Deep Learning in Medical Field: Nerve Segmentation

Chayan Gulati
chayan18csu054@ncuindia.edu

Ashish Chhabra
ashish18csu038@ncuindia.edu

Abstract

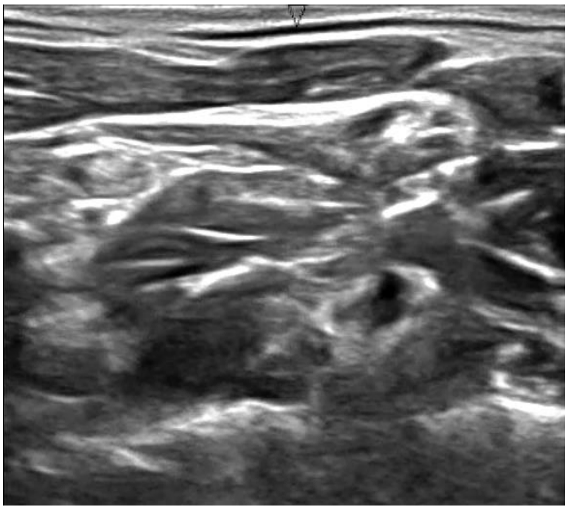
After surgery, patients have typically prescribed sedatives in order to reduce pain and aid in the recovery process. Sedatives, however, have many adverse health effects such as nausea/vomiting, liver damage, and physical dependence. Alternative solutions for pain mitigation exist but require radiologists to locate nerve clusters in ultrasound images. Our goal is to be able to accurately locate these nerve clusters using a convolutional neural network so that this alternate pain management solution can be accessible without needing a radiologist to locate nerve clusters.

Keywords: Ultrasound Image, Nerve Segmentation, Deep Learning, Convolution Neural Network

1. Introduction

Neck Pain is a complicated problem nowadays due to prolonged straining, uncomfortable sitting position and long working hours at corporate. This needs to be solved by an effective method based on deep learning. Recently, some study of ache recognition have employed machine learning for this approach. This find out about focuses on the deep models for pain recognition. With the development and popularization of scientific imaging analysis equipment like ultrasound, have become indispensable gadgets for medical institutions to carry out disorder diagnosis, surgical planning and prognosis evaluation. Ultrasound is the most widely used technique in the field of radio imaging. One of the exceptional features of ultrasound imaging is the wide variety of imaging sequences. The ultrasound picture is produced based on the reflection of the waves off of the body structures. The strength (amplitude) of the sound sign and the time it takes for the wave to travel through the body grant the information necessary to produce an image. Ultrasound imaging (ultrasonography) uses high-frequency sound waves to view internal of the body. Unlike CT and imaging, the resolution of the ultrasound pictures is comparatively low. Medical photograph segmentation is a critical step in the field of medical photograph analysis. In order to provide a reliable basis for medical diagnosis, it need to segment the components of medical images we focus and extract applicable features. Initially, medical image analysis was once done with sequential application of low-level pixel processing (e.g., region primarily based method or threshold based method) and mathematical modelling construct compound rule-based structure that solved particular tasks. The segmentation results of this period are commonly not semantically labelled. In the era of deep learning, image segmentation commonly denotes sematic segmentation, which refers to recognition of images at the pixels level.

For example, in the figure below the medical image within the left consists of ultrasound of the area around the neck and therefore the right image is that the results of its linguistics segmentation, that divides the picture element linguistics object i.e., the white labelled region is that where catheter is placed similarly alternatives organisations square measure thought of as background and t be marked black. The most prosperous variety of deep learning models for image analysis so far square measure convolution neural networks (CNNs).



(a)

Figure 1. (b)
Nerve segmentation example results in deep learning methodology. (a) is that the ultrasound of neck image, (b) is that the segmentation result.

The mutual promotion of deep studying and big data and cloud computing has brought lot of development to computer vision. CNN is the most widely used neural community in the field of computer vision which is proposed to resolve image classification problems. Image segmentation is a common task in each natural and medical image analysis. In this paper, convolution neural network architecture will be discussed to solve the problem of locating brachial plexus (network of nerves) from ultrasound images.

