International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Pruned Cascade Neural Network Image Classification

G.D.Praveenkumar, M.Dharmalingam

Abstract- In this paper we propose a new model of deep neural network to build in deeper network. The convoluational neural network is one of the leading Image classification problem. The vanishing gradient problem requires us to use small learning rate with gradient descent which needs many small steps to converge and its take long time to proceed . By using GPU we can process more than one dataset (CIFAR-100) in a particular session. To overcome vanishing gradient problem by using the prune cascade correlation neural network learning algorithm compared to the deep cascade learning in CNN architecture. We improve the filter size, to reduce to the problem by training algorithm that trains in the network from bottom to top approach and its performing attain the task for better image classification in Google Net. We reduce the time complexity (training time), storage capacity can be used pre training algorithm.

Index Terms: Deep Learning, Vanishing gradient, GoogleNet, CNN, Image classification.

I. INTRODUCTION

Artificial Neural Network(ANN) have achieved the performance across a wide range of application are computer vision, natural language processing, speech recognition[15][18][20]. The field of machine learning comes from artificial neural network. Artificial neural network is a collection of multiple hidden layer has a powerful feature learning ability. The artificial neural network are based on non-linear activation function approximation which make them suitable for most of the application[16]. The major difference between a artificial neural network and convolutional neural network is that only in the lasts of fully connected layer but in artificial neural network each neuron is connected to every other neuron. Artificial neural network is not suitable for image because the network leads to over fitting quickly due to the size of the images. The convolutional neural network are commonly used in the field of image processing and pattern recognition. The cascade correlation neural network architecture with supervised learning implemented in resilient back propagation algorithm to train the data[21].
Convolutional Neural Network(CNN) is one of the image classification architectures for hierarchical feature extraction. Deep convolutional neural network become more complex consisting of more layers size etc.

Manuscript published on 30 September 2019
* Correspondence Author
G.D.Praveenkumar, Research Scholar , Department of Computer Science, Bharathiar University Arts and Science College-Modakkurichi, Erode ,India
dr.M.Dharmalingam, Assistant Professor, Department of Computer Science, Bharathiar University Arts and Science College-Modakkurichi, Erode ,India

To build deep layer network in during training method the network fall appalling of an issues known as vanishing gradient problem[4,8] in our novel approach to handle the gradient weight using training algorithm that trains the network from bottom to top layers incrementally with pruned cascade correlation algorithm.
This pruned cascade correlation developed by Scott Fahlman[22] proposed the pruned cascade neural network architecture as an approach used learning algorithm like back propagation, quick propagation or Rprop. It is an constructive learning rule , starts with single layer consist only of an input and an output layer. Cascade correlation is a supervised learning architecture which build near minimal multi layer topology as an approach to sequentially train perceptron and connect this output to perform single classification[15,17]. In our proposed method using an same computation with different cascade algorithm to develop deep cascade layer by pruned cascade correlation neural network. This algorithm reduces the memory and time requirements .Convolutionalneural network are very flexible because they are trained as feature extractor and not only a classification devices[13]. Now , CNN are increased layer by layer. GoogleNet architecture contain 22 layers in comparison to VGG which held a 19 layer also made a approach called inception module shown in Fig 1. In a single layer multiple types of feature extractors or present . The network perform better as the network training itself has many option to choose from when solving the task, GoogleNet train faster than VGG[5,6].VGG model can have greater than 500 MB, GoogleNet has a size of only 96 MB.

