International Journal of Engineering & Technology, 7 (2.7) (2018) 614-617

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper

Image classification using Deep learning

M Manoj krishna[1]*, M Neelima[2], M Harshali[3], M Venu Gopala Rao[4]

[1,2]B tech Student, Dept. of ECE, K L E F, India
[3]SRF-CSIR, Dept. of ECE, K L E F, India
[4]Professor, Dept. of ECE, K L E F, India
**Corresponding author E-mail: manoj.krishna4301@gmail.com*

Abstract

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 use AlexNet architecture with convolutional neural networks for this purpose. Four test images are selected from the ImageNet database for the classification purpose. We cropped the images for various portion areas and conducted experiments. The results show the effectiveness of deep learning based image classification using AlexNet.

Keywords: *AlexNet; Convolutional Neural Networks; Deep Learning; Image Classification; ImageNet; Machine Learning*.

1. Introduction

Classification is a systematic arrangement in groups and categories based on its features. Image classification came into existence for decreasing the gap between the computer vision and human vision by training the computer with the data. The image classification is achieved by differentiating the image into the prescribed category based on the content of the vision. Motivation by [1], in this paper, we explore the study of image classification using deep learning.

The conventional methods used for image classifying is part and piece of the field of artificial intelligence (AI) formally called as machine learning. The machine learning consists of feature extraction module that extracts the important features such as edges, textures etc and a classification module that classify based on the features extracted. The main limitation of machine learning is, while separating, it can only extract certain set of features on images and unable to extract differentiating features from the training set of data. This disadvantage is rectified by using the deep learning [2]. Deep learning (DL) is a sub field to the machine learning, capable of learning through its own method of computing. A deep learning model is introduced to persistently break down information with a homogeneous structure like how a human would make determinations. To accomplish this, deep learning utilizes a layered structure of several algorithms expressed as an artificial neural system (ANN). The architecture of an ANN is simulated with the help of the biological neural network of the human brain. This makes the deep learning most capable than the standard machine learning models [3, 4].

In deep learning, we consider the neural networks that identify the image based on its features. This is accomplished for the building of a complete feature extraction model which is capable of solving the difficulties faced due to the conventional methods. The extractor of the integrated model should be able to learn extracting the differentiating features from the training set of images accurately. Many methods like GIST, histogram of gradient oriented and Local Binary Patterns, SIFT are used to classify the feature descriptors from the image.

The basic artificial neural network is outlined in Section-II. Section-III describes about AlexNet. The implementation and results are discussed in Section-IV. We conclude in section-V and finally the references are given at the end.

2. Artificial Neural Networks

A neural network is a combination of hardware bonded or separated by the software system which operates on the small part in the human brain called as neuron. A multi layered neural network can be proposed as an alternative of the above case. The training image samples should be more than nine times the number of parameters essential for tuning the classical classification under very good resolution. The multi-layered neural network is so complicated task with respect to its architecture in the real world implementations[14-17]. The multi-layered neural network is at present expressed as the Deep Learning.

In deep neural networks every node decides its basic inputs by itself and sends it to the next tier on behalf of the previous tier.

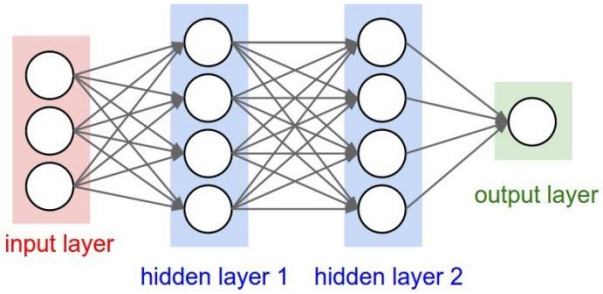


Fig. 1: Basic Deep Neural Network

We train the data in the networks by giving an input image and conveying the network about its output. Neural networks are expressed in terms of number of layers involved for producing the inputs and outputs and the depth of the neural network. Neural networks are involved in many principles like fuzzy logic, genetic algorithms and

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Bavesian methods. These layers are generally referred to as hidden layers. They are expressed in terms of number of hidden nodes and number of inputs and outputs every node consists. The Convolutional Neural Network (ConvNet) is most popular algorithm used for implementing the deep learning technique. The ConvNet consists of Feature detection layers and classification. A ConvNet is composed of several layers, and they are convolutional layers, max-pooling or average-pooling layers, and fully-connected layers.

