UNIVERSITY KASDI MERBAH OUARGLA

Faculty of New Technologies of Information and Communication

Department of computer science and Information Technology

MASTER Thesis

Domain: Mathematics and Computer science

Specialty: Industrial

Presented by:

Nermine Dina Zekri

Khaoula Yazi

Title:

Colored multi-scale texture classification using morphological mathematic operations and CNN

Publicly discussed:

14/10/2020

Before the jury:

Dr. Khadidja Amor	MCB	President	UKM Ouargla
Dr. Belal KHALDI	MCB	Supervisor	UKM Ouargla
Dr. Oussama AIADI	MCB	Examiner	UKM Ouargla

Academic year: 2019-2020

Acknowledgment:

First and for most ,we have to thank GOD for providing us with everything that we required in completing this project and guided us always to the right path our life .

We would like to sincerely thank our parents and families for their love and support during our carrier since the childhood tell now.

Endless thanking for our supervisor Dr: Khaldi Belal who did not keep any effort in encouraging us to do a great job, and providing us with information and advices to be better each time.

We have to thank the jury members starting from the president Dr:Khadidja AMER and examining board Dr:Oussama AIADI .

We would like to express our gratitude for everyone who helped us during the graduation project: sisters, brothers, friends and classmates.

In addition, our gratitude to all the professors of department of computer science, university KASEDI MERBAH OUARGLA, and my colleges promoting 2019-2020 master computer science.

Abstract :

In recent years, digital image processing became a rapidly growing research area of computer science. Its multidiscipline uses make it a center of interest of many researchers in order to achieve better understand of the image's content. Texture is considered as one of the fundamental descriptors of the image besides color and shape. Different methods have been proposed to analyse and represent texture. Regardless their efficiency, lot of these methods are not able to extract all the features distributed in different scales of interest. In this work, we address this limitation by proposing an improved method based on morphological mathematic operations and CNN classifier to represent different scale of interest. First, we apply on the input color image a series of opening and closing morphology operations to generate different scales of interest presented as new images. The output will then be fit to different models of CNN for purpose of extracting features and classify them. Our method aims to extract both color and multi-scale information from texture images. The obtained results from KTH_TIPS dataset were about (84% in our proposal model, 85% on the VGG model, and 56.54% on the RES-net model). The results were satisfying and promising compared to methods that target only one scale of interest.

Key words: multi-scale, texture representation, CNN, morphological operations, image classification.

Résumé :

Ces dernières années, le traitement d'images numériques est devenu un domaine de l'informatique en pleine croissance. Ses usages pluridisciplinaires en font un centre d'intérêt de nombreux chercheurs afin de mieux comprendre le contenu de l'image.

La texture consiste en l'un des descripteurs fondamentaux de l'image à côté de la couleur et de la forme. Différentes méthodes sont proposées pour évaluer les analyses de texture, bien que leurs lots efficaces ne soient pas en mesure d'extraire toutes les caractéristiques réparties à différentes régions d'intirest dans le patron de l'image. Par ce travail, nous abordons cette limitation en proposant une méthode améliorée basée sur des opérations mathématiques morphologiques et le classificateur CNN.

D'abord, nous appliquons sur les images colorées une série des opérations d'ouverture et de fermeture pour générer différentes échelles et régions d'intérêt présentées comme de nouvelles images. Que seront intégré à différents modèles de CNN dans le but d'extraire les caractéristiques et de classer la texture. Cette méthode vise à extraire à la fois des informations de couleur et multi-scale. Les résultats obtenus après son application sur KTH_TIPS étaient d'environ (84% dans notre méthode de proposition, 85% sur le modèle VGG et 56.54 sur le modèle RES-net) .Les résultats étaient satisfaits et prometteurs donc la méthode proposée est robuste et efficace pour classer les couleurs textures multi-scale.

Mots clés : multi-échelle, texture, CNN, opération morphologique, image.

ملخص:

في السنوات الأخيرة، أصبحت معالجة الصور الرقمية مجالًا سريع النمو لعلوم الكمبيوتر. تجعله استخداماته متعددة التخصصات مركز اهتمام العديد من الباحثين من أجل تحقيق فهم أفضل لمحتوى الصورة.

يعتبر النسيج كواحد من الواصفات الأساسية للصورة بجانب اللون والشكل. تم اقتراح طرق مختلفة لتقييم تحليلات النسيج، وعلى الرغم من فعاليتها، الا ان الكثير منها غير قادر على استخراج جميع الميزات الموزعة في مقاييس مختلفة في نمط الصورة. في هذا العمل نعالج المشكل المطروح من خلال اقتراح طريقة محسنة تعتمد على العمليات الحسابية المورفولوجية ومصنف CNN.

نطبق أولاً على الصورة الملونة المدخلة عدة عمليات فتح وإغلاق لتوليد مقاييس مختلفة ومناطق الاهتمام لتعرض كصور جديدة. هاته الأخيرة تصبح مدخلات لنماذج مختلفة من CNN بغرض استخراج الميزات وتصنيف النسيج. تهدف هذه الطريقة إلى استخراج المعلومات الملونة ومعلومات تخص النسيج في نطاقات متعددة معا كانت النتائج التي تم الحصول عليها بعد تطبيقه على KTH_TIPS حوالي (84٪ في نموذجنا المقترح، 85٪ على نموذج VGG و56,54 ٪ على نموذج RES-net) النتائج كانت مرضية وواعدة وبالتالي فإن الطريقة المقترحة قوية وفعالة في تصنيف الصور الملونة متعددة النطاقات.

الكلمات المفتاحية: متعدد النطاقات , الملمس الشبكة, العصبية للالتفاف ,العملية المورفولوجيا , صورة.

Contents

List of tables

List of figures

List of abbreviation

CNN: Convolution Neural Network.

MMO: Mathematic morphological operations.

RGB: (Red, Green, Bleu).

HVS: (Hue, Saturation, Value).

YCBCR: (Y)green,CBlue,CRed (digital video color space)

CT: Computerized Tomography.

MRI: Magnetic Resonance Imagining.

2,3D: 2, 3 Dimension.

GLCM: gray-Level Co-occurrence Matrix.

LBP : Local Binary Pattern.

MRIR: Multi-scale Rotation-Invariant Representation

MP : Mapping Pattern.

MRCNN: Multi-scale Rotation-invariant Convolution Neural Network.

MIAS: Mammographic Image Analysis Society.

SOI: Scale Of Interest.

SE: Structuring Elements.

FC : Fully Connected.

General introduction

The advanced technologies nowadays make the way of getting a digital image so easy and available to everyone, from our familiar and variant environment such as (people, photographs of landscapes, computer-generated drawings and paintings, images from medical radiology, satellite images, and so on) to become a type of information exchange. This causes a rapid accumulation and huge volume of data collections daily, which triggered a need for efficient and effective schemes to facilities it's classification.

Texture is one of the image's descriptors witch acquired the attention of many researchers for years, to extract the characteristics and features that cannot be observed using only human's eyes.

Different researchers have attempted to characterize and define the texture additionally different methods and descriptors are proposed to evaluate the analysis of texture since 1980. Every work had provided their own framework for defining and characterizing the notion of texture. Although of the efficient and success of many approaches in extracting texture feature from images, each of these approaches has their own advantages and disadvantages that determine their applicability, suitability rapidity and performance. In other side a lot of them are unable to pick up all the features distributed in different scale of interest ,especially that Previous works argued that the texture differs from a scale to another, which causes different extracted features from it too. Besides, that most of these approaches are firstly designed for dealing with gray-scale images.

The convolutional neural network is an efficient classification method witch known as one of the successful methods that provide high accuracy in extracting features from the pattran of texture using the different filters However, it suffers from a limitation that it is the long time in training spatially using deep training .

Multi-scale mathematical morphology is a powerful concept that can explore the multi-scale information from the same image, and multi-scale mathematical morphology have potential to effectively increase the classification accuracy for texture images they were and still a useful tool to represent and describe the image and its features due to its efficacy and robustness against noisy and intensity changes .

Thus, the collective use of the two concepts CNN and multi-scale mathematical morphology have potential to effectively increase the classification accuracy for texture images.

This is what motivated us to add a contribution to the digital image's classification using the multi-scale texture besides the color information using the efficient proprieties of morphological mathematic operation and the robustness of the CNN method.
This work is divided into **4** chapters starting from the **background chapter** which introduces several notions related to the image, and image processing as a primary chapter, passing to **the related work** that summarizes the previous works that done in this topic with their different approaches and methods, then the **proposal methods chapter** witch explain the morphological mathematic operations and CNN classifier and how they work, finally, the **discussion and experimental results chapter** that shows how the combination between the methods work, the results, and the accuracy of the classification.

Chapter2:

Background

This chapter provides some background concepts about the image, image processing and its applications, feature extraction and its descriptors in addition to texture classification.

1. Image and image processing:

Image technology is a rapidly developing, and a multidisciplinary application that uses the knowledge from different fields to fulfill various tasks.

1.1.Digital Image:

An image is defined as a two-dimensional function F(x,y) where x and y are spatial coordinates, and the amplitude of F at any pair of coordinates (x,y) is called the intensity or gray level of that image at that point[2]. When x, y and amplitude values of F are finite, we call it a digital image. In other words, an image can be defined by a two-dimensional array specifically arranged in rows and columns. Digital Images are composed of a finite number of elements, each of which have a particular value at a particular location. These elements are referred to as picture elements or image elements [2]. For each pixel, the imaging device records a number, or a small set of numbers, that describe some property of this pixel, such as its brightness (the intensity of the light) or its color.

There are three main types of images that we may apply processing on:

*. **Binary image**: The binary image as the name suggests, contains only two pixel elements i.e. 0 & 1, where 0 refers to black and 1 refers to white. This image is also known as Monochrome.

*** Gray-level format**: It is the most famous image format. It has 256 different shades of colors and it is commonly known as Grayscale Image. In this format, 0 stands for Black, and 255 stands for white, and 127 stands for gray.

*. **Color format**: can be modeled as three-band monochrome image data, where each band of data corresponds to a different color. Typical, color images are represented as red, green and blue channels. It may be represented in other formats such HVS, YCbCr, ... etc. [2]

1.2.Digital Image processing:

Digital image processing consists of the manipulation and analyses of images using digital computers. It is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it[3]. It is a type of signal processing in which the input is an image and output may be image or characteristics/features associated with that image.

1.3.Origin:

One of the first applications of digital images was in the newspaper industry when pictures were first sent by submarine cable between London and New York. The introduction of the Bart lane cable picture transmission system in the early 1920s reduced the time required to transport a picture across the Atlantic for more than a week to less than 3 hours. Specialized printing equipment coded pictures for cable transmission and reconstructed them at the receiving end. [2]. Figure'2' was transmitted in the way and reproduced on the telegraph printer fitted with typefaces simulating a halftone pattern.

Figure 2*: digital picture in1921 from a coded type by telegraph printer with special type faces. [2]*

1.4.The applications:

Nowadays, we can say that no domain that is not impacted in some way by digital image processing regardless of their different fields, why not and it has become a necessity for human‘s daily life. Effectively, today it is among rapidly growing technologies that considering as one of the important research areas within engineering and computer science disciplines too.

***Face Detection:** face detection can be treated as a specific case of object class detection. The objective of face detection is to find the specified features such locations and sizes of a known number of faces.

***Digital Video Processing:** video processing techniques are used in television sets, VCRs, DVDs, video codec, video players and other devices.

***Remote Sensing:** remote sensing is collecting different data signals using variety of devices for gathering information on a given object or area depend on electromagnetic radiation, force fields, or acoustic energy that can be detected by cameras, radiometers, lasers, radar systems, sonar seismographs, thermal meters, etc..

*** Medical and Biomedical Image Enhancement & Analysis:** It is a technology that can be used to generate images of a human body (or part of it) to be processed or analyzed by experts (Ultrasonic, X-ray, Computerized Tomography (CT) and Magnetic Resonance Imaging (MRI).

*** Biometric Verification:** it refers to the automatic identification, verification or recognition of humans by their behaviors or characteristics.

Underwater Image Restoration & Enhancement: underwater images suffer from different problems such as blurring, non-uniform lightening, noise, low contrast. That is why underwater images restoration is an essential area for research to improve the smoothening and the quality and reduce the noise.

*** Character Recognition:** Either is known as optical character recognition. It serves as an effective electronic translator of handwritten images or printed text (usually captured by a scanner) into machine editable text.

*** Criminology / Forensics:** images can include accidents and crime scenes, items of evidences in a laboratory, victims or suspects that makes the investigators and forensic scientists to develop meaningful information about the crime.

***transportation:** of the technological progresses is the design of automatically driven vehicles, also applied in traffic control and transportation planning.

