Date Fruit Classification using Transfer Learning Techniques

Ghazanfar Latif
Department of Computer Science
Prince Mohammad Bin Fahd
University, Khobar, Saudi Arabia
glatif@pmu.edu.sa

Ghassen Ben Brahim
Department of Computer Science
Prince Mohammad Bin Fahd
University, Khobar, Saudi Arabia
gbrahim@pmu.edu.sa

Taha Houda
Department of Computer Science
Prince Mohammad Bin Fahd
University, Khobar, Saudi Arabia
glatif@pmu.edu.sa

Jaafar Alghazo
Artificial Intelligence Research
Initiative, University of North
Dakota, Grand Forks, ND, USA
jaafar.alghazo@und.edu

***Abstract*— Date fruit industry is considered among the important sectors in Saudi Arabia due to what this particular fruit represents to Saudi citizens in terms of cultural heritage significance. Among the several challenges that this industry is facing is the automation of the sorting process of dates which do exists in a wide variety of types, colors, size, taste, shapes, etc. This paper proposes developing an automatic classification model capable of accurately identifying different types of date fruits and providing detailed information about each type. This initial phase of this work consists of creating a large dataset of 15 types of date fruits by taking high-quality pictures, particularly from markets in Saudi Arabia. These types include the following: Safry, Mabroom almadeena, Sokary, Ajwat almadeena, Sagie, Khlas, Mjdwal, Ratab, Lobana, Rotana, Eedyah, Safawy, Khodhry, Rshodya, and Anbar. In the second phase of this resreach, three different CNN models (AlexNet, GoogleNet, VGG16) based on Transfer Learning were considered with Non-Freeze weights. Along with these models, two levels of fine-tuning were considered: (1) freezing all feature extraction layers and unfreezing the fully connected layers, where classification choices are determined; and (2) freezing beginning feature extraction layers and unfreezing the latter feature extraction and completely linked layers. Performance results affirm the efficacy of transfer learning in fruit classification tasks. The deep-structured VGG16 model, when fine-tuned with frozen weights, demonstrates exceptional validation accuracy and excellent scores in precision, recall, and F1 measure reaching 98.19%, 97.82%, 98.03%, and 98.10%, respectively.**

Keywords*— *Date Fruit, Convolution Neural Network, Deep Learning, Dates Classification, Transfer Learning

I. INTRODUCTION

Machine learning has a multitude of applications, including in both the industrial and commercial sectors. Machine learning algorithms are used in industries such as banking, security, manufacturing, agriculture, and others. To address a critical challenge in the agricultural sector, this paper focuses on the application of machine learning in the date fruit industry, particularly in Saudi Arabia. In Saudi Arabia, the date fruit industry is significant, with the country producing approximately 1.3 million tons out of the 8.5 million tons of worldwide production. Therefore, Saudi Arabia has a substantial market share in the date fruit market. Currently, the date fruit farming business is conducted entirely manually, and thus, depending on the type, date fruits can be very expensive when they reach end customers. This paper aims to tackle the inefficiencies and high expenses associated with manual labor in date fruit farming by introducing machine learning solutions. Research towards using machine learning to automate the farming process of date fruits is crucial; however, the research is still very limited. To the knowledge of the authors, there is still no public dataset for date fruits to aid researchers in this task. Many of the research findings on other fruits may be applicable to date fruits and vice versa. China and India are the largest fruit producers. There is a limited number of skilled workers, even in populous countries such as India and China, and thus it is estimated that approximately 35% of the harvested fruits are wasted.

In date fruit farms, a process of sorting, checking ripeness, and identifying types must be done prior to sale. This process is usually performed manually by workers, who must inspect each tree for ripeness before picking the dates. Once picked, workers need to inspect each date to remove any that are bad or might be infected by diseases. The process then continues with the separation of dates by type before they are boxed and sent to the market. By employing machine learning, each tree can be monitored independently, with an automated alert system indicating which trees are ready for harvesting. This would involve the use of drones, cameras, and a backend machine learning algorithm trained on the various types of dates and their different ripeness levels. Date trees, like many plants, are susceptible to various diseases. Therefore, the machine learning algorithm would also need to be trained to identify multiple diseases that could infect date fruits. During the regular ripeness checks, the system could also detect diseases and use the automated alert system to inform the farm's central control. Another example of machine learning application in agriculture is using machine learning/image processing to provide information about specific date types. Customers could use an app on their phones, powered by a machine learning backend, to take pictures of dates in the market. The app would then provide all necessary information about the date type, including grade, price per kilo, calorie content, and unique qualities, etc.

