DOI: 10.15849/IJASCA.231130.13

Int. J. Advance Soft Compu. Appl, Vol. 15, No. 3, November 2023
Print ISSN: 2710-1274, Online ISSN: 2074-8523
Copyright © Al-Zaytoonah University of Jordan (ZUJ)

Improving Date Fruit Sorting with a Novel Multimodal Approach and CNNs

Ibtissam Boumaraf, Abdelhamid Djeffal, Sarah Setta, and Abdelmalik Taleb-Ahmed
Laboratory of LESIA, Department of Computer Science, University of Biskra, Algeria
e-mail: ibtissam.boumaraf@univ-biskra.dz
Laboratory of LESIA, Department of Computer Science, University of Biskra, Algeria
e-mail: abdelhamid.djeffal@univ-biskra.dz
Department of Computer Science, University of Biskra, Algeria
e-mail : settasarah7@gmail.com
IEMN UMR CNRS 8520, UPHF, Valenciennes, France
e-mail: abdelmalik.talebAhmed@univ-valenciennes.fr

Abstract

Date fruit is a beloved and widely consumed food in the Middle East and North Africa, and its popularity is growing globally. However, sorting these fruits can be a time-consuming and labor-intensive process, particularly when done manually. To address this challenge, we have proposed an innovative approach that uses multimodal data fusion and convolutional neural networks (CNNs) to efficiently classify Algerian date fruit. Our process involves capturing four RGB images of the date fruit from various angles, a thermal image, and the weight of the fruit. We create a new image where the first channel consists of a grayscale image obtained by averaging the four RGB images of the fruit. The second channel contains the thermal image, and the third channel contains the normalized weight data. The new dataset is then divided into training, validation, and testing sets. We conducted experiments using four different models: VGG16, InceptionV3, ResNet50, and Basic CNN. Our findings show that the VGG16 model achieved the highest accuracy during training, validation, and testing, with scores of 99.6%, 90.4%, and 94%, respectively. The InceptionV3 model had the lowest accuracy, while the ResNet50 and Basic CNN models had similar performances. Our results indicate that the VGG16 model is the most suitable for classifying Algerian date fruit. Our proposed approach offers a promising solution to improve efficiency and accuracy, ultimately enhancing the quality of sorted fruit and increasing its market value.

Keywords: *Convolutional Neural Network, Date Fruit, Image Classification, Multiscale Sorting Process, Thermal Image, Transfer Learning, Weight scale.*

1 Introduction

Date fruit plays a crucial role in Algeria’s economy and holds a special place in the hearts of many Algerians. This delicious fruit has a rich history and cultural significance in the country, and it is widely consumed both within Algeria and exported to other countries. Algeria is actually the third-largest producer of dates in the world, after Egypt and Iran [4], and this industry is a vital contributor to the country’s economy.

Received 7 July 2022; Accepted 30 October 2023

The Algerian government has implemented policies to support the growth and development of the date fruit industry. They recognize the importance of this industry to the country's economy and have taken measures to improve the quality and quantity of date production. Thanks to the country's abundant natural resources, including the Sahara Desert, the conditions are perfect for date palm cultivation. With the government's support, the Algerian date industry continues to thrive, providing a source of income for many Algerians and contributing to the country's overall economic growth [5].

Algeria exports date fruit to many countries, including France, Spain, and Italy, with exports estimated at approximately 26,000 tons in 2020 [6]. The importation of date fruit is also prevalent in Algeria, with dates imported from Tunisia, Morocco, and other countries to supplement domestic demand.

However, sorting date fruit is a time-consuming and labor-intensive process that requires manual effort. Therefore, the development of an automated approach for date fruit sorting has become a critical area of research for the Algerian date industry.

A multitude of approaches, spanning the domains of supervised and unsupervised machine learning techniques, have accumulated significant attention from researchers [1,2,3,4]. Their collective aim has been to enhance the efficiency and precision of the automated date fruit sorting and grading processes. However, despite the dedicated efforts, achieving a level of performance that meets the stringent demands of date fruit classification remains an enduring challenge for existing classification algorithms. This enduring challenge is primarily rooted in the intricacies of date fruit's diverse characteristics and appearances.

Our study presents an innovative approach to classifying Algerian date fruits, utilizing advanced techniques like multimodal data fusion and CNNs to significantly enhance accuracy. By incorporating additional data sources such as fruit weight, thermal imagery, and RGB images, our method not only improves sorting efficiency but also offers versatile applications in agriculture and computer vision. It holds the potential to reduce manual labor in date fruit sorting, making it adaptable for sorting various fruits and vegetables, promising superior food quality control, increased productivity, and cost-effectiveness. The significance of this research can be summarized as follows:

- Collecting and preparing a dataset of date fruit images that includes four RGB and thermal images, as well as their weight.
- Preprocessing the image data by undergoing grayscale transformation, image averaging, and setting the pixel value of image channels.
- Developing different CNN models, including VGG16, Inception V3, ResNet50, and Basic CNN, to accurately classify the date fruit images.
- Evaluating the performance of the CNN models using various metrics, such as accuracy, precision, recall, F1-Score, ROC AUC, Cohen's Kappa, and Matthews Correlation Coefficient (MCC) scores.
- Analyzing the results to assess the effectiveness of the proposed system for automated date fruit grading.

The paper is structured into different sections: the related work will be described in Section 2, dataset acquisition, proposed architecture, and data preprocessing in Section 3. Section 4 presents the experimental results and discussion, and finally, in Section 5, conclusions are drawn based on the study's findings.

2 Related Work

In the literature, traditional learning methods were employed for sorting and grading date fruits. Aiadi et al. (2017) [12] presented a comprehensive approach to categorize date fruits into ten distinct classes. They utilized a diverse range of features, including color, texture, cooccurrence matrix (GLCM), and shape attributes, and reduced dimensionality using Principal Component Analysis (PCA). Their research involved experimenting with various methodologies, such as Gaussian Mixture Models (GMM), the method introduced by Haidar et al. [15], Support Vector Machines (SVM), and Decision Stumps (DS).