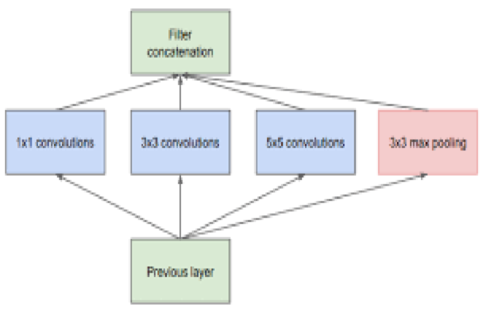

Fig 1. Inception module

II. RELATED WORK

Stefano Squartini(2016) etal[24] to solve the vanishing gradient problem in recurrent neural network. Preprocessing are done by discrete wavelet transformation. Two problems are identified latching problem ,long term memory in series prediction and develop new approach called RMN. Stefano Squartini(2003)etal[25] compare the performance of different recurrent neural network again define a latching problem and to introduce a three architecture to overcome the latching problem ,fully recurrent neural network, recurrent multiscale network, echo state network. MoiHoon Yapetal[19] developed the deep learning approaches to lesion detection on CNN. Three different methods are used in patch based Lenet, Unet, transfer learning approach with a preprocessed FCN-Alexnet with compare the four algorithm and improved by the deep learning network to evaluate the performance of computer aided detection frame work.Yasuwri Ishii etal[26] proposed a deep learning in heterogeneous network with multiple size of filters, filters size are vary from one another and applying zero-padding techniques. Hsin – keriHuang(2016) etal[9] developed a mixture of deep CNN based ensemble model for image classification used in Alexnet and NIN ECNN combines two different architecture.Dharmalingam (2015) etal [15] was solve double dummy bridge problem in gamming theory to neural network to solve the bridge problem on cascade correlation neural network and Elman neural network with work point count system in Rpropagation, cascade correlation is an superior of ENN method. Both methods are produce the better result[16][17][20]. Scott Fahlman(1990) etal[22] is a supervised learning architecture which built a minimal multilayer topology, combines two ideas (i.e) cascade architecture , learning algorithm. It starts with minimal network consisting only of an input and output layer both are fully connected. The small architecture to built automatically and built deep network without back propagation. Learning take place from an ordinary learning algorithm to avoid the residual error and learns quickly. Alex Krizhevsky(2012)etal[1] to classify 1.2 million images in the ImagenetLSVRC into 1000 different classes. Neural network contain 60 million parameter, 6,50,000 neuron consist of eight learned layer, five convolutional and three fully connected layers with softmax to reduce the overfitting multi-GPU and local response normalization ,overlapping, data augementation , dropout[14].YadingYuanetal[28] develop a deep convolutional neural network that is trained end to end learning. Novel approach was loss function based on jacard distance to eliminate need of using entropy as the loss function for image segmentation. Fully convolutionalnetwork to map an RGB input contain 19 layers with 290, 129 trainable parameter and use ReLu with deconvolutional layer. Fully convolutional network model to convert image segmentation into a pixel wise classification problem, improve the efficiency of network training while reducing the overfitting. Adam optimization method used to reduce the overfitting. Alex Krizhevsky (2010)etal[2]is dealing with edge pixels with CNN to reduce the overfitting on CIFAR-10 dataset. CNN is a modification of the basic architecture of Alexnet.ZejiaZheng based on the combination of CNN and hand crafted color histogram feature extractor. It minimize the number of lower level kernel in a convolutionallayerby dividing color information from real image. Sameer Khan(2017) etal[23] proposed evaluate the three milestone architecture is Lenet,Alexnet and GoogleNet and proposed CNN architecture for classifying medical anatomy images. These models overfit due to the number of layers and hyper parameter used in these architecture has been layer set of natural image and produce better result of image classification. Enrique S. Marqueze(2018) etal[7] developed by cascade correlation approach in vanishing gradient problem by proposing training algorithm , that trains the network from the bottom to top layer incrementally. In this proposed system develop two methods are End to End learning and Deep cascade learning on CIFAR-10 & 100 dataset with VGG (Visual Geometry Group)network and all fully connected layer . It reduce the vanishing gradient problem to layer filter to change the initial setting[25]. This algorithm reduce the memory and time requirement ,if training compared with the traditional end to end backpropagation ,solves the 10 and 100 class problem and attain the best accuracy.Zhezhi He(2018) etal[29] is achievedgreat success in many artificial intelligence application ,proposed statistically weight scaling and residual expansion methods to convert the entire network weight parameter to ternary format with objective to greatly reduce model size, computation cost, accuracy on CIFAR-10 and Imagenet.

III. PRUNED CASCADE NEURAL NETWORK

Deep pruned cascade neural network algorithm is the computational advantages of training layer wise manner. It is used to minimize the expected test set error , instead of actual training error . It tries to determine the optimal number of hidden units and to remove unneeded weights after a new hidden unit is installed, stop learning is used to applied to digest away unneeded weights[26] in algorithm 1.

A.PRUNED CASCADE NEURAL NETWORK ALGORITHM

Algorithm 1: pseudo code of pruned cascade correlation learning algorithm derived from cascade correlation learning algorithm [7,11,20, 27].