2. Literature Review

In Tanner Henry, Neil Vranicar, Stephen West [1], for semantic segmentation the convolution network architecture U-Net is used. Split data in train, test and validation set due to less data in training set, augmentation technique is used to maintain a large training set. Activation function used in all layers is ELU (Exponential Linear Unit) except in last layer sigmoid is used. Moreover, post-preprocessing like thresholding is done to improve performance further. In Vidushi Vashishtha and Dr. Aju D [2], the proposed methodology is to divide the processing of image into 3 stages i.e., image preprocessing, nerve segmentation and machine learning. Preprocess the data to remove noise, segmentation algorithm canny and sobel operator are applied to extract the edges detection and lastly machine learning algorithm support vector machine (SVM) is used. In Naoki Yamato [3], semantic segmentation with deep learning to extract nerves from label-free scrutiny pictures obtained mistreatment coherent anti-Stokes Raman scattering (CARS) for nerve-sparing surgery is delineate. we have a tendency to developed a CARS rigid medical instrument so as to spot the precise location of peripheral nerves in surgery. myelinated nerves area unit envisioned with a CARS lipoid signal in an exceedingly label-free manner. as a result of the lipoid distribution includes different tissues moreover as nerves, nerve segmentation is needed to realize nerve-sparing surgery. we have a tendency to propose mistreatment U-Net with a VGG16

encoder as a deep learning model and pre-training with light pictures, that visualize the lipoid distribution kind of like CARS pictures. In Hitesh Garg et al. [4], the proposed methodology is on segmentation of thyroid gland. The approach used to take intensity of pixels and used feed forward neural network for classification. Used image enhancement techniques followed by feature extraction and train the network to classify the results.In Mohammed Ahmed, Hongbo Du and Alaa AlZoubi [5] performed CNN architecture for the detection of breast lesion from ultrasound images. The model is Efficient Neural Architecture Search (ENAS) having ENAS7 and ENAS17, works exceptionally well in comparison to the AlexNet and CNN3. In YU WENG, TIANBAO ZHOU [6], neural architecture search (NAS) has vital progress in up the accuracy of image classification. Recently, some works decide to extend NAS to image segmentation that shows
preliminary practicability. However, all of them concentrate
on looking design for linguistics segmentat ion in natural scenes. during this paper, we have a tendency to style 3 kinds of primitive operation assault search house to mechanically notice 2 cell design DownSC and UpSC for linguistics image segmentation particularly medical image segmentation. galvanized by the U-net design and its variants with success applied to numerous medical image segmentation, we have a tendency to propose NAS-Unet that is stacked by identical variety. Iresha Rubasinghe & Dulani Meedeniya [7], paper discusses associate degree application for analysis of ultrasound nerve segmentation-based medicine pictures.

Our technique uses the probabilistic artificial language Edward with the U-Net model and generative adversarial networks below totally different optimizers. The segmentation method showed the smallest amount Dice loss (-0.54) and therefore the highest accuracy (0.99) with the Adam optimizer within the U-Net model with the smallest amount time consumption compared to alternative optimizers. In laf Ronneberger, Philipp Fischer, and Thomas Brox [8], the authors used an overlapping tile strategy to apply the network to large images, and used mirroring to extend past the image border. Data augmentation included elastic deformations. The loss function included per-pixel weights both to balance overall class frequencies and to draw a clear separation between objects of the same class. In YouLi [9], he attempts to segment medical ultrasound images using convolutional neural networks (CNNs) with a group of noisy activation functions which have recently been demonstrated to improve the performance of neural networks. He reports on the segmentation results using a U-Net-like CNN with noisy rectified linear unit (NReLU) functions, noisy hard sigmoid (NHSigmoid) functions, and noisy hard tanh (NHTanh) function on a small data set. In Cong Liu1,∗ , Feng Liu1 , Lang Wang1 , Longhua Ma1 and Zhe-Ming Lu1,2 [10], they develop a deep adversarial neural network. Firstly, set up a segmentation network based on well-established deep neural network. Secondly, the anatomical dependencies are ensured by a discriminator network that assesses the segmentation quality and punishes the segmentation network accordingly. Thirdly, the elastic deformation and its byproduct, small object issue, are handled by deformation data augmentation and diluted convolutions respectively. Comparing their approach to estimates made by experts in brachial plexus diagnosis shows significant performance gain over state-of-the-art models. The proposed deep adversarial networks perform best among all four other deep models. This suggests the high order dependencies captured by the proposed model play a key role in this task.