Kamal-Eldin et al. (2018) [13] employed cluster analysis and Principal Coordinate Analysis (PCoA) to classify Emirati date fruits. Their focus was on attributes like color, size, and malic acid content, enabling the differentiation of 20 Emirati date varieties.
In a more recent study in 2021, Murat et al. [14] introduced a stacking method that combined Logistic Regression (LR) and Artificial Neural Networks (ANN) to classify a variety of date fruit types, including Barhee, Deglet Nour, Sukkari, Rotab Mazafati, Ruthana, Safawi, and Sagai. Impressively, this stacking model achieved an accuracy rate of 92.8%, surpassing the individual LR and ANN models, which achieved accuracy rates of 91% and 92%, respectively.

Nasiri et al. [7] used a controlled environment and the VGG16 architecture on a relatively small dataset of four date fruit classes. They achieved an impressive 98.49% classification accuracy but acknowledged limitations, including the small dataset size and lack of real-world testing. Altaher et al. [2] employed a fine-tuned VGG-16 model on a substantial dataset of over 8,000 images of date fruits from various varieties. Their deep learning model achieved an overall accuracy of 95.25%. Limitations included a focus on visual characteristics and a lack of consideration for non-visual factors and viewing angles.

Alhamdan et al. [3] used CNNs on a dataset of over 5,000 images of date fruits from six different types, captured in a controlled setting. The CNN model achieved an overall accuracy of 94.8% and excelled in identifying different date fruit varieties. Pérez-Pérez et al. [1] utilized transfer learning on a dataset of over 3,000 images of Medjool dates sorted manually into "good" and "poor" categories. They achieved an overall accuracy of 92.53% in classifying ripe Medjool dates based on external appearance.

While these studies primarily focused on the visual characteristics of date fruits and did not consider the impact of different viewing angles, based on various studies, it has been found that convolutional neural networks (CNNs) have delivered positive outcomes in sorting date fruits using different image acquisition techniques, datasets, and training methodologies. In this context, a fresh research approach has been employed for the classification of Algerian date fruits. The method involves multimodal data fusion and CNNs, resulting in improved accuracy during the classification process. This is achieved by integrating additional data sources, such as fruit weight, thermal images, and four RGB images of four angle views of date fruit.

3 The Proposed Method

In this research paper, the architecture of the proposed method is described in detail, as illustrated in Figure 1. The first step is the Dataset Acquisition phase, where a dataset of date fruits is collected, including 4 RGB images, 1 thermal image, and weight data for each fruit. The Data Preprocessing phase follows, where the data is transformed into a suitable format through grayscale transformation, image averaging, and customizing image channel values. The dataset is then divided into three sets: training, validation, and testing, with 70%, 20%, and 10% of the overall data allocated to each set, respectively. To select the best-performing model based on the evaluation results, various CNN models such as VGG16, InceptionV3, ResNet50, and Basic CNN are trained and validated using the Training and Validation sets. The trained models are then evaluated using the Testing set and various performance measures during the Testing and Evaluation stage. Finally, an application is created that utilizes the trained CNN models to classify date fruits into Deglet-Noor and Mech-Degla varieties. Each variety is graded into five based on quality.

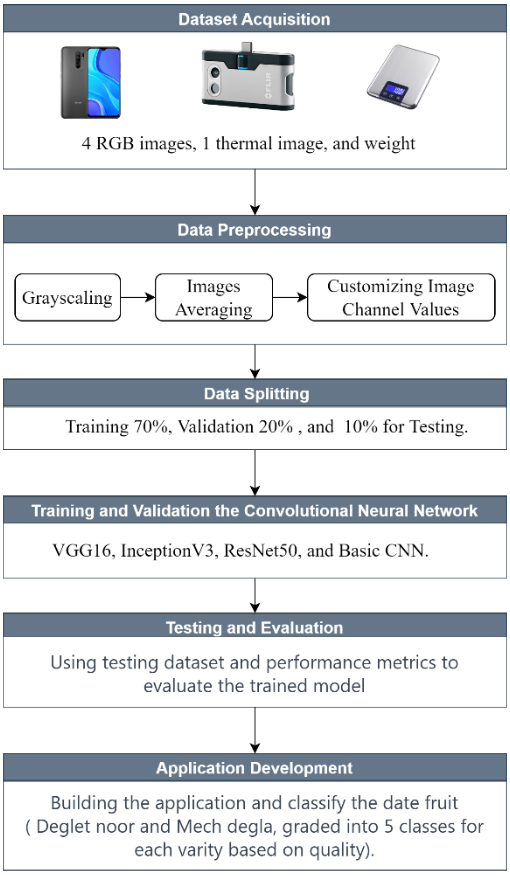


Figure 1 A diagram presenting the overview architecture.

3.1. Dataset acquisition

In this phase of our research, we assembled a dataset containing two date fruit varieties: Deglet Noor and Mech Degla. Each variety was categorized into five quality classes. Deglet Noor is a renowned date fruit variety characterized by its semi-moist texture, glossy and smooth appearance, and beige-brown color. In its highest quality state, it is free of impurities and blemishes. However, as its quality decreases, it exhibits an increased percentage of blemishes, altered luster, and reduced moisture content. Mech Degla, another Algerian date fruit variety, features a cylindrical shape with a slight taper at the tip. It boasts a light beige hue with a faint brownish tint and a non-glossy skin that's free of imperfections. The mesocarp of Mech Degla is fleshy and possesses a dry consistency with a fibrous texture. As its quality diminishes, it experiences greater shrinkage, resulting in the formation of wrinkles, decreased humidity levels, increased fragility, and the appearance of spots on the dates. The lowest-quality type is considered inedible and is referred to as a glans. In Figure 2 , we present an example from our dataset, showcasing both date fruit varieties and their respective quality grades. Each date fruit was captured from four different angles using an RGB camera, resulting in four RGB images per fruit. Additionally, we recorded a thermal image using a Flir camera and noted the weight scale reading for each fruit, as illustrated in Figure 3.

