Comparison of novel activation functions applied in U-Net architecture for cell segmentation

Aadya Chinubhai
BTech CSE
School of Engineering and Applied Sciences

Abhishu Oza
BTech CSE
School of Engineering and Applied Sciences

Razin Karimi
BTech CSE
School of Engineering and Applied Sciences

Nihar Jani
BTech CSE
School of Engineering and Applied Sciences

***Abstract*—Cell segmentation is a fundamental task in image analysis, which aims to identify and delineate cells and their components in microscopic images. It is a challenging problem due to variations in cell morphology, staining, and imaging conditions. In recent years, deep learning-based methods have shown promising results for cell segmentation tasks. U-Net is a widely used deep learning architecture that has demonstrated excellent performance in various biomedical image segmentation problems.**

In this paper, we discuss the application of U-Net for cell segmentation in microscopic images. We first provide a brief overview of U-Net's architecture and its characteristics. We then describe the novel activation functions used for comparison. Then we train the model on a dataset of human cells and compare the different activation functions on the basis of accuracy, training time and convergence.

***Index Terms*—Index Terms—Cell Segmentation, Convolutional Neural Networks, U-Net, Activation Functions**

I. Introduction

Cell segmentation is a crucial task in biomedical image analysis, which involves identifying and delineating cells and their components in microscopic images. It is a fundamental step in various applications, including cell counting, cell tracking, and disease diagnosis. Accurate cell segmentation is challenging due to variations in cell morphology, staining, and imaging conditions. Traditional methods for cell segmentation rely on handcrafted features and rule-based algorithms, which often require extensive human intervention and are prone to errors.

In recent years, deep learning-based methods have shown significant progress in various biomedical image segmentation tasks, including cell segmentation. Deep learning-based methods can automatically learn features from data, making them more robust and accurate than traditional methods. Among deep learning architectures, U-Net has gained considerable attention for its outstanding performance in biomedical image segmentation tasks, including cell segmentation.

The U-Net architecture is a convolutional neural network (CNN) that uses a contracting path and an expanding path to capture the spatial information at different scales. In this paper, we show the implementation of U-Net architecture for the cell segmentation task, including the pre-processing steps, training process, evaluation metrics, and the current state-of-the-art results. We also implement U-Net with the newly proposed activation functions cell segmentation.

II. Literature Survey

We looked at the following papers and resources for our literature survey of cell segmentation task and computer vision architectures - 1. ”Deep learning for nuclei segmentation: A review” by Bhavika Tekwani et al. (2020): This review paper provides an overview of recent deep learning-based methods for nuclei segmentation. The paper discusses various approaches, including U-Net, Mask R-CNN, and Watershed-based methods. 2. ”U-Net: Convolutional Networks for Biomedical Image Segmentation” by Ronneberger et al. (2015). This was the original paper that introduced the U-Net architecture for Biomedical Image Segmentation. 3. ”2018 Data Science Bowl” which is a dataset for the task of finding the nuclei in divergent images to advance medical discovery. 4. ”Comparison of different convolutional neural network activation functions and methods for building ensembles” by Loris Nanni et al. This paper proposed new activation functions for convolution neural networks which could improve the performance of the U-Net architecture.

III. Implementation

For the report, we have implemented the U-Net architecture on the problem of finding the nuclei of cells from the dataset ”2018 Data Science Bowl”

A. U - Net

U-Net gets its name from its U-shaped architecture, which consists of a contracting path and an expanding path.

The contracting path of the U-Net consists of several convolutional and pooling layers, which extract high-level features from the input image. The purpose of this path is to reduce the spatial dimensions of the input image while increasing the number of feature channels. This helps capture the global context of the image and the high-level features that are necessary for accurate segmentation.

The output of the U-Net is obtained using a softmax layer, which produces a probability map for each pixel in the input image. The probability map indicates the likelihood of each pixel belonging to a specific class, such as the cell nucleus or background.