***Satellite Imaging:** to extract useful information from images captured using satellite such as weather and environmental conditions monitoring. [3]

2. Feature extraction:

Feature extraction is a fundamental step where the input data transform into a set of useful features. It reduces the amount of data that must be processed, while still accurately and completely describing the original data set.

A feature, or descriptor, is defined as a function of one or more measurements, specifying some quantifiable property (i.e.: color, texture, or shape) of the whole image or sub-image or of single object. [4]. Its goal is to extract useful characteristics and capture the essential data subset. It also reduces the number of resources needed for processing without losing important or relevant information. "The reduction of the data and the machine's efforts in building variable combinations (features) facilitate the speed of learning and generalization steps in the machine learning process" [5]. Feature extraction in image processing can be used to detect features such as shaped, edges, or motion in a digital image or video. Figure '3' present a simple operation of extracting features and how it captures specific characteristics from the image.

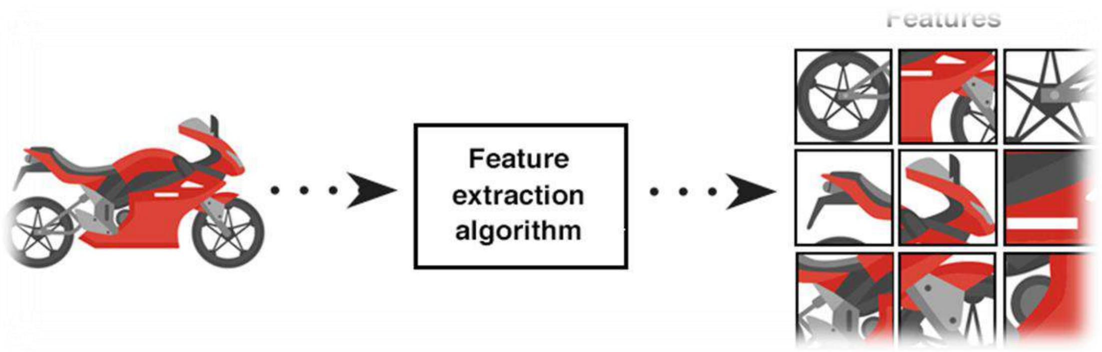


Figure3: an example that clarifies feature extraction process [6]

Feature extraction can often be divided into two independent steps:
Feature detection and **Feature description**. The main objective of a feature detector is to find a set of stable (invariant) distinctive interest points or regions, while the description encodes information in spatial neighborhoods of the determined regions mathematically. Features may describe either shape, color, texture or other information.

Descriptors are mathematical functions, which are applied to an image and produce numerical values, which are representative of a particular characteristic of the image. These numerical values can then be processed in order to provide some information about the image**. [7]**

Features descriptors must have some essential properties such as:

- **Identifiability:** shapes, which are found perceptually similar to humans, have the same features that are different from the others.
- They need to be stable under changing viewing conditions, such as changes in * Illumination,*shading, *highlights and *rotation
- **Noise resistance**: features are better to be robust against noise, clutter and occlusions.
- **Statistically independent**: two features must be statistically independent. This represents the compactness of the representation.
- **Reliability:** as long as one deals with the same pattern, the extracted features must remain the same.[8]

2.1.Shape:

Shape can be defined as the feature surface configuration of an outline, object or contour. It is used to separate objects from the background and surroundings by its outline representation. Figure '4' shows different possibilities of presenting the shape of a sphere in 3 dimensions. Despite the same shape, the features differ from presentation to another.

- Examples of shape descriptors: Fourier descriptors, Curvature Scale Space descriptors, Angular Radial Transform, 2D and 3D descriptors...etc. [6]

Figure 4: sphere shape with different possible representations on 3D witch assure different features extracted from each shape [9]

2.2. Color :

It is one of the most important low-level visual features that describe how the intensities of Red, Green and Blue channels (commonly referred to as RGB color space) of a pixel in an image under illumination change. Compared with shape and texture feature, color shows better stability and is more insensitive to the rotation and zoom of the image [9]. It does not add beauty to objects only but also adds more information; due to this fact, color features contain a significantly larger amount of discriminative information. [10]. Examples of color features: Dominant Color, Scalable Color, Color Structure, Color Layout…etc.[6]

2.3.Texture:

Texture features is the collective behavior of a group of pixels arising as a result of physical surface properties or orientation of the constituent particles which in turn give differences in reflective power, thereby color differences. [14]

There are two main approaches to define exactly what texture is:

*Structural: The texture consists of texture primitives or texture elements named “texels” occurring in some regular or repeated pattern in an image or selected region of an image.

*Statistical: texture is a quantitative measure of the arrangement of intensities in a region; it is more general and easier to compute and is used more often in practice. [11]

- figure ‘5’ shows the approaches that define texture where ‘A’ represents the statistical texture while ‘B’ represents the structural texture.

A

B

Figure 5: presentation of statistical and structural texture from (google images).

Texture has many properties such as:

– Texture is constituted by a large number of objects, which are not perceived individually.

_ Texture cannot be defined for a point.

– Texture is a regional attribute. In an image, it is defined concerning regional neighborhood in contrary to this texture of a single primitive is undefined.

– Texture within an image can be observed at various scales or levels of resolution.

– Texture arises because of the spatial distribution of intensity values or variation of colors.

– Texture may arise because of the reflective power of the imaged surface.

– The terms uniformity, density, coarseness, roughness, regularity, linearity, directionality, frequency and phase are associated with the textures. [12]

3. Image classification:

Classification is one of the most active research topic and application areas of image processing consequently a number of texture analysis methods had been proposed over the years. The goal is to predict categorical class labels for new samples this means to assign an unknown sample image to one of a set of known texture classes according to their features by finding common traits or characters. [12]

3.1.Types of classification:

There are two main classification types:

- **Unsupervised Classification** (without human involvement):

It performs clustering of pixels in a data set based only on their statistics without using previous knowledge about the spectral classes present in the image. Within the computer uses techniques to determine which pixels are related and groups them into classes. The user can specify which algorithm will be used and the desired number of output classes but otherwise does not aid in the classification process.[13]

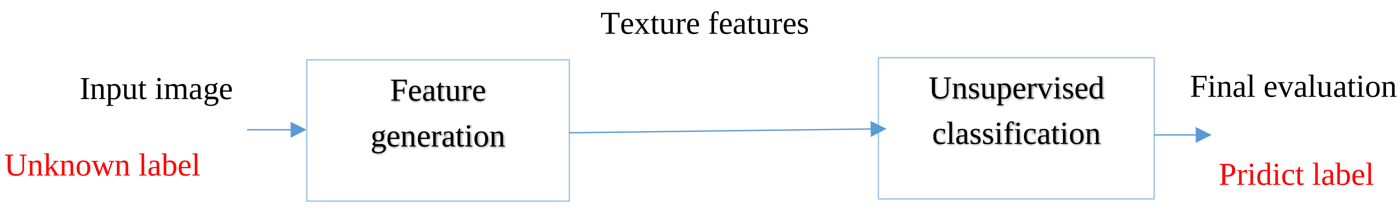


Figure 6: A general framework of unsupervised classification

- **Supervised classification** (human-guided):

Is the process of using samples of known identity or training data to classify pixels of unknown identity, the training is by using data which is well "labeled".[13]

Some of the methods that use supervised classification:

KNN, SVM, and Neural Networks.

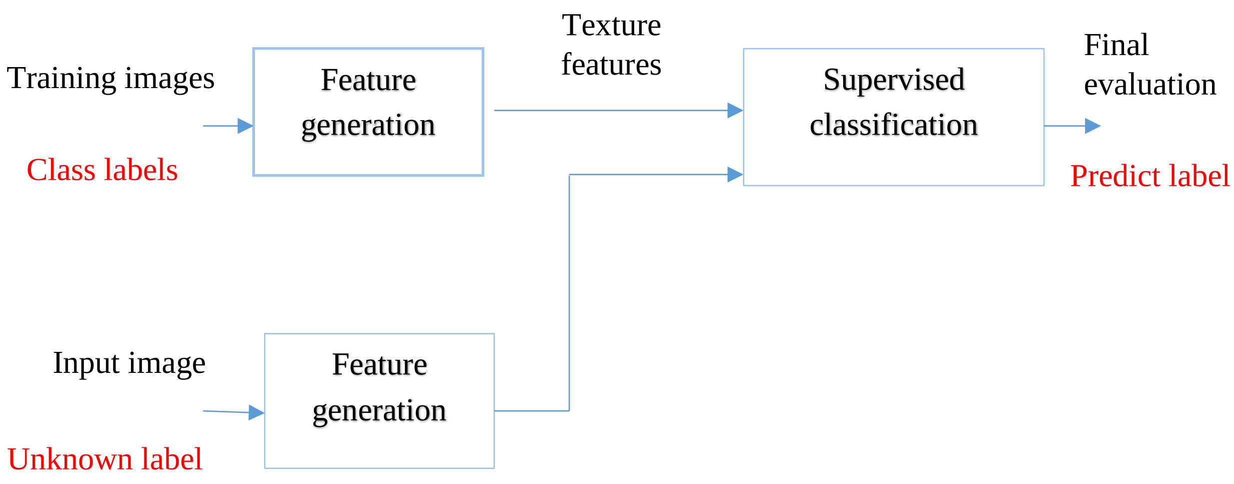


Figure7: A general framework of supervised classification.

- **Semi supervised classification:** Semi-supervised machine learning is a combination of supervised and unsupervised learning. It uses a small amount of labeled data and a large amount of unlabeled data. That means you can train a model to label data without having to use as much labeled training data.[34]

Chapter3: Related works

The general objective of texture study is to recognize texture categories based on their perceptual attributes. In literature, a number of textural attributes have been studied which may allow the detailed description of scenes.

This chapter surveys previous works that used multi-scale texture attributes in diffirenet approaches for recognition or classification texture,as summarised in table 1.

table1: related works

The work	The group of researchers	Year
A multi-scale Hierarchical Threshold-Based Completed Local Entropy Binary Pattern for Texture Classification	Xiaochun Xu , Yibing Li1 ,and Q. M. Jonathan Wu	2019
Multi-scale sampling based texture image classification	Yongsheng Dong, Jinwang Feng, Lingfei Liang, Lintao Zheng, and Qingtao Wu	2017
Multi-scale Rotation-Invariant Convolution Neural Networks for lung Texture classification	Qiangchang Wang, Yuanjie Zheng, Gongping, Weidong Jin, Xinjian Chen∗, Yilong Yin	2017
Classification of Mammogram Images Using Multiscale all Convolutional Neural Network (MA-CNN)	S. Akila Agnes et al.	2019
A Machine Learning Approach on Multi-scale Texture Analysis for Breast Micro-calcification Diagnosis	Annarita Fanizzi1 et al.	2018
Efficient Color Texture Classification Using Color Monogenic Wavelet Transform	Shan Gai	2017
Efficient texture retrieval using multiscale local extrema descriptors and covariance embedding	Minh-Tan Pham	2018

Texture description using multiscale morphological GLCM	Mudassir Rafi, Susanta Mukhopadhyay	2017
Performance of two multiscale texture algorithms in classifying silver gelatin paper via k-nearest neighbors	Kirsten R. Basinet et al.	2018

1. A multi-scale Hierarchical Threshold-Based Completed Local Entropy Binary Pattern for Texture Classification: (in 2019):

This work is a contribution to texture classification considering in multi-scale thresholding based on hierarchical partitions to binary encoding and completed local entropy binary pattern.

Problematic: most of LBP methods are sensitive to noise that's what makes it lose some texture's information, in addition they give simple and direct representations make it difficult to extract deep information from them, this is effectively what makes a question about how to describe the different local texture and how to extract information from them?

The proposal of this work is about, first propose a simple robust multi-scale thresholding framework depend on a hierarchical adaptive local partition, second keeping the simplicity of completed LBP and focus on its weakness by proposing an efficient completed local entropy binary pattern, then third a combination between the two methods a CLEBP extended to multi-scale thresholding framework.

Experimental results: For testing, they used Outex, UIUC, and KTH-TIPS databases, which show the powerful, and effectiveness of the framework. It gave high accuracy of classification comparing with recent methods.

This study is very good and effective but if it take into consideration color image only, it will be general and better.

2. Multi-scale sampling-based texture image classification:(in 2017):

The purpose of the study in this work is considered as a contribution on texture image classification by a multi-scale rotation-invariant representation (MRIR) of texture using multi-scale sampling.

Problematic: - most of the methods in the spatial domain-based are disrespect the directional information of textures by changing the scale of the local neighborhood, In addition, the transform-based methods are not effective to the rotated textures.

The proposal: First, it samples the multi-scale by Wavelet Based Multi-scale Decomposition of Magnitude Pattern Mapping (MP) than samples its Sign (SP). Second, build a multi-scale invariant representation to describe texture. Finally, obtain RIRs of different scales constructed by MRIR and use the minimum distance classifier for classification.

