Brand Detection System using Deep Learning

Saumya Singh, Saumya Dubey, Paridhi Singhal, Rajiv Kumar

Abstract*: In this paper a method of recognizing logos of the brand of cosmetic products using deep learning. There are several of hoax product which easily copies the famous brand's logo and deteriorates the company's image. The machine learning has proved to be useful in various of the fields like medical, object detection, vehicle logo recognitions. But till now very few of the works have been performed in cosmetic field. This field is covered using the model sequential convolutional neural network using Tensorflow and Keras. For the visual representation of the result Tensorboard is used. Work have been started with two of the brands-Lakme and L'Oreal. Depending upon the success of this technique, further brands for logo may be added for recognition. The accuracy of approximately 80% was obtained using this technique.

***Keywords* : Brand detection, deep learning, tensorflow, CNN, overfitting**

I. INTRODUCTION

Logo recognition for videos and images is the main problem in huge range of applications. One such area of application was introduced in Deep learning for Logo Recognition[1]. Deep learning for various logo recognition like Adidas, Texaco, Hewlett Packard. The Convolutional Neural network was trained using FlickrLogos-32 and Logos-32 plus dataset. The authors have used selective search algorithm to generate object proposals and the images collected are further processed by augmentation, rotation to check the variation of result. In vehicle logo recognition for intelligent traffic control systems[1], the regions are extracted using the idea first investigated by Girshick et al[12].Pre-trained CNN is used as a feature extractor, Selective search for extracting object proposals and a linear SVM is used for logo recognition and classification. Similarly in Logo Recognition Using CNN Features[13], unsupervised segmentation algorithm is used to get candidate sub windows and the SVM based classification. The similar approach was followed in Image Recognition with Deep learning[7] to classify food from images. All of the above work used similar kind of pipeline for the classification and feature extraction .Later in Image recognition and processing using Artificial Neural Network[6] technique of filtering the image for classification was introduced. The images are filtered and then average error is calculated which helps in further classification.

The work of the ANN was simplified by the introduction of CNN for image processing. The above works have shown that CNN gave us promising results for image detection. The simple convolutional neural network was created by Guo, Tianmei, et al [7] with less computational cost. Various of the optimization algorithm, use of the proper activation functions like ReLu, use of dropout layer was introduced. Thus the sequence to sequence learning has shown result with less computational cost. This was again proved in CNN is all you need [10] where Recurrent Neural Network was challenged

The above mentioned works have their pre built dataset but the cosmetic field lacked the prebuilt dataset.

II. DATASET DEVELOPMENT

Among the several techniques for dataset acquisition, in this paper, Digital and Social Media techniques is used to build dataset. Since, no pre-existing dataset is available for the images of the cosmetic brand products; dataset is built by collecting images from various online authentic sites, such as Nykaa.

The authenticity for the products available is proven by the authentication certificates stating that all the brands product are original, hence, the images provided are also of the original product.

Out of various available brands, it was decided to go for only 2 brands for the beginning, so L'Oreal Paris and Lakme is chosen. The table 1, below shows the final and the total number of images for each brand product, the label associated with each brand, also the number of dataset split for training and validation purpose.

Table 1. Dataset Detail

Details/dataset	Lakme	L'Oreal	Total count
Label	0	1	-
Total images	389	267	656
Training set (70%)	272	187	459
Validation set (30%)	117	80	197

III. PROPOSED METHODOLOGY

For building the model for classification, libraries and functions is used to build a model through Tensorflow and Keras. Convolutional layer, dense layer, number of nodes through layers, optimization of the model varies as per the task to carry.

Revised Manuscript Received on July 22, 2019.

***** Correspondence Author

Saumya Singh*, Computer Science Engineering department, G L Bajaj Institute of Technology and Management, Greater Noida, U.P. Email: saumyasingh111@gmail.com

Saumya Dubey, Computer Science Engineering department, G L Bajaj Institute of Technology and Management, Greater Noida, U.P. Email: saumyadubeygkp@gmail.com

Paridhi Singhal, Computer Science Engineering department, G L Bajaj Institute of Technology and Management, Greater Noida, U.P. Email: singhalparidhi5@gmail.com

Dr. Rajiv Kumar, Professor, Department of Computer Science and Engineering, G L Bajaj Institute of Technology and Management, Greater Noida, UP, India.

