

An Effective approach for implementing COVID-19 CT Lung Segmentation using Low-Cost System On Chip

Arvinth Srinivasan
Graduate Student: Dept. ECE
California State University, Fresno
California, USA
arvinth@mail.fresnostate.edu

Dr. Nan Wang
Professor: Dept. ECE
California State University, Fresno
California, USA
nwang@csufresno.edu

Abstract— Computer vision techniques always had played a salient role in numerous medical fields, especially in image diagnosis. Amidst a global pandemic situation, one of the archetypal methods assisting healthcare professionals in diagnosing various types of lung cancers, heart diseases, and COVID-19 infection is the Computed Tomography (CT) medical imaging technique. Segmentation of Lung and Infection with high accuracy in COVID-19 CT scans can play a vital role in the prognosis and diagnosis of a mass population of infected patients. Most of the existing works are predominately based on large private data sets that are practically impossible to obtain during a pandemic situation. Moreover, it is difficult to compare the segmentation methods as the data set are obtained in various geographical areas and developed and implemented in different environments. To help the current global pandemic situation, we are proposing a highly data-efficient method that gets trained on 20 expert annotated COVID-19 cases. To increase the efficiency rate further, the proposed model has been implemented on NVIDIA - Jetson Nano (System-on-Chip) to completely exploit the GPU performance for a medical application machine learning module. To compare the results, we tested the performance with conventional U-Net architecture and calculated the performance metrics. The proposed state-of-art method proves better than the conventional architecture delivering a Dice Similarity Coefficient of 99%.

Keywords— U-Net, Computed Tomography, COVID-19, Medical Image Diagnosis Introduction

I. INTRODUCTION

The COVID-19 pandemic, otherwise called the Coronavirus pandemic, is an existing global pandemic of coronavirus disease 2019, which is caused due to extremely intensive respiratory syndrome SARS-CoV-2. The first recognition of the covid virus was in December 2019 in Wuhan, China. The World Health Organization (WHO) proclaimed it as an emergency health crisis Worldwide in Jan 2020 and later pronounced it a pandemic in Mar 2020. Starting in Jun 2021 more than 181 million cases have been affirmed, with more than 3.93 million affirmed deaths ascribed to Coronavirus, making it probably the deadliest pandemic ever.

Despite the release of vaccination against the disease, most of the countries in the world are facing an increasing number of daily cases. In some countries like India, Bangladesh the COVID-19 virus got mutated into a much stronger strain which led to the increasing number of affected cases and deaths in a very short period worrying the entire world.

Over the last few decades, there have been various improvements in the clinical tools used to aid medical diagnosis, and one of the most important is the Computerized Tomography (CT) scan which enables visualization of internal organs, such as the lungs and their structures[1]. There is no computerized model available to delineate COVID 19 virus in chest CT scans which can be used for quantifiable assessment in predicting the severity [2]. The first challenge for the clinical diagnosis of lung diseases using CT scans is to demarcate the region of the lungs[1]. To support future diagnostic processes a radiological specialist must mark the lung regions manually which is a tedious process when handling a mass population of infected people.

In this work, we developed a Deep Learning Module which uses Convolution Neural Networks to automatically segment the areas of COVID-19 infection in a CT lung scan to predict the severity of the infection. A Convolutional Neural Network (CNN) is one of the artificial neural networks that is used to extract useful information from an image input. To achieve this process, it uses convolutional layers where the original image is applied with various filters to get a deep understanding of the image at the pixel level. These filters can be customized based on parameters that can be learned to extract the typical information for a particular task. One of the major advantages of this proposed work is its data efficiency in training the module as it only requires a total of 20 experts annotated COVID CT Lung scans. The proposed network architecture uses a deep partial U-net architecture.

The module is implemented on a system-on-chip called Jetson Nano manufactured by NVIDIA which is a small yet powerful computer that supports the running of parallel neural networks for image segmentation. The second advantage is the computational power required to run a machine learning module is minimized as less than 10 watts for maximum performance. Moreover, the proposed model proves in delivering the performance and accuracy compared to the preexisting works without compromising its efficiency.

