A deep learning approach for Moroccan dates types recognition.

AOULALAY Ayoub[1], Abderrahim El MHOUTI[1], Abdellah El zaar[2], Rachida Assawab[2], and Mohammed MASSAR[3]

[1] ISISA, FS, Abdelmalek Essaadi University, Tetouan, Morocco
[2]Laboratory of R&D in Engineering Sciences, FSTH, Abdelmalek Essaadi University, Tetouan, Morocco
[3]Analyse Non Linéaire Appliquée

E-mail address: aoulalayayoub@gmail.com, aelmhouti@uae.ac.ma, , abdellah.elzaar@gmail.com, rachida.assawab@gmail.com, massarmed@hotmail.com.

Received ## Mon. 20##, Revised ## Mon. 20##, Accepted ## Mon. 20##, Published ## Mon. 20##

Abstract: the growth of the date fruit market, it becomes necessary to use artificial intelligence techniques to recognize date fruits categories. In this work, we use computer vision approaches to classify date fruits produced in Morocco according to their type. To do this, A dataset has been curated, comprising images of the seven most prevalent types of date fruit found in Morocco. Distinguishing itself from other prominent datasets in the field, our dataset poses a challenge to the model due to its inclusion of images captured under varying conditions. For the recognition algorithm, two computer vision approaches are compared and evaluated in terms of performance. Both approaches are based on transfer learning of a convolutional neural network CNN. These two approaches are standard feature extraction where deep features are used to train a machine learning classifier, we compare four classifiers and show that SVM gives the best results. The second approach is fine-tuning where we fit the pre-trained model to our dataset. The approaches used in this work achieve outstanding performance on our dataset, with a classification precision of 97 %. We employ the GradCam technique to visualize the features of our model, revealing that the model primarily emphasizes the texture of the date fruit in its predictions.

Keywords: dates fruit classification; transfer learning; fine tuning; features extraction; deep convolutional neural network

1. INTRODUCTION

In recent years, deep learning has emerged as a breakthrough in artificial intelligence techniques that has overwhelmed the field of pattern recognition and computer vision research by delivering cutting-edge results. In the agricultural field, this technology is widely used to improve productivity, including the classification of agricultural products from images [2], [17]. Recently, Convolutional Neural Network (CNN) has emerged as a powerful tool for image processing tasks, achieving remarkable results, making it the leading technique for vision applications. In the present study, we will use recent methods based on CNN applied to the classification and recognition of Moroccan dates according to their type. Indeed, Morocco is one of the largest producers of date fruit. According to statistics from the Ministry of Agriculture, Morocco is the seventh-largest area of date palms. The regions that produce the most dates in Morocco are mainly located in the eastern south, in the regions of Erachidia Ouarzazate, Tata Zagora, and Figuig. The date palm chain covers about 59,600 hectares (1 % of the arable land nationwide). Daraa-Tafilalt region is responsible for producing an average of 100 000 Tons of dates yearly.

Morocco produces a variety of dates fruit types with 455 types, but across Morocco, only a few are well known. These types range from the best quality to less quality. Al Majhoul is the date fruit that is considered to have the best quality. And this is what makes it the most expensive date fruit. Erachidia and Ouarzazate are responsible for 90 % of the production of this type. However, the production of this type of dates fruit in Morocco is low (0.3 %) compared to other countries. The most produced dates fruit in Morocco are Bofgous and Jihle with 12.2 % and 11.9 %. The most consumed type of dates fruit is the Tathmout, due to it's high-quality while being very cheap. Bouzkri is widely used

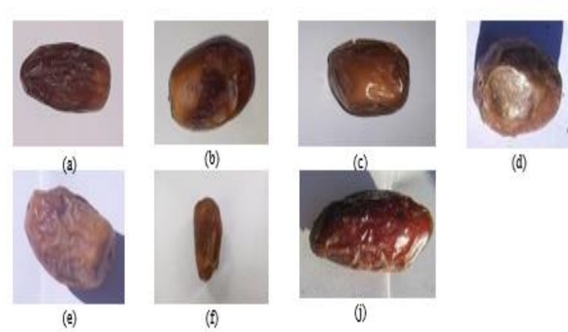

Figure 1: images of the seven Moroccan cultivars of dates used in this work (a): al majhoul ,(b): aziza, (c): bofgous, (d): bouzkri , (e): jihl, (f): sokari , (j): tathmout.

in pastries when it is dry and unlike the other types, Bouzkri reaches its best quality when it ages a little. Aziza is rare in Morocco, and it represents a percentage of 0.2 % of the date production in Morocco, it is located in the region of Fuigig mainly. Although it is imported and not cultivated in Morocco, Sokari is very popular in Morocco and has earned its place on Moroccan tables alongside other types of dates. In this work, we aim to automatically recognize the seven types of dates that are produced in morocco.

Different types of dates are produced in Morocco with high quality types like almajhoul which is well known internationally and other local types only known in their local places. Each type of dates has a price that defines its quality and rarity, Almajhoul is the most expensive because of its high quality. The other types each have their own price. The type of dates is one of the factors that define their marketable price. And thus, it is important to know the type of dates before buying it.

