Deep Inside Feature Learning for Image Classification

Ranjini Surendran[1] and D. Jude Hemanth[1]

[1,2]Dept. of ECE, Karunya University, Coimbatore, India
E-mail: jude_hemanth@rediffmail.com

Abstract. A scene consists of manifold objects and visualizing them manually is relatively easier when compared to the task of machines. Training a machine to perform visualize an image automatically is an important task in self driving cars, robotics, scene understanding at the moment. Importance is for classifying image whatever be their variants with respect to their scaling, rotation, luminance, depth but belonging to the same class. What features do they possess could be generated by training a CNN by varying their number of hidden layers? Deeper the layers more will be the features learned so that we can easily understand what pixels are on the background and what on the objects. Success of classification lies in the success of feature learning. In visual recognition it is still the features from low – level to high- level which matters. Deep network can showcase its efficacy when large data set trains the network.

Keywords: Convolutional Neural Network, Deep Learning, Feature Learning, Image Classification, Visual Recognition

1 Introduction

Scene understanding in deep representation have changed the revelation of computer vision tasks in classifying different objects in a complex scene by training a Convolutional Neural Network [1,2,3,4] which learns the features of the image . Understanding an image by classifying image level description as in past using hand-engineered features is meager as compared to machine learned features to classify the different images in the scene which are much complex.

Classification of an image transpires accurately by the lower layers of the network. The top level layers don't put in much contribution in feature learning in the classification progression. As the hidden layers are augmented various pertinent features may get condensed and the layers will get over fitted. Lower layers of the network learns more low level features which bear the minuscule details of the image and decides the feature learning effectiveness of the network. Training a CNN [1] for extracting features of an image can improve the performance of image classification.

2 Previous Works

One of the key errands in the field of computer visual task is still image classification. This come up with the emerging developments in the feature learning tasks of CNN as compared to the traditional hand-engineered features using HOG [5], BoW [6], SIFT [6], Spatial Pyramid [7] to extract features to describe an image. These features were then given to any of the machine learning classifiers like Decision trees, Random Forest or SVM [8]. Learning features automatically with large image set paved the way to increasing demand of CNN [10] with deep network.

3 Methodology

The proposed method aims at classifying an image with following stages: training image set, testing image set, preprocessing and Deep Network using CNN for feature extraction. The steps for the projected system can be summarized as follows:

- Selecting the required image data set
- Splitting the data set into testing and training categories
- Preprocessing both datasets with methods like cropping unwanted regions
- Feeding the preprocessed image to the proposed CNN model where the layers get trained by updating the weights with the input images
- Once network is trained new set of image from testing data set is given to network

3.1 Preprocessing

The image data set taken for both training and testing need to be preprocessed before applying to network. This will include removing noise and selecting the region of interest which can help the network to learn features in minimum training time and avoiding problems like over fitting.

3.2 Feature Extraction & Classification

CNN is trained with a set of images which has different rotation angle, scaling and blurring. The model is proposed with three convolutional layers and two fully connected layers. The preprocessed image is given to the input layer . This layer does not for study any features of the image. Each convolutional layer is followed by nonlinear layer and pooling layer. Convolutional layer maps the input pixels to certain neurons in different regions. Non-linear layer turns negative values to zero and can fix vanishing gradients. The pooling layer down samples data to reduce the number of inputs to the next layer. Each layer learns some feature from the image. Features learned by the first hidden layer is transferred to the second hidden layer where it learns new features which are carried over to the next layer with the weights updated. The top fully connected layers get trained with higher level features of the image.

Once the model gets trained new image is tested. In feature extraction, the top layers of the network get trained with rest of the network remaining fixed. Higher level features that are learned from original image set will be transferred to the new image set and the image is classified.

