



A NEW METHOD OF COLORED IMAGE CLASSIFICATION USING UNSUPERVISED CLUSTERING METHOD

Thara Devi M¹, Pooja J.P², Ramya.K³, Sai Sowmya.B⁴, Shreeya Naik⁵

¹Assistant Professor, Department of Computer Science and Engineering,

Dr.T.Thimmaih Institute of Technology KGF

^{2,3,4,5}Students, Department Of Computer Science and Engineering,

Dr.T.Thimmaih Institute of Technology KGF

Emailid: ¹tharamani005@gmail.com

Abstract

In this paper, we introduce a new method of colored image classification using unsupervised clustering method. The proposed approach aims to train the image and partition into highly relative cluster. Two method are used such as 1) Image processing and 2) Classification stage. In image preprocessing stage the images are trained to limited colored images through the feature extraction and feature selection. In the feature extraction stage, mean, standard deviation and variance is applied to the each individual block and the three features are extracted from the each block in the digital image. In the feature selection stage, the size of the image feature is limited. In the classification stage, images are trained and it is portioned into highly relative clusters through the k mean technique. The experimental result shows the proposed approach is used to classification of colored image using unsupervised clustering method.

INTRODUCTION

Supervised and unsupervised way is used for image classification. Information class on the image can be specified when supervised classification starts. An algorithm is then used to form class signatures which summarize multispectral information from the specified areas on the image. This process is known as supervised training. In case of unsupervised classification, an algorithm is first applied to the image and some spectral classes (also called clusters) are formed. The image analyst tries to assign a spectral class to the desirable

information class an efficient method to identify and classify the exudates as soft and hard exudates was introduced.

Candidate exudates were detected using K-means clustering technique. In this system, classification of multispectral images introduces a hybrid clustering which is combination of k-means and PSO clustering. Here, initially k-means clustering was done and the result was used to see initial swarm. Also proposed a new unsupervised classification approach for automatic analysis of polar metric synthetic aperture radar (SAR) image. Here, an optimization problem was addressed as the classification of multidimensional SAR data space by dynamic clustering. A different approach was proposed, they used supervised features in the context of image classification and retrieval yields to excellent result and demonstrated how these supervised features can be effectively used for unsupervised image categorization that is for grouping semantically similar images.

In this paper, the results of some tests conducted to assess the K-means algorithm performance. The effective K-means method is used in producing good clustering results for many practical applications. K-means method is well known for its relatively simple implementation and proper results. A direct algorithm of k-means method requires time proportional to the product of the number of vectors and number of clusters per iteration. This is computationally very expensive for large datasets. The k-means algorithm is an iterative procedure and requires the number of clusters k to be given a priori.

Suppose that the k initial cluster centers are given, then the algorithm follows the steps given below:

1. Euclidean distance can be computed from each of the documents to each cluster center. A document is associated with a cluster, its distance from that cluster is the smallest one among all the distances.
2. After this association of all the documents with one of k clusters is done, then each cluster center is recomputed as to reflect the true mean of its constituent documents. Until the convergence gets achieved repeat both the steps.

In this paper, we have focused on a new method for image feature vector classification using unsupervised k -means clustering. Proposed approach partitions the trained image feature vector into highly relative clusters. The major goal of the proposed method is to partition and classify image feature vector with good accuracy.

OVERVIEW OF IMAGE PROCESSING

In the image processing the digital image are processed, i.e. the noises are removed and it also removes irregularities of the digital image. The images will be creped if the noise and irregularities are present this will happen either during the formation and transformation etc. In the mathematical analysis the images are defined by two dimensional function $f(a, b)$ a, b are the spatial coordinators the f is the amplitude the coordinate the coordinate (a, b) is called the colored level of image at point. If a, b and intensity value is finite discrete quantities, is called as the digital image. The digital image is composed of finite number of elements, the particular location and value is given. The elements consisting of all these features are called the picture element, image element, peels and pixels.

Image processing have been developed many techniques in last four to five decades. Image Processing are very famous because of availability of powerful personnel computers, graphics, software, memory devices. Image processing is used in various application such as Medical field, transmission and encoding and Microscopic Imaging.

The digital image is done generally in two dimensional picture by a digital computer. In

broader context, it implies two dimensional in the digital processing. An array of real numbers represented by a finite number of bits is called as the digital image. The advantages of Digital Image Processing is noise free digital imaging.

