Deep Learning for Detection of Tumours from CT Images of Multiple Organs

MSc Research Project
MSc in Data Analytics

Karan Mahendra Singh
Student ID: 18192726

School of Computing
National College of Ireland

Supervisor: Dr. Rashmi Gupta

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Karan Mahendra Singh
Student ID:	18192726
Programme:	MSc in Data Analytics
Year:	2020
Module:	MSc Research Project
Supervisor:	Dr. Rashmi Gupta
Submission Due Date:	17/12/2020
Project Title:	Deep Learning for Detection of Tumours from CT Images of Multiple Organs
Word Count:	XXX
Page Count:	19

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	31st January 2021

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Deep Learning for Detection of Tumours from CT Images of Multiple Organs

Karan Mahendra Singh
18192726

Abstract

Computer aided detection of tumors is yet a challenging problem since it is difficult to differentiate between tumors and other tissues or bodily components. Early detection of cancer could mean a matter of life and death as it allows more time for treatment and leads to higher odds of the patient surviving. With this in mind, several researchers have developed enhanced methods of detecting tumors. However, each method differs with the organ concerned and only a few studies have tried developing a universal system. In this study, we propose a model fit for using universally for detecting lesions from eight organs. The model was trained on augmented CT scan images using an adapted U-Net. It was finally evaluated by calculating specificity, dice coefficient, binary accuracy and loss. We achieved competitive results with the proposed model. Lastly, this study also applies watershed algorithm and color based segmentation for segmenting lesions, they were evaluated by comparing the predicted and true segment.

1 Introduction

Cancer starts developing when certain changes in the body affect the normal life cycle of cells. Their numbers surge to an extent that other normal cells are overpowered and cannot function normally. This leads to the formation of a tumour[1][2]. These tumors may be malignant or benign. In cases of a malignant tumor, early detection of cancer leaves more time for treatment and a higher probability of the patient surviving. Thus, methods of diagnosis had to be not only well performing, but also time bound. Thus, image processing found its application in medical sciences as body scans like CT (Computed Tomography), MRI (Magnetic Resonance Imaging), X-rays and so on were developed to initiate computer-aided diagnosis.

Eventually, computer-aided diagnosis started becoming more prominent for detecting cancer from scan images as robust models showed promising performances and automatic detection of cancer cells became its essential part Tang, Yan, Tang, Liu, Xiao and Summers (2019). Deep learning and CNNs performed better than other conventional machine learning approaches. However, such automated approaches required labeled data which requires medical expertise Yan, Wang, Lu and Summers (2018). Notwithstanding that, many researches such as Kopelowitz and Engelhard (2019) and Liu, Dong, Dong, Yu and Qi (2018) for detection of lung nodules, Furuzuki, Lu, Kim, Hirano, Mabu, Tanabe

[1]https://www.cancer.org/cancer/cancer-basics/what-is-cancer.html
[2]https://www.cancer.net/navigating-cancer-care/cancer-basics/what-cancer

and Kido (2019) for liver cancer, Jamil, Abbas, Umair, Habib, Hussain et al. (2020) for breast cancer and so on were performed. These and many other researches came forth with novel approaches and promising results, thereby refining the process.

A majority of approaches concerned themselves with finding solutions for an organ as per their choice. However, with all the developments and refinements in methods, there could be a unified solution for detection of tumors from all organs. This could be a major breakthrough for machine learning and computer-aided diagnosis. Yan et al. (2018) put together a dataset comprehensive of eight organs containing lesions, with the aim of helping researchers and forming a strong base for further research by developing a universal detector based on faster RCNN. Tang et al. (2019) soon followed by developing a ULDOR - Universal Detector with Mask RCNN. With the state of the art model, they achieved brilliant results. However, it is clear that there has been limited work of this sort. Such systems could be primary tools before the concerned specialist is visited. This study is inspired by this idea of developing a universal detector of tumors and therefore, proposes a model to see how well can deep learning be applied to a mixed group of CT scan images of different organs to detect tumors. Mask RCNN uses deep architecture and feature extractors for segmentation. This study tries a shallower model for the same task and achieve competitive results.

This research applies an altered U-Net to a dataset from Kaggle that contains CT scans of eight organs containing lesions. It further applies watershed and color based segmentation to find the best segmenting method.

This paper is divided into subsections as follows: Section 2 - Related work contains contributions of past researches in the same field, it is further organized into subsections to sort the works based on their approaches, Section 3 - Methodology describes our technical approach, Section 4 - Evaluation explains the evaluation metrics used for evaluating the model proposed in this study and presents the results, Section 5 - Discussion talks about the aim and findings and finally, Section 6 - Conclusion and Future Work - provides conclusion and brings suggestions for future work.

2 Related Work

This section holds an account of related work. It is further divided into subsections; grouping works according to their methods.

The importance of an early detection of cancer cannot be emphasized enough. Chaudhary and Singh (2012) used CT images for detection of lung cancer. Various enhancement and segmentation techniques were used with MATLAB to obtain the highest accuracy. Processing of images was divided into stages as follows: first enhanced using Fast Fourier, Gabor and auto enhancement. Further, watershed and thresholding techniques were used for segmentation. Features such as perimeter, area and roundness were extracted which helped establish the presence of tumours. Scales of these features also helped in determining the stage of cancer. Works are now usually divided into three stages described this research.

Botterill, Lotz, Kashif and Chase (2014) provided another perspective by proposing a technique using 3D surface motion for detecting breast tumours. The breast is hanged into the machine while being vibrated. Cameras take pictures of it in motion at various stages followed by the identification of its 3D profile. The skin surface motion is estimated and compared with the actual tracked surface. Some points are tracked throughout the movement to identify errors. The system detects smaller phase difference near the tumour as the organ is stiffer there. Results show reconstruction errors amount to almost 1 mm, and the approach is more accurate than its predecessors.

2.1 Systems using Support Vector Machine

Alam, Alam and Hossan (2018) detected and predicted the probability of lung cancer with multi-class SVM. After enhancement, segmentation, and detection of cancer from images, the stage of cancer was identified. The systems check for cancerous cells, if present, it identifies the stage of cancer, if not present, it computes the probability of cancer. Detection and prediction accuracies were 97% and 87% respectively. SVM works well with hyperplane classification which is fundamental to classifying cancer and hence the good results.

