IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

INTELLIGENT IMAGE CLASSIFIER

A. Sai Charan [1], P. Meghana [2], G. Harsha Vardhan [3] B.Tech IV year, ECE
Dr. S.N. Chandra Shekar, Assistant Professor

Department of Electronics and Communication Engineering, Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India.

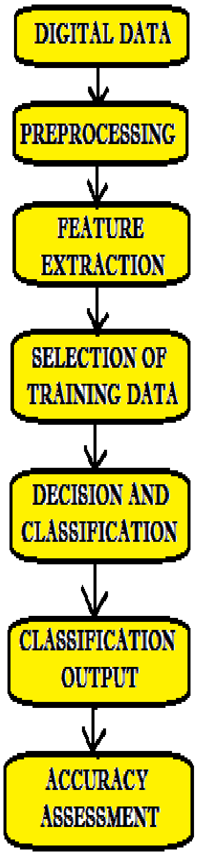
ABSTRACT
We know that all humans have neurons to store the data. Humans can say what is in an image. For example, we classify the images based on the terrains or landscapes present in them. So, can we do the same with a computer? Yes. The Intelligent Image Classifier is an Artificial Neural Network (Convolutional Neural Network) model that classifies the images based on the landscapes present in them. The model after its creation is linked with a front-end webpage and an android application in order to provide the user with a GUI. The user can choose an image that he wants to classify. Then, the model runs based on the classification of images into labels. Finally, the output is displayed on the screen. This model can be used to sort the images in the mobile phone based on the user's choice.

INTRODUCTION

When a human looks at an image, at the very next second, the human identifies the category of the image. For example ,if the image is of an apple ,it identifies the category as fruits .In Image Classification ,we assign the pixels of a particular picture of image to specific categories or classes ,based on the dataset which we train.

Some of the example images can be considered as categories of fruits , gadgets , automobiles etc. To classify the set of images into different categories ,the link between the data and the classes in which they are classified has to be clearly stated or understood .This can be achieved by training your computer .Here ,we train the system using the nueral networking to achieve the task .Classification Techniques were brought into light with the help of Pattern Recognition fields . Computer classification of remotely sensed images involves the process of the computer program learning the relationship between the data and the information classes. Traditional machine learning methods (such as multilayer perception machines, support vector machines, etc.) mostly use shallow structures to deal with a limited number of samples and computing units. When the target objects have rich meanings, the performance and generalization ability of complex classification problems are obviously insufficient. The deep neural network developed in recent years has been widely used in the field of image processing because it is good at dealing with image classification and recognition problems and has brought great improvement in the accuracy of many machine learning tasks. It has become a powerful and universal deep learning model.

