Improving Deep Neural Network's Segmentation Efficiency Of Medical Images By Using Image Features Explicitly

Kajal Gupta

Department of Computer Science

New Jersey Institute of Technology

Newark, NJ 07102, USA

Contact: kg494@njit.edu

ABSTRACT

Use of Deep Neural Networks for segmenting images of lungs, blood vessels, brain MRIs and many others have become a popular choice in the medical imaging. Dependence on these results for careful analysis and prediction of diseases has lead to great stress for improving the quality of results produced by these models. This leads to an increase in research work revolving around their architecture with little focus given to its driving engine i.e "input image" behind it. Since these networks are robust enough to handle all kinds of data in its raw form eliminating the need of selecting the features manually[1], the focus given on the data side has been reduced tremendously. Since mistakes made in the medical imaging might be crucial and more impacting, this paper explores and presents research work in improving image segmentation of Deep Neural Network by including image features as an added input channel along with the original image. This paper presents the techniques of preparing such data and compares the experimental results of training a model with this newly prepared data, with a baseline model on

three different medical datasets. Our research showcases improvement in the segmentation efficiency of the model when trained with the data that we prepared by incorporating image features explicitly as input channels.

I. INTRODUCTION

Inspecting the available medical images to find out diseases is a long and tedious process done by doctors and radiologists. With the advent of GPUs and software tools for training a DNN models and migration of medical industry[3] for relying on these technologies allows doctors and experts to identify many diseases at levels exceeding certified doctors and experts. However, the notion continues that the architecture of Deep Neural Network[2] is the sole criteria for accurate segmentation along with minor data preprocessing. Hence more work is focused on improving the model architecture to make the segmentation better, however, no research work is centred around data that is fed to these models and whether those data could be highlighted in some way and fed as extra features to the model to improve its accuracy. This paper is dedicated to explore the use of data content of an image to produce new features and feeding it to the model explicitly as additional channels to improve its accuracy without changing the architecture of Deep Neural Network and thus achieving better results.

II. PROBLEM STATEMENT AND EXISTING METHODS

Before the use of deep learning many techniques were employed during classical machine learning to extract useful features from the medical images. Preprocessing were done to reduce technical variability across images which occurred from different scanning instruments, signal drifts and other calibration related issues via resampling of dimensions to isometric voxels to homogenize image resolutions. Other medical images like that of MRI, required normalization of intensity pixels and sometimes were complemented with spatial normalization. Conversion from one color space to another like RGB to HSV and operating in it is another popular notion in medical image processing. Additionally the use of engineered features, to quantify specific radiographic characteristics of diseased tissue were some other measures. Sometimes, domain-specific knowledge about the diseased tissue were incorporated and automated to extract the features. Apart from feature engineering, data augmentation like cropping, reflection and rotation were applied to increase robustness.

Recently, promising statements were achieved in Chest X-Ray[4], Retinal vessel[5] and cell membrane[6] segmentation using deep learning. Various datasets for all these medical analysis has been released publicly for research. Various architecture like U-Net[7], DUNet[8], focuses on improving the segmentation accuracy. Feature extraction are not done explicitly for these networks as they themselves have a higher level of feature abstraction in the layers and provide better prediction performance. The processing steps applied to these are normalization, mean shift, cropping, PCA whitening, and augmentation. These networks work directly on raw images convolving them with kernels and rely on it to learn each of the useful features as it moves higher in the network layer. Sometimes a post processing step such as Conditional Random Fields [9] is applied to model output in order to improve the fineness of model segmented output. However the challenge imposed by all these networks are that they require huge dataset in order to bring out the accuracy so that it's worth retaining.

With the advent of Deep Neural Networks and shifts towards it from traditional models to Convolutional Neural Networks, one key thing that all these research work fails to understand is that model behaves as per the input data. There might be some important features [10, 11] of images that might be useful for one problem and suited to one kind of data, and might differ from other problems and other kinds of data. Also, there is a limit to which model params can be tuned and architecture can be changed, as when our data changes, all those changes will become obsolete and will require re-tuning of parameters and model architecture to suit the new data. Since images are key players here, we need to focus on using the underlying content information in some way so that it can boost model learning capabilities. Hence this research work tries to explore whether for a fixed Convolutional Neural Network Architecture, can we improve the segmentation efficiency, localization and finesses just by incorporating certain features of images so that instead of relying on network to learn on its own, these features could be provided as an additional channel as an input along with the original image itself.

