

Segmentation of Liver in CT images using 3D-Res-UNet

Lakshya Gupta

Bachelor of Technology in Computer Science

Atal Bihari Vajpayee Indian Institute of Information Technology and Management

Gwalior, India

lakshyag9775@gmail.com

Abstract—Medical segmentation is a method for locating the borders of internal organs on images taken using the two modern imaging modalities that are most frequently used: MRI and 3D computer tomography. The U-net illustrates a deep neural network architecture with the ability to segment medical pictures. Its basic structure is made up of two components: encoders and decoders. To construct the Modified U-Net architecture that will be utilised in this project, we used the free source Monai Framework.

Index Terms—Deep learning, Convolutional neural network, UNET, RES-UNET, MONAI, Liver segmentation, Residual learning.

I. INTRODUCTION

With the aid of Monai, a PyTorch framework, we were able to complete the 3-D segmentation of liver volumes for this project. ResUNet is a CNN architecture that has been updated. A method for locating the borders of interior human organs in 3D CT scans is called medical segmentation. [12]. Image

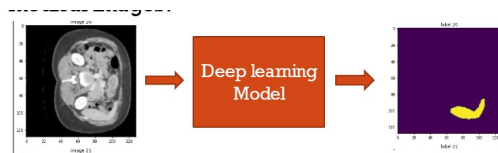


Fig. 1. Basic object of segmentation model

classification, one of the first deep learning applications, is the most frequent deep learning problem in computer vision. When classifying an image, all we do is determine if the object we're seeking for is visible in the image or not. Image detection follows image classification and involves enclosing the item we're looking for in the image with a box. The next step in classification and detection is segmentation, and this approach not only finds an item inside an image but also accurately separates the pixels that make up the object. Instance segmentation and semantic segmentation are the two categories of segmentation. In our project, we are using semantic segmentation, since just have to find whether the slice is having liver or not. We'll utilise the U-Net model for segmentation, which is based on the open-source MONAI framework, which is built on Pytorch. With only a little

quantity of training data, the medical image analysis tool U-Net can segment pictures. All imaging modalities, including CT and MRI, can be segmented. Before training the data we are going through some preprocessing, normalisation and visualisation of data, visualisation and normalisation is done by software named 3D slicer.

II. RESEARCH WORK FLOW

It is crucial to leverage low-level details in **semantic segmentation** while keeping high-level semantic information in order to provide a finer outcome. It is crucial to leverage low-level details in semantic segmentation while keeping high-level semantic information in order to provide a finer outcome [11]. For the data related to medical field and segmentation process mostly available in the form of 3d volumes mainly in CT format or MRI images. In this paper we are using a general and effective method which can be used for segmentation of various parts of body and can be directly applied to Nifti formate of data means directly on the CT scans [12]. In old methods we generally convert these Nifti files into various Dicoms slices to train the model, Which also works fine, but takes extra bit of time and work. We are usign Monai framework which is a open-source framework of Pytorch [2]. Using this framework we are preprocessing our data and and taining our 3D UNet model which will use ReLU as activation function and also add the dropout layers. This activation fuction adaptively learns the paremeters of recti-fires and improves the accuracy at minimum cost [3]. UNet model is specially designed for medical segmentaion with good accuracy, but with some changes we use the same model but after adding residual layer and some minor changes .We will be using Residual Unet model for the segmentation task [16] which is giving good results [10]. We demonstrate enhanced performance over several of the existing models on a broad range of benchmark tasks like segmentation of different organs of a body.

III. METHODOLOGY

A. Proposed hypothesis

The 3D-ResUNet architecture of Monai was specifically created to address segmentation issues in the medical area by enhancing previous research to get the best outcomes. In

our project, we partition liver CT images using this approach. The 3D-RES-UNET model is an improvement on the current UNET models, which just include CNN lanes and no activation function. However, we include a residual layer in our proposed model that addresses a number of issues with the prior UNET, such as vanishing gradient, overfitting, and the curse of dimensionality. As the model gets deeper, it aids in the creation of skip connections, which is the primary function of residual learning.

