DOI: 10.15849/IJASCA.230320.03

Int. J. Advance Soft Compu. Appl, Vol. 15, No. 1, March 2023
Print ISSN: 2710-1274, Online ISSN: 2074-8523

Classification of Moroccan decorative patterns based on computer vision approaches using complex datasets

Ayoub Aoulalay[1], Abderrahim El Mhouti[1], Mohammed Massar[2], Mohamed Fahim[3] and Yassine El Borji[3]

[1] ISISA, FS, Abdelmalek Essaadi University, Tetouan, Morocco
e-mail: aoulalayayoub@gmail.com
[1] ISISA, FS, Abdelmalek Essaadi University, Tetouan, Morocco
[2]ANLA, FSTH, Abdelmalek Essaadi University, Tetouan, Morocco
[3]ISISA, FS, Abdelmalek Essaadi University, Tetouan, Morocco
[3]SOvIA, ENSAH, Abdelmalek Essaadi University, Tetouan, Morocco

Abstract

Because of its geography and long history, Morocco has seen many populations cohabit or succeed each other on its soil. Consequently, Morocco inherits very rich decorative arts which are traditions and know-how still very much alive. Archaeological sites, mosques, madrasas and palaces, medinas, Berber casbahs, ceramics and multicolored carpets, etc. ... Morocco is rich in a unique artistic tradition. However, given the variety and richness of Moroccan decorative arts, the classification of decorative motifs is a major issue today. Indeed, there are many types of decorations that consist of decorative units borrowed from nature with an abstraction or simplification and some additions or modifications. In this con-text, the classification of motifs according to type is therefore an important aspect for better understanding this art. The objective of this work is to use computer vision approaches for the classification of Moroccan decorative motifs. Thus, three computer vision approaches are compared and evaluated in terms of their performance. For this purpose, two datasets are used with a different level of challenge, one containing texture images and the other containing patterns in complex scenes. The three approaches studied are hand-created features, training a convolutional neural network from scratch and transfer learning. The best results are obtained with transfer learning on both datasets, reaching 95% accuracy. Transfer learning is also generalizable, giving good results in both datasets used.

Keywords: *Texture classification, Moroccan decorative motifs, material recognition, hand-crafted features, transfer learning, deep convolutional neural network*

1 Introduction

The Moroccan decorative style has become very popular in recent years because of its richness and harmony. Due to the particular aesthetics and elegance, it brings, many people outside of Morocco have a tendency to decorate their homes using the Moroccan decorative style. The principle of Moroccan decorating is based on the organization of basic patterns to create a very refined textured decor. The Moroccan ornamental style consists of the

Received 20 October 2022; Accepted 23 January 2023

organization and repetition of five fundamental elements that are "ceramic, fabric, wood, plaster and glass". These are simple elements, but Moroccan artists, over the course of history, have found a way to organize these elements to create excellent visual scenes. Each of these elements has unique characteristics, but they are not identified by the machine. In this work, we propose a new method using computer vision for detecting and classifying these components.

The classification of Moroccan decorative patterns is a material recognition challenge, it is difficult to distinguish the material that makes up the decorative patterns. The five patterns can sometimes be visually alike when used in decorative scenes. The material can be used to create the same decorative scenes. for example, fig1 shows that a rosette (an Islamic decoration very popular in Morocco) is made of ceramics and plaster, the pattern in the two images is similar but the material is different.

(a) (b)

Fig1: an example of decorative scenes made of (a): ceramic (b): plaster from [25].

In this work, we used supervised machine learning techniques for the automatic recognition of five basic Moroccan decoration patterns. For this purpose, three approaches are used, these are 1-handcrafted features, 2- training a convolutional neural network from scratch, and 3-transfer learning. To cover different scenarios, we tested their performance in two databases, one texture database that contains texture images of the decorative pattern and the other one contains the objects in scenes.

In the first strategy, we used machine learning approaches based on handcrafted approaches. Handcrafted approaches require the extraction of features that must be discriminative and sufficiently represent the image content. the classification comes next, where the classifier takes the feature vector and gives us the class of the pattern based on the discrimination it has learned from other samples called the training dataset. We propose that the texture and color features are sufficient to tell which of the five Moroccan decorative patterns the image contains [1]. For this, we tried to classify the Moroccan decorative patterns using feature extraction techniques the gray level co-occurrence matrix GLCM [2] and the local binary pattern LBP [3] combined with the color histogram [4], and using an SVM [5] and KNN [6] classifier. The best performance was obtained using the grayscale co-occurrence matrix combined with the color histogram and using an SVM classifier. Our results indicate that the performance of the handcrafted feature approach differs depending on the difficulty level of the dataset. This approach performs well on the texture dataset, reaching 77%, but performs poorly on the more complex dataset with 54% precision.

In the second strategy, we evaluated the performance of a convolutional neural network CNN. CNN has become very popular in recent years due to its capability and high performance in various tasks such as image classification, speech recognition, natural language analysis, etc. On the image classification task, since 2012, the classification accuracy, using CNN outperforms the accuracy using other machine learning techniques such as SVM [5] and Random Forest [7]. Since then, the state of the art of image classification is obtained with different CNN architectures such as VGG16[8] and ResNet [9]. Some of the more recent CNNs and their accuracy are shown in Fig2.

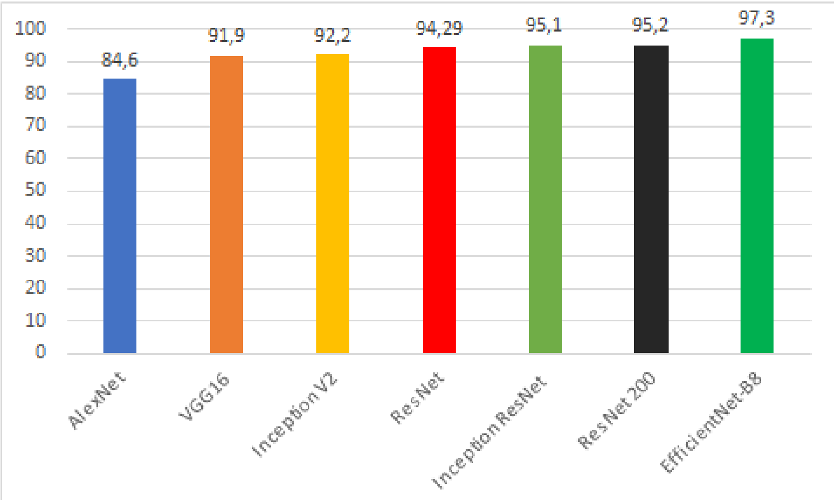

Fig2: the accuracy of different state of art CNN architectures

The problem is that CNN requires a large amount of data to be used as a classification tool, sometimes thousands or even millions of data are needed to achieve acceptable classification accuracy. We study this problem by training a CNN from scratch to classify Moroccan decorative patterns, noting that the two databases used have 130 and 120 images per class respectively, which is insufficient to train a CNN. To solve this problem, we followed the work [10] by using a shallow CNN with a reduced number of parameters. After many experiments, we came to some interesting conclusions.