Kido, Hirano and Hashimoto (2018) used CNN and RCNN for diagnosis and detection of lung abnormalities, respectively. A small image dataset was used with no feature extractor hence, the pretrained AlexNet in conjunction with SVM as its classifier formed the model for computer aided diagnosis. Further, the computer aided detection system detected abnormalities in lungs with RCNN. Accuracy for diagnosis was 99% however, due to a lack of a comprehensive dataset, nodules could not be detected. Using the pretrained AlexNet was a smart decision, given they had few images to train with. This must have saved time required for training.

As it is difficult to narrow down the causes for the occurrence of lung carcinoma to prevent it from occurring, it should be detected in the initial stages Roy, Chaudhury, Burman, Ganguly, Dutta, Banik and Banik (2019). The user uploads a CT scan image to the website which is sent to the MATLAB-based server where it is processed, and the patient is diagnosed. Results of two classifiers based on SVM and Random Forest were compared to conclude that the SVM-based classifier performed better. For processing, the image was turned into grayscale and binarized (0,1). Subsequently, segmentation using watershed was performed to yield segmented lungs. The Speeded Up Robust Features (SURF) extracted features.

2.2 Systems using Deep Learning

Kumar, Jatti and Narayanappa (2019) used a two layered CNN where the first layer detected probable regions for identification of breast cancer. The second layer performed segmentation and attempted to reduce false positives. This approach successfully diminished the rate of false positives to 0.39. The segmentations are obtained considering real boundaries of objects in view thereby, more accurate than previous works.

An approach using Deep Convolutional Neural Networks (DCNN) for classification of tumours in lungs was designed by Mohanapriya, Kalaavathi and senthil Kuamr (2019). The classifier works in four segments: (1) tumour detection, (2) detection of malignancy, (3) classification of tumour into of benign or malignant tumour and (4) classifying the patient by calculating the probability of him having a tumour. The CNN was built with two convolutional and max-pooling layers, each and one softmax classifier and fully connected layer, each. The above model was then evaluated by calculating precision, recall, dice-similarity coefficient, and accuracy. Performances of three other architectures were compared, results of the proposed model showed that it performed better than them. The DCNN network learns feature maps by transforming first the input image by using a linear-based filter. Later the model uses a non-linear function to add a bias term.

Maximum Intensity Projection (MIP) images enhance the detection of tumours in lungs. Thus, a system involving the use of MIP images and CNN was proposed by Zheng, Guo, Cui, Veldhuis, Oudkerk and Van Ooijen (2019). The model augmented images of varying thicknesses to achieve better spatial information and differentiate well between tumours and blood vessels. The system operated in two steps: candidate detection for nodules and reduction of false positives. The first stage involved detection of candidate regions for the presence of nodules. Four CNNs are used for this purpose and subsequently, the four candidate regions are merged. Second step classified tumours. The research ended by determining that MIP images result in lesser false positives and higher sensitivity. Using MIP is beneficial for detecting tumours with lesser scans for evaluation of images.

Random forests and deep neural networks were combined for obtaining a higher degree of performance by Katzmann, Muehlberg, Suehling, Nörenberg, Holch and Gross (2020). RFs typically use a few features to determine a split for classifying objects into categories. This approach applies this idea of using limited features at a time and uses semantic features obtained from deep neural networks. The DNNs which are at the root nodes of the decision trees, split the data with extracted features and pass it on ahead to the further nodes. The decisive classification if performed by either leaf nodes or those of only one class. Results showed this method performed better with a small dataset than approaches with ResNet, ConvNet and RECIST.

Jamil et al. (2020) proposed a unique deep neural network CanNet and found it to be performing better than other state-of-the-art methods. The system is a feed-forward neural network which is trained with the traincgp function and tested with the cross-entropy function. The performance of the model is evaluated by calculating sensitivity, specificity, and accuracy, all scoring above 98%.

Jacob and Gopakumar (2020) bring a survey of existing technologies applied to detect tumours in lungs. Existing algorithms suffered from a high rate of false positives in detecting nodules. Deep learning algorithms resolved this issue. Further, it was noted that 3D dense networks performed better than others. When images are used, image pre-processing is necessary to bring them into a more understandable form for the machine. Techniques such as Median filtering, Gaussian filtering, Adaptive Histogram Equalization and so on are applied for pre-processing images. Furthermore, execution time may be reduced when the organ of interest is segmented, and surrounding body parts are removed. Segmentation approaches are automatic and semi-automatic. These methods leverage different aspects of the image for segmentation. Threshold-based approaches use the contrast between objects, while the fractal geometry method works on the shape of boundaries of objects. Lastly, classifiers are used to classify the nodules. Traditional machine learning methods do not perform satisfactorily. Consequently, deep learning algorithms were developed. However, the performance of these algorithms was restricted with the limited availability of medical images. This led to the development of transfer learning methods. This is how new methods and challenges were summarized.

Li, Zhou, Chen, Sun, Fu, Gong and Zhang (2020) studied the working of two models. One model consisted of a wavelet dynamic analysis used for extracting the lung parenchyma and achieve higher accuracy in detecting lung tumours. Another method had a genetic algorithm optimizing the CNN to extract features of the images. Parameters of the CNN must be optimized before the training starts. Further, the learning speed was increased by creating small batches of sample groups where, the weights were updated with every group. The research concluded by observing that the optimized CNN had higher accuracy.

Furuzuki, Lu, Kim, Hirano, Mabu, Tanabe and Kido (2019) proposed a technique of achieving a rectangular region of interest for a probable cancer site. After pre-processing, faster R-CNN is used to obtain the rectangular region of interest containing cancer. Faster R-CNN works in three parts: backbone, region proposal network (RPN) and head. Backbone extracts features; VGG and ResNet 50 and 101 are used here. RPN discovers object regions. Sliding window along with anchors were used in this layer. RoI pooling layer collects region candidates and outputs a feature map. Finally, the position of the object is determined in the head. Average precision of 64% proved ResNet101 detected better than ResNet50. Therefore, a method of detecting cancer is developed and a strong base for further research involving segmentation and/or classification of the stage of cancer could be done hereafter.

Giddwani, Tekchandani and Verma (2020) used 3D images to learn volumetric information from medical image data using a multi-rate deep dilation network in V-Net as it performs well. The research used a Weighted-Fusion Loss (WFL) to prevent data imbalance between bright and dark pixels. The initial weights for the model were set with Kaiming initializer. Sensitivity and precision of 0.87 and 0.97 was attained. This approach achieved a leading-edge performance in segmentation of pancreas.