With the rapid development of mobile Internet technology, more and more image information is stored on the Internet. Image has become another important network information carrier after text. Under this background, it is very important to make use of a computer to classify and recognize these images intelligently and make them serve human beings better. In the initial stage of image classification and recognition, people mainly use this technology to meet some auxiliary needs, such as Baidu's star face function can help users find the most similar star. Using OCR technology to extract text and information from images, it is very important for graph-based semi-supervised learning method to construct good graphics that can capture intrinsic data structures. This method is widely used in hyperspectral image classification with a small number of labeled samples. Among the existing graphic construction methods, sparse representation (based on SR) shows an impressive performance in semi-supervised HSI classification tasks. However, most algorithms based on SR fail to consider the rich spatial information of HSI, which has been proved to be beneficial to classification tasks. Yan et. proposed a space and class structure regularized sparse representation (SCSSR) graph for semi-supervised HSI classification. Specifically, spatial information has been incorporated into the SR model through graph Laplace regularization, which assumes that spatial neighbors should have similar representation coefficients, so the obtained coefficient matrix can more accurately reflect the similarity between samples. In addition, they also combine the probabilistic class structure (which means the probabilistic relationship between each sample and each class) into the SR model to further improve the differentiability of graphs. Hyion and AVIRIS hyperspectral data show that our method is superior to the most advanced method. The invariance extracted by Zhang et al., such as the specificity of uniform samples and the invariance of rotation invariance, is very important for object detection and classification applications. Current research focuses on the specific invariance of features, such as rotation invariance. In this paper, a new multichannel convolution neural network (mCNN) is proposed to extract the invariant features of object classification. Multi-channel convolution sharing the same weight is used to reduce the characteristic variance of sample pairs with different rotation in the same class. As a result, the invariance of the uniform object and the rotation invariance are encountered simultaneously to improve the invariance of the feature. More importantly, the proposed mCNN is particularly effective for small training samples. The experimental results of two datum datasets for handwriting recognition show that the proposed mCNN is very effective for extracting invariant features with a small number of training samples. With the development of big data era, convolutional neural network (CNN) with more hidden layers has more complex network structure and stronger feature learning and feature expression ability than traditional machine learning methods. Since the introduction of the convolutional neural network model trained by the deep learning algorithm, significant achievements have been made in many large-scale recognition tasks in the field of computer vision. Chaib et al. [18] first introduced the rise and development of deep learning and convolution neural network and summarized the basic model structure, convolution feature extraction, and pool operation of convolution neural network. Then, the research status and development trend of convolution neural network model based on deep learning in image classification are reviewed, and the typical network structure, training method, and performance are introduced. Finally, some problems in the current research are briefly summarized and discussed, and new directions of future development are predicted. Computer diagnostic technology has played an important role in medical diagnosis from the beginning to now. Especially, image classification technology, from the initial theoretical research to clinical diagnosis, has provided effective assistance for the diagnosis of various diseases. In addition, the image is the concrete image formed in the human brain by the objective things existing in the natural environment, and it is an important source of information for a human to obtain the knowledge of the external things. With the continuous development of computer technology, the general object image recognition technology in natural scene is applied more and more in daily life. From image processing technology in simple bar code recognition to text recognition (such as handwritten character recognition and optical character recognition OCR etc.) to biometric recognition (such as fingerprint, sound, iris, face, gestures, emotion recognition, etc.), there are many successful applications. Image recognition (Image Recognition), especially (Object Category Recognition) in natural scenes, is a unique skill of human beings. In a complex natural environment, people can identify concrete objects (such as teacups) at a glance (swallow, etc.) or a specific category of objects (household goods, birds, etc.). However, there are still many questions about how human beings do this and how to apply these related technologies to computers so that they have humanoid intelligence. Therefore, the research of image recognition algorithms is still in the fields of machine vision, machine learning, depth learning, and artificial intelligence.



Literature survey

Image classification is an important step in the object detection and image analysis. The output of the image classification step can be the final output or the intermediate output. A lot of image classification techniques have been proposed till date. Various studies have been conducted in order to conclude about the best satellite image classification technique. It is hard to decide any one technique as the best technique among all, because the results and its accuracy depend on a number of factors [2]. Over the last few decades, there is a constant modification in the conventional methods as well as invention of new image classification techniques in order to get maximum accurate results. Each of the classification technique has its own advantages and disadvantages. The research now concentrates on combining the desired features of these techniques in order to increase the efficiency. As the hard classifiers cannot handle the problem of mixed pixels, the soft classifiers are used. But, soft classifiers have their own disadvantages. [9] presents a study that combines the desirable features of a soft classifier and a hard classifier. It makes the use of LSMM, a soft classifier and SVM, a hard classifier and compared the results with the ones produced by LSMM and SVM separately. The results showed that the combination of both classifiers produced better results when compared to either of them. Another study was conducted on PoISAR data combined the Markov random field smoothness constraint with supervised Softmax regression model [6]. The experiments conducted during this study proved that the combination of supervised and unsupervised algorithms produced better results as compared to the ones produced by either of these techniques separately. [16] presents the combination of fuzzy logic and neural networks in order to design a system that can detect the face and fingerprints of the person. This is done in order to determine the authenticity of the person. This system can be used for various security purposes

The image classification is a classical problem of image processing, computer vision and machine learning fields. In this paper we study the image classification using deep learning.

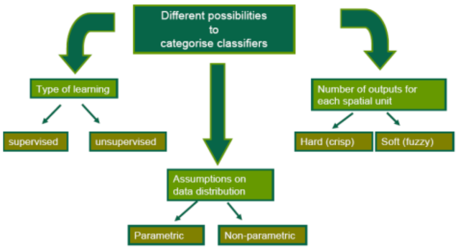
We are creating a dataset ,which contains information of different categories ,for example ,category of automobiles ,fruits ,animals etc. When an input is given ,the system identifies the image and categorizes.Also,when there are n number of images in a certain site ,this sytem could classify all those images into their respective categories.This could be clearly explained like this,in fashion sites,the items are categorized based on their style i.e.,dress style like skirts,t-shirts,tops etc,also,beauty items like foundation box,concealers etc all are categorized so that it would be easy for the customers to shop in that site.These are the areas where we could clearly observe the application of Image classification.

