



Comparative Study of Various Deep Learning Models For Medical Image Segmentation

P.S Anu Rakhi¹, R.S Rajesh²

¹Research Scholar, Department of Computer Science and Engineering, Manonmaniam Sundaranar University, Tirunelveli, India-627012. anurakhi@msuniv.ac.in

²Professor, Department of Computer Science and Engineering, Manonmaniam Sundaranar University, Tirunelveli, India-627012.

ABSTRACT

Image processing plays a vital role in the detection of medical images segmentation. Accurate detection will help the radiologist to predict the images masses. Using Deep Learning methods the segmentation of medical images can be possible with best results when compared with state o art methods. In this paper we have compared the deep learning methods like a) U-net, b) U-net attention, c) Dense-net, d) Attention Dense U-net for the segmentation of mammography images. Performance metrics like Accuracy, Sensitivity, Specificity, F1 Score, and AUC are measure for the above mentioned models. Attention Dense U-net outperforms the best in accuracy with 78.38%, sensitivity is 77.89%, F1 score with 82.24 , and AUC with 86.05%.

1 Introduction

Medical image segmentation attempts to increase the visibility of anatomical or diseased structural changes in images; it frequently plays a crucial part in computer-aided diagnosis and smart medicine due to the significant improvement in diagnostic efficacy and accuracy. Liver and liver-tumor segmentation, brain and brain-tumor segmentation, optic disc segmentation, cell segmentation, lung segmentation, pulmonary nodules, and cardiac image segmentation are examples of common medical image segmentation tasks [1, 2]. The requirement for automated processing of scanned imaging data is fairly great due to the increased accessibility and use of modern medical imaging techniques like Magnetic Resonance Imaging (MRI) or Computed Tomography (CT) [3]. Physicians now use a manual approach to assess medical images and examine with more image data when there are more slices present, especially with the growing use of high-resolution medical imaging. An automatic pre-segmentation of abnormal characteristics in medical images would be necessary to reduce the time-consuming inspection and evaluation process [4, 5].

Image segmentation remains one of the most challenging subjects in the field of computer vision due to the difficulty of feature representation, despite the fact that many algorithms have been described and they are effective in some situations. Particularly, it is more challenging to extract distinguishing characteristics from medical photographs compared to normal RGB images because the former frequently has issues with blur, noise, low contrast, etc. Convolutional neural networks (CNN) successfully achieve hierarchical feature representation



of images, making them the hottest research topic in image processing and computer vision. Medical image segmentation will no longer require hand-crafted features due to the rapid development of deep learning techniques. CNNs employed for feature learning produce good segmentation results for medical images since they are insensitive to image noise, blur, contrast, etc. [6].

In several visual recognition tasks, deep convolutional networks have outperformed the state of the art. Convolutional networks have been around for a while, but their effectiveness has been constrained by the size of the training sets that were available and the size of the networks that were being taken into consideration. The innovation resulted from the supervised training of a large network with millions of parameters over eight layers on the Image Net dataset with one million training photos. Since then, even more substantial and extensive networks have been trained. Then, larger and more intricate networks were trained [7]. The goal of image semantic segmentation is to classify a picture's pixels. Researchers recommended the fully convolution network (FCN), U-Net, and Deep encoder-decoder architecture, one of the most widely used end-to-end designs, for this purpose. In these structures, an encoder is frequently used to extract image features, while a decoder is frequently used to output the final segmentation findings and restore extracted features to the original picture size. Although the end-to-end form is practical for segmenting medical images, it lessens the models' interpretability [8].

2 Related Works

In order to segregate breast masses from mammograms, Dhungel et al. [9] suggested a novel method that combined structured learning and deep learning. Based on generally acknowledged accuracy and running time measurements on open-source datasets, the author proposed a quantitative analysis that is simple to replicate in order to evaluate the effectiveness of breast mass segmentation approaches. The obtained segmentation accuracy was measured quantitatively using the Dice index, which was 89%.

On digital mammograms, a novel method of diagnosing breast masses was put out by Boulehmi et al. [10]. The next step in the analysis of enhanced mammograms involves computing many primitives related to GGD measurement of overlapped blocs using a 64 by 64 pixels sliding window. With the use of a cross-validation and a three-layered Bayesian back-propagation neural network, these features are used to detect regions with masses on the mammography. With the MIAS database, the detection rate of the masses has reached 97.08%.