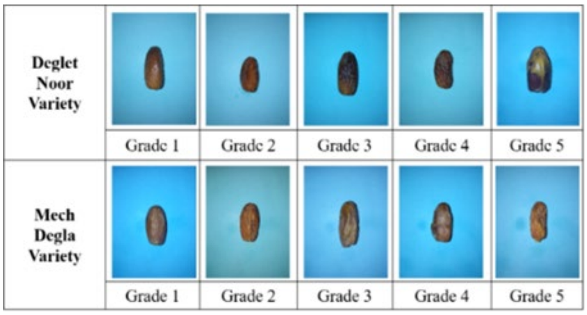


Figure 3 An example of our dataset includes two varieties of Algerian date fruit and their grades

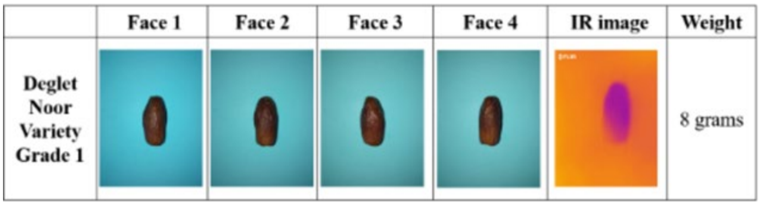


Figure 2 An example of Deglet Noor Grade 1, showcasing the features used in our dataset.

To ensure high-quality image capture, we employed four fluorescent lamps to provide consistent lighting and minimize any potential image quality issues. The images were captured against a uniformly colored background with a chosen blue hue, which standardized the images and simplified subsequent image processing. There were no size restrictions during image capture, allowing us to represent the fruits in their natural state. The dataset comprises approximately 1,103 date fruits, consisting of two varieties: Deglet-Noor and Mech-Degla. Each variety is further subdivided into five grades based on quality. Deglet-Noor has 109 date fruits in Grade 1, 105 in Grade 2, 69 in Grade 3, 104 in Grade 4, and 80 in Grade 5. Mech-Degla is also classified into five grades with 63 date fruits in Grade 1, 120 in Grade 2, 203 in Grade 3, 140 in Grade 4, and 110 in Grade 5.

3.2. Data Preprocessing

To prepare data for machine learning algorithms, a sequence of processes is required to convert raw data into a format suitable for further analysis. This process is referred to as data preprocessing. In our study, we carried out three key operations during the data preprocessing phase: image grayscaling, image averaging, and customizing image channel values.

3.2.1. Image Grayscale Transformation

In the data preprocessing step, we apply grayscale transformation to the five images captured for each date fruit, including the four RGB images of each face (view) and a thermal image (IR). Grayscaling reduces the color channels from three (RGB) to one (grayscale), simplifying the image data and reducing the input size required for further image processing steps. This process is illustrated in Figure 4 and Eq. 1 with parameters and variables for grayscale conversion of different images:

$$Gray(img)=0.2989\times R(img)+0.5870\times G(img)+0.1140\times B(img). \qquad \text{Eq. 1}$$

Gray(img) represents the grayscale image obtained from the input image, where R(img), G(img), and B(img) represent the red, green, and blue color channels of the input image, respectively. This equation can be applied to all images by replacing 'image' with the specific image, such as $Gray_{F1}$, $Gray_{F2}$, $Gray_{F3}$, $Gray_{F4}$, or $Gray_{IR}$. Where F=Face number.

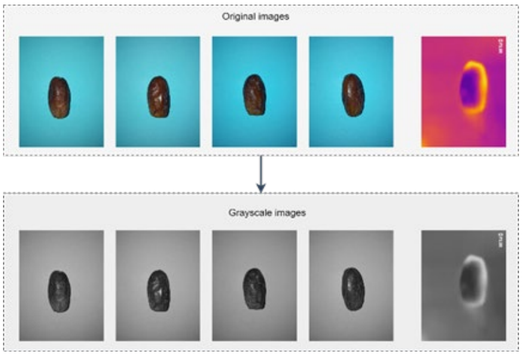


Figure 4 The image grayscale transformation is applied to all the images

3.2.2. Image Averaging

During the image averaging step in our data preprocessing, we calculate the average of the four grayscale images corresponding to each side of the date fruit. This results in a single grayscale image that represents the average appearance of the fruit from all four sides. This step helps to reduce noise and ensure consistency in the input data for the subsequent image processing steps. The output of this step is a single grayscale image per fruit, this process is illustrated in Figure 5. Eq. 2 describes the process of averaging the grayscale images of the four sides of the date fruit to obtain a single averaged grayscale image.

$$Gray(x,y)_{avg}=(Gray(x,y)_{F1}+Gray(x,y)_{F2}+Gray(x,y)_{F3}+Gray(x,y)_{F4})/4. \qquad \text{Eq. 2}$$

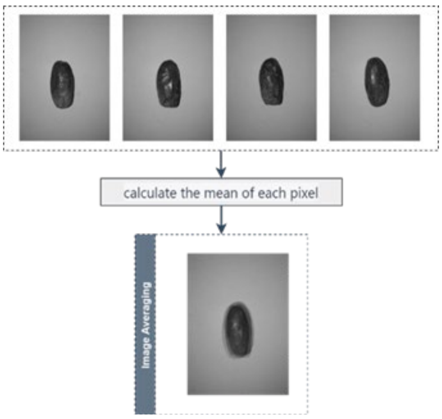


Figure 5 The process of image averaging using four input images.

3.2.2. Customizing Image Channel Values

The customizing of image channel values is the final step in our data preprocessing phase, where we create the resulting image that serves as the input for our CNN network. In this step, we standardize the size of the IR image with the image obtained from the concatenation of the four faces of the date fruit, which was calculated in the previous step. Next, we create a new image where each pixel is defined as follows: the red value corresponds to the grayscale value of the averaged image obtained from the previous step, the green value corresponds to the grayscale value of the FLIR image, and the blue value corresponds to the weight of the date fruit, as present in Eq. 3.Weight represents the normalized weight of the date fruit and then scaled to an integer value between 0 and 255 as Eq. 4 where 20.0 represents the maximum value of the date fruit weight. This process is repeated for all pixels of the resulting image. This step is crucial in creating a standardized input format for our CNN network as it provides a balanced combination of color, thermal, and weight information for each date fruit. The process is illustrated in
Figure 6.

$$\begin{cases} Resulting\_Image(x,y) = (R\_Value, G\_Value, B\_Value). \\ R\_Value = Gray(x,y)_{avg}. \\ G\_Value = Gray(x,y)_{IR}. \\ B\_Value = Scaled\_Weight. \end{cases} \quad \text{Eq. 3}$$

$$\begin{cases} Normalized\_Weight = (Weight\_original - 1)/20.0 \\ Scaled\_Weight = round(Normalized\_Weight * 255) \end{cases} \quad \text{Eq. 4}$$

