2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Automated Detection of COVID-19 Cases using Recent Deep Convolutional Neural Networks and CT images*

Mohamed Chetoui, Moulay A. Akhloufi, *Senior Member IEEE*

***Abstract*— COVID-19 is an acute severe respiratory disease caused by a novel coronavirus SARS-CoV-2. After its first appearance in Wuhan (China), it spread rapidly across the world and became a pandemic. It had a devastating effect on everyday life, public health, and the world economy. The use of advanced artificial intelligence (AI) techniques combined with radiological imaging can be helpful in speeding-up the detection of this disease. In this study, we propose the development of recent deep learning models for automatic COVID-19 detection using computed tomography (CT) images. The proposed models are fine-tuned and optimized to provide accurate results for multiclass classification of COVID-19 vs. Community Acquired Pneumonia (CAP) vs. Normal cases. Tests were conducted both at the image and patient-level and show that the proposed algorithms achieve very high scores. In addition, an explainability algorithm was developed to help visualize the symptoms of the disease detected by the best performing deep model.**

I. INTRODUCTION

Radiographic patterns on CT chest scans have shown higher sensitivity (SN) and specificity (SP) compared to RT-PCR detection of COVID-19 which, according to the World Health Organization (WHO) has a relatively low positive detection rate in the early stages [1].

One study found that the sensitivity of COVID-19 detection on chest CT was 97% (95% CI, 95-98%, 580/601 patients), based on positive RT-PCR performance, out of 1041 cases from China [2]. It is possible to reliably detect the differences between viral forms of pneumonia through the use of AI and deep learning (DL), making it an important screening method in the fight against the COVID-19 pandemic. The use of convolutional neural networks (CNN) for detecting pneumonia (including COVID-19) has been investigated in several recent studies. The results show that deep learning-based diagnostic systems can achieve a high performance in detecting chest diseases, especially COVID-19 and pneumonia.

Xu *et al.* [3] presented the detection of COVID-19, Influenza viral pneumonia and normal (healthy cases) using CT images. The authors used ResNet [4] architecture with Location Attention. They tested their model on 219 CT images of COVID-19, 224 of viral pneumonia and 175 of normal. Their model achieved an accuracy (ACC) of 86.7%, Area Under Curve (AUC) of 99.6% for Coronavirus vs Non-coronavirus cases. They obtained a sensitivity (SN) of 98.2% and a specificity (SP) of 92.2%.

Song *et al.* [5] developed a deep learning-based CT diagnosis system named DeepPneumonia to identify COVID-19. The authors used 88 patients diagnosed with the COVID-19 from hospitals in two provinces in China, 101 patients infected with bacterial pneumonia, and 86 healthy persons. The experimental results showed that their model could discriminate between the COVID-19 infected patients and bacteria pneumonia-infected patients with an AUC of 95%, SN of 96% and ACC of 86%.

Wang *et al.* [6] collected 1,065 CT images of COVID-19, 325 of normal and 740 images with viral pneumonia. The authors fine-tuned the Inception network [7] with transfer learning for COVID-19 detection. Their algorithm achieved an ACC of 89.5%, a SP of 88% and a SN of 88%.

Zheng *et al.* [8] developed a weakly supervised deep learning-based system using 3D CT volumes to detect COVID-19. The authors used a pre-trained UNet [9] to segment the lung region, then the segmented 3D lung region was fed into a 3D CNN to predict the probability of COVID-19. 499 CT volumes were used for training and 131 CT volumes were used for testing. The authors obtained an AUC of 95.90%, a SN of 90.70%, a SP of 91.10% and an ACC of 90.10%.

Mishra *et al.* [10] presented various Deep CNN for detecting the presence of COVID19 in CT images. Their approache combined predictions from multiple individual models to produce a final prediction. The authors used a dataset containing 360 positive COVID-19 cases and 397 negatives. Experimental results showed that the proposed decision fusion-based approach achieved above 86% results across all the performance metrics under consideration, with an average AUC and F1-Score of 88.30% and 86.70%, respectively.

Harmon *et al.* [11] used a series of CNNs, trained in a CT of 1,280 patients to localize parietal pleura/lung parenchyma followed by classification of COVID-19 pneumonia. They achieved an ACC of 90.8%, a SN of 84% and a SP of 93%, evaluated in an independent test set of 1,337 patients.

