EasyChair Preprint

№ 6323

An Intelligent Approach to Identify the Date Palm Varieties Using Leaves and Fruits

Neji Ines, Hamza Hammadi and Ejbali Ridha

August 17, 2021

An intelligent approach to identify the date palm varieties using leaves and fruits

Ines NEJI [1,*], Hammadi HAMZA[2], and Ridha EJBALI[3]

[1] *Faculty of Sciences of Gabes, University of Gabes, Tunisia*
[2] *Arid and Oases Cropping Laboratory, Arid Area Institute, Medenine 4119, Tunisia*
[3] *Research Team in Intelligent Machines, Engineering School of Gabes, University of Gabes*
nejiines1997@gmail.com, hamzapalmier@yahoo.fr, ridha_ ejbali@ieee.org

Abstract

Tunisian people lack the knowledge to determine the type of date fruit, many consumers struggle when they go to buy fruit every day, even novice farmers find it difficult to know the type of dates they produce. It is very difficult for the human eye to classify similar things such as date fruit. Thus, it is necessary to develop an accurate solution that can accomplish this task. The goal of this research is to prove the feasibility and test the capability of our solution, we focus on identifying the varieties of dates fruits not only using dates images but also palm leaves images. We implement three CNN models for date fruit and leaf classification which the first classifies fruit and leaf image (binary classification), the second classify fruit varieties, and the third classify the leaves varieties. We have come up with good accuracy 99.97% for the binary model, 99.82% for the fruit model, and 99.73% for the leaf model.

keywords : *Phoenix dactylifera* L, Date palm leaves, Date palm fruits, Classification, Convolution neuron network.

1 Introduction

The date palm (*Phoenix dactylifera L.*) is one of the oldest fruit crops cultivated in the Middle East and North Africa. The fruits of the date palm and by products have contributed to the food security and livelihoods of population in the region for over 5,000 years. The crop has been linked to ancient civilizations, including Sumerians, Akkadians and Babylonians, and is mentioned in Islamic, Jewish and Christian holy texts[1].

Currently, more than 100 varieties of date palm exist in Tunisia. Many fresh fruits of the date palm are available for 8 months of the year. The date palm production is constantly evolved. In fact, 242 thousand tons of dates have been resulted in 2017. In 2021, the producing of dates in Tunisia was forecast at 345 thousand tons [2].

The date palm tree is considered a slow-growing species, it takes four to eight years before they produce fruit. In addition, and for lack of knowledge, many farmers can't know the varieties growing in the oasis. At the end of the production chain, dates should be sorted based on type and degree of freshness so that they can be packaged and priced. Moreover, sorting through large amounts of dates is very time-consuming and requires a lot of experienced and knowledgeable labor. It is difficult to tell the different types apart, as there are many dates types that have similar properties and are slightly different in color, shape, and texture.

Currently, there are not many date fruit and date leaf data sets available to work with [3]. This work contributes to the computer vision and agricultural technology fields by providing a new

image data-set in the area and exploring machine learning for the classification of these images. This then connects the gap between newer technologies and date farms to introduce the smart farming concept in date farms, smart farming is an emerging concept that refers to managing farms using technologies like IoT, robotics, drones, AI to increase the quality and quantity of products while optimizing the human labor required by production. The work can be built on to create a more extended date fruit and leaf classification app which might be useful for consumers and owner to check the labeling of products

We organize this paper as follows: Section 2 takes an overview of deep learning, neuron network, and convolution neuron network. Section 3 gives the related work to date palm classification subject. Section 4 gives a description of our methodology. Section 5 covers the result and test of our proposed solution. And section 6 gives the conclusion and our future work.

2 Background

In this section, we gave an overview of artificial neuron networks, deep learning, and convolution neuron network (CNN) architecture.

2.1 Artificial Neuron Network

Artificial Neural Networks (ANN) are simplified models of the biological nervous system and therefore have drawn their motivation from the kind of computing performed by a human brain. The ANN is an approach addressing purely human functions such as perception, decision, learning and reasoning.