3. Proposed Methodology

Apart from classification, CNN is employed nowadays for additional advanced issues like image segmentation, object detection, etc. Image segmentation may be a method in a computer vision wherever the image is segmented into different segments representing every different category within the image. Segmentation helps to spot wherever objects of various categories are present in an image. U Net is a convolutional neural network architecture that swollen with few changes within the CNN architecture. It was invented to deal with biomedical images where the target is not only to classify whether there is a damaged nerve or not but also to identify the area of damaged nerve.

This article can demonstrate however we are able to build a image segmentation model using U-Net that will predict the mask of an object present in an image. The model can localize the object within the image using this methodology. The U Net design had two main components that were encoder and decoder. The encoder is all regarding the covenant layers followed by pooling operation. It's accustomed extract the factors within the image. The second half, decoder uses transposed convolution to allow localization. It is once more an F.C connected layers network. The SegNet works on encoder network and a corresponding decoder network, then the final results is pass through final pixelwise classification layer. The ultimate decoder output is fed to a multi-class soft-max classifier to supply class probabilities for every pixel independently. In Trasnfer Res U Net we used the segmentation_models library to work on our model. We use the layers of a deep neural network (ResNet) and therefore the parameters found training on image classification (ImageNet) and use them for our U Net.

How Seg Net work :

Seg Net has associated degree encoder network and a corresponding decoder network, followed by a final pixelwise classification layer. This design is illustrated in higher than figure. The encoder network consists of thirteen convolutional layers that correspond to the first thirteen convolutional layers inside the VGG16 network designed for object classification. They discard the absolutely connected layers in favour of retentive higher resolution feature maps at the deepest encoder output. This additionally reduces the number of parameters inside the Seg Net encoder network considerably (from 134M to 147M) as compared to alternative recent architectures. Every encoder layer options a corresponding decoder layer and thence the decoder network has thirteen layers. the final word decoder output is fed to a multi-class soft-max categorifier to provide class possibilities for each constituent severally.

How U-Net work for our model:

U-Net consists of Convolution Operation, Max Pooling, ReLU Activation, Concatenation and Up Sampling Layers and 3 sections: contraction, bottleneck, and expansion section. the contractions section has four contraction blocks. Each contraction block gets an input, applies 2 3X3 convolution ReLU layers and then it apllies 2X2 max pooling. The quantity of feature maps gets double at every pooling layer. The bottleneck layer uses 2 3X3 Convolutional layers and 2X2 up convolution layer. The enlargement section consists of many enlargement blocks with every block passing the input to two 3X3 Conv layers and a 2X2 up sampling layer that halves the quantity of feature channels. It additionally includes a concatenation with the correspondingly cropped feature map from the acquiring path. At last, 1X1 Conv layer is utilized to create the quantity of feature maps as same as the amount of feature maps that are desired in the output. U-net uses a loss function for each element of the image. This helps in straightforward identification of individual cells inside the segmentation map. Sigmoid is applied to every pixel followed by a loss function. This converts the segmentation drawback into a classification drawback where we'd prefer to classify every element to a minimum of in every of the categories.

