ELSEVIER

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

COVID-CAPS: A capsule network-based framework for identification of COVID-19 cases from X-ray images

Parnian Afshar[a], Shahin Heidarian[b], Farnoosh Naderkhani[a], Anastasia Oikonomou[c], Konstantinos N. Plataniotis[d], Arash Mohammadi[a],*

[a] Concordia Institute for Information Systems Engineering, Concordia University, Montreal, QC, Canada
[b] Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada
[c] Department of Medical Imaging, Sunnybrook Health Sciences Centre, University of Toronto, Canada
[d] Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada

ABSTRACT

Novel Coronavirus disease (COVID-19) has abruptly and undoubtedly changed the world as we know it at the end of the 2nd decade of the 21st century. COVID-19 is extremely contagious and quickly spreading globally making its early diagnosis of paramount importance. Early diagnosis of COVID-19 enables health care professionals and government authorities to break the chain of transition and flatten the epidemic curve. The common type of COVID-19 diagnosis test, however, requires specific equipment and has relatively low sensitivity. Computed tomography (CT) scans and X-ray images, on the other hand, reveal specific manifestations associated with this disease. Overlap with other lung infections makes human-centered diagnosis of COVID-19 challenging. Consequently, there has been an urgent surge of interest to develop Deep Neural Network (DNN)-based diagnosis solutions, mainly based on Convolutional Neural Networks (CNNs), to facilitate identification of positive COVID-19 cases. CNNs, however, are prone to lose spatial information between image instances and require large datasets. The paper presents an alternative modeling framework based on Capsule Networks, referred to as the COVID-CAPS, being capable of handling small datasets, which is of significant importance due to sudden and rapid emergence of COVID-19. Our results based on a dataset of X-ray images show that COVID-CAPS has advantage over previous CNN-based models. COVID-CAPS achieved an Accuracy of 95.7%, Sensitivity of 90%, Specificity of 95.8%, and Area Under the Curve (AUC) of 0.97, while having far less number of trainable parameters in comparison to its counterparts. To potentially and further improve diagnosis capabilities of the COVID-CAPS, pre-training and transfer learning are utilized based on a new dataset constructed from an external dataset of X-ray images. This is in contrary to existing works on COVID-19 detection where pre-training is performed based on natural images. Pre-training with a dataset of similar nature further improved accuracy to 98.3% and specificity to 98.6%.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Novel Coronavirus disease (COVID-19), first emerged in Wuhan, China [1], has abruptly and significantly changed the world as we know it at the end of the 2nd decade of the 21st century. COVID-19 seems to be extremely contagious and quickly spreading globally with common symptoms such as fever, cough, myalgia, or fatigue resulting in ever increasing number of human fatalities. Besides having a rapid human-to-human transition rate, COVID-19 is associated with high Intensive Care Unit (ICU) admissions resulting in an urgent quest for development of fast and accurate diagnosis solutions [1]. Identifying positive COVID-19 cases in early stages helps with isolating the patients as quickly as possible [2], hence breaking the chain of transition and flattening the epidemic curve.

Reverse Transcription Polymerase Chain Reaction (RT-PCR), which is currently the gold standard in COVID-19 diagnosis [1], involves detecting the viral RNA from sputum or nasopharyngeal swab. The RT-PCR test is, however, associated with relatively low sensitivity (true positive rate) and requires specific material and equipment, which are not easily accessible [1]. Moreover, this test is relatively time-consuming, which is not desirable as the positive COVID-19 cases should be identified and tracked as fast as possible [2]. Images [3] in COVID-19 patients, on the other hand,

* Corresponding author.
E-mail addresses: arash.mohammadi@concordia.ca, arashmoh@encs.concordia.ca (A. Mohammadi).

have shown specific findings, such as ground-glass opacities with rounded morphology and a peripheral lung distribution. Although imaging studies and theirs results can be obtained in a timely fashion, the previously described imaging finding may be seen in other viral or fungal infections or other entities such as organizing pneumonia, which limits the specificity of images and reduces the accuracy of a human-centered diagnosis.

