International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-2, July 2019

Categorized Image Classification uing CNN Features with ECOC Framework

Shameem Fatima, M. Seshashayee

Abstract: ***Image Classification technique is used to classify images into categories. In this study, an application is presented to examine category based image classification by combining Support Vector Machine with error correcting output codes (ECOC) framework. The ResNet50 used as Network architecture, our image dataset include caltech101 images from 9 categories (classes) which builds our classification task a multiclass problem. ECOC is a commonly used framework to model multiclass classification problem. We present one-verses-all coding design of ECOC and apply to SVM classifier. A pre-trained CNN (convolution neural network) is used for extracting image feature and as a classifier Multiclass Support Vector Machine is used. The extracted features are then passed for classification via ECOC approach. The final classification result predicts the class labels. The application is implemented in Matlab using pre-trained CNN. The prediction accuracy of each category is evaluated and presented. The experimental result shows an accuracy of 97.6%. Further experiments are carried out on different dataset which showed that best accuracy is achieved using CNN with ECOC for multiclass problem.***

Index Terms: ***Convolution Neural Network, ECOC, Image classification, SVM, ResNet50.***

I. INTRODUCTION

The Humans are gifted with a great capability to effortlessly recognize an object in image or video by their human vision system. They can distinguish between different object categories, but for a computer application however there exist an issue in detection and classification of objects which depends on various factors such as lightning condition, weather, shape, color, etc. Image Classification is considered as a computer vision branch where images are categorized [1].Image can be classified into one of the two broad categories they are Supervised and Unsupervised techniques. Image classification have been extensively studied for various purposes in literature such as in [2] performs the classification of vehicle make and model using SVM which uses a combination of SIFT and bag-of-words. X. Li and X. Guo [3] developed a forward vehicle detection system that uses HOG descriptor and SVM. In recent research literature, convolutional neural network (CNN) has been used as one of the standard model for image classification. CNN, which was originally introduced by Le cun *et al* (1998) [4], has a specially designed architecture for image classification. We can create and train a new network or use pre trained network. The main purpose of classifying an image is to identify the labels for the given images. When image are categorized into more than two classes it is considered as multiclass problem. A multiclass problem with Error correcting output codes framework [5] with a classification method is studied in this paper. In literature, classification methods such as SVM, Discriminant analysis, Nearest Neighbor are used for Supervised Learning Technique [6]. In this study, image categorized classification is implemented by combining Support Vector Machine (SVM) with ECOC framework. The Experiments are carried out on image dataset using ResNet50 Network. Extracting Image features are done using CNN. The ability of CNN to handle huge data and can perform extracting features automatically, it have been used as a feature extractor with multiclass SVM. In this work, one-verses-all coding design approach is used to categories images. Overall the SVM is combined with ECOC framework to achieve the category classification. The experiment evaluates the performance metrics of the classifier used.

II. RELATED WORK

Sharma *et al*(2018)[7]presents a study for analyzing the performance and predicting accuracy of three different convolutional neural networks with two datasets(CIFAR10 and CIFAR100)on ten different class of objects The result analysis showed that GoogLeNet and ResNet50 performed better in precision compared to AlexNet. It also revealed that transfer learning performed better using trained network along with greater accuracy. It was stated that neural networks is considered as a new and best emerging techniques for solving categorization problems.

Kibria and Hasan(2017)[8]presents an study to investigate selected four methods such as Bag of Words (BoW), Histogram of Oriented Gradients (HOG) extractor with SVM classifier, Convolution Neural Network (CNN) and pre-trained CNN with SVM for image classification to find the accuracy to identify knives from image dataset. The methodology involves the following steps, the first step deals with extracting features from images. In the second step, a classifier is trained with these features to recognize and then classify them correctly and then, the classifier predict image category. The classifier is then tested with a test set to view its accuracy. All these algorithms have been tested in two ways. In the first case the dataset is divided into two equal partitions as training and testing set.