The rest of the paper is organized as follows. Section II discusses the related works developed for segmentation tasks. Section III debriefs the hardware used in this experiment. Section IV explains the proposed work which includes a description of the dataset. A result comparison of the proposed work with conventional U-Net architecture is given

in section V and the paper ends with a conclusion and future scope in section VI.

II. RELATED WORKS

A wide variety of Computer Vision Techniques such as region growing [3], active contours [4], watershed [5], etc. are available in the literature that performs segmentation tasks.

Supervised and unsupervised[6] are the two broad categories for implementing segmentation tasks. Apart from these another method called semi-supervised models are available. the semi-supervised model utilizes mixed intervention of humans and software to perform segmentation tasks. These methods are used in different areas such as predicting the defects in software, sensor quality checks in manufacturing industries, etc. Especially, in the field of medicine, these artificial intelligence methods are used to perform segmentation tasks like Brain tumor segmentation [7], identification of white matter lesions [8], lung cancer segmentation [9], etc. All these segmentation tasks require complete or partial human contact to implement the models.

Current studies show that the U-net architecture can deliver promising results if a good quantity of labeled training cases is available [10]. In a model developed by Shan et al, their V-net structure achieved 91.6 dice with 249 CT scans [11]. The model proposed by Huang et al using U-Net [12] for segmentation of lung and infection on 774 experts annotated CT scans showed that the model can monitor the disease progression and reduce the burden of disease. These experts annotated CT scans require a professional diagnosis from radiologists which is merely impossible as it consumes a lot of manpower and time which the world doesn't have in the present-day scenario.

The main goal of this paper is to perform lung and infection mask prediction with limited training data. The major problems solved on this work are:

1) Although there are not many experts annotated CT scan data available on open source, this proposed model delivers a high-quality segmentation mask with a limited number of training cases.

2) A comparison between conventional U-net[14] and the proposed work in terms of accuracy and performance metrics has been given for readers to understand and deploy in their research areas since there is no proper works are available which provide a clear benchmark of these models.

Figure 1 represents the conventional U-Net architecture for biomedical image segmentation.

3) Power consumption is a major problem in running these machine learning models and this proposed work is implemented on the CUDA core of Jetson Nano which consumes only 10 watts of power to deliver high-quality results.

III. HARDWARE

A. Hardware Setup

To utilize the power of artificial intelligence from a low-cost device NVIDIA Jetson Nano developer kit is used. It is supported by Jetpack SDK and can handle the performance capabilities of modern artificial intelligence workloads. The software development kit (SDK) Jetpack includes the following,

- 1) NVIDIA drivers supporting full desktop Linux environment.
- 2) Libraries and APIs for computer vision and artificial intelligence.
- 3) Tools and documentation for researchers and developers.

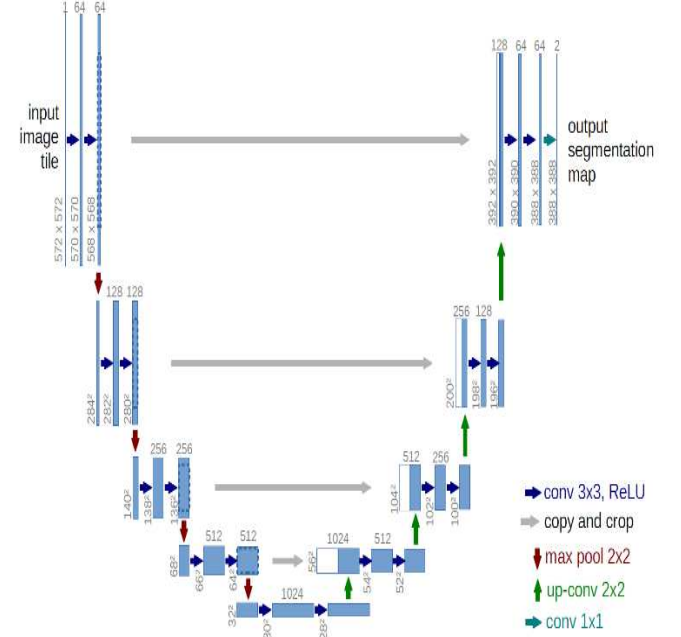


Fig. 1. Conventional U-Net Architecture [14]

The developer kit can be set up by flashing a microSD card with Jetpack components and operating systems by following the instruction in reference [16].