Full-resolution convolutional network (FrCN), a novel method for segmenting mammography images, was proposed by Al-antari et al. [12]. For automated mass segmentation in mammography, the total accuracy of the breast lesion segmentation based on FrCN was 92.97 percent, M Matthews' correlation coefficient was 85.93 percent, F1-score was 92.69 percent, and the Jaccard similarity coefficient was 86.37 percent.

Abdelhafiz et al. [13] presented a convolution neural network. The model is based on the semantic segmentation U-Net model, which was created for biomedical image segmentation tasks originally. The proposed U-Net model had a 95.10 percent mean Dice coefficient index



and a 90.90 percent mean IOU. Furthermore, using data augmentation improves the results, as the Dice coefficient index rises from 92.20 percent to 95.10 percent and from 85.0 percent to 90.90 percent, respectively.

Hossain et.al. [14] proposed microcalcification segmentation using a modified U-Net segmentation network from mammogram images. The proposed method was trained with images acquired from the DDSM database, and noise was removed using the Laplacian filter. The method was divided into five stages: image preprocessing, segmentation of breast regions, extraction of suspicious patches, selection of positive patches, and training of the segmentation network. The method produced 98.50% of the F-measure and a 97.80% Dice score

3 Methods

In this paper, we have compared the various deep-learning method for medical image segmentation. The models like a) U-net, b) U-net attention, c) Dense-net, and d) Attention Dense U-net.

3.1 U-Net

Ronneberger *et al.* [15] presented a U-Net segmentation architecture based on FCN, which has been demonstrated to be more appropriate to perform segmentation task than FCN. It consists of a decoder and an encoder. The encoder uses the CNN representational architecture, which performs the functions of the network's feature extractor and down sampling. A maximum pool layer (stride = 2) and two convolutional layers (without padding) are present in each down sampling module. With each downsampling, the number of feature channels increases and the size of the feature maps is decreased by half. The feature maps from the shallow and deep levels are combined via skip connections. After each convolutional layer, activation functions for Rectified Linear Units (ReLU) are used. The network may take into account both the participation of shallow information and deep information at the final output since it concatenates the outputs of the shallow layer and the deep layer. This yields better segmented outcomes for the segmentation of medical images as compared to FCN.

3.2 U-net with attention Gates

Oktay et al. [16] developed Attention Gates (AGs) to segment the pancreas without the use of additional networks, and these AGs outperformed a single U-Net in terms of segmentation results. Additionally, AGs are occasionally utilized in natural image analysis and language processing. The two types of trainable attention are hard attention and soft attention. The training procedure for hard-attention is more challenging because it uses reinforcement learning to update parameters, which is similar to other attention strategies. On the other hand, soft-attention is stochastic and updates parameters via conventional back-propagation. We were able to equal the translation performance of the most advanced system by incorporating additive soft attention into the encoder-decoder neural machine translation model.



3.3 Dense-net

Fully Convolutional DenseNet (FC-DenseNet) is constructed using the DenseNet architecture. FCNs add an upsampling path made up of convolution, upsampling operations, and skip connections in order to restore the input spatial resolution. In this convolution technique in FC-DenseNets with a dense block and an upsampling operation known as transition up. The previous feature maps are upsampled by a transposed convolution in the transition up modules. The upsampled feature maps are then concatenated to the ones coming from the skip connection to form the input of a new dense block. The input and output of a dense block are not concatenated in order to get around this restriction. As a result, just the feature maps acquired by the most recent dense block are subjected to the transposed convolution and not all feature maps that have been previously concatenated. At the same resolution, the last dense block compiles all the data from the earlier dense blocks.

3.4 Attention Dense U-net

Li, Shuyi, et al. [17] proposed the network with an encoder-decoder architecture which is composed of an encoder (a contracting path on the left side) and a decoder (an expansive path on the right side). The proposed network is shown as Figure 1. The encoder is a densely-connected network. Firstly, the input image is subjected to a 2×2 convolution calculation, and then the output is sent to the first dense block containing five densely-connected convolutional layers (a BN operation, a ReLU activation function and a 3×3 convolution). The input of every layer in the dense block is the feature maps output from all the previous layers of this dense block. There are five dense blocks in the proposed network.

4 Result and Discussion

In this paper the comparative study was done by using various deep-learning methods for medical image segmentation. The analysis was based on parametric metrics like Accuracy, Sensitivity, Specificity, F1 Score, and AUC. The table 1 shows the comparison of various methods using parametric metrics.