The benefits of automated systems in the fields of farming and agriculture are of paramount importance. The cost of production directly affects the sale price to the end customer. If

979-8-3503-0950-8/24/$31.00 ©2024 IEEE

production costs are reduced, the benefits will reflect on both the farmers and end customers, and ultimately on the country's economy, especially in countries that rely heavily on agriculture as a main source of income. Automated systems would allow for the expansion of farm sizes at minimal cost, as most of the processes would be handled by machines and software. This leads to increased harvest yields at reduced costs. Therefore, these benefits serve as a strong motivation for researchers in the fields of machine learning and image processing to develop techniques with high accuracy that could be commercialized and implemented in practice. The Internet of Things (IoT) and drone technology are already developed technologies that could be seamlessly integrated with high-accuracy image processing techniques.

This research aims to develop an automatic classification system capable of accurately identifying different types of date fruits and providing detailed information about each type. The first phase of this research involves creating a dataset of date fruits by taking high-quality pictures of various dates, particularly from markets in Saudi Arabia. A wide range of date types have been identified in Saudi Arabia, including Safry, Mabroom almadeena, Sokary, Ajwat almadeena, Sagie, Khlas, Mjdwal, Ratab, Lobana, Rotana, Eedyah, Safawy, Khodhry, Rshodya, and Anbar. For a diverse dataset, ten samples of each type from three different shops will be selected. The captured images will be of high quality, with dimensions of 780x1024. Upon completion, to the knowledge of the authors, this will be the first dataset of its kind developed for the various kinds of date fruits grown in Saudi Arabia, representing a significant contribution to this field of research. The development of an automatic classification technique for these date fruit types is a primary goal of this paper. Although the initial dataset will not include diseased fruits, the second phase plans to capture real-time images of various infected date fruits from farms across the country.

II. Literature Review

In [1], the authors developed a method to classify four types of dates (Ajwah, Sagai, Sellaj, Sukkar) using a modified SVM classifier with various visual descriptors. They trained and validated this classifier using 800 images (200 per type), all captured with the same camera and a uniform background to reduce noise and variable factors. The approach yielded a 98.1% classification accuracy, leveraging SVM, texture descriptors, and shape-size features. The high accuracy is not surprising for a small dataset.

In [2], a classification approach for different types and grades of date fruits like Fard, Khalas, and Nagal was introduced, using statistical analysis of 39 RGB features. Two models were used: the first, using LDA, classified dates into three hardness categories (hard, semi-hard, and soft), while the second, using SDA, categorized them into two (hard and soft, with semi-hard included as soft). Validation was performed on a dataset of 3300 images, achieving 68% to 86% accuracy in the first model and 83% to 96% in the second. [3] presented a grading method for dates using the Matlab MFIS tool, focusing on length and freshness to produce a single grade evaluation, with a 91% accuracy rate tested on 500 Mozafati date images. In [4], a vision system for classifying date maturity, detecting defects, and counting dates was proposed, utilizing Matlab's 'regionprops', 'RGB2HSV' functions, and the FLIR tool from Mac OS. The dataset comprised ten images for counting, 20 for maturity and type classification, and thermal images for defect detection. [5] introduced a system for grading date fruits using two backpropagation neural network models based on five features: flabbiness, shape, size, defect, and intensity, with accuracies ranging from 55% to 80%. Lastly, [6] described a method for automatic calorie counting in foods using image recognition with three CNN models of varying layers, each trained on two different custom datasets.

The study in [7] introduced a method to determine the hardness of dates using edge detection with two ANN models. These models used 36 edge detection features, with the first model classifying dates into hard, semi-hard, and soft, and the second model distinguishing only between hard and soft. Trained on 1800 monochrome images (600 per category), the first model achieved 75% accuracy, and the second model 87%. In [8], authors tested two CNN models for automatic fruit classification. The first model had six layers, and the second used a visual geometry group with 16 pre-trained deep learning models. They trained these models on two datasets: one with clear fruit images and another with fruits in challenging conditions. The models achieved 99.49% and 99.75% accuracy on the first dataset, and 85.43% and 96.75% on the second. The study in [9] evaluated date classification on the Zynq-7000 platform using an ANN. This ANN considered 15 features including color and shape, applied to a dataset of 600 dates (100 per class: Khalas, Fardh, Khunaizi, Qash, Naghal, and Maan). The classification accuracy reached 97.26% in real time. In [10], authors proposed an automated system for harvesting dates based on ripening stages, using two models based on AlexNet and VGG-16. This system includes three classifiers for maturity, type, and harvesting decisions, trained on a dataset of 8072 images from 29 palm varieties. The VGG-16 model achieved accuracy rates of 99.01%, 97.25%, and 98.59% for the classifiers, with classification times under 35 milliseconds.

