Global Transitions Proceedings 2 (2021) 382–385

Contents lists available at ScienceDirect

Global Transitions Proceedings

journal homepage: http://www.keaipublishing.com/en/journals/global-transitions-proceedings/

Image classification and prediction using transfer learning in colab notebook

J Praveen Gujjar[a,*], H R Prasanna Kumar[b], Niranjan N. Chiplunkar[c]

[a] CMS Business School, Jain (Deemed-to-be University), Bengaluru, Karnataka, India
[b] PES Institute of Technology and Management, Shivamogga, Karnataka, India
[c] Principal, NMAMIT, Nitte, Karnataka, India

ARTICLE INFO

Keywords:
MobileNet
MobileNetV2
ResNet50
Transfer learning
VGG16
VGG19

ABSTRACT

Transfer learning is used to reuse the pre-trained model. Transfer learning uses the knowledge which was gained from the previous task. Transfer learning is most generally used in image classification, image prediction and natural language processing. Some of the example for the natural language processing it includes sentiment analysis, text auto complete etc. The literature shows that deep learning performance is relatively more when compared with machine learning technique for the large data set. In this paper pre trained models such as MobileNet, MobileNetV2, VGG16, VGG19 and ResNet50 has been used for image classification and prediction. For the image classification and prediction Google Colab notebook has been used. The performance of the system depends on the GPU system hence results are tested in Google colab notebook. The result shows that MobileNetV2 performance is relatively better than other pre trained models. MobileNetV2 uses the less number of parameters as compared with other trained model. ResNet50 accuracy is more when it compared with other trained model with the ImageNet dataset. In the future enhancement transfer learning may be used for natural language processing to obtain highest accuracy.

1. Introduction

Artificial intelligence is becoming a buzzword in the current scenario. Machine learning is the subset of artificial intelligence and it deals about the statistical tools to carry out the real time application. Deep learning is a subset of machine learning. The performance of the deep learning is more when it compared with the machine learning for the large data set [1]. Widespread use of machine learning and deep learning artificial intelligence is becoming more popular. Using the data set machine learning and deep learning predicts the new instance of the data. Deep learning is also referred to as deep neural network. Convolution neural network eliminated the manual interventions of feature extraction. In convolution neural network input is called as feature vector [1]. To detect the pattern, it is going to use filter or detector and the resultant vector is called feature map. Automated feature extraction gives accurate learning models for computer vision tasks. Deep learning models neural network extract, the features without programmer interventions [2]. Transfer learning is used to reuse the pre trained model for image classification, image prediction and natural language processing. The paper is organized in four parts. Part 1 is the introduction; Part 2 Convolutional neural network; Part 3 Transfer learning and related work; Part 4 presents the Results and discussion; Part 5 presents conclusions. References are given after Part 5.

2. Related works

The model trains with a large volume of data. This model is used to test new data set [2]. Transfer learning is used to leveraging the knowledge. Leveraging knowledge takes place from the source task. Leveraging the knowledge is considered as one of the main goal of transfer learning. Transfer learning can improve learning by three important measures. First learning measure is transferred knowledge. Second is the transferred knowledge how much time it takes to learn. Third measure is comparing the performance level without transfer.

VGG is an acronym for Visual Geometry Group [3]. The transfer learning is used to compare the performance VGG19 and AlexNet [4]. VGG16 and VGG19 use 3×3 filters. Transfer learning is used to classify the different dog breeds [5]. In order to enhance the accuracy for image identification combined a convolutional neural network with Adaboost algorithm [6]. The transfer learning is used to detect the object in the image and also it is used for image classification [7,8]. Deep learning model was proposed using an auto encoder for MNIST data sets [9]. DCNN autonomous learning algorithm was proposed based on the genetic algorithm to improvised the accuracy [10]. Deep learning model is

* Corresponding author.
E-mail address: dr.praveengujjar@cms.ac.in (J. Praveen Gujjar).

https://doi.org/10.1016/j.gltp.2021.08.068
Received 8 June 2021; Accepted 7 July 2021
Available online 12 August 2021

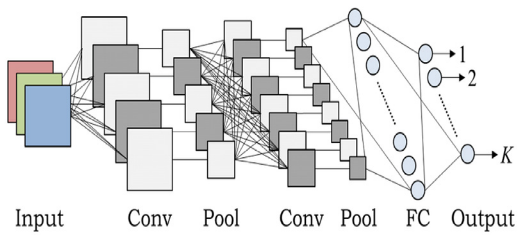

Fig. 1. CNN General Architecture.

