Deep Fuzzy Multi-Object Categorization in Scene

S.Kumaravel, S.Veni

Abstract: ***Object Categorization is the process of, identifying and labelling the various distinct Classes (Categories), in the given input image. The Deep Fuzzy Multi-Object Categorization (DFMOC) model, combines the learning capability of Convolution Neural Networks (CNN) and the uncertainty-managing ability of Fuzzy system, for carrying out the categorization task. This work starts with Background Elimination process for ensuring the image clarity, followed by Fuzzification and Fuzzy Entropy computation. Simple fuzzy sets are to be framed, by employing Fuzzy C-Means (FCM) algorithm, for fuzzification of the input image. Thresholding Block is incorporated, for determining the clusters . The Fuzzy Entropy Computation (FEC) is done, to minimize the Fuzziness rate of the acquired input and consequently, the layers of CNN are trained in accordance with that. Caltech-101 Dataset is been utilized for analysis. Average Precision Rate of Categorization (APRC), along with other metrics namely Time taken and Error Rate, shows that DFMOC model performs better than other models.***

Keywords: Deep Fuzzy Model, Fuzzy Entropy, Object Categorization, Thresholding.

I. INTRODUCTION

For understanding a particular image in a better way, the objects in the images are required to be isolated and their bonding analyzed. This process is technically termed as Image Segmentation [3]. **Figure 1** describes the typical operations involved in digital image processing on Object Recognition. As given in the figure, the input images are acquired from knowledge data base, and the quality of the images is needed to be enhanced by denoising. Consequently, the images are restored for further processing and compression is made when there is a need. Image segmentation is the process of detailing the objects present in a particular image and its features are being extracted by some derivatives. By some training and testing mechanism, the objects in an obtained image are recognized, classified and categorized.

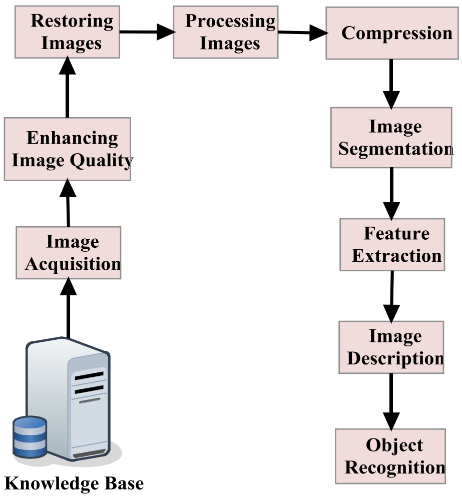

Figure 1: Operations in Digital Image Processing

Deep Learning paradigm differ from the conventional techniques, in that, the feature learning process are carried out directly from the solid pixels of the input image, without the use of basic approaches like SURF and HOG. The features like corners, lines and Boundaries are getting extracted from the low-level layers and strong features like object parts are extracted from the higher layers.

In particular, since there are some uncertainties present in many phases of image processing, fuzzy logic model is the most required in computer vision. The acquired image may have some additive noise in low level processing, ambiguity in the consideration on algorithms and some imprecision at the high levels [6]. The neural network (NN) structure framed in this paper is based on Adaptive Neuro Fuzzy Inference System (ANFIS), and is given in **Figure 2**. Layer 1 is the input layer that gets the input image and Layer 2 presents firing of rules through the Product operation. Layer 3 is for Normalization of rules, from the previous layer. Layer 4 nodes represent the consequent part of fuzzy rules and layer 5 performs Defuzzification, by summing up the various outputs. In this Deep Fuzzy Model, fuzzy sets are framed using FCM algorithm for fuzzification, and thresholding has been done for clustering. Later, training the NN is made on the basis of Entropy Evaluation. The trained NN is based on fuzzy logic and result in an effective object categorization.

Revised Manuscript Received on November 01, 2020.

S.Kumaravel, Associate Prof., in Computer Science, SRMV College of Arts and Science , Coimbatore-20. Corresponding Author, Email Id : ara28vel@gmail.com.