Literature Review: Since revealing the potentials of computed tomography (CT) scans and X-ray images in detecting COVID-19 and weakness of the human-centered diagnosis, there have been several studies [5–7] trying to develop automatic COVID-19 classification systems, mainly using Convolutional Neural Networks (CNNs) [4]. Xu et al. [1] have first adopted a pre-trained 3D CNN to extract potential infected regions from the CT scans. These candidates are subsequently fed to a second CNN to classify them into three groups of COVID-19, Influenza-A-viral-pneumonia, and irrelevant-to-infection, with an overall accuracy of 86.7%. Wang et al. [2] have first extracted candidates using a threshold-based strategy. Consequently, for each case two or three regions are randomly selected to form the dataset. A pre-trained CNN is fine-tuned using the developed dataset. Finally, features are extracted from the CNN and fed to an ensemble of classifiers for the COVID-19 prediction, reaching an accuracy of 88%. CT scans are also utilized in Reference [8] to identify positive COVID-19 cases, where all slices are separately fed to the model and outputs are aggregated using a Max-pooling operation, reaching a sensitivity of 90%. In a study by Wang and Wong [9], a CNN model is first pre-trained on the ImageNet dataset [10], followed by fine-tuning using a dataset of X-ray images to classify subjects as normal, bacterial, non-COVID-19 viral, and COVID-19 viral infection, achieving an overall accuracy of 83.5%. In a similar study by Sethy and Behera [11], different CNN models are trained on X-ray images, followed by a Support Vector Machine (SVM) classifier to identify positive COVID-19 cases, reaching an accuracy of 95.38%.

Contributions: All the studies on deep learning-based COVID-19 classification have so far utilized CNNs, which although being powerful image processing techniques, are prone to an important drawback. They are unable to capture spatial relations between image instances. As a result of this inability, CNNs cannot recognize the same object when it is rotated or subject to another type of transformation. Adopting a big dataset, including all the possible transformations, is the solution to this problem. However, in medical imaging problems, including the COVID-19 classification, huge datasets are not easily accessible. In particular, COVID-19 has been identified only recently, and large enough datasets are not yet developed. Capsule Networks (CapsNets) [12] are alternative models that are capable of capturing spatial information using routing by agreement, through which Capsules try to reach a mutual agreement on the existence of the objects. This agreement leverages the information coming from instances and object parts, and is therefore able to recognize their relations, without a huge dataset. Through several studies[13–18], we have shown the superiority of the CapsNets for different medical problems such as brain tumor [13–17] and lung tumor classification [18]. In this study, we propose a Capsule Network-based framework, referred to as the COVID-CAPS, for COVID-19 identification using X-ray images. The proposed COVID-CAPS achieved an accuracy of 95.7%, a sensitivity of 90%, specificity of 95.8%, and Area Under the Curve (AUC) of 0.97.

To potentially and further improve diagnosis capabilities of the COVID-CAPS, we considered pre-training and transfer learning using an external dataset of X-ray images, consisting of 94,323 frontal view chest X-ray images for common thorax diseases. This dataset is extracted from the NIH Chest X-ray dataset [21] including 112,120 X-ray images for 14 thorax abnormalities. From existing 15 diseases in this dataset, 5 classes were constructed with the

help of a thoracic radiologist, with 18 years of experience in thoracic imaging (A. O.). It is worth mentioning that our pre-training strategy is in contrary to that of Reference [9] where pre-training is performed based on natural images (ImageNet dataset). Intuitively speaking, pre-training based on an X-ray dataset of similar nature is expected to result in better transfer learning in comparison to the case where natural images were used for this purpose. In summary, pre-training with an external dataset of X-ray images further improved accuracy of COVID-CAPS to 98.3%, specificity to 98.6%, and AUC to 0.97, however, with a lower sensitivity of 80%. Trained COVID-CAPS model is available publicly for open access at https://github.com/ShahinSHH/COVID-CAPS. To the best of our knowledge, this is the first study investigating applicability of the CapsNet for the problem at hand.