better than the machine learning model in accuracy for the large data set [11]. Deep learning and machine learning was compared; it was shown that deep learning performance is increased while machine learning model gives constant performance [13]. Sentimental analysis places a very important role in the decision making of the product by the customer [12]. Analysis of the financial data can be done using latest technology [14]. Active learning can be used for image classification. Active learning gives the better result when it compared with other techniques for the image classification [15]. Transfer learning model was compared GoogleNet and AlexNet, transfer learning gives the better result for image classification [16]. [17] TCA is a acronym for Transfer component analysis which is considered as fault diagnosis. TCA shows that algorithm performance is gradually increasing [18]. Inductive learning performance is relatively high. Predictive model which comes from the inductive bias can make accurate predictions. Hence, it has been urge that to have generalization capability it is better to have inductive bias [19-21]. Transfer learning can be used in the natural language processing to analyze the sentiment of the user statement [22-24].

3. Methods

Convolutional neural network (CNN) consists of input layer and out layer. In between these two layer there is a convolutional layer, pooling layers and fully connected layer. Additional layer may be used for the complex set of models. Matrix vector multiplication is at the heart of how data and weights are represented [1]. CNN general architecture is as shown in the fig 1.

CNN matches the portion of the input image (feature) rather than the whole image. A 3×3 grid is used to represent the feature extraction by the CNN. After completing the convolution process max polling was invoked which is going to shrink the image stack. Window size is defined for the max poling along with the stride. Then the window is filtered across the image in stride. The max pooling reduces the dimensionality of each feature map. The normalization process in convolutional neural network is usually carried out by rectified linear unit (ReLU). ReLU changes all the negative value in the filtered image into zero. This ReLU process is increases the non-linearity property of the model. In the next layer fully connected layer was observed in the convolutional neural network which is also referred as classifier. In this paper image classification and predictions using transfer learning are done in the Google. Google Colaboratory is also called as Colab it is a cloud based service. Colab is based on the Jupyter Notebook for doing machine learning and deep learning operation. The Google Colab provides a free of costs access GPU in the runtime. The Google colaboratory is useful for computer vision and other GPU centric applications. Mobilenet and MobilenetV2 architecture is as shown below

The general architecture of MobileNet is shown in the Fig 2 indicates that first layer is depthwise convolution which is used for lightweight filtering which applies a single convolutional filter per input channel. In the same way second layer is of 1×1 convolutions which are also called as pointwise convolution. Second layer in the MobileNet is used for creating new features by linear combination computing of the input channels. The activation function ReLU is used because of its robustness with low precision computation. MobileNetV2 is making use of CNN architecture so that it can perform well on mobile devices. MobileNetV2 uses residual structure as shown in Fig 3. Further, MobileNetV2 is having the bottleneck layers in the residual connections. Lightweight depthwise convolutions are used by intermediate expansion layer to filter features as the source of non linearity.

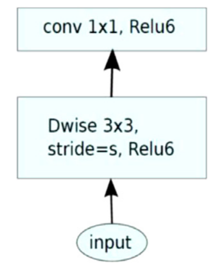

Fig. 2. MobileNet Architecture.

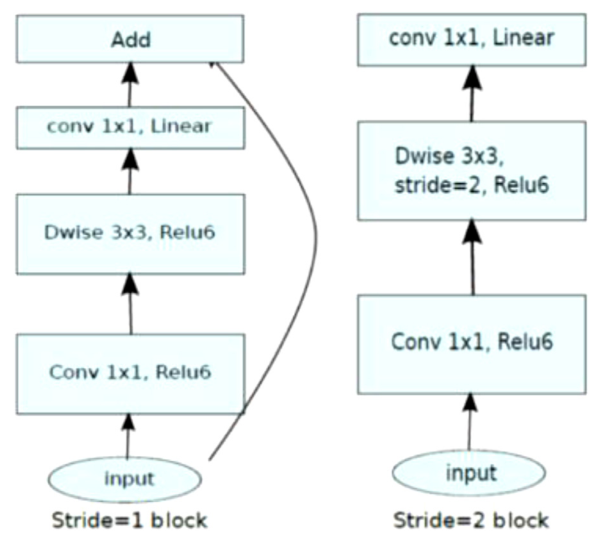

Fig. 3. MobileNetV2 Architecture.

MobileNetV2 is having the 32 filters initial fully connected convolution layer followed by 19 residual bottleneck layers. Numbers in VGG 16 and VGG 19 indicates that there are total 16 and 19 layers are available respectively. The architecture for VGG 16 and VGG 19 is as shown in the Fig 4 and Fig 5 respectively [25-26].