The main problem in distinguishing the type of dates is that dates can be very similar, and are often confused with each other. Distinguishing one type from another requires a certain level of expertise. Figure 1 shows the seven types discussed in this work. some dates fruit share the same visual properties. For example, the texture of Majhoul and Jihl is similar. Bouzkri and Bofgous also have a similar texture. And for the case of shape, Tathmout and Almajhoul are similar. And in terms of color, Almajhoul, Bofgous, and Sokari are similar. another factor that makes dates fruit classification very challenging is that dates fruit from the same class can be very different as in the example in figure2. Here we can see two images of the type Bofgous that are visually very different. This makes classifying dates fruit into types a very difficult process. In this work we use computer vision to automatically classify these seven types of dates fruit. Our approach is based on transfer learning of convolutional neural network CNN.

Figure 2: difference between two dates belonging to bofgous cultivar.

In recent years CNN has shown a great performance in a variety of visual recognition tasks, like face recognition [9], handwritten character recognition [6], medical image classification [22] , Vehicles identification [25] and deepfake detection [26] . CNN architectures like AlexNet [13] VGG [19], ResNet [8] have reached a precision that is close to perfection in the ILSVRC challenge.

It has been shown that increasing the depth of the CNN network leads to increased performance [13] [19]. And as shown in many works The CNN relies on training data to build its image description. Each CNN filter is adjusted during training to extract a discriminative feature type for the current task. This makes the CNN a data-intensive model, since the essential part of the CNN, as well as its depth, is the data. To train a deep CNN, a large amount of data is needed, as adding more layers means having more trainable parameters. with limited data, a state-of-the-art architecture will suffer from overfitting [5]. The need for a large amount of data is the main drawback of CNN. In many tasks, data collection can be expensive and is not available to everyone. In our case, we have 700 images split between training and testing. This is too little compared to ImageNet (14 million images), and it is not enough to train a state-of-the-art network. In this work, to overcome this problem, we use transfer learning [21] [24]. It is the process of using knowledge from one model to improve the other. In tasks with limited data, it is useful to use a model that has performed well on a task closer to the task at hand. In object recognition tasks the models used to transfer knowledge are trained on ImageNet. The weights of these models are used to extract the saliency representations from the limited data of the task at hand. This can be done in two ways. The first is to extract off the shelf features that will be used to train a new classifier. The second is to adapt the source network to our task, replacing the last layers with new ones and fine-tuning them to our task.

In this work, we compare the two transfer learning strategies for the classification of Moroccan dates fruit. With transfer learning, we were able to achieve a very good performance with a classification precision of 97 % using both approaches.

This paper is organized as follows: firstly, the document sets out the related work. Then the paper presents the materials and methods. In this section, we describe the data, and the theory of the approaches used for the automatic recognition of Moroccan dates fruits. The paper then presents and discusses the results obtained. Finally, the article ends by discussing future work.

2. RELATED WORKS

The categorization and classification of dates fruit have attracted the attention of many researchers in recent years. Especially in countries that produce dates. The categorization of dates is very important because it is a factor in the marketing of dates. thus, many works have proposed the use of chemical or organic characteristics of dates to categorize them.

The authors in [11] suggested the use of chemical characteristics as well as morphological and textural characteristics to classify date fruit. They studied the most important characteristics to classify 20 types of dates grown in the United Arab Emirates (UAE). Their results show that textural characteristics are not essential for date fruit classification.

Although classifying date fruit is a fine-grained classification challenge, distinguishing between them is mostly a visual task. It only takes a glance for experts to know which type of date fruit it is. For this reason, much work has focused on date fruit classification using visual information alone.

The authors in [12] used machine learning to automatically classify dates fruit into categories. Their study was conducted on dates produced in the Middle East. To do so, they used three image description methods, these methods are based on extracting color, shape, and morphology features. The extracted features are then fed into logistic regression and artificial neural network models for classification. Their study shows that the best results were obtained by combining logistic regression and neural network, with an accuracy of 92.8 %.

The authors in [14] proposed a new approach for the detection and classification of date fruit. Their approach is based on the extraction of color, shape, and size features. These features will be used to train a neural network classifier. They obtained an accuracy of 97.2 % on the classification of three types of dates that grow in Pakistan.

The authors in [7] used shape, color, size, and texture feature extraction for the classification of seven date fruit produced in the UAE. They compared three classification techniques: KNN, LDA, and neural network. They found that the best results were obtained using the neural network with 98.6 % accuracy.

The authors in [15] proposed a pipeline for the classification of four types of date fruit grown in the Arabian Gulf countries. Their system is based on combining texture, size, and shape features to describe date fruits images. In their work, texture features are extracted from three color components. for this they compared several texture descriptors and achieved the best performance using a combination of the three feature extraction methods, with 98.1 % accuracy.

Transfer learning was utilized in [1], where the authors used a model based on MobileNetV2 to classify dates from Saudi Arabia. To train their model they used a database comprising images captured under stable lighting conditions, obtained using a 48 cm, 55 W 5500 K Dimmable LED Ring Light outdoor mobile tripod. This device was also employed to remove shadows. The camera-to-date fruit distance was fixed for all images, and the images were captured in daylight to avoid texture changes. Consequently, the database contains high-quality images, and their model achieves an accuracy of 99 %. The model is an excellent tool for classifying date fruits in industrial applications where such conditions and devices are available. However, such conditions are not always feasible. In our study, we use images with various lighting conditions, varying image qualities, and the presence or absence of shadows, as shown in the figure. We aim to collect images in different scenarios and qualities to present a challenge to the classification model. We believe this will lead to a model that generalizes better in real-world scenarios where stable conditions are not always available.