- The performance of the CNN is higher if trained on the texture database than if trained on the objects databases.
- Data augmentation improves the performance of the CNN, when trained on the objects database, but does not help the classification performance when using texture databases.

In the third strategy, we experimented with another solution to the problem of not having enough data to train a convolutional neural network, namely transfer learning [11]. The basic idea of transfer learning is that a network pre-trained on a large-scale database such as (ImageNet), called a source or teacher network, has knowledge that can be transferred to another network called a target or student network, which will be trained on a small-scale database. This knowledge is in the form of learned features.

In this work, we use two different methods of knowledge transfer, the first being feature extraction using a pre-trained network and then using these features to train an SVM classifier. The second is the finetuning technique, which consists of replacing the fully connected layer of the pre-trained network with a randomly initialized one, then training the target database and performing the target task with it. Transfer learning gave us the best results and outperformed the first two strategies with 94% precision in the texture database and 95% in the object database. More details on these two techniques are presented in the materials and methods section.

The contribution of this work is:

- We studied material recognition on Moroccan decorative patterns with two different types of input data (texture and patterns in scenes).
- We also studied the problem of lack of training data where we compare different methods and the effect of the nature of the input data.

This paper is organized as follows: first, the paper introduces the related work, then the paper presents the materials and methods used, describing the datasets used, and the theory of the three machine learning approaches used for the classification of Moroccan decorative patterns. Then, the paper presents and discusses the results obtained. Finally, the article concludes with a discussion of future work.

2 Related Work

- <u>Image classification with limited data</u>

Machine learning and deep learning algorithms are data-driven models. The essential elements of these algorithms are training and testing data. These algorithms are designed to simulate the perceptual system of humans and some mammals. The data can be thought of as experiments that the model faces to better understand the problem. while it is not certain that these models work exactly like the human eye and brain (i.e., they read and analyze data in the same way), the results obtained by the theme have become closer to human performance in a variety of computer vision tasks such as handwriting recognition [14], object recognition [15], facial recognition [16], and various applications. One of the problems facing deep learning and machine learning algorithms is data dependency. These models require a lot of training data to perform well. In many real-world applications, data collection can some-times be very expensive. Many studies have addressed this problem by looking for solutions to train machine learning and deep learning algorithms despite the lack of necessary data.

In the work [10], the authors showed that a low-complexity convolutional network (i.e., a CNN that has a small number of trainable parameters) performs better when data is limited. They also showed that standard data augmentation techniques im-prove performance.

In the work [24], the authors studied the performance of a variety of CNN architectures when the training dataset is small. They showed that the XeptionNet model has the best performance in the limited data problem. They also showed that data augmentation increases the classification accuracy, but its effect differentiates with different networks. In some CNN architectures, it increases the accuracy by a large margin, but in others, it has a marginal effect. This shows that the choice of CNN architecture plays a huge role in image classification with small data.

In our work, we also address this problem by studying the performance of image classification using small data in different scenarios. We studied the effect of the level of data complexity using two small-scale datasets. A less challenging dataset consisting of texture images where texture fills the images, and a more challenging dataset of objects in scenes where patterns are in complex scenes.

- Transfer learning

One solution for the lack of necessary data is to transfer the knowledge from one model to another. It is well known that a human being can use the prior knowledge he or she possesses to perform a new task faster or better [11] e.g., a rugby player needs less training to become an American football player than a non-athlete. Transfer learning is based on this theory to solve the problem of limited training data. A model trained on a large amount of available data has some knowledge that will help improve a model that has access to only a few data. Deep transfer learning is performed by transferring the image representation of a source model to a target model. This image representation is so powerful that it is the first thing to be tried to solve a visual recognition problem [12][13].

- Material recognition

Material and texture recognition is a very challenging topic even for the human eye. Materials often share the same visual proprieties and thus it becomes hard to distinguish between them. For that, many researchers are based on the idea that additional information about the material beside the image will improve material recognition systems. In the work [17], the authors proposed a new deep learning method for surface material recognition based on fully connected networks (FCN). They combined visual information with depth sensory information to improve the classification performance. They used two different inputs, an "acceleration signal" from a sensor that carries relevant information about the surface of the material, and an image. Their experiments showed that using both inputs is better than using only the visual or sensory input.

In the work [18], the authors proposed an approach with a machine learning algorithm that uses the information provided by a sensor system composed of a laser and a camera. They achieved an accuracy of 97% with a decision tree.

Although these systems are very accurate, human-used tools, such as smartphones, do not always have these sensor systems. Various research has been conducted on texture and material recognition using only images as inputs. In the work [19] the authors studied the performance of a set of handcrafted algorithms in parallel with some pre-trained CNN architectures for texture and material classification. They tested these algorithms on the T1K+ database which contains texture im-ages of 1129 diverse classes. They found that the features extracted by the pre-trained models far outperformed the manually created descriptors.

In the work [22], the authors showed that improved FISHER vectors computed on SIFT features can increase the performance of texture classification. On their Describable Texture Dataset (DTD) this type of algorithm reaches an accuracy of 61.5%, outperforming specialized texture descriptors. FISHER vectors were also used in the work [23], in a combination with convolutional neural networks. In this work, the authors studied texture description and material classification using images in a clutter. In their work [21], the authors combined statistical features with a CNN mod-el to improve the performance of a texture classification task. Their results on the describable texture dataset (DTD), and the flicker material dataset (FMD) showed that this combination is better than using a CNN alone. In the work [20], the authors used transfer learning to classify texture images, they used a three-step process in which a CNN is trained on a source task, then it is used to project the data into an-other feature space and finally use it to train a target model. On the Brodatz-32 texture dataset this method showed a 6% improvement in accuracy over training a CNN directly. In our work, we have studied a variety of computer vision approaches for the classification of Moroccan decorative patterns.

