Arab American University

Faculty of Graduate Studies

Image Classification in Cultural heritage

By

Muhammed Derar Ali Saffarini

Supervisor

Dr. Muath Sabha

This thesis was submitted in partial fulfillment of the requirements for the Master's degree in computer science

March / 2019

II

Image Classification in Cultural heritage

By

Muhammed Saffarini

This thesis was defended successfully on 9-3-2019 and approved by:

Committee members	Signature
1. Supervisor Name: Dr. Muath Sabha	……………………
2. Co- Supervisor Name:	……………………
3. Internal Examiner Name: Dr. Mohammad maree	……………………
4. External Examiner Name: Dr. Yousef Awwad	……………………

Declaration

I declare that this thesis entitled " Image classification in cultural heritage " is my own work and has been composed solely by myself and does not contain any work from other researchers and has not been submitted for any other degree or scientific qualification, and I except that all the references are done correctly.

Dedication

Proudly, I dedicate my thesis to my parents, as I always feel their prayers in all aspects of my life, I also dedication my thesis to my sisters, friends, colleagues who are always willing to provide any support.

Acknowledgements

My deepest gratitude goes to Allah Almighty; who is providing me with strength and health, with wisdom and inspiration, with support and patience to let me overcome obstacles throughout my difficult times.

I am immensely indebted to my supervisor, Dr. Muath Sabha for the invaluable advice, patient and guidance he provided throughout my period. I would never be able, by myself, to accomplish this work without his support and guidance.

My gratitude goes also to my mother, to my father, who prayed days and nights to enthusing me during my study.

To all friends and relatives who have supported so genuinely pleased and interested over my studying period.

Abstract

In this thesis, an automated supervised image classification technique specifically for classifying images in the cultural heritage domain is developed. In general, most of the image classification techniques are used for known and semi-visible models that has known objects to detect, while the developed technique classifies images according to a particular date, culture, people and historical age. The proposed technique is comprised of two stages, the first is extracting features using unsupervised segmentation technique, then the unsupervised classification stage. The developed technique uses only hue from the CIE LAB color space for segmentation and K-means is used for Clustering. Some segments are merged to get the result of the cultural heritage to which it has the most relevance for these segments. In the learning phase, common features were extracted and then compared their histograms, then they were categorized accordingly. Finally, adjacent columns in the histograms were merged to reduce the complexity of the algorithm. Finally following the technique and applying it on a repository of cultural heritage images, reduced the complexity of the algorithm and get accuracy approximately to 70% for 20 cultural heritages, each of which has about 150 images from 3200 images in whole dataset.

Table of Contents

Table of Figures

List of Tables

Table of Abbreviations:

K-NN	K nearest neighbor
KMC	K Mean Clustering
SVM	Support Vector Machine
HSV	Hue, Saturation, Value
CIE	International Commission on Illumination
LAB	Lightness, green–red and blue–yellow color components
BST	Binary Search Tree

Chapter One: Introduction.

This thesis related to five main topics is the computer vision, color space, machine learning, image classification and culture heritage.

Introduction to Computer vision:

Computer Vision is a growing field of image processing and artificial intelligence (machine learning and deep learning). Its aim is to make the machines and computers "see" and "recognize" the visual world in a smart and meaningful way in the same way that how the humans see and recognize. The objective of this field is to develop the ability to recognize, classify and clustering of these images. And to get to know people and recognize their behavior and the interaction between other objects and people.

There are many techniques and models for Computer Vision that make the computers smartly explain, understand and recognize the images and videos under different scenarios. Nowadays, computers have the ability to detect and recognize people faces with accuracy that is considered perfect [9]. Additionally, computers can detect and recognize various actions in real time. The increase of using computer vision is growing at an exponential rate and the domain is beginning to mature, with a huge influence on human life. Currently, Computer Vision is present in computer games that recognize player actions to interact with the game which has been very profitable. Moreover, Computer Vision is also present in the security field with face recognition, eye recognition or voice recognition as an authentication way instead of password. Computer Vision links the practical and theoretical results in many fields, like computer science, mathematics, statistics, machine learning, and neuroscience (brain / cognitive science).

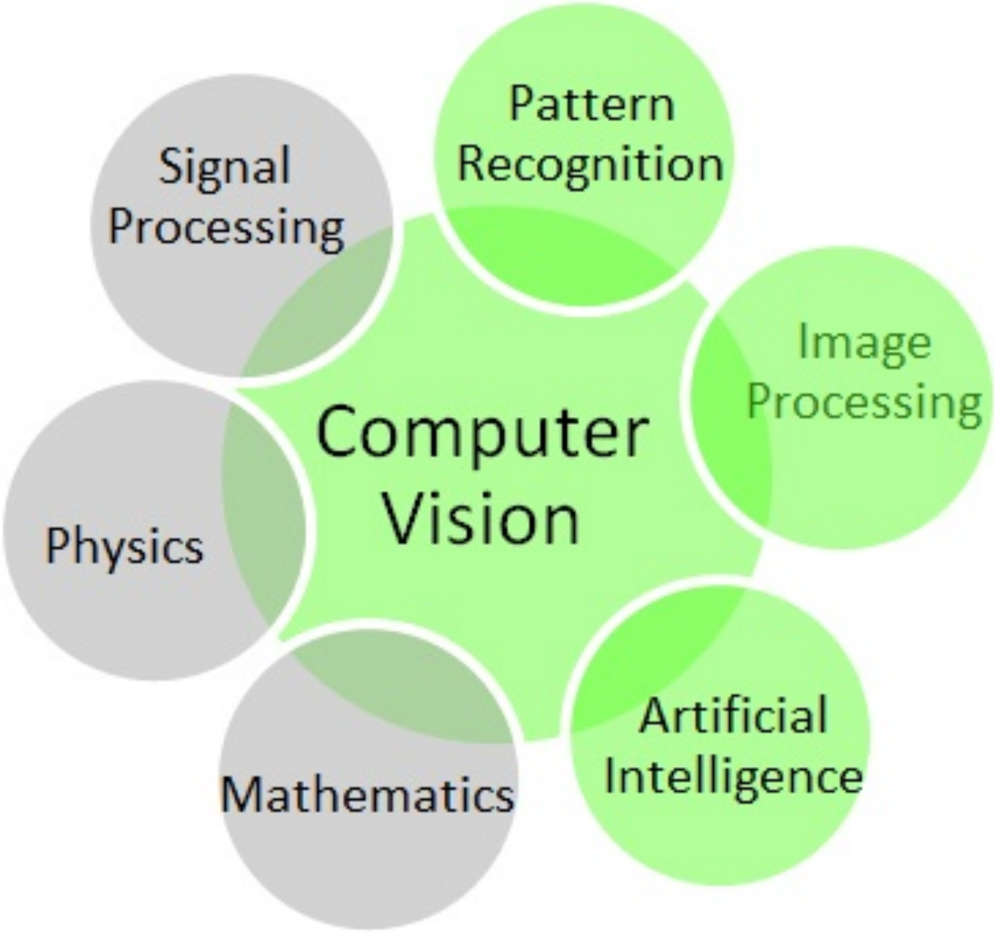

Figure 1 Branches of Computer Vision

As show in the figure 1, computer vision combines several branches like Image Processing, Mathematics, Artificial Intelligence and pattern recognition the main branch of this research.

Introduction to Color space:

A color space is a space which can be made to view and represent many colors similar to the human vision (but, most often, they just represent a subgroup of these colors that we can see). The colors that are often referred to, are the colors that the humans' visual system can perceive and all the possible mixtures of all the colors from the visible spectrum.

The domain of potential colors that may be viewed and represented with a special system is called the gamut. A color space in its simplest form, can be seen and viewed as the mixture of three primary colors (Red, Green and Blue) in color scheme. Mixing any of these colors with each other may create a wide domain of colors (this is the core of the RGB color space).

Any color space can't represent and view all the colors like how human see these colors, color spaces is more restricted or more limited than any color space that can view and represent all perceivable colors. The full gamut is the gamut that would represent all the colors the human can see. In fact, technically, this gamut is called the gamut of human vision.

The color space is the range of colors that can be created by the primary colors of pigment and these colors then define a specific gamut. Color spaces are very important in computer vision as they represent every pixel inside the image which has a value from these color spaces. The RGB color space can't give the hue (hue is the true color), it just gives the color of three components R, G and B. however the problem in the light that falls on the color that effect on the true colors, that goes to the lightness effect on the results of classification and segmentation. In this research, we use another color space, the CIE Lab that define bellow, and that is because it is closer resemblance to human perception and it has three channels, the first is Blue towards Yellow (Opponents), the second is Green towards Red (Opponents), and the third is the lightness. The lightness has special channel and the hue get along without lightness.

CIE LAB color space is one of the best color spaces, Lab color space is a 3-dimentional color space with the first dimension is lightness (L) and the rest are A and B for the color dimensions. In this study, LAB color space is used because that this color space includes all the colors in the spectrum full gamut as shown fin figure15. Which means all color (even colors outside of human perception) are included. The LAB color space is the most accurate means of color representation and it is also device independent. This accuracy, precision and portability makes it appropriate in several applications such as; automotive in self driving, printing, textiles, and plastics. The LAB color space created three channels. that has a huge gamut. It channels Blue towards Yellow (Opponents) and Green towards Red, which is almost as human perception of things.

Introduction to Machine Learning:

Machine learning (ML) is a field of Artificial Intelligence (AI). The aim of ML in generally is to understand the data structure and produce suitable models that are used and understood by humans.

Although ML is employed in various computer science application domains, it is different from classical computational techniques. In classical computing, algorithms are groups of explicitly programmed orders utilized by computers and machines to calculate or solve problems. While ML algorithms and techniques main goal is to allow machines and computers learn from training data inputs and use analytical and statistical methods in order to produce output values that is within a specific range of values. Subsequently, ML allows the computers and machines to make decisions based on the training set as shown in the below figure.

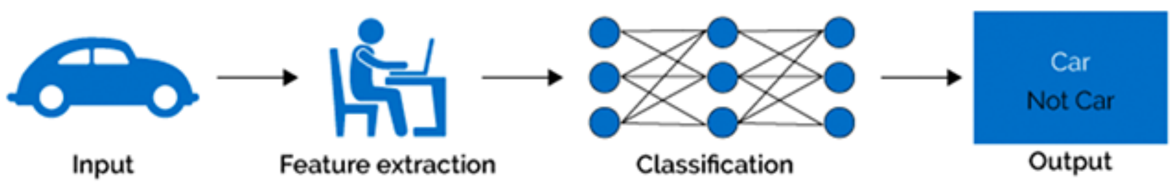

Figure 2 Machine learning procedure (supervised)

Many fields have benefited from ML, such as; face recognition which is used in social media platforms to help users share photos and tag their friends automatically. Another technology is Optical Character Recognition (OCR) which converts images that contain a text into texts and string values like google translate. Recommendation Search Engines also benefited from ML, which suggest programs or films to watch based on user preferences. Self-driving cars is another field that also relies on ML.

Introduction to Image Classification:

Image classification is the process of putting an image with unknown label to one of labeled classes. This problem is considered as one of the root problems in computer vision. There are two main types of classification; unsupervised clustering where the system analyses the image without the need of the user to provide samples, where it takes number of cluster and it clusters the image by using specific equation like Euclidian distance, and supervised learning which is based on selecting samples by the user applying the training set, after that the system will predict the class that an image belongs to. For humans, the classification method consists of many stages, starting by image segmentation, extracting the features from the image, and then recognition and classification of the image, that is, assigning the testing image to the right category.

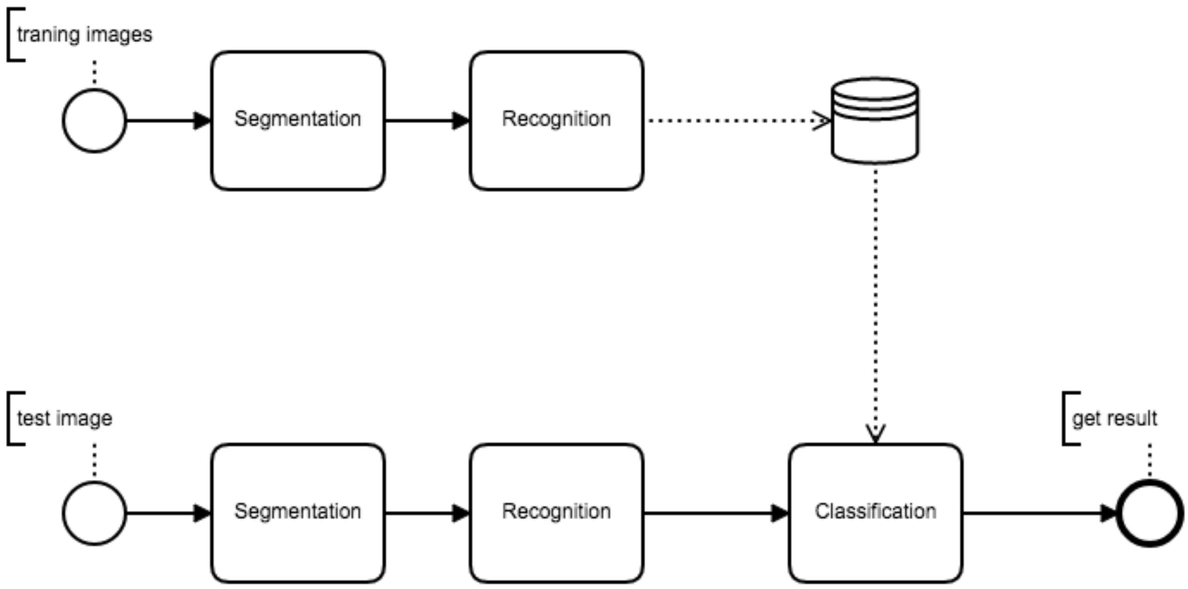

Figure 3: Block diagram for image classification

As shown in figure 3, in general, image classification consists of the stages below:

- *Training set*: the input will be N images; each image is labeled with one of K distinguish classes. This data is referred to as the *training set or dataset*.
- *Learning*: use the dataset to identify the features in each class. This step is called *training a classifier* or *learning*.

- *Evaluation*: estimate the quality of the classifier by providing prediction labels from dataset for a new set of images that it has never seen before. Then the true labels of these images are compared to the ones predicted by the classifier.

Introduction to Cultural Heritage:

Cultural heritage is the heritage of the material and intangible artifacts of a group or society inherited from previous generations, preserved in the present and given to the benefit of future generations. Each cultural heritage has many features such as Pharaonic culture.

Many of the image classifications are used for known and semi-visible object and shapes, while the proposed model classifies images related to a particular date, culture, and historical ages. The challenge is defining a particular culture, where there are many objects and symbols that are unique for each era (like the figure 4 for a roman remnant). Another challenge is finding some symbols that can help identifying image class. Every category has common features that distinguish this culture from another. For example, The Roman culture has its own sleepwear and common architectural forms, and the images is classified accordingly these features.

Figure 4 Roman Cultural heritage, features that distinguish the Roman (clothes and building)

Our algorithm utilizes an unsupervised model for segmentation the images using K-Means clustering for finding the number of clusters that closer to features and use the supervisor (feature in our methodology) to classify the test image to correct class (specific cultural heritage) using K-NN algorithm and Binary Search Tree (BST) is used in recognition and classification stage which reduces the time complexity for some process in whole process.

Classification vs Clustering

Classification and Clustering are the two kind of learning techniques which identify objects into sets by one or more features. It seems that these processes are similar, but there is a difference between them in context of data mining.

the difference between classification and clustering is classification is used in supervised learning technique where predefined class or labels are assigned to instances by properties, on the contrary, clustering is used in unsupervised learning where similar instances are grouped, based on their features or properties.

When training is provided for the system, the class label of the exercises is known and then tested, and this is known as supervised learning. On the other hand, unsupervised learning does not include training or learning, and the training sample is not previously known

COMPARISON	CLASSIFICATION	CLUSTERING
Basic	This function compiles the data into one of many already selected categories.	This function maps data in a multiple group where the order of data elements depends on similarities.
Involved in	Supervised learning	Unsupervised learning
Training sample	Provided	Not provided

Table 1 comparison between Classification and Clustering

Problem statement:

Many of the image classifications are used for known and semi-visible object and shapes, but we will classify images related to a particular date, culture, peoples and historical ages by extraction the common features like clothes and building of the image and then classify the features for which culture it belongs to.

The problem is that the process of defining a particular culture that there are many things and symbols that indicate each age and also not found some symbols and the clear through which the identification of any era belongs to this picture.

Since there are many cultures that have things in common with other cultures and also often cultural images have a large size ratio and therefore the process of learning or examination will need more time than others images for small features.

Purpose of the research:

In this thesis, an automated image classification technique specifically for the cultural heritage images is developed. In general, Image classification is labeling an image with a specific category from a set of categories. Most of the image classification techniques are used for known and semi-visible models, while the developed technique classifies images according to a particular date, culture, people and historical age

Our contribution:

an automated supervised image classification technique specifically for classifying images in the cultural heritage domain is developed, that we used CIE LAB color space instead on RGB color space that has full gamut (large set of color) and use Local Min-Max algorithm to find the number of clusters inside the image and use it inside K-Means Clustering to segment the image , after that apply merging histogram column for layer to get 29 that these 29 values represent the segment and then give constant time complexity after that put it in BST that give small time (Log n), after that apply K-NN for each 29 values that represent the datasets to classify the segment to best culture , this technique give accuracy approximately 69% and fast time.

Outline:

The content of this thesis is arranged as follows:

Chapter two: Literature review. In this chapter we introduce segmentation, classification and introduce many algorithms for classification for cultural heritage

Chapter three: Methodology. In this chapter, we discuss our method that is proposed to solve the problem of Image Classification in Cultural Heritage.

Chapter Four: Results. This chapter includes a brief description about the database used in our experiments, in addition to all conducted experiments. The produced results in comparison with the related algorithms and the evaluation are discussed here.

Chapter Five: Conclusion and Future Work. A conclusion of all work done is given in this chapter

Chapter Two: Literature Review

The segmentation, recognition and classification topics are common and there are many papers and techniques discussed in one of them (recognition, segmentation and classification) or combine them.

Image Segmentation:

One of these papers called A Color Space Performance Comparison in the Processing of Color Textured Images: RGB vs. LAB [1] that show the performance between two color spaces (LAB and RGB) that treatment with it in image segmentation by color texture that the topic is compares between RGB and CIE LAB color space in expression of its influence in the ideal image processing task.

The RGB color space is nearly globally accepted by the image processing research society to display images, the prime purpose being that RGB data is easily ready as the explicitly data created by the camera. There are, however, cognitive more regular spaces, such as Lab and Luv, where the difference in color measured in analogy with the human cognitive of such a variation. Whether the use of such a color space would supply better outcomes in image processing functions (segmentation/classification).

In this research RGB and Lab in the expression of their performance in the ideal image processing function, such as, the segmentation of images. There has been only restricted act on the color sides of textured images.

And there is the comparison between different color spaces [20], since it has the capability to show the outcomes in a way that is much most nearest to the human eyes. Selecting a suitable color space is a very necessary case for color image segmentation process. mostly LAB and HSV are the two much selection color spaces. In this research a comparative test is performed between the HSV and LAB color spaces with respect to image segmentation. For scale their performance.

Another algorithm “Color image segmentation using global information and local homogeneity” [2] this research suggests a new technique of color image segmentation taking both global information and local homogeneity. The method performed the average shift algorithm in hue and density sub-space of HSV. The cyclic property of the hue component is also considered in the suggest technique. Tests on natural color images display promising outcomes.