2.3 Systems using Mask RCNN

Anantharaman, Velazquez and Lee (2018) applied the state of the art – Mask RCNN for the detection and segmentation of oral diseases. The implementation employs two combinations: ResNet 101 and FPN and ResNet 50 and FPN. Thereafter, the top anchors of RPN predictions were picked and passed to the last stage: ROI classifier regressor. Finally, the segmentation masks are created for the regions selected by the ROI. Dice coefficients are calculated for comparing the results of the two models. Asper the results obtained, ResNet 101 performed better than ResNet 50. Results showed accuracies of more than 70% for both cold and cankers, with more images for training. The study was limited with few images to train from and inclusion of more images to bring in variation would broaden the scope.

Mask RCNN was used to segment lung nodules by Liu et al. (2018). Processing such as segmentation of lungs, pore filling and corrosion expansion was performed to obtain a smooth image. The Mask R-CNN backbone consisted of classification Network: ResNet 101 and detection network: FPN. Model was trained with: (1) using own data to tune all layers on the trained model. (2) train the heads and fine tune all layers. Higher accuracy of 0.80mAP was obtained in the later. The (2) method of fine tuning the model is a smarter way and could be used to improve the accuracy of the model.

Qadir, Shin, Solhusvik, Bergsland, Aabakken and Balasingham (2019) state that polyps are sometimes missed by physicians during colonoscopy thus, automatic detection and segmentation is required to be performed. Mask RCNN is used along with various other CNNs as feature extractors. By adding newer polyp images, the performance change is observed to check whether better datasets or deeper and more complex CNNs are required. Various feature extractors such as ResNet50, ResNet101 and Inception ResNet are deployed to evaluate their performances. Resnet50 performed better in all evaluation criteria since deeper nets require larger datasets. Inception Resnet improved the most after adding newer images which shows that it can extract more knowledge from larger training data. Results conclude that given a large dataset, a deeper CNN is not necessary. Deeper models could lose some spatial information because of the contraction and pooling layers but obtain higher level of features.

Kopelowitz and Engelhard (2019) used Mask RCNN on 3D images to detect and segment lung nodules from CT images. It detects objects in 2 steps: RPN and R-CNN. Inception ResNet v2 is applied in 3D to obtain a feature map. RPN proposals having scores less than 0.1 are fed to the ROI Align layer. RCNN employs convolution layers on the aligned proposals. Evaluation was done with the Competition Performance Metrics (CPM). Results confirm that Mask-RCNN could be applied to 3D images to achieve good results (CPM score of above 85%) for detection as well as segmentation. Inception v2 was appropriate as deeper models could lose some information Qadir et al. (2019).

2.4 State of the art

Yan et al. (2018) built a universal detector which detects lesions in organs from across almost the entire body with faster RCNN. CT images of organs containing lesions and RECIST were combined to form a comprehensive dataset containing annotations of tumours. The backbone of the system was formed by VGG-16. The last two pooling layers were discarded to yield feature maps of better resolution and increase the sampling ratio of samples that contained tumours as they are small and scarce in an image. The feature maps were then given to a Region Proposal Network followed by RoI. Two FC layers then established confidence scores. The system achieved 81% sensitivity with five false positives per image.

Detection of lesions from CT images is a tough task due to similar appearances of lesions and tissues. A universal detector based on Mask RCNN was built by Tang et al. (2019). Pseudo-masks were used for regions with tumours. Moreover, a hard-negative example mining method was proposed for reducing false positives. Based on the RECIST annotations, pseudo-masks are generated which makes it feasible and easy to train the model. A higher degree of sensitivity than Yan et al. (2018) with five false positives per image was obtained. The pseudo masks used here helped the machine understand lesions better. ResNet-101 has fewer parameters than VGG, limiting learning of features. However, with the high number of parameters in deep learning, the model could overfit.

Zhang, Gao, Xing, Chen and Zhang (2019) deployed a model called DualRes-UNet. This model applied continuous down-sampling layers to obtain structures in the image. Since down sampling results in loss of intricate details, concatenation is used to preserve features, but it still leads to a downside of localization in U-Nets. Researchers altered the down sampling process by incorporating addition operations. This helps retain the low-level features. Results conclude that the approach performs well on angle artifact reduction.It is stated that artifacts are globally distributed which is why receptive layers of CNN are essential, but this leads to loss in details. Therefore, this research studies how the change in dilation rate helps or not in preserving details resulting in better segmentations.

Table 1: Summary of related work

Author(s)	**Scope**	**Title**	**Findings**
Yan et al. (2018)	Build a universal lesion detector with faster RCNN.	DeepLesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning.	Developed an integrated framework for detecting lesions of all types.
Tang et al. (2019)	Universal detector based on Mask RCNN.	Uldor: a universal lesion detector for CT scans with pseudo masks and Hard negative example mining.	Improved the above approach by using Mask RCNN.
Zhang et al. (2019)	DualRes-UNet for obtaining better low level features.	DualRes-UNet: Limited Angle Artifact Reduction for Computed Tomography.	Performs effectively on limited angle artifact reduction.
Chaudhary and Singh (2012)	Reduce the time required for diagnosis of cancer.	Lung Cancer Detection on CT Images by using Image Processing.	Calculated area, perimeter, eccentricity, and intensity of tumours
Botterill et al. (2014)	Imaging of surface motion of breasts is used to detect tumours.	Reconstructing 3-D Skin Surface Motion for the DIET Breast Cancer Screening System.	Showed a way of finding tumors with 3D surface motion.
Kido et al. (2018)	Developed computer aided diagnosis and detection.	Detection and Classification of Lung Abnormalities by Use of Convolutional Neural Network (CNN) and Regions with CNN Features (R-CNN).	Feature extractors are not essential for computer aided detection and diagnosis.
Alam et al. (2018)	SVM for detection and prediction of lung cancer.	Multi-Stage Lung Cancer Detection and Prediction Using Multi-class SVM Classifier.	Detected tumor and calculated probability of developing, with good results.
Liu et al. (2018)	Detected edges of lung nodules with ResNet-101 and FPN.	Segmentation of Lung Nodule in CT Images Based on Mask R-CNN.	Detected nodules along with their contour information.
Anantharaman et al. (2018)	Identified cold and canker sores in the mouth.	Utilizing Mask R-CNN for Detection and Segmentation of Oral Diseases.	Achieved promising results with Mask RCNN.
Kopelowitz and Engelhard (2019)	Used 3D CT scan images for detection of lung nodules.	Lung Nodules Detection and Segmentation Using 3D Mask-RCNN.	Mask RCNN can achieve competitive results.