3 Methodology

3.1 Data

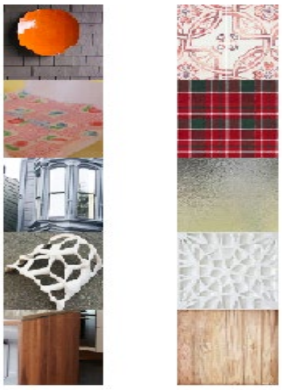

Fig. 3. example of samples used for the five models, the left column represents the samples from DATASET02 (object dataset), the right column represents the samples from DATASET01 (texture dataset).

In this work, we tested our models in two different situations: when the data are in the form of texture and when the data are images where the motif is seen in a setting surrounded by other objects. For this purpose, we created two datasets: the first one is the "DATASET 01" composed of texture images of the five elements of the Moroccan decoration. We used a hybrid data source with images from the site "motif in Islamic art" [25] and some images are taken from Moroccan salons. The second dataset is a"DATASET2" composed of hybrid images from the FMD dataset [26], from the MNIC dataset [27], and some images taken in Moroccan houses. The reason for using two data sets is that we are looking for a model that is generalizable, and gives acceptable results in different situations.

An example of images from both datasets is shown in Fig3. we can see that the DATASET02 is more difficult because the pattern has a background that can sometimes be complex, or the patterns are overlaid by other objects, and materials. in the DATASET01 dataset, the texture fills the image, that is all the pixels of the image belong to the pattern which makes it less challenging.

3.2 Handcrafted features

The handcrafted feature approach is the first strategy we used to classify the Moroccan patterns. Handcrafted features are a set of algorithms specially designed to extract certain low-level image features such as edges, color, saturation, texture, etc..... In our work, as shown in fig4, the handcrafted features, strategy is based on extracting texture and color features, and combining them into a feature vector to train a machine learning classifier. For texture features, we used the gray level co-occurrence matrix and the local binary pattern. And for color, we used the color histogram. All three techniques are briefly explained in what follows.

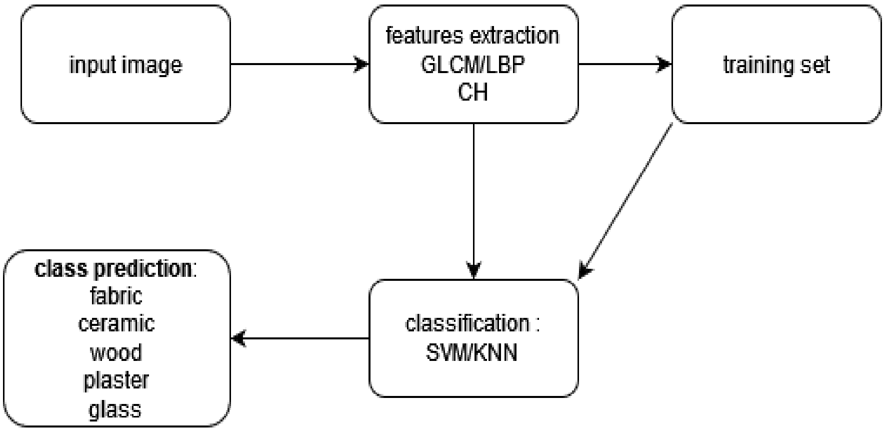

Fig. 4. proposed hand-crafted features approach for the classification of Moroccan decorative motifs

3.2.1 GLCM.

The gray level co-occurrence matrix [2] is a powerful texture descriptor first introduced in 1973 for rock classification. It is a two-step feature extraction technique. The first step is to construct the co-occurrence matrix. The elements of this matrix represent the number of times a pixel of intensity i is adjacent to a pixel of intensity j at a particular distance and orientation. The distance is fixed and the matrix is calculated in four orientations (vertical, horizontal, and two diagonals) as shown in fig5. We then obtain four matrices.

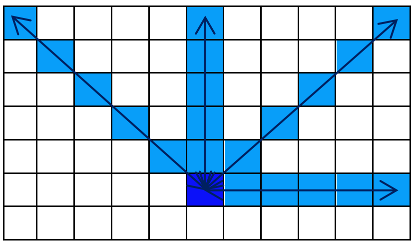

Fig. 5. directions of the GLCM

The second step is to extract the texture measures from the constructed matrices. The work [2] proposed fourteen measures that characterize the texture information of the image. Some of these measures have a direct relationship with the texture such as "correlation, homogeneity and contrast" others contain discriminative information about the texture. More details about the fourteen measures of Haralick are explained in the work [2]. For each of the four matrices constructed, the Haralick measures are calculated. We take the average and the range to finally have a vector of 28 features.

3.2.2 LBP.

The local binary model LBP [3] is a popular method for texture feature extraction. LBP is a statistical method that computes local texture information. In our work, we used the extended version of LBP developed in 2002 in the work [28]. This approach consists in detecting "uniform" patterns, which are most often included in the texture compared to other patterns. Calculating the occurrence of uniform patterns in an image gives a powerful texture descriptor.

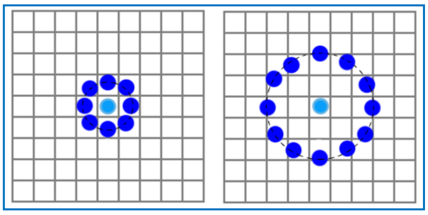

Fig. 6. an example of LBP operator for (P,r)=(8,1) and (P,r)=(16,2)

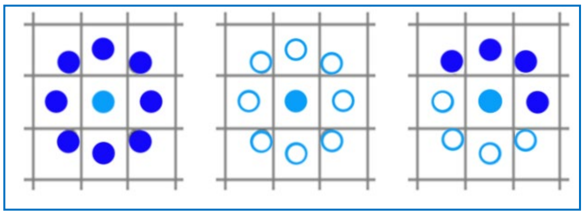

Fig. 7. Example of uniform patterns

Uniform patterns are detected using a circular window placed in the image with two parameters: the radius of the circle r and the number of pixels in the window P as shown in figure 6. The LBP calculates the difference between each pixel in the window and the central pixel and takes into account the singularity of this difference by assigning to the pixel in the window, 0 if the difference is negative and 1 if it is positive.

$$LBP_{p,r} = \begin{cases} \sum_{P=0}^{p-1} S(gp - gc) \text{ if } u(LBP_{P,r}) \leq 2 \\ P + 1 \qquad \text{otherwise,} \end{cases} \tag{1}$$

with u(LBP) is the number of transitions in the LBP.