Convolution layer is used in the first two layers with 3×3 filters, and in the these two convolution layer uses 64 filters so that it ended up in $224 \times 224 \times 64$. (CONV64) $\times$ 2 represents it uses 2conv layers with 64 filters. The filters use 3×3 with stride of 1 and implemented with the same convolutions. Pooling layer in this architecture which will reduces height and width of a volume $224 \times 224 \times 64$ down to $112 \times 112 \times 64$. Then couple more convolutional layers used with 128 filters. Then, a pooling layer is added so new dimension will be $56 \times 56 \times 128$. Then 2conv layers with 256 filters used. And Finally$7 \times 7 \times 512$ into Fullyconnected layer (FC) with 4096 units, and in a softmax output one of a 1000 classes obtained.

The architecture for the residual network is as shown in Fig 6. In Artificial neural network a special kind of network has been created which is known as residual neural network. Skip connections are widely utilized by residual network. Skip connects is also called as shortcut jump over layer.

4. Results and discussion

In this paper the five transfer learning model was used. Namely MobileNet, MobileNetV2, VGG16, VGG19, ResNet50. The

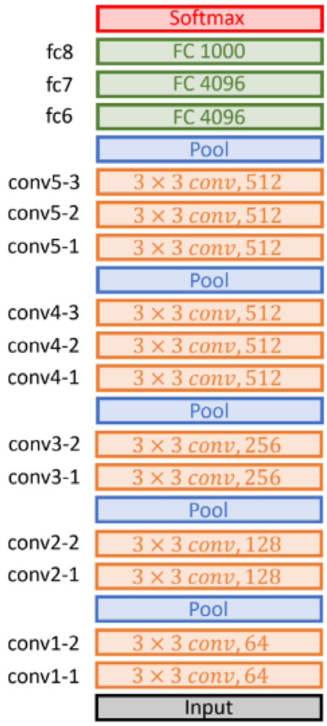

Fig. 4. VGG16 Architecture.

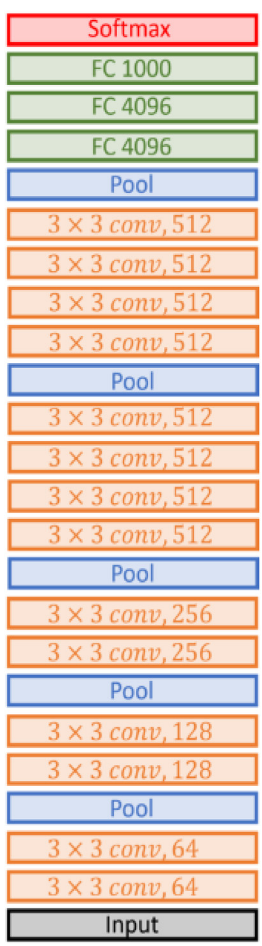

Fig. 5. VGG19 Architecture.

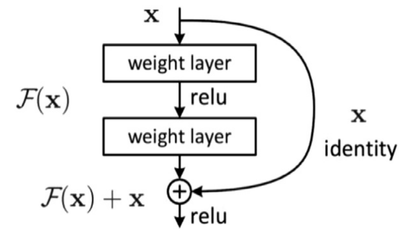

Fig. 6. ResNet Architecture.

Table 1
Transfer learning state-of-the-art model.

Sl.No	State-of-the-art Model	Disk Space	Parameters
1.	MobileNet	16 MB	4,253,864
2.	MobileNetV2	14 MB	3,538,984
3.	VGG16	528 MB	138,357,544
4.	VGG19	549 MB	143,667,240
5.	ResNet50	98 MB	25,636,712

MobileNet uses the concept of depthwise convolution method followed by point wise convolutional method to carry out the process. MobileNet model was invoked using keras framework as keras.applications.mobilenet.MobileNet(). Image for the MobileNet transfer learning has been given as input to predict the object through Google drive in Google Colab. MobileNet uses only 16MB of disk space and accuracy varies between 70 percent to 89.5 percent for the ImageNet data set.

