Optimal Accuracy Zone Identification in Object Detection Technique - A Learning Rate Methodology

Chamarty Anusha, P S Avadhani

Abstract: In the recent past, Deep Learning models [1] are predominantly being used in Object Detection algorithms due to their accurate Image Recognition capability. These models extract features from the input images and videos [2] for identification of objects present in them. Various applications of these models include Image Processing, Video analysis, Speech Recognition, Biomedical Image Analysis, Biometric Recognition, Iris Recognition, National Security applications, Cyber Security, Natural Language Processing [3], Weather Forecasting applications, Renewable Energy Generation Scheduling etc. These models utilize the concept of Convolutional Neural Network (CNN) [3], which constitutes several layers of artificial neurons. The accuracy of Deep Learning models [1] depends on various parameters such as ‘Learning-rate’, ‘Training batch size’, ‘Validation batch size’, ‘Activation Function’, ‘Drop-out rate’ etc. These parameters are known as Hyper-Parameters. Object detection accuracy depends on selection of Hyper-parameters and these in-turn decides the optimum accuracy.

Hence, finding the best values for these parameters is a challenging task. Fine-Tuning is a process used for selection of a suitable Hyper-Parameter value for improvement of object detection accuracy.

Selection of an inappropriate Hyper-Parameter value, leads to Over-Fitting or Under-Fitting of data. Over-Fitting is a case, when training data is larger than the required, which results in learning noise and inaccurate object detection. Under-fitting is a case, when the model is unable to capture the trend of the data and which leads to more erroneous results in testing or training data.

In this paper, a balance between Over-Fitting and Under-fitting is achieved by varying the ‘Learning rate’ of various Deep Learning models. Four Deep Learning Models such as VGG16, VGG19, InceptionV3 and Xception are considered in this paper for analysis purpose. The best zone of Learning-rate for each model, in respect of maximum Object Detection accuracy, is analyzed. In this paper a dataset of 70 object classes is taken and the prediction accuracy is analyzed by changing the ‘Learning-rate’ and keeping the rest of the Hyper-Parameters constant. This paper mainly concentrates on the impact of ‘Learning-rate’ on accuracy and identifies an optimum accuracy zone in Object Detection.

Keywords: Learning-rate, Deep Learning, Object Detection, Hyper-Parameter, Convolution Neural Network, VGG-16, VGG-19, InceptionV3, Xception.

I. INTRODUCTION

Hyper-Parameters are a set of functional parameters in Object Detection models, whose values are to be fixed or initialized before starting of the training process and the same Parameters needs to be

Fine-Tuned for better accuracy. Important Hyper-parameters that affect the fine-tuning accuracy are:

- Learning Rate
- Number of Epochs
- Training Batch Size & Validation Batch Size
- Activation Function [3]
- Drop-out rate
- Optimization Algorithm [3]

A. Learning Rate

Learning rate [4] determines the weights of the network and update the weights during the training process. It regulates the speed with which the model learns the weights in the network. Large learning rate results in a fast learning model, requires less number of epochs but results in sub-optimal weights. On the other hand, a small learning rate makes the model to learn slower but requires more number of epochs to train and results in optimal weights. Learning rate is the most important parameter to be analyzed. Learning rate also impacts the training speed because larger learning rates result in faster convergence. On the other hand, too large or too small learning rates results in huge training time [11].

Finding a better value of learning rate [4] for a particular model and for a particular dataset is important for Fine-Tuning a Neural-Network.

B. Number of Epochs

Number of Epochs is a count of “Number of times the entire training sample dataset is passed through the learning algorithm”. An epoch is an iteration, in which one forward and one backward pass of all the training samples complete. After each epoch, the sample in the training dataset updates the weights of the Neural Network.

C. Batch Size

Batch size is the number of samples of a dataset passed through a CNN in a particular iteration. A dataset is commonly sub-divided into Training and Validation Datasets. There are three types of Learning algorithms [12] based on the batch size:

i. Batch Gradient Descent – In this algorithm, batch size is equal to the size of training dataset.

ii. Stochastic Gradient Descent – In this algorithm, Batch size consists of only one image.

iii. Mini Batch Gradient Descent – Here, Batch size is in between the above two algorithms. Generally the sizes chosen in this algorithm are 32, 64, 128 etc.