The works described above have focused on the classification of dates from the East and the Middle East. They achieved good results, sometimes close to optimal classification accuracy, using only handmade descriptors. This indicates that in this domain, perfect hand-designed features have been discovered. Moroccan dates have some similarities with the middle east date fruit types and some differences. Many types that are grown in Morocco can be very similar to each other. Besides, a type of dates fruit may have different subtypes like Bofgous, and Tathmout with black and red Tathmout. This makes the classification of Moroccan date types a very difficult task. In our work, we investigate the power of deep learned features for Moroccan dates fruit categorization. For this purpose, we use transfer learning of a deep convolutional neural network.

3. DATA

Our dataset was collected manually by taking photos of seven types of Moroccan dates. The photos of Al majhoul, Jihl, bofgous, Tathmout, Sokari Aziza, and Bouzkri are taken using cameras and smartphones. Each image in the dataset represents a single dates fruit on a white background. Our dataset contains seven classes with 100 images per class. Each class has a large intra-class variance because each class contains dates with different harvest dates. Each date cultivar appears different as it ages. And the appearance as a function of harvest date differs from one cultivar to another. For example, Bozkri and jihl harden over time, and other cultivars may look the same with just a small difference. Dates are purchased in local markets, so we are not sure of the dates and places of harvest, nor of their post-harvest history. The dataset we collected contains images of date cultivars with different picking dates. The images are taken under different conditions (different lighting, different image quality, different cameras) in the training and evaluation sets. The reason we use images with different conditions is to impose some challenge on our model. This will make the model robust to changes in the input and image conditions, as we want it to be usable in real-world applications, such as smartphones. To ensure that the model makes the decision based only on the discriminating features of the date presented in the image, and is not sensitive to the state of the input image with different lightning and conditions. We use the model visualization tool, gradcam.

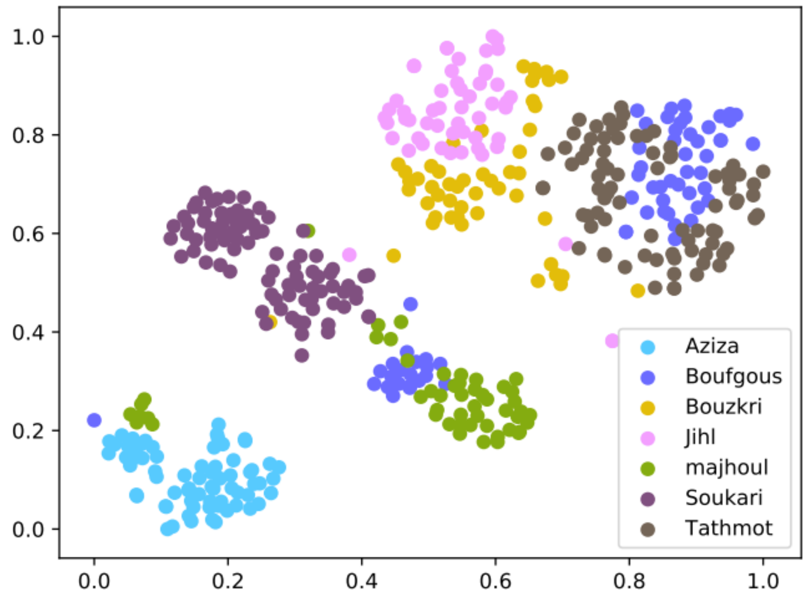

Figure 3: : T-distributed stochastic neighbor embedding(t-sne) visualization of the dataset using ResNet50 deep features. different colors represent the different cultivars. Samples that are similar are close to each other in the graph.

Grad-cam is a neural network visualization technique introduced by [16]. As the name implies, Grad-cam uses the gradient of the classification prediction against the feature maps of the last convolutional layers to show the parts of the input image that lead to the decision made by the CNN, and that are most important for distinguishing one class from another.

We divide the dataset into 80 % for training and 20 % for testing. To analyze the distribution of the data, we use t-distributed stochastic neighbor embedding T-SNE visualization. We use as input to T-SNE, the features extracted by the ResNet50 model pre-trained on ImageNet. The results are presented in figure4. We can see that the most similar date fruit types are clustered next to each other.

4. Convolutional Neural Network and Transfer Learning

In recent years, CNN has become the state of the art in various image classification and pattern recognition tasks. CNN is designed to represent the image content by extracting a hierarchical representation that starts at a low level by extracting low-level features such as edges and color blobs, up to high-level features that are class-specific. Since 2012, CNN has outperformed other machine learning techniques such as random forest [3] and SVM [4] in a variety of image classification challenges like the ImageNet Large Scale Visual Recognition Challenge ILSVRC. The CNN has benefited from advances in computing resources and technology, with advanced systems such as GPUs for deep learning applications. These systems are able to compute operations in tensors, which are the building blocks of a neural network, more quickly. The main difference between CNN and other machine learning models is that the image representation part is learned from with convolutional layers in CNN, unlike other methods where the image representation has to be extracted using hand-made descriptors. It is believed that a deep network should be able to discover by hidden layer the effective representation of a given task [5]. As shown in many works the parameter that led to better accuracy is the depth of the networks. A deeper CNN network will succeed in extracting features that are highly related to the task and that are discriminative enough to separate classes. The only problem is that a deeper network has a large number of parameters, and they need a large training set to be trained. In our work, we only have 100 images per class, which is not sufficient to train a deep CNN.

One solution for this problem is transfer learning. Transfer learning [21], [24], [18], is the process of using the knowledge of a source network to improve the performance of another one. This is based on the hypothesis that a network that has been trained on a dataset to perform a task has built some knowledge that allows him to perform well on another task close to the one it's trained on. Transfer learning is defined as an operation of using a model to improv another.