The rest of the manuscript is organized as follows: Section 2 briefly introduces the Capsule networks. The COVID-CAPS is presented in Section 3. Utilized dataset for evaluation of the proposed COVID-CAPS, and our results are presented in Section 4. Finally, Section 5 concludes the work.

2. Capsule networks

Each layer of a Capsule Network (CapsNet) consists of several Capsules, each of which represents a specific image instance at a specific location, through several neurons. The length of a Capsule determines the existence probability of the associated instance. Similar to a regular CNN, each Capsule i, having the instantiation parameter $\boldsymbol{u}_i$, tries to predict the outputs of the next layer's Capsules, using a trainable weight matrix $\boldsymbol{W}_{ij}$, as follows

$$\hat{\boldsymbol{u}}_{j|i} = \boldsymbol{W}_{ij}\boldsymbol{u}_i, \tag{1}$$

where $\hat{\boldsymbol{u}}_{j|i}$ denotes the prediction of Capsule i for Capsule j. The predictions, however, are taken into account based on a coefficient, through the "Routing by Agreement" process, to determine the actual output of the Capsule j, denoted by $\boldsymbol{s}_j$, as follows

$$a_{ij} = \boldsymbol{s}_j.\hat{\boldsymbol{u}}_{j|i}, \tag{2}$$

$$b_{ij} = b_{ij} + a_{ij}, \tag{3}$$

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})}, \tag{4}$$

$$\text{and } \boldsymbol{s}_j = \sum_i c_{ij}\hat{\boldsymbol{u}}_{j|i}, \tag{5}$$

where a_{ij} denotes the agreement between predictions and outputs, and c_{ij} is the score given to the predictions. In other words, this score determines the contribution of the prediction to the output. Routing by agreement is what makes the CapsNet different from a CNN and helps it identify the spatial relations.

The CapsNet loss function, l_k, associated with Capsule k, is calculated as follows

$$l_k = T_k \max(0, m^+ - ||\boldsymbol{s}_k||)^2 + \lambda(1 - T_k)\max(0, ||\boldsymbol{s}_k|| - m^-)^2, \tag{6}$$

where T_k is one whenever the class k is present and zero otherwise. Terms m^+, m^-, and λ are the hyper parameters of the model. The final loss is the summation over all the l_ks. This completes a brief introduction to CapsuleNets, next we present the COVID-CAPS framework.

3. The proposed COVID-CAPS

The architecture of the proposed COVID-CAPS is shown in Fig. 1, which consists of 4 convolutional layers and 3 Capsule layers. The inputs to the network are 3D X-ray images. The first layer is a convolutional one, followed by batch-normalization. The second layer
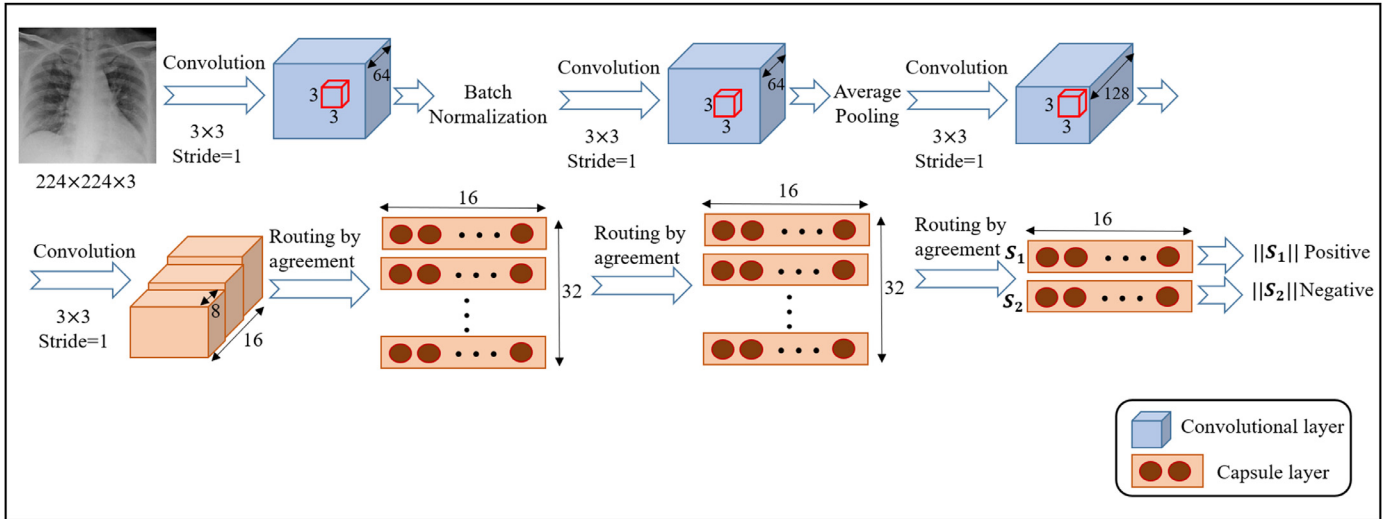
Fig. 1. The proposed COVID-CAPS architecture.