The following definitions are applied to all of the metrics:

$$\text{Sensitivity (True Positive Rate, TPR)} = \frac{TP}{TP+FN} \quad (1)$$

$$\text{Specificity (True Negative Rate, TNR)} = \frac{TN}{TN+FP} \quad (2)$$

$$\text{Overall accuracy} = \frac{TP+TN}{TP+FN+TN+FP} \quad (3)$$

$$\text{F1-score (Dice Coefficient)} = \frac{2 \times TP}{2 \times TP + FP + FN}$$

Table 1: *Parametric metrics of different methods*

Methodology	Accuracy	Sensitivity	Specificity	F1score	AUC
Densenet	77.93	76.19	80.82	81.36	83.36
Unet	74.83	68.01	86.24	77.32	82.77



Unet attention		74.37	66.92	87.62	75.71	83.26
Attention Dense Unet		78.38	77.89	84.69	82.24	86.05

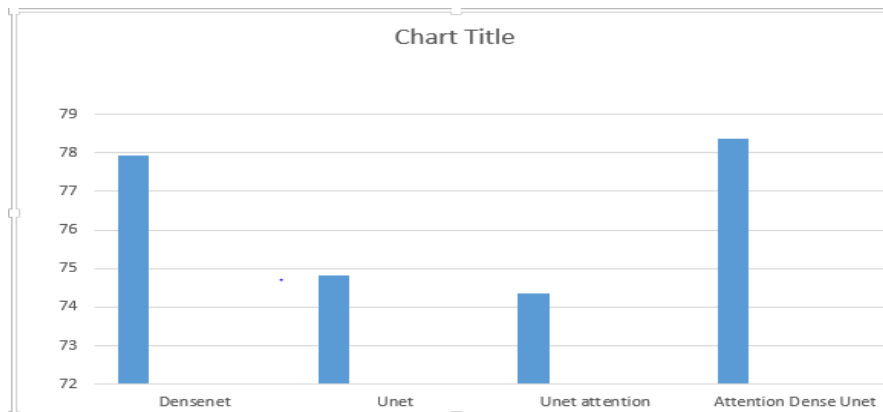


Figure 1: Accuracy

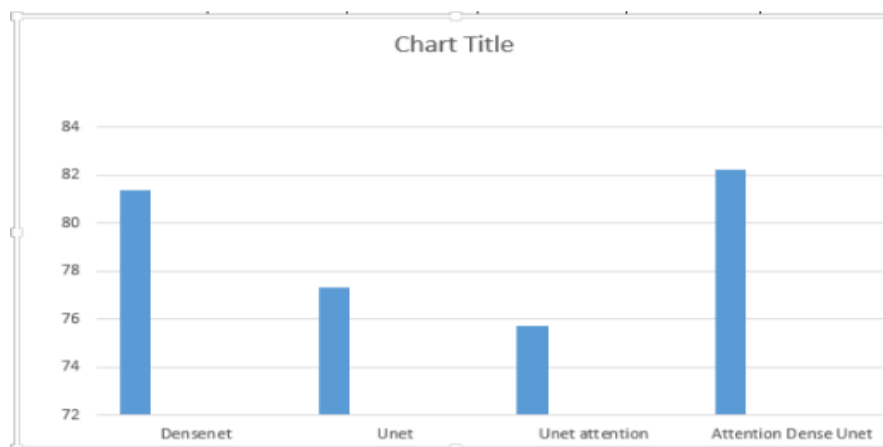


Figure 2: F1 Score

5 Conclusion

Medical Image diagnosis may require CDA based detection in order to perform accurate decision for medical expert. Image processing technique takes a major role in the prediction of results. Now a days deep learning techniques takes an important part in the diagnosis of tumor detection. In this paper we have compared different segmentation models like a) U-net, b) U-net attention, c) Dense-net, d) Attention Dense U-net. We have compared the performance metrics such as like Accuracy, Sensitivity, Specificity, F1 Score, and AUC. Attention Dense U-net outperforms the best in accuracy with 78.38%, sensitivity is 77.89%, F1 score with 82.24 and AUC with 86.05%.