Another paper is Hill-Climbing Algorithm for Efficient Color-Based Image Segmentation [11]. In this paper considered as a new way to segment the images that based in hill-climbing algorithm ,did the segmentation by finding the histogram for three layer of the image ,after that find the local maximum values in the histogram for three layer H,S and V in image , then the algorithm connects image pixels with local maximum values , this research assume that the system doesn’t know anything about number of segment or cluster or the content inside the images.

And this paper use HSV color space because it simulates the human perception of the colors that can get the pure colors and treats with hue and saturation and lightness and not using RGB gradient, it will create a histogram for these three components and find the local maximum value.

However, this technique has many drawbacks like A flat area of the search space in which all neighboring states have the same value and Make a big jump to try to get in a new section.

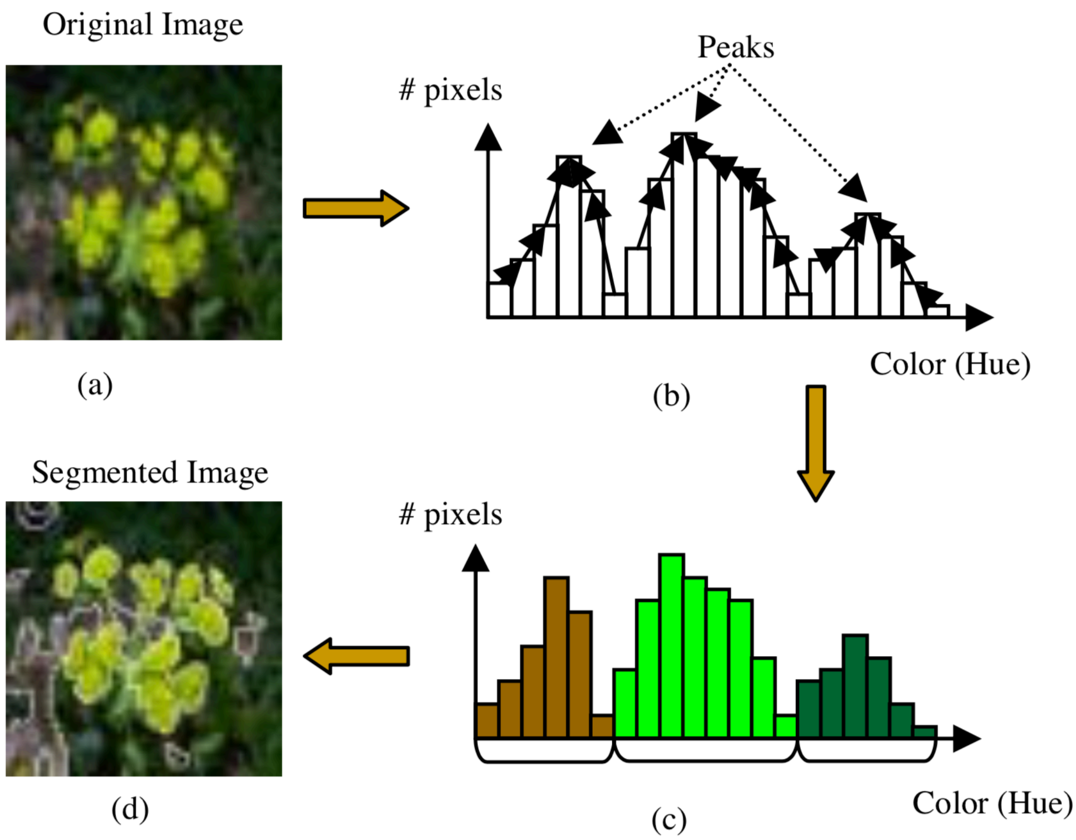

Figure 5: the paper strategy of segment the image: (a)the original image (b)find the histogram for just the hue and find the peaks (c)segment the histogram into classes depends local max value(d)the segment image

In another research “image Color Thresholding Method for Image Segmentation of Natural Images” [4] Most of the thresholding procedures involved setting of boundaries based on grey values or density of image pixels. In this research, the thresholding is to be completed depend on color values in natural images. The color thresholding methodology is being executed depending on the acclimation and little altering of the grey scale thresholding techniques. Multilevel thresholding has been linked to the RGB color data of the get object from the surrounding area and another object. various natural images have been used in the research of color data. Outcomes displayed by using selected threshold values, the image fragmentation methodology was able to divide the object from the surrounding background.

There is many research's that make use ML techniques to segments the images like "K-Means Cluster Analysis for Image Segmentation" [16] applies K-Means on a number of 2, 3 and k-cluster color images for (k>3) in images, images have been used to achieve efficient function. Silhouette analysis help the peaks for set k-cluster image however there is some drawbacks like Selection of optimal number of clusters is difficult and Selection of the initial centroids is random. and more and more depends in texture of color [18] [19].

Another paper is "Segmentation using saturation thresholding and its application in content-based retrieval of images" [5] that analysis several of the visual characteristics of the HSV color space and improves image segmentation methodology utilizing the outcomes of their analysis. In this technique, features are taken either by selecting the hue or the density as the dominant property depend on the saturation value of pixel. We perform content-based image retrieval by object-level matching of segmented images. A freely usable web- enabled application has been developed for demonstrating our work and for performing user queries.

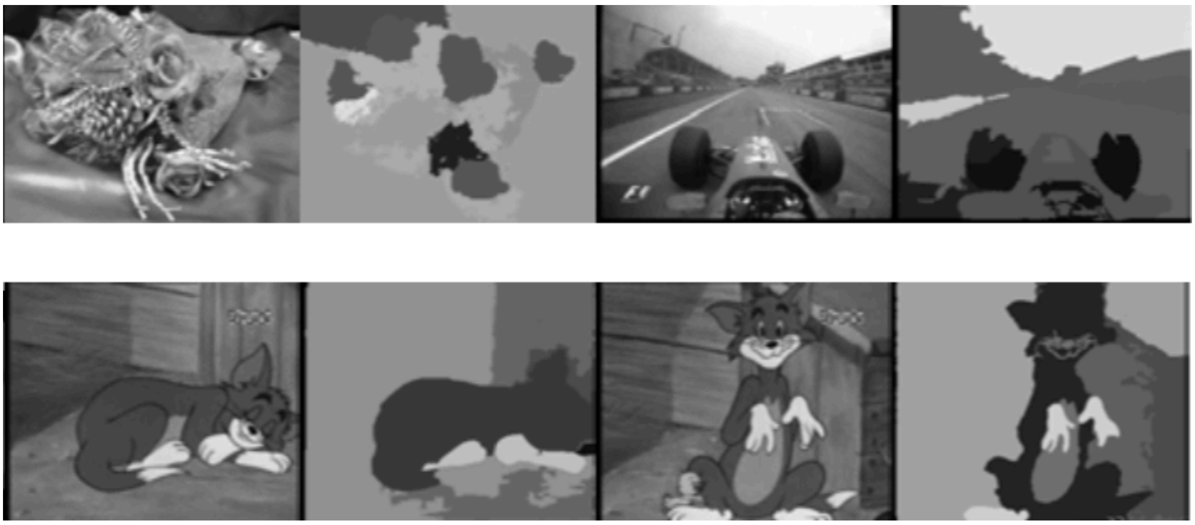

Figure 6 Segmentation using saturation thresholding

Another research for image segmentation called Statistical Region Merging (SRM) [10]. The algorithm explores a statistical basis for a process often described in computer vision, Segment the image by region combined after a specific order in the selected areas. The algorithm is used to estimate the values in the regional extension and classify them together based on the standardized standard resulting from a smaller list. Some best examples would be makes in image processing classifying a class of adjacent pixels based on their shadows that are placed in a particular threshold (Qualification Criteria) However it has some drawbacks like It is a local method with no global view of the problem, Sensitive to noise, computationally expensive and Unless the image has had a threshold function applied to it, a continuous path of points related to color may exist which connects any two points in the image.

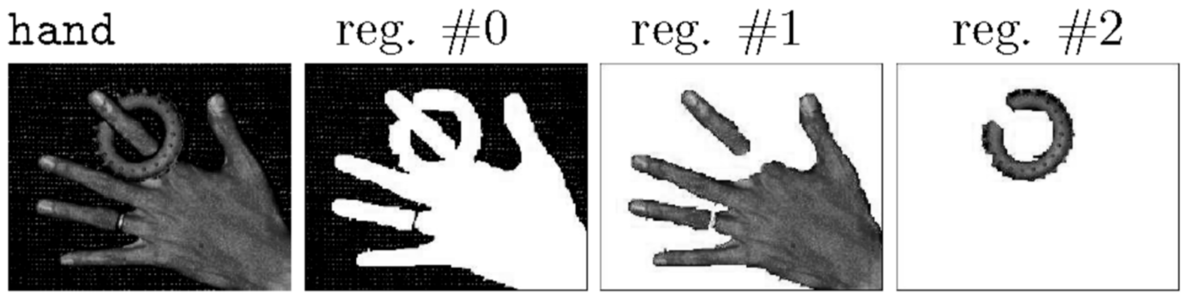

Figure 7 SRM result

Another research for segmentation is “Color Image Segmentation” [22] a new way to automatically segment images, called JSEG, is presented. First, colors in image are quantized to several representing classes that can be used to differentiate regions in the image. Image pixel colors are exchanged by their identical color class labels, so shaping a class-map of the image. A standard for "good" segmentation is suggested using this map. Apply the criterion to local separators in the grade-map results in "Image-J", where high and low values correspond to potential region boundaries and regional centers, respectively, respectively. An area increasing technique is then used to split the image depend on the multi-scale J-images. tests display that JSEG good segmentation outcomes on images.

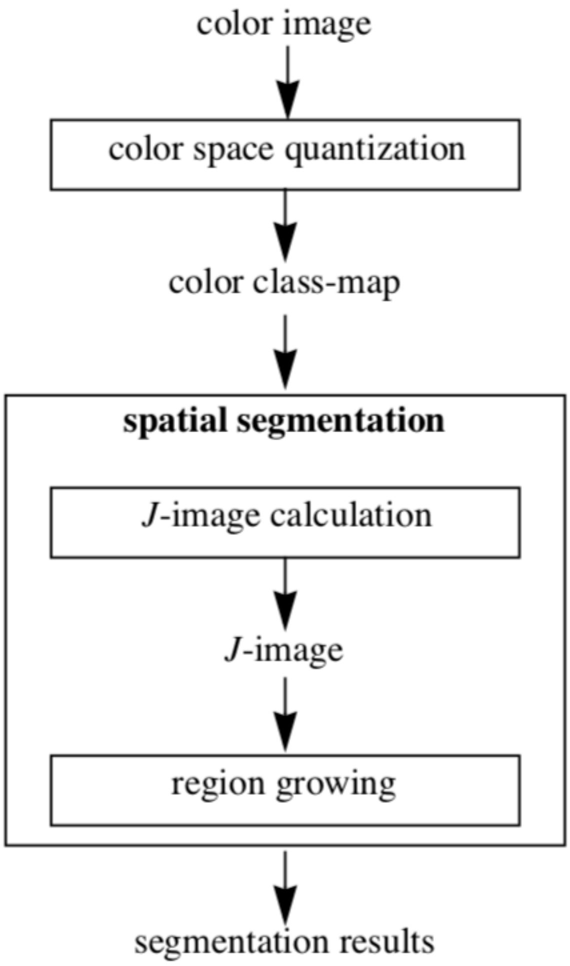

Figure 8 JSEG steps

Another algorithm "Automatic Image Segmentation using Threshold Based Parameter and Genetic Algorithm" [23]. In this article, discussed new threshold-based information parameter for image segmentation over Genetic Algorithm. This paper applied Genetic Algorithm because it is able to select the best number of areas of segmentation. The processing method utilize an evolutionary standard and create best segment selection. In suggested method, Genetic Algorithm is applied in order to select evolutionary best segmented image based on a new information-based parameter via threshold. This technique is performed only on gray level images.

Another research for segmentation is "Grayscale Image Segmentation Using Multilevel Thresholding and Nature-Inspired Algorithms" [3] Recently Multi-level image thresholding plays a critical part in analyzing and explaining the digital images. In the last research, it was discovered that traditional tried research methodology is the time taken to increase the number of thresholds. To fix the issue, many nature-inspired algorithms (NAs) which can provide high-quality solutions in time have been used for multi-level thresholding. This article discusses three ideal types of NAs and their hybridizations in fixing multi-level image thresholding. A novel hybrid algorithm of gravity search algorithm (GSA) with genetic algorithm (GA), called GSA-GA, is suggested to discover the optimal threshold values. The selection objective functions in this article are Kapur's entropy and Otsu standard. This article linked testing on two famous test images and two real satellite images using different numbers of thresholds to estimate the performance of various NAs.

Image Classification:

Another algorithm for recognition the images is Histogram and phase - only correction [9] that based on the plot of frequency of occurrence of various gray levels of an image is known as histogram. Since images have many types, so their histograms depend on these types binary images that resented as black white image and gray scale images, for obtaining face image histograms, in the first place, the 256-level histogram must be created, and then each 8 contiguous levels are represented as one value in order to simplify calculations and comparison as well as to accelerate the image processing without affecting the quality of the image. After this process, there will be 32 pins (peaks) only instead of 256 pins, which peaks are the sum of the 8 levels frequencies, this algorithm in general based in face recognition but we will enhancement it and using it in segment that I'll take from images and it contain second part call phase shift that depend in 2D Discrete Fourier Transform (2D DFTs) .

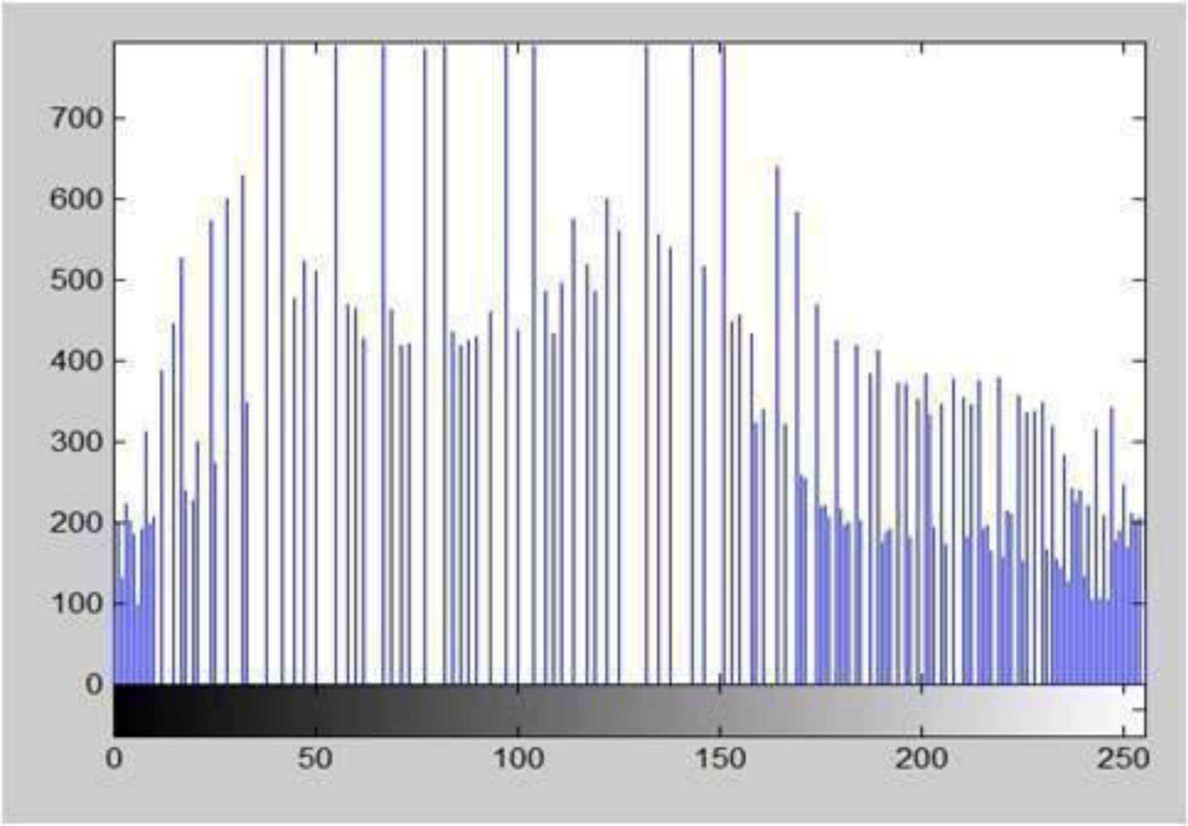

Figure 9 histogram of poc

Another algorithm for recognition called Object recognition from local scale-invariant features [12] [13] that depend on SIFT (Scale invariant features transform) it a new technique that that make distinct fixed features from the image that we can used to match the different views of the objects, the extraction of the feature can be done after some operation like rotation and scaled.

Another algorithm for recognition called Object Detection Combining Recognition and Segmentation[14] ,, in this paper divide the operation into two steps the first is the top-down approach used for recognition and the second is bottom -up approach used for segmentation , then this two steps called as the first is the hypothesis and the second is the verification in hypothesis they improve a Shape Context Feature that that create the group of hypothesis like object location and ground mask ,..etc. , and in the verification step compute list of possible segments can create by the hypothesis step .

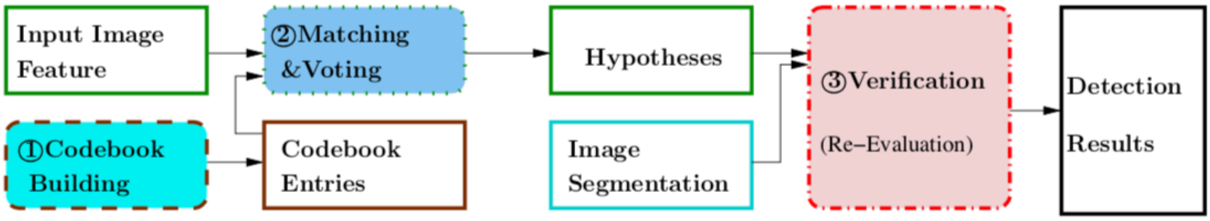

Figure 10 Block diagram for two step hypotheses and verifications

Another research for classification that uses K-NN algorithm is called "Texture Features and KNN in Classification of Flower Images" [15], the proposed algorithmic model is based on textural features such as Gray level co-occurrence matrix and Gabor responses. A flower image is segmented using a threshold-based method. The data set has different flower species with similar appearance (small inter class variations) across different classes and varying appearance (large intra class variations) within a class However we faced face the trade-off between sample size and accuracy, that applicable only to small texture class's and can produce boundary artifacts if the input texture is not tillable.

