EasyChair Preprint

№ 11384

Exploring Diverse Approaches to Image Classification: a Comprehensive Review

Savita Sangam and Rishov Saha

November 25, 2023

Exploring Diverse Approaches to Image Classification: A Comprehensive Review

Dr. Savita Sangam, Guide   Mr. Rishov Saha, Author

Dr. S. Sangam is with the Computer Engineering Department, as the head of the department, Shivaji Rao Jondhale College of Engineering, Asangaon, Shahapur, Maharashtra, PIN 421601 INDIA and is a guest lecturer at School of Engineering and Applied Sciences, Kalyan, University of Mumbai (e-mail: savita.sangam@gmail.com).

R. Saha is with the Computer Engineering Department, School of Engineering and Applied Sciences, Kalyan, University of Mumbai, Maharashtra, PIN 421301 INDIA (e-mail: rishov42@gmail.com).

Abstract

Image classification is a study which is a part of computer vision and is defined as categorization of images into various classes or categories using the software. Out of so many image classification techniques available, it becomes difficult to choose a desired image classification technique to achieve our desired purpose. To understand when we should apply which image classification technique, it is important to understand how the individual image classification techniques work. The purpose of this research paper is to discuss some of the image classification techniques. There are many image classification techniques available but this research paper covers two image classification techniques each from the two main categories of image classification which are supervised image classification and unsupervised image classification. Here, we theoretically understand what are the techniques, how the techniques are implemented and the advantages and disadvantages of the techniques. Finally, we conclude this research paper by stating that a single image classification technique should not be considered when classifying images; rather, a mixture of various image classification techniques should be used to get lucid and vivid results.

Keywords: computer vision, image classification technique, lucid, vivid

Statements and Declarations

I have submitted this research paper for the fulfillment of the requirement of Master of Technology in the field of Computer Engineering from University of Mumbai and does not hold any financial interests either directly or indirectly. No funding of any sort has been received in order to carry out this research. I have properly cited the references when and wherever required in this research paper.

Acknowledgement

In this paper I acknowledge the valuable inputs received from the respective research papers and articles whose references have been listed with proper citations wherever required. Also, this paper acknowledges the valuable recommendations received from Dr. Savita Sangam, guide and HOD of Computer Science Department of Shivajirao S. Jondhale College of Engineering and a guest lecturer at University of Mumbai. Also I would like to acknowledge the continuous support and blessings received from my parents.

1. 1. INTRODUCTION

Image classification refers to the process of categorization of images into specific categories. These are achieved with the help of specific algorithms applied with the aid of softwares. It finds its applications in multiple areas such as healthcare, security, autonomous driving systems, manufacturing (for quality control and defect detection), etc. Mainly, image classification has two types of techniques which are as follows:

1. Supervised classification
2. Unsupervised classification

Examples of image classification include detection of objects and classification of obstacles in and around an autonomous vehicle, detection of tumors during X-ray of a patient, target identification and surveillance in defense by automatically recognizing and categorizing objects in pictures and videos taken by satellites or unmanned aerial vehicles (UAVs), etc [1].

To achieve the task of image classification, mainly two types of datasets are required namely, test data and train data. Test data is that data which is used to test the software or model whereas the train data is that data which is used to train the software or model.

2. 2. TYPES OF CLASSIFICATION

a. 2.1. **Supervised Classification**

This process takes place under the guidance of a supervisor. The principle is that the supervisor selects sample pixels from the given dataset containing specific classes and then instructs the image processing software to use these training data as reference data for classifying other pixels in the image. The supervisor sets limits for how similar pixels must be so that they can be grouped together. The limits are set on the basis of the spectral characteristics of the training data. A tolerance for a certain increment is added in the training data. This tolerance is caused due to brightness or strength of reflection in some spectral bands. The supervisor also determines the number of classes that the image should be classified into.

The important types of supervised classification techniques are described below.

2.1.1. Convolutional Neural Network

It is a type of neural network architecture commonly used in computer vision to understand and interpret the image or visual data. It is considered as an extended version of the artificial neural network (ANN). It has multiple layers namely, convolutional layer, pooling layer and fully connected layer. The filter is applied to input images and the features are extracted, the pooling layer reduces computation by downsampling the images. Then, the fully connected layer makes the full prediction. Through backpropagation and gradient descent algorithms, the network learns the optimal filters [1].

2.1.2. Maximum likelihood

It is also known as Bayesian classifier.

This method applies probability theory for the task of classification. It does the following [2]:

Applies statistics.
Examines the probability function of a pixel for each class.
Assign the pixel to that class which has the highest probability.

Mathematically, it is defined as follows -

$$D = \ln(ac) - [0.5 * \ln(|Covc|)] - [0.5 * (X - Mc)^T * (1 / Covc) * (X - Mc)] \quad (1)$$

Where,

D = likelihood,

c = a particular class,

X = measurement vector of candidate pixel,

Mc = the mean vector of sample of class c,

ac = percent probability that any candidate pixel is a member of class c,

Covc = the covariance matrix of the pixels in sample of class c,

|Covc| = determinant of Covc,

1 / Covc = inverse of Covc.

b. 2.2. **Unsupervised Classification**

This process essentially implies that it will be guided by the software. The outcomes of this process are based on the analysis of the image done by the software without the need of any user intervention.. Using this technique, the pixels are grouped into classes. The user can specify the type of algorithm to be used by the software and the desired number of output classes. The only prerequisite is that the user needs to have full knowledge on the area in which classification will happen so that when the groupings of pixels with common characteristics produced by the computer are to be related to actual features, the user must understand as to what actually happened [2]. The methods that we will discuss in this type of classification are described below.

i.

ii.2.2.1. K-means Clustering

K-means Clustering determines the number of clusters k and assumes a midpoint of this cluster. Then any random object is taken and made as the initial centroid. This algorithm finally performs three steps for classification of images which are as follows:

1. Determine the coordinates of the centroid.
2. Determine the distance of each object to the centroids.
3. Grouping the object based on minimum distance.