is also a convolutional one, followed by average pooling. Similarly, the third and forth layers are convolutional ones, where the forth layer is reshaped to form the first Capsule layer. Consequently, three Capsule layers are embedded in the COVID-CAPS to perform the routing by agreement process. The last Capsule layer contains the instantiation parameters of the two classes of positive and negative COVID-19. The length of these two Capsules represents the probability of each class being present.

Since we have developed a Capsule Network-based architecture, which does not need a large dataset, we did not perform any data augmentation. However, since the number of positive cases, N^+, are less than the negative ones, N^-, we modified the loss function to handle the class imbalance problem. In other words, more weight is given to positive samples in the loss function, where weights are determined based on the proportion of the positive and negative cases, as follows

$$\text{loss} = \frac{N^+}{N^+ + N^-} \times \text{loss}^- + \frac{N^-}{N^+ + N^-} \times \text{loss}^+, \qquad (7)$$

where loss^+ denotes the loss associated with positive samples, and loss^- denotes the loss associated with negative samples.

As stated previously, to potentially and further improve diagnosis capabilities of the COVID-CAPS, we considered pre-training the model in an initial step. In contrary to Reference [9] where ImageNet dataset [10] is used for pre-training, however, we constructed and utilized an X-ray dataset. The reason for not using ImageNet for pre-training is that the nature of images (natural images) in that dataset is totally different from COVID-19 X-ray dataset. It is expected that using a model pre-trained on X-ray images of similar nature would result in better boosting of the COVID-CAPS. For pre-training with an external dataset, the whole COVID-CAPS model is first trained on the external data, where the number of final Capsules is set to the number of output classes in the external set. From existing 15 disease in the external dataset, 5 classes were constructed with the help of a thoracic radiologist, with 18 years of experience in thoracic imaging (A. O.). To fine-tune the model using the COVID-19 dataset, the last Capsule layer is replaced with two Capsules to represent positive and negative COVID-19 cases. All the other Capsule layers are fine-tuned, whereas the conventional layers are fixed to the weights obtained in pre-training.

In summary, COVID-CAPS architecture contains the following modifications applied to the original Capsule Network presented in Reference [12]:

- The Capsule Network presented in Reference [12] originally works on a dataset of digital numbers, which are black-and-white and small in size compared to X-ray images. To make the Capsule Network applicable in the problem at hand, we have extended the Capsule layers and the number of routing procedures to be able to extract useful patterns from X-ray images.
- The dataset originally used for the development of the Capsule Networks is completely balanced in terms of the number of instances available for each class label. The COVID-19 identification problem, however, is restricted to highly unbalanced datasets, as COVID-19 is a new disease. To account for this unbalanced dataset, we modified the original margin loss to assign more penalty to mis-classified positive cases.
- We pre-trained the Capsule Network to compensate for the small available dataset. The pre-training is performed on an external dataset with 5 classes, reflected in 5 final Capsules. These 5 Capsules are then collapsed into two Capsules, and all the Capsule layers are fine-tuned on the main COVID-19 dataset.