Experimental results: It used three rotated BRODATZ- data set by respectively rotating three original Brodatz texture datasets in terms of nine angles (0°, 5°, 10°,15°,30°,45°,60°,75°and 90°). the KTH-TIPS2-a dataset, and Outex TC 00012 under different rotation angles (0°, 5°, 10°,15°,30°,45°,60°,75°, and 90°) and illumination conditions (Inca, t184, and horizon) to evaluate the accuracy of the proposed method which gives a good and gratification results .

3. **Multi-scale Rotation-Invariant Convolution Neural Networks for lung Texture classification** : (in 2017)

The purpose of this study is lung texture classification by a multi-scale rotation-invariant convolution neural network (MRCNN).

Problematic: motivating by the distinguishing between lung tissue patterns like: (the emphysema) on High-resolution computed tomography (HRCT), others propose a method to extract effective features and find which intra-class patterns are more similar and inter-class patterns are more distinctive.

The proposal: First, the Gabor filter method for optimizing resolution in space and frequency domains, second Local Binary Pattern (LBP) to describe the various types of lung tissue patterns and the result of the concatenation of Gabor-LBP considered as input of Convolutional Neural Network (CNN) to learn multi-scale-high level features.

Experimental results: For testing, they use ILD database and evaluate the accuracy of the proposed method, which gives acceptable results but comparing to results based on handcrafted features are less.

The result of the purposed method is not robust enough, may be because the training is not enough or we can replace Gabor filter with other multi-scale filters and we can apply MRCNN to other medical and biomedical image classification problems as another suggested.

4. Classification of Mammogram Images Using Multi-scale all Convolutional Neural Network (MA-CNN): (in 2019)

This work is a medical contribution to make the classification of mammogram more easy and efficient by proposing an automated classification model of the normal and abnormal patterns using the approach: multiscale all conventional neural network (MA-CNN).

Problematic: the traditional CNN uses 3*3 size of the kernel to extract features, and sometimes this is not sufficient to detect and extract a wider range moreover the pooling layer has a risk to delete important information during the training.

The proposal: For that, Akila et al proposed to extract spatial features from a larger area of interest keeping the efficacy and speedily. They propose an approach of CNN, contains only conventional layers, in addition to multi-scale dilated convolution starting from the pre-processing to improve the quality of mammogram images. Than multiple dilated convolution develops the area of the receptive field without disturbing the feature map resolution, finally Striding for feature reduction used for subsampling.

Experimental results: For testing, the proposed approach used the mini-Mammographic Image Analysis Society (MIAS) database, and it showed an accuracy of 96.47 in 20s as time of classification, which confirms the effectiveness of the method in mammography photos classification.

5. A Machine Learning Approach on Multi-scale Texture Analysis for Breast Micro-classification Diagnosis: (in 2018):

Problematic: - breast cancer is one of causes of death of women in the earth, that is what it used mammography as a diagnostic tool for detecting this cancer, but it is still difficult for radiologist to understand clearly the results.

The proposal: First, Textural Feature Extraction: in this step, a wavelet transform is based on a multi-scale texture analysis approach. Second, Interest Point/Corner Detection: used two methods, 1-SURF method (Speed up Robust Feature) consist of 3 parts interest point detection, local neighborhood description and matching. 2-MinEigenAlg method to identify the interesting corner of an image. Finally, the classification model: for evaluating the kind of feature selection techniques, such as the filter method.

Experimental results: in this work, they obtained ROI (Region of Interest) from the image ofthe BCDR (Breast Cancer Digital Repository) database to evaluate the purpose method. Therefore, the accuracy was better compared with the results of other related works.

6. Efficient Color Texture Classification Using Color Monogenic Wavelet Transform: (in 2017):

This paper showing the effectiveness of taking into consideration the color proprieties in image classification, it propose a multi-scale color texture classification algorithm that adds a contribution to non-marginal representation for the color monogenic signal.

Problematic: the new presentation of the image is composed of intensity, color and motion features that's why the exploitation of this proprieties makes the analyses more robust, this work based on the CMWT (color monogenic wavelet transform) to propose a new presentation of the features vector and a new classifier.

The proposal: depending on the CMWT method that obtains better texture analysis results, it extracted information at different scales of the directional sub-bands such as mean value, standard deviation, energy, and entropy this data considered as a new presentation of features vector, and for finding the appropriate class it used the Euclidean distance.

Experimental results: The proposal was tested on three different databases: VistTex, USPTex and Outex and it showed good results with high accuracy 98.67, 99.08 and

99.89% for VisTex, USPTex and Outex respectively that's what confirms the strength of the algorithm and the advantage of using color proprieties in image classifications.

7. **Efficient texture retrieval using multi-scale local extreme descriptors and covariance embedding: (in 2018).**

The proposed method in this work used multi-scale feature extraction and embedding based on the local extreme key points to describe the texture of the image.

The proposed method: The color image is divided into SLED vectors that least are regular blocks that have the same size and each one is characterized by a feature vector constructed using the radiometric (i.e. color) geometric and structural information of its local extrema (local max and min pixels). Those vectors established into feature covariance matrix, and to implement the algorithm for different scales they proposed to sample up and sample down each image then use Riemannian, distance to measure the dissimilarity of the proposed method within the retrieval task

Results: good results yielded by this method on different databases witch are MIT Vistex, the Salzburg Texture (Stex) , the Outex TC-00013 and the USPtex .It gave high accuracies of 94.95% , 79.87% , 76.15% and 89.74% orderly, more of its simplicity and easy.

8. **Texture description using multi-scale morphological GLCM: (in 2017)**

Problematic: the GLCM is one of the powerful descriptors of textures but it suffers from a limitation of capturing information at one single scales.

This paper is written to confirm that the classification accuracies can be improved if multi-scale mathematical morphology is applied with GLCM.

Methodology: based on four observations drown by the others the frameworks are inspired. MMGLCM1: consist of two parallel phases opening and closing that split the original imager into a set of filtered images keeping the resolutions unchanged. Then the normalized features are combined using the concatenation of the results of opening and closing. MMGLCM2 decomposes the textural image into component images based on MMGLCM1.

Results: the methods are applied for three databases (Brodatz, UIUC, UMD), The obtained classification results prove the accuracy and confirm the effectiveness of combining mathematic morphology and GLCM even different scales of interest.

9. Performance of two multiscale texture algorithms in classifying silver gelatin paper via k-nearest neighbors: (in 2018).

Problematic: As in other domains, texture analyses provides facilities to the community of museum and other collecting community institution to comprehend different images than classify them, this paper’s classification is based on visual inspection by art conservators and curators. To more understand the silver gelatine images and its purpose and make its classification more easy and efficient, two multi-scale approaches combined with the KNN approach are implicated to the real world database.

Method: anisotropic multistate analysis (AMA): this method proposed in the context of analysis of scale-invariant (free) texture. Pseudo_area_scale analysis: to decompose the surface into a patchwork of triangles of a given size. Passed to the KNN method that usees AMA and PASA as a distance function to provide features to classification.

Results: by applying this method on Yale Lens Media Lab (LML) Reference Collection of Photographic papers, more than 69% accuracy of classification has been achieved.

Chapter4: Methodology

As it is mentioned in the previous chapter, texture classification refers to assigning an image into one of a set of predefined categories. Texture classification tasks involve two main steps: **feature extraction step**, where texture features are extracted from the image and **classification step** where texture class membership is assigned according to the extracted texture features. In this work, we choose the mathematical morphological operations as a method to generate different scales from the input image and CNN as a method for the feature extraction and classification.

This chapter provides a presentation of major concepts as: multi-scale texture, mathematic morphological operations and CNN (definitions, applications and how they work), moreover to the role of color in extracting features and our proposal texture classification method.

- The Plan in figure ’8’ shows the steps of the proposed classification system.

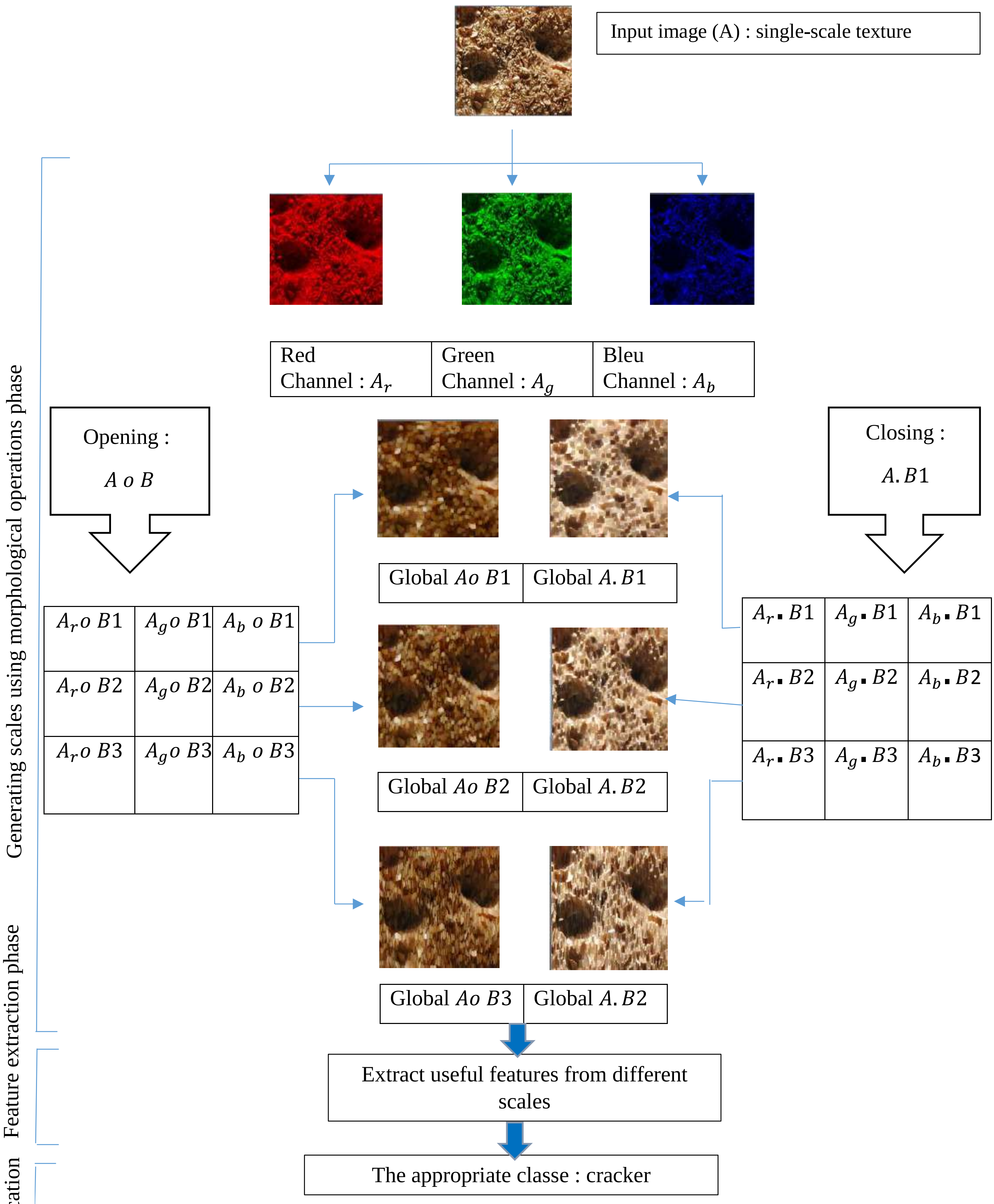


Figure 8: the steps of our proposed method of texture classification

1. Multi-scale texture:

Petrou et al. argued that: “Texture is the variation of gray levels or color that have a specific relative arrangement, at scales smaller than the scale of interest (SOI)”. Since the image is a function of gray level’s light intensities or color channel’s variation, any region or scale of interest should include a scale smaller than it, to provide the variation of the features. In other words, we can say that different region of interest in the texture generates different and diversity of the features extracted from each region to explore the multi-scale information from the same image.