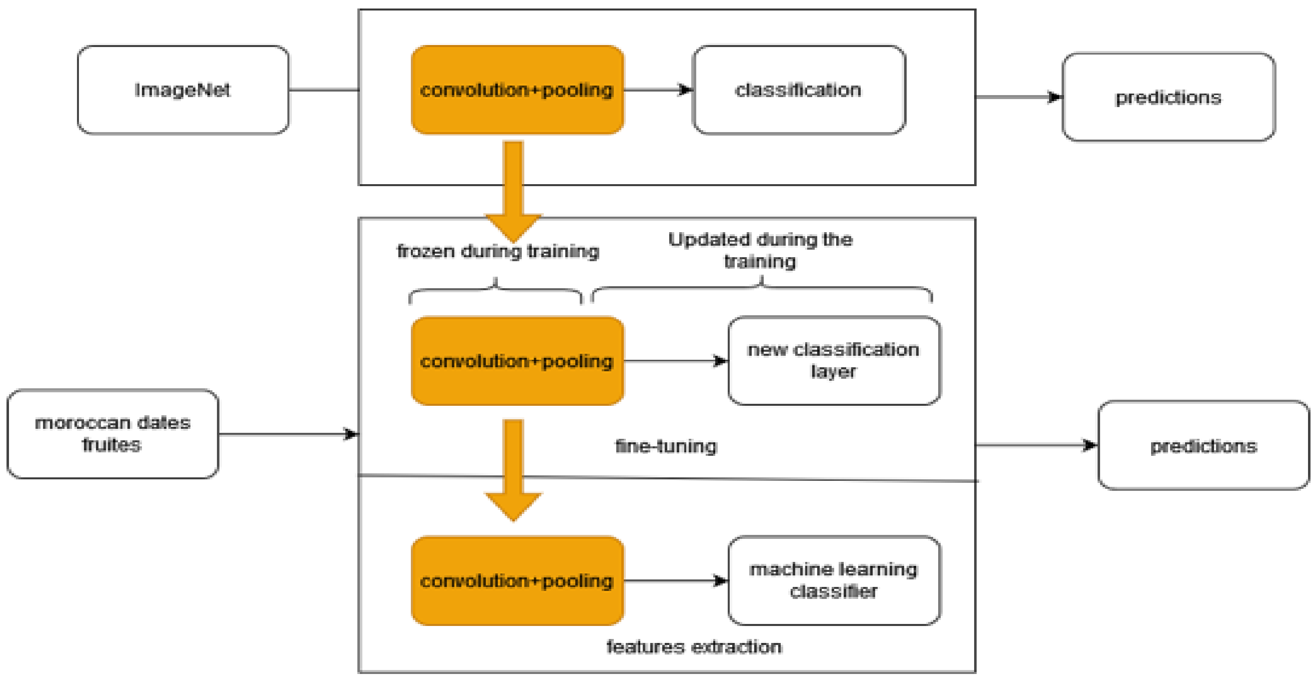

Figure 4: Representation of transfer learning, the convolutional part of a model trained on ImageNet is transferred to our model which will be trained on our dataset. For ne-tuning, some of the first layers are frozen and the others are updated during training. For feature extraction, the features extracted using the transferred part are used to train a machine learning classifier.

a domain D is composed of a features space χ and a marginal distribution P(X), D= {χ , P(X)} where X is the set of instances X= {x1, x2,. xn}, and a learning task is composed of a label space γ, and a prediction function f. T={γ,f}. note that the prediction function is learned from the data. Transfer learning involves two elements the first one is the source model that has a source domain Ds and a source task Tt. The second element is the target model that has the target domain Dt and the target task Tt. Transfer learning can be defined as the process of using the knowledge gained from training the source model trained on the source domain Ds and the source task Ts, to improve the prediction function of the target model Ft on the target domain Dt. The condition of transfer learning is that both source and target domains have some similarities. For example, a professional swimmer will need less training to be a Water polo player, than a regular person. In transfer learning, the learning model from one task is reused as a starting point for a second task [5] In our case, Moroccan dates fruit classification, the source model is a state of art network, trained on the ImageNet. The source and the target domain have some similarities as they are all composed of images. The tasks also have similarities as they focus on the classification of images. ImageNet is also very close to our dataset, as it contains a variety of fine-grained classes like different breeds of dogs for example. It was shown that ImageNet is a powerful source for transfer learning tasks. And many schools of thought have been raised on the reason why ImageNet performs well on transfer learning tasks [10]. The overall structure of Image Net is what makes it suitable for the transfer learning application, as it has a large number of images divided into 1000 classes. These classes vary from coarse-grained to fine-grained classes. The diversity of classes as well as the number of images and classes makes ImageNet a perfect data source for transfer learning.

In this work, we investigated two transfer learning strategies. The first one is by using off-the-shelf features of the source model to extract features from our dataset. These features are then used as an input of a machine learning classifier. Image representation extracted with a pre-trained network has shown a great performance in fine-grained image classification challenges [18].