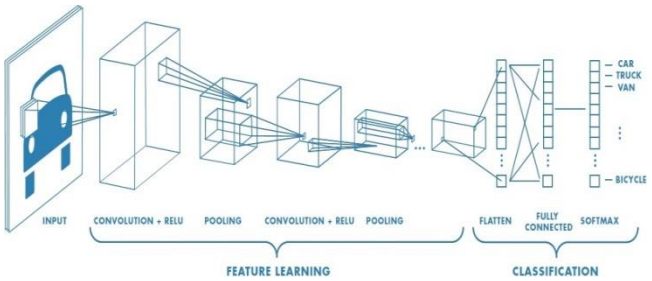


Fig. 2: Architecture of CNN

3. Alexnet

The ConvNet is categorized into two types named LeNet and AlexNet. The LeNet is expressed as the Shallow Convolutional Neural Networks which is designed to classify the hand-written digits. The LeNet comprises of 2 convolutional layers, 2 subsampling layers, 2 hidden layers and 1 output layer [5]. The AlexNet is expressed as the deep convolutional neural networks which are used for classifying the input image to one of the thousand classes.

AlexNet is used to solve many problems like indoor sense classification which is highly seen in artificial neural intelligence. It is a powerful method of knowing the features of the image with more differential vision in the computer field for the recognition of patterns. This paper discuss about the classification of a particular size of image of required choice. It can very effectively classify the training sample of images present in the AlexNet for better vision.

The AlexNet comprises of 5 convolutional layers, 3 sub sampling layers and 3 fully connected layers. The main difference between the LeNet and AlexNet are the type of Feature Extractor. We use the non-linearity in the Feature Extractor module in AlexNet whereas Log sinusoid is used in LeNet. AlexNet uses dropout which is not observed in any other data sets of networking.

4. Implementation, Results and Discussions

We selected four images Sea Anemone, Barometer, Stethoscope and Radio Interferometer from the ImageNet database for experimentation purpose (See Fig. 3) [6]. The block diagram of the architecture shown in Fig.4 and the corresponding implementation is illustrated below [7].



Fig. 3: Test images (a) Sea Anemone of size 375x500x3(b) Barometer of size 500x375x3 (c) stethoscope of size 375x500x3 (d) Radio interferometer of size 375x500x3

In the first layer, there are 96 11x11 filters are used at stride 4. The output volume size is 55x55x96. The AlexNet is trained on the GPU named GTX580 which is having a small amount of 3GB of memory. So, the CONV1 output will be halved and sent to two GPU's i.e. 55x55x48 is sent to each GPU. The 2nd, 4th, and 5th convolutional layers bits are related just to the part maps in the previous layer which dwell on the same GPU said in the figure. The kernels of the 3rd convolutional layer are associated with all kernel maps in the 2nd layer. The neurons in the fully connected layers are associated with all neurons in the past layer.

The 3rd, 4th, and 5th convolutional layers are associated with each other with no interceding pooling or standardization layers. The 3rd convolutional layer has 384 parts of size $3 \times 3 \times 256$ associated with the (standardized, pooled) yields of the 2nd convolutional layer. The fourth convolutional layer has 384 kernels of size $3 \times 3 \times 192$ and the fifth convolutional layer has 256 kernels of size $3 \times 3 \times 192$. The first two fully connected layers have 4096 neurons each.