Revised Manuscript Received on 30 July 2019.
***** Correspondence Author
Shameem Fatima*, Department of CS, GITAM (Deemed to be University), Visakhapatnam, India.
Dr. M. Seshashayee, Department of CS, GITAM (Deemed to be University), Visakhapatnam, India.

The second case, the sets are reversed and the first is used for testing while the second for training. The result showed that deep learning CNN (pre-trained, untrained) based methods performed best in terms of accuracy. It was also stated that the use of pre-trained Alexnet along with SVM performed best among all.

Tien Vo *et al* (2017) [6] proposes an model for classifying image into two classes YES or NO.YES refer to the advertisement being displayed clearly, "Clear" means that the user can view the content of the advertisement. No refers to not displayed or not clear. The methodology involves using Convolutional Neural Network with two parameters (number of layers, number of filters in Conv layer). The proposed model involves four stages that are input, capturing, classification, and output. The input is URL of a website which is saved as screenshot. These images are then resized (32x32 pixels) for processing by CNN. The output is one of two conclusions YES or NO. The results showed an accuracy of 85.74% which reflects feasibility of the proposed model.

Farren (2017) [9] presents retail product classification approach on grocery shelves and uses a novel dataset. It uses a greedy algorithm to improve the network performance by changing the architecture. The parameters include filter size, no of filters of each convolution layer, the no of convolution layers, size and number of fully connected layer. The architecture use a batch normalize layer after first two convolutional max pool layers. The activation function was Relu with a dropout rate of 50%on every Relu layer to avoid over fitting. The network was trained using softmax loss function. Each iteration uses back propagation to calculate the gradient with respect to the loss at each layer, It then uses Adam algorithm to adapt the learning rate and change in layer weight overtime. The experimental results show that considerable increase in the model accuracy. When the layer was dropped the filter size did not make much of the relative difference in improving the model performance.

Sunderhauf *et al* (2014)[10] presents a supervised plant classification system that uses features from CNN. The different parts of the plants such as branch, stem, leaf, leaf scan, fruit and flower has been referred to as content categories. The methodology involves using features that are pre trained. Feature extraction is performed by examining one of the initial two fully connected layers (Layer 17, Layer 19). Each category uses separate classifier to train the dataset. The classification result is obtained by averaging output of all image features that gives a probability distribution for each image. The experimental result shows that Layer19 performed better for plant classification.

III. METHODOLOGY

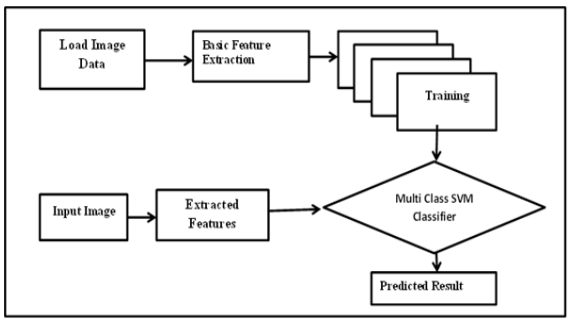


Image Database

Our Image Database contains 1,533 images which involves 9 categories: Camera, Cellphone, Cup, Headphone, Human, Lamp, Motorbikes, Scissors and Wheelchair are taken from the image archive known the Caltech101 [11]. The Table 1 shows the total number of image in each category along with its Index. The Caltech 101 dataset which consist of 9,146 images and it has been divided into 101 distinct object categories, whose image size is about 300x200 pixels [12].

Table 1: No of Image for each category and its Index

Category	Camera	Cellphone	Cup	Headphone	Human	lamp	Motorbikes	Scissors	Wheelchair
No of Image	368	798	50	59	57	42	61	39	59
Index	1	2	3	4	5	6	7	8	9

Feature Extraction using CNN

The first and most important step deals with extracting features from all images. This is performed by creating a data store of image object which holds images along with its category labels. The labels have been assigned by taking the names of the folders that contain the image files. The feature extraction technique used is CNN, with its ease to train a database, we used a pre trained CNN as a feature extractor. The initial layers of the network are responsible for capture basic image features (edges and blobs). Only few of the starting layers with in network (CNN) deals with feature extraction. The input image receives a response from each layer produced by CNN. To capture advanced features the early features are combined and processed by deeper layers of network. In ResNet50, the training features are extracted using layer "fc1000" [13]. The table 2 shows the first section of ResNet 50 Network.