We used Adam optimizer with an initial learning rate of 10^{-3}, 100 epochs, and a batch size of 16. We have split the training dataset, described in Section 4, into two sets of training (90%) and validation (10%), where training set is used to train the model and the validation set is used to select a model that has the best performance. Selected model is then tested on the testing set, for the final evaluation. The following four metrics are utilized to represent the performance: Accuracy; Sensitivity; Specificity, and Area Under the Curve (AUC). Next, we present the obtained results.

4. Experimental results

To conduct our experiments, we used the same dataset as Reference [9]. This dataset is generated from two publicly available chest X-ray datasets [19,20]. As shown in Fig. 2, the generated dataset contains four different labels, i.e., Normal; Bacterial; Non-COVID Viral, and; COVID-19. As the main goal of this study is to identify positive COVID-19 cases, we binarized the labels as either positive or negative. In other words, the three labels of normal, bacterial, and non-COVID viral together form the negative class.

Using the aforementioned dataset, the proposed COVID-CAPS achieved an accuracy of 95.7%, a sensitivity of 90%, specificity of 95.8%, and AUC of 0.97. The obtained receiver operating characteristic (ROC) curve is shown in Fig. 3. *In particular, false positive cases have been further investigated to have an insight on what types*

Table 1
Results obtained from the proposed COVID-CAPS, along with the results from Reference [11]. Pre-trained CNN refers to a CNN with the same front-end as the COVID-CAPS.

Method	Accuracy	Sensitivity	Specificity	Number of Trainable Parameters
COVID-CAPS without pre-training	95.7%	90%	95.8%	**295,488**
Pre-trained COVID-CAPS	**98.3%**	80%	**98.6%**	**295,488**
Reference [11]	95.4%	**97.3%**	93.5%	23,000,000
Pre-trained CNN	96.2%	50%	97%	368,508,226
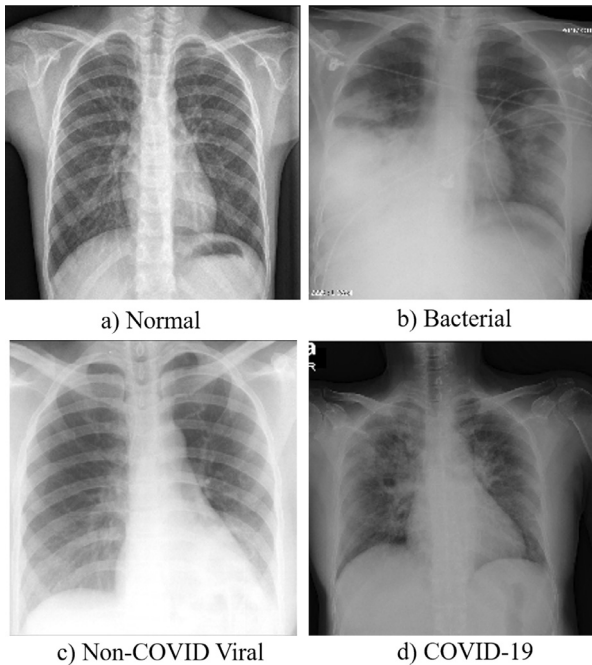
Fig. 2. Labels available in the dataset.

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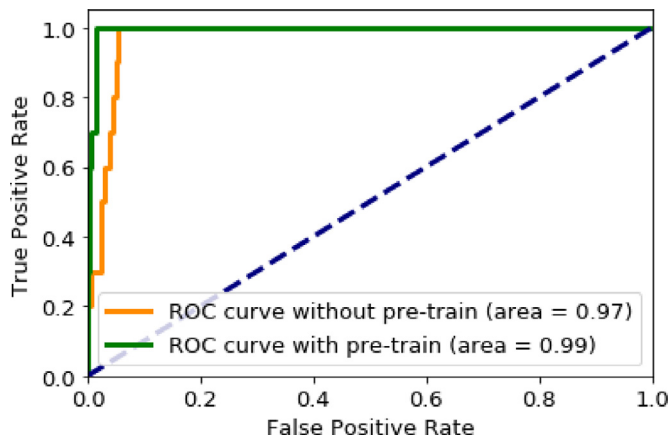

