H3DNN: 3D DEEP LEARNING BASED DETECTION OF COVID-19 VIRUS USING LUNGS COMPUTED TOMOGRAPHY

ABDULLAH AMAN KHAN[1], SIDRA SHAFIQ[2], RAJESH KUMAR[1], JAY KUMAR[1], AMIN UL HAQ[1]

[1]School of Computer Science and Engineering, University of Electronic Science and Technology of China, China
[2]Department of Computer Science, The Women University Multan, Punjab, Pakistan
E-MAIL: abdkhan@std.uestc.edu.cn, sid.siduuu@gmail.com, rajakumarlohano@gmail.com, jay@std.uestc.edu.cn, khan.amin50@yahoo.com

Abstract:

With the rapid spread of the novel COVID-19 virus, there is an increasing demand for screening COVID-19 patients. Typical methods for screening coronavirus patients have a large false detection rate. An effective and reliable screening method for detecting coronavirus is required. For this reason, some other reliable methods such as Computed Tomography (CT) imaging is employed to detect coronavirus accurately. In this paper, we present a 3D-Deep learning based method that automatically screens coronavirus patients using 3D volumetric CT image data. Our proposed system assists medical practitioners to effectively screen out COVID-19 patients. We performed extensive experiments on two datasets i.e., CC-19 and COVID-CT using various state-of-the-art 3D Deep learning based methods including 3D ResNets, C3D, 3D DenseNets, I3D, and LRCN. The results of the experiments show the competitive effectiveness of our proposed approach.

Keywords:

COVID-19; Coronavirus; 3D deep learning; Deep learning; Artificial intelligence

1. Introduction

Due to the rapid spread of the novel COVID-19 virus, the Artificial Intelligence (AI) research community explored many ideas for diagnosing lung infection by analyzing Computed Tomography (CT) imaging [1-3]. The initial reason behind this attention towards chest CT imaging was the lack of nucleic acid-based CoVID-19 detection kits. After observing the high false-negative rate of nucleic acid test, clinical practitioners started to prefer screening of COVID-19 patients via chest CT imaging [4]. Specifically, for early-stage detection, CT imaging offers a glass-like clarity to highlight lesions of the lung. However, according to the radiologists, clinical screening of COVID-19 is still unsatisfactory [5-7]. Therefore, automated screening with the help of AI can assist clinical practitioners to improve screening accuracy.

Over the past few months, many deep learning based screening approaches have been proposed to detect infected lesions from 2D CT imaging [7-9]. Unfortunately, they either demand a high percentage of annotated areas of lesions or lack of interpretability. Generally, the standard pioneer input for the classifiers is either patch-based or lesion-based. However, the 3D volume of CT imaging is still not well explored. The exploitation of 3D volume is under consideration to improve the screening accuracy. Compared with classic 2D CT images, the generated 3D volume of CT usually contains hundreds of slices which are more difficult to analyze even for the clinical practitioners.

Previously proposed deep learning models, such as [5][6][10], cannot be directly applied for 3D imaging. Therefore, there exists a need to design an automated model to detect lung infection caused by COVID-19 from 3D chest CT imaging.

In this paper, we propose a Hybrid 3D Deep Neural Network model (H3DNN) to classify chest CT imaging. The model is developed by ensembling the Inflated inception (I3D) and 3D ResNet 50 to build a common architecture for capturing the Spatio-temporal dimension including the inception block. Unlike previous approaches, H3DNN can semantically generate deep 3D samples with permutation-invariance to improve network accuracy. The main contributions of our work are given as follows:

- We propose an automated 3D deep learning model (H3DNN) to classify 3D chest CT imaging to screen out infected patients.
- Our approach can easily diagnose the early stage of COVID-19 patients by considering Spatio-temporal features and constructing 3D filters.
- We conducted extensive experiments on two available datasets. The results of the experiments boldly show the significance of H3DNN.

2. Proposed Method

In this section, we elaborate on our proposed 3D deep learning network. 3D deep learning based models efficiently make use of Spatio-temporal information that utilizes 3D convolutions and other relevant blocks. Such networks have achieved higher performance in many Artificial Intelligence (AI) applications. AI based automation has already proven itself in other domains [11-13].

The proposed network can be considered as a hybrid of a single stream I3D [14] and a 3D ResNet 50 [15] network model. 3D deep learning models bootstraps 3D convolution filters based on 2D convolutions. I3D utilizes a receptive field for an artificial neural network (ANN). However, when a temporal dimension is included it requires finding an optimal receptive field.

