ResCapsNet: Detection of COVID-19 from Chest X-Ray Images Using ResNet and Capsule Network

Parvin Malekzadeh
Yuxuan Chen
University of Toronto

Abstract

The coronavirus disease (COVID-19) is spreading rapidly and has drastically affecting the health of people over 200 countries. COVID-19 is spreading fast; therefore; a critical step in controlling the epidemic curve is its early detection. While current diagnosis tests in clinics and hospitals require specific requirements, it has been shown that COVID-19 affects the lungs of patients and displays the symptoms of pneumonia, where can be diagnosed from Computed Tomography (CT) scans and X-ray images [1]. It, therefore, necessities the need for detection of COVID from other lung diseases. Deep learning-based algorithms especially Convolutional Neural Networks (CNNs) are among the most interested methods in this regards. However, CNNs-based frameworks suffer from being unable to find the spatial relations between image instances and the need for a large number of data requirement, which is unavailable due to sudden emergence of COVID-19. To tackle the mentioned problems of CNNs, this paper proposes a Capsule Network-based framework, referred to as ResCapsNet, which uses capsule networks as the main body and ResNet50 for feature extraction. The proposed ResCapsNet framework has been performed on a dataset of X-ray images and an accuracy of 98%, sensitivity of 98.07%, specificity of 99.15%, and area under the curve of 99% were resulted. The achieved results demonstrate the advantage of the proposed method in comparison to CNN-based framework proposed in [2] and capsule network-based COVID-CAPS framework proposed in [3]. In order to further improve the performance of ResCapsNet, we applied the proposed ResCapsNet framework on the augmented data and accuracy of 98.9% and sensitivity of 98.99% , and specificity of 98.30% were achieved. The code is available at: https://github.com/sherrychen127/CSC413_project.git.

1 Introduction

The novel Coronavirus Disease (COVID-19), which first appeared in December 2019, is an infectious disease that causes respiratory disease, and it can easily spread if infected cases are not diagnosed and treated in time. Due to the rapid transition rate of the Coronavirus through human interactions, the first step for stopping this epidemic is detection of COVID-19 cases in an early stage. The current method of COVID diagnosis, Transcription Polymerase Chain Reaction (RT-PCR), requires high-performance diagnosis equipments, which are expensive and are not accessible in all countries. Moreover, RT-PCR test is a time consuming process whose result can be reported within 24 hours at the earliest. It has been shown that computed tomography (CT) scans and X-ray images have visual cues that can be used to distinguish COVID-19 infected cases from normal cases [3, 4, 5]. Hence, we propose an automatic COVID-19 detection framework which leverages X-ray images that are readily available in most of the hospitals. This can be used as a quick and cheap alternative for COVID diagnosis.

1.1 Related Works

Since the potentials of X-ray images in COVID-19 detection has been revealed, there have been several recent studies that apply Convolution Neural Networks (CNNs)-based learning methods for automatic COVID-19 detection on CT scans and chest X-ray images [6, 7, 8]. Wang et al. [5] proposes a deep CNN framework, known as COVID-Net, referred to as a CNN model, which is first pre-trained on the ImageNet dataset [9] and then followed by fine-tuning using a dataset of X-ray images, for COVID-19 screening. Wang et al. [5] also contribute an open source dataset of X-ray images, COVIDx, which has the largest number of COVID-19 positive cases in comparison to other COVID dataset . They achieved the overall accuracy and sensitivity from applying COVID-Net on COVIDx data equals 93.3% and 91%, respectively. Sethy et al. [2] also conduct a similar study where a ResNet50 network is trained on X-ray images. In addition to ResNet50, this method also adopts a Support Vector Machine (SVM) classifier to detect the affected X-ray images from the others using the deep features. They reached an accuracy of 95.38%. Although CNNs are commonly used as feature extractor for object detection and classification tasks, they have certain drawbacks: 1- Being unable to capture the spatial relationship between features; therefore, cannot predict the same object when it is rotated, and: 2- The need to a large dataset to be trained on, which is difficult to obtain extensive amount of image data on COVID-19 due to its recent occurrence. Geoffrey E. Hinton et al. [10] present the Capsule Network (CapsNet) as an alternative model of CNNs, which is made of a group of related neurons indicating various features of an entity, such as posture and texture features. CapsNet is capable of capturing image features and spatial relationships using neuron packaging while reducing its dependency on a large dataset. Parnian Afshar et al.[3] propose a capsule network-based methodology, referred to as COVID-CAPS, to detect COVID-19 cases from X-ray images. They used 4 convolutional layers followed by 3 capsule layers and achieved an overall accuracy of 95.7% and sensitivity of 90%. These results prove the feasibility of capsule networks for COVID-19 detection.