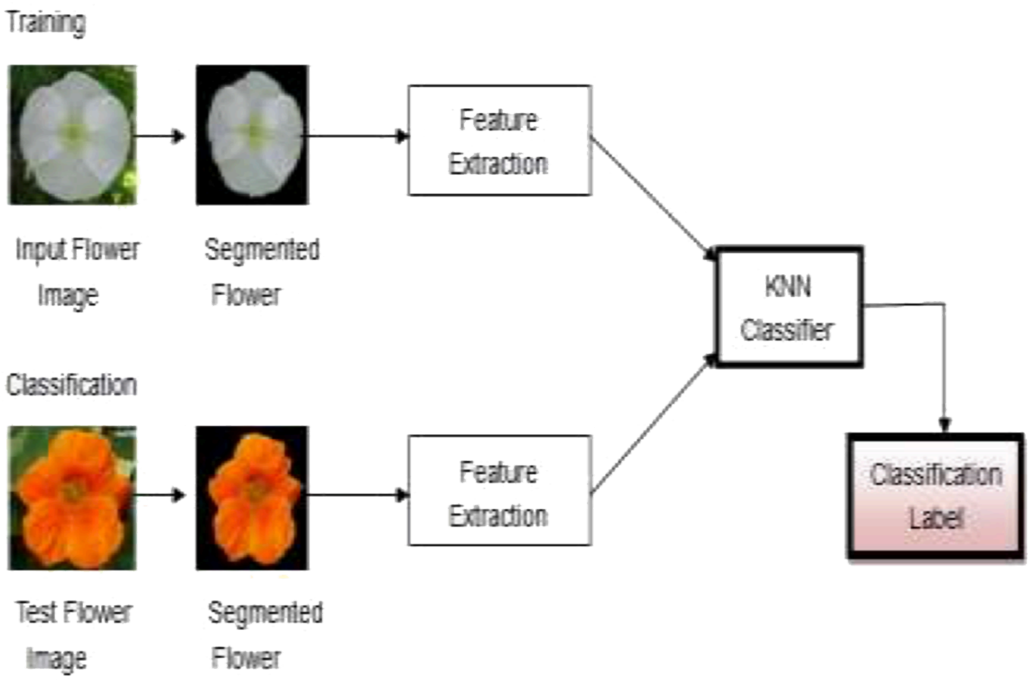

Figure 11 model of paper using K-NN algorithm [15]

Another research for image classification is called “A Framework for image classification”[7] that depends on annotation of the image and the feedback that comes from the humans ,then at the first the human provides list image annotations ,after that the classifier apply these on the training set ,after that the classifier will take a new test image and trying to predict the annotation on this image ,after that human give a new feedback to the new image were added in the trainset which lead to the group training again and so on.

Another algorithm for classification is “Deep Multiple Instance Learning for Image Classification and Auto-Annotation” [24] recent development in learning deep representations has demonstrated its wide applications in traditional vision tasks like classification and detection. In this paper, attempt to model deep learning in a weakly supervised learning (multiple instance learning) framework. the setting in this research, each image follows a dual multi-instance assumption, where its object proposals and possible text annotations can be regarded as two instance sets. This research thus designs efficient systems to use the MIL property with deep learning technique from the two edges; this research also attempts to jointly learn the link between object and entourage suggestions. This research linked wide test and show that this research powerless supervised deep learning framework not only realize masking execution in vision missions including classification and image entourage, on both widely applied benchmarks such as PASCAL VOC and MIT Indoor Scene 67 as a figure 12.

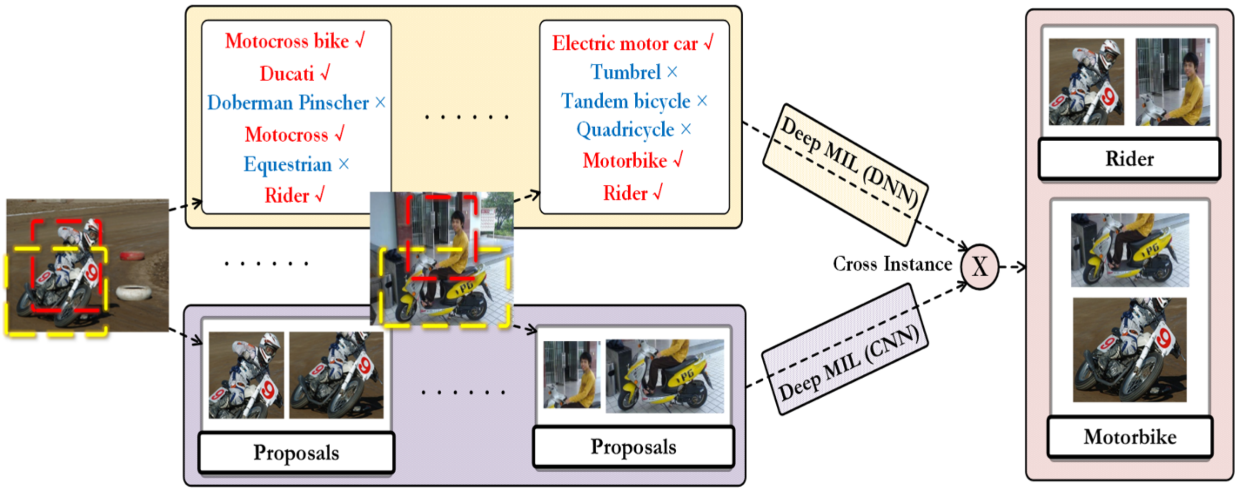

Figure 12 design efficient systems to use the MIL property with deep learning technique from the two edges

Texture synthesis also is used for pattern recognition, and so image segmentation, by recognizing the textures filling the features. [21]

Chapter Three: Methodology

Introduction:

Many of the image classifications are used for known and semi-visible objects and shapes, but we will classify images related to a particular date, culture, peoples and historical ages.

In our algorithm we have two types of images, the first is the training set that describes a specific cultural heritage where each of them will have a “group label” that each group of these images will have a known label (known cultural heritage), and the second type of images is the test images that we don’t know any information about it and don’t know for any cultural heritage they belong to, however these images have some features that we can determine for which cultural heritage they belong, then our methodology will decide for any cultural heritage these test images belong to.

We aim to make the classification for cultural heritage by extracting the common features from each heritage and each collection is specifically placed in a group for each culture. Our method will put the common features in groups and use these groups to determine the new images to which group they belong, and we use many steps and equations to reduce the time complexity for every step to get a best result and less time comparing to another method that checks the whole image.

We created a dataset of different groups of images where each group represents a specific cultural heritage, each group has a collection of images that represents the common features of these cultural heritage. For example, if we take the roman culture, there are many features in this culture such as the column style and the amphitheaters ... etc. We distinguish any new image by extracting the common features in the image and compare them with the most close group that has most common features. In later steps we will explain how we represent these images in groups in the dataset and how to extract the features.

Our methodology -as shown in figure 14- is divided into three main stages, the first is segmentation that will segment the image and extract the common features from it using machine learning by applying K Means clustering (unsupervised), but before this stage we convert the original color space to CIE LAB color space. The second is the recognition that analyses the segments and convert them to vectors. And the third is classification, to determine to which group this image belongs, using machine learning by applying K-NN algorithm (supervised).

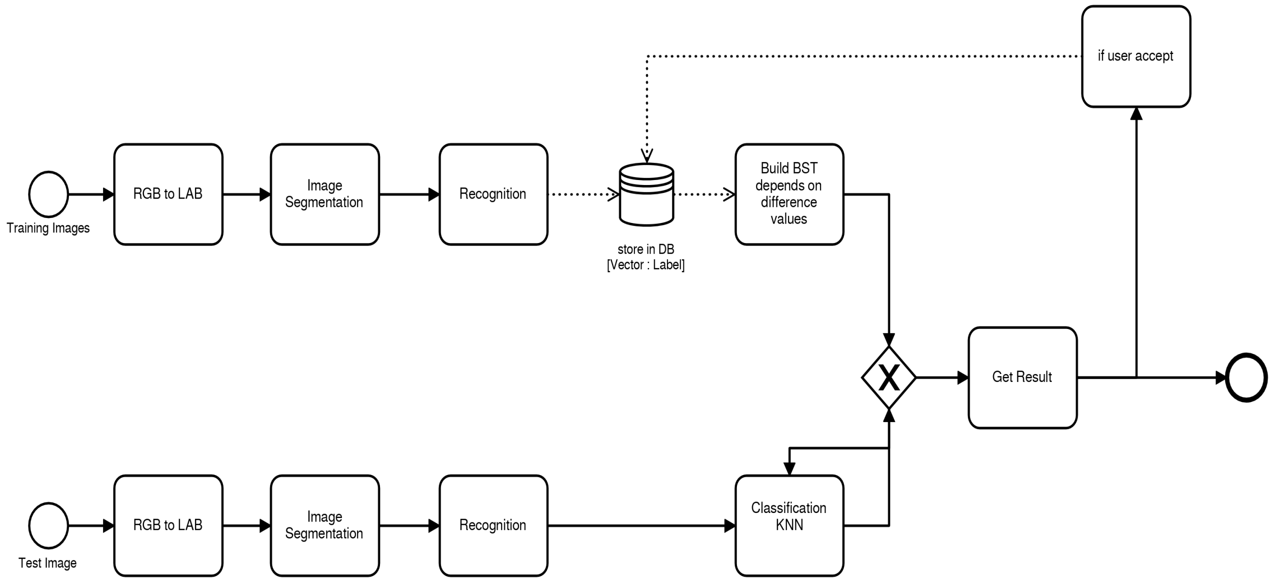

Figure 13 Block Diagram for our Methodology

Convert to LAB Color space:

At the beginning we convert the image color space from RGB to LAB because in LAB color space we have two main channels, the first channel is channel Blue towards Yellow (Opponents) and second is channel Green towards Red which it almost closer to how humans perceive images. By using ab from Lab, we can extract the lightness and the darkness from the image that towards to reduce the effect of shadows, darkness and lightness inside the images., We use pure color and use the mixed of colors inside each layer of the image that will get more accurate in segmentation. Moreover, the LAB gamut (Range of colors) is larger than RGB gamut, so we get a larger set of new colors not found in RGB. Finally, the histogram of each channel of the Lab color space will show us the existence of each color inside this layer by local min and max value in each channel.

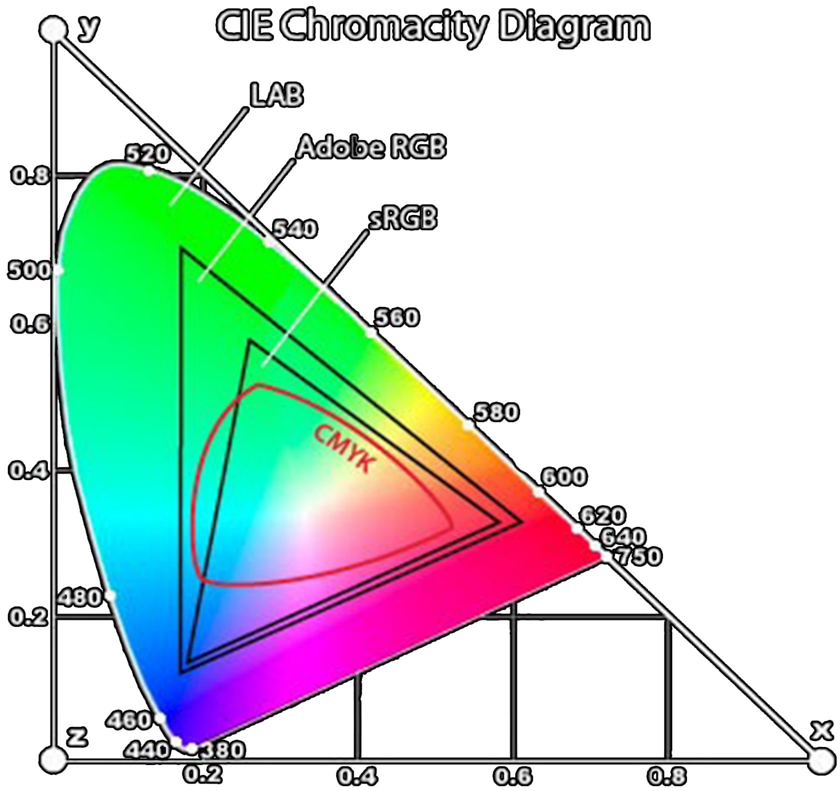

Figure 14 CIE Chromacity Diagram

As shown in figure 15, which explains the CIE Chromacity for LAB and RGB gamut and the difference of color range in each color space.

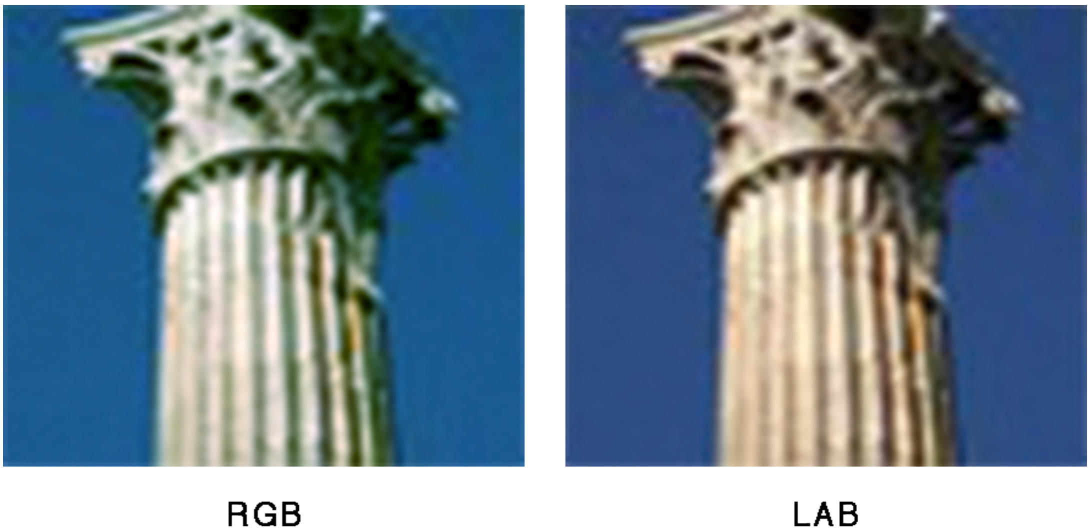

Figure 15 the different between image using RGB color space and image using LAB color space

As shown in figure 15 and figure 16 the color inside the LAB color space has more accuracy of true colors and more precise and vital, which enables us to extract features in a more accurate way.

Segmentation step:

Figure 17 explains the steps of our segmentation process.

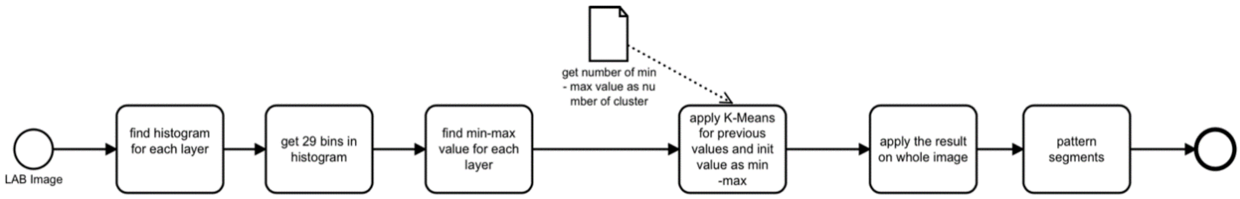

Figure 16 Block Diagram for Segmentation Stage

After convert image from RGB color space to LAB color space, we perform segmentation. At this step, we segment images and take the common features inside them, then at the first we calculate the histogram for each Lab channel of the image. The LAB color space has three layers L,A,B that the range value for L is [0 - 100] (but don't use it), , range value for A is [-86.185 – 98.254] and the range value for B is [-107.863 – 94.482]. Each value in each layer will have a set of values such as in B channel, we may have 30 pixels who have the value 10., then the result will be in every layer group on columns depend on the range for this layer like if we get the lightness layer the number of columns will be 201 columns (range value of B layer). After this step we merge every 3 adjacent columns for every layer that will reduce number of columns in each channel. (that will reduce the time needed to recognize the feature as shown in the next stage from N*M to Constant value) which we will apply the merging until the number of total column be 29 columns for each layer. These 29 values in calculation doesn't take too much processing, and so takes less time and it is not that small for getting less accuracy. The merging done by taking the average of adjacent columns, after that we create vector of 29 value that replaces the old columns in the original layer with new columns in the new vectors, which leads to a new histogram that consist of 29 columns. This will reduce the time complexity in the next step as these vectors will be the new representation for each layer instead of the previous histogram for A and B layer.

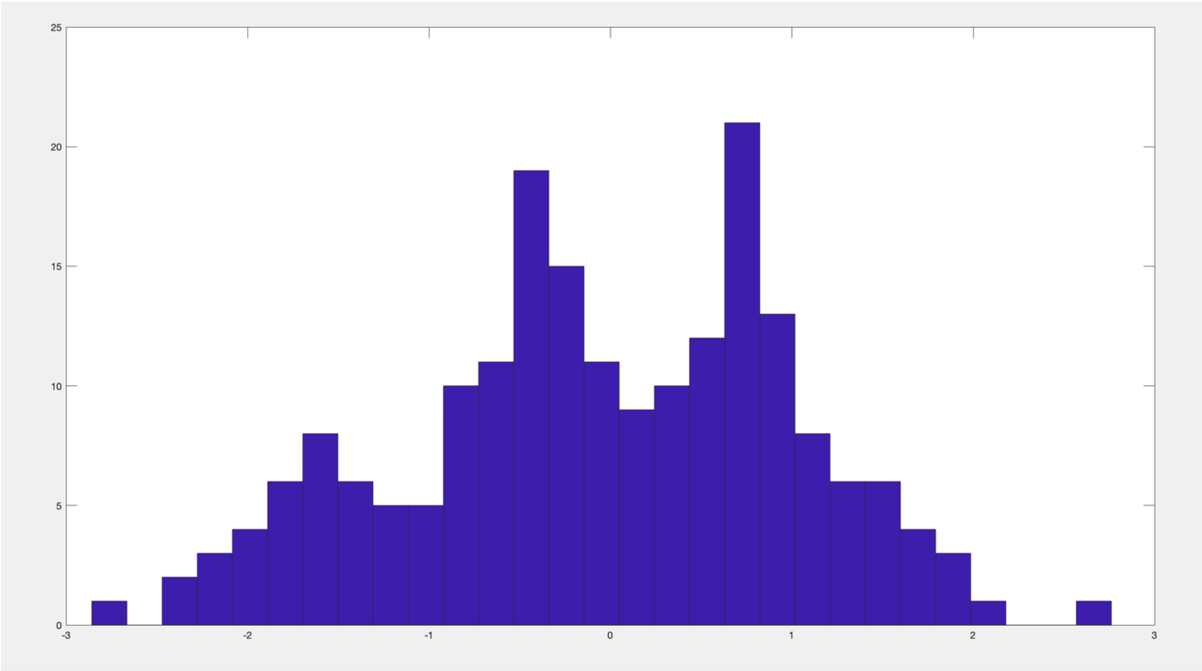

Figure 17 Histogram of the image in figure 16.

Figure 18 shows the result of getting 29 bins [9] out of the whole 256 columns histogram, this means that each column in the new vector (the new histogram) will have a representation value of the original one.

After that we enter these new vectors (the new histogram) to equations that calculate the local minimum and maximum values in these histograms by finding the centers of the previous histograms to convert the columns to groups of points. That means, first we convert the histogram from the columns to the linear chart that contains continuous values, then we apply the local min-max value ,so that we can calculate the local min-max values, the aim of calculating the local min-max values is to see when the color inside LAB layers changes from one color to another. After calculating the min-max values, we store these values in vectors, thus producing these vectors of points that we call clusters (C) that consist of local min-max values for three layers in LAB color histogram of the image. We assume that every cluster will have the values between every contiguous local min and max values in the histogram as shown in figure 19. At this stage, we found that the time for computation will be short and constant, where the time complexity is O(C). We will apply equation of merging columns that gives 20 column not for all the columns in the first histogram, then the number of the local min-max values will the number of clusters that we will use to segment the images into segments to use these segments for recognition and classification as we will see in the next steps.

