COVID-19 Detection from CT lungs dataset using Attention Augmented Capsule Network

A Project Report

Submitted in Partial Fulfilment of the Requirements for the Award of

Degree of

Master of Technology

In

COMPUTER SCIENCE AND ENGINEERING

Submitted by:

B. LAKSHMI VARA PRASAD
Roll.no - 21CSM1R14
MTech CSE

Department of Computer Science and Engineering

National Institute of Technology Warangal

TABLE OF CONTENTS

1.ABSTRACT

COVID-19 is an infectious disease caused by the Coronavirus. Coronaviruses are a large family of viruses that cause infection ranging from the common flu or cold to more severe disease that primarily affects the lungs and leads to death in the severe stage. Primary identification and isolating are the key to stop the spreading of this disease. Detecting COVID-19 disease is mainly carried out using methods like Coronavirus screening detection strips, COVID-19 antibody detection kits and CT scanning detection. However, in all these methods there exists the problem of low efficiency as these reports needed to be examined by a medical professional, which might not be possible at some times. To overcome this problem, convolution neural network model by using lung CT images is proposed which can classify covid and normal case.

In this paper, we propose a capsule network model in which we use a two-dimensional relative self-attention mechanism to execute convolution. It can be utilized to increase COVID 19 case classification accuracy. The combination of traditional medical imaging diagnosis and deep learning technology helps medical personnel to make more rapid, accurate and effective diagnosis.

2.INTRODUCTION

The first identification of outbreak of disease COVID19 was in Wuhan, China 2019. COVID19 is a respiratory dysfunction which caused due to existence of SARS-CoV-2 virus in human lungs. The risk of COVID-19 is stated as very high at the global level by WHO in February 2020. 190 million cases and more have been reported in 192 countries and regions around the world, and more than 4 million patients have died until July 2021. World Health Organization declared the outbreak of COVID19 as Public Health Emergency of International Concern on 30 January 2020, and it later declared as pandemic on 11 March 2020.

Identification of this virus in human body can be done by RTPCR tests which is a microbiological test, that highly needs expertise Pathologist which may not be convenient in the hour of emergency. Another method of identification is by computed tomography (CT) scans of lungs. In the diagnosis and therapy of COVID-19, chest computed tomography is preferred over chest radiography. Physicians need to know whether the CT symptoms will progress, stay the same, or regress in the coming days based on the initial diagnosis. Slight findings in the early stage, i.e., less than 5 days from onset, peak findings in the middle stage, i.e., around 5 to 13 days from onset, and recovered findings in the late stage, i.e., more than 14 days from onset, have all been recorded during COVID-19. The idea of ground-glass opacity is used to detect lung illness (GGO) [1]. Hence, we can differentiate the GGOs features of COVID-19 affected patient and normal subjects by using the deep learning approach on lung CT scans.

The CT scan images used in this, is an open-source dataset of 2 classes namely COVID-19 lung CT frames 1252, and NON COVID19 lung CT frames 1229 [2]. We imposed preprocessing techniques like converting RGB CT images to greyscale, resizing the images and normalization of pixel's values.

Dataset Characteristics:	Covid	Non-Covid	Count
Number of Classes:			2
Number of Training samples:	877	860	1737
Number of Testing samples:	375	369	744

