

DETECTING COVID-19 AND COMMUNITY ACQUIRED PNEUMONIA USING CHEST CT SCAN IMAGES WITH DEEP LEARNING

Shubham Chaudhary[◇], Sadbhawna[◇], Vinit Jakhetiya[◇], Badri N Subudhi[◇],
Ujjwal Baid[†], Sharath Chandra Guntuku[†]

[◇] Indian Institute of Technology, Jammu, India

[†] University of Pennsylvania, PA, USA

ABSTRACT

We propose a two-stage Convolutional Neural Network (CNN) based classification framework for detecting COVID-19 and Community Acquired Pneumonia (CAP) using the chest Computed Tomography (CT) scan images. In the first stage, an infection - COVID-19 or CAP, is detected using a pre-trained DenseNet architecture. Then, in the second stage, a fine-grained three-way classification is done using EfficientNet architecture. The proposed COVID+CAP-CNN framework achieved a slice-level classification accuracy of over 94% at identifying COVID-19 and CAP. Further, the proposed framework has the potential to be an initial screening tool for differential diagnosis of COVID-19 and CAP, achieving a validation accuracy of over 89.3% at the finer three-way COVID-19, CAP, and healthy classification. Within the IEEE ICASSP 2021 Signal Processing Grand Challenge (SPGC) on COVID-19 Diagnosis, our proposed two-stage classification framework achieved an overall accuracy of 90% and sensitivity of .857, .9, and .942 at distinguishing COVID-19, CAP, and normal individuals respectively, to rank first in the evaluation. Code and model weights are available at https://github.com/shubhamchaudhary2015/ct_covid19_cap_cnn

Index Terms— COVID-19, CAP, Chest CT, Deep Learning.

1. INTRODUCTION

As of March 2021, there have been more than 119 million confirmed cases of the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) infection, the virus that causes the novel coronavirus disease (COVID-19), resulting in over 2.6 million reported deaths [1]. Chest Computed Tomography (CT) images have shown to be an essential method for detecting interstitial pneumonia, a distinctive feature of COVID-19 [2]. Deep learning based computational imaging techniques are promising at the evaluation of positive COVID-19 cases [3].

Several works have studied the feasibility of automating COVID-19 detection using Convolutional Neural Networks

(CNN) [4–11]. Harmon *et al.* showed that AI-based algorithms trained on a multinational cohort could classify CT images between pneumonia associated with COVID-19 and non-COVID-19 pneumonia with over 90% accuracy [5]. Wang *et al.* proposed a Deep tailored CNN for chest X-ray images to detect COVID-19 [6] and released an open-source dataset of 13975 chest X-ray images. On another dataset of 349 COVID-19 CT images from 216 patients and 463 non-COVID-19 samples, an automatic diagnosis system using multi-task and self-supervised learning techniques reported accuracy 89% [7]. Zheng *et al.* used 499 3D CT volumes to predict COVID-19 infections by training a weakly-supervised deep learning model [8]. COVID-CAPS [9] is a capsules network-based framework trained on X-ray images to identify COVID-19 infections. COVID-FACT [10] is a two-stage network where the first stage identifies slices with infection and classifies them further into COVID-19 and other infections in the second stage.

Manual analysis of chest CT scans by professional medical professionals is a resource and time-intensive process. An automated risk assessment system could potentially serve as an initial screening tool to aid medical professionals' diagnosis. One of the challenges associated with the automatic detection of COVID-19 is its differential diagnosis compared to CAP [12]. To address this, we propose a two-stage Convolutional Neural Network (CNN) based framework, termed COVID+CAP-CNN, for the differential detection of COVID-19 and CAP.

In Stage-1, we train two different CNN architectures for labeling COVID-19 and CAP CT scans. In Stage-2, we propose to use EfficientNet architecture to classify individuals and CT scans into three classes, i.e., Normal, COVID-19, and CAP. We use the SPGC-COVID dataset [4], which contains CT scans of individuals with a combination of COVID-19 and CAP, along with CT scans of individuals with no infections. The dataset has slice-level as well as patient-level labels for several individuals. In Fig. 1, example CT scans for normal individuals and patients affected with COVID-19 and CAP in the SPGC-COVID dataset are shown. Sample descriptives of the SPGC dataset are given in Table 1.

