European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 07, Issue 3, 2020

Automatic Identification of Covid-19 regions on CT-images using Deep Learning

Dr.R.Devika Rubi[1*], Dr.Rajasekaran Subramanian[2], Srinivas Yadav[3], Vijaya Bhargavi[4]
[1] Associate Professor, R & D Department, Keshav Memorial Institute of Technology, Hyderabad, Telangana State, India
[2*]Professor, R & D Department, Keshav Memorial Institute of Technology, Hyderabad, Telangana State, India
[3] Research Intern, R & D Department,Keshav Memorial Institute of Technology, Hyderabad, Telangana State, India
[4]Research Intern, R & D Department, Keshav Memorial Institute of Technology, Hyderabad, Telangana State, India
Corresponding Author E-mail: rajasekarans@kmit.in
Orcid ID: 0000-0002-6572-3934

Corresponding Author: Dr.RAJASEKARAN SUBRAMANIAN, R&D DEPARTMENT, KESHAV MEMORIAL INSTITUTE OF TECHNOLOGY, 3-5-1026 NARAYANAGUDA, HYDERABAD, TELANGANA 500029, rajasekarans@kmit.in

Abstract
Covid-19 is a contagious respiratory illness caused by a new coronavirus called SARS-COV-2, spreads all around the world and death rate increases at an exponential rate. Covid-19 can be diagnosed either by laboratory base approaches such as nucleic acid testing, antigens test and serology (antibody) tests or by medical imaging tools such as X-ray and Computed Tomography (CT). RT-PCR remains the primary and gold standard for diagnosing Covid-19 but due to shortages of RT-PCR kit, CT images can be used as an alternative early detection toolkit of Covid-19 as a simpler, quicker and more reliable diagnosis of Covid-19. As increase in bandwidth of CT images as well as the new Covid-19 virus consumes a lot of time and workload of the radiologist increases substantially. Deep Learning models can assist the radiologist by learning the features of Covid-19 by their annotated CT images. This paper proposes novel deep learning models for the three main tasks namely 1. Binary classification of Covid-19 2. Automatic lung segmentation and 3. Covid-19 region segmentation. The proposed deep learning models produce an accuracy of 97%, 98% and 99% respectively. The results of the deep learning models show that the models can assist radiologists for quick, accurate and unbiased diagnosis for Covid-19.
Key-words: Covid-19, CT-Images, Lung region in CT-images, Covid-19 region segmentation from CT Images.

Introduction

Covid-19 is a contagious respiratory illness caused by a new coronavirus called SARS-COV-2. This new coronavirus is an RNA virus which is like flu and measles are prone to changes and mutations compared to DNA viruses such as herpes, small box and human papillomavirus (HPV) [1]. Till date there is no vaccine identified to prevent Covid-19. Since December 2019, corona virus has spread throughout the world and around 6.5 million number of patients have been affected. The diagnostic of Covid-19 can be done by broadly two categories either by laboratory base approaches such as nucleic acid testing, antigens test and serology (antibody) tests or by medical imaging tools such as X-ray and computed tomography [2].

Early identification of Covid-19 helps to isolate people from the infected and to reduce the spread of the disease. RT-PCR remains the primary and gold standard for diagnosing Covid-19 but some clinicians use chest CT scans as a simpler, quicker and more reliable way to diagnose Covid-19. CT scan will help the early detection of Covid-19 quickly. Due to the shortage of RT-PCR test kits, CT-Scan can be used as an alternative for screening and diagnosing of Covid-19.

Covid-19 affects many organs like heart, blood vessels and lungs. The surface receptor ACE2 (Angio tension-Converting Enzyme2) found in alveoli (Tiny air sacs took oxygen from our breath) of human lungs. Covid-19 virus binds with ACE2 and makes breathing difficult. CT scan helps to identify this kind of abnormality.

Deep learning methods can assist in the accurate CT image diagnosis decision making process of Covid-19. This paper proposes three deep networks models for binary classification of Covid-19, automatic CT lung segmentation and Covid-19 region segmentation on CT images.