Uniform patterns are those that have at most two 0-1 or 1-0 transitions, for example, a window of (00000000, zero transitions) and (01000000, two transitions) are uniform patterns as shown in Figure 7. but (01010000, four transitions) are not. Finally, a histogram that represents the occurrence of each of the uniform patterns forms the final feature vector. In this histogram, the non-uniform patterns are grouped into a single label.

3.2.3 Color histogram.

Texture information is an effective way to describe the content of decorative images, but it is not a robust tool for the classification of multicolor images. To increase the classification performance, we use the texture descriptor combined with the color descriptor. For color, we use the color histogram [4], which is a color feature extraction tool developed in 1991. The color histogram represents the occurrence of each discrete color in the image. We set the number of bins in the histogram to 8 and choose the HSV color system. The final size of the feature vector will be 512.

3.2 Training a CNN from scratch

In the second strategy, we explore the performance of training a convolutional neural network from scratch. The CNN represents the state of the art in image classification with architectures like AlexNet [29], VGG16[8], ResNet [9], Inception [30], ... etc. On the challenging ImageNet dataset, these architectures achieve near-human accuracy. The problem is that all these architectures have a large number of parameters that need to be trained, and they benefit from the large amount of data offered by ImageNet (1.5 images) to adjust all these parameters and reach good accuracy.

In our case, we have only 130 images per class in DATASET01 and 120 in DATASET02, which is too few to train one of these state-of-the-art architectures. For this, we propose to reduce the number of parameters by using a shallow CNN. We followed the work10 where the authors show that a shallow CNN with few parameters is more appropriate when data is limited.

After several experiments, we found that the CNN that gives the best results is a CNN with three convolutional layers with the number of filters is 16 32 64 respectively, and with a spatial max-pooling after each convolutional layer. The fully connected layer contains two layers: the first one has 16 neurons activated with the Relu [29]. The second one is the output layer with five outputs activated with SoftMax. The details of our CNN are shown in figure 8. The number of trainable parameters is (32 816), which is too little compared to some architectures like vgg16 (138 million) and ResNet (23 million).

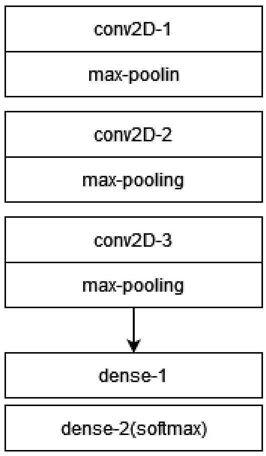

Fig. 8. CNN architecture

To increase the amount of data, we used data augmentation [31]. This technique increases the number of training images by artificially creating new versions of an original image. The motivation behind data augmentation is that the more a machine algorithm has access to data, the more it is effective. In this work, data augmentation is done by some traditional transformations such as random transitions, rotations, and scale changes. We tested the effectiveness of data augmentation in both datasets, and the results are reported in the section of experiments and results

3.3 Transfer learning

One of the well-known solutions to train a CNN with limited data is transfer learning [11][12][13] It is a technique that consists in using knowledge from another network to improve performance. In transfer learning, the learning model from one task is reused as a starting point for a second task [36].

Consider a source model Ms trained on a source domain Ds and a target model Md with a target domain Dt, transfer learning aims to improve the performance of Mt using the knowledge gained from training Ms on Ds. This is based on the assumption that if Dt and Ds have some similarity, Ms will have useful knowledge for Mt. In our case "Moroccan decorative image classification" the source model is a state-of-the-art architecture; we compare between two architectures VGG16[8] and ResNet [9].

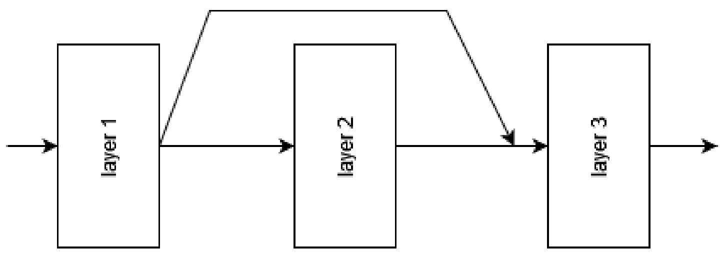

Fig. 9. An example of a shortcut connection used by the ResNet architecture.

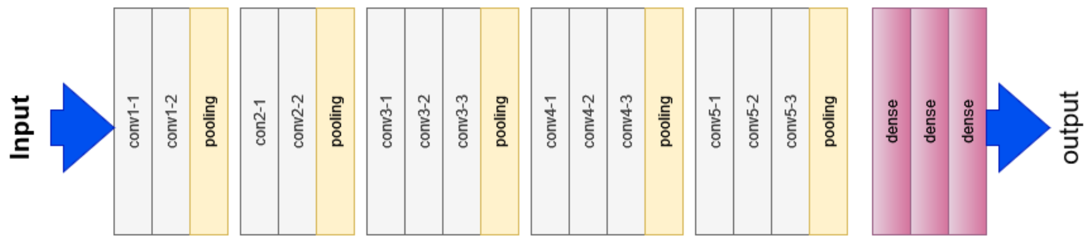

Fig. 10. vgg16 architecture

Vgg16: is a CNN architecture first introduced in 2014 in the work [8] the main idea behind this architecture is that increasing the depth of the network will improve performance. For this, small size filters (3x3) are used in the convolutional layers, to increase the depth while maintaining a small number of parameters. vgg16 architecture is shown in fig10.

ResNet: increasing the depth of the network improves performance. But at some point, in the process of adding layers, the network suffers from an accuracy degradation problem. the ResNet [9], residual neural network came to solve this problem. These networks use shortcut connections that cause the layers to learn the residual mapping instead of the underlying mapping, as shown in fig9.

In our case, Ds is ImageNet [32] a large-scale dataset. The reason why ImageNet is very popular in transfer learning tasks is that it contains many training images that will be sufficient to train a very deep network like VGG16 and ResNet. It is also a difficult dataset because it contains 1000 classes divided into fine-grained classes (different breeds of dogs for example) and coarse-grained classes (cats, dogs). Many schools of thought on why ImageNet performs well in transfer learning tasks have been raised, however, the study done by [33] has answered some key questions on this topic.

