International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 06, June 2017
www.ijarse.com

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

WASTE IDENTIFICATION AND CLASSIFICATION THROUGH EXISTING CONVOLUTIONAL NEURAL NETWORKS

Saifali Noorani[1], Milind Fernandes[2]

[1,2]Electronic communication and instrumentation, Goa college of Engineering, Goa University (India)

ABSTRACT

To make the waste disposal process easier, waste can be classified using existing Deep Convolutional Neural Networks. Hence we aim at studyingexisting Convolutional Neural Networks and see how they identify the object present in the waste.

Keywords - AlexNet, Convolutional Neural Networks (CNN/ConvNets), Caffe, Google Net, Top-1 error rate, Top-5 error rate

II. INTRODUCTION

Waste management association, administration and transfer issues are not simply issues of a specific nation or a landmass. It is a worldwide issue which ought to be dealt with instantly. Nations and governments are indicating worry over issues with their waste recognizable proof, characterization, administration and transfer. As per the United Nations, there are around 60 percent of nations around the globe that communicated their worry about overseeing squanders and other ecological worries in the 2015 Earth Summit. Therefore for proper waste administration, waste gathering, waste recognizing, classification, is the important step. Appropriate gathering is the foundation for ensuring that the aggregation, transportation, stockpiling and treatment of waste is done in a way that offers protection to the earth and human prosperity and in consistence with authentic essentials.

Subsequently this venture goes for identifying the leftovers and classifying them into various classes .The most ideal way is to do this by using Deep Convolutional Neural Networks. Hence different convolutional neural networks are trained and it classifies the remaining data into appropriate classes.

II. LITERATURE SURVEY

At the point when Alex Krizhevsky, IlyaSutskever, Geoffrey E. Hinton [1] arranged a sweeping, significant convolutional neural framework to bunch the 1.2 million high-assurance pictures in the ImageNet LSVRC-2010 test into the 1000 assorted classes. On the test data, they accomplished TOP 1 and TOP 5 Error rates of 37 percent and 17.07 percent which is fundamentally better than the past front line. The neural framework, which has 60 million parameters and 650,000 neurons, contains five convolutional layers, some of which are trailed by maxpooling layers, likewise, three totally connected layers with a last 1000-way softmax. To make get ready speedier system, they used non-splashing neurons and a to a great degree gainful GPU execution of the convolution operation. To decrease overfitting in the totally related layers they used a starting late made regularization technique called "dropout" that wound up being to a great degree capable. We in like manner entered a variety of this model in the ILSVRC-2012 contention and fulfilled a triumphant TOP 5 test mistake rate of 15.3.

Christian Szegedy,Wei Liu , YangqingJia, Pierre Sermanet, Scott Reed, DragomirAnguelov, DumitruErhan, Vincent

Vanhoucke, Andrew Rabinovich at created google[5] .GoogleNet is similar to other neural networks except for the fact that they use inception modules instead of convolutional layers. Inception module uses 1x1, 3x3, and 5x5 convolutions along with a 3x3 max pooling. Max pooling layer is included to the Inception module since good networks include pooling for parameter reduction. The bigger convolutions are all the more computationally costly, so the paper recommends first doing a 1x1 convolution diminishing the dimensionality of its feature map, which is given to a ReLu layer, and afterward doing the bigger convolution (for this situation, 5x5 or 3x3). The 1x1 convolution is important because it will be used to reduce the dimensionality of its feature map.