The remaining paper is organized as follows: The related work section discusses the existing CNN model related to Covid-19, proposed model section describes three proposed CNN models and their performance evolution and conclusion section concludes the research work, their limitations and future research enhancement.

Materials and Methods:

A. Problem Definition

CT-scan investigation helps to identify pre dominant patterns of lung abnormalities like unilateral, multifocal and peripherally ground glass opacities (GGO). The early infection sign for Covid-19 is ground glass opacity (GGO) for later stage it is pulmonary consolidation. CT slices can provide the above important decision for diagnosing Covid-19. Manual annotations of lung infections is a tedious and time-consuming task to the radiologist who is influenced by their bias decision and clinical experiences. At the peak time of Covid-19, due to the high bandwidth of CT images, the radiologists are unable to read the number of CT images timely and also due to the newly developed virus they require additional training for diagnosis. Otherwise, it affects the accuracy of diagnosis decision.

Deep learning models automatically learns the features of Covid-19 from the expert (radiologist) annotated Covid-19 CT images and automatically predicts the Covid-19 regions by using the learned features of the model.

B. Related work

There are many CNN models available for Covid-19. Due to the tremendous growth of Covid-19 population and shortage of RT-PCR kit, X-Rays and CT-images are well played as a secondary diagnosis for Covid-19. CNN models can be viewed as assistants to pathologists and help to identify Covid-19 patients quickly.

Narhes and Shervin et al. [3] proposed a deep learning approach for Covid-19 segmentation on CT images. They used U-net with a modified loss function which is a combination of binary cross entropy loss and 2D regularization of total variation loss. They used 929 images from medseg.si of size 630*630 and the model produces a dice score of 86% and average precision of .94. Athanasis and Eftychios et.al.[4] proposed CT segmentation Covid-19 deep learning model using U-net which is built completely on fully convolutional layers with dense layers and a FCN8 network. They used Radiopaedia [5] consists of 939 cross sectional CT-images of size 630*630. The model claims that the proposed custom U-net produced more coarse boundaries whereas U-net is producing smoother and smaller than that of the original annotated area.

Arnab kumar Mishra,Sujit kumar Das et.al [6] explored various CNN approaches such as VGG16, Inception V3, Resnet 50 and Densenet 121/201 for identifying Covid-19 presence in CT images presents in the public Dataset [7]. They proposed a model in which the convolution part of the above baseline models are kept as Imagenet model and a three layers of fully connected layers with the Relu activation function. They also improved the efficiency of predictions of the baseline model with decision fusion approach with F1 accuracy of .883.

Amine Amyar,Romain modzelewski et al [8] proposes an automatic CT screening tool for classification and segmentation of Covid-19. They proposed a multitask learning model which combines the architecture based on the three tasks such as classification, segmentation and image feature reconstruction for various public datasets like Covid-CT-dataset [9], medical segmentation [10] and Henri Becquerel cancer center (HBCC), France with a collection of 1069 images of size 256*256. The Model produces a dice coefficient of .88 for segmentation and an area under the

ROC curve higher than 97% for the classification.

Qingsen Yan,BO Wang et al [11] designed the Deep learning model Covid-segNet segment the Covid-19 infectious in CT images. The model contains a method called Feature variation (FV) which automatically adjusts the confusing boundaries of Covid-19 region and enhances the features of the annotated images. The model adjusts the shape variation of the images by a method called progressive atrous spatial pyramid pooling [12]. The model used 731 images for training and 130 images for testing which were collected from five Chinese hospitals. The model produced .987 dice coefficient for lung segmentation and .726 dice coefficient for Covid-19 lesion detection.

Deng-pingfan, Tao Zhou et al [13] developed a model called Inf-Net for the automatic detection of Covid-19 regions. Inf-Net aggregates high level features by using parallel partial decoders. InfNet uses a semi-supervised framework which requires a minimum number of labelled images for training.

Dilibag singh,vijay kumar et.al [14] classify chest CT Covid-19 images using CNN, ANN and ANFIS methodologies. They conclude that among them CNN will be the best with a limitation that hyper parameters tuning. They proposed multi objective differential evolution (MODE) based CNN for classifying Covid-19. They tuned the hyperparameters of CNN such as kernel size, kernel type, stride, padding, activation functions, learning rate and batch size by using DE (differential evolution) algorithm [15]. The model used 20 fold cross validation to present the overfitting. F1 accuracy of the proposed system is 90.