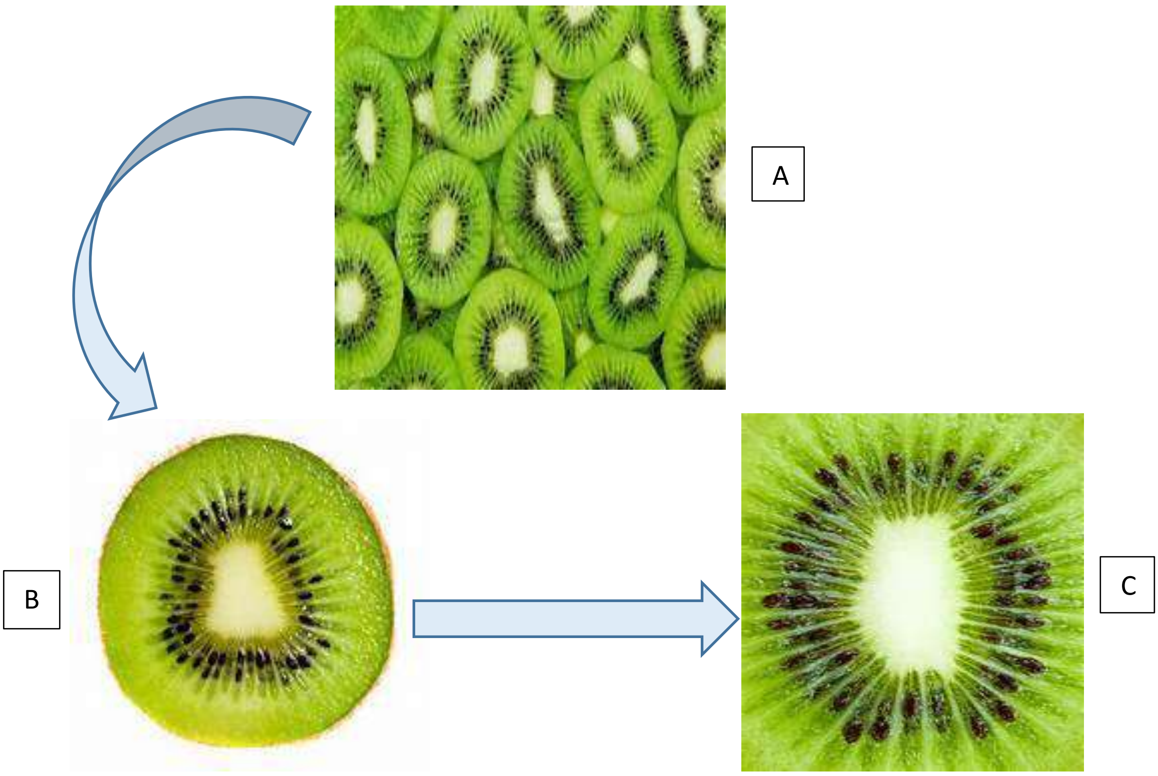


Figure 9: different scales of interest of kiwi's texture

In Fig 9, it can be seen clearly that switching texture from one scale to another when the scale of interest is large (A), the variation of color of kiwis pattern termed as texture; When the scale has been reduced (B) containing only one kiwi, it shows the surface of another type of texture.

2. Mathematical Morphology Operations :

The word morphology refers to form, shape and structure. In theory, mathematical morphology is a tool for extracting image components that are useful for representation and description; it is used for the analysis of spatial structures in images that says in simple word, “The keyword for morphology is the shape”.

Based on literature: the mathematical morphology is a theory that dates back to Georges Matheron (1964) and Jean Serra’s (1982) studies. Both initiated the development of binary images. Then, the theory has been extended to greyscale images and color images too. [14]Uses of morphological analysis in image processing: the mathematical morphology can be applied in any field of image processing where shape plays some role, this can be an object:

- Noise reduction and filtering: reduce noise as much as possible without eliminating essential features.
- Shape description
- Feature detection and extraction.
- segmentation
- texture analyses
- Classification
- Analysis of connectivity of components
- Object selection using geometric features and so on.

2.1.Structuring Elements (SE):

Morphological operations use a small shape or template known as a structuring element.

2.1.1. Definition and objective:

Structuring elements are a sub-image or small set in binary matrix form used to interact with the image to be probed. The mechanism of the structuring element resembles that of masks used in spatial filtering, it is moved all around the image and is positioned at all possible locations on it. Moreover, the structuring element is compared to the corresponding neighborhood of pixels to find hit or miss at that location, it also helps to differentiate image objects or features. [15]

- The important characteristics that should take into consideration for a structuring element are its shape, size and origin. The **Shape** of the structuring element is the arrangement of ones and zeros in a pattern within the matrix. The **Size** of the structuring element acts as a ‘window’ over which the interaction takes place. **The origin** of the structuring element identifies the pixel of interest that is to be processed. [15]

2.1.2. Shape of structuring element:

The structuring element can be shaped like a disk, a square, vertical rectangular, horizontal rectangular, diagonal, or irregular form. Better details can be obtained from the image by choosing the suitable structuring element. Figure 10 represent different shapes of structuring elements for binary image

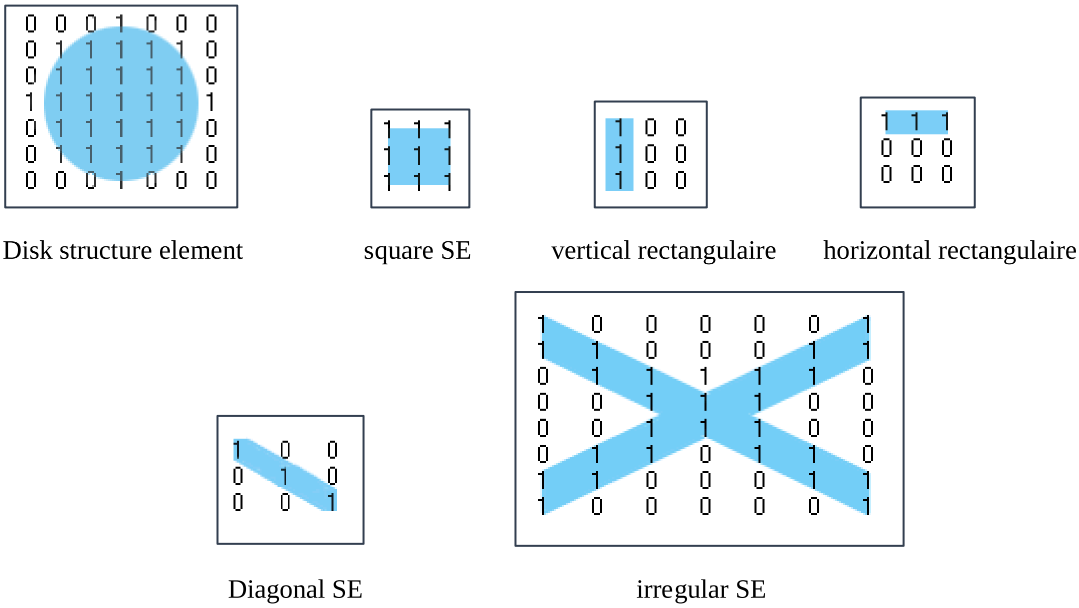


Figure 10 : different Shape of SE [16]

2.1.3. Hits and fits:

Fit: All pixels in the structuring element cover on pixels in the image

Hit: Any pixel in the structuring element covers one pixel in the image.

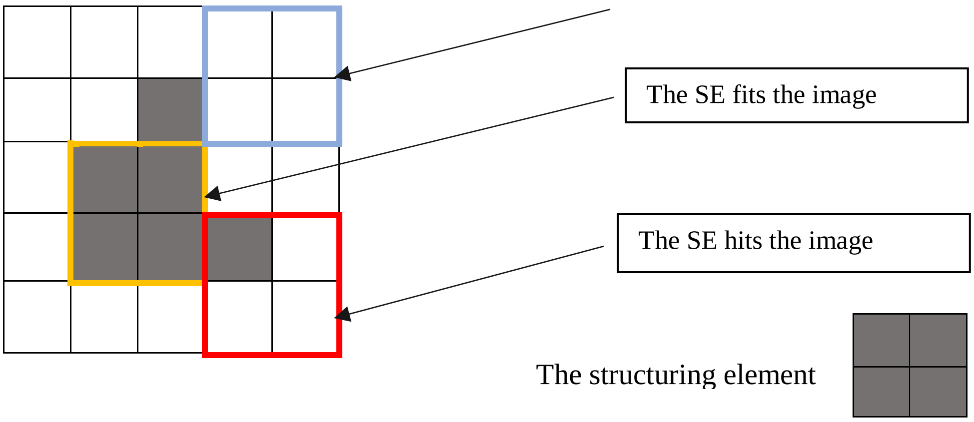


Figure 11: fits and hits using structuring elements

2.2.Erosion and dilation:

The two basic morphological set of transformations are erosion and dilation. Dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels at the boundaries of objects. The number of pixels added or removed from the objects in an image depends on the size and shape of the one used to process the image. [17]

2.2.1. Erosion:

The morphological erosion is the operation of making an object lose size and makes the gaps between different regions larger. It is typically applied to binary images, but some versions work on grayscale images. The erosion of an image **A** (set) by structuring element **B** is defined as:

$$A \ominus B = \{Z|(B)_z \subseteq A\}$$

The erosion of an image **A** by structuring element **B** is the set of all points **z** such that the structuring element **B** is translated by **z** is a subset of the image. This operation results in the loss of boundary pixels of the object.

The erosion process enlarges the background of the object (number of pixels with value zero in binary image) and regresses the foreground of the object (number of pixels with value one). It removes the structures which are smaller than that of the structuring element. So it can be used to remove the noisy “connection” between two objects in the image. [18]

• The structuring element **S** is positioned with its origin at **(x, y)** and the new pixel value is determined using the rule:

$$g(x,y) = \begin{cases} 1 & if\ S\ fits\ the\ pixels\ in\ image\ f(x,y) \\ 0 & otherwise \end{cases}$$

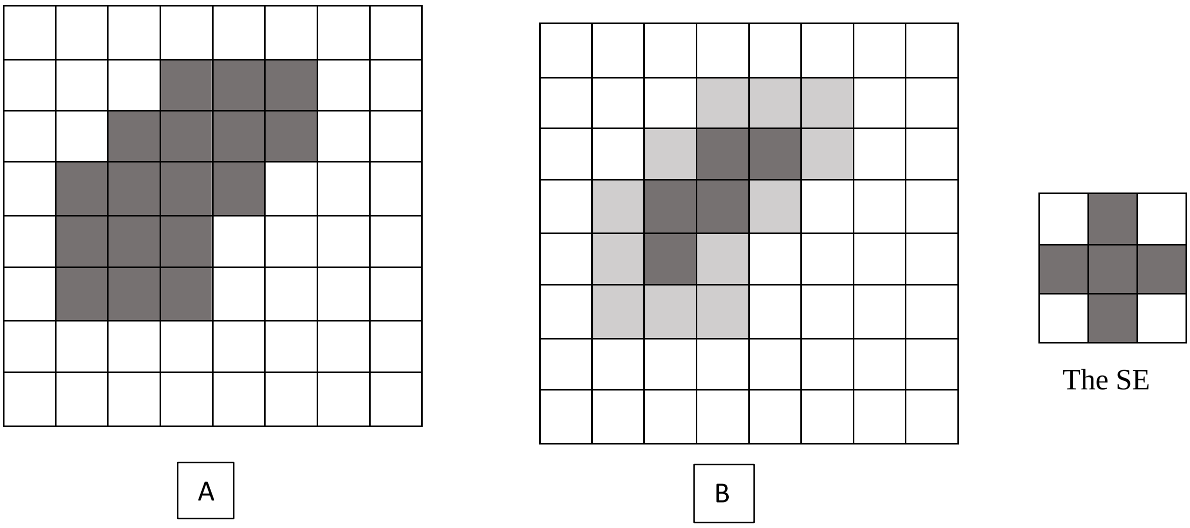


Figure 12: Example for erosion transformation where A is the original image and B is the processed one

Figure 12 shows the shape of the object in the image before erosion in A and after it in B using 3-3 structuring element where the shape becomes smaller than the original one.

1.3. Dilation:

The morphological dilation is a complement of the erosion operation. It is the operation of making an object grow by size, to be more visible and fills small holes (missing pixels) in continuous objects depends on the nature and shape of the structuring element.

The dilation of an image **A** (set) by structuring element **B** is defined as:

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset \}$$

If a set **B** is reflected about its origin and translated by **z**. then the dilation of **A** by **B** is the set of all displacements z such that **B^** and **A** have at least one common element. Dilation, as said above, adds pixels to the object's boundary. The dilation process enlarges the foreground (the number of pixels with value one in binary image) and regress the background (number of pixels with value zero).[18]

Dilation affects the intensity in the image, as a result blurring effect can be observed. Therefore, it can be said that it is analogous to smoothing spatial low-pass filters that are used in linear filtering of the image.

