3D U-NET SEGMENTATION LAYER EFFECTED ON SPECT IMAGING VOLUME TARGETED

Mohd Akmal Masud [1,*], Mohd Zamani Ngali [2], Siti Amira Othman [1]
[1] Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor.
[2] Faculty of Mechanical and Manufacturing, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor.
[*] Corresponding author: mohdakmalmasud@gmail.com

ABSTRACT

Functional imaging, particularly SPECT Iodine-131 ablation imaging, has gained recognition as a useful clinical tool for diagnosing, treating, assessing as well as avoiding a variety of disorders, which includes metastasis. Nonetheless, SPECT imaging is conspicuously characterized by low resolution, high sensitivity, limited specificity, and a low signal-to-noise ratio. This is caused by the imaging data's visually similar characteristics of lesions amongst diseases. Concentrating on the automated diagnosis of diseases with SPECT Iodine-131 ablation imaging, in this work, three types of segmentation layers are used. This comprises a pixel classification layer, dice classification layer, and focal loss layer that will be tested to determine which segmentation layer is high for auto-segmentation lesions on SPECT Iodine-131 ablation imaging. The data pre-processing, which mostly entails data augmentation, is initially carried out to address the issue of small SPECT image sample sizes by using the geometric transformation operation. Deep Designer Network App was used to develop a 3D U-Net Convolutional Neural Network (CNN). The dice classification layer shows the highest accuracy for the thyroid uptake data set, which is 42.34, 0.7333, and 0.5789 for RMSD, DSC, and IoU, respectively. There is significance in using the dice classification layer in data sets that have various forms of ground truth labeling. On the other hand, the pixel classification layer is promising and workable for the multi-disease, multi-lesion classification task of SPECT Iodine-131 ablation imaging with a huge data set training.

KEYWORD
SPECT, automated diagnosis, image classification, deep learning, CNN

INTRODUCTION

Deep learning algorithms have become increasingly popular in the medical imaging field since 2018 until now [1]. However, some things need to be taken seriously to implement deep learning, namely, the amount of data that needs to be trained. For example, for organ segmentation, the U-Net CNN plateaued at 160 cases [2], while 1000 slice images are required to obtain more than 75% accuracy and above for lesion SPECT imaging [3]. In lesion segmentation, especially in SPECT images, volume calculation needs to be discussed in most of the U-Net models utilized; average only evaluates dice similarity coefficient, accuracy, and precision [4].

Another thing is that the type of segmentation layer used for lesion segmentation in the image at the end of the layer in the U-net model is scarcely discussed. Hence, this paper will focus on lesion volume calculation using the 3D U-Net model for SPECT images using Iodine-131 as a radiotracer. In addition, three types of segmentation layers at the end of the 3D U-net model layer, namely the pixel classification layer, dice classification layer, and focal loss classification layer, will be employed as a comparative study.

For training data, the Whole Body Ablation Iodine-131 patient image will be scanned using the Philips BrightView SPECT modality. In addition, the National Electrical Manufacturers Association (NEMA) phantom image with the sphere volume will be used to compare the three-layer segmentation layer. Subsequently, Matlab 2022a will be used as a tool for the ground truth

labeling process using segmenter apps. Meanwhile, deep network designer apps will be used to develop a 3D U-Net network with 40 layers and 41 connections.

MATERIAL AND METHODOLOGY

Whole Body Ablation Iodine-131 SPECT Image

Whole Body Ablation Iodine-131 SPECT images were gathered in diagnosing differentiated thyroid carcinoma using Philips BrightViewXCT imaging equipment at National Cancer Institute, Putrajaya. When performing a SPECT test, the equipment recorded the patient's intravenous administration of the radiotracer Iodine-131. Clinical metastases were found in 131 patients, ranging in age from 24 to 67. A Whole Body Ablation Iodine-131 SPECT image may display the majority of a patient's body due to its 130 (width) x 130 (height) x 90 (slices) size. Additionally, a radiation dosage matrix was expressed by a 16-bit unsigned integer, and every bone SPECT image was kept in a DICOM file (.dcm).