The Extensive review of the literature reveals that chest CT-images can be used as an alternative tool for Covid-19 diagnosis and deep learning models can be used for the decision making of Covid-19 diagnosis.

c. Data

The chest CT images are collected from medical segmentation created by medseg DlinRadiology which contains 10 axial volumetric CTS in which 1 axial volumetric CT contains 100 slices of Covid-19 positive images and masks of size 512*512 from more than 40 patients and other 9 axial volumetric CTs contains 829 masks and images of size 630*630 which is a NIfTI format used with python Nibabel packge [16]. The model uses 929 CT images for training and testing.

d. Methods

This paper proposes a model which performs three tasks namely 1. Binary classification of Covid-19 2. Automatic lung segmentation and 3. Covid-19 region segmentation.

The three main tasks of the proposed model are discussed here below.

1) ***Binary classification Model***

This proposed model predicts whether given CT Covid-19 is positive or negative. As the CT images in the dataset consists of different sizes by using nearest neighbour interpolation the CT images are resized into the same size of 630*630. The pre-processing method maintains uniformity locally during resizing which is needed because the Covid-19 CT features are like GGO (Ground glass opacities), cloudy uniform structure. The dataset is divided into 585,65 and 275 for training, validation and testing with equal proportion of positive and negative samples in each. As the CT images are of NIfTI format which doesn't require any noise reduction techniques.

This model used VGG16 [17] [18] and Resnet-50 [19][20] deep learning architecture for the Covid-19 classification task. VGG16 is a simple neural network consisting of different blocks containing convolutional layer forward by max pooling and fully connected layers (dense layers) at the backend with output layer at the end. Convolutional layers at the beginning stage learn the high-level features and the dense layers at the backend learns low level features but increase the parameters drastically due to the dense layers and increased training time. Model size is around 528 MB and has 138,359,544 parameters. The architecture of the VGG-16 model is illustrated in Figure 1.

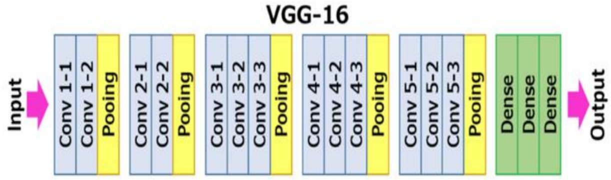


Fig. 1. Architecture of VGG16

Another deep learning network Resnet-50 uses 50 layers deep with skip connections to learn low level features and remember the features to avoid overfitting. As there are no dense connections at the end reduces the size of the network and training time significantly compared to VGG16.

Resnet-50 architecture and its skip connections are illustrated in Figure 2.

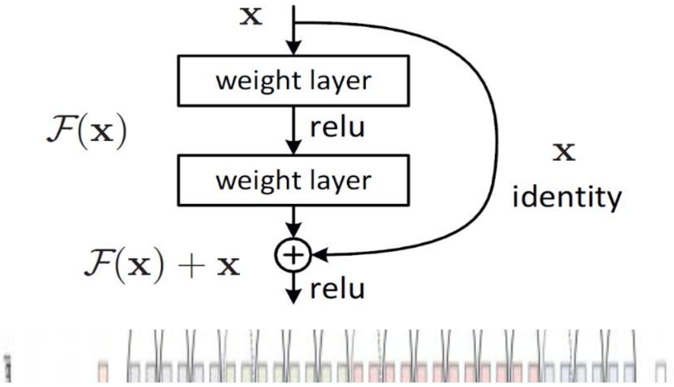


Fig. 2. Resnet Architecture and Skip connections

1) Lung Detection Model

This model automatically detects the lung region in the CT-images. The Deep learning model U-net [21] [22] is used for this proposed semantic segmentation. In U-net model the input images pass through multiple stages of convolutional and pooling which reduces the height and width of image as increase in depth after each convolution in down sampling then followed by fully convolutional and through multiple stages of up sampling to produce mask of the image. The architecture of the U-Net model is illustrated in Figure 3.