TABLE2.1 Description about our dataset

2.1 DATASET:

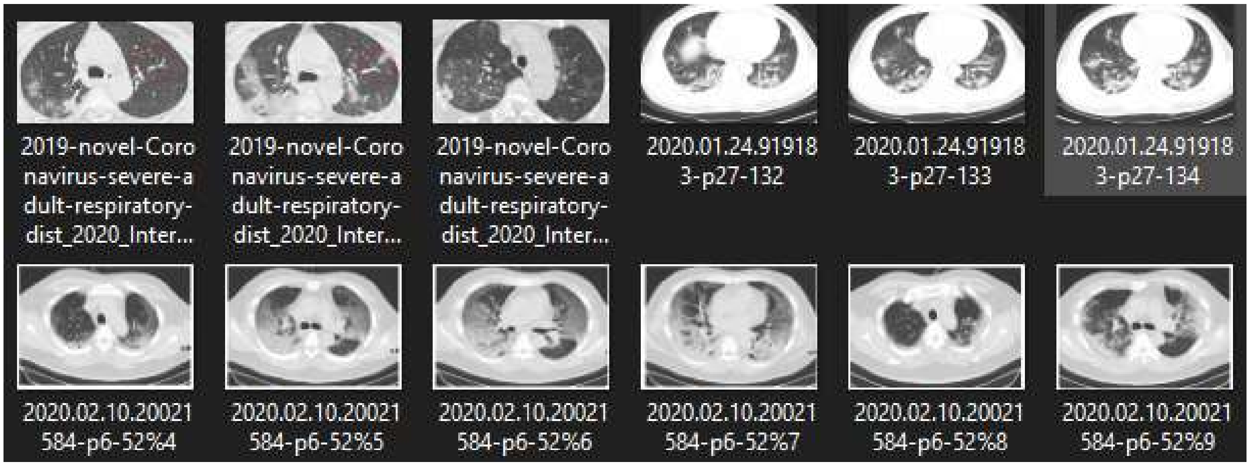


Fig2.1 CT Images of COVID19 Dataset

3.RELATED WORK

Researchers have assessed some prediction models for diagnosis purposes of COVID-19 on lung CT [3-8].

- Rakibul Islam. [3] introduced a LeNet5 CNN based Architecture to detect COVID19 from CT Images.
- Loey et al. [4] introduced an artificial intelligent methodology by a deep transfer learning-based concept on CT digital images.
- Fang et al. [5] proposed a system for predicting COVID19 by differentiating the sensitivity of RT-PCR and lung CT. They identified that lung CT is more sensitive for COVID-19 finding.
- Xie et al. [6] also revealed that sensitivity calculation of lung CT took a higher position over RT-PCR in the prediction of COVID-19 at an early stage.
- Horry et al. [7] imposed a deep learning based COVID19 detection system through transfer learning on multiple types of medical image (x-ray, CT, ultrasound).
- Recently, some deep learning methodologies have been introduced in the diagnosis purposes of pneumonia in chest CT images related to COVID-19 [8].

Now we have the direct lung CT images of COVID-19 affected subjects and not affected subjects, and our aim is to introduce a Capsule Network architecture based COVID-19 classifier model based on COVID-19 lung CT findings.

4.PROPOSED MODEL

4.1 CAPSULE NETWORK

Geoffrey Hinton and his team published two papers that introduced a completely new type of neural network-based model referred to as capsules. The team published an algorithm, called dynamic routing between capsules, that allows to train such a network.

A capsule is a group of neurons. The activity vector of those neurons represents the instantiation parameters of a specific type of entity such as an object or part of an object. Length of the activity vector represent the probability that the entity exists, and its orientation represent the instantiation parameters. At one level the Active capsules make predictions, via transformation matrices, for the instantiation parameters of higher-level capsules. Whenever multiple predictions agree, a higher level capsule becomes active.

Convolution neural networks have made remarkable success in the field of computer vision. They are one of the reasons deep learning is so popular today. Despite their success, CNNs have some fundamental drawbacks. A convolutional layer is the major component of a CNN, and its function is to discover essential features in an image. Deeper layers will learn to recognize simple characteristics like edges and color gradients, while higher layers will combine simple features to create more complex features.CNN requires massive amounts of data to learn. The layers in CNN reduce the spatial resolution and the output of the networks never change even with a small amount of change in the inputs. It cannot directly relate to the relation of parts and requires additional components. This is where Capsule Networks comes into play and overcomes all the drawbacks that are present on CNN.

Capsule Networks (CapsNet) are networks that may retrieve spatial information and other critical aspects in order to avoid the information loss experienced in CNN's pooling operations. The main distinction between a capsule and a neuron is that a capsule produces a vector with a direction as an output. When we change the orientation of an image, the vector moves in the same direction, whereas the output of a neuron in CNN is a scalar quantity that provides no direction information.

4.2 DYNAMIC ROUTING

Dynamic routing replaces the max-pooling for delivering the information of capsules. The vector-output capsule, in which the vector's direction indicates the entity's attributes and the length of the vector represents the entity's probability. All child capsules in the lower layer i have identical connection probability b_{ij} with the following layer j at the start of dynamic routing. The connection probability in the low-level capsule layer is normalized by the coupling coefficient c_{ij}. The anticipated capsule and the coupling coefficient are then multiplied by the parent capsule. The agreement between the parent capsules and the prediction capsules for k iterations also updates the b_{ij} and c_{ij}. A more accurate coupling coefficient can be found over multiple iterations.