III. CONVOLUTIONAL NEURAL NETWORKS (CNNS / CONVNETS)WORKING

The input given to the CNN has the size 224 × 224 pixel image. This image is RGB.NO preprocessing is done except, mean RGB is calculated on the training set image and this value is subtracted from each pixel of input image. Resultant image is then given to a chain of layers known as the convolutional layerswhich basically have filters. These filters have a tiny receptive field: 3 × 3. In one of the designs 1 × 1 convolution channels, which can utilized for linearity of the info channels which is then trailed by non-linearity layers. The stride is 1 pixel for convolution; the convolutional layers are spatially padded to conserve spatial resolution,which is the padding is 1 pixel for 3 × 3 convolutional layers. To incorporatespatial pooling there are five max-pooling layers, max-pooling layers are after the convolutional layers (not all the convolutional layers are followed by max-pooling). Max-pooling is refined over a 2 × 2 pixel window, with stride of 2. For case, AlexNet has a heap of convolutional layers is trailed by three Fully-Connected (FC) layers: the initial two fully connected layers have 4096 channels each, the third fully connected layer performs 1000-way ILSVRC order and therefore has 1000 channels, that is one channel for each class. The last layer is the soft max layer. The design of the fully connected layers is comparative for all systems. Every hidden layer are given the ReLu layer (correction) .This is for non-linearity.

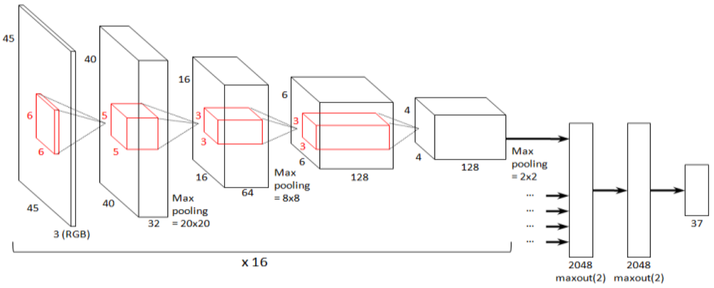


Fig. 3.1.Basic Convolutional Neural Network which has neurons arranged in 3D.

IV. NETWORKS STUDIED

4.1. AlexNet[1]

Let us observe the AlexNet and its engineering. As appeared in Figure4.1 [1], the system contains eight layers with weights; the initial five are convolutional and the staying three are fully connected(FC) layers. The yield of the last fully connected layer is given to a 1000-way softmax which makes dispersion over the 1000 classes. This framework extends the multinomial vital backslide objective, which is indistinguishable to intensifying the typical transversely over get ready examples of the log-probability of the correct stamp under the estimate flow. The filter on the second, fourth, and fifth convolutional layers are related just to those bit maps in the past layer which harp on the same GPU. The bit of the third convolutional layer is associated with all filter maps in the second layer. The neurons in the completely associated layers are associated with all neurons in the past layer. Response-standardization layers is put after the first and second convolutional layers. Max-pooling layers are obliged after both response-standardization layers and in addition the fifth convolutional layer. The ReLu non-linearity is put after each fully connected layer or convolutional layer.

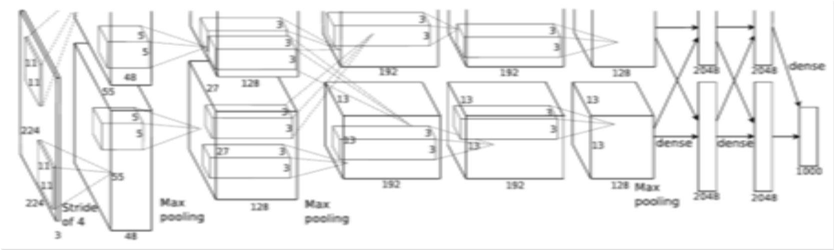


Fig. 4.1.Architecture Of Alexnet.

4.2 Berkeley Model

The Berkeley model was developed by YangqingJiafor caffe frame work[2]. It is alike to the AlexNet, that is it has eight layers. But the Berkeley model is trained on different set of images with respect to AlexNet. Berkeley model has eight layers,the first five are convolutional layer which are followed by three convolutional layers. The last fully connected layer is followed by the softmax layer. Max-pooling layer may be placed after a convolution layer or a fully connected layer to reduce the parameters after the operation. Fully connected layer or the convolutional layer may also be followed by ReLu layerwhich increases the non-linearity to the output.