Figure 2 and 3 [15] shows the front view, rear view, top view and bottom view of the Nano board respectively. The configuration of the NVIDIA Jetson nano is described in Table I [16].

B. Power Consumption

One of the main advantages of implementing this module on Jetson nano is its power consumption is as less as 5 watts for less energy use and 10 watts for maximum performance [15]. The developer environment demands

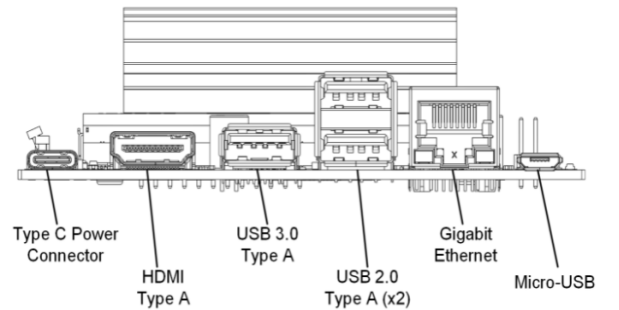


Fig. 2. Front view of Jetson nano

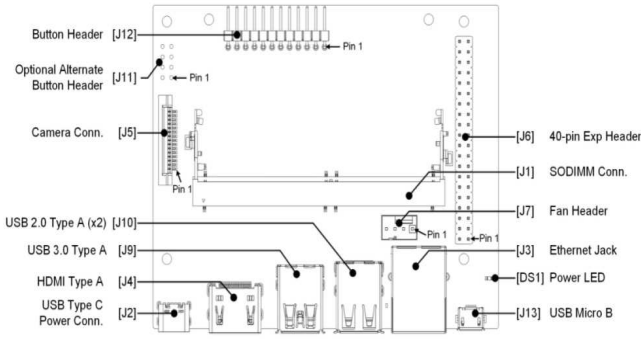


Fig. 3. Top view of Jetson nano carrier board

power usage only for the module, peripherals, and the carrier board. This allows us to deploy the model in low economic countries where power barrier is a major concern.

C. GPU

The data-parallel architecture of NVIDIA Jetson nano is an added advantage in handling the segmentation task. Since the device is contrived of 128 CUDA cores that are capable of 472 GFLOPS, the device can process each thread parallelly. These thread processors inside the GPU core performs similar instructions monitored by a shared controlled unit. These thread processors are called CUDA cores. This allows the device to accomplish calculation on individual pixels effectively thereby delivering high-quality segmentation masks.

TABLE I. JETSON NANO CONFIGURATION

S.No	Component	Description
1	GPU	128-core NVIDIA Maxwell™
2	CPU	Quad-core ARM® A57 @ 1.43 GHz
3	Memory	2 GB 64-bit LPDDR4 25.6 GB/s
4	Storage	microSD (64 GB)
5	Video Encode	4Kp30 4x 1080p30 9x 720p30 (H.264/H.265)
6	Video Decode	4Kp60 2x 4Kp30 8x 1080p30 18x 720p30 (H.264/H.265)
7	Connectivity	Gigabit Ethernet, 802.11ac wireless
8	Camera	1x MIPI CSI-2 connector
9	Display	HDMI
10	USB	1x USB 3.0 Type A, 2x USB 2.0 Type A, USB 2.0 Micro-B
11	Others	40-pin header (GPIO, I2C, I2S, SPI, UART) 12-pin header (Power and related signals, UART) 4-pin Fan header
12	Mechanical	100 mm x 80 mm x 29 mm

IV. PROPOSED WORK

A. Data Set

The data set consists of COVID-19 CT scans of 20 patients obtained from Coronacases Initiative and Radiopaedia [1] from an open-source [13] with CC BY-NC-SA license [1]. The COVID-19 infection is present in all the 20 CT scans with infection proportions varying from 0.01% to 59%. The data set is verified by radiologists in three different faces. Initially, it has been refined by 1 to 5 years of experience followed by cross verification of radiologists with 5 to 10 years of experience, and finally, annotation results were finetuned by senior radiologists experienced in chest radiology with almost 10 years of experience. Both pathological and normal regions are included in the whole

lung masks. ITK-SNAP performed the annotations manually slice-by-slice manner axially [1]. A total of around 300 infections with above 1800 slices are available with this dataset which on average takes 400 to 445 minutes for delineating each CT scan.

