International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2019 IJSRCSEIT | Volume 5 | Issue 2 | ISSN : 2456-3307
DOI : https://doi.org/10.32628/CSEIT195298

Image Classification using CNN and Machine Learning

G. Keerthi Devipriya[1], E. Chandana[2], B. Prathyusha[3], T. Seshu Chakravarthy[4]

[1,2,3] B.Tech, Department of Computer Science and Engineering, VVIT, Guntur , Andhra Pradesh, India

[4] Assistant Professor, Department of Computer Science and Engineering, VVIT, Guntur , Andhra Pradesh, India

*Corresponding Author: keerthidevipriya@gmail.com

ABSTRACT

Here by in this paper we are interested for classification of Images and Recognition. We expose the performance of training models by using a classifier algorithm and an API that contains set of images where we need to compare the uploaded image with the set of images available in the data set that we have taken. After identifying its respective category the image need to be placed in it. In order to classify images we are using a machine learning algorithm that comparing and placing the images.

Keywords :- Classifier, Feature Extraction, BoF, Supervised Learning

I. INTRODUCTION

The purpose of this paper is to classifying and placing the images by using some certain number of steps like Feature Extraction and Feature Analysis and Visualization. Hence this paper attempts to give the different feature extraction and classification algorithms.

The paper flow will goes on this way first the concept of related work then there is explanation regarding the process of machine learning and then the Image classification process has been explained and then the experiment and evaluation has been derived then conclusion.

II. RELATED WORK

D G Lowe et al. [4] This uses labeling and can deal with the large intensity variations where the image is taken into binary pixels where there is no direct interaction between pixels and hence the detection is bit difficult. Hecht Neilson R et al [11] This uses a feed forward technique for processing where is a trivial task and it may take very long time. This approach will use some random weight and transforms into a form that is used by the output layer. Graham B et al [8] This approach uses a globally connected network where the images are reduced in the matrix form but there is the problem of over fitting so that this technique is less efficient. Wan L et al [7] This states that for the classification it sends some random size of data to the training model where this approach also ends with the problem of over fitting and the process done in the linear way that is from left to right. Matthew D et al [2] It uses a multilayered De-convolution network and it considers the images as objects. where it need to know about the orientation and accurate computer reconstruction.

III. METHODS AND MATERIAL

3.1 MACHINE LEARNING PROCESS

Machine Learning is an algorithm set especially suited to prediction. These ML methods are easier to implement and perform better than the classical statistical approaches. Instead of starting with a data model, ML learns the relationship between the

response and its predictors by the use of algorithms. During the learning phase ML algorithms observe inputs and responses in order to find the dominant patterns. The Machine Learning Workflow is shown in Fig 1.

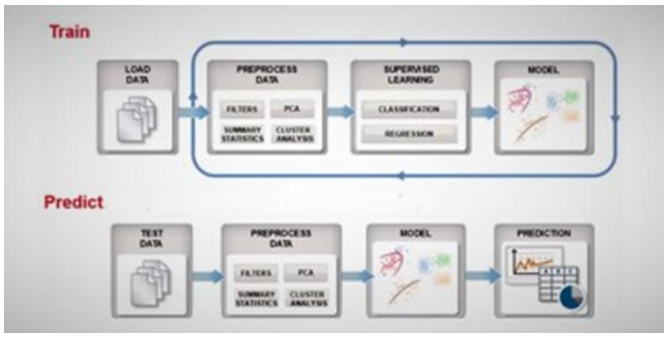

Fig 1 : Workflow of Machine Learning

Many fields of modern society use Machine Learning technologies like web searches, content filtering on social networks, recommendations on e-commerce websites, computer vision, etc. In this we are interested in implementing the classification of image by using a machine learning algorithm utilizing the strategy of supervised learning. The image that is taken to classify has undergone the classification process. Then based on the result the image is placed into its category.

3.2 IMAGE CLASSIFICATION USING BAG OF FEATURES

The concept of Bag of Features has been inspired by Bag of words. In this model just like the Bag of Words every feature of the image has been considered as the word and comparison has been applied similar to that of the word comparison in the documents.