 # S.Veni, Prof., & HOD, Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore-21. Email Id : venikarthik04@gmail.com

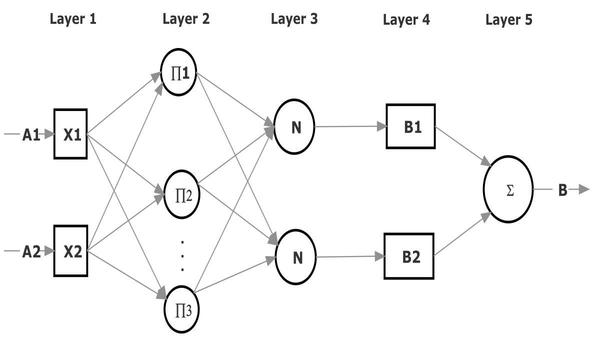

Figure 2: Layers in Deep-fuzzy System

The organization of this paper follows: Section 2 explains the related works. Section 3 outlines the detailed mechanism of the proposed Deep Fuzzy Model.. Finally, Section 4 presents the Conclusion .

II. RELATED WORK

In this section, a comprehensive study is given about the review papers, that lead to the current work.

In [1], a model has been developed for denoising the captured images using Fuzzy-logic which can determine the expert knowledge, using rule-based linguistic labels in a direct way. The work process of Neural Network (NN) has also been combined with this, for not only considerably reducing the cost and time, but also enhancing the system performance. The model involves reducing unwanted noise by preserving image details and sharpness of the edge, along with improving the image contrast factor.

A new Tri-State Median (NTSM) [2] filter is used in Neuro-fuzzy network for getting the enhanced output image. The architecture includes hybrid filtering technique for having noise- free images for processing. It is stated in the paper that the filter is more adaptive specifically against impulse noise than the traditional ones.

There are many applications of Artificial Intelligence such as medical system, economic system, traffic control, student modelling, forecasting and social sciences [3,4]. Most of them depend on the Neuro-fuzzy models for attaining effective results.

While discussing applications, ANFIS system (Adaptive Neuro-fuzzy Inference System) has been proposed for MR Brain Image Classification [5]. In the method, the image features are extracted and then fed to the classifiers as a complicated architecture for multiple object categorizations of brain structures from MR outcome by adaption of Neuro-fuzzy logic named as ANFIS, operating through the precise knowledge depiction of fuzzy sets with the training ability of artificial feed forward neural network. Using Neuro-fuzzy logic provides accurate results with a reduced time convergence rate.

. The efficiencies of Convolutional Neural Networks (CNN) [6] and Support Vector Machine (SVM) have been examined on object classification for LANDSAT 8 images. It has been stated that, a machine learning process is required for processing huge data, since the acquired LANDSAT images are abundant and complicated to compute.

In another dimension, image classification has been done with the employment of softmax activation for attaining reduced cross entropy loss [7]. L2-SVM has been developed for the replacement of softmax layer on the archetype of linear SVM. Back propagation algorithm is used to learn the lower level features of the input image that explicitly reduces the primary problems over SVM.

Object recognition has been achieved using deep neural networks instead of Convolutional Neural Networks [8]. Further, it produces most precise results with less time consumptions. The authors have analysed the supervised learning method by the execution of decision trees.

Xiaogang Wang [9] discusses the imperative role of deep learning in operations such as object detection, classification, segmentation and recognition. The overall conversation on the paper is prepared on object recognition framework on ImageNet, image classification, video processing, biometrics, etc. The paper has given a valuable resource on approaches of deep learning process.

The speed of image classification has been improved by boosting the fuzzy classifiers [10]. The authors have framed a new technique for object classification in accordance with the local-image-feature-based simple fuzzy classifier. It helps to differentiate among classes. The most derivative local image features are recognized by boosting the Meta learning model, given in the paper.

In [11], image segmentation and appropriate edge detection have been done using Adaptive Neuro-fuzzy System in an automated way. The architecture comprises multilayer perception that involves segmentation process based on the adaptive thresholding. Fuzzy entropy has also been utilized for determining the errors in the segmentation with potential edge pixels. Moreover, fuzzy entropy relaxation process has also been monitored for effectively finding the edge pixels.