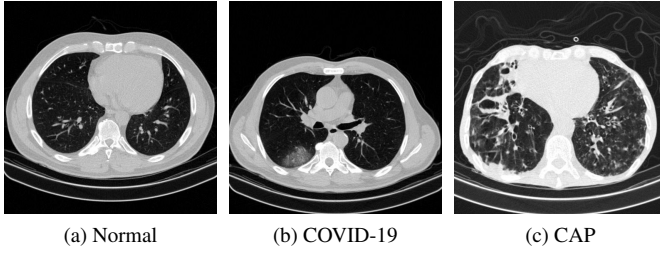


Fig. 1: Segmented slice images from three example CT volumes of a) normal, b) COVID-19, c) CAP patients

Table 1: The description of SPGC dataset.

Disease Type	Number of Individuals	Number of Patients with slice-level labels
Normal	76	—
CAP	60	25
COVID-19	171	55

2. PROPOSED MODEL

Our proposed COVID-19 and CAP detection system using deep learning, termed as COVID+CAP-CNN, consists of two stages. In the first stage, unlabelled CT scans in the SPGC dataset are labeled using a pre-trained CNN-based algorithm (detailed in subsection 2.2). In the second stage, labeled slices are pooled into individual-level estimates of COVID-19, CAP infected patients, and normal individuals using the state-of-the-art EfficientNet architecture (detailed in subsection 2.3). The complete architecture of the two stages of the proposed COVID-CNN model is shown in Figs. 2 and 3. We describe data pre-processing and implementation of these two stages in the following subsections.

2.1. Dataset Pre-processing

Each patient's data in a chest CT scan is given in the form of a 3D volume, a combination of many image slices captured from different angles. Images provided in the SPGC-COVID dataset were mapped from Hounsfield Units (HU) to $[0, 255]$ using a window centered at -500 HU with a width of 1300 HU. Further, we observed that the crucial signal associated with infection is often present in the 3D chest CT volumes' central slices. Consequently, we selected the middle slices from the 3D volumes to fine-tune the proposed COVID-CNN architecture. For COVID-19 and CAP patients with over 80 slices in the 3D chest CT volume, we select the middle 80 slices, and if the number of slices is less than 80, then we take the middle 40 slices. The distribution of adopted data for Stage-1 architectures is shown in Table 2. To overcome class imbalance, a similar number of CT scans are selected from different classes.

Table 2: Data used for fine-tuning Stage-1

Disease Type	Number of Patients	Total Selected Slices	
		Infectious	Non-Infectious
CAP	55	910	910
COVID-19	25	3046	3046

2.2. Stage-1: Slice Label Prediction

All slices for each patient in the dataset do not contain labels. Consequently, it is required to differentiate the slices with infections and without infections. To automatically detect the slices with infection and without infection, we use Stage-1 slice level label prediction framework (Fig. 2). We propose to independently label the unlabelled slices of COVID-19 and CAP 3D volumes using two different CNN architectures. We first used the SPGC dataset labels to create separate classes of infectious and non-infectious slices. Using transfer learning [13], we extracted features from the pre-trained CNN model for Stage-1. Several works have utilized transfer learning to extract features successfully for tasks both within and across domains [14–17]. We utilized the weights of DenseNet-121 [18] architecture to transfer learning across domains. The pre-trained features from Densenet-121 are obtained using Global Average Pooling (GAP) on the last layer. Then a fully connected layer with 1024 features is used. The fully connected layer was unfrozen, and its weights were trained. The extracted features were then fed into a Sigmoid activation function to obtain the two-class probabilities.

We used Adam optimizer for the training process. The chosen parameters values for β_1 was 0.5. The adopted learning rate for our model was $2 * 10^{-4}$. We obtained an accuracy of over 94 % while fine-tuning.