4.3.GoogleNet[5]

GoogleNet is like other neural systems aside from the way that they utilize inception modules rather than convolutional layers. Inception module[5] utilizes 1×1, 3×3, and 5×5 convolutions alongside a 3×3 max pooling. Max pooling layer is incorporated to the Inception module since great systems incorporate pooling for parameter decrease. The greater convolutions are more computationally expensive, so the paper prescribes first doing a 1×1 convolution reducing the dimensionality of its feature map, which is given to a ReLu layer, and a while later doing the greater convolution (for this circumstance, 5×5 or 3×3). The 1×1 convolution is vital in light of the fact that it will be utilized to lessen the dimensionality of its feature map.

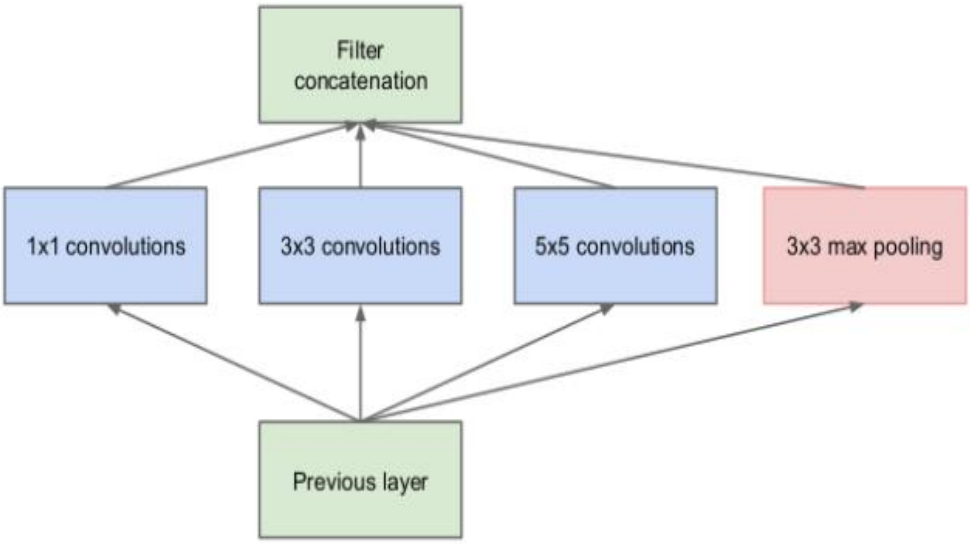


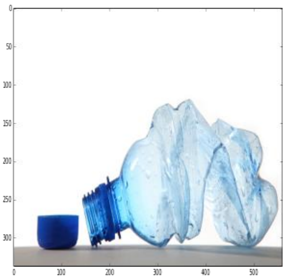
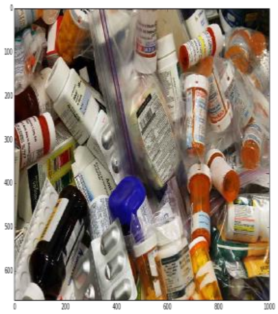
Fig. 4. Inception Module

V. OUTPUT AND RESULT

International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 06, June 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

The below table shows the output achieves for given input. It is seen that these networks give top five results for a given input. Form the output top-1 and Top-5 error rates are achieved. It was seen that TOP-5 error rate was less than the TOP-1 error rate for each network