4.3 SQUASH FUNCTION

The length of output vector of a capsule represents the probability that the entity represented by the capsule is present in the current input. We therefore use a non-linear "squashing" function to ensure that short vectors get shrunk to almost zero length and long vectors get shrunk to a length slightly below 1.

$$\mathbf{v}_j = \frac{||\mathbf{s}_j||^2}{1+||\mathbf{s}_j||^2}\frac{\mathbf{s}_j}{||\mathbf{s}_j||} \tag{1}$$

4.4 LOSS FUNCTION

$$L_c = T_c\left(\max(0, m^+ - \|v_c\|\right)^2 + \lambda(1-T_c)(\max\left(0, \|v_c\| - M^-\right)^2$$

For Capsule networks, the loss function is a linear combination of the margin loss function and the reconstruction loss. The first part of the margin loss is calculated for correct prediction; it is 0 if the output has a probability of greater than 0.9, and non-zero otherwise. The second element is used to calculate inaccurate predictions: 0 when an incorrect prediction with a probability of less than 0.1 is predicted, and non-zero otherwise.

4.5 ATTENTION AUGMENTED CONVOLUTION

Attention-augmented Convolution is a two-dimensional relative self-attention mechanism that can be used instead of convolutions as a stand-alone computational primitive for image classification. As with Transformers, it uses scaled-dot product attention and multi-head attention [9]. It works by concatenating convolutional and attentional feature map.

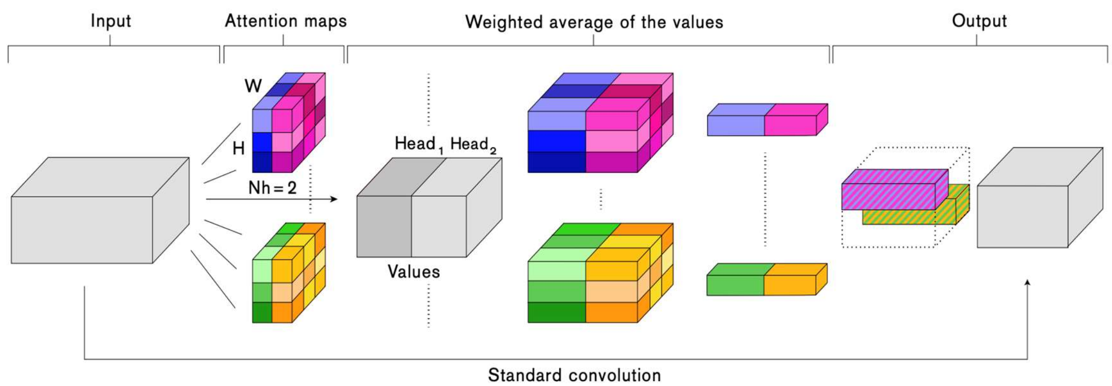


Figure 2. **Attention-augmented convolution**: For each spatial location (h, w), N_h attention maps over the image are computed from queries and keys. These attention maps are used to compute N_h weighted averages of the values V. The results are then concatenated, reshaped to match the original volume's spatial dimensions and mixed with a pointwise convolution. Multi-head attention is applied in parallel to a standard convolution operation and the outputs are concatenated.