We use the local response normalization in the normalization layer. There are two normalization layers present in the AlexNet architecture. The Deep Neural Network with ReLU Nonlinearity can train very fast than with the identical of the function tanh units. The ReLU considers quicker and more compelling training by mapping the negative esteems to zero and keeping up positive esteems. Signifying by the movement of a neuron figured by applying kernel i at position (x, y) and after that applying the ReLU nonlinearity, the response-normalized movement is expressed as

$$c^{i}_{(x,y)} = d^{i}_{(x,y)} \Big/ \left(k + \alpha \sum_{j=max(0,i-n/2)}^{min(N-1,i+n/2)} \left(d^{i}_{(x,y)} \right)^2 \right)^{\beta} \quad (1)$$

This kind of response standardization actualizes a type of parallel hindrance roused by the sort found in genuine neurons, making rivalry for huge exercises among neuron yields registered utilizing different kernels. The test images are cropped to various portion areas and applied for classification. The results are shown in Fig. 5, Fig. 6, Fig. 7 and Fig. 8. From the results, it is observed that in all cases of the cropped data, the classification is successful.

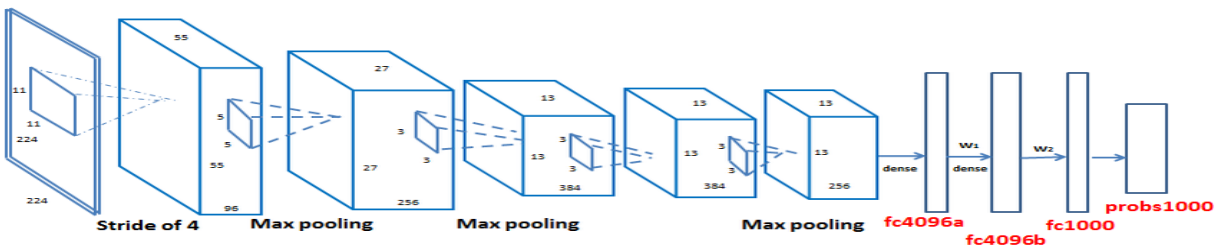


Fig. 4: AlexNet Architecture

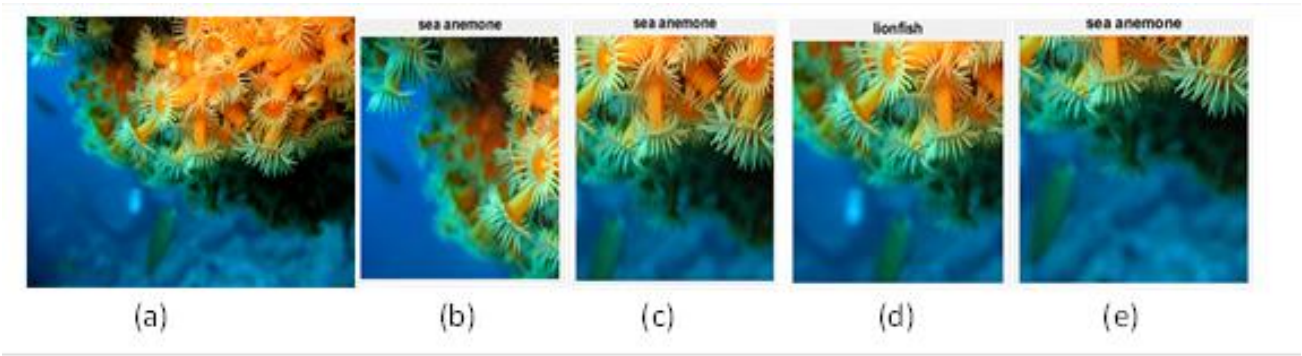


Fig. 5: Sea Anemone cropped to various areas

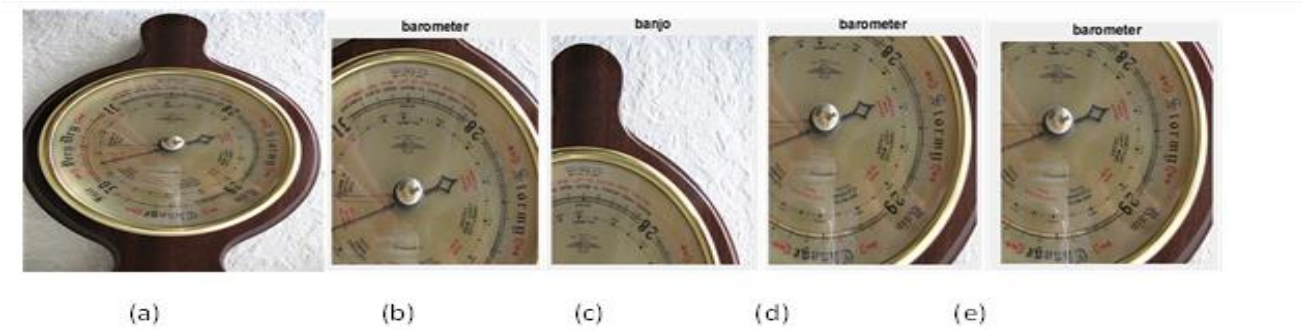


Fig. 6: Barometer cropped to various areas

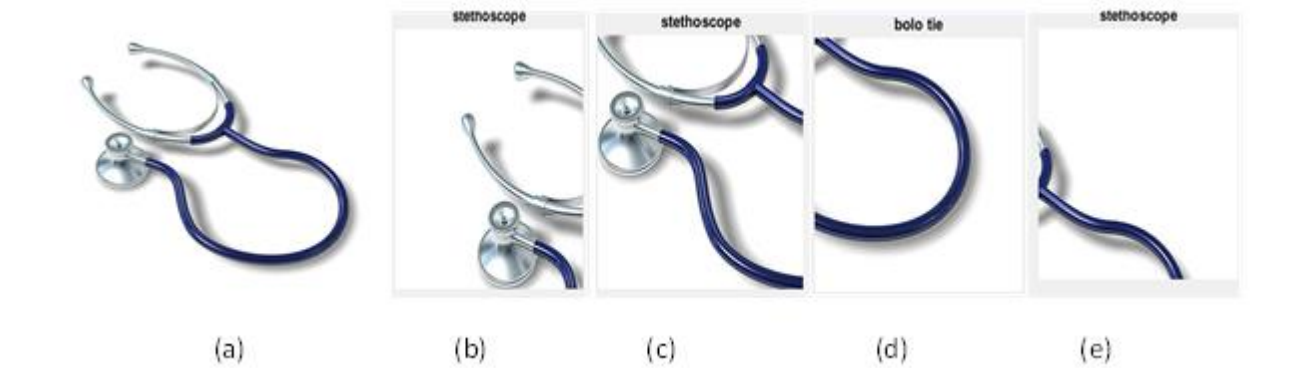


Fig. 7: Stethoscope cropped to various areas

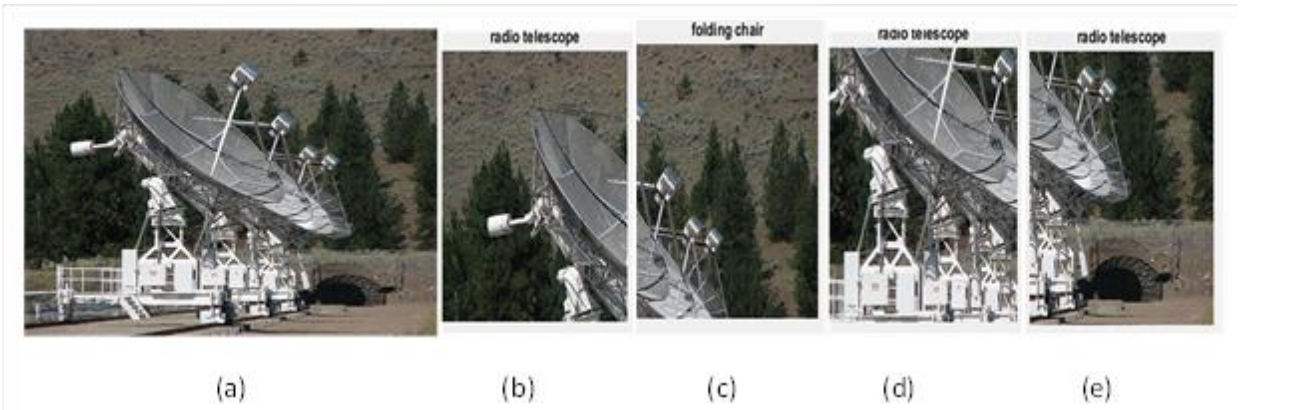


Fig. 8: Radio Interferometer cropped to various areas

5. Conclusion