Inflated inception: I3D utilizes the inception block as shown in Figure 1. The main motivation of this module is to allow the network to grow wider instead of deeper. The I3D network model is represented by the block C in Figure 2. The I3D network starts with a 3D convolution layer with a stride of 2 followed by a MaxPool layer having a stride of (1,2,2). Further, two convolution layers are followed by a MaxPool layer having a stride of (1,2,2). The resultant of this Maxpool layer is then fed to two inception blocks. Further, another MaxPool layer is followed by five more inception blocks. The resultant is then forwarded to a MaxPool with a stride of 2 attached to two more inception blocks. Finally, an average pooling is carried out and the resultant is passed on to a convolution layer followed by a fully connected layer.

3D ResNet: A unitary ResNet block consists of two convolution layers fed to a batch normalization (BN) and ReLU [16] layers. A bypass connects the top block to the layer before ReLu. The ResNet 3D block extracts more valuable features as compared to 2D. The ResNet architecture is represented by block A in Figure 2. The first convolution layer with a stride of 64 is followed by six consecutive convolution layers. Further, eight more convolution layers are added with a stride of 128. The resultant is forwarded to a block of twelve convolution layers with a stride of 256 followed by a block of six convolution layers. Finally, average pooling is carried out followed by a 400d fully connected layer.

Hybrid model (H3DNN): I3D and 3D ResNet 50 are capable of extracting prominent features from Spatio-temporal data. We feed both the network models with a sequence of CT scan images. In our case, we feed a series of 60 and 35 CT scan images. The reason for choosing these numbers is that some of the patients' CT scan images in the datasets contain a minimal of 60 and 35 CT scan slices/images. The input CT scan sequence was first resized $224 \times 224 \times 3$ to match the input profile of the targeted 3D feature extraction network. The slices are fed to I3D and 3D ResNet block separately. Each feature extraction extracts Spatio-temporal features. Further, each 3D block i.e. Block A and Block C in Figure 2 are used to train fully connected layers [17]. The probability output of the fully connected layers is then added and the resultant fed to a Softmax layer [18], which, finally classifies the suspect as positive or negative.

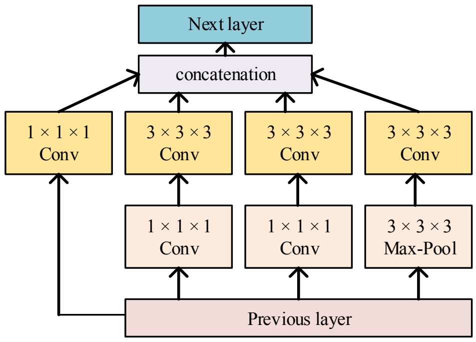

Fig. 1 Inception block of I3D network. The parameters from the previous layer are the input to this block where the next layer represents the output of the inception block.

3. Experiments and results

We performed extensive experiments on two publically available CC-19 [8] and COVID-CT [19] datasets. First, we provide some details about the CC-19 and COVID-CT datasets. We used these datasets as they contain the required series of CT scan images for 3D analysis.

3.1. Datasets

CC-19 dataset: contains about 34,000 CT scan images for 89 subjects. Out of 89, 68 subjects are confirmed COVID-19 patients. CC-19 dataset contains a huge amount of data. The data was recorded on a day to day basis for every subject. CC-19 dataset was collected from various hospitals in Sichuan, China.

COVID-CT: This dataset contains CT scan images for 349 scans from 216 patients and 463 CT scans from non-covid subjects. Both of these datasets were confirmed by professional radiologists.

3.2. Experiment setup

All the experiments were performed on an Intel Xeon 40 core processor equipped with Ubuntu 20.04.1 LTS operating system, 128 GB RAM, hard drive 6 Gbps data bandwidth, and 08 Tesla K80 graphic processing units

(GPU). All the experiments were carried out using Keras with a TensorFlow backend.