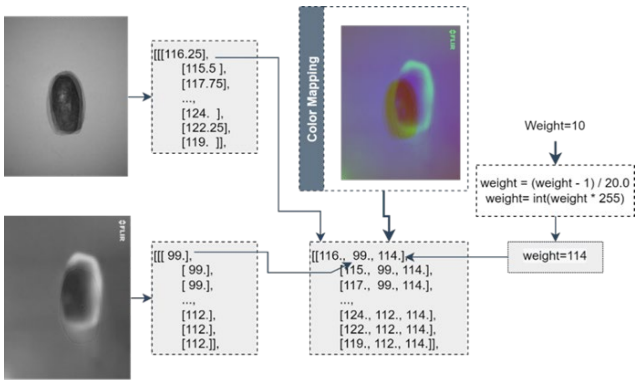


Figure 6 Illustration of the customizing of image channel values using the average image, IR image, and weight.

3.3. Model Conception

After completing the data preprocessing stage, we proceeded to the model conception phase in our research paper. Our objective was to classify dates based on their quality using various deep learning models, including VGG16, InceptionV3, ResNet50, and a Basic CNN model:

- **VGG16** is a variant of the VGGNet architecture, designed to reduce the number of parameters in convolutional layers and improve training time. This model comprises a total of 138 million parameters, featuring 3x3 convolutional kernels and 2x2 max-pooling kernels with a stride of two. It is configured to accept input images of size 224x224 pixels [16][17].

- **InceptionV3**, also known as GoogLeNet, leverages the inception module to approximate a sparse convolutional neural network (CNN) with a normal dense construction. The InceptionV3 architecture, as implemented in Keras, incorporates updates proposed by Szegedy et al. [18] to enhance ImageNet classification accuracy. It is designed to process input images of size 299x299 pixels [17] [19].

- **ResNet50** is a variant of the ResNet architecture, renowned for its depth and utilization of residual connections to combat the vanishing gradient problem. Comprising 50 layers, ResNet50 is one of the widely adopted versions of the ResNet architecture. It is configured to handle input images of size 224x224 pixels [16][17].

- **A basic Convolutional Neural Network (CNN)** is a foundational deep learning model frequently employed for image classification tasks. It comprises multiple convolutional layers, pooling layers, and fully connected layers responsible for learning and feature extraction from input images. The architecture and performance of a basic CNN can vary depending on specific design choices and hyperparameters during training [20].

3.3.1. Transfer Learning Models

In our work, we utilized transfer learning with pre-trained models, namely VGG16, InceptionV3, and ResNet, while maintaining the same hyperparameters for all models except for the base model selection. We performed fine-tuning by unfreezing all block layers for each model. Additionally, we augmented the pre-trained convolutional neural network (CNN) by incorporating a global average pooling 2D layer, a dropout layer, and two dense layers. The GlobalAveragePooling2D layer calculated the average of each feature map in the output tensor, reducing its spatial dimensions while preserving crucial features. To prevent overfitting, dropout layers were introduced during training, and the final image classification was conducted using dense layers with the ReLU activation function. For all transfer learning models, the dropout rate was set to 0.2. The architecture employed is depicted in Figure 7.

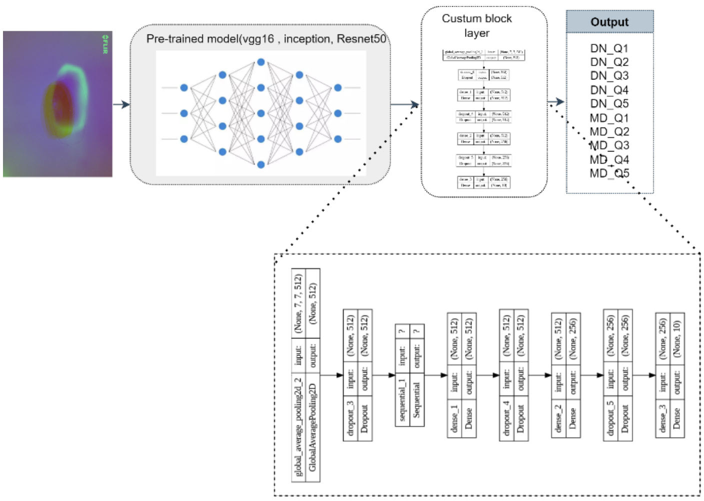


Figure 7 Transfer learning model architecture.

3.3.2. Basic CNN Model

The Basic CNN model is a convolutional neural network that consists of four convolutional layers and max-pooling layers to reduce the spatial dimensions of the output feature maps. During training, two dropout layers were included to mitigate overfitting. The first convolutional layer employs 64 filters of size (3,3) with a ReLU activation function, and the input shape is (224, 224, 3). The subsequent convolutional layers have 64 and 128 filters of size (3,3) with the ReLU activation function, respectively. The final convolutional layer comprises 128 filters of size (3,3) with a ReLU activation function, followed by a dropout layer with a rate of 0.3. The output of the last max-pooling layer is flattened before being fed into the hidden layer of a dense neural network with 512 neurons and a ReLU activation function. To address overfitting, an additional dropout layer with a rate of 0.4 is applied. Finally, the output layer consists of 10 neurons with a softmax activation function. The architecture employed is depicted in Figure 8.