2.2 Deep Learning

Deep learning is a kind of an automatic learning algorithm and it is a new area of machine learning research. Deep learning algorithms use a cascade of several layers to create knowledge. The input layer receives the output layer from the previous layer as input. Generally, the architecture of deep learning is formed not only by several layers of the input and the input but also by parameterized non linear modules. The parameters are the subject of learning. Each layer allows a higher level representation than the previous one.

2.3 Convolutional Neural Network Architecture

In the field of Deep learning, Deep Convolutional Neural Network (CNN) is a particular variety of Neural Networks. It is one of the best learning algorithms for understanding image content and has presented exemplary performance in image segmentation, classification [4], recognition [5], detection, and retrieval related tasks. The main benefit of CNN compared to its antecedents is that it automatically identifies the relevant features without any human supervision.The design of CNNs was stimulated by neurons in human and animal brains, much the same to a conventional neural network[6].

The CNN architecture consists of a number of layers, each layer in the CNN architecture, including its function

- **Convolutional Layer** Consists of a store of convolutional filters (called kernels). The input image, expressed as N-dimensional metrics, is convolved with these filters to generate the output feature map.
- **Pooling Layer** The main task of the pooling layer is the sub-sampling of the feature maps. These maps are generated by following the convolutional operations.

- **Activation Function** (non-linearity) Mapping the input to the output is the core function of all types of activation function in all types of neural network.
- **Fully Connected Layer** Commonly, this layer is located at the end of each CNN architecture.

3 Related work

Recent research papers and journals were studied to implement a modern and better approach to the problem of classification in agriculture and commercial fields at hand while maximizing the accuracy of the deep learning model. A recent work in fruit classification, aiming to the problem of great similarities between some apple varieties and pears and peaches. Wu et al.[7] proposes a method of fruit automatic recognition and classification based on convolutional neural network. Moreover, Rajesh et al. [8] have used CNN to classify four different fruits, and have achieved 90% accuracy with a data-set of 200 images.

Furthermore, in the some goal of our paper, Alhamdan et al.[9] explores the use of CNN in classifying images of date fruits as one of 9 varieties, create several models with the highest achieving 97% accuracy. It contributes an original dataset of 1658 images taken in a controlled environment. In addition, Sakibul Hasan et al.[10] survey various pre-trained CNN model to classify 8 date fruit varieties by capturing the image from the market using mobile phone and attend 2246 images in their data-set to achieving 82.67% accuracy .

In [11], convolution neuron network are used to classify five levels of growth of dates, providing a framework for a robot in an orchard environment that can quickly decide if they were ready for harvest or not. This paper used more than 8000 images, however, the images were not suitable for single date classification in a controlled environment.

Another related paper [12] collected over 1300 images of 4 different growth stages. It used a CNN model to classify the 4 stages, including a defective stage. However, the growth stages were different enough in colour alone to tell them apart, unlike in this work where many date types are very close in colour or shape. The images in both works were not in a controlled environment, in terms of date position, camera focus, angle of capture, lighting conditions, and camera distance. In [13] the author propose a method for automatic classification of date fruits based on computer vision and pattern recognition. This method was implemented, and empirically tested on an image data spanning seven different categories of dates.

In the previous works, *Phoenix dactylifera* varieties are classified using the date fruit image, these solutions can benefit the consumer more than the owner of the feather palm farm. Our proposed solution aims to prevent both the consumer and the owner, with the use of the date fruit and date palm leaf to predict the date varieties.

4 Methodology

The prime aim of our research work is to build a model that can predict the dates varieties from an input image. Before predicting the variety our CNN model classifies the input image as date fruit or date leaf image (binary classification). In addition, from the produced result, the date variety prediction is done. Our proposed methodology workflow is depicted in figure 1.