B. Dataset description

The dataset we use is the LiTS dataset [6] from the Liver Tumor Segmentation Challenge from MICCAI 2017. It contains ct scans of 131 patients and the mask of liver and lesions. In this project we were doing liver segmentation and due to computation limitations we use subset of this dataset. we take 25 patients CT scans for training purpose and 10 CT scans for testing purpose.

C. Procedure followed

Our LiST dataset includes the segmented Nifti files (CT scans) of 130 individuals. We use a tiny piece of the dataset to train our model because the entire dataset would be too large to train. After that, we separate it into training data and testing data. These Nifti files or CT scans are non-uniform because each file contains a different number of slices. To visualise the data and determine which Nifti file contains how many Dicom slices, we converted the Nifti files into Dicom files and then grouped them into Nifti files of a uniform number of slices (we chose 65 no. of slices per nifti file). Then we remove the empty files which does not contain any foreground(mask) using the `f_data` of CT scans. After these data cleaning and visualisation steps, we apply several transforms to our Nifti files, including AddChannel, Spacing, ScaleIntensityRange, CropForegroundd, Resize, and others that sharpen and clarify our image for training by adjusting the contrast from that dense vision into something more visible. The second step is to normalise the voxel values and place them between 0 and 1 so that the training will be faster, and crop down the empty region.

D. Architecture of model

UNet: CNN is only used to classification problem and only returns the label as the result but in case of biomedical field it is not only stop to classification but also we have to return the mask of interest part of that organ after classify it correctly. So this problem is successfully solved by UNet architecture which consist of two parts majorly to paths:

- i) Encoder/Contracting path
- ii) Decoder / expansive path.

Contracting path is simply uses the convolutional process and the decoder path is uses transposed convolutional layers(you can think it as an upsampling technique for now).

Residual Unet: Going deeper would improve the performance of a multi-layer neural network, however could hamper

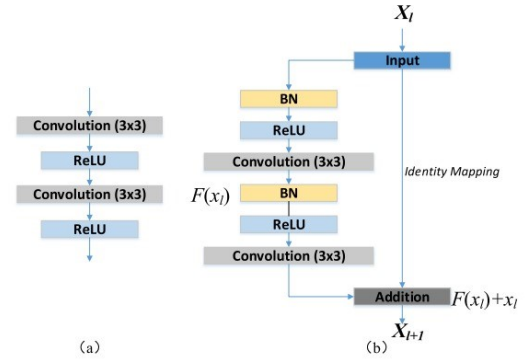


Fig. 2. (a)Neural unit used in U-Net (b) Residual block with after adding some layers in UNET model. [16]

the training, and a degradation problem maybe occur. To overcome these problems, we proposed the residual neural network to facilitate training and address the degradation problem. The residual neural network consists of a series of stacked residual units. Each residual unit can be illustrated as a general form: $y_l = h(x_l) + F(x_l, W_l)$, $x_{l+1} = f(y_l)$, where x_l and x_{l+1} are the input and output of the l -th residual unit, $F(\cdot)$ is the residual function, $f(y_l)$ is activation function and $h(x_l)$ is a identity mapping function, a typical one is $h(x_l) = x_l$. Fig. 4.1 shows the difference between a plain and residual unit [16].

Deep ResUnet, a semantic segmentation neural network, which combines the strengths of both U-Net and residual neural network. This combination brings us two benefits: 1) the residual unit will ease training of the network; 2) the skip connections within a residual unit and between low levels and high levels of the network will facilitate information propagation without degradation, making it possible to design a neural network with much fewer parameters, however, could achieve comparable ever better performance on semantic segmentation.