Layer (type)	Output Shape	Param #
conv2d_98 (Conv2D)	(None, 222, 222, 64)	1792
max_pooling2d_8 (MaxPooling2D)	(None, 111, 111, 64)	0
conv2d_99 (Conv2D)	(None, 109, 109, 64)	36928
max_pooling2d_9 (MaxPooling2D)	(None, 54, 54, 64)	0
conv2d_100 (Conv2D)	(None, 52, 52, 128)	73856
max_pooling2d_10 (MaxPooling2D)	(None, 26, 26, 128)	0
conv2d_101 (Conv2D)	(None, 24, 24, 128)	147584
max_pooling2d_11 (MaxPooling2D)	(None, 12, 12, 128)	0
dropout_14 (Dropout)	(None, 12, 12, 128)	0
flatten_1 (Flatten)	(None, 18432)	0
dense_14 (Dense)	(None, 512)	9437696
dropout_15 (Dropout)	(None, 512)	0
dense_15 (Dense)	(None, 10)	5130

Figure 8 Summarized Basic CNN Model

3.3.3. Evaluation metrics

In the assessment of classification models, several essential metrics are utilized. Accuracy measures the proportion of correctly classified instances out of all instances and serves as the primary evaluation metric [8]. Precision quantifies the ratio of true positive predictions among all positive predictions, while recall calculates the percentage of true positive predictions among all instances of positive outcomes [8]. The F1 score, a harmonic mean of precision and recall, provides a balanced evaluation of these two metrics [8]. Furthermore, the ROC AUC score assesses binary classifier performance at different thresholds [9], while the Cohen Kappa score evaluates agreement between predicted and actual class labels, accounting for chance agreement [10]. Lastly, Matthew's correlation coefficient (MCC), with values ranging from -1 to 1, considers true positives, true negatives, false positives, and false negatives to gauge classifier performance, with 1 indicating perfection, 0 indicating randomness, and -1 indicating complete incorrectness [11]. Table 1 displays the formulas for these evaluation metrics:

Table 1 Evaluation Metrics Formula. Probability of agreement (P_o), probability of random agreement (P_e)

Assessments	Formula
Accuracy	$TP + TN/TP + TN + FP + FN.$
Precision (P)	$TP/TP + FP.$
Recall (R)	$TP/TP + FN.$
F1-Score	$2 \times (P \times R/P + R).$
MMC	$\frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}.$
Cohen Kappa score	$(P\_o - P\_e)/1 - P\_e.$

4. Results and Discussion

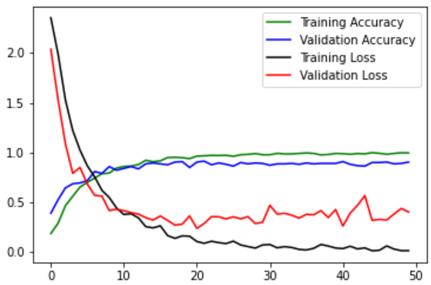
For the project, Python 3.9.6 was used on a Windows 10 Pro machine equipped with an Intel(R) Core (TM) i5-6200U CPU running at 2.30 GHz and 2.40 GHz, and 4.00 GB of RAM. The installed versions of TensorFlow and Keras were 2.6.0. Following the training and validation stages, we achieved specific results for four distinct models, using 70% of the data for training and 20% for validation. We trained all models using the Adam optimizer with a learning rate of 0.00001, a batch size of 16, and an image size of 244x244 pixels, except for InceptionV3, which used a picture size of 299x299 pixels:

- The VGG16 model (Figure 9 (a)) achieved a training accuracy of 99.6% and a validation accuracy of 90.4%, with a loss of 0.0153 during training. The loss decreased significantly, while the accuracy increased, indicating effective learning. Validation accuracy remained consistently high at around 88% to 91% after epoch 17.

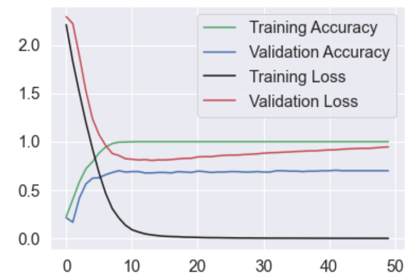
- The Inception V3 model (Figure 9 (b)) achieved 100% training accuracy but had a lower validation accuracy of 69.9%. Training started with a high loss and low accuracy, gradually improving as training progressed. Validation accuracy was generally lower, suggesting some degree of overfitting. The model achieved 100% training accuracy after 7 epochs but had limited generalization to new data.

- The results of Resnet50 (Figure 9 (c)) showed that the model achieved 100% training accuracy but had a stabilized validation accuracy of around 78%, indicating overfitting. Both the loss and validation loss decreased initially, but after epoch 10, the validation loss started to rise, indicating overfitting due to the model's complexity with 50 layers.

- The basic CNN model (Figure 9(d)) achieved an 87.5% training accuracy and an 81.4% validation accuracy with a loss of 0.3893 during training. As epochs increased, the training loss decreased, and accuracy improved, as seen in Figure 8d. Validation metrics also showed improvement, suggesting effective generalization to new data.

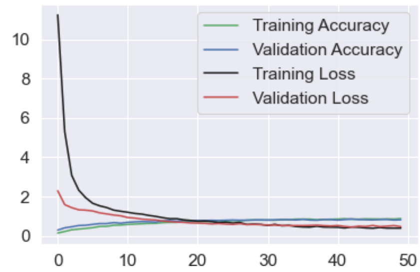


(a) Accuracy and Loss of the VGG16 model.



(b) Accuracy and Loss of the Inception V3 model.

(c) Accuracy and Loss of the ResNet50 model.



(d) Accuracy and Loss of the Basic CNN model.

Figure 9 Curves of Loss and accuracy during the model training for four models

We evaluate the efficiency of our models by subjecting them to a 10% test dataset and gauging their performance through different performance metrics, which are elaborated on in Section 3.3.3, with the aim of predicting how well the model will perform on new.

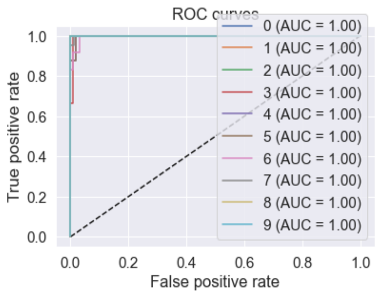
- The VGG16 Model (Figure 10(a)) achieved exceptional performance across all classes, with ROC AUC values approaching 1.0, indicating high accuracy. Additionally, the Kappa and MCC scores both yielded 0.93, signifying almost perfect agreement between predicted and actual labels (Figure 12(a)). based on Table 3, the F1-score of 93.69% demonstrated a desirable balance between precision and recall. Precision and recall were 93.75% and 93.69%, respectively, with an overall accuracy of 93.69%.

