Performance Evaluation of U-Net Architectures for Biomedical Image Segmentation

Sushant S. Borkar[1], Shreya A. Bharati[2], Bhavik O. Gade[3], Deepali M. Kotambkar[4], Jagdish D Kene[5]

[1]Student, Shri Ramdeobaba College of Engineering and Management , Nagpur, India

[2]Student, Shri Ramdeobaba College of Engineering and Management , Nagpur, India

[3]Student, Shri Ramdeobaba College of Engineering and Management , Nagpur, India

[4]Assistant Professor, Shri Ramdeobaba College of Engineering and Management , Nagpur, India

[5]Assistant Professor, Shri Ramdeobaba College of Engineering and Management , Nagpur, India

Abstract

In order to detect numerous diseases in their early stages, medical image processing is crucial. A convolutional neural network (CNN) achieves high accuracy in the semantic segmentation of medical image analysis. This strategy has a high computational cost. It is essential to use proper architecture during the training process to overcome this problem. For this, several CNN architectures have been described in the literature during the past ten years. This paper conducts a comparative study of four CNN architectures (U-Net, U-Net++, MultiRestUnet, and Dual Channel U-Net.) The accuracy and

Keywords: Biomedical Segmentation; Neural Network architecture ; U-Net; Semantic segmentation; loss of mention architectures for tuberculosis data base segmentation were explored.

DOI Number: 10.14704/nq.2022.20.9.NQ44239 **Neuro Quantology 2022; 20(9):2065-2071**

1. Introduction

Deep learning is frequently used in the interpretation of medical pictures as a result of recent developments in computer vision during the past ten years. While the use of deep learning in computer vision has spread throughout many industries, there are still challenges in the field of medical imaging. Over the years, numerous innovative techniques have been created to address these various issues. Convolutional networks are advancing recognition technology. In addition to whole-image classification, convolutional neural networks are improving at local tasks with structured output [1, 2, and 3]. Several researchers converted the existing classification networks (AlexNet [1], GoogleLeNet [3], and VGG net [2]) into fully convolutional networks and then applied their newly acquired representations to the segmentation issue by fine-tuning [4]. The basic purpose of most convolutional networks is to assign a single label to the entire image. In order to analyse medical images, pixel-level information cannot be provided by these convolutional networks. In response to such problems, a variety of strategies and methods have been created. One such method will be discussed in a work in [5] in 2015. The U-net is a deep learning strategy frequently employed in the field of medical imaging. A serious hazard to global health is tuberculosis (TB). The creation of computer-aided diagnostic systems for tuberculosis detection has benefited from the advent of fresh, very effective hardware and software approaches (TB). Even with a modest quantity of labelled training data, U-Net has demonstrated such promise in segmenting medical images that it has established itself as the industry standard [6]. Skin lesions [8, 9], colon histology [10], lung nodule [11], and other biological pictures of neural structures have all been successfully segmented using U-Net and U-Net-like models [5, 7].

In order to analyse biological images, UNet, which developed from the traditional convolutional neural network, was developed in 2015. However, in biological applications, we must not only evaluate whether there is a disease but also localise the region of abnormalities. The goal of UNet is to fix this problem. It can localise and identify borders by classifying every pixel, resulting in an output that is the same size as the input.

In this analysis, we utilize a standard image database for tuberculosis to compare the performance and losses of four neural networks used in biomedical image analysis: UNet [5], UNet++ [12], MultiRestUnet [14], and Dual Channel U Net [15].

2. Architecture of Neural Networks

2.1 U-Net

U-net is an image segmentation-focused neural network architecture. [5][12] . This paper investigated several U-net variations and their various applications on a wide range of image modalities. It also delineates the primary deep learning methodologies and application areas for each publication in this study. It was determined that U-net-based architecture is genuinely new and helpful in the interpretation of medical images.

The U-Net architecture is U-shaped, as shown in Figure 1, and consists of two blocks, the Encoder and Decoder, followed by a final segmentation output layer. Each step in the encoder block consists of two 3x3 convolutional layers with ReLu activation followed by a 2x2 max pooling layer for down sampling. This procedure is repeated four times. The decoder on the right side consists of a 2x2 up sampling layer followed by an up-convolution layer that halves the feature filters. The output is concatenated with the contracting path's corresponding output and two 3x3 convolution layers with ReLu as the activation function.

