International Journal of Scientific Research & Engineering Trends
Volume 6, Issue 6, Nov-Dec-2020, ISSN (Online): 2395-566X

A Student Network Model for Image Classification

Saleema N P, Sruthy K G
Dept. of Computer Science&Engg.
MGM College of Engineering and
Pharmaceutical Science, Kerala, India
saleemaperidot@gmail.com, sruthydas.abhi@gmail.com

***Abstract*- Nowadays, Convolutional Neural Network achieved remarkable changes in image classification. To obtain the demanding precision of the categorized image we need a high-performance GPU and Huge datasets are necessary for training images. Image classification is computationally very expensive since it requires high-performance hardware devices and memory storage. So to tackle this issue we propose a student network model by utilizing the Knowledge Distillation technique, Let small networks learn from big teacher networks with accuracy compete with teacher networks.There exist many methods for trimming the convolutional neural network and designing a small network. But all these techniques are very expensive and have not achieved great results. The proposed model addresses the aforementioned issues effectively.**

***Keywords*– Convolutional Neural Network, GPU, knowledge distillation.**

I. INTRODUCTION

In 2010 the ImageNetLargescale Visual Recognition challenge is selected as a benchmark for image classification challenges. From there, onwards many types of research are going on in the field of object category identification and detection. Every year competition has been conducted in the same field. The organization, researchers, university students participated in the competition. Winners are awarded a cash prize. In 2010 AlexNet[1] won the competition.AlexNet was the first tried convolution neural network for image classification. It classified 1.2 million high-resolution images with 1000 different classes. Researches have gained extensive use of deep learning technology and keeps changing the record every year. Some of them are VGGnet, GoogleLeNet, ResNet, etc.

However, Excellent performance requires the support of a huge amount of computation. Portable devices and mobile devices do not have large storage capacity and computational resources, which blocks them from taking full advantage of deep neural networks.so, Network with less hardware requirement while still have similar accuracies are of great importance to machine learning and computer vision.meaning. Where as in word embedding feature extraction both semantic as well as syntax willconsidered.

Teacher-student learning technology makes utilizes the technology of knowledge distillation, which is the most popular approach to model compression and acceleration. Taking a well-trained network such as GoogleNet or ResNet, which is already been trained with bigdata and massive computational resources as the teacher network. Astudent network can be trained with a teacher's guidance. Which inherits the advantages of the teacher network.There exist many methods of trimming networks by using vector quantization and hashed nets by utilizing a hashing function. Other techniques include designing a small network. But those techniques are not efficient for solving the current issues. So we propose a student-teacher model by using the knowledge distillation technique.Hence, the robustness of the student network can resist noise on the test set can be improved. Experimental analysis shows that the proposed method for learning small and deep neural networks with comparable accuracy.Existing methods of the teacher-student network have experimented with clean data set. But in this paper, we examined noisy image still perform well.

The paper is categorized as follows. Section II describes the literature survey of the previous methods which is explained different text classification methods to detect various kind of speech detection technique. Section III explain the proposed method. Finally, the Section IV gives the conclusion.

II. LITERATURE SURVEY

Several studies had conducted to reduce the sizSeveral studies had conducted to reduce the size of the neural network, to reduce computation cost, storage capacity reduction. So as to use a deep neural network in portable and mobile devices. In this section, as a part of the survey, we mainly focused on teacher networks, Network trimming methods, and designing small networks.

1. Image Net Large Scale Visual Recognition Challenge

In 2010 Image Net [1] Visual Recognition challenge contributes to Image classification Identification , which was then accepted as a benchmark in the same field. The ImageNet Large Scale Visual Recognition Challenge aims at the detection of images and identification of the object in the image. It classifies 22 millions of images and 22 thousands of object categories.

ILVRC conducted a competition in the fields of computer vision technology every year. This attracts participants from many organizations and institutions. Image Net was developed by taking images from the internet and crowdsourcing. Arranged in WordNetformat.Word Net was developed manually in1985 and used in natural language processing researches. Amazon Mechanical Turk (MTurk) worked in research , they are part to remove unnecessary or irrelevant images from the dataset. Automatically download images from the internet and M Turk removes unwanted images. So, For classifying images large dataset, heavy computational resource and storage capacity is required. Therefore a new model is proposed for portable and mobile devices is proposed by using a limited number of the dataset, lightweight computational resources, and with a lesser storage device.

2. Image Net Classification with Deep Convolutional Neural Networks

AlexNet[2] is a large deep convolutional neural network developed by Alex Krizhevsky.In 2012 AlexNet won the ImageNet Large Scale Visual Recognition Challenge.AlexNet makes utilize of Graphical Processing Unit(GPU) for image classification purpose.AlexNet architecture contains a total of 8 layers among them 5 are convolutional layers and 3 fully connected layers.AlexNet uses DropOut, ReLu, and Preprocessing to reduce overfitting in the network, Which was a great achievement in performance and contribute to computer vision task.ALexNet had achieved a great resolution on a highly challenging dataset. Performance gets decreased when any convolutional layer is removed.AlexNet uses high computational power to increase performance. The larger network means higher performance. However, AlexNet uses GPU to get high performance, But we aim to design a small and compatible size network for image classification. without using GPU and reduce the memory storage.

