Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

Multi-class Image Classification Using Deep Learning Algorithm

To cite this article: W A Ezat *et al* 2020 *J. Phys.: Conf. Ser.* **1447** 012021

Multi-class Image Classification Using Deep Learning Algorithm

W A Ezat[1], M M Dessouky[2] and N A Ismail[3]

Department of Computer Science and Engineering, Faculty of Electronic Engineering
Menoufia University, Menouf, Menoufia, Egypt

Email: [1] waelezat2@hotmail.com
Email: [2] mohamed.moawad@el-eng.menofia.edu.eg
Email: [3] nabil_is@hotmail.com

Abstract. Classifying images is a complex problem in the field of computer vision. The deep learning algorithm is a computerized model simulates the human brain functions and operations. Training the deep learning model is a costly process in machine resources and time. Investigating the performance of the deep learning algorithm is mostly needed. The convolutional neural network (CNN) is most commonly used to build a structure of the deep learning models. In this paper convolutional neural network (CNN) model pre-trained on Image-Net is used for classification of images of the PASCAL VOC 2007 data-set. The transfer learning approach is used to improve the performance of the deep learning CNN model where classification works fairly well with the smallest amount of computation time and fewer machine resources. The behavior of the Deep learning CNN model is studied and the performance has been measured. The obtained results are compared with the obtained test results from the Super-vector coding of local image descriptors method, SVM method, and Region Ranking SVM method, which tested with the PASCAL VOC 2007 data-set. The final results evaluate the deep learning algorithm as a state-of-the-art method for an image classification task.

1. Introduction

Image classification using deep learning algorithm is considered the state-of-the-art in computer vision researches. In the deep learning algorithm, the object feature extracted engineering is done by the algorithm automatically. Engineering the object feature extraction is a difficult process and time-consuming. Operations of the feature extraction process required a domain expert operator for design and testing. The deep learning algorithm solves wide problems of classification tasks. The ability to process large clusters of images quickly, state the deep learning algorithm as the most important method for images classification. The ability and flexibility for changeable in the deep learning model with a wide range of data-sets make the deep learning algorithm as the most important technique for the classification task. The most difficult problem in the deep learning algorithm is the cost of hardware and consuming time in training processes. Training the deep learning model may need weeks with costly GPU machine. Performing practical research focused on the performance of the deep learning algorithm is the most important to evaluate the deep learning method. This study points out the success of the deep learning algorithm in solving of multi-class image classification task problem. This paper, studied the performance of the deep learning convolutional neural network CNN model which pre-trained on large data-set ILSVRC and test it on other smaller data-set PASCAL VOC 2007 [1].

In the experiment, the performance of the deep learning CNN model is tested with four classes from the PASCAL VOC 2007 data-set before training cycle [2-4]. The model weights are adjusted in the training process by the transfer learning [5]. The performance of the deep learning CNN model is measured and the obtained results are compared with the super-vector coding of local image descriptors method [6], SVM method [7] and region ranking SVM method [8].

The paper has been approved that performance of the deep learning CNN model is better than super – vector coding of local image descriptors method in the image classification task and nearly equal performance of the SVM method and region ranking SVM method. The resulted out approved that the deep learning algorithm is the future choice to solve a wide range of the image classification task problem easily with minimum cost.

The contribution of this paper is

a) Comparison of the performance of the deep learning algorithm for multi-class image classification with the performance of other classification algorithms and methods.
b) Demonstrate simplicity of implementing the deep learning algorithm

2. The data-set

For the task of training and validation for classification considered four classes from the PASCAL VOC 2007 data-set. Many computer researchers and scientific papers use PASCAL 2007 as data-set for training and validation. PASCAL 2007 data-set is small-scale with 20 object classes. The PASCAL 2007 data-set has text files label every image as the image name and image class. All images are JPG format [9].

3. The experiment

3.1 Environment

Training the deep learning model carried out on the Amazon EC 2 CPU instance called t2.large running a Bitfusion boost Ubuntu 14 Caffe-2016-07-operating system [AWS].