D. Activation Functions [3]

These functions are very important for Deep Learning models for establishing a complex functional mapping between inputs and outputs.
They introduce non-linearity in the network and convert input signals to output signals. The main purpose of these functions is to make the network powerful by learning complicated data like audio, video, images etc. Some most popular activation functions in use are Sigmoid, Tanh, ReLU etc.

E. Drop-Out
Drop-out means dropping neurons or nodes, which are chosen at random during the training process. These are not considered during a particular forward or backward pass. The main purpose of Drop-out is to prevent Over-Fitting. The Deep Learning model reduces the interdependent learning among neurons, called Over-Fitting, by adding penalty to the loss function. Regularization is a way to reduce Over-Fitting.

F. Optimization Algorithm
These algorithms [3] are used to calculate the optimal values for internal parameters like weights, bias etc., for a Deep Learning Model. The main purpose of these algorithms is to minimize the Objective Function or Error Function. The Error function is dependent on the internal learnable parameters as mentioned. There are two major types of Optimization algorithms in use.

- First Order Optimization Algorithms
- Second Order Optimization Algorithms

These algorithms minimize the loss function using their gradient values. The first order derivative indicates, whether the loss function is increasing or decreasing. A Gradient is generally represented by a Jacobian Matrix [5]. These algorithms are easy to compute and leads to faster convergence even on large datasets.

- Second Order Optimization Algorithms:
These algorithms uses Hessian Matrix or Second order derivative to minimize the loss function. They are costly and slower in execution, as they need complex computation of second order derivative. These algorithms have a better edge over First Order Algorithms in a way that the Second Order Algorithms indicate the increasing or decreasing functional properties and curvature of a surface.

Some commonly used Optimization algorithms used are

- Stochastic Gradient Descent
- Mini Batch Gradient Descent
- Adam
- RMSprop

Fine-Tuning [6][7] a Convolutional Neural Network [8][9][10] Model is a process of changing the Hyper-Parameters and analyzing the prediction accuracy. Various Convolutional Neural Network Models are used for investigation purpose. All the models were simulated by varying the Learning Rate and keeping other Hyper Parameters intact. The optimum band of learning rate is identified, at which maximum accuracy zone of Object detection is achieved.

II. METHODOLOGY
In this paper, a dataset of 7308 images consisting of animals, plants, food items etc. are taken for training purpose. Also, a separate dataset of 966 images are considered for testing purpose. These datasets consists of 70 classes of testing and training images.

Numerous studies and experimentations are carried out in trial and error basis for freezing the best suitable indices for hyper-parameters for the considered dataset. Table-1 is a collection of values of various hyper-parameters which are fixed in this paper for analysis purpose.

| S.No | Hyper-Parameters | Values   |
|------|------------------|----------|
| 1    | Training batch size | 100      |
| 2    | Validation batch size | 40       |
| 3    | Drop-out rate      | 0.5      |
| 4    | Activation Function | Relu [13] |
| 5    | Number of Epochs   | 10       |
| 6    | Training Layers    | Top-4 Layers |

In this study, the number of epochs are limited to ‘10’, because there is no much variation observed in prediction accuracy with respect to increment in Epochs.

III. EXPERIMENTAL RESULTS
Object Detection Accuracies and number of Wrong Predictions of each Deep Learning Model are analyzed by varying ‘Learning rates’ in the range $10^{-6}$ to $10^{-3}$. The following are the abbreviations used for representation of parameters in the tables defined below:

- WP: Number of Wrong Predictions (Out of 966 Testing images)
- LR: Learning-rate
- \%PA: Percentage of Prediction Accuracy

A. Visual Geometry Group16 (VGG16):
The experimental results of 966 testing images shows a variation in Object detection accuracy as the learning rate changes from $10^{-6}$ to $10^{-3}$. Also, the results indicate that, as the percentage accuracy increases the number of wrong predictions decreases.

The experimental results are tabulated in Table-2 and it is observed that the number of wrong predictions are minimum at a ‘Learning-rate=10^{-3}’, and a maximum Object Detection accuracy of 92.45% is achieved. It also surpassed the Top-1 accuracy of 70.5%, as per the available literature [14].