A. AFE(Accurate Feature Extraction) Technique

For Feature detection and Feature Extraction the technique of AFE has been used. To classify the image there is a step by step mechanism to be done like Feature extraction, Feature Description and visualization.

1)Feature Extraction: In Machine Learning any processing technique whether pattern recognition or image processing feature extraction starts from an initial set of measured data and build derived values intended to be informative and non- redundant [2].

Feature Extraction is dimensionally reduction process where an initial set of raw variables is reduced to more manageable groups for processing while still accurately and completely.

When the input data to an algorithm is too large to be processed and it is suspected to be redundant just like the same measurement in both feet and meters or the repetitiveness of images presented as pixels. Then it can be transformed into a reduced set of features.

Deep learning classifiers figure out when they go through training with the added bonus that not only they are differentiating features but they are also determining what the actual features are in the first place.

In any image if one go deeper and deeper layers of neural network those layers combine the information they get from the earlier layer and from the higher level image. Perhaps things like objects overall shape ,well feature extraction does required good knowledge of Machine learning techniques unlike other classifier construction methods done entirely with narrow network feature extraction is super quick requires very little data and one don't need good hardware to get good results.

In Multi-class SVM classifier disposal and these techniques run and learn really release fast but they require an initial list of features that you have to supply.

A convolutionary neural network which is trained to classify thousands of objects so we are going to import Alexa net and with this line of code we pull in all its layers. Generally the deepest layer of network right before the layer that actually starts classifying objects is a good place to start for Feature Extraction.

2)K Means Clustering for description: After extracting descriptors from the training images unsupervised learning algorithms such as k-means are used in grouping them into N clusters of visual words. The metric used to categorize the descriptor into its cluster centroid is the Euclidean Distance. For this purpose each image extracted descriptor is assigned to its closest cluster centroid. In order to generate the histogram of counts the cluster centroid number of occupants is incremented each time a descriptor mapped into it. At the end of this process each image is categorized by a histogram vector of length N to ensure the invariance of this method with respect to the number of descriptors used, it is essential to normalize each histogram by its L2-Norm. To group the descriptors and construct the N- visual words we use K- means clustering. This approach is selected over expectation maximization as many experimental methods have conformed the computational efficiency of k-means with respect to EM.

3) Recognition based on Bag of Features model: Research in computer vision field have led to many learning approaches to leverage the BoF model for the purpose of Image recognition for multiple label classification problems the evaluation metric which is used is the confusion matrix [3].

A Confusion matrix is defined as a particular table making it possible to visualize the accuracy of a supervised learning algorithm. Matrix columns symbolize the instance in a predicted class whereas rows represent the instances in an actual class or vice-versa. The appellation is due to the fact that it makes simple to see if the system confuses two categories (i.e mislabeling one as other).

In this work we investigate many supervised learning algorithms such as: K-nearest neighbour, boosted regression trees and Instance based learning to classify an image. Each image in dataset is encoded by its Bag of Features histogram vector as shown in the Fig 2.

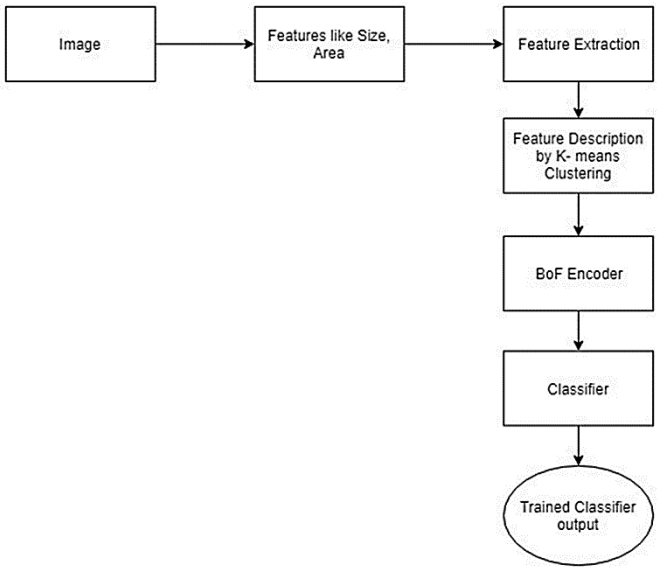

Fig 2 : Step by step in Image classification

4) Visualization: Visualization has always been used to study relationship between anatomical structure and function where there are several types of techniques in it hence volume visualization is more beneficial in comparison than the 2D technique.