1.2 Contributions

Due to the mentioned advantages of CapsNet over CNNs, in this study, we use capsule networks for COVID-19 detection based on X-ray images. We propose a model, referred to as ResCapsNet, which takes the capsule network as its main framework and uses ResNet50 as the feature extractor. For comparison purpose, we use the same available X-ray dataset [5] used by COVID-CAPS [3] and COVID-Net [5]. The achieved results prove the efficacy of our proposed ResCapsNet framework over COVID-CAPS and COVID-Net frameworks in terms of overall accuracy for COVID-18 detection and sensitivity.

2 Methodology

As previously stated, most of the studies on deep learning-based COVID-19 detection have so far used CNNs, which although being powerful image processing techniques, they are unable to capture spatial relations between image instances. In this work, therefore, we focus on the combination of capsule network as the main encoder and ResNet50 [4] as the feature extractor. The main objective is to train a detection models on limited COVID-19 datarie to distinguish COVID-19 cases from normal cases using chest X-ray images from COVIDx dataset, collated by Wang et al. [5]. We also augment the data to generate more meaningful data by rotating, vertically flipping, and cutting the training images. In Sub-section 3.1, we will report the performance of our proposed model in terms of overall accuracy, sensitivity and specificity and will compare the results to COVID-Net [5] and COVID-CAPS [3].

2.1 COVIDx Dataset

For this project, we use the COVIDx dataset. The dataset consists of 13,975 chest X-rays of 13,870 patients. Fig. 1 shows one sample of the dataset. There are four categories in the COVIDx [5] dataset: normal (without pneumonia infection), bacterial pneumonia, viral pneumonia, and COVID-19. We divide the dataset into two main categories, positive (COVID-19) and negative (other viral pneumonia, bacterial pneumonia and normal). The dataset has been divided into a training set and a test set with a ratio of 7:3.

(a) Normal (b) Bacterial (c) Non-COVID Viral (d) COVID-19

Figure 1: Data Labels in COVIDx dataset. [5]

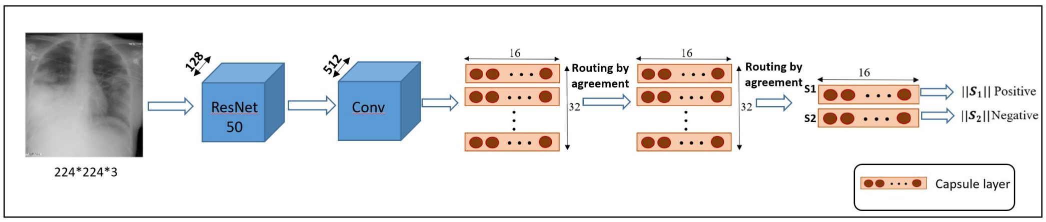

Figure 2: Overall structure of the proposed ResCapsNet framework.

2.2 Data Augmentation

In our project, we perform rotation, random vertical flip, and random crop and resize to create more virtual training samples from the existing data. Applying data augmentation would help to the generalization of the model. The positive effect of data augmentation on the performance of the proposed ResCapsNet framework will be proved based on the achieved results in Sub-section. 3.1.

2.3 Architecture

The overall architecture of the proposed ResCapsNet network consists of ResNet50 and three Capsule layers. Fig. 2 depicts the overall structure of ResCapsNet framework.

ResNet50 We use ResNet50 to extract and aggregate features to perform binary classification on the input chest X-ray images. The ResNet50 is pretrained on ImageNet [9] and is fine-tuned on the COVIDx dataset. We applied data augmentation on the data, then converted it into tensors. After normalization, we pass the input to the ResNet50 network for feature extraction.

Capsule Network The capsule network consists of several capsule layers, where each capsule layer is composed of multiple capsules. The output of the ResNet50 is reshaped to a 512 dimensional vector using a convolutional layer. The output of convolutional layer is then passed to the primary capsule layer. Each Capsule i has the instantiation parameter u_i, and tries to predict the outputs of the next layer's Capsules using a trainable weights W_{ij}. The prediction of Capsule i for Capsule j is defined as:

$$\hat{u}_{j|i} = W_{ij} u_i \tag{1}$$

The actual output of the Capsule j, denoted by s_j, is determined through the "Routing by Agreement" process as follows:

$$a_{ij} = s_j \hat{u}_{j|i}, \quad b_{ij} = b_{ij} + a_{ij}, \quad s_j = \sum_i c_{ij} \hat{u}_{j|i}, \quad \text{and} \quad c_{ij} = \frac{exp(b_{ij})}{\sum_k exp(b_{ik})}, \tag{2}$$

where a_{ij} represents the agreement between predictions and outputs, and c_{ij} denotes the score of the predictions, which determines the contribution of the prediction to the output. The last Capsule layer outputs the instantiate parameters of the two classes of positive and negative COVID-19, and the vector lengths of these two Capsules represent the probability of the occurrence of each class.

