IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Segmentation of the Carotid Artery Using Deep Learning U-Net technique

[1]Anitha M, [2]Vijayalaxmi Inamdar, [3]Chethana T S, [4]Deepa N C,[5]Gurijala Pranathi

[1]Assistant Professor, [2]Assistant Professor,[3]Student,[4]Student,[5]Student

Department of Computer Science and Engineering

Dayananda Sagar Academy of Technology and Management, Bengaluru, Karnataka, India

[1]Dayananda Sagar College of Engineering, Bengaluru, Karnataka, India

ABSTRACT: Deep learning and image segmentation techniques are widely used in various sectors nowadays but, these methods play a predominant role in diagnosis of various health conditions. Automated and detailed medical image segmentation model have gained popularity since the advent of deep learning techniques specifically fully convolutional neural networks (FCNN). U-Net is one such fully convolutional neural network (FCNN) based image segmentation model, which has proven its efficiency in medical image segmentation over recent years. In this paper, we present the traditional U-Net model in comparison with a novel U-Net model with an additional dropout layer in each convolutional layer. Here, we compare these models considering the dice average values obtained for two different data inputs containing ultrasound images of a carotid artery in various patients. The proposed U-Net model with dropout layer is observed to have better dice average values in comparison with the traditional U-Net model.

Keywords: Medical Image Segmentation, FCNN, Deep learning, U-Net.

I. INTRODUCTION

Carotid arteries are the most vital and sensitive arteries in a human body. These Carotid arteries are predominant arteries in a human as they carry out a prime duty to supply oxygenated blood to the parts of neck, face and brain from the heart. As they control the blood supply to the head, these arteries are located on either sides of the neck and are made up of glomus cells.

These arteries contain a lumen, which is the inner region of the blood vessels, of an average diameter ranging from 4.4mm to 7.8mm. Fatty deposits also referred to as plaques, in the lumen of a carotid artery to a greater extent can cause a clog with in the blood vessel and hence resulting in disturbance in the normal flow of the blood with in the artery. This can cause serious strokes which occurs due to the loss of blood supplied to the brain. These strokes can also occur due to the decrease in the diameter of the carotid artery. And this condition is called carotid artery stenosis.

Image processing methods are best used for obtaining an accurate and automated segmentations of these carotid arteries and hence obtain a precise border of the wall of a carotid artery. This can be used to determine the changes in the diameter of the blood vessel in a patient and therefore the doctor can easily analyze and predict the occurrence of strokes and prescribe the necessary care to be taken by the patient.

Image segmentation, specifically sematic segmentation using deep learning techniques, has proved to be the best approach for clinical diagnosis in the present world. U-net otherwise, the U-shaped neural network is one such model based on Fully convolutional neural network. A convolutional neural network is an artificial neural network that contains different layers, to generalize it, an input layer which takes the input image, a hidden layers where the feature maps, maximum pooling and activation functions are deployed to perform feature extraction, and an output layer, where all the extracted features are combined to perform semantic segmentation which produces an output image which highlights the specific area of interest within the input image.

II. METHODOLOGY

2.1 Data

There are two training datasets. The first dataset(dataset 1) is comprised of 283 images of carotid artery. These images are cross-section of common carotid artery of the human body. The images are in PNG format with size 430 x 430. The ground truth values of the carotid artery are in the form of length and width of ellipse. The position of center is also given.

The second dataset(dataset 2) consists of 1100 images of cross-section of common carotid artery. These images are taken form 11 patients in different angles. They were taken in Mindary UMT-500Plus ultrasound machine with an L13-3s linear probe. These images are in the form of aquistified ultrasound images in PNG format. Each patient has 100 images which are converted from time series format to PNG format. The ground truth masks are also enclosed in the dataset.

2.2 Data preprocessing

The images are resized to 128 x 128 to train to model. The resized images are normalized by subtracting the pixel value with the minimum pixel value and then dividing it by the difference between the maximum and minimum pixel values. Normalization reduces the intensity of pixel values which are more than the range and increases the intensity of pixel values which are lesser than the range. It helps in optimal comparisons of the pixel values and texture of the image. It also makes computation easier by reducing the range of the pixel values.