$\boldsymbol{Procedure}: prunedcascadelearning\ (Layers, n, epochs, epochsupdate\ , h, out)$
$\boldsymbol{Input} Layers: modellayerparameter$
$n:\ learningrate$
$epochs:\ starting number of epochs$
$k: epochsupdateconstant$
$h: hiddenneuron\ /layer$
$out:\ outputblockspecification$
$\boldsymbol{output} w:\ wl\ outputblock$
$forLayerIindex\ = 1$
$: Ldo\quad /$
$/\ performcascadelearningthroughtrainablelayer$
$InitnewLayerandconnectoutputblock$
$h = convl * pool$
$il \leftarrow epochs\ + k * Layer\_index\ /$
$/\ computetheselectioncriterion.$
$//\ trainedthenewLayerwithweight\ (w)$
$Fori = 0; i + +; i < ildo$
$wnew \leftarrow wold - h\nabla j\ (w)$
$* h/$
$/\ updateweightbygradientdesecent$
$Ifwnew \leftarrow last \leftarrow 0\ then$
$il \leftarrow epochs + k * layer\_index * h$
$else$
$wnew \leftarrow lastwli$
$wnew \leftarrow wli$
$endif$
$endif$
$ifil >\ 1\ then$
$hnew \leftarrow wnew$

Retrieval Number F2929037619/2019©BEIESP
DOI: 10.35940/ijrte.F2929.098319
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$il \leftarrow epochs + k * layer\_index * hnew$
$if valaditionerrorplateausthen$
$n \leftarrow n/10$
$endif$

The pruned cascade neural network algorithm splits the network into the layers and trains each layer one by one until all the layer in the input architecture have been trained. To reduce the vanishing gradient problem by network to learn feature correlated with output on each and every layer[8]. This algorithm takes the input from the cascade learning and computed by the layer by layer to update the weight by gradient descent. The training follows bottom to top layer are trained by CNN the layer 1 output value is passing through next layers as inputs and so on. In our proposed algorithm can satisfied by cascade learning algorithm and introduced the$h$(hidden layer/neuron),wl is passing an input to train the connection of layers and compute the selection criterion, determine optimal number of hidden layer to remove unwanted weights after hidden units install , a set each weights of the last inserted unit to zero and compute . The selection criterion a $hi(epochs) = k$, if it is >1 then compute before inserting the new hidden unit[7].

B.TIME COMPLEXITY

The time complexity of the convolutional neural network layer is

$$o\left(\sum_{l=1}^{d} n_{l-1}\, s_l^2 n_l m_l^2 i\right) \quad (1)$$

Here$l$ is the index of a convolutionalal layer d the last layer index,n is the number of filters, s is the spatial size of the filter (input), $m$is the spatial size of output. $Ni - l$is also known as the number of input channel of the lth layer,il represent the number of training iteration for the lth layer. In prune cascade algorithms inherits from deep cascade learning, so we have directly implement an same equation 1[7]. There is no modification done in that equation because in existing system used in VGGNet[7] and proposed system implement in GoogleNet, compare the time complexity of both system in CNN. The time complexity applies to both training and testing with different scale. The training time image and testing image are processed by both forward and backward propagation in convolutionalal layers.

IV. IV.EXPERIMENTS

A. DATASET

We are used unbalanced dataset isCIFAR-100. The CIFAR -100 dataset contain 60000, 32*32 color image in three sets of 45,000 images for training , 5,000 images for validation and 10,000 images for testing.

B.GOOGLENET

GoogleNet is the flexible deep learning architecture consisting of the early layer, middle layer,final layer[3]. In our proposed system to developed by GoogleNet show in Fig2. , image are split into 32*32 patches. The input patches are processed by convolutional layer set as kernel size of filter are 1*1, 3*3 and 5*5 as like the middle layer is formed by repetitive structure called inception layer values passed

$endfor$
$disconnectoutputblockandgetnewinputs$
$endfor$
$endprocedure$

into the maxpool set as filter is concatenate with early layer and finally connected as dense layer or fully connected layer with softmax produce an final classification of image. In this level of training the network with mean while window , the filter weight are randomly initialized by non linear and biases are set to 0.0. The learning rate is 0.001. Training 32*32 image is randomly accepted from full image is used , no other data augmentation such as scale jittering is used[23,28]. A network goes deeper problem such as vanishing gradient may become more which make it more difficult to tune the parameter of the earlier layer. The dataset are reduce the over fitting. The training weight are update from output layer to input layer. The random weight is used to decrease the vanishing gradient problem.