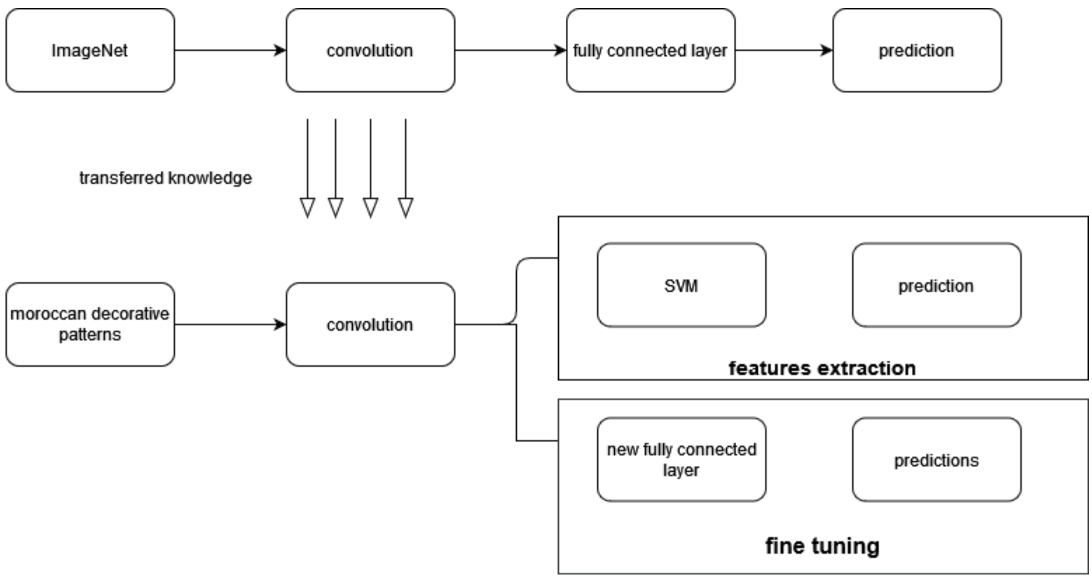

Fig. 11. transfer learning approach

the target domain Dt is our dataset DATASET01 and DATASET02 on which we train our Mt model to classify Moroccan decorative patterns.

In this work, as shown in Figure 11, we used two types of transfer learning: feature extraction and fine tuning. Feature extraction is the process of using the convolutional part of a pre-trained model to extract features from the image. These features are used to train a machine learning classifier. Fine-tuning involves freezing the first layers of the pre-trained model, which extract general features [35], and updating the last layers during training on the target dataset.

For the fine-tuning process, we have frozen all convolutional layers and replace the fully connected layers with new randomly initialized layers that will be updated during training.

4 Results, Analysis and Discussions

In this section, we present the results and analysis of the three strategies used in this work on two Moroccan decorative pattern datasets. To evaluate the classification performance, we divided our database into 80% training images and 20% test images. We then extracted the precision of the five classes and other evaluation measures.

4.1 hand crafted features

We combined the GLCM and LBP with the color histogram to extract features that will be used to train an SVM and KNN classifier. Some of the main results of our experiment are shown in Table 1. The performance of the SVM classifier is better than KNN in both datasets and algorithms used. In the case of using the local binary pattern combined with the color histogram, the classification accuracy, using SVM, exceeds that of KNN by a significant margin. In general, the combination of texture and color features is better in DATASET01 and less in DATASET02. This result can be explained by the nature of these algorithms, they are not invariant to the position of the pattern in the image. We believe that a prior segmentation process of the pat-tern before applying the algorithm will give good results.

Table 1. precision of the hand-crafted approaches

dataset	*Features extraction*	*SVM*	*KNN*
DATASET01	*GLCM+CH*	*77%*	*71%*
	LBP+CH	*77%*	*58%*
DATASET02	*GLCM+CH*	*54%*	*48%*
	LBP+CH	*52%*	*50%*

4.2 Training a CNN from scratch

Table 2. The performance of training a CNN from scratch.

dataset	*Data augmentation?*	*SVM*
DATASET01	*yes*	*67%*
	no	*75%*
DATASET02	*yes*	*61%*
	no	*58%*

We used a shallow CNN to classify decorative patterns as a solution to the small amount of data available. To explore the performance of the CNN in different situations, we used two datasets, one containing texture images (DATASET01) and the other containing the five patterns in complex scenes (DATASET02).

The results of the proposed CNN on the two datasets are shown in table2. The performance of a shallow CNN on the texture data is generally good with an accuracy of 80%, however, the accuracy on DATASET02 is lower with only 61%. The shallow CNN shows the same behavior as the handcrafted approach where the classification performance of less difficult texture images is better than the classification performance of more complex images. This behavior can be explained by the fact that color and texture information is captured in the first convolutional layers of the convolutional neural network [35]. As we know, color and texture are the basic features of the texture data used in this work. On the other hand, the performance of object classification is inferior to that of texture, which is due to the challenge of this dataset where the patterns are in complex scenes and requires a deep CNN with many layers to represent the information well. In general, the performance of CNN is better than that of hand-crafted features on both datasets despite the lack of data.