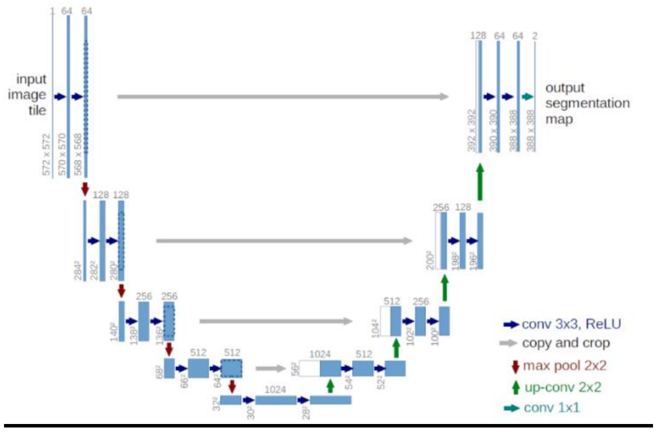


Fig. 3. Architecture of U-Net Model

1) Covid-19 Region Segmentation

This model identifies the Covid-19 region in the lung CT image. The model uses U-net with the same modified loss function to segment the Covid-19 region in CT images. The model uses 373 Covid-19 positive annotated images with 240:75:58 for training, testing and validation respectively and all the images are resized to 624*624. The model predicts the Covid-19 region for the non-annotated images of Covid-CT-dataset which is the collection of data from medRxvv [23], bioRxiv [24] and pyMuPDF [25].

Results and Discussion:

This section lists out results obtained from the various tasks as proposed in the model.

A. *Binary Classification Model*

Various CNN models and their evaluation metrics like Learning Rate, Accuracy, Precision, Recall and F1 Score are listed out in the table 1.

Table 1. Evaluation Metrics of various CNN Models

Model	Learning Rate (Adam)	Accuracy	Precision	Recall	F1 Score
VGG16	0.001	45.87	45.8	1.00	62.8
RESNET50	0.001	94.24	92.42	95.31	93.85
VGG16	0.0001	97.132	96.15	97.66	96.90
RESNET50	0.0001	97.132	96.88	96.88	96.88

The confusion matrix for various CNN models are listed out in Table 2 and their training times are reported in Table 3.

Table 2. Confusion Matrix

Model	Learning Rate	TN	FP	FN	RP
VGG16	0.001	0	0	151	128
RESNET50	0.001	141	10	6	122
VGG16	0.0001	146	5	3	125
RESNET50	0.0001	145	6	4	124

Table 3. Training Time

Model	Time (Approx)	Epochs
VGG16	1 Hour 30 Mins	100
RESNET50	2 Hour 30 Mins	100

All the models were trained using Google Colab Tesla K80 Free GPU 12 GB platform. Training Accuracy, Training Loss, Validation Accuracy and Loss for the automatic identification of Covid-19 regions on CT-images using VGG16 net model with a learning rate of 0.0001 is as illustrated in Figure 4.

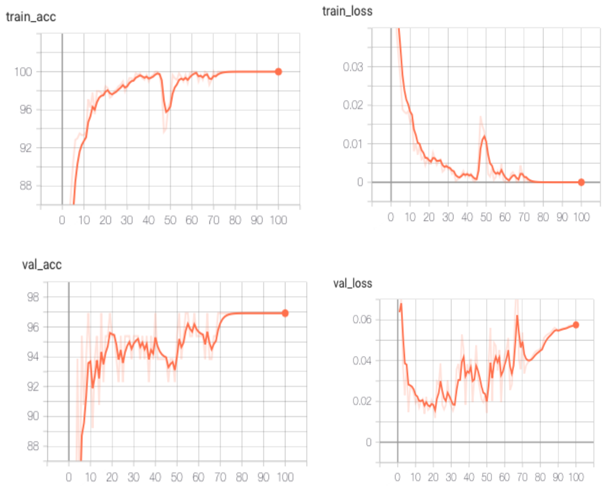


Fig.4. Classification Accuracy, Training Loss for VGG-16 Model

Training Accuracy, Training Loss, Validation Accuracy and Loss for the automatic identification of Covid-19 regions on CT-images using Resnet50 model with a learning rate of 0.0001 is as illustrated in Figure 5.