Table 2: VGG16 Accuracy for various Learning-rates

| S.No | LR   | \%PA | WP  |
|------|------|------|-----|
| 1    | $10^{-1}$ | 2.07 | 946 |
| 2    | $10^{-2}$ | 1.04 | 956 |
| 3    | $10^{-3}$ | 1.04 | 956 |
| 4    | $10^{-4}$ | 92.45 | 77  |
| 5    | $10^{-5}$ | 70.50 | 285 |
| 6    | $10^{-6}$ | 70.50 | 285 |

B. Visual Geometry Group19 (VGG19):
VGG19 model is executed at various Learning-rates, and from the results it is identified that the number of wrong predictions are minimum at a ‘Learning rate=10^{-3}’ and the maximum Object Detection accuracy of 92.96% is achieved. It also surpassed the Top-1 accuracy of 75.2% available in literature [14]. The results are tabulated in Table-3.
Table 3: VGG19 Accuracy for various Learning-rates

| S.No | LR  | %PA  | WP  |
|------|-----|------|-----|
| 1.   | 10^-1 | 1.04 | 956 |
| 2.   | 10^-2 | 1.04 | 956 |
| 3.   | 10^-3 | 1.04 | 956 |
| 4.   | 10^-4 | 92.96 | 68 |
| 5.   | 10^-5 | 67.39 | 315 |
| 6.   | 10^-6 | 7.66  | 892 |

C. InceptionV3:
InceptionV3 is an advanced model as compared to VGG16 and VGG19. This model has more number of hidden layers as compared to the previous models. The Object Detection Accuracy is tested for this model for the considered dataset. The output results indicate a better accuracy at a ‘Learning-rate=10^-5’ and the number of wrong predictions are also found to be minimum. Table-4 indicates the experimental results of Inception V3 model.

The experimental results indicates a maximum Object Detection accuracy of 93.69% and the number of wrong predictions are only 61 out of 966 images. These results surpasses the Top-1 accuracy of 78.8% as per the available literature [15].

Table 4: InceptionV3 Accuracy for various Learning-rates

| S.No | LR  | %PA  | WP  |
|------|-----|------|-----|
| 1.   | 10^-1 | 2.07 | 946 |
| 2.   | 10^-2 | 1.86 | 948 |
| 3.   | 10^-3 | 1.86 | 948 |
| 4.   | 10^-4 | 91.90 | 78 |
| 5.   | 10^-5 | 93.69 | 61 |
| 6.   | 10^-6 | 84.06 | 154 |

D. Xception:
Xception is the most advanced Object Detection Model out of the four models taken for analysis. The results of this model are far better and the percentage accuracy is at power with other models.

The experiment is carried out by considering the dataset of 966 images and the number of wrong prediction are only 27. This leads to the Object Detection accuracy as high as 97.65% at a ‘Learning-rate=10^-4’. The present experiment clearly surpasses the Top-1 accuracy of 79% mentioned in literature [16].

Table 5: Xception Accuracy for various Learning-rates

| S.No | LR  | %PA  | WP  |
|------|-----|------|-----|
| 1.   | 10^-1 | 1.04 | 956 |
| 2.   | 10^-2 | 1.04 | 956 |
| 3.   | 10^-3 | 28.78 | 689 |
| 4.   | 10^-4 | 97.65 | 27 |
| 5.   | 10^-5 | 95.47 | 48 |
| 6.   | 10^-6 | 91.31 | 88 |
Optimal Accuracy Zone Identification in Object Detection Technique - A Learning Rate Methodology

Figure 4: Xception WP by $10^{-4}$ Learning-rate

IV. DISCUSSIONS

It can be seen from the tabulated results that the accuracy is maximum at learning rate ranges from $10^{-4}$ to $10^{-5}$ in the considered networks.

A. VGG16:
The variation of Accuracy with respect to Learning-rate for VGG-16 Model is shown in Figure-5. Maximum accuracy of 92.45% is achieved at a Learning-rate=$10^{-4}$.

Figure 5: VGG16 Accuracy at different Learning rates.