The mask values in dataset 1 are converted to annotations on the images. These annotations are then used to obtain masks of the carotid artery. The masks are in the form of an ellipse. These masks can be used to train the model. The dataset 2 is provided with masks based on the boundary and location of common carotid artery.

2.3 Model architecture

We present with two model architectures which are suitable for each of the datasets. These architectures are comprised of different number of layers and blocks. Below is the description of each model.

2.3.1 U-Net

The u-net architecture consists of repeating blocks of encoder and decoder. The encoder is left part of symmetric u-net which consists of 3 blocks. Encoder decreases the spatial dimensions of the image and increases number of channels with each block. This reduces the computation cost of the model. Each block in the encoder consists of convolutional layer with kernel size of 3 x 3 with 2 x 2 stride. It is followed by a ReLu layer with kernel size of 2 x 2 and another convolutional layer. The use of ReLu layer in each block increases the speed of the model.

Encoder has max polling layer which decrease the spatial dimensions of the image after each block. This layer also provides basic translational invariance to the model. The last part of encoder is connected to a block which is also called the bottleneck layer is used to learn the compression of input data. This layer contains same layer as the block. The diagram below shows the UNet architecture.

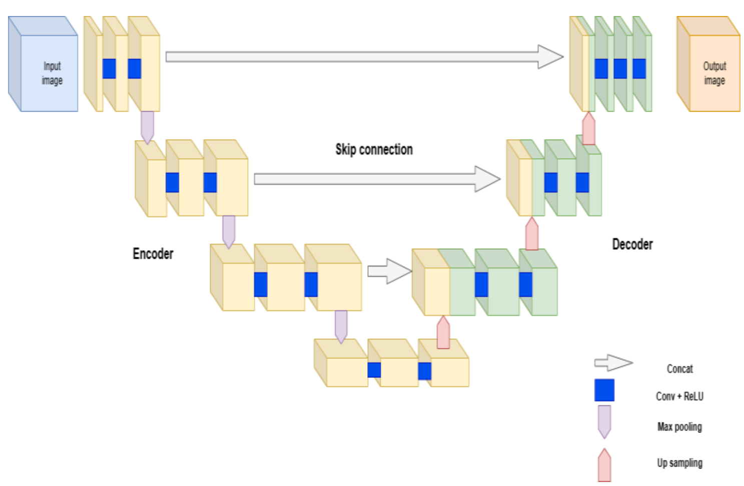

Fig.1. U-Net architecture

The decoder contains repetitive up sampling blocks which increase the spatial dimension of the image. There are skip connections between the encoder and the decoder. They are used to concatenate data from encoder with decoder. The decoder block implements the convolutional transpose layer. Transpose convolutional layer is reverse of convolutional layer where input slides over the kernel and elements wise multiplication and summation takes place. The decoder then crops the output from encoder which was stored earlier and concatenates with the up-sampling output. The output obtained from concatenation is then passed through convolution plus ReLU layer. The segmentation mask is obtained by passing the output through regression head.

3.3.2 Dropout u-net

The Dropout u-net is a u-net with dropout layer in each block. The encoder starts with two convolutional layers and ReLu layer which is then passed through dropout layer. The dropout layer helps in preventing neurons from converging to the same goal. It helps in regularization of the model and overfitting of data.

There is dropout layer at the end of the bottleneck layer since the bottleneck layer is similar to the encoder blocks. The decoder begins with up sampling layer which increases the dimensions of the input and concatenated with cropped output from encoder layer. This prevents loss of information when the network becomes deeper.

The encoder has 4 blocks in this network which increases the efficiency of the model. The decoder also has 4 blocks which is later converted to a segmentation map.