In this work, we present the use of various recent CNN models for a multiclassification of COVID-19 vs. Normal vs. Community Acquired Pneumonia (CAP). The models are fine-tuned in order to detect COVID-19, all models are trained and tested on SPGC-COVID dataset [12].

* This work was partially supported by the Microsoft AI For Health program, and by support from Atlantic Canada Opportunities Agency (ACOA), Regional Economic Growth through Innovation - Business Scale-Up and Productivity (project 217148), Natural Sciences and Engineering Research Council of Canada (NSERC), Alliance Grants (ALLRP 552039-20), and New Brunswick Innovation Foundation (NBIF), COVID-19 Research Fund (COV2020-042).

M. Chetoui and M.A. Akhloufi are with the Perception, Robotics, and Intelligent Machines Research Group (PRIME), Dept of Computer Science, Université de Moncton, Moncton, NB, Canada {emc7409, moulay.akhloufi}@umoncton.ca

II. Proposed approach

A. deep learning models

Various CNN models are used in our experiments for detecting CAP, COVID-19 and normal cases. A DenseNet-121 and 3 of EfficientNet models (B0, B4 and B5). The DenseNet [13] architecture decreases the connection between input and output, which helps solve the issue of vanishing gradient. Each layer in the DenseNet has a reduced size of the feature map which is useful for training the CNNs on a small dataset, leading to a less likelihood of facing over-fitting issues. The DenseNet consists mainly of Dense Block, Transition Layer and Rate of Growth. By setting the required composite ratio coefficient to match the width, depth and resolution of the network, EfficientNet [14] incorporated the characteristics of the above network, so that when extending the three dimensions of the network, the model gives a better efficiency using the compound coefficient ϕ using the following equation:

$$\begin{aligned} depth &: d = \alpha^{\phi} \\ width &: w = \beta^{\phi} \\ resolution &: r = \gamma^{\phi} \\ \alpha \geq 1, &\beta \geq 1, \gamma \geq 1 \end{aligned} \tag{1}$$

Where α, β, γ are constants that a limited grid search can be used to evaluate. ϕ is a user-defined coefficient that regulates the amount of additional resources necessary for model scaling, while α, β, γ specify how these additional resources can be allocated to network width, resolution, and depth, respectively.

In this study, we used various models of EfficientNet: EfficientNet-B0, EfficientNet-B4 and EfficientNet-B5. The DenseNet model was also used in our experiment, it has 121 layers (DenseNet-121). The models used in this study are pre-trained using ImageNet dataset. We fine-tuned each model by adding the Global Average Pooling, and 2 Dense layers of size 512 and 128 respectively. Each Dense layer is followed by Batch Normalization and ReLu Activation. To reduce overfitting, we set 25% for Dropout. Finally, a Softmax layer gives the probability prediction scores for detecting CAP, COVID-19 and Normal.

B. SPGC-COVID dataset

For training and testing our CNN models, we used SPGC-COVID dataset [12]. It contains volumetric chest CT scans of patients positive for COVID-19 infection, Community Acquired Pneumonia (CAP), and normal patients. The dataset includes slices in DICOM format with the size of 512 × 512. The COVID-19, CAP, and normal cases are collected from April 2018 to December 2019 and January 2019 to May 2020. COVID-19 infection diagnosis is based on positive findings of the real-time Reverse Transcription Polymerase Chain Reaction (rRT-PCR) examination, clinical criteria, and CT scan symptoms reported by three thoracic radiologists. Labels given by the three radiologists indicated a high degree of consensus (more than 90 percent). Diagnosis for CAP and normal cases was confirmed using clinical parameters and CT scans. The dataset is acquired in a retrospective study with the use of automated exposure control and automated tube potential selection. The dataset contains volumetric CT scans slices of 171 patients positive for COVID-19, 60 CAP, and 76 normal cases.

In our work, we kept only the slices where the lung area is present. Thus removing the upper and lower slices from the CT scans. We finally got 9,171 CT images for training and 4,411 for the test (no data augmentation was used). Figure 1 shows examples of the removed images, as we can see they do not represent the lung area.