3.2 Experiment producers

The experiment is carried out on the CNN CAFFE [10] Image-net model using the data-set PASCAL VOC 2007 and with four different classes and (702 images) from these four classes.

a) In the local machine with python scripting code a descriptors text files of the data-set have been read and converted into a comma-separated values files CSV, The CSV files generated contains the images numbers and images classes and labels.

b) A CSV file has been created to contain an index for all images from classes [Bird - Cat- Train – TV-monitor] and (702 images) have been chosen randomly.

c) With python scripting code three CSV files have been generated to index the chosen images and partition images randomly into three equal subset [G1 – G2 – G3] contain the images name, image classes, and images file number.

d) A generated CSV files have been uploaded to remote machine AWS.

e) JPG images of the data-set have been downloaded to the remote machine.

f) A python scripting code in the AWS remote machine has been used to separate JPG images files of subset [G1-G2-G3], separated them in independent three folders at the workspace storage of the AWS remote machine and labeling every image to every class belong to.

g) The cross-validation [11] technique has been used in training and validation. In the first training and validation cycle, the subsets G1 and G2 have been used for training the model and the subset G3 has been used for the validation as shown in Figure 1.

h) In the second training and validation cycle subsets G1 and G3 have been used for training and the subset G2 are used for the validation process as shown in Figure 2. In the third training and validation cycle, subsets G2 and G3 have been used for the training process and subset G1 for the validation process as shown in Figure 3.

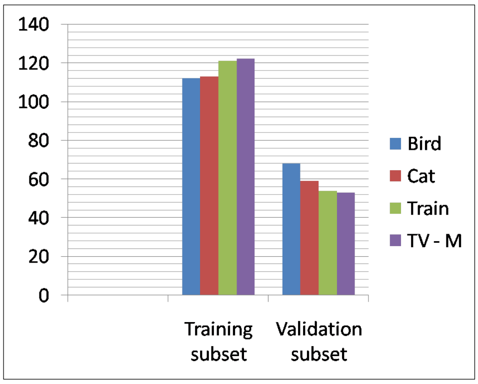

Figure 1. Subsets images number over classes for training cycle-1

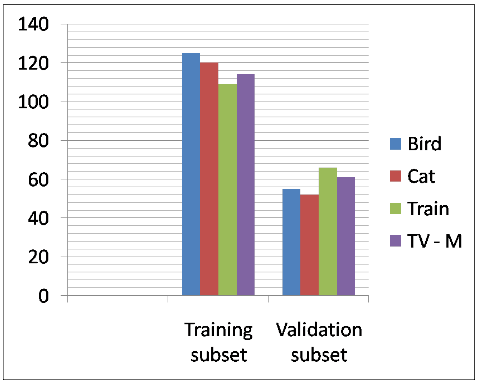

Figure 2. Subsets images number over classes for training cycle-2

4. Experimental results

The following experimental results are obtained by deep learning framework CAFFE. The model has been trained and accuracy from the three training cycles is the accuracy of (87 %, 93 %, and 92.7 %) with an average accuracy value of 90.9 %.
To compare accuracy before training the model and after training process cycles as in Figure 4 and Figure 5.
The average computational time for every training and validation cycle is 15 hours.