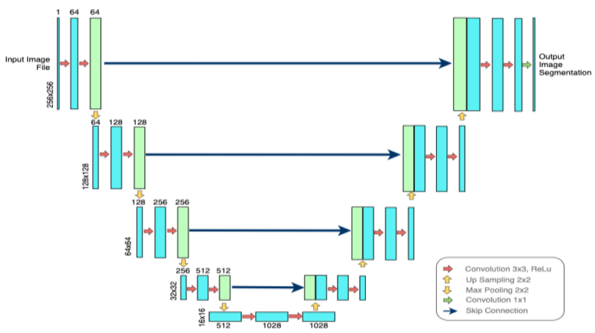

Figure 1. U-Net architecture

The applications of the UNet experiment demonstrate how the u-net may be utilized to address three distinct segmentation issues: wrapping error, rand error, and pixel error. [5] The use of this architecture to 2D biological images was quite successful. Network speed is quick. A 512x512 image can be segmented in under a second with a contemporary GPU. Accurate localization is made possible by the expanding and contracting paths being symmetric [12].

2.2 UNet++

U-net++ [12] is another effective U-net design version that takes DenseNet as its source of inspiration. It utilizes a complex network of skip links as a transitional grid between the contracting and growing paths [13]. As a result, additional semantic information is transferred between the two channels in order to aid the network in segmenting photos more precisely.

In 2018, in [13] proposed a modified UNet architecture that uses densely connected nested decoder sub-networks to improve processing and outperform UNet in medical image segmentation applications. Figure 2 shows UNet++ architecture.

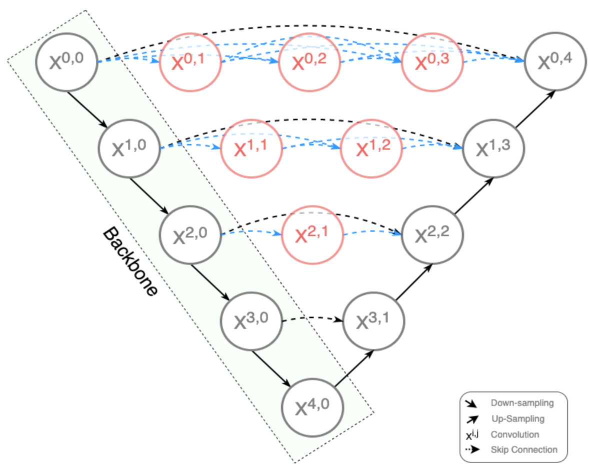

Figure 2 UNet++ architecture.

2.3 MultiResUNet

The MultiResUNet architecture, proposed in 2019, investigates the potential scopes of improvement to the original UNet architecture. [14] In this architecture, the proposed MultiRes block replaces the original UN-et's sequence of two convolutional layers at each level. MultiResUNet improves significantly with ResPath and Resblock as shown in Figure 3.

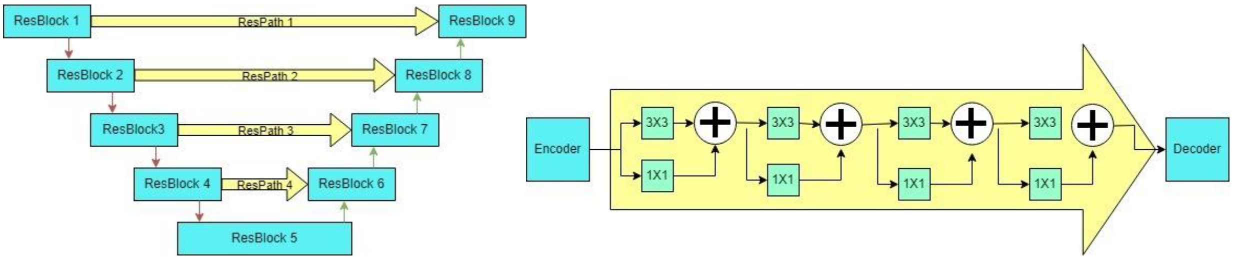

Figure 3. (a) MultiResUNet architecture, (b) ResPath

2.4 Dual-Channel UNet(D-C Unet)

The DC UNet architecture offers a way to address the shortcomings of MultiResUNet architecture by making improvements that are more effectively planned. Few additional spatial features are provided by the residual connection, which could not be sufficient for some of the most difficult tasks. The potential of the different-scale feature in medical image segmentation has already been demonstrated. Thus, in [15], a series of 3 x 3 convolutional layers were used to substitute the residual link in MultiRes blocks in order to solve the issue of insufficient spatial information that is referred to as a Dual-Channel block as shown in Figure 4.(a),(b)