IV. EXPERIMENT AND EVALUATION

In the following we provide a summary of different experiments that we use to evaluate the performance of our image classification machine learning framework. Our results are reported on CIFAR10 image dataset to which we have added some new images of existing categories. We are interested in sofa category recognition in Fig 3.

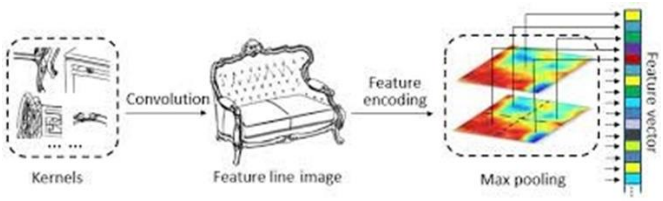

Fig 3 : Detection of Image

A. Evaluation Process

1)Encoding for Feature Extraction: We use Bag of Features to encode each image of dataset into a vector feature which represents the histogram of visual word occurrences contained in the following Fig 4(b).

2)Classifying Process: The encoded training images are fed into the classifier training process to generate a predictive model. In this section, we are interested in measuring the classifier average accuracy and its confusion matrix [3].

Fig 4 : (a)

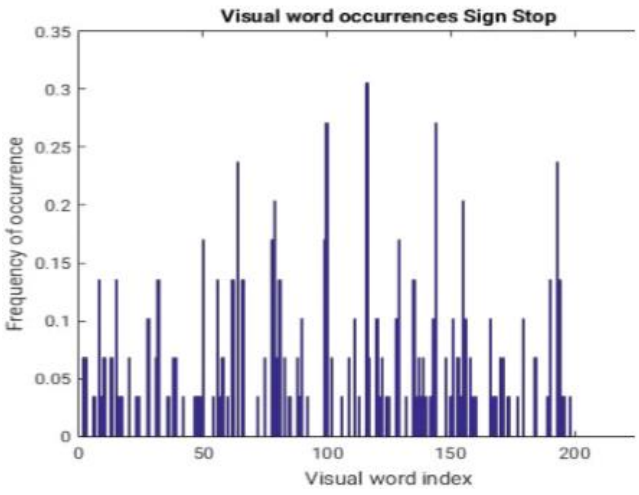

Fig 4: (b)

For tests we use this technique and the Linear classifier technique. We choose 30% among images for training and remaining for validation. The obtained confusion matrices are shown below.

	Predicted	
Known	Sofa	Dog
Sofa	0.94	0.06
Dog	0.08	0.92

Fig 5 : Confusion matrix for two image category classification

As shown in the Fig 5 we noticed that the average accuracy of the classifier is influenced by the number of categories in training dataset. This metric is lower when the number of sets increases.

3)Evaluating Image Category Classifier: The linear classifier is applied on sofa, dog, laptop, tiger categories. We use k-means of clustering algorithm instead of other technique. It is reported that the average accuracy is 0.77.

	Predicted		
Known	Sofa	Dog	Laptop
Sofa	0.92	0.02	0.04
Dog	0.06	0.93	0.01
Laptop	0.05	0.01	0.94

Fig 6: Confusion matrix for three Image category classification

	Predicted			
Known	Sofa	Dog	Laptop	Tiger
Sofa	0.92	0.02	0.04	0
Dog	0.01	0.93	0.01	0.05
Laptop	0.04	0.01	0.94	0.01
Tiger	0.02	0.12	0.01	0.85

Fig 7: Confusion matrix for four Image category classification

We noticed that in our approach is better to use a local feature extractor than global feature extractor. This result is expected as global feature extraction technique is better with scene categorization and not for object classification.

	Predicted		
Known	Sofa	Dog	Laptop
Sofa	0.84	0.04	0.12
Dog	0.1	0.67	0.23
Laptop	0.02	0.22	0.76

Fig 8: Confusion matrix for feature Extractor

4)Training Learner Evaluation: We then increase the number of categories into AFE technique and evaluate models on varying the classifier algorithm: SVM, KNN and Assemble classified categories. We then generate the histogram of the average accuracy based on the training classifier.