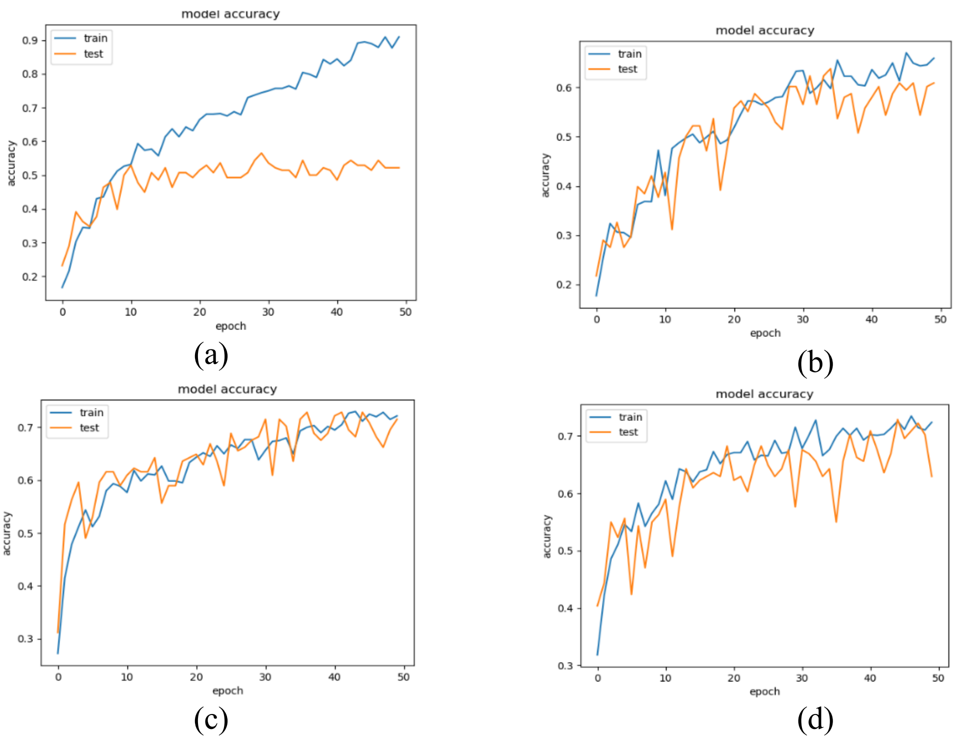

Fig. 12. differences in accuracies curves of the CNN (a): on DATASET02 before data augmentation (b): on DATASET after data augmentation, (c): on DATASET01 before data augmentation, (d): on DATASET01 after data augmentation

To increase the amount of data, we used data augmentation. The difference that data augmentation techniques make is illustrated in Table 2 and Figure 12. In DATASET01, traditional data augmentation does not have a positive effect on CNN performance, and it decreases the accuracy by 8%. This can be explained by the nature of the data where performing random geometric transformations produces similar copies of the original image, which does not help the network generalize better. Besides the fact that in this situation data augmentation is not necessary, the model already has good performance without data augmentation. On the other hand, in-creasing the data increases the classification accuracy on DATASET02 by 3%. Data augmentation also reduces the overfitting between training and validation, as shown in figure11, and makes the model more generalizable.

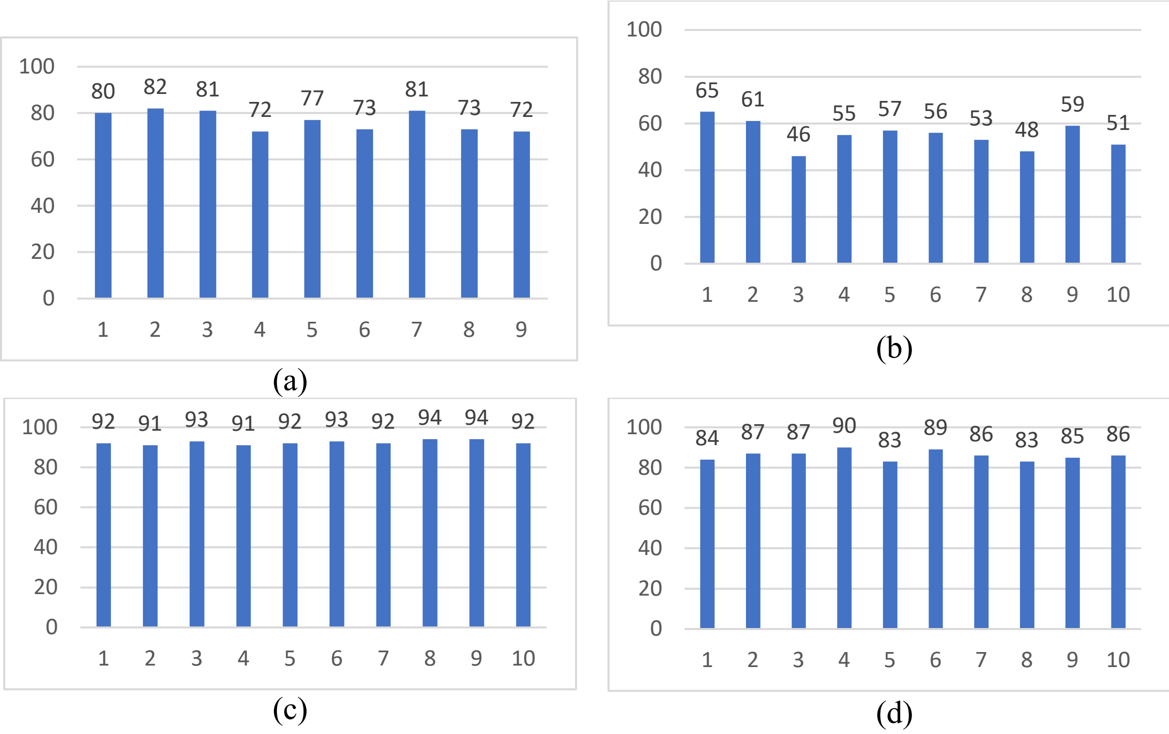

Fig. 13. variance of the mAP of, (a): CNN from scratch on DATASET01, (b): CNN from scratch on DATASET02, (c): fine-tuning on DATASET01, (d): fine-tuning on DATASET02.

One of the observations we made is that the accuracy differs each time we run the network. This variance can be explained by the stochastic nature of CNN, where the initial weights and biases are chosen stochastically, meaning that CNN has a different starting point at each run. Another key contributor to this variance is the use of a random train-test split to evaluate our model. We run each CNN ten times and report the accuracy as shown in Figure 13. The figure shows that the accuracy of the net-work ranges from 72% to 82% on DATASET01 and 46% to 65% on DATASET02. The variance of the classification accuracy using the texture dataset is lower than that of the object dataset, meaning that the level of challenge represented by the dataset affects this variance.

4.3 Transfer learning

We use two transfer learning techniques: the first is the feature extraction method, where a pre-trained CNN (VGG16 and ResNet50) is used to extract features that will be used to train an SVM model. The second is fine-tuning, where we replace the fully connected layer of the pre-trained models with new randomly initialized layers, and train it for a few epochs. The new classifier consists of four layers with 128, 64, 32 and 5 filters respectively. The first three layers are activated with the Relu activation function, and the last one is activated with SoftMax. the network is trained for 30 epochs with the Adam optimizer, the learning rate is decreasing with training. The results are presented in Table 3. Transfer learning outperforms all previous methods by a significant margin. On DATASET01, the accuracy of ResNet50 is slightly better than VGG16 by 3% using feature extraction, and by 1% using fine tuning. On DATASET02, ResNet performs better using both methods, by 5% for feature extraction and 9% for fine tuning.