Four test images sea anemone, barometer, stethoscope and radio interferometer are chosen from the AlexNet database for testing and validation of image classification using deep learning. The convolutional neural network is used in AlexNet architecture for classification purpose. From the experiments, it is observed that the images are classified correctly even for the portion of the test images and shows the effectiveness of deep learning algorithm.

Acknowledgement

The authors would like to express their deep gratitude towards the Department of ECE and the management of K L E F for their support and encouragement during this work.

References

[1] https://in.mathworks.com/matlabcentral/fileexchange/59133-neural-network-toolbox-tm--model-for-alexnet-network
[2] H. Lee, R. Grosse, R. Ranganath, and A.Y. Ng. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In Proceedings of the 26th Annual International Conference on Machine Learning, pages 609–616. ACM, 2009
[3] Deep Learning with MATLAB – matlab expo2018
[4] Introducing Deep Learning with the MATLAB – Deep Learning E-Book provided by the mathworks.
[5] https://www.completegate.com/2017022864/blog/deep-machine-learning-images-lenet-alexnet-cnn/all-pages
[6] Berg, J. Deng, and L. Fei-Fei. Large scale visual recognition challenge 2010. www.imagenet.org/challenges. 2010.
[7] Fei-Fei Li, Justin Johnson and Serena Yueng, "Lecture 9: CNN Architectures" May 2017.
[8] L. Fei-Fei, R. Fergus, and P. Perona. Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories. Computer Vision and Image Understanding, 106(1):59–70, 2007.