2.3. Stage-2: Classification Model for Diagnosis

After segmenting the 3D volumes into middle slices that exclusively contain lung data, as explained in the pre-processing step, we perform Stage-1 of the framework, slice-wise labeling of unlabelled slices. Our CNN architecture classifies CT scan slices into three fine-grained classes in the second stage: COVID-19, CAP, and healthy. The architecture designed for diagnosis purposes is illustrated in Fig. 3.

In order to classify CT scan images into three classes, we used several pre-trained architectures such as EfficientNet [19], InceptionV3 [20], ResNet [21], and DenseNet [18]. EfficientNet architecture has several advantages over other deep learning architectures such as compound scaling (in dimensions such as width, depth, resolution of an image, etc.), reduced set of parameters that make the training process efficient. For this reason, we fine-tuned EfficientNet architecture to extract features for our proposed classification model. These features are extracted using Global Average Pooling (GAP) and then fed into two fully connected dense layers

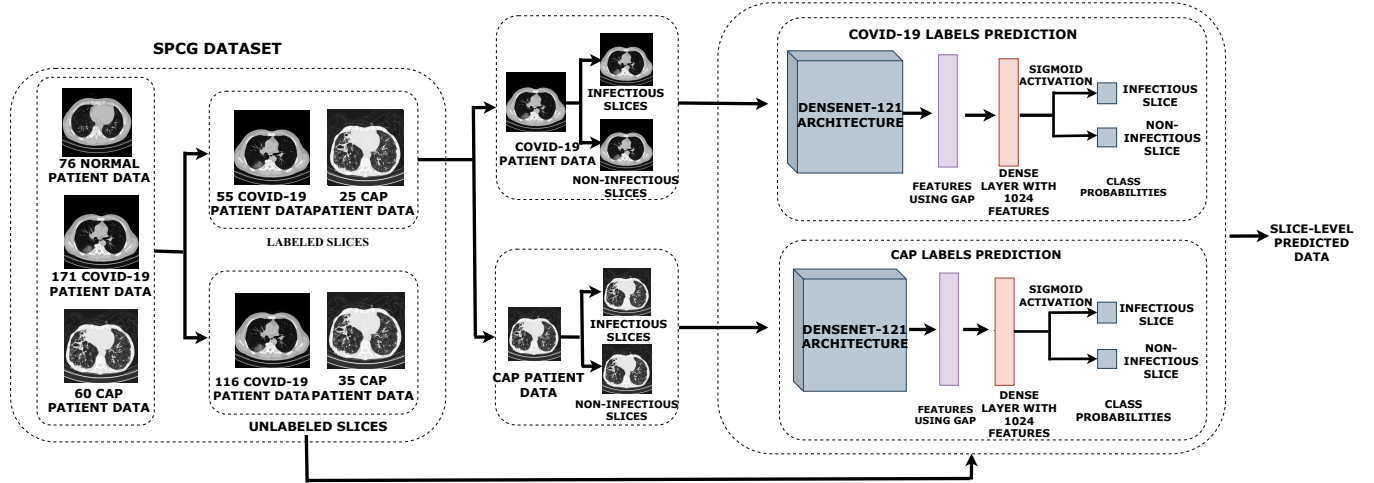


Fig. 2: Overview of Stage-1 of the proposed COVID+CAP-CNN Architecture

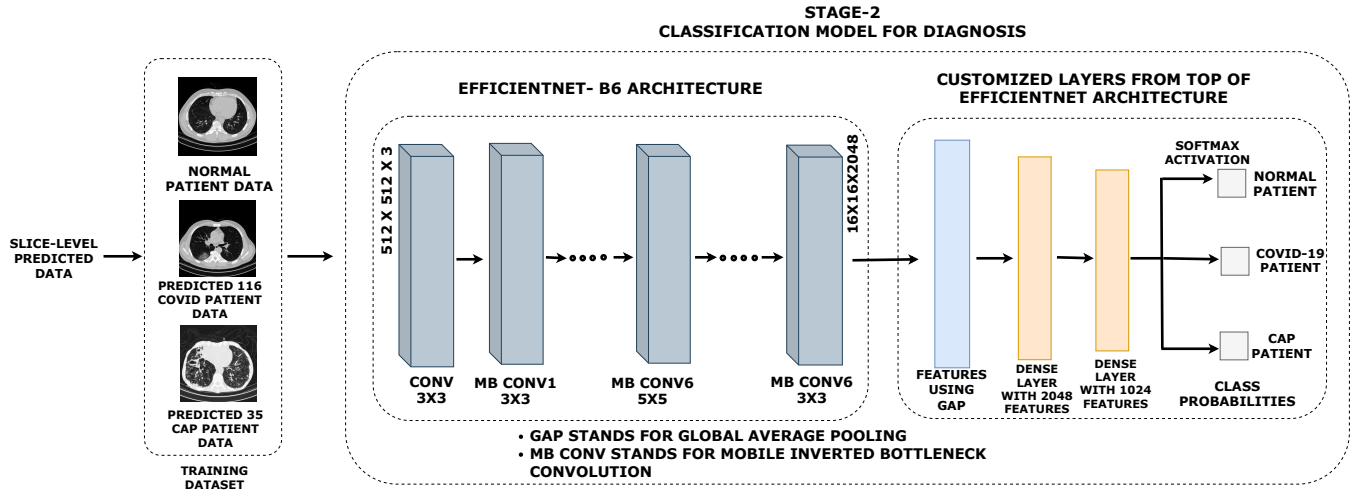


Fig. 3: Overview of Stage-2 of the proposed COVID+CAP-CNN Architecture.

with 2048 and 1024 trainable parameters, respectively. For three-class classification, a softmax activation function is used to predict the final class probabilities. The original and predicted labels (for COVID-19 and CAP) from Stage-1 are used for each of the three training classes. We make sure that there is no overlap in training and validation data and ensured out-of-sample patient validation. For training, we adopted the same optimizer as Stage-1 architecture i.e. Adam Optimizer with $\beta_1 = 0.5$ and learning rate of $2 * 10^{-4}$.