The second strategy is fine-tuning. Fine-tuning consists of using the weights of the pre-trained model as a starting point for the target model. CNN is a hierarchical features extractor, as it extracts different levels of features. We can categorize these features into two types, general and specific [23]. general features are the ones that can be used in all image classification tasks. Specific features represent the characteristics that are related to the task at hand. General features as shown in many works are extracted in the first and the mid-layers of the network while the specific features are extracted in deeper layers. The basic idea of fine tuning is using the layers of the source network that it's believed extract features related to both tasks (ILSVRC and Moroccan dates fruits classification) meaning that these features are general. And we replace the task specific layers with new ones. For this we use a state of art network trained on ImageNet as a source network. The first layers that extract general features are frozen, (i.e. they are not updated during the training) and we replace the fully connected layers with a new randomly initialized one, and we train the whole network on our dataset.

5. RESULTS AND DISCUSSION

In this section, we present the results and analysis of the two transfer learning strategies used in this work on Moroccan date fruit categories datasets. To evaluate the classification performance, we divided our database into 80 % training images and 20 % test images. We then extracted the precision and the accuracy of the five classes along with other evaluation measures. We used the [8:2] split due to the small size of the data set. We can confirm that human-selected hyperparameter bias did not contribute to our results because in our experiment we use a standard set of hyperparameters. This set was not changed at any point in the experiment. Before feeding the images to the model we preprocess them by subtracting the mean RGB pixel intensity from the ImageNet dataset.

To evaluate the performance, we use the classification report tool of scikit-learn library in python, it provides us with three metrics.

- the precision is the ratio of correct forecasts to supposedly positive ones.

$$precision = \frac{TP}{TP + FP} \quad (1)$$

- the recall determines the detection capability of all images that represent the target class in a dataset.

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

- the F1-score allows to establish the balance between precision and recall.

$$\mathrm{F1-score} = 2\mathrm{x}\frac{Precision\ x\ Recall}{Precision + Recall} \quad (3)$$

A. Features extraction results

Feature extraction involves using the pre-trained model to extract features from the dataset. These features will be used to train a machine learning classifier. We have tested several pre-trained source models, which are trained on ImageNet. And we compare the performance of a variety of machine learning classification algorithms. The models tested are VGG16, VGG19, ResNet50, and Inception V2. We compared four classifiers: KNN, SVM, logistic regression, and Random Forest. The results are presented in Table 1.

The best precision is obtained with VGG19 as feature extractor and SVM as a classifier. on the other hand, the lowest precision was obtained with Inception V2 as features extractor and KNN and random forest as a classifier. Random forest and KNN have the lowest prediction precision compared to SVM and logistic regression which perform better. And for feature extraction, Inception V2 performs worse than other source models.

TABLE 1: RESULTS OF FEATURES EXTRACTION APPROACH USING DIFFERENT SOURCE MODELS AND DIFFERENT CLASSIFIERS.

pretrained model	**The machine learning classifier**			
	SVM	**KNN**	**LR**	**Random-Forest**
VGG19	97%	83%	96%	88%
VGG16	95%	94%	95%	91%
ResNet50	95%	92%	95%	91%
Inception V2	87%	82%	88%	82%

TABLE 2: THE PRECISION OF NE-TUNING DIFFERENT CONVOLUTIONAL NEURAL NETWORK MODELS IN %. VGG16 GIVES THE BEST PERFORMANCE AND INCEPTIONV3 IS THE WORST PERFORMING.

model	**precision**
VGG19	94%
VGG16	97%
ResNet50	95%
Inception V2	93%

B. Fine-tuning results

Fine-tuning involves replacing the task-specific layers of a pre-trained CNN with new layers and training the entire network for the target task. As we did in the feature extraction approach, we compare four architectures, VGG16, VGG19, Inception V3, and ResNet. We replace the set of fully connected layers with a new one whose weights are chosen stochastically. The new classifier is composed of four layers whose number of filters is respectively (128, 64, 32, and 7).