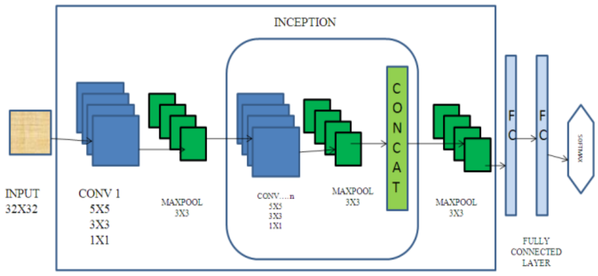

Fig.2. Architecture of the proposed GoogleNet.

C.EXPERIMENT RESULT

These experiments performed on CIFAR-100 dataset to evaluate and compare learning capabilities from vanishing gradient problem . First the learning pattern of pruned cascade learning differ from the cascade learning. Building block are train the complex value network such as patch size and weight initialization. The pruned cascade learning algorithm follows the back propagation passing output value to next layer as input. The existence of shorter pathway help gradient back propagation during training multipath is also used in but the extra objective function are used in hidden layer to difficult in train deep network[12]. The result of our proposed method attained better performance than existing system show in Table I.Convolutional neural network models at constrained time cost during both training and testing stages investigation involve the depth, width. Filter size and strides of the architecture because the time cost constraint differences the architecture [10,11].

Table I. comparison on accuracy in GoogleNet and VGG.

LAYERS	1	2	3	4	5	6
GoogleNet	0.3	0.37	0.48	0.53	0.57	0.6
VGG	0.35	0.39	0.5	0.5	0.53	0.59

The space complexity and memory usage of a network consider batches of data used for the training, so the amount of data at every layer is product by the patch size. All the experiments are run on single CPU or GPU. The input images 32*32 and output block consisting fully connected layer size of parameter like 64,128,256,512 convolutionalfilter with a size 1*1,3*3,5*5. The relevant maxpooling between a block .GoogleNet is reduce the number of parameter that does not affect any convolutionallayer . In pruned cascade neural network algorithm to reduce the size of parameter and attain the accuracy of 0.9% in image classification shown in Fig.3.

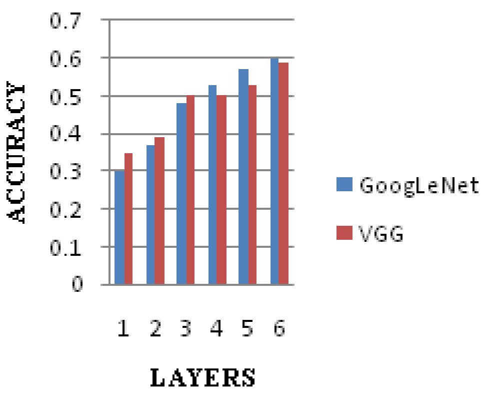

Fig.3Performance of accuracy in GoogleNet and VGG.

V. V.CONCLUSION

In this work , we proposed pruned cascade learning based on deep convolutional neural network for image classification on CIFAR-100 data set . GoogleNet architecture are developed by convolutional layer with max pool layer in depth concatenate and final layer called dense layer or fully connected layer with softmax to display classification output.GoogleNet consume more memory than deep learning model.The deep convolutional layer is to decrease vanishing gradient problem ,reduce the time complexity, space complexity to compare the cascade learning in VGG .