In [11], a comparison of various classification algorithms using texture, size, and color features was conducted on four Saudi Arabian date types. The Support Vector Machine (SVM) showed the best performance with 73.8% accuracy, outdoing Decision Trees (DT) and Random Forest (RF), which ranged between 60%-69%. [12] explored a Smart City system integrating big data, IoT, cloud, fog, edge computing, and 5G for date fruit classification, achieving 99.2% accuracy with a 4000-image dataset. In [13], an intelligent system comprising three subsystems (DTES, DMES, DWES) was proposed for estimating date fruit class, weight, and maturity using deep learning methods like ResNet and VGG-19, with accuracies up to 99.175% for DTES. [14] introduced a smart harvesting system based on date fruit maturity, employing deep learning for detecting seven maturity levels and achieving 99.4% to 99.7% accuracy. [15] developed a method to distinguish healthy from defective date fruits and assess ripeness using a Deep Convolutional Neural Network (CNN) with VGG-16 architecture, reporting 96.98% accuracy. In [16], a deep learning model classified date fruits based on shape, color, and size, with an average accuracy of 97.2% at epoch 4, using a custom dataset and the fruit-360 dataset. [17] released a new dataset for date fruit classification and automatic harvesting, including videos, available on the IEEE DataPort. Finally, [18] proposed an automatic classification system using the Gaussian Mixture Model (GMM) for maturity and shape analysis, showing effective results with a dataset of 5,000 images from ten varieties.

The agricultural industry of date fruits holds significant importance in many countries, leading to ongoing research and development of systems tailored to this sector. Continuous efforts are being made in date fruit agriculture, but there is also a growing interest in other fruits and vegetables globally. We envision a future where comprehensive systems capable of automatic harvesting, classification, ripeness detection, and disease diagnosis are developed for all types of fruits, vegetables, and agricultural produce, creating a universal solution for the agricultural sector.

Research in using machine learning algorithms in the agriculture industry has expanded to include a variety of fruits. These systems and algorithms may be integrated into combined systems for farms with diverse fruit trees, including dates. This existing research is valuable for developing comprehensive mobile applications not limited to date fruits. In [19], a review of machine learning techniques for fruit classification, identification, and grading from 2010-2019 was presented, highlighting limitations and achievements. [20] featured a review on fruit and vegetable grading systems using image processing and various machine learning techniques, proposing a hybrid system for disease and grading. [21] analyzed five multivariate methods for sorting the visual ripeness of Cape gooseberry fruits, with PCA showing the highest accuracy. [22] introduced a tomato grading system based on RGB images, achieving 95.15% accuracy in defect detection and suggesting its use in automatic sorting. In [23], an intelligent system for classifying four fruit types and their quality was proposed, using 30 features and achieving a maximum accuracy of 98.48% with SVM. Machine learning algorithms are increasingly being integrated with technologies like drone technology, IoT, and cloud computing. This integration aims to develop practical systems for use in the agriculture industry. References [24-26] in previous literature have proposed such innovative systems [24-26].

III. METHODOLOGY

This paper presents the development of an innovative system for the recognition of date fruits, accessible through a smartphone application. The system is designed to provide users with extensive information about date fruits, including health benefits, geographic origins, and caloric content, as illustrated in Figure 1. The foundation of this research is the 'Date Fruits Dataset', a comprehensive collection of various date fruit varieties, each represented as a distinct class within the dataset. This collection includes, but is not limited to, varieties such as Ajwa, Anbara, and Deglet Noor, ensuring a wide coverage for accurate fruit recognition.

Prior to model training, preprocessing of the dataset images was undertaken. This involved the removal of background to focus on the fruit, application of image enhancement techniques for improved clarity, and resizing of images to standardize the input for the model. The core of the recognition system is a Convolutional Neural Network (CNN) employing the VGG16 architecture, chosen for its proven effectiveness in image classification tasks. To enhance the model's performance and its ability to generalize from training data, data augmentation techniques were applied during the training process. For evaluation, the model was tested using a separate dataset. Performance metrics such as accuracy, precision, recall, and the F1 score were used to assess the model's effectiveness.