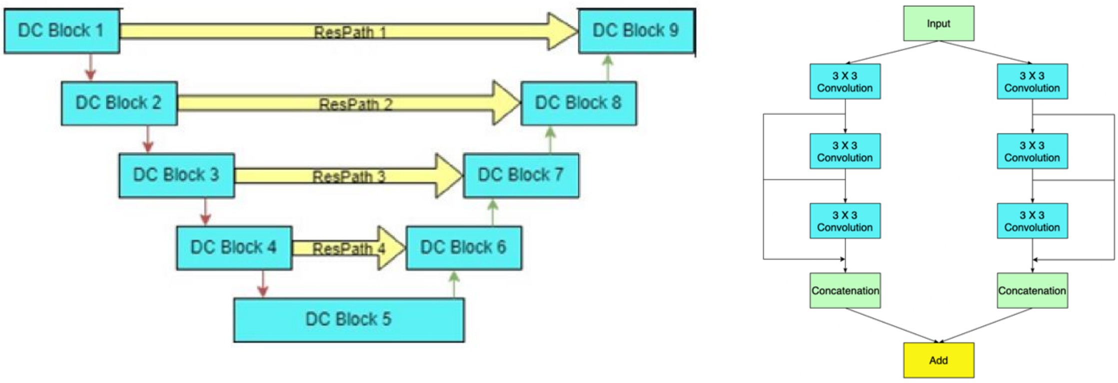

Figure 4. (a)Dual-Channel Unet architecture, (b)Dual-channel block

3 Methodology

3.1 Dataset

Montgomery County - Chest X-ray database from kaggle is used for the segmentation task. The national library of medicine and the department of health and human services collaborated to produce this standard picture database for tuberculosis. The collection includes 138 chest X-ray scans, including 58 tuberculosis patients and 80 healthy cases. There are 12 bits of grayscale in the images, which are either 4020x4892 or 4892x4020 in size. [16, 17].

3.2 Pre-processing and Training

As the X-rays in the Montgomery dataset are relatively large, we first downscale the images from 4892×4890 to 256x256 in order to reduce number of computations of training the model. Direct resizing of these photographs could affect the actual size proportion. To solve this problem, the image is first placed in the middle of the background image, which is 4892 x 4890, and then it is scaled down to the final resolution of 256 x 256 pixels. The Montgomery dataset contains only 138 X-rays, which is insufficient to train the deep learning model adequately. As a result, we employ data augmentation. Data augmentation is necessary to teach the network the needed invariance and robustness properties when there are only a limited number of sample images. [18]

The training was carried out using the python3 programming language. All models were trained on a Google colab Nvidia Tesla T4 Gpu with 16GB VRAM. All models were trained for 19 epochs at 300 steps per epoch, with a dropout rate of 0.5, the Adam optimizer, and binary cross entropy as the loss functions. The number of trainable parameters for U-Net, U-Net++, MultiRes-Unet, and DC-Unet are compared in Table 1.

Table 1: Number of Parameters trained for each Model

Model	Parameters
U-Net	31,031,685
U-Net++	9,041,601
MultiResUNet	7,237,982
DC-UNet	10,040,236

4 Results

As seen in Figure 5 and Table 2, throughout the neural network training process, the MultiResU-Net and DC net accuracy grow more quickly than the U-Net and Unet++ accuracy. The most accurate neural network is reached for each of the four architectures before 20 epochs.

Table 2: Comparison of accuracy for each model

Epochs	U-net	U-net++	MultiResU-Net	DC-Unet
1	82.91%	82.78%	96.32%	96.81%
2	97.39%	93.04%	98.61%	98.75%
3	98.15%	94.82%	99.03%	99.10%
4	98.62%	96.05%	99.61%	99.18%
5	98.84%	97.52%	99.04%	99.03%
6	98.97%	98.01%	99.12%	98.74%
7	97.81%	98.18%	98.73%	99.08%
8	98.35%	98.34%	99.06%	99.09%
9	98.86%	98.46%	98.94%	98.94%
10	98.92%	98.47%	98.97%	98.97%
11	99.04%	98.56%	98.94%	98.57%
12	98.80%	98.67%	98.72%	98.57%
13	99.02%	98.57%	98.95%	98.87%
14	98.89%	98.69%	98.55%	98.56%
15	99.15%	98.75%	98.55%	98.82%
16	99.20%	98.79%	98.82%	98.98%
17	99.20%	98.78%	98.91%	99.00%
18	99.20%	98.60%	98.62%	98.49%
19	99.26%	98.81%	98.90%	99.98%