3 Methodology

This section describes our methodology. It is essential to detect cancer in preliminary stages to leave more time for treatment and improve chances of survival of the patient. Therefore, this research builds a general model based on UNet to identify cancer from more than one specific organ. This section describes the approach in detail.

We modified the basic UNet approach applied on Kaggle[3] to obtain better segmentation results. Later we applied Watershed and OpenCV based approaches to detect lesions and find the best segmenting approach.

Figure 1 shows the steps followed in the process.

Figure 1: Methodology for detecting lesions

3.1 Data collection

The research was conducted on the DeepLesion dataset originally from NIH[4]. However, the dataset was too large and was made available in a much smaller size on Kaggle[5]. The original dataset contained over 32,000 CT scan images of lesions in eight organs (Abdomen, Bone, Kidney, Liver, Lungs, Mediastinum, Pelvis and soft tissue). The subset of this dataset, used in this study contained over 1300 images. In addition, it had a csv file containing information such as file names, types of lesions, patient gender, bounding box coordinates for forming segmentations and so on. Certain analyses such as 2 and 3, and more described in the configuration manual, were performed in the research to find out more information about the patient data.

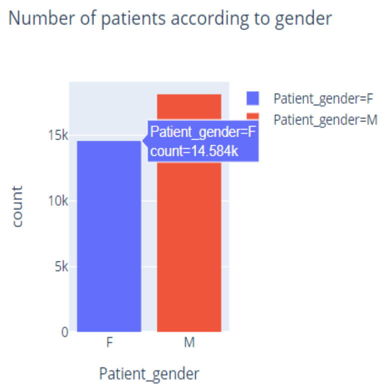

Figure 2: Patients according to gender

[3]https://www.kaggle.com/kmader/simple-lesion-segmentation-u-net

[4]https://www.nih.gov/news-events/news-releases/nih-clinical-center-releases-dataset-32000-ct-images

[5]https://www.kaggle.com/kmader/nih-deeplesion-subset/data

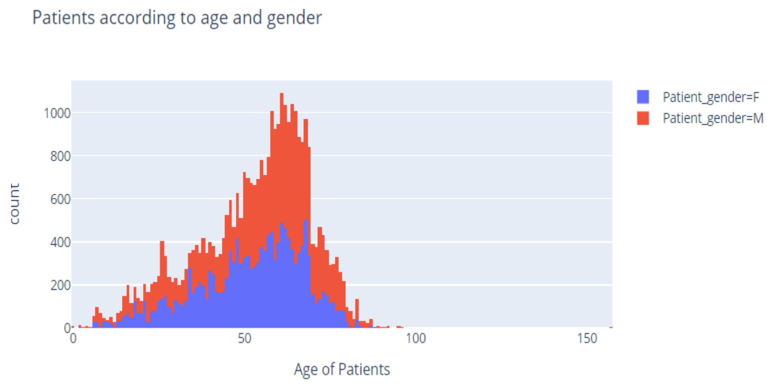

Figure 3: Age group and gender having most cases of lesions

3.2 Data preparation

The lession segmentations were provided in a column in a csv file. The file contained another column containing file names which helped track the file name and its corresponding lesion coordinates. Bounding boxes were extracted from the csv and stored in a list. The images and segmentation points were read in a numpy array. This was later used to train the model. We faced major challenges in mapping the image file to the corresponding segmentation. Thus, the procedure described by a notebook on Kaggle[6] proved majorly helpful in this task. After plotting the bounding boxes in the image with a filled yellow rectangle, it resulted in an image showing the location of the lesion, Figure 4.

Figure 4: Marked lesion

There were a total of 1350 images and the final shape of the lists holding images and down sampled segments was (1350, 512, 512, 1) and (1350, 128, 128, 1), respectively. Further configurations were as follows:

- Environment setup: This study made use of Google Colab and its GPU environment for the development and execution of the code. The GPU majorly helped in

[6]https://www.kaggle.com/kmader/deeplesion-overview

To be used for generating images for Water Segmentation model

```
[ ] img_read = lambda x: imread(x).astype(np.float32)-32768

    _, test_row = next(dl_df.sample(1, random_state=86).iterrows())
    fig, ax1 = plt.subplots(1, 1, figsize = (10, 10))
    c_img = img_read(test_row['location'])
    ax1.imshow(c_img, vmin = -1200, vmax = 600, cmap = 'gray')

    # This line may be excluded for generating image without bounding boxes of lesions
    ax1.add_collection(PatchCollection(masks(test_row), alpha = 0.25, facecolor = 'yellow'))

    ax1.set_title('{File_name} Gender:{Patient_gender}, Age:{Patient_age}, File:{File_name}'.format(**test_row))
    print('{File_name}'.format(**test_row))
    fig.savefig('/content/drive/My Drive/unsegmented_img.png', dpi = 300)
```

Figure 5:

reducing the training time by accelerating procedures. Nonetheless, the entire code may be run on base machines, but with configurations as follows:
RAM: 8GB, NVIDIA GeForce GTX 1050 Ti graphics and Processor: Intel(R) Core(TM) i7-8750H CPU. Other configurations might as well work, but the base machine used to develop this project had the above specifications and they were found to be sufficient. Furthermore, the dataset was stored on Google Drive and it was therefore essential to load it at the time of execution.

- Data Transformation: Certain columns in the csv file were transformed and some others were created to perform necessary tasks and visualizations. The Bounding_boxes column contained entries with "," which made it difficult to map segmentations to the image. Therefore, a new column was formed by extracting the coordinates. Further, the dataset was split into two sets of training and testing, of 70% and 30% composition, respectively. The model was trained using the training set and evaluated with the testing set.

- Data Augmentation: Just before the model was fit, the training set was augmented using ImageDataGenerator from keras.preprocessing.image. The images were augmented using parameters such as rotation_range, width_shift_range, brightness_range and zoom_range.

Lastly, Figure 5 shows the code segment necessary to run for generating images used for Watershed algorithm and color based segmentation algorithms.

3.3 Model

The research applied a basic UNet as well as its own proposed model which was an altered U-Net on the data described above and compared the results. Further, it applied the Watershed algorithm and a color based segmentation approach to segment lesions, by using OpenCV.

Zhang et al. (2019) states that artifacts are globally distributed in U-Nets which is why receptive layers of CNN are essential, but this leads to loss in details. Therefore, this research studies how the change in dilation rate helps or not in obtaining better segmentations.