Table 3. Precision of transfer learning approach

dataset	*Source model*	*Features extraction*	*Fine tuning*
DATASET01	*VGG16*	*91%*	*93%*
	ResNet50	*94%*	*94%*
DATASET02	*VGG16*	*90%*	*86%*
	ResNet50	*95%*	*95%*

The first thing we noticed is that the performance of transfer learning methods is good on both datasets, unlike the previous strategies where the performance is only good in the texture dataset. The pre-trained model ResNet did a better job than VGG16, which raises the hypothesis that when comparing two pre-trained models, the one that performed better in the source task will perform better in the target task (ResNet performs better than VGG in ImageNet). In general, transfer learning far outperforms the handcrafted feature approach and training CNNs from scratch. This means that the features provided by the transferred layers are very powerful and discriminative.

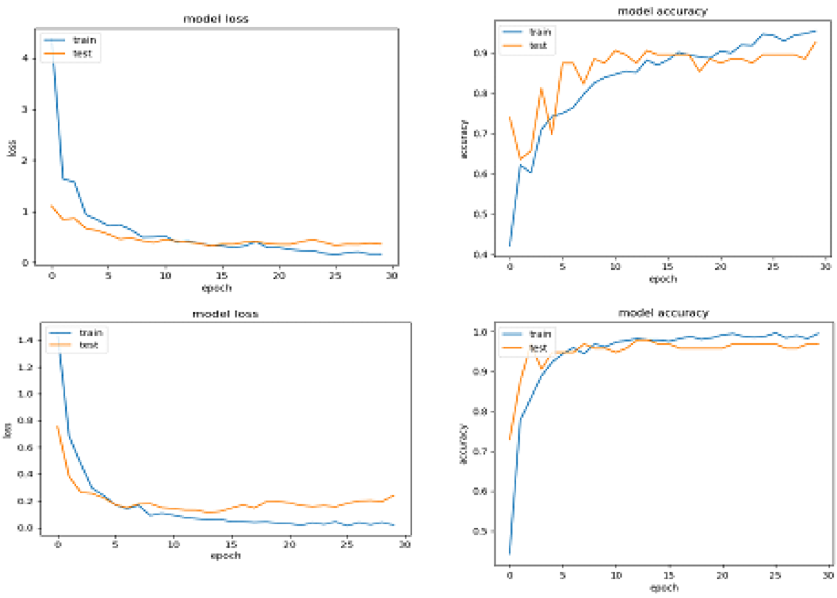

Fig. 14. accuracy and loss curves of the fine tuning of the pre-trained model vgg16 the upper curves are for the training on DATASET01, the lower curves are for DATASET02

The accuracy and loss curves in Figure 14 show that the network behaves differently if it has been provided with knowledge about image classification than if it al-lows the knowledge to be learned by itself. The validation and learning accuracies converge faster in a few epochs, the curves do not show too much oscillation, and the overfitting is low. Moreover, the features extracted using the convolutional part that learned the optimal image representation from ImageNet are more generalizable to the two datasets used than the handcrafted features.

As we did for training from scratch, we tested the variance of the fine-tuning by running a model ten times. The model we tested is created by placing a new fully connected layer on top of the pre-trained VGG16 convolutional layers. We report the variance of the model on both datasets in Figure 13. The variance on DATASET01 is limited compared to training a CNN from scratch, the accuracy ranges from 91% to 94%. This shows that the model is not greatly affected by the stochastic nature of the CNN and the random division of the dataset when trained on texture images. Even if the model starts from a different point, in the end it will produce the same predictions. On the other hand, the variance on DATASET02 is larger, but still less than training a CNN from scratch. The variance measures the quality of the machine learning model and shows that the model predictions are stable when the network has some knowledge of the task, as in the case of ImageNet image representation transfer.

5 Conclusion

The classification of Moroccan decorative patterns is very important to better under-stand this art and to distinguish the materials that make up these patterns. In this paper, we used three computer vision approaches and compared them. We evaluated their performance on two small-scale datasets with different levels of challenge. The first dataset contains texture images, and the other is more challenging, with patterns in complex scenes. The three strategies we used are handcrafted features, training a convolutional neural network from scratch, and transfer learning. The best results are obtained by transfer learning, with 95% accuracy using a pre-trained ResNet model. Transfer learning is also generalizable, i.e., it is not affected by the difficulty of the dataset, compared to hand-created features and convolutional neural network learning from scratch, where the performance on texture data is much better than on the more complex data. In future studies, we will make further improvements that may increase the performance of the hand-created feature approach and training a CNN from scratch.

ACKNOWLEDGEMENTS
I would like to express my gratitude to Abdellah LAZZER and Rachida ASWAB, my colleagues at Abdelmalek Essadi University, who revised my texts and answered with tireless patience many questions concerning the writing of an article on computer vision.

References

[1] Aoulalay, A., Makhfi, N.E., Abounaima, M.C., & Massar, M. (2020). Classification of Moroccan Decorative Patterns Based on Machine Learning Algorithms. 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), 1-7.

[2] Haralick, R.M., Shanmugam, K.S., & Dinstein, I. (1973). Textural Features for Image Classification. IEEE Trans. Syst. Man Cybern., 3, 610-621.

[3] Ojala, T., Pietikäinen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. Pattern Recognit., 29, 51-59.

[4] Swain, M.J., & Ballard, D.H. (2004). Color indexing. International Journal of Computer Vision, 7, 11-32.

[5] BurgesChristopher, J.C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery.

[6] Cover, T.M., & Hart, P.E. (1967). Nearest neighbor pattern classification. IEEE Trans. Inf. Theory, 13, 21-27.

[7] Breiman, L. (2004). Random Forests. Machine Learning, 45, 5-32.

[8] Liu, S., & Deng, W. (2015). Very deep convolutional neural network-based image classification using small training sample size. 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), 730-734.

[9] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778.

[10] Chandrarathne, G., Thanikasalam, K., & Pinidiyaarachchi, A. (2020). A comprehensive study on deep image classification with small datasets. In Advances in Electronics Engineering (pp. 93-106). Springer, Singapore.

[11] Weiss, K.R., Khoshgoftaar, T.M., & Wang, D. (2016). A survey of transfer learning. Journal of Big Data, 3, 1-40.