Fig 4.1 Attention Augmented Convolution Architecture

4.5 ARCHITECTURE OF PROPOSED MODEL

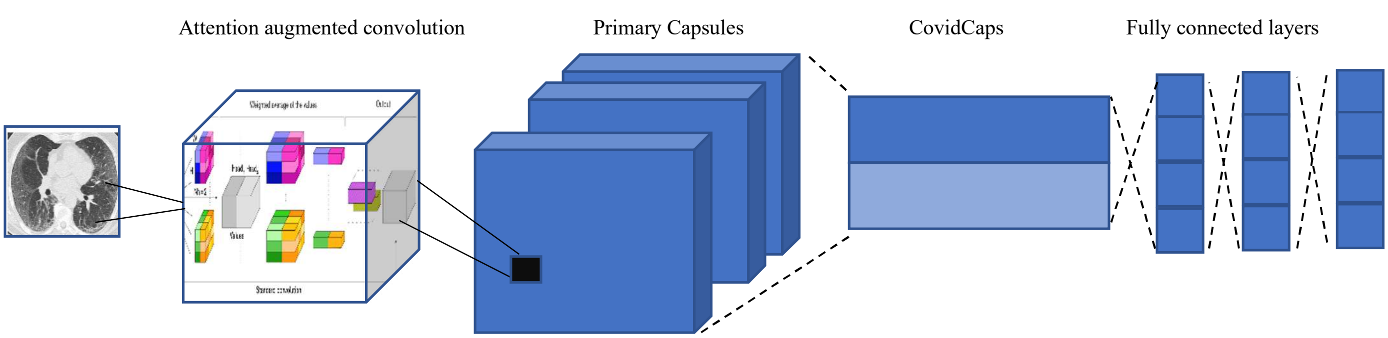


Fig 4.2 Overall proposed capsule network architecture

We are introducing Attention Augmented convolution operation to normal convolution as novelty to our capsule network. Best results were obtained when combining both convolutions and self-attention. We therefore propose to augment capsule network with this self-attention mechanism by concatenating convolutional feature maps with a set of feature maps produced via self-attention [10].

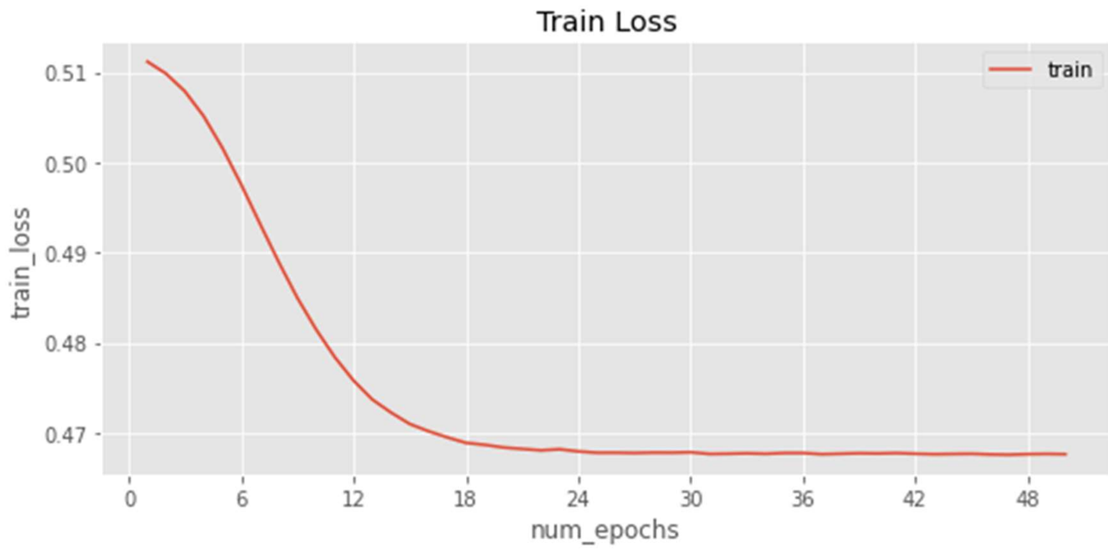
5. EXPERIMENT RESULTS

5.1 CAPSULE NETWORK: The Capsule Network Architecture shown below is the base Architecture. Attention Augmentation Convolution Layer is to be introduced before the convolution layers in the Capsule Network.

