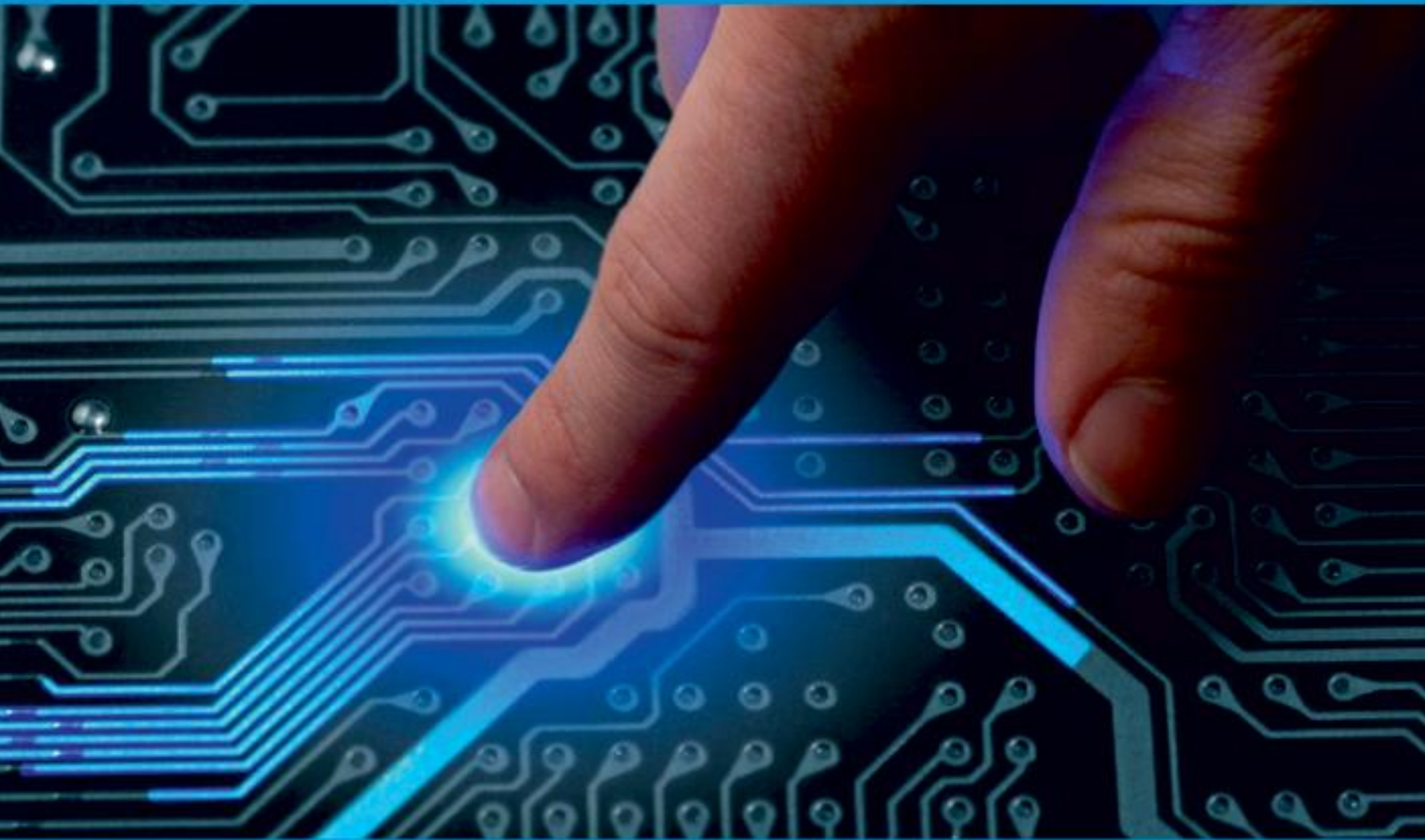
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 5, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105309 |

Literature Survey on Different CNN models

Prof. Sharda Thete[1] , Harshwardhan Ahire[2] , Mariya Ansari[3]