Measurements show that the image classification process performs better when we use a machine learning algorithm. It is reported that the image yields average accuracy which reaches 90 percent. The KNN technique offers an average accuracy around 65 percent among the assemble classifier trainers. The bagged trees achieve the best accuracy [3].

V. CONCLUSION

In this paper we related different techniques and algorithms used in our Machine Learning Framework for Image Classification. We presented Machine Learning state of the art applied to computer vision. We introduced Bag of Features paradigm and highlighted the step by step process involved in Image Classification.

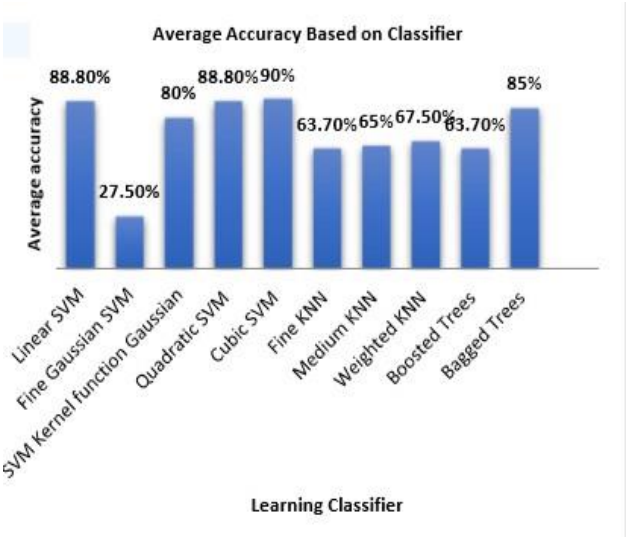

Fig 9 : Average accuracy based on the learning classifier

VI. REFERENCES

[1] A.Chergui, A.Bekkhoucha, W.Sabbar, "Video scene segmentation using the shot transition detection by local characterization of the points of interest". SETIT'**2012**,Tunisie DOI:**10.1109**/SETIT.**2012.6481949, 2012**.

[2] M. S.Nixon, A. S.Aguado, "Feature Extraction and Image Processing". Elsevier Ltd, ISBN: **978-0-12372-538-7, 2008**.

[3] C.Domeniconi, D.Gunopulos, J.Peng, "Large margin nearest neighbor classifiers". IEEE Transactions on Neural Networks, **16(4):899–909**, **2005**.

[4] D. G.Lowe, "Distinctive image features from scale invariant keypoints". IJCV, **60(2):91–110**, **2004**.

[5] Tianmei Guo, Jiwen Dong, Henjian Li, Yunxing Gao, "Simple Convolution Neural Network on Image Classification".

[6] J. H.Friedman, "Greedy Function Approximation: A Gradient Boosting Machine", **1999**.

[7] L.Rokach, O.Maimon, "Data mining with decision trees: theory and applications". World Scientific Pub Co Inc. ISBN 978-9812771711, 2008.

[8] Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2013, 35(8): 1798-1828.

[9] Graham, B., Fractional Max-Pooling. 2014. Ar Xiv e-prints, December 2014a.

[10] Deeply supervised nets. In Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics, AISTATS 2015, San Diego, California, USA, May9-12, 2015.

[11] Hecht-Nielsen R. Theory of the backpropagation neural network[M]// Neural networks for perception (Vol. 2). Harcourt Brace & Co. 1992:593-605 vol.1.

[12] Y. Boureau, J. Ponce, and Y. Le Cun, "A theoretical analysis of feature pooling in visual recognition," in ICML, 2010, pp. 111–118.

Cite this article as :

G. Keerthi Devipriya, E. Chandana, B. Prathyusha, T. Seshu Chakravarthy, "Image Classification using CNN and Machine Learning", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 5 Issue 2, pp. 575-580, March-April 2019. Available at doi : https://doi.org/10.32628/CSEIT195298
Journal URL : http://ijsrcseit.com/CSEIT195298