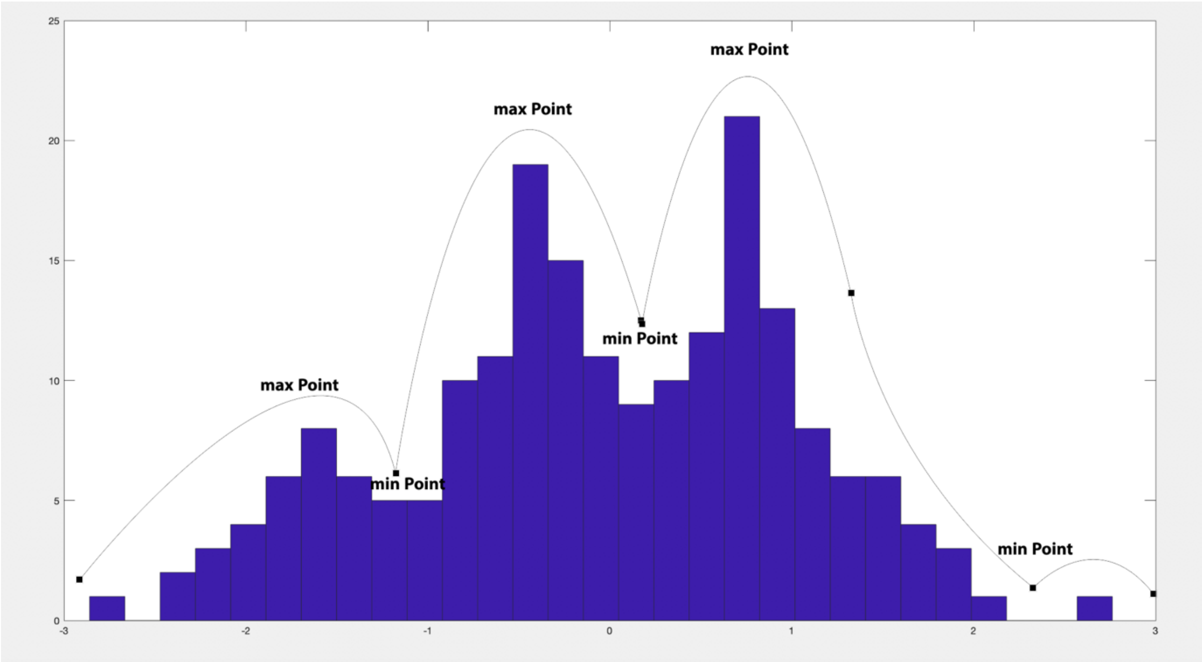

Figure 18 local min-max values

After finding min-max vectors, we use these vectors as the clusters by applying this entries this vectors into K-Mean Clustering (this technique is unsupervised technique for clustering images without the user providing sample classes or labels, and uses an equation to get results such that Euclidian Distance). After that we apply the k-means for whole original image by applying the K-Means in the original image which is two dimensional array. Then we take the number of local min-max values that will be the number of groups and the number of segments in the image which we want to apply this algorithm on, and the values for the vectors that is the outcomes of local min-max values as initial values for centroid in each cluster. We finally produce indices for clusters, which will enhance the final outcomes of the K-Means that will represent the segments in the image we are working on.

The K-Means will use an equation to compute and distribute the points for every specific cluster. Our methodology depends on the Euclidian distance equation to find the distance from each centroid to the points that belong to this cluster of an image. The equation is:

$$d(p_1, p_2) = \sqrt{(x_1 - x_2)^2 + (y_{1-}y_2)^2}$$

Equation 1 distance between two points

That means the system will take each pixel inside the original image and calculate the distance between it and every centroid to get the minimum distance and save the cluster for this minimum value. We then place this pixel for this cluster that has the minimum distance for its centroid point. Depending on that, we investigate to find that pixel belongs to any cluster using the equation 1 (Euclidian distance). After finding the minimum distance for each pixel, we should update the centroid for every cluster that has new pixel added to it, then to calculate and find the new centroid for every cluster by taking the average of all pixels inside each cluster and get new value for this centroid. After that we check if there are new values for any centroid. We reapply this algorithm from the beginning, but using these new centroids. Which means that we will find the distance from every pixel to every centroid to get the minimum. Then if there is any pixel with distance shorter to another centroid, this means that this pixel doesn't belong to this centroid. This pixel is transferred to the cluster that has the minimum distance to its centroid. After that, we check if there's any new value pixel to be transferred to another cluster, update the centers of clusters (as the first steps), until there's no any transfers between pixels from any cluster to another. Then the result will be the vector contains the indices to that clusters that will represent the segments in the image that we apply this algorithm on it. That means this step will produce new image with group of segments where each segment is represented inside the image with a limited number of colors where each color represents the segment from the original image.

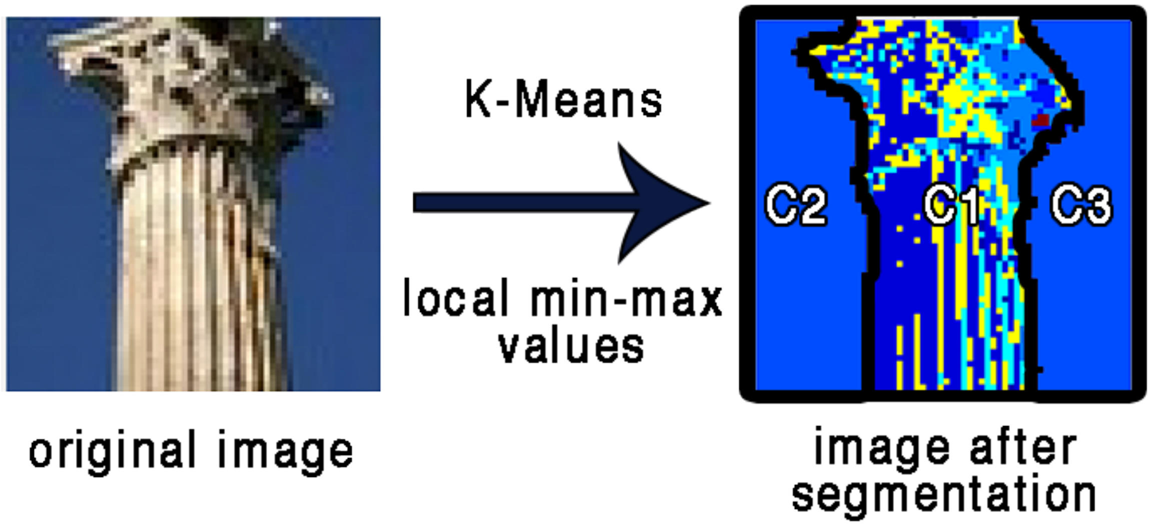

Figure 19 Result of apply K-Means using local min-max value

As shown in figure 20, the new image will have a limited number of colors (that each color represents a segment inside the image). These segments will be use in the recognition and in the classification steps. Then at the first, We trim the area adjacent to this image by trimming the area of the same color (so that the image contains a finite number of colors to represent each color segment) but apply the trimming process in the original image with LAB color space by comparing each area of the new image (Containing a finite number of colors) with the original image containing all colors, then we get the all pixels that has color with value C that found in the index (i's , j's) and go to the original image and get all pixel in the same indices but it will have group of colors that found in index (i's , j's) in the original image, this operation will apply for all three layers L,A and B all in the (i's ,j's) regions

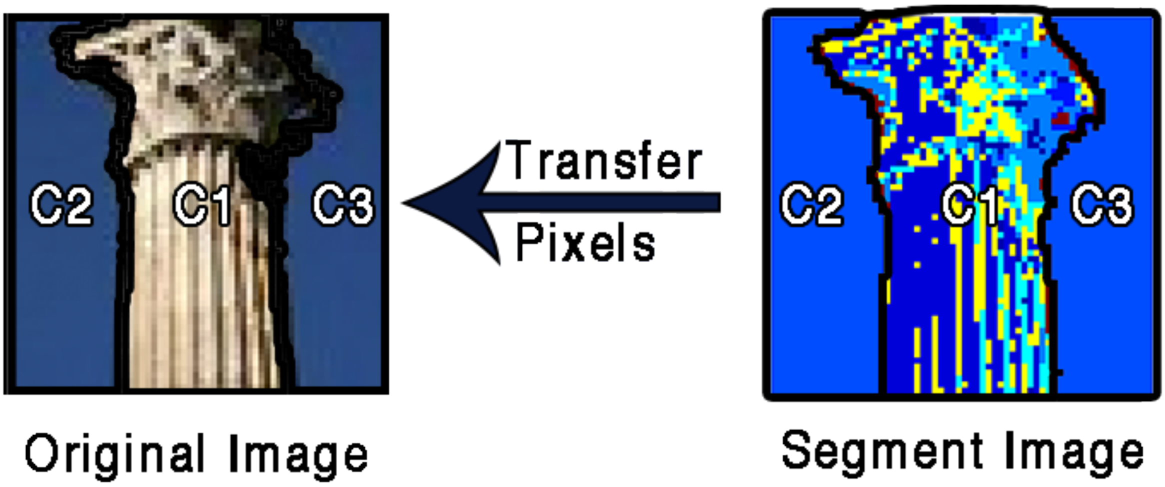

Figure 20 segment original image by transfer the indices

As shown in the figure 21 the segmented regions of the K-Means outputs are transferred to the original image in the same place for all layer, Then the areas that were trimmed in the previous stages will be used as images and the transfer of these new images to the recognition and classification of calls in the next stage.

In the second stage, we start the recognition steps, after segments the images we take each segment separately (these segments that should represent important meaningful else it will reject as we will see in future steps).

Recognition step:

As shown in the 22 figure, we present that steps of the recognition process.

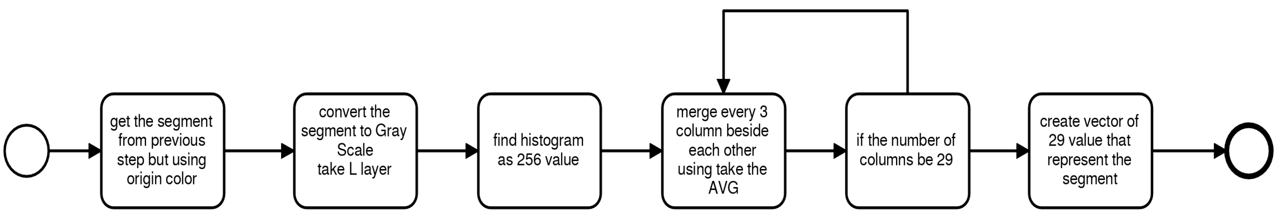

Figure 21 second step: recognition step

At this stage, we recognize the segments that we have trim it from the original in the previous stage. As we discussed in the previous sections all segments will be treated as new image. Then we first change the color space for the original color space of these segments into the gray scale that produces images that consist of just one layer where every pixel inside this image will have a value in range from 0 to 255 . After that, we apply the histogram technique on this segment that will create a vector of 256 cells (the size of the x-axis in the histogram), and each cell has the number of frequency value for this specific Gradient that mean all pixels have values between 0 and 255 where all pixels having value 0 will be stored in a cell with index 0 and so on. Then every index inside this vector has the number of repetitions for each pixel value as shown in the figure 23.

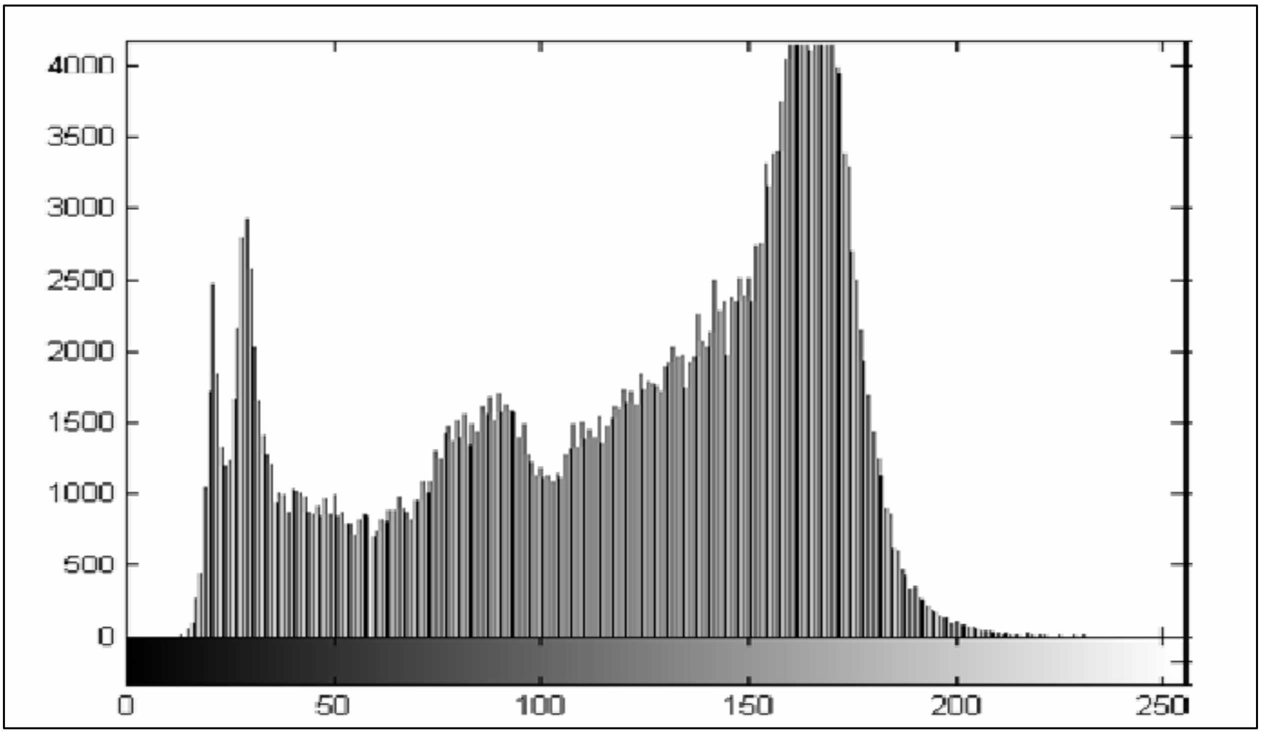

Figure 22 gray scale histogram

After creating this vector of 256 value we merge every 3 adjacent cells with each other by taking the average for these 3 cells and replacing them by one cell that is their average as shown in the figure 24.

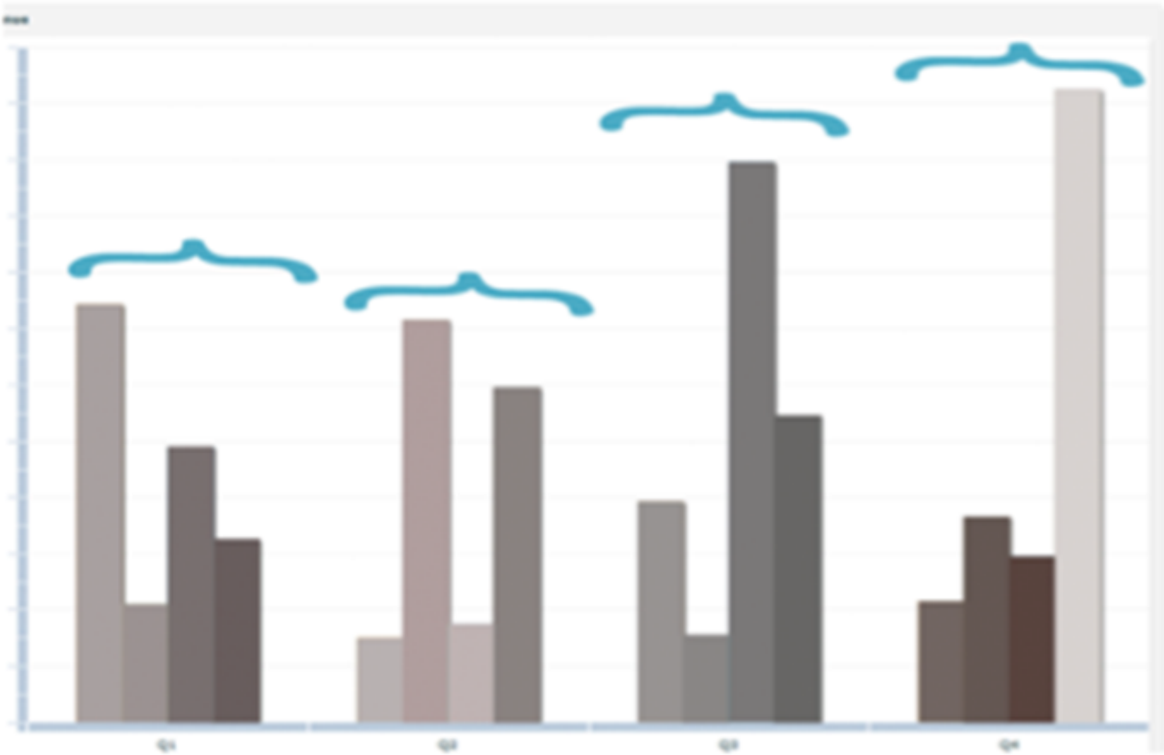

Figure 23 Merged Columns

Then applied this operation until the number of the columns that are produced from merging and take the average is 29 column that will produce a vector be 29 cells[9] by applying the merging two times. That means the result will be just 29 cells instead of 256 cells, these 29 values will represent the segments are trimmed from the original image, and this operation will be Appling for all images in the dataset that every image will represented by vector of 29 cells that mean instead of segment that has N*M of pixels that I want to treatment with it we just need a vector of limit size of 29 value that all operation of this segment will convert from complexity of O(N * M) to O(C) that mean convent from the system from quadratic system to constant system that will increase the speed and the performance and make it more powerful , then we will use this vector in classification step instead of original segment.

Figure 24 vector with 29 value that represent the segment

After shown in the figure above that it's the result vector of 29 value that produced in the merging columns from histogram.

K-NN step:

We will apply K-Nearest Neighbors Algorithm (K-NN algorithm) to classify these segmentations algorithm to determine for which class each segments are belong of heritage culture that stored in dataset , then we will take the vector that are produced in the previous step and get all vector that are stored inside the dataset that are hold labels and start see which of these vectors that are nearest for the new vector by using this formula:

$$d_1(I_1, I_2) = \sum_P \left| I_1^p - I_2^p \right|$$

Equation 2 K-NN equation

Where *P: is cell number or the index inside the vector, $* I_{1,2}$: represent the images or segments, *d: is the result that has the label of specific cultural heritage that the new image is belong to.

We subtract the vector from the 29 values that represent the segment produced in the last step with some vectors from a set of vectors in the complete dataset. For each subtraction there are 29 values to be produced and then we will combine these values within the result vector. After that, the last step produce group of values that each value represent sun of all 29 value that produced from subtract operation then we will find the minimum value of these values and look at the image that will get the min value for these operation then this image will hold the goal cultural heritage for this segment but not the whole original image because the whole original image has many segments.