To identify criminals from a blurred image,this is also the one problem we are going to address through this project.Also,in medical imagimg image classification system eases the work of doctors.The problem location could be found using this image classification system.The approach for our project is building a dataset first,then we need to create a front end page,from where we could give the input.This front end page can be optional,we can directly take the input from the google and also could give to the system.The next step would be integrating the classification model that means the datset and the front end page for ease of the users.At the end step,the testing of the project is carried out.

Image classification plays an important role in remote sensing images and is used for various applications such as environmental change, agriculture, land use/land planning, urban planning, surveillance, geographic mapping, disaster control and object detection, and also it has become a hot research topic in the tech industry now.

Existing methodology/Background

4.1 IMAGE CLASSIFICATION APPROACHES



There exists different types of techniques to achieve image classification.The major techniques used in deep learning to achieve this are discussed in this article.

1.UNSUPERVISED CLASSIFICATION

In unsupervised classification, it first groups pixels into "clusters" based on their properties. Then, you classify each cluster with a land cover class.

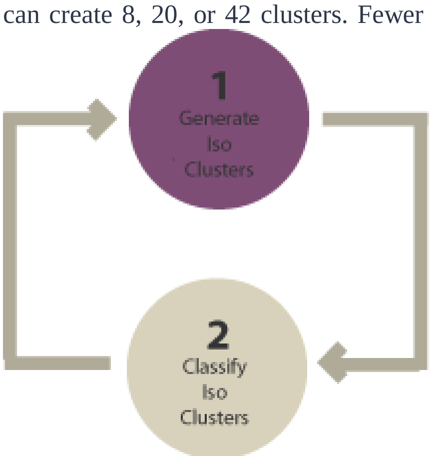
Overall, unsupervised classification is the most basic technique. Because you don't need samples for unsupervised classification, it's an easy way to segment and understand an image.

The two basic steps for unsupervised classification are:

- Generate clusters
- Assign classes

- Using **remote sensing software**, we first create "clusters". Some of the common image clustering algorithms are:
- K-means
- ISODATA

After picking a clustering algorithm, you identify the number of groups you want to generate. For example, you can create 8, 20, or 42 clusters. Fewer clusters have more



resembling pixels within groups. But more clusters increases the variability within groups.

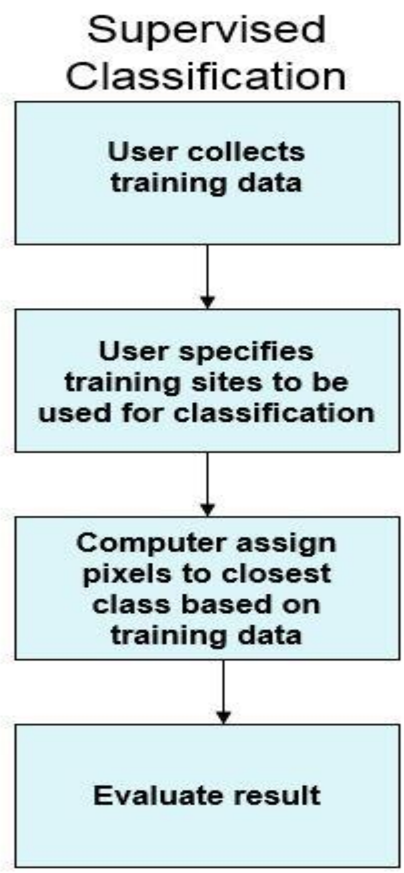
To be clear, these are unclassified clusters. The next step is to manually assign land cover classes to each cluster. For example, if you want to classify vegetation and non-vegetation, you can select those clusters that represent them best.

2.Supervised Classification

In supervised classification, you select representative samples for each land cover class. The software then uses these "training sites" and applies them to the entire image.

The three basic steps for supervised classification are:

- Select training areas
- Generate signature file
- Classify
- For supervised image classification, you first create training samples. For example, you mark urban areas by marking them in the image. Then, you would continue adding training sites representative in the entire image



For each land cover class, you continue creating training samples until you have representative samples for each class. In turn, this would generate a signature file, which stores all training samples spectral information.

Finally, the last step would be to use the signature file to run a classification. From here, you would have to pick a classification algorithms such as:

- Maximum likelihood
- Minimum-distance

- Principal components
- Support vector machine (SVM)
- Iso cluster

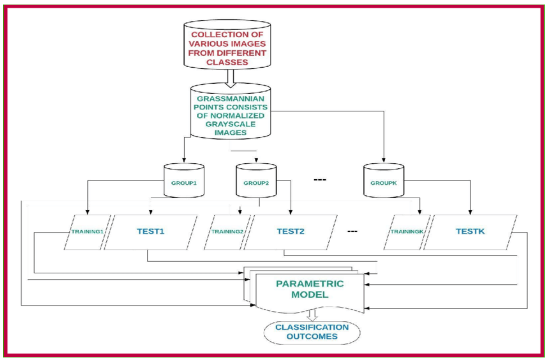
3.PARAMETRIC CLASSIFICATION

The performance of a parametric classifier depends largely on how well the data match the pre-defined models and on the accuracy of the estimation of the model parameters. The parameters like mean vector and covariance matrix are used.