Training Image for category classification using CNN and SVM

There are various popular pre trained networks such as AlexNet, GoogLeNet, ResNet50, VGG-16 etc. These networks are trained on ImageNet dataset [14], [15].The present work uses the ResNet50 model. The input image dimension (224-by-224-by-3) is defined by the first layer. The next middle layers are formed by chain of convolution layer and three fully connected layers, which are combined with rectified linear units (ReLU) and max-pooling layers [16].The classification layer is the last layer consists of 1000 classes. The machine learning is divided into supervised and unsupervised learning [1], classification is specified to be supervised learning, where it learns from the model to put the right category, before categorizing an unidentified image or observation. A pre trained CNN is used for the purpose of extracting features, the starting network layer capture the primary image feature (edges, blobs). A convolutional layer contains network weight. A pre trained CNN extract image features. The preparation of the training and test data set is performed by dividing the dataset into 30 percent for training and 70 percent for validation data. The processing of training and testing sets is done by CNN model. A SVM classifier is trained by the featured that are being extracted using CNN present in the training data set.

During the testing process, using same model the test feature are extracted and passed to classifier to perform testing [13].

Support vector machine for classification

Support Vector Machine is one of the Machine learning Algorithm used for classification. It is based on supervised learning which classifies points to one of the two disjoint planes [17]. SVM is used in many applications such as face recognition, text categorization, pattern recognition etc. [18]. SVM was mainly used in binary classification where there are only two target categories or values. Now a days, we mostly have huge amount of data which we want to classify, with the need to classify more than two target classes, Multiclass classification is needed [19]. Multiclass SVM is used when there are more than two target categories. A framework was proposed by Dietterich et al (1995) [20] in which the error correcting output code was used for converting multiclass problem into several binary problems. The combination of ECOC with Svm enhances the fault tolerance of classification model when solving multiclass classification problem [5].

Error Correcting Output Codes (ECOC) Framework:

Error correcting output codes model is considered as one of the powerful framework when dealing with multiclass classification problem [21]. An error correcting output codes (ECOC) model is an approach that does reduction of the classification problem with more than two classes (categories) to a collection of binary classifiers. The ECOC approach consists of two steps: Coding scheme and Decoding scheme [5].

Coding Scheme: This approach needs a coding design for ECOC classification. The coding design describes ways to reduced multiclass problems to a collection of binary class problems. The classes are defined on which the binary Learners are trained . Coding Design is in the form of a matrix where each row is related to an individual class and each column relates to a binary learner [22]. Two widely used coding design approaches are “one-versus-all” classifier (usually referred to as OVA classification) and “one-versus-one” (usually referred to as OVO classification) classifier. In one –verses-all for each binary learner there is a class which is positive and the rest are negative. It divides a K class problem into collection of K binary class problem. K refers number of individual classes. In one-verses-one for each binary learner there exists a class which is positive other is negative and the rest are ignored. It divides a K class problem into K $(K-1)/2$ binary class problems [23].

Decoding Scheme: It describes ways to combine result obtained from binary learners.

Error Correcting Output Code (ECOC) SVM: It has been stated that when solving multiclass classification problem ECOC Svm enhances the ability of fault tolerance which makes it suitable for image classification [5]. The steps is as follows

Step 1: Load Images
Step 2: Feature Extraction
Step 3: Apply SVM binary classifier via ECOC framework, to perform Multiclass SVM with ECOC Framework use fitcecoc and predict function available in Matlab.
Step4:Classification Results with ECOC framework and Performance Analysis.