A significant aspect of the project was the integration of user interface design with the technical machine learning processes. This ensured a seamless user experience, enabling users to effortlessly capture the image of the fruit and obtain detailed information about it. The use of advanced image processing and deep learning techniques is central to the system, providing accurate and informative results to users.

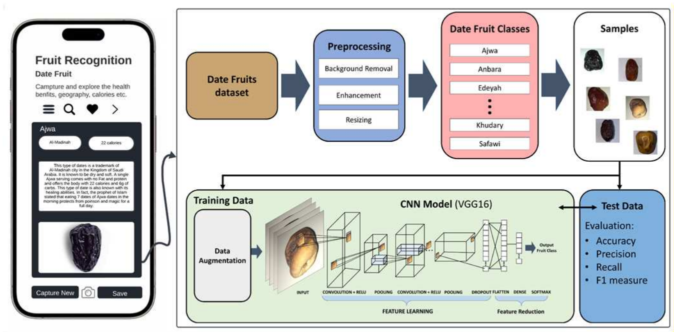

Fig. 1. Proposed method framework and mobile Application Interface

A. *Experimental Dataset*

This research work considers a new dataset we built ourselves. It consists of a collection of images that include 15 types of the most famous dates. Each type is considered as a "label" that the proposed classification model will attempt to predict. These images were taken using smart phones equipped with digital camera with a resolution of 12 MP. Our dataset comprises a set of 1350 images equally partitioned among the 15 types of dates (90 images per class). The dates fruit were collected from different sources. The images were captured in various environment and postures settings such as brightness, saturation, dates angles, and contrast. Table 1 shows a detailed description about each class of dates, the assigned label, and an image sample of each type. During data collection process, several challenges were faced such as dates fruit used to have distinct visual characteristics (size, shape, color) at different ripening stages, and capturing their images with lighting and maturity stage can also effect. Furthermore, accurate labelling requires domain knowledge about date cultivars, ripening stages, and diseased dates fruit.

B. *CNN based Transfer Learning*

Typically, CNN models that were trained on top of common machine learning classifiers like Support Vector Machines and Random Forests are used to extract features in transfer learning. The CNN models are improved by doing network surgery or fine-tuning in the alternative transfer learning technique. Existing CNN models can be adjusted using a variety of techniques, such as redesigning the architecture, retraining the model, or freezing particular layers of the model to make use of pre-trained weights.