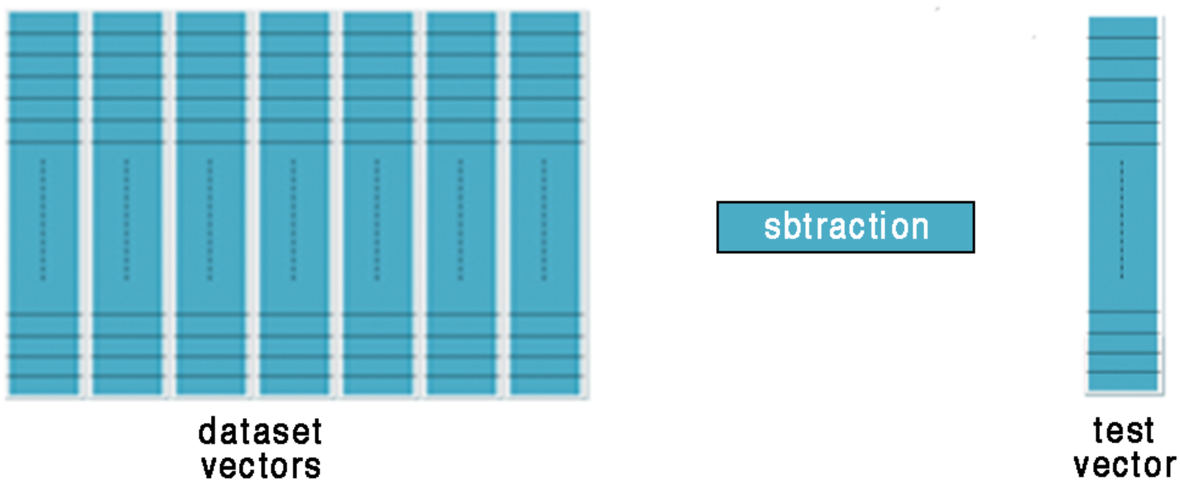

Figure 25 subtract the result vector with all vectors inside the dataset

As shown in the figure 26, instead of applying the K-NN algorithm and subtract whole image of size x*y*3 (three layers) we just need to subtract 29 values, then the time complexity will be powerful for number of pixel that we will subtraction that the time complexity is O(C) instead O(N*M).

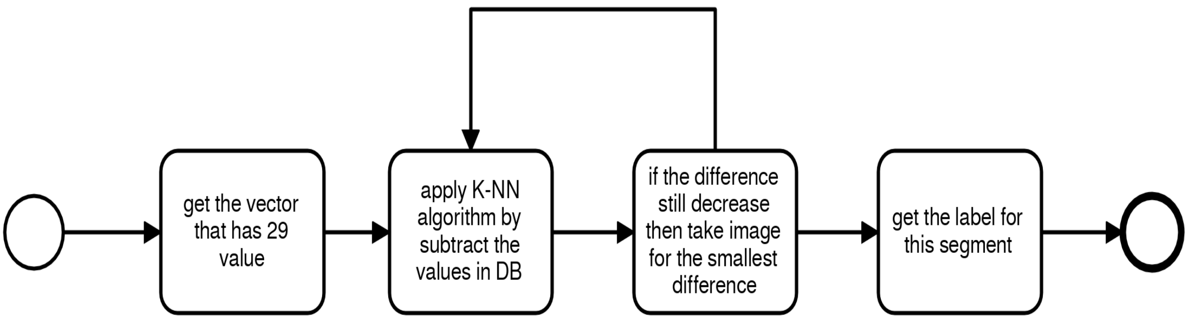

Figure 26 K-NN steps

Classification Step and get the result:

Now, we will start at the third and the last step in our technique that is the classification step that we will see cultural heritage to which the test image will belong, as described in the methodology in the figure 27.

Every culture has common features then the classification way is to take every segment for test image and find where is the minimum value found in each segment from the vector and get the image and get the label that holed in that image then we can say that this segment belongs to this group that represent the cultural heritage, after that we get each segments labels then we will create vector of size the number of all cultural heritage appears from the image segments, after that calculate the count of all redundant groups and store it in the vector then the vector will has a value from 0 to number of all groups in the dataset.

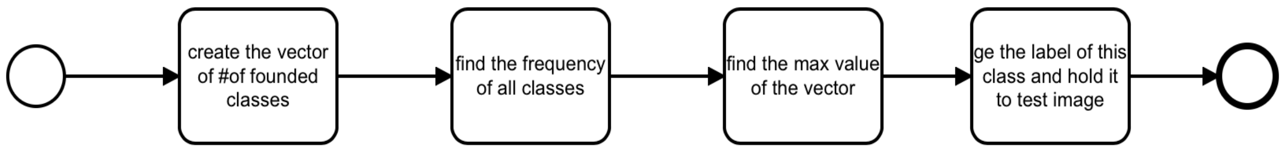

Figure 27 Final Step in Classification

As shown in the above figure that shows the final step of classification stage, Then after we fill the vector we should to determine the whole test image for which cultural heritage group that is belong to, then our methodology is to take the cell that has maximum number of segments that mean this cultural heritage is the goal cultural heritage group, this is the our way to find the goal group of all segments vectors in test image with all segment in all dataset, However the time complexity will be O(K*N) that K will be the number of segments in each image in dataset and N is the number of images in dataset , it will be huge for huge datasets such as if we have 100000 images in the dataset and for each image has 50 segments then the number of iterative will be 100000 * 50 that will produce 5000000 and its very huge in real time system, then to solve this problem we will use another methodology to store data that is a Binary Search Tree (BST) for represent the segments for each image in dataset, that we will create Binary Search Tree for every group of cultural heritage that consists of segments in shape as vector from 29 cells that we are create it in the previous steps.

That mean each image inside specific cultural heritage all segments will store inside Binary Search Tree that group of Binary Search Trees will represent this cultural heritage, then the aim of use Binary Search Tree is to apply its concept that the value of the root is larger than the left child and less than the right child that give flexible in our search instead our search take O(K) by using linear search or using search item by item the complexity will be O(Log K) then it will take log K of titration instead of K iterations , that mean we don't need to test whole tree just log k of elements from this tree (that the BST is just test the left or right branch), then the test will be by make the different between vector the get the summation and test it with the tree ,that will take decide to go to left or right or stop if the value be lager that the previous level, that every time we go deep of left levels the value should be lower than previous left level and vice versa for right side , then in every test we will reject half of vectors from vectors inside the tree that represent specific image then its value of search will be smaller than the level that stand in it before like shown in the figure 29.

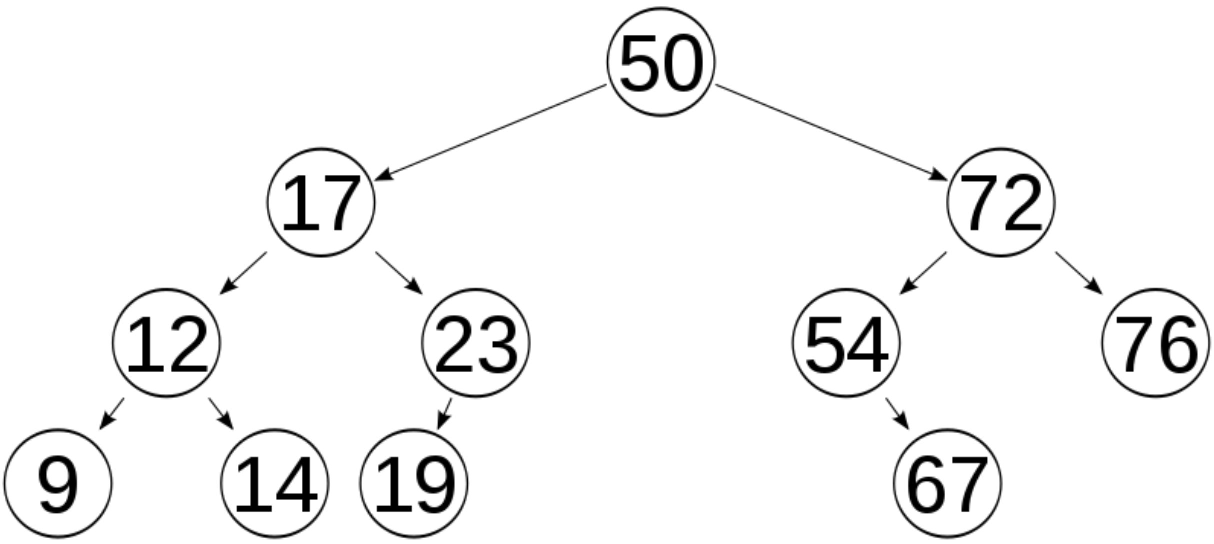

Figure 28 BST with vectors value

Algorithm : Get The Closest Class for Test Segment

Result: Closest Class
TestSegment //segment that need to test as a vector of 29 value;
DataSet //Array Of BST for each Image in Dataset;
for $I \leftarrow 1$ *to* N **do**
  $BST \leftarrow Dataset[I]$ //Binary Search Tree has Segments for Image I Of Dataset;
  $diff \leftarrow MaxInt$;
  while $diff$ *Decreases AND* $BST \neq null$ **do**
    $DLeft \leftarrow BST.left - TestSegment$;
    $DRight \leftarrow BST.right - TestSegment$;
    if $DLeft \leq diff$ **then**
      $diff \leftarrow DLeft$;
    end
    if $DRight \leq diff$ **then**
      $diff \leftarrow DRight$;
    end
    if $DLeft \leq DRight$ **then**
      $BST \leftarrow BST.left$;
    else
      $BST \leftarrow BST.right$;
    end
  end
end

Figure 29 Pseudocode for K-NN and BST

Then as shown in the figure 29, this is the pseudocode to get the specific class using K-NN and BST ,this method will result in time complexity for a segments of search within the group of segments will take (log K) then the time complexity of whole search will reduce for good time that the time will be O (N * log K) for classification instead of O (N * K).

Chapter Four: Results

Segmentation Result:

After approximately 150 experiments on many cultural heritage (roman , pharaonic , Islamic ,….) with dataset of 20 culture that each culture have approximately over than 100 image that this dataset get from Pinterest that the number of images inside the dataset is 3200 image, we noticed that the segmentation step has several attribute to consider such as the depends on colors difference inside the image that when there is an edge between colors then we can cut the segment , another thing that the segmentation depends on image resolution , that means if the image resolution increased then the segmentation will increase too, and vice versa. There is another attribute that determines the resolution of the color space segmentation of this image in our methodology. We used the channels created by the color space in the laboratory, making the selection of the section more powerful and solving the problem created by the lightness, which has a specific channel of light that can help to solve the parts. Division of certain areas, as shown in our methodology (that discussed in the previous chapter) we segment the images by taking the close regions of that segment like the figure 30.

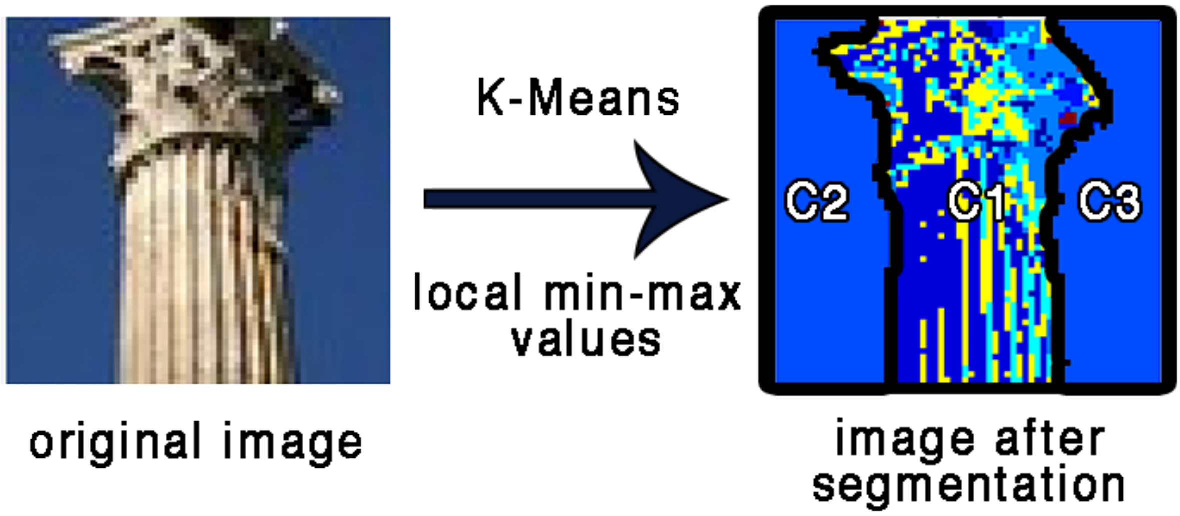

Figure 30: the result form segmentation using our method

Recognition result:

After we finish the segmentation step , we apply recognition and classification steps that we use histogram and K-NN algorithm to get the result ,that we convert each image to vector from 256 value come from histogram for gray scale level and convert it to vector from 29 value then this step will reduce the complexity that make the complexity of subtraction be fixed that mean just 29 value need to subtract not the whole image (N rows * M columns) as shown in figure 31.

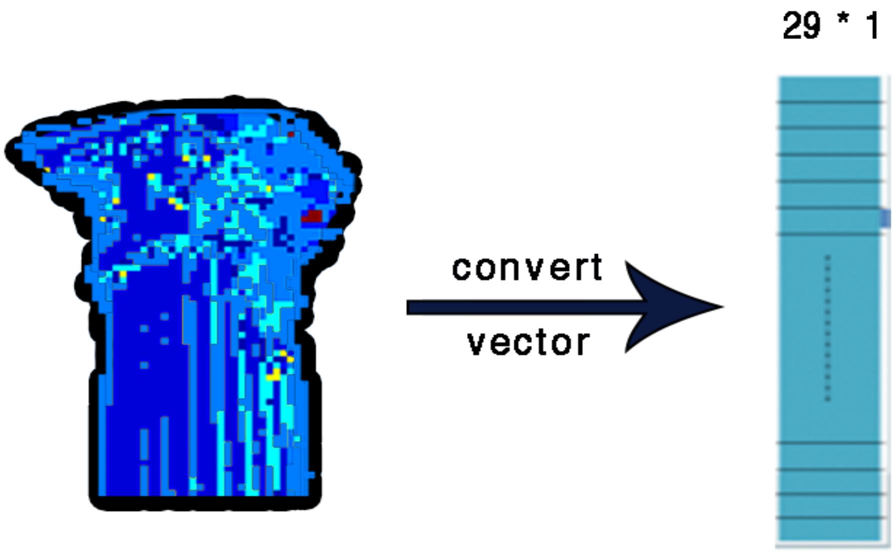

Figure 31 convert segments to vectors

Classification result:

After that we will go to the next step which is the classification step that determines for which cultural heritage an image belongs, But as we see at the first the time the complexity be O(N*K) N number of images in the dataset and K the number of segments in images because we will test each segments inside the test image will all segments in each images in the dataset by subtract each segment with all segments in the dataset and it will take too much time to get the result of our classification to get the specific cultural heritage, however we enhance this method by put each cultural heritage segments inside Binary Search Tree that make the search more powerful because in every iteration we will cut group of segment inside the same path then the time complexity for searching will be O (N log K) then we don't need to check all segment in the dataset just the closest value to segment from test image as shown in the figure 32.

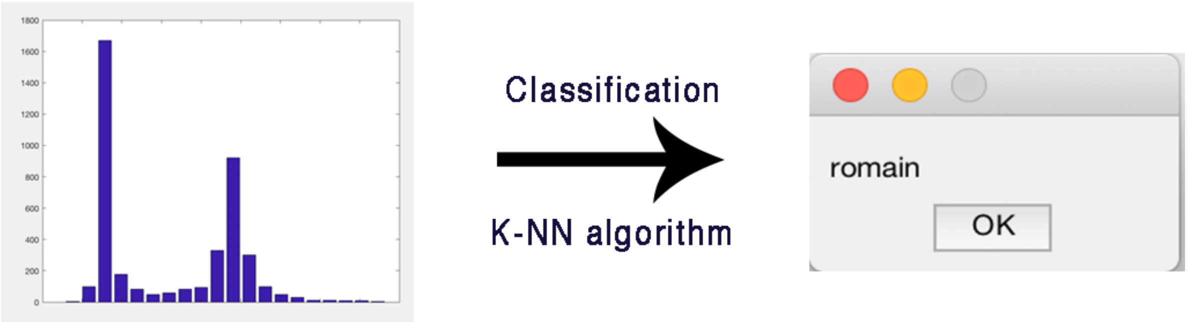

Figure 32: the result from recognition using histogram and convert it to vector from 29 values and classification using K-NN algorithm that applied in the BST.

Some of pass results:

We apply our algorithm on dataset of groups of cultural heritage that each group of each cultural heritage has range between [70 - 150] images that describe its cultural heritage and we use more than 100 experiments as test for different cultural heritage to test our methodology.

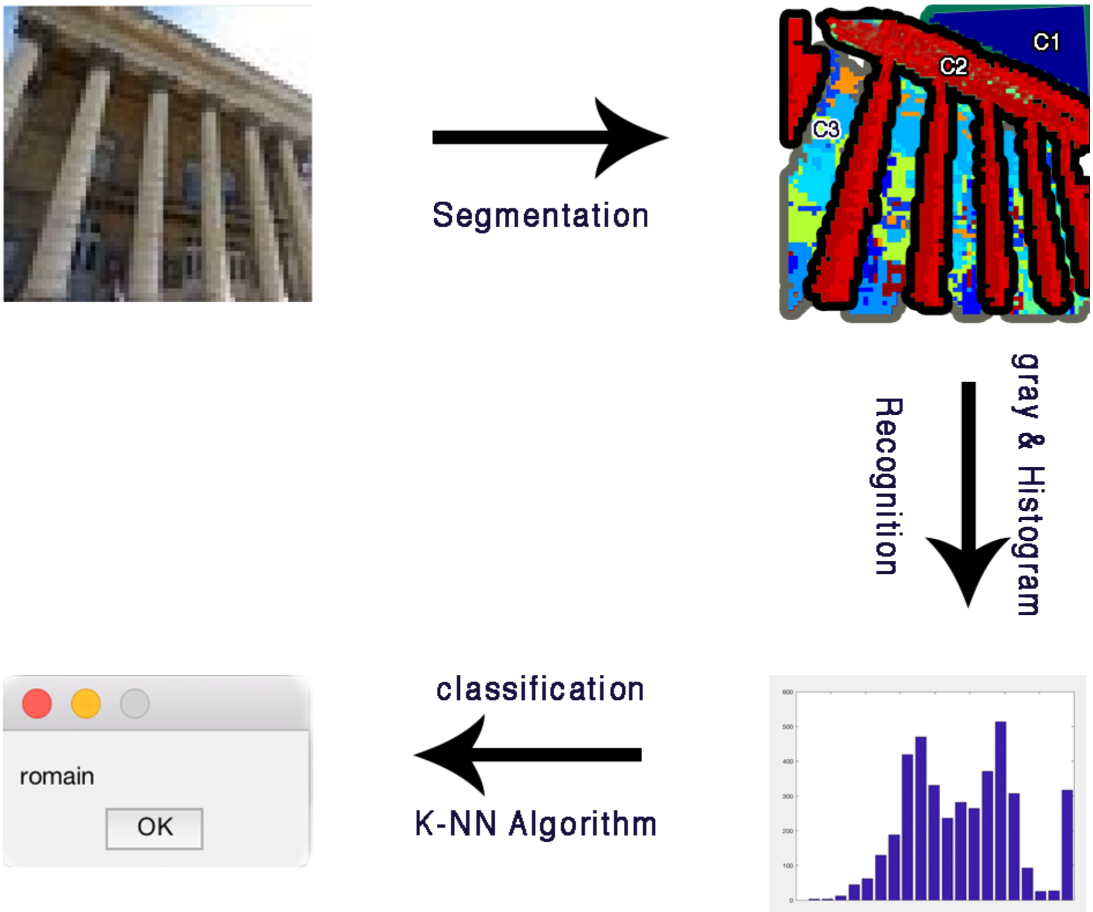

Figure 33 the result from test image from roman culture

As shown in the figure 33 the number of segments is 3 the columns segment and another two segments that the first segment will recognize the column as roman culture the other segments will get unknown label then the result dependent on labeled segment is cultural heritage be the roman.