B. Proposed Architecture

Most of the classification tasks result in a class label for an input image. In medical image processing, the module should output the class label as well as localize the individual pixels present in the input image. As mentioned earlier it is difficult to obtain a medical image dataset containing a mass volume of training cases. In this method, we modified the U-Net [14] architecture to make it trainable with fewer training cases at the same time maintain the segmentation performance.

The architecture is designed in such a way the resolution of the image output is increased to get the most precision. The up-sampling operators replace the pooling layers in the contraction path of the succeeding layers. Once the resolution of the output is increased, the localization of the pixel takes place. This is implemented by combining the output from the up-sampling layer with the enhanced resolution features. Now, the succeeding convolution layers learn to arrange a more precise output. In this model, the number of convolution layers has been increased to enable the model to get the most accurate output.

The number of feature channels has also increased with the successive convolution layers, thus the growth path of the layers with high resolution gets the information without any loss in context. Ultimately, the architecture becomes symmetrical in contracting as well as the expansive path exhibiting a U-net shape but deeper in the architectural level with modified pooling layers which has never been done before.

Since the size of the image is larger, to prevent the model from losing the resolution due to GPU memory limitations, a tiling strategy [17] is used. The tiling strategy is used in predicting the pixels in the corner regions by mirroring the input image.

Figure 4 illustrates the network architecture of the proposed work. The left side represents the contracting paths, and the right side represents the expanding path. CNN is implemented on the contracting path. It consists of two 3x3 convolutions that are repeated twice with the same padding. The activation function used is “ReLU” which is abbreviated as a rectified linear unit. For downsampling, a max-pooling operation of 2x2 with stride 2 is performed. The feature channels are doubled in the downsampling step.

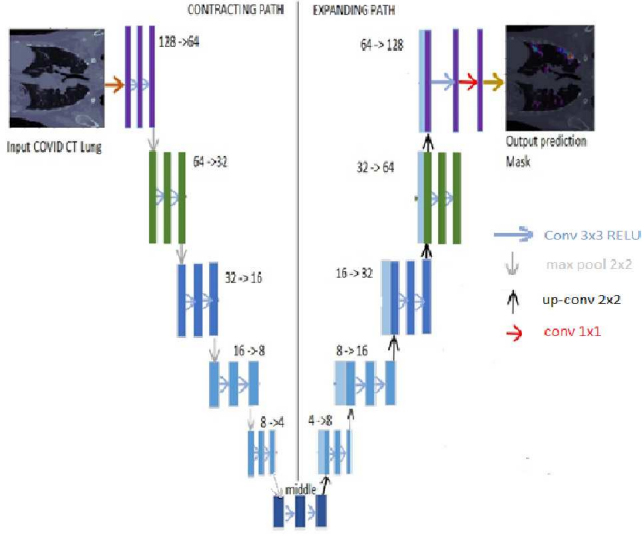


Fig. 4. Proposed deep U-Net architecture

In the expanding path, the feature map is upsampled before a 2x2 up-convolution. The number of feature channels is cut in half with this convolution layer, which includes two 3x3 convolutions, concatenation of featured map cropped from left part, and 2 'ReLU'. Since the border pixels get lost in every convolution layer, the cropping process is carried out. At last, a 1x1 convolution layer is used to extract and map the feature vector to and label the infection mask precisely.

C. Training

The training process consists of three tasks namely lung segmentation, infection segmentation, and both lung and infection segmentation. The fivefold cross-validation results are described for each of the tasks in the dataset [13]. Table II [1] represents the number of training cases available for each fold in the dataset.