Due to the effectiveness of Neuro-fuzzy system on image segmentation, it has been applied for tumour detection in brain images [12]. The hybrid nature of NN and fuzzy logic provides a high accuracy rate of image classification or segmentation in various medical applications. The overall process comprises the following phases :

1. Image Acquisition
2. Pre-processing Images
3. Neuro-fuzzy application
4. Error or Abnormality detection.

The tumour cells have been recognized using the Fuzzy logic system and the backpropagation algorithm of NN.

Further, a comparative analysis has been made between the SVM and Neuro-fuzzy system in pattern recognition or classification [13]. The paper has given another dimension to application studies, and it has also analyzed the limitations of the procedures.

In [14], an elaborated survey is given about the various fuzzy clustering approaches for pattern recognition. The paper categorizes the analysis under three classifications as given below:

1. Fuzzy based Approaches
2. Genetic Algorithm based Approaches
3. Neuro-fuzzy based approaches

The third category has provided a satisfying informative section about the fuzzy based clustering approaches, that have been used for various image processing areas such as filtering, segmentation, detection, clustering and so on.

III. PROPOSED WORK

The combination of fuzzy logic and NN has developed into a potential research field, which integrates the efficiencies of both the error handling ability of fuzzy logics and the learning capability of NN. The proposed work has designed and implemented a Deep Fuzzy Model , which performs object categorization in an efficient way. The object segments termed as clusters are identified automatically using FCM (Fuzzy C-Mean clustering) and, the CNN is trained to categorize the multiple objects in a particular image.

Figure 3 portrays the generic work flow of the proposed model. The overall process of Deep Fuzzy model comprises the following functions :

1. Generation of Background Model and Background Elimination. (**Figure 3**)
2. Defining Error Function – Fuzzification .
3. Thresholding Block (**Figure 4**)
   - **A.** *Fuzzy Entropy Computation*
   - **B.** *NN Training &Tuning*

A. *Generation of Background Model and Background Elimination*

In certain images, the background is dynamic based on an instant, and the system will not have any knowledge about that. Hence, the first process is to generate a background model based on the training phase that needs some initial image sequence with objects. Moreover, in an image, each pixel is designed as a Gaussian Mixture Model Distribution. With that, the probability of each pixel can be stated as,

$$P(A_t) = \sum_{i=1}^{N} W_{i,t}\, F_i(A_t, \mu_{i,t}, \Sigma_{i,t}) \qquad (1)$$

In the equation, μ and 'W' denote Mean and weight respectively. 'N' is the number of distributions that is evaluated by the computational memory which is available at that instant. Further, 'F' represents the probability density function and the $\Sigma_{i,t}$ is the covariance matrix of 'N' th component. With those parameter settings, the Gaussian Function is given as,

$$F_i(A_t, \mu_{i,t}, \Sigma_{i,t}) = \frac{1}{(2\pi)^{\frac{D}{2}}|\Sigma_{i,t}|^{1/2}} \frac{e^{\frac{-1}{2}(A_t-\mu_{i,t})^T}}{\Sigma_{i,t}^{-1}(A_t-\mu_{i,t})} \qquad (2)$$

After initializing parameters, the background model has been accomplished and the model is utilized for evaluating the variations with the tested frames. In general, the background pixels are more focused than the foreground pixels. If any pixel is matching with the results of the distribution model and convinced with the following equation (3), it will be categorized under background pixel. The same procedure will be followed for each pixel in an image and the equation is as follows:

$$C = arg\left[min_c\left(\sum_{i=1}^{c} W_{i,t} > T\right)\right] \qquad (3)$$

Where, T states the Threshold value. Since, the background model is generated at each level pixelwise, the thresholding process in the foreground segmentation phase will have some advantages over cost complexities, which is required more for attaining real time performance in a better way.