B. VGG19
The output result of VGG-19 Model is shown in Figure-7. This figure indicates the variation of Accuracy with respect to Learning-rate in VGG-19 Model. The maximum accuracy obtained is 92.96%, at a Learning-rate=$10^{-4}$. The number of objects wrongly predicted with respect to various Learning-rates is indicated in Figure-8. From the figure, it is clear that the Number of wrong predictions reduced considerably to a minimum value of 68 at $10^{-4}$ Learning-rate and it is the minimum value of WP that could be achieved with this model from the considered data set.

Figure 6: VGG16 WP at different Learning rates

Figure 7: VGG19 Accuracy at different Learning rates

Figure 8: VGG19 WP at different Learning rates
C. InceptionV3

Figure 9 shows the rate of change of Accuracy with respect to Learning-rate in InceptionV3 Model. Highest accuracy is obtained at a Learning-rate of $10^{-5}$ for the considered dataset.

![InceptionV3: Learning Rate Vs Accuracy](image)

Figure 9: InceptionV3 Accuracy at different Learning rates

![InceptionV3: Learning Rate Vs Wrong predictions](image)

Figure 10: InceptionV3 WP at different Learning rates

Figure 10 shows the number of objects wrongly predicted in InceptionV3 model with respect to various Learning-rates. From this figure, it is clear that the Number of wrong predictions are only 61 out of 966 at $10^{-5}$ Learning-rate. For the given dataset, it is the minimum value of WP that could be achieved with this model.

D. Xception:

The Accuracy trend with respect to Learning-rate in Xception model is shown in Figure 11. As per the trend curve, it is observed that the accuracy reached a peak value at a Learning-rate of $10^{-4}$.

![Xception: Learning Rate Vs Accuracy](image)

Figure 11: Xception Accuracy at different Learning rates

The number of objects wrongly predicted by the Xception model with respect to various Learning-rates is depicted in Figure 12. From this figure, it is clear that the Number of wrong predictions reduced considerably to a minimum value of 27 at $10^{-4}$ Learning-rate and it is the minimum value of WP that could be achieved with this model.

![EXCEPTION-Learning Rate Vs Wrong predictions](image)

Figure 12: Xception WP at different Learning rates

![Learning Rate Vs Wrong predictions](image)

Figure 13: WP of all Deep Learning Models at different Learning-rates

Figure 13 is a plot of number of wrong predictions of all Deep Learning Models with respect to Learning-rates. The plot clearly shows that all the considered models have hardly any wrong predictions (WP) obtained near $10^{-4}$ and $10^{-5}$ learning-rates. The corresponding Accuracies of all Deep Learning Models with respect to the Learning-rates are shown in Figure 14. As per this figure, the highest accuracy zone is identified in the range of $10^{-4}$ to $10^{-5}$ learning-rate. It is also observed that the accuracy goes in a decreasing trend once it reaches a peak value, after a learning rate of $10^{-5}$.
surpassed the Top-1 accuracies in literature. Hence it can be concluded that the Optimal Accuracy Zone is identified at a learning rate ranging from $10^{-4}$ to $10^{-5}$ by keeping the other Hyper-Parameters constant.

VI. FUTURE SCOPE

To achieve further improved accuracy, the best values of all hyper-parameter needs to be calculated and analyzed by rigorous Fine-Tuning process. The relationship between Hyper-Parameters and Prediction accuracies needs to be formulated based on the size of dataset, which simplifies the Object Detection techniques, without going through the regular iterative methods.

VII. ACKNOWLEDGEMENT

This work is done under the scholarship of Visvesvaraya PhD Scheme for Electronics and IT, Government of India. Sincere thanks to CS&SE Dept. and Oceanography and Meteorology Dept. of Andhra University for providing the laboratory facility.