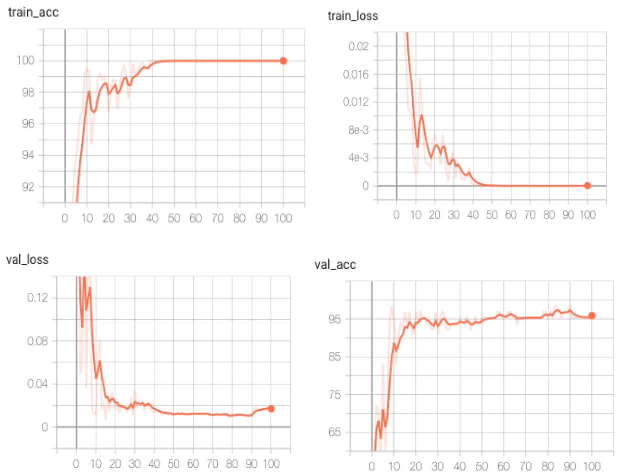


Fig 5. Classification Accuracy, Training Loss for Resnet50

Actually, predicted results against the ground truth given using VGG-16 Model, are illustrated in the Figure 6.

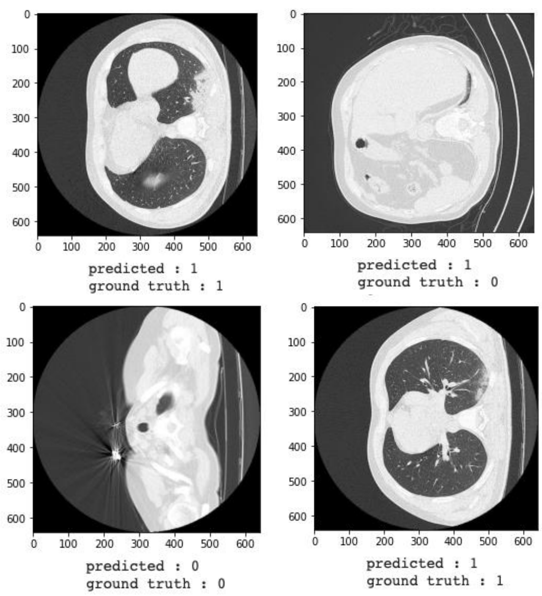


Fig.6. Classification prediction results against ground truth using VGG-16

Actually, predicted results against the ground truth given using the Resnet50 Model, are illustrated in Figure 7.

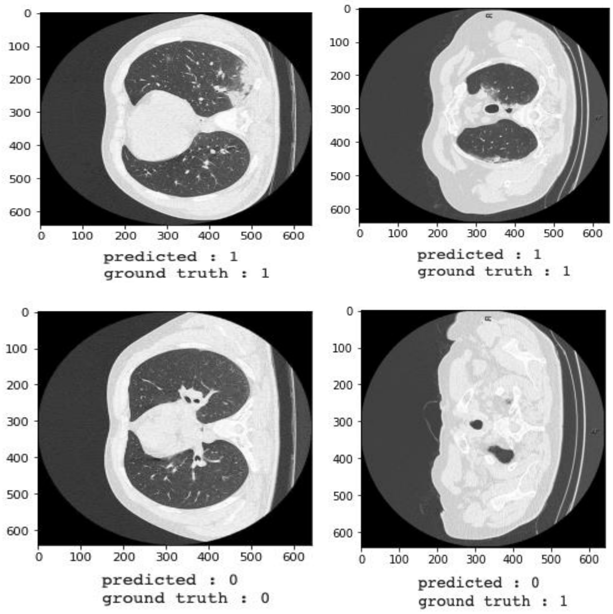


Fig.7. Classification prediction results against ground truth using Resnet50

With the Binary Classification model, VGG-16 network has produced an F1 score of 96.90 and Resnet50 network has produced an F1 score of 96.88.