Fig. 3. ROC curve from the proposed COVID-CAPS. Here, "without pre-train" label refers to training from scratch.

are more subject to being mis-classified by COVID-19. It is observed that 54% of the false positives are normal cases, whereas bacterial and non-COVID cases form only 27% and 19% of the false positives, respectively.

As shown in Table 1, we compare our results with Reference [11] that has used the binarized version of the same dataset. COVID-CAPS outperforms its counterpart in terms of accuracy and specificity. Sensitivity is higher in the model proposed in Reference [11], that contains 23 million trainable parameters. Reference[6] is

another study on the binarized version of the same X-ray images. However, as the negative label contains only normal cases (in contrast to including all normal, bacterial, and non-covid viral cases as negative), we did not compare the performance of the COVID-CAPS with this study. *It is worth mentioning that the proposed COVID-CAPS has only 295,488 trainable parameters. Compared to 23 million trainable parameters of the model proposed in Reference [11], therefore, COVID-CAPS can be trained and used in a more timely fashion, and eliminates the need for availability of powerful computational resources.*

In another experiment, we pre-trained the proposed COVID-CAPS using an external dataset of X-ray images, consisting of 94,323 frontal view chest X-ray images for common thorax diseases. This dataset is extracted from the NIH Chest X-ray dataset [21] including 112,120 X-ray images for 14 thorax abnormalities. This dataset also contains normal cases without specific findings in their corresponding images. In order to reduce the number of categories, we classified these 15 groups into 5 categories based on the relations between the abnormalities in each disease. The first four groups are dedicated to No findings, Tumors, Pleural diseases, and Lung infections categories. The fifth group encompasses other images without specific relations with the first four groups. We then removed 17,797 cases with multiple labels (appeared in more than one category) to reduce the complexity. The adopted dataset is then used to pre-train our model. Table 2 demonstrates our classification scheme and distribution of the data. Results obtained from fine-tuning the pre-trained COVID-CAPS is also shown in Table 1, according to which, pre-training improves accuracy and specificity. The ROC curve is shown in Fig. 3, according to which, the obtained AUC of 0.99 outperforms that of COVID-CAPS without pre-training.

Based on an inclusive study reported in Reference [22], human-centered COVID-19 detection from chest radiography leads to a high sensitivity, whereas specificity remains as low as 25%. The low specificity can lead to excessive expenses to isolate and treat false positive cases. The obtained specificity of 98.6% using the proposed COVID-CAPS can significantly assist radiologists to lower the number of reported false positives. Furthermore, the ROC curve can provide physicians with a means to calibrate and balance the sensitivity and specificity. In other words, by changing the probability threshold, above which the positive label is assigned to a subject, physicians are able to form the desired balance between sensitivity and specificity. To make this point more clear, we changed the probability threshold based on the ROC curve from the default value of 0.5 to 0.44. This new threshold increases the sensitivity to 100%, while specificity remained, more or less, intact (98.4%).

To further elaborate on the effectiveness of the proposed model, we designed a CNN that has the same front-end as that of the COVID-CAPS. In other words, it has the same convolutional layers (the first four main layers of the COVID-CAPS). The Capsule layers, however, are replaced with three fully-connected layers, the first two of which have 256 neurons and the last one, having a Sigmoid activation, has two neurons representing the two classes of positive and negative COVID-19 cases. It is worth noting that we considered fully-connected layers after the front-end, because to some

Table 2
Description of the external X-ray images dataset used for pre-training COVID-CAPS.

