Fuzzy U-Net Neural Network Design for Image Segmentation

Mark Kirichev, Todor Slavov, and Galina Momcheva(✉)

Department of Computer Science, Varna Free University "Chernorizets Hrabar", Yanko Slavchev str 84, Chayka resort, 9007 Varna, Bulgaria
galina.momcheva@vfu.bg

Abstract. One of the problems in biomedical image analysis is the problem of nuclei segmentation. The annotation of nuclei by hand has proven itself very time consuming with varying results depending on many factors. More recently convolutional neural networks have made the problem of automatic image segmentation easier, faster and more reliable. In this article, an extension to the standard U-Net is proposed which aims at improving the quality of biomedical image segmentation. By integrating Fuzzy computations in the standard U-Net architecture we have achieved even better accuracies than the ones reached by the base architecture. The Fuzzy Layers resemble a sense of uncertainty, which is already seen in the real world. This allows for a more precise detection and instance segmentation of cellular nuclei. When the original U-Net was first developed one of its focuses was to be trainable with a small dataset. This quality has been proved useful when undertaking the task of Biomedical Image Segmentation. The model was trained with the Kaggle 2018 Dataset. The segmentation process comes right after, including the following two steps: 1) creating a prediction matrix, 2) thresholding the matrix to attain a visual result. The second step exhausts the following techniques: Manual Thresholding; Adaptive Thresholding; Gaussian Thresholding, and Otsu Thresholding. The results point out which thresholding technique, combined with a certain set of Fuzzy Layers, yields the best possible results in terms of accuracy of the models.

Keywords: Deep learning · Fuzzy layer · U-net · Semantic segmentation · Otsu thresholding · Hyperparameter optimization · Keras tuner · Grid search · Automated ML

1 Image Segmentation

1.1 Description

In digital image analysis, image segmentation is the process of dividing an image into multiple segments (pixel clusters). The goal of segmentation is to give a better perspective over an image making it easier to analyze it. It can also be used for quantitative measurements. The technique can be used for object recognition and tracking. On a more basic level the process of assigning a value to every pixel in an image, such that pixels that share the same value have some qualities or characteristics in common [1].

S. S. Sotirov et al. (Eds.): BioInfoMed 2020, LNNS 374, pp. 177–184, 2022.
https://doi.org/10.1007/978-3-030-96638-6_19

1.2 Problems

Manual image segmentation is a very time consuming task – a single image can take minutes of an experienced professional, which is not desirable in the field of medicine [2]. Automated techniques significantly reduce this time – a standard CNN takes less than a second [3]. The segmentations of microscopic images are highly sensitive to experimental conditions (such as human error, level of expertise, fatigue and subjectivity), which often leads to increased variability of the final images even in slides processed in one batch. To reduce this variance we introduce Fuzzy Logic in image segmentation.

2 Fuzzy U-Net Model Architecture

2.1 The Standard U-Net

The U-Net is a convolutional neural network designed for fast and precise semantic segmentation of images. It was released in 2015 [4] and has proven itself repeatedly to be one of the best currently available methods for image segmentation. The aim of the design is to perform well even with a minimal amount of training data and still yield satisfactory results (Fig. 1).

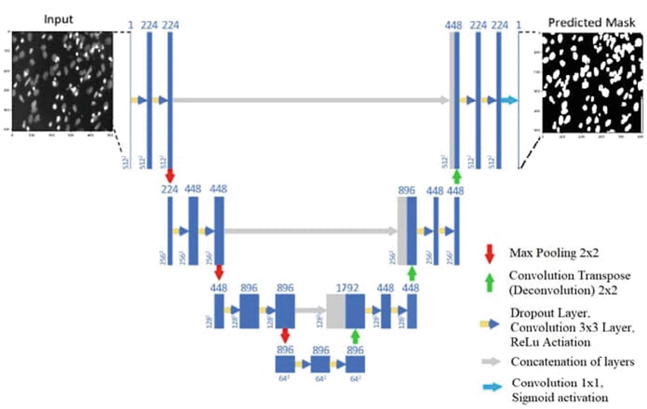

Fig. 1. Graphical representation of the design of our U-Net model, used for obtaining the hereinafter results

2.2 Initial Cross Validation

For the optimization of the parameters of the model used for predictions and for its testing, the Kaggle Dataset was divided in a ratio 9:1 [5]. The larger part has been introduced to the model in the training phase, while the smaller was used in the process of validation and comparison of true results versus predicted ones.