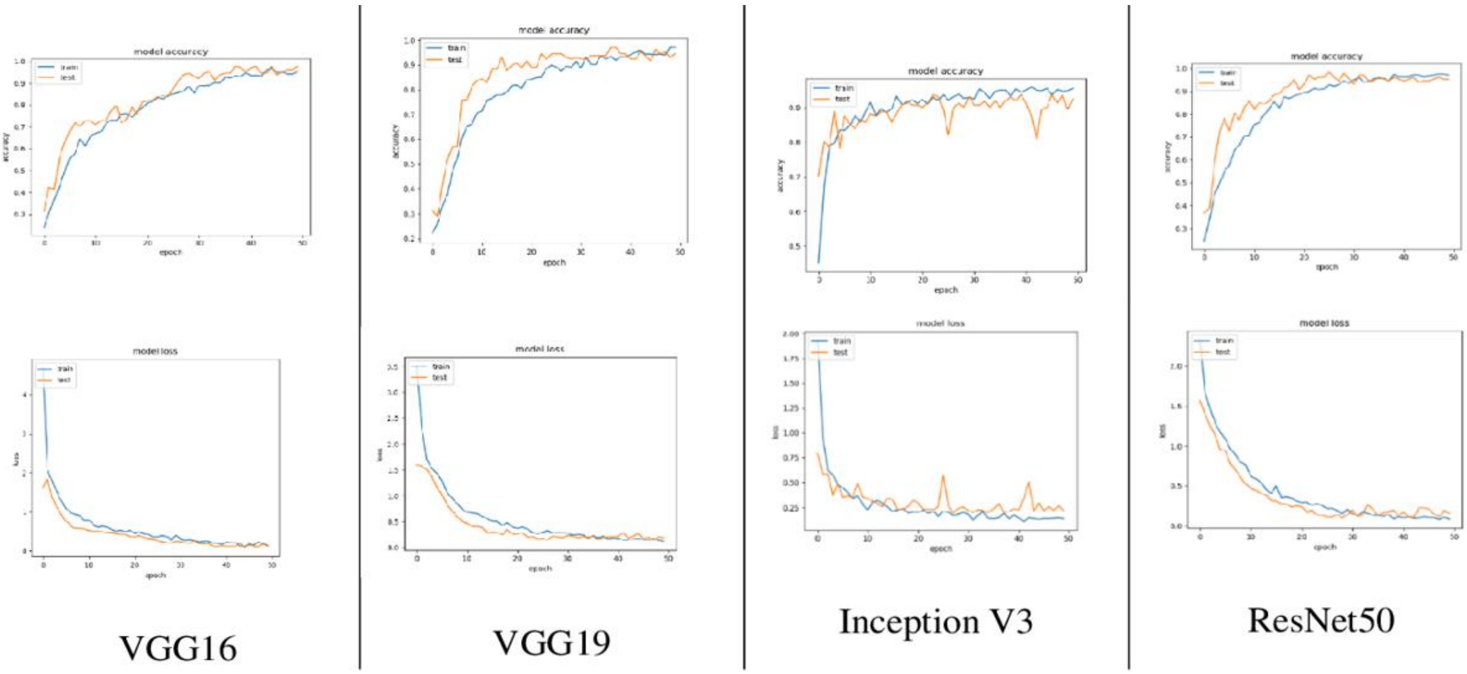

Figure 5: accuracy and loss curves of ne-tuning different models the first line belongs to VGG16 the second for VGG19 the third for INCEPTION and the last one for RESNET50

The last layer is a Soft-Max layer, while the others are activated with Relu[13]. Dropout [20] is used to reduce overfitting and the whole network is trained for 50 epochs with stochastic gradient descent as an optimizer. The learning rate is reduced during training.

the results are reported in fig5 the best precision is obtained with VGG16 as the source model with 97 %, and the lowest when using InceptionV3 with 93 %. The fine-tuning approach shows the same performance as the feature extraction approach, achieving the same accuracy of 97 %. The only difference is in which model has the best accuracy,

In the fine-tuning approach, VGG16 is the best performing model, while it is better to use VGG19 in the features extraction approach. The accuracy and loss curves in figure7 show that the model does not suffer from overfitting despite having limited data. In the first few epochs, the validation accuracy is better than the training accuracy, but after a few epochs, they start to be equal.

Transferring the image representation of a network trained on ImageNet helps in date fruit classification. Keep in mind that date fruit classification is a fine-grained challenge, and one of the key elements that make ImageNet ideal for transfer learning tasks is that it has a large number of fine-grained classes. A network trained on ImageNet always learns to distinguish subclasses, and filters learned from ImageNet has specialized in this way. Fine-tuning this filter towards dates fruit is more efficient and faster than learning this representation from scratch.

C. Class precision

We extract classification precision off each of the Moroccan date fruit types used in this work. for this purpose, we use the best-performing model in the features extraction approach and fine-tuning approach. For features extraction approach, we use VGG19 and SVM as it gives the best results. And for fine-tuning we use VGG16. We extract

precision, recall, and f1-score and show them in the table2 and table3. table2 shows the class-performance off the features extraction approach .it shows that only two classes don't have perfect precision. These classes are Tathmout and Bouzkri which obtain respectively 91 % and 86 %. The other classes are classified perfectly. For the recall, only two classes are not detected perfectly. These are al Majhoul and Bofgous with 85 % and 90 % respectively.

Table3 represent the fine-tuning approach. the precision of the classes is perfect except for the precision off Tathmout and Boufgous that reaches 95 % and 86 % respectively.

Three classes have a recall lower than 100 %. these are al majhoul with 90 %, tathmout 95 % and boufgous 95 %. Bofgous and Tathmout represent a challenge for our approach, as these are the classes that our model has difficulty recognizing using both methods. Bofgous have high intra-class variance because there are a variety of

Bofgous types and thus there will be subclasses within the same class.

TABLE 3: THE PRECISION, THE RECALL AND F1 SCORE O EACH CLASS USING FEATURES EXTRACTION OF VGG19 AND SVM CLASSI ER IN %. BOUZKRI , SOUKARI , JIHL AND AZIZA HAVE THE PERFECT PRECISION AND RECALL.

Class	precision	Recall	F1-score
Almajhoul	100	85	91
Tathmout	91	100	95
Bouzkri	100	100	100
Boufgous	86	90	88
Soukari	100	100	100
Jihl	100	100	100
aziza	100	100	100

TABLE 4: THE PRECISION, THE RECALL AND F1 SCORE O EACH CLASS OF NE TUNING VGG16 IN % . BOUZKRI , SOUKARI , JIHL AND AZIZA HAVE THE PERFECT PRECISION AND RECALL.

Class	precision	Recall	F1-score
Almajhoul	100	90	95
Tathmout	95	95	95
Bouzkri	100	100	100
Boufgous	86	95	90
Soukari	100	100	100
Jihl	100	100	100
aziza	100	100	100

For Tathmouts, we can explain the model's performance by the fact that Tathmouts vary visually with their level of maturity as well as their age. aging only few days can make a huge difference in the appearance of the Tathmout.

D. *The level of features*

It is believed that the general image representation suitable for both the source and target task resides in the intermediate layers [5]. We tested this assumption by testing the power of the extracted features at different levels of CNN. For this, we use different blocks of the VGG16 network pre-trained on ImageNet.