The various Image Processing techniques are:

- Image preprocessing: In image preprocessing, image data recorded by sensors on satellite restrain errors related to geometry and brightness values of the pixels. These errors are corrected using appropriate mathematical models which are either definite or statistical models. Image enhancement is the modification of image by changing the pixel brightness values to improve its visual impact. Image enhancement involves a collection of techniques that are used to improve the visual appearance of an image, or to convert the image to a form which is better suited for human or machine interpretation.
- Feature extraction: The feature extraction techniques are developed to extract features in synthetic aperture radar images. This technique extracts high-level features needed in order to perform classification of targets. Features are those items which uniquely describe a target, such as size, shape, composition, location etc.
- Image classification: It is the labeling of a pixel or a group of pixels based on its grey value. Classification is one of the most often used methods of information extraction. In Classification, usually multiple features are used for a set of pixels i.e., many images of a particular object are needed.
- Image Segmentation: It is the process that subdivides an image into its constituent parts or objects. The level to which this subdivision is carried out depends on the problem being solved, i.e., the segmentation should stop when the objects of interest in an application have been isolated.

MOTIVATION

Clustering as a technique of finding sub-groups within observations which is widely used in applications like market segmentation in which

we try and find some structure in the data. Although an unsupervised machine learning technique, in supervised machine learning model the clusters can be used as features. K-means is a clustering algorithm which divides observations into k clusters. Since, we can dictate the amount of clusters, it can be easily used in classification where we divide data into k clusters which can be equal to or more than the number of classes.

PROPOSED SYSTEM

This proposed technique approaches for image feature vector classification using unsupervised k-means clustering. The proposed approach consists of two stages namely (1) Image preprocessing stage (2) Classification stage. The image preprocessing stage aims to train the limited image feature vector set from the set of colored images through the feature selection and feature extraction. In feature extraction stage, three spatial statistical operators are applied such as mean, variance and standard deviation over the each individual block in the digital image and extracted three features from each block. In feature selection stage, size of the image feature vector set is limited. In classification stage, trained image feature vector set is partitioned into highly relative clusters through the k-means technique.

PRE-PROCESSING

Image enhancement can be done in two methods, Spatial and frequency domain. In the spatial domain approach, the pixel composing of image details are considered and the various procedures are directly applied on these pixels. The spatial technique approach are simple and easy to implement and speed of operation is high.

FEATURE EXTRACTION

In feature extraction method, the input image is divided into (8*8) blocks of smaller size. On each block, the applied three spatial statistical operators and extracted three features from each individual blocks such as mean, variance and standard deviation.

FEATURESELECTION

In feature selection step, the feature vector size is limited. Among all the features only mean values of all the input images are taken.

CLASSIFICATION STAGE

In classification stage, the trained image feature vector set is partitioned into highly relative clusters through the k-means technique.

ALGORITHM

K-MEANS ALGORITHM

K-means is the simplest unsupervised learning algorithms which solves the clustering problems. The method follows a simple and easy way to classify a given data set through a certain number of k clusters fixed to a priority. The idea is to define k centroids, to each cluster. These centroids should be placed in a greedy way because of different resultscauses different location. So, the best choice is to place them as far as much away from each other. The next step is to take the data set of each point associated with it to the nearest centroid. When there is no point pending, early grouping is done and first step is completed. At this stage we need to recalculate k new centroids of the clusters that occurred in the previous step. After obtaining the k new centroids, a new binding has to be done between the nearest new centroids and the same data set points. Generation of loop has taken place. After the generation of this loop wenote that the k centroids change their location step by step until there is no more changes. That is, the centroids do not move any more.

The main aim of this algorithm is to minimize an objective function; in this method a sum squared error function.

Even though it is proved that the procedure will terminate all the time, the k-means algorithm does not in all the situation find the most optimal configuration, corresponding to the global objective function minimum. The algorithm is also slightly sensitive to the beginning aimlessly selected cluster centers. The k-means algorithm can be run multiple times to reduce this effect.

The algorithm is composed of the following steps:

□ Place K points into the space that are represented by the objects that are clustered. These points are called as starting group centroids.

□ Allocate each object to the group that has the nearest centroid.

□ When all the objects are allocated, the positions of the K centroids has to be recalculated.

• Steps 2 and 3 should be repeated until the centroids can no longer move.

Input: Trained Image feature vector set of n images

X_i (i=1 to n) Number of clusters "m"

Output: "n" images clustered into "m" Clusters.

Begin

1. Randomly select $K = \{K_1, K_2 \dots K_m\}$ from image feature vector set X as candidate representative representing initial clusters.
2. Repeat
3. Calculate the distance between X_i ($i=1$ to n) and all „m“ cluster centers K_j ($j=1$ to m) using Euclidean distance and assign X_i to nearest cluster.
4. For each cluster, recalculate the cluster center using average mean function using
5. Until no changing of cluster centers or criterion function becomes minimum using

End

SYSTEM REQUIREMENT SPECIFICATION

System requirement specification is combined by extracting the appropriate information to implement system. It is an elaborative conditions which the system need to attain. The SRS delivers a complete knowledge of the system to understand what this project is going to achieve without any constraints on how to achieve this goal. This SRS not only provides the information to outside characters but also hides the plan.