Table 1: Results obtained from the proposed ResCapsNet compared to baseline models

Methods	Accuracy	Sensitivity	Specificity	# of Trainable Parameters
ResCapsNet (With Data Augmentation)	**98.9**%	**98.99**%	98.30%	23,300,000
ResCapsNet (No Data Augmentation)	98.0%	98.07%	**99.15**%	23,300,000
COVID-CAPS [3]	95.70%	90.00%	95.80%	**295,488**
COVID-Net [5]	93.3%	91.0%		11,750,000

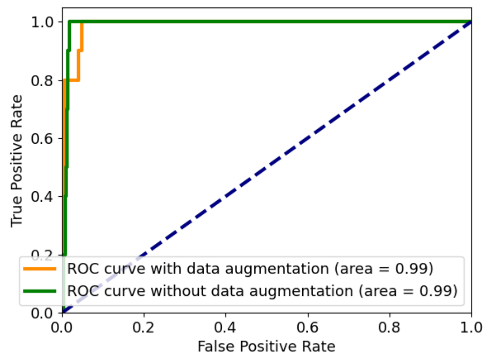

Figure 3: ROC curve from the proposed ResCapsNet

The loss function for each Capsule k in the network is defined as:

$$l_k = T_k max(0, m^+ - ||s_k||)^2 + \lambda(1 - T_k) max(0, ||s_k|| - m^-)^2, \quad \text{and} \quad L = \sum_k l_k, \tag{3}$$

where T_k is evaluated to 1 when the class k is present and 0 otherwise.

3 Experiments

We set the total training iteration to be 30 epochs. We use ADAM optimizer with an initial learning rate of 0.0001. The model is trained with a batch size of 128 on 1 GPU.

3.1 Results & Discussion

We combine ResNet50 with capsule networks to generate ResCapsNet. We use the dataset COVIDx to train the network once with data augmentation; another time without data augmentation. The experimental results are shown in Table 1. The best performance of the proposed ResCapsNet is reached with data augmentation with accuracy of 98.9%, a sensitivity of 98.99%, and a specificity of 99.30%. The obtained receiver operating characteristic (ROC) curve is shown in Figure 3, according to which, the proposed method achieved 0.99 with or without data augmentation. Our proposed method outperforms both the COVID-Net[5] and COVID-CAPS[3] methods by a noticeable margin in terms of accuracy, sensitivity and specificity. The obtained high performance of our model can effectively assist radiologists to lower the number of false positives and false negatives in COVID-19 testing. The proposed method has a larger number of trainable parameters compared to the two references due to large parameters of ResNet. However, despite COVID-CAPS and COVID-Net, ResNet is pre-trained on the natural dataset and does not need specific dataset for pertraining [11].

4 Conclusion

In this paper, we propose a deep learning framework for COVID-19 detection from X-ray images based on ResNet50 and CapsNet. We benefited of pretrained ResNet50 on natural ImageNet dataset for feature extraction and of capsule networks for packaging features into capsules when small number of data is available for training. Based on the achieved results, the proposed ResCapsNet outperforms two state of the art frameworks. It also has been shown that our algorithm performance does not require specific data for pre-training, which greatly reduces the training time.

5 Authors Contributions

Sherry(Yuxuan) and Parvin conceived and designed the project.
Sherry and Parvin wrote the code for the algorithms together.
Sherry did the simulations, and the experimental predictions for algorithm with data augmentation.
Parvin did the simulations, and the experimental predictions for algorithm without data augmentation.
Parvin and Sherry interpreted the results.
Parvin and sherry wrote and revised the manuscript together.

References

[1] S. Stephanie, T. Shum, and H. C. et. al, "Determinants of chest x-ray sensitivity for covid- 19: A multi-institutional study in the united states," 2020.

[2] P. K. Sethy, S. K. Behera, P. K. Ratha, and P. Biswas, "Detection of coronavirus disease (covid-19) based on deep features and support vector machine," 2020.

[3] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K. N. Plataniotis, and A. Mohammadi, "Covid-caps: A capsule network-based framework for identification of covid-19 cases from x-ray images," *Pattern Recognition Letters*, vol. 138, pp. 638–643, 2020.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015. [Online]. Available: http://arxiv.org/abs/1512.03385

[5] L. Wang and A. Wong, "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images," 2020.

[6] O. Gozes, M. Frid-Adar, H. Greenspan, P. D. Browning, H. Zhang, W. Ji, A. Bernheim, and E. Siegel, "Rapid ai development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis," *arXiv preprint arXiv:2003.05037*, 2020.

[7] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," *arXiv preprint arXiv:2003.10849*, 2020.

[8] M. Farooq and A. Hafeez, "Covid-resnet: A deep learning framework for screening of covid19 from radiographs," 2020.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, pp. 1097–1105, 2012.

[10] G. E. Hinton, S. Sabour, and N. Frosst, "Matrix capsules with EM routing," in *International Conference on Learning Representations*, 2018. [Online]. Available: https://openreview.net/forum?id=HJWLfGWRb

[11] C. Zhang, P. Benz, D. M. Argaw, S. Lee, J. Kim, F. Rameau, J.-C. Bazin, and I. S. Kweon, "Resnet or densenet? introducing dense shortcuts to resnet," 2020.