The MobileNetV2 uses the concept of 1×1 expansion layer followed by 3×3 depthwise convolution method followed by 1×1 projection layer to carry out the process. MobileNetV2 model was invoked using keras framework. The keras frame work has been given as tf.keras.applications.mobilenet.preprocess_input(x[tf.newaxis,...]). MobileNetV2 method has been invoked as tf.keras.applications.MobileNetV2(). Image for the MobileNetV2 transfer learning has been given as input to predict the object through Google drive in Google Colab. MobileNetV2 uses only 14MB of disk space and accuracy varies between 70 percent to 91 percent for the ImageNet dataset. The VGG16 uses the concept of 16 layer of convolutional neural network to carry out the task. VGG16 model uses the input target size 224×224. Image for the VGG16 transfer learning has been given as input to predict the object through Google drive in Google Colab. VGG16 uses only 528MB of disk space and accuracy varies between 70 percent to 90 percent for the ImageNet dataset. The VGG19 uses the concept of 19 layer of convolutional neural network to carry out the task. VGG19 model uses the input target size 224×224. Image for the VGG19 transfer learning has been given as input to predict the object through Google drive in Google Colab.VGG19 uses only 549MB of disk space and accuracy varies between 71 percent to 90 percent for the ImageNet data set. The ResNet is also called as Residual network. The ResNet50 uses the concept of short cut or skipping in the convolutional neural network to carry out the task. ResNet50 model uses the input target size 224×224. Image for the ResNet50 transfer learning has been given as input to predict the object through Google drive in Google Colab. Table 1 shows the transfer learning model comparison. Which shows the MobilenetV2 takes only 14 MB of disk space and accuracy is also relatively high. ResNet50 uses only 98MB of disk space and accuracy varies between 75 percent to 92 percent for the ImageNet dataset. MobileNet uses 4,253,864 parameters where as MobileNetV2 uses 3,538,984. VGG16 and VGG19 use 138,357,544 and 143,667,240 number of parameters respectively. Table 2 shows the state of the art comparison for the transfer learning MobileNet, MobileNetV2, VGG16, VGG19 and ResNet50 with the ImageNet dataset. MobileNetV2 uses less number of parameters compared to other transfer learning as

Table 2
Comparison to the state-of-the-art model.

State-of-the-art Model	Dataset	Accuracy (Percentage)
MobileNet	ImageNet	70 to 89.5
MobileNetV2	ImageNet	71 to 90
VGG16	ImageNet	71 to 90
VGG19	ImageNet	71 to 90
ResNet50	ImageNet	74 to 92

shown in Table 1. Further, MobileNetV2 uses relatively less number of Disk space. Hence, MobileNetV2 can be used in the mobile devices.

Table 2 shows MobileNetV2 uses the less number of parameters as compared with other trained model. ResNet50 accuracy is more when it compared with other state of the art trained model with the ImageNet dataset. It has been observed that MobileNet is having 70 to 89.5 percent accuracy. MobileNetV2 is having 71 to 90 percent accuracy. VGG16 and VGG19 are having the 71 to 90 percent accuracy. ResNet50 is having 74 to 92 percent accuracy.