Table 1. Summary of the recent studies on 6G development

#	Date Name	Sample	Description & Benefits
1	Ajwa / عجوة		Ajwa dates, from Al-Madinah, Saudi Arabia, are dry, soft, and nutritious, offering 22 calories and 6g of carbs per serving with no fat or protein. They are also revered for their healing properties. In fact, the prophet of Islam stated that eating 7 dates of Ajwa dates in the morning protects from poinson and magic for a full day.
2	Anbara / Anbarah / عنبرة		Anbara dates, larger than Ajwa and favored for their fleshy texture, big size, and small seed, are a brown, very sweet variety. Like other dates, they are free from cholesterol, fat, and protein, with one serving providing 66 calories. Cultivated primarily in Madinah, Saudi Arabia, Anbara dates are among the most expensive types available.
3	Edeyah /		These dates are heavy, sweet, and come in various sizes, ranging from medium to large. They start as a reddish color and turn amber when ripe. They bear resemblance to Bahri dates and are typically cultivated during the late season.
4	Ekhlas / Kholas / خلاص		Kholas dates, originating in Saudi Arabia, have a butter caramel flavor when cured. They are sticky when fresh and commonly consumed with Arabic coffee. These dates are of average size, with dark brown and reddish skin. They provide energy, are low in calories and fat, high in fiber, and rich in minerals.
5	Khudary / Khudari / خضري		These dates, one of the most common varieties, come in various sizes, from small to large, and have a dark red or brown color. They are dry, sweet, with chewy flesh and flaky, wrinkled skin. Known for their reasonable price and freshness, they are commonly exported and found in shops and hotel reception desks. They are a great energy source, with one date providing 20 calories.
6	Lubana / لبانة		This type of dates is known to be very nutritious. It comes with high quantity of fiber and disease fighting antioxidants. It is also considered as an excellent natural sweetner and is recommended to be added to somoone's diet. It is also recommeded to stimulate natural labor for pregnant ladies. This date is known to be dry and with dark yellow and colour and wrinkled skin.
7	Mabroom / مبروم		This type of dates is among the most delicious dates. It has firmer, slender, and elongated body compared to other dates. Clolour wise, the Mabroom dates are dark brown. Its outer soft skin membrane comes with spacious wrinkles providing nices looks. Compared to other type of dates, the mabroom dates are less sugary. They are also knows to be very nutritious in terms of vitamins, minerals, antioxidants, and fiber. A single date serving offers 22 claories.
8	Medjool / مدجول		Medjoul dates, often referred to as the "Queen of Dates," are distinguished by their large size, reaching up to 3 inches in length. They have a cushiony texture and a strong, sweet flavor. Popular in the United States, they are frequently added to shakes, smoothies, and energy bars. As they ripen, Medjoul dates change in color from amber to reddish-brown. One Medjoul date contains 66 calories, 18 grams of carbs, and 16 grams of sugar.
9	Ratab		Ratab dates, highly renowned in the Kingdom of Saudi Arabia, are typically small in size and available at the beginning of summer for a short duration of a few weeks. These dates are juicy, soft, and have thin papery skin. They are characterized by their dark reddish skin and yellowish side. Ratab dates are rich in nutrition, including vitamins and fiber.
10	Rotana		The Rotana dates are also a characteristic of Saudi Arabia dates. This type of dates is of a small size. It comes with 2 colours yellowish and light brown. They are usually better preserved in freezer when unripe to preserve their firm and hard skin. Such skin becaomes softer and stickier with time. This type of dates is also knwoen to be very rich in terms of fibers and other nutritions.
11	Rashudiah / رشوديه		The Rashudiah date is a pitted fruit. It is characterized with a non soft skin. A date of this kind shows colours light yellow and light brown. Its skin has wrinkles. The Rashudiah is relatively elongated date. Nutrition wise, 1 date serving offers 25 calories and 6 grams of sugar. Like all dates, it is also considered as a good source of fiber and energy.
12	Safawi / صفاوي		Safawi dates, cultivated mainly in Madinah, Saudi Arabia, are sweet, black, and fleshy. They are rich in various vitamins, tender, and soft. Safawi dates are believed to have curative properties for several diseases and provide balanced nutrition and minerals. It is recommended to consume them on an empty stomach for better effectiveness, as they can help eliminate stomach worms and bacteria. One Safawi date contains 15 calories.
13	Suqa'ey / Saqy / صقعي		Crisp dates, primarily cultivated in Riyadh, Saudi Arabia, are brown with a distinctive light yellow tip. They are wrinkled but not flaky and have an elongated and slightly dense shape. These dates are known for their benefits, including aiding in the recovery of intestinal disorders, supporting heart health, quick recovery from intoxication, and relieving constipation. A single serving of these dates provides 34 calories.
14	Suffry / Zefari / صفري		Suffry dates, found in the Al-Madinah region of Saudi Arabia, are a long variety with a dark brown color. They are soft, sweet, and flavorful. A single date serving provides 31 calories, with no fat or saturated fat, and they are a good source of fiber.
15	Sukary / Sukkari / سكري		Sukkary dates, with a golden color, originate from the Al-Qassim region in Saudi Arabia. They are sweet, high-quality, and crispy, deriving their name from the Arabic word for sugar, "sukkar." Sukkary dates are renowned for their healing and nutritious properties, including lowering cholesterol levels, providing energy, and aiding in preventing tooth decay. When dried, they have a crystallized sugar layer on the outside and a soft, caramel-like flavor.

A CNN-based architecture called VGG16 was suggested for the classification of substantial amounts of visual data. Tiny convolution filters are used in these designs to deepen the network. These networks receive fixed-size, 224 by 224 pictures with three color channels as inputs. A sequence of convolutional layers with max pool layers and tiny receptive fields (3 x 3) receive the input. Conv3-64 and conv3-128, respectively, with a ReLU activation function are used in the first two sets of VGG. Conv3-256, conv3-512, and conv3-512, respectively, are used in the final three sets with a ReLU activation function. Every set of convolutional layers in VGG16 is succeeded by a max-pooling layer with window size of 2x2 and stride of 2. It is noticeable that the convolutional layers' channel counts range from 64 to 512.

LeNet is extended by AlexNet, which has a far more complex architecture. Five convolution layers and three completely connected layers make up its total of eight layers. Every layer is linked to an activation function of ReLU. AlexNet employs drop-out and data augmentation approaches to prevent overfitting issues that may result from using too many parameters. DenseNet can be compared to an extension of ResNet in which a layer above adds its output to a subsequent layer. Concatenation of the outputs of earlier layers with those of later layers was suggested by DenseNet. Concatenation increases efficiency by improving the distinction in the input of subsequent layers.