Single-slice prediction for final diagnosis does not guarantee the individual's diagnosis [10]. Hence, we utilized the slice-level predictions from Stage-2 of our proposed model for patient-level prediction of diagnosis based on a simple voting technique. All the slices of a patient are tested for classification into three classes. As discussed in Section 2.1, central slices of the 3D volume are important because they contain crucial information regarding infection and

were used for patient-level prediction. Let x, y, z be the total number of slices from the centre of the 3D volume predicted as COVID-19, CAP, normal, respectively. Further, let x', y', z' be the total number of slices other than centre slices of 3D volume predicted as COVID-19, CAP, normal, respectively. We calculated the final Patient Label (PL) as: $PL = (\max((x + 0.7x'), (y + 0.7y'), (z + 0.5z')))$, where 0.7 and 0.5 are the weight factors obtained heuristically.

3. RESULTS

The proposed COVID+CAP-CNN model is trained and validated on different out-of-sample folds of the SPCG-COVID dataset. All analyses performed in Stage-2 are evaluated on non-overlapping out-of-sample sets. For an ablation study, we evaluated both the stages of our model for training and validation accuracy. Table 3 shows the accuracy

of the proposed model for Stage-1, which is over 94 %.

Table 3: Performance of Stage-1 for binary classification: infected and non-infected

Classification Type	Training accuracy	Validation accuracy
COVID-19	99.29 %	97.7 %
CAP	99.17 %	94.7 %

Stage-2 accuracy is shown in Tables 4 and 5. In Table-4, we show the variation of performance on different CNN architectures. EfficientNet is seen to outperform others. In Table-5, we evaluate the performance of the proposed model on different train-validation splits. Our proposed model achieved over 89.3 % accuracy on the validation set for different splits.