2.1 HARDWARE REQUIREMENTS

- Processor: Windows 10
- RAM: 2GB and above
- Speed: 2.2 GHZ and above
- Secondary device: 25 GB and above

**SOFTWARE REQUIREMENTS
MATLAB**

High-performance language of MATLAB is used for multi-paradigm numerical computing environment and proprietary programming are done in easiest way, and the problems and solutions are expressed in mathematical way.

GENERAL ARCHITECTURE

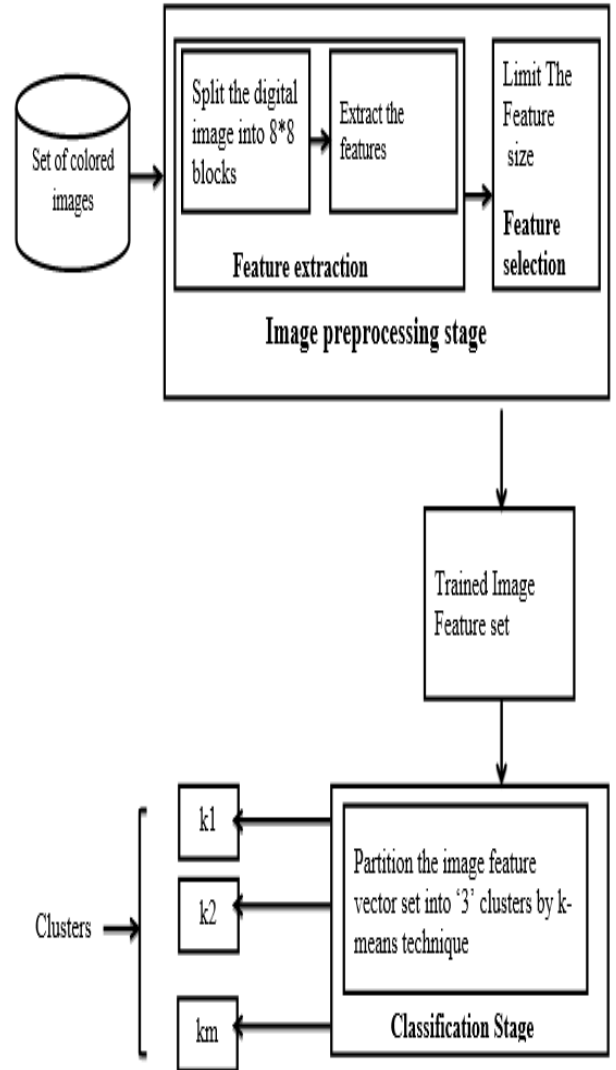


Fig 1. General architecture

EXPERIMENTAL RESULTS

In this section we tested the proposed system over 30 sample images as shown in the fig2 and these sample images are colored and the size is 200*200. And these datasets are divided into 3 clusters that is 1. Rose 2. Dinosaurs and 3. Horses