Inception modules, the foundation of GoogleNet architecture, feature convolution processes with various filter sizes operating at the same level. This essentially broadens the network's width as well. The architecture has nine stacked inception modules and 27 layers (22 layers with parameters). A fully connected layer with the SoftMax loss function serves as the classifier for the four classes at the conclusion of the inception modules.

To train the aforementioned models from scratch, computing power and data resources are needed. Using transfer learning in one experimental context and applying it to other settings that are similar is perhaps a superior strategy. Transferring all learned weights as it may not work effectively in the new situation. Therefore, it is preferable to freeze the first few levels and use random initializations for the latter layers. To learn the additional data classes, this partially modified model is retrained using the available dataset. The amount of processing capacity and accessible dataset determines how many layers are fine-tuned or frozen. We can unfreeze more layers and adjust them for the particular issue if there is enough data and processing power available.

We employed two levels of fine-tuning in this study: (1) freezing all feature extraction layers and unfreezing the fully connected layers, where classification choices are determined; and (2) freezing beginning feature extraction layers and unfreezing the latter feature extraction and completely linked layers. The latter should yield better results but requires more training time and data. In Case 2, only the first 10 layers of VGG16 were frozen, while the remaining layers were retrained for fine-tuning.

IV. RESULTS AND DISCUSSIONS

The data presented in Tables 2 and 3, as well as the confusion matrix in Figure 2, offer detailed insights into the performance of various Convolutional Neural Network (CNN) models in the task of classifying date fruits. These models employ transfer learning techniques, utilizing both non-frozen and frozen weight configurations.

Table 2 shows the results of the experiments of three different CNN models based Transfer Learning with Non-Freeze weights (AlexNet, GoogleNet, VGG16). The results demonstrate promising outcomes. Notably, VGG16 exhibits superior performance metrics, achieving a validation accuracy of 97.41%, overall accuracy of 97.04%, precision of 97.13%, recall of 97.28%, and an F1 score of 97.78%. This table indicate VGG16's remarkable adaptability to the date fruit classification task, attributable perhaps to its deeper architecture and enhanced capability for learning complex features, especially when all weights are adjustable during training.

Table 2. Experimental Results of different CNN models based Transfer Learning with Non-Freeze weights

	Validation Accuracy	Accuracy	Precision	Recall	F1 Measure
AlexNet	97.04%	96.67%	95.93%	96.30%	95.93%
GoogleNet	94.81%	94.44%	94.81%	94.44%	94.93%
VGG16	97.41%	97.04%	97.13%	97.28%	97.78%

Conversely, the frozen weights methodology outlined in Table. 3 also positions VGG16 as the leading model, albeit with marginally improved metrics compared to the non-frozen approach. It shows a validation accuracy of 98.19%, accuracy of 98.04%, precision of 97.82%, recall of 98.03%, and F1 score of 98.10%. This improvement suggests that freezing the initial layers of VGG16, which likely encapsulate more general features acquired from pre-training datasets, and fine-tuning the latter layers for this specific task, enhances generalization and slightly boosts performance across all metrics.

Table 3. Experimental Results of different CNN models based Transfer Learning with Freeze weights

	Validation Accuracy	Accuracy	Precision	Recall	F1 Measure
AlexNet	96.30%	95.56%	95.48%	96.27%	95.61%
GoogleNet	92.22%	91.11%	92.17%	91.34%	91.11%
VGG16	98.19%	98.04%	97.82%	98.03%	98.10%

The confusion matrix for the VGG16 model presented in Figure 2 underscores this high degree of classification accuracy. It reveals that most date fruit types are correctly identified with high precision, often reaching 100%, and instances of misclassification are rare. This graphical representation corroborates the numerical findings, emphasizing VGG16's robustness in this context, particularly when utilizing the frozen weights technique. The minimal misclassifications, indicated by the non-zero off-diagonal elements in the confusion matrix, could potentially be further minimized with additional strategies such as increased training data, data augmentation, or further hyperparameter optimization. The higher accuracies were achieved using VGG16 model due to the fact that the model extracts both low-level and high-level features from the dates fruit images. In the early convolutional layers, VGG16 focuses on extracting basic features like edges, lines, and corners. As the network progresses through deeper layers, it starts to combine the low-level features to identify color combinations, textures, and basic shapes that are relevant to the image content.