We also evaluated patient-level accuracy for the three-way classification of the proposed framework and observed that the proposed algorithm achieves an 84% accuracy across three fine-grained classes.

3.1. Class-wise Sensitivity Analysis

We show the confusion matrix considering the three classes: Normal, CAP, and COVID-19 in Fig. 4 and Fig. 5. The data provided in the confusion matrix in Fig. 4 is at slice-level, and that of Fig. 5 is at patient-level. Our fine-tuned EfficientNet model obtains a class-wise sensitivity of 89.2 %, 81.9 % and 91.49 % for Normal, CAP and, COVID-19 classes, respectively.

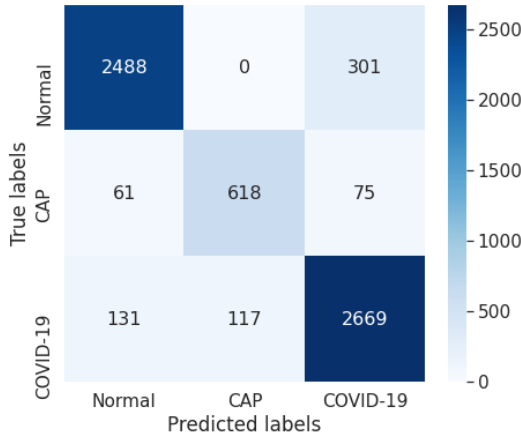


Fig. 4: Confusion Matrix for proposed proposed COVID+CAP-CNN model at slice-level.

3.2. Execution Time

We performed all the experiments on an NVIDIA Tesla V100 Tensor Core GPU. The proposed model takes less than 0.16 seconds for testing one CT scan image slice of shape $512 \times 512 \times 3$.

Table 4: Performance of Stage-2 for fine-grained classification: with different architectures.

Architecture	Training Accuracy	Validation Accuracy
Proposed(EfficientNet)	99.00 %	89.3 %
InceptionV3	98.78 %	82.9 %
ResNet	99.61 %	78 %
DenseNet	98.66 %	80.6 %

Table 5: Performance of Stage-2 for fine-grained classification: with different train-validation splits

Train-Validation Split	Training Accuracy	Validation Accuracy
70-30	99.00 %	89.3 %
80-20	98.54 %	90.25 %
90-10	99.3 %	91 %

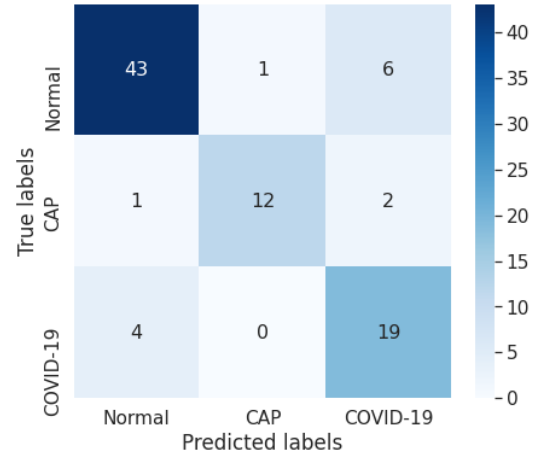


Fig. 5: Confusion Matrix for proposed proposed COVID+CAP-CNN model at patient-level.

4. CONCLUSION

In this work, we proposed a two-stage framework to detect COVID-19 and CAP using CT scan images. In the first stage, individual slices of CT scans are labeled using fine-tuned DenseNet based deep-learning architecture. In the second stage, a fine-grained differential classification in three classes, i.e., COVID-19, CAP, and healthy individuals, by fine-tuning the EfficientNet architecture. The proposed two-stage framework achieved over 94% accuracy for classifying the CT scan images in the binary classification task: infectious vs. non-infectious, and accuracy of 89.3% for the fine-grained three-class classification: COVID-19, CAP, and normal.