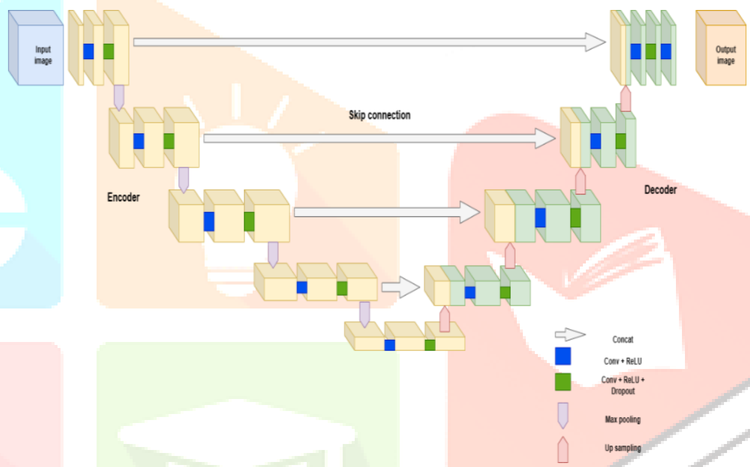

Fig.2. U-Net architecture with dropout

III. RESULTS

4.1 Implementation

Our code is written in pytorch. The model is implemented using layers from pytorch such as conv, convtranspose, maxpooling, ReLu and dropout. The implementation also uses other libraries of python for implementation of the model. The model can be run on any system with proper space in the memory.

4.2 Training schedule

This model is run for maximum of 300 epochs to increase efficiency in time and reduce memory consumption. The results both the models are different for different datasets. The optimizer used for the model is Adam optimizer. The loss function used is Binary cross entropy with logits loss function which is also implemented using pytorch library.

4.3 Experiments

The datasets are experimented on different models. This following table shows the average dice scores of dataset1 which are calculated using intersection of predicted segmentation mask and ground truth.

Table 1.1 Comparison of dice average scores

Epochs	u-net	Dropout u-net
50	72.77	65.28
180	70.33	78.56
300	79.82	87.86

The above results show that the average dice scores are higher for the Dropout UNet architecture with 4 blocks. The images for the best dice values obtained are shown below.

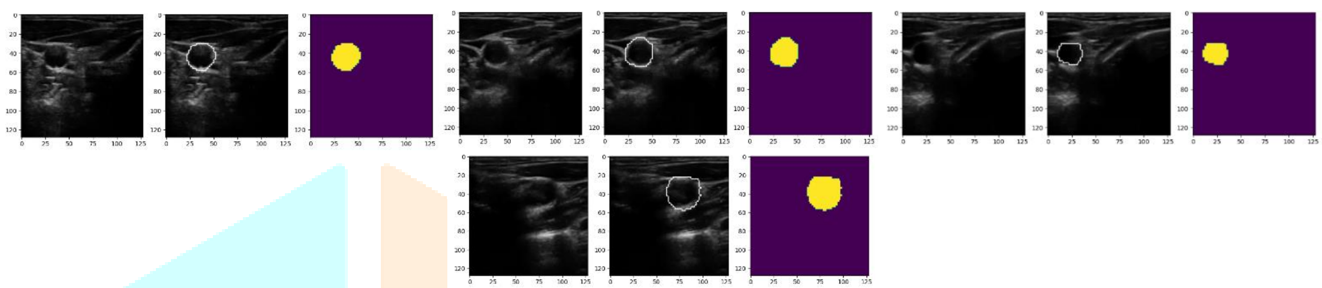

Fig. 3. Outputs for the best dice average scores

The first image is the original ultrasound image, second is the annotation of predicted mask on the image and the third is the segmentation mask obtained by the model. The above images show the segmentation of the carotid artery for different types of images. The model is trained with encoder levels of 3, 16, 32 and 64. The decoder levels are 64, 32 and 16. The training loss is represented as a graph for 300 and 180 epochs for dataset 1 in the below diagrams.

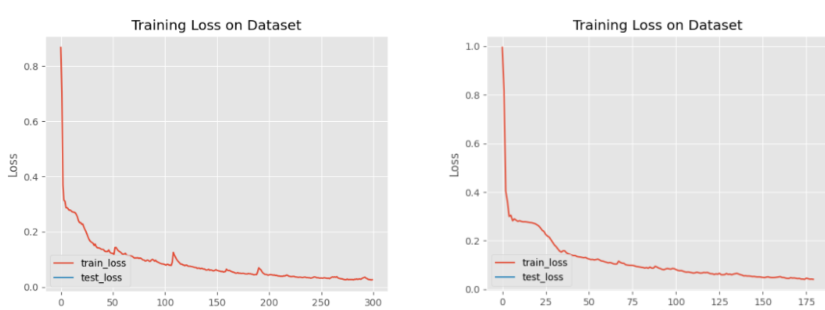

Fig. 4. Graphical representation of training loss for 300 and 180 epochs, dataset 1

These experiments are also done on dataset 2 which contains 1100 images and has a refined ground truth. The dataset 2 has a greater number of training images compared to dataset 1. The below table shows the average dice scores for first model for dataset 2.