- The Inception V3 Model (Figure 10(b)) obtained a ROC AUC score of 0.9609, signifying strong performance in categorizing all classes. However, the confusion matrix (Figure 11(b)) showed variability in performance across several classes. The Cohen Kappa score and Matthew's correlation coefficient of 0.7332 and 0.7359, respectively, indicated substantial but not perfect agreement (Figure 12(b)). based on Table 3, The model achieved a moderate balance between precision and recall, with an F1-score of 76.58%. Precision was 80.56%, recall was 76.58%, and accuracy was 76.58%.

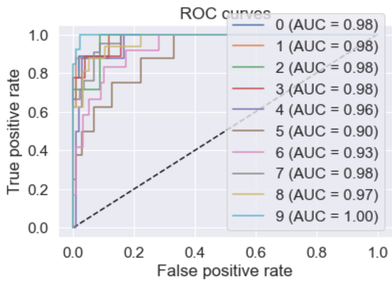
- Results for the ResNet50 Model (Figure 10(c)) revealed high performance in most classes, with perfect ROC AUC values for five classes and values exceeding 0.98 for the remainder. The confusion matrix further demonstrated high accuracy across all classes (Figure 11(b)). The Cohen Kappa score and Matthew's correlation coefficient of 0.8259 and 0.8285, respectively, indicated substantial agreement but not perfection (Figure 12(c)). based on Table 3, The model achieved a satisfactory trade-off between precision and recall, with an F1-score of 84.68%. Precision was 87.50%, recall was 84.68%, and accuracy was 84.68%.

- The Basic CNN Model (Figure 10(d)) displayed excellent performance with ROC AUC values exceeding 0.96 for all classes. It effectively distinguished between different classes . The confusion matrix illustrated moderate predictions for each class (Figure 11(d)). The Cohen Kappa score and Matthew’s correlation coefficient of 0.8263 and 0.8277, respectively, indicated substantial but not perfect agreement (Figure 12(d)). based on Table 3, The model achieved a moderate balance between precision and recall, with an F1-score of 84.68%. Precision was 80%, recall was 84.68%, and accuracy was 84.68%.

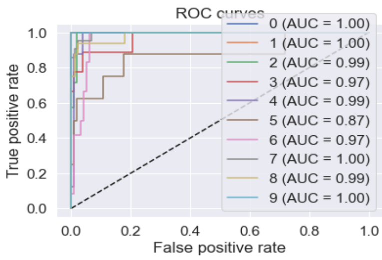
The Table 2 shows the training and validation results for four different models, including VGG16, InceptionV3, ResNet50, and Basic CNN. The training results show that all models achieved high accuracy during training, with VGG16 model having a 99.6%, InceptionV3 and Res-Net50 models having a 100% accuracy rate, and Basic CNN model having a 87,5% accuracy rate. However, during validation, the VGG16 achieved the highest accuracy rate of 90.4%, while the InceptionV3 model had the lowest accuracy rate of 69.9%. The results of testing the four models are presented in Table 3 and Figure 13.
These results indicate that the VGG16 model outperformed the other models in terms of the F1-Score, precision, recall, and accuracy, indicating that it was the most effective in classifying the images. On the other hand, the InceptionV3 model had the lowest F1-Score, precision, recall, and accuracy, indicating that it was the least effective. The ResNet50 and Basic CNN models had comparable F1-Score, precision, recall, and accuracy values, with the Basic CNN model having a slightly lower F1-Score and precision but a higher ROC AUC value.



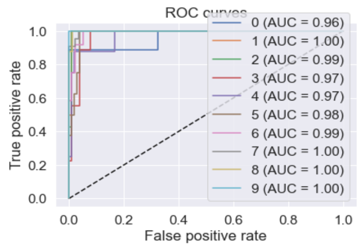
(a) The ROC AUC curve of the VGG16 model.



(b) The ROC AUC curve of the Inception V3 model.

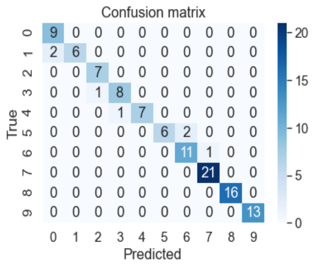


(c) The ROC AUC curve of the ResNet50 model.

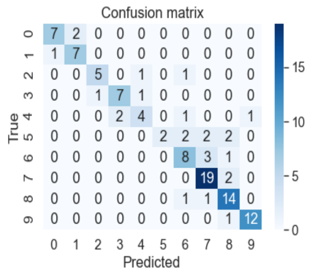


(d) The ROC AUC curve of the Basic CNN model.

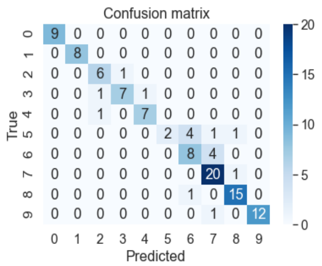
Figure 10 The ROC AUC curve for various models



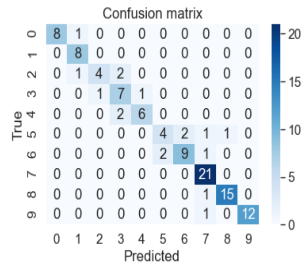
(a) The confusion matrix of the VGG16 model.



(b) The confusion matrix of the Inception V3

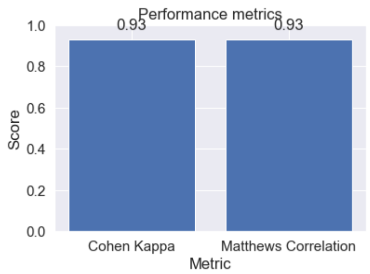


c) The confusion matrix of the ResNet50 model.

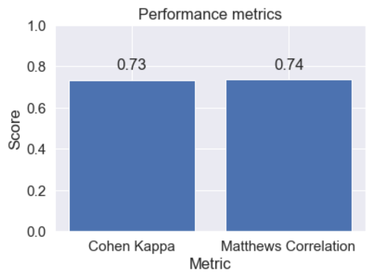


(d) The confusion matrix of the Basic CNN model.