Final Category	Initial Categories	Number of Images
No Findings	No Findings	60361
Tumors	Infiltration, Mass, Nodule	16103
Pleural Diseases	Effusion, Pleural Thickening, Pneumothorax	8042
Lung Infection	Consolidation, Pneumonia	1668
Others	Atelectasis, Cardiomegaly, Edema, Emphysema, Fibrosis, Hernia	8149

extent they resemble Capsule layers in the sense that there is no shared weights or kernels. This CNN is pre-trained on the same external dataset. In the fine-tuning phase, the convolutional layers are kept fixed and only the fully-connected layers are retrained. Furthermore, the cross-entropy loss function is modified (similar in nature to the modifications introduced on the margin loss of the COVID-CAPS in Eq. (7)) to give more penalty to mis-classified positive cases. All other hyper-parameters, including the optimizer and learning rate, exactly resemble the hyper-parameters of the COVID-CAPS. The training, validation, and test sets are also the same as the ones used in COVID-CAPS. Based on the obtained results, which are presented in Table 1, the designed CNN, having 368,508,226 trainable parameters, achieves an accuracy of 96.24%, a sensitivity of 50%, and a specificity of 96.97%. The lower performance of the CNN, and the fact that it has exactly the same front-end with only the Capsule layers replaced with fully-connected ones support the effectiveness of the Capsule layer with the routing by agreement mechanism.

Finally, it is worth providing some intuition on COVID-CAPS time and space complexity. In particular and following the literature [23], we model the time complexity as a function of the number of required multiplications in both Capsule and fully-connected layers. Generally speaking, a fully-connected layer involves a matrix multiplication. Considering $m \times d_1$ and $n \times d_2$ neurons in two consecutive fully-connected layers, the required matrix multiplication involves $m \times d_1 \times n \times d_2$ multiplication operations. Reshaping the two fully-connected layers into two consecutive Capsule layers leads to m Capsules of dimension d_1 making predictions for n Capsules of dimension d_2. Each single prediction involves $d_1 \times d_2$ multiplications, as each lower layer Capsule i with dimension d_1 should be multiplied by the weight matrix $\boldsymbol{W}_{ij}$ to form the prediction $\hat{\boldsymbol{u}}_{j|i}$ for the higher layer Capsule j of dimension d_2. In other words, $\boldsymbol{W}_{ij}$ has d_1 rows and d_2 columns. Considering n Capsules in the lower layer and m Capsules in the higher layer, the total number of operations is $m \times d_1 \times n \times d_2$, which is exactly the same as the fully-connected scenario. However, based on Eq. (5), each parent Capsule is calculated as a weighted average over the predictions. Weighting each prediction $\hat{\boldsymbol{u}}_{j|i}$ by the coupling coefficient c_{ij} involves d_2 (dimension of the prediction and parent Capsule) multiplications. Again having n Capsules in the lower layer and m Capsules in the higher layer, one routing by agreement process includes $d_2 \times n \times m$ multiplications. In conclusion, even with one round of routing by agreement, which means equal contribution of all the predictions, a Capsule layer has $d_2 \times n \times m$ multiplications more than a fully-connected layer. In practice, however, Capsule Networks require far less layers to have comparable performance with CNNs. To illustrate this point we calculated the time needed to predict the outcome of one single subject using the proposed COVID-CAPS. Our TITAN Xp GPU computer takes almost 0.16 s to calculate the outcome, whereas this time is approximately 1.62 s for the ResNet-50 model utilized in Reference [11]. Finally, regrading the space complexity, as we showed in the Table 1, COVID-CAPS contains far less trainable parameters compared to its counterparts. In particular, while trained COVID-

CAPS occupies almost 1.5 Megabytes, the ResNet-50 requires 98 Megabytes.