For example:- maximum likelihood classifier

The parametric and non-parametric image classification techniques come under the supervised learning. The parametric classifiers use the algebraic possibility for allocation to each class. Some of the parametric classifiers are Bayesian Classifier, Naïve Bayes classifier and decision tree. The parameters required are taken from the training data. Parameters like mean and co-variance are used in these classifiers Models of data with a categorical response are called *classifiers*. A classifier is built from *training data*, for which classifications are known. The classifier assigns new *test data* to one of the categorical levels of the response.

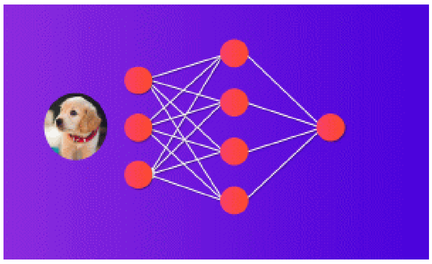
Parametric methods, like Discriminant Analysis Classification, fit a parametric model to the training data and interpolate to classify test data.



Flowchart of the general parametric Classification Model defined in this thesis. The model consists of the shape and appearance of the distribution, where the shape is represented by the normalizing constant for the location of features and enables the proposed model to achieve scale and translation invariance. Each appearance is a point on the Grassmann manifold. The decision is based on calculating the log-likelihood function of trained normalizing constant (shape of the distribution) and appearance points (or rapping distribution). For all of the experiments, the normalizing constant or shape of the distribution was trained for each class with some reserved data, while the rest of the data points were used in the appearance part of the distribution for a few random trials for the final outcomes in the Bingham distribution via Bayesian classifier by straightforward MLE.

Challenges with Artificial Neural Network (ANN)

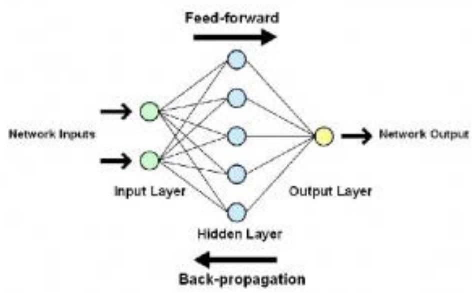
- While solving an image classification problem using ANN, the first step is to convert a 2-dimensional image into a 1-dimensional vector prior to training the model. This has two drawbacks:
- The number of trainable parameters increases drastically with an increase in the size of the image



In the above scenario, if the size of the image is 224*224, then the number of trainable parameters at the first hidden layer with just 4 neurons is 602,112. That's huge!

ANN loses the spatial features of an image. Spatial features refer to the arrangement of the pixels in an image. I will touch upon this in detail in the following sections

- One common problem in all these neural networks is the Vanishing and Exploding Gradient. This problem is associated with the backpropagation algorithm. The weights of a neural network are updated through this backpropagation algorithm by finding the gradients:



So, in the case of a very deep neural network (network with a large number of hidden layers), the gradient vanishes or explodes as it propagates backward which leads to vanishing and exploding gradient.

ii)DECISION TREE

DT are knowledge based (i.e. a method of pattern recognition that simulates

the brains inference mechanism).

DT are hierarchical rule based approaches and use non-parametric approach.

DT predict class membership by recursively partitioning a dataset into

homogeneous subsets.

Different variables and splits are then used to split the subsets into further

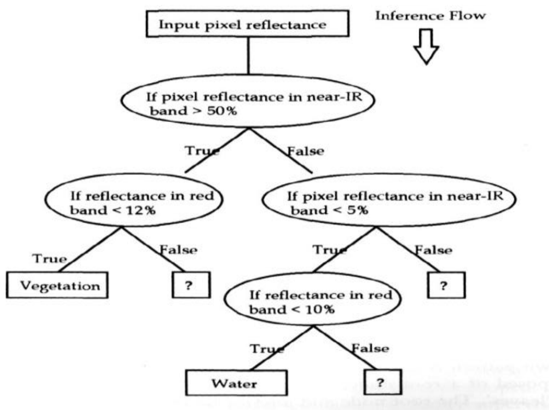
subsets.