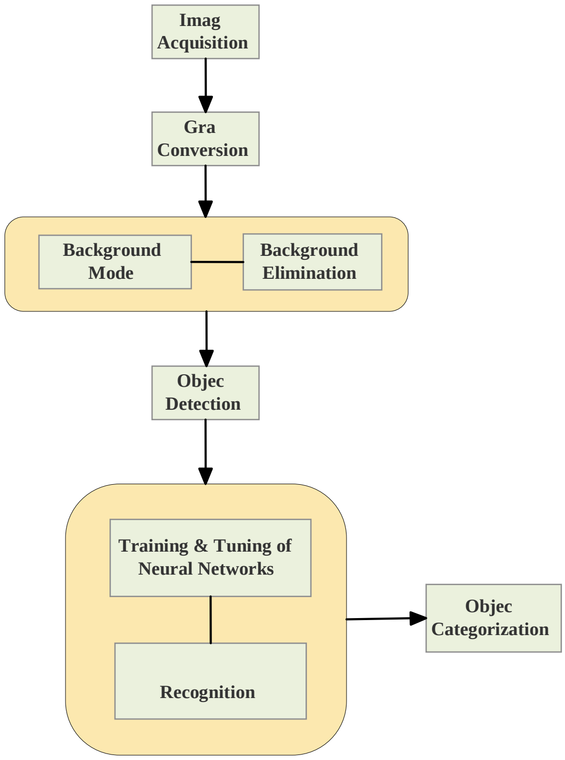

Figure 3: Generic Work Flow of the Proposed Deep Fuzzy Model

B. *Defining Error function*

This block technically involves validating clusters and fuzzification process. The main contribution of this block is to produce an intent error function which is further utilized by the robust thresholding. For evaluating an error function, the number of objects present in all planes in the desired image should be determined by cluster validation. In the adduced model, **Fuzzy C-means (FCM) Clustering algorithm** is incorporated, in which, the data points are grouped into N clusters, with every data point belonging to other clusters, to a certain degree.

- **FCM Algorithm :**

The Algorithm Steps are, as follows :

1.Initialize the data points into desired number of clusters.

2.Find the Centroid of each cluster.

$V_{ij} = \sum_1^n(\mu^m_{ik} * x_k)/\sum_1^n \mu^m_{ik}$

Where, μ = Fuzzy membership value of the data point,
m = Fuzziness parameter (= 2)
xk = data point

3.Find the distance of each point from both the centroids, using Euclidean Distance.

4.Update the membership values.

$\mu = \sum_1^n [(d^2_{ki} / d^2_{kj})^{1/m-1}]^{-1}$

5.Repeat the 2 nd, 3 rd and 4 th steps, unless centroids are not changing . Thus, the fuzzification block involves dividing the desired input into different fuzzy partitions. Iterative optimization takes place in accordance with the weighted comparison between the image pixels and each cluster point. The iteration process is done for determining the number of objects in the respective image for a range of determined clusters and predicting the appropriate option based on validating clusters. From the FCM algorithm, the minimized objective function is given below:

$$W_n(A,B) = \sum_{i=1}^{k}\sum_{k=1}^{C}(\mu_{ki})^n (p_{ki})^2 \qquad (4)$$

From the above equation (4), 'A' denotes the fuzzy C-set of the image and 'B' the cluster set centres. 'n' manages the cluster nature, μ_{ki} means the membership value of fuzzy in ith pixel and kth cluster and p_{ki} is the norm metric product.

C. *Thresholding Model*

The functions involved in Deep Fuzzy Model are given in **Figure 4**. Moreover, the thresholding block comprises two significant operations: Fuzzy Entropy Computation (FEC) and Neural Network Training and Tuning. In the overall process, the input images are acquired, and fuzziness measure has been produced with error function determination, and the output is provided as the segmented image. For appropriate image segmentation, it is very vital to describe the cluster numbers in an image. Further, the target values and thresholds are attained from the previous section, as the minimum and maximum fuzzy levels are demonstrated. The NN activation function has been framed, once the target and the threshold values are attained by the following equation (5):

$$f(x) = \sum_i \left(\frac{B_i - B_{i-1}}{1+e^{-(A-\theta_i)}} + B_{i-1}\right) \times \left[\mu\left((A - B_{i-1}) \times s^2\right) - \mu((A-Bi)\times s2)\right. \qquad (5)$$

Here, θ_i is representing the threshold value and B_i points the target rate of each sigmoid respectively. μ is considered as the step function and 's' denotes the dimension of the neighbourhood.