TABLE II. TRAINING CASES

S.No	Task	Training and Validation
1	Fivefold cross validation	Fivefold cross validation
2	Lung	4 cases (20% for training)
3	Lung and infection	16 cases (80% for validation)

D. Implementation

As mentioned earlier, the model has been implemented on Jetson Nano, a system on a chip manufactured by NVIDIA. The Linux environment of the system helps in better navigation of the resources with ease of use. The libraries used for the design include the following :

- 1) *os* – standard library for python 3
- 2) *glob* – to import and export data trade information
- 3) *Pandas* – mainly used for data analysis
- 4) *numpy* – to create an array which is then passed to *pandas*
- 5) *nibabel* – to support operations on neuroimaging file formats [18]
- 6) *matplotlib* – to plot the performance metrics [19]
- 7) *tensorflow* – to support numerical computation and large-scale machine learning
- 8) *cv2* – used to resolve computer vision-related computations

Initially, the data is read and explored by reading the metadata file to gain a better understanding of the dataset. All the four different types of images, for instance, original CT lung, lung mask, covid virus mask, lung, and covid infection mask are available in this case. To read these images which are in 'NifTi' [19] format a separate function can be used. Since the image is rotated by 90 degrees, data reorientation is executed in the process. The above-mentioned 'nibabel' library is used to assist this. Once the volumetric CT scans with COVID / non-COVID infected original CT, COVID masks, lung masks as well as lung and infection masks are obtained the image slices from the sample CT are read. Figure 5 represents one of such images in the dataset.

In Figure 5 the first image is the original CT lung followed by the lung mask being marked in white and green color in the second image. The infection masks are represented in white color in the third image followed by the combination of both lung and infection masks which are visible on the fourth image.

The original CT lung and infections masks are considered for further analysis in this segmentation task. It is observable that each CT image has several slices and they are of different sizes.

Regardless of the CT scan's ability to give improvised resolution output of the tissue volume, the air and density volume are the two main quantities to get the region of interest. The Hounsfield unit [21] is the scale used to measure this. It varies from 1,000 HU to 0 HU approximately for air and water. The HU is 1,000 for bone.

The lung is composed of air density at 1000 HU and blood, water, and cells at 0 HU respectively. Figure 6 represents the HU scale plotted for one of the CT lungs images. This HU is converted into tissue and air content for the Region of Interest (ROI). Extracting only the ROI will helps in increasing the processing speed.

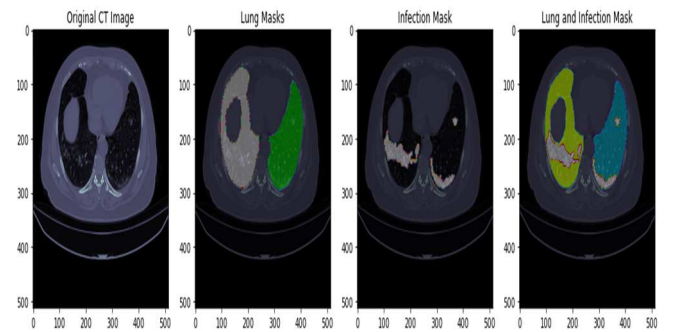


Fig. 5. Original CT, COVID masks, lung masks as well s lung and infection masks

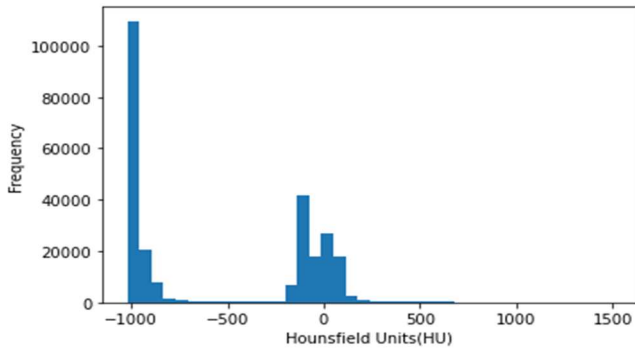


Fig. 6. HU vs Frequency plot of the lung content in CT scan

V. RESULT AND COMPARISON

The most common and widely used performance metrics for machine learning algorithms are the training loss and the validation loss over time. The compatibility of the model in fitting the training data is measured by training loss whereas the compatibility of the model in fitting the new data is measured using validation loss. In this work, we calculated the validation loss over the number of epochs.

In machine learning, the term epoch represents the total number of passes the model takes for the training dataset to run the algorithm. The number of epochs is the number of iterations when the batch size is equal to the entire training dataset.