Professor ,Computer Engineering Department, ISB&M College of Engineering Nande, SPPU, Pune, India [1]

Student ,Computer Engineering Department, ISB&M College of Engineering Nande, SPPU, Pune, India [2]

Student ,Computer Engineering Department, ISB&M College of Engineering Nande, SPPU, Pune, India [3]

ABSTRACT: Convolutional Neural Networks (CNNs) as whole have made some pretty big changes, or we could say revolutionized the domain of computer vision and image processing , by showing outstanding performance metrics in various tasks such as semantic segmentation, image classification and object detection. The aim of this literature survey is to discuss various architectures and Application of CNN, Such as Alexnet , VGGNET , GoogLeNet (Inception) , ResNet , DenseNet . By studying and analyzing a collection of research studies , this survey is made in regards to make better decisions regarding the choice of CNN models and their impact on the Field of Deep Learning.

KEYWORDS: Convolutional Neural Networks; ResNet; Alexnet; MNIST

I. INTRODUCTION

CNN models are robust highly efficient when it comes to image recognition, etc With this survey, we strive to further our knowledge regarding different aspects and architectures of CNN as well as study aspects from models made by different researchers in the process.

Background and Purpose: There are a number of reasons behind the rapid advancement of CNN models. First, the ability to train deeper and more sophisticated architectures has been made possible by the availability of enormous amounts of labelled training data, such as ImageNet. Second, improvements in technology have made it possible to train and use CNN models quickly, including the use of graphics processing units (GPUs) and specialized accelerators. Last but not least, advanced architectural layouts and training methods have improved CNNs' performance and generalization powers even more.

II. LITERATURE SURVEY

1) **AlexNet**: Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton proposed a paper on AlexNet as a deep CNN architecture that achieved a very impressive top-5 error rate of 15.3% in the (ILSVRC) 2012. It helped make the use of convolutional layers, ReLU activation, and GPU acceleration mainstream in deep learning.it also introduced LRN concepts.

2) **VGGNet**: Karen Simonyan and Andrew Zisserman proposed a paper on VGGNet as a deep CNN architecture as a simple yet deep Architecture. its main components consists of several convolutional layers used with very small filters/kernels. VGGNet achieved a pretty competitive performance in the ILSVRC 2014.

3) **GoogLeNet (Inception)**: Christian Szegedy et al proposed a paper on GoogLeNet that introduced the concept of inception modules, it consists of "multiple parallel convolutional layers" with different kernel sizes which promote its efficiency and accuracy .The goal of the model was to achieve a higher accuracy while reducing the number of parameters to address the tradeoff between computational efficiency and model depth.

4) **ResNet**: Kaiming He et al proposed a paper on ResNet as a deep CNN architecture that introduced "residual connections", which allowed the training of very deep networks .it helped resolve the problem as vanishing gradients as well as show promising generalization outputs due to "shortcuts.

5) **DenseNet**: Gao Huang et al proposed a paper on DenseNet that introduced the concept of densely connected blocks, where each layer is directly connected to every other layer in a feed-forward manner . It allows feature reuse and helps minimize the vanishing gradient problem. Densenet has shown promising results on all sorts of data making it versatile in DL operations.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105309 |

A. TECHNOLOGY USED.

1)Tensor Flow: A popular open-source machine learning framework used for building and training deep learning models.

2)Keras: It is a high-level neural networks API that is integrated with TensorFlow. It provides an easy-to-use interface for defining and training deep learning models.

3)Convolutional Neural Networks (CNN): The code implements a CNN architecture for image classification tasks. CNNs are a class of deep learning models particularly effective for image-related tasks.

4)MNIST dataset: The code uses the MNIST dataset, which is a widely used benchmark dataset for handwritten digit recognition. It consists of 60,000 training images and 10,000 test images of handwritten digits from 0 to 9.