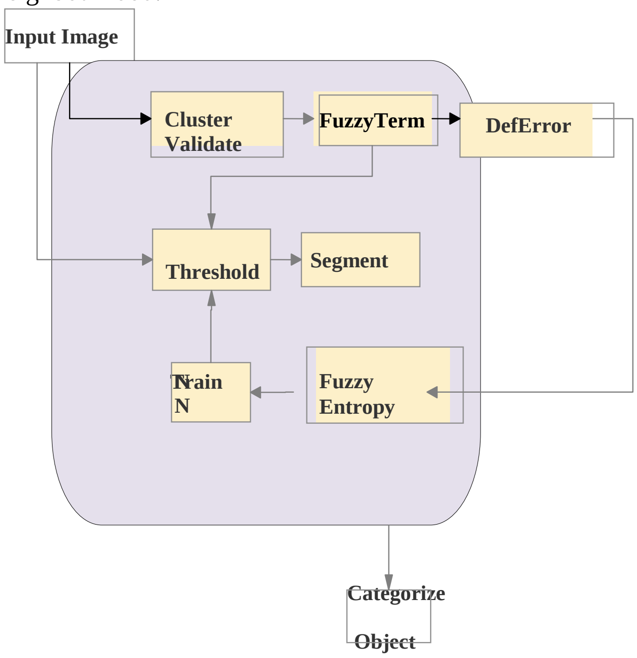

Figure 4: Operations in Deep Fuzzy Model.

1) *Fuzzy Entropy Computation (FEC)*

A function of fuzzy set is termed as fuzzy entropy here. It can also be denoted as the uncertainty measure or fuzziness measure. In the proposed Deep Fuzzy Model, the fuzzy entropy is incorporated to compute an error in the performance or prediction of the system. Because of this criterion, it handles the system based on the results generated by the error function definition (section 3.2.) at each training level. This section endeavours to minimize the fuzziness rate of the acquired input. Hence, the partition entropy is calculated on the basis of the following equation:

$$FEC = -\frac{1}{n \ln\left(\frac{1}{C}\right)} \sum_{i=1}^{k}\sum_{k=1}^{C}[\mu_{ki} \, ln\,(\mu_{ki})] \qquad (6)$$

2) *Neural Networks Training and Tuning*

The main aim of this block is to renew the weight connection as depicted in equation (7), with the concern of output error in NN. Moreover, the NN is trained based on the backpropagation algorithm. In each epoch of training, error is evaluated on the basis of the difference produced between the actual and the expected output of the network. As per the motive of this work, the NN is to be trained and tuned to reduce the errors in object categorization.

$$\Delta W_{xy} = n \left(\frac{-\partial E}{\partial O_x}\right)\frac{\partial O_x}{\partial I_x} O_y \Delta \quad \text{For output layer,}$$

$$n\left(\sum z\left(\frac{-\partial E}{\partial O_x}\frac{\partial O_z}{\partial I_z} W_{xy}\right)\right)\frac{\partial O_x}{\partial I_x} O_y \quad \text{For Hidden layer} \qquad (7)$$

In the above equation, 'n' denotes the learning rate of the network and 'E' the error predicted from the output. O_x the ith neuron x in one layer to neuron y in the adjacent layer and W_{xy} weight connection between the neurons from one layer to the next layer. Finally, I_x denotes the total input to the y neuron. When the training process is revised, the pixels are determined with more precision and the results obtained.

IV. RESULTS AND DISCUSSIONS

In the proposed Deep Fuzzy Model for Multi-Object Categorization in Scene (DFMOC), the analysis is carried out with the Caltech-101 dataset[12], which is considered to be one of the familiar and complex datasets and a synthesised dataset with 16 categories and 471 images.. Evaluation process has been done through VIfeat and MatConvNet tool boxes of MATLAB on Pentium IV. The built-in function for FCM in MATLAB tool, is used for effective clustering. For all computations, the proposed model incorporated the second –order 3✕3 neighborhood scheme for connectivity of neurons. The results obtained by the proposed DFMOC Model are depicted in **Figure 5** and also compared with L2-SVM, CNN and ANFIS approaches. The main aim of the evaluation is to compute, the Average Precision Rate of Categorization (APRC) and test the adaptability of the proposed model. Further, it is also been compared with some significantfactors like time, Error rate and computational complexity. APRC is evaluated using equation (8) , for testing the accuracy provided by the model.