Fig I. given below depicts the activity diagram for the algorithm of k-means clustering.

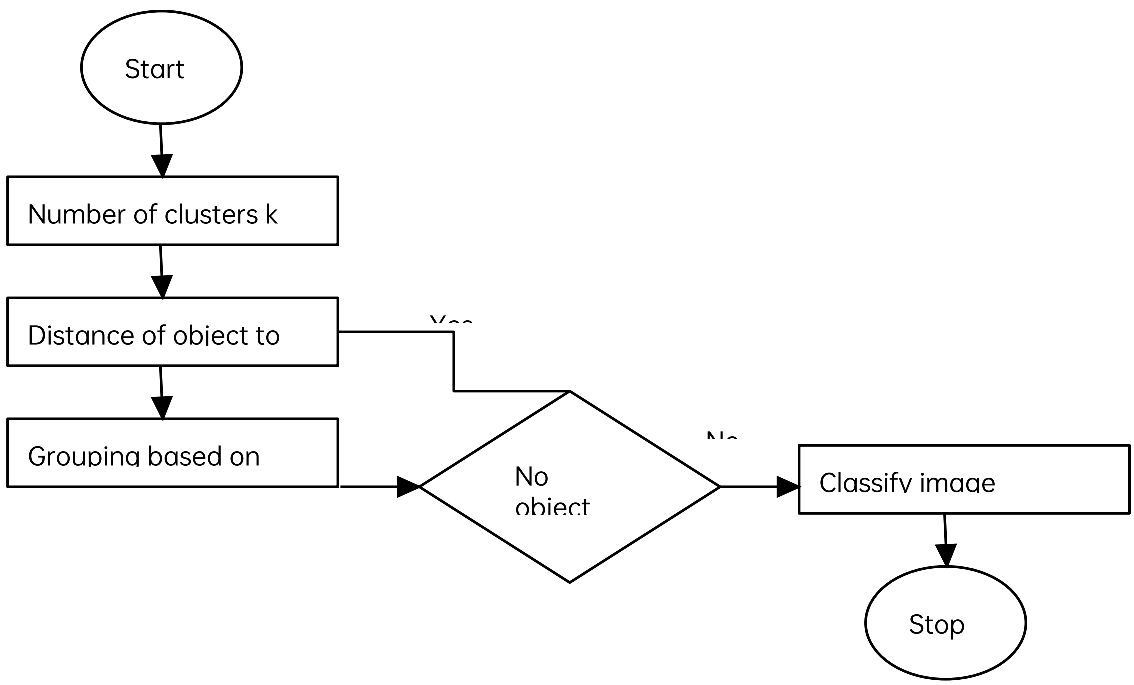

Fig I. Activity diagram showing k-means clustering algorithm

The condition for ending the iteration is the magnitude of the change of the mean from iteration i-1 to i summed over all K clusters [3].

2.2.2. Expectation Maximization Algorithm

It is a standard model whose purpose is to fit finite mixed models into observed data [3]. It is iterative in nature, and converges to a local maxima of the probability function.

The following equation defines this algorithm:

$$P(\theta/x) = P(x/\theta) * P(\theta) \quad (2)$$

where, θ = a set of unknown parameters from x.
It is a general method for estimating the features of a given data set, provided that the data are incomplete or have missing values [4]. Being an iterative procedure, its cost of computation is high.

The advantages and disadvantages of each of the methods have been tabulated in Table I.

Metrics	CNN	K-means	Maximum Likelihood	Expectation Maximization
Advantages	It has high accuracy rates.	It scales to large datasets.	It has the ability to provide the user with a consistent and flexible approach.	The steps are easy to implement.
	It has little dependency on preprocessing.	It is simple to implement.	It produces unbiased outcomes for larger samples.	It is guaranteed that likelihood increases with each iteration.
	It has the ability to handle large datasets.	It easily adapts to new examples.	It gets applied in different types of applications where other model's assumptions fail.	The E - step and M - step are easy to implement.
Disadvantages	It has difficulty with small datasets.	The mode of choosing k should be manual.	It becomes highly sensitive while choosing the initial values of a model.	It converges very slowly.
	It does not encode the position and orientation of objects.	It is sensitive to initial conditions.	It produces biased outcomes for smaller samples.	It converges to the local optima only.
	It requires a large amount of training data to train itself.	It is unable to handle categorical data.	Based on the complexity of the MLE (Maximum Likelihood Estimation) function , the numerical computation becomes quite expensive.	It requires both forward and backward probabilities.

Table I depicts the various advantages and disadvantages of the methods discussed here.

The author states that there is no conflict of interest.

References

3. https://research.aimultiple.com/image-classification/.

4. S. Domadia, Dr. T. Zaveri Comparative Analysis of Supervised and Unsupervised Image classification techniques, National Conference on Recent Trends in Engineering & Technology.

5. T. Lee, Md. Faizal, A. Fauzi and R. Komiya Segmentation of CT Brain Images Using K-means and EM Clustering. 5th IEEE International Conference on Computer Graphics, Imaging and Visualization, 2008.

6. T. Yamazaki Introduction of EM Algorithm into Color Image Segmentation. ATR Adaptive Communications Research Laboratories, 1998.