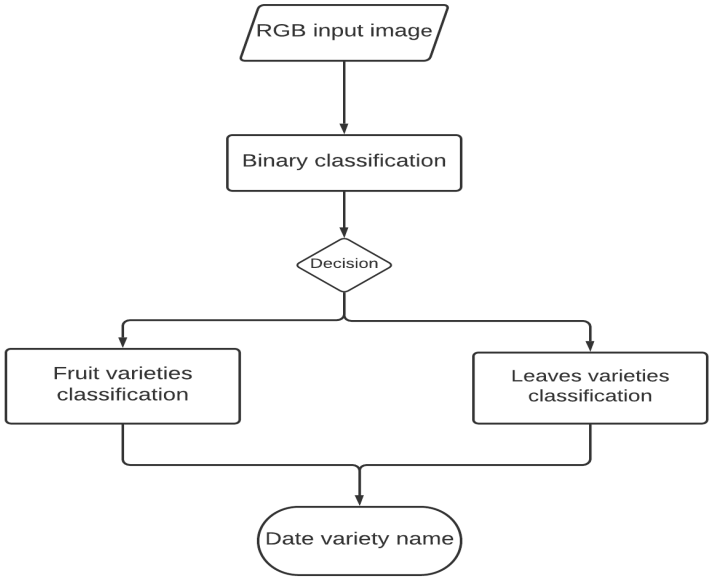

Figure 1: Proposed methodology

In our proposed methodology we have implemented three different CNN models, a Binary model to classify the input image as fruit or leaf, a fruit model to classify fruit varieties, and a leaf model to classify leaves. Each model has been trained using a specific data-set. Figure 2 present the different steps of our proposed solution.

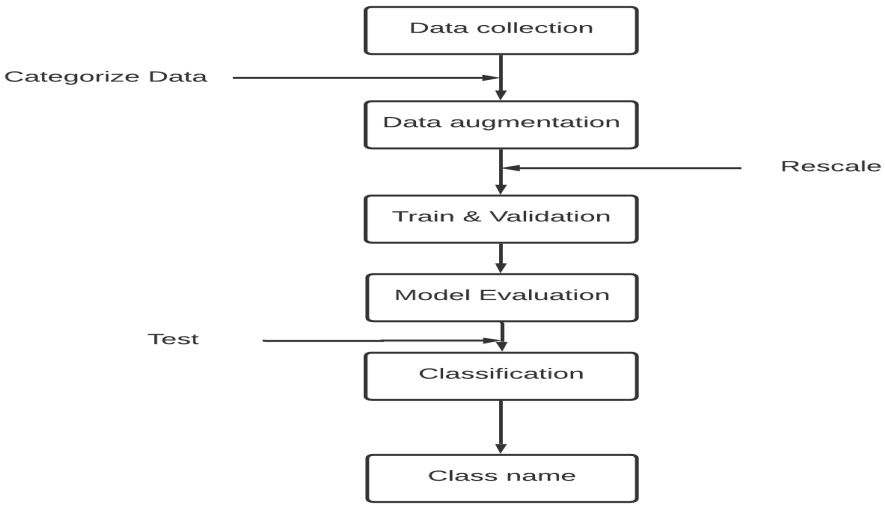

Figure 2: General workflow

4.1 Data-set creation

Five different types of dates were considered: Zekri, Alig, DegletNour, Khwat Alig, and Kenta. The fruit image present a lab data that it was taken in a laboratory environment (Arid Regions Institute Laboratory) with a smartphone camera while the leaves images were taken in a real environment where the leaf image present a field data with no controlled background. A small dataset was constructed, one image of each fruit date type and three of each leaf's date type.

4.2 Data augmentation

Data starvation cannot be accomplished may raise learning issues and alter the accuracy of the model. The small raw data-set provided by the Arid Regions Institute cannot fulfil our aims also may lead to data-starvation making the model not learning enough resulting in poor prediction results. To solve this problem we use data augmentation techniques to increase the diversity of our training set by applying random geometric transformation [14]. For each class of date fruit and palm leaf we use the geometric transformation given below:

- **Rotation** rotation augmentations where done by rotating the image right or left on an axis 90 °.
- **Translation** shifting images left, right, can be a very useful transformation to avoid positional bias in the data.
- **Flipping** horizontal axis flipping is much more common than flipping the vertical axis but in our work we used both vertical and horizontal.
- **Shear** create or rectify the perception angles in input image.
- **Zoom** randomly zoom the image and adds new pixels for the image.