One of the other performance metrics calculated in this work is validation accuracy. This is also known as testing accuracy which represents the precision calculated on the dataset not been used for training rather been used for testing. This often helps in the ‘early stopping’. This denotes that the model reaching its highest accuracy level.

The proposed model took a training time of 12,960 seconds in total. The graph for validation accuracy was maximum in epoch 9 with a value around 0.9942. An early stopping at epoch 9 can be performed to reduce the timing. In this work, a total of 10 epochs has been run to understand how the accuracy changes over time. Figure 7 represents the graph plot for Accuracy variation with the number of training epochs. The validation loss metrics obtained in this model are around 0.085. Figure 8 represents the validation loss with several training epochs.

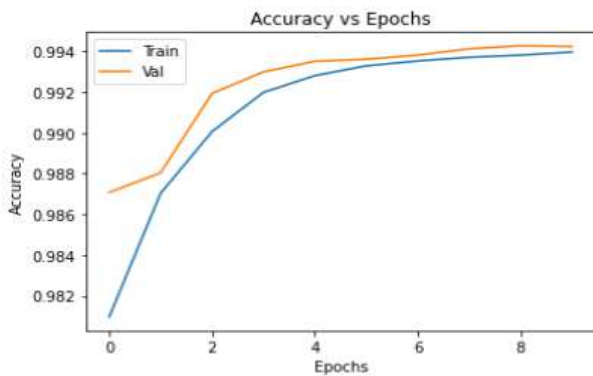


Fig. 7. Accuracy variation vs number epochs

Figure 9 represents the final output of the original COVID CT lung, infection mask, and predicted mask.

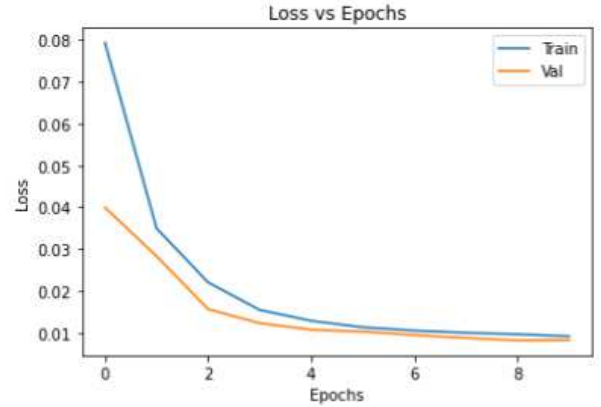


Fig. 8. validation loss vs number of epochs

The performance metrics were compared with the conventional U-net architecture and the network took around 16,000 seconds in total to accomplish the same task with an accuracy of around 0.95. The comparison results are tabulated in Table III. It is clear that the proposed model works the best for the segmentation task undertaken in this work and delivers better performance and accuracy. On top of that, the power consumed for the entire network is 10 watts without compromising the performance, thus making the model more reliable.

TABLE III. PERFORMANCE METRICS COMPARISON

S.No	Performance metrics	U-Net	DP-U-Net
1	Validation loss	0.12	0.08
2	Validation accuracy	0.95	0.99
3	Training period	16,000	12,960

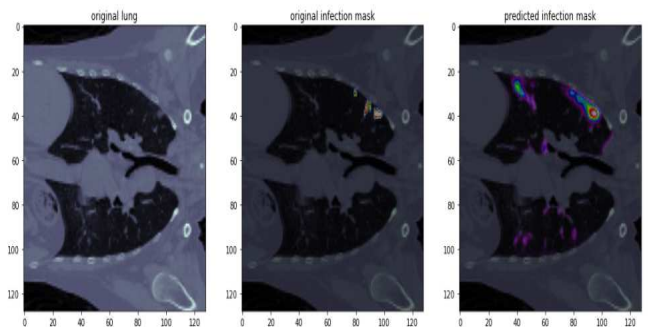


Fig. 9. Final predicted mask output

In Figure 9, the left image represents the original lung CT image which has the COVID 19 infection. The middle image contains the infection mask mapped based on the Hounsfield unit. The image on the right represents the predicted infection mask. The dark red circle in the image represents the core infection tissue and the blue and violet tint represents the presence of tissue in the developing phase.