In order for this overfitting problem to be avoided, we have implemented the following two approaches: 1) using dropout for every layer, ensuring a relative randomness in the model. The dropout values vary with every layer, rising with the increase of depth in terms of the convolution layers. Moreover, the dropout values represented an arithmetic progression model with a first term of 0.1 and a common difference of 0.1; 2) the 90% fraction of the original dataset is kept constant, however, the particular images varied with each iteration of the training process. This random split into training and testing sets is easily achieved using the Sci-kit library and the train_test_split function. The 90% selection of the original 700 segmented nuclei images, on which the training is performed, coupled with the dropout ensure that a set of different enough inputs is introduced to the model in order for it to be optimal. Through multiple epochs, this method proves to be a key element in evaluating the best possible weights in the neural network architecture.

2.3 Integration of Fuzzy Layers into the Standard U-Net

Concatenating Fuzzy Layers with the standard U-Net architecture has proven to yield better results both accuracy-wise and visually [6]. Adding Fuzzy Layers in some key points of the model allows for a level of uncertainty, which better represents a nature-like behavior [7]. Therefore, the placement of the Fuzzy Layers in the model serves for a way for it to display different results when the Layers are differently arranged. Depending on the positioning of the Fuzzy Layers, they are taken with a different weight in the final calculations, which means there is room for optimization in the placement of said layers [8].

Table 1. Summary of trained models with 1–5 Fuzzy Layers concatenated to the U-Net.

Number of layers	1	2	3	4	5
Number of such Fuzzy architectures	15	30	30	15	3
Average accuracy	0.854288	0.852531	0.849569	0.849902	0.837194
Maximum accuracy	0.871554	0.8838943	0.877656	0.869461	0.851802
Minimum accuracy	0.823202	0.8133097	0.82635	0.829388	0.812643

In Table 1 shown above is shown the summary of all the trained models consisting of 1, 2 3, 4 or 5 integrated Fuzzy Layers. We can see that in the average case best accuracy is achieved when 1 FL is used but out of all trained models, the one that performed the best had two Fuzzy Layers.

Table 2. Summary of all the models consisting of the first through fifth Fuzzy Layer

FL activated	A	B	C	D	E
Number of such architectures trained	48	48	48	48	48
Average accuracy	0.852984	0.848939	0.85071	0.850761	0.846178
Maximum accuracy	0.877656	0.883894	0.883894	0.873607	0.877313
Minimum accuracy	0.81264	0.81264	0.81264	0.81264	0.81264

In Table 2 layers A, B, C, D, and E correspond to the Fuzzy Layers, shown on Fig. 2 starting from the layer which is visually represented as the highest one and continuing in the down direction. In Table 2. all models that have with the aforementioned (A-E) layers are summarized. The conclusion that can be drawn is that the first Fuzzy Layer has the best impact on the overall accuracy of the model. In contrast, models consisting of the fifth layer usually perform worse.

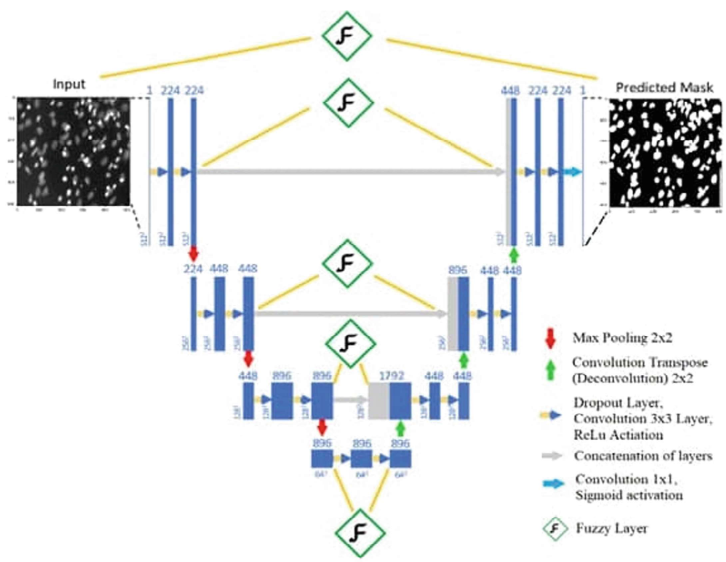

Fig. 2. Graphical representation of all the possible placements of the Fuzzy Layers around the original U-Net