$$APRC = \frac{N(AccurateCategorization)}{N(Test)} \qquad (8)$$

APRC is the fraction of Accurate Categorizations among the Total Tests done by the model. For making evidence to the predominance of the proposed method, the APRC values produced by the traditional models such as ANFIS, L2-SVM, CNN and the proposed DFMOC are compared in the following **Figure 5.** APRC comparative results are obtained, by increasing the input samples and depicted in **Table 1**. The results depicted from the **Table 1** along with **Figure 5** shows that, DFMOC model provides better precision rate than the existing models.

Table 1 Comparative Results of APRC

Number of Input Samples	Average Precision Rate of Categorization			
	DFMOC	ANFIS	L2-SVM	CNN
20	83	73.4	70.6	71.8
40	81.3	80.7	73.2	74.2
60	79.6	73.1	74.1	72
80	78	70.3	76.2	69.1
100	77	69	72	68

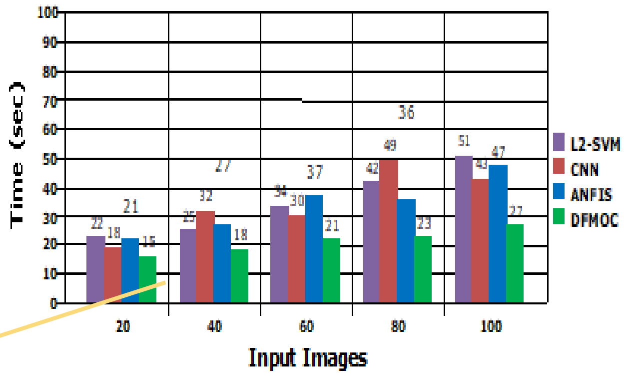

Figure 7 Time Consumption of DFMOC with other models.

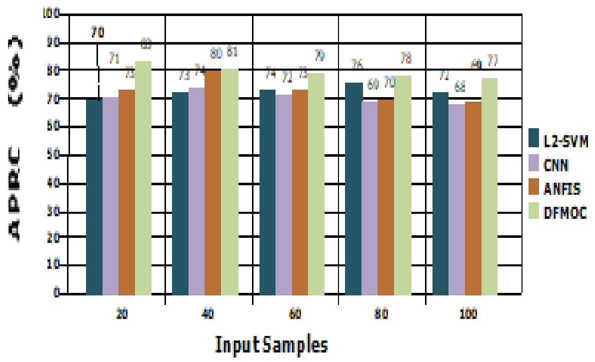

Figure 5 Comparison of Models based on APRC

The average Error rate in object categorization is evaluated on the input images. From the results obtained, it is apparent that, the error rate occurring in the results of the proposed work, is considerably less. This provides an evidence that, the proposed DFMOC outperforms the existing models. The following **Figure 6** portrays the results in accordance with the obtained input images. Moreover, when there is a higher precision rate, the error rate and the false positive values would be lower. At the time of experimentation, the DFMOC achieved this and also the time consumption has been significantly reduced. The graph presented in **Figure 7**depicts time utilization of DFMOC model with traditional models in object categorization. The graph points out that, the DFMOC model takes less time, on comparison to other models. Empirical evidence shows that, the proposed novel approach provides better object categorization results with more accuracy rate, less error and time consumption. **Figure 8**, **Figure 9**, **Figure 10** shows the Multi-Object Categorization, in the input scene.

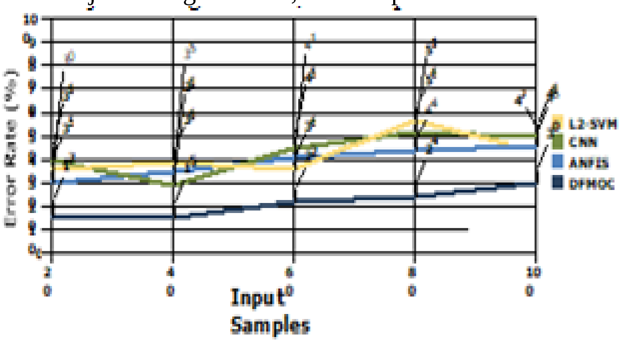

Figure 6 Error Rate of DFMOC with other models.