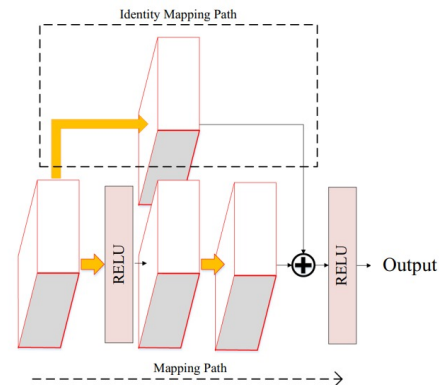


Fig. 3. Rssidual block [15]

We train this model with ADAM optimizer and Dice Coefficient as loss function

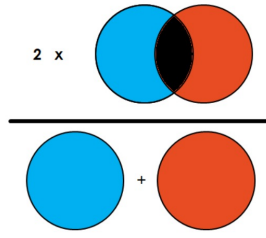


Fig. 4. Dice Coefficient as loss function. [1]

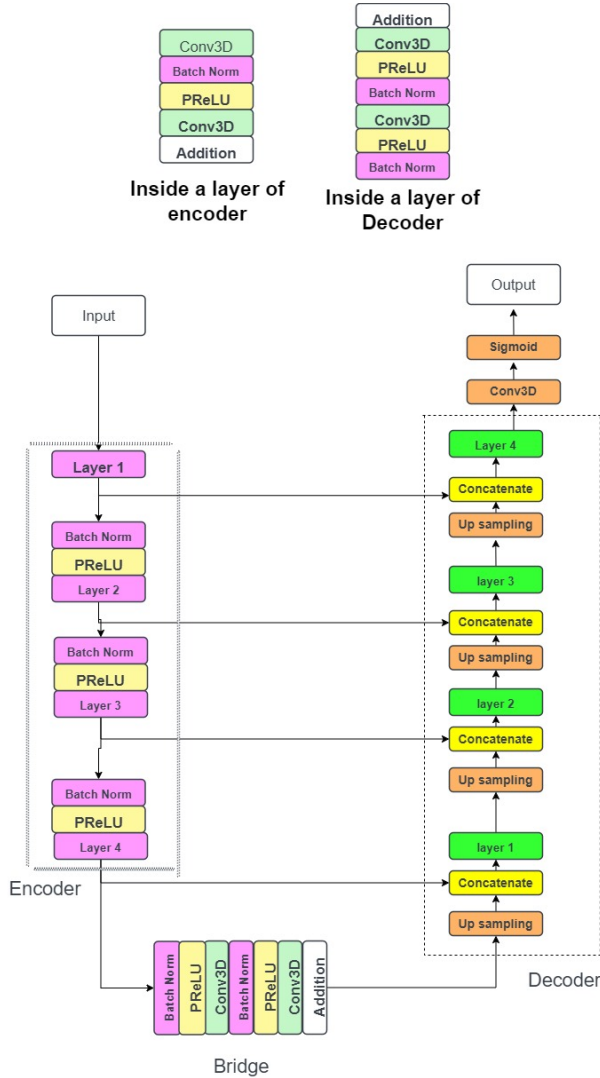


Fig. 5. Proposed Architecture of 3D-ResUNet model used in this project

IV. EXPERIMENTS AND RESULTS

To demonstrate the accuracy and efficiency of proposed 3D-RES-UNET model we test it on LITS dataset [6] and compare our results to the previously present methods like **Graph cut**

segmentation.. In graph cut method we take every pixels as the the vitices and the weigths are represented using edges.

A. Implemetation details

This proposed model is implemented using a pytorch's open source framework MONAI [2]. The data normalisation and visulization is done using 3rd party software named 3D slicer. After this step we do some data transformation using MOANI transforms, before giving CT scan to the model for training.