5. Conclusion

Transfer learning is used to reuse the pre-trained model. Transfer learning uses the knowledge which was gained from the previous task. Transfer learning is most generally used in image classification, image prediction and natural language processing. Some of the example for the natural language processing it includes sentiment analysis, text auto complete etc. In this paper, researcher presented image classification and image prediction using transfer learning in Google colaboratory for the ImageNet data set. Transfers learning used in this paper are MobileNet, MobileNetV2, VGG16, VGG19 and ResNet50. For the image classification and prediction Google Colab notebook has been used. The performance of the system depends on the GPU system hence results are tested in Google colab notebook. MobileNetV2 uses the less number of parameters as compared with other trained model. ResNet50 accuracy is more when it compared with other trained model with the ImageNet dataset. The limitation of this paper is that, only transfer learning pertaining to MobileNet, MobileNetV2, VGG16, VGG19 and ResNet50 has been taken for the measuring the transfer learning accuracy. The result shows that MobileNetV2 performance is relatively better than and also it uses less disk space when it compared with other pre trained models. In the future enhancement transfer learning may be used for natural language processing to obtain highest accuracy.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] N. Sharma, V. Jain, A. Mishra, An analysis of convolutional neural networks for image classification, Proc. Comput. Sci. 132 (2018) 377–384.
[2] Fukushima &Kunihiko, Neocognitron: a hierarchical neural network capable of visual pattern recognition, Neural Netw. 1 (2) (1988) 119–130.
[3] Simonyan, Karen, & Andrew Zisserman (1998)," Very deep convolutional networks for large-scale image recognition", arXiv preprint arXiv, pp. 1409.1556.
[4] M. Shaha, M. Pawar, Transfer learning for image classification, in: 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), IEEE, 2018, pp. 656–660.
[5] P. Devikar, Transfer learning for image classification of various dog breeds, Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET) 5 (12) (2016) 2707–2715.
[6] S.J. Lee, T. Chen, L. Yu, C.H. Lai, Image classification based on the boost convolutional neural network, IEEE Access 6 (2018) 12755–12768.
[7] R. Aarthi, S. Harini, A survey of deep convolutional neural network applications in image processing, Int. J. Pure Appl. Math. 118 (7) (2018) 185–190.
[8] G. Ososkov, P. Goncharov, Shallow and deep learning for image classification, Opt. Memory Neural Networks, Springer 26 (4) (2017) 221–248.
[9] E. Maggiori, Y. Tarabalka, G. Charpiat, P. Alliez, High-resolution image classification with convolutional networks, in: Geoscience and Remote Sensing Symposium (IGARSS), 2017 IEEE International, IEEE, 2017, pp. 5157–5160.
[10] Ma, B., & Xia, Y. (2018). “Autonomous Deep Learning: A Genetic DCNN Designer for Image Classification”. arXiv preprint arXiv:1807.00284.
[11] S. Pouyanfar, S. Sadiq, Y. Yan, H. Tian, Y. Tao, M.P. Reyes, S.S. Iyengar, A survey on deep learning: algorithms, techniques, and applications, ACM Comput. Surv. (CSUR) 51 (5) (2018) 92.
[12] S. Loussaief, A. Abdelkrim, Deep learning vs. bag of features in machine learning for image classification”, In 2018, in: International Conference on Advanced Systems and Electric Technologies (IC_ASET), IEEE, 2018, pp. 6–10.
[13] Praveen Gujjar, H R Prasanna Kumar, Sentimental analysis for running text in email conversation, Int. J. Comput. Sci. Eng. (IJCSE) 9 (4) (2020) 67–68.
[14] T Manjunatha, Praveen Gujjar, Performance analysis of indian information technology companies using DuPont model, IUP J. Manag. Res. 17 (Issue 4) (2018) 7–14.
[15] H. Ranganathan, H. Venkateswara, S. Chakraborty, S. Panchanathan, Deep active learning for image classification, in: Image Processing (ICIP), 2017 IEEE International Conference, IEEE, 2017, pp. 3934–3938.
[16] P. Ballester, R.M. de Araújo, On the performance of GoogLeNet and AlexNet applied to sketches, Ass. Adv. Artif. Intell. (AAAI) (2016) 1124–1128.
[17] T. Han, C. Liu, W. Yang, D. Jiang, Deep transfer network with joint distribution adaptation: a new intelligent fault diagnosis framework for industry application, ISA Trans. 97 (2020) 269–281.
[18] B. Yang, Y. Lei, F. Jia, S. Xing, An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings, Mech. Syst. Signal Process. 122 (2019) 692–706.
[19] T.N. Nguyen, V.V. Le, S.I. Chu, B.H. Liu, Y.C. Hsu, Secure localization algorithms against localization attacks in wireless sensor networks, Wireless Personal Commun. (2021) 1–26.
[20] T. Kowsalya, R.G. Babu, B.D. Parameshachari, A. Nayyar, R.M. Mehmood, Low area PRESENT cryptography in FPGA using TRNG-PRNG key generation, CMC-Comput. Mater. Continua 68 (2) (2021) 1447–1465.
[21] Fan Yang, Wenjin Zhang, Laifa Tao, Jian Ma, Transfer learning strategies for deep learning-based PHM algorithms, Appl. Sci. 10 (2020) 2361, doi:10.3390/app10072361.
[22] B.D. Parameshachari, H.T. Panduranga, S. liberata Ullo, Analysis and computation of encryption technique to enhance security of medical images, IOP Conference Series: Materials Science and Engineering, 925, IOP Publishing, 2020, September.
[23] G. Chandrasekaran, T.N. Nguyen, D, J. Hemanth, Multimodal sentimental analysis for social media applications: a comprehensive review, Wiley Interdiscipl. Rev. (2021) e1415.
[24] J Praveen Gujjar, H R Prasanna Kumar, Sentiment analysis:textblob for decision making, Int. J. Sci. Res. Eng. Trends Issue 2 (7) (March-April-2021) 1097–1099.
[25] R.K. Dash, T.N. Nguyen, K. Cengiz, A. Sharma, Fine-tuned support vector regression model for stock predictions, Neural Comput. Appl. (2021) 1–15.
[26] T.G. Nguyen, T.V. Phan, D.T. Hoang, T.N. Nguyen, C. So-In, Efficient SDN-based traffic monitoring in IoT networks with double deep Q-network, in: International Conference on Computational Data and Social Networks, Cham, Springer, 2020, December, pp. 26–38.