References

1. Wen Li, Year: 2015, “Automatic segmentation of liver tumor in CT images with deep convolutional neural networks”, *J. Comput. Commun.*, Vol: 3, no: 11, pp. 146 – 151.
2. Chen Chen; Chen Qin; Huaqi Qiu; Giacomo Tarroni; Jinming Duan; Wenjia Bai; Daniel Rueckert, Year: 2020, “Deep learning for cardiac image segmentation: a review”, *Front. Cardiovas. Med.*, Vol: 7, pp. 25.
3. Syed Muhammad Anwar; Muhammad Majid; Adnan Qayyum; Muhammad Awais; Majdi Alnowami; Muhammad Khurram Khan, Year: 2018, “Medical image analysis using convolutional neural networks: a review”, *J Med Syst.*, Vol: 42, pp. 226.
4. Alex Krizhevsky; Ilya Sutskever; Geoffrey E. Hinton, Year: 2017, “Imagenet classification with deep convolutional neural networks”, *Commun. ACM*, Vol: 60, no: 6, pp. 84 – 90.
5. Alex Krizhevsky; Ilya Sutskever; Geoffrey E. Hinton., Year: 2012, “Imagenet classification with deep convolutional neural networks”, In: *NIPS.*, pp. 1106 – 1114.
6. Jonathan Long; Evan Shelhamer; Trevor Darrell, Year: 2015, “Fully convolutional networks for semantic segmentation”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Piscataway, pp. 3431 – 3440.
7. Yanming Guo; Yu Liu; Theodoros Georgiou; Michael S. Lew, Year: 2018, “A review of semantic segmentation using deep neural networks”, *Int J Multimed Inf Retr.*, Vol: 7, pp. 87 – 93.
8. Liang-Chieh Chen; George Papandreou; Florian Schroff; Hartwig Adam, Year: 2017, “Rethinking atrous convolution for semantic image segmentation”, Preprint, arXiv: 1706.05587.
9. Neeraj Dhungel; Gustavo Carneiro; Andrew P. Bradley, Year: 2015, “Deep structured learning for mass segmentation from mammograms”, in *2015 IEEE International Conference on Image Processing (ICIP)*, pp. 2950 – 2954.
10. Hela Boulehmi; Hela Mahersia; Kamel Hamrouni, Year: 2016, “Breast masses diagnosis using supervised approaches”, in *2016 International Image Processing, Applications and Systems (IPAS)*, pp. 1 – 6.
11. Sami Dhahbi; Walid Barhoumi; Ezzeddine Zagrouba, Year: 2015, “Breast cancer diagnosis in digitized mammograms using curvelet moments”, *Computers in biology and medicine*, Vol: 64, pp. 79 – 90.
12. Mugahed A. Al-Antari; Mohammed A. Al-Masni; Tae-Seong Kim, Year: 2020, “Deep Learning Computer-Aided Diagnosis for Breast Lesion in Digital Mammogram”, in *Advances in Experimental Medicine and Biology*, Springer, Cham, pp. 59 – 72.
13. Dina Abdelhafiz; Jinbo Bi; Reda Ammar; Clifford Yang; Sheida Nabavi, Year: 2020, “Convolutional neural network for automated mass segmentation in mammography”, *BMC*



Article Title: Comparative Study of Various Deep Learning Models For Medical Image Segmentation

Bioinformatics, Vol: 21, no: S1, pp. 1 – 19.

14. Md Shamim Hossain, Year: 2019, “Microcalcification segmentation using modified u-net segmentation network from mammogram images”, Journal of King Saud University-Computer and Information Sciences.
15. Olaf Ronneberger; Philipp Fischer; Thomas Brox, Year: 2015, “U-NetConvolutional Networks for Biomedical Image Segmentation”, In: Nassir Navab; Joachim Hornegger; William M. Wells; Alejandro Frangi, eds. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015 – 18th International Conference Munich, Germany, October 5 – 9, 2015, Proceedings, Part III Springer, pp. 234 – 241.
16. Ozan Oktay; Jo Schlemper; Loic Le Folgoc; Matthew Lee; Mattias Heinrich; Kazunari Misawa; Kensaku Mori; S. McDonagh; N. Y. Hammerla; B. Kainz; B. Glocker, Year: 2018, “Attention U-Net: Learning where to look for the pancreas”, [Online]. Available: <https://arxiv.org/abs/1804.03999>.
17. Shuyi Li; Min Dong; Guangming Du; Xiaomin Mu, Year: 2019, “Attention dense-u-net for automatic breast mass segmentation in digital mammogram”, IEEE Access, pp. 59037 – 59047.