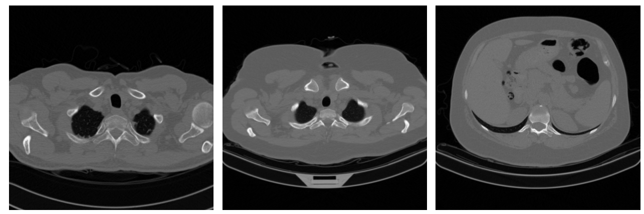

Fig. 1: Example of removed CT images from SPGC-COVID dataset

III. Experimental results

A. Training parameters

For each model, we used the provided data with 9,171 CT images for training and 4,411 for validation, including 2,329 CT images of COVID-19, 999 for CAP and 1,083 for normal. Keras Library [15] was used to develop the proposed models and training was carried out using 3 Nvidia Tesla K80 [16] in Microsoft Azure servers. Adam [17] was used as optimizer, Batch size was fixed to 32 and we trained the models for 200 epochs. All measures are calculated on the basis of sensitivity (SE), specificity (SP), area under curve (AUC) and accuracy (ACC).

B. Image-level and patient-level results

The proposed fine-tuned DenseNet-121 gives the best performance. The model achieved an ACC of 93.62%, a SP of 96%, a SN of 94%, and an AUC of 98.81% using the 4,411 CT images in the validation set. An interesting score was obtained by a fine-tuned EfficienNet-B4 with an AUC of 97.01%. Table I shows the obtained results. Figure 3 shows the ROC curve for the COVID-19 detection. We can see the high performance achieved by fine-tuned DenseNet-121 model followed by EfficienNet-B4. The confusion matrix of the validation set for the best model is given in figure 2 (left), only 92 CT of COVID-19 images from 2,329 are misclassified, this shows the performance of our model in the detection of COVID-19 cases.

To confirm how the models learned to detect COVID-19 signs, we developed an explainability model based on the use of Gradient-weighted Class Activation Mapping (Grad-CAM) [18]. This approach was used to generate a visual description of the outcomes of the proposed CNN models. Grad-CAM uses any target's gradients, flowing into the final convolutional layer to generate a coarse map of localization

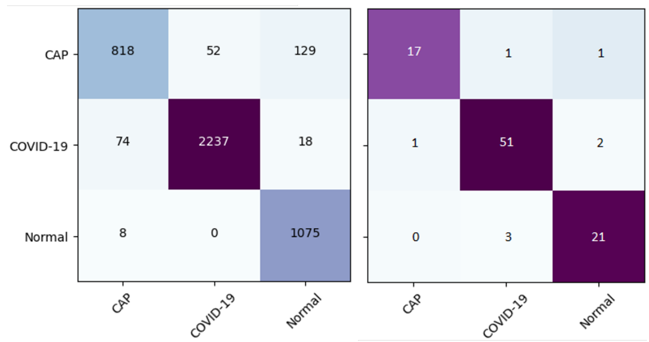

Fig. 2: Confusion matrix for image-level (left) and patient-level (right) classification

TABLE I: ACC, AUC, SN, and SP of our deep learning models for detecting COVID-19, CAP and Normal

Model	ACC%	AUC%	SN%	SP%
EfficienNet-B0	88.23	96.55	88.00	90.00
EfficienNet-B4	91.95	97.01	93.00	91.00
EfficienNet-B5	88.36	95.98	92.00	94.00
DenseNet-121	**93.62**	**98.81**	**94.00**	**96.00**

highlighting important regions in the predictive image. Grad-CAM is applicable to our proposed CNN model without any architectural changes or re-training. The proposed technique combines Grad-CAM with fine-grained visualizations to create a high-resolution class-discriminative visualization. Figure 4 shows positive cases of COVID-19 detected by our best model. We can see that the heatmap is accurately located around the signs of COVID-19, such as opacity.

A comparison with recent COVID-19 detection methods is provided in table II. We can see that for COVID-19 detection using CT images, our model outperforms most of the recently published work. Moreover, relative to other studies, the number of COVID-19 images used in this analysis is higher, showing the degree of generalization of the proposed model for COVID-19 detection. The proposed model enhances our previous model [19], which was created using a limited amount of positive COVID-19 CT images (347 CT images). Our previous model was based on EfficientNet-B0 and obtained an AUC of 85%, a SP of 81% a SN of 75% and an ACC of 79%. The results show that the current proposed model is robust and able to detect COVID-19 with a high sensitivity and specificity. In addition, our explanability model achieves a very detailed localization of COVID-19 symptoms which can be used in a CAD tool to further assist physicians in their diagnosis.