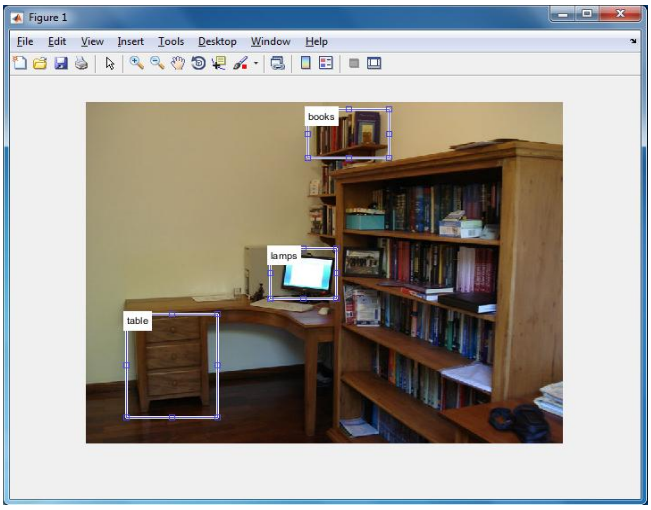

Figure 8 Multi-Object Categorization – Eg 1.

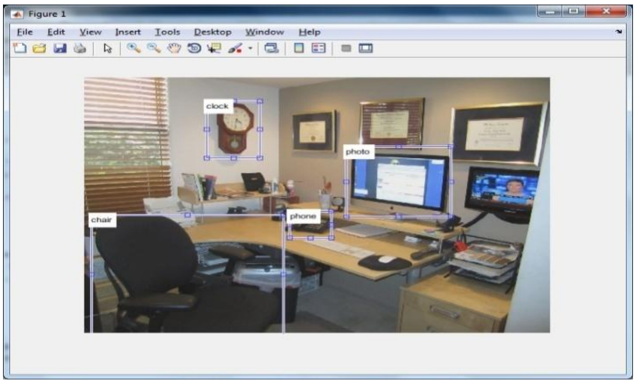

Figure 9 Multi-Object Categorization – Eg 2.

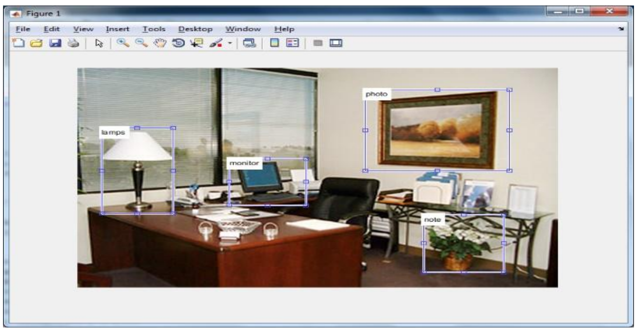

Figure 10Multi-Object Categorization – Eg 3

V. CONCLUSION

In this work , a Deep Fuzzy Model for Multi-Object Categorization in Scene has been developed and implemented. For efficient object detection at initial stage, background model has been framed by using Gaussian distribution. Following that, background elimination has also been done for ensuring the image clarity. Moreover, the model used the infused predominance of Convolution Neural Networks and Fuzzy Logic. Fuzzification was carried out using Fuzzy C- Means algorithm . The Thresholding Block performs two main functions in the proposed model namely, Fuzzy Entropy Computation and Neural Network Training and Tuning. The use of FCM Algorithm and consequently Fuzzy Entropy usage, has reduced the Computational complexity (computational cost), which was a bottleneck in ANFIS due to the large number of parameters (large inputs). The multilevel layers in Convolution NeuralNetwork have been trained and tuned for effective categorization, based on the threshold values. Furthermore, the Average Precision Rate of Categorization (APRC) has been determined, in order to prove the result accuracy. The proposed DFMOC model has produced better results than the other models (L2-SVM, CNN and ANFIS).