In the basic UNet model, dilations are performed using SpatialDropout2D() layers. In this study, we bring a comparison of performances by altering the dilations performed by UNet. The images augmented are zoomed in to a greater extent, brightness is lowered to avoid over lighting certain parts of the image and dilation rates are reduced. Such augmentation is expected to work against the lower dilation rates in convolutions. The system is run with varying combinations of dilations to see what difference it makes, or not. The working of the model and its architecture is described in Figure 6 and 7, respectively.

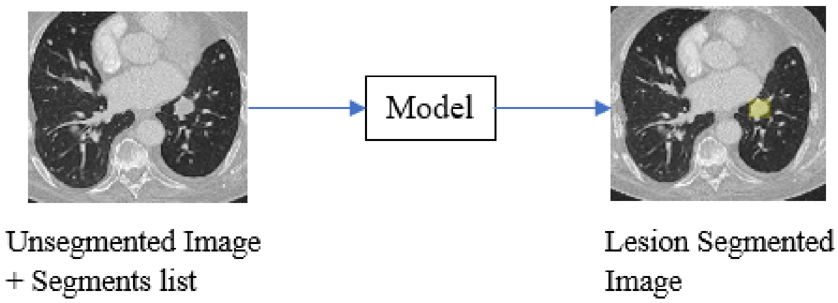

Figure 6: Working of the model

4 Evaluation

This section explains how the systems were evaluated. Metrics such as loss, true positive rate, dice coefficient and binary accuracy were calculated to study the performance of the systems based on U-Net. Graphs of how these metrics changed through the training process were plotted and the model was evaluated against the test dataset. Metrics used are as follows:

- Loss
- Binary Accuracy
- Dice coefficient: (2 * Area of overlap) / (Total pixels in both images)
- True positive rate or Sensitivity: (TP / Total) * 100

4.1 Basic U-Net

Values of the above mentioned metrics during training were plotted to gain insights on how the machine gained knowledge with the passage of every epoch, Figure 8.

After the machine was trained, it was evaluated on the test dataset to find the results shown in Figure 9. Furthermore, the machine was used to predict masks of images in the test dataset which it did as shown in Figure 10 .

4.2 Proposed model

The proposed model was evaluated using the same evaluation method. However, this model was tested with more than one dilation rates. The first time, it was run with dilations 2, 4, 6, 8, 12 and 18, Figure 11, Figure 12.

Later, the system ran with dilations 2, 4 and 6, Figure 13, Figure 14.

Lastly, the system ran with just dilations 2 and 4, Figure 15, Figure 16. The evaluation metrics are shown in Figure 17.

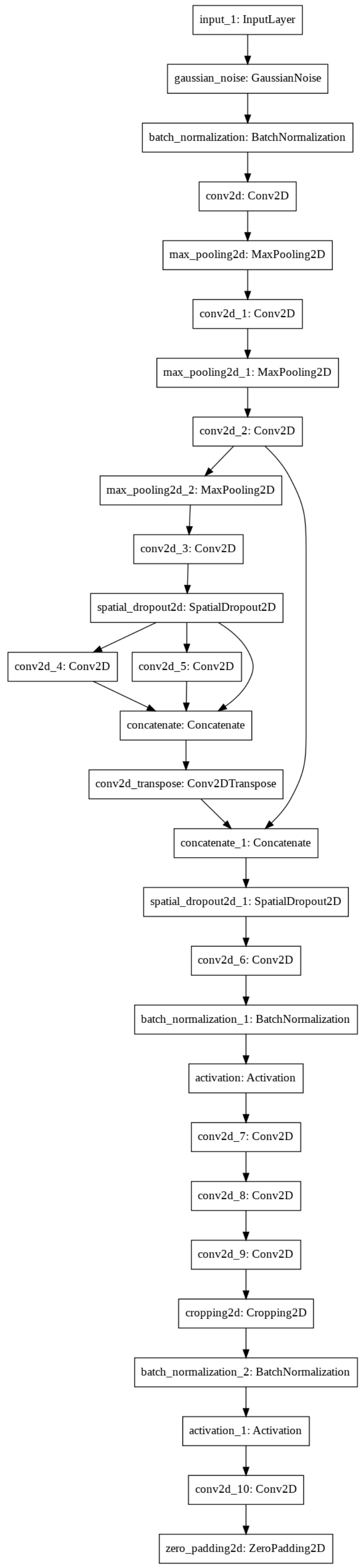

Figure 7: Architecture of the proposed model

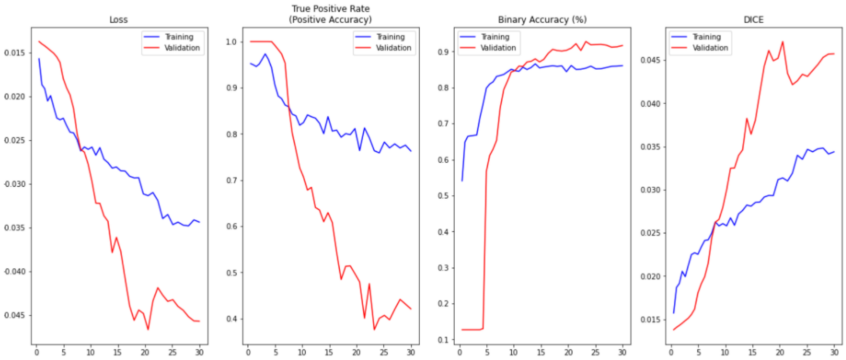

Figure 8: Training of Basic U-Net

```
print("test loss:", results[0])
print("test dice:", results[1])
print("test binary accuracy:", results[2])
print("test true positive rate:", results[3])
```

```
test loss: -0.05511686950922012
test dice: 0.055907074362039566
test binary accuracy: 0.9261903166770935
test true positive rate: 0.40928414463996887
```