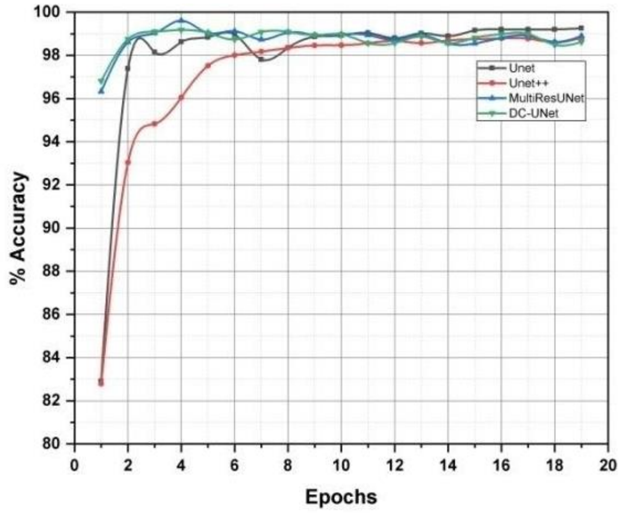

Figure 5: Comparison of accuracy for each model

A thorough comparison of the loss metrics on the validation dataset for the four neural networks is shown in Figure 6 and Table 3.

Table 3: Comparison of loss for each model

Epochs	Unet	Unet++	MultiResU-Net	DC Unet
1	0.9635	1.0026	0.3772	0.3732
2	0.1753	0.6015	0.3146	0.3098
3	0.1396	0.4261	0.2797	0.2788
4	0.1132	0.3188	0.2591	0.2587
5	0.0945	0.2263	0.2480	0.2479
6	0.0805	0.1775	0.2403	0.2434
7	0.0991	0.1490	0.2401	0.2371
8	0.0777	0.1266	0.2356	0.2354
9	0.0607	0.1092	0.2351	0.2350
10	0.0545	0.1004	0.2333	0.2332
11	0.0474	0.0898	0.2335	0.2362
12	0.0508	0.0797	0.2334	0.2350
13	0.0417	0.081	0.2337	0.2342
14	0.0423	0.0737	0.2328	0.2331
15	0.0338	0.0678	0.2351	0.2328
16	0.0309	0.0639	0.2349	0.2336
17	0.0292	0.0618	0.2329	0.2324
18	0.0281	0.0681	0.2337	0.2345
19	0.0253	0.0612	0.2329	0.2346

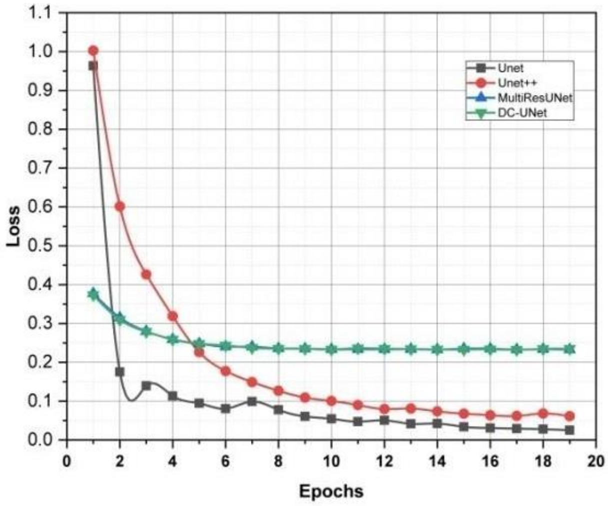

Figure 6: Comparison of loss for each model

The final accuracy and loss for all four models during training are summarized in Table 4 for the training procedure.

Table 4:

Models	Accuracy	Loss
U-Net	99.26%	0.0253
U-Net++	98.81%	0.0612
MultiResUNet	98.90%	0.2329
DC-UNet	99.98%	0.2346

5 Conclusion

In this work, we compare the accuracy and loss of the U-Net, U-Net++, MultiRestUnet, and Dual Channel U-Net architectures. In order to do the comparison for semantic segmentation, a database of images for tuberculosis was used. This database contains 138 chest X-ray scans, 58 of which are of tuberculosis patients and 80 of which are of healthy individuals. Due to the short size of the database, data augmentation is also carried out during the training period. The same data base and epochs are used to train all four models. After evaluating the accuracy and loss of mentioned U-Net variations, it was discovered that U-Net has a significant deal of potential to affect clinical applications involving automated biomedical imaging segmentation.