REFERENCES

1. Amaninder Kaur, Brar Vikas Wasson," Image Denoising Using Improved Neuro-fuzzy Based Algorithm: A Review",International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 4, April 2014, pp. 1072-1075.
2. R.Pushpavalli, G.Sivarajde, "Image Enhancement Using Adaptive Neuro-fuzzy Inference System", International Journal of Scientific & Technology Research , Vol 2, Iss 6, june 2013, pp. 256-262.
3. Kumaravel.S and Veni.S, "Recognizing Object Categories in Natural Scenes using ROI Segmentation in SVM with Spatio-Geometric constraints and Visual Similarity", Karpagam Journal of Computer Science, KAHE, Vol 12, Issue 6, 2018, pp 258-266.
4. Samarjit Kar, Sujit Das, Pijush Kanti Ghosh, "Applications of neuro fuzzy systems: A brief review and future outline", Springer publications, Applied Soft Computing 15, 2014, pp. 243–259.
5. R.Arun, D.Murugan, "An Innovative Method For Mr Brain Image Classification Corroborated On ANFIS System", International Journal of Computer Application, Vol 3, Iss 3 , May-June 2013, pp. 15-27.
6. Tomohiro Ishii, Ryosuke Nakamura, Hidemoto Nakada, Yoshihiko Mochizuki and Hiroshi Ishikawa, "Surface Object Recognition with CNN and SVM in Landsat 8 Images", 14th IAPR International Conference on Machine Vision Applications (MVA), 2015, pp. 341-344.
7. Yichuan Tang, "Deep Learning using Linear Support Vector Machines", International Conference on Machine Learning 2013: Challenges in Representation Learning Workshop, 2013.
8. Malay Shah, Rupal Kapdi," Object Detection Using Deep Neural Networks" International Conference on Intelligent Computing and Control Systems ICICCS 2017, pp. 787-790.
9. [9] Xiaogang Wang," Deep Learning in Object Recognition, Detection, and
10. Segmentation" Foundations and Trends R in Signal Processing, Vol. 8, Iss. 4 , 2014, pp. 217–382.
11. Marcin Korytkowski, Leszek Rutkowski, Rafał Scherer," Fast image classification by boosting fuzzy classifiers", Springer, Information Sciences 327, 2016, pp. 175–182.
12. Shweta Lawanya Rao and C.L. Chandrakar, "An Adaptive Neuro-Fuzzy
13. System For Automatic Image Segmentation And Edge Detection", IJDW Volume 4, Iss 1, January-June 2012, pp 33-40.
14. Chalumuri Revathi and B. Jagadeesh," Segmentation Of Brain Images For Tumor Detection Using Neuro-FuzzyInference System" Issn (Print): 2393-8374, (Online): 2394-0697, Vol 4, Iss 12, 2017, pp. 42-48.
15. Maya Nayak1 and Er. Jnana Ranjan Tripathy, "Pattern Classification Using Neuro Fuzzy and Support Vector Machine (SVM) – A Comparative Study" International Journal of Advanced Research in Computer andCommunication Engineering, Vol. 2, Iss 5, May 2013 , pp. 2301-2306.
16. G. Nagalakshmi and S. Jyothi, "A Survey on Pattern Recognition using Fuzzy Clustering Approaches" , International Refereed Journal of Engineering and Science(IRJES), Vol 2, Iss 5, May 2013 , pp .24-31.
17. Fei-Fei L, Fergus R and Perona P. One-Shot learning of object categories. IEEE Trans Pattern Analysis and Machine Intelligence, 28, 2016, pp. 594–611.

AUTHORS PROFILE

Dr S. Kumaravel completed his Ph.D from the Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore – 21 and is working as Assoc.Prof of Computer Science in SRMV College of Arts and Science, Coimbatore-20. He has presented 20 Papers in National and International Conferences and has published 4 Journal papers. His Research interests are in the Domain of Soft Computing and Image Analysis.

***Dr. S. Veni* is** Professor & HOD, Department of Computer Science, at Karpagam Academy of Higher Education, Coimbatore - 21. She has completed her Doctoral degree from Bharathiar University. She has published 47 Research articles and has attended various National and International conferences. Her Research areas includes Networks and Data Mining .

.