5. Conclusion

In this study, we proposed a Capsule Network-based framework, referred to as the COVID-CAPS, for diagnosis of COVID-19 from X-ray images. The proposed framework consists of several Capsule and convolutional layers, and the lost function is modified to account for the class-imbalance problem. The obtained results show that the COVID-CAPS has a satisfying performance with a low number of trainable parameters. Pre-training was able to further improve the accuracy, specificity, and AUC. Trained COVID-CAPS model is available publicly for open access at https://github.com/ShahinSHH/COVID-CAPS. As more and more COVID-19 cases are being identified all around the world, larger datasets are being generated. We will continue to further modify the architecture of the COVID-CAPS and incorporate new available datasets. New versions of the COVID-CAPS will be released upon development through the aforementioned link.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] X. Xu, X. Jiang, et al., A deep learning system to screen novel coronavirus disease 2019 pneumonia, Engineering (2020).

[2] S. Wang, B. Kang, et al., A deep learning algorithm using CT images to screen for corona virus disease (COVID-19), medRxiv (2020).

[3] P. Afshar, A. Mohammadi, K.N. Plataniotis, A. Oikonomou, H. Benali, From handcrafted to deep-learning-based cancer radiomics: challenges and opportunities, IEEE Signal Processing Magazine 36 (4) (2019) 132–160. July

[4] R. Yamashita, M. Nishio, et al., Convolutional neural networks: an overview and application in radiology, Insights Imaging 9 (4) (2018) 611–629.

[5] O. Gozes, M. Frid-Adar, et al. Rapid AI development cycle for the coronavirus (COVID-19) pandemic: initial results for automated detection & patient monitoring using deep learning CT image analysis, arXiv:2003.05037 (2020).

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

[7] M. Farooq, A. Hafeez, COVID-ResNet: a deep learning framework for screening of COVID19 from radiograph, arXiv:2003.14395 (2020).

[8] L. Li, L. Qin, et al., Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT, Radiology (2020).

[9] L. Wang, A. Wong, COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images, arXiv:2003.09871 (2020).

[10] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Neural Inf. Process. Syst. (NIPS) (2012).

[11] P.K. Sethy, S.K. Behera, Detection of coronavirus disease (COVID-19) based on deep features, 2020, Preprints 2020, 2020030300.

[12] G. Hinton, S. Sabour, N. Frosst, Matrix capsules with EM routing, ICLR (2018).

[13] P. Afshar, A. Mohammadi, K.N. Plataniotis, Brain tumor type classification via capsule networks, in: IEEE International Conference on Image Processing (ICIP), 2018, pp. 3129–3133.

[14] P. Afshar, K.N. Plataniotis, A. Mohammadi, Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries, in: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 1368–1372.

[15] P. Afshar, K.N. Plataniotis, A. Mohammadi, Capsule networks' interpretability for brain tumor classification via radiomics analyses, in: 2019 IEEE International Conference on Image Processing (ICIP), 2019, pp. 13816–13820.

[16] P. Afshar, K.N. Plataniotis, A. Mohammadi, BoostCaps: a boosted capsule network for brain tumor classification, Accepted in *IEEE Engineering in Medicine and Biology Society (EMBC)*, 2020.

[17] P. Afshar, K.N. Plataniotis, A. Mohammadi, A bayesian approach to brain tumor classification using capsule networks, *IEEE Signal Processing Letters* (2020). Submitted

[18] P. Afshar, A. Oikonomou, F. Naderkhani, P.N. Tyrrell, K. Farahani, K.N. Plataniotis, A. Mohammadi, 3d-MCN: A 3d multi-scale capsule network for lung nodule malignancy classification, Nat. Sci. Rep. 10 (2020).

[19] J.P. Cohen, COVID chest x-ray dataset, 2020, https://github.com/ieee8023/covid-chestxray-dataset.

[20] P. Mooney, Kaggle chest x-ray images (pneumonia) dataset, 2020, https://github.com/ieee8023/covid-chestxray-dataset.

[21] X. Wang, Y. Peng, et al., ChestX-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 3462–3471.

[22] B. Xu, Y. Xing, et al., Chest CT for detecting COVID-19: a systematic review and meta-analysis of diagnostic accuracy, Eur. Radiol. (2020) 1–8.

[23] V. Sze, Y. Chen, et al., Efficient processing of deep neural networks: a tutorial and survey, Proc. IEEE 105 (12) (2017) 2295–2329.