IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Kids Interactive Learning Platform

Nandini Ankam
Dept. Computer Science and Technology
Usha Mittal Institute of Technology
Santacruz (W), Mumbai
Email: nandiniankam1@gmail.com

Sneha Kale
Dept. Computer Science and Technology
Usha Mittal Institute of Technology
Santacruz (W), Mumbai
Email: snehakale987@gmail.com

Rashi Kocheta
Dept. Computer Science and Technology
Usha Mittal Institute of Technology
Santacruz (W), Mumbai
Email: rashikocheta99@gmail.com

Prof. Sumedh Pundkar
Dept. Computer Science and Technology
Usha Mittal Institute of Technology Santacruz (W), Mumbai
Email: sumedhpundkar@gmail.com

Abstract—Education is a prime factor of growth of the individual, community and the world. The children use mobile phones for playing games, watching cartoons or only for entertainment purpose. The basic objective of writing this paper to utilise children time and energy in productive works which help them in developing their skills. This platform will provide children with primary knowledge at one platform that is going to entertain them and also help them in self-development and nation-building.

Index Terms—Computer Vision, Image Processing, Image Synthesis, Neural Network, Convolution model, Speech synthesis, Blind, Pattern recognition.

I. INTRODUCTION

Children around us of any age group use mobile phones better than us for playing games, watching cartoons. This gives a clear observation that young minds are brilliant nowadays and can be utilize in learning key things. The main aim of the paper is to explore a platform for children of age group 3-8. An interactive learning application that provides them basic education right from Rhymes, Alphabets, Handwriting improvement with regular tests and parental login systems.

The younger generation this world commonly known referred to as the 'Gen Z', is seen to be very well acquainted and comfortable with all the modern-day technologies. These kids are tremendously intrigued by, and attracted to all the modern-day electronic devices. Therefore, with firm belief that technology could indeed prove to be the best in true to these kids by combining their proficiency with the technology with the immense power of knowledge. The vision is to promote early learning in the kids between the age group of 3 years to 8 years by making learning fun, interesting and extremely promising with the help of this learning platform. This platform will make good quality learning material accessible to the kids at the comfort of their homes, and thereby, fostering the habit of learning at a very young age.

II. LITERATURE REVIEW

Convolution neural network (CNN) also known as shift invariant is the most popular deep learning architecture which is fully connected [1]. These networks were encouraged by the idea of 'neurons', a biological term specialized in transmitting information that resides in the visual cortex in the mind. In ConvNet the neurons are arranged in three dimensions which make it outstanding from the regular neural network, the three main layers in ConvNet are (a) Convolutional layer(b) Pooling layer © Fully connected layer CNN has successfully reached down in the applications used in everyday life for instance in analysing documents [1]. In ImageNet, there are some effective CNN models, one particular model, that this paper states is the "GoogleNet" which is a 22 layer deep network. The focus of the paper was on a deep neural architecture codenamed Inception that can achieve a new range in classification and detection in the ImageNet challenge for computer vision. Inception layers when repeated many times lead to a 22- layer deep model that is the GoogleNet model [2]. The Inception architecture used filter sizes of 1×1, 3×3, and 5×5 which are based on convenience. It implies that the architecture depends on the input of the next stage. Layer after layer construction is used in which the last layer is examined and its correlation statistics are analysed so that it

can be grouped with higher correlation units. To reduce the expenditure of a convolution network as after pooling units the merging leads to the sudden increase in the outputs, the Inception architecture reduces dimensions whenever the computational requirement increases [2]. The training of the GoogleNet network used asynchronous gradient descent of 0.9, Polyak averaging to create the final model used at inference time, the images were resized to 4 scales. The output of GoogleNet for model performance detection was 38.02 percent, which gives stronger results with an ensemble that motivates for further use of Inception architecture [2]. This approach is similar to R-CNN but here the Inception model is used as the region classifier. Architectures of CNN noisome

One other paper uses in Web-based Interactive Virtual Classroom instead of using Adobe Flash or Microsoft Silverlight, this project utilizes new technology called WebRTC to create video conference systems. WebRTC (Web Real-Time Communication) is a free, open-source project providing web browsers and mobile applications with real-time communication (RTC) via simple application programming interfaces (APIs)[3]. It allows audio and video communication to work inside web pages The communication is done through SDP (Session Description Protocol). After successfully exchanging, two peers can set up local and remote descriptions where they can see each other. Screen sharing only works on SSL domains, which as the application must be accessed through HTTPS [3]. Result shows that the teacher (host machine) assigns into the classroom and appears in main box surrounded by sub boxes as students along with ppt presentation and also web services for learning.

Summary of Different CNN architecture.

