghani, R. Weissleder, and B. Fischl, "Unsupervised medical image segmentation based on the local center of mass," *Sci. Rep.*, vol. 8, no. 1, p. 13012, Dec. 2018.

[42] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Comput. Electron. Agric.*, vol. 161, pp. 272–279, Jun. 2019.
[43] M. Momeny, A. Jahanbakhshi, K. Jafarnezhad, and Y. D. Zhang, "Accurate classification of cherry fruit using deep CNN based on hybrid pooling approach," *Postharvest Biol. Technol.*, vol. 166, Aug. 2020, Art. no. 111204.
[44] K. Deeba and B. Amutha, "ResNet-deep neural network architecture for leaf disease classification," *Microprocess. Microsyst.*, Oct. 2020, Art. no. 103364. Accessed: Mar. 12, 2021, doi: 10.1016/j.micpro.2020.103364.
[45] Z. Qiang, L. He, and F. Dai, "Identification of plant leaf diseases based on inception V3 transfer learning and fine-tuning," in *Proc. Int. Conf. Smart City Informatization*, Nov. 2019, pp. 118–127.
[46] G. Rohith and L. S. Kumar, "Remote sensing signature classification of agriculture detection using deep convolution network models," in *Proc. Int. Conf. Mach. Learn. Image Process. Netw. Secur. Data Sci.*, Jul. 2020, pp. 343–355.
[47] H. Wang, Y. Shi, Y. Yue, and H. Zhao, "Study on freshwater fish image recognition integrating SPP and DenseNet network," in *Proc. IEEE Int. Conf. Mechatronics Autom. (ICMA)*, Oct. 2020, pp. 564–569.
[48] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, Stateline, NV, USA, Dec. 2012, pp. 1097–1105.
[49] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 770–778.
[50] M. Lin, Q. Chen, and S. Yan, "Network in network," 2013, *arXiv:1312.4400*. [Online]. Available: http://arxiv.org/abs/1312.4400
[51] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1–9.
[52] F. Emmert-Streib, Z. Yang, H. Feng, S. Tripathi, and M. Dehmer, "An introductory review of deep learning for prediction models with big data," *Front. Artif. Intell.*, vol. 3, p. 4, Feb. 2020.
[53] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 2261–2269.
[54] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, "Show and tell: Lessons learned from the 2015 MSCOCO image captioning challenge," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 652–663, Apr. 2017.
[55] M. Féré, C. Gobinet, L. H. Liu, A. Beljebbar, V. Untereiner, D. Gheldof, M. Chollat, J. Klossa, B. Chatelain, and O. Piot, "Implementation of a classification strategy of Raman data collected in different clinical conditions: Application to the diagnosis of chronic lymphocytic leukemia," *Anal. Bioanal. Chem.*, vol. 412, no. 4, pp. 949–962, Feb. 2020.
[56] A. Wu, J. Zhu, and T. Ren, "Detection of apple defect using laser-induced light backscattering imaging and convolutional neural network," *Comput. Electr. Eng.*, vol. 81, Jan. 2020, Art. no. 106454.
[57] H. K. Suh, J. IJsselmuiden, J. W. Hofstee, and E. J. van Henten, "Transfer learning for the classification of sugar beet and volunteer potato under field conditions," *Biosyst. Eng.*, vol. 174, pp. 50–65, Oct. 2018.
[58] W. Mumtaz and A. Qayyum, "A deep learning framework for automatic diagnosis of unipolar depression," *Int. J. Med. Informat.*, vol. 132, Dec. 2019, Art. no. 103983.
[59] C. Kwan, B. Chou, J. Yang, A. Rangamani, T. Tran, J. Zhang, and R. Etienne-Cummings, "Deep learning-based target tracking and classification for low quality videos using coded aperture cameras," *Sensors*, vol. 19, no. 17, p. 3702, Aug. 2019.
[60] C. Kwan, B. Chou, J. Yang, A. Rangamani, T. Tran, J. Zhang, and R. Etienne-Cummings, "Target tracking and classification using compressive sensing camera for SWIR videos," *Signal, Image Video Process.*, vol. 13, no. 8, pp. 1629–1637, Nov. 2019.
[61] J. Zhang, C. Lu, J. Wang, L. Wang, and X.-G. Yue, "Concrete cracks detection based on FCN with dilated convolution," *Appl. Sci.*, vol. 9, no. 13, p. 2686, Jul. 2019.
[62] K. Dutta, P. Krishnan, M. Mathew, and C. V. Jawahar, "Towards accurate handwritten word recognition for Hindi and Bangla," in *Proc. Nat. Conf. Comput. Vis. Pattern Recognit. Image Process. Graph.*, Dec. 2017, pp. 470–480.
[63] Q. Sun and Z. Ge, "A survey on deep learning for data-driven soft sensors," *IEEE Trans. Ind. Informat.*, vol. 17, no. 9, pp. 5853–5866, Sep. 2021, doi: 10.1109/TII.2021.3053128.
[64] J. W. Webb and R. A. Reis, *Programmable Logic Controllers: Principles and Applications*. 5th ed. Delhi, India: PHI Learning, 2002.