Figure 9: Evaluation of Basic U-Net

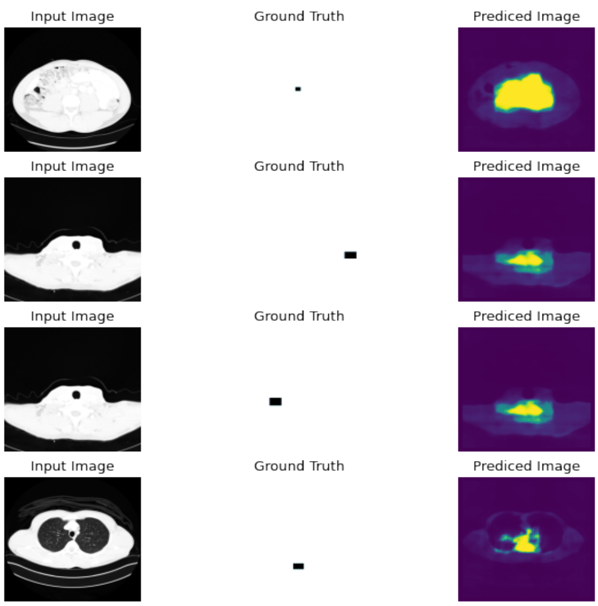

Figure 10: Segment prediction of basic U-Net

4.3 Watershed Algorithm

Watershed algorithm was run on one image at a time which was generated by using the code in Figure 5. Further, this algorithm was evaluated by observing its segmentations,

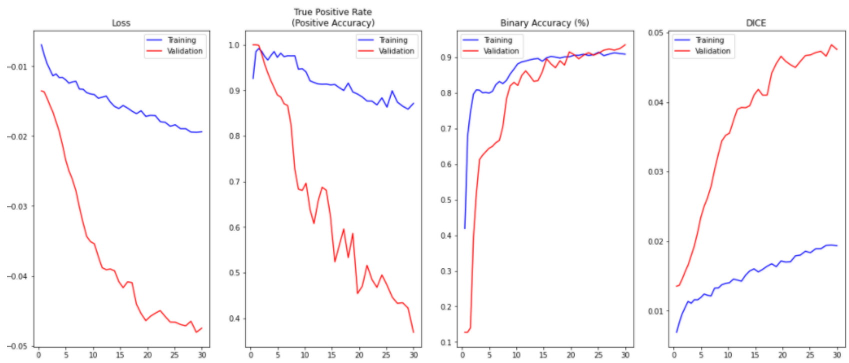

Figure 11: Training with dilations 2, 4, 6, 8, 12 and 18

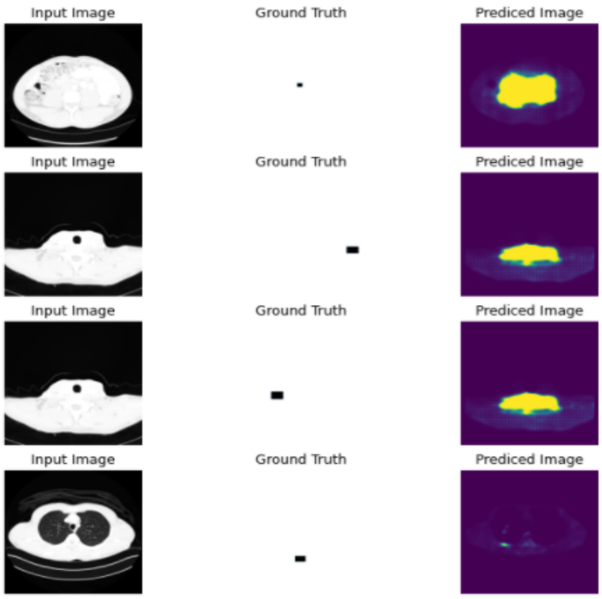

Figure 12: Segment prediction with dilations 2, 4, 6, 8, 12 and 18

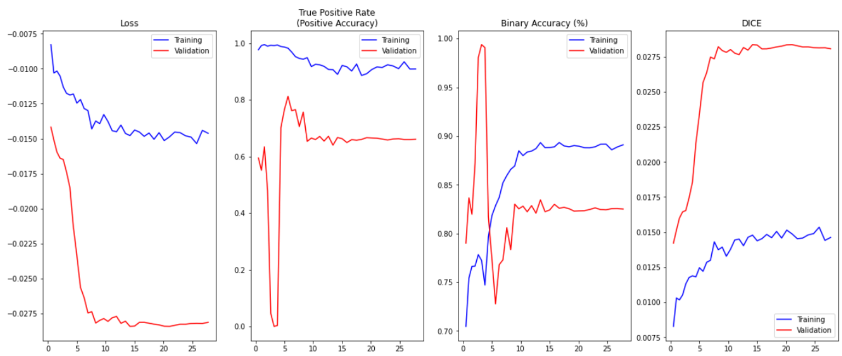

Figure 13: Training with dilations 2, 4 and 6

shown by green lines against the expected lesion segmentations, Figure 18.

4.4 Color based segmentation

This method was evaluated like Watershed Algorithm. Shown in Figure 19.

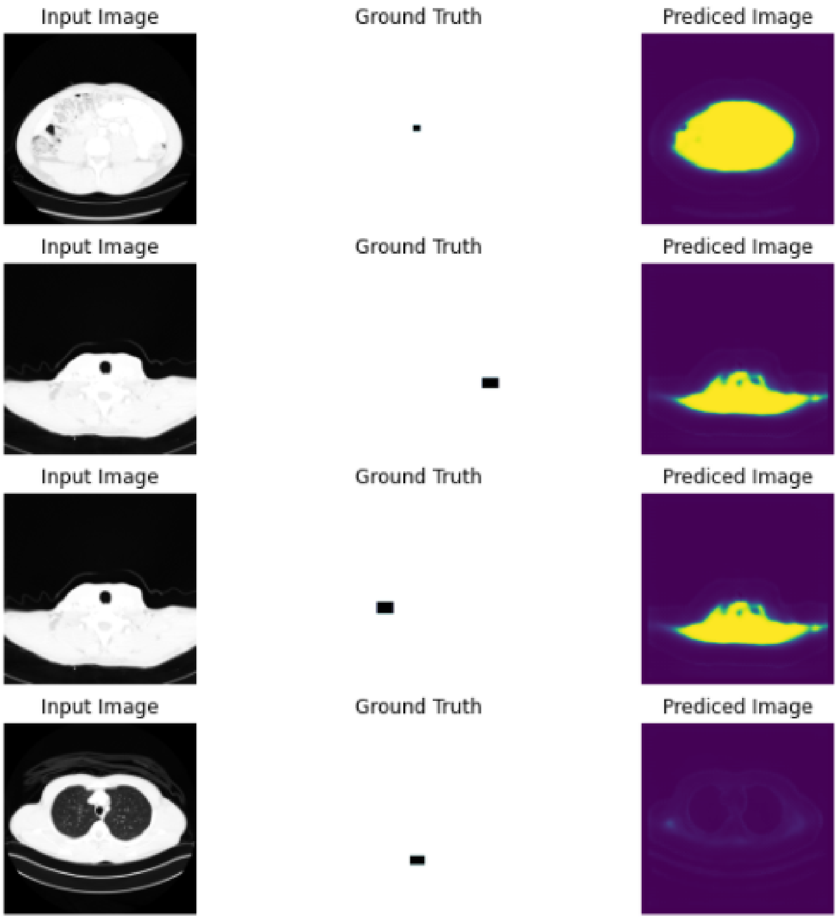

Figure 14: Segment prediction with dilations 2, 4 and 6

Figure 15: Training with dilations 2 and 4

5 Discussion

We presented a deep learning approach to detect tumours from CT scan images of multiple organs. The study aimed to explore how well deep learning can be applied to detect tumours. It intended to find a universal approach to finding tumours from all organs and used a set of eight organs for training and evaluation of the proposed model.