5. REFERENCES

- [1] E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track covid-19 in real time," *The Lancet infectious diseases*, vol. 20, no. 5, pp. 533–534, 2020.
- [2] F. Grillet, J. Behr, P. Calame, S. Aubry, and journal=Radiology volume=296 number=3 pages=E186–E188 year=2020 publisher=Radiological Society of North America Delabrousse, E., "Acute pulmonary embolism associated with covid-19 pneumonia detected with pulmonary ct angiography," .
- [3] 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*, 2020.
- [4] P. Afshar, S. Heidarian, N. Enshaei, F. Naderkhani, R. Moezedin J., A. Oikonomou, F. Fard, K. Samimi, K. Plataniotis, and A. Mohammadi, "Covid-ct-md: Covid-19 computed tomography (ct) scan dataset applicable in machine learning and deep learning," *arXiv*, 09 2020.
- [5] S. A. Harmon, T. H. Sanford, S. Xu, E. B. Turkbey, H. Roth, ..., M. Cariati, G. Carrafiello, P. An, B. J. Wood, and B. Turkbey, "Artificial intelligence for the detection of covid-19 pneumonia on chest ct using multinational datasets," *Nature Communications*, vol. 11, pp. 2041–1723, 2020.
- [6] L. Wang, Z. Q. Lin, and A. Wong, "Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images," *Scientific Reports*, vol. 10, 2020.
- [7] X. Yang, X. He, J. Zhao, Y. Zhang, S. Zhang, and P. Xie, "Covid-ct-dataset: A ct scan dataset about covid-19," *arXiv*, 2020.
- [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*, 2020.
- [9] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, Ko.s 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.
- [10] S. Heidarian, P. Afshar, N. Enshaei, F. Naderkhani, A. Oikonomou, S. F. Atashzar, F. Fard, K. Samimi, K.s Plataniotis, A. Mohammadi, and R. Moezedin J., "Covid-fact: A fully-automated capsule network-based framework for identification of covid-19 cases from chest ct scans," *arXiv*, 10 2020.
- [11] P. Afshar, F. Naderkhani, A. Oikonomou, M. J. Rafiee, A. Mohammadi, and K. N. Plataniotis, "Mixcaps: A capsule network-based mixture of experts for lung nodule malignancy prediction," 2020.
- [12] F. Shi, L. Xia, F. Shan, B. Song, D. Wu, Y. Wei, H. Yuan, H. Jiang, Y. He, and Y. Gao, "Large-scale screening of covid-19 from community acquired pneumonia using infection size-aware classification," *Physics in Medicine & Biology*, 2021.
- [13] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "Cnn features off-the-shelf: An astounding baseline for recognition," in *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 512–519.
- [14] P. Mahajan, V. Jakhetiya, P. Abrol, P. K. Lehana, B. N. Subudhi, and S. C. Guntuku, "Perceptual quality evaluation of hazy natural images," *Transaction on Industrial Informatics*, 2021.
- [15] Sadbhawna, V. Jakhetiya, D. Mumtaaz, and S. Jaiswal, "Distortion specific contrast based no-reference quality assessment of dibr-synthesized views," in *IEEE Workshop on Multimedia and Signal Processing (MMSP)*, 2020, pp. 1–6.
- [16] S. C. Guntuku, S. Roy, and L. Weisi, "Evaluating visual and textual features for predicting user 'likes'," in *2015 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2015, pp. 1–6.
- [17] S. C. Guntuku, D. Preotiuc-Pietro, J. C Eichstaedt, and L. H Ungar, "What twitter profile and posted images reveal about depression and anxiety," in *Proceedings of the international AAAI conference on web and social media*, 2019, vol. 13, pp. 236–246.
- [18] G. Huang, Z. Liu, and K. Q. Weinberger, "Densely connected convolutional networks," *arXiv*, vol. abs/1608.06993, 2016.
- [19] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," *CoRR*, vol. abs/1905.11946, 2019.
- [20] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," *CoRR*, vol. abs/1512.00567, 2015.
- [21] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," *CoRR*, vol. abs/1603.05027, 2016.