- The new pixel's value is determined using the rule:

$$g(x,y) = \begin{cases} 1 & if\ S\ hits\ the\ pixels\ in\ image\ f(x,y) \\ 0 & otherwise \end{cases}$$

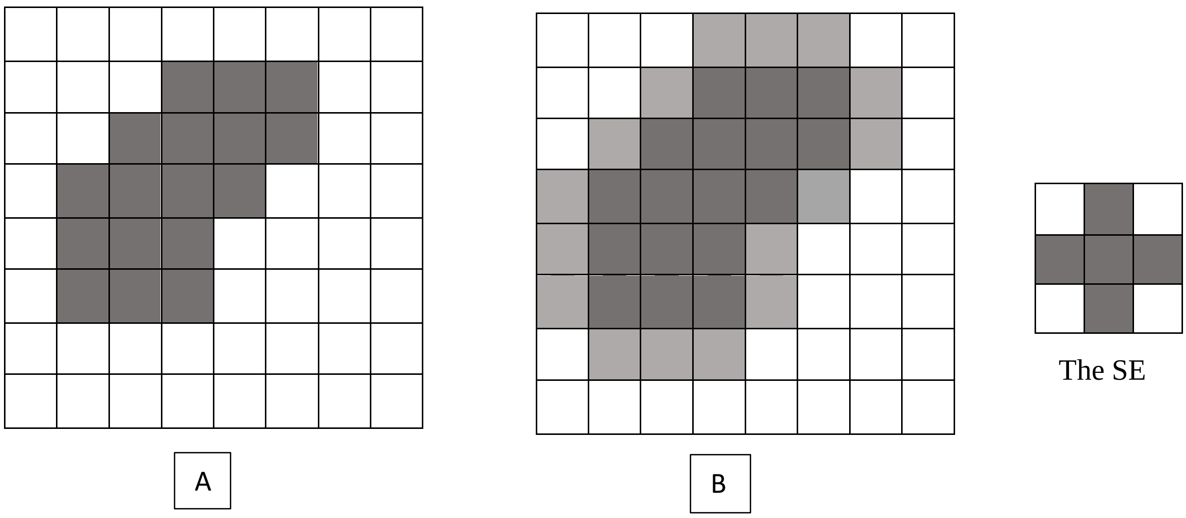


Figure 13: Example of dilation transformation where A is the original image and B is the processed one

Figure 13 shows the shape of the object in the image before dilation (A) and after it (B) using 3-3 structuring element where the shape becomes bigger than the original one.

- What about if we want: to remove structures and fill holes without losing the remaining parts?

  SOLUTION: Is to combine erosion and dilation using the same structuring element.

  From here came the idea of opening and closing operations.

2.3.The Opening and closing operations:

Opening and closing are the most useful of these for morphological filtering.

2.3.1. Opening:

The opening of an image involves the application of erosion, followed by dilation using the same structuring element. The name tells that the operation can create an opening between two structures that are connected only in a thin bridge, without shrinking the structures (as erosion would do). [19]

The morphological opening of an image A by the pixels of a SE B is given by: A O B

and defined as: $A\ o\ B = (A\ \ominus B) \oplus B$ that means $A\ o\ B = \cup \{Bx \coloneqq Bx \subseteq A\}$.

the opening of A by B is obtained by taking the union of all translates of B that fit inside A. Parts of A that are smaller than B are removed.

The effect of the opening operation is: to smooth boundaries, to break narrow isthmuses, to eliminate small noise regions (small objects), and to separate connected objects.

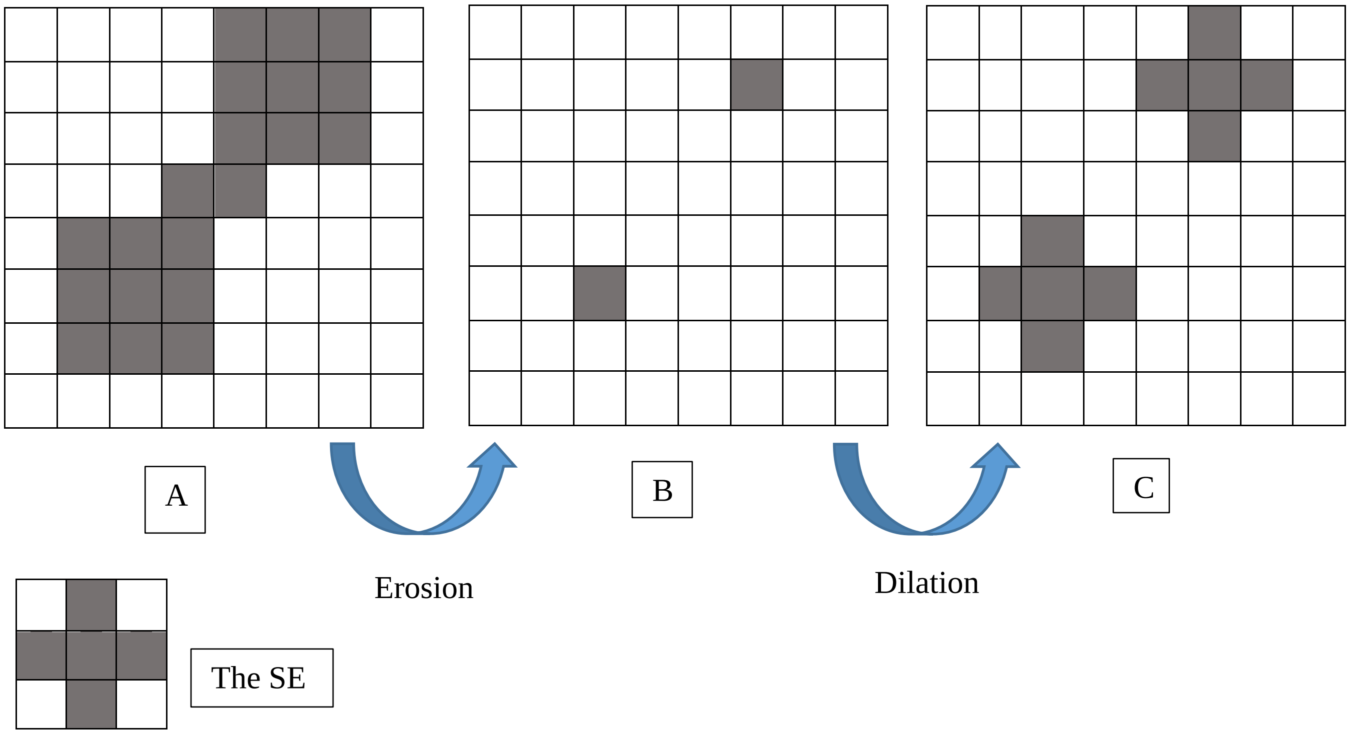


Figure 14: A is the original image; C is the result image the opening operation using the same SE

figure 14 represents the opening operation where the shape in the image A separated into two shapes in image C after a consecutive eriosion and dilation using 3-3 structuring element.

2.3.2. Closing:

The closing of an image involves the application of dilation, followed by erosion using the same structuring element.The name tells that the closing operation can close gaps between two structures without growing the size of the structures as dilation would. [19]

Morphological closing of an image A by the pixels of a SE B is given by: A . B

and defined as: $A \,.\, B = (A \oplus B) \ominus B$

The closing operation in general fills narrow breaks and thin gaps in addition to eliminates small holes in the object’s boundaries3.

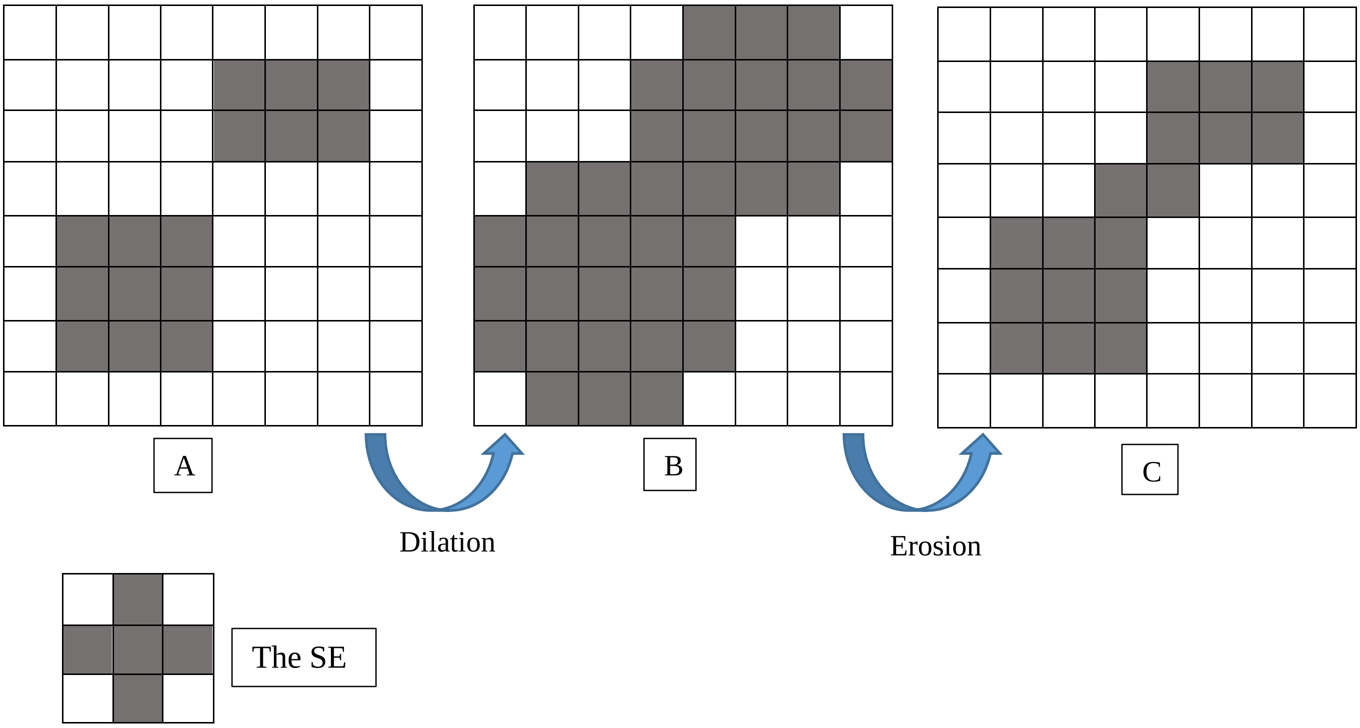
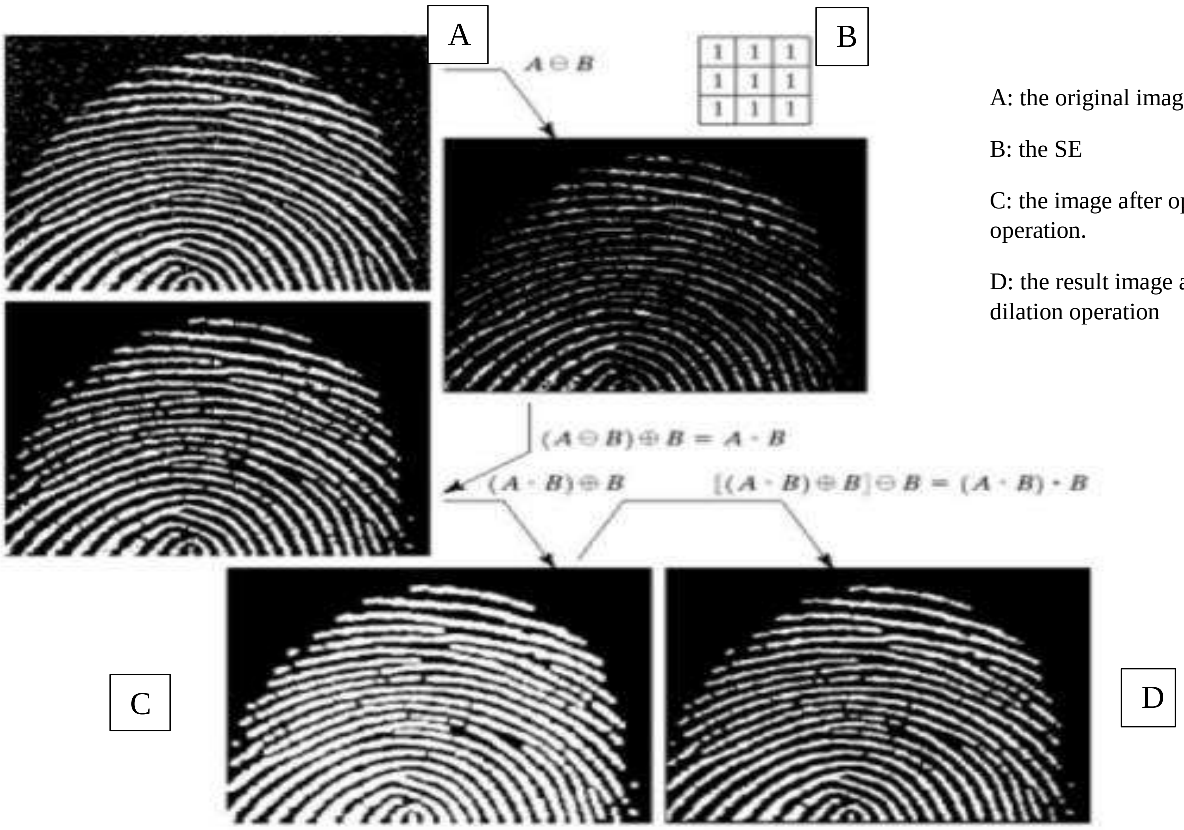


Figure 15: A is the original image, C is the resulting image the closing operation using the same SE

Figure 15 represents the closing operation where the two shapes in the image A becomes one as is showed in image C after a successive dilation and erosion using 3-3 structuring element .An example based to morphological operators (opening, closing):



A: the original image

B: the SE

C: the image after opening operation.