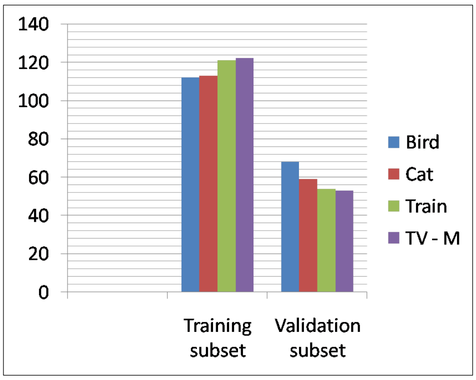

Figure 3. Subsets images number over classes for training cycle-3

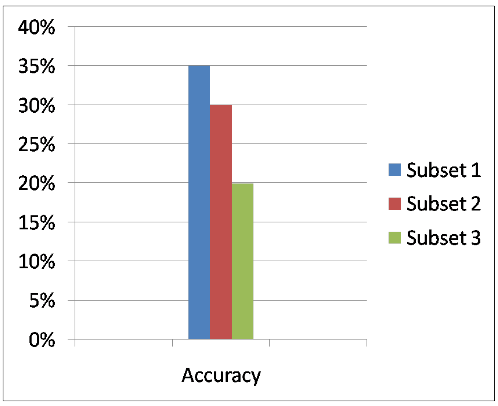

Figure 4. Accuracy before the training

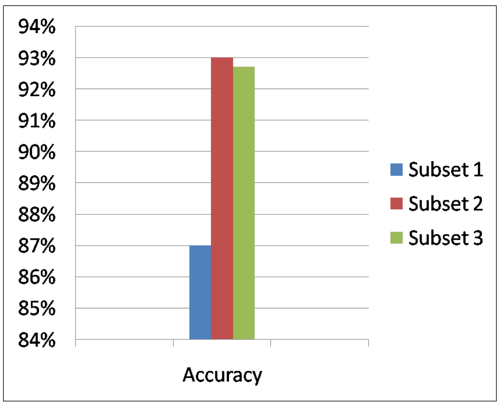

Figure 5. Accuracy after the training

The following experimental results depicted in Tables 1, 2, 3 and 4. are obtained on the three training and testing cycles by evaluating the classification accuracy, measured by average precision using python script code and based on regular mathematical terms.

Table 1. The deep learning CNN precision performers in PASCAL VOC 2007

P (%)	subset 1	subset 2	subset 3	average
Bird	90	94	89	91
Cat	82	90	94	88.6
Train	94	98	90	94
TV	84	90	95	89.66

Table 2. The deep learning CNN error rate performers in PASCAL VOC 2007

ER (%)	subset 1	subset 2	subset 3	average
Bird	7.2	2.9	4.7	4.9
Cat	8.9	5.1	5.1	6.3
Train	2.1	.4	2.9	1.8
TV	5.5	4.2	2.5	4

Table 3. The deep learning CNN accuracy performers in PASCAL VOC 2007

AC (%)	subset 1	subset 2	subset 3	average
Bird	92	97	95	94.66
Cat	90	94	94	92.66
Train	97	99	97	97.66
TV	94	95	97	95.3

Table 4. The deep learning CNN performers in PASCAL VOC 2007

Parameter	Rate %
Precision	90.81
Error Rate	4.25
Accuracy	95.07

Experiment results are compared with Super–vector coding of local image descriptors method, SVM method and RRSVM method as in Table 5. The deep learning CNN model has more accuracy than Super–vector coding of local image descriptor, and approximately accuracy value of SVM and RRSVM methods as shown in Figure 6 and Figure 7.

Table 5. Comparison of the deep learning CNN model with SVC, SVM and RRSVM methods in PASCAL VOC 2007 data-set

	Bird	Cat	Train	TV/Monitor	AP (%)
D.L	88.6	91	94	89.66	90.81
SVC [6]	55.6	63.3	83.3	58.9	65.27
SVM [7]	97.1	96.1	97.4	89.9	95.1
RRSVM [8]	98.1	97.4	98.6	93.1	96.8