3. Very Deep Convolutional Networks for Large-Scale Image Recognition

VGGNet [3] Start their research by increasing the depth of the neural network and calculate the effect on large-scale image classification. They increase the depth of the convolutional layer (16-19) and reduce the size of filters 3 by 3.VGGNet input size is fixed in 224 by 224 RGB image. The image is passed through several convolutional layers. The final layer is the softmax layer. They won thecompetition of ImageNet in the year 2014. They marked performance in image localization and classification , They achieved 1st and 2 nd price respectively. They are considered as the state-of-art result in image classification. Many version of VGGNet is available with different depth of the network.VGGNet is trained on Nvidia Titan GPU. It is very slow to train and takes 2-3 weeks for training. Memory consumption and computational time are high.

4. Rethinking the Inception Architecture for Computer Vision

By increasing the neural networks model size and computational cost there is a great chance of efficiency. But for portable and mobile devices we need to reduce the number of parameters, less computational cost. so designing a model for mobile and portable devices is a challenging task. Mobile vision and Bigdata scenario computational cost and storage is still a question mark. GoogleLeNet in 2014 came up with a deep network model with lower parameter and less computational cost still accurate.

Google Le Net efficiently factorizes convolutions and aggressive regularization had to make the network available for mobile and portable devices.GoogleLeNet won the competition of the ImageNet Large Scale Vision Recognition Challenge in 2014. The computational cost of GoogleLeNet is much lower than the VGGNet.Which is best suited for bigdata scenario and reduce memory use. They scale up the networks still maintaining the quality of classification.

5. Deep Residual Learning for Image Recognition

It is very difficult to train a neural network with deep network architecture or having many layers. Very deeper network suffers from degradation error as it passes through several layers. So, ResNet came into the machine learning environment with a residual function to learn the neural network. ResNet is deeper than other deep learning architecture. Which is easy to optimize and accurate. ResNet has depth up to 152 layers. Comparing with VGGNet ,ResNet is 8 times deeper. They won first prize in ILSVRC 2015 classification task.ResNet has 60 million parameters and 11B floating-point operations.

6. Compressing Deep Convolutional Networks Using Vector Quantization

Convolutional neural networks are widely used nowadays for image classification tasks. However deep convolutional neural networks are difficult to train, required high computational cost have a large dataset for training, required a large number of parameters, and required heavy computational cost. In cellphones and embedded devices, we do not have strict contains as mentioned above. So much researches are going on to make a neural network to make available for small and embedded devices.

Vector Quantization [6] is a method proposed to reduce the computational cost by reducing the parameters of the convolutional neural network. A good balance between model size and recognition accuracy is achieved by product quantization. This method compresses only densely connected layers but compressing

7. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications

Another major approach is to design a small network by designing a lightweight convolutional neural network to make it feasible for mobile and portable devices. Real-world applications such as self-driving cars and augmented reality require high computational cost and high efficiency. So MobileNet[7] proposes depthwise separable convolutions. The experimental result shows that these networks are suited for small devices. Unlike training large model MobileNet uses less regularization and data augmentation technique.

III. PROPOSEDSYSTEM

In the literature survey, we explained variously teacher network available. Teacher networks such as ImageNet, AlexNet, VGGNet, GoogleLeNet, Res Net are developed by making utilize of heavy network architecture. To train these network high computational resources. Vector quantization mentioned here is used to compress the densely connected layers.MobileNet is a network architecture used to design small networks. But all these have some disadvantages. In this paper student network model is proposed for small networks and portable devices without using a Graphical Processing Unit. The robust student network model aims to classify images by making utilize of teacher network by using a method called Knowledge Distillation. The existing Student-teacher paradigm works well with clean data, but the proposed method can work efficiently with noisy data as well..

IV. CONCLUSION

The paper focused on the literature survey that gets to a final discussion of student learning model with the help of teacher network. This paper discussed about 5 teacher networks [1]-[5].Vector quantization [6] method for compressing convolution network.Finally MobileNet[7],an example of small network. The proposed student learning paradigm differ from all these survey paper is that the proposed method is light weight but accuracy itcompatible with teacher network designed with heavy networks.

REFERENCES

[1]. O. Russakovsky et al., "ImageNet large scale visual recognition challenge," Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, Dec. 2015.

[2]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proc. Adv. Neural Inf. Process. Syst., 2012, pp. 1097–1105.

[3]. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," Apr. 2014, arXiv: 1409. 1556. [Online].Available:https:// org/abs/1409.1556

[4]. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 2818–2826.

[5]. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770–778.

[6]. Y. Gong, L. Liu, M. Yang, and L. Bourdev, "Compressing deep convolutional networks using vector quantization," Dec. 2014, arXiv:1412.6115. [Online]. Available: https://arxiv.org/abs/1412.6115

[7]. A. G. Howard et al., "MobileNets: Efficient convolutional neural networks for mobile vision applications," Apr. 2017, arXiv:1704.04861. [Online]. Available: https://arxiv.org/abs/1704.04861

About Author

Saleema N P received the B.Tech degree in Computer Science and Engineering from University of Calicut, Kerala, India, in 2014. Presently, she is doing her Post Graduation in the Department of Computer Science and Engineering, at MGM College of Engineering and Pharmaceutical Science, Valanchery, Kerala. Her latest research include Image processing, Cloud Computing, and Machine Learning

Sruthy KG obtained her Bachelor's degree in Computer Science and Engineering from MEA Engineering College Perinthalmanna, Kerala, India and Master's degree in Computational Linguistics from Governmennt Engineering College Sreekrishnapuram, Kerala, India. Currently, she is working as Assistant Professor in the department of Computer Science and Engineering, Cochin College of Engineering and Technology, Valanchery, Kerala, India. Her current research include Artificial Intelligence, Natural language Processing and InformationRetrieval