The position of the Fuzzy Layers in the U-Net model in Fig. 2 has been determined to be optimal in between the different sets of convolution-dropout layers and deconvolution-dropout layers. This is due to the fact that these sets of layers form a frame of images, carrying the same number of pixels and, therefore, the same quantity of information for the neural network architecture. The Fuzzy Layers, in this case, represent an alternative path in the model, which exploits the use of the Fuzzy membership function instead of the usual Conv2D or Conv2DTranspose transformations. By concatenating the results from both paths the neural network has information of both the normal U-Net, as well as the Fuzzy Layers, which surround it.

3 Thresholding

3.1 Thresholding Methods Tested

To produce as good as possible bitmasks from the acquired predictions (Fig. 3.1) we tried different thresholding techniques such as manual thresholding, Global isodata thresholding, Adaptive Gaussian, Adaptive mean and Otsu thresholding [10]. The first method would return satisfying results only in some cases, while in others there would be no distinguishable segmentation between some nuclei. In the average scenario, the second method would yield good results but in some extreme cases would also not perform well. The third and fourth approach would also perform well but we found out that the final approach proved to be effective for all occasions. In conclusion, global methods of thresholding appear to give good bitmasks overall but not as good as Otsu's method [11] (Fig. 3.2).

Finally, a thresholding was implemented, using the following mathematical expression:

$$\begin{cases} 1, P(x, y) \geq \sigma_{\omega}(x, y) + G(x, y) * W \\ 0, P(x, y) < \sigma_{\omega}(x, y) + G(x, y) * W \end{cases},$$

where $P(x, y)$ is the value of the current pixel, $\sigma_{\omega}(x, y)$ is the Otsu value, $G(x, y)$ is the Gaussian blur factor and W is a weight, used to negate biases (Figs. 4, 5 and 6).

After experimenting with different values of W, we found out that this method gave us the most visually desirable output, which we accomplished to derive from the input. The following results prove it by highlighting better segmentation, better nuclei separation and sharper object boundaries.

Fig 3.1.
Original Image

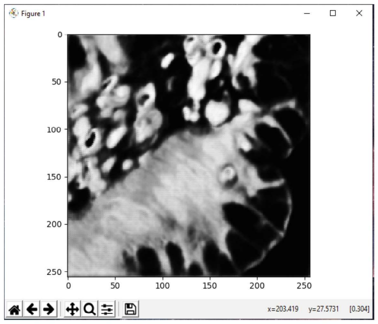

Fig. 3.2.A
Otsu Thresholding

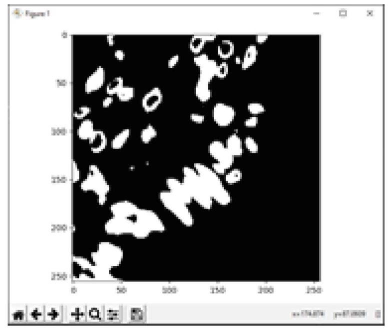

Fig.3.2.B
Adaptive Thresholding

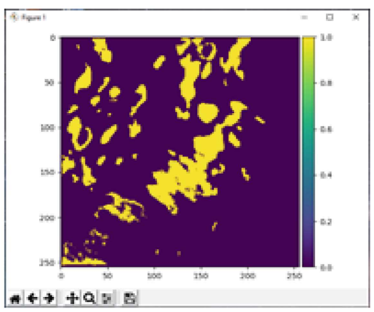

Fig. 3.2.C
Adaptive Gaussian

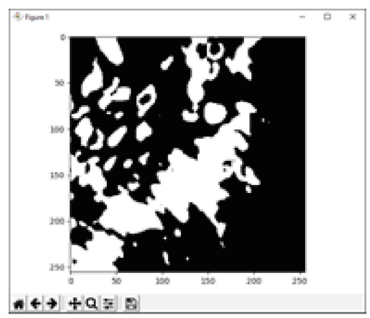

Fig.3.2.D
Manual
(with high T value)