Model	Year	Top-5 Error	Filter Sizes	# of Layers	# of Weights	# of GFLOPs
AlexNet [9]	2012	18.1%	3,5,11	5	61M	0.72
VGGNet-16 [11]	2014	9.9%	3	13	138M	15.5
GooLeNet [12]	2014	7.89%	1,3,5,7	21	7M	1.43
ResNet-50 [2]	2015	7.6%	1,3,7	49	25.5M	3.86
DenseNet-161 [3]	2016	6.34%	1,3,7	160	28.9M	7.7
DPN-131 [13]	2017	5.12%	1,3,7	130	79.5M	16.0
SE-ResNet-50 [14]	2017	4.47%	1,3,7	49	27.5M	3.88

Fig. 1 Top 5 error is tested on the ImageNet[15] validation set with single crop

III. Problem Statement

Objective is to help child to build a strong foundation for their future academic success by providing a comprehensive and engaging online curriculum to greatly assist early learners to succeed in kindergarten and in early elementary school programs. Platform should able to explore with their parents, or on their own safely. Kids Interactive Learning is Platform for learning giving test by drawing numbers which will be checked from the drawn input for this different available architecture can be used such as Neural Net, convolution Neural Network, VGG 19 etc. amongst which the highest giving accuracy model will be selected based on their performance index.

IV. Result and Discussion

The Convolutional Neural Network algorithm can be used for digits or texts recognition purpose with its different architectures. The performance is based on the efficiency of each architecture to recognize. Many similar methods can be deployed for digits or texts recognition purpose with different features, extraction process and different classifiers. The graph in the fig. shows the performance of the proposed method and other methods algorithms namely GoogLeNet, ALexNet, VGG 19 which are currently available and described. Out of these the output obtained during the testing phase of these algorithm. It can be seen that GoogleNet has the highest accuracy rate of 92 percent.

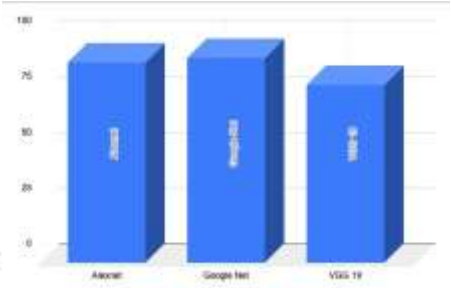

Fig. 2 Accuracy in percentage

The goal of the research done in this paper was to create a model having interactive functionality for kids. The prototype was successfully implemented and tested where it has shown an accuracy of 93 percent after fine tuning some layers of the network and even freezing some layer so to get better training speed. Analysis was done using confusion matrix which gives output based on the correct and incorrect prediction summary, it shows how the classification model gets confused when it makes prediction. The calculation done using confusion matrix is shown in the fig. 2 where 100 inputs have been taken.

Accuracy is calculated by the formula:

(TP + TN) / total = (70+23)/100=0.093

where TP stands for true positive, TN stands for true negative

n= no. of observations

n=100	Predicted: Wrong	Predicted: Correct	
Actual: Wrong	TN=23	FP=3	26
Actual: Correct	FN=4	TP=70	74
	27	73	

Fig. 3 Confusion Matrix

Load Pretrained GoogLeNet Network In this step, first load the pre-trained GoogLeNet network. After that fine-tune some layers of the network.

Even freezing the learning of layers for few iterations which leads to increases in training speed. There is an extraction of features from the input images via the convolutional layer which are then analysed by the classification layer into the

mistakes being made by a classifier however more significantly the sorts of blunders that are being made.

For implementing the model in the first step, data is collected online which consists of different images for the training purpose of our algorithm. It consists of more than a thousand pictures which were further divided in the ratio of 7:3 for training and testing.

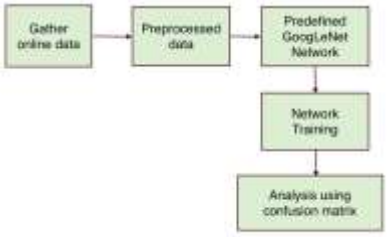

Fig. 4 Procedural flow of proposed model

In the next step, Pretrained network is used to increase the speed by tuning some layers. The pre-processing of the picture was done by using a convolution layer followed by a classification layer. For the GoogleNet algorithm, the image should be of size 224 x 224 But all the data gather cannot have this particular size so we used an augmented image datastore where resizing can be done. After this the trained network is used for classifying the authorized images using predicted labels and forecasting the probabilities of the images that hold these labels. The next step is the use of a confusion matrix which is used to visualize the performance of the algorithm. There is two predicted class Correct and Wrong.

V. Conclusion

In this paper, the variations of accuracies for canvas written digit was observed for 3 different algorithms 100 times. The algorithm used is GoogleNet which is a Convolution Neural Network for the model training that is required in training and practice phase of the model because it gives the highest accuracy based on the research, the digits or letters images as a feature vector and GoogleNet, Alex Net, VGGNet as a classifier for canvas written digits or letters recognition. The stored database is used in the model for evaluating the experiments. The accuracy graph was generated for 3 algorithms. From the results, it can seem that the experiment result achieved 93 percent recognition accuracy by using GoogleNet.

Acknowledgment

The authors would like to thank to our guide Prof. Sumedh Pundkar and all the Computer Science and Technology department for the constant support and encouragement they have given. We are very thankful for all the faith they had in us.

References

[1] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.

[2] C. Zhang and W. Xu, "Neural networks: Efficient implementations and applications," in *2017 IEEE 12th International Conference on ASIC (ASICON)*, pp. 1029–1032, IEEE, 2017.

[3] V. Sze, Y.-H. Chen, T.-J. Yang, and J. S. Emer, "Efficient processing of deep neural networks: A tutorial and survey," *Proceedings of the IEEE*, vol. 105, no. 12, pp. 2295–2329, 2017.