Function to Train and Predict

Multiclass SVM with ECOC approach uses fitcecoc and predict function to train multiclass SVM and predict label using Matlab[22].

fitcecoc and predict function: We make use of”fitcecoc” function for training a multiclass model for SVM (composed of binary SVM learners). Using this function we train the SVM classifier on the extracted features. The function returns a trained ECOC model. The present work uses K binary SVM Algorithm with one-verses-all coding approach, K refers distinct class. To predict label predict function is used which returns a predicted class labels for the predictor data in the table or Matrix based on error-correcting output code (ECOC) model.

IV. IMAGE CLASSIFICATION STEPS

Step 1: Input image from one of the category folder.
Step2: Load Database: Load images use a function Image DataStore which operates on image location to hold images and labels that are associated with each image category. An image data store allow us to store large image data, it also divide the data into 70% training and 30% test data. As each folder has different number of images per category. Using the function Count Each label determines the smallest number of images in category in order to get exactly the same number of images in each category. Load pre-trained ResNet50 Network using functionresnet50.
Step 3: Image Preprocessing: The CNN model processes both trainingSet and testSet, where these datasets are divided into training and validation data. Thirty percent (30%) of images are randomly chosen for training and the remaining seventy percent (70%) as validating data. Image Preprocessing for CNN depending on the network used, it is performed by, Resize the image according to the network (224-by-224) and converting grayscale images into RGB images using the function augmented Image Datastore.
Step 4: Extract Training Features Using CNN using function Activation with feature layer fc1000.
Step 5: Training of a multiclass SVM Classifier is done using CNN Features.
Step 6: Evaluation of the Classifier is performed by extracting image features from test dataset then these features are passed to the classifier for measuring the accuracy of the trained classifier.
Step 7: Predicted Category:
The function used in the above steps is available in [13].

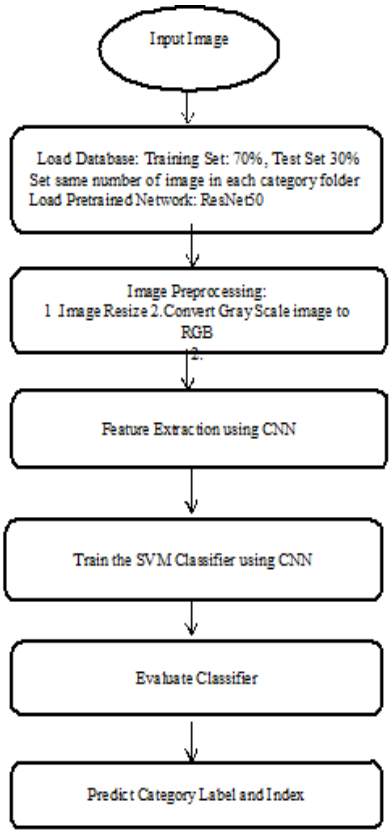


Fig 1: Image Classification Steps

V. PERFORMANCE MEASURES

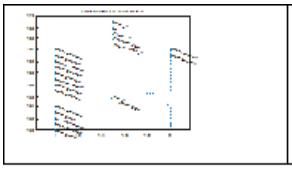
One of the important aspects for an application is to measure the performance. In literature there exist various performance metrics for image classification. Classification technique uses different metrics depending on each domain. The three metrics used for classification are Accuracy, Precision and Recall [19]. Table 2 presents the measure with its formulae (TP, TN, FP, FN refers to true positive, true negative, false positive, false negative respectively).

Accuracy: It is the most widely used measure. It represents the overall classification result [19]**.** It is displayed at the bottom right most cell of the confusion matrix table.

Precision: It is a fraction of positively true occurrences to overall positive occurrences of categories of class. It is a measure that represents the prediction for an individual class or category [19]. It is displayed in the last column at the right side of the confusion matrix table.

Recall: It is a percentage of positively true occurrences to the total occurrences of true positives and false negatives. It is a measure that represents the prediction only for individual class or category [19].It is displayed in the last row at the bottom of the confusion matrix table.