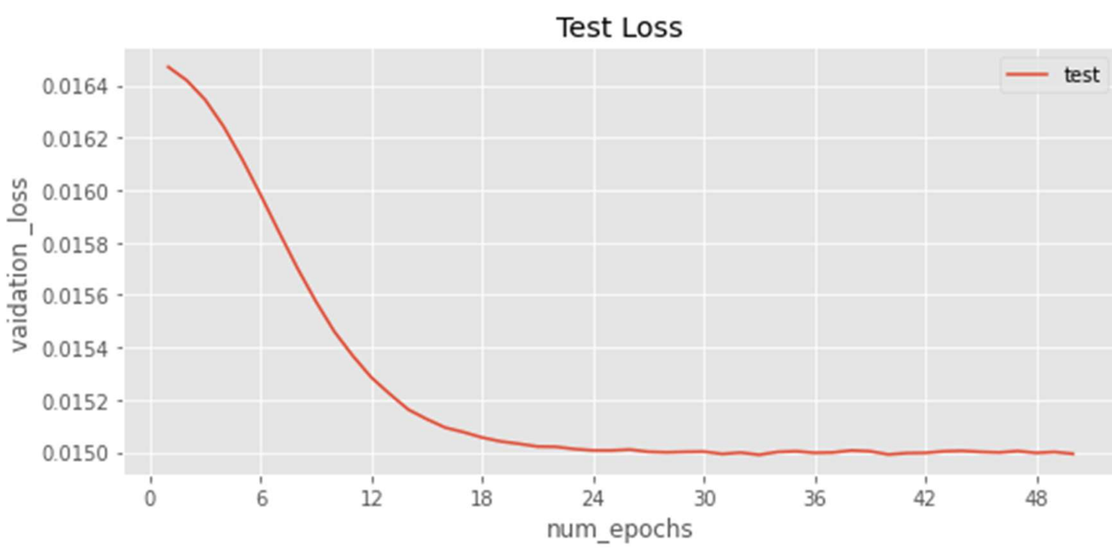
```
summary(CapsuleNetwork().cuda(), input_size=(1, 28, 28))

----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv2d-1          [-1, 256, 20, 20]          20,992
            Conv2d-2             [-1, 32, 6, 6]         663,584
            Conv2d-3             [-1, 32, 6, 6]         663,584
            Conv2d-4             [-1, 32, 6, 6]         663,584
            Conv2d-5             [-1, 32, 6, 6]         663,584
            Conv2d-6             [-1, 32, 6, 6]         663,584
            Conv2d-7             [-1, 32, 6, 6]         663,584
            Conv2d-8             [-1, 32, 6, 6]         663,584
            Conv2d-9             [-1, 32, 6, 6]         663,584
     CapsuleLevel-10             [-1, 1152, 8]               0
     CapsuleLevel-11               [-1, 2, 16]               0
           Linear-12                  [-1, 512]          16,896
             ReLU-13                  [-1, 512]               0
           Linear-14                 [-1, 1024]         525,312
             ReLU-15                 [-1, 1024]               0
           Linear-16                  [-1, 784]         803,600
          Softmax-17                  [-1, 784]               0
================================================================
Total params: 6,675,472
Trainable params: 6,675,472
Non-trainable params: 0
----------------------------------------------------------------
```

After training and testing the loss for each epoch is shown below.

5.2 TRAINING:



5.3 TESTING:



6. CONCLUSION

We can further reduce the loss values by increasing the epoch number. The proposed model trained for 50 epochs. And further performing data augmentation and applying different datasets the model can perform better.

7.REFERENCES

[1] Y. Pan, H. Guan, S. Zhou, Y. Wang, Q. Li, T. Zhu, et al., "Initial CT findings and temporal changes in patients with the novel coronavirus pneumonia (2019-nCoV): a study of 63 patients in Wuhan China", European radiology, pp. 1-4, 2020.

[2] X. Yang, X. He, J. Zhao, Y. Zhang, S. Zhang, and P. Xie," COVIDCT-dataset: a CT scan dataset about COVID-19," arXiv, pp.arXiv-2003, 2020.

[3] M. R. Islam and A. Matin," Detection of COVID 19 from CT Image by The Novel LeNet-5 CNN Architecture," 2020 23rd International Conference on Computer and Information Technology (ICCIT), 2020, pp. 1-5, doi: 10.1109/ICCIT51783.2020.9392723.

[4] M. Loey, G. Manogaran, and N.E.M. Khalifa," A deep transfer learning model with classical data augmentation and cgan to detect covid19 from chest ct radiography digital images," Neural Computing and Applications, pp. 1-13, 2020.

[5] Y. Fang, H. Zhang, Y. Xu, J. Xie, P. Pang, and W. Ji," CT manifestations of two cases of 2019 novel coronavirus (2019-nCoV) pneumonia," Radiology, vol. 295, no. 1, pp. 208-209, 2020.

[6] X. Xie, Z. Zhong, W. Zhao, C. Zheng, F. Wang, and J. Liu," Chest CT for typical 2019-nCoV pneumonia: relationship to negative RTPCR testing," Radiology, pp. 200343-200343, 2020.

[7] M.J. Horry, S. Chakraborty, M. Paul, A. Ulhaq, B. Pradhan, M. Saha, and N. Shukla," COVID-19 detection through transfer learning using multimodal imaging data," IEEE Access, vol. 8, pp. 149808-149824, 2020.

[8] A. Bernheim, X. Mei, M. Huang, Y. Yang, Z.A. Fayad, N. Zhang, K. Diao, B. Lin, X. Zhu, K. Li, and S. Li," Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection," Radiology, pp. 200463.

[9] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin,” Attention Is All You Need,” arXiv:1706.03762, 2017.

[10] Irwan Bello, Barret Zoph, Ashish Vaswani, Jonathon Shlens, Quoc V. Le, “Attention Augmented Convolutional Networks,” arXiv.1904.09925, 2019.