III. NEW METHOD FOR PREPARING TRAINING AND TESTING DATA

Inorder to test whether our approach worked better as compared to base model, we tested it on three medical datasets. All these dataset were downloaded from Kaggel site and there original sites where it was available freely to conduct research. The three datasets are as follows: Chest X Rays and masks label[12]: this dataset contains 704 train and 96 test images. Cell membrane[13]: this dataset contains 30 train and 30 test images. DRIVE Digital Retinal Images for Vessel Extraction[14]: this dataset contains 20 train and 20 test images. We start by observing what features should be useful that could help model visualize and understand the features more. If we look at these medical images, then for some we can definitely identify the masks ourselves, but for others even we get confused how to segment them out. For example in fig. a chest x ray image, we are able to clearly see the boundaries, but for fig. b, it's hard to see the exact boundary of the chest x ray. Similarly for cell membrane and retina, it's even harder for humans to see the segments clearly.

Fig. a

Fig. b

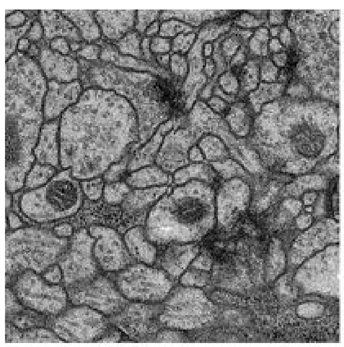


Fig. c

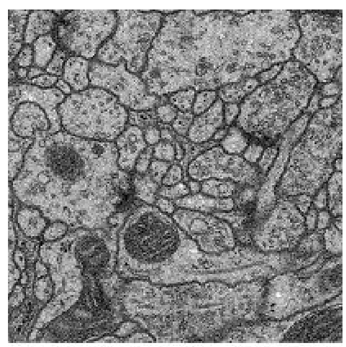


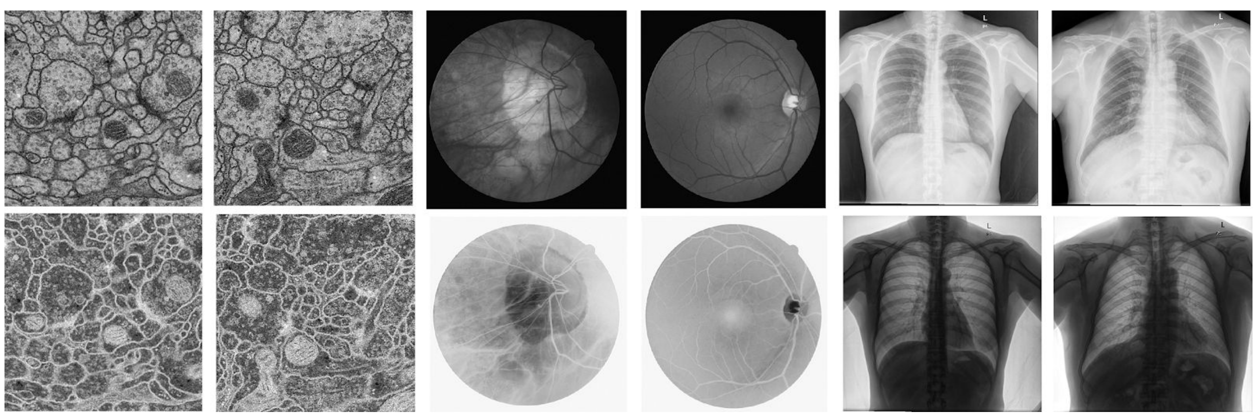
Fig. d

Fig. e

Fig. f

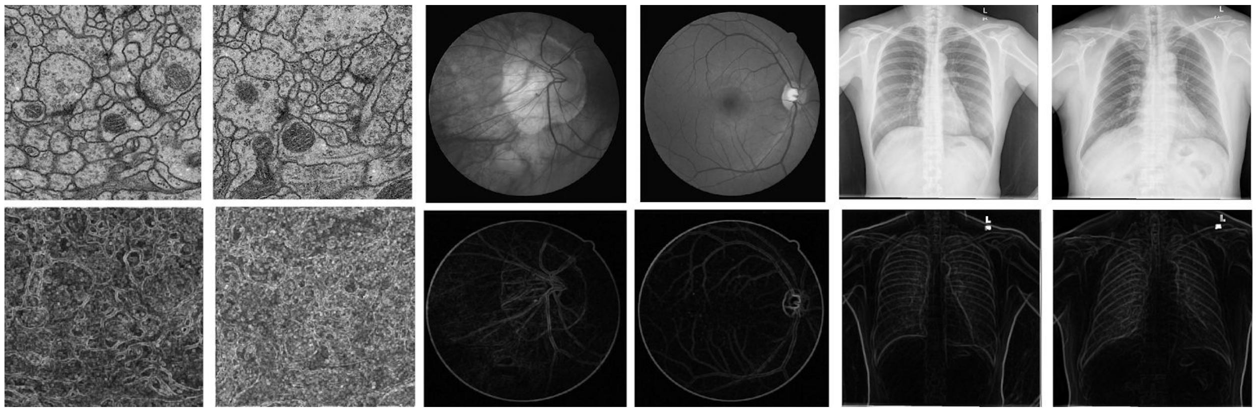