Table 2: First section of ResNet 50 Network and Performance metrics formulae

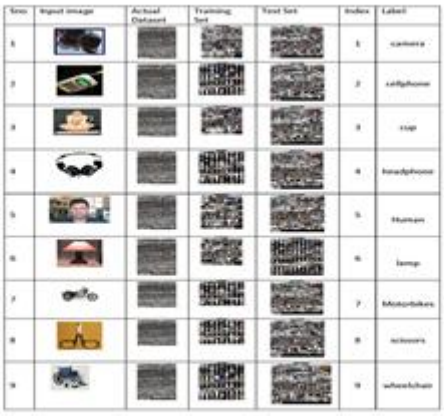


Measure	Formula
Accuracy	(TP +TN)/ (TP +TN - FP + FN)
Precision	TP/ (TP + FP)
Recall	TP/ (TP +FN)

VI. EXPERIMENTAL RESULTS AND ANALYSIS

To demonstrate the multiclass image classification, experiment was conducted with nine image categories taken from caltech101 dataset [11].each category is assigned an index ranging from 1 to 9.The application was implemented in matlab2018[24].The experiment take input image data, extract feature using CNN, image preprocessing is done as per network requirement and then a multiclass SVM is trained on training feature, finally the classifier is evaluate to predict the categories label along with its index. The three performance measures evaluated are accuracy, precision and recall. The confusion matrix in Table4 shows performance measure evaluated for each category. The Experimental Results of the image classification with predicted Image label and Index are presented.

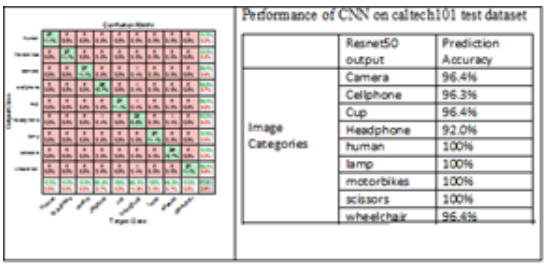
Table 3: Results of Experiment1on Caltech 101 dataset



Evaluation of the Model

As we have multiple saved models, which can be used to evaluate the loss/highest accuracy. The confusion matrix is used to show the correct vs. incorrect class labels. The first nine rows in the Confusion Matrix table related to the predicted category class (Output Class) and the first nine columns relate to the true class (Target Class). The slanting (diagonal) cells relate to the number of categories that have been correctly classified. The off-diagonal cells relates to incorrectly classified categories. The result shows that they are best suitable for image classification.

Table 4: Confusion Matrix and Performance of CNN for caltech101 dataset



Performance of CNN on caltech101 test dataset

	Resnet50 output	Prediction Accuracy
Image Categories	Camera	96.4%
	Cellphone	96.3%
	Cup	96.4%
	Headphone	92.0%
	human	100%
	lamp	100%
	motorbikes	100%
	scissors	100%
	wheelchair	96.4%

Table 5: Confusion Matrix forExperiment2, Experiment3, Experiment4.

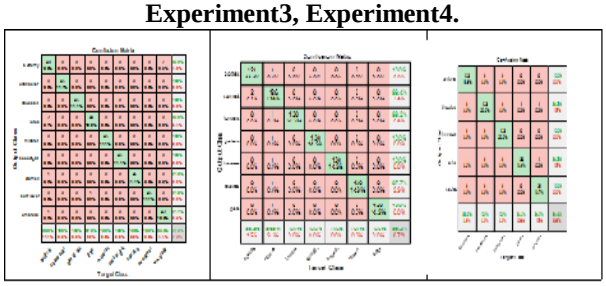


Table 6: Accuracy comparison for different models and dataset.