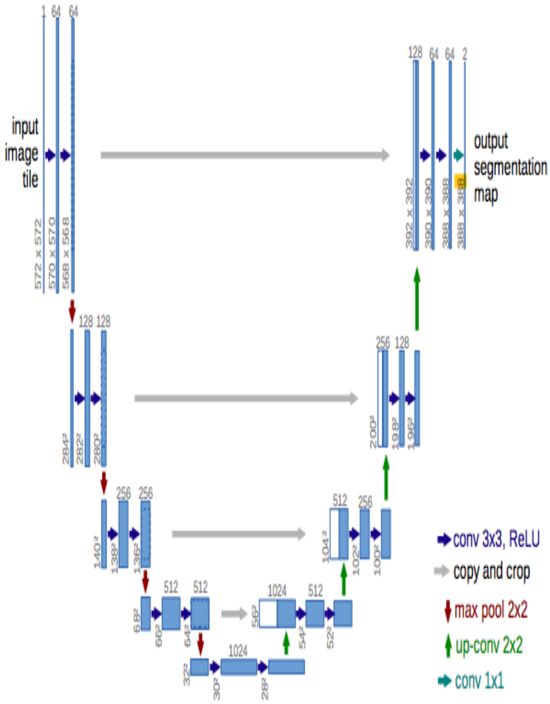


Figure 2 U- NET

How Transfer Res U Net work for our model:

We have created two lists one for storing masks and therefore the alternative for storing the image (mask & img). Once storing the image and mask we've picked solely 10281 images with their corresponding masks.

Sorting has been done we tend to read the image and label mask in X and y. We've captured the index of the image file and keep the directory of that index image. After that, we tend to open that image and size it conjointly then we tend to convert that image into a grayscale image and store its index. We tend to then store the mask of the image appreciate the index we stored the grayscale image. After that, we offer the directory of the mask and browse the mask. Finally, we've pre-processed the mask image by resizing it and normalizing the pixel value then keep it at the pre-processed mask image at the output array at the equivalent index position. X[n] stores the image and y[n] stores the corresponding mask.

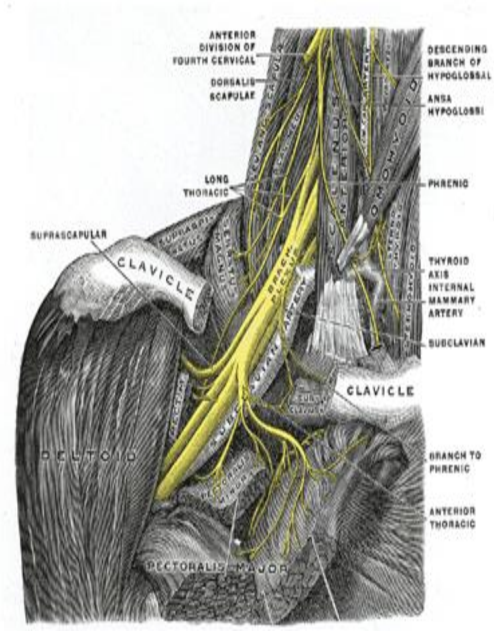


Figure 3 BP Nerve

Import the U Net model where Res Net works as a backbone network and helps to load weights of image net. We've then outlined the input form that's expected by the bottom model and the custom layer that takes that base mode input whose output is then passed to the U Net model. The output of U Net model is then passed to alternative outlined Conv Net layers having activation as ReLU. The output is then reshaped to 28X28. At last, we've outlined the model that takes input (inp) and provides us the output (x_out) victimisation the base_model.

Analytical Part: After saving the model we tend to create predictions on X_train and X_test victimization the trained model and keep it. Once creating prediction we tend to have an outlined function to visualize the prediction created by the model. The function expects input array and output array therefore the predictions. We've outlined k to be none so it will pick random images from the training data and for the equivalent index of the picked training image we have taken the mask. We've then outlined the figure size and premeditated all three that are image, mask, and predicted mask. Image segmentation is a most useful part in computer vision which will be applied to a variety of use-cases in medical field to capture different segments or different classes in real-time.

4. Experimental Setup and Result Analysis

Dataset. A dataset consists of high resolutions ultrasound scans images and reciprocate mask images of brachial plexus. The segmentation module was trained over 8224 high resolution images with annotations form doctors which enable large scale brachial plexus prediction for real clinic images and in order to test the model performance it was tested over another 2057 high resolution images. As in the dataset, where the brachial plexus is located the corresponding mask is shown and where it is not present no mask is shown.