Hence, our assumption while doing this research was if any feature helps us in distinguishing the segment clearly it will help the machine too. We start the process of identifying the features that would help us see the boundaries clearly, identify similar regions and distinguish one region from others. Hence for this research we chose to play with **Image Negatives, Image Skeletons, Pseudo Color Image Processing**[15].

First we began by generating image negatives for all the datasets. Thus for each of the image above, our output feature will look like these:



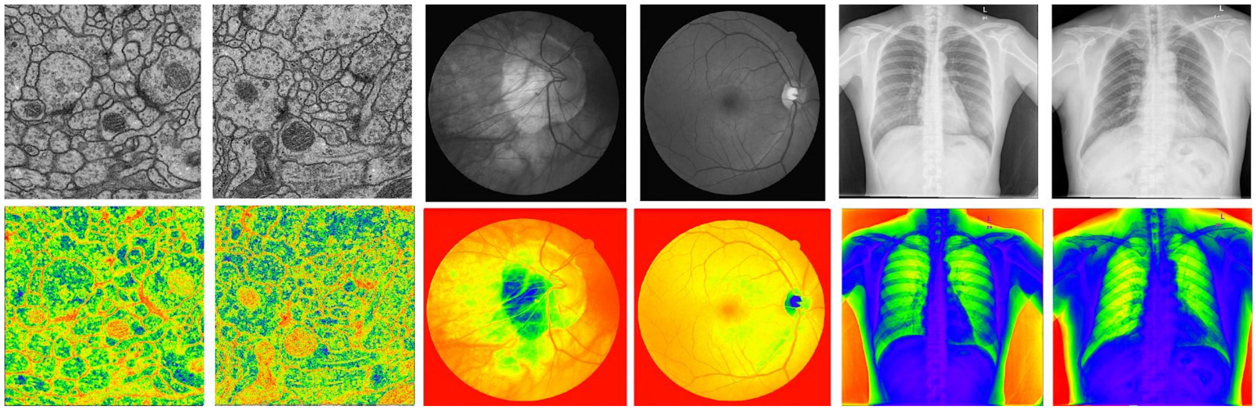
As you can see the duals complement each other. The information that seem to be hidden in original image, seems to contrasted in the negatives and help us to make out the features more to some extent. In cell membrane, the white region are more prevalent in negatives and so is for chest x ray. Lungs appears to be more visually embossed in negatives as compared to original image. Hence we take negatives as **first additional feature** channel for our model.

Next, we work on identifying the boundaries of the image so that model is able to learn more about the shape of the segments and image as a whole. Although lower layers of neural networks are said to learn these features, we wanted to try whether giving them in as an additional inputs will help the model in some way or not. Hence we computed the gradient of image to determine the skeleton. For each of the image above gradient image looks like these:

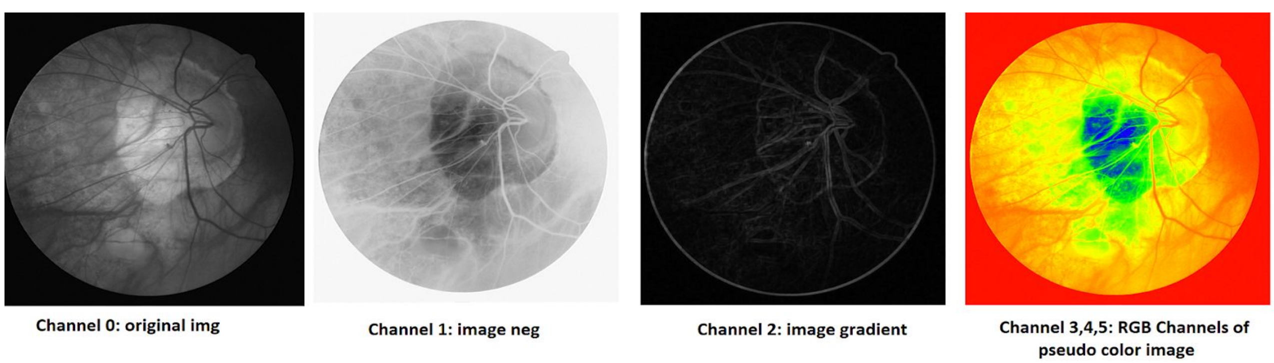


You can notice that although it looks like noise in cell membrane and almost indistinguible to human eyes, the pixel values of the gradients will make much better sense to machines while dealing with it. For retinal vessel and chest x ray images, one can easily see the outlines of retinal vessels and lungs along with body outline. Our intuition is that this boundary/skelton will help refine the segmentation masks of the images by the model. Hence, we choose these gradient images as **second feature** channel of our image.

For our **final feature** image we focus on pseudo coloring techniques that help us to assign colors to gray value based on specific criteria. By varying the number of colors and the span of intensity colors, one can quickly determine the characteristics of intensity variation in grayscale images. This will help to visualize the medical images such as chest x rays and retinal vessels in new ways that will help inspectors and us to discriminate and identify regions more significantly. For this technique we simply use pseudo coloring function available in Matlab and prepare the color images for each dataset. For each of the above images, we see our final features as below:



Once all the channels are ready for all the three datasets, we concatenate these individual channels in one so that it can be used for training and testing purposes. For doing so we pick the newly created channel 1, channel 2, channel 3 and original grayscale image image and stack them together one by one along column to create six channel final image which we input to the model. The shape of the final image will be 256x256x6. The additional 5 channels are 1 for negative, 1 for gradients and 3 RGB channel of pseudo color image map. Hence total of 6 channels for inputs.



IV. EXPERIMENTAL RESULTS

A. Training

The above prepared data for each dataset were used to train a Deep Neural Network Model. The architecture used is U-NET. U-NET [8] is a Convolutional Networks for biomedical image segmentation. A baseline model was trained alongside the experimental model for comparison against all the three models. All models were trained on the same architecture, hyper parameters, batch size and epochs. Same set of training and test sets were used for all three models respectively.

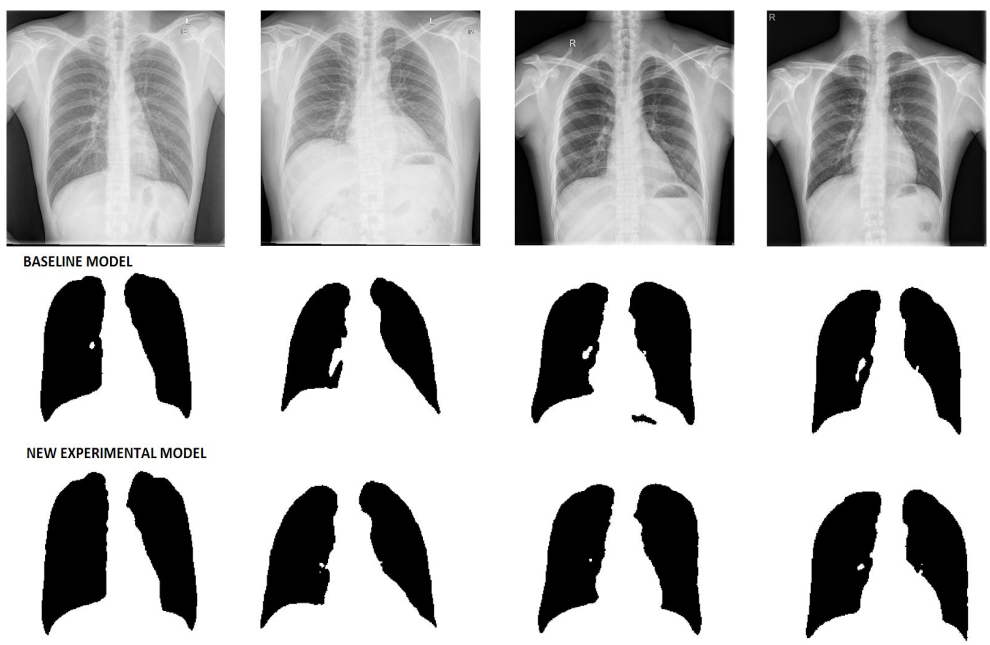
B. Performance Analysis and Results

All the trained models with newly prepared features were compared against baseline model and following accuracy have been reported. From the table one can clearly see that the model trained with our method of incorporating features explicitly as additional channel performed well as compared to the baseline model. Along with accuracy, three additional gains were observed: improved in fineness of generating the segmentation boundaries, increase in learning speed during early epochs and improved localization of objects, removal of false positives.