D: the result image after dilation operation

Figure 16: an example for opening than closing to the same image A

In figure 16, we can see that the image A after the opening and closing morphological operations becomes clearer without mess and noise to extract only the useful features from the input image.

2.3.3. Why we choose mathematic morphological operations?

As mentioned before, the morphological is a powerful tool in image processing, and it has been applied successfully to generate different levels of scales. Thus, generate different features associated in the resulting image in different regions of interest in the image.

That is why the MMOs are the suitable tool in our work for make new multi-scale textures images to use them in the extraction and classification steps.

3. Convolutional neural network (CNN):

Convolutional neural network (CNN or ConvNet for short) is one of the most popular tools for deep learning. It is a particular kind of supervised learning used to analyze data.

3.1. Applications:

CNN has shown excellent results in:

- identifying faces
- powering the vision in robots
- Self-driving cars and traffic signs.
- Object recognition, segmentation, detection, classification…
- Natural language processing
- Health Risk Assessment
- Video analysis …..

3.2. Brief history and inspiration:

The CNN was inspired by the natural visual perception mechanism of the living creatures. In 1959, **Hubel & Wiesel** found that cells in the animal visual cortex are responsible for detecting light in receptive fields. Inspired by this discovery, **Kunihiko Fukushima** proposed the neocognitron in 1980, which could be regarded as the predecessor of CNN. In 1990, **LeCun et al** published the seminal paper establishing the modern framework of CNN, and later improved it moreover, they developed a multi-layer artificial neural network called LeNet-5 which could classify handwritten digits. [20]. Since 2006, many methods have been developed to overcome the difficulties encountered in training deep CNNs, with the success of those methods many works have been proposed to improve their performance by getting deeper such as:

GoogleNet, and ResNet.

By increasing depth in recent works, the network can better approximate the target function with getting better feature representations.

3.3.What makes the CNN useful?

The use of CNNs for deep learning has become increasingly popular due to three important factors:

- CNNs eliminate the need for manual feature extraction: CNN itself learns the features directly.
- CNNs generate effective recognition results.
- CNNs can be retrained for new recognition tasks from existing networks.[21]

3.4.The architecture of CNN:

The Regular Neural Networks receive an input and transform it through a series of hidden layers. Each hidden layer is made up of a set of neurons, where each neuron is fully connected to all neurons in the previous layer. Neurons in a single layer function completely independently and do not share any connections. The last fully-connected layer is called the "output layer" and in classification settings, it represents the class scores. [22]

3.4.1. Types of layers in CNN:

CNN consists of three types of layers, namely convolutional, pooling, and fully-connected layers which are the key parts of almost all CNN models.

1- The convolutional layer: This is the central layer in the context of a convolutional neural network, which gives the network its name, a convolution is a linear operation that involves the multiplication of a set of weights with the input, much like a traditional neural network. The multiplication is performed between an array of input data and the filter or the **kernel** to get a feature map as an output**.**

- **The kernel**: is a matrix that moves over the input data, performs the dot product with the sub-region of input data, and gets the output as the matrix of dot products. Kernel moves on the input data by the stride value [23], In short, the kernel is used to extract high-level features like edges from the image.

Convolution operation can be expressed as follows:

$$f_l^k(p,q) = \sum_c \sum_{x,y} i_c\,(x,y).\,e_l^k(u,v)$$

Where: $i_c(x,y)$an element of the input image multiplied by is: $e_l^k(u,v)$ index of the k^{th} convolutional kernel of the l^{th} layer, whereas output feature-map of the k^{th} convolutional operation can be expressed as: $F_l^k = [f_l^k(1{,}1), \dots f_l^k(p,q), \dots f_l^k(P,Q)]$ [24].

Figure 17, shows the convolutional layer where the input pixel with value 1 multiplicated with a kernel (or filter) to become a new value (-8), and the combination of all input pixels after the multiplication makes up a new matrix knowing as feature map.

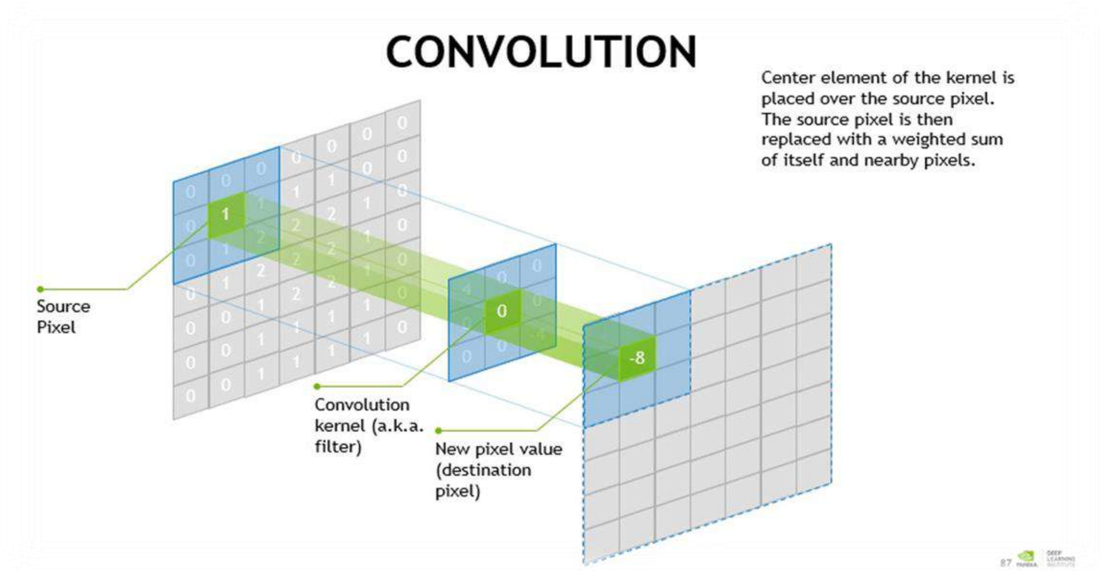


Figure 17 : convolution layer [25]

- **Parameters of convolution layer :**

Any convolution layer applies the same parameters across different regions in the image that we can specify:

*** Filter size:** This corresponds to how many inputs features in the width and height dimensions one neuron takes in (in other way is: kernel size).

*** Stride:** Stride is simply how many pixels we want to move (towards the right/down direction) when we apply the neuron again [26] with the same parameters of the kernel.

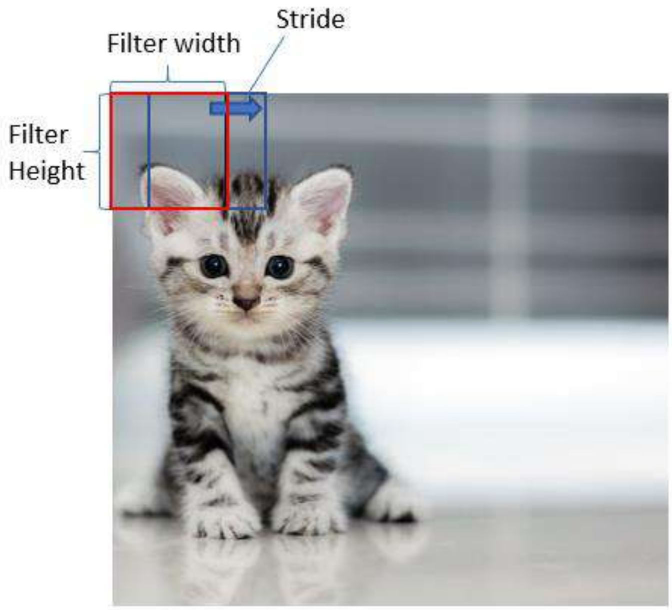


Figure 18: A convolution layer with filter width, filter height and stride [26]

The last figure shows the parameters of the kernel besides the moves of the stride on the image.

*** Depth:** refers to the fact how many different neurons applied to each part of the image.

* **Padding:** Recall that the center of the 3x3 filter started at pixel 2 (instead of at pixel 1) and ended at pixel 255 (instead of at pixel 256). To make the center the filter starts at pixel 1. We can pad the image with a border of ‘0’s. [26]

2- Pooling layer (downsampling):

This type of layer is often placed between two layers of convolution: it receives several feature maps and applies the pooling operation to each of them. The pooling operation consists in reducing the size of the images while preserving their important characteristics. To do this, we cut the image into regular cells, and then we keep the maximum value within each cell. In practice, small square cells are often used to avoid losing too much information. [27]

Pooling operation can be expressed as follows:

$$Z_l^k = g_p(F_l^k)$$

Equation (2) shows the pooling operation in which Z_l^k represents the pooled feature-map of l^{th} layer for k^{th} input feature-map F_l^k , whereas $g_p()$ defines the type of pooling operation. [27]

3- The Rectified Linear Units layer (ReLU):

The ReLU correction layer replaces all negative values received as inputs by zeros keeping positive values [26]. It acts as an activation function. As only activated features continue on their way to the next layer and defined by

$a_{i,j,k} = \max(Z_{i,j,k}, 0)$. Where $Z_{i,j,k}$ is the input of the activation function at location (i, j) on the k^{th} channel.Visually, it looks like the following (figure 19):

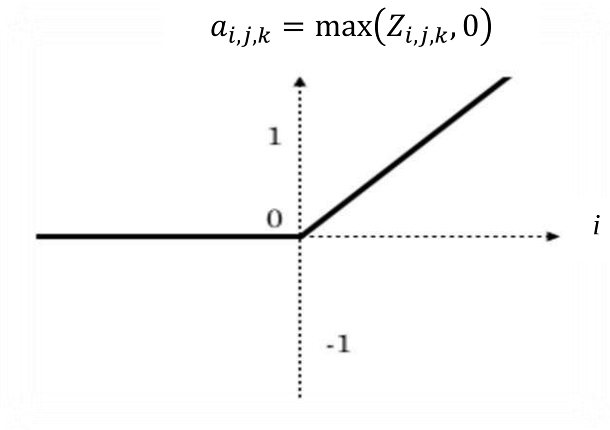


Figure 19: the ReLU function [25]

4- Fully connected layer (FC) :

The fully connected layer is mostly used at the end of the network for classification. On that layer, Neurons have full connections to all activations in the previous layer. It takes input from feature extraction stages and globally analyses the output of all the preceding layers; consequently, it makes a non-linear combination of selected features, which are used for the classification of data. [28] .So, the ConvNets need not be limited to only one Convolutional Layer. Effectively, the first ConvLayer is responsible for capturing the Low-Level features such as edges, color, gradient orientation, etc. With added layers, the architecture adapts to the High-Level features as well, giving us a network that has the wholesome understanding of images in the dataset, similar to how we would.[29]

Figure 20 represents conventional neural network architecture from the input image passing truth many convolutional and pooling layers arriving at FC layer to get an output class.