The experiments affirm the efficacy of transfer learning in fruit classification tasks. The deep-structured VGG16 model, when fine-tuned with frozen weights, demonstrates exceptional validation accuracy and excellent scores in precision, recall, and F1 measure. These results have significant implications for the development of automated systems in agricultural contexts, where rapid and accurate fruit classification is crucial. The generalizability of CNN model's findings for date fruit classification to other regions or fruit types depends on several factors related to variations based on the fruit characteristics (shape, color, and size). While the trained CNN model for date classification might not directly translate to other fruit types, however, the convolutional layers can still learn low-level and mid-level features relevant to fruit images in general and the higher-level features specific to dates would need to be re-learned with a new dataset for a different fruit type.

V. Conclusion

This research proposes an automatic classification model which is capable of accurately identifying different types of date fruits and providing detailed information about each type. We have created a large dataset of 15 types of date fruits (Safry, Mabroom almadeena, Sokary, Ajwat almadeena, Sagie, Khlas, Mjdwal, Ratab, Lobana, Rotana, Eedyah, Safawy, Khodhry, Rshodya, and Anbar) by taking high-quality pictures, particularly from markets in Saudi Arabia. We have considered three different CNN models based on Transfer Learning were considered (AlexNet, GoogleNet, VGG16) along with two levels of fine-tuning: (1) freezing all feature extraction layers and unfreezing the fully connected layers; and (2) freezing beginning feature extraction layers and unfreezing the latter feature extraction and completely linked layers. In both scenarios, VGG16 model outperformed the other 2 models. Best results was achieved when the model was fine-tuned with frozen weights. In our future work, we propose increasing the scope of our dataset by considering diseased fruits. We plan in capturing a large set of real-time images of various infected date fruits from farms across the country and assess the performance of the proposed classification models.

	Ajwa	Anbara	Edeyah	Ekhlas	Khudary	Lubana	Mabroom	Medjool	Ratab	Rotana	Rashudiah	Safawi	Saqy	Suffry	Sukary
Ajwa	94%	0%	0%	6%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Anbara	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Edeyah	0%	0%	88%	0%	0%	6%	0%	0%	0%	0%	0%	6%	0%	0%	0%
Ekhlas	6%	0%	0%	82%	0%	6%	0%	0%	0%	0%	6%	0%	0%	0%	0%
Khudary	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Lubana	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Mabroom	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%
Medjool	0%	0%	0%	0%	0%	0%	0%	94%	0%	0%	6%	0%	0%	0%	0%
Ratab	0%	0%	5%	0%	0%	0%	0%	0%	85%	0%	0%	5%	0%	5%	0%
Rotana	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%
Rashudiah	6%	0%	0%	0%	0%	0%	0%	0%	0%	0%	94%	0%	0%	0%	0%
Safawi	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%
Saqy	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	94%	6%	0%
Suffry	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%
Sukary	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%

Fig. 2. Confusion Matrix of the for different types of date fruits classification using VGG16 based Transfer Learning

References

[1] Muhammad, G. (2015). Date fruits classification using texture descriptors and shape-size features. *Engineering Applications of Artificial Intelligence*, *37*, 361-367.

[2] Manickavasagan, A., Al-Mezeini, N. K., & Al-Shekaili, H. N. (2014). RGB color imaging technique for grading of dates. *Scientia Horticulturae*, *175*, 87-94.

[3] Alavi, N. (2013). Quality determination of Mozafati dates using Mamdani fuzzy inference system. *Journal of the Saudi society of agricultural sciences*, *12*(2), 137-142.

[4] Najeeb, T., & Safar, M. (2018, March). Dates maturity status and classification using image processing. In *2018 International Conference on Computing Sciences and Engineering (ICCSE)* (pp. 1-6). IEEE.

[5] Al Ohali, Y. (2011). Computer vision based date fruit grading system: Design and implementation. *Journal of King Saud University-Computer and Information Sciences*, *23*(1), 29-36.

[6] Latif, G., Alsalem, B., Mubarky, W., Mohammad, N., & Alghazo, J. (2020, April). Automatic Fruits Calories Estimation through Convolutional Neural Networks. In *Proceedings of the 2020 6th International Conference on Computer and Technology Applications* (pp. 17-21).

[7] Manickavasagan, A., Al-Shekaili, H. N., Thomas, G., Rahman, M. S., Guizani, N., & Jayas, D. S. (2014). Edge detection features to evaluate hardness of dates using monochrome images. *Food and bioprocess technology*, *7*(8), 2251-2258.