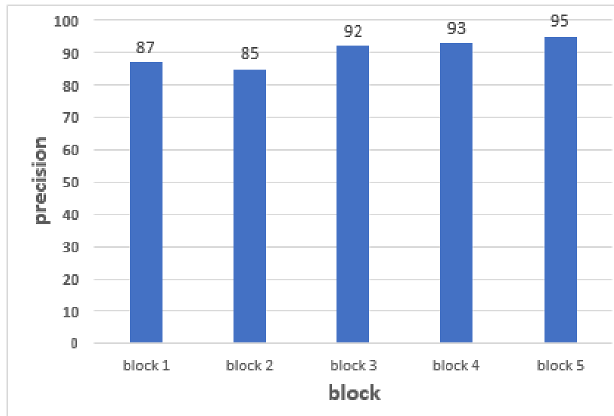

Figure 6: the precision of features extracted from different blocks of VGG16 network in %. Extracting features from deeper layers (block 5) gives the best accuracy, compared to earlier layer (block1 and block2).

The convolutional part of VGG16 is divided into five blocks as shown in the figure. Each block ends with a maximum pooling layer. we extract features from different levels of CNN by using the output off the max polling layer at the end of each block. these features will be fed into an SVM model for classification. The model's blocks indicate the level of the extracted features. From low-level features extracted in the first block to high-level task-specific features extracted in block5. figure8 shows the classification accuracy corresponding to each block. The best results are obtained by using the features of the last convolutional block with 95 %. At the same time, the accuracy increases with increasing feature levels. This indicates that the optimal representation shared by ImageNet classification and Moroccan date fruit classification is in deep layers close to the classification part. We can say that the similarity between the two tasks is huge. these also shows the power of ImageNet deep features because even using the low-level representation in block1, we still get a good precision with 87 %.

E. *model visualization*

We use gradCam to visualize the concentration of the last convoluted layer. For this purpose, we used a ResNet50 model trained on our dataset. The results of applying gradCam to images of the seven date cultivars used in this work are shown in Figure8. The model focuses on the surface of the date. This means that date texture and color are the main features that contribute to the final decision.

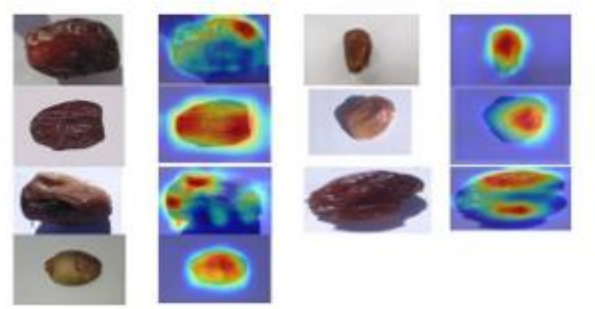

Figure 7: Visualization of the gradient concentration of the last convolution layer of a ResNet50 model using Grad-Cam. The images represent seven date cultivar samples used in this work, the corresponding Grad-cam map is to the right of each sample, and shows that the algorithm focuses on the date surface.

F. *deep features Vs hand-crafted features*

Many works have used hand-created features prior to classification and have performed well on different datasets. The problem with this approach is that we have to use improved image quality in the training and test dataset, either by using a system to obtain images with the same lighting and conditions as in [12], which achieves 92.8 % accuracy. This is a drawback because the ordinary everyday user would not have access to the same conditions. The same is true for [14], which achieved 97.2 % accuracy, but used the same conditions for data collection. They used the same camera and manual preprocessing before feeding images to the classification algorithm. [7] achieved an accuracy of 98.6 %. But a thresholding was performed to segment the date from its background in order to optimally apply the feature extraction. In our work, the only preprocessing we use is image normalization by subtracting the average RGB pixel intensity from the ImageNet dataset. We use images of different conditions and qualities to impose some challenge on the model, and to exploit real-world scenarios. For this, we use deep learning features that we believe are robust to changes in input conditions, and we obtained an accuracy of 97 % in a challenging dataset.

6. CONCLUSION

Dates are an important product in the countries that produce them, this type of fruit has become in recent years an important aspect of the diet. Indeed, the market for dates has developed rapidly throughout the world. One of the important factors that define the price and quality of dates is their type. In this work, we introduce an approach for automatic recognition of Moroccan dates fruit types based on transfer learning of a convolutional neural network. The first step is the collection of data on which we built a dataset of seven well-known date types in Morocco. Our model is used to classify the images of this dataset. We compare two approaches: feature extraction and fine-tuning. Our results show that both approaches achieve a very good performance with a classification precision of 97 %. We find that in the feature extraction approach, VGG19 combined with SVM performs best, while in the fine-tuning approach, VGG16 is the best performing model. In future studies, we will make further improvements to increase the classification performance.

References

[1] Khalied Albarrak, Yonis Gulzar, Yasir Hamid, Abid Mehmood, and Arjumand Bano Soomro. 2022. A deep learning-based model for date fruit classification. Sustainability 14, 10 (2022), 6339. https://doi.org/10.3390/su14106339

[2] Pham Thi Quynh Anh, Dang Quoc Thuyet, and Yuichi Kobayashi. 2022. Image classification of root-trimmed garlic using multi-label and multi-class classification with deep convolutional neural network. Postharvest Biology and Technology 190 (2022), 111956. https:
//doi.org/10.1016/j.postharvbio.2022.111956.

[3] Leo Breiman. 2001. Random forests. Machine learning 45, 1 (2001), 5–32. https://doi.org/10.1023/A:1010933404324.