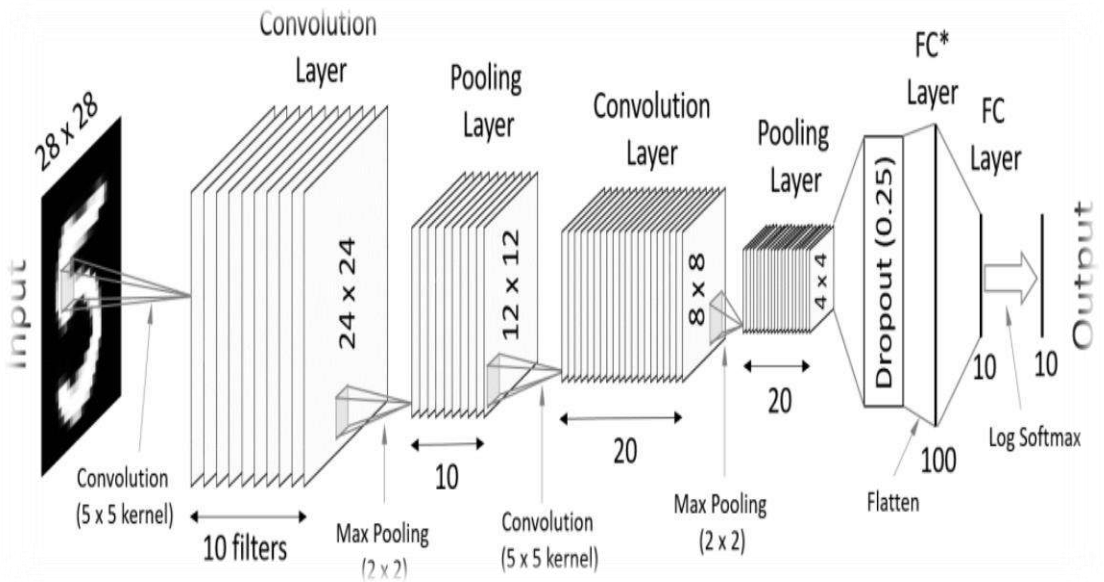


Figure 20: a general Convolutional neural network architecture [30]

3.4.2. How does the CONVnet work:

1- The input image pass to the first convolutional layer. The convoluted output is obtained as an activation map. The filters (kernels) applied in the convolution layer extract relevant features from the input image to pass further.

2- Each filter gives a different feature to aid the correct class prediction.

3- Pooling layers are then added to reduce the number of parameters

4- Several convolutions and pooling layers are added before the prediction is made. Convolutional layer help in extracting features. As we go deeper into the network more specific features are extracted.

5- In the fully connected layer, the input from the other layers is flattened and sent so as the transform the output into the number of classes as desired by the network.
6- The output is then generated through the output layer and is compared to the output layer for error generation. A loss function is defined in the fully connected output layer to compute the mean square loss. The gradient of error is then calculated.
7- The error is then backpropagated to update the filter (weights) and bias values.
8- And like that One training cycle is completed in a single forward and backward pass.[22]

4. The role of color in image classification:

Although color is important information in visual object recognition. Many texture features do not take in consideration the color information because many of their calculations are based on the gray scale image.

Color features is shown to be advantageous and to achieve higher success rate than grayscale features in texture analysis, they can give an additional information of the intensities, and this is what provide an effective results and high accuracy of texture classification.

5. How our methodology work ?

In the first step, we split our input image into three channels (Red, green, and blue) to take the advantage of all color texture information. Then for each channel we apply morphological operations (opening and closing) using three different kernels (showing in figure 21) to lead us after merging them to six new images. In order to generate ifferent scales of interest and to facilities the pickup of features from different levels in the image.

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

K1 K2 K3

Figure 21 : the kernels used in our methodology

In the second phase, the 6 images are fitted to CNN to extract features and capture the information contained on the texture using its robust layers. This also transforms the image from a matrix into a useful feature vector to use it in the classification phase. By corresponding each image with its appropriate class after training and test validation.

Chapter 5:

Experimental result and Discussion

In this section, comprehensive experiments are conducted to demonstrate the effectiveness of our proposed method. We use morphology operations to generate multi-scale and CNNs for extract features and classification.

We shall describe the experiment setup. Then, a comparative study with single scale vs multi-scale and handcrafted texture vs learning-based descriptors using KTH-TIPS dataset.

1. Setup:

In this step, we will describe the conditions of our experiment.

1.1. Dataset (KTH-TIPS):

KTH consists of 10 texture classes. Images are captured at 9 scales. We divide it into two parts, from each class and each scale we took the last image (part 1 for testing) and the rest we divided into two parts. 25% for test validation and 75% for training.[31]

Figure22 : classes of images in the KTH-TIPS database[31].

1.2. Metrics:

Accuracy: is a famous measure to evaluate the performance of a color texture classification system.

$$\text{Accuracy} = \frac{true\ positives + true\ negatives}{all\ simples} \times 100$$

1.3. Descriptors and networks:

a **LBP**: Is a very powerful method to describe **texture** and shape of a digital image. It appeared to be suitable for feature extraction. [32]. It is a vector of binary codes holding information about the local neighborhoods. The grey value of that pixel is compared to the grey values of pixels in its neighbourhood [33]. In our experiments, we used uniform-LBP with eight radius.

b **GLCM**: a statistical method of examining texture that considers the spatial relationship of pixels is the gray-level co-occurrence matrix, also known as the gray-level spatial dependence matrix. In our experiment, we used distance= 3 and the angle =45 and consider the following statistics: correlation, contrast, energy and homogeneity.

c **ANN**: In our experiment, we used 3 simple layers (with 4624, 2312 and 10 neurons respectively) for classification purpose:

d **VGG**: is a convolutional neural network that is 19 layers deep. We load a version of the network pre-trained on more than a million images from the ImageNet database. The pretrained network can classify images into 1000 object categories. As a result, the network has learned rich feature representations for a wide range of images [34]. In the experiment, we added 5 layers to the pre-trained model in reason of flattening the output of the pooling layer than fit them into the next layers before the last fully connected layer that contains 10 neurons (classes) , and to make the learning deep also. In this case, we train only the new layers to suites the last outputs. The new architecture is shown in figure 24.

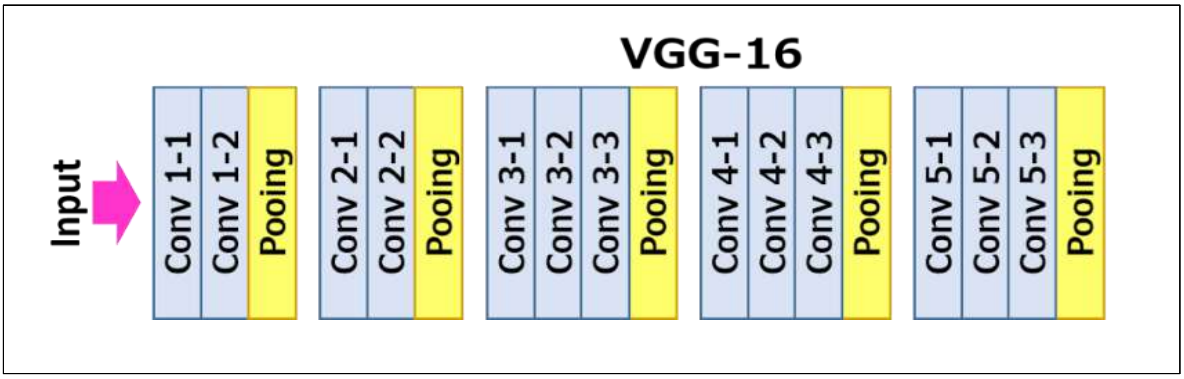


Figure 23: the pertained VGG model

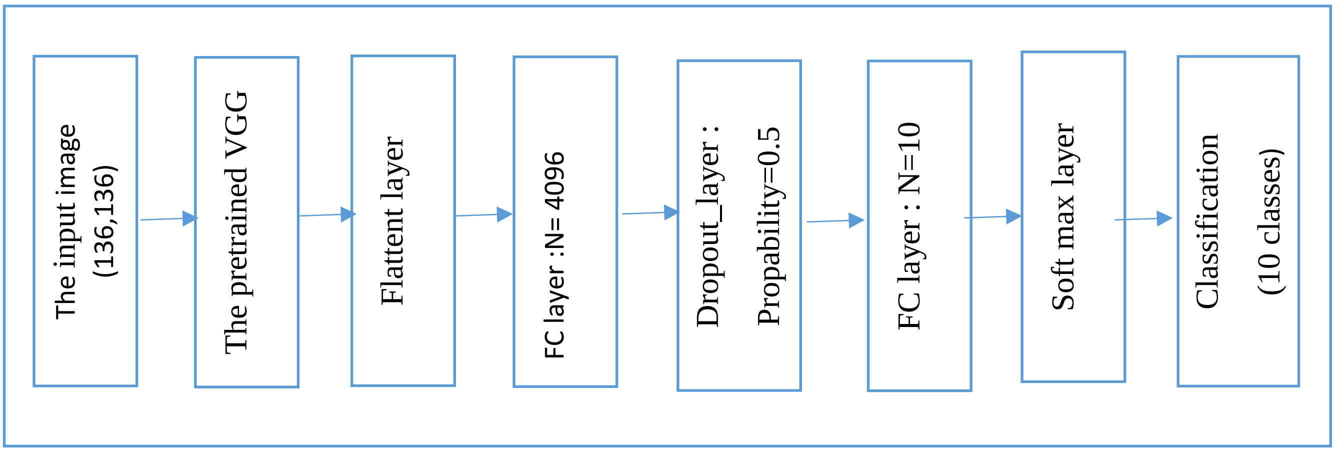


Figure 24: the VGG model used in the work

e **ResNet**: ResNet-50 is a convolutional neural network that is 50 layers deep trained on more than a million images from the imageNet dataset. As a result, the network can classify images into 1000 object categories such as key bored and animals. Figure 24 shows the architecture of RESnet-50.

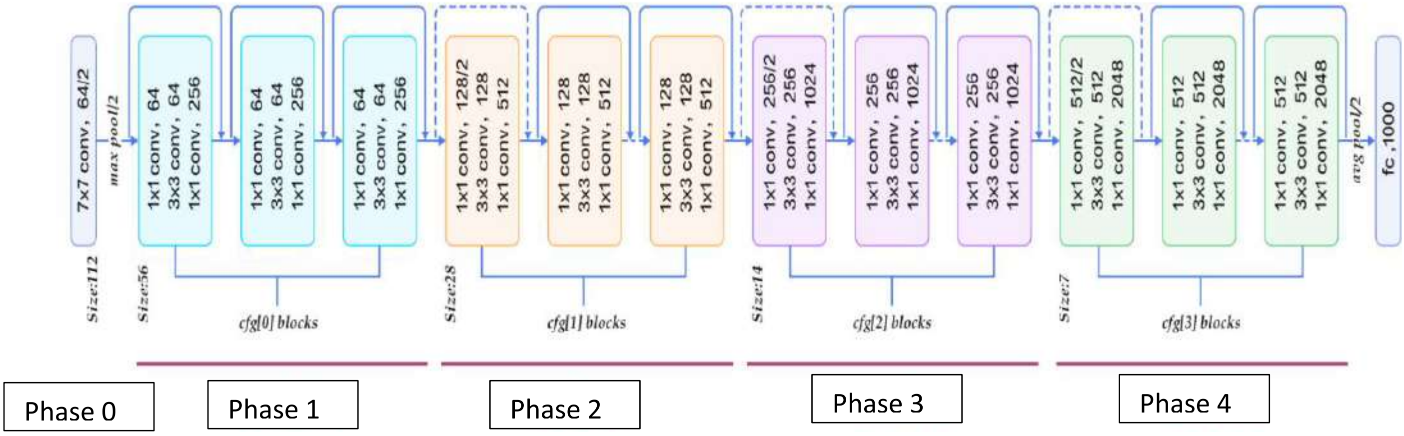


Figure 22: the archetecture of RESnet

To use this architecture in our experimentation, we keep the same layers but change the sizes of the input image into (136,136,3) and the parameters of the fully connected layer into 10 (number of classes).

f Our proposed CNN model :

We proposed also a CNN model that contained 12 layers, its architecture showed in figure 25. This model has been trained from scratch using KTH-TPS dataset.

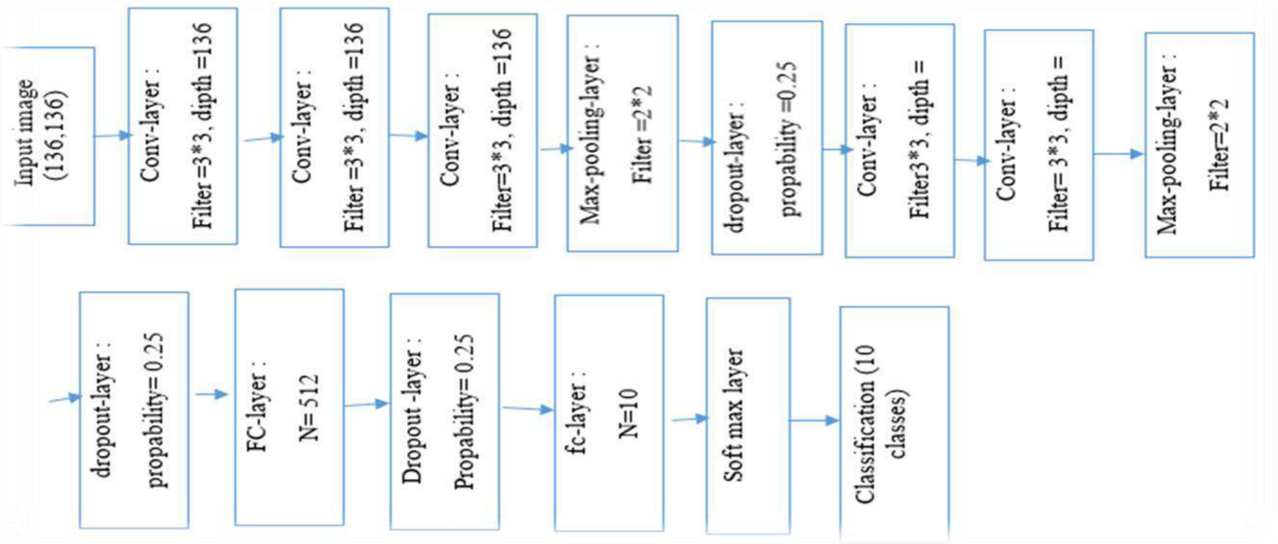


Figure25: Our proposed CNN model

1.4. Hardwar and software:

a **Hardwar**: all tests were performed on a personal computer HP I7 with a 8Go capacity memory (RAM), and central process units Intel (R) Core (TM) 2 GHz 2.6GHz, with windows 7 edition integral, service pack 1, 32 bit system type as an installed operating system.

b **Software**: In order to implement our application, we have used the environment Jupyter Notebook. It is an open-source web application that allows us to create and share documents that contain live code, equation, visualizations and narrative text. Its uses include data learning and transformation, numerical simulation, statistical modelling, data visualization, machine learning and more [35].