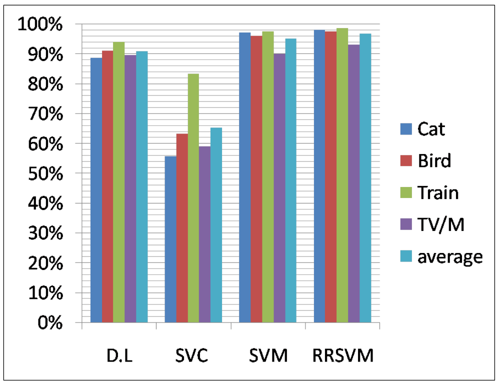

Figure 6. Precision rate over classes

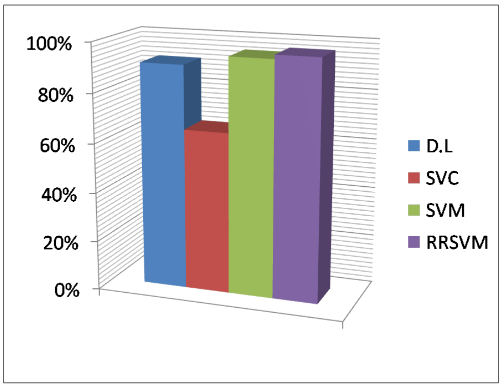

Figure 7. Average precision rate over classification methods

5. Conclusion

The results of the experimental work has been evidenced that the deep learning algorithm CNN model is a simple and effective solution for the multi-class images classification task. The deep learning algorithm CNN model has the ability to flexible change regarding a requirement with a wide verity of data-sets. As seen in local machine and with python scripts, a simple database for training data has been made and uploaded to a powerful machine for the training process. As pointed out the deep learning algorithm CNN model has been trained and the model weight has been fine-tuned. The model has been adjusted for the new data-set. With a few training samples the accuracy of the model has been increased. As pointed the CNN deep learning algorithm in multi-class image classification task have the more accurate value than Super–vector coding of local image descriptors method and accuracy nearly equal accuracy of RRSVM method.

6. Future work

The experiment has been done with four classes due to lack of time. Future work concerns of particular training and test the deep learning model with more number of the data-set classes. Proposals for apply the deep learning algorithm to solve problem of classification Egyptian vegetables and fruits.

References

[1] Cengil, E., Cinar, A., & Ozbay, E. (2017). Image classification with caffe deep learning framework. 2017 International Conference on Computer Science and Engineering (UBMK). http://doi.org/10.1109/ubmk.2017.8093433.

[2] CS231n: Convolutional Neural Networks for Visual Recognition. Retrieved May 21, 2019, from http://cs231n.stanford.edu/.

[3] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Rabinovich, A. (2015). Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). http://doi.org/10.1109/cvpr.2015.7298594.

[4] JT Turner, David Aha, Leslie Smith, & Kalyan Moy Gupta. Convolutional Architecture Exploration for Action Recognition and Image Classification (publication). Convolutional Architecture Exploration for Action Recognition and Image Classification. Knexus Research Corporation. Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/a624847.pdf.

[5] Gupta, D., & Dishashree. (2019, June 25). Master Transfer learning by using Pre-trained Models in Deep Learning. Retrieved July 23, 2019, from https://www.analyticsvidhya.com/blog/2017/06/transfer-learning-the-art-of-fine-tuning-a-pre-trained-model.

[6] Zhou, X., Yu, K., Zhang, T., & Huang, T. S. (2010). Image Classification Using Super-Vector Coding of Local Image Descriptors. Computer Vision – ECCV 2010 Lecture Notes in Computer Science, 141–154. http://doi.org/10.1007/978-3-642-15555-0_11.

[7] Akata, Z., Perronnin, F., Harchaoui, Z., & Schmid, C. (2014). Good Practice in Large-Scale Learning for Image Classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(3), 507–520. http://doi.org/10.1109/tpami.2013.146.

[8] Wei, Z., & Hoai, M. (2016). Region Ranking SVM for Image Classification. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). http://doi.org/10.1109/cvpr.2016.326.

[9] "The PASCAL Visual Object Classes Challenge 2007." The PASCAL Visual Object Classes Challenge 2007 (VOC2007). Retrieved May 21, 2019, from http://host.robots.ox.ac.uk/pascal/VOC/voc2007/index.html.

[10] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., … Darrell, T. (2014). Caffe. Proceedings of the ACM International Conference on Multimedia - MM 14. http://doi.org/10.1145/2647868.2654889.

[11] Adi Bronshtein. (2017, May 17). Train/Test Split and Cross Validation in Python. Retrieved May 21, 2019, from https://towardsdatascience.com/train-test-split-and-cross-validation-in-python-80b61beca4b6.