It consists of 3 parts:

Partitioning the nodes

Find the terminal nodes

Allocation of class label to terminal nodes



Decision tree is a tree based classification, widely used in data mining, which classifies the input data set into predefined classes. Decision tree approach is used here to train the image understanding system to perform supervised machine learning. The various low level characteristic features (color, shape, texture) of the image form the various attributes of the decision tree among others.

.

Advantage: -

Can handle non-parametric training data.

Does not required an extensive design and training.

Simple and computational efficiency is good.

Disadvantage: -

Becomes complex calculation when various values are undecided and/or when various outcomes.

Conclusions and Future Directions

In this work we have explored the problem of bottom-up figure-ground segmentation, both as an image segmentation task, and as a perceptual grouping problem. We presented an comprehensive overview of current research in both fields, and discussed reasons why despite a vast research effort, image segmentation and perceptual grouping on unconstrained images continue to be extremely challenging. Up to this point, we have not discussed the relationship between image segmentation and perceptual grouping. It is clear from our research, that our current implementations of each of these methodologies are particularly well suited to certain classes of images. Our contour extraction algorithm can be expected to yield good results in reasonably structured environments, and on images that contain simple objects. SE-MinCut proves particularly useful for dealing with less structured environments, and with irregularly shaped objects. Figure 9.1 illustrates some of these points. Image segmentation and perceptual grouping have traditionally relied on different image cues. Segmentation is often based mostly on pixel appearance, be it by using brightness, colour, or some measure of texture similarity (though the issue of cue integration for segmentation has received a reasonable amount of attention, see [62, 63]); whereas perceptual grouping usually relies on the information provided by image edges, and on grouping principles that exploit two images with different characteristics, and results from SE-MinCut, and our contour extraction method. Notice that on the top image which comes from a structured environment, contour extraction performs particularly well. Conversely, on the bottom image, which comes from a relatively un-structured environment, SE-MinCut works best. In general, contour extraction using only edge information is not well suited for un-structured environments, or objects whose boundaries are too irregular. It should be noted that SE-MinCut performed segmentation on down sampled versions of each image. In this thesis, we have made contributions to the fields of image segmentation and perceptual grouping. Within the field of image segmentation, we have presented spectral embedding a general technique for clustering data. We have shown a direct connection between spectral embedding and anisotropic, smoothing kernels. We've used spectral embedding to generate seed regions for min-cut, proposed an algorithm for combining seed regions into source/sink pairs, and shown that the resulting partitions capture salient image structure. We proposed a complete segmentation algorithm, and showed a visual comparison of segmentation results between SE-MinCut and three of the leading segmentation algorithms. We also carried out a thorough quantitative evaluation of segmentation quality over the Berkeley Segmentation Database. We proposed a suitable algorithm for matching the boundaries of two segmentations, and used precision/recall metrics based on this boundary matching to measure segmentation quality. Our results indicate that SE-MinCut produces better segmentations across its range of input parameters, in particular, SE-MinCut is capable of capturing a higher fraction of the salient boundaries of an image with less over-segmentation than competing algorithms, and the boundaries themselves are more accurately localized.

REFERENCES

[1] Arnon Amir and Michael Lindenbaum. A generic grouping algorithm and its quantitative analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(3):168–186, February 1998.

[2] Adrian Barbu and Song-Chun Zhu. Graph partition by Swendsen-Wang cuts. In IEEE International Conference on Computer Vision, pages 320–327, 2003.

[3] M. Belkin and P. Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering. In ”T.G. Diettrich, S. Becker, and Z. Ghahramani”, editors, ”Adv. in Neural Inf. Proc. Sys. 14”, pages 585–591. MIT Press, 2002.

[4] Serge Belongie, Charless Fowlkes, Fan Chung, and Jitendra Malik. Spectral partitioning with indefinite kernels using the Nystrom¨ extension. In European Conference on Computer Vision, pages 21–31, 2002.

[5] Serge Belongie and Jitendra Malik. Finding boundaries in natural images: A new method using point descriptors and area completion. In European Conference on Computer Vision, pages 751–766, 1998.

[6] J. Ross Beveridge, Joey Griffith, Ralf R. Kohler, Allen R Hanson, and Edward M Riseman. Segmenting images using localized histograms and region merging. International Journal of Computer Vision, 2(3):311–347, jan. 1989.

[7] C. M. Bishop. Neural Networks for Pattern Recognition. Oxford Univ. Press, 1995.