Image classification using Machine Learning Approaches

Dr. H S Prasantha

Professor, K.S.Institute of Technology, Bangalore, Karnataka, India

drhsprashanth@gmail.com

ABSTRACT

In machine learning, classification is a supervised learning concept which basically categorizes a set of data into classes. The most common classification problems are – speech recognition, face detection, handwriting recognition, and image classification. Image Classification is one of the most fundamental tasks in computer vision. It has revolutionized and propelled technological advancements in the most prominent fields, including the automobile industry, healthcare, manufacturing, and more. Image Classification (often referred to as Image Recognition) is the task of associating one (single-label classification) or more (multi-label classification) labels to a given image. Image classification is the process of predicting a specific class, or label, for something that is defined by a set of data points. Image classification is a subset of the classification problem, where an entire image is assigned a label. Grouping images into semantically meaningful categories using low-level visual features is a challenging and important problem in content-based image retrieval. Based on these groupings, effective indices can be built for an image database

Keywords—Image, machine learning, Classification

#  INTRODUCTION

An image is a 2-dimensional signal. It defined by the mathematical function where x is the horizontal coordinate and y is the vertical coordinate. At any point the value of f(x,y) gives the pixel value of the image at that point. Operations done on image to manipulate and extract useful information from image is called image processing. In general, it is classified into few groups, they are checking for presence, object detection and localization, measurement as well as identification and verification. Fundamental steps in digital image processing are image acquisition, image enhancement, image restoration, color image processing, image classification, compression, segmentation, and morphological processing. Classification is used to identify the category of new observations on the basis of training data. It predicts ‘Categorical class labels’ like yes or no, male or female, true or false, red or blue, disease or no disease.

Ex: When filtering emails “spam” or “not spam”, When looking at transaction data, “fraudulent”, or “authorized”

Image classification is the process of categorizing and labelling groups of pixels or vectors within an image based on specific rules. Image classification is where a computer can analyse an image and identify the image falls under. Algorithms used for classification are Logistic Regression, Naive Bayes, K-Nearest Neighbors, Decision Tree, and Support Vector Machines (SVM), Spam filtering, Random forest, Decision Tress.

**II. LITERATURE SURVEY**

|  |  |  |  |
| --- | --- | --- | --- |
| TITLE | OBJECTIVE | MACHINE LEARNINGALGORITHM | TOOLSUSED, REMARKS |
| Classification of IRISDataset using Classification Based KNN Algorithm in supervised learning. -By Thirunavukkarasu K, Ajay S. Singh, Prakhar Rai, Sachin Gupta | The objective is tobuild a model that is able to automatically recognize the iris species (Setosa, Versicolor and Virginica) and classify them. | Supervised learning, K-Nearest Neighbors (KNN) classification Algorithm. | Numpy, Pandas, Matplotlib and machineLearning library Scikit-learn. Able to recognize the iris species accurately on the basis of 3 classes, but some samples provide the misclassified result. Prediction for class 0 and class 2 was 100% correct but prediction for class 1 was 4% wrong. |
| Automatic fruit classification usingrandom forest algorithm-By Hossam M Zawbaa,Maryam Hazman, MonaAbbass, Aboul Ella Hassaien | The aim of this paper is to develop an effective classification approach based on Random Forest (RF) algorithm to classify fruits. | The features extraction stage uses two algorithms for extracting the fruit images feature which are shape and colour algorithm and ScaleInvariant Feature Transform (SIFT) algorithm. The classification is done using the Random Forest (RF) algorithm. | Matlab RF based algorithms provided better accuracy when compared to the other well known machine learning techniques such as K-NN and SVMAlgorithms. |
| Fruit Classification using Statistical Features in SVM Classifier -By R. S. S. Kumari & V. Gomathy | Proposes a technique uses colour and texture features for fruit classification. | Support Vector Machine(SVM) classifier is used to classify the different kinds of fruits. | The overall classification accuracy is 95.3%. |
| Classifications of High Resolution Optical Images using Supervised algorithmsBy B. Battula, L. Parayitam, T. S. Prasad, P. Balakrishna and C.Patibandla | To perform land use classification of optical data | Maximum likelihood (ML) and Support vector machines (SVM). | Comparison showed anOverall superiority of the Support Vector Machine over maximumLikelihood. |
| PCA-Based Animal Classification System-By Emre Dandil, Rukiye Polattimur | In this study, a Principal componentAnalysis (PCA) based application was developed for the recognition andclassification of different species of animals. | Principal ComponentAnalysis (PCA | Experimental studies oncow, cat, dog, goat andrabbit animal species shows a success rate of 92% in the first nearestrecognition and 83% inthe second nearestrecognition. |
| On Image Classification: City vs Landscape -By Aditya Vailaya, Anil Jain,and Hong Jiang Zhang | how a specific high levelClassification problem (city vs. landscape classification) can be solved from relatively simple low-level features suited for the particular classes. | k-Nearest Neighbour classifier | Classification systemresults in an accuracy of93:9% when evaluatedon an image database of2,716 images using theleave-one-out method. |
| Bacteria Classificationusing Image processingand deep learning -By Treesukon Treebupachatsakul1 , Suvit Poomrittigul | To research thepossibility to useimage classificationand deep learningmethods for classifyinggenera of bacteria. | LeNet CNN method PythonProgramming and the KerasAPI with TensorFlow Machine Learning framework | The experimental results have confirmed that two species of bacteria in different cell shape, Staphylococcus aureus (spherical or round shaped) andLactobacillus delbrueckii (long-rod shaped) are able to automatically predict using machine learning by image classificationand deep learning method. |
| Improvement ofClassification accuracyusing image fusiontechniques-By Rajesh Singh, Rajiv Gupta | In this paper, study area has been classified into three classes i.e. settlement, trees and agricultural by classification of an image which has been enhanced using fusion of two images. | The Brovery, Multiplicative and Principal ComponentAnalysis (PCA) method have been used for imagefusion. The resultant images have been classified using the supervised classification with maximum likelihoodparametric rule | The fused images give better classification results than the original Image. The Brovery fused image provides better results than other Fused images. Out of three fusion techniques the classification accuracy of Broveryfusion has the highest accuracy as 99.67 % than other techniques  |
| Classification of X-rayImages Using Grid Approach-BY Bertalya, Prihandoko, DjatiKerami, Tb. Maulana Kusuma | Dealing with theclassification ofmedical image to the image classes | Euclidean distance to obtain image similarity. FreemanCode to represent the shape of X-ray images and JeffreyDivergence technique, image classification processby using grid approach. | The result of this experiment for a number of images shows that there is a different score for Euclidean Distance and Jeffrey Divergence techniques of similarity measurement. The recognition rate of Jeffrey Divergence is better than Euclidean Distance. |
| Image Classification Skin Cancer Detection with CNN -By Eman Othman Fathy, DinaAlaa Ahmed | To classify skin cancer to detect if it’s benign or malignant. | Convolutional neural network (CNN) | Keras Helps doctors to determine if it isdangerous or not and to maintain its way of treatment. |