The model's working is also affected by the type of dataset it gets to learn on. Thus it is very important to improve dataset's quality before passing it to network.

A. Data Preprocessing

Image processing techniques helps in increasing the efficiency of the system [4], the methods that we here choose are:

a. Resizing the image- All the images here are resized to 110pixel which is perfect for network. Now the size selected to pass the network should be such that it should be the smallest recognizable figure. Too large size may increase the complexity of network and too small will cause problem in feature extraction from the image.

b. Data Augmentation- Data augmentation helps in increasing the efficiency of network. By rotating images by 180 degree and cropping, has increased dataset size from 300 to more than 600.

Table 2. Images before and after data augmentation

Data Augmentation technique	Original Image	Augmented Image
Crop		
Rotation		
Flip		

B. Network Architecture

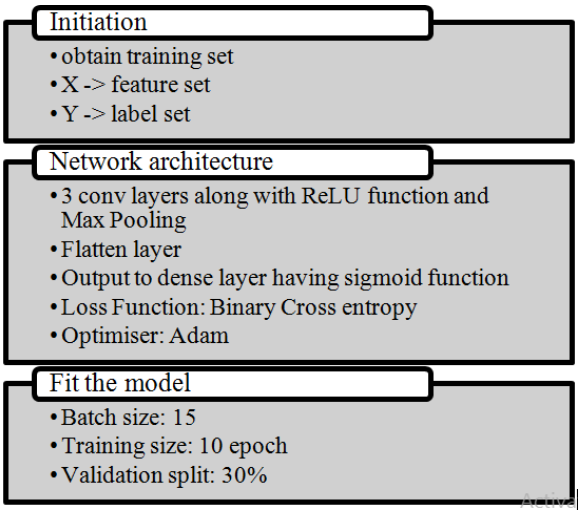

Figure 1. Flowchart showing model's summary

The above flowchart, shown in figure 1, represents the basic flow of working of this model. The model is trained on the training set, which is 70% of the total dataset.

The output of all this sent to the dense layer also known as Fully Connected Layer, having sigmoid function. Furthermore, Binary Cross Entropy is used as the loss function and Adam Optimizer as well.

For the training of model, a batch size of 15 images was selected and training size was set to 10 epoch and validation split was of 30%. All of these were determined on experiment basis.

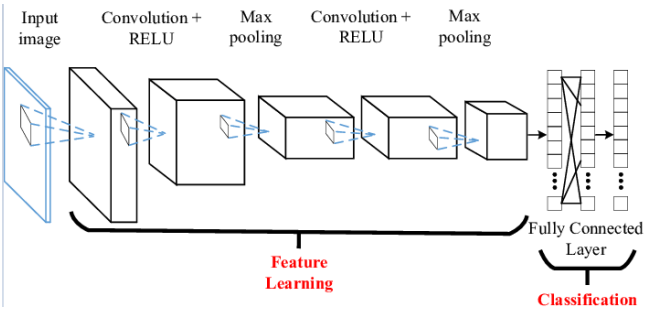

Figure 2. Architecture of sequential convNet

As shown in figure 2 above, a set of convolution layer with ReLU activation function and filter size of 3x3 followed by Pooling Layer with Max Pool operation is taken; this whole combination is repeated 3 times. In this set each layer contains a total of 64 nodes. The input of this fully connected layer is the output passed from the final convolutional layer and pooling layer. The task of this layer is to flatten the output; it means that multidimensional matrix is converted to vector. After passing this layer, the output vector is finally passed to a softmax activation function, which is used to get the probabilities of the input belonging to a particular class, hence the classification is done. In this architecture, a single dropout layer is also used to reduce overfitting found in the model. Further, for compiling the model, the loss function which was used was, Binary cross entropy, this is the default function that is used for the binary classification problem. While the optimizer that was chosen for the model was the Adam Optimizer, because it adjusts the learning rate throughout training.

C.Evaluation

Evaluation for Accuracy and Loss function is calculated with the help of Binary cross-entropy loss function. Training set and validation set is used while training of the model. Accuracy is calculated by using True positive(TP), True negative(TN), False positive(FP) and False Negative(FN) as shown in eq(i).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad \ldots \text{Eq(i)}$$

IV. RESULT AND DISCUSSION

A. Obtained Result:

After training the model with dataset, result is obtained on four parameters: epoch_accuracy, epoch_loss, validation_acc and val_loss as shown in below table 3 and graphs are shown in figure 3. Validation result is obtained on validation dataset and epoch result are obtained on training dataset.