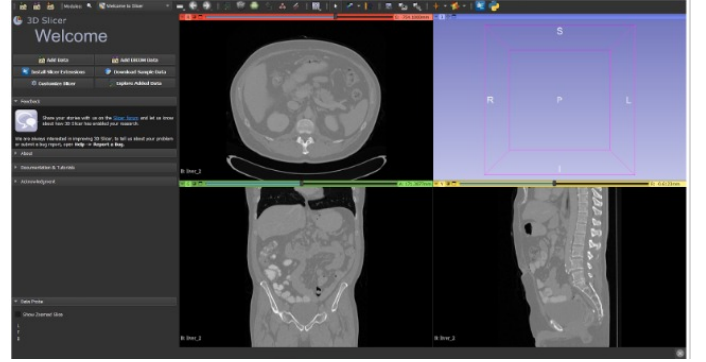


Fig. 6. visualization and Normalisation the CT scans(Nifti) of patients

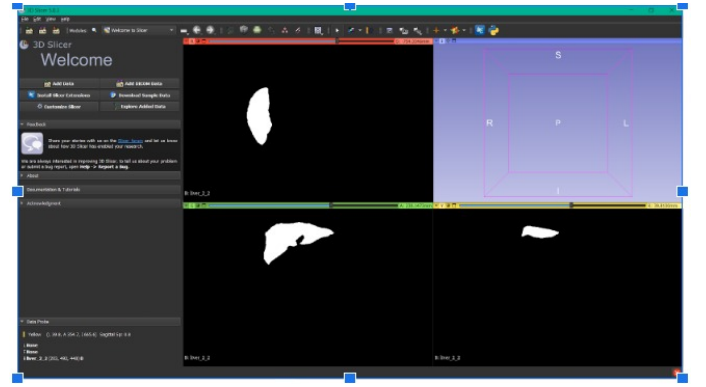


Fig. 7. visualization and Normalisation the mask of patients

B. Experiment 1

We have CT scans to train our model on, so our input CT scans having dimensions of 512 X 512 and also the output mask as the same dimensions as the property of UNET tells us.

Parameters and architecture of our model:

- dimensions=3,
- in_channels=1,
- out_channels=2,
- channels=(16, 32, 64, 128, 256),
- strides=(2, 2, 2, 2),
- residual units = 2,
- norm=batch normalisation
- activation function = PReLU(with parameter 1)
- drop-out layer with probability of 0.4

we are Using adam optimizer with dice coefficient as loss function and learning rate of 0.0001 and used 120 epochs.

Adam optimizer: An algorithm for gradient descent optimization is called adaptive moment estimation. When dealing with complex problems involving a lot of data or factors, the strategy is incredibly effective. It is effective and uses little memory. It combines the "gradient descent with momentum" algorithm and the "RMSP" algorithm, intuitively. Building on the advantages of earlier models, Adam optimizer performs far better than those models and outperforms them in producing an optimised gradient descent.

1) *Results:* This parameter setting and model gives us the mean Dice value of 0.712 and Dice loss = 0.30 after 120 epochs.

Mean Dice value: 0.712

Dice Loss: 0.30

C. Experiment 2

Input image as of same dimensions with some changes in parameters like dropout probability and some hyperparameter tuning and setting our learning rate to 0.00001 this time.

Parameters and architecture of our model:

- dimensions=3,
- in_channels=1,
- out_channels=2,
- channels=(16, 32, 64, 128, 256),
- strides=(2, 2, 2, 2),
- residual units = 2,
- norm=batch normalisation
- activation function = PReLU(with parameter 1)
- drop-out layer with probability of 0.5

we are Using adam optimizer with dice coefficient as loss function and learning rate of 0.0001 and used 216 epochs.

1) *Results:* This parameter setting and model gives us the mean Dice value of 0.8095 and Dice loss = 0.24 after 216 epochs.

Mean Dice value: 0.8095

Dice Loss: 0.24

D. Conclusion and results

We saw the greatest results around epoch 216 when our learning rate was 0.00001, and then overfitting started, according to the experiment data. Using the 3D-Res-Unet model, an improved version of the UNet model, we were able to obtain accurate dice results. By creating skip connections in the residual block, many problems with UNET—such as vanishing gradient and the curse of dimensionality—were alleviated.

As the graph shown in fig 8 telling us about the dice loss and the mean dice of our model in some final epochs during testing. The graph of dice loss is decreasing as the epochs are increasing and reached a minimum level after 216 epochs.