REFERENCES

1. Chamarty Anusha, P. S Avadhani, “Object Detection Using Deep Learning”, in International Journal of Computer Applications, vol. 182(32), Dec. 2018, pp. 18-22.
2. YannLeCun, YoshuaBengio, Geoffrey Hinton, “Deep Learning” in Nature, vol. 521, no. 7553, May 2015, pp. 436-444.
3. Ross Girshick, Jeff Donahue, Trevor Darrell, Jotendra Malik, “ Region based Convolutional Networks for Accurate Object Detection and Segmentation”, in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, 2016, pp. 142-158.
4. IgiirChinwe Peace, Anyama Oscar Uzoma, Silas Abasiamalta, “Effect of Learning Rate on Artificial Neural Network in Machine Learning”, in International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 4, Issue 02, Feb-2015.
5. Jian-Xun Peng, Kang Li, George W Irwin, “A New Jacobian Matrix for Optimal Learning of Single-Layer Neural Networks”, in IEEE Transactions on Neural Networks, Vol. 19, Issue 1, Jan. 2008, pp. 119-129.
6. Chamarty Anusha, P. S Avadhani, “Fine-Tuning Convolutional Neural Network Models for improvement of Object Detection Accuracy”, in International Journal of Engineering and Advanced Technology, ISSN: 2249-8958, Vol-8 Issue-5, June-2019.
7. Xuesong Zhang, Fei Yan, Yan Zhuang, Huosheng Hu, Chunguang Bu, “Using an Ensemble of Incrementally Fine-Tuned CNNs for Cross-Domain Object Category Recognition”, in IEEE Access, vol. 7, March 2019, pp. 33822-33833.
8. Zhong-Qui Zhao, Peng Zhong, Shou-tao Xu, Xindong Wu, “Object Detection with Deep Learning: A Review”, in arXiv preprint arXiv: 1807.05511. Apr 2019.
9. Juuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Gang Wang, Jianfei Cai, Tszhan Chen, “Recent Advances in Convolutional Neural Networks”, in Pattern Recognition, vol. 77, May 2018, pp. 354-377.
10. Ross Girshick, “Fast R-CNN”, in arXiv preprint arXiv: 1504.08083v2, sept. 2015.
11. D.R. Wilson, T.R. Martinez, “The need for small learning rates on large problems”, in IEEE Conference on International Joint Conference on Neural Networks, july 2001.
12. Hyung Jin Son, T B Trafails, M B Richman, “Determination of the optimal batch size in incremental approaches: an application to tornado detection”, in IEEE International Joint Conference on Neural Networks, 2005.
13. Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, “ImageNet Classification with Deep Convolutional Neural Network”, in the conference on Neural Information Processing Systems, 2012.
14. K. Simonyan and A. Zisserman “Very Deep Convolutional Networks for Large-Scale Image Recognition”, in International Conference on Learning Representations 2015.
15. Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke and Andrew Rabinovich, “Going Deeper with Convolutions”, in the Conference on Computer Vision and Pattern Recognition 2015.

16. Francois Chollet, “Deep Learning with Depthwise Separable Convolutions”, in the Conference on Computer Vision and Pattern Recognition, 2017.

17. Matthew D. Zeiler, Rob Fergus, “Visualizing and Understanding Convolutional Networks”, in arXiv preprint arXiv: 1311.2901v3, Nov. 2013.

AUTHOR PROFILE

Mrs. Chamarty Anusha has received B. Tech degree from Jawaharlal Nehru Technological University, Andhra Pradesh and M.Tech from Andhra University, Visakhapatnam. The author is currently pursuing her Ph.D under Visvesvaraya Ph.D Scheme, Government of India from the Department of Computer Science and Systems Engineering, Andhra University, Visakhapatnam.

Her main research work focuses on Object Detection using Deep Learning, Cyber Security and Machine Learning. She has published some research papers in various International Journals and attended numerous workshops and conferences during her career. The author has five years of experience in teaching field and dealt with various subjects in Computer Science like Computer Networks, Data Structures, Database designing etc.

Prof. P S Avadhani pursued M.Tech and Ph.D from IIT Kanpur. He is currently working as a Professor in the Department of Computer Science and Systems Engineering Andhra University, Visakhapatnam. His areas of interest are Cyber Security, Fuzzy Logic, Cyber Forensics, Computer Algorithms, Public Cryptography, and Data and Network Security. The author has written 6 books in Computer Science Engineering on various subjects. The author has published 58 papers in International Journals, 8 papers in National Journals, 38 papers in International Conferences and 9 papers in National Conferences. He has received several awards for his excellence in Teaching, Research and Administration like “State Best Teacher award”, “Best Researcher award”, “Distinguished Principal award” from the Government of Andhra Pradesh. The author had trained various PhD, M.Tech & B.Tech Engineering students during his 35 years of long teaching career.