Operating System:
Windows or Linux
Google Colab(Optional)
Python 3.9
IDE (of your choice)
Jypter notebook (if preferred)

Hardware:
Processor (CPU): A multi-core processor like Intel Core i5 or AMD Ryzen 5
RAM: 8 to 16 is ideal but we have run colab with 4gb of ram so bare minimum is 4
Storage: if we decide to use Google colab , we don't need much storage as everything is implemented on the cloud still we can choose to opt for 500gb harddrive or SSD
GPU:A dedicated GPU is always preferred but to prove we can run and build a project on bare minimum hardware we have chose to use a chromebook for all our testing using google colab and successfully build a model and ran all the necessary test , regaredless a gpu is always preferred , would recommend a NVIDIA GTX 1060 for an entry level system

III. METHODOLOGY

As we conducted the Literature survey , we have come to the conclusion to use , Res-net as our proposed architecture, we plan to implement a simple yet complex model , we will use the MNIST Dataset for our project to classify handwritten digits 0-9 . Our plan is to implement all the best practices we know to the said model to achieve a very accurate model with very less loss , which will avoid overfitting with the help of data Augmentation as well as schedule learning

For our res-net model we followed the following steps

Data Pre processing: this step included processing the input images and applying batch normalization technique
Defining our res-net blocks: in this step we are defining our resnet block as well as how many layers as well as specified the dropout rate with the help of strides sizes.
Use of data augmentation: By using Random rotation and zoom in our architecture we introduced ways to further increase our chance to generalize our model and avoid over-fitting .
Implementing schedule learning tuning: the right schedule learning is one of our main focus in this project , we tried many approaches and tested multiple time to come out with a satisfactory result regarding our resnet architecture learning rate which helped us further our accuracy and minimize our losses
Evaluating our model on test data: in the end we evaluated our model on testing set on the mnist dataset