4.3 Model development

The work was coded using Colaboratory, or "Colab" for short, which uses TensorFlow 2.7 by default us open-source library for a number of various tasks in machine learning with python 3 and Keras as neural network library and minor use of other libraries. We implemented three CNN models which one binary CNN model and two other multi-class models. The table 1 gives more details for each CNN models.

Models name	Model type	Model description
FruitORLeaf model	Binary model	This model contains 2 **Conv2D** layers with 32 and 16 respectively as numbers of filters, 3 as kernel size (224,224,3) as input shape, and use '**Relu**' as activation function followed by **Max-pooling** layer.This model uses '**sigmoid**' as classifier activation function, '**Adam**' as an optimizer, and '**binary_crossentropy**' as loss function.
Fruit model & Leaf model	Multi-class model	These 2 models contain 4 **Conv2D** layers with 16,32,64 and 128 respectively as filters numbers, 3 as kernel size (224,224,3) as input shape, and use '**Relu**' as activation function followed by **Max-pooling** layer.These model use '**Relu**' as classifier activation function, '**Adam**' as an optimizer, and '**categorical_crossentropy**' as loss function.

Table 1: CNN models details

5 Result and test

In this section, we discussed the obtained data-set after preparation and we showed our models performances using the accuracy and confusion matrix plots.

5.1 Data-set

Three different data-sets were created; the first was based on the images of both dates and leaves, the second contains the fruit images of the most 5 popular dates in Tunisia and the third one contains the images of the leaves of these 5 aforementioned fruit types. It is worth noting, the high similarity between the leaf and fruit types has complicated the classification extremely hard. A total of 300 images if high quality pictures with was collected for each type. Figure 3 and figure 4 showed 5 samples of date palm leaves and date palm fruits included in our data-set, respectively. For our analysis the image size was fixed on 224 * 224 in both train and validation.

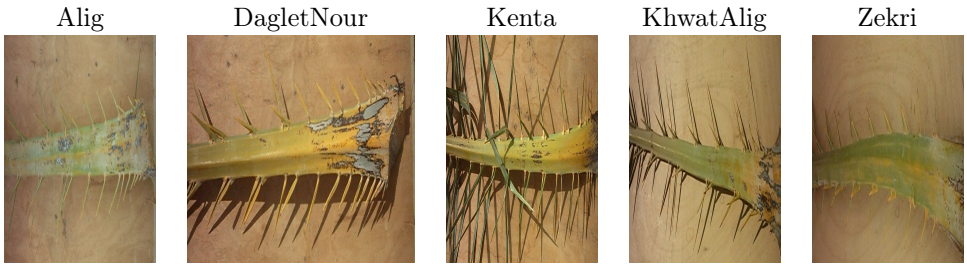

Figure 3: Samples from leaves dataset, resized and cropped[15]

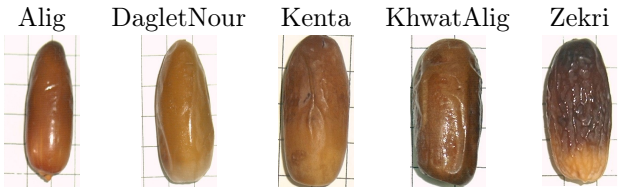

Figure 4: Samples from fruit dataset, resized and cropped[15]

5.2 Model performance

Our results revealed that the three models used in the analysis were highly effective, with an accuracy of 99.97%, 99.82%, 99.73% for the binary model, fruit model, and leaf model,respectively. The plots depicting the training and validation accuracy of leaf and fruit models were shown in figure 5.