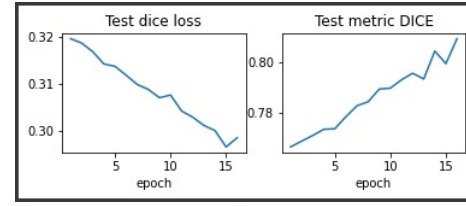


Fig. 8. left: Test dice loss with no of epoch(in final 20 epochs) right: Test metric dice vs epoch

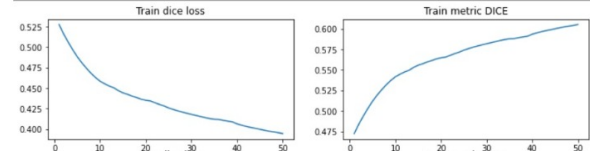


Fig. 9. Train mean dice and train loss vs epoch

fig 9 is showing us the the train mean dice and dice loss vs epochs. We achieved 0.8095 mean dice value after 216 epochs.

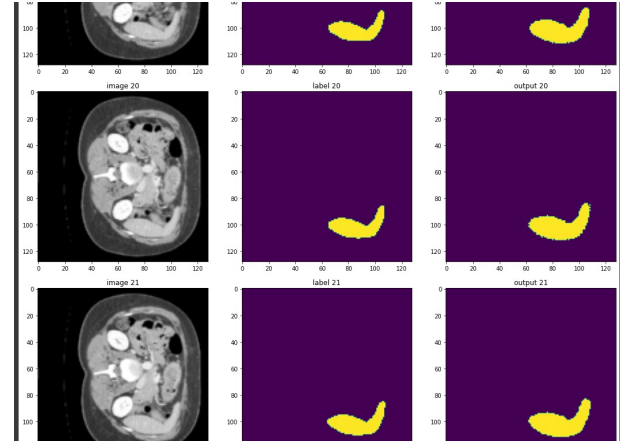


Fig. 10. Comparison between the expected ground truth segmentation (label 1) and the output segmentation of the ResUNet (output1) for the 3D CT Liver dataset

fig 10 showing the segmentation done by our 3D-RES-UNET model with good precision and the results are almost overlapping the ground truth.

TABLE I
COMPARISON OF OUR MODEL WITH OTHER METHOD PRESENT FOR SEGMENTATION

| S.No. | Model | Mean Dice | Dice loss |
|-------|----------------------------|-----------|-----------|
| 1 | Graph cut segmentation [5] | 0.78 | 0.26 |
| 2 | ResUnet (Ours) | 0.8095 | 0.24 |

The data in the table of comparisons show how our model surpasses the outdated method, which had a lower mean dice and a greater loss value than our method. The liver task is segmented with more accuracy and precision by our model. Results may be improved even more if we used the complete dataset instead of training our model on a subset of it due to technical limitations.

V. LIMITATIONS

The algorithm used in this endeavour needs a huge dataset with many more patient CT images in order to train effectively and predict suitable masks to the liver volumes. Due to the relatively little dataset we used, our model does not predict particularly good masks. Additionally, because of restricted computing resources and the lengthy training procedure for this model, we can only train it on a small dataset. In the near future, it would be helpful to incorporate tumour segmentation and categorization as our technique only segments the liver, not tumours.

VI. FUTURE SCOPE

The mask of various human organs may be predicted using this model in a variety of medical segmentation applications. This is crucial since manually segmenting organs takes a lot of time and isn't very accurate, thus automatically segmenting organs, particularly the liver, is very beneficial for the medical field's ongoing study. Because the liver is located in the abdominal region, which also contains several bones and other muscles, segmenting the liver is the most challenging task among all human organ segmentations. Given that our model was trained on a tiny sample of data, it is currently not providing very accurate results. The model can be trained on a big dataset to provide better results, and it can also be used to forecast the tumour mask in different human organs, solving one of the main problems of detecting malignant tumours, particularly in the liver.