Device. The model was executed on a Acer Predator workstation using Nvidia graphics card with 4 GB graphic memory, 16 GB internal RAM and 1 TB ROM with rotational speed of 7200 RPM.

Evaluation Metrics. The primary metric for evaluation is the Sørensen–Dice Coefficient[11]. It is also commonly called the similarity. Today, it is a common method for analyzing pixel-wise agreement between two images.

[Equation 1.]

$$\text{Dice} = 2\ |A \cap B|\ /\ |A| + |B|$$

Intersection over Union (IOU) [12] is an evaluation metric used to measure the accuracy of an object detector on a particular dataset. Intersection over union, also known as Jaccard index, is the percent overlap between the target mask and the prediction output. Computing Intersection over Union can therefore be determined via:

[Equation 2.]

$$\text{IOU} = \text{Area of Overlap} \ /\ \text{Area of Union}$$

or

$$J(X,Y) = |X \cap Y|\ /\ |XUY|$$

Training. In this project, a few different models were implemented: Seg Net, U Net, Trans Res U Net. The segmentation network is trained alternatively from scratch. The experimental results showed that training became not much unstable in Seg Net and U Net but Trans Res U Net performed exceptionally well. Few losses are defined in form of graphs in Fig 3 for Seg Net, in Fig 4 for U Net and in Fig 5 for Trans Res U Net.

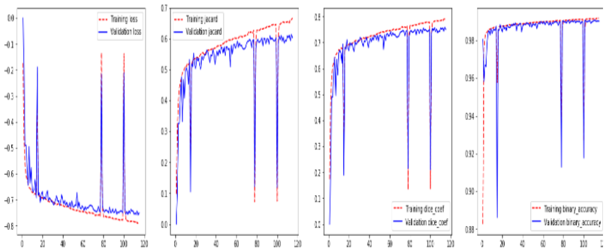


Figure 4 Seg Net Model results

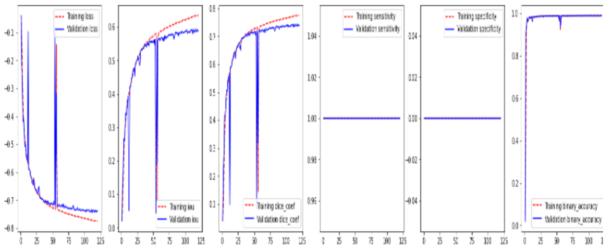


Figure 5 U Net Model results

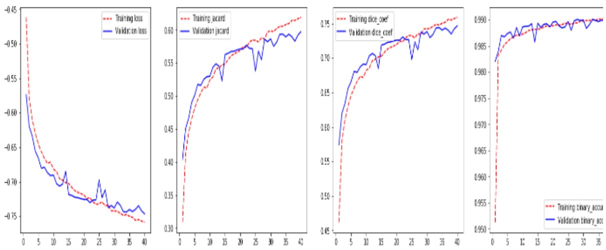


Figure 6 Trans Res U Net Model results

Results. The model has been implemented various Deep Convolution Networks Fig 7 shows examples of three different segmentation cases: (a) true positive case, where the algorithm correctly detects and locate the region of interest (ROI); (b) false negative, where the algorithm does not detect the nerve in the image; and (c) False positive case. The algorithm indicates the existence of nerves in images where the nerves do not exist. Of all cases, (a) is the ideal one, and (b) is the one that needs to be avoided at all costs, because it may lead to injuries to patients.