[4] Gayani Chandrarathne, Kokul Thanikasalam, and Amalka Pinidiyaarachchi. 2020. A comprehensive study on deep image classification with small datasets. In Advances in Electronics Engineering. Springer, 93–106. https://doi.org/DOI:10.1007/978-981-15-1289-6_9.

[5] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. 2014. Decaf: A deep convolutional activation feature for generic visual recognition. In International conference on machine learning. PMLR, 647–655. https://doi.org/ 10.48550/arXiv.1310.1531.

[6] Abdellah Elzaar, Rachida Assawab, Ayoub Aoulalay, Lahcen Oukhoya Ali, Nabil Benaya, Abderrahim El Mhouti, Mohammed Massar, and Abderrahim El Allati. 2021. Amazigh Handwriting Recognition System—Multiple DCNN Strategies Applied to New Wide and Challenging Database. In International Conference on Advanced Technologies for Humanity. Springer, 241–252. https://doi.org/10. 1007/978-3-030-94188-8_23.

[7] Abdulhamid Haidar, Haiwei Dong, and Nikolaos Mavridis. 2012. Image-based date fruit classification. In 2012 IV International Congress on Ultra Modern Telecommunications and Control Systems. IEEE, 357–363. https://doi.org/10.1016/j.postharvbio.2019.04.003

[8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition. 770–778. https://doi.org/10.48550/arXiv.1512.03385

[9] Guosheng Hu, Yongxin Yang, Dong Yi, Josef Kittler, William Christmas, Stan Z Li, and Timothy Hospedales. 2015. When face recognition meets with deep learning: an evaluation of convolutional neural networks for face recognition. In Proceedings of the IEEE international conference on computer vision workshops. 142–150. https://doi.org/10.48550/arXiv.1504.02351

[10] Minyoung Huh, Pulkit Agrawal, and Alexei A Efros. 2016. What makes ImageNet good for transfer learning? arXiv preprint arXiv:1608.08614 (2016). https://doi.org/10.48550/arXiv.1608.08614

[11] Afaf Kamal-Eldin and Sami Ghnimi. 2018. Classification of date fruit (Phoenix dactylifera, L.) based on chemometric analysis with

multivariate approach. Journal of Food Measurement and Characterization 12, 2 (2018), 1020–1027. https://doi.org/10.1007/s11694-018-9717-4

[12] Murat Koklu, Ramazan Kursun, Yavuz Selim Taspinar, and Ilkay Cinar. 2021. Classification of date fruits into genetic varieties using image analysis. Mathematical Problems in Engineering 2021 (2021). https://doi.org/10.1155/2021/4793293

[13] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. 2012. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems 25 (2012). https://doi.org/10.1145/3065386

[14] Aurangzeb Magsi, J Ahmed Mahar, Shahid Hussain Danwar, et al. 2019. Date fruit recognition using feature extraction techniques and deep convolutional neural network. Indian Journal of Science and Technology 12, 32 (2019), 1–12. https://doi.org/10.17485/ijst/2019/v12i32/146441.

[15] Ghulam Muhammad. 2014. Automatic date fruit classification by using local texture descriptors and shape-size features. In 2014 European Modelling Symposium. IEEE,174–179.https://doi.org/DOI:10.1109/EMS.2014.63.

[16] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. 2017. Grad-cam: Visual explanations from deep networks via gradient-based localization. (2017), 618–626. https://doi.org/10.1109/ICCV.2017.74

[17] Anindita Septiarini, Andi Sunyoto, Hamdani Hamdani, Anita Ahmad Kasim, Fitri Utaminingrum, and Heliza Rahmania Hatta. 2021. Machine vision for the maturity classification of oil palm fresh fruit bunches based on color and texture features. Scientia Horticulturae 286 (2021), 110245. https://doi.org/10.1016/j.scienta.2021.110245

[18] Ali Sharif Razavian, Hossein Azizpour, Josephine Sullivan, and Stefan Carlsson. 2014. CNN features off-the-shelf: an astounding baseline for recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 806–813. https://doi.org/10.48550/arXiv.1403.6382

[19] Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014). https://doi.org/10.48550/arXiv.1409.1556

[20] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research (2014). https://doi.org/10.5555/2627435.2670313

[21] Karl Weiss, Taghi M Khoshgoftaar, and DingDing Wang. 2016. A survey of transfer learning. Journal of Big data 3, 1 (2016), 1–40. https://doi.org/10.1186/s40537-016-0043-6

[22] Samir S Yadav and Shivajirao M Jadhav. 2019. Deep convolutional neural network based medical image classification for disease diagnosis. Journal of Big Data 6, 1 (2019), 1–18. https://doi.org/10.1186/s40537-019-0276-2

[23] Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson. 2014. How transferable are features in deep neural networks? Advances in neural information processing systems 27 (2014). https://doi.org/10.48550/arXiv.1411.1792

[24] Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He. 2020. A comprehensive survey on transfer learning. Proc. IEEE 109, 1 (2020), 43–76. https://doi.org/10.48550/arXiv.1911.02685.

[25] Zohreh Dorrani, Hassan Farsi, and Sajad Mohammadzadeh. Edge detection and identification using deep learning to identify vehicles. Journal of Information Systems and Telecommunication (JIST), 3(39):201, 2022.

[26] Adeep Biswas, Debayan Bhattacharya, and Kakelli Anil Kumar. Deepfake detection using 3d-xception net with discrete fourier transformation.Journal of Information Systems and Telecommunication (JIST), 3(35):161, 2021.