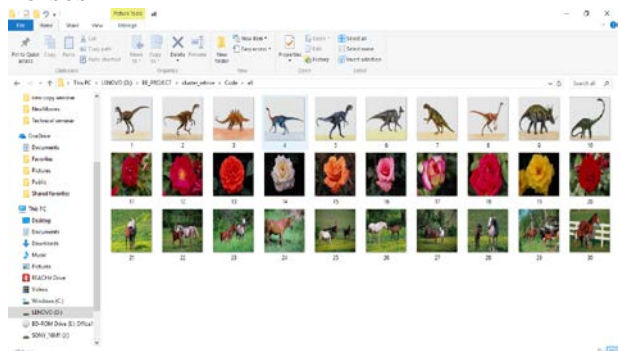


Fig2. Datasets

Fig 3 depicts the image of 30 colored samples getting trained

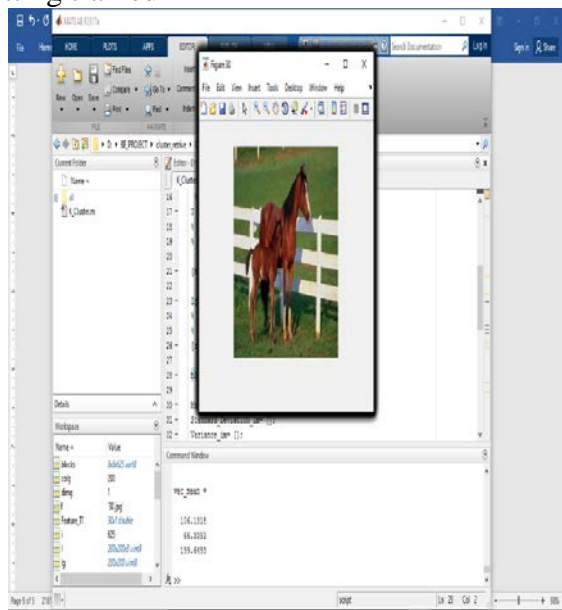


Fig 3. Training of images

CONCLUSION

In this paper, we introduced a new method for colored image classification using unsupervised clustering method. The proposed approach is aimed to partition the trained image feature vector set into highly relative clusters. The proposed approach consists of two stages viz.

- (1) Image preprocessing stage.
- (2) Classification stage.

The image preprocessing stage aims to train the limited image feature vector set from the set of gray scale images through the feature extraction and feature selection. In the feature extraction stage, applied three spatial statistical operators such that mean, standard deviation and variance over the each individual block in the digital image and extracted three features from each block respectively. In the feature selection stage, the size of the image feature vector set is limited. In the classification stage, the trained image feature vector set is partitioned into highly relative clusters through the k-means technique. We tested our proposed system over the 30 sample medical images and partitioned into three highly relative clusters or classes with high accuracy. According to the experiment results, our proposed system is better suitable for partitioned the images into highly relative classes for the image classification. In the future work we will improve our system for classifying

the new sample images that belong to which class or cluster in the existing classes.

REFERENCES

- [1] Rajput G.G, Patil and Preethi "Detection and classification of exudates using k-means clustering in color retinal images" Signal and Image Processing (ICSIP), 2014 Fifth International Conference 2014, pp126 - 130, 2014.
- [2] Venkatalakshmi.k Dept of I.T "Classification of multispectral images using vector machines based on PSO and K-means clustering" Intelligent sensing and information processing, proceeding of 2005 International Conference on Date 4-7Jan 2005, pp 127-133, 2005.
- [3] Turker ince "Unsupervised classification of polarimetric SAR image with dynamic clustering: An image processing approach Advances in Engineering Software", Science Direct Volume 41, Issue 4, pp 636- 646, 2009.
- [4] Gianluigi Ciocca, Claudio Cusano, Simone Santini, Raimondo Schettini "On the use of supervised features for unsupervised image categorization: An evaluation Computer Vision and Image Understanding", Volume 122, Issue null, pp155-171 (science digest) 2013.
- [5] Sanghoon Lee, Kyunggi-Do, Crawford, M.M "Unsupervised multistage image classification using hierarchical clustering with a bayesian similarity measure" Image Processing, IEEE Transactions Volume: 14, Issue: 3, pp 312-320, 2005.
- [6] Genitha, C.H, Vani, K "Classification of satellite images using new Fuzzy cluster centroid for unsupervised classification algorithm" Information & Communication Technologies (ICT), 2013 IEEE Conference on 11-12 April 2013, pp 203 - 207, 2013.
- [7] Frank Y. Shih, Gwotsong P. Chen "Classification of landsat remote sensing images by a fuzzy unsupervised clustering algorithm" Information Sciences - Applications, Volume 1, Issue 2, pp 97-116, 2013.
- [8] Beaulieu, J.M.; Touzi, R. "Mean-shift and hierarchical clustering for textured polarimetric

SAR image segmentation/classification in Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International Publication Year: 2010 , pp 2519 - 2522,2010.

[9] Kato, Z., Zerubia, J., Berthod, M. "Unsupervised parallel image classification using a hierarchical Markovian model" Computer Vision Proceedings, Fifth International Conference on 20-23 Jun 1995, pp169 - 174, 1995.

[10] Chardin, A. , Perez, P. "Unsupervised image classification with a hierarchical EM algorithm" Computer Vision, 1999. The Proceedings of the Seventh IEEE International

Conference on image processing Volume: 2 , pp 969 - 974 ,1999.

[11] <http://www.google.co.in/imgres?imgrefurl=http%3A%2F%2Fpaulbourke.net%2Fmiscellaneous%2Fcortex%2F&tbid=PTPNHg1rb9OTlM:&docid=iN4bzVS7JKb0M&h=388&w=600>

[12] Dr.M.P.S Bhatia and Deepika Khurana "Experimental study of Data clustering using K-means and modified algorithms" International Journal of Data Mining and Knowledge Management Process(IJDKM) Vol.3,No.3, May 2013.