Table 1.2 Dice average scores for first model and dataset 2

Epochs	u-net
50	57.35
150	72.62
200	68.07

The above results were comparatively lesser than the dice scores for the dataset 1. The model was trained on different epochs on the dropout model based on the loss values recorded during training data. The results for the dropout model are shown below.

Table 1.3 Dice average scores for second model and dataset 2

Epochs	Image size	Dropout u-net
50	128	59.04
62	256	60.97
200	128	68.07
250	256	74.06
300	128	80.84

The segmentation images obtained by the dropout model for dataset 2 are shown below.

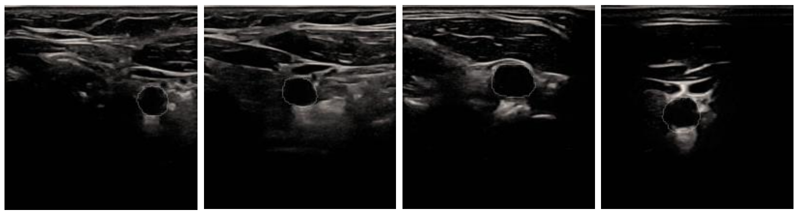

Fig. 5. Outputs obtained by dropout model for dataset 2

IV. CONCLUSION

The u-net architecture gives a good segmentation mask for bio-medical ultrasound images. The model is trained for 5 hours which reduces the training time due to lesser number of blocks in encoder and decoder. The u-net architecture with dropout layer yields better results compared to the previous u-net architecture. This model can be used in many other tasks due to its time efficiency and accurate segmentation.

REFERENCES

[1] Ronneberger, Olaf & Fischer, Philipp & Brox, Thomas. (2015). UNet: Convolutional Networks for Biomedical Image Segmentation. LNCS. 9351. 234-241. 10.1007/978-3-319-24574-4_28.

[2] Zhou, Zongwei & Rahman Siddiquee, Md Mahfuzur & Tajbakhsh, Nima & Liang, Jianming. (2018). UNet++: A Nested Unet Architecture for Medical Image Segmentation: 4th International Workshop, DLMIA 2018, and 8th International Workshop, MLCDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings. 10.1007/978-3-030-00889-5_1.

[3] Huang, Huimin & Lin, Lanfen & Tong, Ruofeng & Hu, Hongjie & Qiaowei, Zhang & Iwamoto, Yutaro & Han, Xian-Hua & Chen, Yen-Wei & Wu, Jian. (2020). UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation.

[4] Cao, Hu & Wang, Yueyue & Chen, Joy & Jiang, Dongsheng & Zhang, Xiaopeng & Tian, Qi & Wang, Manning. (2021). SwinUnet: Unet-like Pure Transformer for Medical Image Segmentation.

[5] ZITAO ZENG , WEIHAO XIE, YUNZHE ZHANG, AND YAO LU (2019).RIC-Unet: An Improved Neural Network Based on Unet for Nuclei Segmentation in Histology Images

[6] Pricing, Jinke Wang, Xiangyang Zhang, Chunlei Ji, Lubiao Zhou, Haiying Wang. An improved residual U-Net with morphological-based loss function for automatic liver segmentation in computed tomography

[7] Yongjin Zhou, Member, IEEE, Weijian Huang, Pei Dong, Yong Xia, and Shanshan Wang, Member, IEEE.D-UNet: a dimension-fusion U shape network for chronic stroke lesion segmentation

[8] Steven Guan, Amir A. Khan, Siddhartha Sikdar, and Parag V. Chitnis. Fully Dense UNet for 2D Sparse Photoacoustic Tomography Artifact Removal.

[9] Qingfeng Zhang, Yun Du,Zhiqiang Wei, Hengping Liu, Xiaoxia Yang, and Dongfang Zhao. Spine Medical Image SegmentationBased on Deep Learning

[10] Guo, Changlu & Szemenyei, Márton & Pei, Yang & Yi, Yugen & Zhou, Wei. (2019). SD-Unet: A Structured Dropout U-Net for Retinal Vessel Segmentation. 439-444. 10.1109/BIBE.2019.00085.