SPECT Image Labelling

Image labeling is crucial in training a dependable deep learning-based segmentation model in the supervised learning domain. Nonetheless, it consumes more time and is arduous to classify a SPECT image, given its low spatial resolution. Therefore, image segmenter apps in Matlab 2022a are used to label the background and lesion in the head-to-neck area [5]. Labeled lesions are on the uptake of the thyroid gland, parotid gland, submandibular and sublingual gland. The diagnostic report in text format and the DICOM file of a head-to-neck SPECT image was previously imported into volume segmenter apps, as shown in Figure 1. Other than that, lesion and background labeling are by default dicominfo for brightness, contrast, and threshold using the smallest paintbrush size in the toolbar. All of the SPECT image manual annotation findings are utilized as experiment ground truth (as a logical.mat file) to create a combined annotation file that is input into the segmentation models.
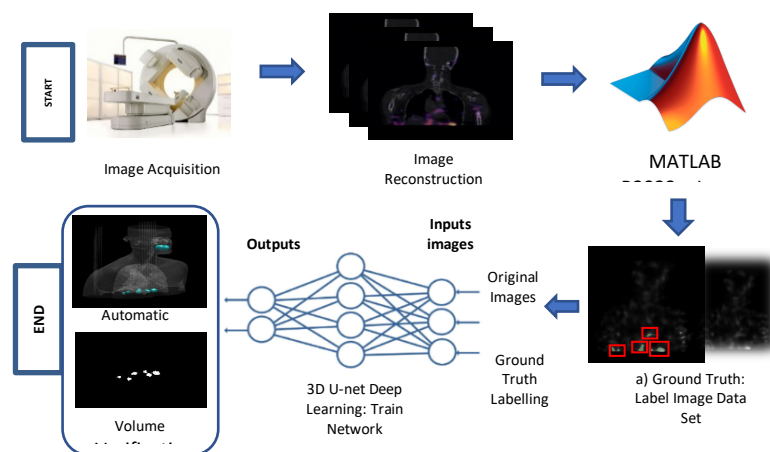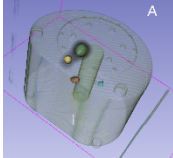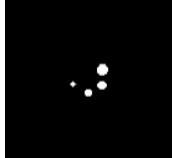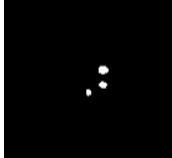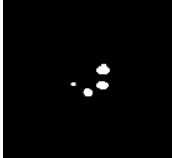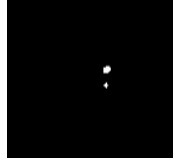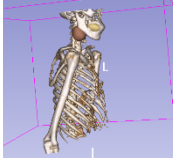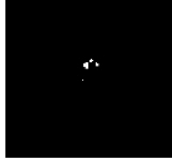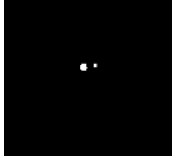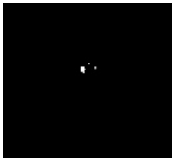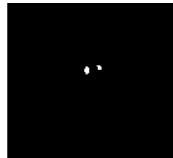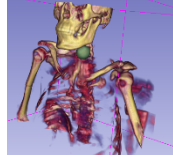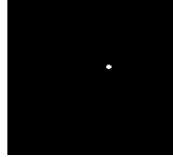

Figure 1: Workflow for image labeling and training.

RESULT AND DISCUSSION

Three data sets were used in this study, namely; 1) NEMA phantom with six sphere insert, 2) set 1 Gland: Thyroid, Parotid, Submandibular, Sublingual, 3) Set 2 Gland: Thyroid only. Note that these sets were tested using three different segmentation layers: pixel classification layer, dice classification layer, and focal loss layer.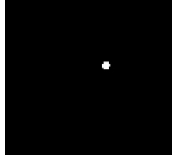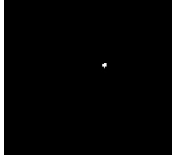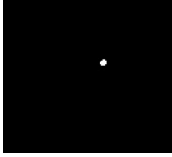 As in Table 1, column 2 is a fusing image between a SPECT image and a CT image. Column 3 shows a picture of the ground truth labeling for the training data set. Finally, Columns 4, 5, and 6 are the auto-segmentation images for three differ of segmentation layers. For image patients, labeling lesions are only taken from the head to the thorax area. This is to test the segmentation of the thyroid gland, parotid gland, and the submandibular and

sublingual gland in a training data set to determine which segmentation layer has the highest accuracy.

Table 1: Data set for groundtruth image labeling.