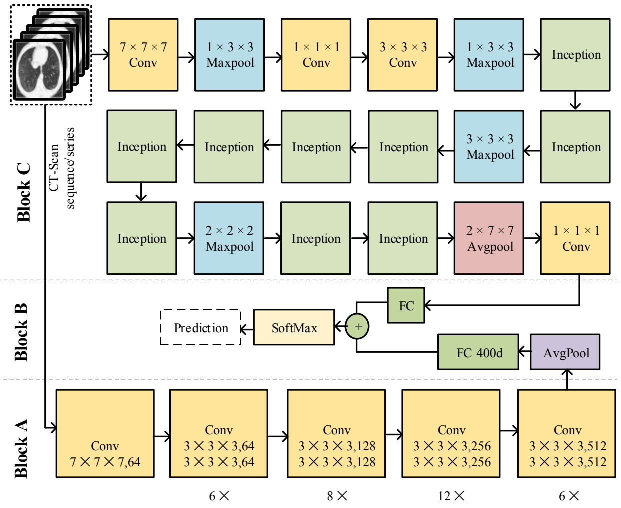

Fig. 2 Our proposed hybrid network model.

We trained I3D and 3D ResNet 50 from scratch. Further, the individual weights of the best-trained models were saved. 3D networks are hard to train as they require more computational resources, the batch size for training was kept 2. These models were trained using Adams' optimizer with a learning rate of 10^{-5} and a decay rate of 10^{-6}. We use these smaller values as we trained the network from scratch. Moreover, we used an early stopping machoism with patience of 5 with 1000 epochs at maximum. For 3D ResNet 50, the regularization factor was set to 2.5×10^{-2}.

3.3. Results

We performed comprehensive experiments to validate the proposed model. The results of the experiments are shown in Table 1. The accuracy is computed as $Acc = (TP + TN)/(TP + TN + FP + FN)$. Where, TP, TN, FP, and FN are the true positive, true negative, false positive, and false negative respectively.

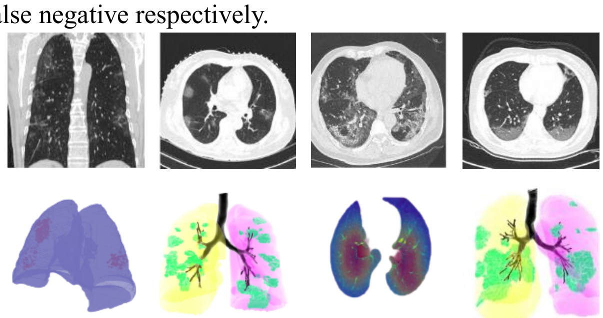

Fig.3 3D visualization with various ways of our proposed model. The top left image shows the XZ plane of the 3D volume.

From Table 1, it can be seen that our proposed model H3DNN showed superior performance. Figure 3 shows some 3D visualization of the 3D CT scans from these datasets. We believe, the reason behind the superiority of our model is that it combines the goodness of both I3D and 3D Resnet 50 deep learning models.

Table 1. Comparison of our proposed technique with state-of-the-art methods. Moreover, S and Acc. Represents the number of CT slices and accuracy respectively. C-CT represents the COVID-CT dataset.

Method	Size	CC-19		C-CT	
		S	Acc.	S	Acc
C3D [20]	150×150	60	0.76	35	0.81
LRCN [21]	150×150	60	0.73	35	0.76
3D Conv [20]	100×100	60	0.76	35	0.75
DenseResNet3D [15]	112x112	60	0.76	35	0.74
DenseNet 3D [15]	112×112	60	0.76	35	0.75
I3D [14]	224×224	60	0.80	35	0.81
R2Plus1D [22]	171×128	60	0.76	35	0.75
3D ResNet 18 [23]	224×224	60	0.76	35	0.77
3D ResNet 34 [23]	224×224	60	0.76	35	0.79
3D ResNet 50 [23]	224×224	60	0.83	35	0.82
3D ResNet 101 [23]	224×224	60	0.78	35	0.77
3D ResNet 151 [23]	224×224	60	0.80	35	0.80
H3DNN	224×224	60	0.85	35	0.84

4. Conclusion

In this paper, we proposed a hybrid 3D deep learning model (H3DNN). H3DNN makes use of I3D and 3D Resnet 50 to screen out COVID-19 patients. The proposed model automatically and effectively detects the COVID-19 patients at a low cost in terms of annotations of CT images. We conducted comprehensive experiments using two available datasets. The results of the experiments reveal the superiority of our proposed method. In the future, we plan to design a full multimedia system for doctors that can effectively segment and point out the infections caused by the COVID-19 virus.

References

[1] D. P. Fan et al., "Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Images," IEEE Trans. Med. Imaging, 2020, doi: 10.1109/TMI.2020.2996645.