For patient-level classification, we want to evaluate the performance of the model in detecting COVID-19, CAP and normal for each patient. Thus, instead of just detecting the disease in individual images, we will use the CT volume of each patient to output a decision in our validation set, where we have 54 COVID-19 positive patients, 19 with CAP, and 24 normal.

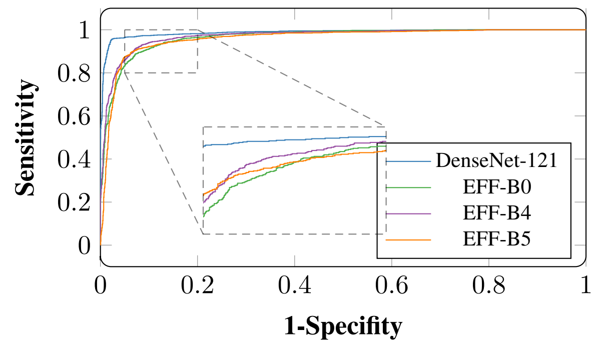

Fig. 3: ROC curves of the proposed deep learning models for detecting COVID-19 on CT images

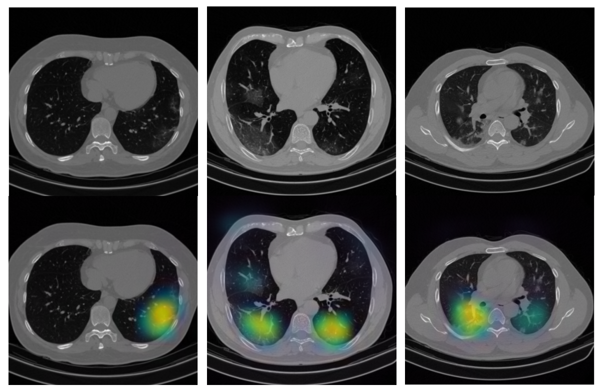

Fig. 4: Example of detected COVID-19 cases. The yellow colored areas show the location of the COVID-19 signs

Little work has been done for a decision at a patient-level. Zheng *et al.* [8] proposed the use of a threshold to classify each patient CT volume. Their best threshold was 0.5 for COVID-19 detection, meaning that if the probability of COVID-19 was larger than 0.5, the patient was classified as COVID-positive, and vice versa.

In this work, we propose an algorithm based on both voting and thresholding as shown in algorithm 1. The algorithm receives the CT volume for patient i as input, classifies the slices and outputs the number of slices detected for each class V_i ={#COVID-19, #CAP, #Normal}. For example if the CT volume has 100 slices and 50 are detected as COVID-19, 30 as CAP, and 20 as Normal, then $V_i = \{50, 30, 20\}$. The output vector V_i is used to decide if a patient is COVID positive. If the number of slices detected as COVID-19 are above a threshold Th and this number is above the number of CAP slices, then the output C_d is COVID-19, else if CAP slices are above a threshold Th then the output C_d is CAP, otherwise the patient is detected as normal. In this work the best performing threshold was 0.35 (It was selected empirically), meaning that we need 35% of positive slices before taking a decision for COVID-19 and CAP.

Figure 2 (right) gives the obtained confusion matrix for our classifications. In average, we obtain a SN of 91.00%, a SP of 96.00%, an ACC of 92.00% and an AUC of 93.00%. Table II gives a comparison with the work of Zheng *et al.* [8] for patient-level COVID-19 classification using CT scans.

Algorithm 1: Patient-level classification

Data: CT(i): CT volume
Output: C_d:{COVID-19, CAP, Normal}
Function
  V_i = Classify(CT(i)) ;
  return V_i ;
if $V_i(0) > Th \cap V_i(0) > V_i(1)$ **then** $C_d =$ COVID-19 ;
else if $V_i(1) > Th$ **then** $C_d =$ CAP ;
else
  $C_d =$ Normal
end

TABLE II: Performance comparison with state-of-the-art methods using CT images and for patient-level classification

Study	ACC%	SN%	SP%	AUC%	# COVID CTs
Xu *et al.* [3]	86.74	98.21	92.28	99.67	219
Song *et al.* [5]	86.00	96.00	-	95.00	777
Wang *et al.* [6]	89.5	88.00	88.00	-	1,065
Mishra *et al.* [10]	86.00	-	-	88.00	360
Harmon *et al.* [11]	90.88	84.00	93.00	-	1,337
Our (DenseNet-121)	**93.62**	**94.00**	**96.00**	**98.81**	**2,329**
Xu *et al.* [3]	86.74	98.21	92.28	99.67	90 patients
Zheng *et al.* [8]	90.19	90.79	91.17	95.95	131 patients
Our (DenseNet-121)	**92.00**	**91.00**	**96.00**	93.00	97 patients