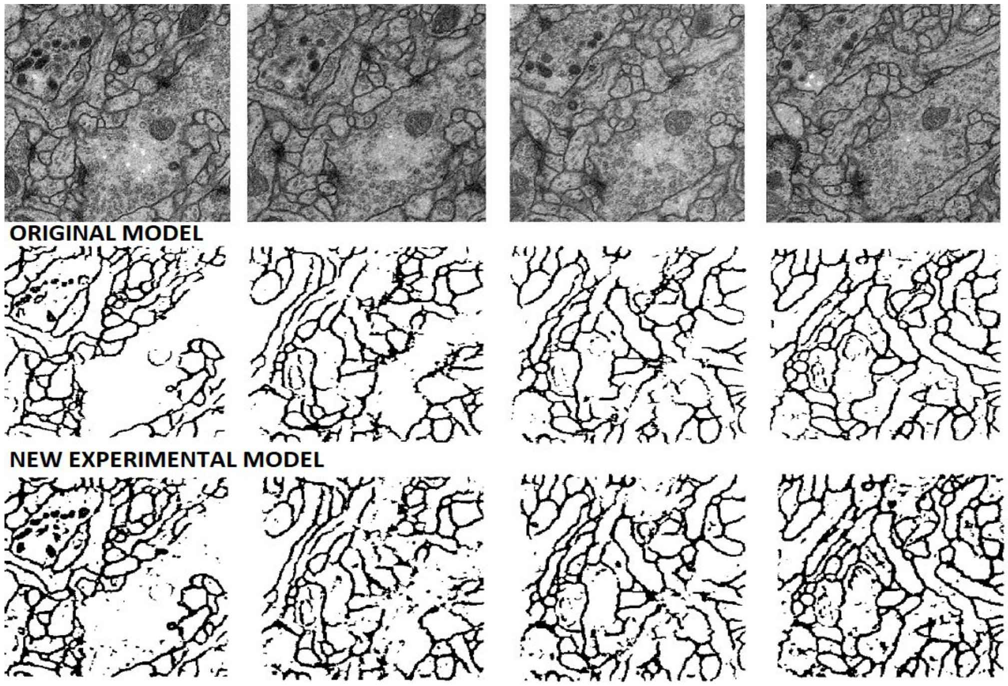
Dataset	**Model**	**Epochs**	**Training Accuracy**	**MSE**	**Loss**
Cell membrane	New Experiment	Steps per epoch: 300 Epoch: 1	0.8872	0.0892	0.30970
	Baseline		0.8619	0.0963	0.37998
Chest X Ray	New Experiment	Steps per epoch: 300 Epoch: 3	0.9782	0.0231	0.11148
	Baseline		0.9660	0.0257	0.13079
Retinal Vessel	New Experiment	Steps per epoch: 300 Epoch: 1	0.9297	0.0058	0.11541
	Baseline		0.9180	0.0072	0.12925

The results can be summarized futher. If we analyse the baseline model with new experimental model for Chest X Ray/ Cell Membrane and Retinal vessel dataset we can see that new model is able to generate fine masks, better localization and no false positive areas as compared to baseline model which generated broken/irregular masks, no smooth boundaries and false positive areas. This shows that our new model trained with explicit features perform much better as compared to the baseline.