[12] Azizpour, H., Razavian, A.S., Sullivan, J., Maki, A., & Carlsson, S. (2015). From generic to specific deep representations for visual recognition. 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 36-45.

[13] Razavian, A.S., Azizpour, H., Sullivan, J., & Carlsson, S. (2014). CNN Features Off-the-Shelf: An Astounding Baseline for Recognition. 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, 512-519.

[14] Elmansouri, M., Makhfi, N.E., & Aghoutane, B. (2020). Toward Classification of Arabic Manuscripts Words Based on the Deep Convolutional Neural Networks. 2020 International Conference on Intelligent Systems and Computer Vision (ISCV), 1-5.

[15] Bui, H.M., Lech, M., Cheng, E., Neville, K.L., & Burnett, I.S. (2016). Object Recognition Using Deep Convolutional Features Transformed by a Recursive Network Structure. IEEE Access, 4, 10059-10066.

[16] Singh N.S., Hariharan S., Gupta M. (2020) Facial Recognition Using Deep Learning. In: Jain V., Chaudhary G., Taplamacioglu M., Agarwal M. (eds) Advances in Data Sciences, Security and Applications. Lecture Notes in Electrical Engineering, vol 612. Springer, Singapore. https://doi.org/10.1007/978-981-15-0372-6_30.

[17] Zheng, H., Fang, L., Ji, M., Strese, M., Ozer, Y., & Steinbach, E.G. (2016). Deep Learning for Surface Material Classification Using Haptic and Visual Information. IEEE Transactions on Multimedia, 18, 2407-2416.

[18] Tamas Aujeszky, Georgios Korres & Mohamad Eid (2019) Material classification with laser thermography and machine learning, Quantitative InfraRed Thermography Journal, 16:2, 181-202, DOI: 10.1080/17686733.2018.1539895

[19] Cusano, C., Napoletano, P., & Schettini, R. (2021). T1K+: A Database for Benchmarking Color Texture Classification and Retrieval Methods. Sensors (Basel, Switzerland), 21.

[20] Hafemann, L.G., Oliveira, L., Cavalin, P.R., & Sabourin, R. (2015). Transfer learning between texture classification tasks using Convolutional Neural Networks. 2015 International Joint Conference on Neural Networks (IJCNN), 1-7.

[21] Jbene, M., Maliani, A.D., & Hassouni, M.E. (2019). Fusion of Convolutional Neural Network and Statistical Features for Texture classification. 2019 International Conference on Wireless Networks and Mobile Communications (WINCOM), 1-4.

[22] Cimpoi, M., Maji, S., Kokkinos, I., Mohamed, S., & Vedaldi, A. (2014). Describing Textures in the Wild. 2014 IEEE Conference on Computer Vision and Pattern Recognition, 3606-3613.

[23] Cimpoi, M., Maji, S., & Vedaldi, A. (2015). Deep filter banks for texture recognition and segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3828-3836.

[24] Zhao, P., Li, C., Rahaman, M.M., Yang, H., Jiang, T., & Grzegorzek, M. (2021). A Comparison of Deep Learning Classification Methods on Small-scale Image Data set: from Convolutional Neural Networks to Visual Transformers. ArXiv, abs/2107.07699.

[25] https://patterninislamicart.com/

[26] Sharan, L., Rosenholtz, R., & Adelson, E.H. (2010). Material perception: What can you see in a brief glance? Journal of Vision, 9, 784-784.

[27] Bell, S., Upchurch, P., Snavely, N., & Bala, K. (2015). Material recognition in the wild with the Materials in Context Database. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3479-3487.

[28] Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002). Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. IEEE Trans. Pattern Anal. Mach. Intell., 24, 971-987.

[29] Krizhevsky, A., Sutskever, I., & Hinton, G.E. (2012). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60, 84 - 90.

[30] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S.E., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1-9.

[31] Shorten, C., & Khoshgoftaar, T.M. (2019). A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6, 1-48.

[32] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248–255).

[33] Huh, M., Agrawal, P., & Efros, A.A. (2016). What makes ImageNet good for transfer learning? ArXiv, abs/1608.08614.

[34] Zeiler, M. D., & Fergus, R. (2014, September). Visualizing and understanding convolutional networks. In European conference on computer vision (pp. 818-833). Springer, Cham.

[35] Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? ArXiv, abs/1411.1792.

[36] Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A Survey on Deep Transfer Learning. ArXiv, abs/1808.01974.

Notes on contributors

Ayoub AOULALAY is a PhD student at Faculty of Science, Abdelmalek Essaadi University, Tetouan, Morocco. His main research interests include: Texture recognition, Pattern recognition, Image classification, Machine Learning, Deep Learning. He has published several research conferences articles in international conferences.

Abderrahim EL MHOUTI received the PhD degree in Computer Science in 2015 from Abdelmalek Essaadi University, Tetouan, Morocco. Mr. EL MHOUTI is Professor of Computer Science at the Faculty of Sciences at same University. His research fields include: Educational Big Data and Learning Analytics, Cloud Computing, Educational Technologies, Machine Learning and Deep Learning. He has published several articles in his areas of research.

Mohammed MASSAR began his academic career at the Faculty of Sciences of Oujda, Morocco. In 2008, he obtained his Master's degree, and he defended his PhD in Applied Mathematics in 2012. Actually, he is associate professor of mathematics at Faculty of Science and Techniques of Al-Hoceima at same University. His research fields include: Applied Mathematics, Nonlinear Analysis, Partial Differential Equations, Analysis and Functional Analysis. He has published several articles in his areas of research.

Yassine EL BORJI is currently Professor at the National School of Applied Sciences of Al-Hoceima at Abdelmalek Essaadi University. He works in the field of applied engineering, didactics and educational technology. His main research interests are serious games, gamification and mixed reality. Due to the continuous dialogue with these disciplines, he is constantly involved in related fields, such as simulations, artificial intelligence, Learning Analytics, integrated data models and data exchange formats.

Mohamed Fahim: Received the Ph.D. degree in Computer Science from Moulay Ismail University, Meknes, Morocco. Mr. Fahim is a Professor of Computer Science at the Faculty of Sciences and Technologies of Al-Hoceima belonging to Abdelmalek Essaadi University. His research fields include technology-enhanced learning, educational technologies, machine learning, and deep learning. He has published several articles in his areas of research.