Sno	Ref	Network	Model	Dataset	Accuracy
1	[5]	-	ECOC SVM	Landsat TM	88.85%
2	[27]	AlexNet	SVM	Caltech	89.7%
3	Experiment 1	Resnet50	Multiclass SVM	Caltech101	97.6%
4	Experiment 2	Resnet50	Multiclass SVM	Caltech101	98.0%
5	Experiment 3	Resnet50	Multiclass SVM	[25]http://www.robots.ox.ac.uk/~vgg/data	99.3%
6	Experiment 4	Resnet50	Multiclass SVM	Caltech256	99.4%

To evaluate the model Accuracy the application was tested on ResNet50 network architecture with Multiclass SVM model on multiple datasets. The Table 5 shows the confusion Matrix of each experiment and Table 3 shows the experimental results. Each experiment was tested with increase in number of image in database. Experiment1 uses nine category of classes discuss in detail in the application.Experiment2 is tested by taking different categories from the same Caltech101 dataset with nine categories namely: 'butterfly', 'hawksbill', 'lotus','soccer_ball', 'sunflower', 'umbrella' ,'minaret', 'chandelier', 'starfish'. The result showed an accuracy of 98.0%. Experiment3 was conducted on[25] with seven categories namely: 'bottles' ,'camels', 'flowers', 'guitars',' houses', 'leaves', 'pets' .The results shows an accuracy of 99.3%. Experiment4was conducted on Caltech 256[26]with five categories namely:' airplanes', 'binoculars', 'lighthouse', 'tshirts', 'watches'. The result shows an accuracy of 99.4%.

VII. DISCUSSION

The image categorization problem is a multiclass classification problem (consisting of nine category image folders), to which SVM cannot be applied directly. They are various frameworks that can be applied to SVM for multiclass classification. The present work uses SVM with ECOC framework. It uses one-verses-all coding design method to train a classifier. For K class this result is K binary classifier. It constructs a set of K binary classifier each trained to separate class from the rest and combines them by performing the multiclass classification. The results of a similar study was also conducted by [27] constructed on the morphological image processing and deep learning using binary SVM model with pre trained CNN. The result showed less accuracy when compared with the present research using multiclass SVM with ECOC approach where one-verses all coding design is used. Another study conducted by [5] for image classification gave less accuracy when compared to remaining all the experiments in the study. The outstanding classification performance of the present research work was mainly achieved by utilizing pre-trained CNN with ResNet50 network architecture. Therefore, classification task can produce good performance when using CNN as feature extractor and SVM classifier with ECOC framework for multiclass classification problem.

VIII. CONCLUSION

Categorization of image into meaningful categories is one of the promising applications of image classification in computer vision. In this study a supervised classification is performed using a SVM algorithm with ECOC framework with a small database containing 1533 images of nine categories taken from caltech101 dataset with ResNet50 network architecture. Image features are extractor using pre-trained CNN, multiclass SVM is used as a classifier. The performance is evaluated by plotting the confusion matrix. The prediction accuracy of each category is computed. It was observed that four categories human, lamp, motorbikes and scissors are classified with 100% prediction accuracy. The CNN model coupled with a multiclass SVM via ECOC framework resulted in an overall accuracy of 97.6%.The application was also tested on different datasets and compared with the result in literature, the results show that it outperformed in term of performance. The research work can be further studied advanced features such as color, shape that can categories the images.

REFERENCES

1. M. S.Hasan, (2017, December). An application of pre-trained CNN for image classification. In *2017 20th International Conference of Computer and Information Technology (ICCIT)*(pp. 1-6). IEEE.
2. M. A. Manzoor and Y. Morgan, "Vehicle Make and Model classification system using bag of SIFT features," in Computing and Communication Workshop and Conference (CCWC), 2017 IEEE 7th Annual, pp. 1–5, IEEE, 2017.
3. X. Li and X. Guo, "A HOG feature and SVM based method for forward vehicle detection with single camera," in Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2013 5th International Conference on, vol. 1, pp. 263–266, IEEE,2013.
4. Y. LeCun, B. Boser, J.S. Denker, Henderson, R.E.Howard, W.Hubbard, L.D.Jackel,: Backpropagation applied to handwritten zip code recognition. Neural Comput. 1(4), 541–551 (1989)
5. Z. Yan., & Y. Yang, (2014). Application of ECOC SVMs in Remote Sensing Image Classification. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *40*(2), 191
6. A. T.Vo, H. S.Tran, & T. H.Le, (2017, October). Advertisement image classification using convolutional neural network. In *2017 9th International Conference on Knowledge and Systems Engineering (KSE)* (pp. 197-202). IEEE.
7. N.Sharma, V.Jain, & A.Mishra, (2018). An Analysis Of Convolutional Neural Networks For Image Classification. Procedia Computer Science, 132, 377-384.
8. S. B. Kibria, & M. S.Hasan, (2017, December). An analysis of Feature extraction and Classification Algorithms for Dangerous Object Detection. In 2017 2nd International Conference on Electrical & Electronic Engineering (ICEEE)(pp. 1-4). IEEE.