Table 3. Obtained Result

Results	Avg loss	Avg val_loss	Avg acc	Avg val_acc
No.of epochs(10)	37%	48%	80.89%	77.36%

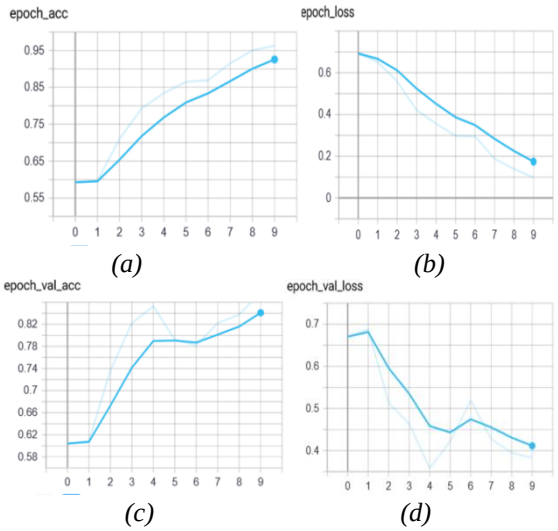

Figure 3. Graphical representation of Results

B. Discussion:

Working of the model can be evaluated on the basis of the difference between avg_loss and val_loss. There must be not much difference between val_loss and epoch_loss. Difference between them causes the problem of overfitting and underfitting as discussed in [15]. To remove it we performed following methods. a. Change in number of epochs: In beginning 20 epochs is used where huge difference was observed between avg_loss and val_loss. When the epochs were reduced to 10 the difference between them decreased; the results of above explained are shown in table 4

Table 4. Obtained result by change in epochs

No. of epochs	Avg_loss	Avgval_loss
20 epochs	20%	57.23%
10 epochs	37%	48.1

b. Change in Dataset size: The initial dataset size was approximately 300 images. This size caused overfitting during training, to reduce it some of the data augmentation techniques were performed. The new size of our dataset was more than 600 images and thus beneficial result was obtained during training. Val_loss decreased after increasing the dataset size and so the difference between val_loss and epoch_loss, results shown in table 5.

Table 5. Obtained result by dataset increamnet

Dataset size	Avg_Val_loss
328	73.05
656	48%

V. EXPERIMENTAL SETUP

Selection of convolution layer, number of nodes depends upon the network's working. For the proper working of model few of the experiments were performed to select the proper network architecture. Result is selected such that least difference between val_loss and epoch_loss can be obtained. Starting with the selection of convolution layer, results of 2 layers and 3 layers were examined. With 2 layers, difference obtained was more than 15% while with 3 layers difference was reduced and also the value of losses. The node selection was made between 64 nodes and 128 nodes with 3 conv layers. With 128 nodes, range of loss were between 40-60 and the difference between them were around 20%. When 64 nodes were tried this range was reduced to 30-50 and the difference obtained were 10%. Now to further reduce the values of loss, dropout function is added. By adding dropout function to 3 conv layer, 64 nodes, there was drop in difference between val_loss and epoch_loss. Thus model with 3 conv layer, 64 nodes, dropout function is selected. The table 6 below, shows the experimental results.

Table 6. Experimental Setup

Parameters	Epoch_loss	Val_loss	Difference
3 conv layer	49%	61.23%	12.23%
2 conv layer	52%	69%	17%
3 conv layer, 128 modes	48%	64%	16%
3 conv layer, 64 nodes	39%	51.22%	12.22%
3 conv layer, 64 nodes, dropout added	37%	48%	11%

VI. CONCLUSION AND FUTURE SCOPE

Result of 80.89% was obtained using this network where it is able to classify the logos of the product thus Convolutional Neural Network is a good approach. Here only two brandsis included, dataset size is also very small. By addition of more brands it can be shifted from binary classification to multi-class classification. The system will be able to identify the product by their images. More the number of data, more accurate will be extracted feature. Thus the system will also be able to identify the difference between real and hoax brand's product which are impossible from naked eyes. By including more of the augmentation techniques, qualities of the images may be further improved and various angle variation of a single image can be obtained.