[2] R. Kumar et al., "An Integration of Blockchain and AI for Secure Data Sharing and Detection of CT images for the Hospitals," Comput. Med. Imaging Graph.,

2020.
[3] X. Ouyang et al., “Dual-Sampling Attention Network for Diagnosis of COVID-19 from Community Acquired Pneumonia,” IEEE Trans. Med. Imaging, 2020, doi: 10.1109/TMI.2020.2995508.
[4] V. M. Corman et al., “Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR,” Eurosurveillance, 2020, doi: 10.2807/1560-7917.ES.2020.25.3.2000045.
[5] Y. Oh, S. Park, and J. C. Ye, “Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets,” IEEE Trans. Med. Imaging, 2020, doi: 10.1109/TMI.2020.2993291.
[6] S. Roy et al., “Deep Learning for Classification and Localization of COVID-19 Markers in Point-of-Care Lung Ultrasound,” IEEE Trans. Med. Imaging, 2020, doi: 10.1109/TMI.2020.2994459.
[7] X. Ouyang et al., “Dual-Sampling Attention Network for Diagnosis of COVID-19 From Community Acquired Pneumonia,” IEEE Trans. Med. Imaging, vol. 39, no. 8, pp. 2595–2605, 2020.
[8] R. Kumar et al., “Blockchain-Federated-Learning and Deep Learning Models for COVID-19 detection using CT Imaging,” arXiv. 2020.
[9] G. Wang et al., “A Noise-Robust Framework for Automatic Segmentation of COVID-19 Pneumonia Lesions from CT Images,” IEEE Trans. Med. Imaging, 2020, doi: 10.1109/TMI.2020.3000314.
[10] L. Zhou et al., “A Rapid, Accurate and Machine-Agnostic Segmentation and Quantification Method for CT-Based COVID-19 Diagnosis,” IEEE Trans. Med. Imaging, 2020, doi: 10.1109/TMI.2020.3001810.
[11] A. A. Khan, J. Shao, W. Ali, and S. Tumrani, “Content-Aware Summarization of Broadcast Sports Videos: An Audio–Visual Feature Extraction Approach,” Neural Process. Lett., 2020, doi: 10.1007/s11063-020-10200-3.
[12] A. A. Khan, H. Lin, S. Tumrani, Z. Wang, and J. Shao, “Detection and localization of scorebox in long duration broadcast sports videos,” 2020, doi: 10.1117/12.2575834.
[13] S. Tumrani, Z. Deng, A. A. Khan, and W. Ali, “PEVR: Pose Estimation for Vehicle Re-Identification,” in Asia-Pacific Web (APWeb) and Web-Age Information Management (WAIM) Joint International Conference on Web and Big Data, 2019, pp. 69–78.
[14] J. Carreira and A. Zisserman, “Quo Vadis, action recognition? A new model and the kinetics dataset,” 2017, doi: 10.1109/CVPR.2017.502.
[15] Z. Qiu, T. Yao, and T. Mei, “Learning Spatio-Temporal Representation with Pseudo-3D Residual Networks,” 2017, doi: 10.1109/ICCV.2017.590.
[16] Y. Li and Y. Yuan, “Convergence analysis of two-layer neural networks with RELU activation,” 2017.
[17] X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” 2010.
[18] B. D. Ripley, Pattern recognition and neural networks. 2014.
[19] J. Zhao, X. He, X. Yang, Y. Zhang, S. Zhang, and P. Xie, “COVID-CT-Dataset: A CT image dataset about COVID-19,” arXiv. 2020.
[20] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, “Learning spatiotemporal features with 3D convolutional networks,” 2015, doi: 10.1109/ICCV.2015.510.
[21] J. Donahue et al., “Long-Term Recurrent Convolutional Networks for Visual Recognition and Description,” IEEE Trans. Pattern Anal. Mach. Intell., 2017, doi: 10.1109/TPAMI.2016.2599174.
[22] D. Tran, H. Wang, L. Torresani, J. Ray, Y. Lecun, and M. Paluri, “A Closer Look at Spatiotemporal Convolutions for Action Recognition,” 2018, doi: 10.1109/CVPR.2018.00675.
[23] K. Hara, H. Kataoka, and Y. Satoh, “Learning spatio-Temporal features with 3D residual networks for action recognition,” 2017, doi: 10.1109/ICCVW.2017.373.