TABLE 5.1. OUTPUTS ACHIEVED FOR CORRESPONDING INPUTS

Input image	GoogleNet	Alexnet	Berkeley reference model
	1.plastic bag 2.diaper,nappy,napkin 3.water bottle 4.Band Aid 5.handkerchief, hankie, hanky, hankey	1.plastic bag 2. West Highland white terrier 3.Ibizan hound, IbizanPodenco 4.water bottle 5.Chihuahua	1.plastic bag 2.jigsaw puzzle 3.conch 4. pinwheel 5. diaper, nappy, napkin
	1.water bottle 2.oxygen mask 3.nipple 4. gasmask, respirator, gas helmet 5. spotlight, spot	1.running shoe 2.nipple 3.gasmask, respirator, gas helmet 4.diaper, nappy, napkin 5.sandal	1.oxygen mask **2.**gasmask, respirator, gas helmet 3.diaper, nappy, napkin 4.nipple 5.plastic bag
	1.pill bottle 2.sunscreen, sunblock, sun blocker 3.hair spray 4. lotion 5. medicine chest, medicine cabinet	1.pill bottle 2. paintbrush 3.thimble 4.syringe 5.hair spray	1.shoe shop, shoe-shop, shoe store 2. paintbrush 3.thimble 4.hair spray 5. perfume, essence
	1.syringe 2.screwdriver 3.carpenter's kit, tool kit 4.whistle 5.paintbrush	1. syringe 2.whistle 3.screwdriver 4.oxygen mask 5. hammer	1.powerdrill 2.carpenter's kit, tool kit 3. hammer 4. syringe 5.dumbbell

International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 06, June 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

TABLE 5.2

TABLE SHOWING THE ACCURACY AND ERROR RATES OF THE NETWORKS USED

Name of the network	Bvlc_model	Alexnet	Googlenet
Top most Accuracy for 100 images	30	31	38
First five suggestion accuracy for 100 images	48	49	57
Top-1 error rate	70%	69%	62%
Top five Error rate	52%	51%	43%

Lets look at first example of table 5.1. It is seen that all the 3 networks identified the waste image of plastic bags correctly. The probably given was 32%, 20% and 19% by GoogleNet, AlexNet and Berkeley model.

GoogleNet has more accuracy and lesser error rate as compare to other models.

Error rates are given in table 5.2

V. CONCLUSION

This methodology successfully developed a system that can classify waste.

It is seen that more the number of layers better is the accuracy of the system , hence it was seen that Googlenet having more number of layers has better accuracy.

It was also seen that if a network is trained with a set of images for a particular object , during validation the network is able to identify the object easily , that is the accuracy further improves .

In ILSVRC-2012 competition AlexNet had achieved a winning top-5 error rate of 15.3% while the GoogleNet has a Top-5 error rate as low as 6.67%.But networks above that is AlexNet, Berkeley model and GoogleNet show Top-5 error rate of 51%, 50% and 43%, hence stating that these networks have not been trained for waste image database or the input images we have fed to it.

VI. FUTUREWORK

Create a database of waste images which is around more than 10,000 in number , so that the network has good convergence. Use this database to train a network with appropriate number of layers. This should improve the accuracy and reduce the error rate of the network.

Future work also involves that waste classification happens in real time that is the project is converted into a real time project.

REFERENCES

[1] Alex Krizhevsky, IlyaSutskever , Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", Neural Information Processing Systems Conference, 2012.

International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 06, June 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

[2] YangqingJia, Evan Shelhamer, Jef Donahue, Sergey Karayev;Jonathan Long; Ross; Girshick; Sergio Guadarrama; Trevor Darrell "'Caffe: Convolutional Architecture for Fast Feature Embedding", ACM MULTIMEDIA 2014 OPEN SOURCE SOFTWARE COMPETITION UC Berkeley EECS, Berkeley, CA 94702, 2014.

[3] Ciresan, D. C., Meier, J., and Schmidhuber, J. Multicolumn, "Deep neural networks for image classification ", In CVPR, 2012

[4] R. B. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation. In Computer Vision and Pattern Recognition ", IEEE Conference on, 2014.

[5] Christian Szegedy1, Wei Liu2, Yangqing Jia1, Pierre Sermanet1, Scott Reed3,DragomirAnguelov1, Dumitru Erhan1, Vincent Vanhoucke1, Andrew Rabinovich, "Going Deeper with Convolutions.", Neural Information Processing Systems Conference,2015

[6] S. Kaiming He, XiangyuZhang, ShaoqingRen, Jian Sun, "Deep Residual Learning forImage Recognition. ", Microsoft Research, 2015.