References

1. A. Krizhevsky, I. Sutskever, and G. E. Hinton, Imagenet classification with deep convolutional neural networks , in NIPS, 2012.
2. K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition. CoRR, abs/1409.1556, 2014
3. Szegedy, Christian & Liu, Wei & Jia, Yangqing & Sermanet, Pierre & Reed, Scott & Anguelov, Dragomir & Erhan, Dumitru & Vanhoucke, Vincent & Rabinovich, Andrew, Going deeper with convolutions. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 1-9. 10.1109/CVPR.2015.7298594, . (2015).
4. J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang,E. Tzeng, and T. Darrell. DeCAF, A deep

convolutional activation feature for generic visual recognition. In ICML, 2014.

5. Ronneberger, Olaf & Fischer, Philipp & Brox, Thomas, U-Net: Convolutional Networks for Biomedical Image Segmentation (2015).

6. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JAWM, van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. Med Image Anal. 2017 Dec;42:60-88. doi: 10.1016/j.media.2017.07.005. Epub 2017 Jul 26. PMID: 28778026.

7. Lin, Bill & Michael, Kevin & Kalra, Shivam & Tizhoosh, Hamid. Skin lesion segmentation: U-Nets versus clustering. 1-7. 10.1109/SSCI.2017.8280804. (2017).

8. Sirinukunwattana K, Pluim JPW, Chen H, Qi X, Heng PA, Guo YB, Wang LY, Matuszewski BJ, Bruni E, Sanchez U, Böhm A, Ronneberger O, Cheikh BB, Racoceanu D, Kainz P, Pfeiffer M, Urschler M, Snead DRJ, Rajpoot NM.,Gland segmentation in colon histology images: The glas challenge contest. Med Image Anal. 2017 Jan;35:489-502. doi: 10.1016/j.media.2016.08.008. Epub 2016 Sep 3. PMID: 27614792.

9. Çiçek, Özgün & Abdulkadir, Ahmed & Lienkamp, Soeren & Brox, Thomas & Ronneberger, Olaf. (2016), 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. 424-432. 10.1007/978-3-319-46723-8_49.

10. Jameson Merkow, Alison Marsden, David Kriegman, and Zhuowen Tu. Dense volume-to-volume vascular boundary detection. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 371–379. Springer, 2016.

11. Patrick Ferdinand Christ, Mohamed Ezzeldin A Elshaer, Florian Ettlinger, Sunil Tatavarty, Marc Bickel, Patrick Bilic, Markus Rempfler, Marco Armbruster, Felix Hofmann, Melvin D'Anastasi, et al. , Automatic liver and lesion segmentation in ct using cascaded fully convolutional neural networks and 3d conditional random fields. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 415–423. Springer, 2016.

12. Siddique, Nahian & Sidike, Paheding & Elkin, Colin & Devabhaktuni, Vijay. (2020),U-Net and its variants for medical image segmentation: theory and applications.

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

14. Ibtehaz, Nabil & Rahman, Mohammad, MultiResUNet : Rethinking the U-Net architecture for multimodal biomedical image segmentation. Neural Networks. 121. 10.1016/j.neunet.2019.08.025. (2019).

15. Lou, Ange & Guan, Shuyue & Loew, Murray,DC-UNet: Rethinking the U-Net Architecture with Dual Channel Efficient CNN for Medical Images Segmentation. (2020).

16. S. Candemir et al., "Lung Segmentation in Chest Radiographs Using Anatomical Atlases With Nonrigid Registration," in IEEE Transactions on Medical Imaging, vol. 33, no. 2, pp. 577-590, Feb. 2014, doi: 10.1109/TMI.2013.2290491.

17. Jaeger, Stefan & Karargyris, Alexandros & Candemir, Sema & Folio, Les & Siegelman, Jenifer & Callaghan, Fiona & Xue, Zhiyun & Palaniappan, K. & Singh, Rahul & Antani, Sameer. (2014). Automatic Tuberculosis Screening Using Chest Radiographs. IEEE Transactions on Medical Imaging. 33. 233-245. 10.1109/TMI.2013.2284099.

18. Alexander Karimov et all, Comparison of UNet, ENet, and BoxENet for Segmentation of Mast Cells in Scans of Histological Slices, arXiv:1909.06840v3 [eess.IV] 22 Nov 2019