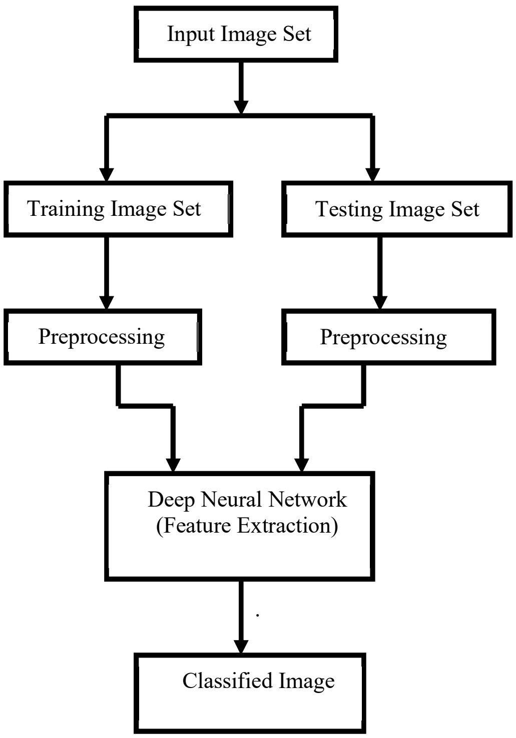

Fig. 15. Block Diagram

3 Experimental Results and Discussions

Proposed CNN model worked with an accuracy of 95% in classifying images during training phase showed an accuracy of 94% during the testing phase of new image. The model is trained 30 times in forward and backward pass and the with the result shown. Fig 2. shows accuracy improved at with 15 epochs even in the training phase. Fig 3. Compared the accuracy achieved during the training and testing phase. The model showed a better result images of different scaling, blurring and rotation.

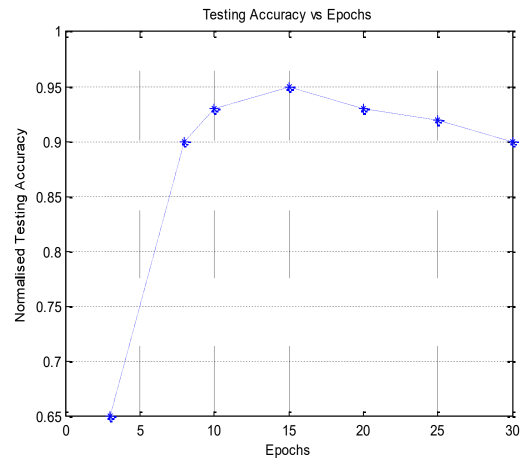

Fig. 2. Testing Accuracy vs epochs.

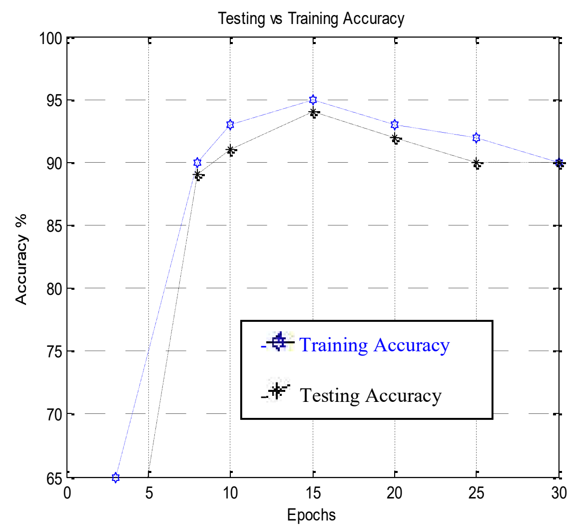

Fig. 3. Training vs Testing Accuracy.

4 Conclusions

The CNN network trained in this paper is carried out on limited number of images with few hidden layers. These limited hidden layers consume much time of the machine in getting trained to the 30 epochs given. The result was considerable with high accuracy on this small image set with varying scaling, flipping and rotation variants. The proposed model could be improved by increasing hidden layers to extract more features and train and test the network in less loss and high accuracy. The number of layers in a network is still a kind of trial and error because hidden layers are limited due to practical constraints like available memory, relevant data and running time.

References

1. Yim, J., Ju, J., Jung, H., & Kim, J. (2015). *Image Classification Using Convolutional Neural Networks With Multi-stage Feature. Robot Intelligence Technology and Applications 3, 587–594.* doi:10.1007/978-3-319-16841-8_52
2. LeCun, Y., Huang, F.J., Bottou, L.: Learning methods for generic object recognition with invariance to pose and lighting. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2004, vol. 2, pp. II–97. IEEE (2004)
3. Jarrett, K., Kavukcuoglu, K., Ranzato, M., LeCun, Y.: What is the best multi-stage architecture for object recognition? In: 2009 IEEE 12th International Conference on Computer Vision, pp. 2146–2153. IEEE (2009)
4. National Y. LeCun, K. Kavukcuoglu, and C. Farabet. Convolutional networks and applications in vision. In Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on, pages 253–256.IEEE, 2010.
5. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, vol. 1, pp. 886–893. IEEE (2005).
6. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2), 91–110 (2004).
7. Lazebnik, C. Schmid, J. Ponce, Beyond bags of features: spatial pyramid matching for recognizing natural scene categories, in: Proceedings of the CVPR, 2006.
8. Y.Lin, F.Lv, S.Zhu, etal., Large-scale image classification: fast feature extraction and svm training, in: Proceedings of the CVPR, 2011.
9. D.Ciresan, U. Meier, J. Schmidhuber, Multi-column deep neural networks for image classification, in: Proceedings of the CVPR, 2012.
10. A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Proceedings of the NIPS, 2012