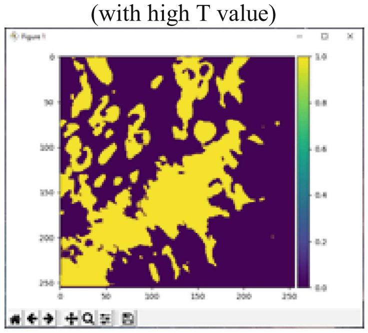

Fig. 3.2.E
Manual
(with low T value)

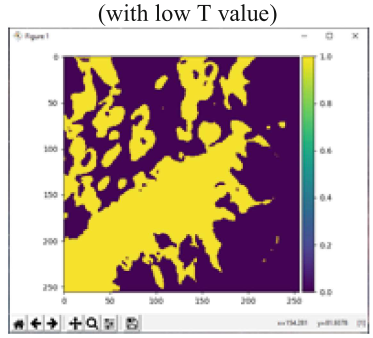

Fig. 3. Visual representation of the Original image and the types of thresholding performed

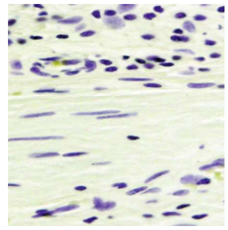

Fig. 4. Original image

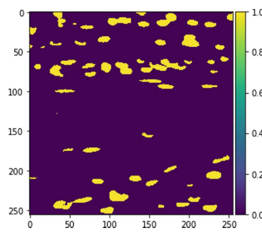

Fig. 5. Otsu method

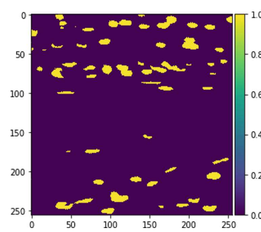

Fig. 6. Our method

3.2 Conclusion

The different placement of the Fuzzy Layers results in different model accuracies and yields varying visual representations as a result. The Fuzzy Logic helps to differentiate the groups of nuclei sharing a common part and makes those differences more visible. The accuracy values for the training of Fuzzy Layers implemented inside the U-Net architecture range from 0.81 (approx.) to 0.88 (approx.). This proves that the models depend on the location of the Fuzzy Layers and the segmentation can be further developed through the addition of a proper set of them. Moreover, the number of included Fuzzy Layers also has a major impact on the results with just above 7% accuracy difference amongst the number of Fuzzy Layers incorporated in the model.

References

1. Haralick, R., Shapiro, L.: Image segmentation techniques. Comput. Vis. Graph. Image Process. **29**(1), 100–132 (1985)

2. Rosas González, S., Birgui Sekou, T., Hidane, M., Tauber, C.: 3D automatic brain tumor segmentation using a multiscale input U-net network. In: Crimi, A., Bakas, S. (eds.) BrainLes 2019. LNCS, vol. 11993, pp. 113–123. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-46643-5_11
3. Falcao, A., Udupa, A., Miyazawa, J.: An ultra-fast user-steered image segmentation paradigm: live wire on the fly. IEEE Trans. Med. Imaging **19**(1), 55–62 (2000)
4. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
5. Kaggle Data Science Bowl 2018 Dataset. https://www.kaggle.com/c/data-science-bowl-2018/data. Accessed 10 Oct 2020
6. Korshunova, K.: A convolutional fuzzy neural network for image classification. In: 3rd Russian-Pacific Conference on Computer Technology and Applications (RPC), Vladivostok, pp. 1–4 (2018)
7. Xi, Z., Panoutsos, G: Interpretable machine learning: convolutional neural networks with RBF fuzzy logic classification rules. In: International Conference on Intelligent Systems (2018)
8. Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., Talwalkar, A.: Hyperband: a novel bandit-based approach to hyperparameter optimization. J. Mach. Learn. Res. **18**, 6765–6816 (2018)
9. Springenberg, J., Klein, A., Falkner, S., Hutter, F.: Bayesian optimization with robust Bayesian neural networks. In: Advances in Neural Information Processing Systems, vol. 29, pp. 4134–4142. Curran Associates, Inc. (2016)
10. Goh, T., Basah, S., Yazid, H., Safar, M., Saad, F.: Performance analysis of image thresholding: Otsu technique. Meas. J. Int. Meas. Confed. **114**, 298–307 (2018)
11. Gong, J., Li, L., Chen, W.: Fast recursive algorithms for two-dimensional thresholding. Pattern Recogn. **31**(3), 295–300 (1998)