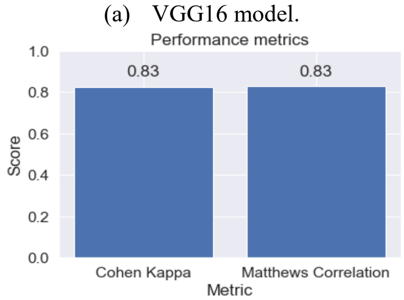
Figure 11 The confusion matrix: (a) VGG16 (b) InceptionV3 (c) ResNet50 (d) Basic CNN.



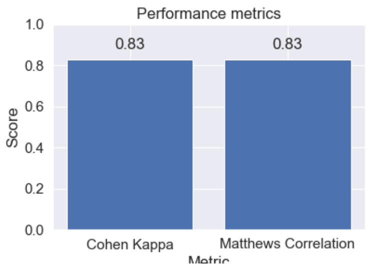
(a) VGG16 model.



(b) Inception V3 model.



(c) ResNet50 model.



(d) Basic CNN model.

Figure 12 The kappa and Matthews scores

Table 2 Training and Validation Results for Various Models

Model Type	Training results		Validation results	
	Loss	Accuracy %	Loss	Accuracy %
VGG16	0.0153	99.6	0.4027	90.4
InceptionV3	0.0010	100	0.9461	69.9
ResNet50	0.0009	100	0.7274	78.64
ResNet50	0.3893	87.5	0.4684	81.4

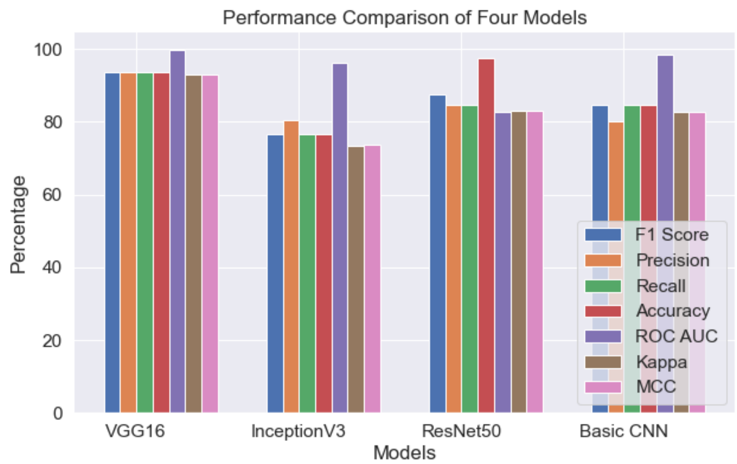
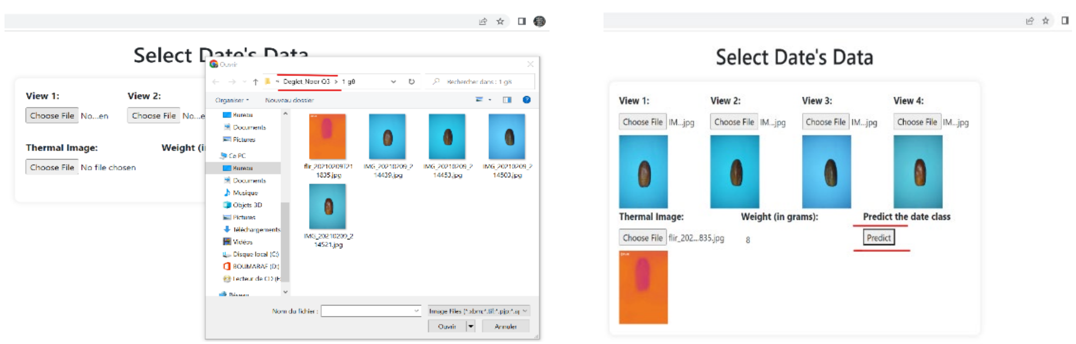


Figure 13 Performance Comparison of Four Models.

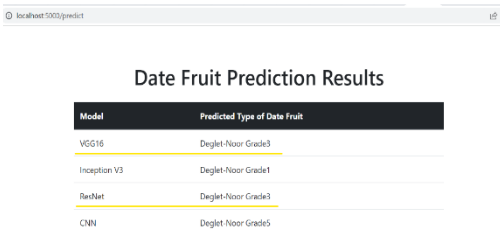
Table 3 Testing Results for Four Models.

Model	F1-Score%	Precision %	Recall %	Accuracy %	ROC AUC	Kappa	MCC
VGG16	93.69	93.75	93.69	93.69	0.998	0.9287	0.9293
InceptionV3	76.58	80.56	76.58	76.58	0.9609	0.7333	0.7359
ResNet50	84.68	87.50	84.68	84.68	0.9757	0.8259	0.8285
Basic CNN	84.68	80.00	84.68	84.68	0.9837	0.8263	0.8277

We developed a Flask–based application to sort dates according to their quality and make predictions. To use our model, we first select the date fruit we want to classify using both thermal and four–face images and input the weight of the fruit. In this case, we used Deglet Noor Q3, as shown in Figure 14(a). Figure 14(b) displaying the data for prediction. This interface presents all the information entered by the user, including the thermal and four-face images of the chosen date fruit, and its weight. This information is used to make a prediction about the class of the date fruit. After making the prediction, we observed the results in a table of Figure 14 (c), which indicated the predicted class by each of the four models. The VGG16 and Resnet50 models correctly predicted the class of the date fruit, while the other models did not, confirming the superior performance of VGG16 in our date fruit classification task. Our results show the benefit of using multiscale date fruit images to improve prediction accuracy.



(a) (b)



(c)

Figure 14 Application Interface for Date Fruit Classification and Prediction.

5 Conclusion

Our research has shown promising results for efficiently classifying Algerian date fruit using a combination of multimodal data fusion and convolutional neural networks (CNNs). To simplify and standardize the input data for our CNN models, we utilized grayscaling, images averaging, and customizing image channel values during the data preprocessing steps. This led to high accuracy rates during training, validation, and testing.

During our experimentation, we found that the VGG16 model was the most effective in classifying Algerian date fruit, achieving a training accuracy rate of 99.6% and a testing accuracy rate of 94%. However, the InceptionV3 model showed the lowest accuracy rate, while the ResNet50 and Basic CNN models performed similarly. By implementing our approach, we can significantly improve the efficiency and accuracy of date fruit sorting, which is typically a labor-intensive and time-consuming process when performed manually. This can enhance the quality of sorted fruit and ultimately increase its market value.