2. Results and discussion:

In this section we will evaluate two main aspects. First, the impact of generating and using multi-scales with handcrafted feature descriptors. Secondly, the impact of using multi-scales with CNNs compared to handcrafted descriptors.

2.1. Scenario 1: Single scale VS multi-scale:

In this experimentation for the single scale evolution, we used 720 images (612 for training and 108 for evaluation) and for multi-scale, we used 4320 (720*6images generated from morphological operations) 75% from them for training (3240 image) and the rest for evaluation.

Table2 : single-scale vs Multi-scale

	The accuracy	
The method	Single scale	Multi-scale (morphological operations)
LBP+ANN	22.13%	58%
GLCM+ANN	34.42%	40.38%

Table2 shows the comparison of classification obtained with LBP+ANN and GLCM+ANN using single scale texture and multi-scale texture on KTHP-TIPS dataset. We can see clearly that, the accuracy of classification using multi-scale texture in both methods (LBP+ANN) and (GLCM+ANN) is higher than classification using single scale. We can see in table 2 that LPB+ANN method raised by more than 100% while the accuracy in GLCM+ANN raised by 17%. We can say that using multi-scale texture in classification is more effective, it generate more useful features and information from the pattern the makes the accuracy of classification high than the single scale pattern.

2.2 Scenario 2: hand-crafted vs learning-based:

For this experiment, we used 720 images from KTHP-TIPS generate 4230 new images using morphological opening and closing operations. 75% from them is used for training (3240 images) and the rest for evaluation (1080 images).

Table3 : the accuracy of methods

The method	Accuracy
LBP +ANN	58%
GLCM +ANN	40.38%
Our proposed model of CNN	84.35%
VGG model (CNN)	85%
RES-net model (CNN)	56.54%

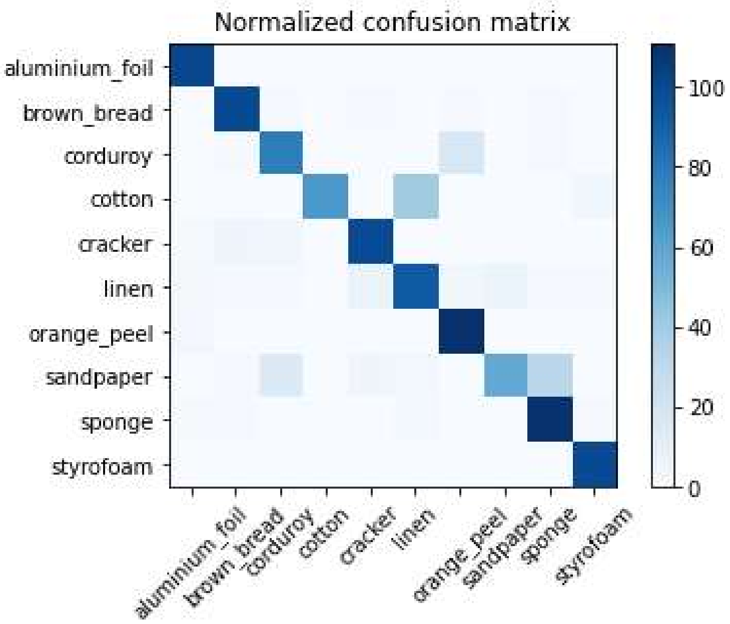
Table3 presents a comparison of the classification results obtained by handcraft methods (LBP, GLCM), and those obtained by based learning methods (VGG, RES-NET and our proposed CNN method) witch have applied on the data set.

The results show that best rates given from learning-based methods where VGG and the proposed model give accuracies (85% , 84.35%) respectively, whereas, RES-net model gives 56.54%, which is an acceptable accuracy even it is low compared to the first ones (may be the depth architecture of RES_net is not suitable to train a multiscale dataset).

For the handcraft methods (LBP and GLCM) respectively gave 58% and 40.38% as an accuracy of the classification. The handcrafted-based results are less than learning-based accuracies.

Using learning-based methods in extract features from the texture is more easy and practical, which makes the classification accuracy more effective and robust.

The confusion matrix is one of the most powerful tools for predictive analysis in machine learning. A confusion matrix gives information about how our classifier has performed, petting properly classified examples against misclassified examples.




```
Normalized confusion matrix
[[1.   0.   0.   0.   0.   0.   0.   0.   0.   0.  ]
 [0.   0.96 0.01 0.   0.01 0.   0.01 0.   0.01 0.  ]
 [0.   0.01 0.8  0.   0.   0.   0.18 0.   0.01 0.  ]
 [0.   0.   0.   0.6  0.   0.36 0.   0.   0.   0.04]
 [0.01 0.05 0.04 0.   0.91 0.   0.   0.   0.   0.  ]
 [0.02 0.01 0.01 0.   0.06 0.8  0.03 0.05 0.01 0.01]
 [0.03 0.   0.   0.   0.   0.   0.97 0.   0.   0.  ]
 [0.   0.01 0.13 0.   0.04 0.03 0.   0.51 0.28 0.  ]
 [0.01 0.01 0.   0.   0.   0.02 0.   0.   0.96 0.01]
 [0.   0.   0.   0.   0.   0.   0.   0.   0.   1.  ]]
```

Figure 26: confusion matrix of KT.HP classes, yielded by VGG model

Figure 26 presents the confusion matrix of KTHP-TIPS dataset using the VGG_model where we can see that most images are assigned to their appropriate classes such as: (aluminum-foil and styrform with accuracy of 100%. where the classifier predicts 97% for orange-peel, 96% for sponge and brown bread and 91% for cracker). But there are some outliers when the classifier misclassify sponge as sandpaper and linen as cotton)

A B

Figure 27: (C) the sande cotton and (D) linen textures from KTHP_TIPS

C D

Figure 28: (C) the sande paper and (D) sponge textures from KTHP_TIPS

The figures 27, 28 show the resemblance of the textures and color of cotton and linen beside the resemblance of the textures and color of sandpaper and sponge that is what causes a miss understanding of the system and the low classification accuracy.

General conclusion

In this work, we present our contribution in improving the performance of the color multi-scale texture using morphological mathematic operations and CNN.

Before proposing our method, we present a quick background for reader to understand image, and digital image processing in the first chapter. Than we show a review of recent works in the same topic with setting their different methods. Than we pass into methods chapter were, we provide a description of the proposed method beside definition and explanation of the concepts used in the work (multi-scale texture, morphological mathematic operations, CNN) and how each technique works. Finally, we describe the used materiel starting by the hardware, software and the tested dataset (KTH_TIPS). Moreover, we discussed the suggested approach and the results obtained after validation and testing using different model of CNN.

Our proposed method was effective and robust in extracting different features from different scales of interest of textures and classify it into the appropriate category too by achieving 85% as an accuracy. Based on this result we are propose perspectives for future research works. the further experimentation can be realized in :

- work to make the calculation faster using deeper training (more useful information in short time).
- The Generation of more scales from the input image causes the generation of more useful features on it.
- Test this method for other datasets to assure the effectiveness and crease the accuracy.
- Combine the morphological mathematic operations with other self-learning classifier.

Bibliography

1- VINH TRUONG HOANG ,These: “Multi color space LBP-based feature selection for texture classification” , FRANCE University, 2018

2- Rafael c. Gonzalez, RECHARD e. Woods , “Digital Image Processing” ,second edition ,1992.

3- Shailendra Kumar Dewangan, “Importance & Applications of Digital Image Processing”, researchGate ,jul 2016.

4- Muhammad Ashraq Salahuddin, Fadzilah Siraj, Shahrul Azmi Mohd Yusof “Digital Image Classification for Malaysian Blooming Flower”, researcheGate,October 2010.

6- Mohamed Elgendy, “The Computer Vision Pipeline, Part 4: feature extraction from Deep Learning for Vision Systems”

7- Ali Ismail Awad, Mahmoud Hassaballah “Image Feature Detectors and Descriptors”, springer,2016.

8- Fred Park, “Shape Descriptor/Feature Extraction Techniques”, UCI iCAMP 2011

9-Jian Kang Wu,Mohan S. Kankanhalli,Joo-Hwee Lim,Dezhong Hong “Perspectives on Content-Based Multimedia Systems” pp 49-67,chapter3, (2002) .

10- Carlos Martinez-Ortiz “2D and 3D Shape Descriptors”, Department of Computer Science, University of Exeter 2011.

11- Vinay Modi, “Color descriptors from compressed images”, School of Informatics, The University of Edinburgh, UK, January 2008.

12- Michael A. Wirth, Ph.D “Texture Analysis”, University of Guelph, Computing and Information Science 2004.

13- Vinh TRUONG HOANG ,“Multi color space LBP-based feature selection for texture classification”, University Lille Nord-De-France, 2018.

14-Mudassir Rafi1, Susanta Mukhopadhyay “Texture description using multi-scale morphological, GLCM”, Springer Science+Business Media, 2018.

15- AlessandroLedda,“MathematicalMorphologyinImageProcessing”,researchGate, aug2014.

16- Ravi Srisha , Am Khan ,“Morphological Operations for Image Processing” : Understanding and its Applications, researcherGate, december2013

18- Robert Fisher, Simon Perkins, Ashley Walker, ErikWolfart “Number plate detection using double segmentation ” researcheGate, January 2010.
19- , Anne Solberg, Gonzalez and Woods, Morphology, Chapter 9, 2011.
20- Jiuxiang Gua et al ,“Recent Advances in Convolutional Neural Networks”, researcheGate,December 2015.
23- William BokuiShen “Convolutional Neural Networks for Visual Recognition” course, spring 2019
25- Jiuxiang Gua, Zhenhua Wangb, Jason Kuenb “Recent Advances in Convolutional Neural Networks”,2017
28- Jianxin Wu,LAMDA Group “Convolutional neural networks”, National Key Lab for Novel Software Technology,Nanjing University, China,may 2020
29- Asifullah Khan1, Anabia Sohail1, Umme Zahoora1, and Aqsa Saeed Qureshi1 “A Survey of the Recent Architectures of Deep Convolutional Neural Networks”, springer,2020.
30- Sumit Saha, “A Comprehensive Guide to Convolutional Neural Networks”, twoards data science, 2018.
31- Mårten Björkman, Computational Vision and Active Perception Laboratory (CVAP) Dept. of Numerical Analysis and Computer Science KTH, SE-100 44 Stockholm, Sweden, may 2004.

32- G.P. Zhang, Neuarel Network for classification —PART C: APPLICATIONS AND REVIEWS, VOL. 41, NO. 6, IEEE, NOVEMBER 2011.
33- ATLILI.S, YAZI.K, ZEKRI.N, “An Ear Recognition System with LBP”, UNIVERSITY KASDI MERBAH – Ouargla, 2018.

Webgraphy:

5 - https://deepai.org/machine-learning-glossary-and-terms/feature-extraction

17- https://northstarwww.dartmouth.edu/doc/idl/html_6.2/Determining_Structuring_Element_Shapes_and_Sizes.html 04-08-2020

22- https://la.mathworks.com/solutions/deep-learning/convolutional-neural-network.html 05-08-2020

24- https://intellipaat.com/deeplerning 05-08-2020

26- https://towardsdatascience.com :visited 05-08-2020

27- medium.com/intuitive-deep-learning , Joseph Lee Wei En ,Feb 2019 visited 10-08-2020

34- www.mathworks.com, /help/deeplearning/ref/vgg19.17/09/2020.23:42., 17

58