IV. CONCLUSION

In this work, we developed multiple CNN models based on EfficienNet-B0, EfficienNet-B4, EfficienNet-B5 and DenseNet-121 for COVID-19 detection. All models were fine-tuned and trained to use CT images. The findings show that the best model (DenseNet-121) outperforms many recent deep learning approaches for COVID-19 detection. On a validation set of 4,411 CT images, including 2,329 of COVID-19, we obtained high AUC scores with 98.81% for DenseNet-121 and 97.01% for EfficienNet-B4 to classify the CT images as COVID-19, CAP or healthy. We also showed the performance of the best algorithm for patient-level classification. The results are promising and achieve high scores with a SN of 91%, a SP of 96%, an ACC of 92% and an AUC of 93%. An explanability algorithm was also developed and shows that the proposed model can efficiently detect most pathology regions of COVID-19 disease. The main limitations of the algorithm are when bones or liver present in some scans look like COVID-19 signs. More work is needed to improve these cases. The proposed approach is an interesting contribution to the development of a Computer Aided Diagnostic capable of detecting COVID-19 cases on CT images. Future work includes testing with other deep architectures and the use of more CT images for training. In addition, the approach will be adapted to the classification of other types of medical images and diseases.

REFERENCES

[1] WHO, "Diagnostic testing for sars-cov-2," https://www.who.int/publications/i/item/diagnostic-testing-for-sars-cov-2, 2020, online; accessed April 2021.
[2] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, and L. Xia, "Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in china: a report of 1014 cases," *Radiology*, vol. 296, no. 2, pp. E32–E40, 2020.
[3] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, L. Yu, Q. Ni, Y. Chen, J. Su *et al.*, "A deep learning system to screen novel coronavirus disease 2019 pneumonia," *Engineering*, vol. 6, no. 10, pp. 1122–1129, 2020.
[4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
[5] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, J. Chen, R. Wang, H. Zhao, Y. Zha *et al.*, "Deep learning enables accurate diagnosis of novel coronavirus (covid-19) with ct images," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2021.
[6] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng, and B. Xu, "A deep learning algorithm using CT images to screen for corona virus disease (COVID-19)," *European Radiology*, Feb. 2021. [Online]. Available: https://doi.org/10.1007/s00330-021-07715-1
[7] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.
[8] C. Zheng, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, and X. Wang, "Deep learning-based detection for COVID-19 from chest CT using weak label," *MedRxiv*, Mar. 2020. [Online]. Available: https://doi.org/10.1101/2020.03.12.20027185
[9] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
[10] A. K. Mishra, S. K. Das, P. Roy, and S. Bandyopadhyay, "Identifying COVID19 from chest CT images: A deep convolutional neural networks based approach," *Journal of Healthcare Engineering*, vol. 2020, pp. 1–7, Aug. 2020. [Online]. Available: https://doi.org/10.1155/2020/8843664
[11] S. A. Harmon, T. H. Sanford, S. Xu, E. B. Turkbey, H. Roth, Z. Xu, D. Yang, A. Myronenko, V. Anderson, A. Amalou *et al.*, "Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets," *Nature communications*, vol. 11, no. 1, pp. 1–7, 2020.
[12] SPGC, "Signal Processing Grand Challenge(SPGC) on COVID-19," http://i-sip.encs.concordia.ca/2021SPGC-COVID19/index.html, online; accessed April 2021.
[13] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.
[14] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," *CoRR*, vol. abs/1905.11946, 2019. [Online]. Available: http://arxiv.org/abs/1905.11946
[15] F. Chollet *et al.*, "Keras," https://keras.io, 2015.
[16] NVIDIA, "NVIDIA TESLA K80," https://www.nvidia.com/en-gb/data-center/tesla-k80/, online; accessed April 2021.
[17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
[18] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618–626.
[19] M. Chetoui, T. Abadarahmane, and M. A. Akhloufi, "Deep learning for covid-19 detection on chest X-Ray and CT scan," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2020, poster.