B. Lung Detection Model

This model resizes the image of size 624*624 using nearest neighbor interpolation for the same reason not to lose local uniformity. The 9 axial volumetric data set contains 713 annotated images out of 829 CT images. The model uses 450 images for training, 214 images for testing and 49 images for validation. The model uses a modified loss function which is a combination of dice loss and binary cross entropy. It also uses a modified Adam optimize with a learning rate of .01 trained for 100 epochs on google Colab, Tesla K40 for an hour. This model produces an accuracy of 98.2, specificity 99.83, sensitivity 98.3, dice coefficient 97.21 and IOU 95.44.

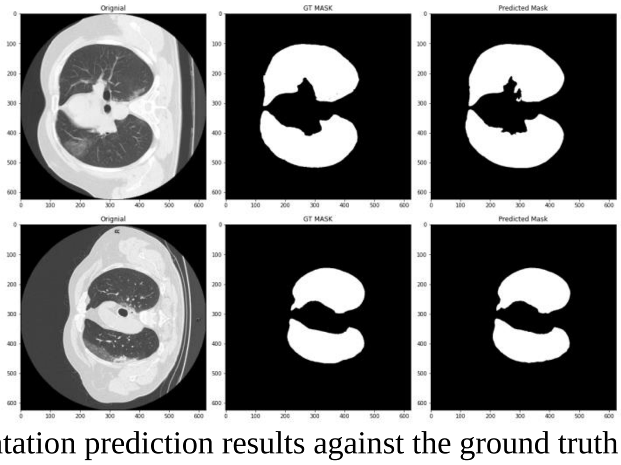
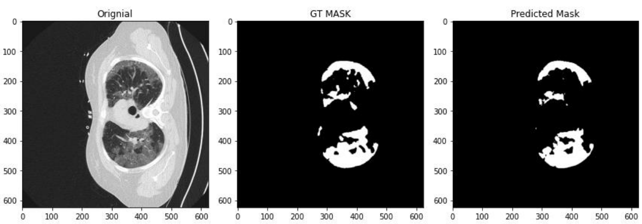


Fig.8. Lung segmentation prediction results against the ground truth mask using the U-Net.

c. Covid-19 Region Segmentation

The performance evaluation parameters of this model are, accuracy 99.4, specificity 99.5, sensitivity 80.83, dice coefficient 72.4 and IOU 61.59.



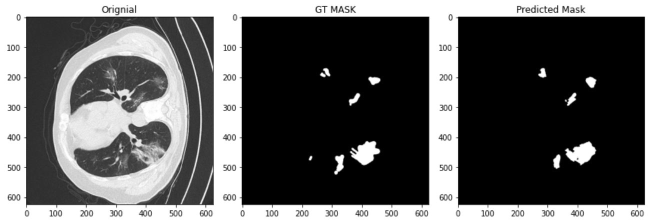


Fig.9. Covid region segmentation prediction results against the ground truth mask using the U-Net .

D.Discussion

The infectious disease Covid-19 spreads all around the world and the death rate increases at an exponential rate. As there is no vaccination available for Covid-19, the infected people need to be separated quickly and stay in quarantine places to avoid the speedy spread of the disease. Due to the shortage of the RT-PCR toolkit, CT images can be used as an alternative diagnosing toolkit for detecting early detection of Covid-19. As increase in bandwidth of CT images as well as the new Covid-19 virus consumes a lot of time and workload of the radiologist increases substantially. Deep Learning models can assist the radiologist by learning the features of Covid-19 by their annotated CT images. This paper proposes three deep learning models for the tasks such as binary classification of Covid-19, detection of lungs region on CT images and Covid-19 region identification on CT images. The proposed deep learning models produce an accuracy of 97%, 98% and 99% respectively. The results of the deep learning models show that the models can assist radiologists for quick, accurate and unbiased diagnosis for Covid-19.

Acknowledgement:

Dr.Rohit Tapadia, MBBS, MD, Director, Tapadia Diagnostic Centre for the Clinical and Biomedical Advisory and Evaluation.

Prof. Neil Gogte, Director, Keshav Memorial Institute of Technology for the Project Guidance, Finance and Material support.

References

[1] Richmen DD, Whitley RJ & Hayden FG. Clinical virology. 4th Edition, ASM Press. 2016.