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105309 |

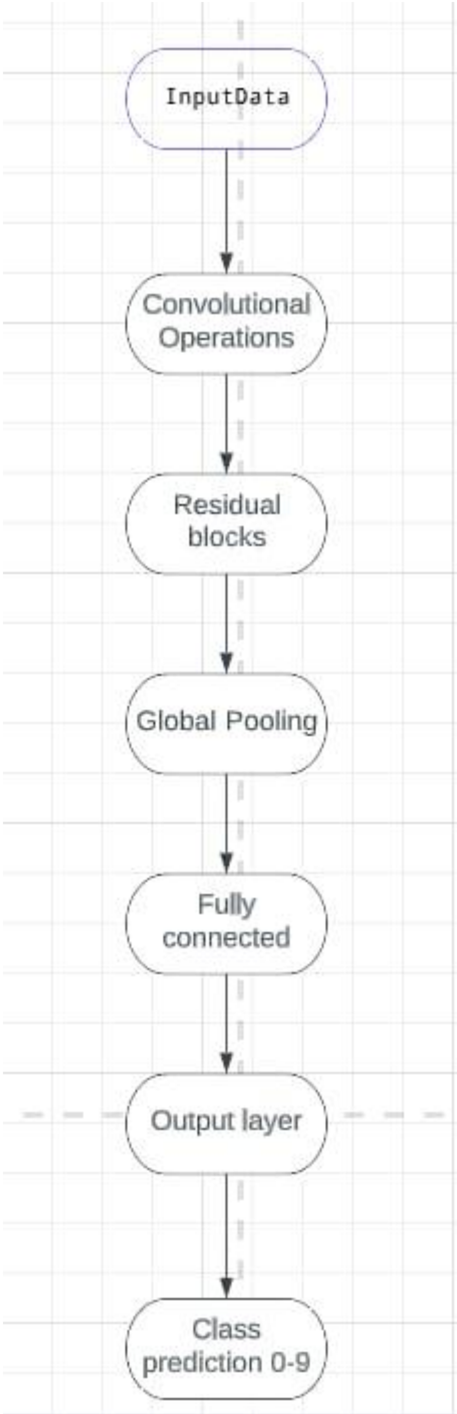


Fig 1. DFD Model of our Resnet model

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105309 |

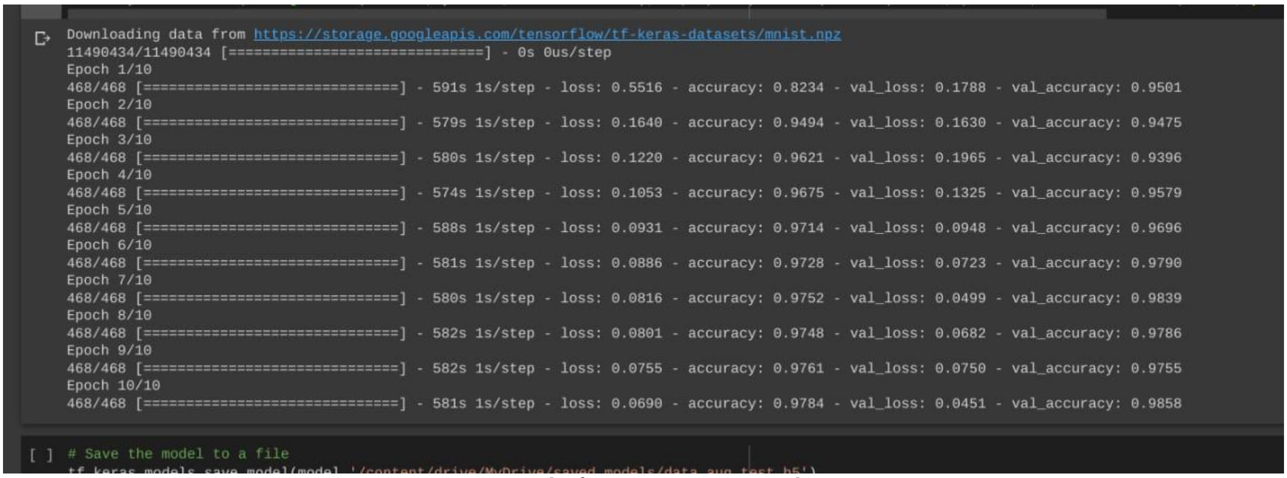


Fig 1. Result of our testing over 10 epochs

IV. CONCLUSION, DISCUSSION AND FURTHER SCOPE

Resnet seems to perform significantly better allowing us to reach 99.01% accuracy on our initial model but we are sure to improve that with given time and more implementation and tuning of our hyper parameters such as learning rate. we have different fundamental approaches to training and testing a model. Readers are urged to explore these sets of architectures on their own as they are fascinating to study and understand. In conclusion, we as a whole have decided to use Res-net as a architecture for our CNN model .
We have also shared our initial result regarding the model proposed initially. we have also Explored Different studies done by people on architectures , these studies have helped us mold our understanding of our domain more , and helped us choose the right model architecture for our exploration of Handwritten Digit Recognition
We are further going to tune our hyper parameters and achieve better results and test it on unseen data to validate our hypothesis

REFERENCES

1. Krizhevsky, Alex & Sutskever, Ilya & Hinton, Geoffrey. (2012). ImageNet Classification with Deep Convolutional Neural Networks. Neural Information Processing Systems. 25. 10.1145/3065386. S.K. Dhurandher, S. Misra, M.S. Obaidat, V. Basal, P. Singh and V. Punia,'An Energy-Efficient On Demand Routing algorithm for Mobile Ad-Hoc Networks', 15 th International conference on Electronics, Circuits and Systems, pp. 958-9618, 2008.
2. Simonyan, K. and Zisserman, A., "Very Deep Convolutional Networks for Large-Scale Image Recognition", <i>arXiv e-prints</i>, 2014. doi:10.48550/arXiv.1409.1556. AlGabri Malek, Chunlin LI, Z. Yang, Naji Hasan.A.H and X.Zhang ,' Improved the Energy of Ad hoc On- Demand Distance Vector Routing Protocol', International Conference on Future Computer Supported Education, Published by Elsevier, IERI, pp. 355-361, 2012.
3. C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015, pp. 1-9, doi: 10.1109/CVPR.2015.7298594.
4. K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.
5. G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 2261-2269, doi: 10.1109/CVPR.2017.243.

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.379

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details