SEYED-HASSAN MIRAEI ASHTIANI received the M.Sc. degree in agricultural machinery engineering from the Ferdowsi University of Mashhad, Mashhad, Iran, in 2012, where he is currently pursuing the Ph.D. degree in biosystems engineering. His research interests include application of image processing and deep learning techniques to solve agricultural problems.

SHIMA JAVANMARDI received the B.S. and M.Sc. degrees in software engineering. She is currently pursuing the Ph.D. degree with the Leiden Institute of Advanced Computer Science (LIACS), Leiden University, The Netherlands. She is working as an Education and Research Staff Member at the LIACS, Leiden University. Her research interests include AI and machine learning, with a focus on image processing and deep learning.

MEHRDAD JAHANBANIFARD received the B.Sc. degree from the University of Isfahan, Isfahan, in 2014, and the M.Sc. degree from Shahid Beheshti University, Tehran, Iran, in 2017. He is currently pursuing the Ph.D. degree with the Leiden Institute of Advanced Computer Science (LIACS), Leiden, The Netherlands. He is currently working as a Guest Researcher at the Naturalis Biodiversity Center, Leiden. His research interests include artificial intelligence, machine learning, computer vision, and deep learning.

ALEX MARTYNENKO received the B.Sc. degree in electromechanical engineering from Ukrainian Agricultural Academy, in 1980, the M.Sc. degree from Moscow Agro Engineering University, in 1986, and the Ph.D. degree from the University of Guelph, Canada, in 2005. He is currently a Professor of bioelectronics and bioinstrumentation at Dalhousie University and a Professional Engineer, working in the area of industrial automation, robotics, machine vision, machine learning, and intelligent control systems. His unique expertise in system analysis, sensor fusion, control, and optimization resulted in multiple IoT applications for agriculture and food processing. He is the author and coauthor of four books and 12 patents. He is also a Reviewer of *Journal of Food Engineering* and *Drying Technology* and an Instructor of industry-oriented short courses.

FONS J. VERBEEK received the Ph.D. degree in applied physics from the Pattern Recognition Group, Delft University of Technology, The Netherlands, in 1995. He is currently a Professor of computational bio-imaging and pattern recognition at the Leiden Institute of Advanced Computer Science (LIACS), Leiden University, The Netherlands. His research is on development of robust methods for large-scale image processing and analysis in the biosciences. Therefore, his research interest includes classical and deep learning machine learning approaches. He works on reconstruction problems in different kinds of imaging techniques, techniques that are applied on a wide variety of organisms for which 3D imaging is important. The zebrafish model system is used in a range of different application areas in his research. In addition to image processing in general, his research involves information visualization and interactive installations for information visualization.