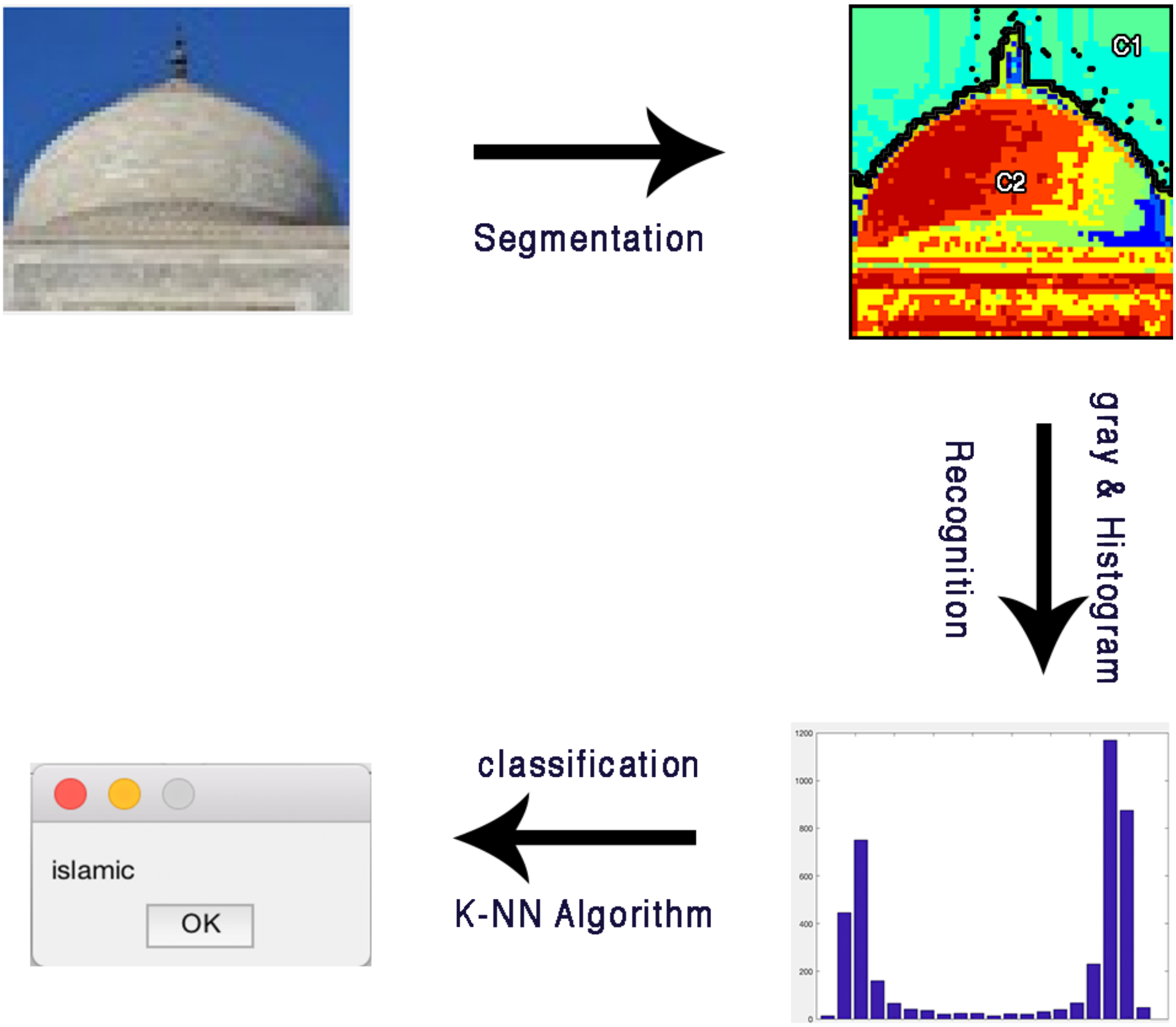

Figure 34 Result for Islamic culture

In the figure 34, the image is split into two segments: the dome and another segment (unknown label) then the dome segment is recognized as Islamic heritage but the second segment is classified unknown then the main cultural heritage will be the Islamic heritage.

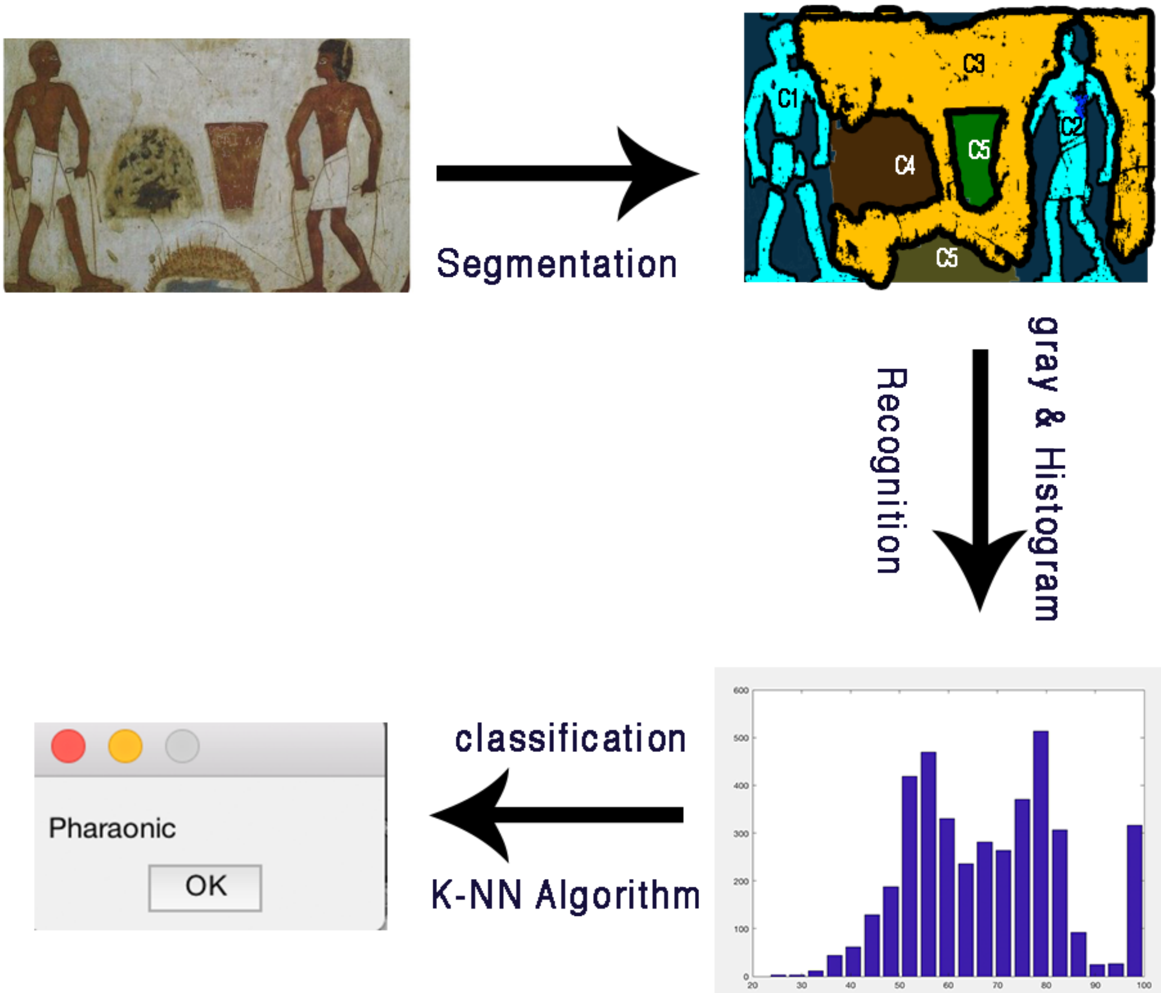

Figure 35 Pharaonic heritage

In figure 35 the original image splits into five segments then the first and second segments will be in pharaonic heritage and the segment five will be in another segment and the last two segments will be unknown then the propriety will be to the higher number then the main heritage will be the pharaonic.

After many experiments in many cultural heritages the ratio of success and knowing the cultural heritage is more that 69% from image inside the test and dataset that will discuss the numbers in next steps.

Failed Result:

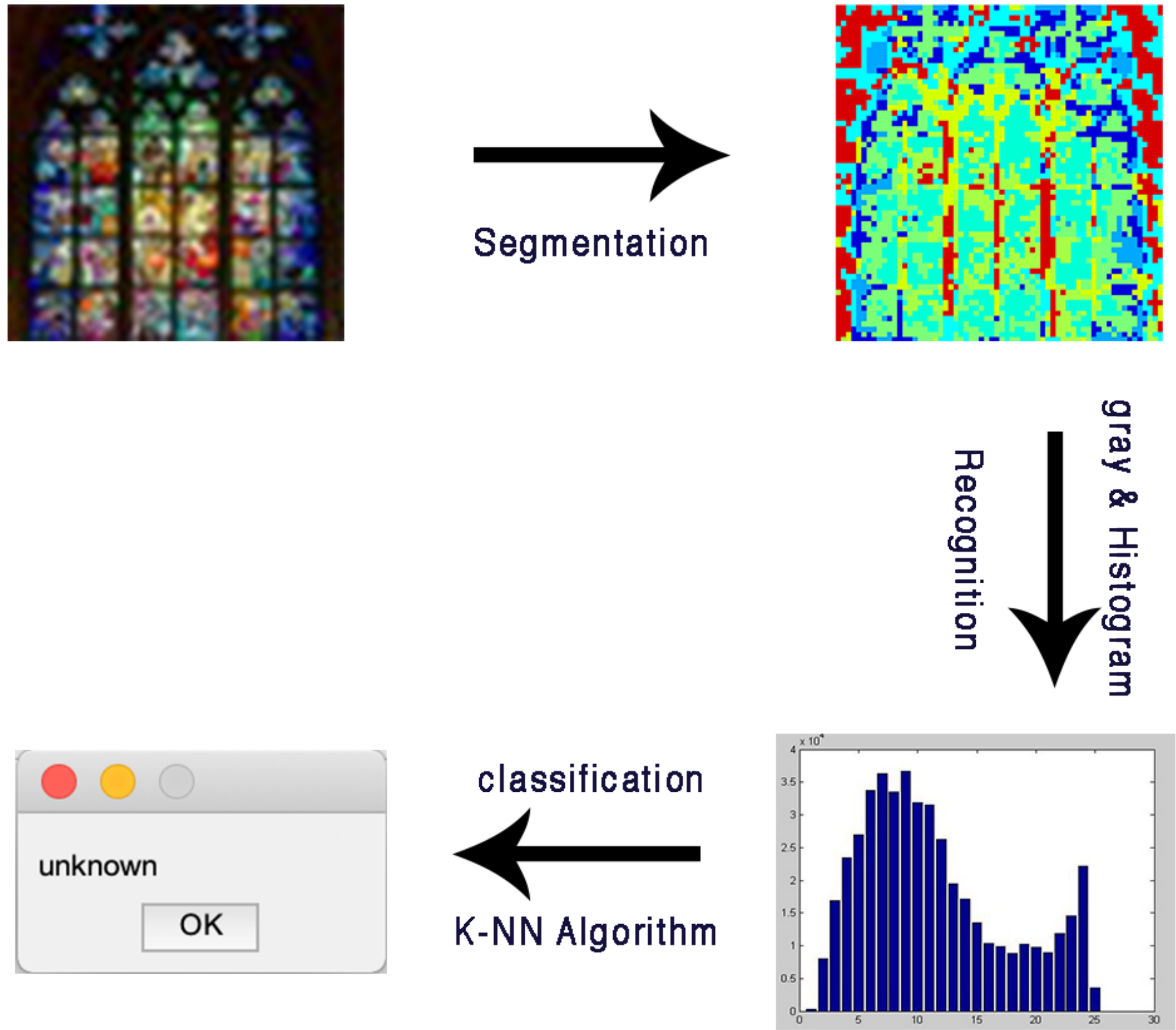

Figure 36 Fail images

In the figure 36, we can see when our methodology fails mostly in the images that have many groups of colors then our system will segment every group of colors alone however may some segments have collection of colors to represent one object.

Discussion results with charts and tables:

We have many algorithms for image segmentation and there also many algorithms for image classification, we will discuss some of the result of these algorithms some of segmentation and some on classification, we will start in segmentation we will discuss three main algorithms on segmentation process the gray thresholding , color thresholding and max value thresholding, at the first we have use gray thresholding algorithm and get the result as the chart 37:

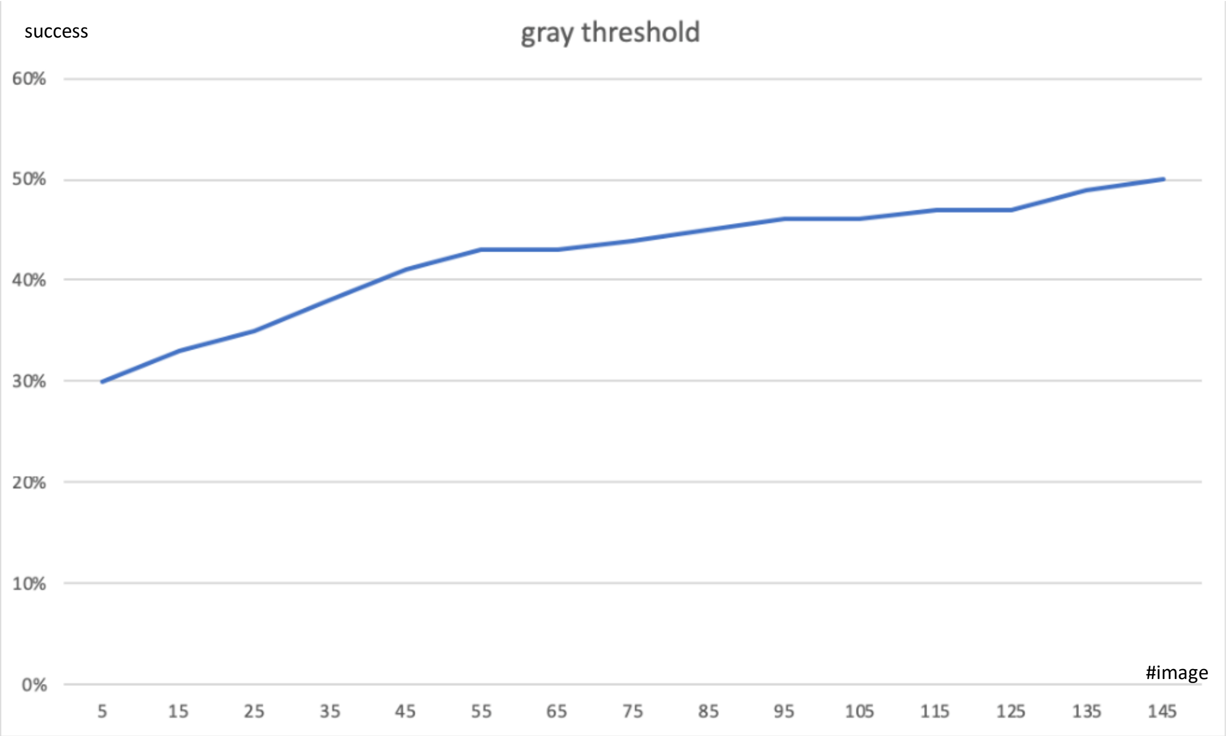

Figure 37 accuracy value when use gray thresholding versa number of images

As shown in the chart 37 for the result of gray thresholding that we apply it in range of images from 5 image to 145 images by increase 10 images in every step as we noticed that the result start from 30 % and the maximum accuracy percent is the 50% that mean half of test images will fail when we using this technique, because when we convert the image into gray scale there are many details that will be loss because we reduce the number of layer from 3 layers (colors layer) to one layer (gray layer) , and as we see the upward shape of the curve seemed to rise very little with the high number of images in each group.

# of images	gray threshold correct percent
5	30%
15	33%
25	35%
35	38%
45	41%
55	43%
65	43%
75	44%
85	45%
95	46%
105	46%
115	47%
125	47%
135	49%
145	50%

Table 2 accuracy value when use gray threshold versa number of images

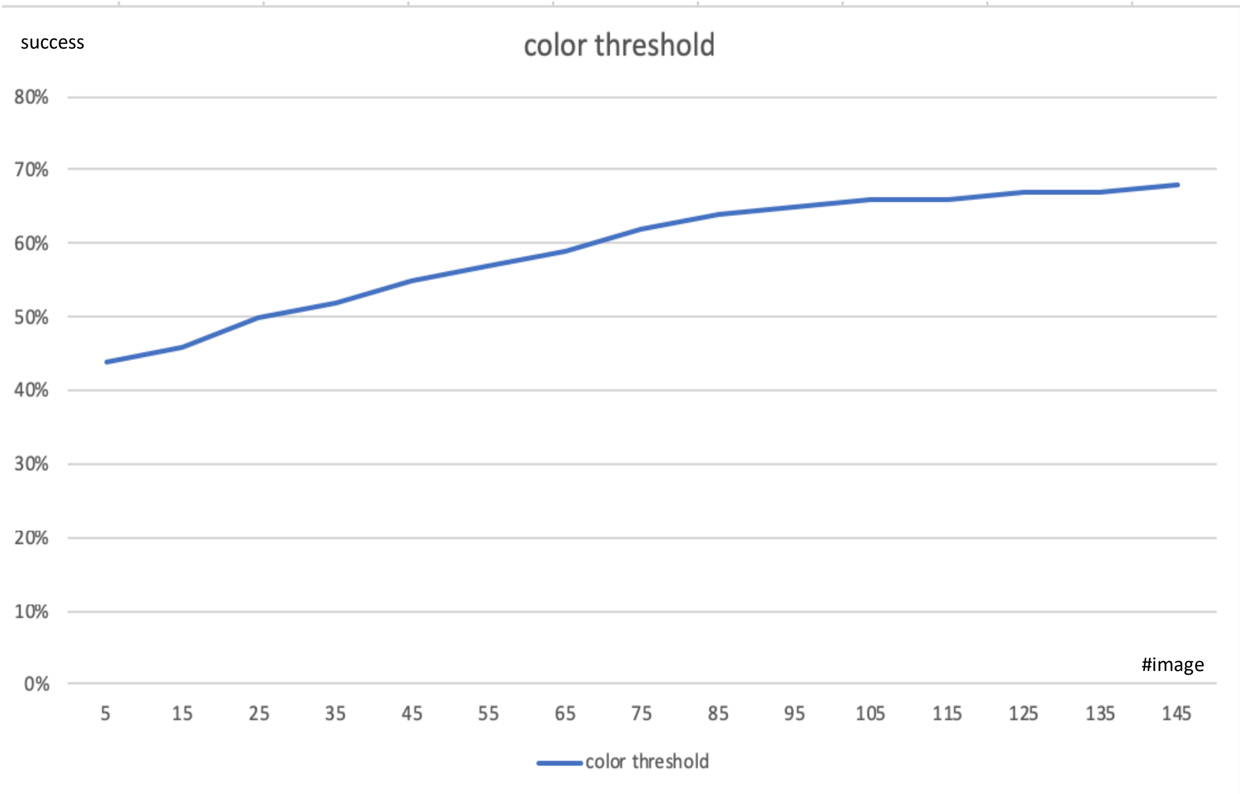

Figure 38 the accuracy of using Texture Feature algorithm

# of images	Texture Feature correct percent
5	44%
15	46%
25	50%
35	52%
45	55%
55	57%
65	59%
75	62%
85	64%
95	65%
105	66%
115	66%

125	67%
135	67%
145	68%

Table 3 accuracy value when using Texture Feature versa number of images

As shown in the chart 38 for the result of Texture Feature that we apply it in range of images from 5 image to 145 images by increase 10 images in every step as we noticed that the result start from 44 % and the maximum accuracy percent is the 68% this means that two third of images will success (we will see that is the most technique that the result be closer to our algorithm), when we using this technique, we will split the image into segment by taking the color threshold and segment the image by number of threshold in each image in this technique we used the three layers, and as we see the upward shape of the curve seemed to rise very more with the high number of images in each group.