[8] Hossain, M. S., Al-Hammadi, M., & Muhammad, G. (2018). Automatic fruit classification using deep learning for industrial applications. *IEEE Transactions on Industrial Informatics*, *15*(2), 1027-1034.

[9] Ammari, A. C., Khriji, L., & Awadalla, M. (2020, April). HW/SW Co-‘esign for Dates Classification on Xilinx Zynq SoC. In *2020 26th Conference of Open Innovations Association (FRUCT)* (pp. 10-15). IEEE.

[10] Altaheri, H., Alsulaiman, M., & Muhammad, G. (2019). Date fruit classification for robotic harvesting in a natural environment using deep learning. IEEE Access, 7, 117115-117133.

[11] Abi Sen, A. A., Bahbouh, N. M., Alkhodre, A. B., Aldhawi, A. M., Aldham, F. A., & Aljabri, M. I. (2020, March). A Classification Algorithm for Date Fruits. In 2020 7th International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 235-239). IEEE.

[12] Hossain, M. S., Muhammad, G., & Amin, S. U. (2018). Improving consumer satisfaction in smart cities using edge computing and caching: A case study of date fruits classification. Future Generation Computer Systems, 88, 333-341.

[13] Faisal, M., Albogamy, F., Elgibreen, H., Algabri, M., & Alqershi, F. A. (2020). Deep Learning and Computer Vision for Estimating Date Fruits Type, Maturity Level, and Weight. IEEE Access, 8, 206770-206782.

[14] Faisal, M., Alsulaiman, M., Arafah, M., & Mekhtiche, M. A. (2020). IHDS: Intelligent harvesting decision system for date fruit based on maturity stage using deep learning and computer vision. IEEE Access, 8, 167985-167997.

[15] Nasiri, A., Taheri-Garavand, A., & Zhang, Y. D. (2019). Image-based deep learning automated sorting of date fruit. Postharvest biology and technology, 153, 133-141.

[16] Magsi, A., Mahar, J. A., & Danwar, S. H. (2019). Date fruit recognition using feature extraction techniques and deep convolutional neural network. Indian Journal of Science and Technology, 12(32), 1-12.

[17] Altaheri, H., Alsulaiman, M., Muhammad, G., Amin, S. U., Bencherif, M., & Mekhtiche, M. (2019). Date fruit dataset for intelligent harvesting. Data in brief, 26, 104514.

[18] Aiadi, O., & Kherfi, M. L. (2017). A new method for automatic date fruit classification. International Journal of Computational Vision and Robotics, 7(6), 692-711.

[19] Behera, S. K., Rath, A. K., Mahapatra, A., & Sethy, P. K. (2020). Identification, classification & grading of fruits using machine learning & computer intelligence: a review. Journal of Ambient Intelligence and Humanized Computing, 1-11.

[20] Kumar, M. P., & Parkavi, A. (2020). Quality grading of the fruits and vegetables using image processing techniques and machine learning: a review. Advances in Communication Systems and Networks, 477-486.

[21] De-la-Torre, M., Zatarain, O., Avila-George, H., Muñoz, M., Oblitas, J., Lozada, R., ... & Castro, W. (2019). Multivariate analysis and machine learning for ripeness classification of cape gooseberry fruits. Processes, 7(12), 928.

[22] Ireri, D., Belal, E., Okinda, C., Makange, N., & Ji, C. (2019). A computer vision system for defect discrimination and grading in tomatoes using machine learning and image processing. Artificial Intelligence in Agriculture, 2, 28-37.

[23] Bhargava, A., & Bansal, A. (2020). Automatic Detection and Grading of Multiple Fruits by Machine Learning. Food Analytical Methods, 13(3), 751-761.

[24] Panday, U. S., Shrestha, N., Maharjan, S., Pratihast, A. K., Shrestha, K. L., & Aryal, J. (2020). Correlating the plant height of wheat with above-ground biomass and crop yield using drone imagery and crop surface model, a case study from Nepal. Drones, 4(3), 28.

[25] James, K., & Bradshaw, K. (2020). Detecting plant species in the field with deep learning and drone technology. Methods in Ecology and Evolution, 11(11), 1509-1519.

[26] Latif, G., Alghazo, J., Maheswar, R., Vijayakumar, V., & Butt, M. Deep learning based intelligence cognitive vision drone for automatic plant diseases identification and spraying. Journal of Intelligent & Fuzzy Systems, (Preprint), 1-12.