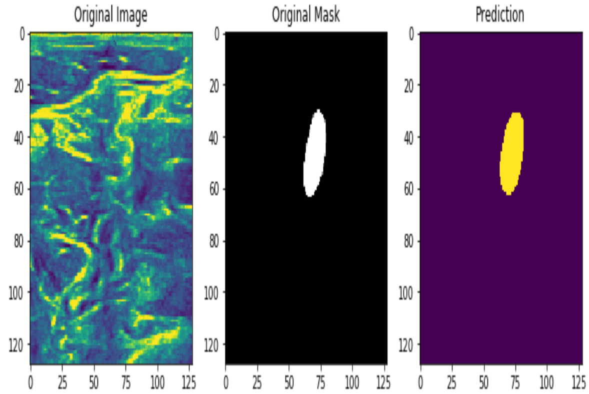


Figure 7 (a). True positive case. The algorithm correctly detects the nerves and achieve high agreement with the segmentation by human volunteers.

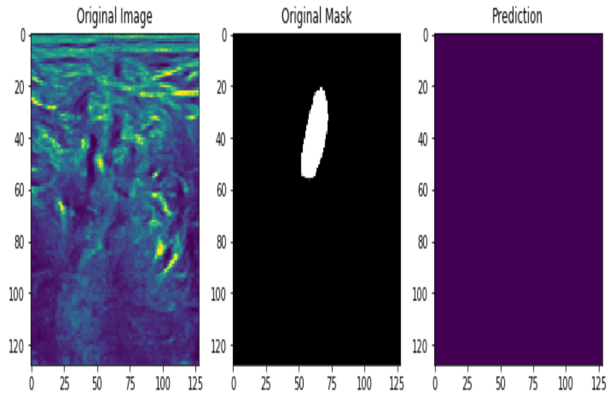


Figure 7 (b). False negative case. The algorithm does not detect the nerve in the image. This is the case that should be avoided at all costs, because it may lead to injuries to patients.

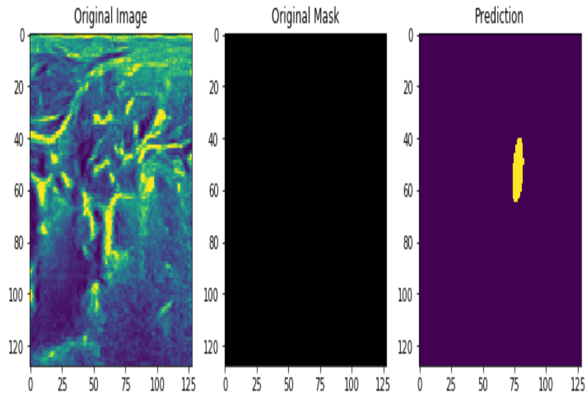


Figure 7 (c). False positive case. The algorithm indicates the existence of nerves in images where the nerves do not exist.

We explore the three different architectures with different fields-of-view and filter channels are explored here. Out of those three, Seg Net gives the highest accuracy but it overfits to the model and when tested on test data set it gives more false positive and false negative however U Net performed well but it also gave some incorrect results so by hyperparameter tuning and with the use of transfer learning Trans Res U Net gave the excellent performance on the dataset.

Table 1. Performances comparison between different discriminator architectures

Model / Metrics	Seg Net	U Net	Trans Res U Net
Avg. Accuracy	0.99	0.98	0.99
Avg. Loss	-0.76	-0.74	-0.74
Avg. Dice coefficient	0.76	0.74	0.74
Avg. IOU/ Jaccard	0.61	0.59	0.59

5. Conclusion and Future Work

In this paper, we attempted Neural Architecture to medical image of nerve segmentation. This paper discusses the different deep learning models which can be useful to get good result in medical nerve segmentation field. As seen from the above results of the experiment, Transfer Res U Net proves to be far better than Seg Net and U Net.

In the future with more time and resources, we would work to expand the network and dataset to allow for classification of a wider variety of nerve clusters. We would also like to apply the model to video stream input rather than a single image to allow for real time image labelling; which is essential to allowing a medical professional to accurately insert the catheter during the procedure.

References

[1] Henry, Tanner, Neil Vranicar, and Stephen West. "Ultrasound Nerve Segmentation–Group 9."

[2] Vashishtha, Vidushi, and D. Aju. "Nerve segmentation in ultrasound images." In *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)*, pp. 1-5. IEEE, 2017.