REFERENCES

1. A.kerizhevsky, "Convolutional deep belif network on CIFAR-10",Aug 2010,pp.1-9.
2. A.krizhevsky, I.Sutskever and G.E.Minton "Imagenetclassification with deep convolutional neural network", in Proc.Adv.NeuralInf.Processing Syst.,2012, pp.1097-1105.
3. B.Y.Lin, C.S.Chen, "Parallel DCNN for pedestrian detection", IEEE,2015.
4. C.S.Zegedy,W.Liu,Y.Jia,P.Sermanet,S,Reed,D.Anguelov,D.Erhan,V. Vanhoucke and A.rabinovich, "Going deep with convolutions", in Proc of IEEE, CVPR 2015 pp.1-9.
5. C.Trabelsi,O.Bilaniuk,Y.Zhang,D.Serdyuk, "Deep Complex Network",ICLR, arvix,Feb 2018 ,pp.1-1 9.
6. C.Wang,Y.Xi "Convolutional neural network for image classification".
7. E.S Marquez,J.S.Hare, M.Niranjan, "Deep Cascade Learning", IEEE trans on Neural network and learning system,2018,pp-1-11.
8. G.D.Praveenkumar, M.Dharmalingam, " Analysis of vanishing gradient problem in deep learning on CIFAR dataset", Proc.,NACA, Jan-2019 ,ISBN-9789387862463,PP.1-3.
9. H.K.Huang,C.Fchiu,C.H.Kuo,Y.C.Wu,N.N.Y.Chu,P.C.Chang, "Mixture of Deep CNNBased Ensemble Model For Image Retrieval", IEEE, 5thGlobal conference on consumer electronics,2015.
10. J.Long,E.Shelhamer,T.Darrel, "Fully convolutional networks for semantic segmentation",IEEE Conf,CVPR,2015,pp.3431-3440.
11. K.He and J.Sun, "convoluational neural network at constrained time cost",inProc,IEEE conf.CVPR,2015,pp.5353-5360.
12. K.He,X.Zhang,S.Ren and J.Sun "Delving deep into rectifiers: surpassing human level performance on imagenet classification" in Proc IEEE,ICCV-2015,pp.1026-1034.
13. K.Simonyan,A.Zisserman, "Very deep convolutional neural network for large scale image recognition",ICLR,2015.
14. M.D.McDonnell, R.G.McKilliam ,P.d.Chazer "On the Importance of Pairwise Feature Correlation for Image Classification",IEEE,IJCNN,2016.
15. M.Dharmalingam and Amalraj, "Artificial Neural Network Architecture for solving double Dummy bridge problem in Impact game",IJCII,vol.5, No.3., 2015., pp.188-197.
16. M.Dharmalingam, "Double Dummy Bridge Problem in Contract Bridge: An Overview", International Journal Artificial Intelligent Systems and Machine Learning, Vol.No 10(01), pp-01-07, January 2018.
17. M.Dharmalingam, "Supervised learning Architectures for solving double dummy bridge problem" First International Conference on Computer Vision, Networks and Informatics conducted at The Gandhigram Rural Institute, Gandhigram, pp 30-37, March 22-23, 2018, (International Journal of Computer Sciences and Engineering)
18. M.Dharmalingam, R.Amalraj, "A Solution to the Double Dummy Bridge Problem in Contract Bridge Influenced by Supervised learning module Adapted by Artificial Neural Network", ICTACT Journal on Soft Computing: Special Issue on Distributed Intelligent System and Applications, Vol. No 05(01), pp 836-843, October 2014.
19. M.H.Yap,etal "Automted Breast ultrasound lesion detection using Convoluational Neural Network" IEEE transaction content mining 2016 pp.1-9.
20. R.Amalraj, M.Dharmalingam, "A Work Point System Coupled with Back-Propagation for Solving Double Dummy Bridge Problem", Neurocomputing (Elsevier) Vol. No 168, pp 160-178, June 2015.
21. R.O.Duda,P.E.Hart and D.G Stork,Pattern classification, 2nd edition Hoboken.NJ,VSA,2012.
22. S.E.Fahlman,C.Lebiere, "The cascade-correlation learning architecture", in Advances in Neural Information Processing Systems (NIPS) D.S. Touretzky (Ed.),Vol. 2, 1990, pp. 524–532.
23. S.Khan,S.P.Yong, " A deep learning architecture for classifying medical: Image of anatomy object",Proc of APSIPA,2017.
24. S.Squartini, S.Paolinelli, F.Piazza, " comparing different recurrent neural architecture on a specific task from vanishing gradient effect perspective", IEEE, ICNSC,2006, pp. 380-385.
25. S.Squartini,A.Hussain,F.Piazza, "Attempting to Reduce the Vanishing Gradient Effect Through A Novel Recurrent Multiscale Architecture" ,IEEE,IJCNN,2003, pp.2819-2824.
26. Y.Ishii ,R.Hagawa,S.Tsukizawa "Deep Learning Using Heterogeneous Feature Map for Maxout Network",3rd IAPR,2015,pp.459-463.
27. Y.Lecun.L.Bottou,V.Bengio and P.Haffener, "Gradient-Based Learning applied to document recognition" ,Proc.IEEE.VOL.86 no.11,pp.2278-2324,1998
28. Y.Yauan,M.chao,Y.C.Lo "Automatic Skin Lesion Segmentations Using Deep fully convolutional Networks with Jccard Distance", IEEE transaction on medical imaging, Vol 36,No.9,2017 pp-1876-886.
29. Z.He,B.Gong,D.Fan "Optimize Deep Convoulational Neural Network with Ternarized weights and high accuracy" , arvix,2018.