**III. TOOLS**

* Python programming higher installed in windows 10 or any distribution of Linux.
* Arduino IDE software
* Training of an ANN model using Coding Source in MATLAB.
* Python open source library called TensorFlow1 is used in order to train the neural network.

**V. APPLICATIONS**

● Medical imaging: Its aim is to classify medical images into different categories to help doctors in disease diagnosis or further research. It addresses the problem of diagnosis, analysis and teaching purposes in medicine.

● Object identification in satellite images: These models involve two steps. In the first step, the regions of presence of objects in the image are detected. In the second step, the objects are classified using convolution neural networks. In this work, a customised convolution neural network is proposed to detect and classify objects in satellite images.

● Machine vision: It allows for the classification of a given image as belonging to one of a set of predefined categories.

● Image classification with localization involves assigning a class label to an image and showing the location of the object in the image by a bounding box. Some examples of image classification with localization include: Labelling an x-ray as cancer or not and drawing a box around the cancerous region. Classifying photographs of animals and drawing a box around the animal in each scene.

● Image Classification for Websites with Large Visual Databases With image recognition, companies can easily organize and categorize their database because it allows for automatic classification of images in large quantities. This helps them monetize their visual content without investing countless hours for manual sorting and tagging.

**VI. REFERENCES**

1. R. S. S. Kumari and V. Gomathy, &quot;Fruit Classification using Statistical Features in SVM Classifier&quot;, 2018 4th International Conference on Electrical Energy Systems (ICEES), 2018, pp.526-529
2. B. Battula, L. Parayitam, T. S. Prasad, P. Balakrishna and C. Patibandla, &quot;Classifications of High Resolution Optical Images using Supervised Algorithms,&quot; 2018 IEEE 8th International Advance Computing Conference (IACC), 2018, pp. 126-130
3. E. Dandil and R. Polattimur, &quot;PCA-Based Animal Classification System,&quot; 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2018, pp. 1-5
4. A. Vailaya, A. Jain and Hong Jiang Zhang, &quot;On image classification: city vs. landscape,&quot; Proceedings. IEEE Workshop on Content-Based Access of Image and Video Libraries (Cat.No.98EX173), 1998, pp. 3-8
5. T. Treebupachatsakul and S. Poomrittigul, &quot;Bacteria Classification using Image Processing and Deep learning,&quot; 2019 34th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), 2019, pp. 1-3
6. K. Thirunavukkarasu, A. S. Singh, P. Rai and S. Gupta, &quot;Classification of IRIS Dataset using Classification Based KNN Algorithm in Supervised Learning,&quot; 2018 4th International Conference on Computing Communication and Automation (ICCCA), 2018, pp. 1-4
7. H. M. Zawbaa, M. Hazman, M. Abbass and A. E. Hassanien, &quot;Automatic fruit classification using random forest algorithm,&quot; 2014 14th International Conference on Hybrid Intelligent Systems, 2014, pp. 164-168
8. R. Singh and R. Gupta, &quot;Improvement of Classification Accuracy Using Image Fusion Techniques,&quot; 2016 International Conference on Computational Intelligence and Applications (ICCIA), 2016, pp. 36-40
9. Bertalya, Prihandoko, D. Kerami and T. M. Kusuma, &quot;Classification of X-Ray Images Using Grid Approach,&quot; 2008 IEEE International Conference on Signal Image Technology and Internet Based Systems, 2008, pp. 314-319
10. N. Aburaed, A. Panthakkan, M. Al-Saad, S. A. Amin and W. Mansoor, &quot;Deep Convolutional Neural Network (DCNN) for Skin Cancer Classification,&quot; 2020 27th IEEE International Conference on Electronics, Circuits and Systems (ICECS), 2020, pp. 1-4