Future work can involve applying our proposed method to other types of date fruit and other fruits in general. Further optimization of the CNN models can also be explored to improve their performance. Our proposed approach presents a promising solution for the efficient and accurate classification of date fruit, contributing to the advancement of the date industry.

ACKNOWLEDGEMENTS

We are sincerely grateful to the chief of the LESIA laboratory at Mohamed Khider University in Biskra, Algeria, for their invaluable assistance throughout our research project. Additionally, we would like to express our appreciation to the Agro-Date factory in Biskra for generously sharing their knowledge and expertise on the date fruit sorting process.

References

[1] Pérez-Pérez, B. D., García Vázquez, J. P., & Salomón-Torres, R. (2021). Evaluation of Convolutional Neural Networks' Hyperparameters with Transfer Learning to Determine Sorting of Ripe Medjool Dates. Agriculture, 11(2), 115. https://doi.org/10.3390/agriculture11020115

[2] Altaheri, Hamdi & Alsulaiman, Mansour & Muhammad, Ghulam. (2019). Date Fruit Classification for Robotic Harvesting in a Natural Environment Using Deep Learning. IEEE Access. PP.

[3] Alhamdan, W. & Howe, J. (2021). Classification of date fruits in a controlled environment using Convolutional Neural Networks. AMLTA: Advanced Machine Learning Technologies And Applications. 1339 (pp. 154-163), https://openaccess.city.ac.uk/id/eprint/25158/

[4] FAOSTAT Production Quantities of Dates in 2020 by Country. (http://www.fao.org/faostat/en/#data/QC,2021)

[5] Ministère de l'Agriculture et du Développement Rural Stratégie Nationale de Développement de la Filière des Dattes. (Http://www.minagri.dz/index.php/fr/component/content/article?id=398,2017)

[6] World's Top Exports Dates Exports by Country in 2020. (http://www.worldstopexports.com/datesexports-by-country/,2021)

[7] Nasiri, Amin & Taheri-Garavand, Amin & Zhang, Yudong. (2019). Image-based deep learning automated sorting of date fruit. Postharvest Biology and Technology. 153. 133-141. 10.1016/j.postharvbio.2019.04.003.

[8] Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. International journal of data mining & knowledge management process, 5(2), 1.

[9] Sokolova, M., Japkowicz, N. & Szpakowicz, S. (2006). Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. In A. Sattar & B.-H. Kang (eds.), Australian Conference on Artificial Intelligence (p./pp. 1015-1021), : Springer. ISBN: 3-540-49787-0

[10] Carletta, J. The Kappa statistic: A second look. Proceedings Of The 17th International ConferenceOn Computational Linguistics-Volume 1. pp. 1-7 (1998)

[11] Chicco, D., Tötsch, N., & Jurman, G. (2021). The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. BioData Mining, 14.

[12] O. Aiadi and M. L. Kherfi. (2017). A new method for automatic date fruit classification. International Journal of Computational Vision and Robotics, vol. 7, no. 6, (pp. 692-711).

[13] A. Kamal-Eldin and S. Ghnimi. (2018). Classification of date fruit (Phoenix dactylifera, L.) based on chemometric analysis with multivariate approach. Food Measure, vol. 12, (pp. 1020-1027).

[14] M. Koklu, R. Kursun, Y. S. Taspinar and I. Cinar. (2021). Classification of Date Fruits into Genetic Varieties Using Image Analysis. Mathematical Problems in Engineering, vol. 2021.

[15] Haidar, A., Haiwei, D. and Mavridis, N. (2012). Image-based date fruit classification. Proc. of the 4th International Congress on Ultra-Modern Telecommunications and Control Systems and Workshops (ICUMT), pp.357–363.

[16] Anwar, A. (2022, January 22). *Difference between alexnet, vggnet, ResNet and inception*. Medium. https://towardsdatascience.com/the-w3h-of-alexnet-vggnet-resnet-and-inception-7baaaecccc96

[17] Rosebrock, A. (2021, June 17). *ImageNet: Vggnet, ResNet, inception, and xception with keras*. PyImageSearch. https://pyimagesearch.com/2017/03/20/imagenet-vggnet-resnet-inception-xception-keras/
[18] *ResNet, Alexnet, vggnet, inception: Understanding various architectures of Convolutional Networks*. CV. (2017, August 9). https://cv-tricks.com/cnn/understand-resnet-alexnet-vgg-inception/
[19] Christian, Szegedy., Wei, Liu., Yangqing, Jia., Pierre, Sermanet., Scott, Reed., Dragomir, Anguelov., Dumitru, Erhan., Vincent, Vanhoucke., Andrew, Rabinovich. (2015). Going deeper with convolutions. 1-9. doi: 10.1109/CVPR.2015.7298594
[20] Mishra, M. (2020, September 2). *Convolutional Neural Networks, explained*. Medium. https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939

Notes on contributors

Ibtissam Boumaraf received the Master's degree in Computer Science from Mohamed Khider University, Biskra, Algeria, in 2016. She is currently pursuing the Ph.D. degree in Applied Computer Science at LESIA Laboratory, Mohamed Khider University, Biskra, Algeria. Her research interests include machine learning, deep learning, Agricultural sector and image processing.

Abdelhamid Djeffal obtained his PHD from Biskra university in 2012, Master degree in Image processing and AI from Biskra university in 2004, Engineer degree in computer science from National Institute for computer Science at Algiers in 1997. Assistant Professor in the Department of Computer science in the University of Biskra since December 2004. Also, he is a member of LESIA laboratory. His research interests are Artificial Intelligence, Image Processing, Websites, Data Mining, Machine Learning.

Sarah Setta received a Master's degree in Computer Science from Mohamed Khider University in Biskra, Algeria, in 2022. Her specialty is Information Systems, Optimization, and Decision (SIOD). Her research interests include machine learning, deep learning, the agricultural sector, and image processing.

Abdelmalik Taleb-Ahmed has been a Full Professor with the University of Polytechnique Hauts-de-France, Valenciennes, France, since 2004. His research interests include segmentation, classification, data fusion, pattern recognition, computer vision, and machine learning, with applications in biometrics, video surveillance, autonomous driving, and medical imaging.