From the evaluation metrics in Figure 9 and Figure 17, the DICE coefficient of Basic U-Net is more significant than the proposed model. Essentially, this means the prior predicted segmentations which overlapped with true segmentations more. However, sens-

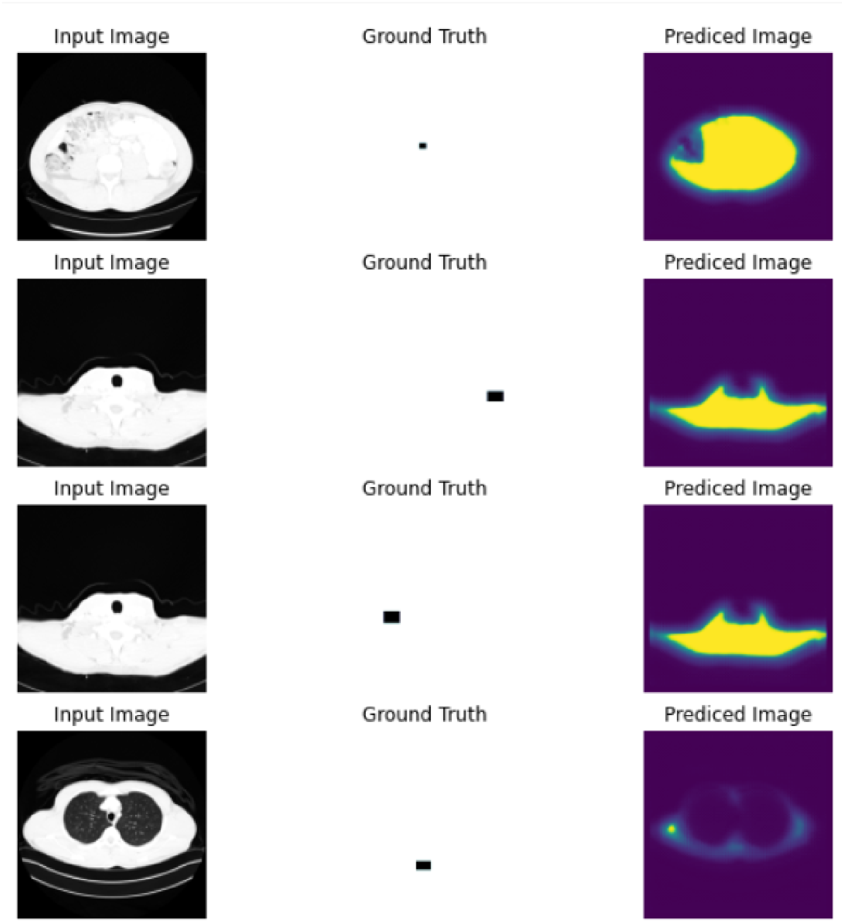

Figure 16: Segment prediction with dilations 2 and 4

```
print("test loss:", results[0])
print("test dice:", results[1])
print("test binary accuracy:", results[2])
print("test true positive rate:", results[3])
```

```
test loss: -0.030674900859594345
test dice: 0.030852258205413882
test binary accuracy: 0.7954323887825012
test true positive rate: 0.7507731318473816
```

Figure 17: Evaluation of proposed model

itivity of the proposed model is significantly higher than the basic model. While the binary accuracy of the proposed model is lesser than the basic U-Net model, it is closer to the sensitivity measure. This means from whatever predictions the proposed model made about the segments, most of them were right. This problem could possibly be overcome with more training. Tang et al. (2019) achieved a higher sensitivity rate by making use of Mask RCNN and pseudo masks in the system.

Figure 10 shows better segmentations of lesions than those in Figure 12 in that, the

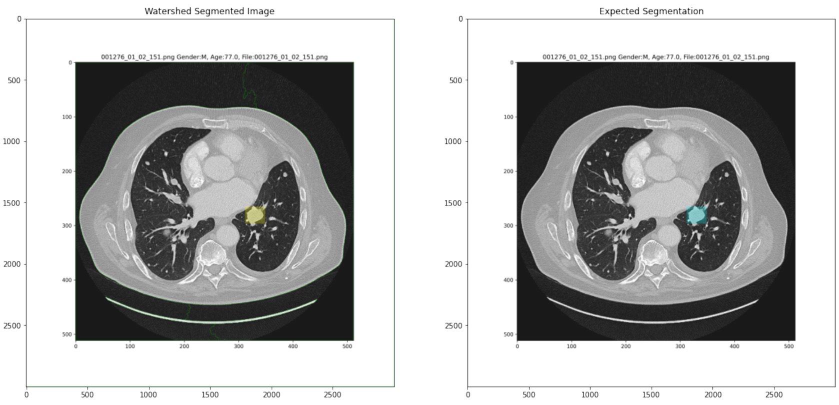

Figure 18: Segmentations with Watershed Algorithm

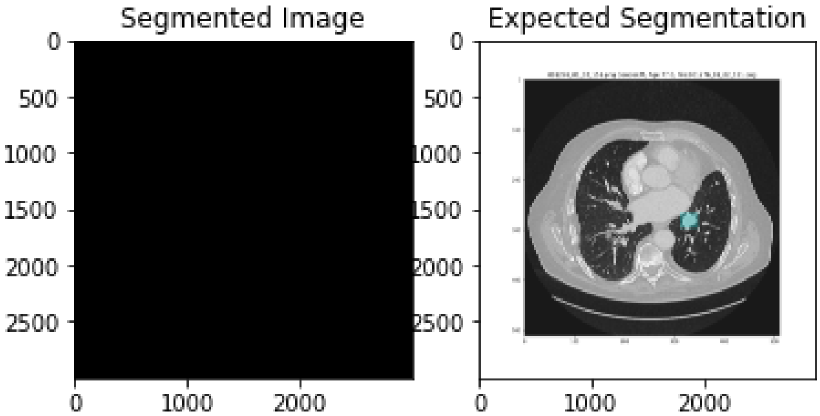

Figure 19: Segmentations with Color based segmentation

segmentations formed are smaller in size and cover area over the true lesions shown in the ground truth column. Segmentations formed by the proposed model although coincide with the area marked by the ground truth, are larger in size. Thus, showing bigger tumours than actual. The proposed model could not neatly identify the lesion in the last image. Whereas the basic U-Net could recognize it in the same image.