[3] Yamato, Naoki, Hirohiko Niioka, Jun Miyake, and Mamoru Hashimoto. "Improvement of nerve imaging speed with coherent anti-Stokes Raman scattering rigid endoscope using deep-learning noise reduction." *Scientific reports* 10, no. 1 (2020): 1-11.

[4] Garg, Hitesh, and Alka Jindal. "Segmentation of thyroid gland in ultrasound image using neural network." In *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, pp. 1-5. IEEE, 2013.

[5] Ahmed, Mohammed, Hongbo Du, and Alaa AlZoubi. "An ENAS Based Approach for Constructing Deep Learning Models for Breast Cancer Recognition from Ultrasound Images." *arXiv preprint arXiv:2005.13695* (2020).

[6] Weng, Yu, Tianbao Zhou, Yujie Li, and Xiaoyu Qiu. "NAS-Unet: Neural architecture search for medical image segmentation." *IEEE Access* 7 (2019): 44247-44257.

[7] Rubasinghe, Iresha, and Dulani Meedeniya. "Ultrasound nerve segmentation using deep probabilistic programming." *Journal of ICT Research and Applications* 13, no. 3 (2019): 241-256.

[8] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234-241. Springer, Cham, 2015.

[9] Li, You. "Segmentation of medical ultrasound images using convolutional neural networks with noisy activating functions." *Semantic Scholar* (2016).

[10] Liu, Cong, Feng Liu, Lang Wang, Longhua Ma, and Zhe-Ming Lu. "Segmentation of nerve on ultrasound images using deep adversarial network." *Int J Innov Comput Inform Control* 14 (2018): 53-64.

[11] Dietenbeck, Thomas, Martino Alessandrini, Denis Friboulet, and Olivier Bernard. "CREASEG: a free software for the evaluation of image segmentation algorithms based on level-set." In *2010 IEEE International Conference on Image Processing*, pp. 665-668. IEEE, 2010.

[12] Rahman, Md Atiqur, and Yang Wang. "Optimizing intersection-over-union in deep neural networks for image segmentation." In *International symposium on visual computing*, pp. 234-244. Springer, Cham, 2016.

[13] Niwattanakul, Suphakit, Jatsada Singthongchai, Ekkachai Naenudorn, and Supachanun Wanapu. "Using of Jaccard coefficient for keywords similarity." In *Proceedings of the international multiconference of engineers and computer scientists*, vol. 1, no. 6, pp. 380-384. 2013.

[14] Wang, Ziyang, Zhengdong Zhang, Jianqing Zheng, Baoru Huang, Irina Voiculescu, and Guang-Zhong Yang. "Deep learning in medical ultrasound image segmentation: A review." *arXiv preprint arXiv:2002.07703* (2020).

[15] Schug, Stephan A., Detlev Zech, and Stefan Grond. "Adverse effects of systemic opioid analgesics." *Drug Safety* 7, no. 3 (1992): 200-213.

[16] Ciresan, Dan, Alessandro Giusti, Luca Gambardella, and Jürgen Schmidhuber. "Deep neural networks segment neuronal membranes in electron microscopy images." *Advances in neural information processing systems* 25 (2012): 2843-2851.

[17] Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3431-3440. 2015.

[18] Halyard Health, "Ultrasound Nerve Segmentation," 6 May 2016. [Online]. Available: https://www.kaggle.com/c/ultrasound-nerve-segmentation/data. [Accessed 2 May 2019].

[19] jocicmarko, "Github," 4 February 2017. [Online]. Available: https://github.com/jocicmarko/ultrasound-nerve-segmentation. [Accessed 11 June 2019].

[20] A. Beam, "Andrew L. Beam: Machine Learning and Medicine," 20 August 2016. [Online]. Available: http://beamandrew.github.io/deeplearning/2016/08/20/kaggle/segmentation.html. [Accessed 11 June 2019].

[21] Dice, Lee R. "Measures of the amount of ecologic association between species." *Ecology* 26, no. 3 (1945): 297-302.

[22] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778. 2016.

[23] Szegedy, Christian, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. "Rethinking the inception architecture for computer vision." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818-2826. 2016.