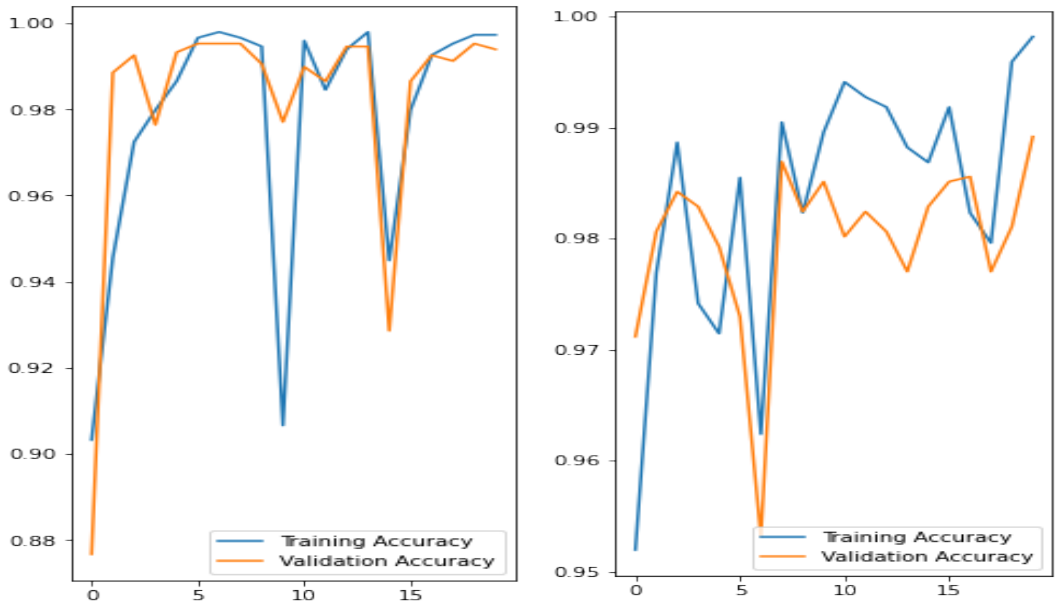

Figure 5: Train and validation accuracy for leaf(left) and fruit(right) classifier

A confusion matrix, known as an error matrix is a specific table layout that allows the visualization of the performance of model/classifier. According to the confusion matrix in figure 6, revealed in our results, the identification of the date variety on the basis of fruit image was higher than 95%, which in agreement with diverse previous studies [9] and [11]. However, our results showed also that the accuracy of identification on the basis of leaf images, although lower than the fruit images, was very high, reaching almost 90% for Zekri variety for example.

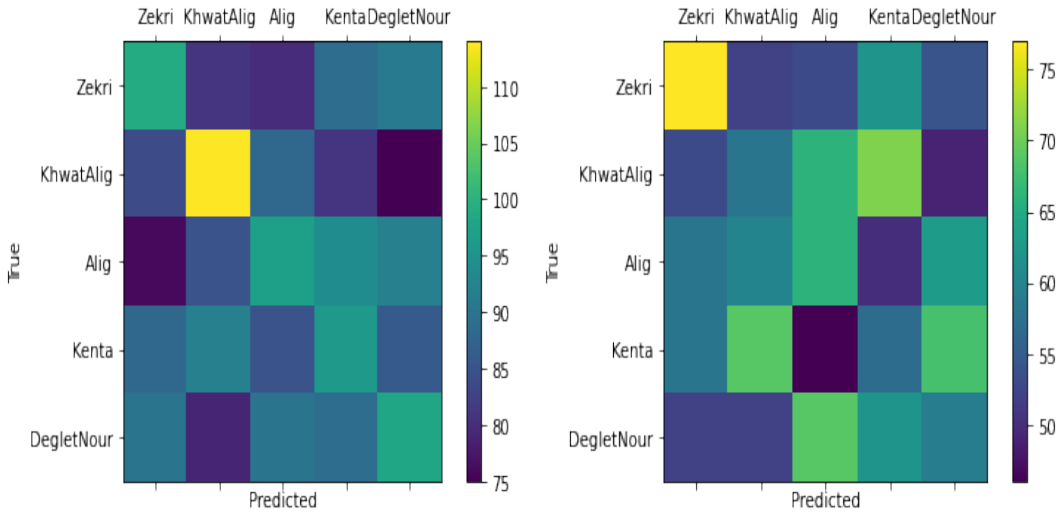

Figure 6: Confusion matrix for fruit(left) and leaf (right) model

This result indicated that the identification of the variety on the basis of the leaf images was very effective and it could be used by farmer at early stage of date growth.