VI. CONCLUSION

This work was focused mainly on implementing a machine learning module in a system on a chip. This work proves that the training models can be deployed on low-cost hardware without compromising performance. Moreover, the new architecture introduced in this paper can be used as a start for researchers to explore the possibilities of bringing machine learning models to low-cost devices.

Also, it is possible to implement different architecture to reduce the training time. Combining the possibilities of the low-cost device and increased performance proves the ultimate efficiency of the machine learning model.

REFERENCES

- [1] Hu, Qinhua, et al. "An Effective Approach for CT Lung Segmentation Using Mask Region-Based Convolutional Neural Networks." *Artificial Intelligence in Medicine*, vol. 103, 2020, p. 101792.
- [2] Fei Shan, et al. "Abnormal Lung Quantification in Chest CT Images of COVID - 19 Patients with Deep Learning and Its Application to Severity Prediction." *Medical Physics*, 2020, pp. Medical Physics, 2020-11-22.
- [3] Qian Zhao, Yu, et al. "Retinal Vessels Segmentation Based on Level Set and Region Growing." *Pattern Recognition*, vol. 47, no. 7, 2014, pp. 2437–2446.
- [4] Rebouças Filho, Pedro Pedrosa, et al. "Novel and Powerful 3D Adaptive Crisp Active Contour Method Applied in the Segmentation of CT Lung Images." *Medical Image Analysis*, vol. 35, 2017, pp. 503–516. J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73
- [5] Grau, V, et al. "Improved Watershed Transform for Medical Image Segmentation Using Prior Information." *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, 2004, pp. 447–458.
- [6] Chen, Xiang, et al. "Software Defect Number Prediction: Unsupervised vs Supervised Methods." *Information and Software Technology*, vol. 106, 2019, pp. 161–181.
- [7] Angulakshmi, M, and Lakshmi Priya, G.G. "Automated Brain Tumour Segmentation Techniques— A Review." *International Journal of Imaging Systems and Technology*, vol. 27, no. 1, 2017, pp. 66–77.
- [8] Kapouleas, I. "Computer Based Identification Of White Matter Lesions." [1990] *Proceedings of the Twelfth Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1990, pp. 396–397.
- [9] Li, Zhang, et al. "Deep Learning Methods for Lung Cancer Segmentation in Whole-Slide Histopathology Images-The ACDC@LungHP Challenge 2019." *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 2, 2021, pp. 429–440.
- [10] Huang, Lu, et al. "Serial Quantitative Chest CT Assessment of COVID-19: A Deep Learning Approach." *Radiology. Cardiothoracic Imaging*, vol. 2, no. 2, 2020, p. e200075.
- [11] Kaiming He, et al. "Deep Residual Learning for Image Recognition." 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
- [12] Hu, Xuegang, and Yang, Hongguang. "DRU-Net: a Novel U-Net for Biomedical Image Segmentation." *IET Image Processing*, vol. 14, no. 1, 2020, pp. 192–200.
- [13] Ma Jun, Ge Cheng, Wang Yixin, An Xingle, Gao Jiantao, Yu Ziqi, ... He Jian. (2020). *COVID-19 CT Lung and Infection Segmentation Dataset (Version Version 1.0)* [Data set]. Zenodo. <http://doi.org/10.5281/zenodo.3757476>
- [14] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.
- [15] Web link: <https://developer.nvidia.com/embedded/learn/jetson-nano-2gb-devkit-user-guide>
- [16] Web link: <https://developer.nvidia.com/embedded/jetson-nano-2gb-developer-kit>
- [17] Yu, Hanwen, et al. "Residues Cluster-Based Segmentation and Outlier-Detection Method for Large-Scale Phase Unwrapping." *IEEE Transactions on Image Processing*, vol. 20, no. 10, 2011, pp. 2865–2875.
- [18] <https://nipy.org/nibabel/gettingstarted.html>
- [19] https://matplotlib.org/2.0.2/users/pyplot_tutorial.html
- [20] <https://radiopaedia.org/articles/nifti-file-format?lang=us>
- [21] Thing, Rune Slot, et al. "Hounsfield Unit Recovery in Clinical Cone Beam CT Images of the Thorax Acquired for Image Guided Radiation Therapy." *Physics in Medicine & Biology*, vol. 61, no. 15, 2016, pp. 5781–5802.