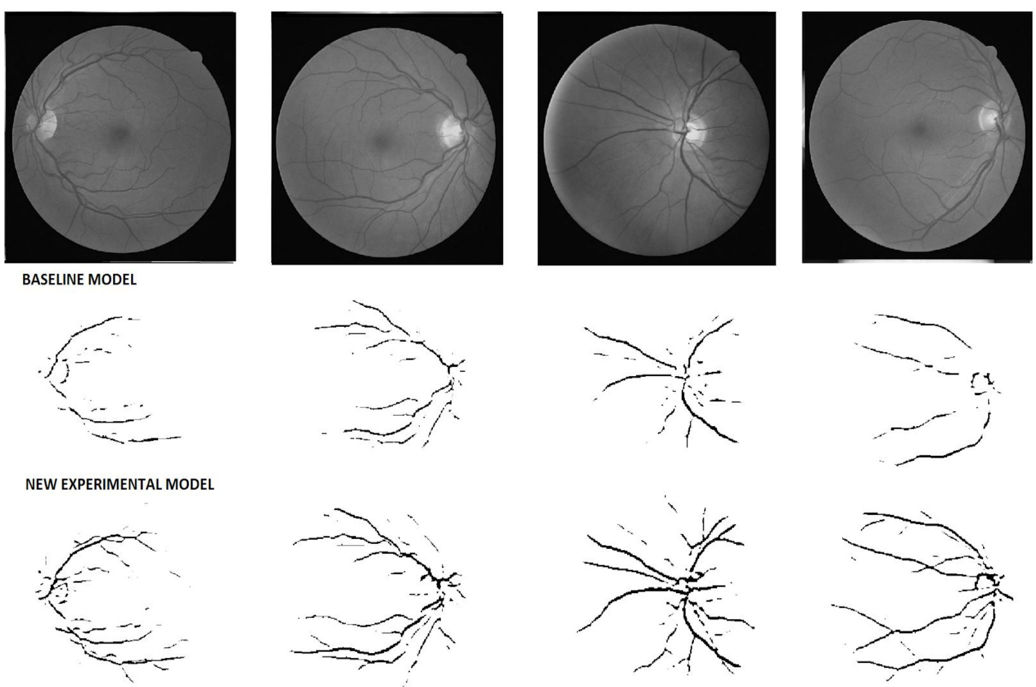
Chest X Ray Results:



Cell Membrane Results:



Retinal Vessel Segmentation Results:



V. CONCLUSION

In this research work we proposed a novel method of preparing training dataset by incorporating image features as an additional channel which lead to improvement in segmentation efficiency of the model. Our main aim was to make use of features that are present in data to improve the segmentation efficiency since there is a limit to how frequently an architecture can be changed when data changes. So instead of focusing on changing the architecture, we proposed to focus on underlying features of medical data so that any kind of Deep Neural Network can be modeled with it to perform better resulting in better segmentation quality.

REFERENCES

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems 25*, 2012, pp. 1097–1105.

[2] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” arXiv preprint arXiv:1409.1556, 2014.

[3] D. Shen, G. Wu, and H. Suk. "Deep learning in medical image analysis." Annual review of biomedical engineering, vol. 19, pp. 221-248, 2017.

[4] Stirenko, Sergii & Kochura, Yuriy & Alienin, Oleg & Rokovyi, Oleksandr & Gang, Peng & Zeng, Wei & Gordienko, Yuri. (2018). Chest X-Ray Analysis of Tuberculosis by Deep Learning with Segmentation and Augmentation.

[5] G. Hassan, N. El-Bendary, A. E. Hassanien, A. Fahmy, A. M. Shoeb, and V. Snasel, “Retinal Blood Vessel Segmentation Approach Based on Mathematical Morphology,” Procedia Computer Science, vol. 65, pp. 612–622, 2015.

[6] Ciresan, D.C., Gambardella, L.M., Giusti, A., Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images. In: NIPS. pp. 2852–2860 (2012)

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

[8] Jin, Qiangguo & Meng, Zhaopeng & Pham, Tuan & Chen, Qi & Wei, Leyi & Su, Ran. (2018). DUNet: A deformable network for retinal vessel segmentation.

[9] S. Zheng, S. Jayasumana, B. Romera-Paredes, V. Vineet, Z. Su, D. Du, C. Huang, and P. H. Torr, “Conditional random fields as recurrent neural networks,” in *Proceedings of the IEEE International Conference on Computer Vision,* 2015, pp. 1529–1537.

[10] Jackowski, K., Cyganek, B. (2015). “A learning-based colour image segmentation with extended and compact structural tensor feature representation. Pattern Analysis
And Applications”, 20,1-14

[11] Jackowski, K., Cyganek, B. (2015). “A learning-based colour image segmentation with extended and compact structural tensor feature representation. Pattern Analysis
And Applications”, 20,1-14

[12] WWW: Kaggel page for Chest X Ray and Masks dataset. https://www.kaggle.com/nikhilpandey360/chest-xray-masks-and-labels

[13] WWW: Web page of the em segmentation challenge, http://brainiac2.mit.edu/isbi_challenge/

[14] WWW: Web page of the Drive: digital retinal images for vessel extraction. https://drive.grand-challenge.org/

[15] Digital Image Processing, fourth edition, R. C. Gonzalez, R. E. Woods