REFERENCES

1. Bianco, Simone, Marco Buzzelli, Davide Mazzini, and RaimondoSchettini. "Deep learning for logo recognition." Neurocomputing 245 (2017): 23-30.
2. Huang, Yue, Ruiwen Wu, Ye Sun, Wei Wang, and Xinghao Ding. "Vehicle logo recognition system based on convolutional neural networks with a pretraining strategy." IEEE Transactions on Intelligent Transportation Systems 16, no. 4 (2015): 1951-1960.
3. Veluchamy, Magudeeswaran, KarthikeyanPerumal, and ThirumuruganPonuchamy. "Feature extraction and classification of blood cells using artificial neural network." American journal of applied sciences 9, no. 5 (2012): 615.
4. Islam, MdTohidul, BM NafizKarimSiddique, SagidurRahman, and TaskeedJabid. "Image Recognition with Deep Learning." In 2018 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), vol. 3, pp. 106-110. IEEE, 2018.

5. Szegedy, Christian, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and ZbigniewWojna. "Rethinking the inception architecture for computer vision." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818-2826. 2016.
6. Quraishi, MdIqbal, J. Pal Choudhury, and Mallika De. "Image recognition and processing using Artificial Neural Network." In 2012 1st International Conference on Recent Advances in Information Technology (RAIT), pp. 95-100. IEEE, 2012.
7. Guo, Tianmei, Jiwen Dong, Henjian Li, and YunxingGao. "Simple convolutional neural network on image classification." In 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(, pp. 721-724. IEEE, 2017.
8. Chandana, N., P. S. Chondamma, M. Harshitha, M. Anusha, and S. Padmashree. "Brand Logo Detection Using Convolutional Neural Network."
9. Begum, Arshiya, Farheen Fatima, and AsfiaSabahath. "Implementation of Deep Learning Algorithm with Perceptron using TenzorFlow Library." In 2019 International Conference on Communication and Signal Processing (ICCSP), pp. 0172-0175. IEEE, 2019.
10. Chen, Qiming, and Ren Wu. "CNN is all you need." arXiv preprint arXiv:1712.09662 (2017).
11. Bukhari, Syed Aftab Hassan. "What is Comparative Study." Available at SSRN 1962328 (2011).
12. Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. "Rich feature hierarchies for accurate object detection and semantic segmentation." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580-587. 2014.
13. Bianco, Simone, Marco Buzzelli, Davide Mazzini, and RaimondoSchettini. "Logo recognition using cnn features." In International Conference on Image Analysis and Processing, pp. 438-448. Springer, Cham, 2015.
14. Nisa, SeharUn, and Muhammad Imran. "A Critical Review of Object Detection using Convolution Neural Network." In 2019 2nd International Conference on Communication, Computing and Digital systems (C-CODE), pp. 154-159. IEEE, 2019.
15. [15] Jabbar, H., and RafiqulZaman Khan. "Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study)." Computer Science, Communication and Instrumentation Devices (2015): 163-172.

AUTHORS PROFILE

Saumya Singh is doing her final year B.Tech from G L Bajaj Institute of Technology and Management, Greater Noida, U.P, India. Her final year project is based on the topic “Brand Detection System using Deep Learning”. She also wishes to pursue her career and research interests in the field of Machine learning and Artificial Intelligence.

Saumya Dubey is doing her final year B.Tech from G L Bajaj Institute of Technology and Management, Greater Noida, U.P, India. Her final year project is based on the topic “Brand Detection System using Deep Learning”. She also wishes to pursue her career and research interests in the field of Machine learning and Artificial Intelligence.

Paridhi Singhal is doing her final year B.Tech from G L Bajaj Institute of Technology and Management, Greater Noida, U.P, India. Her final year project is based on the topic “Brand Detection System using Deep Learning”. She also wishes to pursue her career and research interests in the field of Machine learning and Artificial Intelligence.

Dr. Rajiv Kumar received the Bachelor of Technology degree from Bangalore University, Master of Technology degree in Information Technology from Bengal Engineering and Science University, West Bengal, Ph.D from Sharda University, Greater Noida. Currently he is working as Professor in the Department of Computer Science and Engineering at G L Bajaj Institute of Technology and Management, Greater Noida, UP, India. His research interests include Image Processing, Pattern Recognition, Character Recognition, Agent Based Software Engineering and Soft Computing.