[2] F.Shi, J.Wang, J.Shi et al. Review of Artificial Intelligence in imaging data acquisition, segmentation and diagnosis for Covid-19. IEEE reviews in Biomedical Engineering, 2020.

[3] Narges Saeedizadeh, Shervin Minaee et al. Covid TV-U-Net: Segmenting Covid-19 chest CT Images Using Connectivity Imposed U-net. arXiv:2007.12303v3. Aug 2020. Available from: https://arxiv.org/abs/2007.12303v2.

[4] Athanasias, Eftychios et al. Deep learning models for Covid-19 infected area segmentation in CT images. MedRxiv, 2020. Available from: https://www.medrxiv.org/content/10.1101/2020.05.08.20094664v2.

[5] Radiopaedia. https://radiopaedia.org/. (Accessed May 04, 2020).

[6] Arnab Kumar Mishra, Sujit kumar Das et al. Identifying Covid-19 from chest CT images: A Deep Convolutional Neural Networks based approach. Journal of Healthcare Engineering. Hindawi, Volume 2020.

[7] Xiaowei Xu, Xiangao Jiang, Chunlian Ma, et al. A Deep Learning System to Screen Novel Coronavirus Disease 2019 Pneumonia. https://doi.org/10.1016/j.eng.2020.04.010.

[8] Amine Amyar, Romain Modzelecoski et al. Multi task Deep learning based CT imaging analysis for classification and segmentation. Computers in Biology and Medicine, Elsevier, Volume:126, Nov 2020.

[9] J.Zhao,Y.Zhang et al. Covid-CT-dataset: A CT scan dataset about Covid 19. arXiv:2003.13865,2020. Available from: https://arxiv.org/abs/2003.13865.

[10] Covid19. http://medicalsegmentation.com/covid19/. (Accessed July 04, 2020).

[11] Qingsen Yan, Bo.Wang et al. Covid-19 Chest CT image segmentation - A deep convolutional neural network solution. arXiv:2004.10987v2,EESS IV, 2020. Available from: https://arxiv.org/abs/2004.10987.

[12] L.C.Chen, G.Papandreou et al. DeepLab: Semantic image segmentation with deep convolutional nets, Atrous convolution and fully connected CRFs. arXiv:1606.00915,2016.

[13] Deng-Ping Fan, Tao Zhou et al. Inf-Net: Automatic covid-19 lung infection segmentation from CT images. arXiv:2004.14133v4 (EESS.IV), May 2020.

[14] Dilibag Singh, Vijay Kumar et al. Classification of covid-19 patients from chest CT images using multi objective differential evolution based convolutional neural networks. Eur J Clin Microbiol Infect Dis, Apr 2020.

[15] Hancer E, Xve B, Zhang M. Differential Evolution for filter feature selection based on information theory and feature ranking. Knowledge-based Systems. Volume 140, 15 January 2018, Pages 103-119.

[16] NiBabel Package. Available from: https://nipy.org/nibabel/.

[17] https://neurohire.io/en/popular- networks/vgg16

[18] X. Zhang, J. Zou, K. He and J. Sun. "Accelerating Very Deep Convolutional Networks for Classification and Detection". IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 38, No. 10, pp. 1943-1955, 1 Oct. 2016. doi: 10.1109/TPAMI.2015.2502579.

[19] https://neurohire.io/en/popular- networks/resnet.

[20] Kaiming He, Xiangyu Zhang et al. Deep residual learning for Image Recognition. arXiv:1512.03385v1[cs.cv], 2015. Available from: https://arxiv.org/abs/1512.03385.

[21] Uni-freiburg. U-Net: Convolutional Networks for Biomedical Image Segmentation. Available from: https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/.

[22] Olaf Ronneberger, phillipp Fischer et al. U-net: convolutional networks for biomedical image segmentation. MICCAI, Springer, vol:9351:234-241,2015.

[23] medRxiv. The Preprint server for Health Sciences. https://www.medrxiv.org/.

[24] bioRxiv. The Preprint server for Biology. https://www.biorxiv.org/.

[25] Python bindings for MuPDF’s rendering library. https://github.com/pymupdf/PyMuPDF