Data Set	Original Images	Ground Truth Labelling	Pixel Classification Layer	Dice Classification Layer	Focal Loss Layer
NEMA Phantom					
Set 1 Gland: Thyroid, Parotid, Submandibular, Sublingual					
Set 2 Gland: Thyroid ONLY					

Based on Table 1, the NEMA phantom with six spheres inside has been evaluated through the use of Root Mean Square Deviation (RMSD), the Intersection of Union (IoU) as well as Dice Similarity Coefficient (DSC) using three different types of segmentation layers. The results show that the RMSD for the dice classification layer is the highest accuracy (23.98%) compared to the pixel classification layer (27.16%) and the focal loss layer (42.54%). However, for dice similarity, it shows that the pixel classification layer is the highest for DSC and IoU, with 0.8035 and 0.6716, respectively.

For Table 2, only patients with uptake in the thyroid gland are labeled and trained. This is to see if there is significant use of the dice classification layer for RMSD, DSC, and IoU. The results present that RMSD, DSC, and IoU recorded the highest accuracy for the dice classification layer, which are 42.34%, 0.7333, and 0.5789.

Table 2: Comparison of experimental results of three segmentation layers for set 2 gland data set.

Set 2 Gland: Thyroid ONLY					
Evaluation/Segmentation Layer	No. of Patient	Volume Actual (ml)	Volume Calculation (ml)	Root Mean Square Deviation (RMSD)	Dice Similarity (DSC)	Intersection of Union (IoU)
Pixel Classification	1	9.54	1.5	84.27	0.2727	0.1579
Dice Classification	1	9.54	5.5	42.34	0.7333	0.5789
Focal Loss	1	9.54	4.1	57.02	0.6029	0.4316

However, the values for the pixel classification layer for RMSD, DSC, and IoU are much lower compared to the dice classification layer, which are 84.27, 0.2727, and 0.1579. In addition, the focal loss layer shows 57.02%, 0.6029, and 0.4316 for RMSD, DSC, and IoU, respectively. The dice classification layer displays the highest accuracy because all the labeled thyroid gland shapes seem to have the same shape when the labeling data is created. Thus, when the labeling data is trained, the data accuracy reaches its maximum level at a low iteration, which is at the 20th iteration.

Meanwhile, the pixel classification layer produces the lowest accuracy due to inconsistency in the labeled pixel value. This means that if any labeled data has a high difference between one data and another, the efficiency of data learning is relatively low compared to training data with many labeled pixels and consistency.

CONCLUSION

This study has found that the different types of segmentation layers will provide different segmentation volumes for the same data on 3D-Unet network. In this paper, the dice classification layer displays the highest accuracy because all the labeled thyroid gland shapes seem to have the same shape when the labeling data is created. Meanwhile, the pixel classification layer produces the lowest accuracy due to inconsistency in the labeled pixel value. This means that if any labeled data has a high difference between one data and another, the efficiency of data learning is relatively low compared to training data with many labeled pixels and consistency.

REFERENCE

Namdev, Utkarsh, Shikha Agrawal, and Rajeev Pandey (2022). Object detection techniques based on deep learning: A review. Computer Science and Engineering: An International Journal 12, no. 1: 1-21. https://doi.org/10.5121/cseij.2022.12113.

Bardis, Michelle, Roozbeh Houshyar, Chanon Chantaduly, Alexander Ushinsky, Justin Glavis-Bloom, Madeleine Shaver, Daniel Chow, Edward Uchio, and Peter Chang (2020). Deep learning with limited data: Organ segmentation performance by U-net. Electronics (Switzerland) 9, no. 8: 1199. https://doi.org/10.3390/electronics9081199.

Lin, Qiang, Chuangui Cao, Tongtong Li, Zhengxing Man, Yongchun Cao, and Haijun Wang (2021). dSPIC: a deep SPECT image classification network for automated multi-disease, multi-lesion diagnosis. BMC Medical Imaging 21, no. 1: 122. https://doi.org/10.1186/s12880-021-00653-w.

Zhang, Jingyi, Mengge Huang, Tao Deng, Yongchun Cao, and Qiang Lin (2021). Bone metastasis segmentation based on Improved U-NET algorithm. Journal of Physics: Conference Series, 1848, no. 1: 012027. https://doi.org/10.1088/1742-6596/1848/1/012027.

Palladino, L., B. Maris, and P. Fiorini (2020). 3D Slicer module for semantic segmentation of ultrasound images in prostate biopsy using deep learning techniques. CARS Computer Assisted Radiology and Surgery, 34th International Congress and Exhibition.