VII. ACKNOWLEDGMENT

I am really grateful to Dr. Gaurav Kaushal for giving me the freedom to work on and experiment with new ideas. I'd want to use this opportunity to express my gratitude to him for both his academic and personal guidance. The creation and evolution of the present work was much aided by his insightful counsel, recommendations, sound judgement, and constructive criticism as well as his passion for greatness. He showed interest in my concept and continued to support, build my confidence, and motivate me. By constantly hearing my thoughts, appreciating and enhancing them, and giving me entire flexibility in my project, my mentor never allowed me to feel inexperienced. He was really patient and always answered to all of my queries with a grin. His fervent attention and encouraging attitude are the only reasons why the current effort has advanced thus far.

I'd want to conclude by expressing my thanks to our organisation, whose unwavering support allowed me to focus my attention and energy on this project while maintaining my motivation.

REFERENCES

- [1] Loss functions for semantic segmentation link: <https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2>.
- [2] Monai toolkit, 01 2022, [online] available: <https://www.kitware.com/project-monai-launches-open-source-artificial-intelligence-framework-for-medical-research/>.
- [3] Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, paper link: <https://doi.org/10.48550/arXiv.1502.01852>. 2015.
- [4] Sonali Tulsiram Bodkhe and Sheetal Adarsh Raut. A unique method for 3d mri segmentation. In *2008 International Conference on Computer Science and Information Technology*, pages 118–122, 2008.
- [5] Jung won Cha, M. Mehdi Farhangi, Neal Dunlap, and Amir Amini. Volumetric analysis of respiratory gated whole lung and liver ct data with motion-constrained graph cuts segmentation. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3405–3408, 2017.
- [6] Patrick Christ. Lits – liver tumor segmentation challenge (lits17).
- [7] Zhengquan Chu, Tian Tian, Ruyi Feng, and Lizhe Wang. Sea-land segmentation with res-unet and fully connected crf. In *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pages 3840–3843, 2019.
- [8] Ahmed Harouni, Alexandros Karargyris, Mohammadreza Negahdar, David Beymer, and Tanveer Syeda-Mahmood. Universal multi-modal deep network for classification and segmentation of medical images. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 872–876, 2018.
- [9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 1026–1034, 2015.
- [10] S. Ren K. He, X. Zhang and J. Sun. Deep residual learning for image recognition. In *2008 International Conference on Computer Science and Information Technology*, page 770–778., Jun. 2016.
- [11] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3431–3440, 2015.
- [12] Didi-Liliana Popa, Radu-Teodoru Popa, Lucian-Florentin Barbulescu, Renato-Constantin Ivanescu, and Mihai-Lucian Mocanu. Segmentation of different human organs on 3d computer tomography and magnetic resonance imaging using an open source 3d u-net framework. In *2022 23rd International Carpathian Control Conference (ICCC)*, pages 54–57, 2022.
- [13] Shima Rafiei, Ebrahim Nasr-Esfahani, Kayvan Najarian, Nader Karimi, Shadrokh Samavi, and S.M. Reza Soroushmehr. Liver segmentation in ct images using three dimensional to two dimensional fully convolutional network. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 2067–2071, 2018.
- [14] Xiao Xiao, Shen Lian, Zhiming Luo, and Shaozi Li. Weighted res-unet for high-quality retina vessel segmentation. In *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*, pages 327–331, 2018.
- [15] Zhang L Cao Y Yu X Sun J. Yu H, Li J. Design of lung nodules segmentation and recognition algorithm based on deep learning. In *BMC Bioinformatics.*, 2021.
- [16] Zhengxin Zhang, Qingjie Liu, and Yunhong Wang. Road extraction by deep residual u-net. *IEEE Geoscience and Remote Sensing Letters*, 15(5):749–753, 2018.
- [17] Vasily Zyuzin and Tatiana Chumarnaya. Comparison of unet architectures for segmentation of the left ventricle endocardial border on two-dimensional ultrasound images. In *2019 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)*, pages 110–113, 2019.