6 Conclusion

This research aimed to identify the main distinguishing features of 5 date fruit varieties grown in Tunisia. We succeeded in making 3 classifiers, a binary classifier for fruit or leaf, multi-class classifiers for fruit and leaf. The fruit and leaf classifier were use the same CNN model but with different data-set. The 3 data-sets are important because there are not any date fruit and date leaves data-sets.

As perspective, we will deploy this project handling Docker container, is an open platform for developing and running applications, and it enable to separate the application from infrastructure. In addition, we will ameliorate our project by adding other functions like the ability to predict the date crops quantity and quality after identify date fruit varieties.

Acknowledgement

The authors would like to acknowledge the financial support of this work by grants from General Direction of Scientific Research (DGRST), Tunisia, under the ARUB program.

References

[1] Nazir Hussain. Date palm global development plan, 09 2010.

[2] Food and agriculture organization of the united nations. URL: http://www.fao.org/country-showcase/selected-product-detail/en/c/1287948/. Accessed: 2021-08-03.

[3] Hammadi Hamza, Mokhtar Rejili, Elbekkay Mokhtar, and Ali Ferchichi. New approach for the morphological identification of date palm (*Phoenix Dactylifera L.*) cultivars in tunisia. *Pak J Bot*, 41, 12 2009.

[4] Ramzi Ali, Ridha Ejbali, and Mourad Zaied. A deep convolutional neural wavelet network for classification of medical images. *Journal of Computer Science*, 14:1488–1498, 11 2018.

[5] Salwa Said, Olfa Jemai, Salima Hassairi, Ridha Ejbali, Mourad Zaied, and Chokri Ben Amar. Deep wavelet network for image classification. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 922–927. IEEE, 2016.

[6] Jinglan Alzubaidi, Laith Zhang, Ayad Humaidi, Amjad Al-Dujaili, Ye Duan, Fadhel Mohammed Al-Shamma, Omran Santamaría, Muthana Al-Amidie, and Laith Farhan. Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions. *Journal of big Data*, 8(1):1–74, 2021.

[7] Wu, Zhang Liuchen, Ruibo Hui, Chen, and Junfei Yi. Fruit classification using convolutional neural network via adjust parameter and data enhancement. In *2020 12th International Conference on Advanced Computational Intelligence (ICACI)*, pages 294–301. IEEE, 2020.

[8] Yam, Rajesh parala, Ramaiah Challa, Krishna Kantharao, and Seetha Rama. Computerized classification of fruits using convolution neural network. In *2020 7th International Conference on Smart Structures and Systems (ICSSS)*, pages 1–4. IEEE, 2020.

[9] Wadha, Jacob Alhamdan, and M. Howe. Classification of date fruits in a controlled environment using convolutional neural networks. In *Advanced Machine Learning Technologies and Applications: Proceedings of AMLTA 2021*, pages 154–163. Springer International Publishing, 2021.

[10] Hasan, Sattar Sakibul, and Abdus. Arabian date classification using cnn algorithm with various pre-trained models. In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, pages 1431–1436. IEEE, 2021.

[11] Hamdi Altaheri, Mansour Alsulaiman, and Ghulam Muhammad. Date fruit classification for robotic harvesting in a natural environment using deep learning. *IEEE Access*, 7:117115–117133, 2019.

[12] Nasiri, Taheri Amin, Garavand Amin, Yu Zhang, and Dong. Image-based deep learning automated sorting of date fruit. *Postharvest biology and technology*, 153:133–141, 2019.

[13] Haidar, Dong Abdulhamid, Mavridis Haiwei, and Nikolaos. Image-based date fruit classification. In *2012 IV International Congress on Ultra Modern Telecommunications and Control Systems*, pages 357–363. IEEE, 2012.

[14] Shorten, Khosh goftaar Connor, and Taghi. A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1):1–48, 2019.

[15] Hammadi Hamza, Monia Jemni, Mohamed Benabderrahim, Ahmed Touil, and Salah Othmani. Date palm status and perspective in tunisia. *Date Palm Genetic Resources, Cultivar Assessment, Cultivation Practices and Novel Products*, pages 191–219, 01 2014.