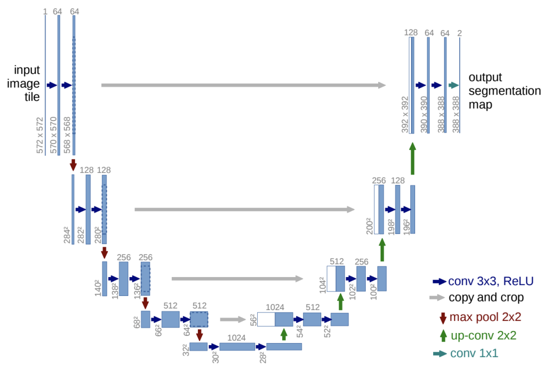

Fig. 1. U-Net architecture (example for 32x32 pixels)

B. Activation Functions

We implemented the following activation functions on UNet and compared their accuracy, convergence and training time. Mexican ReLU -

$$y = f(x) = ReLU(x) + \alpha \cdot gaussian(x, \mu, \sigma)$$

Swish -

$$y = f(x) = x \cdot sigmoid(x) = \frac{x}{1 + e^{-x}}$$

Trainable LeakyReLU -

$$y = f(x) = \{ \alpha \, x : x < 0, x : x \geq 0$$

InverseSquareRootUnit -

$$y = f(x) = \frac{x}{1 + \alpha x^2}$$

Mish -

$$y = f(x) = x \cdot tanh(softplus(\alpha x))$$
$$= x \cdot tanh(ln(1 + e^{\alpha x}))$$

Gated Swish -

$$y = f(x) = sigmoid(x) * swish(x)$$
$$= \frac{x}{(1 + e^{-\beta x})^2}$$

Learnable Swish -

$$y = f(x) = x \cdot sigmoid(\beta x)$$
$$= \frac{x}{1 + e^{-\beta x}}$$

SigmoidReLUShifted -

$$y = f(x) \{ \alpha * (sigmoid(x) - 0.5) : x < 0, x : x \geq 0$$

where α and β are learnable parameters.

C. Code Implementation

We firstly import the libraries and define our UNet class. The images are firstly imported from the dataset. Then for preprocessing, we perform the resizing of training and testing images and masks.

We then implement our selected activation functions using tensorflow.

Then we proceed to train our U-Net model created from our class on our selected dataset. Hyperparameter tuning is conducted to get the optimal values, and results are recorded for 4 Layer and 5 Layer architecture. The testing set is also used as the validation set.

IV. Results

These set of results included below are of a UNet with 4 layers. We were able to achieve an accuracy of about 87 to 90% from our implementation.

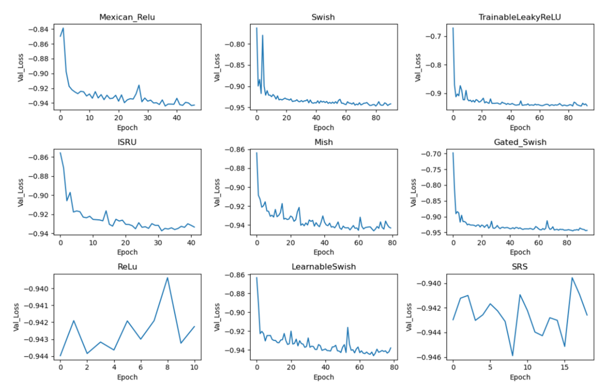

Fig. 2. Validation Loss Graph for activation functions

Activations	*Top-1 Accuracy*	*Epochs till Convergence*	*Training time (s)*
Mexican ReLU	90.08	71	125.46
Swish	89.88	79	82.74
Trainable LeakyReLU	87.59	27	32.94
InverseSquareRootUnit	89.11	61	79.65
Mish	89.93	60	65.71
GatedSwish	90.01	67	78.63
ReLU (Baseline)	89.81	21	22.24
Learnable Swish	90.11	63	75.78
SigmoidReLUShifted (SRS)	90.15	24	27.584

Fig. 3. Accuracy, Convergence and Training Time for activation functions

V. Conclusion

The results showed that the four layered U-Net architecture performed the best among the tested models. Although there were slight variations in the performance of different activation functions, ReLU showed faster convergence due to its non-saturating nature. Moreover, ReLU had the shortest training time, which is crucial in deep learning. When it comes to accuracy, SRS outperformed the other methods. Overall, these findings emphasize the importance of carefully selecting the right architecture and activation function to achieve optimal performance in deep learning models.

References

[1] ”Deep learning for nuclei segmentation: A review” by Bhavika Tekwani et al. (2020)

[2] ”U-Net: Convolutional Networks for Biomedical Image Segmentation” by Ronneberger et al. (2015) https://arxiv.org/abs/1505.04597

[3] ”2018 Data Science Bowl” https://www.kaggle.com/c/data-science-bowl-2018

[4] ”Comparison of different convolutional neural network activation functions and methods for building ensembles” by Loris Nanni et al. (2021) https://arxiv.org/abs/2103.15898