[9] J. Sánchez and F. Perronnin. High-dimensional signature compression for large-scale image classification. In Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, pages 1665–1672. IEEE, 2011.
[10] https://in.mathworks.com/help/vision/examples/image-category-classification-using-deep-learning.html
[11] Alex Krizhevsky, Ilya Sutskever and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks" May 2015.
[12] A. Krizhevsky. Learning multiple layers of features from tiny images. Master's thesis, Department of Computer Science, University of Toronto, 2009.
[13] https://in.mathworks.com/help/nnet/deep-learning-imageclassification.html
[14] KISHORE, P.V.V., KISHORE, S.R.C. and PRASAD, M.V.D., 2013. Conglomeration of hand shapes and texture information for recognizing gestures of indian sign language using feed forward neural networks. International Journal of Engineering and Technology, 5(5), pp. 3742-3756.
[15] RAMKIRAN, D.S., MADHAV, B.T.P., PRASANTH, A.M., HARSHA, N.S., VARDHAN, V., AVINASH, K., CHAITANYA, M.N. and NAGASAI, U.S., 2015. Novel compact asymmetrical fractal aperture Notch band antenna. Leonardo Electronic Journal of Practices and Technologies, 14(27), pp. 1-12.
[16] KARTHIK, G.V.S., FATHIMA, S.Y., RAHMAN, M.Z.U., AHAMED, S.R. and LAY-EKUAKILLE, A., 2013. Efficient signal conditioning techniques for brain activity in remote health monitoring network. IEEE Sensors Journal, 13(9), pp. 3273-3283.
[17] KISHORE, P.V.V., PRASAD, M.V.D., PRASAD, C.R. and RAHUL, R., 2015. 4-Camera model for sign language recognition using elliptical fourier descriptors and ANN, International Conference on Signal Processing and Communication Engineering Systems - Proceedings of SPACES 2015, in Association with IEEE 2015, pp. 34-38.