Figure 39 the accuracy of using hill climbing algorithm depends in # of image in the dataset

# of images	hill climbing correct percent
5	40%
15	42%
25	44%
35	47%
45	49%
55	51%
65	52%
75	55%
85	57%
95	58%
105	58%
115	59%
125	60%
135	60%
145	61%

Table 4 accuracy value when use hill climbing versa number of images

As shown in the chart 39 for the result of max value thresholding that we apply it in range of images from 5 image to 145 images by increase 10 images in every step as we noticed that the result start from 40 % and the maximum accuracy percent is the 61% that mean sixty of images will success ,when we using this technique, we will split the image into segment by taking the max value that are come from the histogram and segment the image by number of max values in each image, in this technique we used the three layers, and as we see the upward shape of the curve seemed to rise very little with the high number of images in each group.

As shown in all previous examples that the increasing in number of images insides each group that represent specific cultural heritage lead to increase the ratio of correct recognition and classification and get the better result.

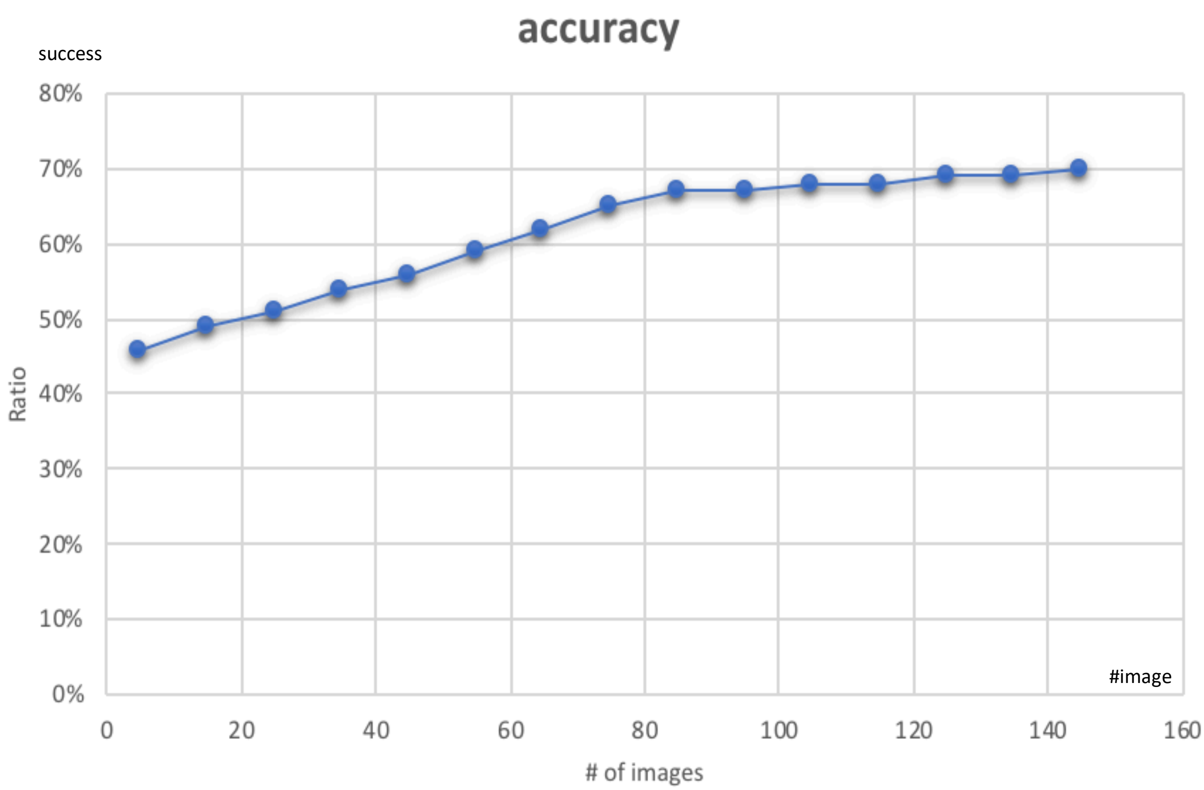

Figure 40 the accuracy of using our algorithm depends in

# of images	min - max with LAB
5	46%
15	49%
25	51%
35	54%
45	56%
55	59%
65	62%
75	65%
85	67%
95	67%
105	68%
115	68%
125	69%
135	69%

145	70%

Table 5 accuracy value when use min-max lab color space versa number of images

As shown in the chart 40 for the result of Min – Max With K-Mean algorithm with LAB color space that we apply it in range of images from 5 image to 145 images by increasing 10 images in every step we noticed that the result start from 46 % and the maximum accuracy percent is the 70% that mean more than two third of images will success (we will see that is the our technique results be more than color thresholding algorithm), when we use our technique, we split the image into segment by taking the number cluster that come from min-max values and send it to K-means and segment the image according to the cluster that is produced from K-Means in each image and in this technique we used the three layers, and as we see the upward shape of the curve seemed to rise very more with the high number of images in each group.
As seen the accuracy increased from 46% to 70% to get correct cultural heritage and as we see that when take the average number of image like 70 images for each group that the difference begins to enter the stage of stability that the different be small.

In the figure 41 is the compare between the time complexity on our method of recognition and classification based when we used BST (that will not check all segments in the dataset just log n of them) the time complexity will be O (N * Log K) and when apply the linear search (that test all images and segments in the datasets) that the method will take O (N * K).

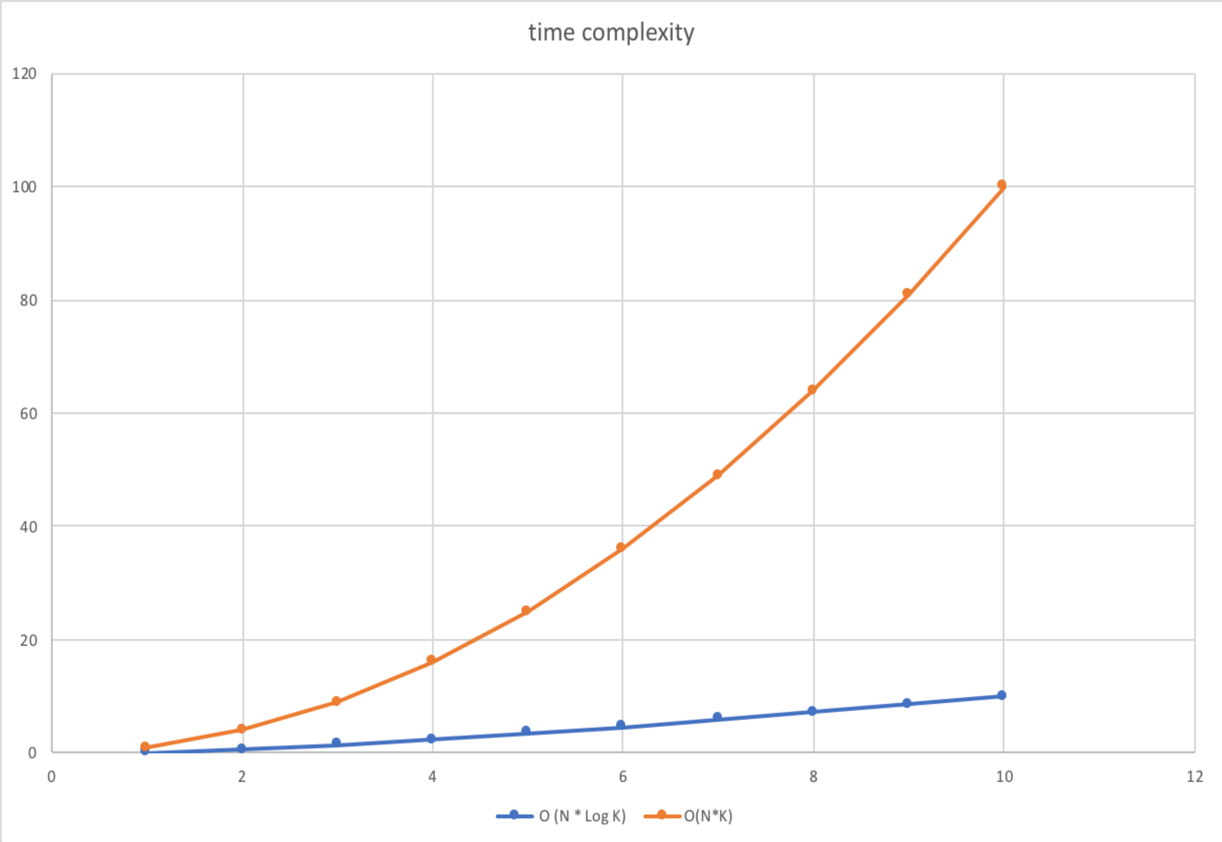

Figure 41 time complexity for our method vs. the default technique in search method

As we see in the chart 41 that there's the gap between two lines increased in huge shape in little difference in x-axis that mean the time will be very faster in large number of segments inside the images this is from the strong points in my methodology.

In the figure 42, we will see the comparison for time complexity between our methodology for subtraction in K-NN algorithm and the default subtraction that the result be constant

O(C).

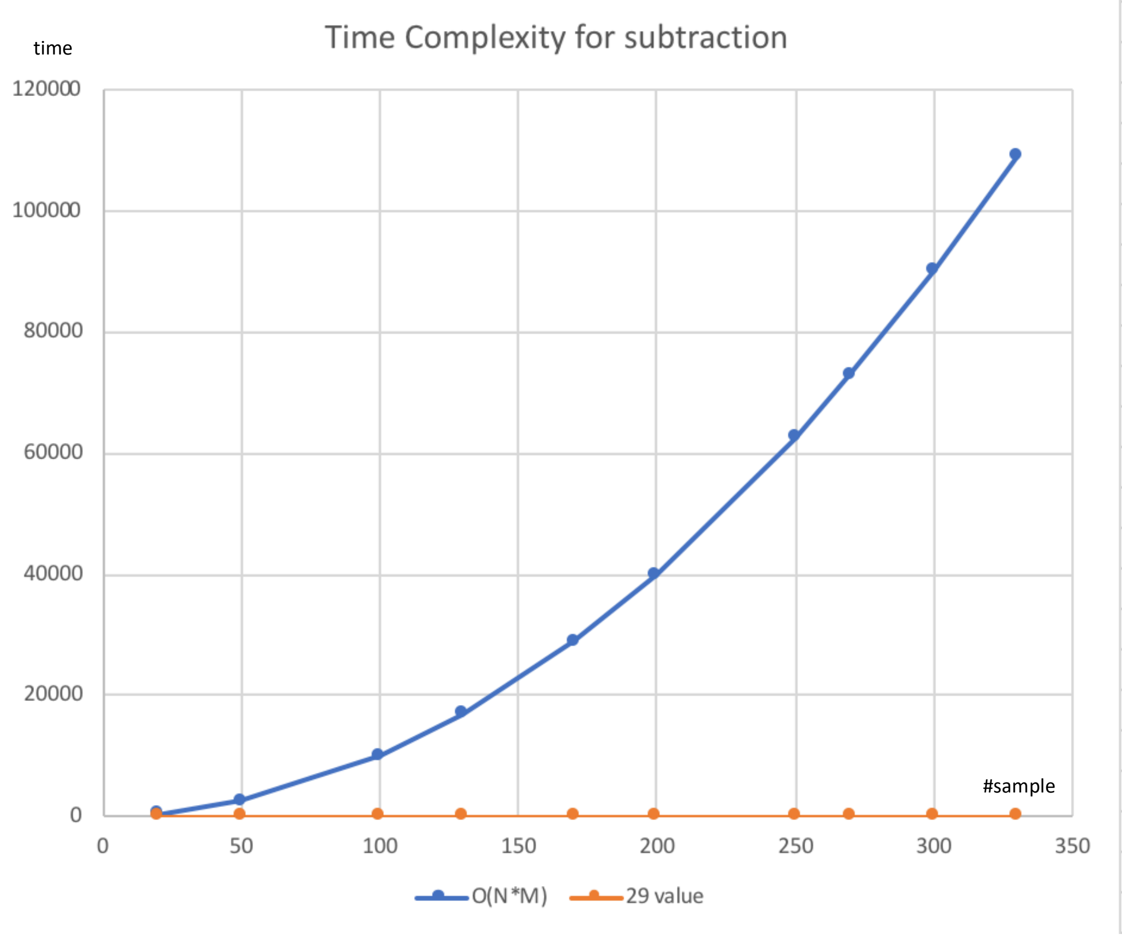

Figure 42 time complexity for our subtraction method vs. default subtraction in K-NN

As known that the K-NN algorithm will subtract the whole test image with images from dataset that will subtract image that has N rows and M columns. However, in our methodology will need just to subtract vector from 29 values that mean the complexity in subtraction will be static for all images inside the dataset.

As seen in the chart 42 That there is a large gap between the two lines in a huge form in a small difference in the small x-axis that means that the time will be much faster in a large number of images because this subtraction has a fixed value only 29 value (that produced from histogram and taking the average for adjacent columns) instead of N rows * M columns this is from the strong points in my methodology.

In the figure 43, we will see the comparison between our technique with other techniques that the ratio of our method be the best of them mostly when the number of image is huge.

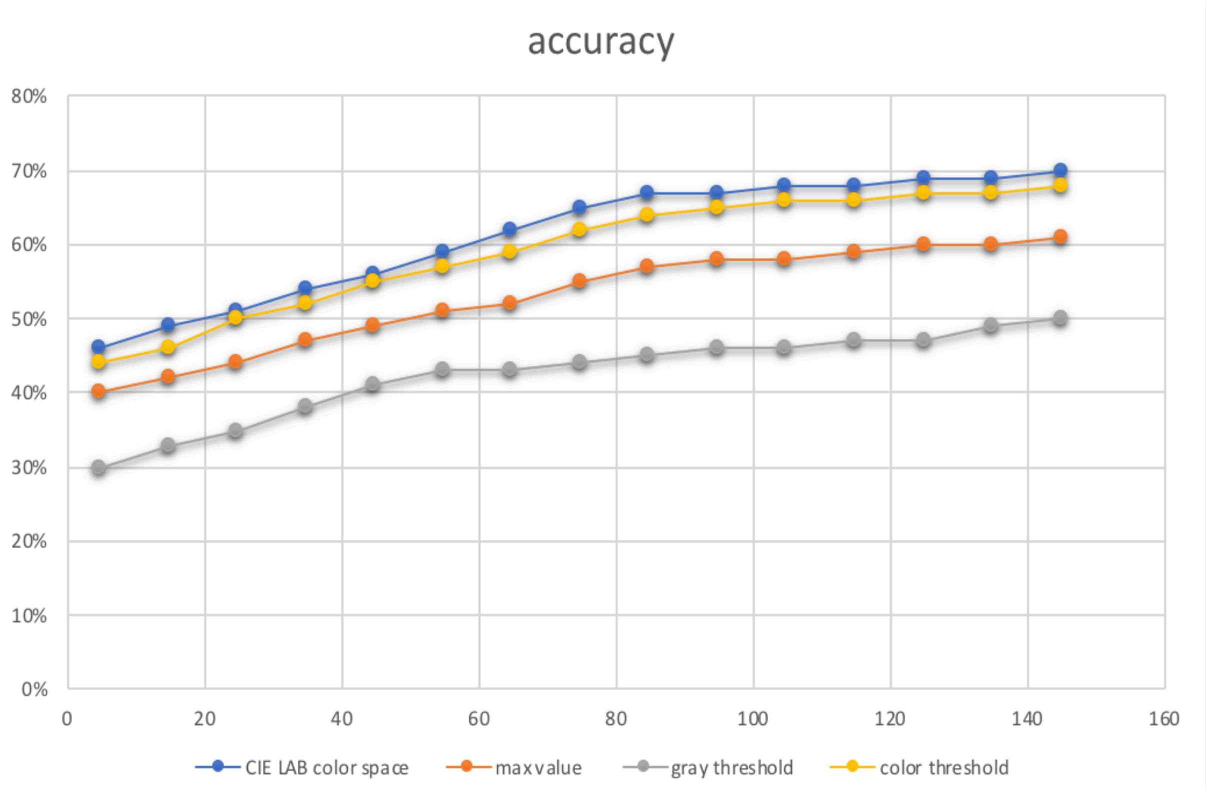

Figure 43 the comparison between our algorithm vs. other algorithms

In the figure 43 we attention that we used at the first the LAB color space that has big gamut and in the segmentation algorithm we used the K-Mean that take the local min-max values that found from histogram for each layer for LAB layers and consider it as the first indices for clusters get the best result because it take many attribute with it segment the images consists that K-Mean is part from unsupervised that it doesn't need that the user providing the sample classes just take the number of classes and start to collect the pixels inside the image to each class and make many iteration to enhancement the indices of the class that mean the pixel may transfer from class to another if it be closer to its indices and by applying it in this image that determined by the number of local min-max values and the K-Mean will determine these pixels for any class in it belong that get the segments by using squared Euclidean distance.

In the figure below, we describe the different between use CIE-LAB color space and another color space such as RGB color space and HSV color space in our application (cultural heritage).

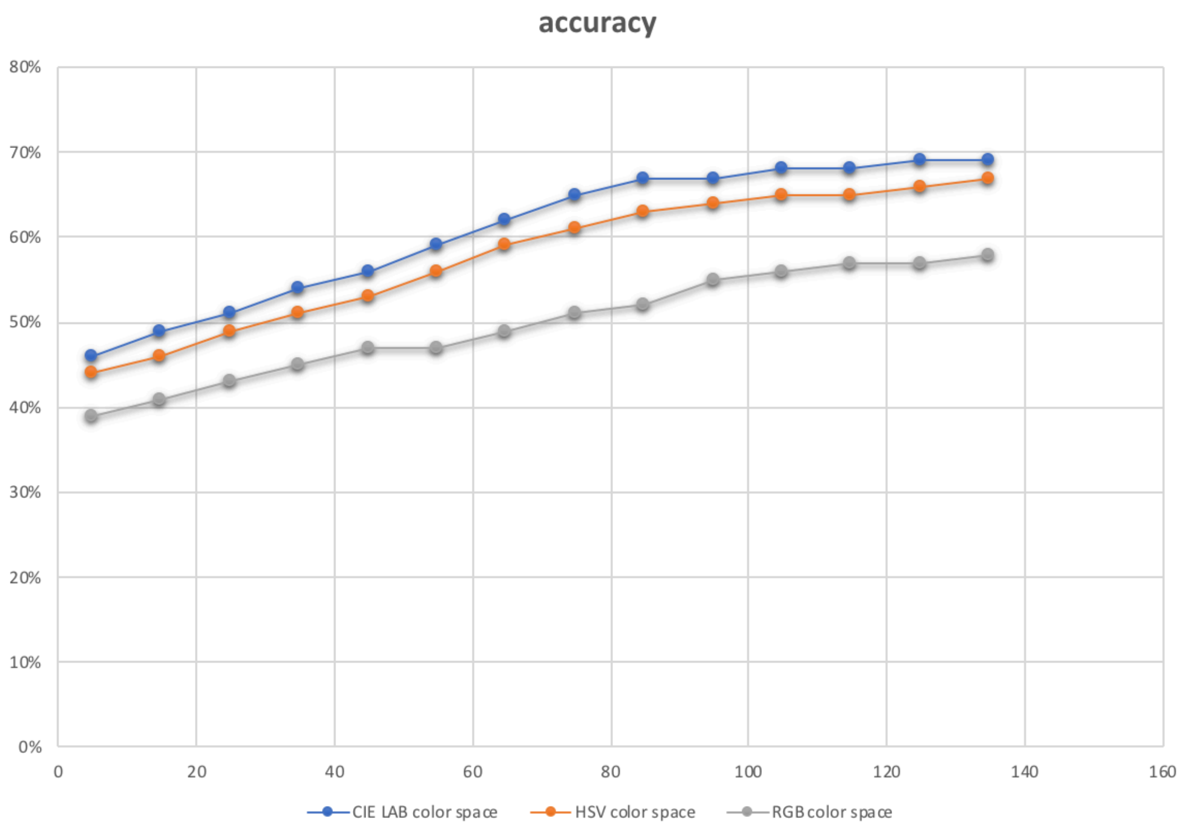

Figure 44 the different of accuracy between the LAB color space, the RGB color space and HSV color space

As we see that the accuracy of using CIE LAB color space get more accuracy than another color spaces the RGB color space and HSV color space that prove our method of using the LAB color space that use three channel two of them is the Blue across Yellow (Opponents) and Green across Red (Opponents) that represent the color in approximately shape how humans see, and the last channel is the lightness then we can take a pure color without it lightness or maintains the light inside the image to get the best result and the CIE LAB color space gamut is larger than the RGB color space then the number of color we can treatment with it is more that will higher accuracy that the RGB color space , and as we see that the worst color space in the RGB that it doesn't segments the image in perfect way because each layer get me the gradient of Red or Green or blue but do not get all other colors when you start or finish, however, the HSV color space makes me approach our approach but still less than the LAB.