9. D. Farren, Classifying food items by image using Convolutional Neural Networks.http://cs231n.stanford.edu/reports/2017/pdfs/607.pdf
10. N.Sünderhauf, C.McCool, B.Upcroft, & T.Perez, (2014). Fine-Grained Plant Classification Using Convolutional Neural Networks for Feature Extraction. In *CLEF (Working Notes)*(pp. 756-762).
11. www.vision.caltech.edu/Image_Datasets/Caltech101 Caltech 101
12. https://en.wikipedia.org/wiki/Caltech_101
13. https://www.mathworks.com/examples/computer-vision/mw/vision-ex77068225-image-category-classification-using-deep-learning
14. Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009.
15. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).
16. Krizhevsky, Alex, IlyaSutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.
17. Y.Ahuja, & S. K.Yadav, (2012). Multiclass classification and support vector machine. *Global Journal of Computer Science and Technology Interdisciplinary*, *12*(11), 14-20.
18. Donahue, Jeff, et al. "Decaf: A deep convolutional activation feature for generic visual recognition." arXiv preprint arXiv:1310.1531 (2013).
19. C. Demirkesen, & H. Cherifi, (2008, October). A comparison of multiclass SVM methods for real world natural scenes. In *International Conference on Advanced Concepts for Intelligent Vision Systems* (pp. 752-763). Springer, Berlin, Heidelberg.
20. T. G. Dietterich, G.Bakiri, 1995. Solving Multiclass Learning Problems via Error-Correcting Output Codes. Journal of Artificial Intelligence Research, 2, pp. 263-286.
21. M. ali Bagheri, G. A. Montazer, & S. Escalera, (2012, May). Error correcting output codes for multiclass classification: application to two image vision problems. In *The 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP 2012)* (pp. 508-513). IEEE.
22. https://in.mathworks.com/help/stats/fitcecoc.html accessed on 6 may 2019.
23. C. W.Hsu, & C. J. Lin, (2002). A comparison of methods for multiclass support vector machines. *IEEE transactions on Neural Networks*, *13*(2), 415-425.
24. https://in.mathworks.com/products/matlab/
25. http://www.robots.ox.ac.uk/~vgg/data
26. http://www.vision.caltech.edu/Image_Datasets/Caltech256/
27. Y. Shima, (2016, December). Extraction of number plate images based on image category classification using deep learning. In 2016 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS) (pp. 19-26). IEEE

AUTHORS PROFILE

ShameemFatima is Research Scholar in the Department *of* CS, GITAM (Deemed to be University), Visakhapatnam, India. She has received her M.C.A in Computer Science from Osmania University (India) in 1998. She has 9 years of teaching experience and 5 years of research experience. She published 9 papers in conferences and journals. Her research interests include Image Processing, convolution neural network.

Dr.M.Seshashayee is working as Assistant Professor, Department of Computer Science, GIS, GITAM (Deemed to be University), Visakhapatnam,India. She holds a doctorate degree in Computer Science and Engineering. She has 13 years of teaching experience. Her area of research is Image Segmentation using Data Mining Techniques and Convolution Neural Networks. She has published 12 research papers in reputed International Journals. She attended 20 conferences and 6 workshops. She is active both in academic and administrative areas. She is member of CSI and IAENG and also reviewer of IJICSE.