This approach performed a little lower than the state-of-the-art method for segmentation of lesions. Mask RCNN is known for achieving superior performance in almost all applications. This approach tried to apply a shallower model to achieve better results. Thereafter, watershed and color based segmentation algorithms could not segment lesions satisfactorily. The color based approach was provided a wide range of yellow color RGB values and yet, it could not find the color from the input image.

6 Conclusion and Future Work

This approach showed a way of applying U-Net, a deep learning method for detection of tumours from CT scan images of multiple organs. The study aimed at developing a universal approach for tumour detection and successfully completed the task with competitive results. The results showed that the reduced dilation rate did not lead to a major difference in performance. Better augmentation of training images should probably help improve performance.

In future scope, application of models like attention-based Multiple Instance Learning (MIL) could be tried. Further, usage of pre-trained models over medical datasets could yield some interesting results.

References

Alam, J., Alam, S. and Hossan, A. (2018). Multi-stage lung cancer detection and prediction using multi-class svm classifie, *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, IEEE, pp. 1–4.

Anantharaman, R., Velazquez, M. and Lee, Y. (2018). Utilizing mask r-cnn for detection and segmentation of oral diseases, *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, pp. 2197–2204.

Botterill, T., Lotz, T., Kashif, A. and Chase, J. G. (2014). Reconstructing 3-d skin surface motion for the diet breast cancer screening system, *IEEE transactions on medical imaging* **33**(5): 1109–1118.

Chaudhary, A. and Singh, S. S. (2012). Lung cancer detection on ct images by using image processing, pp. 142–146.

Furuzuki, M., Lu, H., Kim, H., Hirano, Y., Mabu, S., Tanabe, M. and Kido, S. (2019). A detection method for liver cancer region based on faster r-cnn, *2019 19th International Conference on Control, Automation and Systems (ICCAS)*, IEEE, pp. 808–811.

Giddwani, B., Tekchandani, H. and Verma, S. (2020). Deep dilated v-net for 3d volume segmentation of pancreas in ct images, *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)*, IEEE, pp. 591–596.

Jacob, C. and Gopakumar, C. (2020). Pulmonary nodule detection techniques in ct images: New strategies and challenges, *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, IEEE, pp. 1279–1283.

Jamil, S., Abbas, M. S., Umair, M., Habib, F., Hussain, Y. et al. (2020). A novel deep neural network cannet for malignant detection, *2020 International Conference on Information Science and Communication Technology (ICISCT)*, IEEE, pp. 1–5.

Katzmann, A., Muehlberg, A., Suehling, M., Nörenberg, D., Holch, J. W. and Gross, H.-M. (2020). Deep random forests for small sample size prediction with medical imaging data, *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, IEEE, pp. 1543–1547.

Kido, S., Hirano, Y. and Hashimoto, N. (2018). Detection and classification of lung abnormalities by use of convolutional neural network (cnn) and regions with cnn features (r-cnn), *2018 International Workshop on Advanced Image Technology (IWAIT)*, IEEE, pp. 1–4.

Kopelowitz, E. and Engelhard, G. (2019). Lung nodules detection and segmentation using 3d mask-rcnn, *arXiv preprint arXiv:1907.07676* .

Kumar, M. N., Jatti, A. and Narayanappa, C. (2019). Probable region identification and segmentation in breast cancer using the dl-cnn, *2019 International Conference on Smart Systems and Inventive Technology (ICSSIT)*, IEEE, pp. 1144–1149.

Li, G., Zhou, W., Chen, W., Sun, F., Fu, Y., Gong, F. and Zhang, H. (2020). Study on the detection of pulmonary nodules in ct images based on deep learning, *IEEE Access* **8**: 67300–67309.

Liu, M., Dong, J., Dong, X., Yu, H. and Qi, L. (2018). Segmentation of lung nodule in ct images based on mask r-cnn, *2018 9th International Conference on Awareness Science and Technology (iCAST)*, IEEE, pp. 1–6.

Mohanapriya, N., Kalaavathi, B. and senthil Kuamr, T. (2019). Lung tumor classification and detection from ct scan images using deep convolutional neural networks (dcnn), *2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*, IEEE, pp. 800–805.

Qadir, H. A., Shin, Y., Solhusvik, J., Bergsland, J., Aabakken, L. and Balasingham, I. (2019). Polyp detection and segmentation using mask r-cnn: Does a deeper feature extractor cnn always perform better?, *2019 13th International Symposium on Medical Information and Communication Technology (ISMICT)*, IEEE, pp. 1–6.

Roy, K., Chaudhury, S. S., Burman, M., Ganguly, A., Dutta, C., Banik, S. and Banik, R. (2019). A comparative study of lung cancer detection using supervised neural network, *2019 International Conference on Opto-Electronics and Applied Optics (Optronix)*, IEEE, pp. 1–5.

Tang, Y.-B., Yan, K., Tang, Y.-X., Liu, J., Xiao, J. and Summers, R. M. (2019). Uldor: A universal lesion detector for ct scans with pseudo masks and hard negative example mining, *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, IEEE, pp. 833–836.

Yan, K., Wang, X., Lu, L. and Summers, R. M. (2018). Deeplesion: automated mining of large-scale lesion annotations and universal lesion detection with deep learning, *Journal of Medical Imaging* **5**(3): 036501.

Zhang, T., Gao, H., Xing, Y., Chen, Z. and Zhang, L. (2019). Dualres-unet: Limited angle artifact reduction for computed tomography, *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, IEEE, pp. 1–3.

Zheng, S., Guo, J., Cui, X., Veldhuis, R. N., Oudkerk, M. and Van Ooijen, P. M. (2019). Automatic pulmonary nodule detection in ct scans using convolutional neural networks based on maximum intensity projection, *IEEE transactions on medical imaging* **39**(3): 797–805.