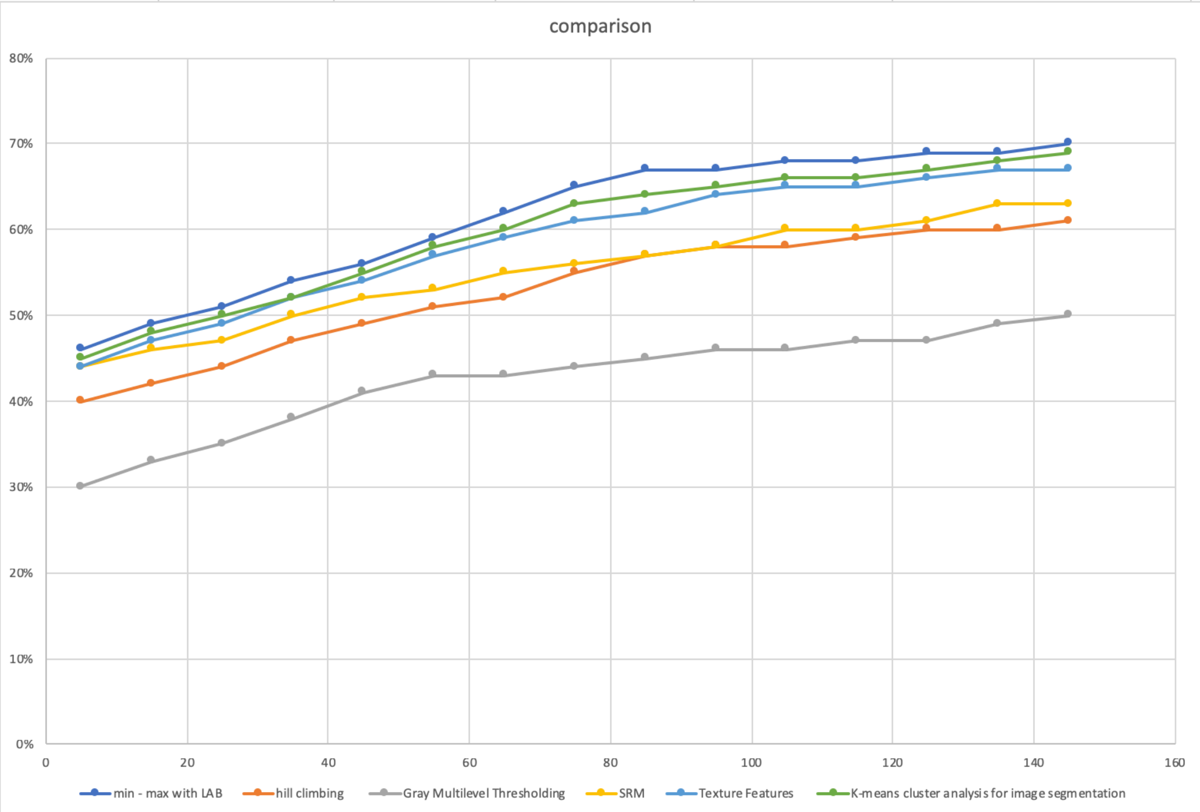

Figure 45 comparison with some algorithm

Finally in the figure 45 , we make some comparison with algorithm from previous work that as we noticed that the most closest algorithms is the texture and K-Means analysis that because in depends in number of clusters that give finite number of segments but they used L layer in their solution but in out solution we remove it to treatment with pure color in the images and the texture technique are mixed color into a texture analysis and recognition plan can be very important that based on the texture feature of objects in the images, and hill climbing algorithm get less result that just depends on max value the number of cluster will be small and it remove too much features in the images then the segmentation will be less than our algorithm that the number of segments about double and in gray thresholding it use one layer different from our algorithm that used two layer of perceptual color space that give true colors , at the last as shown in figure 45 our technique give the best result in cultural heritage features.

Chapter Five: Conclusion and Future Work

Conclusion

We make a new technique for image classification our algorithms will do each stage of classification starting from the segmentation and ending with their classification the test image to correct class as shown in the block diagram in figure 46.

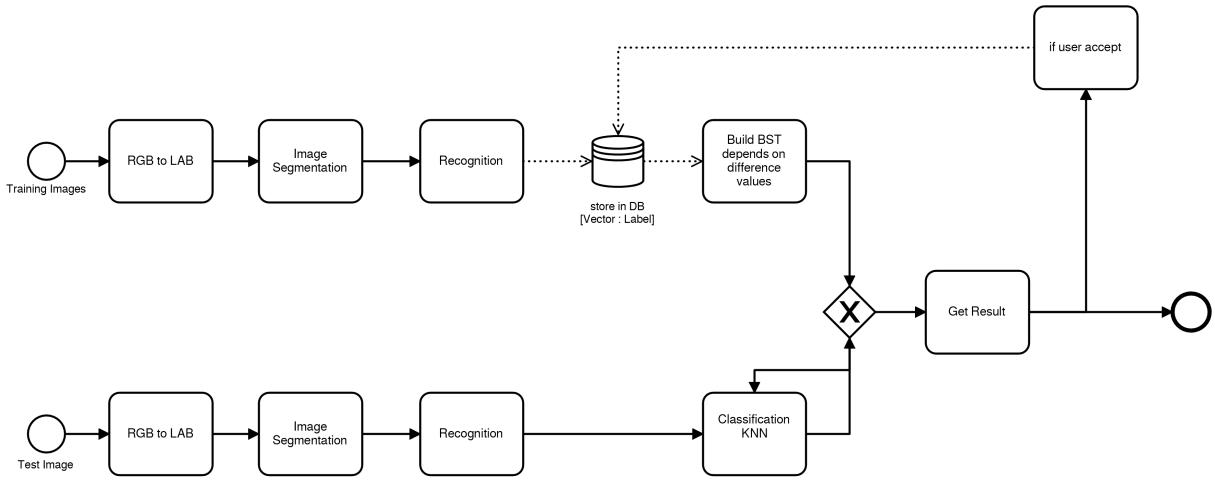

Figure 46 methodology

Our algorithm start with take the input image and convert its color space to the CIE LAB color space then the three channel will created (that has a huge gamut and it's channel Blue towards Yellow (Opponents) and Green towards Red which it almost closer for how human perceive things), and as we see LAB color space get best result in our application, after that we will find the histogram for these channel and we will the find local min-max values from these histograms then we will take the number of these values and towards it to the K-Means Algorithm with the image then the K-Means will segment the image by these local min-max values by make making many iteration until knowing for which each class that are belong to any class ,after that we will extract the features from the image by trim the close regions that created by K-Means in the test image and take each segment to the next step, then the segmentation step is over.

The next step will be the recognition that start by convert these segment to gray scale after that find the histogram for these segments then a vector of 256 cells will created then we will merger every three-column adjacent to each other by taking the average and iterative this until we get vector from 29 cells then this will reduce the timer complexity for the subtraction in the next step that we don't need to subtract the whole image of size X * Y * 3 that just we need to subtract 29 value ,it will make the time complexity more powerful that the default operation that the that will be constant and small.

After that we will enter the last step it's the classification and we use the K-NN algorithm that we all subtract all vectors form dataset with the test vector and get the smallest difference ,as we see if we apply it directly the time complexity will be O(N * K) so, we solve it by using Binary search tree that make the complexity O (N* Log K).

After many experimental we can conclude that there's many parameter we should be considered , one of them is the color space we there's many color spaces and most of the images used RGB color space but as we see the RGB color space get worst result because of the histogram for each layer represent the gradient for red or green or blue but not then combination of them however we use the LAB color space that the two main channel represent combination of colors and it's channel Blue towards Yellow (Opponents) and Green towards Red which it almost closer for how human perceive things and the last layer in the lightness that can get the pure color in the image without any effect of light.

We conclude that K-Means algorithm get better result on segmentation that the color thresholding because K-Means in every adding point to the any class the indices should recalculate and recheck for which the point belong to.

Another conclude is the time complexity in classification that shouldn't need to check all images in the dataset the from the fist that the difference between then are large then you should drop it then we use Binary Search tree that in every iteration it drop path of images and it enhance the time complexity.

Future Work

We can develop our algorithm in many ways because it consists of more than two stages that can develop any one of them.

We can enhance our algorithm and reduce the time complexity by adding all clusters in the Binary Search Tree that every node will represent group. Each group will be also a binary search tree that represent the features of this group, when apply the K-NN algorithm on two Binary Search Tree then the complexity will be O (Log N * Log K) as the figure bellow.

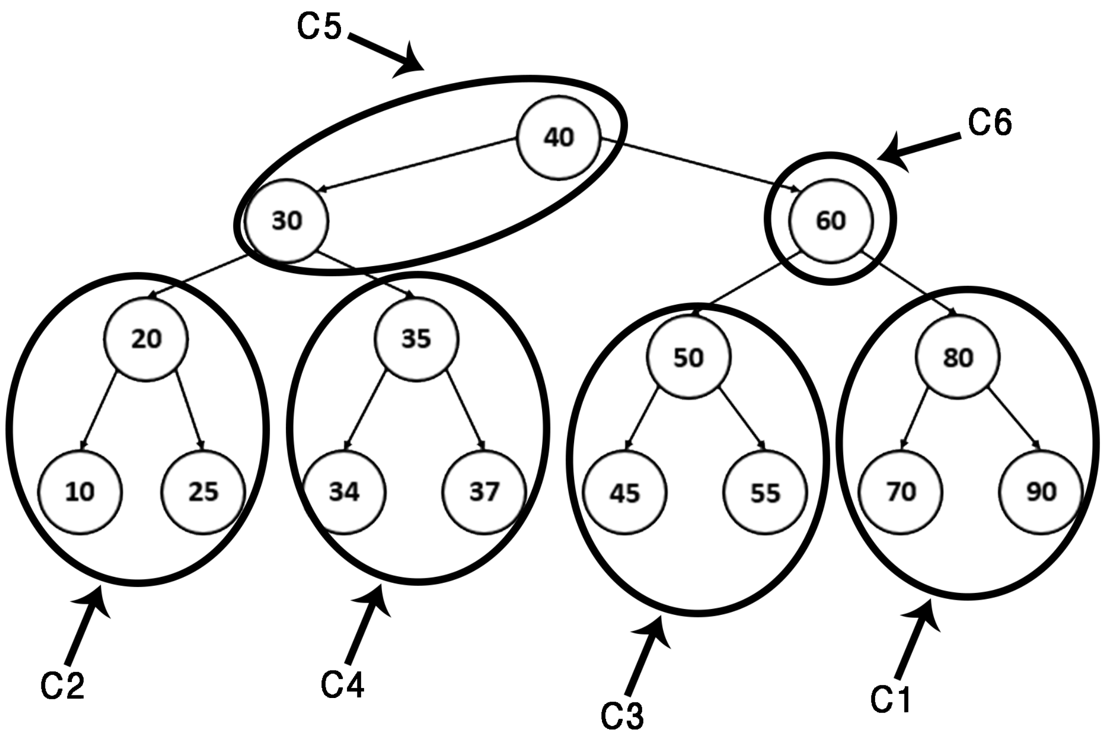

Figure 47 future work

And you can make it distributed group that posting the dataset inside huge binary search tree and every node has label of the class that in belong to and when the smallest difference appears then stop and get the answer.

The second thing that we can split the image to regions and then make all these process in every part that will make add another dimension in our work that is the position on the image.

By using the LAB color space, we can solve the sulhate of the images by take the hue directly without the lightness and reject the light ness the can create the new image with specific colors that it makes many problems in recognition method.

Another thing that use he neural network our system will take time just in the tanning step but when get the result for test input then the result will be fast , and when we use this technique we can make our system learn automatically without any interference from humans that will make our system more intelligence , and if we use it we can enhance the result by doing many tanning for our system that we can make input for many expression and make our system to conclude many features alone from itself.

References

[1] Paschos, George. "A Color Space Performance Comparison in the Processing of Color Textured Images: RGB vs. L* a* b." *PICS*. 1999.

[2] Wang, Hanzi, and David Suter. "Color image segmentation using global information and local homogeneity." *Proceeding of 7th Conf. of Digital Image Computing: Techniques and Applications*. 2003.

[3] Sun, Genyun, Aizhu Zhang, and Zhenjie Wang. "Grayscale Image Segmentation Using Multilevel Thresholding and Nature-Inspired Algorithms." *Hybrid Soft Computing for Image Segmentation*. Springer, Cham, 2016. 23-52.

[4] Kulkarni, Nilima. "Color thresholding method for image segmentation of natural images." *International Journal of Image, Graphics and Signal Processing* 4.1 (2012): 28.

[5] Vadivel, A., et al. "Segmentation using saturation thresholding and its application in content-based retrieval of images." *International Conference Image Analysis and Recognition*. Springer, Berlin, Heidelberg, 2004.

[6] Purohit, Amit D., and S. T. Khandare. "A Survey On Different Color Image Segmentation Techniques Using Multilevel Thresholding." *International Journal of Computer Science and Mobile Computing* 6.4 (2017): 267-273.

[7] Awad, Mamoun, et al. "A framework for image classification." *Image Analysis and Interpretation, 2006 IEEE Southwest Symposium on*. IEEE, 2006.

[8] Burney, SM Aqil, and Humera Tariq. "K-means cluster analysis for image segmentation." *International Journal of Computer Applications* 96.4 (2014).

[9] Javed, Muhammad Younus, and Usman Qayyum. "Face recognition using processed histogram and phase-only correlation (poc)." *Emerging Technologies, 2007. ICET 2007. International Conference on*. IEEE, 2007.

[10] Nock, Richard, and Frank Nielsen. "Statistical region merging." *IEEE Transactions on pattern analysis and machine intelligence*26.11 (2004): 1452-1458.

[11] Ohashi, Takumi, Zaher Aghbari, and Akifumi Makinouchi. "Hill-climbing algorithm for efficient color-based image segmentation." *IASTED International Conference on Signal Processing, Pattern Recognition, and Applications*. 2003.

[12] D. G. Lowe. Object recognition from local scale-invariant features. In ICCV '99: Proceedings of the International Conference on Computer Vision-Volume 2, page 1150, Washington, DC, USA, 1999. IEEE Com- puter Society.

[13] D. G. Lowe. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision, 60(2):91–110, 2004.

[14] Wang, Liming, et al. "Object detection combining recognition and segmentation." *Asian conference on computer vision*. Springer, Berlin, Heidelberg, 2007.

[15] Guru, D. S., Y. H. Sharath, and S. Manjunath. "Texture features and KNN in classification of flower images." *IJCA, Special Issue on RTIPPR (1)* (2010): 21-29.

[16] Burney, SM Aqil, and Humera Tariq. "K-means cluster analysis for image segmentation." *International Journal of Computer Applications* 96.4 (2014).

[17] Khurana, Khushboo, and Reetu Awasthi. "Techniques for object recognition in images and multi-object detection." *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)* 2.4 (2013): pp-1383.

[18] S. Belongie, et. al., "Color- and texture-based image seg- mentation using EM and its application to content-based image retrieval", *Proc. of ICCV*, p. 675-82, 1998. [2] M. Borsotti, P. Campadelli, and R. Schettini, "Quantitative evaluation of color image segmentation results", *Pattern Recogni- tion letters*, vol. 19, no. 8, p. 741-48, 1998.

[19] D. Comaniciu and P. Meer, "Robust analysis of feature spaces: color image segmentation", *Proc. of IEEE Conf. on Com- puter Vision and Pattern Recognition*, pp 750-755, 1997.

[4] Y. Delignon, et. al., "Estimation of generalized mixtures and its application in image segmentation", *IEEE Trans. on Image Processing*, vol. 6, no. 10, p. 1364-76, 1997.

[20] Bora, Dibya Jyoti, Anil Kumar Gupta, and Fayaz Ahmad Khan. "Comparing the performance of L* A* B* and HSV color spaces with respect to color image segmentation." *arXiv preprint arXiv:1506.01472* (2015).

[21] Sabha Muath, Peers Pieter, and Dutre Philip, Texture Synthesis using Exact Neighborhood Matching, Computer Graphics Forum, 26(2):131-142, 2007.

[22] Deng, Yining, B. Shin Manjunath, and Hyundoo Shin. "Color image segmentation." *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*. Vol. 2. IEEE, 1999.

[23] Singh, Dhirendra Pal, and Ashish Khare. "Automatic Image Segmentation using Threshold Based Parameter and Genetic Algorithm." *Science* 3.3 (2012).

[24] Wu, Jiajun, et al. "Deep multiple instance learning for image classification and auto-annotation." *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*. IEEE, 2015.

الملخص

في هذه الرسالة ، تم تطوير تقنية لتصنيف الصور تحت الإشراف الآلي خصيصًا لتصنيف الصور في مجال التراث الثقافي. بشكل عام ، يتم استخدام معظم تقنيات تصنيف الصور في النماذج المعروفة وشبه المرئية التي تعرف الكائنات التي يجب اكتشافها ، بينما تقوم التقنية المتقدمة بتصنيف الصور وفقًا لتاريخ معين وثقافة وشعب وعمر تاريخي. تتكون التقنية المقترحة من مرحلتين ، الأولى هي استخراج الميزات باستخدام تقنية تجزئة غير خاضعة للإشراف ، ثم مرحلة التصنيف غير الخاضعة للإشراف. تستخدم التقنية المطورة صبغة اللون فقط من مساحة ألوان CIE LAB للتجزئة ويتم استخدام -K للوسم التجميعي. يتم دمج بعض القطاعات للحصول على نتيجة للتراث الثقافي الذي له أكثر أهمية لهذه القطاعات. في مرحلة التعلم ، تم استخراج الميزات الشائعة ثم مقارنة الرسوم البيانية الخاصة بهم ، ثم تم تصنيفها وفقًا لذلك. أخيرًا ، تم دمج الأعمدة المجاورة في الرسوم البيانية لتقليل تعقيد الخوارزمية. أخيرًا ، اتبع هذه التقنية وتطبيقها على مستودع لصور التراث الثقافي ، وقلل من تعقيد الخوارزمية وحصل على دقة تقريبًا إلى 70٪ لمدة 20 وراثة ثقافية ، ولكل منها حوالي 150 صورة من 3200 صورة في مجموعة بيانات كاملة.